Open access • Journal Article • DOI:10.1007/S00170-018-3190-4

Intelligent chatter detection using image features and support vector machine

Yun Chen, Huaizhong Li, Xiubing Jing, Liang Hou ...+1 more authors

Institutions: Xiamen University, Griffith University, Tianjin University

Published on: 01 Jun 2019 - The International Journal of Advanced Manufacturing Technology (Springer London)

Topics: Feature (computer vision), Support vector machine, Continuous wavelet transform and Machining

The International Journal of Advanced Manufacturing Technology
https://doi.org/10.1007/s00170-018-3190-4

ORIGINAL ARTICLE

Intelligent chatter detection using image features and support vector machine

Yun Chen[1] · Huaizhong Li[2] · Xiubing Jing[3] · Liang Hou[1] · Xiangjian Bu[1]

Received: 5 June 2018 / Accepted: 12 December 2018

Abstract
Chatter is a self-excited vibration that affects the part quality and tool life in the machining process. This paper introduces an intelligent chatter detection method based on image features and the support vector machine. In order to reduce the background noise and highlight chatter characteristics, the average FFT is applied to identify the dominant frequency bands that divide the time-frequency image of the short-time Fourier transform into several sub-images. The non-stationary properties of the machining conditions are quantified using sub-images features. The area under the receiver operating characteristics curve ranks the extracted image features according to their separability capabilities. The support vector machine is designed to automatically classify the machining conditions and select the best feature subset based on the ranked features. The proposed method is verified by using dry micro-milling tests of steel 1040 and high classification accuracies for both the stable and unstable tests are obtained. In addition, the proposed method is compared with two additional methods using either image features from the continuous wavelet transform or time-domain features. The results present a better classification performance than the two additional methods, indicating the efficiency of the proposed method for chatter detection.

Keywords Micro-milling · Chatter detection · Image features · Dominant frequency bands · Support vector machine

1 Introduction

Chatter is a self-excited vibration due to complex, nonlinear, and non-stationary characteristics of the machining process. Poor surface quality, reduced tool life, and breakage of cutting tools may be induced when chatter occurs. Development of the stability chart defined in the space of the spindle speed and axial depth of cut is an analytical approach to avoid chatter. This approach is based on a series of simplifications of the machining system, such as an assumption of a linear and time-invariant machining system [1]. Hence, it cannot fully model the nonlinear and dynamic machining system with various uncertainties [2]. An alternative approach for avoiding chatter is to monitor the machining process. Given the recent advances in signal processing and pattern recognition, intelligent and automatic chatter detection is preferred in practical industrial applications, as it is the prerequisite for timely chatter suppression and also suitable for fully automated or lightly staffed machining environments.

Due to non-stationary and uncertain characteristics of measured signals generated in machining operations, chatter detection is sometimes too tricky in order to achieve reasonable performance. Numerous algorithms have been applied to process the measured signals for chatter detection, including Fourier transform (FT) analysis [3, 4] and time-frequency analysis [5]. Although the FT analysis has provided acceptable results for chatter detection [4, 6], it is unable to determine the transition in non-stationary signals and hence ineffective for real-time or intelligent chatter detection [7–10]. The time-frequency analysis offers the advantage over a single time or frequency domain analysis, which can reflect the non-stationary properties of the machining operations.

The commonly used time-frequency methods for chatter detection include the short-time Fourier transform (STFT)

Yun Chen and Huaizhong Li contributed equally to this work.

✉ Liang Hou
hliang@xmu.edu.cn

[1] Department of Mechanical and Electrical Engineering, Xiamen University, Xiamen 361005, China

[2] Griffith School of Engineering, Gold Coast Campus, Griffith University, Brisbane, QLD 4222, Australia

[3] Key Laboratory of Mechanism Theory & Equipment Design of ministry of education, Tianjin University, Tianjin 300072, China

Published online: 07 January 2019

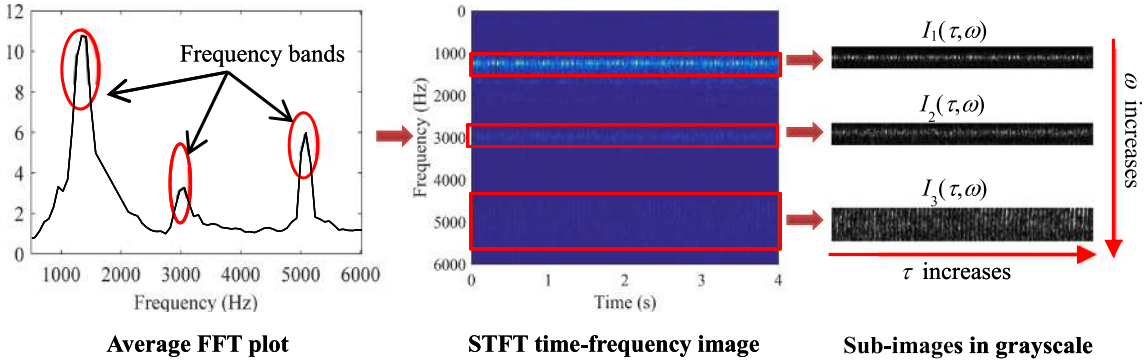

Fig. 1 Dominant frequency bands and sub-images

[11], empirical mode decomposition (EMD) based approach [7], and wavelet-based method [12, 13]. Those methods usually decompose the measured signals into several components, and each component contains information in a narrow frequency band. The components that are sensitive to the abnormalities of the machining operations are selected for feature extraction. Mei et al. [14] used singular spectrum analysis to decompose the signals, and a sensitive component was selected to extract the singular value entropy as chatter indicator. Liu et al. [13] extracted features from the dominant frequency bands decomposed from the wavelet packet transform. Although the decomposition of the original signals into components limited in narrow frequency bands has received reasonable performance, these features are still extracted from one-dimensional components in the time domain, which do not fully explore the advantage of the time-frequency analysis that defines a two-dimensional space. To avoid this problem, the spectra obtained from the time-frequency analysis can be treated as two-dimensional images, and corresponding image features can therefore be used for chatter detection. Statistical approaches are the common tools used to extract image features, and represent the image using statistics collected from the distributions and relationships between the gray levels of an image [15]. Dynamic changes that may be hidden in the single time or frequency domain can be discovered by the time-frequency image and bring about variations of neighboring pixel distributions and relationships in the image. The image features from the time frequency analysis are often used in the biomedical field for brain diseases [16–18]. Boashash et al. [17] employed time-frequency image features for detection of abnormalities in newborn electroencephalogram (EEG) signals, and concluded that the two-dimensional image features performed significantly better than the one-dimensional frequency-domain features. For machining, Khalifa et al. [19] analyzed the machined surface images using the statistical approaches to distinguish the stable and unstable machining conditions.

Chatter is highly related to the dynamic behavior of the machining system. When chatter occurs, energy rises around the specific natural frequencies of the machining system [20, 21]. In order to decrease the high level of environmental noises and keep as much dynamic information as possible, the frequency bands that contain the natural frequencies are selected to highlight chatter-related characteristics. Signal decomposition is commonly used to decompose the signal into a set of components with narrow frequency bands. However, most signal decomposition techniques are not adaptive, such as wavelet packet decomposition. They may divide one continuous frequency band into two individual components or two (or more) individual frequency bands into one component, resulting in less sensitive features related to chatter [22]. EMD is an adaptive decomposition based on the empirical modes, but suffers from mode mixing [23]. Additional assistances from white noises [23] or wavelet decomposition [24] can be used to avoid mode mixing with increased costs of computation time. On the other hand, there are alternative methods to identify the dominant frequency bands. For instance, Attoui et al. [22] processed the FFT of the signal using a short-time window to yield the dominant frequency bands, and Lamraoui et al. [25] used the hammer impact test to identify the natural frequencies of the machining system and therefore the dominant frequency bands.

Although the extracted features can quantify the degree of stability of the machining condition, additional efforts are

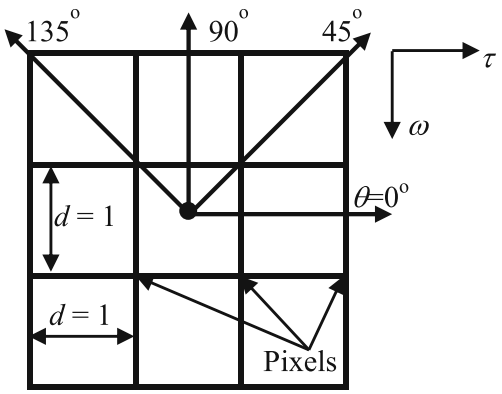

Fig. 2 GLCM construction methods at different angles

Table 1 Image features and their notation for the nth sub-image $I_n(\tau, \omega)$

Feature	Notation	Feature	Notation	Feature	Notation
Mean	F_1^n	Mean contrast	F_5^n	Contrast range	F_9^n
Standard variance	F_2^n	Mean correlation	F_6^n	Correlation range	F_{10}^n
Skewness	F_3^n	Mean energy	F_7^n	Energy range	F_{11}^n
Kurtosis	F_4^n	Mean homogeneity	F_8^n	Homogeneity range	F_{12}^n

required to determine the boundary between the stable and unstable tests in the feature space. Pattern recognition can automatically discriminate the stable and unstable tests through a learning process. Several smart recognition techniques have been used to monitor machining conditions, including neural network [25], fuzzy logic [26], and hidden Markov models [27]. Teti et al. [28] stated that neural network and fuzzy logic techniques are the most widely used ones in monitoring of machining operations. Alternatively, the support vector machine (SVM) was applied for chatter detection in mirror milling [29], as it has simple geometric interpretation and is suitable for small sample sizes [30]. The application of the SVM to a tool breakage diagnosis system also obtained reliable results, as indicated by Hsueh and Yang [31].

In the authors' previous work [32], the time-frequency analysis is an ideal tool for processing the non-stationary signals generated in machining operations, and its image features from the STFT show better performance in discriminating the stable and unstable tests than the time-domain features in macro-milling. This paper is a continuous effort to extend time-frequency images features for intelligent chatter detection in micro-milling. The high-energy frequency bands extracted from the average FFT are used to reduce the background noise and highlight chatter characteristics. The area under the receiver operating characteristics curve (AUC) and the SVM are combined to classify the machining conditions and identify the best feature subset with the highest classification accuracies. The proposed chatter detection method is compared with two additional methods using either time-domain features or image features from the continuous wavelet transform. The results indicate a good classification performance of the proposed method.

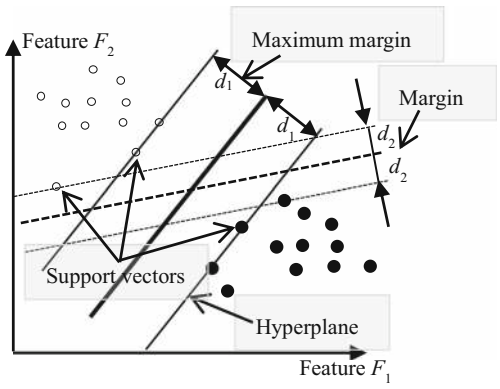

Fig. 3 SVM for a two-class case

2 Methodology

2.1 Feature extraction and ranking

The feature extraction and ranking techniques proposed by the authors [32] are briefly summarized in this section. In order to identify the dominant frequency bands with high energy, a constructed signal $y(t_k)$, that is, a sum of signals selected from the stable and unstable machining conditions is processed using an average fast Fourier transform (FFT).

$$Y(\omega) = \sum_{l=1}^{L} \sum_{k=0}^{K} \mathrm{FFT}[y(t_k) w_0(t_k - a_l)] \tag{1}$$

where t_k is the kth sample time, $w_0(t_k - a_l)$ is a rectangular window function centered at time a_l, and $\{a_l\}$ $(l = 1,\ldots, L)$ is an arithmetic sequence. FFT$[y(t_k)]$ is the FFT of the signal $y(t_k)$. The FFT algorithm is implemented using the function "fft" in Matlab [33]. Equation (1) defines an average FFT $Y(\omega)$ over L segments of the constructed signal $y(t_k)$. The average FFT smooths the FFT spectrum and reduces the background noises, resulting in highlighted dominant frequency bands.

The short-time Fourier transform (STFT) is a commonly used method to discover the information of the non-stationary signal in both the time and frequency domains. The signal $x(t)$ is transformed by applying STFT.

$$I(\tau, \omega) = \int_{-\infty}^{+\infty} x(t) w_1(t-\tau) e^{-j\omega t} dt \tag{2}$$

where $w_1(t\text{-}\tau)$ is a window function centered at time τ. $I(\tau, \omega)$ defines a time-frequency image. This image is then

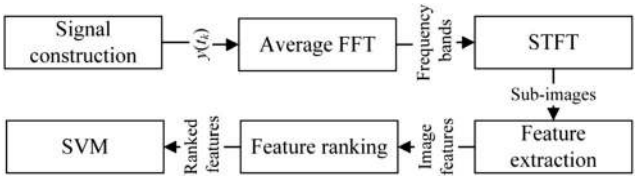

Fig. 4 Proposed chatter detection method

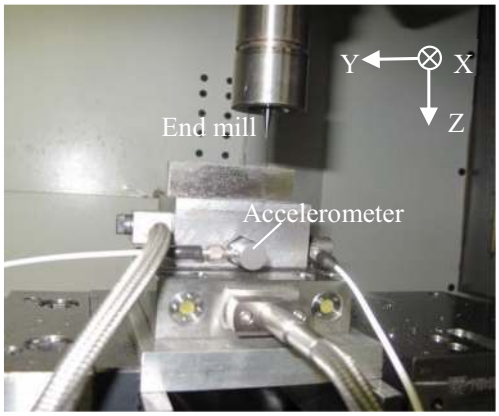

Fig. 5 Micro-milling experimental setup

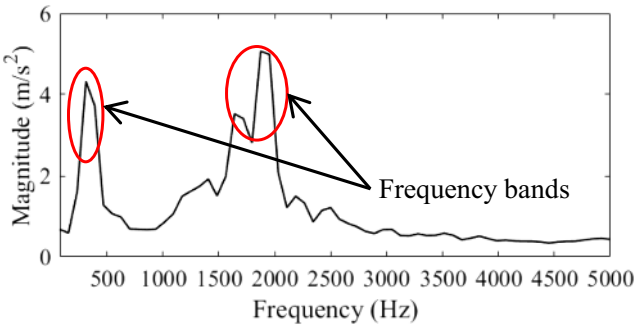

Fig. 6 Average FFT plot

divided into a set of sub-images according to the dominant frequency bands and converted to grayscale sub-images with G gray levels, as shown in Fig. 1. The first-order statistical and second-order statistical approaches are used to extract the features from the grayscale image. Detailed algorithms for image features are given in Appendix. The first-order statistics are calculated from the intensity-level histogram $h(i)$:

$$h(i) = \sum_{\tau}\sum_{\omega}\delta(I_n(\tau,\omega),i) \quad i = 0,\cdots,G{-}1 \tag{3}$$

where i is a gray level value, $I_n(\tau,\omega)$ is the nth sub-image in grayscale (Fig. 1) and δ is the Kronecker delta function. The first-order statistics used for image feature extraction are the mean, standard variance, skewness, and kurtosis of $h(i)$. The second-order statistics are extracted from the gray level co-occurrence matrix (GLCM). GLCM considers the relationship between two pixels, a distance d apart along a given direction θ having co-occurring gray values i and j:

$$GLCM_{d,\theta}(i,j) = \begin{cases} \sum_{\tau}\sum_{\omega} 1\, I_n(\tau,\omega) = i \& I_n(\tau + d\cos\theta, \omega + d\sin\theta) = j \\ 0 \qquad others \end{cases} \tag{4}$$

Four second-order statistical features, that is, contrast, correlation, energy, and homogeneity, are evaluated from the GLCM. Let the distance $d = 1$ pixel and $\theta = 0^o$, 45^o, 90^o, or 135^o (Fig. 2), there are four sets of second-order statistical features. The mean and range of each second-order feature are used for chatter detection. Table 1 lists all the statistical features and their notations.

Irrelevant features should be removed in order to improve the classification performance and decrease the computation time. The area under the receiver operating characteristics curve (AUC) is used to rank the image features in terms of their separability capabilities [34]. To select the best feature subset, an algorithm that iteratively realizes the classification first for the subset with the top ranked feature then for subset with top two features then for top three features, etc. until all features are classified. The best feature subset is chosen as the subset with best classification accuracies.

Table 2 Cutting parameters

Cutting parameter	Values
Spindle speed n_p (rpm)	10,000, 12,000, 14,000, 16,000
Feed per tooth f_t (μm/tooth)	0.5, 1, 2, 3, 5, 7
Axial depth of cut a_p (μm)	50, 80, 110, 140

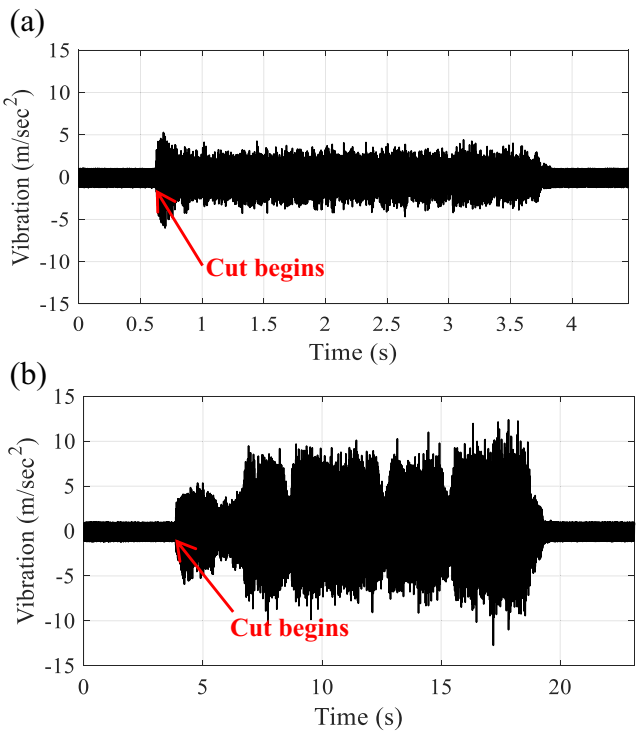

Fig. 7 Vibration signals **a** stable at n_p 16,000 rpm, a_p 50 μm, and f_t 5 μm/tooth **b** unstable at n_p 16,000 rpm, a_p 50 μm, and f_t 1 μm/tooth

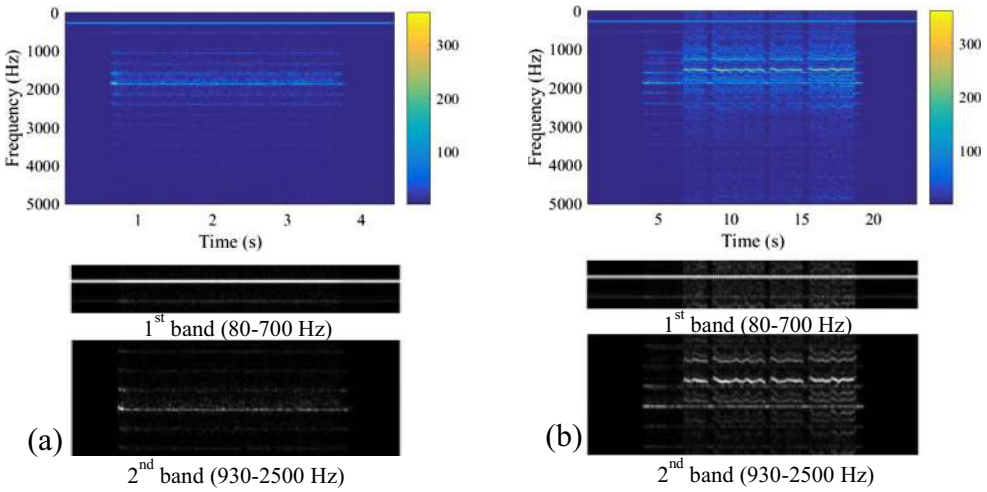

Fig. 8 STFT images and sub-images for tests in Fig. 7 **a** stable, **b** unstable

2.2 Classification

The SVM is suitable to classify data samples with small sizes. It separates classes using the training data by fitting an optimal separating hyperplane in a feature space. The optimization problem being solved aims to maximize the margins between the optimal hyperplanes and the closest training data (i.e., support vectors), as shown in Fig. 3. Let $\{\mathbf{x}\}$ be the feature vectors of the training data set and $\boldsymbol{y}\in\{-1, 1\}$ be the corresponding class labels. The decision function $g(\boldsymbol{x})$ is defined by the weight vector $\boldsymbol{w}$ and the threshold w_0 as.

$$g(\mathbf{x}) = \mathbf{w}\cdot\mathbf{x} + w_0 \tag{5}$$

$$\text{If } \mathbf{w}\cdot\mathbf{x} + w_0 \geq 1\ (\leq -1), \mathbf{x}\in\text{Class } \mathbf{y} = 1(-1) \tag{6}$$

The support vectors lie on two hyperplanes $g(\mathbf{x}) = \pm 1$. The margin maximization between the two hyperplanes is solved by minimizing $\|\mathbf{w}\|^2$. Figure 3 shows a completely separate case with linear hyperplanes in a simple two-dimensional feature space. In more practical cases, features in higher dimensional space are usually used, and the classes are more likely non-separable. Thus, nonlinear kernels may be needed for better classification.

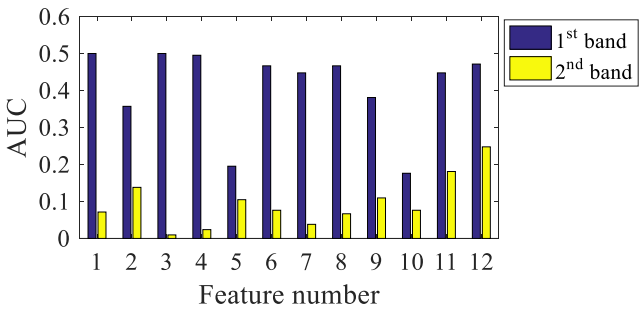

Fig. 9 AUC for each feature from STFT images

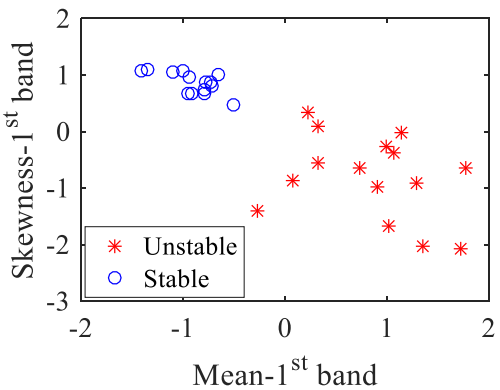

Fig. 10 Test distribution in the space defined by the mean F_1^1 and skewness F_3^1 from the STFT images

2.3 Proposed chatter detection method

An intelligent chatter detection method based on image features and SVM is proposed in this work. The procedures to conduct the proposed method are shown in Fig. 4. Several signals from the stable and unstable machining conditions are summed up to construct a synthesized signal for the average FFT. In order to highlight the chatter-related characteristics, the dominant frequency bands are identified by localizing the frequency bands at which the energy is high

Table 3 Sensitivity analysis of the frequency resolution

FR	NFBS	AC_{sd}	AC_{usd}	AC
2^7	2	100.0%	93.3%	96.6%
2^8	2	100%	100%	100%
2^9	2	100%	100%	100%
2^{10}	2	100%	100%	100%

Table 4 Time domain features used by Lamraoui et al. [25]

Feature	Notation	Feature	Notation
Variance (F_{t1})	$\sum_{k=1}^{K}(x(t_k)-\bar{x})^2/K$ *	Skewness (F_{t6})	$\left[\sum_{k=1}^{K}(x(t_k)-\bar{x})^3/K\right]/F_{t1}{}^{3/2}$
Kurtosis (F_{t2})	$\left[\sum_{k=1}^{K}(x(t_k)-\bar{x})^4/K\right]/F_{t1}{}^2$	Peak value (F_{t7})	$[\max x(t_k)-\min x(t_k)]/2$
Root mean square (F_{t3})	$\sqrt{\sum_{k=1}^{K}x^2(t_k)/K}$	Clearance factor (F_{t8})	$F_{t7}/\left[\sum_{k=1}^{K}\lvert x(t_k)\rvert/K\right]^2$
Crest factor (F_{t4})	F_{t7}/F_{t3}	Shape factor (F_{t9})	$F_{t3}/\sum_{k=1}^{K}\lvert x(t_k)\rvert/K$
Impulse factor (F_{t5})	$F_{t7}/\left[\sum_{k=1}^{K}\lvert x(t_k)\rvert/K\right]$		

*$x(t_k)$ is the envelope of a signal pre-processed by multiband filters

in the average FFT plot. The measured vibration signals are then transformed into the time-frequency images using the STFT. A set of sub-images are selected from the obtained time-frequency images according to the identified dominant frequency bands and converted into grayscale images. Statistical image features are extracted from each sub-image and ranked according to the AUC. Finally, the SVM classifier is employed for the ranked features, and the best feature subset is selected according to the highest classification accuracies.

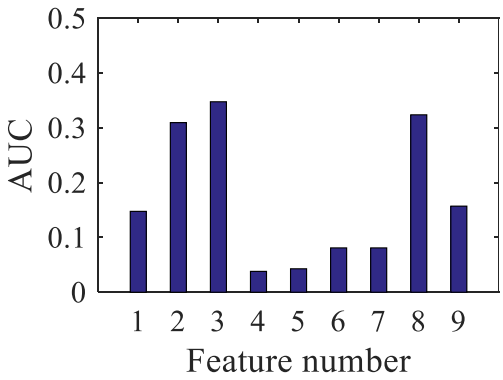

Fig. 11 AUC for each time-domain feature

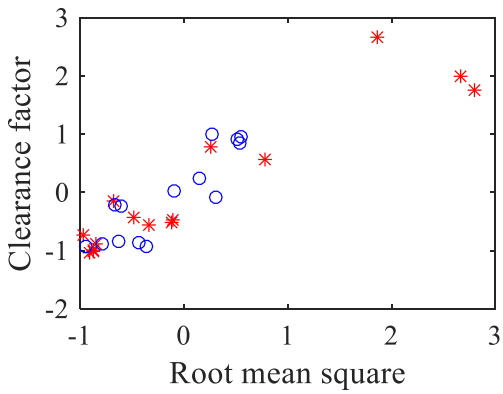

Fig. 12 Test distribution in the space defined by the root mean square F_{t3} and clearance factor F_{t8}

3 Experimental assessment

3.1 Experimental setup

Micro-milling tests are used to assess the proposed chatter detection method using the STFT image features. The experimental setup and the coordinate system of the vibration signals are shown in Fig. 5. The workpiece material is steel 1040. The feed direction is the x direction. The accelerometer was mounted on the workpiece to measure the vibration signals in the x direction. The sample rate was 10 kHz. Slot micro-milling was performed throughout the tests. Two-flute end mill cutters with a diameter of 600 μm and a helix angle of 30° were used. The cutting parameters are given in Table 2, and the total number of tests is 29.

3.2 Detection results

More than five micro-milling tests for each case of machining condition, either stable or unstable, are summed to obtain the average FFT plot as shown in Fig. 6. The average FFT is evaluated over a rectangular window size of 256. Two dominant frequency bands, i.e., 80–700 Hz (the first frequency band) and 930–2500 Hz (the second frequency band), are identified. To examine the difference between the stable and unstable tests, two vibration signals used for the average FFT

Table 5 Classification results

Approach	NFBS	AC_{sd}	AC_{usd}	AC
STFT*	2	100%	100%	100%
TD**	6	100%	93.3%	96.6%
CWT	2	92.9%	100.0%	96.6%

*Frequency resolution is 2^8; ** TD represents time domain

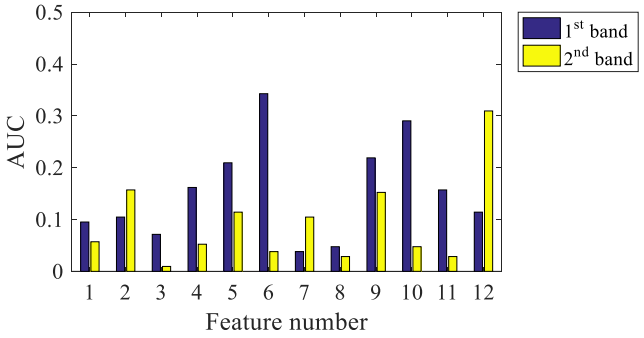

Fig. 14 AUC for each CWT image feature

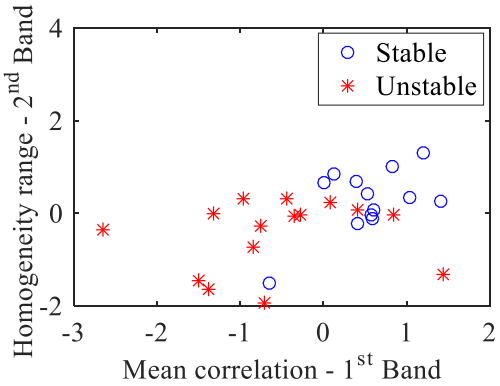

Fig. 15 Test distribution in the space defined by the mean correlation F_6^1 and homogeneity range F_{12}^2 from the CWT images

plot are shown in Fig. 7. It can be seen that the vibration signal for the stable test has a relatively small amplitude variation, whereas the signal for the unstable test shows a significant amplitude variation.

The STFT images and the corresponding sub-images for the two tests in Fig. 7 are given in Fig. 8. The STFT images are obtained using a Hanning window with a size of 2^9 and an overlap of a quarter of the window size. The STFT has a trade-off between the time and frequency resolutions, i.e., the higher the frequency resolution, the lower the time resolution is and vice versa. The time resolution of the STFT is generally much lower than the number of sample points. For example, Fig. 8a has a time resolution of 1340 pixels, whereas the number of sample points is 44,551. Sub-images are firstly identified from the STFT images according to the dominant frequency bands, and then converted into grayscale images. A sensitivity analysis is conducted to choose the total number G of gray levels. It is found that the image features become stable when $G >= 2^8$. Thus, G is set to 2^8. The image patterns significantly differ between the stable and unstable tests. When chatter occurs, all frequency components vary significantly. However, the first and second frequency bands contain a majority of high-intensity pixels, indicating high energies for those frequency bands and large signal-to-noise ratios. Thus, the STFT image features from the two bands are more sensitive to chatter onset.

The AUC that assesses the separability capability of each feature is shown in Fig. 9. It can be seen that most image features from the first frequency band have large AUCs and therefore, high separability capabilities. The top ranked features are the mean F_1^1 and skewness F_3^1 with the same AUC of 0.5. Figure 10 gives the distribution of the micro-milling tests in the two-dimensional space defined by the mean F_1^1 and skewness F_3^1. The test distribution in the two-feature space shows no overlap. A less overlap distribution usually implies a better pattern recognition performance.

A linear SVM is used for classification, as it gives better classification results than the nonlinear SVMs in this study. The performance of the classification approach is assessed using the leave-one-out (LOO) method [35]. The input

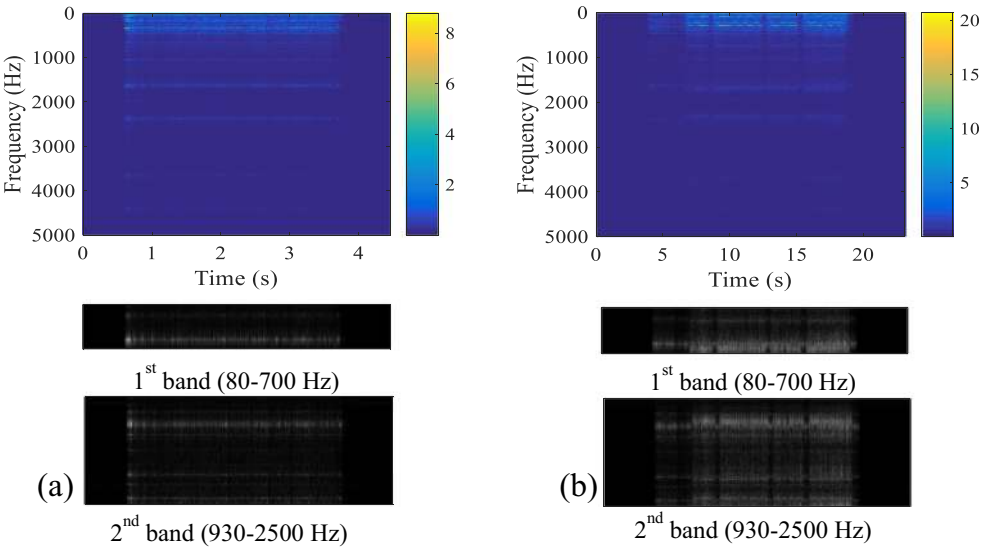

Fig. 13 CWT images for tests in Fig. 7 **a** stable, **b** unstable

Table 6 First-order statistical features for the nth sub-image $I_n(\tau, \omega)$

Feature	Notation	Feature	Notation
Mean (F_1^n)	$\frac{1}{N}\sum_{i=0}^{G-1} i h_n(i)$	Skewness (F_3^n)	$\frac{1}{N}\left(F_2^n\right)^{-3}\sum_{i=0}^{G-1}\left(i-F_1^n\right)^3 h_n(i)$
Standard variance (F_2^n)	$\sqrt{\frac{1}{N}\sum_{i=0}^{G-1}\left(i-F_1^n\right)^2 h_n(i)}$	Kurtosis (F_4^n)	$\frac{1}{N}\left(F_2^n\right)^{-4}\sum_{i=0}^{G-1}\left(i-F_1^n\right)^4 h_n(i)-3$

features are normalized using the z-score method before classification. Three types of accuracies are defined for classification assessment, namely, AC_{sd}, AC_{usd}, and AC, which are the percentages of stable tests classified as stable, unstable tests classified as unstable, and total tests classified correctly, respectively. The best feature subset that gives the highest classification accuracies is identified by using the iterative procedure introduced in Section 2.1. As the resolution of the time-frequency image affects separability capabilities of features and classification performance [17], a sensitivity analysis is conducted to investigate the effect of the resolution on the classification performance. As the numbers of the sample points of the measured signals are fixed, the time resolution is fully determined by the window size or the frequency resolution (FR). Table 3 lists the classification accuracies and the number of features in the best subset (NFBS) under different FRs. All the classification accuracies AC_{sd}, AC_{usd}, and AC reach 100% when the FR is equal or higher than 2^8. As a high FR increases the computation time for obtaining STFT images and decreases the time resolution, a FR of 2^8 pixels is selected for feature ranking and classification, as well as the further discussion presented in Section 3.3.

3.3 Discussion

In order to further assess the performance of the proposed method for chatter detection, the image features based on the STFT are compared with the time-domain features extracted using the approach proposed by Lamraoui et al. [25]. Lamraoui et al. designed a set of multiband resonance filters based on the dominant frequency bands to filter the vibration signals. The TD features in Table 4 are extracted from the envelopes of the filtered signals. Figure 11 shows the AUC for each time-domain feature. The root mean square F_{t3} has the maximum AUC of 0.34, which is much lower than the maximum from the STFT image features. Figure 12 shows the distribution of the micro-milling tests in the space defined by the top two time-domain features F_{t3} and F_{t8}. The test distribution shows significant overlapping between the stable and unstable tests, which indicates the low separability capabilities of the time domain features. Table 5 presents the best feature subsets and the classification accuracies using the time-domain features. Similar to the maximum AUC, the classification accuracies for the time-domain features are lower than the corresponding ones for the STFT image features. This again evidences that the STFT image features are powerful for intelligent chatter detection.

The STFT image features are also compared with the continuous wavelet transform (CWT) image features in terms of their classification performance. The CWT images and their corresponding sub-images for the two tests in Fig. 7 are given in Fig. 13. The CWT images are obtained using a complex Morlet wavelet. The FR of 2^8 for CWT images is equal to that for STFT images. The time resolution for CWT images is much higher, which is the same as the number of sample points of signals. Similar to STFT images, CWT images also discover the non-stationary properties of signals when chatter occurs. Figure 14 gives the AUC for each feature from the CWT images. The mean correlation F_6^1 from the first frequency band has the maximum AUC of 0.34, which is lower than the maximum AUC 0.5 in Fig. 9. Thus, the test distribution (Fig. 15) in the top two-feature space for the CWT shows an overlap between the stable and unstable tests. Similarly, the classification accuracies AC_{sd} and AC based on the CWT image features are lower than their corresponding accuracies based on the STFT image features (Fig. 15). Although the CWT can output images with both high time and frequency resolutions, it does not give better performance for chatter detection than the low-resolution images from the STFT. It is important to select the proper time-frequency analysis method for better classification performance [17].

Table 7 Definitions of variables μ_i, μ_j, σ_i, and σ_j

Variable	Notation	Variable	Notation
μ_i	$\sum_{i=0}^{G-1} i\sum_{j=0}^{G-1} p_{d,\theta}^n(i,j)$	σ_i	$\sqrt{\sum_{i=0}^{G-1}(i-\mu_i)^2\sum_{j=0}^{G-1} p_{d,\theta}^n(i,j)}$
μ_j	$\sum_{i=0}^{G-1}\sum_{j=0}^{G-1} j p_{d,\theta}^n(i,j)$	σ_j	$\sqrt{\sum_{j=0}^{G-1}\left(j-\mu_j\right)^2\sum_{i=0}^{G-1} p_{d,\theta}^n(i,j)}$

Table 8 Second-order statistical features for the nth sub-image $I_n(\tau, \omega)$

Feature	Notation	Feature	Notation
Mean contrast (F_5^n)	$\text{mean}\left(F^{Con}_{d=1,\theta=\{\theta_m\}}\right)^*$	Contrast range (F_9^n)	$\max\left(F^{Con}_{d=1,\theta=\{\theta_m\}}\right) - \min\left(F^{Con}_{d=1,\theta=\{\theta_m\}}\right)$
Mean correlation (F_6^n)	$\text{mean}\left(F^{Cor}_{d=1,\theta=\{\theta_m\}}\right)$	Correlation range (F_{10}^n)	$\max\left(F^{Cor}_{d=1,\theta=\{\theta_m\}}\right) - \min\left(F^{Cor}_{d=1,\theta=\{\theta_m\}}\right)$
Mean energy (F_7^n)	$\text{mean}\left(F^{E}_{d=1,\theta=\{\theta_m\}}\right)$	Energy range (F_{11}^n)	$\max\left(F^{E}_{d=1,\theta=\{\theta_m\}}\right) - \min\left(F^{E}_{d=1,\theta=\{\theta_m\}}\right)$
Mean homogeneity (F_8^n)	$\text{mean}\left(F^{H}_{d=1,\theta=\{\theta_m\}}\right)$	Homogeneity range (F_{12}^n)	$\max\left(F^{H}_{d=1,\theta=\{\theta_m\}}\right) - \min\left(F^{H}_{d=1,\theta=\{\theta_m\}}\right)$

*$\{\theta_m\} = \{0^o, 45^o, 90^o, 135^o\}$

4 Conclusions

In this paper, an intelligent chatter detection method based on STFT image features and SVM is proposed for micro-milling operations. Image features are used to describe the non-stationary properties shown in the STFT time-frequency images. The dominant frequency bands that are identified from the average FFT are used to divide the STFT images, in order to increase the signal-to-noise ratio, as well as the sensitivity of image features. The AUC that quantifies the separability capability of each feature is used to identify the sensitive image features and best feature subset. A linear SVM for classification is used to automatically determine the boundary between the stable and unstable tests in the best feature subset. The classification performance is evaluated using the LOO method.

The proposed chatter detection method is implemented in micro-milling of steel 1040. The two features, namely, mean F_1^1 and skewness F_1^3 from the first dominant frequency band, have the largest AUCs and high capabilities to separate the stable and unstable tests. The classification accuracies AC_{sd}, AC_{usd}, and AC all reach 100% using the best subset consisted of the top two features. For further verification, the proposed method is compared with two additional chatter detection methods. The results show that the STFT image features give a better classification performance than the time-domain features and the CWT image features. Thus, the proposed method based on STFT image features from dominant frequency bands is efficient for chatter detection. It is interesting to find that high-resolution images obtained from the CWT do not improve the classification accuracies as expected. To fully explore the image features for intelligent chatter detection, it is recommended to select the proper time-frequency analysis method for better classification performance.

Acknowledgments This study was jointly supported by the Collaborative Innovation Center of High-End Equipment Manufacturing in Fujian and International Postdocs Exchange Program. The authors would like to express their acknowledgements to the Advanced Manufacturing Laboratory, UNSW, for the support of the experimental work. Comments and suggestions from reviewers are greatly appreciated.

Appendix

The first-order statistics, that is, mean F_1^n, standard variance F_2^n, skewness F_3^n, and kurtosis F_4^n, are defined in Table 6. The variables G and N are the number of gray levels and number of pixels for the sub-image $I_n(\tau, \omega)$, respectively.

The second-order statistical features, i.e., contrast, correlation, energy, and homogeneity, are defined as [36]:

$$\text{Contrast}: F^{Con}_{d,\theta} = \sum_{j=0}^{G-1}\sum_{i=0}^{G-1}(i-j)^2 p^n_{d,\theta}(i,j) \quad (7)$$

$$\text{Correlation}: F^{Cor}_{d,\theta} = \sum_{j=0}^{G-1}\sum_{i=0}^{G-1}\left(ijp^n_{d,\theta}(i,j) - \mu_i\mu_j\right)/\sigma_i\sigma_j \quad (8)$$

$$\text{Energy}: F^{E}_{d,\theta} = \sum_{j=0}^{G-1}\sum_{i=0}^{G-1}\left[p^n_{d,\theta}(i,j)\right]^2 \quad (9)$$

$$\text{Homogeneity}: F^{H}_{d,\theta} = \sum_{j=0}^{G-1}\sum_{i=0}^{G-1} p^n_{d,\theta}(i,j)/1 + (i-j)^2 \quad (10)$$

where $p^n_{d,\theta}(i,j)$ is an approximate second-order statistical probability for changes between gray levels i and j at a pair (d, θ).

$$p^n_{d,\theta}(i,j) = \frac{GLCM^n_{d,\theta}(i,j)}{\sum_j\sum_i GLCM^n_{d,\theta}(i,j)} \quad (11)$$

The variables μ_i, μ_j, σ_i, and σ_j are the means and standard deviations of the marginal probability matrix elements, and defined in Table 7. Since four angles $\{\theta_m | 0^o, 45^o, 90^o, 135^o\}$ are considered for evaluating GLCM, there are four sets of

second-order statistical features. The mean and range of each type of a second-order statistical feature are used for chatter detection (Table 8).

References

1. Jin X, Altintas Y (2011) Slip-line field model of micro-cutting process with round tool edge effect. J Mater Process Technol 211(3):339–355
2. Peng Y (2004) Empirical model decomposition based time-frequency analysis for the effective detection of tool breakage. J Manuf Sci Eng 128(1):154–166
3. Li XQ, Wong YS, Nee AYC (1997) Tool wear and chatter detection using the coherence function of two crossed accelerations. Int J Mach Tool Manu 37(4):425–435
4. Delio T, Tlusty J, Smith S (1992) Use of audio signals for chatter detection and control. J Eng Ind 114(2):146–157
5. Cao H, Yue Y, Chen X, Zhang X (2017) Chatter detection in milling process based on synchrosqueezing transform of sound signals. Int J Adv Manuf Technol 89(9):2747–2755
6. Hashimoto M, Marui E, Kato S (1996) Experimental research on cutting force variation during regenerative chatter vibration in a plain milling operation. Int J Mach Tool Manu 36(10):1073–1092
7. Cao H, Lei Y, He Z (2013) Chatter identification in end milling process using wavelet packets and Hilbert–Huang transform. Int J Mach Tool Manu 69 (0):11–19
8. Cao H, Zhou K, Chen X (2015) Chatter identification in end milling process based on EEMD and nonlinear dimensionless indicators. Int J Mach Tool Manu 92 (0):52–59
9. Pérez-Canales D, Álvarez-Ramírez J, Jáuregui-Correa JC, Vela-Martínez L, Herrera-Ruiz G (2011) Identification of dynamic instabilities in machining process using the approximate entropy method. Int J Mach Tool Manu 51(6):556–564
10. Vela-Martínez L, Carlos Jauregui-Correa J, Rodriguez E, Alvarez-Ramirez J (2010) Using detrended fluctuation analysis to monitor chattering in cutter tool machines. Int J Mach Tool Manu 50(7): 651–657
11. Griffin J, Chen X (2009) Multiple classification of the acoustic emission signals extracted during burn and chatter anomalies using genetic programming. Int J Adv Manuf Technol 45(11–12):1152–1168
12. Yao Z, Mei D, Chen Z (2010) On-line chatter detection and identification based on wavelet and support vector machine. J Mater Process Technol 210(5):713–719
13. Liu Y, Wang X, Lin J, Zhao W (2016) Early chatter detection in gear grinding process using servo feed motor current. Int J Adv Manuf Technol 83(9):1801–1810
14. Mei Y, Mo R, Sun H, Bu K (2017) Chatter detection in milling based on singular spectrum analysis. Int J Adv Manuf Technol 95(9–12):3475–3486
15. Materka A, Strzelecki M (1998) Texture analysis methods–a review. Technical University of Lodz, Institute of Electronics, COST B11 report, Brussels
16. Alcn OF, Siuly S, Bajaj V, Guo Y, Sengur A, Zhang Y (2016) Multi-category EEG signal classification developing time-frequency texture features based Fisher vector encoding method. Neurocomputing 218:251–258
17. Boashash B, Khan NA, Ben-Jabeur T (2015) Time–frequency features for pattern recognition using high-resolution TFDs: a tutorial review. Digit Signal Process 40:1–30
18. Boashash B, Barki H, Ouelha S (2017) Performance evaluation of time-frequency image feature sets for improved classification and analysis of non-stationary signals: application to newborn EEG seizure detection. Knowl Based Syst 132:188–203
19. Khalifa OO, Densibali A, Faris W (2006) Image processing for chatter identification in machining processes. Int J Adv Manuf Technol 31(5):443–449
20. Li H, Jing X, Wang J (2014) Detection and analysis of chatter occurrence in micro-milling process. Proc Inst Mech Eng B J Eng Manuf 228(11):1359–1371
21. Zhang Z, Li H, Meng G, Tu X, Cheng C (2016) Chatter detection in milling process based on the energy entropy of VMD and WPD. Int J Mach Tool Manu 108:106–112
22. Attoui I, Fergani N, Boutasseta N, Oudjani B, Deliou A (2017) A new timefrequency method for identification and classification of ball bearing faults. J Sound Vib 397:241–265
23. Huang NE, Wu Z (2009) Ensemble empirical mode decomposition: a noise-assisted data analysis method. Adv Adapt Data Anal 01(01):1–41
24. Peng ZK, Tse PW, Chu FL (2005) An improved Hilbert–Huang transform and its application in vibration signal analysis. J Sound Vib 286(1–2):187–205
25. Lamraoui M, Barakat M, Thomas M, Badaoui ME (2013) Chatter detection in milling machines by neural network classification and feature selection. J Vib Control 21(7):1251–1266
26. Tansel IN, Wang X, Chen P, Yenilmez A, Ozcelik B (2006) Transformations in machining. Part 2. Evaluation of machining quality and detection of chatter in turning by using s-transformation. Int J Mach Tool Manu 46(1):43–50
27. Yao ZH, Li X, Chen ZC (2007) Prediction of cutting chatter based on hidden Markov model. Key Eng Mat 353-358:2712–2715
28. Teti R, Jemielniak K, O'Donnell G, Dornfeld D (2010) Advanced monitoring of machining operations. CIRP Ann Manuf Technol 59(2):717–739
29. Wang Y, Bo Q, Liu H, Hu L, Zhang H (2018) Mirror milling chatter identification using Q-factor and SVM. Int J Adv Manuf Technol 98(5):1163–1177
30. Burges CJC (1998) A tutorial on support vector machines for pattern recognition. Data Min Knowl Discov 2(2):121–167
31. Hsueh Y-W, Yang C-Y (2009) Tool breakage diagnosis in face milling by support vector machine. J Mater Process Technol 209(1):145–152
32. Chen Y, Li H, Hou L, Bu X (2018) Feature extraction using dominant frequency bands and time-frequency image analysis for chatter detection in milling. Precis Eng. online.
33. The MathWorks Inc Fast Fourier transform
34. Theodoridis S, Koutroumbas K (2009) Chapter 5 - feature selection. In: Pattern recognition, Fourth edn. Academic Press, Boston, pp 261–322
35. Theodoridis S, Pikrakis A, Koutroumbas K, Cavouras D (2010) CHAPTER 4 - feature selection. In: Introduction to pattern recognition. Academic Press, Boston, pp 107–135
36. Albregtsen F (2008) Statistical texture measures computed from gray level coocurrence matrices. Department of Informatics, University of Oslo