ANALYSIS OF THE EXPERIMENTAL POSITRON LIFETIME SPECTRA BY NEURAL NETWORKS

by

Senada AVDIĆ[1], Roumiana CHAKAROVA[2], and Imre PÁZSIT[3]

Received on February 11, 2003; accepted in revised form on May 21, 2003

This paper deals with the analysis of experimental positron lifetime spectra in polymer materials by using various algorithms of neural networks. A method based on the use of artificial neural networks for unfolding the mean lifetime and intensity of the spectral components of simulated positron lifetime spectra was previously suggested and tested on simulated data [Pázsit *et al*, *Applied Surface Science*, *149* (1998), 97]. In this work, the applicability of the method to the analysis of experimental positron spectra has been verified in the case of spectra from polymer materials with three components. It has been demonstrated that the backpropagation neural network can determine the spectral parameters with a high accuracy and perform the decomposition of lifetimes which differ by 10% or more. The backpropagation network has not been suitable for the identification of both the parameters and the number of spectral components. Therefore, a separate artificial neural network module has been designed to solve the classification problem. Module types based on self-organizing map and learning vector quantization algorithms have been tested. The learning vector quantization algorithm was found to have better performance and reliability. A complete artificial neural network analysis tool of positron lifetime spectra has been constructed to include a spectra classification module and parameter evaluation modules for spectra with a different number of components. In this way, both flexibility and high resolution can be achieved.

Key words: positron lifetime spectra, unfolding procedure, spectral parameters, backpropagation network, spectral components, classification, SOM and LVQ algorithms

INTRODUCTION

Positron annihilation lifetime spectroscopy (PALS) is widely used in the study of defects in metals and semiconductors, as well as in the study of free volume in polymer materials. This experimental technique is a powerful tool for probing the microstructure of condensed matter and shows an advantage over other microprobes. Defect spectroscopy should provide information about the kind and concentration of the defects under considerations.

The positron annihilation technique is based on the positron capture in defects. In most applications, the positron is produced in the β^+-decay of the radioactive isotope ^{22}Na. The source is sandwiched between two identical pieces of the sample. The positron loses energy in the interactions with the material and slows down to thermal energy with the mean positron penetration depth of about 100 μm. Compared to the positron lifetime in solids, the thermalization process usually takes a rather short time, up to a few picoseconds, and can be neglected. The positron in thermal equilibrium with the environment begins the diffusion process with the diffusion length of about 100 nm. During the diffusion, the positron interacts with defects in the solid and the trapping of the positron into a localised state may happen. The diffusion length strongly determines the sensitivity of the positron method to detect defects. Finally, the positron annihilates with an electron, resulting mostly in the emission of two

Scientific paper
UDC: 539...17:519.876.5
BIBLID: 1451-3994, *18* (2003), 1, pp. 16–21

Authors' addresses:
[1]Faculty of Science, University of Tuzla, Univerzitetska 4, 75 000 Tuzla Bosnia and Herzegovina
[2]Sjukhusfysik, Boris Lasarett Södra Älvsborg Sjukhus, 50182 Boris, Sweden
[3]Reactor Physics Department, Chalmers University of Technology, S-41296 Göteborg, Sweden

E-mail address of corresponding author:
aadeno@bih.net.ba (S. Avdić)

gamma quanta of 511 keV and, in special cases, the annihilation of the positronium can lead to two or three annihilation gamma quanta. When the positron is trapped in a defect, the annihilation parameters are changed. Positron lifetime is longer due to the lower electron density and the positron-electron pair momentum distribution is wider at the defect site. This method is very sensitive to the open-volume defects such as monovacancies and larger vacancy clusters. The development of the slow-positron-beam technique has provided the examination of the regions close to the surface or the surface itself. Several experimental techniques of positron annihilation, such as the angular correlation of annihilation gamma quanta, Doppler broadening of the annihilation line and positron lifetime spectroscopy were developed for detecting defects in various materials [1].

A fast-fast coincidence spectrometer with a high time resolution is conventionally used for positron measurement. The positron lifetime spectrum is obtained as the number of annihilating positrons per time channel in the multi-channel analyser and is the absolute value of the time derivative of the positron decay spectrum. All positron experimental techniques provide the trapping rate that is proportional to defect concentration.

The construction of a pulsed beam for slow positrons is near to completion at the Reactor Physics Department, Chalmers University of Technology, Göteborg, Sweden. The final beam will be operated at 50 MHz with a 20 ns time window. The design of a sample chamber and an acceleration system which could locate the positron deeper in the samples will be finished soon. Besides the construction of the beam, there is an intention to run the conventional positron measurements *i. e.*, positron lifetime spectroscopy and Doppler broadening spectroscopy of the annihilation line, based on the positrons emitted from the source ^{22}Na [2].

There is a continuously increasing interest in the application of defect spectroscopy, as well as in the improvement of data processing. The main goal of the positron experimental data analysis is the determination of defect concentrations of the dominant positron trapping centres. Several computer codes were developed for this purpose. The computer program PATFIT [3], based on the Gauss-Newton fitting procedure, is used for the conventional data analysis in which the experimental lifetime spectrum is expressed as a sum of several exponential components convoluted with the instrument resolution function. In this method, one has to specify the number of spectral components before analysis.The accurate value of the time resolution is also required to get a reliable solution. The PATFIT analysis gives the average of the defect lifetime, not the lifetime distribution. The program code CONTIN [4, 5], based on the Laplace inversion technique and the program code MELT [6, 7], based on the maximum entropy principle and the Bayes' theorem, can give continuous distribution of positron lifetimes or annihilation rates. The latter program has the advantage that the number of the spectral components and the time resolution of the instrument are not required as the input parameters. Applications of the continuous lifetime analysis reported in literature show that the resolution of CONTIN is lower than that of MELT. In some cases (*e.g.*, probing of semiconductors), CONTIN cannot separate lifetimes close to each other, but gives only a peak broadening.

A new method based on the use of artificial neural networks (ANN) for unfolding the mean lifetime and intensity of the spectral components of simulated positron lifetime spectra was suggested and partially tested in our previous paper [8]. In the present work, the capabilities of the method to analyse experimental data is demonstrated in the case of three component spectra obtained from free-volume measurements of polymer materials [9]. The backpropagation ANN described in [8] has been trained by simulated spectra when expanding the range of the intensities and lifetime spectral components in order to cover the regions of the experimental parameter values. The variation of the parameter values was made much finer in order to obtain a higher precision of the neural network in the evaluation of the spectrum parameters. In general, a similar backpropagation ANN can be designed and tuned for any other number of spectral components. In order to achieve flexibility and to be able to unfold the spectra with an unknown number of spectral components, the standard ANN module has been combined with a special classification model. Classification modules based on self-organizing map (SOM) and learning vector quantization (LVQ) algorithms have been developed and their reliability in identifying the number of the spectral components tested.

EVALUATION OF THE EXPERIMENTAL POSITRON SPECTRA PARAMETERS

A hundred experimental positron spectra with three spectral components obtained from polymer materials were available for the ANN test (paper VI in [9]). A portion of them is shown in fig. 1. PATFIT analysis of the spectra has served as a basis for the evaluation of the validity of the ANN. A neural network of the same type as in the earlier work [8] has been constructed, namely a three-layered feed-forward network with backward error propagation. Simulated lifetime spectra have been generated by the Posgen code [10] by an algorithm based on eq. (1). In the case of discrete lines in the

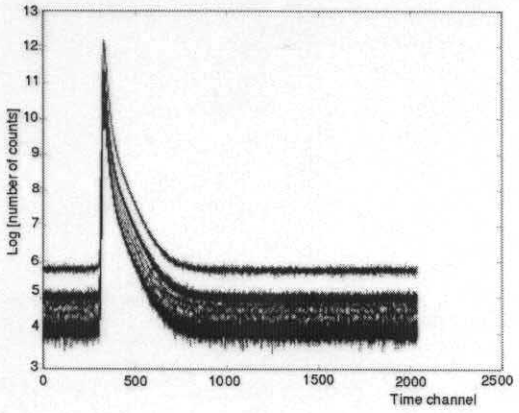

Figure 1. Experimental positron spectra

lifetime profile, a measured spectrum $Y(t)$ can be expressed in the form [11]:

$$Y(t) = R(t) \cdot (N_t \sum_{i=1}^{n} I_i e^{-\lambda t} + B) \qquad (1)$$

where $R(t)$ is the detector resolution function with which the measured signal is convoluted, B is a random background, N_t is a total number of counts, I_i and λ_i are the intensity and positron annihilation rate of the i-th spectral component, respectivelly, and n is the number of spectral components. Actual experimental parameters, such as the time channel width and the FWHM of the Gaussian component of the resolution function, have been included in the Posgen code. In order to obtain a higher accuracy and resolution of the network, it was essential to generate spectra over the whole region of the possible variations of the experimental values of lifetimes and intensities when building a fine mesh. A set of 18200 has been used in the trainings process. Both simulated and experimental spectra have been normalized to the maximum value and the start-up region, where the signal rises very fast to reach maximum, has been excluded.

The number of the input nodes of the network was set equal to the number of time channels. The number of the output nodes is equal to the number of parameters to be determined, *i. e.*, three intensities and three lifetimes in our case. The number of nodes in the hidden layer was found by trial end error. In our case, its optimal value of 40 nodes is the same as in the earlier work [8]. Several faster algorithms were applied for the training of the network [12]. A resilient algorithm was chosen as the optimal one since it shows an advantage in the instance of a large input, when the slope of sigmoid function approaches zero. In that case, a very small magnitude of the gradient can cause small changes in the weights and biases, even though the weights and biases are far from their optimal values. The resilient backpropagation-training algorithm eliminates these effects and takes into account only the sign of the derivative to determine the direction of the weight update. A set of additional 90 spectra, non-identical with any of the training samples, was generated by simulation [10] for testing the accuracy of the network tuning.

Two versions of the ANN code have been used in the training process, namely, a serial code running on a SUN ULTRA 10 station, and a parallel version implemented on the SGI Origin 2000 parallel computer at Chalmers university [13]. The tuned network reproduces the intensities and lifetimes of the 90 test patterns with an accuracy given in tab. 1. The average relative error with respect to the target values and the standard deviation of the error taken over all test samples are shown. The results from the evaluation of the experimental spectra (the recall phase of the ANN) are presented in fig. 2a and fig. 2b for the intensities and the lifetimes, respectively.

Table 1. The results of testing the neural network performance with 90 test patterns

	I_1[%]	τ_1 [ns]	I_2 [%]	τ_2 [ns]	I_3 [%]	τ_3 [ns]
Range of parameters	14-34	0.16-0.24	42-52	0.38-0.48	24-34	1.65-2.85
Average relative error	1.5	0.002	0.4	– 0.02	0.05	– 0.02
Standard deviation of relative error	12.5	0.5	6.3	2.1	2.2	1.6

As can be seen, standard deviations are larger for the intensities than for the lifetimes. Accuracy is better for parameters of the second and third spectral component. In general, precision can be improved by further refining the training parameter mesh and by increasing the number of iterations in

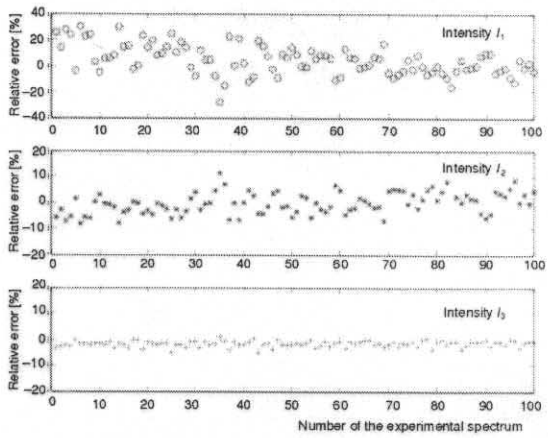

Figure 2a. Results of the experimental spectra evaluation (Recall error of the intensities)

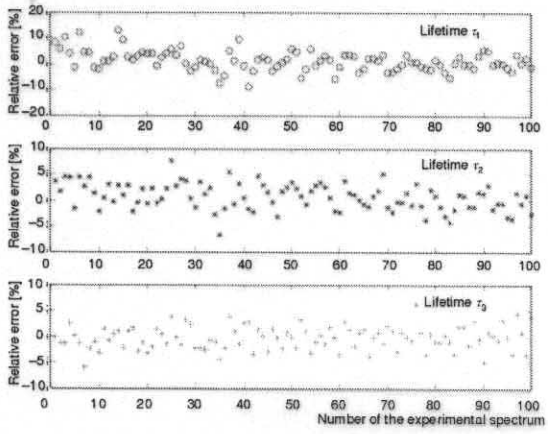

Figure 2b. Results of the experimental spectra evaluation (Recall error of the lifetimes)

the training process. It has to be noted that the performance of the ANN-unfolding technique in practical applications depends very much on the quality of the simulation model. A more adequate simulation model for the training samples would give better agreement between the results of the neural network unfolding procedure and the experimental values.

CLASSIFICATION OF POSITRON LIFETIME SPECTRA BY SOM AND LVQ ALGORITHMS

It was shown [8] that the backpropagation neural network is not appropriate for the identification of both the parameters and the number of spectral components. This would require more training samples, longer training times and a large number of iterations that have prohibitively large values for practical applications. The unknown number of components represents an important input parameter for the Posgen simulation code. In addition, for the evaluation of spectra with a different number of spectral components, separate backpropagation networks with a different structure should be used for practical reasons. For these reasons, a new network was designed only for identifying the unknown number of components. SOM as well as LVQ algorithms have been applied. The complete ANN analysis tool of positron lifetime spectra includes the SOM or LVQ algorithms for selecting the number of spectral componenets and a backpropagation neural network for determining the lifetimes and intensities.

Unsupervised neural networks are typically used to extract features within data. The SOM algorithm [14] does not utilise class information and its results are inherently suboptimal. A self-organizing map is particularly useful for visualisation purposes. It maps a high dimension space to 2D space and maintains the topology mappings. A topological map is simply a mapping that preserves neighbourhood relations. There is only one input and one output layer, no hidden layer. Neurons in the output layer are arranged on a regular flat grid. Each neuron has an associated prototype vector and after training, neighbouring neurons have similar prototype vectors. The SOM algorithm finds the closest matching neuron to the training input and moves the weights of this neuron and those in the neighbouring proximity towards the input vector. The algorithm [15] described above was applied for the extraction of the cluster structure of the simulated positron lifetime spectra with two and three components. Simulated spectra with 2 and 3 components have been generated by the Posgen code to test the classification module.

The input layer of the SOM network consists of the simulated spectral data. The output layer is the map constituted of neurons in 2D space. The neurons learn to recognize groups of similar input vectors in such a way that neurons physically close together in the map grid respond to similar input data. The determination of the map size is based on the amount of data vectors and the principal eigenvectors of the input data. The input samples are normalized so that each component has unit variance. The training is done in two phases: first with large neighbourhood radius and then fine-tuning with small radius. The distance matrix technique can show the cluster structure of the spectral data. The "U-matrix" visualises the cluster structure and shows distances between neighbouring units of the map, as well as the median distance from each map unit to its neighbours. High values on the U-matrix mean large distances between neighbouring map units and indicate cluster borders. The U-matrix obtained for the simulated positron spectra with two and three spectral components shows two different groups of spectral data with cluster borders, fig. 3. The applied technique is appropriate for the visualisation of the cluster structure of spectral data, especially when a huge set of positron spectra should be analysed. However, it cannot give clear information on the classification of spectral data in input space, since the visualisation technique is related to the 2D map of neurons. In that case the map cannot follow the spectral data in details, since the input space dimension is much higher than the 2D map. In order to get a better physical interpretation of the results, the map can be labelled with corresponding labels. Results displayed in fig. 3 show that the SOM visualisation technique is convenient only for a rough estimation of the cluster structure of spectral data. Apart from this, the SOM algorithm does not give correct results in a wide range of spectral data, especially when Euclid distance be-

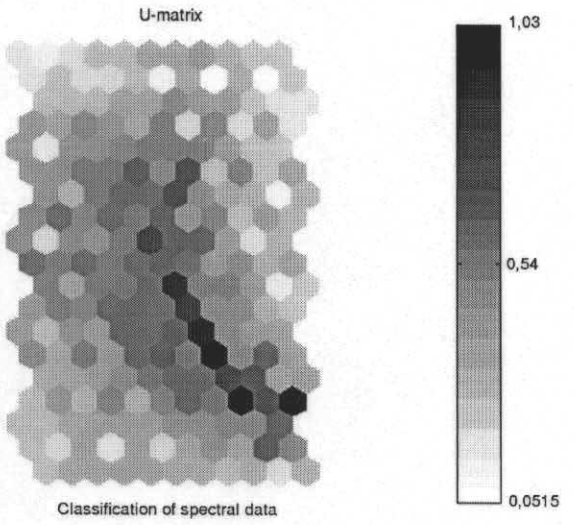

Figure 3. SOM visualisation results and U-matrix for the simulated positron spectra with two and three spectral components

tween the data from different positron spectra are too close to each other.

Therefore, a new network based on the LVQ algorithm has been designed. The unsupervised neural network has been transformed into a supervised LVQ neural network taking into account the class information. The network architecture is similar to a SOM, but without a topological structure. Two types of LVQ neural network were constructed, *i. e.*, with and without visualisation [16]. The input layer contains the spectral data with the class information. The LVQ network without visualisation includes a competitive layer for finding subclasses of input vectors and combining them into target classes. The output layer of the LVQ network with visualisation contains the map of neurons with

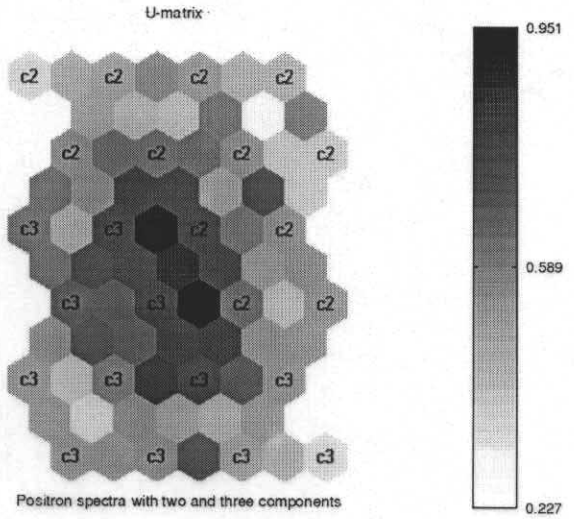

Figure. 4. LVQ visualisation results of the classification of simulated positron spectra with two and three components

the class information. The LVQ technique without visualisation gives correct results in 97% of the cases for all experimental positron spectra. This kind of analysis provides a long output containing the class information for each channel of spectral data. The other kind of LVQ analysis with visualisation gives a faster response, but the results are related to the map of neurons, *i. e.*, not to the input space. In this way one can obtain only a rough information on the classification of positron spectra, depending on the number of spectral components. Results of the LVQ visualisation for the simulated positron spectra with two and three components are presented in fig. 4.

Experimental positron spectra have been analysed by the SOM algorithm with visualisation technique. The number of the input nodes of the network was set equal to the experimental data. The neuron on the 2D map grid, that constitutes the output layer, arranges itself so that neighboring neurons recognize similar input vectors. Analysis of experimental positron spectra by SOM visualisation technique, fig. 5, correctly shows one class of data corresponding to polymer spectra with three spectral components. The map of neurons does not indicate cluster structure. Values on the U-matrix indicate the same result in comparison with U-matrix for the simulated spectra with two and three components. However, in a wider range of positron spectra in which the spectral data are too close to each other, the SOM technique does not give correct results. For more precise information on it, one can apply the LVQ technique without visualisation to get classification results per channel of input space.

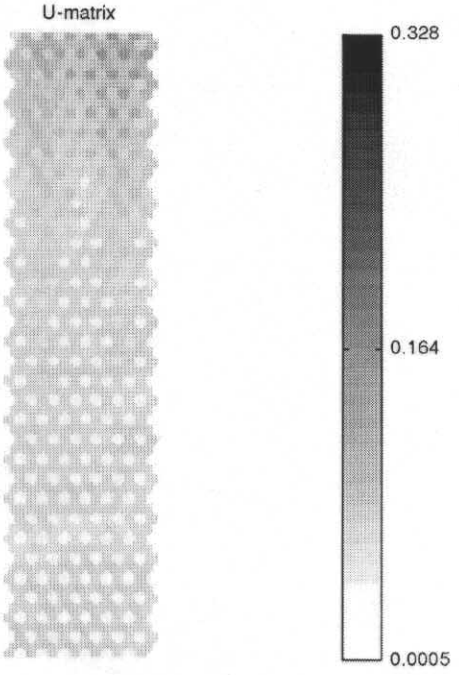

Figure 5. SOM visualisation results for experimental positron spectra

CONCLUSION

The applied techniques, based on backpropagation, SOM and LVQ algorithms, show an advantage in respect to traditional methods for processing positron spectral data, when positron spectra with unknown components should be analysed with high resolution of the lifetimes. There is a further possibility for the improvement of the architecture and the training algorithms of the designed neural networks. A more adequate simulation model for generating the simulated positron spectra would provide a higher quality of input data, better agreement with the target values and, consequently, higher competitiveness with the conventional methods. In general, the ANN method has the potential to analyse other features of the PAL spectra, for example lifetime broadening (the FWHM of the distribution function for each lifetime).

ACKNOWLEDGEMENTS

The authors would like to thank Dr. Marcus Schmidt for making available the experimental spectra for the ANN study and for the valuable discussions.

REFERENCES

[1] Krause-Rehberg, R., Leipner, H. S., Positron Annihilation in Semiconductors, Defect Studies, Springer Series in Solid-State Sciences, 1999
[2] ***, Department of Reactor Physics, Report of the Chalmers University of Technology, Göteborg, Sweden, 2002
[3] ***, PATFIT Package, Riso National Laboratory, Roskide, Denmark, 1989
[4] Provencher, S. W., A Constrained Regularization Method for Inverting Data Represented by Linear Algebraic or Integral Equations, *Comput. Phys. Commun. 27* (1982), pp. 213–227
[5] Gregory, R. B., Zhu, Y., Analysis of Positron Annihilation Lifetime Data by Numerical Laplace Inversion with the Program CONTIN, *Nucl. Instr. and Meth. A., 290* (1990), pp. 172–182
[6] Hoffmann, L., Shukla, A., Peter, M., Barbiellini, B., Manuel, A. A., Linear and Non-Linear Approaches to Solve the Inverse Problem: Application to Positron Annihilation Experiments, *Nucl. Instr. and Meth. A, 335* (1993), pp. 276–287
[7] Shukla, A., Peter, M., Hoffmann, L., Analysis of Positron Lifetime Spectra Using Quantified Maximum Entropy and a General Linear Filter, *Nucl. Instr. and Meth. A, 335* (1993), pp. 310-317
[8] Pázsit, I., Chakarova, R., Linden, P., Maurer, F. H. J., Unfolding Positron Lifetime Spectra with Neural Networks, *Applied Surface Science, 149* (1998), pp. 97–102
[9] Schmidt, M., Macroscopic Volume and Free Volume of Polymer Blends and Pressure Densified Polymers, Ph. D. thesis, Department of Polymer Technology, Chalmers University, Göteborg, Sweden, 2000
[10] Pedersen, N. J., Numerical Code Posgen, 1991
[11] Chen, Z. Q., Wang, Z., Wang, S. J., Application of Positron Lifetime Distribution to the Discrimination of Defects in Semiconductors, *Nucl. Instr. and Meth. in Physics Research B, 160* (2000), pp. 139–148
[12] ***, Neural Network Toolbox for Use with MATLAB, User's Guide, The MathWorks, Inc.
[13] Lindén, P., Chakarova, R., Faxén, T., Pázsit, I., Lecture Notes in Computer Science, Springer Verlag, 1999, Vol. 1593, p. 1194
[14] Kohonen, T., Self-Organization and Associative Memory, 2nd ed., Springer-Verlag, Berlin, 1987
[15] Vesanto, J., Himberg, J., Alhoniemi, E., Parhankangas, J., SOM Toolbox for Matlab 5, Report A57, Helsinki University of Technology, Finland, 2000
[16] Avdić, S., Numerical Code for Parameters Evaluation and Classification of Positron Spectra by SOM and LVQ Algorithms (unpublished paper), University of Tuzla, Bosnia and Herzegovina, 2001

Сенада АВДИЋ, Румјана ЧАКАРОВА, Имре ПАЖИТ

АНАЛИЗА ЕКСПЕРИМЕНТАЛНИХ СПЕКТАРА ВРЕМЕНА ЖИВОТА ПОЗИТРОНА ПОМОЋУ НЕУРАЛНИХ МРЕЖА

У раду се анализирају експериментални спектри времена живота позитрона који су добијени у полимерским материјалима са дефектима помоћу различитих алгоритама неуралних мрежа. Нов метод, заснован на коришћењу вештачких неуралних мрежа за евалуацију времена живота и интензитета спектралних компоненти симуларних позитронских спектара, предложен је и парцијално тестиран у претходном раду [Pázsit *et al*, *Applied Surface Science, 149* (1998), 97]. Овде је верификована применљивост метода за анализу експерименталних позитронских спектара са три спектралне компоненте у полимерским материјалима. Показано је да се помоћу бекпропагационе неуралне мреже могу одредити спектрални параметри са високом тачношћу и да се може извршити декомпозиција времена живота позитрона који се разликују за 10% и више. Пошто стандардна бекпропагациона неурална мрежа није погодна и за евалуацију параметара и за идентификацију непознатог броја спектралних компоненти, конструисана је посебна мрежа базирана на SOM и LVQ алгоритмима за решавање проблема класификације спектара. Утврђено је да LVQ алгоритам има боље перформансе и већу поузданост.