PROBABILITY UPDATING FOR DECODER AND ENCODER RATE CONTROL TURBO BASED WYNER-ZIV VIDEO CODING

Catarina Brites, Fernando Pereira

Instituto Superior Técnico – Instituto de Telecomunicações
{catarina.brites, fp}@lx.it.pt

ABSTRACT

In Wyner-Ziv video coding (WZVC), powerful error correcting codes must be used to achieve high compression efficiency; turbo codes are the most commonly used error correcting codes in WZVC. To improve the turbo coding performance in the context of WZVC, this paper proposes a probability updating technique (PUT) acting as an outer loop of the common turbo decoding operation. Whenever a turbo decoded bitplane is not error-free, the proposed technique attempts to correct bitplane errors by updating the correlation noise probabilities for the most likely in error bits, followed by turbo re-decoding. The new tool is evaluated both in the context of encoder rate control (ERC) and decoder rate control (DRC) turbo based WZVC scenarios with average overall PSNR gains up to about 0.5 dB in ERC and average WZ rate savings up to about 6% in DRC.

Index Terms— Wyner-Ziv coding, turbo codes, probability updating, encoder rate control, decoder rate control

1. INTRODUCTION

Wyner-Ziv video coding (WZVC) allows a flexible complexity allocation between encoder and decoder by exploiting the video statistics, partial or totally, at the decoder; theoretically, WZVC can reach the same rate-distortion (RD) performance as predictive video coding (under certain conditions) [1], where video statistics are jointly explored at both the encoder and decoder. The turbo coding based transform domain Wyner-Ziv (TDWZ) video coding scheme proposed in [2] is, nowadays, the most studied WZVC solution. Since the error correcting code (typically a turbo code) is the core of this TDWZ video coding architecture, it assumes a key role in terms of the overall RD performance. To control the turbo parity rate to achieve good compression, the decoder makes use of a feedback channel (FC); as in the literature, this type of rate control strategy will be referred in this paper as *decoder rate control* (DRC). More recently, FC suppression was proposed for TDWZ video coding [3][4] since FC usage is not possible in certain application scenarios due to FC unavailability and the associated delay implications. Without FC, the decoder cannot perform rate control and, therefore, this task becomes the encoder responsibility; this type of rate control strategy is typically known as *encoder rate control* (ERC). Despite all the efforts, the RD performance of both DRC and ERC based WZVC schemes are still below the state-of-the-art predictive video coding RD performance. Moreover, DRC solutions have typically a higher RD performance than ERC solutions since the rate does not have to be estimated but simply requested when needed.

In this context, this paper proposes a probability updating technique (PUT) to enhance the turbo coding performance in TDWZ video coding and, therefore, to improve the overall WZVC RD performance. The proposed PUT is inspired on the idea presented in [5] of inserting correction impulses in the most likely in error bits of the received systematic data to enhance the turbo code performance in a channel coding scenario. However, the necessary adaptation of the turbo decoder to the WZVC coding scenario (joint source-channel coding) [6], notably in terms of correlation noise (CN) modeling, does not allow to use directly the solution in [5]. The proposed PUT acts over the probabilities fed to the turbo decoder, which are computed from the correlation noise between the WZ and side information (SI) frames, updating them for the most likely in error bits. This probability updating process is performed after the common turbo decoding operation and, therefore, it uses information about the WZ (source) data (available after turbo decoding) to provide more accurate data to the turbo decoder input; by doing this, PUT aims to help the turbo decoder to further correct the bitplane errors without additional parity bits. Whenever the proposed PUT leads to successful turbo decoding, WZ rate can be saved and, thus, better overall RD performance is obtained. For completeness, both DRC and ERC based TDWZ video coding architectures are considered to evaluate the proposed PUT since the same turbo code is used for both rate control strategies. The proposed PUT leads to total PSNR gains up to about 0.5 dB in an ERC scenario and WZ rate savings up to about 6% in the DRC scenario. An improved parity rate estimation metric for the ERC scenario regarding the one presented by the same authors in [3] is also proposed here; although the main parity rate estimator (PRE) agents are similar, they are combined in a way that leads to more accurate encoder rate estimation and, thus, to a better overall RD performance. This paper is organized as follows: Section 2 presents an overview of the TDWZ codec; Section 3 proposes the novel PUT; Section 4 proposes an improved PRE for ERC based TDWZ video coding; Section 5 presents experimental results for both DRC and ERC based TDWZ video coding; and, finally, some final remarks are presented in Section 6.

2. THE TRANSFORM DOMAIN WYNER-ZIV (TDWZ) VIDEO CODEC

When the dotted modules at the encoder are not taken into account, Figure 1 illustrates the DRC based TDWZ video codec architecture proposed in [7] and adopted in this paper; this codec is an advanced TDWZ video coding solution which shares the same architecture with the codec proposed by Aaron *et al.* in [2] but uses more efficient tools which contribute to achieve a better RD performance (for more details the reader should consult [7]). The DRC based TDWZ video codec depicted in Figure 1 works similarly to the codec described in [7]. The major contribution of this paper lies on the turbo decoder (blue highlighted module in Figure 1). Basically, a probability updating tool (PUT) is proposed which acts as an outer loop of the common turbo decoding operation whenever a turbo decoded bitplane is not error-free; 'common turbo decoding' stands for the usual iterative decoding procedure which makes use of two soft-input soft-output decoders.

978-1-4244-7994-8/10/$26.00 ©2010 IEEE

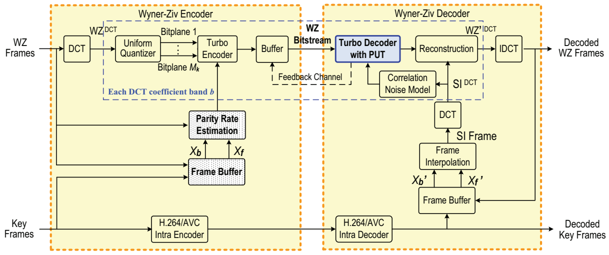

Figure 1 – TDWZ video codec architecture.

When a parity bits chunk is received, it is first commonly turbo decoded. If an error-free bitplane is not obtained, PUT attempts to correct the remaining errors in the turbo decoded bitplane by updating the *correlation noise probabilities* (CNPs) at the turbo decoder input for the most likely in error bits (one bit at a time), followed by turbo re-decoding; CNPs stand for the conditional probabilities that are obtained from the CN model, employed to convert the SI DCT coefficients into *soft* information (probabilities) needed for turbo decoding. This CNPs updating and turbo re-decoding procedure is performed until an error-free bitplane is achieve or the CNPs have been updated for a given number of bits. When an error-free bitplane is obtained, the next DCT band bitplane starts to be turbo decoded in an analogous way; otherwise, the turbo decoder requests for more parity bits through the FC and the overall turbo decoding with PUT process starts again. In case an error-free bitplane is obtained in the initial turbo decoding stage (before using PUT), the next bitplane associated to that DCT band starts to be turbo decoded; this means PUT is skipped since there are no errors to correct in the current bitplane.

Consider again Figure 1 now with the dotted modules at the encoder taken into account and the FC suppressed. This case depicts the ERC based TDWZ video codec architecture presented in [3]; this codec will also be adopted in this paper to have a more complete evaluation of the proposed PUT. In general, the ERC based TDWZ video coding procedure is similar to the DRC based one. The major differences regard the way the rate control operation is performed and, consequently, the turbo decoder behavior. Since no FC is available, the amount of parity data to be transmitted to the decoder is estimated at the encoder side for each bitplane of each DCT band; this bitplane parity rate estimation is performed by the parity rate estimation module using a past X_b and a future X_f reference frames of the WZ frame to be coded and is sent at once to the decoder. The turbo decoder works in a similar way as for the DRC case above except when the estimated parity rate is not enough to correct the SI DCT bitplane and the PUT is not able to correct the bitplane errors; in this case, the decoder simply proceeds to the next bitplane as there is no FC to request for more parity data to correct those errors. It is important to note that the proposed PUT has a higher impact in the ERC scenario than in the DRC one: by correcting the turbo decoded bitplane remaining errors, PUT allows to remove (or reduce) block artifacts which would appear in the decoded frame when the estimated parity rate is not enough or not properly distributed by the DCT bands bitplanes.

3. PROBABILITY UPDATING TECHNIQUE (PUT)

The turbo decoder with the proposed PUT aims to improve the overall TDWZ video codec RD performance. To accomplish this, PUT acts on one of the turbo decoder inputs, notably on the information provided by the correlation noise model, i.e. on the CNPs

$$P_{-1} = P\left(B_n^i = -1 \middle| Y_n, B_n^{i-1}, \ldots, B_n^0\right) \ \wedge \ P_{+1} = P\left(B_n^i = +1 \middle| Y_n, B_n^{i-1}, \ldots, B_n^0\right) \qquad (1)$$

where B_n^i represents the n^{th} bit of a given WZ DCT band bitplane B^i, Y_n stands for the n^{th} SI^{DCT} coefficient of the corresponding band and $P(z)$ is the probability of z; bitplane index i varies from 0 (most significant bitplane) to M-1 (where 2^M is the number of quantizer levels used). Basically, the CNPs are computed for each bit (representing the quantized DCT source symbol) of a given bitplane to be decoded taking into account the SI and all the previously decoded bits, as described in [3]. The more accurate the CNPs, the better is the overall RD performance [8]. Whenever the WZ decoder receives new parity data, the turbo decoder with the proposed PUT proceeds as follows:

1. **Decode the bitplane as usual in WZ video coding** [7]. From now on in this paper, this first turbo decoding operation, as well as all the data involved on it (i.e. parity data, CNPs, the *a posteriori* probabilities, bit hard-decisions), will be referred as *initial* decoding/data.
2. **Determine bitplane error probability using a confidence measure based on the *a posteriori* probabilities ratio [7]**.
   a. If the decoded bitplane is error-free, store the decoded bitplane as the output of the turbo decoder. Go to Step 1 and start decoding the next bitplane.
   b. If the decoded bitplane is not error-free, the proposed PUT will attempt to correct those errors. First, store the decoded bitplane as well as the logarithm of the *a posteriori* probability (LAPP) ratio associated to each bit given by

$$L_{ap}(B_n) = \ln\left(\frac{P\left(B_n = +1 \middle| Y, \mathit{parity\ data}\right)}{P\left(B_n = -1 \middle| Y, \mathit{parity\ data}\right)}\right) \qquad (2)$$

   where $L_{ap}(B_n)$ stands for the initial LAPP ratio of the n^{th} bit, B_n, $P(B_n = \pm 1 | Y, \mathit{parity\ data})$ is the *a posteriori* probability (i.e. the probability of the n^{th} WZ bit being equal to ±1 given the SI^{DCT} coefficient and the received parity data), and ln(.) is the natural logarithm operator.
3. **Sort the LAPP ratio values in increasing order of magnitude.**
4. **Get the *N* bits most likely in error from the sorted LAPP ratio values (obtained in Step 3)**. The lower the magnitude of the decoded bit LAPP ratio, the lower is the turbo decoder confidence on the corresponding bit decision. Thus, lower bit LAPP ratio magnitudes most probably correspond to an erroneous bit, since the probabilities $P(B_n = +1 | Y, \mathit{parity\ data})$ and $P(B_n = -1 | Y, \mathit{parity\ data})$ are close to each other.
5. **Process each one of the *N* bits obtained in Step 4, one at a time, starting with the bit with lowest LAPP ratio magnitude**:
   a. Store the initial CNPs (used in Step 1) for the bit being processed (say n^{th} bit).
   b. Update the n^{th} bit CNPs to force the turbo decoder to change (or flip) the n^{th} bit decision (±1); if a decoded bit is in error, by flipping it the error will be corrected. The initial CNPs are obtained from the CN distribution whose parameter is estimated without having access to the original WZ data [7]. However, at the current decoding stage, some information regarding the original WZ bitplane is already available (i.e. the LAPP) due to the decoding of the parity data already received. In this context, the initial LAPP ratio is used to obtain (better) CNPs for the n^{th} bit, according to the following two cases: 1) $\text{sgn}\left[L_{ap}(B_n)\right] = \text{sgn}\left[L_{cn}(B_n)\right]$ and 2) $\text{sgn}\left[L_{ap}(B_n)\right] \neq \text{sgn}\left[L_{cn}(B_n)\right]$; sgn[.] is the sign function and $L_{cn}(B_n)$ stands for the logarithm of the CNPs ratio given by

$$L_{cn}(B_n) = \ln\left(\frac{P_{+1}}{P_{-1}}\right). \qquad (3)$$

   The sign of $L_{ap}(B_n)$ corresponds to the decoded bit value (±1); by the same reasoning, the sign of $L_{cn}(B_n)$ can be seen as the bit value at the turbo decoder input. Assuming the n^{th} decoded bit is in error, if the $L_{ap}(B_n)$ sign (at the output of the turbo decoding) matches the $L_{cn}(B_n)$ sign (at the turbo decoder input) – case 1) – this means that $L_{cn}(B_n)$ also corresponds to an erroneous bit. In this case, the CNPs are update to the complementary of the initial ones, i.e.

$$\mathrm{P}^{u}_{-1} = \mathrm{P}_{+1} \;\wedge\; \mathrm{P}^{u}_{+1} = \mathrm{P}_{-1} \tag{4}$$

where P^u standing for updated CNP is sufficient to force the turbo decoder to change the n^{th} decoded bit decision and, therefore, to correct the (likely) error. In case 2), and once more assuming the n^{th} decoded bit is in error, the $L_{cn}(B_n)$ sign most likely corresponds to the error-free bit value, following the reasoning above. Therefore, the strategy to update the CNPs in (4) is inadequate since it would flip the $L_{cn}(B_n)$ sign. In this case, the CNPs are updated by reinforcing the magnitude of the corresponding initial CNPs according to

$$\mathrm{P}^{u}_{-1} = e^{L_{ap}} \times \mathrm{P}_{-1} \times c \;\wedge\; \mathrm{P}^{u}_{+1} = \frac{1}{e^{L_{ap}}} \times \mathrm{P}_{+1} \times c \tag{5}$$

which is typically sufficient to force the decoder to flip the n^{th} decoded bit decision, correcting the (likely) error; in (5), c is a normalizing factor used to guarantee that $\mathrm{P}^{u}_{-1} + \mathrm{P}^{u}_{+1} = 1$. The initial *a posteriori* probabilities ratio e^{Lap} reflects the confidence on the decoded bit decision; assuming the decoded bit is in error, the higher the e^{Lap} value, the higher is the reinforcement provided to the turbo decoder input to force it to change the corresponding bit decision.

c. Re-decode the bitplane (as in Step 1) with the updated CNPs computed from (4) or (5) for the n^{th} bit.

d. Determine the bitplane error probability as in Step 2.

  i. If the decoded bitplane is not error-free, restore the n^{th} bit CNPs with the initial ones (used in Step 1) since the bitplane is still in error and, thus, the updated $L_{cn}(B_n)$ value is not guaranteed to be better than the initial one. Then, move to the next bit that is likely in error within the N selected bits and go to Step 5. If all N bits were already processed, the output of the turbo decoder is the initial decoded bitplane (obtained in Step 1) as there is no guarantee that additional errors were not introduced during the probability updating process; go to Step 1 for another bitplane (ERC) or another chunk of parity data (DRC) for the same bitplane. In the ERC case, an undesirable loss in quality occurs since some erroneous bits were not corrected.

  ii. If the decoded bitplane is error-free, perform as in Step 2a. In this case, PUT allowed to further correct bitplane errors leading to WZ rate savings in the DRC scenario or decoded quality improvements in the ERC scenario.

Since DRC and ERC based TDWZ video codecs use the same turbo codec, the PUT description proposed above is the same for both rate control strategies.

4. PARITY RATE ESTIMATOR (PRE)

In this paper, the bitplane parity rate estimation for the ERC based TDWZ video codec is given by

$$R^i = \frac{1}{2} H_{X|Y}^{B^i} \times e^{H_{X|Y}^{B^i}} + w \times \sqrt{p} \tag{6}$$

where $H_{X|Y}^{B^i}$ stands for the i^{th} bitplane conditional entropy of the WZ data (X) given the corresponding SI (Y), p is the bitplane relative error probability and w is a weighting factor; for a detailed description on p and bitplane conditional entropy computation, see [3]. As in [3], p is used to better allocate the parity rate by avoiding the undesired parity rate underestimation case since it introduces a quality penalty. Nevertheless, it may lead to overestimation of the parity rate, introducing a bitrate penalty; it was found experimentally that this is especially true for low motion video sequences. In this context, it is proposed here to affect the p term by the weight w allowing to have a more accurate PRE. Figure 2 illustrates the relationship between the number of parity chunks (NoPC) estimated at encoder (with (6) and with [3]) and the DRC scenario NoPC for the 6^{th} bitplane of the first AC band of the *Hall Monitor* QCIF video sequence at 15 Hz; all the results in Figure 2 refer to the highest quality RD point in [7]. As it can be observed, the proposed PRE (circles) is closer to the ideal PRE than the PRE in [3] (triangles); the ideal PRE line corresponds to the case where the estimated NoPC matches the NoPC really needed to correct the SI errors (the DRC case).

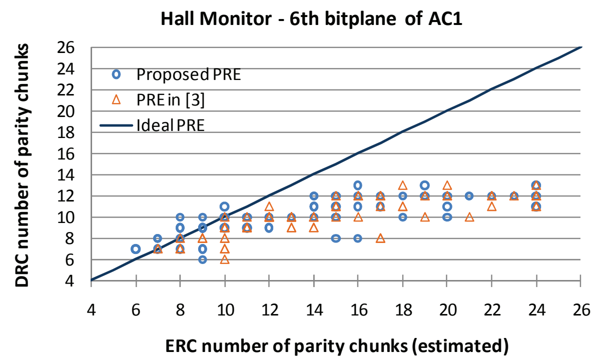

Figure 2 – DRC versus ERC number of parity chunks.

5. EXPERIMENTAL RESULTS

To assess the proposed PUT, two rather different QCIF video sequences are considered: *Hall Monitor*@15Hz and *Soccer*@30Hz; all frames are used, i.e. 165 frames for *Hall Monitor* and 300 frames for *Soccer*; since Soccer has rather high and complex motion, it is appropriate to use a higher frame rate. All the experiments were conducted only for the luminance component, as usual in WZVC. The test conditions for the DCT, quantizer, frame interpolation, turbo codec and reconstruction modules are the same as in [7]; a GOP length of 2 is used. The key frames are Intra coded with H.264/AVC Main profile with a QP depending on the RD point. The weight w in (6) is equal to $\sqrt{0.5}$ since this value was considered a good trade-off between parity rate overestimation and underestimation.

5.1. Decoder rate control performance

Table 1 shows, for the *Hall Monitor* and *Soccer* sequences, the total rate R savings ΔR, in percentage, for the DRC based TDWZ video codec with the proposed PUT regarding the same codec without PUT; this means $\Delta R = 100 \times (R_{DRC} - R_{DRC+PUT})/R_{DRC}$. In Table 1, Q_j represents the j^{th} WZ quantization matrix associated with the j^{th} RD point [7]; the higher the Q_j, the higher is the rate and the quality. Table 1 results were obtained with $N = 15$ (number of bits most likely in error); this value was found experimentally to be a good trade-off between decoder complexity increase (due to turbo re-recoding) and RD performance. The same PSNR values were obtained for the two codec solutions since, in both cases, the turbo decoder requests for parity data until successful decoding is reached; thus, and due to paper length constraints, only one PSNR column is shown. Table 1 also illustrates the WZ rate savings ΔR_{WZ}, in percentage, since the proposed technique applies only to the WZ frames. For bitplanes with initial mean absolute L_{ap} lower than 3.0 (i.e. with a high number of errors), the PUT Steps 3-5 are skipped and the decoder goes to Step 1 for another bitplane or another chunk of parity data, speeding up the decoding process; since PUT acts only in one bit at a time, it is not expectable to correct a huge amount of errors and, therefore, performing turbo re-decoding with updated CNPs would be very likely a waste of resources.

Table 1 – Total and WZ rate savings (in %) for DRC based TDWZ video codec with PUT.

Hall Monitor (QCIF, 15 Hz)				
Q_j	DRC		DRC+PUT	
	Rate [kbps]	PSNR [dB]	ΔR [%]	ΔR_{WZ} [%]
1	84.35	31.46	0.85	6.20
3	97.86	32.07	0.94	4.65
5	137.35	34.38	1.20	4.90
7	200.21	37.36	1.58	5.56
Soccer (QCIF, 30 Hz)				
Q_j	DRC		DRC+PUT	
	Rate [kbps]	PSNR [dB]	ΔR [%]	ΔR_{WZ} [%]
1	96.68	28.59	2.51	4.73
3	139.21	30.01	2.67	4.87
5	246.76	32.53	2.95	5.43
7	473.41	35.81	2.16	4.26

As it can be observed in Table 1, the proposed PUT allows reducing the total rate up to 3% for the *Soccer* sequence. In terms of WZ rate, savings up to 6.2% can be obtained for the *Hall Monitor* and up to 5.4% for the *Soccer* sequence. For *Hall Monitor*, the key frames bitrate is significant when compared to the WZ bitrate and, thus, ΔR improvements are smaller in spite of the ΔR_{WZ} higher gains.

5.2. Encoder rate control performance

Figure 3 illustrates the ERC based TDWZ video codec RD performance with and without the PUT using the PRE metric proposed in Section 4; the RD performance obtained with the same codec using the PRE solution in [3] is also shown. In Figure 3, the "Proposed PRE + PUT" curve was obtained with $N = 70$. Due to the negative quality impact of the errors left in the decoded frame, a higher N value was used in the ERC scenario (compared to the DRC case); a higher N value means that more likely in error bits are processed and, therefore, the decoded quality increases for the allocated rate.

As shown in Figure 3, the novel PRE leads to better RD performance than the one in [3], notably for low motion video sequences, where the proposed PRE reveals to be more effective due to the reduction of parity rate overestimation (see Section 4); consequently, the gap between the ERC and DRC scenarios is reduced. For *Soccer* (high motion sequence), the overall RD performance obtained with the proposed PRE is similar to the one in [3]. Comparing the codec RD performance obtained with and without the proposed PUT (using the PRE in (6)), PSNR gains up to 0.5 dB can be observed for *Soccer*; for *Hall Monitor*, gains up to 0.13 dB are obtained. The highest gains are observed for *Soccer* since there are more errors left to correct due to the bigger difficulty to estimate the rate for high motion sequences (low quality SI is typically generated). Regarding state-of-the-art video codecs with similar encoding complexity, notably H.264/AVC Intra and H.264/AVC Zero-Motion, it can be noticed that: a) for *Hall Monitor* (low motion sequence) the proposed TDWZ video codec is more efficient than H.264/AVC Intra; b) for both sequences, H.264/AVC Zero-Motion codec is more efficient than the TDWZ video codec since it exploits co-located frame differences at the encoder to achieve better RD performance.

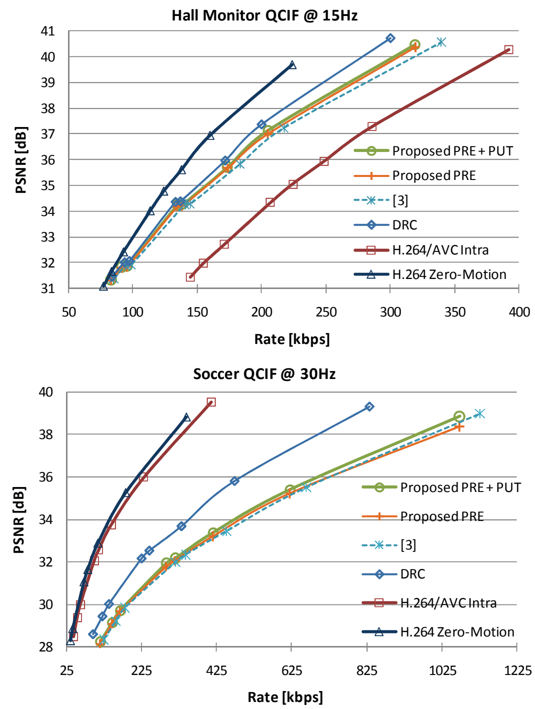

Figure 3 – ERC based TDWZ video codec RD performance.

6. FINAL REMARKS

This paper proposes a novel tool (PUT) to improve both DRC and ERC based TDWZ video coding RD performances; the proposed PUT acts an outer loop of the common turbo decoding operation by updating the CNPs fed to the turbo decoder for the most likely in error bits. Experimental results show: a) overall PSNR gains up to about 0.5 dB in ERC scenarios; and b) WZ rate savings up to about 6% in DRC scenarios.

7. REFERENCES

[1] A. Wyner, J. Ziv, "The Rate-Distortion Function for Source Coding with Side Information at the Decoder", *IEEE Trans. Inf. Theory*, vol. 22, no. 1, pp. 1-10, Jan. 1976.

[2] A. Aaron, S. Rane, E. Setton, B. Girod, "Transform-Domain Wyner-Ziv Codec for Video", *VCIP*, San Jose, California, USA, Jan. 2004.

[3] C. Brites, F. Pereira, "Encoder Rate Control for Transform Domain Wyner-Ziv Video Coding", *IEEE ICIP*, San Antonio, Texas, USA, Sep. 2007.

[4] W. A. R. J. Weerakkody, W. A. C. Fernando, A. M. Kondoz, "Enhanced Reconstruction Algorithm for Unidirectional Distributed Video Coding", *IET Image Process.*, vol. 3, no.6, pp. 329-334, Dec. 2009.

[5] Y. Ould-Cheikh-Mouhamedou, S. Crozier, K. Gracie, P. Guinand, P. Kabal, "A Method for Lowering Turbo Code Error Flare using Correction Impulses and Repeated Decoding", *Int. Symp. Turbo Codes and Related Topics*, Munich, Germany, Apr. 2006.

[6] C. Brites, "Advances on Distributed Video Coding", M.Sc. Thesis, Instituto Superior Técnico, Technical University of Lisbon, Portugal, Dec. 2005.

[7] C. Brites, J. Ascenso, J. Q. Pedro, F. Pereira, "Evaluating a Feedback Channel based Transform Domain Wyner-Ziv Video Codec", *Signal Process. Image Commun.*, vol. 23, no. 4, pp. 269-297, Apr. 2008.

[8] C. Brites, F. Pereira, "Correlation Noise Modeling for Efficient Pixel and Transform Domain Wyner-Ziv Video Coding", *IEEE Trans. Circuits Syst. Video Technol.*, vol. 18, no. 9, pp. 1177-1190, Sep. 2008.