I. V. Denisov, O. T. Kamenev, A. Yu. Kim, Yu. N. Kulchin, A. V. Panov. — *Optical Memory & Neural Networks (Information Optics)*, v. **12**, No. 3, pp. 165–172 (2003).

# Neural data processing method for fiber-optic distributed measuring systems

I. V. Denisov [1,2], O. T. Kamenev [1,2], A. Yu. Kim [1], Yu. N. Kulchin [1,2], A. V. Panov [2]

[1]*Far Eastern State Technical University, 10, Pushkinkaya st., Vladivostok, 690950, Russia*

[2]*Institute for Automation and Control Processes, 5, Radio st., Vladivostok, 690041, Russia*

**ABSTRACT**

We developed a method of output signal processing for distributed fiber-optical measuring systems. This method is based on neural-like principles of data processing. A mathematical model of the three-layered perceptron was used to reconstruct the physical field distribution measured by a distributed interferometric system. We proposed an algorithm that is based on the committee method of recognition of the dynamic object detected by a fiber-optic measuring system.

**Keywords**: fiber optic, sensor, measurement, neural network, perceptron, tomography, pattern recognition, committee method.

## INTRODUCTION

The study of natural and simulated physical objects and fields, distributed in great areas, require the use of an informational-measuring system in which data gathering is carried out by means of a distributed measuring system. A highly promising type of measuring system is the distributed fiber-optical measuring system [1]. It consists of a set of distributed fiber-optical measuring lines, which sensors detect exterior physical actions in an area. Output signals of such an interoferometric system are formed by tomographic principles of data gathering. The intensity of light transmitted through a fiber-optic measuring line and that is detected is proportional to the integral action of a physical field on the sensors of this line [2]. Thus the optical signals on outputs of measuring lines contain the information on the parameters of the physical field explored. The tomographic output data of a measuring system represent the multivariate arrays of the rapidly varying analog information. The problem of creating high-speed computers is real for data processing in real time.

The use of conventional digital computers for these purposes is limited to their information capacity, processing speed and other restrictions, bound with a serial principle of data processing. The application of neural networks has significantly increased the speed of tomographic data processing. Moreover the adaptivity of neural networks allows us to obtain immunity in a processing system to changes of input data called environmental influence. Thus the application of the neural-like processors effectively allows us to reconstruct the physical field distributions measured by a fiber-optic distributed system.

If measurements are carried out with the sensors which have a linear system performance, to reconstruct a desired distribution function, it is enough to use a two-layered linear perceptron, in which learning is carried out using a method of error back-propagation. The two-layered perceptron can be realized on the basis of optics [2] or as a program for a personal computer [3].

Unfortunately, the measurements usually have errors of the same order of neural network reconstruction and fiber-optic interferometric sensors perform as a nonlinear system. Both these both factors lead to an error of reconstruction of a distribution function being required [4]. So it is necessary to develop new principles of data processing utilizing a nonlinear analog neural network with more than two layers.

The information reconstructed by a neural network can be used for object attribute information system definition. The observed object class can be determined in the case of building a recognition algorithm. Thus one can obtain an informational-measuring system registering object quite accurately, describing its behavior and classifying it.

The problem of selection of the information attribute systems and their definition can be solved on the basis of analysis of distribution of the physical field detected by a fiber-optic measuring system. The problem of discriminant analysis can be solved by a committee method [5].

The committee method in this problem can be quite simply realized and can quickly divide numerous objects into classes due to a simple logic of committee construction design proposed. At this stage, it is not important to divide the registered objects into patterns of some permissible set by means of teaching on the basis of a known teaching pattern, but on the contrary, it is essential to treat the problem on the other side and for known images to realize the process of reference of any object in one of them. Such an approach to the solution of the problem allows us to build an algorithm and implement it as a computer program.

## DATA PROCESSING BY NEURAL NETWORK

If the fiber-optic distributed measuring system [6] gathers data through a tomographic method, then it might have the architecture shown in Fig. 1. The measuring network consists of the measuring lines stacked along three directions. So this network forms a square lattice. Figure 1 shows the architecture of the measuring system with configuration $4 \times 4$. If a physical field explored does not act on all the measuring lines but only on some sensors then output signals of the measuring system $x_1 \dots x_{15}$ contain the information on values of the field parameters $y_1 \dots y_{16}$ at places where three lines are intercrossed.

The output tomographic data of the measuring system are determined by expressions:

$$\begin{aligned} x_1 &= f\left(y_1 + y_2 + y_3 + y_4\right); \\ &\vdots \\ x_{11} &= f\left(y_2 + y_7 + y_{12}\right); \\ &\vdots \\ x_{15} &= f\left(y_{13}\right); \end{aligned}$$

where $f$ is nonlinear transfer function defined by the principle of operation of a measuring line, $s_k$ is the corresponding sum of parameters explored $y_{\mathrm{i}}$.

The problem of finding of the solution $y_i$ is incorrect because the number of equations is less than quantity of parameter values reconstructed. This tomographic problem is traditionally solved by iterative methods. When function $f$ is linear the neural network such as a two-layered perceptron can be used to solve this problem [2,3]. However, if the measuring lines are single-fiber low-mode interferometers then

$$f(\zeta) = \mathrm{A} + \mathrm{B}\cos(\mathrm{C}\cdot\zeta), \tag{1}$$

where A, B, C are constants, $\zeta$ is the sum of $y_{\mathrm{i}}$ along a line. This function $f(\zeta)$ can be approximated by linear dependence within a certain range of $\zeta$. So in this work, we chose the perceptron with the nonlinear hidden layer since such networks have universal approximation capability [7,8]. A three-layered perceptron allows to solve tomography problem for function (1) and so it can be used to process output signals of distributed fiber-optic measuring systems.

The architecture of three-layered neural network is shown in Fig. 2. The neurons of the first layer serve as network inputs and feed data from the measuring system to the next layer. In order to reconstruct a physical field distribution from the area $n \times n$, where $n$ is the size of the lattice, the first layer should contain $4n - 1$ neurons. In this paper, we suppose that the known scheme of stacking of fiber-optic lines on three directions is applied (see Fig. 1).

In this case, each neuron of input layer corresponds to the certain fiber-optic measuring line and output potential of this neuron $x_k$ is proportional to intensity of light detected. The second (hidden) layer processes the following non-linear transformation:

$$s_j = \tanh\left(\sum_{k=1}^{4n-1} \bar{w}_{jk} x_k\right),\ k = 1,\ldots 4n-1, \quad (2)$$

where $x_k$ are states of neuron inputs being signals from measuring lines, $s_j$ are states of outputs and $\bar{w}_{jk}$ are synapses. The second layer contains $n^2$ neurons and the matrix of connections $\bar{\mathbf{w}}$ has $n^2(4n-1)$ elements.

The output layer of the neurons takes the linear transformation:

$$y_i = \sum_{j=1}^{n^2} w_{ij} s_j\ ,\ i = 1,\ldots n^2, \quad (3)$$

where $y_i$ are states of third layer neuron outputs, $w_{ij}$ are synapses of third layer. This layer consists of $n^2$ neurons so the matrix of connections $\mathbf{w}$ has $n^4$ elements. The accuracy of the physical field distribution reconstruction is determined by the neural network training error (objective function). The following expression was used as an objective function:

$$D = \frac{1}{2}\sum_{\mu,i}\left(y_i^\mu - \tilde{y}_i^\mu\right)^2, \quad (4)$$

where $\mu$ is superscript indicating number of learning pattern, $\tilde{y}_i$ are output states of the neural network for some learning pattern. The learning patterns $\tilde{y}_i^\mu$ have a peak-alike or smooth distributions on the surface, and were randomly generated by numerical modeling or obtained from experimental data. For each training pattern $\tilde{y}_i^\mu$ are proportional to the parameter of the field investigated and $\tilde{x}_i^\mu$ are proportional to the integral signal measured. In Eq. (4) $y_i$ are calculated by the expression:

$$y_i^\mu = \sum_j w_{ij} \tanh\left(\sum_k \bar{w}_{jk} \tilde{x}_k^\mu\right).$$

We used error back-propagation for the network training, so we had to minimize $D$ with respect to $N = n^4 + n^2(4n-1)$ dimensional vector $\boldsymbol{\omega} = \{\bar{w}_{jk},\ w_{ij}\}$. We utilized Broyden-Fletcher-Goldfarb-Shanno (BFGS) method to minimize the object function since this gradient method is rather simple and fast.

The multidimensional surface of the objective function is complex and the gradient minimization procedure may stick at a local minimum or valley. So we applied the "jog of weights" technique between the series of the gradient optimization in order to avoid local minima of $D$. The "jog of weights" means the random addition ~ $0.1\omega_i$ to the each component of $\boldsymbol{\omega}$. The length of these series was usually chosen as 20000 iterations of the minimized function. We utilized the randomly generated set of $\omega_i$, $i = 1,\ldots N$ as the initial vector for the optimization. The minimization procedure was completed after a certain iteration count, typically several hours of computations on the Pentium III 800 MHz processor.

Examples of distributions modeled with different $n$ reconstructed by the neural network and initially unknown to it are shown in Figs. 3 and 4. The neural networks were learned independently in all the cases with the similar randomly generated patterns. In this simulation, we suggested that each sensor had a randomly generated but permanent weight in range 0.5–1, i. e. the sensibility of modeled sensors varied. The number of the training patterns in this simulation was about 100. One can see from Figs. 3 and 4 that the neural network sufficiently accurately reconstructs an unknown pattern and can be used in practice.

Next we utilized our neural network for reconstruction of experimental data from the mock-up of the fiber-optic system measuring the field of acoustic oscillations [6]. In this experiment the each $\tilde{y}_i$ was proportional to the amplitude of the speaker acoustical oscillation. Figure 5 shows original (a) and reconstructed by the neural network (b) acoustic field distributions. This pattern was absent in the set of training patterns. We utilized 31 learning patterns to train this neural network. One can see from Fig. 5 that the neural network rather accurately reconstructs the unknown distribution. The form of the distributions was reconstructed exactly and the values have errors of about 20%. The errors were mainly introduced by the noise in the measuring lines.

## PATTERN RECOGNITION USING COMMITTEE METHOD

Now when the data measured is reconstructed, we can proceed to information recognition. The information about the number of peaks, the intensity of a peak, its variance and the velocity of its movement along with other attributes can be extracted from a field distribution reconstructed by a neural network.

Let the set $M$ consist of elements which can be called permissible objects and exist as vectors in space $\Re^n$. Let $M \in \Re^n$. We can refer to $M$ as a permissible set. Let $M_1, M_2, \ldots, M_m$ exist and for them we have

$$M = M_1 \cup M_2 \cup \ldots \cup M_m,\ M_i \cap M_j = \varnothing,\ i \neq j. \tag{5}$$

In general, the solution of the problem of discriminant analysis has two steps:

1) the algorithm is made and possible set is divided into patterns using teaching pattern $\tilde{M}$ ( $\tilde{M} \cap M_i = \tilde{M}_i \neq \varnothing$);

2) assuming classification of the $M_c$ elements set is known, the algorithms are applied to object set $M_c$ ($M_c$ is a control pattern $M_c = \{c_1, c_2, ...., c_k\}$ for which $M_c \cap \tilde{M} = \varnothing$ is performed) and recognition quality is checked.

Since information attributes are determined on the basis of distribution analysis of the value characterizing the physical field and images are assumed to be known ones we use another approach: 1) the algorithm is built which refers objects to the known images; 2) the recognition quality is tested by means of control pattern $M_c$.

**Description of the recognition algorithm**

Consider the recognition algorithm in detail for object of recognition $\mathbf{c}_k = \left[x_{k1}^c, x_{k2}^c, \ldots, x_{kn}^c\right]$, $n$ is number of the object attributes, $\mathbf{c}_k \in M$, $1 \le k \le q$, $k \in Z$ and for $\forall\, q$ ($q$ is the number of the control objects). Let the sets $M_1, M_2, \ldots, M_m$ satisfying Eq. (5) be the classes of object recognition.

Assume vectors $\overline{\mathbf{X}}_1, \overline{\mathbf{X}}_2, \ldots, \overline{\mathbf{X}}_m$ of sets $M_1, M_2, \ldots, M_m$ are centers of classes and $\mathbf{D}_1, \mathbf{D}_2, \ldots, \mathbf{D}_m$ are corresponding ranges of variation of the attributes :

$$\overline{\mathbf{X}}_1 = \left[x_1^1, x_2^1, \ldots, x_n^1\right],\ \overline{\mathbf{X}}_2 = \left[x_1^2, x_2^2, \ldots, x_n^2\right],\ \ldots,\ \overline{\mathbf{X}}_m = \left[x_1^m, x_2^m, \ldots, x_n^m\right],$$

$$\overline{\mathbf{X}}_1 \in M_1,\ \overline{\mathbf{X}}_2 \in M_2,\ \ldots,\ \overline{\mathbf{X}}_m \in M_m;$$

$$\mathbf{D}_1 = \left[d_1^1, d_2^1, \ldots, d_n^1\right], \mathbf{D}_2 = \left[d_1^2, d_2^2, \ldots, d_n^2\right], \ldots,\ \mathbf{D}_m = \left[d_1^m, d_2^m, \ldots, d_n^m\right].$$

Define $\mathbf{R}_k^1 = \left[r_{k1}^1, r_{k2}^1, \ldots, r_{kn}^1\right]$, $\mathbf{R}_k^2 = \left[r_{k1}^2, r_{k2}^2, \ldots, r_{kn}^2\right]$, …, $\mathbf{R}_k^m = \left[r_{k1}^m, r_{k2}^m, \ldots, r_{kn}^m\right]$, where each component of these vectors for $1 \le l \le n$, $l \in Z$; $1 \le k \le q$, $k \in Z$ is given by formula:

$$r_{kl}^i = \left|x_l^i - x_{kl}^c\right|,\ 1 \le i \le m,\ i \in Z. \tag{6}$$

Then let us compare given $r_{kl}^i$ with $d_l^i$ correspondingly on the rule:

$$\beta_{Ak}^l\left(x_{kl}^i\right) = \begin{cases} 1, & \text{if } r_{kl}^i \le d_l^i; \\ 0, & \text{if } r_{kl}^i > d_l^i. \end{cases} \tag{7}$$

After determining the coordinates of the information vector $\boldsymbol{\beta}_{Ak} = \left[\beta_{Ak}^1, \beta_{Ak}^2, ...., \beta_{Ak}^n\right]$ the information vector $\boldsymbol{\alpha}_A$ is formed for $\mathbf{c}_k$ as follows:

Test the condition:

$$\alpha_{Ak}^{i}\left(\mathbf{c}_k\right)=\begin{cases}1, \text{ if } & \sum_{l=1}^{n}\beta_{Ak}^{l}>\frac{n}{2};\\ 0, \text{ if } & \sum_{l=1}^{n}\beta_{Ak}^{l}\le\frac{n}{2}.\end{cases} \tag{8}$$

Under this condition, if more than $\frac{n}{2}$ attributes of instance object correspond to one of set $M_i$, then this object belongs to this set.

Then the information vector $\boldsymbol{\alpha}_{Ak}$ of the $\mathbf{c}_k$ object for $m$ classes is defined by the formula:

$$\boldsymbol{\alpha}_{Ak}=\left[\alpha_A^1\left(\mathbf{c}_k\right),\alpha_A^2\left(\mathbf{c}_k\right),\ldots,\alpha_A^m\left(\mathbf{c}_k\right)\right],$$

where $\alpha_{Ak}^{i}\left(\mathbf{c}_k\right)\in\{0,1\}$. After calculation one can conclude that:

$$\begin{aligned}&\text{if } \alpha_A^i\left(\mathbf{c}_k\right)=1, \quad \text{then } \mathbf{c}_k\in M_i;\\ &\text{if } \alpha_A^i\left(\mathbf{c}_k\right)=0, \quad \text{then } \mathbf{c}_k\notin M_i.\end{aligned} \tag{9}$$

The program implementing this algorithm was realized, the results of 2 pattern separation between two classes ($m=2$) with six information attributes of the object ($n=6$) are shown in Table 1.

## CONCLUSIONS

In this paper, we developed a method of output signal processing for distributed fiber-optical measuring systems. This method is based on the application of a three-layered perceptron. A mathematical model of the perceptron was realized as a computer program. The program was used to reconstruct a physical field distribution measured by a fiber-optic distributed system. We showed that the neural network successfully solved the tomographic problem. The algorithm that used the committee decision method and which recognized the information detected by the measuring system and was reconstructed by the perceptron was proposed.

## AKNOWLEDGEMENTS

This work was supported by grant No. 02-01-00710 of the Russian Foundation for Basic Research.

## REFERENCES

1. Yu. N. Kulchin, O. B. Vitrik, O. V. Kirichenko et al., “Tomography Methods for Vector Field Study Using Space-Distributed Fiber Optic Sensors with Integral Sensitivity”. — *Fiber and Integrated Optics*, 17 (1), pp. 75–84, 1998.
2. Yu. Kulchin, O. Vitrik, Yu. Petrov, O. Kirichenko, O. Kamenev, R. Romashko and I. Denisov “Holographic neural network for processing of signals of distributed optical fiber measuring networks with the topographic principle of data gathering”. — *Optical Memory and Neural Networks*, 6 (2) pp. 149–156, 1997.
3. Yu. N. Kulchin, O. T. Kamenev, I. V. Denisov et al., “Computer Neural Networks for Processing of Optical Tomography Information”. — *Pacific Science Review*, 1, pp. 1–4, 1999.
4. Yu. N. Kulchin, I. V. Denisov, O. T. Kamenev, “Neural processing system for optical information measuring systems”. — *Proc. SPIE*, v. 4357, pp. 109–117, 2001.
5. Vl. D. Mazurov, “*Committee method in optimization and classification problems*”, Nauka, Moscow, 1990, 248 p.
6. Yu. N. Kulchin, O. T. Kamenev, Yu. S. Petrov, M. Kantur, “Fiber optical distributed network for dynamic deformation measuring”. — *Proc. SPIE*, v. 5129, pp. 68–73, 2003.
7. K. Funahashi, “On the approximate realization of continuous mapping by neural networks”. — *Neural Networks*, 2, pp. 183–192, 1989.
8. K. Hornik, M. Stinchcombe, H. White, “Multilayer feed-forward networks are universal approximators”. — *Neural Networks*, 2, pp. 359–366, 1989.

# FIGURES

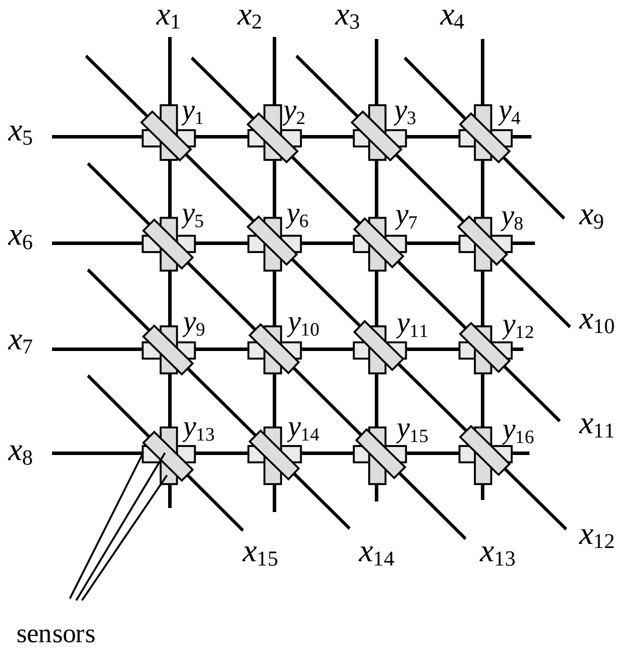



Fig.1. Architecture of the distributed measuring network with configuration $4 \times 4$.

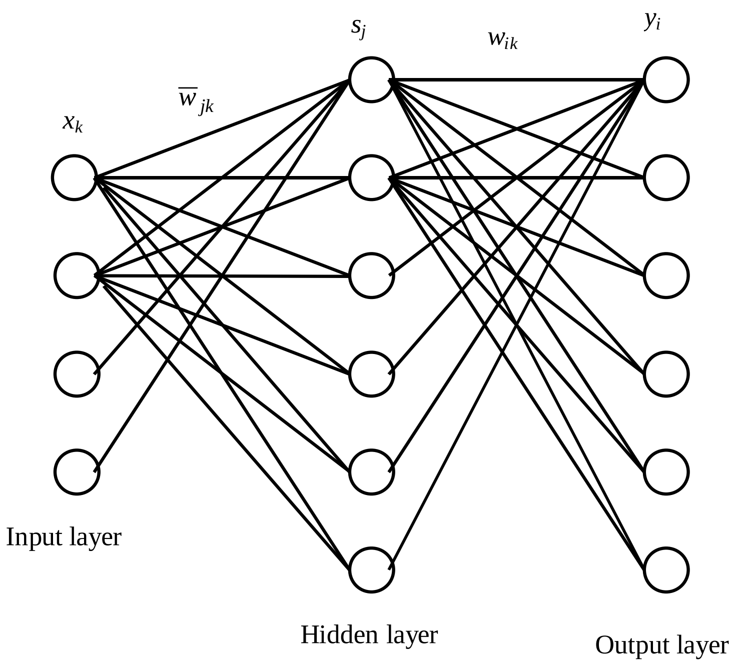
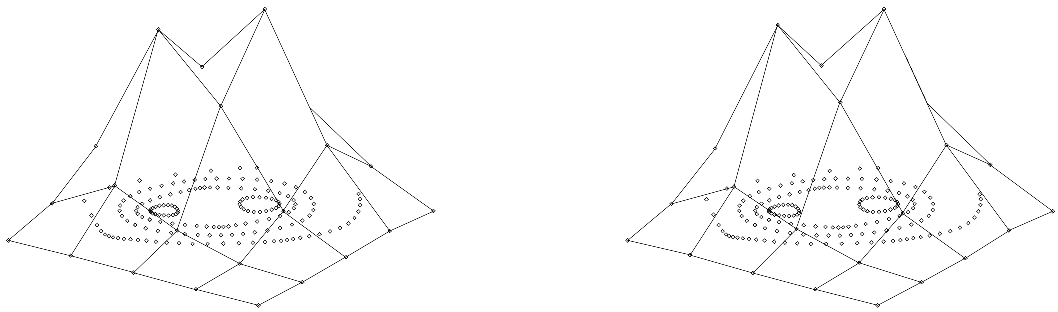
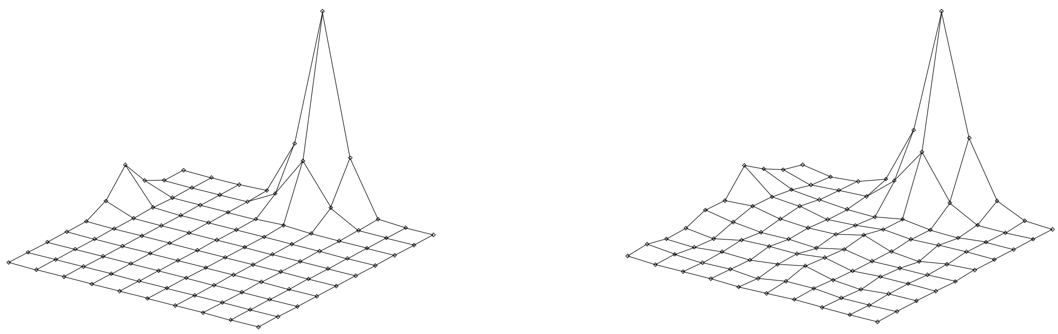



Fig. 2. Schematic diagram of a three-layered neural network.



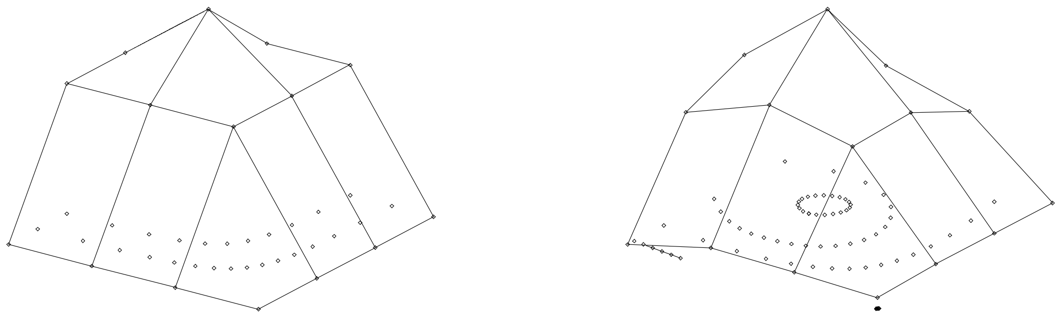
(a) (b)

Fig. 3. The original (a) and reconstructed by the neural network (b) distributions of the modeled physical field with $n = 5$.



(a) (b)

Fig. 4. The original (a) and reconstructed by the neural network (b) distributions of modeled physical field with $n = 10$.



(a) (b)

Fig. 5. The original (a) and reconstructed by the neural network (b) unknown acoustic field distribution. The original distribution was detected with the mock-up of the interferometric fiber-optic measuring system [6].

**TABLE**

Table 1. An example of the separation of two patterns $\mathbf{c}_1$ and $\mathbf{c}_2$ between two sets $M_1$ and $M_2$. $x_l^1$ and $x_l^2$ are the coordinates of class centers, $d_l^1$ and $d_l^2$ are coordinates of class ranges, $r_{1l}^1$, $r_{2l}^1$, $r_{1l}^2$ and $r_{2l}^2$ are components of distances from class centers computed, $\beta_{Ai}^l(\mathbf{c}_k)$ are information vectors.

| $l$ | $\mathbf{c}_1$ | $\mathbf{c}_1$ | $x_l^1$ | $d_l^1$ | $r_{1l}^1$ | $\beta_{A1}^l(\mathbf{c}_1)$ | $r_{2l}^1$ | $\beta_{A1}^l(\mathbf{c}_2)$ | $x_l^2$ | $d_l^2$ | $r_{1l}^2$ | $\beta_{A2}^l(\mathbf{c}_1)$ | $r_{2l}^2$ | $\beta_{A1}^l(\mathbf{c}_2)$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1.0 | 2.0 | 1.0 | 0.0 | 0.0 | 1 | 1.0 | 0 | 2.0 | 0.0 | 1.0 | 0 | 0.0 | 1 |
| 2 | 0 | 0.48 | 0 | 0 | 0.0 | 1 | 0.48 | 0 | 0.50 | 0.05 | 0.5 | 0 | 0.02 | 1 |
| 3 | 50.0 | 40.0 | 45.0 | 5.0 | 5.0 | 1 | 5.0 | 1 | 50.0 | 5.0 | 0.0 | 1 | 10.0 | 0 |
| 4 | 95.0 | 205.0 | 100.0 | 10.0 | 5.0 | 1 | 105.0 | 0 | 200.0 | 20.0 | 105.0 | 0 | 5.0 | 1 |
| 5 | 36.5 | 23.0 | 36.0 | 1.0 | 0.5 | 1 | 13.0 | 0 | 22.0 | 2.0 | 14.5 | 0 | 1.0 | 1 |
| 6 | 7000 | 6800 | 6700 | 500.0 | 300.0 | 1 | 100.0 | 1 | 7000 | 500.0 | 0.0 | 1 | 200.0 | 1 |
| Decision: | | | | | $\mathbf{c}_1 \in M_1$ | | $\mathbf{c}_2 \notin M_1$ | | | | $\mathbf{c}_1 \notin M_2$ | | $\mathbf{c}_2 \in M_2$ | |