University of Southampton
Faculty of Engineering and Applied Sciences
School of Electronics and Computer Science

Wireless Speech and Audio Communications

by

Noor Shamsiah Othman
B.Eng., M.Sc.

A doctoral thesis submitted in partial fulfilment of the requirements for the award of Doctor of Philosophy at the University of Southampton

June 2008

SUPERVISORS:
Professor Lajos Hanzo
M.Sc., Ph.D, FREng, DSc, FIEEE, FIET
Chair of Telecommunications

Dr Soon Xin Ng (Michael)
BEng, PhD, MIEE, MIEEE

School of Electronics and Computer Science
University of Southampton
Southampton SO17 1BJ

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF ENGINEERING AND APPLIED SCIENCE
SCHOOL OF ELECTRONICS AND COMPUTER SCIENCE

Doctor of Philosophy

Wireless Speech and Audio Communications

by
Noor Shamsiah Othman

The limited applicability of Shannon's separation theorem in practical speech/audio systems motivates the employment of joint source and channel coding techniques. Thus, considerable efforts have been invested in designing various of joint source and channel coding schemes. This thesis discusses two different types of Joint Source and Channel Coding (JSCC) schemes, namely Unequal Error Protection (UEP) aided turbo transceivers as well as Iterative Source and Channel Decoding (ISCD) exploiting the residual redundancy inherent in the source encoded parameters.

More specifically, in Chapter 2, two different UEP JSCC philosophies were designed for wireless audio and speech transmissions, namely a turbo-detected UEP scheme using twin-class convolutional codes and another turbo detector using more sophisticated Irregular Convolutional Codes (IRCC). In our investigations, the MPEG-4 Advanced Audio Coding (AAC), the MPEG-4 Transform-Domain Weighted Interleaved Vector Quantization (TwinVQ) and the Adaptive MultiRate WideBand (AMR-WB) audio/speech codecs were incorporated in the sophisticated UEP turbo transceiver, which consisted of a three-stage serially concatenated scheme constituted by Space-Time Trellis Coding (STTC), Trellis Coded Modulation (TCM) and two different-rate Non-Systematic Convolutional codes (NSCs) used for UEP. Explicitly, both the twin-class UEP turbo transceiver assisted MPEG-4 TwinVQ and the AMR-WB audio/speech schemes outperformed their corresponding single-class audio/speech benchmarkers by approximately 0.5 dB, in terms of the required E_b/N_0, when communicating over uncorrelated Rayleigh fading channels. By contrast, when employing the MPEG-4 AAC audio codec and protecting the class-1 audio bits using a 2/3-rate NSC code, a more substantial E_b/N_0 gain of about 2 dB was achieved. As a further design alternative, we also proposed a turbo transceiver employing IRCCs for the sake of providing UEP for the AMR-WB speech codec. The resultant UEP schemes exhibited a better performance when compared to the corresponding Equal Error

Protection (EEP) benchmark schemes, since the former protected the audio/speech bits according to their sensitivity. The proposed UEP aided system using IRCCs exhibits an E_b/N_0 gain of about 0.5 dB over the EEP system employing regular convolutional codes, when communicating over AWGN channels.

In Chapter 3, we proposed and investigated a novel system that invokes jointly optimised ISCD for enhancing the error resilience of the AMR-WB speech codec. The resultant AMR-WB coded speech signal is protected by a Recursive Systematic Convolutional (RSC) code and transmitted using a non-coherently detected Multiple-Input Multiple-Output (MIMO) Differential Space-Time Spreading (DSTS) scheme. To further enhance the attainable system performance and to maximise the coding advantage of the proposed transmission scheme, the system is also combined with multi-dimensional Sphere Packing (SP) modulation. The AMR-WB speech decoder was further developed for the sake of accepting the *a priori* information passed to it from the channel decoder as extrinsic information, where the residual redundancy inherent in the AMR-WB encoded parameters was exploited.

Moreover, the convergence behaviour of the proposed scheme was evaluated with the aid of both Three-Dimensional (3D) and Two-Dimenstional (2D) EXtrinsic Information Transfer (EXIT) charts. The proposed scheme benefitted from the exploitation of the residual redundancy inherent in the AMR-WB encoded parameters, where an approximately 0.5 dB E_b/N_0 gain was ahieved in comparison to its corresponding hard speech decoding based counterpart. At the point of tolerating a SegSNR degradation of 1 dB, the advocated scheme exhibited an E_b/N_0 gain of about 1.0 dB in comparison to the benchmark scheme carrying out joint channel decoding and DSTS aided SP-demodulation in conjunction with separate AMR-WB decoding, when communicating over correlated narrowband Rayleigh fading channels.

In Chapter 4, we proposed two jointly optimized ISCD schemes invoking the optimized soft-output AMR-WB speech codec using DSTS assisted SP modulation. More specifically, the soft-bit assisted iterative AMR-WB decoder's convergence characteristics were further enhanced by using Over-Complete source-Mapping (OCM), as well as a recursive precoder. EXIT charts were used to analyse the convergence behaviour of the proposed turbo transceivers using the soft-bit assisted AMR-WB decoder.

Explicitly, the OCM aided AMR-WB MIMO transceiver exhibits an E_b/N_0 gain of about 3.0 dB in comparison to the benchmark scheme also using ISCD as well as DSTS aided SP-demodulation, but dispensing with the OCM scheme, when communicating over correlated narrowband Rayleigh fading channels. Finally, the precoded soft-bit AMR-WB

MIMO transceiver exhibits an E_b/N_0 gain of about 1.5 dB in comparison to the benchmark scheme dispensing with the precoder, when communicating over correlated narrowband Rayleigh fading channels .

Acknowledgements

I am very grateful to my supervisors Prof. Lajos Hanzo and Dr. Soon Xin Ng (Michael) for their generous guidance, kindness and patience. This research would not have progressed this far without them. All the help, advice, ideas and encouragement are greatly appreciated.

I would also like to thank all my colleagues in the Communications Group of University of Southampton, both past and present, for all their supports and helps throughout this project. Special thanks to Dr. Soon Xin Ng, Dr. Jin Wang, Mohammed El-Hajjar, Anh Pham Quang, Dr. Osamah Alamri, Dr. Jörg Kliewer and Nasruminallah for their technical support and collaborative work. The assistance provided by Denise Harvey and Rebecca Earl, the Communication Group secretaries, is gratefully acknowledged.

The financial support of University Tenaga Nasional Malaysia as well as that of the School of Electronics and Computer Science in the University of Southampton is gratefully acknowledged.

A special thanks also goes to all my friends in Southampton and elsewhere, for their constant kind wishes. There are two people that I would like to thank especially, Dr Rosmila Senik and Dr Azizah Mainal for their friendship and constant encouragement.

I would like to express my appreciation to my parents, my sisters and my brothers for their unconditional love and encouragement. My gratitude goes to my beloved husband Mohd Fadzil and my daughter Khadijah for their patience, sacrifices and encouragement, which motivates me to be steadfast and never bend to difficulties. Last but not least, my sincere gratitude to our newborn Muhammad. I owe my every achievement to all of them.

List of Publications

JOURNAL PAPERS:

1. **N. S. Othman**, M. El-Hajjar, O. Alamri and L. Hanzo "Iterative AMR-WB Source and Channel-Decoding Using Differential Space-Time Spreading Assisted Sphere Packing Modulation", to appear in IEEE Transactions on Vehicular Technology.

2. A. Q. Pham, L. -L. Yang, **N. S. Othman** and L. Hanzo "EXIT-Chart Optimized Block Codes for Wireless Video Telephony", to appear in IEEE Transactions on Circuits and Systems for Video Technology.

CONFERENCE PAPERS:

3. **N. S. Othman**, S. X. Ng and L. Hanzo, " Turbo-Detected Unequal Protection MPEG-4 Audio Transceiver Using Convolutional Codes, Trellis Coded Modulation and Space-Time Trellis Coding", IEEE 61st Vehicular Technology Conference, Stockholm, Sweden, 30 May-1 June 2005, pp. 1600 -1604.

4. **N. S. Othman**, S. X. Ng and L. Hanzo, "Turbo-Detected Unequal Protection Audio and Speech Transceivers Using Serially Concantenated Convolutional Codes, Trellis Coded Modulation and Space-Time Trellis Coding", IEEE 62nd Vehicular Technology Conference, Texas, USA, 25-28 September 2005, pp. 1044-1048.

5. J. Wang, **N. S. Othman**, J. Kliewer, L-L. Yang and L. Hanzo "Turbo Detected Unequal Error Protection General Configuration Irregular Convolutional Codes Designed for the Wideband Advanced Multirate Speech Codec", IEEE 62nd Vehicular Technology Conference, Texas, USA, 25-28 September 2005, pp. 1044-1048.

6. **N. S. Othman**, M. El-Hajjar, O. Alamri and L. Hanzo "Soft-Bit Assisted Iterative AMR-WB Source-Decoding and Turbo-Detection of Channel-Coded Differential Space-Time Spreading Using Sphere Packing Modulation", IEEE 65th Vehicular Technology Conference, Dublin, Ireland, 22-25 April 2007, pp. 2010-2014.

7. **N. S. Othman**, M. El-Hajjar, A. Q. Pham, O. Alamri, S. X. Ng and L. Hanzo "Over-Complete Source-Mapping Aided AMR-WB MIMO Transceiver Using Three-Stage Iterative Detection", IEEE International Conference on Communication, Beijing, China, 19-23 May 2008, pp. 751-755.

8. Nasruminallah, M. El-Hajjar, **N. S. Othman**, A. Q. Pham, O. Alamri and L. Hanzo "Over-Complete Mapping Aided, Soft-Bit Assisted Iterative Unequal Error

Protection H.264 Joint Source and Channel Decoding", to appear in Proceedings of IEEE 68th Vehicular Technology Conference, Calgary, Alberta, Canada, 21-24 September 2008.

9. **N. S. Othman**, M. El-Hajjar, O. Alamri, S. X. Ng and L. Hanzo "Three-Stage Iterative Detection of Precoded Soft-Bit AMR-WB for Speech MIMO Transceiver", to be submitted to IEEE 69th Vehicular Technology Conference, 2009.

BOOK CONTRIBUTIONS:

1. J. Wang, **N. S. Othman**, J. Kliewer, L-L. Yang and L. Hanzo "Section 9.6: Turbo-detected Unequal Error Protection Irregular Convolutional Coded AMR-WB Transceiver", L. Hanzo, C. Somerville and J. Woodard, "Voice and Audio Compression for Wireless Communications, 2nd Edition", Chichester, UK: John Wiley-Sons Inc., 2007, pp. 442-453.

2. **N. S. Othman**, S. X. Ng and L. Hanzo "Section 10.6: Turbo-detected Space-time Trellis Coded MPEG-4 Audio Transceivers", L. Hanzo, C. Somerville and J. Woodard, "Voice and Audio Compression for Wireless Communications, 2nd Edition", Chichester, UK: John Wiley-Sons Inc., 2007, pp. 516-524.

3. **N. S. Othman**, S. X. Ng and L. Hanzo "Section 10.7: Turbo-detected Space-time Trellis Coded MPEG-4 versus AMR-WB Speech Transceivers", L. Hanzo, C. Somerville and J. Woodard, "Voice and Audio Compression for Wireless Communications, 2nd Edition", Chichester, UK: John Wiley-Sons Inc., 2007, pp. 525-534.

Contents