Optical measurement system for real-time estimation of gas pollutants' concentration based on artificial neural network

[1]W.Wójcik, [2]Z.Gotra, [1]M.Duk, [1]A.Kotyra, [1]S.W.Przyłucki, [1]A.Smolarz

[1] Department of Electronics, Technical University of Lublin, 38A Nadbystrzycka St., 20-618 Lublin, Poland
[2] Lviv State Technical University, 12 Bandery St., Lviv, Ukraine

Received 29.11.2000

After revision: 8.02.2001

Abstract

In this paper we present optoelectronic measurement system for monitoring the industrial gas pollutants. The system consists of a source of light, an optical fiber as a data transmission link, a spectrometer, an optical detection system and a neural network unit for a real-time spectral data processing. We have paid the major attention to the neural network structure and its properties in gas recognition and gas concentration estimation.

Keywords: neural network, power industry, spectroscopy, environmental pollution

1. Introduction

During last few years, the interest in developing technologies which can decrease the gaseous pollutants emission to the atmosphere, is growing. New, lower emission limits introduced in many countries, tend to continuous improvement of the existing industrial installations. For the same reason, the increasing demand for flexible, easy to maintain air pollution monitoring systems, is observed on the market. Often, modern solutions of these systems are installed as a part of the technological process control and then the quantity of emission is one of the variables in a control algorithm. In that case, data analysis time is of the same importance as precision of the estimation of pollutants concentration. It is especially true, when the control system needs to know if emission limit was exceeded for several different gases. Among many monitoring techniques, optical measurements seem to be superior, offering high precision, repeatability of measurements in conjunction with non-intrusive interaction with monitoring media. Optical measurement heads together with optical fibers show high reliability and immunity on the industrial disturbances (electromagnetic field, temperature, etc.) and become not only new alternative for "ordinary" methods but have created a new quality standard [1]. Additionally, a single optical measurement path can carry a big amount of data in entirely parallel manner (e.g. spectral response of light in an absorbing media). Unfortunately, even fast electronic chips are not able to handle a real-time (or fast enough) computing a huge number of data sets delivered by an optical system, especially when one wants to analyze several gas pollutants in changing measurement conditions (variable temperature, pressure, humidity, etc.). One of possible solutions, ensuring parallel evaluation of data with a lot of components (i.e. spectral data set) are neural nets (NN) [2]. Among many possible neural architectures, the counterpropagation net

(CP) with a few described modifications have been chosen. The CP net can classify data sets and find correlation between their vector components. Relationships of data can remain unknown what is an important advantage for all cases of variable measurement conditions. The quality of the neural network performance depends on its training process and it includes a number of neurons. A data given by analysts inside the network is entirely parallel and it is fault tolerantly covering noisy data. Conjunction of elements described above, the optical detection unit, the optical fiber and the hardware implemented neural net, gives way to the development of relatively cheap, easy to maintain, reliable measurement system.

2. The measurement system description

2.1 System principles

The physical background of measurement is very well known and based on Beer–Lambert–Bourger law of exponential attenuation for the light passing through an absorbing medium.

$$I(L,\lambda) = I_0(\lambda)\cdot \exp(-c\delta_\lambda L) \tag{1}$$

where $I_0(\lambda)$ is incident light intensity, $I(L, \lambda)$ is a transmitted light intensity after passage through an optical path-length L in the medium, c is the concentration of an absorber, δ_λ is the absorption cross section at the specified wavelength. From the point of view of an environment protection two gases are being considered SO_2 and NO_2.

2.2 System configuration

The system is shown on Fig.1 and consists of the optical detection system, the optical fiber link, the optoelectronic data analysis unit and the graphical user interface.

Depending on expected concentration level, single or double (with retroreflector) optical path can be employed. In order to simultaneous determinated presence and concentration of several gas pollutants, a broad band light source is to be used. Spectral region and wavelength range has influence on the choice of a light source, an optical fiber, a diode array detector and spectrometer parameters. Of course, that choice is to be made taking into consideration the absorption media parameters and features of the monitored gases. In our case SO_2 and NO_2 are being monitored. As it can be seen the required bandwidth is from 270nm to 600nm.

The light, after passing through the optical path in the absorber, is collected into the optical fiber, which is then connected with the measurement head and with the spectrometer. Questions of light source to fibre coupling are not considered here, detailed analysis of the problem can be found in [3]. After passing through an input slit the light is being diffracted on a grating. The spectrometer includes the 32 diode array as a light detector. Diode array forms an input layer of the neural network. Digital signal processor unit normalizes the NN input and serves as the intelligent output interface for a graphical user interface (GUI).

According to the grating properties this spectrometer theoretically can cover a range from 200nm to 730nm. Yet spectral attenuation of the fiber itself sets lower limit to about 220nm. Absorption peaks of oxygen, because of it small concentration in flue gases, does not limit the bandwidth. Spectral resolution results in $\frac{750-120}{32} \approx 17\text{nm}$. The integration time of the device ranges from 10 to 600 milliseconds and resulting signal, thus accuracy, is directly proportional to this time. In harsh industrial conditions such solution has a principle advantage – no movable mechanical parts. Additionally it combines advantages of „in situ” methods and extraction methods.

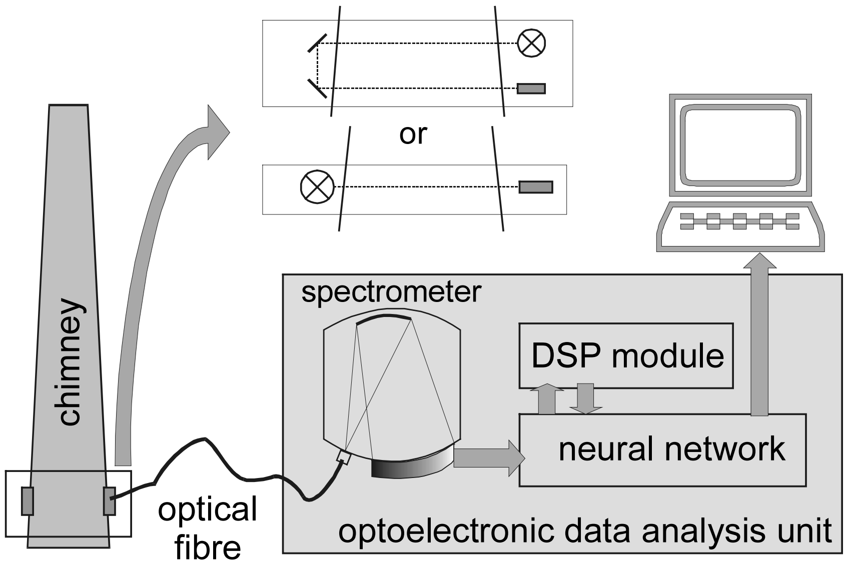

Fig.1. Schematic diagram of example application of the neural based optical measurement system for determination of presence and concentration of gases in the power plant chimney.

3. The counterpropagation neural network algorithm

Conditions changing during the measurement and number of gases taken into account complicate the learning process of the net. We assume that the properties of knowledge generalisation inside neural network will be of a great advantage. Nevertheless, the utilisation of the most popular error backpropagation algorithm and perceptron-like feedforward net could be the time consuming task, especially for bigger structures. Better solution could be obtained if the network would have a fast internal data representation feature. The networks of that kind are called self-organising nets (e.g. Kohonen net). For our purposes, another layer is necessary. The output signals should inform about the detected gas and its concentration. It can not be done basing only on the internal feature map of the self-organising layer. A neural network fulfilling above conditions (among others) is a so called counerpropagation net (CP).

Counterpropagation was originally proposed as a pattern-lookup system that takes advantage of the parallel architecture of neural networks. For our implementation we use a modified version of the CP network. Two types of layers are used (Fig.2): the hidden layer is a Kohonen layer with competitive units that do unsupervised learning, the output layer is a Grossberg layer, which is fully connected with the hidden layer.

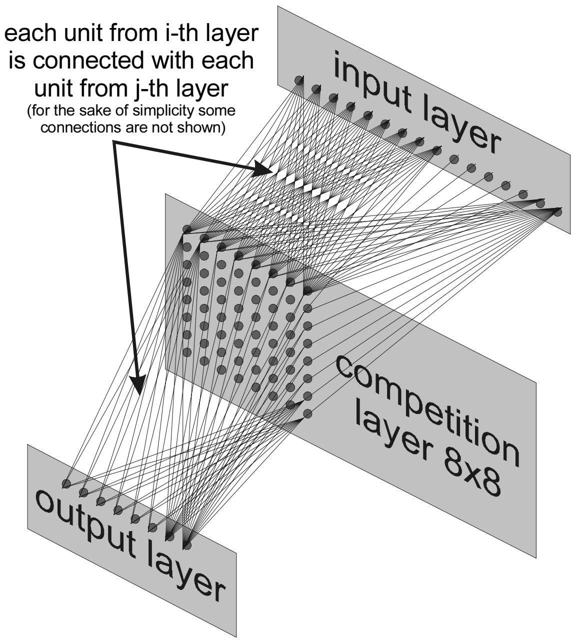

Fig.2. Example diagram of the counterpropagation neural net. The structure consists of 16 input units, 8x8 units competition layer and 4 outputs units.

Additionally, there is the input layer which is fully connected with the hidden layer but there is no computation inside it. Input layer can be integrated with the diode array because the number of inputs is equal to the number of array outputs. When presented with a pattern (e.g. single absorption spectra), the network classifies that pattern using a learned reference vector. The hidden units play a key role in this process, since the hidden layer perform a competitive classification to group the patterns.

3.1 Network operation principles

Being trained, the network works as follows. After presentation of a pattern in the input layer, the units in the hidden layer sum their inputs according to

$$net_j = \sum_i w_{ij} \cdot o_i \tag{2}$$

where i is the number of inputs (number of neurons) in the input layer, j is the number of units in the hidden layer, net_j denotes the input of j-th unit (neuron) in the hidden layer, $w_{i,j}$ the weight of the link from i-th unit to j-th unit, and o_i is output of i-th unit in input layer (i-th element of the input pattern).

In the standard CP algorithm, the unit with the highest net input wins and its activation is set to 1 while all others are set to 0 (winner takes all the rule). After the competition, the output layer does a weight sum on the outputs of the hidden layer.

$$a_k = net_k = \sum_i w_{jk} \cdot o_j \tag{3}$$

where k is a number of outputs (number of neurons in the output layer), a_k is an activation function of k-th output unit (we used the sigmoidal function).

Let c be the index of the winning hidden layer neuron. Since o_c is the only non-zero element in the sum, which in turn is equal to one, this can be reduced to:

$$a_k = w_{ck} \tag{4}$$

Thus the winning hidden unit activates a pattern in the output layer. As it was mentioned above, we have introduced some modifications to the standard algorithm. We replaced Winner-Takes-All layer (hidden layer) by the Kohonen self organizing map. That layer offers two additional features:

– during learning process not only the winning neuron in the hidden layer is able to change its weight but all weights of neurons in the neighborhood of the winning one are changed as well (a neighborhood dimension is defined as a learning parameter and it decreases during the learning process),

– the winner-takes-all rule was modified allowing more than one neuron to be activated at the same time; thanks to this the output layer is not only stimulated by the winning neuron but it can approximate a few links (it expands number of classes, the network can recognize the number of units in the hidden layer).

3.2 Network training process

Before starting the learning process, it is important to initialize the competitive layer with normalized weights. The input vector is to be normalized to ensure proper feature selection, too. The input pattern vectors are presented to all competitive units in parallel and the best matching unit is chosen as the winner. Since the vectors are normalized, the similarity between the normalized input vector $X = (x_i)$ and the reference unit $W_j = (w_{ij})$ can be calculated using the dot product:

$$net_j = X \cdot W = \sum_i x_i w_{ij} \tag{5}$$

The topological order is achieved by using a spatial neighborhood relation between the competitive unit during learning. Not only the best-matching vector, with weight W_c but also its neighborhood, is adapted [4],

$$\Delta w_{i,j}(t) = e(t) \cdot \left(x_i(t) - w_{i,j}(t)\right) \quad \text{for j} \in \text{N}_\text{c} \tag{6}$$

$$\Delta w_{i,j}(t) = 0 \quad \text{for j} \notin \text{N}_\text{c} \tag{7}$$

where $e_j(t) = h(t) \cdot \mathrm{e}^{\left(\frac{d_j}{r(t)}\right)^2}$ is the Gaussian function, d_j is the distance between W_j

and winner W_c , *h(t)* is the adaptation height at time *t*, *r(t)* is the radius of the spectral neighborhood N_c at time *t* and N_c is the definition of the neighborhood.

The adaptation height and radius are usually decreased over time to enforce the clustering process. All described procedures are called an unsupervised learning. The output layer is taught according to supervising learning algorithm. The output of the network is computed (Eqn. (3)) and compared to the target pattern. The weights between the competition layer units and the output units are updated according to

$$w_{jk}(t+1) = w_{jk}(t) + \beta \cdot (o_k - w_{jk}(t)) \tag{8}$$

where β is the learning parameter describing the basic weight adaptation step.

4. Results of performance simulation of the measurement system.

All data sets for the system simulation were taken in „Kozienice" Power Plant. The optical detection system based on OPSIS AB monitoring equipment, which estimate the pollutants emission by the differential optical absorption spectroscopy (DOAS) [1]. That system was also a source of the target data set for testing data pattern during estimation of total system performance. All measurements have been made in the UV region and two main gas pollutants were taken into account: nitrogen dioxide – NO_2 and sulfur dioxide – SO_2. A broad band light source consists of high pressure 150W xenon lamp and a light collimating system. Germanium doped optical fiber, 15m long, was used to transmit the light from the optical head to the spectrometer. The prototype spectrometer includes a monochromator which diffracts the light and focuses it onto 32 elements diode array. Spectral data for pure gases were used to create the teaching patterns. The simulation process requires at least two different pattern sets:

1. the teaching pattern consisting of 250 spectral characteristics of NO_2. and 250 spectral characteristics of SO_2, both with increasing concentration;
2. the validation pattern consisting of 1000 measured gas pollutants spectral characteristics.

Presented results were obtained with the CP network, with a size of 32 units in the input layer, 16x16 units in the Kohonen layer and 2 units in the Grossberg layer. The learning process included 10 000 steps with random choice of one spectrum. After the training of the CP, each unit in the two dimensional competition layer is identified with the type of gas and its concentration represented by this unit. Each output unit represent a specific gas and the value of the unit activation function informs the recognized gas concentration. For our teaching pattern, the Kohonen layer has found two different configurations and begins to arrange spectral data in different regions. The NN is able to distinguish two gases. The area of recognition of the first gas (Fig.3 left) is separated from the area of recognition of the second gas (Fig.3 right). Additionally, different units represent different concentrations (different gray levels). Next, the recognition of two gases is verified by applying the validation pattern. In 91.82% the NN recognizes the right gas with an error less than 5%. Another important question is to be answered, is dealing with the optimal network architecture. Using the same patterns, we estimate the performance of the CP network of different architectures. As the overall quality index of network performance, we choose the root mean square error (RMS). Our test has covered the CP network with 32 input neurons, 2 output neurons and a variable number of neurons in the competition layer, from 4x4 to 16x16 units. The results are presented on Fig.4. Each set of the active neurons in competition layer generate a network output. The number of possible solution depends on the number of neurons in the hidden layer and a kind of adopted neighborhood function. This is the reason why a small number of neurons produces a bigger error.

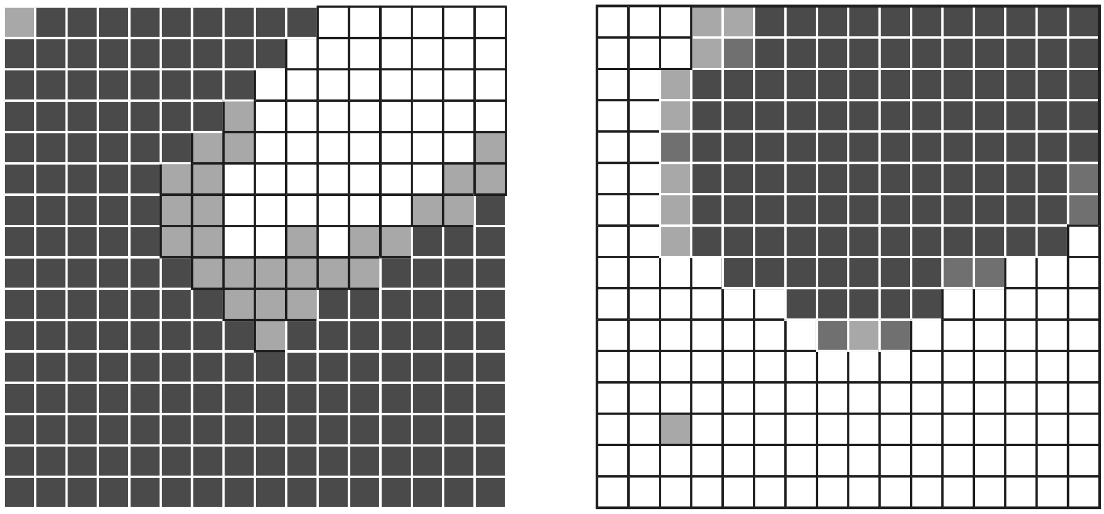

Fig.3. Clustering of the competition layer for NO_2 (left) and for SO_2 (right).

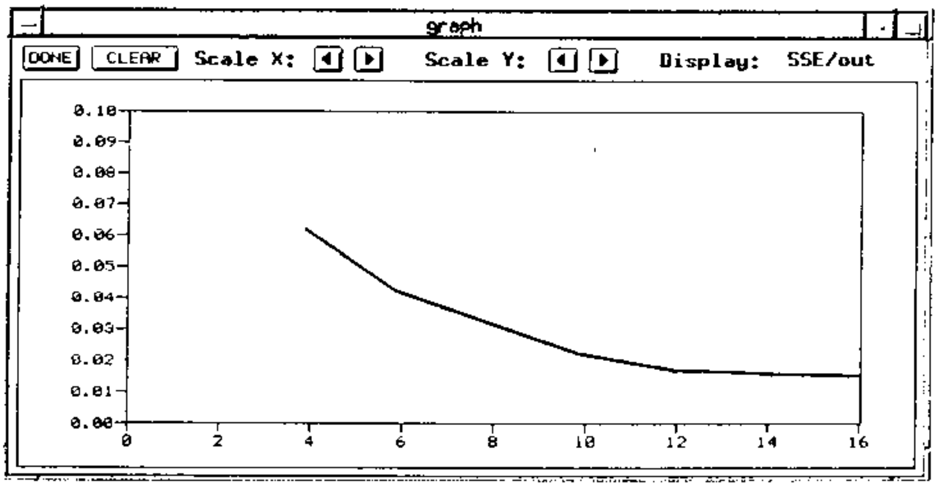

Fig.4. The relative output error as a function of the number of neurons in the competition layer

The network output error decreases as the number of neurons increases but as it can be seen after a certain dimension further adding of neurons to the competition layer do not significantly improve the quality of the output signal.

5. Conclusions

The counterpropagation neural network seems to be a good solution for evaluation of spectral data concerning the problems of flexibility, quality and performance. Spectral data can be recognized in real time by a learned CP net. The network validates noisy data. That is an important fact because every measurement can vary in several parameters. That usually makes "conventional" serial computing time consuming and requires a big amount of memory. The bitstream encoding allows considerable reduction of hardware required. Yet, even very complicated numerical solution can not anticipate all possible measurement conditions. That is the reason why, the reference data sets, which are fitting numerically to real spectral data, are not completed. That usually leads to evaluation errors [5].

References

1. Par Ragnarson "Optical Techniques for Measurements of Atmospheric Trace Gases"

Lund Raports on Atomic Physics LRAP-152, Lund Institute of Technology, Department of Physics, April 1994.
2. Paul E. Keller, Richard T. Kozues, Lars J. Kangas. Procceding of the 8-th Conference on Real-Time Computer Applications in Nuclear, Particle and Plasma Physics (1993).
3. Krolopp W.J., Wójcik W., Kotyra A., Przyłucki S., Smolarz A. Reports of the Lodz Polytechnick, Electrical, **90**, (1998), 5–38.
4. Kohonen T., Self-Organizing Maps, Spriger Series in Information Sciences, (1995).
5. Xiang Ouyang, Philip L. Varghese Appl. Optics, **38**, (1989)