

Application of neural networks for determining optical parameters of strongly scattering media from the intensity profile of backscattered radiation

S.P. Kotova, I.V. Mayorov, A.M. Mayorova

Abstract. We analyse the possibilities of simultaneous measuring three optical parameters of scattering media, namely, the scattering and absorption coefficients and the scattering anisotropy parameter by the intensity profile of backscattered radiation by using the neural network inversion method and the method of adaptive-network-based fuzzy inference system. The measurement errors of the absorption and scattering coefficients and the scattering anisotropy parameter are 20 %, 5 %, and 10 %, respectively.

Keywords: optical parameters of scattering media, intensity profiles of backscattered radiation, Monte-Carlo method, neural networks, hybrid systems.

1. Introduction

The problem of measuring the scattering coefficient μ_s , absorption coefficient μ_a , and scattering anisotropy parameter g (the average cosine of the scattering angle) in strongly scattering media is of current interest because these parameters, along with the refractive index n , allow a complete description of the optical (laser) radiation intensity scattered and absorbed in a medium. Interest in this problem is related to a great extent to the development of optics of biological media and tissues [1–5] and to wide applications of low-intensity laser radiation in medicine both for diagnostics and therapy. From the point of view of medicine, of special interest is the development of non-invasive (nondestructive) methods for *in vivo* diagnostics of the state of biological systems. The methods for determining optical parameters from the spatial characteristics of backscattered radiation can be used for such diagnostics.

We proposed, theoretically substantiated and experimentally verified in [6–8] the method for simultaneous measurements of the scattering and absorption coefficients and scattering anisotropy parameter from the intensity profile of backscattered radiation. Optical parameters are determined in real time by comparing the experimental intensity profiles of backscattered radiation with model

profiles obtained by the Monte-Carlo method in a broad range of optical parameters. The experimental intensity profiles were obtained by using optical fibres for delivering and receiving radiation; a radiation source was a He–Ne laser. The use of continuous-wave radiation sources in diagnostic devices is convenient because such devices are quite simple and do not require the employment of fast radiation detectors and high-frequency devices for signal processing. Optical parameters were determined from the intensity profile of backscattered radiation by two independent methods of full enumeration and regularity (see details in [8]). As a result, the optical properties of media with albedo $\mu_s/(\mu_a + \mu_s) \leq 0.98$ were determined with a high accuracy (about 10 %). However, for media with a higher albedo, the acceptable accuracy was not achieved. According to the data presented in [1, 2], some biological tissues (for example, blood, uterus, skin derma, grey and white brain substance, etc.) are characterised by a high albedo (above 0.98) in the visible and near-IR regions.

In this paper, we analyse the possibility of determining optical parameters from the intensity profile of backscattered radiation by the method of neural network inversion and the method of adaptive-network-based fuzzy inference system (ANFIS). The main task was to improve the measurement accuracy of optical parameter (in particular, for albedo exceeding 0.98) and to select the most simple and convenient mathematical algorithm. Note that, although neural networks are widely used in optics of biological media, for example, for determining the absorption coefficient and transport scattering coefficient [9, 10] or parameters (average radius, refractive indices) of individual particles [11], the aim of our study was to demonstrate the possibility of using the neural network inversion and ANFIS methods for simultaneous measuring all the three optical parameters: the scattering and absorption coefficients and scattering anisotropy parameter. The knowledge of these parameters is important for dosimetry (when information on the light-field distribution near biological tissues is needed) and for diagnostic applications of medical devices in which the distance between the transport and receiving fibres is equal to one-two photon mean free paths $l_{tr} = 1/(\mu_a + \mu_s)$ [12], as well as devices with a limited receiving aperture [8].

2. Creation of the model data base by the Monte-Carlo method

As mentioned above, optical parameters were determined by comparing the experimental intensity profile of back-

S.P. Kotova, I.V. Mayorov, A.M. Mayorova Samara Branch of the P.N. Lebedev Physics Institute, Russian Academy of Sciences, Novo-Sadovaya ul. 221, 443011 Samara, Russia; e-mail: kotova@fian.smr.ru

Received 20 June 2006; revision received 10 October 2006

Kvantovaya Elektronika 37(1) 22–26 (2007)

Translated by M.N. Sapozhnikov

scattered radiation with model profiles constructed for a broad range of optical parameters. Model profiles were constructed by the Monte-Carlo method based on the numerical simulation of propagation of photons in a medium. The random walk of photons inside a biological tissue sample is simulated with a computer by using a random number generator and is monitored from the point of incidence of a photon on a sample until its absorption or escape from the sample through the tissue boundary. A photon experiences random collisions with inhomogeneities inside the sample in which, depending on the absorption and scattering coefficients, it is either absorbed or randomly changes (in accordance with the scattering phase function) its propagation direction [1]. We performed simulations for the experimental setup shown schematically in Fig. 1 [6].

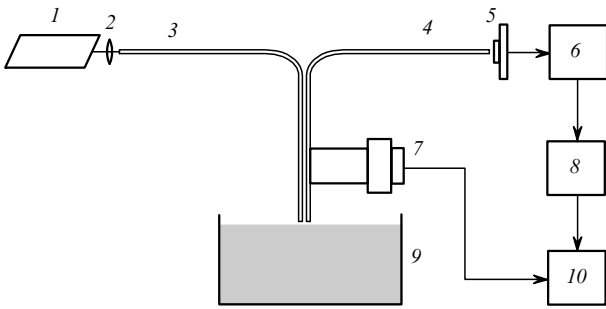


Figure 1. Scheme of the experimental setup: (1) laser; (2) objective; (3) supply fibre; (4) receiving fibre; (5) photodiode; (6) photocurrent–voltage transducer; (7) step motor; (8) voltmeter; (9) cell with a model medium; (10) computer.

A scattering medium with the scattering coefficient μ_s , absorption coefficient μ_a , the anisotropy parameter (the average cosine of the scattering angle) g and the refractive index $n = 1.4$ was restricted by the cell size ($49 \times 77 \times 26.5$ mm). The medium was illuminated by a beam from a $0.63\text{-}\mu\text{m}$, 4-mW GN-4P He–Ne laser through a supply fibre with the core diameter of $400\ \mu\text{m}$ and the aperture angle 35° . The intensity profile $R(\rho)$ of back-scattered radiation (where ρ is the distance from the supply fibre centre to the receiving fibre centre) was measured with the help of the receiving fibre fixed on a platform of a step motor. The receiving fibre had the same parameters as the supply fibre. The four-dimensional data base $R(\rho, \mu_a, \mu_s, g)$ was simulated in a broad range of optical parameters corresponding to the optical parameters of biological tissues: $0.01\ \text{mm}^{-1} \leq \mu_a \leq 0.15\ \text{mm}^{-1}$ with the step $\Delta\mu_a = 0.014\ \text{mm}^{-1}$; $2\ \text{mm}^{-1} \leq \mu_s \leq 20\ \text{mm}^{-1}$ with the step $\Delta\mu_s = 1.8\ \text{mm}^{-1}$; $0.8 \leq g \leq 0.98$ with the step $\Delta g = 0.018$; and $0 \leq \rho \leq 5.6$ mm with the step $\Delta\rho = 0.4$ mm. The Henyey–Greenstein scattering phase function

$$p(\theta) = \frac{1}{4\pi} \frac{1 - g^2}{(1 + g^2 - 2g \cos \theta)^{3/2}}$$

was used, where θ is the scattering angle. Each of the dependences was constructed by using 10^5 photons. Note that the function $R(\rho, \mu_a, \mu_s, g)$ has no local extrema in the region of values of ρ , μ_a , μ_s , and g under study; the profiles $R(\rho)$ rapidly decrease at small distances between fibres

(~ 1 mm), and then the dependence on ρ becomes exponential.

3. Determination of optical parameters by the method of neural network inversion

The problem of determining optical parameters from the characteristics of forward and backward scattered radiation is often called the inverse problem [1–3]. The neural network method is one of the universal methods used for solving inverse problems. The neural network allows one to simulate complex nonlinear relations between input and output parameters, similarly to the ‘black box’ model, i.e. the internal structure of the problem is not described by a set of mathematical equations but is created during training. There exist many neural networks, which have different properties [13–15]. We have selected a single-layer perceptron (Fig. 2) – a feed-forward network. Neurons in such a network are located in several layers. Signals are transmitted only in one direction from the input to output layer. The network in the problem under study consists of the input layer containing three neurons (with optical parameters μ_a , μ_s , g), the hidden layer containing 20 neurons [the intensity $R(\rho)$ is specified at 15 points]. The input and output layers are linear, and neurons in the hidden layer perform nonlinear transformations with the help of the activation function (hyperbolic tangent). The nonlinear function of this type was chosen because it provides simple calculations. Note that the network configuration and its type, the number of layers and the number of neurons in hidden layers are determined experimentally. The more complex is the network, the greater are its possibilities; however, the training time increases considerably in this case. In addition, the single-layer perceptron is used for solving not very complicated problems.

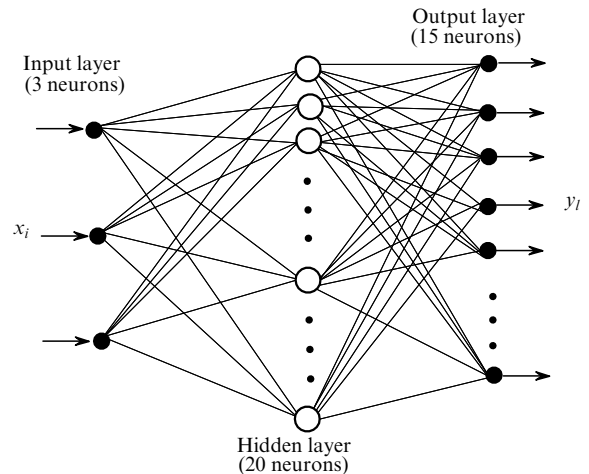


Figure 2. Scheme of a single-layer perceptron.

The network training was performed as follows. Let the input signal x_i be a set of optical parameters (μ_a, μ_s, g) and the output signal $y_l = R(\rho_l)$ be a set of intensity values at 15 points, N be the number of neurons in the hidden layer, and $\sigma(z) = \tanh z$ be a nonlinear activation function of a neuron. Let us denote the weights of neural couplings in the hidden layer by w_{ij}^h (i is the number of the input signal component, j

is the neuron number in the hidden layer), the weights of output couplings by w_{jl}^{out} (j is the neuron number, l is the number of the output signal component, i.e. of the intensity at the l th point). Then, if the signal x_i is fed to the perceptron input, the ‘activation’ signal

$$z_j = \sum_{i=1}^p w_{ij}^h x_i$$

will appear on neurons of the hidden layer (where $p = 3$ is the number of components of the input signal; $j = 1, \dots, N$), which is transformed by the neuron to the function $\sigma(z_j)$. The output signal component y_l is formed by adding all neurons of the hidden layer:

$$y_l = \sum_{j=1}^N w_{jl}^{\text{out}} \sigma(z_j),$$

where $l = 1, \dots, 15$.

Thus, we obtain the intensity values at 15 points at the output. Then, the intensity measurement error was found by comparing the results of measurements with model data, i.e. the intensity values obtained by the Monte-Carlo method. The weights of neurons of the output and hidden layers were corrected by the method of error backpropagation:

$$w_{jl}^{\text{out}} \leftarrow w_{jl}^{\text{out}} + \alpha \sigma(z_j) e_l,$$

$$w_{ij}^h \leftarrow w_{ij}^h + \alpha x_i \sigma'(z_j) \sum_l w_{jl}^{\text{out}} e_l,$$

where α is the training rate constant; e_l is the difference between the true and obtained l components of the output signal; and $\sigma'(z)$ is the derivative of the activation function.

Such a single calculation is called an ‘epoch’. After some number of epochs (as a rule, 500–1000), the error becomes acceptable. Usually, it amounts to 2.5%–10% because in the case of smaller errors a threat of the network ‘retraining’ appears, when the network is well trained by using one data set and poorly reproduces another data set. Note that, except the root-mean-square error after each iteration cycle (epoch), the program realisation allowed us to select the number of neurons in the hidden layer (from 20 to 30), the training rate, the number of iterations (from 100 to 1000), and different algorithms for searching the minimum of the error function: the first-order gradient and the second-order Levenberg–Marquardt algorithms [16]. To avoid the transfer of neurons to the saturation states described by the tails of the function $\tanh x$, all the data were normalised to the range $[-1, 1]$.

As a result, the trained neural network ‘stores’ the model in the numerical coefficients called neuron weights. Now, if the input signal is fed, the network produces the output signal, which differs from the model signal by the value not exceeding the training error. In addition, the trained network has approximating and predicting properties and allows one to obtain the intensity values for any optical parameters in the range under study. Thus, the network performs the nonlinear mapping $R(\rho, x_i)$. Note that in this case, the neural network was used in fact to solve the problem of approximation of a multidimensional nonlinear function. At the second stage of the problem, the input optical parameters were determined from the known values of $R(\rho)$ by using the trained network. Now, the network should operate in the inverse order: we assume that the neuron weights are fixed, the input signal will be 15 intensity values, and the fitting parameters will be the ‘input’ parameters, i.e. μ_a , μ_s , g . This solution method is called the neural network inversion [17]. The network can be trained at once to solve the inverse problem, i.e. to specify $R(\rho)$ and to obtain μ_a , μ_s , and g ; but the results will be much worse in this case. This is explained by the ‘poor’ properties of the inverse transformation (because the problem is ill-posed).

Thus, calculations with the use of the neural network included the following stages:

1. The input of data normalised to the range $[-1, 1]$.
2. The choice of the number of neurons in the hidden layer (20–30).
3. The choice of the training rate (0.1, 0.01, 0.001).
4. The choice of the number of iterations (100–1000).
5. The choice of the training method.
6. The determination of the root-mean-square error after each iteration cycle (epoch). If the error is considered acceptable, the training process is stopped, if not – the training is continued.
7. The verification of several test sets to make sure that the network is not ‘retrained’.
8. The fixation of neuron weights.
9. The data input for solving the inverse problem (from the file or Monte-Carlo data base).
10. The neural network inversion.

Table 1 presents some results of our numerical experiments on determining optical parameters. The neural network was trained on the entire data base. The training rate was set equal to 0.001, the number of hidden neurons was 30, and the number of epochs was 500. The network training was stopped when the root-mean-square error was less than 10% (9.3%). Optical parameters were determined from the profile $R(\rho)$.

Analysis of numerical experiments showed that the

Table 1. Measurement accuracy of optical parameters in the model experiment.

$\mu_s/(\mu_s + \mu_a)$	True values			Obtained values			Measurement errors (%)		
	μ_s/mm^{-1}	μ_a/mm^{-1}	g	μ_s/mm^{-1}	μ_a/mm^{-1}	g	$\Delta\mu_s/\mu_s$	$\Delta\mu_a/\mu_a$	$\Delta g/g$
0.996	9.0	0.04	0.91	8.5	0.03	0.87	5.5	25	4
0.996	11	0.04	0.91	9.7	0.03	0.87	13	25	4
0.995	20.0	0.10	0.80	17.1	0.04	0.92	14.5	60	15
0.993	11.0	0.08	0.89	10.2	0.05	0.91	7	50	2
0.993	9.0	0.06	0.91	8.1	0.04	0.90	10	33	1
0.992	9.2	0.07	0.89	9.23	0.04	0.91	< 1	43	2
0.98	5.0	0.10	0.85	5.8	0.06	0.91	16	40	7
0.93	2.0	0.15	0.80	1.2	0.12	0.85	40	20	6

method of neural network inversion allows one to determine the scattering coefficient μ_s and the scattering anisotropy parameter g with an error of $\sim 10\%$ even for high albedo (above 0.99); however, the measurement error for the absorption coefficient μ_a is 20%–60%. Such a large error is explained by the fact that the intensity profiles (especially for high albedo) are weakly sensitive to variations in the absorption coefficient. In addition, the measurement error of optical parameters considerably increases when the true values of μ_s , μ_a , and g coincide with the boundaries of the range of optical parameters for which the model data base was constructed (the third line in Table 1: $\mu_s = 20 \text{ mm}^{-1}$, the eighth line: $\mu_s = 2 \text{ mm}^{-1}$, $\mu_a = 0.15 \text{ mm}^{-1}$).

Note that, as in the method of full enumeration with averaging and the regularity method, the accuracy of optical parameter measurement by the method of neural network inversion depends considerably on the step of parameter specification in the model network. We constructed the additional model data base $R(\rho, \mu_a, \mu_s, g)$ in which the scattering coefficient μ_s was varied from 1.5 to 6.5 mm^{-1} with the step $\Delta\mu_s = 0.5 \text{ mm}^{-1}$, the anisotropy parameter was varied from 0.7 to 0.9 with the step $\Delta g = 0.02$, and the absorption coefficient μ_a and coordinate ρ were varied within the same range and with the same step as in the main model data base. The neural network was trained on this additional model data base. The training rate was set equal to 0.01, the number of hidden neurons was 20, and the number of epochs was 100. The network training was stopped when the root-mean-square error achieved 5%. Then, optical parameters were determined from the profile $R(\rho)$. The accuracy of their measurements increased considerably. For example, for a profile with the true values of optical parameters $\mu_a = 0.1 \text{ mm}^{-1}$, $\mu_s = 5 \text{ mm}^{-1}$, and $g = 0.85$ (seventh line in Table 1), the error decreased to 1%, 5%, and 3% for μ_s , μ_a , g , respectively. For a profile with $\mu_a = 0.15 \text{ mm}^{-1}$, $\mu_s = 2 \text{ mm}^{-1}$, and $g = 0.8$ (eighth line), errors were 14%, 6%, and 12.5% for μ_a , μ_s , and g , respectively.

Nevertheless, we consider that the method of neural network inversion did not improve in fact the measurement accuracy of optical parameters compared to the methods of full enumeration and regularity [8]. The substantial disadvantages of the neural network method are the difficulty of its realisation at the training stage: time-consuming computer calculations requiring high memory capacities, a considerable arbitrariness in the choice of the geometry of the system (the necessity to perform experiments with various networks and even in one network with different numbers of layers and different numbers of neurons in them) and also the possibility of the neural network ‘retraining’.

4. Solution of the inverse problem by using an adaptive-network-based fuzzy inference system

Due to considerable progress in the theory of fuzzy sets and experiments on combining the advantages of fuzzy logic and neural networks, hybrid systems in which the geometry of a neural network depends on the number of the so-called rules have gained acceptance beginning from the mid-1990s. Hybrid systems, namely, ANFIS [18] operate as follows: the data are fed to the system input and are replaced in the system by fuzzy distributions (fuzzification), and the system

of rules is constructed. Then, the fuzzy distribution of the output parameter values is transformed to the region of real numbers (defuzzification), and the neural network ‘corrects’ the parameters of the rules during training. The trained ANFIS operates externally as the neural network does; however, its internal mechanism is related to the adjustment of the system of fuzzy rules rather than to the fitting of the weights of the neural network. We emphasise that fuzzy sets are used inside the system, while the input and output data are usual quantities. To simplify algorithms, we simulated fuzzy sets by Gaussian functions. Usually, a broad system of functions from linear to bell-shaped is employed, but this leads to more complicated processing algorithms.

The typical fuzzy rule has the form

if x_1 in A_{1k} , x_2 in A_{2k} , ..., x_n in A_{nk} ,

then $y_k = a_{k1}x_1 + \dots + a_{kn}x_n$,

where x_1, x_2, \dots, x_n are input parameters; y_k is the output parameter; A_{nk} are fuzzy sets; a_{kn} are the coefficients of the Takagi–Sugeno model [19]; and k is the rule number. Defuzzification is performed by taking simultaneously into account the action of all the rules:

$$y(x_i) = \sum_k^K \gamma_k(x_i) \sum_i^M a_{ki}x_i,$$

where $\gamma_k(x_i)$ is the so-called activation degree of the k th rule; K is the number of rules; and M is the number of input signal components.

In practice in the ANFIS the Takagi–Sugeno model [19] is used, as a rule. The described model is the multiple inputs–single output (MISO) model. However, we need to use the MIMO model [multiple inputs (three optical parameters) – multiple outputs (intensity values at 15 points)]. The MIMO model can be obtained from several independent MISO models. The input signal for each of the MISO models is three optical parameters, and the output signal is the intensity at one point. After training 15 MISO models, the inverse problem was solved by the neural network inversion method, which was considered earlier.

Let us list all the successive steps used in the ANFIS for solving the inverse problem:

1. The input of the data base, its normalisation to the range $[-1, 1]$ and thinning out to $\sim 1/15$ of the entire volume.
2. The choice of the number of the fuzzy rules (3, 5, 7, ...).
3. The choice of the number of iterations (100–500) and the training rate (0.1–0.01).
4. Control of the root-mean-square error ($\sim 5\%$).
5. The input of data for solving the inverse problem (the input intensity signal is specified from the Monte-Carlo data base).
6. The solution of the inverse problem by the network inversion method.

Note only that, due to a large volume of the data base $R(\rho, \mu_a, \mu_s, g)$ ($15 \times 10 \times 10 \times 10$) obtained in Monte-Carlo simulations, we could not train the ANFIS on all data or on an arbitrary part of the data, as is usually done. Unlike the neural network, in which we successively specified each pair (x_i, y_i) , where x_i is the input signal and y_i is the output signal, and fitted neuron weights at each step, we used in the

ANFIS the so-called batch correction, i.e. a large set of (x_i, y_i) pairs. The batch correction is less sensitive to the local minima of the error function and gives better results; however, it requires much greater memory resources for matrix storage. Due to stricter requirements for computational resources, we could use only the first-order optimisation algorithms (gradient method), whereas in the previous problem the Levenberg–Marquardt algorithms could be also employed.

As for the data base, taking into account the restricted memory resources and providing the acceptable training time (no longer than 1 h by using a 1600-MHz, 512-Mb Pentium IV processor), we selected all 15 elements over ρ from this data base and by four elements for μ_a , μ_s , and g (instead of 10). This reduced data base is approximately 15 times smaller than the entire model data base.

Numerical experiments showed that the inverse problem was solved by using the ANFIS with better accuracy than with the help of the single-layer perceptron. We found that μ_a is determined with an error of 15%–20%, and μ_s and g with an error of less than 10%. Note that the values of μ_a and μ_s were often determined ambiguously [i.e. the specified profiles $R(\rho)$ were coincident, whereas the true and obtained values of μ_s and μ_a were different]. This is explained by the fact that the inverse problem is ill-posed and by the use of a small amount (less than 10%) of initial data for the numerical experiment.

5. Conclusions

We have solved the problem of simultaneous measuring three optical parameters of strongly scattering media from the intensity profile of backscattered radiation by using four mathematical algorithms: the method of full enumeration with averaging, the regularity method [6–8], and the methods of neural network inversion (single-layer perceptron) and Takagi–Sugeno adaptive-network-based fuzzy inference system considered in this paper. Neural networks were trained on the model data base $R(\rho, \mu_a, \mu_s, g)$ obtained by the Monte-Carlo method in a broad range of optical parameters. A specific feature of neural networks and fuzzy networks is that the training process requires considerable memory and processor resources, while the trained system gives results rapidly and does not require great computational resources. The measurements errors for the absorption coefficient, scattering coefficient and scattering anisotropy parameter were 20%, and 5%–10%, respectively [including media with high albedo (above 0.98)]. A comparatively high measurement error of the absorption coefficient is explained first of all by the fact that the intensity profiles of scattered light (especially for high albedo) are weakly sensitive to variations in the absorption coefficient.

A comparative analysis of algorithms has shown that the network-inversion calculation provides virtually the same accuracy as the method of full enumeration with averaging, however, is more time-consuming. The methods of regularity and fuzzy neural system give similar results. An attractive property of neural networks used for determining optical parameters is the possibility of their training on the experimental data base without using the model Monte-Carlo data base. Note that the method of fuzzy logic is very promising, it is quite general and therefore can be used for solving similar problems.

6. References

1. Tuchin V.V. *Lazery i volokonnaya optika v biomeditsinskikh issledovaniyakh* (Lasers and Fibre Optics in Biomedical Studies) (Saratov: Saratov State University, 1998).
2. Tuchin V. *Tissue Optics* (Bellingham: SPIE Press, 2000).
3. Wilson B., Jacques S. *IEEE J. Quantum Electron.*, **26**, 2186 (1990).
4. Gosh N., Mohanty S., Majumder S., Gupta P. *Appl. Opt.*, **40**, 176 (2001).
5. Mourant J., Fuselier T., Boyer J., et al. *Appl. Opt.*, **36**, 949 (1997).
6. Kotova S.P., Mayorova A.M., Yakutin V.V. *Opt. Spektrosk.*, **95**, 452 (2003).
7. Mayorova A.M., Kotova S.P., Jakutkin V.V. Preprint FIAN No. 35 (Moscow, 2002).
8. Mayorova A.M., Kotova S.P., Rakhmatulin M.A., Jakutkin V.V. *J. Russian Laser Research*, **24** (1), 1 (2003).
9. Farrel T., Wilson B., Patterson M. *Phys. Med. Biol.*, **37** (12), 2281 (1992).
10. Kienle A., Ligte L., Patterson M., et al. *Appl. Opt.*, **35**, 2304 (1996).
11. Berdnik V.V., Mukhamedyarov R.D., Loiko V.A. *Opt. Spektrosk.*, **96**, 323 (2004).
12. Bevilacqua F., Piguat D., et al. *Appl. Opt.*, **38**, 4939 (1999).
13. Braspenning P.J., Thuijjsman F., Weijters A.J. *Artificial Neural Networks* (Berlin: Springer-Verlag, 1995).
14. Bishop C.M. *Neural Networks for Pattern Recognition* (Oxford University Press, 1995).
15. Gorban' A.N., Rossiev D.A. *Neironnye seti dlya personal'nykh komp'yutero* (Neural Networks for PC) (Novosibirsk: Nauka, 1999).
16. Hagan M.T., Menhaj M. *IEEE Trans. Neural Networks*, **5** (6), 989 (1994).
17. Jensen C.A., Reed R.D., et al. *Proc. IEEE*, **87** (9), 1536 (1999).
18. Jang J.-S.R. *IEEE Trans. Systems, Man, and Cybernetics*, **23** (3), 665 (1993).
19. Takagi T., Sugeno M. *IEEE Trans. Systems, Man, and Cybernetics*, **15** (1), 116 (1985).