

Analysis of vibration signals using short-time analysis and clustering in parameter space for detection of combustion engine state

The paper presents a short-time analysis of the vibration signals for the diagnosis of Diesel engine of combustion locomotive by recognition of different engine states using the clustering technique. The main aim of the researches was to distinguish between different engine states represent different wear extends. The proposed method of vibration signal analysis consists on sliding a time window along signal in time and observing the changes of some given statistical parameters. The set of this parameter values creates a multidimensional parameter space where the time evolution can be observed. For recognition and detection of different engine system states some clustering techniques in the parameter space were performed. The results show the possibility of distinguishing different cluster centers within the parameter space which can be assigning to different engine states represented the states before and after a general repair.

Key words: vibration signals, Diesel engine diagnostic, short-time analysis, clustering

1. Introduction

Typical combustion engine after a long usage time usually grow old and run down. The engine diagnostic needs new methods for estimation of the engine state in time of its exploitation. The rail vehicle diagnostic has not the same long history as for example car diagnostic. It starts in seventies of XX century when the number of steam locomotives decrease rapidly and they were substituted by Diesel and electric locomotives. From this time the diagnostic methods in the area of diesel locomotive engines started to develop.

The locomotive Diesel engines the same like other combustion engines are nowadays an important source of pollution. To reduce air pollution for passenger cars the OBD norms were introduced. The main function of the OBD system is a continuous monitoring of basic system parameters. The problem of air pollution concerns also the diesel engines and combustion locomotive engines [14]. The rail area is partially regulated with several regulations considering limits on emission of combustion gases (for example, cart UIC 623 1-2-3 in Europe). But all the time there are no obligatory regulations for systems monitoring the emission critical damages that might play a similar role to that of the OBD for cars. But each year the norms and regulations regarding combustion gases emission regarding vehicles with heavy diesel engines like combustion locomotives are develop. This fact gives an impulse to research for new methods for detection of combustion locomotives Diesel engines faults. To achieve this aim the vibration signals were proposed to use.

The vibration signals were taken from an engine and were processed using the short time analysis and clustering method [7, 9, 11, 17]. The measurements were performed on Diesel locomotive 401Da – 427. Unfortunately in this area the classical methods like spectral analysis appear insufficient. In the paper the analyses consider a vibration signal acquired from the combustion locomotive Diesel engine body in two states: before and after a general repairman. The method of vibration signal analysis proposed in the paper consists on sliding a time window along a signal in time. This gives a possibility to calculate the

values of some chosen statistical parameters for a signal in short time for each time moment. This gives also a possibility to observe the changes of these parameters in time. As a result a set of observing parameter values creates a multidimensional parameter space where the time evolution can be observed. For recognition and detection of different system states a clustering in the parameter space can be performed. The results show the possibility of distinguishing different cluster centers in the parameter space corresponding to engine states before and after a general repair of an engine.

2. Measurements

The acceleration measurements before repair of a Diesel engine were done by using acceleration sensors EGCS Entran Devices of the range ± 5 g. The signal was registered by cart PCL-818HD ADVANTECH with the sampling frequency $f_{Hz} = 1004$ Hz/channel. The measurement after repair were done using the sensors EGCS and some new sensors PCB PIEZOELECTRONICS 393B04 where the signal was amplify by 3-channel signal conditioning amplifier and next registered using analogue to digital cart. The main research object was the 14D40 no 8849 diesel engine of Diesel combustion locomotive ST44.

A Diesel locomotive ST44 have six wheel sets placed on two bogies, each driven by a separate traction electric engine. The electric transmission gear is transferred to a torque of diesel engine (14D40; power output 1470kW = 2000 KM) onto locomotive axles set. The ST44 locomotive was a typical combustion traction vehicle for railway in Poland about twenty and thirty years ago. Nowadays they are gradually withdrawing from current usage [6, 8, 9].

The little different ways of performing the measurement before and after a repair was imposed by specific of measurement sets and circumstances provided in the diagnostic station. The measurements before repair were done under load (on water recoil) for adapted powers in the given measurement point (Fig. 1). The sensors were located on engine body in the places which correspond of engine crankshaft bearing. In each measuring point acceleration was registered in two directions: vertical and horizontal transversal. For each point two measurement series were

done. Because of the limited number of sensors the measurements were performed in three points in one time. The measurements were performed for defined settings describing by fixed values of powers and rotational velocities. The periodic inspection repair included service of subassemblies and assemblies, partial disassembling and change of damaged or worn-out elements.

The measurements after repair were done under load (on water recoil) for adapted powers in the given measurement point. This time the measurements were done in six points. Sensors were the same located on engine body in the places which correspond of engine crankshaft bearing. Each measuring point registered acceleration in two directions: vertical (Entran sensors) and horizontal transversal (PCB sensors). For each point two measurement series were done.



Fig. 1. The example vibration sensors mounted on an engine body in a diagnostic station [1]

3. Data analysis

The global characteristics of the signal like Fourier spectrum do not provide any information about local and instantaneous signal alterations. To find some local characteristic some special short-time methods of analysis can be performed. Nowadays the using of wavelets seems most popular in this area. But the main idea of an approach proposed in the paper was taken from the earlier concept of the short-time Fourier spectrum analysis which bases on the short-time Fourier transformation [2, 3, 15]. For discrete case for function $u(i)$ the following definition can be used

$$U(f, n) = \sum_{i=-\infty}^{+\infty} u(i)h(n - i)e^{-j2\pi fTi} \quad (1)$$

where $h(n)$ is a time function called a window function. The spectrum calculated from (1) is a continuous spectrum and is periodic in frequency f . It depends on the window function form and the moment n [2, 3, 13, 15]. In a special and most simple case the window can be considered as the rectangular function. And the considerations performed in the look for changes in such a short-time window. The approach performed in the paper based on calculating particular parameters in the window sliding in time along the signal. The main advantage of this approach is its simplicity what gives us the direct possibility of diagnostic application. These methods were applied so far rather to singular parameters just to make a simple one-dimensional comparison [8, 9, 11]. In this paper the analysis was performed for more parameters and to consider a problem a multidimen-

sional parameter space was taken into account. The general schema of proposed algorithm is as follows [5, 10, 11]:

1. Choose the width of the time window.
2. Choose the parameters which will be calculated for each window position.
3. In a sliding window for the each given window position calculate the chosen signal parameters.
4. The set of parameter values for the given time moment represents a point in a multidimensional data space.
5. Sliding the window in time along the whole signal gives the set of points in the multidimensional parameter space, which represent the evolution in the parameter space for successive time moments.
6. The points in multidimensional parameter space can create some groups which should represent the state of an engine.
7. The comparison of different signals for different engine state base on analyzing the clusters in multidimensional parameter space.
8. Eventually the diagnosis will consist of distinguishing between different engine states basing on the different clusters created in the parameter space (comparison of the cluster centers) (see Fig. 2).

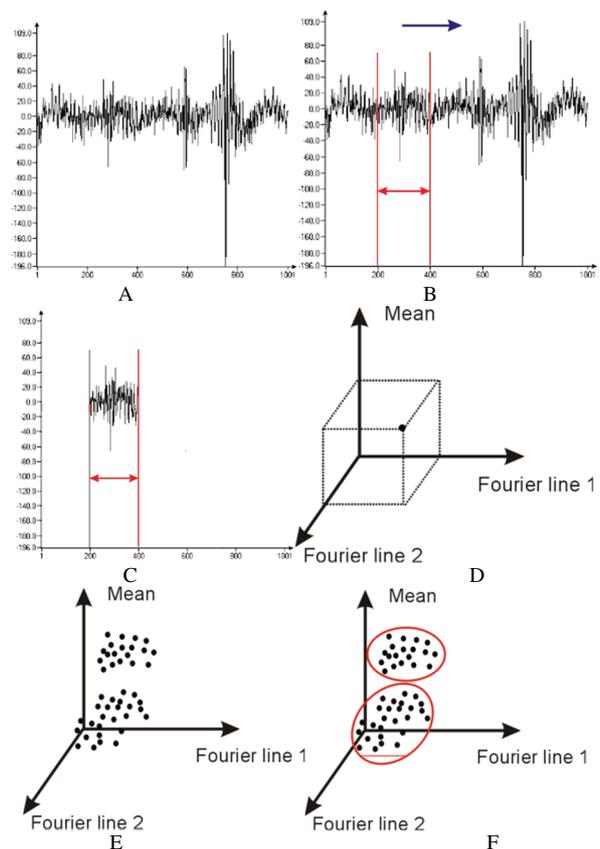


Fig. 2. For the vibration signal (A) the time window is chosen (B) and for the fragment treated as the short-signal (C) the chosen parameters are calculated. The set of parameters values calculated for the given time moment gives the point (vector) in the corresponding parameters space (D). For longer time interval in the parameter space the set of points is observed which represents the time evolution for parameter values (E). Eventually the distribution of points in the parameter space can divide into clusters which can represent different engine states

As a time moment usually a sampled time is taken into account. Many kinds of parameters were taken into consid-

eration: the successive coefficients (lines) of classical Fourier transformation FFT, statistical parameters like median, higher order moments and others. Of course some of them can appear not fully sufficient but from classification point of view the bigger number of parameters gives the higher dimension of the parameter space what make distinguishing between different clusters easier.

Having a point distribution in the parameter space some observation can distinguish between different engine states. To make the observation objective one can perform a clustering in the parameter space. The clustering is one of the most important techniques of pattern recognition. Assume n element data set X . The clustering algorithm finds the number c (lower than the number of elements n) and divide the data set X for c subsets, where elements are most similar to each other [4]. For calculations performed in the paper the classical hard c -means algorithm of clustering was applied. It can be obtained by optimization of the objective function in a form [4, 12]

$$J = \sum_{i=1}^n \sum_{j=1}^c p_{ij} E_{ij} \quad (2)$$

where energy E_{ij} is defined as an Euclidean distance $E_{ij} = |x_i - y_j|^2 = d_{ij}$ between the data point x_i and the centroid of cluster y_j and it is assumed that probabilities (membership values) p_{ij} that associates the data point x_i to cluster j are

$$p_{ij} = \begin{cases} 1 & \text{if } x_i \in \text{cluster}_j \\ 0 & \text{if } x_i \notin \text{cluster}_j \end{cases} \quad (3)$$

4. Results and discussion

The measurements were performed on two ST44 locomotives (number 2045 and 2061). For each signal the set of eleven parameters were calculated for each time moment: 1 FFT line, 2 FFT line, 3 FFT line, 4 FFT line, 5 FFT line, mean, moment 2, moment 3, moment 4, moment 5 and median. This set of eleven parameter values created a point in the corresponding parameter space.

Finally for all time moments one could observe the set of points in the parameter space. The evolution in the parameter space did not create a pattern which could distinguish between signals before and after a repair. Rather the irregular balls of puffs were found [7] (see Fig. 3). Therefore the only way to distinguish the signals was to recognize in which part of parameter space the point were placed. And the best parameter here is a center of points calculated using clustering algorithms.

Then for the set of points its center was calculated using the clustering algorithm HCM (classical hard c -means algorithm) [4, 12].

First of all the examinations which were taken into account tried to distinguish the results for signals registered before and after a repair which should represent the states of worn out and right working engine. The comparison was done for some special cases but at the beginning the researches on window size were performed to choose the width of the time window. The introductory researches on influence of window width on parameters values (see Table 1) did not show any interesting dependencies. Generally the values of parameters (especially FFT parameters) grow

while the window width rises. Eventually the window width was chosen arbitrary taking into account that the window width should include more than one characteristic periodic signal components. The eventual value of window width taken into calculation was 1000 samples. This time corresponds approximately to the rotation period of crankshaft.

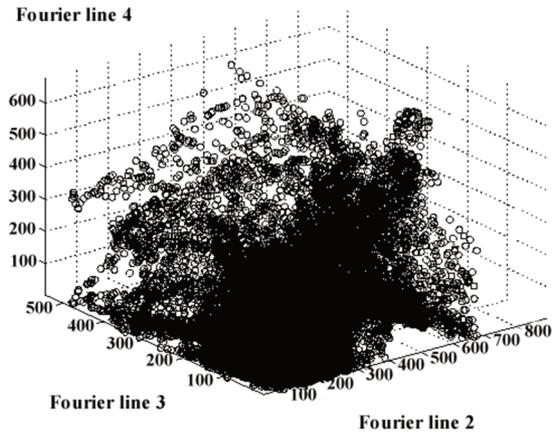


Fig. 3. An example view of a pattern created by points in a parameter space (find out that a dimension parameter space is 9 but at above figure there are only plot of three chosen parameters what can be visualized)

Values of FFT line represent the number of sample in frequency domain which corresponds to the given line. Taking into account that the sample frequency was $f = 1004$ Hz (approximately 1 kHz) it is easy to calculate the frequencies of coordinates for FFT calculation depending on number of samples. But in the approach presented in the paper the values of successive FFT lines are representing the successive coordinates in parameter space. Therefore the results in Hz are not need and the original data will be using by taking the sample frequency 1 „per unit”.

The centers for signals represent wear and good working engines (before and after a repair) were compared in all possible data sets:

- A – all signal,
- A2045 – signals for locomotive 2045,
- A2061 – signals for locomotive 2061,
- V – signals registered in vertical direction,
- V20145 – signals registered in vertical direction for locomotive 2045,
- V20161 – signals registered in vertical direction for locomotive 2061,
- H – signals registered in horizontal transversal direction,
- H2045 – signals registered in horizontal transversal direction for locomotive 2045,
- H2016 – signals registered for horizontal transversal direction for locomotive 2061.

Generally the relative units are used to describe all axis of parameter space. The comparison of the different clusters is performed by calculating the distances between cluster centers corresponding to the given engine states.

Table 1. Dependency of cluster center coordinates obtain using HCM algorithm on number of samples in a window (locomotive ST44-2045 after repair)

Number of samples	FFT 1	FFT 2	FFT 3	FFT 4	FFT 5	Mean	Mom 2	Mom 3	Mom 4	Mom 5	Median
10	7.5	10.4	11.3	15.6	13.9	0.0	20.3	2.9	1769.1	198.7	0.0
20	9.7	13.0	14.9	15.0	15.4	0.0	20.8	6.0	1873.9	1011.9	0.0
50	14.2	16.4	21.3	18.9	27.2	0.0	21.1	6.3	1933.4	916.7	0.0
100	19.4	22.2	23.5	25.6	32.8	0.0	21.2	6.2	1951.2	817.3	0.0
200	27.0	30.4	31.4	30.9	35.3	0.0	21.2	6.2	1959.6	805.9	0.0
500	42.4	48.0	48.5	46.6	48.9	0.0	21.2	6.2	1965.9	798.7	0.0
1000	59.1	66.7	67.1	72.2	68.5	0.0	21.2	6.2	1968.5	802.2	0.0
2000	84.0	92.2	95.8	101.3	92.2	0.0	21.2	6.4	1967.9	848.8	0.0
5000	138.2	134.9	150.8	133.6	153.8	0.0	21.2	6.4	1967.0	841.9	0.0
10000	185.6	216.7	182.5	194.8	233.9	0.0	21.2	6.4	1961.6	726.2	0.0

The Table 2 is showing the distances between the cluster centers corresponding to engine states.

It can be found that the biggest relative distances are for signals registered in horizontal transversal direction. More deep analysis is also showing that the comparison should be performed for signals from the same locomotive.

Table 2. The distances between the data sets for engine working before and after a repair

Signal data set	D
A	817521
A2045	935008
A2061	700295
V	4429
H	1630642
V2045	6402
V2061	2530
H2045	1863622
H2061	1398172

The last step was to perform a trial of a simple classification of an engine state basing on results of clustering in parameter space.

The classification in practice consists of building a classifier which can classify new data and can make a decision about their membership [4, 12, 16]. Taking into account the results from Table 2 only the data of horizontal measurements were taken into account in a classification. In this approach first a training set must be created and allocated. The training set was used just to train a system to distinguish between the engine states before and after a repair. The training set was chosen arbitral taking into consideration a first half of each data set. The applied classification method was classical linear discriminant function approach where a plane in a decision space is found to divide of data space on decision regions, which decide in assigning of an object to the given class. After a training stage, a testing stage was performed. On this stage the second subset of each data set was used to estimate a classification ability of

classifier. The results of the testing stage are presented in Table 3. The classification states in Table 3 are described as proper and improper, where proper means the engine working after the repair (right working engine) and improper means the engine working before the repair (worn out engine).

The general conclusion can be formulated as follow: the signals from horizontal measurements can be used to diagnosis of engine state making classification of vibroacoustic signals basing on short tome analysis and clustering in a parameter space.

5. Conclusion

The short-time analysis was performed to distinguish between different engine states which correspond to engine state before and after a repair. Eleven parameters were selected for each time moment what created the evolution in a multidimensional parameter space. In this space the clustering was performed. The obtained results seem quite promising. However, taking into account the great complexity and variety of possible measurement schemes it must be noted that the presented experiments can be treated only as the introductory stage of the research. We have used only signal examples taken from two locomotives. The set of using parameters seems reasonable but of course there is possibility of choosing higher number of them or taking another type of parameters. Calculated in sliding window parameters values give us the points whose coordinates evolve in the parameter space. After calculating of cluster centers the comparison is done by finding distances between the given cluster centers. Observing the parameter space we can find the massive center and some outside traces that can be treated as the improper behavior of an engine. The next and the last was a classification. The results of classification are not perfect but they are showing. The next step of analysis will be performing a classification with using of Support Vector Machine SVM method [1, 16] which seems quite effective in similar problems [10].

Table 5. The result of classification for data sets H – signals registered in horizontal transversal direction, H2045 – signals registered in horizontal transversal direction for locomotive 204 and H2016 – signals registered for horizontal transversal direction for locomotive 2061

	H		H2045		H2016	
	Percentage of classification as a		Percentage of classification as a		Percentage of classification as a	
	Proper	Improper	Proper	Improper	Proper	Improper
Proper (after the repair)	69%	31%	79%	21%	85%	15%
Improper (before the repair)	75%	25%	83%	17%	93%	7%

The future researches need to perform an experiments on bigger set of data what usually increase an expenses of researches.

It seems also that there are the significant differences for the data taken from different locomotives. This shows

that to classify and real diagnosis using vibroacoustic signal analysis all measurement must be taken from the same locomotive and for the same installation of injector and accelerometers also.

Nomenclature

A	all signals	H2061	signals registered for horizontal transversal direction for locomotive 2061
A2045	signals for locomotive 2045	OBD	on-board diagnostic system
A2061	signals for locomotive 2061	V	signals registered in vertical direction
FFT	Fast Fourier Transformation	V2045	signals registered in vertical direction for locomotive 2045
H	signals registered in horizontal transversal direction	V2061	signals registered in vertical direction for locomotive 2061
H2045	signals registered in horizontal transversal direction for locomotive 2045		

Bibliography

- [1] ABE, S. Support vector machines for pattern classification. *Springer-Verlag*. 2005.
- [2] ALLEN, J.B., RABINER. A unified approach to short-time Fourier analysis and synthesis. *Proceedings of the IEEE*. 1977, **65**, 1558-1564.
- [3] ALLEN, J.B. Short term spectral analysis, synthesis, and modification by Discrete Fourier Transform. *IEEE Transactions on Acoustic, Speech, and Signal Processing ASSP-25*. 1977, 235-238.
- [4] BEZDEK, J.C. Pattern recognition with fuzzy objective function algorithms. *Plenum Press*, Second edition, 1987.
- [5] BOGUŚ, P., LEWANDOWSKA, K. Short-time signal analysis using pattern recognition methods. *Artificial Intelligence and Soft Computing*. 2004, **3070**, 550-555.
- [6] BOGUŚ, P., MERKISZ, J. Short-time analysis of combustion engine vibroacoustic signals through pattern recognition techniques. *SAE Technical Paper*. 2005, 2005-01-2529.
- [7] BOGUŚ, P., MERKISZ, J. Misfire detection by short-time analysis with using clustering techniques. *Congress Proceedings, PTNSS KONGRES 2005*, September 25th – 28th, 2005, Bielsko-Biała/Szczyrk.
- [8] BOGUŚ, P., SIENICKI, A., WOJCIECHOWSKA, E., MERKISZ, J. The comparison of vibroacoustic signals taken from an engine before and after repair. *Combustion Engines*. 2007-SC3, 300-306.
- [9] BOGUŚ, P., SIENICKI, A., WOJCIECHOWSKA, E. Porównanie stanu silnika lokomotywy spalinowej ST44 przed i po remoncie przy użyciu sygnału wibroakustycznego. *Pojazdy Szynowe*. 2007, **2**, 28-36.
- [10] BOGUŚ, P., MERKISZ, J., MAZUREK, S. The prospects of artificial intelligence methods in identification and prevention of critical railway accidents. *L. Rutkowski, R. Tadeusiewicz, L. Zadeh, J. Zurada (eds.). Computational Intelligence: Methods and Applications*. EXIT, Warsaw 2008, 445-453.
- [11] BOGUŚ, P., GRZESZCZYK, R., WRONA, A. et al. Estimation of fuel spraying from diesel engine injector using multiresolution wavelet analysis of vibroacoustic signals. *Combustion Engines*. 2015, **162**(3), 264-270.
- [12] DUDA, R., HART, P. Pattern classification and scene analysis. New York, *Wiley Interscience* 1973.
- [13] HARRIS, F.J. On the Use of windows for harmonic analysis with the Discrete Fourier Transform. *Proceedings of the IEEE*. 1978, **66**, 51-83.
- [14] MERKISZ, J. Ecological aspects of combustion engines (Part 1 and 2). Poznań: *Wydawnictwo Politechniki Poznańskiej*. 1998 and 1999 (in Polish).
- [15] PORTNOFF, M.R. Time-frequency representation of digital signals and systems based on Short-Time Fourier Analysis. *IEEE Transactions on Acoustic, Speech, and Signal Processing*. 1980, **28**, 55-69.
- [16] WANG, L. (ed.). Support vector machines: theory and applications. *Springer-Verlag* 2005.
- [17] YANG, J., PU, L. et al. Fault detection in a diesel engine by analyzing the instantaneous angular speed. *Mechanical Systems and Signal Processing*. 2001, **15**, 549-564.

Prof. Piotr Boguś, DSc., PhD. – Rail Vehicle Institute TABOR in Poznań, Poland.
e-mail: piotr.bogus@gumed.edu.pl



Prof. Jerzy Merkisz, DSc., DEng. – Faculty of Transport Engineering, Poznan University of Technology.
e-mail: jerzy.merkisz@put.poznan.pl

