H.264 Wireless Video Telephony Using Iteratively-Detected Binary Self-Concatenated Coding

Nasruminallah[1], Muhammad Fasih Uddin Butt[1,2], Soon Xin Ng[1] and Lajos Hanzo[1]
[1]School of ECS, University of Southampton, SO17 1BJ, UK.
Email: {n06r, mfub06r, sxn, lh}@ecs.soton.ac.uk, http://www-mobile.ecs.soton.ac.uk
[2]Dept of EE, COMSATS Institute of Information Technology, Islamabad, 44000, Pakistan, http://ciit.edu.pk

***Abstract*— In this contribution we propose a robust H.264 coded wireless video transmission scheme using iteratively decoded self-concatenated convolutional coding (SECCC). The proposed SECCC scheme is composed of constituent recursive systematic convolutional (RSC) codes and an interleaver is used to randomise the extrinsic information exchanged between the constituent RSC codes. Additionally, a puncturer is used to increase the achievable bandwidth efficiency. At the receiver self-iterative decoding is invoked between the hypothetical decoder components. The performance of the system was evaluated using the H.264/AVC source codec for interactive video telephony. Furthermore, EXIT charts were utilised in order to analyse the convergence behaviour of the SECCC scheme advocated. We demonstrate the efficiency of this approach by showing that the video quality is significantly improved, when using the binary SECCC scheme. More explicitly, the proposed system exhibits an E_b/N_0 gain of $6\ dB$ at the PSNR degradation point of $2\ dB$ in comparison to the identical-rate benchmarker carrying out RSC coding and puncturing, while communicating over correlated Rayleigh fading channels.**

I. Motivation and Background

Interactive wireless video telephony constitutes an attractive service. However, the band-limited nature of communication networks imposes the constraint of maintaining a low bit-rate, associated with highly compressed video. Unfortunately, the employment of standard video coding techniques, such as variable length coding, make the coded stream sensitive to channel-induced errors. Therefore, some form of forward error correction (FEC) must be employed to protect the compressed video stream. Various error protection schemes utilising both source and channel coding for video communication using turbo-style transceivers have been designed in [1]. Furthermore, the secure transmission of H.264 coded video using iterative joint source-channel decoding (JSCD) is investigated in [2].

In recent years, intensive research efforts have been invested in the design of concatenated coding schemes, a philosophy, which was originally proposed by Forney [3]. The turbo principle of parallel concatenated convolutional codes (PCCCs or turbo codes) employs two or more constituent codes, which exchange extrinsic information in order to operate near the Shannon limit [4, 5], as detailed in [6]. Since their invention they have found wide acceptance in bandwidth-limited communication systems. Turbo codes have also been adopted by numerous international standards, such as the UMTS/3GPP standard for personal communications [7], the European Telecommunications Standards Institute (ETSI) standard for digital video broadcasting (DVB) [8] and the Consultative Committee for Space Data Systems (CCSDS) standard for deep-space telemetry [9]. The family of serially concatenated convolutional codes (SCCCs) combined with interleavers [10] is also capable of achieving a performance comparable, and in some cases superior to that of classic turbo codes, especially at very low bit-error rates (BERs). Similarly, Benedetto *et al.* [11] and Loeliger [12] proposed another attractive FEC code family, namely iteratively-decoded self-concatenated convolutional codes (SECCC-ID).

The extrinsic information transfer (EXIT) chart concept was proposed by ten Brink [13] as a tool designed for analysing the convergence behaviour of iteratively decoded systems. EXIT charts are capable of predicting the SNR value required for achieving an infinitesimally low BER. In [14] EXIT charts were used to design binary self-concatenated convolutional codes employing iterative decoding for exchanging extrinsic information between hypothetical decoder components.

New concepts and coding tools have been introduced in state-of-the-art video coding standards [15]. Various error resilient features, such as flexible macro-block ordering (FMO), data partitioning (DP) and slice structuring of video pictures were also incorporated in video coding standards such as H.264/AVC [16]. However, in reality the source codes designed without taking into account the presence of channel decoding errors tend to have a poor performance. In a high-compression source codec even a low number of erroneous bits within an entropy coded sequence typically results in a high video distortion in the reconstructed sequence. This imposes stringent bit error ratio (BER) requirements on the channel coding scheme employed in order to attain an acceptable video quality at the receiver. Various error resilient schemes have been proposed in [1] in order to alleviate these problems, but the price paid is a potential reduction of the achievable compression efficiency and an increased computational complexity. Instead of the traditional serial concatenation of the classic Variable Length Codes (VLC) with a channel code, a parallel concatenated coding scheme was presented in [17], where the VLCs were combined with a turbo code. An optimised bit rate allocation scheme using a rate-1 inner channel encoder along with 3 to 6-bit source mapping was proposed in [18], and its performance was evaluated relative to conventional ISCD using a rate-$\frac{1}{2}$ recursive non-systematic convolutional (RNSC) inner code.

In this system design study we use low complexity near-capacity iteratively decoded binary self-concatenated codes, involving only a single encoder and a single decoder for the transmission of source coded stream. This technique is suitable for low-complexity video-telephony, which requires a low transmission power. Furthermore, the practically achievable interactive video performance trends are quantified when using state-of-the-art video coding techniques, such as H.264/AVC. Explicitly, an E_b/N_0 gain of 6 dB is attained using SECCC in comaprison to the identical-rate state-of-the-art benchmarker.

The organisation of the paper is as follows. We present our proposed system model in Section II. The performance of the system is characterised with the aid of EXIT chart analysis in Section III, while the overall performance results are presented in Section IV. Finally, we offer our conclusions in Section V.

II. System Model

The schematic of the proposed system is shown in Figure 1. We considered a rate-$1/2$ SECCC scheme as an example to evaluate its performance benefits relative to those of conventional RSC codes. Additionally, we considered the transmission of Gray-coded Quadrature Phase-Shift Keying (QPSK) modulated symbols over correlated Rayleigh fading channels. As shown in Figure 1, the video source signal is encoded using the H.264 video codec. The error resilient feature of slice structuring of video pictures in the H.264 codec was employed, which results in the partitioning of each video frame into multiple slices, each independently coded from the others. Additionally, we also incorporated data partitioning (DP), which results in three different sensitivity video coded streams per video slice, containing different coding elements and parameters. In our system setup, we organised each type of stream with multiple occurance in each frame for the different slices, into a single stream of each type. Subsequently, the three streams are concatenated, to generate a single stream s_1 as shown in Figure 1, before being encoded using the SECCC encoder considered. The input of the self-concatenated encoder is interleaved using the bit-interleaver Π_1 of Figure 1 to generate the bit-sequence s_2. The employment of the interleaver Π_1 renders the information bits more-or-less uncorrelated. This is a necessary requirement for the beneficial employment of iterative decoding and for accurate EXIT chart analysis, because they require the Log-Likelihood-Ratios (LLRs) of the information bits to be independent Gaussian distributed. The resultant bit-sequences s_1 and s_2 are parallel-to-serial converted, before they are fed to the RSC encoder. More specifically, we used a rate $R_1 = 1/2$ RSC encoder having generator polynomials of $G_0 = 7$, $G_1 = 5$ and a memory of $v = 2$. The SECCC encoded bits are interleaved using the interleaver Π_2 of Figure 1, in order to randomise the coded bits, and are subsequently punctured using the rate $R_2 = 1/2$ puncturer of Figure 1. The employment of puncturing results in an increased bandwidth efficiency η. Each of the puncturer blocks of Figure 1 punctures a single bit out of the two SECCC encoded bits. Hence the overall code rate R_{total} can be calculated as:

$$R_{total} = \frac{R_1}{2 \times R_2} = \frac{1}{2}\left(\frac{1}{2(\frac{1}{2})}\right) = \frac{1}{2}. \qquad (1)$$

The resultant bits are then mapped to the QPSK symbol sequence m before transmission over the communication channel. At the receiver side the additive white Gaussian noise (AWGN) contaminated signal $m' = hx + n$ is received, where h is the channel's non-dispersive fading coefficient and n is the AWGN having a variance of $N_0/2$ per dimension. The resultant signal is then fed into the soft-demapper in order to generate the conditional probability

$$P\left(m' \mid m^{(i)}\right) = \frac{1}{\pi N_0} \exp\left(-\frac{\mid m' - hm^{(i)} \mid^2}{N_0}\right) \qquad (2)$$

of receiving m', provided that $m^{(i)}$ was transmitted, where $i \in \{0, 1, 2, 3\}$.

The resultant probability is then passed to the soft-depuncturer and the resultant values are converted to bit-based Log-Likelihood Ratios (LLRs). Additionally, the puncturer inserts zero LLRs at the punctured bit-positions. The LLRs are then deinterleaved using the deinterleaver Π_2^{-1} of Figure 1 and are fed to the Soft-Input Soft-Output (SISO) maximum a-posteriori probability (MAP) decoder [10]. The self-concatenated decoder of Figure 1 iteratively exchanges extrinsic information, represented in the form of LLRs, to assist the SECC decoder in approaching the point of perfect convergence, represented as the $(1, 1)$ point in the EXIT-chart. The variable $L(.)$ is used to represent the respective bit-LLRs, where the specific type of the LLRs is identified by the superscript a, p and e, corresponding to $a\ priori$, $a\ posteriori$ and $extrinsic$ information, respectively. Observe in Figure 1 that the extrinsic LLR $L^e(s_1)$ is generated by subtracting the $a\ priori$ information $L^a(s_1)$ from the $a\ posteriori$ LLR values $L^p(s_1)$ at the output of the SECCC based soft-input soft-output (SISO) MAP decoder, as shown in Figure 1. The LLRs $L^e(s_1)$ are interleaved using the softbit interleaver Π_1 and are passed to the MAP decoder in order to produce the $a\ posteriori$ LLR values $L^p(s_2)$. Similarly, observe in Figure 1 that the $extrinsic$ LLR values $L^e(s_2)$ of the bit sequence s_2 are obtained by subtracting the $a\ priori$ information LLRs $L^a(s_2)$ input to the MAP decoder from the generated $a\ posteriori$ LLRs $L^p(s_2)$. Following deinterleaving, these $extrinsic$ LLRs $L^e(s_2)$ are fed back to the MAP decoder as the $a\ priori$ information $L^a(s_2)$. The $a\ priori$ information $L^a(s_1)$ and $L^a(s_2)$ is exploited by the self-concatenated decoder for the sake of providing improved $extrinsic$ information in the successive iterations.

III. EXIT Chart Analysis

EXIT charts [13] constitute an effective tool designed for analysing the convergence behaviour of concatenated systems based on the exchange of mutual information amongst the constituent receiver components. They analyse the input/output

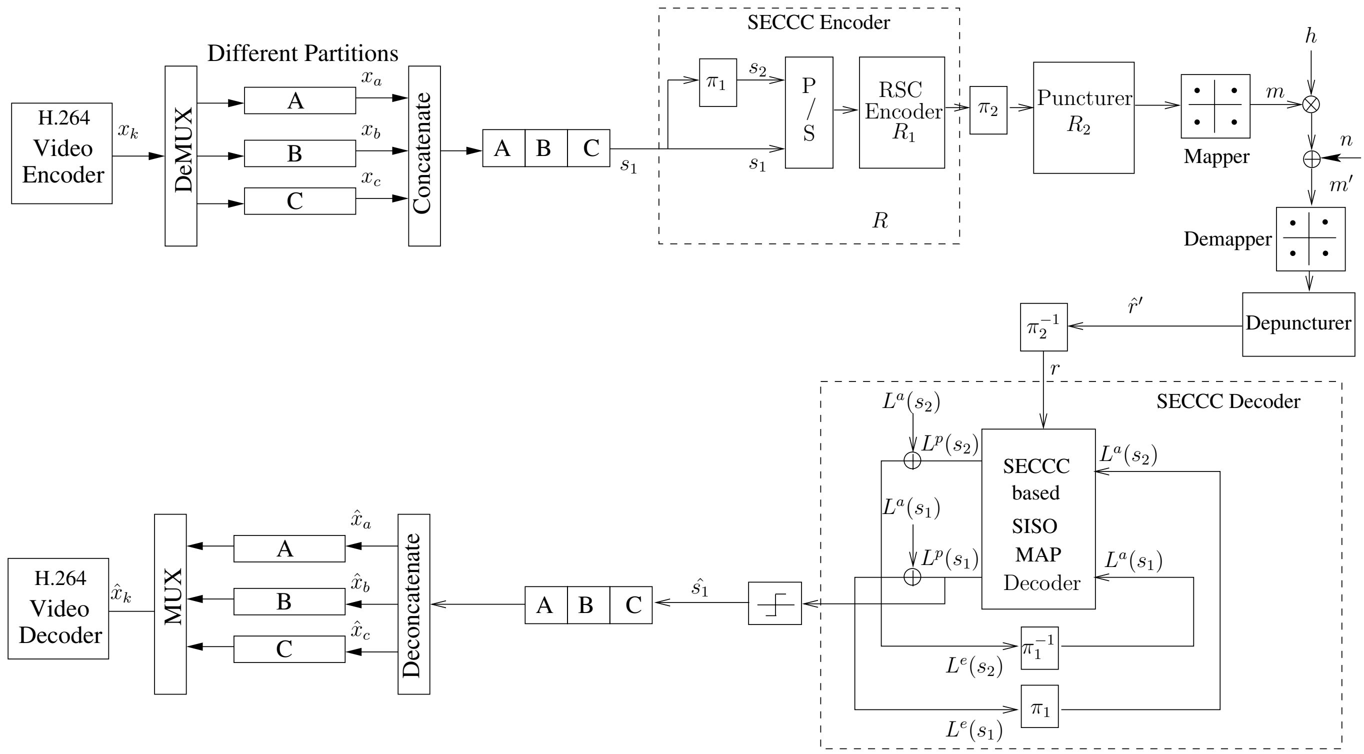

Fig. 1. The Proposed SECCC Aided Iterative System Decoding Model.

TABLE I
SYSTEMS PARAMETERS

System Parameters	Value	System Parameters	Value
Source Coding	H.264/AVC	No of MB's/Slice	11
Source Bit-Rate	64Kbps	Intra-frame MB update/frame	3
Frame Rate	15fps		
No of Slices/frame	9		
Inner module	Puncturer	Over-all Code Rate	1/2
Modulation Scheme	QPSK	Channel	Correlated Rayleigh Fading
Number of Transmitter Antennas	1		
Number of Receiver Antennas	1	Normalised Doppler Frequency	0.01
Interleaver Length	$\approx (128000/15)$		

mutual information characteristics of the SISO decoder's $a\ priori$ and $extrinsic$ LLR values with respect to the corresponding bit-decisions. More explicitly, for the symbol s having an $a\ priori$ LLR value of $L^a(s)$, the mutual information (MI) is denoted by $I_A(x)$, while the MI between the $extrinsic$ LLR $L^e(s)$ and the corresponding symbol x is denoted by $I_E(x)$. The EXIT charts are capable of successfully predicting the SNR, where the turbo-cliff region of the BER curve is expected to be for a concatenated code. This is achieved by identifying the E_b/N_0 value, which results in an open EXIT-chart tunnel between the EXIT curves of the corresponding constituent components.

The EXIT charts recorded for the binary SECCC scheme of Table II are shown in Figure 2. The two curves represents the two hypothetical decoder components of the self-concatenated scheme. Since they are identical components, we only computed the EXIT curve of a single component, while the other is its mirror image with respect to the diagonal line. The stair-case shaped function corresponds to the Monte-Carlo simulation based decoding trajectories of the proposed system at the E_b/N_0 values of 3, 4 and 5 dB for video frame indices 6, 2 and 3 of the 45-frame $Akiyo$ video sequence, respectively. The reason for plotting the EXIT chart together with the corresponding decoding trajectory is to visualise the transfer of extrinsic information between the decoder components.

The EXIT chart obtained for the self-concatenated code employed is similar to that of the parallel concatenated turbo trellis coded modulation (TTCM) schemes [19], where an open EXIT tunnel exists if the EXIT curves do not intersect with the diagonal line connecting the $(0, 0)$ point to the $(1, 1)$ point of the EXIT chart. According to the EXIT chart properties of iterative decoding, an infinitesimally low BER may only be achieved by an iterative receiver, if there is an open EXIT tunnel between the EXIT curves of the two constituent decoders, so that during the process of iterative decoding the outer decoder results in the highest possible extrinsic information of $I_E(outer) = 1$ at the input $a\ priori$ information of $I_a = 1$. Therefore, according to Figure 2, the EXIT curve of the combined SISO module constituted by the SECCC scheme cannot reach the $(1, 1)$ point of perfect convergence in the EXIT chart at the E_b/N_0 value of 3 dB, since the two curves intersect and hence they are unable to provide an open EXIT tunnel. This implies that an infinitesimally low BER cannot be achieved. However, it may be observed from the decoding trajectories of Figure 2, that for an E_b/N_0 value of 3 dB

or higher the SECCC scheme becomes capable of achieving the highest possible extrinsic information of $I_E(outer) = 1$ during the iterative decoding process.

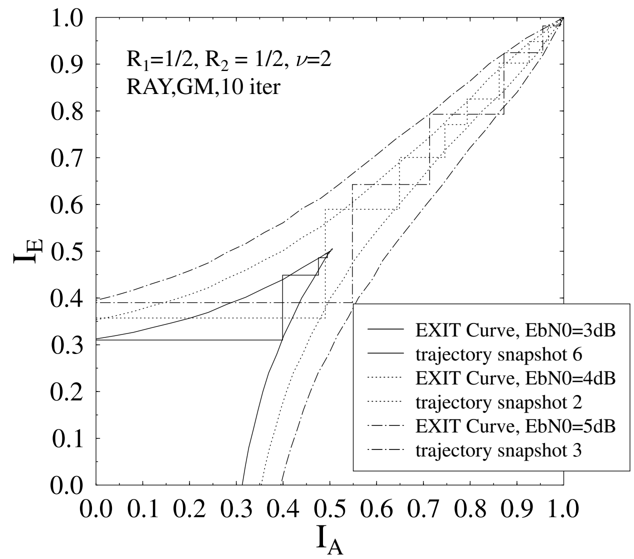

Fig. 2. The EXIT chart and simulated decoding trajectories for the SECCC scheme associated with $R_1 = 1/2$, $R_2 = 1/2$ and QPSK modulated at E_b/N_0 = 3, 4 and 5 dB, for transmission over correlated Rayleigh fading channel.

IV. System Performance Results

In this section we present our performance results characterising the proposed system. We provided the video source signal generated by the H.264/AVC as input to our proposed system model of Figure 1. The performance of this system was evaluated for the $Akiyo$ video test sequence represented in (176×144)-pixel Quarter Common Intermediate Format (QCIF) resolution. The test sequence was encoded using the H.264 codec at 15 frames-per-second (fps), with a target bitrate of 64 $kbps$. Various error resilient features of the H.264 video codec, such as the slice-based structure of the video frame and data partitioning (DP) were also exploited. Each frame was partitioned into 11 macro-blocks (MBs) per slice and there were 9 slices per QCIF frame. Additionally, intra-frame coded MB update of three randomly distributed MBs per QCIF frame was also incorporated, which associated us in avoiding avalanche like error propagation to the successive frames. Each intra or 'I' frame was followed by 44 predicted or 'P' frames, which curtailed error propagation beyond the 45-frame boundary. Therefore, the video sequence was refreshed with the aid of an 'I' frame after every 3-second interval. Additionally, the motion search was restricted to only the immediately preceding frame, in order to reduce the computational complexity of the video decoding process. For reasons of reduced computational complexity, we also avoided the insertion of bi-directionally predicted 'B' pictures, which would result in unacceptable loss of lip-synchronisation owing to its processing delay. Similarly, in order to keep the encoder's complexity realistic for the real-time videophone implementation, the error resilient flexible macro-block ordering (FMO) [16] was turned off, because it typically results in modest video performance improvements in low-motion video sequences, despite its substantial increase in computational complexity. Additional source codec parameters include the employment of quarter-pixel-accuracy motion estimation, intra-frame MB update and the use of Universal Variable Length Coding (UVLC) type entropy coding. The remaining system parameters of our experimental setup are listed in Table I.

The performance of the system was evaluated while considering $I_t = 10$ iterations within the joint SECCC system module. Additionally, for the sake of increasing the confidence in our results, we repeated each 45-frame video transmission experiment 200 times with a specific number of system iterations and averaged the results generated.

The BER performance of the proposed error protection schemes is shown in Figure 3. It can be observed from the Figure 3 that as expected, the SECCC scheme results in the best BER performance, when compared to the identical-rate RSC-coded benchmarker scheme.

Furthermore, the $PSNR$ versus E_b/N_0 curve of the proposed error protection schemes is portrayed in Figure 4. It may be observed in Figure 4 that the SECCC scheme employing $R_1 = 1/2$, $R_2 = 1/2$ and $I_t = 10$ iterations results in the best PSNR performance across the entire E_b/N_0 region considered. Quantitatively, when using the SECCC scheme of Table II, an E_b/N_0 gain of upto 6 dB may be achieved relative to the RSC-coded benchmarker scheme at the PSNR degradation point of 2 dB, as shown in Figure 4.

Finally, the subjective video quality of the error protection schemes employed is characterised in Figure 5. In order to have a pertinent subjective video quality comparison, the video frames presented in Figure 5 were obtained by averaging the results after the repeated transmission of both the luminance and chrominance components of the $Akiyo$ video sequence 30 times. Observe from Figure 5 that an unimpaired video quality is attained by the SECCC scheme at an E_b/N_0 value of $5dB$. However, video impairment persist for the RSC-coded benchmarker scheme even at the relatively high E_b/N_0 values of 6 dB, 7 dB, and 8 dB, as shown in Figure 5.

TABLE II
Code rates for different Error Protection schemes

Error Protection Scheme	Code Rate		
	Outer Code	Inner code	Overall
SECCC Scheme	SECCC $R = \frac{1}{4}$	Puncturer $R_2 = \frac{1}{2}$	1/2
Benchmarker Scheme	RSC $R = \frac{1}{4}$	Puncturer $R_2 = \frac{1}{2}$	1/2

V. Conclusions

In this paper, we presented a SECCC scheme design constituted by self-concatenated and iteratively decoding based components, designed for near-capacity channel coding. The ysed with the aid of EXIT charts. Explicitly, the proposed design example using an SECCC scheme exhibited an E_b/N_0 gain of 6 dB at the PSNR degradation point of 2 dB relative to the identical-rate benchmarker employing equivalent-rate channel coding. Our future work will focus on the design of a SECCC based three-stage system design constituted by serially concatenated and iteratively decoded short block codes (SBCs), a unity rate precoder and a multi-dimensional sphere packing (SP) modulation scheme, operating closer to capacity, while maintaining a high bandwidth efficiency.

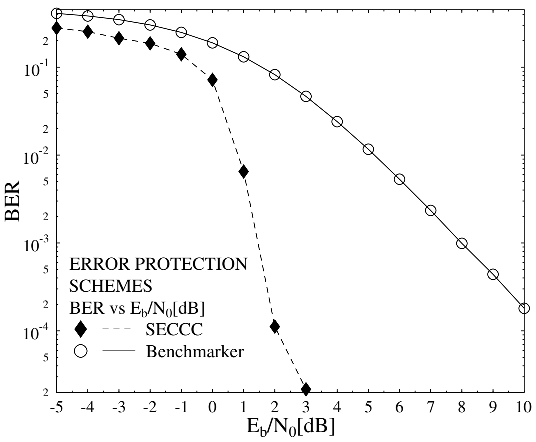

Fig. 3. BER versus E_b/N_0 performance of the various error protection schemes summarised in Table II.

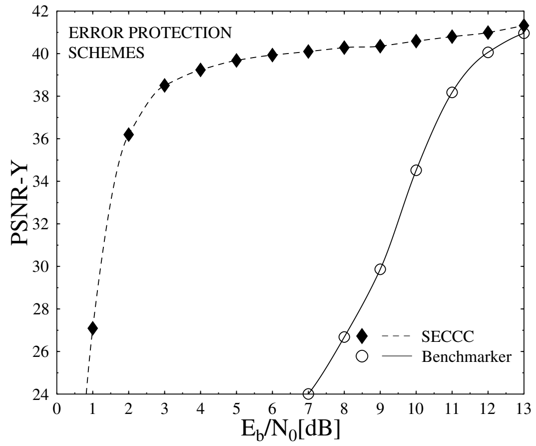

Fig. 4. PSNR-Y versus E_b/N_0 performance of various error protection schemes summarised in Table II.

Fig. 5. Subjective video quality of the 45^{th} "$Akiyo$" video sequence frame using (from top) SECCC Scheme and Benchmarker scheme summarised in Table II at E_b/N_0 values of (from left) 5 dB, 6 dB, 7 dB and 8 dB.

SECCC invoked puncturers, in order to increase the achievable bandwidth efficiency. The employment of the SECCC scheme provided significant improvements in terms of the $PSNR$ versus E_b/N_0 performance, when compared to the benchmarker scheme employing an equivalent-rate RSC code. Additionally, the convergence behaviour of the proposed system was anal-

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.* Wiley-IEEE Press, 2007.

[2] Nasruminallah and L. Hanzo, "Exit-chart optimised short block codes for iterative joint source and channel decoding in H.264 video telephony," *IEEE Transactions on Vehicular Technology*, 2009.

[3] G. Forney, "Concatenated codes," *MIT Press, Cambridge*, 1966.

[4] C. E. Shannon, "A mathematical theory of communication," *The Bell System Technical Journal*, vol. 27, pp. 379–423,623–656, July 1948.

[5] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near Shannon limit error-correcting coding and decoding:turbo-codes. 1," in *IEEE International Conference on Communications*, vol. 2, (Geneva, Switzerland), pp. 1064–1070, May 1993.

[6] L. Hanzo, T. H. Liew, and B. L. Yeap, *Turbo Coding, Turbo Equalisation and Space-Time Coding for Transmission over Fading Channels.* New York, NY, USA: John Wiley & Sons, 2002.

[7] 3rd Generation Partnership Project, *Multiplexing and channel coding*. 3G TS 25.212, 1999.

[8] European Telecommunications Standards Institute (ETSI), *Digital video broadcasting (DVB). Interaction channel for Satellite Distribution Systems.* DVB-RCS001, Feb. 2000.

[9] Consultative Committee for Space Data Systems, *Recommendations for space data systems, telemetry channel coding*. BLUE BOOK, May 1998.

[10] S. Benedetto, D. Divsalar, G. Montorsi, and F. Pollara, "Serial concatenated trellis coded modulation with iterativedecoding," in *Information Theory. 1997. Proceedings., 1997 IEEE International Symposium on*, (Ulm, Germany), June/July 1997.

[11] S. Benedetto, D. Divsalar, G. Montorsi, and F. Pollara, "Self-concatenated codes with self-iterative decoding for power andbandwidth efficiency," in *Information Theory, 1998. Proceedings. 1998 IEEE International Symposium on*, (Cambridge, MA, USA), Aug. 1998.

[12] H. Loeliger, "New turbo-like codes," in *in Proc. IEEE International Symposium on Information Theory*, pp. 9–10, June/July 1997.

[13] S. ten Brink, "Designing iterative decoding schemes with the extrinsic information transfer chart," in *AEU International Journal of Electronics and Communications*, vol. 54, pp. 389–398, Nov. 2000.

[14] M. F. U. Butt, R. A. Riaz, S. X. Ng, and L. Hanzo, "Near-capacity iteratively decoded binary self-concatenated code design using EXIT charts," in *IEEE Global Communications Conference, GLOBECOM '08*, (New Orleans, USA), Nov./Dec. 2008.

[15] J. Ostermann, J. Bormans, P. List, D. Marpe, M. Narroschke, F. Pereira, T. Stockhammer, and T. Wedi, "Video coding with H.264/AVC: tools, performance, and complexity," *IEEE Circuits and Systems Magazine*, vol. 4, no. 1, pp. 7–28, 2004.

[16] T. Stockhammer, M. M. Hannuksela, and T. Wiegand, "H.264/AVC in wireless environments," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 13, pp. 657–673, July 2003.

[17] J. Liu, G. Tu, C. Zhang, and Y. Yang, "Joint source and channel decoding for variable length encoded turbo codes," *EURASIP J. Adv. Signal Process*, vol. 2008, no. 1, pp. 1–10, 2008.

[18] M. Adrat, P. Vary, and T. Clevorn, "Optimized bit rate allocation for iterative source-channel decoding and its extension towards multimode transmission," in *in Proceedings of IST Mobile and Wireless Communications Summit (Dresden, Germany)*, pp. 1153–1157, June 2005.

[19] J. Kliewer, S. X. Ng, and L. Hanzo, "Efficient Computation of EXIT Functions for Nonbinary Iterative Decoding," *IEEE Transactions on Communications*, vol. 54, pp. 2133–2136, Dec. 2006.