



**HAL**  
open science

## Inter-frame dependent rate and distortion models for statistical multiplexing of video programs

Nesrine Changuel, Bessem Sayadi, Michel Kieffer

► **To cite this version:**

Nesrine Changuel, Bessem Sayadi, Michel Kieffer. Inter-frame dependent rate and distortion models for statistical multiplexing of video programs. *IEEE Vehicular Technology Magazine*, 2009, 4 (3), pp.62 - 68. hal-00447002

**HAL Id: hal-00447002**

**<https://hal.science/hal-00447002>**

Submitted on 14 Jan 2010

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Inter-frame dependent rate and distortion models for statistical multiplexing of video programs

Nesrine Changuel, Bessem Sayadi

Alcatel-Lucent  
Bell-Labs France

Route de Villejust, 91620, Nozay, France

Email: {nesrine.changuel, bessem.sayadi}@alcatel-lucent.com

Michel Kieffer

Laboratoire des Signaux et Systemes  
CNRS - Universite - Paris-Sud, 11

3 rue Joliot-Curie, 91192 Gif-sur-Yvette, France,

Email: kieffer@lss.supelec.fr

**Abstract**—Statistical multiplexing (SM) is a useful technique for transmitting multiple video streams over a channel while satisfying multiple simultaneous bandwidth and quality related constraints. Design of efficient SM algorithms is based on Rate and Distortion (R-D) information extracted from successive encoded frames. Good R-D models usually take into account inter-frame dependencies specified in terms of impact of encoding parameters of a given frame over the next one. A highly accurate low complexity rate and distortion model is proposed in this paper. A minimum number of R-D measurement points per stream is shown to be sufficient to predict the effect of inter-frame dependencies. Design of efficient SM algorithms satisfying the above mentioned performance constraints on the basis of the proposed R-D model is described.

**Index Terms**—R-D models, H264/AVC coding, Statistical multiplexing, video broadcasting

## I. INTRODUCTION

Dynamically allocating transmission bandwidth between multiple video streams is one of the major challenges of broadcast systems such as DVB or S-DMB. Statistical multiplexing [1] based on joint control of multiple video coders offers a good quality smoothing solution in this situation. Additionally, the attempt is made to maximize the number of video streams sharing the same transmission channel in the face multiple simultaneous constraints. These are related to bandwidth usage, minimum required video quality, smooth quality transitions over time and balanced quality between streams [2]. Rate and Distortion (R-D) characteristics of the transmitted streams are known to offer a useful tool for solving the resulting constrained optimization problem.

The R-D characteristics for a video stream can be modeled either by considering each frame as an independent unit or by taking into account the inter-frame dependencies. The latter is used for performing *motion compensation* between a reference and a predicted frame. Several parametric models representing R-D behavior of video coders for independent frames have been proposed. For example [3] introduces a quadratic model used to solve the rate control problem for constant quality video. Nevertheless, this model does not take into account the control parameters of the video encoder. In [4], a  $\rho$ -domain model is proposed, where  $\rho$  indicates the number of null coefficients in a DCT encoded/quantified image block. This

model offers a reduced complexity solution due to its linear structure but it does not consider the inter-frame dependencies.

The quality of motion-compensated inter-Predicted (P) or Bidirectionally-predicted (B) frames is strongly impacted by the quality of the corresponding reference frame(s). Thus *dependent* R-D models are more complex than if independence between successive frames is considered. Moreover, the number of R-D measurement points required for a good quality fit is higher for such models. The *dependent* R-D model presented in [5] incorporates the influence of current and previous quantization parameters as well as the effect of texture and motion information. However, a large number of R-D measurements is needed for a good parametrical fit - 12 for the distortion model and 16 for the rate model. The inter-frame dependency is also considered in [6] where R-D models for an I-Frame and a P-Frame are presented. Nevertheless, these models do not take into account the quantization parameter (QP) of the considered video coder.

The *dependent* R-D models proposed in [5] and [6] are based on experimental analysis. An analysis of inter-frame dependency based on the statistical properties of DCT coefficients is proposed by [7]. The *pdf* of DCT coefficients follows a Laplace distribution and it is experimentally shown to be a function of the energy of the residual predicted frames, which is directly impacted by the reconstruction quality of the reference frame.

This paper proposes a new dependent R-D Model that offers a good prediction performance and a moderate computational complexity *i.e.*, requiring much less R-D measurement points per picture than in [5]. It focuses on the dependency between P frames and their reference I or P frames which are assumed to be unique in the video sequence. It analyzes the impact of quantization parameters of the reference frame on the rate and distortion of the predicted frames.

Section II introduces the statistical multiplexing problem. Section III introduces a R-D model considering independent successive frames. It is shown to be sufficiently accurate for R-D modeling of I-Frames. Then, in section IV a dependent R-D model is introduced. Multiple regimes constituting this model are explained in section V in terms of an independent regime switching auto-regressive model with variable quantization parameters. The theoretical R-D curves obtained with this model

are shown in section V to be very satisfying with a dependency behavior similar to that obtained in the experimental R-D curves of H.264 encoded frame. The concluding section VI presents some future perspectives on R-D models based on very few parameters estimation.

## II. PROBLEM STATEMENT

Statistical multiplexing is an efficient technique for maximizing the number of transmitted programs in broadcast systems. Figure 1 presents a broadcast system in which  $N$  video programs are encoded in parallel, the compressed bitstreams are multiplexed and transmitted over a communication channel at a constant rate  $R_c$ . The video programs are encoded using a video coder, *e.g.*, H264/AVC [8]. In this work, only two types of frames are considered: Intra (I) and Predictive (P) frames.

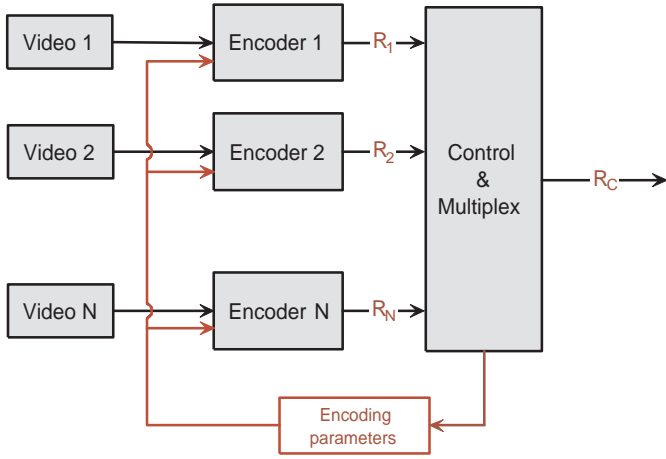


Fig. 1. Structure of a statistical multiplexer

Apart from bandwidth constraints, the control of each encoder involved in the statistical multiplexing has to ensure that:

- Each decoded program offers at least a minimum specified quality (quality constraint),
- programs are decoded with more or less the same quality (*fairness* constraint),
- for each program, the quality of the reconstructed video varies smoothly (*smoothness* constraint).

Satisfying all constraints simultaneously may sometimes be difficult. Due to the non-stationary content of each program. Variations may be due, *e.g.* to scene changes or to high activity within a scene. In [2], a predictive control scheme multiplexing H.264/AVC encoded video programs has been presented, taking all previously mentioned constraints into account. The control may be performed at a frame level (a distinct QP is selected for each frame), or considering several frames simultaneously (all sharing the same QP). If it is done at a frame level, for the  $j$ -th frame, [2] introduces a control window containing the previous, current, and  $w - 2$  future frames. The quantization parameters  $Q_j = (Q_{1j}, Q_{2j}, \dots, Q_{Nj})$  are evaluated for the  $j$ -th GOP by solving the following constrained optimization problem for the rate and quality control:

$$(\hat{Q}_j, \dots, \hat{Q}_{j+w-2}) = \arg \min_{Q_j, \dots, Q_{j+w-2}} \sum_{i=1}^N D_{ij}(Q_{ij}) \quad (1)$$

subject to

$$\begin{cases} \sum_{i=1}^N R_{ij}(Q_{ij}) \leq R_c \\ D_{ij}(Q_{ij}) \leq D_{\max} \\ |P_{ij}(Q_{ij}) - P_j| \leq \Delta P_p \\ |P_{ij}(Q_{ij}) - P_{ij-1}(\hat{Q}_{ij-1})| \leq \Delta P_s, \\ |P_{ij+k}(Q_{ij+k}) - P_{ij+k-1}(Q_{ij+k-1})| \leq \Delta P_s \\ \text{with } k = 1 \dots w - 2 \text{ and } i = 1, \dots, N. \end{cases} \quad (2)$$

$R_{ij}(Q_{ij})$  and  $D_{ij}(Q_{ij})$  denote the rate and distortion for the  $j$ -th frame in the  $i$ -th program when encoded with the quantization parameter  $Q_{ij}$ .  $D_{\max}$  is the maximum tolerated distortion to satisfy the minimum quality constraint. The quality of the  $j$ -th frame in the  $i$ -th program is evaluated by the Peak Signal-to-Noise Ratio (PSNR)

$$P_{ij}(Q_{ij}) = 10 \log_{10} \left( \frac{255^2}{D_{ij}(Q_{ij})} \right) \quad (3)$$

A controlled variation  $\Delta P_s$  of the PSNR between successive frames of a given program can be allowed in view of the smoothness constraint. The fairness constraint of the PSNR between successive frames of a given program, whereas for the fairness constraint, the difference of PSNR between programs is bounded by  $\Delta P_p$ .  $\hat{Q}_{i(j-1)}$  denotes the choice of the quantization parameter for the  $j - 1$ -th frame in the  $i$ -th program.

Performance of the control depends on the models for  $D_{ij}(Q_{ij})$  and  $R_{ij}(Q_{ij})$ . Considering frame  $j$  as Intra coded, if the future frames ( $j + 1, \dots, j + w - 2$ ) are considered independently, their corresponding R-D models will be tuned by the quantization parameters applied to each frame ( $Q_{j+1}, \dots, Q_{j+w-2}$ ). Nevertheless, for a given program, the R-D models of future frames are impacted by the choice of the quantization parameter at frame  $j$ . Therefore, R-D models should involve the quantization parameter used for the current frame and at least the one used in the corresponding reference frame. In the sequel, models for  $D_{ij}$  and  $R_{ij}$  are proposed as function of  $Q_{ij}$  and the QP at the reference frame, usually  $Q_{ij-1}$ . In what follows, the program index  $i$  is omitted.

## III. INDEPENDENT RATE AND DISTORTION MODEL

The aim of this section is to provide a independent R-D model for I frames. Experimental R-D curves for P frames show that the previous model may not be efficiently extended for these frames. The R-D curves of a P frame depend significantly on the QP of the corresponding reference frame.

### A. R-D model for I frames

This section introduces a R-D model suitable for I frames, which are encoded independently of any reference. Exponential R-D models presented in [2] are considered

$$R_j(Q_j) = a_R \exp(b_R Q_j), \quad (4)$$

and

$$D_j(Q_j) = a_D \exp(b_D Q_j) \quad (5)$$

where  $j$  denotes the index of the current frame and  $Q_j$  is the QP used to encoded it. At least two encoding trials per frame are necessary to estimate  $a_D$ ,  $b_D$ ,  $a_R$  and  $b_R$ . The

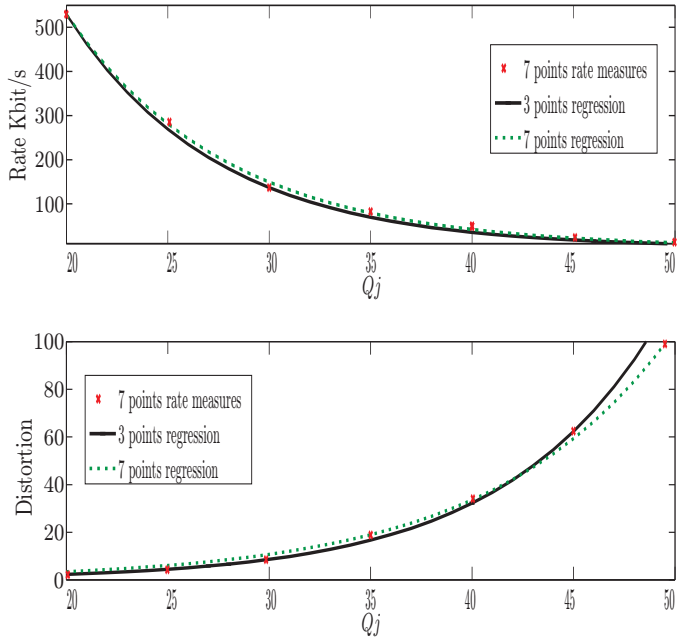


Fig. 2. Rate and distortion model fitting when 3 and 7 experimental values are considered (Foreman)

proposed model is tested on different frames in different video sequences. Figure 2 shows the rate and distortion functions for the first intra frame in *Foreman.cif*. Due to the non-precision of the obtained models with two encoding trials, we represent it with 3 and 7 measurements while  $Q_j$  vary. One sees that 3 measurements are enough to efficiently fit the R-D models.

### B. Experimental R-D curves for P frames

To determine whether the models (4) and (5) may be extended to P frames, several R-D curves have been plotted as a function of  $Q_{j-1}$ , the QP of the reference frame, considering several values  $Q_j$ , the QP of the P frame. The reference frame is an I frame here.

Figure 3 shows the logarithm of the rate as a function of  $Q_{j-1}$ , parametrized with  $Q_j$ , for the first P-frame in *Foreman.cif* sequence. Figure 4 shows the logarithm of the distortion for the same frame. The R-D curves depend significantly on  $Q_{j-1}$ . The parameters of the models (4) and (5) would thus depend on  $Q_{j-1}$ .

Dependent R-D models have thus to be introduced, as done in [5] and [6].

## IV. DEPENDENT RATE AND DISTORTION MODEL

In order to propose dependent R-D models, several regimes are first identified on the experimental curves of Figure 3 and 4

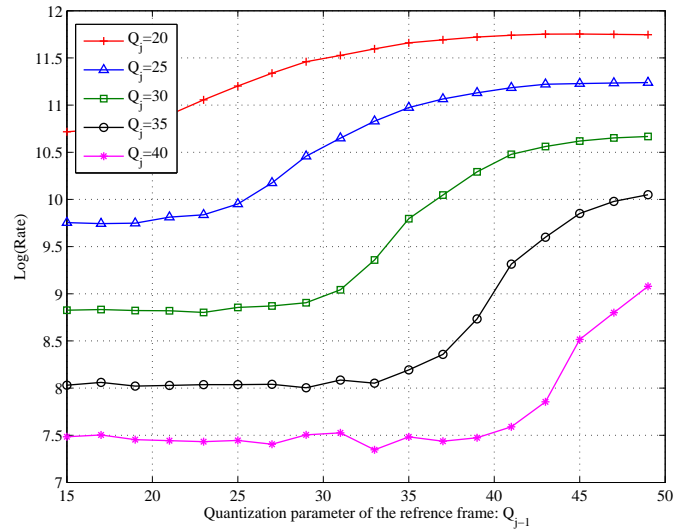


Fig. 3. logarithm of the rate as a function of  $Q_{j-1}$ , parametrized with  $Q_j$  in Foreman sequence

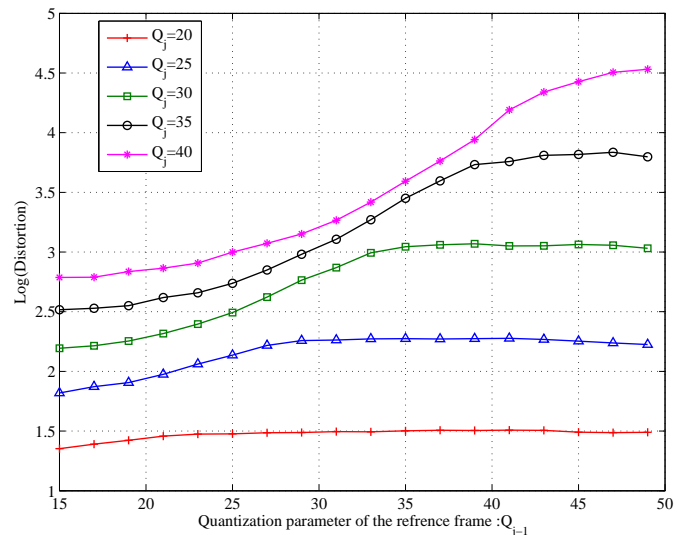


Fig. 4. logarithm of the distortion as a function of  $Q_{j-1}$ , parametrized with  $Q_j$  in Foreman sequence

### A. Rate regimes

All rate curves of Figure 3 decrease when  $Q_{j-1}$  decreases. When the quality of the reference frame increases, the motion compensation is usually more efficient, and the texture (motion compensation residuals) contains less energy, and requires thus less bits to be represented. Besides this tendency, three regimes may be identified for the rate curves.

When  $Q_{j-1} \leq Q_j$ , the rate remains approximatively constant. As long as the quality of the reference frame remains good, it allows to get a texture containing mainly elements which cannot be efficiently predicted from the reference frame, whatever its quality. The rate is thus mainly determined by the amount of elements which are difficult to predict in the P frame.

When  $Q_{j-1}$  is much larger than  $Q_j$  ( $Q_{j-1} \geq Q_j + K$ ), the rate remains approximatively constant. In this regime, the reference frame has been coarsely quantized, and is thus of poor quality. The motion compensation is not efficient and the amount of information in the texture is not significantly affected by a poor or very poor reference frame. The rate is thus essentially determined by  $Q_j$ .

The intermediate regime, when  $Q_{j-1} \in [Q_j, Q_j + K]$ , corresponds to the situation described at the beginning of this section. Increasing the quality of the reference frame reduces the amount of information of the texture, and thus the rate.

### B. Distortion regimes

All distortion curves of Figure 4, decreases when  $Q_{j-1}$  decreases, as for the rate curves. The interpretation is the same: having a better reference increases the efficiency of the motion-compensation, and reduces the amount of information in the texture. Even if the texture is coarsely quantized, the reconstructed frame may be of good quality thanks to the good quality of the reference frame. Now two regimes may be observed.

When  $Q_{j-1} \geq Q_j$ , the distortion remains approximatively constant, since it is essentially determined by  $Q_j$ .

When  $Q_{j-1} \leq Q_j$ , the distortion decreases when  $Q_{j-1}$  decreases. In this regime, the motion-compensation residuals of frame  $j$  are coarsely quantized compared to frame  $j-1$ . When the motion-compensation is performed, many blocks are well predicted, and for those blocks, the motion-compensation residuals are negligible when compared to the quantization step corresponding to  $Q_j$ . Many blocks have a distortion that is thus determined by  $Q_{j-1}$ .

The fact that the R-D curves do not converge to the same point when  $Q_{j-1}$  is small is due to the fact that independently of the value of  $Q_{j-1}$ , some block are not well predicted during motion-compensation. The corresponding texture for those blocks are thus of quite high energy. The R-D characteristics for those blocks are actually determined by  $Q_j$ , and have a behavior similar to I frames.

### C. Proposed dependent model

Based on the analysis of Sections IV-A and IV-B, the following piecewise linear models are proposed for  $\log(R_j(Q_j, Q_{j-1}))$  and  $\log(D_j(Q_j, Q_{j-1}))$

$$\log(R_j(Q_j, Q_{j-1})) = \begin{cases} \alpha_j^r + \beta_j^r Q_j & \text{if } Q_{j-1} \leq Q_j \\ \alpha_j^r + \beta_j^r Q_{j-1} & \text{if } Q_j \leq Q_{j-1} \leq Q_j + K \\ \alpha_j^r + \beta_j^r (Q_j + K) & \text{if } Q_{j-1} \geq Q_j + K \end{cases} \quad (6)$$

Where  $R_j$  is the distortion at the predicted frame. and

$$\log(D_j(Q_j, Q_{j-1})) = \begin{cases} \alpha_j^d + \beta_j^d Q_{j-1} & \text{if } Q_{j-1} \leq Q_j \\ \alpha_j^d + \beta_j^d Q_j & \text{else} \end{cases} \quad (7)$$

Where  $D_j$  is the distortion at the predicted frame.  $\beta_j^r$  and  $\beta_j^d$  are constants, and  $\alpha_j^r$  and  $\alpha_j^d$  correspond to the information linked to the current QP  $Q_j$ . This relationship is similar to

that shown in (4) and (5) for the rate and the distortion. In the logarithmic scale the relationship between  $\alpha_j^d$ ,  $\alpha_j^r$  and  $Q_j$  is so represented by a linear function. This induces to an R-D model involving a total of seven parameters.

## V. ALTERNATIVE DEPENDENT MODEL

The main difficulty with the preceding models (6) and (7) is that they require several coding trials of the P frame and of its reference frame with several distinct values of the pair  $(Q_{j-1}, Q_j)$  to allow their parameters to be identified. The transform DCT coefficients can be classified into DC and AC coefficients. The DC coefficient is the mean value of the image block and carries most of the energy in the image block. The AC coefficients carry energy depending on the amount of detail in the image block. Most of the energy is compacted in the DC coefficient and a few ac coefficients. In this section, an alternative approach is investigated, considering a model in the transform domain of the dependency between the DC coefficients of blocks of the P frame and of its reference frame.

The rate due to signaling and motion vectors is not considered here. Several authors agrees on the Gaussian model of the DC coefficient in the transform domain [9], [10]. For AC coefficients, there is less consensus, *e.g.*, [9] and [10] consider a Laplace distribution, whereas [11] use a Cauchy distribution.

### A. Independent regime-switching autoregressive model

Here, an independent regime-switching autoregressive (IRS-AR) model

$$Y_j = a_{X_j} Y_{j-1} + b_{X_j} U_j, \quad (8)$$

is considered to represent the DC value  $Y_j$  of a given block of frame  $j$  and  $Y_{j-1}$  is the DC value of the corresponding collocated block in the reference frame  $j-1$  (possibly after motion compensation). In (8),  $X_j$  is a sequence of independent and identically distributed (iid) binary-valued random variables with  $\Pr(X_j = 0) = 1 - \rho$  and  $\Pr(X_j = 1) = \rho$ ,  $U_j$  is a sequence of iid zero-mean and unit variance Gaussian variables. The sequences  $X_j$  and  $U_j$  are assumed to be independent. This model allows to take into account the fact that most of the time (when  $X_j = 0$ ), the DC value of collocated blocks is quite similar, in which case  $a_0$  is close to one, and that at some time instants (when  $X_j = 1$ ), there is few correlation, in which case  $a_1$  is close to zero ( $a_1 = 0$  in what follows).

It is assumed that the variance of  $Y_j$  does not depend on  $j$ . This imposes some constraints on the parameters  $a_0$ ,  $b_0$ , and  $b_1$ . If  $X_j = 0$ , one gets

$$\begin{aligned} \sigma_y^2 &= a_0^2 \sigma_y^2 + b_0^2 \\ &= \frac{b_0^2}{1 - a_0^2} \end{aligned} \quad (9)$$

and if  $X_j = 1$ , one obtains

$$\sigma_y^2 = b_1^2. \quad (10)$$

Combining (9) and (10), one gets

$$b_1^2 = \frac{b_0^2}{1 - a_0^2}. \quad (11)$$

The IRS-AR model is thus characterized by three parameters  $a_0$ ,  $b_0$ , and  $\rho$ , which are assumed to be known in what follows.

This model is only a very coarse approximation of the dependency between DC coefficients of successive frames. Nevertheless, at low bitrate, most of rate is due to the low frequency coefficients. The rate and distortion performance of scalar quantizers applied to  $Y_{j-1}$  and  $Y_j$  is studied in the following section.

### B. Rate and Distortion performance

Assume that  $Y_{j-1}$  is simply quantized with a scalar uniform midtread quantizer  $q_{j-1}$  with step size  $\Delta_{j-1}$  (intra coding). For  $Y_j$  a predictive coding is performed with

$$\hat{Y}_j = a_{X_j} q_{j-1}(Y_{j-1}) \quad (12)$$

as prediction for  $Y_j$  (the value of  $X_j$  is assumed to be known). The prediction residual

$$E_j^{X_j} = Y_j - \hat{Y}_j = a_{X_j} (Y_{j-1} - q_{j-1}(Y_{j-1})) + b_{X_j} U_j \quad (13)$$

is then quantized with a stepsize  $\Delta_j$ . Depending on  $X_j$ , (13) may become

$$E_j^0 = a_0 (Y_{j-1} - q_{j-1}(Y_{j-1})) + b_0 U_j$$

when  $X_j = 0$  and

$$E_j^1 = b_1 U_j$$

when  $X_j = 1$ .

The aim of the remainder of this section is to provide rate and distortion curves for the model (8) as a function of  $\Delta_{j-1}$  and  $\Delta_j$ . In the H.264/AVC standard [12], [13], characteristics of the quantizers depend on a quantization parameter  $Q$ . The relation between the quantization stepsize  $\Delta$  and  $Q$  may be approximated as

$$\Delta(Q) = 2^{\frac{Q-4}{6}} / PF, \quad (14)$$

where  $PF$  is a constant which value depends of the subband, see [14].

The distortion for  $Y_j$  may be written as

$$D_j^{X_j}(\Delta_{j-1}, \Delta_j) = \int_{-\infty}^{+\infty} (x - q_j(x))^2 f_{E_j^{X_j}}(x) dx, \quad (15)$$

where  $f_{E_j^{X_j}}(x)$  is the probability density function of  $E_j^{X_j}$ . The rate required to represent the quantized  $Y_j$  is evaluated as the entropy of the output of the quantizer fed with  $E_j^{X_j}$

$$R_j^{X_j}(\Delta_{j-1}, \Delta_j) = - \sum_{k=-\infty}^{+\infty} P_k(\Delta_{j-1}, \Delta_j) \log(P_k(\Delta_{j-1}, \Delta_j)) \quad (16)$$

where

$$P_k(\Delta_{j-1}, \Delta_j) = \int_{(k-\frac{1}{2})\Delta_j}^{(k+\frac{1}{2})\Delta_j} f_{E_j^{X_j}}(x) dx. \quad (17)$$

When  $X_j = 1$ ,  $E_j^1$  is zero-mean Gaussian with variance  $b_1^2 = \sigma_y^2$  and does not depend on  $\Delta_{j-1}$ ,

$$f_{E_j^1}(x) = \frac{1}{\sqrt{2\pi\sigma_y^2}} \exp\left(-\frac{x^2}{2\sigma_y^2}\right). \quad (18)$$

is thus known. High-rate approximations for (15) and (16) are easily obtained, see [15]. At medium to low rates (large values of  $\Delta_j$  compared to  $b_1$ ), such high-rate approximation becomes coarse, but (15) and (16) contain only few significant terms.

When  $X_j = 0$ ,  $f_{E_j^0}(x)$  is the convolution of the pdfs of  $a_0(Y_{j-1} - q_{j-1}(Y_{j-1}))$  and of  $b_0 U_j$ . One may show that

$$f_{E_j^0}(x) = \frac{1}{\sqrt{8\pi\sigma_y^2}} \sum_{k=-\infty}^{+\infty} \exp\left(-\frac{(x + k\Delta_{j-1}a_0)^2}{2\sigma_y^2}\right) G(x, \Delta_{j-1}, k), \quad (19)$$

where

$$G(x, \Delta_{j-1}, k) = \operatorname{erf}\left(\frac{2a_0x + \Delta_{j-1}(1 - 2k(1 - a_0^2))}{2\sqrt{2}\sigma_y\sqrt{1 - a_0^2}}\right) - \operatorname{erf}\left(\frac{2a_0x - \Delta_{j-1}(1 + 2k(1 - a_0^2))}{2\sqrt{2}\sigma_y\sqrt{1 - a_0^2}}\right). \quad (20)$$

Now, since  $f_{E_j^0}(x)$  and  $f_{E_j^1}(x)$  are known, one may evaluate numerically  $D_j^{X_j}(\Delta_{j-1}, \Delta_j)$  and  $R_j^{X_j}(\Delta_{j-1}, \Delta_j)$  using (15) and (16). The expectation of the rate and distortion with respect to  $X_j$  is then

$$D_j(\Delta_{j-1}, \Delta_j) = (1 - \rho) D_j^0(\Delta_{j-1}, \Delta_j) + \rho D_j^1(\Delta_j) \quad (21)$$

$$R_j(\Delta_{j-1}, \Delta_j) = (1 - \rho) R_j^0(\Delta_{j-1}, \Delta_j) + \rho R_j^1(\Delta_j), \quad (22)$$

since  $D_j^1$  and  $R_j^1$  do not depend on  $\Delta_{j-1}$ .

### C. Experimental results

Figures (5) and (6) represent  $\log D_j(Q_{j-1}, Q_j)$  and  $\log R_j(Q_{j-1}, Q_j)$ , with  $a_0 = 0.99$ ,  $b_0^2 = 3$ , and  $\rho = 0.1$ , meaning that 10% of DC coefficients may not be efficiently predicted from the previous frame. The conversion between  $\Delta$  and  $Q$  is done using 14. The chosen quantization parameters are  $Q_{j-1} \in [15, \dots, 50]$  for the reference  $Y_{j-1}$  and  $Q_j \in \{20, 25, 30, 35, 40\}$  for  $Y_j$ .

Theoretical R-D curves shown in Figures (5) and (6) have the same behavior as the R-D curves obtained using H.264/AVC video coder in Figures (3) and (4). The same regimes are observed. However, the two sets of rate curves do not have the same values. A constant element is missing in the theoretical curves to have the same values as in the experimental curves. This is due to the rate of the motion compensation residues which is not considered in the theoretical model. This later has an almost constant evolution while varying the quantization parameter, which perfectly explains the missing part in the total rate in the theoretical curves. Theoretical and experimental distortion curves are very close. In this case, all distortion elements are considered and the proposed model is able to accurately reproduce the distortion values from the H.264/AVC video coder.

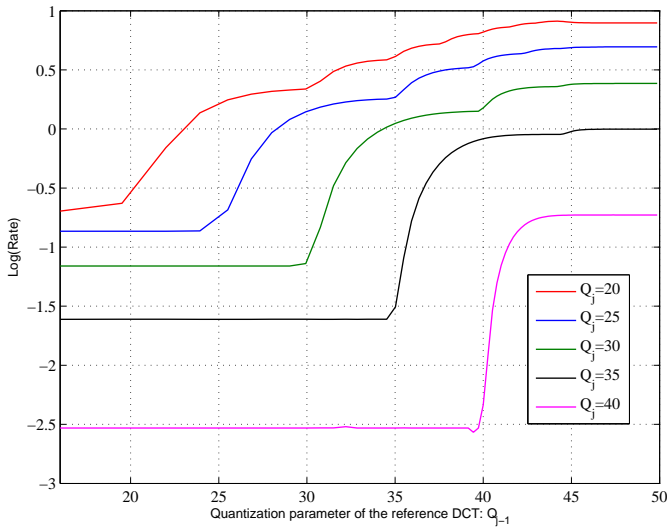


Fig. 5. Rate dependency between current and reference DCT coefficient

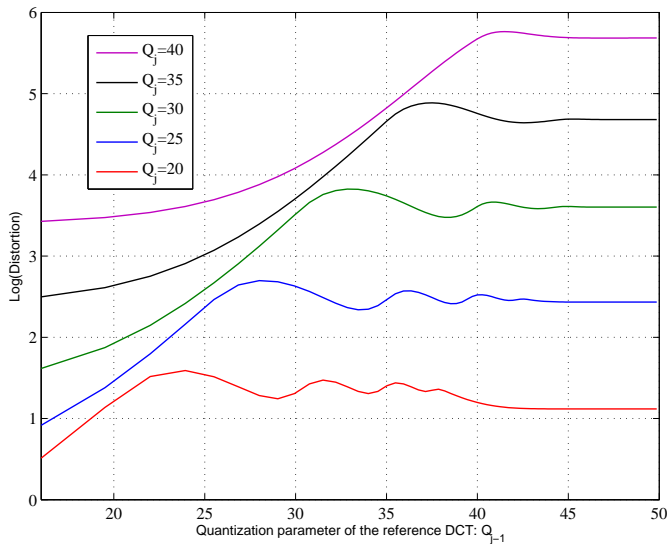


Fig. 6. Distortion dependency between current and reference DCT coefficient

Thus, the model (8), defined by three parameters is a viable alternative to the much more complicated models (6) and (7) involving seven parameters.

The estimation of the parameters of the model (8) may be done with several coding trials for different values of the pair  $(Q_{j-1}, Q_j)$ , as would be done for the models (6) and (7). Nevertheless, since the parameters of the model proposed in this section have a physical meaning, one expects to be able to perform their identification without performing several coding trials.

## VI. CONCLUSION

In this work, we propose a new rate and distortion models for the statistical multiplexing system using H.264/AVC video encoders. These models take into account the inter-frame dependencies used for performing *motion compensation*

between a reference and a predicted frame even I and P frames or P and P frames. In this paper, the dependencies specified in terms of impact of encoding parameters of a given frame over the next one is proved experimentally and interpreted theoretically by considering the quantization of an independent regime-switching autoregressive source model with varying quantization parameters. Theoretical R-D curves obtained for this source model show a behavior which is very similar to the curves obtained using H.264/AVC video. The proposed theoretical models is based on three parameters  $a_0$ ,  $b_0$ , and  $\rho$ . Further work, will be dedicated to the estimation of these parameters based on several coding trials for different values of the pair  $(Q_{j-1}, Q_j)$ .

## ACKNOWLEDGMENT

The authors are deeply grateful to V.Kumar for his helpful comments and continual encouragement.

## REFERENCES

- [1] Z. He and D. O. Wu, "Linear rate control and optimum statistical multiplexing for H.264 video broadcast," *IEEE Transactions on Multimedia*, vol. 10, no. 7, pp. 1237–1249, Nov. 2008.
- [2] N. Changuel, B. Sayadi, and M. Kieffer, "Predictive control for efficient statistical multiplexing of digital video programs," *IEEE International Packet Video Workshop*, 2009.
- [3] B. Xie and W. Zeng, "Sequence-based rate control for constant quality video," *Proc. International Conference on Image Processing*, vol. 1, pp. 1–77–1–80, 2002.
- [4] S. Milani, L. Celetto, and G. A. Mian, "A rate control algorithm for the H.264 encoder," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 18, no. 2, pp. 257–262, Feb 2008.
- [5] J-L. Lin and A. Ortega, "Bit-rate control using piecewise approximated rate-distortion characteristics," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 8, no. 4, pp. 446–459, Aug 1998.
- [6] X. Minghui, A. Vetro, and L. Huifang, S.and Bede, "Rate-distortion optimized bit allocation for error resilient video transcoding," *Circuits and Systems, 2004. ISCAS '04. Proceedings of the 2004 International Symposium on*, vol. 3, pp. III–945–8 Vol.3, May 2004.
- [7] Y. Wang, J. Sun, and W. Ma, S.and Gao, "Theoretic analysis of inter frame dependency in video coding," pp. 935–939, 2008.
- [8] ITU-T and ISO/IEC JTC 1, "Advanced video coding for generic audiovisual services," Tech. Rep., ITU-T Rec. H.264, and ISO/IEC 14496-10 AVC, nov. 2003.
- [9] R. Reininger and J. Gibson, "Distributions of the two-dimensional dct coefficients for images," *Communications, IEEE Transactions on*, vol. 31, no. 6, pp. 835–839, Jun 1983.
- [10] Stephen Smoot and Lawrence A. Rowe, "Study of dct coefficient distributions," pp. 403–411, 1996.
- [11] N. Kamaci, Y. Altunbasak, and R.M. 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, Aug. 2005.
- [12] ITU, *H.264: Advanced video coding for generic audiovisual services*, nov. 2007.
- [13] I. E. G. Richardson, "H.264/MPEG-4 part 10: Transform & quantization," Available: <http://www.vcodex.com>, 2002.
- [14] S. Ma, Wen Gao, and Yan 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, Dec. 2005.
- [15] K. Sayood, *Introduction to Data Compression, Second Edition*, Morgan Kaufmann, San Francisco, 2000.