AD/A-000 445

THEORY OF PATTERN RECOGNITION AND MODERN FORECASTING

V. Kaspin

Foreign Technology Division
Wright-Patterson Air Force Base, Ohio

18 October 1974

DISTRIBUTED BY:

NTIS
National Technical Information Service
U. S. DEPARTMENT OF COMMERCE

UNCLASSIFIED
Security Classification

DOCUMENT CONTROL DATA - R & D

AD/A-000 445

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author)	Foreign Technology Division Air Force Systems Command U. S. Air Force
2a. REPORT SECURITY CLASSIFICATION	UNCLASSIFIED
2b. GROUP	
3. REPORT TITLE	THEORY OF PATTERN RECOGNITION AND MODERN FORECASTING
4. DESCRIPTIVE NOTES (Type of report and inclusive dates)	Translation
5. AUTHOR(S) (First name, middle initial, last name)	V. Kaspin
6. REPORT DATE	1971
7a. TOTAL NO. OF PAGES	15
7b. NO. OF REFS	0
8a. CONTRACT OR GRANT NO.	
b. PROJECT NO.	
c.	
d.	
9a. ORIGINATOR'S REPORT NUMBER(S)	FTD-HC-23-2338-74
9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)	
10. DISTRIBUTION STATEMENT	Approved for public release; distribution unlimited.
11. SUPPLEMENTARY NOTES	
12. SPONSORING MILITARY ACTIVITY	Foreign Technology Division Wright-Patterson AFB, Ohio
13. ABSTRACT	05

DD FORM 1 NOV 65 1473

i

UNCLASSIFIED
Security Classification

EDITED TRANSLATION

FTD-HC-23-2338-74 18 October 1974

THEORY OF PATTERN RECOGNITION AND MODERN FORECASTING

By: V. Kaspin

English pages: 10

Source: Sovet Ekonomicheskoy Vzaimopomoshchi. Postoyannaya Komissiya Po Koordinatsii Nauchnykh i Tekhnicheskikh Issledovaniy. Teoriya i Praktika Prognozirovaniya Razvitiya Nauki i Tekhniki v Stranakh-Chlenakh Sev, Izd vo Ekonomika, Moscow, 1971, pp. 79-84

Country of Origin: USSR
Translated under: F33657-72-D-0854
Requester: FTD/PDTA
Approved for public release; distribution unlimited.

U. S. BOARD ON GEOGRAPHIC NAMES TRANSLITERATION SYSTEM

Block	Italic	Transliteration
А а	*А а*	A, a
Б б	*Б б*	B, b
В в	*В в*	V, v
Г г	*Г г*	G, g
Д д	*Д д*	D, d
Е е	*Е е*	Ye, ye; E, e*
Ж ж	*Ж ж*	Zh, zh
З з	*З з*	Z, z
И и	*И и*	I, i
Й й	*Й й*	Y, y
К к	*К к*	K, k
Л л	*Л л*	L, l
М м	*М м*	M, m
Н н	*Н н*	N, n
О о	*О о*	O, o
П п	*П п*	P, p
Р р	*Р р*	R, r
С с	*С с*	S, s
Т т	*Т т*	T, t
У у	*У у*	U, u
Ф ф	*Ф ф*	F, f
Х х	*Х х*	Kh, kh
Ц ц	*Ц ц*	Ts, ts
Ч ч	*Ч ч*	Ch, ch
Ш ш	*Ш ш*	Sh, sh
Щ щ	*Щ щ*	Shch, shch
Ъ ъ	*Ъ ъ*	"
Ы ы	*Ы ы*	Y, y
Ь ь	*Ь ь*	'
Э э	*Э э*	E, e
Ю ю	*Ю ю*	Yu, yu
Я я	*Я я*	Ya, ya

*<u>ye</u> initially, after vowels, and after ъ, ь; <u>e</u> elsewhere. When written as ё in Russian, transliterate as yё or ё. The use of diacritical marks is preferred, but such marks may be omitted when expediency dictates.

* * * * * * * * * *

GRAPHICS DISCLAIMER

All figures, graphics, tables, equations, etc. merged into this translation were extracted from the best quality copy available.

RUSSIAN AND ENGLISH TRIGONOMETRIC FUNCTIONS

Russian	English
sin	sin
cos	cos
tg	tan
ctg	cot
sec	sec
cosec	csc
sh	sinh
ch	cosh
th	tanh
cth	coth
sch	sech
csch	csch
arc sin	$\sin^{-1}$
arc cos	$\cos^{-1}$
arc tg	$\tan^{-1}$
arc ctg	$\cot^{-1}$
arc sec	$\sec^{-1}$
arc cosec	$\csc^{-1}$
arc sh	$\sinh^{-1}$
arc ch	$\cosh^{-1}$
arc th	$\tanh^{-1}$
arc cth	$\coth^{-1}$
arc sch	sech^{-1}
arc csch	csch^{-1}

---

rot	curl
lg	log

V. Kaspin (USSR)

THEORY OF PATTERN RECOGNITION AND MODERN FORECASTING

The current state of forecasting science is characterized by a rapid growth in the tempo and volume of research and simultaneously by a standard of scientific quality. The range of application of forecasting is growing at an uninterrupted pace, such that it now embraces the most varied areas of human activity and knowledge. It is being applied more and more frequently in such nontraditional areas as automatic control theory and cybernetics, technical and medical diagnostics, long-term planning, international relations, regulation of social processes, etc. This expansion, naturally, has drawn into forecasting research specialists in various sciences, branches of knowledge and industries with their concrete problems and approaches to their solution.

The most important problems in forecasting (as the science of methods of making predictions) at the present time are the utilization of the methods of forecasting for concrete, practical purposes, the search for new methods from the apparatus of other sciences and their reinterpretation in forecasting terms, and the creation on this basis of a characteristic theoretical apparatus.

The theory of pattern recognition should constitute a significant part of this process. The fundamental goal of pattern recognition may be defined as the development of methods of classifying objects and phenomena according to some set of their characteristics. The theory of pattern recognition did not arise accidentally; it arose in response to the rapid growth in the volume of information deriving from scientific research, the advent of digital computers and the evident human inability to mediate between fast computers and the enormous volume of data being processed by them. In the present time pattern recognition is in the process of active development; there does not exist as yet a unified conception of how to approach the problem; there is no complete and general theory or set of methods. Mathematical statistics and game theory, multidimensional analysis and Boolean algebra, regression and factor analysis, information theory, and specific methods developed within the field are all widely used. The development of methods of pattern recognition has made possible the universality of these methods. It is sufficient to note that pattern recognition is now being used in automatic reading and automatic translation, automatic speech perception, medical and technical diagnostics, nuclear physics, criminology, meteorology, paleontology, hydroacoustics, geophysics, radiosonding, etc.

As can be seen even from this short list, the spheres of interest of pattern recognition and forecasting intersect. We will examine the basic concepts and problems arising in the framework of pattern recognition and consider how they relate to the concepts and problems of forecasting. The object of a forecast, as a rule, is a complex or phenomenon described by a fairly large number of variables. In pattern recognition, analogously, the practical problems which arise are most often multidimensional, and so the following geometrical interpretation has become widespread: a pattern is defined as some region of attribute space in which a set of objects, phenomena or states is represented, this set being isolated in accordance with certain considerations into a specific class. In our view, this interpretation is also useful in forecasting research. In it, the object of forecasting is represented as an n-dimensional vector in the space of the variables describing it, and this vector changes in time. At each moment in time the position of the object is characterized by a point; some set of nearby points forms a region of states which are close by some criterion; there exists a transition boundary of the vector, representing the object into a region of qualitatively different states or another situation.

In pattern recognition the boundaries of such regions, like the attribute structure, are defined by the goal of the recognition process, i.e. by how its results will be used. In forecasting the concept of the attribute is also used, but it is not always associated with the goal of the forecast. It is useful, in our view, to define as attributes only those variables, characterizing the process, which are inherent in the goal of the forecast.

Besides primary attributes (obtained directly from the object), pattern recognition uses derivative attributes. These are obtained by various types of transformation of the primary attributes: from the simplest (such as the ratio of two quantities) to the use of special algorithms for recognizing secondary attributes through an aggregate of primary ones. In forecasting the analogs of derivative attributes--such concepts as the "factor," "potential," "indicator," and others--are treated qualitatively and intuitively. It is necessary to make explicit the procedure for obtaining such indicators from the primary attributes.

The principle of dividing attribute space into regions corresponding to various forms or patterns is termed the decision function.

Finally, the concepts of "probability of recognition error" and "recognition reliability" in forecasting may be interpreted as "probability of forecasting error" and "forecasting accuracy."

Let us examine the basic types of recognition problems as a function of the researcher's goal.

1. A set of patterns to be recognized, a set of distinguishing attributes and a permissible error magnitude are given. A decision function providing optimum (in some sense) recognition for the given conditions is required (this is the most widespread type of problem in the theory and practice of pattern recognition). A number of recognition methods and decision functions of both a specialized and a general character have been developed to solve this type of problem: comparison with a standard, correlation methods, potential function methods, statistical testing of hypotheses, and adaptive methods. In our opinion, all of these may be used successfully in forecasting.

The problem may be posed in the following form: there is a set of situations in which the object of the forecast in a given attribute space may find itself, and an allowable probability of forecasting error; what is required is to find the simplest rule relating any given set of atrributes from the given space to one of the known situations with a reliability not less than the one specified. This problem arises when, in planning changes in the parameters of the object of prediction, the investigator wishes to determine the possible consequences of this for the general state of the object. For example, how will labor productivity in some enterprise change as a result of a transition to a five-day working week? A problem of this sort arises when the singular forecasts of individual parameters of a complex object are fulfilled and it is necessary to appraise the situation as a whole, to compare it with some set of known situations.

In the full cycle of generating a forecast there are three basic stages: retrospection, diagnosis and the forecast itself. During the first stage the object of the forecast is defined, the attribute space is constructed, the values of the attributes and the states of the object corresponding to them are defined from the point of view of the previously established goal of the forecast, and the structure of the object and the basic factors influencing the tendencies of its development are determined. During diagnosis the degree of adequacy of the model of the object of the forecast is determined, along with possible forecasting methods, and means for appraising and testing the reliability of the forecast. During the forecasting stage, forecasts of changes in the characteristics of the object or its subsystems are elaborated on the basis of the selected methods; these changes are then combined to yield the final result. In solutions to recognition problems of the first type two basic stages are usually observed: instruction and recognition. During the first stage the attribute space is formed. Then the instructional sequence is given in the form of descriptions of the objects which are known to belong to specific classes. Finally, the simplest decision functions which divide the selected space into regions corresponding to these classes are found.

This stage is in many ways similar to the reconstruction and diagnosis stages in forecasting, and most of the concepts, approaches, functions and instruction algorithms from recognition theory may be successfully applied.

The elaboration of criteria and methods for defining the representativeness of the instructional selection, and also of self-instructional algorithms as a continuation of the instruction of the system when the representativeness of the initial selection is inadequate, is of considerable interest.

The second stage, recognition itself, is the obtaining of a description of an unknown object or phenomenon, the transformation of this description according to the functions found during the instruction stage, and the making of decisions regarding its belonging to a particular class in accordance with the decision functions. This stage can be regarded as corresponding to the synthesis part of the forecasting stage, in which the singular forecasts are combined and synthesized and the resulting forecast specified. The numerous methods for making decisions in pattern recognition may therefore be applied in the generation of forecasts.

The fundamental division of methods of making decisions in pattern recognition is made in accordance with two approaches: the parametric and the probabilistic. In the first approach it is assumed that at any point of the attribute space the probability of the realization of one of the classes is equal to one, and that of the remaining classes, to zero.

In this approach most of the functions reduce to the search for compact regions guaranteeing 100% division of the instructional selection. They are generated on the basis of the construction of various types of "proximity functions" or "appurtenance functions" and are, as a rule, fairly simple and convenient.

The second approach (and the probabilistic methods corresponding to it) assumes that various classes may be realized at a single point in attribute space with a probability different from one.

In this case the theory of the testing of statistical hypotheses is used: the "ideal observer," the "minimax" criterion, the Bayesian principle, the Neumann-Pearson criterion, etc.

This approach and the methods corresponding to it should come to be widely used in the generation of complex forecasts as a result of their probabilistic nature.

2. The problem of finding the most informative system of attributes is one of the most important, and one of the most difficult to solve by formal methods, in pattern recognition. Three fundamental aspects of this problem may be distinguished: the selection of a system of attributes as regards the problem itself, optimization of some system of attributes, and determination of the optimal discretization or quantization of the attributes. It is evident that the solution of these problems is extremely important for modern forecasting now that the systems approach and the complex nature of forecasts have become generally recognized.

The enormous number of parameters characterizing objects of forecasting has caused "accursed dimensionality" to become an insuperable obstacle to the solution of the problem.

At the present time the selection of the initial system of attributes for both recognition and forecasting problems is effected on the basis of experience, intuition and concepts of the goal of the research in question. The study of the capacities of biological systems is carried out basically in the area of visual and auditory perceptions; individual results are obtained (the "frog's eye" in radiosonding), but general principles are not uncovered and the question of the selection of attributes remains the least studied in the field of pattern recognition, and so reliance is placed on the selection of the most informative subsystem of attributes from some initial system of greater dimensions.

The most widely used means of evaluating the information content of attributes in pattern recognition is the Shannon measure of information--the difference between the initial and final entropy of the system utilization of the attribute in question. However, this definition was created for communications theory, as a statistical evaluation of a signal independent of the meaning of the message being communicated.

In our view, in structuring forecasting research from the point of view of pattern recognition, in defining precisely the state of the object and the boundaries between them in the retrospection stage, combined probability distributions of attributes and situations contain information about their role in a given process; evaluations of attributes on the basis of such distributions are therefore of interest.

The reliability of recognition according to training or a test sample is also used as a criterion of the effectiveness of a subsystem of attributes in pattern recognition theory. Clearly, this approach may be used with sufficiently complete retrospective data to select the attributes of the object of the forecast.

With a large number of attributes in the initial set and in the selected subsystem the sorting range may prove to be excessively large. In this case various forms of directed sorting are used, such as random search with adaptation.

Another trend in efforts to increase the effectiveness of systems of attributes is the use of various types of linear and nonlinear transformations of coordinate space in order to increase the discriminating properties of the system of attributes or render the attributes invariant to a given group of allowable transformations of descriptions of the object (normalization of descriptions). This trend has been fairly extensively developed in theoretical and applied work on recognition, and the results are applicable to various problems in increasing the information content of descriptions of objects of forecasting.

The last, and extremely important aspect of the solution of various types of information problems in pattern recognition, is optimal quantization and discretization. The lower limit of discretization of attributes may be clarified with the aid of Kotel'nikov's theorem, but the optimal time interval, as has been shown in the works of B. Varskiy, Yu. Barabash and others, is determined by the characteristics of a random process, and in particular by the magnitude and sign of the autocorrelation function of the process. In pattern recognition, communication theory and its methods, in particular the Shannon method of evaluating information content, are widely used to deal with this problem. Optimal quantization is sought through successive rejection of boundaries of low information content from the preliminary uniform scale of the attribute. This algorithm has yielded good results in an experiment in the recognition of random projections of three-dimensional objects.

Problems of optimal quantization are extremely important in forecasting research, where they are closely associated with problems of measurement and problems of qualitative-quantitative transformations. Discussion of this problem would be beyond the scope of this paper; we note only that the methods and results of pattern recognition are extremely useful in this area.

3. Problems of taxonomy are the least developed in the field of pattern recognition. These problems consist basically in the division of initial situations into classes in accordance with various "proximity" or "similarity" considerations. This division should be useful in some sense, and therefore the concept of an "objective" taxonony has no practical value. In reality the "subjectivity" of automatic taxonomy is defined by the goal of the research in question and the attribute space being used. Taxonomic algorithms are used in specific areas of social and forecasting research.

The basic idea of the algorithm is to unify points in attribute space into groups according to the proximity principle in some metrical system in such a way that the average distances between points within groups are significantly less than the average distances between the points of different groups. The results of experiments have shown the usefulness and promise of taxonomic methods in social research and in forecasting.