Low Complexity Rate Estimators for Low Latency Wyner-Ziv Video Decoders

Mohamed Haj Taieb, Jean-Yves Chouinard and Demin Wang

***Abstract*—Distributed video coding is a video paradigm where most of the computational complexity can be transfered from video encoders to the decoders. This allows for video sequences transmission involving inexpensive encoders and powerful centralized decoders. Unfortunately, due to the typically numerous feedback requests and needed decoders run cycles, this often leads to unacceptably long decoding latencies. One approach to addressing the latency problem consists in estimating an *initial number of parity bit chunks* (INC) that are then sent at once to reduce the number of decoders run cycles. A practical implementation challenge is to properly estimate as accurately as possible the INC, that is, without neither underestimation nor overestimation. This paper proposes two INC estimation techniques based on the temporal correlation between successive Wyner-Ziv frames and on the correlation between the different bit-planes.**

***Index Terms*—Distributed video coding, hybrid rate control, feedback channel, rate estimation**

I. Introduction

DIGITAL video coding standards are evolving to achieve high compression performances using sophisticated and increasingly complex techniques for accurate motion estimation and motion compensation. These techniques are executed at the encoder, resulting in computationally consuming video encoding tasks. The decoder, on the other hand, can easily reconstruct a video sequence by exploiting the motion vectors computed at the encoder. This *computational inbalance* is well suited for common video transfer applications such as broadcasting and video streaming, where the encoder typically benefits from high computational means to compress the video sequence only once and then to send it to many computationally limited low cost devices.

However, with the emergence of locally distributed wireless surveillance cameras, cellular interactive video utilities, and many other applications involving several low cost video encoders, at the expense of high complexity central decoders, traditional video encoding standards (e.g. H.264/AVC standard [2]) have been revised and the encoder-decoder task repartition has been reversed. Slepian and Wolf information-theoretic approach to lossless coding for correlated distributed sources [3] and its extension to lossy source coding with side information at the decoder, as introduced by Wyner and Ziv [4], constitute the theoretical framework for distributed source coding. This gave birth to a wide new field of applications, such as distributed video coding (DVC).

This work is supported by the Communications Research Centre Canada and the Natural Sciences and Engineering Research Council of Canada.

M. Haj Taieb and J.-Y. Chouinard are with the Department of Electrical and Computer Engineering, Laval University, Quebec, QC, G1V 0A8 Canada e-mail: mohamed.haj-taieb.1@ulaval.ca and jean-yves.chouinard@gel.ulaval.ca.

D. Wang is with Advanced Video Systems, Communications Research Centre Canada, Ottawa, ON, K2H 8S2 Canada e-mail: Demin.Wang@crc.ca.

Although the DVC paradigm have initiated in recent years an important body of research developments to achieve competitive R-D performances, the inherently high decoding complexity remains unacceptable for most practical DVC applications. For instance, turbo coding based DVC systems experiences unacceptably long delays caused by the several required runs of turbo decoding using parity bit chunks sent gradually upon feedback requests. Therefore, limited use of the feedback channel is crucial for the design of low latency real-time DVC applications. The DVC paradigm applies also to low density parity check (LDPC) based DVC coding schemes as they also require low computational complexity decoders.

The way the feedback channel is used by the encoder-decoder pair, highlights the trade-off between low latency and video sequence reproduction quality requirements. On one hand, the feedback channel is useful to insure decoder rate control with a minimum forward rate, but this at the price of several decoding loops. On the other hand, the encoder rate control without a feedback channel reduces drastically the system delay: in this case, the encoder needs to estimate the number of parity bits needed by the turbo decoder. If the estimated number of bits exceeds the minimum number of parity bits actually needed, this increases the bit rate while if the number of parity bits is underestimated, the turbo decoding will not converge, leading potentially to visual artifacts in the reconstructed frames.

Between these two rate control schemes, a hybrid (trade-off) technique can be adopted where the encoder and the decoder cooperate to estimate the minimal rate using the feedback channel. This article is a revised and extended version of conference paper [1] in which two INC estimation methods are introduced. In section II, a review of the Discover DVC architecture is presented. Different rate control mechanisms are described in section III. In section IV, the proposed algorithms are proposed to provide accurate estimates of the INC. The first one is a hybrid rate control technique based on the temporal correlation between the *final number of parity bit chunks* (FNC). The second hybrid rate control technique exploits the correlation at the bit-plane levels. A comparative study, supported by simulations on test video sequences, between the different estimators is presented in section V.

II. Discover DVC codec architecture

Figure 1 illustrates the architecture of the Discover DVC Wyner-Ziv (WZ) system. As shown in this figure, the WZ encoder-decoder pair is based on turbo coding [5]. The Discover DVC codec architecture is based on the Stanford WZ codec and includes several means for improving the rate distortion performance. The key frames are H.264/AVC

(Advance online publication: 9 February 2013)

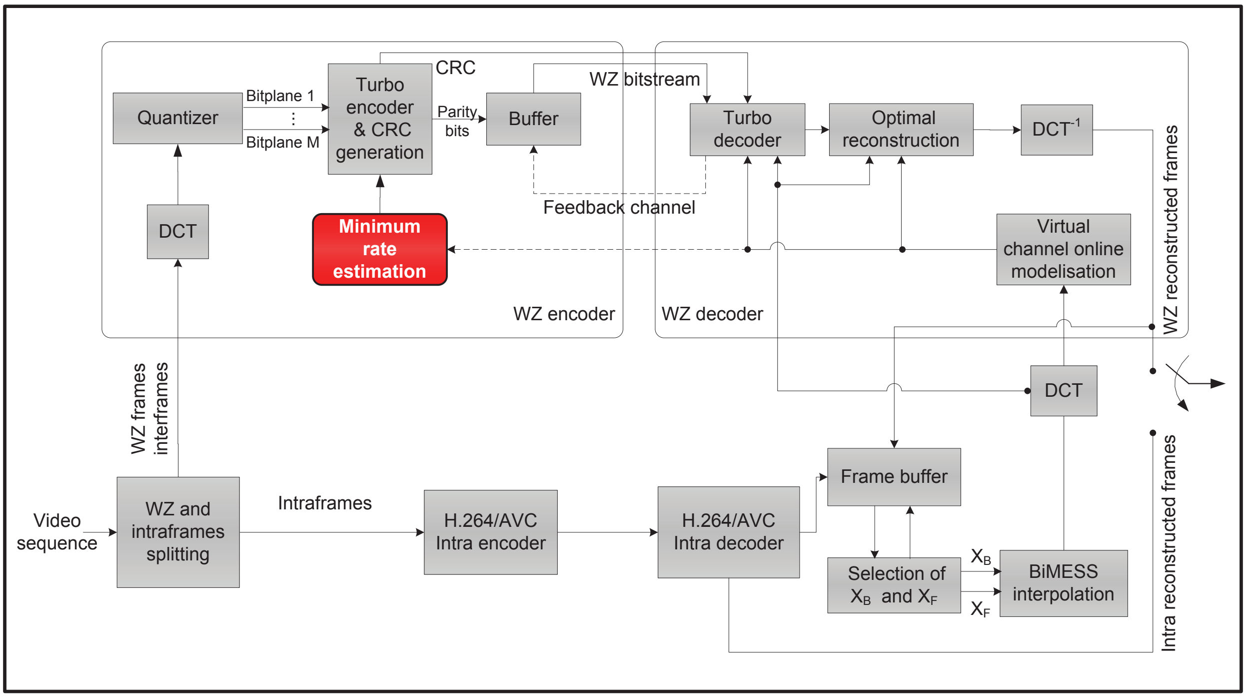

Fig. 1. Transform domain Wyner-Ziv video codec (Discover DVC codec architecture).

intra encoded (refered to as *intraframes*) and transmitted to generate the side information to decode the Wyner-Ziv frames (*interframes*). The interframes are compressed using a 4x4 block discrete cosine transform (DCT). The DCT coefficients are fed to a uniform quantizer. The quantized coefficients are then fed to a turbo encoder consisting of two constituent rate 1/2 recursive convolutional encoders (RSCs). Each RSC associates a parity bit to the quantized DCT-coefficients. To achieve compression of the transmitted data, the systematic bits are discarded since the decoder has already an interpolated version of the even frames (i.e. the intraframes). Moreover, the parity bits are stored in a buffer and sent gradually, packet by packet, upon decoder feedback requests according to a periodic puncturing pattern. The feedback channel allows for adapting the forward transmission rate to the changing virtual channel conditions. This also implies several turbo decoder runs. To alleviate the decoder computational hurdle, an initial number of parity bit packets is estimated by an hybrid encoder/decoder rate control mechanism [6]. These parity bit packets are sent once to the decoder and eventually subsequent packets will be sent if needed.

At the decoder, an interpolated version of the current WZ frame is produced using the already received neighboring key frames. The motion compensated temporal interpolation technique (MCTI) presented in [7], known as *bidirectional motion estimation with spatial smoothing* (BiMESS), was adopted for most DVC architectures. The BiMESS performances are improved using a hierarchical coarse-to-fine approach in bidirectional motion estimation [8] and sub-pixel precision for motion search [9]. The interpolated frame is then DCT transformed and the DCT coefficients represent the side information used to decode the WZ frames. The WZ DCT coefficients are modeled as the input of a virtual channel and the side information as its output. During the turbo decoding process, a Laplacian model is assumed for this virtual channel. The estimation of Laplacian distribution parameter α is based on an online correlation noise modeling technique at the coefficient/frame level: parameter α is estimated for each coefficient band of each frame [7].

The turbo decoder computes the systematic log-likelihood ratios. The systematic information is corrupted by a Laplacian noise whose parameter is, beforehand, online estimated (without using original data). Actually, there are no systematic bits and the side information is used instead. The received parity bits along with the side information are fed to the turbo decoder. After a number of iterations, the log-likelihood ratios are computed and then the bitplane value is deducted. To estimate the decoded bitplane error rate, without access to the original data, these log-likelihood ratios are used to compute a confidence score [6]. If this score exceeds 10^{-3}, then a parity bits request is sent back to the encoder. Otherwise, the decoding process is likely to be satisfactory. However, some errors can still persist even if the confidence score is below 10^{-3}. For this reason, an 8-bit long cyclic redundancy check (8-CRC) code is used to help detecting the remaining bitplane decoding errors. If the decoded bitplane CRC corresponds to the original data CRC, then the decoding process is considered successful, otherwise, more parity bits are requested. Using jointly the confidence score and the CRC code results in error detection performances as good as ideal error detection where the decoded bitplane is directly compared to the original bitplane [6].

After being decoded, these bitplanes are recombined to form the quantized symbols. These symbols and the side information are used to reconstruct the DCT coefficients. An optimal reconstruction function is proposed in [10] to minimize the mean squared error according to the Laplacian correlation model. For coefficients bands that have not been transmitted, the side information is directly considered in the reconstruction. Finally, an inverse $4{\times}4$ DCT is applied to the reconstructed frequency band to restore the WZ frame in the pixel domain.

III. Conventional rate control mechanisms

The purpose of the minimum rate estimation component in the WZ encoder, as shown in figure 1, is to estimate the required number of parity bits chunks needed for the convergence of the turbo decoding process. This initial number of chunks (INC) is sent at once to the decoder and, eventually, additional parity chunks will progressively be sent later until turbo decoding convergence as shown in figure 2. Two undesired situations may arise:

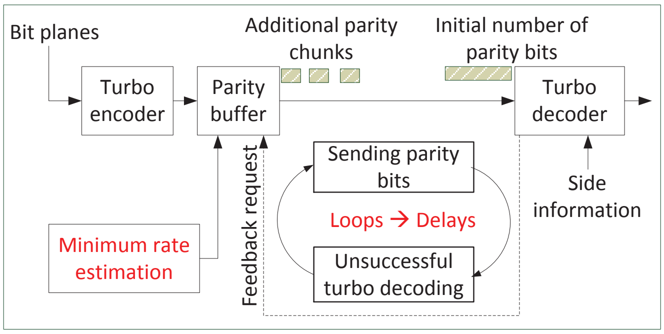

Fig. 2. Rate control in distributed video coding.

1) Underestimation: If the decoder does not converge using the initial number of parity bits chunks, a feedback request for an additional parity bits chunk is sent to the encoder. For each received chunk, the turbo decoder is launched again. Requests-decoding loops are performed until turbo decoding convergence: this may cause long delays at the decoder.
2) Overestimation: If the estimated initial number of parity bits chunks (INC) exceeds the actual needed number, the turbo decoder will converge with no further feedback request. Thus there are no decoding delays: however, the rate-distortion performance will be affected.

These two undesirable situations need to be avoided as much as possible to allow for an efficient rate-distortion performance DVC codec with low latency. This is possible by accurately estimating the needed number of parity bit chunks.

A. Decoder rate control (DRC)

Decoder rate control was adopted for the first DVC implementation [11], because it resulted in the best rate-distortion performances. Excessive execution delays were experienced at the decoder as the technique did not estimate an initial number of parity bit chunks (INC) and involved sending these chunks until the decoder converged. However, there were no overestimation hence leading to the best performances.

In the following, two hybrid encoder/decoder rate control methods to estimate the minimum rate R_{min} (or the INC) are described: the first one is based on the Slepian-Wolf correlated sources coding theorem while the second one is a low complexity method to estimate the minimum rate.

B. Hybrid rate control based on the Slepian-Wolf theorem

Kubasov, Lajnef and Guillemot [6] have proposed a hybrid rate control technique for evaluating the minimum parity rate R_{min} for each bitplane of each DCT band. The decoder must estimate the *correlation noise* between WZ interframe DCT samples and the corresponding interpolated samples from the H.264/AVC encoded intraframes (i.e. the side information). The correlation noise is modeled as a Laplacian process:

$$f_N(n) = \frac{\alpha}{2} e^{-\alpha|n|} \tag{1}$$

The decoder estimates the correlation noise (Laplacian) model parameter α and sends it back to the encoder. Knowing the original data and the Laplacian model parameter α, the encoder first estimates the probability of crossover p_{co} and then the minimal rate R_{min} according to the Slepian-Wolf theorem [6]:

$$\begin{aligned} R_{\min} = H(X|Y) &= -p_{co}\log_2 p_{co} \\ &\quad -(1-p_{co})\log_2(1-p_{co}) \end{aligned} \tag{2}$$

The crossover probability is estimated for each bitplane and corresponds to the probability that bitplane x_{pb} is different from the estimated bitplane at the decoder, $\hat{x}_{pb}$, using the side information y and the previously decoded bitplanes $(x_{pb-1},\ \ldots, x_2,\ x_1)$:

$$\hat{x}_{pb} = \arg\max_{i=0,1} \Pr\left(x_{pb} = i\,|y,\ x_{pb-1},\ \ldots, x_2,\ x_1\right) \tag{3}$$

where $\Pr\left(x_{pb} = i\,|y,\ x_{pb-1},\ \ldots, x_2,\ x_1\right)$ designates the *a posteriori* probability of event $x_{pb} = i$.

An example of $\hat{x}_{pb}$ calculation is depicted in figure 3. After determining $\hat{x}_{pb} = 1$, the crossover probability $\Pr\left(x_{pb} \neq \hat{x}_{pb}\right)$ is computed as:

$$\Pr\left(x_2 \neq \hat{x}_2\right) = \frac{\int_{y-B_1}^{y-B_2} f_N(n)dn}{\int_{y-B_1}^{y-B_3} f_N(n)dn} = \frac{F_N(y-B_2) - F_N(y-B_1)}{F_N(y-B_3) - F_N(y-B_1)} \tag{4}$$

where $F_N(n)$ is the cumulative distribution function (CDF) of the Laplacian probability density function (PDF)[1]:

$$F_N(n) = 0.5\left(1 + \text{sign}(n) - \text{sign}(n)\,e^{-\alpha|n|}\right) \tag{5}$$

The crossover probability p_{co} is computed at the WZ encoder which has no knowledge about side information y: thus the next step consists in integrating over all possible values of y, that is over the range $[V_{\min}, V_{\max}]$. Finally the average over the original WZ DCT coefficients is taken:

$$p_{co} = \frac{1}{N}\sum_{x\in WZ}\left[\int_{V_{\min}}^{V_{\max}} \Pr\left(x_{pb} \neq \hat{x}_{pb}\right)\frac{\alpha}{2}e^{-\alpha|y-x|}dy\right] \tag{6}$$

where N is bitplane length. Thereby, the computation of the crossover probability, p_{co}, requires averaging N relatively complex integrals. This involves considerable computations at the encoder (supposed to be light in the DVC paradigm). For each DCT band, the p_{co} computation is thus given by :

$$p_{co} = \frac{1}{N}\sum_{x\in WZ}\left[\int_{V_{\min}}^{V_{\max}} \frac{F_N(y-B_2) - F_N(y-B_1)}{F_N(y-B_3) - F_N(y-B_1)}\frac{\alpha}{2}e^{-\alpha|y-x|}dy\right] \tag{7}$$

[1]The use of the CDF avoids the need to perform integration.

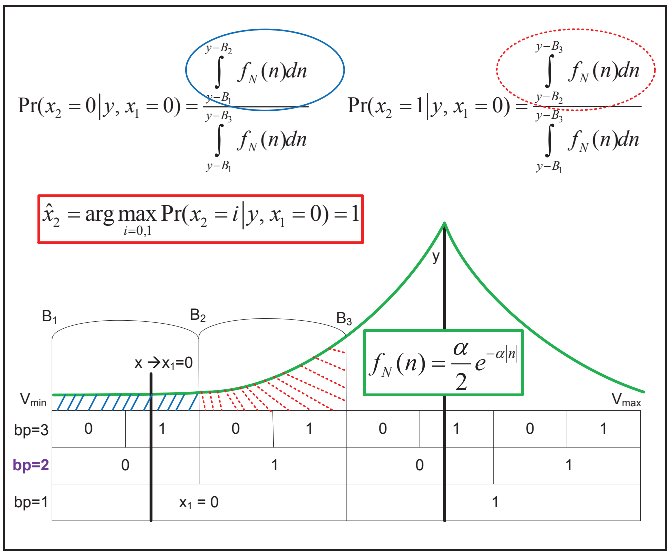

Fig. 3. *Example of the computation of $\hat{x}_{pb=2}$ as a function of conditional probabilities for a given side information y and the previously decoded bitplane $x_{pb=1}$.*

C. Low complexity hybrid rate control

The previous technique incurs some additional encoder complexity to estimate the minimum rate. In [12], Areia, Ascenso, Brites and Pereira proposed a low complexity hybrid rate control technique. For each bit-plane j, of band i, the initial number of parity bits chunks (INC) is estimated using the final number of parity bits chunks (FNC) sent for the same bit-planes in the previous 3 WZ frames:

$$INC\,(i,j) = \text{floor}\left[(1-k)\times \text{median}\left(FNC_{-1}\,(i,j)\,,\; FNC_{-2}\,(i,j)\,,\; FNC_{-3}\,(i,j)\right)\right] \tag{8}$$

where k is a scale factor such as $k = 0.1$ for the first five DCT bands ($i = 1, \ldots, 5$) and $k = 0.05$ for the remaining bands. The term $(1-k)$ prevent from estimation rate saturation.

IV. Proposed Algorithms for Initial Number of Chunks (INC) Estimation

In this section, we propose two minimum rate estimation techniques. These techniques are based on the temporal evolution of the *final number of parity bits chunks* (FNC) for each bit-plane of each DCT band.

A. Estimation algorithm based on temporal correlation (TC)

This algorithm exploits the FNC's temporal stationarity to perform a two-step estimation of the minimum rate. More specifically, consider the estimation of the INC for the sixth bit-plane of the first DCT band of the WZ frame number $t = 36$ as shown in figure 4. This figure displays the temporal evolution of the FNC determined by decoder rate control. The first step consists in computing the INC using the three previous FNC values:

$$INC^{(\text{step } 1)}_{t=36}\,(\text{band}=1,\text{bp}=6) = a\times FNC_{35}\,(1,6) + a^2\times FNC_{34}\,(1,6) + a^3\times FNC_{33}\,(1,6) \tag{9}$$

where a is a scaling factor such that $a + a^2 + a^3 = 1 \Rightarrow a = 0.54$ if there is under-estimation at the previous frame ($t = 35$) and such that $a + a^2 + a^3 = 0.8 \Rightarrow a = 0.47$ if there is over-estimation in the first band sixth bit-plane (DC band) of the WZ frame number $t = 35$. In this manner, the estimated number INC does not grow indefinitely when there is overestimation at some point, thus avoiding saturation. In this case, the turbo decoder converges in one run after receiving the initial number of parity chunks and no feedback request are being sent. Here, two possible cases can actually occur. The first case is actually an overestimation which occur when the estimated INC exceeds the needed number found with the decoder rate control mechanism ($DRC_{35}(1,6)$):

$$INC_{35}\,(1,6) > DRC_{35}(1,6) \tag{10}$$

The second case happens when the estimation is perfect:

$$INC_{35}\,(1,6) = DRC_{35}(1,6) \tag{11}$$

It is not possible to distinguish between these two situations since, in both cases, the turbo decoder converges at the first iteration. In the remainder of this paper, we will refer to these two cases as overestimation.

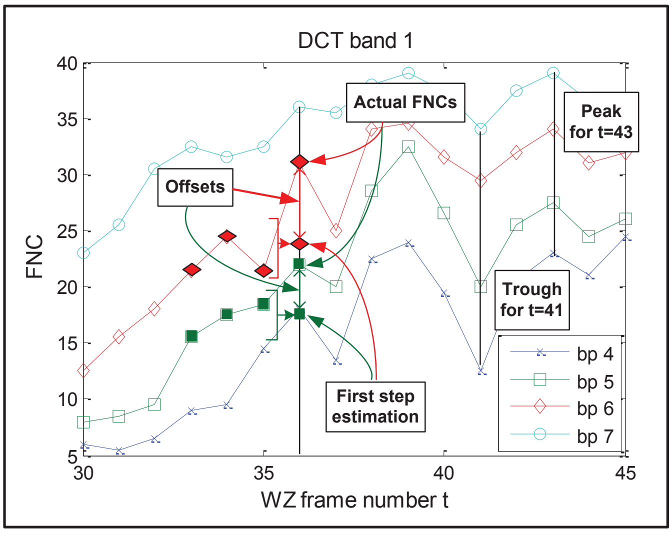

Fig. 4. *FNC temporal variation with peaks and troughs occuring for the different bit-planes. The offset between the first step estimation for bp=5 can be used to adjust the estimation for bp=6.*

The first step estimation calculates a weighted average between the previous FNCs values. However, when there is a peak or a trough, this estimation is not close enough to the actual value. This (a peak or a trough) can be detected by observing the previous bit-plane. For instance, it can be observed from the bit-plane $bp = 5$, that there is a peak, and an offset can be computed between the first step estimated value and the actual value. This offset is expected to occur again for the bit-plane $bp = 6$. Thus the first step estimation can be adjusted as follows:

$$INC_{t=36}^{(\text{step } 2)}(1,6) = INC_{t=36}^{(\text{step } 1)}(1,6) + \text{offset}(bp = 5) \tag{12}$$

B. Estimation algorithm based on bit-plane correlation (BP)

Figure 5 shows the temporal evolution of the FNC offsets. This figure shows that the offset between the FNCs of two successive bit-planes is almost the same at times (frames) t and $t+1$:

$$FNC_t(i,j) - FNC_t(i,j-1) \approx FNC_{t+1}(i,j) - FNC_{t+1}(i,j-1) \tag{13}$$

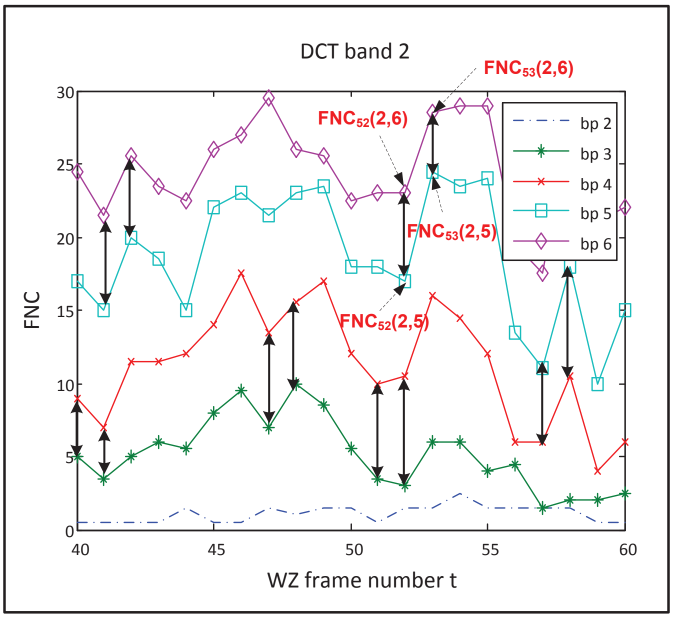

Fig. 5. *FNC offset between two successive bit-planes.*

This observation is used to compute an INC estimation for a bit-plane bp based on the FNC of the previous bit-plane $bp-1$. For instance, the estimation of the INC for the second DCT band sixth bit-plane for the WZ frame $t = 53$, is:

$$INC_{53}(2,6) = FNC_{53}(2,5) + [FNC_{52}(2,6) - FNC_{52}(2,5)] \tag{14}$$

Expression (14) is applied when there is no overestimation in $FNC_{52}(2,6)$ nor in $FNC_{53}(2,5)$. The overestimation for these two bit-planes can be detected during their respective decoding. In fact, if $FNC_{52}(2,6) > INC_{52}(2,6)$ then there is no overestimation and, in this case, $FNC_{52}(2,6) = DRC_{52}(2,6)$. Recall that $DRC_{52}(2,6)$ is the target number of chunks as computed with the decoder rate control technique. If there is an overestimation in one of these two bit-planes, then expression (14) can no longer be applied: otherwise, there is a risk of accumulation of overestimations leading to *saturation*. In this case, (i.e. overestimation at the sixth bit-plane of the WZ frame number $t = 52$ xor[2], overestimation at the fifth bit-plane of the WZ frame number $t = 53$), equation (14) becomes:

$$INC_{53}(2,6) = FNC_{53}(2,5) + a \times [FNC_{52}(2,6) - FNC_{52}(2,5)] \tag{15}$$

[2]xor: *exclusive or*.

where a is a scale factor, set empirically to 0.8. When there is overestimation in both these two bit-planes (i.e. overestimation at the sixth bit-plane of the WZ frame number $t = 52$ and overestimation at the fifth bit-plane of the WZ frame number $t = 53$) the overestimation situation is more likely to happen again in the current bitplane and more caution are taking by changing the scale factor a to a^2. Then, Equation (14) becomes:

$$INC_{53}(2,6) = FNC_{53}(2,5) + a^2 \times [FNC_{52}(2,6) - FNC_{52}(2,5)] \tag{16}$$

Thus, the initial number of parity chunks estimation algorithm based on bit-plane correlation is summarized as follows:

$$\begin{aligned}
&\text{Case 1}: FNC_{52}(2,6) > INC_{52}(2,6) \\
&\quad \text{and } FNC_{53}(2,5) > INC_{53}(2,5) \\
&\text{then} \\
&INC_{53}(2,6) = FNC_{53}(2,5) + \\
&\qquad [FNC_{52}(2,6) - FNC_{52}(2,5)] \\
&\text{Case 2}: FNC_{52}(2,6) = INC_{52}(2,6) \\
&\quad \text{xor } FNC_{53}(2,5) = INC_{53}(2,5) \\
&\text{then} \\
&INC_{53}(2,6) = FNC_{53}(2,5) + \\
&\qquad a \times [FNC_{52}(2,6) - FNC_{52}(2,5)] \\
&\text{Case 3}: FNC_{52}(2,6) = INC_{52}(2,6) \\
&\quad \text{and } FNC_{53}(2,5) = INC_{53}(2,5) \\
&\text{then} \\
&INC_{53}(2,6) = FNC_{53}(2,5) + \\
&\qquad a^2 \times [FNC_{52}(2,6) - FNC_{52}(2,5)]
\end{aligned} \tag{17}$$

Notice here that overestimation is reported when the turbo decoder converges directly after receiving the initial number of chunks INC without any feedback request. Thus, the final number of chunks is equal to the initial number of chunks:

$$FNC_t(\text{b}, \text{bp}) = INC_t(\text{b}, \text{bp})$$

.

V. Simulations and Discussion

In this section, the proposed estimators EST: TC (temporal correlation estimator) and BP (bit-plane correlation estimator) are compared with the correlation estimator used by Kubasov *et al.* [6] and the estimator of Areia *et al.* [12]. Three QCIF video sequences at 15 frames per second are considered for the simulation tests: `Foreman`, `Soccer` and `Coastguard`. These sequences are downloaded from the Discover website [5]. All 149 frames of the sequences are considered, corresponding to 74 WZ frames. The frame size is 144 $\times$ 176 =25344 pixels, leading to bit-planes length of 25344/16=1584 bits for each DCT 4$\times$4 component. The puncturing period length is 48 which results in a chunk size of a $(1584/48) \times 2 = 33 \times 2 = 66$ parity bits sent at each feedback request. This corresponds to 33 parity bits for each of the two recursive systematic convolutional (RSC) encoders. The estimated initial number of chunks (INC) involves sending INC $\times$66 parity bits at once.

A. Estimators comparison criteria

To compare between the estimators' performances, three points have to be considered:

1) Overestimation: it engenders a rate increase. The average number of chunks, over the N WZ frames, sent in excess is given by:

$$\text{Excess} = \frac{\sum_{n=1}^{N} \max\left(\left[INC^{EST}(n) - FNC^{DRC}(n)\right], 0\right)}{N} \quad (18)$$

2) Underestimation: when the INC is below the FNC, the decoder will ask gradually for more parity bits chunks. For each feedback request, the turbo decoding will be launched again, thus causing delays. The average number of feedback requests over the N WZ frames is given by:

$$\text{Request} = \frac{\sum_{n=1}^{N} \max\left(\left[FNC^{DRC}(n) - INC^{EST}(n)\right], 0\right)}{N} \quad (19)$$

3) Accuracy: To assess the accuracy of the estimator as a whole, taking into account both overestimation and underestimation, the *average absolute difference* (estimation error) between the INC and the FNC obtained with decoder rate control is evaluated as:

$$\text{Difference} = \frac{\sum_{n=1}^{N} \left|FNC^{DRC}(n) - INC^{EST}(n)\right|}{N} \quad (20)$$

B. Comparison of the minimum rate estimation algorithms

Prior to comparing the proposed estimators performances from a global point of view, figure 6 shows the temporal behaviour of the estimated initial number of chunks (INC) as well as final number of chunks of the decoder rate control solution DRC for some selected bit-planes. The DRC indicates the target number of chunks to be estimated. According to figure 6 the following remarks are made:

- Areia estimator [12]: This estimator is based on the simple calculation of the weighted median of the 3 previous frames. As can be observed for the `Coastguard` video sequence, it is not able to follow with a sudden variations in the needed number of parity bits chunks. This phenomenon is analoguous to the *slope-overload quantization distortion* observed in the *delta modulation*.
- TC and BP proposed estimators: these two estimators can provide more accurate estimates than the algorithm of Areia *et. al* [12] since they are able to follow more accurately the rapid variations of the DRC. In fact these techniques can adjust the estimation process according to the previously decoded bit-planes (of the same WZ frame or of the previous WZ frames).
- The TC-based estimator benefits from the FNC temporal structure (see equation (9)) as well as from the quasi-stationnarity of the offsets from bit-plane bp to bit-plane $bp+1$ (see (12)).
- The BP estimator is based on the FNC temporal quasi-stationnarity of the gaps between two successive bit-planes of the same WZ frame.
- These two empirical observation based techniques give good estimation and can be adjusted using the forgetting parameter a (see equations (9) and (17)). In fact, when the parameter a decreases, the overestimation occurence is reduced, thus leading to better rate-distortion performances. When parameter a increases, underestimation occurs less frequently, leading to a faster decoding.

A comparison of the different estimators is presented in figure 7 using the `Foreman`, `Soccer` and `Coastguard` video test sequences. The three previously cited criteria, that is the number of feedback requests to the WZ encoder per video frame, the excess number of parity bits chunks per video frame, and the average absolute difference are computed to evaluate the relative performance of the estimators. This figure also shows the *peak signal-to-noise ratio* (PSNR) as a function of the overall bitrate for the same test video test sequences and the same 4 estimators, along with the DRC estimator (without estimation of the initial number of parity bit chunks) for comparison purposes. Eight quantization matrix indices (Qi) are considered. The estimator absolute difference criterion stipulated that the proposed solutions (TC and BP) are more accurate and give closer estimates to the number of chunks required for the decoding convergence. The proposed rate estimation solutions are more accurate and the rate-distortion curves displays a reasonable rate increase caused by overestimation.

The estimator performances are then summarized in tables I and II. Table I provides the decoder complexity reduction percentage compared to the DRC method. The second table presents the percentage of the rate increase again by comparison with the DRC method. These tables shows that the proposed estimators can reduce significantly the decoder latencies (an average reduction of 87.5 % for the TC solution and 88.29% for the BP solution) without a severe impact on the rate-distorsion performances (only 8.93% and 9.37% rate increase, respectively).

TABLE I
Decoder complexity reduction percentage relative to decoder rate control (DRC) method.

	Kubasov	Areia	BP	TC
`Foreman`	53.22%	84.59%	88.02%	87.07%
`Soccer`	55.68%	86.74%	90.21%	89.00%
`Coastguard`	64.28%	84.75%	86.65%	86.43%
Average	57.72%	85.36%	88.29%	87.50%

TABLE II
Rate increase percentage caused by over-estimation compared to decoder rate control method.

	Kubasov	Areia	BP	TC
`Foreman`	0.65%	17.09%	10.88%	10.46%
`Soccer`	0.53%	8.51%	7.92%	7.31%
`Coastguard`	1.04%	15.24%	9.33%	9.03%
Average	0.74%	13.61%	9.37%	8.93%

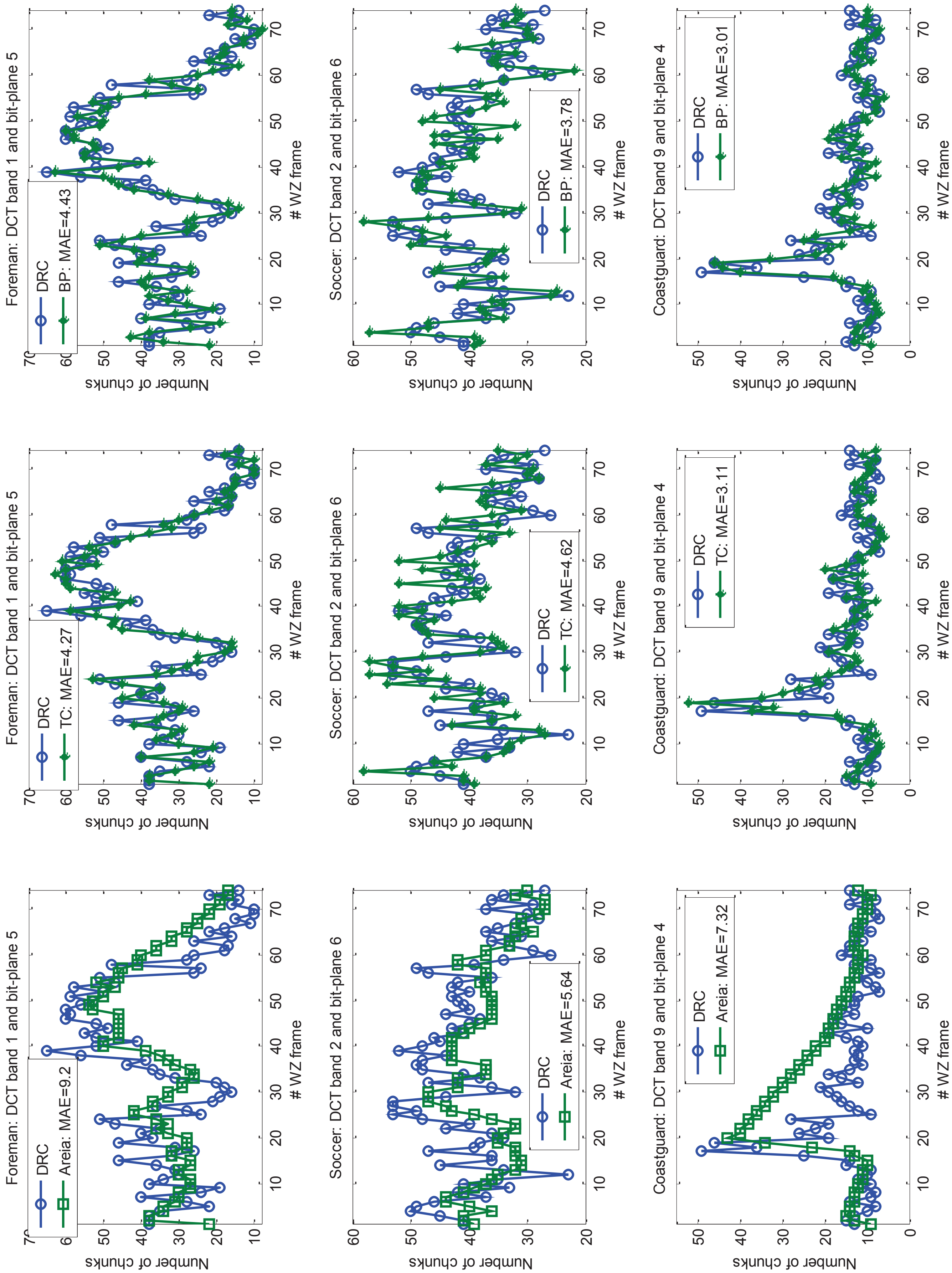

Fig. 6. *Temporal evolution of the estimated INC and the target number of chunks DRC for some bit-planes of the quantization index $Qi = 8$. The performance of the different estimators is compared by computing the mean absolute error (MAE) over the 74 WZ frames of the* `Foreman`, `Soccer` *and* `Coastguard` *video sequences.*

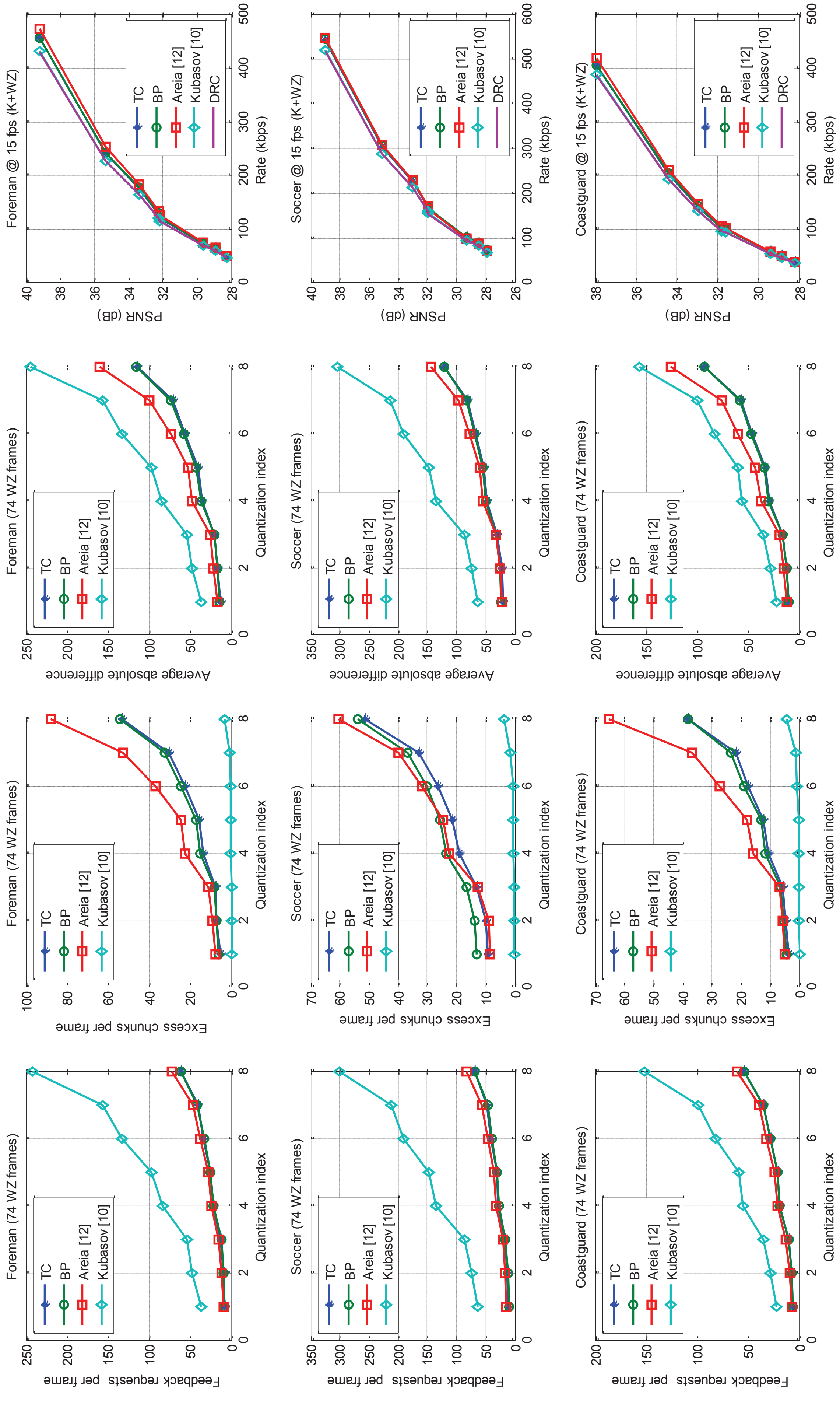

Fig. 7. *Comparison between the different R_{min} estimators performances.*

VI. Conclusion and future work

In this paper, new techniques for low complexity rate control are proposed for low latency Wyner-Ziv video decoders. These mehtods are inspired from the observed temporal behavior of the FNC which displays not only a temporal quasi stationnarity between successive WZ frames but also a correlation between successive bit-planes. More precise estimation, allowing lower decoding delays, are obtained thanks to these techniques, and this at the expense of only a slight overall bit rate increase. These techniques depend strongly on the hypotheses of temporal correlation and the structure of the FNC. However, if in some instances, these hypotheses are not verified, then the estimation can be severely compromised. As future investigations, the rate estimation can be improved by a down-sampled version of the WZ frames being sent and then decoded at the first iteration: from the obtained FNC, the INC for the remaining WZ frame can be estimated.

References

[1] Mohamed Haj Taieb, Jean-Yves Chouinard and Demin Wang, "Low Complexity Hybrid Rate Control Schemes for Distributed Video Coding," in *Lecture Notes in Engineering and Computer Science: Proceedings of The World Congress on Engineering and Computer Science 2012*, 24-26 October 2012, San Francisco, USA, pp. 589–594.

[2] I. E. G. Richardson, *H.264 and MPEG-4 Video Compression: Video Coding for Next Generation Multimedia*. WileyInterscience, 2003.

[3] J. Slepian and J. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. Inf. Theory*, vol. IT-19, no. 4, pp. 471–480, July 1973.

[4] A. D. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," *IEEE Trans. Inf. Theory*, vol. IT-22, no. 1, pp. 1–10, January 1976.

[5] "The DISCOVER codec evaluation," [Online]. Available: http://www.discoverdvc.org/.

[6] K. L. D. Kubasov and C. Guillemot, "A Hybrid Encoder/Decoder Rate Control for Wyner-Ziv Video Coding with a Feedback Channel," in *Proceedings of Int. Workshop on Multimedia Signal Processing*, October 2007, pp. 251–254.

[7] C. Brites and F. Pereira, "Correlation Noise Modeling for Efficient Pixel and Transform Domain Wyner Ziv Video Coding," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 53, no. 2, September 2008.

[8] F. P. J. Ascenso, C. Brites, "Content adaptive Wyner-Ziv video coding driven by motion activity," in *IEEE International Conference on Image Processing*, October 2006.

[9] S. Klomp, Y. Vatis and J. Ostermann, "Side Information Interpolation with Sub-pel Motion Compensation for Wyner-Ziv Decoder," *International Conference on Signal Processing and Multimedia Applications (SIGMAP)*, August 2006.

[10] D. Kubasov, J. Nayak and C. Guillemot, "Optimal Reconstruction in Wyner-Ziv Video Coding with Multiple Side Information," in *Proc. of MMSP, IEEE International Workshop on Multimedia Signal Processing*, October 2007.

[11] B. Girod, A. Aaron, S. Rane, and D. R. Monedero, "Distributed video coding," in *Proceedings IEEE, Special Issue on Advances in Video Coding and Delivery*, vol. 93, no. 1, January 2005, pp. 71–83.

[12] J. D. Areia, J. Ascenso, C. Brites and F. Pereira, "Low Complexity Hybrid Rate Control for Lower Complexity Wyner-Ziv Video Decoding," in *16th European Signal Processing Conference (EUSIPCO)*, August 2008.