RF Vital Sign Sensing Under Free Body Movement

JIAN GONG [*], School of Computer Science and Engineering, Central South University, China
XINYU ZHANG, University of California San Diego, United States
KAIXIN LIN, University of California San Diego, United States
JU REN[†], Department of Computer Science and Technology, BNRist, Tsinghua University, China // School of Computer Science and Engineering, Central South University, ChangSha, China
YAOXUE ZHANG, Department of Computer Science and Technology, BNRist, Tsinghua University, China // School of Computer Science and Engineering, Central South University, ChangSha, China
WENXUN QIU, Samsung Research America, Plano, Texas, United States

Radio frequency (RF) sensors such as radar are instrumental for continuous, contactless sensing of vital signs, especially heart rate (HR) and respiration rate (RR). However, decades of related research mainly focused on static subjects, because the motion artifacts from other body parts may easily overwhelm the weak reflections from vital signs. This paper marks a first step in enabling RF vital sign sensing under ambulant daily living conditions. Our solution is inspired by existing physiological research that revealed the correlation between vital signs and body movement. Specifically, we propose to combine direct RF sensing for static instances and indirect vital sign prediction based on movement power estimation. We design customized machine learning models to capture the sophisticated correlation between RF signal pattern, movement power, and vital signs. We further design an instant calibration and adaptive training scheme to enable cross-subjects generalization, without any explicit data labeling from unknown subjects. We prototype and evaluate the framework using a commodity radar sensor. Under a variety of moving conditions, our solution demonstrates an average estimation error of 5.57 bpm for HR and 3.32 bpm for RR across multiple subjects, which largely outperforms state-of-the-art systems.

CCS Concepts: • **Human-centered computing → Ubiquitous and mobile computing systems and tools**.

Additional Key Words and Phrases: millimeter wave radar, deep learning, heart rate, respiration rate

ACM Reference Format:
Jian Gong, Xinyu Zhang, Kaixin Lin, Ju Ren, Yaoxue Zhang, and Wenxun Qiu. 2021. RF Vital Sign Sensing Under Free Body Movement. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 5, 2, Article 62 (June 2021), 21 pages. https://doi.org/10.1145/3463517

[*]This work was conducted when the author was a visiting scholar at the University of California San Diego.
[†]The corresponding author.

Authors' addresses: Jian Gong, School of Computer Science and Engineering, Central South University, China, gongjian@csu.edu.cn; Xinyu Zhang, University of California San Diego, United States, xyzhang@ucsd.edu; Kaixin Lin, University of California San Diego, United States, kal016@eng.ucsd.edu; Ju Ren, Department of Computer Science and Technology, BNRist, Tsinghua University, China, // School of Computer Science and Engineering, Central South University, ChangSha, China, renju@csu.edu.cn; Yaoxue Zhang, Department of Computer Science and Technology, BNRist, Tsinghua University, China, // School of Computer Science and Engineering, Central South University, ChangSha, China, zyx@csu.edu.cn; Wenxun Qiu, Samsung Research America, Plano, Texas, United States, wenxun.qiu@samsung.com.

2474-9567/2021/6-ART62 $15.00
https://doi.org/10.1145/3463517

1 INTRODUCTION

Continuous tracking of vital signs can be instrumental for eliciting healthy behaviors and assessing physical status during physical exercises. For example, resting HR in athletes was markedly lower than that in normal people [20, 34] and fitness crowds usually display longer exhale and inhale cycle time during long time exercise than normal people. Although current wearable technologies such as wristbands and chest vest can monitor vital signs, they are uncomfortable for continuous use, and unsuitable for people with skin allergy and dementia. Radio frequency (RF) signals, as an alternative medium, can overcome such limitations and enable contactless vital sign sensing.

In the past few decades, radar has been extensively explored as an RF sensor for vital signs [2, 13, 16, 45]. The basic principle is that the rhythmical heart beats and respiration-induced chest movement reflects and modulates the RF signals, which can be detected by radar using phase [12], frequency [18] or Doppler [16] feature extraction. Correspondingly, different hardware and signal processing schemes have been investigated to enhance the sensing range and accuracy. Examples include phased-array radar with high spatial resolution [23], IR-UWB radar with high time/range resolution [48], *etc.*

Most of the existing radar vital sign sensing systems work under controlled settings, where a human subject stays still in front of the radar. Several recent systems investigated restricted mobility scenarios, where the subjects may perform in-place typing/fidgeting activities or move the body regularly back and forth [2, 39]. However, in practical scenarios, the subject may engage in a much wider scale of body activities, *e.g.*, walking, household chores, fitness exercises, and random limb stretching. Such ambulant activities may intensively distort the radar signal pattern, immersing the vital sign induced features, thus inhibiting the parameter estimation. To our knowledge, *RF-based vital sign sensing under ambulant daily living conditions remains an open challenge today.*

In this paper, we propose a general model to remove the motion artifacts and enable RF vital sign sensing in ambulant conditions. Our model was inspired by prior physiological and medical research [25, 30, 31, 35], which indicated that human vital signs can be affected by body movement in a predictable way. We thus use a radar sensor to detect vital signs along with motion intensity. Specifically, we opportunistically call on the radar sensor, using it to obtain a baseline heart rate (HR) and respiration rate (RR) estimation *only* when the subject is static. Otherwise, whenever movement is detected, we predict the HR and RR on top of the baseline, rather than dealing with motion contacted RF signals directly. To incorporate these design principles into a unified framework, we design a deep learning framework which can extrapolate both HR and RR despite the motion artifacts.

Practical materialization of the above working principles entails several non-trivial design challenges. The first lies in model generalization across subjects. The mapping between motion intensity and vital sign intensity may vary among different people [8]. However, it is infeasible to request each user of our system to train a separate model by collecting ground-truth data and labels. To overcome this challenge, we propose a self-calibrated LSTM framework, comprised of two mechanisms: *(i)* an *adaptive training* scheme, aiming to calibrate the long-term error by fine-tuning the LSTM model parameters leveraging the accurate measurements at static instances; *(ii)* an *instant calibration* scheme, designed to opportunistically compensate for each single period of motion artifacts without altering the LSTM parameters, which can take effect even with a few calibration samples.

A second challenge lies in movement intensity detection for different body parts. Due to limited spatial resolution of commodity radar devices, we cannot easily discriminate different body parts as a Lidar/camera does. Instead, we design a movement power detection scheme, which estimates the moving intensity of individual body parts by combining information from reflection power, Doppler speed and distance. Our microbenchmark evaluation shows that the method can faithfully track the intensity and frequency of body movement from body parts, and outputs an estimate of the total movement power, which builds the basis for subsequent HR and RR estimation amid motion artifacts.

To verify the proposed system, we collect vital sign sensing data using a commodity 77GHz FMCW radar,

and further implement our HR and RR estimation models using the Pytorch [27] framework. We evaluate the proposed methods on multiple subjects under daily ambulant conditions. Two state-of-the-art RF based HR/RR sensing systems [2, 29] are also implemented and used as baseline. The results demonstrate an average HR estimation error of 5.57 bpm and average RR estimation error of 3.32 bpm across the subjects, which outperforms the state-of-the-art baseline (average HR error 55 bpm [2] and RR error 5.05 bpm [29]) by a large margin. In addition, our model shows prominent generalization capability across subjects, even with a small number of automatically generated calibration points (without any explicit labeling efforts from the users).

To our knowledge, this is the first RF vital sign sensing system which can work under unrestricted body movement. The main contributions of this paper are summarized as follows.

- We propose a novel paradigm to resolve the motion artifacts in RF vital sign sensing. Our system leverages the empirically verifiable correlation between HR/RR and motion intensity, instead of using time-frequency analysis which fundamentally suffers from motion-induced interference.
- We design a movement power detection module which performs multidimensional fusion of radar signals to estimate the motion intensity of individual body parts under ambulant conditions. We introduce analytical model and LSTM based model to realize explicit or implicit estimation.
- We propose a self-calibrated LSTM model to solve the problem of cross-subject generalization. This customized LSTM model combines instant calibration and adaptive training to resolve short-term and long-term generalization errors. Unlike classical transfer learning schemes, our model can adapt to the vital sign variation patterns of unseen subjects without explicit data labeling.

Note that our model builds on the correlation between vital sign variation and the motion intensity of body parts. So it is not applicable to the cases where this correlation is corrupted, *e.g.*, for people with certain cardiovascular diseases such as arrhythmia. However, it can still satisfy the needs of a wide range of ordinary usage scenarios, such as metabolism tracking at home and fitness tracking on a treadmill, *etc.* The source code and dataset generated are available online to promote further research[1].

2 CHALLENGES OF RF VITAL SIGN SENSING UNDER FREE BODY MOVEMENT

2.1 A primer on RF vital sign sensing

The frequency modulated continuous wave (FMCW) radar is commonly used as an RF sensor for vital sign detection. FMCW radar emits a signal called a *chirp* [37], *i.e.*, a sinusoid whose frequency increases linearly with time. As the distance varies between the target object and radar, the frequency difference between transmitted chirp and reflected chirp varies linearly. Therefore, the distance of object can be detected by calculating the differential signal between transmitted and reflected chirps. This differential signal is called an IF signal and is derived by an analog mixer module on the radar. Higher frequency of the IF signal corresponds to longer range between the radar and the object. After applying an FFT on the IF signal, we can extract a frequency spectrum with different range values, corresponding to different reflection points in the target scene. This is referred to as a *range FFT* representation. Each FFT bin is thus called a *range FFT bin*. The total number of range FFT bins is equal to the total number of samples within a chirp.

The general principle of RF vital sign sensing lies in the radar's ability to sense the minor distance variation caused by inhale and exhale on chest. This distance variation is reflected by the *phase variation* over time of the range FFT bin corresponding to the chest. After applying another FFT on phase variation, we can extract the frequency of respiration by selecting the peak on the frequency spectrum. This operation is referred to as *Doppler-FFT* [38]. Most existing RF vital sign sensing systems [2, 29] use such frequency domain analysis to extract the rhythmic features of respiration. The same principle can be applied to extract heart rate because heart

[1]Project: *https://github.com/jamesdeep/VitalSign*

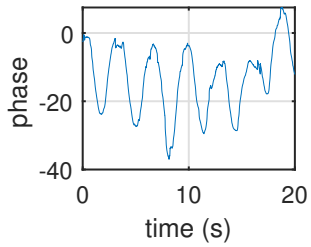

Fig. 1. Original phase variation under static situation.

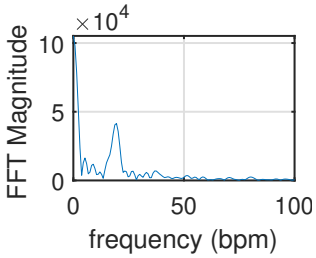

Fig. 2. Frequency spectrum under static situation.

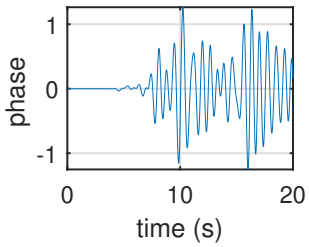

Fig. 3. Filtered phase variation under static situation.

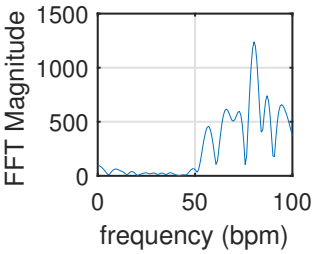

Fig. 4. Filtered frequency spectrum under static situation.

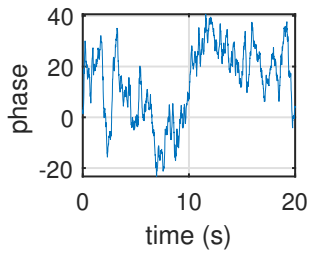

Fig. 5. Original phase variation under body movement (jogging).

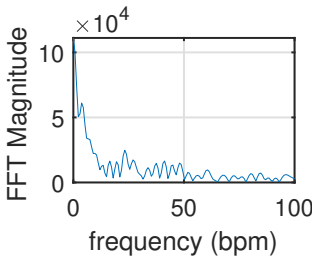

Fig. 6. Frequency spectrum under body movement (jogging).

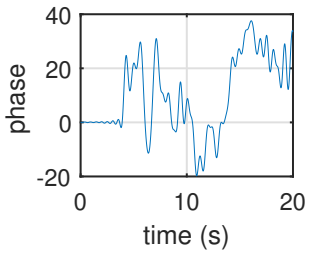

Fig. 7. Filtered phase variation under body movement (jogging).

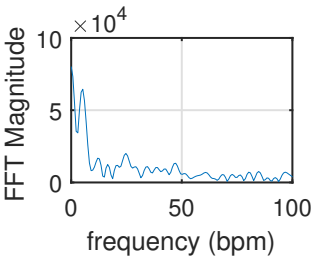

Fig. 8. Filtered frequency spectrum under body movement (jogging).

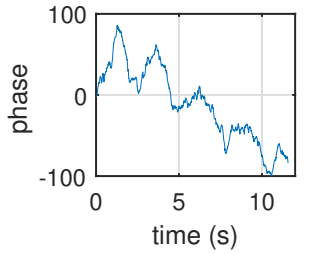

Fig. 9. Original phase variation under restricted body movement (waving arms while sitting).

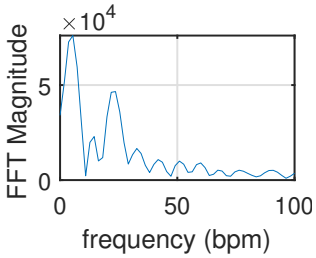

Fig. 10. Frequency spectrum under restricted body movement (waving arms while sitting).

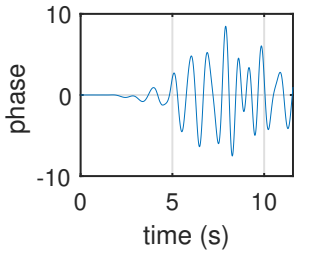

Fig. 11. Filtered phase variation under restricted body movement (waving arms while sitting).

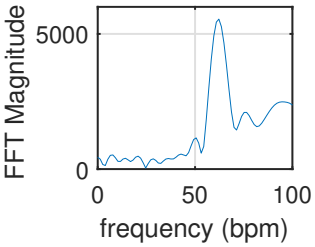

Fig. 12. Filtered frequency spectrum under restricted body movement (waving arms while sitting).

beats also cause chest fluctuation. Since the vibration of heart beats modulates atop the respiration, we can use band pass filters with different frequency range to separate them [2].

2.2 Body movement corrupts vital sign features

For static body, the magnitude of vital organ movement easily outweighs noise, and can be discriminated from high/low frequency noise through a bandpass filter [29]. On the other hand, in order to dodge the impacts of occasional aperiodic movement, such as raising limbs, an alternative strategy [2] is to discard the radar frames containing such movement. However, under ambulant conditions, different body parts' movements can create a wide range of frequency components spanning a relatively long period of time, which obfuscate the vital sign features.

In order to demonstrate the limitations of traditional frequency domain analysis, we collect radar sensing

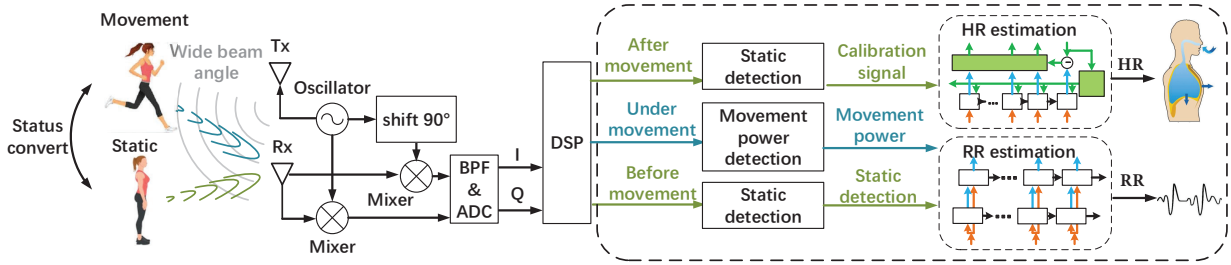

Fig. 13. Overall architecture of the robust RF vital sign sensing system.

data under different ambulant conditions, and attempt to use the frequency analysis methods to detect the vital signs (Details of our sensing platform and data collection process is described in Section 7). For each situation, we respectively demonstrate the phase variation over time, Doppler-FFT spectrum, filtered phase variation and filtered Doppler-FFT spectrum. The phase variation measurement is extracted from the range FFT bin corresponding specifically to the distance from the radar-to-chest distance. Therefore, it reflects the actual distance variation over time. We then apply Doppler-FFT on the time series of phase measurements to generate the frequency spectrum. In order to extract the weak heart beats from the noisy movement, we apply a band-pass filter on the phase variation curve with frequency range of 50-180 bpm. Finally, we apply a Doppler-FFT on the filtered phase variation to observe the frequency spectrum of heart beats. The results are shown in Fig. 1-12.

It is exactly our algorithm

When the body is static, we can clearly observe the inhale and exhale wave of respiration (Fig. 1). The frequency spectrum (Fig. 2) exhibits a sharp peak corresponding to the respiration rate. The heart beat frequency is also visible after applying a band-pass filter with frequency range of 60-90 bpm (Fig. 4).

As a verification of existing methods, we rerun the experiment under restricted movement (waving arms while sitting), which is consistent to the aperiodic assumption in previous work [2]. Although the phase variation has continuous drift and noise caused by movement, we can still observe the inhale and exhale respiration patterns (Fig.9). FFT analysis reveals a sharp peak corresponding to the respiration frequency around 20 bpm (Fig.10). After applying a band pass filter, we could even observe the wave of heart beats (Fig.11). Subsequently, by applying Doppler-FFT, the heart rate could be derived by observing the sharp peak of spectrum (Fig.12).

However, with less restricted movement such as jogging (Fig. 5), we see that the phase pattern is corrupted by noise. In the frequency domain, the heart beats can no longer be observed even after applying the band pass filter (Fig. 8). Additionally, the respiration peak is also overwhelmed by the body movement (Fig. 6). This benchmark experiment shows that the unrestricted body movements can affect the radar signal patterns in the same way as HR and RR. They may cover a similar frequency range with larger magnitude, making it infeasible to extrapolate the vital signs.

3 SYSTEM ARCHITECTURE

The high-level idea of our system is to sense the body movement and vital signs simultaneously, and leverage their correlation to reliably predict HR and RR. Fig. 13 illustrates the system work flow, which consists of 4 major steps.

I implemented only this part

Stage 1: Static detection. Whenever the subject becomes static, we adopt the traditional frequency analysis [2] to estimate the HR and RR.

Stage 2: Movement power detection. We use a movement power detection model to continuously derive the subject's movement intensity, frequency and patterns.

Stage 3: HR and RR estimation. Under ambulant conditions, we employ two customized LSTM models for estimating HR and RR from movement power, respectively. The HR model accepts a closed-form movement power estimation as input and outputs real-time HR. The RR LSTM consists of two stacked LSTMs, one for

movement pattern estimation and another for RR estimation.

Stage 4: Opportunistic self-calibration of the LSTMs. Whenever the subject transitions from mobility to static mode, we use the traditional frequency analysis again to detect the vital signs. The result is treated as ground-truth data and used for the adaptive training scheme to fine-tune the LSTM for model generalization. Additionally, another calibration scheme, instant calibration, also leverages this detection to calibrate the average error for each single period of movement.

4 MOVEMENT POWER DETECTION

The movement detection module, as illustrated in Fig. 14, aims to sense the subject's movement intensity and patterns (*e.g.*, periodical movement like jogging, versus random movement like stretching arms or turning around). More specifically, it needs to detect two movement properties. The first is the relative *mass*, *i.e.*, physical weights, of moving body parts. The second is how much force the subject uses to move different body parts. Unfortunately, the weights cannot be directly detected by radar. Therefore, we design a model that links body parts' mass with the radar samples. We can then estimate the movement power of different body parts which sum up as the total movement power.

4.1 Body mass estimation from range FFT bins

In classical static radar vital sensing scenarios [2, 45], the radar directly points to the subject. The radar's frequency mixer produces an IF signal x_{IF} from TX wave x_1 and reflected wave x_2:

$$x_1 = A_1 sin(\omega_1 t + \phi_1). \tag{1}$$

$$x_2 = A_2 sin(\omega_2 t + \phi_2). \tag{2}$$

$$x_{IF} = \frac{A_1 A_2}{2} sin((\omega_1 - \omega_2)t + (\phi_1 - \phi_2)). \tag{3}$$

where ω represents the frequency of the wave and ϕ represents the phase of the wave. Since A_1 is constant, the amplitude of the IF signal is proportional to the amplitude of received signal A_2.

Now consider a more general situation where the body comprises of multiple reflecting points and radar beam angle is not 0°, *i.e.*, it only covers part of the subject's body. Each body part only receives a small portion of the transmitted signals, so the amplitude of reflected wave of single body part is decreased to:

$$A_2^p = \frac{M_p}{M_{wd}} A_2, \tag{4}$$

where M_p represents the area of body part p, M_{wd} represents the radiation area at distance d. In our model, we assume the body can be approximately decomposed into a set of patch areas with similar sizes. A single patch area does not have to correspond to a specific body part. In contrast, a single body part, e.g., limb, trunk, head, can be divided into multiple patch areas. The assumption of the patch area aims to emphasize that the adjacent body areas can be regarded as a unified patch area where their reflection signals are superimposed. Therefore, all the body parts share the same M_{wd}.

After applying an FFT on the raw FMCW radar signal, we can extract a frequency spectrum with respect to different range values, *i.e.*, the range FFT, corresponding to different reflection points in the target scene. Since

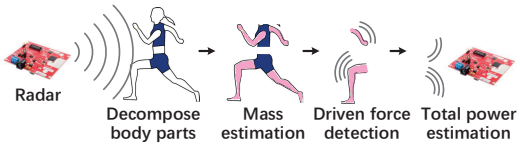

Fig. 14. Movement power detection.

the reflected waves from equidistant body parts bear the same phase, the received power of reflected wave are superimposed. Therefore, the received power of each range FFT bin is linearly proportional to the body area at a specific distance d:

$$S_k = D_k c_k, \tag{5}$$

where c_k is the received power of the k-th range bin, D_k is a converting proportional coefficient which converts received power to body area in the k-th range bin. The calculation of this coefficient will be explained at the end of this section.

Let the body part area resolved in the k-th range bin be S_k. We assume that the body part volume V_k is linear relative to S_k:

$$V_k = \mu S_k, \tag{6}$$

where μ represents an (unknown) proportional coefficient. This assumption means that the thickness of different body parts (chest, head, thigh...) are assumed to be close. Therefore, we can simplify the representation of V_k by μS_k. Although the exact thickness of different body parts does vary due to the variety of body types, they can be cancelled out by the self-calibration scheme introduced later.

After that, the mass of moving body part k is estimated by:

$$m_k = \rho V_k, \tag{7}$$

where ρ represents mass density.

By combining equations (5)-(7), we can get:

$$m_k = \rho \mu D_k c_k, \tag{8}$$

where we can combine the 3 unknown coefficients ρ, μ and D_k into one coefficient Y_k:

$$m_k = Y_k c_k. \tag{9}$$

In practice, we do not have to care about the absolute value of body mass m_k because it is proportional to the reflected power c_k. Since the actual relation between vital signs and movement power will be learned by our LSTM model automatically, the relation between m_k and c_k will be encompassed into the model implicitly. Therefore, we can set Y_k to an arbitrary constant, such as 1.

4.2 Sensing the driving force of movement

The phase variation of IF signal between two radar sensing frames can be used to estimate the distance variation of one body part:

$$\Delta d_k = \frac{\lambda \Delta \phi_k}{4\pi}, \tag{10}$$

where $\Delta \phi_k$ represents the phase difference between two adjacent frames, λ represents wave length, k represents

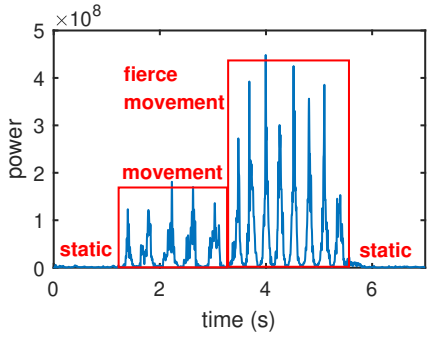

Fig. 15. Real-time power produced by the movement detection module. The subject experiences 4 movement stages: standing still, intermediate movement (jogging), fierce movement (running), standing still again.

range bin index.

Since the adjacent body parts are connected and share similar movements, we make an approximate assumption that the body parts equidistant to the radar have the same distance variation. Even in exceptional cases where adjacent body parts have different distance variation, we consider that there exists one dominant value to represent the distance variation.

With Δd, we can derive the speed v_k and acceleration a_k of body part k:

$$v_k = \frac{\Delta d_k}{T_f}, a_k = \frac{\Delta v_k}{T_f}, \tag{11}$$

where T_f represents the period of radar frames. Since v_k and a_k are 1^{st} and 2^{nd} order derivatives of d_k, they may be more noisy than d_k itself. We can use a first-order low pass filter to smooth them. The cut-off frequency can be simply set to the maximum movement frequency, such as 10 Hz.

For body parts equidistant to the radar, we use F_k to represent the force and calculate it based on the basic physical force formulation:

$$F_k = m_k a_k. \tag{12}$$

4.3 Estimation of the total movement power

Since the radar frame is typically quite short (on the order of milliseconds), we assume the force F that drives the body movement remains constant within the period. So the physical workload of the body parts in the k-th range bin and i-th frame is:

$$W_k^i = F_k^i \Delta d_k^i. \tag{13}$$

The workload for all the body parts of frame i can be calculated as:

$$W^i = \sum_{k=1}^{M} (F_k^i \Delta d_k^i), \tag{14}$$

where M represents the total number of range bins. We can convert the physical workload to movement power by:

$$P^i = \frac{W^i}{T_f}, \tag{15}$$

where T_f represents frame periodicity.

A showcase of movement power estimation. Fig. 15 showcases the real-time output of the movement power detection module when the test subject goes through different phases of movement. The power is nearly zero when the subject is static. Then it increases and manifests regular patterns when the subject performs jogging at different intensities. The power peaks correspond to the maximum speed of the body parts, likely from the limbs. The power level variation over time correspondingly reflects the intensity variation of body movement. This benchmark experiment verifies that our movement power estimation method can faithfully keep track of the relative intensity of body movement.

5 HEART RATE ESTIMATION

5.1 Correlation between HR and motion

Our HR estimation model builds on empirical observations and intrinsic physiology of human body. To investigate the correlation between movement power and HR, we conduct a benchmark experiment to measure a subject's ground-truth HR with different movement patterns, including periodical movement (jogging) and random movement (*e.g.*, swinging arms and turning around). We adopt the Polar H10 heart rate monitor chest strap [26]. The Polar H10 is based on electrocardiograph (ECG) principle, which is also adopted by medical-level equipment [32]. Additionally, it is specifically designed to guarantee robustness under movement scenarios, particularly

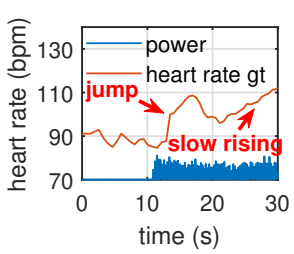

Fig. 16. Multi-stage HR variation[35] under periodical movement. Some common patterns exist in all people.

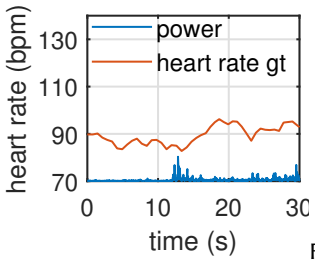

Fig. 17. Heart rate variation under random movement. The variation are mainly affected by amplitude of power.

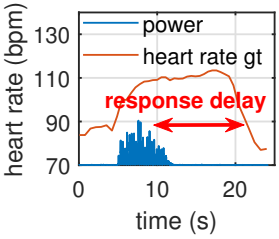

Fig. 18. Heart rate response delay phenomenon after movement. The delay is nearly constant on the same person.

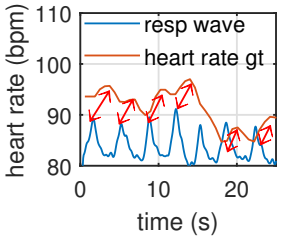

Fig. 19. Respiratory sinus arrhythmia (RSA) [31] phenomenon observed on a static subject.

using chest strap to make sure the sensor contact tightly with the body. Therefore, it can serve as HR ground-truth provider. Fig. 16-19 show the results.

From Fig. 16-17, we observe that the HR rises to a higher level regardless of the type of movement involved. More specifically, the HR increases rapidly at the beginning of movement, followed by slower increase and then plateaus at certain level. At the end of the movement (*i.e.*, movement power drops to nearly 0), the HR shows an inertia effect–staying high for a certain period and then drops to the same level as the static case (Fig. 18). Our observation echoes previous medical research [21, 35] which involves a large group of human subjects.

Note during the subject's movement, the HR value shows minor fluctuations even though the subject attempts to maintain consistency of activities. We conjecture that two factors may cause this phenomenon: respiration and movement intensity. In clinical and basic cardiology, it is known that respiration can affect HR [31]. HR increases during inhale and decreases during exhale, an effect generally referred to as *respiratory sinus arrhythmia (RSA)*. From our measurement in Fig. 19, we also see that the HR exhibits a synchronized fluctuation with the respiration wave. However, when comparing Fig. 16 and Fig. 19, we can find that the HR fluctuation caused by RSA is roughly 10 times lower than movement power ($\sim$ 2 bpm vs $\sim$ 20 bpm). It indicates that the RSA phenomenon has a much lower impact on HR fluctuation, so we may neglect its influence and attribute the HR fluctuation mainly to the movement power variation. There are also other factors that would impact the HR, e.g., gender, ages, weight. We summarize them based on existing medical research and list them in table. 1.

5.2 Self-calibrated LSTM model for HR estimation

Based on the above observations, we propose a customized LSTM structure, which can recover the HR patterns under motion interference.

LSTM [11] is a kind of recurrent neural network (RNN) which learns to map the input data sequence to ground-truth data sequence. Through the design of "gates", LSTM achieves better fitting performance on long sequences than traditional RNN. To map the sequence of noisy radar samples to a sequence of HR estimations, LSTM is a natural fit. The vanilla LSTM suffers from poor generalization when applied directly to robust HR estimation. Since the impact of motion on HR may vary among different subjects, the LSTM needs to be trained separately for each subject. This is unrealistic in practice, as it requires that each subject provide a large training data set with ground-truth HR labels. To overcome the barrier, we propose a *self-calibrated LSTM architecture*, as shown in Fig. 20. To ensure generalization across subjects without per-subject training, we incorporate two mechanisms: instant calibration and adaptive training, for calibrating the error of each single period of movement and long-term generalization error, respectively.

Instant calibration. The basic idea behind instant calibration is to calibrate the HR prediction model when the subject ends a period of movement and becomes static again. When the LSTM begins to apply on a new

Table 1. Relation between HR and various factors.

Age	The resting and average heart rates were not affected by age [14]. In contrast, the age only affects the maximum tolerable HR, which means that the HR in motion should not exceed it.
Gender	The HR of females is averagely higher than male because they have lower global autonomic activity [19].
Weight	The obese people usually have higher HR than normal people [9, 47].
Motion	The exact HR during motion is mainly dependent on motion intensity and personal body physical fitness [20]. The impacts of other attributes (gender, age) are relatively negligible.

subject, its estimated HR would have large error because it has not learned the new HR variation features. For a continuous period of movement, we assume that majority of the estimated HR values are either above or below the ground-truth. The sign of the estimation error for a random sample would likely be the same as the sign of the average estimation error within this period. Since the RF sensing error under static case is negligible [2], we can regard the direct sensing results at static instances as ground-truth. Therefore, we can use the error at a static instance to calibrate the average error of the previous period of movement:

$$H_C'^i = H'^i - D_H, \tag{16}$$

where H'^i represents the estimated HR by LSTM at time i; $H_C'^i$ represents the calibrated HR; D_H represents the estimation error at the moment when the subject becomes static. With a small probability, the sign of D_H may not be the same as the average error. Therefore, we add a coefficient q to reduce the impact of instant calibration:

$$H_C'^i = H'^i - q * D_H. \tag{17}$$

When q is too small (< 0.1), it would take very limited effect for calibration. On the other hand, a large q (> 0.5) may worsen the performance. Our empirical observation shows that a moderate value in the range of [0.2-0.4] would lead to reasonable results.

Adaptive training. Before applying the model to real users, we first train a generic LSTM based on a single dataset, which contains the ground-truth HR of multiple subjects under body movement. Whenever the model is applied to users unseen before, we apply adaptive training to fine tune the parameters, to improve generalization performance. The physical intuition behind adaptive training is that the HR/RR maintain their levels in a nearly constant manner right after the change of movement state [21, 35]. For adaptive training, the ground-truth data is opportunistically obtained at each static moment through direct RF sensing. In practice, the users typically alternate between static and mobile states, which enables such calibration opportunities. Although this may require additional training time, it obviates the need for cumbersome equipment such as chest ECG.

Decision of the status (static, movement) of users: since the instantaneous movement power can show the movement status distinctly, as shown in Fig. 15, we use a simple threshold p_{th} to classify the movement status. We compare the mean value of movement power p_m to the p_{th}. If p_m is larger than p_{th}, it is classified as movement and vice versa. Since the movement power differs greatly between static status and movement, the classification accuracy is not sensitive to the exact value of p_{th}, which can be set in a large range of [1e6, 1e7]. Through this simple method, the movement status can be classified very accurately. We list all the tunable parameters in the proposed model in table. 2.

Table 2. All tunable parameters.

Tunable parameter	Usage	Suggested value
q	instant calibration	0.2-0.4
p_{th}	movement status classification	1e6-1e7

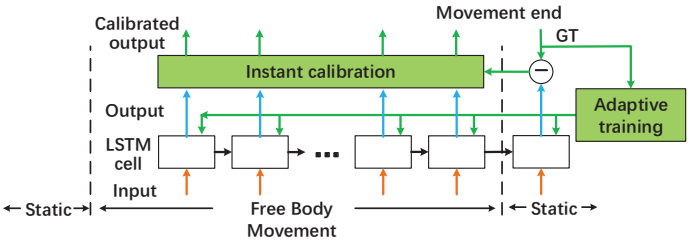

Fig. 20. Self-calibrated LSTM architecture for HR estimation.

6 RESPIRATION RATE ESTIMATION

6.1 RR patterns under different kinds of movement

To gain insights into the correlation between movement power and RR, we measure the ground-truth RR under periodical movement and random movement from different persons, similar to the HR measurement setup in Sec. 5.1. The ground-truth RR is collected by using the Go Direct Respiration Belt [41], which essentially senses the variation of force throughput caused by the respiration fluctuation on the chest. The Belt's sensing accuracy is determined by the force sensor resolution of data collection, *i.e.*, 0.1 Newton. Even under motion, a subject's respiration produces $5-10$ Newton variation on the belt. Therefore, the resolution of the belt is sufficient to monitor respiration despite the motion artifacts.

From the result in Fig. 21, we find that the duration of a full cycle of inhale and exhale contains a fixed number of movement power peaks. It implies that the *respiration pattern is synchronized to the fluctuation of movement power*. The observation corroborates existing physiological research on the mutual influence between respiration and body movement [25, 30], which showed that the mutual influence between respiration and body movement is general in human. During any physical activities, a subject needs to keep her movement and respiration synchronized, to avoid extra energy consumption due to the out-sync. Whenever out-sync occurs, she would adjust them and resync unconsciously to save energy. Additionally, different from HR variation, we find that RR does not increase obviously by the increase of movement intensity. It seems to more related to the movement frequency.

In the random movement experiment (Fig. 22), we find that RR does not change much under low movement intensity. On the other hand, with more intensive movement, the RR value shows a more complicated relation with movement without any tractable patterns.

6.2 Stacked LSTM for movement pattern and RR estimation

To capture the sophisticated correlation between movement and RR especially under random activities, a data-driven model with higher representation capacity is needed. We thus design a stacked LSTM model, whose structure can leverage the aforementioned experimental insights.

As shown in Fig. 23, the model comprises two sub-LSTMs, which first explicitly classifies the movement patterns, and then runs the RR estimation. The reason for using two sub-LSTMs instead of a single comprehensive one for RR estimation is that we find that explicitly estimating the RR patterns can help boost the RR estimation accuracy. The first sub-LSTM is trained to learn the RR patterns explicitly and the second sub-LSTM adopts both the RR patterns and the movement power to give more accurate estimation. The bottom LSTM accepts the original movement power sequence as input. It contains 3 LSTM cells with decreasing kernel size of 128, 64 and 32, respectively. A dropout layer [10] is inserted between each two LSTM cells to prevent overfitting. The final layer is a softmax structure which outputs the probability of each movement pattern. Our current model incorporates two classes of patterns, *i.e.*, periodical and random. As for the top LSTM, the input is the

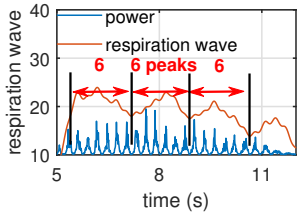

Fig. 21. Mutual influence between respiration and movement [25, 30] under moderate periodical movement.

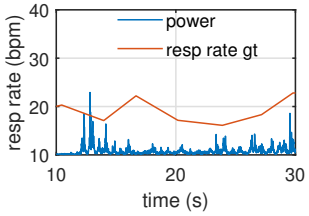

Fig. 22. RR variation under random movement, including turning around and waving arms.

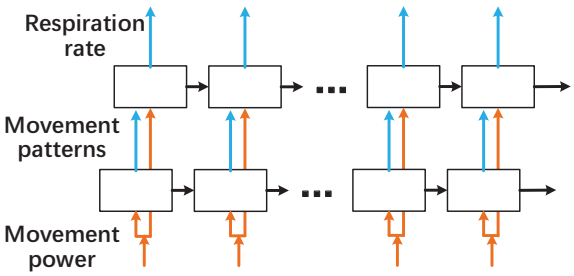

Fig. 23. Stacked LSTM architecture for RR estimation.

concatenation of movement patterns and movement power. It contains 2 LSTM cells with the same kernel size of 32. We find through experiments that using the same kernel size here shows best accuracy/size ratio compared with other structures (such as decreasing kernel size). The final layer is a dense layer that outputs the estimated RR. The estimated RR does not involve normalization because the top LSTM is shallow (2 LSTM cells) and the gradient vanish problem is unlikely to take effect. Additionally, since the length of all data sequences in our dataset (refer to Section 7) is the same, the output length of dense layer is set to the same length as the data sequence (300 samples in our implementation) for faster training.

We train the two sub-LSTMs separately. Specifically, we first collect data for each type of movement to train the bottom LSTM. Then we freeze its parameters and train the top LSTM. To ensure cross-subject generalization, we apply the the adaptive training scheme (Section 5.2) again when applying the model across different subjects.

7 SYSTEM IMPLEMENTATION

RF sensing hardware. We use the TI 77 GHz millimeter-wave (mmWave) FMCW radar as the sensing platform [36], as shown in Fig. 24. The main advantages of mmWave radar, in comparison to low-frequency RF devices, lie in its high range resolution. Owing to the short wavelength, the radar's return signal phase is highly sensitive to tiny subject movement. In addition, the wide bandwidth (4 GHz) translates into high time-of-flight resolution. Such high spatio-temporal resolution ensures that the signal features likely embody the minute movement of body parts and vital organs.

Table 4. Baseline HR and RR errors in static circumstances (bpm). The 'be' represents before movement and 'af' represents after movement.

Model	HR-be	HR-af	RR-be	RR-af
VitalRadio[2]	1.62	2.81	0.26	0.38
DopplerSleep[29]	1.34	5.60	0.25	0.40

Data collection. We collect a dataset on 14 different subjects, covering 5 ethnicities and different height, weight and gender. The statistical information of these subjects are listed in Table 3. For each subject, 18-45 groups of data are collected, each lasting for 30 seconds and containing a mix of static and moving scenarios. During the data collection, a subject performs a wide range of typical daily activities, such as normally walking, doing fitness exercises and stretching limbs. The subject also varies his/her position, distance and relative orientation to the radar. For comparison, we collect the ground-truth HR and RR by using medical-grade equipment: Polar H10 monitor chest strap [26] and Go Direct Respiration Belt[41], as shown in Fig. 25.

Meanwhile, the radar continuously transmits signals towards the subject, which stays within the radar's field of view despite movement. The ground truth HR and RR data are streamed to a PC host through Bluetooth, and eventually synchronized with the radar data through the PC's time stamping. In practice, we can also use the static detection method in [2] to opportunistically obtain accurate HR/RR sensing data when the subject is in a

Table 3. Statistical information of participants.

Age	19-23 (4), 24-28 (7), 28-30 (3)
Height	1.60-1.70m (3), 1.70-1.80m (8), >1.80m (3)
Gender	Male (13), Female (1)
Weight	48-60kg (4), 60-70kg (3), 70-80kg (4), 80-90kg (2), >90kg (1)

static state. The method is reported to achieve an accuracy of 99%, which suffices for our adaptive calibration mechanism.

Since there is only one female in our dataset and all subjects are under 30 years, one may wonder whether gender and age would affect the HR. We think gender and age would not affect the performance of our model for three reasons: First, existing medical research has shown that gender and age would not have major impacts on HR. The exact HR during motion is mainly dependent on motion intensity and personal fitness [20]. The impact of gender is negligible. The resting and average heart rates are not affected by age [14]. In contrast, the age only affects the maximum tolerable HR, which means that the HR in motion should not exceed it. Second, the weaker heart beats of women and elders would not affect our model because our model does not directly sense the vibration of heart beats. Instead, our model senses the body movement intensity, which can be easily detected in people with any ages and gender. Third, although weaker heart beats of women and elders may affect the static HR detection, we note that our model is able to predict HR during movement. The heart beats after movement is much stronger than peace and thus are easier to be detected.

Model implementation. We use Pytorch [27] to implement the deep learning models for HR and RR estimation. For HR estimation, we use a 2-cell LSTM comprised of dense layers with kernel size 128. We use Stochastic Gradient Descent (SGD) with weight decay of $1e^{-4}$. The initial learning rate is $1e^{-2}$ and the momentum parameter is 0.8. Simple mean square error (MSE) is adopted as loss function. The batch size is set to 8 and the total training epoch is 200. For RR estimation, we also use a 2-cell LSTM, but the kernel size is set to 256, the initial learning rate is $5e^{-3}$ and the momentum parameter is 0.9. The average training time of each epoch is 6 minutes for HR estimation and 8 minutes for RR estimation on a server with GPU of Nvidia RTX 2080 and CPU of i9-9980XE. The self-calibrated LSTM model is able to run on a common personal computer with standalone GPU, e.g., LEGION Y7000P with GPU of GTX 1660 and CPU of i7-8750H. The inference time for a 30 seconds input sequence is usually within 5 seconds (on a laptop with NVIDIA 1660), which should suffice for near real-time applications, such as fitness tracking. To identify proper hyper parameters, we empirically balance the model overfitting and underfitting. We adjust the parameter size by observing the training loss and validation loss. When the model is overfitting (low training loss and high validation loss), we decrease the parameter size and vice versa.

Baseline methods. For comparison, we implement two state-of-the-art RF vital sign sensing systems: DopplerSleep [29] and VitalRadio [2]. Since the original source code is not publicly available, we implement them based on their description. DopplerSleep is a representative frequency analysis method. It aims to sense vital signs in sleep scenarios where the subject's movement is constrained to low speed, such as turning over on the bed. Following [29], we set the RR estimation range to 9-20 bpm and the HR estimation range to 45-80 bpm. We use a bandpass filter with stop band frequencies at 0.1 Hz, 0.8 Hz and pass band frequencies at 0.3 Hz, 0.7 Hz for RR noise reduction. Similarly, we use use a bandpass filter with stop band frequencies at 1 Hz, 3 Hz and pass band frequencies at 1.5 Hz, 2.5 Hz to remove the noise in the HR sensing spectrum. For both filters, we set the pass band ripple and stop band attenuation to 1 dB and 60 dB respectively.

We further implement the VitalRadio approach [2], which adopts a periodicity interval method to isolate aperiodic movement artifacts. Following [2], we set the length of observing time window to 30 seconds. We set the band-pass filter with pass frequency range of 40-200 bpm and stop band attenuation to 60 dB.

In order to verify the performance of the implemented baseline models, we evaluate them in static situations. The test datas are divided into two situations: standing still before movement and standing still after movement.

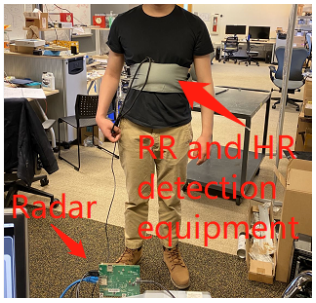

Fig. 24. Experiment setup.

Fig. 25. RR and HR ground-truth devices.

The results in Table. 4 show that the DopplerSleep and VitalRadio have average HR errors of 3.47 and 2.21 bpm and average RR errors of 0.33 and 0.32 bpm, which are pretty close to their original results. The VitalRadio has lower error on 'after movement' due to two reasons: First, the HR and RR after movement is larger than the 'before movement' situation and has higher frequency. Second, the DopplerSleep is designed for sleep monitoring applications, which essentially has an unsuitable filter range for movement cases.

8 SYSTEM EVALUATION

8.1 HR and RR estimation at a fixed position

In this experiment, we vary the distance between the radar and subjects from 1m to 3m. Under each distance setting, the subject performs different kinds of body parts movement at a fixed position, including periodical movement (jogging) and random movement (stretching arms and turning around). The training data and validation data come from the same subject, but are orthogonal to each other. We will evaluate the cross-subject generalization later.

We validate the proposed HR estimation model using a leave-one-out method, *i.e.*, among the N groups of data, one is selected as validation set and the other $N-1$ groups constitute the training set. We repeat the validation experiment N times. For this experiment, we do not use the self-calibration method.

Table 5. HR estimation errors under different motion and different distance (unit: bpm). The 'per' refers to 'periodical movement'. The 'ran' refers to 'random movement'.

model	1 meter			2 meter			3 meter			average
	per	ran	all	per	ran	all	per	ran	all	
VitalRadio[2]	43.57	52.62	48.10	57.83	44.94	51.38	52.50	46.51	49.51	49.66
DopplerSleep[29]	20.24	15.67	17.96	14.54	19.06	16.80	13.22	14.65	13.94	16.23
LSTM(HR)	4.94	4.15	4.54	9.01	5.27	7.14	8.12	4.44	6.28	5.99

Table 6. RR estimation errors under different motion and different distance (unit: bpm). The 'per' refers to 'periodical movement'. The 'ran' refers to 'random movement'.

model	1 meter			2 meter			3 meter			average
	per	ran	all	per	ran	all	per	ran	all	
VitalRadio[2]	6.76	5.22	5.99	5.95	6.78	6.36	5.72	6.03	5.87	6.07
DopplerSleep[29]	5.81	4.13	4.97	4.30	5.92	5.11	4.62	5.51	5.07	5.05
LSTM(RR)	5.61	1.49	3.55	2.93	3.08	3.01	2.14	1.93	2.03	2.86

Table 5-6 show the HR and RR estimation results respectively. The proposed model has an average HR and RR error of 5.34 and 3.63 bpm, which outperforms all the state-of-the-art baselines. In addition, our model performs similarly under periodical and random movement, indicating the design of the movement detection module is robust to the variation of movement patterns. In different distance settings, the proposed model does not show

obvious performance difference, indicating that it is insensitive to signal strength variations.

In comparison, the frequency analysis based DopplerSleep [29] shows an average HR and RR error of 16.23 and 5.05 bpm, which is unreasonably high (around 203% and 39% higher than our model). Since this method is designed for restricted scenarios, it fails even under mild motion artifacts. Additionally, we test the periodicity interval method to isolate movement artifacts in VitalRadio [2]. The average HR error is 49.6 bpm, which is the highest among all methods. This approach assumes that the movement is aperiodic and rare, which is not applicable to certain daily activities like jogging. Additionally, VitalRadio also leads to large errors in random movement scenarios, because our dataset contains normal ambulant activities with frequent limb motion.

8.2 HR and RR estimation under daily ambulant conditions

We now conduct experiments in more realistic situations where the subjects can walk normally in a square area in front of the radar without any restriction of position and direction. The radar-to-subject distance ranges from 1-6 meters and direction varies by ± 60°. In the 'front-back' setting, the subjects walk in a straight line back and forth, so only the *distance varies*. In the 'left-right' setting, the subjects walk in a straight line from left to right in front of the radar, so the *direction variations dominate*. In the 'comprehensive' setting, the subjects are asked to walk along arbitrary trajectories. Similar to the previous experiment, the training data and validation data come from the same subject and the leave-one-out validation method is used.

Table 7. HR errors in ambulant conditions (bpm). The 'Com' refers to 'Comprehensive'.

model	1-3 meters			4-6 meters			average
	Front-back	Left-right	Com	Front-back	Left-right	Com	
VitalRadio[2]	47.83	55.33	55.11	52.54	49.75	53.66	52.37
DopplerSleep[29]	18.57	12.37	13.59	17.85	20.13	19.86	17.06
LSTM(HR)	3.36	3.29	4.62	4.15	5.27	4.66	4.22

Table 8. RR errors in ambulant conditions (bpm). The 'Com' refers to 'Comprehensive'.

model	1-3 meters			4-6 meters			average
	Front-back	Left-right	Com	Front-back	Left-right	Com	
VitalRadio[2]	6.57	6.83	7.22	6.62	7.76	7.25	7.04
DopplerSleep[29]	5.81	4.37	5.05	5.86	6.03	5.79	5.48
LSTM(RR)	1.99	2.59	3.23	2.78	3.25	3.64	2.91

Table 7-8 show the HR and RR estimation results respectively. Our model has an average HR and RR error of 4.22 and 2.91 bpm in all the ambulant conditions, which outperforms all the baseline models. In addition, we find that our model performs similarly under different ambulant situations, indicating its robustness under daily ambulant conditions and the ability to accommodate both direction and distance variations.

Similar to the fixed-position experiment in section 8.1, DopplerSleep and VitalRadio suffer from severe performance degradation under motion, showing large deviation off the ground-truth (around 15% and 35% deviation for HR and RR by DopplerSleep, and 60% deviation for HR by VitalRadio). It indicates that they both fail under motion interference caused by daily ambulant activities.

We also provide a detailed trace showing the instant HR and RR errors in Fig. 26, which corresponds to the 'comprehensive' setup. From the illustration in Fig. 26 we can draw 3 major conclusions. First, the average HR and RR errors within 1-3m are 3.76 and 2.6 bpm, which are relatively lower than over longer distances (over 4m). Second, longer distances correspond to larger HR and RR errors. For HR Estimation, the average error of 4-6m is 4.69bpm, which is 25% higher than 1-3m. for RR Estimation, the average error of 4-6m is 3.22bpm, which is 24% higher than 1-3m. The main reason for the increase in error is the limitation of radar power. As the distance increases gradually, the reflected signal of the target becomes weak and is easily interfered by the ambient noise. By using more powerful mobile radar (such as AWR2243), it is possible to achieve high precision detection over a

long distance. Third, the errors close to both sides are larger than middle areas. For HR estimation, the errors over 90° is about 5.23 bpm, which is 21% larger than within 90°. For RR estimation, the errors over 90° is about 2.79, which is 43% larger than within 90°. The main reason is that the radiation intensity of radar is smaller on both sides but larger in the middle, which makes the reflected signals on both sides vulnerable to environmental noise interference. This phenomenon tends to occur over long distances and can be also resolved by using the more powerful mobile radar.

Existing medical research [24] shows that the HR estimation accuracy of most accurate commercial devices (Polar H7 band, iWatch) are 96%-98%, which are equivalent to 6bpm-3bpm estimation errors. Therefore, the average error 4.22 bpm of our model is sufficient for monitoring accurate HR during fitness. Existing medical research [40] also shows that the average RR detection error of 7.9 bpm leads to 96% of adequate treatment decisions. Additionally, the expected error for 100% adequate treatment decisions corresponds to an RR error of 2bpm. Therefore, the average error of 2.91 bpm of our model is very close to achieving the requirement for clinical judgement.

8.3 Generalization capability of the model

8.3.1 Generalization across subjects. In this section, we verify the proposed self-calibration schemes for improving the generalization performance.

We conduct experiments with the training and validation sets coming from same/different subjects, respectively. Each subject generates 18 groups of data, involving 12 groups of fixed-position movement and 6 groups of ambulant movement. The specific movement types are similar to Section 8.1-8.2. In the 'self' setting in Fig. 27, all the subjects are trained and tested using their own data by the leave-one-out method, which guarantees orthogonality of training and testing sets. In the 'cross' setting in Fig. 27, the training set for each subject comes from another subject. The HR estimation results, marked as "self" and "cross" accordingly in Fig. 27, show that the cross-subject estimation error increases obviously compared to the "self" setup, almost for all the subjects (12 out of 14). In particular, the maximum cross-subject error increases by 15.7 bpm for person 7, which is nearly 2× compared to the "self" error. The results imply that the HR patterns differ across people even under similar motion artifacts. Similarly, the cross-subject RR estimation error (Fig. 28) increases for more than half of the subjects (8 out of 14).

We now verify the effectiveness of the proposed self-calibration methods, *i.e.*, adaptive training and instant calibration. Recall that adaptive training opportunistically leverages the static moments to create calibration instances. We thus truncate the collected data to different lengths to create an enriched dataset to emulate different segments of ambulant conditions. For each segment of data, we only use the last point of data to mimic the static instances in practical scenarios that involve a mix of ambulant and stationary periods. We use the "n-point" term to denote that n points of data are used for adaptive training. All the LSTM models have been pretrained using the aforementioned "cross" setup, denoted as "cross" in Fig. 27. After applying 2-point adaptive training, the HR estimation errors of all subjects are decreased by 18% on average. Subject 13 benefits the largest error reduction of 11.3 bpm, whereas the lowest error reduction is 0.4 bpm on person 1. The difference is likely because the estimation error of person 1 is already quite low. The RR comparison results in Fig. 28 show that the 2-point adaptive training show 9% error reduction compared to the 'cross' setup, indicating that the adaptive training scheme is also effective for generalizing RR estimation. Remarkably, after applying 4-point adaptive training, all subjects achieve even lower HR and RR estimation error, *i.e.*, 6.02 and 3.23 bpm, respectively. Therefore, more calibration points help reduce the generalization error, bringing it closer to the single-subject case. But the improvement becomes marginal as more calibration points are added beyond 4, as the bottleneck may be bounded by the noise and the model capacity itself. Additionally, we also notice that there are some unexpected results. In Fig. 27, the 'cross' errors of subject 2 and 5 are lower than the 'self' setup. We note the error differences are only 0.79 and 0.82 bpm, which are only one seventh of the average error difference (5.53 bpm). Therefore, these

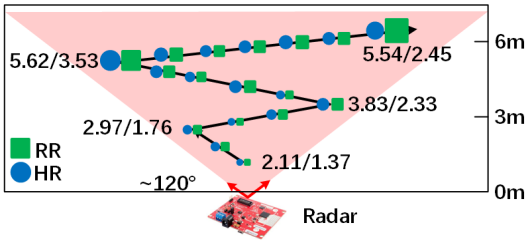

Fig. 26. Visualization of the instant HR and RR errors by trajectory. We use the symbols of blue circles and green rectangles to represent the HR and RR estimation error. The size of the symbols refer to the error value.

error differences can be regarded as normal model noise. Another unexpected result is that the self-calibration errors of subject 2,5 and 14 with 2-point have higher error than 'cross' and 'self' setup. The main reason is that the self-calibration process may introduce errors due to the instant calibration when the calibration iteration is not enough. As explained in Sec. 5.2, a large q of instant calibration may worsen the performance. However, we should notice that the induced errors are only 13% of the 'cross' setup on average, which do not cause obvious performance loss. Additionally, we can find that the 4-point calibration eliminates these errors, indicating that the problem can be addressed by more calibration iterations.

8.3.2 Generalization across environment. Different environment may experience different RF background reflections to radar, and thus may impact the detection performance in different ways. Therefore, we run the vital sign estimation model on 3 different places to check how it performs in new environment. All 3 places are indoor with rich multipath reflections, caused by furniture, appliances, room layout, *etc.* Place 1 has no furniture inside the experimental area; Places 2 and 3 have sofas and chairs. The training data only comes from place 1 and the validation data comes from all 3 places. Table 9-10 show the HR and RR estimation results respectively. We can find that both the HR and RR estimation errors in the other 2 places do not have obvious difference from place 1, indicating that our model model can effectively work across different environment. Besides the learning model itself, this is also attributed to the simple filter preprocessing which mitigate background noise by subtracting static reflections.

Table 9. HR estimation errors in different places (bpm).

	place 1	place 2	place 3
periodical	6.95	4.56	2.27
random	4.76	3.07	3.06
ambulant	4.64	7.53	4.35

Table 10. RR estimation errors in different places (bpm).

	place 1	place 2	place 3
periodical	5.21	1.20	3.44
random	1.56	1.76	3.60
ambulant	1.48	1.68	2.69

8.4 HR and RR estimation in noisy environment

Table 11. HR errors in noisy background (bpm).

	periodical	random	ambulant	All
w/o noise	4.83	3.37	3.03	3.74
w/ noise	4.62	5.37	3.34	4.45

Table 12. RR errors in noisy background (bpm).

	periodical	random	ambulant	All
w/o noise	3.93	4.90	2.90	3.91
w/ noise	5.70	2.14	1.62	3.15

In reality, multiple people may be co-located in the same environment. We now evaluate the performance of our vital sign sensing model subject to interference from ambient activities. During the testing, the subject performs various kinds of movement, such as periodic, random and ambulant activities, similar to those mentioned in sections 8.1-8.2. Another person acts as an interfering user walking around the subject to create ambulant background noise. During the testing, the interfering user continuously walks around the subject with distances of 1.5-2m. The walking area is always within the +- 60° view of radar. Table 11-12 show the resulting HR and RR estimation, respectively. We can find that the largest deviation between w/ and w/o noise in table 11 is 2 bpm of random activity, which is equivalent to only 1.3% of the average HR. Also, there is no obvious performance

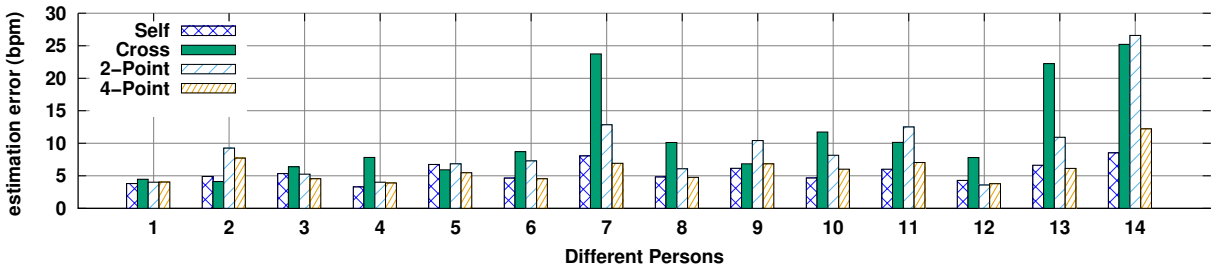

Fig. 27. Cross-subject generalization capability of the HR estimation model.

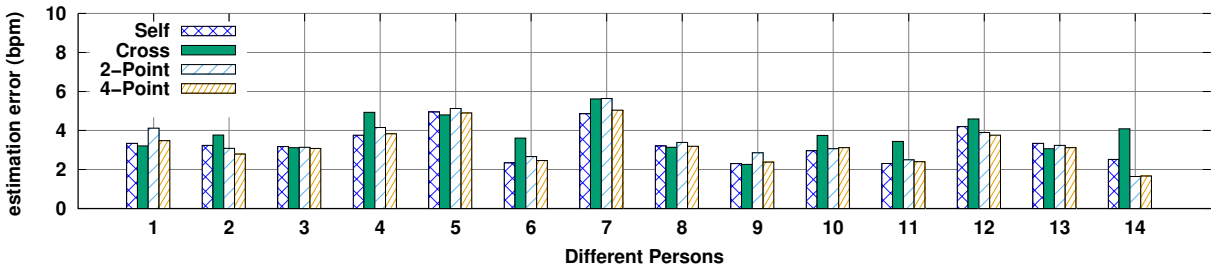

Fig. 28. Cross-subject generalization capability of the RR estimation model.

deviation between w/o and w/ noise in table 12, indicating that the model is robust to background human activities. Additionally, we find that RR estimations with noise gain much lower errors than without noise in 12. The main reason is that the noise does not have an obvious impact on RR estimation. Therefore, the error deviation between w/ and w/o noise in table 12 is not caused by the noise but the model estimation itself.

9 RELATED WORKS

Radar based vital sign sensing technologies date back to the 1970's [17], and have been revisited over the past decades as new RF hardware and use cases emerge [2, 29]. Early studies mostly focused on controlled scenarios with a single subject sitting still near the radar [16]. More recent work addressed more challenging scenarios involving multiple persons [45] and longer range [2]. To fully explore the potential of radar, certain specialized cases have also been investigated, such as recovering the breathing signals of multiple subjects very close to each other [49] and monitoring vital signs by attaching radar in tricky position (waist) of body [7]. Besides high-end radar devices, commodity communication devices such as WiFi can track RR and even HR [43, 46]. In effect, the phase change rate and micro-Doppler profile [28] can be extrapolated from the WiFi channel state information (CSI), and mapped to vital activities.

Nonetheless, conventional RF vital sign detection algorithms only work for static subjects, as vital sign induced signal features can easily become immersed under motion artifacts. Recent work attempted to mitigate the motion artifacts through adaptive filtering techniques. For example, Mostafanezhad *et al.* employed empirical mode decomposition to cancel the motion artifacts [22], but the solution only applies to restricted vibration movement. Sharafi *et al.* [33] used a correlation algorithm to improve the vital sign SNR in the presence of motion interference. Tu *et al.* [39] extended the scenario to 1-D whole-body movement and extracted respiration rate. Adib *et al.* [2] addressed the smart home scenarios where the subject is seated and performing in-place activities such as typing. All these schemes assume restricted movement activities, so the motion artifacts can be isolated through filtering. In contrast, our work tackles the more general and practical scenarios where the subject can move freely, without restricting the movement forms, position and direction.

The impact of random body movement can potentially be alleviated through sensor fusion. Gu *et al.* [12] used

a camera to capture simple back-and-forth body movement, and compensate for the phase variation of radar signals. Wang *et al.* [42] proposed to fuse the signals from two radars placed on the front and back side of the subject, thus suppressing the noise caused by periodic movement (*e.g.*, jogging). These proposals entail more costly hardware and further limit the usage scenarios.

RF based vital sign sensing has been explored in a variety of usage scenarios, such as elder care [6] and sleeping monitoring [29]. The actual sensing problem in these scenarios are similar to those under restricted position and limited body movement.

Besides mmWave radar based methods, alternative sensing paradigms exist for vital signs. Leal *et al.* [15] used polymer optical fiber materials which can be woven into clothes to sense the in-body vibration. Brink *et al.* [4] adopted four high-resolution force sensors under bedposts to sense the vital signs. Wearable devices like smartwatches use Photoplethysmogram (PPG) to realize unobtrusive monitoring of heart rate [1, 3]. Unlike RF based solutions, these systems require close contact between the sensors and human body. Since Kinect [5] can extract distance variation to the subject, it can be used to monitor the respiration fluctuation on the chest. Additionally, by placing color fiducials on the chest of subject, optical flow [44], which is an vector representation of movement extracted from video, can also be used to sense respiration fluctuation from chest. However, these camera based solutions require the subject being fixed in position or direction.

10 LIMITATIONS

Although the proposed methods can address the HR/RR detection under daily ambulant environment, there exist a few limitations that deserve further research.

(1) Impacts of daily objects on users. All subjects in our experiments wear normal clothes and do not have any personal objects on their bodies (*e.g.*, handbag and exercise equipment). Large objects may temporarily increase the estimated weights during the movement power detection, causing higher estimated HR/RR. One possible solution is to combine with other sensors, *e.g.*, smart scale and smart in-home camera, and mitigate the drifting through a learned fusion model.

(2) Application to users with certain cardiovascular diseases. Our approach targets daily usage scenarios such as metabolism and fitness tracking. Our model builds on the correlation between vital sign variation and the motion intensity of body parts. So it is not applicable to the cases where this correlation is corrupted, *e.g.*, for people with certain cardiovascular diseases such as arrhythmia. One possible solution is using static detection method to detect the abnormal HR/RR variation and infer the cardiovascular diseases, which is used as additional information to build more accurate model for movement estimation. This is left for our future work.

11 CONCLUSION

Despite decades of research, RF vital sign sensing systems remain in controlled settings with static subjects or restricted mobility. In this paper, we have demonstrated the feasibility of RF vital sign sensing with free body movement, just as in daily living conditions. Our system pinpoints a new principle to isolate the motion artifacts, by reusing the RF sensor to extrapolate an additional dimension of information (*i.e.*, movement power), and using customized deep learning models to capture its correlation with vital signs. As future work, we plan to explore additional feature dimensions, such as visual information from cameras, which can potentially handle more challenging setup (*e.g.*, multiple co-located subjects).

ACKNOWLEDGEMENT

This research was supported in part by the National Key R&D Program of China under Grant No. 2019YFA0706403, National Natural Science Foundation of China under Grant No. 62072472 and U19A2067, Natural Science Foundation of Hunan Province, China under Grant No. 2020JJ2050, 111 Project under Grant No. B18059, and the Young

Talents Plan of Hunan Province of China under Grant No. 2019RS2001. We also acknowledge the gracious support from Samsung Research America.

REFERENCES

[1] Shyam A, Vignesh Ravichandran, Preejith S. P, Jayaraj Joseph, and Mohanasankar Sivaprakasam. 2019. PPGnet: Deep Network for Device Independent Heart Rate Estimation from Photoplethysmogram. *CoRR* abs/1903.08912 (2019).

[2] Fadel Adib, Hongzi Mao, Zachary Kabelac, Dina Katabi, and Robert C Miller. 2015. Smart homes that monitor breathing and heart rate. In *Proceedings of the 33rd annual ACM conference on human factors in computing systems*. 837–846.

[3] D. Biswas, L. Everson, M. Liu, M. Panwar, B. Verhoef, S. Patki, C. H. Kim, A. Acharyya, C. Van Hoof, M. Konijnenburg, and N. Van Helleputte. 2019. CorNET: Deep Learning Framework for PPG-Based Heart Rate Estimation and Biometric Identification in Ambulant Environment. *IEEE Transactions on Biomedical Circuits and Systems* 13, 2 (2019).

[4] Mark Brink, Christopher H Müller, and Christoph Schierz. 2006. Contact-free measurement of heart rate, respiration rate, and body movements during sleep. *Behavior research methods* 38, 3 (2006), 511–521.

[5] N. Burba, M. Bolas, D. M. Krum, and E. A. Suma. 2012. Unobtrusive measurement of subtle nonverbal behaviors with the Microsoft Kinect. In *IEEE Virtual Reality Workshops (VRW)*.

[6] Giovanni Diraco, Alessandro Leone, and Pietro Siciliano. 2017. A radar-based smart sensor for unobtrusive elderly monitoring in ambient assisted living applications. *Biosensors* 7, 4 (2017), 55.

[7] Biyi Fang, Nicholas D Lane, Mi Zhang, Aidan Boran, and Fahim Kawsar. 2016. BodyScan: Enabling radio-based sensing on wearable devices for contactless activity and vital sign monitoring. In *Proceedings of the 14th annual international conference on mobile systems, applications, and services*. 97–110.

[8] Maria Faurholt-Jepsen, Søren Brage, Maj Vinberg, and Lars Vedel Kessing. 2016. State-related differences in the level of psychomotor activity in patients with bipolar disorder–Continuous heart rate and movement monitoring. *Psychiatry research* 237 (2016), 166–174.

[9] Denise Felber Dietrich, Ursula Ackermann-Liebrich, Christian Schindler, Jean-Claude Barthélémy, Otto Brändli, Diane R Gold, Bruno Knöpfli, Nicole M Probst-Hensch, Fr d ric Roche, Jean-Marie Tschopp, et al. 2008. Effect of physical activity on heart rate variability in normal weight, overweight and obese subjects: results from the SAPALDIA study. *European journal of applied physiology* 104, 3 (2008), 557–565.

[10] Yarin Gal and Zoubin Ghahramani. 2016. Dropout as a bayesian approximation: Representing model uncertainty in deep learning. In *international conference on machine learning*. 1050–1059.

[11] Felix A Gers, Jürgen Schmidhuber, and Fred Cummins. 1999. Learning to forget: Continual prediction with LSTM. (1999).

[12] Changzhan Gu, Guochao Wang, Yiran Li, Takao Inoue, and Changzhi Li. 2013. A hybrid radar-camera sensing system with phase compensation for random body movement cancellation in Doppler vital sign detection. *IEEE transactions on microwave theory and techniques* 61, 12 (2013), 4678–4688.

[13] Te-Yu J Kao, Yan Yan, Tze-Min Shen, Austin Ying-Kuang Chen, and Jenshan Lin. 2013. Design and analysis of a 60-GHz CMOS Doppler micro-radar system-in-package for vital-sign and vibration detection. *IEEE Transactions on Microwave Theory and Techniques* 61, 4 (2013), 1649–1659.

[14] John B Kostis, AE Moreyra, MT Amendo, JOANNE Di Pietro, NORA Cosgrove, and PT Kuo. 1982. The effect of age on heart rate in subjects free of heart disease. Studies by ambulatory electrocardiography and maximal exercise stress test. *Circulation* 65, 1 (1982), 141–145.

[15] Arnaldo G Leal-Junior, Camilo R Díaz, Cátia Leitão, Maria José Pontes, Carlos Marques, and Anselmo Frizera. 2019. Polymer optical fiber-based sensor for simultaneous measurement of breath and heart rate under dynamic movements. *Optics & Laser Technology* 109 (2019), 429–436.

[16] Changzhi Li, Jun Ling, Jian Li, and Jenshan Lin. 2009. Accurate Doppler radar noncontact vital sign detection using the RELAX algorithm. *IEEE Transactions on Instrumentation and Measurement* 59, 3 (2009), 687–695.

[17] J. C. Lin. 1975. Noninvasive microwave measurement of respiration. *Proc. IEEE* 63, 10 (1975).

[18] Lanbo Liu and Sixin Liu. 2014. Remote detection of human vital sign with stepped-frequency continuous wave radar. *IEEE journal of selected topics in applied earth observations and remote sensing* 7, 3 (2014), 775–782.

[19] Mohamed Faisal Lutfi and Mohamed Yosif Sukkar. 2011. The effect of gender on heart rate variability in asthmatic and normal healthy adults. *International journal of health sciences* 5, 2 (2011), 146.

[20] FS Martinelli, MPT Chacon-Mikahil, LEB Martins, EC Lima-Filho, R Golfetti, MA Paschoal, and L Gallo-Junior. 2005. Heart rate variability in athletes and nonathletes at rest and during head-up tilt. *Brazilian journal of medical and biological research* 38, 4 (2005), 639–647.

[21] Dionne Matthew and Anne Delextrat. 2009. Heart rate, blood lactate concentration, and time–motion analysis of female basketball players during competition. *Journal of sports sciences* 27, 8 (2009), 813–821.

[22] Isar Mostafanezhad, Ehsan Yavari, Olga Boric-Lubecke, Victor M Lubecke, and Danilo P Mandic. 2013. Cancellation of unwanted Doppler radar sensor motion using empirical mode decomposition. *IEEE Sensors Journal* 13, 5 (2013), 1897–1904.

[23] Mehrdad Nosrati, Shahram Shahsavari, Sanghoon Lee, Hua Wang, and Negar Tavassolian. 2019. A Concurrent Dual-Beam Phased-Array Doppler Radar Using MIMO Beamforming Techniques for Short-Range Vital-Signs Monitoring. *IEEE Transactions on Antennas and Propagation* 67, 4 (2019), 2390–2404.
[24] Selena R Pasadyn, Mohamad Soudan, Marc Gillinov, Penny Houghtaling, Dermot Phelan, Nicole Gillinov, Barbara Bittel, and Milind Y Desai. 2019. Accuracy of commercially available heart rate monitors in athletes: a prospective study. *Cardiovascular diagnosis and therapy* 9, 4 (2019), 379.
[25] Raffaella Pellegrini and Maria Rita Ciceri. 2012. Listening to and mimicking respiration: Understanding and synchronizing joint actions. *Review of psychology* 19, 1 (2012), 17–27.
[26] Polar. 2019. *Polar H10 heart rate sensor.* https://www.polar.com/us-en/products/accessories/
[27] Pytorch. 2020. *Pytorch website.* https://pytorch.org/
[28] Kun Qian, Chenshu Wu, Zheng Yang, Yunhao Liu, and Kyle Jamieson. 2017. Widar: Decimeter-Level Passive Tracking via Velocity Monitoring with Commodity Wi-Fi. In *Proceedings of the 18th ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc)*.
[29] Tauhidur Rahman, Alexander T Adams, Ruth Vinisha Ravichandran, Mi Zhang, Shwetak N Patel, Julie A Kientz, and Tanzeem Choudhury. 2015. Dopplesleep: A contactless unobtrusive sleep sensing system using short-range doppler radar. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. 39–50.
[30] B Rassler and J Kohl. 2000. Coordination-related changes in the rhythms of breathing and walking in humans. *European journal of applied physiology* 82, 4 (2000), 280–288.
[31] JD Schipke, G Arnold, and M Pelzer. 1999. Effect of respiration rate on short-term heart rate variability. *Journal of Clinical and Basic Cardiology* 2, 1 (1999), 92–95.
[32] José J Segura-Juárez, David Cuesta-Frau, Luis Samblas-Pena, and Mateo Aboy. 2004. A microcontroller-based portable electrocardiograph recorder. *IEEE Transactions on Biomedical Engineering* 51, 9 (2004), 1686–1690.
[33] Azadeh Sharafi, Mehran Baboli, Mohammad Eshghi, and Alireza Ahmadian. 2012. Respiration-rate estimation of a moving target using impulse-based ultra wideband radars. *Australasian physical & engineering sciences in medicine* 35, 1 (2012), 31–39.
[34] Kunsoo Shin, HARUYUKI Minamitani, SHOHEI Onishi, HAJIME Yamazaki, and MYOUNGHO Lee. 1997. Autonomic differences between athletes and nonathletes: spectral analysis approach. *Medicine and Science in Sports and Exercise* 29, 11 (1997), 1482–1490.
[35] Steven W Su, Weidong Chen, Dongdong Liu, Yi Fang, Weijun Kuang, Xiaoxiang Yu, Tian Guo, Branko G Celler, and Hung T Nguyen. 2010. Dynamic modelling of heart rate response under different exercise intensity. *The open medical informatics journal* 4 (2010), 81.
[36] TI. 2019. *AWR1243Boost.* http://www.ti.com/tool/AWR1243BOOST
[37] TI. 2019. *FMCW principle.* http://www.ti.com/lit/wp/spyy007/spyy007.pdf
[38] TI. 2019. *The fundamentals of millimeter wave sensors.* http://www.ti.com/lit/wp/spyy005/spyy005.pdf
[39] Jianxuan Tu, Taesong Hwang, and Jenshan Lin. 2016. Respiration rate measurement under 1-D body motion using single continuous-wave Doppler radar vital sign detection system. *IEEE Transactions on Microwave Theory and Techniques* 64, 6 (2016), 1937–1946.
[40] Kim van Loon, Linda M Peelen, Emmy C van de Vlasakker, Cor J Kalkman, Leo van Wolfswinkel, and Bas van Zaane. 2018. Accuracy of remote continuous respiratory rate monitoring technologies intended for low care clinical settings: a prospective observational study. *Canadian Journal of Anesthesia/Journal canadien d'anesthésie* 65, 12 (2018), 1324–1332.
[41] Vernier. 2019. *Respiration Monitor Belt.* https://www.vernier.com/product/go-direct-respiration-belt/
[42] Fu-Kang Wang, Tzyy-Sheng Horng, Kang-Chun Peng, Je-Kuan Jau, Jian-Yu Li, and Cheng-Chung Chen. 2011. Single-antenna Doppler radars using self and mutual injection locking for vital sign detection with random body movement cancellation. *IEEE Transactions on Microwave Theory and Techniques* 59, 12 (2011), 3577–3587.
[43] Hao Wang, Daqing Zhang, Junyi Ma, Yasha Wang, Yuxiang Wang, Dan Wu, Tao Gu, and Bing Xie. 2016. Human respiration detection with commodity wifi devices: do user location and body orientation matter?. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. 25–36.
[44] S. Wiesner and Z. Yaniv. 2007. Monitoring Patient Respiration using a Single Optical Camera. In *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*.
[45] Zhicheng Yang, Parth H Pathak, Yunze Zeng, Xixi Liran, and Prasant Mohapatra. 2016. Monitoring vital signs using millimeter wave. In *Proceedings of the 17th ACM International Symposium on Mobile Ad Hoc Networking and Computing*. 211–220.
[46] Zhicheng Yang, Parth H Pathak, Yunze Zeng, Xixi Liran, and Prasant Mohapatra. 2017. Vital sign and sleep monitoring using millimeter wave. *ACM Transactions on Sensor Networks (TOSN)* 13, 2 (2017), 1–32.
[47] Saeid Yazdanirad, Habibollah Dehghan, Yaser Rahimi, Mohammad Zeinodini, and Mahnaz Shakeriyan. 2015. The relationship between overweight and heart rate in hot and very hot weather under controlled conditions. *Health Scope* 4, 4 (2015).
[48] Wenfeng Yin, Xiuzhu Yang, Lei Li, Lin Zhang, Nattapong Kitsuwan, and Eiji Oki. 2018. Hear: Approach for Heartbeat Monitoring with Body Movement Compensation by IR-UWB Radar. *Sensors* 18, 9 (2018), 3077.
[49] Shichao Yue, Hao He, Hao Wang, Hariharan Rahul, and Dina Katabi. 2018. Extracting multi-person respiration from entangled RF signals. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2, 2 (2018), 1–22.