

AN ENHANCED RECONSTRUCTION ALGORITHM FOR UNIDIRECTIONAL DISTRIBUTED VIDEO CODING

W.A.R.J. Weerakkody, W.A.C. Fernando, A.M. Kondo

University of Surrey, Guildford, Surrey GU2 7XH, UK
Email: R.Weerakkody@surrey.ac.uk

ABSTRACT

Distributed Video Coding (DVC) is an emerging video coding technology that utilizes the distributed source coding principles to build very low cost video encoders, yet with remarkable error resilience. In the common DVC framework, the reconstruction function plays a vital role that has a direct impact on the output video quality. In this paper, a novel algorithm is proposed for the reconstruction function, particularly focusing on the unidirectional DVC architecture. The proposed technique exploits the variations of the bit error rate of the Wyner-Ziv decoded bit stream and the side information stream. The simulation results show that the proposed algorithm yields a significant improvement of the objective and subjective video quality at no additional bit rate cost.

Index Terms— Distributed Video Coding, Wyner-Ziv coding

1. INTRODUCTION

With the increasing potential for video sensor networks and communications systems for a variety of diverse applications, there is a huge demand for more efficient and cost effective video coding solutions. Distributed Video Coding (DVC) has been widely acclaimed as a potential enabling technology for future video surveillance applications due to its flexibility for building low cost encoders and the significantly high built-in error resilience.

In the existing common DVC framework, *reconstruction* module plays a significant role as much as the *quantization* plays in contributing to the compression efficiency. When the pixel domain DVC architecture is concerned, while quantization effectively means limiting the number of bit planes transmitted per pixel, reconstruction is the inverse process that increase the precision of the decoded quantized pixels without an additional bit rate burden. This process utilizes the *side information* estimated at the DVC decoder for the Wyner-Ziv frames. The simple reconstruction model proposed by Aaron et al. [1] has been widely used in the DVC architectures. However, this approach lacked the

consideration of the statistical properties of the input, decoded, and side information bit streams. Further enhancements to the reconstruction algorithm were proposed in [2] and [3] which concentrated on the calculation of the expected value for the reconstructed pixel (or the DCT coefficient in transform domain coding) based on the Laplacian distribution properties of the residual error in side information with respect to the original video signal.

However, it is observed that these techniques are optimized for a DVC framework with a reverse channel, that formulates a closed loop feedback mechanism for determining the optimum parity bit rate, thus yielding a controlled very low bit error rate (BER) in the Wyner-Ziv decoded bit stream. However, having considered the practical constraints in involving a reverse feedback channel in video communication systems, increasing attention is focused on unidirectional DVC architectures; with no reverse channel. This development inevitably results in a higher, possibly uncontrolled, BER in the decoded bit stream. In this case, the reconstruction process could be better modeled as a statistical process involving the finite BER of the decoded bit stream as well as the side information stream. The objective of this paper is to propose a novel reconstruction algorithm for this scenario. A pixel domain DVC framework is utilized for the simulations, yet the concept is expandable for reconstructing the transform coefficients in transform domain coding.

2. BACKGROUND AND RELATED WORK

DVC technology is designed upon the distributed source coding principles established by Slepian and Wolf in 1973 in the so called Slepian-Wolf theorem [4], which explains the rate conditions for the independent encoding and joint decoding of statistically dependant discrete random sequences for lossless coding. Wyner-Ziv [5] extended this scenario for lossy coding in 1976 considering side information at the decoder, as followed in the current DVC architectures. Based on the above, a number of distinct DVC frameworks have been proposed in literature. The pixel domain design using turbo coding proposed by Aaron et al. commonly referred as the Stanford architecture [1] has been

followed and further enhanced, primarily in the side information estimation process in a number of research publications. Transform domain coding has also been proposed for enhanced compression efficiency even though with a little compromise of encoder complexity [6]. Other possible DVC implementations have also been proposed, including the PRISM architecture [7].

In the Stanford proposal for DVC, a basic algorithm was proposed for the reconstruction of pixels at the decoder utilizing the decoded bit planes of the pixels and the locally estimated side information. In this algorithm, the reconstructed pixel \hat{x} is calculated considering the corresponding side information value y and the corresponding quantization bin boundaries as:

$$\hat{x} = \begin{cases} z^-, & y < z^- \\ y, & z^- < y < z^+ \\ z^+, & y > z^+ \end{cases}$$

where z^- and z^+ represent the lower and upper edges of the quantization bin. This algorithm is widely utilized for its simplicity and *reasonable* performance.

More enhanced approaches for this purpose were presented in [2] and [3] known as the minimum mean square error (MMSE) techniques. In [2], Vatis et al. made use of the average statistical distribution of the transform coefficients. In other notable proposal, Kubasov et al. developed an algorithm to compute the reconstructed value \hat{x} as the expectation $E[x | x \in [z^-, z^+), y]$ of the random variable x . It is noted that both these approaches were optimized for systems with very small BER of the decoded bit stream achieved by involving the reverse channel for optimum parity bit requests.

3. PROPOSED SOLUTION

The proposed modification to the reconstruction algorithm is built upon the hypothesis that, when the BER of the decoded output increases, the relative significance of the side information increases in the reconstruction process. Thus the proposed algorithm is designed with particular focus on unidirectional architectures, unlike the previous proposals. This effect could be quantified considering the noise distributions of side information and the decoded output. Let's consider the case where the decoded quantized pixel falls outside the quantized side information bin. In this case, based on the above hypothesis, it is considered that the reconstructed pixel value can be modeled as a function of the side information value and the corresponding quantization bin boundaries as:

$$\hat{x} = \begin{cases} z^- + k(y - z^-) & , y < z^- \\ z^+ + k(y - z^+) & , y > z^+ \end{cases}$$

where k is an index reflecting the noise distribution of the decoded bit stream and the side information. A gross illustration of the proposed reconstruction model is shown in Figure 1. Based on this model, the derivation of the optimized gradient for the reconstruction line (call this *recon angle*) formulates an important research problem. This problem could be addressed by considering the statistical properties of the decoded and side information bit streams. It is further noted that these statistical properties varies significantly within each frame based on the regional motion activity. To exploit this, the recon angle is optimized for blocks of pixels. A training sequence is performed for experimentally determining the optimum recon angle for different BER levels of the decoded bit stream. Figure 2 shows the scatter plot of recon angles that yielded optimum PSNR for each block with the corresponding BER. Once this learning process is over, the resultant regression line is used to determine the recon angle for each block in a frame depending on the achieved BER for the block.

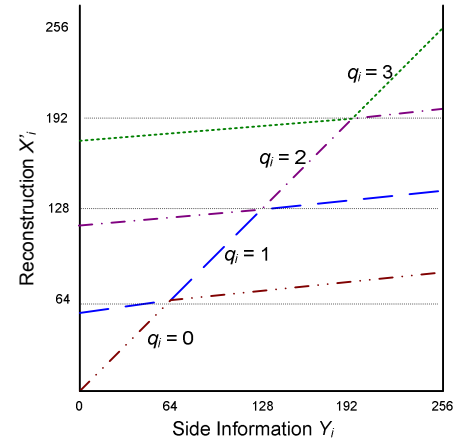


Figure 1: Proposed reconstruction model for 4 quantization levels

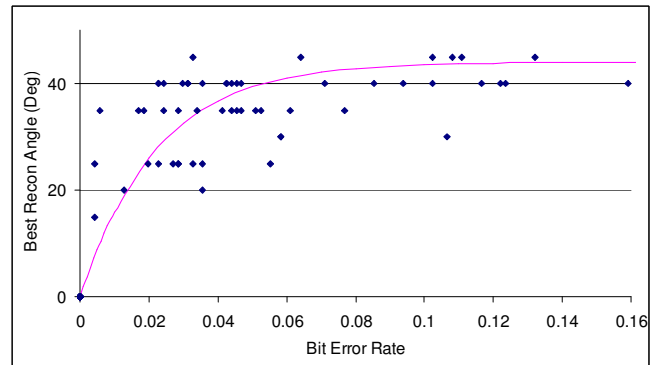


Figure 2: Scatter plot for the optimum recon angle for each block and the corresponding BER.
(Foreman sequence: 20 frames)

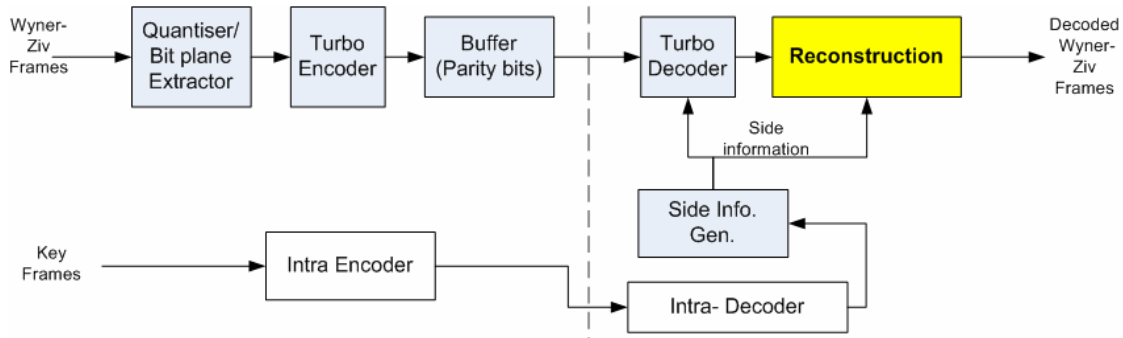


Figure 3: System block diagram

The proposed algorithm is tested on the pixel domain unidirectional DVC framework illustrated in Figure 3. Note that each frame is divided into blocks before encoding to exploit the regional variations of the statistical properties of the noise in the decoded bit stream.

4. SIMULATION RESULTS

The proposed reconstruction algorithm incorporated in to the pixel domain unidirectional DVC codec is tested for *Foreman* and *Hall Monitor* test video sequences, both QCIF at 15 Hz frame rate. Key frames are H.264-I coded (QP=26). The Wyner-Ziv bit rate is fixed at 1 bit per pixel (bpp). The optimum recon angle for each pixel is calculated based on the regression line described above in section 3 for each block. The block size is set to 352 bits. The performance is compared with the same DVC system incorporating the basic reconstruction model proposed in [1]. Figures 4 and 5 depict the comparative results for *Foreman* and *Hall Monitor* sequences respectively. Figure 6 graphically illustrates the perceptual quality improvement using a selected frame from the *Foreman* sequence.

The simulation results show a very promising and consistent improvement in PSNR of up to 4 dB for the *Foreman* and up to 2dB for the *Hall Monitor* sequences. It is noted that, in the unidirectional architecture, the decoded bit stream has residual errors that yield a notable degradation of video quality as seen in Figure 6. The proposed algorithm helps to solve these irritating effects by biasing the pixels with higher error probability towards the side information. This solution improves both objective (Figures 4 and 5) and subjective (Figure 6) video quality.

5. CONCLUSION

It is noted that when the Unidirectional DVC implementations are concerned, the decoded output tend to contain a notable BER that results in significant distortion on the reconstructed frame. In this paper, a novel reconstruction model was presented by considering the

statistical noise distribution of the decoded output of the DVC decoder. The performance of the proposed algorithm was verified for test video sequences and a notable gain in objective and subjective video quality is demonstrated.

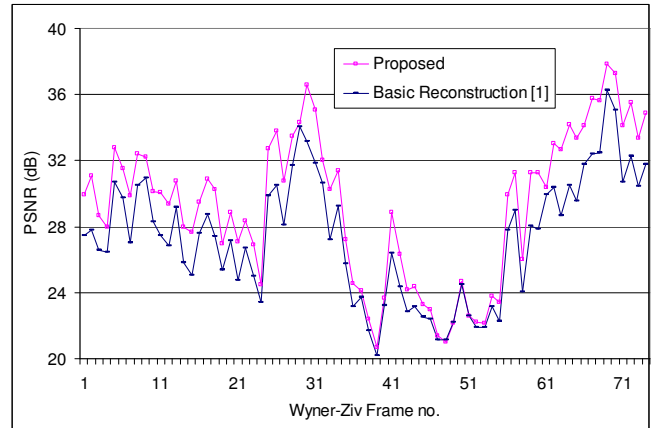


Figure 4: Performance comparison of the proposed algorithm (*Foreman* sequence)

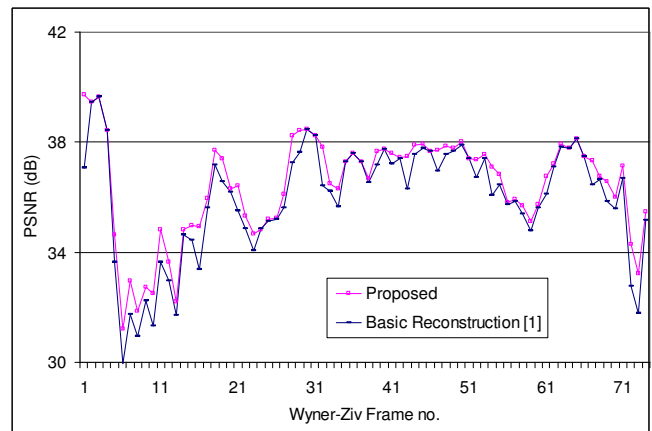


Figure 5: Performance comparison of the proposed algorithm (*Hall Monitor* sequence)



Figure 6: Performance comparison of the proposed algorithm (*Foreman* sequence: frame no.3)
 (a) without proposed model
 (b) with proposed model

ACKNOWLEDGEMENTS

The work presented was developed as a part of VISNET II, a European Network of Excellence (<http://www.visnet-noe.org>), funded under the European Commission IST FP6 programme.

REFERENCES

- [1] A. Aaron, R. Zhang, B. Girod, "Wyner-Ziv coding of motion video," Proceedings of Asilomar Conference on Signals and Systems, Pacific Grove, CA, Nov. 2002.
- [2] Y. Vatis, S. Klomp, J. Ostermann, "Enhanced reconstruction of the quantized transform coefficients for Wyner-Ziv coding", Proceedings of Int. Conf. on Multimedia & Expo (ICME), Beijing, China, July 2007.
- [3] Denis Kubasov, Jayanth Nayak, Christine Guillemot, "Optimal Reconstruction in Wyner-Ziv Video Coding with Multiple Side Information," Proceedings of Int. workshop on MMSP, Greece, October 2007.
- [4] D. Slepian, J.K. Wolf, "Noiseless coding of correlated information sources," IEEE Trans. on Inform. Theory, vol. IT-19, pp. 471–480, July 1973.
- [5] D. Wyner, 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.
- [6] C. Brites, J. Ascenso, F. Pereira, "Improving Transform Domain Wyner-Ziv Video Coding Performance," Proceedings of IEEE ICASSP, Vol 2, pp. II-525 - II-528, May 2006.
- [7] R. Puri and K. Ramachandran, "PRISM: A new robust video coding architecture based on distributed compression principles," Proceedings of Allerton Conference on Communications, Oct 2002.