Chemical Analysis of AES, XPS and XRD Data using Self Organising Maps

K. Obu-Cann*, H. Tokutaka*, K. Fujimura*, K. Yoshihara**, and Metal Materials Group of SASJ.

*Dept. of Electrical and Electronics Engineering, Tottori University, Koyama-Minami 4-101, Tottori 680, Japan
Email:tokutaka@ele.tottori –u.ac.jp
*National Research Institute for Metals,1-2-1 Sengen, Tsukuba 305, Japan
Email:kazuhiro@nrim.go.jp

(Received September 28 1998; accepted February 2 1999)

This paper presents a report on the application of self-Organised Maps developed by Teuvo Kohonen [1] to chemical spectra. With the application of SOM method, the composition of unlabeled spectra can be determined. In this study, the data mining capabilities of SOM are examined using data from Auger Electron Spectroscopy (AES), X-ray Photoelectric Spectroscopy (XPS) and X-ray Diffraction (XRD) for FeNi alloys. The results obtained are compared to determine which data is more adaptive to the SOM.
KEYWORDS: SOM, data mining, and chemical data analysis

1. Introduction

Self-Organising maps which was first developed by Teuvo Kohonen [1] was first applied to information processing. One characteristic of the SOM is its ability to make multidimensional data visible on a 2 dimensional map.
SOM can be considered as a grid with predefined but originally empty nodes. During learning, the pattern of filling of the nodes is determined by the degree of similarity between the data. Thus it is possible to distinguish between similar and dissimilar data on a 2 dimensional SOM. The magnitude of the distance between the nodes is shown by a grey level expression.
The spectral data as well as the alloy composition from the AES, XPS and XRD data are considered as multidimensional characteristics.
The objective of this experiment is to determine the alloy composition from the spectral data and to compare results obtained from the SOM for the three types of data i.e. AES, XPS and XRD.

2. The SOM Algorithm

T. Kohonen [1] developed an equation, which governs the information content of a unit in the grid based on the information processing ability of the brain.

$$m_i(t+1) = m_i(t) + \alpha(t)[x(t) - m_i(t)] \qquad (1)$$

Where $\mathbf{m}_i(t)$ is the information processing ability of the neuron cell (node) **i** at time t and **x** (t) is the input signal. At time t, the cell learns this input signal. During time (t+1), the information processing ability of the cell becomes $\mathbf{m}_i(t+1)$. If **x** (t) is an n-dimensional input vector, then **x** (t) $=[\xi_1, \xi_2, \ldots, \xi_n]$ and the n-dimensional reference vector $\mathbf{m}_i(t) =[\mu_{i1}, \mu_{i2}, \ldots, \mu_{in}]$. $\alpha(t)$ is the learning coefficient factor with values between 0 and 1. When an n-dimensional input vector is introduced to the network, the reference vector in the network (node) that is closest to the input vector is defined as the best-matching node "winner" and its information processing ability is denoted by $\mathbf{m}_c(t)$. Prior to learning, a large reference area surrounding the winner is selected as a neighbourhood region. The reference unit vectors in this neighbourhood region $N_c(t)$ as well as the winner $\mathbf{m}_c(t)$ learn the input vector **x** (t) following eq.1. This forms a typical learning cycle. The next cycle begins with the introduction of the next input vector but with a reduced

neighbourhood region. Learning continues until only the winner is trained. Furthermore, $\alpha(t)$ reduces to 0 as learning progresses. For the case of this experiment, $\alpha(t)$ is a linear decreasing function.

$$\alpha(t) = \alpha_0(1 - t/T), \qquad (2)$$

Where α_0 is the initial value, t is the present learning cycle and T is the maximum number of the learning cycles after which learning is terminated. Other types of decay functions for α can be considered.

3. Chemical Data Mining Analysis

Binding energies of AES data from 498eV to 1000eV in increments of 1eV are considered as dimensional units. Each spectrum is a 503 dimensional input vector. The signal value of the spectrum are then normalised to values between 0 and 1 and used as input vector for the SOM [2]. 5 samples of FeNi alloys (Fe100Ni0, Fe80Ni20, Fe50Ni50, Fe30Ni70 and Fe0Ni100) spectra shown in Fig. 1 are used as input data.

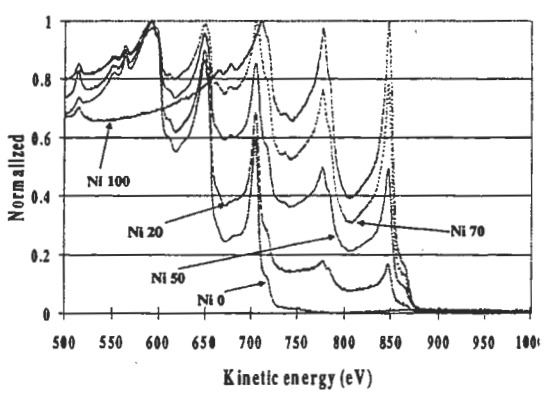

Fig. 1 Normalised AES data of 5 FeNi Alloys from 500 to 1000eV.

For the purposes of data mining, the dimensional values of Fe50Ni50 are excluded from the input data. The compositions of the alloys are also considered as new dimensional units and are also expressed within the values of 0 and 1. For instance Fe80Ni20 alloy has 20% Nickel and is therefore denoted by 0.2.
SOM is constructed from the multidimensional data of Fig. 1 into a 2 dimensional 20X30-neuron unit using grey level expression in Fig. 2

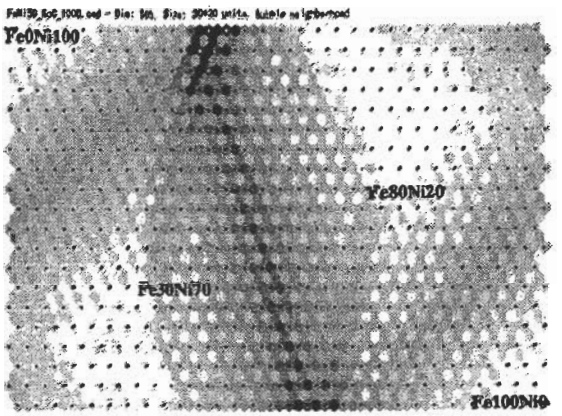

Fig. 2 SOM of AES on 20X30 (600) unit grid. 4 kinds of FeNi alloys (Ni 0, 20, 70, 100%)

After SOM learning, all the 600 units or nodes on the grid are compared by the Error function Err:

$$Err = \sum_{j=1}^{n} (x_j - m_{ij})^2, \qquad (3)$$

Where x_j and m_{ij} are the j-th component of the n-th dimensional input data and the i-th unit (node) respectively. The labelled positions of the 20X30-unit grey scale map are determined by the minimum value of the Error function.
Fe50Ni50 alloy was excluded in the input data and is now used as test data to test the generalising ability of the network. By use of eq. 3, all the 600 nodes or units of Fig. 2 are compared with the test data of Fe50Ni50 alloy. The unit with minimum value of the Error function is identified as the best match unit. For this experiment, the closest unit was identified as Fe50.2Ni49.8 as shown in Fig. 3 with a composition error of 0.2%. Fig.4 illustrates the frequency distribution of the composition of all the 600 units of the SOM of Fig.2.

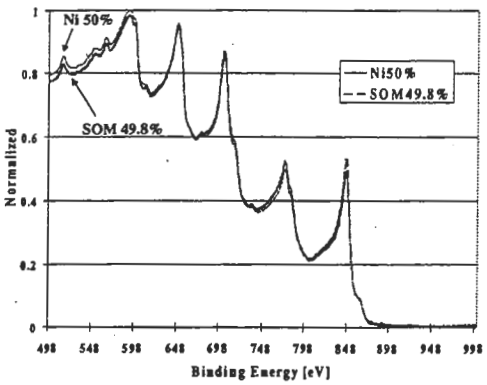

Fig. 3. AES spectra of the original Fe50Ni50 input data and the learned best match spectra.

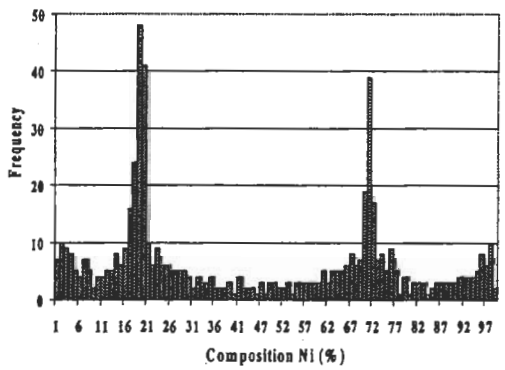

Fig. 4 Frequency distribution of the compositions of all the 600 units in Fig. 2

The distributions were found to be around 10 in the regions around 0 and 100% and slightly higher around 20 and 70% because they constitute the input data, which was learned during SOM. Spectral data for Fe80Ni20 and Fe30Ni70 were also used as test data for the data mining exercise and the best match spectra identified. The above experiment was repeated for the following:

(a). XPS data for Fe2p FeNi alloys with binding energies from 700eV to 735eV in increments of 0.05eV.

(b) XPS data for Ni2p FeNi alloys with binding energies from 840eV to 890eV in increments of 0.05eV.

(c) XRD data for FeNi alloys with diffraction angles from 50 to 116 degrees in increments of 0.02 degrees.

Using the same approach as discussed above. The data that best matched the Ni50% alloy for AES was identified to be Ni49.8% with an error of 0.2%. The same approach was repeated for XPS (Fe2p) and (Ni2p) data. Here, the best match spectrum was 41.4% with an error of 8.6% and 58.3% with an error of 8.3% respectively. Similarly the best match spectra for XRD Ni50% was Ni25.7% with an error of 24.3%.

4. Simulation Results

From the results obtained in the experiment, it is obvious that the SOM is a good tool for data mining. The composition of unlabeled spectra can be determined using SOM. The composition error and mean squared error of the various data used are shown in Figs. 5 and 6.

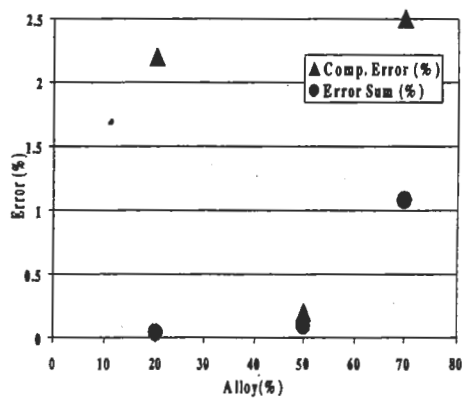

Fig. 5a Mean squared error and composition error of the best match spectra for Ni20, Ni50 and Ni70 after learning for AES Spectra.

Of the 4 different types of data used (AES, XPS {Fe2p}, XPS {Ni2p}, XRD), AES data had the lowest composition error and mean squared error followed by XPS {Fe2p}, XPS {Ni2p} and XRD data respectively. During the process of learning, the similarity between the various input vectors is measured using the Euclidean distance method. This is to determine which cluster or pattern class each vector should be associated with. The input data though independent must

be somewhat related. For spectral data, the relationship is found in the spectral shapes.

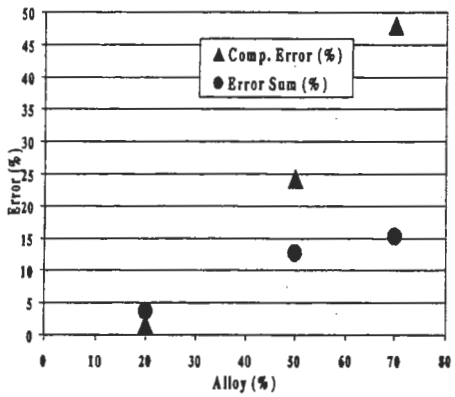

Fig. 5b Mean squared error and composition error of the best match spectra for Ni20, Ni50 and Ni70 after SOM for XRD Spectra.

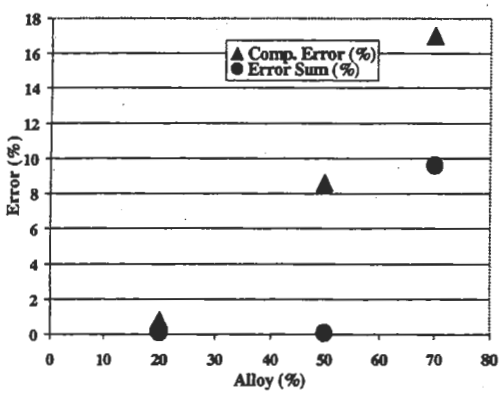

Fig. 6a Mean squared error and composition error of the best match spectra for Ni20, Ni50 and Ni70 after SOM for XPS (Fe2p) Spectra.

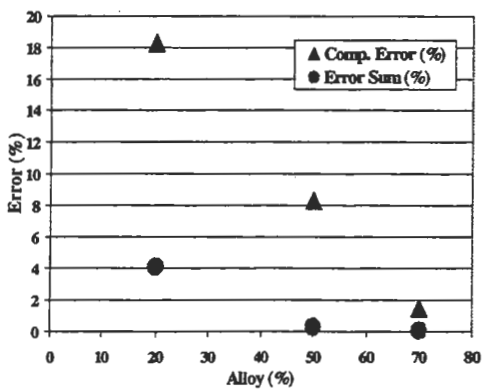

Fig. 6b Mean squared error and composition error of the best match spectra for Ni20, Ni50 and Ni70 after SOM for XPS (Ni2p) Spectra.

XRD data was found to have very different spectral shapes thereby resulting in high Euclidean distances and high errors in the pattern classification.

5. Conclusion

The main advantage of SOM is that multidimensional input data can be sorted and made visible on a 2 dimensional SOM. In total, 4 different types of data were used in this experiment to examine the data mining capabilities of SOM. Results obtained indicate that the composition of unlabeled spectra can be determined from the spectral data learned by the SOM.

Furthermore, if a labelled spectrum with in-complete characteristics or dimensions is placed on the grid, the missing characteristics or dimensions can be derived from the data that has already been assigned to the empty grid through the learning process. Thus it can be concluded that SOM is very effective for chemical data mining

As a first step, it is necessary to normalise the input data. For this experiment, normalisation refers to re-scaling by the minimum and the range of the input spectrum.

({Input–Min.Input}/{Max. Input – Min. input}).

This may have caused the higher composition error and mean squared error that were obtained for the case of XPS signals. The next step of the experiment will be to normalise the XPS data by the minimum and the maximum in the data range of the spectrum.

References

[1] T. Kohonen, Self-Organising Maps, Springer Series in Information Sciences, Vol.30, 1995.

[2] H. Tokutaka, Application of Self Organising Map to Chemical Analysis.. Proceedings of ICONIP'97, pp. 1318-1321 (1997)