Advances in Spatial Science

Editorial Board

Manfred M. Fischer
Geoffrey J.D. Hewings
Peter Nijkamp
Folke Snickars (Coordinating Editor)

For further volumes:
http://www.springer.com/series/3302

Yee Leung

Knowledge Discovery in Spatial Data

Springer

Prof. Yee Leung
The Chinese University of Hong Kong
Dept. of Geography &
Resource Management
Shatin, New Territories
Hong Kong SAR
yeeleung@cuhk.edu.hk

ISBN 978-3-642-02663-8 e-ISBN 978-3-642-02664-5
DOI 10.1007/978-3-642-02664-5
Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: 2009931709

© Springer-Verlag Berlin Heidelberg 2010
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 to prosecution under the German Copyright Law.
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.

Cover design: SPi Publisher Services

Printed on acid-free paper

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

In memory of my father

Wah Sang Leung (梁華生**)**

To

Yuk Lee, my postgraduate advisor
My undergraduate teachers
Sau Kuen Chu, my secondary school teacher

for their initiation, stimulation and guidance in my search for geographical knowledge at various stages of my academic development

Acknowledgements

This monograph contains figures and tables based on copyright figures and tables owned and supplied by China Academic Journal Electronic Publishing House, Elsevier, IEEE, Springer, Taylor and Francis, and Wiley, and are used with their permissions. These comprise of:

Figures 1.2, 2.8–2.12, 2.27–2.34, 4.10–4.14, 5.1–5.7; Tables 4.10–4.13, 5.1 (taken from Springer)

Figures 1.1, 2.1, 2.2, 2.6, 2.7, 2.35–2.37, 5.10, 5.11, 5.14 (taken from IEEE)

Figure 6.5 (taken from China Academic Journal Electronic Publishing House)

Figures 6.6, 6.8, 6.9 (taken from Elsevier)

Figures 1.5, 6.10–6.26; Table 6.2 (taken from Wiley)

Tables 1.1, 4.17–4.28 (taken from Taylor and Francis)

I would like to thank Prof. Manfred M. Fischer who has been encouraging me to write this book for the series. I would also like to thank my research associates, particularly Profs. Z.B. Xu, W.X. Zhang, J.S. Zhang, J.H. Ma, C.L. Mei, J.S. Mi, W.Z. Wu, J.C. Luo, V. Anh and our students who have worked with me over the years to develop the methodologies discussed in this monograph. My appreciation also goes to Ms. Kilkenny Chan and Mr. Eric Wong, particularly Kilkenny, for typing and re-typing the monograph with patience and dedication.

Last but not least, my heartfelt appreciation goes to my wife, Sau-Ching Sherry, for her love and support, and my son, Hei, for giving me a pleasant diversion from work. They both make my life complete and meaningful.

Yee Leung

Preface

When I first came across the term data mining and knowledge discovery in databases, I was excited and curious to find out what it was all about. I was excited because the term tends to convey a new field that is in the making. I was curious because I wondered what it was doing that the other fields of research, such as statistics and the broad field of artificial intelligence, were not doing. After reading up on the literature, I have come to realize that it is not much different from conventional data analysis. The commonly used definition of knowledge discovery in databases: “the non-trivial process of identifying valid, novel, potentially useful, and ultimately understandable patterns in data” is actually in line with the core mission of conventional data analysis. The process employed by conventional data analysis is by no means trivial, and the patterns in data to be unraveled have, of course, to be valid, novel, useful and understandable. Therefore, what is the commotion all about? Careful scrutiny of the main lines of research in data mining and knowledge discovery again told me that they are not much different from that of conventional data analysis. Putting aside data warehousing and database management aspects, again a main area of research in conventional database research, the rest of the tasks in data mining are largely the main concerns of conventional data analysis. Model identification, model construction, and discovery of plausible hypotheses, for example, are not unique to data mining. They are, in addition to model estimation and hypothesis testing, in the agenda of conventional data analysis, such as statistics, also. Searching for clusters, looking for separating surfaces or rules for classification, mining for association rules and relationships, and detecting temporal trends or processes in data constitute the core of the knowledge discovery process that is not unique to data mining. They form the backbone of research in conventional data analysis. From this perspective, there is very little novelty in data mining and knowledge discovery.

On the other hand, if we look at the environment within which data mining and knowledge discovery is taking place, there is something genuine that is worthy of our attention. Though we have traditionally looked for patterns in data by performing clustering, classification, relational analysis, and trend or process analysis, the kinds of data that we are dealing with nowadays are quite different from that targeted by conventional data analysis methods. The sheer volume and complexity of the data that we need to handle nowadays are substantially different from that of the past. Effective discovery of knowledge hidden in data requires novel methods for accomplishing the old tasks. Therefore, it is from this perspective that the mission of data mining and knowledge discovery is justified. This field of research can actually be treated as the continuation of the mission of conventional data analysis into the information and knowledge age. And, our main objective is simply to discover knowledge in data as we have always been doing. Nevertheless, I have no problem in using the term data mining and knowledge discovery adopted by the research community as long as we know exactly what we are doing.

Following up on the literature, I also encountered the term spatial data mining and knowledge discovery. A natural question again is what it is all about, and how it is different from data mining and knowledge discovery in general. An examination of the research activities in this area again tells me that in principle it is more or less similar to that of the general field. The major difference is that data in spatial data mining are mostly geo-referenced and much more complex. And the knowledge to be discovered is often location specific and takes on geometric shapes. Space and time are the two main dimensions along which knowledge discovery is performed. Thus, there is something unique in spatial data mining and knowledge discovery that is worth looking into.

Our main goal is then to discover knowledge in spatial data. It is again in line with conventional spatial data analysis but with special emphasis placed on the nature of spatial data. The idea is to develop novel methods for spatial knowledge discovery. Whether we should call such process spatial data mining and knowledge discovery, or just simply discovery of knowledge in spatial and temporal data is just a matter of terminology. It all involves the discovery of spatial structures, processes, and relationships from spatial and temporal data. As data mining and knowledge discovery have become a commonly employed collective term for such activities, it is used indiscriminately throughout this book. I would not painstakingly point out whether a method should be called a data mining and knowledge discovery method, or just a data analysis method targeting the unraveling of structures, processes and relationships in voluminous and complex spatial and temporal databases.

As a good number of text books and research monographs have been written on data mining and knowledge discovery, one needs a good justification to write another book on the topic. Given the unique features of knowledge discovery in spatial data and the burgeoning growth of research interest in this area, it is an opportune time to make a critical analysis of the field and explore directions for further research. Instead of repeating what has been written in many current books on data mining and knowledge discovery, I would like to write it from the perspective of my own research in this area. So, it is not a text book on data mining and knowledge discovery. It is not a book, like many others, that discusses all aspects of the knowledge discovery process. So there are no discussions on topics such as data warehousing, on-line analytical processing (OLAP), data query, and data mining software. There is no intention to give a comprehensive survey of the literature of the field, although state-of-the-art reviews under relevant topics are made in the book.

This book is intended to be a research monograph on methods and algorithms, conventionally called data mining methods, for the discovery of knowledge in spatial and temporal data. The majority of the methods discussed are based on our own research. So, when I discuss topics such as clustering, classification, relationships and temporal processes, algorithms in the literature are not discussed in detail. Emphasis is placed on the development of our own methods. Nevertheless, it is not difficult to see that some of our methods can, more or less, fit into the family of research methodologies on the same topics. They are developed on the foundation of mathematics, statistics, and artificial intelligence. In brief, the present monograph is not a text book for spatial data mining and knowledge discovery. It is a book for researchers and advanced graduate students who are interested or might have an interest in the methodologies for the discovery of knowledge in spatial and temporal data. The view is more personal, but it fits in with the overall picture of research in the field.

Yee Leung

Contents

1 Introduction 1
1.1 On Spatial Data Mining and Knowledge Discovery 1
1.2 What Makes Spatial Data Mining Different 2
1.3 On Spatial Knowledge 3
1.4 On Spatial Data 4
1.5 Basic Tasks of Knowledge Discovery in Spatial Data 5
1.6 Issues of Knowledge Discovery in Spatial Data 10
1.7 Methodological Background for Knowledge Discovery in Spatial Data 11
1.8 Organization of the Book 12

2 Discovery of Intrinsic Clustering in Spatial Data 13
2.1 A Brief Background About Clustering 13
2.2 Discovery of Clustering in Space by Scale Space Filtering 17
2.2.1 On Scale Space Theory for Hierarchical Clustering 18
2.2.2 Hierarchical Clustering in Scale Space 20
2.2.3 Cluster Validity Check 25
2.2.4 Clustering Selection Rules 29
2.2.5 Some Numerical Examples 31
2.2.6 Discovering Land Covers in Remotely Sensed Images 32
2.2.7 Mining of Seismic Belts in Vector-Based Databases 36
2.2.8 Visualization of Temporal Seismic Activities via Scale Space Filtering 42
2.2.9 Summarizing Remarks on Clustering by Scale Space Filtering 46
2.3 Partitioning of Spatial Data by a Robust Fuzzy Relational Data Clustering Method 49
2.3.1 On Noise and Scale in Spatial Partitioning 50
2.3.2 Clustering Algorithm with Multiple Scale Parameters for Noisy Data 51
2.3.3 Robust Fuzzy Relational Data Clustering Algorithm 54

2.3.4 Numerical Experiments 57
2.4 Partitioning of Spatial Object Data by Unidimensional Scaling 61
2.4.1 A Note on the Use of Unidimensional Scaling 61
2.4.2 Basic Principle of Unidimensional Scaling in Data Clustering 62
2.4.3 Analysis of Simulated Data 64
2.4.4 UDS Clustering of Remotely Sensed Data 66
2.5 Unraveling Spatial Objects with Arbitrary Shapes Through Mixture Decomposition Clustering 70
2.5.1 On Noise and Mixture Distributions in Spatial Data 70
2.5.2 A Remark on the Mining of Spatial Features with Arbitrary Shapes 74
2.5.3 A Spatial-Feature Mining Model (RFMM) Based on Regression-Class Mixture Decomposition (RCMD) 75
2.5.4 The RFMM with Genetic Algorithm (RFMM-GA) 78
2.5.5 Applications of RFMM-GA in the Mining of Features in Remotely Sensed Images 80
2.6 Cluster Characterization by the Concept of Convex Hull 84
2.6.1 A Note on Convex Hull and its Computation 84
2.6.2 Basics of the Convex Hull Computing Neural Network (CHCNN) Model 86
2.6.3 The CHCNN Architecture 89
2.6.4 Applications in Cluster Characterization 94

3 Statistical Approach to the Identification of Separation Surface for Spatial Data 97
3.1 A Brief Background About Statistical Classification 97
3.2 The Bayesian Approach to Data Classification 100
3.2.1 A Brief Description of Bayesian Classification Theory 100
3.2.2 Naive Bayes Method and Feature Selection in Data Classification 101
3.2.3 The Application of Naïve Bayes Discriminant Analysis in Client Segmentation for Product Marketing 102
3.2.4 Robust Bayesian Classification Model 112
3.3 Mixture Discriminant Analysis 113
3.3.1 A Brief Statement About Mixture Discriminant Analysis 113
3.3.2 Mixture Discriminant Analysis by Optimal Scoring 114
3.3.3 Analysis Results and Interpretations 115
3.4 The Logistic Model for Data Classification 117
3.4.1 A Brief Note About Using Logistic Regression as a Classifier 117
3.4.2 Data Manipulation for Client Segmentation 118
3.4.3 Logistic Regression Models and Strategies for Credit Card Promotion 119
3.4.4 Model Comparisons and Validations 125

3.5 Support Vector Machine for Spatial Classification ... 130
3.5.1 Support Vector Machine as a Classifier ... 130
3.5.2 Basics of Support Vector Machine ... 131
3.5.3 Experiments on Feature Extraction and Classification by SVM ... 136

4 Algorithmic Approach to the Identification of Classification Rules or Separation Surface for Spatial Data ... 143
4.1 A Brief Background About Algorithmic Classification ... 143
4.2 The Classification Tree Approach to the Discovery of Classification Rules in Data ... 145
4.2.1 A Brief Description of Classification and Regression tree (CART) ... 145
4.2.2 Client Segmentation by CART ... 148
4.3 The Neural Network Approach to the Classification of Spatial Data ... 156
4.3.1 On the Use of Neural Networks in Spatial Classification ... 156
4.3.2 The Knowledge-Integrated Radial Basis Function (RBF) Model for Spatial Classification ... 159
4.3.3 An Elliptical Basis Function Network for Spatial Classification ... 172
4.4 Genetic Algorithms for Fuzzy Spatial Classification Systems ... 183
4.4.1 A Brief Note on Using GA to Discover Fuzzy Classification Rules ... 183
4.4.2 A General Framework of the Fuzzy Classification System ... 184
4.4.3 Fuzzy Rule Acquisition by GANGO ... 186
4.4.4 An Application in the Classification of Remote Sensing Data ... 194
4.5 The Rough Set Approach to the Discovery of Classification Rules in Spatial Data ... 196
4.5.1 Basic Ideas of the Rough Set Methodology for Knowledge Discovery ... 196
4.5.2 Basic Notions Related to Spatial Information Systems and Rough Sets ... 198
4.5.3 Interval-Valued Information Systems and Data Transformation ... 200
4.5.4 Knowledge Discovery in Interval-Valued Information Systems ... 202
4.5.5 Discovery of Classification Rules for Remotely Sensed Data ... 205
4.5.6 Classification of Tree Species with Hyperspectral Data ... 214
4.6 A Vision-Based Approach to Spatial Classification ... 216
4.6.1 On Scale and Noise in Spatial Data Classification ... 216
4.6.2 The Vision-Based Classification Method ... 218
4.6.3 Experimental Results ... 219
4.7 A Remark on the Choice of Classifiers ... 221

5 Discovery of Spatial Relationships in Spatial Data 223
5.1 On Mining Spatial Relationships in Spatial Data 223
5.2 Discovery of Local Patterns of Spatial Association 225
5.2.1 On the Measure of Local Variations of Spatial Associations 225
5.2.2 Local Statistics and their Expressions as a Ratio of Quadratic Forms 227
5.3 Dicovery of Spatial Non-Stationarity Based on the Geographically Weighted Regression Model 236
5.3.1 On Modeling Spatial Non-Stationarity within the Parameter-Varying Regression Framework 236
5.3.2 Geographically Weighted Regression and the Local–Global Issue About Spatial Non-Stationarity 238
5.3.3 Local Variations of Regional Industrialization in Jiangsu Province, P.R. China 244
5.3.4 Discovering Spatial Pattern of Influence of Extreme Temperatures on Mean Temperatures in China 250
5.4 Testing for Spatial Autocorrelation in Geographically Weighted Regression 254
5.5 A Note on the Extentions of the GWR Model 258
5.6 Discovery of Spatial Non-Stationarity Based on the Regression-Class Mixture Decomposition Method 260
5.6.1 On Mixture Modeling of Spatial Non-Stationarity in a Noisy Environment 260
5.6.2 The Notion of a Regression Class 262
5.6.3 The Discovery of Regression Classes under Noise Contamination 263
5.6.4 The Regression-Class Mixture Decomposition (RCMD) Method for knowledge Discovery in Mixed Distribution 267
5.6.5 Numerical Results and Observations 271
5.6.6 Comments About the RCMD Method 272
5.6.7 A Remote Sensing Application 275
5.6.8 An Overall View about the RCMD Method 276

6 Discovery of Structures and Processes in Temporal Data 277
6.1 A Note on the Discovery of Generating Structures or Processes of Time Series Data 277
6.2 The Wavelet Approach to the Mining of Scaling Phenomena in Time Series Data 279
6.2.1 A Brief Note on Wavelet Transform 279
6.2.2 Basic Notions of Wavelet Analysis 280
6.2.3 Wavelet Transforms in High Dimensions 285
6.2.4 Other Data Mining Tasks by Wavelet Transforms 286
6.2.5 Wavelet Analysis of Runoff Changes in the Middle and Upper Reaches of the Yellow River in China 286

6.2.6 Wavelet Analysis of Runoff Changes of the Yangtze River Basin 289
6.3 Discovery of Generating Structures of Temporal Data with Long-Range Dependence 292
6.3.1 A Brief Note on Multiple Scaling and Intermittency of Temporal Data 292
6.3.2 Multifractal Approach to the Identification of Intermittency in Time Series Data 293
6.3.3 Experimental Study on Intermittency of Air Quality Data Series 297
6.4 Finding the Measure Representation of Time Series with Intermittency 301
6.4.1 Multiplicative Cascade as a Characterization of the Time Series Data 301
6.4.2 Experimental Results 302
6.5 Discovery of Spatial Variability in Time Series Data 307
6.5.1 Multifractal Analysis of Spatial Variability Over Time 307
6.5.2 Detection of Spatial Variability of Rainfall Intensity 309
6.6 Identification of Multifractality and Spatio-Temperal Long Range Dependence in Multiscaling Remote Sensing 312
6.6.1 A Note on Multifractality and Long-Range Dependence in Remote Sensing Data 312
6.6.2 A Proposed Methodology for the Analysis of Multifractality and Long-Range Dependence in Remote Sensing Data 314
6.7 A Note on the Effect of Trends on the Scaling Behavior of Time Series with Long-Range Dependence 317

7 Summary and Outlooks 321
7.1 Summary 321
7.2 Directions for Further Research 322
7.2.1 Discovery of Hierarchical Knowledge Structure from Relational Spatial Data 322
7.2.2 Errors in Spatial Knowledge Discovery 324
7.2.3 Other Challenges 326
7.3 Concluding Remark 327

Bibliography 329

Author Index 351

Subject Index 357

List of Figures

Fig. 1.1 How many clusters are there? 6
Fig. 1.2 How many seismic belts are there? 6
Fig. 1.3 How can the classes be best separated? 7
Fig. 1.4 Is the distribution of mean minimal temperature over 40 years spatially autocorrelated? 9
Fig. 1.5 What is the generating process of this maximum daily concentrations of SO_2? 9
Fig. 1.6 What are the scaling behaviors of these runoffs series? 10
Fig. 2.1 A numerical example of scale space clustering (**a**) Plot of the data set. (**b**) Logarithmic-scale plot of the cluster number $\pi(k)$. (**c**) Logarithmic-scale plot of overall isolation. (d) Logarithmic-scale plot of overall compactness 32
Fig. 2.2 Evolution plot of the scale space clustering in Fig. 2.1 (**a**) Evolutionary tree of cluster centers obtained by the algorithm. (**b**) The partition of the data space obtained by the nested hierarchical clustering algorithm at scales σ_0=0, σ_1=0.99, σ_2=2.38 and σ_3=2.628 (from *bottom* to *top*) 33
Fig. 2.3 Scatter plot of a two-dimensional data set 34
Fig. 2.4 Visualization of the scale-space image obtained from data set in Fig. 2.3 at σ=0.163 (**a**) Scale-space image pseudo-color lot for σ=0.163. (**b**) Mesh plot of scale-space image for σ=0.163. (**c**) Scale-space image contour plot for σ=0.163 34
Fig. 2.5 Visualization of the scale-space image obtained from data set in Fig. 2.3 at σ=1.868 (**a**) Scale-space image pseudo-color plot for σ=1.868. (**b**) Mesh plot of scale-space image for σ=1.868. (**c**) Scale-space image contour plot for σ=1.868 35
Fig. 2.6 Landsat Image of Yuen Long, Hong Kong 36
Fig. 2.7 Land covers revealed by the scale space clustering algorithm 36
Fig. 2.8 Lifetime of the clusterings in Fig. 2.9 39
Fig. 2.9 Mining of seismic belts with MCAMMO (**a**) Original vector-based data set. (**b**) Rasterized image. (**c**) First scale with noises

removed. (**d**) Scale 5. (**e**) Scale 10. (**f**) Scale 13. (**g**) Scale 14. (**h**) Scale 18. (**i**) Scale 25 40
Fig. 2.10 Segmentation after specialization (**a**) Image with the longest lifetime. (**b**) Skeletons. (**c**) Axes of the two longest linear belts. (**d**) Two belts extracted 41
Fig. 2.11 Another seismic area (**a**) Original data set. (**b**) Image at the most suitable scale. (**c**) Skeletons. (**d**) Axes. (**e**) Linear belts. (**f**) Clustering result of Fuzzy C- Lines 41
Fig. 2.12 Lifetime of the clusterings in Fig. 2.11 42
Fig. 2.13 Scale-space clustering for earthquakes ($Ms \geq 6$) 44
Fig. 2.14 Indices of clustering along the time scale for earthquakes ($Ms \geq 6.0$) (**a**) number of clusters. (**b**) lifetime, isolation and compactness of the clustering 45
Fig. 2.15 Ms-time plot of clustering results for earthquakes ($Ms \geq 6$) (**a**) 3 clusters in the 59–95th scale range. (**b**) 17 clusters at the 6th scale step 46
Fig. 2.16 Indices of clustering along the time scale for earthquakes ($Ms \geq 4.7$) (**a**) Number of clusters (The vertical axis just shows the part no larger than 150). (**b**) Lifetime, isolation and compactness of the clustering 48
Fig. 2.17 Ms-time plot of clustering results for earthquakes ($Ms \geq 4.7$) (**a**) 2 clusters in the 74–112th scale range. (**b**) 18 clusters at the 10th scale step 49
Fig. 2.18 Scatter plot of a noisy data set 58
Fig. 2.19 Simulated Experiments of UDS clustering 65
Fig. 2.20 The experimental UDS curves 66
Fig. 2.21 SPOT multispectral image acquired over Xinjing 67
Fig. 2.22 The UDS curve obtained in the remote sensing experiment 68
Fig. 2.23 The histogram of the UDS curve 68
Fig. 2.24 Result obtained by the UDS method 69
Fig. 2.25 Result obtained by the K-means method 69
Fig. 2.26 Result obtained by the ISODATA method 70
Fig. 2.27 Mixture population containing noise and genuine features 73
Fig. 2.28 Process of MDMD algorithm 73
Fig. 2.29 The distributions of various spatial features (**a**) Simple Gaussian class. (**b**) Linear structure. (**c**) Ellipsoidal structure. (d) General curvilinear structure. (e) Complex structure 76
Fig. 2.30 RFMM-GA optimization algorithm 80
Fig. 2.31 Extraction of ellipsoidal feature 81
Fig. 2.32 Extraction of two ellipsoidal features 82
Fig. 2.33 Feature extraction system with RFMM 83
Fig. 2.34 Lineament extraction from satellite imagery (**a**) Original TM5 imagery. (**b**) Results of lineament extraction 83
Fig. 2.35 The $C(S)$ and its inscribed and circumscribed approximations obtained by the CHCNN: case 1 87

Fig. 2.36 The $C(S)$ and its inscribed and circumscribed approximations obtained by the CHCNN: case 2 88
Fig. 2.37 The CHCNN architecture 89
Fig. 3.1 The Radar plot for the selected variables 104
Fig. 3.2 Histograms for the selected variables 107
Fig. 3.3 Experimental separation results with SVM classification. (**a**) A two-class problem. The solid bright dots represent the support vectors. (**b**) A multiple-class problem. The solid bright dots represent the support vectors 137
Fig. 3.4 Original SPOT panchromatic image covering central urban area in Hong Kong 138
Fig. 3.5 The result of urban land cover classification with 5×5 windows 141
Fig. 4.1 A simple tree structure 146
Fig. 4.2 Final binary tree with 46 nodes and 24 terminal nodes at $\alpha = 0.01$ 152
Fig. 4.3 Final binary tree with 113 nodes and 58 terminal nodes at $\alpha = 0.05$ 154
Fig. 4.4 The general architecture of the knowledge-integrated RBF model. (**a**) Data source; (**b**) RBF network; (**c**) Rule-base inference (**d**) Evidence combination 159
Fig. 4.5 The basic architecture of a RBF network 161
Fig. 4.6 Fuzzy ART model for clustering 163
Fig. 4.7 The TM image of the study area. (**a**) The TM image covering the experimental area. (**b**) The three-dimensional display of the same image showing the topographical situation of the area 167
Fig. 4.8 The relationship between average accuracy and the number of kernel unit. (**a**) Land cover map obtained by the MLC classifier. (**b**) Land cover map obtained by the knowledge-integrated RBF model 171
Fig. 4.9 Experimental results. (**a**) Land cover map obtained by the MLC classifier. (**b**) Land cover map obtained by the knowledge-integrated RBF model 171
Fig. 4.10 A mixture distribution of water body sampled from a SPOT-HRV image 173
Fig. 4.11 Architecture of the EM-based EBF classification network 179
Fig. 4.12 Original SPOT image covering the study area 179
Fig. 4.13 Land covers obtained by the EBF network 181
Fig. 4.14 Comparison of average accuracy between the EBF and the RBF networks. (The Curve represents the relationship between the number of hidden nodes and overall accuracy) 182
Fig. 4.15 A fuzzy grid partitioning of a pattern space 185
Fig. 4.16 A schema of fuzzy rule set 185
Fig. 4.17 A fuzzy partition of an axis of spectrum 194

Fig. 4.18 Classification rate of GANGO 196
Fig. 4.19 Lower and upper approximations of a rough concept 199
Fig. 4.20 Discovery of the optimal discriminant function through a blurring process. (**a**) Observing the data set from a very close distance, a discriminant function consisting of the disconnected circles surrounding each datum is perceived. (**b**) Observing the data set from a proper distance, a discriminant function that optimally compromises approximation and generalization performance is perceived. (**c**) Observing the data set from far away, no discriminant function is perceived 217
Fig. 4.21 Simulation result of a spiral classification problem. (The optimal discriminant function is spiral and it is found at σ') 221
Fig. 5.1 The CV score against the parameter θ 245
Fig. 5.2 Spatial distribution of the regression constant in Jiangsu 246
Fig. 5.3 Spatial distribution of the UL parameter in Jiangsu 247
Fig. 5.4 Spatial distribution of the GP parameter in Jiangsu 248
Fig. 5.5 Spatial distribution of the IG parameter in Jiangsu 249
Fig. 5.6 Spatial distribution of the TVGIA parameter in Jiangsu 250
Fig. 5.7 Spatial distribution of the R-Square value in Jiangsu 251
Fig. 5.8 Spatial distribution of the estimates for the coefficient $\beta_1(u_i, v_i)$ of mean maximal temperature over 40 years 253
Fig. 5.9 Spatial Distribution of the estimates for the coefficient $\beta_2(u_i, v_i)$ of mean minimal temperature over 40 years 253
Fig. 5.10 Flowchart of the RCMD method 270
Fig. 5.11 Results obtained by the RCMD method for two reg-classes and one reg-class. (**a**) Scatterplot for two reg-classes. (**a'**) Scatterplot for one reg-class. (**b**) Objective function plot. (**b'**) Objective function plot. (**c**) Contour plot of objective function. (**c'**) Contour plot of objective function 271
Fig. 5.12 Effect of partial model t on the mining of reg-classes. (**a**) $t = 0.001$. (**b**) $t = 0.01$. (**c**) $t = 0.1$. (**d**) $t = 1$. (**e**) $t = 5$. (**f**) $t = 50$ 273
Fig. 5.13 Exact fit property of the RCMD method. (**a**) Scatterplot, with five points exactly. (**b**) Objective function plot located on the line: $y = x$ 274
Fig. 5.14 Identification of line objects in remotely sensed data 275
Fig. 6.1 The Maxican hat wavelet 281
Fig. 6.2 The Haar wavelet 281
Fig. 6.3 The Morlett wavelet 284
Fig. 6.4 Number of months from July, 1919 287
Fig. 6.5 Wavelet coefficient maps of runoff changes 288
Fig. 6.6 Location of hydrological guaging stations in the yangtze river basin 290
Fig. 6.7 Wavelet analysis of the annual maximum streamflow (**a**) and annual maximum water level (**b**) of the Datong station 291

Fig. 6.8 Wavelet analysis of annual maximum streamflow of Datong Station. (**a**) Continuous wavelet power spectrum of the normalized annual maximum streamflow series of Datong station. The thick black contour designates the 95% confidence evel against red noise and the cone of influence (COI) is shown as a lighter shade. (**b**) The cross wavelet transform. (**c**) The squared wavelet coherence result. Arrows indicate the relative phase relationship (with in-phase pointing right and anti-phase pointing left) 292
Fig. 6.9 Wavelet analysis of annual maximum streamflow of Yichang Station (**a**) Continuous wavelet power spectrum of the normalized annual maximum streamflow series of Yichang station. The thick black contour designates the 95% confidence level against red noise and the cone of influence (COI) is shown as a lighter shade. (**b**) The cross wavelet transform. (**c**) The squared wavelet coherence result. Arrows indicate the relative phase relationship (with in-phase pointing right and anti-phase pointing left) 293
Fig. 6.10 Maximum daily concentrations of SO_2 at Queen Mary Hospital ... 294
Fig. 6.11 Maximum daily concentrations of NO at queen mary hospital 294
Fig. 6.12 log periodogram and fitted model (continuous line) of the $QmhSO_2$ series 297
Fig. 6.13 log periodogram and fitted model (continuous line) of the $QmhNO$ series 298
Fig. 6.14 The $\varsigma(q)$ curves for the SO_2 series and fractional Brownian motion 299
Fig. 6.15 The $\varsigma(q)$ curves for the NO series and fractional Brownian motion 299
Fig. 6.16 The $K(q)$ curves for the SO_2 series 300
Fig. 6.17 The $K(q)$ curves for the NO series 300
Fig. 6.18 The $K(q)$ curves and fitted model for the $QmhSO_2$ series 301
Fig. 6.19 The $K(q)$ curves and fitted model for the $QmhNO$ series 302
Fig. 6.20 Maximum daily concentration of SO_2 (parts per billion) at Queen Mary Hospital 303
Fig. 6.21 Maximum daily concentration of NO (parts per billion) at Queen Mary Hospital 304
Fig. 6.22 Maximum daily concentration of NO_2 (parts per billion) at Queen Mary Hospital 304
Fig. 6.23 The $K(q)$ curves of seven SO_2 series 305
Fig. 6.24 The $K(q)$ curves of three NO series and three NO_2 series 305
Fig. 6.25 Fitting of the $K(q)$ curves of SO_2 at the sites ABD, ALC, CHK and WFE 306
Fig. 6.26 Fitting of the $K(q)$ curves of three NO series and three NO_2 series 307
Fig. 6.27 The locations of the 16 stations 309
Fig. 6.28 Normalized rainfall data of the Heyuan station 310

Fig. 6.29 The D_q curves of the 4 stations as examples 310
Fig. 6.30 D_1 and D_2 of the 16 stations 312

List of Tables

Table 1.1 What are the optimal classification rules for the data? 8
Table 2.1 Seismic active periods and episodes obtained by the clustering algorithm and the seismologists 47
Table 2.2 Cluster centers in the experiment 58
Table 2.3 Experimental results of the concordance in languages 59
Table 2.4 Experimental results of clustering of oil types 61
Table 2.5 The error matrix of the numerical experiment 67
Table 2.6 The error matrix of the remote sensing experiment 70
Table 2.7 Diamater of a set S 96
Table 3.1 Descriptive statistics for the bank data set 103
Table 3.2 Selected categorical variables and their values 105
Table 3.3 Selected numerical variables 106
Table 3.4 Classification results obtained by Naive Bayes 110
Table 3.5 Classification results obtained by LDA with available-cases 111
Table 3.6 Classification results obtained by LDA with complete-cases 111
Table 3.7 Classification results obtained by LDA for the whole data set with missing data replaced by the means 111
Table 3.8 Cross validation results of using two assignment criteria 112
Table 3.9 Coefficients obtained by MDA 115
Table 3.10 Results obtained by MDA with feature variables selected by LDA 116
Table 3.11 Results obtained by MDA with feature variables selected by NB 116
Table 3.12 Results obtained by MDA with all feature variables 116
Table 3.13 Comparison of results obtained by MDA with LDA, NB and All 117
Table 3.14 Variable list for the credit card promotion problem 120
Table 3.15 Partial output by SAS logistic procedure for Model-1 122
Table 3.16 Partial output by SAS logistic procedure for Model-2 123
Table 3.17 Target groups of potential clients derived from Model-2 124
Table 3.18 Partial output by SAS logistic procedure for Model-3 125

Table 3.19 Target groups of potential new clients derived from Model-3 126
Table 3.20 Partial output by SAS logistic procedure for Model-4 127
Table 3.21 Comparison of the predicted probabilities and the observed response rate for each group based on Model-2 128
Table 3.22 Comparison of the predicted probabilities and the observed response rate for each group based on Model-3 129
Table 3.23 The correct classification rates of the last 6,000 observations by the respective models fitted with the first 10,000 observations 129
Table 3.24 Comparisons of parameters of the classifiers for land cover classification 139
Table 3.25 The error matrix resulting from the 5×5 window (Accuracy=92.00%, kappa=0.900) 141
Table 4.1 Variables used in the CART 150
Table 4.2 Terminal nodes information for $\alpha = 0.01$ 153
Table 4.3 Terminal nodes information for $\alpha = 0.05$ 155
Table 4.4 Error matrix of classification by the RBF network 168
Table 4.5 Error matrix of classification by the MLC 168
Table 4.6 Error matrix of classification by the BP-MLP 168
Table 4.7 Relationship between accuracy and size of the kernel layer 170
Table 4.8 Error Matrix of classification by the knowledge-integrated RBF model 172
Table 4.9 Land covers of the study area 180
Table 4.10 Error Matrix of classification by the EBF Network 180
Table 4.11 Error matrix of classification by the MLC 182
Table 4.12 Error matrix of classification by the RBF network 182
Table 4.13 Relationship between accuracy and size of the hidden layer 183
Table 4.14 The performance of the proposed training algorithms in five independent runs with p_m=0.00 195
Table 4.15 The performance of the proposed training algorithms in five independent runs with p_m=0.01 195
Table 4.16 A simple decision table 199
Table 4.17 A description of the training samples 206
Table 4.18 A description of the test samples 206
Table 4.19 An interval-valued information system 207
Table 4.20 Discernibility set 207
Table 4.21 Classification accuracy from applying classification reduct $B_1 = \{a_2, a_3\}$ and five rules: r_1, r_2, r_3', r_4, r_5 to the training samples 210
Table 4.22 Classification accuracy from applying classification reduct $B_2 = \{a_3, a_4\}$ and five rules: r_1, r_2, r_3'', r_4', r_5' to training samples 210

Table 4.23 Classification accuracy from applying classification reduct $B_1 = \{a_2, a_3\}$ and five rules: r_1, r_2, r_3', r_4, r_5 to the test samples 211
Table 4.24 Classification accuracy from applying classification reduct $B_2 = \{a_3, a_4\}$ and five rules: r_1, r_2, r_3'', r_4', r_5' to the test samples 211
Table 4.25 Classification accuracy from applying ten rules and three bands (a_2, a_3, a_4) to the training samples 213
Table 4.26 Classification accuracy from applying ten rules and three bands (a_2, a_3, a_4) to the test samples 213
Table 4.27 Spectral bands selected for classification 215
Table 4.28 Comparison of classification accuracies from applying classification reduct B to the training and test tree samples 216
Table 4.29 The statistics of 11 benchmark problems used in simulations ... 220
Table 4.30 Performance of the vision-based classification method 220
Table 5.1 Test statistics of the GWR model 251
Table 6.1 Estimate of p 301
Table 6.2 Values of quantities κ, σ, α and error of all organisms selected 306
Table 6.3 D_1, D_2 for every-5-years rainfall data of Gaoyao station 311
Table 6.4 D_1, D_2 for every-5-years rainfall data of Heyuan station 311
Table 6.5 D_1, D_2 for every-5-years rainfall data of Huiyang station 311
Table 6.6 D_1, D_2 for every-5-years rainfall data of Lianping station 311
Table 6.7 D_1, D_2 of 16 stations using 32 years rainfall data 312