NEW PERSPECTIVES FOR DIGITAL TELEVISION

M. T. Andrade and *A. P. Alves*

INESC Porto, Praça da República, 93 R/C, 4050-497 Porto, Portugal
FEUP, DEEC, Rua Dr. Roberto Frias s/n, 4200-465 Porto, Portugal
telephone: +351 22 2094237 fax number: +351 22 2084172
email: {mandrade, palves}@inescporto.pt
http://telecom.inescn.pt

Abstract

The ability to increase the number of programmes offered to users while simultaneously introducing new interactive features, providing the mechanisms for service customisation and increasing the quality is a requirement of digital television systems. With the convergence of television and internet and the inevitable mixture of types of content traditionally kept separated, it is necessary to devise techniques to deliver high-quality services together with new services typical of the IP world. Statistical multiplexing schemes and VBR video encoding will be important in this process to achieve bandwidth savings thus freeing up resources for new additional multimedia services without compromising the quality of programmes. In this paper we present a framework for the study of statistical multiplexing techniques for digital TV, based on the use of probabilistic modeling of constant quality video sources. This framework explores inference techniques together with content characterization to test statistical multiplexing algorithms. *Keywords: statistical multiplexing, VBR, inference, quality*

1 Introduction

With the introduction of digital services and the widespread convergence between the television and the internet worlds, the TV customer of the near-future will no longer be satisfied with simply increased choice or with the same type of programmes wrapped up with some form of interactivity. TV programme providers will have to be able to deliver not only new interactive multimedia content and access to new services similar to those available with the internet, but also deliver increased picture quality. By doing this it will be possible to accomodate users that expect that digital TV will mean the access to clear crystal pictures and others who mainly see in the transition to digital, the opportunity to add interactivity and other internet-like services. TV services are currently using the MPEG2 standard. Most programmes are transmitted in constant bit rate (CBR) mode. This generally means variable picture quality as the MPEG video encoding scheme generates variable bit rate (VBR) streams when encoding constant-quality video sources. It is common practice to allocate fixed bit rates to each TV channel, sufficient to ensure acceptable picture quality for the majority of time. However this scheme is not efficient in terms of both bandwidth utilization and overall picture quality. Typically, several encoded TV programmes are multiplexed together and sent to different locations for distribution to the public. This operation can be performed using statistical multiplexing techniques and allowing the MPEG codecs to operate in VBR mode. Statistical multiplexing in TV distribution helps to maintain good picture quality, while freeing up bandwidth that can be made available for other services. This advantage has already been demonstrated in terrestrial and satellite TV systems. However, in addition to datacasting services it is now possible to provide additional video services using the excess bandwidth, through, for example, the use of the new video compression framework recently adopted in MPEG4, known as Fine Granular Scalability (FGS). The gain in bandwidth is particularly important for high-definition television. The large amounts of storage currently installed in the receiver, can contribute to the possibility of maintaining a seamless service.

2 System architecture

To be able to offer a constant picture quality service, it is advantageous to use the VBR encoding alternative. To be able to maximise the ratio $number-of-programmes/bandwidth$ it will be necessary to complement the use of VBR with statistical multiplexing techniques. In some cases, TV broadcasters are already using statistical multiplexers to increase the efficiency in the use of resources. However, in these cases, video sources are usually encoded in a restricted VBR mode. This means that the bit rate is allowed to change but only at specified points and within a restricted set of values. These multiplexers also have the drawback of having to be co-located to the encoders. They need to be in constant communication with the encoders, thus preventing remote application and increasing the complexity of the multiplexer and of the rate control mechanism of the video encoders. Our framework is directed to the use of true video VBR encoding and remote operation. The system architecture depicted in figure 1, shows the dissociation between encoders and multiplexer and the use of statistical information describing the video sources, by the multiplexer. Due to this last requirement, the use of this approach for real-time programmes implies that some sort of metadata is inserted in the video stream at coding time. This problem will have to be dealt separately, at a later stage. Because bit rate behaviour will be inferred through the use of prior knowledge obtained from sets of training of data, any misadjustments or errors made in the expected behaviour can be compensated in the receiver through the use of appropriate storage. For content stored in servers, the problem is easily solved using the content classification module represented in figure 1. The necessary statistical information is made available to the statistical multiplexer which uses it to predict the amount of unused bandwidth. The multiplexer may also delay for a few frame periods the transmission of distinct programmes in order to misalign the ocurrence of peak bit rates.

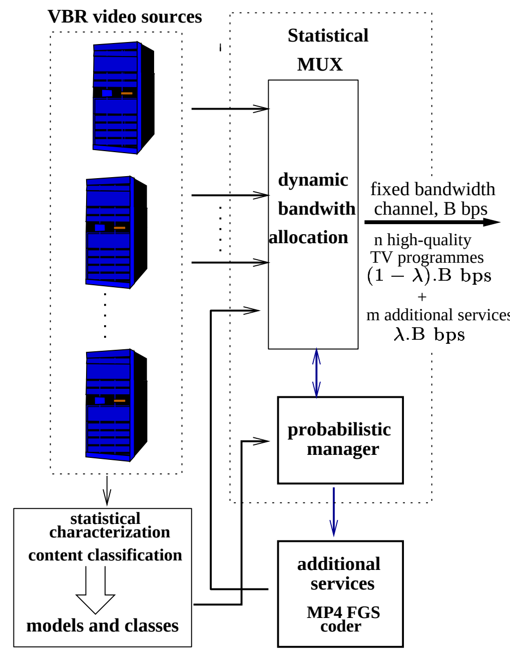

Figure 1: System architecture

3 Statistical framework

3.1 Problem formulation

VBR encoded video sources are statistically analysed in terms of variation of their mean bit rate whithin a GOP (Group Of Pictures) - the deviation of the each GOP's mean bit rate relative to the overall average bit rate (Dev_{GOP}). MPEG video streams usually consist of sequences of repetitive patterns, typically a pattern composed of two B frames and one P frame, with one I frame each $\frac{1}{2}$ s. For this reason and because the human visual system will not normally be able to distincly perceive variations in quality ocurring with less than $\frac{1}{2}$ s spacing, we have decided to analyse video sources on a GOP basis. Another important factor is that both in CBR and VBR modes of operation, there are variations in bit rate at the frame level due to the different types of frames, I, P and B. However, whereas in CBR mode the mean GOP bit rate is almost constant from GOP to GOP, in VBR it is not. Working at the GOP level allows us also to be able to trace the variations in bit rate due to variations in content. Our goal is then to be able to predict the ocurrence of valleys in the mean GOP bit rate, assuming a certain distribution for the duration or frequency of ocurrence of those valleys. Also to be able to correlate them with the type of content. Histograms are generated for the frequency of ocurrence and duration of bit rate valleys in different types of sources. The best matching probabilistic models to be used as priors, (with probability density function $p(\theta)$), are selected by performing the maximum likelihood fit of appropriate distributions to those histograms.

3.2 Inferential process

The statistical characterization of VBR bit streams provides the necessary ingredients to address the problem of predicting the bit rate behaviour as an inferential process. It provides the sample data, facilitates the selection of appropriate family of probability distributions and allows the formulation of inductive statements relative to hypothesis states and the sample data. These statements are in reality posteriori probabilities formulated under the Bayesian approach. Prior knowledge about the bit rate behaviour is included in the statistical model in the form of a prior probability distribution $p(\theta)$ in light of the sample data. Our model uses a variable V to indicate the ocurrence or not of a valley at a certain time t_1. Only two hipothesys can be evaluated at instante t_1 - either there is the beginning of a valley ($V_{t_1,\Delta_{GOP}} = 1$) or not ($V_{t_1\Delta_{GOP}} = 0$). The two possible states are inferred through the statistics extracted from the bit stream. Dev_{GOP} is the bit rate deviation from the expected mean GOP bit rate. It takes a complete GOP period (Δ_{GOP}) to decide whether there is a valley or not. Given a prior distribution for Dev_{GOP}, $p(Dev_{GOP})$, and the likelihood function $\mathcal{L}(V|Dev_{GOP})$, the posteriori distribution for Dev_{GOP} according to Bayes rule is given by $P(Dev_{GOP}|V) \propto p(Dev_{GOP})\mathcal{L}(V|Dev_{GOP})$. The probability that a valley occurs at instant t_2 within a period Δt after the occurrence t_1 of the last valley, and the probability that it does not occur, given the bit rate behaviour within Δt, are:

$\mathbf{V_{t_1,t_2} = 1}$: $\quad (t_2 = t_1 + \Delta t, \quad \Delta t \neq 0)$

$P(V_{t_1,t_2}=1 \mid V_{t_1}=0, Dev_{\Delta t}) = P(Dev_{\Delta t} \mid V_{t_1}=0,\ V_{t_1,t_2}=1) \times P(V_{t_1,t_2}=1 \mid V_{t_1}=0) = P(Dev_{\Delta t} \mid V_{t_1,t_2}=1) \times P(V_{t_1,t_2}=1 \mid V_{t_1}=0)$

$\mathbf{V_{t_1,t_2} = 0}$:

$P(V_{t_1,t_2}=0 \mid V_{t_1}=0, Dev_{\Delta t}) = P(Dev_{\Delta t} \mid V_{t_1}=0, V_{t_1,t_2}=0) \times P(V_{t_1,t_2}=0 \mid V_{t_1}=0) = P(Dev_{\Delta t} \mid V_{t_1,t_2}=0) \times P(V_{t_1,t_2}=0 \mid V_{t_1}=0)$

(*Note* : two valleys can not occur contiguously.)

These formulas can be further processed and used to compute the $posteriori\ log-likelihood$ ratio between the two cases. Using the probability distribution $p(\theta)$ for the period where bit rate valleys are not detected to compute $P(V = v)$, will allow to find the optimal solution to the problem of declaring the occurrence or not of a valley.

4 The importance of content

Constant picture quality in compressed video bitstreams means almost always variable bit rate streams. The number of bits produced by a VBR MPEG2 video encoder depends on the content of the video source being encoded. The same compression scheme applied to a varying content sequence, will consume either a variable amount of bits maintaining a constant picture quality level, or a fixed number of bits producing a varying picture quality stream. The last is the alternative commonly used in TV broadcasting. The overall transmission bandwidth is equally divided among all sources. The share assigned to each one is believed to be sufficient to ensure a minimum picture quality level regardless of the content. However, as the type of content (in terms of degree of difficulty to encode) varies from programme to programme or from scene to scene, the end-user will visualise varying quality pictures. Whithin a TV programme, it is expected that different scenes exhibit different bit rate requirements to achieve the same quality. Likewise, that different types of sources (such as action movies, talk-shows, news, cartoons, sports, etc) present different degrees of variation in bandwidth requirements from scene to scene. This means that it is possible to anticipate distinct behaviours of TV sources regarding their bit rate requirements according to the type of programme. As such, a classification of sources by type of content, will allow a more accurate identification of priors models for the bit rate behaviour of sources. A statistical multiplexer may thus anticipate the needs of each source based on these principles and using the information stored in the database. It may be able to find the best model that maximizes the actual data to predict the required bandwidth for each source. And therefore to predict the amount of released bandwidth that can be used by extra additional services.

5 Some statistical results

The video sequences used to extract statistics and formulate initial approaches to the problem of characterizing the occurrence of valleys in the mean GOP bit rate and thus inferring suitable probabilistic models, were sequences encoded in MPEG2 VBR with DVD quality. They consisted of a set of 1193 GOPs of a movie presenting some variation in the type of content troughout their duration. The histogram with the magnitude of the detected valleys is presented in figure 2. The histogram with the distance in GOPs between two consecutive valleys is shown in figure 3. Both the Weibull and the Gama distributions were used as models having provided good fits for the spacing among valleys.

Regarding the amount of bandwidth expected to be gained by using dynamically multiplexed VBR sources, the results obtained with an aggregation of 4 sources

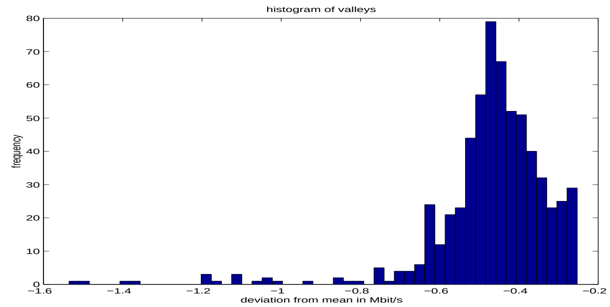

Figure 2: Histogram of valleys (*magnitude under a certain threshold*)

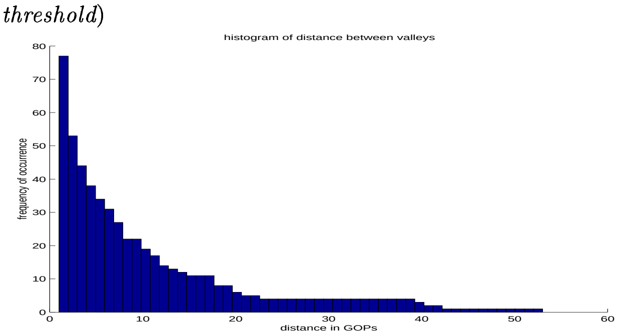

Figure 3: Histogram of distance between valleys

are presented in figure 4. For approximately equivalent qualities, CBR encoding uses almost always considerably more bandwidth than VBR. With 4 sources it is possible to obtain gains up to 7 Mbit/s (*circa* 30%) with an average of around 2,9Mbit/s (*circa* 12%).

6 Future work

This paper described a basic framework for the study of new multiplexing techniques based on the use of probabilistic modeling of VBR video sources, exploring inference techniques and content characterization and allowing the multiplexing operations to be performed independently of the location of the video encoders. Inference techniques may give the multiplexer the ability to predict the amount of bandwidth that will be available for use by other services. The use of VBR encoding and flexible allocation bandwidth algorithms when multiplexing a set of TV programmes gives bandwidth gains compared with static bit rate allocation methods and CBR encoding (with the same target picture quality). The work presented here addresses mainly the principles and objectives of a framework for the evaluation of statistical multiplexing tecnhiques and corresponding gains in bandwidth. The work will proceed applying those principles to longer and diverse training sequences and their aggregations. The system operation with real-time services will have to be addressed, in particular aspects related to buffering in the receiving equipment, required to allow seamless operation.

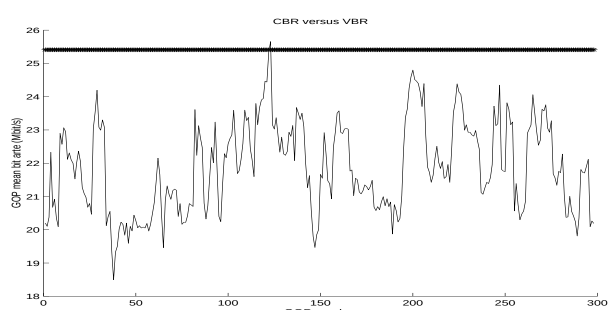

Figure 4: CBR versus VBR

References

[1] A. Stuart and J. K. Ord, "Advance Theory of Statistics: Classical Inference and Relationship", 5th edition, Vol.2, Edward Arnold Editors (1991).

[2] N. Vasconcelos, A. Lippman, "A Bayesian Video Modeling Framework for Shot Segmentation and Content Characterization", MIT Media Laboratory, Cambridge, MA, USA.

[3] M. Krunz, R. Sass and H. Hughes, "Statistical Characteristics and multiplexing of MPEG streams", Michigan State University, Department of Electrical Engineering.

[4] H. Radha, Y. Chen and M. Schaar, "Fine Granular Scalability: a new framework for Internet video coding and transmission", Philips Research, Briarcliff Manor, NY (2000).

[5] S. Huang, "Principles, benefits and applications of variable bit rate coding for digital broadcasting with statistical multiplexing extension", NAB'99 Conference Proceedings, Las Vegas, USA, (1999).

[6] D. Wheeler and J. Jeevanayagam, "Paving the Road to Interactivity", Scientific-Atlanta Inc., USA, (1999)

[7] R. Canonico, L. D'Ardia and G. Ventre, "A simulation based approach for network resources dimensioning in video delivery systems", Telecommunication Systems, Vol. 11 (1999).