State of the art and developments of diesel engine in-line tests monitoring

État de l'art et développements de la surveillance en ligne des essais de moteur Diesel

Luigi Garibaldi
Dipartimento di Meccanica
Politecnico di Torino
C.so Duca degli Abruzzi, 24
10129 Torino
Italy
E-mail : luigi.garibaldi@polito.it

Jerôme Antoni
Université de Technologie de Compiègne
Département du Génie Mécanique
BP 20529
60205 Compiègne CEDEX
E-mail : jerome.antoni@utc.fr

Les développements récents ont permis un progrès significatif dans l'optimisation de l'essai à froid des moteurs, et pour leur diagnostic et leur contrôle en ligne.
Les pas en avant considérables réalisés ces dernières années sont principalement dus aux capacités de progrès des technologies de capteurs, et à leur transfert entre domaines, à l'évolution incroyable du traitement numérique du signal, et à son accélération bénéfique, enfin aux techniques multi-échelles développées pour les différentes échelles de fréquence.
Les techniques et les dispositifs de diagnostic sont déjà présents dans la production des moteurs, mais souffrent toujours de leur faible capacité à traduire des signaux en symptôme de défaut, c'est-à-dire en niveau d'alarme, avec un niveau raisonnable de confiance et, plus fondamentalement, de maintenir un niveau bas de fausses alarmes.
L'association des dispositifs d'émission acoustique et de l'information des signaux à haute fréquence, avec ceux plus classiques d'analyse modale et d'essais de vibration, peut avantageusement conduire à une amélioration des résultats d'identification. Cet objectif peut être atteint en exploitant les effets de modulation sur les signaux à haute fréquence, comme montré par quelques articles récents.
Une revue des techniques mixtes possibles est donnée dans l'article, ainsi que des exemples réels à partir des bancs d'essai de compagnies majeures de fabrication de moteurs de camion.
Quelques commentaires sur les outils mathématiques possibles sont également donnés et analysés.
Un schéma de l'approche itérative suggérée au cours de l'article est donné, en restant encore sous une forme très générale.

Recent progresses in the engine cold test monitoring techniques have been achieved by adopting a multi-disciplinary approach, as for many other engineering tasks.
The considerable steps ahead done in this domain over recent years are mainly due to the growth capacities of sensor technologies and their transfer from other domains, the incredible evolution of digital signal processing and its favourable speed-up, the multi-channel techniques developed for different frequency scales.
Diagnosis techniques and devices are already present in engine production but still suffer for their low capacitance in translating signals into a fault symptom, i.e. a level of alarm, with a reasonable level of confidence and, more fundamental, maintaining a low level of false alarms.
Mixing the acoustic emission devices and high frequency signals information with those from more classical modal analysis and vibration tests can advantageously lead to better identification results. This target can be achieved by exploiting the modulation effects on the high frequency signals as shown by few recent papers.
A review of possible mixed techniques is given in the papers, together with real examples from a leader company manufacturing test rigs for truck engines.
Some comments of possible mathematical tools are also given and analysed.
Comments concerning the future work and the possible obstacles are given at the end of the paper.

Cold Engine Test (CET) monitoring involves a number of different engineering aspects, often correlated one each other, and belonging to a wide range of engineering experiences. As already stated in many articles on the subject, few fundamental steps are pointed out in the procedure to establish the "health state" of a generic structure, i.e. the diagnosis task. This procedure can be adopted for Cold Engine Test too, and mainly concerns the following tasks:

- testing (type of the test)
- sensing (instrumentation needed)
- diagnostics (mainly software tools)
- variability (production variability)
- confidence (setting of confidence levels)

The paper mainly concerns the points number one to three to seek for robust solutions nowadays adopted in the diagnostic environment but also to improve previous existing techniques.
Conversely, task 4 is strictly related with the production output and tolerances accepted in manufacturing, whilst task 5 can be considered as a consequence of point 4.
Back to point 1 to 3, the main steps of the diagnostic procedure can be resumed in the following points :

- Type of physical quantities to be acquired
- Devices to capture the above quantities
- Signals characteristics and their associated analysis tools
- Diagnostic parameters
- Diagnostic principles

Type of physical measurements

Vibrations

Vibrations can be measured through displacement, velocities or accelerations.
For low frequencies and high displacements, displacement measurements are suitable, while for high frequencies and low displacements acceleration measurements are obviously better.
This consideration leads to the fact that low frequency motions are better measured by means of displacement sensors, while high frequency motions are better measured in terms of acceleration, whose quantity result multiplied by ω with respect to velocities. However, in real world, most of the physical quantities are measured as a matter of other facts, such as the industrial production, the cost on the market, the reliability and robustness, more than for pure technical considerations.
The transformation from one physical quantity to another (derivation and integration) is not at all straightforward as proved by the massive number of publications on the subject.

Laser techniques can also be advantageously adopted for Cold Engine Testing, this allowing the overtaking of sensor placement and fixture, always representing a source of problems and noise.

Acoustic Emission (high frequency emission beyond 50 kHz)

Acoustic Emission (AE) is measured by devoted resonant devices, which are sensible to very high acceleration components due to energy releases in the structure and picked-up on the surface of the engine under test. Under this point of view, they simply measure accelerations in a specific range.
Resonance peak frequencies are often between one hundred and 300 kHz.
The scope of these devices is not to directly measure the acceleration signals, but just to measure the number of events overtaking a set of specific soils settled by the operator on the basis of similar structures and under his own experience.

Acoustic Emission (frequencies in the audible range)

Audible frequencies also can be adopted as a monitoring tool, as for many other automated applications.
In spite of its simplicity, some difficulties might rise up, due to the high level of noise usually surrounding the engine test environment. Some specific techniques to separate the sources can be adopted, but the quite often the environment noise is neither white (or harmonic) nor stationary, hence giving further troubles in applying separation techniques.

Torque

Torque measure can be helpful to control the actual input into the system under test. High component might be filtered out, due to the electric motor driving the engine under test.
Transfer function between the input and the output (i.e. pressures and accelerations) might be informative to seek for anomalies.
Unfortunately, quite often, the applied torque is loop controlled to give average constant speed of the engine, so giving origin to extra signals to be identified and filtered along the diagnostic process.

Pressure

In the case of pressure, being the physical quantity “slowly” variable in comparison with other physical quantities, the measure does not pose hard problems, providing a good coupling with the acquisition chain and an opportune, recommended, noise surveillance over the acquisition chain.
Cross coupling between sensors and channels has to be also checked to reduce the spurious signals, due to cross talking, as for the case of all the other measures.
Pressure transducers should be well isolated from the vibrating environment, to uncouple the vibration input they are sensible in some cases.

Devices

Accelerometers: acceleration measures

To detect high frequency contents, accelerometers are ideal, being their piezoelectric version well suited to measure frequencies up to 10 kHz or more.
Accelerometer with larger dimensions (up to 50-100 grams) experience higher sensibility but also a reduced frequency range; their weight might also be not negligible compared with the structure under test, particularly when used for thin structures and plates. Their fixture is posing a number of problems to be solved in case of automated, on line tests.

Magnetic devices: velocity measures

This sensor holds the advantage of non-contact measurements, but is unfortunately limited to ferro-magnetic systems. Its measures the relative velocity between the sensor and the engine.

Capacitive devices: displacement measures

This sensor uses the capacitance between the sensor and the engine to measure the gap (displacement).
The object is limited to a conductor. Capacitance is a function of the conductor area S and the gap D. The sensor and the object should show flat and parallel surfaces.

Eddy Current Sensor: Displacement measures

Since this sensor uses eddy current, the material of the object is limited to metal. If a metal object is close to the high-frequency coil of the transducer, the alternating magnetic field in the coil generates eddy current in the metal. This eddy current depends on the magnetic force at the metal surface, or the distance between the coil and the object.

Laser Trigonometrical (short distance) Displacement measures

Is a displacement sensor using optical trigonometry. Using a laser as a light source ensures high accuracy and high stability. When the laser arrives at the object surface, part of the diffusely reflected light is focused on a position sensing detector (PSD).
We can measure the object displacement by detecting the spot that varies with the displacement.

Laser velocimeter (high distance): velocity measures

Velocity is calculated by means of the Doppler effect on laser beam. High frequency are allowed. Precision is very high but its tuning and maintenance along the engine production chain might be complex. Relative velocity is measured.

Signals characteristics and their associated analysis tools

This section deals with the types of signals generated by internal combustion engines and, accordingly, on the signal processing tools suited to each case.

Periodic signals (torque and pressure measurements)

Low-frequency measurements are essentially periodic in nature, with fundamental period the engine cycle and the firing cycle (firing cycle = engine cycle × number of cylinders). Note this is assuming the engine running at constant speed and torque and, strictly speaking, having the measurements expressed in terms of the crankshaft angle rather than time.
Typical examples of periodic signals are instantaneous torque, pressure, and speed measurements.

Periodic signals are perfectly described by a (small) countable set of Fourier coefficients whose manipulation is always easier than that of the original waveforms. Since the Fourier coefficients contain all the diagnostic information, they can be profitably used in an automatic decision system.

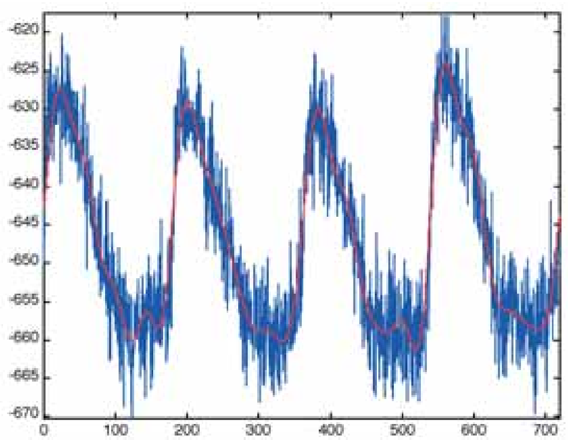

Fig 1 : Example of a Fourier series of an internal pressure measurement: 20 Fourier coefficients are enough to summarise (red curve) and denoise the diagnostic information of a 1440 samples waveform (blue curve)

However many coefficients may be necessary when the signal of interest have discontinuities. In such situations, it may be better advocated to use *wavelet coefficients* which, in addition to sharing the same advantages of the Fourier coefficients are also compactly located in both time and frequency [1].

Cyclostationary signals (vibration measurements)

Measurements performed over a high frequency band usually exhibit a combination of a periodic and a random part. Recent studies have demonstrated that both parts contain diagnostic information, and so should be analysed together [2]. The random part, in spite of having a zero cyclic average, actually has a *periodic* second-order statistical structure. Such signals have been coined *cyclostationary*, i.e. with non-stationary statistics inside a cycle, but stationary from cycles to cycles. This is illustrated in the following figure.

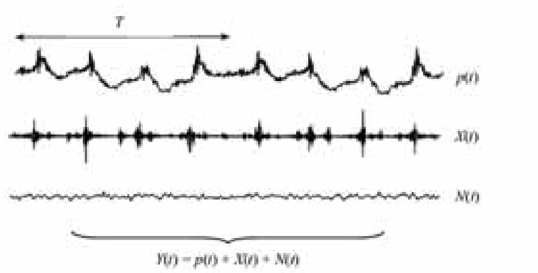

Fig. 2 : Cyclostationary signals are a combination of a periodic part, a random part with periodic statistics, and possibly a stationary part

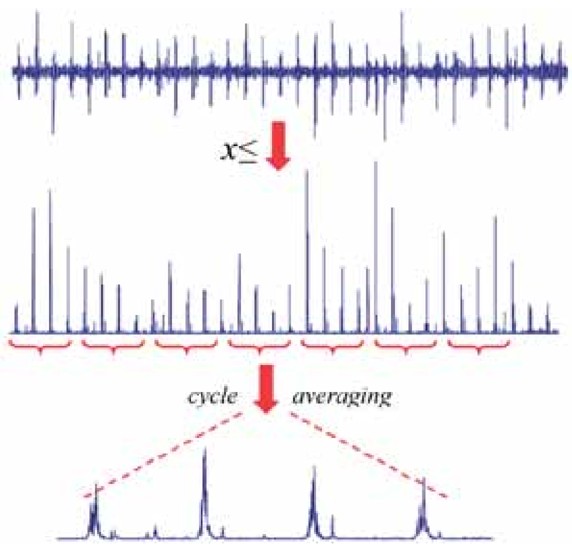

Fig. 3 : Example of a cyclostationary vibration signal: its hidden periodicity is revealed through squaring and cyclic averaging the signal

A recent and efficient theory has been developed for analysing cyclostationary signals. The idea is to apply a non-linear transformation to the signal – such as squaring - which brings to the fore its hidden periodicity. Then cyclic averaging can be used to extract the periodic time waveform, on which the Fourier series or the wavelet series can in turn be applied, as described in the previous subsection. Various transformations (such as filtering, time windowing, etc.) can be applied to the signal before the non-linear transformation.

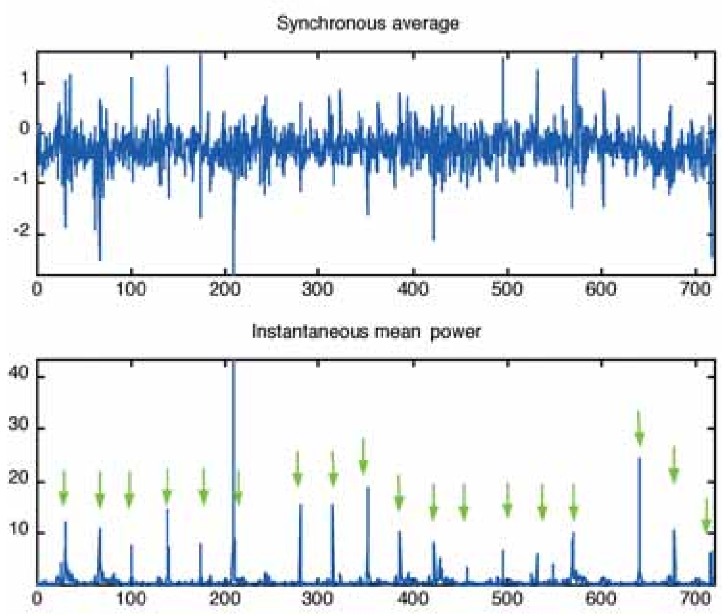

Fig. 4 : Example of analysis of a cyclostationary vibration signals. The instantaneous mean power reveals the presence of abnormal chocks with a rate of 20 occurrences per cycle.

Nonstationary signals (run-up, run-down measurements)

This class contains all measurements which are neither periodic nor cyclostationary, for instance during a run-up or run-down of the engine.
Nonstationary signals are much more complicated to analyse than periodic and cyclostationary signals. However this inherent difficulty should not discourage their consideration, because it is often during the run-up or run-down stages that abnormal phenomena can be detected (this is the technique used by garage mechanics).

Mature techniques based on *time-frequency* representations currently exist which can handle non-stationary signals [4]. When concerned with the diagnostic of rotating and reciprocating machines, time-frequency analysis is often replaced by a "speed-frequency" analysis where the time axis is substituted by a speed axis.

Diagnostic parameters

Primary objective

Diagnostic parameters should summarise the measurements in the most relevant and parsimonious possible way. They should contain enough diagnostic information such as to make possible the easy detection and recognition of a fault.

Secondary objectives

In order to use the diagnostic parameters in an automatic decision system, it is desirable that they fulfil a number of additional properties.

- the diagnostic parameters should be as insensitive as possible to additive noise,
- the diagnostic parameters should be as independent as possible in order to avoid unnecessary information redundancy,
- the diagnostic parameters should be easy and fast to compute, e.g. by means of linear transform or simple non-linear operation,
- the diagnostic parameters should have as regular statistical properties as possible: normal distribution, mutual statistical independence, low standard-deviation-to-mean ratio,

Proposition of time domain parameters

One solution is to consider the whole time waveform or a function of it as series of *instantaneous* parameters. However this solution is highly redundant. Usually, when possible, the time waveform is summarised by means of a small number of scalar parameters which are either *global* or *local*.

• Instantaneous parameters:
- the whole signal
- filtered versions of the signal in frequency bands that maximise the signal-to-noise ratio of some given faults
- the instantaneous power of the signal, i.e. the squared signal as a measure of the energy flow as a function of time
- the instantaneous powers of the signal in given frequency bands
…

• Global parameters:
- mean value (time average), median value
- crest value, crest-to-to crest factor
- energy parameters (rms value)
- dispersion parameters (inter-quartile range, skewness, kurtosis)
…

• Local parameters:
- position and magnitude of the maxima in given intervals
- position and magnitude of the maximum slopes in given intervals
- crest-to-crest factors in given intervals
…

Proposition of frequency-domain parameters

Different phenomena may be characterised by different frequency ranges. Therefore spectral analysis offers a convenient solution to separate them. In the case of periodic measurements, spectral analysis is made particularly simple as it solely consists of computing the coefficients of a Fourier series through FFT algorithms.

The same holds true for cyclostationary measurements, except that the Fourier coefficients should be computed on some non-linear transforms of the signal, such as the instantaneous power of the signal.

The Fourier coefficients have most of the desirable properties to be taken advantage of in an automatic decision system:
- only a few of them are usually enough to summarise the signal,
- they are independent to each other in the sense that the knowledge of one coefficient does not require the knowledge of the others,
- they are fast and easy to compute,
- they tend to be normally distributed according to the Central Limit theorem, i.e. they have a well-behaved statistical distribution.

Because the magnitude of the Fourier coefficients measure the energy of the engine harmonics, they are often used to detect faults that produce some loss in symmetry in the signal: one faulty cylinder is very easy to detect this way as it produces abnormally high harmonics of the engine cycle as compared to the harmonics of the firing cycle. Exact identification of the cylinder requires inspection of the phase of the Fourier coefficients, or parallel inspection of local time-domain parameters.
In the case of non-stationary measurements, this simple approach based on the Fourier coefficients no longer applies. Non-stationarity makes the Fourier coefficients become functions of time or engine speed. Hence one solution is to track the magnitude and the phase of the Fourier coefficients as functions of speed: this is the principle of order tracking.

Time-frequency parameters

A very discriminating way of characterising a fault is by means of its time-frequency signature; the idea is that every fault always occurs in a certain time interval and over a certain frequency band. Knowledge of the time-frequency areas associated with different faults hence allows their individual recognition.
There are different ways to achieve time-frequency analysis. The simplest one is to use time-domain parameters (instantaneous, global, local) computed in frequency bands. The most sophisticated is by means of time-frequency energy distributions [4].
Another very efficient solution is to use the coefficients of a wavelet series which, due to their high localisation in time and in frequency, may be interpreted as a kind of time-frequency representation. Just as with the Fourier series, the wavelet coefficients are very good candidates for diagnostic parameters [1]. They also have the same desirable properties to be taken advantage of in an automatic decision system.

Diagnostic principles

Reduction of the parameters space

The ensemble of selected diagnostic parameters forms the basis of a high dimensional space in which any engine can be represented. The bottom line assumption of diagnostics is that good and faulty engines occupy different regions in the parameter space and can be geometrically separated on this basis. The smaller the dimension of the parameter space, the better for reasons of :

- speed and burden computation,
- statistical significance (minimum probability of false detection),
- statistical regularity (immunity to additive noise).

There are many solutions to reduce the dimension of the space spanned by the selected parameters. The first one is to retain among all the parameters only a small subset that preserve as much as possible of diagnostic information. This can be done :

- on physical considerations (see figure below)
- by means of statistical criteria such as *Fisher's measure* or the *branch and bound algorithm* [5].

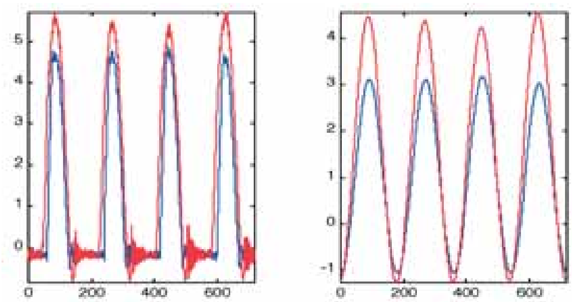

Fig. 5 : These figures illustrates that 6 (left panel) instead 60 (right panel) Fourier coefficients may be enough to discriminate between the good (blue curve) and faulty (red curve) states, even though this is not enough to reconstruct the original signals.

The second solution is to retain a subspace – rather than a subset- of the full space spanned by the parameters. This is usually achieved trhough a singular value decomposition of the covariance matrix of the parameters, with different variants such as [5] :
- the *principal component analysis*
- the *discriminant analysis*

A non-linear approach is also possible which uses neural networks such as *Kohonen's maps*. Both approaches require the values of all the parameters in a large data set of good and faulty engines.

Comparison against a reference

Once a subset or a subspace (i.e. linear or possibly non-linear combinations) of the diagnostic parameters has been selected, it must be compared against a reference in order to discriminate between the good and the faulty state. There are two types of reference signal to be used in a diagnostic procedure:
reference on the good engine
references on specific faulty engines

References are usually obtained from a large data set of individuals. One convenient solution would be to construct this reference progressively and update it each time a new engine is tested. Another solution -which in principle does not need statistics- is to use a theoretical model to simulate a reference. However this assumes the model to be very representative, and may anyway requires some experimental updating.

Statistical decision

Having a reference against which to compare a current measurement is a necessary condition to answer as whether an engine is faulty or not. Into addition, the reference signal should be associated with uncertainty bounds which bracket the "good state" with given probabilities.
Once a reference and its variability are known, then very efficient tests of hypothesis can be applied to answer the questions:

- is the engine good?
- is the engine faulty?
 does the engine have fault X?

The basis of these tests is to put a threshold above the reference such as to yield small probability of false detection:

- *probability of type I error*: the engine is classified as faulty whereas it is good
- *probability of type II error*: the engine is classified as good whereas it is faulty

These tests of hypothesis are very easy to elaborate when the diagnostic parameters of interest are normally distributed (e.g. Fourier and wavelet coefficients):

- *Normal test*
- *Student's test*
- *Fisher's test*
…

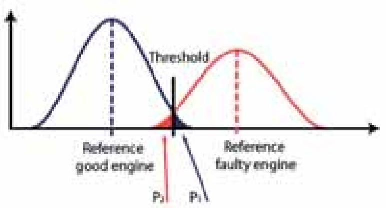

Fig. 6 : Principle of a test of hypothesis: set the comparison threshold such to minimise P1 (probability of deciding the engine is faulty whereas it is not) or P2 (probability of deciding the engine is good whereas it is not).

Typical diagnostics techniques

Linear statistical techniques

These techniques assume the parameters to follow normal distributions. They are relatively easy to implement:

- *simple test of hypothesis*
- *multiple test of hypothesis*
- *sequential test of hypothesis*
- *discriminant analysis*

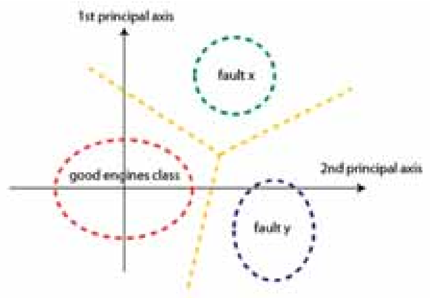

Fig. 7 : Principle of discriminant analysis: the tested individual is projected in the parameter space a space and assigned to the closest class.

Non-linear statistical techniques

These techniques generalise the former ones when the normality assumption of the parameters is relaxed. They are usually more delicate to implement, even though some of them have now become standards.

- *Neural networks*
- *Support Vector Machine*

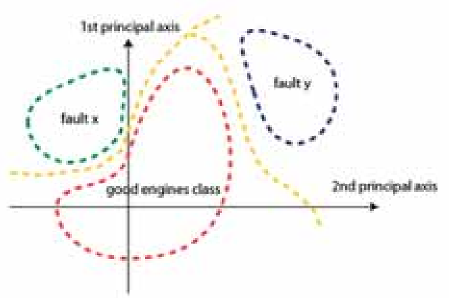

Fig. 8 : Principle of a non-linear statistical diagnostic technique.

Model-based techniques

Contrary to pure statistical approaches, model-based techniques try to make use as much as possible of some *a priori* knowledge of the system. The basic idea is to compare some dynamic equations with current measurements such as to produce a residual signal which is ideally zero when no fault is present and is nonzero otherwise [6]. Some commonly used dynamic equations are:

- that of the system itself
- a functional relationship which describes redundancy between several measurements

Matched filters

The matched filter is a very efficient diagnostic technique for signal-like measurements. It is mid-way between the purely statistical approach and a model-based approach. The principle is summarised in the below figure. The key idea is to design a linear filter (using any digital filter design procedure) that applies to the measurements and ideally returns a zero signal in the absence of fault.

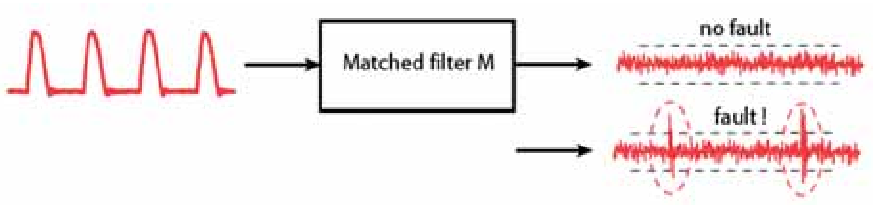

Fig. 9 : Principle of the matched filter.

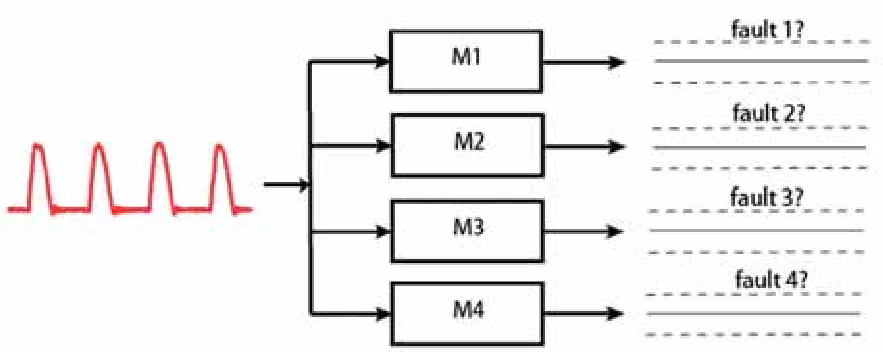

Fig. 10 : Generalisation of the matched filter to discriminate between several faults.

Although this principle is enough to discriminate between the “good” and “not good” states, it can easily be extended to discriminate between several types of faults, by using a bank of matched filters.

Conclusions

The state of the art of monitoring and diagnostic techniques has nowadays reached a mature level, confirmed by the massive amount of publications, revues, journals and conferences on this specific sector of research.
In spite of the tremendous effort spent by the researchers in some strategic domains for diagnostic, such as aerospace, civil engineering (bridges and dams) or military devices, the counterpart of the efforts spent in the automotive industry is definitely poor.
It becomes evident how the knowledge transfer from other disciplines which are nowadays more advanced in diagnostic tasks can only advantage the Cold Engine Test domain performances.
This also implies that research costs are reduced and most of the needed hardware is already available on the market.

Some difficulties or obstacles still should be removed to fully apply the promising performances of these techniques during production:

- time for data acquisition during production is limited and a few cycles only can be monitored, particularly at max speed;
- vibration sensor placement is an important obstacle, due to the automated and quick procedure requested to place the sensors;
- high quality acquisition card is necessary to fulfil the requirements of advanced algorithms;
- new type of measures, such as AE, could be introduced along the procedure and coupled with the other analysis instruments;
- massive use of high level signal processing software and decision maker must be adopted;
- a quasi-real time logic to make correct decisions should be adopted to follow the production variance;
- a complete self-calibration and auto-diagnostic system is necessary to maintain the production tolerances and the false detections.

Acknowledgments

Authors are grateful to Dr Zingali and his staff of AVL Italia for the interesting suggestions and the technical discussions.

References

[1]Stephane Mallat, A Wavelet Tour of Signal Processing, Academic Press 1998.

[2] J.Antoni, J. Danière, F. Guillet “Effective vibration analysis of IC Engines using Cyclostationarity. Part I : A Methodology for Condition Monitoring”, Journal of Sound and Vibration. Vol. 257, No. 5, Nov 2002, pp. 815-837.

[3] J. Antoni, F. Bonnardot, A. Raad and M. El Badaoui , “Cyclostationary Modelling of Rotating Machine Vibration Signals”, Mechanical Systems and Signal Processing, Vol. 18, No. 6, Nov. 2004, pp.1285-1314.

[4] Patrick Flandrin, Joachim Stockler, Time-Frequency/Time-scale Analysis, Academic Press 1999.

[5] Keinosuke Fukunaga, Introduction to Statistical Pattern Recognition, Academic 1990.

[6] Ron Patton, Paul Frank, Robert Clark, Fault Diagnosis in Dynamic Systems: Theory and Application, Prentice Hall 1989.