Optics Communications 287 (2013) 140–144

Contents lists available at SciVerse ScienceDirect

Optics Communications

journal homepage: www.elsevier.com/locate/optcom

Neural networks modeling for refractive indices of semiconductors

A.A. Attia [a,*], M.M. El-Nahass [a], M.Y. El-Bakry [a,b], D.M. Habashy [a]

[a] Physics Department, Faculty of Education, Ain Shams University, Roxy, Cairo 11757, Egypt
[b] Physics Department, Faculty of Science, University of Tabuk, KSA

ARTICLE INFO

Article history:
Received 14 April 2012
Received in revised form
31 July 2012
Accepted 4 September 2012
Available online 1 October 2012

Keywords:
Neural networks
Refractive indices
Semiconductors

ABSTRACT

This paper uses an artificial neural network (ANN) and Levenberg–Marquardt training algorithm to model the nonlinear relationship between refractive index and energy gap in semiconductors. The predicted simulation values of the ANN are in accordance with the experimental data. An error deviation (Δn) was estimated for different models. The lowest deviation is given by the ANN model. The ANN model performance was also tested for some compounds not included in the training and was found to be in good agreement with the experimental data. High precision of the NN model is demonstrated as well as good generalization performance.

1. Introduction

Refractive index is one of the most useful properties of the semiconductor and is important for both fundamental and applied considerations [1–5]. Estimation of this parameter is important for optical wave guiding in optoelectronic structures like heterojunction laser diodes, optical amplifiers or optical fibers. Many attempts have been made to correlate the energy band gap to the optical refractive index of semiconductors. Moss [6] was the first to find a relation between the refractive index n and the energy band gap E_g. Ravindra and Srivastava [7] suggested another relation. Reddy and Anjaneyulu [8] proposed a logarithmical form of n as a function of E_g. Hervé and Vandamme [9] proposed an overall relation based on the classical oscillator theory [10]. For centuries, scientists have attempted to identify and document analytical laws that underlie physical phenomena in nature. Despite the prevalence of computing power, the process of finding natural laws and their corresponding equations has resisted automation. A key challenge to finding analytical relations automatically is defining algorithmically what makes a correlation in observed data important and insightful [11]. Artificial neural network (ANN) is a kind of data mining and artificial intelligence technique (AI) [12–15]. A neural network is a massively parallel-distributed processor that has a neural propensity for storing experiential knowledge and making it available for future uses. Unlike conventional, explicitly programmed computer programs, neural networks (NNs) are trained through the use of the experimental data and then the weights of the neurons iteratively adjusted until the output for a specific network is close to the desired one. Furthermore, NNs possess many excellent properties such as outstanding non-linear approximation, auto-adaptation and association capability. As a complex non-linear system, NN models have been widely employed to map the indeterminate relationship between cause and effect variables in many fields [16–21]. In the present work a universal ANN program is designed on the basis of improvement up on back-propagation training algorithms. This paper uses ANN and Levenberg–Marquardt training algorithm to model the nonlinear relationship between refractive index and energy gap. The experimental data of refractive indices n_{exp} have been collected from literature [8,9,22–25].

Following sections provide analysis of previous models, a brief introduction to NN model, training mechanism and discuss the results.

2. Analysis of previous models

In 1950, Moss [6] proposed a simple relation based on the fact that energy levels of electrons in a crystalline solid are scaled down by a factor $1/\varepsilon_{opt}^2$, where $\varepsilon_{opt}=n^2$ is the optical dielectric constant. He proposed

$$n^4 E_g = K \text{ with } K = 95\ eV \tag{1}$$

Ravindra and Srivastava [7] suggested a revised value of $K=108$ eV for a better curve-fit; hence

$$n^4 E_g = 108 \tag{2}$$

Reddy and Anjaneyulu [8] presented a new formula with an exponential behavior quite similar to the Moss relation:

$$E_g e^n = 36.3 \tag{3}$$

* Corresponding author. Tel.: +20 1001655010.
E-mail address: attia05@yahoo.com (A.A. Attia).

0030-4018/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.optcom.2012.09.016

In 1994, Havé and vandamme [9] proposed an equation based on the oscillatory theory [10]:

$$n=\left[1+\left(\frac{A}{E_g+B}\right)^2\right]^{1/2} \tag{4}$$

where A is equal to the hydrogen ionization energy with value 13.6 eV. Also B is the difference between the UV resonance energy $\hbar\omega_o$ and the energy gap E_g with constant value 3.4 eV as they assumed.

In the previous analysis, all band gaps are indicated by E_g in units of eV and the optical refractive index is in visible light at normal dispersion.

3. Artificial neural networks

Artificial neural networks (ANNs) [26–30] have been developed as generalization of mathematical models of biological nervous systems. The basic processing elements of neural networks are called artificial neurons, or simply neurons or nodes. A typical neuron receives one or more input signals and provides an output signal depending on the processing function of the neuron. An example of a neuron is shown in Fig. 1. This is a simple processing unit and is known as an elementary perceptron. Each input X_i is weighted by a factor W_i and the whole sum of inputs is calculated $\left(\sum_{\text{all inputs}} W_i X_i\right)$; then an activation function f is applied to the result a. The neural output is taken to be f(a):

$$\text{output}=\text{activation function}\left(\sum_{\text{all inputs}} W_i X_i\right)=f(\text{a}) \tag{5}$$

Multi-layer feed forward network (Fig. 2) has become the major and most widely used supervised learning neural network architecture. In the feed forward networks, all connections are unidirectional from input to output layer. More than a hundred different learning algorithms are available, but the most popular method is back-propagation. This method was first proposed in 1969, but was ignored because of its demanding computations. Only in the mid-1980s was the back-propagation learning algorithm rediscovered. The network computes its output pattern, and if there is an error – or in other words a difference between actual and desired output patterns – the weights are adjusted to reduce this error.

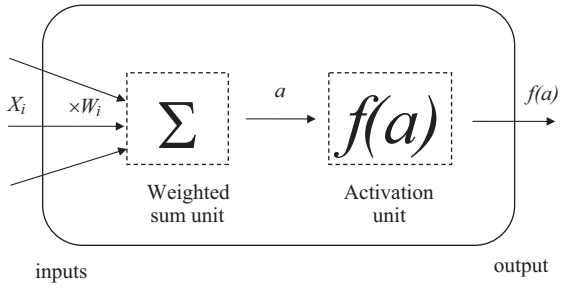

Fig. 1. Neuron model.

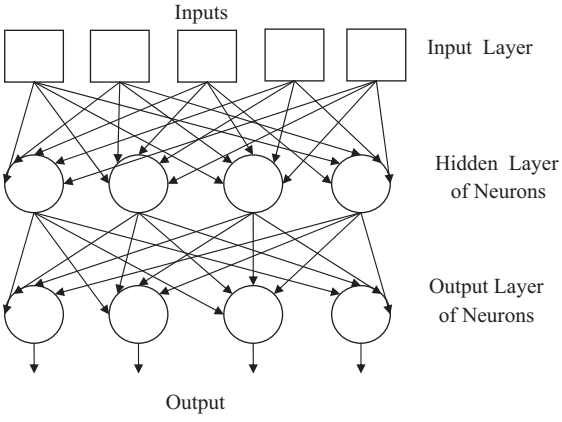

Fig. 2. Multilayer feed forward network with 1 hidden layer.

3.1. Modeling the refractive indices using ANN

The proposed ANN model for refractive indices has one input and one output. The input is energy gap E_g and the output is refractive index n, as shown in Fig. 3. Using this input–output arrangement, different network configurations were tried to achieve good mean square error (MSE) and good performance for the network. The three layers configuration shown in Fig. 3 is chosen. These layers are two hidden layers of 8 and 6 neurons and the output layer consisting of one neuron. The transfer function of the first and second layer was chosen was a logsig while the output to be pureline as in the Appendix.

3.2. Neural network training for refractive indices

The network learning process [31] is supervised, i.e. the network receives (at training phase) both the raw data as inputs (experimental data) and the targets as output (n). The learning involves adjusting weights, based on a comparison of the output and the target, until the network output gets as close as possible to the target value. The proposed ANN was trained using the Levenberg–Marquardt optimization technique. This optimization technique is more powerful than the conventional gradient descent techniques. The Levenberg–Marquardt updates the network weights using the following rule:

$$\Delta W=(J^TJ+\mu I)^{-1}J^Te \tag{6}$$

where J is the Jacobian matrix of derivatives of each error with respect to each weight; $J^T$is the transposed matrix of J; I is the identity matrix that has the same dimensions as those of J^TJ; μ is a scalar changed adaptively by the algorithm and e is an error vector. ΔW is a measure for the rate of learning of the network i.e. when ΔW tends to zero this means that the network has been learnt (ready to predict the unseen values). Then the adjustment for the weights is done by Eq. (6) to reduce the error value.

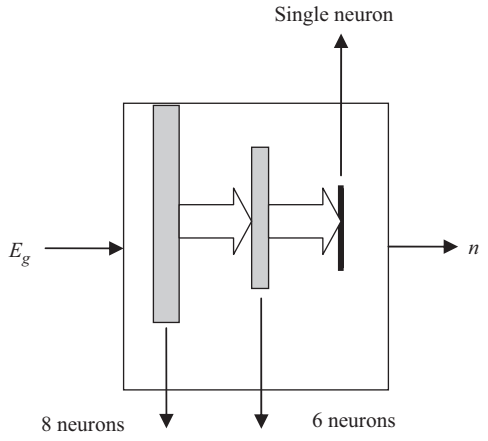

Fig. 3. Block diagram for refractive index (n) and energy gap (E_g) based ANN model.