Unveiling Contactless Sensing with LiDAR Mobility

Junying Hu[‡], Yongjian Fu[‡], Lili Chen[♯], Xinyi Li[◇], Xue Sun[†], Ju Ren[◇], Yaoxue Zhang[◇]
[‡]Central South University, China,
[♯]Northwestern Polytechnical University, China,
[†]National University of Singapore, Singapore,
[◇]Tsinghua University, China,
[‡]{junying.hu, fuyongjian}@csu.edu.cn, [♯]lilichen@nwpu.edu.cn, [†]snowing@nus.edu.sg,
[◇]{xinyili, renju, zhangyx}@tsinghua.edu.cn

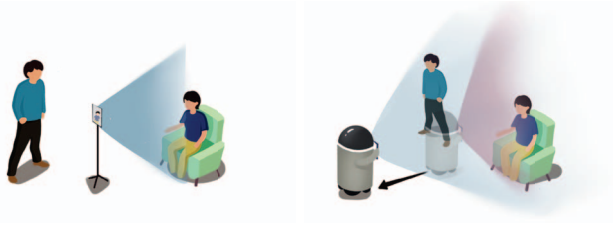
2024 IEEE Smart World Congress (SWC) | 979-8-3315-2086-1/24/$31.00 ©2024 IEEE | DOI: 10.1109/SWC62898.2024.00098

***Abstract*—This paper presents a novel contactless LiDAR-based sensing method that achieves high precision even under device motion. Traditional contactless systems are often compromised by movement interference, which restricts their operational scope and continuous sensing capability. To address this, we employ LiDAR to capture depth images and convert them into point clouds for analysis. This enables the detection of subtle movements by examining changes in point cloud contours. Firstly, we adapt depth images for input into RGB person segmentation models, facilitating accurate segmentation and precise point cloud data extraction for each target. Secondly, we utilize reflections from stationary parts of the target to estimate and compensate for device movement. Our findings reveal that the transition component in the result of Iterative Closest Point (ICP) encompasses both device and subtle movements of interest. Thus, we introduce a Masked Iterative Closest Point (MICP) method that minimizes the influence of the region of interest (RoI), allowing for precise extraction of device motion information. Finally, we extract RoI motion by analyzing the differences in transition components between ICP and MICP results. Experimental validation shows that our method maintains reliable respiration sensing with a relative error below 5%, even at speeds of up to 20 cm/s and distances up to 4 meters. This approach enhances the feasibility of LiDAR in dynamic sensing environments.**

***Index Terms*—wireless sensing, robot sensing systems, respiration sensing, LiDAR**

I. Introduction

Contactless sensing, an area garnering immense interest in various fields, revolutionizes how we interact with our environment. By emitting signals like light, sound, or electromagnetic waves, and analyzing their reflections, it discerns the characteristics of targeted objects without physical contact. This approach significantly differs from conventional methods that necessitate direct contact between the target and sensors, often leading to high costs, complexity, and discomfort. Contactless sensing, in contrast, offers a more economical, convenient, and user-friendly solution. Recent advancements in this field, such as trajectory tracking [1], posture recognition [2], and even fine-grained breathing [3] and heartbeat monitoring [4], demonstrating its profound potential and versatility in practical applications. While promising, most exiting systems presuppose stationary devices. This assumption overlooks numerous real-world scenarios where devices are in motion, as shown in Fig. 1. Examples abound, like handheld gadgets for monitoring vital signs, robots dynamically sensing their surroundings, or drones conduct surveillance and analysis of ground targets.



(a) Only one person in FoV. (b) Second person enters FoV as device moves.

Fig. 1. Motivating example. (a) Only one participant is within the Field of view (FoV) when its position is fixed. (b) Our method maintained continuous observation while positioning both individuals within the FoV.

Addressing this gap is crucial for the evolution and wider applicability of contactless sensing technology.

Note that when the device remains stationary, it accurately captures the target's characteristics or subtle movements. Yet, once in motion, its observations become a blend of the target's features and its own movement. This overlap causes interference in sensi ng the target, making traditional sensing technologies no longer work [5]. Therefore, it is crucial to eliminate the negative impact of device motion on target sensing.

Recent studies have looked into mitigating the effects of device motion on wireless-based contactless sensing. [5] and [6] were proposed for UWB- and acoustic-based sensing, respectively. They employ signals reflected from stationary objects in the environment, which are only affected by device motion, to isolate target motion from the mixed received signals. This method, however, is dependent on static environmental references and can be influenced by environmental changes. Another study [7] investigated eliminating device motion impact in LoRa based sensing. The method involves using two antennas on a robot, which provide similar signals when in the same position. As the robot moves, the second antenna can identify the target's movement by comparing its received signals to those previously received by the first antenna. This technique requires aligning two antennas in the moving direction of the device, making it unsuitable for devices with unpredictable movements. Additionally, due to the still limited spatial resolution of wireless signals, above systems assume the target to be a single point, which makes them susceptible to interference from movements of other

2993-396X/24/$31.00 ©2024 IEEE
DOI 10.1109/SWC62898.2024.00098

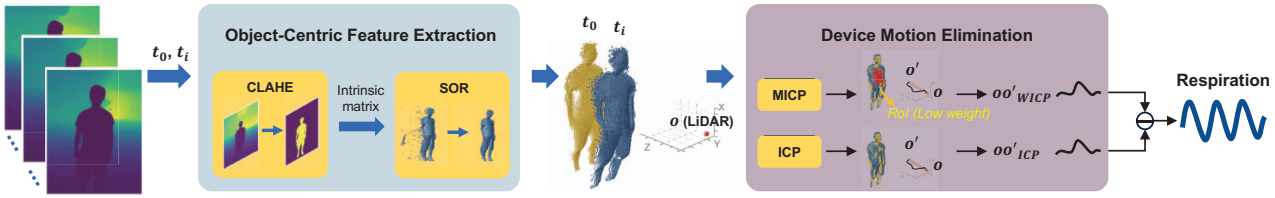


Fig. 2. System overview. Depth images and corresponding intrinsic matrices are used as inputs. We leverage Object-Centric Feature Extraction to create point clouds for each individual, from which respiratory signals are extracted using Device Motion Elimination.

uninterested parts or areas of the target.

In this paper, we aim to introduce an innovative sensing approach that supports device mobility. For the purpose, we choose to employ commercial LiDAR for contactless sensing. Given that LiDAR has been integrated into many mobile devices (like phones, robots, cars, and drones) [8] and offers high spatial resolution, it allows us to discern distance changes of different parts of the target from the observed time-of-flight (ToF) signals. This means we can differentiate various parts of the target, thereby utilizing the stationary parts to obtain the device's motion information. Furthermore, we use it to eliminate device motion from the mixed motion of interested target areas we aim to sense, which allows us to achieve accurate motion sensing for RoI. In this work, we take human respiration sensing as the example task to validate our idea. Specifically, we transform the depth image captured by LiDAR into a point cloud, thereby enabling the fitting of the target's contours in three-dimensional space. By analyzing changes in the point cloud data of the target's contours, we can detect subtle movements of RoI, particularly in the chest region of the target.

Turning our idea into a practical system involves overcoming several challenges. One of the primary issues is that the depth information captured by LiDAR includes both the environmental background and the target. To accurately detect individual targets in settings with multiple targets, it's crucial to separate the depth data specific to each target. Typically, this is achieved using RGB image for effective object segmentation, but this approach incurs extra costs and privacy concerns. Instead, we propose to use the depth image as input to the conventional RGB segmentation model for extraction of target contours. However, employing raw depth image directly for segmentation leads to blurred boundaries between the target and the background. To address this, we perform preprocessing on the depth image, clarifying boundary details and attaining segmentation precision on par with RGB-based methods.

Secondly, we convert the target's depth image into a point cloud [9] and use the ICP [10] registration method to rigidly rotate and translate it. This aligns the point cloud's overall contours at various moments. However, directly detecting subtle movements within a RoI is complex. The challenge stems from the rigid nature of ICP matching - small local changes in target contours can disrupt the process. Upon deeper analysis, we found that ICP matching produces a transformation matrix consisting of rotation and transition. The transition component encompasses both the device motion and subtle movements of the target. Consequently, we extract the subtle movements of the target from the acquired transition component, rather than directly observing the aligned point cloud. We develop a MICP registration method that assigns lower weights to a RoI on the target surface to minimize the impact of its motion, yielding a transition component that includes only device motion. By subtracting the transition component of MICP from that of standard ICP, we can accurately track the motion of a RoI.

In this paper, we make the following contributions. First, we addressed the issue of using depth images as inputs for traditional RGB segmentation models to segment the human body, and utilized the extracted human body mask to retrieve the target point cloud. Second, we analyzed the limitations of observing RoI contour changes directly from the target point cloud after ICP contour matching. Third, recognizing that the transition component in ICP results contains both device and RoI movement, we proposed a MICP approach to minimizes the influence of RoI on contour matching, facilitating the extraction of device motion. We then accurately determined the RoI movement by analyzing the transition differences between ICP and MICP results.

II. System Design

In this section, we take the example of respiration sensing to illustrate our newly designed method for negating the effects of device motion.

A. System Overview

The system flow is presented in Fig. 2. Firstly, we use LiDAR to capture the depth image frames, which include raw depth information from both the environment and the subjects of interest. Then, we use object-centric feature extraction to dispel the environment-related features and distill the object-specific ones (e.g., the human body). To improve image contrast and accurate segmentation, we employ contrast limited adaptive histogram equalization (CLAHE) [11] to overcome the challenge posed by the RGB segmentation models on depth data. Following feature extraction, the enhanced depth images are transformed into point cloud frames in conjunction with the intrinsic matrix. These frames are prepared for further analysis after minimizing background noise through statistical outlier removal (SOR) [12]. For device motion elimination, the system addresses the critical challenge of device motion which can corrupt the respiratory signals. We firstly align the point

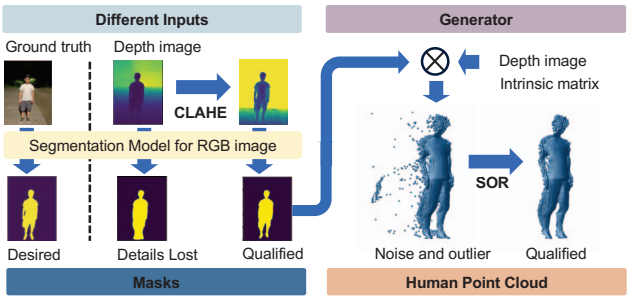
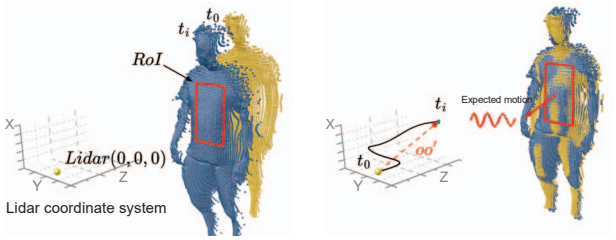


Fig. 3. Person segmentation and point cloud extraction.



(a) Point clouds in two frames. (b) Aligned point clouds.

Fig. 4. Example for two point clouds in different frames. (a) Point cloud before 3D contour matching; (b) Point cloud after 3D contour matching.

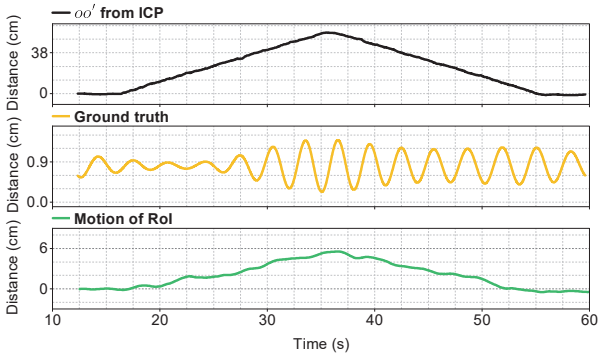


Fig. 5. Motion analysis of RoI following 3D contour matching. First, oo' from ICP represents device motion. The ground truth is the respiration signal obtained via UWB. RoI motion is the variation in the average distance of points within the RoI.

cloud frames at t_i and t_0 using ICP to obtain the transition oo', which encompasses both device and RoI motions. Following that, MICP is proposed to extract pure device motion, which assigns lower weights to points within the RoI to suppress the influence of RoI motion on 3D contour matching. The final output is the isolated respiratory signal, derived by comparing the outputs from the ICP and MICP methods.

B. Object-Centric Feature Extraction

The raw point cloud collected by LiDAR contains reflections from the environment and one or more interested objects (e.g., humans). To enable multi-object sensing independent of the environment, a series of steps is required to isolate features for each object from the raw data. A straightforward solution is to adopt object segmentation approaches inspired by computer vision [13]. However, the advanced segmentation models are typically designed for RGB images instead of depth images. As illustrated in Fig. 3, applying well-designed, off-the-shelf models to depth images results in distorted segmented object sub-images, making it challenging to obtain accurate object information. This limitation is primarily due to the significant depth gradient differences between objects and background. Specifically, different parts of an object (such as the top and bottom) may exhibit significant depth gradient variations, with the gradient at the top against the background being markedly higher than that at the bottom. The relatively minuscule depth gradient of the object's bottom concerning the global depth map blurs the bottom features, hindering convolutional operators designed for RGB images from accurately extracting object shape features.

One could recollect and annotate depth image data to address this challenge of retraining segmentation models. However, this method incurs considerable extra costs due to the need for extensive training data to eliminate the influence of different depth imaging device models, environments, and objects. Therefore, we preprocessed depth images to adapt them for RGB model input, as shown in Fig. 3. our key idea is to bridge the depth gradient gap between different parts of the object and the background by leveraging signal preprocessing, making depth images compatible with existing segmentation models. Through experiments, we found that the most significant factor affecting the segmentation results is the image contrast. Therefore, we employed CLAHE [11] to enhance image contrast, approximating the effect of using RGB images as input. Finally, we applied the obtained human mask to the depth image to acquire the depth values of various parts of the human body, thereby converting it into human point cloud data in conjunction with the intrinsic matrix [9]. However, the obtained human point cloud data still included some points from the background. We further processed the human points using the SOR [12] to achieve more accurate results.

C. Direct RoI Motion Sensing Is Ineffective

When the device is in motion, the position of the target in the depth image changes, making it impossible to set a unified RoI for all depth image frames. So we cannot consistently observe the depth changes at the same position of the target, and this error can easily overwhelm the subtle respiratory signals. By converting the depth map into point cloud data, we found that these issues arise because the position of the target in the LiDAR coordinate system continuously changes with the movement of the device as revealed in Fig. 4(a). Therefore, if we can perform contour matching on the point cloud data from different frames, we can not only unify the RoI but also resolve the issue of coupling between device movement information and depth values as revealed in Fig. 4(b).

The ICP method is a classical method used for 3D coutour matching. Specifically, considering two point cloud frames, P as the initial frame and Q as the target frame, the core idea of ICP is to use optimization to find a transformation function T(.) that minimizes the difference between T(P) and Q. For

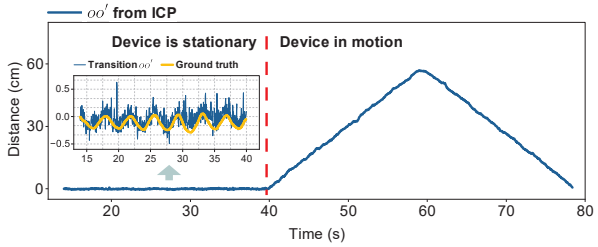


Fig. 6. Analysis of transition oo' during static and moving states of the LiDAR. When LiDAR is stationary, the trend of oo' perfectly match the respiration signal.

simplicity, we only consider T(.) as a rigid transformation, which is a reasonable assumption given the macroscopic relative motion between the device and the human body. Hence, T(.) can be obtained by a rotation matrix R and a transition matrix t, yielding $T(P) = R*P+t$. The estimation problem of the device's trajectory can be formulated as the following optimization problem.

$$E(T) = \sum_{(p,q)\in k} ((p - Tq) \cdot n_p)^2 \tag{1}$$

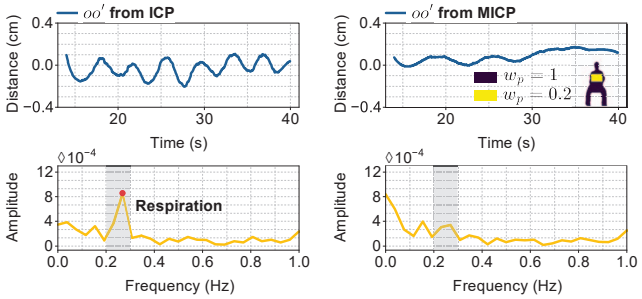
$$T = \begin{bmatrix} R & oo' \\ 0 & 1 \end{bmatrix} \tag{2}$$

In this context, p and q respectively represent points within P and Q. n_p is the normal of point p. T is the matrix including a 3-dimensional rotation matrix R and a transition value oo'. The ICP method used here is the point-to-plane ICP algorithm [14] that has a faster convergence speed than the point-to-point ICP algorithm.

After applying the ICP method, the target's contours in different frames are aligned, which means we can set a unified RoI to monitor the motion caused by exhalation and inhalation. *The expectation is that the motion of points within the RoI region aligns with the target's respiratory signal,* as shown in Fig. 4(b).

To this end, we conducted an experiment where an individual was seated facing the LiDAR mounted on a mobile robot, which moved the equipment forward and backward. We calculate the average distance of points within the RoI and use its variation to estimate the expected RoI motion. The results, depicted in Fig. 5, show that the waveform of the RoI motion signal significantly deviates from the ground truth. This discrepancy arises primarily because the sparsity of the LiDAR point cloud changes with distance (e.g., point separation varies from 5mm to 20mm at distances from 1m to 4m), preventing a consistent match across different frames. Additionally, the standard Iterative Closest Point (ICP) method, which only accounts for rigid body transformations, fails to capture subtle morphological changes in the chest and abdomen regions. Consequently, direct observation of RoI motion using matched point cloud data does not accurately reflect respiratory signals, rendering direct RoI motion sensing for respiratory monitoring unfeasible.

After further analysis of the experimental data, we observed that with the device stationary, changes in transition value oo'



(a) oo' from unmodified ICP (b) oo' from MICP

Fig. 7. Analysis of oo' when device is stationary. (a) oo' from unmodified ICP. Since the device remains stationary, the trend of change in oo' is consistent with the participant's breathing. (b) oo' from MICP with low weights for points in RoI. The respiratory component is suppressed; theoretically, oo' should always equal zero. Fluctuations in oo' mainly stem from involuntary minor movements by the participant and device distortion.

align with the respiratory signal, as shown in Fig. 6. This indicates that the subtle chest and abdominal movements during respiration influence the ICP matching outcomes, confirming that oo' encapsulates the respiratory signal. When the device is in motion, oo' captures both the device's motion and the respiratory signal of the target. Thus, a new concept was formulated: *By eliminating the device's motion from oo', an accurate respiratory signal can be accurately extracted.*

D. Device Motion Extraction

The reason oo' contains the respiratory signal is primarily because ICP considers the contribution of all points in the point cloud, including both the stationary points of various parts of the body and the points on the surface of chest and abdomen that move with respiration. Therefore, the obtained oo' includes not only the device's motion but also the changes in the chest and abdomen. If the points on the chest and abdomen can be removed or given smaller weights, their influence on the ICP can be minimized, resulting in oo' containing only the device's motion. To achieve this goal, we propose a method called MICP. Our intuition is that the minor movements of the RoI caused by breathing have a limited impact on the point cloud (for instance, breathing is confined to the chest area), and thus, the influence of breathing can be mitigated by reducing the weight of the local motion area during the point cloud registration process. Specifically, we added a parameter w_p in the ICP method to represent the weight of point p. The modified optimization problem is:

$$E(T) = \sum_{(p,q)\in k} ((p - Tq) \cdot n_p w_p)^2 \tag{3}$$

By assigning a lower weight to the points located in RoI, we can minimize the impact of their slight movements on the ICP, to get pure device motion. Here, we specify the RoI by manually selecting a mask. As all point cloud frames are aligned to the first frame, we apply a mask to the depth image at t_0 to delineate the RoI. Although the weights set to zero can maximize the discrepancy between MICP and ICP, it will introduce the error in MICP leading to the breath

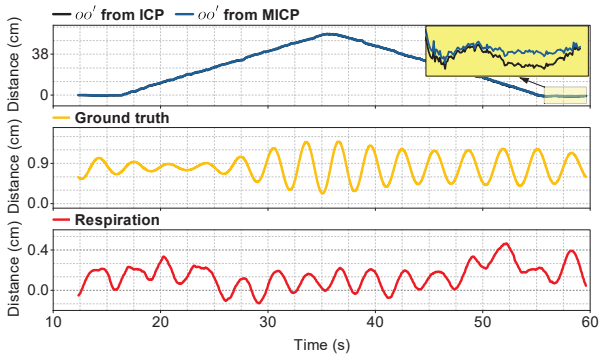


Fig. 8. Respiration extraction by device motion elimination. Due to the large-scale movement of the device, oo' from ICP and MICP are indistinguishable across the entire curve. However, distinct differences emerge in smaller segments. These differences, identified as the respiration signal (RoI motion), align perfectly with the ground truth in the bottom curve.

motion may be submerged. Therefore, we adjust the weights to balance maximizing differences while preserving the accuracy of MICP.

To assess its efficacy, we conducted an experiment where the LiDAR was positioned 2 meters in front of the participants. Using both ICP and MICP methods, we monitored oo' as shown in Fig. 7. While the LiDAR and participant were stationary, oo' captured respiratory components as shown in Fig. 7(a). On contrast, the MICP method to reduce point weights within RoI significantly suppressed these respiratory signals. Despite involuntary minor movements by the participant and device distortion causing fluctuations under 2mm, oo' from ICP were deemed to predominantly reflect device motion rather than participant respiration, as revealed in Fig. 7(b).

E. Device Motion Elimination

We hypothesize that oo' derived via MICP is primarily influenced by the stationary points, thus predominantly reflecting device motion. This value at time t is denoted as $oo'_{MICP}(Q_t, P_{static})$, where Q_t represents the point cloud at time t and P_{static} is the static subset of the initial point cloud at t_0 outside the RoI. In contrast, the standard ICP method incorporates both dynamic and static points, affecting the 3D contour matching. Consequently, oo' via ICP is influenced by the entire target point cloud, capturing both device and RoI motions. This is represented as $oo'_{ICP}(Q_t, P_{static} \cup P_{dynamic})$, where $P_{dynamic}$ includes points within the RoI at t_0.

Given that oo' is a three-dimensional vector describing the trajectory in the LiDAR coordinate system, it is essential to project this vector in a specific direction for analysis. Considering the Z-axis aligns with the LiDAR's directional orientation, we project oo' onto $\vec{Z}$ to isolate the respiratory signal. Then the breathing magnitude can be expressed as:

$$y(t) = \left| oo'_{ICP}(Q_t, P_{static} \cup P_{dynamic}) \cdot \vec{Z} \right| - \left| oo'_{MICP}(Q_t, P_{static}) \cdot \vec{Z} \right| \quad (4)$$

We set up an experiment to demonstrate device motion elimination, as configured in Fig. 5. The result is shown in Fig. 8. The figure initially presents the oo' values derived from ICP and MICP, highlighting their discrepancies. The difference between these values was used as the respiratory signal, which perfectly matched the ground truth.

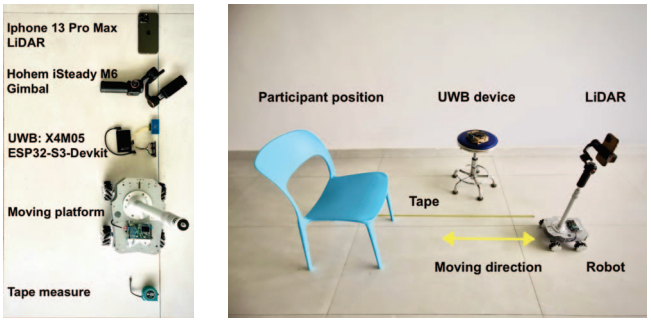


Fig. 9. Experimental hardware (left) and scene (right).

III. Evaluation

Section 3.A details our experimental setup, encompassing LiDAR, ground truth, mobile platform, signal processing, and scene parameters. In Section 3.B, we demonstrate the comprehensive performance of our method, validating its effectiveness in human respiratory sensing under typical conditions. Sections 3.C-D assess the system's robustness to variations in speed, distance, and the multiple targets sensing. Section 3.E further confirms the necessity of the depth image enhancement outlined in Section 2.A.

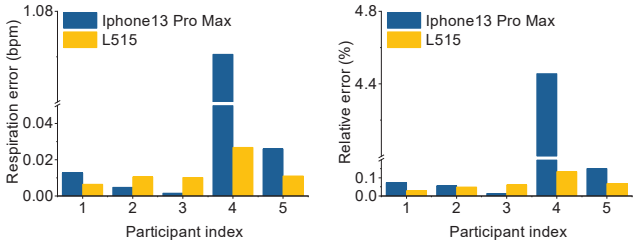
A. Evaluation Setup

LiDAR. We utilized the LiDAR sensor on the iPhone 13 Pro Max and the Intel RealSense LiDAR Camera L515 [15] to capture depth images and intrinsic matrices. These devices deliver depth images at a frequency of 30 Hz, with resolutions of 256×192 for the iPhone and 320×240 for the L515.

Ground truth. In our experiments, we employed a contactless UWB-based method for respiratory monitoring. Compared to contact-based methods requiring subjects to wear devices on their thorax or abdomen, UWB-based monitoring allows subjects to remain relaxed and avoids interfering with the LiDAR sensing. We utilized an ESP32-S3-DevKitC-1 [16] development board connected to an X4M05 [17], positioned statically in front of the subject, to ensure that the UWB data could extract subjects' respiration signals.

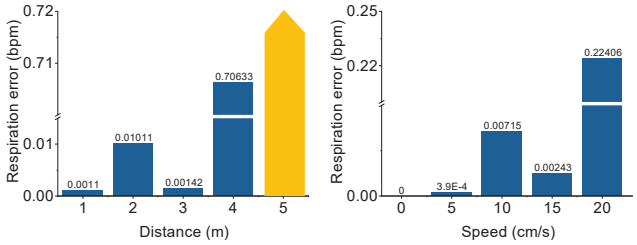
Mobile platform. We mounted the LiDAR on a four-wheeled robot and developed an application to control its movement via Bluetooth. To maintain the stability of the LiDAR and ensure that the subject remained within its field of view at all times, we connected the robot and the LiDAR using a Hohem iSteady M6 [18] gimbal, which features target tracking capabilities.

Signal processing. We transmitted the respiratory signals captured by the X4M05 to the iPhone 13 Pro Max via WiFi, aligning them in time with the LiDAR data. Subsequently, the iPhone 13 Pro Max forwarded both the UWB and LiDAR data to a MacBook for processing. In the experiment involving the L515, we directly captured depth information from the L515



(a) Respiration error (b) Relative error

Fig. 10. Overall performance. (a) Respiration rate error, (b) Relative error that is the percentage of the respiration error relative to the respiration rate.



(a) Impact of distance (b) Impact of device moving speed

Fig. 11. Impact of different factors.

to a laptop and aligned the depth frames with UWB frames on the laptop. We employed the macOS API *VNGeneratePersonInstanceMaskIRequest* for person segmentation, used the *TransformationEstimationPointToPlane* method from the Open3D library [19] for point cloud matching, and modified this method to implement the MICP algorithm described in this paper. Finally, we compared the respiratory signals derived from the LiDAR data with the ground truth provided by the UWB device to validate the effectiveness of our methods.

Scene and Parameters. All experiments were conducted in our laboratory with five participants (three males and two females). Participants were seated facing the LiDAR, supported by a chair with a backrest to minimize motion artifacts from involuntary movements during respiration. The UWB device was positioned laterally between the LiDAR and the participant, facing the subject to avoid the LiDAR's field of view and optimize signal reception from the participant. The hardware and the placement of them is shown in Fig. 9. The comprehensive performance evaluation involved collecting 5 minutes of data from each participant, while other experiments collected 1 minute of data. The experimental environment remains consistent when verifying the effects of specific parameters. In the MICP algorithm, the points within the RoI region were assigned weights of 0.2, with other points weighted at 1.

B. Overall Performances

Our system demonstrated accurate respiratory detection under typical conditions using both the LiDAR of the iPhone 13 Pro Max and the L515, operating at a speed of 20 cm/s. Fig. 10(a) illustrates the respiratory rate error values across participants. All but the fourth participant exhibited minimal error rates. Due to the extremely low respiration error (< 0.02 bpm), there is no discernible trend in performance between the two LiDAR devices under typical conditions. The fourth participant's higher error rate was attributed to lower breathing amplitude, making it challenging to distinguish between some peak and trough values. Conversely, the L515 exhibited better stability with shallow breathing. We also assessed the relative respiratory error rates for each participant, all of which remained below 5%, as shown in Fig. 10(b). These results affirm that our method can reliably monitor respiratory rates at common indoor robot speeds (e.g., a robot vacuum moving at 20 cm/s), proving its practical utility.

C. Parameters Evaluation

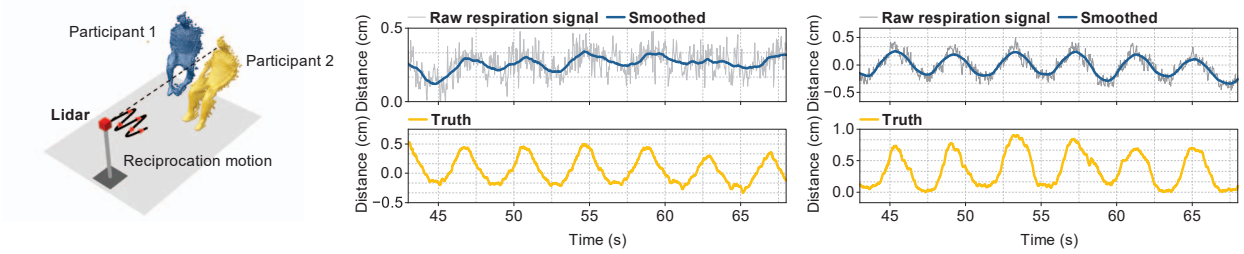
Distance. We investigated the effect of distance between the device and the participant on respiratory detection using iPhone 13 Pro Max. Distance was measured from the target's back to the LiDAR, increasing in 1-meter increments from 0 to 5 meters. As depicted in Fig. 11(a), error rates remained extremely low within a 3-meter distance and showed no correlation with distance, indicating flawless sensing performance. The minor fluctuations observed are likely attributable to environmental variables. At 4 meters, the error rate slightly increased to 0.7063bpm but was still within acceptable limits. Beyond 4 meters, at 5 meters, the method could not provide reliable error rates due to significant data distortion from the LiDAR on the iPhone 13 Pro Max, impairing effective respiratory detection.

Speed. We investigated the effects of device movement speed (0, 5, 10, 15, and 20 cm/s) on respiratory detection accuracy using iPhone 13 Pro Max. As depicted in Fig. 11(b), at a device movement speed of 0 cm/s, the error rate was zero, indicating perfect alignment between the measured respiratory signals and the ground truth, including breath amplitude. At speeds of 5, 10, and 15 cm/s, the error rates were consistently low with no apparent correlation to movement speed. At 20 cm/s, the error rate rose to 0.22406, yet remained within acceptable limits.

This study demonstrated the robustness and practical applicability of our method by investigating the effects of device-to-participant distance and device speed on detection performance.

D. Multiple Targets

We investigated the potential of our method for target perception in a multi-person setting using iPhone 13 Pro Max. As illustrated in Fig. 12(a), the experimental setup involved two participants facing the LiDAR from a distance of 2.5 meters. Participant 1 was seated sideways, while Participant 2 faced directly towards the LiDAR. The device, mounted on a mobile robotic platform, cyclically moved back and forth at 20 cm/s over a distance of 50 cm. Fig. 12(b, c) shows that the respiratory outcomes for both participants aligned with the ground truth. Participant 2 exhibited clearer respiratory signals compared to Participant 1, whose breathing amplitude was lower and some peaks were nearly indiscernible. This was due to Participant 1's sideways posture, which reduced the



(a) Scene (b) Participant 1 (c) Participant 2

Fig. 12. Sensing results for multiple targets.

amplitude of chest and abdominal movements when projected onto the LiDAR's z-axis. This experiment confirmed that our method accurately detects individual breathing patterns in a multi-person environment.

E. Depth Image Enhancement

To assess the efficacy and necessity of depth image enhancement in our methodology, we analyzed single-person scenarios at varying distances using iPhone 13 Pro Max. Fig. 13(a) illustrates the participant mask generated at a 4-meter distance using an RGB image via the macOS API *VNGeneratePersonInstanceMaskIRequest*, serving as ground truth. This baseline facilitated a comparative analysis of results using raw depth images versus CLAHE-enhanced depth images, depicted in Fig. 13(b) and Fig. 13(c) respectively. The comparison clearly shows that the CLAHE-enhanced depth images align more closely with the ground truth, while the raw depth images incorporate excessive environmental noise and lack detail.

For a quantitative evaluation, we computed metrics across distances from 1 to 4 meters. The *Details Lost* metric quantifies the percentage of missing detail pixels relative to the ground truth, and the *Environment* metric measures the excess environmental pixels as a percentage of the ground truth. The *Rel. error <5%* metric assesses whether the relative error in respiratory detection remains below 5%. Table 1 demonstrates that using raw depth images results in greater detail loss and increased environmental interference as distance extends, leading to a relative respiratory detection error rate exceeding 5% at distances of 3-4 meters. In terms of Details Lost and Environment indicators, the performance is not decrease linearly. The reason is that different background and objects in the environment may appear in FoV at different distances, which may slightly affect the performance of image enhancement.

TABLE I
ANALYSIS RESULTS FOR DEPTH AND ENHANCED DEPTH IMAGES AT VARIOUS DISTANCES

Distance	Details Lost (%)		Environment (%)		Rel. error < 5%	
	Origin	Enhanced	Origin	Enhanced	Origin	Enhanced
1	3.285	2.219	3.115	2.034	✓	✓
2	7.684	13.739	8.530	5.281	✓	✓
3	18.878	11.548	19.296	7.940	✗	✓
4	21.653	13.071	90.315	5.682	✗	✓

(a) Ground truth (b) Unmatched (c) Qualified

Fig. 13. Person segmentation at distance of 4m using (a) RGB (as ground truth), (b) Depth, and (c) Depth-Clahe as inputs respectively.

IV. Related Work

Wireless-based respiration sensing. A variety of wireless signals, such as Wi-Fi, sound, and ultra-wideband (UWB), are being utilized for respiration sensing. Sound signals, characterized by their slower propagation speed, can distinguish reflections that are very close to each other (i.e., less than 4.25cm) even under a narrow bandwidth of 4KHz. However, the inherent short-distance propagation characteristic limits their sensing range [20]. Contrastingly, WiFi signals are distinguished by their longer propagation range, which facilitates the detection of respiration across expansive indoor spaces. However, the relatively narrow bandwidth of WiFi signals constrains their effectiveness, particularly for multi-target sensing at close proximities [21]. In recent years, the advent of more advanced UWB signals, now equipped in many smartphones like iPhones and Samsung models, benefit from their large bandwidth, making them suitable for multi-target sensing [22].

Camera-based respiration sensing. Researchers have investigated the use of visual signals for respiration sensing, which enables direct observation of the human body's entire contour [23], [24]. These methods have implemented automatic RoI detection and noise suppression techniques, enabling direct observation of movements in the chest or abdominal area. However, their sensing accuracy can be significantly affected by the involuntary body movements. Recent attempts address motion artifacts by employing multiple cameras [25]. The principle behind this is to observe the changes in the front and back contours of the body, thereby deducing the movements of the chest and abdomen.

V. Discussion and Future Work

Refinement of RoI Selection. The selection of the RoI is closely linked to the characteristics of the clothing, as LiDAR-based respiration monitoring relies on detecting subtle changes in the contours of the fabric during breathing. Loose-fitting clothing presents a challenge since only certain parts of the fabric move consistently with chest movements. In this paper, we manually selected a region encompassing the chest and abdomen to maximize coverage. For future work, we will develop an automatic RoI selection method to enhance the robustness of LiDAR-based respiration monitoring.

Human orientations. The amplitude of the respiration curve may be understated in this study, as it only considers projecting the transitional component onto the Z-axis of the LiDAR coordinate system. This approach potentially introduces errors by overlooking minor chest movements. Future work will explore strategies to mitigate this issue, including the identification of an optimal mapping vector.

Multiple targets sensing. When multiple targets are present, this paper extracts the depth information of each target and derives respiration signals similarly to a single target scenario. The performance largely hinges on the human segmentation model. Segmentation accuracy is influenced by the LiDAR's Field of View (FoV) and the number of targets, which are beyond the scope of this paper. Future work will investigate lightweight segmentation models for LiDAR to optimize multi-target sensing performance.

VI. Conclusion

In this paper, we have introduced a novel contactless sensing approach that utilizes LiDAR mobility to enable precise sensing of RoI while device is in motion. Our method addresses the critical challenge of device motion interference by leveraging reflections from stationary parts of the target to estimate and compensate for device motion. Taking the example of respiration sensing, we enable traditional RGB segmentation models to achieve accurate person segmentation when using depth images as input. Through experiments, we have shown that our proposed MICP registration method can effectively isolate target motion information, enabling accurate sensing of human respiration even when the sensing device is in motion.

Acknowledgment

This research was supported in part by the National Natural Science Foundation of China under Grant No. 62341201, 62122095, 62432004, and by a grant from the Guoqiang Institute, Tsinghua University. Ju Ren is the corresponding author.

References

[1] Yaxiong Xie, Jie Xiong, Mo Li, and Kyle Jamieson. md-track: Leveraging multi-dimensionality for passive indoor wi-fi tracking. In *The 25th Annual International Conference on Mobile Computing and Networking*, pages 1–16, 2019.

[2] Wenjun Jiang, Hongfei Xue, Chenglin Miao, Shiyang Wang, Sen Lin, Chong Tian, Srinivasan Murali, Haochen Hu, Zhi Sun, and Lu Su. Towards 3d human pose construction using wifi. In *Proceedings of the 26th Annual International Conference on Mobile Computing and Networking*, pages 1–14, 2020.

[3] Yuqing Yin, Xu Yang, Jie Xiong, Sunghoon Ivan Lee, Pengpeng Chen, and Qiang Niu. Ubiquitous smartphone-based respiration sensing with wi-fi signal. *IEEE Internet of Things Journal*, 9(2):1479–1490, 2021.

[4] Fusang Zhang, Zhi Wang, Beihong Jin, Jie Xiong, and Daqing Zhang. Your smart speaker can" hear" your heartbeat! *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 4(4):1–24, 2020.

[5] Fusang Zhang, Jie Xiong, Zhaoxin Chang, Junqi Ma, and Daqing Zhang. Mobi2sense: empowering wireless sensing with mobility. In *Proceedings of the 28th Annual International Conference on Mobile Computing And Networking*, pages 268–281, 2022.

[6] Jialin Liu, Dong Li, Lei Wang, Fusang Zhang, and Jie Xiong. Enabling contact-free acoustic sensing under device motion. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 6(3):1–27, 2022.

[7] Binbin Xie, Deepak Ganesan, and Jie Xiong. Embracing lora sensing with device mobility. In *Proceedings of the 20th ACM Conference on Embedded Networked Sensor Systems*, pages 349–361, 2022.

[8] Robert Bogue. The growing importance of lidar technology. *Industrial Robot: the international journal of robotics research and application*, 49(6):1025–1031, 2022.

[9] Haowen Deng, Tolga Birdal, and Slobodan Ilic. Ppfnet: Global context aware local features for robust 3d point matching. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 195–205, 2018.

[10] Zhengyou Zhang. Iterative closest point (icp). In *Computer vision: a reference guide*, pages 718–720. Springer, 2021.

[11] Karel Zuiderveld. *Contrast limited adaptive histogram equalization*, page 474–485. Academic Press Professional, Inc., USA, 1994.

[12] Steven Walfish. A review of statistical outlier methods. *Pharmaceutical technology*, 30(11):82, 2006.

[13] Apple Inc. Segmenting and colorizing individuals from a surrounding scene, 2023. Accessed: 2024-08-12.

[14] Yang Chen and Gérard Medioni. Object modelling by registration of multiple range images. *Image and vision computing*, 10(3):145–155, 1992.

[15] Intel Corporation. Introducing the world's smallest high-resolution lidar depth camera. Accessed: 2024-10-24.

[16] Espressif Systems. *ESP32-S3 User Guide – DevKitC-1*, 2024. Accessed: 2024-08-14.

[17] Novelda AS. X4 - datasheet, 2021. Accessed: 2024-08-14.

[18] Hohem Technology Co., Ltd. isteady m6. Product Page, 2024. Accessed: 2024-08-14.

[19] Qian-Yi Zhou, Jaesik Park, and Vladlen Koltun. Open3D: A modern library for 3D data processing. *arXiv:1801.09847*, 2018.

[20] Shirui Cao, Dong Li, Sunghoon Ivan Lee, and Jie Xiong. Powerphone: Unleashing the acoustic sensing capability of smartphones. In *Proceedings of the 29th Annual International Conference on Mobile Computing and Networking*, pages 1–16, 2023.

[21] Youwei Zeng, Dan Wu, Jie Xiong, Jinyi Liu, Zhaopeng Liu, and Daqing Zhang. Multisense: Enabling multi-person respiration sensing with commodity wifi. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 4(3):1–29, 2020.

[22] Fusang Zhang, Zhaoxin Chang, Jie Xiong, Junqi Ma, Jiazhi Ni, Wenbo Zhang, Beihong Jin, and Daqing Zhang. Embracing consumer-level uwb-equipped devices for fine-grained wireless sensing. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 6(4):1–27, 2023.

[23] Mona Alnaggar, Ali I Siam, Mohamed Handosa, Tamer Medhat, and MZ Rashad. Video-based real-time monitoring for heart rate and respiration rate. *Expert Systems with Applications*, 225:120135, 2023.

[24] Shane Transue, Phuc Nguyen, Tam Vu, and Min-Hyung Choi. Real-time tidal volume estimation using iso-surface reconstruction. In *2016 IEEE First International Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE)*, pages 209–218. IEEE, 2016.

[25] Felix Wichum, Christian Wiede, and Karsten Seidl. Depth-based measurement of respiratory volumes: A review. *Sensors*, 22(24):9680, 2022.