# OSCILLATION-SCALED HISTOGRAM-BASED MARKOV MODELLING OF VIDEO CODECS FOR MULTIPLE ACCESS

*M. DEL BUONO, L. HANZO, P. CHERRIMAN*

Dept. of Electr. and Comp. Sc., Univ. of Southampton, SO17 1BJ, UK.
Tel: +44-703-593 125, Fax: +44-703-593 045
Email: lh@ecs.soton.ac.uk; http://www-mobile.ecs.soton.ac.uk

## ABSTRACT

**A model is proposed for replacing the video codec in networking studies, which is tested for sources with a mean bitrate in the range of 10Kb/s to 10Mb/s. The standard frame sizes investigated are the International Telecommunications Union's (ITU) 352 x 352-pixel Common Intermediate Format (CIF), 176 x 144-pixel Quarter CIF (QCIF), 704 x 576-pixel 4CIF and 128 x 96-pixel Sub-QCIF (SQCIF). The proposed 20-state 'oscillation-scaled' Markov model was found to represent the relevant video codec characteristics adequately.**

## 1. INTRODUCTION

In order to study the behaviour of various multiple access schemes in case of video traffic, a simple, but sufficiently accurate video source model is needed. Reference [1] provides a review of different models commonly used to simulate voice, data and video sources. In the majority of cases Markov models or their derivatives have been favoured for their simplicity. Other common models are the autoregressive models. A comparison beetween these and some other models can be found for example in Reference [2]. Heymann and Lakshman in Reference [3] employed discrete autoregressive (DAR) and Markov models, while Reference [4] has studied the problems associated with the bitrate fluctuation of a video source.

In Section 2 the Markov modelling of video sequences is discussed and the model limitations are highlighted. Section 3 proposes a range of practical imporvements to the basic Markov model, while the performance of the algorithm is characterised in Section 4.

## 2. MARKOV MODELLING OF VIDEO SOURCES

In the spirit of our previous discussions we opted for adopting a Markov modulated process, which can adequately model both the first and the second moment of the bitrate fluctuation of various sources. Furthermore, we found that it was possible to superimpose a number of Markov chains in order to account for particular features of the sequence,

GLOBECOM'98, SYDNEY, AUSTRALIA
ACKNOWLEDGEMENT: THE FINANCIAL SUPPORT OF THE EPSRC, UK IN THE FRAMEWORK OF THE RESEARCH CONTRACT GR/K 74043 AND THAT OF THE EC UNDER THE AUSPICES OF THE EUROPEAN MEDIAN PROJECT IS GRATEFULLY ACKNOWLEDGED.

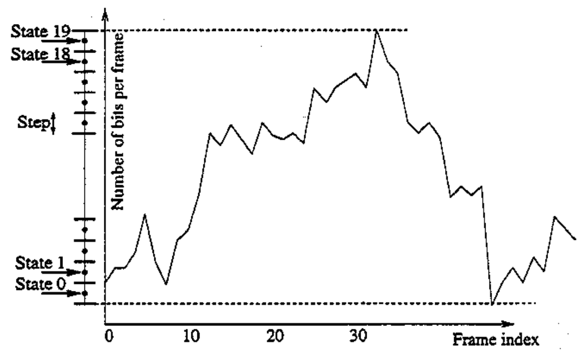


Figure 1: Quantization of the video source-rate fluctuation using a 20 state Markov model.

such as for example spikes in the bitrate histogram and other bitrate irreqularities, as it will be discussed below.

In a Markov Modulated Poisson Process (MMPP) [1] the instantaneous 'arrival rate' of transmitted packets or so-called traffic cells, ie the bitrate generated is 'modulated' by the state of a continuos-time discrete state Markov chain, which will be made explicit during our further discourse. This process is characterised by the arrival rate $\lambda_i$ per each state and the mean sojourn time in each state $1/r_i$. The sojourn time has a negative exponential distribution. The arrival rate $\lambda_i$ simply corresponds to the mean bitrate in state $i$, while the state transition probabilities are denoted by $P_{ij}$.

In video source modelling the first problem is the choice of the number of states in the Markov-chain. In order to match the bitrate histogram of the original source by that of the model sufficiently accurately, a high number of states is required. However, upon increasing the number of states we found many more bitrate histogram spikes in the simulated sequence than there were in the original. This indicates that a high number of Markov states requires very long training sequences for generating an accurate state transition matrix in order to arrive at a statistically meaningful number of state transitions amongst all possible states. This issue will be revisited during our further discussions, but suffice to say here that when experimenting with limited-duration practical video sequences, it was impractical to choose a very high number of states, since then the statistical credibility of the investigations became questionable due to the associated low number of state transitions amongst certain low-probability states. We addressed this problem by con-

0-7803-4984-9/98/$10.00 

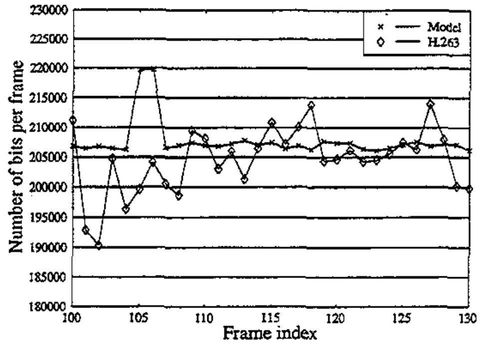


Figure 2: Modelling the H.263-encoded 4CIF Susie sequence at a target bitrate of 7Mbps. For many frames the model resides in the same state resulting in a constant bit rate.

structing mosaic-sequences, constituted for example by four different-length sequences combined to form a quadruple-sized sequence, which will be invoked in Section 4 for algorithmic performance testing. In order to find a good compromise between these two conflicting requirements after a range of experiments we opted for using a 20-state Markov model for the ITU's H.263 codec.

According to the above considerations we investigated a variety of different resolution sequences to be modelled and after identifying the maximum and minimum bitrate, ie the bitrate range of the sequences, we divided this range in 20 uniform bitrate ranges. At the center of each bitrate interval we allocated a state of the Markov chain, as seen in Figure 1. The transition probability from state $i$ to state $j$ has been found by simulation upon observing the sequence after assigning the actual measured bitrate to one of the 20 states.

Our tentative bit-generation model has the following construction. In each bit-generation cycle a random generator is used to determine the next state of the Markov model, which can be any of the 20 states. These transitions are governed by the transition matrix, generated by evaluating the relative frequencies, approximating the probabilities of all possible Markov-state transitions using simulations. Then in each state the actual number of bits generated obeys the Poisson distribution and the corresponding probability density function (PDF) typically overlaps with those of the adjacent states. Having stipulated the basic video model, let us now scrutinize its behaviour in the next Section.

## 3. REDUCED-LENGTH POISSON CYCLES

We note that from a practical point of view operating directly with the number of bits per video frame is inconvenient, since observing the Poisson distribution of

$$P(n) = \frac{(\lambda_i T)^n \exp(-\lambda_i T)}{n!} \qquad (1)$$

we found that the factorial function of the denominator results in an excessive computational demand. Hence the solution is to divide the video frame duration into a number of bit-generation cycles with the advantage that in this way

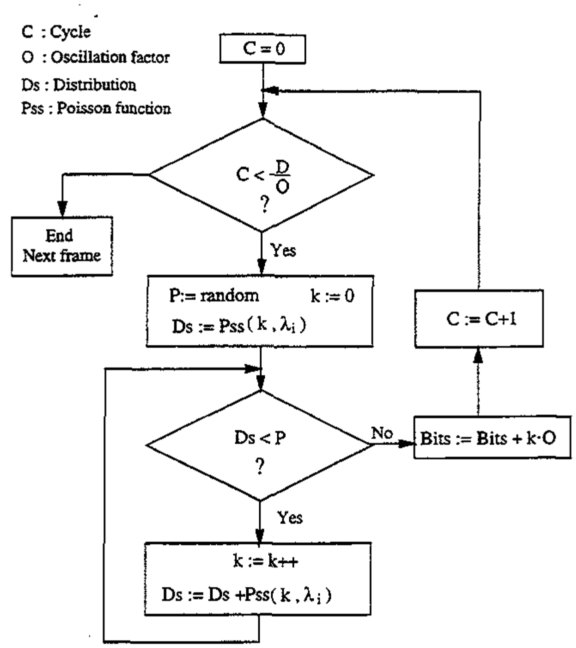


Figure 3: Flow chart of the bit generation algorithm for a video frame

we are able to find the number of bits generated on a more convenient scale, using a granularity more compatible with the typical burst-length of conventional wireless networks.

Hence we decided to divide the video frame in a number of shorter segments in order to reduce the number of bits generated per Poisson cycle to around 300 or less, simply because in case of a higher average number of bits per generation cycle the probability that we must calculate a factorial higher than 1000! is not negligible. Therefore we opted for invoking a division factor of $D = 5000$, which is used to divide the target bit number per video frame in $D = 5000$ smaller bit generation cycles. This choice constitutes a good compromise for source rates from 10Kbps to 10Mbps. For sources at higher bit rates we have to increase the value of $D$.

A consequence is that now the number of bits generated per video frame is the accumulation of the number of bits generated per Poissonian cycle. This means that the distribution is now the convolution of 5000 Poisson distributions. We observed that in this case the number of bits per frame was not sufficiently spread around the average in order to provide a statistically sound model of the bitrate fluctuation for the H.263 codec. This is demonstrated in Figure 2, in comparison to the actual number of bits generated by the H.263 codec for the 4CIF 'Susie' sequence coded with a target bitrate of 7Mbps. Observe that apart from a single excursion to a state corresponding to about 220 000 bits per frame, which occurs at frame index 105, the process resides in a state emitting a Poissonian-rate around 207 000 bits per frame. However, these rate fluctuations appear quite limited.

In order to avoid this near-constant bitrate problem we introduce an 'oscillation factor' $O$, the role of which and the terminology becomes explicit below. The effective number

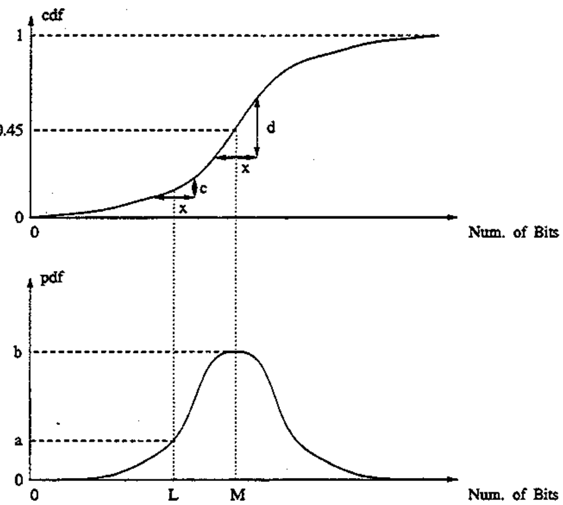


Figure 4: Stylised PDF and CDF

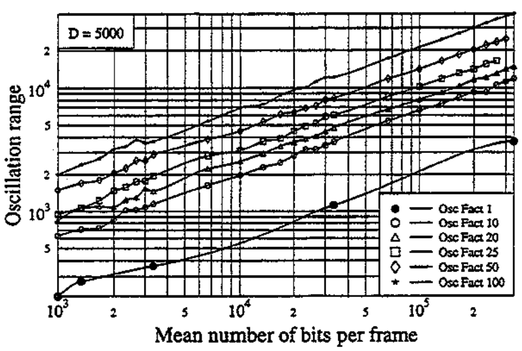


Figure 5: Oscillation range $\Delta R$ versus mean number of bits per frame $R_a$: different ranges can be selected using different $O$ values in order to ensure a better fit of the bitrate histograms of the model to that of the codec modelled.

of bit generation cycles is now computed as $D/O$ but at the end of each modelling cycle we multiply the number of bits generated by the value of $O$. This measure allows us to maintain a better bitrate granularity as it will be demonstrated below. The operation of the model is illustrated in the flow-chart of Figure 3.

Focusing our attention on this flow-chart, let us initially assume an oscillation factor of $O = 1$. Then there are $D = 5000$ Poissonian generation cycles and as long as the cycle index $C$ is lower than $D$, further generation cycles are required for the current video frame. At the beginning a random number is generated and assigned to $P$, where $0 < P < 1$. Then the Poisson PDF is evaluated for the iteration index of $k = 0$ and the returned Poissonian value is assigned to the distribution function $Ds$. If we have $Ds < P$, then the iteration index $k$ is incremented and $Ds$ is updated by adding the Poissonian variable generated using the incremented value of $k$ ie $k = 1$. When

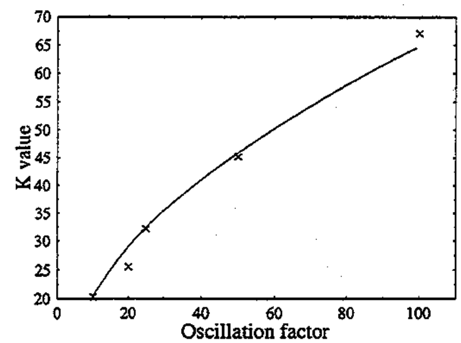


Figure 6: The value of K for different oscillation factors $O$, where the curve was found by minimum mean-squared fitting

$Ds$ reaches the random value $P$, the process is stopped and the current value of the iteration index $k$ is multiplied by the oscillation factor $O$, giving the number of bits generated during the first generation cycle.

In order to study the effect of different oscillation factors $O$ the following experiment was carried out. Initially, $O = 1$ was stipulated, resulting in $D$=5000 Poissonian cycles per videoframe, yielding 5000 Poisson distributed numbers, generated using Equation 1, where $T = FrameDuration/D$ and $\lambda_i$ is the mean bitrate of the Markov chain in state $i$. These simulations were conducted for 3000 video frames. The lowest and the highest number of bits generated were recorded and we refer to their difference as the fluctuation range. This range was then recorded for various average bitrates or mean number of bits per frame. The results are plotted in Figure 5 for a range of $O$ values between 1 and 100.

Explicitly, the curve plotted for $O = 1$ is in fact the original curve, where we used $D$=5000 Poisson cycles per video frame. It is easy to observe that for an average source rate around 5 Mbps, corresponding to 167 000 bits per video frame at 30fps, a fluctuation of 3000 bits around the mean value is almost negligible.

Observe furthermore in Figure 5 that due to the introduction of the oscillation factor $O$ for a 5Mbps, 30 frames/s, 167 000 bits per frame scenario, the fluctuation range now becomes significantly higher, approximately 30 000 bits around the mean value. This is because for $D = 5000$ and $O = 1$ the number of bits per frame was the cumulative value of 5000 Poissonian variables, yielding a near constant value. By contrast, for $O = 100$ the higher oscillation range is a consequence of accumulating only 50 such variables.

At this stage a further step was required in order to complete the model design. Specifically, given an overall average bit rate of $R_a$, a minimum and maximum bitrate of $R_{min}$ and $R_{max}$, respectively, as well as a set of $N$ Markov-model states, the resulting target bitrate $R_i$ (or $\lambda_i$ of state $i$ with $0 \leq i \leq N$ is given by:

$$R_i = R_{min} + i\,\frac{R_{max} - R_{min}}{N} = R_{min} + i\,\Delta R. \qquad (2)$$

While residing in any of the Markov-states, the model will ensure that the range of the instantaneous bitrate-fluctuations

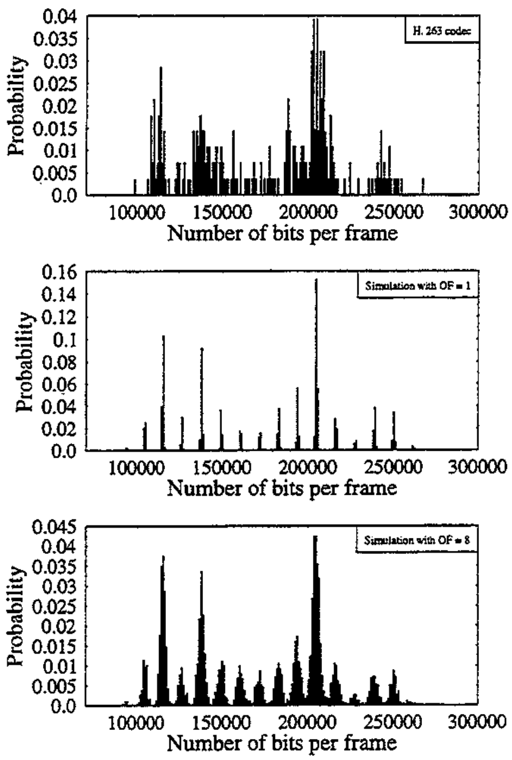


Figure 7: Bitrate histograms for the 4CIF Susie sequence at a 5Mbps target bitrate generated by the H.263 codec and by the proposed model with $O = 1$ and $O = 8$. Observe the better histogram fit due to a higher value of the $O$.

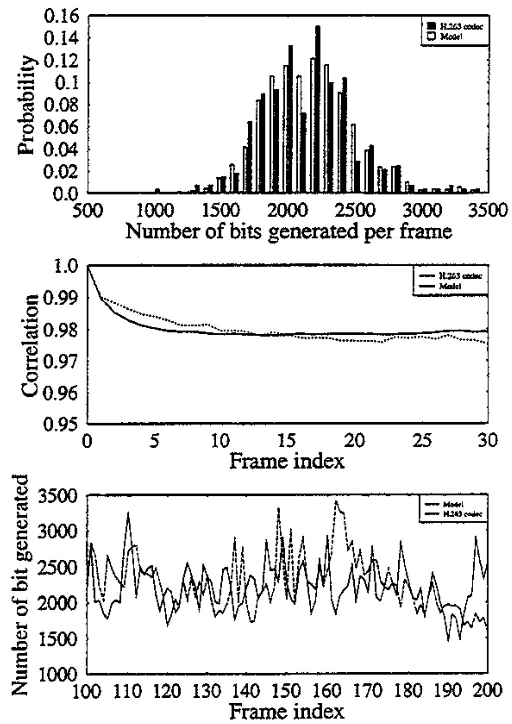


Figure 8: Bitrate histogram, correlation and typical bit rate for the Carphone sequence. Comparison between the H.263 codec at QCIF size , 64kbps target bitrate and the Markov model

is limited to $\Delta R = [R_{min}; R_{max}]$, and the specific bitrate values of each state associated with a certain mean bitrate obey the Poissonian distribution. Given the oscillation range $\Delta R$, we can invoke Figure 5 in order to determine the required oscillation factor $O$.

In order to assist in this, we found an empirical relation between the quantities involved. From Figure 5 we inferred that the relationship between the average value of bits per frame $R_a$ and the oscillation range $\Delta R$ is given by:

$$\Delta R = K\sqrt{R_a}. \tag{3}$$

Minimum mean-squared fitting of the experimental $R_a$ and $\Delta R$ values for various $K$ values revealed the following dependence of $K$ on the oscillation factor $O$:

$$K = a \cdot \sqrt{O} \tag{4}$$

where we have $a = 6.48$. The goodness-of-fit of this matching process is characterised in informal terms by Figure 6.

The above mentioned experimental relationship has been used in our simulations and the corresponding bitrate histograms are depicted in Figure 7 for two different $O$ factors, namely for $O = 1$ and 8, as well as for our experimental data generated by the H.263 codec for the 4CIF 'Susie' sequence, while maintaining an average bitrate of 5Mbps. Observe in the Sub-Figure in the middle that in accordance with our previous experience, for $O = 1$ there is only a very limited bitrate fluctuation or spread within the Markov-states around the target bitrates of the individual states. However, for $O = 8$ a more appropriataly spread Poissonian bitrate distribution is observed in each state. Observe for $O = 8$ at the bottom of the Figure that the PDFs are slightly more spread towards the top end of the bitrate range than in the lower-rate Markov states. This is because for the Poisson distribution the value of the variance is equal to the mean value, which is clearly higher for the the states closer to the top end of the bitrate scale.

## 4. SIMULATION RESULTS

From our simulation results we found that for a source bitrate around 1Mb/s or less an oscillation factor between $O = 1$ and 3 was appropriate. For source rates around 10 Mb/s a value around $O = 50$ was required, depending on the target source rate. Furthermore, we found that the model was quite flexible and allowed us to emulate a range of different video scenes adequately.

Figures 8-10 show a number of model characteristics for various video sequences. Specifically, at the bottom of each of these Figures the typical bitrate fluctuation of the original H.263 codec and that of the model can be seen, as an easily interpreted illustrative example. In the centre of each

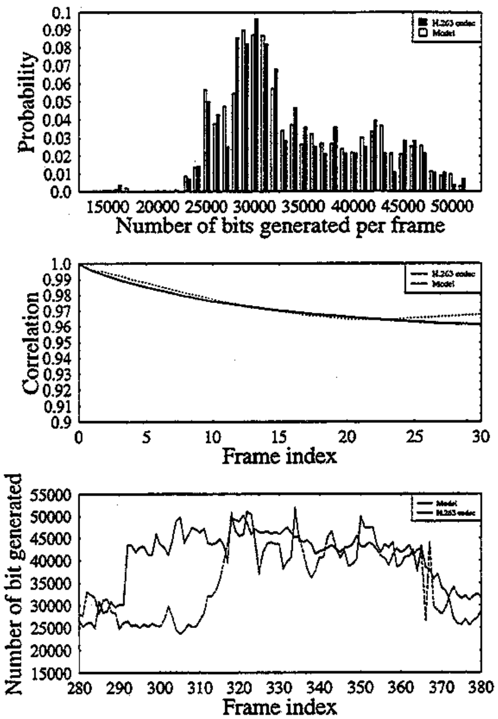


Figure 9: Bitrate histogram, correlation and typical bitrate for the Susie sequence. Comparison between the H.263 codec at 4CIF size, 1Mbps target bitrate and the Markov model

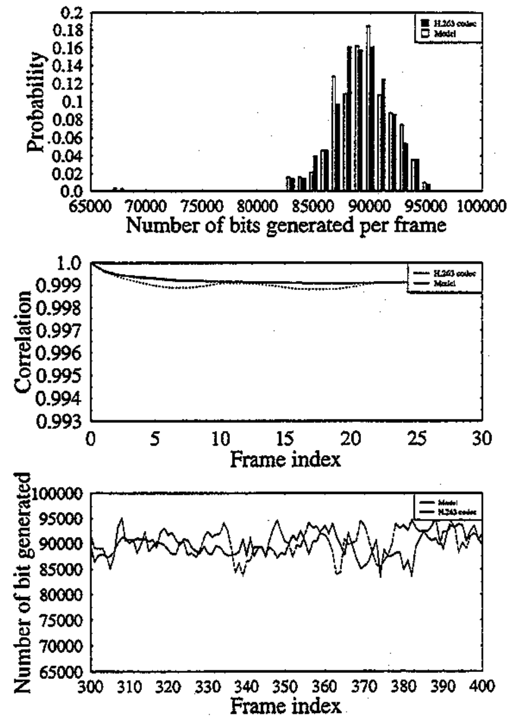


Figure 10: Bitrate histogram, correlation and typical bitrate for the Miss America sequence. Comparison between the H.263 codec at CIF size, 3Mbps target bitrate and the Markov model

of these illustrations the normalised correlation between the bitrates of consecutive frames was plotted, while at the top the bitrate histogram of both the original experimental data and that of the model is displayed.

A representative range of low, medium and high bitrate scenarios were studied using various video sequences encoded at various bitrates. Although the bitrate histograms would not be acceptable at high confidence-level using rigorous goodness-of-fit distribution testing techniques, for practical network modelling purposes they were deemed adequate. The bitrate correlation functions also exibited an adequate match. The observed deviations from experimental features were deemed to be a consequence of the limited-duration training data for the model, which adopted the transition matrix entries of the experimental data. When using these state-transition probabilities, extremely long training sequences and model verification experiments would be required for achieving a better statistical match. Figure 10 shows a very high correlation due to the particular features of the 'Miss America' sequence, which exhibits a rather limited motion activity and hence the number of bits generated per frame is almost constant without large excursions around the mean value.

## 5. CONCLUSIONS

Oscillation-scaled Markov models have been porposed for modelling various video sources for networking studies. A 20-state model was found to reproduce most model features for various video resolutions and frame rates quite accurately.

## 6. REFERENCES

[1] J. Cosmas, G. Petit, R. Lehenert, C. Blondia, K. Kontovassilis, O. Casals, T. Theimer. A review of voice, data and video traffic models for ATM *European Transactions on Telecommunications Vol. 5, No. 2*, pages 139–154, Mar-Apr 1994.

[2] Oliver Rose and Michael R. Frater. A comparison of models for VBR traffic sources in B-ISDN. *Broadband Communications, II*, pages 275–287, 1994.

[3] D. P. Heymann and T.V. Lakshman . Source models for VBR broadcast-video traffic. *IEEE/ACM Transactions on networking Vol 4*, pages 40–48, Feb 1996.

[4] S. S. Lam, S. Chow and D. K.Y.Yau. A lossless smoothing algorithm for compressed video . *IEEE/ACM Tr. on networking Vol. 4*, pages 697–708, Oct 1996.