

Iterative Parallel Interference Cancellation Aided CDMA Based MPEG-4 Video Telephony

J.Y. Chung, H. Wei and L. Hanzo¹

School of ECS, University of Southampton, SO17 1BJ, UK.

Tel: +44-23-8059-3125, Fax: +44-23-8059-4508

Email:lh¹@ecs.soton.ac.uk, <http://www-mobile.ecs.soton.ac.uk>

Abstract— In this contribution we investigate an iterative Parallel Interference Cancellation (PIC) aided CDMA MPEG-4 video phone scheme. The number of iterations used by the PIC-based turbo receiver was four. The family of m -sequence based spreading codes having a length of $N = 15$ chips was employed in our system and the transmission burst length was $T = 120$ bits. A random channel interleaver having a memory of eight transmission bursts was employed. The system considered supported $K = 7$ users and communicated over an asynchronous symbol-spaced, equal-weight two-path uncorrelated Rayleigh fading channel.

I. INTRODUCTION

The Motion Pictures Expert Group's MPEG-4 codec was designed for object-based audiovisual coding, invoking numerous different 'coding tools' [1]. The so-called Video Objects (VO), which are not necessarily rectangular-shaped blocks, are encoded and transmitted separately. The VOs are reconstructed at the decoder, allowing the receiver to choose how a video scene is presented.

Conventional video sequences are coded in MPEG-4 much as they were in the superseded MPEG-1 and MPEG-2 coding schemes. More explicitly, in its classic encoding mode the MPEG-4 codec uses block-based Motion Compensation (MC) and the MC error residual texture is Discrete Cosine Transformed (DCT) before its encoding. Several parts of the MPEG-4 standard address various error resilience issues [2], rendering it suitable for use in error-prone environments, such as wireless video telephone applications.

In reference [3], a Parallel Interference Cancellation (PIC) based iterative Multiuser Detector (MUD) was proposed, which was further developed in [4], employing numerous sophisticated turbo channel coding schemes. The iterative receiver advocated in this treatise obeys a similar structure to that proposed in [3], except for employing half-rate turbo channel coding. In [5] and [6] the proposed iterative multiuser detectors invoked soft estimation of the real-valued BPSK modulated signal. The Log-MAP turbo decoder [7] was employed and the Recursive Systematic Convolutional (RSC) code's memory length was $m = 3$.

The financial support of the Mobile VCE, UK is gratefully acknowledged.

This article is organized as follows. In Section II we briefly introduce the concept of the PIC scheme, while in Section III we provide a brief MPEG-4 video error resilience overview. In Section IV we characterize the achievable performance. Finally, in Section V we offer our conclusions.

II. ITERATIVE MULTIUSER DETECTION ASSISTED DS-CDMA

II-A. Turbo Multiuser Detection

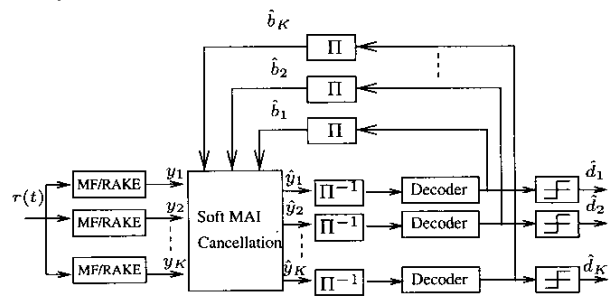


Fig. 1. Schematic of the iterative PIC-aided detector.

Figure 1 shows the schematic of the PIC assisted iterative multiuser detector designed for DS-CDMA communications, where the Matched Filter's (MF) output related to the k users' received signal is given by: $\mathbf{y} = [\mathbf{y}_1, \dots, \mathbf{y}_K]^T$.

Let us assume that the turbo channel decoder generates the received symbol probability $\Pr\{\mathbf{b}_k = \mathbf{x}_m | \hat{\mathbf{y}}_k\}$, $k = 1 \dots K, m = 1 \dots M$, where M is the size of the M -array modulated signal constellation. Then the estimated symbol b_k of user k at a particular time instant may be expressed at the output of the soft Multiple Access Interleaver (MAI) cancellation block of Figure 1 as:

$$\hat{b}_k = \sum_{m=1}^M \Pr\{\mathbf{b}_k = \mathbf{x}_m | \hat{\mathbf{y}}_k\} \cdot \mathbf{x}_m, \quad (1)$$

where $x_m, m = 1 \dots M$ represents the complex-valued M -array symbols. We define the vector $\hat{\mathbf{b}} = [\hat{b}_1, \dots, \hat{b}_K]^T$, seen at the top of Figure 1, which contains the estimated symbol of all the K users. The blocks Π and Π^{-1} in

Figure 1 denote the channel interleaver and deinterleaver, respectively, which arrange the bits at the input of the channel decoders and the MAI canceller in the required order. Once we obtained the soft symbol estimates, soft MAI cancellation is performed, as shown in Figure 2, yielding the decontaminated soft output signal vector $\hat{\mathbf{y}} = [\hat{y}_1, \dots, \hat{y}_K]^T$. Figure 2 shows explicitly that the soft estimates \hat{b}_k , $k = 1, \dots, K$ of the transmitted symbols, which are output by the channel decoder, are utilized for reconstructing the estimates of the transmitted signal of all the K users. Then, for each user, the reconstructed estimated signals of all the interfering users are subtracted from the composite multiuser signal and the resultant decontaminated signal is processed by a Matched Filter (MF) or RAKE receiver, generating the soft output \hat{y}_k . These modulated signal reconstruction, interference cancellation and re-estimation steps are repeated, until the affordable number of iterations is exhausted.

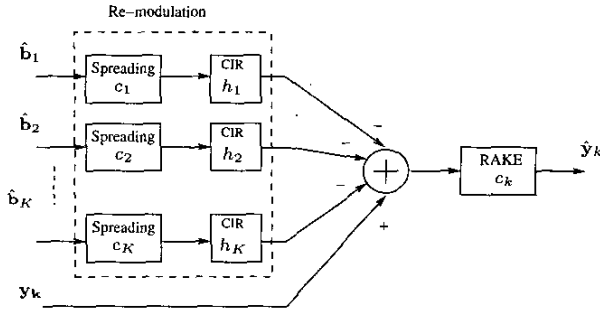


Fig. 2. Schematic of a single PIC cancellation stage.

In our system, half-rate turbo codes [8] are employed. The Log-MAP turbo decoder [7] was employed and the RSC code's memory length was $m = 3$. However, our initial investigations indicated that no significant performance improvements were achieved, when more than four PIC-based outer iterations were invoked. Furthermore, within each outer PIC iteration, two turbo decoding iterations were employed. The family of m -sequence based spreading codes having a length of $N = 15$ was employed in our system and the transmission burst length was $T = 120$ bits. A random channel interleaver denoted by Π in Figure 1 and a memory of eight transmission bursts was employed in our system. The system considered supported $K = 7$ users and communicated over an asynchronous equal-weight symbol-spaced two-path uncorrelated Rayleigh fading channel.

II-B. Asynchronous Interference Cancellation

In an asynchronous system the symbols transmitted during the i th signalling interval are also affected by all the asynchronous interfering signals transmitted during the intervals $(i + 1)$ and $(i - 1)$ owing to the dispersed MUI imposed by other users. Hence, according to [9], for an asynchronous DS-CDMA system communicating over

a non-dispersive AWGN channel the received signal vector \mathbf{y} can also be expressed as:

$$\mathbf{y}[i] = \mathbf{R}^T[\mathbf{1}]\mathbf{A}\mathbf{b}[i+1] + \mathbf{R}[\mathbf{0}]\mathbf{A}\mathbf{b}[i] + \mathbf{R}[\mathbf{1}]\mathbf{A}\mathbf{b}[i-1] + \mathbf{n}[i], \quad (2)$$

where i is the index of time instant and the zero-mean Gaussian channel noise process $\mathbf{n}[i]$ has the crosscorrelation matrix:

$$E[\mathbf{n}[i]\mathbf{n}^T[j]] = \begin{cases} \sigma^2\mathbf{R}^T[\mathbf{1}], & \text{if } j = i + 1; \\ \sigma^2\mathbf{R}[\mathbf{0}], & \text{if } j = i; \\ \sigma^2\mathbf{R}[\mathbf{1}], & \text{if } j = i - 1; \\ \mathbf{0}, & \text{otherwise,} \end{cases} \quad (3)$$

while the matrix $\mathbf{R}[\mathbf{1}]$ and $\mathbf{R}[\mathbf{0}]$ are defined as:

$$\mathbf{R}_{jk}[\mathbf{0}] = \begin{cases} 1, & \text{if } j = k; \\ \rho_{jk}, & \text{if } j < k; \\ \rho_{kj}, & \text{if } j > k, \end{cases} \quad (4)$$

$$\mathbf{R}_{jk}[\mathbf{1}] = \begin{cases} 0, & \text{if } j \geq k; \\ \rho_{kj}, & \text{if } j < k. \end{cases} \quad (5)$$

Let us now define the vectors $\hat{\mathbf{b}}_k[i+1]$, $\hat{\mathbf{b}}_k[i]$, $\hat{\mathbf{b}}_k[i-1]$ as the soft estimates of the k th user during the signalling instants $(i + 1)$, i , and $i - 1$, which have had their k th element set to zero. Therefore, after the soft MAI cancellation stage of Figure 1, the decontaminated signal vector $\hat{\mathbf{y}}_k[i]$ can be represented as:

$$\hat{\mathbf{y}}_k[i] = \mathbf{R}^T[\mathbf{1}]\mathbf{A}(\mathbf{b}[i+1] - \hat{\mathbf{b}}_k[i+1]) + \mathbf{R}[\mathbf{0}](\mathbf{b}[i] - \hat{\mathbf{b}}_k[i]) + \mathbf{R}[\mathbf{1}](\mathbf{b}[i-1] - \hat{\mathbf{b}}_k[i-1]). \quad (6)$$

According to Equation 6, we can observe that the decontaminated signal vector $\mathbf{y}[i]$ will be free from multiple access interference, provided that the symbol estimates $\hat{\mathbf{b}}_k[i+1]$, $\hat{\mathbf{b}}_k[i]$, $\hat{\mathbf{b}}_k[i-1]$ are reliable. The performance of this PIC-aided transceiver was characterized in [4] in conjunction with a suite of channel codecs. In our two-path dispersive channel the effects of the dispersed asynchronous MUI are also aggravated by channel-induced Inter-Symbol Interference (ISI).

III. MPEG-4 ERROR RESILIENCE TOOLS

Owing to the employment of Variable Length Codes (VLC) in the MPEG-4 standard, the compressed video bitstreams are particularly sensitive to channel errors [10]. This is, because the boundary between VLC words is not explicit. Transmission errors typically lead to an incorrect number of bits being used during the VLC decoding process, inflicting loss of synchronization. As a further consequence of VLC encoding, the location where the decoder detects an error in the MPEG-4 encoded bitstream is not the same, as the location where the error has actually occurred. Fortunately, the error resilience tools employed in the MPEG-4 video standard assist in minimizing the

amount of data that has to be discarded, whenever errors are detected. The basic error-resilience tools included in the codec [2] are: Resynchronization markers, Data partitioning, Header Extension Codes (HEC) and Reversible Variable Length Codes (RVLC).

Apart from the tools mentioned above, error concealment [11] has also been implemented in the decoder. Finally, for the sake of limiting the error propagation across consecutive video frames, intra-frame coded images are transmitted at regular intervals.

IV. EXPERIMENTAL RESULTS

Parameter	Value
Channel type	Two-path, symbol-spaced, equal-weight Rayleigh fading Channel Impulse Response (CIR)
Doppler frequency	100Hz
Spreading factor (m-seq.)	15
Receiver type	Iterative PIC
Modulation type	QPSK
Channel code	1/2-rate turbo code
Transmission burst length	120 bits
No. of PIC iterations	4
No. of turbo decoder iterations	2
No. of users	7
Video codec	MPEG-4
Video frame-rate	30 frame/sec
Video format	176 × 144 QCIF
Video bitrate	69kbps
Channel coded bitrate	150kbps
Interleaver length	960 bit

TABLE I

MODULATION AND CHANNEL PARAMETERS

The performance of the proposed MPEG-4 based video telephone scheme is characterized in this section in terms of the average video Peak Signal to Noise Ratio (PSNR) versus Channel SNR. Our simulations were carried out using the (174 × 144)-pixel Quarter Common Intermediate Format (QCIF) "Miss America", "Foreman" and "News" video sequences encoded at a video bitrate of 69 kbps. The MPEG-4 source encoder operated at the video frame rate of 30 frames per second. Table I summarizes the modulation and channel parameters employed. Figure 3 depicts the MPEG-4 codec's video performance in the context of the PIC aided CDMA system considered, expressed in terms of the average PSNR versus the Signal to Noise Ratio (SNR), namely E_b/N_o for transmission over an uncorrelated symbol-spaced two-path fading channel. By contrast, Figure 4 plots the PSNR results from a different perspective, namely as a function of the Bit Error Rate (BER). On the other hand, Figure 5 shows the BER versus channel SNR performance of our proposed scheme.

As seen in both Figure 3 and 4, at a given E_b/N_o or BER, the "Foreman" sequence resulted in the lowest PSNR image quality, followed by the "News" sequence

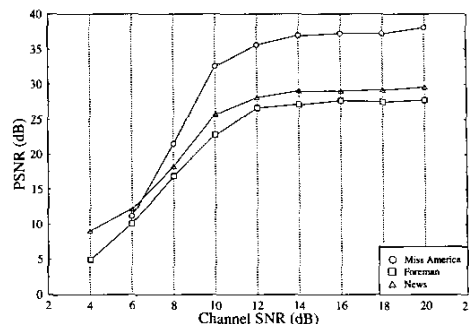


Fig. 3. Average PSNR versus E_b/N_o performance of the PIC assisted MPEG-4 videophone scheme, when supporting $K = 7$ users for transmission over an uncorrelated symbol-spaced, equal-weight two-path fading channel employing the parameters of Table I.

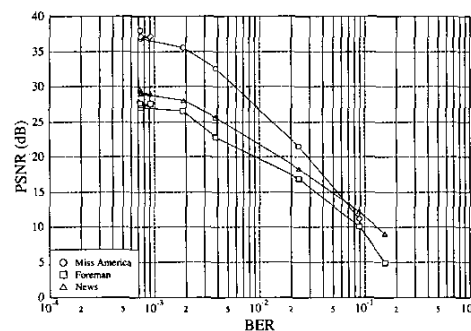


Fig. 4. Average PSNR versus channel BER performance of the PIC assisted MPEG-4 videophone scheme for $K = 7$ and transmission over an uncorrelated symbol-spaced, equal-weight two-path fading channel employing the parameters of Table I.

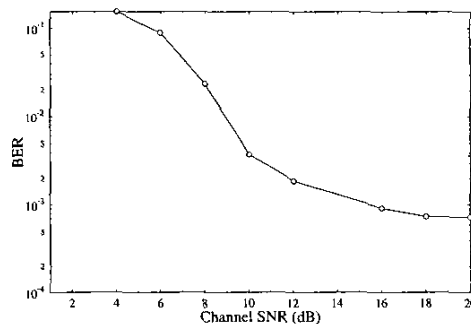
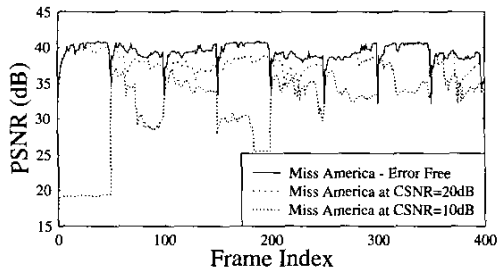
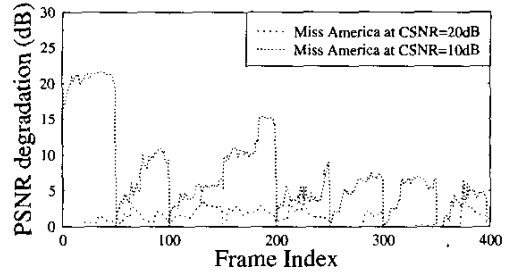


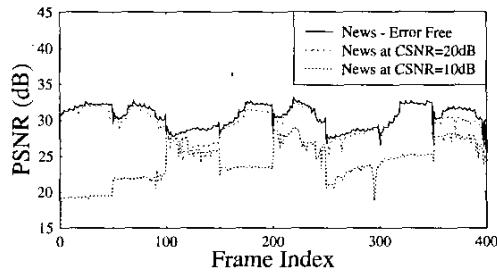
Fig. 5. Channel BER versus channel SNR performance of the PIC assisted MPEG-4 videophone scheme for $K = 7$ and transmission over an uncorrelated symbol-spaced, equal-weight two-path fading channel employing the parameters of Table I.



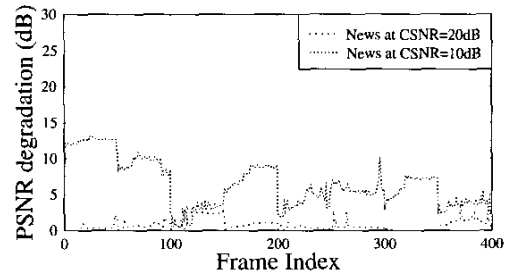
(a) Miss America Video Sequence



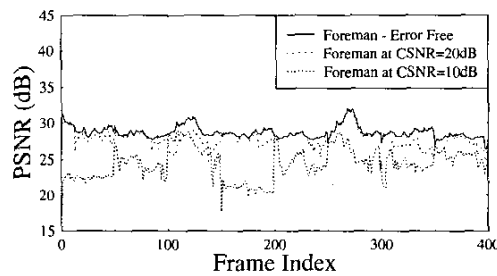
(a) Miss America Video Sequence



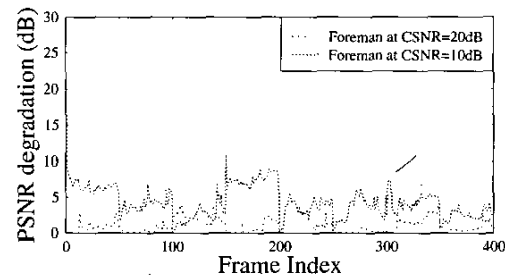
(b) News Video Sequence



(b) News Video Sequence



(c) Foreman Video Sequence



(c) Foreman Video Sequence

Fig. 6. PSNR versus video frame index of the PIC assisted MPEG-4 videophone scheme, when supporting $K = 7$ users for transmission over an uncorrelated symbol-spaced, equal-weight two-path fading channel employing the parameters of Table I.

Fig. 7. PSNR degradation versus video frame index of the PIC assisted MPEG-4 videophone scheme, when supporting $K = 7$ users for transmission over an uncorrelated symbol-spaced, equal-weight two-path fading channel employing the parameters of Table I.

and finally by the "Miss America" sequence. This is, because the "Miss America" sequence is a pure 'head and shoulders' type video sequence, which has the lowest motion activity among the three clips. Therefore less motion vectors have to be transmitted, potentially allowing the MPEG-4 codec to dedicate a larger fraction of the available bitrate budget to the representation of the fine details of texture-encoding, which increases the achievable PSNR. Even if some of the "Miss America" bistream is corrupted, the corrupted VOPs can be readily concealed by the previous VOP of the preceding video frame. The video performance of the proposed scheme is further illustrated in Figure 6, characterizing of each of the consecutive video frames' video quality. The PSNR value recorded at the channel SNR of 20dB (virtually error-free transmission scenario) and at the channel SNR of 10dB were shown in Figure 6. The subjective video quality at SNR=10dB becomes somewhat objectable. Figure 7 portrays the PSNR degradation versus video frame index for the same system.

Figure 8 illustrates the subjective video quality of a particular frame of the three video sequences at E_b/N_o of 20dB, where the average PSNR values of the "Miss America", "News" and "Foreman" sequences are 38.05 dB, 29.55 dB and 27.66 dB, respectively.

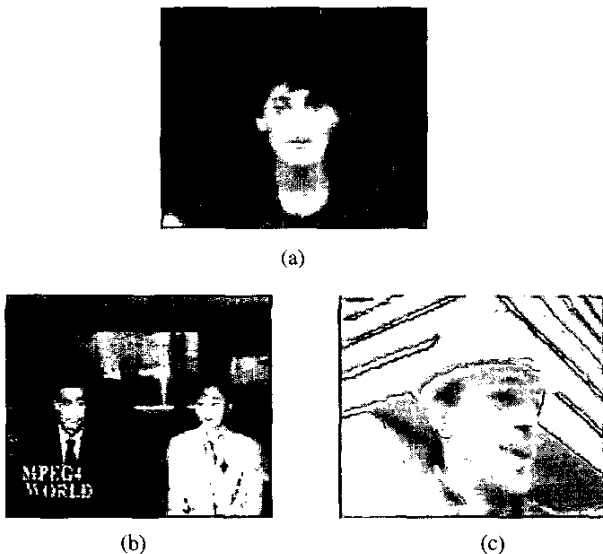


Fig. 8. 10th Frame of the (a) "Miss America", (b) "News" and (c) "Foreman" sequences when supporting $k = 7$ users for transmission over an uncorrelated fading channel at E_b/N_o of 20dB

V. CONCLUSIONS

In this contribution, a PIC-aided CDMA scheme was investigated, employing the MPEG-4 video codec. It was shown in the context of the PIC-CDMA MPEG4 video telephony scheme that at the video coding bitrate of 69kbps, the proposed system was capable of achieving an adequate

video quality.

VI. REFERENCES

- [1] ISO/IEC JTC1/SC29/WG11 W2502, "ISO/IEC 14496-2.", in *Final Draft International Standard. Part 2: Visual*, Atlantic City, 1998.
- [2] R. Talluri, "Error-resilient video coding in the ISO MPEG-4 standard", *IEEE Communications Magazine*, vol. 2, no. 6, pp. 112-119, June 1998.
- [3] M. C. Reed, *Iterative Receiver Techniques for Coded Multiple Access Communication System*, PhD thesis, The University of South Australia, 1999.
- [4] L. Hanzo, L. L. Yang, E. L. Kuan, and K. Yen, *Single- and Multi-Carrier DS-SS-CDMA*, John Wiley and IEEE Press, 2003, 1060 pages.
- [5] Y. Zhang, "Reduced complexity iterative multiuser detection for DS/SS-CDMA with FEC", *International Conference on Universal Personal Communications*, no. 12, pp. 10-14, October 1997.
- [6] A. J. Grant P. D. Alexander and M. C. Reed, "Performance analysis of an iterative decoder for code-division multiple-access", *European Transactions on Telecommunication*, vol. 9, no. 5, pp. 419-426, Sept/Oct 1998.
- [7] L. Hanzo, T. H. Liew, and B. L. Yeap, *Turbo Coding, Turbo Equalisation and Space-Time Coding for Transmission over Fading Channels*, John Wiley-IEEE Press, 2002.
- [8] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near Shannon limit error-correcting coding and encoding: Turbo-codes (1)", in *IEEE International Conference on Communications*, Geneva, Switzerland, May 1993, pp. 1064-1070.
- [9] S. Verdu, *Multiuser Detection*, Cambridge Press, 1998.
- [10] L. Hanzo, P.J. Cherriman and J. Street, *Wireless Video Communications: Second to Third Generation Systems and Beyond*, IEEE Press., NJ, USA, 2001.
- [11] Y. Wang, Q-F. Zhu, "Error control and concealment for video communications: A review", *Proceedings of IEEE*, vol. 86, no. 5, pp. 974-997, May 1998.