

Original Paper

Contactless Micron-Level Vibration Measurement with Millimeter Wave Radar

Renjie Wen¹, Dongheng Zhang^{1*}, Jinbo Chen¹ and Qibin Sun¹

¹*School of Cyber Science and Technology, University of Science and Technology of China*

ABSTRACT

Vibration measurements play a critical role in troubleshooting mechanical equipment and assessing the structural integrity of buildings. However, conventional vibration measurement methods rely on contact-based approaches, such as the attachment of accelerometers to the target object, leading to complex equipment deployment. Therefore, non-contact vibration measurement has attracted great attention but has yet to be fully addressed. In this paper, we propose DeepVib, a non-contact vibration measurement system that enables accurate micron-level vibration monitoring. First, we introduce a series of signal processing algorithms to extract the vibration object motion from mmWave reflection signals. Then, we design a deep neural network to effectively suppress noise interference and achieve outputting higher signal-to-noise ratio data. Finally, we eliminate static reflections with geometric-based method to recover the vibrations of the target. The experimental results show that our non-contact measurement method can accurately measure the vibration at the micron level with the average error of vibration frequency less than 0.1%. For the amplitude below 100 μ m, the median error of estimation is 7.23%. In addition,

*Corresponding author: Dongheng Zhang, dongheng@ustc.edu.cn.

Received 28 October 2023; Revised 16 January 2024

ISSN 2048-7703; DOI 10.1561/116.00000073

© 2024 R. Wen, D. Zhang, J. Chen and Q. Sun

DeepVib reduces the estimation error of 80th-percentile amplitude by 56.60% compared with the conventional method.

Keywords: Wireless sensing, Vibration measurement, deep neural network

1 Introduction

Vibration is a ubiquitous physical phenomenon that conveys crucial information. In industry, object vibrations reflect internal states, and by monitoring their vibration characteristics, amplitude and/or frequency, it is possible to detect damage or failure of equipment at an early stage [15, 1, 19]. In health care, the periodic movement of the human chest and heart is also a vibration that can reflect the health status of the human body, from which information about the body's breathing and heart rate can be mined, and furthermore, it can determine whether certain diseases are present and assist in medical diagnosis [16, 29]. Therefore, it is of great importance to accurately measure the vibration properties of an object.

Significant progress has been made in vibration sensing over the past few decades. Conventional vibration measurement systems utilize specialized sensors, including accelerometers and gyroscopes [22, 12, 21, 18]. These methods necessitate the attachment of sensors to the surface of the vibrating object, which can present practical challenges, such as motors operating in high-temperature environments. As an effort to overcome this limitation, non-contact approaches employing high-speed cameras [24] and lasers [17] have been investigated. Although they offer the advantage of capturing precise measurements of tiny vibrations, their implementation is constrained by costly deployments. Additionally, these methods are contingent upon specific conditions, such as optimal lighting conditions and Line-of-Sight (LoS). Ranging LIDAR [31, 23] presents a more cost-effective solution. However, its reduced resolution and sampling rate impede its applicability for capturing high-frequency micron vibrations. These factors further restrict the practical applicability of these techniques.

Recently, with the rapid advancement of wireless sensing technology, there has been a growing interest in utilizing radio signals for vibration measurement, primarily driven by the distinctive attributes of radio signals, such as their non-contact nature, privacy-friendly properties, and ability to penetrate obstacles. Notably, WiFi and RFID signals have been harnessed for tracking breathing and heartbeat [13, 30, 7, 25], leveraging their capacity to capture vibrational displacements. Furthermore, [28, 3] introduces the utilization of RFID technology for monitoring motor rotation frequency, particularly in noisy environments. However, the accuracy of these techniques is often limited

due to the long wavelength of the signals, making it challenging to accurately measure tiny vibrations.

In contrast, the mmWave radar, with its wavelength in the millimeters (e.g., 3.896 mm at 77 GHz), inherently delivers the necessary resolution to detect tiny vibrations accurately. Consequently, by analyzing the radar signal reflected from the target, both the vibration frequency and displacement can be precisely captured. Several recent studies have explored the utilization of mmWave in diverse sensing applications [26, 9, 6, 14]. However, the performance of these existing systems is still limited. In this paper, we present a mmWave system designed to achieve precise measurement of micron-level vibrations. Building such a mmWave system is non-trivial and there are mainly two challenges needed to be resolved: 1) Extracting vibration signal; 2) Alleviating noise and interference.

Vibration target signal extraction. In practical environments, there exist numerous stationary and moving objects alongside the vibrating targets. While millimeter waves exhibit good directionality, the received signal is usually a complex mixture of reflections from the sensed object and environmental reflectors, resulting in the interference of weaker energy target signals by various interfering signals. To address this issue, we leverage the periodic nature of vibrations, enabling the accurate localization of the target. The system effectively separates vibration signals from clutter signals, facilitating accurate and reliable measurement of vibrations amidst complex environments.

Alleviating the noise interference. The classical displacement-phase model suggests that shorter RF signal wavelengths provide finer perception granularity. While millimeter waves are more sensitive to small displacements compared to other wireless technologies, their effectiveness in detecting micron-level vibrations is limited. For example, with a millimeter wave wavelength of approximately 4 mm, an object with a 50 μm amplitude will cause a mere 0.33 rad phase change in the reflected signal. This small phase change, corresponding to less than 5% of a full circle's arc length, is highly susceptible to noise interference, resulting in inaccurate displacement estimates. To address this, we analyze the impact of target amplitude and noise through extensive simulation data and propose a targeted processing method to enhance the Signal-to-Noise Ratio (SNR). This involves generating simulation data based on the vibration reflection signal model, constructing a deep noise reduction network to learn the noise distribution from data features, and applying the trained network to reduce noise influence in real data estimation.

The main contributions of this paper can be summarized as follows.

- Firstly, to the best of our knowledge, DeepVib is the first work to leverage deep learning techniques for enhancing vibration sensing performance. Notably, DeepVib is trained exclusively using simulated data, eliminating the need for extensive dataset construction efforts.

- Secondly, DeepVib exhibits exceptional computational efficiency, resulting in minimal processing time during the testing phase. Particularly, DeepVib reduces processing time by approximately 80% compared to the frequency group algorithm.
- Thirdly, we conduct comprehensive validation of DeepVib using both simulated and real-world data. The results demonstrate significant improvement, with DeepVib achieving over 40% enhancement in measuring μm -level amplitudes based on the simulated data. Furthermore, despite being solely trained on simulated data, DeepVib achieves a mean amplitude error of 2.1 μm for 100 μm -amplitude vibrations when tested on real-world data.

The remaining sections of this paper are organized as follows. Section II provides an overview of the related work in the field of vibration sensing. In Section III, we introduce a vibration propagation model. The design of DeepVib is presented in Section IV. Section V describes the implementation details of our system, followed by a thorough experimental evaluation. Finally, we conclude the paper in Section VI, summarizing the key findings and discussing potential future directions for research.

2 Related Work

In this section, we introduce in detail the related work on vibration sensing.

RFID based vibration measurement. Several methods have been proposed to utilize RFID technology for vibration measurement. The initial attempt was made by TagBeat [28], which successfully measured the vibration frequency of sub-cm-level vibrations. Building upon this work, TagTwins [3] introduced a two-tag system to mitigate ambient noise. TagSound [11] leveraged the characteristics of harmonic backscatter to detect high-frequency vibration signals. Additionally, TagSMM [27] improved sensitivity by exploiting the coupling effect among tags, enabling sub-millimeter resolution. However, despite these advancements, the current RFID-based vibration sensing methods still fall short of meeting the requirements for achieving μm -level accuracy.

mmWave based vibration measurement. The mmWave technology, with its millimeter-level wavelength, offers enhanced sensitivity to weak vibrations, making it suitable for detecting small displacements. In [2], theoretical lower bounds for object amplitude and frequency estimation are analyzed, along with a simplistic estimation method. However, this method fails to address the practical issue of handling the DC component. To address this limitation, [16] introduces a geometric representation of the signal model in the In-phase and Quadrature (IQ) domain. They propose removing static

reflections using mean-based and fitting-based techniques to recover the object's amplitude. Nevertheless, when the vibration displacement is extremely small, susceptibility to noise becomes a concern. Gao *et al.* in [5] propose a radius correction technique to enhance small displacement accuracy, but it necessitates a lengthy calibration arc. In [8, 10], several methods are proposed to improve fitting accuracy and attenuate noise interference by combining multiple data with different carrier frequencies. However, these methods often require specialized hardware, such as digital phase shifters, and suffer from increased computational effort and longer recovery times due to the larger amount of data. These drawbacks pose challenges for real-time monitoring applications.

3 Theory and Estimation Challenges

3.1 Signal Model for Vibrating Targets

In this paper, we focus on the Frequency Modulated Continuous Wave (FMCW) mmWave radar, as illustrated in Figure 1. This radar operates by transmitting a chirp signal with a linearly varying spectrum across the bandwidth range. The mathematical representation of a single transmitted chirp signal is given as follows:

$$s(t) = A_t \exp \left[j2\pi \left(f_c + \frac{1}{2} K t^2 \right) \right], \quad 0 < t < T_r, \quad (1)$$

where A_t represents the magnitude associated with the transmit power, f_c denotes the carrier frequency, and $K = B/T_r$ is defined as the slope between the signal bandwidth B and the duration of a single chirp period T_r .

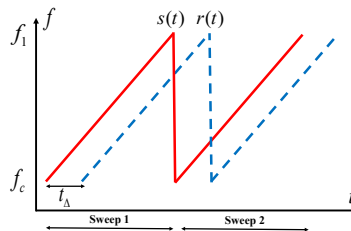


Figure 1: An illustration of FMCW waveform, where the solid line stands for TX signal and dashed line stands for RX signal.

Consider the scenario where a single small object is vibrating at position x_0 relative to the radar, resulting in a time-varying distance between the object

and the radar. The equations describing the vibration of the object and the distance from the object to the radar are given as follows:

$$\begin{cases} x(t) = A_v \sin(2\pi f_v t + \varphi_v), \\ X(t) = x_0 + x(t), \end{cases} \quad (2)$$

where $x(t)$ characterizes the vibration, A_v , f_v , φ_v denote the amplitude, frequency, and initial phase of the vibration target, respectively.

The received signal at the radar receiver is a delayed version of the transmitted signal, represented by $r(t)$, with a delay denoted as $t_\Delta = \frac{2X(t)}{c}$, which is the round-trip propagation time of the chirp. The c denotes the speed of light. Consequently, the intermediate frequency (IF) signal can be derived as follows:

$$\begin{aligned} y(t) &= s(t)r^*(t) \\ &\approx A_t A_r \exp \left[j4\pi(f_c + Kt) \frac{X(t)}{c} \right], \end{aligned} \quad (3)$$

A chirp contains reflected signals at multiple distances. Thus, we need to extract the signal from the right distance to enhance the estimation performance of the vibration. Since the duration of a chirp is typically short, we can assume that the vibrating object remains stationary during this time interval. To extract the target information, we apply Range-FFT on each chirp and combine the range bin results from different chirps to form a slow time sequence. The signal obtained from the range bin associated with the target distance can be mathematically expressed as follows:

$$y'(t) = a \exp \left[j4\pi f_c \frac{X(t)}{c} \right]. \quad (4)$$

From the Equation 4, we can see that the phase of $y'(t)$ encapsulates vital information regarding the target's vibration. The relationship between the phase and the amplitude can be derived as follows:

$$\theta(t) = 4\pi \frac{x_0 + x(t)}{\lambda}, \quad (5)$$

where the signal phase change is directly proportional to the object's displacement variation, forming the core formula for vibration measurement with FMCW radar. By quantifying the phase shift magnitude, we can calculate the object's amplitude.

3.2 Fitting-based Vibration Measurement

In practice, the received signal extracted from the range bin containing the vibrating target encompasses not only information about the target's motion

but also static reflections from surrounding objects and noise interference. Thus, the received signal can be expressed as follows:

$$\begin{aligned} S(t) &= a_0 \exp \left[j4\pi f_c \frac{X(t)}{c} \right] + \sum_{i=0}^n a_i \exp \left[j4\pi \frac{x_i}{c} \right] + w(t) \\ &= a_0 \exp \left[j4\pi f_c \frac{X(t)}{c} \right] + a' \exp \left[j4\pi \frac{x'}{c} \right] + w(t), \end{aligned} \quad (6)$$

where x_i is the distance from i -th stationary object to radar and $w(t)$ is the complex Additive White Gaussian Noise (AWGN). The second equality comes from the fact that the static reflections do not vary with time and thus can be combined together as an aggregated static reflection x' .

As shown in Figure 2, due to the vibration of the object, ideally the received signal $S(t)$ is located on an arc in the complex plane, where the vector \vec{OA} represents the aggregated static reflection. Through identifying the radius and the center of the arc, we can obtain the vibration information. However, in practical scenarios, noise affects the received signal, causing it to be distributed around the arc. Consequently, a direct approach of fitting a circle to the noisy samples is prone to significant noise interference, particularly for small vibrations. For instance, a 50 μm vibration corresponds to a mere 0.33 rad phase change, representing approximately 5% of the complete circle and making it highly sensitive to signal noise.

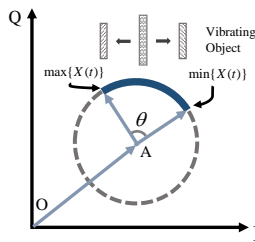


Figure 2: Complex plot representation of $S(t)$, $S(t)$ due to the displacement $X(t)$ is located in the thick arc. The vector \vec{OA} represents the static background reflections.

3.3 Estimation Challenges

To investigate the impact of noise and short arc length (corresponding to small amplitude) on geometric fitting, we conduct two simulation experiments using Python. The simulation parameters for the radar device are set as follows: carrier frequency $f_c = 77.64$ GHz, sampling rate $f_{adc} = 12$ MHz, modulation bandwidth $B = 4$ GHz, and slope $K = 20$ MHz/ μs .

The first experiment involved comparing an ideal noise-free reference signal with a comparison signal with a SNR of 0 dB. In the simulation, the target amplitude is set to 150 μm , and the corresponding IQ sample points form a circular arc, occupying 15% of the entire circle. By adding AWGN to the ideal signal, resulting in an SNR of 0 dB, we obtain the results shown in Figure 3. It is evident that there is a significant discrepancy in the circle center between the two signals, indicating that noise in the reference signal leads to inaccurate estimation of the circle center. When the circle's center is translated to the origin, this deviation is propagated to the arc. From Figure 3, we observe that the arc with noise corresponds to a considerably larger circle center angle, introducing errors in the extraction of the phase angle.

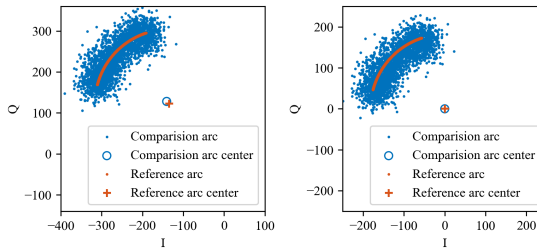


Figure 3: Center estimation results comparison between reference signal (orange) and comparison signal (blue).

The second experiment involved comparing different amplitudes at the same SNR level. Keeping other parameters constant, the SNR is set to 0 dB, and the amplitude is set to 400 μm , corresponding to a theoretical signal arc occupying 41% of the entire circle. The experimental results are presented in Figure 4. It can be observed that the estimated positions of both circle centers are nearly identical. When the center of the circle is translated to the origin, the relative positions of the circular arcs remain unchanged. Comparing with the results of the first experiment, it is evident that the estimation accuracy of the longer circular arc is closer to the ground truth, exhibiting significant improvement over the shorter arc.

To comprehensively investigate the impact of noise level and amplitude on the accuracy of circular center estimation, we conduct a series of simulations using mmWave radar for vibration measurements. In these experiments, we assume a simulated target vibration frequency of 50 Hz, with the amplitude sweeping from 20 μm to 1000 μm and the SNR sweeping from 10 dB to 75 dB. Subsequently, the estimation error is calculated using the following formula:

$$e = \frac{|E_v - A|}{A} \times 100\%, \quad (7)$$

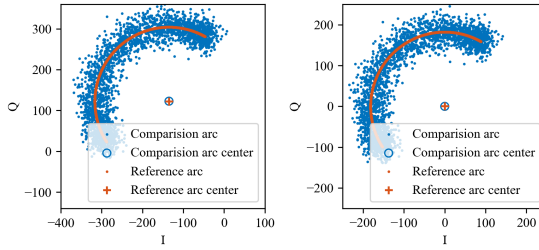


Figure 4: Center estimation results comparison between reference signal (orange) and comparison signal (blue).

where E_v represents the measured amplitude, and A represents the ground truth. Multiple experiments are performed for each set of parameters, and the average value is computed to determine the final estimation error.

The amplitude estimation error is studied by varying both the amplitude and SNR in multiple simulations. The results, shown in Figure 5, indicate that minimizing the error at low SNR requires a larger amplitude, corresponding to a larger arc length. For example, with an amplitude of 150 μm and an SNR of 16 dB, the error is 24.63%. However, increasing the amplitude to 500 μm under the same SNR reduces the error to 5.06%, an improvement of 79.4%. Conversely, at small amplitudes, higher SNR levels significantly reduce the error. Hence, improving the estimation performance involves enhancing the SNR and expanding the amplitude to increase the coverage of circular arcs.

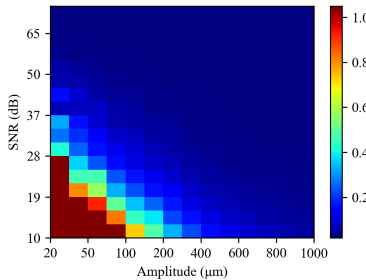


Figure 5: Estimation errors for different combinations of SNR and amplitude.

Based on the findings of the simulation experiments, in order to reduce the error, we adopt the idea of improving the SNR for system design. However, traditional methods face challenges in enhancing the SNR due to fixed radar parameters and target RCS. To overcome this, we propose a deep learning approach to enhance the SNR of vibration reflection signals. By leveraging

a large amount of simulation data, we train a deep learning model to learn the intrinsic features of the data and mitigate the impact of noise on the geometric fit, as illustrated in Figure 6. This approach eliminates the need for actual experimental data acquisition and the setup of a physical environment, resulting in significant reductions in labor costs.

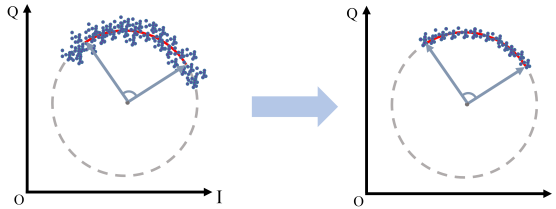


Figure 6: Improving SNR.

4 Proposed Method

This section presents the design of DeepVib. Figure 7 illustrates the workflow of DeepVib, which comprises three key modules.

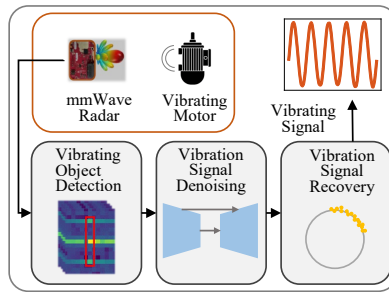


Figure 7: DeepVib Architecture.

- Vibrating Object Detection (VOD):** The VOD module leverages the inherent physical characteristics of the vibrating object and utilizes Range-Doppler FFT to extract the slow time sequence of the vibrating object from the chirp signals.
- Vibration Signal Denoising (VSD):** The VSD module consists of two steps. Firstly, a denoising neural network is trained using the generated data. Then the trained network model takes the noisy sample points as input and outputs the denoised data with higher SNR.

- **Vibration Signal Recovery (VSR):** The VSR module performs circular fitting to eliminate background reflections and subsequently unwraps the phase of the signal to recover the target vibration.

4.1 Vibrating Object Detection

In the VOD module, the radar data is initially processed using Range-FFT, which transforms signals at various distances into distinct range bins. To distinguish the vibrating target from static and other moving objects, we leverage the periodic motion of the vibrating target. When it passes through the center symmetry position of the vibration, the velocity exhibits the same magnitude but opposite direction. For the small amplitude of the vibration, the vibrating target consistently occupies the same range bin. Therefore, with the assistance of Doppler-FFT, we can identify the range bin where the vibration signal is present: the motion characteristics of the vibrating object result in symmetrical positive and negative velocities in the Doppler-FFT spectrum, as depicted in Figure 8. This characteristic enables the exclusion of other moving objects. And the detailed VOD algorithm is described in Algorithm 1, where X denotes a frame of data, ϵ represents the threshold for identifying vibration targets, and R indicates the range bin associated with the vibration signal.

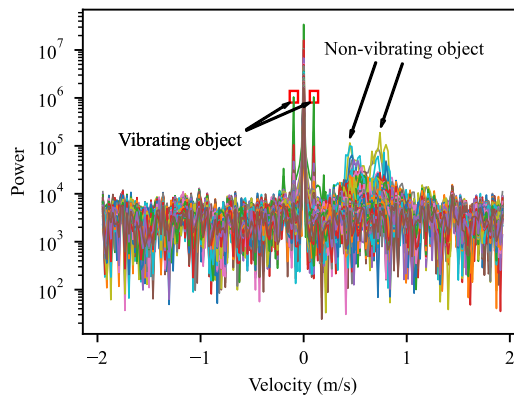


Figure 8: Range-Doppler.

4.2 Vibration Signal Denoising

Following the VOD module, we successfully extract the signals reflected from the vibrating object, which is however noisy due to the noise and imperfect spatial separation of the Range-FFT. Consequently, the presence of noise can obscure micro-meter level vibrations, thereby compromising the accuracy of

Algorithm 1: Detect vibrating target.

Input: one frame data X , ϵ , K

Output: range bin index R

```

1  $X^* \leftarrow$  Range-Doppler and FFTshift on  $X$ 
2  $X^p \leftarrow \mathbf{abs}(X^*)$ 
3  $X^p \leftarrow$  eliminate the component with velocity zero
4  $I \leftarrow$  find index of top  $K$  max values of  $X^p$ 
5 for  $(i, j)$  in  $I$  do
6    $(i', j) \leftarrow$  find symmetric index of  $(i, j)$ 
7    $v \leftarrow \frac{\mathbf{abs}(X^p[i'][j] - X^p[i][j])}{\max(X^p[i'][j], X^p[i][j])}$ 
8   if  $v < \epsilon$  then
9      $R \leftarrow j$ 
10    return

```

vibration estimation. To address this issue, we introduce the VSD module in this subsection. The primary objective of the VSD module is to denoise the extracted signals and effectively suppress the noise components.

Ideally, the IQ samples of the reflected signals exhibit a circular distribution in the complex plane. These samples demonstrate temporal correlation owing to the continuous nature of the vibrating motion, with stronger correlation observed between samples that are closer in time. This correlation pattern bears resemblance to the spatial correlation observed in visual images, where neighboring pixels exhibit high correlation, and proximity in space results in stronger correlation. Leveraging the remarkable capabilities of Convolutional Neural Networks (CNNs) in visual image processing, we adopt CNNs in this study to process the IQ samples of the reflected signals.

Since the underlying structure of the reflected signal is simple, we utilize a shallow U-Net [20] architecture for the network design, as shown in Figure 9. The network architecture consists of multiple 1D-CNN layers, each utilizing a 3x1 kernel. Rather than employing pooling operations for downsampling, we employ convolutional layers with a step size of 2 to achieve feature compression and dimensionality reduction. For upsampling, we employ bilinear interpolation. By employing convolutional layers with different step sizes for expanding or reducing the receptive field, we capture feature information at various scales, enhancing the network's expressiveness and generalization capability. To facilitate optimal network training, we normalize the samples before feeding them into the network as follows:

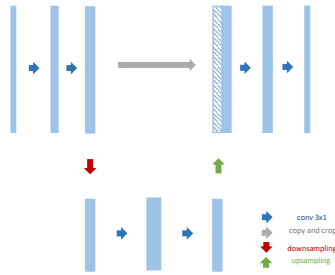


Figure 9: Noise reduction network structure.

$$Q' = \frac{Q}{\max(|\text{mean}(Q)|)}, \quad (8)$$

where Q is the samples extracted from the VOD module.

We use the l_1 -norm sum as the loss function to train the network

$$\mathcal{L} = \sum \|f(Q') - Y\|_1, \quad (9)$$

where Y is the ground truth and $f(Q')$ is the estimated value through the network.

4.3 Vibration Signal Recovery

After the VSD module, the reflected signals are effectively denoised, allowing us to employ a circle fitting method to eliminate background reflections. Let $V = v_{iN}$, $v_i \in \mathcal{R}^2$ denote the IQ samples extracted from each chirp and processed by the VSD module. The circle fitting process is formulated as an optimization problem to minimize the sum of geometric distances from each sample to the circle:

$$u^*, r^* = \arg \min_{u, r} \sum_{i=1}^N (\|v_i - u\|_2 - r)^2, \quad (10)$$

where v_i is the i -th denoised sample, N is the number of samples, u and r stand for the circle center and radius, respectively.

The above problem is a nonlinear least squares optimization problem that lacks an analytical solution. Therefore, it can only be solved using iterative or approximate methods. The Levenberg-Marquardt (LM) algorithm [4] demonstrates the lowest error and the fastest convergence. Hence, we employ the LM algorithm as the solver for circle fitting. Once the radius and center of the circle are estimated, we subtract the circle's center from the sample points to eliminate background reflections:

$$v'_i = v_i - u^*. \quad (11)$$

Then, according to Equation 5, the displacement of the vibration signal D_i can be obtained from the phase of sample v'_i , φ_i , as follows:

$$D_i = \frac{c}{4\pi f_c} \text{unwrap}(\varphi_i), \quad i \in [1, N]. \quad (12)$$

5 Experimental Results

5.1 Implementation Details

5.1.1 Generating simulation data

To train the denoising neural network, we first generate a large amount of simulation data based on the signal model in Equation 3. The output IF signal is influenced by various parameters, such as the object's amplitude, vibration frequency, initial phase, mmWave radar's carrier frequency, and noise intensity, so we can generate enough training data by setting different values for these parameters. The specific parameter settings are as follows:

- $A_v = \{10, 20, 30, 40, 70, 120, 200, 500\}$ μm
- $f_v = \{30, 50, 80, 120, 160, 250\}$ Hz
- f_c uniformly distributed from 77GHz to 78GHz
- φ_v uniformly distributed form 0 to 2π

5.1.2 mmWave radar configuration

In addition to the simulation experiments, we also evaluate the performance of DeepVib in real-world scenarios, where the configuration of the mmWave radar during data acquisition is shown in the Table 1.

5.1.3 Evaluation metrics

We use the Root Mean Square Error (RMSE) and the amplitude estimation error e as evaluation metrics.

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (\hat{y}_i - y_i)^2}, \quad (13)$$

where $\{y_i\}_N$ is the motion of the vibration target, and $\{\hat{y}_i\}_N$ is the result estimated by our model.

Table 1: mmWave radar parameters configuration.

Type	Value
f_c	77GHz
K	50MHz/us
T	75us
B	3.66GHz
N	256
Antenna mode	3Tx4Rx
ADC sample rate	3500ksps
Chirp idle time	20us

$$e = \frac{|\hat{y} - y|}{y} \times 100\%, \quad (14)$$

where y is the ground truth of vibration target amplitude, and \hat{y} is the estimated value of the DeepVib.

5.1.4 Comparison method

We conduct a comparative analysis of our proposed DeepVib with two existing mmWave-based vibration measurement approaches: the circle fitting-based method [16] (referred to as CircleFit) and the multi-signal consolidation model [8] (referred to as mmVib). The biggest difference between our proposed DeepVib and the CircleFit is the addition of the VSD module, which is pre-trained using simulation data, so that the comparison of experimental results with the CircleFit can fully illustrate the superiority of the VSD module. To ensure the fairness of the comparison, all methods use the same data and pre-processing methods, and the experimental results are averaged over multiple experiments. All simulations and experiments are run on a PC host with an Intel i7 10700K CPU @3.7GHz and 32G RAM.

5.2 Simulation Results

Denosing performance. In this subsection, we first evaluate the denoising performance of our VSD module. The IQ samples before and after VSD module are shown in Figure 10. We can see that with the proposed VSD module, the noise is well suppressed and thus the samples fit better with the ground-truth arc, and the sample points are more aggregated with each other. We also illustrate the displacement profile in Figure 11(a), where we can see that the vibration signals reconstructed by the proposed DeepVib matches well

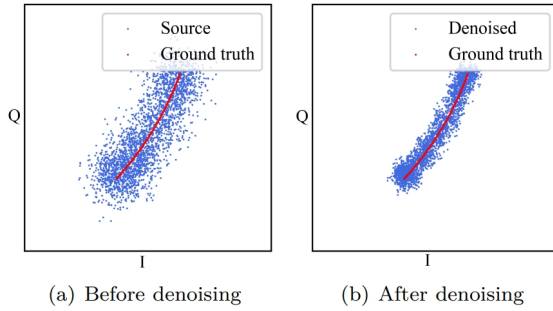


Figure 10: Comparison of samples before and after denoising.

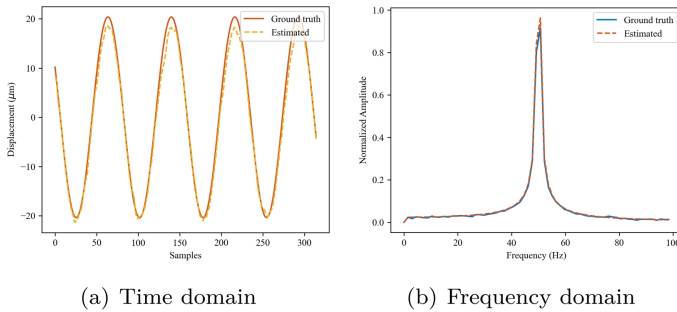


Figure 11: Comparison of extracted vibration signal and ground-truth.

with the ground truth, and the estimated amplitude RMSE is $1.56\mu\text{m}$, which validates the effectiveness of the proposed DeepVib. The frequency-domain signals of vibration target are shown in Figure 11(b). The location of the peak with maximum amplitude, which corresponds to vibration frequency, is approximately the same for the estimated and ground truth spectrum, which demonstrates that the DeepVib could achieve accurate frequency estimation.

Impact of vibration amplitude. We then evaluate the performance of the different methods based on the accuracy of the amplitude estimation. We set the vibration frequency to 50 Hz and reduced the amplitude from $200\mu\text{m}$ to $20\mu\text{m}$. The amplitude estimation errors at different distances are shown in Figure 12. We can see that for all three methods, the performance gradually improves as the amplitude increases and the error rate decreases. Notably, DeepVib consistently achieves the highest accuracy across all cases, while CircleFit exhibits the poorest performance, confirming the effectiveness of the VSD module. Moreover, DeepVib demonstrates an approximately 40% improvement compared to mmVib.

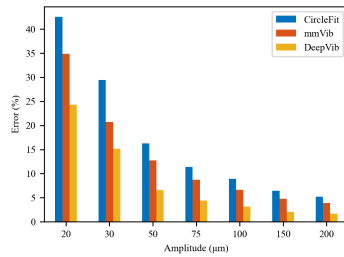


Figure 12: Impact of amplitude.

Impact of SNR. We also evaluate the performance of our system at different SNR, where we keep the frequency at 50 Hz and the amplitude at 100 μm , and then span the SNR from 15 dB to 30 dB. The experimental results are shown in Figure 13. Three methods could achieve high accuracy to estimate the amplitude at high SNR. However, the accuracy of CircleFit for 18dB drops dramatically while DeepVib is still accurate. The performance of all three methods improves as the SNR increases, and the error achieved by the DeepVib is much smaller than that of the comparison method at all SNR levels. When the SNR is 24dB or greater, our system achieves the amplitude estimation error below 3%; the experimental results at 15 dB show that the present system can significantly improve the estimation performance at low SNR.

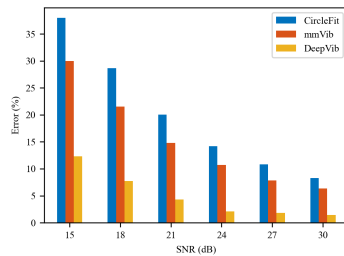


Figure 13: Impact of SNR.

5.3 Vibration Calibrator Experiment

In this subsection, to further illustrate the superiority of the proposed method, we conduct experiments on the vibration calibrator dataset. This dataset, as described in [8], utilizes a Texas Instruments AWR1642 77GHz millimeter wave radar as the acquisition hardware, along with a micron-scale

vibration calibration device serving as both the vibration generation source and the ground truth for vibration parameters. The dataset encompasses multiple sets of data, each characterized by different measurement distances and varying vibration amplitudes. It is important to note that these data are exclusively used for evaluating the model during the testing phase and are not included in the training phase.

Denoising performance. We first evaluate the denoising performance of our VSD module for the real-world data. Figure 14 illustrates the IQ samples of the calibrator with an amplitude of $100\ \mu\text{m}$ at a distance of 640 cm. It is evident that the VSD module effectively suppresses the noise, resulting in more consolidated and refined samples. Moreover, we present the recovered vibration signal with an amplitude of $30\ \mu\text{m}$ at 80 cm in Figure 15. The discrepancy between the peak value and $30\ \mu\text{m}$ is minimal. Additionally, the frequency estimation of this vibration yields a result of 50.05 Hz, with an error rate of less than 0.1% when compared to the actual frequency of 50 Hz. These findings further substantiate the efficacy of our proposed model in accurately capturing real-world vibrations.

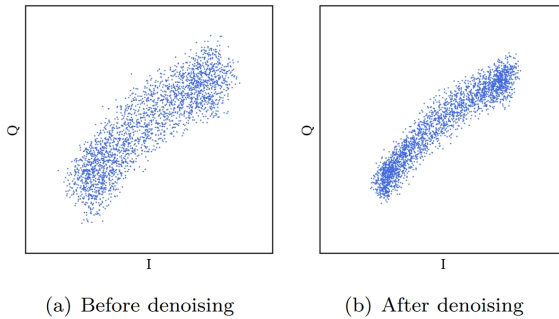


Figure 14: Comparison of samples before and after denoising.

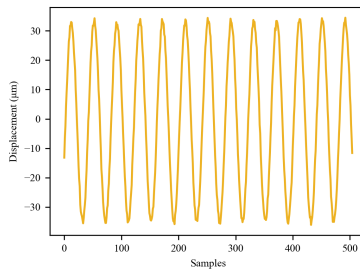


Figure 15: Reconstructed target vibration signal.

Impact of measurement distances. We then verify the capability of our system for vibration measurement, we set the vibration frequency to 50 Hz and perform multiple sets of experiments at amplitudes of $30\mu\text{m}$ and $100\mu\text{m}$ at distances ranging from 80 cm to 640 cm. The results are shown in Figure 16. We can see that the estimation errors of all three methods increase as the distance increases, with our proposed DeepVib having the best performance and CircleFit having the worst performance. At a measurement distance of 80cm, for the amplitude of $100\mu\text{m}$, DeepVib achieves an average amplitude error of $2.1\mu\text{m}$ with a relative error of 2.1%. When the distance is 640cm, DeepVib improves the error in amplitude estimation by about 33% compared to mmVib. Experiments on real data show that our proposed DeepVib can effectively and accurately recover the vibration of the target.

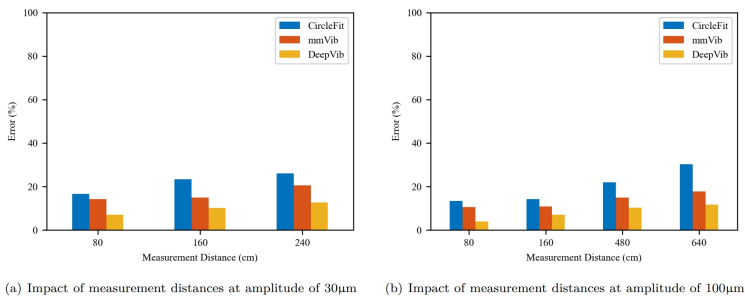


Figure 16: Impact of measurement distances.

Execution time. Finally, we evaluate the performance of the different methods in terms of execution time, and the results are shown in Table 2. In Table 2, we perform 90 experiments, each containing 10 frames of data, and take the average result of the 90 experiments as a display. It is worth noting that DeepVib and CircleFit have similar execution times, and both are about 5 times faster than mmVib. This experiment illustrates the vibration recovery efficiency of our proposed method, which is able to reconstruct the target motion process in real time.

Table 2: Comparison of the computation speeds of various methods.

Algorithm	Time/ms
CircleFit	85.8647
DeepVib	86.6164
mmVib	480.4097

Overall performance. To further illustrate the advantages of the pro-

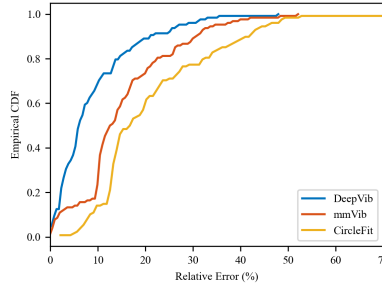


Figure 17: CDF of amplitude estimation.

posed DeepVib system, We adopt the cumulative distribution function(CDF) of amplitude estimation to evaluate performance of three approaches and the comparison is shown in Figure 17. From the figure, we can see that our system can achieve accurate measurement of micron-level vibration, and the relative amplitude error of DeepVib about 80% is less than 14.17%, which reduces the 80% quantile amplitude measurement error by 56.60% compared with the traditional method CircleFit. In addition, the relative error of DeepVib 50% is less than 7.23%. The comparison results show that the DeepVib outperforms the other two approaches.

5.4 Motor Vibration Experiment

In this subsection, we investigate the vibrations occurring on the surface of a motor, employing the experimental setup depicted in Figure 18. The experimental object is the motor on the rail that controls the movement of the platform, the distance between the mmWave radar and the motor is about 60 cm, while the mmWave radar configuration is consistent with the previously described setup.

To verify the effectiveness of the system in measuring motor vibration, we control the platform to perform several sets of experiments with different velocity reciprocal cycle motions in the horizontal direction. Figure 19 shows the time domain signal of motor vibration recovered by DeepVib, in which the platform moves at 5cm/s and 15cm/s, respectively. The amplitude estimation results of DeepVib system are 83.3 μ m and 85.1 μ m, and the frequency estimation results are 10.01Hz and 29.91Hz. The estimated frequencies have small errors compared with the ground truth frequencies of 10Hz and 30Hz, which indicates that DeepVib accurately measures the vibration of the motor at both speeds.

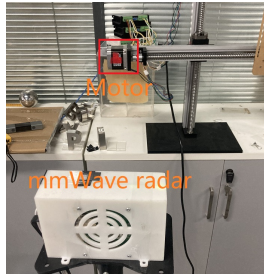


Figure 18: Motor vibration monitoring.

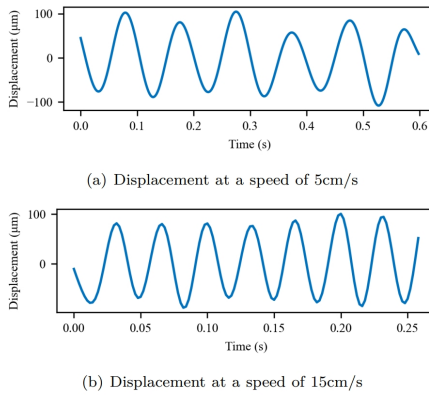


Figure 19: Vibration pattern of motor surface.

Estimation results of the motor at different speeds are shown in the Table 3. We can see that the seventh group of data estimated the frequency of 25.51Hz, which deviates significantly from the ground truth of 30 Hz. This discrepancy arises due to the limited length of the rail. When the platform reaches the ends, it needs to decelerate and reduce speed to zero before changing direction. Consequently, during this period, the motor frequency is lower than 30 Hz. However, for the other groups of data, the experimental results closely align with the ground truth values, demonstrating that DeepVib can accurately recover motor vibrations at the micron level.

5.5 Structure Vibration Experiment

In this subsection, we assess the performance of the DeepVib system in monitoring building structures, which is crucial for evaluating its efficacy in practical applications. The monitoring and analysis of structural responses under dynamic loads, such as earthquakes and wind, play a vital role in de-

Table 3: Motor vibration measurement results.

	Velocity(cm/s)	Estimated Amplitude(μm)	Estimated Frequency(Hz)	Frequency ground truth(Hz)
#1	0.5	81.0	1.00	1.00
#2	1	82.4	2.00	2.00
#3	2	83.0	4.00	4.00
#4	5	83.3	10.01	10.00
#5	10	82.2	20.02	20.00
#6	15	85.1	29.91	30.00
#7	15	75.6	25.51	30.00
#8	20	79.5	40.28	40.00

signing structures capable of withstanding severe catastrophic and weather conditions. In civil engineering, researchers simulate dynamic loads on structures to measure their dynamic responses, including displacement, velocity, and acceleration.

In this experiment, we employ mmWave radar to investigate the effectiveness of DeepVib in measuring the dynamic response of structures. The mmWave radar is securely mounted on a steel beam structure near the ceiling, approximately 2.8 meters above the floor. To excite the floor structure, the experimenters strike it with tools. The estimated displacements of the structure, obtained from the radar signals, are illustrated in Figure 20. Under stationary conditions, the floor response exhibits noise, irregular signal patterns, and displacements of up to $60 \mu\text{m}$ caused by noise. However, when the experimenters strike the floor with different speeds, the floor response becomes regular. Notably, the displacements caused by both motions are approximately $60 \mu\text{m}$. These experimental results demonstrate that DeepVib is capable of accurately sensing the structural response of buildings at the micron level.

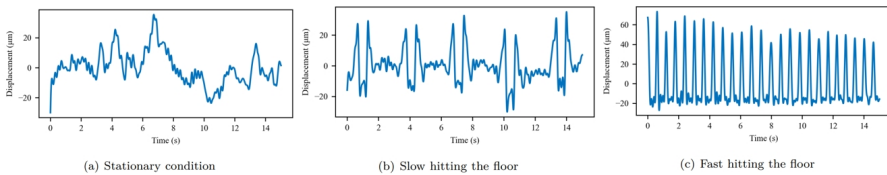


Figure 20: Floor vibration under different excitation.

6 Conclusion

In this paper, we proposed a deep learning based framework, DeepVib, to accurately sense the μm -level vibrating object. The key idea was to exploit the temporal correlations among different IQ samples through a simple yet effective neural network. Both simulation and experimental results demonstrated that DeepVib can extract tiny vibrations robustly, accurately, and efficiently even under low SNR conditions.

Acknowledgments

We would like to thank Prof. Yuan He, Chengkun Jiang, and Junchen Guo for sharing the dataset and the discussions regarding the experimental settings as well as the implementation of mmVib. This work was supported by National Key R&D Program under Grant 2022YFC0869800 and 2022YFC2503405, National Natural Science Foundation of China under Grant 62201542 and 62172381, fellowship of China Postdoctoral Science Foundation under grant 2022M723069, and the Fundamental Research Funds for the Central Universities.

References

- [1] E. P. Carden and P. Fanning, "Vibration based condition monitoring: a review", *Structural health monitoring*, 3(4), 2004, 355–77.
- [2] L. Ding, M. Ali, S. Patole, and A. Dabak, "Vibration parameter estimation using FMCW radar", in *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, 2016, 2224–8.
- [3] C. Duan, L. Yang, Q. Lin, Y. Liu, and L. Xie, "Robust spinning sensing with dual-rfid-tags in noisy settings", *IEEE Transactions on Mobile Computing*, 18(11), 2018, 2647–59.
- [4] W. Gander, G. H. Golub, and R. Strebler, "Least-squares fitting of circles and ellipses", *BIT Numerical Mathematics*, 34, 1994, 558–78.
- [5] X. Gao and O. Boric-Lubecke, "Radius correction technique for Doppler radar noncontact periodic displacement measurement", *IEEE Transactions on Microwave Theory and Techniques*, 65(2), 2016, 621–31.
- [6] F. Guidi, A. Guerra, and D. Dardari, "Personal mobile radars with millimeter-wave massive arrays for indoor mapping", *IEEE Transactions on Mobile Computing*, 15(6), 2015, 1471–84.
- [7] Y. He, Y. Chen, Y. Hu, and B. Zeng, "WiFi vision: Sensing, recognition, and detection with commodity MIMO-OFDM WiFi", *IEEE Internet of Things Journal*, 7(9), 2020, 8296–317.

- [8] C. Jiang, J. Guo, Y. He, M. Jin, S. Li, and Y. Liu, “mmVib: micrometer-level vibration measurement with mmwave radar”, in *Proceedings of the 26th Annual International Conference on Mobile Computing and Networking*, 2020, 1–13.
- [9] W. Jiang, C. Miao, F. Ma, S. Yao, Y. Wang, Y. Yuan, H. Xue, C. Song, X. Ma, D. Koutsonikolas, *et al.*, “Towards environment independent device free human activity recognition”, in *Proceedings of the 24th annual international conference on mobile computing and networking*, 2018, 289–304.
- [10] D. K. Kim and Y. Kim, “Quadrature frequency-group radar and its center estimation algorithms for small vibrational displacement”, *Scientific reports*, 9(1), 2019, 6763.
- [11] P. Li, Z. An, L. Yang, and P. Yang, “Towards physical-layer vibration sensing with rfid”, in *IEEE INFOCOM 2019-IEEE Conference on Computer Communications*, IEEE, 2019, 892–900.
- [12] J. Liu, Y. Chen, M. Gruteser, and Y. Wang, “Vibsense: Sensing touches on ubiquitous surfaces through vibration”, in *2017 14th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON)*, IEEE, 2017, 1–9.
- [13] J. Liu, Y. Wang, Y. Chen, J. Yang, X. Chen, and J. Cheng, “Tracking vital signs during sleep leveraging off-the-shelf wifi”, in *Proceedings of the 16th ACM international symposium on mobile ad hoc networking and computing*, 2015, 267–76.
- [14] C. X. Lu, S. Rosa, P. Zhao, B. Wang, C. Chen, J. A. Stankovic, N. Trigoni, and A. Markham, “See through smoke: robust indoor mapping with low-cost mmwave radar”, in *Proceedings of the 18th International Conference on Mobile Systems, Applications, and Services*, 2020, 14–27.
- [15] K. G. McConnell, *Vibration testing: theory and practice*, John Wiley & Sons, 1995.
- [16] I. V. Mikhelson, S. Bakhtiari, T. W. Elmer, A. V. Sahakian, *et al.*, “Remote sensing of heart rate and patterns of respiration on a stationary subject using 94-GHz millimeter-wave interferometry”, *IEEE Transactions on Biomedical Engineering*, 58(6), 2011, 1671–7.
- [17] K. Otsuka, K. Abe, J.-Y. Ko, and T.-S. Lim, “Real-time nanometer-vibration measurement with a self-mixing microchip solid-state laser”, *Optics letters*, 27(15), 2002, 1339–41.
- [18] S. Pan, T. Yu, M. Mirshekari, J. Fagert, A. Bonde, O. J. Mengshoel, H. Y. Noh, and P. Zhang, “Footprintid: Indoor pedestrian identification through ambient structural vibration sensing”, *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 1(3), 2017, 1–31.

- [19] J. A. Rice, C. Li, C. Gu, and J. C. Hernandez, “A wireless multifunctional radar-based displacement sensor for structural health monitoring”, in *Sensors and Smart Structures Technologies for Civil, Mechanical, and Aerospace Systems 2011*, Vol. 7981, SPIE, 2011, 185–95.
- [20] O. Ronneberger, P. Fischer, and T. Brox, “U-net: Convolutional networks for biomedical image segmentation”, in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5–9, 2015, Proceedings, Part III 18*, Springer, 2015, 234–41.
- [21] N. Roy and R. R. Choudhury, “Ripple {II}: Faster communication through physical vibration”, in *13th USENIX Symposium on Networked Systems Design and Implementation (NSDI 16)*, 2016, 671–84.
- [22] M. Safizadeh and S. Latifi, “Using multi-sensor data fusion for vibration fault diagnosis of rolling element bearings by accelerometer and load cell”, *Information fusion*, 18, 2014, 1–8.
- [23] S. Sami, Y. Dai, S. R. X. Tan, N. Roy, and J. Han, “Spying with your robot vacuum cleaner: eavesdropping via lidar sensors”, in *Proceedings of the 18th Conference on Embedded Networked Sensor Systems*, 2020, 354–67.
- [24] A. Veeraraghavan, D. Reddy, and R. Raskar, “Coded strobing photography: Compressive sensing of high speed periodic videos”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(4), 2010, 671–86.
- [25] C. Wang, L. Xie, W. Wang, Y. Chen, Y. Bu, and S. Lu, “Rf-ecg: Heart rate variability assessment based on cots rfid tag array”, *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 2(2), 2018, 1–26.
- [26] T. Wei and X. Zhang, “mtrack: High-precision passive tracking using millimeter wave radios”, in *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking*, 2015, 117–29.
- [27] B. Xie, J. Xiong, X. Chen, and D. Fang, “Exploring commodity rfid for contactless sub-millimeter vibration sensing”, in *Proceedings of the 18th Conference on Embedded Networked Sensor Systems*, 2020, 15–27.
- [28] L. Yang, Y. Li, Q. Lin, H. Jia, X.-Y. Li, and Y. Liu, “Tagbeat: Sensing mechanical vibration period with cots rfid systems”, *IEEE/ACM transactions on networking*, 25(6), 2017, 3823–35.
- [29] Z. Yang, P. H. Pathak, Y. Zeng, X. Liran, and P. Mohapatra, “Vital sign and sleep monitoring using millimeter wave”, *ACM Transactions on Sensor Networks (TOSN)*, 13(2), 2017, 1–32.
- [30] D. Zhang, Y. Hu, Y. Chen, and B. Zeng, “BreathTrack: Tracking indoor human breath status via commodity WiFi”, *IEEE Internet of Things Journal*, 6(2), 2019, 3899–911.

- [31] Y. Zhang, G. Laput, and C. Harrison, “Vibrosight: Long-range vibrometry for smart environment sensing”, in *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology*, 2018, 225–36.