# A Space-Time Coded OFDM Based MPEG-4 Audio Transceiver

## H.T. How, T.H. Liew 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*— **An MPEG-4 based Space-Time (ST) coded Orthogonal Frequency Division Multiplexing (OFDM) audio transceiver is proposed. The high-quality MPEG-4 audio codec is operated at bit rates between 16 and 64 kbit/s per channel and combined with turbo channel codes as well as space time codes. As expected, the space-time coding scheme, using two transmitters and one receiver outperformed the conventional one transmitter, one receiver arrangement by about 4dB in channel SNR terms, when maintaining an error-free audio quality.**

## I. Introduction

The most severe limitations of wireless communications are imposed by the hostile, time-variant nature of the wireless channel. Hence a range of channel-quality controlled adaptive features have been incorporated in a number of major cellular standards worldwide [1], including IS-95 CDMA, cdma2000 and UMTS W-CDMA, IS-136 TDMA, the General Packet Radio Service of GSM and in the Enhanced Data rates for Global Evolution (EDGE) schemes. Further near-instantaneously adaptive systems have been proposed in [2]-[4].

In this contribution the underlying trade-offs of using the multi-rate MPEG-4 Transform-domain Weighted Interleaved Vector Quantization (TWINVQ) audio codec [5], in conjunction with turbo-coded and space-time coded, adaptive Binary Phase Shift Keying (BPSK), Quadrature Phase Shift Keying (QPSK), 16-level Quadrature Amplitude Modulation (16QAM) assisted Orthogonal Frequency Division Multiplex (OFDM) systems are investigated.

The paper is structured as follows. Section II provides a brief system overview, while the associated system parameters are tabulated in Tables I and II. Section III describes the space-time coding scheme applied in our system. Finally, before concluding, our system performance results are summarised in Section V.

## II. System Overview

Figure 1 shows the schematic of the turbo-coded and space-time coded OFDM system [2]. The source bits generated by the MPEG-4 TWINVQ encoder [5] are passed to the turbo encoder using the half-rate, constraint length three turbo convolutional encoder TC(2,1,3), employing an octal generator polynomial of (7,5). The encoded bits were channel interleaved and passed to the modulator. In order

to counteract the near instantaneous channel quality fluctuations, burst-by-burst adaptive Quadrature Amplitude Modulation (QAM) was combined with space-time coding [2]. Specifically, the choice of the modulation scheme to be used by the transmitter for its next OFDM symbol is determined by the channel quality estimate of the receiver based on the current OFDM symbol. Here, perfect channel quality estimation and perfect signalling of the required modem modes were assumed. In order to simplify the task of signalling the required modulation modes, we employed the subband-adaptive OFDM transmission scheme of [14]. More specifically, the total OFDM symbol bandwidth was divided into equi-width subbands, hosting a given number of subcarriers and having a similar channel quality, where the same modem mode was assigned. The modulated signals were then passed to the encoder of the space-time block code $G_2$ [6], [2], which employs two transmitters and one receiver. The space-time encoded signals were OFDM modulated and transmitted by the corresponding antennas.

The received signals were OFDM demodulated and passed to the space-time decoders. Logarithmic Maximum A Posteriori (Log-MAP) decoding [7], [2] of the received space-time signals was performed, in order to provide soft-outputs for the TC(2,1,3) turbo decoder. The received bits were then channel deinterleaved and passed to the TC decoder, which again, employs the Log-MAP decoding algorithm. The decoded bits were finally passed to the TWINVQ decoder for obtaining the reconstructed audio signal.

Table I and II gives an overview of the proposed system's parameters. The transmission parameters have been partially harmonised with those of the TDD-mode of the Pan-European UMTS system [8]. The sampling rate is assumed to be 1.3 MHz, leading to a 1024 subcarrier OFDM symbol. The channel model used was the four-path COST 207 Typical Urban (TU) channel impulse response (CIR) [9], where each impulse was subjected to independent Rayleigh fading having a normalised Doppler frequency of $2.25 \cdot 10^{-6}$, corresponding to a pedestrian scenario at a walking speed of 3mph.

The channel encoder is a convolutional constituent coding based turbo encoder [10], employing block turbo interleavers and a pseudo-random channel interleaver. Again, the constituent Recursive Systematic Convolutional (RSC) encoder employs a constraint length of 3 and the octal gen-
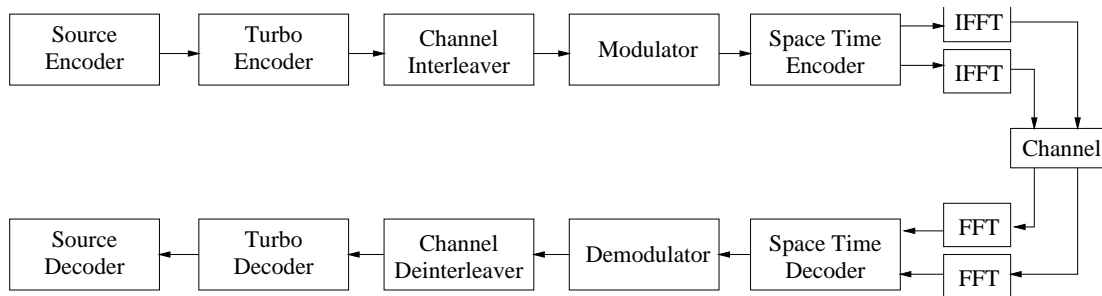
Fig. 1. Schematic overview of the turbo-coded and space-time coded OFDM system.

| Channel Coded Bits/Packet | 928 | 1856 | 3712 |
|---|---|---|---|
| Channel Coded Bit Rate (kbit/s) | 39.9 | 79.9 | 159.9 |
| Source Coded Bits/Packet | 372 | 743 | 1486 |
| Source Coded Bit Rate (kbit/s) | 16 | 32 | 64 |
| Modulation Mode | BPSK | QPSK | 16QAM |
| Minimum Channel SNR for 1% FER (dB) | 4.3 | 7.2 | 12.4 |
| Minimum Channel SNR for 5% FER (dB) | 2.7 | 5.8 | 10.6 |

TABLE I

GENERIC SYSTEM PARAMETERS

erator polynomial of (7,5) [2]. Eight iterations are performed at the decoder, utilising the MAP-algorithm and the Log-Likelihood Ratio (LLR) soft inputs provided by the demodulator [2].

The MPEG-4 audio scheme is the most recent compression standard completed in 1999 [5], which supports the encoding of speech signals at bit rates from 2 kbit/s to 24 kbit/s. It is also capable of encoding general audio at rates ranging from very low bit rates to high rates. Specifically, audio coding is supported at a rate of 8 kbit/s and a bandwidth of 4 kHz, but the standard also includes broadcast quality audio monaural as well as multichannel stereophonic configurations [5]. The TWINVQ scheme [5], [11] is one of the encoding tools designed for coding audio signals in MPEG-4, which was standardised for encoding audio signals at extremely low bit rates. Hence the MPEG-4 TWINVQ [11], [12] codec has been employed in the proposed system, which can be programmed to operate at the bit rates between 16 and 64 kbit/s, as shown in Table I.

## III. SPACE-TIME CODING

Traditionally, the most effective technique of combating fading has been the exploitation of diversity [13]. Diversity techniques can be divided into three broad categories, namely temporal diversity, frequency diversity and spatial diversity. Temporal and frequency diversity introduce redundancy in the time and/or frequency domain, which results in a loss of bandwidth efficiency. Examples of spatial diversity are constituted by multiple transmit- and/or receive-antenna based systems. Transmit-antenna diversity relies on employing multiple antennas at the transmitter and hence it is suitable for downlink transmissions, since having multiple transmit antennas at the base station is certainly feasible. By contrast, receive-antenna diver-

| System Parameters | Value |
|---|---|
| Carrier Frequency | 1.9 GHz |
| Sampling Rate | 3.78 MHz |
| **Channel** | |
| Impulse Response | COST207 |
| Normalised Doppler Frequency | $2.25 \cdot 10^{-6}$ |
| **OFDM** | |
| Number of Subcarriers | 1024 |
| OFDM Symbols/Packet | 1 |
| OFDM Symbol Duration | (1024+64) x $1/(3.78 \cdot 10^6)$ |
| Guard Period | 64 samples |
| Modulation Scheme | BPSK/QPSK/16QAM |
| **Space Time Coding** | |
| Number of transmitters | 2 |
| Number of receivers | 1 |
| **Channel Coding** | Turbo Convolutional |
| Constraint Length | 3 |
| Code Rate | 0.5 |
| Generator Polynomials | 7, 5 |
| Turbo Interleaver Length | 464/928/1856/2784 |
| Decoding Algorithm | Log MAP |
| Number of Iterations | 8 |
| **Source Coding** | MPEG-4 TWINVQ |
| Bit Rates (kbit/s) | 16 - 64 |
| Audio Frame Length (ms) | 23.22 |
| Sampling Rate (kHz) | 44.1 |

TABLE II

SUBSYSTEM PARAMETERS

2
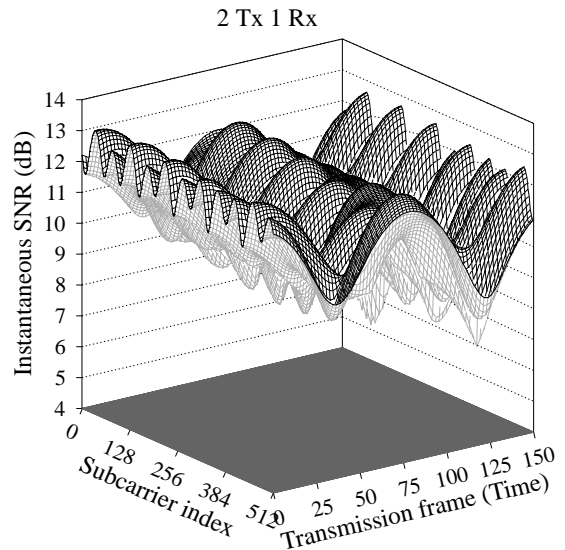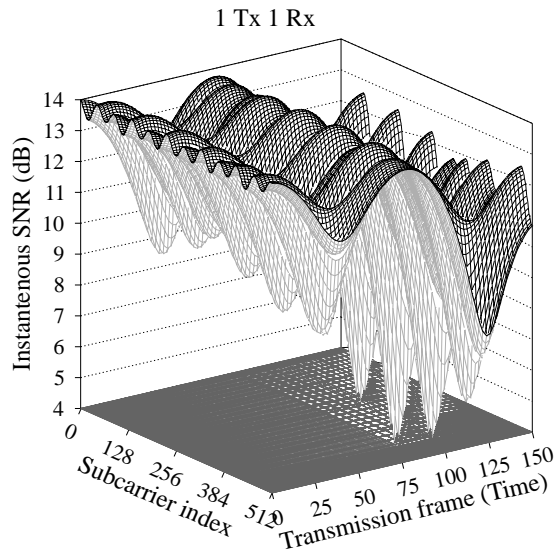
Fig. 2.  Instantaneous channel SNR of 512-subcarrier OFDM symbols for the one-transmitter, one-receiver (1Tx 1Rx) and for the space-time block coded two-transmitter, one-receiver (2Tx 1Rx) scenarios for transmission over the COST207 TU channel at an average SNR of 10 dB.

sity employs multiple antennas at the receiver for acquiring multiple copies of the transmitted signals, which are then combined in order to mitigate the channel-induced fading.

Space time coding [13], [6], [2] is a specific form of transmit antenna diversity, which aims for usefully exploiting the multipath phenomenon experienced by signals propagating through the dispersive mobile channel. This is achieved by combining multiple transmission antennas with appropriate signal processing at the receiver, in order to provide diversity and coding gain over uncoded single-antenna scenarios.

In this system, we employ a two-transmitter and one-receiver configuration, in conjunction with turbo coding [10]. In Figure 2, we show the instantaneous channel SNR experienced by the 512-subcarrier OFDM modem for a one-transmitter, one-receiver scheme and for the space time block code $G_2$ [6] using two transmitters and one receiver for transmission over the COST207 Typical Urban (TU) channel. The average channel SNR was 10 dB. We can see in Figure 2 that the variation of the instantaneous channel SNR for a one-transmitter, one-receiver scheme is severe. The instantaneous channel SNR may become as low as 4 dB due to the deep fades inflicted by the channel. On the other hand, we can see that for the space-time block code $G_2$ using one receiver the variation of the instantaneous channel SNR is less severe. Explicitly, by employing multiple transmit antennas in Figure 2, we have significantly reduced the effect of the channels' deep fades. Whilst space-time coding endeavours to mitigate the fading-related time- and frequency-domain channel-quality fluctuations at the cost of increasing the transmitter's complexity, adaptive modulation attempts to accommodate these channel quality fluctuations by appropriately controlling the AQAM modes, as it will be outlined in the next section.

## IV. ADAPTIVE MODULATION

In order to accommodate the time- and frequency-domain channel quality variations, a multi-mode system is desirable, which allows us to switch between a set of different source- and channel coders as well as transmission parameters, depending on the instantaneous channel quality [14], [2], [3], [4].

In the proposed system, we have defined three operating modes, which correspond to the uncoded audio bit rates of 16, 32 and 64 kbit/s. This corresponds to 372, 743 and 1486 bits per 23.22 ms audio frame. In conjunction with half-rate channel coding and also allowing for check sums and signalling overheads, the number of transmitted turbo coded bits per OFDM symbol is 928, 1856 and 3712 for the three different source-coded modes, respectively. Again, these bit rates are also summarised in Table I. Each transmission mode uses a different modulation scheme, depending on the instantaneous channel conditions. It is beneficial, if the transceiver can drop its source rate, for example from 64 kbit/s to 32 kbit/s and invoke QPSK modulation instead of 16QAM, while maintaining the same bandwidth. Hence, during good channel conditions, the higher throughput, higher audio quality but less robust mode of operations can be invoked, while the more robust but lower audio quality BPSK/16kbit/s mode can be applied during degrading channel conditions.

Figure 3 shows the FER observed for all three audio coding modes versus the channel BER that was predicted by the receiver during the channel quality estimation process. The rationale behind using the FER, rather than the BER for estimating the expected channel quality of the next transmitted OFDM symbol is that the MPEG-4 audio codec has to drop the specific turbo-decoded received OFDM symbols, which contained transmission errors. This is because corrupted audio packets would result in detrimental MPEG-4 decoding error propagation. A FER of 1%
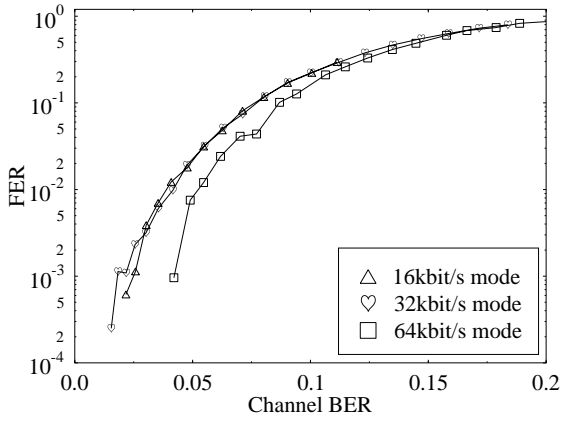
3

Fig. 3. FER against channel BER of the OFDM/$G_2$/TC(2,1,3) system for transmission over the COST207 TU channel.
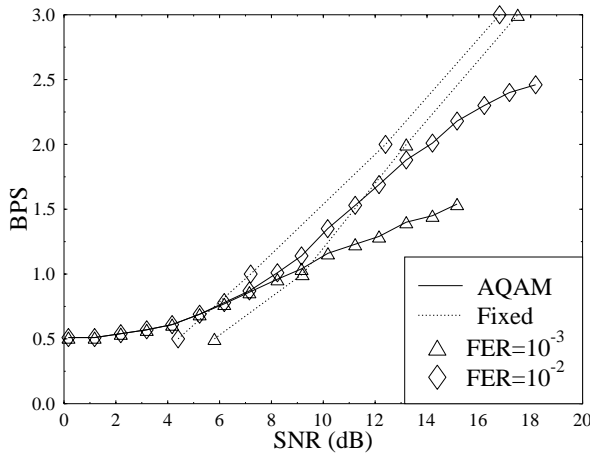


Fig. 4. BPS performance comparison between the adaptive and fixed modulation schemes.

was observed for an estimated channel error rate of about 4% for the 16 and 32 kbit/s modes, while a BER of over 5% was tolerable for the 64 kbit/s mode. This was, because the number of bits per OFDM symbol was quadrupled in the 16QAM mode compared to the BPSK mode, which substantially increased the turbo codec's performance due to benefiting from a longer interleaver without increasing the audio system's delay.

In Figure 4, we show the Bits Per Symbol (BPS) throughput performance comparison between the subband-adaptive and fixed OFDM modulation schemes. From the figure we can see that at a low BPS throughput the adaptive modulation scheme outperforms the fixed modulation scheme. However, as the BPS throughput of the system increases, the fixed modulation schemes are preferred. This is, because adaptive modulation is advantageous, when there are high channel quality variations, as in the one-transmitter, one receiver scheme. However, we

have shown in Figure 2 that the channel quality variation has been significantly reduced by employing two $G_2$ space-time transmitters. Therefore, the advantages of adaptive modulation eroded due to the reduced channel quality variations experienced in the context of the space-time coded system. Two different-complexity system design principles can be proposed as a consequence. The first system is the one-transmitter, one receiver scheme (1T-1R), which mitigates the severe variation of the channel quality by employing adaptive modulation. By contrast, we can design a more complex (2T-2R) space-time system, which employs fixed modulation schemes, since no substantial benefits accrue from employing adaptive modulation, once the fading-induced channel-quality fluctuations have been sufficiently mitigated by the $G_2$ space-time code. In the rest of this contribution, we have opted for investigating the performance of the space-time coded system, requiring an increased complexity in conjunction with different modulation modes, namely BPSK, QPSK and 16QAM.

## V. System Performance

As mentioned before, the detailed subsystem parameters used in our space-time coded OFDM system are listed in Table II. Again, the channel profile used was the COST 207 TU channel [9] having four paths and a maximum dispersion of 4.5 $\mu$s, where each path was faded independently at a Doppler frequency of $2.25 \cdot 10^{-6}$ Hz.

We found that the associated audio quality expressed in terms of the Segmental SNR (SEGSNR) degradation was deemed to be perceptually objectionable for frame error rates in excess of 5%. For the sake of maintaining a high audio quality, a maximum target FER of 1% was stipulated based on our informal listening tests. Again, corrupted frames were dropped and the previous uncorrupted frame was used for replacing the dropped frame.

The BER is plotted versus the channel SNR in Figure 5 for the three different modes of operation. The results using the conventional one-transmitter, one-receiver scenario were also shown in conjunction with the 1, 2 and 4 BPS fixed modem modes as a benchmarker. The employment of space time coding improved the system's performance significantly, giving an approximately 3 dB channel SNR improvement at a FER of 1%. As expected, the lowest throughput BPSK/16kbit/s mode was more robust, than the QPSK/32kbit/s and the 16QAM/64kbit/s configurations, albeit delivering a lower audio quality.

The overall SEGSNR versus channel SNR performance of the proposed audio transceiver is displayed in Figure 6, again, employing $G_2$ space-time coding using two transmitters and one receiver. The benchmarker using the conventional one-transmitter, one-receiver scheme was also characterised in the figure. We observe that the employment of space time coding provides a substantial improvement in terms of maintaining an error free audio performance. Specifically, an SNR advantage of 4 dB was recorded compared to the conventional one-transmitter, one-receiver benchmarker for all three modulation modes. Furthermore, focusing on the three different operating modes using
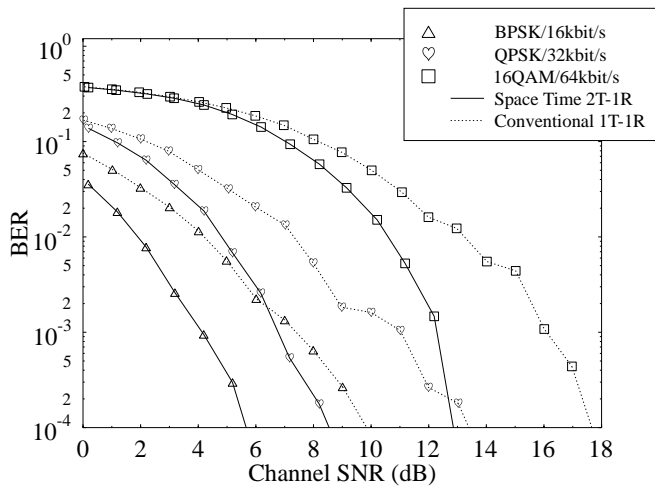
4

Fig. 5. BER versus Channel SNR performance of the $G_2$/OFDM/TC(2,1,3) scheme (2T-2R) for transmission over the COST207 TU channel, in comparison to the conventional one-transmitter, one-receiver (1T-1R) benchmarker.



Fig. 6. SEGSNR versus channel SNR performance of the $G_2$/OFDM/TC(2,1,3) scheme (2T-2R) for transmission over the COST207 TU channel, in comparison to the conventional one-transmitter, one-receiver (1T-1R) benchmarker.
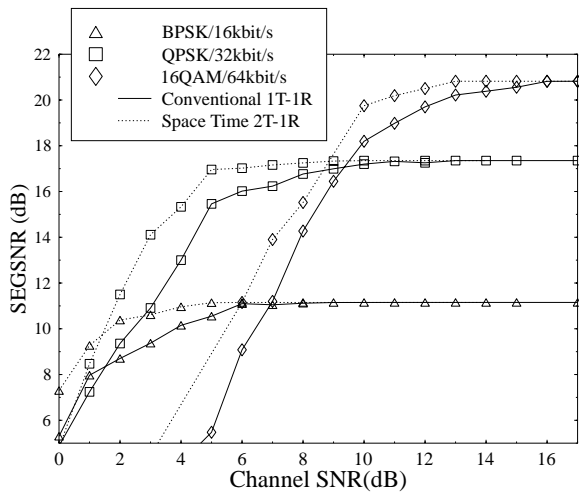
space-time coding, namely on the curves drawn in dotted lines, the 16QAM/64kbit/s mode was shown to outperform the QPSK/32kbit/s scheme in terms of both objective and subjective audio quality for channel SNRs in excess of about 10 dB. At a channel SNR of about 9 dB, where the 16QAM and QPSK SEGSNR curves cross each other in Figure 6, it is preferable to invoke the inherently lower audio quality, but unimpaired QPSK mode of operation. Similarly, at around 5 dB when the QPSK/32kbit/s scheme's performance starts to degrade, it is better to invoke the unimpaired BPSK/ 16kbit/s mode of operation, in order to avoid the channel-induced audio artifacts.

## VI. CONCLUSIONS

In conclusion, a turbo coded and space-time coded adaptive versus fixed modulation based OFDM assisted MPEG-4 audio system has been investigated. The space-time coded, two-transmitter, one-receiver configuration outperformed the conventional one-transmitter, one-receiver scheme by about 4dB in channel SNR terms, when communicating over the highly dispersive COST207 TU channel. In our future work, unequal error protection assisted schemes will be benchmarked against the proposed packet dropping regime of this contribution.

## 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, T.H. Liew, B.L. Yeap: Turbo Coding, Turbo Equalisation and Space-Time Coding for Transmission over Fading Channels, John Wiley, to appear in 2001

[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] ISO/IEC JTC1/SC29/WG11/N2203, MPEG-4 Audio Version 1 Final Committee Draft 14496-3, `http://www.tnt.uni-hannover.de/project/mpeg/audio/documents/`, March 1998.

[6] S. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 16, pp. 1451–1458, October 1998.

[7] L. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," *IEEE Transactions on Information Theory*, vol. 20, pp. 284–287, March 1974.

[8] H. Holma and A. Toskala, *WCDMA for UMTS*. John Wiley-IEEE Press, April 2000.

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

[10] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near Shannon limit error-correcting coding and decoding: Turbo codes," in *Proceedings of ICC*, pp. 1064–1070, 1993.

[11] N. Iwakami, T. Moriya, and S. Miki, "High-quality audio coding at less than 64 kbit/s by using transform-domain weighted interleave vector quantization (TwinVQ)," in *Proceedings of ICASSP*, pp. 3095–3098, May 1995.

[12] T. Moriya, N. Iwakami, A. Jin, K. Ikeda, and S. Miki, "A design of transform coder for both speech and audio signals at 1 bit/sample," in *Proceedings of ICASSP*, pp. 1371–1374, Apr 1997.

[13] V. Tarokh, N. Seshadri, and A. Calderbank, "Space-time codes for high data rate wireless communication: Performance criterion and code construction," *IEEE Transactions on Information Theory*, vol. 44, pp. 744–765, March 1998.

[14] L. Hanzo, W. T. Webb and T. Keller, *Single and Multicarrier Quadrature Amplitude Modulation : Principles and Applications for Personal Communications, WLANs and Broadcasting (2nd Ed.)*. John Wiley-IEEE Press, April 2000.

[15] 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.

[16] T. Keller, M. Muenster, and L. Hanzo, "A Turbo-Coded Burst-By-Burst Adaptive Wideband Speech Transceiver," *IEEE Journal on Selected Areas in Communications*, vol. 18, pp. 2363–2372, Nov 2000.

[1]For detailed contents please refer to http://www-mobile.ecs.soton.ac.uk

[2]For detailed contents please refer to http://www-mobile.ecs.soton.ac.uk

5