Data-driven smart manufacturing: Tool wear monitoring with audio signals and machine learning

Zhixiong Li[a], Rui Liu[b], Dazhong Wu[c],*

[a] *Department of Mechanical and Aerospace Engineering, University of Central Florida, Orlando FL 32816, USA*
[b] *Department of Mechanical Engineering, Rochester Institute of Technology, Rochester, NY 14623, USA*
[c] *Department of Mechanical and Aerospace Engineering, Department of Industrial Engineering and Management Systems, University of Central Florida, Orlando FL 32816, USA*

ABSTRACT

Tool wear in machining could result in poor surface finish, excessive vibration and energy consumption. Monitoring tool wear in real-time is crucial to improve manufacturing productivity and quality. While numerous sensor-based tool wear monitoring techniques have been demonstrated in laboratory environments, few tool wear monitoring systems have been deployed in factories because it is not realistic to install some of the important sensors such as dynamometers on manufacturing machines. To address this issue, a novel audio signal processing approach is introduced. This technique does not require expensive sensors but audio sensors only. A blind source separation method is used to separate source signals from noise. An extended principal component analysis is used for dimensionality reduction. Real-time multi-channel audio signals are collected during a set of milling tests under varying cutting conditions. The experimental data are used to develop and validate a predictive model. Experimental results have shown that the predictive model is capable of classifying tool wear conditions with high accuracy.

1. Introduction

Cutting tools are used to remove material from workpieces by means of shear deformation in various machining processes such as turning, milling, and drilling. According to the U.S. Cutting Tool Institute, manufacturers in the U.S. consumed over $200 million worth of cutting tools during April 2018. Excessive tool wear could result in substantial decreases in dimensional accuracy, significant increases in energy consumption, and eventually total breakage of cutting tools due to excessive cutting forces and vibrations, intensive stresses and temperature, as well as massive fracture at cutting edges. Health monitoring and predictive analytics techniques are crucial to monitoring the health conditions of cutting tools as well as predicting tool wear [1–4].

Current tool wear measurement techniques can be classified into two categories [5]: direct and indirect methods. With the direct methods, tool wear is measured directly using optical or radioactive techniques [6]. The direct methods have two primary limitations: (1) it is not realistic to measure tool wear using imaging systems during machining operations in real-world factory environments; (2) measuring tool wear using imaging systems is very time consuming and expensive. With the indirect methods, tool wear is estimated based on

certain condition monitoring signals such as vibration, acoustic emission, cutting force, spindle power, and temperature signals [7]. Current tool wear indirect monitoring techniques fall into two categories: model-based and data-driven methods [8]. Model-based methods are generally effective for predicting tool wear in the cases where in-depth knowledge of wear mechanisms are available. However, in practice tool wear mechanisms vary depending on cutting conditions, which makes effective tool wear prediction very difficult. In addition, certain statistical distributions must be assumed for model-based methods to develop close-form analytical solutions. These assumptions usually cannot be justified. To complement model-based methods, data-driven methods predict tool wear using predictive models trained by machine learning or pattern recognition algorithms. In addition, existing sensor-based tool wear monitoring techniques have a primary limitation. While various sensor-based condition monitoring signals have been demonstrated to be effective in monitoring and predicting tool wear in laboratory environments, it is not realistic for manufacturers to retrofit CNC machines with expensive sensors such as dynamometers.

To address these issues, a novel signal processing technique is introduced to denoise audio sensor-based signals collected during milling operations. A new feature extraction method is introduced based on

time-frequency domain analysis and adaptive kernel principal component analysis (AKPCA) to process the denoised audio signals. The remainder of this paper is organized as follows: Section 2 presents related work on audio-based tool wear monitoring systems. Section 3 introduces the blind source separation, feature extraction and selection methods. Section 4 presents an experimental setup and the design of milling tests. Section 5 presents experimental results. Section 6 provides conclusions and future work.

2. Related work

2.1. Tool wear monitoring using audio signals

Weller et al. [9] developed an electronic-mechanical system that uses sonic signals to detect the degree of cutting edge wear in turning operations. Experimental results have shown that the system can detect tool wear when cutting AISI 1045 steel with cutting tools made from cemented carbide. Mannan et al. [10] developed a sound analysis-based technique to monitor the conditions of cutting tools. The proposed technique can process audio signals generated during machining processes as well as identify the correlation between tool wear and sound patterns. Experimental results have shown that cutting tools with sharp, semi-dull, and dull conditions can be detected using this technique. Delio et al. [11] introduced an approach that can detect chatter using audio signals. Experimental results have shown that the performance of the audio signal-based chatter detection approach is comparable with other methods based on dynamometers, accelerometers, and displacement probes. Salgado and Alonso [12] developed a tool condition monitoring system for online tool wear monitoring in turning processes. Singular spectrum analysis was used to extract features from feed motor current and sound signals collected during turning. A support vector machine-based method was used to estimate tool wear by analyzing the extracted features. Aliustaoglu et al. [13] developed a tool wear monitoring technique based on two-stage fuzzy logic and audio signals collected via microphones. A set of drilling experiments was conducted on a four-axis CNC machining center. Audio signals were collected through a microphone. Experimental results have shown that the two-stage fuzzy logic is capable of identifying tool wear conditions. Ubhayaratne et al. [14] developed an approach to tool wear monitoring in sheet metal stamping using an audio signal processing technique. A semi-blind signal extraction technique was used to preprocess and denoise the audio signals collected from stamping operations. Seemuang et al. [15] developed a tool condition monitoring system that predicts tool wear by measuring audio signals generated by machine spindles. A low-cost microphone was used to collect the audio signals in drilling. Kothuru et al. [16] developed a tool wear monitoring approach that can classify tool wear conditions using support vector machines (SVM). Audio signals collected during milling processes were transformed to features in the frequency domain. Experimental results have shown that the prediction accuracy is above 90%.

2.2. Noise reduction in tool wear monitoring

According to a literature review [17,18], the most popular signal processing techniques in machining include fast Fourier transform (FFT), Wavelet transform, Hilbert-Huang transform (HHT), principal component analysis (PCA), and independent component analysis (ICA). Fang et al. [19] investigated the effects of tool edge wear on cutting forces and vibrations in high-speed finish machining using wavelet transform analysis. A set of turning experiments was performed to collect vibration signals. Both FFT and wavelet transform were used to denoise the vibration signals. Experimental results have shown that wavelet transform is more effective than FFT. Cao et al. [20] developed a chatter detection method for end milling processes using wavelet package transform (WPT) and HHT. WPT was used to denoise the measured signals. The denoised signals were analyzed using HHT to

obtain time-frequency-energy distribution of the signals. Experimental results have shown that WPT and HHT are capable of detecting chatter effectively. However, FFT, HHT, and WPT are not able to process multiple channel signals. To address this issue, Zhu et al. [21] develop a tool wear monitoring system for micro-milling using ICA. ICA is capable of separating a multivariate signal into additive components by assuming that the subcomponents are non-Gaussian signals and these signals are statistically independent from each other. Experimental results have shown that ICA outperforms traditional denoising methods such as wavelets. Shi et al. [22] developed a sound signal processing technique for tool breakage detection in face milling processes based on empirical mode decomposition (EMD) and ICA. Experimental results have shown that the proposed method is capable of detecting tool breakage by denoising sound signals during face milling. One of the limitations of ICA is that it assumes that subcomponents are independent from each other, which may not hold true [23,24]. To address this issue, bounded component analysis (BCA) has recently been proposed to denoise multivariate signals [23–25].

While previous studies have attempted to develop audio signal-based tool wear monitoring techniques, little research has been reported on monitoring tool wear in milling using multivariate audio signals. In addition, no studies have been conducted to improve tool wear monitoring by denoising multivariate audio signals. To fill the research gap, a novel audio-based tool wear monitoring approach is introduced. This method is capable of denoise multivariate audio signals as well as classify tool wear conditions using data-driven predictive models.

3. Signal processing and machine learning

3.1. Computational framework

Fig. 1 illustrates a computational framework of the proposed audio signal processing approach for tool wear monitoring. First, the raw multi-channel audio signals are decomposed into multiple wavelet sub-
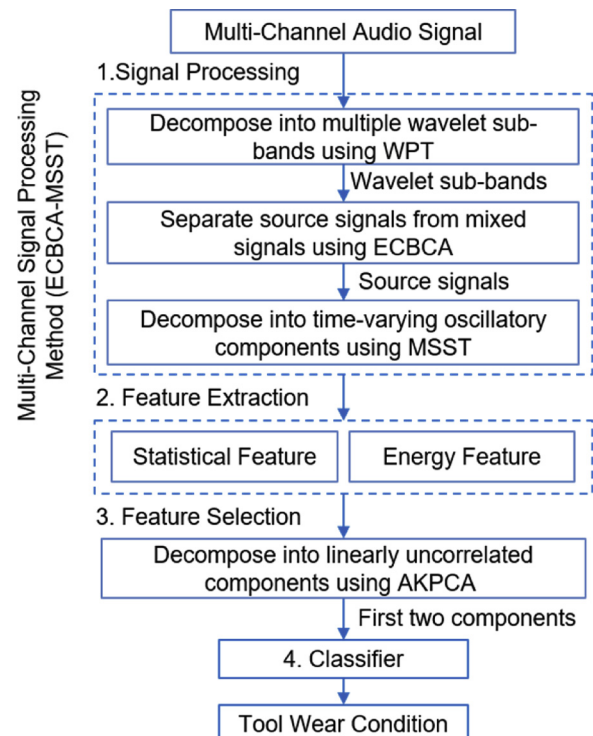

Fig. 1. A computational framework for the proposed audio signal processing approach.

bands using WPT. Second, an extended convolutive bounded component analysis (ECBCA) is used to separate source signals from the wavelet sub-band signals. Third, the separated source signals are decomposed into time-varying oscillatory components using a multivariate synchrosqueezing transform (MSST) in order to denoise. Fourth, eight (8) statistical features (i.e., max, min, mean, median, moment, skewness, kurtosis, and standard deviation) are extracted in the time-frequency domain. Another feature is the sum of frequency-amplitude in a certain frequency band that contains the significant frequency harmonics. Thus, nine features are extracted from the multi-channel audio signal. Fifth, an adaptive kernel principal component analysis (AKPCA) method is used to decompose the extracted features into a set of linearly uncorrelated components. The first two components out of the linearly uncorrelated components are fed into a classifier in order to classify tool wear conditions.

3.2. Extended convolutive bounded component analysis

As shown in Fig. 1, one of the important steps of the computational framework is blind source separation (BSS). BSS refers to the techniques that separate a set of source signals from a set of mixed signals. The most popular BSS method is ICA [23,24] that can separate independent sources. Bounded component analysis (BCA) is an extension of ICA. BCA is capable of separating both independent and dependent sources. The convolutive bounded component analysis (CBCA) was introduced to process signals that are space-time mixtures of the original sources [23,24]. In this paper, an ECBCA is used to separate source signals from audio signals mixed with various noises. The ECBCA is developed based on WPT [26] and CBCA. An audio signal is first decomposed by WPT into multiple wavelet sub-band signals. Then, CBCA is used to extract the sources from the wavelet sub-band signals. A mixed signal can be approximated by the following convolutive law:

$$\mathbf{x}(t) = \sum_{l=0}^{L-1} \mathbf{A}(l)\mathbf{s}(t-l) \tag{1}$$

Where, $\mathbf{x}(t) = [x_1, x_2,...,x_m]^\mathrm{T}$ are the m-channel sensor measurements, $\mathbf{s}(t) = [s_1, s_2,...,s_n]^\mathrm{T}$ are n unknown sources, and $\mathbf{A}(l) = [\mathbf{A}_1(l), \mathbf{A}_2(l),..., \mathbf{A}_n(l)]$ is an $m \times n$ mixing-filter matrix in the lth frequency bin, and the order of the filter is L-1. Eq. (1) describes the convolutive mixing processing in time domain. It can be rewritten in a matrix form as $\mathbf{x}(t) = \widetilde{\mathbf{A}}\widetilde{\mathbf{s}}_L$, where $\widetilde{\mathbf{A}} = [\mathbf{A}(0), \mathbf{A}(1),...,\mathbf{A}(L-1)]$ and $\widetilde{\mathbf{s}}_L = [\mathbf{s}(k), \mathbf{s}(k-1),...,\mathbf{s}(k-L+1)]$.

Convolutive BSS is an effective tool to separate source signals from their mixture by identifying a separator matrix $\mathbf{B}(l) = [\mathbf{B}_1(l), \mathbf{B}_2(l),..., \mathbf{B}_n(l)]$ with an order of M-1. The separated source signals $\mathbf{y}(t) = [y_1, y_2,...,y_n]^\mathrm{T}$ are calculated by

$$\mathbf{y}(t) = \sum_{l=0}^{M-1} \mathbf{B}(l)\mathbf{x}(t-l) \tag{2}$$

Eq. (2) can be rewritten as $\mathbf{y}(t) = \widetilde{\mathbf{B}}\widetilde{\mathbf{x}}_M$, where $\widetilde{\mathbf{B}} = [\mathbf{B}(0), \mathbf{B}(1),...,\mathbf{B}(M-1)]$ and $\widetilde{\mathbf{x}}_M = [\mathbf{x}(k), \mathbf{x}(k-1),...,\mathbf{x}(k-M+1)]$. Convolutive ICA has been used to determine the separator matrix. The CBCA has been used to relax the mutual-independent assumption made in convolutive ICA by iteratively updating the separator matrix [23, 24]. The optimized separator matrix corresponds to the maxima of the following objective function.

$$J(\widetilde{\mathbf{B}}) = \frac{\left(\sqrt{\det(Y_{\mathbf{y}_K})}\right)^{1/K}}{\|\rho(\mathbf{y}(t))\|_r^n} \tag{3}$$

Where, $\mathbf{y}_K = [\mathbf{y}_K(K), \mathbf{y}_K(K+1),...,\mathbf{y}_K(K-M+1)]^\mathrm{T}$ is a set of separated sources, $Y_{\mathbf{y}_K}$ is the covariance matrix of $\mathbf{y}_K$, $\rho(\mathbf{y}(t))$ is the range of $\mathbf{y}(t)$, and $\|\rho(\mathbf{y}(t))\|_r^n$ is a measure of geometrics of the overall separated sources ($r \geq 1$), such as the volume of principal hyper-ellipse [24].

3.3. Multivariate synchrosqueezing transform

The source audio signals separated by ECBCA from the mixed signals can be modeled as a sum of components with a varying amplitude and instantaneous frequency. Therefore, the retrieval of the components of a multicomponent signal is a critical issue in audio signal processing. Time-frequency analysis has been used to process non-stationary signals in both the time and frequency domains simultaneously. While wavelet transform and short-time Fourier transform (STFT) are the most popular time-frequency analysis methods, these methods are not effective in processing signals with oscillatory properties [27]. To address the limitation of wavelet transform and STFT, the multivariate synchrosqueezing transform (MSST) is used to characterize multivariate signals with time-varying oscillatory properties [27,28]. The MSST method is briefly introduced as follows:

For a signal $x(t)$, its wavelet-based synchrosqueezing is expressed as

$$S_w(\omega, b) = \int C(a, b)a^{-\frac{3}{2}}\delta(\omega_x(a, b) - \omega)da \tag{4}$$

Where, $S_w(\omega, b)$ are the synchrosqueezing transform (SST) coefficients, $C(a, b)$ are the wavelet coefficients, a is a scale factor, b is a shift factor, ω is the frequency, and ω_x is the instantaneous frequency.

The SFTF-based synchrosqueezing can be described by

$$S_s(\omega, \tau) = \frac{1}{\omega(0)}\int S_T(\tau, \eta)\delta(\omega - \omega_x(\tau, \eta))d\eta \tag{5}$$

Where, $S_s(\omega, \tau)$ are the SST coefficients, $S_T(\tau, \eta)$ is the STFT component, η is the frequency index, and τ is the time. In order to extend SST, a two-fold method was introduced [27]. A multivariate time-frequency partitioning algorithm is used to partition the SST coefficients of the multi-channel data into multiple frequency bands. Then, a frequency splitting process is conducted to separate the multivariate mono-component signals. The MSST coefficients are determined by summing up the multivariate instantaneous amplitude within each separated frequency band.

3.4. Adaptive kernel principal component analysis

After characterizing the audio signals with time-varying oscillatory components with MSST, multiple features, including eight statistical features extracted in the time-frequency domain and an energy feature, are extracted. The AKPCA is used to convert the nine features into a set of linearly uncorrelated components. AKPCA is an extension of kernel principal component analysis (KPCA). KPCA extracts principal components by mapping the input space into a feature space using a nonlinear mapping function [29]. KPCA is not effective in processing non-stationary signals. Therefore, the AKPCA [30] was introduced by adaptively updating the kernel covariance matrix in KPCA.

For an original feature space $\boldsymbol{f} = (f_1, f_2,...,f_n) \in P^{p \times n}$, with $f_1 \in P^p$ and p is the number of features. The PCA processing can be described by Eq. (6).

$$\mathbf{FF}^\mathrm{T}\mathbf{W} = \lambda\mathbf{W} \tag{6}$$

Where, $\mathbf{F}$ is the covariance matrix of $\boldsymbol{f}$, and λ and $\mathbf{W}$ are the eigenvalues and eigenvectors of $\mathbf{FF}^\mathrm{T}$, respectively. Then q PCs can be calculated using the q-largest eigenvalues.

$$\mathbf{p}_\mathbf{c}(q) = \mathbf{W}^q\mathbf{f} \tag{7}$$

Where, $\mathbf{p}_\mathbf{c}(q)$ is the q th PC, and $\mathbf{W}^q$ are the eigenvectors corresponding to the q-largest eigenvalues. In order to apply PCA to a nonlinear case, KPCA first projects $\boldsymbol{f}$ into a new space $\mathscr{H}$ via a mapping function ϕ, then centers the projected data in $\mathscr{H}$ and subsequently conducts the traditional PCA processing on the centered data. The new covariance matrix in $\mathscr{H}$ is expressed as

$$\mathbf{H} = \frac{1}{n}\sum_{i=1}^{n} \phi(f_i)\phi(f_i)^\mathrm{T} \tag{8}$$

Where, $\mathbf{H}$ is a kernel covariance matrix. The map ϕ adopts kernel to calculate the inner products in $\mathscr{H}$: $k(f_i, f_j) = \phi(f_i) \cdot \phi(f_j)$, where $k(\cdot, \cdot)$ is the kernel function (e.g., the radial basis function). Thus, Eq. (6) can be re-written as

$$K_n \mathbf{a^q} = n\lambda_q \mathbf{a^q} \qquad (9)$$

Where, K_n is the kernel gram matrix, λ_q and $\mathbf{a^q}$ are the first q largest eigenvalues and the corresponding normalized eigenvectors of K_n, respectively. Then, for a new test data f_{n+1}, the PCs can be calculated as $p_c^{n+1}(q) = \mathbf{a^q}\phi(f_{n+1})$. The kernel covariance matrix affects the effectiveness of PCs. AKPCA involves the following steps [30]:

(1) Construct and update the recursive formulation for KPCA by performing the rotation and augmentation operations.
(2) Estimate the kernel covariance matrix using a weighted sliding window.
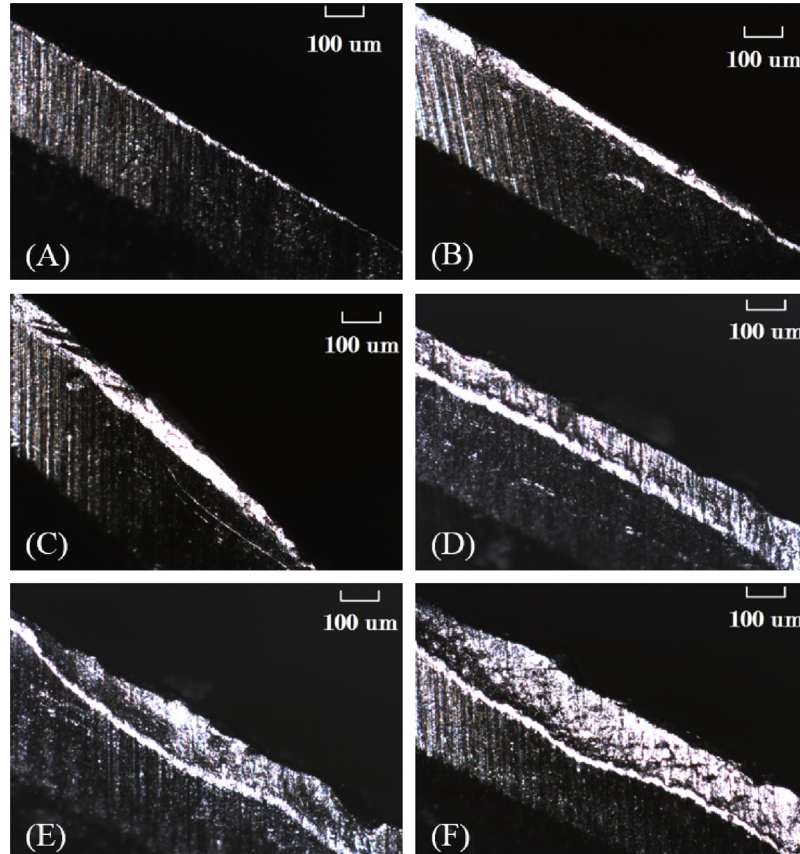(3) Perform KPCA using the estimated kernel covariance matrix.

Because the first two principal components usually contain up to 85% of the original information [30], the first two principal components are selected, and then fed into machine learning algorithms.

4. Experimental setup

To demonstrate the proposed method, a set of milling tests with varying cutting conditions was conducted. All of the milling tests were conducted on a TRACK K3 EMX mill. The workpiece material is 6061 aluminum. The flank wear of a set of end milling tools was characterized using an optical microscope (Olympus BX60 M) according to ISO 8688-1 and ISO 8688-2. As shown in Fig. 2, tool wear conditions were classified into six categories based on the average width of the flank

Table 1
Experimental design for milling tests.

Tool Condition	Spindle Speed (rpm)	Feed Rate (ipm)
Good	1400	14, 20, 30
	1800	9, 13, 17
	2000	10, 15, 17.5
	2300	12, 18, 21
	2400	15, 25, 40
Slight	1400	14, 20, 30
	1800	9, 13, 17
	2000	10, 15, 17.5
	2300	12, 18, 21
	2400	15, 25, 40
Average	1400	14, 20, 30
	1800	9, 13, 17
	2000	10, 15, 17.5
	2300	12, 18, 21
	2400	15, 25, 40
Heavy	1400	14, 20, 30
	1800	9, 13, 17
	2000	10, 15, 17.5
	2300	12, 18, 21
	2400	15, 25, 40
Severe	1400	14, 20, 30
	1800	9, 13, 17
	2000	10, 15, 17.5
	2300	12, 18, 21
	2400	15, 25, 40
Failure	1400	14, 20, 30
	1800	9, 13, 17
	2000	10, 15, 17.5
	2300	12, 18, 21
	2400	15, 25, 40

Fig. 2. Images of six tool wear classes: (a) Good: 0–20 μm in flank wear thickness, (b) Slight wear: 20–40 μm, (c) Average wear: 40–70 μm, (d) Heavy wear: 70–100 μm, (e) Severe wear: 100–150 μm, and (f) Failure: > 150 μm.
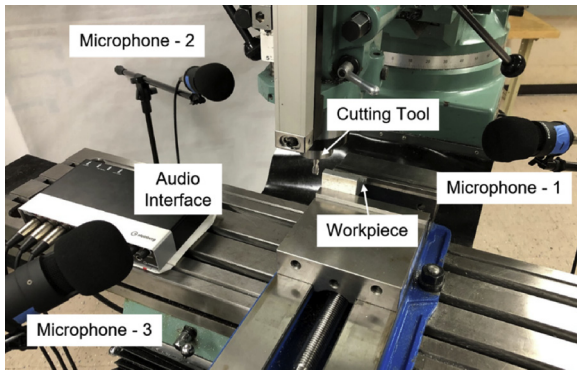
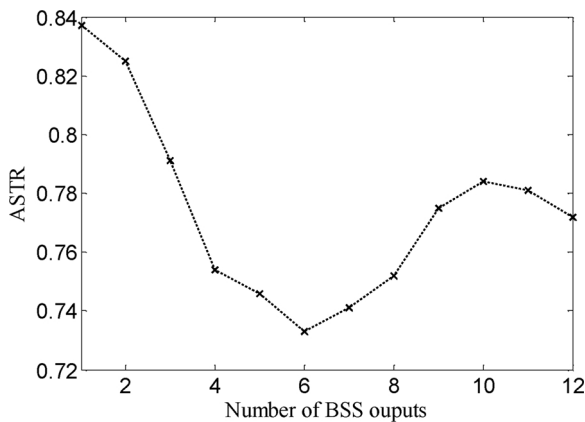**Fig. 3.** Experimental setup for the milling tests.

Fig. 4. The values of ASTR under the good tool wear condition at a rotation speed of 1800 rpm and a feed rate of 13 ipm.

wear land of two flutes. For each flute, the width of flank wear was determined by taking an average of ten measurements at different locations of the cutting edge.

Table 1 lists the details about the cutting conditions. Each tool wear category involves five different cutting speeds and three different feed rates. To collect audio signals during milling tests under varying cutting conditions, three condenser microphones (ECOOPRO EO-200) were used to collect real-time audio signals from different directions and distances as shown in Fig. 3. The sampling frequency of the sound collection was set to 44.1 kHz. More details about the experiments can be found in [16].

5. Results and discussions

5.1. Signal processing using ECBCA-MSST

The proposed ECBCA-MSST method was used to extract useful features in the time-frequency domain from the multi-channel audio signals. First, the audio signal collected from each microphone was decomposed into four (4) wavelet sub-bands using WPT. Therefore, twelve (12) sub-bands were generated from three microphones. The BSS model was then used for denoising. Because the number of the sources in the raw audio signals is unknown, an average signal to total ratio (ASTR) is used to determine the number of source signals.

$$\text{ASTR} = \frac{\sum E_l}{EL} \qquad (10)$$

Where, E_l is the energy of the lth (l = 1, 2, …, L. L is the number of output sources) separated source and E is the energy of original three-channel signals. A smaller ASTR indicates better denoising performance. Fig. 4 shows the ASTR values of different output numbers in the BSS separation process under the good condition (a spindle speed of 1800 rpm and a feed rate of 13 ipm). The smallest value of ASTR was obtained when the number of BSS outputs is six. Similar results were observed for other tool wear conditions. Therefore, the number of source signals is six.

Fig. 5 compares the original time-frequency spectrum generated by STFT against the time-frequency spectrum generated by ECBCA for the good tool condition. As shown in Fig. 5(a), six principal components in the time-frequency domain, including three components ([0 0.05], [0.35 0.4], and [0.75 0.76]) in the frequency domain and three components ([7.5 s 10 s], [11 s 12 s], and [12.5 s 14 s]) in the time domain, were decomposed from the raw signal. The three components in the frequency domain are due to three sources of vibration generated during the milling process. The three components in the time domain are due to three sources of sound waves generated during the milling process. Fig. 5(b) shows strong background noise in red. ECBCA removed some of the noise while retaining six principal components.

In order to extract useful information about tool wear in the time-frequency domain, the ECBCA output was used as the input of the MSST processor. Figs. 6–11 show the time-frequency characteristics of the six different tool wear conditions by using MSST. The SST analysis results were also presented for comparison purpose. The spindle speed was 1800 rpm (i.e., spindle angular frequency f_r = 30 Hz) for each condition. Theoretically, the milling dynamics model is analogical to the dynamics of a rotor rubbing model, where f_r and its harmonics such as 5 or 10 times harmonics (i.e. 150 Hz or 300 Hz) would be the dominating frequency components in the time-frequency domain [31]. Similarly, in the time-frequency plot of the tool cutting dynamics, dominant energy strips would appear at the harmonics of the spindle rotational
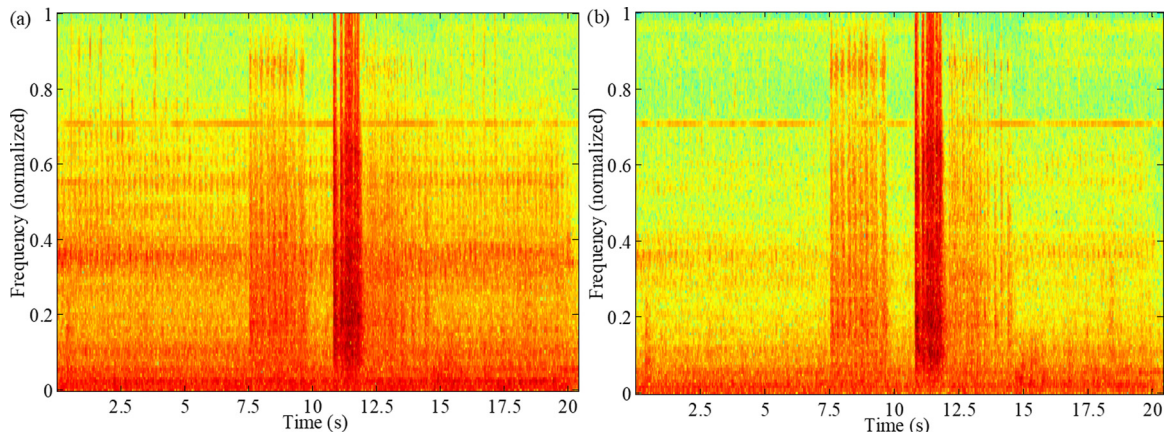

Fig. 5. Time-frequency spectrum generated by (a) STFT and (b) ECBCA.
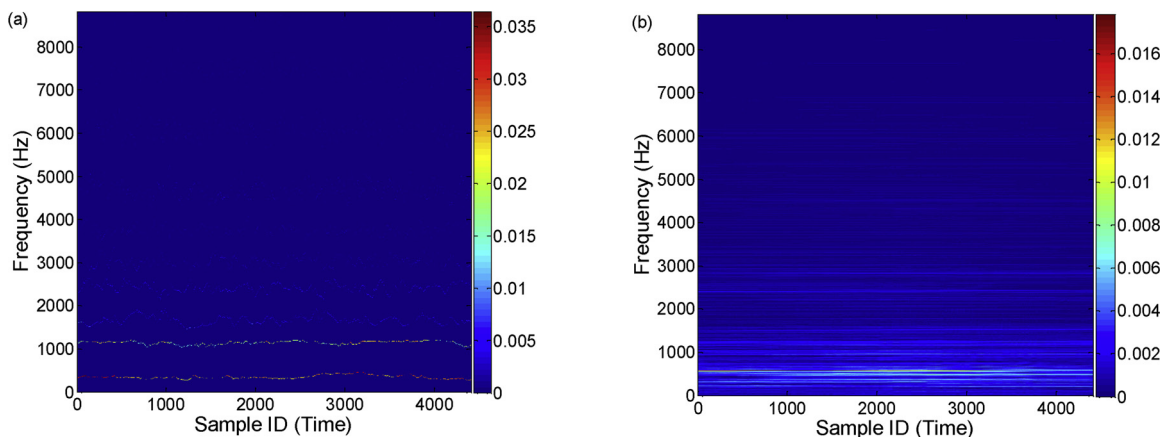
Fig. 6. Time-frequency representation under the good condition: (a) MSST and (b) SST.
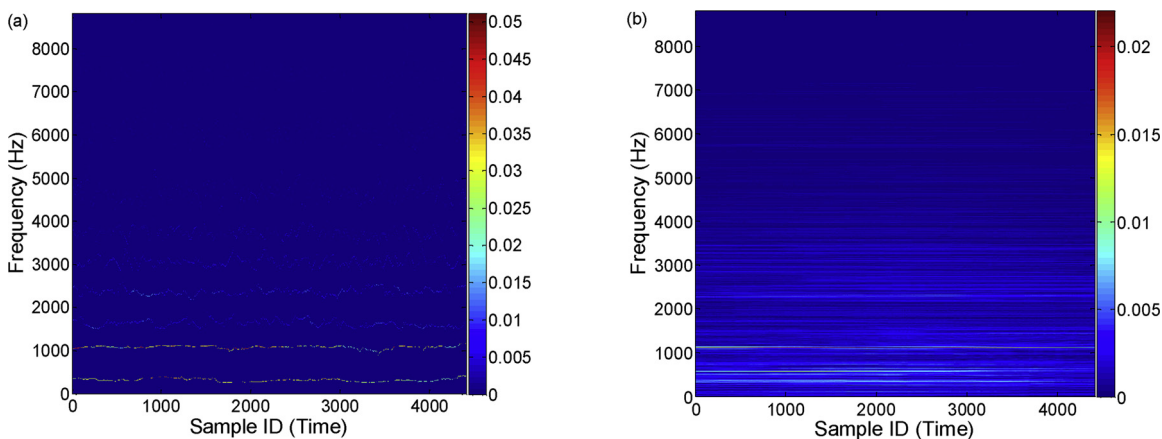

Fig. 7. Time-frequency representation under the slightly tool wear condition: (a) MSST and (b) SST.
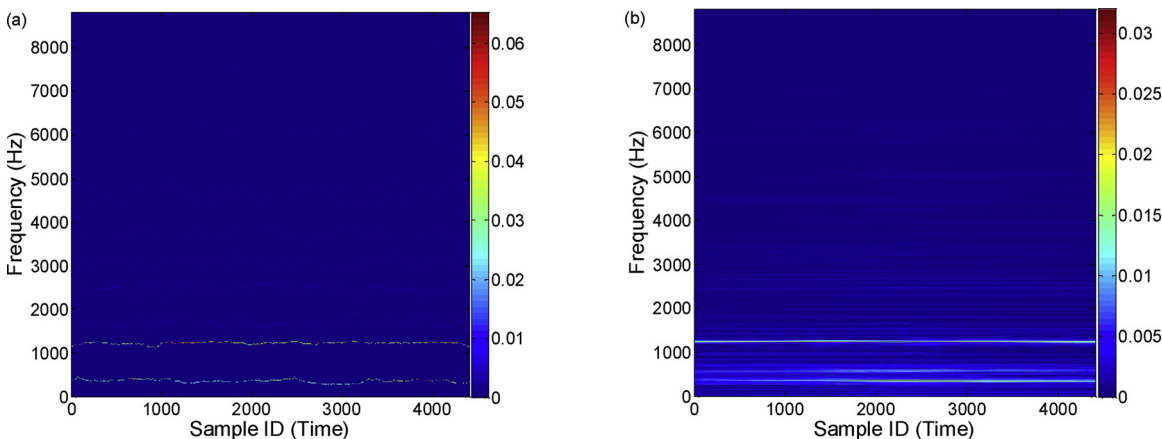

Fig. 8. Time-frequency representation under the average tool wear condition: (a) MSST and (b) SST.

frequency f_r.

During the milling tests, the spindle rotational frequency (f_r = 30 Hz) and the corresponding harmonic frequencies are dominant in the lower frequency band of the audio sound, which is consistent with the observations in previous researches [31–33]. Because the main exciting force in the milling operation is the contact between the cutting tool and workpiece, the dynamical responses of the milling machine with good and worn tools are similar. The health condition of the tool affects the energy distribution of the tool-workpiece dynamics. As shown in Figs. 6(a)–11 (a), the energy distribution at about 5, 10, 35 and 40 times harmonics of the spindle rotational frequency are

prominent in the time-frequency spectrum of the tool-workpiece dynamics under six different tool health conditions. It should be noted that the amplitude of the cutting dynamics varies with different tool health conditions and increases with the severity of the tool wear. The amplitude of good tool is the smallest while the severe tool wear and tool failure generate the largest amplitude. In addition, when comparing the Figs. 6(a)–11 (a) with Figs. 6(b)–11 (b), we observed that the background noise has been significantly removed by MSST. As a result, the signal processing results in Figs. 6–11 demonstrate the effectiveness of the proposed method. It should be noted that the energy magnitude of different tool wear states in Figs. 6(a)–11 (a) are very different.
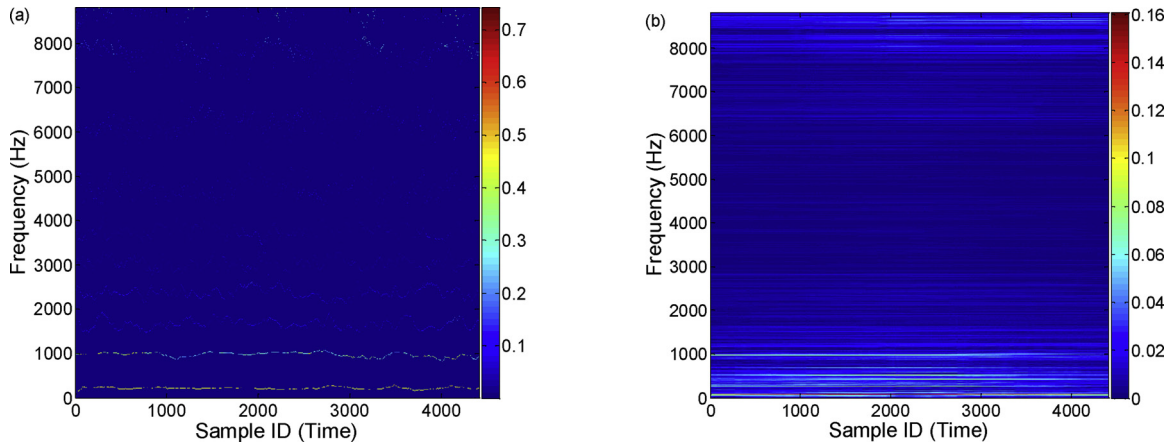
Fig. 9. Time-frequency representation under the heavy tool wear condition: (a) MSST and (b) SST.
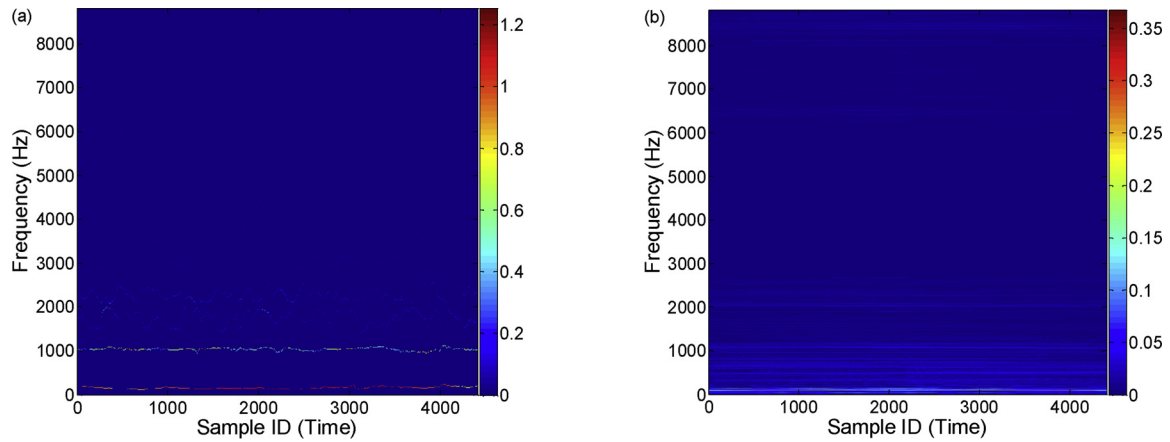

Fig. 10. Time-frequency representation under the severe tool wear condition: (a) MSST and (b) SST.
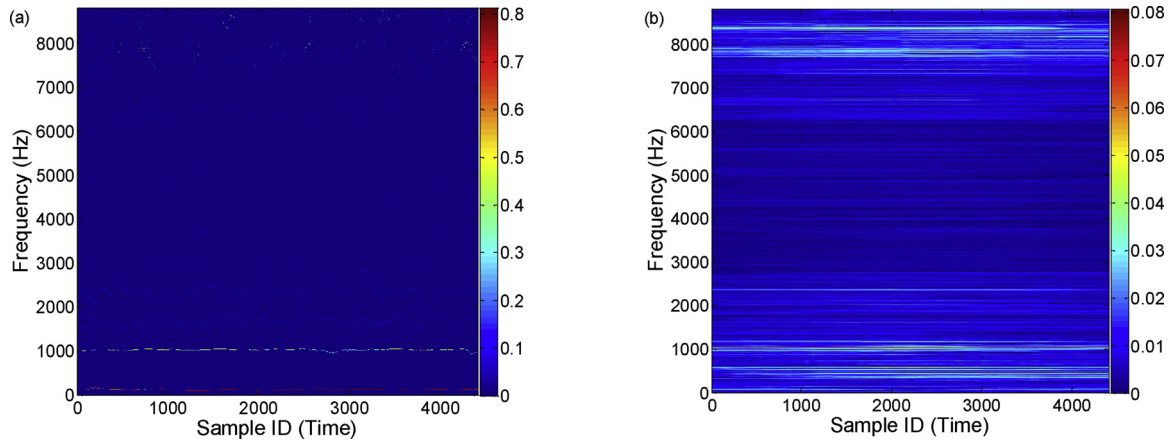

Fig. 11. Time-frequency representation under the failure condition: (a) MSST and (b) SST.

Therefore, the sum of energy amplitudes in the low frequency band (0–1400 Hz) was used as a feature for tool wear classification. Fig. 12 shows the zoomed figures of Figs. 6(a)–11 (a) in the range of 0–1400 Hz. The average values of energy sum features calculated from 15 samples of each tool wear condition are 0.4967, 2.2800, 8.0493, 1.3047, 3.0647, and 17.3893. The average value of energy sum in good tool condition is the smallest while the failure condition produces the largest average value.

5.2. Feature extraction

The audio signal was collected within a time interval of about 30 s during each cutting test. In order to generate sufficient training and testing datasets, the recorded sound wave in each cutting test was segmented into one-second sub-datasets. After the segmentation, there were 333, 351, 407, 439, 627, and 500 sub-datasets under good, wear, average wear, heavy wear, severe wear, and failure conditions, respectively. 2657 sub-datasets were generated.

Eight (8) features (i.e., max, min, mean, median, moment, skewness, kurtosis, and standard deviation) were extracted in the time-frequency
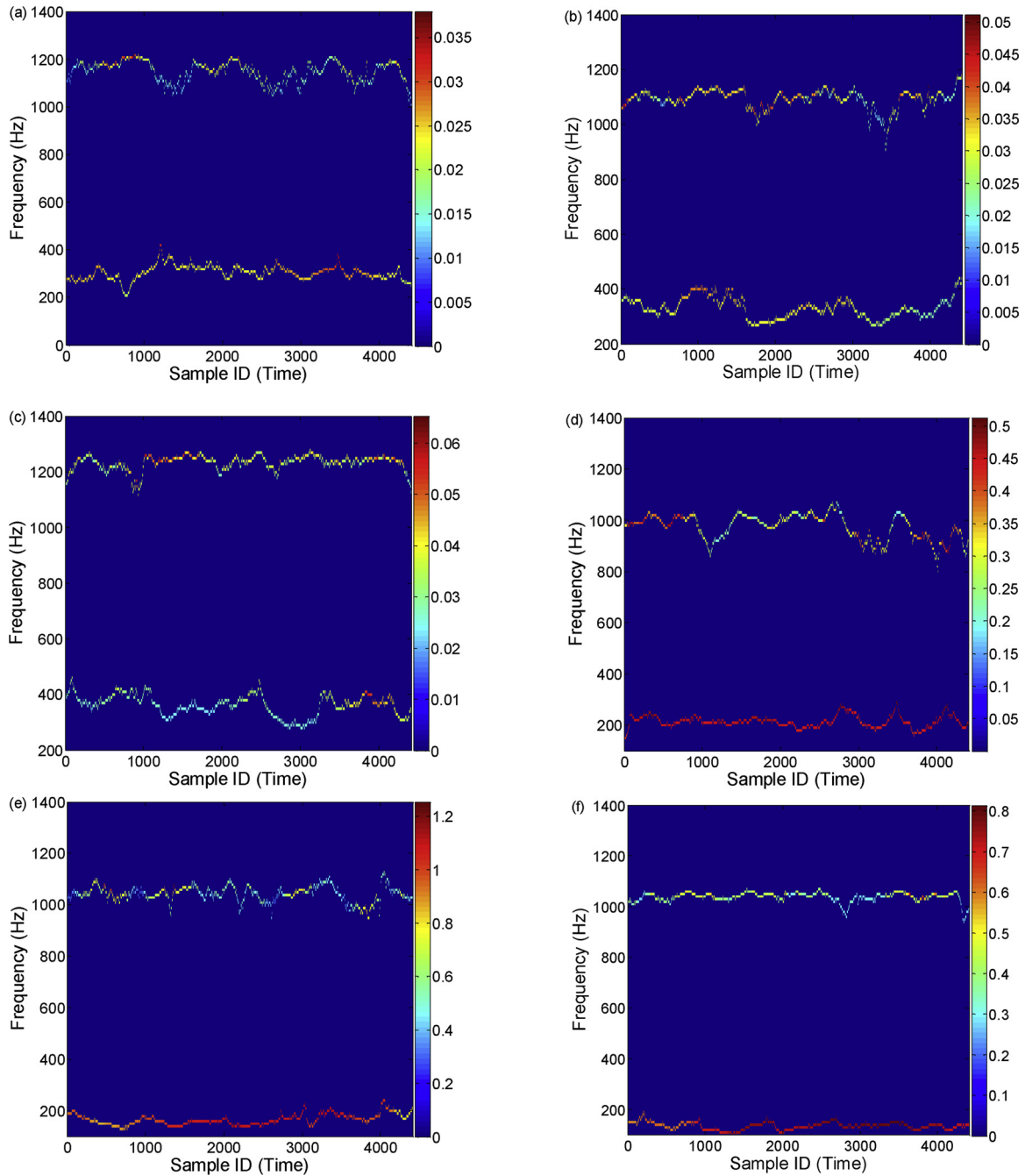
Fig. 12. Time-frequency spectrums under (a) good tool, (b) slight wear, (c) average wear, (d) heavy wear, (e) severe wear, (f) failure conditions.

domain for each sub-dataset using Eq. (11).

$$
\begin{cases}
f_i = g_i(\mathbf{e}) \\
e_j = \sum_{k=0}^{8700} \widetilde{E}_j^k
\end{cases}
\tag{11}
$$

Where, f_i ($i = 1, 2, …, 8$) denotes one of the eight (8) features and $g_i(\cdot)$ is the corresponding calculation function (i.e. the functions of max, min, mean, median, moment, skewness, kurtosis, and standard deviation), $\mathbf{e} = [e_1, e_2, …, e_j]$ ($j = 1, 2, …, N$ and N is the number of sample points of each sub-dataset) is the time series of the energy-amplitude sum at each sample point along the frequency axis, $\widetilde{E}_j^k$ denotes the energy-amplitude at j^{th} sample point and k ($\in$ R) Hz frequency and 8700 (Hz) is the maximum frequency value of each audio signal in the milling tests.

In addition, an energy feature was extracted by summing up the energy-amplitude in the range of 0–1400 Hz as shown in Fig. 12. Thus,

nine (9) features were calculated for each sub-dataset and the total is $\mathbf{F}_{9 \times 2657}$. AKPCA was then employed to fuse the original nine features into two new features. Thus, a new feature space $\mathbf{H}_{2 \times 2657}$ was obtained. Fig. 13 shows the AKPCA feature distribution of half of $\mathbf{H}$. The first AKPCA feature (red dots) can be used for classification of most sub-datasets even if there are overlaps between slight and average wear and between good and heavy wear conditions. The second AKPCA feature (blue circles), however, can provide additional information to distinguish the slight and average wear conditions. The fused features were fed into machine learning algorithms in order to classify tool wear conditions.

5.3. Classification of tool wear conditions

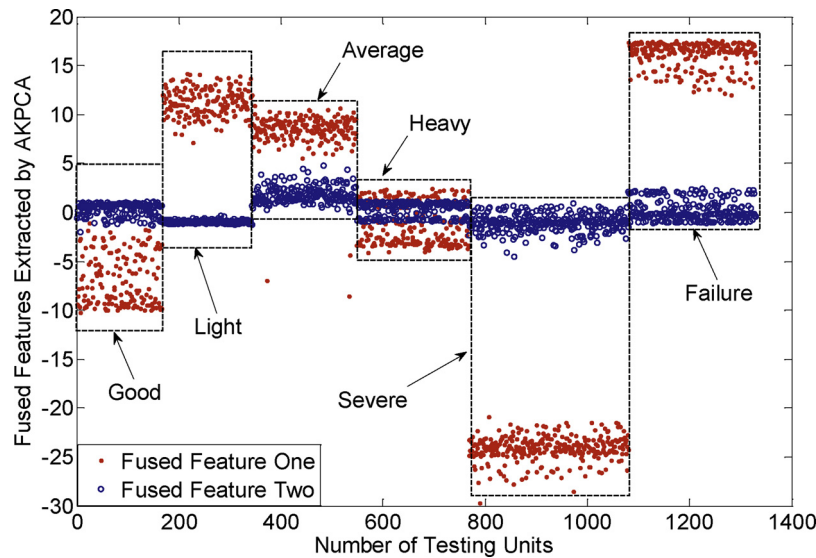Four classifiers, including CART, random forest (RF), kNN and SVM,

Fig. 13. The first two principal components decomposed by AKPCA.

Table 2
Evaluation of the proposed method on 50% dataset.

Condition	Classification Accuracy (%)							
	CART		RF		kNN		SVM	
	Without ECBCA	With ECBCA	Without ECBCA	With ECBCA	Without ECBCA	With ECBCA	Without ECBCA	With ECBCA
Good	96.39	98.19	84.34	86.14	94.58	95.78	98.19	100
Slight	92.57	94.29	85.71	88.00	96.00	97.14	97.14	98.29
Average	96.55	98.52	97.54	98.03	97.54	98.52	96.06	98.03
Heavy	95.43	98.17	87.67	89.50	98.63	99.54	96.80	98.63
Severe	98.73	99.36	99.04	99.36	98.09	98.66	96.24	98.66
Failure	92.00	93.20	98.40	99.36	94.40	97.20	96.00	97.20
Overall Accuracy	95.28	96.96	92.12	93.40	96.54	97.81	96.74	98.47

Table 3
Comparisons of SVM monitoring results with different signal processing methods.

	Results in [16]	SST	MSST	ECBCA-MSST
Good	99.4%	84.34%	98.19%	100%
Slight	97.5%	77.71%	97.14%	98.29%
Average	90.7%	61.08%	96.06%	98.03%
Heavy	92.4%	66.21%	96.80%	98.63%
Severe	98.5%	88.54%	96.24%	98.66%
Failure	97.0%	81.20%	96.00%	97.20%
Overall Accuracy	95.92%	76.51%	96.74%	98.47%

Table 4
Monitoring performance of CART using different training size.

	Size of the training dataset				
	50%	60%	70%	80%	90%
Good	98.19%	96.99%	97.00%	100%	93.94%
Slight	94.29%	95.00%	97.14%	94.29%	100%
Average	98.52%	100%	99.18%	100%	100%
Heavy	98.17%	98.86%	96.21%	94.32%	100%
Severe	99.36%	99.15%	100%	100%	100%
Failure	93.20%	96.50%	94.67%	99.00%	100%
Overall Accuracy	96.96%	97.75%	97.37%	97.93%	98.99%

Table 5
Monitoring performance of RF using different training size.

	Size of the training dataset				
	50%	60%	70%	80%	90%
Good	86.14%	82.17%	86.00%	94.03%	90.91%
Slight	88.00%	86.43%	89.52%	91.43%	94.29%
Average	98.03%	99.39%	98.36%	100%	100%
Heavy	89.50%	94.89%	97.73%	93.18%	97.73%
Severe	99.36%	99.15%	100%	100%	100%
Failure	99.36%	92.50%	92.67%	86.00%	96.00%
Overall Accuracy	93.40%	92.51%	94.05%	94.11%	96.49%

Table 6
Monitoring performance of kNN using different training size.

	Size of the training dataset				
	50%	60%	70%	80%	90%
Good	95.78%	96.99%	98.00%	98.51%	100%
Slight	97.14%	95.00%	99.05%	98.57%	97.14%
Average	98.52%	98.16%	100%	99.18%	98.50%
Heavy	99.54%	99.18%	99.24%	98.86%	99.33%
Severe	98.66%	100%	98.86%	98.15%	99.00%
Failure	97.20%	97.56%	100%	100%	96.00%
Overall Accuracy	97.81%	97.97%	98.94%	98.85%	98.45%

were used to identify tool wear conditions. Table 2 lists the tool wear monitoring results using half of the feature space **H** as the training dataset and the other half as test dataset. The training units were
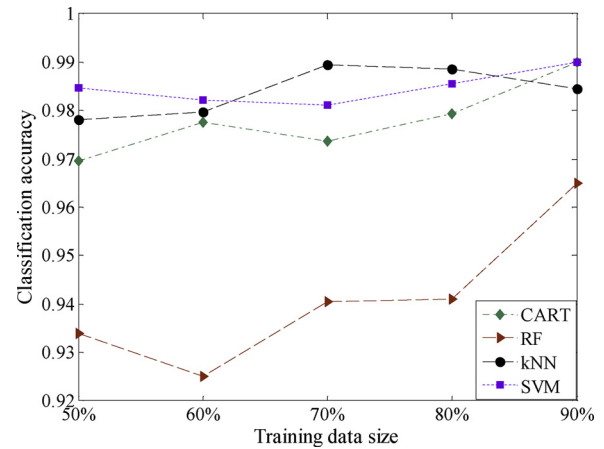
randomly selected for the classifiers and the training-test procedure repeated five times. In the training process of the classifiers, 500 trees

Table 7
Monitoring performance of SVM using different training size.

	Size of the training dataset				
	50%	60%	70%	80%	90%
Good	100%	97.74%	99.00%	97.01%	100%
Slight	98.29%	97.14%	95.24%	97.14%	100%
Average	98.03%	98.77%	100%	100%	100%
Heavy	98.63%	100%	97.73%	100%	100%
Severe	98.66%	99.18%	100%	98.15%	100%
Failure	97.20%	96.50%	96.67%	99.00%	94.00%
Overall Accuracy	98.47%	98.22%	98.11%	98.55%	99.00%

were used for RF, *k* was chosen as 15 for kNN, and the radial basis function with a width parameter 0.5 was used for SVM. The average classification accuracy for five tests is listed in Table 2. For the ECBCA denoising performance, the classification accuracy for each tool wear condition was improved. In addition, the overall classification accuracy for the six tool conditions using ECBCA-based methods is higher than that without ECBCA, and the best result is obtained by the SVM classifier, 98.47% overall.

In a previous study [16], SVM classifier was employed to identify the six tool wear conditions with an overall accuracy of 95.92%. In this study, the frequency domain analysis is extended into the time-frequency domain. The tool wear monitoring performance using an SVM classifier with the signal processing method in [16] and the proposed ECBCA-MSST method is compared in Table 3. It can be seen that the proposed method achieved a 100% accuracy in identifying the good tool condition against the other three methods. The monitoring accuracy of the proposed method for the other five tool wear conditions was also higher. The main reason for the improvement is the contribution of
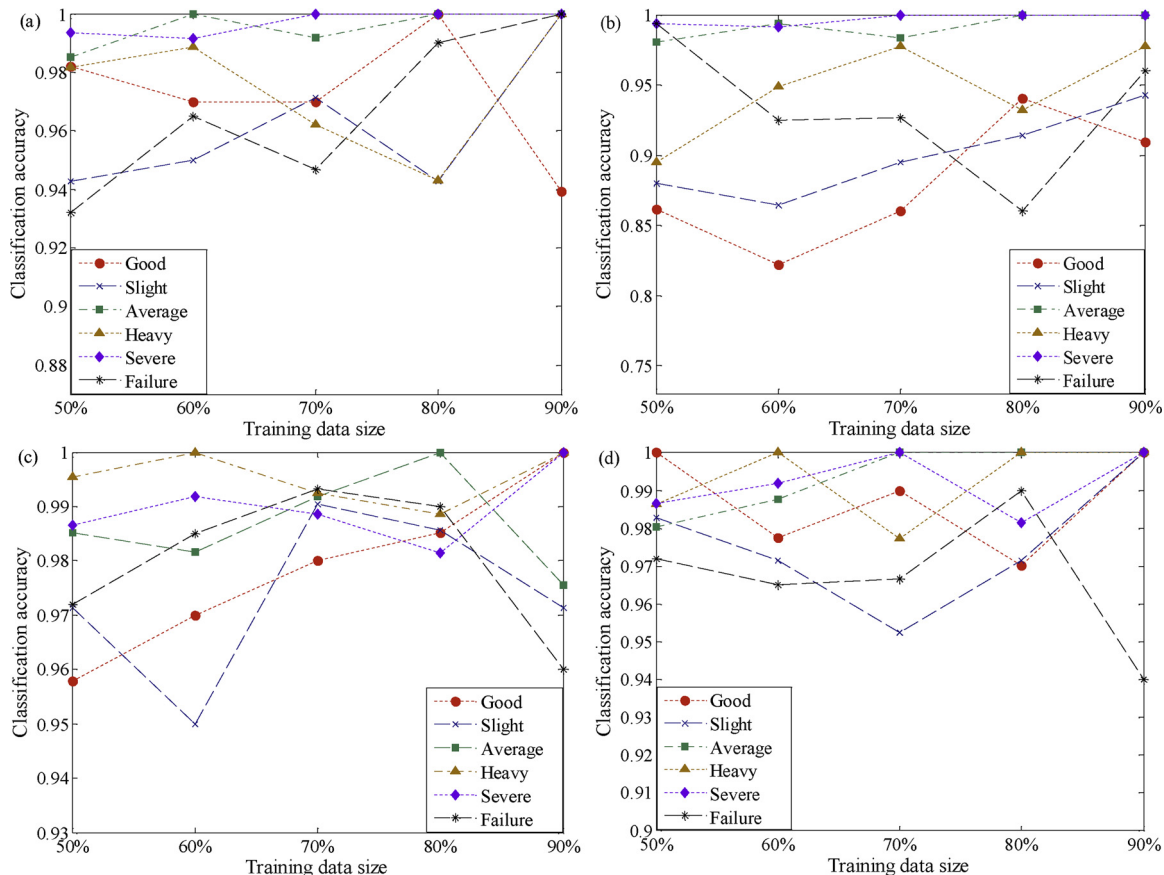

Fig. 15. Overall classification accuracy of four classifiers with varying training size.

the BSS-based multi-channel signal pre-processing. Furthermore, the MSST performed better than the frequency-domain method in [16] with respect to recognition accuracy. This is because MSST is able to extract useful features by analyzing multiple signals in the time-frequency domain. It should be also noted that the SST method is unsuitable for the tool wear monitoring in this study due to ineffective recognition results.

5.4. Discussions

We have demonstrated the proposed method for tool wear monitoring and prediction using half of H as the training dataset. To

Fig. 14. Tool wear monitoring performance using different training size: (a) CART, (b) RF, (c) kNN & (d) SVM.

evaluate the effect of training size on the performance of the predictive model, a varying amount of H was used for training and validation. Tables 4–7 and Fig. 14 show the prediction accuracy of the predictive models trained by CART, RF, kNN, and SVM. As expected, as the amount of training data increases, the classification accuracy increases.

Fig. 15 shows the overall classification accuracy of four classifiers with differing training size. As the training data size increases, the overall classification accuracy increases. The best overall accuracy, 99.0%, was obtained by SVM utilizing 90% of the total dataset.

6. Conclusions and future work

A novel audio-based signal processing approach was presented to improve the accuracy of tool wear prediction. The extended convolutive bounded component analysis (ECBCA) was used to process multivariate audio signals in order to separate a set of source signals from a set of mixed signals. The multivariate synchrosqueezing transform (MSST) was used to characterize multivariate audio signals with time-varying oscillatory properties. The adaptive kernel principal component analysis (AKPCA) was used to transform the denoised signals into a feature space. A set of cutting tests was performed to collect audio signals using multi-channel microphones. These audio signals were used for training and validating the predictive model. The experimental results have shown that the proposed ECBCA-MSST method was able to extract key audio sound sources that are closely related to the dynamic response of tool milling operations. In addition, the AKPCA method was able to improve prediction accuracy while reducing the dimensionality of the extracted features. In the future, the proposed approach will be applied to large volumes of real-time audio signals.

Acknowledgements

References

[1] Wu D, Jennings C, Terpenny J, Kumara S, Gao RX. Cloud-based parallel machine learning for tool wear prediction. J Manuf Sci Eng 2018;140(4):41005.
[2] Ji W, Shi J, Liu X, Wang L, Liang SY. A novel approach of tool wear evaluation. J Manuf Sci Eng 2017;139(9):91015.
[3] Gao R, Wang L, Teti R, Dornfeld D, Kumara S, Mori M, et al. Cloud-enabled prognosis for manufacturing. CIRP Ann Manuf Technol 2015;64(2):749–72.
[4] Wang J, Wang P, Gao RX. Enhanced particle filter for tool wear prediction. Int J Ind Manuf Syst Eng 2015;36:35–45.
[5] Snr DED. Sensor signals for tool-wear monitoring in metal cutting operations—a review of methods. Int J Mach Tools Manuf 2000;40(8):1073–98.
[6] Zhang J, Starly B, Cai Y, Cohen PH, Lee Y-S. Particle learning in online tool wear diagnosis and prognosis. J Manuf Process 2017;28:457–63.
[7] Sick B. On-line and indirect tool wear monitoring in turning with artificial neural networks: a review of more than a decade of research. Mech Syst Signal Process 2002;16(4):487–546.
[8] Wu D, Jennings C, Terpenny J, Gao RX, Kumara S. A comparative study on machine learning algorithms for smart manufacturing: tool wear prediction using random forests. J Manuf Sci Eng 2017;139(7):71018.
[9] Weller EJ, Schrier HM, Weichbro B. What sound can Be expected from a worn tool. J Inst Eng 1969;91(3). 525-&.
[10] Mannan M, Kassim AA, Jing M. Application of image and sound analysis techniques to monitor the condition of cutting tools. Pattern Recognit Lett 2000;21(11):969–79.
[11] Delio T, Tlusty J, Smith S. Use of audio signals for chatter detection and control. J Eng Ind-T Asme 1992;114(2):146–57.
[12] Salgado D, Alonso F. An approach based on current and sound signals for in-process tool wear monitoring. Int J Mach Tools Manuf 2007;47(14):2140–52.
[13] Aliustaoglu C, Ertunc HM, Ocak H. Tool wear condition monitoring using a sensor fusion model based on fuzzy inference system. Mech Syst Signal Process 2009;23(2):539–46.
[14] Ubhayaratne I, Pereira MP, Xiang Y, Rolfe BF. Audio signal analysis for tool wear monitoring in sheet metal stamping. Mech Syst Signal Pr 2017;85:809–26.
[15] Seemuang N, McLeay T, Slatter T. Using spindle noise to monitor tool wear in a turning process. Int J Adv Manuf Tech 2016;86(9–12):2781–90.
[16] Kothuru A, Nooka SP, Liu R. Application of audible sound signals for tool wear monitoring using machine learning techniques in end milling. Int J Adv Manuf Technol 2018;95(9–12):3797–808.
[17] Lauro C, Brandão L, Baldo D, Reis R, Davim J. Monitoring and processing signal applied in machining processes–A review. Measurement 2014;58:73–86.
[18] Zhu K, San Wong Y, Hong GS. Wavelet analysis of sensor signals for tool condition monitoring: a review and some new results. Int J Mach Tools Manuf 2009;49(7–8):537–53.
[19] Fang N, Pai PS, Mosquea S. Effect of tool edge wear on the cutting forces and vibrations in high-speed finish machining of Inconel 718: an experimental study and wavelet transform analysis. Int J Adv Manuf Technol 2011;52(1–4):65–77.
[20] Cao H, Lei Y, He Z. Chatter identification in end milling process using wavelet packets and Hilbert–Huang transform. Int J Mach Tools Manuf 2013;69:11–9.
[21] Zhu K, Hong GS, Wong YS, Wang W. Cutting force denoising in micro-milling tool condition monitoring. Int J Prod Res 2008;46(16):4391–408.
[22] Shi XH, Wang R, Chen QT, Shao H. Cutting sound signal processing for tool breakage detection in face milling based on empirical mode decomposition and independent component analysis. J Vib Control 2015;21(16):3348–58.
[23] Inan HA, Erdogan AT. A convolutive bounded component analysis framework for potentially nonstationary independent and/or dependent sources. IEEE T Signal Proces 2015;63(1):18–30.
[24] Inan HA, Erdogan AT. Convolutive bounded component analysis algorithms for independent and dependent source separation. IEEE Trans Neural Netw Learn Syst 2015;26(4):697–708.
[25] Li Z, Yan X, Wang X, Peng Z. Detection of gear cracks in a complex gearbox of wind turbines using supervised bounded component analysis of vibration signals collected from multi-channel sensors. J Sound Vib 2016;371:406–33.
[26] Tazebay MV, Akansu AN. Adaptive subband transforms in time-frequency excisers for dsss communications-systems. IEEE T Signal Proces 1995;43(11):2776–82.
[27] Ahrabian A, Mandic DP. A class of multivariate denoising algorithms based on synchrosqueezing. IEEE T Signal Proces 2015;63(9):2196–208.
[28] Meignen S, Oberlin T, McLaughlin S. A new algorithm for multicomponent signals analysis based on SynchroSqueezing: with an application to signal sampling and denoising. IEEE T Signal Proces 2012;60(11):5787–98.
[29] Pokharel PP, Liu WF, Principe JC. Kernel least mean square algorithm with constrained growth. Signal Process 2009;89(3):257–65.
[30] Ding MT, Tian Z, Xu HX. Adaptive kernel principal component analysis. Signal Process 2010;90(5):1542–53.
[31] Hu A, Xiang L, Zhang Y. Experimental study on the intrawave frequency modulation characteristic of rotor rub and crack fault. Mech Syst Signal Process 2019;118:209–25.
[32] Reju VG, Koh SN, Soon Y. Underdetermined convolutive blind source separation via time–frequency masking. IEEE Trans Audio Speech Lang Process 2010;18(1):101–16.
[33] Peng D, Smith WA, Randall RB, Peng Z. Use of mesh phasing to locate faulty planet gears. Mech Syst Signal Process 2019;116:12–24.