Models and Tools for Simulation of Video Transmission on Wireless Networks

Jian Zhu Ashraf Matrawy Ioannis Lambadaris
Department of Systems and Computer Engineering, Carleton University
{jianzhu, amatrawy, ioannis}@sce.carleton.ca

Abstract

Robust transmission of video is a dominant requirement of future applications over wireless networks. MPEG4 is an object based video encoding technique which is suitable for wireless applications due to its high compression performance, scalable video coding techniques, error-resilient capability and object-based coding functionalities.

In this paper, we first model traces of MPEG4 traffic. Based on these models, we develop tools for MPEG4 traffic generation. These tools have an adaptive rate control function that is capable of simulating MPEG4's scalable video coding. These tools can be used as source traffic generator in network simulators. This enables the study of MPEG4 transmission performance over wireless networks by using simulation. We model and generate the traffic based on the Transform Expand Sample (TES) methodology.

In our experiment, we generate MPEG4 traffic and show the performance in terms of good matching of the characteristics of the modeled traffic.

Keywords: *Modeling, Simulation, Wireless video.*

1. INTRODUCTION

With the rapid development of mobile communication, robust transmission of video over wireless networks is becoming an increasingly important application requirement. In order to reduce the bandwidth requirements, the source video data is compressed before transmission. Between many video compression standards, MPEG-4 is most suitable for wireless links. It is targeted for low bit rates. Its object-based coding functionalities allow for user interaction with audio-visual objects. The main features of importance are MPEG-4's scalable coding techniques and error-resilient tools. These enhance network utilization and enable MPEG-4 senders to be more responsive to changes in wireless network conditions.

In this work, we first model MPEG-4 traffic traces. The models were developed by using Transform Expand Sample (TES) [1], which is a versatile methodology for modeling any set of given observations in a time series. We model I frames, P frames and B frames using three different TES models as in [2]. Then based on these models, we develop tools for MPEG-4 traffic generation. These tools have an adaptive rate control function that is capable of simulating MPEG-4 encoder's output bit-rate variation due to different quantization level or scalable video coding (except Fine Granularity Scalability). This function is of great relevance in studying video transmission in wireless networks where the network condition is continuously changing. We incorporate these tools into network simulator OPNET. This enable the study of MPEG-4 transmission performance over wireless networks by using simulation. The traffic we generate matches the statistical characteristics (in terms of marginal distribution and auto-correlation function) of an original real trace of video frames that were generated using an MPEG-4 encoder.

The rest of paper is organized as follows: section 2 gives an overview of the MPEG-4 encoding technique. In section 3, we present the Transform Expand Sample (TES) methodology. Section 4 gives the details of how we modeled and simulated MPEG-4 encoded traces using TES and demonstrates the adaptive rate control function. Section 5 concludes the paper.

2. MPEG-4 OVERVIEW

MPEG-4 is an ISO/IEC standard developed by Moving Picture Experts Group (MPEG). Although MPEG-4 is a video and audio compression standard, in this paper we only focus on video.

Designed as an adaptive representation scheme that also accommodates very low bit rate application, MPEG-4 is very appropriate for wireless multimedia applications due to its following features[3] [4]:

1) It is targeted to provide high video quality at relatively low bit rates. The lowest bit rate can be 5k bps.

2) Scalable video coding techniques are introduced into MPEG-4 to provide the varying coding bit rate for the varying channel capacity.

3) Many error-resilient tools are incorporated into MPEG-4, which guarantee the correctness of the data in the error-prone wireless channel.

CCECE 2004- CCGEI 2004, Niagara Falls, May/mai 2004
0-7803-8253-6/04/$17.00 ©2004 IEEE

4) Object-based coding functionalities allow for interaction with audio-visual objects and enable new interactive applications in a wireless environment.
5) Face animation parameters can be used to reduce bandwidth consumption for real-time communication applications in a wireless environment.

MPEG-4 encoders generate three types of frames: I frames, P frames and B frames. They can be described as follows:
1) I frames (intra-coded frames) contain the information that results from encoding a still image. They can be decoded without the need for any other frames.
2) P frames (predictively coded frames) require information from previous I frames and/or P frames for encoding and decoding.
3) B frames (bidirectionally predictively coded frames) require information from the previous and following I frames and/or P frames for encoding and decoding.

MPEG-4 scalable video encoder generates multiple sub-streams in contract to one compressed bit-stream of non-scalable video encoder. One of the compressed sub-streams from the scalable video encoder is the base layer sub-stream, which can be independently decoded and provides coarse visual quality. Other compressed sub-streams, which are called enhancement layer sub-streams, can only be decoded together with the base sub-stream and can provide better visual quality. For example, Temporal scalability is a scheme to compress the raw video data into tow layers at the same spatial resolution, but different frame rates. The base layer is coded at a lower frame rate. In contract, the enhancement layer compresses a video with a higher frame rate providing the missing frames. In the base layer only P-type prediction is used, while in the enhancement layer prediction can be either P-type or B-type from the base layer or P-type from the enhancement layer [4].

3. TRANSFORM EXPAND SAMPLE(TES) METHODOLOGY

TES is a versatile methodology for generating data that match closely (in terms of its marginal distribution and auto-correlation function) any set of given observation of time series. The derivation of TES models is performed in two phases. First a correlated sequence, with uniform marginal in [0,1], (also called background TES process) is formed as follows:

$$U_n^+ = \begin{cases} U_0 & n=0 \\ \langle U_{n-1}^+ + V_n \rangle & n>0 \end{cases}$$

or

$$U_n^- = \begin{cases} U_n^+ & n \text{ even} \\ \langle 1 - U_n^+ \rangle & n \text{ odd} \end{cases}$$

where U_0 is a random variable uniformly distributed in [0,1].{ V_n } is a sequence of iid random variables independent of U_0 , called the innovation sequence (with probability density f_v). The operator < > denotes the modulo-1 operation which for every real x is defined by

$$\langle x \rangle = x - \max\{\text{integer } (n): n \le x\}$$

Both U_n^+ and U_n^- can generate lag-l autocorrelations in the range [0,1] and [-1,0] respectively. Regardless of the probability law of the innovation density f_v, it can be shown [1] that the TES background sequence forms a stationary Markov process with a uniform in [0,1] marginal distribution. Since the < > operator causes discontinuities, a smoothing operation called the stitching transform is used to get rid of it. So in the first phase of the TES model development, the objective is to generate a time series of correlated random variables with uniform marginal in [0,1]. The amount of correlation depends upon the structure of the density function f_v of the innovation sequence { V_n }, which in general may be arbitrary. In the second phase, synthetic sample data called the foreground sequence, which resemble the real samples, may be derived from the background sequence using the inversion technique[1]. This allow us to transform any uniform random variable to one with an arbitrary distribution. More specifically, it is assumed that a histogram H(.) of the empirical video sequence has been built consisting of J cells, where the jth cell is positioned on the interval $[l_j, r_j)$ and is characterized by a probability p_j. Such a histogram can be inverted as follows:

$$H^{-1}(x) = \sum_{j=1}^{J} I_{[C_{j-1}, C_j)}(x) \left[l_j + (x - C_{j-1}) \frac{w_j}{p_j} \right]$$

where I_A is the indicator function of the set A, $w_j = r_j - l_j$ is the width of cell j, and $c_j = \sum_{i=1}^{j} p_i$, $1 \le j \le J$, (C_0 =0, C_J =1). The random variable $\{H^{-1}(x)\}_{n=1}^{\infty}$ has a marginal distribution H(.) and autocorrelation that depends upon the innovation density f_v.

The modeling process is carried out in two steps: the first step essentially captures the autocorrelation of the input time series using innovation density function f_v and applying a smoothing technique using the stitching transform [1]. This would result in the generation of the

background sequence. The second step is to invert the background sequence using the empirical histogram, which always guarantees the matching of the marginal.
The task of finding a suitable f_v that approximates the empirical autocorrelation is carried out via a heuristic search using TES tool [1]. The matching of the empirical histogram is guaranteed, while the matching of the autocorrelation is carried out in an interactive manner with visual feedback for the best f_v.

4. Modeling and Simulation

Traces of MPEG-4 encoded film "STAR WAR IV" [5] are used to present our models and tools for video traffic generation. Due to the periodic nature of the autocorrelation function of the original trace, we first need to separate it into three traces, one for each frame type [2]. This kind of separation also facilitates the development of adaptive rate control function on our traffic generation tools, which we will introduce later . Having done that we model each of the traces using a separate TES model using the methodology described in previous section. Fig 1 illustrated the models of I frames. To generate the MPEG-4 trace for simulation, we interleave generating samples from the three developed models to match the sequence of the original trace (IBBPBBPBBPBBI...)[2]. The final trace generated by interleaving can be compared to the original trace as shown in Fig. 2. Based on the developed models, we also develop adaptive rate control function on our traffic generation tools.

4.1 Variable Quantizer Rate Control

The output bit-rate of MPEG-4 encoder can be controlled by applying different quantization table during encoding. Such control, which could be used to respond to congestion conditions in the network, is highly desirable for network performance improvement. In order to study the effects of such a control scheme, a model of the variable quantizer rate controlled MPEG-4 encoder is required.

Fig 3. illustrate the difference when applying different quantization table on I frames of film "STAR WAR IV". Through observation, we noticed that the shape of the sample path is preserved, but shifted. While the autocorrelation is remain unchanged. The effect on the sample path of changing the quantization number is equivalent to multiplying the size of each frame by a scalar value, which also have been shown in H.261 encoded VBR video [6]. This makes it possible to infer frame size for a wide range of quantization numbers from a single table model. The implication is that a model need only be developed for one of the possible quantization tables. Specifying a different quantization table to the encoder is simulated in the source model by applying a scalar multiple to each value generated. Based on this idea, our traffic generation tools simulate the variable quantizer rate controlled MPEG-4 encoder by using basic models and certain numbers of scalars.

To demonstrate the usefulness of this approach, we developed a model of same sequence at Q=10 by applying a scaling factor to the previously developed I frame model encoded at Q=4. Comparison between the developed model and the original trace of I frame at Q=10 indicated good agreement, as illustrated in Fig 4.

4.2 Variable Layer Rate Control

Multiple layers of streams are generated in MPEG-4 scalable video encoder to accommodate varying channel capacity in wireless applications. Base layer sub-stream can be independently decoded and provides coarse visual quality. Enhancement layer sub-streams can only be decoded together with the base layer sub-stream and provide better visual quality. The scalable video encoder can adjust its output bit rate by sending base layer sub-stream only or base layer sub-stream together with enhancement layer sub-streams according to the feedback of network conditions. To simulate this kind of rate control, we model base layer sub-stream and enhancement layer sub-stream separately according to the TES methodology. When generating the MPEG-4 trace for simulation, the feedback of the network conditions is used to decide whether interleave or not the enhancement layer sub-streams.

Simulation of a type of temporal scalable encoding is introduced here as an example. In this considered temporal scalable encoding type the I and P frames constitute the base layer while the B frames constitute the enhancement layer [7]. We first develop three models each for one frame type. Then by control of combining B frames traffic or not during generation of simulated MPEG-4 traffic, we achieve the variable layer rate control.

5. CONCLUSIONS

This paper presented our models and tools for generating video traffic which could be used for simulation of video transmission on wireless networks. These tools is capable of simulating MPEG-4 encoder's output bit-rate variation due to different quantization level or scalable video coding (except Fine Granularity Scalability). After incorporated into network simulator, this traffic could be used as source traffic in our future study of scheduling over wireless ring networks.

References

[1] Melamed, B. et al., " TES-based traffic modeling for performance evaluation of integrated networks" IEEE INFOCOM '92. , Pages:75 - 84 vol.1,4-8 May 1992

[2] Matrawy, A et al., "MPEG4 traffic modeling using the transform expand sample methodology" Networked Appliances, 2002 IEEE 4th International Workshop on , Gaithersburg. Pages:249 – 256

[3] Rob Koenen, ed. "Overview of the MPEG-4 standard", ISO/IEC JTC1/SC29/WG11 M4030, March 2001

[4] Bo Yan; Ng, K.W.; "A survey on the techniques for the transport of MPEG-4 video over wireless networks" Consumer Electronics, IEEE Transactions on, Volume:48, Issue:4, Nov 2002 ,Pages: 863 – 873

[5] http://trace.eas.asu.edu/

[6] C.M. Sharon et al., "Accurate Modeling of H.261 VBR Video Sources for Packet Transmission Studies", in Proceedings of *IEEE GLOBECOM '95*, Singapore, November 13-17, 1995.

[7] Martin Reisslein et al., "Traffic and quality characterization of scalable encoded video: a large-scale trace-based study" technical report, Dec 2002

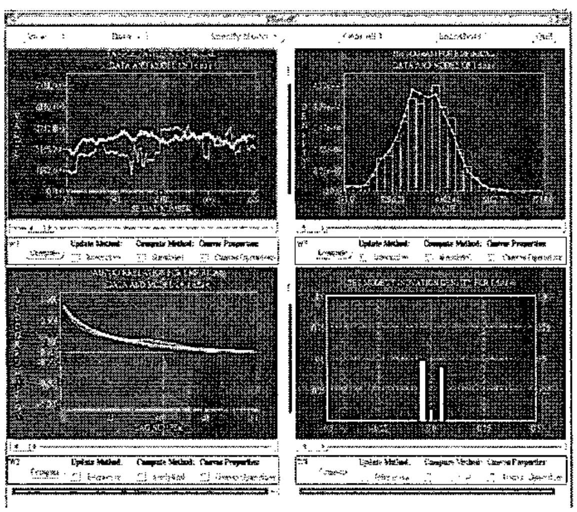

Fig 1. The I frames model

Window 1 (left upper window) shows the sample path of the frame sizes.

Window 2 (left bottom window) shows the autocorrelation of the frame sizes.

Window 3 (right upper widow) shows the histogram of the distribution of the frame sizes.

Window 4 (right bottom window) shows the innovation density f_v that was described in section 3.

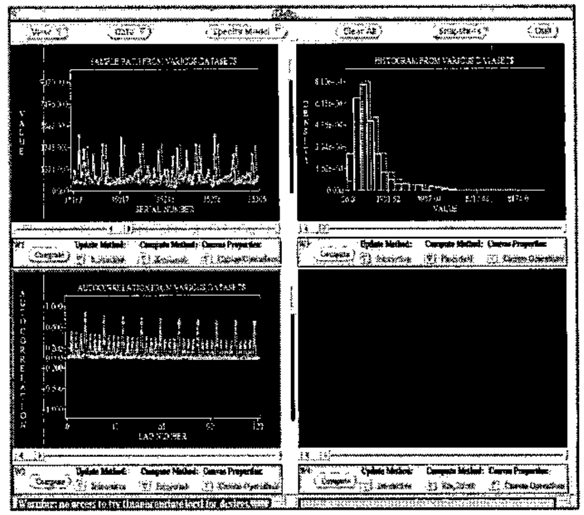

Fig 2. Comparison between the original and the generated traces

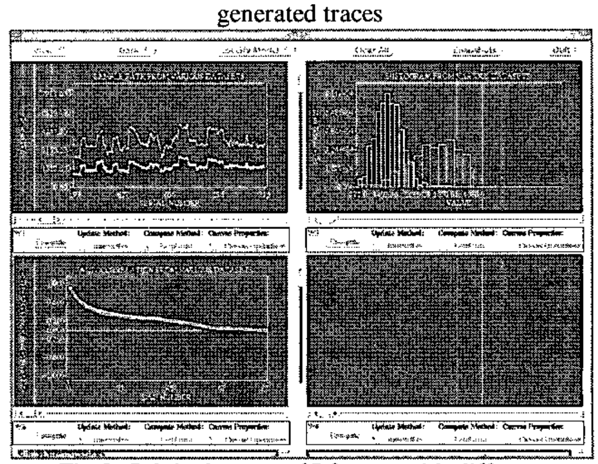

Fig 3. Original traces of I frames with different quantization table

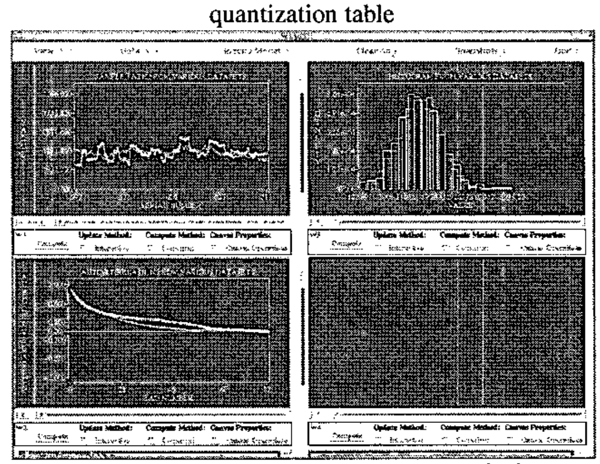

Fig 4. Comparison between original trace of I frame (Q=10) and trace modeled from Q=4.