Book Reviews

Supervised and Unsupervised Pattern Recognition: Feature Extraction and Computational Intelligence by E. Micheli-Tzanakou. (Boca Raton, FL: CRC, 2000, 371 pp., ISBN 0-8493-2278-2). *Reviewed by Ke Chen.*

Pattern recognition is a process of description, grouping, and classification of patterns. A pattern is an entity that could be associated with a name. In a real-world pattern recognition problems in diversified forms are ubiquitous and are critical in most human decision making tasks. In terms of information availability there are two general paradigms for pattern recognition; i.e., supervised and unsupervised schemes. A supervised scheme identifies an unknown pattern as a member of a predefined class, while an unsupervised scheme groups input patterns into a number of clusters defined as classes hereafter. For automatic pattern recognition, the primary task consists of feature extraction and classification. Although some effective technologies have been developed, there exist many challenging problems in this field.

In several past decades great efforts have been made in a number of communities to explore solutions to pattern recognition problems. Micheli-Tzanakou produced a book that describes efforts made by over ten researchers in the Neuroelectric and Neurocomputing Laboratories at Rutgers University. Along with concise introductory materials in pattern recognition this volume presents several applications of supervised and unsupervised schemes to the classification of various types of signals and images. The chapters span a number of challenging pattern recognition problems in signal processing, computer vision and image processing, using mainly neural networks and other sophisticated techniques based on ALOPEX (ALgorithms Of Pattern EXtraction).

The ALOPEX is an optimization procedure originally developed by Tzanakou and Harth in 1973 for receptive field mapping in the visual pathway of frogs, which turns out to be useful in a wide range of optimization problems. Centered on ALOPEX this book treats the theoretical and experimental contributions equally and examines both of them. Unlike other books in neural networks, this book gives an emphasis on feature extraction as well, which provides a systematic way to deal with pattern recognition problems in terms of neural networks and computational intelligence.

There are different stages in pattern recognition. The book consists of a Preface, an Introduction, four sections in terms of different stages, and a detailed Index section. In each section, there are a number of chapters on different topics belonging to the same category. The four sections are organized as follows: 1) Overviews of Neural Networks, Classifiers, and Feature Extraction Methods—Supervised Neural Networks (four chapters, 130 p.); 2) Unsupervised Neural Networks (three chapters, 60 p.); 3) Advanced Neural Network Architectures/Modular Neural Networks (three chapters, 42 p.); and 4) General Applications (seven chapters, 118 p.).

The number of chapters and pages devoted to each section is deliberately listed here in order to provide a sense of the author's balanced approach to each particular set of techniques and applications. Seventeen chapters in this book are relatively independent and contributed by different coauthors of E. Micheli-Tzanakou. In the sequel a review of techniques and applications described in the book in a section-by-section way is given.

Section I introduces the subjects of neural networks, classifiers, and feature extraction methods. Neural-network description is focused mainly on models of the supervised learning type. This overview section contains both traditional and state-of-the-art feature extraction methodologies and classifier techniques with an emphasis on neural networks.

W. Chung and E. Micheli-Tzanakou give an overview of classifiers in Chapter 1. Their overview categorizes many well-known classifiers according to how they learn to classify. First of all, they propose criteria for optimal classifier design as a general description of the classification problem. Then existing classifiers are divided into four categories: Bayesian optimal classifiers, exemplar classifiers, space partition methods, and neural networks. It is followed by a detailed description of different classifiers including design and training strategies. Some issues on learning and capabilities of neural networks are particularly addressed, e.g. universal approximation, generalization, and optimization. This chapter is ended by several statistical system performance assessment methods in order to evaluate effectiveness of different classifiers. Chapter 2 is devoted to the basic description of artificial neural networks. In this chapter, D. A. Zahner and E. Micheli-Tzanakou introduce different single neuron models and the classical multilayer perceptron (MLP) architecture. Besides the famous backpropagation training procedure, they present an alternative procedure, the ALOPEX algorithm, for training an MLP network. Some applications are briefly mentioned in order to demonstrate the effectiveness of artificial neural networks.

In Chapter 3, W. Chung and E. Micheli-Tzanakou present a system for handwritten digit recognition developed by them. There has been a long history of handwritten recognition projects. Unlike most of the existing models they adopt a simple model-free strategy where global moments along with proper statistical analysis are merely used. In terms of modeling characteristics of image patterns, the Zernick moments with statistical analysis are described in detail. Simulation

K. Chen is with National Laboratory of Machine Perception and The Center for Information Science, Peking University, Beijing 100871, China. The work of K. Chen was supported in part by the China National Key Fundamental Research under Grant G1999032708.

V. Kvasnicka and J. Pospichal are with the Department of Mathematics, Slovak Technical University, Bratislava, Slovakia.

P. C. Kainen and V. Kůrková are with the Instutute of Computer Science, Academy of Sciences of the Czech Republic, Prague, Czech Republic.

S. Haykin is with McMaster University, Hamilton, ON, Canada.

Publisher Item Identifier S 1045-9227(01)05427-3.

results reported in the chapter show that their method yields quite acceptable results. On the other hand, this system also illustrates how classical pattern recognition techniques, e.g., feature extraction and dimensionality reduction, can be used to tackle a real-world problem.

In Chapter 4, E. Micheli-Tzanakou, A. Ademoglu, and C. Enderwick introduce several recent feature extraction methods beyond traditional spectral analysis based on Fourier transform. Generic feature extraction methods reported include wavelet-based analysis, invariant moments, entropy, cepstrum analysis, and fractal dimension. For modeling texture features, they give a tutorial on the SGLD matrix originally proposed by Haralick *et al.*However, there is an imbalance in material on the different feature extraction methods. The description of most methods seems too brief while there are many details on the wavelet analysis. Nevertheless the readers should be able to obtain detailed information on most feature extraction methods described in the book from the references listed at end of the chapter.

Section II is dedicated to models with the unsupervised type of learning. This section handles unsupervised neural networks, fuzzy neural networks and their applications to a couple of pattern recognition tasks including handwritten character recognition and recognition of normal and abnormal visual evoked potentials.

Chapter 5 presents the joint work of T. J. Dasey and E. Micheli-Tzanakou on fuzzy neural networks. The chapter covers some concepts and background relevant to pattern recognition in terms of unsupervised learning and typical optimization techniques including the ALOPEX procedure. A tutorial on ideas and early work in fuzzy clustering techniques is presented as well. In particular, a modified ALOPEX algorithm is proposed for optimization and a typical system design method based on neural networks is described for pattern recognition. In this chapter, authors attempt to emphasize the underlying relationship between learning in pattern recognition and optimization techniques at the conceptual level. The material of this chapter does not match the title—fuzzy neural networks. Instead this chapter presents some basic concepts of pattern recognition, clustering, and optimization techniques.

In Chapter 6, T. J. Dasey and E. Micheli-Tzanakou describe an application to handwritten digit recognition in great detail. The author first review background of character recognition and major methodologies and follow by data collection, preprocessing, and classification results of their ALOPEX-based method in the unsupervised learning paradigm. A discussion section is concerned with some basic issues in handwritten character recognition and comparison with other methods. Simulation results indicate that the ALOPEX-trained system yields good generalization in the application to handwritten digit recognition. This chapter provides a good example on how to use the ALOPEX algorithm to perform unsupervised learning.

The same authors continue in the next chapter presenting another application to recognition of normal and abnormal visual evoked potentials (VEPs). VEPs have been used as a diagnostic tool in the clinical environment for many decades. Unfortunately, the information conveyed in VEPs is not well understood. Using the same method as in the previous chapter the authors develop a system to capture the useful information from its original form of VEPs, with both the feature extractor and classifier trained by the ALOPEX algorithm. Simulation results shown here, as well as those reported in the last chapter, show the effectiveness of the ALOPEX algorithm in varying domains.

Section III deals with advanced modular neural network architectures and their applications. The motivation of modular designs is analogous to biological systems. A divide-and-conquer strategy is often used for modular designs, leading to a more effective way to tackle complicated real-world problems. In this section, authors present their work in modular designs, applications to medicine and a three-dimensional neural-network model for simulation of brain functions.

Chapter 8 describes the work on classification of mammograms using a modular neural network, done by L. C. T. Cooley and E. Micheli-Tzanakou. Mammography is a key screening tool for breast abnormalities but mammograms may miss some breast cancers due to the image ambiguities. Therefore detection of abnormal mammograms is a challenging task. Authors develop a system for this purpose by means of modular neural networks. After feature extraction, feature vectors are fed to different modules of different structures. Learning in the modular neural network is based on the ALOPEX algorithm. Results corresponding to different modular designs are reported and show promise for a system of this type to be used in the clinical environment. Some design issues on generalization and model selection are addressed in terms of the aforementioned real-world problem.

In Chapter 9, S. Aleynikov and E. Micheli-Tzanakou describe their work in visual ophthalmology: an automated system for classification of retinal damage. Visual ophthalmologist is a commercial system developed by authors that can classify some diseases resulting from retinal hemorrhage. This system consists of five modules ranging from image acquisition to classification. In particular, a modular neural network is adopted as a classifier. The rationale of system design is briefly presented and reasonably good results are reported in this chapter.

E. Micheli-Tzanakou, T. J. Dasey, and J. Bricker present three-dimensional neural-network architecture in Chapter 10. Their motivation is to create a pattern recognition system with neural components. Their rationale of this research is to construct computational cells with similar properties to biological cells so as to benefit from information conservation and proper utilization of neural architectures. As a result they propose a multilayer neural architecture. Each layer consists of a two-dimensional lattice and the cells may receive lateral connections from neighboring cells located in the same layer, while the interlayer is connected in the feedforward way. This architecture has several biologically plausible properties, e.g. cells of receptive field as well as adaptive and self-organization mechanisms. For synaptic weight adaptation a method combining the Hebbian rule and ALOPEX procedure is used. A network of eight neural layers and an input layer has been created to model Parkinson's disease, based on the original work done by DeLong. Although a discussion of the application is provided no comparisons between medical observations and simulations are given. Thus it is impossible to judge how useful this architecture is in the modeling of Parkinson's disease.

In Section IV, authors discuss general applications of the ALOPEX optimization procedure in various fields of pattern recognition and neural networks. Some issues concerning the implementation of this algorithm in neural networks are also addressed. In a sense an attempt to establish a new perspective on the brain-to-computer connection and to understand some findings from human experiments has been made here.

Chapter 11 describes a feature extraction algorithm using connectivity strengths and moment invariants proposed by T. S. Chon and E. Micheli-Tzanakou. The authors first analyze the ALOPEX process for optimization and then present a generalized ALOPEX algorithm where connectivity strengths between cells are used instead of intensities. The connectionist representation introduces a receptive field with connectivities and can be used to model the similarity of the converging image and the templates via an adaptive process. As described in Chapter 4, moment invariants are combination of moments that have some useful properties, for example assure affine transformation invariance in the corresponding images, and have become one of the most important feature extraction methods in pattern recognition. A connectionist algorithm for template matching is described and issues of its convergence for a large and complicated image in conjunction with moment invariants discussed. Simulation results demonstrate that application of the ALOPEX process to connectivity strengths improves the convergence and provides a feasible way to find global and local extrema.

In Chapter 12, D. A. Zahner and E. Micheli-Tzanakou propose an alternative algorithm to train multilayer perceptions based on their ALOPEX procedure. The weight update rule of ALOPEX is derived for a three-layer perceptron. For instance, they apply the three-layer perceptron with ALOPEX to the recognition of templates corresponding to ten digits. In contrast to the backpropagation algorithm their simulation results indicate that the ALOPEX seems to suffer from slower convergence, in particular, in the early period of training. In addition, one more parameter used to control the additive noise should be tuned for convergence in the ALOPEX rule due to its stochastic nature, though the addition of a random noise component tends to avoid getting stuck in local minima. The authors argue that ALOPEX is ideal for VLSI hardware implementation.

The next chapter investigates various aspects of the implementation of neural networks in silicon. S. Wolpert and E. Micheli-Tzanakou explore the implementation of neural models by VLSI circuits of the CMOS type. After reviewing several well-known neuromorphic models developed by other people the authors present their own work—an alternative implementation of individual artificial neuron cells. Furthermore, they describe the VLSI implementation of neural processes modeling the silicon retina and the silicon cochlea.

Chapter 14 is concerned with an application of ALOPEX and wavelet analysis to a typical pattern recognition task speaker identification. F. Phan and E. Micheli-Tzanakou present their work in speaker identification through wavelet multiresolution decomposition and ALOPEX template matching. Speaker identification is the process that classifies an unlabeled voice token as belonging to one of the reference speakers. Since voice always changes over time it is a challenging pattern recognition task. To tackle the problem the authors adopt an interdisciplinary way by incorporating signal processing via orthogonal wavelets, feature extraction via the ALOPEX optimization procedure, and implementation of neural networks. The authors also report their simulation results for text-dependent speaker identification with a small population in the noise-free condition and in the presence of noise. Since no comparison with other methods are presented it is not clear how effective this method is, though the results look reasonable.

Chapter 15 is devoted to another medical application—face recognition in Alzheimer's disease. In this chapter, E. Micheli-Tzanakou briefly presents neural-network models of the relationship between responses of cells and different facial stimuli in the presence of an impairment of the nervous system. The same technique as described in the previous chapter is applied to the problem. The simulation results imply that the satisfactory performance is achieved and the ALOPEX algorithm can yield faster convergence than the backpropagation method by fine tuning parameters.

In Chapter 16, F. A. Fazal and E. Micheli-Tzanakou present their extended Neocognitron work in pattern clustering. Neocognitron is a layered self-organizing neural network model originally proposed by Fukushima. The authors first briefly review the neocognitron model in terms of the classical pattern recognition techniques. Then they utilize a simplified version of the neocognitron to study the pattern clustering performance. Excluding aspects of tolerance to deformations and shifts of position they attempt to understand only the clustering mechanism underlying the neocognitron and find the dependence of clustering on the form and parameters. For this purpose, they perform two different studies via two different neocognitron structures in terms of character recognition. Their empirical study provides insight into the underlying mechanisms of the neocognitron not discussed in the earlier work.

In the final chapter, Chapter 17, E. Micheli-Tzanakou and R. Lezzi, Jr. propose a neural-network approach to model receptive fields in the human cortex using the ALOPEX optimization technique. They assume that in the model the response of a simple cortical cell is strongly localized in both the spatial and frequency domains and treat vision problems as an optimization of some cost function. Thus, a modified ALOPEX algorithm is proposed to perform optimization in their machine vision system. In order to demonstrate the usefulness of such approach they present an application to template matching by memorizing and retrieving patterns. Finally, they show that significant correspondence exist between computer simulations and real brain functions.

Each chapter contains an extensive bibliography that provides a reliable list of good references. The readers will find this list very useful to understand the material in the book. The editor made a great effort to avoid repetition of algorithms given that papers presented in different chapters were relatively independent in their original form. The book *Supervised and Unsupervised Pattern Recognition: Feature Extraction and Computational Intelligence* should prove very helpful to the broad audience of graduate students, researchers, practicing engineers, and professionals in computer and information science, electrical engineering, and biomedical informatics. In particular this book reflects the long-term continuous endeavors of a research group conducting innovatory research and could provide useful hints to novices in the related fields.

The theory and practice of pattern recognition is a broad field, and the authors address one of the most important issues in the field, feature extraction and classification, from the viewpoint of neural computation. As pointed out by J.D. Irwin, editor of the industrial electronic series, many books on neural networks covering computational intelligence exist, but none of them incorporate both feature extraction and computational intelligence. This book is a pioneering volume presenting works of people paying much attention to the close link between feature extraction and computational intelligence. Without a doubt serious attempt of E. Micheli-Tzanakou to address the aforementioned issues must be welcomed by all interested in the fields of pattern recognition and computational intelligence, therefore this book deserves all the credit.

Multi-Valued and Universal Binary Neurons: Theory, Learning, and Applications by I. N. Aizenberg, N. N. Aizenberg, and J. Vandewalle. (Boston, MA: Kluwer, 2000, 275 pp., ISBN 0-7923-7824-5) Reviewed by *Vlado Kvasnicka and Jiri Pospichal*.

The reviewed book is dedicated to an extension of the perceptron, which in its initial form is able to classify correctly only linearly separable patterns. Minsky and Papert have suggested in their seminal book *Perceptron* (published in 1969) that this serious shortcoming may be surmounted by two different ways: The first way was an introduction of the so-called higher-order input activities that are represented by products of single input activities e.g., $x_1 \cdot x_2$, while the second way employed hidden neurons. Minsky and Papert have rejected both these simple and straightforward extensions of the perceptron theory mainly due to nonexistence of a proper learning algorithm. The reviewed book discusses extensively another alternative way how to generalize perceptrons towards an ability to classify patterns that are not linearly separable. The idea is very simple, authors postulated that weight coefficients may be complex numbers and that a respective activation function is determined as follows: ž

$$f(z) = \begin{cases} 1 & \text{if } 0 \le \arg(z) < \pi/2 \text{ or } \pi \le \arg(z) < 3\pi/2 \\ 0 & \text{if } \pi/2 \le \arg(z) < \pi \text{ or } 3\pi \le \arg(z) < 2\pi. \end{cases}$$

Then it is easy to demonstrate that XOR logical function is realizable by this extension of perceptron. The whole book consists of different extensions of the above simple idea that are able to realize more complicated Boolean functions (in particular the so-called k-valued Boolean threshold functions). The notion of linear separability is extended to the so-called P-realizable functions, then multivalued Boolean threshold functions may be correctly realized. Moreover, it is demonstrated that an incremental perceptron learning may be modified to adjust complex weight coefficients, so that multi-valued Boolean threshold functions are realized. At the end of the book illustrative applications are presented that demonstrate an effectiveness of the proposed method (e.g., an associative memory for gray-scale images processing). The book is written in a highly sophisticated style employing mathematical concepts (e.g., group theory) unusual in neural networks. What does the book offer that is substantial enough to be included in neural-network lectures? The two extensions of perceptron to overcome a linear-separability block suggested by Minsky and Papert may be completed by the third possible extension based on complex weight coefficients. This is an interesting fact, but we would recommend the book only to readers who already know neural networks really well, like mathematics, and are specifically interested in perceptron learning.

Feedforward Neural Network Methodology by T. L. Fine. (New York: Springer-Verlag, 1999, series on Statistics for Engineering and Information Science, 340 pp., hardbound, ISBN 0-387-98745-2). *Reviewed by P. C. Kainen and V. Kůrková*.

This research monograph was developed in lectures to intermediate-level engineering and statistics students. It admirably surveys the field of feedforward networks considering both theoretical and practical aspects. The books attempts to respond to four fundamental questions concerning computational capabilities of feedforward networks:

(Q1) What are the functions *implementable* or *representable* by a particular network architecture? (Q2) What is the *complexity* (e.g., as measured by numbers of weights or nodes) of the network needed to implement a given class of functions? (Q3) How can we *select the architecture, weights*, and *node characteristics* to achieve an implementable function? (Q4) What is the capability of the network and selection/training algorithm for *learning* or *generalization* from training samples? The text is divided into seven chapters: 1) Background and Organization; 2) Perceptrons—Networks with a Single Node; 3) Feedforward Networks I: Generalities and LTU Nodes; 4) Feedforward Networks II: Real-Valued Nodes; 5) Algorithms for Designing Feedforward Networks; 6) Architecture Selection and Penalty Terms; 7) Generalization and Learning, which are followed by an appendix on use of the book as a text. The first chapter contains historical and contextual remarks about artificial neural networks and their neurobiological inspirations. A present-day outline of the characteristics of a mammalian neuron is presented as a motivation for computational models. To prevent overly optimistic conclusions about the neurobiological validity of such models, Fine emphasizes that computational models may not yet have captured the essence of neuronal behavior responsible for the power of biological brains. The remaining sections of the chapter begins the examination of the mathematical properties of such computational models. Chapter 2 focuses on the simplest model—a single perceptron. Its classification capacity, generalization ability and learning algorithms are investigated and alternative models of a single computational unit are discussed. Only binary output units are considered. Chapter 3 also concerns only binary output computational units, but in the more general setting of networks with several such units. At the beginning of this chapter, network terminology is introduced (node types, architecture) and then issues of learnability are studied in terms of Vapnik-Chervonenkis capacity. The study of representational powers of networks continues in Chapter 4, which involves real-valued computational units. Here, a survey of recent results on approximation of multivariable real-valued functions by neural networks is presented including some upper bounds on rates of