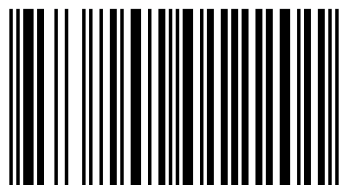


Nowadays, many methods and algorithms have been proposed to improve the compression performance on multichannel audio. This book focuses on performance evaluation of various algorithms on multichannel audio compression. First, we identified and investigated current state-of-the-art audio compression algorithms, both lossless and lossy compression, which can handle mono, stereo, 5.1, and 7.1 multichannel audio. Out of various algorithms available, AC3, AAC, and Ogg have been selected as lossy compression algorithms, while FLAC and MPEG-4 ALS have been chosen as lossless compression algorithms. Two performance measures were used in the experiments, i.e. compression ratio and encoding time. The results showed that among three lossy audio compression algorithms, AC3 has the fastest encoding time while Ogg Vorbis has the highest compression ratio. Furthermore, between FLAC and MPEG-4 ALS, FLAC has faster encoding time and MPEG-4 ALS has higher compression ratio. Overall, in terms of encoding time and compression ratio, it has been found that FLAC is the fastest coder while Ogg Vorbis has the highest compression ratio among five encoders evaluated.



@Siti Aisyah Abdul Rashid obtained her BEng in Communication Engineering from IIUM in 2016. @Teddy Surya Gunawan received his PhD degree from UNSW in 2007 and is currently an Associate Professor at IIUM. @Mira Kartiwi obtained her PhD from UOW in 2009 and is currently an Associate Professor at IIUM.



978-620-2-01220-1

Siti Aisyah Abdul Rashid
Teddy Surya Gunawan
Mira Kartiwi

Multichannel Audio Compression

Performance Evaluation of Various Algorithms

**Siti Aisyah Abdul Rashid
Teddy Surya Gunawan
Mira Kartiwi**

Multichannel Audio Compression

**Siti Aisyah Abdul Rashid
Teddy Surya Gunawan
Mira Kartiwi**

**Multichannel Audio Compression
Performance Evaluation of Various Algorithms**

LAP LAMBERT Academic Publishing

Imprint

Any brand names and product names mentioned in this book are subject to trademark, brand or patent protection and are trademarks or registered trademarks of their respective holders. The use of brand names, product names, common names, trade names, product descriptions etc. even without a particular marking in this work is in no way to be construed to mean that such names may be regarded as unrestricted in respect of trademark and brand protection legislation and could thus be used by anyone.

Cover image: www.ingimage.com

Publisher:

LAP LAMBERT Academic Publishing

is a trademark of

International Book Market Service Ltd., member of OmniScriptum Publishing Group

17 Meldrum Street, Beau Bassin 71504, Mauritius

Printed at: see last page

ISBN: 978-620-2-01220-1

Zugl. / Approved by: Kuala Lumpur, IIUM, 2017

Copyright © Siti Aisyah Abdul Rashid, Teddy Surya Gunawan, Mira Kartiwi

Copyright © 2017 International Book Market Service Ltd., member of OmniScriptum Publishing Group

All rights reserved. Beau Bassin 2017

ABSTRACT

Multichannel audio or surround sound compression is rather more challenging to compress compare to mono and stereo audio. Nowadays, many methods and algorithms have been proposed to improve the compression performance on multichannel audio. This book focuses on performance evaluation of various algorithms on multichannel audio compression. First, we identified and investigated current state-of-the-art audio compression algorithms, both lossless and lossy compression, which can handle mono, stereo, 5.1, and 7.1 multichannel audio. Out of various algorithms available, AC3, AAC, and Ogg have been selected as lossy compression algorithms, while FLAC and MPEG-4 ALS have been chosen as lossless compression algorithms. Two performance measure were used in the experiments, i.e. compression ratio and encoding time. The results showed that among three lossy audio compression algorithms, AC3 has the fastest encoding time while Ogg Vorbis has the highest compression ratio. Furthermore, between FLAC and MPEG-4 ALS, FLAC has faster encoding time and MPEG-4 ALS has higher compression ratio. Overall, in terms of encoding time and compression ratio, it has been found that FLAC is the fastest coder while Ogg Vorbis has the highest compression ratio among five encoders evaluated.

Keywords: Multichannel Audio, lossless compression, lossy compression, encoding time, compression ratio

ACKNOWLEDGMENTS

The authors would like to express their gratitude to the Malaysian Ministry of Higher Education (MOHE), which has provided funding for the research through the Fundamental Research Grant Scheme, FRGS15-194-0435.

TABLE OF CONTENTS

ABSTRACT	I
ACKNOWLEDGMENTS	III
TABLE OF CONTENTS.....	V
LIST OF FIGURES	IX
LIST OF TABLES	XIII
CHAPTER 1 INTRODUCTION	1
1.1 AUDIO COMPRESSION ALGORITHMS.....	1
1.2 PROBLEM STATEMENT	2
1.3 OBJECTIVES	2
1.4 METHODOLOGY	2
1.5 BOOK ORGANIZATION	3
CHAPTER 2 LITERATURE REVIEW	5
2.1 OVERVIEW	5
2.1 DIGITAL REPRESENTATION OF AUDIO SIGNALS	5
2.1.1 <i>Mono audio signal</i>	6
2.1.2 <i>Stereo audio signals</i>	6
2.1.3 <i>Multichannel 5.1 audio signals</i>	7
2.1.4 <i>Multichannel 7.1 audio signals</i>	8
2.1.5 <i>Multichannel 10.2 audio signals</i>	9
2.2 TIME-FREQUENCY REPRESENTATION OF AUDIO SIGNAL	10
2.3 FUNDAMENTAL OF PSYCHOACOUSTICS	11
2.3.1 <i>Lossless and Lossy Audio Coding</i>	11
2.3.2 <i>Perceptual Encoding</i>	11

2.4 PERCEPTUAL ENCODER OF LOSSY ALGORITHM	13
2.4.1 MPEG-1 Layer III (MP3).....	13
2.4.2 Advanced Audio Coding (AAC).....	14
2.4.3 Dolby Audio Codec 3 (AC3).....	15
2.4.4 ITU-T G.719 Encoder	16
2.5 PERCEPTUAL ENCODER OF LOSSLESS ALGORITHM	17
2.5.1 MPEG-4 ALS	17
2.5.2 FLAC.....	18
2.6 RELATED WORKS.....	21
2.6.1 MPEG-4 ALS: Performance, Applications, and Related Standardization Activities	21
2.6.2 Pitch Prediction Algorithm for G.719 Audio Codec.....	22
2.6.3 Down-mix and Op-mix Step-by-step Techniques in Frequency Domain.....	24
2.7 PERFORMANCE MEASURE	27
2.7.1 Compression Ratio	27
2.7.2 Subjective Quality Evaluation.....	27
2.9 SUMMARY	28
CHAPTER 3 RESEARCH METHODOLOGY	29
3.1 INTRODUCTION	29
3.2 AUDIO ENCODER IMPLEMENTATION	29
3.2.1 AAC Encoder	30
3.2.2 OGG Vorbis Encoder.....	31
3.2.3 AC3 Encoder.....	32
3.2.4 FLAC Encoder	33
3.2.5 MPEG-4 ALS Encoder.....	34
3.3 MULTICHANNEL DECORRELATION	35
3.3.1 Encoder Multichannel Decorrelation	35
3.3.2 Decoder Multichannel Reconstruction.....	37
3.4 ENCODING TIME MEASUREMENT USING MATLAB	38
3.5 COMPRESSION RATIO MEASUREMENT USING MATLAB	38

3.6 MATLAB IMPLEMENTATION	38
3.7 SUMMARY	42
CHAPTER 4 RESULTS AND ANALYSIS	43
4.1 INTRODUCTION	43
4.2 AUDIO DATABASE	43
4.3 EXPERIMENTAL SETUP	47
4.4 EXPERIMENTAL RESULTS.....	48
4.4.1 <i>Lossless to lossy compression</i>	48
4.5.2 <i>Lossless to lossless compression</i>	49
4.5 DISCUSSIONS	51
4.6 SUMMARY	52
CHAPTER 5 CONCLUSIONS AND FUTURE WORKS	53
5.1 CONCLUSIONS	53
5.2 FUTURE WORKS	54
REFERENCES	55
APPENDIX A SOFTWARE DOCUMENTATION	57
A.1 NEROAACENC SOFTWARE DOCUMENTATION	57
A.2 OGGENC2 SOFTWARE DOCUMENTATION	58
A.3 FFMPG SOFTWARE DOCUMENTATION	60
APPENDIX B AUDIO TIME WAVEFORMS AND SPECTROGRAMS.....	73
B.1 MONO SIGNALS	73
B.2 STEREO SIGNALS	76
B.3 5.1 MULTICHANNEL SIGNALS.....	78
B.4 7.1 MULTICHANNEL SIGNALS.....	82

LIST OF FIGURES

Figure 2.1: Pulse Code Modulation (Bosi, 1994)	5
Figure 2.2: Stereo Setup (Cheng et al., 2011).....	7
Figure 2.3: ITU 5.1 multichannel loudspeaker arrangements (Cheng et al., 2011).....	8
Figure 2.4: ITU 7.1 multichannel loudspeaker arrangements (Cheng et al., 2011).....	8
Figure 2.5: 10.2 multichannel loudspeaker arrangements (Cheng et al., 2011).	9
Figure 2.6: Perceptual Encoding (Herre & Disch, 2013).....	12
Figure 2.7: Block Diagram of an MP3 Encoder (Lagerströmm.K, 2013)	13
Figure 2.8: Block Diagram of AAC (Cheng et al., 2011).....	14
Figure 2.9: Block Diagram of Dolby AC3 encoder (Liu, 1999).....	15
Figure 2.10: Block Diagram of G.719 encoder (Jansson, 2011).....	16
Figure 2.11: Block diagram of MPEG-4 ALS encoder (Liebchen, 2009).....	18
Figure 2.12: Block diagram of FLAC encoder	19
Figure 2.13: Block diagram for audio archiving tool encoder (Harada et al., 2006).....	21
Figure 2.14: Block diagram of pitch prediction (Jiang et al., 2012).....	22
Figure 2.15: Multi-Channel Audio Encoding Proposed By (Jiang et al., 2013).....	24
Figure 2.16: Decoder block diagram of multi-channel audio (Jiang et al., 2013)	25
Figure 3.1: Encoding uncompressed audio to AAC using NeroAacEnc	31
Figure 3.2: Encoding uncompressed audio to OGG using Oggenc2	32
Figure 3.3: Encoding uncompressed audio to AC3 using FFmpeg	33
Figure 3.4: FFmpeg encode uncompressed audio to FLAC	34
Figure 3.5: FFmpeg encode uncompressed audio to MPEG-4 ALS.....	34
Figure 3.6: Dual Channel Decorrelation.....	35
Figure 3.7: Typical Channel Decorrelation for Multichannel Audio (2.1, 3.1, 5.1, and 7.1)	36

Figure 3.8: Matlab Excerpts to Measure Encoding Time	38
Figure 3.9: Matlab Excerpts to Measure Compression Ratio	38
Figure 3.10: Program Flowchart	39
Figure 3.11: Overall Matlab Implementation	41
Figure 4.1: Time domain waveform and its spectrogram for mono audio File1	44
Figure 4.2: Time domain waveform and its spectrogram for stereo audio File1	45
Figure 4.3: Time domain waveform and its spectrogram for 5.1 multichannel audio File1	45
Figure 4.4: Time domain waveform and its spectrogram for 7.1 multichannel audio File1	46
Figure 4.5: Block Diagram of Lossless and Lossy Audio Coding.....	47
Figure B.1: Time domain waveform and its spectrogram for mono audio File1.....	73
Figure B.2: Time domain waveform and its spectrogram for mono audio File2.....	74
Figure B.3: Time domain waveform and its spectrogram for mono audio File3.....	74
Figure B.4: Time domain waveform and its spectrogram for mono audio File4.....	75
Figure B.5: Time domain waveform and its spectrogram for mono audio File5.....	75
Figure B.6: Time domain waveform and its spectrogram for stereo audio File1	76
Figure B.7: Time domain waveform and its spectrogram for stereo audio File2	76
Figure B.8: Time domain waveform and its spectrogram for stereo audio File3	76
Figure B.9: Time domain waveform and its spectrogram for stereo audio File4	77
Figure B.10: Time domain waveform and its spectrogram for stereo audio File5	77
Figure B.11: Time domain waveform and its spectrogram for 5.1 multichannel audio File1	78
Figure B.12: Time domain waveform and its spectrogram for 5.1 multichannel audio File2.....	79
Figure B.13: Time domain waveform and its spectrogram for 5.1 multichannel audio File3.....	80
Figure B.14: Time domain waveform and its spectrogram for 5.1 multichannel audio File4.....	81
Figure B.15: Time domain waveform and its spectrogram for 5.1 multichannel audio File5.....	82
Figure B.16: Time domain waveform and its spectrogram for 7.1 multichannel audio File1	83
Figure B.17: Time domain waveform and its spectrogram for 7.1 multichannel audio File2.....	84

Figure B.18: Time domain waveform and its spectrogram for 7.1 multichannel audio File3 85

Figure B.19: Time domain waveform and its spectrogram for 7.1 multichannel audio File4 86

Figure B.20: Time domain waveform and its spectrogram for 7.1 multichannel audio File5 87

LIST OF TABLES

Table 2.1: Summary of Comparison for Various Audio Compression Algorithms	26
Table 3.1: Audio Format and Software Implementation	29
Table 3.2: AAC Encoder Configuration	30
Table 4.1: Database information of mono and stereo audio	43
Table 4.2: Database information of 5.1 and 7.1 multichannel audio	44
Table 4.3: Comparison of lossy compression in mono audio	48
Table 4.4: Comparison of lossy compression in stereo audio.....	48
Table 4.5: Comparison of lossy compression in 5.1 audio	49
Table 4.6: Comparison of lossy compression in 7.1 audio	49
Table 4.7: Comparison of lossless compression in mono audio	50
Table 4.8: Comparison of lossless compression in stereo audio.....	50
Table 4.9: Comparison of lossless compression in 5.1 audio	50
Table 4.10: Comparison of lossless compression in 7.1 audio	50

CHAPTER 1

INTRODUCTION

1.1 AUDIO COMPRESSION ALGORITHMS

The audio coding proficiency developments as well as their resultant standard have certainly improved the productivity in audio world. Basically, audio compression can be defined as a way to reduce the size of the audio file. The compression is implied by a codec. Codec is a program that uses unique algorithm to shrink the size of data. The important of compression include keeping space and decreasing transmission period. Begin from the monophonic, multichannel audio compression technology further to stereophonic, quadraphonic, 5.1 channels, 7.1 channels and more.

There are two kinds of audio compression algorithm those are lossy and lossless that will be further explain in next topic. For short, lossy audio compression is known by their well-designed system to shrinks file sizes. Advanced Audio Coding (AAC), MPEG-1 Layer III (MP3), Dolby AC-3, Opus, OGG Vorbis and Windows Media Audio Lossy (WMA lossy) are the examples of prevalent foremost lossy audio coding system. AAC can be considered as the most influential multichannel audio coding algorithm. This is due to its ability to support audio channels up to 48 channels and contribute lossless audio for 5.1 channels at sampling rates 320 kbits/s. Meanwhile, AC-3 provides high audio quality at 384kbit/s.

Meanwhile, the most well-known codec in lossless algorithm are Free Lossless Audio Codec (FLAC), Apple Lossless Audio Codec (ALAC), Waveform Audio File (WAV), MPEG-4 Audio lossless, True Audio (TTA), and more. Each of the codec, have the own domain and advantage to encode and decode the audio. Literally, lossless methods do not have any loss information and provide an exact replica of the original signal.

1.2 PROBLEM STATEMENT

As the improvement of audio history developed from mono to stereo audio and then stereo to multichannel audio, more improvement to audio codec is needed to provide high quality audio performance. There is a strong demand in encoding tools in both lossless and lossy algorithm compression in order to give efficient compression ratio and smallest time encoding.

As we know, the previous methods of multi-channel compression mostly focuses on single-channel audio (mono) and the signal coding rate is fairly high. Specifically, in this book, several audio codec will be evaluated on various audio files to study state-of-the art coding of multichannel audio, mono audio and stereo audio. The multi-channel audio compressions algorithms evaluated are AAC, AC3, OGG Vorbis, and MPEG-4 ALS. Matlab will be used to automate the process of encoding and decoding, as well as performance measure in terms of compression ratio and encoding/decoding time.

1.3 OBJECTIVES

The objectives of this book are listed as follows:

- To investigate various algorithms, including lossy and lossless compression, on mono audio, stereo audio, 5.1 and 7.1 multichannel audio signals.
- To evaluate the performance of lossy and lossless audio compression algorithms in terms of encoding time and compression ratio.

1.4 METHODOLOGY

To achieve the objectives stated previously, we conducted several activities as follows:

1. Literature review on the various topics of multichannel audio compression, surround sound and encoding method. Of the various algorithms found, we classify into two main methods which are

lossy and lossless compression, in which we further selected some algorithms for our benchmarking.

2. Determine and analyze information on various codec techniques and methods for multichannel, mono and stereo audio compression.
3. Understanding on current state of the art of multichannel audio compression.
4. Evaluation on the performance of various algorithms in terms of the encoding/decoding time and compression ratio.

1.5 BOOK ORGANIZATION

This book consists of five main chapters starting with Chapter 1 that focus on the introductory part. This chapter also includes the problem statement, objectives, methodology and scope of this research work. Chapter 2 describes the literature review and theoretical background of the multichannel, mono and stereo audio signals as well as various related audio compression algorithms. Chapter 3 presents the experimental setup and implementation. This chapter explains each method on how to encode audio to AAC, OGG Vorbis and AC3 format (lossy compression) as well as FLAC and MPEG-4 ALS audio format (lossless compression). Chapter 4 discusses the experimental result of the encoding simulation of AAC, AC3, OGG Vorbis, FLAC and MPEG-4 ALS using binary on Matlab. Chapter 5 concludes all the findings of this book and also the future work related to this book.