

Audio Watermark

Yiqing Lin • Waleed H. Abdulla

Audio Watermark

A Comprehensive Foundation
Using MATLAB

Yiqing Lin
The University of Auckland
Auckland, New Zealand

Waleed H. Abdulla
The University of Auckland
Auckland, New Zealand

Additional material to this book can be downloaded from <http://extras.springer.com>

ISBN 978-3-319-07973-8 ISBN 978-3-319-07974-5 (eBook)

DOI 10.1007/978-3-319-07974-5

Springer Cham Heidelberg New York Dordrecht London

Library of Congress Control Number: 2014945572

© Springer International Publishing Switzerland 2015

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed. Exempted from this legal reservation are brief excerpts in connection with reviews or scholarly analysis or material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work. Duplication of this publication or parts thereof is permitted only under the provisions of the Copyright Law of the Publisher's location, in its current version, and permission for use must always be obtained from Springer. Permissions for use may be obtained through RightsLink at the Copyright Clearance Center. Violations are liable to prosecution under the respective Copyright Law.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

Audio watermarking is a technique providing a promising solution to copyrights protection for digital audio and multimedia products. Using this technique, hidden information called *watermark* containing copyrights information is imperceptibly embedded into the audio track of a host media. This watermark may be extracted later on from a suspected media to verify the authenticity. To function as an effective tool to enforce ownership rights, the audio watermarking scheme must satisfy the imperceptibility, robustness, security, data payload, and computational complexity requirements. Throughout this book we will be illustrating in a practical way the commonly used and novel approaches of audio watermarking for copyrights protection. We will also introduce our recently developed methods for objectively predicting the perceptual quality of the watermarked audio signals.

This book is directed towards students, researchers, engineers, multimedia practitioners, and academics who are interested in multimedia authentication and audio pirating control. The theoretical descriptions of the watermarking techniques are augmented by MATLAB implementations to ease understanding of the watermarking principles. A GUI demonstration program for watermarking embedding and extraction under different attacks is also provided to quickly surf through the different aspects of the watermarking attributes.

Book Motivations and Objectives

Motivated by the booming of the digital media applications, plenty of research has been conducted to investigate the methods of audio watermarking for copyrights protection. However, clear and easy to follow information about the audio watermarking subject are still not widely available and scattered among many publications. Currently, it is hard to find an easy pathway to develop research in this field. One main reason to this difficulty is that most of the works are bounded by IP or patent constraints. On the implementation side it is still hard to find or write the implementation programs for the known audio watermarking techniques

to see how the algorithms work. This book is introduced to establish a shortcut to get into this interesting field with minimal efforts. The commonly known techniques are well explained and supplemented with MATLAB codes to get a clear idea about how each technique performs. In addition, the reader can reproduce the functional figures of the book with provided MATLAB scripts written specifically for this purpose.

From the robustness and security perspectives, the commonly used audio watermarking techniques have limitations on the resistance to various attacks (especially desynchronization attacks) and/or security against unauthorized detection. Thus, in this book we develop new robust and secure audio watermark algorithm; it is well explained and implemented in MATLAB environment. This algorithm can embed unperceivable, robust, blind, and secure watermarks into digital audio files for the purpose of copyrights protection. In the developed algorithm, additional requirements such as data payload and computational complexity are also taken into account and detailed.

Apart from the improvement of audio watermarking algorithms, another landmark of this book is the exploration of benchmarking approaches to evaluate different algorithms in a fair and objective manner. For the application in copyrights protection, audio watermarking schemes are mainly evaluated in terms of imperceptibility, robustness, and security. In particular, the extent of imperceptibility is graded by perceptual quality assessment, which mostly involves a laborious process of subjective judgment. To facilitate the implementation of automatic perceptual measurement, we explore a new method for reliably predicting the perceptual quality of the watermarked audio signals. A comprehensive evaluation technique is illustrated to let the readers know how to pinpoint the strengths and weaknesses of each technique. The evaluation techniques are supported with tested MATLAB codes.

Furthermore to what we have just stated that this book extensively illustrates several commonly used audio watermarking algorithms for copyrights protection along with the improvement of benchmarking approaches, we may pinpoint the following new contributions of the current book:

- We introduce a spread spectrum based audio watermarking algorithm for copyrights protection, which involves Psychoacoustic Model 1, multiple scrambling, adaptive synchronization, frequency alignment, and coded-image watermark. In comparison with other existing audio watermarking schemes [1–10], the proposed scheme achieves a better compromise between imperceptibility, robustness, and data payload.
- We design a performance evaluation which consists of perceptual quality assessment, robustness test, security analysis, estimations of data payload, and computational complexity. The presented performance evaluation can serve as one comprehensive benchmarking of audio watermarking algorithms.
- We portray objective quality measures adopted in speech processing for perceptual quality evaluation of audio watermarking. Compared to traditional perception modelling, objective quality measures provide a faster and more

efficient method of evaluating the watermarked audio signals relative to host audio signals.

- We analyze methods for implementing psychoacoustic models in the MPEG standard, with the goal of achieving inaudible watermarks at a lower computational cost. With the same level of minimum masking threshold, Psychoacoustic Model 1 requires less computation time than Psychoacoustic Model 2.
- We identify the imperceptibility, robustness, and security characteristics of audio watermarking algorithms and further use them as attacks in the process of multiple watermarking.
- We propose the use of variable frame length to make the investigated cepstrum domain watermarking, wavelet domain watermarking, and echo hiding robust against time-scale modification.

Organization of the Book

The chapters in this book are organized as follows.

Chapter 1 provides an overview of digital watermarking technology and then opens a discussion on audio watermarking for copyrights protection.

Chapter 2 describes the principles of psychoacoustics, including the anatomy of the auditory system, perception of sound, and the phenomenon of auditory masking. Then two psychoacoustic models in the MPEG-1 standard, i.e., Psychoacoustic Model 1 and 2, are investigated. Through comparisons of the masking effect and the computational cost, the minimum masking threshold from Psychoacoustic Model 1 is chosen to be used for amplitude shaping of the watermark signal in Chap. 4.

Chapter 3 begins with the implementation specifications for perceptual quality assessment and the basic robustness test used in this chapter. Then it describes and evaluates several algorithms for audio watermarking, such as least significant bit modification, phase coding, spread spectrum watermarking, cepstrum domain watermarking, wavelet domain watermarking, echo hiding, and histogram-based watermarking. In the meantime, possible enhancements are exploited to improve the capabilities of some algorithms.

Chapter 4 presents a spread spectrum based audio watermarking algorithm for copyrights protection, which uses Psychoacoustic Model 1, multiple scrambling, adaptive synchronization, frequency alignment, and coded-image watermark. The basic idea is to embed the watermark by amplitude modulation on the time–frequency domain of the host audio signal and then detect the watermark by normalized correlation between the watermarked signal and corresponding secret keys.

In Chap. 5, the performance of the proposed audio watermarking algorithm is evaluated in terms of imperceptibility, robustness, security, data payload, and computational complexity. The evaluation starts with perceptual quality assessment, which consists of the subjective listening test (including the MUSHRA test and SDG rating) and the objective evaluation test (including the ODG by PEAQ and

the SNR value). Then, the basic robustness test and the advanced robustness test (including a test with StirMark for Audio, a test under collusion, and a test under multiple watermarking) are carried out. In addition, a security analysis is followed by estimations of data payload and computational complexity. At the end of this chapter, a comparison between the proposed scheme and other reported systems is also presented.

Chapter 6 presents an investigation of objective quality measures for perceptual quality evaluation in the context of different audio watermarking techniques. The definitions of selected objective quality measures are described. In the experiments, two types of Pearson correlation analysis are conducted to evaluate the performance of these measures for predicting the perceptual quality of the watermarked audio signals.

Auckland, New Zealand
Auckland, New Zealand

Yiqing Lin
Waleed H. Abdulla

Contents

- 1 Introduction** 1
 - 1.1 Information Hiding: Steganography and Watermarking 1
 - 1.2 Overview of Digital Watermarking 3
 - 1.2.1 Framework of the Digital Watermarking System 4
 - 1.2.2 Classifications of Digital Watermarking 5
 - 1.2.3 Applications of Digital Watermarking 7
 - 1.2.3.1 Copyrights Protection 7
 - 1.2.3.2 Content Authentication 7
 - 1.2.3.3 Broadcast Monitoring 7
 - 1.2.3.4 Copy Control 8
 - 1.3 Audio Watermarking for Copyrights Protection 8
 - 1.3.1 Requirements for the Audio Watermarking System 8
 - 1.3.1.1 Imperceptibility 9
 - 1.3.1.2 Robustness 9
 - 1.3.1.3 Security 9
 - 1.3.1.4 Data Payload 9
 - 1.3.1.5 Computational Complexity 10
 - 1.3.2 Benchmarking on Audio Watermarking Techniques 10
 - 1.3.2.1 Perceptual Quality Assessment 11
 - 1.3.2.2 Robustness Test 12
 - 1.3.2.3 Security Analysis 13
- 2 Principles of Psychoacoustics** 15
 - 2.1 Physiology of the Auditory System 15
 - 2.1.1 The Outer Ear 16
 - 2.1.2 The Middle Ear 17
 - 2.1.3 The Inner Ear 19
 - 2.2 Sound Perception Concepts 22
 - 2.2.1 Sound Pressure Level and Loudness 22
 - 2.2.2 Hearing Range and Threshold in Quiet 23
 - 2.2.3 Critical Bandwidth 24

| | | |
|----------|--|-----------|
| 2.3 | Auditory Masking | 27 |
| 2.3.1 | Simultaneous Masking | 27 |
| 2.3.1.1 | Narrowband Noise Masking Tone | 28 |
| 2.3.1.2 | Tone Masking Tone | 30 |
| 2.3.1.3 | Narrowband Noise or Tone Masking Narrowband Noise | 31 |
| 2.3.2 | Nonsimultaneous Masking | 32 |
| 2.3.2.1 | Pre-masking | 32 |
| 2.3.2.2 | Post-masking | 33 |
| 2.4 | Psychoacoustic Model | 33 |
| 2.4.1 | Modelling the Effect of Simultaneous Masking | 33 |
| 2.4.1.1 | Models for the Spreading of Masking | 33 |
| 2.4.1.2 | Implementation of Psychoacoustic Model 1 | 35 |
| 2.4.1.3 | Comparison Between Psychoacoustic Model 1 and Model 2 | 44 |
| 2.4.2 | Modelling the Effect of Nonsimultaneous Masking | 47 |
| 2.5 | Summary | 48 |
| 3 | Audio Watermarking Techniques | 51 |
| 3.1 | Specifications on Performance Evaluation | 51 |
| 3.1.1 | Audio Test Signals Used for Evaluation | 52 |
| 3.1.2 | Implementation of Perceptual Quality Assessment | 53 |
| 3.1.3 | Implementation of Robustness Test | 53 |
| 3.1.3.1 | Basic Robustness Test | 53 |
| 3.1.3.2 | Advanced Robustness Test | 55 |
| 3.2 | Audio Watermarking Algorithms | 56 |
| 3.2.1 | Least Significant Bit Modification | 57 |
| 3.2.1.1 | Algorithm | 57 |
| 3.2.1.2 | Performance Evaluation | 58 |
| 3.2.2 | Phase Coding | 59 |
| 3.2.2.1 | Algorithm | 59 |
| 3.2.2.2 | Performance Evaluation | 60 |
| 3.2.3 | Spread Spectrum Watermarking | 63 |
| 3.2.3.1 | Algorithm | 63 |
| 3.2.3.2 | Performance Evaluation | 65 |
| 3.2.4 | Cepstrum Domain Watermarking | 65 |
| 3.2.4.1 | Algorithm | 68 |
| 3.2.4.2 | Strategies for Improvement | 68 |
| 3.2.4.3 | Performance Evaluation | 74 |
| 3.2.5 | Wavelet Domain Watermarking | 75 |
| 3.2.5.1 | Algorithm | 76 |
| 3.2.5.2 | Performance Evaluation | 77 |
| 3.2.6 | Echo Hiding | 78 |
| 3.2.6.1 | Algorithm | 81 |

| | | |
|----------|--|------------|
| 3.2.6.2 | Performance Evaluation | 84 |
| 3.2.7 | Histogram-Based Watermarking | 88 |
| 3.2.7.1 | Algorithm | 88 |
| 3.2.7.2 | Performance Evaluation | 89 |
| 3.3 | Summary | 93 |
| 4 | Proposed Audio Watermarking Scheme | 95 |
| 4.1 | Preliminaries | 95 |
| 4.1.1 | Selection of Watermarking Regions | 96 |
| 4.1.2 | Structure of the Watermarking Domain | 97 |
| 4.1.3 | Gammatone Auditory Filterbank | 100 |
| 4.2 | Watermark Embedding | 101 |
| 4.2.1 | Embedding Algorithm | 101 |
| 4.2.2 | Multiple Scrambling | 103 |
| 4.3 | Watermark Detection | 104 |
| 4.3.1 | Basic Detection | 107 |
| 4.3.2 | Adaptive Synchronization | 111 |
| 4.3.3 | Frequency Alignment Towards Excessive PITSM and TPPSM | 113 |
| 4.3.3.1 | Frequency Alignment Against TSM and PSM | 113 |
| 4.3.3.2 | Implementation of Frequency Alignment | 114 |
| 4.3.3.3 | Error Analysis Associated with T_{BER} | 116 |
| 4.4 | Coded-Image Watermark | 118 |
| 4.5 | Summary | 120 |
| 5 | Performance Evaluation of Audio Watermarking | 123 |
| 5.1 | Experimental Setup | 123 |
| 5.2 | Perceptual Quality Assessment | 127 |
| 5.2.1 | Subjective Listening Test | 127 |
| 5.2.2 | Objective Evaluation Test | 131 |
| 5.3 | Robustness Test | 132 |
| 5.3.1 | Error Probability | 132 |
| 5.3.2 | Basic Robustness Test | 133 |
| 5.3.3 | Advanced Robustness Test | 139 |
| 5.3.3.1 | Test with StirMark for Audio | 139 |
| 5.3.3.2 | Test Under Collusion | 142 |
| 5.3.3.3 | Test Under Multiple Watermarking | 144 |
| 5.4 | Security Analysis | 151 |
| 5.5 | Data Payload and Computational Complexity | 151 |
| 5.5.1 | Estimation of Data Payload | 151 |
| 5.5.2 | Estimation of Computational Complexity | 153 |
| 5.6 | Performance Comparison | 154 |
| 5.7 | Summary | 157 |




| | | |
|----------|---|------------|
| 6 | Perceptual Evaluation Using Objective Quality Measures | 159 |
| 6.1 | Perceptual Quality Evaluation | 159 |
| 6.2 | Objective Quality Measures | 161 |
| 6.3 | Experiments and Discussion..... | 164 |
| 6.3.1 | Audio Watermarking Techniques Default Settings..... | 164 |
| 6.3.2 | Subjective Listening Tests..... | 165 |
| 6.3.3 | Objective Evaluation Tests | 166 |
| 6.3.4 | Performance Evaluation Using Correlation Analysis | 169 |
| 6.4 | Summary | 175 |
| A | SDMI Standard | 177 |
| B | STEP 2000 | 179 |
| C | StirMark for Audio | 181 |
| D | Critical Bandwidth..... | 185 |
| E | List of Audio Test Files | 187 |
| F | Basic Robustness Test..... | 189 |
| G | Nonuniform Subbands | 191 |
| | References..... | 193 |

List of Figures

| | | |
|----------|--|----|
| Fig. 1.1 | A generic digital watermarking system | 4 |
| Fig. 2.1 | Structure of the peripheral auditory system | 16 |
| Fig. 2.2 | Average pressure levels at auditory canal entrance versus free-field pressure, at six azimuthal angles of incidence. <i>Notes:</i> (1) The sound pressure was measured with a probe tube located at the left ear of the subject. (2) A point source of sound was moved around a horizontal circle of radius 1 m with the subject's head at the center. At $\theta = 0^\circ$, the subject was facing the source, and at $\theta = 90^\circ$, the source was normally incident at plane of left ear | 18 |
| Fig. 2.3 | Anatomy of the cochlea (a) Relative location of the cochlea in the inner ear (b) Schematic of the unraveled cochlea (c) Cross-section through one cochlea turn | 20 |
| Fig. 2.4 | Resonant properties of the basilar membrane (a) Envelopes of vibration patterns on the basilar membrane in response to sound of different frequencies (b) Distribution of resonant frequencies along the basilar membrane..... | 21 |
| Fig. 2.5 | Equal-loudness contours | 23 |
| Fig. 2.6 | Hearing range | 24 |
| Fig. 2.7 | Approximation for the threshold in quiet (a) Frequency on a linear scale (b) Frequency on a logarithmic scale | 25 |
| Fig. 2.8 | Threshold in quiet on Bark scale | 26 |

| | | |
|-----------|--|----|
| Fig. 2.9 | Determination of the critical bandwidth (a) The threshold for a narrowband noise 2 kHz centered between two tones of 50 dB as a function of the frequency separation between two tones (b) The threshold for a tone of 2 kHz centered between two narrowband noises of 50 dB as a function of the frequency separation between the cutoff frequencies of two noises..... | 26 |
| Fig. 2.10 | Two types of masking: simultaneous and nonsimultaneous masking | 27 |
| Fig. 2.11 | Simultaneous masking | 28 |
| Fig. 2.12 | Masking thresholds for a 60 dB narrowband noise masker centered at different frequencies | 29 |
| Fig. 2.13 | Masking thresholds for a 60 dB narrowband noise masker centered at different frequencies in Bark scale | 30 |
| Fig. 2.14 | Masking thresholds from a 1 kHz narrowband noise masker at different SPLs | 30 |
| Fig. 2.15 | Masking thresholds from a 1 kHz narrowband noise masker at different SPLs in Bark scale | 31 |
| Fig. 2.16 | Masking thresholds from a 1 kHz tonal masker at different SPLs | 32 |
| Fig. 2.17 | Spreading function in ISO/IEC Psychoacoustic Model 1 | 35 |
| Fig. 2.18 | Comparison of four spreading functions relative to an 80 dB masker | 36 |
| Fig. 2.19 | Initial and normalized PSD estimates (a) Frequency on linear scale (b) Frequency on Bark scale | 38 |
| Fig. 2.20 | Tonal and nontonal maskers (a) Frequency on a linear scale (b) Frequency on Bark scale | 40 |
| Fig. 2.21 | Individual masking thresholds (a) Frequency on linear scale (b) Frequency on Bark scale | 43 |
| Fig. 2.22 | Global masking threshold and minimum masking threshold (a) Frequency on linear scale (b) Frequency on Bark scale | 45 |
| Fig. 2.23 | Mapping between spectral subsamples and subbands | 46 |
| Fig. 2.24 | Comparison of MMTs from Psychoacoustic Model 1 and 2 | 47 |
| Fig. 2.25 | Modelling the effect of post-masking | 48 |
| Fig. 3.1 | An example of a two-channel stereo signal | 52 |
| Fig. 3.2 | Host signal and a watermarked signal by LSB modification. Note that the watermarked signal is produced by using $L = 6$ and modifying the third and fourth decimal places. (a) Host audio signal. (b) Watermarked audio signal. (c) Difference between the watermarked and host audio signals..... | 58 |

| | | |
|-----------|---|----|
| Fig. 3.3 | Host signal and a watermarked signal by the modified phase-coding method. Note that the watermarked signal is produced by watermarking with $N = 2,048$ and $n_e = 128$. (a) Host audio signal. (b) Watermarked audio signal. (c) Difference between the watermarked and host audio signals..... | 62 |
| Fig. 3.4 | Block diagram of basic SS watermarking scheme. (a) Embedding process. (b) Detection process | 63 |
| Fig. 3.5 | Host signal and a watermarked signal by SS watermarking. Note that the watermarked signal is produced by watermarking with $N = 4,096$, $n_r = 3$ and $\beta = 0.03$. (a) Host audio signal. (b) Watermarked audio signal. (c) Difference between the watermarked and host audio signals..... | 66 |
| Fig. 3.6 | Block diagram of computing the complex cepstrum and the inverse complex cepstrum. (a) Complex cepstrum $\hat{x}(n) = F^{-1} \{\log(F\{x(n)\})\}$. (b) Inverse complex cepstrum $x(n) = F^{-1} \{\exp(F\{\hat{x}(n)\})\}$ | 67 |
| Fig. 3.7 | Distributions of R_{one} and R_{zero} under different attacks. Note that these data are produced by watermarking with $N = 4,096$, $\alpha_w = 0.001$, and $n_r = 3$ | 70 |
| Fig. 3.8 | Host signal and a watermarked signal by cepstrum domain watermarking. Note that the watermarked signal is produced by watermarking with $N = 2,048$, $\alpha_w = 0.0015$ and $n_r = 3$. (a) Host audio signal. (b) Watermarked audio signal. (c) Difference between the watermarked and host audio signals..... | 74 |
| Fig. 3.9 | A three-level DWT decomposition and reconstruction. (a) Wavelet decomposition. (b) Wavelet reconstruction. | 76 |
| Fig. 3.10 | Host signal and a watermarked signal by wavelet domain watermarking. Note that the watermarked signal is produced by watermarking with $N = 2,048$, $n_r = 3$, and $\alpha_w = 0.01$. (a) Host audio signal. (b) Watermarked audio signal. (c) Difference between the watermarked and host audio signals..... | 78 |
| Fig. 3.11 | Impulse response of echo kernels. (a) “One” kernel. (b) “Zero” kernel | 81 |
| Fig. 3.12 | Auto-cepstrum for echo detection | 83 |
| Fig. 3.13 | Host signal and a watermarked signal by echo hiding. Note that the watermarked signal is produced by watermarking with $N = 4,096$, $\alpha = 0.2$, and $n_r = 3$. (a) Host audio signal. (b) Watermarked audio signal. (c) Difference between the watermarked and host audio signals .. | 85 |

| | | |
|-----------|---|-----|
| Fig. 3.14 | Host signal and a watermarked signal by histogram-based watermarking. Note that the watermarked signal is produced by watermarking with $N_w = 40$, $\lambda = 2.2$, and $E_h = 1.5$. (a) Host audio signal. (b) Watermarked audio signal. (c) Difference between the watermarked and host audio signals..... | 90 |
| Fig. 4.1 | Selection of watermarking regions (a) $E_T = 0.01$ (b) $E_T = 0.1$ | 97 |
| Fig. 4.2 | Diagram of blocks in the watermarking domain..... | 97 |
| Fig. 4.3 | Configuration of one block | 98 |
| Fig. 4.4 | Distribution of the watermark bits and synchronization bit. <i>Note:</i> Slots,  ,  and  are used for embedding B_1 , B_2 , and B_s , respectively | 99 |
| Fig. 4.5 | Frequency response of a 16-channel GTF | 100 |
| Fig. 4.6 | Block diagram of watermarking one host frame..... | 102 |
| Fig. 4.7 | Host signal and a watermarked signal by the proposed scheme. (a) Host signal. (b) Watermark signal. (c) Watermarked signal | 104 |
| Fig. 4.8 | Block diagram of watermark detection. <i>Notes:</i> (1) Basic detection works independently. (2) Adaptive synchronization is an improvement technique for block synchronization. (3) Frequency alignment indicated by dashed lines is an additional solution to excess PITSM and TPPSM..... | 106 |
| Fig. 4.9 | Illustration of random samples cropping | 111 |
| Fig. 4.10 | Flowchart of adaptive synchronization..... | 112 |
| Fig. 4.11 | Duality between TSM and PSM..... | 114 |
| Fig. 4.12 | Illustration of frequency alignment (a) Positive PSM (b) Negative PSM | 115 |
| Fig. 4.13 | Coded-image denoising by morphological operations | 119 |
| Fig. 4.14 | Character recognition by the neural network. (a) Letters “C” “O” “P” “Y” “R” “I” “G” “H” “T”. (b) Noisy coded-image watermark. (c) Recovered coded-image watermark | 120 |
| Fig. 5.1 | Determination of watermark strength. α_w (a) $\alpha_w = 50$ for <i>Bass.wav</i> . (b) $\alpha_w = 100$ for <i>Pop.wav</i> | 126 |

| | | |
|----------|--|-----|
| Fig. 5.2 | Screenshot of the MATLAB GUI for the MUSHRA test. The buttons on the GUI have the following functions: “Load,” load the host audio signal to be evaluated. “Start,” start playing the sound from the beginning. “Pause/Stop,” pause or stop the sound that is currently playing. “Resume,” resume the sound from the pause position. “Save,” save the host signal name and the participant name as well as the registered scores into a .txt file. “Reset,” reset the interface for the next trial | 128 |
| Fig. 5.3 | Results of the MUSHRA-based subjective listening test..... | 130 |
| Fig. 6.1 | Objective evaluation via perception modelling | 161 |
| Fig. 6.2 | Evaluation of PEMO-Q, PEAQ, and EAQUAL | 167 |
| Fig. 6.3 | Overall correlation coefficients over audio test signals A_1 , A_2 , and A_3 | 174 |

List of Tables

| | | |
|------------|---|-----|
| Table 1.1 | Classifications of digital watermarking..... | 5 |
| Table 1.2 | Subjective difference grade (SDG) | 11 |
| Table 3.1 | Results of performance evaluation of LSB modification | 59 |
| Table 3.2 | Results of performance evaluation of phase coding | 62 |
| Table 3.3 | Results of performance evaluation of SS watermarking | 67 |
| Table 3.4 | Results of performance evaluation of cepstrum domain watermarking | 71 |
| Table 3.5 | Results of performance evaluation of wavelet domain watermarking | 79 |
| Table 3.6 | Results of performance evaluation of echo hiding | 86 |
| Table 3.7 | Results of performance evaluation of histogram-based watermarking | 91 |
| Table 5.1 | Results of the SDG-based subjective listening test | 131 |
| Table 5.2 | Results of the objective evaluation test | 131 |
| Table 5.3 | Results of error probabilities under $T_{BER} = 20\%$ | 132 |
| Table 5.4 | Results of the basic robustness test on the watermarked <i>Bass.wav</i> signal | 134 |
| Table 5.5 | Results of the basic robustness test on the watermarked <i>Gspi.wav</i> signal | 135 |
| Table 5.6 | Results of the basic robustness test on the watermarked <i>Harp.wav</i> signal | 136 |
| Table 5.7 | Results of the basic robustness test on the watermarked <i>Pop.wav</i> signal | 137 |
| Table 5.8 | Results of combined attacks on the watermarked <i>Bass.wav</i> signal | 139 |
| Table 5.9 | Results of StirMark for Audio attacks | 141 |
| Table 5.10 | Results of the averaging collusion attack | 144 |
| Table 5.11 | Results of multiple self-watermarking | 146 |
| Table 5.12 | Results of inter-watermarking Experiment I | 149 |

| | | |
|------------|--|-----|
| Table 5.13 | Results of inter-watermarking Experiment II on <i>Bass.wav</i> signal | 150 |
| Table 5.14 | Results of the computational complexity estimation | 153 |
| Table 5.15 | Performance comparison of different audio watermarking schemes | 155 |
| Table 6.1 | Comparison of the total computation time (s)..... | 168 |
| Table 6.2 | Pearson correlation coefficients under our proposed scheme | 171 |
| Table 6.3 | Pearson correlation coefficients under cepstrum domain watermarking | 171 |
| Table 6.4 | Pearson correlation coefficients under wavelet domain watermarking | 172 |
| Table 6.5 | Pearson correlation coefficients under echo hiding | 172 |
| Table 6.6 | Pearson correlation coefficients under histogram-based watermarking | 173 |
| Table A.1 | Robustness test items in SDMI | 178 |
| Table B.1 | Robustness test items in STEP 2000 | 180 |
| Table C.1 | Robustness test items in StirMark for Audio | 181 |
| Table D.1 | Critical bands over the frequency spectrum [11]..... | 185 |
| Table E.1 | Descriptions of audio test files for performance evaluation | 187 |
| Table F.1 | Descriptions of basic robustness test | 189 |
| Table G.1 | Thirty-two nonuniform subbands over the frequency spectrum.. | 191 |