ELectronics EXpress

IEICE Electronics Express, Vol.3, No.8, 178–183

4-PSK TTCM for Wyner-Ziv frame coding in DVC

W. A. R. J. Weerakkody[1a)], **W. A. C. Fernando**[1b)], **A. B. B. Adikari**[1c)], and **R. M. A. P. Rajatheva**[2d)]

[1] *School of Engineering and Design, Brunel University, Uxbridge, UB8 3PH, UK*

[2] *Telecommunications Programme, School of Engineering, Asian Institute of Technology, Thailand*

a) *Rajitha.weerakkody@brunel.ac.uk*

b) *anil.fernando@brunel.ac.uk*

c) *adikari.adikari@brunel.ac.uk*

d) *rajath@ait.ac.th*

Abstract: This paper proposes a novel approach for the implementation of Distributed Video coding (DVC) using 4-PSK turbo trellis coded modulation (TTCM). We have adapted the TTCM concept for source coding by generating parity at the encoder and transferring the symbol mapping to the decoder. The parity bits are sent to the decoder with puncturing as in a turbo based DVC codecs and combined with side information to form the symbols used in the TTCM decoder. The side information is generated by basic frame interpolation. The proposed codec was tested with different test video sequences and the results obtained show up to 6 dB improvement in PSNR with less memory for the same bit rate when compared with the turbo coding based DVC codecs.

Keywords: DVC, TTCM, turbo coding, Wyner-Ziv coding

Classification: Science and engineering for electronics

References

[1] D. Slepian and J. K. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. Inform. Theory*, vol. IT-19, pp. 471–480, July 1973.

[2] D. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," *IEEE Trans. Inform. Theory*, vol. IT-22, pp. 1–10, Jan. 1976.

[3] A. Aaron, R. Zhang, and B. Girod, "Wyner-Ziv coding of motion video," *Proc. Asilomar Conf. Signals and Syst.*, Pacific Grove, CA, Nov. 2002.

[4] B. Girod , A. Aaron, S. Rane, and D. R. Monedero, "Distributed Video Coding," *Proc. IEEE Special Issue Advances in Video Coding and Delivery*, vol. 93, no. 1, pp. 1–12, 2003.

[5] A. Aaron, E. Setton, and B. Girod, "Towards practical Wyner-Ziv coding of video," *Proc. IEEE Int. Conf. Image Processing*, ICIP-2003, Barcelona, Spain, Sept. 2003.

DOI: 10.1587/elex.3.178
Received February 03, 2006
Accepted March 20, 2006
Published April 25, 2006

IEICE Electronics Express, Vol.3, No.8, 178–183

1 Introduction

Video sensor networks are used today in a wide range of applications including, but not limited to, security surveillance, welfare services including child and elderly care, disaster zone monitoring, mobile communications and entertainment systems. These applications involve wide deployments of video cameras with compression and transmission capabilities in the field. Therefore they need be low cost in implementation and also be bandwidth efficient to use narrowband channels for transmission. Traditional video coding methods such as H264, MPEG-4 etc. cannot be used for these applications due to their high complexity in the encoder. Distributed video coding (DVC) is a solution to this problem which has a very simple encoder structure, possibly at the expense of a high complex decoder which can be placed at central sites for communication with multiple remote encoders. Wyner-Ziv approach is a solution for DVC problem which designates selected frames as key frames and Wyner-Ziv frames. Key frames are transferred directly to the decoder, using traditional compression techniques, to be used for the side information generation for Wyner-Ziv frames.

The closeness of the decoder output frames to the original video fames (measured in PSNR) and the transmission bit rate are used in the comparison of the goodness of the codecs.

2 Related work

The theoretical framework and the guidelines for Distributed Source Coding were established by Slepian-Wolf et al. [1] and the current work in this field is based on the work by Wyner-Ziv et al. [2]. According to the Slepian-Wolf theorem, two statistically dependent signals X and Y can be separately encoded and be jointly decoded with an arbitrarily small error probability if the following conditions are satisfied.

$$R_X \geq H(X/Y) \tag{1}$$

$$R_Y \geq H(Y/X) \tag{2}$$

$$R_X + R_Y \geq H(X,Y) \tag{3}$$

Where, $H(X,Y)$ is the joint entropy of X and Y, $H(X|Y)$ and $H(Y|X)$ are their conditional entropies.

Let's assume that $X_m(i,j)$ is the current Wyner-Ziv frame and $Y_m(i,j)$ is the correlated side information for the current frame generated based on Eq. (4) given below:

$$\begin{aligned} Y_m(i,j) = g(&X_{coded,m-N_2}(i,j), X_{coded,m-(N_2-1)}(i,j) \\ &\ldots X_{coded,m-1}(i,j), X_{coded,m+1}(i,j), \\ &\ldots X_{coded,m+(N_1-1)}(i,j), X_{coded,m+N_1}(i,j)) \end{aligned} \tag{4}$$

Where, $g(\,)$ is a function to describe the motion compensated prediction carried out using N_2 past reference frames and N_1 future frames. Then,

DOI: 10.1587/elex.3.178
Received February 03, 2006
Accepted March 20, 2006
Published April 25, 2006

EL ectronics
EX press

IEICE Electronics Express, Vol.3, No.8, 178–183

the relationship between $Y_m(i,j)$ and $X_m(i,j)$ can be modelled with a noise term, as shown in Eq. 5.

$$Y_m(i,j) = X_m(i,j) + n_m(i,j) \tag{5}$$

It can be shown that the noise term can be approximated to an additive stationary white noise signal, if the motion estimation is accurate. For most cases, this noise process can either be modelled using a Gaussian or using a Laplacian probability distribution.

Based on the above concept, several algorithms for Wyner-Ziv coding have been proposed recently. Aaron et al. has proposed a turbo coded based Wyner-Ziv codec for motion video [3, 4, 5]. They used simple frame interpolation [3] and motion interpolation and extrapolation techniques [5] to predict the side information. Based on this turbo based Wyner-Ziv codec, several side information prediction algorithms were proposed. It is clear that all these literature mainly considered the side information improvement based on the turbo Wyner-Ziv codec. In this paper, we concentrate on developing a TTCM based Wyner-Ziv codec for DVC.

3 Proposed system

Fig. 1 shows the block diagram of the proposed video codec implementation. The odd frames $\{X_1, X_3 \ ...\}$ are the Wyner-Ziv frames which go through the interframe encoder to generate the parity sequence to be transferred to the decoder. The even frames are directly passed to the decoder as Key-Frames using currently available compression techniques; we will not discuss the transfer of key-frames in this paper. The Wyner-Ziv frames are first passed through the 2^Q level quantiser where the level Q is an independently varied parameter based on the expected quality of output and the available channel bandwidth. Next, the Slepian-Wolf based encoder incorporates the bit plane extractor and then the turbo trellis encoder. Each rate 1/2 component encoder of our implementation has a constraint length $K = M + 1 = 4$ and a generator polynomial of [11 02] in octal form. A Pseudo-random interleaver of length 10 is used in front of the 2nd constituent encoder. Only the parity bit sequence thus generated is retained in the parity buffers and the systematic bits are discarded.

The decoder generates the side information using the Key-frames employing a pixel interpolation algorithm as below:

$$Y_m(i,j) = \frac{1}{2}\left(X_{m-1}(i,j) + X_{m+1}(i,j)\right) \tag{6}$$

This side information together with the parity bits passed from the encoder, upon request, form the 4-PSK symbols to be processed in the TTCM decoder. A multi level set partitioning is done with the constellation mapping of the TCM symbols in order to maintain the maximum Euclidian distance between the information bits. Where ever parity bits are not available due to puncturing being effective, the symbol automatically reduces to a lower modulation level. In the implementation under discussion, a combination

DOI: 10.1587/elex.3.178
Received February 03, 2006
Accepted March 20, 2006
Published April 25, 2006

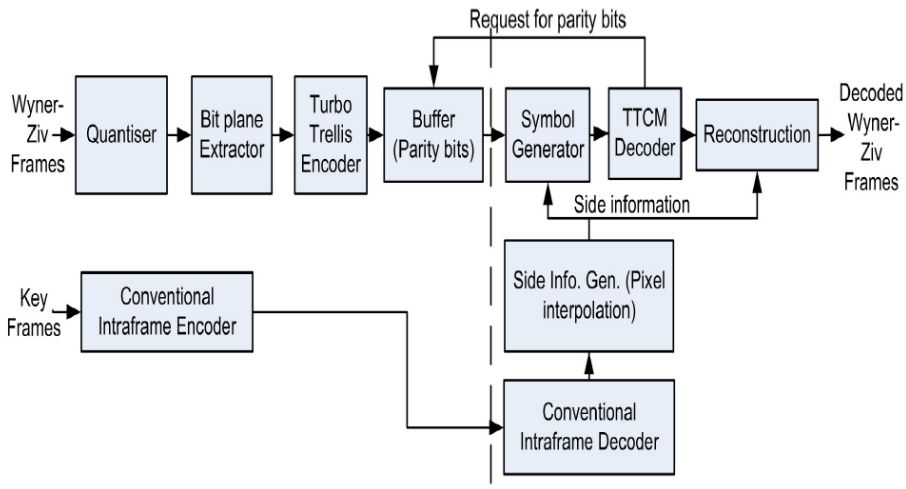

Fig. 1. Block Diagram of the proposed codec

of 4 PSK and Binary-PSK is used based on the availability of the parity bits for the constellation mapping. As commonly understood, Trellis Coded Modulation is conceptually a channel coding technique used to optimize the bandwidth requirements of a channel while protecting the information bits by increasing the size of the symbol constellation. Our effort is to exploit the high coding gain and the noise immunity inherent in this technique.

A symbol based MAP algorithm is used in the turbo trellis decoder which is run for 6 iterations as a complexity performance trade-off. A modification was done to the branch metric calculation to take care of the independent distributions of side information and parity bits. The parity bits are supplied to the decoder through an "on-demand" approach using a reverse channel for passing the request to the parity buffer maintained in the encoder. The parity patterns are predefined and known to both encoder and decoder. The pattern selection decision is taken at the decoder based on an error estimation performed on the decoded output. The process starts with minimum parity bits and additional bits are requested using signaling bits on the reverse feedback channel. The de-puncturer function in the decoder basically watches the parity availability and manipulates the symbols fed to the SISO based MAP decoder accordingly. A reconstruction function similar to the one proposed in [3] is used to smoothen some adverse effects in the output sequence including some contribution by the quantization noise.

As discussed earlier in this paper, the side information generated by the temporal interpolation of two key frames is assumed to be a form of the original Wyner-Ziv frame subjected to noise. The probability distribution of this noise was a part of the detailed study. It was noticed that both the Gaussian noise distribution and the Laplacian noise distribution resembled the interpolation noise with distinct variance parameters. However, most interestingly, it was noted that our implementation of the codec was not susceptible to errors by sub-optimal approximations of the distribution. For the purpose of taking the results, an additive White Gaussian noise (AWGN) with variance 0.125 is assumed.

DOI: 10.1587/elex.3.178
Received February 03, 2006
Accepted March 20, 2006
Published April 25, 2006

ELectronics EXpress

4 Results

The proposed codec was simulated for two test video sequences, "Foreman" & "Carphone". The results are obtained for the PSNR by varying the transmitted bit rate. The bit rate is varied by independently controlling the level M of quantiser where $2^M \in \{2, 4, 8, 16, 64\}$. The puncturing pattern is varied dynamically inside the codec in order to generate an optimum decoded output with respect to the quantized input fed to the encoder as described above. The methods of estimating the accuracy of the decoded bit stream are beyond the scope of this paper. The PSNR of the decoded frames is then calculated and plotted against the transmitted bit rate, averaged over the sequence.

The bit rate and PSNR are calculated for the luminance of the Wyner-Ziv frames, or the odd frames of the sequence, with a frame size of 176 × 144 (QCIF) at a Wyner-Ziv frame rate of 15 fps. The results for the "Foreman" & "Carphone" sequences are illustrated in Fig. 2 & Fig. 3 respectively (For all the simulations, we assumed the other conditions as in [3]). The simulation results for the proposed codec are compared with other video coding techniques currently available including that proposed by Aaron et al [3] using a turbo decoder with side information generated by average interpolation.

The results show up to 6 dB improvement in terms of the PSNR compared to that of a Turbo codec with equivalent side information generation used in our implementation. Our system was not simulated for side information improvement since our main focus is the codec improvement part rather than the overall outcome. Those plots including H263+ and H264 are reproduced only for the completeness in comparison.

Further, it is noted that our implementation of the encoder uses a memory length of 3 whereas the codec proposed in [3, 4] uses length 4. Shorter memory length means reduced computational complexity.

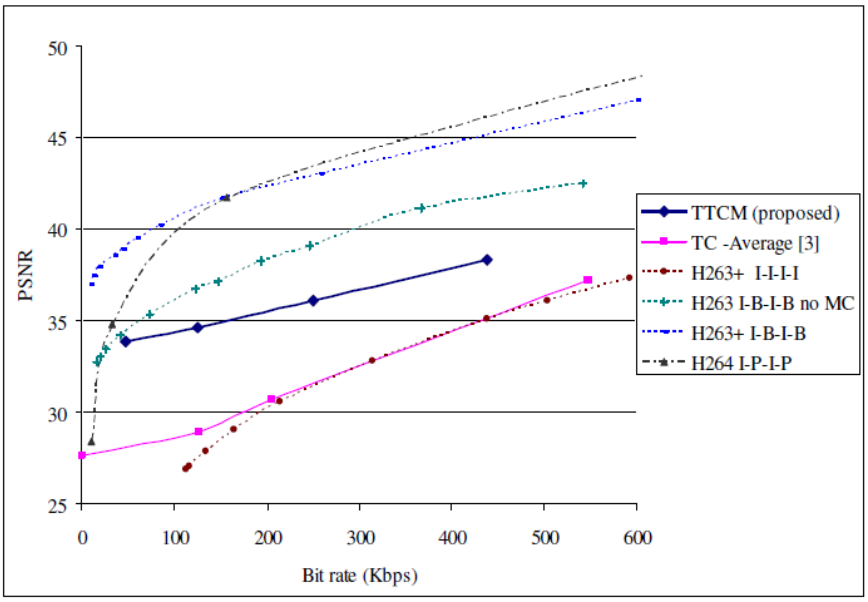

Fig. 2. PSNR vs. Bit rate for "Foreman" sequence

DOI: 10.1587/elex.3.178
Received February 03, 2006
Accepted March 20, 2006
Published April 25, 2006

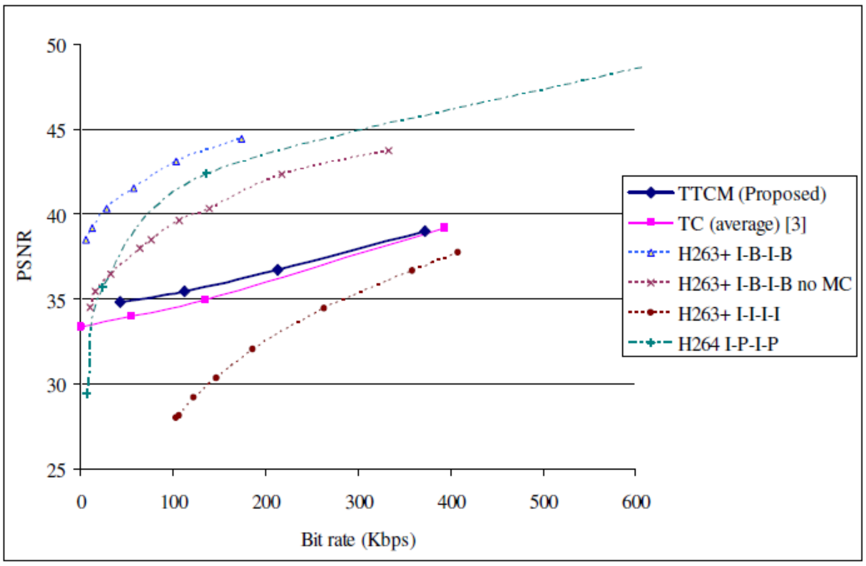

Fig. 3. PSNR vs. Bit rate for “Carphone” sequence

It is seen that the TTCM based codec performs significantly better irrespective of the motion level whereas turbo coding based implementation has a dependency on the motion level as evident when comparing the results for the two test video sequences used.

5 Conclusion

In this paper, we proposed a Wyner-Ziv video codec implementation based on TTCM principle. This is compatible with the DVC concepts with a very simple low cost encoder and a central decoder with high power processing capabilities. The simulations are applied on side information generated at the decoder by basic interpolation (average) of two Key-frames.

We have shown that the proposed codec outperforms the turbo coding based codec with a very high PSNR at a lower data rate, enabling to be used over low bandwidth transmission media with a higher viewing quality. Furthermore, it is noted that the proposed codec performs equally well irrespective of the motion level in the video sequence and is immune to the variations of the probability distribution of the interpolation noise introduced at the side information generation.

The bit rate and PSNR are calculated for the luminance of the Wyner-Ziv frames, or the odd frames of the sequence, with a frame size of 176 × 144 (QCIF) at a Wyner-Ziv frame rate of 15 fps.

DOI: 10.1587/elex.3.178
Received February 03, 2006
Accepted March 20, 2006
Published April 25, 2006