

# Cauchy-Density-Based Basic Unit Layer Rate Controller for H.264/AVC

Sergio Sanz-Rodríguez, *Student Member, IEEE*, Óscar del-Ama-Esteban, Manuel de-Frutos-López, *Student Member, IEEE*, Fernando Díaz-de-María, *Member, IEEE*

**Abstract**—The rate control problem has been extensively studied in parallel to the development of the different video coding standards. The bit allocation via Cauchy-density-based rate-distortion (R-D) modeling of the discrete cosine transform (DCT) coefficients has proved to be one of the most accurate solutions at picture level. Nevertheless, in some specific applications operating in real-time low-delay environments, a basic unit (BU) layer is recommended in order to provide a good trade-off between picture quality and delay control. In this paper, a novel BU bit allocation for H.264/AVC is proposed based on a simplified Cauchy probability density function (PDF) source modeling. The experimental results are twofold: 1) the proposed rate control algorithm (RCA) achieves an average PSNR improvement of 0.28 dB respect to a well known BU layer RCA, while maintaining a similar buffer occupancy evolution; and 2) it achieves to notably reduce the buffer occupancy fluctuations respect to a well known picture layer RCA, while maintaining similar quality levels.

**Index Terms**—H.264/AVC, low-delay, rate control, basic unit, bit allocation

## I. INTRODUCTION

THE inherent variability of video information implies consequently a variable output bit rate, which must be controlled in order to comply with the requirements of real-time video communications through constant bit rate (CBR) channels. For this purpose, several schemes for rate control have been recommended during the standardization process of video codecs, such as TM5 for MPEG-2 [1], VM8 for MPEG-4 [2], TMN8 for H.263 [3], and JM for H.264/AVC [4]. The hypothetical reference decoder (HRD), a normative part of the H.264/AVC standard that describes a set of requirements to transmit and decode bit streams, is taken into account in order to restrict the target bit assignment per picture by means of two (lower and upper) bounds for a virtual buffer fullness. The RCA must keep the buffer in secure levels, avoiding both overflow and underflow and, given that the decoder must wait for enough bits to begin the decoding process of a frame, the maximum allowed delay is directly related to the buffer size. In order to maintain the output rate within the buffer limits without noticeable visual quality degradation, the RCA must assign the most appropriate amount of bits and the corresponding quantization step (Q) to each coding unit. The

traditional RCAs select this Q value using an analytical rate-quantization (R-Q) function. This function is derived by means of a source modeling of the DCT coefficients. For instance, using a Gaussian PDF, a logarithmic R-Q function can be inferred [5]. On the other hand, considering a Laplacian PDF, different linear [6], quadratic [2] or  $\rho$ -domain-based [7] R-Q models have been proposed. Finally, assuming a Cauchy PDF, Kamaci *et al* [8] obtain a simple exponential R-Q model which outperforms the previous algorithms.

In the particular case of low-delay applications, where the buffer size is very small, the selection of Q on a frame basis is not accurate enough, and a finer rate control is desirable. Due to the growing popularity of these scenarios, such as in videophone or videoconferencing, several RCAs have been proposed, which employ an additional BU layer, defined as a group of macroblocks in raster scan order which share the same Q value. In [1], [3], [7] and [9] a single macroblock is proposed as BU size, while in [4] a greater size is recommended for a better trade-off between quality smoothness and target frame size adjustment.

In this paper, a Cauchy-PDF-based BU rate controller is derived, showing that exponential R-Q and R-D models, similar to those applied in [8] on a frame basis, work properly at the BU layer.

The paper is organized as follows. In Section II, the proposed BU layer RCA for H.264/AVC is described. In Section III, a pseudocode summary of the proposed algorithm is given. Experimental results are shown and analyzed in Section IV. Finally, in Section V some conclusions are drawn.

## II. BASIC UNIT LAYER RATE CONTROLLER

In real-time low-delay applications, the feasible buffer size is quite limited. As pointed out in [3] and [9], the picture layer RCAs are generally not suitable approaches for this kind of environments, since the misadjustments in the frame target bits increase the buffer overflow or underflow risk. Therefore, a fine quantization parameter (QP) value variation within each frame is recommended for a better adjustment to the frame target bits.

In the following subsections we describe the proposed Cauchy-PDF-based RCA for low-delay environments. It consists of three layers: the group of pictures (GOP) and picture layers are similar to those described in [4] and [8], while the BU layer is novel and entails a bit allocation algorithm, the consequent QP estimation and, finally, the R-Q and R-D model update.

Manuscript received XX X, 200X; revised XX XX, 2010. This work was supported in part by PRODYS, SL.

Copyright (c) 2010 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending an email to [pubs-permissions@ieee.org](mailto:pubs-permissions@ieee.org).

The authors are with the Department of Signal Theory and Communications, Universidad Carlos III de Madrid, Leganés, Madrid 28911 Spain (e-mail: {sescalona, oaesteban, mfrutos, fdiaz}@tsc.uc3m.es).

### A. GOP Layer

This layer computes the QP value for the first picture of each type,  $(QP_0^I, QP_0^P, QP_0^B)$ , which depends on the average number of target bits per pixel. The amount of remaining bits for the  $j^{\text{th}}$  picture in the  $i^{\text{th}}$  GOP is determined as follows:

$$B_i(j) = \begin{cases} \frac{J \cdot R}{f} - V_i(j), & \text{if } j=1 \\ B_i(j-1) - t_i(j-1), & \text{otherwise,} \end{cases} \quad (1)$$

where  $J$  represents the number of pictures in a GOP,  $R$  is the target bit rate,  $f$  is the frame rate,  $V_i(j)$  is the current encoder buffer fullness, and  $t_i(j-1)$  is the number of bits used to encode the last picture.

### B. Picture Layer

Similarly to [4], the bit allocation algorithm at picture level consists of main two terms, one modeling the target bits and the other the buffer status, and obeys the following expression:

$$T_i(j) = \begin{cases} \hat{T}_i(j), & \text{if I picture} \\ (1-\beta)\tilde{T}_i(j) + \beta\hat{T}_i(j), & \text{otherwise,} \end{cases} \quad (2)$$

with

$$\hat{T}_i(j) = \tilde{b}_i(j) + \tilde{h}_i(j), \quad (3)$$

and

$$\tilde{T}_i(j) = \frac{R}{f} + \delta(S_i(j) - V_i(j)). \quad (4)$$

The first term,  $\hat{T}_i(j)$ , is a target bit model of the remaining bits in a GOP,  $B_i(j)$ . The amount of target texture bits of the  $j^{\text{th}}$  frame in the  $i^{\text{th}}$  GOP,  $\tilde{b}_i(j)$ , is computed following [8]. The number of header and motion data bits,  $\tilde{h}_i(j)$ , is predicted by computing an average of the header and motion data bits of the previously encoded pictures in the GOP.

The second term,  $\tilde{T}_i(j)$ , is derived from the buffer occupancy,  $V_i(j)$ .  $S_i(j)$  is the target encoder buffer level after encoding the  $j^{\text{th}}$  picture in the GOP, which is computed either by means of a linear model for IP...P GOP patterns or a saw-tooth shaped model [10] for GOP patterns with B pictures.  $\delta$  is chosen as a trade-off between QP variation and target buffer level adaptation. And  $\beta$  determines the relative importance between both bit allocation models.

Finally,  $T_i(j)$  is bounded to satisfy the HRD constraints (see [4] for further details).

### C. Basic Unit Bit Allocation

As mentioned in Section I, a BU is a group of macroblocks in raster scan order which share the same quantization value. Therefore, the set of possible BU sizes, generally an entire fraction of the total number of macroblocks in a picture, goes from one macroblock to an entire frame.

We assume that the  $j^{\text{th}}$  picture in the  $i^{\text{th}}$  GOP is composed of  $K$  BUs ( $K > 1$ ), and the previous  $k-1$  BUs in the picture have already been encoded. We have a budget of  $T_{r,i}(j)$  bits to encode the remaining BUs ( $k, k+1, \dots, K$ ), which is calculated as:

$$T_{r,i}(j) = \begin{cases} T_i(j), & \text{if } k=1 \\ T_{r,i}(j) - t_{k-1,i}(j), & \text{otherwise,} \end{cases} \quad (5)$$

where  $t_{k-1,i}(j)$  represents the number of bits used to encode the previous basic unit. Then, we want to find the  $\tilde{b}_{l,i}(j)$  target texture bits for the  $l^{\text{th}}$  BU so that:

$$T_{r,i}(j) = \sum_{l=k}^K (\tilde{b}_{l,i}(j) + \tilde{h}_{l,i}(j)), \quad (6)$$

where  $\tilde{h}_{l,i}(j)$  is a prediction of the number of header and motion data bits of the remaining basic units.

In order to maintain consistent video quality within the frames, the proper bit allocation for the current BU is subject to:

$$D_{l,i}(j) \approx D_{m,i}(j) \quad l \neq m, \quad (7)$$

where  $D_{l,i}(j)$  represents the distortion of the  $l^{\text{th}}$  BU in the  $j^{\text{th}}$  picture. In [8] is assumed that the same QP value in consecutive pictures produces similar distortion, since the global variation between consecutive pictures is usually moderated. However, at BU level, this assumption is no longer valid because of the potentially large differences in complexity of neighboring BUs. Thus, the theoretical approach becomes different from that used in [8]. Specifically, we propose to use the simplified R-D function derived from a Cauchy-type model of the DCT coefficient distribution, i.e.:

$$D(R) \approx cR^{-\gamma}, \quad (8)$$

where  $c$  and  $\gamma$  are the model parameters. Using this R-D model and the constant quality condition (7), the target texture bits for the  $l^{\text{th}}$  basic unit,  $\tilde{b}_{l,i}(j)$ , can be obtained from those of the current one,  $\tilde{b}_{k,i}(j)$ , as follows:

$$\tilde{b}_{l,i}(j) = c_{l,j}^{\frac{1}{\gamma_{l,j}}} c_{k,j}^{-\frac{1}{\gamma_{l,j}}} \tilde{b}_{k,i}(j)^{\frac{\gamma_{k,j}}{\gamma_{l,j}}}. \quad (9)$$

The model parameters  $c_{l,j}$  and  $\gamma_{l,j}$  can take different values according to the type of picture, as it will be described in Subsection II-E. As mentioned before, the complexity of neighboring BUs may be significantly different in those pictures with large spatial heterogeneity. Thus, the assumptions  $c_{l,j} = c_{k,j}$  and  $\gamma_{l,j} = \gamma_{k,j}$ , with  $l \neq k$ , are no longer valid. Therefore, combining the expressions (6) and (9), the following BU bit allocation is obtained:

$$T_{r,i}(j) = \sum_{l=k}^K \left( c_{l,j}^{\frac{1}{\gamma_{l,j}}} c_{k,j}^{-\frac{1}{\gamma_{l,j}}} \tilde{b}_{k,i}(j)^{\frac{\gamma_{k,j}}{\gamma_{l,j}}} + \tilde{h}_{l,i}(j) \right). \quad (10)$$

Finally, assuming the same header and motion data bits for the  $N_r$  remaining BUs within the picture, the expression (10) can be rewritten as:

$$T_{r,i}(j) = \sum_{l=k}^K \left( c_{l,j}^{\frac{1}{\gamma_{l,j}}} c_{k,j}^{-\frac{1}{\gamma_{l,j}}} \tilde{b}_{k,i}(j)^{\frac{\gamma_{k,j}}{\gamma_{l,j}}} \right) + N_r \tilde{h}_{k,i}(j), \quad (11)$$

where  $\tilde{h}_{k,i}(j)$  can be obtained as the average number of header and motion data bits used to encode the previous  $k-1$  basic units. The Newton's method can be applied in (11) to iteratively obtain a solution for  $\tilde{b}_{k,i}(j)$ .

TABLE I  
 $\alpha_{k,P/B}$  COEFFICIENT UPDATE.

	$\alpha_{k,P}$	$\alpha_{k,B}$
$B_{pp,k} < 0.05$	1.60	2.00
$0.05 \leq B_{pp,k} < 0.10$	1.40	1.80
$0.10 \leq B_{pp,k}$	1.20	1.60

TABLE II  
 $\gamma_{k,P/B}$  COEFFICIENT UPDATE.

	$\gamma_{k,P}$	$\gamma_{k,B}$
$B_{pp,k} < 0.07$	0.50	0.45
$0.07 \leq B_{pp,k} < 0.13$	0.70	0.60
$0.13 \leq B_{pp,k} < 0.20$	0.70	0.75
$0.20 \leq B_{pp,k}$	1.00	0.75

#### D. Quantization Parameter Estimation

When the  $j^{th}$  P or B picture within the  $i^{th}$  GOP is to be encoded, the quantization parameter value for the  $k^{th}$  basic unit,  $QP_{k,i}(j)$ , an integer value between 0 and 51, is obtained from the corresponding quantization step,  $Q_{k,i}(j)$ , which is computed using the Cauchy-PDF-based R-Q model:

$$Q_{k,i}(j) = \left( \frac{\tilde{b}_{k,i}(j)}{a_{k,j}} \right)^{-\frac{1}{\alpha_{k,j}}}, \quad (12)$$

where  $a_{k,j}$  and  $\alpha_{k,j}$  are the model parameters. The remaining operations for QP selection at BU layer such as:

- the QP value to be assigned to the first BU in the picture,
- the QP increment to be applied when the frame bit budget is over and there are still some BUs to be encoded,
- and the clipping algorithms to bound the final QP value,

are the same as those proposed in [4].

For intra pictures other than the first one, the QP estimation at BU layer is not recommended, since eventual object movements might reduce the correlation between co-located BUs when the distance between intra pictures is large. Consequently, the picture layer RCA described in [8] is used to obtain a unique quantization parameter value to encode all BUs in these intra pictures.

#### E. Rate-Quantization and Rate-Distortion Model Update

After encoding the  $k^{th}$  BU of the  $j^{th}$  inter picture, the coefficients  $a_{k,P/B}$  and  $c_{k,P/B}$  are updated as follows:

$$a_{k,P/B} = b_{k,i}(j) \tilde{Q}_{k,i}(j)^{\alpha_{k,P/B}}, \quad (13)$$

$$c_{k,P/B} = d_{k,i}(j) b_{k,i}(j)^{\gamma_{k,P/B}}, \quad (14)$$

where  $b_{k,i}(j)$  is the number of texture bits used to encode the  $k^{th}$  basic unit, and  $d_{k,i}(j)$  is the sum of the square error between the original and reconstructed luminance pixels in the basic unit.  $\tilde{Q}_{k,i}(j)$  is the quantization step value obtained from  $QP_{k,i}(j)$ , the quantization parameter for the  $k^{th}$  BU of the  $j^{th}$  frame in the  $i^{th}$  GOP. Furthermore, if the P/B picture is the first one in the GOP, Tables I and II are employed to update the parameters  $\alpha_{k,P/B}$  and  $\gamma_{k,P/B}$  according to the average number of texture bits per pixel of the  $k^{th}$  basic unit,  $B_{pp,k}$ , which obeys the following expression:

$$B_{pp,k} = \frac{b_{k,i}(j)}{N_{p,BU}}, \quad (15)$$

where  $N_{p,BU}$  is the number of pixels per BU (if the BU size is one macroblock,  $N_{p,BU} = 16 \times 16 \times 1.5$  for a 4:2:0 format sequence). Thresholds used to set the  $\alpha_{k,P/B}$  value (see Table I) are the same as those applied to update the parameter  $\alpha_{P/B}$  at picture layer [8], and thresholds used to set the  $\gamma_{k,P/B}$  value have been determined empirically (see Table II).

Finally, after encoding a whole I/P/B picture, the model coefficient at picture layer,  $a_{I/P/B}$ , must also be updated (see [8]) for the next picture bit allocation. The required  $Q_i(j)$  value is obtained as the quantization step associated to  $QP_i(j)$ , the average quantization parameter value of  $QP_{k,i}(j)$ .

### III. ALGORITHM OUTLINE

The complete RCA is summarized as follows:

#### A. GOP and Picture Layers

1. Initialize GOP number:  $i = 1$
2. Compute initial QP values  $QP_0^I/QP_0^P/QP_0^B$
3. For each picture  $j = 1$  to  $J$ , do
  - Update buffer level  $V_i(j)$ : Ref. [4]
  - Update  $B_i(j)$  remaining bits to encode GOP: (1)
  - Compute  $T_i(j)$  target bits: (2), (3), (4), Ref. [8]
  - If I picture
    - Compute  $QP_i(1)$ : Ref. [8]
    - Encode picture
  - Else
    - Go to step B1
  - End
  - Update  $a_{I/P/B}$ : Ref. [8]
  - If first I/P/B picture in GOP
    - Update  $\alpha_{I/P/B}$ : Ref. [8]
  - End
4.  $i = i + 1$
5. Go to step A3

#### B. Basic Unit Layer

1. For each BU  $k = 1$  to  $K$ , do
  - Update  $T_{r,i}(j)$  remaining bits to encode frame: (5)
  - Compute  $\tilde{b}_{k,i}(j)$  target texture bits: (11)
  - Compute  $QP_{k,i}(j)$ : (12), Ref. [4]
  - Encode basic unit
  - Update  $a_{k,P/B}$  and  $c_{k,P/B}$ : (13), (14)
  - If first P/B picture in GOP
    - Update  $\alpha_{k,P/B}$  and  $\gamma_{k,P/B}$ : Tables I, II
  - End
- End

### IV. EXPERIMENTS AND RESULTS

The proposed BU layer RCA was implemented on the Joint Video Team (JVT) H.264/AVC reference software version JM 10.2 [11]. In order to assess its performance, it was compared to the Laplacian-PDF-based scheme adopted by JVT [4] and the Cauchy-PDF-based picture layer RCA proposed in [8].

Two sets of color video sequences were used in the experiments. The first one consisted of the following QCIF sequences: "Carphone", "MotherDaughter", "News", "Salesman", "Silent" and "Paris-Football". The second one consisted of the following CIF sequences: "Coastguard", "Container", "Football", "Foreman", "Paris" and "Bus-Foreman". Note that the last sequence in each group contained at least one scene change. Thus, the behavior of the proposed method in non-stationary situations could be evaluated as well.

The aforementioned sequences were encoded with the following configuration according to a low-delay application:

- **Profile:** main
- **Number of pictures:** 300
- **Frame rate:**  $f = 25$  f/s
- **GOP:** 300 pictures IP...P
- **Target rates (QCIF/CIF):**  $R = 32/128, 64/256, 96/512, 128/1024$  kbits/s
- **Buffer size:** 300 ms
- **BU size:** a row of macroblocks
- **R-D optimization:** enabled
- **Symbol mode:** CABAC

Given a sequence and a target bit rate, the quantization parameter value for the first picture of the sequence,  $QP_0^I$ , was the same for all the compared RCAs. It was selected so that the encoder buffer does not overflow during the encoding process. Moreover, a row of macroblocks is proposed as BU size to achieve a good trade-off among quality, buffer fluctuation and computational cost.

The  $\delta$  value in the expression (4) was set to 0.5 to achieve a fast enough adjustment to the target buffer level without increasing the QP fluctuation. After studying the behaviour of the proposed algorithm for different  $\beta$  values, this parameter was fixed to 0.5 for our buffer-constrained scenario, which is a good trade-off between PSNR and buffer fluctuation, and agrees with the value employed in [4]. Furthermore, for comparison purposes, the RCA described in [8] was also adapted to the low-delay requirement by including the same bit allocation model described in Subsection II-B, with identical  $\delta$  and  $\beta$  values.

Some representative video sequences and target rates have been selected to summarize the encoding results. Luminance PSNR (mean and standard deviation) and output bit rate are shown in Table III. In addition, the evolution of PSNR and encoder buffer occupancy versus picture number are depicted in Figs 1 and 2, respectively. When compared to another scheme including a BU layer [4], our proposal achieves an average PSNR improvement of 0.28 dB while exhibiting a similar buffer occupancy evolution. When compared to a picture layer rate controller [8], the proposed algorithm achieves, as expected, a finer control of both the output bit rate and the buffer occupancy evolution, reducing notably the risk of underflow and overflow (see for example the buffer occupancy evolution around picture #130 for "Paris-Football" or around picture #155 for "Bus-Foreman"), while maintaining similar quality level.

TABLE III  
PSNR (MEAN AND STANDARD DEVIATION) AND OUTPUT RATE FOR SOME VIDEO SEQUENCES AND TARGET RATES.

Sequence Target Rate (kbits/s)	Algorithm on JM 10.2	Mean PSNR (dB)	PSNR Deviation (dB)	Output Rate (kbits/s)
News (QCIF) 64	[4]	37.20	1.07	64.13
	[8]	37.27	1.10	63.92
	Proposed	37.53	1.19	64.20
News (QIF) 128	[4]	41.68	0.93	128.22
	[8]	41.66	0.93	128.01
	Proposed	42.12	1.08	128.06
Paris-Football (QCIF) 64	[4]	28.24	2.85	64.57
	[8]	28.59	3.07	64.16
	Proposed	28.49	3.19	64.02
Paris-Football (QCIF) 128	[4]	32.14	3.54	128.12
	[8]	32.54	3.90	130.16
	Proposed	32.62	3.74	127.98
Silent (QCIF) 64	[4]	35.78	1.02	64.04
	[8]	36.21	1.37	60.01
	Proposed	36.36	1.32	64.04
Silent (QCIF) 128	[4]	40.35	1.07	128.04
	[8]	40.82	1.26	127.94
	Proposed	40.67	1.18	128.02
Bus-Foreman (CIF) 256	[4]	32.31	3.85	256.13
	[8]	32.54	3.79	255.77
	Proposed	32.41	3.81	256.06
Bus-Foreman (CIF) 1024	[4]	38.64	3.30	1024.22
	[8]	38.79	3.26	1023.75
	Proposed	38.63	3.30	1024.05
Football (CIF) 256	[4]	25.84	0.64	255.90
	[8]	25.96	0.66	256.19
	Proposed	25.91	0.63	255.78
Football (CIF) 1024	[4]	31.77	0.71	1023.59
	[8]	31.84	0.72	1024.20
	Proposed	31.96	0.62	1023.69
Paris (CIF) 256	[4]	32.58	0.69	256.12
	[8]	33.24	0.91	255.99
	Proposed	33.47	0.89	256.01
Paris (CIF) 1024	[4]	41.58	0.85	1023.97
	[8]	41.85	0.85	1023.72
	Proposed	41.86	0.89	1023.93

## V. CONCLUSIONS

In this paper, a BU layer RCA has been proposed for H.264/AVC. Starting from a simplified Cauchy PDF for modeling of DCT coefficient distribution, a novel bit allocation has been developed and the exponential R-Q model has been applied on a BU basis.

In order to assess its performance, the RCA was integrated on an H.264/AVC encoder, which was configured according to a real-time low-delay application and compared to two well known RCAs. First, the proposed algorithm achieves a 0.28 dB average PSNR improvement over the Laplacian-PDF-based BU layer rate controller of JM 10.2 [4], with similar buffer occupancy evolutions. And second, our proposal notably reduces the buffer fluctuation in sequences with time varying complexity while maintaining a similar quality level, when compared to the picture layer RCA described in [8]. Therefore, the proposed BU extension is recommended for buffer-constrained scenarios.

The computational cost of the proposed BU layer algorithm obviously increases with respect to that of the picture layer RCA reported in [8], since the rate control operations are

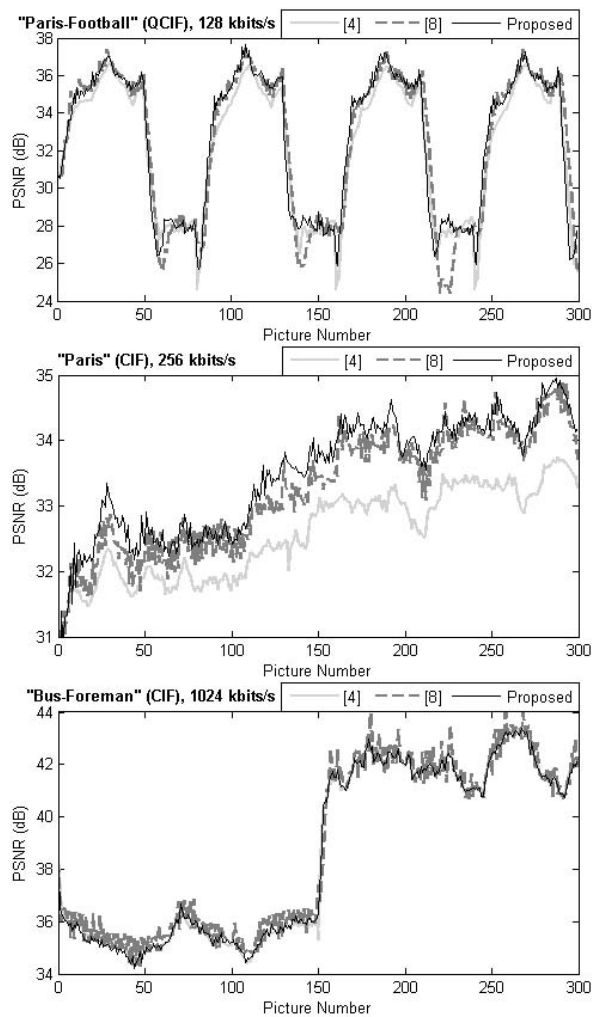


Fig. 1. PSNR evolution for some video sequences and target rates.

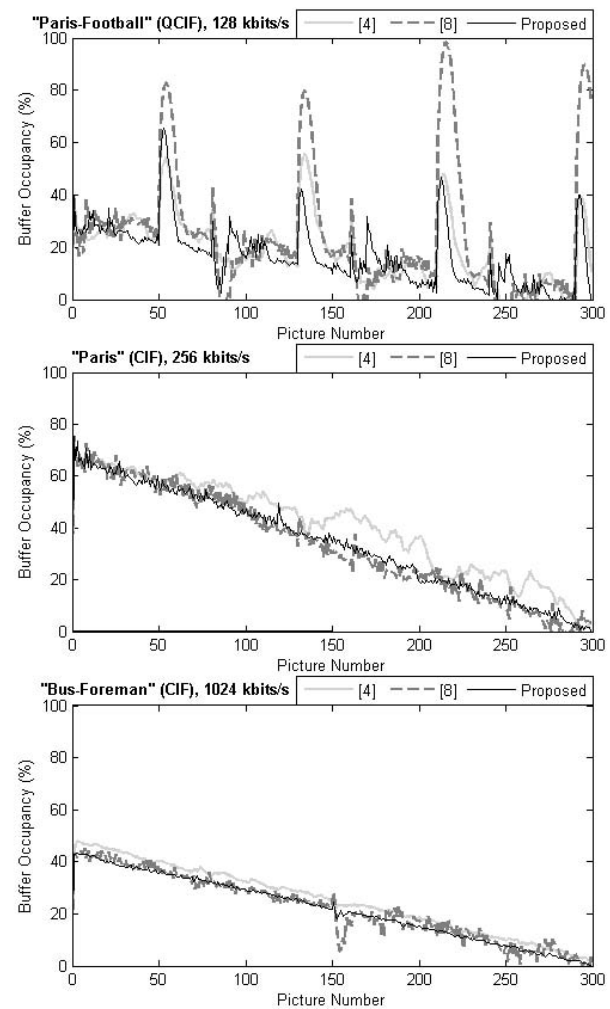


Fig. 2. Encoder buffer occupancy evolution for some video sequences and target rates.

performed on a BU basis. Nevertheless, the computational requirements to update the R-Q and R-D models associated to each BU in P/B pictures,  $a_{k,P/B}$  and  $c_{k,P/B}$ , are low when compared to the Laplacian-PDF-based BU layer scheme used in JM 10.2 [4]. This last RCA involves updating parameters on both the quadratic R-Q and the complexity prediction models by means of a regression analysis of previously encoded BUs. Our proposal simply updates  $a_{k,P/B}$  and  $c_{k,P/B}$  following the expressions (13) and (14), which depend linearly on the number of texture bits, the distortion generated by the  $k^{th}$  BU, and the quantization step value.

## REFERENCES

- [1] Test Model 5 [Online], "http://www.mpeg.org/MPEG/MSSG/tm5/"
- [2] T. Chiang, and Y.-Q. Zhang, "A new rate control scheme using quadratic rate distortion model," in *Image Processing, 1996. Proceedings., International Conference on*, vol. 1, 1996, pp. 73–76 vol. 2.
- [3] J. Ribas-Corbera, and S. Lei, "Rate control in DCT video coding for low-delay communications," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 9, no. 1, pp. 172–185, 1999.
- [4] S. Ma, Z. Li, and F. We, "Proposed draft of adaptive rate control," *JVT-H017, 8th JVT Meeting*, Geneva, Switzerland, May 2003.
- [5] B. Tao, B. Dickinson, and H. Peterson, "Adaptive model-driven bit allocation for MPEG video coding," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 10, no. 1, pp. 147–157, Feb. 2000.
- [6] S. Ma, W. Gao, and Y. Lu, "Rate-distortion analysis for H.264/AVC video coding and its application to rate control," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 15, no. 12, pp. 1533–1544, 2005.
- [7] Z. He, Y. K. Kim, and S. Mitra, "Low-delay rate control for DCT video coding via  $\rho$ -domain source modeling," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 11, no. 8, pp. 928–940, 2001.
- [8] N. Kamaci, Y. Altunbasak, and R. Mersereau, "Frame bit allocation for the H.264/AVC video coder via Cauchy-density-based rate and distortion models," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 15, no. 8, pp. 994–1006, 2005.
- [9] M. Jiang, and N. Ling, "Low-delay rate control for real-time H.264/AVC video coding," *Multimedia, IEEE Transactions on*, vol. 8, no. 3, pp. 467–477, 2006.
- [10] S. Sanz-Rodriguez, M. de-Frutos-Lopez, I. Gonzalez-Diaz, and J. Cid-Sueiro, "A rate control algorithm for low-delay H.264 video coding with stored-B pictures," in *Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference on*, vol. 1, 15–20 April 2007, pp. I-1153–I-1156.
- [11] JM 10.2 [Online], "http://iphome.hhi.de/suehring/tml/download/old\_jm/"