

V.A. Kovalevsky

Image Pattern Recognition



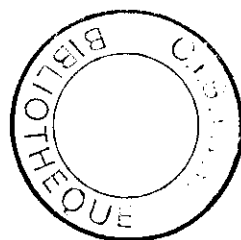
Springer-Verlag
New York Heidelberg Berlin

BIBLIOTHEQUE DU CERIST

V. A. Kovalevsky

Image Pattern Recognition

Translated from the Russian by Arthur Brown



Springer-Verlag
New York Heidelberg Berlin

Preface

During the last twenty years the problem of pattern recognition (specifically, image recognition) has been studied intensively by many investigators, yet it is far from being solved. The number of publications increases yearly, but all the experimental results—with the possible exception of some dealing with recognition of printed characters—report a probability of error significantly higher than that reported for the same images by humans.

It is widely agreed that ideally the recognition problem could be thought of as a problem in testing statistical hypotheses. However, in most applications the immediate use of even the simplest statistical device runs head on into grave computational difficulties, which cannot be eliminated by recourse to general theory.

We must accept the fact that it is impossible to build a universal machine which can learn an arbitrary classification of multidimensional signals. Therefore the solution of the recognition problem must be based on *a priori* postulates (concerning the sets of signals to be recognized) that will narrow the set of possible classifications, i.e., the set of decision functions. This notion can be taken as the methodological basis for the approach adopted in this book.

To specify a set of concrete *a priori* postulates we propose first to define the process that generates the images to be recognized by means of parametric models—i.e., we define the stochastic dependence of the images on various parameters, of which some are essential and others are nuisance parameters. The recognition problem can then be seen as that of adopting

an optimal decision as to the values of the essential parameters. The concept of the parametric model allows us to adopt a single point of view with regard to various recognition methods and various formulations of the problems of learning and self-learning. In particular, it allows us to develop an approach to the problem of structural analysis of composite images as a generalization of the linguistic approach to the problem of recognizing images distorted by noise. We propose an optimization approach to the problem of structural analysis, and offer a method for solving the problem.

In all of the cases that we shall consider, the criterion of optimality is either the probability of making a wrong decision or the likelihood function. The latter is usually easier to compute than the former. It is well known that under certain conditions the maximization of the likelihood function is equivalent to minimizing the probability of making a wrong decision. This is usually the goal of pattern recognition in practical applications. The methods we describe in this book are based on the solution of optimization problems, and the optimality criterion is justified from the point of view of the applications.

The methods we describe are mainly oriented toward the recognition of graphic images in the presence of noise. They have been applied to the development of character readers and to systems for the automatic processing of photographs of particle tracks resulting from physical experiments. Together with modern methods for preprocessing halftone images, they can be used to solve problems in the analysis of aerial and satellite photographs of the earth's surface, images of micro-objects found in biological and medical research, etc.

Most of the chapters of this book are based on the unifying idea of parametrizing the process that generates the images to be recognized, and obtaining the maximum-likelihood estimates for the parameters.

There are nine chapters. The first is devoted to a survey of the literature on both theory and technique of image recognition. The material cited is classified according to the approach used for formulating the problem. The first section of the chapter defines the basic concepts and terms for the field of image processing. After analysis of the results obtained by various authors, it is concluded that effective algorithms for image recognition should be constructed on the basis of a study of the specific features of concrete problems.

In the second chapter we describe a parametric model of a pattern-generating process and show how, using this model, we formulate the problems of recognition, learning and self-learning under various circumstances. We first analyze the obstacles that hinder the successful application of the known general theoretical methods and of several empirical approaches to the image recognition problem. In essence, the second chapter contains a general formulation of the problem, which comprehends most of the special cases and marks out a way to surmount the obstacles that have to be met.

The third chapter is devoted to the problem of teaching a machine to recognize patterns, or learning pattern recognition. We consider a new and more general formulation of the parametric learning problem, taking into account the mutual dependence among realizations of the training sample and the influence of nuisance parameters. It is shown that in the presence of nuisance parameters the statistical problem of learning becomes one of estimating parameters from a mixed sample. The well-known statistical formulation of the self-learning problem is a special case of the formulation given here. We also present a formulation that has undeservedly escaped the attention of research workers in this field—namely one based on minimization of the posterior Bayesian risk. This formulation is the most adequate to meet practical demands and leads to a new class of decision rules, distinct from the traditional ones.

In Chapter IV we consider questions connected with the choice of a system of features, i.e., a means for describing a pattern. We introduce the concept of the insufficiency of a description; this concept, as a criterion, provides a quantitative estimate of the distance between the description and a sufficient statistic. We also indicate the role of entropy as a measure of insufficiency, and we prove some theorems on the quantitative relationship between the conditional entropy and the minimum probability of recognition error. These theorems may be useful in solving problems connected with the choice of optimal systems of features.

The remaining chapters of the book are devoted to the solution of several image recognition problems, using the parametric models defined in Chapter II.

In Chapter V we consider a strategy for solving the recognition problem in the presence of nuisance parameters, which we call the *method of admissible transformations*. We use the method of maximum likelihood to obtain an estimate of the parameters. Using two formally defined notions—prototype image and similarity—we formulate the recognition problem as a search for those values of the parameters yielding a prototype that is most similar to the given signal. We also examine the case for which the set of admissible prototype images is a linear subspace of the signal space, of small dimensionality, and we develop the so-called correlation method, which is optimal in this case. Some experimental results are adduced.

In Chapter VI we solve the recognition problem for the case in which the number of values assumed by nuisance parameters is large and maximization of the likelihood function by an exhaustive search is very cumbersome. We consider a piecewise linear function of bounded complexity, and seek the values of its parameters that will minimize an upper bound of the probability of a recognition error. This task reduces to the search for the minimax of a function of very many variables, and can be implemented by a generalized gradient method. We describe an application to the computation of optimal templates for a character reader.

In Chapter VII we describe a method for recognizing composite images, called the *reference sequence method*. Like the correlation method, it derives

from a system of admissible transformations. It is much more powerful, however; that is, it is applicable in cases where the images in one class may differ from one another by their dimensions and other geometric features. Each of the recognition classes is defined by describing a sequential image-generating process. The process consists of building an image from elementary pieces, according to established rules, and then subjecting it to distortion by random noise. It is shown that under given circumstances, the maximum-likelihood sequence of elementary pieces can be found by dynamic programming. The method yields not only a classification of composite images but also a structural analysis of them. Some algorithms are given for the solution of concrete problems.

In Chapter VIII the reference sequence method is used to solve the recognition problem for a one-dimensional sequence of images. We consider the important problem of recognizing a typewritten line of characters that are not separated by blanks and that are distorted by noise and random injuries. We prove a theorem on the necessary and sufficient conditions for indicating a portion of the optimal sequence before the analysis of the whole sequence is completed.

Chapter IX is devoted to a description of a character reader that implements the concepts, methods, and algorithms described in this book. We describe a technological implementation of the algorithms substantiated in Chapters V and VI. We also discuss the operational characteristics of the reader and the results of experiments in reading a large mass of documents.

It is clear from the foregoing remarks that the book deals with problems of both theory and application in image recognition. It is useful for readers at several levels of preparation. For an overview of the problems one may limit one's reading to the first and second chapters, to Section 3.5 of Chapter III, to Chapter V, and to the first three sections of Chapter VII. Those who are interested primarily in the theory and methodology may omit Chapter V (but not Section 5.1) and Chapter IX, as well as Sections 7.5 and 7.6 and the greater part of Chapter VIII (but not Section 8.4). Those readers who are already familiar with the problem and who are interested in applications should first familiarize themselves with the basic ideas presented in Chapters II and V (Sections 5.1 and 5.2) and then go to Chapter VI and the following chapters.

The author wishes to thank M. I. Šlezinger most warmly for many fruitful discussions.

V. Kovalevsky

Contents

I	The Current State of the Recognition Problem	1
1.1	Basic Concepts and Terminology	2
1.2	Heuristic Paths	11
1.3	Methods Based on Assumptions on the Family of Decision Functions	16
1.4	Methods Based on Assumptions about the Properties of the Signals	23
1.5	Applications and Results	34
1.6	Conclusions	38
II	A Parametric Model of the Image-Generating Process	40
2.1	Difficulties in Image Recognition	40
2.2	A Parametric Model with Additive Noise	43
2.3	The General Parametric Model	49
2.4	Recognition and Learning Viewed as Problems of Optimal Decision with Respect to Parameter Values	51
2.5	Recognition in the Absence of Nuisance Parameters	52
2.6	Recognition with Nuisance Parameters	52
2.7	Optimization of the Decision Rule over a Prescribed Class	53
2.8	Learning and Self-Learning	53
2.9	Problems with Nonstatistical Criteria	55
2.10	Conclusions	56
III	The Parametric Learning Problem	57
3.1	Learning with a Sample of Mutually Dependent Signals in the Presence of a Nuisance Parameter	58
3.2	Learning with Independent Nuisance Parameters	61

3.3	Learning as the Minimization of the Conditional Risk	63
3.4	Conclusions	66
IV	On the Criteria for the Information Content of a System of Features	67
4.1	On the Choice of Primary and Secondary Features	67
4.2	Sufficient Statistics	69
4.3	A Measure of Insufficiency	71
4.4	Generalization of the Measure of Insufficiency to the Case of Probabilistic Transformations	73
4.5	Entropy as a Measure of Insufficiency	75
4.6	On the Kullback Divergence	76
4.7	Entropy and Error Probability	78
4.8	The Information Content of the Optimal Decision	81
4.9	Theorems on the Relation between the Conditional Entropy and the Error Probability	82
4.10	Conclusions	90
V	The Method of Admissible Transformations	91
5.1	Sets that are Closed under Transformations	91
5.2	Admissible Transformations and the Formalization of the Notion of "Similarity"	92
5.3	Peculiarities of the Method of Admissible Transformations	96
5.4	Recognition by the Correlation Method	98
5.5	Experimental Results	104
5.6	Potential Applications of the Correlation Method	111
5.7	Conclusions	114
VI	Optimization of the Parameters of a Piecewise Linear Decision Rule	117
6.1	The Adequacy of a Piecewise Linear Rule and Formulation of the Optimization Problem	118
6.2	The Linear Decision Rule as a Special Case	121
6.3	The Optimization Problem and Its Solution	125
6.4	Solution of the Optimization Problem for a Piecewise Linear Rule	130
6.5	An Application to the Recognition of Alpha-numeric Characters by a Character Reader	135
6.6	Conclusions	143

VII	The Reference-Sequence Method	145
7.1	Formal Statement of the Structural- Description Problem	146
7.2	Formal Syntactical Rules for Constructing Composite Images	147
7.3	Solution of the Problem of Maximum Similar- ity	151
7.4	Images on a Two-Dimensional Retina	158
7.5	Recognition of Lines with Restricted Change of Direction	163
7.6	Recognition of Handwritten Characters	171
7.7	Conclusions	174
VIII	The Recognition of Sequences of Images	177
8.1	Mathematical Model of a Typewritten Line, and Formulation of the Problem	178
8.2	Solution of the Problem	180
8.3	Solution of the Problem with a Correlation Criterion for the Similarity	186
8.4	Recognition of Sequences of Unbounded Length	188
8.5	Examples and Experiments	192
8.6	Scope of the Algorithm and Possible General- izations	196
8.7	Conclusions	198
IX	The "ČARS" Character Reader	199
9.1	The Operational Algorithm for a Charac- ter Reader	200
9.2	The Technical Implementation of the Line Recognition Algorithm	205
9.3	The Choice of the Hardware for Finding and Scanning the Lines	213
9.4	Block Diagram of the ČARS Reader	219
9.5	Tests of the Reader	222
9.6	Conclusions. On the Use of Character Readers	226
	References	228
	List of Basic Notations	235
	Index	239