Z.-Q. Liu, J. Cai, R. Buse

# Handwriting Recognition

Springer-Verlag Berlin Heidelberg GmbH

Studies in Fuzziness and Soft Computing, Volume 133
http://www.springer.de/cgi-bin/search_book.pl?series=2941

**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

Vol. 114. M. Fitting and E. Orowska (Eds.)
Beyond Two: Theory and Applications of Multiple Valued Logic, 2003
ISBN 3-7908-1541-1

Vol. 115. J.J. Buckley
*Fuzzy Probabilities, 2003*
ISBN 3-7908-1542-X

Vol. 116. C. Zhou, D. Maravall and D. Ruan (Eds.)
*Autonomous Robotic Systems, 2003*
ISBN 3-7908-1546-2

Vol 117. O. Castillo, P. Melin
*Soft Computing and Fractal Theory for Intelligent Manufacturing, 2003*
ISBN 3-7908-1547-0

Vol. 118. M. Wygralak
*Cardinalities of Fuzzy Sets, 2003*
ISBN 3-540-00337-1

Vol. 119. Karmeshu (Ed.)
*Entropy Measures, Maximum Entropy Principle and Emerging Applications, 2003*
ISBN 3-540-00242-1

Vol. 120. H.M. Cartwright, L.M. Sztandera (Eds.)
*Soft Computing Approaches in Chemistry, 2003*
ISBN 3-540-00245-6

Vol. 121. J. Lee (Ed.)
*Software Engineering with Computational Intelligence, 2003*
ISBN 3-540-00472-6

Vol. 122. M. Nachtegael, D. Van der Weken, D. Van de Ville and E.E. Kerre (Eds.)
*Fuzzy Filters for Image Processing, 2003*
ISBN 3-540-00465-3

Vol. 123. V. Torra (Ed.)
*Information Fusion in Data Mining, 2003*
ISBN 3-540-00676-1

Vol. 124. X. Yu, J. Kacprzyk (Eds.)
*Applied Decision Support with Soft Computing, 2003*
ISBN 3-540-02491-3

Vol. 125. M. Inuiguchi, S. Hirano and S. Tsumoto (Eds.)
*Rough Set Theory and Granular Computing, 2003*
ISBN 3-540-00574-9

Vol. 126. J.-L. Verdegay (Ed.)
*Fuzzy Sets Based Heuristics for Optimization, 2003*
ISBN 3-540-00551-X

Vol 127. L. Reznik, V. Kreinovich (Eds.)
*Soft Computing in Measurement and Information Acquisition, 2003*
ISBN 3-540-00246-4

Vol 128. J. Casillas, O. Cordón, F. Herrera, L. Magdalena (Eds.)
*Interpretability Issues in Fuzzy Modeling, 2003*
ISBN 3-540-02932-X

Vol 129. J. Casillas, O. Cordón, F. Herrera, L. Magdalena (Eds.)
*Accuracy Improvements in Linguistic Fuzzy Modeling, 2003*
ISBN 3-540-02933-8

Vol 130. P.S. Nair
*Uncertainty in Multi-Source Databases, 2003*
ISBN 3-540-03242-8

Vol 131. J.N. Mordeson, D.S. Malik, N. Kuroki
*Fuzzy Semigroups, 2003*
ISBN 3-540-03243-6

Vol 132. Y. Xu, D. Ruan, K. Qin, J. Liu
*Lattice-Valued Logic, 2003*
ISBN 3-540-40175-X

Zhi-Qiang Liu
Jinhai Cai
Richard Buse

# Handwriting Recognition

## Soft Computing and Probabilistic Approaches

Springer

Professor Dr. Zhi-Qiang Liu
City University of Hong Kong
School of Creative Media
Tat Chee Avenue 83
Kowloon, Hong Kong
PR China
E-mail: smzliu@cityu.edu.hk

and

Department of Computer Science
and Software Engineering
The University of Melbourne
Victoria 3010
Australia
E-mail: zliu@csse.melbourne.edu

Dr. Jinhai Cai
School of Software Engineering
& Data Communications
Queensland University of Technology
GPO Box 2434
Brisbane QLD 4001
Australia
E-mail: j.cai@qut.edu.au

Richard Buse
Department of Computer Science
and Software Engineering
The University of Melbourne
Victoria 3010
Australia

ISBN 978-3-642-07280-2 ISBN 978-3-540-44850-1 (eBook)
DOI 10.1007/978-3-540-44850-1

Library of Congress Cataloging-in-Publication-Data applied for

A catalog record for this book is available from the Library of Congress.

Bibliographic information published by Die Deutsche Bibliothek
Die Deutsche Bibliothek lists this publication in the Deutsche Nationalbibliographie; detailed bibliographic data is available in the internet at <http://dnb.ddb.de>.



http://www.springer.de


Originally published by Springer-Verlag Berlin Heidelberg New York in 2003
Softcover reprint of the hardcover 1st edition 2003



Typesetting: camera-ready by authors
Cover design: E. Kirchner, Springer-Verlag, Heidelberg
Printed on acid free paper 62/3020/M - 5 4 3 2 1 0

# Abstract

Off-line handwriting recognition systems can be broadly divided into three categories: statistical, syntactic, and soft computing-based approaches. Most current systems use traditional statistical approaches that make decisions based on the statistical information present in data. This book takes a fresh look at the problem of unconstrained handwriting recognition and introduces to the reader new techniques using statistical and soft computing approaches. The text discusses in detail the types of uncertainties and variations present in handwriting data. Since handwritings are 2-D data, this book presents several algorithms that use modified hidden Markov models and Markov random field models to model the handwriting data statistically and structurally in a *single* framework. As it is well-known that many uncertainties in handwritings cannot be modeled adequately by traditional statistical methods, the recognition of a word in different styles may be best accomplished by fuzzy logic. The book explores methods that use fuzzy logic and fuzzy sets for handwriting recognition. The effectiveness of these techniques is demonstrated through extensive experimental results on real handwritten characters and words.

# Acknowledgments

The authors would like to thank the staff of the Department of Computer Science and Software Engineering, the University of Melbourne for their generous support and members of the Computer Vision and Machine Intelligence Laboratory for many scholarly debates and inspiring discussions. Zhi-Qiang Liu would like to thank the School of Creative Media, City University of Hong Kong for its support during the preparation of the final draft of this book. Jin-Hai Cai is grateful for all the people in the School of Software Engineering and Data Communications, Queensland University of Technology for their help. We would also like to thank Yee-Cheng Ng for her proofreading of Chapter 4 of the book. Finally, we are very much obliged to Helen Chu Lee for her meticulous proofreading of the entire manuscript. It is her tireless effort that has greatly improved the book.

# Preface

Over the last few decades, research on handwriting recognition has made impressive progress. The research and development on handwritten word recognition are to a large degree motivated by many application areas, such as automated postal address and code reading, data acquisition in banks, text-voice conversion, security, etc. As the prices of scanners, computers and handwriting-input devices are falling steadily, we have seen an increased demand for handwriting recognition systems and software packages. Some commercial handwriting recognition systems are now available in the market. Current commercial systems have an impressive performance in recognizing machine-printed characters and neatly written texts. For instance, High-Tech Solutions in Israel has developed several products for container ID recognition, car license plate recognition and package label recognition. Xerox in the U.S. has developed TextBridge for converting hardcopy documents into electronic document files.

In spite of the impressive progress, there is still a significant performance gap between the human and the machine in recognizing off-line unconstrained handwritten characters and words. The difficulties encountered in recognizing unconstrained handwritings are mainly caused by huge variations in writing styles and the overlapping and the interconnection of neighboring characters. Furthermore, many applications demand very high recognition accuracy and reliability. For example, in the banking sector, although automated teller machines (ATMs) and networked banking systems are now widely available, many transactions are still carried out in the form of cheques. To integrate conventional and e-business models and save operational costs, it is desirable to have automated systems that are able to recognize, interpret, and verify what has been written on the cheque and carry out the required transactions, e.g., bill payments, deposits, and so on. In such applications, the handwriting recognition system must be able to perform *at least* as well as the bank teller; otherwise it will result in too many returned cheques or transaction errors, which would be disastrous in some cases.

In general, a distinction is made between on-line and off-line script in handwriting recognition. For on-line systems, the dynamic (temporal) information about the writing speed and acceleration, and about the order of line-segments making up a word or character is available to the recognizer. That is, in on-line systems we may use spatial-temporal information. As a result, on-line handwriting recognition systems often have higher recognition rates than that of their off-line counterparts. Several well-established algorithms, such as hidden Markov models and dynamic programming, can be applied readily in the on-line case.

In off-line systems, variations in handwritings may result not only from different writing styles but also from different writing media such as pen and paper, etc. Broadly, an off-line handwriting recognition system includes three parts: image pre-processing, feature extractor, and classifier. Pre-processing is primarily used to reduce variations of handwritten characters. It usually involves noise reduction, slant correction, and size normalization. Feature extractor is essential for efficient data representation and extracting, hopefully, most discriminant features to be used in the recognition process. The classifier makes a final decision according to the features and some additional knowledge. In general, off-line handwriting recognition approaches can be divided into two basic categories:

- Segmentation-based approaches that first segment the word into characters or sub-character parts. The recognition process is then accomplished with the aid of a dictionary. Many segmentation algorithms are available in the literature [26, 48, 43, 166]. However, characters in unconstrained handwritten words often connect or even overlap with neighboring characters, which makes it difficult to tell where one character ends and another begins. As a result, such approaches are susceptible to segmentation errors.

- Word-based (holistic) approaches that use no segmentation. Such systems do not rely on individual characters, instead, they recognize words as entities [82, 192, 228]. Although these systems do not require segmentation, they need at least one model or template for each word.

In terms of recognition techniques, off-line handwriting recognition systems can also be divided into three main categories: statistical, syntactic, and soft computing-based approaches. Statistical approaches make a decision by statistical information derived from the input data. In most systems, the probability density functions are assumed to be Gaussian functions or Gaussian mixtures mainly because they are mathematically tractable.[1] Approaches in this category are relatively insensitive to noise and distortions caused by the feature extraction process; however, it is difficult to use statistical model to describe structural information. On the other hand, syntactic approaches model handwritten images by structures that are composed of primitive patterns according to some structural descriptions or a set of rules. There are two major disadvantages in these approaches: First, it is difficult to include statistical information present in the handwriting in the

[1] Although most traditionalists would insist that the Gaussian model represents the data distribution, the lack of robustness in real applications demonstrates that such an assumption has little relevance to reality.

recognition process. This may result in a recognition system that is sensitive to noise. Second, the syntactic descriptions may not able to model distortions in data, such as broken lines, filled holes, etc. which unfortunately happen rather frequently in handwriting images. Neural network approaches have become popular in handwriting recognition since the 1980s. This is mainly attributed to neural network's powerful learning capability and the flexibility of input features. Recently, genetic algorithms have been adopted in handwriting recognition by many researchers. Genetic algorithms are stochastic global search methods, which mimic the mechanism of natural biological evolution. More recently researchers have developed soft computing techniques for handwriting recognition and received encouraging results.

In this book, we present several new techniques for the recognition of handwritten words and characters using both statistical and soft computing approaches. The book is organized as follows:

Chapter 1 introduces the general area of handwritten character recognition, discusses major problems and techniques available in the literature, and overviews the state-of-the-art research and development. More specifically, we discuss some major techniques for feature extraction, which are divided broadly into two categories: features from binary images and from gray-scale images. We will present some of the most popular and important classification methods for handwriting recognition, and discuss their properties.

In Chapter 2, we present pre-processing techniques for slant and correction, size normalization, and baseline determination. We also present in detail the use of the Gabor filter in feature extraction. The most prominent merits of the Gabor filter-based feature extraction method are that it is robust to noise and is able to extract features from both binary images and gray-scale images. In order to extract useful features regardless of the average line-width of handwritten characters or words, we present an algorithm for adaptive estimation of the parameters of Gabor filters. We also introduce several new techniques for extracting such features as skeletons, outer contours, and oriented line segments used in handwriting recognition.

Chapters 3 to 5 present Markov model-based approaches to recognizing handwritten numerals. Chapter 3 describes a new technique for recognizing totally unconstrained handwritten numerals based on hidden Markov models. This approach integrates the structural and statistical information to make a decision. The statistical information is modeled using hidden Markov models, where state-duration-adapted transition probabilities are used to improved the modeling of state durations in the conventional hidden Markov models. To improve the ability of statistical methods in modeling pattern structures, we present the concept of *macro states* for capturing the structural information. In order to speed up the recognition process, we present a modified Viterbi algorithm to avoid the logarithmic and exponential operations that are computationally intensive. The experimental results show that this approach can achieve high performance in terms of speed and accuracy. In Chapter 4 we present a method using a 2-D hidden Markov process to model spectral features for recognizing handwritten numerals. Because Fourier descriptors can achieve invariance to rotation, reflection and scale, they have been widely used for recognizing 2-D shapes such as tools and handwritten numerals. This chapter demonstrates that the 2-D hidden Markov models with features extracted using Fourier descriptors can obtain good results. Since handwritings are 2-D data, we may use Markov random field (MRF) models to model handwritings in their natural data format. In this way, we are able to model the handwriting data statistically and structurally in a single framework. Chapter 5 discusses the important issues of Markov random field models including the framework of MRFs, the neighborhood systems, the design of clique functions and calculation of the global optimum. We apply MRFs to recognizing handwritten numerals.

Chapters 6 to 8 present new techniques for handwritten word recognition. In Chapter 6, we present a method using Markov random field model for unconstrained handwritten word recognition. Again we use the Gabor filter to extract directional features. We use the MRFs to model the relationships between line-segments that have similar orientations within the neighborhood system. Based on these relationships, we define a similarity measure between the templates and images. We establish fuzzy neighborhood systems and design fuzzy matching measures to handle large variations in the data. We use relaxation labeling to maximize the global compatibilities of the MRF models. This chapter also discusses the influence of neighborhood sizes and iteration number of relaxation labeling on recognition rates. Chapter 7 describes a new off-line word recognition system to recognize unconstrained handwritten words based on structural and relational information in handwritten words. We use Gabor filters to extract features from words, and then use an evidence-based approach for word classification. We also present a new method for estimating the Gabor filter parameters, enabling the Gabor filter to be automatically tuned to the word image properties. We introduce two new methods for correcting the slant and tilt present in handwritten words. Throughout the chapters, we present some experimental results.

When writing a given word, say, "*greetings*" there may be large variations between the written words by different people. However, after preprocessing, these different versions of "*greetings*" share some *core* common features that provide the essential clues for recognition. In human perception, such core features play a critical role in identifying the word written in different styles. We argue that the majority of writing variations cannot be modeled adequately by statistical methods and that the recognition of the word in different styles is a process of gathering the *degree* (fuzzy) of evidence from the core features to make a decision. In Chapter 8, we present an off-line word recognition system using fuzzy logic. This system uses structural information in the unconstrained written word. Oriented features in the word are extracted with the Gabor filter. Based on the oriented features, we associate each word with a set of fuzzy word features. We present a two-dimensional fuzzy word classification system where the spatial locations and shapes of the fuzzy membership functions are derived from the training words. The system is able to achieve an average recognition rate of 74% for the word being correctly classified in the top position, and an average of 96% for the word being correctly classified within the top five positions.

Finally, in Chapter 9 we summarize the handwriting recognition techniques presented in this book and discuss some possible and promising directions for further research. In particular, multi-level recognition systems, Markov random field models, and soft computing for handwriting recognition.

Handwriting recognition is not a simple signal processing task, rather, it is an intelligent process that requires learning, reasoning, and knowledge. If we were ever to develop a reliable and accurate recognition system, we cannot rely on traditional computational techniques. We need to use soft computing techniques for developing the new generation of handwriting recognition systems.

# Contents