Rotating Machine Diagnosis using Acoustic Imaging and Artificial Intelligence
SURVISHNO Conference 2019

Abdelhakim Darraz[1,2], Jérôme Antoni[3], Pierre Mollon[2], Gabriel Kirié[2], Jean-Hugh Thomas[1], Anthony Larcher[4]

[1] Université du Mans/ENSIM, LAUM, UMR CNRS 6613, F-72085 Le Mans, France
[2] VALEO Equipements Electriques Moteur, F-38070 Saint-Quentin-Fallavier, France
[3] Univ Lyon, INSA Lyon, Laboratoire Vibration Acoustique, F-69621 Villeurbanne, France
[4] Université du Mans/ENSIM, LIUM, EA 4023, F-72085 Le Mans, France

Corresponding author: abdelhakim.darraz@valeo.com

Abstract

Mass production of quality equipment in the automotive industry requires controls throughout the production line. These controls are done through monitoring and validation tools for both production and finished products. The use of signal processing methods, applied to acoustic and vibratory recordings collected during the operating cycle, aims to ensure that they are in good working order, to maintain them and to guarantee the quality of the service provided by a manufacturer to its customers. However, sometimes the techniques used do not reach the expected performance, which of course depends on the defect to be recognized but also on the conditions under which the measurements were made. This paper introduces a preliminary study on machine diagnosis by combining both signal processing methods and artificial intelligence. This work is dedicated to develop a system which allows us to measure several high frequency channels and to transmit them to a computer via USB interface.

1 Introduction:

In a production environment many parts are unfairly detected as defective when monitoring is based on indicators from the literature. The causes of these errors are often related to the not conducive noisy environment to such a diagnosis by records sensitive to disturbance. Moreover, from one production site to another, it is not possible to apply the same default detection thresholds because of a different environment involving a variation of the structures and of the product frequency responses. Therefore, today it remains difficult to do a relevant diagnosis in a noisy environment and particularly on non-stationary signals. The aim of the study is to improve this diagnosis by first using a microphone antenna and then operating an artificial intelligence process on a database acquired on production benches. The microphone array leads to the provision of a spatial map of the acoustic field generated by the monitored system. An acoustic imaging approach allows the addition of a new spatial dimension in the data representation. The preliminary study presented consists in differentiating several states of the system to be monitored from the simultaneous exploitation of information expressed in the time-frequency-space domain.

1.1 Rotating Machine

The system considered for the study is a starter. A starter is an electric motor used to rotate a thermal engine until the self-combustion in the engine takes over. It is composed by several subassemblies such as a planetary gear reducer, a drive shaft, an armature permitting under the effect of a magnetic field the rotation of the motor and a current transmission subassembly allowing the current to flow to the armature by brushes.

Each subset radiates its own acoustic signal which, by interactions, contributes to the overall signal emitted by this motor. These interactions associated to the resonance phenomena in the transient operating phases of the machine make the diagnosis more difficult. Indeed, this diagnosis allows to focus the acoustic measurement on the rotating machine by a beamforming process while freeing of the disturbing sources coming from other directions. First results of comparisons will be presented.

There are three states of the electric motor to recognize:

- Healthy state
- Armature imbalance
- Gear defect

Figure 1: Electric motor

1.2 Classification

Automatic classification is a branch of artificial intelligence. Artificial intelligence, commonly known as AI, is defined as the set of theories and techniques used to make machines capable of simulating human intelligence [1]. This field was born in the middle of the 20th century with the development of computer science and the ambition to create machines with the ability to think in a similar way in their functioning to the human mind in terms of perception, understanding and even in taking decision.

Automatic classification is used in all pattern recognition systems. A pattern in the broad sense is an object of very varied nature. It can be a bar code, a face, a fingerprint and more generally a digital data suite that will constitute the signature of the belonging of an object to a family.

Indeed, these systems allow the algorithmic characterization of objects and consist of assigning an object to a class or category based on prior learning.

Learning is the process of constructing a general model based on particular observations of the real world in order to predict a behaviour or a decision in front of new unseen data. The second definition of learning, taking more the sense of training, consists of improving the performance of the model in a progressive way by being confronting to the exercise of an activity.

This idea of improving the accuracy of models through training is easy to perceive for humans since the ability of a person to perform a certain task is often judged by his experience in the field.

2 Experimentation

2.1 Simple sensor analysis

2.1.1 Machine learning

Supervised Machine Learning requires expert know-how in the intended field application. Indeed, it is a question of being able to label the samples correctly on the one hand but also to define relevant indicators to characterize samples regarding the classification to be carried out in the sense that these indicators must be representative for class distinction.

Here the chosen features are indicators coming from signal processing such as:

- RMS value
- Peak to peak
- Kurtosis
- Partial levels around kinematic frequencies
- Global level
- Energy in one-third octave band

In order not to bias the classification for the algorithms used in this part, a step of standardization of these different parameters is necessary. This consists in subtracting a value of a parameter by the average of the values of this parameter and then by dividing the obtained value by the standard deviation of the parameter. For most classification algorithms it is also important to use the same number of samples per class, i.e. to have a balanced dataset. Indeed, Bayesian algorithms introduce the probability of belonging to a class in the calculation of the conditional probability that a sample belongs to a class knowing its characteristics.

In this part we are primarily interested in a diagnosis of the electric motor based on a single microphone recording. The setup consists on the analysis of a signal coming from a single omnidirectional microphone located at 50 centimeters from the electric motor. The dataset is composed by 25 starters per class.

	Bayesian	Neural Network	Decision Tree	Support Vector Machine	Random Forest
True Healthy	24	21	21	21	25
False Healthy	2	5	2	7	1
True Imbalance	25	23	25	21	25
False Imbalance	0	1	1	2	0
True Gear	23	22	22	22	24
False Gear	1	3	4	2	0
# Error	3	9	7	11	1
Performance	96%	88%	90.67%	85.33%	**98.67%**

Table 1: Classification results

The classification mode used for the results obtained above is the so-called cross-validation mode. Cross-validation consists in dividing the population into N groups of constant size. Then, N-1 groups are randomly selected and used to learn and create a model while the Nth group is used as a test population. This step is repeated N times so that all samples are used for learning and testing in order to create N models.

The global performance is an average of the N classification results created. Cross-validation thus makes it possible to obtain a faithful and global performance of the tested algorithm but has the disadvantage of non-negligible cost in terms of computing time.

Table 1 indicates a very good fault recognition performance for all algorithms, particularly for Bayesian and Random Forest, which respectively achieved 96% and 98.67% of recognition. These results confirm the interest of using such methods for machine diagnosis. However, these measurements were carried out in a healthy environment without external disturbance.

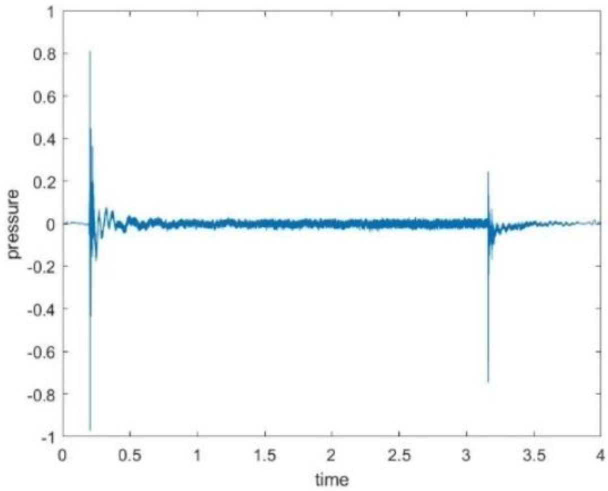

Figure 2: Acoustic pressure signal of the electric motor in unload condition

In the first case, the Mel Frequency Cepstrum Coefficients (MFCC) are extracted from the time signal. This extraction of the coefficients is developed using the Fast Fourier Transform (FFT) and the Discrete Cosine Transform (DCT) on the Mel scale which is a perceptual parametric model. These are the most used criteria in Automatic Speech Recognition (ASR) systems.

First, the signal is split into N windows of a few milliseconds with overlapping (usually in ASR the length of the window is 25 ms and the overlap is 10 ms). A Hamming window is then applied to the signal to limit the spectral distortion (appearance of parasitic high frequencies) related to the overlap and the disturbance at the beginning and the end of the window before passing to the frequency domain via the FFT.

A conversion of the frequency scale f to the Mel scale is performed according to

$$mel(f) = 2595 * \log(1 + \frac{f}{700}). \quad (1)$$

A triangular response filter bank with variable frequency bandwidth is applied to simulate the response of the human ear in the best possible way. This band variation represents the capacity of the human being to be able to easily distinguish two near frequencies at low frequencies than at high frequencies. For each triangular filter, the sum of the energies is calculated so we get as many coefficients as filters.

Finally, we convert the logarithmic spectrum of Mel obtained to the time domain with the DCT, and then we usually keep the first 12 coefficients for each of the windows.

The coefficients are calculated according to

$$C_k = \sum_{i=1}^{N} \log(E_i) * \cos[\frac{\pi k}{N}(i - 0.5)] \quad (2)$$

- N: number of filters.
- E_i: energy calculated with the i^{th} filter.
- C_k: k^{th} Mel Frequency Ceptsrum Coefficient.

Short windows are used for the transient operation phases of the motor while longer windows are used on the stationary phase (this phase is not used in its entire duration; an average and a standard deviation are calculated to take into account any variation of the remaining signal). In addition, to add more information and thus improve the signal recognition, the differential coefficients Δ and acceleration ΔΔ are implemented. These coefficients, calculated directly from the MFCC coefficients taking respectively the derivative of the first and the second order, make it possible to consider the dynamics of the signal.

According to [2], the addition of these differential coefficients and accelerations increases the recognition by about 20%. At least 702 features were extracted from each time signal. Here the dataset is reduced to 15 starters per class.

	Bayesian	Decision Tree	Support Vector Machine	Random Forest
True Healthy	13	11	11	14
False Healthy	7	7	4	3
True Imbalance	14	14	15	14
False Imbalance	0	2	2	0
True Gear	9	9	10	13
False Gear	2	2	3	1
# Error	9	11	9	4
Performance	80%	75.56%	80%	**91.11%**

Table 2: Classification results using MFCCs

Table 2 shows the performance achieved by algorithms. Random Forest remains more efficient than the other algorithms with a recognition performance of 91.11%. However, the overall performance achieved with this approach is less than the performance obtained previously in Table 1.

2.1.2 Deep Learning

The aim here is not to extract previously mentioned indicators, but rather to provide complete acoustic data to the algorithms. Spectrogram as input was studied.

The classification consists in identifying the class of each starter by recognizing the spectrogram of the acquired signal. For this study, an image processing approach based on a pre-learnt Convolutional Neural Network (CNN) AlexNet is used. AlexNet was developed at Toronto university [3] for the ImageNet LSVRC-2010 contest, a competition for which it is proposed to classify 1.2 million images into 1000 different classes. The authors achieved a winning top-5 test error rate of 15.3% for the ILSVRC-2012 contest.

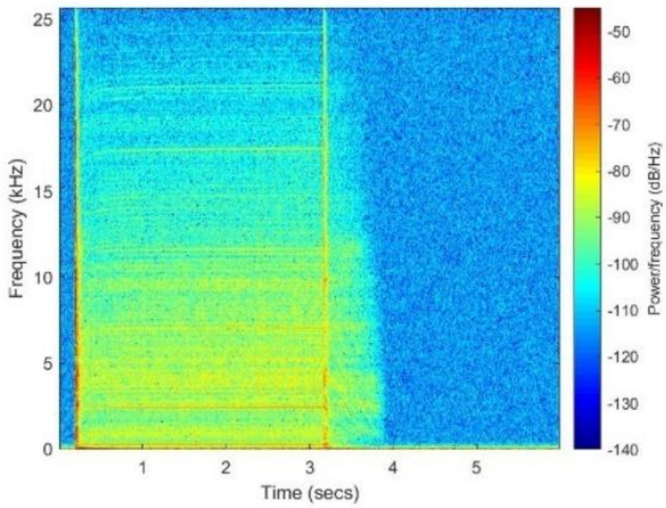

Figure 3: Acoustic signal spectrogram of the electric motor in unload condition

Transfer learning consists in using a trained Neural Network to extract features, which achieved very good performance in recognition on a specific field with thousands of samples for the learning phase, and to transfer the knowledge learnt in that field to another field. This is what we have done here, AlexNet has been trained on 1 million images and performs very well on object recognition in images.

This allows us to extract relevant features for image recognition even if we only have a low amount of data. A modification of the final layers and a training phase are necessary in order to adapt the classification to our field by changing the number of outputs and the weights of each neuron.

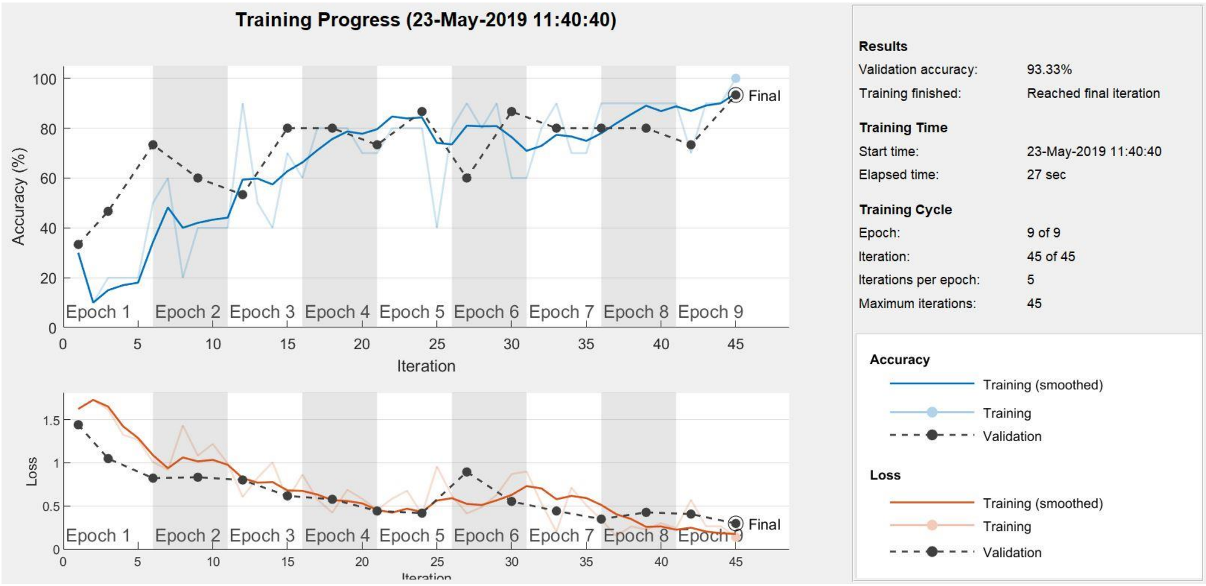

Figure 4: Acoustic pressure by time of the electric motor in unload condition

We achieved a good validation accuracy of 93.33% with this method compared to results obtained by MFCCs. However, it is possible that the model overfits the data because of a large number of parameters (60 million features) compared to the number of samples. In statistics, overfitting corresponds to a model or an analysis which fits perfectly with a dataset. In machine learning, overfitting is one phenomenon to avoid because these models perform very well on the data used to create the model but cannot generalize on new data which means low performances.

2.2 Acoustic localization

Unfortunately, it is not possible to obtain such a recognition rate in a production environment. Indeed, the random ambient noise generates ambiguity perturbing the recognition of defects. This is why we have chosen to introduce acoustical imaging to our work in order to locate the sound sources in a space, focus the measurement towards a privileged direction and thus filter the signal emitted by the electric motor from the ambient noise.

An acoustic antenna represents a distribution of microphones according to a specific geometry and number of sensors. The resulting geometry formed by the repartition of these microphones as well as their number condition the performance of the system, which is characterized by the frequency range of use, the resolution and the capacity of localization. For this project, we made the choice of digital MEMS (Micro electro-Mechanical Systems).

This technology introduced for the first time in 1967 [4] corresponds to miniaturized sensors or actuators which couple several physical principles including mechanics and electronics. The miniaturization allowed by this technology responds to a growing need related to congestion, sensitivity, mass production and to the complexification of systems by allowing the addition of new functions in a non-intrusive way.

The method used to represent the acoustic field is a Beamforming-based method.

2.2.1 Acquisition System

The acquisition system is composed of MEMS microphones. The working of these systems remains faithful to those of conventional microphones because in a general way the physical laws governing the different domains are unchanged. A MEMS consists of a fixed substrate (a semiconductor generally made of silicon) and a moving part.

In the case of the MEMS microphone, the moving part is represented by a membrane. A conductor measures the impedance variation induced by the deformation of the membrane and this allows to obtain a minimum size while not skimping on the performance of the sensor.

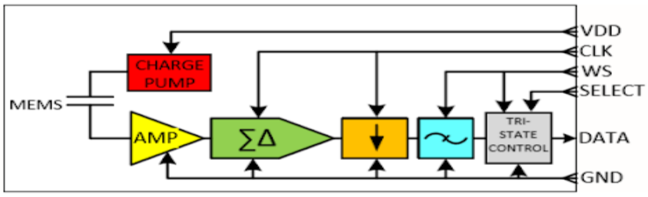

Figure 5: Block diagram of the I2S digital mic

- The charge pump (in red) allows the supply voltage to be raised to a level necessary to polarize the transducer.

- The MEMS is responsible for converting the measured pressure into a voltage.

- The amplifier (in yellow) stores the voltage from the MEMS and amplifies the signal.

- The Sigma Delta converter (in green) converts the analog signal in memory into a pulse-density modulation (PDM) signal on 1-bit resolution.

- The decimator (in orange) converts the PDM signal into a 24-bit pulse coded modulation (PCM) signal by down-sampling with a factor of 64.

- A low pass filter (in blue) removes the remaining high frequency components.

- The three-state mode (in gray) makes it possible to associate two MEMS on the same I2S line for a stereo recording.

Pre-assembled I2S digital MEMS microphones with welds on PCB plate already made were chosen for convenience because the size of the microphone involves a high precision in the welding process.

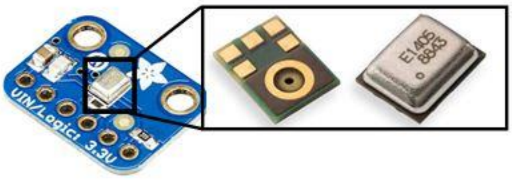

Figure 6: I2S MEMS microphone

An I2S output format MEMS implies that the microphone integrates a large part of the acquisition chain and a MiniDSP acquisition board is used to measure up to 8 synchronous I2S channels by board.

2.2.2 Beamforming

Beamforming [5] is a signal processing tool used in the field of antennas for the directional signal transmission or reception. The main principle of this method is to combine the elements of a sensor array in such a way that in a particular direction signals interfere constructively whereas in other directions the interferences are destructives. It is a question of measuring and applying a delay on the signals acquired by the microphones.

Beamforming is used with radio or sound waves and has many applications in radar, sonar, seismology and acoustics. It allows a representation of a wave field by an estimate of the directions of arrival.

Let an antenna formed by 10 microphones (red stars in FIG. 6), a sinusoidal source (green circle in FIG. 6) with a frequency of 2500 Hz and a white noise (blue circle in FIG. 6). The sources are both placed 0.20 meters from the plane of the antenna and spaced from each other by 0.05 meters.

Through a Beamforming algorithm, we are able to distinguish these two sources.

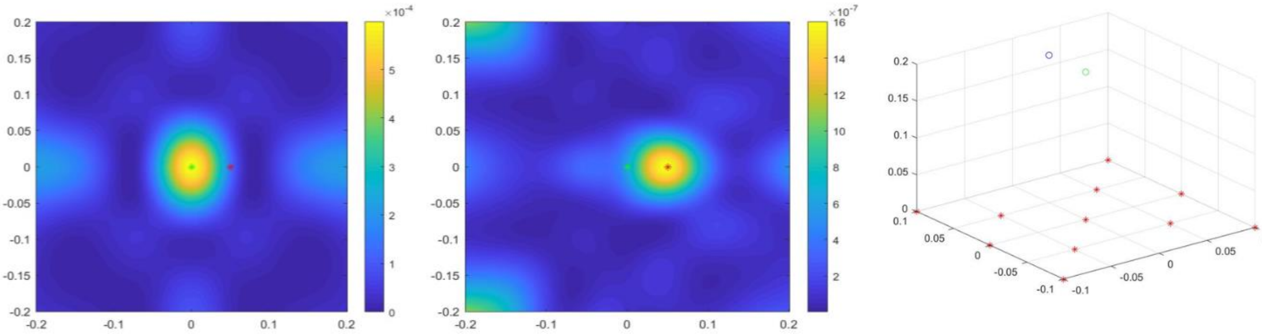

Figure 7: Measurement of the pressure field at 2500 Hz on the left, at a different frequency in the middle

Using an antenna makes the diagnosis less sensitive to disturbance and thus more reliable. We can from this observation reconstruct the acoustic signal coming only from the direction of the source in order to apply the diagnosis to a non-noisy signal. Therefore, we hope to find at least the results obtained in Table 1. An improvement in these results will then be considered from a more advanced imaging diagnosis.

Conclusion and further works

The results introduced above show that the use of classification algorithm is relevant to diagnose electric motors. However, even if these results seem good it is not possible to consider an implementation of these tools in a production bench because 1.33% of False Positive misclassification, achieved with Random Forest in 2.1.1, on a production of 5000 starters per day first means that 67 electric motors are misclassified and in this particular case 67 electric motors are in a bad condition but classified as healthy could be sent to customers.

We will first focus the work on collecting data with the microphone array in a production environment and then improve the recognition by using classification algorithms developed on time signals and spectrograms. By using algorithms near from our application field and adding spatial informations with Beamforming, we hope increase the recognition tasks.

References

[1] C. Bishop, Pattern Recognition and Machine Learning, New York, Springer, 2006.

[2] Hanson, Brian, and Ted Applebaum. "Robust speaker-independent word recognition using static, dynamic and acceleration features: Experiments with Lombard and noisy speech." Acoustics, Speech, and Signal Processing, 1990. ICASSP-90., 1990 International Conference on. IEEE, 1990.

[3] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. "ImageNet Classification with Deep Convolutional Neural Networks", Pages 1097-1105 NIPS'12, 2012.

[4] H.C.Nathanson, et al., "The Resonant Gate Transistor", IEEE Trans. Electron Devices, March 1967, vol. 14, 1967.

[5] J.J. Christensen, J. Hald, “Beamforming”, Brüel & Kjaer Technical Review N°1, 1-48, 2004.

[6] J.-H. Thomas, Diagnostic d’un système non stationnaire à partir d’une approche de reconnaissance des formes floue. Application au diagnostic de cliquetis moteur, Traitement du Signal, 16 (3), 203-216, 1999.

[7] M. Mein, Surveillance acoustique de moteurs diesel, thèse de doctorat de l’Université du Maine, Le Mans, 1997.