

Pushing the Limit of Radar-Based Vibration Measurement With Deep Learning

Renjie Wen, Dongheng Zhang, Jinbo Chen, Qibin Sun, and Yan Chen
University of Science and Technology of China, Hefei, China

Email: {rjwen, jinbochen}@mail.ustc.edu.cn, {dongheng, qibinsun, eecyan}@ustc.edu.cn

Abstract—Vibration is a widespread physical phenomenon that often carries important information such as the internal state of the devices. Thus, vibration measurement is of great importance in the field of modern engineering and has drawn much attention. While achieving promising performance, existing methods fail when the vibration amplitude is tiny, e.g, smaller than 50 μ m. To address such a challenge, in this paper, we propose a contactless method with deep learning, denoted as DeepVib, to sense the tiny vibration using millimeter wave radar. Specifically, DeepVib first makes full advantage of the physical characteristics of the vibrating object and combines Range-Doppler FFT to find the range bin of vibrating objects. Then, DeepVib trains a denoising neural network using a large amount of simulated data, which takes the noisy sample points as input and outputs the denoised data with better SNR. Finally, the vibration status is recovered through the phase variation of the extracted signal. Simulation results show that DeepVib achieves over 40% improvement in measuring μ m-level amplitudes with over 5x faster processing time, while real experimental results show that DeepVib achieves a mean amplitude error of 2.1 μ m for the 100 μ m-amplitude vibration.

Index Terms—Wireless Sensing, Vibration Measurement, Deep Neural Network

I. INTRODUCTION

Vibration is a widespread physical phenomenon that often carries important information. In industry, the vibration of objects can reflect their internal state. By monitoring their vibration characteristics, amplitude and/or frequency, it is possible to detect damage or failure of device at an early stage [1]. In healthcare, the periodic motion of the human chest and heart is also a vibration that can reflect the health status of the human body [2]. Therefore, it is of great importance to accurately measure the vibration properties of an object.

Existing systems employ dedicated sensors such as accelerometers [3] and gyroscopes [4] to measure vibration. These approaches require the sensors to be attached to the surface of the vibrating object, which may not be possible in practice. High-speed cameras [5] and lasers [6] have also been explored for non-contact vibration sensing. While achieving accurate measurements of tiny vibration, both approaches are too expensive and limited to certain conditions such as good lighting and/or Line-of-Sight (LoS) conditions.

Recently, the researches of using radio signals to measure vibrations have been emerged mainly due to the unique characteristics of radio signals that is contactless, privacy friendly, and can penetrate obstacles. For example, WiFi signals have been utilized to track breathing and heartbeat [7]–[9], which

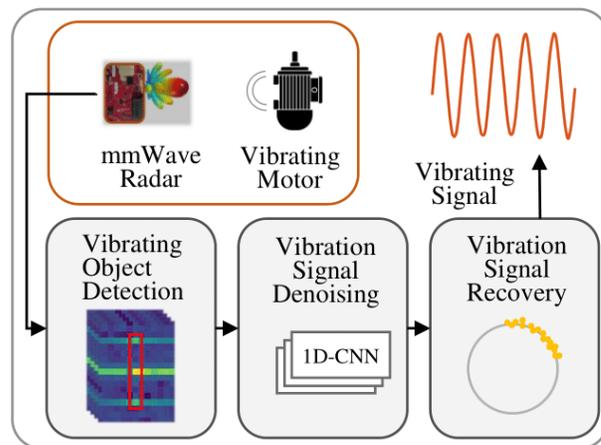


Fig. 1. DeepVib Architecture

can be considered as small vibrational displacement. [10] proposes to utilize RFID technology to monitor the rotation frequency of motors in noisy environment.

Because of the relatively short wavelength, e.g., 3.896 mm at 77 GHz, millimeter wave (mmWave) signals have been utilized for various sensing tasks [11]–[13]. In [14], the Cramer-Rao Lower Bound (CRLB) of the vibration parameter estimation with FMCW radar is derived, which however does not consider the multipath effect. To handle the multipath problem, [2] presents the geometric form of the signal model and proposes to remove the static reflections by geometric fitting. Nevertheless, such a method is very susceptible to noise when the vibration displacement is small. In order to address weak vibrations under low signal-to-noise ratio (SNR) scenarios, several approaches have been proposed to improve the fitting accuracy by combining multiple data with different frequencies [15]–[17]. However, these methods require extra hardware such as digital phase shifter with extra data. The increase of amount of data leads to a significant increase in computational complexity for vibration signal recovery, which limits the real-time applications.

Through the aforementioned discussions, there are two major challenges to be resolved for tiny vibration measurements as follows. 1) It is difficult to extract the vibration signal in a rich multipath environment. In a typical industrial environment, there are multiple objects besides the vibrating target,

due to which the received signal is a mixture of reflections from all objects. In such a case, the vibration-induced phase changes are often distorted, making the extraction of vibration signals extremely difficult. 2) It is difficult to recover the micrometer-level vibration. The vibration amplitude of most industrial devices is tiny, e.g., 20-100 μm . Such micrometer-level vibrations only lead to subtle phase change at the Intermediate Frequency (IF) signal. For example, a vibration of 50 μm may only lead to about 0.06 rad phase change for the mmWave with 3.896 mm wavelength, which can be easily submerged by the noise.

To resolve the aforementioned challenges, in this paper, we propose a contactless vibration sensing system with deep learning, denoted as DeepVib, to push the limit of sensing mechanical vibrations. Unlike existing methods that acquire extra data to suppress the noise, DeepVib utilizes a denoising neural network to reduce the noise level. As shown in Fig. 1, DeepVib consists of three main modules: vibrating object detection (VOD) module to extract the signals reflected from the vibrating object, vibration signal denoising (VSD) module to denoise the vibration signals, and vibration signal recovery (VSR) module to recover the vibration signals. The main contributions of this paper can be summarized as follows:

- To the best of our knowledge, DeepVib is the first work to utilize deep learning techniques to improve vibration sensing performance. We would like to emphasize that DeepVib is trained completely based on the simulated data and thus does not require great efforts for dataset construction.
- DeepVib is computationally efficient, and the processing time is extremely short during the testing phase. In particular, DeepVib reduces the processing time by about 80% over the frequency group algorithm.
- We validate DeepVib on both simulated and real-world data. The results show that DeepVib achieved over 40% improvement in measuring μm -level amplitudes on the simulated data. Although only trained by simulated data, DeepVib achieves a mean amplitude error of 2.1 μm for the 100 μm -amplitude vibration when testing on the real-world data.

The remainder of this paper is structured as follows. In Section II, we introduce the system model. Then, Section III presents the design of DeepVib in detail. Section IV shows the simulation and experimental results. Finally, we draw the conclusion in Section V.

II. SYSTEM MODEL

As shown in the Fig. 2, in this paper, we consider Frequency Modulated Continuous Wave (FMCW) mmWave radar, where the signal spectrum varies linearly over the bandwidth range. Specifically, a single transmitted chirp signal can be written as

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

where A_t is the magnitude related to the transmit power, f_c is the carrier frequency, and $K = B/T_r$ with B being the bandwidth of the signal.

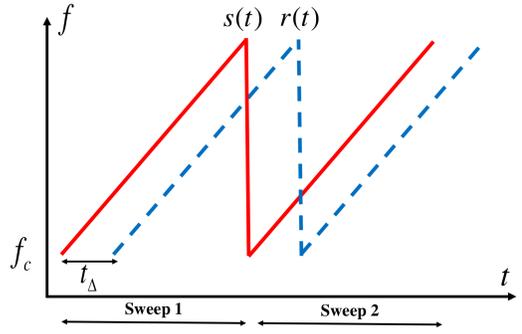


Fig. 2. An illustration of FMCW waveform, where the solid line stands for TX signal and dashed line stands for RX signal.

Suppose there is only one small object vibrating at the position x_0 from the radar. The time-varying distance $X(t)$, due to the vibration, from the object to the radar can be written 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 are the amplitude, frequency, and initial phase of the vibration target, respectively.

The reflected wave off the object at the receiver is the delayed version of $s(t)$ with a delay $t_\Delta = 2X(t)/c$, which is the round-trip propagation time of the chirp. The c is the light speed throughout the whole paper. Therefore, we can obtain the IF signal as

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

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. Note that the time of a chirp is generally short, and the vibrating object can be considered stationary during this time slot. To get the information of the target, we perform Range-FFT on each chirp, and combine the range bin of different chirps into a slow time sequence. The signal extracted from the range bin of the target distance can be expressed as

$$\tilde{y}(t) = a \exp[j4\pi f_c \frac{X(t)}{c}]. \quad (4)$$

From (4), we can see that the phase of $\tilde{y}(t)$, $\theta_{\tilde{y}}(t)$, contains all the vibration information of the target, from which we can obtain

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

where we can see that the phase change is proportional to the change of the distance to the vibrating object, which could be utilized to measure the vibration.

In practice, there are multiple reflected objects in the environment. Thus, the signal extracted from the range bin

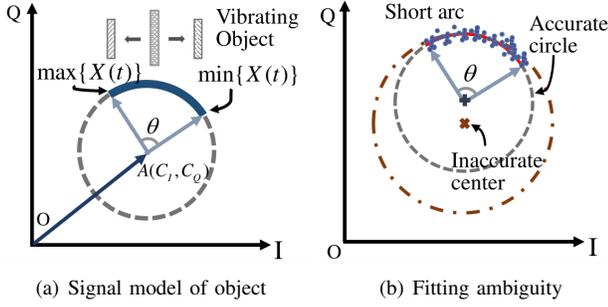


Fig. 3. (a) Complex plot representation of $S(t)$, $S(t)$ due to the displacement $X(t)$ is located in the thick arc. The vector \vec{OA} represents the static background reflections. (b) A small signal arc causes fitting ambiguity and induces inaccuracy. The small (gray) circle is the ground truth, and the large (brown) circle is the estimated incorrect circle.

where the vibrating target is located includes the information about the target's motion, the reflections from static objects in the environment, as well as the noise. In such a case, the received signal can be written as

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

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 S_s 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 Fig. 3(a), 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. In practice, due to the noise, the received signal $S(t)$ would be distributed around the arc. In such a case, a straightforward method would be fitting a circle with the noisy samples, which is however severely affected by the noise, especially for small vibrations. For example, a 50 μm vibration only results in a phase change of 0.06 rad, less than 1% of the entire circle, which is very sensitive to signal noise, as shown in Fig. 3(b).

III. PROPOSED METHOD

In this section, we will introduce the proposed DeepVib framework in detail. As shown in Fig. 1, DeepVib consists of three main modules as follows:

- **Vibrating Object Detection (VOD):** VOD takes full advantage of the physical characteristics of the vibrating object and combines Range-Doppler FFT to extract the slow time sequence of the vibrating object from the chirps.
- **Vibration Signal Denoising (VSD):** VSD first trains a denoising neural network 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):** VSR utilizes the circular fitting method to remove background reflections, and then unwraps the phase of the signal to recover the target vibration.

A. Vibrating Object Detection

In the VOD module, we first process the radar data through Range-FFT to transform signals at different distances into different range bins. To differentiate the vibrating object from the static objects, we observe that the vibrating target generally makes a periodic motion, due to which the velocity has the same magnitude but opposite direction when it passes through the center symmetry position. Thus, we can determine the range bin where the vibration signal is located with Doppler-FFT. As shown in Fig. 4, symmetrical velocities appear in the spectrum of the Doppler-FFT, which corresponds to the vibrating object. Here, we utilize such a feature to exclude other moving objects, and the detailed VOD algorithm is shown in Alg. 1, where X denotes a frame of data, ϵ is the threshold value for screening vibration targets, and R denotes the output range bin where the vibration signal is located.

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 $X^p \leftarrow \text{abs}(X^*)$
 - 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 $v \leftarrow \frac{\text{abs}(X^p[i'][j] - X^p[i][j])}{\max(X^p[i'][j], X^p[i][j])}$
 - 8 **if** $v < \epsilon$ **then**
 - 9 $R \leftarrow j$
 - 10 **return**
-

B. Vibration Signal Denoising

After the VOD module, we can extract the signals reflected from the vibrating object, which is however noisy due to the noise and imperfect spatial separation of the Range-FFT. In such a case, the micro-meter level vibration may be buried in the noise, leading to the inaccurate vibration estimation. Therefore, in this subsection, we introduce the VSD module to perform denoising on the extracted signals to suppress the noise.

Ideally, the IQ samples of the reflected signals are distributed on a circle in the complex plane. These samples are correlated in the temporal domain due to the continuity of the vibrating motion, and the closer in time, the stronger the correlation. Such a phenomenon is very similar to the visual

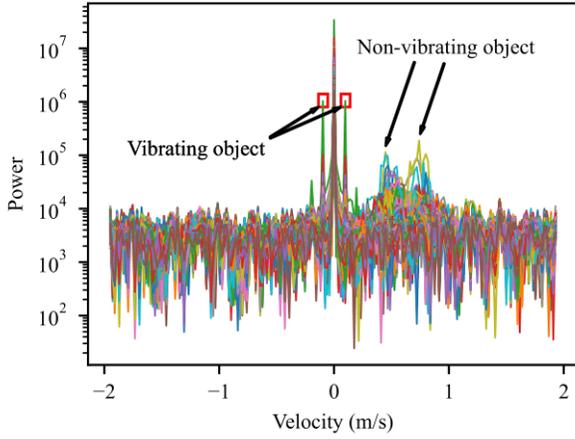


Fig. 4. Vibration object detection

image where the neighboring pixels are highly correlated and the closer in space, the stronger the correlation. Since convolutional neural network (CNN) has shown great advantages in processing the visual images, in this paper, we also utilize the CNN to process the IQ samples of the reflected signals.

Since the underlying structure of the reflected signals is simple, we utilize a network with three 1D-CNN layers, where the kernels are all 3x1 and repeated values are padded after each convolution to ensure the equal dimensions. Moreover, since the samples have real geometric significance, no activation function is used throughout the network. To better train the network, the samples are normalized before being fed into the network as follows:

$$H' = \frac{H}{\max(|\text{mean}(H)|)}, \quad (7)$$

where H is the samples extracted from the VOD module.

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

$$\mathcal{L} = \sum \|f(H') - Y\|_1, \quad (8)$$

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

C. Vibration Signal Recovery

After the VSD module, the noise of the reflected signals has been suppressed. Thus, we can utilize the circle fitting method to estimate the vibration by minimizing the geometric distances from the samples to the circle

$$u^*, r^* = \arg \min_{u, r} \sum_{i=1}^N (\|v_i - u\|_2 - r)^2, \quad (9)$$

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-square optimization problem, which has no analytical solution and can only be solved by iterative or approximated methods [16]. In this paper, we adopt the Levenberg-Marquardt algorithm [18],

which has been shown to have low error and fast running time [16]. After estimating the radius and the center of the circle, the sample points are subtracted from the center of the circle to eliminate static reflections using

$$v'_i = v_i - u^*. \quad (10)$$

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

$$D_i = \frac{c}{4\pi f_c} \text{unwrap}(\varphi_i), \quad i \in [1, N]. \quad (11)$$

IV. SIMULATION AND EXPERIMENTAL RESULTS

A. Implementation

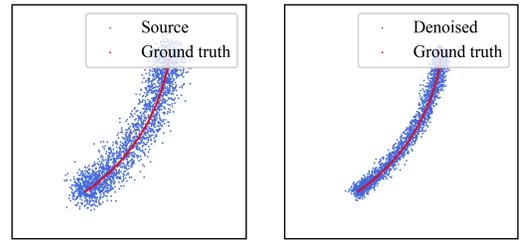
In order to train the denoising network for the VSD module, we generate the samples according to the signal model in (6) with the following settings

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

In this paper, the optimizer used for training the model is Adam, and the initial learning rate is set to 0.001, with an iteration period of 25. After every 5 iterations, the learning rate is reduced to one-tenth of the original rate. All the simulations and experiments are run on a computer with an Intel Core™ i7 10700k CPU @3.7GHz and Nvidia GeForce RTX 3090 graphics card.

We compare the proposed DeepVib with two baselines, which are the circle fitting-based method [2] (denoted by “CircleFit”) and the multi-signal consolidation model [15] (denoted by “mmVib”). For fair comparisons, all approaches use the same data and pre-processing methods.

B. Simulation Results



(a) Initial samples

(b) Denoised samples

Fig. 5. Comparison of samples before and after denoising.

Denoising Performance. We first evaluate the denoising performance of our VSD module. Fig. 5 show the initial IQ samples and those after the VSD module. 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. We also illustrate the displacement profile in Fig. 6, where we can see that the vibration signals reconstructed by the proposed

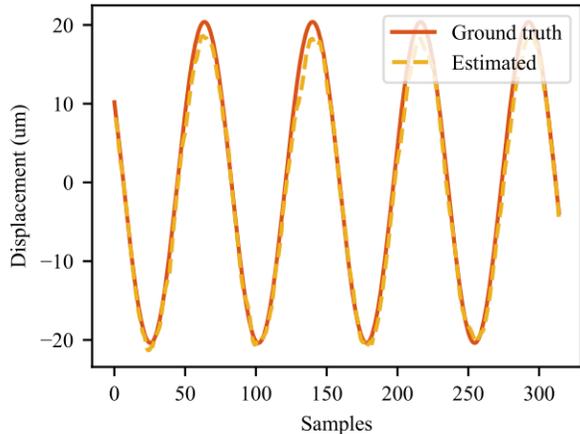


Fig. 6. Comparison between ground truth displacement and estimated

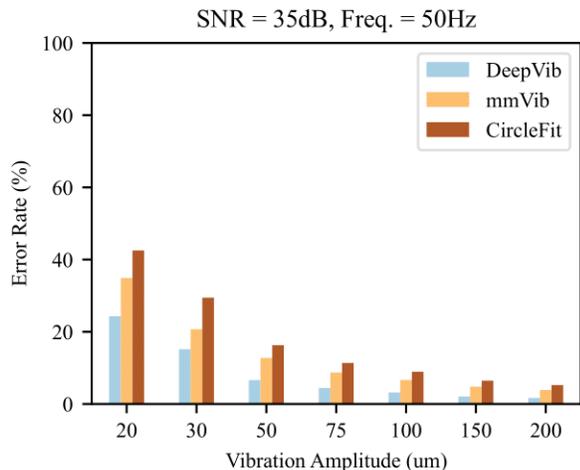


Fig. 7. Impact of vibration amplitude

DeepVib matches well with the ground truth, which validates the effectiveness of the proposed DeepVib.

Accuracy of Amplitude Estimation. We then evaluate the performance of different methods in terms of amplitude estimation accuracy. We set the vibration frequency at 50Hz and decrease the amplitude from 200um to 20um. The results are shown in Fig. 7. We can see that for all three methods, as the amplitude increases, the performance becomes better. DeepVib achieves the highest accuracy at all cases, while the CircleFit performs the worst. The improvement of DeepVib over mmVib is about 40%.

C. Experimental Results

In this subsection, we evaluate the effectiveness of the proposed DeepVib in real-world data. Specifically, we utilize the dataset provided in [15], which uses a Texas Instruments (TI) 77 GHz mmWave radar to acquire data and a vibration calibrator to generate vibration with ground truth in a $2.4\text{m} \times 10\text{m}$ hallway. Notice that these data are only used to evaluate the model in the testing phase, and not included in the training phase.

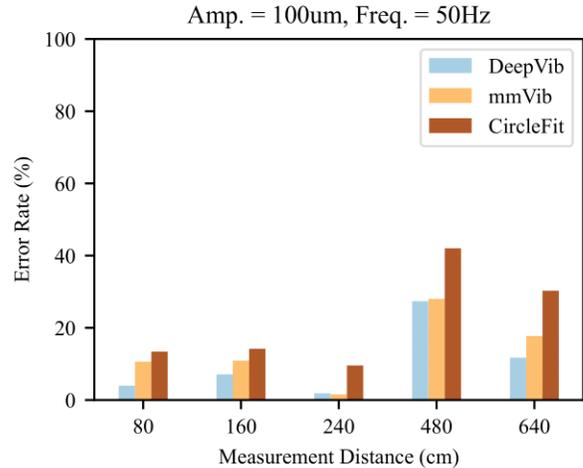


Fig. 8. Impact of measurement distance

Accuracy at Different Distances. In this experiment, we set the vibration frequency at 50 Hz and vary the distance from 80 cm to 640 cm at an amplitude of 100 um. The results are shown in Fig. 8. We can see that the experimental results are consistent with the simulation result in the previous subsection, where the proposed DeepVib achieves the best performance while the CircleFit performs the worst. For a typical case of 240 cm distance, DeepVib achieves an average amplitude error of 2.1 um with a relative error of 2.1%, which satisfies the measurement requirements of real industries. The improvement of DeepVib over mmVib can be up to 33% when the distance is 640 cm.

TABLE I
COMPARISON OF THE COMPUTATION SPEEDS OF VARIOUS METHODS.

Algorithm	Time/ms
CircleFit	85.8647
DeepVib	86.6164
mmVib	480.4097

Execution Time. Finally, we evaluate the performance of different methods in terms of execution time, and the results are shown in Table I, where we average over 90 experiments and each experiment contains 10 frames of data. We can see that the proposed DeepVib has similar execution time with the CircleFit, both are about 5x faster than mmVib.

V. CONCLUSION

In this paper, we proposed a deep learning based framework, DeepVib, to accurately sense the um-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.

VI. ACKNOWLEDGEMENT

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.

REFERENCES

- [1] K. G. McConnell and P. S. Varoto, *Vibration Testing: Theory and Practice*, John Wiley & Sons, Hoboken, N.J, 2nd edition, 2008.
- [2] I. V. Mikhelson, S. Bakhtiari, T. W. Elmer, II and A. V. Sahakian, "Remote Sensing of Heart Rate and Patterns of Respiration on a Stationary Subject Using 94-GHz Millimeter-Wave Interferometry", *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 6, pp. 1671-1677, 2011.
- [3] M.S. Safizadeh, S.K. Latifi, "Using multi-sensor data fusion for vibration fault diagnosis of rolling element bearings by accelerometer and load cell", *Information Fusion*, vol. 18, pp. 1-8, 2018.
- [4] J. Liu, Y. Chen, M. Gruteser and Y. Wang, "VibSense: Sensing Touches on Ubiquitous Surfaces through Vibration", *IEEE SECON*, 2017.
- [5] 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*, vol. 33, no. 4, pp. 671-686, 2011.
- [6] Kenju Otsuka, Kazutaka Abe, Jing-Yuan Ko, and Tsong-Shin Lim, "Real-Time Nanometer-Vibration Measurement with a Self-Mixing Microchip Solid-State Laser", *Opt. Lett.*, vol. 27, no. 15, pp. 1339-1341, 2002.
- [7] J. Liu, Y. Wang, Y. Chen, J. Yang, X. Chen, and J. Cheng, "Tracking vital signs during sleep leveraging off-the-shelf wifi", *ACM MobiHoc*, 2015.
- [8] D. Zhang, Y. Hu, Y. Chen and B. Zeng, "BreathTrack: Tracking Indoor Human Breath Status via Commodity WiFi", *IEEE Internet of Things Journal*, vol. 6, no. 2, pp. 3899-3911, 2019.
- [9] 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*, vol. 7, no. 9, pp. 8296-8317, 2020.
- [10] C. Duan, L. Yang, H. Jia, Q. Lin, Y. Liu and L. Xie, "Robust Spinning Sensing with Dual-RFID-Tags in Noisy Settings", *IEEE INFOCOM*, 2018.
- [11] T. Wei and X. Zhang, "mTrack: High-Precision Passive Tracking Using Millimeter Wave Radios", *ACM MobiCom*, 2015.
- [12] 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", *ACM MobiCom*, 2018.
- [13] F. Guidi, A. Guerra and D. Dardari, "Personal Mobile Radars with Millimeter-Wave Massive Arrays for Indoor Mapping", *IEEE Transactions on Mobile Computing*, vol. 15, no. 6, pp. 1471-1484, 2016.
- [14] L. Ding, M. Ali, S. Patole and A. Dabak, "Vibration Parameter Estimation using FMCW Radar", *IEEE ICASSP*, 2016.
- [15] C. Jiang, J. Guo, Y. He, M. Jin, S. Li, Y. Liu, "mmVib: Micrometer-Level Vibration Measurement with mmWave Radar", *ACM MobiCom*, 2020.
- [16] D. K. Kim and Y. Kim, "Quadrature Frequency-Group Radar and its Center Estimation Algorithms for Small Vibrational Displacement", *Scientific Reports*, vol. 9, no. 1, pp. 6763-6780, 2019.
- [17] J. -H. Park and J. -R. Yang, "Multiphase Continuous-Wave Doppler Radar With Multiarc Circle Fitting Algorithm for Small Periodic Displacement Measurement", *IEEE Transactions on Microwave Theory and Techniques*, vol. 69, no. 11, pp. 5135-5144, 2021.
- [18] W. Gander, G. H. Golub, and R. Strebler, "Least-Squares Fitting of Circles and Ellipses", *BIT Numerical Mathematics*, vol. 34, no. 4, pp. 558-578, 1994.