A. Benninghoven, P. Bertrand, H.-N. Migeon and H.W. Werner (Editors).
Proceedings of the 12th International Conference on Secondary Ion Mass Spectrometry,
Brussels, Belgium, 5-11 September 1999



# NEUROSPECTRANET - A SELF-ORGANISING NEURAL NETWORK MECHANISM FOR INTERPRETATION OF COMPLEX SSIMS SPECTRA.

O.D.Sanni, A.Henderson, D.Briggs and J. C.Vickerman

Surface Analysis Research Centre, UMIST, Sackville St, Manchester M60 1QD, UK
email: debo.sanni@stud.umist.ac.uk

## 1. Introduction

Interpretation of SSIMS spectra requires a great deal of experience and costly expertise. Human experts often support their knowledge by searching a library of known spectra for a suitable match to the unknown spectrum in a process that can be very tedious, time-consuming and uneconomic for industrial surface analysis. To make the power of SSIMS more widely and easily accessible to R&D and Quality Control laboratories, an intelligent automated search engine capable of recognising and classifying SSIMS spectral patterns is required to enhance rapid and efficient interpretation of spectra.

*NeuroSpectraNet* is a novel and unique neural networks mechanism designed to overcome these difficulties by providing an efficient search engine requiring little or no human assistance for an effective determination of the unknown spectrum.

### 1.1 Why Neural Networks For SSIMS?

Neural networks are powerful computational software and hardware techniques designed to parallel the operations of the human brain in various tasks of pattern recognition, classification and intelligent data analysis. Currently regarded as the greatest technological advancement since the advent of transistors, neural networks consist of computational processing elements, called neurons, which carry out some specific computational task integral to the overall objective. Unlike classical data analysis, neural networks achieve their objectives by learning the data (which describes the unknown) and inductively reasoning out a solution, rather than being pre-programmed for a specific solution. Each of the various existing neural network paradigms combine the abilities and power of several statistical mechanisms into one powerful algorithm providing robustness against noise and outliers, high speed of operation, adaptive capability and better power of discrimination and prediction.

SSIMS (Static SIMS)spectra are often complex and generally impure in nature. Furthermore, the molecular-ion is not observed in many SSIMS spectra, and ion re-combinations can complicate analysis. For these reasons, neural networks may well become invaluable for intelligent interpretation of SSIMS by fully exploiting the following neural power:

*i.* Ability to learn, adapt and remember;
*ii.* Efficiency at modelling non-linearities;
*iii.* Robustness against inadequate and inconsistent data
*iv.* Extensive knowledge indexing and automatic data abstraction;
*v.* Ability to self-organise and readily generalise solutions;
*vi.* Ability to self-generate own model.

## 2. NeuroSpectraNet

*NeuroSpectraNet* was designed to take advantage of both the self-organising and supervised properties of neural networks. The self-organising mechanism enables the network to organise spectral data according to their chemical similarities, while the supervised mechanism supports the ability to interpret a spectrum that is somewhat unique and not directly similar to any spectrum in the database. The supervised mechanism achieves this by detecting some specific chemical properties that the spectrum exhibits in common with some other spectra in the database. Thus, if the self-organising mode fails to find a match for the spectrum, interpretation can be assisted by switching to detection of chemical functionalities that may be evident in the spectrum.

Unlike other methods used for mass spectra analysis including MsNet [1], PCA [2], and STIRS [3], *NeuroSpectraNet* does not require knowledge of the molecular-ion or any specific significant peaks. Instead, it self-generates its own model of significant micro-features (peaks) from the raw composite spectrum of positive-ion and negative-ion spectra of a material.

### 2.1 Design Principle

The backbone of the design philosophy of *NeuroSpectraNet* is based on a self-organising mechanism similar to the ART2 [4] neural network which is capable of organising input data into categories of similar patterns without *a priori* training. The mechanism is also able to mimic the technique of spectral library searching more as an intelligent search engine. Initially, *NeuroSpectraNet* models various SSIMS patterns into categories of chemical similarities. An unknown spectrum may then be tested for a match with the models of the generated categories. If no match is found, the unknown sets up a new category with its model. The model generated from such a unique spectrum is then analysed to identify any chemical functionalities that may be present in the spectrum. Fig.1 illustrates the principle of this self-organising mechanism.

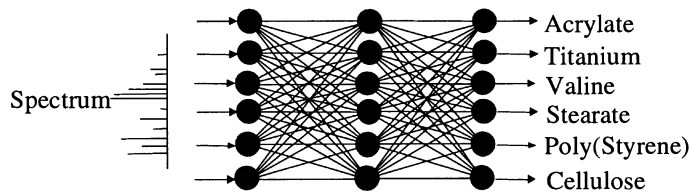


Fig.1. Illustration of NeuroSpectraNet

### 2.2 Data & Computation

The positive-ion and negative-ion stick spectra of a material are normalised to the total intensities of their respective masses. The normalised positive-ion spectrum is then appended to the negative-ion counterpart to form a composite spectrum used as input to *NeuroSpectraNet*.

#### 2.2.1 Prototype Generation and The Self-Organising Mechanism

The first task is to generate a prototype that contains all the salient features of the original spectrum. The prototype generation is performed by a multi-layer set of neurons. Each of the neurons in a layer is dedicated to computing the spectral information of a mass position. The composite spectrum is modified within the layers by a process that successively re-normalise the composite spectrum and provide contrast enhancements until the prototype generation process stabilises. Since low-intensity peaks at high mass are significant in SSIMS, *NeuroSpectraNet's* contrast enhancement provides a

modified sigmoid function to boost low-intensity peaks at high mass. Automatic bias filteration is also implemented within the mechanism to set a noise-free threshold for each prototype.

The first prototype to be generated is stored as an archetype representing the category of materials the spectrum belongs to. On encountering the second and any subsequent spectrum, the new prototype is compared for a match with the stored archetype of each of the existing categories. *NeuroSpectraNet's* matching process involves several numerical tests of similarity and dissimilarity in which the contribution of each mass in the unknown spectrum is aggregated to produce a unique validation index that is assessed to confirm whether or not the spectra are chemically and structurally similar. The various tests are designed to numerically emphasise the spectral differences between dissimilar materials while similarities of spectra of like materials are enhanced. As a result of the complexity of many SSIMS spectra no single test suffices to cover all the observed spectral differences. Therefore, the various tests operate to produce a joint decision, thus fully utilising the advantage of breaking a complex task into easier sub-tasks that are handled in parallel by different processing elements.

If the archetype of a category is found largely similar to the new spectrum's prototype, it is entered into a pool of potential winners while the archetypes of remaining categories are similarly evaluated for a match. If more than one potential winner exists after all the categories have been tested, the potential winners are then re-assessed in an arbitration module to determine which of them has the best correlation to attract the unknown into it's category's membership. If only one winner exists in the pool, its category automatically becomes the winner.

However, if no match is found, the prototype of the unknown sets up a new category for attracting spectra of materials similar to it. Additional tests are then undertaken to detect pre-defined functionalities that may be present in the new spectrum via a process described in section §2.2.2.

### 2.2.2 Supervised Training and Functionalities Identification

The prototypes generated from the various spectra are ideal candidates for training a network for detection of some specific functionalities which may be hidden in the spectra of many materials. Through a supervised extension, the self-organising *NeuroSpectraNet* learns via examples, discovers and models the salient characteristics of a functionality. The functionalities present in an unknown spectrum then becomes identifiable through the models thus generated. Thus, *NeuroSpectraNet* is able to utilise collective information available throughout the database to efficiently model functionalities by self-discovery of the characteristics of a functionality, to make it ideal for SSIMS which often contain complex mixture of functionalities.

This added advantage to *NeuroSpectraNet's* self-organising mechanism makes it a powerful tool for providing information on specific chemical functionalities that may be present in the unknown.

## 3. Results and Discussion

Our exploratory design based on ART2 [4] network correctly grouped closely-related materials together when tested on a small range of materials. However, the classification efficiency degraded over a broader range of materials. The phenomenon of re-grouping was also very evident as membership of categories readjusted when similar spectra of some materials that were previously categorised into some groupings decamped to some more recently-formed categories to which they were indeed closer. This, we found, was due to the original design of ART2 [4] being unsuitable for

SSIMS where there is a very strong commonality amongst spectra of dissimilar materials. With *NeuroSpectraNet*, the re-grouping phenomenon is not observed; efficiency of classification is good over an unlimited range of materials, as the discrimination tests are comprehensive and numerically efficient. A typical classification such as Fig. 2 shows that there is a good correlation between spectra of many materials that have structural similarities.

| | |
|---|---|
| Poly(2-chloroethyl methacrylate) | 1 |
| Poly(methacrylic acid) (ToF) | 0.940338 |
| Poly(n-lauryl methacrylate) | 0.928526 |
| Poly(n-butyl methacrylate) | 0.924523 |
| Poly(methyl methacrylate - co - PEG 1000 methacrylate), (12.5% PEGMA) | 0.911585 |
| Poly(methacrylic acid) (Quad) | 0.910734 |
| Poly(methyl methacrylate) (ToF) | 0.898974 |
| Poly(cyclohexyl methacrylate) | 0.898097 |
| Poly(ethyl methacrylate) | 0.897175 |
| Poly(2-hydroxyethyl methacrylate) | 0.893147 |
| Poly(n-hexyl methacrylate) | 0.886860 |
| Poly(2-hydroxypropyl methacrylate) | 0.880651 |
| Poly(methyl methacrylate) (Quad) | 0.862834 |
| Poly(n-propyl methacrylate) (Quad) | 0.794553 |

Fig.2. A Typical NeuroSpectraNet Grouping of Spectra of Similar Materials, Showing Their Correlation Factors Against Poly(2-chloroethyl methacrylate).

NeuroSpectraNet's classification is fast with an average classification time of 10 sec for a 2x1500 a.m.u spectrum search of 450 spectra in *The Static SIMS Library* [5].

## 4. Conclusion

NeuroSpectraNet has demonstrated that neural processing power can be harnessed for analysis of SSIMS spectra. If a spectrum similar to an unknown exist in NeuroSpectraNet's database, it would efficiently and correctly identify the unknown in a matter of seconds. If an unknown is so unique that it is completely new to NeuroSpectraNet's self-organising mechanism, NeuroSpectraNet is designed to help the analyst to correctly identify pre-defined functionalities that may be present in the unknown. Because the analysis is fast and reliable, the spectrometrist can focus his attention on more complex tasks better armed.

## References

[1] B. Curry and D. E. Rumelhart, Msnet: A Neural Network Which Classifies Mass Spectra, Tetrahedron Computer Methodology, 3 (1990) 213.
[2] K-S. Kwok, R. Venkataraghavan and F. W. McLafferty, J. Am. Chem. Soc. 95 (1973) 4185.
[3] J. Arunachalam and S. Gangadharan, Feature extraction from spectral and other data by principal components and discriminant function techniques, Anal. Chim. Acta 157 (1984) 245.
[4] G.A. Carpenter and S. Grosberg, ART2: Self-Organization of Stable Category Recognition Codes For Analog Input Patterns. IEEE 1st International Conference on Neural Networks, II (1987) II-735.
[5] The Static SIMS Library, Surface Spectra Ltd, Manchester, UK.