

# JOINT SPATIAL AND TEMPORAL CORRELATION EXPLOITATION FOR WYNER-ZIV FRAMES CODING IN DVC

A.B.B. Adikari, W.A.C. Fernando, W.A.R.J. Weerakkody, A.M.Kondoz

Centre for Communication Systems Research

University of Surrey, Guildford, Surrey GU2 7XH, UK

adikari@ieee.org; W.Fernando@surrey.ac.uk; R.Weerakkody@surrey.ac.uk

## ABSTRACT

Source coding by exploiting the temporal and spatial correlations in an input video stream is well established in conventional video coding. However, when Distributed Video Coding (DVC) is concerned, shifting the source coding from the encoder to the decoder made this a more complicated and possibly a sub optimal process. Exploiting the temporal correlations has so far been attempted utilizing the key frames, by motion compensation using the interpolated (or extrapolated) motion field. However, exploiting the spatial correlations within the DVC framework is still an unsolved problem. In this paper, we propose to exploit both spatial and temporal correlations to generate the side information at the decoder. The proposed architecture involves a bit plane level side information refinement mechanism extracting both temporal and spatial information available at the decoder. Simulation results show that, the proposed coding scheme can achieve about 40% bit rate saving over 3D refinement algorithm proposed in [1] which is utilized as the basis for implementing the proposed technique.

**Key words:** Distributed video coding, Video coding, Digital signal processing

## 1. INTRODUCTION

It is reasonable to anticipate an increased demand in the near future for inexpensive video capturing mechanisms such as video sensor networks for security surveillance, monitoring of disaster zones, design of entertainment systems and some domestic applications like monitoring children and elderly people by the guardians. The low cost video conferencing requirements in mobile communications are also on the rise. Traditional video coding algorithms such as H.264/AVC, MPEG-2, and MPEG-4 are not ideal candidates for such requirements due to their computationally complex encoders that results in a higher manufacturing cost.

Distributed video coding (DVC) has emerged as a technology capable of reducing the processing complexity of the encoder, enabling low cost implementation, while majority of the computations are shifted to the decoder. This is an ideal solution for the requirement of capturing

video streams at multiple remote sites to be decoded at a central site as necessary in the above mentioned applications. The main advantages of the DVC encoder include low requirement of memory, computational capacity and power which are generally scarce resources at the remote sites.

In DVC, the source coding is performed in the decoder. This is in contrast to other conventional video compression schemes where the source coding happens in the encoder. The conventional approaches involve a complex encoder structure and an easy to implement-low cost decoder which was motivated by applications such as broadcasting, video on demand, and video streaming. The source coding at the decoder in DVC is done by means of generating a side information stream that is statistically correlated with the original data. In order to generate side information in the decoder, a sequence of selected original frames is generally passed to the decoder over the channel using existing intra coding techniques. These reference frames are called 'key-frames'. The existing intra coding tools in MPEG-2[4] or H.264/AVC [5] can be utilized for this. However, in one of our papers [6], we have proposed a Wyner-Ziv coding based key frame coding scheme that exploits spatial correlations.

The objective of this paper is to present an enhanced algorithm for side information estimation by exploiting both temporal and spatial correlations in contrast to the existing DVC codecs that exploit only the temporal correlations. Rest of the paper is organized as follows. In section 2, we summarize some related work on DVC. The proposed coding scheme is presented in section 3. Section 4 presents simulation results and a detailed analysis. Finally, section 5 presents the conclusion.

## 2. RELATED WORK

Distributed source coding concept was presented by Slepian and Wolf in 1973 [3]. They discussed the rate conditions for the independent decoding and joint decoding of statistically dependant discrete random sequences, for the lossless coding. Later in 1976, Wyner and Ziv extended this concept for the lossy coding scenario, presenting the concept of the use of *side information* at the decoder [2]. The side information ( $Y_m$ ) is commonly modeled

considering the original bit stream ( $X_m$ ) and an additive white noise term ( $n_m$ ) as,

$$Y_m = X_m + n_m$$

For most of the cases, this noise process can either be modeled using a Gaussian or Laplacian probability distribution.

Based on the above concept, a number of DVC solutions have been proposed; many of them based on the most dominant Stanford DVC framework [7]. Introducing the Stanford framework, Aaron et al. has proposed a Turbo coding based Wyner-Ziv codec [7][8]. They proposed simple frame interpolation [7] and motion interpolation and extrapolation techniques [8] to predict the side information. Improving the side information estimation has been the focus of much of the subsequent research in DVC. In some notable achievements, Tagliasacchi et al. proposed a motion compensated temporal filtering technique [9]. Ishwar et al.[10] presented an information theoretic study of video codecs based on source coding with side information. Nataro et al. proposed an algorithm to generate side information based on motion field smoothing to provide improved performance [11]. Ascenso et al., presented a scheme using motion interpolation to derive the side information [12]. They used forward and bidirectional motion estimation and a spatial motion smoothing algorithm to generate the side information. They also proposed a motion refinement algorithm using weighted motion estimation to further improve the side information [13]. This algorithm can be considered as the best available pixel domain Wyner-Ziv codec algorithm discussed in the research literature. Later-on, we consider this algorithm to compare the performance of proposed algorithm.

In other related work, we have proposed a TTCM based coding scheme for DVC and showed that a significant coding gain can be achieved compared to the more conventional Turbo coding based DVC codecs [14]. Furthermore, we also proposed several techniques to improve the performance of the DVC codec [15][16].

### 3. PROPOSED CODING SCHEME

The proposed DVC codec is based on the Stanford DVC framework in the pixel domain. The encoder structure is unaffected which include the quantizer and bit plane extraction followed by the Turbo encoder. The parity bit stream produced in Turbo coding is buffered for dynamic transmission to the decoder based on the closed loop parity request mechanism involving the feedback channel. The bit stream is segmented to small blocks before Turbo coding to improve the efficiency of parity request mechanism and also to help in the spatial prediction process discussed later in

this section. In the proposed solution, three side information streams are generated at the decoder, two of which are generated using motion extrapolation and compensation (ME-C) as in [16] and the third stream is generated by means of spatial prediction.. Since now the decoder handles three side information streams, the computational cost of the decoder is increased. To compensate for the additional computational complexity, we limit the constraint length of the Turbo coder to 4 whereas in [11] a length of 5 is proposed. Furthermore, to reduce the computational cost, we have reduced the number of iterations in the Turbo decoder to 3 compared to 18 iterations used in [11].

#### *Spatial Prediction*

The spatial prediction is performed by exploiting the spatial correlation within the partially decoded Wyner-Ziv frames. Since each bit plane is segmented into smaller blocks before Turbo coding, previously decoded blocks of the current frame are available for the prediction of side information for the current block. This process is an extension of the spatial prediction technique proposed for Wyner-Ziv Intra coding algorithm presented in our previous work [6]

#### *Decoding the most significant bit plane*

Figure 1 illustrates the major functional elements of the proposed codec for first bitplane coding, as shown, MC-E 1 and MC-E 2 generate two side information streams that are predicted by motion extrapolating the previous two closest key frames and immediate key frame and the closest Wyner-Ziv frame. This prediction and decision criteria are as in [16].The third side information stream generated by means of spatial prediction and in the figure this module is show as SP. It should be noted that this algorithm is implemented to perform bit stream segmentation before turbo coding similar to WZ-I coding scheme as in [6]. Therefore, even in the first bitplane, after decoding at least one block, information is available for the spatial prediction module to make a prediction to be used as side information for the next decoding cycle. Similar to other parallel DVC research, we assume that the decoder is capable of determining the error probability of the decoded block. Based on the error probability, the turbo decoder decides which side information stream is used for decoding a given block. Since the turbo decoder operates in multiple cycles depending on the length of the bitstream, it decodes one block of the bitstream independent of the other blocks.

#### *Decoding subsequent bit planes*

Figure 2 illustrates the active functional elements of the codec for decoding the higher order bitplane i.e., the bitplanes from the second bitplane onwards. As shown in the figure MC-E 1 and MC-E 2 are now inactive but 3D Refinement II[1] has become active that undertakes the side

information processing. This arrangement also uses three side information streams, first is the output of the 3D refinement, and the second is the composite side information stream from the 2-SI as in[1]. The third side information stream is again the output of SP module. In this stage SP module receives more information since the first bitplane is already decoded and therefore, it tends to perform even better.

#### 4. SIMULATION RESULTS

In all simulations presented in this paper, we used a turbo coder with two rate  $\frac{1}{2}$  component encoders with a constraint length (K) of 4. Turbo decoding is done for 3 iterations. All other simulation parameters are set as in[7].

The proposed DVC codec is tested for several test video sequences with the spatial the resolution of 176x144 (qcif). All picture parameters are set as in [7]. Results for four test video sequences are illustrated in figure 3. The illustrated results are averaged over first 100 frames of the sequences. The bitrate is varied by independently controlling the granularity of the quantizer as in the previous sections. A frame rate of 15 fps used with a GOP of 2 and the results are shown for the Wyner-Ziv frames only. The bitrate and the PSNR shown in the plot are only for Y signal of the sequence. Results of the proposed technique are shown together with several other techniques for the first test sequence. It is evident from the illustrated results that the proposed enhancement has resulted in a significant reduction of bitrate compared to the 3D refinement technique and a much higher gain can be observed compared to other techniques. The bitrate saving can be seen in the plot as a horizontal left-ward shift of the curve. The bitrate reduction is up to about 40% in some cases as it can be seen in the figures.

#### 5. CONCLUSIONS

In this paper we used both temporal and spatial properties of the image in Wyner-Ziv frame coding in DVC. We used the concept of multiple side information concept to improve the performance. To maintain the computational cost at a similar level, we reduced the constraint length of the turbo encoder and reduced the number of iterations. Based on the simulation results, it can be concluded that, the proposed algorithm provides significant reduction in bit rates in pixel based DVC codecs.

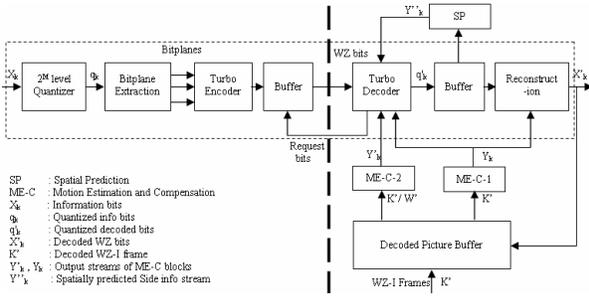
#### ACKNOWLEDGEMENT

The work presented was developed as a part of VISNET II, a European Network of Excellence ([http://www.visnet-](http://www.visnet-noe.org)

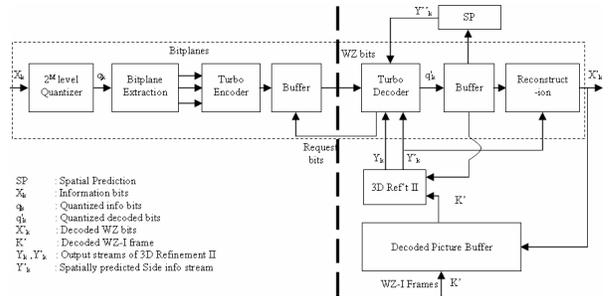
[noe.org](http://www.visnet-noe.org)), funded under the European Commission IST FP6 programme.

#### REFERENCES

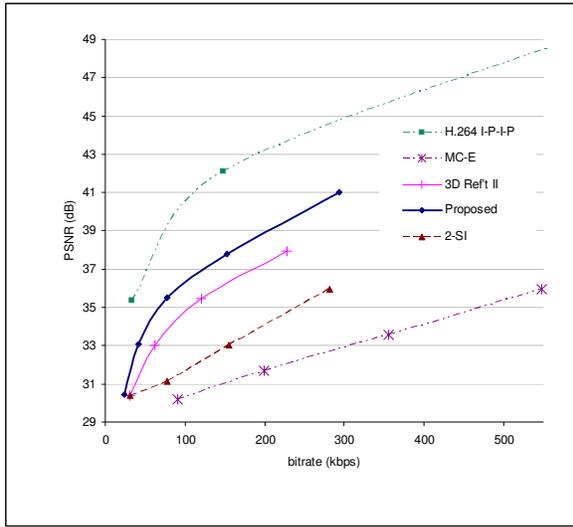
- [1] A.B.B. Adikari, W.A.C. Fernando, W.A.R.J. Weerakkody, "Side Information Improvement in DVC with Two Side Information Streams and 3D Motion Refinement," Proc. CCECE 2007, pp.32-35, Canada, April 2007.
- [2] D.Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," IEEE Trans.on Inform. Theory, vol. IT-22, pp. 1-10, Jan. 1976.
- [3] D. Slepian and J.K. Wolf, "Noiseless coding of correlated information sources," IEEE Trans. on Inform Theory, vol. IT-19, pp. 471-480, July 1973.
- [4] R. Sch'af'er and T. Sikora, "Digital video coding standards and their role in video communications," Proc. IEEE, vol. 83, pp. 907-924, June 1995.
- [5] Thomas Wiegand, Gary J. Sullivan, "Overview of the H.264/AVC Video Coding Standard", IEEE trans on circuits and systems for video technology, Vol.13, No.7, July 2003
- [6] A.B.B. Adikari, W.A.C. Fernando, W.A.R.J. Weerakkody, "Independent Key Frame Coding Using Correlated Pixels in Distributed Video Coding," IEE Electronics Letters, Vol. 43, Issue 7, pp. 387-388, March 2007.
- [7] A. Aaron, R. Zhang and B. Girod, "Wyner-Ziv coding of motion video," Proc. Asilomar Conference on Signals and Systems, Pacific Grove, CA, Nov. 2002.
- [8] B. Girod , A. Aaron, S. Rane and D.R. Monedero, "Distributed Video Coding," Proc. of IEEE Special Issue on Advances in Video Coding and Delivery, Vol.93, No. 1, pp.1-12, 2003.
- [9] M. Tagliasacchi and S. Tubaro, "A MCTF Video Coding Scheme Based on Distributed Source Coding Principles," Visual Communication and Image Processing, July 2005.
- [10] P. Ishwar, V. Prabhakaran and K. Ramchandran, "Towards a Theory for Video Coding Using Distributed Compression Principles," Proc. of International Conference on Image Processing, Vol.3, pp.687-690, 2003.
- [11] L. Natario, C. Brites, J. Ascenso and F. Pereira, "Extrapolating Side Information for Low-Delay Pixel-Domain Distributed Video Coding," Int. Workshop on Very Low Bitrate Video Coding, Sept 2005
- [12] J. Ascenso, C. Brites and F. Pereira, "Motion Compensated Refinement for Low Complexity Pixel Based Distributed Video Coding," IEEE International Conference on Advanced Video and Signal Based Surveillance, Como, Italy, Sept 2005
- [13] J. Ascenso, C. Brites and F. Pereira, "Improving Frame Interpolation With Spatial Motion Smoothing For Pixel domain Distributed Video Coding", 5th EURASIP Conference on Speech and Image Processing, July 2005.
- [14] W.A.R.J. Weerakkody, W.A.C. Fernando, A.B.B. Adikari, R.M.A.P. Rajatheva, "4-PSK TCM for Wyner-Ziv frame coding in DVC," IEICE ELEX Electronics Letters, Vol. 3, No 8, April, 2006.
- [15] A.B.B. Adikari, W.A.C. Fernando, H.K. Arachchi, W.A.R.J. Weerakkody, "A Sequential Motion Estimation Based Refinement Technique for Distributed video coding of Wyner-Ziv frames," IEE Electronics Letters, Vol. 42, Issue 7, pp 396-397, March,2006.
- [16] A.B.B. Adikari, W.A.C. Fernando, W.A.R.J. Weerakkody, "Multiple Side Information Streams for Distributed Video Coding," IEE Electronics Letters, Vol. 42, Issue 25, pp. 1447-1449, December 2006.



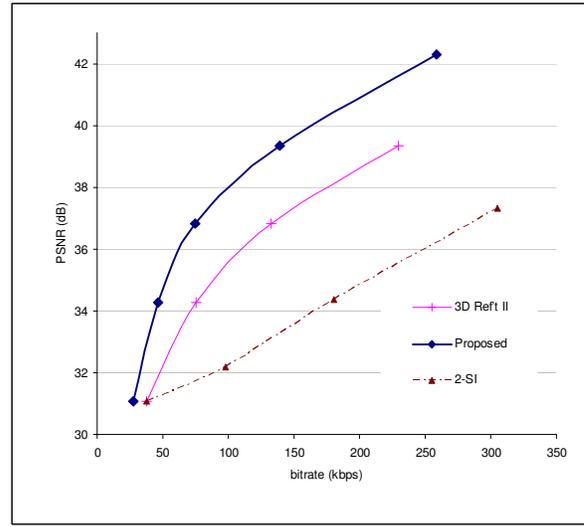
**Figure 1:** Functional elements of the algorithm for first bitplane coding



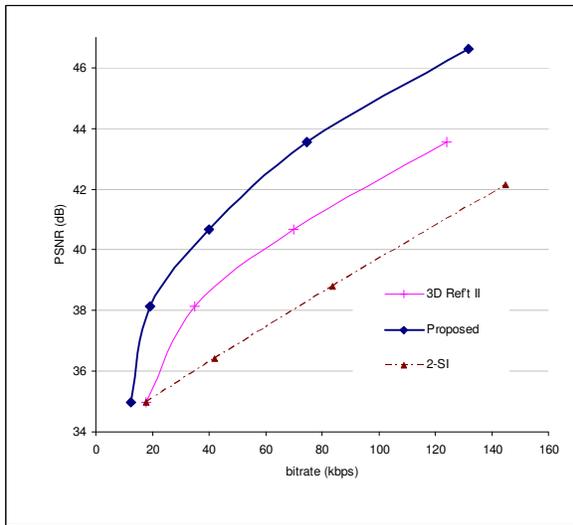
**Figure 2:** Functional elements of the algorithm for higher order bitplane coding



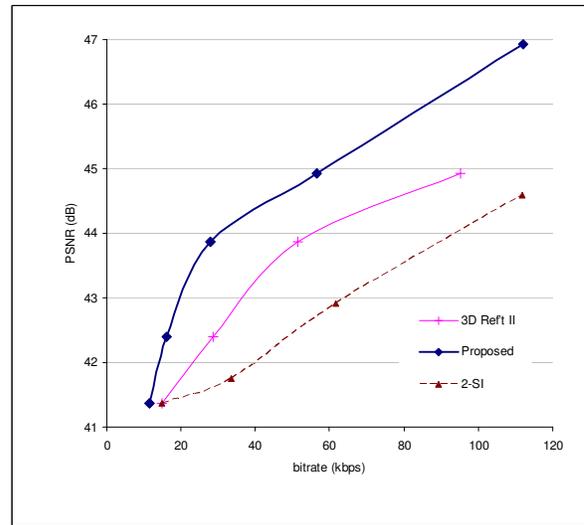
(a) "foreman"



(b) "carphone"



(c) "news"



(d) "claire"

Where; 2-SI: Ref. [16], 3D Ref't II: Ref. [1], MC-E: Ref [11].

**Figure 3:** Performance comparison of the proposed algorithm