PATTERN RECOGNITION USING GENERALIZED PORTRAITS

V. N. Vapnik and A. Ya. Lerner

(Moscow)

Translated from Avtomatika i Telemekhanika, Vol. 24, No. 6,
pp. 774-780, June, 1963
Original article submitted December 26, 1962

An axiomatic definition of pattern is given. The concepts of "generalized portrait", "distinction", and "recognition" are introduced. Algorithms are proposed for learning recognition and distinction on the basis of finding generalized portraits of patterns.

A number of studies have appeared in recent years on the subject of modelling pattern recognition processes [1-3]. A number of effective algorithms have been proposed for learning to distinguish visual patterns in special-purpose or universal digital computers. The posing of these problems and the results obtained constitute an important stage in the development of self-organizing systems.

In the present paper an attempt is made to formalize certain concepts connected with pattern recognition. The authors start from the principle that a pattern is defined by the objective properties of the set of objects under consideration and the subjective properties of the machine perceiving them.

This approach has permitted us to introduce the concept of "generalized portrait" and threshold of recognition, the ensemble of which characterizes the system of machine patterns, and to show that the problem of pattern recognition in general consists of the two subsidiary problems of recognizing and distinguishing patterns.

Let us consider a machine consisting of a perception device PD, a transformation device TD and a recognition device RD (Fig. 1). We shall term the state of its perception device the image ϕ_i of the ith object presented to the machine, and the outputs f_i of the conversion device we term the description of the object. The following considerations are based on the definition of the concept "pattern".

Let there exist a certain set of objects H. We shall assume that the set of their images for the given machine can be divided into n patterns if the set of objects H can be divided into n subsets $H_1, \ldots, H_n$, such that after a certain sufficient number of objects of each subset has been shown to the machine, it can divide the entire set H into the same subsets $H_1, \ldots, H_n$. It is assumed that for each of these subsets $H_1, \ldots H_n$ there exists a description permitting an estimate of the degree of correspondence to it of the descriptions of each object.

The unique assignment of an object to a subset is then possible when, if the descriptions of two objects belonging to different subsets are compared to the description of one of the subsets, a larger value of the estimation parameter is obtained for the description of the object which belongs to that subset.

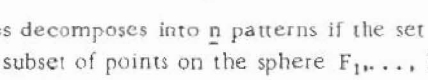

Fig. 1.

1. Definition of pattern

Let there exist a set of images T. We shall consider a certain subset of the images $\Phi \subset T$ and a certain single-valued transformation $\mathcal{F} \in U$, where U is a given set of single-valued transformations, each of which places in correspondence with each image $\phi \in \Phi$ a point on the unit sphere in Hilbert space.

We shall say that the set of images decomposes into n patterns if the set Φ can be divided into n subsets $\Phi_1, \ldots, \Phi_n$ such that the corresponding subset of points on the sphere $F_1, \ldots, F_n$ has the following property.

For each subset F_i a point on the sphere φ_i can be found such that for any $f_i \in F_i$, $f_j \in F_j$ the inequality

$$(f_i\varphi_i) > (f_j\varphi_i) \quad (i \neq j) \tag{1}$$

is satisfied.

If the set Φ is such that there does not exist an image $\phi \in T \setminus \Phi$, whose corresponding point f^* satisfies the inequalities

$$(f^*\varphi_i) > \min_{f_j} (f_j\varphi_j), \ (f^*\varphi_i) < \min_{f_*} (f_i\varphi_i) \quad (i \neq j),$$

we shall say that the images Φ divide into defined patterns $^*\Phi_1, \ldots, ^*\Phi_n$. In the contrary case we say that the images belong to the indefinite patterns $^*\Phi_1, \ldots$

We shall say that the images Φ do not divide into n patterns if in the given set of transformations U there is no transformation on the sphere $\mathcal{F}$, for which condition (1) is satisfied*.

It is obvious that several transformations may be found for which conditions (1) are satisfied.

We shall say that the transformation $\mathcal{F}_i \in U$ is similar to the transformation $\mathcal{F}_j \in U$, and write $\mathcal{F}_i \to \mathcal{F}_j$, if each set Φ, dividing into n patterns $^*\Phi_1, \ldots, ^*\Phi_n$ under the transformation $\mathcal{F}_i$, can be decomposed into the same n patterns $^*\Phi_1, \ldots, \Phi_n$ under the transformation $\mathcal{F}_j$. The set of patterns $^*\Phi_1, \ldots, ^*\Phi_n$ into which a given set of images decomposes under the transformation $\mathcal{F}_i$ will be termed a system of $\mathcal{F}_i$-homogeneous patterns, and will be denoted by $[\mathcal{F}_i(\Phi)]$.

2. Recognition and Distinction

Let us consider a system of $\mathcal{F}_i$-homogeneous patterns $[\mathcal{F}_i(\Phi)]$. Let the patterns of this system be $^*\Phi_1, \ldots, ^*\Phi_n$. To the subset of images $\mathcal{F}_i$ under the transformation $\mathcal{F}_i$ let there correspond the subsets of points on the sphere $F_1, \ldots, F_n$.

We shall consider a certain image $\phi \in T$ and the point f corresponding to it. We call the establishment of the membership of f in the subset F_i pattern distinction if it is known in advance that the image ϕ belongs to the set Φ. In the contrary case establishment of the membership of the point f in one of the subsets $F_1, \ldots, F_n$ is termed pattern recognition. The recognition problem can be defined also for a set of indefinite patterns. The following geometric interpretations of recognition and distinction can be proposed.

Conditions (1) may be written in the form $(f_i\varphi_i) \geqslant C_i$, where $C = \min_{f_i} (f_i\varphi_i)$. This signifies that in some functional space in the sphere with center at the point φ_i and radius R_i no description of an object belonging to pattern $^*\Phi_j$ $(j \neq i)$ can occur.

In the case of distinction incidence in the sphere is sufficient for recognition of the pattern. However the mutual dispositions of the spheres may be such that certain of them may intersect (Fig. 2). The portion of the points belonging simultaneously to two spheres cannot be recognized by inequalities (1) (i.e., these points are not images belonging to the sample).

If it is not known in advance if the given image belongs to one of the n patterns, incidence of its description within one of the spheres is not sufficient for recognition. It is further necessary to determine if the description of the description of the object is incident within a single sphere only.

Let us consider the conditions

$$(f_i\varphi_i) > (f_j\varphi_i), \ (f_j\varphi_j) > (f_i\varphi_j).$$

We term $k_i = \min_{f_i} (f_i\varphi_i)$ the recognition threshold of the pattern $^*\Phi_i$, and correspondingly, $k_j = \min_{f_j} (f_j\varphi_j)$ the recognition threshold of the pattern $^*\Phi_j$.

Let the system of $\mathcal{F}_i$-homogeneous patterns $[\mathcal{F}_i(\Phi)]$ be such that the definite patterns composing it are

*This definition of pattern can be given in metric space, replacing conditions (1) by the conditions $\rho(f_i, \varphi_i) < \rho(f_j, \varphi_j)$.

$^*\Phi_1, \ldots, {}^*\Phi_n$ (a subset of all images belonging to the patterns $\Phi_1, \ldots, \Phi_n$, and the corresponding subset of points on the sphere $F_1, \ldots, F_n$).

According to the definition of pattern, for each subset there exists a certain point of the sphere φ_k. We term this point on the sphere φ_k the generalized portrait of the pattern $^*\Phi_k$.

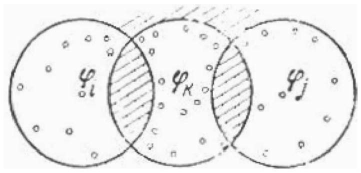

Fig. 2.

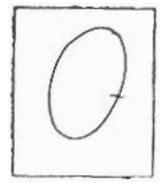

Fig. 3.

It is obvious that the generalized portraits of the patterns $\varphi_1, \ldots, \varphi_n$ and the set of their recognition thresholds $k_1, \ldots, k_n$ fully characterize such a system of $\mathcal{F}_i$-homogeneous patterns*.

Let $k_1 \neq 0, \ldots, k_n \neq 0$; we put $s_j = (1/k_j)\varphi_j$. The inequalities (1) may be rewritten in the form

$$(f_i s_i) \geqslant 1 > (f_j s_i), \qquad (f_j s_j) \geqslant 1 > (f_i s_j),$$

whence

$$(f_i s_i) > (f_i s_j).$$

The last inequality can be taken as the condition for distinguishing the patterns $^*\Phi_1, \ldots, {}^*\Phi_n$ of the system $[\mathcal{F}_i(\Phi)]$.

If a certain object f_j does not belong to any of the patterns, $F(\Phi)$ can be estimated from the quantity $^*\Phi_i$, and it can be assigned to that pattern (f_j, s_j) for which the value $^*\Phi_k$ is maximum.

Let us put $\theta(f s_i) = x_i^k$,

$$\theta(z) = \begin{cases} 0 & \text{for} \quad z < 1, \\ 1 & \text{for} \quad z \geqslant 1. \end{cases}$$

Here the upper index indicates the index of the system of $\mathcal{F}_k$-homogeneous patterns containing the sample $^*\Phi_i$. Then the condition for recognizing the pattern $^*\Phi_i$ consists in determining such x_i^k, that

$$\bar{x}_1^k \ldots \bar{x}_{i-1}^k x_i^k \bar{x}_{i+1}^k \ldots \bar{x}_n^k = 1. \tag{2}$$

The existence of this requirement indicates that the operation of recognition does not terminate after the recognition threshold has been passed, while the operation of distinguishing terminates with passage of the threshold.

Thus, to determine the membership of some image in a pattern, it is necessary to:

1) compare the description of the object presented to all the generalized portraits of the homogeneous systems in the memory;

2) put out the signal of passage of the threshold;

*It should be kept in mind, that with increase of learning (i.e., with increase in the number of patterns entering into the $\mathcal{F}_i$-homogeneous system) patterns can be redefined. However at each given stage there exists its corresponding system of true $\mathcal{F}_i$-homogeneous patterns.

Thus, it is likely that a person ignorant of the Latin alphabet will term the image in Fig. 3 the letter O. A person who knows the Latin alphabet, on being asked what the drawing represents, will probably answer: "I don't know."

3) verify if requirement (2) has been satisfied in the case of recognition.

It is useful to introduce one further definition. We term the order of distinguishability of two patterns $^*\Phi_i$ and $^*\Phi_j$ the quantity $I = 1 - (\varphi_i\varphi_j)$, and the order of distinguishability of the system of $\mathcal{F}_i$-homogeneous patterns the quantity

$$I = 1 - \max_{i,j} (\varphi_i\varphi_j),$$

where φ_i, φ_j are the generalized portraits of the system $\{\mathcal{F}_i(\Phi)\}$. It is possible to consider the order of distinguishability as the characteristic degree of similarity of the patterns in the system of homogeneous patterns.

3. One Representation

As a specific representation we shall consider the excitation of the receptor field, i.e., a certain sequence $\alpha_1, \ldots, \alpha_n$, where $0 \leq \alpha \leq 1$.

Let us consider the set of transformations U over the receptor field. Each transformation $\mathcal{F}_{i_k} \in U$ puts into correspondence with the sequence $\alpha_1, \ldots, \alpha_n$ a certain unit vector.

Assume that by means of the transformation $\mathcal{F}_{i_k}$ to the sequence $\alpha_1, \ldots, \alpha_n$ is placed in correspondence the unit vector $(\beta_1, \ldots, \beta_e)$, and by means of the transformation $\mathcal{F}_{i_e}$ to the same sequence $\alpha_1, \ldots, \alpha_n$ is placed in correspondence the unit vector $\gamma_1, \ldots, \gamma_d$.

We define a certain transformation $\mathcal{F}_{i_k+i_e}$ as the "sum" of transformations $\mathcal{F}_{i_k+i_e} = \mathcal{F}_{i_k} \oplus \mathcal{F}_{i_e}$:

$$\mathcal{F}_{i_k+i_e}(\alpha_1, \ldots, \alpha_n) = \mathcal{F}_{i_k}(\alpha_1, \ldots, \alpha_n) \oplus \mathcal{F}_{i_e}(\alpha_1, \ldots, \alpha_n) = \left(\frac{\beta_1}{\sqrt{2}}, \ldots, \frac{\beta_e}{\sqrt{2}}, \frac{\gamma_1}{\sqrt{2}}, \ldots, \frac{\gamma_d}{\sqrt{2}}\right).$$

We define the operation "multiplication by a number":

$$\mathcal{F}(\alpha_1, \ldots, \alpha_n) = (c_1\mathcal{F}_{i_k} \oplus c_2\mathcal{F}_{i_e})(\alpha_1, \ldots, \alpha_n) = \left(\frac{c_1\beta_1}{\sqrt{c_1^2+c_2^2}}, \ldots, \frac{c_1\beta_e}{\sqrt{c_1^2+c_2^2}}, \frac{\gamma_1}{\sqrt{c_1^2+c_1^2}}, \ldots, \frac{\gamma_d}{\sqrt{c_1^2+c_2^2}}\right).$$

These operations are neither commutative nor associative, but any transformation $\mathcal{F}_{(i_k+i_e)+i_n}$ can always be represented as

$$\mathcal{F}_{(i_k+i_e)+i_n} = c_1\mathcal{F}_{i_k} \ominus c_2\mathcal{F}_{i_e} \oplus c_3\mathcal{F}_{i_n}.$$

It is easily shown that this transformation $\mathcal{F}^* = c_1\mathcal{F}_{i_1} \oplus \ldots \oplus c_n\mathcal{F}_{i_n}$ is similar to an arbitrary transformation

$$\mathcal{F}_{i_k} \subset U^* \quad (k = 1, 2, \ldots, n), \quad \mathcal{F}_{i_k} \rightarrow \mathcal{F}^*.$$

In the transformation $\mathcal{F}_{i_k}$ let a certain set of images Φ divide into the patterns $^*\Phi_1, \ldots, {}^*\Phi_n$, whose generalized portraits are $\varphi_1, \ldots, \varphi_n$.

Let us consider two images. Image $\underline{a}$, belonging to pattern $^*\Phi_1$, and image $\underline{b}$, not belonging to pattern $^*\Phi_1$. Let, in the transformation $\mathcal{F}_{i_k}$, their vectors and the generalized portrait of the pattern $^*\Phi_1$ take the forms

$$\mathcal{F}_{i_k}(a) = f_1 = (d_1, \ldots, d_{h_i}), \mathcal{F}_{i_k}(b) = f_2 = (\varepsilon_1, \ldots, \varepsilon_{h_i}), \varphi = (\nu_1, \ldots, \nu_{h_i}).$$

Conditions (1) for these vectors take the form

$$\sum_i d_i\nu_i > \sum_i \varepsilon_i\nu_i.$$

These same images in the transformation $\mathcal{F}^*$ will be

$$\mathcal{F}^*(a) = f_1^1 = (c_1 d_1, \ldots, c_1 d_{h_1}, c_2 x_2^1, \ldots, c_2 x_{h_2}^2, \ldots, c_n x_1^n, \ldots, c_n x_{h_n}^n),$$
$$\mathcal{F}^*(b) = f_2^1 = (c_1 e_1, \ldots, c_1 e_{h_1}, c_2 y_2^1, \ldots, c_2 y_{h_2}^2, \ldots, c_n y_1^n, \ldots, c_n y_{h_n}^n).$$

Let us consider the vector $\varphi^1 = (\nu_1, \ldots, \nu_h, 0, \ldots, 0)$. Then

$$(f_1^1\varphi^1) = c_1 \Sigma d_i \nu_i, \quad (f_2^1\varphi^1) = c_1 \Sigma e_i \nu_i$$

and the inequality $(f_1^1\varphi^1) > (f_2^1\varphi^1)$ follows from inequality (A). Let $\mathcal{F}_1(\alpha_1, \ldots, \alpha_n) = (\beta_1, \ldots, \beta_n)$, $\mathcal{F}_2(\alpha_1, \ldots, \alpha_n) = (0, \ldots, 0, \beta_2, 0, \ldots, 0)$.

We define the operation "subtraction":

$$\mathcal{F}_3(\alpha_1, \ldots, \alpha_n) = \mathcal{F}_1(\alpha_1, \ldots, \alpha_n) \ominus \mathcal{F}_2(\alpha_1, \ldots, \alpha_n) = \\ = \left(\frac{\beta_1}{c}, \ldots, \frac{\beta_{i-1}}{c}, \frac{\beta_{e+1}}{c}, \ldots, \frac{\beta_h}{c}\right),$$

where

$$c^2 = \beta_1^2 + \ldots + \beta_{i-1}^2 + \beta_{i+1}^2 + \ldots + \beta_n^2.$$

Let us consider a machine consisting of n transformation devices each of which has the structure defined by the transformation $F_i \in U$. The machine can realize one of n transformations.

According to the above, such a machine may be replaced by a machine whose entire structure is given by a single transformation

$$\mathcal{F}^* = c_1\mathcal{F}_1 \oplus \ldots \oplus c_n\mathcal{F}_n,$$

where $\mathcal{F}_1, \ldots, \mathcal{F}_n$ is the system of transformations in the set U.

Then each set of images Φ, dividing into m patterns for the first machine, divides into the same patterns for the second machine also.

The structure (system of transformations) of each machine should be chosen according to the purpose of the machine and its image should be joined in a pattern.

The problem of finding a transformation providing effective pattern recognition apparently cannot be solved by direct calculation. Nevertheless there exists a means of finding the required transformation, corresponding to the goals set before the machine. This is the method of directed selection.

By directed selection we understand a progressive improvement of an arbitrary initial structure such that inefficient (with respect to a given criterion) portions of the structure are eliminated and replaced by other random structures. It is obvious that this will always lead to progress of the structure.

It should be emphasized that this method is not that of random selection over all possible structures, and therefore each succeeding modification of the machine does not inherit inefficient components.

If as the criterion of efficiency of the structure we take the order of pattern distinguishability and organize in the machine the directed selection of transformations*, then after a certain number of selection operations the quality of the transformations will be brought to the required level of perfection.

4. Finding the Generalized Portrait

According to inequality (1), the generalized portrait $^*\Phi_i$ is the center of a sphere in a certain functional space within which all points belonging to a certain subset F_i of the pattern $^*\Phi_i$ fall, and within which no point of the subset F_j of the system containing the pattern $^*\Phi_j$, falls.

Let us consider the set η of points s_i, for which the inequality

---

* i.e., remove from the initial set those transformations such that in the new structure the order of distinguishability be increased, and add new transformations taken at random from a certain set U.

$$(f_i s_i) \geqslant 1, \quad (f_j s_i) \leqslant 1. \tag{3}$$

is valid.

If the set η is not empty, the following propositions* hold.

1. There exists an unique point $s_i^0 \in \eta$ such that

$$s_i^0 = \sum_k \alpha_k f_{ik} + \sum \beta_e f_{je}, \quad \alpha \geqslant 0, \quad \beta \leqslant 0,$$

where the coefficients α and β are nonvanishing only for those vectors f_i and f_j for which inequality (1) passes into equality.

2. The vector s_i^0 has minimum length among all vectors $s \in \eta$. This signifies that if for the generalized portrait we take the unit vector $\varphi_i = s_i^0 / \| s_i^0 \|$, the threshold of recognition $k_i = \min (f_i \varphi_i)$ will be maximum.

3. $\| s_i^0 \| = \sqrt{\Sigma \alpha_k + \Sigma \beta_e}$.

Correspondingly the threshold of recognition k_i for the generalized portrait

$$\varphi_i = s_i^0 / \| s_i^0 \|$$

is equal to

$$k_i = \frac{1}{\sqrt{\Sigma \alpha_k + \Sigma \beta_e}}.$$

4. The coefficients α_i, β_i can be found as the coordinates of a stable singular point of the system of equations

$$\dot{\alpha} = -\varepsilon \alpha_k + F_1 (1 - (\varphi_i f_{ik})), \quad \dot{\beta}_d = -\varepsilon \beta_d + F_2 (1 - (\varphi_i f_{jd})),$$

where

$$F_1(z) = \begin{cases} z & \text{for} \quad z > 0, \\ 0 & \text{for} \quad z < 0, \end{cases} \qquad F_2(z) = \begin{cases} z & \text{for} \quad z < 0, \\ 0 & \text{for} \quad z > 0. \end{cases}$$

The generalized portrait can be generated in the following way.

A description of the object, which in the first approximation is taken as the generalized portrait, is shown to the machine. If when a second object belonging to the same pattern is shown to the machine, it is not recognized by the machine. the description of this object is included in the generalized portrait generated as the second approximation. It is clear that the descriptions of recognized objects enter with zero weight into the generation of the generalized portraits.

During the learning process it is necessary to store in the machine memory not only the generalized portrait but those descriptions of objects which have taken part in the generation of this generalized portrait.

Using the concept of "generalized portrait" it is also possible to solve a number of other problems.

5. Certain Problems

The problem of autonomous learning may be formulated in the following manner.

Let there exist a certain set of objects. It is required to divide it into subsets such that in the image $\mathcal{F} \in U$ the descriptions of the objects satisfy inequality (1). If there are several such divisions, it is possible to require that division to be found for which the order of distinguishability is maximum.

With the discovery of the generalized portrait it is possible to solve the problem of the "shifting" image.

Let there exist a certain shifting image with description $f(x, t)$, which for $t \to \infty$ can be assigned to one of the

*The proofs of these propositions will be published separately.

patterns $^{*}\Phi_1, \ldots, {}^{*}\Phi_n$. It is necessary to determine as soon as possible to which of the patterns the shifting image belongs.

The solution of this problem reduces to the following extrapolation problem: there exist n functions $1_i(t) = (f(x, t)\varphi_i(x))$ $(i = 1, 2, \ldots, n)$, defined on the interval $[0, T]$. Determine which of the functions first (for the smallest value of t^*) passes the recognition threshold (which can be solved by existing methods).

LITERATURE CITED

1. F. Rosenblatt, Generalized perception of transformation groups. Cybernetics Coll. No. 4 [Russian translation] Izd. inostr. lit. (1962).
2. E. M. Braverman, Experiments in teaching a machine to recognize visual patterns. Avtomatika i telemekhan Vol. 23, No. 3 (1962).
3. M. M. Bongard, Modelling recognition processes on a digital computer. Biofizika, Vol. 4, No. 2 (1961).

All abbreviations of periodicals in the above bibliography are letter-by-letter transliterations of the abbreviations as given in the original Russian journal. *Some or all of this periodical literature may well be available in English translation.* A complete list of the cover-to-cover English translations appears at the back of this issue.
