2008

# Fault Classification in Hermetic Compressors Using Self-Organizing Map

Atilla Kaya
*Arcelik A.S*

Emin Germen
*Arcelik A.S*

Ümit Ünlü
*Arcelik A.S*

Suha Toprak
*Arcelik A.S*

# FAULT CLASSIFICATION IN HERMETIC COMPRESSORS USING SELF-ORGANIZING MAP

Atilla KAYA[1], Emin GERMEN[2], Ümit ÜNLÜ[1], Suha TOPRAK[3]

[1] ARCELIK A.S. Eskişehir Compressor Plant, Department of Product Development
Eskişehir-TURKEY
Tel:+90-222-2134659 & 90-222-2134676, Fax: +90-222-2134344
e-mail: atilla.kaya@arcelik.com , umit.unlu@arcelik.com
[2] Anadolu University Electrical & Electronics Engineering Department
Eskişehir-TURKEY
Tel:+90-222-3350580 Fax:+90-222-3357942
e-mail: egermen@anadolu.edu.tr
[3] TEI-TUSAS Engine Industries Inc.
Eskişehir-TURKEY
Tel:+90-222-2112467, Fax: +90-222-2112101
e-mail: Suha.Toprak@tei.com.tr

## ABSTRACT

The Self Organizing Map (SOM) is used to classify the possible five very common faults of hermetic compressors that can occur either after the production line or during the R/D phase of new types. The main concern is discriminating the faulty compressors from the healthy one. In order to identify these faults, the SOM network has been trained with the feature vectors that are obtained from sound pressure data, vibration data, pressure pulsations that are gathered from numerous faulty and healthy compressors. After the training phase, the Learning Vector Quantization (LVQ) algorithm is used to scrutinize the classification borders. The results obtained after the classification denotes that SOM is very useful tool in order to discriminate the healthy compressors from the faulty ones Also it is worth to mention that, the classification of the different type of the faults has been achieved.

## 1. INTRODUCTION

Reciprocating compressors are the most important noise and vibration sources in household refrigeration systems. Not only the overall level but also the quality of this noise and vibration is important. Due to faults during production processes or resulting from the origin of the used material, sound and vibration characteristics of the products may vary. These unknown faults sometimes create noise problems that are observed in periodic sampling tests. On the other hand, because of the mechanical structure, the sources of these faults cannot be determined unless the compressor is disassembled. The goal of the present study is to identify the common faults by means of acoustical measurements without disassembling the compressors.

Both for design phase and serial production process, the identification of the faults has vital importance. Since it is not possible to reuse the compressor that is disassembled due to some unknown fault in serial production, detecting the source of the problem may prevent loss of time to investigate the problem and loss of money by taking some precautions. Also for the prototypes of new designs, the neural network analysis may save plenty of time and can supply important information that can shape the design.

The neural network tools are very convenient for fault analyses in different disciplines. (Germen, *et. al.* 2005) By neural network tools, it is aimed to recognize common faults by investigating the compressors in a controlled manner. In order to gather the data to be used in neural network analysis, some common serious faults have been implemented one by one on the compressors that are chosen such that they do not have any of these faults at the beginning. After performing a series of detailed acoustic and vibration tests on these base compressors, controlled faults related to muffler, shock loop tube, motor and springs are implemented on them one by one. At each time the fault is implemented, the series of tests are conducted and these tests are repeated at each step. 17 compressors were selected for the tests. All the experiments were done with these compressors. Six series of experiments (one with the normal and one with each of the five faulty conditions) for each of the 17 compressors, those are 102 experiments in total, were designed and conducted.

All of the measurements are carried out in a semi-anechoic chamber up to ISO 3745. The sound power level of each of these compressors is measured by 10 microphones as described in ISO 3745 standard after the compressors reach steady state conditions.1/3 octave sound spectra for all of these 10 microphones is measured. There exist two different approaches to analyze the data. First one uses a weighted average of all 10 microphones' measurements. Second approach uses all data from each microphone separately, which means larger input files and higher computation time, when compared to the first approach but has the ability to investigate the fault regions in a more detailed way. In this study, since the faults could not reach high separation ratios using the first method, the second method is used. According to the results, each fault can satisfactorily be identified using the second approach when sufficient training is done.

In section 2, a very brief introduction of SOM (Self Organizing Map) has been introduced which is used to find the possible clusters for different data sets. Also the proposed map dimensions and the characteristics of the feature vectors in order to train the map will be introduced in this section. In section 3 the results obtained will be discussed and the section 4 concludes the paper.

## 2. SOM (SELF ORGANIZING MAP) AND LVQ (LEARNING VECTOR QUANTIZER)

Kohonen's SOM (Self Organizing Map) is an impressive tool to visualize the possible classes occurs in high dimensional data sets. The theory of SOM has inspired from the structural organization of neurons in cerebral cortex in human neural system. (Kohonen, 1995) It is observed that some specific areas of brain tissue are organized according to the types of the input signals in adaptive and automatic nature. Similarly, SOM shows same kind of organization in unsupervised manner. SOM in general provides a projection of high dimensional data set which has a character $\Lambda = \Re^n$ into *m* many codebook vectors of size *n* to two dimensional domains. It is also worth to mention that, the organization of codebook vectors with connection in two-dimensional planar surfaces, keeps the relational information between the input data that provides us clustering information.

In SOM, the learning period is described as :

$$M_i(k) = M_i(k-1) + \alpha(k).\beta(i,c,k)(\Lambda(k) - M_i(k-1)) \tag{1}$$

where $\alpha(k)$ is the learning rate parameter which is changed during the adaptation phase and $\beta(i,c,k)$ is the neighborhood function around *c* where *c* is the Best Matching Unit index which can be found during training as:

$$c = \arg min_i \left\| \Lambda(k) - M_i(k) \right\| \tag{2}$$

The interpretation of above equation requires explanation of parametric learning rate and neighboring function. Learning rate has decreasing characteristic during the learning period that effects the changing positions of the neurons in lattice. For the most of the applications, the general approach is fast at the beginning and slow at the end of the learning phase. The neighborhood function describes the impact area around BMU that describes how the neighboring neurons will be drawn near to BMU. The BMU describes the winning neuron in the training phase where index *c* is determined by equation 2.

In this work the feature vectors $\Lambda = \Re^n$ where $n$ = 250 are used to train the 10x10 SOM neurons which has connected in Hex-lattice manner. The 250 components of feature vectors are obtained from 10 different microphones. Each microphone is used to record the noise spectrum between 40 Hz- 10KHz. divided to 25 distinct regions. Here it has been observed that various frequency components in this range don't have contributions in discrimination of clusters like the others. In order to inspect which frequency components do have effects on determination of the classification borders, plenty many experiments have been carried out. The effects of different frequency bands on discrimination of the classification borders will be the subject of another paper.

After obtaining the possible codebook vectors using SOM training algorithm that is an unsupervised technique, the possible classification regions should have to be obtained in supervised manner. In literature Learning Vector

Quantization 3 (LVQ3) algorithm (Kohonen, 1990) is the one of the most suitable technique in order to delineate and adjust the crossing borders of the possible classification regions.

The LVQ-3 algorithm can be explained as:

$$\begin{aligned} M_i(k+1) &= M_i(k) - \mu(k)\left(\Lambda(k) - M_i(k)\right) \\ M_j(k+1) &= M_j(k) + \mu(k)\left(\Lambda(k) - M_j(k)\right) \end{aligned} \tag{3}$$

Where $M_i$ and $Mj$ are the two closest codebook vectors to $\Lambda(k)$, whereby $\Lambda(k)$ and $Mj$ belongs to the same class, while $\Lambda(k)$ and $Mi$ belongs to different classes respectively; furthermore $\Lambda(k)$ must fall zone of a *window* defined as;

$$\min\left(\frac{d_1}{d_2} \quad \frac{d_2}{d_1}\right) > s \quad \text{where } s = \frac{\text{1-window}}{\text{1+window}} \tag{4}$$

where $d_1$ and $d_2$ are the distance between codebook vectors $M_i$ - $\Lambda(k)$ , and $M_j$ - $\Lambda(k)$. Also it is necessary to have:

$$\begin{aligned} M_i(k+1) &= M_i(k) + \varepsilon(k)\mu(k)\left(\Lambda(k) - M_i(k)\right) \\ M_j(k+1) &= M_j(k) + \varepsilon(k)\mu(k)\left(\Lambda(k) - M_j(k)\right) \end{aligned} \tag{5}$$

where $M_i$ and $Mj$ are the two closest codebook vectors to $\Lambda(k)$, whereby $\Lambda(k)$ and $Mj$ and $Mi$ belong to same classes. The $\mu(k)$ and $\varepsilon(k)$ parameters are learning rates in the algorithm.

## 3. EXPERIMENTS AND THE RESULTS

The primary concern of using a Neural Network Tool (especially SOM), is discriminating the healthy compressors from the defective ones without regarding the types of faults they have. However during the separation phase of the compressors it has been observed that SOM not only gives clues for differentiating the healthy ones from the others but also signifies the types of faults. As it is noted in the introduction, in the experiments the data have been collected from 10 different microphones. During the data collection phase the center frequencies in order to build up the feature vectors are given in the table 1.

Table 1. Center frequencies obtained in experiments

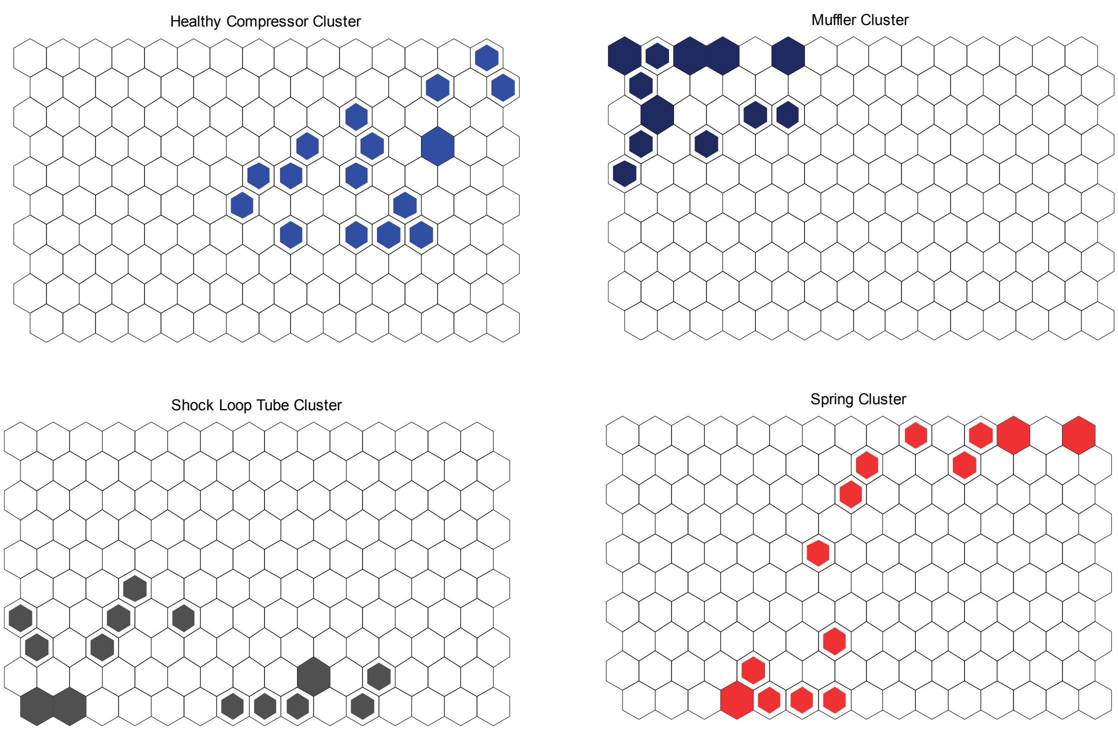
| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 40 Hz. | 50 Hz. | 63 Hz. | 80 Hz. | 100 Hz. | 125 Hz. | 160 Hz. | 200 Hz. | 250 Hz. | 315 Hz. | 400 Hz. | 500 Hz. | 600 Hz. |
| 630 Hz. | 800 Hz. | 1.0 KHz. | 1.6 KHz. | 2.0 KHz. | 2.5 KHz. | 3.15 KHz. | 4.0 KHz. | 5.0 KHz. | 6.3 KHz. | 8.0 KHz. | 10.0 KHz. | |

For 17 different compressors with 4 major fault types and a healthy data, the 10x15 hexagonally connected lattice SOM Map has been trained with the data obtained from 10 different microphones with the 25 different center frequencies values. After the training phase of SOM, the resultant maps are investigated with the faulty compressors data and the healthy ones. At the end of the experiments, although the results were quite reasonable from the point of view of discrimination of types of faults, however they were not so impressive. In this data range it has been observed that some center frequency data has no importance to differentiate the clusters as stated in Section 2. So for each center frequency range data, the same kind of SOM Map has been trained and the qualities of formation of clusters have been measured by Leave one out (LOO) method. In this technique, 102 different SOM training experiments have been done leaving one of the experimental data out for each center frequency with 10 microphone set. After training the map with 101 experimental data, Learning Vector Quantization (LVQ3) technique is used to denote the classification borders. After formation of the border, the data, which is taken out of the experiment, has been tested to locate its class. After those experiments, the results of the LOO tabulated and the most prominent data which belong 10 center frequencies which have more impacts on formation of classification border as desired have been chosen to form the resultant data vector. In table 2, the chosen center frequencies have been given.

Table 2. Center frequencies used in training the SOM in experiments

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 125 Hz. | 250 Hz. | 315 Hz. | 630 Hz. | 1.25 KHz. | 1.6 KHz. | 2.0 KHz. | 5.0 KHz. | 6.3 KHz. | 8.0 KHz. |

The trained SOM with the cluster structure of each type of faults and the healthy compressor has been shown in Figure 1. By investigating the formation of the clusters for both healthy compressor and the faulty ones, it is easy to deduce that the muffler and the motor faults are the most distinctive types of faults compared to the others. They have localized in very specific regions in the SOM lattice with not interfering with the other clusters. However Shock Loop Tube and Spring types of faults have dispersed characteristics in the resultant map. The scattered formation of the Spring and Shock Loop types of faults denote that those types of faults produce different types of noise characteristics in experiments, and the shape of the spectral formation of noise is more close to the healthy compressor than the other types of faults. However it is worth to mention that SOM is very powerful tool to discriminate each type of fault from the normal compressor. Although the visual inspection gives idea about the quality of the classification of results, a quantitative analysis is also required.



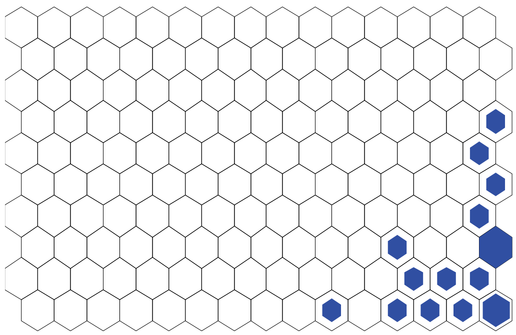



Figure 1. SOM Maps for classification of healthy and faulty compressors

In table 3, the results of LOO method after formation of classification borders using LVQ, has been given. This makes sense from the point of quantative analyses of the method that is proposed.

Table 3. LOO Results

| | Classified as Healthy | Classified as Muffler fault | Classified as Shock Loop fault | Classified as Spring fault | Classified as Motor fault |
|---|---|---|---|---|---|
| Healthy Data | 16 | 0 | 0 | 1 | 0 |
| Muffler Data | 0 | 17 | 0 | 0 | 0 |
| Shock Loop Data | 0 | 0 | 15 | 2 | 0 |
| Spring data | 0 | 0 | 1 | 16 | 0 |
| Motor data | 0 | 0 | 0 | 0 | 17 |

## 4. RESULTS & CONCLUSIONS

The results of both qualitative and quantative analyses after training the noise data for different types of faulty compressors and the healthy one are quite notable in order to use the proposed method in fault classification for compressors. Here the muffler and motor fault types of errors have been discriminated with 100% success rate. It is not difficult to deduce that the Shock loop fault and Spring faults produce some kind of noise spectrum. One of the reasons for this could be the difficulty of producing uniform kinds of faults for each compressor during the experiments. Here the Shock Loop faults have been classified at a rate of 88% and the success rate of clustering the Spring fault is 94%. Similarly the main concern was the discrimination of healthy compressors from the faulty ones, and by scrutinizing the whole experiments with 85 different data, the classification success rate is 98.8% which is quite impressive.

## REFERENCES


Germen E, Ece D. G., Gerek Ö. N., “Self Organizing Map (SOM) approach for classification of power quality events,” LNCS 3696: 403-408, (2005),

Kohonen T, “The Self Organizing Map”, Springer Verlag, 1995,

Kohonen T, “The Self Organizing Map”, Proceedings of IEEE, 78, 9, (1990), 1464-1480

Toprak S., Iftar A., “Fault Diagnosis on Hermetic Compressors Based on Sound Measurements”, 16th IEEE International Conference on Control Applications, Singapore, 1-3 October 2007, 788-793