Available online at www.sciencedirect.com
ScienceDirect

Information Sciences 00 (2013) 1–29

Information Sciences

Improved process monitoring and supervision based on a reliable multi-stage feature-based pattern recognition technique

Lou'i Al-Shrouf, Mahmud-Sami Saadawia, Dirk Söffker

Chair of Dynamics and Control, University of Duisburg-Essen, Lotharstr. 1-21, 47057, Duisburg, Germany

Abstract

This paper investigates two new multisensor data fusion algorithms for object detection in monitoring of industrial processes. The goals were to reduce the rate of false detection and obtain reliable decisions on the presence of target objects. The monitoring system uses acceleration sensors and is used as a sensor-cluster. In principle the approach can include arbitrary data acquisition techniques. Two approaches were proposed. The first uses a short-time Fourier transform (STFT) as a prefilter to extract relevant features from the acceleration signals. The features extracted from different sensor channels are first classified using support vector machine (SVM)-based filters. A novel decision fusion process to combine individual decisions was developed. The second approach uses a continuous wavelet transform (CWT) as a prefilter to extract relevant features from the acceleration signals. The features extracted from different sensor signals are subjected to further prefiltering processes before SVM-based classification. The individual decision functions are then combined in a decision fusion module. The classification system was trained and validated using real industrial data. The two approaches were tested using the same data and their performance and modeling complexity are compared. The developed approaches show strong improvements in detection and false alarm rates.

Keywords: System state monitoring, Information fusion, Short-time Fourier transform (STFT), Continuous wavelet transform (CWT), Support vector machine (SVM)

1. Introduction

Monitoring and supervision systems are essential for the automation of industrial processes. Individual process-related variables are usually considered and thresholds are defined and used to distinguish regular and abnormal operations. In addition, abnormal conditions need to be identified for fault detection and diagnostic tasks. Such methods are not generally applicable for highly complex systems because the multidimensionality and interrelations

Email addresses: loui.alshrouf@uni-due.de (Lou'i Al-Shrouf), mahmud-sami.saadawia@uni-due.de (Mahmud-Sami Saadawia), soeffker@uni-due.de (Dirk Söffker)

involved cannot be handled by low-dimensional approaches that use classical thresholds. Model-based monitoring systems are often not suitable for complex systems because precise models of the mechanical system considered are required for reliable monitoring [6]. Model-based methods usually require complex modeling of the process with detailed process parameters and additional information on changes in the system states.

Signal-based diagnostic methods are based on an analysis of measured (physical) signals. They are useful when the measured variables contain direct or implicit information about possible faulty behavior. Signal-based diagnostic methods are easy to use and are widely adopted to extract relevant process characteristics from analyzed sensor data in combination with further knowledge. Feature extraction can be performed in either the time or frequency domain of the signal. The extracted features should be able to represent the regular state of the system, as well as non-regular behaviors; in other words, they should indicate changes in system states. Thus, signal-based methods when combined with machine learning techniques can be used to distinguish system states. Depending on the machine and process complexity, suitable sensors have to be used to define suitable mappings between machine operating states and sensor data.

In this work, a new monitoring approach is developed. The developed approach is inspired by the monitoring task of production processes that suffer from various drawbacks of approaches applied before [4, 35]. The goal is to detect the presence of target objects within the material transported during a production process. For such specific applications, the presence of target objects in transported material (overburden) has to be detected (target object present yes/no) to avoid resulting disturbances and failures during the continuous transportation process. The approach developed must be practical and suitable for real-time use for standard industrial hardware.

A few studies have considered the development of monitoring systems for similar production processes. Petrich and Köhler [38] developed a monitoring system based on georadar to detect target objects in overburden before excavation. A detection rate of less than 60% was achieved. Their system was not able to distinguish non-critical objects such as frozen overburden from the target objects of interest. Such objects lead to false alarms, disturbing the production process. Their system could also not identify the position of target objects accurately. Petrich and Köhler also installed a radiometric measuring system above a transport belt to detect target objects in the material flow. A detection rate of less than 70% was achieved. Changes in the petrography and elemental composition of the overburden, other objects such as clay chunks, and small objects led to false alarms.

Nieß developed an automatic monitoring system to detect target objects in a material flow using acceleration sensors [35]. Several acceleration sensors were mounted in the area of impact along the production line. The amplitude of the acceleration signals and vibration durations were considered to determine the presence of target objects. A detection rate of approximately 75% was achieved. However, the production process was often disturbed by false alarms for this system [41].

Owing to the complexity and variety of detection schemes in production processes, here is assumed that the task cannot be solved satisfactorily using just one sensor technique. Moreover, due to the complexity of the transportation process, no single sensor technique can achieve the task directly. Thus, it is concluded that only indirect measurements or suitable combinations of individual measurements of relevant events (physical effects) of transported objects can be performed. The individual signals measured have to be preprocessed, evaluated, and combined to develop a reliable monitoring system to determine the presence of target objects.

One of the physical effects considered is force, which induces impact responses during transportation of target objects. These responses should be measured by sensors (five acceleration sensors here) in the impact area along the transportation line. These signals were investigated to build a detection module (an acceleration module here). Other physical effects measured will be used to develop further detection modules. The different output statements of the individual detection modules have to be fused using an appropriate method to obtain a reliable and accurate decisions on the presence of target objects. The different detection modules and the fusion module constitute the proposed monitoring system.

This paper investigated the suitable design concepts and multisensor data and decision fusion principles. Two different detection approaches for the intended acceleration module were developed and compared. Signals from the acceleration sensors were considered first. The inevitable time shift between the object impact stimulations of the individual sensors varies making the fusion of the process information difficult. Therefore, signal preprocessing, feature extraction, and classification for the individual acceleration sensors were used. Then a decision fusion process based on specific decision criteria was applied to combine the preliminary individual decisions of different classifiers (target object present yes/no) (Fig. 1).

Two different detection approaches for the acceleration module were investigated. The first uses a short-time Fourier transform (STFT) as a prefilter and a support vector machine (SVM) as a classifier. The second approach uses a continuous wavelet transform (CWT) as a prefilter and SVM as a classifier. The decision fusion process in both approaches is realized using different criteria. Both approaches are described and discussed in the following sections.

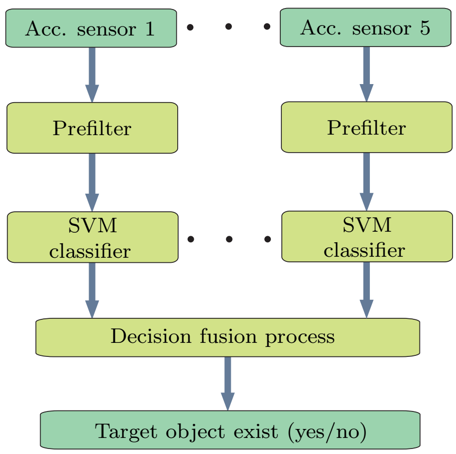


Figure 1. Detection approaches based on STFT-SVM and CWT-SVM.

2. Theoretical background

For successful classification of a system state, the data should be prepared by careful transformation to extract classification indicators. Irrelevant and redundant information should be excluded to avoid deterioration of the accuracy. The classification indicators containing useful information should be presented in a recognizable structure suitable for classification, referred to as features. The essential task in fault detection and identification is to classify features into a number of classes.

In the next chapter, a brief theoretical revision of the methods of feature extraction and classification used in this work are presented and discussed.

2.1. Feature extraction transforms

The main reason for combining feature extraction with classification is to make the classification of different classes or states easier.

Signals obtained during system monitoring usually consist of three major components: a periodic component resulting from interaction between elements of the cycling dynamics in a mechanical system; a transient component caused by non-stationary events such as crack initiation; and broadband background noise. In applications involving vibrational signals such as the system discussed here, these signal components are typically associated with a large variety of related frequencies. Thus, feature extraction should be based on a suitable domain-specific transform module such as STFT or wavelet transformation. The goal of such signal processing algorithms is to transform a time-domain signal into a suitable domain to extract those characteristics that are embedded in the time series which cannot be directly observed in the original form [7, 22, 46]. Mathematically, this can be achieved by representing a signal $x(t)$ as a series of parameters and inner product coefficients according to comparison of the signal to a set of known template functions [9, 15, 19]. However the monitoring problem considered here involves detecting non-stationary events in acceleration signals resulting from the presence of target objects within the transported material. The time-frequency analysis is used for the analysis of the concerned signals [20, 24, 30]. This technique is widely used in different application areas, such as machinery fault detection and diagnosis [14, 25, 27], medical applications [16, 17, 33], and defense and security systems [21, 29, 32].

For Fourier transforms, the similarity of a time series signal to a series of sine and cosine template functions is evaluated. The Fourier transform represents only the average frequency information for the entire period of the signal analyzed and not the variation of its content over time.

To overcome the limitations of Fourier transforms, STFT uses a window function ($g(t)$ centered at τ) that slides over the signal along the time axis to perform a localized window-based Fourier transform according to

$$STFT(\tau, f) = < x, g_{\tau,f} > = \int x(t)g(t-\tau)e^{i2\pi ft}dt. \tag{1}$$

As a result, STFT transforms a time series signal into a two-dimensional time–frequency representation (Fig. 4(a)). Once the window function in STFT is chosen, the time and frequency resolutions over the entire time–frequency plane are fixed, leading to a trade-off between the time resolution and frequency resolution [3, 8].

The continuous wavelet transform (CWT) is an alternative method for generating a time–frequency representation of a time series. The wavelet transform allows for variable window sizes in analyzing different frequency components within the signal [8, 11, 28]. This allows good frequency resolution at low frequencies and good time resolution at high frequencies. In wavelet analysis, signals are compared to a set of template functions obtained from scaling (stretching or squeezing) and time shifting of a mother wavelet function $\psi(t)$. These scaled and shifted functions represent localized frequencies of varying durations of a sound signal or image details, for example [36, 39]. The superiority of wavelets is more tangible in the case of non-stationary measurements and the existence of non-stationarities in time [28, 45]. The CWT of a signal $x(t)$ projected into a two-dimensional, time-scale plane is represented as

$$wt(s, \tau) = < x, \psi_{s,\tau} > = \frac{1}{\sqrt{s}} \int_{\text{inf}}^{\text{inf}} x(t)\psi^*(\frac{t-\tau}{s})dt. \tag{2}$$

2.2. Classification with Support Vector Machine (SVM)

Since early applications in fault diagnosis [40, 42], SVM has yielded better results than other techniques such as neural networks, decision trees, and model-based reasoning approaches [43, 44, 51]. The method introduced by Cortes and Vapnik [13] is based on statistical learning theory and is considered one of the best techniques for pattern recognition.

Support vector machine (SVM) implementations have been demonstrated in a wide range of applications, including economics [23], text mining [49], medicine and biology [31, 37], remote sensing [34], image segmentation [50], in addition to machine fault diagnosis and condition monitoring [7, 47, 48].

The SVM was used for classification because of its good generalization ability and its robustness to outliers. The SVM generalization ability can be improved using the concept of large margin classification [18]. Unlike typical classification methods, SVM uses information on the separating margin while learning from a data set, which leads to improved separability between classes. The SVM is trained to maximize the margin, and thus the generalization ability is better under conditions such as scarce training data. Moreover, SVM training always finds a global solution, in contrast to neural networks, for example, for which many local minima usually exist [1, 2, 10]. The SVM training also appears to be easier and requires less parameter tuning. Moreover, geometric interpretation of the separating hyperplane in the SVM feature space provides better transparency and interpretability of the results than neural networks do.

For signal fusion tasks, the SVM feature space is used as a tool to realize a complimentary transformed description in which a combination of signals provides better insight into the problem and therefore better accuracy than direct consideration of individual signals. Another advantage of SVM is its robustness to outliers. Proper setting of the penalty parameter C, which controls the misclassification error, can suppress outliers and reduce the effect of increased noise. In neural networks, by contrast, outliers need to be eliminated before training [1].

The importance of the SVM robustness to outliers is more emphasized by high-dimensional data sets with large number of features. The performance of traditional classification methods such as neural networks often decreases as the number of features increases, which is referred to as the curse of dimensionality. To deal with this problem, dimensionality reduction and feature subset selection techniques are often applied as a data preprocessing step prior to classification. In case of SVM, the learning complexity is independent of the dimensionality of the input space [26]. Therefore, dimensionality reduction methods do not significantly increase SVM accuracy. A support vector machine (SVM) classifier with a small number of support vectors has good generalization ability, even in very high-dimensional spaces [26].

The learning task in SVM [1] involves finding the unknown nonlinear dependency mapping between high-dimensional input vectors x and scalar outputs y. In general, no information about the underlying joint probability function is available. The only information available is implicitly included as features in the training data set used. The solution for SVM problems correspond to minimizing the cost function

$$J = \frac{1}{2}W^T W + C\sum_{i=1}^{l}\zeta_i \tag{3}$$

with respect to

$$y_i(W^T\phi(x_i) + b) \geq 1 - \zeta_i \tag{4}$$

and

$$\zeta_i \geq 0, \tag{5}$$

where W is the coefficient vector for the separating hyperplane, C is a penalty parameter, ζ_i a slack variable associated with x_i, l is the number of data points, b is a scalar representing the bias term of the separating hyperplane, and ϕ is a mapping function. The SVM kernel functions are used to map input data from the input space to a higher-dimensional feature space, where the maximum margin separating the hyperplane is constructed by training (Fig. 2).

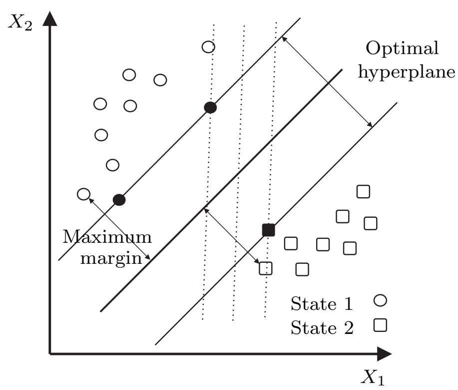


Figure 2. Feature space of a support vector machine.

3. Approach I: Detection system based on STFT and SVM

In the STFT-SVM approach, STFT is used to extract relevant information from the signals obtained from different acceleration sensors. The SVM is then used to classify the extracted features. In addition, a specific fusion process based on SVM and experimentally-based decision rules is developed and applied to combine the preliminary decisions of the individual sensors. The signals are individually prefiltered with STFT (Fig. 1 and Fig. 3). This prefiltering process is used to extract relevant features of the acceleration signals (Section 3.1). A set of supervised classification filters, denoted SVM I, is developed to classify the features extracted from each sensor signal. An adjustable decision fusion process is developed to combine the preliminary individual decisions of the different classifiers (Section 3.3). Feature extraction, classification, and decision fusion processes are described in detail below.

3.1. STFT-based feature extraction

The STFT extracts relevant information on system states. It serves to classify the information related to a single information path based on previously observed phenomena for a sensor signal. Fig. 4(a) (raw signal) shows a raw sample acceleration signal during the production process. At 12.5 s, a target object was manually classified. The impact of this object results in strong acceleration signal amplitudes. The other peaks in the signal are caused by other unknown events and are not object-induced signal changes. From the raw acceleration signals in the time domain it is very difficult to distinguish target object's effect from other unknown events. The different events can be classified on the basis of features extracted using prefilters (STFT). As shown by the spectrogram in Fig. 4(a), the target object causes strong excitation of low frequencies. By contrast, higher frequencies are excited more by unclassified or unknown events. This effect is used to distinguish events due to target objects from those related to other events.

3.2. Classification process

Three classification modules are included in the detection system (Fig. 3). The SVM-based algorithms are used to detect the system states. The Libsvm algorithm [12] is used to realize the SVM classifiers.

3.2.1. The Module SVM I

Due to the inevitable time delay between the excitation of the sensors which is constantly varying as a result of the structural dynamical behavior between the impact's and the sensor's locations. This affects the feature vectors. A classifier based on the SVM algorithm (SVM I) is developed for each individual acceleration signal. Data clearly indicating the presence of target objects are used to train SVM I. This should limit the false alarm rate, which is directly affected by the strength and intensity of the indicators used for training. Weak indicators lead to a higher rate of false alarms and vice versa. The decision functions of the individual classifiers generated by SVM I are input as preliminary decisions into subsequent stages to confirm the assumed system state.

3.2.2. The Module SVM II

When SVM I does not provide sufficient information, the decision on the presence of target objects is uncertain. In such a case, and in the cases mentioned in the following sections, a more local and precise investigation is necessary. Data with weak indications of the presence of target objects are used to train module SVM II; the limited area of application of the signal allows more flexible detection criteria.

The SVM II is trained with data consisting of two states. State 1, "target object present", is represented by training data with weak indications of the presence of target objects. State 2, "uncertain", is represented by training data with all other indications except state 1. The output statement of SVM II is either "target object present" or "uncertain".

3.2.3. The Module SVM III

A classification process based on SVM III is performed in cases for which SVM II provides an uncertain output statement. The SVM III provides further data classification for uncertain output statements. The SVM III is trained using data with clear indication of no target objects (state 1) and data with uncertain indications (state 2). The output statement of SVM III is either "target object not present" or "uncertain".

The SVM II and SVM III classifiers are used for more accurate trained classification locally in cases where further assessment of unclear decisions is required.

3.3. Adjustable decision fusion process

A new decision fusion process was developed to combine individual preliminary decisions and generate a final decision on the system state. The decision method is based on knowledge derived from analysis of experimental data from different acceleration sensors. It is designed to obtain the highest possible detection rate for the lowest possible false alarm rate. Therefore, tuning parameters are used to systematically adjust the fusion process.

The decision fusion filter consists of two rule-based filter levels and classification levels SVM II and SVM III (Fig. 3).

Classification level SVM I provides preliminary decision functions for the individual classifiers (Section 3.2.1). Depending on the individual preliminary decisions, the final decision of the acceleration module could be met by either rule-based filter I or by rule-based filter II.

The rule-based filter I consists of predefined rules that govern the final decision of the acceleration module. When it is impossible to achieve a reliable final decision from the acceleration module based on current values of the individual preliminary decision functions, specific rules to trigger further classification levels (SVM II and SVM III) are included in rule-based filter I. The rules for rule-based filter I are as follows:

- *Rule I*: At least two simultaneous positive individual decisions lead to a final decision of "positive: target object present".

- *Rule II*: Weak individual positive decisions with a decision value less than the experimentally defined threshold value T_1 are ignored, leading to a final decision of "negative: no target object present".

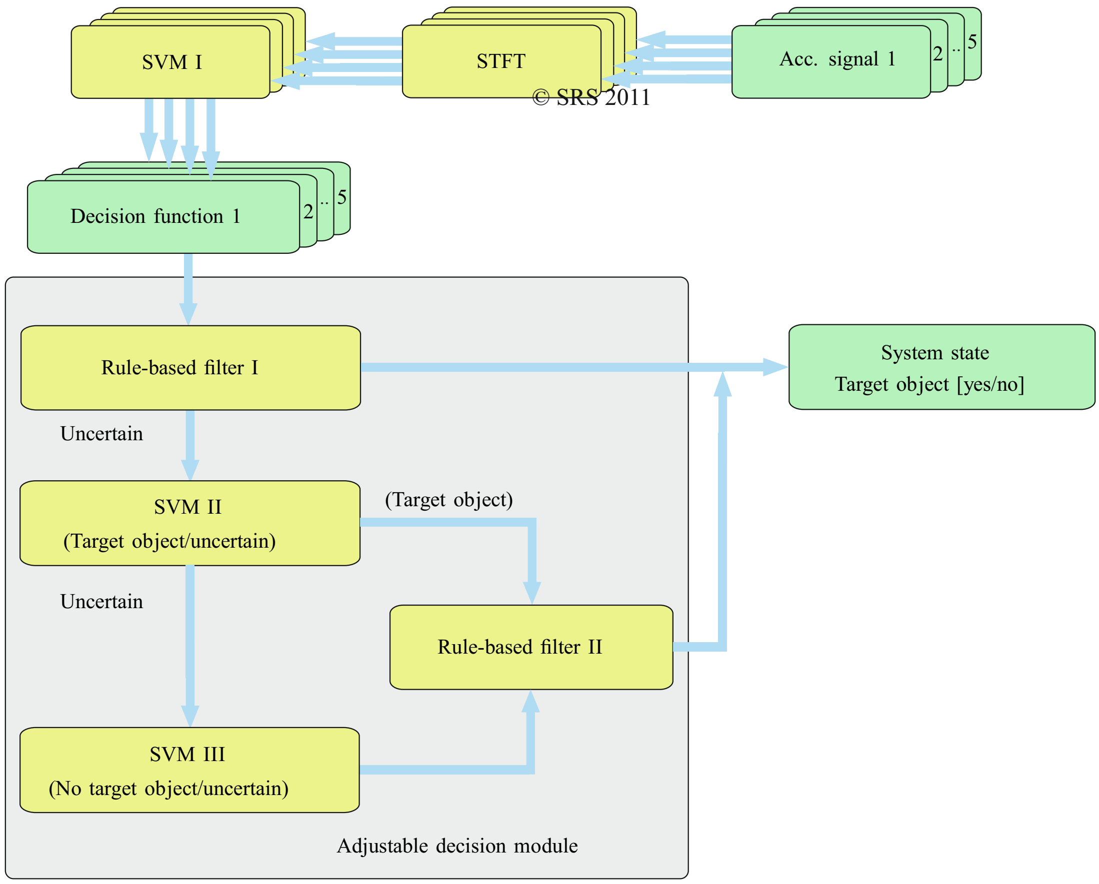


Figure 3. Adjustable decision fusion process.

Experimental evidence shows that weak individual positive decisions (with decision values less than T_1) are generally caused by events denoted as no target objects. Therefore this kind of decision is ignored and not considered as an indicator of a target object.

- *Rule III*: Individual positive decision with a decision value greater than the experimental threshold T_2 ($T_2 > T_1$) leads to a final decision of "positive: target object present".

- *Rule IV*: Individual positive decision with a decision value greater than T_1 and less than T_2 triggers further classification levels.

As mentioned before, classification levels SVM II and SVM III are activated if rule IV is fulfilled. When an individual positive decision has a decision value greater than T_1 and less than T_2, the other four acceleration signals that provide negative decision values (parallel to the single individual positive decision) are subjected to more accurate evaluation using either SVM II alone or both SVM II and SVM III. Further classification is performed locally on neighboring areas of the corresponding acceleration signals to confirm or disprove the correctness of the single individual positive decision.

The acceleration signals are first evaluated using SVM II. The SVM II output statement is either "target object present" or "uncertain". Thus, SVM II confirms the presence of target objects. The SVM III evaluation is performed for acceleration signals that yield an uncertain output statement from SVM II. These acceleration signals are evaluated for the presence of events denoted as "no object". The output statement of SVM III is either "target object not present" or "uncertain". Thus, SVM III confirms the absence of the target object.

The single individual preliminary decision from SVM I and the other four decisions provided by SVM II, SVM II and SVM III, or SVM III are combined using rule-based filter II. The rules for rule-based filter II are as follows:

- *Rule I*: If the number of output statements "target object present" is greater than the number of "uncertain" statement provided by SVM II and SVM III, the final decision is "positive: target object present" (majority rule).
- *Rule II*: An individual output statement "target object not present" from SVM III leads to a final decision of "negative: no target object".

The decision generated by the fusion module is self-adaptable and depends on re-evaluation of individual partial decisions. Owing to the complexity of the system to be monitored, it is difficult to provide these benefits using classical fusion techniques. The main reason for using the proposed method for fusion is the inevitable and varying time shift between the object impact stimulation of the individual sensors, and the fact that the effect of noise and disturbance signals to the system and sensors (both individual and group sensors) is inevitable. Therefore, the importance of individual decisions is retained by combining and comparing them with other decisions that do not necessarily coincide in time. This data handling requires a floating decision window, which increases the computational load.

The improvement in quality resulting from the developed fusion technique over single-stage SVM classifiers demonstrates the validity of the approach. This multi-stage technique with additional stages focuses mainly on classes that cannot be identified with suitable reliability, and therefore have to be considered separately and in detail.

It should be noted that the floating window and decision re-evaluation require buffer savings for the range of data considered. These buffer savings should cover the processing window and transformed data for the different channels and their decisions, as well as decisions generated by the fusion module. These buffer savings increase the memory requirements of the system and the time required for the final decision owing to the inefficient buffering time.

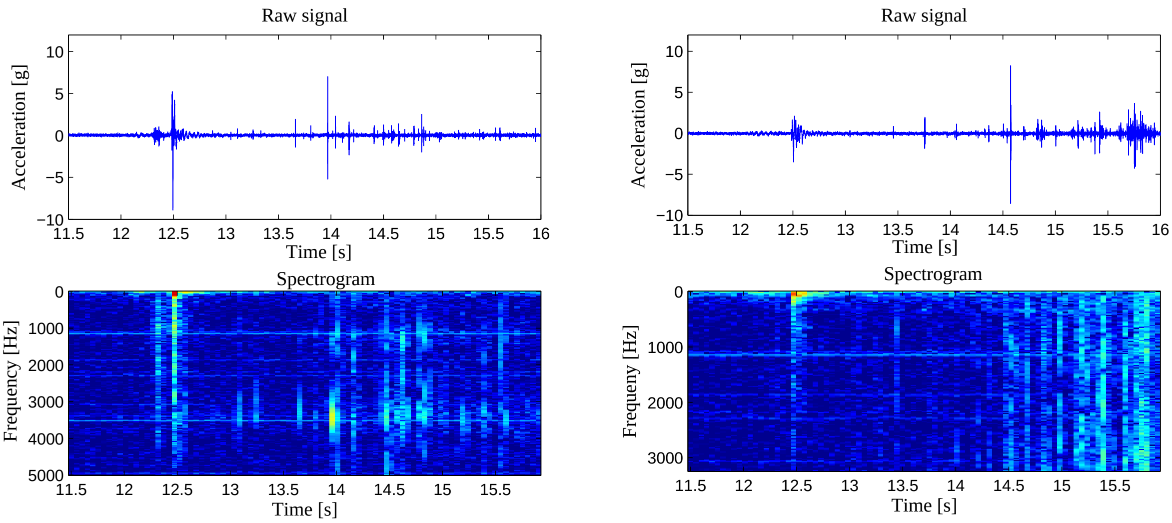
In spite of the complexity of the detection system, the main requirement is that the final decision should at least be faster than subsequent events to provide enough time to isolate target objects that are detected. This requirement is fulfilled according to the implementation results presented below.

3.4. First industrial implementation and results

The STFT algorithm is used as a prefilter to extract relevant features from acceleration signals. The spectrograms generated (Fig. 4(a) and 4(b)) show the features extracted (frequencies in the range up to 3250 Hz) as functions of time. The impulse intensities for each feature are represented by a suitable color map. The feature vector is based on 511 features, one for each acceleration signal.

The spectrogram in Fig. 4(a) reveals that frequencies between 1 and 200 Hz are dominant at 12.5 s (red) owing to the impact of a target object. The specific behavior depends on the structural dynamic characteristics of the contact surface on which the sensor is mounted and on the impact position. Higher frequencies at 12.5 s show less energy (light blue) than frequencies below 200 Hz.

The amplitude behavior of the acceleration signal at 13.95 s (Fig. 4(a)) in the raw time-domain signal is in principle similar to that at 12.5 s; however the range of frequencies excited in the STFT features is different (2560–4200 Hz). Thus, it is experimentally observed



(a) Acceleration signal for sensor 1 (b) Acceleration signal for sensor 2

Figure 4. Acceleration signal for sensors 1 and 2

that the machine structure reacts differently to target object events and allows a statement about the presence of objects. Therefore, target objects and other events can be detected, classified, and separated. If the resonance properties of the structure at the collision point, taking into consideration the sensor position, are known and are considered to allow reliable distinction between relevant frequency ranges and related impact power, the distinction is considered reliable. In reality, this distinction has been observed as very robust for different sensor locations.

In the raw signal in Fig. 4(b), the event observed at 12.5 s and the disturbance effects between 15.6 and 16 s show similar amplitudes and behavior in the time domain. It is expected that the excited structural dynamics at 12.5 s responds differently (in the frequency domain 0–200 Hz (Fig. 4(b))).

The approach was tested using an experimental set of real industrial data. The results for preliminary application to the system are summarized in Table 1. The best individual detection accuracy is 58.3% (classifiers 1 and 4). Classifier 5 leads to the lowest accuracy and false alarm rates, although it has the smallest number of support vectors, indicating comparatively low levels of noise. This result demonstrates a typical compromise: an increase in the detection rate leads to an increase in the false alarm rate (Table 1).

Table 1. Classification results for the STFT-SVM approach

Training data					
Sensor/classifier number	1	2	3	4	5
Target objects	17	19	17	17	16
No. of support vectors	193	210	177	158	79
SVM kernel			Linear		
Individual results for the test data					
Target objects			36		
Objects detected	21	20	19	21	16
Accuracy [%]	58.3	55.5	52.8	58.3	44.4
False alarms	18	14	7	4	2
Fusion results for the test data					
Objects detected			27		
Accuracy [%]			75		
False alarms			7		
False alarms/number of objects [%]			19.4		

The accuracy of the system based on fused decisions for the acceleration sensor network is 75%, which represents an improvement of at least 16.7% over the individual accuracy rates. This improvement in accuracy indicates that the individual sensors have different views, depending on their mounting position and their relation to the materials transported, including target objects. This means that each individual classifier can detect target objects that possibly went undetected by other classifiers. The rate of false alarms can be compromised, however, because the false alarms for individual sensor paths are not necessarily identical. The fusion approach not only improves the detection rate, but also leads to a strong reduction in the number of false alarms (Table 1). During development of the detection system, and considering the requirements of the mechanical system, a compromise between accuracy and rate of false alarms must be achieved.

4. Approach II: Detection system based on CWT and SVM

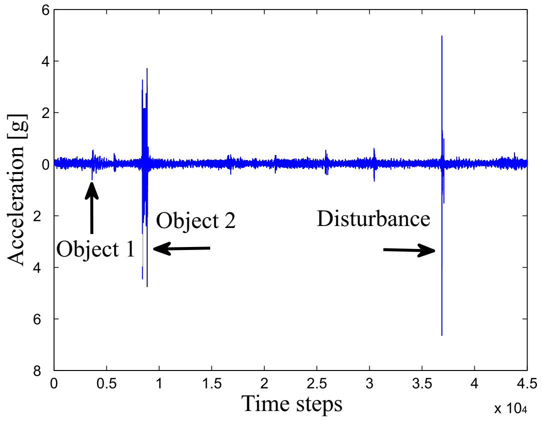
The CWT-SVM approach uses CWT to extract relevant information from acceleration signals for the different sensor channels. The SVM is used to classify the features extracted. A fusion process is applied to combine the individual decisions of the different SVM classifiers (Fig. 1).

The individual sensor signals are prefiltered separately using CWT (Fig. 1). The features extracted for individual sensors are subjected to multistage filtering. The feature extraction and decision fusion processes are described below.

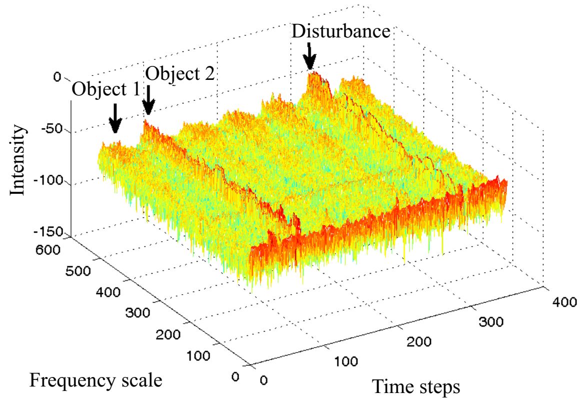
4.1. CWT-based feature extraction

As an example, Fig. 5(a) shows the acceleration signal for sensor 1. To illustrate the solution concept and for the purpose of comparison [4], the signal was filtered using STFT (Fig. 5(b)) and CWT (Fig. 5(c)). The signal has two events, denoted as objects 1 and 2 at time points 4000 and 8500, respectively, that were manually classified as target objects. These events appear at time points 30 and 70 in the STFT extracted feature space (Fig. 5(b)) and at time points 4000 and 8500 in the CWT extracted feature space (Fig. 5(c)). A third event resulting from an unknown disturbance is evident at time point 37000 in Fig. 5(a). Since this event was not classified manually, it cannot denote a target object. It is also evident in the STFT and CWT extracted feature spaces [5].

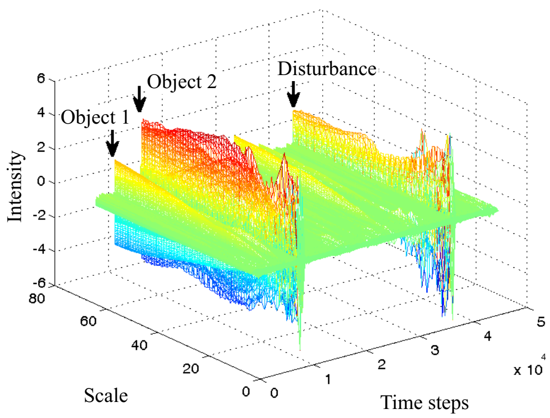
The object at time point 8500 and the event at time point 37000 can be clearly recognized in the STFT and CWT results. In the STFT results the two events seem to be similar, whereas in the CWT results they appear to be different. In the case of CWT, the higher scales (low frequencies) for the second object are excited more strongly than the lower ones, while the lower scales (higher frequencies) are more strongly excited in the case of the disturbing event.



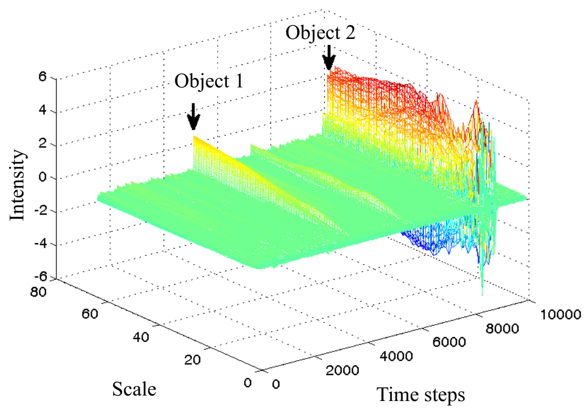
(a) Acceleration signal for sensor 1



(b) STFT decomposition for sensor signal 1



(c) CWT decomposition for sensor signal 1



(d) First part of the CWT decomposition

Figure 5. Comparison of STFT and CWT

Unlike the case in which a target object is present, the higher scales (lower frequencies) of the disturbance are related to lower energy than the lower scales (higher frequencies). The reason is that the range of the activated frequency and accordingly the center of frequency, which coincides with the energy peak are different for the disturbance and target object. This advantage of the CWT approach is used as a base rule for further filtering steps.

To illustrate this, consider the first target object in the sample data (object 1 at time 4000) that cannot be clearly distinguished from the time series signal (Fig. 5(a)). The object is difficult to detect because of its impact on the mechanical structure, which is obviously dampened by the accompanying materials (overburden). In the STFT results (Fig. 5(b)), the presence of the object is characterized by weak excitation of low frequencies. In the CWT results (Fig. 5(c)), the object can be better recognized and characterized by a longer band of high scales (low frequencies) of a specific shape. The effects of objects 1 and 2 are evident in Fig. 5(c) and are magnified in Fig. 5(d).

Several noise and disturbance sources are involved in this complex and unstable production process. These lead to difficulties in recognizing target objects. The disturbance can be stationary background noise or non-stationary noise, with large or rapid spectral changes over time, and can therefore resemble events resulting from target objects.

Fig. 6(a) shows an acceleration signal resulting from CWT as a function of time and frequency. The signal includes different events. The marked event is the only one that needs to be detected. All other events are caused by different noise events.

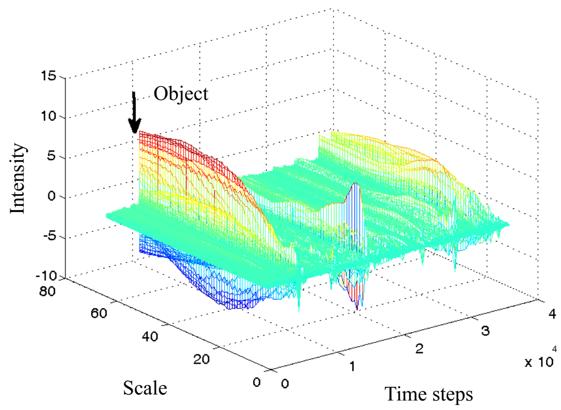
For efficient learning and reliable classification, further filtering is applied to reduce the data complexity. This filtering eliminates known noisy events, as described below.

4.1.1. Prefilter I: Prefiltering the background noise

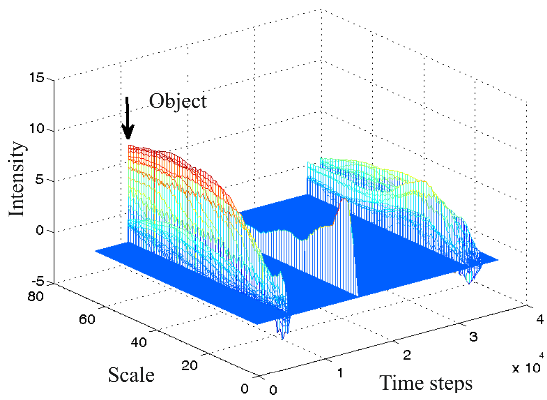
An acceleration signal contains permanent background noise in both the time and frequency domains (Fig. 6(a)). The presence of events in the training data set complicates SVM training.This complexity results due to difficulty in labeling training data because of the interpenetration of the different events. Inaccurate labeling of the training data affects the classifier performance. The aim of prefilter I is to eliminate stationary background noise from the extracted features to avoid this problem and simplify SVM classifier training. This involves eliminating events in the acceleration signals for time points at which the maximum intensity of low-frequency excitation is less than an experimental threshold according to the prefiltering rule. If this condition is fulfilled, the intensity for all frequencies at such time points is set to zero (Fig. 6(b)).

4.1.2. Prefilter II: Prefiltering specific known noise behavior

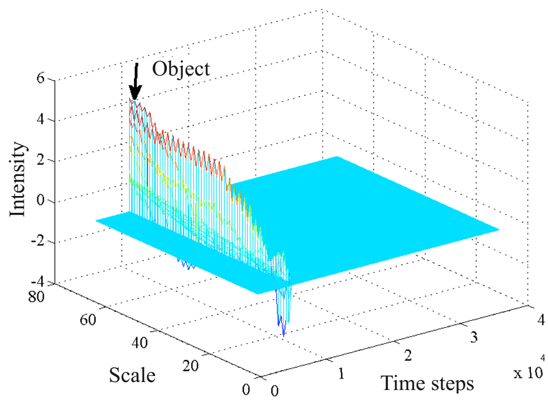
The distinguishing characteristics are used to filter process-related non-stationary disturbances. The main characteristics differ for detection cases and for disturbances related



(a) CWT decomposition of an acceleration signal



(b) CWT decomposition of the acceleration signal after applying first prefiltering rule



(c) CWT decomposition of the acceleration signal after applying second prefiltering rule

Figure 6. Prefiltering results of prefilter I and II

to gradients and event forms. The gradient from the low-frequency to the high-frequency domain is estimated at each time point and compared to an experimental threshold. The frequency intensity is set to zero at those time points for which the gradient does not exhibit the same behavior as target events. In addition, the relative signal intensity for the low scales is checked and compared to the behavior of the target events. Experimental evidence has shown that certain behaviors for the starting scales can be known as they do not belong to target objects but to specific noise events. This characteristic was confirmed using manual classification of the target objects to be removed.

It should be noted that event strength is not necessarily a reliable characteristic for distinguishing between events to be detected and disturbances owing to the nature of the production process. Some target objects cause weaker effects than those caused by disturbances. Therefore, scale-invariant recognition based on the event form is more reliable and effective.

An example data set filtered using prefilter II is presented in Fig. 6(c). The second condition checks whether the gradient at each time point exhibits the behavior of the target events.

4.2. Classification process

The Libsvm algorithm [12] was used for numerical realization of the SVM classifiers. An experimental data set was prepared using wavelet-based prefilters to build the training data set. The SVM classifier model was then built based on this set and used for classification of the five different acceleration signals.

4.3. Decision fusion process

The decision fusion process based on experimental data combines individual preliminary decisions to reach a final decision. The decision fusion filter is a rule-based filter. The rules are as follows:

- *Rule I*: At least two simultaneously positive individual decisions lead to a final decision of "positive: target object present".
- *Rule II*: Several strong positive decisions from at least one of the five classifiers within a specific floating decision window lead to a final decision of "positive: target object present". Several relatively strong positive decisions imply that the number of positive peaks within the floating decision window is greater than an experimental threshold. In addition, the decision value should be greater than an experimental threshold.

4.4. First industrial implementation and results

Results for CWT-based detection are summarized in Table 2. The best individual accuracy is 58.3% (classifiers 1 and 2), but this corresponds to a higher rate of false alarms. It should be noted that increases in the training accuracy for individual classifiers will lead to improvements in the detection accuracy; the rate of the false alarms would also increase accordingly. During development and depending on the system requirements, a compromise between detection accuracy and false alarms must be achieved.

Table 2. Classification results for the CWT-SVM approach

Training data					
Sensor/classifier number	1	2	3	4	5
Target objects			16		
No. of support vectors			1197		
SVM kernel			RBF		
Individual results for test data					
Target objects			36		
Objects detected	21	21	20	18	13
Accuracy [%]	58.3	58.3	55.5	50.0	36.1
False alarms	7	11	12	6	1
Fusion results for test data					
Objects detected			23		
Accuracy [%]			63.9		
False alarms			4		
False alarms/number of objects [%]			11		

The individual and fused results reveal that classifier fusion leads to a reduction in the false alarm rate by approximately 89%. The final detection rate is approximately 6% better than for the best individual classifier.

5. Discussion and comparison of approaches

The efficiency of any monitoring system is generally evaluated according to the detection accuracy and the false alarm rate. The challenge for any detection approach is to achieve the highest possible detection rate with the lowest possible false alarm rate. The complexity of the monitoring system should also be considered in evaluations. Complicated systems involve more unexpected defects and flaws than simple ones because of the unpredictable nature of disturbances which is not rigorously perceptible. In addition, implementation of more complicated systems usually requires more effort to realize and can involve greater difficulties in real-time applications.

According to the results in Tables 1 and 2, which use the same test data set to realize the testing phase, the STFT-SVM approach detects at least 10% more target objects compared to the CWT-SVM approach. The detection rate is 75% for the STFT-SVM approach and only 64% for the CWT-SVM approach. This does not necessarily mean that the STFT-SVM approach is better than CWT-SVM, because the highest number of false alarms for STFT-SVM is approximately double that for CWT-SVM. It should be noted that the sensitivity of the training process affects the rate of detection and the rate of false alarms in various ways. Increasing the sensitivity of the model can increase the rate of detection, as well as the number of false alarms. In fact, a trade-off exists in the relationship between detection and false alarms. This means that both approaches have the same improvement potential with respect to detection and false alarm rates.

Other important issues regarding realization and implementation are as follows. In the first approach (STFT and SVM) a number of 7 SVM classification algorithms (5 of them in parallel) are developed in order to realize the proposed fusion modules, whereas in the second one (CWT and SVM) only one SVM classification algorithm is developed and used for the five individual preliminary classifiers in the system. Moreover, the two sets of multiple rules required for STFT-SVM and the backward evaluation scheme are more complicated than the simple combination rules for the CWT-SVM approach. These are strong factors in judging whether the CWT-SVM approach is more efficient and reliable for real-time applications.

The STFT-SVM approach is indeed difficult to design and develop and much processing effort is required for implementation. In spite of the comparable results for detection and false alarm rates, the CWT-SVM approach is more convenient to realize and implement and more appropriate for real-time applications.

For the two approaches, it is evident that the system was successful in isolating the required objects from noise events. After processing and transformation, many noise events acquire different shapes and become visually different from the target object; however, some noise events do not. Such noise events might be detected simultaneously by many sensors, which can lead to false alarms. Such noise events cannot be avoided. They are usually caused by objects that are similar to the targets but smaller in size.

Since the object size does not necessarily coincide with the event intensity, it is difficult to differentiate these smaller objects by simple adjustment of the system sensitivity, which could lead to deterioration in the performance and detection rate.

It should be noted that the accuracy achieved is not always that targeted for the intended monitoring system. These accuracies represent individual accuracies for the acceleration sensors among other detection modules based on other data acquisition techniques. All these modules are fused together to realize the monitoring system. In this contribution, the principles behind are developed and compared to previous results for similar applications.

In the following, some additional results justify the fusion of all five sensor channels instead of using more simple combinations of two decisions, for instance.

The classifiers based on sensors 1 and 2 were considered and the STFT-SVM approach was applied. In contrast to the 27 objects detected after fusion of five sensors, only 16 target objects were detected as the best achievable result; however, the number of false alarms decreased from seven to six. After fusion of sensors 3 and 4, the detection rate was 11 objects out of 36, with two false alarms. To compare the sensors 1 and 2 with the sensors 3 and 4, the individual numbers of false alarms are added for sensors 1 and 2 as 32 false alarms, and for sensors 3 and 4 as 11 false alarms. The ratios of the fused number of false alarms (6 for sensors 1 and 2, and 2 for sensors 3 and 4) to the sum of the individual number of false alarms (32 for sensors 1 and 2, and 11 for sensors 3 and 4) are almost similar (approximately 0.18).

For the CWT-SVM approach, sensors 1 and 2, and sensors 3 and 4 were fused. The results were as follows. Only 19 objects were detected by fusing sensors 1 and 2 and the number of false alarms was reduced to 1. For fusion of sensors 3 and 4, 18 objects out of 36 were detected, with no false alarms. The ratios of the fused number of false alarms (1 for sensors 1 and 2, and 0 for sensors 3 and 4) to the sum of the individual number of false alarms (18 for sensors 1 and 2, and 18 for sensors 3 and 4) are much smaller than that obtained through STFT-SVM approach. These results clearly indicate the better fusibility of the individual sensors and the lower redundancy.

In general, the individual sensors deliver a high number of false alarms for both approaches, so detection is unreliable. By combining two individual sensors, better reliability can be achieved with fewer false alarms. The detection reliability can be improved by decreasing the number of false alarms and increasing the rate of detection. This is achieved with a fusion process for the five sensors and related filters, in order to achieve deep insight into the problem diversities and distinctions. Although fusion leads to a slight increase in the number of false alarms, this deterioration is negligible compared to the overall improvement in system performance. The deterioration can be explained by redundancy and the accumulation of disturbances when many sensors are considered together. The improvement potential of fusion lies in the quality of the sensor data obtained and the feature extraction process used.

6. Conclusion

Two different approaches for feature-based multisensor fusion were developed for a monitoring system. The goal was to monitor a production process for online detection of a target object with the lowest possible false alarm rate. The application is a stationary process with a continuous stream of sensor signals. The two approaches were compared in terms of performance and complexity.

The STFT-SVM approach uses STFT for feature extraction. The individual feature vectors for different acceleration sensors are classified using SVM to provide individual preliminary decisions about the presence of a target object. These decisions are fused using a novel fusion module to provide a reliable decision about the system state.

The CWT-SVM approach uses CWT as a prefilter to extract features, which are subjected to further prefiltering processes before SVM classification.

The results for both approaches reveal improved detection rates and fewer false alarms. Approach I leads to a better improvement in detection rate compared to individual preliminary decisions. The number of false alarms is also lower. Approach II leads to an even better improvement in the false alarm rate than approach I.

References

[1] S. Abe, Support vector machines for pattern classification, Springer-Verlag London Limited, London, England, 2010.

[2] G. Acuna, M. Curilem, Comparison of neural networks and support vector machine dynamic models for state estimation in semiautogenous mills, in: A. Aguirre, R. Borja, C. Garci (Eds.), MICAI 2009: Advances in artificial intelligence, volume 5845 of *Lecture Notes in Computer Science*, Springer Berlin Heidelberg, 2009, pp. 478–487.

[3] F. Al-Badour, M. Sunar, L. Cheded, Vibration analysis of rotating machinery using time-frequency analysis and wavelet techniques, Mechanical Systems and Signal Processing 25 (2011) 2083–2101.

[4] L. Al-Shrouf, M. Saadawia, N. Szczepanski, D. Söffker, Adaptive classification based on multisensoric decision fusion, in: F. Chang (Ed.), 8th Int. Workshop on Structural Health Monitoring 2011, Stanford, USA, pp. 127–134.

[5] L. Al-Shrouf, M. Saadawia, N. Szczepanski, D. Söffker, Event detection using multisensor fusion and filtering techniques based on CWT and SVM, in: C. Boller (Ed.), 6th European Workshop on Structural Health Monitoring 2012, Dresden, Germany, pp. 206–213.

[6] H. Aljouma, D. Söffker, Multi-class approach based on fuzzy-filtering for condition monitoring, IAENG International Journal of Computer Science 38 (2011) 66–73.

[7] T. Banerjee, S. Das, Multi-sensor data fusion using support vector machine for motor fault detection, Information Sciences 217 (2012) 96–107.

[8] G. Bhatnagar, Q. Wu, B. Raman, Fractional dual tree complex wavelet transform and its application to biometric security during communication and transmission, Future Generation Computer Systems 28 (2012) 254–267.

[9] G. Bhatnagar, Q. Wu, B. Raman, Discrete fractional wavelet transform and its application to multiple encryption, Information Sciences 223 (2013) 297–316.

[10] C. Burges, A tutorial on support vector machines for pattern recognition, Data Mining and Knowledge Discovery 2 (1998) 121–167.

[11] M. Canal, Comparison of wavelet and short time fourier transform methods in the analysis of EMG signals, Journal of Medical Systems 34 (2010) 91–94.

[12] C. Chang, C. Lin, LIBSVM: A library for support vector machines, ACM Transactions on Intelligent Systems and Technology 2 (2011) 1–27. Software available at http://www.csie.ntu.edu.tw/ cjlin/libsvm.

[13] C. Cortes, V. Vapnik, Support vector networks, Machine Learning 20 (1995) 273–297.

[14] Z. Feng, M. Liang, F. Chu, Recent advances in time-frequency analysis methods for machinery fault diagnosis: A review with application examples, Mechanical Systems and Signal Processing 38 (2013) 165–205.

[15] R. Gao, R. Yan, Wavelet theory and applications for manufacturing, Springer Science + Business Media, New York, USA, 2011.

[16] A. Gramfort, D. Strohmeier, J. Haueisen, M. Hämäläinen, M. Kowalsk, Time-frequency mixed-norm estimates: Sparse M/EEG imaging with non-stationary source activations, NeuroImage 70 (2013) 410–422.

[17] L. Guo, D. Rivero, J. Dorado, C. Munteanu, A. Pazos, Automatic feature extraction using genetic programming: An application to epileptic EEG classification, Expert Systems with Applications 38 (2011) 10425–10436.

[18] I. Guyon, I. Elisseeff, An introduction to variable and feature selection, Journal of Machine Learning Research 3 (2003) 1157–1182.

[19] D. Hammonda, P. Vandergheynst, R. Gribonval, Wavelets on graphs via spectral graph theory, Applied and Computational Harmonic Analysis 30 (2011) 129–150.

[20] F. Hlawatsch, F. Auger, Time-frequency analysis concepts and methods, John Wiley & Sons, Inc, London, England, 2008.

[21] H. Hu, Augmented DT-CWT feature based classification using regularized neighborhood projection discriminant analysis for face recognition, Pattern Recognition 44 (2011) 519–531.

[22] C. Huang, D. Yu, H. Guan, G. Li, Q. Wu, Feature extraction of frictional vibration based on CWT time-frequency image, in: Intelligent Control and Information Processing (ICICIP), 2010 International Conference on, Dalian, China, pp. 192–196.

[23] K. Hung, K. Lin, Long-term business cycle forecasting through a potential intuitionistic fuzzy least-squares support vector regression approach, Information Sciences 224 (2013) 37–48.

[24] A. Jardine, D. Lin, D. Banjevic, A review on machinery diagnostics and prognostics implementing condition-based maintenance, Mechanical Systems and Signal Processing 20 (2006) 1483–1510.

[25] D. Jena, S. Panigrahi, R. Kuma, Gear fault identification and localization using analytic wavelet trans-

form of vibration signal, Measurement 46 (2013) 1115–1124.

[26] V. Kecman, Support vector machines - an introduction, in: L. Wang (Ed.), Support vector machines: Theory and applications, volume 177 of *Studies in Fuzziness and Soft Computing*, Springer Berlin Heidelberg, Berlin, Germany, 2005, pp. 1–47.

[27] P. Konar, P. Chattopadhyay, Bearing fault detection of induction motor using wavelet and support vector machines (SVMs), Applied Soft Computing 11 (2011) 4203–4211.

[28] J. Lee, S. Lee, I. Kim, H. Min, S. Hong, Comparison between short time fourier and wavelet transform for feature extraction of heart sound, in: Y.K. Kim (Ed.), Tencon 99. Proceedings of the IEEE Region 10 Conference, volume 2, The Silla Cheju, Cheju Island, Korea, pp. 1547–1550.

[29] P. Lei, J. Wang, P. Guo, D. Cai, Automatic classification of radar targets with micro-motions using entropy segmentation and time-frequency features, AEU - International Journal of Electronics and Communications 65 (2011) 806–813.

[30] P. Li, F. Kong, Q. He, Y. Liu, Multiscale slope feature extraction for rotating machinery fault diagnosis using wavelet analysis, Measurement 46 (2013) 497–505.

[31] H. Lin, L. Tseng, Prediction of disulfide bonding pattern based on a support vector machine and multiple trajectory search, Information Sciences 199 (2012) 167–178.

[32] L. Liu, D. McLernon, M. Ghogho, W. Hu, J. Huang, Ballistic missile detection via micro-doppler frequency estimation from radar return, Digital Signal Processing 22 (2012) 87–95.

[33] Y. Liu, U. Aickelin, J. Feyereisl, L. Durrant, Wavelet feature extraction and genetic algorithm for biomarker detection in colorectal cancer data, Knowledge-Based Systems 37 (2013) 502–514.

[34] G. Mountrakis, J. Im, C. Ogole, Support vector machines in remote sensing: A review, ISPRS Journal of Photogrammetry and Remote Sensing 66 (2011) 247–259.

[35] T. Nieß, Automatic metal and stone detection on bucket-wheel excavators used in RWE power AGs opencast lignite mines, World of Mining 61 (2009) 226–233.

[36] H. Nobuhara, D. Trieu, T. Maruyama, B. Bede, Max-plus algebra-based wavelet transforms and their FPGA implementation for image coding, Information Sciences 180 (2010) 3232–3247.

[37] G. Orru, W. Pettersson-Yeo, A. Marquand, G. Sartori, A. Mechelli, Using support vector machine to identify imaging biomarkers of neurological and psychiatric disease: A critical review, Neuroscience & Biobehavioral Reviews 36 (2012) 1140–1152.

[38] F. Petrich, U. Köhler, Stone detection and stone excavation - results and current directions of work with regard to technical stone handling in lusatian opencast mines, World of Mining 57 (2005) 412–420.

[39] S. Rein, M. Reisslein, Identifying the classical music composition of an unknown performance with wavelet dispersion vector and neural nets, Information Sciences 176 (2006) 1629–1655.

[40] M. Rychetsky, S. Ortmann, M. Glesner, Support vector approaches for engine knock detection, in:

Neural Networks, 1999. IJCNN '99. International Joint Conference on, volume 2, Washington, USA, pp. 969–974.

[41] J. Saß, F. Matare, T. Schubert, Metallsuchgerät und Steinerfassungssystem, Bagger 258 und 261 Tagebau Garzweiler, Bagger 259, 260 und 290 Tagebau Hambach, Bagger 255 Tagebau Inden, Internal report, Siemens AG, 2007.

[42] C. Saunders, A. Gammerman, H. Brown, G. Donald, Application of support vector machines to fault diagnosis and automated repair, in: 11th Int. Workshop on Principles of Diagnostics (DX '00), Beijing, China, pp. 8–10.

[43] M. Seyagh, M. El-Mostapha, A. Jarid, D. Cherqaoui, A. Schmitzer, D. Villemin, Pattern recognition: Application of support vector machines, artificial neural networks and decision trees for anti-HIV activity prediction of organic compounds, in: Multimedia Computing and Systems (ICMCS), 2011 International Conference on, Marrakech, Morocco, pp. 1–6.

[44] Y. Shao, R. Lunetta, Comparison of support vector machine, neural network, and CART algorithms for the land-cover classification using limited training data points, ISPRS Journal of Photogrammetry and Remote Sensing 70 (2012) 78–87.

[45] R. Supangat, N. Ertugrul, W. Soong, D. Gray, C. Hansen, J. Grieger, Broken rotor bar fault detection in induction motors using starting current analysis, in: Power Electronics and Applications, 2005 European Conference on, Adelaide SA, Australia, pp. 1–10.

[46] G. Vachtsevanos, F. Lewis, M. Roemer, A. Hess, B. Wu, Intelligent fault diagnosis and prognosis for engineering systems, Wiley, 2006.

[47] A. Widodo, B. Yang, Support vector machine in machine condition monitoring and fault diagnosis, Mechanical Systems and Signal Processing 21 (2007) 2560–2574.

[48] J. Xiao, H. Wang, X. Yang, Z. Gao, Multiple faults diagnosis in motion system based on SVM, Machine Learning and Cybernetics 3 (2012) 77–82.

[49] S. Yen, Y. Wu, J. Yang, Y. Lee, C. Lee, J. Liu, A support vector machine-based context-ranking model for question answering, Information Sciences 224 (2013) 77–87.

[50] Z. Yu, H. Wong, G. Wen, A modified support vector machine and its application to image segmentation, Image and Vision Computing 29 (2011) 29–40.

[51] E. Zanaty, Support vector machines (SVMs) versus multilayer perception (MLP) in data classification, Egyptian Informatics Journal 13 (2012) 177–183.