Feature extraction using dominant frequency bands and time-frequency image analysis for chatter detection in milling

Yun Chen[1]*, Huaizhong Li[2]*, Liang Hou[1†], and Xiangjian Bu[1]

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

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

Abstract

Chatter is a cause of low surface quality and productivity in milling and crucial features need to be extracted for accurate chatter detection and suppression. This paper introduces a novel feature extraction approach for chatter detection by using image analysis of dominant frequency bands from the short-time Fourier transform (STFT) spectrograms. In order to remove the environmental noises and highlight chatter related characteristics, dominant frequency bands with high energy are identified by applying the squared energy operator to the synthesized fast Fourier transform (FFT) spectrum. The time-frequency spectrogram of the vibration signal is divided into a set of grayscale sub-images according to the dominant frequency bands. Statistical image features are extracted from those sub-images to describe the machining condition and assessed in terms of their separability capabilities. The proposed feature extraction method is verified by using dry milling tests of titanium alloy Ti6Al4V and compared with two existing feature extraction techniques. The results indicate the efficiency of the time-frequency image features from dominant frequency bands for chatter detection and their better performance than the time domain features and wavelet-based features in terms of their separability capabilities.

* Yun Chen and Huaizhong Li contributed equally.

† Corresponding author: hliang@xmu.edu.cn.

Keywords: chatter, feature extraction, image analysis, dominant frequency bands, short time Fourier transform.

Nomenclature

a_e	Radial depth of cut
a_p	Axial depth of cut
d	Pixel distance for gray levels (i, j)
f	Frequency
f_t	Feed per tooth
F_0	Feature threshold
$F_1^n, F_2^n, F_3^n, F_4^n$	Mean, standard variance, skewness and kurtosis
$F_5^n, F_6^n, F_7^n, F_8^n$	Mean contrast, correlation, energy and homogeneity
$F_9^n, F_{10}^n, F_{11}^n, F_{12}^n$	Contrast, correlation, energy and homogeneity ranges
$F_{d,\theta}^{Con}, F_{d,\theta}^{Cor}, F_{d,\theta}^{E}, F_{d,\theta}^{H}$	Contrast, correlation, energy and homogeneity
G	Gray level
$GLCM_{d,\theta}^n(i, j)$	Entry of gray level co-occurrence matrix at gray levels (i, j)
$h_n(i)$	Intensity-level histogram for the nth sub-image
i, j	Gray levels
K	Number of vibration signals
m	Time
M, N	Sizes of the nth sub-image $S_n(m, f)$
n	Index for nth sub-image
n_p	Spindle speed
N_p	Number of data samples for one spindle rotation
$p_{d,\theta}^n(i, j)$	Second-order statistical probability at gray levels i and j
$p_n(i)$	Probability density of occurrence of gray level i
$p_{\omega 1}(F), p_{\omega 2}(F)$	Overlapping probability density functions of a feature F in classes ω_1 and ω_2
SEO	Squared energy operator
$S(m, f)$	Spectrogram of vibration signal

t	Time
$w_1(f)$	Smoothing window function for SEO
$w_2(t\text{-}m)$	a window function centered at time m
$\underline{x_k}(t)$	kth sample vibration signal
$X(m,f)$	STFT of vibration signal
$y(f)$	Synthesized FFT spectrum
$Y(f)$	Squared energy
$\alpha(F_0)$, $\beta(F_0)$	Probabilities for wrong decisions concerning classes ω_1 and ω_2
δ	the Kronecker delta function
θ	Pixel angle gray levels (i,j)
ω_1, ω_2	Class label

1 Introduction

Current advances towards productivity improvement and cost saving require accurate condition monitoring of the manufacturing process in order to decrease the amount of the unplanned downtime. In fully automated or lightly staffed machining environments, a requirement of accurate detection of the machining condition is significantly necessary. Chatter which could damage tools, machines or workpieces is a cause of unplanned downtime in a machining environment. Many manufacturers deal with chatter by setting machining parameters low, instead of tool strength and spindle horsepower defining the metal removal rate. Therefore, chatter becomes a limiting factor that keeps the process from reaching its potential. An accurate and reliable machining condition monitoring system is expected to reduce downtime by allowing maintenance to be scheduled in advance, and increase a cutting speed by 10–15% and save the cost by 10–40% [1].

In order to monitor the machining condition, crucial features are extracted from the physical signals, such as cutting forces [2], vibrations [3], sound [4] and acoustic emission [5, 6], etc. The extracted features play a key role in identifying the machining condition. Vela-Martinez et al. [7] applied the Hurst exponent to identify

the dynamical changes in milling operations. Two features, the energy ratio and the ratio of the mean and standard deviation, were proposed by Fu et al. [8] to describe the machining condition, and a Gaussian mixture model was applied to determine the thresholds of the two features for the stable and unstable tests. The above works adopted one or two feature(s) as scalar indicators to diagnose the machining condition. On the other hand, Chen et al. [5] obtained a better diagnosis in a high dimensional feature space from the frequency domain. It is noted that not all features can adequately reflect the dynamic changes in the machining conditions. A high dimensional space with irrelevant features affects the computation time. Salient features need to be selected from the original feature set. Separability capability ranking of features is a commonly-used technique for feature assessment and selection. Lamraoui et al. [9] adopted Kullback-Leibler distance to rank nine time-domain features extracted from vibration signals and selected four top-ranked features to develop a time-effective monitoring system. Boashash et al. [10] applied the area under the receiver operating characteristics curve to assess the performance of different feature sets for EEG seizure detection.

Due to inherent complexity and variability of machining mechanisms, the characteristics of the signals obtained in machining can be very complex, showing non-stationarity and a variety of uncertainties [11]. The effectiveness of a naive signal processing technique in a single domain, either the time domain or frequency domain, is not sufficient to provide reliable results for practical applications. The time-frequency methods can reflect the frequency variations of a non-stationary signal over time. Liu et al. [12] used a wavelet-based algorithm for chatter detection in grinding, and Li et al. [13] applied the wavelet packet transform to investigate chatter occurrence in the micro-milling process. Although wavelet-based methods perform well in detecting transitions of nonstationary data, they rely on the specified basis functions [14]. To overcome the dependence on the basis functions, Susanto et al. [3] used the Hilbert-Huang transform for chatter detection in end milling. Perez-Canales et al. [15] implemented a multi-scale approximate entropy analysis to identify the

randomness of the signals when chatter occurs.

Although the time-frequency transform is powerful to process the non-stationary signals, it alone is unable to quantify the machining condition and cannot be directly processed by classification methods for intelligent chatter detection. Additional scalar indicators or features must be extracted from the time-frequency analysis for quantification. Scalar features are usually extracted from the decomposed components of the time-frequency analysis [16, 17]. Although the decomposed components have narrower frequency bands than the original signal, features from those components are still based on one-dimensional space and may not sufficiently provide enough information for efficient chatter detection. On the other hand, the time-frequency spectrum can be treated as a two-dimensional image. Features can be extracted to describe different image characteristics between the stable and unstable machining conditions. 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 [18]. The most popular statistical image features are derived from the gray level co-occurrence matrix (GLCM). The GLCM assesses co-occurrence of gray levels of two pixels located at any given distance and angle in an image. Abnormalities that may be hidden in the traditional time domain can be discovered by the time-frequency image, and bring about variations of neighboring pixel distributions and relationships in the image. Thus, the GLCM that considers the relationships among pixels at different distances and angles is powerful to detect abnormalities in non-stationary signals [10]. The statistical image features have been used in the neurocomputing science [19] and scene segmentation [20]. In tool condition monitoring, Dutta et al. [21] employed statistical image features to machined surface images for on-machine tool prediction of flank wear.

This paper applies time-frequency image features from dominant frequency bands for chatter detection in milling. The dominant frequency bands with high energy are

identified to reduce the environmental noise. Image features from dominant frequency bands are applied to describe complex and non-stationary properties in STFT spectrograms. The area under the receiver operating characteristics curve is used for assessment, so that the most sensitive image features related to chatter can be identified. The proposed method is verified in milling and compared with two additional methods for chatter detection.

2 Methodology

2.1 Frequency band selection

The stability of milling operations highly depends on the dynamic behavior of the machine. When the milling process is unstable, dominant vibration frequencies or chatter frequencies with high energy are in most cases close to natural frequencies of the milling system [22]. Identification of dominant frequency bands related to natural frequencies can decrease the high level of environmental noises, and improve the sensitivities of features extracted from those bands. Lamraoui et al. [9] used resonance values of the frequency response function (FRF) to identify dominant frequency bands for chatter detection. The hammer impact test is a common approach to determine the FRF, but it requires additional sensors (e.g. hammer) for experiments. In order to avoid additional sensors, a synthesized FFT of vibration signals from all cases of machining conditions is defined in Eq. (1) to localize dominant frequency bands. Direct localization of dominant frequency bands from the synthesized FFT is difficult. Energy operators based on moving windows can smooth the synthesized FFT spectrum and magnify the local peaks, which help identification of dominant frequency bands. Candidate energy operators include Teager energy operator [23], nonlinear energy operator [24] and squared energy operator (SEO) [25], etc. This work uses the SEO that depends on only the window function, instead of the other two candidates that rely on both the window function and lag parameter.

Sample training vibration signals $\{x_k(t)\}$ (k=1, 2… K) from each case of machining

condition, either stable or unstable, are firstly analyzed by using FFT algorithm, and then summed up to obtain a synthesized FFT spectrum

$$y(f) = \sum_{k=0}^{K-1} \mathrm{FFT}[x_k(t)] \tag{1}$$

where $y(f)$ is the synthesized FFT spectrum, f is the frequency and FFT[$\underline{x_k}(t)$] is the FFT of signal $\underline{x_k}(t)$. The FFT algorithm is implemented using the function "fft" in Matlab [26]. In order to identify dominant frequency bands with high energy, the SEO is applied to process $y(f)$

$$Y(f) = E(y^2(f)) = y^2(f) \otimes w_1(f) \tag{2}$$

where $\otimes$ is the convolution operator, and $w_1(f)$ is a smoothing window function. The smoothed and squared energy in Eq. (2) magnifies the local frequency components with high energy, as the non-dominant frequency band has a significantly smaller energy value.

A better choice of window length can properly separate the dominant bands from non-dominant ones in the synthesized FFT spectrum. A small window length cannot sufficiently smooth the synthesized FFT spectrum, and a large window length may combine two closely-located dominant frequency bands. Therefore, the length of the smoothing window should be selected carefully. The proposed procedure for identification of frequency bands is based on the work given by Attoui et al. [27]. Instead of the short-frequency energy defined by Attoui et al., a more widely-used squared energy operator is employed for frequency band selection in this study. It is noted that Attoui et al. did not state the method for determining the upper and lower limits of frequency bands. Therefore, the cut-off frequencies that are commonly used in filter design are introduced to determine the limits of each frequency band. The cut-off frequencies are the frequencies where the squared energy $Y(f)$ drops to $1/\sqrt{2}$ of its local peak value.

2.2 Time-frequency image features

The time-frequency spectrogram by the STFT contains the information of the vibration signal in both time and frequency domains. The vibration signal is transformed by applying STFT

$$X(m,f)=\sum_{t=-\infty}^{\infty}x(t)w_2(t-m)\exp(-2\pi ftj) \tag{3}$$

where $x(t)$ is a vibration signal, and $w_2(t\text{-}m)$ is a window function centered at time m. The spectrogram of $x(t)$ is the squared magnitude of its STFT

$$S(m,f)=|X(m,f)|^2 \tag{4}$$

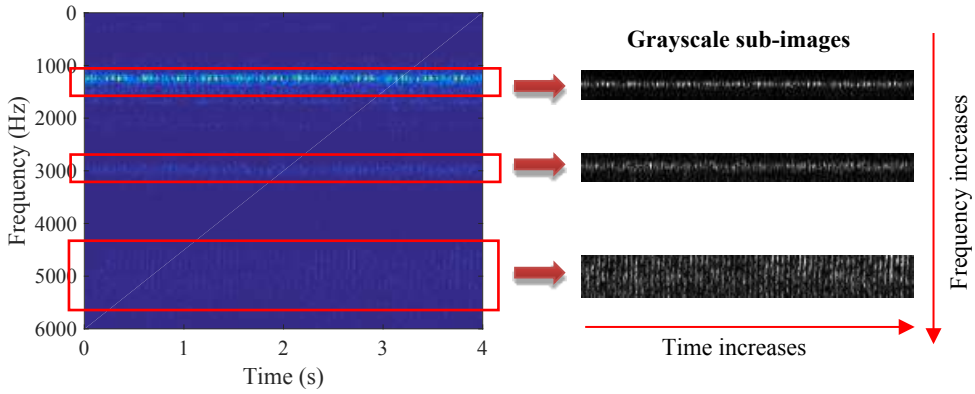

Fig. 1. Spectrogram and its grayscale sub-images.

The spectrogram is divided into several sub-images according to the dominant frequency bands, and converted to G gray-level sub-images, as shown in Fig. 1. Statistical approaches are used to extract the features from the grayscale sub-images. The first-order statistics are calculated from the gray-level histogram showing the number of pixels in each sub-image [18]. The gray-level histogram is defined as

$$h_n(i)=\sum_f\sum_m\delta(S_n(m,f),i) \qquad i=0,\cdots,G-1 \tag{5}$$

where i is a gray level, δ is the Kronecker delta function, and $S_n(m,f)$ is nth grayscale sub-image divided from the whole spectrogram $S(m, f)$. The approximate probability density of occurrence of the intensity levels is obtained as

$$p_n(i)=\frac{h_n(i)}{MN} \qquad i=0,\cdots,G-1 \tag{6}$$

where M and N are the sizes of the sub-image $S_n(m,f)$. The first-order statistics used

are listed in Table 1.

Table 1. First-order statistical image features.

Feature	Equation
Mean (F_1^n)	$\sum_{i=0}^{G-1} i p_n(i)$
Standard variance (F_2^n)	$\sqrt{\sum_{i=0}^{G-1}(i-F_1^n)^2 p_n(i)}$
Skewness (F_3^n)	$(F_2^n)^{-3}\sum_{i=0}^{G-1}(i-F_1^n)^3 p_n(i)$
Kurtosis (F_4^n)	$(F_2^n)^{-4}\sum_{i=0}^{G-1}(i-F_1^n)^4 p_n(i)-3$

The first-order statistics ignore the pixel neighborhood relationships. On the other hand, the gray level co-occurrence matrix (GLCM) estimates image properties related to second-order statistics that consider the relationship between pixels. GLCM is an estimate of the joint probability distributions of two pixels, a distance d apart along a given direction θ having co-occurring gray values i and j [28]

$$GLCM_{d,\theta}^{n}(i,j)=\begin{cases}\sum_{m}\sum_{f}1 & S_n(m,f)=i \;\&\; S_n(m+d\cos\theta, f+d\sin\theta)=j \\ 0 & others\end{cases} \tag{7}$$

The matrix element $p_{d,\theta}^{n}(i,j)$ is an approximate second-order statistical probability for changes between gray levels i and j at a particular pair (d, θ)

$$p_{d,\theta}^{n}(i,j)=\frac{GLCM_{d,\theta}^{n}(i,j)}{\sum_{j}\sum_{i}GLCM_{d,\theta}^{n}(i,j)} \tag{8}$$

The second-order statistical features (i.e. contrast, correlation, energy, and homogeneity) can be extracted from the GLCM

$$\text{Contrast:}\quad F_{d,\theta}^{Con}=\sum_{j=0}^{G-1}\sum_{i=0}^{G-1}(i-j)^2 p_{d,\theta}^{n}(i,j) \tag{9}$$

$$\text{Correlation:}\quad F_{d,\theta}^{Cor}=\sum_{j=0}^{G-1}\sum_{i=0}^{G-1}(ijp_{d,\theta}^{n}(i,j)-\mu_m\mu_f)/\sigma_m\sigma_f \tag{10}$$

$$\text{Energy:}\quad F_{d,\theta}^{E}=\sum_{j=0}^{G-1}\sum_{i=0}^{G-1}[p_{d,\theta}^{n}(i,j)]^2 \tag{11}$$

$$\text{Homogeneity:}\quad F_{d,\theta}^{H}=\sum_{j=0}^{G-1}\sum_{i=0}^{G-1}p_{d,\theta}^{n}(i,j)/1+(i-j)^{2} \tag{12}$$

where μ_m, μ_f, σ_m and σ_f are the means and standard deviations of the marginal probability matrix elements $p_{m,d,\theta}^{n}(i)$ and $p_{f,d,\theta}^{n}(j)$, respectively.

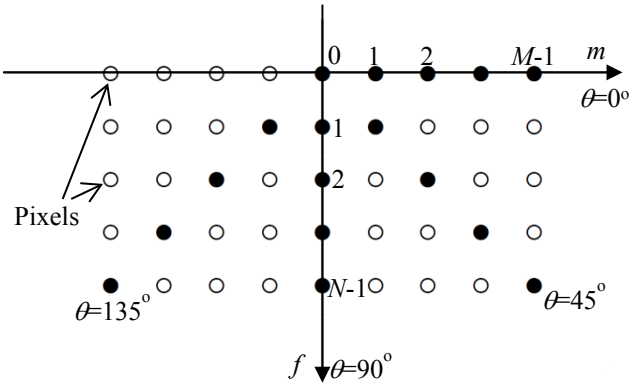

Fig. 2. GLCM construction methods at different angles.

Table 2. Second-order statistical image features.

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

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

Due to the intensive nature of computations involved, often the distances d = 1 and 2 pixels with angles $\{\theta_i\,|\,0^o,45^o,90^o,135^o\}$ are considered as suggested by Haralick et al. [29], as shown in Fig. 2. This study uses a distance d = 1 pixel, and the symmetric form of co-occurrence matrix. As four angles $\{\theta_i\,|\,0^o,45^o,90^o,135^o\}$ are considered, there are four sets of second-order statistical features. Instead of the whole four sets of second-order features, the mean and range of each type of a second-order feature are

used for chatter detection (Table 2).

2.3 Feature assessment

The first-order and second-order image features may differ in their ability to contain as much information to describe the machining condition. Thus, the receiver operating characteristics (ROC) curve and the area under this curve (AUC) are used to assess image features in terms of their separability capability [30]. Fig. 3(a) illustrates two overlapping probability density functions $p_{\omega 1}(F)$ and $p_{\omega 2}(F)$ describing the distribution of a single feature F in two classes ω_1 and ω_2, together with a threshold F_0. It is assumed that values on the left of the threshold belong to class ω_1 and values on the right belong to class ω_2. This results in two probabilities $\alpha(F_0)$ and $\beta(F_0)$ of reaching wrong decisions concerning class w_1 and class w_2, respectively

$$\alpha(F_0) = \int_{F_0}^{\infty} p_{\omega 1}(F), \quad \beta(F_0) = \int_{-\infty}^{F_0} p_{\omega 2}(F) \tag{13}$$

The ROC curve plots parametrically $1-\beta(F_0)$ versus $\alpha(F_0)$ with F_0 as the variable, as Fig. 3(b) demonstrates. If the two probability density functions have complete overlap, $\alpha(F_0)$ equals $1-\beta(F_0)$ for any threshold. As the two distributions move apart, the corresponding ROC curve departs from the straight line $\alpha(F_0) = 1-\beta(F_0)$, as shown in Fig. 3(b).

The AUC is defined as the shaded area between the ROC and the straight line. This area varies between zero for complete overlap, and 1/2 for complete separation. It is a measure of the class separability capability of the specific feature [10]. A high AUC implies that a feature is good in discriminating between two classes, whereas a low AUC implies that a feature fails to discriminate. Thus, the time-frequency image features are ranked according to their corresponding AUCs.

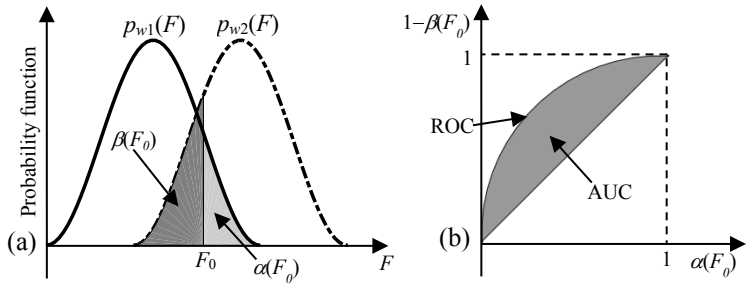

Fig. 3. Feature assessment (a) overlapping probability density functions of the same feature in two classes and (b) the resulting ROC curve and AUC.

2.4 Proposed chatter detection method

The performance of a chatter detection method depends on crucial features extracted from the measured signals. Noise reduction is critical to extract the most informative features related to chatter. The extracted feature type determines its sensitivity to chatter. As the time-frequency method is powerful to reflect non-stationary properties, image features are employed to describe the properties revealed in time-frequency images. In order to increase the signal-to-noise ratio, this paper employs the SEO for localization of dominant frequency bands that are used to pre-process STFT time-frequency images. Image analysis is used to extract features from time-frequency images, and the sensitivity of image features is assessed using the AUC. The procedure for the proposed method is summarized in Fig. 4.

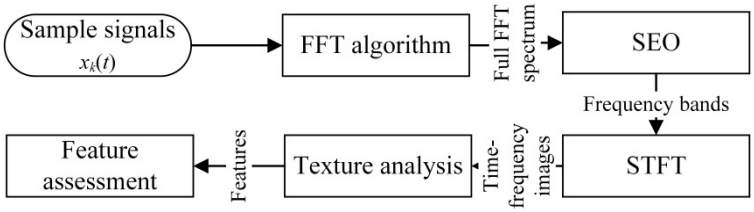

Fig. 4. Proposed chatter detection method.

3 Experimental results

3.1 Experimental setup

The proposed method was assessed in milling of titanium alloy Ti6Al4V under dry cutting conditions. The experimental setup and coordinate system of the vibration signals are shown in Fig. 5. The feed direction is the x direction. An accelerometer was mounted on the workpiece clamp to measure the vibration signals in the y direction. The sample rate for the tests was 20 kHz. Shoulder milling was performed throughout the tests. The cutter was a two-flute end mill with a diameter of 8 mm and a helix angle of 30°. The total number of experiments is 82 under spindle speeds n_p from 1100 to 7600 rpm, feed per tooth f_t = 0.08, 0.1, 0.12 mm/tooth, axial depth of cut a_p = 4, 8, 12 mm, and radial depth of cut a_e = 0.5, 1, 1.5mm.

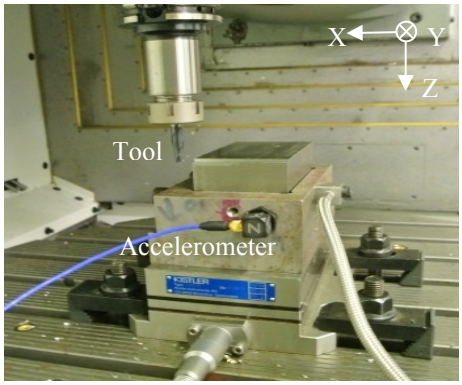

Fig. 5. Milling experimental setup.

3.2 Feature extraction and assessment

A set of training samples that include four tests for each case of machining condition, either stable or unstable, are summed to localize the dominant frequency bands. Two of the tests used for selection of the dominant frequency bands are shown in Fig. 6. The vibration signal for the stable test has a small amplitude variation, whereas the signal for the unstable test shows a significant variation. The Hanning window function is used for w_1 in Eq. (2). Fig. 7 shows the synthesized FFT spectrum and its squared energy under different window lengths. The squared energy for each window length is normalized using its maximum. A small window length (e.g. 50 and 200)

cannot clearly reveal the dominant frequency bands, whereas a large window length (e.g. 800 and 1,000) may combine two dominant frequency bands that are closely located. In this study, a window length between 300 and 700 is proper for localization of the dominant frequency bands.

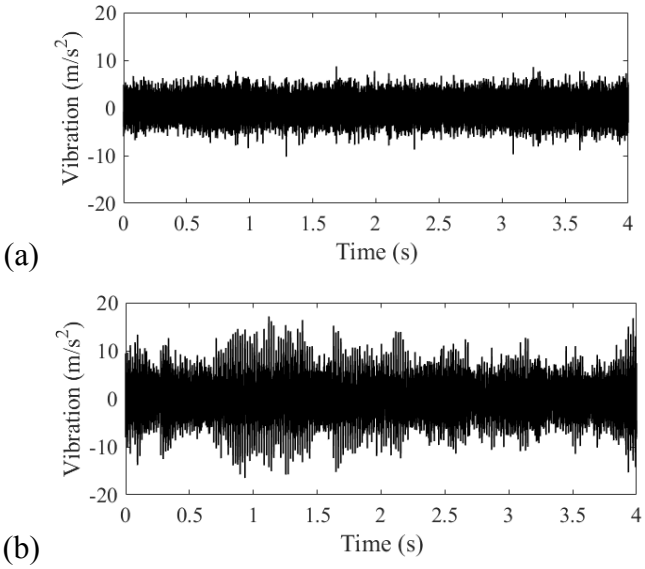

Fig. 6. Vibration signals (a) stable at n_p 2100 rpm, a_p 12 mm, f_t 0.1 mm/tooth and a_e 0.5 mm (b) unstable at n_p 1600 rpm, a_p 8 mm, f_t 0.1 mm/tooth and a_e 1 mm.

For further analysis, a window length of 400 is selected, and its corresponding normalized energy is shown in Fig. 8. Five dominant frequency bands are identified using the cut-off frequencies, which are located between 180 and 350 Hz, 1050 and 1450 Hz, 1650 and 1900 Hz, 2800 and 3050 Hz and 4800 and 5200 Hz. Table 3 lists the natural frequencies that are obtained by hammer impact testing. Detailed experimental setup and results for the hammer impact test are given in Ref. [31]. Interestingly, except the first frequency band 180-350 Hz, the left four bands cover or close to multiple natural frequencies. The first frequency band may be formed by the high energy of the tooth passing frequencies of the selected tests. Because of such a relationship between the dominant frequency bands and natural frequencies, Lamraoui et al. [9] have used the FRF from hammer impact testing to identify the dominant frequency bands. The squared energy provides an alternative method to identify dominant frequency bands, and no hammer impact experiments are required.

Although the cut-off frequencies are determined by the half power method, the width of each frequency band can be narrowed or broadened, which do not significantly affect the AUCs of the most informative image features.

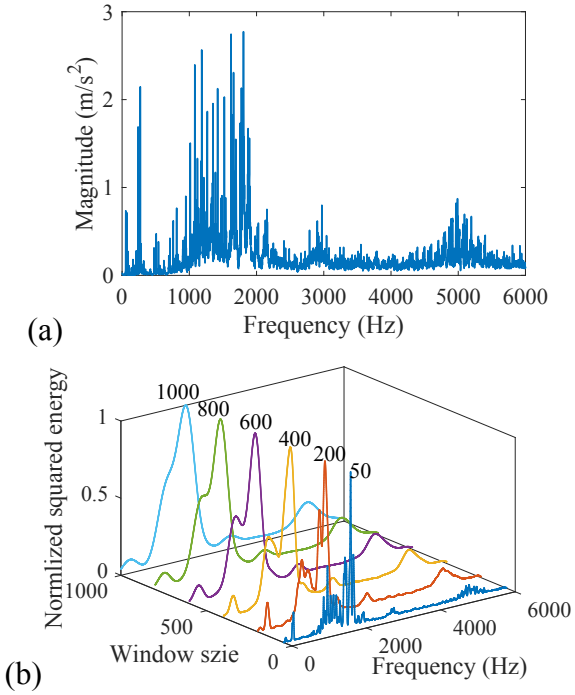

Fig. 7. Effect of window length (a) synthesized FFT spectrum (b) normalized squared energy with different window lengths.

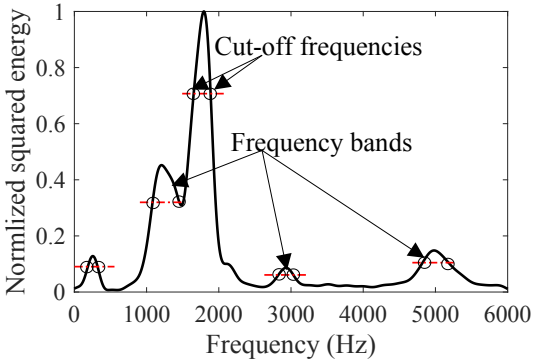

Fig. 8. Normalized squared energy with window length 400.

Table 3. Natural frequencies for the workpiece and tool [31].

	Direction	Natural frequency (Hz)
Workpiece	*x*	1107
	y	1220
	z	1111, 2945
Tool	*x*	1355, 1900, 2780, 4912
	y	1308, 1875, 2719, 5011

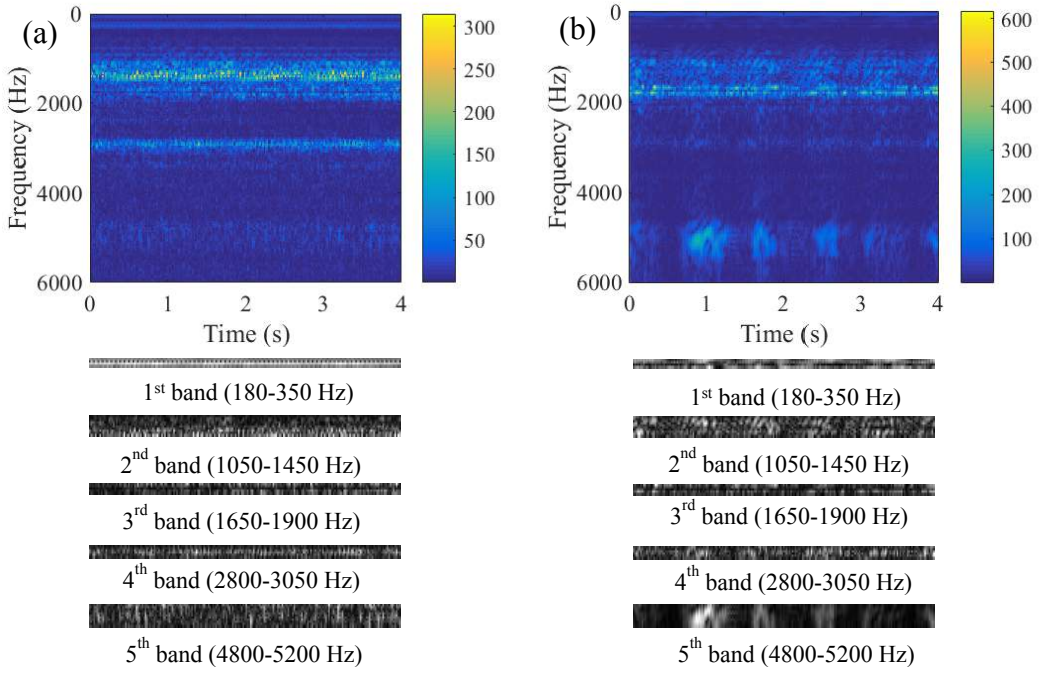

Fig. 9. Spectrograms of tests in Fig. 6 (a) stable (b) unstable.

The spectrograms and sub-images for the tests in Fig. 6 are presented in Fig. 9. A sensitive study is conducted before determining the proper length for the window function w_2. The investigated window length ranges from 2^7 to 2^{10}. A window length of 2^9 is finally selected as it balances the time and frequency resolutions of the STFT spectrograms and gives the largest AUCs for the top ranked features. The window function w_2 for STFT has a length of 2^9. Sub-images are firstly extracted from the spectrograms according to the dominant frequency bands and then converted into grayscale sub-images. A sensitivity analysis is conducted to choose the total number G of gray levels. It is found that all the image feature values become stable when $G >= 2^8$. Thus, the total number $G=2^8$ is used to normalize each sub-image into a range from 0 to 2^8-1. As aforementioned, chatter is related to energy rise around specific natural frequencies. The fifth frequency band extracted from the synthesized FFT spectrum automatically covers two natural frequencies 4912 and 5011 Hz in the x and y directions, respectively. When the test becomes unstable, pixel intensities rise in the fifth band. This indicates that the two natural frequencies are important for chatter occurrence in the investigated milling system.

The image features are extracted from the grayscale sub-images, instead of the whole spectrograms. The AUC is used to rank the separability capability for each feature, as shown in Fig. 10. The results show large AUCs for most features from the fifth band. The two features from the fifth band, mean correlation F_6^5 and correlation range F_{10}^5, have large AUCs of 0.45 and 0.42, respectively, and therefore high capabilities to separate the stable and unstable tests. Obviously, it is the big difference of the fifth band between the stable and unstable tests that results in large AUCs for image features from this band. Those results evidence the efficiency of image features to describe characteristics implied by time-frequency images. Separability capabilities of different features are compared in Fig. 11. In Fig. 11(a), the top two ranked features define a two-dimensional space and the location of each test in this space is fully determined by its feature values. Similarly, Fig. 11(b) gives the test distributions in the space defined by the last-two ranked features. The test distributions in the top two feature space show much less overlapping than those in the last two feature space. A less overlapping distribution usually implies better pattern recognition performance for intelligent chatter detection.

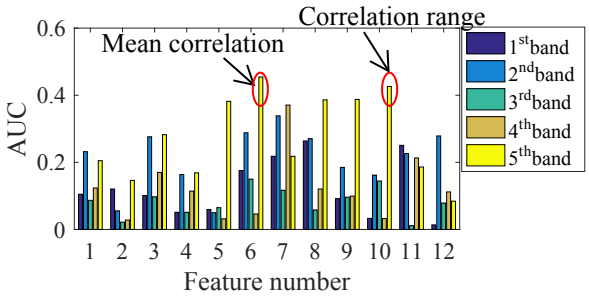

Fig. 10. AUC for each image feature from sub-images.

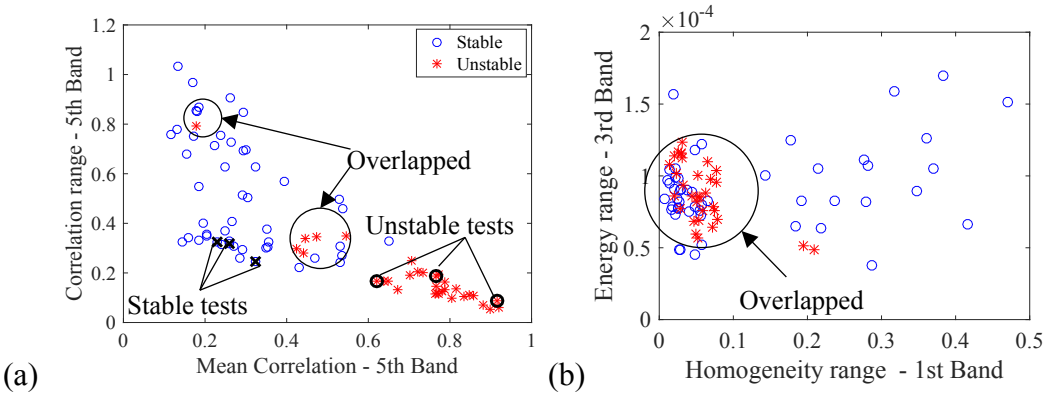

Fig. 11. Two-dimensional feature spaces (a) top two ranked features with high AUCs (b) last two features with low AUCs.

Table 4. Cutting parameters for selected six tests.

Test #	Condition	n_p rpm	f_t mm/tooth	a_p mm	a_e mm
1	Unstable	1600	0.1	8	1
2	Unstable	2100	0.1	8	1
3	Unstable	2600	0.1	8	1
4	Stable	6100	0.1	12	1
5	Stable	6600	0.1	12	1
6	Stable	7100	0.1	12	1

Six tests in Table 4 are used to examine the capacity of the proposed method to distinguish the forced vibration condition with high level of vibration in respect to chatter condition. The unstable tests 4-6 are with large spindle speeds and high level of vibration due to forced vibration. Fig. 12 compares the machined surfaces, time domain signals and spectrograms for tests 3 and 6. The stable test 6 had a high level of vibration and showed no chatter marks. The spectrogram for test 3 shows abnormalities in the fifth frequency band. The six tests are highlighted in Fig. 11(a). The three stable tests are represented with black "x", and the three unstable tests with black "o". Those stable and unstable tests are well separated in Fig. 11(a). As the sub-images are normalized in the range from 0 to 2^8-1, features from the normalized sub-images are little related with the absolute level of vibration, but depend on the time-frequency image patterns or relative gray levels of pixels. Hence, the proposed method is efficient to distinguish forced vibration conditions from chatter conditions.

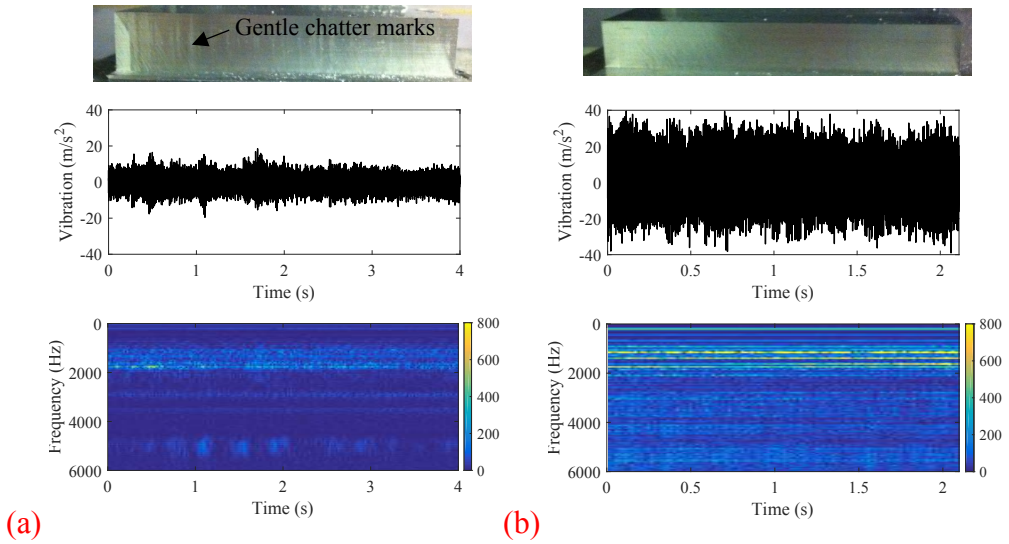

Fig. 12. Machined surfaces, time signals and spectrograms for (a) test 3 (b) test 6 in Table 4.

Three typical tests in Table 5 are selected in order to further examine the top two features on qualification of the time-frequency images. The machined surfaces, vibration signals, spectrograms and grayscale sub-images are shown in Fig. 13. The amplitudes of the vibration signals increase and become chaotic, as the surface quality decreases from test 1 to test 3 or the severity of chatter increases. Different image patterns, especially in the fifth frequency band, are reflected on the spectrograms and sub-images. The top two features, mean correlation F_6^5 and correlation range F_{10}^5, either decrease or increase with the increase of the severity of chatter, as listed in Table 6. As the correlation $F_{d,\theta}^{Cor}$ defined in Eq.(10) is a measure of gray level linear dependence between the neighboring pixels at the specified directions [28]. A large correlation means highly correlated neighboring pixels, and vice versa. The increase of the mean correlation F_6^5 and the decrease of the correlation range F_{10}^5 imply a high correlation among the neighboring pixels over the four directions $\{0^o,45^o,90^o,135^o\}$. In both Fig. 9 and Fig. 13, the pixel intensities of the fifth frequency band in the stable tests are relatively randomly distributed, whereas the high-intensity pixels of the fifth frequency band are gathered up due to chatter. This

explains why F_6^5 increases and F_{10}^5 decreases when chatter occurs. On the other hand, the last two features, homogeneity range F_{12}^1 and energy range F_{11}^3, are not proportional to the severity of chatter, which cannot fully quantifying the stability of machining operations.

Table 5. Cutting parameters for three typical tests.

Test #	Condition	n_p rpm	f_t mm/tooth	a_p mm	a_e mm
1	Stable	1100	0.10	4	1.0
2	Unstable	2100	0.10	8	1.0
3	Unstable	2600	0.12	12	1.0

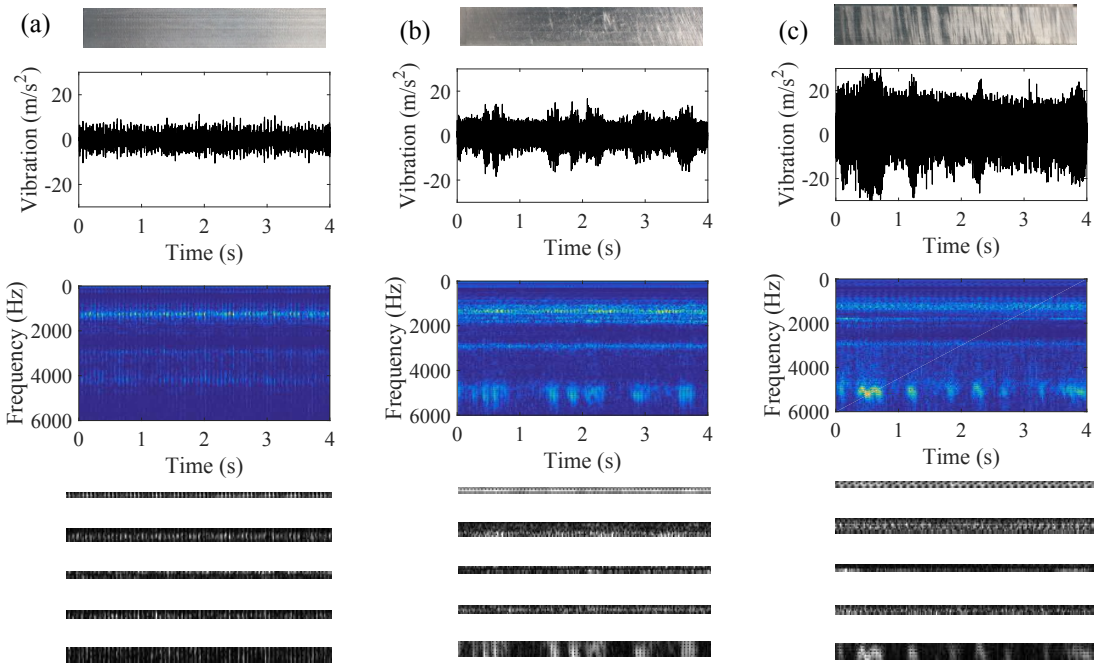

Fig. 13. Machined surface, vibration signals, spectrograms and sub-images (a) test 1 (b) test 2 (c) test 3 in Table 5.

Table 6. Top two and last two features for tests 1-3.

Test #	Top two features		Last two features	
	F_6^5	F_{10}^5	F_{12}^1	F_{11}^3
1	0.13	1.03	0.05	3.28e-04
2	0.72	0.20	0.08	2.91e-05
3	0.76	0.12	0.02	6.33e-05

Two additional tests in Fig. 14 are presented to investigate the capability of the mean correlation F_6^5 and correlation range F_{10}^5 as potential indices for online chatter detection. A moving hanning window with a length of 5,000 is applied to the whole vibration signals to extract data samples for STFT, and each pair of successive windows has an overlap of 5,000-N_p (N_p is the number of data samples for one spindle rotation) data samples. The corresponding F_6^5 and F_{10}^5 are given in Fig. 14. It can be seen that the unstable test has a larger mean correlation F_6^5 than the stable one, and also a smaller correlation range F_{10}^5. Zoomed views of the two tests are given in Fig. 15. The vibration signal for the unstable test shows significant amplitude variations, implying a highly complex and non-stationary milling process. The features F_6^5 and F_{10}^5 sufficiently track the amplitude variations. As STFT is a relatively fast time-frequency analysis method, the calculation of a single pair F_6^5 and F_{10}^5 over 5,000 data samples takes about 0.03 s using Matlab on a desktop computer with CPU 2.8 GHz and RAM 8GB. Therefore, the pair of features can be used as chatter indices for real-time condition monitoring. The current discussion of image features for chatter detection is limited in a laboratory environment, and more efforts are required to explore them in practical applications.

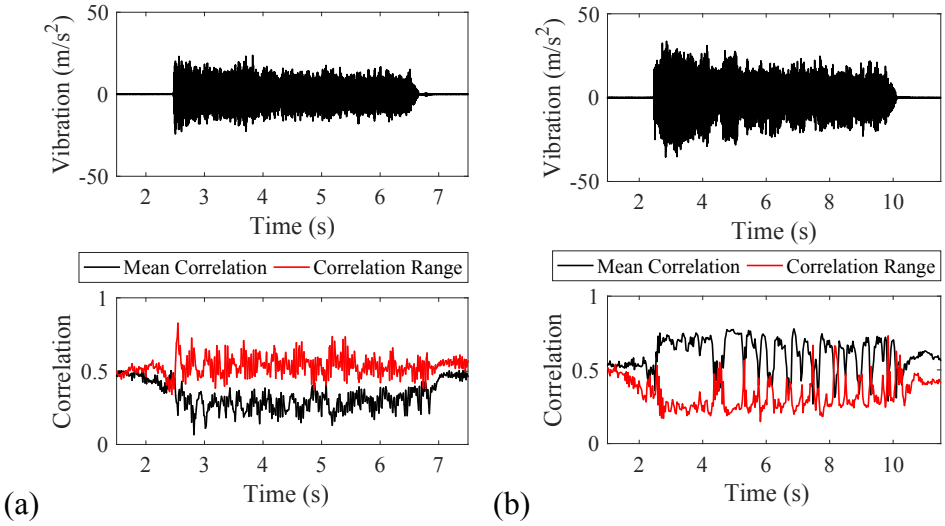

Fig. 14. Vibration signals and corresponding mean correlation F_6^5 and correlation range F_{10}^5

(a) stable at n_p 4600 rpm, a_p 8 mm, f_t 0.1 mm/tooth and a_e 1 mm (b) unstable at n_p 2100 rpm, a_p 12 mm, f_t 0.1 mm/tooth and a_e 1 mm.

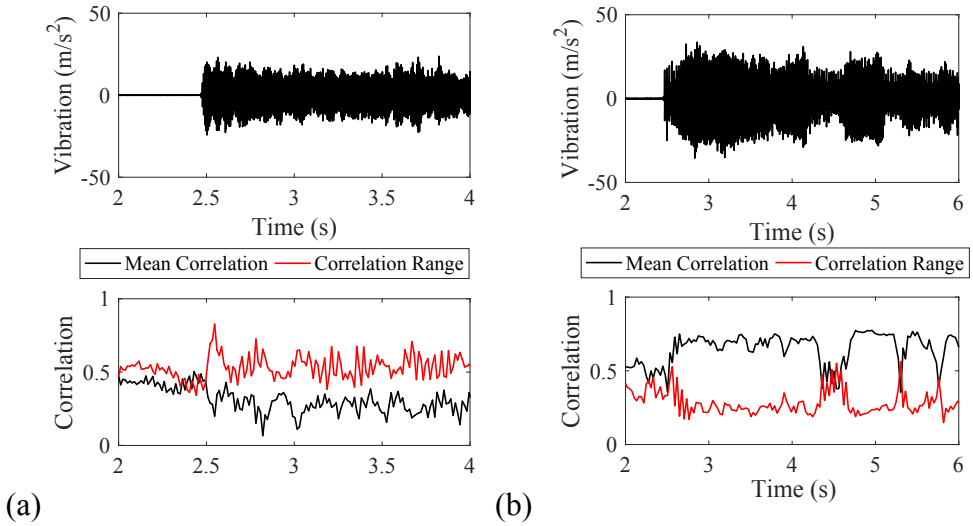

Fig. 15. Zoomed views of (a) stable test and (b) unstable test in Fig. 14.

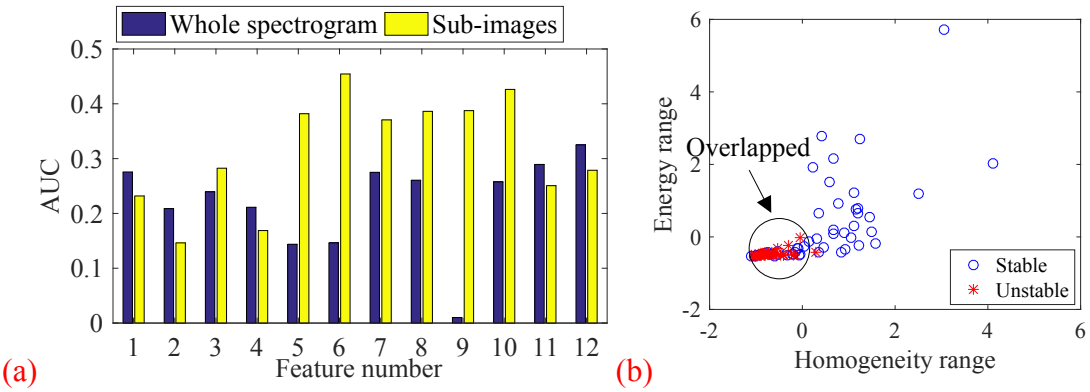

Fig. 16. Performance of image features from the whole spectrograms (a) AUCs for each feature (b) test distributions in the top two ranked feature space.

3.3 Effect of dominant frequency bands

In order to further validate the proposed method, the image features from the sub-images and from the whole spectrograms are compared using the same feature assessment method. The AUC for each feature from the whole spectrogram is shown in Fig. 16(a). For comparison, the largest AUC among the five sub-images or dominant frequency bands for each type of feature is also given in Fig. 16(a). The

feature, homogeneity range F_{12}^{1} from the whole spectrogram, has a maximum AUC of only 0.32, which is lower than the maximum AUC 0.45 for the mean correlation F_{6}^{5} from the sub-images. To further examine the separability capabilities of the image features from the whole spectrogram, Fig. 16(b) shows the stable and unstable tests in the space defined by the top two ranked features from the whole spectrogram. The stable and unstable tests show significant overlapping.

3.4 Comparison

The proposed feature extraction method is compared with the method proposed by Lamraoui et al. [9]. In Lamraoui's approach, a set of multiband resonance filters based on the dominant frequency bands are designed to pre-process vibration signals. The time-domain (TD) features in Table 7 are extracted from the envelopes of the pre-processed signals. Fig. 17(a) shows the AUC for each time-domain feature. The root mean square F_{t3} has the maximum AUC of 0.43, which is lower than the maximum in Fig. 10 but higher than that in Fig. 17(a). Similar results can be observed in Fig. 17(b) that gives the test distributions in the top two ranked feature space, which also shows a relatively large overlapping area.

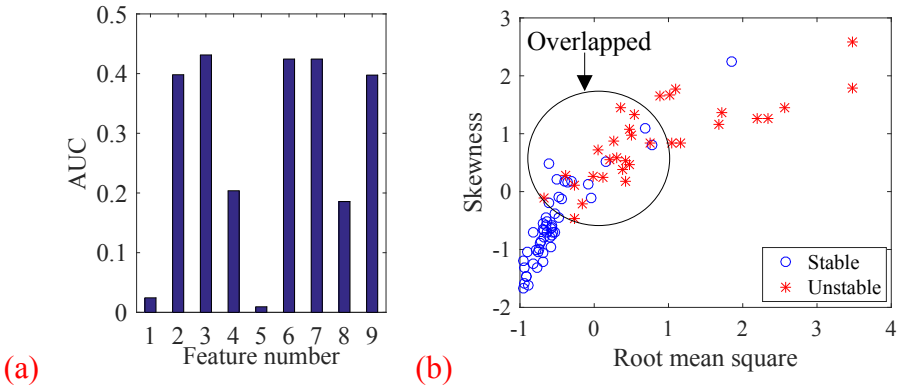

Fig. 17. Performance of time domain features (a) AUCs for each time domain feature (b) test distributions in the top two ranked feature space.

Table 7. Time domain features used by Lamraoui et al. [9].

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

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

A feature extraction method based on wavelet transform is also used for comparison. Yao et al. [16] extracted two wavelet-based features, standard variation (T1) and energy ratio (T2), from the most sensitive component obtained by using discrete wavelet and wavelet package decomposition, respectively. Similar to Yao's work, 3-level decomposition and wavelet db10 are used to decompose the signals. The sensitive component is selected according to the fifth frequency band (4800 and 5200 Hz). The AUCs for T1 and T2 are 0.24 and 0.38, respectively, indicating their low separability capabilities. Hence, the stable and unstable tests show a significant overlap in Fig. 18.

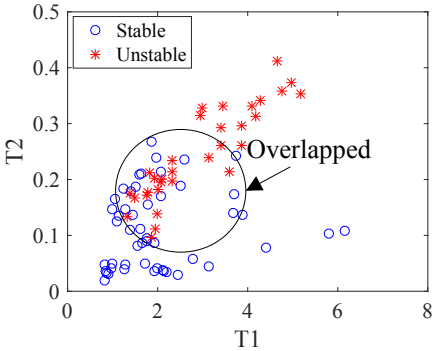

Fig. 18. Test distributions in the space defined by wavelet-based features T1 and T2.

A significant overlap between the stable and unstable tests in the feature space usually

means a poor classification performance for intelligent chatter detection. Although both time domain features and wavelet-based features are extracted from components with frequency bandwidths narrower than the raw signal, the two types of features are still based on one-dimensional components. On the other hand, image features represent the relationships between neighboring pixels in the time-frequency image, and therefore have a better performance in discriminating the stable and unstable tests.

4 Conclusions

This paper applies time-frequency image features from dominant frequency bands for chatter detection. The time-frequency analysis is powerful to discover the non-stationary properties of vibration signals in milling. Scalar features are needed to directly quantify the properties revealed by the time-frequency analysis. This study treats STFT spectrograms as images and applies image features from those images for quantification of machining condition. The dominant frequency bands with high energy are identified by the SEO and used to pre-process the STFT spectrograms, which increases the signal-to-noise ratio and the sensitivity of generated image features to chatter. In order to identify the most sensitive image features, the separability capability for each feature is assessed using the AUC.

The proposed feature extraction method is verified under various cutting parameters in milling. When chatter occurs, energy rises at chatter frequencies that are close to certain natural frequencies of the milling system. In this study, the fifth dominant frequency band that covers two natural frequencies 4912 and 5011 Hz in the x and y directions respectively shows a big difference in the energy distributions between the stable and unstable tests. Thus, most image features from the fifth dominant frequency band have large AUCs or high separability capabilities. The two features, mean correlation F_6^5 and correlation range F_{10}^5 from the fifth band, have the largest AUCs

and high capabilities to separate the stable and unstable tests. F_6^5 increases and F_{10}^5 decreases with the severity of chatter, which imply an increasing correlation among the neighboring pixels over the considered four directions $\{0^o,45^o,90^o,135^o\}$ in the fifth band. Correspondingly, the pixel intensity in the fifth band changes from a relatively random to concentrated distribution when the test becomes unstable. Therefore, F_6^5 and F_{10}^5 can well quantify the stability of machining condition. For further validation, the proposed method is compared with two existing feature extraction methods. The results show that image features give a better performance than time domain features and wavelet-based features. This work indicates that the proposed method is efficient for discriminating between the stable and unstable tests, and will be further applied for intelligent chatter detection in the next paper.

5 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 acknowledgments to the Advanced Manufacturing Laboratory, UNSW, for the support of the experimental work. Comments and suggestions from reviewers are greatly appreciated.

References

[1] Lamraoui M, Thomas M, Badaoui ME, Girardin F. Indicators for monitoring chatter in milling based on instantaneous angular speeds. Mechanical Systems & Signal Processing. 2014;44:72-85.

[2] Xu J, Yamada K, Seikiya K, Tanaka R, Yamane Y. Effect of different features to drill-wear prediction with back propagation neural network. Precision Engineering. 2014;38:791-8.

[3] Susanto A, Liu C-H, Yamada K, Hwang Y-R, Tanaka R, Sekiya K. Application of Hilbert–Huang transform for vibration signal analysis in end-milling. Precision Engineering. 2018;53:263-77.

[4] Zhang X, Chen H, Xu J, Song X, Wang J, Chen X. A novel sound-based belt condition monitoring method for robotic grinding using optimally pruned

extreme learning machine. Journal of Materials Processing Technology. 2018;260:9-19.
[5] Chen J, Chen H, Xu J, Wang J, Zhang X, Chen X. Acoustic signal-based tool condition monitoring in belt grinding of nickel-based superalloys using RF classifier and MLR algorithm. The International Journal of Advanced Manufacturing Technology. 2018;98:859-72.
[6] Zeng H, Chen X. Acoustic emission sensing and signal processing for machining monitoring and control. Advanced Automation Techniques in Adaptive Material Processing: World Scientific, 2002. p. 91-124.
[7] Vela-Martínez L, Jáuregui-Correa JC, Álvarez-Ramírez J. Characterization of machining chattering dynamics: An R/S scaling analysis approach. International Journal of Machine Tools and Manufacture. 2009;49:832-42.
[8] Fu Y, Zhang Y, Zhou H, Li D, Liu H, Qiao H, et al. Timely online chatter detection in end milling process. Mechanical Systems and Signal Processing. 2016;75:668-88.
[9] Lamraoui M, Barakat M, Thomas M, Badaoui ME. Chatter detection in milling machines by neural network classification and feature selection. Journal of Vibration and Control. 2013;21:1251-66.
[10] Boashash B, Barki H, Ouelha S. Performance evaluation of Time-Frequency image feature sets for improved Classification and Analysis of Non-Stationary Signals: Application to Newborn EEG Seizure Detection. Knowledge-Based Systems. 2017.
[11] Peng Y. Empirical Model Decomposition Based Time-Frequency Analysis for the Effective Detection of Tool Breakage. Journal of Manufacturing Science and Engineering. 2004;128:154-66.
[12] Liu Y, Wang X, Lin J, Zhao W. Correlation analysis of motor current and chatter vibration in grinding using complex continuous wavelet coherence. Measurement Science and Technology. 2016;27:115106.
[13] Li H, Jing X, Wang J. Detection and analysis of chatter occurrence in micro-milling process. Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture. 2014;228:1359-71.
[14] Vela-Martínez L, Carlos Jauregui-Correa J, Rodriguez E, Alvarez-Ramirez J. Using detrended fluctuation analysis to monitor chattering in cutter tool machines. International Journal of Machine Tools and Manufacture. 2010;50:651-7.
[15] Pérez-Canales D, Álvarez-Ramírez J, Jáuregui-Correa JC, Vela-Martínez L, Herrera-Ruiz G. Identification of dynamic instabilities in machining process using the approximate entropy method. International Journal of Machine Tools and Manufacture. 2011;51:556-64.
[16] Yao Z, Mei D, Chen Z. On-line chatter detection and identification based on wavelet and support vector machine. Journal of Materials Processing Technology. 2010;210:713-9.
[17] Cao H, Zhou K, Chen X. Chatter identification in end milling process based on EEMD and nonlinear dimensionless indicators. International Journal of Machine

Tools and Manufacture. 2015;92:52-9.
[18] Materka A, Strzelecki M. Texture analysis methods–a review. Technical University of Lodz, Institute of Electronics, COST B11 report, Brussels1998. p. 9-11.
[19] Alcn OF, Siuly S, Bajaj V, Guo Y, Sengur A, Zhang Y. Multi-category EEG signal classification developing time-frequency texture features based Fisher Vector encoding method. Neurocomputing. 2016;218:251-8.
[20] Huang X, Zhang L. An SVM Ensemble Approach Combining Spectral, Structural, and Semantic Features for the Classification of High-Resolution Remotely Sensed Imagery. IEEE Transactions on Geoscience and Remote Sensing. 2013;51:257-72.
[21] Dutta S, Pal SK, Sen R. On-machine tool prediction of flank wear from machined surface images using texture analyses and support vector regression. Precision Engineering. 2016;43:34-42.
[22] Insperger T, Stépán G, Bayly PV, Mann BP. Multiple chatter frequencies in milling processes. Journal of Sound and Vibration. 2003;262:333-45.
[23] Kaiser JF. Some useful properties of Teager's energy operators. 1993 IEEE International Conference on Acoustics, Speech, and Signal Processing1993. p. 149-52 vol.3.
[24] Mukhopadhyay S, Ray GC. A new interpretation of nonlinear energy operator and its efficacy in spike detection. Biomedical Engineering, IEEE Transactions on. 1998;45:180-7.
[25] Dimitriadis D, Potamianos A, Maragos P. A Comparison of the Squared Energy and Teager-Kaiser Operators for Short-Term Energy Estimation in Additive Noise. IEEE Transactions on Signal Processing. 2009;57:2569-81.
[26] The MathWorks Inc. Fast Fourier transform.
[27] Attoui I, Fergani N, Boutasseta N, Oudjani B, Deliou A. A new timefrequency method for identification and classification of ball bearing faults. Journal of Sound and Vibration. 2017;397:241-65.
[28] Albregtsen F. Statistical Texture Measures Computed from Gray Level Coocurrence Matrices. Department of Informatics, University of Oslo2008.
[29] Haralick RM, Shanmugam K, Dinstein I. Textural Features for Image Classification. IEEE Transactions on Systems, Man, and Cybernetics. 1973;6:610-21.
[30] Theodoridis S, Koutroumbas K. Chapter 5 - Feature Selection. Pattern Recognition (Fourth Edition). Boston: Academic Press, 2009. p. 261-322.
[31] Chen Y, Li H, Hou L, Wang J, Bu X. An intelligent chatter detection method based on EEMD and feature selection with multi-channel vibration signals. Measurement. 2018;127:356-65.