Journal of the Operations Research
Society of Japan
Vol. 42, No. 1, March 1999

A MARKOVIAN MODEL OF CODED VIDEO TRAFFIC WHICH EXHIBITS LONG-RANGE DEPENDENCE IN STATISTICAL ANALYSIS

Toshiyasu Kurasugi Kazutomo Kobayashi
NEC Corporation

Yukio Takahashi
Tokyo Institute of Technology

(Received April 30, 1997; Final August 3, 1998)

Abstract The purpose of this paper is to construct a Markovian model generating a sequence having almost the same statistical characteristics as a real video traffic process. We deal with measured traffic data from a certain video source. Taking scene changes into account we analyze the data and construct a model, called Markov-AR model, composed of three submodels; a Markov transition model for scene changes, an AR model for spikes of scenes, and an AR model with random parameters for bit rate sequences of individual scenes. A simulation study shows that statistical characteristics of a sequence generated by this model are very similar to the actual video traffic. Especially, in the variance-time analysis and in the R/S analysis, these two sequences give similar estimates for the Hurst parameters and exhibit long-range dependence even though the model consists of only Markovian type, short-range dependent processes.

1. Introduction

Video traffic will be the main traffic in communication networks in the future and understanding its statistical characteristics is one of the most important and urgent tasks in designing future networks.

Long-range dependence is said to be an inherent feature of processes of video traffic. A wide-sense stationary stochastic process $\{W(n)\}$ with autocorrelation function $\gamma(k)$ is *long-range dependent* if

$$\gamma(k) \sim k^{2H-2} L(k), \qquad \text{as } k \to \infty, \tag{1.1}$$

where $1/2 < H < 1$ and $L(k)$ is a slowly varying function in the sense $\lim_{t\to\infty} L(tx)/L(t) = 1$, for all $x > 0$ [2]. H in (1.1) is called the *Hurst parameter*. The autocorrelation function of a long-range dependent process is non-summable, i.e., $\sum_{n=0}^{\infty} \gamma(k) = \infty$. On the contrary, a wide-sense stationary stochastic process $\{W(n)\}$ is *short-range dependent* if the autocorrelation function decays geometrically as $\gamma(k) \sim c\rho^{-k}$ where $0 < \rho < 1$. The autocorrelation function of a short-range dependent process is summable, i.e., $\sum_{k=0}^{\infty} \gamma(k) < \infty$. The discussion on the long-range dependence started with a paper [1] by J. Beran et al. They statistically analyze 17 data sets of actual VBR (variable bit rate) video traffic using variance-time analysis and R/S analysis, and find that all of them show the long-range dependence. Relying on this result, they insist that the long-range dependence is an inherent characteristic of VBR video traffic. Along this line some processes which have long-range dependent natures are proposed in order to model video traffic [1, 6, 11].

In this paper, a Markovian type model is constructed so that it generates a sequence whose statistical characteristics are almost the same as the actual video traffic data. In another words, the authors try to construct a model so that it generates a sequence as close as possible to the original one in the statistical analyses. Then in the variance-time and

R/S analyses, we see that the generated sequence as well as the actual traffic that exhibits long-range dependence though the model consists of only short-range dependent processes. This fact will help us to consider practical meanings of the long-range dependence in the traffic analyses and performance evaluations of information networks.

We deal with actual traffic data extracted from an entertainment video. It is coded by a variant protocol of H.261 which is modified so that the sequence becomes a VBR one. The data sequence has spikes at moments of scene changes and statistical behavior differs from scene to scene. First we classify scenes into five types by camera work and by degree of motion of the subjects in the pictures, and then we examine their statistical properties for each scene type. Based on the results, we develop a model consisting three submodels, a Markovian scene transition model, a spike model, and an AR model for bit rates following the spike. We will refer to this model as *Markov-AR model*. We generate a sequence by this model and analyze it for comparing its statistical characteristics with those of the original traffic data. The comparison includes marginal distribution and autocorrelation function as well as results of the spectral analysis, the variance-time analysis and the R/S analysis. These results show that the simulated sequence has almost the same statistical characteristics as the actual one. The estimated values of the Hurst parameter are almost the same, and hence both sequences exhibit long-range dependence.

A number of Markovian type models are proposed so far for video traffic. Teleconference traffic is modeled in [3, 4, 8, 24], though this kind of traffic has no scene changes. In [24] and [9], entertainment video traffic is modeled and scene changes are taken into models. The model introduced in [24] is a motion-classified AR model and resembles our Markov-AR model. But our model is more sophisticated and tuned up so that it generates sequences having closer statistical properties. It is also reported in [10] that a Markovian type model might generate a sequence which shows long-range dependence in the variance-time and R/S analyses.

The remainder of the paper is organized as follows. In Section 2 we explain our original data, and in Section 3 we introduce the Markov-AR model. Statistical results of the analysis of scenes in the original source are summarized in Appendix A. In Section 4 statistical characteristics of a sequence generated by the model are compared with those of the original sequence. Finally, several simplified models of the Markov-AR model are introduced and estimated values of the Hurst parameter in the variance-time and R/S analyses are compared in Section 5.

2. VBR Video Traffic Data

We analyze traffic data taken from a video of an American travel movie. To capture statistical characteristics of the video traffic, it is compressed by a modified protocol of H.261 in which intrafield/interframe DCT (Discrete Cosine Transform) and motion compensation are used and smoothing of bit rates is removed. The resultant sequence becomes to a VBR sequence.

We denote the sequence by $\{x(n)\,;\, n = 1, 2, \cdots, N\}$ where $x(n)$ is the bit rate of the n-th frame in Mbps. The value of $x(n)$ ranges from 0 to 6.114 Mbps and the number of frames of the sequence is $N = 79686$. The frame interval is $1/30$ seconds. So the time length of the whole sequence is about 43 minutes. Figure 1 shows a part of the sequence. Referring to the original video pictures we see spikes appear at moments of scene changes in the video.

We divide the sequence $\{x(n)\}$ into subsequences corresponding to scenes as $\{\{x_k(n)\,;\, n = 1, 2, \cdots, N_k\}\,;\, k = 1, 2, \cdots, K\}$, where K is the number of scenes. We will refer

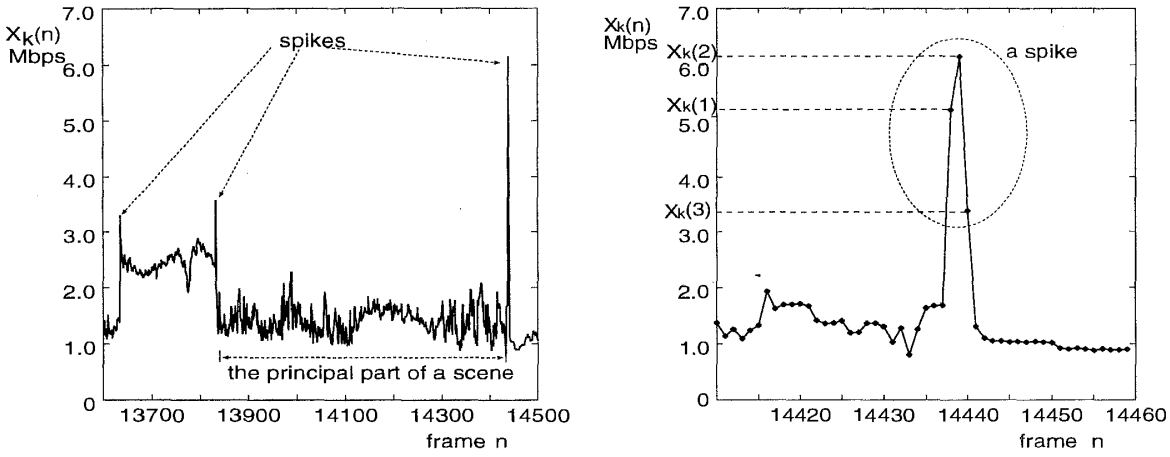


Figure 1: Scenes and spikes

to the subsequence $\{x_k(n)\,;\, n = 1, 2, \cdots, N_k\}$ as the *k-th scene sequence*. In each scene sequence, a few frames at the head of the sequence are seen to form a *spike*. We denote by $\{x_k(1), \cdots, x_k(M_k)\}$ the subsequence forming the spike, and refer to the remainder $\{X_k(n)\,;\, n = M_k+1, \cdots, N_k\}$ as the *principal part* of the scene sequence. There are $K = 163$ scenes in the video, but for statistical analysis of scenes we use first 100 scenes covering 43526 frames.

We observe that the behavior of a scene sequence $\{x_k(n)\}$ largely depends on camera work and on the degree of motion of subjects in the pictures of the scene. The 100 scenes are classified into five scene types:

Type 1 : *Scene with subjects moving slightly* (26 scenes)
Type 2 : *Scene with subjects moving mildly* (31 scenes)
Type 3 : *Scene with subjects moving greatly* (16 scenes)
Type 4 : *Scene with a moving camera* (9 scenes)
Type 5 : *Scene with a panning camera* (18 scenes)

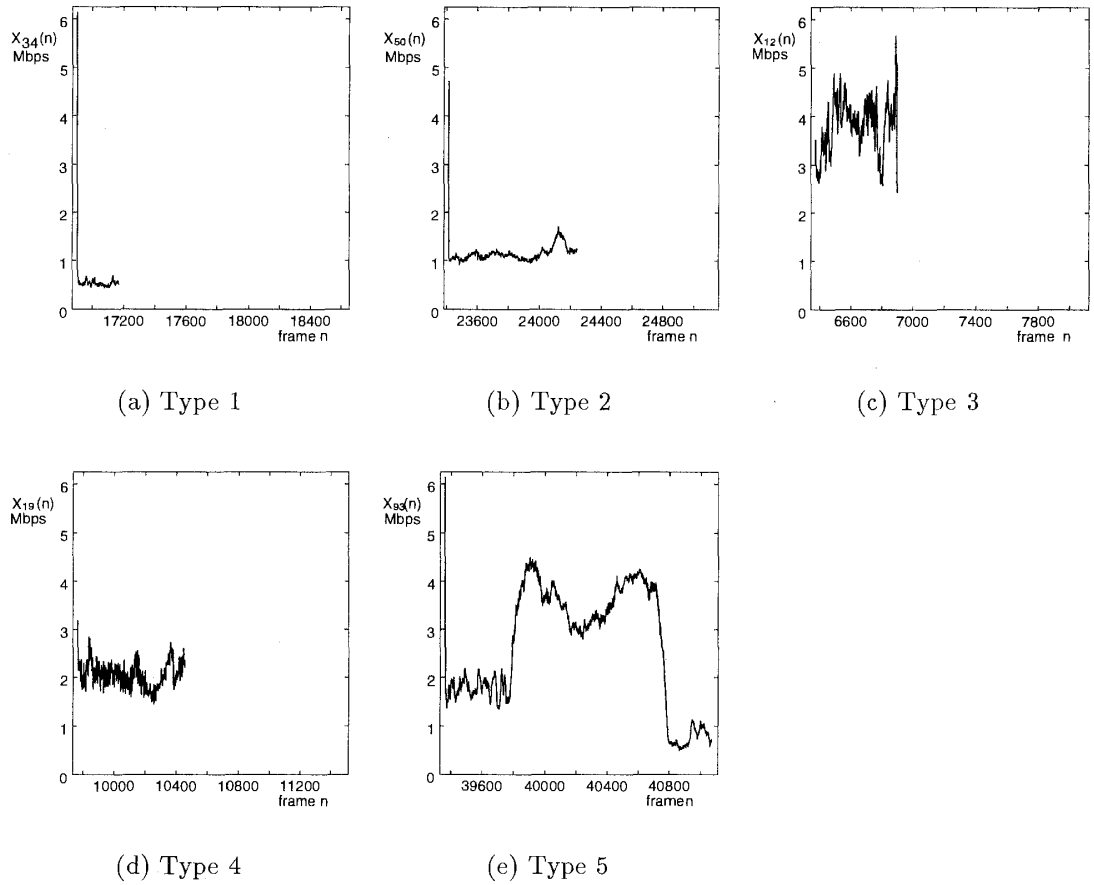
Figure 2 shows an example of $\{x_k(n)\}$ for each of the five scene types. The graphs exhibit different patterns of the principal parts; namely, a flat curve in (a), a mountain-like curve in (e), and waved curves in (b), (c), and (d). Actually, (a) represents the data sequence of a scene of refueling at a gas station, (b) a scene of a ship coming alongside a pier, (c) a scene at a sports meeting in an elementary school, (d) a scene of a town from a moving car, and (e) a scene of a restaurant with many people captured by a panning camera.

We analyze these scene sequences statistically for each type. The results are summarized in Appendix A. For more detailed descriptions, refer to [15, 16].

3. Markovian Model for VBR Video Traffic

Based on the statistical analysis of scene sequences, we introduce a Markovian model which we call Markov-AR model for VBR video traffic. Our aim is to construct a Markovian model which generates a sequence $\{y(n)\,;\, n = 1, 2, \cdots, N\}$ having almost the same statistical characteristics as those of the actual video traffic sequence $\{x(n)\}$. For the purpose we use parameters of the model estimated from the actual sequence.

Our model consists of three submodels; a Markovian scene transition model, a spike model and an AR model for bit rates following the spike. We denote by $\{Y(n)\}$ the process



(a) Type 1 (b) Type 2 (c) Type 3

(d) Type 4 (e) Type 5

Figure 2: Examples of scene sequences $\{X_k(n)\}$

generated from the model and by $\{Y_k(n)\,;\, n = 1, 2, \cdots, N_k\}$ the subsequence of $\{Y(n)\}$ corresponding to the k-th scene for $k = 1, 2, \cdots, K$. To abbreviate descriptions, we use the same symbols N, K, and N_k here as in the actual traffic sequence though their values might be different. In the subsequent model description, all random variables generated are assumed to be mutually independent unless otherwise stated.

Markov transition model for scene types

Successive scene types are selected according to a Markov chain with a transition probability matrix

$$R = (r_{ij}) = \begin{bmatrix} .308 & .308 & .000 & .192 & .192 \\ .258 & .355 & .129 & .032 & .226 \\ .250 & .188 & .562 & .000 & .000 \\ .222 & .111 & .111 & .222 & .333 \\ .176 & .471 & .118 & .059 & .176 \end{bmatrix}, \tag{3.1}$$

which is estimated from the actual data.

Spike model

The spike length M_k in the actual data ranges between 0 and 5, and it is equal to two or three in most of scenes. The maximum value x_k^{max} among the k-th spike sequence $(x_k(1), \cdots, x_k(M_k))$ is taken mostly at the second frame and its value distributes almost

uniformly over the interval $(A_k, 6.144)$, where A_k is the average bit rate of the principal part of the k-th scene.

Relying on these observations, we introduce a spike model as follows. We set the spike length constant and equal to three. Hence the k-th spike sequence is written as $(Y_k(1), Y_k(2), Y_k(3))$. First we choose the average bit rate of the principal part A_k randomly according to a scene-type dependent gamma distribution with mean μ_{Ai} and standard deviation σ_{Ai} given in Table 1, where i represents the scene-type of the k-th scene. Then the second value $Y_k(2)$, the largest one among the three, is randomly selected according to the uniform distribution over the interval $(A_k, 6.144)$, and the other two are given by $Y_k(1) = 0.83\,Y_k(2) + 0.17\,A_k$ and $Y_k(3) = 0.42\,Y_k(2) + 0.58\,A_k$. In these equations parameters are taken from the results of the analysis of actual data.

AR model for principal part

We observe that principal parts of scene sequences for scenes of Types $1 \sim 4$ fit well to AR (auto-regressive) processes. Hence we shall use AR models for them. The movements of bit rates of Type 5, however, are large and not stable. It seems difficult to derive a suitable model which is simple and generates similar sequences to them. Since our aim is to construct a model which generates a sequence having almost the same statistical characteristics as the actual sequence, we shall give up seeking such a well-fitted model for Type 5 sequences and be satisfied with using an AR model.

An AR model for the k-th scene sequence requires these parameters: the mean bit rates A_k, the length of the sequence N_k, the order of the regression m_k, the coefficients of the regression $a_{k,j}$, and the standard deviation of the error terms σ_k. If these parameters are chosen and if the initial values $Y_k(4), \cdots, Y_k(m_k+3)$ are given, the remaining of the principal part of the k-th scene $(Y_k(m_k+4), \ldots, Y_k(N_k))$ is generated according to the relations

$$\begin{aligned} Z_k(n) &= \sum_{j=1}^{m_k} a_{k,j} Z_k(n-j) + \epsilon_k(n), \\ Y_k(n) &= Z_k(n) + A_k, \qquad \text{for} \quad n = m_k+4, \ldots, N_k, \end{aligned}$$

where $\epsilon_k(n)$, $n = m_k+4, \ldots, N_k$, are independent random variables subjected to the normal distribution with mean 0 and variance σ_k^2.

The parameters and the initial values are selected in the following manner. A_k has already been selected at the use of the spike model. m_k and $a_{k,j}$'s are selected randomly not depending on the scene-type because the difference of the distribution among types are not observed. In the following, index i stands for the type of the k-th scene.

- N_k is randomly selected according to a scene-type dependent gamma distribution with mean μ_{Ni} and standard deviation σ_{Ni} given in Table 1.
- m_k is randomly selected from two possible values, 1 with a probability of 17/81 and 2 with a probability 64/81.
- $a_{k,j}$, $j = 1, \cdots, m_k$, are generated as follows.
  1. When $m_k = 1$, $a_{k,1}$ is randomly selected so that $1 - a_{k,1}$ is subjected to the exponential distribution with a mean of 0.141.
  2. When $m_k = 2$, first $a_{k,2}$ is randomly selected according to the normal distribution with mean 0.010 and variance 0.096^2, and then $a_{k,1}$ is randomly selected so that $1 - (a_{k,1} + a_{k,2})$ is subjected to the exponential distribution with mean 0.216.

Table 1: Estimated parameters of gamma distributions

	Scene length		Average bit rate		v_k for standard deviation	
Scene type	μ_{Ni}	σ_{Ni}	μ_{Ai}	σ_{Ai}	μ_{vi}	σ_{vi}
Type 1	207	110	0.573	0.446	.0595	.0366
Type 2	413	320	1.161	0.596	.0528	.0523
Type 3	374	295	3.616	0.934	.0535	.0425
Type 4	434	145	1.856	0.788	.0799	.0307
Type 5	859	364	2.364	1.517	.0528	.0523
Scene type	μ_{Ni}	σ_{Ni}	μ_{Ai}	σ_{Ai}	μ_{vi}	σ_{vi}
Type 1	207	110	0.573	0.446	.0595	.0366
Type 2	413	320	1.161	0.596	.0528	.0523
Type 3	374	295	3.616	0.934	.0535	.0425
Type 4	434	145	1.856	0.788	.0799	.0307
Type 5	859	364	2.364	1.517	.0528	.0523

- For σ_k we first randomly select a value v_k according to a scene-type dependent gamma distribution with mean μ_{vi} and standard deviation σ_{vi} given in Table 1. Then we set $\sigma_k = v_k\sqrt{A_k} + 0.02A_k$. In Table 1, μ_{v5} and σ_{v5} for Type 5 are set equal to μ_{v2} and σ_{v2} for Type 2 since the magnitudes of fluctuations of bit rates are similar between the two types.
- The initial values $Y_k(4), \cdots, Y_k(m_k+3)$ are randomly selected according to the normal distribution with mean A_k and variance σ_k^2.

4. Comparison Between Actual Data and Simulated Data

The purpose of this section is to show that a sequence $\{y(n)\}$ generated by the Markov-AR model in the preceding section has similar statistical characteristics to the actual sequence $\{x(n)\}$, and also to especially demonstrate that both of the sequences exhibit the long-range dependence in the variance-time analysis and the R/S analysis. In this section, we do such a statistical analysis for the actual sequence $\{x(n)\,;\, n = 1, 2, \cdots, 79686\}$ and a generated sequence $\{y(n)\,;\, n = 1, 2, \cdots, 79686\}$ of the same length regarding them as realizations of stationary processes $\{X(n)\}$ and $\{Y(n)\}$. The logarithms used in this section are all common logarithms, i.e. the bases of logarithms are all 10.

4.1. Marginal distributions

Figure 3 compares marginal distributions of the sequences $\{x(n)\}$ and $\{y(n)\}$. To get a better estimate, we use a prolonged sequence of the length 2×10^6 for $\{y(n)\}$. We see that the simulated sequence has less mass at small bit rates and at the upper limit of 6.144 Kbps. This is because we use an AR model for principal part of scene sequence of Type 5.

4.2. Autocorrelation function

The autocorrelation function of a stationary process $\{W(n)\}$ is estimated from a data sequence $\{w(n) : n = 1, \cdots, N\}$ by the equation

$$\hat{\gamma}(k) = \sum_{n=1}^{N-k} \frac{(w(n) - \hat{\mu})(w(n+k) - \hat{\mu})}{N\hat{\sigma}^2},$$

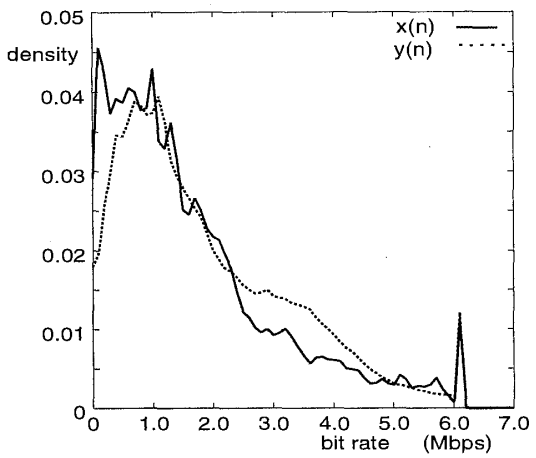


Figure 3: Marginal distributions

where

$$\hat{\mu} = \sum_{n=1}^{N} \frac{w(n)}{N}, \qquad \hat{\sigma}^2 = \sum_{n=1}^{N} \frac{(w(n) - \hat{\mu})^2}{N}.$$

In Figure 4 autocorrelation functions estimated from $\{y(n)\}$ and $\{x(n)\}$ are presented together with the 95% confidence limits of $\gamma(k)$ for $\{x(n)\}$. The confidence limits indicate that the null hypothesis $\gamma(k) = 0$ is rejected if $\gamma(k)$ is out of it [25].

In the region $k < 1000$, the estimated two autocorrelation functions are almost the same. In the region $k > 1000$, they are much different but they are within the confidence belt. This suggests that it is difficult to discuss asymptotic behaviors of autocorrelation functions for large k, and consequently we cannot judge whether the processes $\{X(n)\}$ and $\{Y(n)\}$ are long-range dependent or not. Hence we check the long-range dependence in the variance-time analysis and the R/S analysis.

4.3. Variance-time analysis

One of the methods commonly used for testing the long-range dependence is the variance-time analysis [1, 2]. For a stationary process $\{W(n)\}$, we introduce a stationary process $\{W^{(m)}(n)\}$ for each $m = 1, 2, \ldots$ by averaging the original process over nonoverlapping blocks of size m:

$$W^{(m)}(n) = \frac{1}{m} \sum_{h=nm-m+1}^{nm} W(h), \qquad n = 1, 2, \ldots.$$

If $\{W(n)\}$ is long-range dependent with autocorrelation function (1.1), then $(\sigma^{(m)})^2 = \mathrm{Var}(W^{(m)}(n)) \sim m^{2H-2} L_1(m)$ as $m \to \infty$, where $L_1(m)$ is a slowly varying function. If the process is short-range dependent, the variance decays as the reciprocal of m.

In the variance-time analysis, we calculate the estimates $(\hat{\sigma}^{(m)})^2$ of the variances $(\sigma^{(m)})^2$ from the data, and then plot $\log(\hat{\sigma}^{(m)})^2$ against $\log m$. Usually the plots decreases linearly and we can get an estimate $\hat{H}^{vt}$ of the Hurst parameter from the gradient of the line.

Figure 5 shows the normalized variance-time plots of $\{x(n)\}$ and $\{y(n)\}$, where normalization is done by the variance of the original sequence. We estimate the Hurst parameter by the least square method using points in the region $2.6 < \log(m) < 4$. Then

$$\hat{H}_X^{vt} = 0.68 \qquad \text{and} \qquad \hat{H}_Y^{vt} = 0.64$$

for $\{X(n)\}$ and $\{Y(n)\}$.This result indicates both processes are long-range dependent though the estimated values of the Hurst parameter are not so large.

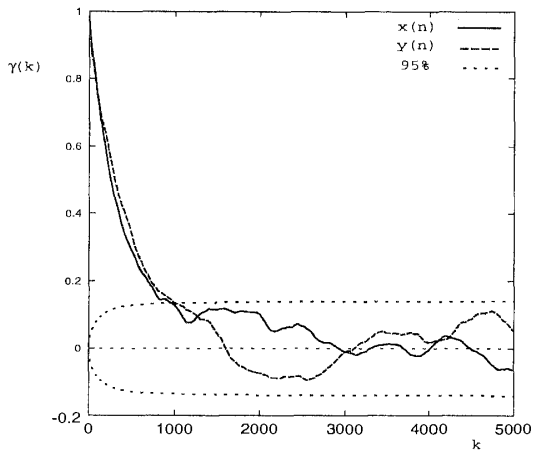


Figure 4: Autocorrelation functions

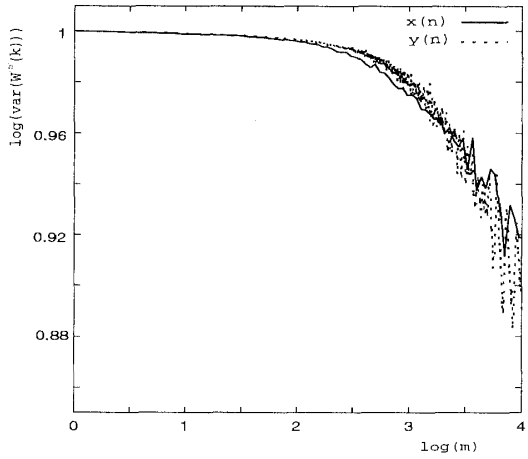


Figure 5: Variance-time plots of $\{x(n)\}$, $\{y(n)\}$

4.4. R/S analysis

Another commonly used method for checking the long-range dependence is the R/S analysis [1, 2]. Let $\overline{W}(u)$ be the average and $S(u)$ the standard deviation of the first u values, $W(1), \cdots, W(u)$, of a stationary process $\{W(n)\}$, and introduce a statistic

$$\frac{R(u)}{S(u)} = \frac{1}{S(u)} \left[\max\{0, V_1(u), V_2(u), \cdots, V_u(u)\} - \min\{0, V_1(u), V_2(u), \cdots, V_u(u)\} \right], \quad (4.1)$$

where $V_i(u) = \sum_{n=1}^{i} \{W(n) - \overline{W}(u)\}$, $1 \leq i \leq u$. We refer to the function $R(u)/S(u)$ in (4.1) as the R/S *function*. If the process is long-range dependent, then the expectation of the statistic $R(u)/S(u)$ asymptotically decays as $u^H L_2(u)$. In the R/S analysis, we estimate the R/S function from a data sequence $\{w(n)\}$ and plot $\log \hat{R}(u)/\hat{S}(u)$ against $\log u$. The gradient of a fitted line is used as the estimation of the Hurst parameter H.

Figures 6 and 7 show plots for the data $\{x(n)\}$ and $\{y(n)\}$, and the method of least squares for u in the region $2.6 < \log(u) < 4.5$ leads estimate values

$$\hat{H}_X^{R/S} = 0.88 \qquad \text{and} \qquad \hat{H}_Y^{R/S} = 0.90$$

for $\{X(n)\}$ and $\{Y(n)\}$. These values are near to 1 and clearly indicate that the sequences are long-range dependent.

4.5. Spectral analysis

The spectral density function $S_p(f)$ of a process $\{W(n)\}$ can be estimated from a data sequence $\{w(n)\}$ of length N through the periodogram in the following manner.

1. Using the Fast Fourier Transform (FFT), obtain the Fourier transform of the sequence $\hat{w}(f_j) = \sum_{n=1}^{N} w(n)\, e^{i2(n-1)\pi f_j}$, $j = 1, \cdots, N/2$, where $f_j = j/\Delta N$ and $\Delta = 1/30$.
2. Obtain the periodogram $I(f_j) = |\hat{w}(f_j)|^2/N$, $j = 1, \cdots, N/2$.
3. Estimate $S_p(f_k)$ by averaging $I(f_j)$'s with bandwidth $l/\Delta N$ centered at f_k.

If the process is long-range dependent, the behavior of the spectral density function $S_p(f)$ in the neighborhood of the origin is represented by the form $S_p(f) \sim L(f) f^{-\nu}$ where $0 < \nu < 1$ and $L(f)$ is slowly varying as $f \to 0$. The Hurst parameter is given by $H = (1 + \nu)/2$. On the contrary, if the process is short-range dependent then ν is zero. If $\nu \geq 1$ then the process is non-stationary since the autocorrelation function cannot exist[23].

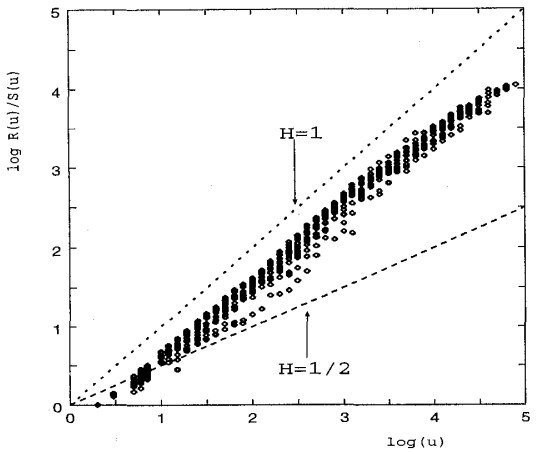


Figure 6: R/S function of $\{x(n)\}$

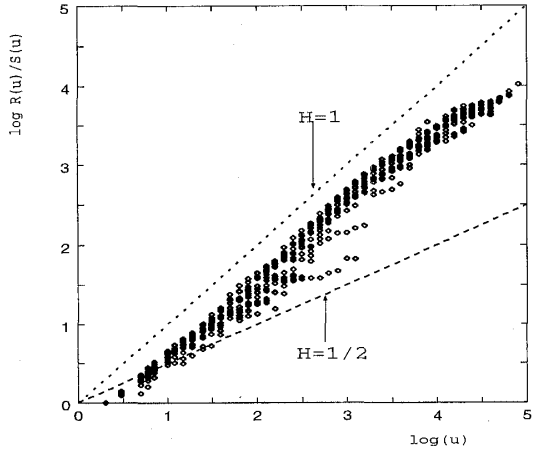


Figure 7: R/S function of $\{y(n)\}$

For the sequences $\{x(n)\}$ and $\{y(n)\}$, we set $N = 65536(= 2^{16})$ and $l = 32$ and plot $\log S_p(f)$ against $\log f$. The resulting curves are presented in Figure 8 and the spectral density functions of $\{x(n)\}$ and $\{y(n)\}$ are very similar to each other. The estimates of ν are greater than one. This means that the sequences are judged to be non-stationary sequences rather than stationary ones.

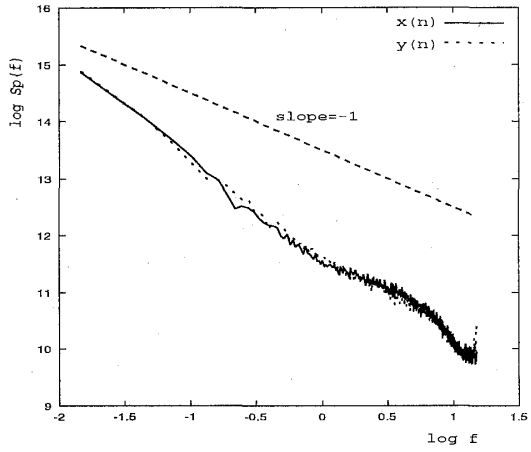


Figure 8: Spectral density functions of $\{x(n)\}$ and $\{y(n)\}$

5. Model Complexity and Long-Range Dependence

In order to see the reason why our Markov-AR model exhibits long-range dependence in the R/S analysis and the variance-time analysis, we introduce six simpler models and compare sequences generated by them. These models are constructed from the Markov-AR model by replacing submodels with the following simpler ones.

1. **Markov-AR**: This is our model introduced in Section 3. Hence the scene length is randomly selected and subjected to a gamma distribution, the bit rates of a spike are generated by the spike model, and bit rates of each scene following the spike are generated by the AR model.
2. **Markov-AR without spikes**: This model is the same with the Markov-AR model except that spikes are omitted.

Table 2: Comparison of estimated values of Hurst parameters

	scene length	spikes	bit rates	R/S	VT
1. Markov-AR	gamma	with	AR	0.90	0.64
2. Markov-AR without spikes	gamma	–	AR	0.82	0.68
3. SMMIR with spikes	gamma	with	MMIR	0.82	0.60
4. SMMIR without spikes	gamma	–	MMIR	0.84	0.42
5. MMIR 100	constant 100	–	MMIR	0.65	0.52
6. MMIR 1	constant 1	–	MC	0.55	0.56

3. **SMMIR with spikes**: This model is the same with the Markov-AR model except that the bit rates of principal part remain constantly equal to A_k for the k-th scene. The value A_k is randomly generated as the average bit rate in Section 3. This model is considered as a SMMIR (semi-Markov modulated input rate) model with a spike model.
4. **SMMIR without spikes**: This is a SMMIR model constructed from the above model by omitting spikes. The scene type changes according to the Markov chain (3.1), the scene length is randomly selected and subjected to a gamma distribution, and the bit rates are constantly equal to A_k though A_k is randomly generated and subjected to a scene-type dependent distribution.
5. **MC 100**: This is the same as the above model except that the scene lengths are not random variables but constantly equal to 100. This model can be considered as a Markov chain model with time unit equal to 100, or a semi-Markov process with constant sojourn times, though the state space is not finite but equal to $\{1,2,3,4,5\} \times (0,\infty)$.
6. **MC 1**: This is the same model as the above model except that the scene lengths are constant equal to 1. This is a typical Markov chain model.

The estimated values of the Hurst parameter for sequences generated by these models are listed in Table 2. We see that there is a tendency that the estimated Hurst parameter in the R/S analysis becomes larger if the variation of the scene length is large. Also there is a tendency that the estimated Hurst parameter in the variance-time analysis becomes larger if there is large variation of the bit rates in scenes. Thus, the R/S analysis and the variance time analysis might capture different features of the long-range dependence, and it is difficult to say whether a sequence is long-range dependent or not even if the estimated Hurst parameter becomes greater than $1/2$ in the variance-time and R/S analyses.

6. Concluding Remarks

In this paper we analyzed a sequence of traffic data of an entertainment video, and constructed a Markovian type model, called the Markov-AR model, which generates a sequence having almost the same statistical characteristics. The model is Markovian type and hence inherently short-range dependent. However the generated sequence showed the long-range dependence in the variance-time and R/S analyses. The estimated value of the Hurst parameter was 0.64 in the variance-time analysis and 0.90 in the R/S analysis. In the spectral analysis, the sequences were judged as non-stationary sequences, and we could not estimate the Hurst parameter.

In the paper by J. Beran et al. [1], and in other related papers, the variance-time analysis and the R/S analysis were used for judging if sequences were long-range dependent or

not. However, the above result shows that the long-range dependence may appear from short-range dependent models. This is not surprising since the definition of the long-range dependence is for a sequence of infinite length but sequences we deal with are of finite length.

In Section 5 we compared estimated values of the Hurst parameter for sequences from various models with different complexities. The comparison showed that more complex models, roughly speaking, lead larger estimated values and that the variance-time analysis and the R/S analysis detect different features of long-range dependence.

These results will help us to consider the practical meanings of the long-range dependence of video traffic. One of the main concerns of the authors' is how we should construct (queueing type) models for performance evaluation of information networks with video traffic. The first problem we encounter is how we model input processes which may exhibit long-range dependence in the variance-time and R/S analyses. Probably, traditional Markovian type models would be inadequate, since they are too simple and the sequences derived from them will not show long-range dependence. Another possibility is to use self-similar processes such as fractional Gaussian processes to model the input processes. However, at the moment, we don't have any confident reason for the use of such models. The results of this paper seem to indicate that models with scene change mechanisms are possible candidates though the number of parameters to be estimated is too large. Rather, we might seek another type(s) of models. This is a big theme for further study.

A. Statistical Analysis of Scenes

A.1. Transition of scene types

The frequency q_{ij} of transitions from Type i to Type j are given as in the following matrix. The transition probability matrix (3.1) is formed from it.

$$Q = (q_{ij}) = \begin{array}{c} \nearrow \\ \text{Type1} \\ \text{Type2} \\ \text{Type3} \\ \text{Type4} \\ \text{Type5} \end{array} \begin{array}{c} \begin{array}{ccccc} \text{Type1} & \text{Type2} & \text{Type3} & \text{Type4} & \text{Type5} \end{array} \\ \begin{bmatrix} 8 & 8 & 0 & 5 & 5 \\ 8 & 11 & 4 & 1 & 7 \\ 4 & 3 & 9 & 0 & 0 \\ 2 & 1 & 1 & 2 & 3 \\ 3 & 8 & 2 & 1 & 3 \end{bmatrix} \end{array},$$

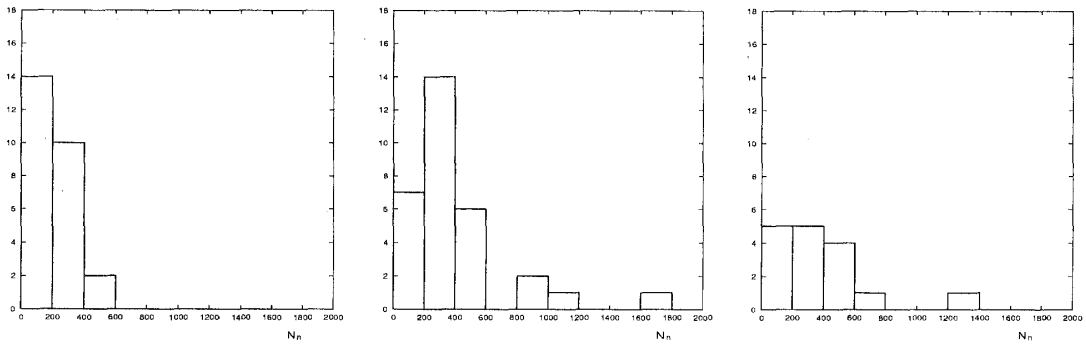
A.2. Scene length and average bit rate of principal part

The sample mean μ_{Ni} and the sample standard deviation σ_{Ni} of the scene length N_k of scene sequences of Type i ($i = i, \cdots, 5$) are given as listed in Table 1. The histogram of N_k is given in Figure 9. The figure suggests that a gamma distribution with mean μ_{Ni} and standard deviation σ_{Ni} fits well to the sample distribution of the scene length. The gamma distributions are accepted with 10% significance in a test of goodness of fit.

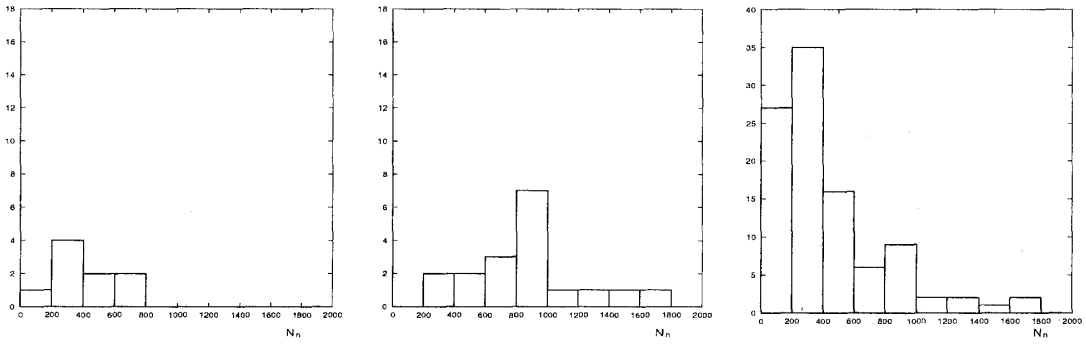
The sample mean μ_{Ai} and the sample standard deviation σ_{Ai} of average bit rate A_k of the principal parts of sequences for scenes of Type i ($i = i, \cdots, 5$) are also listed in Table 1. A gamma distribution with mean μ_{Ai} and standard deviation σ_{Ai} fits well to the sample distribution of A_k of Type i ($i = 1, \cdots, 5$). The gamma distributions are accepted with 10% significance in a test of goodness of fit.

A.3. Spike

For the k-th scene sequence, we identify the spike sequence in the following manner. Let A'_k and D'_k be the average and the standard deviation of the bit rates of the whole scene sequence of the k-th scene. We let M_k be the maximum number h (≤ 5) such that



(a) Type 1 (b) Type 2 (c) Type 3



(d) Type 4 (e) Type 5 (f) Total

Figure 9: Histogram of the scene length N_k

$x_k(n) > A'_k + 2D'_k$ for all n in the interval $[0, h]$, and we let $(x_k(1), \ldots, x_k(M_k))$ be the spike sequence of the k-th scene.

The spike length M_k ranges from 0 to 5. Among 100 scenes, the numbers of scenes having $M_k = 0, 1, 2, 3, 4$ and 5 are respectively 22, 34, 34, 3 and 2. The maximum value X_k^{max} of the bit rate in a spike sequence distributes almost uniformly between A_k and 6.144Mbps. This phenomenon is common among all scene types. If the spike length M_n is three, then $X_n(2)$ mostly takes the largest value and $X_n(3)$ takes the smallest value among the three.

When the spike length M_k is three, $X_k(1)$ and $X_k(2)$ are positively correlated and $X_k(1) - A_k$ is near to $0.83(X_k(2) - A_k)$. Also $X_k(2)$ and $X_k(3)$ are positively correlated and $X_k(3) - A_k$ is near to $0.42(X_k(2) - A_k)$.

A.4. Principal part

In Figure 2, we see that most of principal parts of scene sequences behave like AR processes. The only exception is the graph in Figure 2(e) of a scene of Type 5. Here we analyze 82 scenes which are of Types $1 \sim 4$.

An AR process $\{W_k(n) : n = M_k + 1, \ldots, N_k\}$ is represented in a form;

$$\begin{aligned} Z_k(n) &= \sum_{j=1}^{m_k} a_{k,j} Z_k(n-j) + \epsilon(n), \\ W_k(n) &= Z_k(n) + A_k, \end{aligned}$$

where m_k is order of the process, A_k is average rate, $a_{k,j}$'s are coefficients, and $\epsilon_k(n)$'s are independent and identically normally distributed random variables with zero mean and variance σ_k^2.

From a data sequence $(w_k(M_k+1),\ldots,w_k(N_k))$, we estimate the order m_k together with parameters A_k, σ_k and $a_{k,j}$'s by using the MDL (Minimum Description Length) principle [22]. The procedure is as follows. First we estimate A_k by the average bit rate $\hat{A}_k = \sum_{n=M_k+1}^{N_k} w_k(n)/(N_k - M_k)$, and set $z_k(n) = w_k(n) - \hat{A}_k$ $(n = M+1,\cdots,N)$. For coefficients $a_{k,j}$'s let m_k be a candidate order and $\hat{a}_{k,j}(m_k)$'s be the solution of the equations

$$\sum_{n=M_k+1}^{N_k} z_k(n-h)\left(z_k(n) - \sum_{j=1}^{m_k} \hat{a}_{k,j}(m_k) z_k(n-j)\right) = 0, \qquad h = 1,\cdots,m_k,$$

$\hat{a}_{k,j}(m_k)$'s are maximum likelihood estimates in the sense that they make the estimate

$$\hat{\sigma}_k(m_k)^2 = \sum_{n=1}^{N_k}\left(z_k(n) - \sum_{j=1}^{m_k} \hat{a}_{k,j}(m_k) z_k(n-j)\right)^2$$

of σ_k^2 minimum. Then the value of MDL for the candidate value m_k is defined by

$$MDL(m_k) = (N_k - M_k)\log \hat{\sigma}_k^2(m_k) + \frac{m_k+1}{2}\log(N_k - M_k). \tag{A.1}$$

An m_k which makes $MDL(m_k)$ smallest is chosen as the estimate $\hat{m}_k$ of the order for the process. We denote by $\hat{\sigma}_k$ and $\hat{a}_{k,j}$'s instead of $\hat{\sigma}_k^2(\hat{m}_k)$ and $\hat{a}_{k,j}(\hat{m}_k)$.

We estimate the parameters for the AR process fitted to the principal part of the k-th scene in the way stated above. The estimated order $\hat{m}_k$ ranges from 0 to 9.

Unfortunately, we can't find any specific tendency for $\hat{a}_{k,j}$'s and $\hat{\sigma}_k$. Probably the number of parameters is too large. So we make an estimation by restricting the order less than or equal to two. Let $\hat{m}'_k$ be the value of m_k which makes $MDL(m_k)$ in (A.1) smallest among $m_k = 0, 1$ or 2. Then there are 64 scenes for which $\hat{m}'_k = 2$, and 17 scenes for which $\hat{m}'_k = 1$. There is only 1 scene for which $\hat{m}'_k = 0$. We will denote the resulting estimates as $\hat{\mu}'_k$, $\hat{a}'_{k,i}$ and $\hat{\sigma}'_k$.

A.4.1. AR coefficients

For the AR coefficients $\hat{a}'_{k,i}$'s, we don't observe any difference among distributions for four scene types. However the distributions are much different in the cases with $\hat{m}'_k = 1$ and $\hat{m}'_k = 2$. Hence we discuss here the distribution of AR coefficients with classification by $\hat{m}'_k$ but without classification by the scene type.

Figure A.4.1 shows the distribution of the coefficient $\hat{a}'_{k,1}$ when $\hat{m}'_k = 1$. From this figure, the distribution of $1 - \hat{a}'_{k,1}$ seems to be exponential. Its sample average is 0.141.

Figure A.4.1 is the scatter diagram of $\hat{a}'_{k,1}$ and $\hat{a}'_{k,2}$ when $\hat{m}'_k = 2$. The distribution of $\hat{a}'_{k,2}$ is shown in Figure A.4.1 and the distribution of $\hat{a}'_{k,1} + \hat{a}'_{k,2}$ is shown in Figure 13. From Figure 13, the distribution of $1 - (\hat{a}'_{k,1} + \hat{a}'_{k,2})$ seems to be exponential. Its expectation is estimated as 0.216 from the sample mean. These exponential distributions for $\hat{m}'_k = 1$ and 2 are not rejected with 10% significance by a test of goodness of fit.

The distribution of $\hat{a}'_{k,2}$ in Figure A.4.1 seems to be bimodal with modes near to 0.15 and -0.30. Its sample mean and sample deviation are 0.010 and 0.096, respectively.

As shown in Figures A.4.1 and A.4.1, almost all of the estimated coefficients satisfy the stable condition of AR processes;

$$\begin{aligned} -1 < a'_{k,1} < 1 \qquad & \text{when } \hat{m}'_k = 1, \\ a'_{k,1} + a'_{k,2} < 1,\ a'_{k,1} - a'_{k,2} < 1,\ -1 < a'_{k,2} < 1 \qquad & \text{when } \hat{m}'_k = 2. \end{aligned}$$

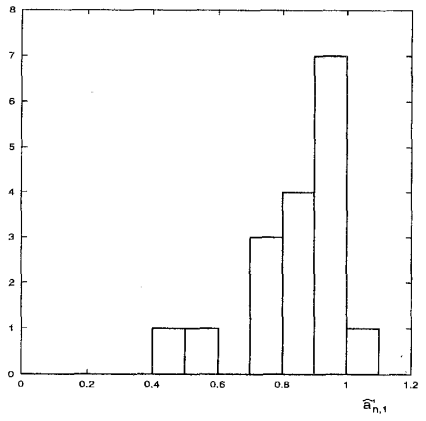


Figure 10: Histogram of $\hat{a}'_{k,1}$ when $\hat{m}'_k = 1$

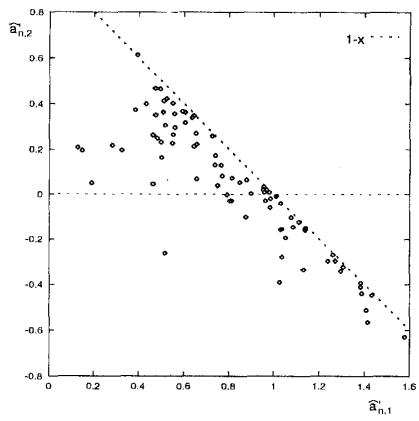


Figure 11: Scatter diagram of $\hat{a}'_{k,1}$ and $\hat{a}'_{k,2}$ when $\hat{m}'_k = 2$

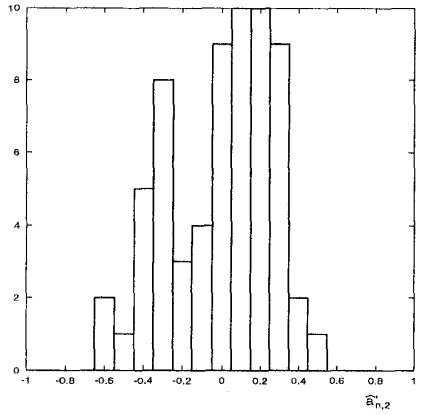


Figure 12: Histogram of $\hat{a}'_{n,2}$ when $\hat{m}'_n = 2$

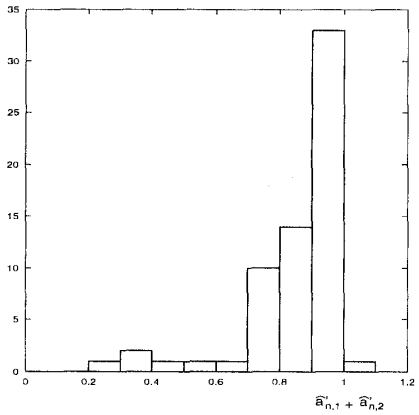


Figure 13: Histogram of $\hat{a}'_{n,1}+\hat{a}'_{n,2}$ when $\hat{m}'_n = 2$

A.4.2. Standard deviation $\hat{\sigma}'_k$ of $\epsilon_k(n)$'s

Figure 14 is the scatter diagram of the average rate A_k and the standard deviation $\hat{\sigma}'_k$ of $\epsilon_k(n)$ for the AR process. There exists a weak positive correlation between A_k and $\hat{\sigma}'_k$.

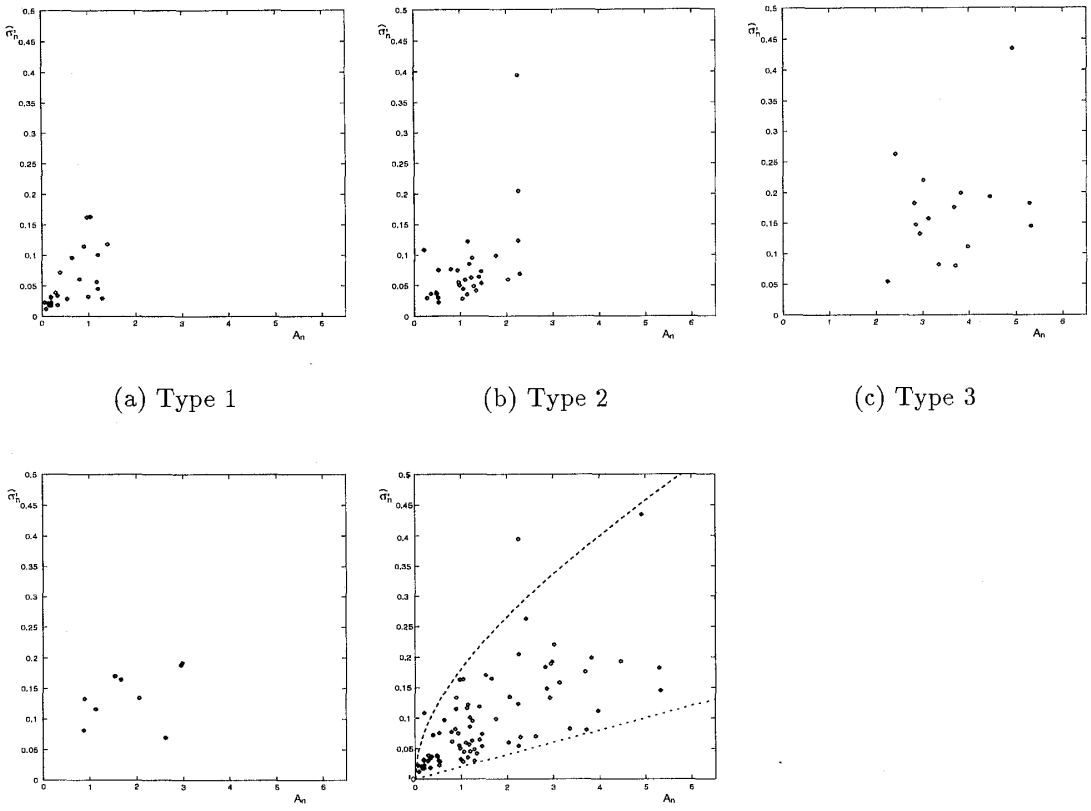
The standard deviation $\hat{\sigma}'_k$ seems to range between $0.02A_k$ and $0.16\sqrt{A_k} + 0.02A_k$ (see Figure 14(e)). The average μ_{v_i} and the standard deviation σ_{v_i} of the statistic

$$v_k = (\hat{\sigma}'_k - 0.02A_k)/\sqrt{A_k}$$

of Type i $(i = 1, \cdots, 4)$ are listed in Table 1. The histogram of v_k is given in Figure A.4.2. A gamma distribution with mean μ_{v_i} and standard deviation σ_{v_i} is fitted to the distribution of v_k of Type i $(i = 1, \cdots, 4)$. The gamma distributions with these estimated parameters are not rejected with 10% significance in a test of goodness of fit.

A.5. Dependence between parameters

In subsection A.4.2, we showed the correlation between the average rate A_n and the standard deviation $\hat{\sigma}'_n$ of the AR process. We can't find any clear correlations among other estimated parameters.



(a) Type 1 (b) Type 2 (c) Type 3

(d) Type 4 (e) Total

Figure 14: Scatter diagram of A_n and $\hat{\sigma}'_n$

Acknowledgements

The authors thank anonymous referees for their comments which improve the description of the paper very much.

References

[1] J. Beran, R. Sherman, M. S. Taqqu and W. Willinger: Long-range dependence in variable-bit-rate video traffic. *IEEE Trans. Commun.*, **43**-2/3/4 (1995) 1566–1579.

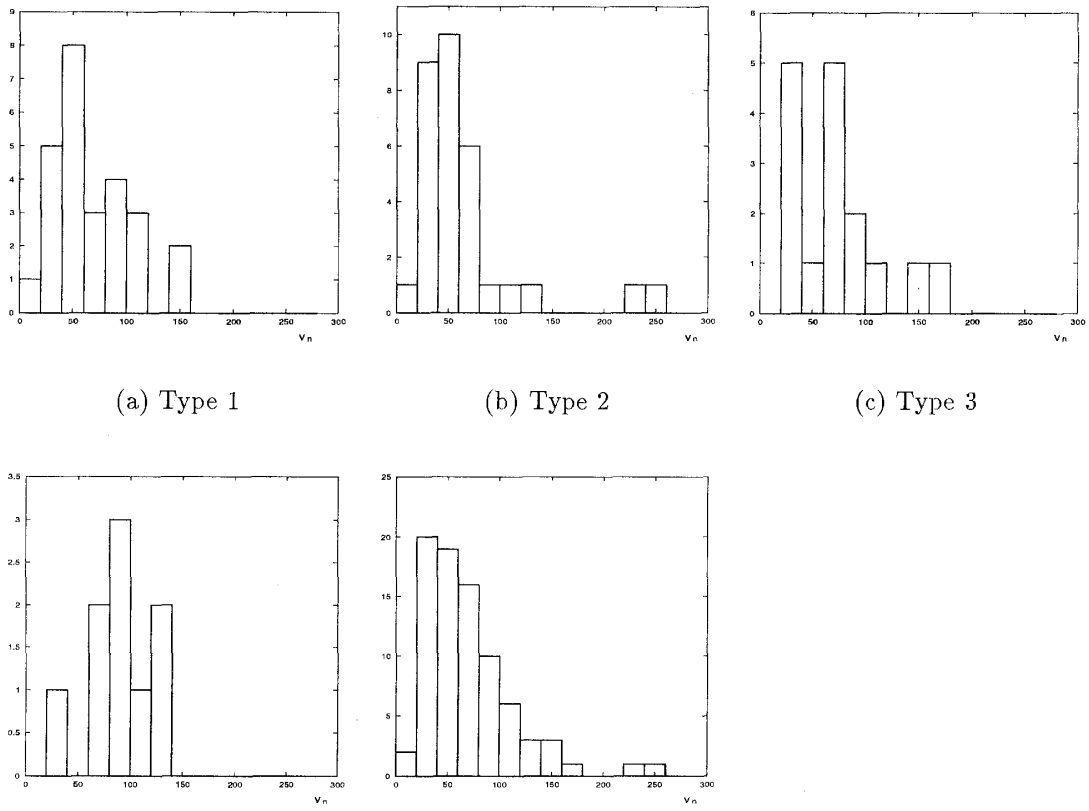
[2] D. R. Cox: Long-Range Dependence. H. A. David and H. T. David (eds.): *Statistics, an appraisal* (The Iowa State University Press, Ames, Iowa, 1984), 55–74.

[3] A. Elwalid, D. Heyman, T. V. Lakshman, D. Mitra and A. Weiss: Fundamental results on the performance of ATM multiplexers with applications to video teleconferencing. *SIGMETRICS'95* (Ottawa, Ontario, Canada, 1995), 86–96.

[4] M. R. Frater, J. F. Arnold and P. Tan: A new statistical model for traffic generated by VBR coders for television on the broadband ISDN. *IEEE Trans. Circuits and Systems for Video Tech.*, **4**-6 (1994) 521–526.

[5] C. W. Granger and R. Joyeux: An introduction to long-memory time series models and fractional differencing. *J. Time Series Anal.*, **1** (1980) 15–29.

[6] M. Grasse, M. R. Frater and F. Arnold: Origins of long-range dependence in variable bit rate video traffic. *ITC 15* (1997), 1379–1388.



(a) Type 1 (b) Type 2 (c) Type 3

(d) Type 4 (e) Total

Figure 15: Histogram of v_n

[7] R. Grünenfelder, J. P. Cosmas, S. Manthorpe and A. Odinma-Pkafor: Characterization of video codecs as autoregressive moving average processes and related queuing system performance. *IEEE J. Select. Areas Commun.*, **9**-3 (1991) 284–293.

[8] D. P. Heyman, A. Tabatabai and T. V. Lakshman: Statistical analysis and simulation study of video teleconference traffic in ATM networks. *IEEE Trans. Circuits and Systems for Video Tech.*, **2**-1 (1992) 49–59.

[9] D. P. Heyman and T. V. Lakshman: Source models for VBR broadcast-video traffic. *IEEE/ACM Trans. on Networking*, **4**-1 (1996) 40–48.

[10] D. P. Heyman and T. V. Lakshman: What are the implication of long-range dependence for VBR-video traffic engineering?. *IEEE/ACM Trans. on Networking*, **4**-3 (1996) 301–317.

[11] C. Huang, M. Devetsikiotis, I. Lambadaris and A. R. Kaye: Modeling and simulation of self-similar variable bit rate compressed video: a unified approach. *SIGCOMM '95* (Cambridge, 1995), 114–125.

[12] K. Kobayashi: *Performance Modeling and Analysis of ATM Multiplexers* (Ph.D.thesis, Tokyo Institute of Technology, 1996).

[13] K. Kobayashi and Y. Takahashi: The tail probability of a Gaussian fluid queue under finite measurement of input processes. *PMCCN'97* (1997), 57–71.

[14] K. R. Krishnan, G. Meempat: Long-range dependence in VBR video streams and ATM

traffic engineering. *Performance Evaluation*, **30** (1997) 45–56.

[15] T. Kurasugi, K. Kobayashi and Y. Takahashi: Data Analysis and Modeling of ATM Coded Video Traffic with Scene Changes. *Proceedings of the Symposium on Performance Models for Information Communication Networks* (1996), 371–382.

[16] T. Kurasugi: *Data Analysis and Modeling of ATM Coded Video Traffic with Scene Changes.* (Master thesis, Tokyo Institute of Technology, 1996).

[17] B. Maglaris, D. Anastassiou, P. Sen, G. Karlsson and J. D. Robbins: Peformance Models of Statistical Multiplexing in Packet Video Communications. *IEEE Trans. Commun.*, **36**-7 (1988) 834–843.

[18] B. B. Mandelbrot and J. W. Van Ness: Fractional Brownian motions, fractional noises and applications. *SIAM Review*, **10**-4 (1968) 422–437.

[19] M. Nomura, T. Fuji and N. Ohta: Basic characteristics of variable rate video coding in ATM environment. *IEEE J. Select. Areas Commun.*, **7**-5 (1989) 752–760.

[20] I. Norros: A storage model with self-similar input. *Queueing Systems*, **16** (1994) 387–396.

[21] G. Ramamurthy and B. Senupta: Modeling and analysis of a variable bit rate video multiplexer. *IEEE INFOCOM'92* (1992), 817–827.

[22] J. Rissanen: Modeling by shortest data description. *Automatica*, **14** (1978) 465–471.

[23] H. Takayasu: *Fractals in the Physical Sciences.* (John Wiley & Sons, 1989).

[24] F. Yegenoglu, B. Jabbari and Y. Zhang: Motion-classified autoregressive modeling of variable bit rate video. *IEEE Trans. on Circuits and Systems for Video Techonology*, **3**-1 (1993) 42–53.

[25] W. Vandaele: *Applied Time Series and Box-Jenkins Models.* (Academic Press, 1983).

Toshiyasu Kurasugi:
C&C Media Research Laboratories
NEC Corporation
4-1-1, Miyazaki, Miyamae-ku, Kawasaki
Kanagawa, 216-8555, Japan
E-mail: tk@ccm.cl.nec.co.jp