



Exploring the effectiveness of using a smartphone for position-oriented process monitoring

Kenan Deng¹ · Dong Gao¹ · Qinghe Guan¹ · Fenglei Wang¹ · Shoudong Ma¹ · Chang Zhao² · Yong Lu¹

Received: 25 October 2022 / Accepted: 22 January 2023 / Published online: 6 February 2023
© The Author(s), under exclusive licence to Springer-Verlag London Ltd., part of Springer Nature 2023

Abstract

Process monitoring is a practical approach for improving machining quality. However, expensive sensors and complex communication technologies render process-monitoring challenging. Smartphones with high-performance processors, open-source operating systems, and multi-sensors are promising portable process monitoring systems. This study proposed a smartphone-based position-oriented process monitoring and chatter detection method. First, a smartphone was used to video tape a machine tool's human-machine interface (HMI) and tool motion during the machining process. The tool cutting position can be correlated with the sound signal using tool motion images or numerical control (NC) program positions detected using the optical character recognition method. Open-source models are used for speech recognition and human-voice separation. Subsequently, the sound signal after removing the periodic components and human voice was used for chatter detection. The residual signal in the specified frequency band was extracted based on the multiple-tool modal frequencies. Accordingly, the ratio of the residual signal energy to the total signal energy was used to detect chatter. Finally, the proposed method was validated using robotic milling and deep-hole drilling experiments.

Keywords Smartphone · Process monitoring · Chatter detection · Position-oriented · Robotic milling · Deep-hole drilling

1 Introduction

Process monitoring is essential for improving the quality and efficiency of machining because the information obtained enables online control or offline optimization of

process parameters. Process monitoring systems integrate professional software, acquisition systems, and sensors. Moreover, they are integrated with machine-tool controllers through communication protocols [1–3]. However, the cost and technology of software and industrial sensors limit their application to small- and medium-sized enterprises (SMEs). Therefore, there is an urgent need for low-cost, portable, and easy-to-operate monitoring methods. Smartphones have become essential devices in people's personal and professional lives. Owing to the use of high-performance micro-processors, multiple embedded sensor types, and open-source software resources, smartphones are ideal tools for developing low-cost portable monitoring systems, especially for SMEs, which are cost-sensitive [4, 5]. Technicians can use smartphone apps to collect signals, such as images, sounds, and acceleration data, to process and analyze monitoring data. They can send data to the analysis device via the Web, Bluetooth, and instant messaging software (<https://phyphox.org/>, <https://ww2.mathworks.cn/products/matlab-mobile.html>) [6]. Stacks et al. (<https://phyphox.org/>, <https://ww2.mathworks.cn/products/matlab-mobile.html>) developed a mobile app called Phyphox to collect and analyze data that contained signal processing algorithms such as the fast Fourier transform (FFT) and time-frequency analysis method.

✉ Chang Zhao
zhaochang@connect.hku.hk

✉ Yong Lu
luyong@hit.edu.cn

Kenan Deng
1110810922@hit.edu.cn

Dong Gao
gaodong@hit.edu.cn

Qinghe Guan
18B908122@stu.hit.edu.cn

Fenglei Wang
22B908132@stu.hit.edu.cn

Shoudong Ma
neusdma@163.com

¹ School of Mechatronic Engineering, Harbin Institute of Technology, Harbin 150001, China

² Department of Mechanical Engineering, The University of Hong Kong, Pokfulam, Hong Kong

Smartphones have been used for vibration monitoring and can obtain measurements with an accuracy close to that of professional vibration sensors [7, 8]. Therefore, smartphones have been used for structural vibration performance testing [7], rail transit vibration comfort tests [8], and bridge health monitoring [9, 10]. However, process monitoring requires industrial sensors with high sampling frequencies (>1 kHz) and large measurement ranges (>20 g), which are difficult to achieve with triaxial accelerometers on smartphones.

Recently, machine vision has been used to monitor the condition of machine tools. Remote technicians can observe the machining site using cameras installed inside or outside the machine [11, 12]. Deep learning-based text recognition techniques are used to obtain machine tool information from human-machine interface (HMI), enabling automatic monitoring. Kim et al. [13] proposed a low-cost machine tool monitoring system based on a commercial camera and open-source software to monitor online machine operation data by gathering information on the HMI of CNC tools. Lee et al. [14] used a webcam to monitor the HMI of a five-axis machine, and adjusted the spindle speed by using the machine's control panel. Xing et al. [15] developed a data recording system for machine tool measurements based on a low-cost camera and an open-source computing platform to reduce human error, machining costs, and time. Compared with webcams, smartphones have advantages in terms of camera performance, data transmission, software resources, and computing performance. They are widely used in portable microscopic inspection [16] and tool-wear-type determination (<https://www.sandvik.coromant.cn/zh-cn/knowledge/machining-calculators-apps/pages/tool-wear-analyzer.aspx>). Structural vibrations can be analyzed from videos recorded by smartphones using motion alignment techniques in image processing. Gupta et al. [17] proposed a modal parameter identification method based on the high-speed photography function of a smartphone. André et al. [18] presented methods for performing advanced mechanical analyses by using smartphones. Using recorded video data, they estimated the instantaneous angular velocity of a rotating fan and the natural frequency of a cantilever beam structure.

The acoustic emission characteristics in the air can reflect the machine and tool conditions. Moreover, acoustic acquisition systems can be installed far from the workpiece owing to their sound propagation characteristics, thus providing a quick and easy method to diagnose faults [19, 20]. Smartphones can be developed as portable acoustic acquisition and fault diagnosis devices owing to their low cost and simple operation. Akbari et al. [21] used a smartphone to collect sound signals from tool strikes and identify the modal parameters of the holder-tool system. Xu et al. [22] proposed a mobile device-based fault diagnosis method for extracting fault-related features using quasi-tachometer signal analysis and envelope order spectroscopy. Huang et al. [23] proposed

an acoustic fault diagnosis method based on time-frequency analysis and machine learning for automotive power seats using a smartphone as an acoustic signal acquisition instrument. Sound signal is a standard signal used for process monitoring, especially for chatter detection. Chatter stability can be determined by extracting features related to the chatter component in the time, frequency, and time-frequency domains of the sound signal [20, 24–27]. The effectiveness of smartphones in chatter detection has not yet been explored. In contrast to traditional industrial microphones, technicians can use their smartphones for process monitoring and recording process parameters and experimental phenomena. Speech can be recognized as textual information by using open-source AI models. However, the human voice signal overlaps with the sound signal generated by the cutting process in the frequency domain, which may lead to false positives in chatter detection.

The tool status is related to the cutting position [27]. For example, chatter can be effectively detected by relying only on acoustic signals; however, traditional industrial sensors and instruments cannot obtain the tool position information related to chatter occurrence without collecting controller information. Smartphones can simultaneously collect multiple types of monitoring information by integrating multiple sensors. The combined measurement results obtained from different sensors can help technicians analyze machining process anomalies more accurately. This study proposes a position-oriented mobile monitoring method that utilizes image and sound information simultaneously. The main contributions of this study are as follows:

- 1) This study explored the effectiveness and extendibility of using smartphones for process monitoring at manufacturing sites. Compared to industrial sensors and acquisition instruments, the AI-based monitoring method can obtain machine operation information from the HMI without retrieving it from the machine tool controller, and enables the operator to record the experimental details easily.
- 2) Proven open-source AI and signal processing algorithms facilitates reduced development cycles and costs, which facilitate rapid deployment by SMEs or research institutions.
- 3) Simultaneously acquired video and sound signals are used to enable position-oriented process monitoring. The image contains the actual tool movement or tool position displayed on the HMI and can help establish an association between the sound signals and tool position.
- 4) Robotic milling and deep hole drilling experiments were conducted to verify the effectiveness of the proposed method at different manufacturing sites.

The remainder of this paper is organized as follows. In Section 2, a process monitoring scheme using a smartphone

was proposed, in which open-source toolkits and algorithms were used for optical character recognition (OCR), speech recognition, voice elimination, motion tracking, and tool status monitoring. In addition, the tool modal parameter was identified using the operational modal analysis (OMA). The sound signal, with the human voice removed, was used to calculate the energy ratio and time-domain characteristics to help the operator determine the chatter and tool wear status. In Section 3, the effectiveness of the proposed method was verified using robotic milling and deep-hole drilling experiments.

2 Materials and methods

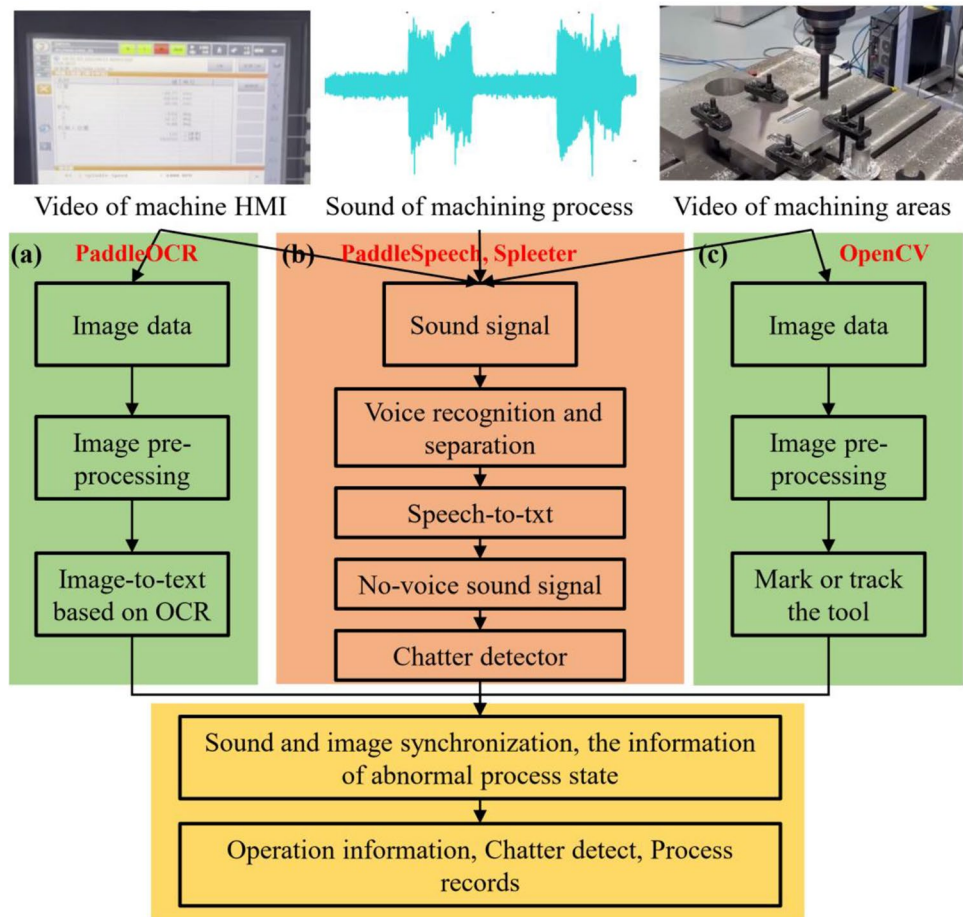
This section presents a smartphone-based position-oriented machine-tool monitoring method using open-source software and models written in Python. Open-source AI techniques have been used for image text recognition, speech recognition, human voice cancellation, and motion tracking. The time-domain characteristic and signal process methods of sound signals are used for modal parameter identification, chatter detection, and tool status determination. Figure 1

illustrates the process monitoring scheme based on a recorded video of the HMI or machining area during the machining process. The images and sound signals were captured simultaneously. Because the image contains processing position information, the sound signal and tool position can be correlated effectively. This helps the technician better understand the machining status with respect to different positions and modify the cutting parameters in the corresponding area.

2.1 Processing monitoring using AI methods

Various types of information, including the program position, spindle speed, and feed rate, can be extracted from the HMI using OCR technology. Subsequently, the sound signal is correlated with the identified program instructions over time. A mapping relationship was established between the sound signal and its image when using a smartphone to film the tool movement during machining. We can then locate the corresponding cutting area with respect to the monitoring anomaly. Tool position can be automatically identified via motion alignment for batch processing or long-term monitoring scenarios. The implementation method for each part is described as follows:

Fig. 1 Scheme of smartphone-based process monitoring



The image information in the video provides a visual record of the machine HMI and machining area information. The machine HMI displays the necessary information about the machine's operational status, which can be monitored by taking videos using a smartphone. As shown in Fig. 1a, the video is converted into an image, and subsequently, the appropriate area is selected using image-processing techniques. Subsequently, the HMI information was extracted using the OCR method. The OCR method can recognize the original information based on training data related to the recognized characters. In this study, PaddleOCR (<https://github.com/PaddlePaddle/PaddleOCR>) was chosen as the OCR algorithm, of which Baidu is the most accurate currently available free and open-source OCR engine. Note that when the operating panel characteristics are minor or light, fixing the handset next to the machine's operation panel is recommended to improve the recognition accuracy.

The cutting tool status and position were recorded and analyzed using a smartphone to film the machining area. The magnification adjustment function of the video recording allows the filmmaker to observe the cutting area from afar. As shown in Fig. 1c, the tool position can be marked in a video or image using a target recognition algorithm. It should be noted that target tracking is not required in most cases. Target tracking can help the operator discover the tool position more visually when the phone is fixed to monitor batch processing. In this study, we used the CascadeClassifier with OpenCV (https://docs.opencv.org/4.x/d6/d00/tutorial_py_root.html) for target recognition using Hal and local binary features as recognition features. The images were grayed and resized to improve the training and recognition efficiency. Note that when the operator takes a video using a smartphone, the video will shake, thereby affecting the recognition results. In addition, environmental factors such as lights, fixtures, and tools at the experimental site reduce the recognition accuracy.

The sound signal was captured simultaneously during video recording, which is a commonly used process for monitoring signals. In contrast to traditional industrial sensors and acquisition systems, technicians can collect sound signals while recording experimental details using smartphones. Therefore, speech recognition is required to eliminate the influence of human voices on process monitoring. Figure 1b shows the process of speech recognition and human voice separation using open-source software. Speech recognition is a widely used artificial intelligence (AI) technology. The open-source toolkit, PaddleSpeech (<https://github.com/PaddlePaddle/PaddleSpeech>), and its lightweight speech recognition model were used for speech recognition in this study. Large recognition errors may occur because the model is trained without considering noise during processing. Because the frequency range that contains the human voice (< 1000 Hz) overlaps with the signal frequency of the cutting process, system misclassification may occur if the human voice signal is not eliminated.

This study used Spleeter (<https://github.com/deezer/spleeter>) to separate human voice from noise, which is an open-source separation library and pre-trained model developed by Deezer, to separate human voices from noise. Note that the eliminated vocal sound signal still contains residual human vocal signals, which may interfere with process monitoring. Therefore, we marked the location of the vocal signal to help the technician perform a better assessment. A human voice signal separated from the original signal cannot be used for speech recognition, mainly because the necessary information is removed during the separation process.

The video and audio signals captured synchronously using a smartphone are aligned in time, and the number of points in the sound signal corresponding to the video signal for N_s frames can be calculated using the following equation:

$$N_s = \frac{N_v f_s}{f_{s_v}}, \quad (1)$$

where N_v is the number of video frames. f_{s_v} and f_s are the sampling frequencies of the video and sound signals, respectively.

2.2 Modal analysis and chatter detection using sound signal

Figure 2 describes the scheme of tool modal parameter identification and chatter detection. The operator determines the time and position of chatter occurrence based on the sound signals and images collected by the smartphone.

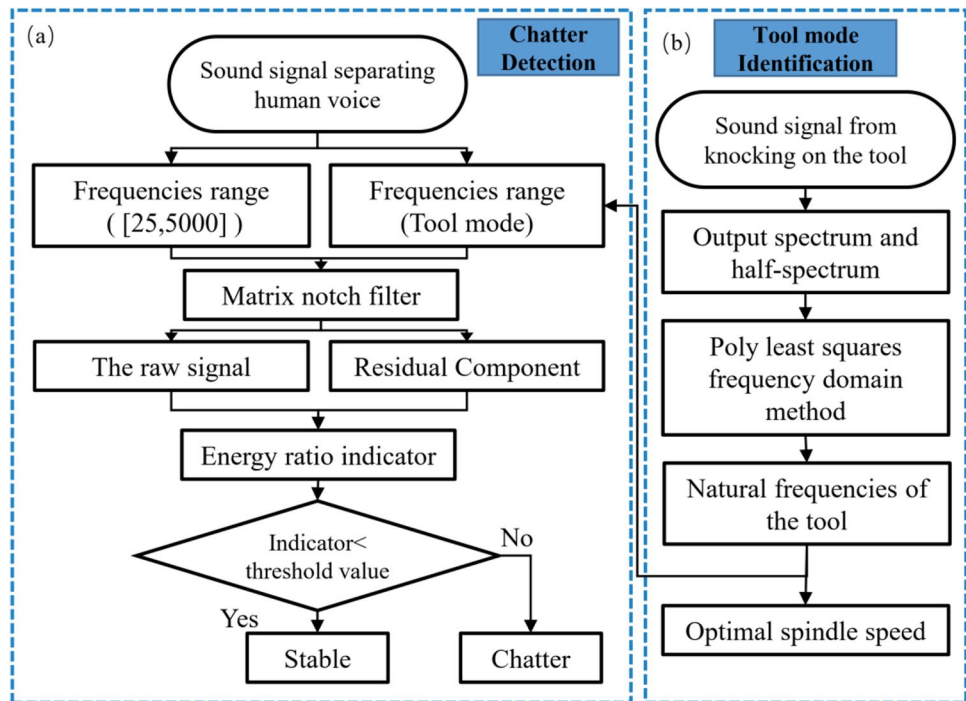
2.2.1 Identification of tool modal parameters by OMA

Tool mode parameters can help the operator select appropriate cutting parameters to avoid chatter. However, the modal hammer test requires specialized acquisition instruments and analysis software. This paper proposes a tool modal identification method based on sound signals recorded by a smartphone, as shown in Fig. 2b. First, the tool tip was struck using a metal tool commonly used at the machining site, and the sound of the strike was captured using a smartphone. The sound signal contains information about the tool tip vibration; therefore, the natural frequency and damping ratio can be estimated using the OMA method. OMA is a modal parameter estimation method that uses only the output signal. Accordingly, the relationship between the output power spectral density (PSD) matrix of the sound signal and the frequency response function matrix $\mathbf{H}(\omega)$ of the tool tip is obtained as follows:

$$S_{yy}(\omega) = H^*(\omega)S_{xx}(\omega)H(\omega)^T \quad (2)$$

where $S_{xx}(\omega)$ is the PSD matrix of excitation, which is the diagonal matrix of white-noise excitation with constant real

Fig. 2 Scheme of tool modal parameters identification and chatter detection



values. To reduce the computational effort, a modal decomposition equation based on the semi-spectral method was used, which is expressed as follows:

$$S_{yy}^+(\omega) = \sum_{r=1}^{N_m} \frac{\Phi_r g_r^T}{j\omega - \lambda_r} + \frac{\Phi_r^* g_r^H}{j\omega - \lambda_r^*} \quad (3)$$

where Φ_r , g_r , and λ_r are the vibration mode, operating factor, and stability pole, respectively. Finally, the modal parameters were identified through a poly-reference least-squares frequency domain method (p-LSCF); a more detailed explanation of the p-LSCF method is described in [28]. The least-squares method was used to solve the equations and determine the model parameters during the calculation. The natural frequency and damping ratio are calculated as follows:

$$\omega_r = \frac{|\lambda_r|}{2\pi}, \xi_r = -\frac{\text{Re}(\lambda_r)}{|\lambda_r|} \quad (4)$$

The physical and computational poles of stability were determined according to the stability criteria. Subsequently, the modal parameters were automatically extracted using a clustering approach. The frequency response function estimation results obtained using the OMA method cannot be directly used for chatter stability prediction because of the lack of an input force signal. However, the operator can select an appropriate spindle speed based on the identification results to ensure that the spindle rotation and harmonic frequencies avoid the

natural frequency of the tool tip. This is indicated by the remainder of the selected tool tip natural frequency and spindle rotation frequency, as shown in the following equation:

$$\text{Inx} = \text{mod}\left(\frac{\omega_{\text{tool},r}}{\omega_{\text{sp}}}\right), \quad (5)$$

where $\omega_{\text{tool},r}$ is the r th-order natural frequency of the tool tip and ω_{sp} is the spindle rotation frequency. The threshold value is determined empirically; typically, the first three orders of tool tip frequency are used to determine the spindle speed.

2.2.2 Chatter detection method

The sound signal is effective for detecting cutting chatter. Industrial microphones can be positioned away from the cutting area to collect sound signals, which are analyzed using specialized acquisition systems and software. In this study, a smartphone was used as the sound acquisition instrument. The sound signal during the cutting process was recorded next to the machine, and was composed of

$$x_{\text{raw}} = x_{\text{chatter}} + x_{\text{period}} + x_{\text{voice}} + x_{\text{noise}}, \quad (6)$$

where x_{period} is the periodic component associated with spindle rotation and is the dominant signal during stable cutting. x_{voice} is a human voice, which has not been considered in other chatter detection studies. x_{chatter} and x_{noise} are the chatter and noise components of a signal, respectively. The

chatter frequency associated with the natural frequency of the spindle tool system appears in the milling signal when the chatter occurs. Moreover, the energy of the chatter frequency increases dramatically as the chatter develops. The chatter detection procedure, based on the chatter mechanism, is illustrated in Fig. 2a.

First, the raw signal is filtered using a bandpass filter according to the modal parameters of the tool to reduce the effects of the high- and low-frequency components. The sampling frequency of the phone exceeded 40 kHz, whereas chatter generally occurred below 5 kHz. As mentioned above, the component near the natural frequency of the tool tip is the focus of the chatter detection. However, there are multiple frequencies in the vibration spectrum distributed around the central chatter frequency owing to the signal modulation. Moreover, multiple modes of tool tip frequency may cause multi-central chatter frequencies under certain spindle speeds. Hence, the cutoff frequency of the passband filter is determined based on the lowest and highest modal natural frequencies of the tool tip as follows:

$$\begin{cases} \omega_{lp} = \max(\omega_{tool,1} - 3.5\omega_{sp}, \omega_{sp}) \\ \omega_{hp} = \min(\omega_{tool,max} + 3.5\omega_{sp}, 5000) \end{cases}, \quad (7)$$

where ω_{lp} is the lowest cutoff frequency in the passband, and is determined by the first-order natural frequency of the tool tip $\omega_{tool,1}$ and spindle rotation frequency ω_{sp} . ω_{lp} is the highest cutoff frequency in the passband, and $\omega_{tool,max}$ is the highest order tool tip frequency.

In this study, a matrix notch filter was used to suppress the periodic signals in the acquired signal to reduce interference with chatter detection [29]. A matrix notch filter is typically used to suppress specific frequencies in a signal and is employed to eliminate narrowband interference while keeping the broadband signal unchanged. The output signal filtered by the notch filter can be expressed as follows:

$$\mathbf{x}_{res} = \mathbf{H}_{\omega_s} \mathbf{x}_{raw,tool}, \quad (8)$$

where $\mathbf{x}_{raw,tool}$ denotes the input signal obtained by using the designed bandpass filter. \mathbf{x}_{res} is the output signal, and is called the residual signal. \mathbf{H}_{ω_s} is a matrix trap filter characterized by the following constraints:

$$H_{\omega_s} V(\omega) = \begin{cases} 0 & \omega = \omega_s \\ V(\omega) & \omega \neq \omega_s \end{cases}, \quad (9)$$

where the vector $\mathbf{V}(\omega)$ is defined as $V(\omega) = [1, e^{j\omega}, e^{j2\omega}, \dots, e^{j(N-1)\omega}]^T$ and the passband of the filter is $P = [0, \omega_s - \varepsilon] \cup [\omega_s + \varepsilon, \pi]$, where ω_s is the trap frequency, which is related to the

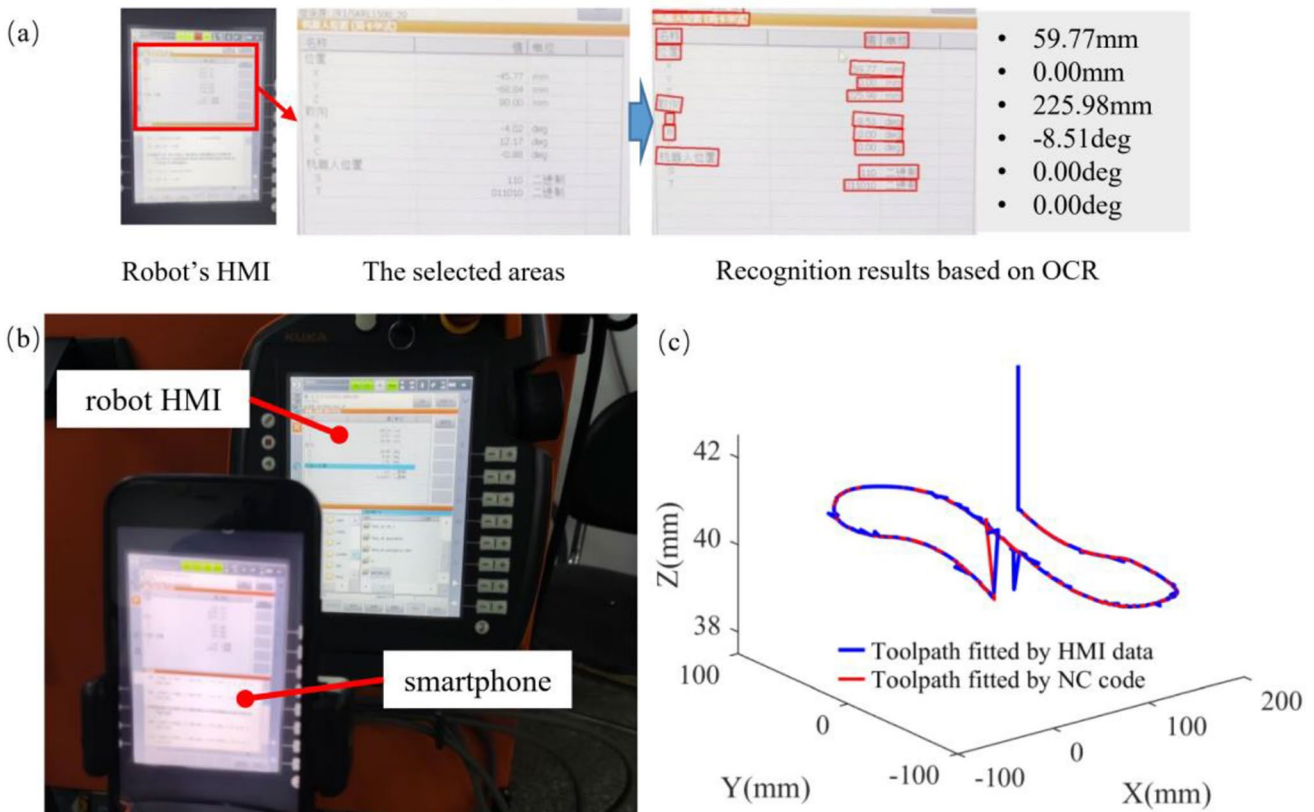


Fig.3 Process monitoring based on the recorded HMI video: **a** image recognition procedure, **b** the fixed smartphone was used to record robot HMI, **c** comparison between the toolpath fitted by NC code and that fitted by measured values

rotation and harmonic frequencies. ε is a small positive-valued constant. A detailed design of the filter is presented in [29].

Subsequently, the raw signal $\mathbf{x}(n)_{raw}$ is filtered using a bandpass filter within a passband frequency range of [25–5000 Hz], thereby eliminating the direct component and high-frequency components to obtain the signal $\mathbf{x}(n)_{raw,bp}$.

Finally, the ratio of the residual signal energy to the total signal energy is used as an indicator for chatter detection using:

$$ER = \frac{\sum_{n=1}^N x(n)_{res}^2}{\sum_{n=1}^N x(n)_{raw,bp}^2}. \quad (10)$$

The indicator determines the threshold value for chatter occurrence through cutting experiments. When the energy ratio of the sound signal exceeded this threshold, the algorithm adjudicated the occurrence of chatter. According to the results of the milling experiments, the chatter detection threshold in this study was 0.15. To improve the frequency resolution, six image frames and the corresponding sound signals were used to calculate the chatter indicator.

3 Experiments and results

3.1 Process monitoring in robotic milling

An idle operation test was conducted using the KUKA KR 160 robot to verify the effectiveness of HMI monitoring. Figure 3a shows the results of text recognition based on the OCR method using the selected HMI area. As shown in Fig. 3b, the smartphone was mounted on a mobile stand to record the robot HMI during robotic operation. A comparison between the toolpath fitted by the NC program and that fitted by the recognition values is shown in Fig. 3c. Text-recognition errors cause errors in some areas of the recognition curve. Because the operator must hold the operation panel in robotic milling, which is not conducive to photographing the HMI with a phone for text recognition, subsequent milling experiments were conducted using a smartphone to record a video of the machining area. A computer (Intel Core i7-6500U 2.5 GHz, 16 GB RAM) with Python software was used to process the signals.

Milling tests were performed using a KUKA KR500 MT-2 robot equipped with a mechanical spindle (Fig. 4). A Sandvik R390-022A20L-11L milling cutter, mounted with two inserts, was used in the experiments. The insert is T3 16M-KM H13A; the detailed tool parameters are listed in Table 1.

The raw sound signal of the knock and its PSD are shown in Fig. 5. The sound of a metal rod knocking on the tool tip was captured using a Redmi K40 phone. To reduce the

estimation error, data collected from five knocking tests were used for the OMA. The natural frequencies and damping ratios in the different directions of the tool can be neglected so that only the sound signals when tapping the tool in the x-direction are collected. Figure 6 shows the stabilization diagram of the OMA, where the natural frequency and damping ratio were automatically calculated based on the stability pole selection criteria. The dominant mode is typically used for chatter prediction. However, it is difficult to determine the tool mode with the lowest modal stiffness, owing to the lack of mass scaling in the OMA method. Therefore, this study chose the multiorder mode of the tool for chatter detection. The identification results of the tool are listed in Table 2.

The feed direction and tool path are shown in Fig. 4. The milling test contained two toolpaths: slot milling and side milling. The milling experiments on aluminum alloys with the experimental parameters are listed in Table 3. The spindle speed was 4000 rpm, and the feed per tooth was 0.225 mm. Note that the 5th-order tool natural frequency was chosen as the highest cutoff frequency $\omega_{tool,max}$ in Eq. (9) because of the low spindle speed.

Figure 7a presents a comparison of the raw signal with the separated vocal signal in test no. 5. The raw sound signal contained three vocal signal regions, marked as V1, V2, and V3. The sound signal of V1 can be used to accurately identify the cutting parameters and corresponding textual

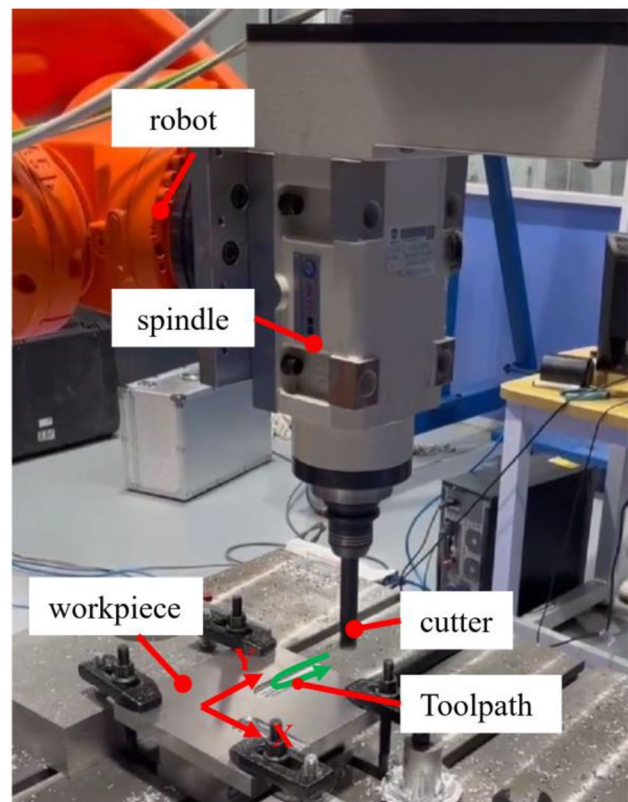


Fig. 4 Robotic milling experiment setup

Table 1 Parameters of milling cutter and insert

Diameter of cutter (mm)	Number of teeth	Overhang (mm)	Cutting edge angle (°)	Corner radius (mm)
22	2	120	90	1.6

removed according to the method proposed in [30]. Severe chatter occurred during period 28.6–30.2 s.

Figure 8 shows the procedure of the method proposed in this study for test no. 2. An analysis of comparison between the residual and raw signals is shown in Fig. 8a. According to the chatter indicator curve shown in Fig. 8b,

Fig. 5 Tool modal testing: **a** raw sound signal and **b** power spectral density of the sound signal

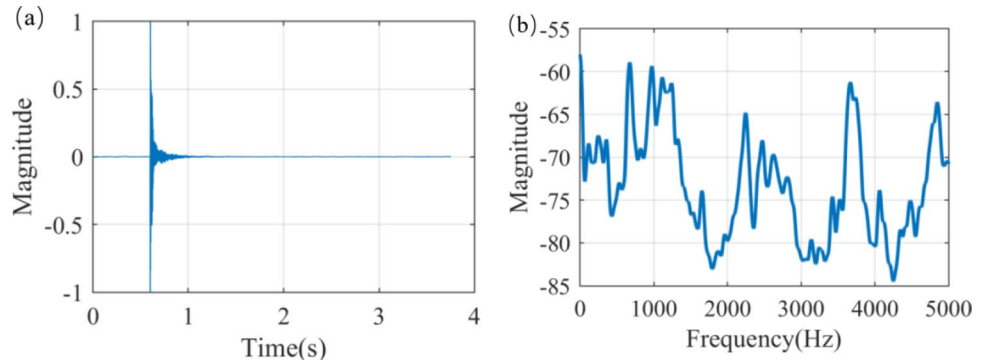


Fig. 6 Stabilization diagram for modal identification

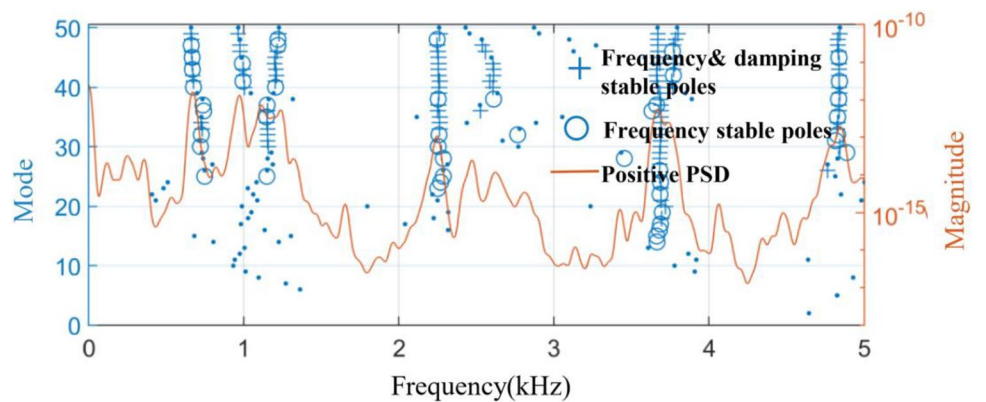


Table 2 Tool modal parameters

Natural frequency (Hz)	674.74	975.94	1109.21	1242.22	2239.86	3675.78	4841.98
Damping ratio	0.034	0.028	0.038	0.032	0.015	0.013	0.010

content. By contrast, the sound signals in the V2 and V3 regions could not be recognized as statements with a clear meaning. It should be noted that spindle rotation and noise from the milling process can interfere with the recognition results, particularly if the operator speaks at a fast speed. The failure in speech recognition is related to spindle rotation noise and unintelligible speech of the operator during the cutting process. Figure 7b shows the chatter index calculated from the sound signal, in which the spindle noise was

the tool has the most severe chatter at 5.6 s (P1), while the chatter disappears at 6.6 s (P2). Figure 8c presents the FFT diagram for different frequency ranges at 5.6 s. The chatter frequencies can be clearly observed and are related to the first-order natural frequency of the tool. The signal energy is mainly concentrated in 600–1000 Hz at the time. Figure 8d shows the tool position when the most severe chatter occurred and when the chatter disappeared, as determined by the chatter indicator curve.

Table 3 Cutting parameters and experimental settings

Test No.	Feed direction	Axial depth of cut (mm)	Number of cuts	Signal types	Phone model
1	Y	0.4	1	video and sound	iPhone 13
2	Y	0.6	1	video and sound	iPhone 13
3	Y	0.7	1	video and sound	iPhone 13
4	Y	0.8	1	video and sound	iPhone 13
5	X	1.0	2	video and sound	iPhone 13
6	X	0.8	1	sound	Redmi k40

Severe chatter occurred before the tool entered the corner and disappeared when the tool cut into the corner. There was no chatter during the side-milling process.

Fig. 9 compares the chatter indicator curves of tests no. 1, no. 3, no. 4, and no. 6. The maximum stability axial depth of the cut is 0.4 mm when the tool feeds along the y-direction. In contrast, the chatter is improved when the tool is fed along the x-direction. The feed directions in tests no.4 and no.6 were different. Comparing Fig. 9 a and b, tool chatter occurred when the depth of the cut was increased from 0.2 to 0.4 mm. Comparing Fig. 9 c and d, severe chatter occurred in test no. 4 during the slot milling process, which means that the chatter component was the main sound signal component. A lower chatter level was observed in test no.6 compared to test no. 4. The severe chatter occurred during the period 7.8–8.2 s in test no. 6. However, it is impossible to determine the tool position when the chatter occurred because only the sound signal was recorded.

The corresponding tool positions were extracted from the video recordings according to the initiation and disappearance of chatter, as shown in Fig. 9c. Figure 10 indicates that chatter occurs when the tool starts to cut into the workpiece and continues until

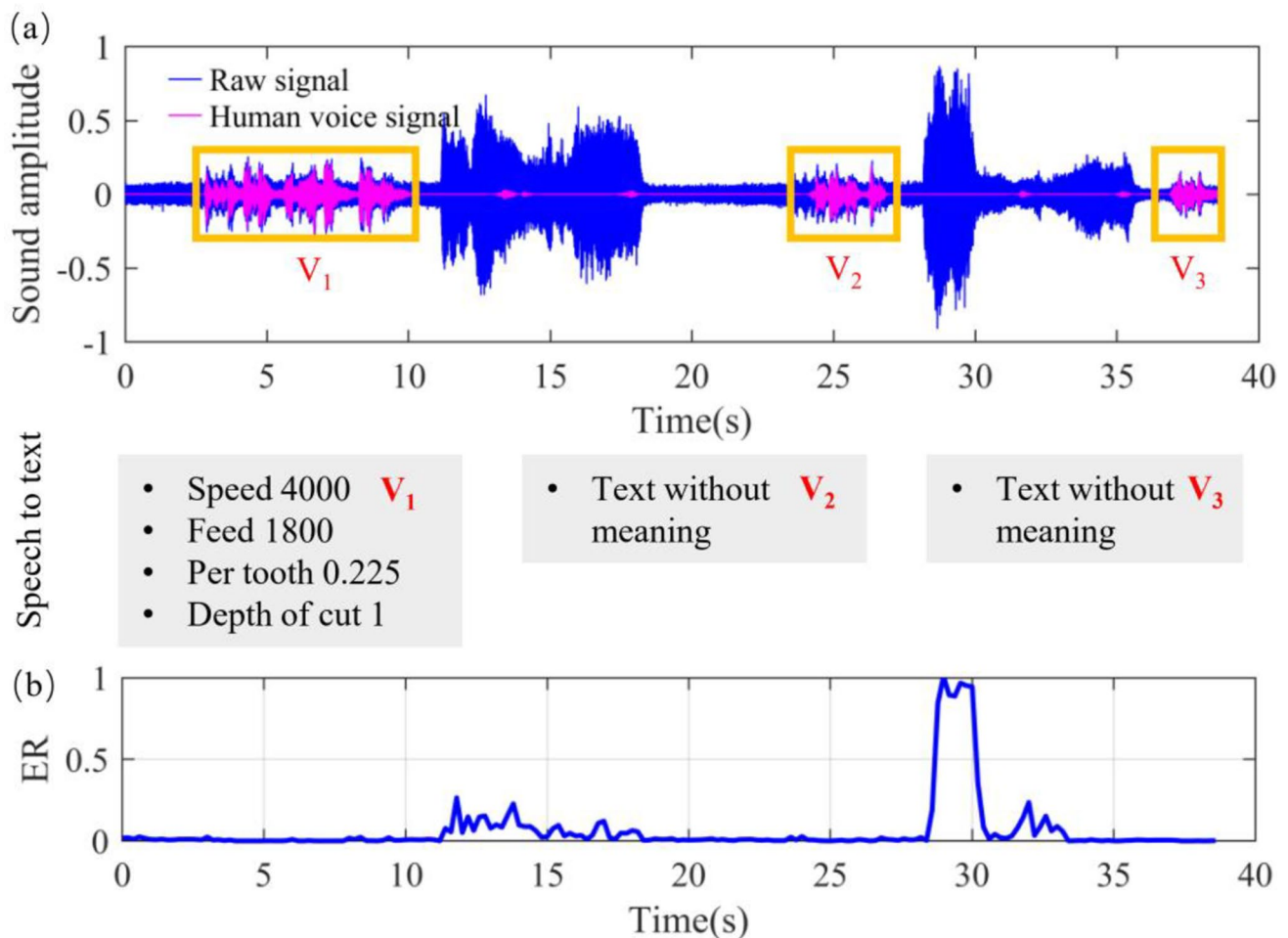


Fig. 7 Sound signal and chatter indicator curve in test no. 5. **a** Comparison of raw sound signal and human voice signal, the result of speech recognition; **b** chatter indicator curve

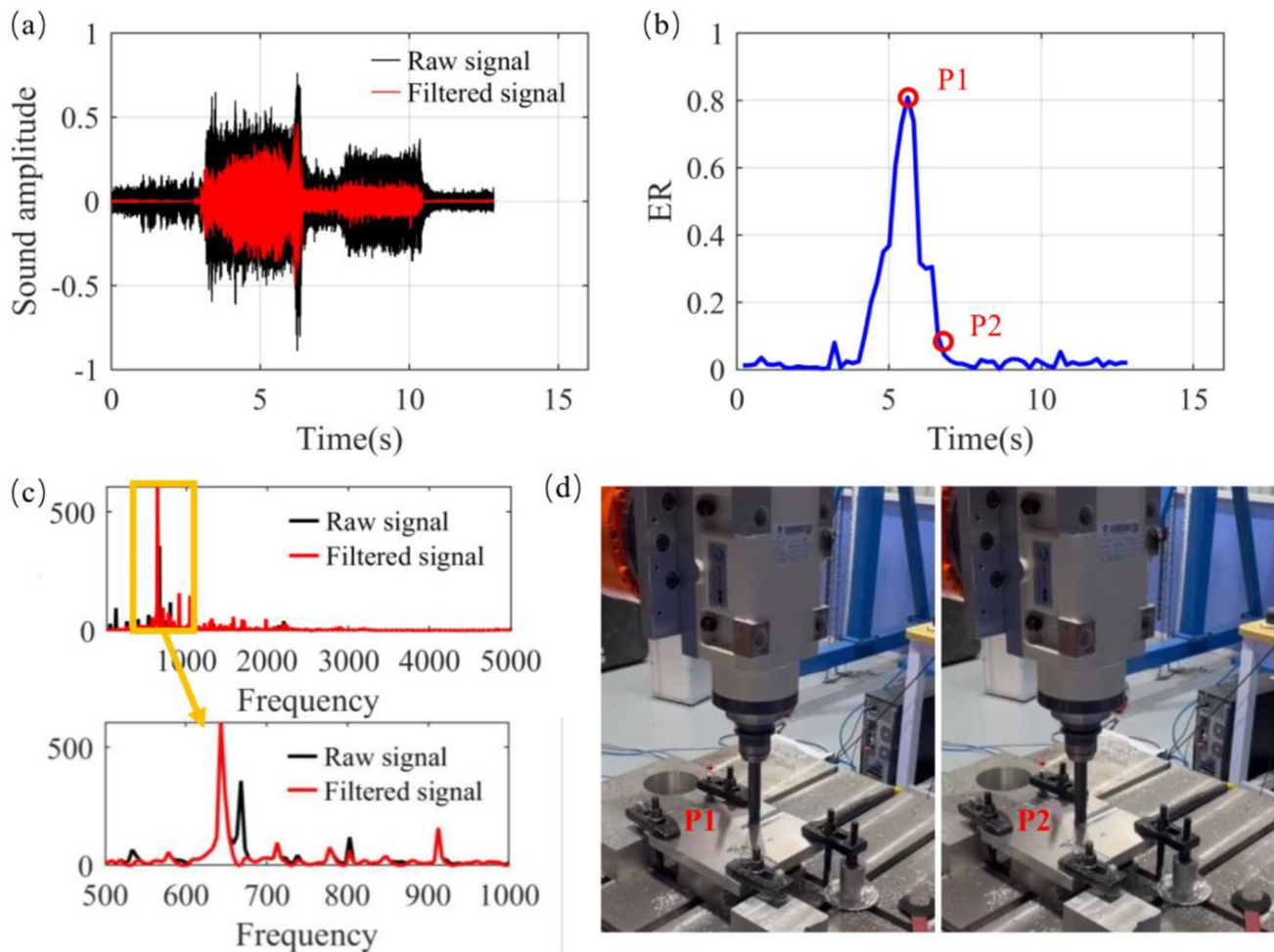


Fig. 8 Chatter detection associated with the cutting position of the tool in test no. 2: **a** comparison between the raw signal and residual signal, **b** chatter indicator curve, **c** FFT at P1, and **d** tool position at P1 and P2

the tool cuts into the corner. In addition to the FFT diagram of the sound signal, the technician can detect the chatter by observing the surface quality. Compared with the machined surface without chatter, shown in Fig. 10b, the chatter marks on the surface of the workpiece used in test no.4 can be clearly observed in Fig. 10a. It extended from the cut-in area to the corner. Hence, the chatter marks in Fig. 11a demonstrate the accuracy of the position of the chatter occurrence, as shown in Fig. 10.

The aforementioned experiments proved the effectiveness of the position-oriented chatter detection method proposed in this study. Fusing the image and sound signals can help the operator discover the tool position related to chatter occurrence.

3.2 Process monitoring in the deep hole drilling process

To verify the method proposed in this study, the machine HMI shown in Fig. 12 was filmed in a factory environment

using a smartphone during the deep-hole drilling process. The machine could drill three holes simultaneously.

The machining object was a composite sheet of Inconel 625 and FeCr alloy. The spindle speed was 95 rpm during the drilling of Inconel 625 and was increased to 500 rpm as the drilling depth increased. Figure 13 shows the identification process of the machine HMI, which displays the spindle speed, feed speed, tool position, axial thrust, spindle power, and other information for each spindle. The recognized failure rate related to machining information with a small font was high; therefore, this study mainly recognized the spindle power and thrust values. The spindle speed displayed on the HMI was determined based on the programmed spindle speed.

Experienced technicians can assess the tool wear state based on the sound and workpiece vibrations during the cutting process. Figure 14a shows the sound signal collected from the deep-hole drilling processing site, which contains spindle sound, cutting sound, cooling system sound, and human voice. In addition, it is difficult to

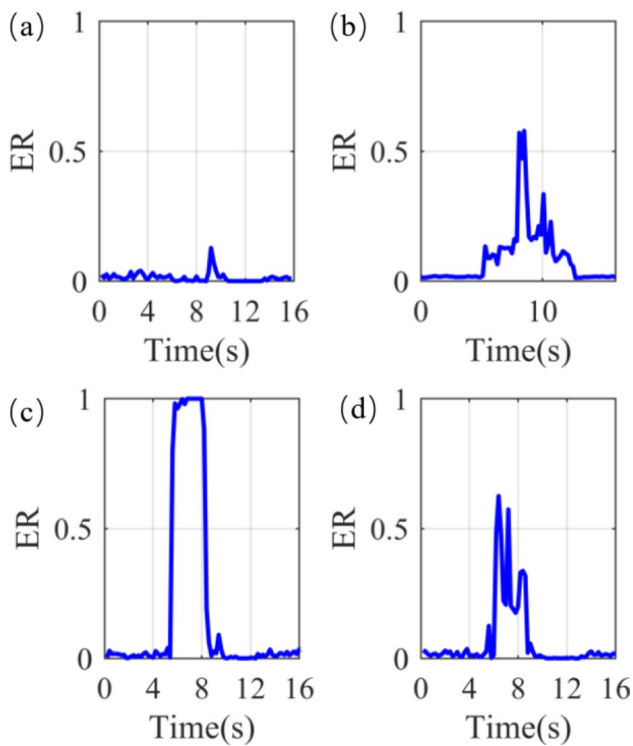


Fig. 9 Chatter indicator curve in tests. **a** no. 1, **b** no. 3, **c** no. 4, and **d** no. 6

isolate the cutting sound signal related to each spindle accurately when the three spindles operate simultaneously. In this study, the raw signal was processed using a bandpass filter with a set passband frequency range [25–1500] Hz, and the filtered sound energy was used as the reference data. Figure 14b shows a comparison

between the spindle speed and signal energy curve during the drilling process. It should be noted that the signal energy curve was calculated using the filtered signal after removing vocals. The spindle speed had the greatest effect on sound. Moreover, the sound energy curve increased abruptly at 320 s, which may be related to the tool wear state.

The thrust signals of the three spindles are compared in Fig. 15, where the z-directional thrust force is primarily related to the cutting edge. Based on the thrust signals, we determined that the thrust forces of the three spindles were higher when drilling the sheet of Inconel 625, and lower when the cutter cut into the sheet of the FeCr alloy. The thrust force of tool 3 was higher than that of tools 1 and 2 owing to the more severe tool wear detected in tool 3. The thrust force of tool 3 suddenly increased at 275 s, which may have been caused by chipping of the tool cutting edge.

Figure 16 presents a comparison between the power curves of the three spindles during the drilling process. Exhibiting a trend similar to that of the thrust signal, the spindle power signals were higher when machining Inconel 625 and gradually decreased to steady values after the tools started to cut into the sheet of FeCr alloy. The spindle power required by tool 3 was higher than that required by tools 1 and 2. The exception is that Fig. 15 does not depict an abrupt change in the spindle power curve for tool 3. The spindle power requirement is primarily related to the radial cutting forces caused by the guide pad. Therefore, by comparing the thrust and spindle power curve tool 3, it can be observed that the cutting edge was severely worn or chipped when the guide block was in a normal wear condition. Tool 3 was replaced

Fig. 10 Tool positions captured from the video recorded during test no. 4: **a** tool position when chatter occurs and **b** tool position when chatter disappears

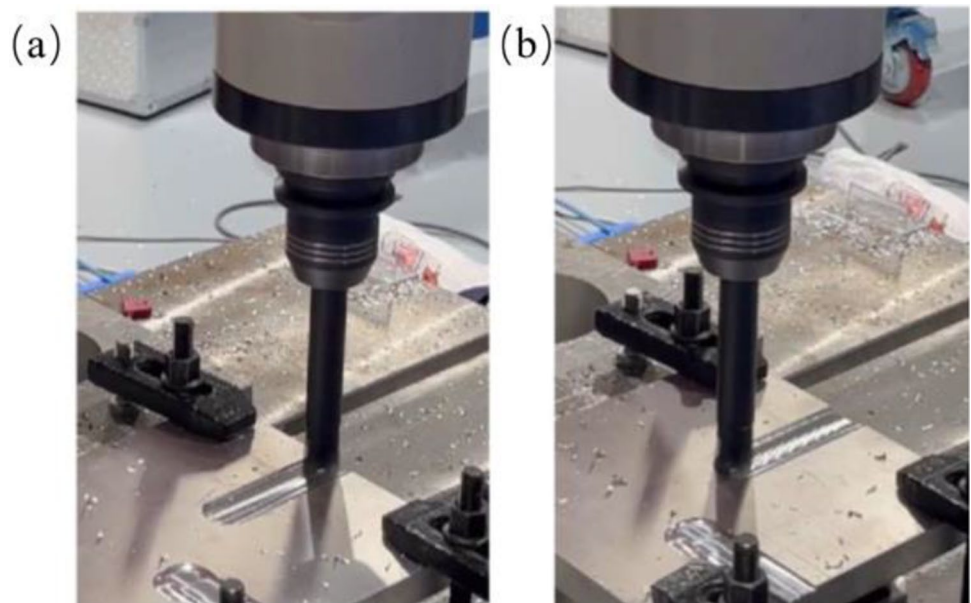


Fig. 11 Experimental surface characteristic: **a** surface characteristic after test no. 5 and **b** surface characteristic after stable milling



Fig. 12 **a** Deep hole drilling machine and **b** schematic structure of the drill piece

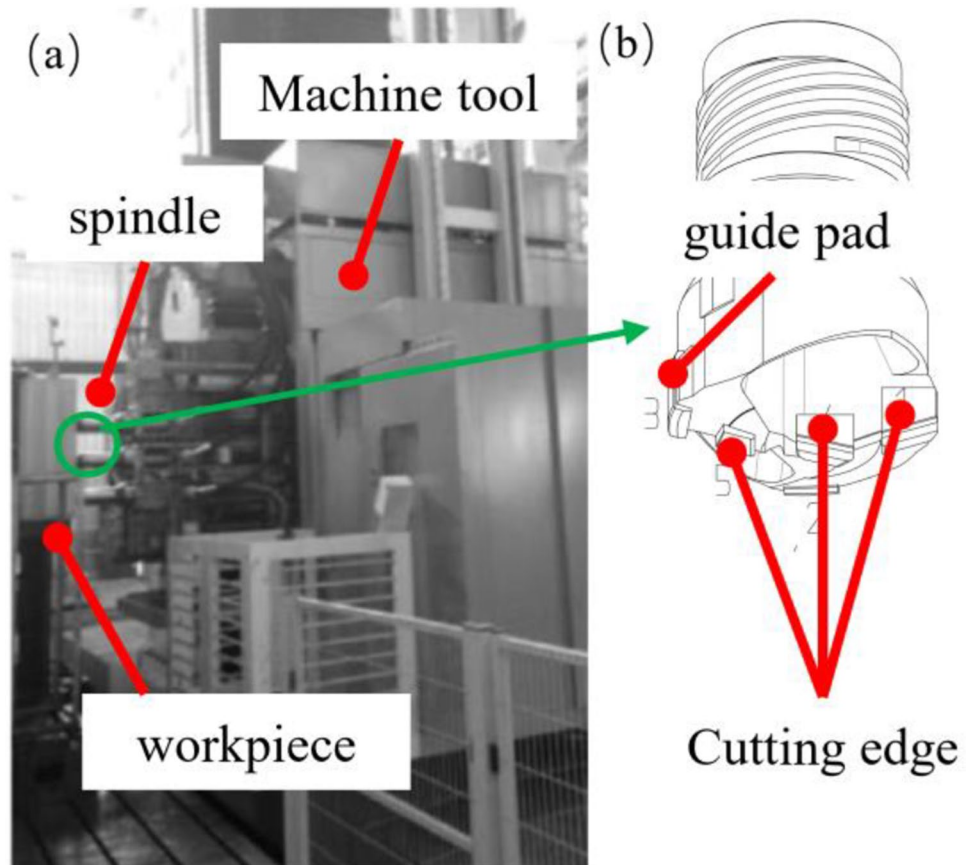


Fig. 13 Image recognition procedure of deep hole drilling machine HMI

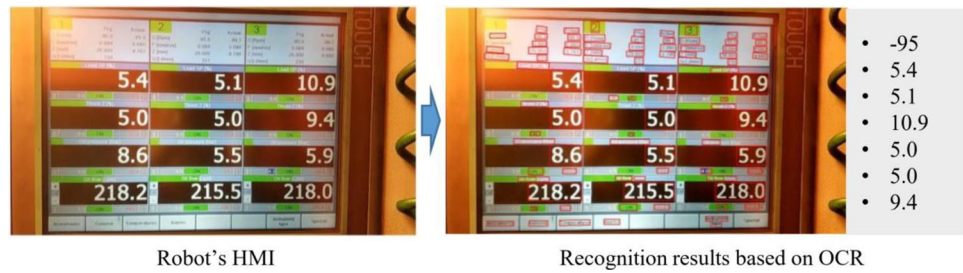


Fig. 14 Sound signal and spindle speed recognition results: **a** raw sound signal and human voice signal and **b** synchronization of spindle speed and sound signal energy

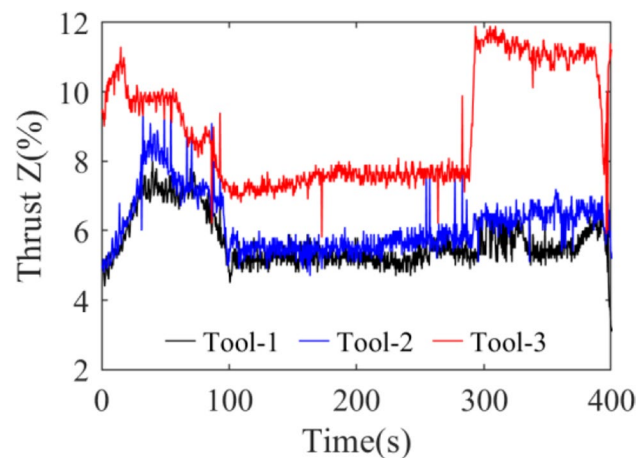
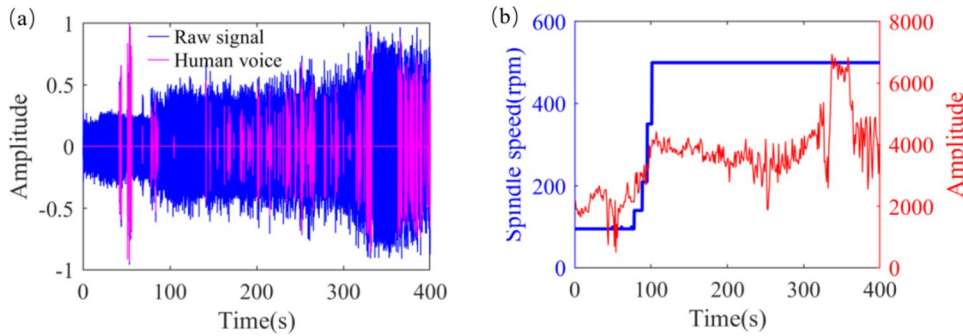


Fig. 15 The z-directional thrust values of the three machine spindles during the drilling process

by the operator because of severe wear after machining, whereas tools 1 and 2 remained in use for drilling.

4 Conclusions

This study explored the effectiveness of using smartphones for position-oriented process monitoring. The proposed method uses recorded videos of the HMI or machining area during the machining process to determine the tool status based on AI and signal processing techniques. First,

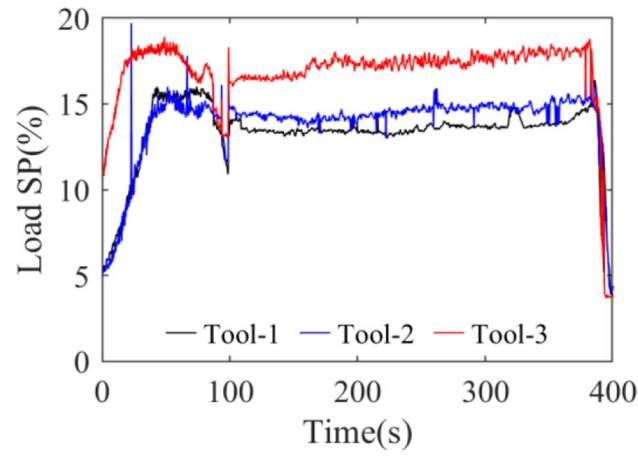


Fig. 16 Spindle power curves during the drilling process

information on the tool position and machine tool status was obtained through HMI recognition based on PaddleOCR. Then, the open-source toolboxes PaddleSpeech and Spleeter were used for speech recognition and human voice elimination based on the simultaneous acquisition of sound signals. In addition, sound signals were used for modal identification by oma, and the residual sound signal, which was obtained after eliminating the human voice and periodic components, was used to calculate the chatter indicator.

Finally, robotic milling and deep hole drilling experiments were conducted to validate the effectiveness of process monitoring using a smartphone. The simultaneous

acquisition and analysis of tool motion images and sound signals in robotic milling enables tool-position-oriented chatter detection. The chatter marks on the workpiece surface match the extracted tool position. The status information of the robot and machine tool can be efficiently obtained from the HMI video based on the proposed methods. For deep hole drilling machine tools, spindle speed, thrust, spindle power, and acoustic signals are available from the HMI videos. Operators can use monitoring data to visually analyze the wear status of different tools. In summary, mobile phone-based process monitoring helps operators gain insight into tool status during cutting without changing the machine control cabinet.

In the future, a smartphone application will be developed based on the proposed method, which can be used for signal collection, data processing, and chatter detection.

Author contributions KND, YL, and DG conceived and designed the study. KND, QHG, FLW, and SDM performed the experiments. KND and CZ wrote the paper. KND, YL, and DG reviewed and edited the manuscript. All authors have read and approved the manuscript.

Funding This study was supported by the National Key Research and Development Project of China (2018YFB1306803).

Data availability Publication-related datasets are available from the corresponding author upon reasonable request.

Code availability (software application or custom code) Not applicable.

Declarations

Ethics approval (include appropriate approvals or waivers) Not applicable.

Consent to participate (include appropriate statements) Not applicable.

Consent for publication (include appropriate statements) All authors approved the manuscript and gave their consent for submission and publication.

Competing interests The authors declare no competing interests.

References

- Wang WK, Wan M, Zhang WH, Yang Y (2022) Chatter detection methods in the machining processes: a review. *J Manuf Process* 77:240–259. <https://doi.org/10.1016/j.jmapro.2022.03.018>
- Yang K, Wang G, Dong Y, Zhang Q, Sang L (2019) Early chatter identification based on an optimized variational mode decomposition. *Mech Syst Signal Process* 115:238–254. <https://doi.org/10.1016/j.ymsp.2018.05.052>
- Liu D, Luo M, Pelayo GU, Trejo DO, Zhang DH (2021) Position-oriented process monitoring in milling of thin-walled parts. *J Manuf Syst* 60:360–372. <https://doi.org/10.1016/j.jmsy.2021.06.010>
- Qurat ul ain Z, Mohsan SAH, Shahzad F, Qamar M, Qiu BS, Luo ZF, Zaidi SA (2022) Progress in smartphone-enabled aptasensors. *Biosens Bioelectron* 114509. <https://doi.org/10.1016/j.bios.2022.114509>
- Grossi M (2019) A sensor-centric survey on the development of smartphone measurement and sensing systems. *Measurement* 135:572–592. <https://doi.org/10.1016/j.measurement.2018.12.014>
- Staacks S, Hütz S, Heinke H, Stampfer C (2018) Advanced tools for smartphone-based experiments: phyphox. *Phys Educ* 53(4):045009
- Wang L, He H, Li S (2022) Structural vibration performance test based on smart phone and improved comfort evaluation method. *Measurement* 111947. <https://doi.org/10.1016/j.measurement.2022.111947>
- Rodríguez A, Sañudo R, Miranda M, Gómezc A, Benaventea J (2021) Smartphones and tablets applications in railways, ride comfort and track quality. *Trans Zones Anal Meas* 182:109644. <https://doi.org/10.1016/j.measurement.2021.109644>
- Yu Y, Han R, Zhao X, Mao X, Hu W, Jiao D, Li M, Ou J (2015) Initial validation of mobile-structural health monitoring method using smartphones. *Int J Distribut Sensor Netw* 11(2):274391. <https://doi.org/10.1155/2015/274391>
- Castellanos-Toro S, Marmolejo M, Marulanda J, Cruz A, Thomson P (2018) Frequencies and damping ratios of bridges through operational modal analysis using smartphones. *Constr Build Mater* 188:490–504. <https://doi.org/10.1155/2015/274391>
- José Álvares A, Oliveira LES, Ferreira JCE (2018) Development of a cyber-physical framework for monitoring and teleoperation of a CNC lathe based on MTconnect and OPC protocols. *Int J Comput Integr Manuf* 31(11):1049–1066. <https://doi.org/10.1080/0951192X.2018.1493232>
- Liu C, Zheng P, Xu X (2021) Digitalisation and servitisation of machine tools in the era of Industry 4.0: a review. *Int J Prod Res.* <https://doi.org/10.1080/00207543.2021.1969462>
- Kim H, Jung WK, Choi IG, Ahn SH (2019) A low-cost vision-based monitoring of computer numerical control (CNC) machine tools for small and medium-sized enterprises (SMEs). *Sensors* 19(20):4506. <https://doi.org/10.3390/s19204506>
- Lee W, Cheng H, Wei CC (2018) Development of a machining monitoring and chatter suppression device[C]//2018 IEEE Industrial Cyber-Physical Systems (ICPS). IEEE 404–408. <https://doi.org/10.1109/ICPHYS.2018.8387692>
- Xing K, Liu X, Liu Z, Mayer JRR, Achiche S (2021) Low-cost precision monitoring system of machine tools for SMEs. *Procedia CIRP* 96:347–352. <https://doi.org/10.1016/j.procir.2021.01.098>
- Bueno D, Munoz R, Marty JL (2016) Fluorescence analyzer based on smartphone camera and wireless for detection of Ochratoxin A. *Sensors Actuators B Chem* 232:462–468. <https://doi.org/10.1016/j.snb.2016.03.140>
- Gupta P, Rajput HS, Law M (2021) Vision-based modal analysis of cutting tools. *CIRP J Manuf Sci Technol* 32:91–107. <https://doi.org/10.1016/j.cirpj.2020.11.012>
- André H, Leclere Q, Anastasio D, Benaïcha Y, Billon F, Bireme M, Bonnardot F, Chin ZY, Combet F, Daems PJ, Daga AP, De Geest R, Elyousfi B, Griffaton J, Gryllias K, Hawwari Y, Helsen J, Lacaze F, Larocheb L, Li X, Liu C, Mauricio A, Melot A, Ompusunggu A, Paillot G, Passos S, Peeter C, Perez M, Qi J, Sierra-Alonso EF, Smith WA, Thomas X (2021) Using a smartphone camera to analyse rotating and vibrating systems: Feedback on the SURVISHNO 2019 contest. *Mech Syst Signal Process* 154:107553. <https://doi.org/10.1016/j.ymsp.2020.107553>
- Henriquez P, Alonso JB, Ferrer MA, Travieso CM (2013) Review of automatic fault diagnosis systems using audio and vibration signals. *IEEE Trans Syst Man Cybern Syst* 44(5):642–652. <https://doi.org/10.1109/TSMCC.2013.2257752>

20. Yue CX, Gao HN, Liu XL, Liang SY, Wang LH (2019) A review of chatter vibration research in milling. *Chin J Aeronaut* 32(2):215–242. <https://doi.org/10.1016/j.cja.2018.11.007>
21. Akbari VOA, Postel M, Kuffa M, Wegener K (2022) Improving stability predictions in milling by incorporation of toolholder sound emissions. *CIRP J Manuf Sci Technol* 37:359–369. <https://doi.org/10.1016/j.cirpj.2022.02.012>
22. Xu X, Li W, Zhao M, Hu H (2022) Mobile device-based bearing diagnostics with varying speeds. *Measurement* 200:111639. <https://doi.org/10.1016/j.measurement.2022.111639>
23. Huang X, Teng Z, Tang Q, Zhou Y, Hua J, Wang X (2022) Fault diagnosis of automobile power seat with acoustic analysis and retrained SVM based on smartphone. *Measurement* 202:111699. <https://doi.org/10.1016/j.measurement.2022.111699>
24. Schmitz TL (2003) Chatter recognition by a statistical evaluation of the synchronously sampled audio signal. *J Sound Vibr* 262(3):721–730. [https://doi.org/10.1016/S0022-460X\(03\)00119-6](https://doi.org/10.1016/S0022-460X(03)00119-6)
25. Weingaertner WL, Schroeter RB, Polli ML, Gomes JDO (2006) Evaluation of high-speed end-milling dynamic stability through audio signal measurements. *J Mater Process Technol* 179(1–3):133–138. <https://doi.org/10.1016/j.jmatprotec.2006.03.075>
26. Caliskan H, Kilic ZM, Altintas Y (2018) On-line energy-based milling chatter detection. *J Manuf Sci Eng* 140(11):111012. <https://doi.org/10.1115/1.4040617>
27. Gao J, Song Q, Liu Z (2018) Chatter detection and stability region acquisition in thin-walled workpiece milling based on CMWT. *Int J Adv Manuf Technol* 98(1):699–713. <https://doi.org/10.1007/s00170-018-2306-1>
28. Magalhães F, Cunha Á (2011) Explaining operational modal analysis with data from an arch bridge. *Mech Syst Signal Process* 25(5):1431–1450. <https://doi.org/10.1016/j.ymsp.2010.08.001>
29. Tao TJ, Zeng H, Qin C, Liu C (2019) Chatter detection in robotic drilling operations combining multi-synchrosqueezing transform and energy entropy. *Int J Adv Manuf Technol* 105(7):2879–2890. <https://doi.org/10.1007/s10845-019-01509-5>
30. Li X, Wan S, Huang XW, Hong J (2020) Milling chatter detection based on VMD and difference of power spectral entropy. *Int J Adv Manuf Technol* 111(7):2051–2063. <https://doi.org/10.1007/s00170-020-06265-y>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.