A Method for the Generation of Correlated Random Processes

Maurizio Andronico, Salvatore Casale and Aurelio La Corte

Istituto di Informatica e Telecomunicazioni, University of Catania, Italy

In this paper the authors propose a method which will allow the generation of random discrete variables with a given probability distribution and autocorrelation sequence. The method is applicable in cases where, once experimental measurements have established the statistical characteristics of a stationary random process, an algorithm is to be implemented to generate random discrete variables with the same statistical properties in terms of amplitude distribution and correlation among the values. The method proposed allows a correlated random process to be generated by combining two ergodic independent statistical uncorrelated random processes. A case study is given to apply the method proposed to modelling of variable bit rate video sources by simulation.

Keywords: random processes, simulation, modelling

1. Introduction

The construction of general random processes is a problem to be dealt with in several fields such as detection, classification and control systems [Johnson, 1994] [Middleton, 1966] [Van Trees, 1968] and traffic source modelling [Frost and Melamed, 1994] [Habib and Saadawi, 1992]. While it is not difficult to generate uncorrelated discrete random variables with a given distribution function, if this is a continuous, reversible function [Van Trees, 1968][Frost and Melamed, 1994], or particular correlated random processes [Johnson, 1994], it is sometimes quite hard to generate variables correlated according to a given law. In this paper the authors propose a method which allows the generation of discrete random variables, fixing their probability distribution and autocorrelation sequence independently. The method can be applied in cases where, once the statistical characteristics (in terms of probability distribution and autocorrelation sequence) of a discrete random process are known from experimental measurements, one wants to implement a simple algorithm allowing the generation of discrete random variables with the same statistical properties. A case in point is the simulation modelling of a variable bit rate (VBR) traffic source.

Section 2 below gives a description of the method proposed by the authors, demonstrating statistical properties of the random process obtained by suitably combining two uncorrelated, statistically independent random processes. Section 3 makes some brief remarks on the results obtained in simulations performed to validate the proposed method. In Section 4 the authors describe how to apply the proposed method to model a VBR traffic source by simulation. Section 5 gives the authors' conclusions and some final remarks.

2. Generation of correlated random processes

This section shows how to construct sets of random numbers with particular amplitude distributions and correlation among values.

The authors' basic idea is to generate a set of correlated discrete random variables starting from two sets of statistically independent discrete random variables [Andronico et al., 1994]. By fixing the probability function of each of

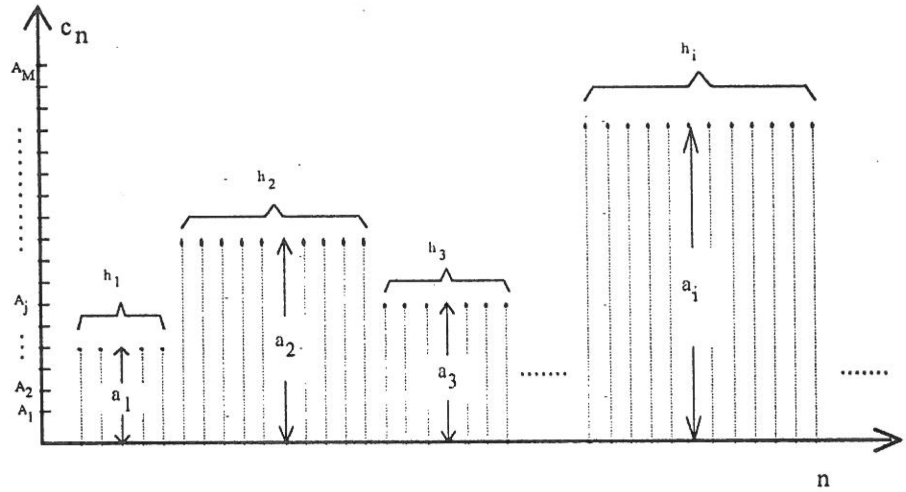

Fig. 1. Generation of the compound process $\{C\}$ from processes $\{A\}$ and $\{H\}$.

these two sets, one obtains a set of discrete random variables with a given probability function and autocorrelation sequence. The problem of generating a correlated random process thus boils down to the simpler one of generating two ergodic independent statistical uncorrelated random processes.

In the following the generation of a correlated random process is illustrated and its statistical properties in relation to the two component processes are outlined.

Let us consider two discrete random processes, that is, two ordered sets of quantized time-discrete random variables $\{a_n\}$ and $\{h_n\}$. The values assumed by the variables a_n belong to a finite, numerable set $\{A\} = \{A_1, A_2, \ldots, A_M\}$, which is called *the set of amplitude values*. Likewise, the values assumed by the variables h_n belong to a finite, numerable set $\{H\} = \{H_1, H_2, \ldots, H_N\}$, called *the set of holding times*. The two random processes $\{a_n\}$ and $\{h_n\}$ are both wide-sense stationary, ergodic and statistically independent. It is also assumed that both the variables a_n and h_n are identically distributed, and that the variables a_n are uncorrelated. Henceforward we will use the terms process $\{A\}$ and process $\{H\}$ to indicate the random process of quantized variables $\{a_n\}$ and $\{h_n\}$ respectively.

The two sets $\{A\}$ and $\{H\}$ are characterized by the corresponding probability functions of the discrete random variables, respectively indicated as $p_{a_n}(x)$ and $p_{h_n}(x)$, where $p_{a_n}(x)$ is the probability that a_n is equal to "x" and $p_{h_n}(x)$ is the probability that h_n is equal to "x".

Starting from these two sets, which are also called "the component processes", the "compound" process $\{C\}$ is built, considering repetition of the generic variables a_n a number of times equal to the variable h_n. In other words, each element a_n which assumes a value belonging to the set $\{A\}$, representing the amplitude of the compound process, is associated with a duration given by the value assumed by the variable h_n belonging to the set $\{H\}$ (see Fig. 1). That is, the compound process is formed by the ordered sequence:

$$\{c_n\} = \Big\{a_1, \underbrace{\ldots}_{h_1 \text{ times}}, a_1, a_2, \underbrace{\ldots}_{h_2 \text{ times}}, a_2, \ldots, a_n, \underbrace{\ldots}_{h_n \text{ times}}, a_n, \ldots\Big\}$$

The values assumed by the random variable c_n are the same as those assumed by the random variable a_n, so $\{C\} = \{A_1, A_2, \ldots, A_M\}$. As for the processes $\{A\}$ and $\{H\}$, henceforward the term process $\{C\}$ will be used to indicate the random process of quantized time-discrete random variables $\{c_n\}$.

The statistical parameters of process $\{C\}$ are obtained from those of processes $\{A\}$ and $\{H\}$. The following properties in fact hold:

Property 1

The probability function of process $\{C\}$ is equal to that of process $\{A\}$, i.e.

$$p_{c_n}(x) = p_{a_n}(x) \tag{2.1}$$

where $p_{c_n}(x)$ is the probability that c_n is equal to "x".

Proof. Consider "k" extractions from process $\{A\}$ and process $\{H\}$ and the corresponding compound process $\{C\}$. We also define the probabilities calculated on the finite-length sequences obtained from the above extractions as *experimental probabilities*, indicating

$p^{(e)}_{a_n}(x)$ = probability that $a_n = x$,

$p^{(e)}_{h_n}(x)$ = probability that $h_n = x$,

$p^{(e)}_{c_n}(x)$ = probability that $c_n = x$.

As the set of values process $\{C\}$ can assume coincides with the set of values process $\{A\}$ can assume, for the *experimental probabilities* of process $\{C\}$ we have [Oppenheim and Schafer, 1974]:

$$p^{(e)}_{c_n}(A_x) = \frac{h_{A_x}}{\sum\limits_{i=1}^{k} h_i} = \frac{h_{A_x}}{\sum\limits_{i=1}^{M} h_{A_i}}$$

where h_{A_x} is the number of samples of process $\{C\}$ with an amplitude of A_x during an observation of the process corresponding to "k" events of processes $\{A\}$ and $\{H\}$.

Assuming that process $\{C\}$ is wide-sense stationary and ergodic, like the two component processes $\{A\}$ and $\{H\}$, we have:

$$p_{c_n}(A_x) = \lim_{k\to\infty} p^{(e)}_{c_n}(A_x) = \lim_{k\to\infty} \frac{h_{A_x}}{\sum\limits_{i=1}^{k} h_i}$$

$$= \lim_{k\to\infty} \frac{\left(\sum\limits_{j=1}^{N} H_j \cdot p^{(e)}_{h_n}(H_j)\right) \cdot k \cdot p^{(e)}_{a_n}(A_x)}{\sum\limits_{i=1}^{M}\left(\sum\limits_{j=1}^{N} H_j \cdot p^{(e)}_{h_n}(H_j)\right) \cdot k \cdot p^{(e)}_{a_n}(A_i)}$$

$$= \lim_{k\to\infty} \frac{\left(\sum\limits_{j=1}^{N} H_j \cdot p^{(e)}_{h_n}(H_j)\right) \cdot k \cdot p^{(e)}_{a_n}(A_x)}{\left(\sum\limits_{j=1}^{N} H_j \cdot p^{(e)}_{h_n}(H_j)\right) \cdot k \cdot \sum\limits_{i=1}^{M} p^{(e)}_{a_n}(A_i)}$$

$$= \lim_{k\to\infty} p^{(e)}_{a_n}(A_x) = p_{a_n}(A_x).$$

(2.1) is thus demonstrated. Hence, the mean, mean quadratic value and variance are equal, i.e.:

$$m_c = \sum_C c \cdot p_{c_n}(c) = \sum_A a \cdot p_{a_n}(a) = m_a \tag{2.2}$$

$$E[c_n^2] = \sum_C c^2 \cdot p_{c_n}(c) = \sum_A a^2 \cdot p_{a_n}(a) = E[a_n^2] \tag{2.3}$$

$$\sigma_c^2 = \sigma_a^2 \tag{2.4}$$

Property 2

The autocorrelation sequence of the compound process $\{C\}$ is linked to the variance of process $\{A\}$ and the probability distribution of process $\{H\}$ by the following relation:

$$\Phi_{cc}(m) = \begin{cases} \Phi_{aa}(0) = \sigma_a^2 + m_a^2, & m=0 \\ \Phi_{cc}(0) - \dfrac{\sigma_a^2}{m_h} \cdot \sum\limits_{i=0}^{m-1} [1 - P_{h_n}(i)], & m>0 \end{cases} \tag{2.5}$$

where m_h is the mean of process $\{H\}$ and $P_{h_n}(i)$ = probability $[h_n \le i] = \sum\limits_{j=1}^{i} p_{h_n}(j)$.

Proof. The autocorrelation sequence for process $\{C\}$ is by definition given by [Oppenheim and Schafer, 1974]:

$$\Phi_{cc}(n, n+m) = \sum_i \sum_j x_i \cdot x_j \cdot p_{c_n, c_{n+m}}(x_i, x_j) \tag{2.6}$$

The stationarity of the moments of the first order of process $\{C\}$, used to obtain relations (2.1)–(2.4), derives from the stationarity of process

$\{A\}$. Admitting that when $n > h_{\min}$ (the first $h_{\min}$ samples of the process $\{C\}$ are certainly equal to c_1) is

$$p_{c_{n+k},c_{m+k}}(x_i, x_j) = p_{c_n,c_m}(x_i, x_j) \tag{2.7}$$

we have:

$$\Phi_{cc}(n, n+m) = \Phi_{cc}(m) \tag{2.8}$$

Consider the autocorrelation sequence $\Phi_{cc}(m)$. Obviously

$$\begin{aligned}\Phi_{cc}(0) &= E[c_n^2] = E[a_n^2] \\ &= \sigma_a^2 + m_a^2 = \lim_{k\to\infty} \Phi_{cc}^{(e)}(0) \\ &= \lim_{k\to\infty} \frac{\sum_{i=1}^{k} h_i \cdot a_i^2}{\sum_{i=1}^{k} h_i}\end{aligned} \tag{2.9}$$

where $\Phi_{cc}^{(e)}(.)$ indicates the experimental autocorrelation sequence of process $\{C\}$, i.e. the autocorrelation sequence values obtained by observing a finite number of process $\{C\}$ events. The first part of (2.5) is thus demonstrated. Likewise, as regards $\Phi_{cc}(1)$, we have:

$$\begin{aligned}\Phi_{cc}(1) &= \lim_{k\to\infty} \Phi_{cc}^{(e)}(1) \\ &= \lim_{k\to\infty} \frac{\sum_{i=1}^{k}(h_i - 1) \cdot a_i^2 + \sum_{i=1}^{k-1} a_i \cdot a_{i+1}}{\sum_{i=1}^{k} h_i} \\ &= \lim_{k\to\infty} \frac{\sum_{i=1}^{k} h_i \cdot a_i^2}{\sum_{i=1}^{k} h_i} - \lim_{k\to\infty} \frac{\sum_{i=1}^{k} a_i^2}{\sum_{i=1}^{k} h_i} \\ &\quad + \lim_{k\to\infty} \frac{\sum_{i=1}^{k-1} a_i \cdot a_{i+1}}{\sum_{i=1}^{k} h_i}\end{aligned} \tag{2.10}$$

The second and third term in the second member of expression (2.10) represent the contribution of the "border effect", that is, the product of two samples belonging to two adjacent holding times. Of course the higher the value of m_h, the less accentuated this contribution is.

(2.10) can be written as:

$$\begin{aligned}\Phi_{cc}(1) = \lim_{k\to\infty} \frac{\sum_{i=1}^{k} h_i \cdot a_i^2}{\sum_{i=1}^{k} h_i} - \lim_{k\to\infty} \frac{\sum_{i=1}^{k} a_i^2}{k} \cdot \frac{k}{\sum_{i=1}^{k} h_i} \\ + \lim_{k\to\infty} \frac{\sum_{i=1}^{k} a_i \cdot a_{i+1}}{k-1} \cdot \frac{k-1}{\sum_{i=1}^{k} h_i}\end{aligned} \tag{2.11}$$

and, as $\lim_{k\to\infty} \frac{\sum_{i=1}^{k} h_i}{k} = m_h$, we have:

$$\begin{aligned}\Phi_{cc}(1) &= \Phi_{cc}(0) - \frac{\Phi_{aa}(0)}{m_h} + \frac{\Phi_{aa}(1)}{m_h} \\ &= \Phi_{cc}(0) - \frac{[\Phi_{aa}(0) - \Phi_{aa}(1)]}{m_h}\end{aligned} \tag{2.12}$$

As process $\{A\}$ is by assumption uncorrelated, we have $\Phi_{aa}(0) - \Phi_{aa}(1) = \sigma_a^2$ and so

$$\Phi_{cc}(1) = \Phi_{cc}(0) - \frac{\sigma_a^2}{m_h} \tag{2.13}$$

As a particular case of relation (2.13), it can be observed that if $h_i = 1$, $\forall i$ integer and greater than zero, which means that processes $\{C\}$ and $\{A\}$ coincide, $\Phi_{cc}(1) = \Phi_{aa}(1) = \Phi_{aa}(0) - \sigma_a^2 = m_a^2$, as it should be since process $\{A\}$ is uncorrelated.

Just as (2.13) was obtained, the expression of $\Phi_{cc}(2)$ can be obtained in the case in which $h_i \geq 2$, $\forall i$ integer and greater than zero. We have, in fact:

$$\begin{aligned}\Phi_{cc}(2) &= \lim_{k\to\infty} \Phi_{cc}^{(e)}(2) \\ &= \lim_{k\to\infty} \frac{\sum_{i=1}^{k}(h_i - 2) \cdot a_i^2 + 2 \cdot \sum_{i=1}^{k-1} a_{i+1}}{\sum_{i=1}^{k} h_i}\end{aligned}$$

$$= \lim_{k\to\infty} \frac{\sum_{i=1}^{k} h_i \cdot a_i^2}{\sum_{i=1}^{k} h_i} - 2 \cdot \lim_{k\to\infty} \frac{\sum_{i=1}^{k} a_i^2}{\sum_{i=1}^{k} h_i}$$

$$+ 2 \cdot \lim_{k\to\infty} \frac{\sum_{i=1}^{k} a_i \cdot a_{i+1}}{\sum_{i=1}^{k} h_i}$$

$$= \lim_{k\to\infty} \frac{\sum_{i=1}^{k} h_i \cdot a_i^2}{\sum_{i=1}^{k} h_i} - 2 \cdot \lim_{k\to\infty} \frac{\sum_{i=1}^{k} a_i^2}{k} \cdot \frac{k}{\sum_{i=1}^{k} h_i}$$

$$+ 2 \cdot \lim_{k\to\infty} \frac{\sum_{i=1}^{k-1} a_i \cdot a_{i+1}}{k-1} \cdot \frac{k-1}{\sum_{i=1}^{k} h_i}$$

$$= \Phi_{cc}(0) - 2 \cdot \frac{\sigma_a^2}{m_h}$$

This reasoning can be repeated to calculate $\Phi_{cc}(m)$ for all the values of m such that $h_i > m - 1$, $\forall i$ integer and greater than zero, thus obtaining:

$$\Phi_{cc}(m) = \Phi_{cc}(0) - m \cdot \frac{\sigma_a^2}{m_h}, \quad m \leq h_{\min} \quad (2.14)$$

or:

$$\Phi_{cc}(m) = \Phi_{cc}(m-1) - \frac{\sigma_a^2}{m_h}, \quad m \leq h_{\min} \quad (2.14')$$

where $h_{\min}$ indicates the minimum value assumed by the stochastic variable h_n (this value exists according to the definition of process $\{H\}$).

Let us now consider the case in which at least one value of "i" is such that $h_i \leq m - 1$. If $h_{\max}$ indicates the maximum value assumed by the stochastic variable h_n (a value which exists according to the definition of process $\{H\}$), if $m > h_{\max}$ considering how process $\{C\}$ is constructed, and process $\{A\}$ being made up of uncorrelated variables, the variables which are considered in calculation of the autocorrelation sequence are uncorrelated; so:

$$p_{c_n, c_{n+m}}(x_i, x_j) = p_{c_n}(x_i) \cdot p_{c_{n+m}}(x_j) \quad (2.15)$$

and thus from (2.6)

$$\Phi_{cc}(m) = m_c^2 = m_a^2 \quad (2.16)$$

As (2.16) is verified for $\forall m > h_{\max}$, then

$$\Phi_{cc}(m+1) = \Phi_{cc}(m), \quad m \geq h_{\max} \quad (2.17)$$

On the basis of the considerations made above it can be stated that if $h_{\min} < m < h_{\max}$ (the intermediate case between those pointed out in (2.14') and (2.17)), we have:

$$\Phi_{cc}(m) = \Phi_{cc}(m-1) - \frac{\sigma_a^2}{m_h} \cdot p(h_n > m-1)$$
$$h_{\min} \leq m \leq h_{\max} \quad (2.18)$$

Corresponding to $h_i : h \leq m - 1$, in fact, on account of (2.17) the contribution made to the difference $\Phi_{cc}(m) - \Phi_{cc}(m-1)$ is null; on the other hand, corresponding to $h_i : h_i > m - 1$ on account of (2.13) the contribution made to the difference $\Phi_{cc}(m-1) - \Phi_{cc}(m)$ is equal to σ_a^2/m_h.

In conclusion, from (2.18) we can obtain the expression of the autocorrelation sequence of processs $\{C\}$, which is:

$$\Phi_{cc}(m) = \Phi_{cc}(0) - \frac{\sigma_a^2}{m_h} \sum_{i=0}^{m-1} p(h_n > i) \quad m \geq 1 \quad (2.19)$$

As $p(h_i > i) = 1 - P_{h_n}(i)$ it is possible to express (2.19) in the equivalent form:

$$\Phi_{cc}(m) = \Phi_{cc}(0) - \frac{\sigma_a^2}{m_h} \sum_{i=0}^{m-1} [1 - P_{h_n}(i)] \quad m \geq 1 \quad (2.19')$$

Let us finally verify the congruence of (2.19) with the properties which, by definition, must characterize the autocorrelation sequence $\Phi_{cc}(m)$:

a) certainly $|\Phi_{cc}(m)| \leq \Phi_{cc}(0)\ \forall m$ integer and greater than zero. The second term of the second member of (2.19') is, in fact, certainly a non-negative quantity.

b) when $m = 0$ we have $\Phi_{cc}(0) = E[c_n^2] = E[a_n^2]$;

c) $\lim_{m\to\infty} \Phi_{cc}(m) = (E[c_n])^2 = m_c^2 = m_a^2$.

Let us demonstrate that relation c) is also verified. Considering that $\Phi_{cc}(0) = E[c_n^2] =$

$E[a_n^2] = \sigma_a^2 + m_a^2$, verifying relation c) is equivalent to verifying that $\lim_{m\to\infty} \frac{\sigma_a^2}{m_h} \sum_{i=0}^{m-1} p(h_n > i) = \sigma_a^2$, that is $\sum_{i=0}^{\infty} p(h_n > i) = m_h$.

Let us note that, as h_n can only assume integer values greater than zero, we have:

$$\sum_{i=0}^{\infty} p(h_n > i) = \sum_{i=0}^{\infty} p(h_n \geq i)$$
$$= \sum_{i=1}^{\infty}\sum_{j=i}^{\infty} p(h_n = j) = \sum_{j=1}^{\infty}\sum_{i=1}^{j} p(h_n = j)$$
$$= \sum_{j=1}^{\infty} j \cdot p(h_n = j) = m_h$$

Thus, as relation c) is also verified, (2.5) is demonstrated.

The method presented for the generation of discrete random variables has comparatively general characteristics, but at the same time it has two inherent limits. From (2.5), in fact, we obtain the following two relations:

i) $\Phi_{cc}(m) - \Phi_{cc}(m-1) = -\frac{\sigma_a^2}{m_h} \cdot [1 - P_{h_n}(m-1)] \leq 0, m \geq 1$

ii) $[\Phi_{cc}(m) - \Phi_{cc}(m-1)] - [\Phi_{cc}(m-1) - \Phi_{cc}(m-2)] = -\frac{\sigma_a^2}{m_h}[P_{h_n}(m-2) - P_{h_n}(m-1)] \geq 0, m \geq 2$.

The autocorrelation sequence $\Phi_{cc}(m)$ is always not-increasing, that is, it is not possible to represent random processes with fluctuations in the autocorrelation sequence, for instance the non-stationary process of a video sequence. The assumption of the stationary nature of processes $\{A\}$ and $\{H\}$ and the compound process $\{C\}$ underlies the whole proposal being made. Moreover, the second derivate of the autocorrelation sequence is always non-negative, that is, the autocorrelation sequence graph is represented by a broken line which always has a downward convexity.

3. Tests on the proposed method

On the basis of the method described in the previous section, we have written a simulation model which allows us to compare the analytical results with those obtained by means of simulation.

In writing this simulation model the problem was how to generate the variables relating to the two processes $\{A\}$ and $\{H\}$, and, in particular, what probability functions to consider in order to compare the theoretical results expressed by equations (2.1)–(2.5) in the previous section to those obtained through simulation. The approach followed was to consider probability functions as samples from a probability density function, multiplying these samples by a scale factor in such a way that the sum of the probabilities is equal to 1. The use of sampled values involves an approximation of the probability functions of the component process, to which a further approximation is added, because the tails of the probability density functions are truncated. This approximation is necessary because the computer operates with finite blocks of data so as to limit the set of values of the variables of the two component sets.

In the simulation study we generated samples from the following probability density functions: Uniform, Gaussian and Exponential. We varied the sampling interval and truncation of the probability density functions. The simulation results were obtained in steady-state conditions with a 95% confidence interval of the true value. All the results obtained from simulations and the corresponding comparisons with the theoretical values obtained from (2.1)–(2.5) confirmed the properties demonstrated in the previous section.

4. An example of application of the proposed method

The problem of modelling a VBR traffic source by simulation consists of generating correlated discrete random variables which represent the number of packets offered to the network, or the burst interarrival times, or again the burst duration times, according to the statistical characteristics of the source to be modelled [Catania et al., 1994] . In this section we will illustrate how the proposed method can be applied to model a source.

Let us suppose that we have experimental measurements of the statistics of a source, more specifically the histogram of the experimental probabilities $p_{c_n}^{(e)}(x)$, for all the values "x" the output of the source can assume, and the experimental autocorrelation sequence $\Phi_{cc}^{(e)}(m)$ for $m = 0, \ldots, m_{\text{MAX}}$. Let us see how to determine the statistics of the processes $\{A\}$ and $\{H\}$ in such a way that for the compound process $p_{c_n}(x) = p_{c_n}^{(e)}(x)$ and $\Phi_{cc}(m) = \Phi_{cc}^{(e)}(m)$.

From Property 1 in Section 2 the probability function of process $\{A\}$ has to be equal to the experimental one for the source. That is:

$$p_{a_n}(x) = p_{c_n}^{(e)}(x)$$

If $p_{c_n}^{(e)}(x)$ can be approximated to a sequence of samples of a probability density function whose corresponding probability distribution function is a continuous reversible function, it is possible to generate random uncorrelated numbers which represent the values assumed by the variable a_n through the generation of uniformly distributed numbers [Therrien, 1992]. If that is not possible, it is necessary to use an ad hoc algorithm for the generation of these variables [Johnson, 1994][Therrien, 1992].

Now let us determine the statistics of process $\{H\}$. From (2.5) when $m = 1$ and $h_{\min} \geq 1$, we obtain:

$$m_h = \frac{\sigma_a^2}{\Phi_{cc}(0) - \Phi_{cc}(1)}$$

Substituting in (2.5) we have:

$$\Phi_{cc}(m) = \Phi_{cc}(0) - [\Phi_{cc}(0) - \Phi_{cc}(1)] \cdot \sum_{i=0}^{m-1} [1 - P_{h_n}(i)] \qquad 1 \leq m \leq m_{\text{MAX}} \quad (4.1)$$

As the variables are discrete we have:

$$p_{h_n}(i) = P_{h_n}(i) - P_{h_n}(i-1) = \frac{\Phi_{cc}(i+1) - 2 \cdot \Phi_{cc}(i) + \Phi_{cc}(i-1)}{\Phi_{cc}(0) - \Phi_{cc}(1)} \qquad 1 \leq i \leq m_{\text{MAX}} - 1 \quad (4.2)$$

By imposing

$$\Phi_{cc}(i) = \Phi_{cc}^{(e)}(i)$$

we obtain the probabilities of process $\{H\}$

$$p_{h_n}(i) = \frac{\Phi_{cc}^{(e)}(i+1) - 2 \cdot \Phi_{cc}^{(e)}(i) + \Phi_{cc}^{(e)}(i-1)}{\Phi_{cc}^{(e)}(0) - \Phi_{cc}^{(e)}(1)} \qquad 1 \leq i \leq m_{\text{MAX}} - 1 \quad (4.3)$$

It can be verified that the condition of congruence of the probabilities

$$\sum_{i=h_{\min}}^{h_{\max}} p_{h_n}(i) = 1$$

where $h_{\min} \geq 1$ and $h_{\max} \leq m_{\text{MAX}} - 1$ is demonstrated similarly to the relation (c) in Section 2.

As $p_{h_n}(i) \geq 0$ has to be true, from (4.3) we obtain

$$\frac{\Phi_{cc}^{(e)}(i+1) - 2 \cdot \Phi_{cc}^{(e)}(i) + \Phi_{cc}^{(e)}(i-1)}{\Phi_{cc}^{(e)}(0) - \Phi_{cc}^{(e)}(1)} \geq 0 \implies \Phi_{cc}^{(e)}(i+1) - 2 \cdot \Phi_{cc}^{(e)}(i) + \Phi_{cc}^{(e)}(i-1) \geq 0 \qquad \forall i \geq 1 \quad (4.4)$$

This condition is satisfied if the autocorrelation sequence graph is represented by a broken line which always has a downward convexity, that is, if the second derivate of the autocorrelation sequence is non-negative.

It should be pointed out that generation of the variables h_n, once the corresponding probability function is known, is easy to achieve as it is not necessary for them to be uncorrelated. As the set $\{H\}$ is made up of a finite number of elements, a set can be built containing the possible values h_n can assume, each of them being repeated a number of times that is proportional to the corresponding probability $p_{h_n}(i)$; generation of the variables of the process $\{H\}$ is achieved by taking values from this set at random.

In order to facilitate understanding of application of the proposed method, we consider the following example: the source to be modelled is a video source which generates 30 frames per second, with a conditional replenishment

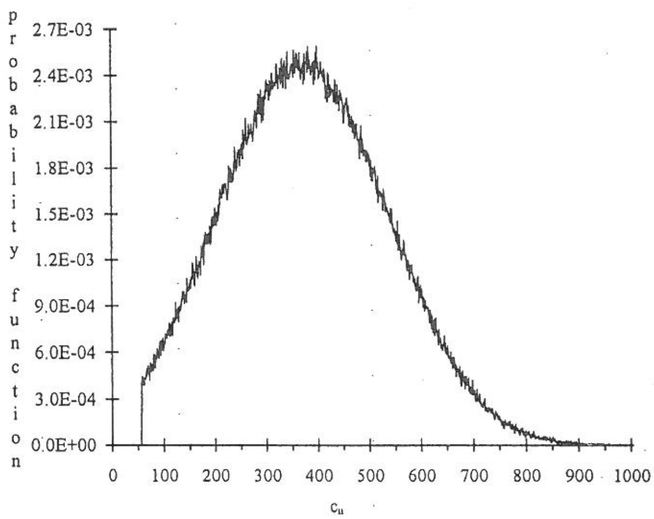

Fig. 2. Probability function of process $\{C\}$ obtained by means of simulation. The probability function of process $\{A\}$ derives from the sampling of a Gaussian probability density function with mean $\mu = 370$ and standard deviation $\sigma = 164$ in the interval $[57, 1002]$ and the probability function of process $\{H\}$ derives from the sampling of an exponential probability density function with mean $\mu = 7.2$ in the interval $[1, 49]$. The quantization interval of the variables is equal to 1.

interframe coding scheme. From experimental measurements [Maglaris et al., 1988 — La Corte et al., 1991 — Nomura et al., 1989 — Gruenfelder et al., 1991 — Karlsson and Vetterli, 1988], calling the bit rate of the source during the n-th frame c(n), it was observed that, approximately:

- the steady-state distribution of c(n) is Gaussian with mean $m_c^{(e)} = 130000$ bit/frame and standard deviation $\sigma_c^{(e)} = 57500$ bit/frame;
- the maximum and minimum values of $c(n)$ are respectively $c_{\max} = 352500$ bit/frame and $c_{\min} = 20000$ bit/frame;
- the autocovariance of $c(n)$ is exponential according to the law:

$$C_{cc}^{(e)}(i) = \sigma_c^{(e)^2} \cdot e^{-0.13 \cdot i} \quad i = 0, 1, 2, \ldots, 50$$

Assuming that the flow of output data from the video coder is organized in packets of 44 octets [La Corte et al., 1991], we have:

$$m_c^{(e)} = 370 \text{ packets/frame}$$

$$\sigma_c^{(e)} = 164 \text{ packets/frame}$$

$$c_{\max} = 1002 \text{ packets/frame}$$

$$c_{\min} = 57 \text{ packets/frame}$$

The set $\{A\}$ is thus made up of the set of integers between c_{min} and c_{max}, whose probabilities are the sampled values of a Gaussian probability density function with a mean equal to 370 and a standard deviation equal to 164.

The probabilities of the variables of process $\{H\}$ are obtained from (4.3) and are:

$$p_{h_n}(i) = \frac{C_{cc}^{(e)}(i+1) - 2 \cdot C_{cc}^{(e)}(i) + C_{cc}^{(e)}(i-1)}{C_{cc}^{(e)}(0) - C_{cc}^{(e)}(1)}$$
$$= (e^{0.13} - 1) \cdot e^{-0.13 \cdot i} = 0.1388 \cdot e^{-0.13 \cdot i}$$
$$i = 1, 2, \ldots, 49$$
$$p_{h_n}(i) = 0 \quad i > 49$$

Process $\{H\}$ has the probability function which can be approximated with the samples of an exponential distribution function with a mean of 1/0.13=7.2, which assumes values in the interval of integers between 1 and 49.

Some simulations were performed during which the variables of processes $\{A\}$ and $\{H\}$ were generated in agreement with the statistical properties calculated above. Fig. 2 shows the trend of the probability function of the variables for the compound process $\{C\}$, while Fig. 3 gives the normalized autocorrelation sequence of the compound process $\{C\}$ and the experimental autocorrelation sequence of the source being

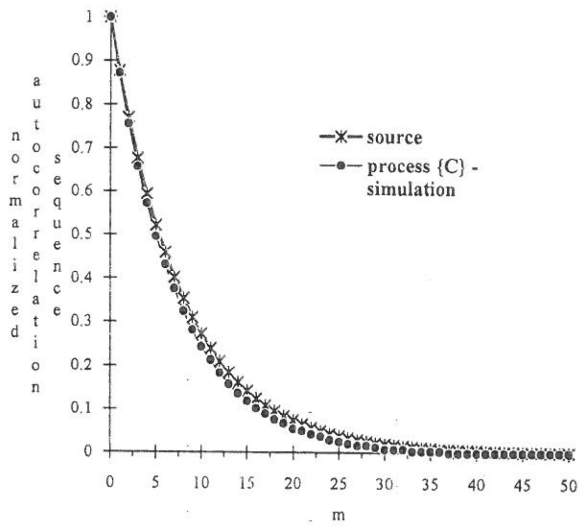

Fig. 3. Normalized autocorrelation sequence of the source being modelled and normalized autocorrelation sequence of the compound process $\{C\}$ obtained by means of simulation. Process $\{A\}$ has a probability function deriving from the sampling of a Gaussian probability density function with mean $\mu = 370$ and standard deviation $\sigma = 164$ in the interval $[57, 1002]$ and the probability function of the process $\{H\}$ derives from the sampling of an exponential probability density function with mean $\mu = 7.2$ in the interval $[1, 49]$. The quantization interval of the variables is equal to 1.

modelled. Here again, the results of the simulation show the effectiveness of the method proposed in modelling the source.

5. Conclusion

In this paper the authors have presented a method which allows the generation of discrete random variables with a given probability function and autocorrelation sequence. The method can be applied to all cases where the statistical characteristics of a random process are known, in terms of the probability histogram and autocorrelation sequences, and the latter is decreasing with downward convexity. The method can easily be used to simulate traffic sources which meet its validity conditions in statistical terms.

The authors are currently working on modifying the method so as to represent sources which are cyclostationary or pseudo-stationary and they are studying techniques for the generation of the variables of the two component processes when a finite number of points in the experimental autocorrelation sequence does not satisfy the conditions for the validity of the method.

References

M. ANDRONICO ET AL., A model for simulation of correlated VBR video sources. Presented at the Proceedings of the *International Conference on Communication Techonology*, (1994), Shangai, China.

V. CATANIA ET AL., A simulation study of DQDB metropolitan area network. *Simulation Journal*, **62** (1994), 58–70.

V. S. FROST, B. MELAMED, Traffic modelling for telecommunications networks. *IEEE Communications Magazine*, **34** (1994), 70–81

R. GRUNENFELDER ET AL., Characterization of video codecs as autoregressive moving average processes and related queueing system performance. *IEEE Journal on Selected Areas in Communications*, **3** (1991), 284–293.

I. W HABIB, T. N. SAADAWI, Multimedia traffic characteristics in broadband networks. *IEEE Communications Magazine*, **7** (1992), 48–54.

G. E. JOHNSON, Construction of particular random processes. *Proceedings of the IEE*, **2** (1994), 270–285.

G. KARLSSON, M. VETTERLI, Sub-band coding of video for packet networks. *Optical Engineering*, **2** (1988), 574–586.

A. LA CORTE ET AL., Modelling activity in VBR video sources. *Image Communication Journal*, **3** (1991), 167–178.

B. MAGLARIS ET AL., Performance models of statistical multiplexing in packet video communications. *IEEE Transanction on Communication*, **7** (1988), 834–844.

D. MIDDLETON, Canonically optimum threshold detection. *IEEE Transaction on Information Theory*, **2** (1966), 230–243.

M. NOMURA ET AL., Basic characteristics of variable rate video coding in ATM environment. *IEEE Journal on Selected Areas in Communications*, **5** (1989), 752–760.

A. OPPENHEIM, R. SCHAFER, *Digital Signal Processing*. Prentice–Hall, New Jersey, 1974.

C. W. THERRIEN, *Discrete Random Signals and Statistical Signal Processing*, Prentice–Hall, New Jersey, 1992.

H. L. VAN TREES DETECTION, *Estimation and Modulation Theory: Part 1. Wiley*, New York, 1968

Received: October, 1994
Accepted: April, 1995

Contact address:
Maurizio Andronico
Istituto di Informatica e Telecomunicazioni
V.le A. Doria 6
95125 Catania, Italy
phone: +39 95 339449
fax: +39 95 338280
e-mail: mandroni@iit.unict.it

MAURIZIO ANDRONICO received the degree in Electrical Engineering from the University of Catania, Italy, in 1993. Since 1993 he has been co-operating in research on Modelling and Simulation Techniques with the Inst. of Informatics and Telecommunications of the University of Catania. From October 1993 he is a Ph.D. student in Computer Science. His interests include reference parameter recovery algorithms and digital signal processing.

SALVATORE CASALE received his Electrical Engineering degree from the University Palermo, Italy, in 1969. Since 1974 he has been an Associate Professor of Telecommunication Systems at the Inst. of Informatics and Telecommunications of the University of Catania. His current research interest is in High Speed Data Networks, B-ISDN, Multimedia Communications and Digital Signal Processing techniques. He is a member of the IEEE Communication Society.

AURELIO LA CORTE received the degree in Electrical Engineering from the University of Catania, Italy, in 1988. In 1991 he received the Ph.D. degree in Computer Science. From 1993 he is a researcher at the Inst. of Informatics and Telecommunications of the University of Catania. His interests include traffic modelling and digital signal processing.