Development of bearing fault diagnostics methodology based on signal processing and machine learning tools

PhD Thesis

Krisztián Deák

Supervisor: Imre Kocsis PhD, dr. habil.

UNIVERSITY OF DEBRECEN

Doctoral Counsil of Natural Sciences and Information Technology

Doctoral School of Informatics

Debrecen, 2020

This thesis has been made during the Doctoral Program Industrial and Scientific Applications of Informatics Doctoral Counsil of Natural Sciences and Information Technology, University of Debrecen in order to obtain PhD degree at University of Debrecen.

Debrecen, 24th February, 2020.

candidate's signature

I certify that Krisztian Deák PhD candidate worked under my supervision from September 2013 to June 2016 as a PhD student of the Doctoral Counsil of Natural Sciences and Information Technology, Industrial and Scientific Applications of Informatics at Doctoral School of Informatics, Science and Informatics Doctoral Council, University of Debrecen. The creative activity of the candidate decisively contributed to the preparation of the results in this thesis. I recommend acceptance of the thesis.

Debrecen, 24th February, 2020.

supervisor's signature

Development of bearing fault diagnostics methodology based on signal processing and machine learning tools

Dissertation submitted in partial fulfilment of the requirements
for the doctoral (PhD) degree in Informatics

Written by: Deák Krisztián, certified mechanical engineer

This thesis has been made during the Doctoral Program Industrial and Scientific Applications of Informatics at Doctoral School of Informatics, Science and Informatics Doctoral Council, University of Debrecen in order to obtain PhD degree at University of Debrecen.

Dissertation advisor: Dr. Kocsis Imre

The comprehensive examination board:

chairperson:	Dr. Hajdu András	.....
members:	Dr. Vesselényi Tibor	.....
	Dr. Zsidai László	.....

The date of the comprehensive examination: 25th October, 2019.

The official opponents of the dissertation:

	Dr.	.....
	Dr.	.....
	Dr.	.....

The evaluation committee:

chairperson:	Dr.	.....
members:	Dr.	.....
	Dr.	.....
	Dr.	.....
	Dr.	.....

The date of the dissertation defence: …… 2020

Acknowledgement

I would like to express my deepest gratitude to my Supervisor, Dr. habil. Imre Kocsis PhD for the patient guidance, great help, encouragement and advices he has provided me. His patronage, mathematical and technical guidance was crucial to establish deep researches in this field.

I also would like to express my gratitude to the Head of the Doctoral School Prof. Dr. János Sztrik DSc to his help. I would like to thank the work all Members of the Committee. I wish my thank for the Chairman Prof. Dr. András Hajdu DSc. I thank the precise work and advices of the Reviewers Prof. Dr. László Tóth DSc and Dr. habil. László Zsidai PhD.

I also wish to say thank for the help of the Dean of the Faculty of Engineering Prof. Dr. Edit Szűcs PhD, the Vice-Dean of the Faculty of Engineering Dr. habil. Géza Husi PhD and the Head of the Department of Mechanical Engineering, Dr. Tamás Mankovits PhD.

I would like to thank the patience and the encouragement of my Family who supported me during the whole period of my studies.

Finally, I would like to say special thank for the work and guidance of those who contributed to my dissertation.

Contents

List of Acronyms

AE Acoustic Emission
ANN Artificial Neural Network
BP Back-propagation
BPFI Ball Pass Frequency Inner Race
BPFO Ball Pass Frequency Outer Race
BSF Ball Spin Frequency
CWT Continuous Wavelet Transform
DFT Discrete Fourier Transform
DWT Discrete Wavelet Transform
E/S Energy-to-Entropy ratio
FFT Fast Fourier transform
FTF Fundamental Train Frequency
GD Gradient descent
HFRT High Frequency Resonance Technique
HMM Hidden Markov Model
IDWT Inverse Discrete Wavelet Transform
IR Inner-ring
IRB Inner-ring back surface
IT Informatics
L-M Levenberg-Marquardt
MLP Multilayer percepton
MRA Multiresolution Analysis
MSE Mean Square Error
OR Outer-ring
PSD Power Spectral Density
QMF Quadrature Mirror Filter
R Roller
RF Roller fault
RMS Root Mean Square
SC Scalogram
SCG Scaled conjugate back propagation
SNR Signal-to-Noise Ratio
SC Scalogram
SP Spectrogram
SPM Shock Pulse Method
STFT Short-Time Fourier Transform
TSA Time Synchronous Average
WT Wavelet Transform

1 Introduction

Modern technical diagnostics plays an important role in maintenance engineering and the industrial production. It applies wide range of condition monitoring methods such as vibration diagnostics, acoustic emission measurements, thermography, endoscopy, chemical oil analysis which are integrated in Expert Systems using artificial intelligence methods.

Vibration analysis is crucial part of diagnostics that can be used for continuous as well as intermittent condition monitoring with the best sensitivity and immediate response to structure change. Vibration analysis refers to processing and analysing machine vibration signals obtained with vibration transducers like proximity probes, velocity pickups, accelerometers, shaft encoders and laser vibrometers. Vibration analysis is by far the most prevalent method for machine condition monitoring, especially for rotating equipment. Vibration signals contain a significant amount of information related to the machine components. This field of engineering originates in 1970's in the production industry answering the requirements of predictive maintenance and developed parallel to informatics. The diagnostics toolset was always determined by the actual capacity of the available IT tools. As an integration of data acquisition, data processing, process and equipment evaluation and automatic intervention to the processes, technical diagnostics can be considered as an important source of knowledge applied today in digital manufacturing in the framework of cyber physical systems as part of the Industry 4.0 revolution. Potential advantages of condition monitoring include increased machine availability and reliability, improved efficiency, reduced costs, extended operational life, and improved safety.

Certain types of faults create short time impulses, called transients and results in strong harmonics of the fault frequencies in the spectrum of vibration signals. These fault frequencies are sometimes masked by adjacent frequencies in the spectra due to their little energy. Fourier transform is useful in identifying periodic and harmonic components of signals but not for transient detection. To reduce drawbacks of traditional FFT analysis it is well advised to search for more precise and efficient methods such as wavelet transform.

Wavelets-based representations offer an important option for capturing localized effects in many signals. This is achieved by employing representations continuous transforms or discrete transforms which contain the scaling and shifting of a mother function. Over a period of several decades, wavelet analysis has been set on a rigorous mathematical framework and has been applied to

quite diverse fields. Wavelet families associated with specific mother functions have proven quite appropriate for a variety of problems. In this context, fast decomposition and reconstruction algorithms ensure computational efficiency over rival classical spectral analysis algorithms such as the fast Fourier transform. Wavelet transform provides powerful multi-resolution in time–frequency analysis for characterizing the transitory features of non-stationary signals. Moreover, wavelet transform can decompose an inspected signal into a family of elementary functions. This ability renders the analysis of the inspected signal easier for machine operators.

To make the decision process faster automatic fault classification is developed applying machine learning tools, namely Support Vector Machine (SVM) which is supervised learning model and artificial neural network (ANN) that are also efficient tools for bearing fault classification.

Methods based on IT tools can be integrated in an Expert System that can be used for bearing fault diagnostics in industrial applications as the part of Industry 4.0 activities.

1.1 Overview of researches in the field of bearing fault diagnostics

Literature overview focus on the relevant publications in connection with bearing fault diagnostics besides methods of signal analysis, feature extraction with wavelet analysis and automatic fault detection using machine learning techniques.

Patel et al. [1] used envelope methods to reveal local faults on the races of deep groove ball bearings. This paper deals with detection of local defects existing on races of deep groove ball bearing in the presence of external vibrations using envelope analysis and Duffing oscillator. Experiments have been carried out using a test rig for capturing the vibration signals of test bearing. The external vibration has been imparted to the housing of the test bearing through electromechanical shaker. In envelope analysis the centre frequency has been selected using the spectral kurtosis for the filters length of 32 and 64 for different bandwidths. The spectra of selected centre frequency with several bandwidths have been studied and compared for identification of defective frequency. The system defined by the Duffing equation entered into the periodical state from the chaotic state at the critical value of disturbing periodic force in the presence of defective bearing signal.

Kalman and H filter were applied by *Khanam et al.* [2] to measure bearing faults especially in noisy condition with low signal-to-noise ratio when it was difficult to identify the useful components of the vibration signal. This work

attempts to examine the scope of applying Kalman filter and H∞ filter individually on the vibration signal acquired for identifying local defects in a rolling element bearing. This is essentially a system dynamic approach, which is another choice, examined to be a better one in comparison with few other signal analysis approaches reported in the literature. Kalman and H∞ filters are optimal state estimators; Kalman filter is the minimum variance estimator while H_∞ filter minimizes the worst case estimation error. Experiments have been carried out to investigate the performance of Kalman and H_∞ filters each with the Envelope Analysis technique, a popular one for identification of bearing faults, in a noisy environment. Envelope Analysis is performed by taking a Hilbert transform of the band pass filtered signal, whose centre frequency and bandwidth are to be properly selected for satisfactory performance of the algorithm.

Acoustic emission measurement is a powerful method to detect cracks inside the bearing material which are the initial reasons of fatigue spalling. *Al-Ghamd and Mba* [3] applied this method combined with the traditional vibration analysis to determine the bearing outer race defect width directly from the raw signal. The experimental investigation reported in this paper was centred on the application of the AE technique for identifying the presence and size of a defect on a radially loaded bearing. An experimental test rig was designed such that defects of varying sizes could be seeded onto the outer race of a test bearing. Comparisons between AE and vibration analysis over a range of speed and load conditions are presented. In addition, the primary source of AE activity from seeded defects is investigated. It is concluded that AE offers earlier fault detection and improved identification capabilities than vibration analysis. Furthermore, the AE technique also provided an indication of the defect size, allowing the user to monitor the rate of degradation on the bearing; unachievable with vibration analysis.

Elforjani and Mba [4] emphasized the effectiveness of acoustic emission method in the case of slow-speed bearings. This paper presents results of an investigation to assess the potential of the Acoustic Emission(AE) technology for detecting and locating natural defects in rolling element bearings. To undertake this task a special purpose test-rig was built that allowed for accelerated natural degradation of a bearing race. It is concluded that sub-surface initiation and subsequent crack propagation can be detected using a range of data analysis techniques on AE's generated from natural degrading bearings. The paper also investigates the source characterisation of AE signals associated with a defective bearing whilst in operation. This study also attempted to identify the size of a natural defect on bearings using AE technology.

Sawalhi and Randall [5] researched to determine the fault size of the bearings from the vibration signal by analysing the entry and exit impulses. studies indicated that there might be two parts to the defect vibration signal of a spalled bearing, the first part being originating from the entry of the rolling element into the fault (de-stress) and the second part being due to the departure of the rolling element from the fault (re-stress). This is investigated in this paper using vibration signatures of seeded faults at different speeds. The acceleration signals resulting from the entry of the rolling element into the spall and exit from it were found to be of different natures. The entry into the fault can be described as a step response, with mainly low frequency content, while the impact excites a much broader frequency impulse response. The latter is the most noticeable and prominent event, especially when examining the high pass filtered response or the enveloped signal. In order to enable a clear separation of the two events, and produce an average estimate of the size of the fault, two approaches are proposed to enhance the entry event while keeping the impulse response.

Because of its flexibility and computational efficiency wavelets are perfect tools for fault feature extraction, singularity detection for signals, denoising and extraction of the weak signals from the vibration signals. These applications were presented by *Peng and Chu* [6]. The purpose of this review is to present a summary about the application of the wavelet in machine fault diagnostics, including the following main aspects: the time–frequency analysis of signals, the fault feature extraction, the singularity detection for signals, the denoising and extraction of the weak signals, the compression of vibration signals and the system identification.

Discrete wavelet transform with Daubechies-4 (db04) mother wavelets to analyse the combination of different faults on the races of ball bearings were used by *Prabhakar et al.* [7]. Vibration signals from ball bearings having single and multiple point defects on inner race, outer race and the combination faults have been considered for analysis. The impulses in vibration signals due to bearing faults are prominent in wavelet decompositions. It is found that the impulses appear periodically with a time period corresponding to characteristic defect frequencies. It has been shown that DWT can be used as an effective tool for detecting single and multiple faults in the ball bearings.

Combination of envelope spectrum and wavelet transform for extraction of defect problems in bearings were used by *Shi et al.* [8]. To overcome the shortcomings in the traditional envelope analysis in which manually specifying a resonant frequency band is required, a new approach based on the fusion of the wavelet transform and envelope spectrum is proposed for detecting and localizing defects in rolling element bearings. This approach is capable of completely extracting the characteristic frequencies related to the defect from

the resonant frequency band. Based on the Shannon entropy of wavelet-based envelope spectra, a criterion to select optimal scale to monitor the condition of bearings is also presented.

Nikolaou and Antoniadis [9] applied complex shifted Morlet wavelets to analysing vibration signals generated by rolling element bearings. The envelope detection or demodulation methods have been established as the dominant analysis methods for this purpose, since they can separate the useful part of the signal from its redundant contents. This paper proposes an effective demodulation method, based on the use of a complex shifted Morlet wavelet family. The method is designed in a way that can fully exploit the underlying physical concepts of the modulation mechanism, present in the vibration response of faulty bearings, using a time–frequency representation of the signal. A key element of the proposed method is the systematic introduction of selection criteria for the automated choice of the critical parameters that characterise the Morlet wavelet family used.

Qiu et al. [10] successfully used wavelet filter-based weak signature detection method and its application for diagnosis of rolling element bearings. In this paper, the performance of wavelet decomposition-based de-noising and wavelet filter-based de-noising methods are compared based on signals from mechanical defects. The comparison result reveals that wavelet filter is more suitable and reliable to detect a weak signature of mechanical impulse-like defect signals, whereas the wavelet decomposition de-noising method can achieve satisfactory results on smooth signal detection. In order to select optimal parameters for the wavelet filter, a two-step optimization process is proposed. Minimal Shannon entropy is used to optimize the Morlet wavelet shape factor. A periodicity detection method based on singular value decomposition (SVD) is used to choose the appropriate scale for the wavelet transform. The signal de-noising results from both simulated signals and experimental data are presented and both support the proposed method.

Junsheng et al. [11] pointed out the effectiveness of impulse response wavelet to the fault diagnosis of rolling element bearings. To target the characteristic of roller bearing fault vibration signals, the impulse response wavelet is constructed by using continuous wavelet transform to extract the feature of fault vibration signals, based on which two methods namely scale-wavelet power spectrum comparison and auto-correlation analysis of time-wavelet power spectrum are proposed. The analysis results from roller bearing vibration signals with outer race or inner race fault show that the two proposed methods can detect the faults of roller bearing and identify fault patterns successfully.

Symlet wavelets were used efficiently in the study of *Kumar and Singh* [12]. In their study tapered roller bearings were analysed to determine the fault size

on the outer ring. Symlet-5 wavelet represents the entry and impact events as the roller hits the defects during operation of the bearing. A detailed study was presented about the decomposition of the vibration signals using discrete wavelet transform with Symlet-5. Experiments were carried out on a customized test setup. Seeded defects of different size were introduced separately in the form of an axial groove on the outer race of taper roller bearings using laser engraving technique. It is not only difficult but ambiguous to detect the entry point in the groove defect by making use of the signal. The ambiguity gets reduced by using Symlet5 wavelet due to its linear phase nature which maintains sharpness in the signal even when there is a sudden change in signal.

Symlet wavelet is an effective tool for noise reduction because it could filter out the useful components of the complex signal from the noisy background [13]. Noise reduction is important for getting useful bio-medical signal such as, ECG signals. Because ECG signal can be corrupted by various types of noise which leads to incorrect diagnosis. However, details of the signal must be conserved very well after the de-noising for proper diagnosis. In this paper, Symlet wavelet filter as well as different kinds of moving average and S-G filter was studied for de-noising of ECG signal. R-peak automatically identifies via R-peak detection algorithm.

Analytical Wavelet Transform (AWT) based acoustic emission technique for identifying inner race of radial ball bearing were applied by *Kumar et al.* [14]. An acoustic emission technique based on Analytical Wavelet Transform (AWT) has been investigated to identify the presence and severity of defect in the inner race of bearing. The defect was in the form of an axial groove. The nature of AE burst and correlation between AE burst generated due to defect is estimated. Experimental investigations reveal that AWT is an effective tool in analysing the acoustic signal transmitted from bearing in order to characterize and find out the size of the defect. The results were verified from octave analysis.

Yan and Gao [15] revealed localized structural defects and made experiments in their studies using multi-scale enveloping spectrogram for vibration analysis of bearings. Compared to the conventional enveloping spectral analysis technique in which the bandwidth of the signal components of interest needs to be known a priori to obtain consistent results under varying machine operating conditions, the new technique enables simultaneous multi-scale decomposition to extract and separate envelopes of the repetitively excited mechanical vibrations with different frequency coverage, thus improving the robustness in signal analysis. Complex wavelet was investigated as the base wavelet on its ability in combining band-pass filtering and enveloping into a single-step operation. The new technique was experimentally evaluated using vibration

signals measured on rolling element bearings that contain localized structural defects.

Patil et al. [16] developed an analytical model and simulation to predict the effect of a localized defect on the ball bearing vibrations by considering the contact between the ball and the races as non-linear springs. In the analytical formulation, the contacts between the ball and the races are considered as non-linear springs. The contact force is calculated using the Hertzian contact deformation theory. A computer program is developed to simulate the defect on the raceways with the results presented in the time domain and frequency domain. The model yields both the frequency and the acceleration of vibration components of the bearing. The effect of the defect size and its location has been investigated.

Optimal wavelet filtering and sparse code shrinkage was presented by *He et al.* [17]. To extract the impulsive features buried in the vibration signal, a hybrid method which combines Morlet wavelet filter and sparse code shrinkage (SCS) is proposed. First, the wavelet filter is optimized by differential evolution (DE) to eliminate the interferential vibrations and obtain the fault characteristic signal. Then, to further enhance the impulsive features and suppress residual noise, SCS which is a soft-thresholding method based on maximum likelihood estimation (MLE) is applied to the filtered signal.

To extract the impulsive features buried in the vibration signal, a hybrid method which combines Morlet wavelet filter and sparse code shrinkage was proposed. Simulations and signal processing techniques to track the spall size after Time Synchronous Average (TSA) were used by *Sawalhi and Randall* [18]. The entry into the fault can be described as a step response, with mainly low frequency content, while the impact excites a much broader frequency impulse response. The latter is the most noticeable and prominent event, especially when examining the high pass filtered response or the enveloped signal. Cepstrum analysis can be used to give an average estimate of the spacing between the entry and impact events, but the latter can also be assessed by an arithmetic estimation of the mean and standard deviation of the event separation for a number of realizations.

Khanam et al. [19] estimated the fault size in the outer race of ball bearing using discrete wavelet transform of the vibration signal. This work presents the decomposition of the vibration signal by using discrete wavelet transform assisted by Symlet_5 wavelet. Symlet wavelet has a linear phase nature which maintains sharpness in the signal even when there is sudden change in the signal. The decomposed signal evidently splits the peak corresponding to the ball entry into and exit from the fault, enabling in an estimation of the defect size present in the bearing.

Tóth and Tóth [20] and [21] revealed artificial faults of the inner rings of deep groove bearings by wavelet analysis. A realistic signal model of ball bearings with inner race fault was created to design a new wavelet to reveal the defect more efficiently from the vibration signature. The paper considers the comparison of the Meyer and Morlet wavelet for bearing fault diagnosis. It was created a wavelet based upon a transient vibration signal model established for signals generated in deep-groove ball bearings with pitting (spalling) formulation on their inner race. The wavelet creation used the sub-optimal algorithm that matches a Meyer wavelet to a band limited signal in two steps. Morlet wavelet was used as a benchmark for evaluating the performance of the matched wavelet.

Borghesani et al. [22] applied cepstrum pre-whitening for diagnostics of rolling element bearings. Due to its moderated computational requirements it was an appropriate tool for automatic damage recognition algorithm. A comparison with the traditional pre-whitening techniques revealed that cepstrum pre-whitening was more suitable and efficient tool for automatic fault detection. a new technique for pre-whitening has been proposed, based on cepstral analysis: the so-called cepstrum pre-whitening. Owing to its low computational requirements and its simplicity, it seems a good candidate to perform the intermediate pre-whitening step in an automatic damage recognition algorithm. In this paper, the effectiveness of the new technique will be tested on the data measured on a full-scale industrial bearing test-rig, able to reproduce the harsh conditions of operation.

Figlus and Stanczyk [23] presented a method of diagnosing damage of rolling bearings in the vicinity of toothed gears of processing lines. Vibration response was measured by laser vibrometer. Discrete wavelet transform was successfully applied to detect damages. The research has shown the usefulness of vibration signal measurements performed with a laser vibrometer and of the method of denoising signals by means of a discrete wavelet transform in detecting damage to bearings. The application of the method of analysis of the characteristic frequencies of changes in the vibration signal amplitude made it possible to draw conclusions about the type of damage to the bearings.

Slavič et al. [24] used force measurement instead of the traditional acceleration measurement to identify bearing faults. Signal was processed by envelope technique. The research showed that frequency domain analysis could successfully applied to identify both amplitude and frequency of the force signal. The procedure was applied to high-series production line as well. The research discusses the use of force measurements to identify bearing faults. A force sensor is fixed between the rigid surroundings and the bearing to measure all of the reactive forces due to the vibration excitation. Using a force measurement, systematically prepared samples with the five typical faults that can occur during

the assembly process (axial, radial, bending moment, contamination and shield defect) were investigated. The samples were prepared with low, medium and high fault ratings.

Abboud et al. [25] characterized bearing fault vibrations and explored angle/time cyclostationary property. They experimentally validated their results on real vibration signals and the possible application for bearing fault detection. the assumption of cyclostationarity of rolling-element bearing vibrations is jeopardized. The emitted signal comprises an interaction between time-dependent components related to the system dynamics and angle-dependent mechanisms related to the system kinematics

Antoni [26] applied cyclic spectral tool for incipient fault diagnosis of rolling element bearings. They demonstrated the optimality of cyclic coherence. It was proved that the diagnostic information is perfectly preserved in the cyclic frequency domain as a symptomatic pattern of spectral lines. The paper discusses which cyclic spectral tools should be considered for that purpose. Specifically, it demonstrates the optimality of the cyclic coherence, which can not only evidence the presence of a fault in high levels of background noise, but can also return a relative measure of its severity. The estimation issue of the cyclic coherence is addressed in detail, as well as its use in a statistical test, and sub-optimal simplifications.

Tewfik et al. [27] worked out a design method that matches a wavelet to the time domain form of a signal. Two techniques for finding the discrete orthogonal wavelet of support less than or equal to some given integer that leads to the best approximation to a given finite support signal up to a desired scale are presented. The techniques are based on optimizing certain cost functions. The first technique consists of minimizing an upper bound that is derived on the L norm of error in approximating the signal up to the desired scale. The second technique is based on maximizing an approximation to the norm of the projection of the signal on the space spanned by translates and dilates of the analysing discrete orthogonal wavelet up to the desired scale.

Chapa and Rao [28] developed an algorithm that searches for a matching wavelet in frequency domain. Their method is capable of designing Meyer wavelets that approximate the wavelet amplitude and phase spectra separately. The cost function is the minima of the Mean Squared Error (MSE), calculated from the amplitude spectra and group delay of the signal and the wavelet. In this paper, two sets of equations are developed that allow us to design the wavelet directly from the signal of interest. Both sets impose band-limited feature, resulting in closed form solutions. The first set derives expressions for continuous matched wavelet spectrum amplitudes. The second set of equations

provides a direct discrete algorithm for calculating close approximations to the optimal complex wavelet spectrum.

Mallat [29] created algorithm to analyse bearing fault signal and make the correlation signal to the original parameters to reveal the defects more efficiently.

In bearing diagnostics several optimization methods are applied successfully. Among them the Support Vector Machines and Neural Networks used in the following contributions.

Saravanan et al. [30] measured gear tooth breakage, gear with crack at root with face wear by Morlet wavelet. Statistical features standard deviation, variance, kurtosis, range were added to SVM to achieve bearing fault classification. This paper deals with the effectiveness of wavelet-based features for fault diagnosis using support vector machines (SVM) and proximal support vector machines (PSVM). The statistical feature vectors from Morlet wavelet coefficients are classified using J48 algorithm and the predominant features were fed as input for training and testing SVM and PSVM and their relative efficiency in classifying the faults in the bevel gear box was compared.

Rafiee et al. [31] used machine learning for defects classification of ball, cage, outer race, and on gears with 324 mother wavelets from various wavelet families like Haar, Symlet, Daubechies, Morlet, Gaussian. Time domain parameters such as variance, kurtosis, and skewness were fed to artificial neural network. They recommended Daubechies_44 wavelets. Four statistical features were selected: standard deviation, variance, kurtosis, and fourth central moment of continuous wavelet coefficients of synchronized vibration signals (CWC-SVS). In this research, the mother wavelet selection is broadly discussed. An automatic feature extraction algorithm is introduced for gear and bearing defects. It also shows that the fourth central moment of CWC-SVS is a proper feature for both bearing and gear failure diagnosis. Standard deviation and variance of CWC-SVS demonstrated more appropriate outcome for bearings than gears.

Kankar et al. [32] [33] [34] [35] combined SVM fault classification with wavelet transform for fault detection of ball bearings in the researches. The papers present a feature-recognition system for rolling element bearings fault diagnosis, which utilizes cyclic autocorrelation of raw vibration signals. Cyclostationary analysis of non-stationary signals clearly indicates the appearance of several distinct modulating frequencies. The coefficients of wavelet transform are calculated using six different base wavelets, after calculating cyclic autocorrelation of vibration signals. The base wavelet that maximizes the Energy to Shannon Entropy ratio is selected to extract statistical features from wavelet coefficients. Finally, a comparative study is carried out with the calculated statistical features as input to soft computing techniques.

Three soft computing techniques are used for faults classifications, out of which two are supervised techniques i.e. Support vector machine, Artificial Neural Network and other one is an unsupervised technique i.e. Self-Organizing Maps. The Complex Gaussian wavelet is selected based on maximum Energy to Shannon Entropy ratio. A comparative experimental study of the effectiveness of ANN and SVM is carried out. The results show that the machine learning algorithms mentioned above can be used for automated diagnosis of bearing faults.

Purushotham et al. [36] investigated the complex cepstrum of rolling element bearings with Daubechies wavelets. Bearing race faults have been detected by using discrete wavelet transform (DWT). Vibration signals from ball bearings having single and multiple point defects on inner race, outer race, ball fault and combination of these faults have been considered for analysis. Wavelet transform provides a variable resolution time–frequency distribution from which periodic structural ringing due to repetitive force impulses, generated upon the passing of each rolling element over the defect, are detected. It is found that the impulses appear periodically with a time period corresponding to characteristic defect frequencies. In this study, the diagnoses of ball bearing race faults have been investigated using wavelet transform. These results are compared with feature extraction data and results from spectrum analysis. It has been clearly shown that DWT can be used as an effective tool for detecting single and multiple faults in ball bearings. This paper also presents a new method of pattern recognition for bearing fault monitoring using hidden Markov Models (HMMs).

Honghu et al. [37] investigated the bearing faults by using one-dimensional convolutional neural networks and long-short term memory recurrent neural networks methods by using test-rig for creating efficient training results.

Zhuang Li et al. [38] applied wavelet transform with artificial neural network for the diagnosis of gearboxes. A new approach using discrete wavelet transform and an adaptive resonance theory neural network for crack fault detection of a gearbox is proposed. With the use of a multi-resolution analytical property of the discrete wavelet transform, the signals are decomposed into a series of sub-bands. The changes of sub-band energy are thought to be caused by the crack fault. Therefore, the relative wavelet energy is proposed as a feature. An artificial neural network is introduced for the detection of crack faults. Due to differences in operating environments, it is difficult to acquire typical, known samples of such faults. An adaptive resonance theory neural network is proposed in order to recognize the changing trend of crack faults without known samples on the basis of extracting the relative wavelet energy as an input eigenvector. The

proposed method is applied to the vibration signals collected from a gearbox to diagnose a gear crack fault.

Strączkiewicz et al. [39] used supervised and unsupervised learning as pattern recognition methods for damage classification and clustering of rolling bearings. Clustering analysis was found to be effective for determining the bearing state conditions. The investigation is performed on the data collected using an experimental test grid and rolling element bearing with deteriorating condition of an outer race.

Gligorijevic et al. [40] presented an automated technique for early fault detection of rolling element bearings by dividing the signal to sub-bands by means of wavelet decomposition. An automated technique for early fault detection and diagnosis in rolling-element bearings based on vibration signal analysis. Following the wavelet decomposition of vibration signals into a few sub-bands of interest, the standard deviation of obtained wavelet coefficients is extracted as a representative feature. Then, the feature space dimension is optimally reduced to two using scatter matrices. In the reduced two-dimensional feature space, the fault detection and diagnosis is carried out by quadratic classifiers. Accuracy of the technique has been tested on four classes of the recorded vibrations signals, i.e., normal, with the fault of inner race, outer race, and ball operation. The overall accuracy of 98.9% has been achieved.

Strączkiewicz et al. [41] applied supervised and unsupervised pattern recognition methods for damage classification and clustering of rolling bearings. Clustering analysis was effective for determining the number of bearing state conditions. The paper selected supervised and unsupervised pattern recognition methods are employed for this purpose. The attention of the authors is given to assessment of selection, performance benchmarking and applicability of selected pattern recognition methods. The investigation is performed on the data collected using an experimental test grid and rolling element bearing with deteriorating condition of an outer race.

Paya et al. [42] analysed drive-lines with multiple faults. The paper presents the investigation carried out in order to study both bearing and gear faults introduced first separately as a single fault and then together as multiple faults to the drive-line. The real time domain vibration signals obtained for the drive-line were pre-processed by wavelet transforms for the neural network to perform fault detection and identify the exact kinds of fault occurring in the model drive-line. It is shown that by using multilayer artificial neural networks on the sets of pre-processed data by wavelet transforms, single and multiple faults were successfully detected and classified into distinct groups.

Wang et al. [43] applied dynamic wavelet neural networks for fault diagnostics. This paper attempts to address this challenging problem with

intelligence-oriented techniques, specifically dynamic wavelet neural networks (DWNNs). DWNNs incorporate temporal information and storage capacity into their functionality so that they can predict into the future, carrying out fault prognostic tasks. Such fundamental issues as the network structure, learning algorithms, stability analysis, uncertainty management, and performance assessment are studied in a theoretical framework. An example is presented in which a trained DWNN successfully prognoses a defective bearing with a crack in its inner race.

Ray et al. [44] used wavelet coefficients with genetic algorithms for automatic feature extraction in machine fault analysis. A method is presented that can be used for automatic extraction of high quality features from wavelet coefficients without a priori knowledge of features. Pre-processing of the wavelet coefficients is necessary to obtain a measurable set of features. The pre-processing is suitable for the Morlet wavelet. Genetic algorithms are used in combination with learning vector quantization neural networks to select the relevant features from the processed wavelet coefficients. A simple variation of the traditional feature selection genetic algorithms is used as it applies to this method. The method has been applied on different signals for classification and has shown high classification rates with a small number of features. Results from different signal classification problems are also presented.

Chen et al. [45] produced the classification of wavelet map patterns using multi-layer neural networks for gear fault detection. A multi-layer perceptron pattern classifier is defined for wavelet map interpretation and its application is described as a tool for mechanical fault detection. As a key step, an instantaneous scale distribution is introduced for quantifying pattern features. Instead of directly inspecting complicated wavelet patterns in time–scale domains with limited human experience and availability, automated classification of the localised features related to gear faults, therefore, can be implemented. The detail of constructing, training and testing the multi-layer perceptron based classifier has been described with application to a gearbox.

Nikolau et al. [46] used demodulation of vibration signals which are generated by rolling bearing element defects. The envelope detection or demodulation methods have been established as the dominant analysis methods for this purpose, since they can separate the useful part of the signal from its redundant contents. The paper proposes an effective demodulation method, based on the use of a complex shifted Morlet wavelet family. The method is designed in a way that can fully exploit the underlying physical concepts of the modulation mechanism, present in the vibration response of faulty bearings, using a time–frequency representation of the signal. A key element of the proposed method is

the systematic introduction of selection criteria for the automated choice of the critical parameters that characterise the Morlet wavelet family used.

Antoniadis et al. [47] combined the physical model with Support Vector Machines for rolling bearing analysis in industrial environment. A hybrid two stage one-against-all Support Vector Machine (SVM) approach is proposed for the automated diagnosis of defective rolling element bearings. The basic concept and major advantage of the method, is that its training can be performed using simulation data, which result from a well-established model, describing the dynamic response of defective rolling element bearings. Then, vibration measurements, resulting from the machine under condition monitoring, can be imported and processed directly by the already trained SVM, eliminating thus the need of training the SVM with experimental data of the specific defective bearing. A key aspect of the method is the data pre-processing approach, which among others, includes order analysis, in order to overcome problems related to sudden changes of the shaft rotating speed. Moreover, frequency domain features both from the raw signal as well as from the demodulated signal are used as inputs to the SVM classifiers for a two-stage recognition and classification procedure. At the first stage, a SVM classifier separates the normal condition signals from the faulty signals. At the second stage, a SVM classifier recognizes and categorizes the type of the fault. The effectiveness of the method tested in one literature established experimental test case and in three different industrial test cases, including a total number of 34 measurements.

Wensheng et al. [48] used optimal Morlet wavelet filter and autocorrelation enhancement for the diagnosis of rolling element bearing faults. A hybrid method based on optimal Morlet wavelet filter and autocorrelation enhancement is presented. First, to eliminate the frequency associated with interferential vibrations, the vibration signal is filtered with a band-pass filter determined by a Morlet wavelet whose parameters are optimized by genetic algorithm. Then, to further reduce the residual in-band noise and highlight the periodic impulsive feature, an autocorrelation enhancement algorithm is applied to the filtered signal. In the enhanced autocorrelation envelope power spectrum, only several single spectrum lines would be left, which is very simple for operator to identify the bearing fault type. Moreover, the proposed method can be conducted in an almost automatic way. The results obtained from simulated and practical experiments prove that the proposed method is very effective for bearing faults diagnosis.

Manpreet et al. [49] investigated thrust bearing groove race defects with accurate measurements by wavelet decomposition of pre-processed vibration signal. In this work measurement of groove race defect width in thrust bearing is presented by processing vibration signal using wavelet based technique. In the

first step, pre-processing of vibration signal is done by multiplying the signal amplitude with its own absolute values. It helps to overcome the problem of detection of weak burst in normal vibration signal retaining sign of the original signal. In the second step, Symlet5 wavelet based decomposition is applied to the pre-processed signal to measure the size of seeded defect in outer groove race of thrust bearing. The decomposition based on Symlet wavelet is preferred because it reduces the ambiguity in detecting the entry point and exit point of the defect as it is having linear phase in nature. Measurement has been done for the groove defect width over the range of 0.4399–1.4854 mm.

1.2 Conclusions of the literature overview

Wavelet transform was successfully applied for the detection of sudden changes in the signal and reveal singularity. Continuous and discrete wavelets are effective tools for feature extraction from the vibration signal.

Daubechies, Symlet and Morlet wavelets were applied for bearing diagnostics in some related articles.

Since the center frequency and the bandwidth of complex Morlet wavelet could be adjusted in flexible way, it provides better correlation to the measured signal.

Researchers applied the Energy-to-Shannon Entropy Criteria for ranking the wavelets which were the more efficients for a certain application.

In the relevant literature a wide range of optimization and machine learning procedures were applied for general bearing diagnostics.

As for critism of the researches, the conclusions that the related papers mostly focused on ball bearings and the roller bearings are not completely analysed however they have significant role in mechanical engineering. Moreover, the operational and wear defects were the main areas of the investigations, while the manufacturing bearing faults were not discussed in the scientific papers. Grinding faults were not examined at all especially in the case of bearings which had rectangular shape contact areas.

In most cases, outer ring defects of roller bearings were analysed and there were some researches which investigated the inner ring and roller problems. Although, the back-surface defects of the rollers could cause serious problem under high-speed and high axial load applications, however there were no literatures which examined them from this view.

Besides of the wear defects, artificial faults and spalls were investigated which were created by artificial way of electro discharging machines or laser beam technology but grinding defects from manufacturing technologies were not investigated.

1.3 Purposes of my research

Detecting of the manufacturing faults are very important be able to control the manufacturing process. Today methods, which are frequently applied by the manufacturing companies are able to reveal the most common problems of the bearings e.g. waviness problems, roller defects, ring defects. However, there are bearing problems which can not be detected efficiently with the conventional time domain and frequency domain methods because they cause short-time transients in the vibration signature and they have their unique manner and morphology. Small size rectangular shape surface defects are very difficult to reveal especially if they are on the back-surface of the roller in tapered roller bearings under noisy conditions. Therefore, my main purpose is to develop new reliable bearing fault diagnostics methods based on vibration signal processing for detecting rectangular shape grinding faults related to bearings manufacturing.

My further purpose is to analyse the entry and exit events of the roller and fault interaction and determine a geometrical range where the method is applicable for accurate fault size estimation.

It is needed to realize an object-oriented measurement system which is suitable to achieve precise measurements to reveal the bearing defects efficiently. The system would contain a designed test-rig, vibration transducer, data acquisition unit, algorithms for fault detection and fault classification and software programs. In my concept, the test-rig would be flexible to analyze roller bearings of different types and sizes.

My purpose is to find methods to efficiently reveal the useful information content from the noisy signal different methods are tested and compared which are able to detect the unique bearing problems under low signal-to-noise ratio (SNR<30 dB) environment. It is necessary to create a signal model for the analysis of different size and types of bearings with similar rectangular shape grinding problem. In this way the transient wave is analysed which is caused by faults with similar morphology.

My goal is to test different wavelets which are determined as efficient ones for bearing fault diagnostics and analyse how efficient they are for the actual manufacturing grinding process fault detection. To choose the best efficient wavelets the Maximum Energy-to-Shannon Entropy ratio principle is used. Furthermore, I would like to design new continuous and discrete wavelets which are better for the feature extraction of the grinding defects than the customized wavelets. Complex Morlet wavelet is preferred because of its flexibility to design for a problem. As for the discrete wavelet design, I plan to apply the wavelet FIR

filters using Parks McClellan and Chapa Rao algorithm to determine the ideal filter coefficients which ensure the best feature extraction.

My purpose is to validate the measurement data with precise contact devices and optical devices to check the differences between the calculated and the measured values and prove the efficiency of the applied methods.

Furthermore, my aim is to execute automatic fault classification. Using the proper feature vectors it possible to classify the bearing not only as healthy or faulty but in multi-classes with high efficiency above 90% which is sufficient level according to the related scientific papers. I plan to compare different machine learning techniques and determine the best efficient for classifying the special grinding faults based on SVM and ANN because they seem to be effective methods in the machine learning theory. As for the input parameters I test the efficiency of time-domain statistical parameters of the vibration signal to teach the machine learning systems. My plan is to connect the statistical data set with wavelet multiresolution analysis (MRA) to fed the networks with the data set after MRA. It is supposed that statistical data input vectors after MRA provide higher efficiency rates than if the statistical parameters are used is raw form for teaching the machine learning systems. Here, my further purpose is to analyse different training algorithms, make comparisons and tests, determine the proper training parameters for the machine learning methods with global or stochastic optimization methods.

All of these methods are integrated in the object-oriented measurement method and system which is developed for the experiments.

2 Manufacturing faults of roller bearings and methods of the analysis

A bearing is a machine element that constrains relative motion to only the desired motion, and reduces friction between moving parts.

There are many types of roller bearings: cylinder, tapered, thrust, spherical roller bearings. All of these type of bearings have similar contact area which can be defined as “line contact” as opposed to the “point contact”. Because of this similarity of the contact area between the roller and the ring they can be analysed with the same mechanical principles.

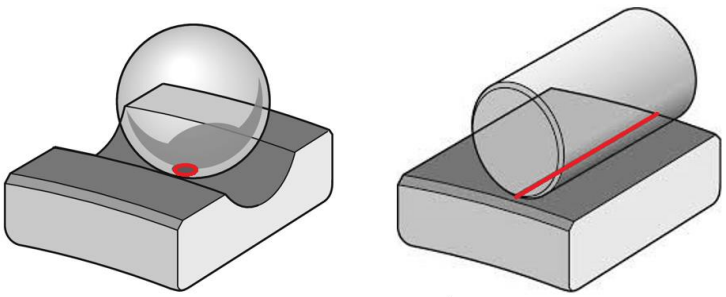


Figure 1. Morphological shape of contact areas: point contact (left) and line contact (right) [84]

Balls and rollers are different in how they make contact with the raceways. Balls make point contact with the ring raceways (Figure 1). With increasing load acting on the bearing, the contact point becomes an elliptical area. The small contact area provides low rolling friction, which enables ball bearings to accommodate high speeds but also limits their load-carrying capacity.

Rollers make line contact with the ring raceways. With increasing load acting on the bearing, the contact line becomes rectangular in shape. Because of the larger contact area and the consequently higher friction, a roller bearing can accommodate heavier loads, but lower speeds, than a same-sized ball bearing.

Tapered roller bearings are bearings that can take large axial forces as well as being able to sustain large radial forces. The inner and outer ring raceways are segments of cones and the rollers are also made with a taper so that the conical surfaces of the raceways and the roller axes if projected, would all meet at a common point on the main axis of the bearing. This geometry makes the motion of the cones remain coaxial, eliminating sliding motion in the bearing. Tapered roller bearings are separable and have the following components: outer ring, inner ring, and roller assembly containing the rollers and a cage (Figure 2). The non-separable inner ring and roller assembly is called the cone, and the outer

ring is called the cup. Internal clearance is established during mounting by the axial position of the cone relative to the cup.

Figure 2. Roller bearings: cylinder (left), tapered (middle), thrust (right)

In many applications tapered roller bearings are used in back-to-back pairs so that axial forces can be supported equally in either direction. Pairs of tapered roller bearings are used in car and vehicle wheel bearings where they must cope simultaneously with large vertical (radial) and horizontal (axial) forces. Applications for tapered roller bearings are commonly used for moderate speed, heavy duty applications where durability is required. Common real-world applications are in cutting machines, construction and mining equipment, axle systems, gear box, engine motors and reducers.

2.1 Vibrations generated in bearings

Bearing defects generate shock pulses when the rollers pass through the defects. Shock pulses spread inside the material towards the vibration transducer that measures the vibration parameter.

Figure 3 represents the spreading of the shock waves caused by bearing fault. The attenuation of the materials sufficiently influences the energy level of the waves that reach the vibration transducer. Therefore, it is important to use sensitive vibration sensor and to filter the noise, to surpass the harmful structural vibrations of the test-rig with adequate design. In this way even the small size defects can be revealed.

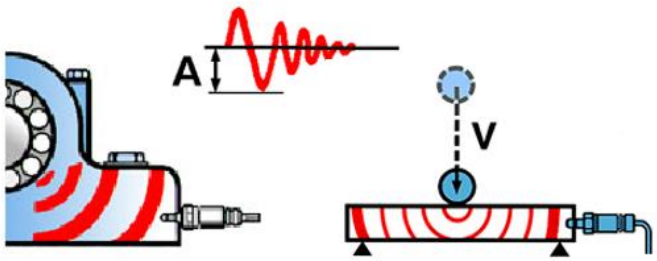


Figure 3. Shock waves in bearing caused by defects [74]

The faults cause certain fault frequencies determined by the bearing geometry and the rotational speed of the shaft. Basically, four types of them are distinguished: bearing pass frequency of outer race (BPFO), bearing pass frequency of inner race (BPFI), fundamental train frequency (FTF), ball spin frequency (BSF) which can be calculated by numerical way. [5][7]

$$BPFO = \frac{n \cdot f_r}{2}\left\{1 - \frac{d}{D}cos\Phi\right\} \qquad BPFI = \frac{n \cdot f_r}{2}\left\{1 + \frac{d}{D}cos\Phi\right\}$$

$$FTF = \frac{f_r}{2}\left\{1 - \frac{d}{D}cos\Phi\right\} \qquad BSF = \frac{D \cdot f_r}{2d}\left\{1 - \left(\frac{d}{D}cos\Phi\right)^2\right\}$$

where f_r is the speed of the shaft, n is the number of rolling elements, φ is the contact angle, d is the ball diameter, D is the pitch diameter.

2.2 Typical bearing faults

Bearing defects are caused by several factors such as material, environment, human factors, machine, technology. Defects can be classified in three groups as follows.

Operational defects: Wear, impact marks, smearing, spalls, fatigue cracks, corrosion, electric faults, contamination.

Machine problems: The proper condition of the manufacturing machines is necessary for the bearing companies to achieve the acceptable level of asset reliability, productivity and profit. There are several factors that could describe the machine health. In the bearing manufacturing process, the final surface quality of bearing elements is determined by the accuracy of grinding which is very sensitive to machine vibrations and environmental noises. The grinding problems are generally due to the grinding wheel wear process and misalignment of the wheel. The proper angle to the bearing parts is necessary. One of the critical items in the grinding machine is the high-speed integrated spindle with built-in motor (Figure 4). Maintenance of these spindles are important, the position of rotors, stators, and shafts has to be checked on regular basis. Since the grinding wheels are placed on the tool interface, it is important to keep the standard shape and diameter of the grinding wheels to avoid the faulty products. Spindle condition monitoring and machine vibration monitoring get important role nowadays. Using vibration transducers to monitor the spindle, certain parts of the machine problems are avoided. However, according to my research, not all problems can be solved in this way, therefore machine problems are always badly influence the quality of the product.

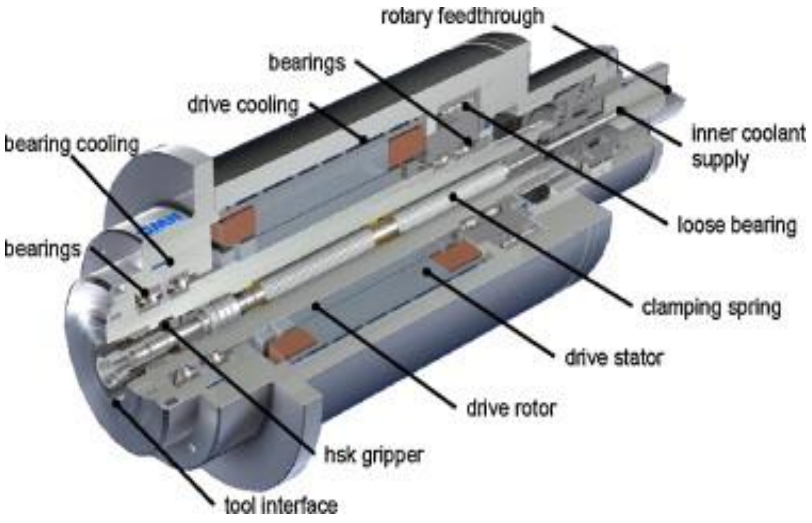


Figure 4. High-speed integrated spindle with built-in motor [82]

Technological defects: material problems, cracks inside the material, grinding problems, pulling difficulties, improper handling of the bearing parts, waviness problems.

Grinding is the last technological step and it has crucial role in the determination of the surface quality. High-speed grinding technology rotational speed of $45000 - 60000\ 1/min$ is necessary to ensure right morphology of outer rings, inner rings, rollers. It is a widespread production process and has long been a fixed part of almost every industrial production environment. During grinding, the bearing parts are literally given the final surface quality. Grinding operations contribute greatly to the quality of the finished product. However, problems can often occur during the grinding process.

One of the typical problem is the grinding burn, which is thermal damage to the rim zone of the part. Grinding burn occurs when too much heat is channelled into the part. Micro cracks and brittle surfaces are often the result. Therefore, it is important to apply sufficient amount cooling liquid for the technological procedure. To ensure that the coolant lubricant optimally reaches the machining zone, a precise ratio between the rotational speed of the grinding wheel and the exit speed of the coolant lubricant from the nozzle needs to be set. Once this operating point has been found, the machining task runs in an optimum way.

If a grinding wheel is highly loaded, the pore spaces become clogged. This means that the grinding wheel can then no longer transport any coolant lubricant and removed chips are no longer led away reliably from the machining site. In such a case, cleaning nozzles are recommended in order to make optimum use of the grinding wheel. Cleaning nozzles are nozzles which have been designed

especially for removing machining residues from the grinding wheel with the aid of coolant lubricant.

Other sources of technological problems can be improper technological parameters, faults of the CNC program.

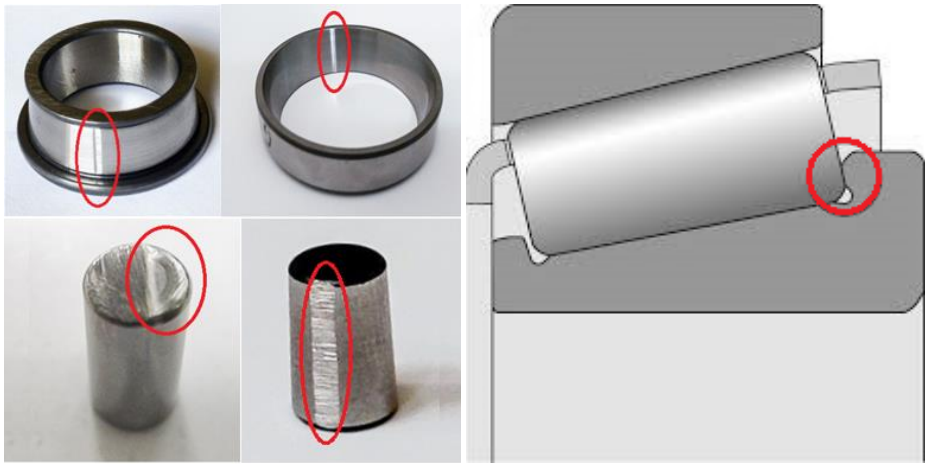


Figure 5. Outer ring (OR) fault, inner ring (IR) fault, roller (R) fault and inner-ring back surface (IRB) fault as linear rectangular shape topology grinding faults [84]

There are defects which has similar shapes, typically linear, longitudinal line defects. Certain type of grinding defects are cracks, pulling traces, false brinelling, etching pits. Some of them are coming from manufacturing, other from operation. The result of the process is mostly longitudinal (linear) shape, narrow fault with well-defined entry and exit contact points with the rollers.

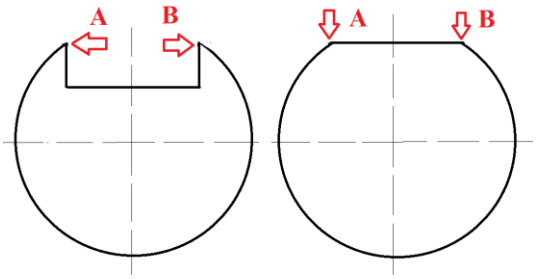


Figure 6. Roller defect model of crack (left) and grinding fault on the roller (right) with A and B interaction event contact points

There are other technological faults which could produce similar transients in the vibration signal than the grinding fault. They are coming from the pulling technology, cracks, false brinelling, etching pits as Figure 7 demonstrates. However, it is not allowed to say that the transient wave is the same in all cases. The shape of a grinding fault is not exactly the same as Figure 6. demonstrates the visual difference in case of a roller. However, both defect have linear,

rectangular shape the formation of the fault is different as Figure 6 presents. That is the situation in case of the rings too. It is only an assumption that similar fault types produce a similar transient wave with well-defined entry and exit contact points between the roller and the fault. It is necessary to analyse the transient signal model of the grinding fault and analyse the entry and exit interaction events between the roller and the fault for the fault detection and the fault size estimation.

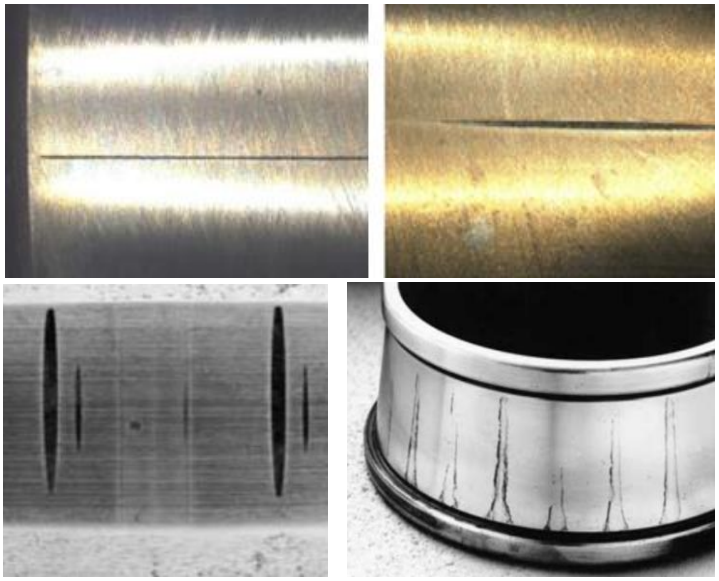


Figure 7. Other types of linear shape defects: cracks (left above), pulling defects (right above) and false brinelling (left down) and etching pits (right down) of similar morphology of linear defects [85]

Adhesive wear and fatigue damages can cause distributed surface damages (Figure 8). These failures have different morphological shape that the above mentioned linear shape defects, thus they are producing other kind of vibration signatures, especially the adhesive wear and the fatigue damage which are distributed on the raceway. Distributed defects have no direct edges that the roller can hit during operation and no direct entry and exit points defined which cause the short time transient signature in the vibration spectrum.

Therefore, it is important to classify the surface problems, referring to their shapes. In the Thesis, it is a focus on the longitudinal, linear and rectangular shape morphological defects which produce a clear entry and exit events so that the defect width can be measured precisely. The parameters of the transients are analysed in Chapter 6. such as period, duration, duty cycle, undershoot, overshoot, slew rate. In this way, and producing a transient signal model, the analysis method of the grinding faults of tapered roller bearings can be

generalized to be applicable to investigate similar surface faults on different kind of roller bearings.

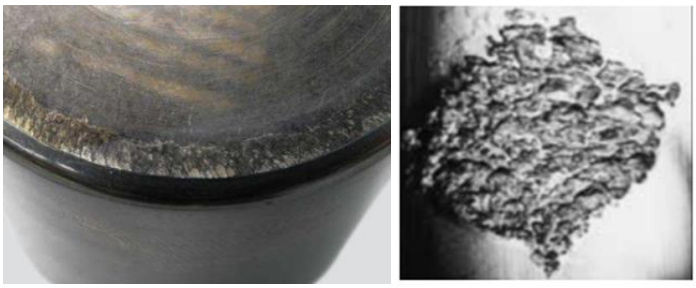


Figure 8. Different morphology defects: adhesive wear (left) and fatigue damage (right) [85]

Possible reasons related of the problem and their morphological appearance on the bearing elements are summarized in Table 1.

Table 1 Fault analysis of defects referring to their morphological shapes [S.7]

Problem	**Possible Causes**
Chatter marks	Incorrect grinding wheel dressing
Regularly spaced marks	Vibration
Short, regularly spaced marks	Spindle pulley loose
Wide, regularly spaced marks	Vibration from flat belt Pulley loose or out of balance
Regular, widely spaced disclosed chatter marks	Grinding problem: wheel specification unsuitable
Long, rectangular shaped marks	Grinding problem: wheel out of balance Grinding problem: wheel face out of round
Long, widely spaced chatter marks	Drive gear backlash
Regular or irregular marks	Faulty thrust bearing
Even and regularly or regularly space marks	Belt joint (belt lacing)
Deep, narrow, regular scratches	Wheel specification unsuitable
Wide, irregular scratches	Wheel specification unsuitable
Mottle marks	Faulty wheel face
Irregular marks of various lengths	Dirty coolant
Grain marks	Wheel specification unsuitable. Incorrect grinding conditions
Uneven traverse lines	Table traverse mechanism worn

Linear (longitudinal), rectangular shape marks of similar manners.	Table traverse mechanism worn. Technological parameters wrong. Etching pits. False brinelling. Pulling problems of the roller. Grinding defects. Cracks.

As the result of the geometrical width analysis of 593 grinding faults (184 outer-ring faults, 162 inner-ring faults, 247 roller faults) on the outer rings, inner rings and on the rollers, the range of the grinding faults occurred in the 0.58 mm to the 1.68 mm interval. Fault sizes followed Gaussian distribution and 95.46 % of the cases were between 0.63 mm to 1.63 mm as Figure 9 presents. Mean (μ) of the measured data was 1.13 mm and std. deviation (σ) was 0.25. Therefore, for the later experiments I chose four sizes of width which represented the whole range: 0.6311 mm, 1.2492 mm, 1.4751 mm, 1.6236 mm. As for the inner ring back-surface faults I measured lower values of grinding faults as small as 0.1127 mm by measuring 138 rollers. Thus, for machine learning classifications in Chapter 8 and Chapter 9 all these values are used to follow the real circumstances of the manufacturing process.

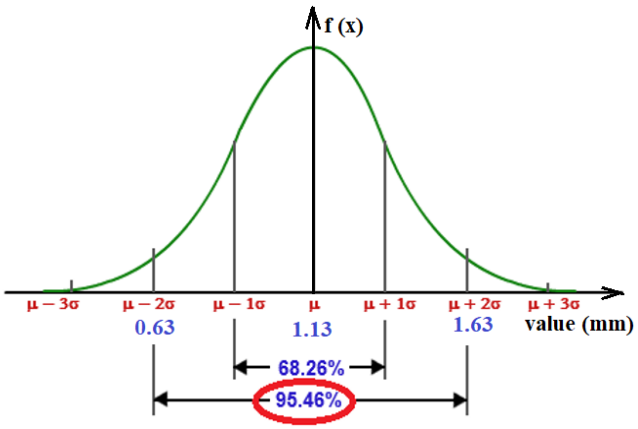


Figure 9. Gaussian distribution of the analysed fault range

2.3 Traditional methods of bearing fault analysis

2.3.1 Time-domain methods

Time domain analysis extracts features from the time domain signal and works with the statistical parameters of the signal.

Crest Factor is the ratio of the peak acceleration to the RMS value. Crest factor is a good indicator of small size defects; although, when localized damage grows, the value of the crest factor decreases significantly because of the increasing RMS.

Skewness is a measure of the asymmetry of the probability distribution of a real-valued random variable about its mean. The skewness value can be positive or negative, or even undefined. Normal distribution has a skewness of zero. But in reality, data points may not be perfectly symmetric. Negative values for the skewness indicate data that are skewed left and positive values for the skewness indicate data that are skewed right.

Kurtosis, the fourth normalized statistical moment, corresponds to the peak value of the data. Kurtosis for a standard normal distribution is 3. Distribution with kurtosis less than 3 are said to be platykurtic. An example of a platykurtic distribution is the uniform distribution, which does not have positive-valued tails. Distributions with kurtosis greater than 3 are said to be leptokurtic. An example of a leptokurtic distribution is the Laplace distribution, which has tails that asymptotically approach zero more slowly than a Gaussian. ([69], [72]). Kurtosis value is more useful, when it is compared with the RMS, crest factor, and peak value. Significant skewness and kurtosis clearly indicate that data are not normal.

Impulse factor is peak divided by arithmetic mean. Shape factor is RMS value divided by arithmetic mean.

2.3.2 Frequency-domain methods

Besides time domain parameters, data revealed from the frequency spectrum are also important information in bearing monitoring. The analysis of the frequency spectrum with Fourier Transform (FT) a common way of bearing diagnosis in the industry, but it is not appropriate for non-stationary cases and the detection of transients in the signal.

The Fourier transform

$$\hat{f}(\omega) = \frac{1}{2\pi} \cdot \int_{t=-\infty}^{\infty} f(t) \cdot e^{-i \cdot \omega \cdot t} \, dt, \quad \omega \in \mathbb{R}$$

of integrable functions $f: \mathbb{R} \to \mathbb{R}$ is a basic concept in the theoretical background of the frequency analysis. Because of the digital measuring systems are used today in industry, the discrete Fourier transform

$$X[k] = \frac{1}{N} \cdot \sum_{n=0}^{N-1} x[n] \cdot e^{-i \cdot k \cdot n \cdot \frac{2\pi}{N}}, \quad k = 0, \dots, N-1$$

of discrete samples $x[0], x[1], \dots, x[n]$, and its faster versions, the fast Fourier transforms (FFT) are the main tool of calculations.

Bearing problems cause frequency peaks with harmonics in the spectrum. Basically, four sections (A, B, C, D) can be distinguished (Figure 10). The main 1× and 2× frequencies are in Section A and are connected to the rotational speed of

the shaft. In Section B, there are the bearing fault frequencies, which are related to the outer ring, inner ring, the roller or the cage of the bearing. They appear with their harmonics in the spectrum. In Section C, the structural frequencies of the bearing developed which are connected to the size and the rigidity of both the bearing and the bearing house. In Section D, the ultrasound frequencies appear. Initial bearing problems are derived from the high frequency regions. As the fault size increase, the frequencies which describe the fault phenomena become lower in frequency and more harmonics developed with side-bands with increasing amplitudes as Figure 11 shows.

It is important to identify these frequencies in the spectrum which is generated by the Fourier transform. FFT is fast way to produce the spectrum, however, nowadays digital signals need Discrete Fourier Transform (DFT).

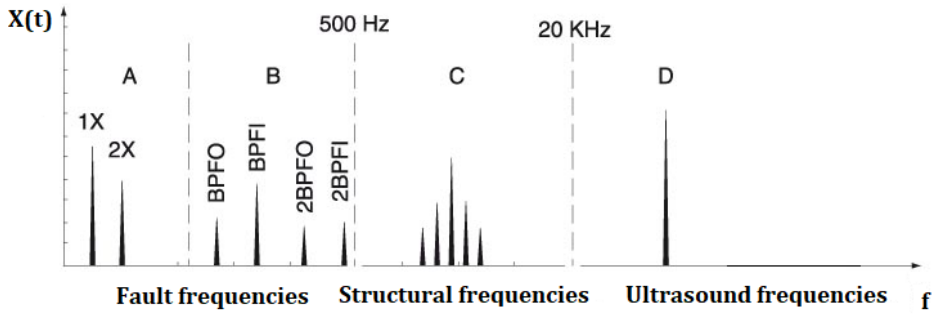


Figure 10. Initial fault frequencies in the vibration spectrum [66]

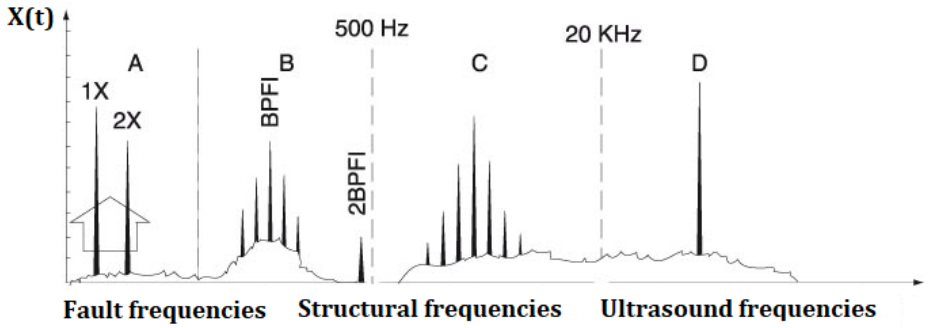


Figure 11. Increased fault frequencies in the vibration spectrum [66]

To choose the proper DAQ unit and adequate test-rig design, it is necessary to know the excepted values of the frequencies in the case of the examined bearings. The vibration signal of the bearing is sampled by the DAQ unit. Shannon-Nyquist criteria defines that the sampling frequency (f_s) should be at least two times higher than the maximum frequency (f_{max}) that we suppose to measure in our experiments ($f_s > 2f_{max}$).

That is typically approximated by filtering the original signal to attenuate high frequency components before it is sampled. If the sampling frequency does not satisfy the Shannon-Nyquist criteria alien signals appears in the vibration signature which are not part of the measured signal. This phenomenon causes serious measurement error that is necessary to avoid. Attenuated high frequency components still generate low-frequency aliases, but typically at low enough amplitudes that they do not cause problems. A filter in anticipation of a certain sample frequency is called an anti-aliasing filter.

3 Time-frequency analysis based on wavelets

Vibrations signals in engineering diagnostics are non-stationary which means frequency-domain representation (frequency spectrum) changes over time. The detection of low-energy transients in the signal requires information about frequencies and also on the time when a particular frequency component present. Several time-frequency analysis methods are known which are able to provide both types of data, for instance the Windowed Fourier Transforms and the Wavelet Transforms (WT). Because of their flexibility, the wavelet transforms (and the related multiresolution analysis) can be used more effectively in the condition monitoring of machine elements.

Wavelets enable to analyse several timescales of the local properties of complex signals that can present non-stationary zones. A wavelet is a function oscillating as a wave but quickly damped. Being well localized simultaneously in time and frequency it makes it possible to define a family of analysing functions by translation in time and dilation in scale. Wavelets constitute a mathematical "zoom" making it possible to simultaneously describe the properties of a signal on several timescales.

3.1 The concept of wavelets

A function $\psi \in L^1 \cap L^2$ will be called a wavelet if it verifies the admissibility condition

$$K_\psi = \int_0^{+\infty} \frac{\left|\hat{\psi}(\omega)\right|^2}{|\omega|} d\omega = \int_{-\infty}^{0} \frac{\left|\hat{\psi}(\omega)\right|^2}{|\omega|} d\omega < +\infty,$$

where $\hat{\psi}$ indicates the Fourier transform of ψ.

The admissibility condition involves, that the wavelet integrates to zero, that is, $\int_{\mathbb{R}} \psi(x)\, dx = 0$. It is often reinforced by requiring that the wavelet has m vanishing moments, i.e.

$$\int_{\mathbb{R}} x^k \cdot \psi(x)\, dx = 0, \quad \text{for } k = 0, \ldots, m.$$

The oscillation of a wavelet is measured by the number of vanishing moments and its localization is evaluated by the interval where it takes values significantly different from zero.

Using translation and dilation a family of functions $\{\psi_{a,b}\}$ is defined by

$$\psi_{a,b}(x) = \frac{1}{\sqrt{a}}\, \psi\left(\frac{x-b}{a}\right)$$

for any scale $a \in \mathbb{R}^+$ and any position $b \in \mathbb{R}$. If ψ has a norm 1 then all the functions $\psi_{a,b}$ have a norm of 1.

A signal f of finite energy can be analysed by its wavelet coefficients

$$C_f(a,b) = \int_{\mathbb{R}} f(x) \cdot \psi^*_{a,b}(x)\, dx, \quad a \in \mathbb{R}^+, b \in \mathbb{R}.$$

The calculation of wavelet coefficients C_f is called continuous wavelet transform. The $C_f(a,b)$ measures the fluctuation of function f at scale a, it depends on the values of f in a neighbourhood of b with a length proportional to a. Large values of $C_f(a,b)$ provide information on the local irregularity of f around position b and at scale a.

The squared magnitude of wavelet coefficients

$$\left|C_f(a,b)\right|^2 = \left| \int_{-\infty}^{\infty} f(t) \cdot \psi^*_{a,b}(x) \right|^2$$

constitutes the so-called scalogram, which is an important tool in of evaluations based on wavelet transforms [81].

For instance, a ranking of wavelets can be generated with respect to their efficiency in bearing fault detection using the so-called Energy-to-Shannon entropy criteria using the scalograms. Values in the scalogram are related to the energy content of signal components.

The representation of scalogram (Figure 12) in the time-frequency plane is useful source of visual information e.g. in the identification of transients in vibration signal caused by surface defects of bearings.

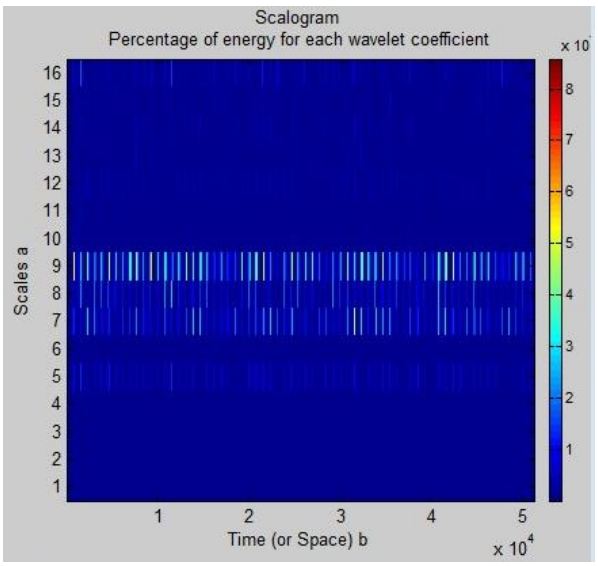


Figure 12. A representation of the scalogram of a defected bearing

Supposing certain stronger properties than merely the admissibility condition we limit ourselves to the values [81]

$$a = 2^j, b = k \cdot 2^j, \quad j,k \in \mathbb{Z}.$$

This idea leads to the discrete wavelet transform which is closely related to the so-called multi-resolution analysis (MRA).

3.2 Multi-resolution analysis, orthonormal wavelet bases

For signals f the following two operations

$$(\tau_a f)(x) = f(x+a), \quad a,x \in \mathbb{R}$$
$$(\delta_s f)(x) = f(s \cdot x), \quad s,x \in \mathbb{R}, s > 0$$
$$(E_1 f)(x) = \sum_{k\in\mathbb{Z}} f(x+k) = \sum_{k\in\mathbb{Z}} (\tau_k f)(x), \quad x \in \mathbb{R}$$

will be used in this section [81].

The concept of multi-resolution is based on translation-invariant Riesz-bases. A system $\{\varphi_k\}_{k\in\mathbb{Z}} \subset X$ is a Riesz-basis of X if for all $x \in X$ there exists a unique $(c_k) \in \ell_2$ such that $x = \sum_{k\in\mathbb{Z}} c_k \varphi_k$.
It is important to note that all orthonormal bases are Riesz-bases as well.

Let $\{\varphi_k\}_{k\in\mathbb{Z}} = \{\tau_k \varphi\}_{k\in\mathbb{Z}} \subset X$ be a Riesz-basis with the generator function φ. Then for all $n \in \mathbb{Z}$

$$\varphi_k^n(x) = \frac{1}{\sqrt{2^n}} \cdot \varphi\left(\frac{1}{2^n} \cdot x - k\right), \quad x \in \mathbb{R}, k \in \mathbb{Z}$$

is also a Riesz-basis for X_n, furthermore, if $\{\varphi_k\}_{k\in\mathbb{Z}}$ is orthonormal, then $\{\varphi_k^n\}_{k\in\mathbb{Z}}$ is orthonormal as well.

A sequence of closed sub-spaces $\{X_n\}_{n\in\mathbb{Z}} \subset X$ is called multi-resolution analysis of X if

1. $X_n \subset X_{n-1}, n \in \mathbb{Z}$ — *monotonicity property*
2. $\overline{\bigcup_{n\in\mathbb{Z}} X_n} = X$ — *density property*
3. $\bigcap_{n\in\mathbb{Z}} X_n = \{0\}$ — *separability property*
4. $f \in X_n \quad \Leftrightarrow \quad \delta_2 f \in X_{n-1}, n \in \mathbb{Z}$ — *scaling property*
5. X_0 is generated by a Riesz-basis given by
   $$\varphi_k = \tau_k \varphi, \quad k \in \mathbb{Z}$$
   (The basis is invariant to integer translations).

The *scaling* property characterizes the multi-resolution aspects: the X_j spaces are obtained by dyadic dilation or contraction of the functions of the single space X_0 through the assumptions

$$\big(x \to f(x)\big) \in X_j \quad \Leftrightarrow \quad \big(x \to f(2 \cdot x)\big) \in X_{j-1}, \quad j \in \mathbb{Z}.$$

Property 5 states that there exists $\varphi \in X_0$ such that $\{\varphi(x-k)\}_{k\in\mathbb{Z}}$ is a Riesz-basis of X_0. Function φ is also called scaling function of the MRA.

Let us denote the sub-space generated by a Riesz-basis φ_k^n as

$$X_n^{\varphi} = \{\sum_{k\in\mathbb{Z}} c_k \cdot \varphi_k^n \,|(c_k) \in \ell_2\}.$$

Each multi-resolution analysis generates an approximation process. If $P_n: X \to X_n, n \in \mathbb{Z}$ denote the othogonal projection of f onto the closed sub-space X_n then $\lim_{n\to\infty} \|P_n f - f\| = 0,\ f \in X$.

It can be proved that, beside general assumptions, there exists an orthonormal basis $\{\varphi_k^n\}_{k\in\mathbb{Z}}$ of X_n, and then P_n can be written as

$$(P_n f)(x) = \sum_{k\in\mathbb{Z}} \langle f, \varphi_k^n\rangle \cdot \varphi_k^n(x) = 2^n \cdot \int_{\mathbb{R}} f(t) \cdot K(2^n \cdot x, 2^n \cdot t)\, dt$$

where the kernel function K has the form

$$K(x,t) = \sum_{k\in\mathbb{Z}} \varphi(x-k) \cdot \varphi^*(t-k)\,, \quad x,t \in \mathbb{R}.$$

Let $\mathcal{M}_0$ be the set of non-negative, even, integrable functions decreasing on [0,1[.

It can be proved that if an MRA is generated by a function φ majorated by an element of $\mathcal{M}_0$ then the equality

$$\sum_{k\in\mathbb{Z}} \varphi(x-k) = 1, \quad x \in \mathbb{R}$$

is equivalent with the density property.

Theorem [81]: The following three statements are equivalent:

i) $\{\varphi_k^n\}_{k\in\mathbb{Z}}$ is a Riesz-basis of X_n for all $n \in \mathbb{Z}$.

ii) $\{\varphi_k^0 = \tau_k\varphi\}_{k\in\mathbb{Z}}$ is a Riesz-basis of X_0.

iii) There exist $0 < m \le M < \infty$ real numbers such that $m \le \sqrt{E_1(|\hat{\varphi}|^2)} \le M$.

Let $\mathcal{R} = \left\{\varphi \in X \,\middle|\, m \le \sqrt{E_1(|\hat{\varphi}|^2)} \le M\right\}$. According to the previous theorem $\mathcal{R}$ contains the functions $\varphi \in X$ for which system $\{\varphi_k^n\}_{k\in\mathbb{Z}}$ is a Riesz-basis of X_n for all $n \in \mathbb{Z}$.

Theorem [81]: Let $\varphi \in \mathcal{R}$. System $\{\tau_k\varphi\}_{k\in\mathbb{Z}}$ is orthonormal if and only if $E_1(|\hat{\varphi}|^2) = 1$.

Let $\varphi \in \mathcal{R}$ and let X_n^{φ} denote the Riesz-basis generated by $\{\varphi_k^n\}_{k\in\mathbb{Z}}, n \in \mathbb{Z}$.

Theorem: The monotonicity property $X_n \subset X_{n-1}, n \in \mathbb{Z}$ is equivalent with the

$$\hat{\varphi}(2x) = \alpha(x) \cdot \hat{\varphi}(x), \quad x \in \mathbb{R},$$

scaling equation, where $\alpha \in X$ is a suitable 1-periodic function. α is called the high-pass filter belonging to φ.

Furthermore, the scaling equation is equivalent with the equality

$$\frac{1}{2} \cdot \varphi\left(\frac{1}{2} \cdot x\right) = \sum_{k\in\mathbb{Z}} a_k \cdot \varphi(x+k)\,, \quad x \in \mathbb{R}$$

where (a_k) is the sequence of the Fourier coefficients of α.

Let $\mathcal{M}$ denote the set of $\varphi \in \mathcal{R}$ functions which have majorant in $\mathcal{M}_0$ and satisfies the condition

$$\hat{\varphi}(0) = \int_{\mathbb{R}} \varphi(t)\, dt = 1$$

If $\varphi \in \mathcal{M}$ then $\hat{\varphi}(x) = \prod_{n=1}^{\infty} \alpha\left(\frac{1}{2^n} \cdot x\right), x \in \mathbb{R}$.

Orthonormal wavelets

From functions φ generating MRA an orthonormal wavelet basis $\left\{X_n^{\psi}\right\}_{n\in\mathbb{Z}}$ of X can be constructed as follows.

In the following we suppose that the normality conditions

$$E_1(|\hat{\varphi}|^2) = 1, \quad E_1\left(\left|\hat{\psi}\right|^2\right) = 1$$

hold, that is, systems $\{\varphi_k^n\}_{k\in\mathbb{Z}}$ and $\{\psi_k^n\}_{k\in\mathbb{Z}}$ are orthonormal for any $n \in \mathbb{Z}$.

It can be proved that there exists a 1-periodic filter $\beta \in X$ such that for ψ defined by

$$\delta_2\left(\hat{\psi}\right) = \beta \cdot \hat{\varphi}$$

$$\psi_k^n(x) = \frac{1}{\sqrt{2^n}} \cdot \psi\left(\frac{1}{2^n} \cdot x - k\right), \quad x \in \mathbb{R}, k \in \mathbb{Z}$$

we have the orthogonal decomposition $X_0^{\varphi} = X_1^{\varphi} \oplus X_1^{\psi}$.

Consequently, for all $n \in \mathbb{Z}$ we have $X_n^{\varphi} = X_{n+1}^{\varphi} \oplus X_{n+1}^{\psi}$, furthermore

$$X_j^{\varphi} = X_k^{\varphi} \oplus \left(\bigoplus_{i=k}^{j+1} X_i^{\psi}\right), \quad j < k,$$

$$X_j^{\varphi} = \bigoplus_{i=j+1}^{\infty} X_i^{\psi}$$

$$X = X_j^{\varphi} \oplus \left(\bigoplus_{i=-\infty}^{j} X_i^{\psi}\right)$$

$$X = \bigoplus_{i=-\infty}^{\infty} X_i^{\psi}$$

This result shows that system $\{\psi_k^n\}_{k\in\mathbb{Z}}$, where functions

$$\psi_k^n(x) = \frac{1}{\sqrt{2^n}} \cdot \psi\left(\frac{1}{2^n} \cdot x - k\right), \quad x \in \mathbb{R}, k \in \mathbb{Z}$$

are generated by ψ satisfying $E_1(|\hat{\varphi}|^2) = 1, E_1\left(\left|\hat{\psi}\right|^2\right) = 1$ and $X_0^{\varphi} = X_1^{\varphi} \oplus X_1^{\psi}$ is orthonormal, that is

$$\langle \psi_k^n, \psi_l^m \rangle = \delta_{kl} \cdot \delta_{nm}, \quad k, l, n, m \in \mathbb{Z}.$$

Function ψ is called the mother wavelet of the MRA.

Sub-spaces $X_j^{\varphi}, k \in \mathbb{Z}$ are called approximation spaces, while $X_j^{\psi}, k \in \mathbb{Z}$ are called detail spaces.

The equality $X_n^{\varphi} = X_{n+1}^{\varphi} \oplus X_{n+1}^{\psi}$ says that an element of the approximation space of level n is decomposed into an approximation at level $n+1$, which is less accurate, and a detail at level $n+1$.

According to $X = \bigoplus_{i=-\infty}^{\infty} X_i^{\psi}$ any signal is the sum of all its details, namely its orthogonal projections onto X_j^{ψ}.

The projection equations are

$$g_j(x) = \sum_{k=-\infty}^{\infty} d_k^j \cdot \frac{1}{\sqrt{2^j}} \cdot \psi\left(\frac{1}{2^j} \cdot x - k\right), \qquad d_k^j = \langle f_{j-1}, \psi_{j,k} \rangle$$

$$f_j(x) = \sum_{k=-\infty}^{\infty} c_k^j \cdot \frac{1}{\sqrt{2^j}} \cdot \varphi\left(\frac{1}{2^j} \cdot x - k\right), \qquad c_k^j = \langle f_{j-1}, \varphi_{j,k} \rangle$$

where d_k^j and c_k^j are the projection coefficients.

The wavelet coefficients of a signal f are provided by

$$\alpha_{k,n} = C_f(k,n) = \int_{\mathbb{R}} f(x) \cdot \psi_k^n(x)\, dx, \quad k,n \in \mathbb{Z}$$

Considering the decomposition $X = X_j^{\varphi} \oplus \left(\bigoplus_{i=-\infty}^{j} X_i^{\psi} \right)$ the orthonormal system

$$\{\varphi_k^0, \psi_k^n\}_{k,n \in \mathbb{Z}, n<0}$$

is used instead of $\{\psi_k^n\}_{k \in \mathbb{Z}}$.

Theorem [81]: If the continuously differentiable 1-periodic function $\alpha \in X$ satisfies

$$\alpha(0) = 1, \ |\alpha(x)|^2 + \left|\alpha(x) + \frac{1}{2}\right|^2 = 1, \qquad x \in \mathbb{R}$$

and

$$|\alpha(x)| > 0, \quad if \quad |x| \le \frac{1}{4}$$

then function φ defined by $\hat{\varphi} = A$, where

$$A(x) = \lim_{m \to \infty} \prod_{n=1}^{m} \alpha\left(\frac{1}{2^n} \cdot x\right)$$

generates an MRA of X.

Wavelets are organized using two parameters time and scale. Time k makes it possible to translate the forms for a given level; scale 2^j makes it possible to pass from a level j to the immediately lower level in the underlying tree represented in Figure 13.

In the first column we find the dyadic dilates of the scaling function φ and in the second column, those of the wavelet ψ.

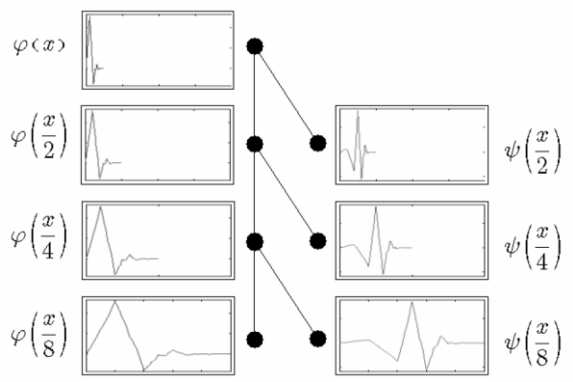


Figure 13. MRA of the signal [62]

The functions in the first column are used for calculating the coefficients of approximation

$$\beta_{k,n} = \int_{\mathbb{R}} f(x) \cdot \psi_k^n(x)\, dx,$$

which define local averages of the signal $f(x)$.
The signal

$$A_j(x) = \sum_{k\in\mathbb{Z}} \beta_{j,k} \cdot \varphi_k^n(x)$$

is an approximation.
The functions in the second column are associated with the calculation of wavelet coefficients

$$\alpha_{k,n} = \int_{\mathbb{R}} f(t) \cdot \psi_k^n(x)\, dx, \quad k, n \in \mathbb{Z}$$

which relate to the differences between two successive local averages. These are the details of the form:

$$D_j(x) = \sum_{k\in\mathbb{Z}} \alpha_{j,k} \cdot \psi_k^n(x)$$

To present certain approximations and details we use the so-called wavelet tree illustrated in Figure 14 for a signal s.

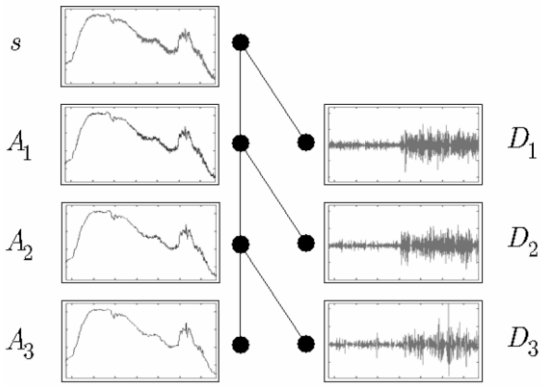


Figure 14. The wavelet tree for a signal [62]

At the root of a we find the signal. The first column in the figure yields three approximations, from the finest A_1 to the coarsest A_3. The differences between two successive approximations are captured in the details denoted D_1 to D_3. More precisely, we have $D_1 = s - A_1$, $D_2 = A_1 - A_2$ and, thus, $s = A_2 + D_2 + D_1$. This representation help us to understand the basic relations between approximations and details, for instance $A_{j-1} = A_j + D_j$, $A_j = \sum_{i>j} D_i$ and $s = A_j + \sum_{i \leq j} D_i$.

The following theorem says that from any scaling function of a MRA a mother wavelet of a complete orthonormal wavelet system can be constructed.

Theorem: Suppose that for $\varphi \in X$ $E_1(|\hat{\varphi}|^2) = 1$ and $\delta_2 \hat{\varphi} = \alpha \cdot \hat{\varphi}$ hold, and define functions β and ψ as

$$\beta = \epsilon \cdot \tau_{\frac{1}{2}}(\alpha^*), \quad \delta_2 \hat{\psi} = \beta \cdot \hat{\varphi}.$$

Then $\{X_n^{\varphi}\}_{n \in \mathbb{Z}}$ is an MRA of X and $\{\psi_k^n\}_{k,n \in \mathbb{Z}}$ complete orthonormal system.

The Fourier coefficients of the 1-periodic function $\beta \in X$ can be expressed by the Fourier coefficients of α, namely if $\alpha = \sum_{k \in \mathbb{Z}} a_k \cdot \epsilon_k$, then $\beta = \sum_{k \in \mathbb{Z}} \alpha_{1-k}^* \cdot \epsilon_k$.

3.3 Wavelets with compact support

In the applications the MRAs are used first of all whose generator functions has compact support. The filters belonging to these generator functions are trigonometric polynomials [81].

Filter α can be found in the form $\alpha = \left(\frac{1+\epsilon_{-1}}{2}\right)^N \cdot T$, where T is a trigonometric polynomial, and additionally, $\alpha(0) = 1$ and $|\alpha(x)|^2 + \left|\alpha\left(x + \frac{1}{2}\right)\right|^2 = 1, x \in \mathbb{R}$ hold. Then

$$|\alpha(x)|^2 = \cos^{2N}(\pi x) \cdot |T(x)|^2, \quad x \in \mathbb{R}$$

where $|T|^2$ a trigonometric polynomial. We can suppose that $|T|^2$ is an even function. Then

$$|T(x)|^2 = P(\cos 2\pi x), \quad x \in \mathbb{R}$$

where P is an algebraic polynomial, and using $\cos 2\pi x = 1 - 2\sin^2 \pi x\,, x \in \mathbb{R}$ we have

$$|T(x)|^2 = Q(\sin^2 \pi x), \quad x \in \mathbb{R}$$

where $Q(x) = P(1-2x), x \in \mathbb{R}$.

Based on the formulas above we have to find the solutions of equation

$$\alpha(0) = 1, |\alpha(x)|^2 + \left|\alpha\left(x + \frac{1}{2}\right)\right|^2 = 1, x \in \mathbb{R}$$

in the form

$$|\alpha(x)|^2 = \cos^{2N}(\pi x) \cdot Q(\sin^2 \pi x), \quad x \in \mathbb{R}$$

Introducing the notation $y = \sin^2 \pi x$ our equation is equivalent with equation

$$(1-y)^N \cdot Q(y) + y^N \cdot Q(1-y) = 1, \quad 0 \le y \le 1$$

where Q is an algebraic polynomial.

Its solutions are

$$Q_N(y) = \sum_{k=0}^{N-1} \binom{N+k-1}{k} \cdot y^k, \quad y \in \mathbb{R}$$

If R is a polynomial for which $R(y) + R(1-y) = 0,\ y \in \mathbb{R}$ then (non-negative) polynomials

$$Q(y) = Q_N(y) + y^N \cdot R(y), \quad y \in \mathbb{R}$$

provide functions α in the form

$$|\alpha(x)|^2 = \cos^{2N}(\pi x) \cdot Q(\sin^2 \pi x), \ \ x \in \mathbb{R}.$$

Calculating T we have the filter α.

3.4 A wavelet design procedure

Although, the user-ready wavelets are effective in many cases, the also the design of new wavelets meeting certain criteria is necessary in some cases for the more efficient diagnosis in specific cases. A natural requirement is to find wavelet functions having special shape, for instance being 'similar' to a given transient in the analyzed signal.

Some direct calculation algorithms are available in the literature, for instance in [28] Chapa and Rao introduce an algorithm for designing a mother wavelet ψ such that it matches a signal of interest and wavelets $\psi_k^n = \frac{1}{\sqrt{2^n}}\psi\left(\frac{1}{2^n}x - k\right)$ form an orthonormal Riesz-basis of X. Supposing band-limited spectrum of the scaling and wavelet functions they give the scaling function by discrete matching procedure from the discrete spectrum of the desired transient appearing in the

signal of interest. In this part I summarize briefly the main steps of the method. In this part the notations of [28] are used.

Suppose that we have a sample

$$\mathbf{W} = \left\{ \left| F\left(k \cdot \frac{2\pi}{2^{\ell}}\right)\right|^2 \middle| k = \left[\frac{2^{\ell}}{3}\right], \dots, \left[\frac{2^{\ell+2}}{3}\right] \right\}$$

where F is the spectrum of the signal.

Let us denote the matched wavelet spectrum as Ψ. Through a least squares optimization process, Theorem 5 in [28] gives values

$$\mathbf{Y} = \left\{ \left| \Psi\left(k \cdot \frac{2\pi}{2^{\ell}}\right)\right|^2 : k = \left[\frac{2^{l}}{3}\right], \dots, \left[\frac{2^{l+2}}{3}\right] \right\}$$

using the error function

$$E(a, \mathbf{Y}) = \frac{(\mathbf{W} - a\mathbf{Y})^T \cdot (\mathbf{W} - a\mathbf{Y})}{\mathbf{W}^T\mathbf{W}}.$$

According to the theorem, the optimal wavelet power spectrum is given by

$$\mathbf{Y} = \frac{1}{a}\mathbf{W} + \mathbf{A}^T(\mathbf{A}\mathbf{A}^T)^{-\mathbf{1}}\left(\mathbf{1} - \frac{1}{a}\mathbf{A}\mathbf{W}\right)$$

where $\mathbf{A}$ is an $L \times 2^l$ matrix and

$$a = \frac{\mathbf{1}^T(\mathbf{A}\mathbf{A}^T)^{-\mathbf{1}}\mathbf{A}\mathbf{W}}{\mathbf{1}^T(\mathbf{A}\mathbf{A}^T)^{-\mathbf{1}}\mathbf{1}}.$$

Theorem 4 in [28] says that, in an orthonormal MRA, values of $|\Phi|$ can be calculated from $|\Psi|$ values using the equality

$$\left|\Phi\left(\frac{k\pi}{2^{\ell}}\right)\right|^2 = \sum_{i=0}^{\ell} \left|\Psi\left(\frac{2k\pi}{2^{i}}\right)\right|^2 \quad \text{for } k \neq 0.$$

The necessary and sufficient condition on $Y(k) = \left|\Psi\left(k \cdot \frac{2\pi}{2^{\ell}}\right)\right|^2, k \in \mathbb{Z}$ to guarantee that $|\Phi(k)|$, provided by Theorem 4, generates an orthonormal MRA is

$$\sum_{i=0}^{\ell} \sum_{m=-\infty}^{\infty} Y\left(\frac{2^{\ell}}{2^{i}} \cdot (k + 2^{\ell+1} \cdot m)\right) = 1$$

where

$$\frac{2^{\ell-1}}{3} < \frac{2^{\ell}}{2^{i}} \cdot \left(k + 2^{\ell+1} \cdot m\right) < \frac{2^{\ell+2}}{3}$$

To determine the specific set of constraint equations, first expand the summation over i.

Condition $\sum_{i=0}^{\ell} \sum_{m=-\infty}^{\infty} Y\left(\frac{2^{\ell}}{2^{i}} \cdot (k + 2^{\ell+1} \cdot m)\right) = 1$ generates a set of L linear equality constraints in $Y(k)$ of the form

$$\sum_{i=1}^{L} \alpha_{ik} \cdot Y(k) = 1 \quad \text{for} \quad k = \left[\frac{2^{\ell}}{3}\right], \dots, \left[\frac{2^{\ell+2}}{3}\right]$$

where $\alpha_{ik} \in \{0,1,2\}$. The matrix form of the condition is

$$\mathbf{AY} = \mathbf{1}$$

where $\mathbf{A} = (\alpha_{ik})$ is an $L \times 2^l$ matrix and $\mathbf{1}$ is a $L \times 1$ vector given by $\mathbf{1} = (1, \dots, 1)$.

Let $\theta_\Phi(\omega)$, $\theta_\Psi(\omega)$, $\theta_H(\omega)$, and $\theta_F(\omega)$ be the phase functions of Φ, Ψ, H and F, respectively, where H is the spectrum of sequence of coefficients (h_n) in equality

$$\varphi(x) = 2 \cdot \sum_{k=-\infty}^{\infty} h_k \cdot \varphi(2x - k).$$

Let us introduce functions

$$\Gamma_\Phi(\omega) = \Lambda_\Phi(\omega) + \frac{1}{2}, \quad \Gamma_\Psi(\omega) = \Lambda_\Psi(\omega) + \frac{1}{2}, \text{and } \Gamma_F(\omega) = \Lambda_F(\omega) + \frac{1}{2}$$

where

$$\Lambda_\Phi(\omega) = \frac{d\theta_\Phi(\omega)}{d\omega}, \quad \Lambda_\Psi(\omega) = \frac{d\theta_\Psi(\omega)}{d\omega}, \quad \lambda(\omega) = \frac{d\theta_H(\omega)}{d\omega}$$

are the so-called group delays of Φ, Ψ, and H, respectively.

A least squares optimization procedure is presented in [81] for matching Γ_Ψ to Γ_F which provides Λ_Φ and Λ_Ψ as well. In calculations the periodic function λ has a central role, its period λ_T is modelled with an R-degree polynomial $\lambda_T(\omega) = \sum_{r=0}^{R/2} c_r \cdot \omega^{2r}$, $\omega \in [-\pi, \pi]$ having only even exponents.

The discrete form for λ can now be written in vector notation

$$\boldsymbol{\lambda} = \mathbf{B}\boldsymbol{c}$$

where $\boldsymbol{\lambda}$ is an $N \times 1$ vector, $\mathbf{c}$ is an $\left(\frac{R}{2} + 1\right) \times 1$ vector, and $\mathbf{B}$ is an $N \times \left(\frac{R}{2} + 1\right)$ matrix whose elements depend on the parameter settings (sampling time T and sample size N) used when sampling F. Using this form of $\boldsymbol{\lambda}$ we have

$$\boldsymbol{\Gamma}_\Psi = \mathbf{D}_\Psi \mathbf{c}$$

where matrix $\mathbf{D}_\Psi$ can be calculated from **B.**

$\boldsymbol{\Gamma}_\Psi$ matching $\boldsymbol{\Gamma}_F$ can be obtained minimizing the error function

$$\gamma = \sum_{n=-N/2}^{N/2-1} (\boldsymbol{\Gamma}_F(n) - \boldsymbol{\Gamma}_\Psi(n))^2$$

in a least squares sense.

To consider the passband for spectra, the error function needs to be normalized by the weighting function $\Omega(n) = \frac{Y(n)}{\sum Y(n)}$, where $Y(n)$ are the elements of $\mathbf{Y}$ provided by the amplitude matching algorithm:

$$\gamma_{\Omega} = \sum_{n=-\frac{N}{2}}^{\frac{N}{2}-1} \left(\Omega(n) \cdot (\mathbf{\Gamma}_{\mathrm{F}}(n) - \mathbf{\Gamma}_{\Psi}(n))\right)^2$$

The vector $\tilde{\boldsymbol{c}}$ minimizing γ can be given as

$$\tilde{\boldsymbol{c}} = \left(\bar{\mathbf{D}}_{\Psi}{}^{T}\bar{\mathbf{D}}_{\Psi}\right)^{-1}\bar{\mathbf{D}}_{\Psi}{}^{T}\bar{\mathbf{\Gamma}}_{\mathrm{F}},$$

where the elements of $\bar{\mathbf{\Gamma}}_{\mathrm{F}}$ are the non-zero values of $\{\Omega(n) \cdot \mathbf{\Gamma}_{\mathrm{F}}(n)\}$ and the elements of $\bar{\mathbf{D}}_{\Psi}$ are the corresponding non-zero values of $\{\Omega(n) \cdot d_{n,r}\}$.

Using $\tilde{\boldsymbol{c}}$ functions $\boldsymbol{\lambda}$, $\mathbf{\Lambda}_{\Psi}$ and $\mathbf{\Lambda}_{\Phi}$ and then the discrete phases of Ψ and Φ can be calculated.

Combining these discrete phases with the magnitudes we get the estimate of Ψ and Φ which satisfy all conditions for an orthonormal MRA.

The impulse responses, h and g, corresponding to the matched wavelet and its scaling function can be found using $\Phi(\omega) = H\left(\frac{\omega}{2}\right)\Phi\left(\frac{\omega}{2}\right), \Psi(\omega) = G\left(\frac{\omega}{2}\right)\Phi\left(\frac{\omega}{2}\right)$ and the inverse Fourier transform.

4. Object-oriented measurement system for bearing analysis

4.1 Test-rig design

For the measurements test rig was designed as the part of the object-oriented measurement system. The shaft in the test rig is supported by two tapered roller bearings. Several types of bearings can be tested if the bearing house is changed. The shaft is driven by an alternating current motor of 0.75 kW capacity (Made: Crompton, frequency 50 Hz, current 4.2 A and speed 2880 $1/min$) with the help of rubber V-belt which provides extremely smooth running. The arrangement provides option of different speeds if it is needed controlled by Schneider ATV32HU22M2 2.2 kW variable speed drive [78]. In the experiments speed of the shaft is measured using an optical tachometer (stroboscope). The vibration piezo electronics accelerometer is placed on the top of the bearing casing perpendicular to the axis of the rotation of the shaft in such a way so that it can acquire vertical acceleration. Test rig can be used for acoustic measurements because an optimized anechoic chamber was installed around the test bearing house with an appropriate feature to supress outside noises and reduce reverb time. Test bearing is spanned by screw mechanism to supply the sufficient axial force to the measurements. Force is measured by strain gauge in Wheatstone-bridge. Constant force during the measurement is ensured with glue (Loctite 243) on the thread of the spanning mechanism.

The bearing test rig was designed in a sophisticated way to oppress the resonance phenomena. Structural frequencies of the test-rig were determined by experimental modal analysis. Impact hammer with a FT analyser was used to perform this test. It is important that the natural frequencies of the structure does not match the frequency of expected from the operation of the test-rig. If a natural frequency matches a frequency of the structure it may start to resonate and causes inaccurate measurements.

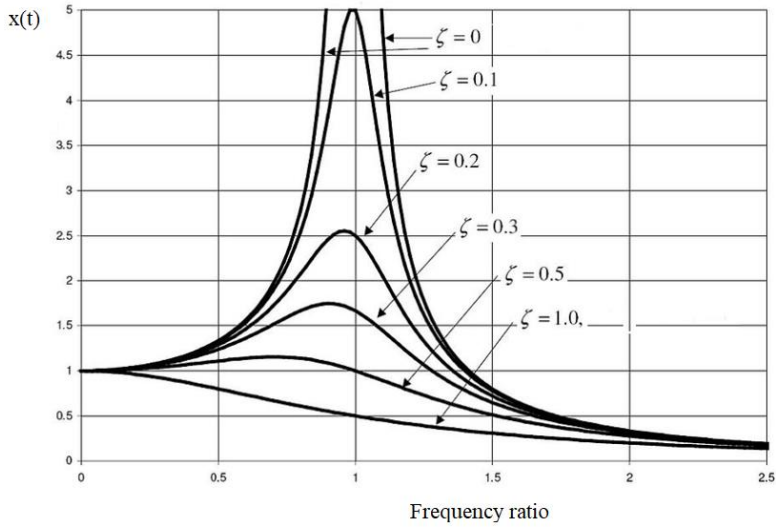


Figure 15. Structural resonance phenomena [71]

To mitigate resonance, the test-rig has a four-layer vibration isolation basement (12) with two-layer rubber plates of 20 mm width. Between the rubber plates soft foam helps to reduce the low frequencies coming from the environment. Rubber bumpers were installed to reduce vibration of the electric motor to the bearing housing in order to minimalize harmful vibrations. The main purpose was to operate the test-rig at 1800 $1/min$ (30 $1/s$) with minimal vibration level to detect bearing faults and estimate the bearing fault size for the precise measurements.

This value satisfies the specifications of American ANSI and German DIN standards (1800 $1/min \pm 2\%$) concerning bearing vibration measurements. ANSI Standard: American National Standard ANSI/AFBMA Std 13-1970, ANSI B3.13-1970, Rolling Bearing Vibration and Noise (Methods of Measuring)
DIN Standard: Deutsches Institut für Normung DIN5426, Laufgeräusche von Wälzlagern, Prüfverfahren.

Figure 15 demonstrates the vibration $x\,(t)$ amplitude in the function of a structure the frequency ratio and the ζ damping factor. Measurement frequencies has to be avoided where the resonance frequency of the structure is the same. In practice, a range of $0.8 \cdot f_s < f_0 < 1.3 \cdot f_s$ frequency-window should be avoided, $24\ Hz < f_0 < 39\ Hz$, to surely avert the vibration phenomena. Therefore, the structure was designed that the main structural frequencies were not in this frequency range. Experimental modal analysis revealed the modal frequencies of the structure (Figure 16) which were 1st modal=13 Hz, 2nd modal=49 Hz, 3rd modal=84 Hz, 4th modal=109 Hz, 5th modal=117 Hz, 6th modal=136 Hz, 7th modal=163 Hz, 8th modal=191 Hz. There is no modal

frequency between $24\ Hz < f_0 < 39\ Hz$ range which was the aim of the design to ensure smooth measurements which need minimal after filtering of the signal.

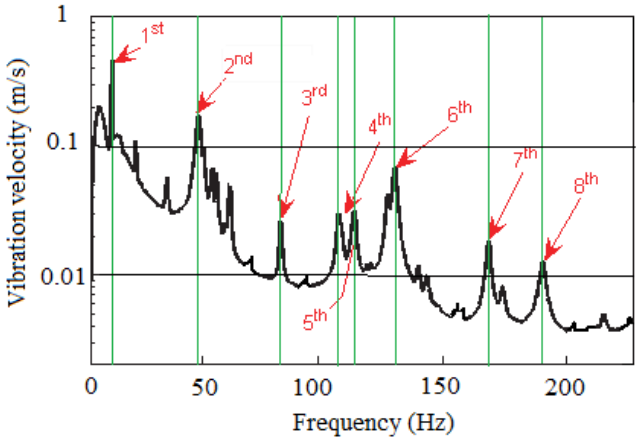


Figure 16. Measured modal frequencies of the test-rig

The (6) engine holder steel structure is made of $40 \times 40 \times 3$ mm rectangular shape welded section. The engine plate (5) is placed on the steel structure via (4) vibration isolation rubber plates which helps to reduce the vibration towards the (1) electric engine. The electric engine was chosen a special type with low vibration level where the bearings of the engine are known, therefore fault frequencies of the electric engine are calculated. In this way, they can be distinguished from the vibration fault frequencies of the tapered roller bearing under testing. Ribbed V-belt (3) produce a very smooth transmission from the engine to the shaft. Radial force caused by the belt force can be neglected. The oscillation of the belt creates a frequency which is different from the bearing fault frequencies and it is easy to filter out from the vibration signal if it is necessary.

The top of the test bearing enclosure is carefully grinded to ensure the $R_a = 1.6\ \mu m$ average surface roughness. Transducer are placed by an $M6 \times 1.5$ screw on the bearing enclosure. This is the best way of fixing compared to the use of glue, bee-wax or magnet because they do not provide smooth frequency transition. The IMI603C01 transducer is pulled by $6.8\ Nm$ torque according to its technical specification which means constant force between the bearing enclosure and the transducer. Between these surfaces thin layer of coupling fluid (oil layer) is used to ensure excellent vibration transition with low damping.

The transmission of the (9) shaft is driven by soft ribbed belt with low vibration feature. It offers smooth driving toward the test-bearing. Tension sleeve (10) ensures the constant axial force of the test bearing during the measurement. The constant tension force is measured with load cells in Wheatstone bridge which was glued on the tension sleeve. Nut (11) is pulled fix

torque to ensure the proper axial loading of 50 … 450 *N* to the examined tapered roller bearing. By adjusting the torque, the axial loading of the bearing could be modified. Furthermore, it is a possibility to create numerous input vectors for machine learning methods as it was executed in Chapter 8 and Chapter 9 for SVM and ANN classification. Maximum axial force was 450 N in order not to exceed the Hertzian contact stress in the contact point between the roller and the faults that cause residual deformation. The purpose is to execute the experiments in the elastic region of the stress-strain curve as Figure 17. illustrates. With optical microscope the contact areas were examined to check the deformations in the contact areas not to exceed the maximum level. Figure 18 and 19 present the parts and basic dimensions of the system.

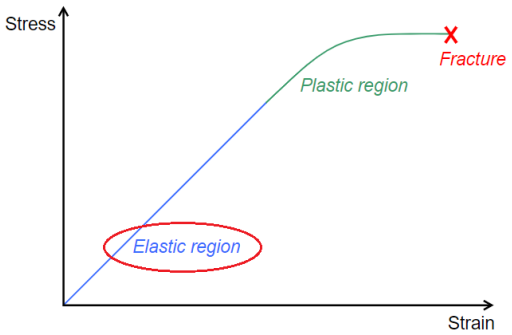


Figure 17. An idealized uniaxial stress-strain curve showing elastic and plastic deformation of the bearing elements

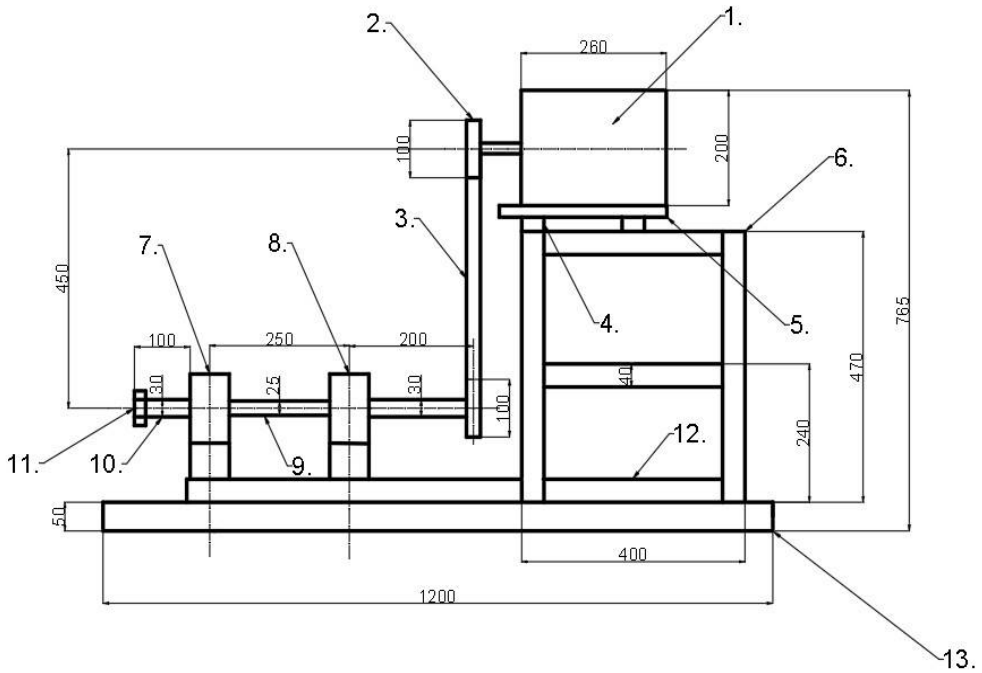


Figure 18. Schematic view of the test-rig

1. Electric motor (type: Cemer, power consumption: $0.75kW$)
2. Ribbed belt disk (dimension: $100 \times 20mm$)
3. Ribbed belt (six-ribbed belt, dimension: $20mm$)
4. Vibration isolation rubber plate (dimension: $40 \times 25mm$)
5. Engine plate
6. Engine-holder steel structure (dimension: $40 \times 40 \times 3mm$)
7. Test-bearing enclosure
8. Holder-bearing enclosure
9. Shaft (length: $600mm$, diameter: from $25mm$ to $30mm$)
10. Tension sleeve (dimension: $30 \times 80mm$)
11. Nut ($M24 \times 1.5mm$)
12. Holder steel structure (dimension: $40 \times 40 \times 3mm$)
13. Vibration isolation rubber basement (dimension: $1200 \times 600mm$)

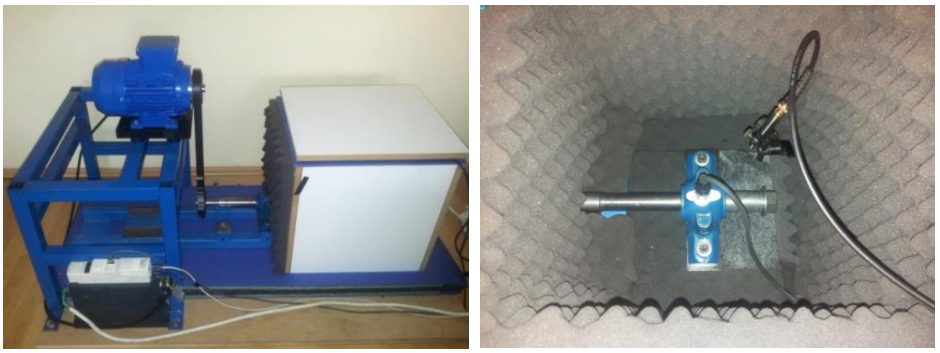


Figure 19. Bearing test rig (left) test bearing in acoustic chamber with pyramidal RAM and electric microphone (right)

4.2. Transducers for vibration acceleration measurement

The accelerometer measures acceleration in the vibration signal. In the experiments PCB IMI 603C01 vibration transducer (Figure 20) is platinum Stock Products; Low-noise, industrial, ceramic shear ICP® accelerometer, 100 mV/g, 0.27 kHz to 10 kHz, top exit, 2-pin connector, single point ISO17025 accredited calibration. The accelerometer has a linear response over a wide frequency range (0.5 Hz to 20 kHz) which is linear up to 50 kHz [35].

Figure 20. PCB IMI 603C transducer for the experiments [77]

4.3. Data acquisition unit

NI 9234 DAQ is applied which is a 4-channel C Series dynamic signal acquisition module (Figure 21) for making high-accuracy audio frequency measurements from integrated electronic piezoelectric (IEPE) and non-IEPE sensors with NI CompactDAQ or CompactRIO systems. The NI 9234 delivers 102 dB of dynamic range and incorporates software-selectable. The four input channels simultaneously digitize signals at rates up to 51.2 kHz per channel with built-in anti-aliasing filters that automatically adjust the sampling rate [76]. 64 bit Intel(R) Core(TM) i5-8600K 3.60 GHz with 32 GB RAM system was used as hardware components.

Figure 21. NI 9234 DAQ for the experiments [76]

Further parts of the object-oriented measurement system are: algorithm for surface fault detection and fault size estimation (designed wavelet and filters), developed Labview VIs for measurements and data processing which are presented in the next Chapters in connection with the experiments.

5 Optimal wavelet selection for grinding defect size estimation

5.1 Fault size estimation

Analysis of the entry point and the exit point are necessary for fault size estimation [10]. Figure 22 represents the entry and exit events when the roller contacts the edges and the bottom of the fault. Ball stress varies during the process, however in this experiment not the mechanical stress is measured but the vibration acceleration value which is proportional with the stress according the Newton's second law and Hooke's law. Linear and isotropic material model is supposed with Poisson values of 0.33 of the 100Cr6 material. Bearing behaves as a mass-spring-damper system with weak damping which creates transient vibration waves when hit the fault.

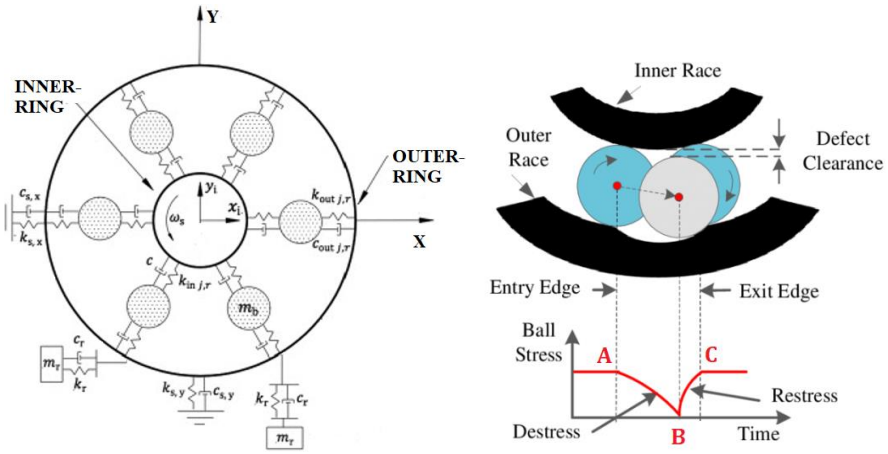


Figure 22. Illustration of the mass-spring-damper mechanical model of the bearing (left) and the entry and exit events with stress response (right)

At point *A*, the roller strikes the rectangular shape grinding defect with high impact which results in re-stressing and high impulse in signal. After this event the roller remains in contact with the defect for some time as Figure 23 presents. When the roller comes in contact with the point *B* it again generates high amplitude in the signal and beyond point *B* progressive decrease in amplitude of signal is observed due to elastic damping of the bearing element.

Fault size is calculated by the time "distance" between points *A* and *C*. This method is very useful because the defects width of the bearing can be determined only from the vibration signature. Bearing defects generate transient impulses in the vibration signal when the rollers pass through the defects. The bearing fault frequencies can be calculated by numerical way: bearing pass frequency of

outer race (BPFO), bearing pass frequency of inner race (BPFI), fundamental train frequency (FTF), ball spin frequency (BSF). For instance, FTF can be calculated as [72]

$$FTF = \frac{f_r}{2}\left(1 - \frac{d}{D}cos\Phi\right)$$

where f_r is the rotational frequency of the shaft, Φ is the contact angle, d is the inner ring diameter, D is the outer ring diameter. The defect can be calculated, where D_{OI} is the outer ring diameter on the raceways, Δt is the time duration between the contact points of the bearing elements [12].

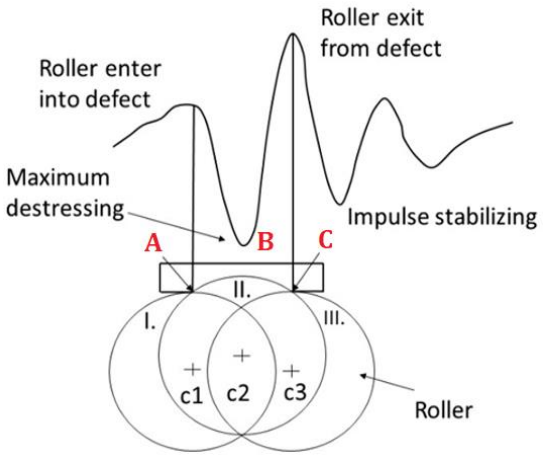


Figure 23. Analysis of the transient impulse presenting the entry and exit events of the roller into the grinding defects on the outer ring [12]

For fault size estimation fault frequencies are calculated [10] which in this experiment are $BPFO = 206.18\ Hz, BPFI = 287.15\ Hz, FTF = 12.88\ Hz, BSF = 89.96\ Hz$ in this experiment at 1800 1/min in case of the No. 30205 bearing.

$$L_{OD} = \pi \cdot \Delta t \cdot D_{OI} \cdot FTF = 1713.74 \cdot \Delta t$$

5.2 Wavelet selection with Energy-to-Shannon Entropy Criteria

The Energy-to-Shannon Entropy Criteria is used to rank wavelets on the basis of scalograms. The Energy-to-Shannon Entropy ratio is a combination of the energy content

$$E(n) = \sum_{i}^{m} \left|C_{n,i}\right|^2$$

and the Shannon entropy

$$S(n) = -\sum_{i=1}^{m} p_i \log_2 p_i,$$

related to the wavelet coefficients $C_{n,i}$, where m is the number of the wavelet coefficients of n-th scale and $(p_1, \dots, p_n)$ is the energy distribution of the wavelet coefficients defined by

$$p_i = \left|C_{n,i}\right|^2 / E(n).$$

The indicator

$$\xi(n) = E(n)/S(n)$$

is used to choose the best wavelet for diagnosis of a special fault. [58]

Nine different wavelets are compared to reveal the grinding faults. According to the literature overview it was found that these wavelets were efficient for bearing fault analysis in general bearing fault diagnostics. Values of the Energy-to-Shannon Entropy ratios are in Table 2.

Table 2. Values of Energy-to-Shannon Entropy ratio belonging to the different wavelets, F_c=2.09 kHz (Morlet)

E/S	OR1	OR2	OR3	OR4	Mean
Sym2	59.96	80.20	100.87	109.16	87.55
Sym5	65.58	95.37	117.07	119.92	99.48
Sym8	82.81	113.77	120.16	118.48	108.81
db02	60.91	81.09	101.12	113.46	89.14
db06	71.74	89.02	120.42	117.40	99.65
db10	77.76	104.69	120.34	120.45	105.81
db14	85.02	120.03	121.37	123.78	112.55
Meyer	92.31	160.31	126.20	105.70	121.13
Morlet	**113.15**	**194.15**	**142.18**	**138.14**	**146.90**

It is observed that Morlet wavelet provided the highest value that indicates to be the most efficient wavelet from the nine wavelets for both fault detection and fault size estimation of the special grinding defect.

5.3 Fault size estimation procedure

To determine the defect size MRA is applied by filter banks which is a design method of most of the practically relevant discrete wavelet transforms.

In the case of No. 30205 tapered roller test bearing BPFO is 206.18 Hz. Down to 3rd level, where transient impulse is analysed for defect width estimation, wavelet band is 1.25 kHz which is more than 3 times bigger than BPFO.

The raw signal is too noisy to detect entry and exit points of the defect but wavelet decomposition makes it possible to analyse the entry and exit events. Using Energy-to-Shannon Entropy Criteria we obtain the best wavelet to determine the fault size from the vibration signature.

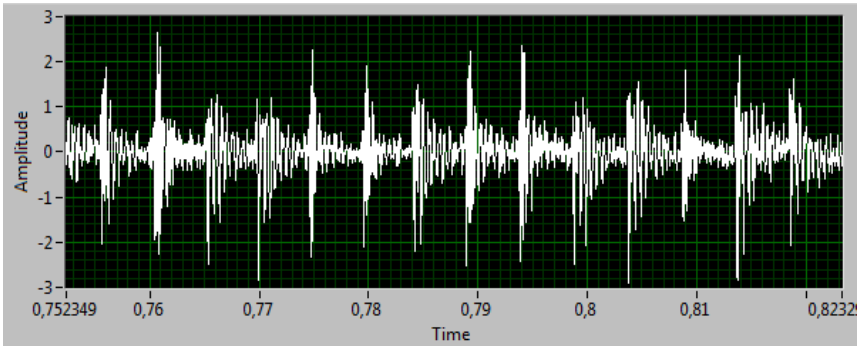


Figure 24. Typical raw time domain signal of bearing having 0.6311 mm of grinding fault width on the outer race

Figure 24 presents the time-domain signal of outer race defect of 0.6311 mm. The highest periodic transient impulse related energy content of the burst occurs at 2.09 kHz that causes 5 ms rate of periodicity which is equal to 206.18 Hz BPFO frequency. The spectrum was determined in all outer rings with different fault sizes and they showed similar manner around the peak at 2.09 kHz as it can be seen in Figure 25.

Multiresolution analysis is made in order to obtain precise frequency analysis. Figure 26 presents the wavelet decomposition tree. Higher decomposition is not necessary because it might not reveal any further information of the signal. Regarding the BPFO frequency analysis was made at 3rd detail level (cD_3) from 1.25 kHz to 2.5 kHz.

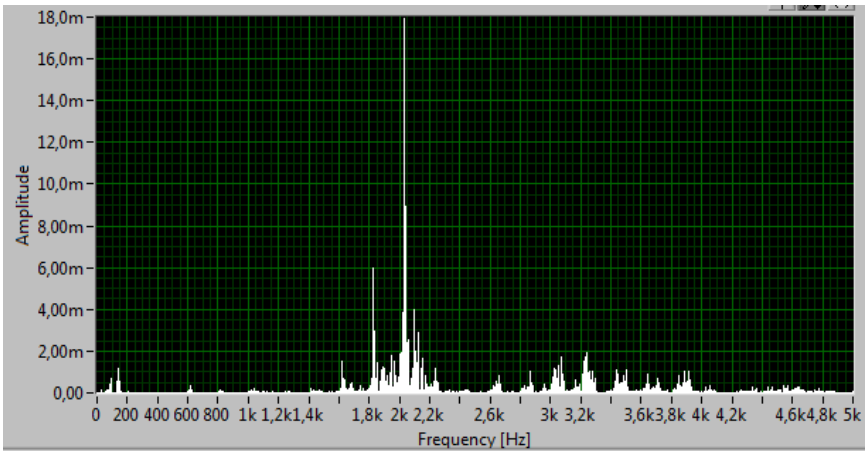


Figure 25. Frequency domain spectrum of bearing having 0.6311 mm of grinding fault width on the outer race, transient frequency of 2.09 kHz

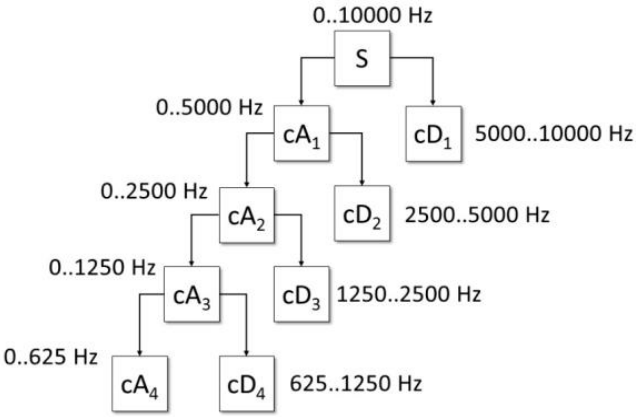


Figure 26. Wavelet decomposition graph of the original vibration signal by MRA with frequency ranges

Figure 27 illustrates the measured transient signal. In the experiment 20 impulses were measured and the average time values were determined between the entry and exit points.

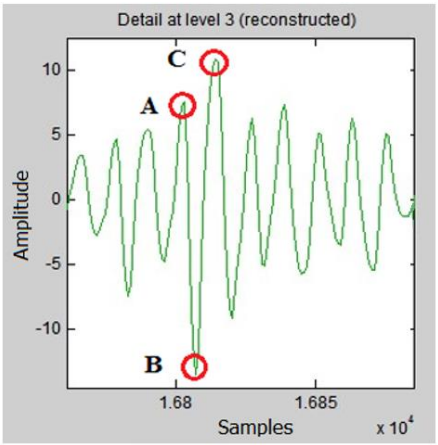


Figure 27. A section of the original vibration signal of 0.6311 mm fault width after MRA at detail 3rd level, entry point (A), de-stressing point (B) and exit point (C) are highlighted

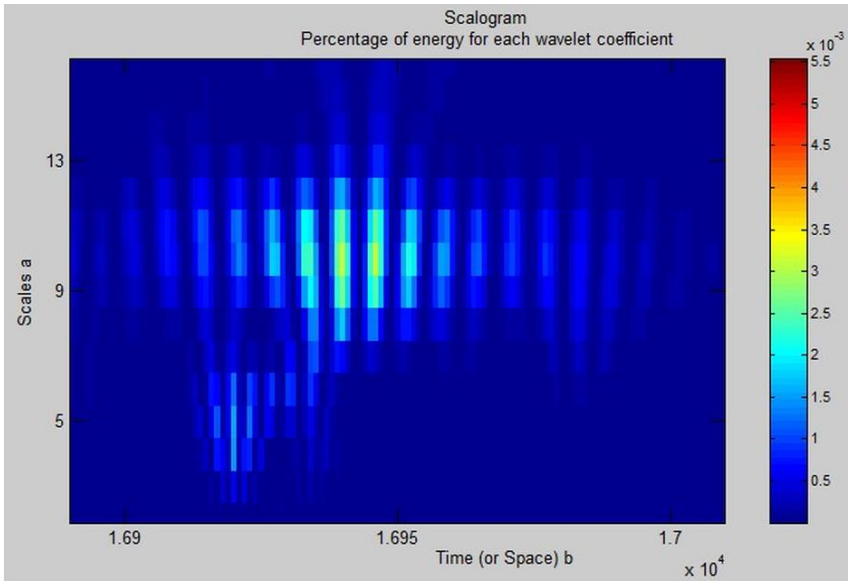


Figure 28. Scalogram of the Morlet wavelet of the outer ring with 0.6311 mm fault width

There is variation in data points as each roller cross over the defect. Average data points are calculated for estimating the time taken by roller to pass over the grinding defect. Figure 28. represents the scalogram by Morlet wavelet which provided the highest wavelet coefficient values.

5.4 Verification of the fault size

To verify the precision of the fault size estimation a contact and optical measurements were applied to reveal the exact geometrical size of the ground defect on the outer race. Garant MM1-200 video microscope and Mahr MMQ 200 contact equipment are used to measure geometrical sizes of the faults as Figures 29-30 present with the results of the measurements as Figure 31-34 illustrate.

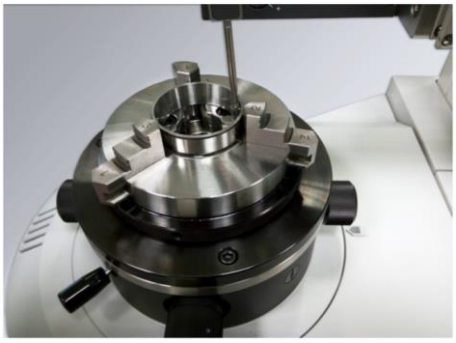


Figure 29. Contact measurement of the outer race defect width with Mahr MMQ 200

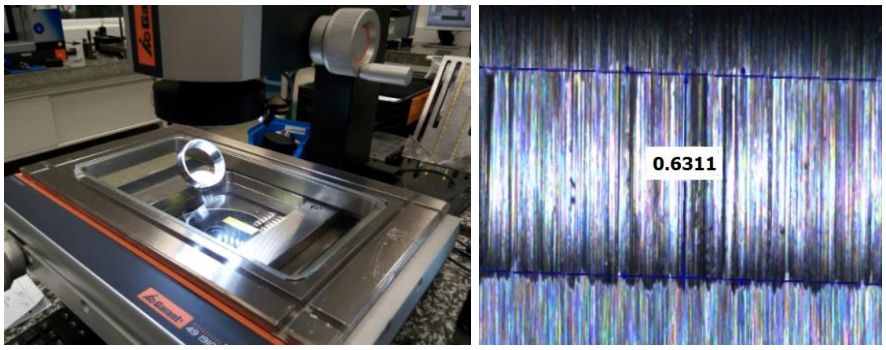


Figure 30. Image of the optical measurement of the outer race defect width (in mm) with Garant MM1-200 video microscope

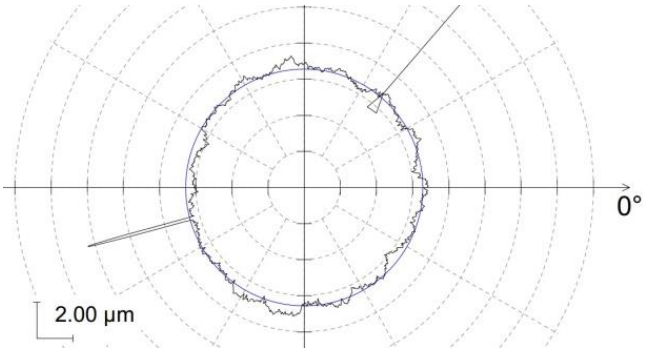


Figure 31. Result of the roundness measurement of the outer race defect with Mahr MMQ 200 equipment, OR1 defect

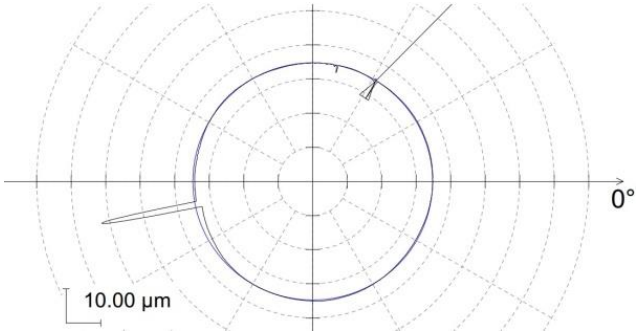


Figure 32. Result of the roundness measurement of the outer race defect with Mahr MMQ 200 equipment, OR2 defect

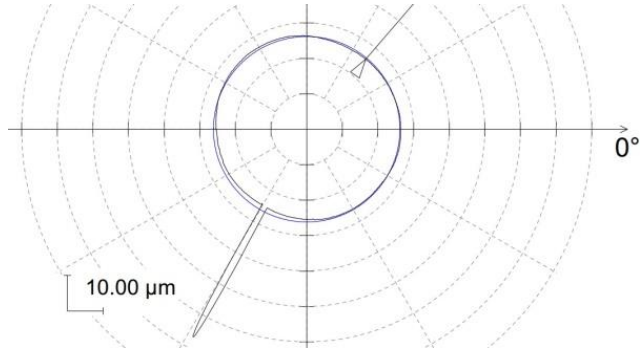


Figure 33. Result of the roundness measurement of the outer race defect with Mahr MMQ 200 equipment, OR3 defect

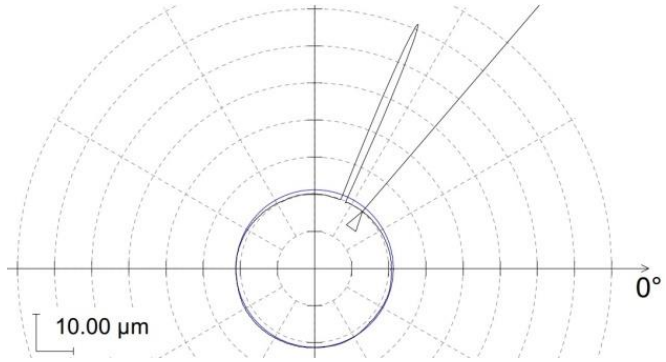


Figure 34. Result of the roundness measurement of the outer race defect with Mahr MMQ 200 equipment, OR4 defect

The maximum difference was obtained *4.12 %* for defect width of *0.6311 mm* by db02 wavelet. Morlet wavelet provided the most accurate width estimation from the vibration signal. Tables 3-5 present the calculated and measured widths and their deviations of the Daubechies_02 (db02), Symlet-5 and the Morlet wavelet.

Table 3. Deviation between calculated and measured fault widths with db02 wavelet

Parameters	**OR1 defect**	**OR2 defect**	**OR3 defect**	**OR4 defect**
Calculated defect widths (mm)	0.657	1.277	1.521	1.653
Optically measured defect widths (mm)	0.6311	1.2492	1.4751	1.6236
Deviation between calculated and measured data (%)	4.12	2.24	3.11	1.78
Mean deviation (%)	2.81			

Table 4. Deviation between calculated and measured fault widths with Symlet-5 wavelet

Parameters	**OR1 defect**	**OR2 defect**	**OR3 defect**	**OR4 defect**
Calculated defect widths (mm)	0.646	1.269	1.4481	1.639
Optically measured defect widths (mm)	0.6311	1.2492	1.4751	1.6236
Deviation between calculated and measured data (%)	2.41	1.56	1.83	0.92
Mean deviation (%)	1.68			

Table 5. Deviation between calculated and measured fault widths with Morlet wavelet

Parameters	**OR1 defect**	**OR2 defect**	**OR3 defect**	**OR4 defect**
Calculated defect widths (mm)	0.644	1.266	1.502	1.637
Optically measured defect widths (mm)	0.6311	1.2492	1.4751	1.6236
Deviation between calculated and measured data (%)	2.06	1.38	1.69	0.84
Mean deviation (%)	1.49			

5.5 Results

A passband filter was applied to suppress noise and enhance the useful information from the vibration signature. Nine real-valued wavelets were analysed and the Morlet wavelet was proved to be the best for the rectangular shape manufacturing grinding fault detection in all cases on the basis of the Shannon Entropy Criteria. Furthermore, the width estimation of the outer ring grinding fault was executed with all nine wavelets. The best wavelet previously chosen by the Energy-to-Shannon Entropy Criteria enables to determine the grinding defect width in the most accurate way. The proposed technique has been successfully implemented for measuring defect width over a range of $0.63\,mm$ to $1.63\,mm$. The defect width has been also verified by optical microscope and contact roundness measurement device using image processing techniques. The maximum deviation between the calculated and the measured values was 4.12 % of the defect width of $0.6311\,mm$.

6 Transient model and wavelet design procedure

Transients of different linear shape surface problems of similar morphology are analysed which were generated by the defects. Discrete wavelet was designed to create a wavelet function which is similar to the signal caused by the special grinding fault. Wavelet Design VI in Labview and the Chapa-Rao algorithm were used [28].

Linear shape surface fault causes an amplitude modulation that is more emphasized in the load zone. One of these impulses can be described as

$$x(t) = A \cdot t^n \cdot e^{-C \cdot t} \cdot \sin(\omega \cdot t),$$

where $\omega = 2\pi \cdot f$, and f is the natural frequency of the transient, A is the initial amplitude, the damping factor C and the exponential n show the behaviour of the transient. Figure 35 shows the parameters of a measured transient signal.

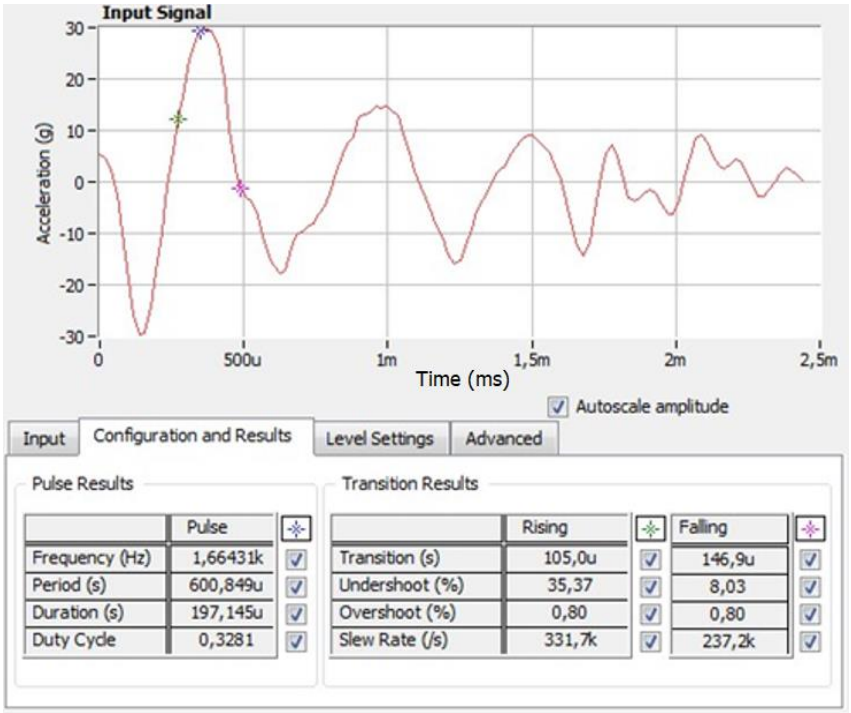


Figure 35. Parameters of a transient signal of the measurement for model creation

It is supposed that the transient can be described by the $x(t)$ function in case of similar shape faults.

Several different types of roller bearings were investigated: cylindrical roller bearings, tapered roller bearings, spherical roller bearings, needle roller bearings. The common feature of these bearings that all of them have rectangular shape contact area.

Experiments with different size of roller bearings were executed. In case of tapered roller bearing the No. 30205, No. 30206, No. 30208 bearings were analysed. The result was that the bearings of different sizes with linear grinding defects produced similar transients. The frequency of the transients and the amplitude were different but typically it was a short-lived burst of energy in a system caused by a sudden change. Only some parameters of the transients were different from each other such as the frequency of the transient, period, duration but the shape of the signal was the same.

Therefore, for the direct measurement No. 30205 tapered bearing was used with rectangular shape manufacturing grinding faults.

In order to create a new wavelet basis function, 1024 sample points of the transient were used. The transient could be described with $A = 28.74$, $n = 2.748$, $C = 17.34$ and $\omega = 265.28$ equation parameters. For the signal model it is useful to follow the transient with an envelope curve and determine the amplitude spectra. It showed close correlation to the transient with minimal mean-square error as Figure 36-37 present.

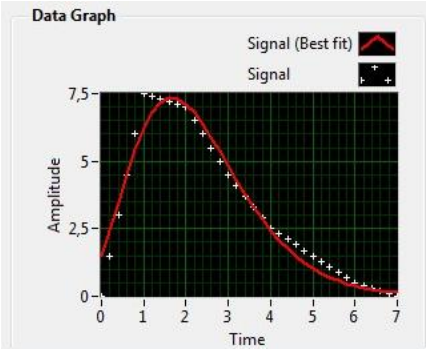


Figure 36. Envelope model of the transient signal

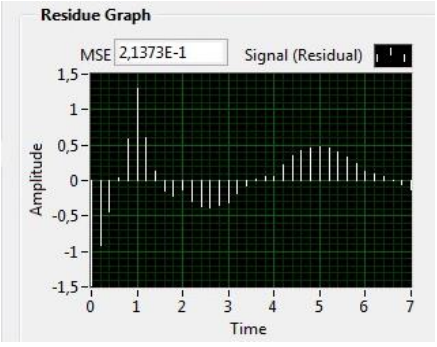


Figure 37. Mean Square Error of the transient from the envelope values

The new wavelet amplitude spectra match the amplitude spectra of the transient with the MSE 0.2137 value as Figure 37 illustrates.

6.1 Application of Labview Wavelet Design

The Wavelet Analysis Tools in Labview provides a collection of commonly used continuous wavelets, such as Mexican hat, Meyer, and Morlet, and a collection of commonly used discrete wavelets, such as the Daubechies, Haar, Coiflet, and biorthogonal wavelets. It is possible to create a discrete wavelet that best matches the signal.

The design procedure consists of three steps as Figure 34 shows.

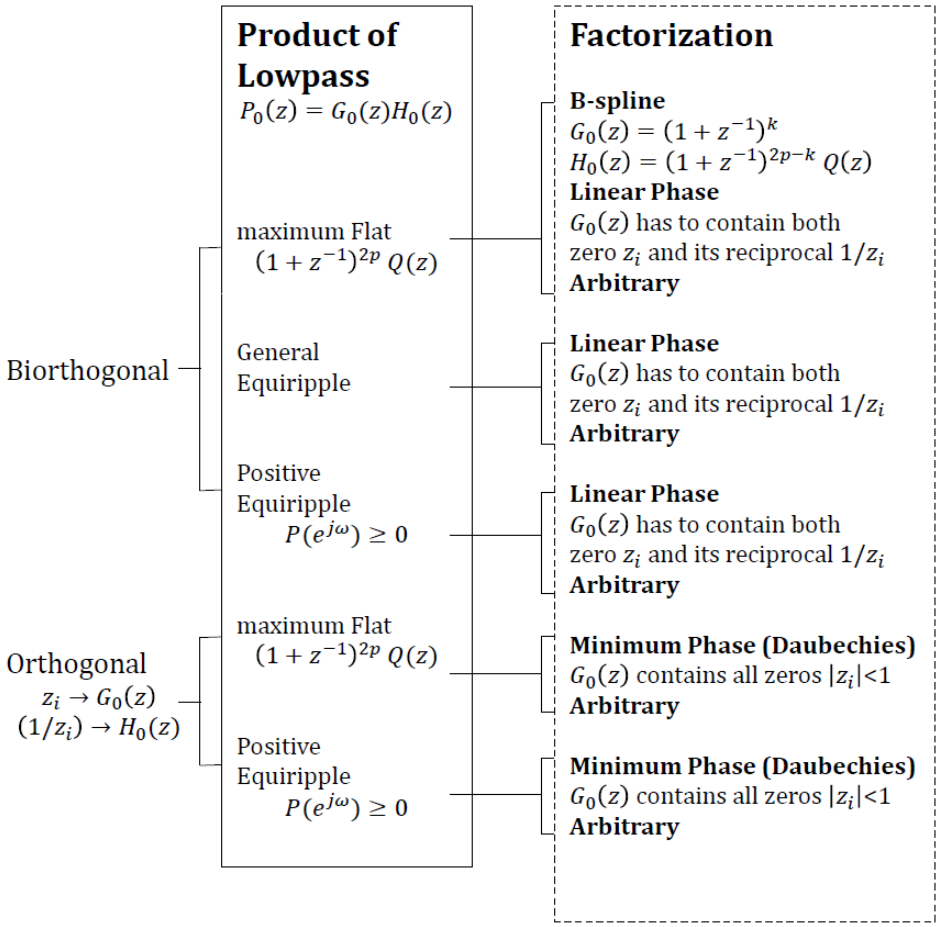


Figure 38. Design procedure for wavelets and filter banks in Labview

Figure 38 represents the wavelet design configuration. The combination of zeros is not unique. Because all filters act as real-valued FIR filters, the zeros of $P_0(z)$, $G_0(z)$ and $H_0(z)$ are symmetrical in the z-plane. Zeros of G_0 and H_0 of the

new-designed wavelet are chosen in a way to provide the best result for the diagnosis of the grinding problem.

For both orthogonal and biorthogonal wavelets and filter banks, either maximum flat or equiripple filters for the product of lowpass filters $P_0(z)$ can be used. The maximum flat filters have good frequency attenuation, but wider transition band. In the experiment, the positive equiripple is used which is a halfband filter, namely a special case of general equiripple halfband filters. It proved to be more efficient than the maximum flat filter. The Fourier transform of the positive equiripple filter is always non-negative. Positive equiripple halfband filter is appropriate for orthogonal wavelets because the auxiliary function $P_0(z)$ must be non-negative.

Remez exchange algorithm was used as the pert of the Parks-McClellan method to find an optimal equiripple set of coefficients which is an iterative algorithm used to find simple approximations to functions. The algorithm then finds the set of $N + 1$ coefficients that minimize the maximum deviation from the ideal. Intuitively, this finds the filter that is as close as possible to the desired response given that only $N + 1$ coefficients can be used. Parks-McClellan VI generates a set of linear-phase FIR multiband digital filter coefficients using the number of taps, sampling frequency: fs, band parameters and filter type.

Two parameters are associated with equiripple filters number of taps and Passband. Use the number of taps control to define the number of coefficients of $P_0(z)$. Because $P_0(z)$ is a type-I FIR filter, the length of $P_0(z)$ must be odd. Use the Passband control to define the normalized passband frequency, ω_p, of $P_0(z)$. The value of ω_p must be less than 0.5 and the value of 0.3 was chosen here. Longer filters improve the sharpness of the transition band and the magnitude of the attenuation in the stopband at the expense of extra computation time for implementation. This lowpass filter with 31 taps provided the wavelet with sufficient correlation to the transient under analysis (Figure 31). In a general equiripple halfband filter, halfband refers to a filter in which $\omega_s + \omega_p = \pi$, where ω_s denotes the stopband frequency and ω_p denotes the passband frequency, as Figure 39 shows.

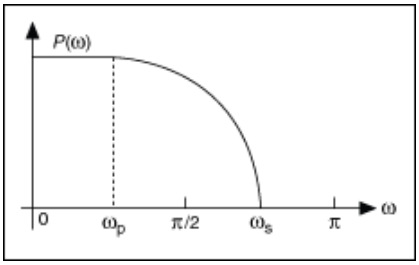


Figure 39. General equiripple halfband filter [61]

The filter has the form

$$P_0(z) = (1 + z^{-1})^{2p} Q(z)$$

and as many zeros are imposed at $\omega = \pi$ as we like. The halfband equiripple filters only can have a pair of zeros at $\omega = \pi$, which gives the equiripple type filters slower convergence rates. However, it is easier to balance the frequency attenuation and transition band for an equiripple filter therefore it is used in this experiment. For a given transition band, the attenuation is proportional to the filter order of $P_0(z)$. The larger the order, the better the attenuation. The zero pairs at π specifies the value of the parameter p, which determines the number of zeros placed at π on the unit circle. The more the zeros at π, the smoother the corresponding wavelet. The value of p also affects the transition band of the frequency response. A large value of p results in a narrow transition band. In the time domain, a narrower transition band implies more oscillations in the corresponding wavelet.

Once $P_0(z)$ is determined it is necessary factorize it into the lowpass filters, $G_0(z)$) and $H_0(z)$. Factorization filter type was chosen arbitrary.

Using these setting parameters (Figure 40) the designed wavelet provided higher E/S ratio than the conventional Symlet and Daubechies wavelets which are frequently used in bearing fault diagnostics. In Chapter 6.2 the results are shown in detail.

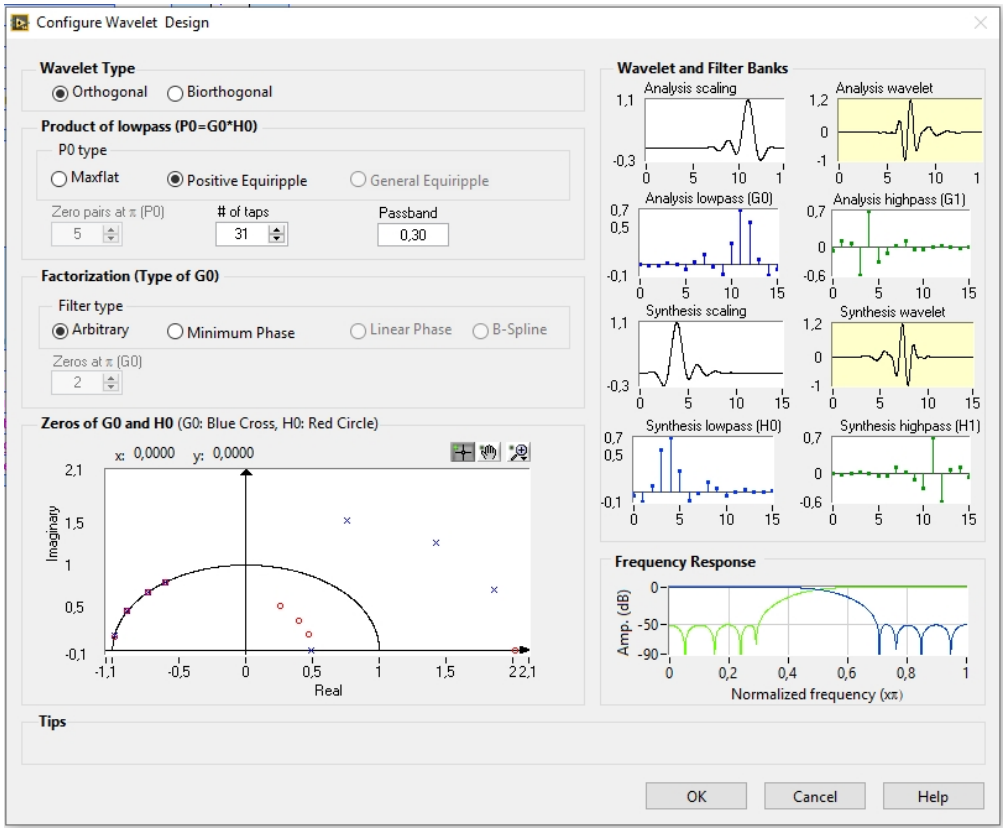


Figure 40. Configuration of the wavelet design

Since the new wavelet basis cannot be given in closed form it is given by the filter coefficients in Table 6.

Table 6. New-designed wavelet filter coefficients

	Synthesis filters		**Analysis filters**	
No.	**Lowpass**	**Highpass**	**Lowpass**	**Highpass**
0	-0.07	0.0032	0.00	-0.0657
1	-0.14	-0.0069	-0.01	0.142
2	0.07	0.0025	0.00	0.0698
3	0.57	0.0258	0.03	-0.5680
4	0.73	0.0023	0.00	0.7290
5	0.28	-0.0529	-0.05	-0.2811
6	-0.12	-0.0375	0.04	-0.1241
7	-0.03	0.0129	0.13	0.0328
8	0.13	0.0330	-0.03	0.1300
9	0.04	-0.1230	-0.12	-0.0373
10	-0.05	-0.2810	0.28	-0.0531

11	0.00	0.7280	0.73	0.0022
12	0.03	-0.5670	0.57	0.0258
13	0.00	0.0704	0.07	0.0026
14	-0.01	0.1420	-0.14	-0.0068
15	0.00	-0.0658	-0.07	-0.0032

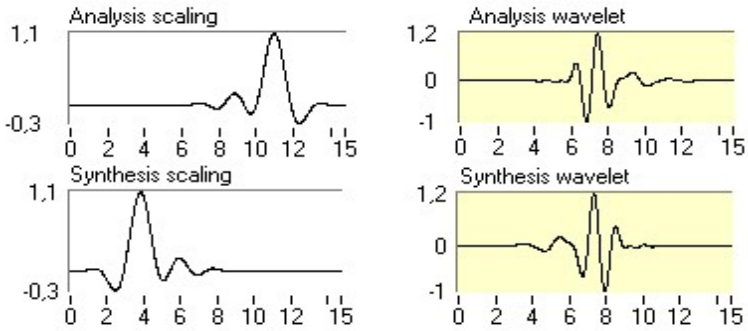


Figure 41. Analysis and synthesis functions of the new-designed wavelet

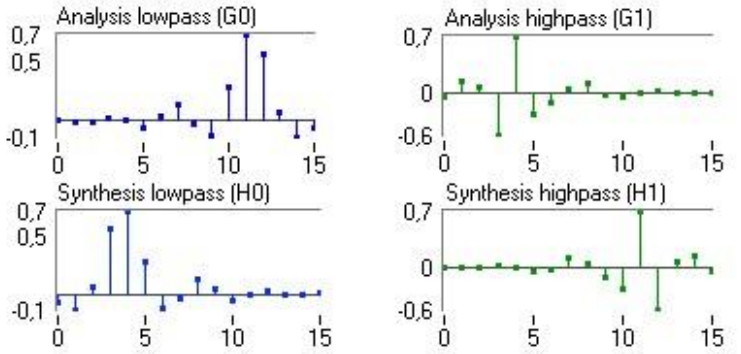


Figure 42. Values of the analysis and synthesis filters of the new-designed wavelet

The coefficients of analysis and synthesis filters are in Figures 41-42. The filter coefficients were checked with Chapa and Rao's method, where the Meyer wavelet amplitude and phase spectra were matched independently to the signal. Since the two wavelets provided the same efficiency in fault detection and width estimation in practice, the filter coefficients provided by Labview were applied. The new wavelet was compared with five generally used discrete wavelets using the Maximum Energy-to-Shannon Entropy ratio criteria.

6.2 Ranking by Energy-to-Shannon Entropy Criteria

Figures 43-45 are a short section of the time domain signal containing one transient which is similarly repeated through the whole measurement. The figures show that the Daubechies_08 gives the worst correlation to the measured

signal therefore it is assumed to be not effective for detecting the bearing fault. By visual inspection it is clear that the new-designed wavelet provides excellent correlation to the measured signal. However, visual comparison is not precise method to determine the efficiency of a wavelet. For this purpose, the Energy-to-Shannon Entropy criteria provided more accurate result.

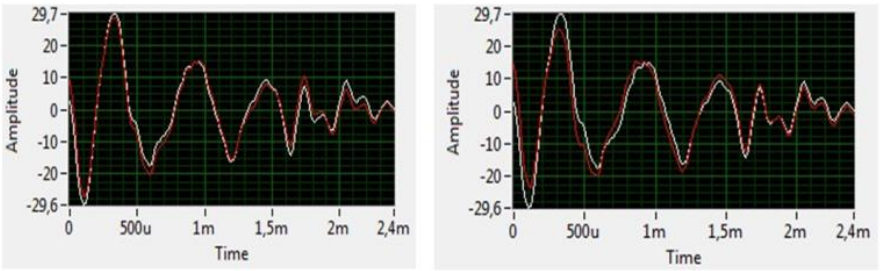


Figure 43. Symlet_08 wavelet (left) and db_06 (right) matching with the original signal

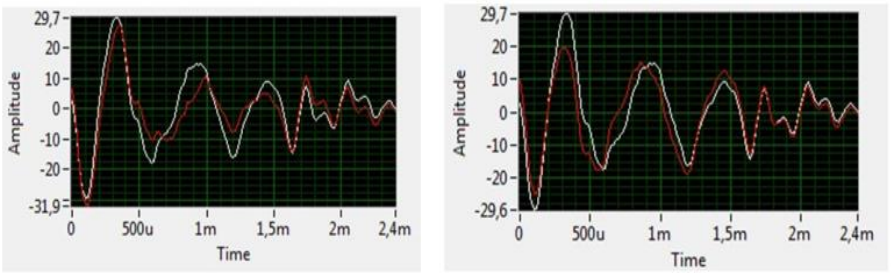


Figure 44. db_04 wavelet (left) and Symlet_05 (right) matching with the original signal

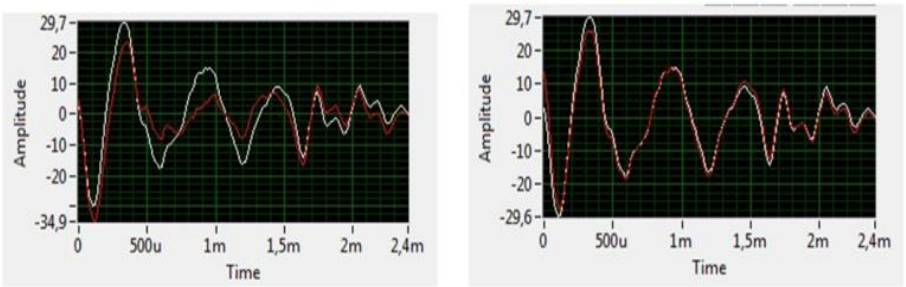


Figure 45. db_08 wavelet (left) and new-designed (right) matching with the original signal

Table 7. E/S values of the wavelets by the number of measurements

E/S values	1st	2nd	3rd	Mean
Sym5	99.65	101.12	100.67	100.48
Sym8	121.85	122.81	123.05	122.57
db04	88.18	87.21	86.84	87.41
db06	117.01	115.94	116.07	116.34
db08	84.28	82.86	84.50	83.88
New_Designed	**155.12**	**154.18**	**155.23**	**154.84**

It is observed that new-designed wavelet provided the highest value that indicated to be the most efficient wavelet for both fault detection and fault size estimation.

Fault detection procedures based on time-frequency methods usually rely on visual observation of contour plots. It is known that if the wavelet matches well with the shape of the signal at a specific scale and location a large wavelet scalogram coefficient value is obtained. The scalogram (Figure 46) shows the time-frequency location of transients.

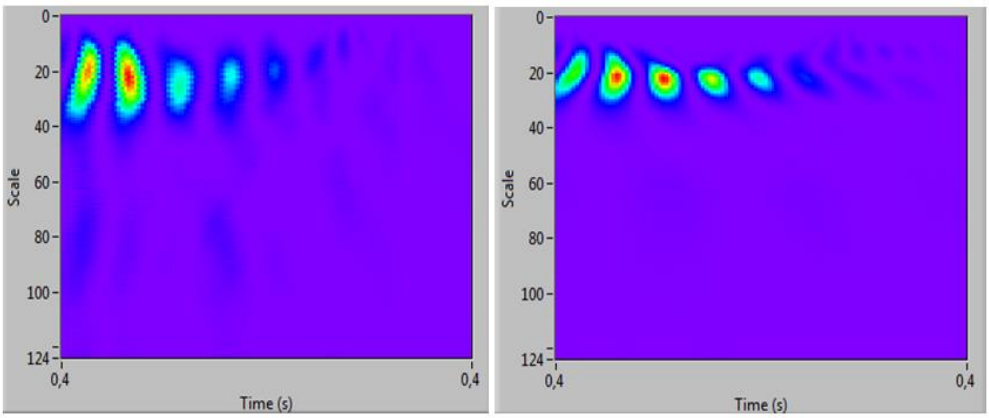


Figure 46. Comparison of the scalograms of the Daub_08 (left) and the new-designed wavelet (right)

Comparing the scalograms it is noticed that the new-designed wavelet provided a more realistic result and provided better energy concentration. The scalogram calculated by the matching wavelet provided better time localization, but its frequency localization is less accurate than that of the designed wavelet.

6.3 Validation of the grinding fault width estimation

To verify the precision of the measurements optical and contact inspection is applied to reveal the exact geometrical size of the ground defect on the outer race. Garant MM1-200 video microscope and Mahr MMQ 200 contact equipment are used to analyze and measure the grinding marks. (Figure 43).

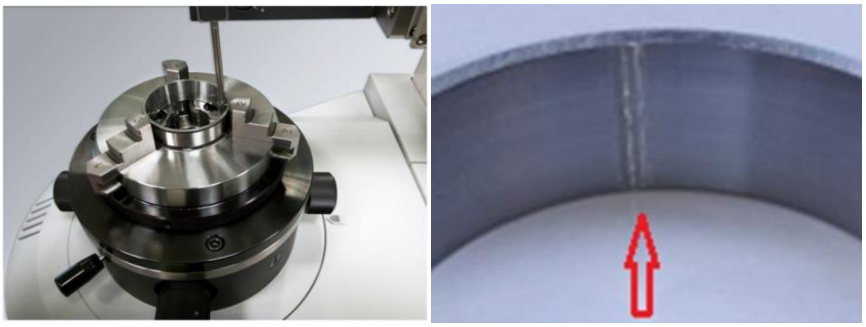


Figure 47. Measurement of the outer ring of the tapered roller bearing with grinding defect of 1.2492 mm (left) and its enlarged image (right)

As Table 8 shows, the new-designed wavelet provides the most accurate geometrical fault determination with the less deviation from the measured value. In this experiment four different widths (0.6311 mm, 1.2492 mm, 1.4751 mm, 1.6236 mm) were measured and the method provided the same result, namely the new design wavelet produced the highest accuracy.

Table 8. Comparison of the calculated and measured fault widths of the faults in case of the 1.2492 mm fault

Wavelet	**Calculated width (mm)**	**Width deviation (%)**
Sym_05	1.2875	3.07
Sym_08	1.2713	1.77
Daub_04	1.2930	3.51
Daub_06	1.2765	2.19
Daub_08	1.2975	3.87
New_designed	**1.2639**	**1.18**

6.4 Results

Wavelet design was established in Labview software. It createsda FIR filter that determines the shape of the designed wavelet. It was shown that the new wavelet can be used for detecting transient pulses generated in a bearing. It was found that the new wavelet had higher efficiency to reveal minor faults on the bearings. Five typical wavelets were compared to the new-designed wavelet. Optical and contact measurements validated that higher Energy-to-Shannon Entropy values of the wavelets resulted in more precise determination of the geometrical defects. Wavelet creation is executed by the Chapa and Rao's method as well, where the Meyer wavelet amplitude and phase spectra are matched independently to the signal. The calculation provided the same result in practice that in Labview application.

7 Complex Morlet wavelet design with global parameter optimization

7.1 Complex Morlet wavelet for flexible analysis

In this experiment complex Morlet wavelet was designed to effective rectangular shape griding fault analysis in noisy environment (SNR<30 dB) where the additional noise was simulated by Duffing oscillator.

A complex Morlet wavelet is defined by:

$$\psi(x) = \frac{1}{\sqrt{\pi\beta}} e^{2i\pi f_0 x} e^{-\frac{x^2}{\beta}}$$

and its value depends on two parameters: β is the bandwidth and f_0 is the central frequency of the wavelet. It has a real part and an imaginary part (Figure 48). The complex Morlet wavelet after parameter optimization is able to reach higher E/S ratio than the real-value Morlet wavelet or other discrete wavelets (e.g. Symlet, Daubechies) where the filter coefficients are pre-determined and there is no possibility to change them.

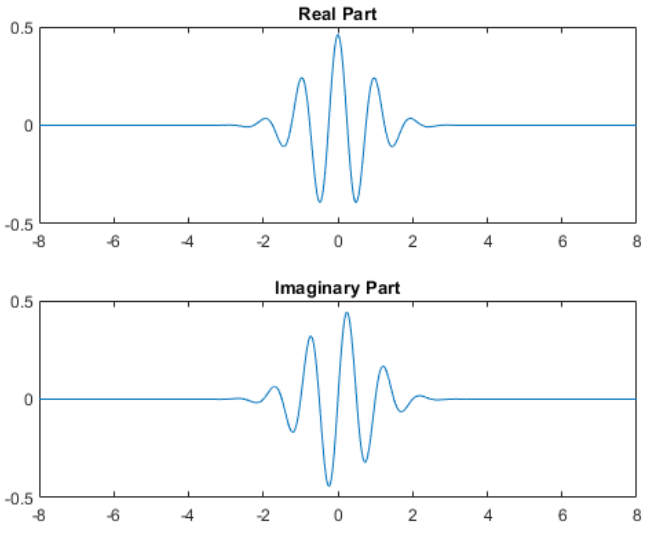


Figure 48. Real part and imaginary part of the complex Morlet wavelet

7.2 Wavelet parameter optimization with genetic algorithm

Genetic algorithm (GA) is commonly used to generate sufficiently precise solutions to optimization and search problems by relying on operators such as mutation, crossover and selection. GA is applied to optimize the Morlet wavelet filter parameters, the center frequency and the bandwidth.

For starting the GA operation, the initial population can be randomly determined (initialization of the population). The size of the genotype population is determined by experience. The size of population is set to 50 according to the experiences (Table 9).

The fitness function estimates how good an individual in the current population (fitness evaluation). In this case, the objective function is the E/S value. The fitness of each chromosome can be calculated by applying linear ranking of objective function value. The chromosome with higher fitness value is selected at greater probability than that with lower fitness value.

The selection operation selects two individuals. The stochastic universal selection is applied in this study.

Crossover is a probabilistic process that exchanges information between two parent chromosomes for generating two child chromosomes. A single-point crossover is applied, the typical crossover probability value is 0.75.

Mutation is used to avoid local convergence of the GA. Mutation occurs with typical mutation probability of 0.03 value.

The maximum number of generations was adopted as the termination criterion for the solution process. Number 60 is chosen in this experiment.

Table 9. Parameters of GA for complex Morlet wavelet optimization

Parameters	**Parameter values**
Population scale	50
Probability of mutation	0.03
Probability of crossover	0.75
Terminal iteration times	60
Length of binary code	
f_0	18
β	16

With genetic algorithm the center frequency and the bandwidth of the complex Morlet wavelet (Figure 49) were optimized to $f_0 = 2118\,Hz$ and $\beta = 642\,Hz$. For wavelet parameter optimization some important requirements should be satisfied.

Wavelet is needed to satisfy the admissibility condition [60], [62]

$$C_\Psi = \int_{-\infty}^{\infty} \frac{|\Psi(\omega)|^2}{|\omega|} d\omega < \infty$$

where Ψ is the Fourier transform of Ψ, ω is the angular frequency. In practice, Ψ will always have sufficient decay so that the admissibility condition reduces to the requirement

$$\int_{-\infty}^{\infty} \Psi(t)dt = \Psi(0) = 0$$

Morlet wavelet does not satisfy this zero-mean requirement. However, the mean value can become small enough if the term f_0/β is sufficiently large. When $f_0/\beta > 3.5$ then

$$\Psi(0) < 4.2146 \cdot 10^{-8},$$

therefore, the admissibility condition is approximately satisfied.

The upper cut-off frequency of Morlet wavelet filter is $f_0 + \beta/2\ Hz$. According to the Shannon Sampling Theorem, the upper cut-off frequency must satisfy the following condition

$$f_0 + \frac{\beta}{2} < \frac{f_s}{2.56}$$

where f_s is the sampling frequency.

The lower cut-off frequency of Morlet wavelet filter is $f_0 - \beta/2\ Hz$. To reduce the interfering effects of the harmonics, the lower cut-off frequency has to be sufficiently large. Therefore, the lower cut-off frequency has to satisfy

$$f_0 - \frac{\beta}{2} \geq N \cdot f_r$$

where f_r is the rotational speed and $N = 35$ is chosen in this experiment.

To fully extract the impulsive feature, the bandwidth of the Morlet wavelet filter must be adequately wide. In this experiment, we choose the bandwidth

$$\beta > 3f_b,$$

where f_b is the ball-pass frequency outer ring (BPFO) in this experiment, because the manufacturing fault is on the outer ring of the tapered roller bearing.

Complex Morlet wavelet parameter optimization is summarized in the conditions:

$$\beta < f_0/3.5$$
$$f_0 + 0.5\beta < 0.39f_s$$
$$f_0 - 0.5\beta \geq 35f_r$$
$$\beta > 3f_b$$

Based on the wavelet inequalities the domain of the optimization process is given by the following: $\beta \in [602\ Hz, 706\ Hz]$, $f_0 \in [1805\ Hz, 9631\ Hz]$.

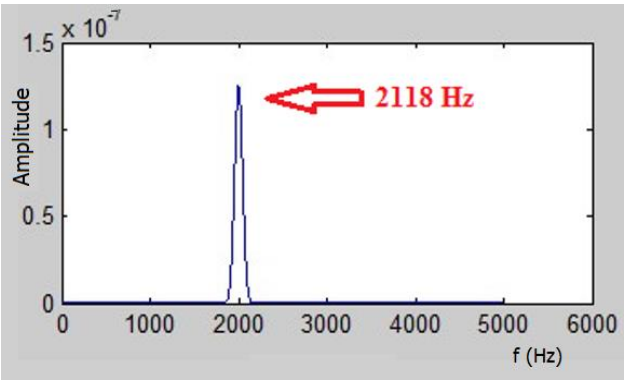


Figure 49. The optimized complex Morlet wavelet

Ten different wavelets are considered for the present study. For the best diagnosis the wavelet with the highest E/S ratio is obtained. The comparison of the wavelets on the basis of their E/S values are in Table 10.

Table 10. Calculated values of Energy-to-Shannon-Entropy ratio of the ten wavelet functions

E/S	OR1	OR2	OR3	OR4	Mean
Sym2	59.86	82.69	103.26	112.52	89.58
Sym5	69.58	98.31	114.01	112.34	98.56
Sym8	79.43	117.24	120.56	113.51	107.69
db02	59.86	82.69	103.26	112.54	89.59
db06	72.64	95.03	116.37	121.17	101.30
db10	77.69	101.8	121.01	123.05	105.89
db14	85.81	118.47	123.14	124.31	112.93
Morlet	114.01	194.25	144.72	142.14	148.78
Cmor	138.26	235.19	247.45	227.89	**212.12**
Meyer	97.56	160.37	127.08	103.33	122.09

It is observed that the optimized complex Morlet wavelet provides the highest E/S value that indicates the Morlet wavelet is the most efficient for both fault detection and fault size estimation.

7.3 Autocorrelation envelope power spectrum for signal enhancement

Autocorrelation is a tool for finding repeating patterns, such as the presence of a periodic signal obscured by noise. In order to enhance the signal in the frequency band, an autocorrelation enhancement algorithm is proposed. It increases the involved periodic impulsive feature so periodic impulsive signal

component while stochastic noise signal component is weakened. The autocorrelation of the wavelet transform is expressed as [48]:

$$r(l) = E\left[WT_{f_0,\beta}(k)WT^*_{f_0,\beta}(k+l)\right]$$
$$l = 0,1,2,\dots,k-1,$$

where l is the lag index, $E[\cdot]$ denotes the mathematical expectation operator.

7.4 Identification of the fault frequencies

Fault frequencies are useful indicators of bearing health condition monitoring. By using the actual parameters of the bearing and operational conditions the fault frequencies are calculated. BPFO frequency should be used for fault detection because the grinding faults are on the outer ring of the bearings.

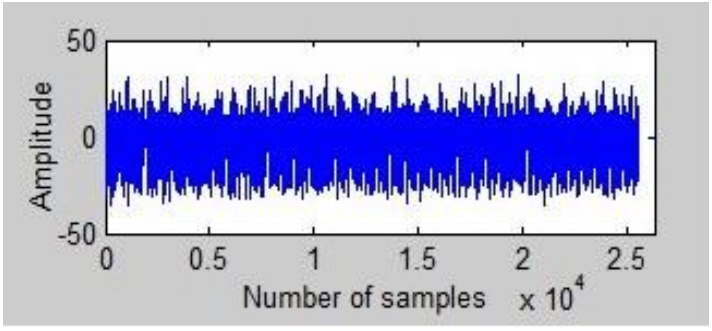


Figure 50. Time domain raw signal of the ORD

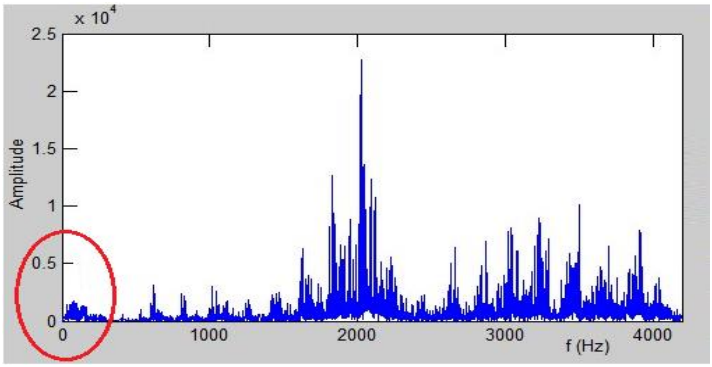


Figure 51. FFT spectrum of the ORD

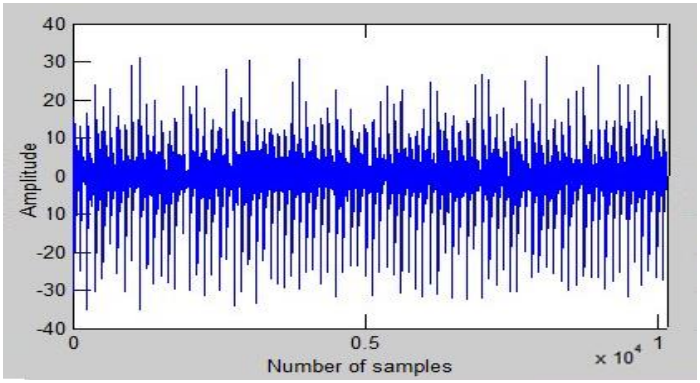


Figure 52. Time domain spectrum of the ORD by the parameter optimized complex Morlet wavelet

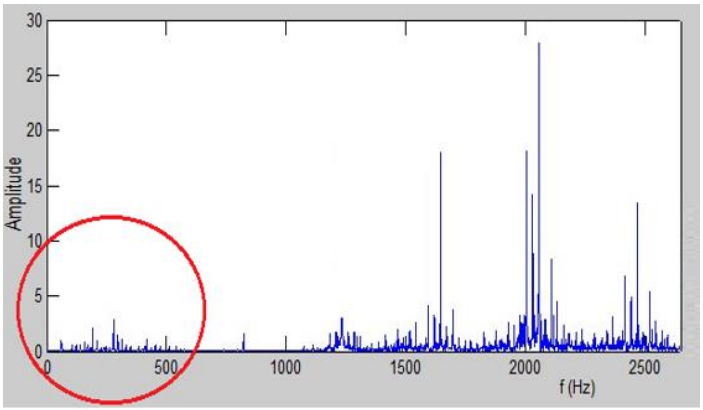


Figure 53. FFT spectrum of the ORD filtered by the parameter optimized complex Morlet wavelet

Figure 50 presents the raw signal in time domain, then Figure 51 shows its vibration signature of outer race defect of bearing with 1.2492 mm fault width on its outer ring. The highest transient impulse related energy content of the burst occurs at 2.09 kHz which is the relevant structural resonance frequency of the test-rig as a mechanical system caused by the periodic transients. The spectrum was determined on the basis of each outer ring with different fault sizes in three separate measurements for high accuracy. Each separate measurement showed similar manner and signal peak of 2.09 kHz. After filtering the raw signal by the optimized complex Morlet wavelet it is much more visible (Figure 52) that the outer race defect causes transients with 5 ms periodicity which are in connection with the calculated 206.18 Hz BPFO fault frequency. In the FFT spectrum it is not easy to find the outer race defect frequency at 206.18 Hz (Figure 53) because of its noisy manner and low amplitude (red circle). Spectrum offers clearer image after filtering by the optimized complex Morlet wavelet compared to the raw FFT spectrum. However, it is not applicable for detailed BPFO analysis yet, because the frequency around the 206.18 Hz is

strongly meshed (red circle) in the spectrum. Therefore, further techniques such as MRA, envelope spectrum and autocorrelation enhancement techniques are applied to analyse this low frequency range for precise diagnosis.

Multi-resolution analysis is executed down to 3rd level in order to obtain precise and detailed frequency analysis. Figure 54 presents the wavelet decomposition tree and the multiresolution analysis from signal to reveal BPFO frequency for outer race grinding fault identification.

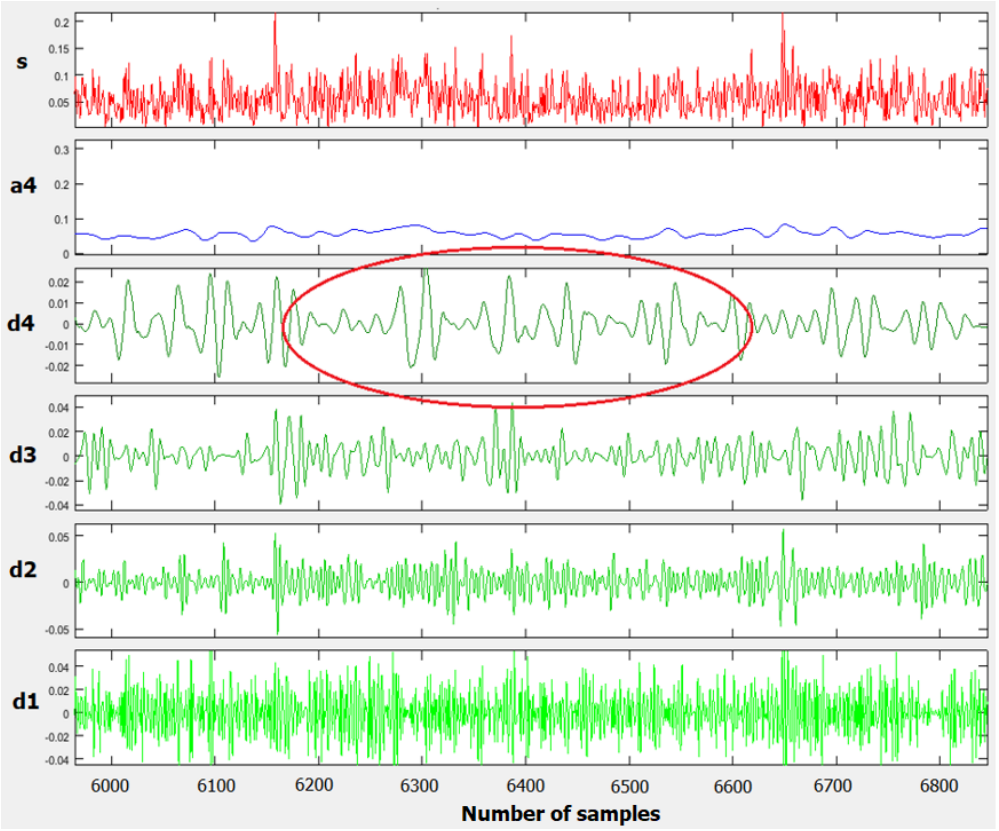


Figure 54. Multi-resolution analysis of the signal for finding transients

Figures 55-57 show the comparison of the envelope spectrum of the grinding fault with three chosen wavelets (Symlet-5, Morlet, Cmor) from the ten wavelets for illustration. The Morlet wavelet provided better feature extraction, therefore higher BPFO peak values, than the Symlet-5 wavelet. However, the optimized complex Morlet wavelet provided the highest peak value for outer race fault identification.

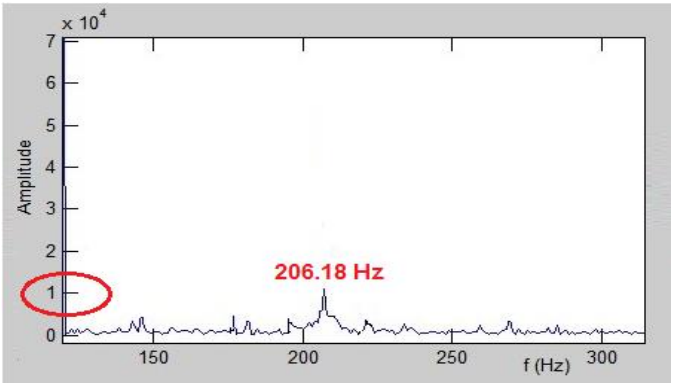


Figure 55. Envelope spectrum of the ORD filtered by Symlet-5 wavelet

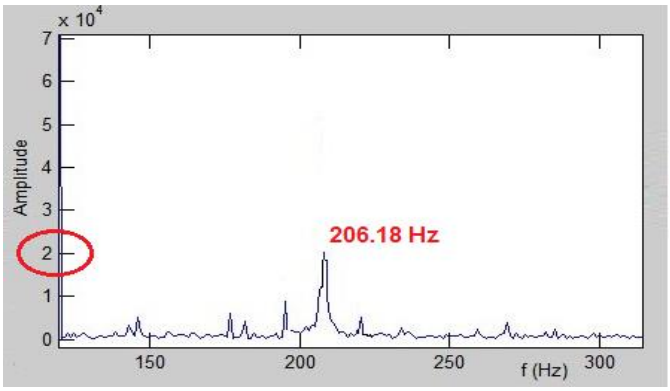


Figure 56. Envelope spectrum of the ORD filtered by Morlet wavelet

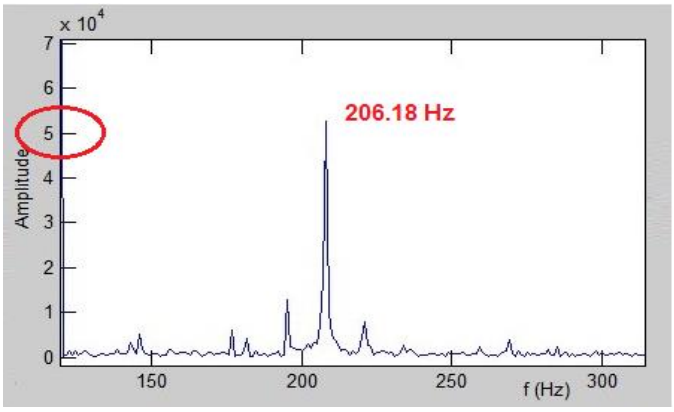


Figure 57. Envelope spectrum of the ORD filtered by the parameter optimized Cmor wavelet

Using the autocorrelation enhancement, it is clear from Figure 58 that the non-cyclic frequencies are suppressed in the vibration spectrum and the signal-to-noise ratio increased providing more efficient fault diagnosis.

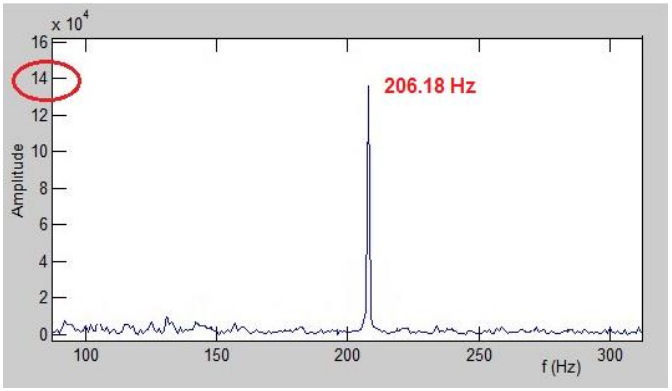


Figure 58. Envelope spectrum of the ORD filtered by the parameter optimized Cmor wavelet with autocorrelation enhancement

Comparing the scalograms of the Symlet-5, Morlet and the optimized complex Morlet wavelet it is visible that the optimized complex Morlet wavelet provides the best result for diagnosis and fault size estimation (Figures 59-61).

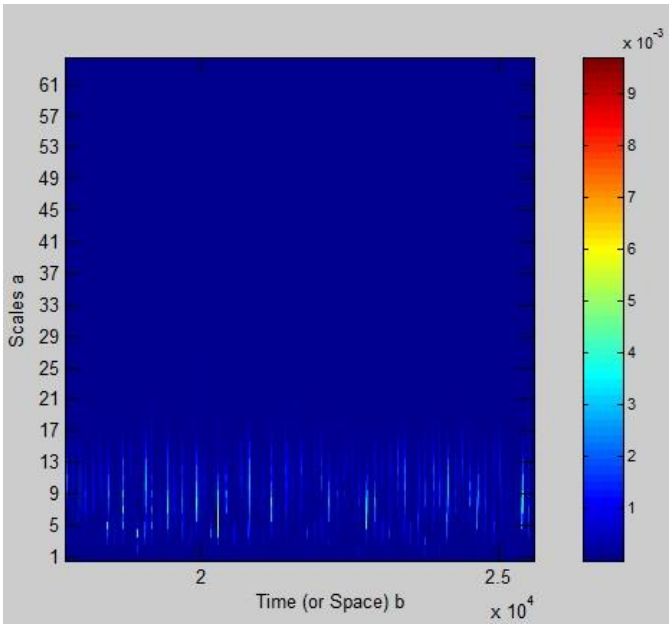


Figure 59. Scalogram of the Symlet-5 wavelet of the ORD

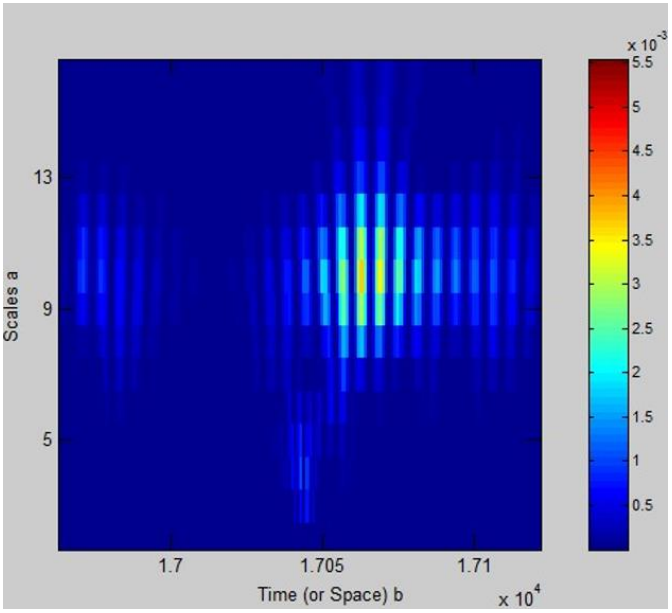


Figure 60. Scalogram of the Morlet wavelet of the ORD

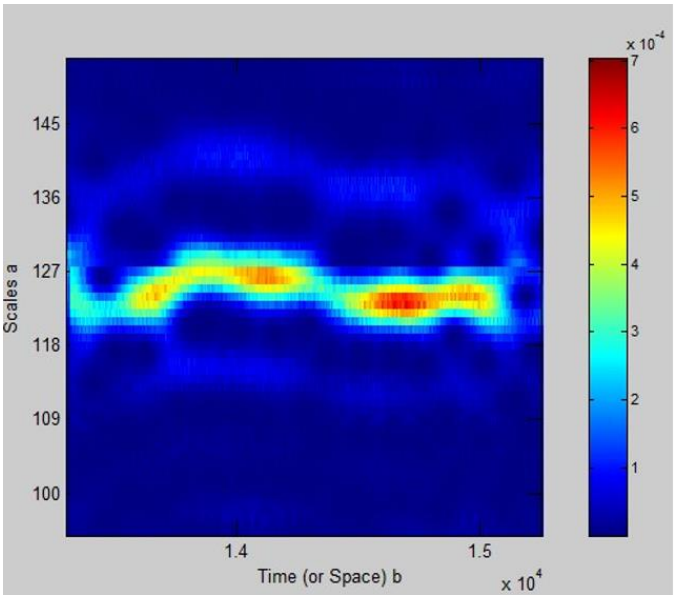


Figure 61. Scalogram of the optimized Cmor wavelet of the ORD

7.5 Fault size estimation based on scalogram of Cmor filter

Analysing the transients, the typical points where the roller enters into and exits from the defect can be defined as Figure 62 shows ([5], [23]).

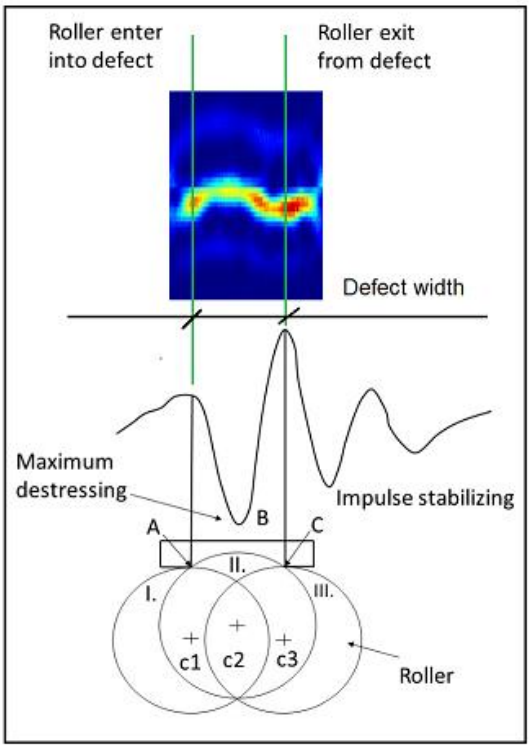


Figure 62. Entry and exit point analysis with wavelet scalogram

At point *A*, the roller strikes the groove base with high impact which results in re-stressing and high impulse in signal. When the roller comes in contact with the point *B* it again generates high amplitude in the signal and beyond the point progressive decrease in amplitude of signal is observed due to elastic damping of bearing element. At point *C* roller comes out of the groove and creates high amplitude.

For peak detection in the scalogram, local maxima finding was applied. With this procedure entry and exit points can be exactly determined for fault width measurement (Figure 63). This technique is applicable for fault width measurement.

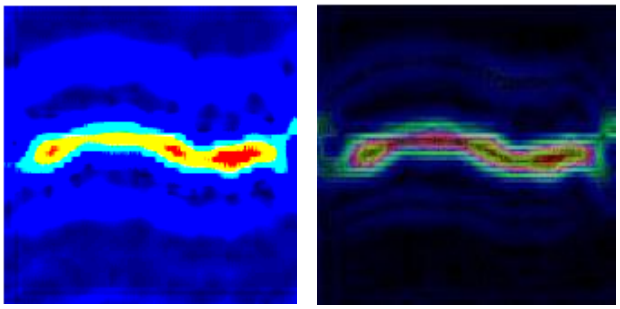


Figure 63. Scalogram after contrast enhancement (left) and then after Sobel edge detection (right)

Compared to the different wavelets the results are in the Table 11. The optimized complex Morlet (Cmor) wavelet provides the most accurate result with only 1.19% deviation from the measured value by Garant MM1-200 video microscope. In the experiment four different widths (0.6311 mm, 1.2492 mm, 1.4751 mm, 1.6236 mm) were measured and the method provided the same result that the Cmor wavelet was the most accurate. Maximum values of the wavelet coefficients were revealed by maximum searching. These areas are represented with red in the scalogram.

Table 11. Comparison of the calculated and measured fault widths of the ORD

Wavelet	Calculated width (mm)	Width deviation [%]
Sym2	1.277	2.24
Sym5	1.269	1.56
Sym8	1.2713	1.77
db2	1.277	2.24
db6	1.2765	2.19
db10	1.2681	1.52
db14	1.2679	1.5
Morlet	1.266	1.38
Cmor	**1.264**	**1.19**
Meyer	1.2674	1.46

7.6 Results

With conventional time domain and frequency domain techniques the vibration signal was analysed. Measurements were executed in noisy condition, where the low signal-to-noise ratio made difficult the feature extraction. An electromechanical shaker as chaotic Duffing oscillator was applied to create additional stochastic vibration. The purpose was to obtain the useful components from the vibration spectrum and enhance them for more precise fault detection and fault width estimation. Wavelet transform was used for

detecting the transients in the spectrum because it had the capability to detect the sharp edges caused by the roller and fault interaction during the rotation of the bearing. Complex Morlet wavelet was designed, its center frequency and bandwidth were optimized by genetic algorithm. The efficiency of the designed complex Morlet wavelet was compared to nine different wavelets, namely Symlet-2, Symlet-5, Symlet-8, Daubechies-2, Daubechies-6, Daubechies-10, Daubechies-14, Morlet and Meyer wavelets. Energy-to-Shannon-Entropy ratio criteria was used for more effective comparison. The designed wavelet provided the highest efficiency. Without filtering, the FFT spectrum was blurred by the noise making impossible to find the local BPFO fault frequency of the defected bearing. Wavelet filtering through multi-resolution analysis down to 3rd level with the best-suited wavelet provided clear vibration spectrum with less noise, higher signal-to-noise ratio. However, it was needed to further emphasize the cyclic manner of the signal by using autocorrelation enhancement. The amplitude value of the BPFO frequency improved and the non-cyclic frequencies were successfully suppressed. Finally, the fault width was determined from the vibration signature by using the optimized complex Morlet wavelet. The fault width measurement procedure was executed with the other nine wavelets for the purpose to compare them to the optimized wavelet. Analysing the spectrogram with Sobel edge detection and local maxima peak detection of the wavelet coefficients, the entry point and the exit points were exactly determined where the roller entered into and exited from the defect. Results were verified by optical and contact measurements. The deviation was 1.19 % in case of the optimized complex Morlet wavelet.

8 Bearing grinding fault classification based on wavelet decomposition

8.1 Fault classification with machine learning methods

Machine learning algorithms build a mathematical model based on sample data, known as training data, in order to make predictions or decisions without being explicitly programmed to perform the task. Classification algorithms are used when the outputs are restricted to a limited set of values. The most frequently applied machine learning methods are Support Vector Machines (SVMs), Artificial Neural Networks (ANNs), Adaptive Neuro-fuzzy Inference System (ANFIS), decision trees, deep learning methods, Bayesian networks, genetic algorithms.

In the experiments the special bearing manufacturing grinding defects were classified by SVM and ANN which were proven to be efficient for bearing diagnosis. It is important to detect and distinguish the bearing defects efficiently in order to control the process and to ensure the sufficient quality of the product. Using machine learning methods for the bearing fault classification, it is possible to exclude the human intervention from the technological process, therefore it can be faster and more efficient. The goal is to develop a procedure which is effective in classification of these faults.

8.2 Bearing grinding faults and feature vectors for fault classification

Five types of bearing defects, the outer-ring (OR), inner ring (IR), roller fault (RF), back-surface fault on the roller (IRB), multi-fault (MF) defects are analysed which are the most common manufacturing bearing defects caused by the grinding process.

The geometrical parameters of the bearing defects can be seen in the Table 12.

Table 12. Geometrical parameters of bearing defects

Type	Width [mm]	Depth [µm]
OR1 defect	0.6311	6.5
OR2 defect	1.2492	33.6
OR3 defect	1.4751	42.3
OR4 defect	1.6236	51.4

Type	Width [mm]	Depth [µm]
IR1 defect	0.7187	7.9
IR2 defect	1.3945	36.8
IR3 defect	1.5268	46.2
IR4 defect	1.7943	58.3

Type	Width [mm]	Depth [µm]
RF1 defect	0.6418	6.8
RF2 defect	1.3245	35.2
RF3 defect	1.5512	48.3
RF4 defect	1.9425	71.2

Type	Width [mm]	Depth [µm]
IRB1 defect	0.6257	12.1
IRB 2 defect	1.6432	14.7
IRB 3 defect	2.1719	19.3
IRB 4 defect	2.8546	15.2

It is necessary to find parameters which constitute a suitable feature vector for classification. Two data sets of statistical parameters are defined. The first data set contains the statistical parameters related to the original measuring results and the second one contains the processed data after wavelet multiresolution analysis (MRA).

Table 13. Sample feature vectors for machine learning of 450 N axial loading

Rpm	Range	Kurtosis	Skewness	Crest	Class
300	0,574379	4,628502	-0,01291	9,450025	H
540	0,38525	3,051092	0,004352	4,51125	H
960	0,741462	3,103492	-0,00135	5,294187	H
1500	1,076	3,173338	-0,0186	5,641156	H
1800	1,212532	3,110764	-0,03074	5,551787	H
300	0,516784	5,565125	-0,04815	11,6029	OR
540	1,087148	5,367559	-0,07931	12,50741	OR
960	1,358686	5,080899	0,003699	8,457317	OR
1500	3,050263	10,21743	0,012411	10,25849	OR
1800	3,433396	11,89574	0,049938	11,62614	OR
300	1,406267	28,72231	1,062394	17,26	IR
540	4,010562	38,75714	1,280824	17,46529	IR
960	9,704434	33,5947	1,068235	15,08711	IR
1500	18,95775	30,99675	0,636304	13,47101	IR
1800	24,06174	23,22122	0,388427	11,52667	IR
300	1,169618	17,00247	0,336159	15,20818	RF
540	3,804007	33,16292	0,467517	18,45741	RF
960	9,047048	39,46635	0,889048	20,07125	RF
1500	22,02965	51,2848	0,98737	22,78385	RF

1800	28,90865	45,26531	0,84332	19,44094	RF
300	0,508437	3,6923	-0,04353	7,898066	IRB
540	1,356502	6,698953	-0,03621	11,9048	IRB
960	1,807607	5,911665	-0,04589	9,621231	IRB
1500	3,380811	8,776434	0,001007	10,71775	IRB
1800	3,889218	7,176219	-0,00471	9,558532	IRB
300	1,260107	10,3169	0,296837	11,68768	Multi
540	3,531542	14,29995	0,183552	15,9122	Multi
960	9,195838	21,28492	0,328307	17,71342	Multi
1500	19,47964	23,02062	0,31904	15,52958	Multi
1800	27,19352	17,17268	0,239156	14,50251	Multi

There were 5 classes (healthy, outer ring fault, inner ring fault, roller fault, roller back-surface fault, multi fault) which are classified. Two main parameters were changed in the experiment: engine rotation speed and the axial spanning force of the bearing. Experiments were executed from 60 to 2880 $1/min$ which is 48 different speed of revolution. Experiments were executed with different spanning force of the test-rig screw mechanism from 25 N to 450 N, namely 25, 50, 75, 100, 125, 150, 175, 200, 225, 250, 275, 300, 325, 350, 375, 400, 425, 450 N. The maximum value of the axial force by the screw mechanism of the test-rig is 450 N because of avoiding the risk of excessive Hertz-tension between the contact areas even in operation when dynamic forces are produced. Thus, 864 input vectors were created and each vectors contained 8 different calculated parameters (range, kurtosis, skewness, crest, mean, std. deviation, Parseval energy, entropy) that were 6912 data for each bearing element. Statistical data were calculated by the Labview VI. In order to obtain representative statistical data, outer-ring fault (184 pieces), inner ring fault (162 pieces), roller fault (247 pieces), back-surface fault (147 pieces) and healthy bearings (225 pieces) were used. For teaching vectors 4 main parameters are used: range, kurtosis, skewness, crest producing 3456 data. Table 13 represents a sample feature vectors for illustration of the 450 N axial force in case of certain revolution speeds. A sample training and testing vector is shown in Appendix 2.

8.3 Application of the SVM method

Support Vector Machine (SVM) is a classification and regression method which is effective in computational science. Non-Linear classifiers such as Artificial Neural Networks (ANN) and SVM can be used in supervised and unsupervised learning for bearing fault diagnostics.

Given data points belong to one of two classes, and the goal is to decide which class a new data point is in. The best hyperplane represents the largest separation or margin in other word, between the two classes. In this case the distance from the hyperplanes to the nearest data point on each side is

maximized. If such a hyperplane exists, it is known as the maximum-margin hyperplane and the linear classifier it defines is known as a maximum margin classifier.

Support Vector Machine method can be interpreted as a transformation to put the lower dimensional data to a higher dimension space. Support Vector Machine constructs a hyperplane or set of hyperplanes in a high or infinite-dimensional space, which can be used for classification and regression. The hyperplanes in the higher dimensional space are defined as the set of points whose dot product with a vector in that space is constant. SVM transforms non-separable patterns to separated patterns the existing failure or incipient failure is getting more identifiable in the higher dimensional space. The core method of SVM is to use the maximal margin method to defeat the overfitting problem that makes the model fit special data sets. [37]

Support Vector Machines uses hypothesis space of a linear functions in a high dimensional feature space. SVM can be trained with a learning algorithm from optimization theory.

Taking a set of training data $\{(\boldsymbol{x_i}, y_i)\}$ in $\mathbb{R}^n \times \mathbb{R}$ sampled the maximum margin is given as [57]:

$$\text{margin} = \arg\min d(\boldsymbol{x}) = \arg\min \frac{|\boldsymbol{x} \cdot \boldsymbol{w} + b|}{\sqrt{\sum_{i=1}^{d} w_i^2}}.$$

For calculating the SVM we see that the goal is to correctly classify all the data. For mathematical calculations we have,

$$\boldsymbol{wx_i} + b \geq 1 \quad \text{if} \quad y_i = +1;$$
$$\boldsymbol{wx_i} + b \leq 1 \quad \text{if} \quad y_i = -1$$

This boundary can be expressed with the use of the Support Vectors as:

$$(\boldsymbol{wx} + b) = 0, \quad w \in \mathbb{R}^n, \quad b \in \mathbb{R}$$

where the vector $\boldsymbol{w}$ defines the boundary, $\boldsymbol{x}$ is the input vector of dimension N and b is a scalar threshold. At the margins, where the SVs are located, the equations for classes A and B, respectively, are

$$(\boldsymbol{wx} + b) = 1 \quad \text{and} \quad (\boldsymbol{wx} + b) = -1$$

Good separation is achieved by the hyperplane that has the largest distance to the nearest training data point of any class, the functional margin. The SVM decision function is an application of the kernel function. The Lagrangian optimization method is used to obtain the optimal decision function from the training data [57]. SVM is generally suitable for two-class tasks.

As Support Vectors (SVs) correspond to the extremities of the data for a given class, a decision function can be created to specify whether a given data point belongs to either A or B. This is defined as [56]

$$f(\boldsymbol{x}) = \text{sign}(\boldsymbol{wx} + b).$$

The optimal hyperplane can be obtained as a solution to the optimization problem:

$$\tau(\boldsymbol{w}) = \frac{1}{2}\|\boldsymbol{w}\|^2 \quad \rightarrow \boldsymbol{min}$$

subject to:

$$y_i(\boldsymbol{w}\boldsymbol{x_i} + b) \geq 1, \;\; i = 1, \dots, l,$$

where l is the number of the training data.

The solution of the constrained quadratic programming (QP) optimization problem can be obtained as

$$\boldsymbol{w} = \sum_{i=1}^{l} v_i \boldsymbol{x_i},$$

where $\boldsymbol{x}_\text{i}$ are SVs obtained from training, furthermore

$$f(\boldsymbol{x}) = \text{sign}\left(\sum_{i=1}^{l} v_i(\boldsymbol{x}\boldsymbol{x_i}) + b\right).$$

In cases where the linear boundary in input spaces is not able to separate the two classes accurately, a hyperplane is created that allows linear separation in the higher dimension feature space. Vapnik suggested a way to create non-linear classifiers by applying the kernel trick to maximum-margin hyperplanes.

In the non-linear case this is achieved through the use of a transformation Φ, which transforms the data from an N-dimensional input space to Q-dimensional feature space. Here

$$f(\boldsymbol{x}) = \text{sign}\left(\sum_{i=1}^{l} v_i(\Phi(\boldsymbol{x})\Phi(\boldsymbol{x_i})) + b\right).$$

The function must be continuous and positive definite. The kernel function $K(x, y)$ is defined as:

$$K(\boldsymbol{x}, \boldsymbol{y}) = \Phi(\boldsymbol{x})\Phi(\boldsymbol{y}).$$

The decision function is accordingly modified as

$$f(\boldsymbol{x}) = \text{sign}\left(\sum_{i=1}^{l} K(\boldsymbol{x}, \boldsymbol{x_i}) + b\right).$$

The parameters v_i are used as weighting factors to determine which of the input vectors are actually Support Vectors $0 < v_i < \infty$.

The following kernel functions are used generally.

Linear kernel function:

$$K(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{x}\boldsymbol{y}$$

Polynomial kernel, where p is the power of polynomial.

$$K(\boldsymbol{x}, \boldsymbol{y}) = (\langle \boldsymbol{x}\boldsymbol{y} \rangle + 1)^{\boldsymbol{p}}$$

Radial Basis Kernel (RBF):

$$K(\boldsymbol{x}, \boldsymbol{y}) = \exp\left(-\frac{|\boldsymbol{x} - \boldsymbol{y}|^2}{2\sigma^2}\right) = \exp(-\gamma|\boldsymbol{x} - \boldsymbol{y}|^2)$$

Sigmoid Kernel:

$$K(\boldsymbol{x}, \boldsymbol{y}) = \tanh(\boldsymbol{v}\langle \boldsymbol{x}\boldsymbol{y}\rangle + c)$$

In these non-separable cases, a constraint is proposed:

$$0 < \boldsymbol{v_i} < C$$

The parameter $cost$ (C) is a penalty constant for those sample points which are wrongly separated by the optimal separation plane. Its role is to strike a proper balance between the calculation complexity and the separating error. For separable case, C is infinity while for non-separable case, it may be varied, depending on the number of allowable errors in the trained solution: few errors are permitted for high C, while low C allows a higher proportion of errors in the solution. To control the generalization capability of SVM, there are a few free parameters like the limiting term C and the kernel parameters [54].

In this content with C penalty constant the SV classification is the following

Minimize: $\frac{1}{2}\|\boldsymbol{w}\|^2 + C\sum_{i=1}^{l}\xi_i$

subject to:

$$\begin{Bmatrix} y_i(w^T x_i + b) \geq 1 - \xi_i \\ \xi_i \geq 0, \quad i = 1,2 \ldots\ldots, N \end{Bmatrix}$$

where C is the cost parameter.

C parameter controls the trade-off between classification of training points accurately and a smooth decision boundary or in a simple word, it suggests the model to choose data points as a support vector. If the value of C is large then model choose more data points as a support vector and we get the higher variance and lower bias, which may lead to the problem of overfitting.

Gamma defines how far the influence of single training example reaches. If the value of Gamma is high, then the decision boundary will depend on points close to the decision boundary and nearer points carry more weights than far away points and the decision boundary becomes unstable. If the value of Gamma is low, then far away points carry more weights than nearer points and thus our decision boundary becomes more like a straight line. The value of gamma and C should not be very high because it leads to the overfitting or it shouldn't be very small (underfitting). High value of Gamma leads to more accuracy but biased results and vice-versa. Similarly, a large value of C indicates poor accuracy but low bias and vice-versa. Thus it was necessary to choose the optimal value of C and Gamma in order to get a good fit.

8.4 Fault classification by SVM method

R Cran software is a language and environment for statistical computing and graphics [79]. A sample training and testing vectors were created to R Cran. The 20% rule was applied for SVM classification to distinguished the training data and the testing data (80% for training and 20% for testing).

Table 14. Confusion matrix for SVM with raw data

OR	IR	IRB	RF	MF	Classified
25	0	0	0	0	OR
0	24	0	0	0	IR
0	0	22	0	0	IRB
1	0	0	24	0	RF
0	0	0	0	26	MF

Table 15. Confusion matrix for SVM with data by 3rd level wavelet decomposition

OR	IR	IRB	RF	MF	Classified
27	0	0	0	0	OR
0	26	0	0	0	IR
0	0	25	0	0	IRB
0	0	0	26	0	RF
0	0	0	0	27	MF

Table 16. Results of the SVM classification

Classified element	**Classification rate with raw data without filtering**	**Classification rate with data after 3rd level wavelet decomposition**
OR	89.3 %	96.4 %
IR	85.7 %	92.8 %
IRB	78.5 %	89.3 %
RF	85.7 %	92.8 %
MF	92.8 %	96.4 %
Mean:	**86.4 %**	**93.5 %**

Effectiveness of the SVM classification is in Tables 14-16. Sigmoid kernel has been applied for SVM classification. Wavelet multiresolution analysis successfully removed the additional noise components from the vibration signature, behaved as a band-filter and enhanced the useful frequency content that involved the unique feature of faults themselves. The filtered signal provided higher classification rates. As the result shows it is a sophisticated task to classify inner race back support faults. The explanation could be that the axial force depends on the spanning of the bearing and usually it is lower than the

radial forces where higher acceleration values are generated therefore fault classification is easier due the the more energy content of the signal.

Figures 64 and 65 represent the software window with script which was created for the analysis. In Appendix 1 there are the full R script program which was written for the bearing fault diagnosis. Recognition rate is calculated:

$$\text{Recognition rate} = \frac{\text{Correct classification samples}}{\text{Total number of testing samples}} \times 100\ \%$$

Results in effectiveness of the SVM classification:
Effectiveness on the basis of raw measurement input data: 24 cases was classified properly out of 28 which means 86.7% classification efficiency of all cases for IR backfaults.
Effectiveness on the basis of 3rd level wavelet decomposition input data: 27 cases was classified properly out of 28 which means 97 % classification efficiency of all cases for OR, IR, roller, multi faults.

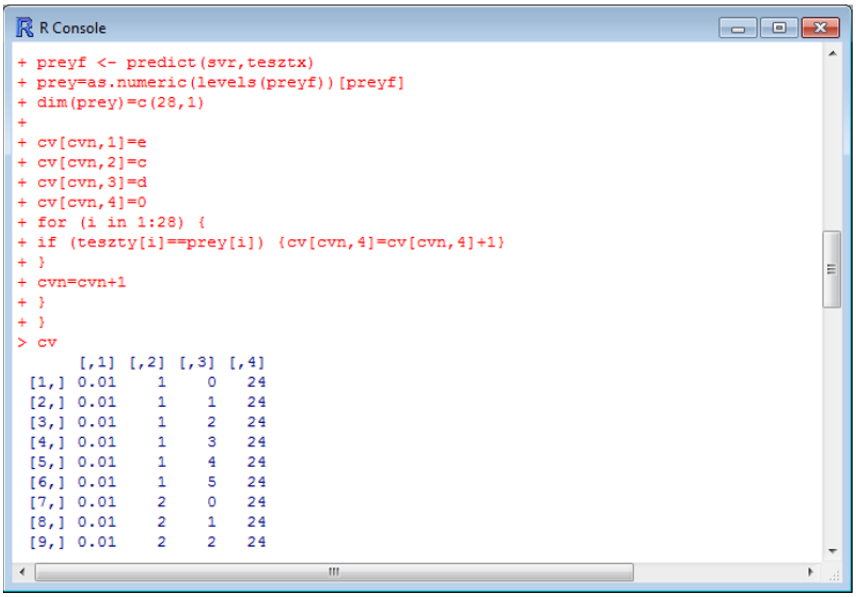


Figure 64. Input script in R for SVM classification [61]

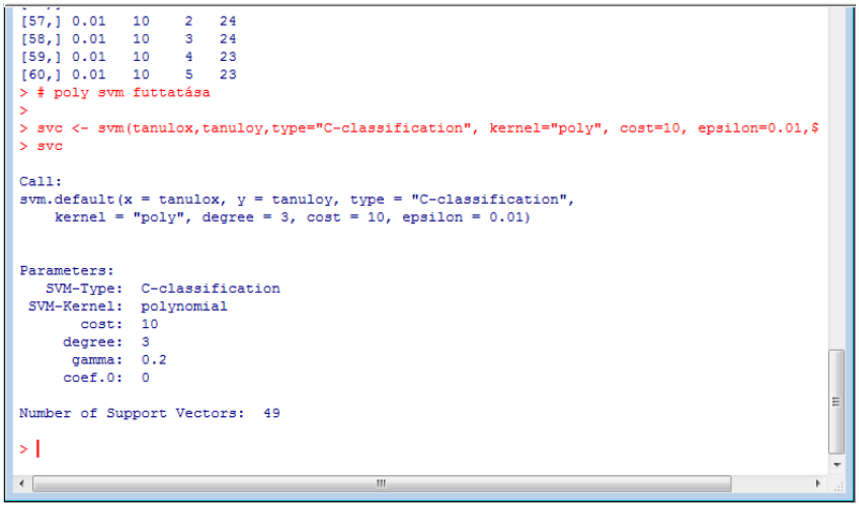


Figure 65. Result of the classification in R environment [61]

The applied kernel function was polynomial. SVM parameters were $C = 1$, $\gamma = 0.2$ in the case of the RBF kernel. Tune.svm of 1071 package was applied in the R Cran software to optimization of the parameters to obtain the best result.

By adjusting C, it was possible to define how hard or soft was the margin. With a low C, (soft margin) samples inside the margins are penalized less than with a higher C. With a C of 0, samples inside the margins are not penalized. Low value C tends to make decision surface smooth, while a high C tries all training examples correctly by giving the model freedom to select more samples as support vectors. The larger C (hard margin) is the less the final training error will be. If C increased too much, I risked losing the generalization properties of the classifier, because it will have tried to fit as best as possible all the training points (including the possible errors of the dataset). In addition, a large C, usually increased the time needed for training. C parameter was adjusted by a grid search. The range of feasible values for C was set in [0,10]. Then a coarse search was executed in laps of 1: 1,2,3,4,...,10 and looked for the average error using a 5 or 10 fold "cross validation" using the training set and kept the best value. Then the same procedure was performed but on a finer search using tune.svm of 1071 package in R.

8.5 Results

A method was presented for classification of tapered roller bearings with grinding faults from the manufacturing process using SVM classifier and optimal wavelet decomposition according to the Energy-to-Shannon entropy principle. Therefore, they could be used for modelling the real problem and for teaching

the SVM classifier. However, the raw vibration signature contained additional noise so wavelet decomposition up to the 3rd level has been applied to obtain clear signal for teaching the SVM classifier. Statistical parameters of the time domain data were calculated such as kurtosis, skewness, crest factor and range. All of these parameters were good indicators of the bearing status and present the faults of the bearing elements. Both raw data set and the filtered data set after wavelet decomposition were added to the SVM classifier which parameters were optimized. Classification effectiveness was higher with the filtered data set than with the raw data in all cases. Classification rate of 96.4 % was given for outer race fault that is remarkable for industrial application. Averaged value of the classification rate was 93.5% using wavelet decomposition. The method provided sufficient rates for the special bearing grinding fault classification in all cases.

9 Artificial Neural Network for grinding fault classification

9.1 MLP neural networks and algorithms

A multilayer perceptron (MLP) network is designed as each output neuron is connected to all the previous layer neurons and there is no connection between neurons of the same layer [64]. Perceptron computes a single output from multiple real-valued inputs by forming a linear combination according to its input weights. An MLP is a supervised network (Figure 62) requires a desired response to be trained. The net learns how to transform input data into a desired response, so it is widely used for pattern classification. An MLP can approximate virtually any input-output map. Most ANN applications involve MLP in practical applications.

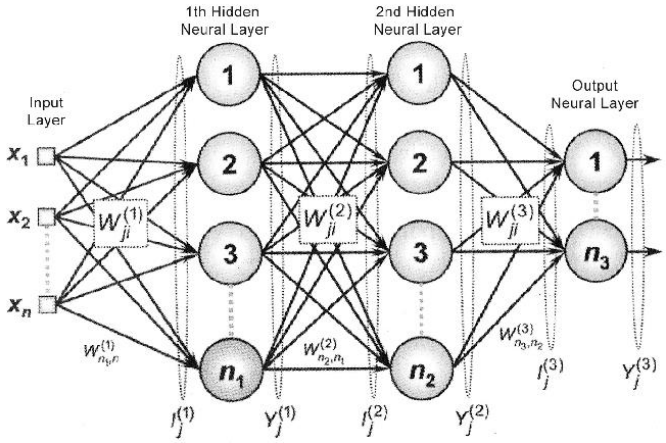


Figure 66. MLP network [16]

The input layer processing functions are all linear, but in the hidden layer non-linear sigmoidal functions such as hyperbolic tangent, logistical function or any sigmoid function are usually used.

In modelling of systems by multilayer neural network, after defining the structure, the neural network weights should be designed in a way that with applying the inputs, a very close output to the real output is achieved. This is called neural network training, which means setting the weights to decrease the errors between network output and real output. In order to increase the training speed, a linear function is usually selected for the output layer. In the following three types of training algorithms will be described which were applied in this research.

Gradient descent method (gd) designated as a first order iterative optimization algorithm for finding the minimum of a function. To find a local

minimum of a function using gradient descent, it is necessary to take steps proportional to the negative of the gradient of the function at the current point. If the steps are proportional to the positive of the gradient, it approaches to a local maximum of that function [67].

Gradient descent is based on the observation that if the multi-variable function F is defined and differentiable in a neighbourhood of a point $\mathbf{a}$, then F decreases fastest if one goes from $\mathbf{a}$ in the direction of the negative gradient of F at $\mathbf{a}$, $-\nabla F(\mathbf{a})$. It follows that, if

$$\mathbf{a}_{n+1} = \mathbf{a}_n - \gamma \nabla F(\mathbf{a}_n)$$

for $\gamma \in \mathbb{R}_+$ small enough, then $F(\mathbf{a}_n) \geq F(\mathbf{a}_{n+1})$. In other words, the term $\gamma \nabla F(\mathbf{a})$ is subtracted from a because we want to move against the gradient, toward the minimum. With this observation in mind, one starts with a guess $\boldsymbol{x}_0$ for a local minimum of F, and considers the sequence $\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \ldots$ such that

$$\mathbf{x}_{n+1} = \mathbf{x}_n - \gamma_n \nabla F(\mathbf{x}_n), n \geq 0.$$

Analysing a monotonic sequence

$$F(\mathbf{x}_0) \geq F(\mathbf{x}_1) \geq F(\mathbf{x}_2) \geq \cdots,$$

it converges to the desired local minimum. The value of the step size γ is allowed to change at every iteration. With certain assumptions on the function F (F is convex and ∇F is Lipschitz) and particular choices of γ,

$$\gamma_n = \frac{|(x_n - x_{n-1})^T [\nabla F(\mathbf{x}_n) - \nabla F(\mathbf{x}_{n-1})]|}{\|\nabla F(\mathbf{x}_n) - \nabla F(\mathbf{x}_{n-1})\|^2}$$

convergences to a local minimum can be guaranteed [17].

As for its limitations, gradient descent is relatively slow close to the minimum: technically, its asymptotic rate of convergence is inferior to many other methods. For poorly conditioned convex problems, gradient descent increasingly 'zigzags' as the gradients point nearly orthogonally to the shortest direction to a minimum point [64].

The gradient descent can take many iterations to compute a local minimum with a required accuracy, if the curvature in different directions is very different for the given function. For such functions, which changes the geometry of the space to shape the function level sets like concentric circles the convergence is slow. However, constructing and applying preconditioning can be computationally difficult.

Levenberg–Marquardt algorithm (lm) is an improvement to Newton's method [64]. Suppose that performance index of the network is the average of the sum of error squares.

$$E(\mathbf{x}) = \frac{1}{2P} \sum_{j=1}^{P} \sum_{q=1}^{m} \left\{ e_q^2(\mathbf{x}) \right\}_j$$

$$e_q = y_q^d - y_q^L;\ q = 1, \ldots, m$$

where P is the total number of training pattern, m is the number of outputs, y_q^d is the desired output, and y_q^L is the network output.

The performance index should be minimized with respect to the parameter vector $\mathbf{x}$. According to Newton's method parameter update is

$$\Delta\mathbf{x} = -[\nabla_{\mathbf{x}}^2 E(\mathbf{x})]^{-1}[\nabla_{\mathbf{x}} E(\mathbf{x})]$$

where $\nabla_{\mathbf{x}}^2 E(\mathbf{x})$ is the Hessian matrix and $\nabla_{\mathbf{x}} E(\mathbf{x})$ is the gradient.

It can be shown that

$$\nabla_{\mathbf{x}} E(\mathbf{x}) = \mathbf{J}^T(\mathbf{x})\mathbf{e}(\mathbf{x})$$

$$\nabla_{\mathbf{x}}^2 E(\mathbf{x}) = \mathbf{J}^T(\mathbf{x})\mathbf{J}(\mathbf{x}) + \mathbf{R}(\mathbf{x})$$

where $\mathbf{J}(\mathbf{x}) = \partial\mathbf{e}(\mathbf{x})/\partial(\mathbf{x})$ is the Jacobean matrix and

$$\mathbf{R}(\mathbf{x}) = \sum_{j=0}^{P} e_j(\mathbf{x})\nabla_{\mathbf{x}}^2 E(\mathbf{x})$$

For the Gauss-Newton method it is assumed that $\mathbf{R}(\mathbf{x}) \cong 0$, and then Hessian matrix becomes

$$\nabla_{\mathbf{x}}^2 E(\mathbf{x}) \cong \mathbf{J}^T(\mathbf{x})\mathbf{J}(\mathbf{x})$$

In order to prevent the matrix $\mathbf{J}^T(\mathbf{x})\mathbf{J}(\mathbf{x})$ from being ill-condition, the weights are obtained based on the modified Levenberg-Marquardt algorithm. Then the parameter update based on this method becomes.

$$\Delta\mathbf{x} = [\mathbf{J}^T(\mathbf{x})\mathbf{J}(\mathbf{x}) + \mu_n\mathbf{I}]^{-1}\nabla_{\mathbf{x}} E(\mathbf{x})$$

where μ_n is the learning factor. For small μ_n the training algorithm becomes Gauss-Newton. Less time and epochs are needed in order to attain a definite error because of the high-speed convergence in this method [20].

Conjugate gradient algorithms require a line search at each iteration. This line search is computationally expensive, since it requires that the network response to all training inputs be computed several times for each search. The scaled conjugate gradient algorithm (scg), developed by Moller, was designed to avoid the time-consuming line search. The scaled conjugate gradient algorithm may require more iterations to converge than the other conjugate gradient algorithms, but the number of computations in each iteration is significantly reduced because no line search is performed. So, scaled conjugate gradient is a variation of a conjugate gradient method which avoids the line-search per learning iteration by using a Levenberg-Marquardt approach in order to scale the step size. By using a step size scaling mechanism, this method avoids a time-consuming line-search per learning iteration.

The scaled conjugate gradient algorithm defined by Moller [21] denotes the quadratic approximation to the error E in a of a neighbourhood point w by:

$$E_{qw}(y) = E(w) + E'(w)^T y + \frac{1}{2}y^T E''(w)y$$

Scaled conjugate gradient belongs to the class of conjugate gradient methods, which show super-linear convergence on most problems. By using a step size scaling mechanism scaled conjugate gradient avoids a time-consuming line-search per learning iteration, which makes the algorithm faster than other second order algorithms. Also we got better results than with other training methods and neural networks tested, as standard back-propagation.

9.2 Realization of ANN for the analysis

9.2.1 Feature extraction and flow-chart

The goal of the experiments is to classify the bearing faults with ANN under the same circumstances it was written in Chapter 8.2. The feature parameters were extracted by discrete wavelet multiresolution analysis. Time domain parameters after multi-resolution analysis, namely range, kurtosis, skewness and crest factor were provided for the MLP neural network using gradient-descent, Levenberg-Marquardt, and scaled-conjugate back propagation algorithms. It was important to extract the proper parameters to increase the classifying efficiency of the neural network. The procedure in flow chart can be seen in Figure 67. The table of the feature vectors is in the Appendix 2.

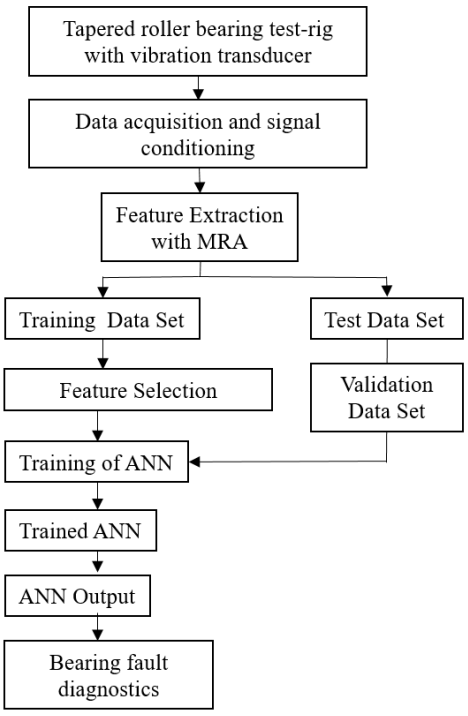


Figure 67. represents the flow chart of the diagnostic procedure

9.2.2 Designing MLP neural network for fault classification

The choice of training parameters is critical to the success of the neural network training process. The choice of these values is generally problem-dependent and there is no generic formula that can be used to choose these parameter values. Table 17 contains the used parameters of the MLP networks. Parameters were optimized with Bayesian method (bayesopt) and experimental way following the advices of the related scientific literatures.

The training of an MLP network is achieved by modifying the connection weights and biases iteratively to optimize a performance function which is the mean square error (MSE).

As for MLP, the ANN was trained and tested with gradient-descent, Levenberg-Marquardt, and scaled-conjugate back propagation training algorithms. MSE is 10^{-5}, maximum iterations (epochs) are 10000. The training process would stop if any of these conditions were met. The initial weights and biases of the network were generated randomly.

One of the critical parameters in neural network training is the speed of convergence, which is determined by the learning rate (lr). In general, it is desirable to have fast learning, but not so fast that cause instability of learning iterations (Figure 68).

The performance of the algorithm was very sensitive to the proper setting of the learning rate. If the learning rate is too small, the algorithm will take too long to converge ([64], [67]).

A too high learning rate γ makes the learning jump over minima but a too low learning rate either takes too long to converge or makes it impossible the global minima. Learning rate highly depends on the problem. The performance of the algorithm was very sensitive to the proper setting of the learning rate.

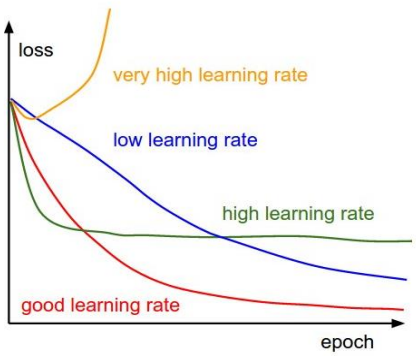


Figure 68. Learning rate scales [64]

Table 17. MLP network design parameters for training with gradient descent and Levenberg-Marquardt algorithms

Network type	Feed-forward BP
Training algorithm 1st neural network	Gradient descent
Training algorithm 2nd neural network	Levenberg-Marquardt
Training algorithm 3rd neural network	Scaled-conjugate back propagation
Training	70%
Validation	15%
Testing	15%
Learning type	Supervised
Learning rate of gd	0.05...0.80
Learning rate of lm	0.05
Learning rate of scg	0.05
Activation function	tansig
Performance function	MSE
Weights initialization	Random
Stopping criteria	
Max. iterations	10000 epochs
Limit MSE	10%
Minimum Gradient	10^{-5}
Maximum number of fails	6
Time	inf

9.3 Comparison of the algorithms

In case of all three algorithms $lr = 0.05$ is used, gradient descent algorithm misses the local minima because lr learning rate is too small, the algorithm takes long time to converge making up to 10000 iterations without stopping (Figure 66). Therefore, lr was optimized to find the shortest CPU running time to converge (4852 iterations) (Figure 71) and the highest percentage of classification ratio as Table 7 shows. Levenberg-Marquardt and scaled-conjugate algorithms performed 36 and 50 iterations which are small enough in CPU running time, however they provided high classification ratio (Figures 69 and 72). To compare the efficiency, Levenberg-Marquardt algorithm is the best for this special task of the research. Parameteres were kept the same value when the training algorithms were compared.

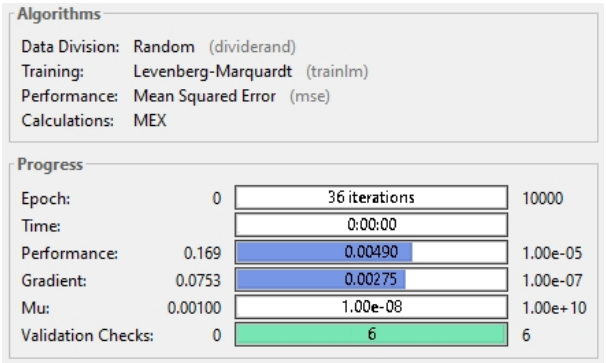


Figure 69. Progress parameters of Levenberg-Marquardt algorithm (lr=0.05)

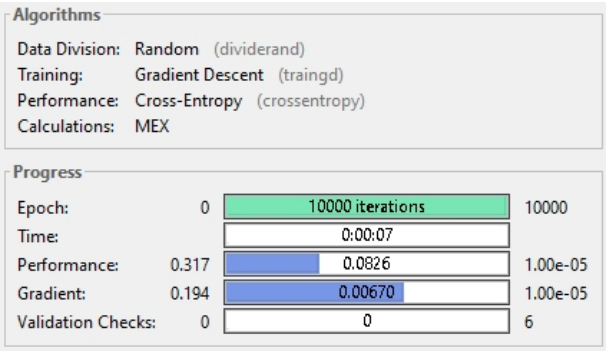


Figure 70. Progress parameters of gradient-descent algorithm (lr=0.05)

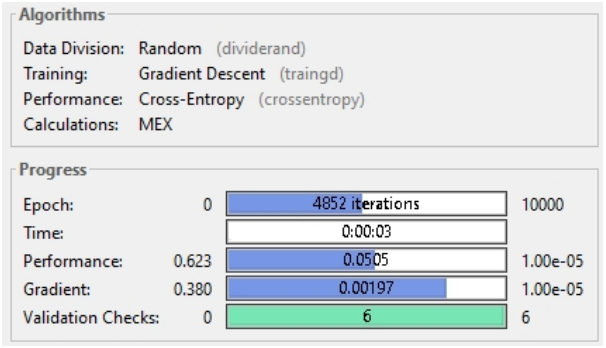


Figure 71. Progress parameters of gradient-descent algorithm (lr=0.78)

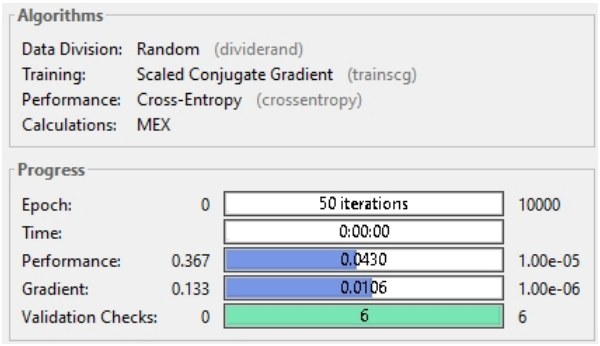


Figure 72. Progress parameters of scaled-conjugate algorithm (lr=0.05)

Validation method performance depends on which algorithm was applied for the classification. In case of Levenberg-Marquardt algorithm the performance function is the MSE value, the gradient descent and scaled-conjugate algorithms apply the cross-entropy as performance function. Figures 73-76 represent the MSE and cross-entropy parameters in the function of epochs in all three training algorithms for the training data, validation data and test data.

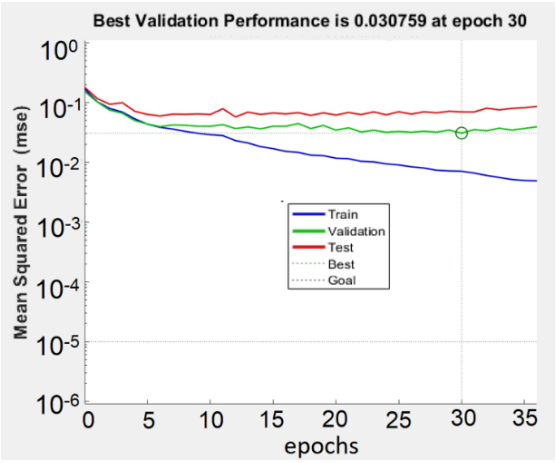


Figure 73. MSE parameter in 30 epochs of Levenberg-Marquardt algorithm (lr=0.05)

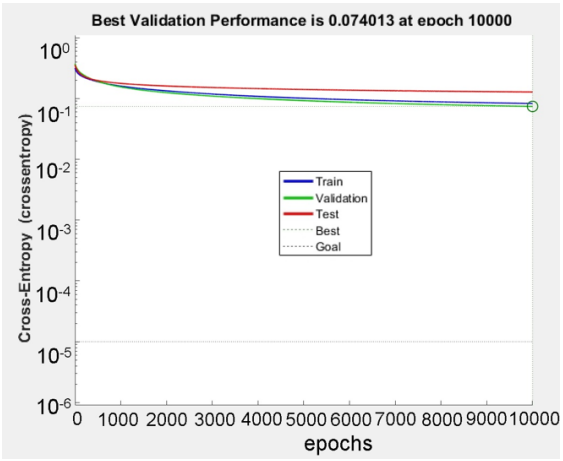


Figure 74. Cross-entropy parameter in 10000 epochs of gradient descent algorithm (lr=0.05)

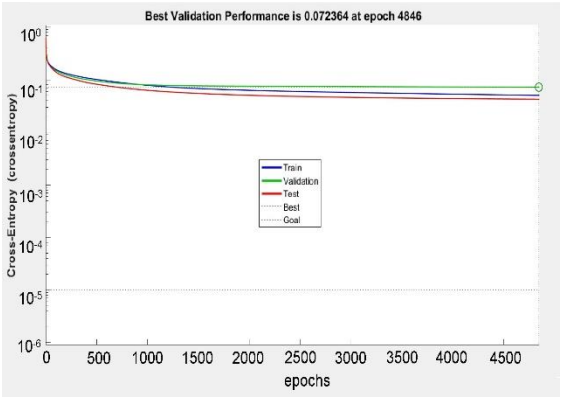


Figure 75. Cross-entropy parameter in 4852 epochs of Gradient Descent algorithm (lr=0.78)

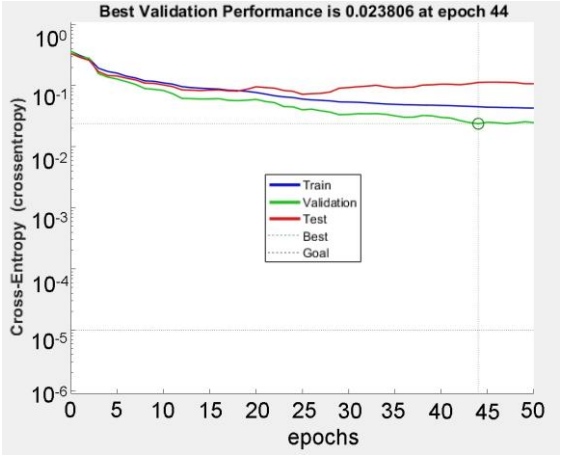


Figure 76. Cross-entropy parameter in epochs of Scaled-conjugate algorithm (lr=0.05)

Gradient values (Figures 77-80) are ideal to decrease fast in the starting epochs and then are gradually decreasing in the later epochs. Scaled-conjugate algorithm provided an even decreasing in the gradient values. In case of the Levenberg-Marquardt algorithm, the decrease of the gradient values is not even, they oscillate in the epochs. Gradient descent algorithm provided a relatively smooth gradient curve, however the number of iterations is quite high.

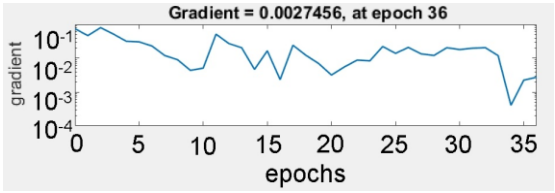


Figure 77. Gradient value in epochs of Levenberg-Marquardt algorithm (lr=0.05)

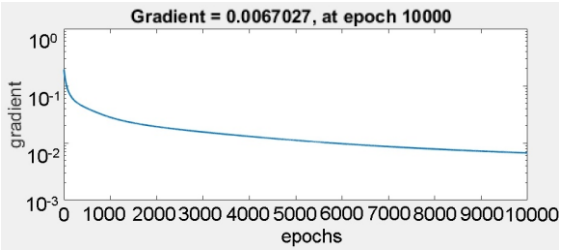


Figure 78. Gradient value in epochs of Gradient Descent algorithm (lr=0.05)

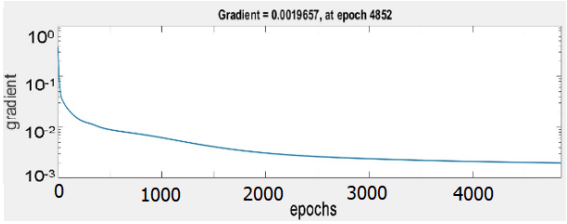


Figure 79. Gradient value in epochs of Gradient Descent algorithm (lr=0.78)

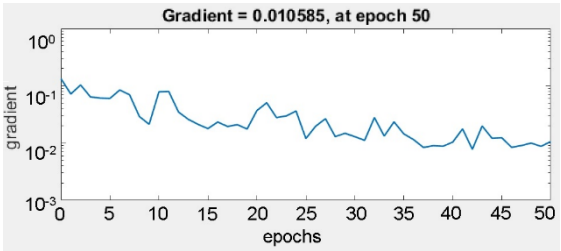


Figure 80. Gradient value in epochs of Scaled-conjugate algorithm (lr=0.05)

Tables 18-21 show the confusion matrices of the three applied algorithms in our experiment. In Table 22 the classification rates can be seen in percentage.

Overall, Levenberg-Marquardt algorithm is the best for classifying percentage as high as 92.0% the line grinding faults. In bearing diagnostics, it is very difficult to identify and classify the small back-surface defects on the rollers. Gradient descent algorithm with lr=0.05 provided only 20% classification rate, therefore it is useless to identify this defect type. Increasing the learning rate up to the optimized level of 0.78 it provided 60.0% classification ratio that is quite lower than the accuracy of the Levenberg-Marquardt algorithm. Scaled conjugate algorithm can not provide sufficient classification rate for the industrial classification of the roller back-surface defects. The changing of the learning ratio did not provide higher classification ratios in case of the Levenberg-Marquardt and the scaled conjugate algorithms.

Table 18. Confusion matrix for ANN with Levenberg-Marquardt algorithm with data by 3rd level wavelet decomposition (lr=0.05)

H	OR	IR	RF	IRB	MF	Class
45	0	0	0	2	0	H
3	18	0	0	1	0	OR
0	0	19	0	0	2	IR
0	0	0	19	0	0	RF
0	2	0	0	17	1	IRB
0	0	0	0	0	17	MF

Table 19. Confusion matrix for ANN with gradient descent algorithm with data by 3rd level wavelet decomposition (lr=0.05)

H	OR	IR	RF	IRB	MF	Class
44	2	0	0	8	0	H
4	16	0	0	8	3	OR
0	0	18	0	0	0	IR
0	0	0	17	0	1	RF
0	1	0	0	4	0	IRB
0	1	1	2	0	16	MF

Table 20. Confusion matrix for ANN with gradient descent algorithm with data by 3rd level wavelet decomposition (lr=0.78)

H	OR	IR	RF	IRB	MF	Class
46	0	0	0	3	0	H
0	16	0	0	5	1	OR
0	0	19	0	0	0	IR
0	0	0	16	0	0	RF
2	3	0	0	12	0	IRB
0	1	0	3	0	19	MF

Table 21. Confusion matrix for ANN with scaled-conjugate algorithm with data by 3rd level wavelet decomposition (lr=0.05)

H	OR	IR	RF	IRB	MF	Class
44	0	0	0	2	0	H
1	16	0	0	4	1	OR
0	0	18	0	0	1	IR
0	0	0	17	0	0	RF
3	3	0	0	14	0	IRB
0	1	1	2	0	18	MF

Table 22. Classification rates of ANN classification with the applied training algorithms

Class	Gd (lr=0.05)	Gd (lr=0.78)	L-M (lr=0.05)	Scg (lr=0.05)
OR	80.0 %	80.0 %	90.0 %	80.0 %
IR	94.7 %	100.0 %	100.0 %	94.7 %
RF	89.5 %	84.2%	100.0 %	89.5 %
IRB	20.0 %	60.0 %	85.0 %	70.0 %
MF	80.0 %	95.0 %	85.0 %	90.0 %
Mean	**72.8 %**	**83.9 %**	**92.0 %**	**84.8 %**

Gd: gradient descent algorithm
L-M: Levenberg-Marquardt algorithm
Scg: scaled conjugate algorithm

For real-time applications it is very important to perform the algorithms in short CPU time (Table 23). It is more important in real-time diagnostics process.

Table 23. CPU times of ANN classification with the applied training algorithms

Class	Gd (lr=0.05)	Gd (lr=0.78)	L-M (lr=0.05)	Scg (lr=0.05)
Time	7813 ms	3248 ms	314 ms	548 ms

9.4 Results

In this experiment, five type of bearing defects were classified by an MLP-ANN: outer-ring fault, inner-ring fault, roller fault, roller back-surface fault and multi-fault Multi-fault was the combination of the roller fault and the roller back-surface fault.

Three different training algorithms were applied: Levenberg-Marquardt, gradient descent and scaled-conjugate gradient. The feature extraction for the artificial neural network was executed by MRA. Feature vectors after MRA consisted of four main statistical parameters of the vibration signal; range, kurtosis, skewness, crest factor. Network was planned by considering the topology, layers, activation functions, number of neurons, training functions and

the learning rate. Parameteres were optimized with Bayesian optimization and experimental way. Parameteres were kept the same value when the training algorithms were compared. Gradient descent algorithm with lr=0.05 provided only 20% classification rate, therefore it is useless to identify this defect type, therefore the learning rate was optimized to reach the proper speed of convergence and high pattern recognition efficiency for the classification. Levenberg-Marquardt algorithm was the best for classifying percentage as high as 92.0% the linear grinding faults. Increasing the learning ratio to the optimized level of 0.78 it provided 60.0% classification ratio that is quite lower than the accuracy of the Levenberg-Marquardt algorithm. Scaled conjugate algorithm could not provide sufficient ratio for the industrial classification of the roller back-surface defects. The changing of the learning ratio did not provide higher classification ratios in case of the Levenberg-Marquardt and the scaled conjugate algorithms. CPU time was analysed to measure the shortest running time, Levenberg-Marquardt provided 314 ms that is small enough for practical applications compared to the other training algorithms.

10. New scientific results

THESIS 1

An object-oriented measurement system and method was developed to analyse certain type of manufacturing surface faults of roller bearings with rectangular shape contact areas. The main parts of the special measurement system are a designed test-rig which is suitable to analyse roller bearings of different sizes and types according to the requirements of the analysis, a vibration transducer and a data acquisition system, algorithms for surface fault detection and fault size estimation, designed wavelets and filters with developed Labview VIs for measurements and data processing. It was proven on tapered roller bearings that the measurement system and the method is effective for the detection and size estimation of the investigated type of bearing faults. It was proved with the Maximum Energy-to-Shannon Entropy Criteria that the optimized wavelets are appropriate for the analysis.

[DK3] K. Deák, T. Mankovits, I. Kocsis, Optimal wavelet selection for manufacturing defect size estimation of tapered roller bearings with vibration measurement using Shannon Entropy Criteria, Strojniski Vestnik-Journal of Mechanical Engineering, vol. 63, no. 1, 2017, p. 3–14.

Further related publications:

[DK10] K. Deák, I. Kocsis, Rotary machine and roller bearing fault measurement with FPGA embedded system, in Tavaszi Szél / Spring Wind, II. kötet Állam- és jogtudomány, közigazgatás tudomány, had- és rendészettudomány, 2014.

[DK18] K. Deák, I. Kocsis, Applied vibration measurement methods and data extraction for bearing fault diagnosis, in Proceedings of the International Scientific Conference on Advances in Mechanical Engineering, 2016, p. 122–126.

[DK20] K. Deák, I. Kocsis, Condition Monitoring by Applying Advanced Diagnostic Method, in Proceedings of the International Scientific Conference on Advances in Mechanical Engineering, 2017, p. 102–108.

[DK5] K. Deák, J. Menyhárt, L. Czégé, Defect analysis of bearings with vibration monitoring and optical methods, International Journal of Engineering and Management Sciences, vol. 3, no. 4, 2018, p. 1–12.

[DK 22] K. Deák and I. Kocsis, Applied diagnostic methods in engine and bearing health monitoring. In: 7th International Scientific Conference on Advances in Mechanical Engineering, 2019.

THESES RELATED TO WAVELET DESIGN

A discrete and a continuous wavelet were designed to detect rectangular shape manufacturing grinding faults of roller bearings and to estimate the fault sizes. It was proven with measurements on tapered roller bearings that the designed wavelet in the range from 0.63 mm to 1.63 mm is more effective than the traditional wavelets which are generally applied for bearing fault detection. The effectiveness of both methods in the fault size estimation was proven by optical and contact validation measurement methods.

THESIS 2A

A new discrete wavelet was developed using the transient signal model referring to the entry and exit interactions between the roller and the defect. It was proven on tapered roller bearings that the designed wavelet is more effective than the traditional wavelets which are generally applied for bearing fault detection.

THESIS 2B

A new complex Morlet wavelet was designed to determine the fault frequency of the investigated fault type, and to estimate the fault size in noisy environment. It was proven with measurements that the designed wavelet is more effective up to SNR=26 dB than the traditional wavelets which are generally applied for bearing fault detection.

[DK1] K. Deák, Design of Discrete Wavelet by Using Transient Model for Exact Measurement of Manufacturing Faults of Tapered Roller Bearings, Periodica Polytechnica Mechanical Engineering, vol. 63, no. 2, 2019, p. 113–122.

[DK2] K. Deák, I. Kocsis, Complex Morlet wavelet design with global parameter optimization for diagnosis of industrial manufacturing faults of tapered roller bearing in noisy condition, Diagnostyka, vol. 20, no. 2, 2019, p. 77–86.

Further related publications:

[DK8] K. Deák, I. Kocsis, Gépek károsodása által előidézett kifáradásos csapágyhibák által generált tranziens impulzusok jelfeldolgozása ablakozott fourier és wavelet transzformációk segítségével = bearing fault detection and

signal processing with STFT and wavelet transform in maintenance engineering, in Műszaki Tudomány az Észak-kelet Magyarországi Régióban, 2014, p. 67–72.

[DK13] Z. Keviczki, K. Deák, I. Kocsis, Machine condition monitoring in bearing manufacture using vibration analysis and intelligent approaches, in Proceedings of the International Scientific Conference on Advances in Mechanical Engineering, 2014, p. 55–69.

[DK14] K. Deák, I. Kocsis, Manufacturing of tapered roller bearings, defects and fault detection, in Proceedings of the International Scientific Conference on Advances in Mechanical Engineering, 2015, p. 44–53.

[DK19] K. Deák, I. Kocsis, Selection of wavelet function for detection of bearing defects by Shannon entropy, in Proceedings of the International Scientific Conference on Advances in Mechanical Engineering, 2016, p. 127–134.

[DK21] K. Deák, I. Kocsis, Wavelet design for diagnosis of tapered roller bearing manufacturing defects, in Proceedings of the International Scientific Conference on Advances in Mechanical Engineering, 2018, p. 39–40.

[DK17] K. Deák, I. Kocsis, Diagnosztikai és karbantartási vizsgálat csapágyhiba esetében wavelet transzformáció alkalmazásával, in Műszaki tudomány az Észak-Kelet Magyarországi régióban, 2016, p. 720–727.

THESIS 3

A new method was developed to classify rectangular shape manufacturing grinding defects of roller bearings. The classifications were executed with SVM method and Artificial Neural Network. Input vectors of the machine learning systems were derived from the 3rd level multiresolution decomposition of the vibration signal using four statistical parameters. It was proven with measurements that the efficiency of the new classification method satisfies the requirements set in the relevant literature. Artificial Neural Network with optimized parameters using Levenberg-Marquardt algorithm resulted in a classification efficiency of an average of 92 %. The SVM method produced a classification efficiency of 93.5 %.

[DK4] K. Deák, I. Kocsis, Support Vector Machine with Wavelet Decomposition Method for Fault Diagnosis of Tapered Roller Bearings by Modelling Manufacturing Defects, Periodica Polytechnica Mechanical Engineering, vol. 61, no. 4, 2017, p. 276–281.

Further related publications:

[DK6] K. Deák, I. Kocsis, A. Vámosi, and Z. Keviczki, Failure diagnostics with SVM in machine maintenance engineering, Annals of the University of Oradea Fascicle of Management and Technological Engineering, vol. XXIII (XIII), no. 1, 2014, p. 19–24.

[DK12] K. Deák, I. Kocsis, A. Vámosi, Z. Keviczki, Failure diagnostics with SVM in machine maintenance engineering, in IMT Oradea, Proceedings of the annual session of scientific papers: volume XIII (XXIII), 2014, p. 23–28.

[DK7] K. Deák, A. Vámosi, I. Kocsis, Csapágy meghibásodások méréstechnikája és rezgésdiagnosztikája mesterséges neurális hálók segítségével = Measurement techniques and diagnostics of bearings based on artificial neural networks, in Műszaki Tudomány az Észak-kelet Magyarországi Régióban, 2014, p. 58–66.

[DK9] K. Deák, I. Kocsis, A. Vámosi, Application of machine vision in manufacturing of bearings using ANN and SVM, in The 9th International Conference on Applied Informatics, vol. 2., 2014, p. 1.

[DK11] K. Deák, I. Kocsis, Z. Keviczki, Machine fault diagnosis, prognosis and condition monitoring based on machine learning, in IMT Oradea, Proceedings of the annual session of scientific papers: volume XIII (XXIII), 2014, p. 29–34.

[DK15] K. Deák, I. Kocsis, A. Vámosi, Gépek és gépelemek diagnosztikája, élettartam becslése gépi tanulás segítségével, in Informatika a felsőoktatásban, 2014, p. 470–476.

[DK16] I. Kocsis, T. Mankovits, A. Vámosi, and K. Deák, SVM Variants Used in the Investigation of some Engineering Optimization Problems, in Book of Abstracts of the 4th International Conference on Engineering Optimization, 2014, pp. 213–214.

11 Conclusions and Future Works

11.1 Summary and conclusion

In the research an object-oriented measurement system and method was realized which could detect rectangular shape manufacturing grinding faults of roller bearings and estimate the fault size. The system and the method were built on vibration analysis using vibration transducer, data acquisition unit, specialized software scripts for feature extraction and fault size classification. It was revealed that the roller bearing grinding fault detection and fault size estimation related to manufacturing were not investigated in the related researches.

For the experiments a test-rig was designed and constructed as the part of the object-orinted measurement sytem considering its structure and design. The test-rig was designed with vibration isolation basement to minimalize the vibration level and the modal frequencies were measured by experimental way to avoid resonances.

The mechanical signal model of the roller bearings with line-contact and rectangular shape grinding problems was defined which described the transient.

The generated transient wave was analysed and the main features of the transient were calculated. The transient was compared to other transient signals which were derived from different size of bearings. It was stated that similar signals were produced, namely transients which were different only in their parameters.

A fault size estimation procedure is developed for the special grinding faults. The entry and the exit points related to the motion of the roller at the place of the fault are analysed considering the mechanical model of the bearing. The geometrical size of the fault is estimated with wavelet multiresolution analysis. To find an efficient wavelet for this procedure the Symlet, Daubechies, Meyer and Morlet wavelets are compared with the Energy-to-Shannon Entropy ratio criteria. According to the study the Morlet wavelet provided the best result in the case of the grinding fault. The entry and the exit points are determined accurately to the fault size estimation in the certain geometrical size range.

A discrete wavelet was designed which proved to be more efficient in detection of the specific manufacturing faults than the user-ready wavelets. Maximum Energy-to-Shannon Entropy ratio criteria is used to compare the efficiency of the wavelets in the experiment. The designed wavelet is compared with Symlet and Daubechies wavelets which are generally used for bearing fault diagnosis. The geometrical size of the defect is estimated accurately by the designed wavelet and the results are validated by contact and optical

measurements. The result is checked by the Chapa-Rao algorithm which provided practically the same result as the wavelet which is designed in Labview.

Complex Morlet wavelet was designed with global parameter optimization. It was revealed as an effective filtering procedure because of its flexibility to change its center-frequency and bandwidth in order to make better match to the transient. Genetic algorithm was used for the parameter optimization. The designed wavelet was compared with Symlet, Daubechies, Morlet, Meyer wavelets by the Energy-to-Shannon Entropy ratio criteria to prove its efficiency in the grinding fault detection. After filtering the information content of the signal is further increased by autocorrelation enhancement. The results provided that the fault frequencies could be more efficiently determined with this procedure and, therefore, the calculation of the geometrical size of the defect was more accurate. Results of the fault size estimation was validated with optical and contact measurement devices.

Machine learning procedure using SVM was introduced to classify five different bearing faults. Besides the widely investigated ring and roller defects, back-surface faults were classified which could cause difficulties during quality inspection in the manufacturing process. Teaching data was created by huge quantity of bearings in order to get representative data set. The rotational speed of the test-rig and the axial spanning force of the bearing were adjusted to realize the sufficient data for the teaching procedure. At first, a fault classification method was presented using Support Vector Machine method. Parameters were chosen to reach the best classifying efficiency in the case of the bearing faults studied. Feature vectors derived from the wavelet decomposition by multiresolution analysis. The method enhanced the statistical parameters of the signal to achieve better bearing fault classification. Outer ring faults, inner ring faults, roller faults, back-surface defects and multi-type defects were classified with this method.

A multilayer neural network (ANN) was planned for bearing fault classification. The main purpose was to find a machine learning method that provides the best pattern recognition capability with the least CPU time for the classification. Three types of algorithm were compared; gradient-descent, Levenberg-Marquardt, scaled-conjugate back propagation. The parameters of the network were chosen to achieve the goal efficiently. Learning rate was optimized to get ideal classification.

Machine learning techniques produced sufficient level of multi-class fault classification even in the case of the small-sized back-fault defect.

11.2 Future works

As for the future work I plan to develop the measurement methods to analyse other kind of manufacturing faults with IT tools.

Electrical marks are many times a problem of the bearings of electrical motors which has similar rectangular shape under certain circumstances. If the fault has a well-defined entry and exit points, the fault could be detected similar way that the grinding defects. My purpose is to analyse how to detect the suffered bearings of the electric motors by using the vibration data of the system.

Further purpose is to test my method for detecting similar shape of problems on other machine elements, e.g. crane raceways or linear motion devices.

The optimization of the wavelet parameters is very subtle method, therefore my plan is to investigate other optimization methods besides the genetic algorithms such as differential algorithms, simulated annealing and compare their results.

I plan to analyse the efficiency other training algorithms in case of deep learning which can be used in neural networks for fault classification, such as BFGS Quasi-Newton, Resilient Backpropagation, Fletcher-Powell Conjugate Gradient, Polak-Ribiére Conjugate Gradient.

Further idea is to combine the SVM and ANN and analyse the fault classification efficiency of this hybrid system.

In some case measurements can not be executed in low noise level conditions, therefore my idea to test Adaptive Noise Cancellation Using RLS Adaptive Filtering to extract useful information from a noisy signal.

I plan to analyse the acoustic signals of the bearings picked by sensitive microphones and compare the results of fault detection and fault size estimation to the vibration measurement methods.

Realizing an expert system which gathers an overall data from the faulty items is an ultimate purpose of my research which can be implemented in the engineering serial production and quality inspection processes.

Publications of Krisztián Deák related to the Dissertation

[DK1] K. Deák, Design of Discrete Wavelet by Using Transient Model for Exact Measurement of Manufacturing Faults of Tapered Roller Bearings, Periodica Polytechnica- Mechanical Engineering, vol. 63, no. 2, pp. 113–122, 2019. Ranking: Q3

[DK2] K. Deák and I. Kocsis, Complex Morlet wavelet design with global parameter optimization for diagnosis of industrial manufacturing faults of tapered roller bearing in noisycondition, Diagnostyka, vol. 20, no. 2, pp. 77–86, 2019. Ranking: Q2

[DK3] K. Deák, T. Mankovits, I Kocsis, Optimal wavelet selection for manufacturing defect size estimation of tapered roller bearings with vibration measurement using Shannon Entropy Criteria, Strojniski Vestnik-Journal of Mechanical Engineering, vol. 63, no. 1, pp. 3–14, 2017. Ranking: Q2, IF: 1.182

[DK4] K. Deák and I. Kocsis, Support Vector Machine with Wavelet Decomposition Method for Fault Diagnosis of Tapered Roller Bearings by Modelling Manufacturing Defects, *Periodica Polytechnica-Mechanical Engineering*, vol. 61, no. 4, pp. 276–281, 2017. Ranking: Q3

Publications in international scientific journals

[DK5] K. Deák, J. Menyhárt, and L. Czégé, Defect analysis of bearings with vibration monitoring and optical methods, *International Journal of Engineering and Management Sciences / Műszaki és Menedzsment Tudományi Közlemények*, vol. 3, no. 4, pp. 1–12, 2018.

[DK6] K. Deák, I. Kocsis, A. Vámosi, Z. Keviczki, Failure diagnostics with SVM in machine maintenance engineering, *Analele universitatii din oradea fasciola management si inginerie tehnologica / Annals of the university of oradea fascicle of management and technological engineering*, vol. XXIII (XIII), no. 1, pp. 19–24, 2014.

Publications in conference proceedings

[DK7] K. Deák, A. Vámosi, and I. Kocsis, Csapágy meghibásodások méréstechnikája és rezgésdiagnosztikája mesterséges neurális hálók segítségével = Measurement techniques and diagnostics of bearings based on Artificial Neural Networks (ANN), in *Műszaki Tudomány az Észak-kelet Magyarországi Régióban 2014*, 2014, pp. 58–66.

[DK8] K. Deák, I. Kocsis, Gépek károsodása által előidézett kifáradásos csapágyhibák által generált tranziens impulzusok jelfeldolgozása

ablakozott fourier és wavelet transzformációk segítségével = bearing fault detection and signal processing with stft and wavelet transform in maintenance engineering, in *Műszaki Tudomány az Észak-kelet Magyarországi Régióban 2014*, 2014, pp. 67–72.

[DK9] K. Deák, A. Vámosi, A. Vámosi, I. Kocsis: Application of machine vision in manufacturing of bearings using ANN and SVM, in: *9th International Conference on Applied Informatics*, ICAI 2014.

[DK10] K. Deák, I. Kocsis, Rotary machine and roller bearing fault measurement with FPGA embedded system, in *Tavaszi Szél 2014 / Spring Wind 2014: II. kötet Állam- és jogtudomány, közigazgatás tudomány, had- és rendészettudomány*, 2014.

[DK11] K. Deák, I. Kocsis, Z. Keviczki, Machine fault diagnosis, prognosis and condition monitoring based on machine learning, in *IMT Oradea - 2014 : Proceedings of the annual session of scientific papers : volume XIII (XXIII) : may 29 - may 31, 2014, Oradea, Romania*, 2014, pp. 29–34.

[DK12] K. Deák, I. Kocsis, A. Vámosi, Z. Keviczki, Failure diagnostics with SVM in machine maintenance engineering, in *IMT Oradea - 2014 : Proceedings of the annual session of scientific papers : volume XIII (XXIII) : may 29 - may 31, 2014, Oradea, Romania*, 2014, pp. 23–28.

[DK13] Z. Keviczki, K. Deák, and I. Kocsis, Machine condition monitoring in bearing manufacture using vibration analysis and intelligent approaches, in *Proceedings of the International Scientific Conference on Advances in Mechanical Engineering*, 2014, pp. 55–69.

[DK14] K. Deák and I. Kocsis, Manufacturing of tapered roller bearings, defects and fault detection, in *Proceedings of the 3rd International Scientific Conference on Advances in Mechanical Engineering*, 2015, pp. 44–53.

[DK15] K. Deák, I. Kocsis, and A. Vámosi, Gépek és gépelemek diagnosztikája, élettartam becslése gépi tanulás segítségével, in *Informatika a felsőoktatásban 2014*, 2014, pp. 470–476.

[DK16] I. Kocsis, T. Mankovits, A. Vámosi, and K. Deák, SVM Variants Used in the Investigation of some Engineering Optimization Problems, in *Book of Abstracts of the 4th International Conference on Engineering Optimization*, 2014, pp. 213–214.

[DK17] K. Deák and I. Kocsis, Diagnosztikai és karbantartási vizsgálat csapágyhiba esetében wavelet transzformáció alkalmazásával, in *Műszaki tudomány az Észak-Kelet Magyarországi régióban 2016*, 2016, pp. 720–727.

[DK18] K. Deák and I. Kocsis, Applied vibration measurement methods and data extraction for bearing fault diagnosis, in *Proceedings of the 4th International Scientific Conference on Advances in Mechanical Engineering*, 2016, pp. 122–126.

[DK19] K. Deák and I. Kocsis, Selection of wavelet function for detection of bearing defects by Shannon entropy, in *Proceedings of the 4th International Scientific Conference on Advances in Mechanical Engineering*, 2016, pp. 127–134.

[DK20] K. Deák and I. Kocsis, Condition Monitoring by Applying Advanced Diagnostic Method, in *Proceedings of the 5th International Scientific Conference on Advances in Mechanical Engineering*, 2017, pp. 102–108.

[DK21] K. Deák and I. Kocsis, Wavelet design for diagnosis of tapered roller bearing manufacturing defects, in *Proceedings of the 6th International Scientific Conference on Advances in Mechanical Engineering*, 2018, pp. 39–40.

[DK22] K. Deák and I. Kocsis, Applied diagnostic methods in engine and bearing health monitoring. In: 7th International Scientific Conference on Advances in Mechanical Engineering, 2019.

References

[1] Patel, V. N., Tandon, N., Pandey, R., K., Defect Detection in Deep Groove Ball Bearing in Presence of External Vibration Using Envelope Analysis and Duffing Oscillator, Measurement 45, 2012, p. 960.

[2] Khanam, S., Tandon, N., Dutt, J. K., Fault Identification of Rolling Element Bearings from Vibration Signals: An Application of Kalman and H LOWHUV10th International Conference on Vibrations in Rotating Machinery (VIRM10), IMechE London, UK, Woodhead Publishing Limited, Institution of Mechanical Engineers (IMechE), 2012, pp. 703-713.

[3] Al-Ghamd, A.M., Mba, D., A Comparative Experimental Study on the Use of Acoustic Emission and Vibration Analysis for Bearing Defect Identification and Estimation of Defect Size, Mechanical Systems and Signal Processing 20, p. 1537.

[4] Elforjani, M., Mba, D., Accelerated Natural Fault Diagnosis in Slow Speed Bearings with Acoustic Emission, Engineering Fracture Mechanics 77, 2010, p. 112.

[5] Sawalhi, N., Randall, R.B., Vibration Response of Spalled Rolling Element Bearings: Observations, Simulations and Signal Processing Techniques to Track the Spall Size, Mechanical Systems and Signal Processing 25, 2011, p. 846.

[6] Peng, Z.K., Chu, F.L., Application of the Wavelet Transform in Machine Condition Monitoring and Fault Diagnostics: A Review with Bibliography, Mechanical Systems and Signal Processing 18, 2004, p. 199.

[7] Prabhakar, S., Mohanty, A.R., Sekhar, A.S., Application of Discrete Wavelet Transform for Detection of Ball Bearin g Race Faults, Tribology International 35, 2002, p.793.

[8] Shi, D. F., Wang, W. J., Qu, L. S., Defect detection for bearings using envelope spectra of wavelet transform, ASME Journal of Vibration and Acoustics 126 (4), 2004, p. 567.

[9] Nikolaou, N. G., Antoniadis, I. A., Demodulation of Vibration Signals Generated by Defects in Rolling Element Bearings Using Complex Shifted Morlet Wavelets, Mechanical Systems and Signal Processing 16(4), 2002, p. 677.

[10] Qiu, H., Lee, J., Lin, J., Yu, G., Wavelet Filter-Based Weak Signature Detection Method and its Application on Rolling Element Bearing Prognostics, Journal of Sound and Vibration 289, 2006, p. 1066.

[11] Junsheng, C., Dejie, Y., Yu, Y., Application of an Impulse Response Wavelet to Fault Diagnosis of Rolling Bearings, Mechanical Systems and Signal Processing 21, 2007, p. 920.

[12] Kumar, R., Singh, M., Outer Race Defect Width Measurement in Taper Roller Bearing using Discrete Wavelet Transform of Vibration Signal, Measurement 46, 2013, p. 537.

[13] Awal, M. A., Mostafa, S. S., Ahmad, M., Quality Assessement of ECG Signal Using Symlet Wavelet Transform, Proceedings of International Conference on Advances in Electrical Engineering, 2012, p.129-134.

[14] R. Kumar, D.P. Jena, M. Bains, Identification of inner race defect in radial ball bearing using acoustic emission and wavelet analysis, in: Proceedings of ISMA 2010 including USD 2010 Leuven (Belgium), 2010, pp. 2883–2891.

[15] R. Yan, R.X. Gao, Multi-scale enveloping spectrogram for vibration analysis in bearing defect diagnosis, Tribology International 42, 2009, 293–302.

[16] M.S. Patil, J. Mathew, P.K. Rajendrakumar, S. Desai, A theoretical model to predict the effect of localized defect on vibrations associated with ball bearing, International Journal of Mechanical Sciences 52 (9), 2010, 1193–1201.

[17] W. He, Z. Jiang, K. Feng, Bearing fault detection based on optimal wavelet filter and sparse code shrinkage, Measurement 42, 2009, 1092–1102.

[18] N. Sawalhi, R.B. Randall, Vibration response of spalled rolling element bearings, observations, simulations and signal processing techniques to track the spall size, Mechanical Systems and Signal Processing 25, 2011, 846–870.

[19] S. Khanam, N. Tandon, J.K. Dutt, Fault size estimation in the outer race of ball bearing using discrete wavelet transform of the vibration signal. 2nd International Conference on Innovations in Automation and Mechatronics Engineering, ICIAME 2014

[20] L. Tóth, T. Tóth, Construction of a realistic signal model of transients for a ball bearing with inner race fault. Acta Polytechnica Hungarica, Vol 10., No 1., 2013.

[21] L. Tóth, T. Tóth: On finding wavelet basis for bearing fault detection, Vol 10., No 3., 2013.

[22] Borghesani, P., Pennacchi, P., Randall, R.B., Sawalhi, N., Ricci, R., Application of cepstrum pre-whitening for the diagnosis of bearing faults under variable speed conditions. Mechanical Systems and Signal Processing, vol. 36, no. 2, 2013, p. 370-384,

[23] Figlus, T., Stanczyk, M., A method for detecting damage to rolling bearings in toothed gears of processing lines. Metalurgija, vol. 55 no. 1, 2016, p. 75-78.

[24] Slavič, J., Brković, A., Boltežar, M, Typical bearing-fault rating using force measurements: application to real data. Journal of Vibration and Control, vol. 17, no 14, 2011, p. 2164-2174.

[25] Abboud, D., Antoni, J., Eltabach, M., Sieg-Zieba, S., Angle\time cyclostationarity for the analysis of rolling element bearing vibrations. Measurement, vol. 75, 2015, p. 29-39.

[26] Antoni, J., Cyclic spectral analysis of rolling-element bearing signals: Facts and fictions. Journal of Sound and Vibration, vol. 304, no. 3-5, 2007, p. 497-529,

[27] A. H. Tewfik, D. Sinha, P. Jorgensen, On the Optimal Choice of a Wavelet for Signal Representation, IEEE Transactions on Information Theory, Vol. 38, 1992, p. 747-765.

[28] J. O. Chapa, R. M. Rao. Algorithms for Designing Wavelets to Match a Specified Signal, IEEE Transactions on Signal Processing, Vol. 48, No. 12, 2007, p. 3395-3406, 2000.

[29] S. Mallat. A Wavelet Tour of Signal Processing, Second Edition, Academic Press, 1998.

[30] N. Saravanan, V.N.S. Kumar Siddabattuni, K.I. Ramachandran. A comparative study on classification of features by SVM and PSVM extracted using Morlet Wavelet for fault diagnosis. Expert Systems with Applications, 35, 2008, p. 1351–1366.

[31] J. Rafiee, M.A. Rafiee, P.W. Tse. Application of mother wavelet functions for automatic gear and bearing fault diagnosis. Expert Systems with Applications, 37, 2010, p. 4568–4579.

[32] P.K. Kankar, Satish C. Sharma, S.P. Harsha: Fault diagnosis of rolling element bearing using cyclic autocorrelation and wavelet transform. Neurocomputing 110, 2013, p. 9-17.

[33] P.K. Kankar, Satish C. Sharma, S.P. Harsha. Fault diagnosis of ball bearings using machine learning methods. Expert System with Applications, 38, 2011, p. 1876-1886.

[34] P.K. Kankar, Satish C. Sharma, S.P. Harsha. Fault diagnosis of ball bearings using continuous wavelet transform. Applied Soft Computing, 11, 2011, p. 2300-2312.

[35] P.K. Kankar, Satish C. Sharma, S.P. Harsha. Rolling element fault diagnosis using wavelet transform. Neurocomputing, 74, 2011, p. 1638-1645.

[36] V. Purushotham, S. Narayanana, A.N. Prasad Suryamarayana. Multi-fault diagnosis of rolling elements using wavelet analysis and Hidden Markov Model based fault recognition. NDT and E International, 38, 2005, p. 654-664.

[37] Honghu, P., Xingxi, H., Sai, T., Fanming, M., An Improved Bearing Fault Diagnosis Method using One-Dimensional CNN and LSTM. Strojniski vestnik-Journal of Mechanical Engineering, 2018, 64, p. 443-452.

[38] Li, Z., Ma, Z., Liu, Y., Teng, W., Jiang, R. Crack, Fault Detection for a Gearbox Using Discrete Wavelet Transform and an Adaptive Resonance Theory Neural Network. Strojniski vestnik-Journal of Mechanical Engineering, 61, 2015, p. 63-73.

[39] Strączkiewicz, M., Czop, P., Barszcz, T., Supervised and unsupervised learning process in damage classification of rolling element bearings. Diagnostyka, vol. 17, no. 2, 2016, p. 71-80.

[40] Gligorijevic, J., Gajic, D., Brkovic, A., Savic-Gajic, I., Georgieva, O., Di Gennaro, S., Online condition monitoring of bearings to support total productive maintenance in the packaging materials industry. Sensors, vol. 16 no. 3, 2016, p. 316.

[41] Strączkiewicz, M., Czop, P., Barszcz, T., Supervised and unsupervised learning process in damage classification of rolling element bearings. Diagnostyka, 2016, 17, p. 71-80.

[42] Paya, B. A., Esat, I. I., Artificial neural network based fault diagnostics of rotating machinery using wavelet transforms as a preprocessor. Mechanical Systems and Signal processing, Vol. 11, No. 5, 1997, p.751-765.

[43] Wang, P., Vachtsevanos, G., Fault diagnostics using dynamic wavelet neural networks. Artificial Intelligent for Engineering Design Analysis and Manufacturing, Vol.15, 2001, p. 349-365.

[44] Ray, S., Chan, A., Automatic feature extraction from wavelet coefficients using genetic algorithms. Proceedings of the 2001 IEEE Signal Processing Society Workshop, 2001, p. 233-24.

[45] Chen, D., Wang, W. J., Classification of wavelet map patterns using multi-layer neural networks for gear fault detection. Mechanical System and Signal Processing, Vol.16, No. 4, 2002, p. 695-704.

[46] N.G. Nikolaou, A.I. Antoniadis, Demodulation of vibration signals generated by defects in rolling element bearings using complex shifted Morlet wavelets, Mechanical Systems and Signal Processing, Vol. 16, 2002, p. 677-694.

[47] K.C. Gryllias, I.A. Antoniadis, A Support Vector Machine approach based on pyhisical model training for rolling element bearing fault detection in industrial environments. Engineering Applications of Artificial Intelligence 25, 2012, p. 326-344.

[48] Wensheng Su, Fengtao Wang, Hong Zhu, Zhixin Zhang, Zhenggang Guo, Rolling element bearing faults diagnosis on optimal Morlet wavelet filter

and autocorrelation enhancement. Mechanical Systems and Signal Processing 24., 2010, p. 1458-1472.

[49] Manpreet Singh, Rajesh Kumar: Thrust bearing groove race defect measurement by wavelet decomposition of pre-processed vibration signal. Measurement 46., 2013 p. 3508-3515.

[50] Bennett, K. P. & Campbell, C., Support Vector Machines: Hype or Hallelujah? SIGKDD 2., 2000.

[51] Luenberger, D. G. & YE, Y., Linear and non-linear programming, New York, 2008, Springer

[52] Nguyen, M. H. & DE LA Torre, F., Optimal feature selection for support vector machines. Pattern Recognition, 43, 2010, p. 584-591.

[53] Burges C., A tutorial on support vector machines for pattern recognition, In Data Mining and Knowledge Discovery. Kluwer Academic Publishers, Boston, 1998, Volume 2.

[54] Theodoros Evgenuiu and Massimilliano Pontil, Statistical Learning Theory, 1998.

[55] Nello Cristianini and John Shawe-Taylor, An Introduction to Support Vector Machines and Other Kernel-based Learning Methods, Cambridge University Press, 2000.

[56] J.P. Lewis, Tutorial on SVM, CGIT Lab, USC, 2004.

[57] Vapnik V., Statistical Learning Theory, Wiley, New York, 1998.

[58] R.A. Wiggins, Minimum entropy Deconvolution, Geoexploration, Vol. 16, 1978, p. 21–35.

[59] Addison, P.S., The Illustrated Wavelet Transform Handbook, Institute of Physics Publishing, Bristol and Philadelphia

[60] Georges Oppenheim, Wavelets and their applications. Hermes Science/ Lavoisier, 2003.

[61] Shreve, D.H., Signal processing for effective vibration analysis, IRD Mechanalysis, Inc, Colombus, Ohio

[62] Micheil-Yves-Georges-Jean, Wavelets and their applications, 2007. ISTA Ltd.

[63] Allan-Thomas: Harris's Shock and Vibration Handbook. McGrawHill, 2010.

[64] Tom M. Mitchell: Machine learning. McGraw-Hill Science. 1997.

[65] I. Kocsis, K. Deák, T. Szabó, R. Kvasz: Diagnosztika és állapotfelügyelet. 2015.

[66] I. Nagy: Műszaki diagnosztika I. Rezgésdiagnosztika. Delta-3N Kft., Paks, 2006

[67] I. Fazekas, Neurális hálózatok. Debreceni Egyetem Informatikai Kar. 2013.

[68] Gy. Kovács: A jelfeldolgozás matematikai alapjai. Debreceni Egyetem. 2014.

[69] F. Dömötör: A rezgésdiagnosztika elemei, SKF, Budapest, 1996

[70] Gy. Lipovszki, Jelfeldolgozás és számítógépes irányítás. Edutus Főiskola 2012.

[71] S. Graham Kelly: Mechanical Vibrations. Theory and Applications. The University of Akron. 2012

[72] Alena Bilosova, Jan Bilos: Vibration Diagnostics. Ostrava, 2012.

[73] Jaroslav Ramík: Soft Computing: Overview and Recent Developments in Fuzzy Optimization. Ostravská univerzita. Listopad 2001

[74] SPM INSTRUMENT BUDAPEST KFT: Az SPM Módszer, 2000. http://www.spminstrument.hu/index.php?lid=0&gcf=24&sidx=1

[75] http://visl.technion.ac.il/documents/wavelet_ug.pdf.tomma.ewing Downloaded: 9th March, 2019.

[76] NI 9234 datasheet: http://www.ni.com/datasheet/pdf/en/ds-316. Downloaded: 18th February, 2019.

[77] PCB IMI 603C01 transducer: https://www.pcb.com/contentstore/docs /PCB_Corporate/IMI/Products/Manuals/603C01.pdf, Downloaded: 20th April, 2019.

[78] Schneider frekvenciaváltó: http://datasheet.octopart.com/ATV32HU22M2-Schneider-Electric-datasheet-14430726.pdf Downloaded: 8th April, 2019.

[79] R Software page: https://cran.r-project.org/. Downloaded: 12th April, 2019.

[80] High Speed spindle for grinding: https://cnc-machine-tools.com/cnc-router-spindle/

[81] F. Schipp, Waveletek, egyetemi jegyzet, ELTE 2003.

[82] E.Abele, Y. Altintas, C.Brecher: Machine tool spindle units. CIRP Annals Manufacturing Technology, Vol. 59, Issue 2, 2010, p. 781-802.

[83] Danobat Machine Tools: https://www.danobatusa.com/bearings-rings-internal-cylindrical-grinding, Downloaded: 20th September, 2019.

[84] SKF Product Catalogue: www.skf.com, Downloaded: 12 th September, 2019.

[85] FAG Rolling Bearing Damage, Recognition of damage and bearing inspection: http://rolling.hu/pdf/FAG/FAG_meghibasodas.pdf, Downloaded: 2 th June, 2019.

Standards

[S.1] ISO 2372:1974 Mechanical vibration of machines with operating speeds from 10 to 200 rev/s -- Basis for specifying evaluation standards

[S.2] ISO 3945:1985 Mechanical vibration of large rotating machines with speed range from 10 to 200 r/s -- Measurement and evaluation of vibration severity in situ

[S.3] ISO 10816-1:1195-(E) Mechanikai rezgések – Géprezgések kiértékelése a nem forgó részeken történő méréssel. – Nemzetközi szabvány

[S.4] American National Standard ANSI/AFBMA Std 13-1970, ANSI B3.13-1970, Rolling Bearing Vibration and Noise (Methods of Measuring)

[S.5] Deutsches Institut für Normung DIN 5426, Laufgeräusche von Wälzlagern, Prüfverfahren.

[S.6] SKF főkatalógus, katalógusszám: 8200, Svéd golyóscsapágy Rt., 810937, Kossuth nyomda Bp.

[S.7] FAG Gördülőcsapágyak Főkatalógus, FAG OEM und Handel AG, 1996.

Appendix

Appendix 1.: SVM script for fault classification

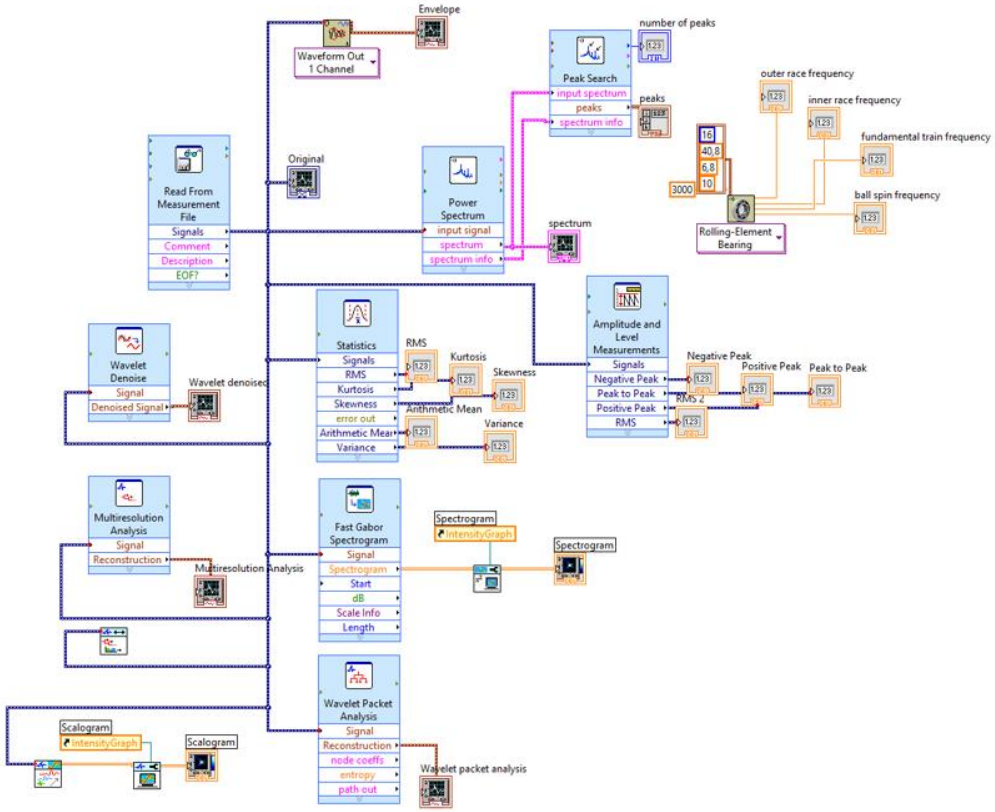
```
#SVM loading

library(e1071)

#teaching and test data

ta <- read.csv("G:\\2015márcPhD\\tanulo.csv", header=FALSE, sep=";")
tanulox=matrix(data=0,nr=118,nc=5,byrow=TRUE)
tanulox[,1]<-data.matrix(ta$V1, rownames.force = NA)
tanulox[,2]<-data.matrix(ta$V2, rownames.force = NA)
tanulox[,3]<-data.matrix(ta$V3, rownames.force = NA)
tanulox[,4]<-data.matrix(ta$V4, rownames.force = NA)
tanulox[,5]<-data.matrix(ta$V5, rownames.force = NA)
tanuloy<-data.matrix(ta$V6, rownames.force = NA)

te <- read.csv("G:\\2015márcPhD\\teszt.csv", header=FALSE, sep=";")

tesztx=matrix(data=0,nr=28,nc=5,byrow=TRUE)
tesztx[,1]<-data.matrix(te$V1, rownames.force = NA)
tesztx[,2]<-data.matrix(te$V2, rownames.force = NA)
tesztx[,3]<-data.matrix(te$V3, rownames.force = NA)
tesztx[,4]<-data.matrix(te$V4, rownames.force = NA)
tesztx[,5]<-data.matrix(te$V5, rownames.force = NA)
teszty<-data.matrix(te$V6, rownames.force = NA)

# searching the optimal parameters (RBF kernel)

n=110
cv=numeric(length=n*4)
dim(cv)=c(n,4)
cvn=1

e=0.01
for (c in 1:10) {
for (g in 0:10) {

svr <- svm(tanulox,tanuloy,type="C-classification", kernel="radial",
cost=c*6, epsilon=e, gamma=0.01*g)
preyf <- predict(svr,tesztx)
prey=as.numeric(levels(preyf))[preyf]
dim(prey)=c(28,1)
```

```
cv[cvn,1]=e
cv[cvn,2]=c
cv[cvn,3]=0.10*g
cv[cvn,4]=0
for (i in 1:28) {
      if (teszty[i]==prey[i]) {cv[cvn,4]=cv[cvn,4]+1}
}
cvn=cvn+1
}
}
cv

#running svm

svr <- svm(tanulox,tanuloy,type="C-classification", kernel="radial",
cost=10, epsilon=0.01, gamma=0.5)
preyf <- predict(svr,tesztx)
prey=as.numeric(levels(preyf))[preyf]
dim(prey)=c(28,1)

#searching optimal parameters (poly kernel)

n=60
cv=numeric(length=n*4)
dim(cv)=c(n,4)
cvn=1

e=0.01
for (c in 1:10) {
for (d in 0:5) {

svr <- svm(tanulox,tanuloy,type="C-classification", kernel="poly",
cost=c, epsilon=e, degree=d)
preyf <- predict(svr,tesztx)
prey=as.numeric(levels(preyf))[preyf]
dim(prey)=c(28,1)

cv[cvn,1]=e
cv[cvn,2]=c
cv[cvn,3]=d
cv[cvn,4]=0
for (i in 1:28) {
      if (teszty[i]==prey[i]) {cv[cvn,4]=cv[cvn,4]+1}
cvn=cvn+1

# running svm

svc <- svm(tanulox,tanuloy,type="C-classification", kernel="poly",
cost=10, epsilon=0.01, degree=3)
```

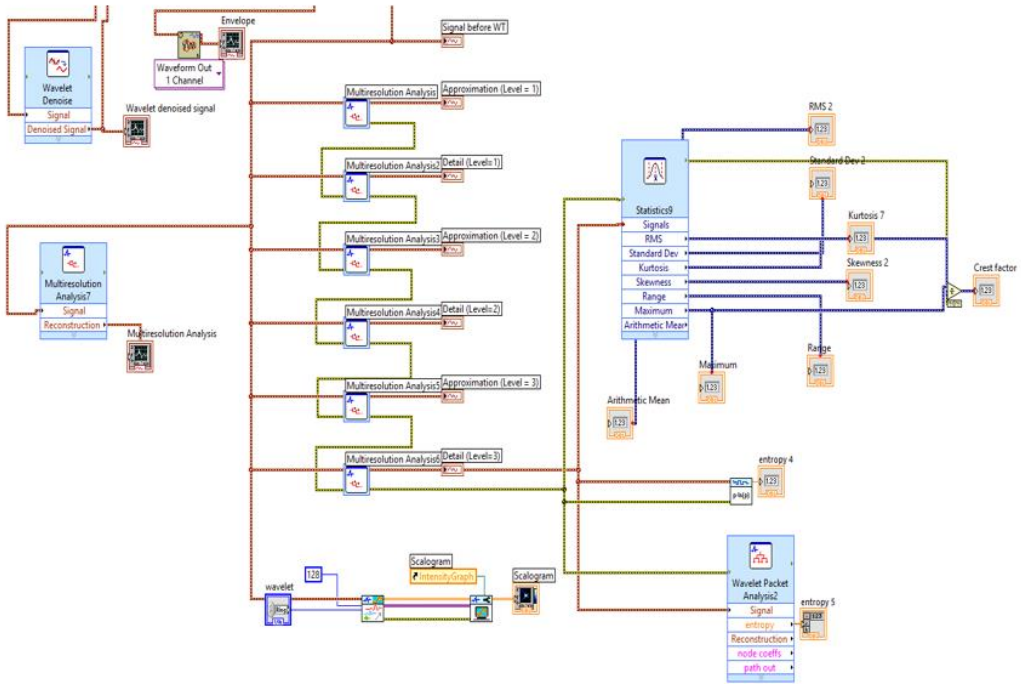
Appendix 2.: Feature vectors axial loading of healthy bearing data set for SVM and ANN classification of 450 N

	Healthy bearing data 3th level wavelet decomposition									
1/min	Hz	Range	Mean value	Std. Dev.	Kurtosis	Skewness	Crest factor	Parseval	Entropy	P/E
60	1	0,018847	-0,00082	0,002264	4,081449	-0,09829	2,981569	6719,5	6,511915	1031,878
120	2	0,018749	-0,000701	0,002772	2,939331	-0,04747	2,875317	4908,128	8,984309	546,3
180	3	0,027093	-0,000716	0,003747	3,150093	-0,0453	3,564365	5131,637	15,09428	339,9723
240	4	0,042218	-0,000866	0,004836	3,493317	-0,15007	3,424472	7502,245	23,71926	316,2934
300	5	0,040454	-0,000813	0,005593	3,310499	-0,18189	3,204644	6609,207	30,43005	217,1934
360	6	0,063872	-0,000827	0,007311	4,272462	0,140571	4,493483	6841,948	48,20512	141,9341
420	7	0,073486	-0,000827	0,009487	3,532236	-0,00693	3,700488	6846,116	76,65694	89,30849
480	8	0,072323	-0,000854	0,008842	3,80974	-0,06155	3,895327	7296,888	67,73188	107,732
540	9	0,067252	-0,000916	0,009812	2,750087	-0,03362	3,151972	8385,763	83,02191	101,0066
600	10	0,076848	-0,000828	0,011545	3,011838	-0,04588	3,329296	6860,713	109,5121	62,64797
660	11	0,108247	-0,000905	0,012636	4,031234	-0,13713	3,359039	8193,394	126,1569	64,94605
720	12	0,089078	-0,000852	0,013013	3,075932	-0,09551	3,30437	7262,358	134,7359	53,90068
780	13	0,101192	-0,000915	0,014952	3,197802	0,000125	3,28679	8370,379	170,7185	49,03029
840	14	0,110178	-0,000871	0,01657	3,061923	-0,05365	3,071873	7592,61	205,0667	37,02507
900	15	0,12057	-0,000937	0,018728	2,799756	-0,17792	2,715597	8780,352	254,6538	34,47957
960	16	0,117035	-0,000868	0,017783	2,709391	-0,14138	3,417192	7533,905	233,7886	32,22528
1020	17	0,143648	-0,000796	0,021634	3,065857	-0,06395	3,17631	6332,704	323,1429	19,59723
1080	18	0,134741	-0,000703	0,01989	3,06941	-0,20346	2,983795	4946,353	281,1854	17,59107
1140	19	0,135903	-0,000787	0,020694	3,04322	-0,19987	3,061897	6187,222	300,1242	20,61554
1200	20	0,140612	-0,000827	0,021626	2,910767	0,059307	3,18775	6844,988	324,7494	21,07775
1260	21	0,137882	-0,000864	0,022926	2,742244	-0,08278	2,897853	7473,141	360,8653	20,70895
1320	22	0,199004	-0,000738	0,027383	3,325325	-0,18368	3,230011	5452,173	482,0188	11,31112
1380	23	0,19111	-0,000453	0,026597	3,081708	-0,1841	3,507912	2052,977	460,0098	4,462898
1440	24	0,210145	-0,000452	0,029225	3,332805	-0,06725	3,535085	2045,343	535,5319	3,819274
1500	25	0,212431	-0,000429	0,030968	3,278124	-0,02197	3,485028	1838,94	589,6311	3,118797
1560	26	0,214854	-0,000506	0,030932	3,785122	-0,05338	3,557816	2558,703	577,8027	4,428333
1620	27	0,263381	-0,000581	0,034852	3,840612	-0,00164	3,640912	3379,216	704,293	4,798025
1680	28	0,280401	-0,000524	0,04207	3,133561	0,204965	3,746952	2746,161	987,2474	2,781634
1740	29	0,268769	-0,000623	0,037258	3,573581	0,237893	4,064713	3885,339	795,4331	4,884558
1800	30	0,245388	-0,000426	0,038042	3,264307	0,153307	3,327976	1811,147	829,7793	2,182685
1860	31	0,244666	-0,000551	0,037747	3,204139	0,121229	3,578486	3039,354	821,3385	3,700489
1920	32	0,226446	-0,000745	0,036252	3,039657	-0,10752	3,117138	5554,346	771,5094	7,199324
1980	33	0,218833	-0,000556	0,035326	2,912621	-0,20885	2,74461	3086,794	744,1189	4,148254
2040	34	0,225838	-0,000525	0,035454	2,914394	0,016655	3,347183	2757,337	749,0188	3,681266
2100	35	0,262278	-0,000665	0,038743	3,137744	0,109817	3,607592	4427,263	861,668	5,138015
2160	36	0,305419	-0,000757	0,050979	2,713718	0,019061	3,169454	5733,026	1370,596	4,182871
2220	37	0,347619	-0,000562	0,054193	2,947048	0,121309	3,261088	3158,619	1496,175	2,11113
2280	38	0,399032	-0,000382	0,06346	2,837045	0,078718	2,931435	1458,583	1932,502	0,754764
2340	39	0,457685	-0,000728	0,072985	2,714054	-0,01189	3,014642	5305,474	2425,619	2,187266
2400	40	0,56376	-0,000694	0,092619	2,600805	0,077268	3,159299	4816,786	3522,263	1,367526
2460	41	0,605841	-0,000719	0,107031	2,793614	0,270008	3,257952	5166,828	4316,854	1,196897
2520	42	0,797444	-0,000709	0,120707	3,027249	0,38206	3,809479	5020,32	5086,45	0,986999
2580	43	0,815401	-0,000539	0,130765	2,873556	0,289278	3,555913	1709,962	5725,623	0,298651
2640	44	0,833066	-0,00059	0,126828	2,835058	0,07388	3,418097	1608,555	5512,676	0,291792
2700	45	0,805706	-0,000493	0,126846	2,773021	0,033412	3,327424	1608,996	5537,521	0,290562
2760	46	0,799241	-0,000886	0,130312	2,678681	-0,10423	3,107404	1698,189	5774,285	0,294095
2820	47	0,754221	-0,00046	0,140225	2,622329	-0,10646	2,552839	1966,302	6418,031	0,306371
2880	48	0,802923	-0,000299	0,145243	2,645185	-0,10946	2,681502	2109,54	6717,008	0,314059
	átlag	**0,278982**			**3,123826**	**-0,01246**	**3,30877794**	**4846,163**	**1436,043**	**70,95211**
	szórás	0,252445			0,405342	0,139734	0,36579156	2303,033	2027,729	175,8715

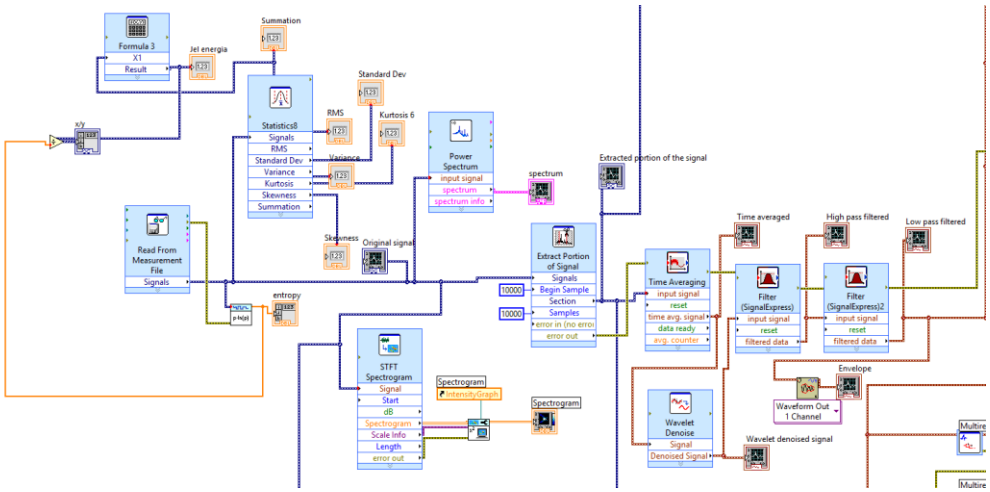
Appendix 3.: Details of the Labview VIs for the experiments



Labview VI for vibration analysis of the bearings



Labview VI for multiresolution analysis



Labview VI with filtering for signal processing

Appendix 4.: Script for Neural Network fault classification and parameter optimization

```
% Solve a Pattern Recognition Problem with a Neural Network
% Script generated by Neural Pattern Recognition app
% Created 30-Jun-2019 20:33:50
%
% This script assumes these variables are defined:
%
%  input - input data.
%  target - target data.

x = input;
t = target;

% Choose a Training Function
% For a list of all training functions type: help nntrain
% 'trainlm' is usually fastest.
% 'trainbr' takes longer but may be better for challenging problems.
% 'trainscg' uses less memory. Suitable in low memory situations.
trainFcn = 'trainlm'; % gradient descent

% Create a Pattern Recognition Network
hiddenLayerSize = 10;
net = patternnet(hiddenLayerSize, trainFcn );

% Choose Input and Output Pre/Post-Processing Functions
% For a list of all processing functions type: help nnprocess
net.input.processFcns = {'removeconstantrows','mapminmax'};
net.output.processFcns = {'removeconstantrows','mapminmax'};

% Setup Division of Data for Training, Validation, Testing
% For a list of all data division functions type: help nndivide
net.divideFcn = 'dividerand'; % Divide data randomly
net.divideMode = 'sample'; % Divide up every sample
net.divideParam.trainRatio = 70/100;
net.divideParam.valRatio = 15/100;
net.divideParam.testRatio = 15/100;

% Choose a Performance Function
% For a list of all performance functions type: help nnperformance
net.performFcn = 'crossentropy'; % Cross-Entropy

% Choose Plot Functions
% For a list of all plot functions type: help nnplot
net.plotFcns = {'plotperform','plottrainstate','ploterrhist', ...
  'plotconfusion', 'plotroc'};

% Train the Network
```

```
[net,tr] = train(net,x,t);

% Test the Network
y = net(x);
e = gsubtract(t,y);
performance = perform(net,t,y)
tind = vec2ind(t);
yind = vec2ind(y);
percentErrors = sum(tind ~= yind)/numel(tind);

% Recalculate Training, Validation and Test Performance
trainTargets = t .* tr.trainMask{1};
valTargets = t .* tr.valMask{1};
testTargets = t .* tr.testMask{1};
trainPerformance = perform(net,trainTargets,y)
valPerformance = perform(net,valTargets,y)
testPerformance = perform(net,testTargets,y)

% View the Network
view(net)

% Plots
% Uncomment these lines to enable various plots.
%figure, plotperform(tr)
%figure, plottrainstate(tr)
%figure, ploterrhist(e)
%figure, plotconfusion(t,y)
%figure, plotroc(t,y)

% Deployment
% Change the (false) values to (true) to enable the following code
blocks.
% See the help for each generation function for more information.
if (false)
  % Generate MATLAB function for neural network for application
  % deployment in MATLAB scripts or with MATLAB Compiler and Builder
  % tools, or simply to examine the calculations your trained neural
  % network performs.
  genFunction(net,'myNeuralNetworkFunction');
  y = myNeuralNetworkFunction(x);
end
if (false)
  % Generate a matrix-only MATLAB function for neural network code
  % generation with MATLAB Coder tools.
  genFunction(net,'myNeuralNetworkFunction','MatrixOnly','yes');
  y = myNeuralNetworkFunction(x);
end
if (false)
  % Generate a Simulink diagram for simulation or deployment with.
  % Simulink Coder tools.
```

```
  gensim(net);
end

[m,n] = size(bearingdata) ;
% Split into train and test
P = 0.7 ;
Training = bearingdata (1:round(P*m),:) ;
Testing = bearingdata (round(P*m)+1:end,:);
XTrain = Training(:,1:n-1);
YTrain = Training(:,n);
XTest = Testing(:,1:n-1);
YTest = Testing(:,n);
% Define a train/validation split to use inside the objective
function
cv = cvpartition(numel(YTrain), 'Holdout', 1/3);
% Define hyperparameters to optimize
vars = [optimizableVariable('hiddenLayerSize', [1,20], 'Type',
'integer');optimizableVariable('lr', [1e-3 1], 'Transform', 'log')];
% Optimize
minfn = @(T)kfoldLoss(XTrain', YTrain', cv, T.hiddenLayerSize, T.lr);
results = bayesopt(minfn, vars,'IsObjectiveDeterministic', false,...
    'AcquisitionFunctionName', 'expected-improvement-plus');
T = bestPoint(results)
% Train final model on full training set using the best
hyperparameters
net = feedforwardnet(T.hiddenLayerSize, 'traingd');
net.trainParam.lr = T.lr;
net = train(net, XTrain', YTrain');
% Evaluate on test set and compute final rmse
ypred = net(XTest');
finalrmse = sqrt(mean((ypred - YTest').^2))
function rmse = kfoldLoss(x, y, cv, numHid, lr)
% Train net.
net = feedforwardnet(numHid, 'traingd');
net.trainParam.lr = lr;
net = train(net, x(:,cv.training), y(:,cv.training));
% Evaluate on validation set and compute rmse
ypred = net(x(:, cv.test));
rmse = sqrt(mean((ypred - y(cv.test)).^2));
end
```