ITMO UNIVERSITY

Pattern recognition techniques

Natalia Khanzhina

Senior Researcher

Machine Learning Lab

nehanzhina@corp.ifmo.ru

Outline

- What is an image?
- Pattern recognition
- Potentials approach

Supervised learning problem

X is **object set**, or input set;

Y is **label set**, or **answer set,** or output set;

$y : X \to Y$ is unknown **target function** (**dependency**).

$\{x_1, \ldots, x_m\} \subset X$ is **training sample**;

$y_i = y(x_i),\ i = 1, \ldots, m$ are **known values** of the function.

Problem: find $a : X \to Y$ is **solving function** (decision function), which approximate y on X. We are going to speak only about algorithms.

Features

$f_j : X \to D_j$, $j = 1, \dots, n$ are **features** or **attributes**.

$\big(f_1(x), \dots, f_n(x)\big)$ is feature description of an object x.
Object is its feature description.

Data is represented with a matrix

$$F = \|f_j(x_i)\|_{m\times n} = \begin{pmatrix} x_1^1 & \dots & x_1^n \\ \dots & \dots & \dots \\ x_m^1 & \dots & x_m^n \end{pmatrix}.$$

4 / 35

Supervised learning problem types

Classification:

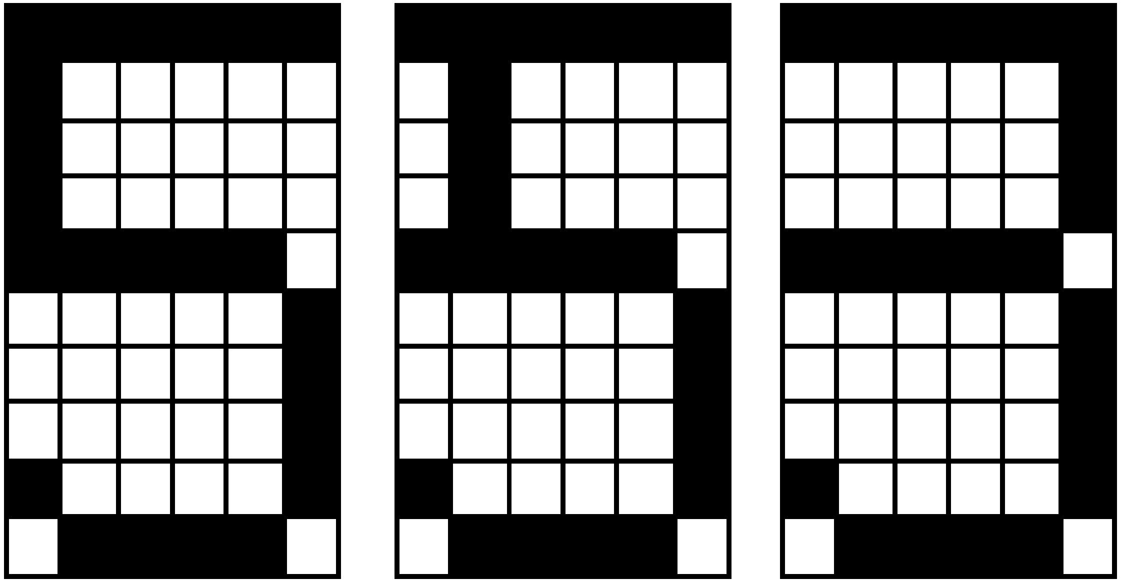
- $Y = \{-1, +1\}$ is binary;
- $Y = \{1, \dots, M\}$ is M non-overlapping classes;
- $Y = \{0, 1\}^M$ is M classes that can overlap.

Regression:

- $Y = \mathbb{R}$ or $Y = \mathbb{R}^m$.

What is an image?

What is an image?



Historical background

- The method of potential functions was originally proposed for solving the problems of learning machines for **pattern recognition.**
- The method was developed for:
  - the geometric formulation of the problem (when training is reduced to constructing in a certain space a surface separating two sets corresponding to two patterns);
  - probabilistic formulation of the problem (which is related to the restoration of a function characterizing the probability of an object belonging to a particular image).

Object Features

To characterize the elements of a set, various methods can be used:

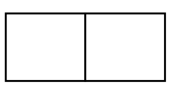
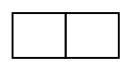
1. **quantitative** (when the value of the characteristic is obtained as a result of measuring a certain physical quantity [length, mass, temperature, electrical resistance, etc.]);
2. **probabilistic** (for which each element of the set is associated with the probability of occurrence of an event);
3. **binary** (when it comes to the presence or absence of any property).

The receptor space

<u>We introduce some concepts:</u>

- Let's start with a very simple field of receptors, consisting of 2 squares.
- Objects that can be represented on this field consist of only 4 flat figures.

These figures, together with their codes, are in the following table:

The receptor field	*Figures*	Codes
x_1 x_2 	x_1 x_2	x_1 x_2
		0 0
		1 0
		0 1
		1 1

The receptor space

- We introduce the rectangular coordinates x1 and x2 on the x1x2 plane.
- In this plane, which we will call the plane of receptors, each of the four figures will correspond to the vertex of the unit square.
- The coordinates of each vertex will be equal to the corresponding figures of the figure code (Fig. 1).

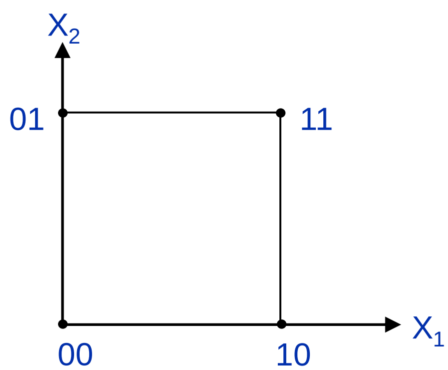


Fig. 1

Thus, the set of figures depicted on a two-element field of receptors can be associated with a set of points in two-dimensional space - on the plane of the receptors. Each of these points corresponds to a specific object from the set of figures.

Note: The distances in the receptor space are calculated as the square root of the sum of the squared differences of the corresponding coordinates. (For example, the distance between the first and last figure above will be equal).

The idea of the method of potential functions

Let in some receptor space X:

- *each object to correspond a point;*
- *classes of objects to correspond to disjoint areas;*

The task is to construct, according to the points shown and the information reported about them, a surface that separates these areas. Those. function that takes positive values at points from one area and negative values at points from the second area.

We introduce the function of two variables **K(x,y)**, where x and y are points of the space **X.**

If we fix the point **y=x***, then the function **K(x,x*)** becomes a function of the point of the space **X** and will depend on **x***.

An example of such a function in physics is the potential. By following this analogy, we call the function **K(x,y) a potential function**.

Selection of a potential function

Before applying the method of potential functions to the problem of pattern recognition, in general, it is necessary to select a receptor space in which any of the objects corresponds to a certain point.

If the recognition object is an image, then the space of n-dimensional points can be taken as the receptor space, where n is the number of pixels in the matrix represening the image

The choice of the receptor space X also determines the choice of the potential function.

Potentials in the receptor space

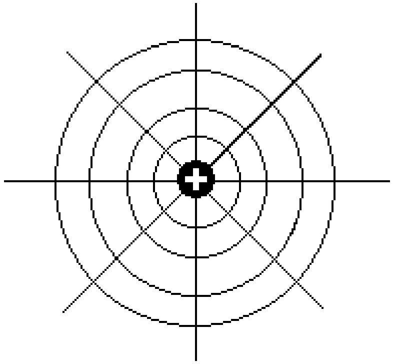


Fig. 2

A point electric charge located separately in a homogeneous medium forms the electric field shown in Fig. 2.

The radial lines in this figure are the lines of force of the electric field.
Concentric circles are lines of potential.

For the one shown in Fig. 2 the potential at each point of space is given by:

$$p = a \frac{q}{r^2},$$

where a is some constant coefficient, q is the value of the charge, r is the distance from the given point to the charge.

Potentials in the receptor space

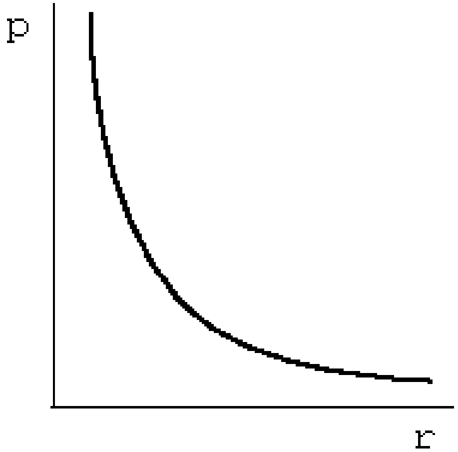


Fig. 3

The change in potential with distance from the charge is represented by the curve shown in Fig. 3.

This curve and formula (3.1) make it possible, knowing the magnitude of the charge and the distance from the charge to a given point of space, to determine the value of the potential at this point.

Thus, the potential can serve as a measure of the distance of the point to the charge.

Potentials in the receptor space

- When a field is formed by several charges, the potential at each point of the field is equal to the sum of the potentials created at this point by each of the charges.
- If the charges forming the field are located in a compact group, the potential will have the largest value inside the group of charges.

The potential of any point of space in this case will determine the distance of this point to the entire group of charges.

Potentials in the receptor space

Suppose that there are two isolated compact groups of charges in space. In one group, the charges are negative, in the other are positive.

In Fig. 4 shows the potential distribution in the vicinity of these charges, the upper curve representing the potential formed by positive and the lower one by negative charges.

Some point is spatially closer to that set of charges whose potential at this point is larger in absolute value.

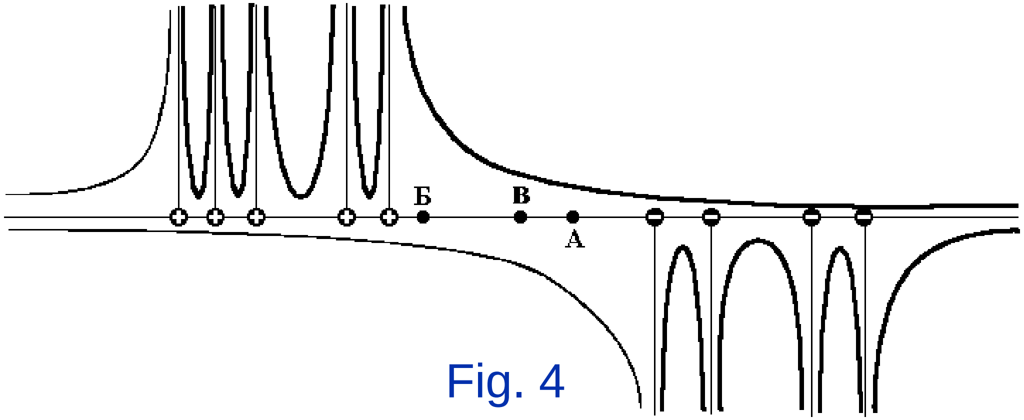


Fig. 4

Potentials in the receptor space

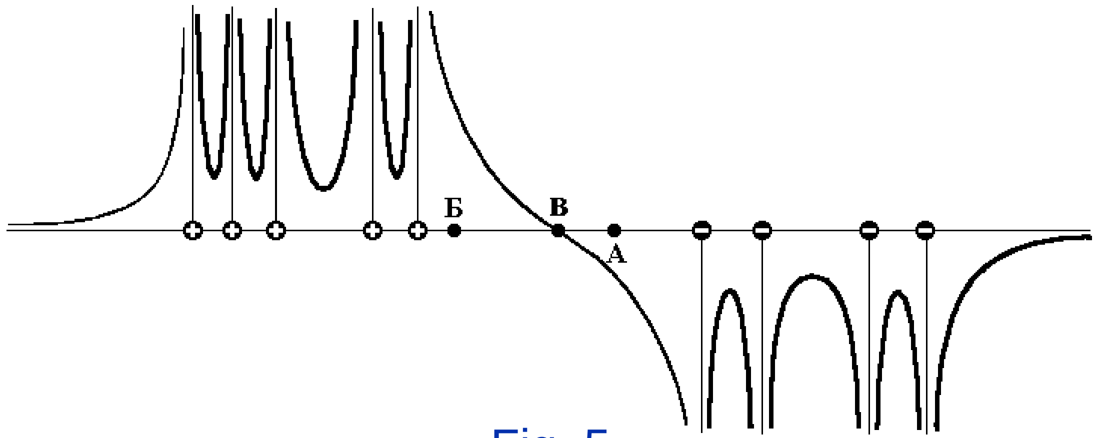


Fig. 5

In Fig. 5 potentials, formed by both sets of charges, are summed. In this case, the points can be assigned to one or another set by the sign of the total potential at these points. The point at which the potential curve passes through zero lies on the boundary of two sets.

Potentials in the receptor space

We apply the above representations for points in the receptor space.

We associate with each point some function analogous in form to the electric potential, including the maximum at this point and decreasing in all directions from it. This function can be, for example, a function:

$$\varphi(R) = \frac{1}{1 + \alpha R^2},$$

Where $\boldsymbol{\alpha}$ is the coefficient of the decay rate $\boldsymbol{\varphi}$, $\boldsymbol{R}$ is the distance, determined in some way, between the source point and the point at which the potential is calculated. (For $\boldsymbol{R}$ we can take, for example, the Euclidean distance between points or the Hamming distance).

Potentials in the receptor space

Suppose that the source is a group of points, for example, the set obtained during training corresponding to the image ***a***:

Then the total potential divided by the number of sources will determine the distance of the given point to the whole image (dividing by the number of sources is needed so that the result does not depend on the number of sources).

Based on this representation, as well as on the hypothesis of compactness, we can make the following assumption:

Let the memory of the machine fix two sets of points in the learning process, the corresponding images of ***a*** and ***b*** types, after which a point to be recognized appeared.

We define the average potentials created at this point by the sets ***a*** and ***b***. The new point will refer to the image, whose potential at this point is greater than the others.

Potentials in the receptor space

The simplest recognition algorithm, based on the method of potentials:

Training.

In the learning process, the codes of all the points appear and the indication of which image each point belongs to.

Recognition.

a) For the point to be recognized, the potentials of each image are calculated, that is, the sums

$$\Phi_a = \frac{1}{n_a}\sum_{i=1}^{i=n_a}\varphi_{ai}, \quad \Phi_b = \frac{1}{n_b}\sum_{i=1}^{i=n_b}\varphi_{bi}, \quad \ldots\ldots\ldots\ldots \quad \Phi_m = \frac{1}{n_m}\sum_{i=1}^{i=n_m}\varphi_{mi},$$

where **a, b, …,m**- patterns

$\mathbf{n_a, n_b, \ldots, n_m}$- the number of points of each image perceived by the machine in the learning process,

$\varphi_{ai} = \dfrac{1}{1+\alpha R_{ai}^2}$ - The potential formed at the recognized point by the i-th point of the image,

$\varphi_{bi} = \dfrac{1}{1+\alpha R_{bi}^2}$ - the potential formed at the recognized point by the i-th point of the image, and so on to

$\varphi_{mi} = \dfrac{1}{1+\alpha R_{mi}^2}$ - potential formed at the recognized point by the i-th point of the image m.

Potentials in the receptor space

b) A comparison is made and the recognizable point refers to the image that creates the greatest potential at this point.

In the simplest case, when two images a and b are to be recognized, they are distinguished by the sign of the function:

$$\Delta\Phi = \Phi_a - \Phi_b,$$

which takes positive values in the region of one image, negative in the region of the other and passes through zero near the boundary between them.

ITMO UNIVERSITY

Potentials in the receptor space

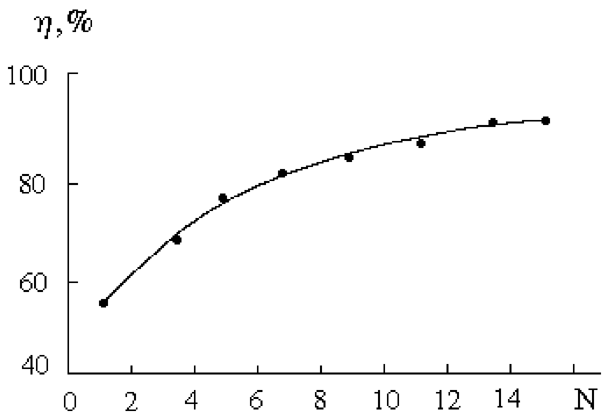


Fig. 6

The graph of Fig. 6 shows the results of experiments on the recognition of ten digits (0,1, ..., 9) using the algorithm described.

The ordinate of the graph shows the percentage of correct answers, along the abscissa - the number N - the number of images of each figure presented in the learning process.

Note: It was possible to achieve a fairly high percentage of correct answers (more than 85% on average for all ten digits). However, further increase in the reliability of recognition was not possible. Starting from N = 13, the recognition reliability curve becomes practically parallel to the abscissa axis, and further increase in the learning duration does not increase the recognition reliability.

Potentials in the receptor space

Cases of misidentification of "known" points actually occurred when testing the described algorithm.

The algorithm can be improved with the following operation: after the machine is presented with all the figures selected for training, the machine is forced to recognize the same figures and monitor the correctness of their work.

If an error occurs, that is, the figure of one of the images is treated by the machine to another image, then it is written in the machine's memory that the "weight" of the corresponding point should be increased by some amount, for example, by one. (This means that in the future the potential created by this point will be doubled).

Potentials in the receptor space

Table 1 shows the results of experiments with an improved algorithm. For comparison, the results of the simplest algorithm are given here. Letter ***N*** denotes the number of images of each digit presented in the training process. At the same time, the greatest increase in recognition reliability occurs in those images that are recognized worst of all.

	Number of objects used in the experiments	Percent of correct answers			
		The simplest algorithm		Improved algorithm	
		N=12	N=21	N=12	N=21
1	122	100	100	100	100
2	157	99,5	100	99,5	100
3	176	88,2	92,0	88,0	94,2
…	…	…	…	…	…
8	134	76,2	71,7	77,1	82,1
9	140	42,0	49,3	51,2	65,3
0	180	94,0	94,0	94,3	93,0
В среднем для всех образов	**-**	**85,0**	**85,0**	**86,1**	**89,3**

Potentials in the receptor field

Using potentials in the receptor space, we took as the distance between two points the number of distinguished bits in the codes of the corresponding figures, or the Euclidean distance between the points. These values, however, do not always correctly reflect the difference between the figures.

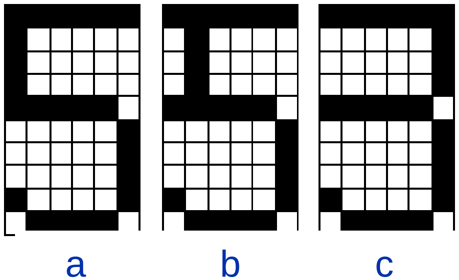


Fig. 10

In Fig. 10, and the five shown on the field of receptors is shown. Moving its vertical line one square to the right, we get the "worst", but still undoubted five (Figure 10, b). If you move the vertical line to the right by five cells, you get an equally certain triple (Figure 10, c).

But both figures differ from the original by the same number of cells in the receptor field, and in the codes of these figures the same number (six) digits does not coincide with the code of the original five. In other words, the point of the five of Figure 10, b and the point of three of Figure 10, c are at equal distances from the original figure.

Potentials in the receptor field

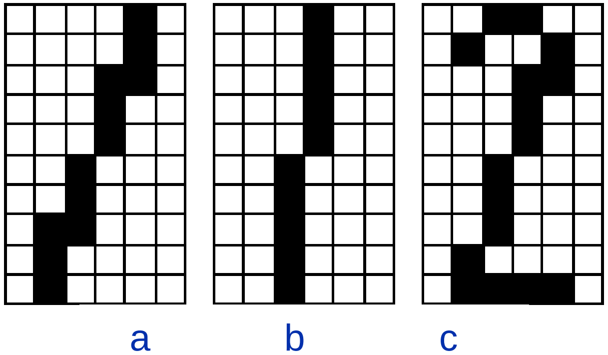


Fig. 11

In Fig. 11, a figure is shown, which can be taken with sufficient confidence as a unit. The “one” of Fig. 11, b and the “two” of Fig. 11, in differ from it by the same number (ten) digits. Again we are dealing with the case where different figures are from the original figure at equal distances.

Potentials in the receptor field

These examples show that the encoding method adopted by us does not convey the differences in figures well enough.

The reliability of recognition can apparently be improved if one resorts to such a method of coding, which takes into account the amount of displacement of the elements of the figures along the field of the receptors.

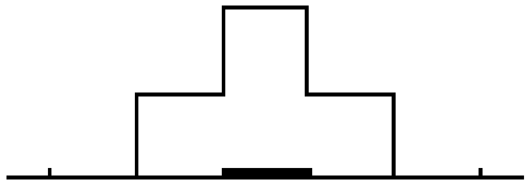


Fig. 12

<u>The following improvement in the coding method may be proposed:</u>
We associate with each firing element of the field of receptors a certain function equal to one on this element, on adjacent elements and to zero - on the sites of the receptor field, distant from the given one by more than one unit (Fig. 12)

Potentials in the receptor field

Suppose that each digit of the figure code, except for zero and one, can take other, including fractional, values equal to the potential value on the corresponding element of the receptor field.

Then, for example, a figure consisting of one excited element on the 18-element element of the receptor field (Figure 13, a) appears in the form of a step function, shown in Fig. 13, b.

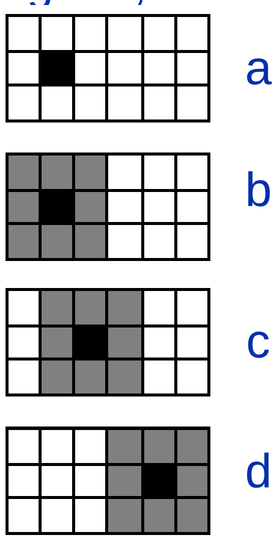


Fig. 13

The code for this shape will look like this:

$$\frac{1}{2}\frac{1}{2}\frac{1}{2}000\frac{1}{2}1\frac{1}{2}000\frac{1}{2}\frac{1}{2}\frac{1}{2}000.$$

Let’s move the firing element one cell to the right (Fig. 13, c). Get the figure shown by the code

$$0\frac{1}{2}\frac{1}{2}\frac{1}{2}000\frac{1}{2}1\frac{1}{2}000\frac{1}{2}\frac{1}{2}\frac{1}{2}00.$$

Potentials in the receptor field

We define the Euclidean distance between the figures in Fig. 13, b and c, as the square root of the sum of the squares of the differences in the values of each digit of the code. Obtained value is √2.

Let us move the excited element two more cells to the right. A figure will be obtained (Fig. 13, d), represented by a code:

$$000\frac{1}{2}\frac{1}{2}\frac{1}{2}000\frac{1}{2}1\frac{1}{2}000\frac{1}{2}\frac{1}{2}\frac{1}{2}.$$

The distance between this figure and the figure in Fig. 13, b is equal to $\sqrt{6} = 2.45$.

As we see, with the method of coding described above, greater distances in the receptor space correspond to larger distances between the elements of the figures on the receptor field.
With the "old" method of coding the distance between any two figures in Fig. 13, b, c and d would be the same and equal.

Potentials in the receptor field

Let's spread the coding method outlined above into more complex figures:

<u>We will use the following rule</u>: every firing element of the receptor field has an "own" potential equal to unity, and increases the potentials of all (including the firing ones) elements adjacent to it horizontally, vertically and diagonally.

2	2.5	2	2	2	1.5
2.5	2.5	1.5	1.5	1.5	1
2	1.5	0	0	0	0
2.5	2.5	1.5	1.5	1	0.5
2	2.5	2	2	2	1
1	1.5	1.5	1.5	2	2
0	0	0	0	1.5	2
0.5	0.5	0	0	1.5	2
1.5	1.5	1.5	1.5	2	2
1	2	2	2	2	1

2	2.5	2.5	2	2	1.5
2	3	2.5	1.5	1.5	1
1.5	2	1.5	0	0	0
2	3	2.5	1.5	1	0.5
2	2.5	2.5	2	2	1
1	1.5	1.5	1.5	2	2
0	0	0	0	1.5	2
0.5	0.5	0	0	1.5	2
1.5	1.5	1.5	1.5	2	2
1	2	2	2	2	1

1.5	2	2	2	2.5	2
1	1.5	1.5	1.5	2.5	2.5
0	0	0	0	1.5	2
1	1.5	1.5	1.5	2	2
1.5	2	2	2	2	1.5
1	1.5	1.5	1.5	2	2
0	0	0	0	1.5	2
0.5	0.5	0	0	1.5	2
1.5	1.5	1.5	1.5	2	2
1	2	2	2	2	1

Fig. 14

Then the figures shown in Fig. 10, a, b and c, will correspond to that shown in Fig. 14 distribution of potentials over the field of the receptors.

The distance between the two fives (see Fig. 10, a and b) will then be 2.25, and the distance between the five pins in Fig. 10, a and the three of Fig. 10, c - 5.22.

Applying this coding method to Fig. 11, you can verify that the distance between ones is 3.3, between the left one and the two is 6.1, between the average one and the two is 4.87.

Potentials in the receptor field

When potentials are used on the receptor field, the distance between objects of different images turns out to be much larger than the distance between two objects of the same image, although with the "old" coding method these distances were the same.

This is because the new encoding method better reflects the mutual distance of the elements of the figures on the receptor field, and therefore better conveys the differences between the figures.

Potentials in the receptor field

Image	Percent of correct answers	
	The simplest algorithm	Algorithm with potentials at the receptor field
1	100	100
2	99,5	100
3	88,2	100
4	90,0	97,5
5	92,1	100
6	86,2	98,5
7	80,5	87,5
8	76,2	100
9	42,0	64,0
0	94,0	91,9
On average for all images	**85,0**	**94,0**

The table gives a comparison of the results of the simplest algorithm and the algorithm with the potentials at the receptor field at N = 12.

The average reliability of recognition after the introduction of potentials on the receptor field increased by 9% and reached 94%.

For individual images (recognized worse than others), the recognition reliability increased to 20-24%.

Sources

English:

1. M. A. Aizerman, E. M. Braverman, and L. I. Rozonoer, "Theoretical foundations of potential function method in pattern recognition," Avtomat. i Telemekh., vol. 25 (1964): 917-936
2. Meisel, William S. "Potential functions in mathematical pattern recognition." IEEE Transactions on Computers 100.10 (1969): 911-918.

Russian:

1. Айзерман М.А., Браверман Э.М., Розоноэр Л.И. «Метод потенциальных функций в теории обучения машин». Москва, 1970.- 384 с.

2. Васильев В. И., Коноваленко В. В., Горелов Ю. И. «Имитационное управление неопределенными объектами» Киев, 1989.- 216с.