Functionalization and prediction of end milling surface topography based on a quantitative indicator of chatter and forced vibration

Feng Feng[a], Zhen Zhu[a], Meng Yuan[a], Kai Zhou[a], Julian Blumberg[b], Enlai Jiang [a,c], Eckart Uhlmann[b], Pingfa Feng[a,d],*

[a] *Shenzhen International Graduate School, Tsinghua University, Shenzhen 518055, China*
[b] *Institute for Machine Tools and Factory Management, Technische Universität Berlin, Berlin D-10587, Germany*
[c] *Xiamen Golden Egret Special Alloy Co., Ltd., Xiamen 361006, China*
[d] *Department of Mechanical Engineering, Tsinghua University, Beijing 100084, China*

Correspondence (P. Feng*): E-mail: fengpf@tsinghua.edu.cn; Tel./Fax: +86-755-26036356.

Abstract

Both chatter and forced vibrations can significantly degrade surface quality in milling. However, the traditional assessment method base on Ra cannot distinguish the contribution of the two types of vibrations, leading to a lack of guidance for specific vibration control. In this study, a novel vibration energy indicator that captures Ra variations and quantifies the contributions of chatter and forced vibrations to surface quality is proposed. Afterwards, the milled surface topography is reconstructed from acceleration signals collected during end milling. A one-dimensional convolutional neural network (1D CNN) processes acceleration data to predict three spatial parameters (amplitude A, spatial frequency ξ, and superposition angle φ), which are then used to reconstruct the milled surface. 1160 surface topography photos scanned by white-light interferometer are employed to develop the 1D CNN model (800 for training and 360 for testing, respectively). The reconstructed surfaces can closely match real surfaces, with predicted Ra values achieving an error less than 0.5 μm. The total vibration energy indicator correlates strongly with Ra^2 ($r = 0.98$), confirming its effectiveness for surface quality evaluation. Analyzing chatter and forced vibration energy distributions across Ra^2 values reveals that for common roughness (typically, Ra below 1.6 μm), both vibration types coexist or alternate, indicating a demand for targeted suppression strategies beyond simple spindle speed or feed rate adjustments. The accuracies of both the indicator and reconstructed surfaces with real-time computing capabilities imply a promising feasibility to realize a framework for integrating online surface quality assessment into digital twin systems for manufacturing.

Keywords: Milling chatter; Forced vibration; Surface topography; Vibration energy; Acceleration signal; Convolutional neural network.

1 Introduction

Milling processes are heavily influenced by vibrations, which significantly impact surface topography and quality. Two main types of vibrations dominate surface topography marks: forced vibrations, which produce predictable patterns at the tooth passing frequency, and chatter, which arises from regenerative effects and causes unmitigated, irregular surface topography [1,2]. Effective vibration control is key to maintaining high surface quality, with precise quantification of vibration effects enabling better suppression strategies, especially in developing real-time digital twin systems.

Traditionally, the relationship between vibration and surface quality is oversimplified: chatter is associated with poor surface quality, while forced vibrations are often overlooked. Tools like stability lobe diagram (SLD) [3–6] and chatter detection techniques [7–9] focus on chatter occurrence but fail to quantify the relationship between vibration intensity and surface quality, as well as the complexities of surface formation, making them inadequate for assessing workpiece quality when both chatter and forced vibrations are present. For example, in milling thin-walled aluminum engine cylinder heads, acceptable surface roughness ($Ra < 1.6$ μm) can be achieved despite chatter, while forced vibrations alone can lead to unacceptable roughness.

To link vibration with surface quality, analyzing surface topography is essential, as it serves as the "fingerprint" of the cutting tool's motion. Surface topography, affected by factors such as rigid-body kinematics [10], spindle runout [11,12], tool and workpiece flexibility[11,13], and surface plastic deformation [14], has been extensively studied in terms of both its generation mechanisms and prediction [15]. While detailed mechanism analysis offers theoretical insights, these methods rely on complex mathematical models and iterative calculations, making them impractical for real-time applications. A simplified spectral-domain analysis of surface topography is more suitable for online scenarios. Stepan et al. [16] applied two-dimensional fast Fourier transform (2D FFT) to differentiate stable and unstable cutting but did not fully investigate chatter patterns and spatial frequencies. Grossi et al. [17] provided an efficient method to classify chatter and forced vibration marks on surfaces from a spectral perspective, though it was only validated on a limited set of experiments.

With analysis methods in place, one key challenge is obtaining or reconstructing surface topography from measured process quantities. Optical metrology system

provides the most accurate reconstruction, as demonstrated by Shao et al. [18], who developed a support vector regression model using high-definition metrology data, with prediction times between 1.5s and 4.3s. However, it is unsuitable for industrial use due to interference from cutting fluid, chip flow, and fast-moving workbenches. Vibration signals, such as acceleration and force, are easier to collect online. Liu et al. [19] used a simplified geometric cutting path and sine vibration analytical model to reconstruct surface topography online with force and displacement signals, achieving relative mean errors of 9.6% for surface roughness in peripheral milling, though tested only on seven cuts and applied solely to stable cutting. To encompass a wide range of vibration scenarios, addressing the nonlinear relationship between 1D time-series vibration signals and 2D spatial topography is essential. Data-driven methods address this. Cao et al. [20] employed a knowledge-embedded sparse Bayesian regression model to predict machining errors, achieving a prediction accuracy with a global RMSE of 8.387 μm, an R^2 of 0.949, and prediction times under 10 ms, though the generalization was limited by a small dataset of 18 surfaces at a fixed spindle speed of 6950 rpm. Neural networks and deep learning techniques, such as CNN, LSTM, RNN, and GAN are more capable of handling nonlinear mapping but have mostly focused on surface roughness prediction [21–24], tool wear [25], surface morphology variations [26] and fault diagnosis [27]. Only Cao et al. [28] developed a CNN-GAN model for online surface topography prediction using cutting forces and vibration signals, but provided only qualitative results. Furthermore, embedding mechanisms in data-driven models have proven effective for predicting machining errors in thin-walled workpieces [20]. However, few works have explored the integration of deep learning with generation mechanisms for online surface topography reconstruction.

Another key challenge is quantifying vibration's contribution to surface quality after online surface topography reconstruction. Typical evaluation metric surface roughness *Ra* only provides roughness information without vibration intensity and source analysis, offering limited real-time guidance for targeted suppression and quality improvement. Moreover, using an indicator based on full 3D surface data rather than profile-based *Ra* is crucial for gaining deeper insights. Kumar et al. [29] introduced a chatter index from denoised signals to quantify chatter severity, achieving a maximum chatter index of 1.965 and an average model prediction accuracy of 93.29%, but as the index is computed from signals rather than surface data, it may not be fully accurate and reliable. Zhou et al. [30] proposed an algorithm for chatter quantification and milling instability

classification, achieving 100% accuracy on 107 milled surfaces with an 8.1 ms processing time, but its reliance on optical imaging limits practical use due to fluid and chip interference. Therefore, computing a vibration quantitative indicator from surfaces reconstructed from vibration signals during milling is a promising approach to ensure both accuracy and real-time monitoring.
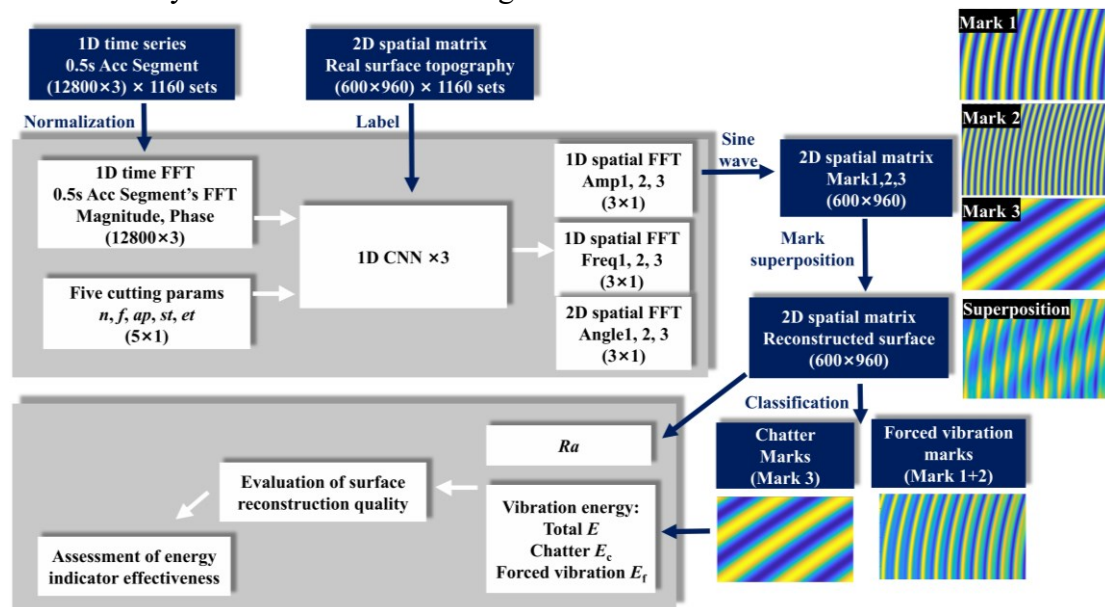

Figure 1: The flow diagram of surface reconstruction and indicator computation

To address the above limitations, this paper introduces a novel vibration energy based indicator for real-time surface quality assessment in end milling, combining the strength of Ra evaluation with vibration source analysis, with the computation flow outlined in Figure 1. By distinguishing between the contributions of chatter and forced vibrations to surface quality, the indicator offers actionable insights for vibration suppression and quality control. A real-time computation framework is developed, integrating a spectral surface topography generation mechanism with 1D CNNs to predict three distinct reconstruction parameters: amplitude A, spatial frequency ξ, and superposition angle φ, generating sine waves that superimpose to form the complete surface, instead of directly predicting topography. This strikes a balance between analytical mechanisms and data-driven modeling, avoiding reliance solely on one method. The FFT of acceleration signals serves as the input, and three separate 1D CNNs are trained to predict each parameter independently. The reconstructed surfaces are compared to real surfaces (scanned by a white-light interferometer) and measured by Ra to assess accuracy. Vibration energies, total E, chatter E_c, and forced vibration E_f, are computed and correlated with Ra^2 values to evaluate the energy indicator's effectiveness. The distribution of these energies across different Ra^2 values, along with surface topography,

provides guidance for targeted vibration suppression. This approach offers a potential framework for integrating online vibration and surface quality assessment and improvement into digital twin systems for manufacturing.

2 Methods

2.1 Vibration marks generation

According to Grossi's spectral interpretation of milled surface [17], the tool's vibration trajectory, although not harmonic but periodic, can be interpreted as a sum of harmonic sine functions with different spatial frequencies. This follows Fourier analysis, which states that any periodic function can be decomposed into harmonic functions. Consequently, the tool's vibration trajectory, reflected on the surface as forced vibration marks and chatter marks, forms the surface topography and can be represented by harmonic sine functions with different spatial frequencies ξ, i.e., the inverse of the wavelength of the marks. These spatial frequencies can be identified by applying Fourier transform (FT) on the surface profile.

Each mark's topography height (z values) can be simulated as a sine wave on a XY mesh grid, where x is the feed direction, y is the cross-feed direction, and z is the tool axial direction. The sine waves are computed based on their amplitude, spatial frequency, and superposition angle, and then superimposed to reconstruct the surface. Typically, three primary sets of marks, that is, three dominant spatial frequencies, are sufficient to capture more than 90% of the surface energy, ensuring high-fidelity reconstruction. This is supported by Parseval's theorem, which states that the total energy in the spatial domain equals the total energy in the frequency domain, implying that the top frequencies capture most of the surface's energy. The formula used for the simulated marks is:

$$z_i = A_i \cdot \sin(2\pi \cdot \xi_i \cdot (\cos(\varphi_i) \cdot (x - \Delta) + \sin(\varphi_i) \cdot y)) \tag{1}$$

$$\Delta = r - \sqrt{r^2 - y^2} \tag{2}$$

$$z = \sum_{i=1}^{3} z_i \tag{3}$$

where A is the amplitude, ξ is the spatial frequency, φ is the superposition angle, Δ is the phase shift which controls the mark curvature, shaping the marks into arcs instead of strips, r is the tool radius.

The surface topography height map, scanned by a Rtec white-light interferometer, is used to determine the amplitude A, spatial frequency ξ, and superposition angle φ. The FFT of each data line along the feed direction (x-direction) is computed for a 600×960 surface data matrix, and the results are averaged to obtain a representative spectrum. From this spectrum, the spatial frequencies ξ and their corresponding

magnitudes ($A/2$, half of amplitude) are identified, with the top three frequencies by magnitude selected as the dominant ones, as shown in Figure 2. In this case, the dominant frequencies are 1.44, 8.64, and 17.27 cycles/mm, with magnitudes of 8.55×10^{-7}, 8.12×10^{-7}, and 2.96×10^{-7}, respectively. The superposition angle φ is derived from the 2D FFT of the entire surface data matrix by connecting the spatial frequency pairs (e.g., 1.44 and -1.44) and measuring the angle between the connecting line and the horizontal reference line (positive for anticlockwise), as shown in Figure 3. For this case, the angles are -46.87°, -10.67°, and -15.78°. These three parameters are then used to build up three marks and superimposed to reconstruct the whole surface, as shown in Figure 4.

Note that the FFT magnitude must be multiplied by 2 to accurately reconstruct the amplitude, as the FFT splits energy between positive and negative frequencies, and only the positive frequency's magnitude is used. In addition, FFT is performed along the x-direction (feed direction), as it inherently contains the spatial frequency information of forced vibration marks, appearing as regular arc segments. The y-direction lacks this information. Chatter marks, in contrast, form an angle with the feed direction and do not align completely with x or y, making either direction suitable for their analysis.
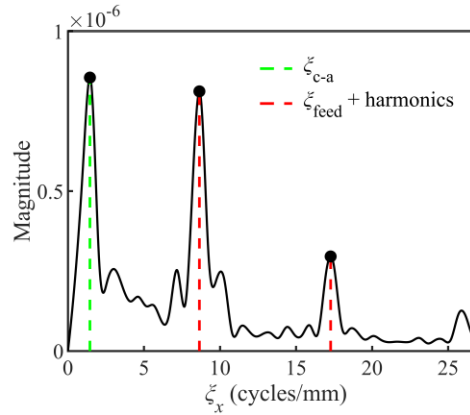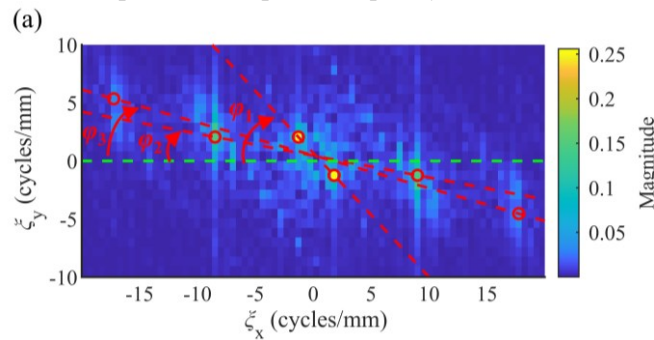

Figure 2: Dominant amplitude and spatial frequency identification from FFT spectrum
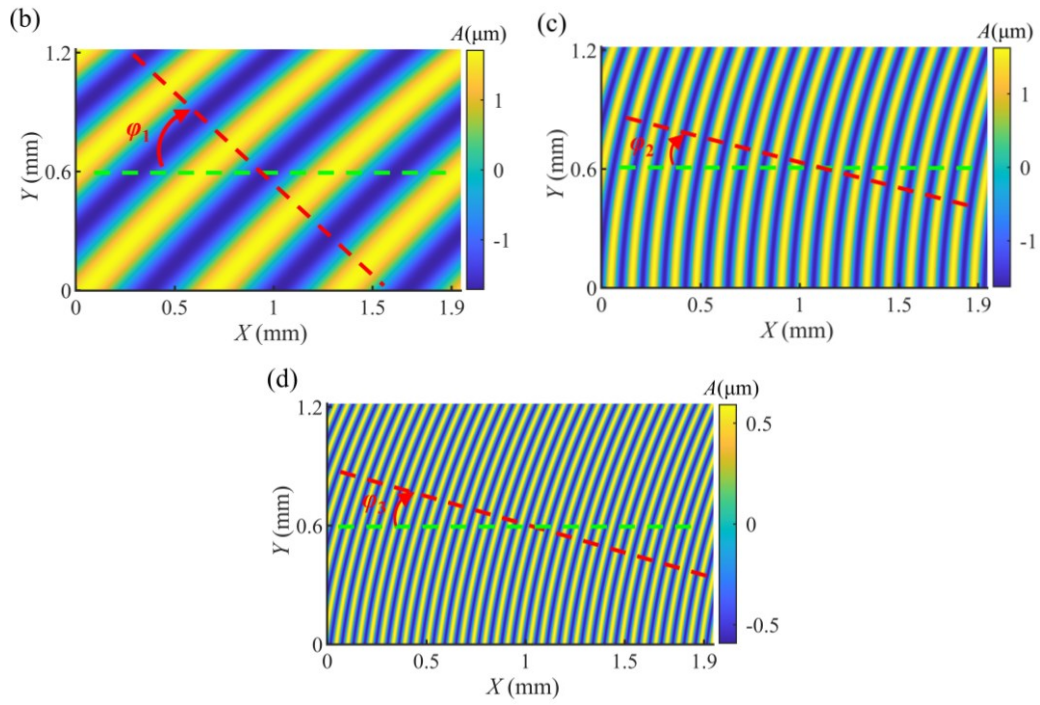
Figure 3: Superposition angle identification from 2D FFT spectrum
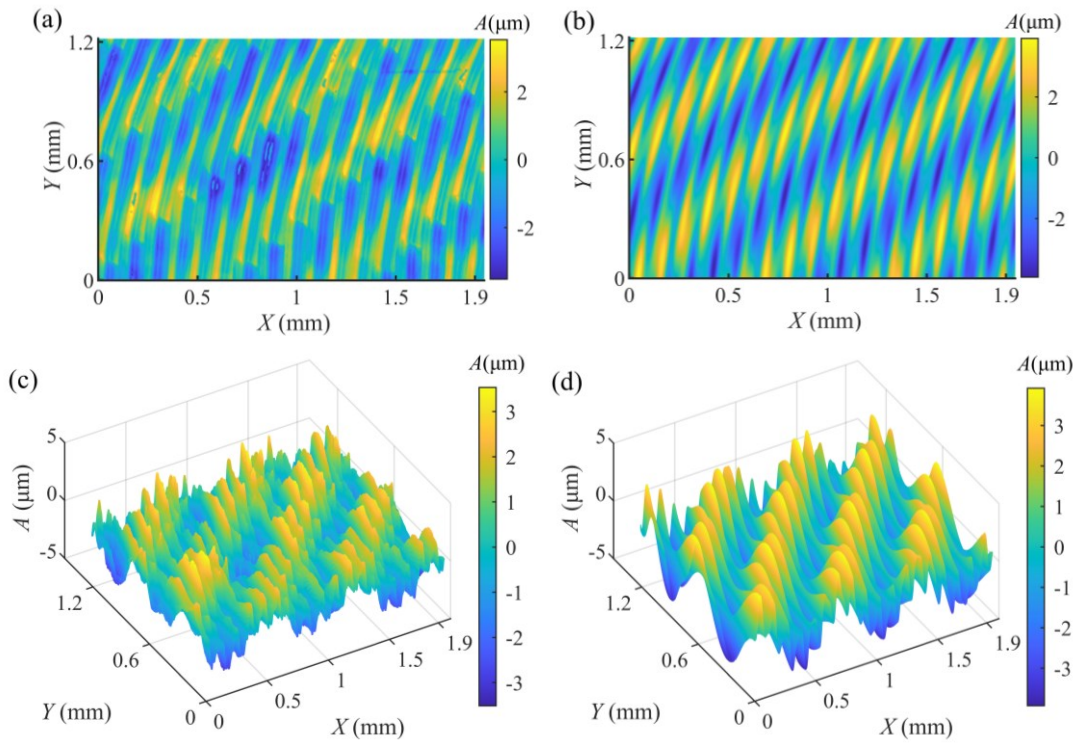(a) 2D FFT spectrum; (b) Mark1; (c) Mark2; (d) Mark3

Figure 4: Superposition of three dominant marks
(a) Scanned 2D surface; (b) Reconstructed 2D surface;
(c) Scanned 3D surface; (d) Reconstructed 3D surface

2.2 Data preparation and model structure

Section 2.1 details a method for reconstructing surface topography using three key parameters derived from actual measurements, making it an offline approach. To enable online prediction, a 1D CNN is used to predict these parameters directly from real-time acceleration signals during end milling. The network is trained with parameters derived from scanned surface topography, establishing a link between the acceleration signals and the reconstruction parameters.

The neural network requires three datasets: three parameters derived from actual surface topography, acceleration signals during milling, and cutting parameters. The first serves as the true value labels, while the last two are used as input layers. The label data consists of 1160 sets, each containing three subsets of amplitude A, spatial frequency ξ, and superposition angle φ, derived from a scanned surface, resulting in a final shape of (1160, 3). The first input layer uses the FFT of each acceleration segment, capturing both magnitude and phase components across the x, y, and z directions. Each FFT segment is shaped as (12800, 3), resulting in a total shape of (1160, 12800, 3), where 12800 corresponds to the sampling rate of 25600 samples per second. The second input layer contains five cutting parameters: spindle speed n, feed rate f, axial depth of cut a_p, segment start time st, and segment end time et, formatted as (5,1).

The choice of deep learning technique over traditional statistical, optimization, and common machine learning methods is driven by the need to capture complex nonlinear relationships, especially the spatiotemporal coupling and inconsistent dimensional mapping between acceleration signals and vibration marks. The model should be able to predict 1D and 2D spatial reconstruction parameters from 1D time-domain signals. After testing various deep learning models, including CNN, LSTM, RNN, and GAN, CNN was ultimately selected for its automatic feature extraction through convolutional layers and ease of training and deployment.

To enhance prediction accuracy, three separate 1D CNNs are constructed to individually predict amplitude A, spatial frequency ξ, and superposition angle φ. This approach is more effective than a single model, as it avoids inconsistent loss minimization across different physical quantities. Additionally, using the FFT of acceleration data rather than raw time segments significantly improves performance. The optimal configuration includes one Conv1D layer with 128 filters, supported by MaxPooling and Dropout layers to prevent overfitting. Keras is used as the framework for this research. All the above data and model structures are summarized in Figure 5.
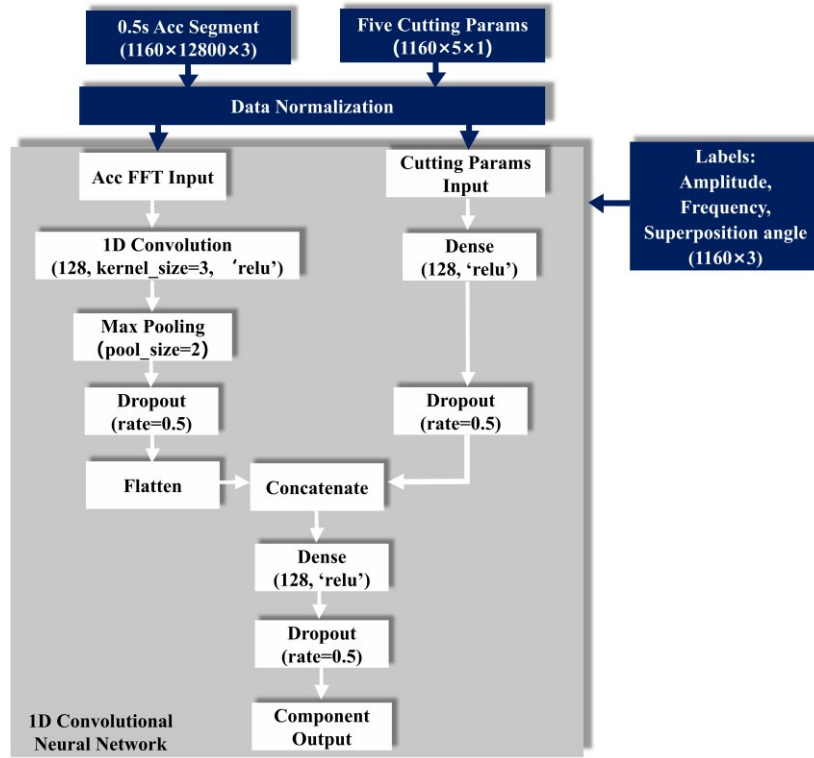
Figure 5: Structure of 1D CNN

2.3 Vibration energy estimation

The cutting tool, basically an axially loaded beam, can be modeled as a spring-mass system [31]. In such a system, the elastic potential energy is given by $E=1/2Kx^2$, where K is the total stiffness, and x is the displacement at the tool tip. Since the tool tip's displacements are accurately reflected in the milled surface's topography, the z-values (heights) of the reconstructed surface can replace x to compute the vibration energy E using Equation (4). The total stiffness K is a sum of the tension-induced stiffness K_T and the bending stiffness K_B of the cutting tool, as expressed in Equation (5) and (6), where E_m is cutting tool's Young's modulus, D and L are tool's diameter and stick-out length. To eliminate the effects of different surface sizes and resolutions, the vibration energy is normalized by $1/N$, where N represents the total number of surface grid points $n \times m$. All calculations are performed using SI units: length in meters (m), spatial frequency in cycles per meter (cycles/m), elastic modulus in Pascals (Pa), and energy in Joules (J).

$$E = \frac{1}{N} \times \frac{1}{2} K \sum_{i=1}^{n} \sum_{j=1}^{m} z_{ij}^2 \qquad (4)$$

$$K_T = E_m \times \frac{\pi D^2}{4L} \qquad (5)$$

$$K_{\mathrm{B}} = \frac{3E_{\mathrm{m}} \cdot I}{L^3}; \quad I = \frac{\pi D^4}{64} \tag{6}$$

2.4 Experimental details

Slot milling experiments are conducted on an EUMA DU810 vertical machining center (Figure 6a) using an 8 mm diameter carbide end mill cutter with a 40 mm stick-out length and 2 teeth, with a Young's modulus E_{m} of 2.06×10^{11} Pa. The workpiece material is 7075 aluminum alloy, measuring 150 mm × 150 mm. Cutting parameters vary as follows: spindle speed n (3000-10000 r/min), feed rate f (300-2000 mm/min), axial depth of cut a_{p} (1 mm), and width of cut a_{e} (8 mm, slot milling). Table 1 details the 58 slot milling cuts performed, with consistent radial and axial depths of cut across all experiments, focusing solely on the impact of different spindle speed and feed rate combinations. Acceleration data is collected using a PCB piezotronics triaxial vibration sensor, sampled at 25600 samples per second. Each cut provides 20 evenly spaced 0.5-second segments, totaling 1160 segments. Data is acquired with an LMS SCADAS III multichannel system (Figure 6c), processed in LMS TestLab 17 and analyzed in MATLAB R2023a. Stable and chatter conditions are shown in Figure 6d and 6e.

The 3D topography of the workpiece surface is scanned with a Rtec white-light interferometer (WLI), covering an area of 1.946 mm × 1.216 mm, resulting in a 600 × 960 data matrix with a resolution of 493.3 cycles/mm, as shown in Figure 7. A total of 1160 WLI surface topography photos are obtained from 58 cuts, with 20 evenly spaced locations per cut. To align acceleration segments with the corresponding scanned surface photo, the exact cutting time is determined by locating the scanned surface on the cut, calculating the distance from the start of cutting, and dividing by the feed rate. A 0.5-second segment is sufficient to capture the scanned surface length (1.946 mm) without overlap.

The 1D CNN training is conducted using TensorFlow 2.9.0 framework on Python 3.8 (Ubuntu 20.04), running on an NVIDIA RTX 3090 GPU with 24GB memory and CUDA 11.2 for acceleration. The system is powered by a 14-core Intel® Xeon® Platinum 8362 CPU @ 2.80GHz, with 45GB of RAM.
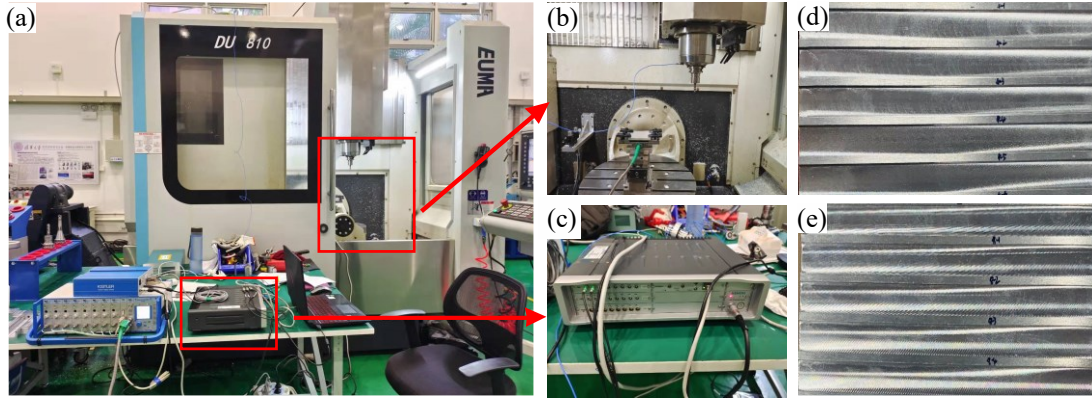
Figure 6: Experimental equipment and data analysis illustration

(a) EUMA DU810 vertical CNC machining center; (b) Workbench; (c) LMS SCADAS III;
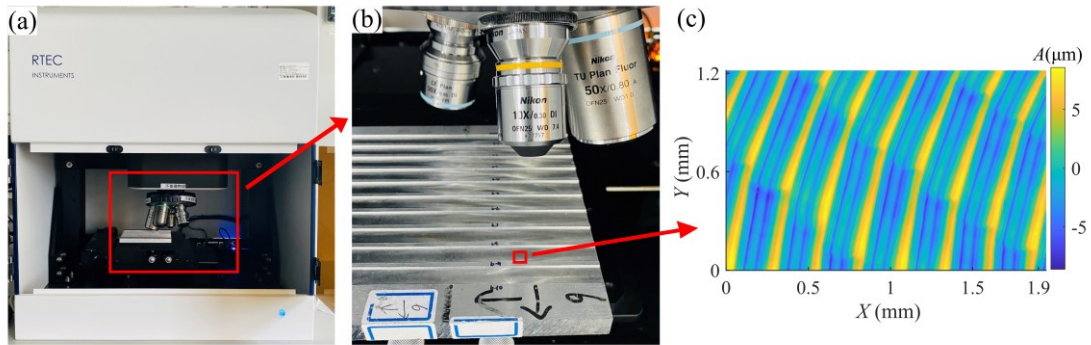
(d) Stable surfaces; (e) Chatter surfaces

Figure 7: Rtec white-light interferometer (WLI) illustration

(a) Rtec white-light interferometer; (b) Workbench; (c) Scanned real surface topography

Table 1: Spindle speed and feed rate combinations

Spindle speed n range (rpm)	Spindle speed step (rpm)	Feed rate f range (mm/min)	Feed rate step (mm/min)
3000-6000	1000	300-1300	100
7000, 10000	/	300-2000	100

3 Results and discussion

Section 3.1 presents the vibration energy indicator's close correlation with Ra^2 and its effectiveness in quantifying chatter and forced vibration contributions to surface quality. Section 3.2, evaluates the accuracy of the surface reconstruction model to validate the indicator's reliability.

3.1 Quantitative indicator of vibration energy

According to Grossi, surface marks can be categorized into four types: feed spatial frequency ξ_{feed}, aliased spatial frequency ξ_{c-a}, pseudo-moire effect spatial frequency ξ_{pm}, and beating effect spatial frequency ξ_{c-a_pm} [17]. Marks with ξ_{feed} are considered as forced vibration marks (feed marks), while the others are classified as generalized "chatter" marks since they originate from chatter aliasing spatial frequency. This classification allows the surface to be divided into chatter and forced vibration sub-surfaces by superimposing their corresponding marks. Using Equation (4), total energy, forced vibration energy, and chatter energy are computed from the corresponding surfaces. The relationship between real Ra and energy indicator of each surface is analyzed to assess the effectiveness of the indicator, along with the variation in chatter energy and forced vibration energy. To better illustrate the relationship between E and Ra, Ra^2 is used for plotting since E is proportional to z^2 while Ra to z. The correlation fit is performed using a smoothing spline method.

Figure 8(a) shows a strong positive correlation between total vibration energy E and square surface roughness Ra^2, with a Pearson correlation coefficient of $r = 0.98$ indicating that the energy indicator closely tracks variations in Ra and serves as a reliable quantitative measure of surface quality. In contrast, both forced vibration energy E_f and chatter energy E_c display some nonlinear behavior, as shown in Figure 8(b) and (c). E_f has a $r = 0.95$, which also indicates a strong correlation, but fluctuations are noticeable in the range of $Ra^2 \leq 1.6^2$ μm^2. Similarly, E_c also exhibits irregular behavior in this region, with a $r = 0.33$. These fluctuations suggest that chatter and forced vibrations coexist and both contribute to surface roughness in this phase. It is evident that chatter can occur even within the typical Ra range for milling, but its impact on surface quality is not severe.

When $Ra^2 > 1.6^2$ μm^2, E_f rises steadily while E_c stabilizes with small dips. This suggests that forced vibrations become the primary factor impacting surface roughness under high Ra conditions, likely due to improper cutting parameters (e.g., low spindle

speed and high feed rate) amplifying vibrations in machine components or workpiece. Conversely, chatter energy appears to reach a limit and does not grow anymore due to the regenerative effect, where energy accumulates from varying chip width and cutting forces. However, the chip width has a physical threshold, preventing further energy growth.
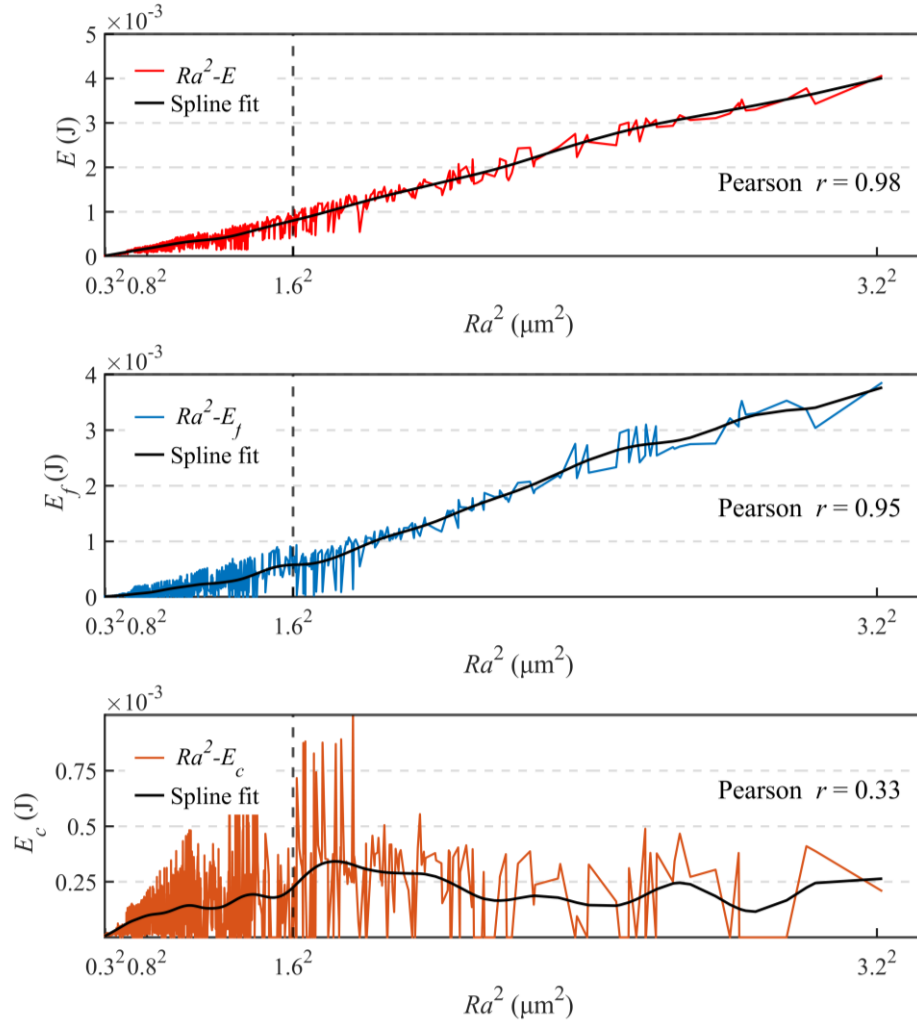

Figure 8: Correlation between vibration energy E, E_f, E_c and surface roughness squared Ra^2

Figure 9 illustrates the contributions of forced vibration energy E_f and chatter energy E_c to surface quality across different ranges of Ra^2. In the low roughness range (a), both energies are balanced and stable, with surface examples showing uniform patterns and only arc-shaped forced vibration marks, reflecting smooth machining with minimal defects. As roughness increases to the moderate range (b), large fluctuations occur, with chatter and forced vibrations alternately dominating. The surfaces in (b) exhibit a mix of chatter and forced vibration marks, leading to a less uniform texture, suggesting real-time adjustments are needed to stabilize the process. In the high roughness range (c), E_f consistently dominates, and the surfaces display pronounced

periodic forced vibration marks, distorted by underlying chatter marks, making them less arc-shaped than in (a). These patterns highlight the need for controlling machine imbalances, such as tool balancing and fixture security, to improve surface quality. Based on these observations, the following guidance is proposed: for low roughness, no adjustments are needed as the process remains stable. In the moderate roughness range, real-time adjustments to spindle speed and feed rate are essential. For high roughness, controlling machine imbalances is critical. This emphasizes the importance of addressing different vibration contributions to optimize surface finish in real-time manufacturing. Figure 10 provides the composition of E_f and E_c, quantifying vibration energy and offering insight into the transition from acceptable to severe vibrations that can harm surface quality. For instance, for Ra^2 values slightly above 1.6^2 (orange points), focusing on reducing forced vibrations would be more effective than addressing chatter. With further investigation, this energy indicator could help predict severe vibration events, making it a valuable tool for real-time surface quality control and optimization in digital twin systems.
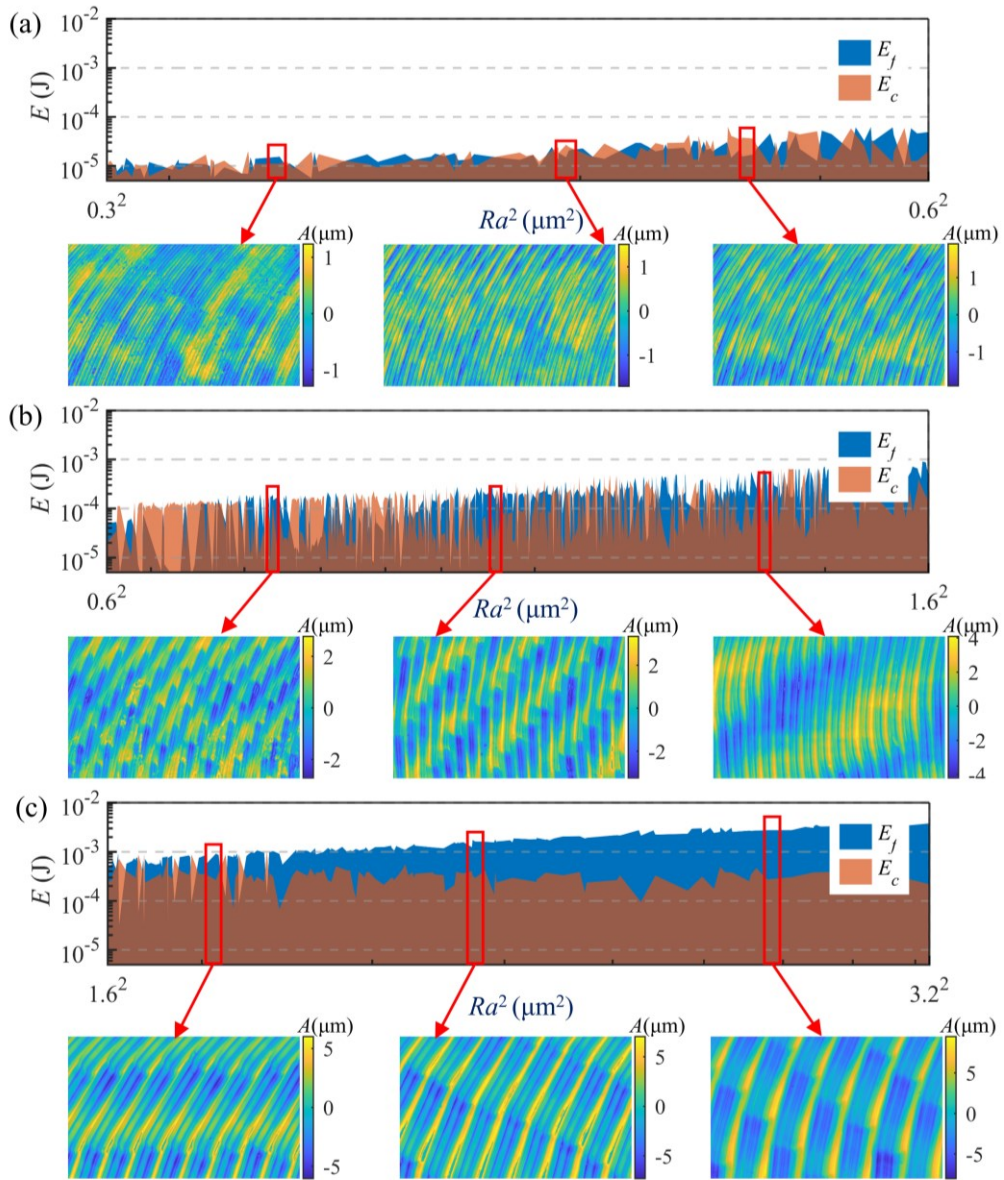
Figure 9: Distribution of vibration energy E_f and E_c across surface roughness squared Ra^2
(a) $0.3^2{\leqslant}Ra^2{\leqslant}0.6^2$; (b) $0.6^2{\leqslant}Ra^2{\leqslant}1.6^2$; (c) $1.6^2{\leqslant}Ra^2{\leqslant}3.2^2$

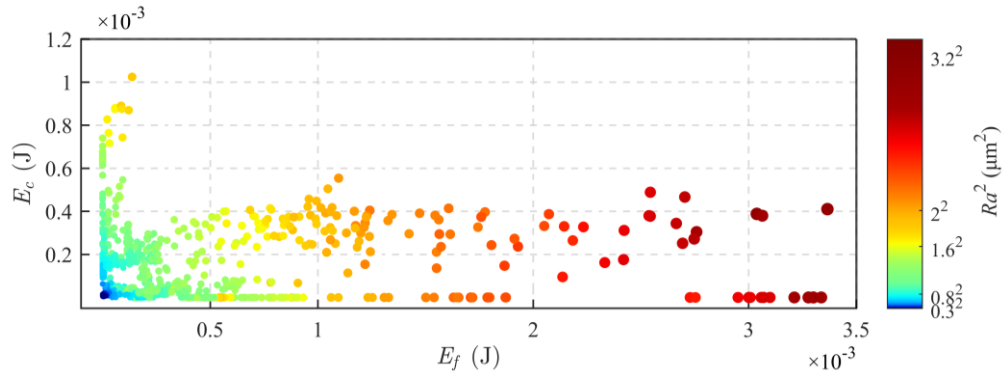

Figure 10: Constitution of forced vibration energy E_f and chatter energy E_c

3.2 Evaluation of surface reconstruction quality

A total of 1160 data sets are fed into the 1D CNN, with 800 sets for training and 360 for testing, maintaining a 7:3 ratio. The network's quality is assessed by calculating the accuracy as (1 - test loss)×100%, which measures the difference between the predicted and actual values for three surface reconstruction parameters: amplitude A, spatial frequency ξ, and superposition angle φ. This measurement can be done using four different metrics in statistical terms, or loss functions in neural network terms: MAE, MSE, Huber, and log-cosh. The following discussion adheres to the term "loss functions". Figure 11 compares the prediction accuracies of these loss functions. Notably, log-cosh and Huber perform best, both achieving 98% accuracy for amplitude, 78% and 76% for spatial frequency, and 76% and 73% for superposition angle. In contrast, MAE and MSE yield much lower accuracies, particularly in predicting frequency and superposition angle, hovering around 40%. This stark contrast highlights the superior effectiveness of log-cosh and Huber in handling non-Gaussian distributions, as also concluded in reference [23]. Given this, log-cosh emerges as the most suitable loss function for this application.

This performance is closely tied to data distribution, with amplitude following a Gaussian distribution, leading to high accuracy. However, non-Gaussian distributions for frequency and orientation angle limit accuracy to around 75%, even with log-cosh. Improving accuracy will require a more diverse and well-distributed dataset, likely involving carefully designed experiments with varying axial and radial depths of cut, a focus for future work.
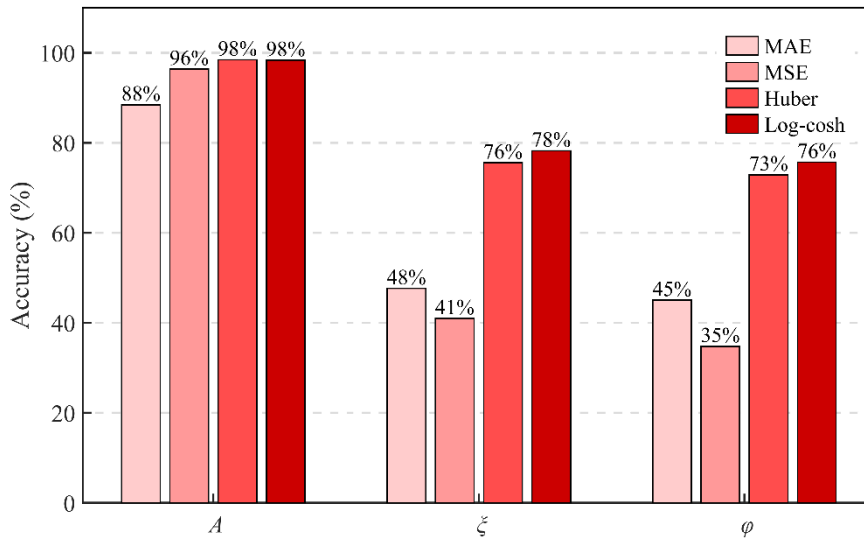

Figure 11: Prediction accuracies (1-test loss) ×100% for amplitude A, spatial frequency ξ, and superposition angle φ across four loss functions: MAE, MSE, Huber, Log-cosh

Reconstructed surfaces are generated by applying the trained 1D CNN to 1160 acceleration segments, using the predicted parameters to reconstruct the surfaces. The real surfaces, scanned by WLI, serve as a reference. Predicted and real Ra values, denoted as Ra_{pred} and Ra_{real}, are calculated from reconstructed and real surfaces respectively, and then compared. Absolute errors $|Ra_{\text{pred}} - Ra_{\text{real}}|$ is computed to quantitatively assess reconstruction accuracy.

Figure 12 shows that while Ra_{pred} values generally follow the trend of the real values with the absolute error mostly under 0.5 µm, some deviations occur. Specifically, in the medium Ra range of 0.8 to 1.6 µm, which is common in end milling and covers about half of the surfaces, stronger deviations are observed, particularly around index 800. The model tends to underestimate in the small range of 0 to 0.8 µm and overshoot in the large range 1.6 to 3.2 µm. These issues likely stem from the training data distribution, where half of the data falls within the medium Ra range, and the other half is split between the small and large ranges, which may affect the model's ability to accurately predict medium Ra values. Constructing separate networks for different Ra ranges might help improve accuracy by minimizing these effects and prevent underestimation and overshooting.

Figure 13 provides an intuitive comparison between part of the real and reconstructed surfaces. High similarities can be observed, with most features being accurately captured by the reconstructed surfaces. All the above evaluations collectively demonstrate a strong performance in surface reconstruction, impressive but potentially limited in extreme cases due to suboptimal data distribution.
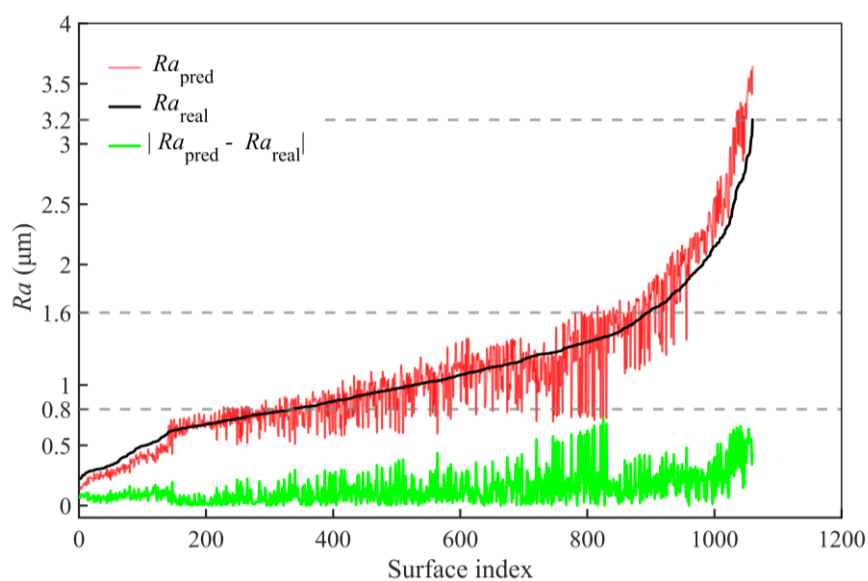

Figure 12: Comparison of Ra values for reconstructed and real surfaces, Ra_{pred} and Ra_{real}
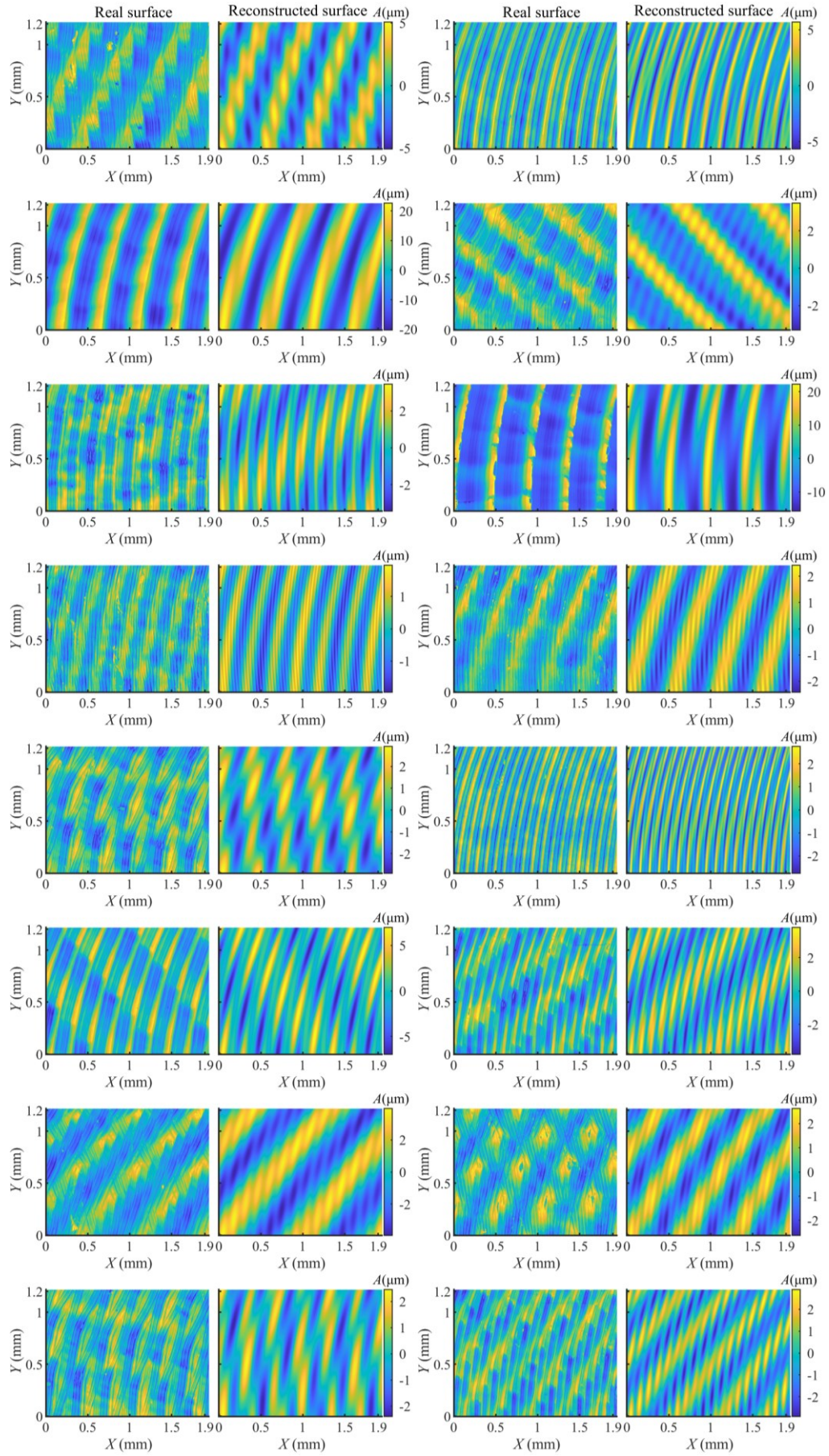
Figure 13: Intuitive comparison of real (scanned by WLI) and reconstructed surfaces

4 Conclusions

1. The proposed vibration energy based quantitative indicator combines the strengths of Ra evaluation with vibration source analysis, addressing the limitations of equating chatter with poor surface quality or relying on Ra solely for acceptability without offering guidance on vibration suppression and surface quality control. It achieves a high correlation with Ra^2 ($r = 0.98$), indicating a similar capability to Ra in surface quality evaluation. By analyzing the distribution of chatter and forced vibration energy E_c, E_f across different Ra^2 ranges, the indicator provides actionable insights. For instance, in common surface roughness cases during end milling (Ra below 1.6), chatter and forced vibrations often coexist or alternate in dominance, requiring tailored suppression strategies rather than simply adjusting spindle speed or feed rate, which may be ineffective when forced vibrations dominate.

2. A real-time computation framework is developed, integrating a spectral surface topography generation mechanism and neural network techniques for online indicator computation and surface quality assessment. Acceleration signals are processed for surface reconstruction, and the resulting surface is used to compute the indicator. A pre-trained 1D CNN converts acceleration data into three spatial parameters (amplitude A, spatial frequency ξ, and superposition angle φ), generating sine waves that superimpose to form the complete surface, balancing data-driven modeling with analytical mechanisms rather than directly predicting topography. The reconstructed surfaces closely match real surfaces, with predicted Ra values showing an absolute error of mostly less than 0.5 μm. Deviations likely arise from the non-Gaussian distribution of label data, affecting the model's reconstruction performance. Nevertheless, the method's accuracy and real-time capability demonstrate its potential for online surface quality assessment for end milling, offering a robust framework for digital twin systems in manufacturing.

References

[1] Y. Altintas, M. Weck, Chatter stability of metal cutting and grinding, CIRP Ann. 53 (2) (2004) 619–642. https://doi.org/10.1016/S0007-8506(07)60032-8.

[2] C. Yue, H. Gao, X. Liu, S.Y. Liang, L. Wang, A review of chatter vibration research in milling, Chin. J. Aeronaut. 32 (2) (2019) 215–242. https://doi.org/10.1016/j.cja.2018.11.007.

[3] D. Hajdu, F. Borgioli, W. Michiels, T. Insperger, G. Stepan, Robust stability of milling operations based on pseudospectral approach, Int. J. Mach. Tools Manuf. 149 (2020) 103516. https://doi.org/10.1016/j.ijmachtools.2019.103516.

[4] M. Sanz-Calle, J. Munoa, A. Iglesias, L.N. López De Lacalle, Z. Dombovari, Semianalytic stability algorithm in the frequency domain for interrupted milling, Int. J. Mach. Tools Manuf. 187 (2023) 104005. https://doi.org/10.1016/j.ijmachtools.2023.104005.

[5] V. Ostad Ali Akbari, M. Kuffa, K. Wegener, Physics-informed Bayesian machine learning for probabilistic inference and refinement of milling stability predictions, CIRP J. Manuf. Sci. Technol. 45 (2023) 225–239. https://doi.org/10.1016/j.cirpj.2023.07.004.

[6] T. Schmitz, Milling stability modeling by sample partitioning with chatter frequency-based test point selection, J. Manuf. Mater. Process. 8 (3) (2024) 109. https://doi.org/10.3390/jmmp8030109.

[7] M. Postel, D. Aslan, K. Wegener, Y. Altintas, Monitoring of vibrations and cutting forces with spindle mounted vibration sensors, CIRP Ann. 68 (1) (2019) 413–416. https://doi.org/10.1016/j.cirp.2019.03.019.

[8] P. Seyrek, B. Şener, A.M. Özbayoğlu, H.Ö. Ünver, An evaluation study of EMD, EEMD, and VMD for chatter detection in milling, Procedia Comput. Sci. 200 (2022) 160–174. https://doi.org/10.1016/j.procs.2022.01.215.

[9] K. Jeong, Y. Seong, J. Jeon, S. Moon, J. Park, Chatter monitoring of machining center using head stock structural vibration analyzed with a 1D convolutional neural network, Sensors 22 (14) (2022) 5432. https://doi.org/10.3390/s22145432.

[10] S.A. Sajjady, H. Nouri Hossein Abadi, S. Amini, R. Nosouhi, Analytical and experimental study of topography of surface texture in ultrasonic vibration assisted turning, Mater. Des. 93 (2016) 311–323. https://doi.org/10.1016/j.matdes.2015.12.119.

[11] E. Artetxe, Dv. Olvera, L.N.L. De Lacalle, F.J. Campa, Dn. Olvera, A. Lamikiz, Solid subtraction model for the surface topography prediction in flank milling of

thin-walled integral blade rotors (IBRs), Int. J. Adv. Manuf. Technol. 90 (1–4) (2017) 741–752. https://doi.org/10.1007/s00170-016-9435-1.

[12] S.E. Layegh K, I. Lazoglu, 3D surface topography analysis in 5-axis ball-end milling, CIRP Ann. 66 (1) (2017) 133–136. https://doi.org/10.1016/j.cirp.2017.04.021.

[13] Y. Ding, L. Zhu, X. Zhang, H. Ding, On a numerical method for simultaneous prediction of stability and surface location error in low radial immersion milling, J. Dyn. Syst. Meas. Control 133 (2) (2011) 024503. https://doi.org/10.1115/1.4003374.

[14] D. Yang, Z. Liu, Surface plastic deformation and surface topography prediction in peripheral milling with variable pitch end mill, Int. J. Mach. Tools Manuf. 91 (2015) 43–53. https://doi.org/10.1016/j.ijmachtools.2014.11.009.

[15] Y. Sun, Y. Liu, M. Zheng, J. Xu, Q. Guo, A review on theories/methods to obtain surface topography and analysis of corresponding affecting factors in the milling process, Int. J. Adv. Manuf. Technol. 127 (7–8) (2023) 3097–3131. https://doi.org/10.1007/s00170-023-11723-4.

[16] G. Stepan, M. Toth, D. Bachrathy, S. Ganeriwala, Spectral properties of milling and machined surface, Mater. Sci. Forum 836–837 (2016) 570–577. https://doi.org/10.4028/www.scientific.net/MSF.836-837.570.

[17] N. Grossi, A. Scippa, L. Sallese, F. Montevecchi, G. Campatelli, On the generation of chatter marks in peripheral milling: A spectral interpretation, Int. J. Mach. Tools Manuf. 133 (2018) 31–46. https://doi.org/10.1016/j.ijmachtools.2018.05.008.

[18] Y. Shao, S. Du, L. Xi, 3D machined surface topography forecasting with space-time multioutput support vector regression using high definition metrology, in: Vol. 1 37th Comput. Inf. Eng. Conf., American Society of Mechanical Engineers, Cleveland, Ohio, USA, 2017: p. V001T02A069. https://doi.org/10.1115/DETC2017-67155.

[19] C. Liu, L. Gao, G. Wang, W. Xu, X. Jiang, T. Yang, Online reconstruction of surface topography along the entire cutting path in peripheral milling, Int. J. Mech. Sci. 185 (2020) 105885. https://doi.org/10.1016/j.ijmecsci.2020.105885.

[20] L. Cao, X.M. Zhang, T. Huang, H. Ding, Online monitoring machining errors of thin-walled workpiece: a knowledge embedded sparse bayesian regression approach, IEEEASME Trans. Mechatron. 24 (3) (2019) 1259–1270. https://doi.org/10.1109/TMECH.2019.2912195.

[21] S. Ghosh, R. Knoblauch, M. El Mansori, C. Corleto, Towards AI driven surface roughness evaluation in manufacturing: a prospective study, J. Intell. Manuf. (2024). https://doi.org/10.1007/s10845-024-02493-1.

[22] H.C. Möhring, S. Eschelbacher, P. Georgi, Machine learning approaches for real-time monitoring and evaluation of surface roughness using a sensory milling tool, Procedia CIRP 102 (2021) 264–269. https://doi.org/10.1016/j.procir.2021.09.045.

[23] A.P. Rifai, H. Aoyama, N.H. Tho, S.Z. Md Dawal, N.A. Masruroh, Evaluation of turned and milled surfaces roughness using convolutional neural network, Measurement 161 (2020) 107860. https://doi.org/10.1016/j.measurement.2020.107860.

[24] W.J. Lin, S.H. Lo, H.T. Young, C.L. Hung, Evaluation of deep learning neural networks for surface roughness prediction using vibration signal analysis, Appl. Sci. 9 (7) (2019) 1462. https://doi.org/10.3390/app9071462.

[25] X. Zhu, G. Chen, C. Ni, X. Lu, J. Guo, Hybrid CNN-LSTM model driven image segmentation and roughness prediction for tool condition assessment with heterogeneous data, Robot. Comput.-Integr. Manuf. 90 (2024) 102796. https://doi.org/10.1016/j.rcim.2024.102796.

[26] P. Rao, S. Bukkapatnam, O. Beyca, Z. (James) Kong, R. Komanduri, Real-time identification of incipient surface morphology variations in ultraprecision machining process, J. Manuf. Sci. Eng. 136 (2) (2014) 021008. https://doi.org/10.1115/1.4026210.

[27] K.B. Lee, S. Cheon, C.O. Kim, A convolutional neural network for fault classification and diagnosis in semiconductor manufacturing processes, IEEE Trans. Semicond. Manuf. 30 (2) (2017) 135–142. https://doi.org/10.1109/TSM.2017.2676245.

[28] L. Cao, T. Huang, X.-M. Zhang, H. Ding, Generative adversarial network for prediction of workpiece surface topography in machining stage, IEEEASME Trans. Mechatron. 26 (1) (2021) 480–490. https://doi.org/10.1109/TMECH.2020.3032990.

[29] S. Kumar, B. Singh, Quantification of tool chatter and metal removal rate using wavelet denoising and statistical approach, Noise Vib. Worldw. 49 (2) (2018) 62–81. https://doi.org/10.1177/0957456518763159.

[30] G. Zhou, M. Yuan, F. Feng, Z.M. Han, X. Song, X.H. Wang, P.F. Feng, M. Zhang, A new algorithm for chatter quantification and milling instability classification based on surface analysis, Mech. Syst. Signal Process. 204 (2023) 110816. https://doi.org/10.1016/j.ymssp.2023.110816.

[31] J. Valle, D. Fernández, J. Madrenas, Closed-form equation for natural frequencies of beams under full range of axial loads modeled with a spring-mass system, Int. J. Mech. Sci. 153–154 (2019) 380–390. https://doi.org/10.1016/j.ijmecsci.2019.02.014.

Appendix A

The spectral interpretation of feed and chatter marks proposed by Grossi in reference [17] is experimentally validated against theoretical computations, with all calculations following the reference. An experiment with an axial depth of cut varying from 0.5 mm to 4 mm is used for illustration. Additional cutting details are provided in Table 2. The tooth passing frequency f_{tp}, calculated as:

$$f_{tp} = \frac{n}{60 \times z} = 116.67 \quad (\text{Hz}) \tag{7}$$

$$\xi_{feed} = \frac{1}{f_z} = \frac{s}{f} = 14 \quad (\text{cycles/mm}) \tag{8}$$

where n is spindle speed in unit of rpm, f is feed rate in unit of mm/min, f_z is feed per tooth, z is the number of flutes. The f_{tp} and its harmonics, 223 Hz, 1166 Hz, 1400 Hz, 4445 Hz, 4682 Hz, 4913 Hz are clearly observed as peaks in Figure 14(b). No chatter frequency is detected, and the magnitude remains far low. Therefore, only the feed spatial frequency ξ_{feed}, calculated as Equation (8), and its harmonics, caused by forced vibration due to periodic tool-workpiece impacts, are clearly visible in the 1D spatial spectrum as distinct peaks at 14, 28, 42, 56 cycles/mm when stable cutting, as shown in Figure 15(a).

When chatter dominates, chatter frequencies appear in the acceleration signal with prominent peaks, as shown in Figure 14(c). The frequency with the highest magnitude (close to the harmonics of natural frequency of 878 Hz for the tool-holder-spindle system) is typically identified as the chatter frequency $f_c = 4386$ Hz, leading to an aliasing effect that produces low spatial frequency marks on the surface. In Figure 15(b), the aliased spatial frequency ξ_{c-a}, fourth, sixth, seventh-order envelop spatial frequencies ξ_{pm4}, ξ_{pm6}, ξ_{pm7} caused by pseudo moiré effect, and the feed spatial frequency ξ_{feed}, as well as their harmonics are detected. The calculations are as follows:

$$\xi_{c-a} = \frac{|m \times f_{tp} - f_c|}{f_z \times f_{tp}} = 5.75 \quad (\text{cycles/mm}) \tag{9}$$

$$\xi_{pm} = \frac{h \times |f_{c-a} - f_{tp}/h|}{f_z \times f_{tp}}; \quad h = 4, 6, 7$$

$$\xi_{pm4} = 8.2, \quad \xi_{pm6} = 19.9, \quad \xi_{pm7} = 25.7 \quad (\text{cycles/mm}) \tag{10}$$

These computed frequencies are accurately predicted by Grossi's method and closely match the observed peaks in Figure 15(b). In addition, different surfaces may exhibit different types of spatial frequency, such as other-order envelope frequencies or

those from beating effects, which are not shown here. For most surfaces in this work, Grossi's method effectively captures and interprets chatter-induced marks, demonstrating its suitability for large-scale experimental slotting milled surfaces.
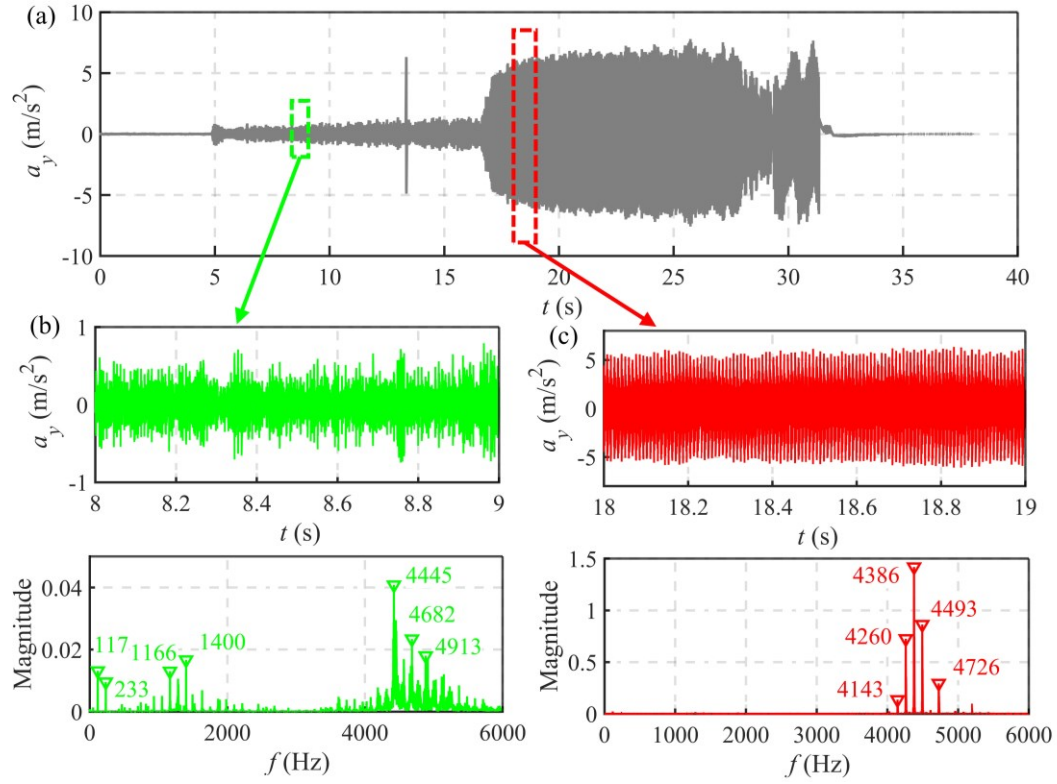

Figure 14: Acceleration signal a_y along cross-feed direction in varying depth of cut end milling (a) Full signal; (b) Stable cutting signal and spectrum; (c) Chatter cutting signal and spectrum
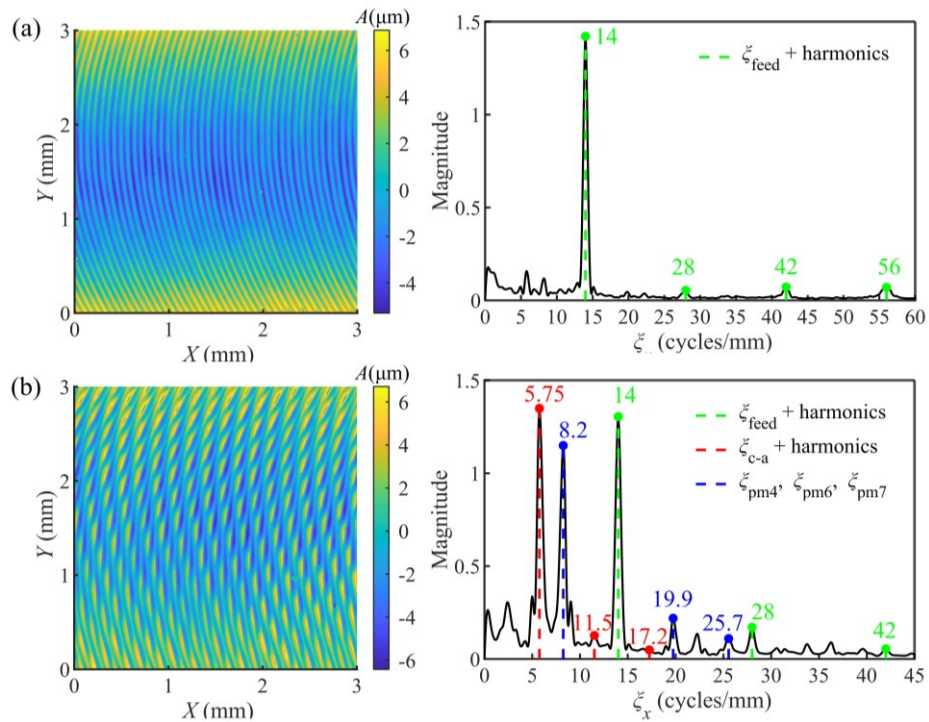

Figure 15: Surface topography and its 1D spatial spectrum

(a) Stable cutting; (b) Chatter cutting

Table 2: Cutting parameters

Spindle speed n (rpm)	Feed rate f (mm/min)	Axial depth of cut a_p (mm)	Width of cut a_e (mm)	Flute
7000	600	0.5-4	8	1