

# IMPROVED QUALITY VIDEO CODING FOR CBR TRANSMISSION: Bit production control and pre-analysis

MARIA M. SIMON (M.) (\*), LUIS CASAMAYOU(\*), PAULO VILLEGAS(+), MIGUEL ROSER(+)

(\*)Facultad de Ingenieria, Montevideo, Uruguay  
Post address: Julio Herrera y Reissig 565, POBox 30, 11300 Montevideo, Uruguay  
E-mails: msimon@iie.edu.uy, casa@iie.edu.uy  
Phone: +5982 710974 Fax +5982 717435  
(+)Telefonica, Investigacion y Desarrollo  
Post address: c/Emilio Vargas 6, 28043 Madrid, Spain  
E-mails: pvillega@tid.es, mroser@tid.es  
Phone: +341 3374226 Fax: +341 337 4491

## ABSTRACT

This work presents a MPEG-compatible coding method designed for the production of visual information. Constant Bit Rate transmission is imposed. Some buffer control strategies, well suited for CBR and adapted to the specific applications are presented. In distributive applications, the edition is supposed to be off-line. This is exploited by performing a pre-analysis, detecting for instance the scene cuts or others characteristics to be taken into account when coding. The results show that the tradeoff between quality and CBR is good, compared with Test Models, for the specific application.

## 1. Introduction

MPEG standards [1, 2] are a widely accepted coding method. The syntax and semantics of the bit stream are fixed, but coding improvements are open.

In most video transmission applications, at least in earlier implementations, Constant Bit Rate (CBR) transmission is envisaged. It allows simpler implementation over current platforms (i. e. copper wires in low bit rate services). In addition, synchronization may be implemented through buffer filling. In fact, most of existent decoders work on roughly CBR bit streams.

On the other hand, CBR imposes strong requirements, which cause generally quality variations during the sequence. The goal of consistent quality is discussed in section 2.

Test models strategies [3, 4] are presented in section 3. They are short term buffer control strategies, that result in a variable quality. In section 4 a model that seeks for a tradeoff between CBR and consistent quality is presented.

In reference models [3, 4] the Group of Pictures (GOP) size is normally constant. However, the MPEG syntax allows to vary it during the play. The model presented in this work includes a pre - analyzer which detects characteristics of the material to be coded, as scene cuts. In this way, Intra frames

are located in convenient places, following rules that are listed in section 5. The whole coding process is performed in two steps: sequence analysis and source coding. This structure in independent modules allows parallelization between computers or in multiprocessor computers.

## 2. The goal of consistent quality

In previous works [6] we have studied a constant quality (CQ) model, whose operation looks for a constant Luminance Signal to Noise Ratio (LNSR) in each slice (i. e. region) of the frames. It was proved that this method gives almost no improvements (regarding the perceptual evaluation) in front of a constant quantization parameter ( $qp$ ) which is much simpler.

It should be noted that LNSR may be not closely related with perceptual quality. A subjective weighting of the material to be coded is done through the quantization matrixes. It is assumed that, with a good election of those matrixes, the value of  $qp$  is correlated with quality. So, a good approach to the goal of consistent quality is a stable  $qp$ . The problem is to adapt this parameter to the nominal bit rate and to the sequence to be coded. In this work a model is presented, whose control strategy is to choose a stable, **adaptive quantization parameter (QPA)**. In the presented versions, the quantization parameter value is updated only every frame.

In the following, we evaluate the consistency of quality upon the stability of  $qp$ .

We evaluate the near CBR operation upon the real buffer fullness, supposing that the flushing rate is constant. If  $qp$  changes often, the quality is variable. If it is controlled in a long period, the control is slow and buffer overflow may occur. Our work shows that the QPA model is a good tradeoff between a moderated buffer size (and so, a moderated delay) and a consistent quality. The results are compared with

the Test Models [3,4].

When scene cuts arise, the requirement of a near - constant  $qp$  is not possible. In addition, the prediction mode used in MPEG causes disturbances until an Intra frame arises. A scene cut detection was implemented by a pre-analysis of the sequence. This kind of procedures give a potentially important improvement in front of hardware real time coders which, in the other hand, are faster.

General principles about buffer control in CBR are discussed in [5]. A particular implementation of this kind of control is presented in [7].

### 3. Test Model strategy. GOP basis

Test models [ 3, 4] manage virtual buffers on a Group of Pictures (GOP) basis. We will review these strategies.

The determination of  $qp$  follows the steps:

1. At the beginning of each GOP, the target generation in the GOP is established upon the nominal bit rate and the remaining of the precedent GOP.

2. At the beginning of each frame a target generation is found as :

$$NT_A = \frac{Nbits\_rem\_GOP \times \frac{X_A}{K_A}}{\frac{nr_I \cdot X_I}{K_I} + \frac{nr_P \cdot X_P}{K_P} + \frac{nr_B \cdot X_B}{K_B}} \quad (1)$$

being:

$Nbits\_rem\_GOP$ : remaining bits in the GOP

$NTA$ : target generation in the frame A

$X_I, X_P, X_B$  : complexities of the last frame of each type

$nr_I, nr_P, nr_B$ : remaining frames (to be coded) in the

GOP, of each type.

$K_I, K_P, K_B$  : Model factors.

3. At the end of each macro block (MB) the buffer level is updated.

4. At the beginning of the next MB,  $qp$  is calculated from a linear dependence on the buffer level.

An additional step is performed.  $Qp$  is corrected upon the activity of the image. This correction looks for avoiding block or contour effects. As the activity correction is independent from the buffer control, we present our results without activity correction, in order to show clearly the effects of each model.

### 4. Consistent qp strategy. (QPA)

An important point to be noted in the previous description is the lack of symmetry in the processing of the frames within a GOP or in the macro blocks within a frame. The compliance with the target generation affects specially the last elements of a group.

The QPA model uses a common  $qp$  value in each frame (which may be activity corrected) going around the lack of

symmetry in the macro blocks processing. This procedure gives a smooth quality and a good buffer control.

The underlying assumption in the TM strategy is that the image complexity may be estimated as:

$$\text{complexity} = qp \cdot \text{generation} \quad (2)$$

and that the complexity is approximately constant, at least between scene cuts.

In the following, three versions of a model are presented. The  $qp$  value is always updated each frame, distributing the target generation on different basis. Distribution is based on "One GOP" available bits and remaining frames, like in the TM models, on "Two GOPs" or on a fictitious "sliding GOP". This last criterion gave the best results, as a tradeoff between quality and buffer stability.

The strategies were tested over several sequences.

Simulations shown in this work were performed on the test sequence "Table Tennis", which has panning and zoom effects and so causes potential buffer control instability.

#### 4.1 One GOP basis.

On the assumption of (2), for a frame  $i$  to be coded, we find the basic  $qp$  from the generation and the  $qp$  of the last frame of the same type. In (3), A stays for a generic frame type.

$$qp_A(i) = \frac{qp_A(i-1) \times \text{generation}_A(i-1)}{\text{Target\_generation}_A(i)} \quad (3)$$

The Target\_generation is found, in this version, like in the TM models.

The Fig. 1 compares the real buffer level in the QPA and TM cases. The step ( $qp$ ), shown in Fig. 3, is less variable, giving a more consistent quality. The control is robust even in front of a rapid panning.

#### 4.2 Two GOP basis.

The previous version improves the quality stability within each frame. However, Fig. 3 shows  $qp$  variations in the last frames of each GOP for both TM and QPA-1GOP models.

In order to go around this point we fixed the target generation of each frame using the remaining bits, remaining frames, etc., of two GOPs. Quality is improved, but the control is proven to be too slow.

#### 4.3 Sliding GOP.

As an alternative, target is fixed using always a fictitious, sliding GOP, which is adjusted taking into account the "carry" of the previous one.

This strategy proved to be more consistent than the GOP basis, and to give a robust buffer control.

Figs. 3 and 4 compare the QPA models on a GOP basis

and sliding GOP with the Test Model. The sliding GOP gives a very stable election of **qp**, even during an important scene change.

### 5. Scene cuts detection

Scene cuts are detected in order to put Intra frames in those places. In this way, the Group of Pictures (GOP) size is variable.

A frame distance is calculated, between original and motion compensated estimation. Normally, a scene cut produces

- absence of scene cuts.
- when two or more cuts occur in a time less than a minimum GOP, the last change is taken, for perceptual reasons
- when a “fusion” of images is detected, an Intra frame is put at the end of the fusion
- cuts prevail on fusions

#### 5.1 Cut detection algorithm

The key element in detection is the pre-analyzer of the material to be coded. It produces motion vectors files and an information file, which contains indications for the source coder. This information file is written in an editable form: a

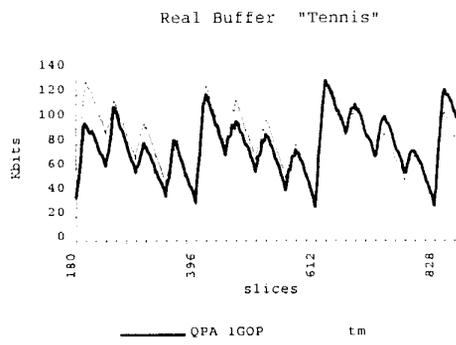


Fig. 1 : Real Buffer occupancy. TM vs. QPA 1GOP control strategies

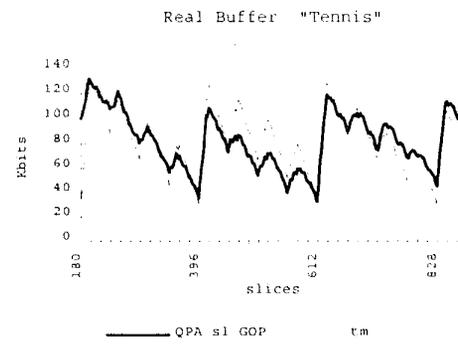


Fig. 2 : Real Buffer occupancy. TM vs. QPA Sliding GOP control strategies

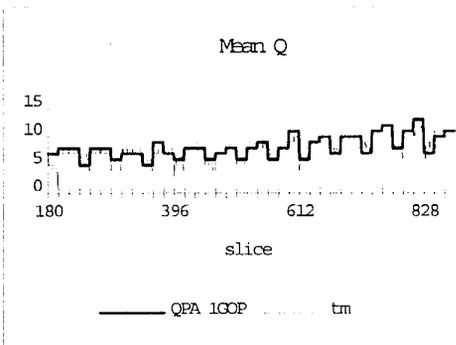


Fig. 3 : Mean Quantization Parameter. TM vs. QPA 1GOP control strategies

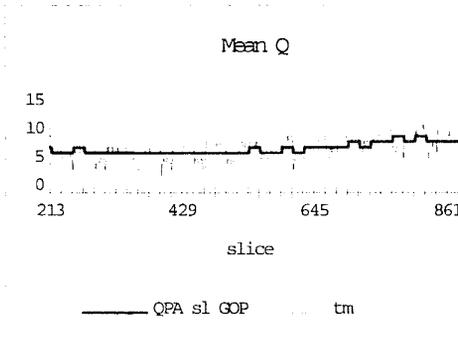


Fig. 4 : Mean Quantization Parameter. TM vs. QPA Sliding GOP control strategies.

a growth in this distance which is easy to detect. Reliable detection requires a preanalysis of the sequence. Some assumptions are obvious:

- the work performed to detect cuts should be useful for the ulterior coding process
- a minimum GOP size (minGOP) is imposed, to avoid buffer overflow or poor quality of Intra frames.
- a maximum (maxGop) is also respected, in order to provide access points in the video stream
- a target may be given, which is the GOP size (targetGop) in

coding expert can modify it or introduce additional information.

During pre analysis, forward motion vectors are calculated using a block matching method. Motion vectors are stored, to be used in coding. Motion estimation implies to calculate the distance between a Macroblock and its best predictor. Those distances are accumulated by frame. Frame distances depend on the complexity of the sequence. In a scene cut, the behavior of the distances presents a high, well localized peak. A fusion corresponds to a flat and wide peak.

The decision rule is:

```
if((dist(n) - dist(n+1)) / dist(n)) > THRES1)
cutDetected = n+1;
else if((dist(n) - dist(n+1)) / dist(n)) > THRES2)
fusionDetected = n+1;
```

An Intra frame is put in a cut or fusion detection only after a forward analysis in a range of minGOP+maxGOP frames, in order to optimize the places of Intra frames for the coding process and for the perceptual impact.

Figure 5 compares the Luminance Signal to Noise Ratio during a scene cut around the frame 25. After this change the sequence is more complex, giving a poorer LSNR. When the change is detected, no transient exists. It is to be pointed off that perceptual impact is much more impressive than figures of LSNR.

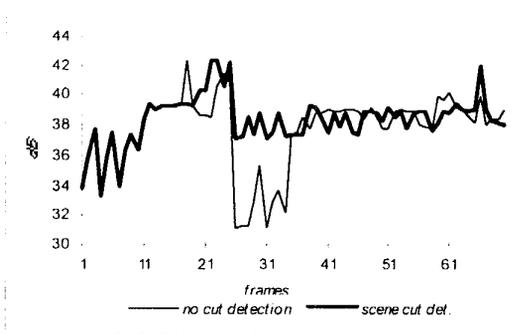


Figure 5. Luminance Signal to Noise Ratio during a scene cut, with and without preanalysis

## 6. Conclusions

The proposed model acts on  $qp$  for each frame. This gives an image free from quality variations, specially between upper and lower regions of the frame. It is well seen in images with a flat - detailed pattern, as "Flower Garden". Figs. 1 and 2 show that the buffer requirements are not stronger than in the TM case; in fact the total excursion is almost the same.

The quantization parameter, that gives a good measure of

the quality consistency, is stable, specially in the sliding GOP control. In Fig. 4 it can be seen the smooth adaptation of this parameter, during a rapid zoom and panning. This adaptation prevents buffer overflow. The sliding GOP version avoids also quality inconsistencies within the GOP.

The detection of scene cuts provides a good mechanism to avoid both disturbances in image quality and buffer instabilities.

## ACKNOWLEDGMENTS

This work was supported by Telefónica I+D, Madrid and the CONICYT (Consejo Nacional de Investigacion Cientifica y Tecnologica, Uruguay).

## REFERENCES

- [1] ISO/IEC JTC 1, IS 11172, (MPEG 1).
- [2] D. Le Gall. "The MPEG Video Compression Algorithm: A Review". SPIE Vol. 1452, Image Processing Algorithms and Techniques II. 1991.
- [3] SM3 - ISO/IEC - JTC1/SC2/WG11, Simulation Model Editorial Group "MPEG Video Simulation Model Three", July 1990.
- [4] TM5 - ISO/IEC - JTC1/SC29/WG11, Test Model Editing Committee "Test Model 5".
- [5] J. Zdepski et al: "Statistically based buffer control policies for constant rate transmission of compressed digital video". *IEEE Trans. on Com.* Vol. 39, No. 6, June 1991.
- [6] M. Simon, P. Villegas, J. Caballero, M. Roser: "A general approach to output rate control in video coding", presented at the SPIE, IS&T Symposium, California, Feb. 1993.
- [7] C.Chen, A.Wong. "A self governing rate buffer control strategy for pseudo constant bit rate video coding", *IEEE Transactions on Image Processing*, January 1993.