A BURST-BY-BURST ADAPTIVE JOINT-DETECTION BASED CDMA SPEECH TRANSCEIVER

H.T. How, T.H. Liew, E.L Kuan and L. Hanzo

Dept. of Electr. and Comp. Sc.,Univ. of Southampton, SO17 1BJ, UK.
Tel: +44-1703-593 125, Fax: +44-1703-594 508
Email: lh@ecs.soton.ac.uk http://www-mobile.ecs.soton.ac.uk

ABSTRACT

A burst-by-burst adaptive speech transceiver is proposed, which can drop its source coding rate and speech quality under transceiver control in order to invoke a more error resilient modem mode amongst less favourable channel conditions. The novel, high-quality, Adaptive Multi-Rate (AMR) speech codec [5], operated at bit rates of 4.75 and 10.2 kbps and combined with source-sensitivity-matched Redundant Residue Number Systems (RRNS) based channel codes. Burst-by-burst adaptive Joint-Detection based Code-Division Multiple Access (JD-CDMA) is used for transmitting the dual-rate bitstream generated by the AMR speech codec.

1. SYSTEM OVERVIEW

In recent years the concept of near-instantaneously adaptive transceivers has reached a state of maturity and various adaptive features have found their way in to standard systems [1]. The time-variant quality fluctuations of the mobile channel result in a time-variant bitrate [2] and hence special attention has to be devoted to contriving interactive multimedia systems, which are capable of efficiently accommodating these fluctuating bit rates [3, 4]. In this contribution, we propose a dual-mode burst-by-burst adaptive speech transceiver scheme, based on the Advanced Multi Rate (AMR) speech codec [5, 6], Redundant Residue Number System (RRNS) assisted channel coding [7] and Joint Detection aided Code-Division Multiple Access (JD-CDMA) [8]. The schematic of the proposed adaptive JD-CDMA speech transceiver is depicted in Figure 1. The mode switching is controlled by the channel quality fluctuations imposed by the time-variant channel. This is not a desirable scenario. However, we will endeavour to contrive measures in order to mitigate the associated perceptual speech quality fluctuations. The underlying trade-offs associated with employing two speech modes of the AMR standard speech codec in conjunction with a reconfigurable, unequal error protection BPSK/4QAM modem are investigated.

2. THE AMR SPEECH CODEC

The AMR codec employs the Algebraic Code-Excited Linear Predictive (ACELP) model [9]. Here we provide a brief overview of the AMR codec following the approach of [5, 6]. The AMR codec's complexity is relatively low and hence it can be implemented cost-efficiently. This codec operates on a 20ms frame of 160 speech samples, and generates encoded blocks of 95, 103, 118, 134, 148, 159, 204 and 244 bits/20ms. This leads to bit rates of 4.75, 5.15, 5.9, 6.7, 7.4, 7.95, 10.2 and 12.2 kbps, respectively. Explicitly, the AMR speech codec provides eight different modes. Multi-rate coding [10] allows a variation in the total allocation of bits for a speech frame, adapting the rate to the local phonetic character of the speech signal, the channel quality or network conditions. This is particularly useful in digital cellular communications, where one of the major challenges is that of designing a codec that is capable of providing high quality speech for a wide variety of channel conditions.

The codec mode adaptation is a key feature of the new AMR standard that has not been used in any prior mobile standard. At a given fixed gross bit rate, this mechanism of adapting the source coding rate has the potential of altering the partitioning between the speech source bit rate and the redundancy added for error protection. For a detailed description and the associated bit allocations of the AMR codec the interested readers are referred to [5]. Let us now briefly focus our attention on the robustness of the AMR codec against channel errors.

3. SPEECH CODEC'S ERROR SENSITIVITY

Some bits are significantly more sensitive to channel errors than others, and hence have to be better protected by the channel coder [11]. A commonly used approach in quanti-
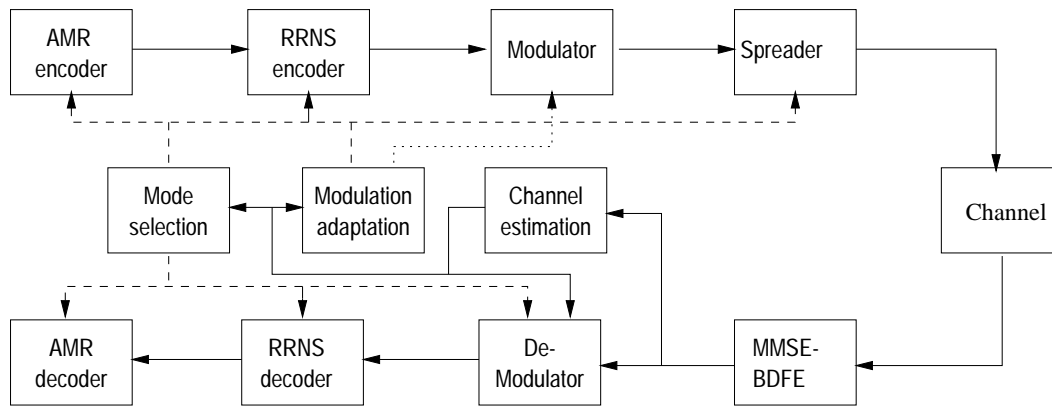
Figure 1: Schematic of the adaptive dual-mode JD-CDMA system
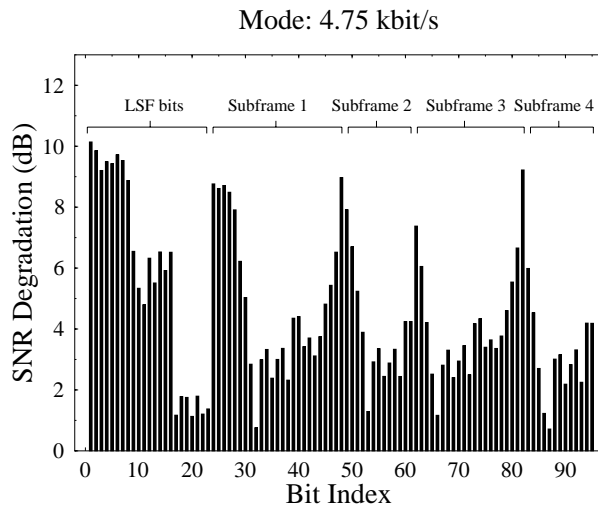
Figure 2: The SEGSNR Degradations due to 100% Bit Error Rate in the 95-bit, 20 ms AMR speech frame.

Class	RRNS Code	databits	Total databits	Total codedbits
4.75kbps/BPSK				
I	RRNS$(8,4)$	40		
II	RRNS$(8,5)$	25	95	160
III	RRNS$(8,6)$	30		
10.2kbps/4QAM				
I	RRNS$(8,4)$	60		
II	RRNS$(8,5)$	25	205	320
III	RRNS$(8,6)$	120		

Table 1: RRNS codes designed for two different modulation modes.

fying the sensitivity of a given bit is to invert this bit consistently in every speech frame and evaluate the associated Segmental SNR degradation. The error sensitivity of various bits for the AMR codec determined in this way is shown in Figure 2 for the bit rate of 4.75 kbps. It can be observed from Figure 2 that the most sensitive bits are those of the LSF subvectors.

The error sensitivity of the adaptive codebook delay is the highest in the first subframe, commencing at bit 24, as shown in Figure 2. The next group of bits is constituted by the quantization gains in decreasing order of bit sensitivity, as seen in Figure 2 between indices 40 and 49. The least sensitive bits are related to the fixed codebook pulse positions, seen at positions 54-61 in Figure 2. This is because, if one of the fixed codebook index bits is corrupted, the codebook entry selected at the decoder will differ from that used

in the encoder only in the position of one of the non-zero excitation pulses, ie the corrupted codebook entry will be similar to the original one. Hence, the algebraic codebook structure used in the AMR codec is inherently quite robust to channel errors. The information obtained here will be used to design the bit mapping procedure in order to assign the channel encoders according to the bit error sensitivities.

Let us now consider the channel coding aspects of our transceiver in the next section.

4. REDUNDANT RESIDUE NUMBER SYSTEM (RRNS) CHANNEL CODING

In order to improve the performance of the system, we employ the novel family of so-called the Redundant Residue Number System (RRNS) codes [7] for protecting the speech bits unequally, depending on their respective error sensitivities . Three different RRNS codes having different code rates are used to protect the three different classes of speech bits. In addition, the RRNS codes employed are also switched in accordance with the modulation modes and speech rates used in our system. The error sensitivity of the 4.75 kbps AMR codec's source bits was evaluated in Section 3 and the same procedures were applied in order to obtain the

2

Parameter	Value
Channel type	COST 207 Bad Urban (BU)
Paths in channel	7
Doppler frequency	80 Hz
Spreading factor	16
Chip rate	2.167 MBaud
JD block size	26 symbols
Receiver type	MMSE-BDFE
AQAM type	Dual-mode (BPSK, 4QAM)
Channel codec	Triple-class RRNS
Channel-coded Rate	8/16 kbps
Speech Codec	AMR (ACELP)
Speech Rate	4.75/10.2 kbps
Speech Frame Length	20 ms

Table 2: Transceiver Parameters

error sensitivity for the source bits of the 10.2 kbps AMR codec. Three protection classes were deemed to constitute a suitable trade-off between the system's complexity and performance.

5. JOINT DETECTION BASED ADAPTIVE CODE DIVISION MULTIPLE ACCESS (JD-CDMA)

Here we propose to combine joint detection CDMA [8] with AQAM, by modifying the approach used by Wong *et al.* [12]. Joint detection is particularly suitable for combining with AQAM, since the implementation of the joint detection algorithms does not require any knowledge of the modulation mode used [8]. The joint detection algorithm utilizes only the CIR estimates and the spreading sequences of all the users. Therefore, the joint detection receivers are suitable for combining with AQAM, since they do not have to be reconfigured each time the modulation mode is switched. Hence the associated complexity is independent of the modulation mode used.

The conditions invoked for switching between the two AQAM JD-CDMA modes were set according to their target BER requirements as:

$$\text{Mode} = \begin{cases} \text{BPSK} & \text{SINR} < t_1 \\ \text{4QAM} & t_1 \leq \text{SINR} \end{cases}, \qquad (1)$$

where t_1 represents the switching threshold between the two modes.

With the system elements described, we now focus our attention on the performance of the adaptive transceiver proposed.

6. SYSTEM PERFORMANCE

The simulation parameters used in our AQAM/JD-CDMA system are listed in Table 2. The channel profile used was the COST 207 Bad Urban (BU) channel [13] consisting of
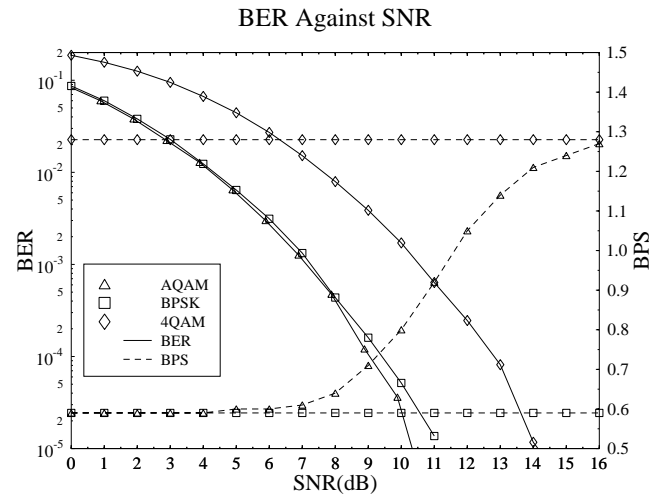

Figure 3: BER and BPS comparisons for fixed mode BPSK and 4QAM as well as for the AQAM/JD-CDMA system, using the RRNS codes of Table 1. The switching threshold for AQAM was set to 10.5 dB and the simulation parameters are listed in Table 2.

seven paths, where each path was faded independently at a Doppler frequency of 80 Hz. The switching threshold t_1 to change the modulation mode is set to 10.5dB.

In Figure 3, the average BER performance of the coded fixed-mode BPSK/JD-CDMA and 4QAM/JD-CDMA systems is presented together with that of the twin-mode AQAM assisted JD-CDMA system supporting two users. The performance of the AQAM scheme was evaluated by analyzing the BER and the throughput expressed in terms of the average number of bits per symbol (BPS) transmitted. At low channel SNRs the BER of the AQAM/JD-CDMA scheme mirrored that of BPSK/JD-CDMA. However, as the channel SNR increased, the BER performance of AQAM/JD-CDMA became better, than that of BPSK/JD-CDMA, as shown in Figure 3. This is because the 4QAM mode is employed more often, reducing the probability of using BPSK. Since the mean BER of the system is the ratio of the total number of bit errors to the total number of bits transmitted, the mean BER will decrease with decreasing number of bit errors or with increasing number of transmitted bits. For a fixed number of symbols transmitted, the total number of transmitted bits in a frame is constant for the fixed mode BPSK/JD-CDMA, while for AQAM/JD-CDMA the total number of transmitted bits increased, when the 4QAM/JD-CDMA mode was used. Consequently, the average BER of the AQAM/JD-CDMA system was lower than that of the BPSK/JD-CDMA scheme.

The Bits Per Symbol (BPS) throughput performance curve is also plotted in Figure 3. As expected, the number of BPS of both BPSK and 4QAM is constant for all channel

3

SNR values. They are limited by the modulation scheme used and the coding rate of the RRNS codes seen in Table 1. For example, for 4QAM we have 2 BPS, but the associated channel code rate is $205/320$, as shown in Table 1, hence the effective throughput of the system is $2 \times \frac{205}{320} = 1.28$. For AQAM/JD-CDMA, we can see from Figure 3 that the throughput is similar to that of BPSK/JD-CDMA at low channel SNRs. However, as the average channel SNR increased, more and more frames were transmitted using 4QAM/JD-CDMA and the average throughput increased gradually. At high average SNRs, the throughput of AQAM/JD-CDMA became similar to that of the 4QAM assisted JD-CDMA scheme.

The overall SEGSNR versus channel SNR performance of the proposed speech transceiver is displayed in Figure 4. Observe that the source sensitivity-matched triple-class RRNS-coded 4.75 kbps BPSK/JD-CDMA system requires a channel SNR in excess of about 8 dB for nearly unimpaired speech quality over the COST207 BU channel of Table 2. When the channel SNR was in excess of about 12 dB, the 10.2 kbps 4QAM/JD-CDMA system outperformed the 4.75 kbps BPSK/JD-CDMA scheme in terms of both objective and subjective speech quality. Furthermore, at channel SNR around 10 dB, where the BPSK and 4QAM SEGSNR curves cross each other in Figure 4 it was preferable to use the inherently lower quality but unimpaired mode of operation. In the light of these findings, the application of the AMR speech codec in conjunction with AQAM constitutes an attractive trade-off in terms of providing users with the best possible speech quality under arbitrary channel conditions. Specifically, the 10.2kbps 4QAM/JD-CDMA scheme has the highest source bit rate and thus exhibits the highest SEGSNR under error-free conditions. The 4.75kbps BPSK/JD-CDMA scheme exhibits a lower source bit rate and correspondingly lower speech quality under error-free conditions. However, due to its less robust modulation mode, the 10.2kbps 4QAM/JD-CDMA scheme is sensitive to channel errors and breaks down under hostile channel conditions, where the 4.75kbps BPSK/JD-CDMA scheme still exhibits robust operation, as illustrated in Figure 4.

The SEGSNR performance of the AQAM system is also displayed in Figure 4. We observe that AQAM provides a smooth evolution across the range of channel SNRs. At high channel SNRs in excess of 16dB, the system operates predominantly in the 4QAM/JD-CDMA mode. As the channel SNR degrades below 16 dB, some of the speech frames are transmitted in the BPSK/JD-CDMA mode, which implies that the lower quality speech rate of 4.75 kbps is employed. This results in a slightly degraded average speech quality, while still offering a substantial SEGSNR gain compared to the 4.75kbps BPSK/JD-CDMA scheme. At channel SNR below 10dB, the performance of the 10.2kbps 4QAM/ JD-CDMA mode deteriorates due to the occurence of a high number of errors, inflicting severe SEGSNR degradations. In these hostile conditions, the 4.75kbps BPSK/ JD-CDMA
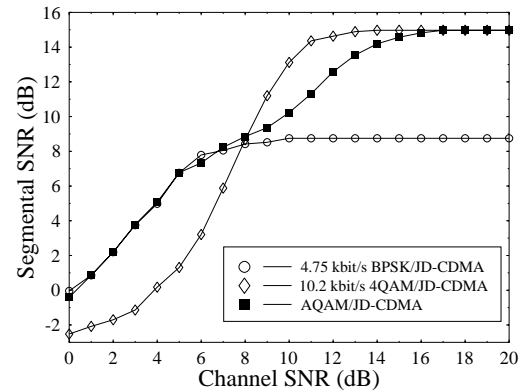

Figure 4: Segmental SNR versus Channel SNR

scheme provides a more robust performance associated with a better speech quality.

The benefits of the proposed dual-mode transceiver are further demonstrated by Figure 5, consisting of three graphs plotted against the speech frame index, giving an insightful characterisation of the adaptive speech transceiver.

When transmitting in the less robust 4QAM mode using the higher-rate speech mode of 10.2 kbps, a sudden steep drop in the channel conditions - as portrayed at Frame 1 in Figure 5 - results in a high number of transmission errors. This happens to occur during the period of voice onset in Figure 5, resulting in the corruption of the speech frame, which has the effect of inflicting impairments to subsequent frames due to the error propagation effects of various speech bits, as alluded to in Section 3. It can be seen in Figure 5 that the high number of errors inflicted in the 4QAM mode during voiced speech segments caused a severe SEGSNR degradation at frame index 10 and the 10.2 kbps speech codec never fully recovered, until the channel conditions expressed in terms of the SINR in Figure 5(c) improved. On the other hand, the significantly more robust 4.75kbps BPSK/JD-CDMA scheme performed well under these hostile channel conditions, encountering a low number of errors, while transmitting at a lower speech rate, hence at an inherently lower speech quality. For the sake of visual clarity, the performance curves of BPSK/JD-CDMA and AQAM/JD-CDMA were not displayed in Figure 5(b) for erroneous scenario because their respective graphs are almost identical to that of the error-free case.

Informal listening tests were conducted, in order to assess the performance of the AQAM/JD-CDMA scheme in comparison to the fixed-mode BPSK and 4QAM assisted JD-CDMA schemes. Through the listening tests we found that for the fixed-mode BPSK scheme unimpaired perceptual speech quality was achieved for channel SNRs in excess of 7dB. With reference to Figure 4, when the channel conditions degraded below 7dB, the speech quality be-
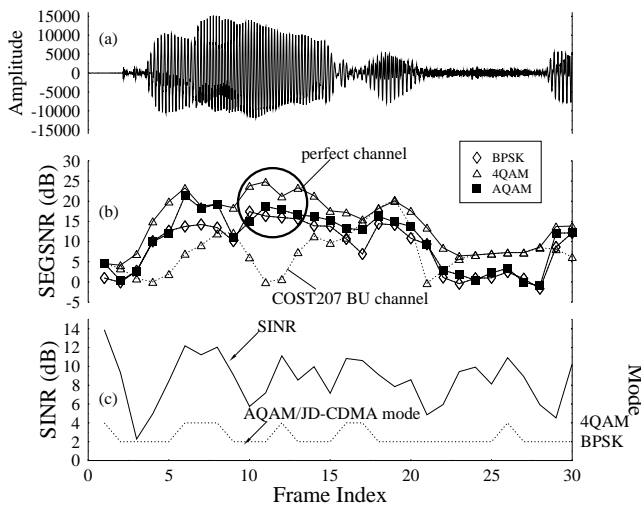
4

Figure 5: Characteristic waveforms of the adaptive system. (a) Time-domain speech signal; (b) SEGSNR in various transceiver modes; (c) SNR versus time and transceiver modes versus time.

came objectionable due to the preponderence of channel errors. For the fixed mode 4QAM/JD-CDMA scheme, the channel SNR threshold was 11dB, below which the speech quality started to degrade. The perceptual performance of AQAM/JD-CDMA was found superior to that of 4QAM/JD-CDMA at channel SNRs below 11dB.

7. CONCLUSIONS

In this contribution a joint-detection aided adaptive CDMA speech transceiver has been designed that allows us to switch between a set of different source and channel coders as well as transmission parameters, depending on the overall instantaneous channel quality. The benefits of the multi-mode speech transceiver clearly manifest themselves in terms of supporting unimpaired speech quality under hostile time-variant channel conditions, where a fixed-mode transceiver's quality would become severely degraded by channel effects. The proposed AQAM/JD-CDMA scheme acheived the best compromise between unimpaired error-free speech quality and channel robustness, which has been verified by our informal listening tests.

Our future research will be focussed on improving the performance of Burst-by-Burst AQAM/CDMA transceivers using wideband speech codecs operated at multiple modes. Furthermore, more robust, turbo space-time coded multi-carrier, frequency-hopped Burst-by-burst AQAM / CDMA transceivers will be invoked.

8. REFERENCES

[1] S. Nanda, K. Balachandran, and S. Kumar, "Adaptation techniques in wireless packet data services," *IEEE Commun. Mag.*, pp. 54–64, Jan. 2000.

[2] L. Hanzo, C.H. Wong, M.S. Yee: Adaptive wireless transceivers: A design paradigm for software-controlled TDMA, CDMA and OFDM systems, John Wiley, to appear

[3] L. Hanzo, F.C.A. Somerville, J.P. Woodard: Voice Compression and Communications: Principles and Applications for Fixed and Wireless Channels; IEEE Press-John Wiley, [1]

[4] L. Hanzo, P. Cherriman, J. Streit: Wireless Video Communications: Second to Third Generation and Beyond, IEEE Press, 2001 [2]

[5] GSM 06.90: Digital cellular telecommunications system (Phase 2+); Adaptive Multi-Rate (AMR) speech transcoding, version 7.0.0, Release 1998

[6] S. Bruhn, E. Ekudden and K. Hellwig, "Adaptive Multi-Rate: A new speech service for GSM and beyond", *Proceedings of 3rd ITG Conference on Source and Channel Coding*, Technical Univ. Munich, Germany, pp. 319-324, 17th-19th, January 2000.

[7] T.H. Liew, L-L. Yang, L. Hanzo, " Soft-decision Redundant Residue Number System Based Error Correction Coding," VTC'99, Amsterdam, The Netherlands, pp. 2546-2550, 19-22 Sept. 1999.

[8] E.L. Kuan, C.H. Wong and L. Hanzo, "Burst-by-burst adaptive joint-detection CDMA", *Proc. of IEEE VTC'99* , pp. 1628-1632, Houston, USA, May 1999

[9] R.A. Salami, C. Laflamme, J.P. Adoul, D. Massaloux, "A Toll Quality 8 kbps Speech Codec for the Personal Communications System(PCS)," *IEEE Transactions on Vehicular Technology*, vol. 43, pp. 808–816, Aug 1994.

[10] W.B. Kleijn and K.K. Paliwal, *Speech Coding and Synthesis* , Elsevier,1995

[11] L. Hanzo and J.P. Woodard, "An Intelligent Multimode Voice Communications System For Indoors Communications," *IEEE Transactions on Vehicular Technology*, vol. 44, pp. 735–749, Nov 1995.

[12] C.H. Wong, L. Hanzo: Upper-bound performance of a wideband burst-by-burst adaptive modem, to appear in IEEE Transactions on Communications, March, 2000

[13] M. Failli, "Digital land mobile radio communications COST 207", *Technical Report*, European Commission, Luxembourg, 1989.