ELSEVIER

Computers & Geosciences 33 (2007) 821–828

COMPUTERS & GEOSCIENCES

www.elsevier.com/locate/cageo

Clustering of volcanic ash arising from different fragmentation mechanisms using Kohonen self-organizing maps

Orkun Ersoy[a,b,*], Erkan Aydar[a], Alain Gourgaud[b], Harun Artuner[c], Hasan Bayhan[a]

[a]*Department of Geological Engineering, Hacettepe University, 06532, Beytepe-Ankara, Turkey*
[b]*Universitè Blaise Pascal, UMR-CNRS 6524, 5 rue Kessler, 63038 Clermont-Ferrand, France*
[c]*Department of Computer Science & Engineering, Hacettepe University, 06532, Beytepe-Ankara, Turkey*

Received 31 October 2005; received in revised form 17 October 2006; accepted 25 October 2006

Abstract

In this study, we present the visualization and clustering capabilities of self-organizing maps (SOM) for analyzing high-dimensional data. We used SOM because they implement an orderly mapping of a high-dimensional distribution onto a regular low-dimensional grid. We used surface texture parameters of volcanic ash that arose from different fragmentation mechanisms as input data. We found that SOM cluster 13-dimensional data more accurately than conventional statistical classifiers. The component planes constructed by SOM are more successful than statistical tests in determining the distinctive parameters.

Keywords: Neural networks; Kohonen self-organizing maps; Volcanic ash; Nemrut

1. Introduction

Fragmentation of magma may occur due to exsolution of gas phases as a result of decompression (magmatic) or by an interaction between external water and magma (phreatomagmatic) (Cashman et al., 2000). These two fragmentation processes produce ash particles with end-members displaying characteristic morphology and surface features (Wohletz, 1983; Heiken and Wohletz, 1985). Scanning electron microscopy (SEM) provides a suitable method for classifying volcanic ash based on surface morphology and texture (Wohletz and Krinsley, 1982; Büttner et al., 1999; Ersoy et al., 2006), with the most extensive SEM studies of pyroclast shapes having been presented by Heiken (1972, 1974), Wohletz (1983) and Heiken and Wohletz (1985). A study of explosive fragmentation dynamics by examination of the morphological features of natural and experimental ash particles has demonstrated the significance of morphological microfeatures on ash grains (Wohletz, 1983). However, the lack of analogue particles produced by scaled experiments at that time allowed only a qualitative assessment. Although a study of natural pyroclastic sequences combined with scaled laboratory experiments identified the different fragmentation

*Corresponding author. Department of Geological Engineering, Hacettepe University, 06532, Beytepe-Ankara, Turkey. Fax: +90 312 2992034.
E-mail address: oersoy@hacettepe.edu.tr (O. Ersoy).

0098-3004/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.cageo.2006.10.008

mechanisms related to the water/magma mass ratios during their interaction (Büttner et al., 1999), the surface features of the ash were not quantified. Classification of volcanic ash surfaces is still limited to descriptive terms such as stepped, planar, crack patterns, and hydration skin. The qualitative data of volcanic ash need to be expressed in quantitative ways, supported by supplementary methods such as statistical analysis and artificial intelligence.

Ersoy et al. (2006) introduced a new method of ash particle characterization based on a quadtree decomposition approach and gradient analysis of SEM micrographs of volcanic ash particles from Mount Nemrut, eastern Turkey. They calculated 13 different surface parameters for volcanic ash particles to assess the fragmentation mechanisms operating during eruption. The calculated quadtree variables, such as the number of blocks (nQT), mean block size (mQT), and standard deviation of block sizes (sQT), as well as surface descriptors derived from gradient analysis, are suitable for quantifying structural changes of the ash surface associated with variable explosion conditions. Ersoy et al. (2006) presented these parameters in volcanology as key parameters for identifying different eruption types. In the present paper, we use the data of the 13 surface parameters with 10 replicates from each of three different samples analyzed by Ersoy et al. (2006).

Visualizing multi-dimensional data are critical to understanding complex relationships in natural systems; consequently, earth scientists are faced with increasingly large amounts of data. As the dimensionality of the data increases, the complexity of visualizing relationships also increases. To understand natural systems, data from diverse sources must be integrated in a manner that is comprehensible to human interpretation (Penn, 2005). Several approaches have been presented to convey high-dimensional data (e.g., Tukey, 1977; Levoy, 1988; Wegman and Solka, 2002). Methods, which preserve certain properties of the structure of the data set as faithfully as possible, are called 'projection methods' (Ripley, 1996).

A projection method that categorizes or groups similar data items is termed 'clustering'. Clustering methods (Anderberg, 1973; Hartigan, 1975; Jain and Dubes, 1988; Jardine and Sibson, 1971; Sneath and Sokal, 1973; Tryon and Bailey, 1973) can be divided into two basic types: hierarchical and partitional clustering (Kaski, 1997).

Ersoy et al. (2006) performed the data reduction method (linear projection method) factor analysis (principal component analysis) on their data to visualize different types volcanic ash. The utility of factor analysis is plotting the singular samples on a factor diagram as factor scores ([Fr]). In this case, two factors explained 98% of the variance. The authors used the three most suitable parameters for differentiating between samples, with resulting p-values well below 0.05 at a 95% confidence interval (CI) in analysis of variance (ANOVA).

For the present paper, we used the data of Ersoy et al. (2006) to cluster the different types of volcanic ash particles via self-organizing maps (SOM), which classify data more accurately than conventional statistical classifiers (Moline and Bahr, 1995; Dolmatova et al., 1997; Kocjancic and Zupan, 1997). Although the data-mining tools described above are divided into two categories, projection and clustering methods, SOM are special cases where data are simultaneously reduced and projected onto a lower-dimension.

2. Artificial neural networks

Artificial neural networks (ANN) are relatively crude electronic models based on the neural structure of the brain. ANN attempt to mirror the brain functions in a computerized way by resorting to the learning mechanism as the basis of human behavior (Lipman, 1987; Anderson and McNeil, 1992; Hagan et al., 1996; Kartalopoulos, 1995).

ANN can generally be defined as a structure composed of a number of interconnected units (Skapura, 1996). Each unit has an input/output (I/O) characteristic and implements a local computation or function. The output of each unit is determined by its I/O characteristic, its interconnection to other units and (possibly) external inputs, and its internal function. The network usually develops an overall functionality through one or more forms of training. The fundamental unit or building block of the ANN is the artificial neuron (termed neuron from here on). The neuron has a set of inputs (X_i) weighted before reaching the main body of the processing element. In addition, it has a bias term, a threshold value that has to be reached or exceeded for the neuron to produce a signal, a non-linearity function (f_i) that acts on the produced signal (R_i), and an output (O_i) (Lipman, 1987; Haykin, 1994; Hagan et al., 1996; Kartalopoulos, 1995). The basic model of a neuron is illustrated in Fig. 1.

ANN are classified into two major types on the basis of learning modes: supervised and unsupervised. In supervised mode, the actual output of a neural network is compared to the desired output. The network then adjusts weights, which are usually initially randomly set, so that the next iteration or cycle produces a closer match between the desired and actual output. The learning method seeks to minimize the current errors of all processing elements. This global error reduction is created over time by continuously modifying the input weights until acceptable network accuracy is achieved. With supervised learning, someone should train the network before it becomes useful. Training consists of presenting input and output data to the network. This training is considered complete when the neural network reaches a user-defined performance level. This level signifies that the network has achieved the desired statistical accuracy as it produces the required outputs for a given sequence of inputs. When no further learning is necessary, the weights are typically frozen for the application. Some network types allow continual training while in operation, although at a much slower rate. This helps the network to adapt to gradually changing conditions.

Unsupervised learning holds great promise for the future, with the possibility that computers could some day learn by themselves. These networks use no external influences to adjust their weights; instead, they monitor their performance internally. These networks search for regularities or trends in the input signals and make adaptations according to the function of the network. Even without being told whether it is right or wrong, the network still requires some information concerning how to organize itself. This information is built into the network topology and learning rules. An unsupervised learning algorithm might emphasize cooperation among clusters of processing elements; in such a scheme, the clusters work together. If an external input activated any node in the cluster, the cluster's activity as a whole is increased. Likewise, a decrease in the external input to nodes in the cluster has an inhibitory effect on the entire cluster. Competition between processing elements also forms a basis for learning. Training of competitive clusters amplifies the responses of specific groups to specific stimuli. As such, it associates these groups with each other and with a specific appropriate response. Generally, when competition for learning is in effect, only the weights belonging to the winning processing element are updated.

Because all ANN are based on the concept of neurons, connections, and transfer functions, there is a similarity between the different structures or architectures of neural networks. The majority of the variations stem from the various learning rules and the way in which these rules modify a network's typical topology (Aleksander and Morton, 1990; Anderson and McNeil, 1992; Fausett, 1994; Schalkoff, 1997). Because of the advantages and accuracy of Kohonen SOM for clustering problems of geological materials (Chang et al., 2002), in the present study we used SOM for our clustering problem.

2.1. Kohonen SOM

Kohonen SOM are unsupervised ANN developed by Kohonen (1982). They represent the result of a vector quantization algorithm that places a number of reference or codebook vectors into a high-dimensional input data space to approximate its data sets in an ordered fashion. When local-order relations are defined between the reference vectors, the relative values of the latter are made to depend on each other as if their neighboring values lie along an 'elastic surface'. By means of the self-organizing algorithm, this 'surface' becomes defined as a kind of nonlinear regression of the reference vectors through the data points. Mapping from a high-dimensional data space R^n onto, for example, a two-dimensional lattice of points, is thereby also defined.

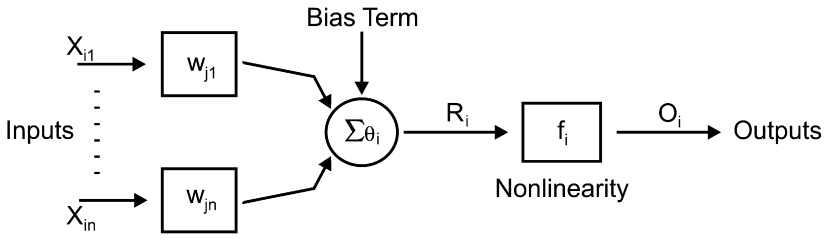

Fig. 1. Basic model of artificial neuron.

Such mapping can be used to effectively visualize metric ordering relations of input samples. In practice, mapping is obtained as an asymptotic state in a learning process (Kohonen et al., 1996a).

3. Source of data

The data include calculated quadtree variables such as the number of blocks (nQT), the mean block size (mQT), the standard deviation of block sizes (sQT), and the surface descriptors derived from gradient analysis of SEM micrographs of volcanic ash particles. Ten replicates were acquired from three samples: N1, N2, and N3. These three samples are volcanic ash particles derived from different fragmentation mechanisms (for details, see Ersoy et al., 2006). ANOVA was performed on the data to test hypotheses concerning differences between means. ANOVA is used to test the significance of differences among several means without increasing the Type I error rate. The successful parameters for discrimination were determined previously by Ersoy et al. (2006); however, here all parameters are considered to document the limits of SOM in visualizing and clustering high-dimensional data. The number of variables is 13, which is also the codebook vector dimension. All values in the input data are numeric and stored in ASCII form.

4. Architecture and learning algorithm of SOM

4.1. Map construction

There are many versions of the SOM. Here SOM define mapping from the input data space R^n onto a regular two-dimensional array of nodes, which can be visualized as a sheet-like neural-network array. We preferred a hexagonal lattice. The x-coordinates of the map (column numbers) range from 0 to $n-1$ where n is the x-dimension of the map, while the y-coordinates (row numbers) vary from 0 to $m-1$, where m is the y-dimension of the map (Fig. 2). Our map dimension is 30×20 in the x- and y-directions, respectively. We used the Self-Organizing Map Program Package (SOM_PAK) (Kohonen et al., 1996a) for the correct application of the SOM algorithm.

A parametric reference vector $m_i \in R^n$ is associated with every node i. An input vector $x \in R^n$ is compared with the m_i, and the best match is defined as 'response'; the input is thus mapped onto this location. SOM is a nonlinear projection of the

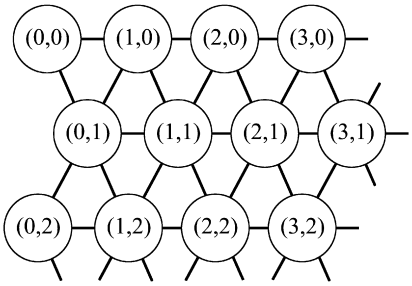

Fig. 2. Locations of units in hexagonal topological structure. The distance between two units in map is computed as a Euclidean distance in (two-dimensional) map topology. Reproduced from Kohonen et al. (1996a).

probability density function of the high-dimensional input data onto the two-dimensional display (Kohonen, 1989). The smallest of the Euclidean distances $\|x-m_i\|$ usually define the best-matching node (c):

$$\|x-m_c\| = \min_i\{\|x-m_i\|\} \quad \text{or,}$$
$$c = \underset{i}{\operatorname{argmin}}\{\|x-m_i\|\}. \tag{1}$$

Thus, x is mapped on the c relative to the parameter values m_i. During learning, the nodes that are topographically close in the array, up to a certain distance, activate each other to learn from the same input. The useful values of m_i can be found as convergence limits of the following learning process, whereby the initial values of the $m_i(0)$ are random in our map:

$$m_i(t+1) = m_i(t) + h_{ci}(t)[x(t) - m_i(t)], \tag{2}$$

where t is an integer, the discrete-time coordinate, and $h_{ci}(t)$ is the so-called neighborhood kernel, which is a function defined over the lattice points. Usually, $h_{ci}(t) = h(\|r_c - r_i\|, t)$, where $r_c \in R^2$ and $r_i \in R^2$are the radius vectors of the nodes c and i, respectively, within the array. With increasing $\|r_c - r_i\|, h_{ci} \to 0$. The average width and form of h_{ci} define the 'stiffness' of the 'elastic surface' to be fitted to the data points. For the definition of h_{ci}, we used a neighborhood set of array points around node c. When we denote this set as N_c, $h_{ci} = \alpha(t)$ if i N_c and $h_{ci} = 0$ if $i \notin N_c$, where $\alpha(t)$ is some monotonically decreasing function of time ($0 < \alpha(t) > 1$). This type of kernel is nicknamed 'bubble' because it relates to certain activity 'bubbles' in laterally connected neural networks (Kohonen, 1989; Kohonen et al., 1996a).

4.2. Training of the map

We trained the map in two phases. In the first phase, we ordered the reference vectors of the map units. It was a rough training for ordering with a large neighborhood radius. The neighborhood radius was five for this phase, and it decreased to one during training. Therefore, in the end only the nearest neighbors were trained. The learning rate for the first phase was 0.07, decreasing to zero, while the training time was 200 000. The learning rate, $\alpha(t)$ is some monotonically decreasing function of time as described above. The training time is the number of steps in training. During the second phase, the reference vectors were fine-tuned and for each unit they converged to their 'correct' values. For the second phase, the training time was longer and the learning rate and neighborhood radius smaller, as described in Kohonen et al. (1996a). Values used for training time, learning rate, and neighborhood radius in the last phase were 400 000, 0.02, and 2, respectively. Therefore, in the beginning the units up to a distance of two were covered. In our map, the training time of the second phase was 2 times longer than in the first phase.

4.3. Quality of learning

Strongly contrasting learning processes can be defined by starting with different initial values $m_i(0)$ and applying different sequences of the training vectors $x(t)$ and different learning parameters. It is clear that an optimal map must exist for certain input data (Kohonen et al., 1996a), but the best map must have a minimum quantization error. The average quantization error (mean of $\|x - m_c\|$) is a useful performance index because it defines the map that best fits the data. In our map, 500 random initializations of the $m_i(0)$ were undertaken, and the map with the minimum quantization error (0.003956) was selected.

5. Results and discussion

We generated a list of coordinates that correspond to the best-matching unit in the map for each data sample. The SOM successfully discriminated the N1 sample from others (Fig. 3). There are two ash particles from each sample, N2 and N3, which fall into the other sample's fields. These ash particles appear to be negatively affecting the success of the clustering. 'A' and 'B' in particular are almost

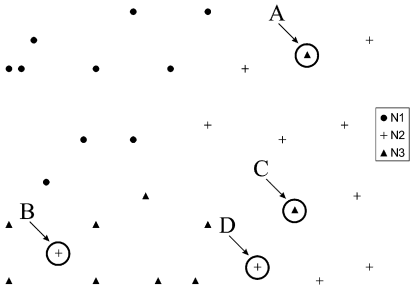

Fig. 3. Best-matching units in map for each data sample. The diagram is constructed from the parameters sGL, nQT, and FF, which were determined as most successful distinctive parameters from statistical analysis (e.g., ANOVA). A, B, C, and D correspond to ash particles with similar surface textures to those of other groups.

members of other groups. The positioning of 'C' and 'D' may reflect the effect of transition between groups, and consequently fragmentation mechanisms; however, this scattering is not the fault of clustering. These ash particles have similar surface textures to other samples. In addition, they are in other sample fields on the discrimination diagram (Fig. 4) of Ersoy et al. (2006).

We also constructed the component planes of our map. These planes visualize the values of the components using gray levels. Each component plane shows the values of one variable in each map unit. From component planes, we can observe the discriminate performance of each variable, namely parameters. Planes of the gray-level standard deviation (sGL), the number of quadtree blocks (nQT), and the mean block size (mQT) appear to be most suitable for visualizing the discrimination (Fig. 5). The gray-level standard deviation (sGL) and the number of quadtree blocks (nQT) were also the most successful parameters for distinguishing fragmentation process in Ersoy et al. (2006), however, the form factor (FF) is not as suitable for discrimination (Fig. 6). We propose the mean block size (mQT) instead of the FF as a parameter for distinguishing fragmentation process by keeping with the component planes of ANN.

We construct the discrimination diagram from factor analysis of three variables: sGL, nQT, and mQT (Fig. 7). The factors are the diagram axes, and the single analyzed surfaces are data points on the diagram. In this case, two factors explain 97% of the variance. We achieve a better-organized diagram by sticking to component planes. Only the ash particle labeled 'A' plotted in the field of a different group.

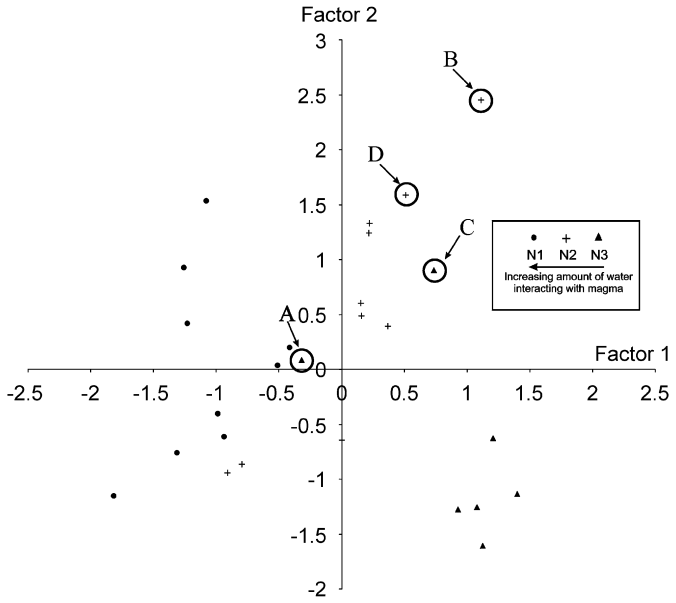

Fig. 4. Discrimination diagram from Ersoy et al. (2006). The diagram is constructed from parameters sGL, nQT, and FF, which were determined as most successful distinctive parameters from statistical analysis (e.g., ANOVA). A, B, C, and D correspond to particles shown in Fig. 3. The inset figure details sample numbers and different fragmentation mechanisms related to amount of water that interacted with magma.

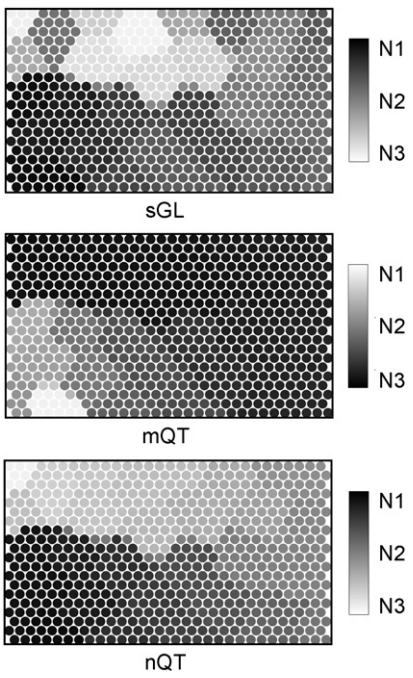

Fig. 5. Component planes of gray-level standard deviation (sGL), number of quadtree blocks (nQT), and mean block size (mQT). The mQT parameter appears as a distinctive parameter instead of form factor parameter, as determined from component planes of self-organizing maps.

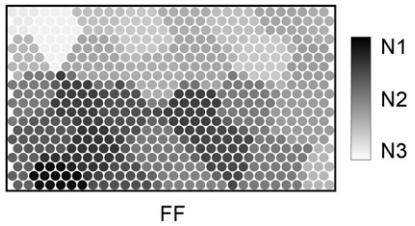

Fig. 6. Component plane of form factor (FF). Form factor parameter is unsatisfactory as a distinctive parameter.

We constructed a new map with the new three-dimensional data, using sGL, nQT, and mQT parameters as variables; the size of the new map was 20×10. We used the settings of the first map for training. After 500 trials, we used the map with the smallest quantization error (0.004572) for visualization. The new distinctive parameters determined by the component planes optimized the new discrimination (Fig. 8).

6. Conclusions

In this study, we present the visualization and clustering capabilities of self-organizing maps (SOM) for analyzing high-dimensional data. SOM

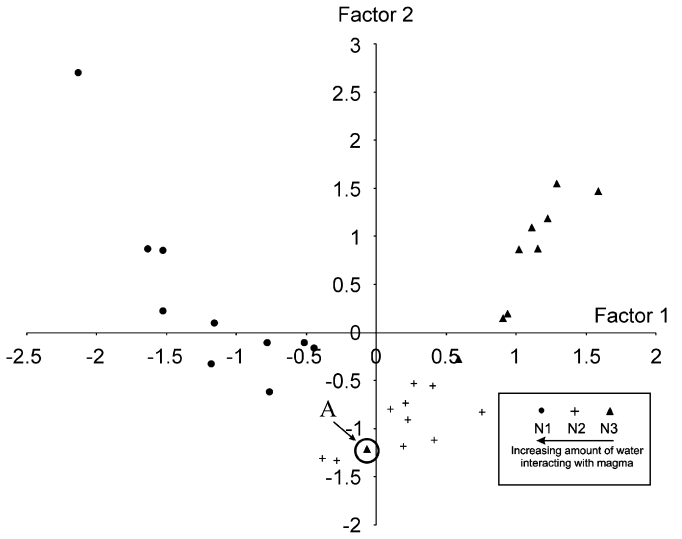

Fig. 7. Discrimination diagram using new parameters derived from component planes of SOM. Scattering of particles is tidier and well organized than that in Fig. 4. Ash particle labeled 'A' is only point that plotted in another group's field.

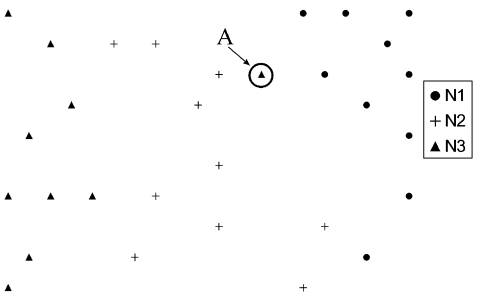

Fig. 8. Best-matching units in new map for each data set. Diagram is constructed from parameters sGL, nQT, and mQT, which were determined as most successful distinctive parameters from component planes of SOM. Only ash particle labeled 'A' is positioned in wrong field, due to its similarity in surface texture with this group. Note optimization as compared with that in Fig. 3.

were used because they implement the orderly mapping of a high-dimensional distribution onto a regular low-dimensional grid and are thereby able to convert complex nonlinear statistical relationships between high-dimensional data items into simple geometric relationships on a low-dimensional display.

We used surface texture parameters of volcanic ash that arose from different fragmentation mechanisms as input data. The component planes constructed by SOM were more successful than statistical tests in determining parameters for distinguishing the different fragmentation mechanisms. Component planes helped to determine the discriminate performance of each variable. Based on these planes, we modified the parameters to attain the absolute discrimination.

The SOM solves difficult high-dimensional and nonlinear problems such as feature extraction and classification of images and acoustic patterns, adaptive control of robots, and the equalization, demodulation, and error-tolerant transmission of signals in telecommunications. A new area of application is the organization of very large document collections (Kohonen, 2001).

In many applications, sensor failures, recording errors, and resource limitations may prevent data collection from completing each input vector; however, such incomplete training examples still contain useful information. For example, partial data can still be used to determine the distribution statistics of the available vector components. For incomplete input data vectors, the distance calculations and reference vector modification steps using the available data components would be undertaken by a supervised learning algorithm; for example, the learning vector quantization (LVQ), which is related to the SOM (Kohonen et al., 1996b). Finally, we note that the SOM are one of the most realistic models of biological brain function (Kohonen, 2001).

Acknowledgments

This work benefited from a research grant from the Hacettepe University Research Foundation

(Project No. 01 01 602 020). The French Government supported the research activities of the first author in France. Pierre Boivin and anonymous reviewers are thanked for constructive suggestions that helped to improve the paper. International Science Editing improved the English of paper.

References

Aleksander, I., Morton, H., 1990. An Introduction to Neural Computing, Second ed. Chapman & Hall, New York, NY, 240pp.

Anderberg, M.R., 1973. Cluster Analysis for Applications. Academic Press, Inc., New York, NY, 359pp.

Anderson, D., McNeil, G., 1992. Artificial Neural Network Technology. Data & Analysis Center for Software, Rome, 83pp.

Büttner, R., Dellino, P., Zimanowski, B., 1999. Identifying modes of magma/water interaction from the surface features of ash particles. Nature 401, 688–690.

Cashman, K.V., Sturtevant, B., Papale, P., Navon, O., 2000. Magmatic fragmentation. In: Sigurdsson, H., Houghton, B.F., McNutt, S.R., Rymer, H., Stix, J. (Eds.), Encyclopedia of Volcanoes. Academic Press, San Diego, pp. 412–430.

Chang, H-C., Kopaska-Merkel, D.C., Chen, H-C., 2002. Identification of lithofacies using Kohonen self-organizing maps. Computers & Geosciences 28, 223–229.

Dolmatova, L., Ruckebusch, C., Dupuy, N., Huvenne, J.P., Legrand, P., 1997. Quantitative analysis of paper coatings using artificial neural networks. Chemometrics and Intelligent Laboratory Systems 36 (2), 125–140.

Ersoy, O., Chinga, G., Aydar, E., Gourgaud, A., Cubukcu, H.E., Ulusoy, I., 2006. Texture discrimination of volcanic ashes from different fragmentation mechanisms: a case study, Mount Nemrut stratovolcano, eastern Turkey. Computers and Geosciences 32, 936–946.

Fausett, L.V., 1994. Fundamentals of Neural Networks, US ed. Prentice Hall, NJ, 461pp.

Hagan, M.T., Demuth, H.B., Beale, M., 1996. Neural Network Design. PWS Publishing Company, Boston, MA, 730pp.

Hartigan, J.A., 1975. Clustering Algorithms. Wiley, New York, NY, 351pp.

Haykin, S., 1994. Neural Networks: A Compherensive Foundation. New York, Macmillan.

Heiken, G., 1972. Morphology and petrography of volcanic ashes. Geological Society of America Bulletin 83, 1961–1988.

Heiken, G., 1974. An atlas of volcanic ash. Smithsonian Contributions to the Earth Sciences 12, 1–101.

Heiken, G., Wohletz, K.H., 1985. Volcanic Ash. University of California Press, Berkeley, California, 245pp.

Jain, A.K., Dubes, R.C., 1988. Algorithms for Clustering Data. Prentice-Hall, Englewood Cliffs, NJ, 320pp.

Jardine, N., Sibson, R., 1971. Mathematical Taxonomy. Wiley, London, 304pp.

Kartalopoulos, S.V., 1995. Understanding neural networks and fuzzy logic: basic concepts and applications. Wiley, IEEE Press, New York, 232pp.

Kaski, S., 1997. Data exploration using self-organizing maps. Ph.D. Dissertation, Acta Polytechnica Scandinavica, Mathematics, Computing and Management in Engineering Series No. 82, Espoo. Published by the Finnish Academy of Technology, 57pp.

Kocjancic, R., Zupan, J., 1997. Application of a feed-forward artificial neural network as a mapping device. Journal of Chemical Information and Computer Sciences 37 (6), 985–989.

Kohonen, T., 1982. Self-organized formation of topologically correct feature maps. Biological Cybernetics 43, 59–69.

Kohonen, T., 1989. Self-Organization and Associative Memory, third ed. Springer, New York Inc, NY, 312pp.

Kohonen, T., 2001. Self-organizing maps, third ed. Springer, Berlin, 501pp.

Kohonen, T., Hynninen, J., Kangas, J., Laaksonen, J., 1996a. SOM_PAK: the self-organizing map program package. Technical Report A31, Helsinki University of Technology, Laboratory of Computer and Information Science, FIN-02150 Espoo, Finland, 27pp.

Kohonen, T., Hynninen, J., Kangas, J., Laaksonen, J., Torkkola, K., 1996b. LVQ_PAK: the learning vector quantization program package. Technical Report A30, Helsinki University of Technology, Laboratory of Computer and Information Science, FIN-02150 Espoo, Finland, 30pp.

Levoy, M., 1988. Display of surfaces from volume data. IEEE Computer Graphics & Applications 8 (5), 29–37.

Lipman, R.P., 1987. An Introduction to Computing with Neural Nets. IEEE ASSP Magazine, April, 4–22.

Moline, G.R., Bahr, J.M., 1995. Estimating spatial distributions of heterogeneous subsurface characteristics by regionalized classification of electrofacies. Mathematical Geology 27, 3–22.

Penn, B.S., 2005. Using self-organizing maps to visualize high-dimensional data. Computers and Geosciences 31, 531–544.

Ripley, B.D., 1996. Pattern Recognition and Neural Networks. Cambridge University Press, Cambridge, Great Britain, 415pp.

Schalkoff, R.J., 1997. Artificial Neural Networks. McGraw-Hill, New York, 448pp.

Skapura, D., 1996. Building Neural Networks. ACM Press, Addison-Wesley, NY, 282pp.

Sneath, P.H.A., Sokal, R.R., 1973. Numerical Taxonomy. Freeman, San Francisco, CA, 573pp.

Tryon, R.C., Bailey, D.E., 1973. Cluster Analysis. McGraw-Hill, New York, NY, 347pp.

Tukey, J.W., 1977. Exploratory Data Analysis. Addison-Wesley, Reading, MA, 688pp.

Wegman, E.J., Solka, J.L., 2002. On some mathematics for visualizing high dimensional data. The Indian Journal of Statistics 64 (A-2), 429–452.

Wohletz, K.H., 1983. Mechanisms of hydrovolcanic pyroclast formation: size, scanning electron microscopy, and experimental studies. In: Sheridan, M.F., Barberi, F. (Eds.), Explosive Volcanism, Journal of Volcanology and Geothermal Research, vol. 17, pp. 31–63.

Wohletz, K.H., Krinsley, D., 1982. Scanning electron microscopy of basaltic hydromagmatic ash. Los Alamos National Laboratory Report, LA-UR 82-1433, 43pp.