Three Notes on the SAM Video Statistical Model

Mário Macedo

INESC-ID, 1000-029 Lisboa, Portugal; and Faculdade de Ciências e Tecnologia da Universidade Nova de Lisboa, 2829-516 Caparica, Portugal

Phone: (351) 918293671

Fax: (351) 212948532

e-mail: mmm@fct.unl.pt

Abstract The SAM model is a seasonal ARIMA (Autoregressive Integrated Moving Average) time series model that has been proposed to capture and reproduce the behavior of video sources, namely those that follow the different MPEG-4 Part 2, Advanced Video Coding, and Scalable Video Coding techniques. The statistical model is simple, elegant, and is apparently suitable to the simulation of video traces. However, three problems were found that needed to be enlightened. Firstly, the SAM creators derived a mathematical expression for its video frame generation process that simply does not work, and seems to be wrong. In this paper, the author tries to fix the problem, by deriving another mathematical expression. The two competing expressions are compared through exhaustive computer simulations, and the results seem to confirm the validity of the new expression. Secondly, even with the right expression, the SAM model generates traces with mean frame sizes, maximum frame sizes, and average absolute deviations from the mean frame size that are also substantially different from the originals. Thirdly, and finally, the SAM model does not model adequately most video traces, namely those presenting structures like steps, and impulse trains.

Keywords: Video Models, Time Series Analysis, Seasonal ARIMA, R project

1 Introduction

Several different statistical models have been proposed to capture and reproduce the behavior of video sources. Among them, a seasonal ARIMA process (seasonal Autoregressive Integrated Moving Average, SARIMA, see [1][2]) has been developed with the purpose of emulating video sources that use the MPEG-4 Part 2, the Advanced Video Coding (AVC), and the Scalable Video Coding (SVC) techniques. This statistical model was named SAM model (Simplified Seasonal ARIMA Model, see [3] through [9]), and its frame generator expression is given in [7], and [8]:

$$X_t = X_{t-1} + \varphi X_{t-1} - \varphi X_{t-2} + \Phi_s X_{t-s} - \varphi \Phi_s X_{t-s-1} - \Phi_s X_{t-s-1} + \\ + \varphi \Phi_s X_{t-s-2} - \theta \varepsilon_{t-1} - \Theta_s \varepsilon_{t-s} + \theta \Theta_s \varepsilon_{t-s-1} + \varepsilon_t \quad (1)$$

Where X_t represents the size in bits of the frame generated at time t, and ε_t is the moving average term, a random variable with a normal distribution, of zero mean, and standard deviation σ. ε is also named as the innovation term, or error term. s stands for the seasonal period. The four parameters of the expression are φ, the autoregressive coefficient; Φ_s, the seasonal autoregressive coefficient; θ, the moving average coefficient; Θ_s, the moving average seasonal coefficient. However, for a computational purpose, the fifth parameter σ, the standard deviation of the moving average term, ε, has also to be considered.

Expression (1) is very appealing because a video trace could be generated by a simple mathematical process, with only those five parameters, and two variables: the past values of the size in bits of the frames, the variable X; and the present and past values of the moving average term, the variable ε. Those five parameters can be extracted from the real video traces, by appropriate statistical software, like the *arima* time series analysis function of the R open-source statistical and graphics project [10], which was used by the SAM authors. The SAM model framework also includes one website [11], with the traces of 54 H264/AVC high definition videos, and their respective parameters.

Expression (1) captured immediately our attention, for its simplicity, and for the prompt availability of its five parameters in the SAM model website, and simulations were run in order to evaluate its behavior.

2 First note: expression (1) is inappropriate

Fig. 1 shows a video trace of a typical video, namely the third of the site (the reason for this choice will be clear later). The trace consists of 3106 frames. It can be seen that there is a seasonal effect with a period s equal to 24 frames, revealed by the peak values of the series, which correspond to the IDR frames of the video trace. All the 54 videos of the SAM model website present a period of 24 frames.

Simulations were run with the application of expression (1). As there is one regressive term of $t - s - 2$, the first s+2=26 values of the simulation were taken from the 26 initial values of the real video trace file. Then the remaining 3106-26 values were successively computed by expression (1). The simulator was written in the C language, and comprised a Mersenne Twister pseudo-Random Number Generator [12].

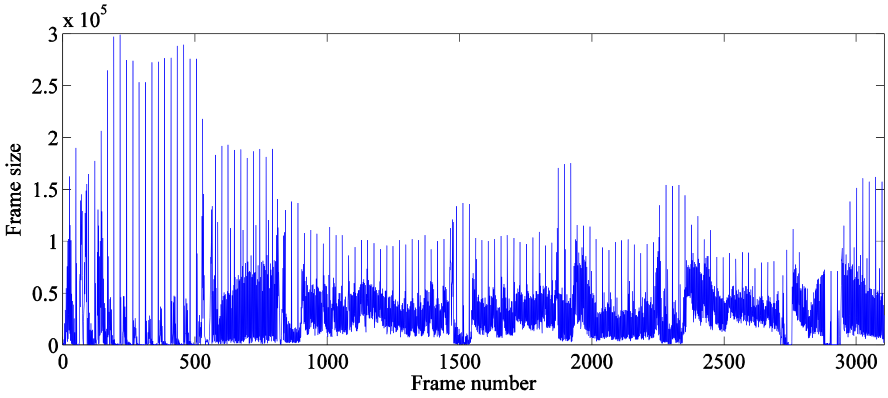

Fig. 1. Frame trace of video 3, cars_lg_framesonly.txt.

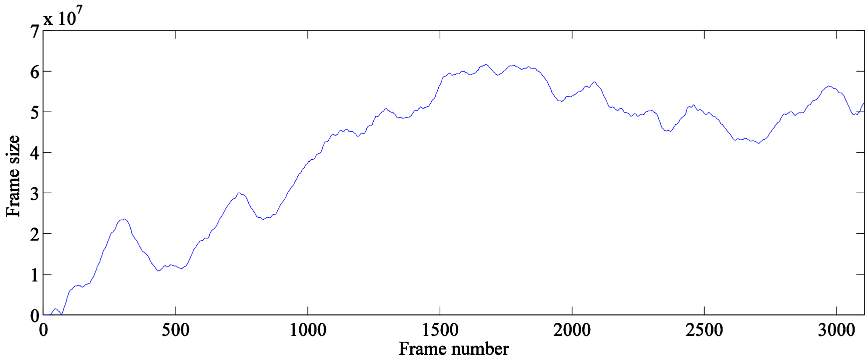

Fig. 2. Frame trace obtained by expression (1) for video 3.

Fig. 2 shows a typical generated trace obtained by expression (1). The trace presents a strong trend, and the frame sizes reach very high values, around 600,000,000 bits. After some simulation runs with other videos, it seemed that there was a problem with expression (1), as almost all (actually, all but video 33) the simulations resulted into strong trends, and very high frame sizes.

3 How expression (1) is derived

Reference [7] shows how expression (1) was derived. Without lack of accuracy, here is presented a shorter version of that process. First, the backward operator B is presented:

$$B^j X = X_{t-j} \qquad (2)$$

And also the differencing operator:

$$\nabla^d X_t = (1-B)^d X_t \qquad (3)$$

A $SARIMA(p,d,q) \times (P,D,Q)^s$ process can be expressed by (see also [1]):

$$\varphi_p(B)\Phi_P(B^s)\nabla^d \nabla_s^D X_t = \theta_q(B)\Theta_Q(B^s)\varepsilon_t \qquad (4)$$

Where $\varphi_p(B) = 1 - \varphi_1 B - \cdots - \varphi_p B^p$, $\Phi_P(B^s) = 1 - \Phi_1 B^s - \cdots - \Phi_P B^{Ps}$, $\theta_q(B) = 1 - \theta_1 B - \cdots - \theta_q B^q$, and $\Theta_Q(B^s) = 1 - \Theta_1 B^s - \cdots - \Theta_Q B^{Qs}$, are the polynomials.

Now, the SAM model authors claim that the respective video traces can be modeled by a $SARIMA(1,0,1) \times (1,1,1)^s$ equation, i.e. that:

$$SAM^s = SARIMA(1,0,1) \times (1,1,1)^s \qquad (5)$$

With s being equal to 24, for all the 54 video traces of the website.

Now, expression (5), applied to expression (4), makes this last one equal to:

$$\varphi(B)\Phi(B^s)\nabla_s X_t = \theta(B)\Theta(B^s)\varepsilon_t \qquad (6)$$

Expression that values:

$$(1 - \varphi B)(1 - \Phi_s B^s)(1 - B^s)X_t = (1 - \theta B)(1 - \Theta_s B^s)\varepsilon_t \qquad (7)$$

Then, expression (7) must give expression (1) after solving it.

The first problem that was noticed with expression (1) is the following: expression (7) only gives as result expression (1), if a term equal to $(1 - B)$ is considered, in the place of the term$(1 - B^s)$, as it is shown next.

Firstly, expression (7) is written, while being modified to include the term $(1 - B)$, in the place of the term $(1 - B^s)$:

$$(1 - \varphi B)(1 - \Phi_s B^s)(1 - B)X_t = (1 - \theta B)(1 - \Theta_s B^s)\varepsilon_t \qquad (8)$$

By multiplying the first two terms of the left side of expression (8), and the first two terms of the right side, it is obtained:

$$(1 - \Phi_s B^s - \varphi B + \varphi\Phi_s B^{s+1})(1 - B)X_t = (1 - \Theta_s B^s - \theta B + \theta\Theta_s B^{s+1})\varepsilon_t \; (9)$$

Now, the first two terms of the left side of expression (13) are multiplied:

$$(1 - \Phi_s B^s - \varphi B + \varphi\Phi_s B^{s+1} - B + \Phi_s B^{s+1} + \varphi B^2 - \varphi\Phi_s B^{s+2})X_t =$$
$$(1 - \Theta_s B^s - \theta B + \theta\Theta_s B^{s+1})\varepsilon_t \qquad (10)$$

Carrying out the multiplications of both sides of expression (14), and applying the backward operator:

$$X_t - \Phi_s X_{t-s} - \varphi X_{t-1} + \varphi\Phi_s X_{t-s-1} - X_{t-1} + \Phi_s X_{t-s-1} + \varphi X_{t-2} -$$
$$-\varphi\Phi_s X_{t-s-2} = \varepsilon_t - \Theta_s \varepsilon_{t-s} - \theta\varepsilon_{t-1} + \theta\Theta_s \varepsilon_{t-s-1} \qquad (11)$$

This expression, after rearranging the terms, results in expression (1).

4 Deriving a new expression

A second problem with expression (1) was also found. The notation adopted by the *arima* function has plus signs on the moving average coefficients. This is the notation followed, for instance, by the Chatfield's book [2], as opposed to the notation of the Box et al.'s book [1], which has minus signs on the moving average coefficients. And this last one was the one used by the SAM model authors.

Therefore, another alternative expression was derived for the SAM model. Firstly, expression (7) is written with the positive signs on the moving average coefficients:

$$(1-\varphi B)(1-\Phi_s B^s)(1-B^s)X_t = (1+\theta B)(1+\Theta_s B^s)\varepsilon_t \qquad (12)$$

By multiplying the first two terms of the left side of expression (12), as well as the first two terms of the right side, it is obtained:

$$(1-\Phi_s B^s - \varphi B + \varphi\Phi_s B^{s+1})(1-B^s)X_t =$$
$$(1+\Theta_s B^s + \theta B + \theta\Theta_s B^{s+1})\varepsilon_t \qquad (13)$$

By multiplying the first two terms of the left side of expression (13), it is obtained:

$$(1-\Phi_s B^s - \varphi B + \varphi\Phi_s B^{s+1} - B^s + \Phi_s B^{2s} + \varphi B^{s+1} - \varphi\Phi_s B^{2s+1})X_t =$$
$$(1+\Theta_s B^s + \theta B + \theta\Theta_s B^{s+1})\varepsilon_t \qquad (14)$$

Carrying out the multiplications of both sides of expression (14), and applying the backward operator:

$$X_t - \Phi_s X_{t-s} - \varphi X_{t-1} + \varphi\Phi_s X_{t-s-1} - X_{t-s} + \Phi_s X_{t-2s} + \varphi X_{t-s-1} -$$
$$-\varphi\Phi_s X_{t-2s-1} = \varepsilon_t + \Theta_s\varepsilon_{t-s} + \theta\varepsilon_{t-1} + \theta\Theta_s\varepsilon_{t-s-1} \qquad (15)$$

And, finally, rearranging the terms of expression (15), it is derived:

$$X_t = \varphi X_{t-1} + X_{t-s} + \Phi_s X_{t-s} - \varphi X_{t-s-1} - \varphi\Phi_s X_{t-s-1} - \Phi_s X_{t-2s} +$$
$$+\varphi\Phi_s X_{t-2s-1} + \theta\varepsilon_{t-1} + \Theta_s\varepsilon_{t-s} + \theta\Theta_s\varepsilon_{t-s-1} + \varepsilon_t \qquad (16)$$

Expression (16) is not as intuitive as expression (1), because it also contains regressions of the -2*s* order, i.e. dependency on past values also distancing the double of the period, but it produces much better results, as will be seen later.

5 Assessing the validity of the new expression

Fig. 3 shows a trace resulting from a simulation run that used expression (16). In this case, a vector of 2*s*+1=49 initial values was taken from the real video trace file.

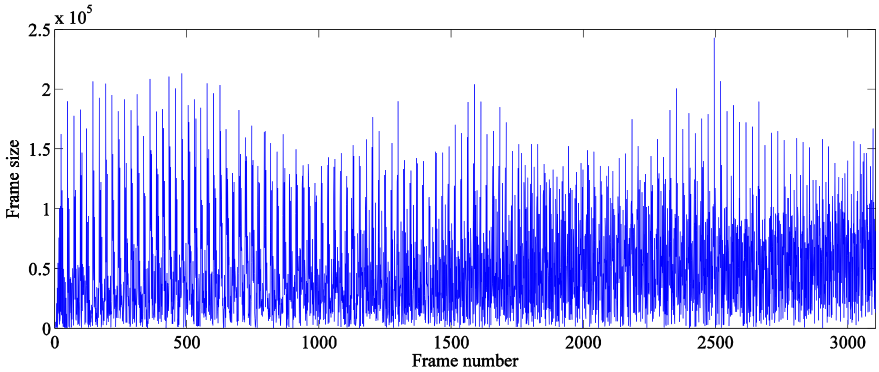

Fig. 3. Frame trace obtained by expression (16) for video trace 3.

The trace is different from the original (Fig. 1), but is reasonable. It presents the same peaks with a period of 24, and the sizes of the frames reach a maximum of around 250,000 bits, when compared to the original of around 300,000 bits. However, the mean value of these new frame sizes seems to be higher. Actually, an average frame size of 49,274 bits was computed, in comparison with the 28,000 bits of the initial trace. And the trace also has a greater randomness than the original trace.

In order to assess the validity of expression (16), simulations were run for each of the 54 videos, and for both expressions (1), and (16). For each pair of the video number and one of the two expressions, 5,000 simulation runs were executed. Then the simulation software computed three statistics, named respectively as *A*, *M*, and *D* (where the word average as the meaning of the arithmetic mean).

The ***A* (after average) statistic** is the ratio between the average frame size of all the frames generated by the 5,000 simulation runs, and the average frame size of the original video frame trace. Its mathematical expression is given by:

$$A = \frac{\frac{\sum_{s=1}^{5000} \sum_{i=1}^{N} FSim_{s,i}}{5000 \times N}}{\overline{FReal}} \qquad (17)$$

With:

$$\overline{FReal} = \frac{\sum_{i}^{N} FReal_{i}}{N} \qquad (18)$$

Where s is the index of the simulation, i is the index of the frame in the video trace, and N is the number of frames of the video trace. $FSim_{s,i}$ is the size in bits of the frame i of simulation s. $FReal_i$ is the size of the frame i of the real original video trace. So, $\overline{FReal}$ is the mean frame size of the original video trace.

The ***M*** **(after maximum) statistic** is the ratio between the maximum frame size obtained in all 5,000 simulation runs, and the maximum frame size of the original video trace, i.e. given by the following expression:

$$M = \frac{\max_{s,i}\{FSim_{s,i}\}}{\max_{i}\{FReal_{i}\}} \qquad (19)$$

Finally, the most complex statistic, the ***D*** **(after deviation) statistic**, which is also a ratio, can be computed as follows: for each simulation run, the average absolute deviation from the mean frame size of the simulation run is computed; then the average absolute deviations obtained by all 5,000 simulations runs are averaged; and finally this final average value is divided by the average absolute deviation from the mean frame size of the original video trace. It can be described by expression (20):

$$\mathrm{D} = \frac{\frac{\sum_{s=1}^{5000}\left(\frac{\sum_{i=1}^{N}\left|\mathrm{FSim}_{s,i} - \overline{\mathrm{FSim}_{s}}\right|}{N}\right)}{5000}}{\frac{\sum_{i=1}^{N}\left|\mathrm{FReal}_{i} - \overline{\mathrm{FReal}}\right|}{N}} \qquad (20)$$

With $\overline{FSim_s}$, being the mean frame size of the simulation with index s, i.e., given by the following expression:

$$\overline{FSim_s} = \frac{\sum_{i=1}^{N} FSim_{s,i}}{N} \qquad (21)$$

Table I shows the results of this set of simulations. Note that the ideal value for the statistics is 1. This value means that the simulated traces have the same average values as the original video traces. Table I reveals the inadequacy of expression (1), as this expression generates traces with high average frame sizes, with high maximum frame sizes, and also with high absolute deviations from the mean values. These values can be hundreds and even thousands of times greater than the originals. For only one video (the 33rd), expression (1) is stable and gives apparently acceptable results.

Video	Expression (1)			Expression (16)		
	A	*M*	*D*	*A*	*M*	*D*
1	4750.482	4481.136	2675.275	6.960	2.215	1.720
2	1221.510	634.939	627.658	3.114	1.883	1.821
3	512.763	316.563	311.870	2.554	1.745	1.815
4	868.190	577.273	435.826	6.293	4.520	3.907
5	804.215	1628.151	744.294	0.852	1.706	0.853
6	775.427	697.872	405.688	1.630	1.731	1.076
7	710.665	271.885	212.216	2.232	1.038	0.909
8	270.990	647.167	226.877	0.610	1.601	0.594
9	81.588	50.693	42.378	3.897	3.019	2.210
10	536.223	683.832	329.272	1.145	1.363	1.133
11	662.102	356.059	371.922	2.543	1.523	1.770
12	2554.604	1299.122	1821.847	1.658	0.748	1.068
13	1045.703	677.042	642.094	1.284	0.534	1.030
14	832.554	663.955	433.509	1.454	1.008	0.828
15	1795.638	1902.078	1120.836	1.183	1.321	0.894
16	1160.099	1540.930	766.953	1.115	1.392	1.383
17	571.566	445.343	231.950	4.346	4.014	2.094
18	461.454	718.580	440.155	1.493	2.024	1.331
19	2861.832	2027.570	1681.472	1.103	0.651	1.012
20	1406.302	965.791	902.972	1.113	0.926	0.854
21	3768.559	5093.408	2638.142	8.879	3.076	2.287
22	663.490	1373.622	566.367	0.678	1.377	0.670
23	1190.396	1265.379	959.214	2.000	1.271	1.513
24	1519.289	2672.692	1067.845	10.692	20.593	9.547
25	1157.240	857.612	676.873	1.693	1.637	1.316
26	58.877	60.758	40.089	3.281	2.848	1.867
27	805.512	628.783	453.484	1.052	0.995	0.876
28	1242.284	1205.203	701.882	1.772	1.831	0.973
29	2373.664	1526.639	1106.222	4.347	1.439	1.513
30	63.973	67.450	48.947	2.848	3.615	2.208
31	566.603	712.589	428.172	1.090	1.297	1.035
32	2325.458	1979.097	4041.080	3.818	3.362	7.297
33	1.267	2.550	0.642	1.318	2.026	0.817
34	3007.540	4553.295	1828.234	1.093	1.065	0.988
35	1339.815	1100.135	456.799	1.785	1.494	0.546
36	340.252	853.652	429.686	0.453	0.816	1.040
37	1358.881	1721.126	926.336	1.853	2.179	1.474
38	1594.144	1418.185	853.734	1.979	1.539	1.166
39	1461.281	914.752	744.632	1.491	0.883	0.797
40	966.623	1548.624	713.936	1.652	1.633	0.922
41	1798.031	2166.339	1389.816	1.162	1.035	0.906
42	592.241	673.629	330.032	2.358	1.683	1.533
43	787.371	300.447	314.545	3.304	1.740	1.729
44	1486.187	1314.709	808.749	1.308	1.047	0.913
45	2369.062	1335.586	1397.847	2.212	1.488	1.694
46	2768.325	1251.339	1409.690	5.112	0.968	2.267
47	4105.210	2166.247	2585.053	4.775	1.026	1.684
48	1105.222	374.183	335.520	3.850	1.747	1.727
49	202.338	195.872	129.539	5.589	5.605	3.878
50	5016.360	1586.988	2648.440	4.238	0.932	2.532
51	2936.933	1401.891	1679.987	2.987	0.798	1.720
52	1180.232	876.893	671.970	1.900	1.521	1.392
53	2770.327	1574.625	1778.304	4.223	1.248	2.651
54	4089.241	1059.786	2050.961	2.085	0.542	1.198

Table I. Average frame size ratios, *A*, maximum frame size ratios, *M*, and average absolute deviations from the mean frame size ratios, *D*, obtained by the two competing expressions.

But this occurrence may be due to a coincidence, because the *arima* function computes coefficients for a valid expression, which are fed into the wrong expression (1), but these coefficients may stabilize expression (1), and give apparently acceptable values. In all other cases, expression (1) generates traces that are totally different from the original traces. Table I shows that expression (16) generates traces that are much more similar to the original traces, as the three statistics are much closer to 1 than those produced by expression (1). Note also that video 3 is an average video with respect to those statistics.

These results demonstrate that expression (16) is more adequate than expression (1), in order to be the right SAM model video trace generator.

In my opinion the SAM model creators never used expression (1) in their simulations, and, by the contrary, they used sophisticated software which calculated the simulated frame traces without the need to use the referred expression, while however being configured with the parameters calculated by the *arima* function. In fact, in references [4], and [5], the authors refer that they used some R project functions to build the SAM model frame generator, and also mention the use of the SAS software [13] in their analysis. But they don't explain why they don't use expression (1) directly as was done here.

6 Second note: the average frame sizes, and the average absolute deviations from the mean frame sizes, generated by the SAM model, are substantially different from the originals

Moreover, the simulation results also show that the SAM model, even with the new expression (16), is not able to reproduce adequately a significant number of video traces. The statistics of the average values (A), and the statistic of the absolute deviations from the mean values (D), the most important of our statistics, are sufficient to show that the SAM model is far from ideal, as many video traces present values that are substantially far from the ideal value of 1. Note that having the statistics *A*, *M*, and *D*, close to 1, is a necessary condition for a model reproducing adequately the respective video traces (but not a sufficient condition).

Returning back to the results, it is clearly unacceptable that a model gives an average frame size that is 6.96 times higher than the original, which is the case of video trace 1. Or even that the average absolute deviation from the mean frame size is 1.72 times greater than the original (video 1). The same can be said about video traces 21, and 24, which present even worse values. And a total of 45 videos, among all the 54 videos of the website, present an *A* statistic, or a *D* statistic, that is out of the range of between 0.75 and 1.25.

Another drawback is the fact that the two statistics, *A* and *D*, are in general substantially different from each other (being the *A* statistic usually higher than the *D* statistic), meaning that the obtained frame sizes are not even proportional to the originals, but simply different. These results mean that the traces generated by the SAM model may fail the goal of reproducing satisfactorily the video traces by means of simple mathematical simulations.

None of the references [3] through [9] present such an exhaustive study of the SAM model frame generator behavior, with respect to its ability to reproduce in a suitable way the 54 video traces of the website. Alternatively, references [3], [4], [5], [6], [8], and [9], usually compare the SAM model to other different video mathematical models, with respect to the values obtained in some given statistics, but these values don't inform us about the absolute ability of the models to reproduce the video traces, they only serve to rank the models regarding to the fit to the data scores that are only meaningful for comparison purposes. For instance, in reference [9], a comparison is made of the SAM model to the Autoregressive (AR), and to the Autoregressive Integrated Moving Average (ARIMA) models, on the scores obtained by the following statistics: mean absolute error (MAE), mean absolute relative error (MARE), normalized mean square error (NMSE), and the Akaike's Information Criterion (AIC). Note that an AR model of order p is defined by the following expression:

$$X_t = \sum_{i=1}^{p} \varphi_i X_{t-i} + \varepsilon_t \qquad (22)$$

And an ARIMA model of orders (p,d,q) can be deduced from the following expression:

$$\left(1 - \sum_{i=i}^{p} \varphi_i B^i\right)(1-B)^d X_t = \left(1 + \sum_{i=i}^{q} \theta_i B^i\right)\varepsilon_t \qquad (23)$$

Where B is the lag operator defined before.

The MAE (mean absolute error) statistic is defined according to the following expression:

$$\mathrm{MAE} = \frac{1}{N} \sum_{i=1}^{N} |e_i| \qquad (24)$$

Where e_i is the error, i.e., the difference between the size of the i-th frame, calculated by the mathematical model, and the size of the corresponding frame of the original video; and N is the number of frames of the video trace.

The MARE (mean absolute relative error) statistic is given by:

$$MARE = \frac{1}{N}\sum_{i=1}^{N}\frac{|e_i|}{x_i} \qquad (25)$$

Where x_i is the size of the i-th frame.

NMSE (normalized mean square error) can be described by:

$$NMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(e_i)^2} \qquad (26)$$

And the Akaike's Information Criterion (AIC) is a complex statistic that scores different models with respect to their ability to fit a given data series, defined by (see, [4]):

$$\mathrm{AIC} = 2\mathrm{k} - 2\ln(\mathrm{L}) \qquad (27)$$

Where k is the number of the parameters of the statistical model, and L is the maximized value of the likelihood function for the estimated model (and note that having a lower AIC statistic means that the corresponding model is more adequate).

In all these four parameters, the SAM model has better scores than the AR, and the ARIMA models. But these statistics does not inform us about the true ability of the SAM model to reproduce the video traces of the website.

Of course, the typical cumulative distribution functions (CDF) are also presented in references [3],[4],[5],[6],[8], and [9], in order to compare the distributions of the frame sizes generated by the SAM model with the data from the true video traces, and with those obtained by other mathematical models. And, the SAM model appears to be better than the alternative video models (see again, for instance, reference [9], for a comparison to the AR, the ARIMA models, and also to the true original trace). But the CDF statistics are less intuitive in order to show the statistical distributions of the frames sizes, and their match to the true video trace distributions.

Besides, the SAM model papers give the impression of having a contradiction, because sometimes the model can reproduce almost exactly a given video trace, but in other occasions it seems to fail that goal, without being explained why this happens. Examples of this aspect can be found in references [4], and [5]. In Fig. 7 of the reference [4] it appears a comparison between the trace of the true Star Wars IV video, and the corresponding trace generated by the SAM model, and it seems that the two traces match very closely. But in Fig. 8 a comparison is made between the trace of the Matrix 3 video, and the respective SAM model trace, with the purpose of testing the random shocks introduced in the SAM model in order to emulate the generation of sudden high size frames, a drawback of the model recognized by its authors [4]. But the frame sizes generated by the SAM model are much different, generally lower, than the originals, with smaller deviations from the average value, and with very random frame sizes. In my opinion, is hard to believe that a random process like the SAM model could match closely any video trace, but perhaps it is just my fault.

7 Third note: the SAM model does not reproduce adequately most video traces, namely the more structured

Our simulations have also shown that the SAM model, even with the use of expression (16), behaves poorly with respect to reproducing the majority of the website videos. In particular, those presenting some kind of structures like steps, and impulse trains. The SAM model has a tendency to generate traces with more randomness. These phenomena are considered to be due to its stochastic nature and its reduced set of parameters. In order to illustrate these points, the video traces 39, and 48 are depicted in Fig. 4, through 7. Video 39 has a prolonged step (Fig. 4) that the SAM model can't reproduce (Fig. 5), but the remaining part of the trace, after the step, has similarities to the original video trace. In Fig. 6, there are some steps that the SAM model can't reproduce, but the impulse train is more or less well defined (Fig. 7), although has higher values in the basis of the impulse train. Sometimes the SAM model also generates traces presenting an oscillatory behavior, or an unstable condition, as will be seen later. But there are also better behaved video traces, such as those obtained for video traces number 10, and 16.

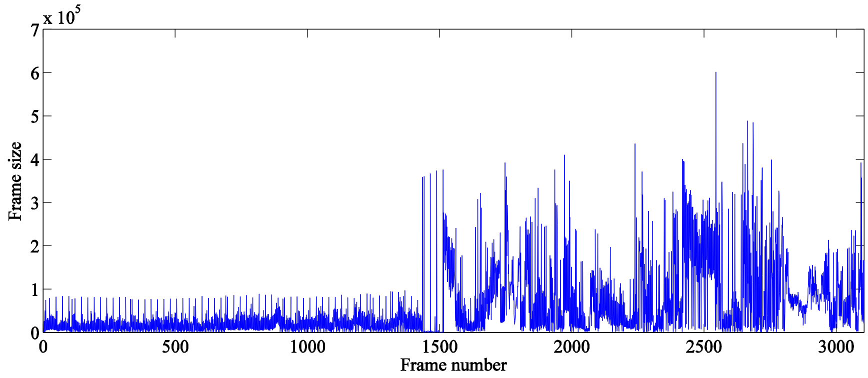

Fig.4. Original trace of video 39.

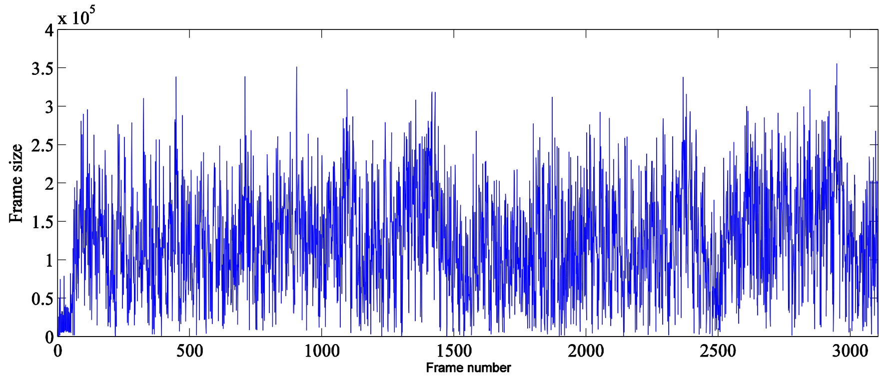

Fig.5. One trace of video 39 obtained using expression (16).

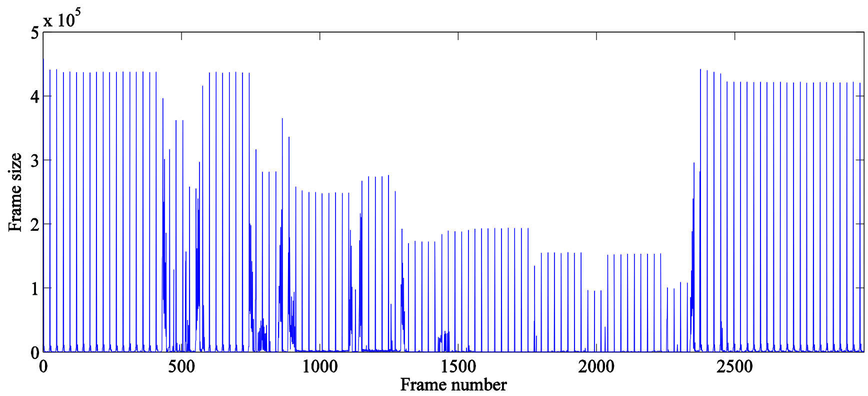

Fig.6. Original trace of video 48.

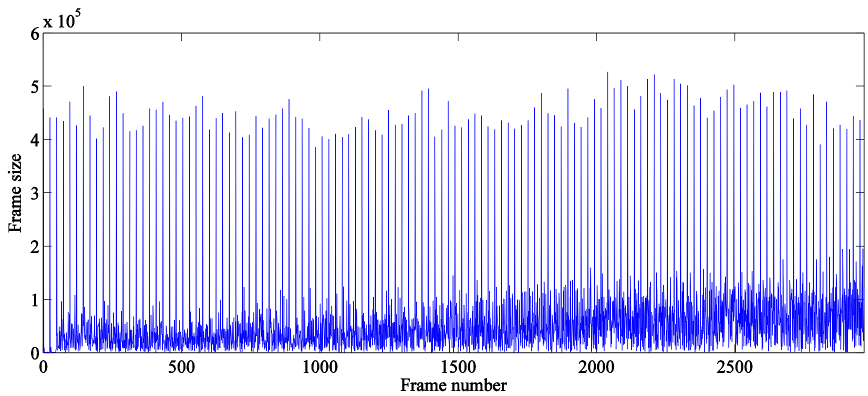

Fig.7. One trace of video 48 obtained using expression (16).

In Fig. 8, and 9, are shown the original trace and the simulated trace respectively of the video trace number 10.

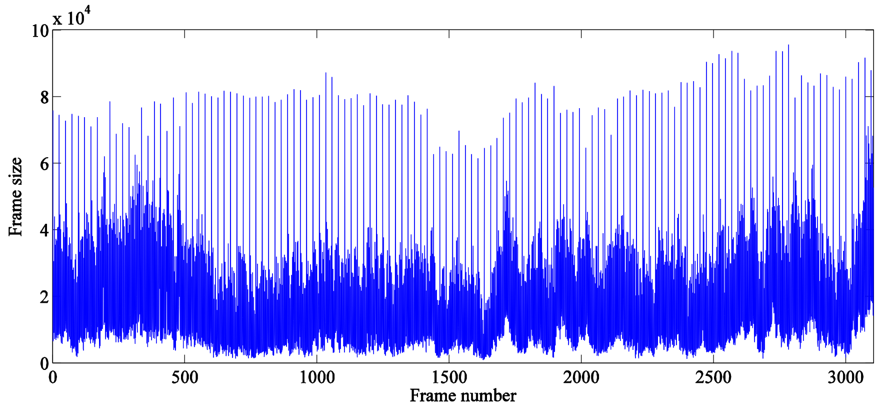

Fig.8. Original trace of video 10.

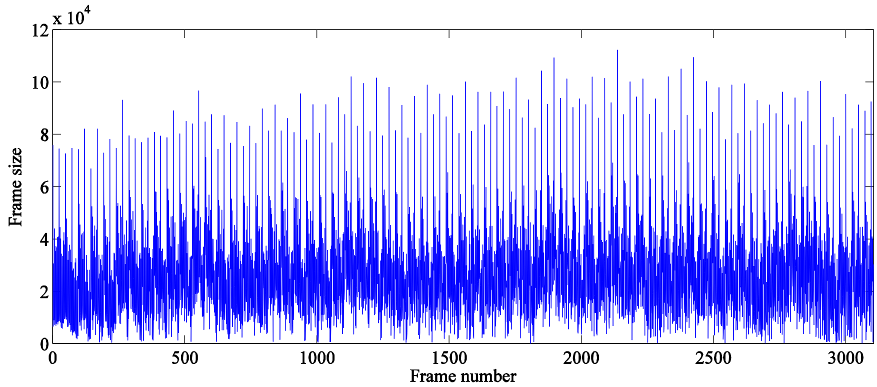

Fig.9. One trace of video 10 obtained using expression (16).

As it can be seen, there is a relatively good match between the original trace and the simulated trace. These few matches come also in support to the argument that expression (16) is the right expression of the SAM model.

In order to systematize all these points, Table II shows the behavior of the SAM model, when the expression (16) is used to generate the simulated video traces, and these traces are compared qualitatively to the originals.

In my opinion, Table I, and Table II are sufficient to show also that the SAM model is not sufficiently accurate to reproduce adequately most of the website video traces.

I have exchanged e-mails with all the three SAM model authors, because some aspects of the SAM model seemed to be less clear, but I have never received an answer to the specific query on expressions (1), and (16).

I think that the questions that were addressed in this paper deserve the attention, and the answer from the SAM model creators, and the attention of the scientific community.

Video	Qualitative observations
1	Some oscillatory behavior.
2	More randomness. Rising values/unstable expression?
3	More randomness.
4	Rising values/unstable expression?
5	Fails to reproduce isolated impulses. More randomness.
6	More randomness. Rising values/unstable expression?
7	Fails to reproduce two steps. Well defined impulses.
8	Same randomness (but lower values).
9	Fails to reproduce two steps and the impulse trains. Rising values/unstable expression?
10	Good match.
11	Fails to reproduce steps and the impulse trains.
12	More randomness.
13	More randomness.
14	Near good match. More randomness.
15	Some oscillatory behavior.
16	More randomness.
17	Fails to reproduce steps. Rising values/unstable expression?
18	More randomness.
19	Fails to reproduce steps. More randomness.
20	Fails to reproduce steps. More randomness.
21	Some oscillatory behavior.
22	Same randomness.
23	Same randomness.
24	Same randomness. Rising values/unstable expression?
25	Fails to reproduce a step. But well defined impulses.
26	More randomness. Not well defined impulses.
27	More randomness.
28	More randomness. Fails to reproduce steps and impulse trains.
29	More randomness.
30	Some oscillatory behavior. Fails to reproduce structures.
31	Presents impulses not present in the original trace.
32	More randomness. Rising values/unstable expression?
33	Some oscillatory behavior.
34	Some oscillatory behavior. Fails to reproduce structures.
35	More randomness.
36	More oscillatory behavior.
37	More oscillatory behavior. Fails to reproduce structures.
38	More randomness. Fails to reproduce structures.
39	Fails to reproduce a wide step.
40	More randomness.
41	Some oscillatory behavior. Fails to reproduce structures.
42	Impulse train well defined.
43	Fails to reproduce steps. More randomness.
44	Fails to reproduce structures. More randomness.
45	Some oscillatory behavior. Fails to reproduce structures.
46	More randomness.
47	Some oscillatory behavior and also more randomness.
48	Fails to reproduce steps. But well defined impulse train, with higher sizes in the basis.
49	Rising values/unstable expression?
50	Some oscillatory behavior and also more randomness.
51	More randomness. Fails to reproduce structures.
52	More randomness. Fails to reproduce structures.
53	More randomness. Fails to reproduce structures.
54	More randomness. Fails to reproduce structures.

Table II. Qualitative observations made when the simulated video traces are compared to the originals. The SAM model used expression (16).

Conclusions

The SAM model is a simple mathematical framework developed to capture and to reproduce video traces which follow the different MPEG-4 Part 2, Advanced Video Coding, and Scalable Video Coding techniques. However, the SAM model authors derived a mathematical expression that is wrong, as we have shown. That expression perhaps has not ever been used by its authors, as more sophisticated software was set for the simulations. A new mathematical expression was derived, and the results obtained by simulations support the idea that it may be the right SAM model equation.

The adequacy of this new expression was assessed through an exhaustive simulation campaign, and the results obtained, in the forms of both quantitative data and qualitative observations, show also that the SAM model is not sufficiently accurate to reproduce satisfactorily most of the video traces of the SAM model website, even if it produces better results than other mathematical models, as it is shown by its authors.

The traces produced by the SAM model usually present the following different behaviors when they are compared with the original traces: difficulty of reproducing more structured video traces, such as presenting steps, and impulse trains; traces with more randomness; traces with some oscillatory behavior, or even showing steadily increasing values, possibly as result of an unstable equation.

Acknowledgments: this work was partially supported by the FCT (Fundação para a Ciência e Tecnologia) under the project MPSat, through grant PTDC/EEA-TEL/099074/2008.

References

[1] G. E. P. Box, G. M. Jenkins, G. C. Reinsel, "Time Series Analysis. Forecasting and Control", Third Edition, Prentice-Hall International, Inc., 1994.

[2] C. Chatfield, "The Analysis of Time Series. An Introduction", Sixth Edition, Chapman & Hall, CRC, 2004.

[3] A. K. A. Tamimi, R. Jain, C. So-In, "SAM: A Simplified Seasonal ARIMA Model for Mobile Video over Wireless Broadband Networks", Proceedings of the IEEE International Symposium on Multimedia (ISM2008), Berkeley, California, December 2008.

[4] A. K. A. Tamimi, R. Jain, C. So-In, "Modeling and Generation of AVC and SVC-TS Mobile Video Traces for Broadband Access Networks", Proceedings of the First Annual ACM SIGMM Conference on Multimedia Systems (MMSys'10), Phoenix, Arizona, USA, February 2010.

[5] A. K. A. Tamimi, C. So-In, R. Jain, "Modeling and Resource Allocation for Mobile Video over WiMAX Broadband Wireless Networks", IEEE Journal on Selected Areas in Communications, Volume 28, Issue 3, pp. 354-365, April 2010.

[6] A. K. A. Tamimi, C. So-In, R. Jain, "Modeling and Prediction of High Definition Video Traffic: A Real-World Case Study", Proceedings of the Second International Conference on Advances in Multimedia (MMEDIA), Athens, Greece, June 2010.

[7] A. K. A. Tamimi, R. Jain, C. So-In, "Dynamic Resource Allocation Based on Online Traffic Prediction for Video Streams", Proceedings of the IEEE 4th International Conference on Internet Multimedia Services Architecture and Application (IMSAA), Bangalore, India, December 2010.

[8] A. K. A. Tamimi, "Modeling and Dynamic Resource Allocation for High Definition and Mobile Video Streams", PhD dissertation, Washington University in St. Louis, August 2010, Report 2011-95.

[9] A. K. A. Tamimi, R. Jain, C. So-In, "High Definition Video Streams Analysis, Modeling and Prediction", Advances in Multimedia, Hindawi Publishing Corporation, (accepted for publication).

[10] R Project Website: http://www.r-project.org/.

[11] WUSTL High-Definition Video Trace Library Website: http://www1.cse.wustl.edu/~jain/sam/index.html.

[12] A Mersenne Twister website: http://www.math.sci.hiroshima-u.ac.jp/~m-mat/MT/VERSIONS/C-LANG/mt19937-64.c.

[13] SAS: Business Analytics and Business Intelligence Software, http://www.sas.com, July 2009.