CHARACTERISATION OF VARIABLE RATE VIDEO CODECS IN ATM TO A GEOMETRICALLY MODULATED DETERMINISTIC PROCESS MODEL.

by J Cosmas and A Odinma-Okafor
Queen Mary and Westfield College, Dept of Electronic Engineering, Mile End Road, London E1 4NS, England.

The modeling of a conditional replenishment (CR) codec is described. The characterisation of data generated from the codec, as a Geometrically Modulated Deterministic Process (GMDP), is presented. Measurements of the sample correlation are made and observations from results regarding the behaviour of the output of codec is compared with the presented model.

1.0 INTRODUCTION

It is expected that video data will be a major traffic on the Asynchronous Transfer Mode (ATM) networks. This will place very large bit rate requirements on the network. A more efficient utilisation of network resources is obtained by compressing the video data, since this reduces the mean load. However, this is at the expense of variance load. In order to study how packet arrivals from codecs affect switching networks, mathematical models of the codecs are required which accurately represents the behaviour of the arrival process.

The objective of this paper is to present the work we have done towards the characterisation of video codecs as Geometrically Modulated Deterministic Processes. Parameters have been obtained for three to eight states models using the CR codec with a head and shoulders, selling products and view from a plane of a city scenes which correspond to low, medium and high activity scenes, respectively.

2.0 DETERMINING THE PARAMETERS OF THE GMDP MODEL

2.1 Description of Model

A GMDP model has been proposed for traffic studies on switch elements. The parameters of a N state GMDP model consists of N(N-1) parameters, namely:
- di - interarrival time for state i (discretely distributed)
- d_i E[S_i] - mean sojourn time in state i (geometrically distributed)
- P_ij - transition probability from state i to j

where i = 0 to N-1

Realistic GMDP parameters are being sought for proposed video codecs so that more realistic traffic studies can be made.

2.2 Unbiased Estimation of GMDP Parameters

Unbiased estimation is the preferred method of determining the GMDP parameters. They can be determined by equating the measured mean, variance and autocovariance to theoretical expressions of the mean, variance and autocovariance for the GMDP model. Derivations of the GMDP model interarrival first and second moments are given in [1] as:

\[ E[d] = \sum_{i=0}^{i=(N-1)} d_i P_i \]

\[ E[d^2] = \sum_{i=0}^{i=(N-1)} d_i^2 P_i \]

Each state i generates cells with an interarrival time of d_i. After a lag of k, d_j(k) is the a new interarrival time. P_i is the steady state probability of being in a state i. P_j\|i(k) is the probability of being in a state j after a lag of k, given state i.

The autocorrelation of the GMDP model is given in [1] as

\[ R(d,d(k)) = E[d \cdot d(k)] \]

\[ = \sum_{i=0}^{i=(N-1)} \sum_{j=0}^{j=(N-1)} d_i d_j(k) P_i P_j\|i(k) \]
The transition matrix defines the probability of transfer to state \( j \), given that present state is \( i \) and is obtained from the mean sojourn times \( d_i E[S_i] \) and the transition probabilities \( p_{ij} \) of the \( N \)-state model.

The transition matrix for the \( N \) state model is given as

\[
X = \begin{bmatrix}
    P_0 & P_1 & 0 & \cdots & 0 \\
    P_0 & P_1 & 0 & \cdots & 0 \\
    \vdots & \vdots & \ddots & \vdots & \vdots \\
    P_0 & P_1 & 0 & \cdots & 0 \\
    0 & 0 & \cdots & P_1 & P_0
\end{bmatrix}
\]

where

\[
P_i = \frac{(E[S_i] - 1)}{E[S_i]}, \\
P_i = (1 - P_i)P_i, \\
E[S_i] \text{ is the expected number of cells produced in state } i.
\]

The transpose of the transition matrix is used to solve for \( P_i \) (the steady state probability of being in a state \( i \)).

The elements of this matrix is used to obtain \( P_\backslash i(k) \). The autocovariance is thus given as

\[
C(d,d(k)) = R(d,d(k)) - E[d]E[d(k)] \\
= C(d,d(k)) = R(d,d(k)) - E[d]^2 
\]

since \( E[d] = E[d(k)] \)

The normalised correlation is given as

\[
C(d,d(k)) / C(d,d)
\]

Suitable expressions of the GMDP model mean and variance can be obtained. However, the autocorrelation expression (equation 3) can only be suitably expressed in matrix form and so an autocovariance expression (equation 4) is not suited for use in an unbiased estimation of GMDP model parameters.

2.3 Biased Estimation of GMDP Parameters

Since an unbiased estimation for the GMDP parameters can not be used, an acceptable biased estimator must be conceived. A state of a video codec is defined as a condition during which there is a specific constant cell arrival rate. A typical video codec could have as many as 6144 different rates and thus would require 6144 states to represent these rates. It is impractical to characterise a codec with this number of states and it has been suggested that the output of a video codec could be approximated to three or more states by grouping inter-packet arrivals.

For characterising the GMDP model, the task of deciding an optimum grouping of inter-cell arrivals into different states is not an easy one as there may be many different combinations that would yield better results. However, experience in running the simulation seem to suggest a necessary criteria for grouping the arrivals into states. The criteria fall into three categories:

1. The intensity of each inter-cell arrival
2. The orderliness in a sequence of arrivals
3. Noisy or erratic behaviour of the arrivals

These criteria provided a starting point for grouping the inter-cell arrivals into states. The intervals once chosen are then varied within its neighbourhood. A search process is used to determine the intervals which yield the best mean arrival rate for a model that best matches the mean of the codec. Those simulations whose differences between the mean for the model and codec is less than a statistical tolerance of 99% confidence [2] are then selected.

After defining the state intervals the model parameters can be measured directly. The method of measuring the model parameters is now described. The transitional probability from state \( i \) to \( j \) for an \( N \)-state model,

\[
\text{mean sojourn time in state } i \\
\text{mean interarrival time for state } i
\]

\[
d_i = T_i / V_i
\]

\( n_{ij} \) is the number of transitions from state \( i \) to state \( j \), \( T_i \) is the total time spent in state \( i \), \( V_i \) is the total number of visits to state \( i \), and \( N_i \) is the total number of packets generated in state \( i \).
2.4 Statistical Measurements of Video Codec Output

The video codec sample mean, second moment and autocorrelation were also directly measured so that subsequent comparison with the GMDP model mean, second moment and autocorrelation can be made. For a sample space of \( N \), the sample mean is

\[
E[d_1] = \frac{1}{N} \sum_{n=0}^{N} d(n)
\]

the sample second moment is

\[
E[d^2] = \frac{1}{N} \sum_{n=0}^{N} d^2(n)
\]

the sample autocorrelation is

\[
R(d(n),d(n+k)) = \frac{1}{N-k} \sum_{n=0}^{N-k} d(n)d(n+k)
\]

where \( d \) = packet inter-arrival time,

With a little manipulation the mean, second moment and autocorrelation can be converted to the variance and normalised correlation. The latter was generated in the experimental results.

3.0 SIMULATING THE VIDEO CODEC

3.1 Experimental Set up

Three main equipments are used, namely: a Halequin image frame grabber with transputer, VP405 Philips video disc player and IBM compatible Epson host PC. The image frame grabber consists of a single T800 transputer with a 1M byte RAM memory and two 256k byte image buffers. The video disc player is connected to the frame grabber via a composite video link and externally controlled by the PC through an RS232 link. The host PC computer controls with the aid of development tools (such as Turbo Pascal and Transputer Development Systems (TDS2)) the transputer and the video disc player. Occam programs are developed on the TDS2 on the PC and loaded onto the transputer. For further details of the experiment refer to [3].

The simulation is performed using a sequence of frames from a typical video phone scene which is processed a frame at a time. Each frame is captured and is processed. The monochrome intensities are sampled at a pixel frequency in accordance with the algorithm being simulated. Thereafter, coded results are then passed to the PASCAL programs on the host PC for characterisation and statistical analysis.

3.2 Conditional Replenishment Codec

The description of the algorithm is given by Chin and Goode [4]. In the algorithm frames are examined pixel by pixel for areas where significant spatial changes of temporal differences has occurred and only those changes between two consecutive frames in a sequence of frames are transmitted. Those areas with no changes are still transmitted as a run length code.

The conditional replenishment codec stores three frames for coding purposes: new, filtered previous and a reference frame. Frames are coded with the aid of a run length coder, temporal filter, activity detector and packetiser for statistical purposes. The operation of the codec is such that a new frame is passed through a temporal filter which compares it with a previous filtered image frame. The new image frame is then filtered for noise. This is necessary in order to avoid unnecessary information being sent since the conditional replenishment can not differentiate between a pixel change due to noise and a pixel change due to movement. The filtered new frame is then used to update the filtered previous frame and also fed into an activity detector which is made up of a pixel comparator with modulus output, a five tap filter and threshold comparator. The comparator compares the pixel differences between the filtered new image frame and the reference frame and its absolute value fed into the five tap filter. The output of the five tap filter which represents spatial activity of temporal change is compared with a threshold value. If the filtered output is greater than the threshold then sufficient spatial activity of temporal change is detected at the center pixel. The reference frame is then updated and packetised. On the contrary, if the filtered output is less than the threshold value then insufficient spatial activity of temporal change at the center pixel is detected and a run length is accumulated for duration of inactivity.
3.3 Organisation of Cells

Packets are organised as 32 bytes information field and 4 bytes signalling field. The information field is subdivided as follows: 24 bytes are used for changed image and run length coded data, 3 bytes for end of line/end of frame (EOL/EOF) data, 3 bytes for changed image and a run length coded flag and the remaining two bytes for error detection and correction of the EOL/EOF. Flag bits are used to differentiate between either changed image and run length coded data which may be packetised. If EOL or EOF occurs then the appropriate bit is set.

4.0 RESULTS

Three simulation runs of the conditional replenishment codec on typical scenes have produced parameters for 3, 5 and 8 state models and also direct measurements of the sample mean, variance and normalised correlation. These scenes have been given the following names which are mnemonic:
(1) Head and Shoulders (low activity)
(2) Selling Products (medium activity)
(3) Aeroplane over City (high activity)

Tables 1, 2 and 3 gives the measured 3, 5 and 8 state model parameters and the measured and model, first and second moments for scenes (1), (2) and (3) respectively.

Graph 1 gives the normalised correlation for (1). Graph 2 gives the 3, 5 and 8 state model normalised correlation for (1). Graph 3 gives the normalised correlation for (2). Graph 4 gives the 3, 5 and 8 state model normalised correlation for (2). Graph 5 gives the normalised correlation for (3). Graph 6 gives the 3, 5 and 8 state model normalised correlation for (3).

The results show that as the number of states are increased from 3 to 8, the matching of the GMDP model variance with the sample variance increases but does not significantly improve. However the number of parameters required increases with the number of states used in accordance with the relation $N(N+1)$. 
Graph 3: Measured Sample Normalised Correlation–Lag Selling Products Scene

Graph 4: State Model Normalised Correlation–Lag Selling Products Scene

Graph 5: Measured Sample Normalised Correlation–Lag Plane over City Scene

Graph 6: State Model Normalised Correlation–Lag Aeroplane over City Scene
A number of observations on the characterisation of the N-state model were made:

(1) The sample mean, variance and normalised correlation for the three scenes differ considerably.

(2) The sample normalised correlation, recorrelated at lags of unit frame intervals (measured in cells/frame). Since the number of cells/frame differed from frame to frame, the location of recorrelation on the lag axis also differs from frame to frame. This phenomenon contributed to the spreading of the recorrelation on the lag axis and the reduction of the recorrelation effect.

(3) The normalised correlation spans values from 1 to near 0, without approaching -1. This means that the successive cell rates from the video codec changes gradually rather than abruptly across cell rate extremes and that successive rates are grouped by size (there are groups of large rates and groups of low rates which keep separate).

(4) The theoretical normalised correlation does not match the sample normalised correlation for the 3, 5 and 8 state models. The theoretical normalised correlation decorrelates very quickly.

(5) No recorrelation occurs for the theoretical normalised correlation.

(6) The degree of variation or dispersion about the average is much greater for the samples than for the 3, 5 and 8 state models. This is reflected by the considerable differences in the sample and theoretical variances.

(7) It is impossible to derive the parameters of the N state model from measures of mean, variance and autocorrelation (unbiased estimation) because the expression of the autocorrelation is too complicated.

5.0 OBSERVATIONS AND CONCLUSIONS

In conclusion, the present method of parameterising the GMDP model is not suited to characterising video sources at the action level because

(1) It does not exhibit recorrelation.

(2) Too many states and therefore far too many parameters are required to reproduce the slow decorrelation.

(3) Unbiased estimation of parameters cannot be obtained.

However periodic GMDP or periodic discrete time discrete state models, as described in [1], do exhibit recorrelation. Their transition matrix are highly diagonal. Some periodic state models can be efficiently implemented using filters and formally parameterised using signal processing techniques as described in [6].

ACKNOWLEDGEMENTS

The authors gratefully acknowledge support from the CEC RACE programme, project R1022 "Technology for ATD" and to Professor Paul Kuehn, University of Stuttgart for suggesting the GMDP model.

REFERENCES


### Table 1: Head and Shoulder Scene

<table>
<thead>
<tr>
<th>States</th>
<th>Inter-arrival time (pixel interval = 0.6104μs)</th>
<th>( \text{mean(μs)} )</th>
<th>( \text{σ(μs)}^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>measured</td>
<td>225.14</td>
<td>121541.55</td>
</tr>
<tr>
<td></td>
<td>model</td>
<td>225.04</td>
<td>56349.66</td>
</tr>
<tr>
<td></td>
<td>( i ) (pixels)</td>
<td>( d_i [E[S_i] (μs) )</td>
<td>( d_{i-1}^{-1} ) (cells/sec)</td>
</tr>
<tr>
<td>0</td>
<td>24 - 30</td>
<td>23.40</td>
<td>64106.85</td>
</tr>
<tr>
<td>1</td>
<td>31 - 55</td>
<td>30.32</td>
<td>39415.38</td>
</tr>
<tr>
<td>2</td>
<td>56 - 6144</td>
<td>2350.19</td>
<td>3811.95</td>
</tr>
</tbody>
</table>

### Table 2: Selling Products Scene

<table>
<thead>
<tr>
<th>States</th>
<th>Inter-arrival time (pixel interval = 0.6104μs)</th>
<th>( \text{mean(μs)} )</th>
<th>( \text{σ(μs)}^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>measured</td>
<td>73.48</td>
<td>73828.06</td>
</tr>
<tr>
<td></td>
<td>model</td>
<td>73.48</td>
<td>7217.89</td>
</tr>
<tr>
<td></td>
<td>( i ) (pixels)</td>
<td>( d_i [E[S_i] (μs) )</td>
<td>( d_{i-1}^{-1} ) (cells/sec)</td>
</tr>
<tr>
<td>0</td>
<td>24 - 30</td>
<td>24.14</td>
<td>63940.86</td>
</tr>
<tr>
<td>1</td>
<td>31 - 55</td>
<td>35.02</td>
<td>39434.93</td>
</tr>
<tr>
<td>2</td>
<td>56 - 6144</td>
<td>332.06</td>
<td>9239.11</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>States</th>
<th>Inter-arrival time (pixel interval = 0.6104μs)</th>
<th>( \text{mean(μs)} )</th>
<th>( \text{σ(μs)}^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>measured</td>
<td>77.03</td>
<td>76843.18</td>
</tr>
<tr>
<td></td>
<td>model</td>
<td>77.03</td>
<td>75843.18</td>
</tr>
<tr>
<td></td>
<td>( i ) (pixels)</td>
<td>( d_i [E[S_i] (μs) )</td>
<td>( d_{i-1}^{-1} ) (cells/sec)</td>
</tr>
<tr>
<td>0</td>
<td>24 - 30</td>
<td>25.48</td>
<td>63697.29</td>
</tr>
<tr>
<td>1</td>
<td>31 - 55</td>
<td>34.92</td>
<td>39233.95</td>
</tr>
<tr>
<td>2</td>
<td>56 - 6144</td>
<td>804.72</td>
<td>2810.37</td>
</tr>
</tbody>
</table>
Table 3: Aeroplane over City Scene

1 States
Inter-arrival time (pixel interval = 0.6104μs)

<table>
<thead>
<tr>
<th>i (pixels)</th>
<th>$d_i[E[S_j]]$ (μs)</th>
<th>$d_j^{-1}$ (cells/sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 24 - 30</td>
<td>24.05</td>
<td>61893.16</td>
</tr>
<tr>
<td>1 31 - 55</td>
<td>43.10</td>
<td>40167.26</td>
</tr>
<tr>
<td>2 56 - 6144</td>
<td>177.81</td>
<td>12805.83</td>
</tr>
</tbody>
</table>

$\Pi_0$ $\Pi_1$ $\Pi_2$
0 0.0000 0.5278 0.3722
1 1.3484 0.0000 0.6516
2 2.2522 0.7478 0.0000

5 States
Inter-arrival time (pixel interval = 0.6104μs)

<table>
<thead>
<tr>
<th>i (pixels)</th>
<th>$d_i[E[S_j]]$ (μs)</th>
<th>$d_j^{-1}$ (cells/sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 24 - 30</td>
<td>23.88</td>
<td>61981.08</td>
</tr>
<tr>
<td>1 31 - 55</td>
<td>42.78</td>
<td>40216.89</td>
</tr>
<tr>
<td>2 56 - 6144</td>
<td>51.90</td>
<td>24397.59</td>
</tr>
<tr>
<td>3 241 - 1300</td>
<td>122.70</td>
<td>12980.94</td>
</tr>
<tr>
<td>4 1301 - 6144</td>
<td>1038.43</td>
<td>2772.51</td>
</tr>
</tbody>
</table>

$\Pi_0$ $\Pi_1$ $\Pi_2$ $\Pi_3$ $\Pi_4$
0 0.0000 0.6237 0.1592 0.2057 0.0013
1 1.3484 0.0000 0.2909 0.3590 0.0017
2 2.1690 0.5162 0.0000 0.3091 0.0057
3 3.1561 0.4945 0.3061 0.0000 0.0433
4 4.0752 0.0074 0.0280 0.8894 0.0000

8 States
Inter-arrival time (pixel interval = 0.6104μs)

<table>
<thead>
<tr>
<th>i (pixels)</th>
<th>$d_i[E[S_j]]$ (μs)</th>
<th>$d_j^{-1}$ (cells/sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 24 - 30</td>
<td>24.50</td>
<td>61964.18</td>
</tr>
<tr>
<td>1 31 - 55</td>
<td>43.63</td>
<td>40168.01</td>
</tr>
<tr>
<td>2 56 - 6189</td>
<td>52.09</td>
<td>24528.09</td>
</tr>
<tr>
<td>3 81 - 240</td>
<td>110.81</td>
<td>13777.55</td>
</tr>
<tr>
<td>4 241 - 280</td>
<td>214.26</td>
<td>6397.73</td>
</tr>
<tr>
<td>5 281 - 530</td>
<td>416.03</td>
<td>4178.30</td>
</tr>
<tr>
<td>6 531 - 1300</td>
<td>819.71</td>
<td>2149.62</td>
</tr>
<tr>
<td>7 1301 - 6144</td>
<td>1158.90</td>
<td>878.16</td>
</tr>
</tbody>
</table>

$\Pi_0$ $\Pi_1$ $\Pi_2$ $\Pi_3$ $\Pi_4$ $\Pi_5$ $\Pi_6$ $\Pi_7$
0 0.0000 0.6519 0.1852 0.1610 0.0001 0.0008 0.0005 0.0006
1 1.3831 0.0000 0.3174 0.2971 0.0010 0.0004 0.0002 0.0009
2 2.1800 0.5389 0.0000 0.2746 0.0020 0.0007 0.0003 0.0036
3 3.1494 0.4835 0.3095 0.0000 0.0347 0.0189 0.0015 0.0025
4 4.0000 0.0131 0.0392 0.5744 0.0000 0.3198 0.0522 0.0013
5 5.0014 0.0055 0.0006 0.3581 0.3402 0.0000 0.2782 0.0069
6 6.0035 0.0000 0.0000 0.0517 0.1724 0.7448 0.0000 0.0276
7 7.4513 0.0177 0.0089 0.1770 0.0708 0.1239 0.1504 0.0000