WIRELESS VIDEO

An Interlayer Error-Protection-Aided Multilayer Approach

Yongkai Huo, Mohammed El-Hajjar, and Lajos Hanzo

Video clips captured from real-world scenes exhibit intraframe correlation among their pixels. This correlation can be removed by applying video compression to reduce the required the storage space, transmission bandwidth, bitrate, and power. **<AU: Kindly check that the preceding edited sentence conveys the intended meaning.>** Layered video coding separates the video sequence into partitions having unequal importance, hence allowing the decoder to progressively refine the reconstructed video quality, when an increased bandwidth is available. On the other hand, compressed video signals are sensitive to channel errors. Therefore, forward error correction (FEC) must be applied when communicating over hostile wireless channels. In addition, based on the fact that the different layers have unequal importance, different-rate FEC codes may be applied to the different layers, leading to unequal error protection (UEP). Our new contribution is that we propose an interlayer (IL) FEC coding technique combined with UEP, where the lower-importance layers are used for protecting

Digital Object Identifier 10.1109/MVT.2014.2334411
Date of publication:

the higher-importance layers in the data-partitioned mode of H.264/AVC **<AU: Please confirm whether "AVC" may be spelled out. If so, please provide the expansion.>** video coding. Explicitly, our simulation results show that the IL-coded system outperforms the traditional UEP system by providing a better video quality for transmission over a wireless channel having E_b/N_0 of 0 dB, when using our multifunctional multiple-input, multiple-output (MIMO) array.

Layered Video Compression

Uncompressed video sequences captured from a real-world scene exhibit a high intraframe correlation amongst pixels. Intuitively, the correlation residing in an uncompressed video sequence should be removed to represent the original video with the aid of fewer bits, yet without any substantial reduction of the perceived visual quality. Video compression will reduce the required storage in a hard drive, for example, or the transmission bandwidth and the transmission power required for distributing the video. A number of video-compression standards [1] have been designed during the past decades for the sake of achieving a high compression ratio.

Moreover, layered video coding was proposed for the sake of generating multiple layers of unequal importance, which has been adopted by a number of existing video-coding standards [2]–[4]. In general, the most-important layer is referred to as the *base layer* (*BL*), while the less-important layers are referred to as the *enhancement layers* (*ELs*). A layered video decoder may output a low-/medium-quality video by decoding only the BL, while a higher-quality video may be reconstructed, when decoding both the BL and ELs. In addition, the layered video decoder relies on the BL for decoding the ELs. For example, a layered video-coding standard referred to as *scalable video coding* (*SVC*) [2], [3] was recently developed as an extension of H.264/AVC [3], which encodes a video sequence into multiple layers, where a reduced-size subset of the bitstream may be extracted to meet the users' specific preferences, such as bandwidth, frame per second (FPS) scanning rate, video-frame size, or visual quality. For example, a mobile TV receiver might decode the BL only, while a high-definition TV receiver would decode both the BL and all the ELs. Moreover, the less-important ELs may be dropped during network congestion or buffer overflow. In layered video transmission, when the BL is corrupted or lost due to channel impairments, the ELs must also be dropped by the video decoder, even if they are perfectly received. More details on the state of the art in layered video communication techniques will be introduced in the "Standardized Layered Video Techniques" section.

UEP for Layered Video

However, the compressed bits become extremely sensitive to the errors introduced during the video distribution, where a single bit error may corrupt the whole video sequence, similarly to a zipped data file. To combat this problem, typically FEC or channel coding is employed as a technique of controlling and correcting errors induced during video distribution, which is achieved by incorporating redundant bits, depending on the compressed video bits. Numerous FEC codes, such as recursive systematic convolutional (RSC) codes, and turbo codes [5], have been used in video applications [1]. Since the ELs depend on the BL for decoding in layered video coding, it is intuitively prudent to protect the BL more strongly than the ELs, while keeping the overall protection redundancy fixed for layered video communication over unreliable channels. In other words, we may encode the BL and ELs using FEC coding rates lower than $R = 0.5$ and higher than $R = 0.5$, respectively, while keeping the overall channel coding rate at $R = 0.5$ in a realistic half-rate coding system. In the literature, this technique is referred to as *UEP*, which was originally proposed in [6]. In UEP, stronger FEC is allocated to the more-significant video bits, while dedicating weaker FEC to the less-important video parameters. Since then, numerous contributions have been made in the field of UEP video communications relying on realistic video signals [7]–[9].

The performance of data-partitioning (DP) [3] aided H.264/AVC video streaming using RSC coded UEP was evaluated in [9]. A novel UEP method was proposed in [7] for scalable video streaming over networks subject to packet-loss events, where the authors presented an efficient performance metric for quantifying the error propagation effects imposed by packet-loss events. Maani and Katsaggelos [8] proposed cross-layer operation-aided scalable video streaming, which aimed for the robust delivery of the scalable video over error-prone channels. The achievable video quality was further improved with the aid of content-aware bit-rate allocation, and a powerful error-concealment method was invoked at the receiver.

IL-Coded Layered Video

In traditional UEP schemes conceived for layered video communication, variable-rate FEC was invoked for the different layers. When the BL is corrupted or lost, the ELs also have to be dropped, regardless of whether they are perfectly received or not, which implies that the transmission power assigned to the ELs was wasted. Motivated by this fact, we seek to efficiently exploit the valuable transmission power allocated to the ELs for recovering the error-infested BL, even if the ELs are sacrificed. The so-called layer-aware FEC philosophy [10] using a Raptor code c1 was invoked for video transmission over binary erasure channels. (A Raptor code belongs to the family of packet-erasure-filling codes, which encode a number of packets by incorporating redundant packets for correcting packet erasures, whilst classic FEC codes encode bits by incorporating redundant bits for correcting bit errors.) At the transmitter, the channel encoding was consciously performed right across the BL and the ELs. As a benefit, at the receiver cross-layer decoding may be invoked for recovering the erased bits within the BL. Motivated by these advances, in [11], we developed an IL operation-aided FEC (IL-FEC) scheme relying on RSC codes, where the systematic bits of the BL were carefully implanted into the ELs without increasing their bitrate. At the receiver, the above-mentioned bits implanted into the ELs may then be beneficially exploited for assisting in decoding the BL. The IL-FEC technique of [11] was also combined with the UEP philosophy for the sake of further improving the attainable system performance, where different layers were encoded by different RSC coding rates. In addition, more advanced turbo codes were also applied in [12] and [13].

At the time of writing, multimedia content is evolving from traditional content to a range of rich, heterogeneous media content, such as traditional TV, and streaming audio and video as well as images and text messaging. In this article, we describe the philosophy of transmitting an IL-FEC encoded compressed video bitstream employing an RSC codec with the aid of a MIMO [14] transceiver structure (IL-RSC-MIMO). We note, however, that this philosophy is directly applicable to arbitrary FEC and transceiver schemes. This scheme may be considered as an evolution of the traditional UEP schemes exemplified by [7] and [8]. The DP mode of the H.264/AVC video codec is employed, where the type B and type C partitions are utilized for protecting the type A partition 2 (for brevity, we will often simply refer to them as A, B, and C). Finally, different rate and different protection channel codecs will be employed as FEC codes for improving the attainable system performance. Against this background, our novel contribution is that we conceive a design methodology for IL-FEC coded layered video systems. Explicitly, since the BL's information is implanted into the EL, this information is extracted from the ELs at the receiver to assist in decoding the BL, when the BL cannot be perfectly decoded on its own right.

Again, we use the H.264/AVC DP mode in our simulations, but our proposed scheme is not limited to partitioning-based video—it may be readily applied to any arbitrary system relying on layered video coding.

Standardized Layered Video Techniques

Layered video compression [2], [10] encodes a video sequence into multiple layers, which enables us to progressively refine the reconstructed video quality at the receiver, when the network's throughput allows this. In general, the most-important layer is referred to as the *BL*, and the less-important layers are termed *ELs*, which rely on the BL. Furthermore, an EL may be relied upon by less-important ELs. Again, when the BL or an EL is lost or corrupted during its transmission, the dependent layers cannot be utilized by the decoder and must be dropped. A layered video scheme is shown in Figure 1, where the video sequence captured from the scene is encoded into four layers by the layered video encoder, i.e., $L_0 \sim L_3$, where layer L_i $(0 < i \leq 3)$ depends on layer L_{i-1} for decoding, while layer L_i improves the video quality of layer L_{i-1}. In other words, layer L_0 is the BL and layers $L_1 \sim L_3$ are ELs depending on the BL. Furthermore, as shown in Figure 1, the ELs L_2 and L_3 rely on the EL L_1. In other words, if layer L_1 is corrupted, then layers L_2 and L_3 are dropped by the decoder. Given only the layer L_0 having a bitrate of 128 kb/s, the corresponding layered video decoder of Figure 1 reconstructs the video with a resolution of quarter common intermediate format (QCIF) at 7.5 frames/s. By contrast, a common intermediate format (CIF) (CIF and QCIF indicate a resolution of 352×288 and 176×144, respectively) based video sequence scanned at 30 frames/s can be reconstructed with the aid of layers L_0, L_1, and L_2, which require bitrates of 128,

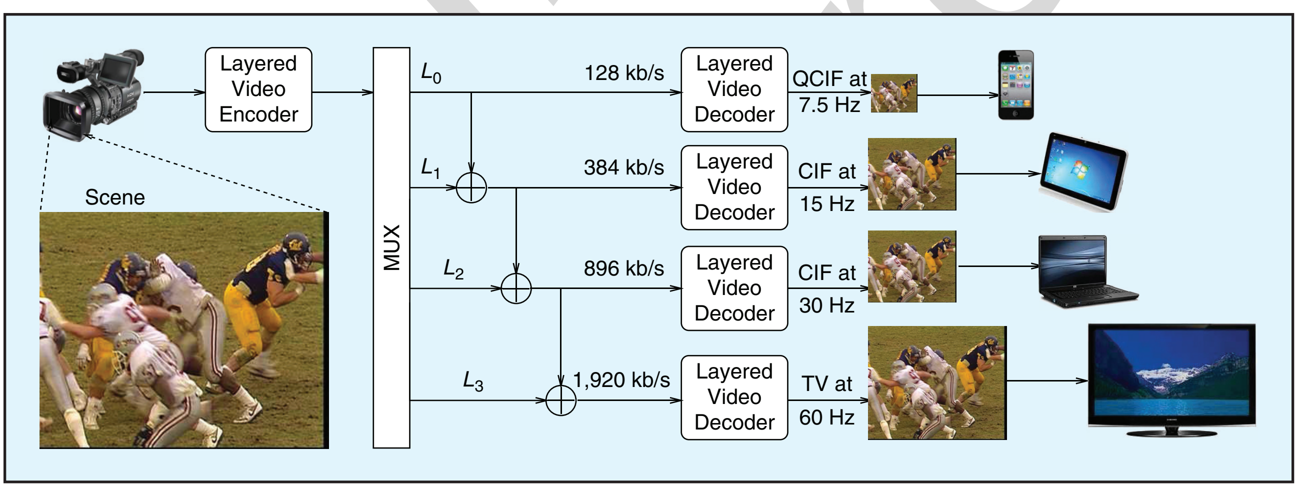

FIGURE 1 The architecture of a layered video scheme, where the video quality is refined progressively. ***<AU: Please confirm that permission has been granted to print these photos, and please provide proper credit for them.>***

256, and 512 kb/s, respectively. If the TV screen of Figure 1 is utilized by the user, all four layers $L_0 \sim L_3$ may also be streamed for achieving the highest video quality. In practice, the different video streaming scenarios of Figure 1 require different bandwidths and, hence, achieve different visual quality. The users may rely on different video screens, such as those of mobile phones, tablets, PCs, and TV screens, as observed in Figure 1, for example.

The subject of SVC [2] has been an active research field for over two decades. This terminology is also used in the Annex G extension of the H.264/AVC video-compression standard [3]. Indeed, SVC is capable of generating several bitstreams that may be decoded at a similar quality and compression ratio to that of the existing H.264/AVC codec. When, for example, low-cost, low-quality streaming is required by the users, some of the ELs may be removed from the compressed video stream, which facilitates flexible bitrate control based on the specific preferences of the users. An H.264/AVC scalable video stream contains a sequence of network abstraction layer units (NALUs) [3], which are the most basic elements of the H.264/AVC encoded bitstream.

Recently, the Joint Video Team (JVT) proposed multiview video coding (MVC) [4] as an amendment to the H.264/AVC standard [3]. Apart from the classic techniques employed in single-view coding, MVC invokes the so-called interview correction technique by jointly processing the different views for the sake of reducing the bitrate. Hence, the first encoded view may be termed the BL, while the remaining views may be treated as the ELs.

A number of layered video-coding schemes have been developed [2], [3], and some of them are adopted by recent video-coding standards, for example, the SVC [2] and DP [3]. In this article, we use DP-based layered video coding in our simulations, which is a beneficial feature of the H.264/AVC codec [3]. In the DP mode, the data streams representing different semantic importance are categorized into a maximum of three bitstreams/partitions [3] per video slice, i.e., type A, type B, and type C partitions. The header information is carried by the A partition, which contains the compression parameters of the current video slice. The B and C partitions carry the intraframe- and interframe-coded data, respectively. Amongst these three partitions, the type A partition may be deemed to be the most important one, which may be treated as the BL. Correspondingly, the B and C partitions may be interpreted as a pair of ELs, since they are dependent on the A partition for decoding. Although the information in partitions B and C cannot be used in the absence of A, partition B and C can be used independently of each other, again, given the availability of A. In this article, we will employ the PM **<AU: Kindly spell out PM.>** of H.264/AVC for characterizing our system.

Wireless Video Architecture

In this section, we will briefly introduce the architecture of the IL-RSC scheme [11] conceived for layered video transmission over a MIMO system. The system's structure is shown in Figure 2, where a layered video codec is employed, while the structure of the check node decoder (CND) [15] is based on the box plus operation ⊞ **<AU: What does this symbol represent? Can it be deleted?>** described in [15]. In the "Transmitter Model" section, we first detail the techniques employed at the transmitter. Then, our IL-RSC decoding techniques will be illustrated in the "Receiver Model" section, with special emphasis on how the RSC decoders 0 and 1, shown in Figure 2, exchange their IL redundancy using the CND for

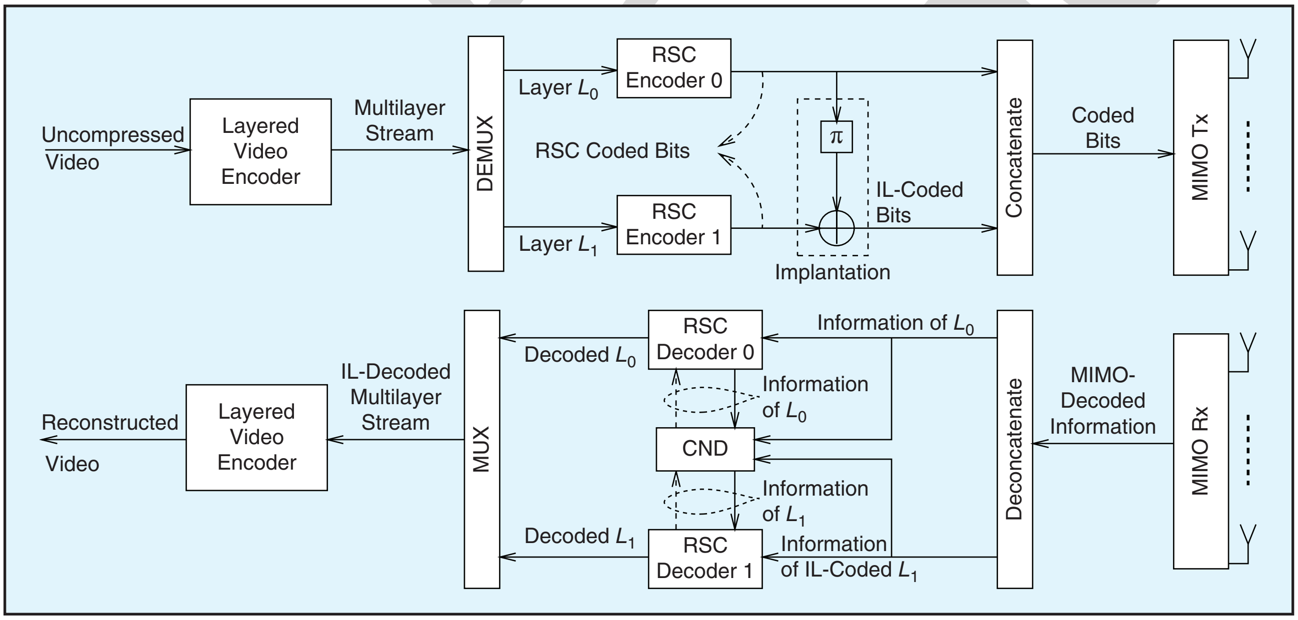

FIGURE 2 The IL-RSC encoding/decoding architecture for layered video coding, where π indicates an interleaver. The IL decoding is detailed in Figure 3.

improving the overall performance of the system. For the sake of simplifying our description, we assume that only two layers are generated by the layered video encoder, namely the BL L_0 and the EL L_1.

Transmitter Model

At the transmitter, the uncompressed video is compressed using the layered video encoder of Figure 2, generating a multilayer stream containing layers L_0 and L_1. Then, the multilayer stream is demultiplexed into layers L_0 and L_1 by the DEMUX block of Figure 2, which are encoded as follows.

1) The BL L_0 will be encoded by the RSC encoder 0 of Figure 2 using the classic encoding method.
2) The EL L_1 will first have been coded by the RSC encoder 1 of Figure 2. Then, the exclusive-OR (XOR) operation will be utilized by the "implantation" process of Figure 2 to implant the coded bits of layer L_0 into the coded bits of layer L_1. Note that interleaving is performed on the coded bits of layer L_0 before the XOR operation.

Finally, the resultant RSC coded bits for L_0 and IL-coded bits for L_1 are concatenated into a joint bitstream for transmission. We assume that the layers L_0 and L_1 carry an identical number of bits. However, as detailed in [11] and [13], our method may be readily extended to the practical scenarios, where the layers contain different number of bits. Following the IL-RSC encoding procedure, the resultant bits are transmitted over a MIMO transmitter architecture.

Receiver Model

At the receiver, the MIMO decoding process is carried out as detailed in [14], generating the information for the layers L_0 and L_1. Following the MIMO decoding process, the IL RSC decoding will be performed on the two layers.

The IL aided RSC decoding process is illustrated by the flowchart of Figure 3. First, the RSC decoder 0 will decode the information of layer L_0 generated by the MIMO decoder for estimating the video layer L_0. Then, the information generated for layer L_0 will be input to the CND block of Figure 2 for extracting the information of layer L_1 using the information of L_1, which is generated by the MIMO decoder.

Then, the RSC decoder 1 of Figure 2 will decode the information of layer L_1 with the aid of the related information of the MIMO decoder of Figure 2. Afterward, a classic cyclic redundancy check (CRC) is employed for detecting, whether the recovered BL L_0 is error-free or not, as shown in Figure 3. We may then invoke one of two possible decoding processes, as shown in Figure 3 and described below.

Operation Without IL Feedback

When the layer L_0 is successfully recovered, the layers L_0 and L_1 will be estimated by the hard decision decoder block shown in Figure 3. Afterward, the receiver may discard the layer L_1 depending on whether it is error-free or not, according to the CRC check. In this case, only the solid lines shown in Figures 2 and 3 will be activated. Note that, given the perfectly recovered L_0, the CND operation involved in the process of extracting the information of L_1 from the information of IL-coded L_1 of Figure 2 is essentially a simple XOR-like operation [11], prohibiting further degradation of the EL L_1. This implies that in this scenario our proposed IL technique is equivalent to the traditional UEP techniques, where the two layers L_0 and L_1 are encoded and decoded independently. Moreover, since the layer L_0 is decoded independently without feedback from the layer L_1, the two layers are only decoded once, i.e., without imposing any extra complexity on the receiver. In addition, in practical applications, the BL L_0 can be reconstructed immediately when it is received, i.e., without waiting for EL L_1. **<AU: Please confirm whether the deletion of quotes in the sentence beginning "Note that, given..." is OK.>**

Operation Using IL Feedback

When the layer L_0 is not successfully decoded, the iterative IL technique of Figures 2 and 3 will be activated for utilizing the information of layer L_0 fed back from the RSC decoder 1. In this case, both the solid lines and the dashed lines shown in Figures 2 and 3 will be activated. More explicitly, the CND block of Figure 3 will be utilized for extracting the information of the layer L_0 based on the information of the IL-coded L_1. After this stage, improved information of the layer L_0 is generated, which concludes the current IL decoding iteration. Finally, the receiver will return to the beginning of the flowchart shown in Figure 3. The iterative IL decoding process continues until the number of affordable iterations was exhausted or the BL L_0 is perfectly recovered, as shown in Figure 3.

Again, the receiver may successfully reconstruct the BL L_0 independently of layer L_1, and the information of layer L_1 can be extracted from the perfectly decoded BL L_0 and from the information of the IL-coded L_1. However, when the receiver fails to reconstruct the BL L_0 without the layer L_1, the iterative IL decoding technique exchanging information between the RSC decoder 0 and decoder 1 will be activated. Note that when the BL L_0 cannot be recovered correctly, the EL L_1 must be dropped by the receiver. Furthermore, since the IL encoding process does not require any extra coded bits, we do not reduce the overall code rate compared with the traditional UEP methods.

Performance Study

Let us continue by characterizing our proposed IL-RSC-MIMO system against the traditional UEP-aided RSC-MIMO, dispensing with IL coding. Three 30-frame video sequences, i.e., the *Foreman*, *Football*, and *Bus* clips, represented in (352 × 288)-pixel CIF were encoded by a H.264/AVC video codec operated in its DP mode. The video scanning rates expressed in FPS were 30, 15, and 30

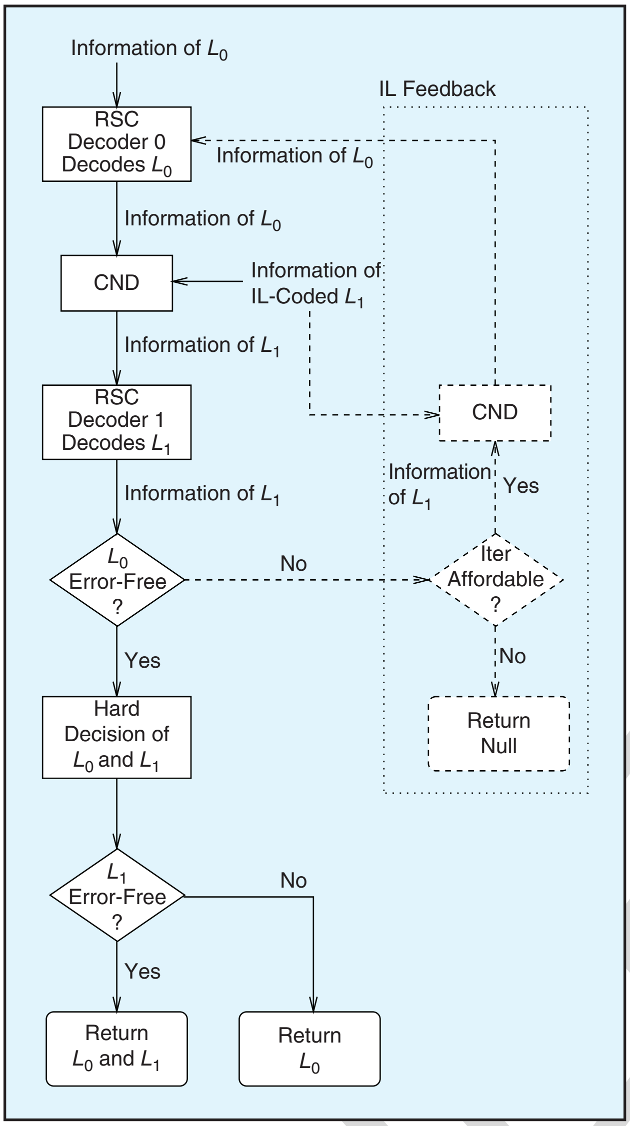

FIGURE 3 The flowchart of IL-aided RSC decoding in Figure 2.

for the *Foreman*, *Football*, and *Bus* sequences, respectively. All the 30-frame video sequences were encoded into an intracoded (I) frame, followed by 29 predicted (P) frames. The bidirectionally predicted (B) frame was disabled due to the fact that it relies on both previous and future frames for decoding, which may introduce precipitated error propagation as well as additional delay. The motion-copy-based error concealment tool built into the H.264/AVC reference codec was employed for the sake of combating the effects of residual channel impairments. All the above configurations jointly result in a bitrate of 655 kb/s and an error-free peak signal-to-noise ratio (PSNR) of 38.4 dB for the *Foreman* sequence, where the PSNR of a $(M \times N)$-pixel frame is calculated as follows:

$$\mathrm{PSNR_{dB}} = 20 \log 10 \left\{ \frac{255}{\sqrt{\frac{1}{M \times N} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} [I(i,j) - K(i,j)]^2}} \right\} \mathrm{dB}, \quad (1)$$

where $I(i, j)$ and $K(i, j)$ represent the value of the original and estimated pixels, respectively. We employed the *Foreman*, *Football*, and *Bus* sequences to show the suitability of our scheme for the transmission of both low-motion and high-motion video. **<AU: Please note that the text "We will first...system of Figure 2" has been deleted as the description of the same is presented later.>**

Error-Protection Arrangements

In the simulations, we employ the overall coding rate of 1/2 for both the equal error protection (EEP) and UEP schemes. For each compressed bitstream, all NALUs were scanned for calculating the total number of bits for the A, B, and C partitions. Let us assume that the A, B, and C partitions have a total of N_a, N_b, and N_c bits, respectively, and the A, B, and C streams have coding rates of r_a, r_b, and r_c, respectively. Then, $2 \cdot (N_a + N_b + N_c) = N_a/r_a + N_b/r_b + N_c/r_c$ must be satisfied for the sake of guaranteeing that the overall coding rate remains 1/2. Again, the A stream is the most-important layer, while the B and C bitstreams are the ELs, where the bitstream B and C are similarly important. Hence, in all the error-protection arrangements, we have $r_b = r_c$. More specifically, we first select a specific value for r_a, then the value of $r_b = r_c$ may be calculated as $r_b = N_b + N_c / 2 \cdot (N_a + N_b + N_c) - N_a/r_a$.

Note that the total number of bits for each partition of the different video sequences may be different, which results in different protection arrangements. Based on the above, the three error-protection arrangements conceived for the *Football* and *Foreman* sequences are shown in Table 1, which may be readily combined with arbitrary EEP or UEP schemes, where variable-rate puncturers were designed and employed for the sake of achieving a specific coding rate. For example, for the *Football* sequence, the UEP1 arrangement encodes the A, B, and C partitions using coding rates of 0.65, 0.47, and 0.47, respectively, while keeping an overall coding rate of 0.5.

Numerical Results

The PSNR versus E_b/N_0 performance recorded for the *Football* sequence is shown in Figure 4(a), where we observe that the video quality quantified in terms of the PSNR increases upon increasing the E_b/N_0 encountered. This is because an increased power results in a reduced bit error ratio for the received video packets. Observe in Figure 4(a) that the EEP aided scheme achieves the best performance among all the arrangements of the traditional RSC coded schemes, because the A partition carries only the video header information and fails to assist the H.264/AVC decoder in concealing the residual errors, when the B and C partitions are corrupted. Furthermore, the systems using our proposed IL coding technique outperform their corresponding benchmarkers. Specifically, the UEP2-IL-RSC-MIMO constitutes the best protection arrangement among all IL-RSC schemes, which achieves a

TABLE 1 ***The coding rates of different error-protection arrangements for the Football/Foreman/Bus sequence. The code rates were adjusted by variable-rate puncturers.***

Error-Protection Arrangements	***Code Rates*** ***Type A***	***Type B***	***Type C***	***Average***
EEP	0.50/0.50/0.50	0.50/0.50/0.50	0.50/0.50/0.50	0.50/0.50/0.50
UEP1	0.65/0.40/0.45	0.47/0.65/0.52	0.47/0.65/0.52	0.50/0.50/0.50
UEP2	0.85/0.60/0.65	0.44/0.43/0.47	0.44/0.43/0.47	0.50/0.50/0.50

power reduction of about 3 dB compared with the EEP-RSC-MIMO scheme at a PSNR of 36 dB. Alternatively, about 3.7 dB of PSNR video quality improvement may be observed at a channel SNR of 0 dB. In comparison to the IL-RSC-MIMO systems, both the UEP1-RSC-MIMO and UEP2-RSC-MIMO schemes dispensed with the IL technique, which resulted in severely distorted video quality. Note that the UEP2-RSC-MIMO scheme is not shown in Figure 4(a), since its performance curve is beyond the visible range of Figure 4(a). A subjective comparison of the UEP2-IL-RSC-MIMO and EEP-RSC-MIMO arrangements for the Football sequence is presented in the Figure 5(a).

For providing further insights for video scenes having different motion activity, the PSNR versus E_b/N_0 performance of the IL-RSC coded systems is included in Figure 4(b) and (c) using the *Foreman* and *Bus* sequences, when employing the protection arrangements of Table 1. In excess of 2 dB of power reduction is achieved by the UEP2-IL-RSC-MIMO arrangement compared with the EEP-RSC-MIMO scheme at a PSNR of 37 dB for both the *Foreman* and *Bus* sequences. Viewed from a different perspective, in excess of 3.2 dB of PSNR video quality improvement may be observed at a channel SNR of –1 dB. A subjective comparison of the UEP2-IL-RSC-MIMO and EEP-RSC-MIMO arrangements for the *Foreman* and *Bus* sequences are portrayed in Figure 5(b) and (c), respectively.

Coding-Rate Optimization

Observe from Figure 4 that for different video sequences, different FEC code rates result in different performance due to their different video characteristics. In [16], the FEC code rates are optimized in real time at the transmitter.

Specifically, we designed an algorithm for estimating the expected video distortion at specific coding rates at the transmitter. Intuitively, the coding rates having the minimum expected video distortion will be selected as our optimal solution. Considering Table 2 as an example, where we test six different coding rates for a video stream containing the layers L_0 and L_1. For each case of Table 2, we estimate the expected video distortion based on the sequence considered and on the available coding rates at the transmitter. Then, the transmitter will encode the layers L_0 and L_1 using the code rates of 0.6 and 0.43, respectively, since Case 2 induces the lowest expected video distortion.

Conclusion

In this article, a brief description of novel protection arrangements conceived for layered video coding was presented. It was shown how UEP can be utilized for enhancing the video quality at the receiver. Furthermore, we described our proposed IL coding technique that can

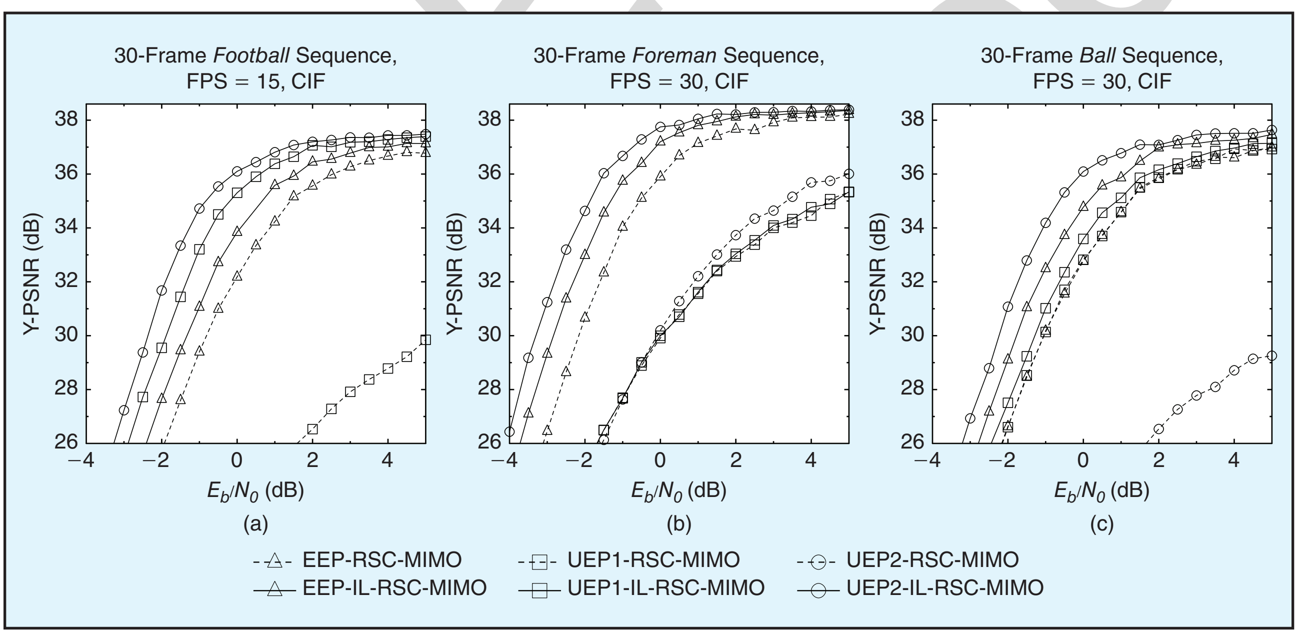

FIGURE 4 The PSNR versus E_b/N_0 performance for the (a) *Football*, (b) *Foreman*, and (c) *Bus* sequences using the RSC coding schemes of Table 1. ***<AU: Please confirm whether the top line of each artwork can be deleted and whether 30-Frame can be added to the caption.>***

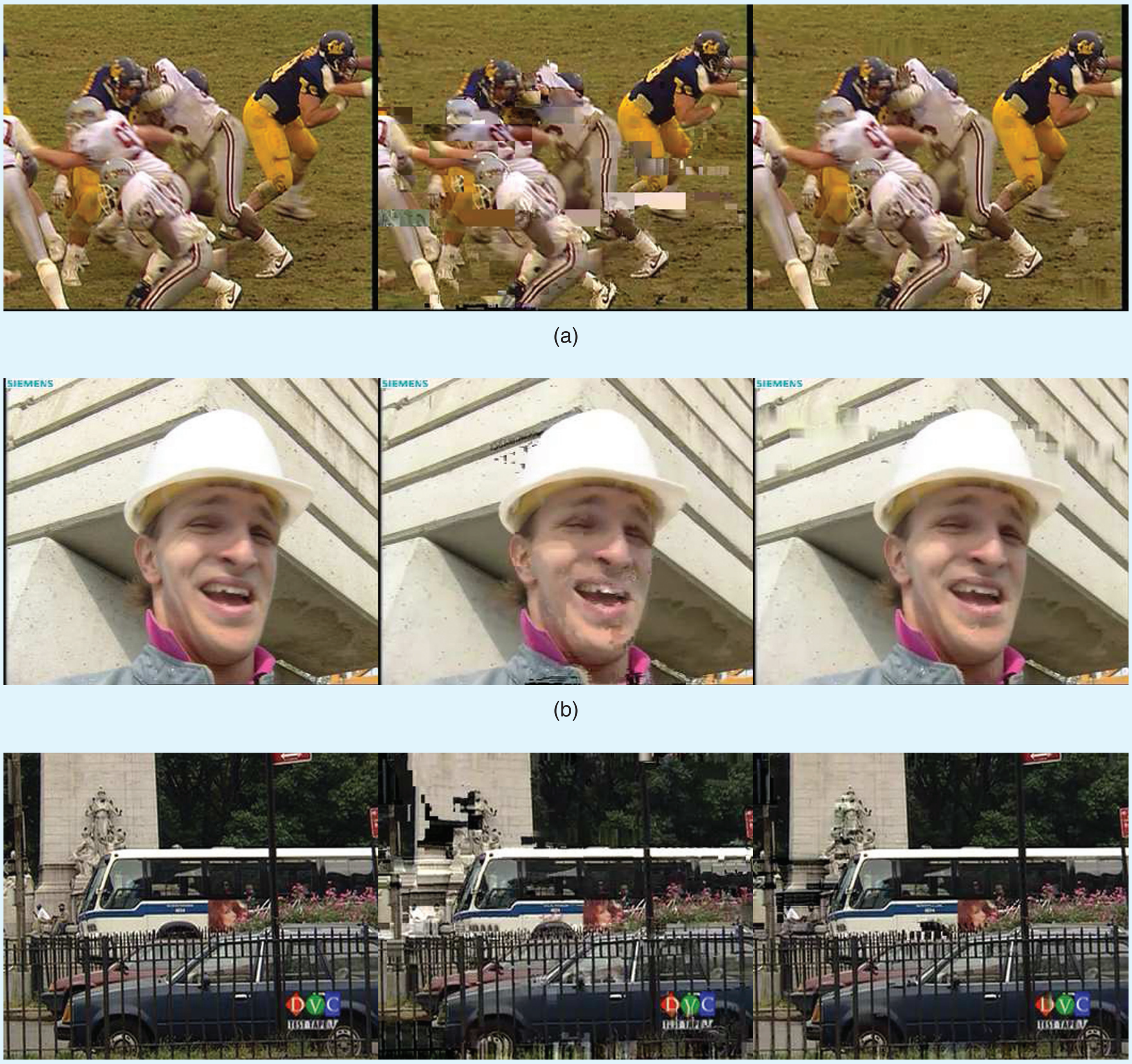

FIGURE 5 A video comparison at $E_b/N_0 = -2.5$ dB for the (a) *Football*, (b) *Foreman*, and (c) *Bus* sequences. The first column indicates the original frames. The second column indicates the EEP-RSC-MIMO decoded frames. The third column represents the UEP2-IL-RSC-MIMO decoded frames for the *Football*, *Foreman*, and *Bus* sequences, respectively. ***<AU: Please confirm that permission has been granted to print this photo, and please provide proper credit for it.>***

TABLE 2 An example of a coding rate versus video distortion table.

Error-Protection Arrangements	*Code Rates*			*Distortion*
	L_0	L_1	*Average*	
Case 0	0.5	0.5	0.5	5 dB
Case 1	0.4	0.67	0.5	2 dB
Case 2	0.6	0.43	0.5	1 dB
Case 3	0.7	0.39	0.5	3 dB
Case 4	0.8	0.36	0.5	7 dB
Case 5	0.9	0.35	0.5	10 dB

be combined with UEP in layered video coding to further improve the attainable system performance. Our simulation results showed that the proposed IL-coded system achieved a gain of about 3 dB of E_b/N_0 or 3.7 dB of PSNR over the traditional UEP system, when employing an RSC codec. A fundamental design guideline for layered video transmission is to optimize the coding-rate allocation of all the layers for the sake of improving the attainable video quality in the face of channel impairments.

Furthermore, the corresponding conclusion of this IL-coded video transmission design guideline is to jointly encode multiple layers for the sake of exploiting their mutual dependency at the receiver to improve the robustness against transmission errors.

Acknowledgments

The authors would like to thank the European Union's Concerto project, the European Research Council's Senior Fellow Grant, and the RC-U.K. **<AU: Kindly spell out RC-U.K.>** under the India–U.K. Advanced Technology Center for their financial support.

Author Information

Yongkai Huo <AU: please provide a contact e-mail address.> received his B.E. degree with distinction in computer science and technology from Hefei University of Technology, Hefei, China, in 2006 and his M.E. degree in computer software and theory from the University of Science and Technology of China, Hefei, in 2009. In 2014, he received his Ph.D. degree in wireless communications from the Communications, Signal Processing, and Control Group, University of Southampton, United Kingdom, where he is currently working as a senior research assistant. He received a scholarship under the China–United Kingdom. Scholarships for Excellence Program. His research interests include distributed video coding, multiview video coding, robust wireless video streaming, and joint source-channel decoding.

Mohammed El-Hajjar <AU: please provide a contact e-mail address.> is a lecturer in electronics and computer science at the University of Southampton, United Kingdom. He received his B.E. degree in electrical engineering from the American University of Beirut, Lebanon, in 2004. He then received his M.Sc. degree in radio-frequency communication systems and Ph.D. degree in wireless communications, both from the University of Southampton in 2005 and 2008, respectively. Following receiving his Ph.D. degree, he joined Imagination Technologies as a research engineer, where he worked on designing and developing the BICM **<AU: Kindly spell out BICM.>** peripherals in Imagination's multistandard communications platform, which resulted in several patent applications. In January 2012, he joined the Electronics and Computer Science Department in the University of Southampton as a lecturer in the communications, signal processing and control research group. He is the recipient of several academic awards and is the author of a Wiley-IEEE book and in excess of 30 journal and international conference papers. His research interests are mainly in the development of intelligent communications systems for the Internet of Things including massive multiple-input, multiple-output systems for millimeter-wave communications, cooperative communications, and radio over fiber systems.

Lajos Hanzo <AU: please provide a contact e-mail address.> received his degree in electronics in 1976, his doctorate degree in 1983, and his D.Sc. degree. During his 35-year career in telecommunications, he has held various research and academic posts in Hungary, Germany, and the United Kingdom. Since 1986, he has been with the School of Electronics and Computer Science, University of Southampton, United Kingdom, where he holds the chair in telecommunications. He has successfully supervised 80 Ph.D. students; coauthored 20 John Wiley/IEEE Press books on mobile radio communications; totaling in excess of 10,000 pages; published 1,300 research entries in IEEE *Xplore*; acted both as technical program chair and general chair of IEEE conferences; presented keynote lectures; and been awarded a number of distinctions. Currently, he is directing a 100-strong academic research team, working on a range of research projects in the field of wireless multimedia communications sponsored by industry, the U.K. Engineering and Physical Sciences Research Council, the European IST Program, and the U.K. Mobile Virtual Center of Excellence. He is an enthusiastic supporter of industrial and academic liaison, and he offers a range of industrial courses. He is also a governor of the IEEE Vehicular Technology Society. From 2008 to 2012, he was the editor-in-chief of the IEEE Press, and, since 2009, he has been a chaired professor at Tsinghua University, Beijing. For further information on research in progress and associated publications, please refer to http://www-mobile.ecs.soton.ac.uk. He is a Fellow of the IEEE.

References

[1] L. Hanzo, P. Cherriman, and J. Streit, *Video Compression and Communications: From Basics to H.261, H.263, H.264, MPEG2, MPEG4 for DVB and HSDPA-Style Adaptive Turbo-Transceivers*. New York: Wiley, 2007.

[2] H. Schwarz, D. Marpe, and T. Wiegand,"Overview of the scala blevideo coding extension of the H.264/AVC standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 17, pp. 1103–1120, Sept. 2007.

[3] Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG, *ITU-T Rec. H.264/ISO/IEC 14496-10 AVC: Advanced Video Coding for Generic Audiovisual Services*, Mar. 2010. **<AU: Kindly specify the reference type.>**

[4] A. Vetro, T. Wiegand, and G. Sullivan, "Overview of the stereo and multiview video coding extensions of the H.264/MPEG-4 AVC standard," *Proc. IEEE*, vol. 99, pp. 626–642, Apr. 2011.

[5] L. Hanzo, T. Liew, B. Yeap, and R. Tee, *Turbo Coding, Turbo Equalisation and Space-Time Coding*. New York: Wiley, 2011.

[6] B. Masnick and J. Wolf, "On linear unequal error protection codes," *IEEE Trans. Inform. Theory*, vol. 13, pp. 600–607, Oct. 1967.

[7] H. Ha and C. Yim, "Layer-weighted unequal error protection for scalable video coding extension of H.264/AVC," *IEEE Trans. Consum. Electron.*, vol. 54, pp. 736–744, May 2008.

[8] E. Maani and A. Katsaggelos, "Unequal error protection for robust streaming of scalable video over packet lossy networks," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 20, pp. 407–416, Mar. 2010.

[9] Nasruminallah and L. Hanzo, "Near-capacity H.264 multimedia communications using iterative joint source-channel decoding," *IEEE Commun. Surveys Tuts.*, vol. 14, no. 2, pp. 538–564, 2012.

[10] C. Hellge, D. Gomez-Barquero, T. Schierl, and T. Wiegand, "Layer-aware forward error correction for mobile broadcast of layered media," *IEEE Trans. Multimedia*, vol. 13, pp. 551–562, June 2011.

[11] Y. Huo, M. El-Hajjar, and L. Hanzo, "Inter-layer FEC aided unequal error protection for multilayer video transmission in mobile TV," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 23, no. 9, pp. 1622–1634, 2013.

[12] Y. Huo, X. Zuo, R. G. Maunder, and L. Hanzo, "Inter-layer FEC decoded multi-layer video streaming," in *Proc. IEEE Global Telecommunications Conf.*, Anaheim, CA, Dec. 2012, pp. 2113–2118.

[13] Y. Huo, M. El-Hajjar, M. F. U. Butt, and L. Hanzo, "Inter-layerdecoding aided self-concatenated coded scalable video transmission," in *Proc. IEEE Wireless Communications Networking Conf.,* Shanghai, China, Apr. 2013, pp. 4647–4652.

[15] S. T. Brink, G. Kramer, and A. Ashikhmin, "Design of low-density parity-check codes for modulation and detection," *IEEE Trans. Commun.*, vol. 52, pp. 670–678, Apr. 2004.

[16] Y. Huo, M. El-Hajjar, R. G. Maunder, and L. Hanzo, "Layered wireless video relying on minimum-distortion inter-layer FEC coding," *IEEE Trans. Multimedia*, vol. 23, pp. 319–331, Jan. 2014.

VT

VIDEO CLIPS CAPTURED FROM REAL-WORLD SCENES EXHIBIT INTRAFRAME CORRELATION AMONG THEIR PIXELS.

THE JOINT VIDEO TEAM (JVT) PROPOSED MULTIVIEW VIDEO CODING (MVC) [4] AS AN AMENDMENT TO THE H.264/AVC STANDARD [3].

THE RSC DECODER 0 WILL DECODE THE INFORMATION OF LAYER L_0 GENERATED BY THE MIMO DECODER FOR ESTIMATING THE VIDEO LAYER L_0.

THE VIDEO SCANNING RATES EXPRESSED IN FPS WERE 30, 15, AND 30 FOR THE FOREMAN, FOOTBALL, AND BUS SEQUENCES, RESPECTIVELY.

ALL NALUS WERE SCANNED FOR CALCULATING THE TOTAL NUMBER OF BITS FOR THE A, B, AND C PARTITIONS.

WHEN THE BL IS CORRUPTED OR LOST, THE ELS ALSO HAVE TO BE DROPPED, REGARDLESS OF WHETHER THEY ARE PERFECTLY RECEIVED OR NOT.

WE DESIGNED AN ALGORITHM FOR ESTIMATING THE EXPECTED VIDEO DISTORTION AT SPECIFIC CODING RATES AT THE TRANSMITTER.