Rho-Domain for low-complexity rate control on MPEG-4 Scalable Video Coding

Yohann Pitrey, Yann Serrand, Marie Babel, Olivier Déforges

ρ-domain for low-complexity rate control on MPEG-4 Scalable Video Coding

Yohann Pitrey   Yann Serrand   Marie Babel   Olivier Déforges

Institut d'Electronique et de Telecommunications de Rennes
INSA de Rennes, 20 av. des Buttes de Coesmes, 35043 RENNES, France.
{*ypitreyh, yserrand, mbabel, odeforge*}*@insa-rennes.fr*

Abstract

Scalable Video Coding was designed in response to the growing need for flexibility in video transmission over networks and channels. MPEG-4 Scalable Video Coding (SVC) is a recently finalized standard which introduces new coding tools such as spatial, temporal and quality scalability, to produce a layer-based scalable video stream. Additionally, inter-layer prediction allows a layer to use information from other layers as a basis for motion and texture prediction, improving the overall coding efficiency. Rate control is a capital issue in video coding, as it is designed to regulate the bitrate at the output of the encoder and keep it close to a specified constraint. Whereas rate control has been extensively studied for non-scalable video coding, only few propositions were made for scalable video coding. In this paper, we adapt an attractive rate control approach, based on a bitrate modeling framework called ρ-domain, for scalable video coding. We show that this model performs well on all spatial, temporal and quality scalabilities, and handles inter-layer prediction quite accurately. After validating the approach in MPEG-4 SVC, we use the ρ-domain model to build a simple-accurate rate control scheme. Results show that the mean frame bitrate error is below 7% on a representative set of configurations, while the impact on the complexity of the encoder is very low.

1. Introduction

In a world where video is omnipresent, with television, internet broadcasting and video-capable mobile devices, efficient video coding becomes more and more crucial. More than fifteen years of research have led to the release of the MPEG-4 AVC/H.264 standard in 2003 [12]. Designed for video communications over networks, it overpasses the compression performance of previous standards such as MPEG-2 by nearly 50%.

Although very effective when addressing some given set-up requirements, this standard is not suited to deal with heterogeneous networks and devices. Current video transmitting services have to face various decoding contexts, with various device screen sizes, computing abilities and network bandwidths. Existing solutions re-encode one specific video stream for each target set-up, causing storage and bandwidth wastage. The new MPEG-4 Scalable Video Coding (SVC) standard [10] provides more flexible video streams, which contain several layers. A base layer is first encoded, then improved using spatial, temporal and fidelity/quality enhancement layers. Spatial enhancement layers increase the frame resolution and address variable screen sizes. Temporal layers increase the number of frames per second and add smoothness to the video motion. Quality layers refine motion and texture information and increase the signal-to-noise ratio in the reconstructed video stream. The standard also introduces a new tool called inter-layer prediction, which allows a layer to use information from the lower layers to predict motion and texture information. This way the redundancy between layers is removed, and the overall coding efficiency is improved.

Rate control is a capital issue when dealing with video transmission. It is intended to control the bitrate at the output of the encoder, to fit with a given bandwidth constraint while maximizing the quality of the reconstructed video stream. The quantization parameter (QP) determines the amount of information lost during the encoding process and has a direct impact on the bitrate and the quality of the reconstructed data. The main challenge concerning rate control is to find a way to predict the output bitrate from the value of the QP. Some approaches have been proposed for different video coding standards. In [5], a simple rate control scheme is proposed for the MPEG-2 standard. This early approach is based on a simplified rate-distorsion model that does not allow fine tunning of the encoder output. In [9], the authors extend the algorithm from [5] to the H.264 standard. In particular, following the original technique, a second optimization step is performed to fur-

ther improve the rate-distorsion tradeoff. The method in [6] is more elaborate and enables better performance at the expense of a higher complexity. An interesting approach called ρ-domain was introduced in [3]. ρ-domain is a bitrate modeling framework, where ρ is defined as the percentage of zero-valued coefficients in a frame after quantization. It has been demonstrated that the relationship between ρ and the output bitrate is quite linear. This enables good bitrate prediction with very low complexity using ρ as an intermediate between the QP and the output bitrate.

Although rate control has been widely studied on single layer video, only few propositions were made for scalable video. In [1], the authors present some early work and focus on the regulation of the bitrate in the base layer. In [13], the method from [9] is adapted for spatial and temporal scalabilities only. These two approaches fail to deal with all the features provided by MPEG-4 SVC. A complete rate control approach is proposed in [8]. The method deals with the three types of scalability and handles inter-layer prediction. However, the good performance of the method is at the expense of an increase of the complexity.

In this contribution, we adapt a ρ-domain based approach for scalable video coding. We first validate the approach in the context of MPEG-4 SVC by studying the relationship between ρ and the bitrate on the three types of scalability and on inter-layer prediction. Then, this approach is used to build a simple effective rate control scheme that performs well on spatial, temporal and quality scalabilities, while being able to handle inter-layer prediction. The effectiveness of the proposed algorithm is assessed on a representative set of set-ups proposed by MPEG-4 SVC. Results show that it controls the bitrate accurately, while the complexity increase is very low.

The paper is organized as follows. Section 2 introduces the ρ-domain bitrate modeling framework. In section 3, we describe the low-complexity rate control scheme derived from the ρ-domain model. In section 4, we illustrate the validity of ρ-domain for MPEG-4 SVC and discuss the performance of the proposed rate control scheme.

2. ρ-domain based rate model

Aside from classical rate control approaches, an original rate control framework called ρ-domain has shown interesting performances on single-layer DCT-based coders [3]. We will now describe this framework and extend it to scalable video coding.

2.1. ρ-domain for single-layer video coding

Let ρ be the percentage of zero coefficients in a frame after quantization. Due to the entropy coding that ends the encoding process, this parameter is directly related to the output bitrate of the encoder [3, 2]. Whether a transform coefficient is coded as a zero or conserved after quantization depends on its position in the macroblock (MB) and the QP value. Let c_{ij}^m be the coefficient at position (i, j) in the macroblock m. This coefficient is coded as a zero if its value is below a certain dead zone threshold. Let $z(c_{ij}^m, i, j, q)$ be an indicator function defined as follows:

$$z(c_{ij}^m, i, j, q) = \left\{ \begin{array}{ll} 1 & \text{if } c_{ij}^m < t(i, j, q) \\ 0 & \text{otherwise,} \end{array} \right. \qquad (1)$$

where $t(i, j, q)$ is the value of the dead zone threshold for a coefficient at position (i, j) and QP value q [11]. From 1, the value of ρ for a frame is computed as:

$$\rho(q) = \frac{1}{M} \sum_{m,i,j} z(c_{ij}^m, i, j, q). \qquad (2)$$

where M is the total number of coefficients in the frame. An interesting point about ρ is that its relation with the bitrate can be approximated quite accurately by means of a *linear* model. This assumption has been validated in various standards such as MPEG-2, H.263 [4] and MPEG-4 AVC/H.264 [11]. In [11], this relation is such as:

$$\rho(R) = \frac{R_0 - R * (1 - \rho_0)}{R_0}, \qquad (3)$$

where R_0 and ρ_0 are two parameters of the model which can be easily evaluated by encoding the frame once with an initial value of QP. Equations 2 and 3 allow us to find a relation between the bitrate and the quantization parameter, which is crucial for the design of a rate control scheme, as we show in the next section.

2.2. Validation of the ρ-domain for scalable video coding

The linearity of the relationship between ρ and the output bitrate has been demonstrated for MPEG-4 AVC/H.264 in [11]. Although, the behaviour of ρ in the MPEG-4 SVC scalable context has never been studied. The new features such as the three types of scalability and inter-layer prediction imply some changes in the very nature of the bitstream. To be able to use a ρ-domain based approach in MPEG-4 SVC, we first need to verify that the relationship between ρ and the output bitrate is still linear.

We tested this assumption on each spatial, temporal and quality scalabilities. Each type of scalability was tested separately, first without inter-layer prediction, then with inter-layer prediction. We measured the values of ρ and output bitrate for each frame and each value of QP between 0 and 51. Figure 1 shows a brief view of the typical relationship between ρ and the bitrate. It is quite visible that the relationship is still linear. Table 1 further confirms this observation and shows the correlation coefficients between ρ and
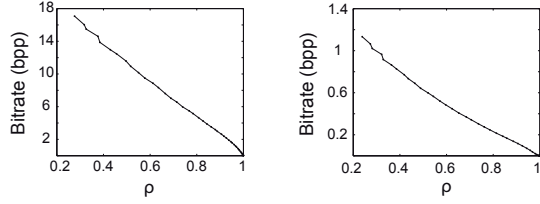
Figure 1. Relationship between ρ and the bitrate *(Left: P frame in the base layer for each value of QP. Right: spatial B frame in the highest enhancement layer with inter-layer prediction)*.

Table 1. Correlation coefficients between ρ and the bitrate.

	SPATIAL		TEMPORAL		QUALITY	
	P	B	P	B	P	B
layer 0	-0.9992	-0.9950	-0.9992	-0.9955	-0.9987	-0.9959
layer 1	-0.9960	-0.9942	-0.9964	-0.9945	-0.9988	-0.9942
layer 2	-0.9957	-0.9955	-0.9974	-0.9949	-0.9991	-0.9954

USING INTER-LAYER PREDICTION

layer 0	-0.9987	-0.9949	-0.9993	-0.9923	-0.9992	-0.9919
layer 1	-0.9964	-0.9947	-0.9987	-0.9935	-0.9977	-0.9934
layer 2	-0.9961	-0.9945	-0.9978	-0.9941	-0.9982	-0.9943

the output bitrate for P and B frames in three layers for each type of scalability. There is a strong correlation between the two variables, which allows us to assert that their relationship is quite linear.

Our tests show that the linearity of the relationship between ρ and the output bitrate is still valid in the scalable context of MPEG-4 SVC. In next section, we use this observation to build a simple rate control scheme and show the interest of a ρ-domain based approach for MPEG-4 SVC.

3. Rate control scheme for SVC

The main objective of rate control is to regulate the output bitrate of the encoder. A rate control scheme classically contains two modules. First, a bit allocation policy is used to dispatch the available bitrate among groups of pictures (GOPs), frames or macroblocks. Then, achieving the target bitrate can be accomplished by acting on the quantization parameter (QP).

In our scheme, a target bitrate is specified for each layer separately. Each frame in a layer l is allocated the same target bitrate R_t:

$$R_t = \frac{T_l}{F_l}, \qquad (4)$$

where T_l is the target bitrate for the layer l and F_l is the number of frames per second in the layer l. Once bit alloca-

tion is performed, the optimal value of QP is computed for each frame. Assuming that the model in 3 is valid, there is a univoque relation between ρ and the bitrate. Therefore, for a target bitrate R_t, there exists a unique value of ρ, say ρ_t. From equation 2, there is a one-to-one mapping between ρ and the QP. Thus, finding the QP that generates the closest bitrate to R_t is equivalent to finding the QP that generates the closest value of ρ to ρ_t. This search is performed by computing equation 2 for each value of QP between 0 and 51. Computing the value of ρ_t first requires to evaluate the parameters R_0 and ρ_0 in equation 3. To do this, we encode the frame with an initial value of QP, as mentioned in section 2.

This rate control scheme was designed to be as simple as possible, to evaluate the performances of the linear rate model itself. It can still be greatly improved by including some more advanced features such as a wise bit allocation strategy and joint rate-distorsion optimisation.

Our rate model and rate control scheme both operate at frame-level. As displayed by figure 2, reaching ρ_t would imply to consider a real QP value. Then, using only integer QP values introduces a restriction in terms of achievable bitrates. Refining the rate control at macroblock-level such as in [3] could alleviate this problem. Although, we wanted to build a low-complexity scheme, and controling the QP at macroblock level this would add large complexity and make it inapplicable, especially in a scalable video encoding context with several layers such as MPEG-4 SVC.
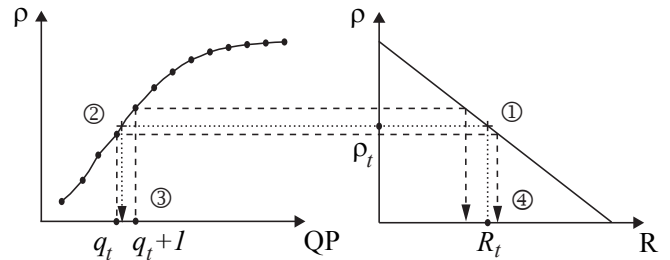

Figure 2. Frame-level rate control threshold effect. ① *the value of ρ_t is processed, which corresponds to R_t;* ② *ρ_t corresponds to a float value of QP, which lies between two integer QP values;* ③ *choose q_t as the value of QP which corresponds to the closest value of ρ from ρ_t;* ④ *the QP thresholding operation also yields a thresholding on the achievable bitrate.*

4. Experimental results

We tested the rate control scheme proposed in Section 3 for each type of scalability enabled by MPEG-4 SVC. Our tests were performed on the JSVM Reference Software 8.6.

First, inter layer prediction is disabled, while a second set of tests evaluates the impact of this coding tool on the accuracy of the ρ-domain rate model. To evaluate the performance of the scheme, we measure the error between the target bitrate for a frame and the real number of bits produced by the encoder after rate control. As stated in Section 3, our frame-level rate-control scheme only allows us to reach a close value to the specified target bitrate for each frame. From equations 2 and 3, we compute the bitrate generated by encoding a frame with the chosen QP, say q_t:

$$\hat{R}_t = \frac{R_0 * (1 - \rho(q_t))}{(1 - \rho_0)}. \tag{5}$$

Then, the bitrate error percentage δ is computed at frame level from $\hat{R}_t$ and the real achieved bitrate R_f:

$$\delta = \frac{R_f - \hat{R}_t}{\hat{R}_t} * 100. \tag{6}$$

We also calculate the error percentage between the achieved bitrate and the original bitrate constraint:

$$\Delta = \frac{R_f - R_t}{R_t} * 100. \tag{7}$$

4.1. Spatial scalability

The generated bitstreams contain three spatial layers (QCIF, CIF and 4CIF) at 30 frames per second. Only dyadic spatial scalability is enabled. GOPs contain 4 frames (*i.e.:* PBBB). Each enhancement layer uses the lower layer as a dependency ID, and all quality IDs were all set to 0. The base layer is encoded at 100 kbits per second, while the two spatial enhancement layers are encoded at 300 and 900 kbits per second, respectively.
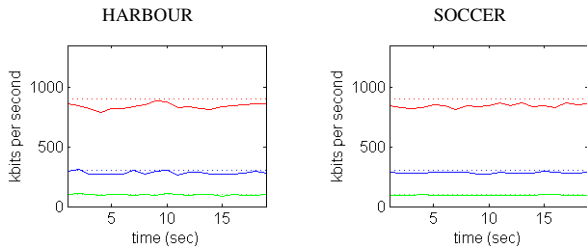

Figure 3. Bitrate per second for three spatial layers. *From bottom to top : QCIF base layer; CIF spatial layer; 4CIF spatial layer. Constraint :* $\cdots$ *, achieved bitrate :* — *.*

The bitrate per second evolution is presented in figure 3. It can be seen that the bitrate constraint is respected over the three layers with very small error. Table 2 presents the frame error between achieved and target bitrates over two

test sequences. μ_δ and μ_Δ denote the mean value of δ and Δ as defined in equations 6 and 7, while σ_δ and σ_Δ denote their standard deviation. l_0 stands for base layer, l_1 and l_2 stand for the enhancement layers.

Table 2. Target and constraint frame error for three spatial layers.

	HARBOUR			SOCCER		
	l_0	l_1	l_2	l_0	l_1	l_2
μ_δ	-0.38%	-3.96%	-5.87%	-3.44%	-4.26%	-5.67%
σ_δ	12.53	11.19	7.37	6.79	7.42	8.80
μ_Δ	-1.20%	-5.50%	-6.80%	-3.87%	-5.76%	-6.37%
σ_Δ	9.58	8.95	6.02	10.01	10.46	12.24

Figure 4 shows achieved bitrate for each frame. The variations of the frame target bitrate $\hat{R}_t$ *(dashed lines)* we observe around the constraint R_t *(dotted lines)* are due to the frame level of our rate control scheme, as stated in Section 3. The achieved bitrates are actually close to their targets, especially for P frames which have been marked with a dot on the figure.
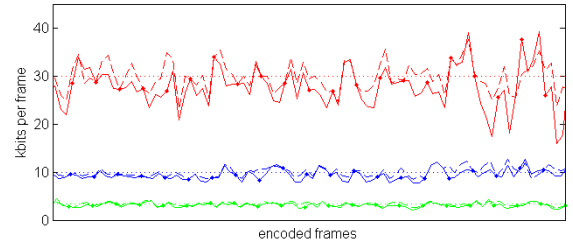

Figure 4. Frame bitrate on HARBOUR for three spatial layers. $R_t : \cdots, \hat{R}_t : \text{- - -}, R_f : \text{—}.$

4.2. Quality scalability

Quality scalability is quite similar to spatial scalability. Instead of changing the dimensions of pictures, residual information is refined. Our tests were run on three CIF layers at 30 frames per second with Coarse Grain Scalability (CGS). The same GOP structure as for spatial scalability was used. Dependency IDs are set to 0 and each enhancement layer use the previous layer as a quality ID.

Figure 5 presents the bitrate per second obtained for quality scalability. Our rate control scheme has a similar behaviour with CGS and spatial scalabilities, as confirmed by the results in table 3. Figure 6 shows achieved bitrates at frame level with quality scalability.
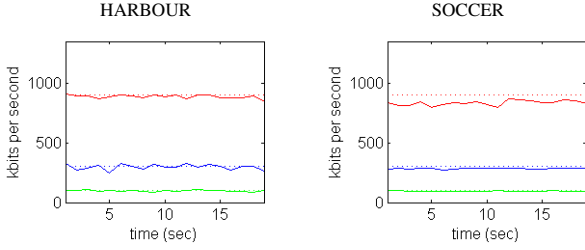
Figure 5. Bitrate per second for three CGS layers. *From bottom to top: base layer encoded at 300 kbps; CGS layer encoded at 600 kbps; CGS layer encoded at 900 kbps. Constraint : ⋯ , achieved bitrate : — .*

Table 3. Target and constraint frame error for three CGS layers.

		HARBOUR			SOCCER		
		l_0	l_1	l_2	l_0	l_1	l_2
μ_δ		-0.29%	-1.48%	-2.61%	-2.99%	-4.98%	-4.10%
σ_δ		10.74	9.07	4.30	8.52	6.01	5.17
μ_Δ		-0.90%	-0.84%	-1.00%	-4.51%	-5.02%	-4.90%
σ_Δ		8.53	10.90	11.04	10.57	9.98	10.43

Figure 6. Frame bitrate on HARBOUR for three CGS layers. $R_t : \cdots , \hat{R}_t : \text{- - -}, R_f : \text{—} .$

4.3. Temporal scalability

Temporal scalability is based on a hierarchical GOP structure. Each enhancement layer increases the number of frames per second of its base layer by two. This is performed by adding a B frame between each two frames from the base layer. Our tests on temporal scalability were performed on three CIF layers at 15 Hz, 30 Hz and 60 Hz. In the base layer, GOPs only contain P frames. In the first enhancement layer they contain one P frame and one B frame (PB—PB...), in the highest layer they contain one P frame and three B frames (PBBB—PBBB...). Results are shown in figure 7 and table 4.

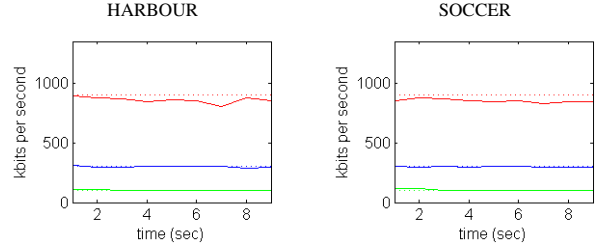On temporal scalability, our ρ-domain rate control is

Figure 7. Bitrate per second for three temporal layers. *From bottom to top : base layer at 15 fps; temporal layer at 30 fps; temporal layer at 60 fps. Constraint : ⋯ , achieved bitrate : — .*

Table 4. Target and constraint frame error for three temporal layers.

		HARBOUR			SOCCER		
		l_0	l_1	l_2	l_0	l_1	l_2
μ_δ		2.05%	0.59%	-2.98%	2.77%	0.60%	-3.99%
σ_δ		3.21	4.31	8.56	9.21	9.15	6.06
μ_Δ		-0.30%	-0.06%	-4.33%	2.27%	-0.04%	-4.40%
σ_Δ		9.48	7.74	10.14	10.57	7.93	11.13

quite accurate. The error is particularly small on the base layer and on the first enhancement layer. Figure 8 shows achieved frame bitrates for temporal scalability. It confirms the results from table 4, as the achieved bitrate is very close to the target for the two lower layers.
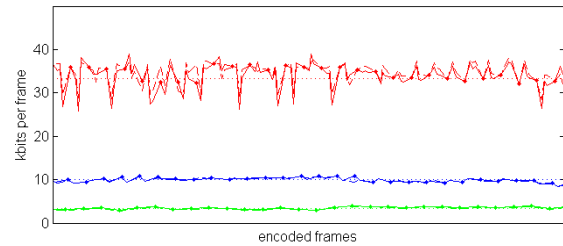

Figure 8. Frame bitrate on HARBOUR for three temp. layers. $R_t : \cdots , \hat{R}_t : \text{- - -}, R_f : \text{—} .$

4.4. Inter Layer Prediction

We reproduced the previous tests over spatial, quality and temporal scalabilities, this time using inter layer prediction. Adaptive mode was enabled, which means that the mode decision and motion estimation choose the best prediction among macroblocks from either inside the frame, other frames in the same layer or in the base layer. We used the same configurations as before : the output streams contain three layers, encoded at 100, 300 and 900 kbits per second.

Figure 9 displays the achieved bitrate per second for spatial scalability with inter layer prediction. Frame allocation error is reported in table 5, while figure 10 presents the achieved frame bitrate.
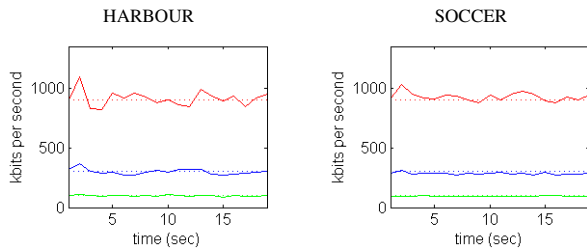

Figure 9. Bitrate per second for three spatial layers with inter layer prediction. *From bottom to top : QCIF base layer; CIF layer; 4CIF layer. Constraint : · · · , achieved bitrate : — .*

Table 5. Target and constraint frame error for three spatial layers with inter-layer prediction.

	HARBOUR			SOCCER		
	l_0	l_1	l_2	l_0	l_1	l_2
μ_δ	-0.38%	-0.03%	0.91%	-3.44%	-4.73%	-4.78%
σ_δ	8.53	13.11	14.19	6.79	9.71	11.29
μ_Δ	-1.20%	-0.21%	-4.95%	-3.87%	-4.86%	-4.96%
σ_Δ	12.58	14.52	15.99	10.01	12.30	14.63

We observe that the error variation is made higher by the use of inter layer prediction. However, the mean bitrate per second remains relatively close to the constraint for each layer. Figures 11 to 14 and tables 6 to 7 display the achieved bitrates and frame errors for quality and temporal scalabilities with inter layer prediction. Our rate control scheme shows the same behavior for these configurations.

4.5. Results interpretation

Our rate control scheme achieves a mean frame error below 7% of the allocated bitrate on each type of scalabil-
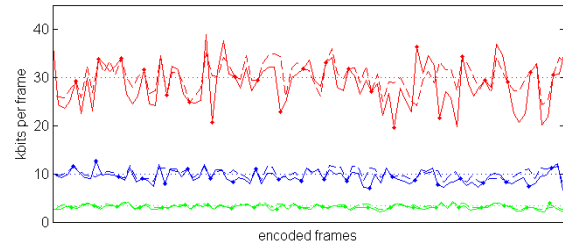

Figure 10. Frame bitrate on HARBOUR for three spatial layers with inter layer prediction. $R_t : \cdots , \hat{R}_t : \text{- - -} , R_f : \text{—} .$

Table 6. Target and constraint frame error for three CGS layers with inter layer prediction.

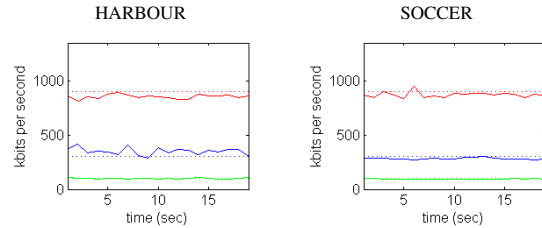|	HARBOUR			SOCCER		
	l_0	l_1	l_2	l_0	l_1	l_2
μ_δ	0.63%	-1.91%	-4.77%	-2.97%	-4.99%	-3.16%
σ_δ	15.71	14.83	6.23	8.27	11.27	5.60
μ_Δ	0.57%	4.30%	-5.23%	-4.48%	-5.40%	-3.85%
σ_Δ	16.12	16.66	13.70	10.63	13.27	9.77

Figure 11. Bitrate per second for three CGS layers with inter-layer prediction. *From bottom to top : base layer encoded at 300 kbps; CGS layer encoded at 600 kbps; CGS layer encoded at 900 kbps. Constraint : · · · , achieved bitrate : — .*

Table 7. Target and constraint frame error for three temporal layers with inter layer prediction.

	HARBOUR			SOCCER		
	l_0	l_1	l_2	l_0	l_1	l_2
μ_δ	2.05%	1.22%	1.58%	2.77%	0.08%	1.04%
σ_δ	3.21	13.31	16.17	11.21	6.85	16.69
μ_Δ	-0.30%	5.05%	2.90%	2.27%	0.88%	1.52%
σ_Δ	9.10	14.77	15.69	10.57	6.89	16.20

ity, including when inter layer prediction is enabled. This makes it quite suitable to fit with a bitrate per second constraint. We would like to report the particular effectiveness on P frames, for which the mean frame error below 3% of
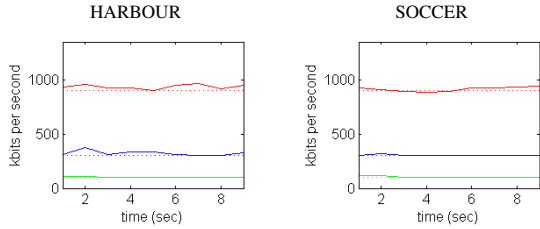
Figure 12. Bitrate per second for three temporal layers with inter-layer prediction. *From bottom to top : base layer at 15 fps; temporal layer at 30 fps; temporal layer at 60 fps. Constraint :* ⋯ *, achieved bitrate :* — *.*
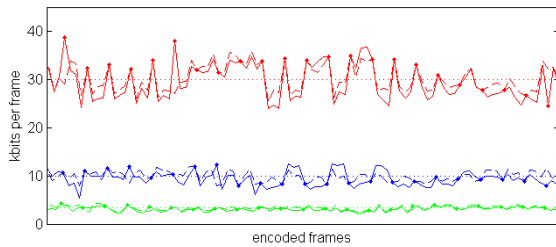

Figure 13. Frame bitrate on HARBOUR for three CGS layers with inter layer prediction. $R_t : \cdots, \hat{R}_t : \text{- - -}, R_f : \text{—}.$
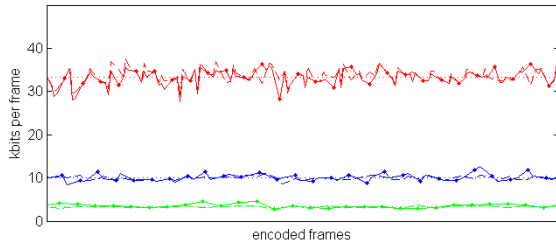

Figure 14. Frame bitrate on HARBOUR for three temporal layers with inter layer prediction. $R_t : \cdots, \hat{R}_t : \text{- - -}, R_f : \text{—}.$

the target bitrate. On B frames though, the performances are not as good, with a mean frame error around 4% to 8%. Inter layer prediction causes the error variation to increase, but the mean frame error is not significantly affected.

At frame level though, the error variation is relatively important as we can see in figures 4 and 10. This is caused by a well-known problem called the chicken and egg dilemma [7]. In MPEG-4 AVC/H.264 and SVC, the QP is involved in the choice of prediction modes and motion vectors. It
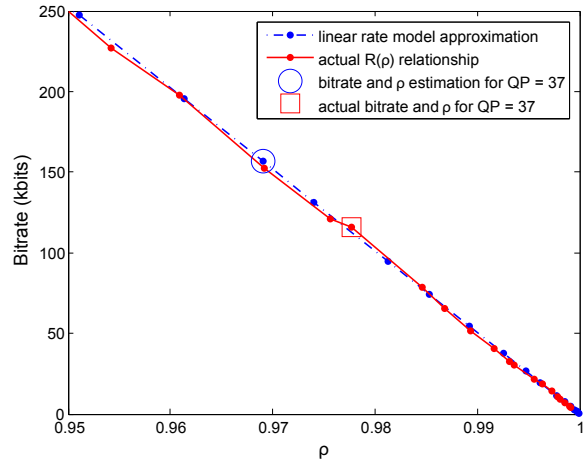

Figure 15. Linear approximation of the $R(\rho)$ **function.**

means that the contents of the residual frame will vary from one QP to another, not only because of the quantization, but also because the QP affects the motion and texture prediction. As a result, if the target QP used for the second pass is too different from the QP used for the first pass, the bitrate prediction loses its accuracy. In figure 15, we can see that the linear rate model from 3 fits with the actual $R(\rho)$ function with very small error. This further confirms that the linear ρ-domain based rate model from [11] is quite efficient on MPEG-4 SVC. Although, there is a slight shift from the ρ and rate predicted by the model to their actual values. As an example, two points were marked on the figure. The first one (inside the circle) was predicted by the linear rate model. The first pass used a QP of 41, and a target QP of 37 was chosen as the best QP to fit with the target bitrate. The second point (inside a square), corresponds to the actual ρ and rate generated after encoding the frame with a QP equal to 37. The gap between these two points show the loss of precision of the rate model when the QPs are too different. Actually, the expected bitrate for this frame was 156 kbits, whereas the final output bitrate was 117 kbits, which is about 25% wrong.

This problem justifies that our rate control scheme is less performant on B frames and with inter layer prediction. Indeed, B frames can use more prediction modes and motion vectors than P frames, which means that changing the QP in a B frame is more likely to have a great impact on the coefficients. As well, inter layer prediction provides additional prediction modes, thus augmenting the possible combinations and perturbing the efficiency of bitrate prediction.

5   Conclusion

In this paper we studied the relevance of a ρ-domain based linear rate model in MPEG-4 Scalable Video Coding. After checking that the relationship between ρ and the bitrate was remains valid for scalable video, the model was incorporated into a simple and efficient rate control scheme.

Our tests on spatial, temporal and quality scalabilities showed that ρ-domain is an interesting approach for rate control in MPEG-4 SVC. We stated that the ρ-rate model manages to fit with the actual relationship with little imprecision on small QP variations. The mean frame error was below 7% of the allocated bitrate, and even lower on P frames. We also observed that inter layer prediction had a limited influence on the performance of this model.

At the moment, our rate control scheme is very simple. Each frame is allocated the same number if bits, whereas taking into account its type (I, P or B) and residual activity can help improving the smoothness of the reconstructed stream. Wise QP initialization should also be added to alleviate the precision loss issue we observed for large QP variations.

In our future work, the presented linear rate model will be incorporated into a more elaborate rate control scheme with frame-type dependant bitrate allocation, error beedback at both frame and GOP levels and distorsion considerations. Cross layer rate control will also be widely investigated. First, we will focus on a wise bitrate dispatching among the scalability layers, then the possibilities offered by inter layer prediction will be fully studied in terms of bitrate reusability and flexibility. We will also be interested in perceptual video quality measures over scalable video coding, which are better user-felt quality indicators.

References

[1] A. M. T. A. Leontaris. Rate control reorganization in the joint model soft reference. Technical report, Joint Video Team, 2007.

[2] Z. He and T. Chen. Linear rate control for JVT video coding. *Information Technology: Research and Education, International Conf. on*, pages 65–68, 2003.

[3] Z. He and S. Mitra. ρ-domain bit allocation and rate control for real time video coding. *Image Processing, International Conf. on*, 3:546–549, 2001.

[4] Z. He and S. Mitra. Optimum bit allocation and accurate rate control for video coding via ρ-domain source modeling. *Circuits and Systems for Video Tech., IEEE Trans. on*, 12(10):840–849, Oct 2002.

[5] ISO/IEC JTC1/SC29 WG11/93-400. MPEG-2 test model 5. Technical report, MPEG Commitee, 1993.

[6] Z. Li, F. Pan, K. Lim, G. Feng, and X. Lin. Adaptive basic unit layer rate control for JVT. Technical report, Joint Video Team, 2003.

[7] Z. Li, F. Pan, K. Lim, X. Lin, and S. Rahardja. Adaptive rate control for h.264. *Image Processing, International Conf. on*, 2:745–748 Vol.2, 2004.

[8] Y. Liu, Z. G. Li, and Y. C. Soh. Rate control of h.264/AVC scalable extension. *Circuits and Systems for Video Tech., IEEE Trans. on*, 18(1):116–121, 2008.

[9] S. W. Ma, W. Gao, Y. Lu, and H. Q. Lu. Proposed draft description of rate control on JVT standard. Technical report, Joint Video Team, 2002.

[10] J. Reichel, H. Schwarz, and M. Wien. Scalable video coding - joint draft 4. Technical report, Joint Video Team, 2005.

[11] I. Shin, Y. Lee, and H. Park. Rate control using linear rate-ρ model for h.264. *Signal Processing - Image Communication*, 4:341–352, 2004.

[12] T. Wiegand, G. Sullivan, G. Bjontegaard, and A. Luthra. Overview of the h.264/avc video coding standard. *Circuits and Systems for Video Technology, IEEE Transactions on*, 13(7):560–576, July 2003.

[13] L. Xu, S. Ma, D. Zhao, and W. Gao. Rate control for scalable video model. In *Visual Communications and Image Processing.*, volume 5960, pages 525–534, 2005.