Embedded lossless audio coding using linear prediction and cascade coding

Kevin Adistambha
University of Wollongong

UNIVERSITY OF WOLLONGONG
COPYRIGHT WARNING

You may print or download ONE copy of this document for the purpose of your own research or study. The University does not authorise you to copy, communicate or otherwise make available electronically to any other person any copyright material contained on this site. You are reminded of the following:

This work is copyright. Apart from any use permitted under the Copyright Act 1968, no part of this work may be reproduced by any process, nor may any other exclusive right be exercised, without the permission of the author.

Copyright owners are entitled to take legal action against persons who infringe their copyright. A reproduction of material that is protected by copyright may be a copyright infringement. A court may impose penalties and award damages in relation to offences and infringements relating to copyright material. Higher penalties may apply, and higher damages may be awarded, for offences and infringements involving the conversion of material into digital or electronic form.

Unless otherwise indicated, the views expressed in this thesis are those of the author and do not necessarily represent the views of the University of Wollongong.

Recommended Citation
NOTE

This online version of the thesis may have different page formatting and pagination from the paper copy held in the University of Wollongong Library.

UNIVERSITY OF WOLLONGONG

COPYRIGHT WARNING

You may print or download ONE copy of this document for the purpose of your own research or study. The University does not authorise you to copy, communicate or otherwise make available electronically to any other person any copyright material contained on this site. You are reminded of the following:

Copyright owners are entitled to take legal action against persons who infringe their copyright. A reproduction of material that is protected by copyright may be a copyright infringement. A court may impose penalties and award damages in relation to offences and infringements relating to copyright material. Higher penalties may apply, and higher damages may be awarded, for offences and infringements involving the conversion of material into digital or electronic form.
Embedded Lossless Audio Coding using Linear Prediction and Cascade coding

A thesis submitted in fulfillment of the requirements for the award of the degree

Master of Engineering Research

From

The University of Wollongong

By

Kevin Adistambha

Bachelor of Science, Master of Internet Technology

School of Electrical, Computer and Telecommunications Engineering

2005
Abstract

This thesis studies the techniques and feasibility of embedding a perceptual audio coder within a lossless compression scheme. The goal is to provide for two step scalability in the resulting bitstream, where both a perceptual version of the audio signal and a lossless version of the same signal are provided in the one bitstream.

The focus of this thesis is the selection of the perceptual coder to be used as the perceptual base layer and the techniques to be used to compress the lossless layer by using backward linear prediction followed by entropy coding. The perceptual base layer used is MPEG-4 AAC, chosen based on entropy measurements of the residual signal. Results of the work in this thesis show that the embedded lossless coding scheme could achieve an average compression ratio of only 6% larger compared to lossless only coding. Performing decorrelation on the AAC residual signal by means of backward linear predictive coding and measuring the entropy of the resulting LPC residual signal of various orders revealed that an 8% decrease in coding rate is achievable using 15th order prediction.

Furthermore, this thesis also investigates an entropy coding technique known as cascade coding which is originally designed to compress hydroacoustic image data and is modified to compress audio data. Cascade coding is an entropy coding technique that uses multiple cascaded stages where each stage codes a specific range of integers and is used to perform entropy coding of the backward linear prediction residual signal. The cascade coding technique explored in this thesis includes using a frame based approach and trained codebooks.
Statement of Originality

This is to certify that the work described in this thesis is entirely my own, except where due reference is made in the text.

No work in this thesis has been submitted for a degree to any other university or institution.

Signed

Kevin Adistambha

6 December 2005
Acknowledgements

I would like to thank my supervisors Dr. Christian Ritz, Dr. Jason Lukasiak and Assoc. Prof. Ian Burnett. Their support and patience made this work possible. My never ending questions and problems don’t seem to deter them to help me finish this thesis.

I would also like to thank my family and friends for their morale support and providing relief when the stress or boredom (or both) is too much for me to handle. Especially my parents, whose patience seems impossible to exhaust.

Last but not least, thanks to Ibanez guitars, Marshall amplification and Line 6 for providing me with much needed distractions.
Table of Contents

1 Introduction .................................................................................................................. 14

1.1 Digital representation of audio data ................................................................. 15

1.2 Perceptual coding ............................................................................................. 15

1.3 Lossless coding ............................................................................................... 16

1.3.1 Scalable lossless coding ............................................................................. 16

1.3.2 Approaches to scalable to lossless audio coding ...................................... 17

1.4 Motivation ......................................................................................................... 19

1.5 Thesis outline .................................................................................................. 20

1.6 Contributions of this Thesis ............................................................................ 20

2 Entropy and decorrelation of audio signals ....................................................... 22

2.1 Introduction ..................................................................................................... 22

2.2 Signal entropy and coding theory .................................................................... 22

2.2.1 Entropy theory ......................................................................................... 22

2.2.2 Entropy coding methods ........................................................................... 23

2.2.3 Entropy coding techniques comparison .................................................. 32

2.3 Scalable entropy coding techniques ............................................................... 33

2.3.1 Fine grain scalability via bit plane coding ............................................... 34
4 Decorrelation and entropy coding ................................................................. 56

4.1 Decorrelation of the AAC residual ............................................................... 56

4.1.1 Overview of decorrelation methods in lossless audio coding ............... 56

4.1.2 Decorrelation results ............................................................................. 61

4.1.3 Statistical characteristics of the AAC residual ..................................... 62

4.2 Entropy coding ......................................................................................... 63

4.2.1 Overview of entropy coding methods ............................................... 63

4.2.2 Entropy coding via Rice coding .......................................................... 63

4.2.3 Entropy coding via cascade coding ..................................................... 64

4.3 Cascade Codebook Design and Testing ..................................................... 70

4.3.1 Cascade coding codebook training algorithm ................................... 70

4.3.2 Training database .............................................................................. 71

4.3.3 Cascade coder codebook training results ......................................... 72

4.3.4 Test database coding result ................................................................. 75

4.4 Cascade Codebook Performance and Discussion ..................................... 76

4.5 Chapter summary .................................................................................... 79

5 Conclusion .................................................................................................. 81

5.1 Limitations of techniques ....................................................................... 82

5.2 Future work ............................................................................................. 82
List of Figures

Figure 1.1. Example of perceptual coding vs. objective coding SNR function. ........18
Figure 1.2. Block diagram of an embedded lossless coder. ...............................19
Figure 2.1. Example of Huffman coding ............................................................24
Figure 2.2. The n\textsuperscript{th} stage of a cascade coder ............... .................31
Figure 2.3. Example of a cascade coder. ............................................................32
Figure 2.4. General operation of the bit plane coding scheme. .........................33
Figure 3.1. Block diagram of the embedded lossless coder. ..........................42
Figure 3.2. General block diagram of a perceptual coder. ..............................43
Figure 3.3. Entropy and compression of the residual signal using gzip and Monkey's Audio. Line A represents residual compressed with gzip, line B represents residual entropy and line C represents the residual compressed with Monkey’s Audio. 48
Figure 3.4. Signal-to-noise ratio and entropy of the signal. ...............................49
Figure 3.5. Total bitrate (perceptual layer + lossless layer) comparison .........53
Figure 4.1. Block diagram of the embedded lossless coder. .........................56
Figure 4.2. The windowing function used in the backward LPC process ..........59
Figure 4.3. LPC gain plot for backward LPC of order 1-40 using Hamming window and Hybrid windows. .................................................................60
Figure 4.4. SFM measurements for the LPC residual signal of order 1-40 using Hamming window and Hybrid window.................................................................60

Figure 4.5. One frame of an AAC residual signal histogram before and after order 15 LPC processing, compared to Laplacian p.d.f. .................................................................62

Figure 4.6. 95% confidence interval plot for cascade coding, Rice coding and entropy values, using no LPC and LPC order 15 on the residual signal..........................67

Figure 4.7. Histogram plot of the absolute values of samples in the codebook 8 frames..................................................................................................................74

Figure 4.8. 95% confidence interval plot of test database coding. Based on the data shown in Table 4.7........................................................................................................76
List of Tables

Table 2.1. Example of LZ77 encoding.................................................................26

Table 2.2. Example of symbol probabilities and interval values for arithmetic coding.................................................................27

Table 2.3. Example of arithmetic coding.............................................................28

Table 3.1. Genre breakdown of test signals ..........................................................46

Table 3.2. Comparison of total lossless bitrate and residual signal-to-noise ratio with residual compressed using Monkey's Audio.........................................................51

Table 3.3. Percentage of filesize increase of two embedded lossless coding methods.................................................................54

Table 4.1. Bits/sample comparison of cascade coding with and without frame adaptive method.................................................................65

Table 4.2. The 15 entry codebook for cascade coding, based on [24]. .................66

Table 4.3. Results of cascade and Rice coding with the entropy of the AAC residual signal, with and without LPC. Figures are in bits/sample.........................................................69

Table 4.4. Proposed cascade coder eight-entries codebook.................................71

Table 4.5. Result of the new proposed codebook training and testing....................72

Table 4.6. Result of the codebook entry 8 training using the test portion of the Training database.................................................................72

Table 4.7. Test database coding result, arranged by signal....................................75
Table 4.8. Test database coding result, arranged by codebook entry........................ 75
### List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAC</td>
<td>Advanced Audio Coding</td>
</tr>
<tr>
<td>AAZ</td>
<td>Advanced Audio Zip</td>
</tr>
<tr>
<td>BPGC</td>
<td>Bit Plane Golomb Code</td>
</tr>
<tr>
<td>BSAC</td>
<td>Bit Sliced Arithmetic Coding</td>
</tr>
<tr>
<td>EBCOT</td>
<td>Embedded Block Coding with Optimize Truncation</td>
</tr>
<tr>
<td>EZW</td>
<td>Embedded Zerotree Wavelet</td>
</tr>
<tr>
<td>IID</td>
<td>Independent and Identically Distributed</td>
</tr>
<tr>
<td>JPEG</td>
<td>Joint Photographic Experts Group</td>
</tr>
<tr>
<td>LPC</td>
<td>Linear Predictive Coding</td>
</tr>
<tr>
<td>LSB</td>
<td>Least Significant Bit</td>
</tr>
<tr>
<td>LTP</td>
<td>Long Term Prediction</td>
</tr>
<tr>
<td>MAC</td>
<td>Monkey’s Audio Coder</td>
</tr>
<tr>
<td>MDCT</td>
<td>Modified Discrete Cosine Transform</td>
</tr>
<tr>
<td>MGE</td>
<td>Multigrid Embedding</td>
</tr>
<tr>
<td>MPEG</td>
<td>Motion Picture Experts Group</td>
</tr>
<tr>
<td>MSB</td>
<td>Most Significant Bit</td>
</tr>
<tr>
<td>PCM</td>
<td>Pulse Code Modulation</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>PDF</td>
<td>Probability Density Function</td>
</tr>
<tr>
<td>PNS</td>
<td>Perceptual Noise Substitution</td>
</tr>
<tr>
<td>SFM</td>
<td>Spectral Flatness Measure</td>
</tr>
<tr>
<td>SMR</td>
<td>Signal to Mask Ratio</td>
</tr>
<tr>
<td>SNR</td>
<td>Signal to Noise Ratio</td>
</tr>
<tr>
<td>SPIHT</td>
<td>Set Partitioning in Hierarchical Tree</td>
</tr>
<tr>
<td>TWINVQ</td>
<td>Transform-domain Weighted Interleave Vector Quantization</td>
</tr>
</tbody>
</table>