Studies in Computational Intelligence, Volume 3

Editor-in-chief
Prof. Janusz Kacprzyk
Systems Research Institute
Polish Academy of Sciences
ul. Newelska 6
01-447 Warsaw
Poland
E-mail: kacprzyk@ibspan.waw.pl

Further volumes of this series
can be found on our homepage:
springeronline.com

Vol. 1. Tetsuya Hoya
Artificial Mind System – Kernel Memory Approach, 2005
ISBN 3-540-26072-2

Vol. 2. Saman K. Halgamuge, Lipo Wang (Eds.)
Computational Intelligence for Modelling and Prediction, 2005
ISBN 3-540-26071-4

Vol. 3. Bożena Kostek
Perception-Based Data Processing in Acoustics, 2005
ISBN 3-540-25729-2

Bożena Kostek

Perception-Based Data Processing in Acoustics

Applications to Music Information Retrieval and Psychophysiology of Hearing

Springer

Professor Bożena Kostek

Multimedia Systems Department
Faculty of Electronics,
Telecommunications and Informatics
Gdansk University of Technology
ul. Narutowicza 11/12
80-952 Gdansk, Poland
and
Institute of Physiology and Pathology of Hearing
ul. Pstrowskiego 1
01-943 Warsaw, Poland
E-mail: bozenka@sound.eti.pg.gda.pl
http://sound.eti.pg.gda.pl

Library of Congress Control Number: 2005926348

ISSN print edition: 1860-949X
ISSN electronic edition: 1860-9503
ISBN-10 3-540-25729-2 Springer Berlin Heidelberg New York
ISBN-13 978-3-540-25729-5 Springer Berlin Heidelberg New York

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilm or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer. Violations are liable for prosecution under the German Copyright Law.

Springer is a part of Springer Science+Business Media
springeronline.com
© Springer-Verlag Berlin Heidelberg 2005
Printed in The Netherlands

The use of general descriptive names, registered names, trademarks, 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.

Typesetting: by the authors and TechBooks using a Springer LaTeX macro package

Printed on acid-free paper SPIN: 11412595 89/TechBooks 5 4 3 2 1 0

To my Father

FOREWORD

Recent years brought up many new techniques combining various aspects in so-called cognitive processing. This way of computing can be used with success in many areas of science and engineering by offering better analogy to human-like processing of information. Such an approach may be especially interesting in acoustics, where we deal with very inaccurate perceptions of phenomena due to the hearing sense characteristics which are highly imprecise with regard to time and spectral resolution.

This book demonstrates in which way soft computing methods (e.g. rough set-based methods) can be applied to provide flexible information processing capabilities for handling ambiguous decision making problems such as for examples musical signal processing, pattern classification, feature and rule generation. Methods of integrating rough sets, fuzzy and artificial neural networks for efficient knowledge discovery are also shown.

Fuzzy logic provides yet another tool that seems one of the best solutions for processing such inaccurate information as can be found in acoustics. In many domains building up membership functions could be problematic, however in acoustics the so-called subjective testing provides a good solution to this problem. Even, if such a testing is time consuming, it falls into the realm of human expertise, thus providing a class of perceptual membership functions. In addition, rules are quite obvious in the acoustic domain, and if not, they can be mined using other soft computing techniques, such as, for example, rough sets. In such a way it is possible to mimic human hearing perception and the way of processing perceived information in human brains.

The book addresses a number of topics such as the fundamentals of hearing and music perception, musical data representation and analysis, automatic classification of musical instrument sounds and musical phrases, musical sound separation, automatic recognition of musical styles, sound processing in hearing prostheses based on artificial neural networks, rough sets fuzzy logic principles, and others – based on cognitive approach. A review of soft computing and data mining techniques is provided, including all mentioned methods and others such as decision trees, evolutionary processing, and genetic algorithms. This book provides however a bal-

anced mixture of both theory and review of applications along with extensive bibliography.

The author has attempted to address a broad group of individuals with the common interest – music and hearing perception and sound processing based on cognitive approach. This is an important issue because of increasing specialization in science, however this book may be read by specialists from various domains and give them a comprehensive view on presented subjects.

This is a book of many new ideas and results, that will be of interest to all those interested in modern approach to imperfect data analysis. The author deserves highest appreciation for her valuable work.

Zdzisław Pawlak

PREFACE

The past can be pondered, the future must be created

E. Schillebeeckx

The emerging concept of human-centered computing or anthropomorphic approach represents a significant move towards intelligent systems, and affords a new perspective on information technology. Relationships between the human brain, mind and perception that have the potential of enhancing peoples' cognitive performance can be found in many domains, examples of which will be shown in relation to music processing and classification. On the other hand, it would be advisable to design systems capable of imitating perceptual processes that are best adapted to specific technological problems.

The objective of this monograph is to provide novel insights into perceptual mechanisms underlying the processing of sound and music in different environments. A solid understanding of these mechanisms is vital for numerous technological applications such as, for example, information retrieval from distributed musical databases. In order to investigate the cognitive mechanisms underlying music perception some soft computing methods will be used. The system proposed by the Author, based on the rough set method and fuzzy logic, provides knowledge on how humans internally represent such notions as quality and timbre and therefore it allows the human-like automatic processing of musical data. In addition, the automatically extracted knowledge on the above processes can be compared to fundamentals of hearing psychophysiology and to principles of music perception. Also other applications of hybrid decision systems to problem solving in music and acoustics will be exemplified and discussed in this book based not only on the review of some literature sources, but also on the experimental results obtained in the Multimedia System Department, Gdansk University of Technology.

The aim of this book is to show examples of the implementation of computational intelligence methods in musical signal and music analysis, as well as in the classification tasks. A part of this book contains a short review of perceptual bases of hearing and music. Then methods and techniques that can be classified as computational intelligence or machine-learning are shortly introduced. The presented methods are applied in the areas considered to be most relevant to music information retrieval (MIR) and acoustics. Accordingly, methods based on such learning algorithms as neural networks, rough sets, fuzzy-logic, and genetic algorithms were conceived, implemented and tested on musical data. In addition, the above-mentioned methods were applied to the analysis of musical duets, musical phrases and audio signals. Another problem discussed within the framework of this book is the 'computing with words' concept applied to both acoustics and psychophysiology. Perception-based analysis applied to psychophysiology focuses on the evaluation of hearing impairments. Application of neural networks to the processing of beamformer signals is another issue reviewed in this book. The last application described is devoted to the problem of audio-visual correlation search. This is based on a hybrid system consisting of rough-fuzzy and evolutionary computation methods.

ACKNOWLEDGMENTS

I am indebted to Professors Z. Pawlak, H. Skarzynski and A. Skowron for their encouragement and interest in this research work.

I am very thankful to Professor J. Kacprzyk who supported the idea of this new book from the beginning.

Much gratitude I owe my distinguished Teachers of Sound Engineering and Acoustics – Professors M. Sankiewicz and G. Budzyński for introducing me to these interesting, interdisciplinary scientific domains.

I thank all my colleagues and students from the Multimedia Systems Department of the Gdansk University of Technology for the discussions and support.

I would also like to thank my husband – Prof. A. Czyżewski for his encouragement and for sharing scientific interests with me.

Finally, I would like to express my gratitude to the members of the Editorial of the Springer-Verlag, in particular Shanya Rehman, Heather King Dr. Thomas Ditzinger, who supported me in a very professional way.

January, 2005
Gdańsk, Poland

Bożena Kostek

CONTENTS

1 INTRODUCTION 1
2 PERCEPTUAL BASES OF HEARING AND MUSIC ..... PERCEPTION 7

2.1 Perceptual Bases of Hearing Perception 7
2.1.1 Human Auditory System 7
2.1.2 Auditory Perception 9
2.1.3 Masking 13
2.2 Perceptual Bases of Music Perception 19
2.2.1 Music Perception 19
2.2.2 Musical Instrument Sounds 23
2.2.3 Musicological Analysis 26
2.2.4 Rhythm Perception 30

References 34

3 INTELLIGENT MUSICAL INSTRUMENT SOUND..... CLASSIFICATION 39

3.1 Problem Overview 39
3.2 Musical Signal Processing 40
3.2.1 Spectral Analysis 40
3.2.2 Wavelet Analysis 49
3.2.3 Pitch Detection Algorithms 52
3.2.4 Parametric Representation 64
3.2.5 MPEG-7 Standard-Based Parameters 84

3.3 Artificial Neural Networks 92
3.3.1 Neural Network Design 95
3.3.2 Recurrent Networks 98
3.3.3 Cellular Networks 99
3.3.4 Gradient-Based Methods of Training Feedforward Networks 100
3.3.5 Application of Pruning Weight Algorithms 102
3.3.6 Unsupervised Training Methods 105
3.3.7 Application of Neural Networks to Musical Instrument Sound Classification 107
3.4 Rough Set-Based Classifier 127
3.4.1 Rough Sets 127
3.4.2 Discretization 134
3.4.3 Application of Rough Sets to Musical Instrument Sound Classification 136
3.5 Minimum Distance Method-Based Classifier 144
3.5.1 Nearest-Neighbor Method and *k*-Nearest-Neighbor Method 144
3.5.2 K-Means Cluster Analysis 146
3.5.3 Application of *k*-Nearest-Neighbor Method to Musical Instrument Sound Classification 147
3.6 Genetic Algorithm-Based Classifier 151
3.6.1 Evolutionary Computing 151
3.6.2 Evolutionary Programming 152
3.6.3 Genetic Algorithms 153
3.6.4 Application of Genetic Algorithms to Musical Instrument Sound Classification 159
3.7 Other Technique-Based Classifiers 165
3.7.1 Decision Trees 165
3.7.2 Hybrid Analysis 167

References 171

4 COGNITIVE APPROACH TO MUSICAL DATA ANALYSIS ... 187

4.1 Musical Signal Separation 187
4.1.1 Musicological and Psychoacoustical Overview 188
4.1.2 Signal Separation Algorithms 194
4.1.3 Automatic Separation of Musical Duets 216
4.2 Musical Phrase Analysis 232
4.2.1 Description of Musical Phrase 238
4.2.2 Parametrization of Musical Phrases 242

4.2.3 Neural Musical Predictor 256
4.2.4 Classification of Musical Phrases 260
4.2.5 Musical Phrase Classification 261
4.2.6 Sequence-Based Classification of Musical Events 263
4.2.7 Feature Vector-Based Classification of Musical Phrases 265
4.2.8 Rough Set-Based Approach 267
4.2.9 Automatic Retrieval of Rhythmic Patterns 270
4.3 Music Analysis 281
4.3.1 Database Organization 282
4.3.2 CDDB Database Organization and Searching Tools 285
4.3.3 Data Mining in CDDB Database 291

References 303

5 COGNITIVE PROCESSING IN ACOUSTICS 315

5.1 Fuzzy Sets and Fuzzy Logic 315
5.1.1 Fuzzy Reasoning in Control 318
5.1.2 Fuzzy Control of Pipe Organ 327
5.2 Rough-Fuzzy-Based Classifier of Musical Timbre 336
5.2.1 Musical Instrument Sound Subjective Descriptors 337
5.2.2 Methodology 338
5.2.3 Rough-Fuzzy Processing of Test Result Processing 339
5.3 Evaluation of Hearing Impairment Using Fuzzy Approach 346
5.3.1 Problem Overview 346
5.3.2 Multimedia Hearing Aid Fitting System (MFHAS) 350
5.3.3 Evaluation of Hearing Impairment Using Fuzzy Approach 352
5.4 Neural Network-Based Beamforming 368
5.4.1 Problem Overview 368
5.4.2 Data Processing 371
5.4.3 System Description 374
5.4.4 Test and Results 377

References 383

6 SYNESTHETIC ANALYSIS OF AUDIO-VISUAL DATA 389

6.1 Data Acquisition 389
6.1.1 Problem Overview 389
6.1.2 Subjective Test Principles 392

6.1.3 Data Processing 395
6.2 Genetic Algorithm-Based Processing 400
6.2.1 Problem Overview 400
6.2.2 Knowledge Base 402
6.2.3 Pattern Searching 405
6.2.4 Rule Generation 406
6.3 Rough-Neuro Processing 407
6.3.1 Neuro-Rough System Principles 407
6.3.2 Experiments 411

References 415

7 CONCLUDING REMARKS 419