Signal Processing and Machine Learning
with Applications

Michael M. Richter •
Sheuli Paul • Veton Këpuska • Marius Silaghi

Signal Processing and Machine Learning with Applications

Springer

Michael M. Richter (deceased)
Electrical and Computer Engineering
University of Kaiserslautern
Kaiserslautern, Germany

Sheuli Paul
FB Elektrotechnik & Informationstechnik
Technische Universität Kaiserslautern
Kaiserslautern, Germany

Veton Këpuska
Florida Institute of Technology
Melbourne, FL, USA

Marius Silaghi
Florida Institute of Technology
Melbourne, FL, USA

ISBN 978-3-319-45371-2 ISBN 978-3-319-45372-9 (eBook)
https://doi.org/10.1007/978-3-319-45372-9

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Switzerland AG 2022
This work is subject to copyright. All rights are solely and exclusively licensed 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.
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.
The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

Learning has always a goal. For this reason, for a given process, the question is what can we learn about it? Despite the diversities in processes, one can find out the common principles underlying them. Therefore, we discuss stochastic processes from a general point of view, first. As for practical applications, we will consider them through the lens of historical developments by considering the following aspects:

1. The invention of numbers and number systems is our starting point. From numbers and their representations, humans invented the theory of signals and systems. Signals do not occur isolated but within a process.
2. The naturally occurring processes are not deterministic but rather stochastic. Hence, the underlying states are often unknown. That means those processes are mostly described with hidden stochastic models (HSM).
3. Signal processes contain information. However, this information is not directly visible or understandable to humans. The information is hidden in properties of the processes. Properties of signal processes are mostly not directly measurable, and we have to find them out, at least approximately.
4. Information about the unknown elements is necessary. One needs methods to capture them, and here lie the major challenges. To address these challenges, we concentrate on machine learning techniques.

Real-world problems are conceptualized, explained in applied domains. The book is intended for persons interested in the techniques and foundations related to signal processing. As mentioned previously, signals occur within processes and are considered not to be fully observable and predictable; hence, they are ruled by probabilities. Therefore, they are termed stochastic processes. We concentrate on properties of signals that are useful for practical applications. One has rather to manipulate the information available about processes using computations. Doing this properly is a major challenge. The computations needed depend very much on the type of signals and the application domain.

The development of practical applications was foremost in our mind when writing this book. principles where the problems and differences are exhibited

in their realizations. today one is deploying robotic systems to perform various tasks. This in turn requires solving communication problems between humans and machines. communicate. Humans and machines can communicate, for instance, using speech, images, and videos. These all have in common that a physical exchange has to take place. At the lowest level, it requires use of elementary signals that could be re-interpreted on higher levels of abstraction. We describe our methods as organized in a machine learning process. There, one records many examples of signals in order to use them when applying machine learning techniques. The techniques involve several transformations. A major difficulty is to find out which transformation one should apply in which situation.

- First, audio typically uses cascaded IIR filters to obtain high performance with minimal latency. Noting that each filtering stage propagates the errors of previous stages, a high degree of precision in both the signal and coefficients is required in order to minimize the effects of these propagated errors.
- Second, the signal accuracy must be maintained, even as its magnitude approaches zero (this is necessary because of the sensitivity of the human ear).

The signals occur in sets, and these sets contain information and have special structures. Important structures are governed by logic. Often, the signals arise from some events and are not part of a systematic communication. Sometimes they are represented in a linear ordering. However, typically they are not ordered and are often described through general stochastic processes.

Humans are interested in the information contained in any process, be it discrete (e.g., digital) or continuous and modeled by deterministic or stochastic processes. One can employ machine learning methods to find the information governing stochastic processes. Here, one needs to employ some combinatorial, probabilistic, and learning algorithms. Learning in particular is needed if the signal sets are complex, or the signals are corrupted by noise. is for efficiency reasons very restricted by the large number of occurring signals and by the noise. explore systematically these topics. Therefore, we show their usefulness in a number of widely reaching application areas. We intend to represent the whole picture of signal processing instead of restricting it to special aspects. Then we look at recognition problems. For both, machine learning provides a wealth of important techniques. Machine learning is a wide area. In this book, we discuss aspects of it related to signal processes. These are mainly supervised and unsupervised learning methods. Finally, we discuss real applications. Noise is one of the main issues when we apply theoretical aspects to solve real-world problems. take place in a clean (noise free) environment and noise often corrupts it. Therefore, we included the study and handling of noise. Then, we discussed multiple application areas. A very important such area is speech processing.

The book is written for students, scientists, researchers, engineers, and practitioners. The readers should have an understanding of basic computational methods as they are taught in the first three semesters of computer science undergraduate curricula. For usage as a textbook, lectures are supported by a systematic presenta-

tion of the material with examples and exercises including background information accompanied with literature discussions.

This work was produced by the authors and was put together in a coherent form. We did our best to give credit to the materials owned by others. We have a reference section at the end of each chapter. In case of any unintentional omission of any work or any part of the text owned by others, we will remain deeply apologetic for this. We read, used, and followed many materials from earlier works done by many other contributors in signal processing and machine learning and generated our work uniquely. We acknowledge these learning and knowledge sharing and remain thankful to all these many contributors.

We would like to thank Ronan Nugent for all supportive arrangements, Bijoy Paul and Michael Sintek for the help from time to time as needed. We also would like to thank the involved Springer team, Celine Chang, Paul Drougas, and Shina Harshavardhan, during this critical transition period for their understanding and cooperation as needed to complete the task.

Kaiserslautern, Germany — Michael M. Richter
Kaiserslautern, Germany — Sheuli Paul
Melbourne, FL, USA — Veton Këpuska
Melbourne, FL, USA — Marius Silaghi
December, 2021

Problems, Challenges, and Principal Methods

A very first question is what a signal is. There are different interpretations of this term. The definitions are mainly motivated by the area in which one is interested. Examples are

- Digital signals, such as Boolean values, integers, or real numbers
- Arbitrary objects like words, documents, messages, and images, etc

In this book, the focus is on digital signals. They have no parts and cannot be decomposed, they are simply units. However, several signals, when viewed together, gives rise to a signal process. Signals usually do not occur isolated but rather organized in processes. Mostly, such processes are ruled by probabilities, i.e., they are stochastic processes. These occur in two ways: with fixed probabilities (e.g., Markov processes) and with unknown probabilities (Hidden stochastic processes, HSM). We concentrate mainly on the latter ones because this is what we get from the applications. Signal processes contain information. However, this information is not directly visible and is not directly understandable to humans. The information is hidden in properties of the processes. Properties of signal processes are in the initial phase mostly unknown, and we have to find them out, at least approximately. For this, one does not only consider signal processes as they are given, one has rather to manipulate the processes using computations. It is a major challenge to do this properly and depends very much on the type of signals and the application domain.

We view our methods as organized in a machine learning process. To perform this task, one has to deploy many examples of signals. The deployed signals undergo several transformations as necessary. A major difficulty is to find out which transformation one should apply in which situation.

From the hardware point of view, a digital signal processor is a specialized microprocessor with its architecture optimized for the operational needs of digital signal processing. However, this is not the main topic of the book.

Machine learning is a wide area. In this book, we discuss aspects of it related to signal processes. These are mainly unsupervised learning methods.

The properties of signals lead to abstractions. On a more abstract level, one can define properties interesting to applications. In evolution, living organisms have

developed such levels and abstractions over time. Of course, one cannot repeat the evolution. However, it can still be studied because many of the primitive organisms are still alive today. Here, we do not consider evolution, but we can still make use of the observations and results.

The problem is that even a perfect solution to a sub problem is not very useful if it does not help with solutions to other sub problems. The reason is that the sub problems are not contained in modules that can be treated in isolation. Each solution of a sub problem is affected to a degree by the system, or as a whole part.

In a book, one faces therefore the following difficulty:

- How to formulate the overall perspective?
- How to combine solutions to the sub problem in the view of the overall perspective?

The readers should have understanding of basic computational methods as they are taught in the first three semesters of computer science.

Some History The origins of signal processing go back to previous centuries, around the seventeenth century. At that time, it was considered as a part of mathematics, in particular, numerical analysis. In modern times (after 1970) several digital processors have been developed, like in Lincoln Laboratories. In the 1980s, the first chips were developed as Intel 2920, TMS32010 from Texas Instrument, and NEC μPD7720. In each chapter, there is a section on background information. It contains mainly historical remarks and references for further reading. Here, the discussed choice was the "Dynamic Automatic Noisy Speech Recognition System" (DANSR). The system was devoted to tackle the problem of noises of different kinds that occur simultaneously. This is a major task in the dissertation of the author Sheuli Paul.

A Guide to Different Chapters In each chapter, one finds illustrations of different kinds. The illustrations typically are figures. They may be derived from the peak period when the machine was functional, and also when the machine is working at standard level. Sometimes, illustration is not understood by itself. Therefore, we added some explanations in the figure, for instance in terms of colors. There will be many such figures in the book. For most illustrations, MatLab was used.

Applications are an important part of the book. They are given in the second half of the book. Part C contains the description of several real situations where stochastic processes play a role. In real situations, there is always noise disturbing the stochastic process. The noise can be of a different level. Based on the noise level, we termed the noise as mild, strong, and steady-unsteady. Therefore, we see the probability density function (pdf) of such noisy commands, classified in the legend as mild, unsteady, and strong.

The structure of all chapters is essentially the same. They are structured as

- Overview
- Main text
- Background information

- Exercises
- Reference

We comment shortly on the just-mentioned structure. The overview contains indications of what the reader should study for better understanding before reading the chapter. In addition, this allows finding topics of interest quite fast and it helps a lecturer to select topics for a discussion in class.

The main text contains the contribution of the chapter. In the background, we give the origins to the discussed topics as well as suggestion for further reading. Here, we also present the links to the references. In the main text, there are no references. Please observe that each chapter contains its own list of references. This means that references are listed at the end of each chapter, for the related work discussed in the corresponding chapter text.

The exercises do not contain solutions so that a lecturer can use them in class. In a short way, we now comment on each chapter. This is intended for the readers who want to see what is close to their special interests.

Part I: Realms of Signal Processing

The first part consists of seven chapters, and provides a survey of mathematical and engineering foundations assumed to be known by the reader.

Chapter 1: Digital Signal Representation The details of digital representations of discrete-time signals are presented in this chapter. This includes bridging the gap between the abstract discrete-time signal notation presented in the book, $x[n]$, and its representation in a digital processor.

Chapter 2: Signal Processing Background This chapter introduces briefly the signal processing discipline. It contains the problems and the principal solution methods. In addition, some history is given as well as a short description for the intended readers.

Chapter 3: Fundamentals of Signal Transformation This chapter gives an overview of processing and transforming signals. Here, some basic signal processing formulations are introduced. It is said how the signal processing is an integral part of machine learning and how these two areas are related and how the related outcome of these two areas generates a successful application. In this chapter, we deal with the decomposition of a signal and techniques used. The signal is decomposed into some basic functions and the coefficients. Furthermore, the coefficients then can be used as linear combinations of the signals representing the signal. Transformations such as the z-transform, Fourier transform, discrete cosine transform, local trigonometric transform, and wavelet transform are briefly discussed.

Chapter 4: Digital Filters Digital filters are some mathematical operations to manipulate or analyze the signals. The filter coefficients are generally not variable like in adaptive filters. We introduce the chapter with fundamentals of finite impulse response (FIR) filter and infinite impulse response (IIR) filter, and discuss the applications of some advanced filters.

Chapter 5: Estimation and Detection This chapter discusses some commonly used filters, methods, and techniques to analyze the signal adaptively. The filters use optimization approach in order to analyze the signals. In this chapter, the optimization techniques with examples are discussed. Such techniques are minimum variance unbiased estimation (MVU), best linear unbiased estimator (BLUE), maximum likelihood estimator (MLE), least squares estimator (LSE), minimum least squares estimator (LMSE), and the Bayesian approach.

Chapter 6: Adaptive Signal Processing This chapter has general information about signal modeling using parametric and non-parametric types. It introduces adaptive filters mainly Wiener filter, Kalman filter, and Particle filter.

Chapter 7: Spectral Analysis This chapter considers the problem of the spectral content of the signals by means of parametric and non-parametric signal modeling. In a parametric signal modeling, a signal is analyzed using only few parameters. A non-parametric method has no such parameters quantification. Techniques such as linear prediction analysis using auto-regressive (AR) signal modeling, moving average signal modeling (MA), and auto regressive moving average (ARMA) signal modeling are introduced. Non-parametric signal modeling such as subspace signal analysis applying sub-band coding, discrete cosine transform, and discrete Fourier transform for spectral signal analysis are outlined.

Part II: Machine Learning and Recognition

The ideas of learning with a particular insight into the influence of the learning goals are discussed in eight chapters in this part.

Chapter 8: General Learning This chapter includes the general aspects and principles of learning. An important aspect is the improvement of the learner after learning. In this part, we emphasize on aspects related to stochastic processes as applied to learning.

Chapter 9: Signal Processes, Learning, and Recognition Learning is a very wide and complex area. We cannot give a complete introduction into the subject. Therefore, we restrict ourselves to aspects and methods related to signal processes and signal recognition. The learning considers in particular hidden stochastic

models. Besides a general introduction into the topic, the main topic is Bayesian Learning. Fundamentals of machine learning and its role for signal processing are briefed here.

Chapter 10: Stochastic Processes Here, we introduce the readers to signal processes, in particular to stochastic signal processes. These are partially ordered sets of signals. In the stochastic situation, there is no deterministic rule governing the process. This is rather of statistical character. Here, the main topic is hidden stochastic model. Their main characteristic is that the involved probabilities are unknown.

Chapter 11: Feature Extraction Features are short vectors describing processes. They have to be extracted. This requires techniques described in this chapter. Basic techniques include spectral shaping, windowing the signals, pre-emphasis, folding operators, unfolding, and spectral analysis and its envelope extraction. These techniques are again used in discussed methods such as linear prediction, search, mean squared error minimization, the Baum–Welch algorithm, and the Viterbi algorithm.

Chapter 12: Unsupervised Learning Unsupervised learning techniques such as clustering are introduced here. They are related to learning methods as applied to signal processes. Included in this discussion are K-Means, VQ, EM, and GMM.

Chapter 13: Markov Model and Hidden Stochastic Model (HSM) The topics such as Gaussian mixture models (GMM), the autocorrelation, the Yule-Walker, the co-variance, and the unconstrained least squared (ULS) approach are discussed in this chapter.

Chapter 14: Fuzzy Logic and Rough Sets When dealing with signals, uncertainty is unavoidable. One general method is to use fuzzy sets. In order to obtain decisions that are mostly of binary character, rough sets are introduced. We also discuss fuzzy clustering and fuzzy probabilities.

Chapter 15: Neural Networks This chapter introduces neural networks. In addition to the basics, it includes feed forward networks. Simulation of how the human brain processes the information (biology) is presented briefly, as well as the use of neural networks in machine learning (computers).

Part III: Real Aspects and Applications

Here, the reader is informed about many topics that occur in applications. It contains the required extensions or modifications of techniques that have been introduced in earlier parts.

Chapter 16: Noisy Signals Noise can corrupt the message contained in a signal to a large extent. We discuss different kinds and properties of noise. This is strongly connected to some measurements criteria. Signal-to-noise relation is a basic noise property. This is measured by the average relation of the noise signals to the original signal. There are different kind of noises that require individually tailored techniques.

Chapter 17: Reasoning Methods and Noise Removal Often, an aim is a noise-free signal, thus a main task here is to minimize or remove the noises. An important part of this is noise removal. Noise removal depends very much on noise types. We restrict ourselves to the noise types introduced so far. Each type requires different removal techniques. A major problem occurs when different types of noises occur simultaneously, which requires appropriate removal techniques.

Chapter 18: Audio Signals and Speech Recognition The handling of audio signals has motivated many progresses in signal processing. Speech processes have specific aspects. These are introduced in some details in this chapter.

Chapter 19: Noisy Speech Speech is a major example where noise plays a role. Noise disturbs the reception of speech independent of the type of receiver, which can be a human being or a machine. If a machine is the receiver, some response is expected.

Chapter 20: Aspects of Human Hearing A main characteristic of hearing is its temporal character. This is quite different from dealing with images. The topic is studied because one can obtain elements of recognition of machines from looking at the human ear.

Chapter 21: Speech Features The speech features contain a short and compact representation of the speech. The main condition is that no important information is lost. The term "important" differs from situation to situation and therefore there are several feature extraction algorithms.

Chapter 22: Hidden Stochastic Model for Speech Speech is a typical example for hidden stochastic models. If a person speaks the same words twice, the uttered signals are not the same. The probabilities over the speech elements are unknown and they need to be determent. This chapter delves into mathematical techniques used to model those processes.

Chapter 23: Different Speech Applications—Part A Here, one finds some typical applications. The applications have in common that the speech receiver is a machine. This machine is supposed to execute a command that is contained in the speech. Therefore, the machine has to be embedded in another complex structure that is responsible for the execution as discussed here.

Chapter 24: Different Speech Applications—Part B Different aspects of machine learning of a few practical examples such as Wake-Up-Word (WUW) are introduced. In addition, the Speech Analysis and Sound Effects Laboratory (SASE Lab), Wake-Up-Word: Tool Demo, and Elevator Simulator are presented in this chapter. The aim is to give an essence of understanding, namely how humans communicate using spoken language.

Chapter 25: Biomedical Signals: ECG, EEG These signals are generated by the human body and contain hidden information about it. A major task is to extract such information and to derive the information they contain. There two types of signals that are studied:

- ECG signals
- EEG signals

Chapter 26: Seismic Signals The source of seismic signals is an event that occurs inside the Earth. They contain important information; we can predict the future behavior of the Earth using digital signal processing techniques. A major aspect is to use seismic signals for prediction of earthquakes.

Chapter 27: Radar Signals Radar is an electromagnetic system that detects and locates objects using electromagnetic waves. Transmitter, receiver, and target are basic components of a radar. Radar is categorized in different ways based on areas of application, system types, wave-forms, frequency bands, and antenna types. This chapter is an introduction to radar sensing attached with an application.

Chapter 28: Visual Story Telling is a visual perception of the world. An informed decision about the situation can be made through an effective communication following the perception. Understanding the context, selecting the right tool to display, right selection of data visualization by removing redundancy, orientation, and focus are important for data visualization. This chapter provides essence of VST.

Chapter 29: Digital Processes and Multimedia At the low level, one presents multimedia by using signals. Traditionally, they can be one, two, or three dimensional. Images are a specific example of this, where the signal corresponds to pixels. A major task is to transform signal processes into multimedia representations. Examples of multimedia, techniques, and principles can be seen in this chapter.

Chapter 30: Visualizations for Emergency Operation Centre A visual story is told to manage an emergency environment. This chapter is based on actual emergency situations. The importance of visualization in emergency situations is outlined here. How the visualizations and visual analytics are applied in a typical emergency situation management center, such as an emergency operation center (EOC), is briefly introduced in this chapter.

Chapter 31: Intelligent Interactive Communications Multi-modal communication combines multiple data sources for an effective and meaningful interaction. A concept of multi-modal intelligent interactive machine is proposed in this chapter. This applies automation capabilities to communicate interactively applying verbal and non-verbal multimodalities of human-to-human and human-to-machine communication. The modes of communication can be speech, text, gesture, images, graphics, and visualization.

Chapter 32: Comparisons This chapter presents the main distinctions of the presented chapters.

Contents

Part I Realms of Signal Processing

1 Digital Signal Representation 3
1.1 Introduction 3
1.2 Numbers 4
1.2.1 Numbers and Numerals 5
1.2.2 Types of Numbers 7
1.2.3 Positional Number Systems 8
1.3 Sampling and Reconstruction of Signals 9
1.3.1 Scalar Quantization 11
1.3.2 Quantization Noise 15
1.3.3 Signal-to-Noise Ratio 19
1.3.4 Transmission Rate 22
1.3.5 Nonuniform Quantizer 23
1.3.6 Companding 24
1.4 Data Representations 24
1.4.1 Fixed-Point Number Representations 25
1.4.2 Sign-Magnitude Format 26
1.4.3 One's-Complement Format 27
1.4.4 Two's-Complement Format 28
1.5 Fix-Point DSP's 29
1.6 Fixed-Point Representations Based on Radix-Point 31
1.7 Dynamic Range 34
1.8 Precision 34
1.9 Background Information 35
1.10 Exercises 38
References 38

2 Signal Processing Background 39
2.1 Basic Concepts 39
2.2 Signals and Information 41
2.3 Signal Processing 41

2.4 Discrete Signal Representations 42
2.5 Delta and Impulse Function 43
2.6 Parseval's Theorem 45
2.7 Gibbs Phenomenon 47
2.8 Wold Decomposition 47
2.9 State Space Signal Processing 49
2.10 Common Measurements 52
2.10.1 Convolution 52
2.10.2 Correlation 53
2.10.3 Auto Covariance 54
2.10.4 Coherence 55
2.10.5 Power Spectral Density (PSD) 56
2.10.6 Estimation and Detection 59
2.10.7 Central Limit Theorem 60
2.10.8 Signal Information Processing Types 61
2.10.9 Machine Learning 61
2.10.10 Exercises 65
References 66

3 Fundamentals of Signal Transformations 69
3.1 Transformation Methods 71
3.1.1 Laplace Transform 72
3.1.2 Z-Transform 73
3.1.3 Fourier Series 76
3.1.4 Fourier Transform 76
3.1.5 Discrete Fourier Transform and Fast Fourier Transform 77
3.1.6 Zero Padding 78
3.1.7 Overlap-Add and Overlap-Save Convolution Algorithms 78
3.1.8 Short Time Fourier Transform (STFT) 79
3.1.9 Wavelet Transform 88
3.1.10 Windowing Signal and the DCT Transforms 91
3.2 Analysis and Comparison of Transformations 92
3.3 Background Information 93
3.4 Exercises 94
References 95

4 Digital Filters 97
4.1 Introduction 97
4.1.1 FIR and IIR Filters 99
4.1.2 Bilinear Transform 104
4.2 Windowing for Filtering 104
4.3 Allpass Filters 106
4.4 Lattice Filters 107
4.5 All-Zero Lattice Filter 111

4.6 Lattice Ladder Filters 111
4.7 Comb Filter 111
4.8 Notch Filter 112
4.9 Background Information 114
4.10 Exercises 114
References 115

5 Estimation and Detection 117
5.1 Introduction 117
5.2 Hypothesis Testing 118
5.2.1 Bayesian Hypothesis Testing 118
5.2.2 MAP Hypothesis Testing 119
5.3 Maximum Likelihood (ML) Hypothesis Testing 120
5.4 Standard Estimation Techniques 120
5.4.1 Minimum Variance Unbiased (MVU) Estimator 122
5.4.2 Best Linear Unbiased Estimator (BLUE) 123
5.4.3 Maximum Likelihood Estimator (MLE) 124
5.4.4 Least Squares Estimator (LSE) 125
5.4.5 Linear Minimum Mean Square Error Estimator (LMMSE) 128
5.5 Exercises 130
References 130

6 Adaptive Signal Processing 131
6.1 Introduction 131
6.2 Parametric Signal Modeling 132
6.2.1 Parametric Estimation 132
6.3 Wiener Filtering 132
6.4 Kalman Filter 135
6.4.1 Smoothing 136
6.5 Particle Filter 137
6.6 Fundamentals of Monte Carlo 139
6.6.1 Importance Sampling (IS) 140
6.7 Non-Parametric Signal Modeling 143
6.8 Non-Parametric Estimation 143
6.8.1 Correlogram 144
6.8.2 Periodogram 144
6.9 Filter Bank Method 145
6.10 Quadrature Mirror Filter Bank (QMF) 148
6.11 Background Information 149
6.12 Exercises 150
References 150

7 Spectral Analysis 151
7.1 Introduction 151
7.2 Adaptive Spectral Analysis 152

7.3 Multivariate Signal Processing 152
7.3.1 Sub-Band Coding and Subspace Analysis 155
7.4 Wavelet Analysis 156
7.5 Adaptive Beam Forming 159
7.6 Independent Component Analysis (ICA) 162
7.7 Principal Component Analysis (PCA) 164
7.8 Best Basis Algorithms 166
7.9 Background Information 166
7.10 Exercises 167
References 167

Part II Machine Learning and Recognition

8 General Learning 173
8.1 Introduction to Learning 173
8.2 The Learning Phases 175
8.2.1 Search and Utility 175
8.3 Search 176
8.3.1 General Search Model 176
8.3.2 Preference Relations 177
8.3.3 Different Learning Methods 178
8.3.4 Similarities 179
8.3.5 Learning to Recognize 179
8.3.6 Learning Again 180
8.4 Background Information 181
8.5 Exercises 182
References 182

9 Signal Processes, Learning, and Recognition 183
9.1 Learning 183
9.2 Bayesian Formalism 186
9.2.1 Dynamic Bayesian Theory 186
9.2.2 Recognition and Search 186
9.2.3 Influences 188
9.3 Subjectivity 189
9.4 Background Information 189
9.5 Exercises 191
References 191

10 Stochastic Processes 193
10.1 Preliminaries on Probabilities 194
10.2 Basic Concepts of Stochastic Processes 194
10.2.1 Markov Processes 195
10.2.2 Hidden Stochastic Models (HSM) 196
10.2.3 HSM Topology 198
10.2.4 Learning Probabilities 198
10.2.5 Re-estimation 199

10.2.6 Redundancy 199
10.2.7 Data Preparation 200
10.2.8 Proper Redundancy Removal 200
10.3 Envelope Detection 201
10.3.1 Silence Threshold Selection 203
10.3.2 Pre-emphasis 203
10.4 Several Processes 206
10.4.1 Similarity 207
10.4.2 The Local-Global Principle 209
10.4.3 HSM Similarities 214
10.5 Conflict and Support 215
10.6 Examples and Applications 215
10.7 Predictions 216
10.8 Background Information 217
10.9 Exercises 218
References 219

11 Feature Extraction 221
11.1 Feature Extractions 222
11.2 Basic Techniques 223
11.2.1 Spectral Shaping 223
11.3 Spectral Analysis and Feature Transformation 229
11.3.1 Parametric Feature Transformations and Cepstrum 229
11.3.2 Standard Feature Extraction Techniques 230
11.3.3 Frame Energy 230
11.4 Linear Prediction Coefficients (LPC) 231
11.5 Linear Prediction Cepstral Coefficients (LPCC) 234
11.6 Adaptive Perceptual Local Trigonometric Transformation (APLTT) 235
11.7 Search 235
11.7.1 General Search Model 235
11.8 Predictions 237
11.8.1 Purpose 237
11.8.2 Linear Prediction 237
11.8.3 Mean Squared Error Minimization 239
11.8.4 Computation of Probability of an Observation Sequence 240
11.8.5 Forward and Backward Prediction 242
11.8.6 Forward-Backward Prediction 244
11.9 Background Information 248
11.10 Exercises 249
References 250

12 Unsupervised Learning 251
12.1 Generalities 251
12.2 Clustering Principles 252

12.3 Cluster Analysis Methods 253
12.4 Special Methods 254
12.4.1 K-Means 255
12.4.2 Vector Quantization (VQ) 256
12.4.3 Expectation Maximization (EM) 257
12.4.4 GMM Clustering 258
12.5 Background Information 259
12.6 Exercises 260
References 260

13 Markov Model and Hidden Stochastic Model 261
13.1 Markov Process 261
13.2 Gaussian Mixture Model (GMM) 262
13.3 Advantages of Using GMM 263
13.4 Linear Prediction Analysis 263
13.4.1 Autocorrelation Method 267
13.4.2 Yule-Walker Approach 268
13.4.3 Covariance Method 269
13.4.4 Comparison of Correlation and Covariance Methods 270
13.5 The ULS Approach 271
13.6 Comparison of ULS and Covariance Methods 273
13.7 Forward Prediction 274
13.8 Backward Prediction 275
13.9 Forward-Backward Prediction 276
13.10 Baum-Welch Algorithm 276
13.11 Viterbi Algorithm 277
13.12 Background Information 277
13.13 Exercises 278
References 278

14 Fuzzy Logic and Rough Sets 281
14.1 Rough Sets 281
14.2 Fuzzy Sets 283
14.2.1 Basis Elements 283
14.2.2 Possibility and Necessity 285
14.3 Fuzzy Clustering 286
14.4 Fuzzy Probabilities 287
14.5 Background Information 288
14.6 Exercises 288
References 289

15 Neural Networks 291
15.1 Neural Network Types 292
15.1.1 Neural Network Training 293
15.1.2 Neural Network Topology 294

15.2 Parallel Distributed Processing 297
15.2.1 Forward and Backward Uses 298
15.2.2 Learning 299
15.3 Applications to Signal Processing 299
15.4 Background Information 300
15.5 Exercises 301
References 301

Part III Real Aspects and Applications

16 Noisy Signals 307
16.1 Introduction 307
16.2 Noise Questions 311
16.3 Sources of Noise 311
16.4 Noise Measurement 311
16.5 Weights and A-Weights 312
16.6 Signal to Noise Ratio (SNR) 315
16.7 Noise Measuring Filters and Evaluation 317
16.8 Types of Noise 317
16.9 Origin of Noises 320
16.10 Box Plot Evaluation 320
16.11 Individual Noise Types 321
16.11.1 Residual 321
16.11.2 Mild 321
16.11.3 Steady-Unsteady Time Varying Noise 322
16.11.4 Strong Noise 323
16.12 Solution to Strong Noise: Matched Filter 324
16.13 Background Information 325
16.14 Exercises 326
References 326

17 Reasoning Methods and Noise Removal 327
17.1 Generalities 327
17.2 Special Noise Removal Methods 328
17.2.1 Residual Noise 328
17.2.2 Mild Noise 328
17.2.3 Steady-Unsteady Noise 330
17.2.4 Strong Noise 330
17.3 Poisson Distribution 331
17.3.1 Outliers and Shots 331
17.3.2 Underlying Probability of Shots 332
17.4 Kalman Filter 332
17.4.1 Prediction Estimates 335
17.4.2 White Noise Kalman Filtering 335
17.4.3 Application of Kalman Filter 336

17.5 Classification, Recognition and Learning 337
17.5.1 Summary of the Used Concepts 338
17.6 Principle Component Analysis (PCA) 338
17.7 Reasoning Methods 340
17.7.1 Case-Based Reasoning (CBR) 341
17.8 Background Information 341
17.9 Exercises 342
References 342

18 Audio Signals and Speech Recognition 345
18.1 Generalities of Speech 345
18.2 Categories of Speech Recognition 347
18.3 Automatic Speech Recognition 349
18.3.1 System Structure 350
18.4 Speech Production Model 350
18.5 Acoustics 351
18.6 Human Speech Production 351
18.6.1 The Human Speech Generation 352
18.6.2 Excitation 353
18.6.3 Voiced Speech 355
18.6.4 Unvoiced Speech 357
18.7 Silence Regions 359
18.8 Glottis 360
18.9 Lips 360
18.10 Plosive Speech Source 361
18.11 Vocal-Tract 361
18.12 Parametric and Non-parametric Models 363
18.13 Formants 366
18.14 Strong Noise 366
18.15 Background Information 367
18.16 Exercises 367
References 367

19 Noisy Speech 369
19.1 Introduction 369
19.2 Colored Noise 369
19.2.1 Additional Types of Colored Noise 371
19.3 Poisson Processes and Shots 372
19.4 Matched Filters 374
19.5 Shot Noise 375
19.6 Background Information 377
19.7 Exercises 378
References 378

20 Aspects of Human Hearing 379
20.1 Human Ear 380
20.2 Human Auditory System 380

20.3 Critical Bands and Scales 382
20.3.1 Mel Scale 382
20.3.2 Bark Scale 383
20.3.3 Erb Scale 384
20.3.4 Greenwood Scale 384
20.4 Filter Banks 385
20.4.1 ICA Network 386
20.4.2 Auditory Filter Banks 387
20.4.3 Filter Banks 387
20.4.4 Mel Critical Filter Bank 388
20.5 Psycho-Acoustic Phenomena 389
20.5.1 Perceptual Measurement 389
20.5.2 Human Hearing and Perception 389
20.5.3 Sound Pressure Level (SPL) 390
20.5.4 Absolute Threshold of Hearing (ATH) 391
20.6 Perceptual Adaptation 392
20.7 Auditory System and Hearing Model 392
20.8 Auditory Masking and Masking Frequency 393
20.9 Perceptual Spectral Features 394
20.10 Critical Band Analysis 394
20.11 Equal Loudness Pre-emphasis 395
20.12 Perceptual Transformation 396
20.13 Feature Transformation 396
20.14 Filters and Human Ear 396
20.15 Temporal Aspects 397
20.16 Background Information 398
20.17 Exercises 399
References 399

21 Speech Features 401
21.1 Generalities 401
21.2 Cost Functions 402
21.3 Special Feature Extractions 403
21.3.1 MFCC Features 403
21.3.2 Feature Transformation Applying DCT 407
21.4 Background Information 412
21.5 Exercises 413
References 413

22 Hidden Stochastic Model for Speech 415
22.1 Generals 415
22.2 Hidden Stochastic Model 417
22.3 Forward and Backward Predictions 421
22.3.1 Forward Algorithm 421
22.3.2 Backward Algorithm 422
22.4 Forward-Backward Prediction 423

22.5 Burg Approach 424
22.6 Graph Search 426
22.6.1 Recognition Model with Search 427
22.7 Semantic Issues and Industrial Applications 428
22.8 Problems with Noise 430
22.9 Aspects of Music 430
22.10 Music Reception 433
22.11 Background Information 433
22.12 Exercises 434
References 434

23 Different Speech Applications—Part A 437
23.1 Generalities 437
23.2 Example Applications 437
23.2.1 Experimental Laboratory 439
23.2.2 Health Care Support (Everyday Actions) 441
23.2.3 Diagnostic Support for Persons with Possible Dementia 444
23.2.4 Noise 446
23.3 Background Information 447
23.4 Exercises 449
References 449

24 Different Speech Applications—Part B 451
24.1 Introduction 451
24.2 Discrete-Time Signals 454
24.3 Speech Processing 456
24.3.1 Framing 457
24.3.2 Pre-emphasis 458
24.3.3 Windowing 460
24.3.4 Fourier Transform 462
24.3.5 Mel-Filtering 465
24.3.6 Mel-Frequency Cepstral Coefficients 467
24.4 Speech Analysis and Sound Effects Laboratory (SASE_Lab) 470
24.5 Wake-Up-Word Speech Recognition 475
24.5.1 Introduction 476
24.5.2 Wake-Up-Word Paradigm 476
24.5.3 Wake-Up-Word: Definition 478
24.5.4 Wake-Up-Word System 479
24.5.5 Front-End of the Wake-Up-Word System 480
24.6 Conclusion 492
24.6.1 Wake-Up-Word: Tool Demo 493
24.6.2 Elevator Simulator 494
24.7 Background Information 494
24.8 Exercises 496
24.9 Speech Analysis and Sound Effects Laboratory (SASE_Lab)" 497
References 497

25 Biomedical Signals: ECG, EEG 499
25.1 ECG Signals 500
25.1.1 Bioelectric Signals 500
25.1.2 Noise 502
25.2 EEG Signals 503
25.2.1 General Properties 503
25.2.2 Signal Types and Properties 504
25.2.3 Disadvantages 504
25.3 Neural Network Use 505
25.4 Major Research Questions 506
25.5 Background Information 506
25.6 Exercises 507
References 508

26 Seismic Signal 509
26.1 Generalities 510
26.2 Sources of Seismic Signals 510
26.3 Intermediate Elements 511
26.4 Practical Data Sources 511
26.5 Major Seismic Problems 512
26.6 Noise 513
26.7 Background Information 514
26.8 Exercises 514
References 515

27 Radar Signals 517
27.1 Introduction 517
27.2 Radar Types and Applications 519
27.3 Doppler Equations, Ambiguity Function (AF) and Matched Filter 520
27.4 Moving Target Detection 521
27.5 Applications and Discussions 523
27.6 Examples 523
27.7 Background Information 524
27.8 Exercises 526
References 529

28 Visual Story Telling 531
28.1 Introduction 531
28.1.1 Common Visualization Approaches 533
28.2 Analytics and Visualization 533
28.2.1 Exploratory Data Analysis 534
28.2.2 Visual Data Mining 535
28.3 Communication and Visualization 535
28.4 Background Information 536

28.5 Exercises 539
References 539

29 Digital Processes and Multimedia 541
29.1 Images 541
29.1.1 Digital Image Processing 542
29.1.2 Images as Matrices 549
29.1.3 Gray Scale Images 550
29.2 Spatial Filtering 552
29.2.1 Linear Filtering of Images 552
29.2.2 Separable Filters 553
29.2.3 Mechanics of Linear Spatial Filtering Operation 555
29.3 Median Filtering 556
29.4 Color Equalization 557
29.4.1 Image Transformations 559
29.4.2 Examples of Image Transformation Matrixes 564
29.5 Basic Image Statistics 567
29.6 Abstraction Levels of Images and Its Representations 569
29.6.1 Lowest Level 570
29.6.2 Geometric Level 571
29.6.3 Domain Level 571
29.6.4 Segmentation 572
29.7 Background Information 572
29.8 Exercises 573
References 574

30 Visualizations of Emergency Operation Centre 575
30.1 Introduction 575
30.2 Communications in Emergency Situations 576
30.3 Emergency Scenario 577
30.3.1 Classification and EOC Scenario 577
30.4 Technical Aspects and Techniques 578
30.4.1 Classification 578
30.4.2 Clustering 579
30.5 Background Information 579
30.6 Exercises 580
References 581

31 Intelligent Interactive Communications 583
31.1 Introduction 583
31.2 Spoken Dialogue System 585
31.3 Gesture Based Interaction 588
31.4 Object Recognition and Identification 588
31.5 Visual Story Telling 589
31.6 Virtual Environment for Personal Assistance 591
31.7 Sensor Fusion 592

31.8 Intelligent Human Machine for Communication: Application Scenario 593
31.9 Background Information 593
31.10 Exercises 595
References 595

32 Comparisons 597
32.1 Generalities 597
32.1.1 EEG and ECG 600
32.1.2 Speech and Biomedical Applications 601
32.1.3 Seismic and Biomedical Signals 601
32.1.4 Speech and Images 601
32.2 Overall 602
32.3 Background Information 603
32.3.1 General 603
32.4 Exercises 603

Glossary 605

List of Figures

Fig. 1.1 The interfacing with a continuous signal that is optionally conditioned by the sensor conditioner. Analog-to-Digital Conversion: (**a**) Continuous signal $x(t)$, (**b**) Sampled signal $x_a(t)$ with sampling period T satisfying Nyquist rate as specified by Sampling Theorem, (**c**) Digital sequence $x[n]$ obtained after sampling and quantization 10

Fig. 1.2 Digital-to-analog conversion. (**a**) Processed digital signal y[n]. (**b**) Continuous signal representation ya(nT). (**c**) Low-pass filtered continuous signal y(t) 11

Fig. 1.3 Conceptual representation of ADC 12

Fig. 1.4 Example of a uniform quantization for $\boldsymbol{L = 16}$ levels. As it is discussed in the following sections, $\boldsymbol{L = 16}$ levels, require 4 bits of the codeword to represent each level. Because the distribution of the input values is uniform the decision and reconstruction levels are uniformly spaced: $\hat{x} = Q(x)$ 13

Fig. 1.5 Example of pdf's (probability distribution function) 15

Fig. 1.6 Subplots are enumerated from top to bottom 17

Fig. 1.7 Example of slowly varying signal that causes quantization error to be correlated 18

Fig. 1.8 Example of speech signal demonstrating the effect of step size to the degree of correlation of quantization error 19

Fig. 1.9 Histogram of quantization error for speech signal. Note the reduction of error magnitude as well as increase of uniformity of the distribution with increase of number of quantization levels 20

Fig. 1.10 Histogram of quantization error for speech signal after adding Gaussian noise with zero mean and variance of 5 21

Fig. 1.11 Uniform probability distribution function $p(e)$ of error signal in the range $[-\Delta/2, \Delta/2]$ 22

Fig. 1.12 Block diagram of companding in the transmitting and receiving DSP system. The $x[n]$ is an unquantized input sample with a nonuniform pdf of values; $y[n]$ is the value obtained after nonlinear transformation with uniform pdf values, and $\tilde{y}[n]$ is quantized sample, $\boldsymbol{c}[n]$ is encoded binary representation of this sample value. This binary encoded stream is typically transmitted to a receiving system where it is converted back to the original by decoding the input encoded samples $\boldsymbol{c'}[n]$ to $\boldsymbol{y'}[n]$, and applying the inverse of the non-linear transformation T^{-1} obtaining sequence $\boldsymbol{x'}[n]$. If $\boldsymbol{c'}[n]$ = $\boldsymbol{c}[n]$, then $\boldsymbol{x'}[n]$ differs from $\boldsymbol{x}[n]$ by amount of introduced quantization noise 25
Fig. 1.13 Signed magnitude format 26
Fig. 1.14 The 16-bit unsigned and signed data type's representation 31
Fig. 1.15 Possible $\mathbb{Q}_{p.q}$ format representations of 16-bit data length 36
Fig. 2.1 Signal is processed from continuous to discrete representation 42
Fig. 2.2 Delta function 43
Fig. 2.3 Rectangular waveform and Gibbs phenomenon applying successive approximations 40 sine waveforms 48
Fig. 2.4 Rectangular waveform and Gibbs' phenomenon approximating $n = 1, 3, 7, 19, 49$ and 70 sine wave-forms 49
Fig. 2.5 Continuous and discrete SISO system 50
Fig. 2.6 SISO and MIMO model analysis 50
Fig. 2.7 The state space representation of a system 52
Fig. 2.8 Input, system and output signals 56
Fig. 2.9 Standard error 61
Fig. 2.10 Time range 61
Fig. 2.11 Exploring problems in machine learning 64
Fig. 3.1 The 64 point Hanning window 80
Fig. 3.2 The 64 point Hamming window 81
Fig. 3.3 The 64 point Kaizer window with parameter $\alpha = 7$ 81
Fig. 3.4 The 64 point Chebyshev window with attenuation parameter 100..... 82
Fig. 3.5 Rising cut-off function 92
Fig. 3.6 Trapezoidal signal 94
Fig. 4.1 The figure shows an ideal low-pass filter specification. The F_{pass} depicts pass-band, F_s depicts the sampling period, while the other notation is self-explanatory 98
Fig. 4.2 The figure depicts the steps involved in periodic discrete-time sequence spectrum analysis using DFT 104

Fig. 4.3 A basic structure of one state allpass filter 107
Fig. 4.4 A two stage allpass filter 107
Fig. 4.5 A basic structure of allpass filter..... 108
Fig. 4.6 Transformation of a lattice filter and its transfer function from A(z) to D(z) and D(z) to A(z)..... 108
Fig. 4.7 M stage allpass lattice filter 108
Fig. 4.8 An M stage lattice filter 109
Fig. 4.9 Comb filter magnitude and frequency response..... 112
Fig. 4.10 Notch filter and its frequency response. 113
Fig. 5.1 Standard adaptive signal processing approach 118
Fig. 6.1 Stochastic signal processing 132
Fig. 6.2 A basic structure of two channel filter bank 145
Fig. 7.1 Multivariate signal processing 153
Fig. 7.2 Decimation and Interpolation 153
Fig. 7.3 Sub-band coding 155
Fig. 7.4 Basic Structure of a filter bank 160
Fig. 7.5 ICA separates the mixtures into its three different sources 165
Fig. 8.1 Learning 174
Fig. 8.2 The learning phases 175
Fig. 8.3 Simulated stock value 180
Fig. 9.1 Typical machine learning process 184
Fig. 9.2 General supervised learning 184
Fig. 9.3 Bayes rule for computing probabilities 187
Fig. 10.1 The figure shows 1st order magnitude response of low-pass and high-pass filters. (**a**) *High-Pass Filter*. (**b**) *Low-Pass Filter*..... 195
Fig. 10.2 Example of a Markov chain 196
Fig. 10.3 Envelop detection by LPF method and Hilbert transform 202
Fig. 10.4 The (**a**) original frequency representation and its (**b**) Hilbert Transform representation..... 203
Fig. 10.5 Example of redundancy removal 204
Fig. 10.6 Block diagram of a pre-emphasis Filter with a_{em} factor 205
Fig. 10.7 Pre-emphasis filter characteristics given by amplitude and phase response 205
Fig. 10.8 DTW frame depicting allowable paths 213
Fig. 11.1 Main processing steps before feature extraction and prediction 222
Fig. 11.2 Pre-filtering and pre-emphasized signal. (**a**) Redundancy removal speech. (**b**) Prefiltered redundancy removal speech. (**c**) Speech spectrum. (**d**) Prefiltered emphasized spectrum 224
Fig. 11.3 Frame of Speech with corresponding Spectrum, and its pre-emphasized versions 225
Fig. 11.4 Spectral shaping and analysis 226
Fig. 11.5 Windowing and folding in wavelet packet analysis 228

Fig. 11.6 Block diagram of procedure deployed computing cepstrum 229
Fig. 11.7 Example of 256 tap (points) Hanning window 232
Fig. 11.8 Example of 256 tap (points) Hamming window 233
Fig. 11.9 Block diagram depicting computation of LPC features 233
Fig. 11.10 LPCC feature extraction 234
Fig. 11.11 Viterbi decoding in speech pattern recognition 236
Fig. 11.12 Linear prediction analysis 238
Fig. 11.13 Viterbi Algorithm as applied to the network of the type shown 246
Fig. 11.14 Viterbi Algorithm as applied to text 247
Fig. 12.1 Unsupervised learning 252
Fig. 12.2 Results of application of K-means procedure 255
Fig. 12.3 Vector quantization 256
Fig. 13.1 Gaussian Mixture example 262
Fig. 13.2 Short-time prediction widow 265
Fig. 13.3 Covarience method applied to a frame of speech 271
Fig. 13.4 ULS based Prediction Coefficients and pole-zero locations of the filter and inverse filter 272
Fig. 13.5 Residual of the inverse ULS filter 273
Fig. 13.6 Frequency response of ULS and covariance approach 273
Fig. 14.1 Decisions using rough sets 282
Fig. 14.2 Visualisation of rough sets 282
Fig. 14.3 Fuzzy set with one dimensional data 286
Fig. 14.4 Representation of strict membership 287
Fig. 14.5 Sigmoid membership function 287
Fig. 15.1 An example of artificial neural network 292
Fig. 15.2 An example of an artificial neuron 296
Fig. 15.3 Example of activation function utilized by a neuron 297
Fig. 15.4 An example of feed-forward neural network 299
Fig. 16.1 Example of the waveform (e.g. speech Data), pitch (e.g., pitch data), and spectographic display (e.g. spectrogram) of clean spoken utterance ‘oeffne die tuer’ in a standard/ quiet environment 308
Fig. 16.2 Example of the waveform (e.g. speech data), pitch (e.g., pitch data), and spectographic display (e.g. spectrogram) of speech ‘oeffne die tuer’ recorded in an noisy industrial environment. Various types of noises were present due to heavy duty of machinery and running motors providing noisy environment. These machines were in **resting state** 308

Fig. 16.3 Example of the waveform (e.g. speech data), pitch (e.g., pitch data), and spectographic display (e.g. spectrogram) of speech 'oeffnedietuer' recorded in an noisy industrial environment. Various types of noises were present due to heavy duty of machinery and running motors providing noisy environment. These machines were in **running state** generating varying load sounds. Such industrial environment have always varying types of noisy sounds. In such environments, the noise can be so high that speaker can not distinguish the speech from noise 309
Fig. 16.4 Linear microphone array for signal processing in a typical noisy environment. Number of microphones m_i for $i = 1, \cdots, 8$ capturing the signal. Arrival times t_i for $i = 1, \cdots, 8$ of the signals in the array. Signal is captured at different times due to varying paths of arrivals of the wave pattern 310
Fig. 16.5 Example of the Noisy Waveform and its Spectrographic display 310
Fig. 16.6 Example of the A weighted filter measurement using varying noisy signal 314
Fig. 16.7 Example of the A weighted filter measurement of the strong noisy signal 315
Fig. 16.8 The A weighted filter measurement of different noises 316
Fig. 16.9 Hybrid noise 318
Fig. 16.10 Energy for the different noisy types 319
Fig. 16.11 Different noisy signal energies 319
Fig. 16.12 Boxplot chart of sound level measurement by A-weighting filter 321
Fig. 16.13 Boxplot chart of sound level measurement by A-weighting filter 323
Fig. 16.14 Noisy and enhanced signal spectrum 324
Fig. 16.15 Example of speech signal mixed with a strong noise and its handling 325
Fig. 17.1 Aliased-free down-sampler 328
Fig. 17.2 Flow diagram of the Kalman prediction 334
Fig. 17.3 Steady-unsteady nosy signal enhanced by Kalman filtering. (**a**) Noisy spoken spectrum: Oeffne die Tuer. (**b**) Enhanced spoken spectrum applying M-band Kalman filter Oeffne die Tuer 336
Fig. 17.4 Example of applied PCA to noisy signals. (**a**) Standard enhanced spectrum applying PCA: Oeffne die Tuer. (**b**) Special enhanced spectrum applying PCA: Oeffne die Tuer 340
Fig. 18.1 The speech cycle from human perspective 346
Fig. 18.2 Speech recognition by a machine 347

Fig. 18.3 General speech recognition system 348
Fig. 18.4 A more detailed speech recognition system 348
Fig. 18.5 Bayes rule in speech recognition system problem 349
Fig. 18.6 The human parts involved in speech production 353
Fig. 18.7 Basic US English sounds classification 354
Fig. 18.8 Delta functions 355
Fig. 18.9 Periodic pulses 356
Fig. 18.10 Source model of periodic pulses 356
Fig. 18.11 Voiced speech in the source excitation model 356
Fig. 18.12 Unvoiced speech in the source excitation model 358
Fig. 18.13 Voiced and unvoiced speech example 359
Fig. 18.14 Plosive sound generation in source excitation speech model 361
Fig. 18.15 Discrete Vocal Tract model of a number of uniform tubes 362
Fig. 18.16 Source excitation speech model 363
Fig. 18.17 Parametric speech source modeling using ARMA model 364
Fig. 18.18 Moving average and regressive processes 365
Fig. 19.1 Gaussian model for colored noise. (**a**) Colored noise. (**b**) Autocorrelation of colored noise. (**c**) Frequency information of colored noise 370
Fig. 19.2 Shaping filter where $y[n] = ay[n-1] + (1-a)x[n]$ 371
Fig. 19.3 Output of matched filtering 374
Fig. 19.4 Additive noise corrupted signal 375
Fig. 20.1 Figure depicting human ear. https://commons.wikimedia.org/w/index.php?curid=1678037 381
Fig. 20.2 Mel scale 383
Fig. 20.3 Bark scale 384
Fig. 20.4 Erb scale 385
Fig. 20.5 Comparison of different scales 386
Fig. 20.6 Threshold of hearing 392
Fig. 20.7 Filters in the human Ear 397
Fig. 21.1 MFCC feature extraction 404
Fig. 21.2 Example of pre-emphasised data computed from TIMIT corpus: TIMIT/TRAIN/DR1/FSJK1/SI696.WAV 405
Fig. 21.3 Example of FFT data computed from TIMIT corpus: TIMIT/TRAIN/DR1/FSJK1/SI696.WAV 405
Fig. 21.4 Triangles of different bandwidths following Mel scale 406
Fig. 21.5 Block diagram of cepstrum features generation 407
Fig. 21.6 Block diagram of PLP features generation 408
Fig. 21.7 Linear Prediction following the source excitation model 409
Fig. 21.8 Block diagram of LPC feature extraction 410
Fig. 21.9 Block diagram of LPCC feature extraction 410
Fig. 21.10 Block diagram of PLPCC feature extraction 411
Fig. 21.11 Block diagram of SILTT feature extraction 412
Fig. 22.1 Signal modeling 420

Fig. 22.2 An hypothetical example of a 3 state HMM 421
Fig. 22.3 Forward-backward prediction 423
Fig. 22.4 The Burg approach 424
Fig. 22.5 Visualization of forward and backward prediction 425
Fig. 22.6 Visualization of hypothetical incorrect classification 426
Fig. 22.7 HSM using hidden Markov modeling (HMM) for acoustic and language search in recognition 427
Fig. 22.8 HSM in DANSR model definition 427
Fig. 22.9 The DANSR system 428
Fig. 22.10 Overall view of ambient assistive living room for a nursing home 430
Fig. 22.11 Musical notation in a score 432
Fig. 23.1 Analysis component of health care support 443
Fig. 23.2 Probability distribution of DTW distance score using 100 training and 25 test data 447
Fig. 24.1 Search and decoding process is depicted in the figure above 452
Fig. 24.2 Figure depicting human anatomy and the organs involved in speech production. http://training.seer.cancer.gov/head-neck/anatomy/overview.html,https://commons.wikimedia.org/w/index.php?curid=1678037 453
Fig. 24.3 Figure depicting block diagram of speech recognition system 453
Fig. 24.4 Continuous and discrete time signal 455
Fig. 24.5 Continuous and discrete time signal with reduced sampling rate 456
Fig. 24.6 Continuous and discrete time signal sampled at the higher rate of 44.1 kHz 456
Fig. 24.7 Front-end stage of a typical Mel-filter Cepstral Coefficients based speech recognition system 457
Fig. 24.8 Speech signal with overlapping frames 458
Fig. 24.9 Pre-emphasis filter block diagram 458
Fig. 24.10 Pre-emphasis filter frequency response 459
Fig. 24.11 Original and pre-emphasised signal 459
Fig. 24.12 Hamming Widow of length 256 and its frequency characteristics depicted to the right 461
Fig. 24.13 Figure depicting: (**a**) Speech signal with overlaid Hamming Window. (**b**) Speech signal after the windowing operation 461
Fig. 24.14 Spectrum of a speech frame depicting its content in frequency 463
Fig. 24.15 Mel-frequency transformation curve 466
Fig. 24.16 Normalized triangular shaped filters following Mel-scale 467
Fig. 24.17 Linear scale vs Mel scale spectrum 468

Fig. 24.18 Example of a mel-frequency Cepstral Coefficients of order 12 of a frame of speech 469
Fig. 24.19 Image depicting the interfacing provided by the MATLAB tool SASE_Lab "Speech Analysis and Special Effects Laboratory" 471
Fig. 24.20 Waveform and Spectrograph analysis performed on the data 472
Fig. 24.21 3 seconds of data selected from the spectrogram depicted in red-color from the waveform 473
Fig. 24.22 3 seconds of data depicted with corresponding spectrogram computed from the selected section of waveform data 473
Fig. 24.23 Vowels, 'a', 'e', 'i', 'o', 'u' uttered and captured 474
Fig. 24.24 Vowels, 'a', 'i', 'e', 'u' 475
Fig. 24.25 Wake-up-word speech recognition system 480
Fig. 24.26 Speech signal with voice activity detector segmentation, spectrogram, and enhanced spectrogram generated by the front end module (Këpuska and Klein 2009) 481
Fig. 24.27 Support vector machine decision surface as computed with linear SVM 482
Fig. 24.28 Scoring of plain WUW recognition 484
Fig. 24.29 Score1 vs. Score2 example (Këpuska and Klein 2009)..... 485
Fig. 24.30 Score1 vs. Score3 example (Këpuska and Klein 2009) 486
Fig. 24.31 Score1 vs. Score2 linear discrimination surface (Këpuska and Klein 2009) 487
Fig. 24.32 Score1 vs. Score3 linear discrimination surface (Këpuska and Klein 2009) 487
Fig. 24.33 Score1 vs. Score2 vs. Score3 hyper-plane discrimination surface (Këpuska and Klein 2009) 488
Fig. 24.34 Distribution of Scores for triple-scoring and linear SVM (Këpuska and Klein 2009) 488
Fig. 24.35 Scatter plot of INV and OOV distributions of Scores1 vs Score2 with RBF discriminating function (Këpuska and Klein 2009) 489
Fig. 24.36 Scatter plot of INV and OOV distributions of Scores1 vs Score3 with RBF discriminating function (Këpuska and Klein 2009) 490
Fig. 24.37 Discriminating surface presented in 3D (Këpuska and Klein 2009) 491
Fig. 24.38 PDF of INV and OOV distributions for RBF (Këpuska and Klein 2009) 491
Fig. 24.39 PDF of INV and OOV distributions for RBF utilizing more data (Këpuska and Klein 2009)..... 492

Fig. 24.40 Example of "no Wake-up-Word"/some other word spoken, and the case when "Wake-up-Word" was spoken not triggering off and triggering the system as indicated with different color as well as audibly played sound "yes". (**a**) Wake-up-word demo "off". (**b**) Wake-up-word demo "on" 493
Fig. 24.41 Example of usage of WUW technology in elevator simulator. (**a**) Initial state of the elevator simulator. (**b**) Elevator simulator acquiring voice command. (**c**) Elevator simulator arriving at the destination floor 495
Fig. 24.42 Executing SASE_LAB from MATLAB 496
Fig. 25.1 Depiction of the typical process involving a patient, a machine, a physician 500
Fig. 25.2 Sample recording and power spectrum of EEG signal 505
Fig. 26.1 Signal produced by earthquake 512
Fig. 27.1 Basic components of a radar system {https://www.tutorialspoint.com/radar_systems/radar_systems_overview.htm} 518
Fig. 27.2 Monostatic radar and bistatic radar http://www.telecomabc.com/b/bistatic-radar.html 519
Fig. 27.3 Basic pulse Doppler radar overview {http://doi.org/Overviewofradars.S.Cruz-PolINEL6069} 520
Fig. 27.4 Moving target detection http://www.jpier.org/PIERM/pier.php?paper=16122501 522
Fig. 27.5 Adaptive spectral M-subband Kalman filter in passive radar signal analysis 524
Fig. 27.6 S-band signal in time domain 525
Fig. 27.7 (**a**) Direct and (**b**) Target signal in time and frequency domain ... 525
Fig. 27.8 Spectral analysis of S-band signal for passive radar 526
Fig. 27.9 Cross-correlation of direct and target signal 527
Fig. 27.10 Ambiguity function at zero delay 528
Fig. 27.11 Adaptive spectral radar signal analysis 528
Fig. 28.1 Image 28.1.: Covid-19 cases over time in Canada (https://covid-19-canada.uwo) 537
Fig. 28.2 Image 28.2.: Cumulative confirmed cases (https://covid-19-canada.uwo) 537
Fig. 28.3 Image 28.3.: Visualization of Covid-19 cases using exploratory data analysis (https://covid-19-canada.uwo) 538
Fig. 29.1 Human image perception 542
Fig. 29.2 Digital image representation 543
Fig. 29.3 Visible wavelengths 543
Fig. 29.4 Gray Scale representation (MATLAB) 544
Fig. 29.5 RGB color space 546
Fig. 29.6 RGB Color representation of a computer image 547
Fig. 29.7 Individual RGB colors of an image 548
Fig. 29.8 Individually depicted RGB colors of an image 548

Fig. 29.9 Gray scale image representation 551
Fig. 29.10 Gaussian Smoothing Kernel 554
Fig. 29.11 Discretization of an Image and Filter Kernel 554
Fig. 29.12 Filtering operation 555
Fig. 29.13 Noisy and its decomposed (RGB) Image 557
Fig. 29.14 Median filtered image 558
Fig. 29.15 Original and histogram Equalized Image 558
Fig. 29.16 Original Image and its Decomposed versions of Red, Green and Blue (RGB) channel 559
Fig. 29.17 Original Image and Decomposed Images versions of Red, Green and Blue channel presented in Gray scale 560
Fig. 29.18 Original and "Red" Channel Equalized Image 560
Fig. 29.19 Original and "Green" Channel Equalized Image 561
Fig. 29.20 Original and "Blue" Channel Equalized Image 561
Fig. 29.21 Image of "lena" and its transpose 562
Fig. 29.22 Image of "lena" and its flipped version 562
Fig. 29.23 Copped image from the original 563
Fig. 29.24 Example of mage "translation" 564
Fig. 29.25 Example of the resized image 565
Fig. 29.26 Rotation of a pixel around an arbitrary origin by an angle θ 566
Fig. 29.27 Discrete nature of a digital picture can create problems with some transformations if not handled properly 567
Fig. 29.28 Histogram of a image of 'lena' 568
Fig. 29.29 Abstraction levels of an Image 570
Fig. 29.30 Descriptive elements of an Image 570
Fig. 29.31 Sketch image of a house 571
Fig. 30.1 Risk visualization example in emergency situation (Eide 2012) 576
Fig. 31.1 Intelligent interactive machine for multimodal human-machine communication 584
Fig. 31.2 Multimodal autonomous interactive interface 585
Fig. 31.3 Interdisciplinary intelligent interactive multimodal communication 586
Fig. 31.4 SSDS components 587
Fig. 31.5 Chatbot for text based interaction using RASA https://rasa.com 587
Fig. 31.6 Gesture recognition for communication 588
Fig. 31.7 Object detection and recognition (Jain and Kasturi 1995) 589
Fig. 31.8 Object detection and recognition https://www.geeksforgeeks.org/object-detection-vs-object-recognition-vs-image-segmentation/ 589
Fig. 31.9 Visualization 590
Fig. 31.10 Visualization and prediction using visual story telling 591
Fig. 31.11 Abstract view of a VE system (Kim 2005) 592
Fig. 31.12 Application Scenario of intelligent interactive machine 594

List of Tables

Table 1.1 Example of 4-bit number representations 26
Table 1.2 Range of values represented by a 16 bit and 32 bit DSP's 30
Table 1.3 Example of 4 bit signed magnitude number operations with 4 bit resulting number 31
Table 1.4 Example of 4 bit signed magnitude number operations with 8 bit resulting number 32
Table 1.5 Example of 4 bit signed magnitude number operations with 4 bit resulting number 32
Table 1.6 Example of 4 bit signed magnitude number operations with 8 bit resulting number 32
Table 1.7 Dynamic range and precision of 16-bit signed and unsigned integer and fractional representations 35
Table 1.8 Maximal, minimal, and precision values for integer and fractional 16-bit fixed-point representations 37
Table 1.9 Maximal, minimal, and precision values for integer and fractional 16-bit signed fixed-point representations 37
Table 16.1 A-weights in dB in frequency range from 63 to 2000 Hz 313
Table 18.1 The list of American English Phonemes 354
Table 23.1 Recognition result using 100 training samples and 25 test samples 446