Fusing UWB and Depth Sensors for Passive and Context-Aware Vital Signs Monitoring

Zongxing Xie
Stony Brook University
zongxing.xie@stonybrook.edu

Bing Zhou
IBM Research
bing.zhou@ibm.com

Xi Cheng
Stony Brook University
cheng.xi@stonybrook.edu

Elinor Schoenfeld
Stony Brook University
elinor.schoenfeld@stonybrook.edu

Fan Ye
Stony Brook University
fan.ye@stonybrook.edu

***Abstract*—This demonstration presents a working prototype of *VitalHub*, a practical solution for longitudinal in-home vital signs monitoring. To balance the trade-offs between the challenges related to an individual's efforts thus compliance, and robustness with vital signs monitoring, we introduce a passive monitoring solution, which is free of any on-body device or cooperative efforts from the user. By fusing the inputs from a pair of co-located UWB and depth sensors, *VitalHub* achieves robust, passive, context-aware and privacy-preserving sensing. We use a COTS UWB sensor to detect chest wall displacement due to the respiration and heartbeat for vital signs extraction. We use the depth information from Microsoft Kinect to detect and locate the users in the field of view and recognize the activities of the respective users for further analysis. We have tested the prototype extensively in engineering and medical lab environments. We will demonstrate the features and performance of *VitalHub* using real-world data in comparison with an FDA approved medical device.**

***Index Terms*—Vital signs monitoring, non-touch sensing, longitudinal in-home data collection, aging, health**

I. Introduction

Longitudinal vital signs data (such as respiratory rate and heart rate) collected in the home environment are invaluable as they play an important role in health status assessment for either self or clinical care. Continuous vital signs monitoring helps detect changes in an individual's health (e.g., onset of flu or COVID) and track changes in chronic diseases. It provides insights for clinical management, and is integral to facilitating aging in place.

Recently, Photoplethysmography (PPG) and electrocardiograms (ECG) become available on wearables (e.g., Apple Watch, Fitbit) and mobile devices (e.g., KardiaMobile), making vital signs monitoring more accessible in our daily life. However, there is a burden on the individual user who most remember to frequently charge, wear the wearable and interact with mobile Apps to obtain data. In populations who may benefit the most from having these data collected (e.g., older adults, individuals with dementia), manual dexterity and cognitive challenges will hinder wear and charging.

This work is supported in part by NSF grants 1652276, 1951880, 2028952, 2119299. Dr. Bing Zhou contributed to this work when he was with Stony Brook University.

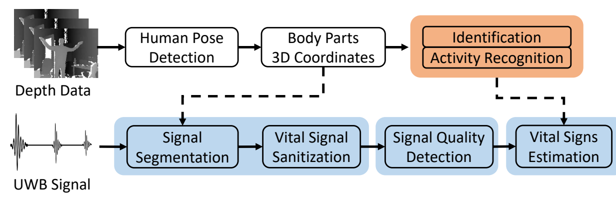

Fig. 1. The overall framework of *VitalHub*. It consists of two parts, the vital signs monitoring module (highlighted in blue) and the context annotation module (highlighted in orange).

Radio frequency (RF) sensors make it possible for non-contact monitoring, which mitigates the challenges to maintaining compliance for continuous monitoring. Recent radio-based designs, such as Wi-Fi, millimeter-wave (mmWave), and ultra-wideband (UWB) solutions demonstrate the feasibility of vital signs extraction by sensing the chest wall displacement due to physiological movements (e.g., respiration and heartbeat). When the radio sensor repeatedly emits radio signal towards the chest wall, it will receive the echoes that bounce off the chest wall continuously. Therefore, the travel distance of the radio signal changes periodically according to the chest wall displacement due to the respiration and heartbeat, and it allows us to extract respiratory rate and heart rate from the periodicity of the changes in the time of flight (ToF) of the radio signal. However, there are three sources of challenges for radio-based vital signs monitoring. First, the sampling time offset, due to the imperfect synchronization between the transmitter and the receiver, will introduce noise to the vital signal. Second, random body movement will introduce large disturbance overwhelming the vital signals. Third, the non-linearity of the channel will introduce the intermodulation components, making it difficult to detect the true vital signs.

To address the aforementioned challenges, we have developed *VitalHub* [1], a robust vital signs monitoring system using a pair of co-located UWB and depth sensors. To meet the design principles for in-home monitoring, *VitalHub* is developed to be robust, passive, context-aware and privacy-preserving, by fusing the information from the UWB and depth sensors. *VitalHub* automatically monitors the vital signs of the users, who neither wear any device nor perform any cooperative effort to collect these data. In this demonstration we show the non-touch vital signs estimation capability of *VitalHub*, which indicates a suitable solution for longitudinal in-home monitoring.

II. Design & Implementation of VitalHub

An overview of *VitalHub* [1] is shown in Figure 1. It works by detecting and segmenting the RF signal reflected from the individual to extract vital signs. Meanwhile, it recognizes the physical activities of the user to annotate the estimated vital signs for further context-aware analysis.

A. System Design

The *VitalHub* system has two major modules: vital sign monitoring and context annotation, obtained by a pair of co-located UWB and depth sensors.

The vital sign monitoring module takes as input the UWB pulse echoes bounced off the objects at different distances, and uses the distance estimated by the depth sensor to select the proper segments of the UWB signal for vital signs extraction. We have built a signal pre-processing pipeline to denoise and extract the vital signal (i.e., the phase change in UWB signal due to the respiration and heartbeat). We have developed a signal quality detector [2] to detect and exclude signals disturbed by random body movement. To achieve robust vital sign monitoring, we have designed a probabilistic weighted framework (PWF) to specifically deal with the challenges from the non-linearity of the channel in heart rate estimation.

The context annotation module takes depth images as input in which pixels indicate ToF, thus distances (or depths) of objects to the depth sensor. We leverage a human pose recognition model from Microsoft Kinect to detect the body parts. Thus the individual's pose is present in the monitoring zone. This model outputs 3D positions of body joints as a representation of skeletal pose. We use the predicted position of torso to help locate the segments of UWB signals corresponding to the chest wall. Moreover, we leverage spatial and temporal features of the consecutive skeletal poses to further recognize the context (e.g., identities and activity).

B. Implementation

The *VitalHub* system consists of three hardware components. First, a COTS IR-UWB System on a Chip (SoC) from Novelda works as the radio frequency (RF) front end. It repeatedly emits UWB pulses within frequency band 7.25-10.2 GHz, and samples the received signal at the frequency of 23.328 Ghz. The UWB SoC is configured to output frames at an update rate of 10 frame-per-second (fps), and each frame includes the echo pulses reflected from the objects within the range of 10 m. Second, the Microsoft Kinect for Windows v2 Sensor, with the RGB camera covered to preserve privacy, serves as the depth sensor. The Kinect SDK is used to detect and locate human bodies present in the field of view at 30 fps. The depth sensor is co-located with the UWB sensor, and their coordinate system is aligned. Both modalities stream data to the same backend PC via serial port. Third, we process the data from the UWB and depth sensors on a backend PC, configured with an Intel Core i7-8750 2.2GHz CPU, 16GB RAM and NVIDIA RTX 2060 GPU. We implemented all algorithms of *VitalHub* in Python.

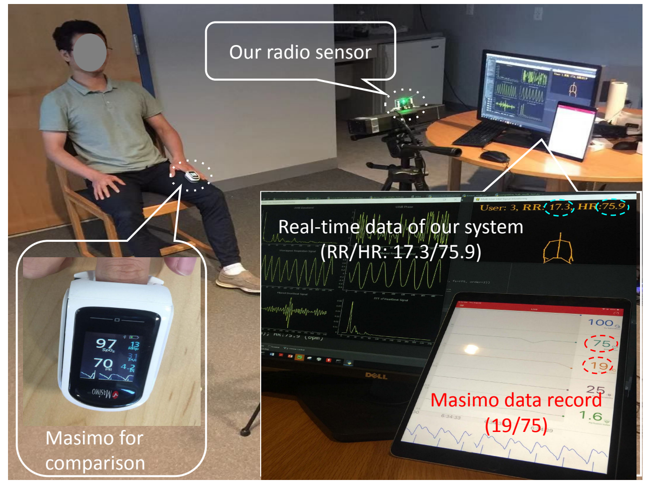

Fig. 2. A typical setup for non-contact vital sign monitoring. The subject sits on the chair at an arbitrary distance from the radio sensor. The pulse oximeter is used to provide the ground truth for comparison.

III. Demo Description

We demonstrate the features of *VitalHub*, which support passive, context-aware, non-touch vital signs monitoring. As shown in Figure 2, the prototype of our system for demonstration is composed of a pair of co-located UWB and depth sensors. When the user walks into the monitoring zone, our system detects the presence of the user and starts to track the movement of the user. When the user stays quasi-stationary at an arbitrary location, the system outputs the respiratory and heart rates of the user once the signal quality detector captures the available signals for valid vital signs estimation.

References

[1] Z. Xie, B. Zhou, X. Cheng, E. Schoenfeld, and F. Ye, "Vitalhub: Robust, non-touch multi-user vital signs monitoring using depth camera-aided uwb," in *IEEE International Conference on Healthcare Informatics*, 2021.

[2] Z. Xie, B. Zhou, and F. Ye, "Signal quality detection towards practical non-touch vital sign monitoring," in *Proceedings of the 12th ACM Conference on Bioinformatics, Computational Biology, and Health Informatics*, 2021.