

Traffic Characterisation and Modelling of VBR Coded MPEG Sources

*A. D. Doulamis, N. D. Doulamis, G. E. Konstantoulakis
and G. I. Stassinopoulos*

*National Technical University of Athens
Division of Computer Science
Zographou, GR-157 73 Athens, Greece
tel: + 30 1 772 1494, fax: + 30 1 772 2530
e-mail: gkonst@telecom.ntua.gr*

Abstract

This paper intends to study the characteristics of VBR MPEG sources and performs models for their traffic behaviour. Our analysis focuses on long-time MPEG video films, approximately 27 min. Known probability density functions (pdf) are presented which fit well with real the pdf of I, P, B and Group of pictures (GOP) frames. In order to approximate the traffic behaviour, we study the autocovariance functions of the previous frames and we propose two different models so as to achieve the necessary fit. The first model is based on GOP frame layer which is able to approximate the cell probability loss at utilisation about 0.7-0.85 but it slightly overestimates the frame loss probability. The other model correlates the error of I, P and B AR models without spoiling the i.i.d. (independent identical distributed) property. Our results show that the latter model approximates well the cell as well as the frame loss probabilities.

INTRODUCTION

Digital video, such as video telephone, video teleconference and switched TV as well as the forthcoming Video on Demand, is expected to become a major traffic component on integrated services digital networks. Since all the above applications impose very large bandwidth requirements, several coding algorithms have been proposed in order to perform efficient and effective video compression without losing

the quality. Thus the required bandwidth to transmit coded digital video is reduced. One of the foremost coding techniques is MPEG-I (Moving Picture Expert Group I) which has been standardised in 1992 and now appears in real applications. It was developed basically for storage of compressed video, but because of its flexibility it can be used for many other applications, varying from multimedia workstations to high definition television (HDTV).

Two main different modes are used for encoding any video source, namely Constant Bit Rate (CBR) and Variable Bit Rate (VBR). In a CBR scheme the video quality cannot be maintained constant for all scenes, since a resolution-reduction mechanism during high video activity is activated to achieve the required constant bit rate. However the users of video applications desire invariable quality regardless the complexity of scenes. Therefore they prefer a VBR scheme, which maintains video quality almost constant (varying the output bit rate), to a CBR one. Since in this scheme the output bit rate changes from extremely low values up to peak values, which can be many times greater than the average, it is very wasteful to choose the allocated bandwidth of the network equal to peak bit rate. To avoid the waste, we are compelled to place a buffer between the video source and the network to convert the variable bit rate to an almost constant one (mainly to smooth out peaks).

Statistical analysis and performance models of different kinds of video traffic are essential for proper network design. These models can be used to estimate the network resources i.e. buffer at a determined cell loss probability or bandwidth allocation. Since a significant fraction of video traffic is anticipated to be generated by MPEG sources, it is quite important to propose models that approach its statistical characteristics and describe the behaviour of buffer size in network multiplexing points.

Several VBR video models have been reported in bibliography. Two markovian chains, one in discrete time and one in continuous were studied in (Maglaris, 1988) for sources without sudden rate jumps due to error recovery procedure (scenes change and camera zooming). In (Sen, 1989) it is suggested a multilayer markovian chain to approximate the variations within a single scene more precisely than the previous model. A gamma distribution is considered accurate for a VBR teleconference video sequence (three people talking with no scenes changes and cuts) and a multirate markovian chain is reported in (Heyman, 1992). A new model which is called scenic is performed by Frater and Arnold in (Frater, 1994). An autoregressive moving average (ARMA) was proposed to characterise the VBR-coded data with more general scene content and correlation structure (Grunenfelder, 1991). However it seems that an autoregressive model is more convenient than an ARMA one. Three different methods such as the stationary interval (SI), asymptotic (ASM) and hybrid queuing network analysis were used to approximate the average queue size of VBR sources (Nasser, 1994).

Although many statistical models have been performed for VBR video traffic (Maglaris, 1988, Nasser, 1994), very little work has been done to characterise MPEG-I video sources. As we describe below (section The MPEG encoder), any MPEG sequence consists of a set of three types of frames I (Intraframe), P (Predictive), B (Interpolative) which have different statistical properties. Since the MPEG encoder repeats these frames periodically, it follows that traditional stochastic models cannot identify the data of these sources. Another difficult problem is the fact that there is relation among the I, P, B frames due to the algorithm of MPEG encoder. This dependence means that a high size of one of these frames is usually followed by high sizes of the other two.

In this paper we survey the statistical properties of some long MPEG coding sequence (40 000 frames or about 27 min film). We use these characteristics to find models which describes accurately the MPEG sources and predicts the cell loss probability under certain buffer utilisation. Our sequences have been recorded using video films and consequently high video activity, scenes changes and camera zooming have been encountered. In the next section a description of MPEG algorithm and some statistical properties are presented. Then, we study the histograms of I, B, P and GOP (Group Of Picture) frames and their autocovariance functions respectively as well as the total sequence (IPB). Finally, we propose two models for approximating the probabilities loss. One at GOP layer and another at frame layer. Finally the last section concludes this paper.

MPEG ALGORITHM AND CHARACTERISTICS.

The MPEG encoder

Since an uncompressed digital video stream needs hundreds of Mbits/sec to be transmitted, video compression has been studied a lot and several coding algorithms have been proposed to reduce the required bandwidth. For instance, ITU adopted H.261 (Liou, 1991) for video telephone and video teleconference. Another attempt resulted in standardisation of MPEG-I algorithm for storage of moving pictures. However, due to the flexibility of its algorithm, it can be used for various applications such as in multimedia workstations, video communication and so on.

The MPEG coding algorithm uses three different types of frames; Intraframe (I), Predictive (P) and Interpolative (B) frames. Initially, an I frame contains a two-dimensional 8×8 discrete cosine transform (DCT) of the original image. Secondly the coefficients of DCT are fed to a quantizer and a 8×8 matrix obtains the quantization step size for of each DCT coefficient. Some psycho physics experiments have defined

the values of this matrix. (ISO/IEC, 1994) Low frequency DCT coefficients are quantized with more accuracy (small step size) than the high frequency ones. As a result a significant number of quantized coefficients will have zero value. Then the algorithm codes this block (8x8) of coefficients using run length and modified Huffman coding. The MPEG standard has defined the most frequent combinations of zero run lengths that follow a non-zero coefficient value. If a specific combination does not exist in the MPEG standard the coding uses a fixed word following by the length of zero and the non-zero value of DCT coefficient (ISO/IEC, 1994).

In P frames, macroblocks are coded with or without motion estimation. The algorithm searches a square area around each macroblock of the previous I or P frame in order to find a motion vector that minimises the absolute difference between the current macroblock and the chosen one in the previous frame. If the absolute difference is less than a threshold, the motion vector is coded and transmitted. Then a DCT transform is applied to the prediction error of each macroblock. The coefficients of transformation are quantized with constant step size instead of intraframe coding. The quantized coefficients are coded as in intraframe mode. If the absolute difference is greater than the above threshold, the motion estimation cannot be used and the coding of the current macroblock is the same as the intraframe. B frames coding is similar to the procedure that have been previously described for (P ones). The only difference is that the motion vector can be estimated with respect to the previous I, P frames, or the following I, P frames or an interpolation between them.

Since the MPEG coding algorithm has a fixed quantizer matrix for each mode (I, P, B), it can scale the quantization level using a parameter q . When MPEG video tries to maintain the output rate constant (CBR mode), it is necessary to vary the q parameter dynamically, therefore the quality of coded moving pictures. If q remains constant, the quality does not change but the output rate becomes to depend on the video activity (Pancha, 1994).

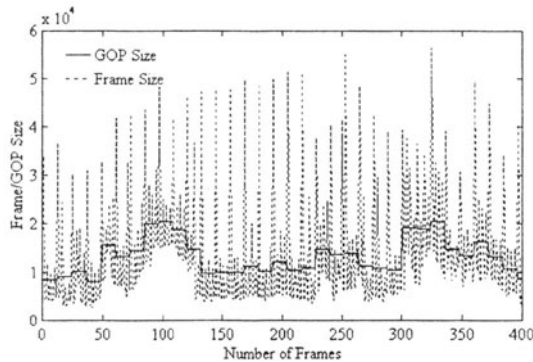


Figure 1 : Frame and GOP size of an MPEG video sequence.

In a MPEG coding there is a constant number N of frames where I, P and B are repeated periodically. The most typical cyclic frame pattern is:

IBBPBBPBBPBBI

The pattern period consists of 1 I, 3 P and 8 B frames. We call the mean value of this cyclic frame pattern Group of Picture (GOP) in the rest of this paper. That means that the GOP value represents the average size of I, P and B frames within one pattern. Figure 1 illustrates the first 500 frames' size of an MPEG video sequence where $N=12$. The solid line presents the GOP size. It is evident that the GOP size follows the video activity, meaning that in case of a lot of motion within a scene the GOP size is high and vice versa.

	<i>Terminator</i>	<i>Jurassic Park</i>	<i>Asterix</i>
Max/Min (I- P- B)	4.8-61.4-162.4	6.8-61.2-78.9	9.4-82.6-288.8
Max/min (GOP)	16.4	12.6	34.9
Mean(I)/Mean(Total)	3.4	4.2	3.1

Table 1 : Some characteristics of MPEG video data

MPEG Characteristics

In Intraframe mode, the size of I frames, has greater value than the size of P and B frames due to the lack of motion estimation. However, the volume of P and in particular of B frames is not always very small. As we have stated above, it is possible for the coding of P and B frames not to use the motion compensation for all macroblocks. This occurs when the prediction error is greater than a defined threshold. Therefore it is anticipated that a high activity causes great sizes of P and B frames. Table 1 shows some characteristics of MPEG video sequences using PAL system (25 frames per second). The table gives these characteristic for these movies (Terminator, Jurassic Park, Asterix), that will be used as throughout this work. We survey 40 000 frames (approximately 27 min) to achieve accurate statistical characteristics.

We observe that the fluctuation of B and P frames is much greater than the of I B frames have the highest ratio of maximum to minimum frame size. Another important characteristic is the ratio Mean(I)/Mean(total) because it affects the aggregate traffic behaviour. Since I frames have a mean value three or four times greater than the average total frames' size (including I, P, B) it is very difficult to achieve, without multiplexing gain, low loss probabilities with small buffer size and for utilisation close to 0.75 or above. Although the small size of P and B frames, it seems that they play a

major role to loss probabilities due to their large fluctuation (see section Traffic behaviour). Therefore it is not correct to ignore them so as to study the traffic characteristics of MPEG video.

Moreover, due to the motion estimation there is a significant correlation between the I, P and B frames (Heyman, 1994). Let B_i ($i=1$ to 8) and P_i ($i=1$ to 3) be the i^{th} B and P frame respectively in a cyclic frame pattern. According to the MPEG algorithm, B_1 and B_2 can have very different values only if a video scene change occurs between them. In addition, P_1 is related to I frame and P_2 to P_1 and so on. The same relation has been noticed among B_1 and P_1 , B_2 and P_1 or B_1 , B_2 and I frames. These relations are consequence of the motion compensation that the MPEG algorithm

	$B_1 B_2$	$B_1 P_1$	$B_1 I$	$P_2 P_1$	$P_1 I$
Correlation Coefficient Terminator	0.83	0.71	0.25	0.61	0.28
Correlation Coefficient Jurassic Park	0.93	0.83	0.32	0.78	0.34
Correlation Coefficient Asterix	0.86	0.82	0.59	0.78	0.59

Table 2: The correlation coefficients of some B_i , P_i I frames

performs. In particular it is observed that the strongest dependence is between B frames while the most weak between P, I and B, I ones. It is therefore difficult to find appropriate models that can fit accurately with a real video MPEG source. Table 2 shows the correlation coefficients of the previous films.

THE DISTRIBUTION OF AN MPEG SOURCE.

Modelling of an MPEG sequence is a very useful work for designing a telecommunication network especially due to the great evolution of multimedia services along with the advances in VLSI technology that made possible the realisation of real time MPEG systems. It can be used to determine the loss probabilities or the required network resources (bandwidth, buffers) without having to witness the real

video data. We examine statistical models which characterise the properties of an MPEG video. Our analysis is based on a long real video sequence (approximately 27 min) with scene changes and camera zooming that are considered as the most difficult sequences for the MPEG to handle.

As we mentioned above, any MPEG video sequence consists of three types of frames; Intraframe (I), Predictive (P) and Interpolative (B). The size histograms of these frames are shown in Figure 2 (a)-(e) respectively as well as the size of GOP. Since these histograms and in particular the ones of P and B frames have a long right tail, it seems that the lognormal and the gamma distributions fit well with the experimental data.

Gamma distribution

The probability density function for gamma distribution is given by

$$f_g(x) = \frac{a^p}{\Gamma(p)} x^{(p-1)} e^{-x}. \quad (1)$$

where $\Gamma(p)$ is the known gamma function defined as $\Gamma(p) = \int_0^{\infty} x^{(p-1)} e^{-x} dx$.

It can be easily proved (Papoulis, 1984) that the mean and the variance of gamma distribution are related to the parameters a, p as following

$$E_g(x) = \frac{p}{a} \quad \text{and} \quad V_g(x) = \frac{p}{a^2}. \quad (2)$$

In order to estimate the unknown parameters of gamma pdf which fit well with the real data we use the method of moments. If MEAN and VAR are the mean and the variance of an MPEG video sequence then the estimated parameters are calculated as

$$\hat{a} = \frac{MEAN}{VAR} \quad \text{and} \quad \hat{p} = \frac{MEAN^2}{VAR}. \quad (3)$$

Lognormal distribution

The pdf function of log normal distribution is

$$f_L(x) = \frac{1}{\sqrt{2\pi\sigma^2}} x^{-1} \exp\left\{-\frac{1}{2\sigma^2}(\ln(x) - \mu)^2\right\}. \quad (4)$$

where its mean and variance are given by $E_{LN}(x) = e^{\mu + \frac{\sigma^2}{2}}$ and $V_{LN}(x) = e^{2\mu + \sigma^2} (e^{\sigma^2} - 1)$. The estimated parameters which are calculated by the method of moments are given by the following formulas

$$\hat{\sigma}^2 = \ln\left\{\frac{VAR}{\exp(2 \ln(MEAN))} + 1\right\} \quad \text{and} \quad \hat{\mu} = \ln(MEAN) - \frac{\sigma^2}{2}. \quad (5)$$

Based on the equations (3) and (5) we can find the appropriate estimated parameters of gamma and lognormal pdfs which fit well with the real video data. Table 3 presents the estimated parameters for the three different films we use.

	Terminator	Jurassic Park	Asterix
<i>Lognormal pdf</i>			
I Frame	$\hat{\mu}=4.56 \quad \hat{\sigma}^2=0.048$	$\hat{\mu}=4.95 \quad \hat{\sigma}^2=0.043$	$\hat{\mu}=5.19 \quad \hat{\sigma}^2=0.070$
P Frame	$\hat{\mu}= 3.51 \quad \hat{\sigma}^2 =0.23$	$\hat{\mu}=3.46 \quad \hat{\sigma}^2=0.36$	$\hat{\mu}=4.11 \quad \hat{\sigma}^2 =0.304$
B Frame	$\hat{\mu}=2.70 \quad \hat{\sigma}^2=0.29$	$\hat{\mu}=2.82 \quad \hat{\sigma}^2=0.31$	$\hat{\mu}= 3.46 \quad \hat{\sigma}^2 =0.37$
<i>Gamma pdf</i>			
I frame	$\hat{a}=0.21 \quad \hat{p}=20.30$	$\hat{a}=0.16 \quad \hat{p}=22.54$	$\hat{a}=0.074 \quad \hat{p}=13.82$
P frame	$\hat{a}=0.11 \quad \hat{p}=3.94$	$\hat{a}=0.06 \quad \hat{p}=2.30$	$\hat{a}=0.04 \quad \hat{p}=2.82$
B Frame	$\hat{a}=0.17 \quad \hat{p}=2.97$	$\hat{a}=0.145 \quad \hat{p}=2.83$	$\hat{a}=0.06 \quad \hat{p}=2.25$

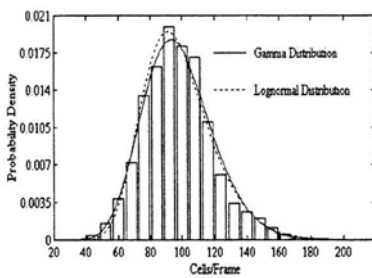
Table 3 : The estimated parameters of gamma and lognormal pdf

Fractile diagram (Q-Q plot).

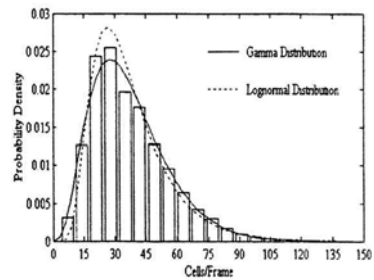
An alternative way for illustrating the fit of a histogram with the previous pdfs is the fractile diagrams method. This method plots the quantiles of the data versus the quantiles of the fitted known distributions (Quantile-Quantile plot). We describe briefly the method but more details can be found in (Kobayashi, 1981). Let $f(x)$ be a known distribution function. Then it can be easily be found (i.e. using an arithmetic method) the quantile x_0 which corresponds to a specific probability density $P_0 = f(x_0)$ that is $x_0 = f^{-1}(x_0)$. This quantile is illustrated versus the quantile y_0 of the real data with the same probability P_0 . Thus the absolute fit is on a line with 45° slope. The advantage of this method is that it shows all the differences with the same accuracy regardless of the value of probability density. For instance it is quite difficult for someone to determine the accuracy of the fit at very large frame size due to the small value of the probability density. Nevertheless, when using Q-Q plots we are able to comprehend the behaviour of fit of any known distribution with our experimental data. As a result we can easily find the most appropriate pdf.

Figure 2 (f)-(j) show the fractile diagram (Q-Q plot) for I, P, B, GOP as well as the total video data for a long MPEG sequence (Terminator film). Similar results have been found for other long MPEG sequences. It seems that the aggregate sequence follows neither gamma nor lognormal distribution function. This occurs because it merges I, P, and B frames which characterise by different statistical properties each. As a consequence, the pdf declines very slowly at large frame size because of I frames which usually consist of a large number of cells and increases rapidly at small frame size due to the large number of P and B frames.

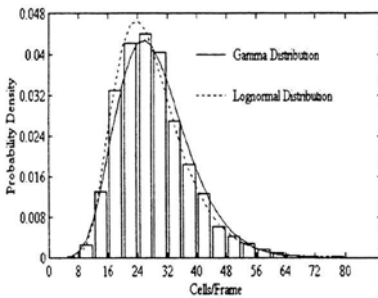
Both of lognormal and gamma distributions fit well with I frames but we can conclude that gamma is slightly better than lognormal at large frames. P frames seem to follow gamma pdf but lognormal is also quiet good (see Figure 2(a), 2(g)). However lognormal distribution gives a better approximation to B and GOP frames than gamma does without excluding the latter of course. In conclusion, lognormal and gamma pdfs fit well with the separated I, P, B and GOP frames but P follows better gamma pdf, B and GOP lognormal one and I frames both of them with gamma slightly better. If a model (e.g. DAR (Jacobs, 1983)) needs a discrete time probability density function it can be used the negative bipolar instead of gamma.



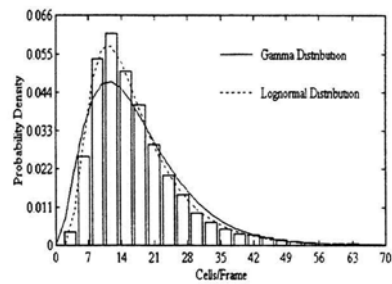
(a)



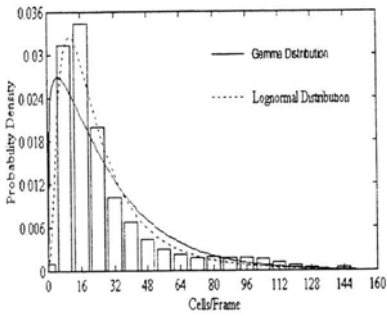
(b)



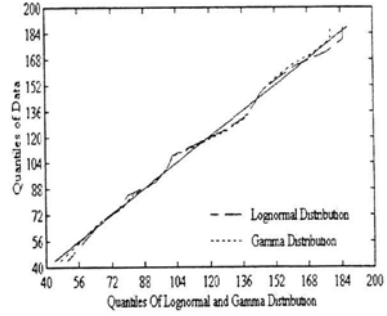
(c)



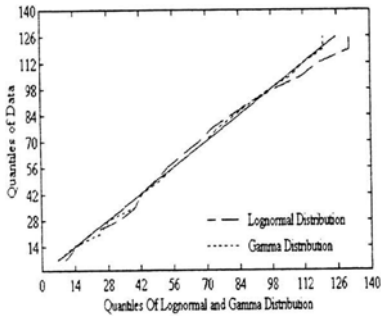
(d)



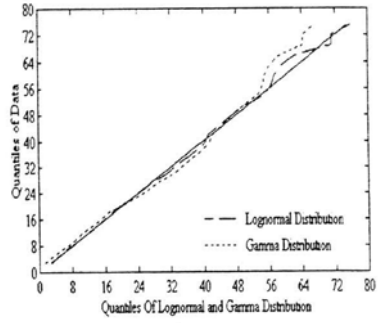
(e)



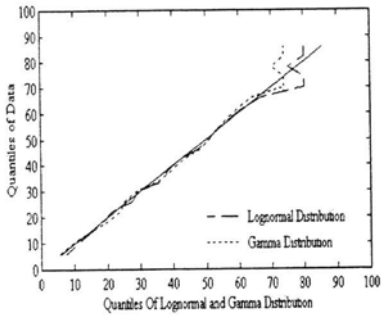
(f)



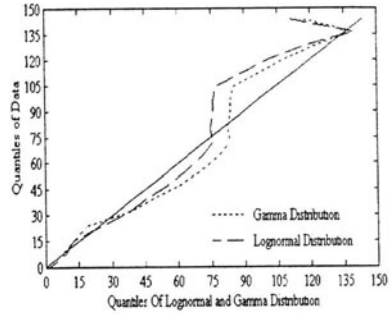
(g)



(h)



(i)



(j)

Figure 2: Histograms and Q-Q plots diagrams for Terminator sequence. (a)-(e) Histograms of I, P, B, GOP and IPB (Total) sequence respectively (f)-(j) Q-Q plots for I, P, B, GOP, IPB Frames.

AR MODELLING OF MPEG

The models that we have described in the previous section cannot approximate the traffic behaviour of a VBR MPEG video source. This occurs because each sample that is generated according to the probability density functions is independent on the others therefore no correlation among them exists. However, they are used for some other models (see DAR models (Jacobs, 1983)) which require to know the pdf of the source. In this section we describe linear models based on the correlation function (or autocovariance) of an MPEG sequence.

Analysis.

Let $x(n)$ be a stochastic process, with $n=0,1,\dots,N$. A k -order autoregressive model for $x(n)$ is defined as

$$x(n) = -\sum_{i=1}^k a_i x(n-i) + b \cdot e(n). \quad (6)$$

where $e(n)$ is an Independent and Identically Distributed (i.i.d.) variable with mean m_e , variance 1, and a_i , b constants. In order to minimise the square value of error, that is the $\min\{E(e(n)^2)\}$ we conclude to the Yule-Walker equations (Papoulis, 1984)

$$\mathbf{R} \cdot \mathbf{a} = -\mathbf{r}_k. \quad (7)$$

where \mathbf{R} is a toeplitz matrix with elements the autocovariance values of $x(n)$ which is denoted as $r(j) = E\{(x(n) - m_x)(x(n-j) - m_x)\}$ where \mathbf{a} the vector of the unknown parameters a_i and $\mathbf{r}_k = [r(1), r(2), \dots, r(k)]^T$.

The mean value of error m_e and the unknown parameter b are given by

$$m_e = \frac{m_x + \sum_{i=1}^k a_i m_x}{b} \quad \text{and} \quad (8)$$

$$b^2 = r(0) + \mathbf{a} \cdot \mathbf{r}_k.$$

The autocovariance function of an AR model for $l > k$ is

$$r(l) = -\sum_{i=1}^k a_i r(i). \quad (9)$$

In case of $k=1$ it can be shown that (Papoulis, 1984)

$$r(n) = \frac{b}{(1 - a_1^2)} a_1^n \quad \text{meaning that autocovariance function has exponential behaviour. If}$$

$k > 1$ we can find that the autocovariance consists of the sum of exponential functions.

Let $H(z)$ be the transfer function of an AR model. Then in order to generate the error $e(n)$ from $x(n)$ we have to pass $x(n)$ from a system with transfer function $1/H(Z)$. The histogram of this error indicates its probability density function.

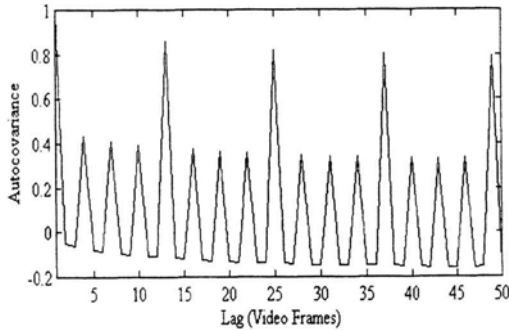


Figure 3: Autocovariance of IPB (total) MPEG sequence (Terminator).

Results

Figure 3 shows the autocovariance function of the Terminator film MPEG-I sequence. The large positive peaks stem from the I frames, the negative ones from the B frames and the others from P. Since the MPEG algorithm repeats the cyclic pattern frame periodically, there is no linear model, e.g. AR, which can fit to the autocovariance function of any MPEG video sequence. Another significant result is that the pattern between two peaks are repeated with very slow decaying amplitude.

On the contrary, it seems that the autocovariance function of I, P, B and GOP frames follows an exponential behaviour with a long-range dependence. This means that they can be characterised by an AR model but its order should be high. As we can notice from Figures (4,5) the autocovariance function of I and GOP frames have a similar behaviour. B and P frames appear to have longer-range dependence than I and GOP frames (Figures 4,5). This is logical since in the typical pattern frame (IBBPBBPBBPBB) there are eight B and only three and one P and I frames respectively. Therefore it is expected the order of B-frame AR model to be eight times higher than the order of I frame model and P-frame AR model three times greater. However it seems that the orders of P and B frames are smaller than the expected ones. In next Figures the autocovariance functions of I, P, B and GOP frames with AR models are illustrated. The error $e(n)$ of the equation 3 is calculated as follows. Firstly we find the error $e(n)$ by using the filter $1/H(Z)$ and estimate its histogram. Then we generate a random white noise which has the particular pdf based on a method which is describe in (Kobayashi, 1981). We can notice that high order AR models are required to fit well the real autocovariance function. Models of low order do not decay very slowly, thus cannot approximate the long range dependence.

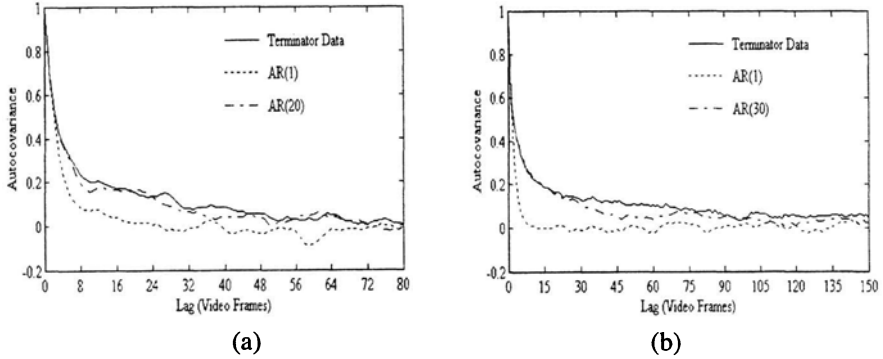


Figure 4: Autocovariance function of (a) I and (b) P frames (Term. Seq.).

We have observed that error pdf is not gaussian but has a bell shape. This is not very significant for providing a good fit for the autocovariance function but plays a major role in the study of traffic behaviour as we present it in the next section. Therefore the choice of the appropriate error distribution appears to be important.

TRAFFIC BEHAVIOUR

In this section we study the traffic behaviour of MPEG video sources. Our attention concentrates on two different layers; GOP layer and frame layer. In GOP layer, we examine a high order AR model and we conclude that this model is able to approximate cell loss probability but gives an over estimate of frame loss probability. In frame layer, high order AR models of I, P, and B, lead to severe under estimates both of cell and frame loss since there is a significant cross-correlation among I, P and B frames (see section MPEG characteristics). To solve this problem we introduce a modulated AR model which correlates I, P and B frames and we find that this model is a very good approximate for of both cell and frame loss probabilities.

To evaluate our models we fed a buffer with MPEG video data and simulated its traffic behaviour. Since we use PAL system the interframe period is 40ms (that is 25 frames per second). Each frame is converted into an equivalent integer number of ATM cells (each cell consists of 48 bytes plus 5 bytes for the header) and the generated cells are transmitted instantaneously to the buffer. If the number of new cells plus the number of cells in the buffer are greater than the defined buffer size, the incoming cells will be lost. We also consider that a frame can enter the buffer only if all the cells of this frame can go into the buffer, otherwise all the cells of this frame will be rejected, thus preserving the buffer and the network from useless and corrupted traffic. The output

link of the buffer transmits the cells with constant bit rate (smoothing peaks) despite the variable input rate.

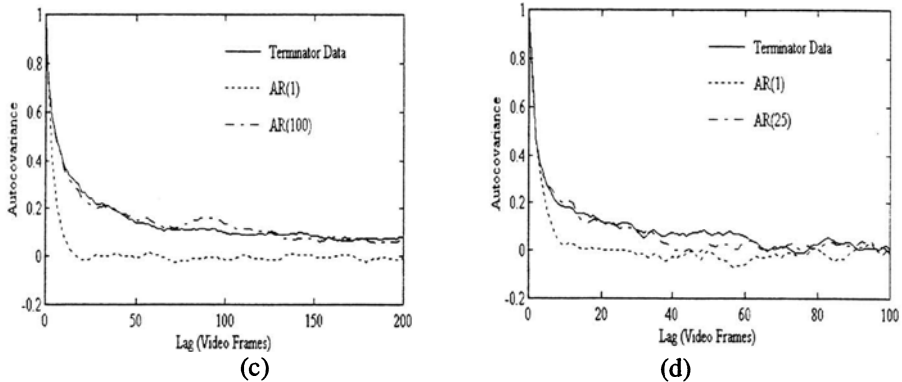


Figure 5: Autocovariance function of (a) B and (b) GOP frames (Term. Seq.).

GOP layer

At GOP layer, our attempt is to find a good fit for cell as well as frame loss probability based on the behaviour of GOP. The number of cells in each GOP frame is computed using high order AR and pdf models and transmitted to the buffer. We additionally assume that in a cyclic frame pattern all frames have the same size which is equal to the GOP size. Figures 6, 7 show the MPEG stream cell and frame loss probability respectively at different utilisation for the Terminator film. In these Figures we have not illustrated the pdf models because they lead to extreme under estimates of the loss probabilities due to the independence of each generated GOP frame. Therefore the pdf models are not convenient to approach the traffic behaviour of MPEG-I.

Although we used 25th order AR models in Figures 4 and 5 it is possible to reduce this order without significant effect in the loss probabilities. However, if the order of AR is very low (for example 1 or 2), the model will not approximate the loss probabilities especially for the very low ones. The choice of 25th order is a consequence of the good fit of autocovariance function as we have stated in the previous section. In Figures 6, 7 we also present the loss probabilities of the RealGOP (that is the GOP frames produced by the real video data) so as to show that a high order AR model of GOP can approach the RealGOP.

We can conclude that an AR model for GOP frames is a convenient way to estimate the cell loss probability of an MPEG video data stream. However, this model overestimates the frame loss. This occurs because the GOP size has neither the large values of I frames nor the small values of P and B. Since it presents the same video

activity as the real sequence (Figure 1), it can approximate the cell loss probability but not the number of the lost frames.

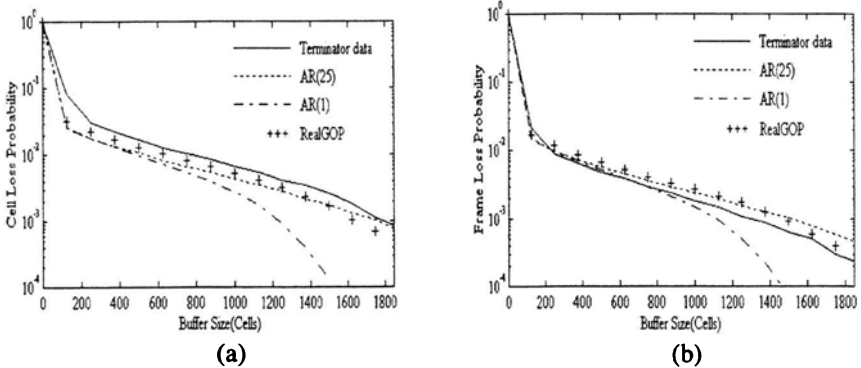


Figure 6 : Comparison of cell loss (a) and frame loss probability (b) - AR(1) and AR(25) of GOP ($U=0.75$) .

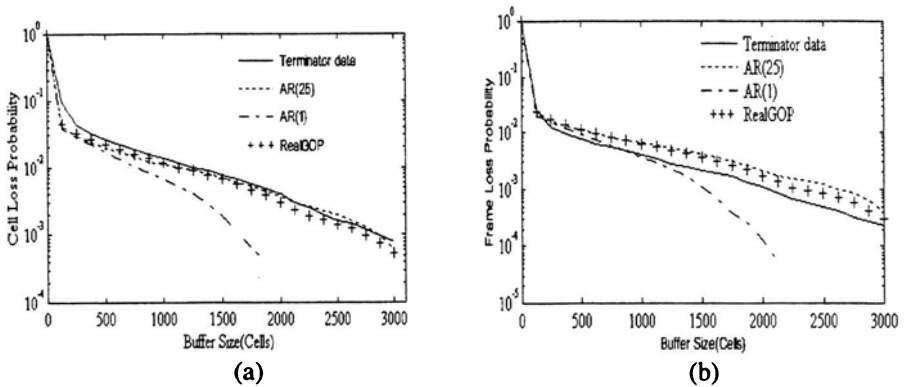


Figure 7 : Comparison of cell loss (a) and frame loss probability (b) - AR(1) and AR(25) of GOP ($U=0.8$).

Frame layer

In frame layer, AR and pdf models cannot give a good estimate of loss probabilities. As we have mentioned before there is a cross-correlation among I, P, and B frames due to the MPEG algorithm. This means that a scene change or a camera zooming causes large volumes for I, P and B frames and overflows the buffer, since a large I frame is usually followed by large values of P and B frames. If we produce models for I, P, and B frames (i.e. AR) and then merge them together, it is difficult for the total

sequence to approach the traffic behaviour. Since AR can estimate well I, P and B separately we are compelled to find a mechanism to achieve the required cross-correlation among I, P, and B frames.

The MPEG video sequences which we examine presents instant peaks in the bit rate (scene changes or camera zooming). In MPEG each scene is expected to differ from the previous or the following in its properties. Although the parameters of AR model have been found by the whole sequence (of I, P or B respectively), this model can estimate the different characteristics of each scene because of the error $e(n)$. When $e(n)$ takes values around the average the scene does not change. Otherwise, when $e(n)$ takes the remote values from the average, a scene change has been occurred. Since we have produced $e(n)$ according to the real pdf and not based on Gauss distribution the probability of scene changes is almost identical to the real sequence.

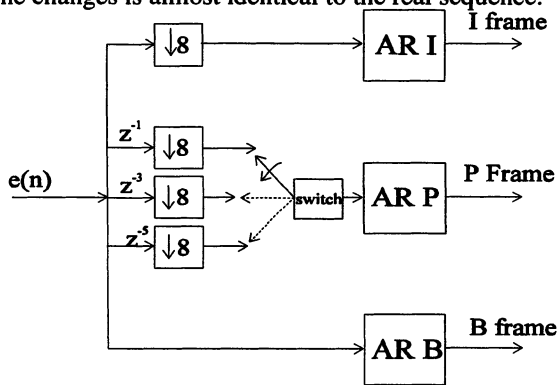


Figure 8 : The block diagram of the modulated AR.

To correlate I, P, and B frames together we propose a parallel mechanism of $e(n)$ which is shown in Figure 8. This procedure simply correlates the error of I, P and B AR models without changing the parameters a_i and b and without spoiling the i.i.d. characteristics of the $e(n)$, therefore the properties of AR model. We have only assumed that the distribution functions of I, P, and B frames are identical. The last condition appears to be very close to the real measurements experimental.

In figure 8 we have correlated the I and P_1 , B_2 and P_1 , B_4 and P_2 , B_6 and P_3 . This has relied on the operation of the MPEG algorithm. The input-output relation of a decimator is given by (more details can be found in (Vaidyanathan, 1990)). $y_d(n) = x(Mn)$, where M is a constant integer parameter. This equation states that the output at time n is equal to the input at time nM and as a result the size of $y_d(n)$ is M times smaller than the input $x(n)$.

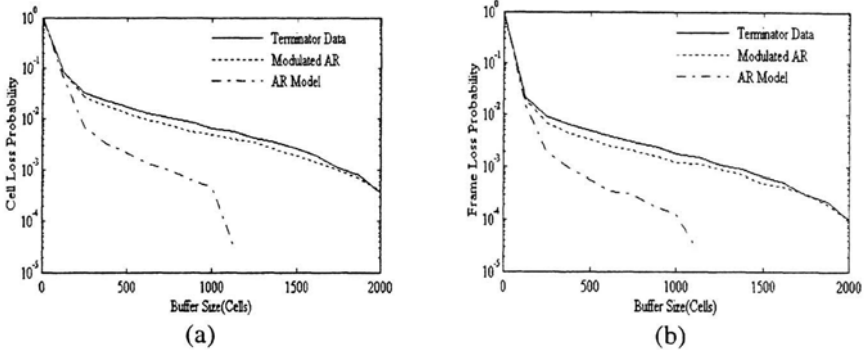


Figure 9 : Comparison of cell loss and frame loss probability - Modulated AR - AR model ($U=0.75$).

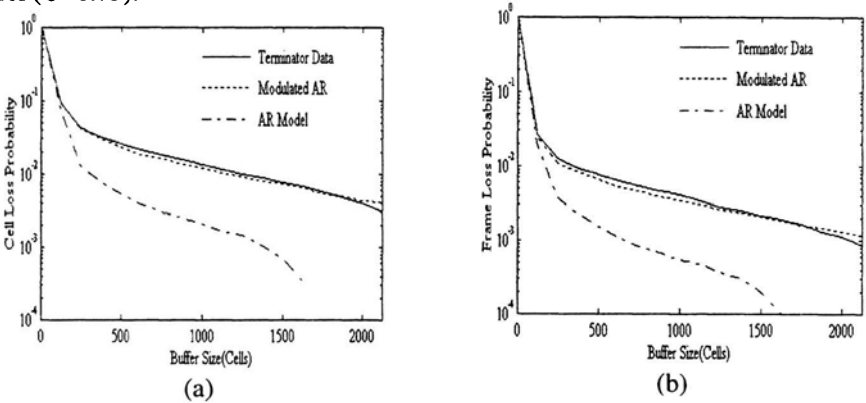


Figure 10 : Comparison of cell loss and frame loss probability - Modulated AR - AR model ($U=0.8$).

The error $e(n)$ that triggers AR I is produced by the error of AR B after the latter is delayed and decimated so that I frames correlate to B_2 ones. A switch is used to select the noise which triggers AR P. Therefore P_1 is related to B_2 , P_2 to B_4 and finally P_3 to B_6 . Since the pdfs of errors of separated frames (I, P, and B) are similar this procedure does not spoil the i.i.d. of the noise which is needed to produce the Yale-Walker (Papoulis, 1984) equations. We call this model modulated AR in the next and we see how it approximates the loss probabilities.

Figures 9, 10 illustrate the simulation of the buffer results using the previous modulated AR model at two different utilisations. In these Figures we also indicate the non modulated AR. As it is expected the behaviour of frame and cell loss probabilities are similar. This is quiet logical because our model is founded on the frame layer.

CONCLUSION

Development of models for MPEG sequences is an important task since these sequences will be an major component over B-ISDN. In this paper, we studied both the statistical characteristics and the traffic behaviour of VBR MPEG sources.

Initially, we observed that although B and P frames have small average bit rate, they play a significant role to the traffic behaviour. We have also seen that there is a strong dependence on I, P and B frames within a cyclic pattern frame that affects the loss probabilities. The pdf of I, P, B and GOP seems to follow a Gamma distribution function (in discrete time negative bipolar) or a Lognormal one. However Lognormal approximates better B and GOP frames while Gamma I, P frames. The total sequence cannot be estimated accurately by a known distribution function since it is not fit well at the large values of frames.

Moreover we observed that the autocovariance function of an MPEG video source presents periodical peaks with very slow decay. Only I, P, B, and GOP separately can be modelled based on linear systems such as AR. It seems that the autocovariance functions of all frames have a long range dependence therefore a high order AR is required to fit them well. Besides, the histogram of error $e(n)$ is not a gaussian distribution and affects strongly the loss probabilities.

At GOP layer a simple AR model of high order can approximate the cell loss probabilities but slightly overestimates the frame loss ones especially when high utilisation is used. An AR(1) underestimates the probabilities when buffer size is large (low loss probability). At frame layer a modulated AR model is proposed to achieve the cross-correlation among I, P and B frames and hence to approach the traffic behaviour. This model satisfies well both the cell and frame loss probabilities.

Acknowledgement We would like to mention that these films have been taken from <ftp-info3.informatik.uni-wuerzburg.de/pub/MPEG>.

REFERENCES

- Frater, M. R., Arnold, J. F and Tan, P. (1994) A new Statistical Model for Traffic Generated by VBR Coders for Television on the Broadband ISDN. *IEEE Trans. On Circ. and Syst. for Video Technol.*, vol. 4, pp.521-526.
- Grunenfelder, R, Cosmas, J, Manthorpe, S and Odima-Okafor, A.. (1991) Characterization of Video Codecs as ARMA Process and Related Queuing System Performance. *IEEE J. Selected Areas in Communications*, vol. 9, no. 3, pp. 284-293.
- Heeke, H. (1993) A Traffic Control Algorithm for ATM Networks. *IEEE Trans. On Circ. and Syst. for Video Technol.*, vol. 3, pp.183-189.

- Heyman, D, Tabatabai, A. and Lakshman, T. V. (1992) Statistical analysis and simulation study of video teleconference traffic in ATM networks. *IEEE Trans. On Circ. and Syst. for Video Technol.*, vol. 2, pp.49-59.
- Heyman, D. P, Tabatabai, A and Laskman, T. V. (1994) Statistical Analysis of MPEG2-Coded VBR Video Traffic, PV '94 B2.1
- ISO/IEC 13818-2. (1994) Generic Coding of Moving Pictures and Associated Audio. Recommendation H.262, Committee Draft.
- Jacobs, P. A and Lewis, P. A. w. (1983) Time Series Generated by Mixtures J. Time Series Analysis, vol 4, no.1, pp. 19-36.
- Kobayashi, H (1981) Addison-wesley Modeling and Analysis.
- Liou, M. (1991) Overview of the px64Kbit/sec Video Coded Standard. Communications of ACM , vol. 34, no. 4, pp 59-63.
- Maglaris, B, Anastassiou, D, Sen, P, Karlsson, G and Robbins, J. D. (1988) Performance models of statistical multiplexing in packet video communication. *IEEE Trans. Communications*, vol. 36, pp.834-843.
- Nasser , Marafih, M., Ya-Qin Zhang, and Pickholtz, R. (1994) Modeling and Queuing Analysis of Variable-Bit-Rate Coded Video Sources in ATM Networks. *IEEE Trans. On. Circ. and Syst. for Video Technol.*, vol. 4, pp.121-128.
- Pancha, P. and Zarki, M. E. Bandwidth-Allocation Schemes for Variable-Bit-Rate MPEG Sources in ATM Networks. *IEEE Trans. On. Circ. and Syst. for Video Technol.*, vol. 3, pp.1991.
- Papoulis, A. (1984) McGraw Hill Probability, Random Variables, and Stochastic Processes, New York.
- Sen, P, Maglaris, B, Rikli, N and Anastassiou, D. (1989) Models for packet switching of variable bit rate video sources. *IEEE J. Selected Areas Commun.*, vol. 7, pp. 865-869.
- Vaidyanathan, P. P. (1990) Multirate Digital Filters, Filter Banks, Polyphase Networks, and Applications: A Tutorial. *Proceedings of the IEEE*, vol. 78, no. 1.

Anastasios Doulamis was born in Athens, on September 14 1972. He received B.S. degree in electrical engineering from National Technical University of Athens (N.T.U.A.) with the highest honor (1995). Now, he works towards his Ph.D thesis in N.T.U.A. He is author for more than five articles. He has received several awards during his studies - diploma thesis honor (1997), best Greek student award during the academy year 1994-1995 (1997), student awards for being upper 1% of electrical engineering students ('94, '95,'96) by Technical Campus of Greece (1997). He has

been also awarded three times by the Greek Government. His interests focuses ATM networks, communication theory, low bit rate coding and neural network classification.

Nikolaos Doulamis was born in Athens, on September 14 1972. He received B.S. degree in electrical engineering from National Technical University of Athens (N.T.U.A.) with the highest honor (1995). Now, he works towards his Ph.D thesis in N.T.U.A. He has published more than six papers and has been several times awarded by Technical Campus of Greece and Greek Government (diploma thesis honor (1997), best Greek student award during the academy year 1994-1995 (1997), student awards for being upper 1% of electrical engineering students ('94, '95,'96)). His interests focuses on video transmission over high speed networks, image processing and coding.

G.E. Konstantoulakis: Born in Chania, Greece (1968) and obtained the Degree in EE from NTUA in 1992. In 1992 he joined the Telecommunications Laboratory of NTUA where he finished his thesis and now finishes his Ph.D. His research interests are in the area of broadband communication networks, high speed-real time architectures and algorithms, performance evaluation. He has over fifteen publications in the above areas. Member of the IEEE and the Technical Chamber of Greece by which he took the reward as the new engineer with the best graduate thesis at national level. He was nominated for the young engineer award of the RACE I and II programmes. He is fluent in English, Greek.

G.I. Stassinopoulos. Born 1951 in Athens Greece he has obtained his degree in EE from ETH-Zuerich (1974) and his Ph.D. from Imperial College London (1971). Since 1981 he is with the National Technical University of Athens (NTUA), where he is now professor at the Division of Computer Science. He is interested in broadband networks services and applications and has worked in the ATM area since 1986. He has over 50 publications in the above and related areas and is presently chairman of IEEE Section Greece. He is fluent in Greek, English, German and France.