A new neural-based procedure for gear fault detection

A. Fernández del Rincón, F. Viadero Rueda, R. Sancibrián Herrera & P. García Fernández
Department of Structural and Mechanical Engineering, University of Cantabria, Cantabria, Spain

Abstract

In this work a new gear fault detection procedure based on the time synchronous averaging (TSA) signal is proposed. In the first place an auto-associative neural network (AANN) is used to obtain the relationship among the TSA signals from several channels. During training the AANN is used to map a data vector with itself so in this way the net provides practically the same output that arrives at the input and the non-linear relationship between channels is embedded in net weights and bias. In order to avoid a direct mapping between input and output, one of the hidden layers must have a lower dimension than the input vector. Once the AANN is trained it is used with new data registers as a linear prediction error filter. If the new register contains the same data structure as the training set the prediction error will be lower and the machine is working properly. Otherwise, when the new register differs from the training set, as a consequence of a fault, prediction error will be increased in each channel. At the end the product of the error signal derivative for each channel is calculated using the resulting vector for detection and diagnostic purposes. The possibility of generalizing the net prediction capabilities using a training data set which contains several load cases is also explored. The technique is validated using experimental data taken on a laboratory gear set simulating several faults. An example is presented in order to show the advantages of the application of the proposed procedure comparing it with the most traditional gear processing tools based on TSA as residual and regular signals.

Keywords: gear, neural networks, diagnostics, TSA, condition monitoring, data fusion.
1 Introduction

Gearboxes are one of the most critical components in machinery sets such as rolling mills, helicopters or wind generators. A fault in one gear could be catastrophic for the whole system. To avoid this, several vibration-based condition monitoring techniques have been developed [1], frequency decomposition being the most widely used [2]. Nevertheless, their application in real machines is complex and needs a high level of expertise. Furthermore, more than one signal channel must be supervised in order to be able to detect faults. Normally, each sensor location is defined by the expert to detect faults in a specific component. In this way, experts tend to focus their attention on only one channel rejecting the information provided by the others. Moreover, as the machinery may change its load or velocity it is necessary to define different vibration data sets as a reference for normal condition, setting up the corresponding alarm limits. Therefore, there is a great interest in developing new processing tools [3] that increase fault detection capability, reducing the amount of data used by the expert in order to carry out a good diagnosis. The development of new processing procedures applicable for any level of load or speed is also desirable.

One of the most popular tools for gear condition monitoring is time synchronous averaging (TSA) [4] which isolates gear vibration features from the other components. This procedure carries out a mapping from the time to the angle domain using a reference signal in order to define each shaft turn. In this way several shaft turns are averaged together rejecting any vibration component that is not a multiple integer of the shaft rotation frequency. Once the TSA signal is available several post-processing procedures are proposed in the literature such as amplitude and phase demodulation [3], residual and difference signal [4-5] or more recently the application of autoregressive models [6]. As TSA is not dependent on the working speed, this tool is also very useful in the condition monitoring of machinery working with variable speed. Unfortunately, TSA is not independent of the load level and as a consequence its application in real machinery requires the definition of different alarm levels for each load level [7].

In this work the development of a new processing procedure based on the application of Autoassociative Neural Networks (AANN) is explored. The aim is to combine the vibration signals coming from a certain number of accelerometers located on the gearbox case. Moreover, the proposed technique should provide satisfactory results for any speed or load level improving the detection capability obtained with the application of conventional tools. To achieve these objectives an AANN is trained in order to capture the non-linear relationship between the instantaneous TSA vibration amplitude recorded for each channel. AANNs are feed-forward networks whose inputs and outputs are identical. The aim of this network is to achieve an identity mapping between input and output vectors. Using this kind of network, all the available data are combined and used to obtain the output vector that should be equal to the input one. In this way, the relationship between each component in the input data vector is embedded in the net weight matrix.
In a first stage an AANN is trained with TSA signals obtained when the gearbox is working properly. At the end of the training the relationship between the TSA amplitude for each channel will be embedded in the net weight matrix as well as in its own specific structure (number of neurons in each layer and their activation functions). Once trained, the AANN is used to process the new TSA registers. If the new registers don’t present divergences from the registers used during training, the net prediction error, defined as the difference between the real value and the net output value, will be reduced and the machine working condition will be correct. Otherwise, it is expected that a fault will modify the relationship between channels and therefore the net output. In this case the net prediction error will be higher in those angular positions where the fault is present. In this way, the information from not only one channel but from all available channels is used for fault detection as the net prediction error will depend on the TSA values coming from all the channels. Selecting properly the training set it would be possible to train the AANN in such a way as to span all the working load range. Furthermore, as the data used are TSA signals the proposed procedure should be insensitive to the working speed and load.

2 The proposed technique

Our interest focuses on the definition of a relationship between the TSA amplitudes obtained at each measurement location around the gearbox case. Figure 1 shows a general flow scheme of the proposed procedure. The original vibration records coming from (p) sensors are time averaged taking as a reference a tachometric signal from the interest shaft [1-5].

![Figure 1: Proposed procedure scheme.](image)

At the end of the synchronous averaging the TSA signal has N equi-spaced samples in the angular domain covering a whole turn of the shaft. Afterwards, N input data vectors containing the angular vibration value for each channel are formed and sent to an AANN. The AANN is previously trained using a vector data set corresponding to normal conditions of the gearbox covering several load levels. If the net is well trained and the interrelationship between TSA amplitudes has not changed, a new data vector at the input should provide the same data vector at the output. Finally the Prediction Error (PE) is calculated by subtraction of input and output vectors. This procedure provides a new register...
that will be used for fault detection and diagnostic purposes. If the gearbox
condition has not suffered any change, the PE signal should be reduced and
similar to the one obtained during training. The development of a fault should
increase the PE value. Depending on the fault type, this increment should be
limited to a specific angular position, for localized faults such as tooth pitting, or
extended to the whole gear turn in case of a global fault such as wear.

2.1 Autoassociative neural networks

Neural networks applied to vibration data registers is a topic that has been
approached by several researchers. The most common application is to classify
the machinery condition on the basis of several vibration features extracted from
the original vibration records. This procedure has been applied in the analysis of
roller bearing condition [8], in the classification of shaft loading [9], as well as in
gear fault detection [10-12]. In all of these cases the vibration data coming from
only one channel are used in the classification task which is based on feature data
vectors usually formed with some signal statistical characteristic. Neural
networks have also been used as a pre-processor of raw data in order to facilitate
detection and diagnosis [13]. Among, the possible network structures the most
popular in condition monitoring are feed-forward networks [8-10] but there are
also other possibilities such as radial basis networks [11].

![Figure 2: Autoassociative neural network layout.](image)

Instead of a classification task, in this work a different approach is proposed
since the objective is to apply a neural network to the extraction of the
relationship between the signals coming form more than one sensor. In order to
achieve the proposed objective a particular feed-forward architecture is used that
is called Autoregressive Neural Network. It consists of an input layer, three
hidden layers called Mapping, Bottleneck and Demapping layers and an output
layer, see figure 2. The network must be symmetrical about the dimensions of
each layer as well as the corresponding activation transfer function. The
Bottleneck layer dimension is fundamental for the application of this kind of
networks. It should be the lowest dimension in the net in order to avoid a direct
connection between input and output ports. On the other hand mapping-
demapping layers should have the highest dimension. Their mission is to
facilitate the capability of extraction of the inter-relationship between the input
vector values.
As AANNs are in essence feed-forward neural networks their working principles and training procedures are those applied in this kind of networks. One of the main applications of this kind of networks is related to sensor surveillance [13] exploiting their efficiency in data feature information extraction in the form of process fault of malfunction signatures. This is the capability applied in this work where the interest is to find the relationship between several vibration channels.

3 A case study

The proposed technique has been validated using a set of experimental registers acquired in a laboratory gear test set-up. The available gear test design was based on a real machine working in a steel rolling mill. The experimental set up (see figure 3) has a 1.1 Kw electric drive, two reduction stages (Z1(20)/Z2(60) and Z3(31)/Z4(31)), two output shafts, and two pneumatic brakes located in the output shafts, which act as loads, simulating the lamination rollers resistant torque. The electric drive provides a nominal input velocity of 32 rpm that yields the nominal gearmesh frequencies indicated in Table 1, where the teeth number of each one also appears.

![Gear test set-up.](image)

Table 1: Gear teeth and gearmesh frequencies.

<table>
<thead>
<tr>
<th>Pair</th>
<th>Teeth</th>
<th>Rotation frequency (Hz)</th>
<th>GMF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Z1-Z2</td>
<td>20/60</td>
<td>0.533/0.177</td>
<td>10.666</td>
</tr>
<tr>
<td>Z3-Z4</td>
<td>31/31</td>
<td>0.177/0.177</td>
<td>5.5111</td>
</tr>
</tbody>
</table>

Data acquisition has been carried out using four piezoelectric accelerometers B&K model 4398 located in points called 1 to 4 as appears in figure 3. Position 1 registers axial vibrations while positions 2, 3 and 4 do the same in the radial shaft.
direction. In addition, a one-per-wheel-revolution tachometer signal from shaft 3 was also acquired. Five minutes of vibration were taken at a sampling frequency of 10 kHz recording all signal channels in a Difa SCADAS III unit. Records were post processed with the LMS CADA-X time signature monitor software in order to obtain 1024 TSA samples for each vibration channel. Finally the TSA registers are exported to the MATLAB® environment, where they are used in the training and testing of several AANN structures.

3.1 Test carried out

Firstly, two no fault condition tests called No Fault I (NFI) and No Fault II (NFII) were carried out in order to define the training data set. Shaft number 3 was dismounted and mounted again at the beginning of the test NFII with the aim of studying the effect that maintenance activities have on the net behaviour. The proposed technique exploits the interrelationship among the TSA magnitudes of all channels. Therefore, it should be expected that a maintenance intervention would not have significant effects on the final PE, avoiding in this way a new training process after each maintenance action. Afterwards, a local defect in one tooth on gear number 4 was simulated, drilling several holes into the tooth face located around 65 degrees after the shaft’s tachometric reference. Two registers were recorded called Specific Defect I (PtDI) and Specific Defect II (PtDII), the latter acquired after several operation hours with the objective of smoothing out the initial defect. Finally, the lubrication pump was switched off, recording a new register called Lubrication Fault (LF).

For each case, four tests were carried out applying different load levels at 100%, 75%, 50% and 25% of the available maximum load. Load is controlled by modifying the brake pressure using the power drive consumption as an indicator of the achieved load level. Thus a total of 20 vibration data registers with five channels for each one (four accelerometers and one tachometer) were acquired and processed. It should be highlighted that the available testing facility works at very low speed, around 10 r.p.m. at the output shafts, with a low load level, far away from the gear load limit. Therefore, it is expected that the results obtained during the application of the proposed procedure in this case will be improved in the case of higher speed or load.

3.2 Training data set and network structures

As was mentioned before, two different training sets were built based on tests NFI and NFII that were carried out when the machine is working properly in the absence of faults. The original TSA signal was resampled at 512 data points per turn and high-pass filtered on the 10th order with the aim of eliminating low frequency components. Therefore, for each load level there are four TSA registers, one for each channel, with 512 samples for each one. As one of the objectives is that the trained AANN should be able to work properly for any load level, the TSA registers for each load were concatenated arriving at a total of 2048 data points per channel covering all the range of possible loads. Afterwards, the available data set were expanded by noise addition in order to
improve the network generalization capability. The contaminating noise was obtained from the difference between the TSA signals of test NFII and PtDI, rejecting the TSA samples near the local fault. Those tests were selected because the only difference between them is the presence of the local fault on PtDI and therefore their differences should be related with the random variation from one TSA to the other. Following this procedure, the original data are duplicated by noise addition obtaining two training sets with 4096 input vectors.

The dimension of the net input-output data vector will be four as long as that is the number of available channels. Several network structures were trained two times using the training data set based on NFI and NFII test cases. The networks studied had 10 or 15 neurons in the mapping-demapping layer and 2 or 3 neurons on the bottleneck layer. In all cases, the linear transfer function was applied to input-output layers while the hyperbolic tangent sigmoid transfer function was used on hidden layers.

4 Experimental results and discussion

Once trained, each available register is sent to the AANN and the PE is calculated. As an example, Figure 4 presents the results obtained for NFII, load 100% using the simplest AANN (4-10-2-10-4) trained with data based on NFI. It is clear that the trained AANN is able to follow the original data arriving at a different PE depending on the channel.

Figure 4: Input (Left), Output (Centre) and PE (Right) for NFII-100% using an AANN (10-2-10) training based on NFI.
Figure 5 shows the results obtained with **PtDI** (load 100%) applying a trained AANN (4-10-2-10-4) based on **NFI**. There is a clear anomaly in the PE for each channel near the angular position of the defect, around 65 degrees.

As can be appreciated in figure 4 and 5, the PE shows a periodic structure that was investigated in the frequency domain revealing that it is dominated by harmonics of 31st order, the number of teeth on the fault gear. Therefore, it should be possible to improve the results using the difference signal [3-5] that is obtained by removing the shaft rotation orders and their harmonics, the gear mesh frequency, their harmonics and the first order sidebands.

The resulting difference signals for the original TSA records as well as those obtained from the PE are presented in figure 6. The PE difference clearly shows evidence of the existence of a fault around 65 degrees in all channels. On the other hand, the TSA difference only shows the presence of a fault in channels 1 and 2 and their existence is not as clear as in the case of PE.

Finally, in order to summarize all the information available, the PE errors for each channel are multiplied obtaining the Prediction Error Product (PEP). Figure 7 shows the PEP results obtained, for 100% load, using the trained (4-10-2-10-4) **NFI**. The local defect is clearly identified on the register where it is present. The results are similar when the training is carried out using **NFII**.
5 Conclusions

In this work a new data processing procedure for fault detection in gearboxes is introduced. The proposed technique is based on the use of an AANN for capturing the relationship among the instantaneous TSA values coming from several measurement points. Once trained, the AANN enables the calculation of
the Prediction Error that is used for detection purposes. The efficiency of the proposed method has been demonstrated analysing several experimental cases. The results were encouraging, showing that the PE signal enables the identification of local faults for any load level. Different processing was applied to PE signals showing their superiority for detection versus the application to TSA records. The information from all available channels could be condensed by multiplication obtaining the PEP which is a simple but very sensitive tool.

The proposed tool is applicable for any speed or load level, condensing the available vibration data records coming from several sensors located on the gearbox case and allowing an earlier detection of local faults. Although the results obtained have were very promising, still more efforts should be carried out on the experimental validation as well as on the post processing of PE signals in order to exploit their periodic structure.

Acknowledgment

This paper has been developed in the framework of the Project DPI2003-01845 funded by the Spanish Ministry of Science and Technology.

References


