Available online at www.sciencedirect.com

ScienceDirect

Procedia Manufacturing 19 (2018) 42–49

Procedia MANUFACTURING

www.elsevier.com/locate/procedia

6th International Conference on Through-life Engineering Services, TESConf 2017, 7-8 November 2017, Bremen, Germany

An Image Processing Approach to Machine Fault Diagnosis Based on Visual Words Representation

Jianjing Zhang, Peng Wang, Robert X. Gao*, and Ruqiang Yan

Department of Mechanical and Aerospace Engineering, Case Western Reserve University, Cleveland, OH, 44120, USA

Abstract

Machine fault diagnosis and remaining service life prognosis provide the basis for condition-based maintenance, and is key to operational reliability. Accurate assessment of machine health requires effective analysis of vibration data, which is typically performed by examining the change in frequency components. One limitation associated with these methods is the empirical knowledge required for fault feature selection. This paper presents an image processing approach to automatically extract features from vibration signal, based on visual words representation. Specifically, a time-frequency image of vibration signal is obtained through wavelet transform, which is then used to extract "visual word" features for recognizing fault related patterns. The extracted features are subsequently fed into sparse representation-based classifier for classification. Evaluation using experimental bearing data confirmed the effectiveness of the developed method with a classification accuracy of 99.7%.

Peer-review under responsibility of the scientific committee of the 6th International Conference on Through-life Engineering Services.

Keywords: Condition Monitoring, Pattern Recognition, Reliability Engineering

Introduction

Through-life Engineering Services (TES), which provides support through the product design, manufacturing, redesign, remanufacture, recycle and reuse stages, has become an integrated part of the product lifecycle itself, as illustrated in Fig. 1 [1]. It takes input from the knowledge of design function, manufacturing method, degradation

* Corresponding author. Tel.: 1-216-368-6045; fax: +1-216-368-6445.
E-mail address: Robert.Gao@case.edu

2351-9789
Peer-review under responsibility of the scientific committee of the 6th International Conference on Through-life Engineering Services.
10.1016/j.promfg.2018.01.007

mechanism, fault diagnostics and remaining service life prognostics, and brings benefits to all product lifecycle stages.

Being a key element in the use phase of the TES, machine condition monitoring contributes to the overall manufacturing sustainability, by 1) providing technical base for operation and maintenance decision making for improved operational reliability and efficiency and 2) providing information to optimize product design and extend its service life based on the usage experience. The theoretical analysis of the machine failure mechanism has helped researchers gain the knowledge of fault detection and diagnosis when a machine deviates from its normal operating condition, leading to the traditional diagnosis techniques based on the manual extraction of the fault-related patterns in the signal's time-frequency spectra that match the analytical or empirical knowledge of the fault [2].

As the manufacturing industry embraces Big Data era, the specific analytical and empirical knowledge of fault mechanism and corresponding reflections on measured sensor data quickly become elusive with respect to arbitrary application scenarios. The task can be further complicated by the addition of fault severity level diagnosis [3]. Furthermore, with the continuous increase of the sensing data size in today's manufacturing industry, it is imperative to have an automated system to facilitate data processing [4-9].

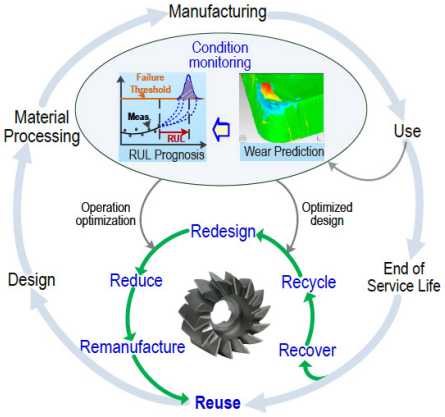

Fig. 1 TES as the concept of service support within the product lifecycle [1]

In this paper, a new integrated method based on wavelet transform and visual words representation, an image processing technique for image spatial pattern analysis, is proposed as an attempt to extend the traditional feature extraction method and represent a potential path to effective, automated time-frequency analysis for fault diagnosis, as images share common properties with manufacturing data such as high dimensionality and high variation [7-9].

The contribution of this research is to extend the capability of detecting and representing local patterns in time-frequency spectra, which are different among various machine health status, from the traditional manual procedure to two automated steps: point of interest detection and pattern description. Analogous to finding the peaks in frequency spectrum, the detection stage involves searching for local extremum in pixel intensity as the point of interest, representing signal characteristics in the time-frequency domain. The pattern description stage involves building a descriptor for each point of interest to describe the distribution of the pixel intensity within its neighbourhood, enabling the capture of a more complete morphological pattern of the signal in the time-frequency domain. Finally, k-means is applied as "feature fusion" to produce visual words clusters as sparse "visual words" feature vector. Visual words representation is a shift, scale and rotation invariant technique, enabling robust point of interest detection and feature comparison [11-13]. By integrating wavelet transform with visual words representation, the advantages of wavelet transform in capturing fault related patterns in the time-frequency domain, given its multi-resolution decomposition capability [14-16], and visual words representation of the converted 2D images are synergistically integrated. The generated visual words feature vectors are subsequently fed into a sparse representation based classifier (SRC) for simultaneous fault type and severity classification due to its simplicity and reliability [17-19].

The remainder of the paper is organized as follows. The theoretical framework of the continuous wavelet transform, the visual words representation and SRC are first introduced in Section 2. Section 3 investigates the effectiveness of the proposed method for bearing fault diagnosis. Conclusions are provided in Section 4.

2. Theoretical Framework

2.1 Continuous Wavelet Transform

Analysis in time-domain or frequency-domain alone is generally not suitable for complex fault diagnosis, due to the difficulties in time series modelling and the inability to characterize the non-stationarity in signals. Time-frequency analysis has been subsequently developed to represent the signal in both time and frequency domain to better reveal the data pattern related to machine status, fault types, and severity [20]. However, the traditional method such as the Short Time Fourier Transform (STFT) cannot achieve high resolution in both domains, which is often necessary for fault-related pattern recognition. This can be achieved by using wavelet transform which converts a signal into a series of wavelet coefficients in time-scale domain. Mathematically, the continuous wavelet transform (CWT) is performed through the convolution between the signal and the complex conjugate of a family of wavelet function [21]:

$$cwt(s,\tau)=\frac{1}{\sqrt{s}}\int x(t)\psi^{*}(\frac{t-\tau}{s})dt \tag{1}$$

where $x(t)$ is the original signal in the time domain, s denotes the parameter of scale and τ is the parameter of translation. $\psi^{*}$is the complex conjugate with scale and translation of the wavelet function ψ. The spectra of the wavelet transform will be used as time-frequency image for visual words representation.

2.2 Visual Words Representation

Image processing in computer vision is often linked to analyzing image local patterns, with the goal of fine-tuning the pattern description such that they're as distinctive as possible while being robust to noise and transformations. However, the goal of being distinct and robust is often constrained by each other [8,9]. Visual words representation aims to strike a balance between the two goals by first detecting sparse, salient points of interest from the image. Detection process is also conducted in different scales to achieve scale-invariance. Then the descriptor is built to describe the spatial distribution of pixel intensity around point of interest, which cannot be captured using common statistical features. Finally, descriptors are fused to produce "visual words" feature vector for each image [11-13].

2.2.1 Point of Interest Detection

The detection of point of interest in an image can be understood through the analogy to the detection of the local extremum of a function. For a function of two variables $f(x,y)$ that has continuous second partial derivatives, from the results of the second derivative condition, the determinant of Hessian matrix is sufficient to find local extrema. For an image I with pixel intensity $I(x,y)$, the point of interest is often associated with the local extrema of pixel intensity. Thus, the Hessian-based detector is proposed for its detection [11], at pixel (x,y) and scale σ:

$$H(x,y,\sigma)=\begin{bmatrix} L_{xx}(x,y,\sigma) & L_{xy}(x,y,\sigma) \\ L_{xy}(x,y,\sigma) & L_{yy}(x,y,\sigma) \end{bmatrix} \tag{2}$$

where four $L(x,y,\sigma)$ are the second order partial derivatives of the image smoothed by a Gaussian kernel to reduce its sensitivity to noise. It can be calculated through the convolution of the corresponding Gaussian second order derivative filter (Fig. 2) and I at pixel (x,y), in the domain of the filter G:

$$L(x,y,\sigma)=(I*g)(x,y,\sigma)=\sum_{m,n\in G} I(x-m,y-n)g(m,n,\sigma) \tag{3}$$

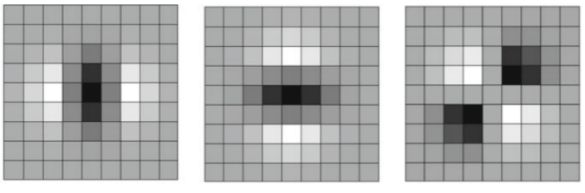

Fig. 2. Gaussian second order partial derivative filter for calculating L_{xx}, L_{yy} and L_{xy} [13].

It can be seen from Eq. (3) that the convolution can be considered as the sum of weighted image pixels, with the weight being the corresponding pixel in the filter. The computational complexity of the convolution $O(q^2)$ increases drastically with filter size. As proposed in [13], the SURF (Speeded Up Robust Features) method can improve the computation speed by using box filter to approximate the Gaussian second order derivative filter and by using integral image of I. Box filter is defined as:

$$bf(m,n)=\begin{cases} 1,(m,n)\in B \\ 0,(m,n)\notin B \end{cases} \tag{4}$$

where B is the domain of box filter, and it is a "filter" of constant values. The integral image of I is defined as:

$$U(x,y)=\sum_{0\le m\le x}\sum_{0\le n\le y} I(m,n) \tag{5}$$

Box filter assigns equal weight to each corresponding pixel in image during the summation, allowing convolution to be computed as simple additions and subtractions of the integral image, which can be pre-computed. The computational complexity is only determined by the number of vertices of the "boxes" appeared in the filter, rather than the absolute size of the filter. For example, in the illustration in Fig. 3, the box filter only has 4 vertices, thus, the computation of the convolution only requires 4 additions or subtractions. The efficiency of using box filter will become more obvious in the scale analysis as the number of vertices remains the same after the scaling of the filter.

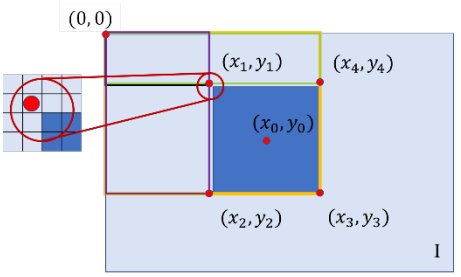

Fig. 3 Illustration of the convolution of image I (light blue) and box filter bf (dark blue) at (x_0, y_0). All the coordinates are pixel coordinates.

Using the same idea, a series of box filters can be obtained to approximate the corresponding Gaussian second order partial derivative filters, as shown in Fig. 4. Each entry in the Hessian matrix can then be written as simple additions and subtractions of the integral image, enabling simplification of the determinant of Hessian matrix.

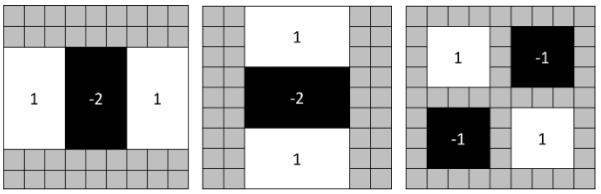

Fig. 4 Box filters to approximate the second order Gaussian partial derivative filters for calculating L_{xx}, L_{yy} and L_{xy}.

The scale analysis is conducted by applying a series of scaled box filters to the image, creating the scale-space [22-25]. The scale-space is divided into octaves. Each octave corresponds to a scaling factor of 2 and is further divided into levels. Thus, a series of points of interest can be extracted through the convolution between the original image and the filter at different scales. The points of interest are determined as local maximum of the determinant of Hessian matrix through a predefined threshold and a 3 x 3 x 3 local neighbourhood verification [13]. The point of interest is represented as triplet (x, y, σ) with both (x, y) coordinates information and the scale information σ.

2.2.2 Descriptor for Point of Interest

To investigate the similarity among points of interest and find the patterns that correspond to each health status, a descriptor is required to describe the pixel intensity distribution in the neighbourhood. The first step is to determine

the dominant orientation to allow rotation invariant comparison. This is done by first calculating the gradient vector $\boldsymbol{\phi}_p(x,y) \in R^2$ at all points (x,y), confined within a circle of radius $6\sigma_p$ centered at point of interest (x_p, y_p). All $\boldsymbol{\phi}_p(x,y)$ are weighted by the distance between the point (x,y) and the center (x_p, y_p) as illustrated in Fig. 5-a).

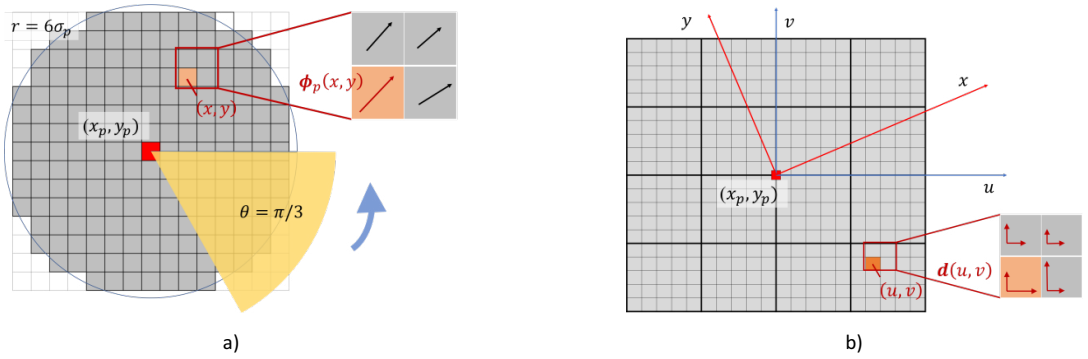

Fig. 5 a) Illustration of calculating dominant orientation ; b) Illustration of calculating descriptor of (x_p, y_p).

All $\boldsymbol{\phi}_p(x,y)$ are then used to calculate the dominant orientation by "marching" a window of $\frac{\pi}{3}$, originated from the point of interest (x_p, y_p), at constant step over 2π as illustrated in Fig. 5-a). In every step, each $\boldsymbol{\phi}_p(x,y)$ is checked to verify whether its direction is within the window. For each marching window, a "cumulated gradient vector" $\boldsymbol{\Phi}_p(\theta)$ is obtained by the sum of all $\boldsymbol{\phi}_p(x,y)$ that fall in the window. The dominant orientation $\theta_p$associated with (x_p, y_p) is the orientation of vector $\boldsymbol{\Phi}_p(\theta)$ that has the largest l_2 norm among all $\boldsymbol{\Phi}_p(\theta)$. To generate the descriptor for each point of interest, a square region of size $20\sigma_p$, centered at (x_p, y_p) and oriented at $\theta_p$is used. The rotation coordinates are (u,v). The square region is split into 16 blocks with each size $5\sigma_p$ and each block is further split into 25 points as shown in Fig. 5-b). The weighted gradient of each point is calculated as vector $\boldsymbol{d}(u,v) \in R^2$. Within each block i, $i = 1, 2 \ldots, 16$, $\sum \boldsymbol{d}(u,v)$ and $\sum|\boldsymbol{d}(u,v)|$ are concatenated to form a vector $\boldsymbol{b}_i \in R^4$. The final descriptor $\boldsymbol{b} \in R^{64}$ is obtained by first concatenating all $\boldsymbol{b}_i$, followed by normalization.

2.2.3 Cluster Descriptors for Visual Words

With the descriptors determined from all images, k-means clustering, based on the similarity between different descriptors, is used to cluster descriptors into k visual words. Then an image can be represented by a normalized histogram vector. The histogram represents the normalized occurrence of each of the k visual words in the image. This procedure can be considered as "feature fusion", which increase the discriminative power by fusing the large number of descriptors into visual words space to produce final feature vectors.

2.3 Sparse Representation Classifier

Sparse Representation Classifier (SRC) is a pattern classification technique based on sparse representation of the input data. Its simplicity in setup and high accuracy in classification has been reported in literature recently [17-19]. To classify C classes with c_j training samples from the jth class, $j = 1, 2 \ldots, C$, a dictionary $A \in R^{dxh}$ is defined as $[\boldsymbol{a}_1\ \boldsymbol{a}_2 \ldots\ \boldsymbol{a}_c]$, where each $\boldsymbol{a} \in R^d$ represents one training sample. $c = \sum_{j=1}^{C} c_j$ is the total number of training samples. Each column in A corresponds to a training sample from one of the C classes. For an input test data $\boldsymbol{z}$, sparse representation first solves the sparse vector $\boldsymbol{w} \in R^h$ such that $A\boldsymbol{w} = \boldsymbol{z}$ and $\boldsymbol{w}$ has the minimum non-zero entries. Greedy pursuit algorithm is used in this study [26]. In the classification stage, the post partitioning of the sparse vector $\boldsymbol{w}$ is utilized. $\delta_j(\boldsymbol{w})$ is defined as a vector created by setting all entries of $\boldsymbol{w}$ into zero except those corresponding to the jth class. Input test data $\boldsymbol{z}$ can be approximated as $\boldsymbol{z_j} = A\delta_j(w)$ for all j. The final classification is obtained by finding the class that leads to the smallest reconstruction error of $\boldsymbol{z}$ [17].

3. Experimental Evaluation

Bearings are one of the most critical parts in rotating machines and their failure often leads to immediate breakdown of the machine itself. Decades of research dedicated to the bearing diagnosis have produced enormous amount of literature, yet, due to the limitation in empirical and analytical knowledge, in a recently published

benchmark study, it is shown that the benchmark diagnosis method still struggles to effectively identify certain fault type, resulting in non-diagnosable outcomes [27]. While these obstacles continue to encourage the researchers to look for a deeper understanding of the failure mechanism, it underscores the necessity to address the gap and solve the problem from a new perspective. By integrating the continuous wavelet transform, visual word representation with sparse representation classifier, a complete bearing diagnosis framework is illustrated in Fig. 6-a).

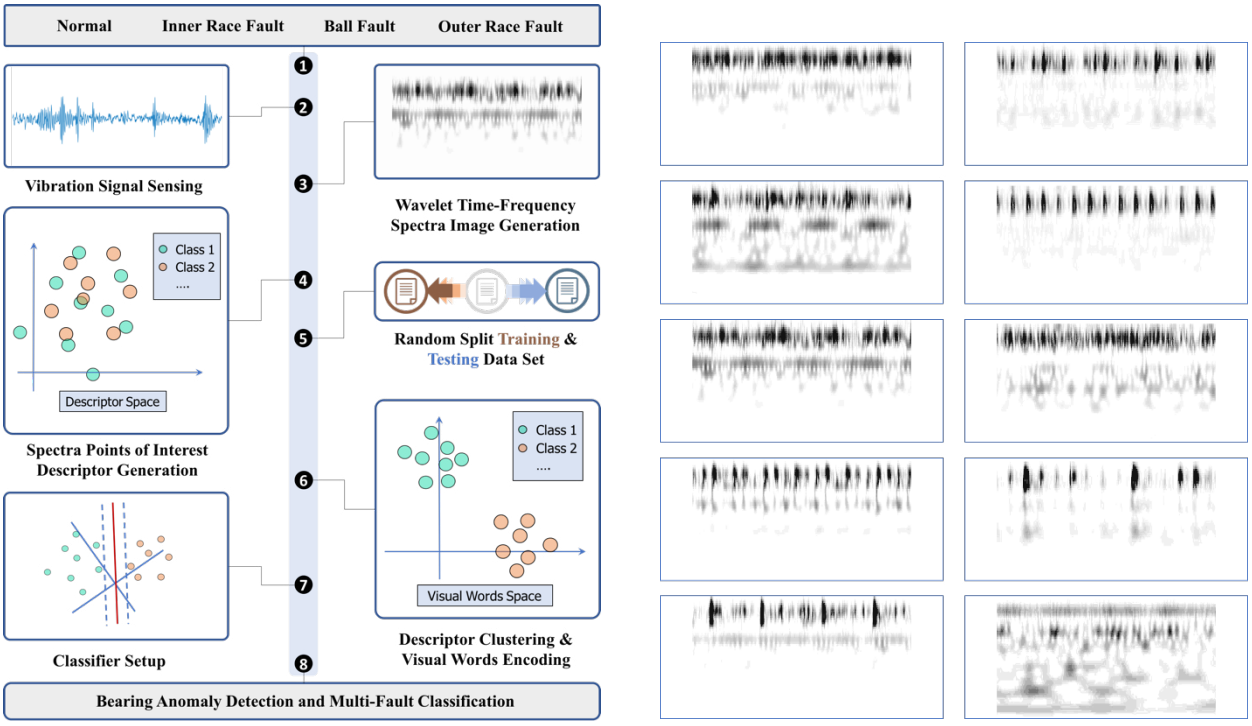

Fig. 6-a). Complete framework for bearing fault type and severity level classification based on wavelet transform and visual words representation; b) Wavelet spectra generated from all fault classes with normal condition. Health status, left column from top to bottom: ball 7 mils, 14 mils, 21 mils, inner race 7 mils, 14 mils. Right column from top to bottom: Inner race 21 mils, outer race 7 mils, 14 mils, 21 mils, normal condition.

Effectiveness of the proposed method has been evaluated using the bearing data provided by CWRU bearing center [28]. Single point faults were introduced to the drive-end bearing at inner race, ball and outer race, each with three fault diameters of 7, 14 and 21 mils. The vibration was measured at a sampling rate of 12 kHz. In total, 3 fault severity levels with 3 fault types were evaluated together with vibration data from the normal bearing condition.

50 samples were collected from each bearing health status and the spectra of the wavelet transform were generated for each sample. Fig. 6-b) is the matrix showing the spectra from each fault class and the normal condition. These patterns pose considerable difficulties for manual fault diagnosis, not only due to the similarity between fault types, but also within each fault type. For each health status, the 50 samples were randomly split into training data of 20 samples and testing data of 30 samples. Visual words were first generated from 200 training samples, then they were used to represent all 500 samples for the feature vector generation. 10 random tests were conducted at k=50, 100, 200, 300, where k is the number of visual words. The threshold is set at 200, 300, 500 and 1000, for point of interest screening. Classification results from the SVM are also reported as comparison as tabulated in Table 1.

The performance of the integrated method shows that, the classification accuracy improves with small threshold and large k. When the number of visual words is larger than 100, the classification results are fairly consistent. With the threshold dropped to 200, SRC and SVM reach mean classification rate of 99.7% and 99.2% respectively at certain k-value, meaning on average, no more than 2 test samples are misclassified out of 300. The threshold filters out the small extremum in the spectra and the k value fuses similar pattern descriptors into individual feature. The results indicate that some small extremum and difference among pattern descriptors contribute to a better separation of the

bearing health status. The consistency in the results suggests a good parameter range at around k=200, threshold=200.

Table 1. Classification accuracy based on SURF detector

	k	Threshold			
		200	300	500	1000
SRC	50	97.5%	96.3%	91.7%	90.1%
	100	99.3%	98.3%	94.6%	95.7%
	200	99.7%	98.3%	96.8%	96.9%
	300	99.4%	98.5%	97.4%	96.6%
	k	**Threshold**			
		200	300	500	1000
SVM	50	94.2%	91.8%	88.2%	87.5%
	100	98.2%	95.5%	91.3%	93.3%
	200	98.5%	97.0%	94.7%	94.1%
	300	99.2%	97.1%	95.7%	93.7%

As a comparison of the effectiveness of feature extraction technique, traditional statistical feature selection was carried out per [29]. The experimental results show that the proposed method based on wavelet transform and visual words representation not only outperforms the statistical feature-based method, but also is computationally more efficient, with training time of only 21.23s compared to 23.44s. A literature survey also verifies that the proposed method achieves similarly high classification accuracy compared with the more sophisticated and complex deep neural network approach while holding lead over PCA and LDA techniques, as shown in Table 2. The training time in comparison with DCNN is 21.23s versus 30 mins [31].

Table 2. Classification accuracy comparison

Classifier	Accuracy	Feature Extraction Technique
KNN	86.8%	PCA [30]
BPNN	89.3%	BPNN [30]
SVM	94.3%	Statistical feature in time and frequency
KNN	98.4%	LDA [30]
DCNN	99.4%	Convolutional Neural Network [30]
DNN	99.7%	Deep Neural Network [31]
SRC	99.7%	Visual words

4. Conclusion

In this paper, an image processing technique, visual words, is explored to process and analyse vibration spectra images, for automatic identification and classification of machine condition. The “words”, visually represented by the distribution of brightness and similarities between adjacent pixels, reflects the time-frequency patterns associated with each machine condition. These “words” and patterns are automatically extracted and detected in the view of machines, to broaden the scope of traditional spectral analysis based on human eyes. Subsequently, the extracted “words” are classified by sparse representation, which has high computational efficiency and does not require prior knowledge. The effectiveness of the proposed technique is further confirmed through evaluation of experimental bearing data. The experimental results indicate that the developed method outperforms most machine learning techniques. Furthermore, it has equivalent performance as deep learning techniques while demonstrating considerable saving in training-time. Finally, it is found that this classification can reduce the overfitting problem that is with other machine learning techniques, for more reliable performance.

Acknowledgements

This research has been supported by DMDII-15-14-01.

References

[1] Gao RX, Wang P. Through Life Analysis for Machine Tools: From Design to Remanufacture. Procedia CIRP. 2017;59:2-7.

[2] Gao Z, Cecati C, Ding SX. A survey of fault diagnosis and fault-tolerant techniques—Part I: Fault diagnosis with model-based and signal-based approaches. IEEE Transactions on Industrial Electronics. 2015;62(6):3757-67.

[3] Gunal S, Gerek ON, Ece DG, Edizkan R. The search for optimal feature set in power quality event classification. Expert Systems with

Applications. 2009;36(7):10266-73.

[4] Geramifard O, Xu JX, Pang CK, Zhou JH, Li X. Data-driven approaches in health condition monitoring—a comparative study. In Control and Automation (ICCA), 2010 8th IEEE International Conference on 2010 (pp. 1618-1622).

[5] Wang J, Liu S, Gao RX, Yan R. Current envelope analysis for defect identification and diagnosis in induction motors. Journal of Manufacturing Systems. 2012;31(4):380-7.

[6] Widodo A, Yang BS. Support vector machine in machine condition monitoring and fault diagnosis. Mechanical systems and signal processing. 2007;21(6):2560-74.

[7] Hinton GE. To recognize shapes, first learn to generate images. Progress in brain research. 2007;165:535-47.

[8] Wright J, Yang AY, Ganesh A, Sastry SS, Ma Y. Robust face recognition via sparse representation. IEEE transactions on pattern analysis and machine intelligence. 2009;31(2):210-27.

[9] Wang P, Yan R, Gao RX. Virtualization and deep recognition for system fault classification. Journal of Manufacturing Systems. 2017.

[10] Csurka G, Dance C, Fan L, Willamowski J, Bray C. Visual categorization with bags of keypoints. In Workshop on statistical learning in computer vision, ECCV 2004 (Vol. 1, No. 1-22, pp. 1-2).

[11] Lindeberg T. Feature detection with automatic scale selection. International journal of computer vision. 1998;30(2):79-116.

[12] Lowe DG. Distinctive image features from scale-invariant keypoints. International journal of computer vision. 2004;60(2):91-110.

[13] Bay H, Tuytelaars T, Van Gool L. Surf: Speeded up robust features. Computer vision–ECCV 2006. 2006:404-17.

[14] Luo GY, Osypiw D, Irle M. On-line vibration analysis with fast continuous wavelet algorithm for condition monitoring of bearing. Modal Analysis. 2003;9(8):931-47.

[15] Rubini R, Meneghetti U. Application of the envelope and wavelet transform analyses for the diagnosis of incipient faults in ball bearings. Mechanical systems and signal processing. 2001;15(2):287-302.

[16] Yan R, Gao RX, Chen X. Wavelets for fault diagnosis of rotary machines: A review with applications. Signal processing. 2014;96:1-5.

[17] Wright J, Ma Y, Mairal J, Sapiro G, Huang TS, Yan S. Sparse representation for computer vision and pattern recognition. Proceedings of the IEEE. 2010;98(6):1031-44.

[18] Liu S, Gao RX, John D, Staudenmayer J, Freedson PS. Classification of physical activities based on sparse representation. In Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE 2012 (pp. 6200-6203). IEEE.

[19] Sun C, Wang P, Yan R, Gao RX. A sparse approach to fault severity classification for gearbox monitoring. In Information Fusion (FUSION), 2016 19th International Conference on 2016 (pp. 2303-2308).

[20] Cohen L. Time-frequency distributions-a review. Proceedings of the IEEE. 1989;77(7):941-81.

[21] Rioul O, Vetterli M. Wavelets and signal processing. IEEE signal processing magazine. 1991;8(4):14-38.

[22] Oyallon E, Rabin J. An Analysis of the SURF method. Image Processing On Line. 2015;5:176-218..

[23] Neubeck A, Van Gool L. Efficient non-maximum suppression. In Pattern Recognition, 2006. ICPR 2006. 18th International Conference on 2006 (Vol. 3, pp. 850-855).

[24] Witkin A. Scale-space filtering: A new approach to multi-scale description. In Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP'84. 1984 (Vol. 9, pp. 150-153).

[25] Evans C. Notes on the opensurf library. University of Bristol, Tech. Rep. CSTR-09-001, January. 2009.

[26] Pati YC, Rezaiifar R, Krishnaprasad PS. Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition. InSignals, Systems and Computers, 1993. 1993 Conference Record of The Twenty-Seventh Asilomar Conference on 1993 (pp. 40-44).

[27] Smith WA, Randall RB. Rolling element bearing diagnostics using the Case Western Reserve University data: A benchmark study. Mechanical Systems and Signal Processing. 2015;64:100-31.

[28] Case Western Reserve University Bearing Data Center, http://csegroups. case.edu/bearingdatacenter/home

[29] Shukla S, Yadav RN, Sharma J, Khare S. Analysis of statistical features for fault detection in ball bearing. In Computational Intelligence and Computing Research (ICCIC), 2015 IEEE International Conference on 2015 (pp. 1-7).

[30] Ding X, He Q. Energy-fluctuated multiscale feature learning with deep convnet for intelligent spindle bearing fault diagnosis. IEEE Transactions on Instrumentation and Measurement. 2017 Mar 17.

[31] Jia F, Lei Y, Lin J, Zhou X, Lu N. Deep neural networks: A promising tool for fault characteristic mining and intelligent diagnosis of rotating machinery with massive data. Mechanical Systems and Signal Processing. 2016;72:303-15.