ENHANCED RECONSTRUCTION OF THE QUANTISED TRANSFORM COEFFICIENTS FOR WYNER-ZIV CODING

Yuri Vatis, Student Member IEEE, Sven Klomp, Student Member IEEE, Jörn Ostermann, Fellow IEEE

Institut für Informationsverarbeitung
Leibniz Universität Hannover, Germany

ABSTRACT

In recent years, the Distributed Video Coding (DVC) has become a more and more popular research area. In the current work, the framework of transform domain Wyner-Ziv coding of video frames is considered, following the scheme developed in the EU Project DISCOVER [1]. While some frames are conventionally coded as key frames, other frames are Wyner-Ziv encoded using turbo codes. After the generation of side information and turbo decoding process at the decoder, the quantisation indices of the quantised transform coefficients of Wyner-Ziv frames are available. In this paper, we devote our attention to the reconstruction of quantised transform coefficients resulting in a minimised quantisation error energy, where coding gains of up to 0.9 dB are obtained.

Index Terms**—** Distributed algorithms, Video coding, Wyner-Ziv.

1. INTRODUCTION

Current video coding solutions, such as MPEG or ITU-T H.26x standards, perform well for broadcasting, streaming and other applications, wherein a video is encoded once and decoded several times. The encoder of such a solution exploits the source statistics, whereby the decoder can be kept very simple. For opposite scenarios with many encoders, Distributed Video Coding might be more suitable than conventional video coding since the decoder performs the complex task of exploiting the source statistics.

DVC is based on the Slepian-Wolf [2] and Wyner-Ziv [3] theorems. These theorems state that it is possible to compress two statistically dependent signals in a distributed way (separate encoding, joint decoding) using a rate equal to that used in a system, where the signals are encoded and decoded together. Current approaches mostly implement the asymmetrical case, where the two signals are coded with different bit rates. A general block diagram of an asymmetrical Wyner-Ziv (WZ) video codec is shown in Figure 1. At the encoder, the sequence is divided into key frames and Wyner-Ziv frames controlled by the group-of-picture (GOP) size (e.g. at GOP size 4 every fourth frame is coded as key frame). The key frames (here, X_{2i-1} and X_{2i+1} for GOP size 2) are coded with a H.264/AVC intra frame coder and the Wyner-Ziv frames with a distributed coder. Starting from key frames, the decoder generates the side information for the Wyner-Ziv frames. For that purpose, the framework presented in [4] and enhanced in [5] is used. Then, the side information, transformed in the frequency domain, is fed to the turbo decoder, which regards it as corrupted information bits. Furthermore, the turbo decoder gets the parity bits from the Wyner-Ziv encoder, so that after the decoding process the quantisation indices of the transform coefficients of the Wyner-Ziv frames are known at the decoder. This paper deals with the reconstruction method of the transform coefficients, based on the minimisation of the quantisation error energy. The paper is organised as follows: Section 2 gives an overview about current quantisation method of the Wyner-Ziv encoder while Section 3 is split into two parts. The current reconstruction scheme is described in the first part. In the second part, an enhanced method is proposed. The results are given in Section 4. This paper finishes with conclusions in Section 5.

2. QUANTISATION OF THE TRANSFORM COEFFICIENTS

In the transform domain scenario of the reference codec [6], a H.264/AVC 4x4-integer transform is applied on each Wyner-Ziv frame X_{2i}. The transform coefficients of the entire frame X_{2i} are then grouped together, according to the position occupied by each transform coefficient within the 4x4 blocks, forming the transform coefficient bands. After the transform coding operation, each coefficients band, labeled as b_k, is uniformly quantised with 2^{M_k} levels (where the number of levels 2^{M_k} depends on the coefficient band $b_k (k = 1, \dots, 16)$). For the quantisation, it is distinguished between AC and DC coefficients:

1) The DC coefficients, which are only positive values, are quantised using a uniform scalar quantiser.

2) The remaining AC coefficients are quantised using the uniform scalar quantiser with a symmetric quantisation interval around zero (Figure 2).

In the reference codec, the dynamic range for each transform coefficients band of each Wyner-Ziv frame is transmitted to the decoder. Dependent on the transmitted

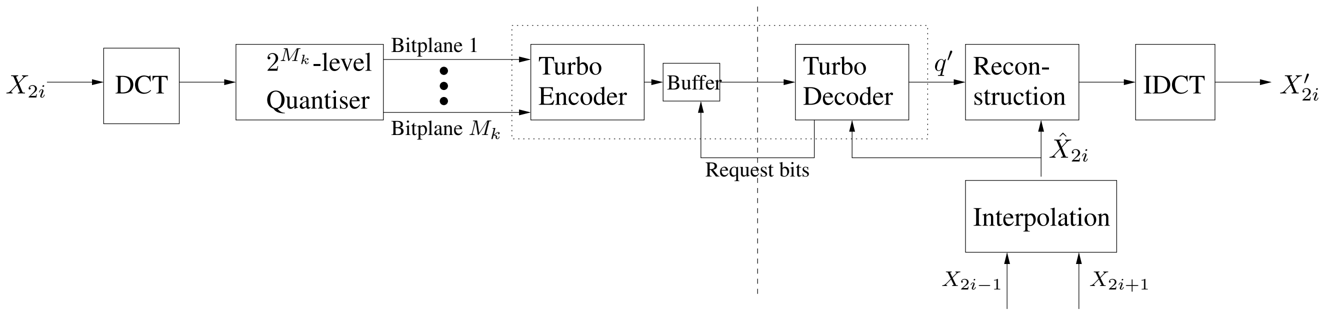

Figure 1: DVC Architecture

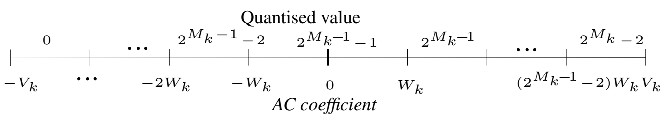

Figure 2: Quantisation of AC coefficients

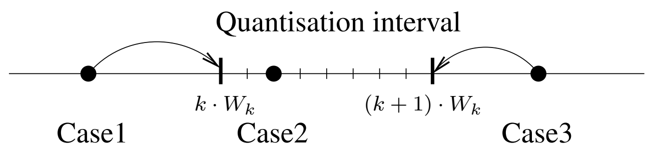

Figure 3: Different cases for the alignment of the side information value and decoded quantisation index

dynamic range, the quantisation step size W_k is adjusted. The quantisation step size W_k for the coefficients b_k is obtained from

$$W_k = \begin{cases} \frac{2 \cdot V_k}{2^{M_k}-1} & \text{for AC coefficients and} \\ \frac{V_k}{2^{M_k}} & \text{for DC coefficients} \end{cases} \quad (1)$$

whereas V_k is the highest absolute value of b_k, and 2^{M_k} is the number of quantisation levels for every b_k.

3. RECONSTRUCTION OF THE TRANSFORM COEFFICIENTS

3.1. State of the art

After the turbo decoding process, the decoder has a quantisation index of each coefficient. It is assumed that the transmission of the quantisation indices from the encoder to the decoder is error-free. At the reconstruction process, it is distinguished between 3 cases, depicted in Figure 3: *Case1*, when the side information value is smaller than the smallest value inside the decoded quantisation interval ($[k \cdot W_k \dots (k+1) \cdot W_k]$), *Case2*, when the side information is inside of the quantisation interval and *Case3*, when the side information value is greater than the greatest value of the current quantisation interval. In *Case1*, the smallest value inside of the quantisation interval of the decoded quantisation index is taken. In *Case2*, the current side information value is used as a reconstruction value. In *Case3*, the greatest value of the current quantisation interval is taken.

3.2. Proposal

Let us first focus on *Case1* and *Case3*. Although the outer values are the most common ones (which is to expect, since the side information values, used for the decoding process, are a good approximation of the real transform coefficients), the optimal strategy to minimise the quantisation error energy is to take the expected value of the original transform coefficients. For that purpose, we generated statistics using several sequences. Furthermore, we separated the quantisation intervals into two groups, *groupB* for big intervals and *groupS* for small ones, due to the observation of a slight dependency between the size of a quantisation interval and the expected value of the original transform coefficient. In Figure 4, the frequency of occurrence of the original AC and DC coefficients within the current quantisation interval is depicted for *Case1* for both groups (*groupS* left, *groupB* right) and for different sequences (such as test sequences: Flowergarden, above, and Soccer, below), which have quasi exponential distribution (Equation 2). The fitted exponential distributions are also depicted in Figure 4. The distribution of the transform coefficients for *Case3* is symmetrical to *Case1*:

$$f_\lambda(x) = \begin{cases} 0 & : \quad x < 0 \\ c\lambda e^{-\lambda x} & : \quad 0 \le x \le 1 \\ 0 & : \quad x > 1 \end{cases} \quad (2)$$

where c is the scaling factor, so that the integral of $f_\lambda(x)$ is equal to 1.

One can observe that the expected value of the original transform coefficients for the sequence Flowergarden, where the side information is much more accurate, is closer to the lower end of the quantisation interval, compared to the sequence Soccer. However, the similarities in both sequences are that for small quantisation intervals, the expected value of the original transform coefficient tends to shift towards the centre of the quantisation interval in comparison to big quantisation intervals, i.e. $\lambda_S > \lambda_B$. Since it is difficult to predict how accurate the side information is, we use an average of the expected values of all tested sequences distinguishing both, big and small quantisation intervals.

In our observation of *Case2* (the side information value is inside of the quantisation interval), the expected value of the original transform coefficient is slightly shifted towards the centre of the quantisation interval in comparison to the side information value. This means, the method for the generation of the side information is not an unbiased predictor. Therefore, in order to compensate the bias, we compute the reconstructed value r as

$$r = a \cdot S + (1 - a) \cdot C \tag{3}$$

where S is the side information value, $C = \frac{1}{2}(k \cdot W_k + (k + 1) \cdot W_k)$ is the centre of the quantisation interval and a is empirically determined as $\frac{7}{8}$.

4. EXPERIMENTAL RESULTS

For the evaluation of the rate-distortion performance, the new method (called here as *enhanced rec.*) is compared with the current software (*discover*). The performance of the H.264 intra frame coder is also considered for comparison. The sequences are coded at QCIF resolution with 15Hz and GOP size II.

Using the proposed method, coding gains of up to 0.9 dB are achieved for the test sequences Foreman, Soccer and Coastguard - the sequences, for which the motion compensated prediction or side information generation performs poorely. For these sequences, the rate-distortion performance is partially or completely below the H.264 intra performance. However, for other sequences, such as test sequence Flowergarden, the coding gains up to 0.4 dB are observed.

The major part of the coding gains is achieved, when the side information value is outside of the current quantisation interval, as in *Case1* and *Case3*. Yet, the enhanced reconstruction improves the current state of the art in case the side information value is inside of the current quantisation interval (*Case2*) of e.g. up to 0.15 dB for Soccer or up to 0.1 dB for Flowergarden, respectively.

5. CONCLUSIONS

We present a method for reconstructing transform coefficients of Wyner-Ziv video frames. The reconstructed value is not only determined by the boundaries of the decoded quantisation interval. We propose to improve reconstruction by using expected values of the original transform coefficients, dependent on boundaries, centre and size of the current quantisation interval.

The results in Section 4 show that enhanced reconstruction of the quantised transform coefficients improves the current software for all sequences and all bit rates. A coding gain of up to 0.9 dB for the same bit rates is achieved. The highest gains are observed for the sequences, which rate-distortion performance is partially or completely below H.264/AVC intra, which is especially important for achieving a consistent video coding performance. For other sequences, the gain is lower but not insignificant.

6. ACKNOWLEDGEMENTS

The work presented was developed within DISCOVER (Distributed Coding for Video Services), an European Commission Future and Emerging Technologies (FET) project, funded under the European Commission IST FP6 programme. Further information can be found at www.discoverdvc.org. The DISCOVER software started from the IST WZ software developed at the Image Group from Instituto Superior Técnico (IST) of Lisbon by Catarina Brites, João Ascenso and Fernando Pereira.

7. REFERENCES

[1] C. Brites, J. Ascenso, and F. Pereira, "Improving transform domain wyner-ziv video coding performance," in *IEEE Int. Conf. on Acoustics, Speech, and Signal Processing*, Toulouse, France, May 2006.

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

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

[4] J. Ascenso, C. Brites, and F. Pereira, "Improving frame interpolation with spatial motion smoothing for pixel domain distributed video coding," in *5th EURASIP*, Slovak Republic, July 2005.

[5] S. Klomp, Y. Vatis, and J. Ostermann, "Side information interpolation with sub-pel motion compensation for wyner-ziv decoder," in *5th EURASIP*, Portugal, August 2006.

[6] C. Brites, J. Ascenso, and F. Pereira, "Studying temporal correlation noise modeling for pixel based wyner-ziv video coding," in *IEEE International Conference on Image Processing (ICIP 2006)*, Atlanta, USA, October 2006.

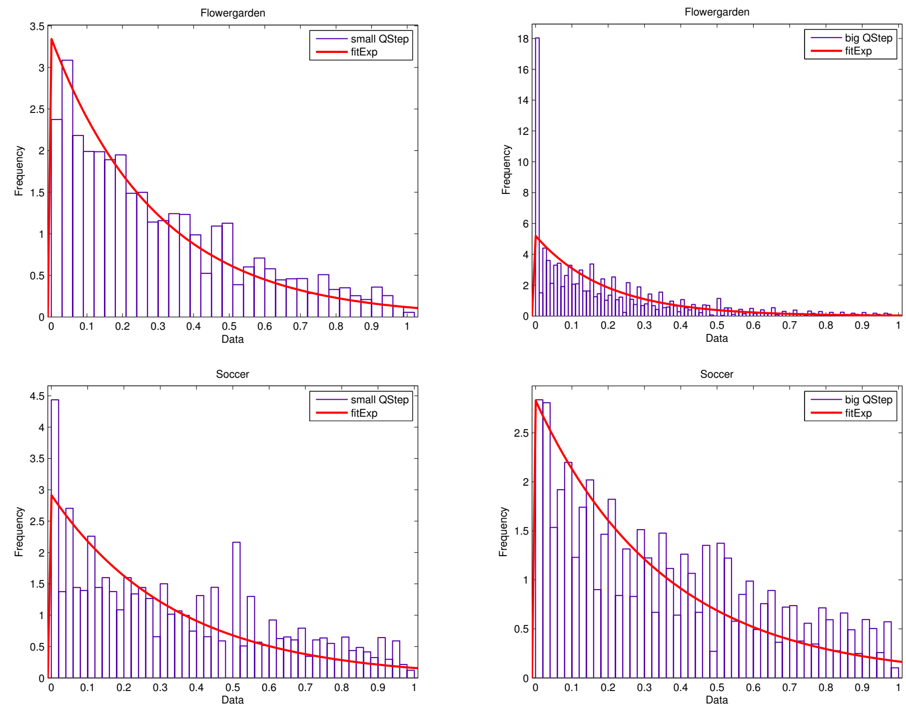

Figure 4: Frequency of occurrence of original transform coefficients within a current quantisation interval for sequences Flower (above) and Soccer (below) for small (left) and big(right) quantisation intervals and corresponding fitted exponential distributions (red line).

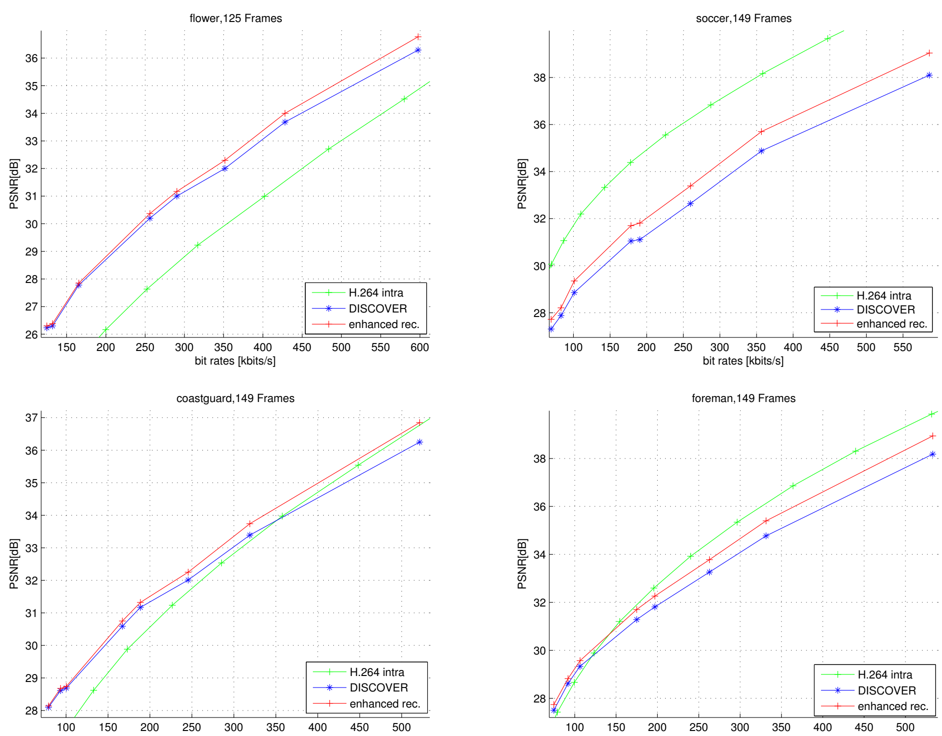

Figure 5: Rate distortion performance, provided by the H.264/AVC Intra (green), by the reference codec (red) and by enhanced codec (blue) for QCIF-sequences Flowergarden (top left), Soccer (top right), Coastguard (down left) and Foreman (down right).