Transactions Letters

An Improved Decoding Algorithm for DVC Over Multipath Error Prone Wireless Channels

S. L. P. Yasakethu, W. A. R. J. Weerakkody, W. A. C. Fernando, F. Pereira, and A. M. Kondo

Abstract—This letter presents an improved decoding algorithm for distributed video coding (DVC) for enhanced performance over error prone wireless channels with multipath fading. The effects of the channel errors on both Wyner–Ziv and key frame bit streams are considered and novel noise models are proposed together with the associated modifications to the decoding algorithm. Simulations are performed using a W-CDMA wireless channel and the results are analyzed to determine the effect of each individual modification. State-of-the-art H.264/AVC is also considered under similar conditions for the comparison. The simulation results show that the proposed modifications provide a significant improvement in the DVC codec performance under unfavorable channel conditions.

Index Terms—Distributed video coding, multipath channels, Wyner–Ziv coding.

I. INTRODUCTION

DISTRIBUTED video coding (DVC) is an emerging video coding approach, particularly attractive due to its flexibility for building extremely simple and low cost video encoders. This feature could be very effectively exploited in several application domains including wireless sensor networks for security surveillance and mobile video communications, where the conventional video coding technologies with highly computational intensive algorithms become weak candidates.

It is a common scenario to deploy wireless networks for video surveillance networks, where the adverse effects of channel noise and time varying multipath fading on the reconstructed video quality are a significant problem. When DVC is concerned, even though its performance over noisy channels

has been the focus of some recent research, notably for packet-based networks [1], it is noted that the performance degradation due to fading with wireless channels has received limited attention. This letter considers this fading wireless channel scenario and modifies the DVC decoding algorithm to enhance the error resilience of the transform domain DVC codec. This discussion is based on a DVC codec that follows the *Stanford* framework [2] with turbo coding, where the Wyner–Ziv coded bit stream and the key frame stream traverses the communication channel. Thus the effects of the channel errors on both these streams are considered herein. Appropriate noise models will be developed for each component at the decoder and the suitable compensative solutions will be proposed.

II. SYSTEM OVERVIEW

The considered transform domain DVC codec architecture [3] with the WCDMA wireless channel simulation system is presented in this section. A block diagram of the system is shown in Fig. 1. The DVC encoder generates the punctured parity bit stream (Wyner–Ziv coded bit stream) [3] that is transmitted over the fading wireless channel. As shown in Fig. 1, a quadrature phase shift keying (QPSK) modulator and a demodulator with a built-in rake receiver are used in the transmitting and receiving nodes of the wireless channel, respectively.

In wireless communications, it is common to use a discrete time Rayleigh fading channel model consisting of multiple narrow-band channel paths separated by delay elements. With these channel models, it is common in communication systems to use rake receivers to improve the performance by collecting the energy dispersed over multiple fading paths. If $R_i(t)$ and $Y(t)$ correspond to the inputs and the output of the rake receiver and $G_1, G_2, \dots, G_N$ are the complex fading coefficients of each path, the output of the rake receiver $Y(t)$ with respect to signals received from N fading paths is computed as

$$R_i(t) = \sum_{j=1}^N (G_j X(t + 2 - i - j)) + n(t) \forall i \in \{1, 2, \dots, N\}$$

$$Y(t) = \sum_{i=1}^N G_i^* R_i$$

Manuscript received April 21, 2008; revised August 7, 2008 and November 10, 2008. First version published July 7, 2009; current version published September 30, 2009. This work, developed as a part of VISNET II, a European Network of Excellence (<http://www.visnet-noe.org>), was supported by the European Commission IST FP6 program. This paper was recommended by Associate Editor D. S. Turaga.

S. L. P. Yasakethu, W. A. R. J. Weerakkody, W. A. C. Fernando, and A. M. Kondo are with the Center for Communications Systems Research (CCSR), University of Surrey, Guildford, GU2 7XH, U.K. (e-mail: s.yasakethu@surrey.ac.uk; r.weerakkody@surrey.ac.uk; w.fernando@surrey.ac.uk; a.kondo@surrey.ac.uk).

F. Pereira is with the Instituto Superior Técnico, Instituto de Telecomunicações, Lisboa 1049-001, Portugal (e-mail: fp@lx.it.pt).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TCSVT.2009.2026799

Fig. 1. Block diagram of the DVC codec with the noisy channel models.

$$Y(t) = \left(\sum_{i=1}^N |G_i|^2 \right) X(t) + \sum_{i=1}^N \sum_{\substack{j=1 \\ i \neq j}}^N G_i^* \{ G_j X(t+2-i-j) + n(t) \}$$

$$Y(t) = \left(\sum_{i=1}^N |G_i|^2 \right) X(t) + n'(t) \quad (1)$$

where $X(t)$ is the transmitted symbol, G_i^* represent the complex conjugate values of the fading coefficient of path index i ($i = 1, 2, \dots, N$), and $n'(t)$ is the composite noise in the rake receive output.

Then, the QPSK demodulator computes the soft channel output for the parity bit stream to be forwarded to the turbo decoder for processing with the side information. The parity bit streams (y_k^p) for the decoder are taken from the soft output of the channel. The systematic component (y_k^s) is taken from the side information stream. The channel reliability factor (L_C) for the parity stream is derived from channel estimation in a practical implementation. For the simulations discussed in this letter, perfect channel estimation is assumed for the wireless channel. Any channel estimation algorithm proposed in the literature, such as [4], can be utilized for this purpose.

III. PROPOSED SOLUTION

In this section, we present the mathematical design of the novel noise model for the systematic and parity bit streams to be used in iterative MAP decoding of Wyner-Ziv frames in DVC.

The conditional log likelihood ratio (LLR) $L(u_k|y_k)$ of the data bit u_k is defined as [5]

$$L(u_k|y_k) = \ln \left(\frac{P(u_k = +1|y_k)}{P(u_k = -1|y_k)} \right) \quad (2)$$

where y_k is the received parity information. Incorporating the code's trellis, this may be written as

$$L(u_k) = \ln \left(\frac{\sum_{s^+} p(s_{k-1} = s', s_k = s, y)/p(y)}{\sum_{s^-} p(s_{k-1} = s', s_k = s, y)/p(y)} \right) \quad (3)$$

where $s_k \in S$ is the state of the encoder at time instance k , S^+ is the set of ordered pairs (s', s) corresponding to all state transitions $(s_{k-1} = s') \rightarrow (s_k = s)$ caused by data input $u_k = +1$, and S^- is similarly defined for $u_k = -1$.

From the BCJR algorithm,

$p(s', s, y) = p(s_{k-1} = s', s_k = s, y)$, is computed as

$$p(s', s, y) = \alpha_{k-1}(s') \cdot \gamma_k(s', s) \cdot \beta_k(s) \quad (4)$$

where $\alpha_k(s)$ and $\beta_k(s)$ are the forward and backward trellis state probabilities respectively, and $\gamma_k(s', s)$ is the state transition probability, defined as follows:

$$\alpha_k(s) = \alpha_{k-1}(v_1) \gamma_k(v_1, s) + \alpha_{k-1}(v_2) \gamma_k(v_2, s) \quad (5)$$

$$\beta_k(s) = \beta_{k+1}(w_1) \gamma_{k+1}(w_1, s) + \beta_{k+1}(w_2) \gamma_{k+1}(w_2, s) \quad (6)$$

$$\gamma_k(s', s) = P(s|s') p(y_k|s', s)$$

$$= P(u_k) p(y_k|u_k) \quad (7)$$

where v_1 and v_2 denote valid previous states and w_1 and w_2 denote valid next states for the current state.

It is reminded that, in the case of DVC, the parity stream for the turbo decoder is transmitted over a noisy wireless channel and the systematic stream is taken from the side information frame locally generated at the decoder which assumes a hypothetical Laplacian noise model. But in real world scenarios, both the transmitted parity and the intra coded key frame information suffer the same noisy wireless channel conditions. Thus, under practical situations, the effects of corrupted parity information, corrupted intra coded key frame information, and the hypothetical channel characteristics have been considered in our discussion.

Considering the noise distribution in parity and systematic components of the input to the turbo decoder to be independent, $p(y_k|u_k)$ in (7) can be written as

$$p(y_k|u_k) = p(y_k^p|u_k^p) p(y_k^s|u_k^s) \quad (8)$$

where $y_k = \{y_k^p, y_k^s\}$, $u_k = \{u_k^p, u_k^s\}$ are the received and transmitted parity and systematic information respectively. The mathematical modeling of received parity, and systematic information are elaborated in the following subsections.

Fig. 2. Noise modeling of systematic information (*Foreman* test sequence). (a) Original hypothetical Laplacian noise model for side information (DC coefficients). (b) Original pdf of DC coefficients. (c) Experimentally observed overall noise in side information, $h(x)$.

A. Modeling the Parity Information

For the parity bit sequence received through a multipath fading wireless channel, the probability of the received bit y_k^p conditioned on the transmitted bit u_k^p can be written as

$$p(y_k^p | u_k^p = \pm 1) = \frac{E_b}{\sqrt{2\pi\sigma^2}} \exp \left(-\frac{[y_k^p \mp \sum_{i=1}^N |a_i|^2]^2}{2\sigma^2} \right) \quad (9)$$

where a_i is the complex fading coefficient, σ is the standard deviation in the Gaussian probability distribution, E_b is the energy per bit, i is the path index, and N is the total number of multipaths.

The conditional LLR for parity information $L(y_k^p | x_k^p)$ can be written as

$$L(y_k^p | x_k^p) = \ln \left(\frac{P(y_k^p | x_k^p = +1)}{P(y_k^p | x_k^p = -1)} \right). \quad (10)$$

Substituting (9) into (10), we get

$$\begin{aligned} L(y_k^p | x_k^p) &= \ln \left\{ \frac{\exp \left(-\frac{E_b}{2\sigma^2} \left(y_k^p - \sum_{i=1}^N |a_i|^2 \right)^2 \right)}{\exp \left(-\frac{E_b}{2\sigma^2} \left(y_k^p + \sum_{i=1}^N |a_i|^2 \right)^2 \right)} \right\} \\ &= \left\{ \left(-\frac{E_b}{2\sigma^2} \left(y_k^p - \sum_{i=1}^N |a_i|^2 \right)^2 \right) - \left(-\frac{E_b}{2\sigma^2} \left(y_k^p + \sum_{i=1}^N |a_i|^2 \right)^2 \right) \right\} \\ &= \frac{E_b}{2\sigma^2} 4 \sum_{i=1}^N |a_i|^2 y_k^p = L_c \cdot y_k^p \end{aligned} \quad (11)$$

where $L_c = (E_b)/(2\sigma^2)4 \sum_{i=1}^N |a_i|^2$ is defined as the channel reliability value, and depends only on the channel SNR and the fading amplitude.

B. Modeling the Systematic Information

Under the error free transmission of intra coded key frames, the correlation of the original Wyner–Ziv frames to be encoded

and the estimated side information is assumed to be Laplacian distributed as

$$f(x) = \frac{\alpha}{2} \exp(-\alpha |y_k^s - u_k^s|) \quad (12)$$

where α is the distribution parameter. But experimentally we have found that in the presence of a noisy wireless channel, the correlation between the estimated side information and the original information no longer exhibits a true Laplacian behavior for the DC band of the DCT coefficients, even though the distributions of the AC bands can still be approximated as a Laplacian model. With the loss of information due to channel conditions, the new noise pdf of the DC band [denoted as $h(x)$ in Fig. 2(c)] is identified as the resultant additive effect of the original pdf of the DC band and the above Laplacian distribution, $f(x)$. This phenomenon is illustrated for *Foreman* test video sequence in Fig. 2. Tests were performed for a set of more than ten standard test video sequences and consistent results were observed. Thus, the conditional probability, $p(y_k^s | u_k^s)$ in (8), is derived from this additive channel model. However, the original DC coefficients of the encoded sequence for the current Wyner–Ziv frames are not available at the DVC decoder. Therefore, the pdf of the DC band distribution is estimated at the DVC decoder using the previously decoded Wyner–Ziv frame and is continuously updated during the decoding process.

Thus the conditional probabilities, $p(y_k^s | u_k^s)$, calculated by modeling the side information, and $p(y_k^p | u_k^p)$, calculated by modeling the parity information, as discussed in Section A, are used in (8) and the result is used in (7) to calculate the branch metric $\gamma_k(s', s)$.

Now, the SISO decoder is modified incorporating the proposed noise models for the MAP decoding algorithm as derived above. This modified turbo decoder is incorporated into the existing DVC decoder architecture [3].

When simulating the channel noise for the intra coded key frames transmitted over identical channel, the decoded key frames are too degraded for use as reference frames for side information estimation. Therefore, a channel coder [6] is incorporated into the intra codec in key frame coding. A low complex channel encoder [rate: 1/3, generator matrix $G(D) = [1 \ 1/(1+D)]$] is chosen not to overburden the DVC encoder and suitable channel encoders are incorporated, for both I and

B frames in H.264/AVC codec which will be discussed later in the comparison of results.

The reverse channel used in DVC for parity-request transmission is assumed to be free of channel errors. This assumption is justified by the very small bit rate utilized in the reverse channel and the possibility of building very strong error correction coding to protect it using the sufficiently available reverse channel bandwidth, in a common symmetric bi-directional channel.

IV. SIMULATION RESULTS

Simulations results of the DVC codec with the proposed modifications are discussed in this section. The test conditions are first explained followed by an analysis of results.

A. Test Conditions

The common test conditions used for all simulations are listed here. Specific conditions adopted in each analysis scenario are described in the following sections.

- 1) Three test video sequences are selected to represent different motion activity levels. *Soccer*: high motion, *Foreman*: medium motion, and *Hall Monitor*: low motion.
- 2) Frame format: YUV 4:2:0, resolution: QCIF (176×144), frame rate: 30 frames/s, no. of coded frames: 100, GOP: 2.
- 3) The transform (DCT) coefficients are quantized using the quantization matrices in [3].
- 4) Key frames are H.264/AVC Intra coded with the quantization parameter (QP) optimized for each rate-distortion (RD) point for comparable picture quality in Wyner-Ziv and key frames.
- 5) The PSNR is calculated for the luminance component.
- 6) H.264/AVC codec [7] is simulated in IBIB GOP structure (JM10.1/Main profile) for identical channel conditions. Here, a channel coder identical to that used for intra coding the key frames, as described in Section III, is used.
- 7) For both H.264/AVC and DVC codecs, the PSNR and bit rates are calculated for both intra coded and inter coded frames. The channel coding overhead bit rate is accounted.
- 8) A W-CDMA wireless channel is used for all simulations as summarized in Table I. A slow fading channel is considered to closely resemble a practical scenario.

B. Analysis of Results

The simulation results are analyzed in the next sections for the effects of proposed noise models, effects of channel coder selection for key frame coding, and the effects of the multiple fading paths, followed by a comparison with the state of the art in conventional video coding.

C. Effects of Noise Models

The performance of the proposed solution with respect to each modification (i.e., parity and side information) is shown

TABLE I
CONFIGURATION PARAMETERS OF THE W-CDMA CHANNEL

Channel model	W-CDMA uplink and downlink channels (slowly-varying freq.-selective fading)
Modulation	QPSK
Chip rate	3.84 Mchip/s
Spreading factor	32
Spreading sequence	OVSF sequences
Carrier freq.	2 GHz
Doppler speed	100 km/h

Fig. 3. RD results for the 2 path fading model (channel SNR = 3 dB), *Foreman*.

in Fig. 3 for the *Foreman* sequence, with multipath fading (two paths) and a channel SNR of 3 dB. It is observed that the DVC codec RD performance is improved by each individual modification. First, the proposed models for the parity (Wyner-Ziv bit stream) and systematic (side information) bit streams have been applied separately to verify the performance. The plots in Fig. 3 are described here in the order of performance improvement (bottom to top).

- 1) The original result without the proposed modifications.
- 2) The proposed model applied for the systematic (side information) bit stream only. The parity bit stream (Wyner-Ziv bit stream) noise model is not used. Here, up to 0.6 dB gain in PSNR is observed for the same bit rate when compared with the no-model result.
- 3) The proposed noise model applied only for the parity bit stream (Wyner-Ziv bit stream). Noise model for systematic bit stream is not used. Here, a PSNR gain of up to 1.5 dB is evident compared with the no-model result.
- 4) Both proposed noise models applied. An overall PSNR gain of up to 2 dB is evident here.

Performance of the proposed model with the *Hall Monitor* and *Soccer* sequences is shown in Figs. 4 and 5 where a PSNR gain of up to 2 dB is evident when both parity and systematic bit models are incorporated. Thus, the observations for the three different sequences are consistent proving that the positive gain of the proposed modification comes irrespective of the motion level in the video content.

Fig. 4. RD results for the 2 path fading model (channel SNR = 3 dB), *Hall Monitor*.

Fig. 5. RD results for the 2 path fading model (channel SNR = 3 dB), *Soccer*.

Fig. 6. Error free key frames using a strong channel code for intra coding (channel SNR = 3 dB), *Foreman*.

D. Effects of Channel Coder Selection for Key Frame Coding

Figs. 6 and 7 illustrate the effect of channel errors on the H.264/AVC Intra coded key frames. The overall performance of the codec is particularly sensitive to the quality of the key frames used in the generation of side information. Since a channel coder is utilized to protect the intra coded bit stream as discussed in Section III, this is effectively an analysis of the effect of channel coder selection on the proposed model. In Fig. 6, the intra coded frames are more protected using a stronger channel coder (code rate of $\frac{1}{4}$), thus enabling error free transmission. The other extreme scenario with highly

Fig. 7. Highly noisy key frames using a weak channel code for intra coding (channel SNR = 3 dB), *Foreman*.

Fig. 8. RD results for single path in fading model (channel SNR = 3 dB), *Foreman*.

corrupted intra frames as a result of less protection in using a very weak channel coder (code rate of $\frac{1}{2}$) is illustrated in Fig. 7. From Figs. 6 and 7, it is evident that a significant performance gain could be achieved through the proposed modifications irrespective of the key frame quality. It further quantifies the effect of the choice of channel coder, which is a compromise in terms of the overhead bit rate as well as the processing complexity.

E. Effects of the Multiple Fading Paths

The verification of the proposed model for the multi-path fading channel model is further extended by varying the number of transmission paths for a single path and a 4-path case. Simulation results are shown in Figs. 8 and 9, respectively, for the single path and 4-path fading channels and it is evident that the performance improvement from the proposed modifications is consistent for each case.

F. Comparison with the State-of-the-Art (H.264/AVC)

Finally, the results of the proposed model are compared with the state-of-the-art H.264/AVC codec (with default error concealment), for different channel SNR values. From the results presented in Fig. 10, it is evident that the proposed DVC codec demonstrates a very reliable and consistent performance over various channel conditions. It is further noted that the H.264/AVC codec is very sensitive to adverse channel conditions. The low complex channel encoder [rate:

Fig. 9. RD results for four paths in fading model (channel SNR = 3 dB), *Foreman*.

Fig. 10. DVC and H.264/AVC performance comparison for *Foreman* (2 path fading).

$1/3$, $G(D) = [1 \ 1/(1 + D)]$ used for DVC did not yield reliable performance for H.264/AVC. Accordingly, suitable new channel coders (only for H.264/AVC) were empirically determined for different SNR levels (for SNR = 3 dB: code rate = $1/3$, $G(D) = [1 \ (1 + D + D^2 + D^3)/(1 + D + D^3)]$, for SNR = 5 dB: code rate = $1/3$, $G(D) = [1 \ (1 + D^2)/(1 + D + D^2)]$ and for SNR = 9 dB: code rate = $1/3$, $G(D) = [1 \ 1/(1 + D^2)]$. Note here that the use of larger constraint lengths means significantly increased coding complexity. This approach was chosen over increasing the code rate (e.g., $1/4$) to keep the overall bit rate at a minimum.

The fact that the DVC codec with the proposed noise model gives very reliable performance at minimal cost of channel coding yields very promising prospects for the target applications discussed earlier in the introduction. This is, particularly, significant since many of those applications are implemented over a multipath fading wireless channel.

V. CONCLUSION

In this letter, novel noise models are proposed for a DVC codec and the decoder is accordingly modified for the use on noisy wireless channels. The performance of the proposed modifications is verified for a W-CDMA wireless channel with multipath fading. The simulation results clearly show that the two proposed noise models for the parity (Wyner–Ziv bit stream) and systematic (side information) bit streams have resulted in a significant gain in the RD performance of the

codec when operated over noisy and fading channel conditions. The improvement is proven to be significant irrespective of the motion level in the video content, the conditions of the fading channel or the transmission accuracy of the key frames. Due to its very stable and consistent behavior, DVC even outperforms the H.264/AVC codec when the channel effects are adverse. Furthermore, it is noted that these results are obtained with no additional complexity burden to the already extremely low complex DVC encoder.

REFERENCES

- [1] J. Pedro, L. Ducla Soares, C. Brites, J. Ascenso, F. Pereira, C. Bandeira, S. Ye, F. Dufaux, and T. Ebrahimi, "Studying error resilience performance for a feedback channel based transform domain Wyner–Ziv video codec," in *Proc. Int. Picture Coding Symp.*, Lisbon, Portugal, Nov. 2007.
- [2] A. Aaron, R. Zhang, and B. Girod, "Wyner–Ziv coding of motion video," in *Proc. Asilomar Conf. Signals Syst.*, Pacific Grove, CA, Nov. 2002, pp. 240–244.
- [3] C. Brites, J. Ascenso, and F. Pereira, "Improving transform domain Wyner–Ziv video coding performance," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, vol. 2, May 2006, pp. 525–528.
- [4] G. Montalbano and D. T. M. Slock, "Joint common-dedicated pilots based estimation of time-varying channels for W-CDMA receivers," in *Proc. IEEE Vehicular Technol. Conf.*, vol. 2, 2003, pp. 1253–1257.
- [5] W. E. Ryan, "A turbo code tutorial," in *Proc. IEEE Int. Conf. Global Telecommun.*, 1998. Available: <http://www.ece.arizona.edu/~ryan/publications/turbo2c.pdf>
- [6] S. Lin and D. J. Costello, Jr., "Turbo coding," in *Error Control Coding*, 2nd ed. New York: Prentice-Hall, 2004, ch. 16, pp. 810–811.
- [7] Joint Video Team of ISO/IEC MPEG and ITU-T VCEG, *H.264/AVC*, Reference Software JM10.1 [Online], Available: http://iphome.hhi.de/suehring/ttml/download/old_jm/jm10.1.zip