


Efficient decoding algorithm for 3D video over wireless channels

S. L. P. Yasakethu¹ · C. T. E. R. Hewage² 

Received: 29 September 2017 / Revised: 21 March 2018 / Accepted: 16 May 2018 /

Published online: 31 May 2018

© The Author(s) 2018

Abstract In wireless communication networks, additive channel noise and time varying multipath fading are considered amongst the main challenges and these effects will bring significant quality degradation to the transmitted 2D/3D video streams. This paper presents an efficient decoding algorithm for Distributed Video Coding (DVC) for enhanced performance of colour and depth based 3D video over error prone wireless channels. The maximum a-posteriori (MAP) algorithm for turbo decoding is modified considering the effects of the channel errors on both Wyner-Ziv and key frame bit streams of colour and depth videos. The proposed codec is simulated using a W-CDMA wireless channel model and the results are analyzed to determine the effect of the modifications. The performance of the state-of-the-art H.264/AVC video codec is also presented for comparison under similar conditions. The results show that the proposed modifications provide a significant improvement in the DVC codec performance under unfavourable channel conditions for colour plus depth based 3D video.

Keywords 3D video · DVC · Wireless channels · Quality of Experience (QoE)

1 Introduction

The 3D video has gone one step beyond the conventional 2D video by providing the depth perception of the scene to the users. These formats of enhanced multimedia content have the potential to render high quality immersive experiences (i.e., immersive QoE) for a variety of applications in home, in workplace, and in public spaces. Distributed Video Coding (DVC) is

✉ C. T. E. R. Hewage
chewage@cardiffmet.ac.uk

S. L. P. Yasakethu
lasithy@sltc.edu.lk

¹ Department of Electrical, Electronic and Computer Science, Sri Lanka Technological Campus (SLTC), Padukka, Sri Lanka

² Cardiff School of Technology, Cardiff Metropolitan University, Cardiff, UK

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 3D video application domains including wireless sensor networks for 3D mobile video communications and future 3D video surveillance systems, where the conventional video coding technologies, with highly computational intensive algorithms become weak candidates (the usage in sensor networks).

A 3DTV broadcasting approach was proposed by the European Information Society Technologies (IST) project ‘Advanced Three-Dimensional Television System Technologies’ (ATTEST), [9]. The 3D image/video representation used in this scenario is based on a monocular colour video and an associated per-pixel depth map. An example of this representation is illustrated in Fig. 1 for the “Orbi” test 3D video sequence. Depth Image Based Rendering (DIBR) methodology utilizes this colour and its depth videos to generate artificial stereoscopic views, one for each eye of a viewer. The main advantage of DIBR technique compared to traditional representation of 3D video with the left-right views is that it provides high quality 3D video with smaller bandwidth required for transmission [6–8, 10]. It is a common scenario to deploy wireless networks for 3D video sensor networks, where the adverse effects of channel noise and time varying multipath fading on the reconstructed video quality is 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 [15], it is noted that the performance degradation due to fading with wireless channels has received limited attention. This paper considers this fading wireless channel scenario and modifies the DVC decoding algorithm to enhance the decoded quality of 3D video (rendered left and right videos) of the transform domain DVC codec. This discussion is based on a DVC codec that follows the Stanford framework [1] with turbo coding, where the Wyner-Ziv coded bit stream and the key frame stream of both colour and depth videos traverses the communication channel. Thus the effects of the channel errors on both Wyner-Ziv and key frame streams are considered herein. Appropriate noise models will be developed for each stream at the decoder and the suitable compensative solutions will be proposed.

2 System overview

The considered transform domain DVC codec architecture [4] with the W-CDMA wireless channel simulation system is presented in this section. Colour and depth videos are encoded

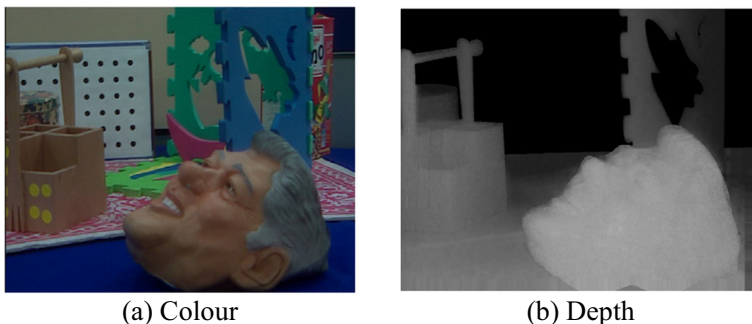


Fig. 1 A Frame from the “Orbi” test sequence

independently at the DVC encoder and are transmitted over the wireless channel. After independently decoding both colour and depth videos DIBR is used to generate stereoscopic views (left and right) at the decoder. A block diagram of the system is shown in Fig. 2. It should be noted that this architecture is common to both colour and depth videos of 3D content. The DVC encoder generate separate punctured parity bit streams (Wyner-Ziv coded bit stream) [4] for colour and depth videos and are transmitted over the fading wireless channel. As shown in Fig. 2, a quadrature phase shift keying (QPSK) modulator and a demodulator with 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 as illustrated in Fig. 3. 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. In Fig. 4, which illustrates the rake receiver, $R(t)$ and $Y(t)$ correspond to the inputs and the output of the rake receiver. $G_1, G_2 \dots G_N$ 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:

$$\begin{aligned}
 R_i(t) &= \sum_{j=1}^N (G_j X(t + 2^{-i-j})) + n(t), \quad \forall i \in \{1, 2, \dots, N\} \\
 Y(t) &= \sum_{i=1}^N G_i^* R_i \\
 Y(t) &= \left(\sum_{i=1}^N |G_i|^2 \right) X(t) + \sum_{i=1}^N \sum_{\substack{j=1 \\ i \neq j}}^N G_i^* \{G_j X(t + 2^{-i-j}) + n(t)\} \\
 Y(t) &= \left(\sum_{i=1}^N |G_i|^2 \right) X(t) + n'(t)
 \end{aligned}
 \tag{1}$$

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.

3 Proposed solution

In this section, we present the mathematical design of the novel noise model for the systematic and parity bit streams of colour and depth videos to be used in iterative MAP decoding of

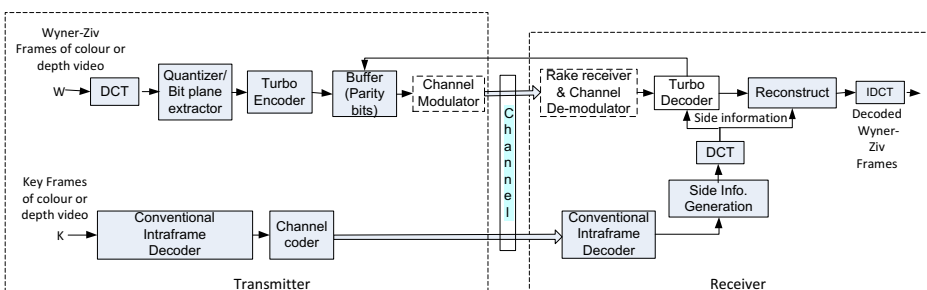


Fig. 2 Block diagram of the DVC codec with the noisy channel models

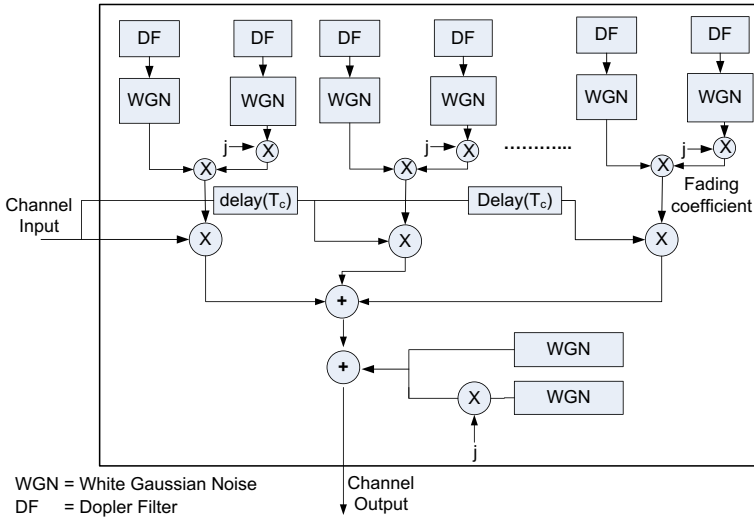


Fig. 3 Rayleigh fading channel model

Wyner-Ziv frames in DVC. It is important to note that all the notations and derivations in the paper are common to both colour and depth videos, unless or otherwise stated.

The conditional Log Likelihood Ratio (LLR) $L(u_k|y_k)$ of the data bit u_k , is defined as [16]:

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

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

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

where $s_k \in S$ is the state of the encoder at time instance k , S^+ is the set of ordered pairs (s', s)

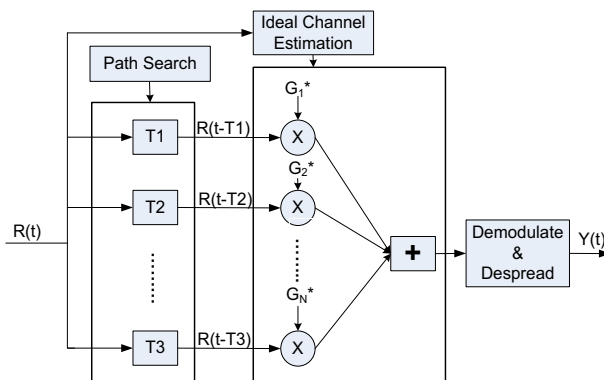


Fig. 4 Block diagram of the rake receiver

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 Bahl-Cocke-Jelinek-Raviv (BCJR) algorithm [16], $p(s', s, y) = p(s_{k-1} = s', s_k = s, y)$, is computed as,

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

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:

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

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

$$\begin{aligned} \gamma_k(s', s) &= P(s|s') p(y_k | s', s) \\ &= P(u_k) p(y_k | u_k) \end{aligned} \tag{7}$$

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 for both colour and depth videos.

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 Eq. (7) can be written as:

$$p(y_k | u_k) = p(y_k^p | u_k^p) p(y_k^s | u_k^s) \tag{8}$$

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 modelling of received parity, and systematic information are elaborated in the following sub-sections.

3.1 Modelling the parity information for colour and depth videos

For both colour and depth parity bit sequences 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:

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

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:

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

Substituting (9) in (10), we get:

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

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

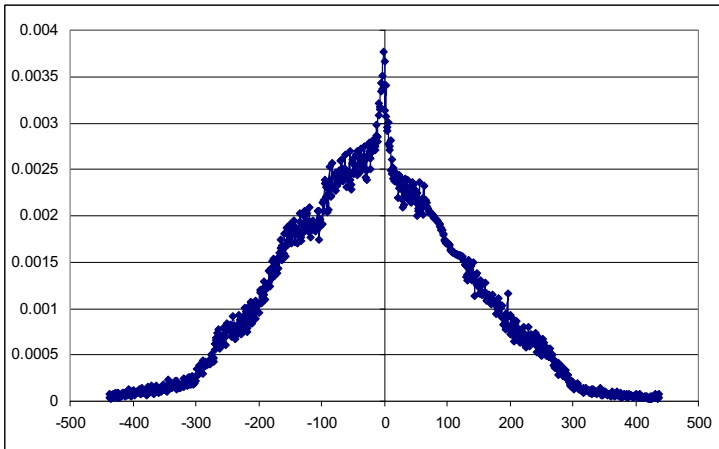
3.2 Modelling the systematic information for colour and depth videos

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,

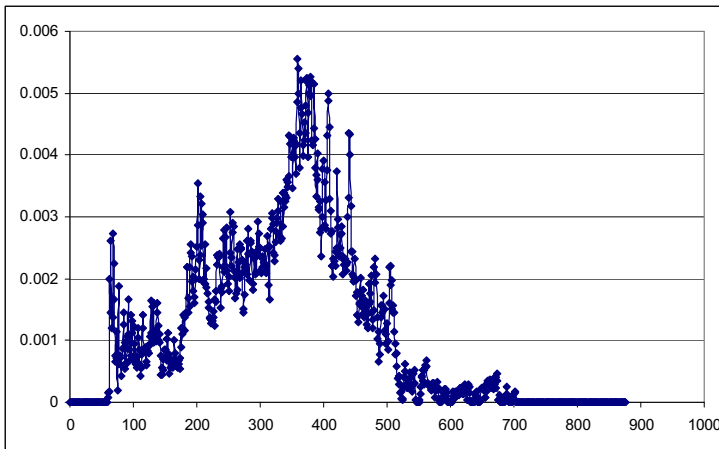
$$f(x) = \frac{\alpha}{2} \exp(-\alpha |y_k^s - u_k^s|) \tag{12}$$

where α is the distribution parameter. But experimentally we have found that for the colour video, 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 behaviour 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 of the colour video (denoted as $h(x)$ in Fig. 5c) 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 *Interview* test video sequence in Fig. 5. Tests were performed for several standard test video sequences and consistent results were observed. Thus, the conditional probability, $p(y_k^s|u_k^s)$ in Eq. (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. For the depth video, it is observed that, the correlation between the estimated side information and the original information for all the DC and AC bands can still be approximated as a Laplacian model.

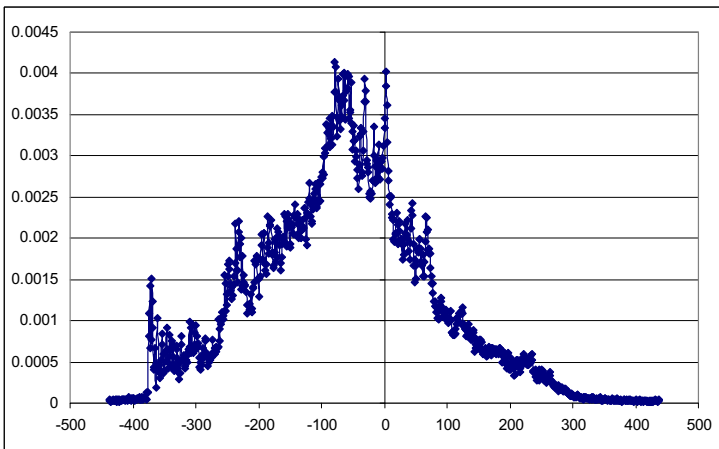
Thus the conditional probabilities, $p(y_k^s|u_k^s)$, calculated by modelling the side information, and $p(y_k^p|u_k^p)$, calculated by modelling the parity information, for colour and depth videos as discussed in section 3, are used in Eq. (8) and the results are used in Eq. (7) to calculate the branch metric $\gamma_k(s', s)$.



(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, $h(x)$

Fig. 5 The noise modelling of systematic information (Colour video of *Interview* test sequence)

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 to the existing DVC decoder architecture [4].

When simulating the channel noise for the intra coded key frames, of both colour and depth videos transmitted over identical channel, the decoded key frames are too degraded for use as reference frames for side information estimation. Therefore, a channel coder [12] is incorporated to 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.

4 Simulation results and discussion

Simulation 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.

4.1 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. Three test video sequences are selected to represent different motion characteristics. “*Breakdancers*” (high motion/texture variation with static camera, at 15fps) “*Orbi*” (medium to high motion/texture variation and parallel camera motion, at 25fps) and “*Interview*” (low motion/texture variation with static camera, at 25fps). Colour video is represented in YUV 4:2:0 and the depth video is represented using the luminance component with 8-bit grey values. Where grey value 0 specifies the furthest value (i.e. away from camera) and the grey value 255 specifies the closest value (i.e. closer to the camera). Resolution is 176×144 and number of coded frames is 100 with a GOP of 2. The transform (DCT) coefficients are quantized using the quantization matrices in [4]. Key frames of colour and depth videos 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. H.264/AVC codec [11] 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 3.2, is used. For both H.264/AVC and DVC codecs, the PSNR and bit rates are calculated for both intra coded and inter coded frames of colour and depth videos. Then the average PSNR of the rendered left and right videos and the total bit rate for colour and depth videos (The channel coding overhead bit rate is accounted.) are calculated and are presented in the simulation results. Previous studies by the authors have demonstrated the relationship between average PSNR and true 3D image quality [6–8]. These studies provide correlation figures (mapping using a logistic function proposed in ITU-R/BT 500–13 [10]) to show how much the average PSNR of

left and right video is correlating with MOS provided by the users. However, to provide the performance parameters against true 3D video quality, subjective quality evaluation tests were conducted. This will illustrate how significant the achieved results are in terms of true 3D user perception. The subjective test conditions and parameters are listed below.

The 42" Philips multi-view auto-stereoscopic display was used in the experiment to display the stereoscopic material. The optics of this display were optimized for a viewing distance of 3 m. Hence, the viewing distance for the observers was set to 3 m. This viewing distance was in compliance with the Preferred Viewing Distance (PVD) of the ITU-R BT.500-13 recommendation [10], which specifies the methodology for subjective assessment of the quality of television pictures. The 3-D display was calibrated using a GretagMacbeth Eye-One Display 2 calibration device. Peak luminance of the display was 200 cd/m². The measured environmental illumination was 180 lx, which was slightly below the ITU-R BT.500-13 recommended value for home environments (i.e. 200 lux) [10]. The background luminance of the wall behind the monitor was 20 lx. The ratio of luminance of inactive screen to peak luminance was less than 0.02. These environmental luminance measures remained the same for all test sessions, as the lighting conditions of the test room were kept constant. Thirty two non-expert observers (ten female observers and 22 male observers) volunteered to participate in the experiment. They were divided into two groups in order to assess one attribute per group. The attribute assessed in this experiment was overall perceived image quality. The observers were mostly research students and staff with a technical background. Their ages ranged from 20 to 40. Fourteen observers had prior experience with 3-D video content using different viewing aids such as shutter glasses and red-blue anaglyph glasses. All participants had a visual acuity of ≥ 1 (as tested with the Snellen chart), good stereo vision < 60 s of arc (as tested with the TNO stereo test), and good colour vision (as tested with the Ishihara test).

A W-CDMA wireless channel is used for all simulations. A slow fading channel is considered to closely resemble a practical scenario (Modulation: QPSK, Chip rate: 3.84Mchip/s, Spreading factor: 32, Spreading sequence: OVVSF, Carrier freq.: 2GHz, Doppler speed: 100 km/h).

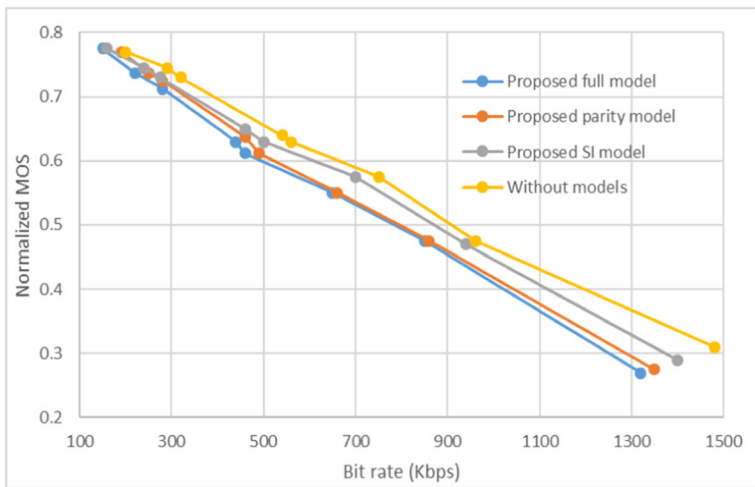
4.2 Analysis of results

The simulation results are analyzed in the following sub-sections for the effects of proposed noise models and the effects of the multiple fading paths, followed by a comparison with the state of the art in conventional video coding.

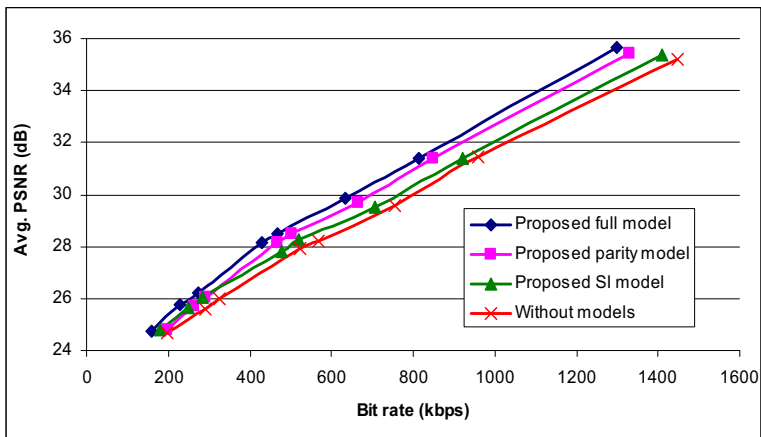
4.2.1 Effects of noise models

The performance of the proposed solution with respect to each modification (i.e. parity and side information) is shown in Fig. 6a and b for the *Orbi* sequence, with multipath fading (2 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 of colour and depth videos have been applied separately to verify the performance. The plots in Fig. 6a and b are described here in the order of performance improvement (bottom to top):

1. The original result without the proposed modifications



(a)



(b)

Fig. 6 a RD results (Normalized MOS) for the 2 path fading model (channel SNR = 3 dB), *Orbi* test sequence. b. RD results for the 2 path fading model (channel SNR = 3 dB), *Orbi* test sequence

2. The proposed model applied for the systematic (side information) bit stream only. Parity bit stream (Wyner-Ziv bit stream) noise model is not used. Here, up to 0.5 dB gain in average PSNR of rendered views is observed for the same bit rate when compared with the no-model result (see Fig. 6b).
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, an average PSNR gain of up to 1.5 dB of rendered views is evident compared with the no-model result.
4. Both proposed noise models applied. An overall average PSNR gain of up to 1.8 dB is evident here. Similar gains can be visible in terms of true 3D video quality measures in Normalized MOS (see Fig. 6a)

Performance of the proposed model with the *Interview* and *Breakdancers* sequences are shown in Figs. 7a/b and 8a/b, where a similar Normalized MOS/Average PSNR gain 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 3D video content.

4.2.2 Effects of the multiple fading paths

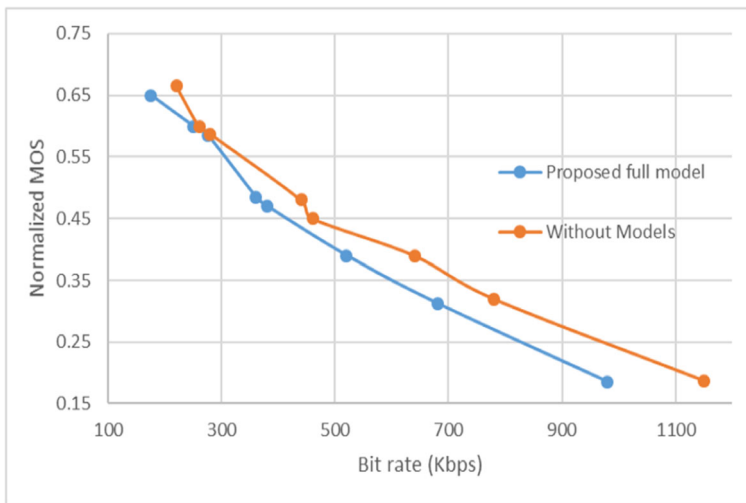
The verification of the proposed model for the multipath fading channel model is further extended by varying the number of transmission paths for a single path and a 4-path case. Simulation results are shown in Figs. 9a/b and 10a/b, respectively for the single path and 4-path fading channels and it is evident that the performance improvement from the proposed modifications is consistent for each case.

4.2.3 Comparison with the state-of-the-art (H.264/AVC)

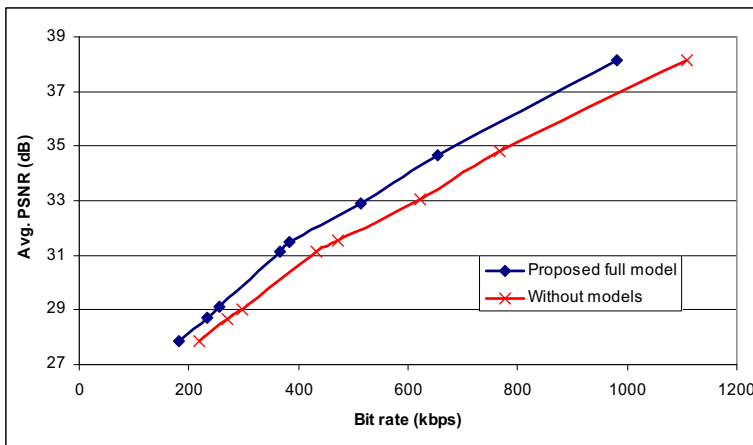
Finally, the results of the proposed model, are compared with the state-of-the-art H.264/AVC codec (with default error concealment), for different channel SNR values. From the results presented in Fig. 11a and b, it is evident that the proposed DVC codec demonstrates a very reliable and consistent performance 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: $1/3$, $G(D) = [1 \ 1/(1+D)]$) 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.

4.3 Related work

Several optimization frameworks have been proposed in the literature for distributed processing or handling of data to maximize the performance under given resource limitations [2, 3, 13, 17–19]. AutoReplica [17] provides a solution to effectively balance the trade-off between I/O performance and fault tolerance using SSD-HDD tier storages. Primarily this explains replication across multiple storages to achieve fault-tolerance. However, this does not explain how high end applications such as next generation multimedia applications (e.g., 3D video) could use similar distributed architectures for distributed processing. This optimization framework proposed in [2] presents a performance approximation approach to model the computing performance of iterative, multi-stage applications running on a master-computer framework. Whilst 3D video could be considered as a multi-stage application, this work is mainly for performance evaluation of such applications rather than end user quality, which is the main focus of our article. Authors in [18] explains automated framework for data placement in multi-tier all-flash



(a)

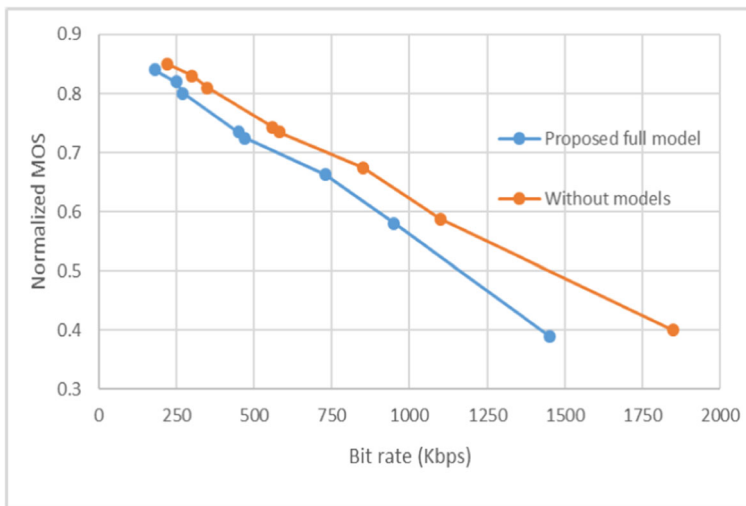


(b)

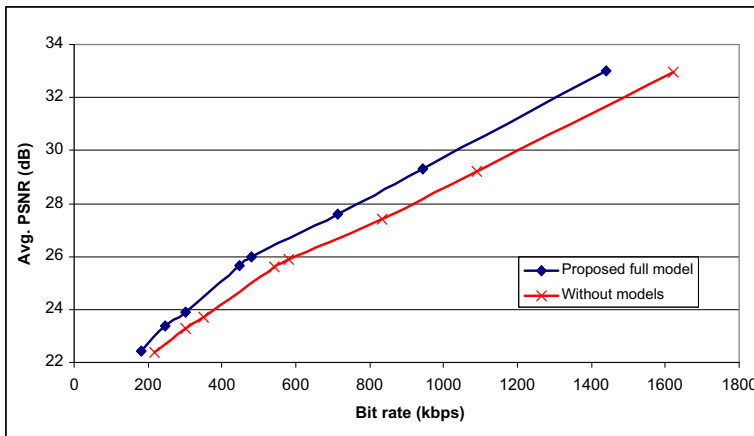
Fig. 7 a RD results (Normalized MOS) for the 2 path fading model (channel SNR = 3 dB), *Interview* test sequence. b. RD results (Average PSNR) for the 2 path fading model (channel SNR = 3 dB), *Interview* test sequence

data centres. This could be useful in multi-view 3D video applications where the importance of different views arise at different times and the proposed work can be extended to be used with such back end framework.

3D and multi-view video applications in proposed DVC architectures can be found in the literature [5, 14]. These work explain how DVC can be used instead of conventional video coding architectures which has high complexity at the encoder side due to motion prediction etc. At the same time, these articles highlight the challenges faced by DVC architectures to enable 3D/multi-video video. For example Side-Information (SI) generation for the same views as well as for adjacent views.



(a)

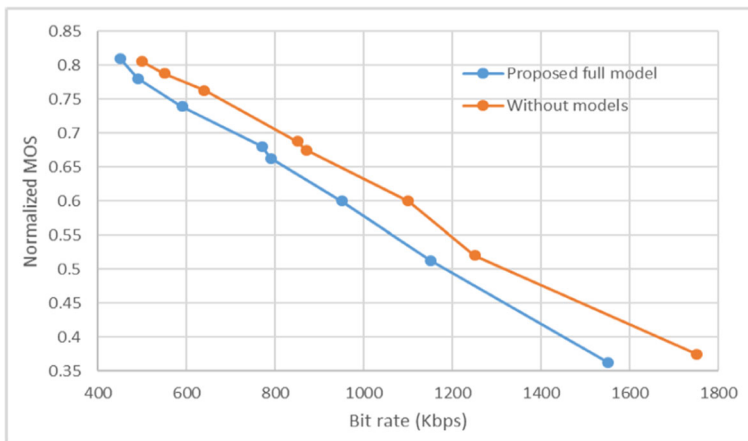


(b)

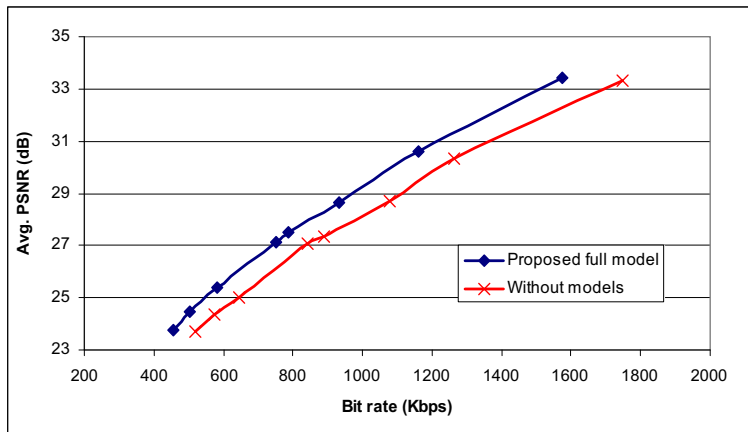
Fig. 8 **a** RD results (Normalized MOS) for the 2 path fading model (channel SNR = 3 dB), *Breakdancers* test sequence. **b** RD results (Average PSNR) for the 2 path fading model (channel SNR = 3 dB), *Breakdancers* test sequence

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 the applications are implemented over multipath fading wireless channel.

The improvement is proven to be significant irrespective of the motion level in the video content, or the conditions of the fading. Due to its very stable and consistent behaviour, 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.



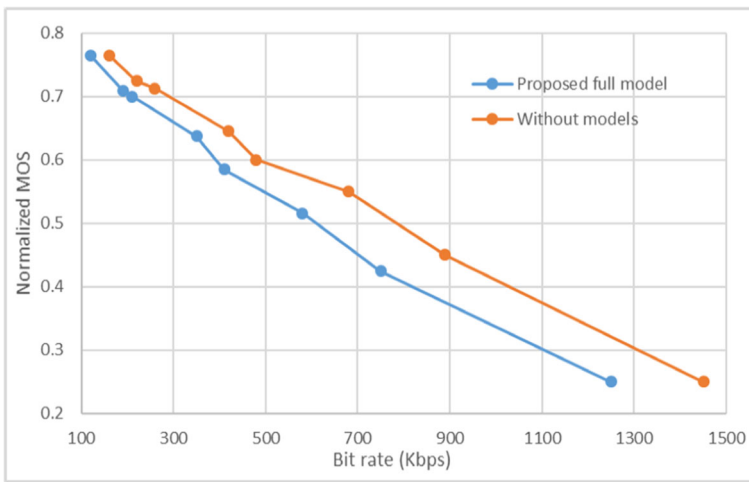
(a)



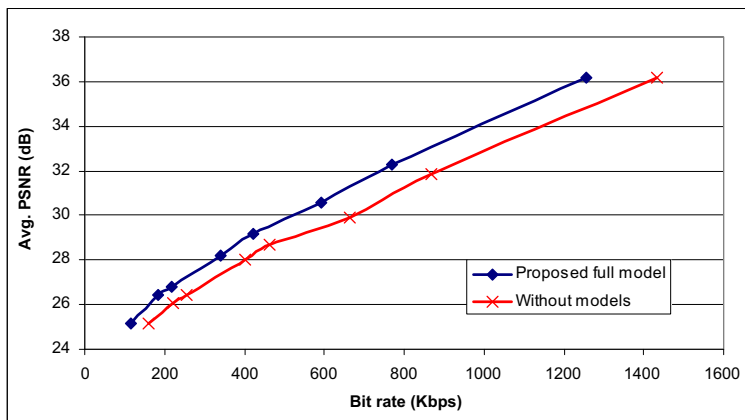
(b)

Fig. 9 a. RD results (Normalized MOS) for single path in fading model, (channel SNR = 3 dB), *Orbi* test sequence. b. RD results (Average PSNR) for single path in fading model, (channel SNR = 3 dB), *Orbi* test sequence

Although the proposed solution shows improvements in the results there are few limitations due to the nature of the DVC codec. The main issue of DVC is its open-loop encoding architecture compared to traditional codecs. The latter got a perfect idea about the encoding artefacts and minimise them at the encoder side (closed-loop encoder). However, in the DVC architecture, neither encoder nor decoder got any idea about decoding artefacts. The delay is another issue. The decoder has to send a request for more data depending on the decoded quality. This repeat-request nature adds delay. However we can think of mechanisms to minimise the delay by sticking to an out-of-order transmission architecture with a buffer. Also, at the decoder in order to request parity bits from the encoder the decoded quality of the video frames needs to be measured. Currently this is done by statistical measures such as PSNR and MSE. However when comes to 3D video statistical measures such as PSNR and



(a)



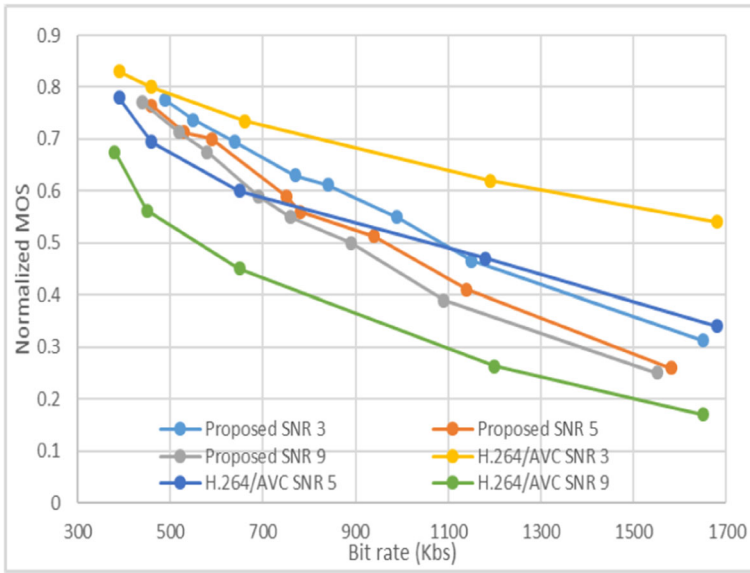
(b)

Fig. 10 **a** RD results (Normalized MOS) for 4 path in fading model, (channel SNR = 3 dB), *Orbi* test sequence. **b**. RD results for 4 path in fading model, (channel SNR = 3 dB), *Orbi* test sequence

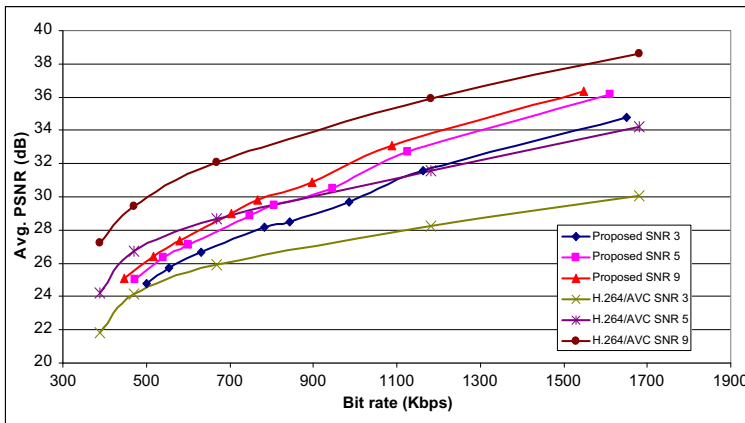
MSE fail to address the perceptual attributes of 3D video such as depth quality and visual discomfort [20]. Thus developing a no reference 3D video quality measure still remains as a challenge in the context of DVC.

5 Conclusions

In this paper, novel noise models are proposed for a DVC codec for colour and depth 3D video and the decoder is accordingly modified for the use on noisy wireless channels. The main contributions of the paper are the design and implementation of noise models for parity information (as discussed in section 3.1) and systematic information for colour and



(a)



(b)

Fig. 11 **a** DVC and H.264/AVC Performance comparison (using Normalized MOS) for *Orbi* test sequence (2 path fading). **b** DVC and H.264/AVC Performance comparison (using average PSNR) for *Orbi* test sequence (2 path fading)

depth videos (as discussed in section 3.2). 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, or the conditions of the fading. Due to its very stable and consistent behaviour, 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.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

1. Aaron A, Zhang R, Girod B (2002) Wyner-Ziv coding of motion video. Proc. Asilomar conference on signals and systems, Pacific Grove, CA, Nov. 2002
2. Bhimani J, Mi N, Leeser M, Yang Z (2017) FIM: performance prediction for parallel computation in iterative data processing applications. 2017 IEEE 10th International Conference on Cloud Computing (CLOUD), Honolulu, CA, pp 359–366
3. Bhimani J, Yang Z, Leeser M, Mi N (2017) Accelerating big data applications using lightweight virtualization framework on enterprise cloud. 2017 IEEE High Performance Extreme Computing Conference (HPEC), Waltham, MA, pp 1–7
4. Brites C, Ascenso J, Pereira F (2006) Improving transform domain Wyner-Ziv video coding performance. Proc IEEE ICASSP 2:II-525–II-528
5. Guillemot C, Pereira F, Torres L, Ebrahimi T, Leonardi R, Ostermann J (2007) Distributed Monoview and Multiview video coding. IEEE Signal Process Mag 24(5):67–76
6. Hewage CT (2008) Perceptual quality driven 3-D video over networks, Thesis, University of Surrey
7. Hewage CTER (2013) Perceptual quality driven 3D video communication. Scholars' Press, Kingston upon Thames
8. Hewage CT, Worrall S, Dogan S, Kondo AM (2009) Quality evaluation of colour plus depth map based stereoscopic video. IEEE J Sel Top Signal Process 3(2):304–318 ISSN 1932-4553
9. Ijsselstein WA, Seutens PJH, Meesters LMJ (2002) State-of-the-art in human factors and quality issues of stereoscopic broadcast television. Technical report D1, IST-2001 34396 ATTEST 2002
10. International Telecommunication Union/ITU Radio communication Sector (2012) Methodology for the subjective assessment of the quality of television pictures. ITU-R BT.500-13, Jan 2012
11. Joint Video Team of ISO/IEC MPEG and ITU-T VCEG, H.264/AVC, Reference Software JM10.1 (online), http://iphome.hhi.de/suehring/tml/download/old_jm/jm10.1.zip
12. Lin S, Costello DJ Jr (2004) Error control coding, 2nd edn. Pearson Prentice Hall, Upper Saddle River, pp 810–811
13. Montalbano G, Slock DTM (2003) Joint common-dedicated pilots based estimation of time-varying channels for W-CDMA receivers. Proc IEEE VTC 2:1253–1257
14. Ouaret M, Dufaux F, Ebrahimi T (2006) Fusion-based multiview distributed video coding. In Proceedings of the 4th ACM international workshop on video surveillance and sensor networks (VSSN '06). ACM, New York, 139–144
15. Pedro J, Ducla Soares L, Pereira F (2007) Studying error resilience performance for a feedback channel based transform domain Wyner-Ziv Video Codec. Picture coding symposium, Lisbon – Portugal, November 2007
16. Ryan WE (1998) A turbo code tutorial. Proc. IEEE Globecom 1998, At: www.ece.arizona.edu/~ryan/publications/turbo2c.pdf
17. Yang Z, Wang J, Evans D, Mi N (2016) AutoReplica: automatic data replica manager in distributed caching and data processing systems. 2016 IEEE 35th International Performance Computing and Communications Conference (IPCCC), Las Vegas, NV, pp 1–6
18. Yang Z et al (2017) AutoTiering: automatic data placement manager in multi-tier all-flash datacenter. 2017 IEEE 36th International Performance Computing and Communications Conference (IPCCC), San Diego, CA, USA, pp 1–8

19. Yang Z et al (2017) H-NVMe: a hybrid framework of NVMe-based storage system in cloud computing environment. 2017 IEEE 36th International Performance Computing and Communications Conference (IPCCC), San Diego, CA, USA, pp 1–8
20. Yasakethu SLP, Hewage CTER, Fernando WAC, Kondo A (2008) Quality analysis for 3D video using 2D video quality models. *IEEE Trans Consum Electron* 54(4):1969–1197



S. L. P. Yasakethu received his BSc. Engineering degree (First Class Hons.) in Electrical and Electronic Engineering from the University of Peradeniya, Sri Lanka, in 2007. He was awarded the prize for best performance in Electronic Communication Engineering by the University of Peradeniya for his achievements in undergraduate studies. In Oct. 2007 he was awarded the Overseas Research Scholarships Award by the Higher Education Funding Council of England to pursue PhD at the University of Surrey UK. After completing his PhD he worked as a Research Engineer for Technicolor Research & Innovations (formerly known as THOMSON R&D), in Rennes France, from Oct. 2010 to March 2012. Then he worked as a Research Fellow and a Senior Research Fellow at University of Surrey and Kingston University respectively. UK. Afterwards he worked as a Senior Lecturer for Faculty of Engineering, The South Asian Institute of Technology and Medicine, Sri Lanka and for Faculty of Engineering University of Peradeniya, Sri Lanka. Currently, he works as a Senior Lecturer at Dept. of Electrical, Electronic and Computer Science, Sri Lanka Technological Campus (SLTC), Padukka, Sri Lanka. His research interests include Quality of Experience (QoE) in multimedia communications, 2D/3D video processing and transmission, Content creation for 3D cinema and 3DTV, Cyber-security and Machine Learning. He has worked for several EU FP6 and FP7 projects in the above fields. He is a member of IEEE.



C. T. E. R. Hewage Chaminda Hewage is a Senior Lecturer in the Department of Computing and Information Systems in Cardiff School of Management in Cardiff Metropolitan University, Cardiff, where he also leads the

Networking and Data Security disciplines at both UG/PG levels. He received the B.Sc. Engineering (First Class Honours Degree) in Electrical and Information Engineering from the Faculty of Engineering, University of Ruhuna (Sri Lanka) in 2004 and the Ph.D. in Multimedia Communications from the University of Surrey (UK) in 2009. He was awarded the Gold Medal for best performance in Engineering (2014) by the University of Ruhuna for his achievements in undergraduate studies at the General Convocation held in 2004. In 2014, he received Post-Graduate Certification in HE Teaching and Learning from Kingston University – London. He is a Fellow of the Higher Education Academy (HEA), UK. After graduation, he joined Sri Lanka Telecom (Pvt.) Ltd. (Sri Lanka) as a Telecommunication Engineer (2004). In Sep. 2005 he was awarded the Overseas Research Scholarships (ORS) award by the Higher Education Funding Council of England (Universities UK) to pursue Ph.D. at University of Surrey, UK. After completing his Ph.D. in 2008, he joined University of Surrey as a Research Fellow. In 2009, he joined WMN Research Group at Kingston University – London as a Senior Researcher. He is attached to the Computing Department at Cardiff Metropolitan University as a Senior Lecture since October 2015. He worked as the Senior Researcher in a number of national and international research projects, funded by the European Commission (e.g., FP6 Visnet II, FP7 OPTIMIX, FP7 CONCERTO, FP7 Qualinet, FP7 3DConTourNet), UK research councils, Innovate UK, and industries. An IEEE Member and he is part of international committees and expert groups, including the IEEE Multimedia Communications technical committee, where he serves as Web Manager of the 3D Rendering, Processing, and Communications Interest Group. His research interests include Quality of Experience (QoE) in multimedia communications, 2D/3D video processing and transmission, Content creation for 3D cinema and 3DTV, Cyber-security and Machine Learning.