RADIO-FREQUENCY SENSOR APPLICATIONS FOR ASSISTED LIVING

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Pragya Sharma

December 2020

RADIO-FREQUENCY SENSOR APPLICATIONS FOR ASSISTED LIVING

Pragya Sharma, Ph. D.

Cornell University 2020

The demographic and socio-economic trends have signaled an overwhelming need for improved assisted living facilities with non-intrusive technologies that address prominent user needs of *wellness* by continuous vital-sign sensing and *safety* and *accessibility* by occupant-detection. This home-centric approach can greatly improve quality-of-life for the elderly suffering from diseases like sleep apnea and chronic obstructive pulmonary disorder (COPD) by long-term continuous monitoring. Moreover, with the ability to track occupant presence, location, and posture, the response time can be shorter in case of emergencies like fall-events.

To solve the problem of *wellness* monitoring, this work focuses on using an over-clothing wearable radio-frequency (RF) sensor to accurately measure heartbeat and respiration. While ambient and wearable chest-motion monitoring sensors exist, they only measure the weak surface motion. This work uses near-field coherent sensing to strongly couple to internal dielectric boundary motion. As the choice of antenna placement and sensor setup has a significant impact on the user comfort level and signal quality, different sensor design variations are considered. Both thorax and abdominal respiratory patterns are measured independently to monitor paradoxical respiration during obstructive apnea. A calibration-based approach is used to estimate respiratory volume, a key parameter to estimate respiratory effort, and results are presented for

different simulated breathing disorders. Finally, results are validated under posture and gender variations in a study conducted on 20 participants, showing a high correlation of presented sensing setup with the reference devices: heart rate ($r = 0.95$), respiratory rate ($r = 0.93$), and respiratory volume ($r = 0.84$), along with high accuracies of 96% and 83% for simulated central and obstructive sleep apnea detection respectively. The work has been extended for attention vs relaxation state classification resulting in a satisfactory accuracy of 93% on the unseen test subjects.

While vital sign monitoring is an active task, which has strong advantages by a wearable or on-the-furniture sensor approach, *safety* and *accessibility* need to be embedded in the ambient, where it can monitor occupant presence to trigger responses even if the occupant is device-free. As ambient sensors need to have sufficient observation coverage in arbitrary room layout and furnishing, ambient passive radio frequency identification (RFID) tags are chosen to provide scalability cost-effectively and reliably. However, the indoor RF signal suffers from heavy multipath and unknown phase offsets due to cables and reader transceiver circuitry. In this work, a novel background calibration algorithm is presented that works well under significant background clutter. A linearized inverse model is used to generate an 'RF image' based on the occupant reflectivity and absorption, using novel sparsity approximation algorithms, from which occupant count and location can be extracted. As the number of occupants increases, the approximations fail and counting becomes challenging. In this scenario, a deep-learning-based solution is presented that can count with good accuracy of 82%, with learning transference to different rooms and setup variations, significantly reducing the cost overhead from training data collection.

BIOGRAPHICAL SKETCH

Pragya Sharma was born in Udaipur, a beautiful historic lake city nestled in the Aravalli mountain range in Northwestern India. She attended St. Gregorios school and was the Indian Prime Minister's Guest, at the Republic Day parade in 2009 for excellent all-India performance in 10^{th} grade. She then went on to pursue a bachelor's in technology degree in electrical engineering from the Indian Institute of Technology (IIT) Kharagpur, India, and received the best signal processing project award. Graduating in 2015, she enrolled in the PhD program in the School of Electrical and Computer Engineering at Cornell University, Ithaca, NY. In the summer of 2017, she was a research intern at Maxim Integrated, Dallas, TX. Her research interests include RF systems, sensor applications in assisted living solutions, biomedical sensors, signal processing, machine learning, and algorithms. Aside from engineering and technology, she enjoys drawing, painting, lettering, and various crafts.

This document is dedicated to my parents Santosh and Jagdish C. Sharma, my family, and Mayank.

ACKNOWLEDGMENTS

This thesis and research work has been a result of support and encouragement from many wonderful people.

First and foremost, I would like to express my deepest gratitude to my Ph.D. advisor, Prof. Edwin C. Kan, for his excellent guidance and incessant support throughout my PhD journey. His patience and constant belief in me have been a continuous source of motivation and encouragement. I would also like to thank my PhD committee, Profs. Dave Hysell and Chris Studer. Prof. Hysell has been a great mentor through this process. He inspired me to dive into the area of RF imaging, by coming up with ideas, brainstorming, and great suggestions, which directed my Ph.D. research in different interesting dimensions. Prof. Studer's comments have provided a new perspective to the problems I have taken to him, and his class on wireless communication was my first introduction to this field, which has helped strengthen the foundation in the area.

I would also like to thank all the collaborators, specifically Dr. Ana Krieger of the Cornell Center for Sleep Medicine and Prof. Huiju Park of Fiber Science and Apparel Design. Along with her constant support, Dr. Krieger's insightful consultations on respiratory characteristics have been very helpful, and together with her passion to help the patients have inspired me to work in the area of biomedical research. Prof. Park and his group's work on occupant human study data collection has been very helpful in my research.

I would like to express my gratitude to the past and current group members and officemates. I am grateful to Xiaonan Hui for his very helpful comments, suggestions, and experimental help. I am thankful to Guoyi Xu for his extensive experimental work and discussions on RF imaging, Jianlin Zhou for his invaluable conversations on electromagnetics and biomedical research, and Thomas Conroy for his help in the

attention monitoring project. I would also like to thank Yingqiu Cao, Yunye Gong, Zijing Zhang, Yunfei Ma, Phil Gordon, and Nan Xu for their support and friendship.

My thanks also go to friends outside the Kan research group, who have always been there for me throughout this journey, Shraddha Vartak, Pragya Shah, Neeraj Kulkarni, Ved Gund, Mayuri Ukidwe, Akanksha Kapoor, Isha Chaknalwar, Nikita Sengar, Adriana Hernandez, and Prateek Sehgal, for all the great time we have spent together.

Special thanks to my parents for their relentless support and sacrifices. My brother Priyaranjan, sister-in-law Kriti, and my adorable niece Tiana, for their unconditional love. Finally, I thank my fiancée, Mayank, for always being there for me and his persistent support and being a brainstorming-person for all my research work.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

With increased life expectancy, the population proportion of age 60 and above is increasing faster than other age groups worldwide [1] and is more vulnerable to chronic illnesses. Thus, many adults find themselves needing assistance with everyday tasks and care to varying levels. With limited at-home family care options and understaffed and costly nursing facilities, more seniors turn toward some form of residential facility based on consumer choice, requirements, and cost factors. These facilities, that promote independent living while taking care of personal daily needs and unscheduled assistance requirements, are termed as assisted living. With huge diversity in their conceptualization, the services vary depending on consumer affordability. This points to a clear need for technological solutions that can support senior assisted living with low-cost, high reliability, and least privacy invasion.

In this dissertation, radio frequency (RF) sensors are proposed as non-invasive or ambient technology that can support the elderly population to receive quality care at a low-cost. This occupant-centric approach of living can be seamlessly integrated into the daily lives as well as existing housing for higher comfort. Particularly, we have focused on two important aspects, ambient monitoring providing *safety and accessibility*, by turning on/off the light and fall-detection, and vital sign sensing for *wellness* monitoring, both with high accuracy to generate an emergency response from

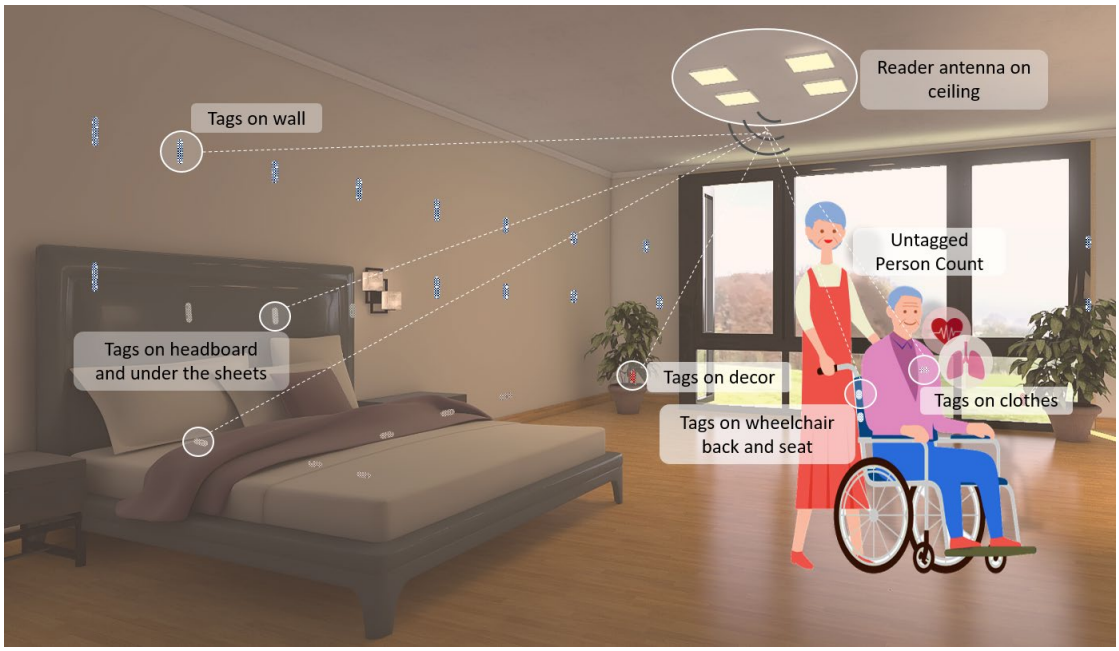


Figure 1.1 RF sensor setup for user wellness and ambient monitoring with RFID tags on the walls, furniture, and user clothes. Clothes tag can monitor vital signs and additionally, the system should be able to detect occupant count and locations, including untagged person, for safety monitoring.

the caregiver or nursing facilities. An overview of possible implementation is shown in Fig. 1.1.

1.1 Vital-sign sensing

In addition to the effective treatment of diseases, healthcare has been extended towards an overall healthier living with improved day-to-day quality of life [2], [3] and integrated end-of-life care not restricted to clinical visits [4], [5]. This requires novel solutions for noninvasive vital-sign sensing without skin contact which can be integrated effectively into various lifestyles for the elderly, as well as for the younger population. Such a technique could potentially help patients with asthma, chronic obstructive pulmonary disease (COPD) [6] by continuous monitoring of physiological parameters including heartbeat and respiration. Furthermore, noninvasive long-term

sleep monitoring and sleep quality analysis is important for wellness and health monitoring, recovery, and improvement of quality of life. Cumulative effects of irregular sleep durations have been associated with increased risk for diseases like diabetes, obesity, heart attack, and stroke [7]. With the high prevalence of sleep-breathing disorders such as central sleep apnea (CSA) and obstructive sleep apnea (OSA) in patients with symptomatic chronic heart failure (CHF) [8], continuous overnight monitoring becomes mandatory. Acute or chronic opioid use has also been linked to severe respiratory disorders, starting from respiratory pauses, delays in expiration and decreased RV that can lead to eventual respiratory arrest and death [9]. In addition to particular disease monitoring, even daily monitoring to track emotional and mental health is of high importance with increased risk of developing anxiety or neurological disorders.

While there is no one-sensor fits-all solution for the patients suffering from various conditions, most diseases and their progression can be monitored as a function of the two important vital signs, heartbeat, and respiratory patterns. Respiratory disorders can be observed in the respiratory waveform patterns, and through derived respiratory effort information, including both respiratory rate (RR) and respiratory volume (RV). Heartbeat motion and the heart rate (HR) estimation provides crucial information on health levels, including obesity. Further, with a clear heartbeat waveform, heart rate variability (HRV) can be derived, which gives insight into the autonomous nervous system (ANS) function with parasympathetic (PNS) and sympathetic nervous system (SNS) balance, that can be used to track mood, emotion, and anxiety [10].

1.1.1 Conventional sensing technologies

In clinical practices, pulmonary function tests are generally assessed by spirometry [11] and physical examinations [12]. The latter might not require any specialized instrument but is limited by the experience of the physician and is not quantitative. Spirometry provides detailed parameters characterizing lung function but requires attentive participation of the patient with forced breathing maneuvers, and is thus not feasible for continuous long-term measurements, or patients suffering from asthma. Whole-body plethysmography [13] is used to measure absolute air volume in the lungs and breathing resistances without extensive voluntary exercises but still requires some patient cooperation. This poses a challenge for respiratory disorder detection in patients that are unable to follow the instructions due to weakness, coma, or other cognitive failures. Other alternatives for respiratory monitoring that do not require active user participation, and are used as a part of polysomnogram (PSG) for sleep monitoring, include nasal probes [14] for airflow information, and strain or inductance belts at thorax and abdomen, which need to maintain reasonable tension in all breathing conditions to measure chest-wall motion. While the nasal flow can tell about the proportion of inhalation and exhalation volume exchange, it cannot provide respiratory volume (RV) information without calibration, similar to the chest belts. The nasal flow meter is typically used in addition to chest belts to identify different sleep apnea accurately. For example, CSA is marked by the cessation of airflow for at least 10 s, lasting up to several minutes, without any muscle activity or cough reflex, as the autonomic nervous system stops giving signals for autonomous respiration [15]. Thus, no change is observed on the chest belts, as well as in the nasal airflow reading. On the

other hand, OSA is more common in the elderly due to respiratory muscle weakness. This apnea is characterized by the asynchronous movement of the thorax and abdomen that can be observed as a time lag of the thoracoabdominal motion or complete out-of-phase motion of thorax and abdomen. In this case, the nasal flow meter shows little or no airflow, however chest belts at thorax and abdomen show thoracoabdominal asynchrony. Proper diagnosis and differentiation of the two apneas are important for correct treatment procedures.

Gold-standard heartbeat monitoring uses an electrocardiogram (ECG) with multiple adhesive skin-contact electrodes to monitor the electrical activity of the cardiac muscle. Pulse oximetry, providing SpO2, blood oxygen saturation information, is a less invasive method of extracting the heartbeat requiring a stable probe in direct contact with the skin [16]–[18], and prone to motion artifacts.

Particularly, for sleep apnea monitoring, overnight PSG is the gold standard requiring a plethora of sensors in addition to the ECG, pulse oximetry, nasal probes, and chest belts. These sensors measure brain activity by electroencephalography (EEG), eye movements by electrooculography (EOG), and muscle activity by electromyography (EMG) to classify different sleep stages and arousals to estimate accurate sleep quality.

1.1.2 Existing non-invasive sensors and their challenges

While conventional gold standard sensors are the norm, their accurate usage often requires specialized nursing staff to carry out the procedure. Also, the sensors are cumbersome to be used in daily life for long-term monitoring.

Respiratory monitoring can be performed with local strain sensors [19] that are comfortable and can get accurate thorax and abdomen respiratory signals separately. A

limitation is direct tight skin contact to measure the chest wall motion. Airflow sensors that estimate the volume of air expired over time can get accurate information on RR [20], but require mouthpieces or facemasks to collect the air for RV, which is not only uncomfortable and conspicuous but may also interfere with the normal breathing process. The pulse oximetry can provide information about blood O_2 content, but has a slow response to apnea-like conditions and suffers from motion artifact [21]. Heartbeat monitoring is typically done with ECG as the small heartbeat chest motion is often more difficult to detect with other noninvasive sensing techniques. While the ambulatory ECG is simple to put-on, it has poorer motion tolerance and thus signal processing can be challenging, while still requiring electrodes to be placed on the body. Thoracic surface and body vibrations due to heartbeat have been studied in detail using seismocardiography (SCG) [22]–[24] and ballistocardiography (BCG) [22], [25]–[27] respectively, for signatures under different abnormalities. Both sensors measure either displacement, velocity, or acceleration originated from the heartbeat but observed at the body surface. Film-based displacement sensors may be affected by artifacts from structural vibrations in a moving wheelchair [28] or ambulance. Integrated solutions like smart shirts [29] with ECG, chest belts, and accelerometers can provide detailed cardiopulmonary characteristics, but have a large device form-factor, and are unsuitable for bedridden and geriatric patients where snug-fit clothing and skin contact electrodes are impractical due to concerns of comfort and bedsores [30], [31]. Fabric electrodes are embedded in these smart-shirts without glued to the skin, however, the fabric must be damp to conduct the electrical signal from the skin which can be inconvenient for a wearable device. Ambient optical [32] and RF sensing [33] record surface motion from

breathing and heartbeat with a limited signal-to-noise ratio (SNR) [34], [35] and require the reader to be in the line of sight (LoS) to the torso [36], [37]. Far-field RF also has limitations on the maximum number of subjects that can be measured simultaneously [38]. RV is even more difficult to retrieve, as it is affected by body posture and orientation variation [37], [39] with respect to the antennas or cameras. While cardiopulmonary signals can be derived with limited SNR, movement of the person under test as well as any other motion in the surrounding can affect the collected data. Additionally, obstructed breathing monitoring is difficult, as this only measures the surface motion of either thorax or abdomen where the radar antenna is directed to. Some noticeable work has been done using RF in the near-field region [40], [41] that can couple to internal dielectric boundary change to clearly measure respiratory motion, but heartbeat can only be extracted during breath-hold.

1.1.3 RF near-field coherent sensing (NCS)

In this study, we employ a non-invasive method based on RF near-field coherent sensing (NCS) [42]–[44] that works by transmitting a low-power continuous wave (CW) RF signal into the body with over-clothing antennas. The near-field coupling to the internal dielectric boundary motion results in a direct measurement of the heart, lung, and diaphragm motion, in contrast to surface motion sensors. This requires a transmitter antenna (Tx) in the near-field of the motion source. The receiver antenna (Rx) can be realized in two different approaches – active [42] and passive [45] sensing. In passive sensing, Rx is in the far-field and the Tx can be a passive unit like an RFID tag placed in a pocket near the thorax or embroidered in the cloth [46]. The active setup is a wearable version with both Rx and Tx on the chest, close to each other. Both setup

designs have been considered and their pros and cons are discussed in this work. Overall, NCS is less sensitive to wearer and ambient motion and can be easily extended for multiple participants in the room, as well as multiple locations on the body. Detailed high-frequency heartbeat characteristics associated with the S1 and S2 sounds can also be potentially extracted from over-clothing placement [47]. However, a detailed comparison with echocardiogram can help clarify this observation, while a further improvement from possible sensor fusion with SCG and phonocardiogram (PCG) needs to be investigated in future studies.

1.2 Occupant detection and counting

The occupant-centered control (OCC) is an emerging concept of smart buildings [48], where real-time occupant monitoring enables various automation functions in assisted living [49] and energy control [50], [51]. In contrast to active observation units or occupant-marker approaches that assume all occupants to wear dedicated devices, ambient passive sensing is critical for applications where device-free occupants must be accounted for and most observation units are passive and maintenance-free. Key requirements for this passive sensing include low cost, low power, low computation, small training overhead, and high accuracy. Many indoor items will have radio-frequency (RF) links in the realization of the Internet of Things (IoT), which can further facilitate the development of a smart environment allowing successful aging of seniors.

1.2.1 Prior art and related work

In this work we have focused on ambient passive sensing approaches, instead of the active approaches, requiring a user to carry some device. Various technologies exist for the latter case, allowing user localization and tracking. In this thesis, we focus on two aspects of occupant monitoring: counting and localization.

The existing WiFi network has been employed as either ambient or active sensing technology [52], [53], and multiple learning methods have been employed. Amplitude-based channel state information (CSI) was extracted from multiple routers in a room, and a classifier was trained to detect [54] and count the number of participants [55], [56]. In [57], randomly moving participants were trained in one room and tested in different rooms based on the features of the received signal strength indicator (RSSI). A probabilistic model of multipath and fading characteristics was developed in [58]. A phase-scattering model was proposed [59] to count people passing through a gate. However, all of these approaches are limited to counting moving occupants and may perform poorly in indoor office or home settings with limited occupant activity. Most Wi-Fi based localization systems do not require user devices to actively interact, but use RSSI information from user devices transmitted as a part of their routine scanning or connection periods [60], which makes it difficult to localize occupants not carrying phone or tablets.

Custom RF sensors using RSSI information are mostly limited to one participant localization and tracking [61], [62]. A large number of an active receiver (Rx) and transmitter (Tx) units were used to count up to four moving participants [63] with only one-person fingerprinting, which however performed poorly when their trajectories were close. Wireless sensor networks (WSN) were distributed in the entire monitoring

area to estimate crowd densities in [64] and classify as low, moderate, or high density of randomly moving occupants. Radio tomography imaging (RTI) based approaches have been used to generate RF images for occupant localization [65] with lots of active links and far away from walls and ceiling to reduce multipath.

Extensive work has been established in the area of optical approaches for indoor occupancy estimate [68], [69], [206], including RGB camera and passive infra-red (PIR) sensors [207], [208]. Solutions have been presented to reduce problem complexity [209] and privacy concerns [210]. These systems can also be very helpful in human motion analysis [211], along with occupant counting and localization. However, camera-based monitoring suffers from occlusion and illumination deficiency [212], and large area coverage and complex layout result in a steep rise in system costs. The performance rapidly degrades with an increasing number of occupants [212], [213], who are either located very close to each other due to limited space or are outside the camera scope of view. Other approaches use data fusion of environmental sensors including temperature, humidity, light, sound, and CO_2 [51], [66] but require a large amount of training data in different environmental conditions, to avoid the bias of varying factors, for example, day and night lighting conditions. A wall-based capacitive sensor was used in [67] but testing was limited to one participant.

With the adoption of radio frequency identification (RFID) technology in commodity items, passive RFID tags are maintenance-free and highly cost-effective for various applications [70] including supply-chain logistics [71], libraries [72], and healthcare [73], [74]. For device-free object localization and tracking [75], passive ultra-high-frequency (UHF) RFID tags offer the unique advantage of increasing the spatial

diversity of RF links without a significant cost. Previous works can be largely classified into static fingerprinting and mobile counting. RSSI fingerprinting requires a database of one participant standing at different cells in the capture volume to enable occupant localization [27 28]. Such systems require tags to be distributed around the region to reduce ambiguity. The other category can count moving participants with fewer tags but have difficulty in handling static occupants [29]. RTI has poor resolution and noisy performance with RFID tag links, demonstrated in [76] using floor tags to detect the location of one occupant, and in [77] showing tests with up to four occupants, but poor occupant detection accuracy for more than one target.

1.2.2 RFID RSSI and carrier phase based ambient monitoring

With the advantages of RFID technology, in this work, we have adopted the ambient passive UHF RFID tags for occupant counting and localization using both carrier phase and RSSI information. As the phase is a function of distance traveled by the EM wave, it has information about the occupants' presence, along with multipath noise leading to constructive and destructive interferences, additional phase from cables, reader transceiver circuitry, and tag modulation. An improved calibration approach is presented to remove phase noise and use it for both occupant localization and counting. As the signal model becomes increasingly non-linear with more number of occupants in the room, we use two separate approaches for counting and localization. The counting of five occupants is shown to be performed with high accuracy using a deep-learning (DL) model. For localization, while a similar training method can be used, the training data increases, requiring fingerprinting at every location in the grid. Thus,

a linear approximation model for RF imaging is proposed, based on reflectivity estimation at each voxel in the room. A commercially available RFID reader has been used for demonstrating the results, which can further allow easier future production at lower costs.

1.3 Thesis organization

This dissertation presents two main contributions in the area of assisted living technologies focusing on vital sign and occupancy monitoring. We first focus on wellness monitoring by presenting details of the RF near-field sensing system, sensor configurations, signal model, and HR, RR and RV extraction algorithms in Chapter 2. Next in Chapter 3, respiratory pattern variation under different healthy and disorder conditions are shown along with motion and other interference analysis. The results of sensor design and algorithms are validated with a human study with 25 participants. Chapter 4 is an exploratory project where heartbeat and respiratory patterns derived from NCS are used for attention and relaxation monitoring. Chapters 5 and 6 focus on the occupant monitoring part, starting from the RF backscatter signal model and novel background calibration approach in Chapter 5. The RF imaging inverse problem and its sparsity-approximation solution are presented. It also compares the signal model with existing RTI model results. Finally, DL-based occupant counting results are presented in Chapter 6 with a detailed comparison of RSSI and phase information and learning transference over unseen setups. Chapter 7 concludes the dissertation with major contributions and future opportunities.

CHAPTER 2

NEAR-FIELD VITAL SIGN SENSING

2.1 Respiration and cardiac motion

Respiratory function is driven by inspiratory muscles, with the diaphragm acting as the major muscle, separating the thoracic cavity from the abdominal cavity. During normal or tidal breathing, the diaphragm contracts and flattens, and thus moves downwards, and ribs are simultaneously being pulled upwards by the intercostal muscles. This results in the enlargement of the thoracic cavity's both vertical and horizontal dimensions, causing the intrathoracic pressure to fall, and the pressure around and in the lungs to be negative relative to the atmosphere, leading to airflow into the lungs. As the diaphragm moves downwards, it pushes the abdominal organs downwards as well. As the pelvic floor prevents the movement in the vertical direction, it causes the abdomen to bulge outwards. This process is called inhalation. During exhalation, the diaphragm and intercostal muscles relax, returning the thoracic and abdominal cavities to their original position, with lungs containing their functional residual capacity of air, generally between 2.5-3.0 liters. The volume of air exchanged during one inhalation and exhalation cycle of normal breathing is termed as tidal volume. This volume change is linearly related to the thorax and abdomen movement [78]. During heavy breathing as seen in people with COPD, there are accessory muscles of inhalation, causing greater contraction of diaphragm than at rest, resulting in a larger increase of rib cage internal volume. The exhalation process also involves additional contraction of the abdomen muscles, causing the diaphragm to move upwards, causing less air in the lungs than the resting functional residual capacity.

Normally, the enlargement of the thorax and abdomen wall occur together in phase, and total respiratory volume can be considered as a sum of the rib cage and abdominal volume. Thus, signals measured as a function of the rib cage or thoracic and abdominal motions can be calibrated to determine overall respiratory volume. In another case, a paradoxical motion of the thorax and abdomen occurs if the diaphragm ceases to contract, or the accessory muscles lose tone and the upper part of the rib cage becomes unstable, or with partial or complete upper airway obstruction. The partial obstruction is mainly a problem for the infants, where the diaphragm contraction creates highly negative intrathoracic pressure, that could result in inward motion of the least stable portions of the rib cage during inspiration. The complete upper airway obstruction generally occurs during sleep and is termed obstructive sleep apnea (OSA), resulting in no movement of air into the lungs. As the volume change is zero, the volume change in the abdomen is equal and opposite in direction to the volume change of the rib cage.

The heartbeat motion is a rhythmic motion performing the cardiac cycle, consisting of relaxation of heart muscles and filling with blood, followed by a period of contraction and pumping of blood. While the heartbeat motion consists of the motion of four chambers, four valves, and other accessory muscles, the largest motion is due to a change in the ventricular volume and can be seen in the RF signal as a periodic motion.

2.2 Sensor configuration

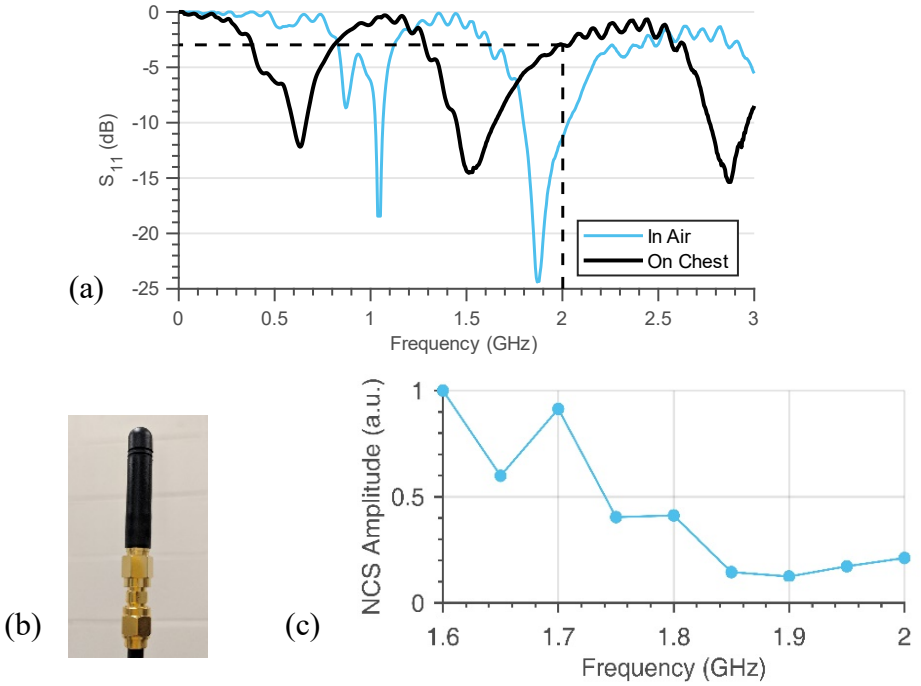


Figure 2.1 (a) S_{11} characteristics in air and on the chest showing shift towards the left. (b) The monopole helical antenna selected for the experiment. (c) The average NCS amplitude trend with frequency, scaled to the value at 1.6 GHz.

We have implemented the sensor prototype by commercial off-the-shelf (COTS) components. UHF monopole antennas are used at both Tx and Rx. The dielectric composition in the near-field region of an antenna will modulate its characteristics. For a single antenna, this change can be measured from the antenna reflection parameter S_{11}. For an antenna pair, this can be derived from the cross-coupling S_{21}. As the Tx and Rx signal chains are better isolated in the S_{21} measurements with less self-interference and higher SNR, we opt to place an antenna pair as part of the NCS sensor to the region of interest, where the intended surface and internal boundary motion can be retrieved after baseband demodulation. Notice that UHF has reasonable penetration into dielectrics in the near-field, and thus the internal dielectric motion during breathing and heartbeat can be locally modulated onto the specific

antenna pair. As the Rx signal consists of both the modulated and unmodulated parts, the relative antenna position and the body placement are optimized for the clearest signal.

2.2.1 Antenna characteristics and radio architecture

Due to the higher permittivity of the human body, the antenna S_{11} characteristics shift towards the left when placed on the chest, compared to in-air, as shown in Fig. 2.1 (a) for a monopole whip antenna shown in Fig. 2.1 (b). The corresponding NCS amplitude vs. frequency plot is shown in Fig. 2.1 (c) for a frequency range of 1.6 to 2.0 GHz. Although S_{11} rises monotonically in the frequency range which implies less RF energy transmitting towards the body, the NCS amplitude shows oscillatory behavior possibly due to additional physical mechanisms including inner tissue impedance matching, level of harmonic self-interference, signal strength balance between the transmitting and re-emitting waves, and wavelength difference that causes different interference features. Nevertheless, in the entire frequency range tested, sufficient RF energy is coupled into the heartbeat motion and reasonable SNR can be maintained.

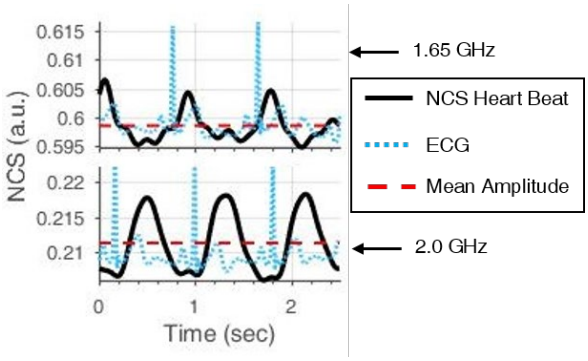


Figure 2.2 Extracted NCS heartbeat waveform for 1.65 GHz and 2.0 GHz, shifted at their respective scaled mean amplitude shown in Fig. 2.1 (c) , with corresponding ECG waveforms.

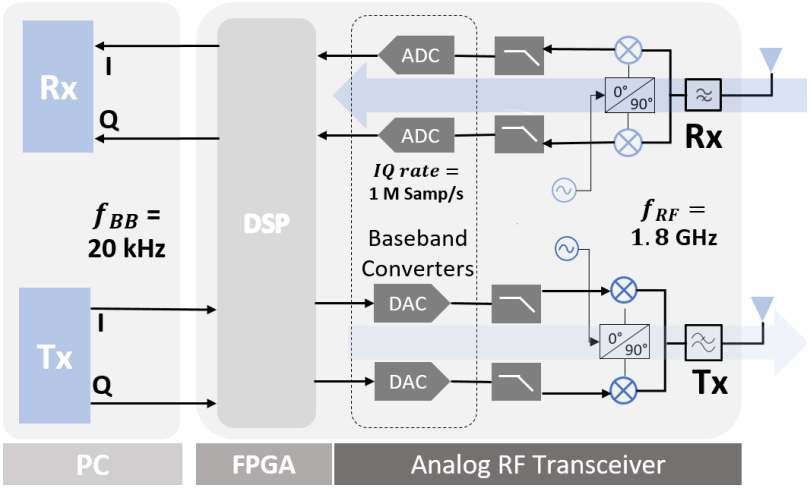


Figure 2.3 Ettus SDR architecture showing quadrature I-Q implementation.

Fig. 2.2 shows the extracted NCS heartbeat waveform, shifted at the scaled-mean amplitude for two frequencies at 1.65 and 2.0 GHz at different incidences, with corresponding ECG waveforms, whose magnitude is normalized for plotting purposes and timing is synchronized for feature comparison.

The RF front end is implemented by a National Instrument Ettus B210 software-defined radio (SDR) [79], with LabVIEW for real-time data recording. The quadrature implementation is shown in Fig. 2.3. The carrier frequency is set between 1 – 1.9 GHz for different antenna selection and multiple antenna placements, with a baseband

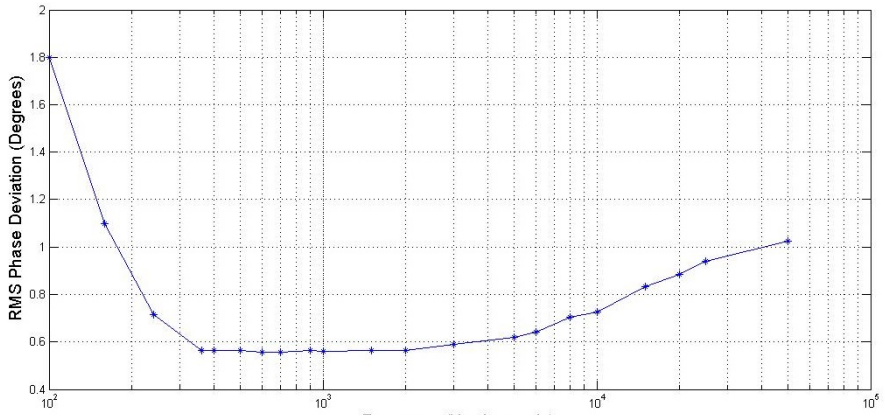


Figure 2.4 The RMS phase deviation vs baseband frequency f_{BB} plot for Ettus USRP B210.

between 5 - 80 kHz, having a typical value of around 20 kHz. The baseband frequency selection has a tradeoff between phase noise, inherent signal frequency content, and filtering requirements. The relation between root mean square (RMS) phase noise and corresponding intermediate baseband frequency measure for universal software radio peripheral (USRP) B210 SDR is shown in Fig. 2.4. The required Tx power is well below the safety limits [80], with -12.84 dBm and -10.42 dBm for the thorax and abdomen sensors used during the human study protocol. The use of baseband modulation allows multiple sensors to share the same radio by assigning different baseband tones at the same carrier frequency. SDR B210 supports a 2×2 multiple-input-multiple-output (MIMO) setup to implement two NCS sensors. Another Ettus SDR B200 mini [81] can be placed entirely on the chest and minimizes RF cable motion during measurement, similar to eventual on-chest sensor placement. Different setup variations are discussed in detail in the next section. The demodulated baseband is sampled at 1 – 2 M samples per second (Sps) by the built-in analog-to-digital converter (ADC).

2.2.2 Setup variations

As our sensing approach modulates the near-field motion onto the RF carrier, the demodulated waveform depends on the antenna chest placement. A major design consideration is passive and active sensing with Rx in far-field and near-field respectively, as discussed earlier. Other variations include antenna and frequency selection, and MIMO implementations.

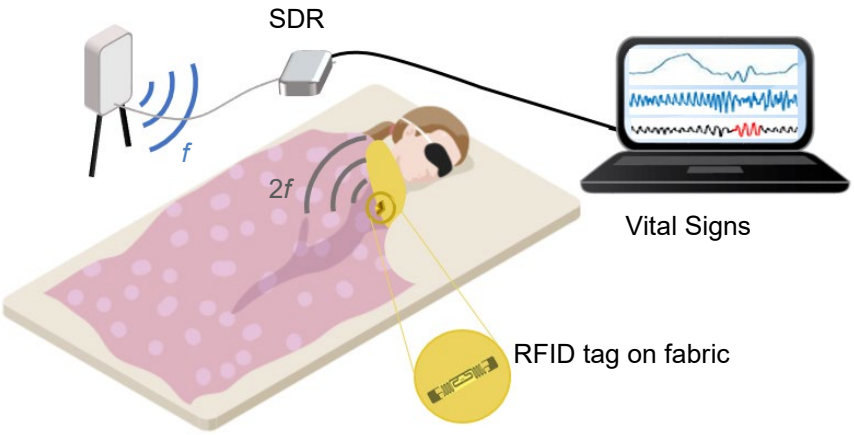


Figure 2.6: Wearable passive NCS setup for breath, heartbeat and motion detection. A passive harmonic RFID tag is deployed in the chest area and backscatters second harmonic frequency to the reader antenna.

The simplest wearable sensing setup is based on passive sensing with an RFID tag on the chest, as shown in Fig. 2.5. Further, the harmonic RFID approach has been used with the tag prototyped on a Wireless Identification and Sensing Platform (WISP) [82] for ease of antenna selection and air protocol modification. The harmonics are

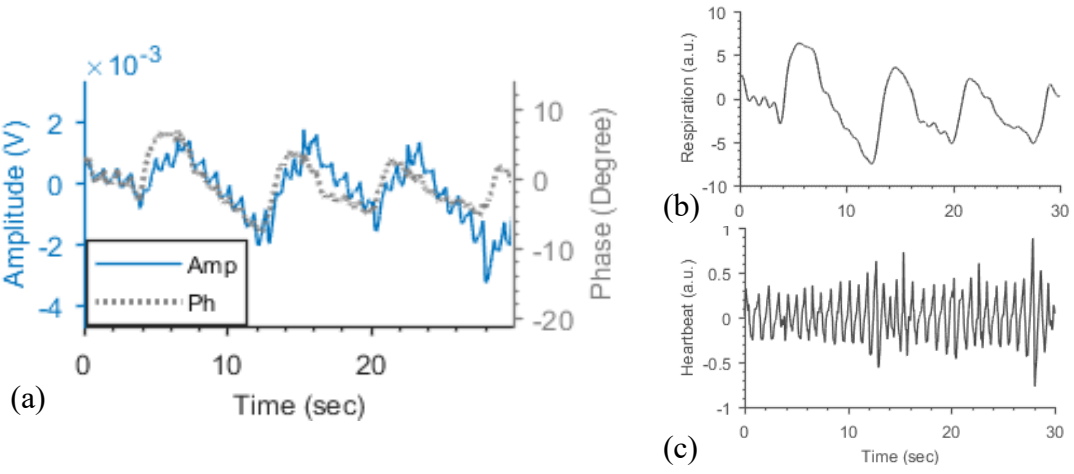


Figure 2.5 (a) Raw NCS amplitude and phase signals showing respiration and heartbeat information. (b) Respiration extraction from phase. (c) Heartbeat extraction from amplitude.

introduced by a nonlinear transmission line (NLTL) design [83]. A commercially available UHF monopole antenna with center frequencies at 1.0 and 1.8 GHz in free space was chosen. The work in [77] shows the separation of downlink and uplink on the harmonics improves received signal quality by improved self-jamming cancellation.

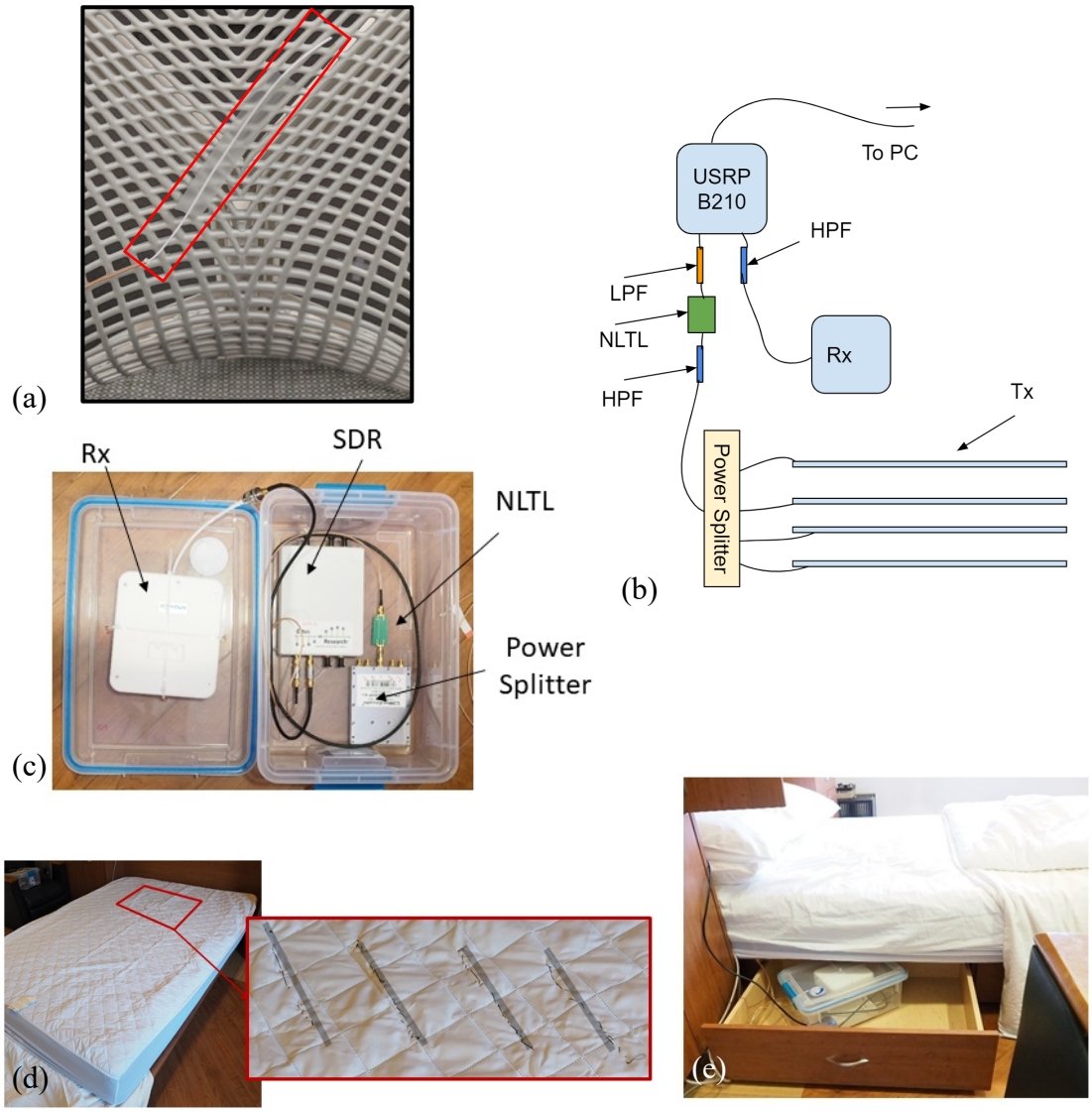


Figure 2.7 Furniture integrated invisible NCS setup for sleep monitoring. (a) Monopole stripped coax cable antenna pasted on back of a chair for testing. (b) The setup showing SDR Tx-Rx chain with harmonic NLTL and 4 Tx antennas powered by a 4-way power splitter. (c) A box containing SDR, power splitter and Rx on the lid. (d) Sleep center bed with coax antennas under the mattress topper. (e) The Rx antenna and the box placed under the bed in a drawer.

The setup has the further advantage of allowing a light-weight passive RFID tag to be placed on the body and even integrated into the fabric [46] with no user discomfort or circadian rhythm disruption from skin touch or motion constraints. The limitations of this setup include sensitivity to participant motion and ambient interference with Rx in the far-field. Thus, this setup is best suited for sleep monitoring, as shown in Fig. 2.5. The waveforms shown in Fig. 2.6 are derived from this setup.

Another variant of this setup is the furniture integrated version, which can be placed on the chair back or on-the-bed. This work particularly focuses on the sleep apnea conditions, and thus a bed version was designed and placed at Cornell Weill Medical Sleep Center, NYC, NY, US. As the antennas need to be thin enough to be unnoticeable, a stripped half-wave monopole co-ax cable antenna was designed as shown in Fig. 2.7 (a). Also, as the antenna need to cover a large area on the bed allowing less sensitivity to user position, it was designed to be operated at 900 MHz and the length was experimentally selected to be 19.4 cm, compared to theoretical $\lambda/2$ =16.7 cm to give optimal performance. Fig. 2.7 (b) shows the setup schematic showing USRP B210 connected to an NLTL on the Tx side to get the harmonic frequency of 900 MHz, transmitted to four antennas connected to a 4-way power splitter. The antennas were sewn under the mattress topper as shown in Fig. 2.7 (d), to provide position tolerance in both horizontal and vertical direction. The components: SDR, NLTL, and Rx were placed in a box, as shown in Fig. 2.7 (c) and placed under the bed as shown in Fig. 2.7 (e), such that the Rx was below the Tx antennas. The respiration pattern can be observed in the NCS signal as shown in Fig. 2.8, which also shows comparing chest belts signal from the PSG and the extracted RR from both the devices. CSA can be extracted from

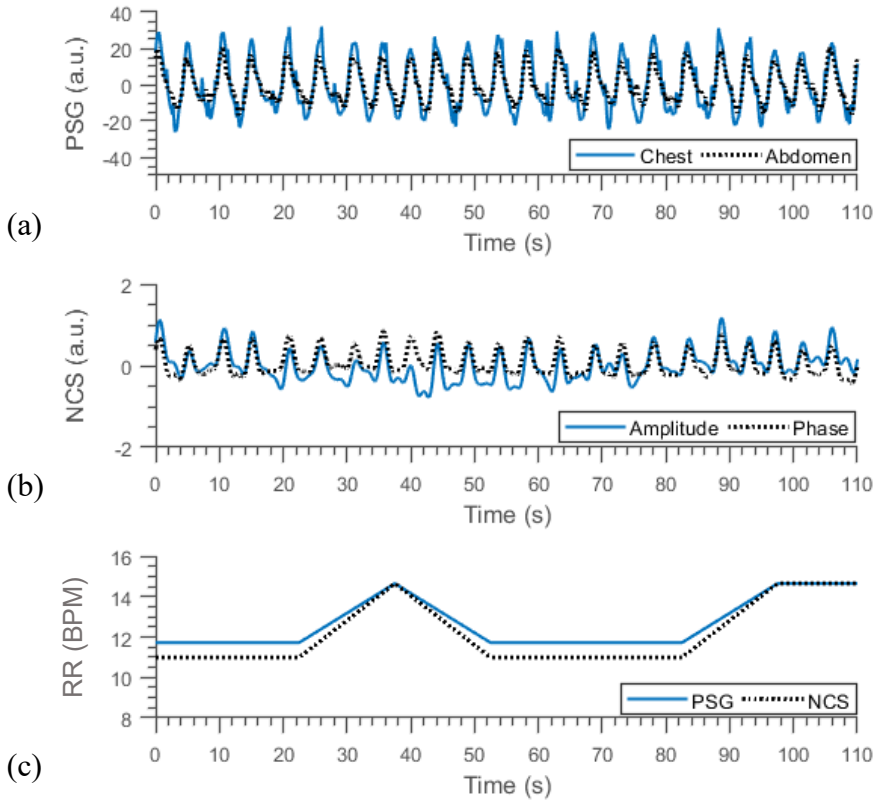


Figure 2.8 Respiration waveform during sleep study on a patient. (a) Thorax (chest) and abdomen waveforms from chest belts in polysomnogram (PSG). (b) NCS amplitude and phase waveforms collected from setup which was shown in Fig. 2.7. (c) Estimated respiratory rate from PSG and NCS using a short-time Fourier transform (STFT).

the NCS signal as shown in Fig. 2.9, where no motion is seen for extended durations in both amplitude and phase readings. However, OSA is difficult as this setup couples to both thorax and abdomen motion without distinguishing between the two. PSG apnea detection shows a slight lag due to poor time synchronization. Fig. 2.10 shows few occurrences of OSA, with some coupling variation between amplitude and phase that could likely be detected by machine learning (ML) and artificial intelligence (AI) based approaches. However, no clear paradoxical motion is visible, and thus difficult to be detected with some pattern matching approaches. A new setup with different thorax and

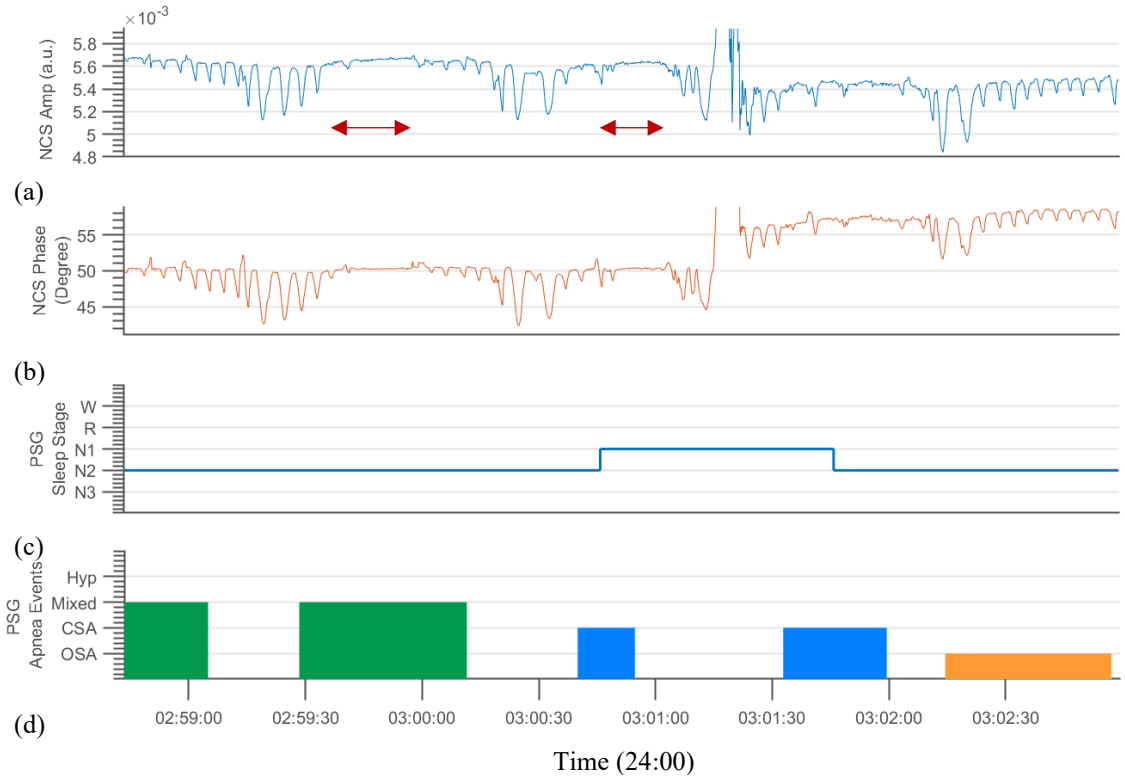


Figure 2.9 NCS and PSG readings for a patient. (a) and (b) show NCS amplitude and phase respectively, with clear no breathing motion during CSA events indicated by arrows. (c) Detected sleep stage. (d) Detected apnea events.

abdomen sensors is implemented [84], also using stripped co-ax cables, acting like notched transmission lines.

A single active sensor can be more conveniently deployed, where the Tx – Rx antenna pair is placed below the xiphoid process to effectively capture the respiratory motion by diaphragm and abdomen, as shown in Fig. 2.11, with the only faint signal from the heartbeat. The sensor can work over layers of clothing without skin contact and does not need to be held tightly against the chest, greatly improving wearer comfort. This setup can correctly capture RR and RV with high accuracy for most breathing patterns and showed low sensitivity to body placement. Tests with the sensor placed 5 cm above and below the xiphoid process showed that respiration can be captured at all

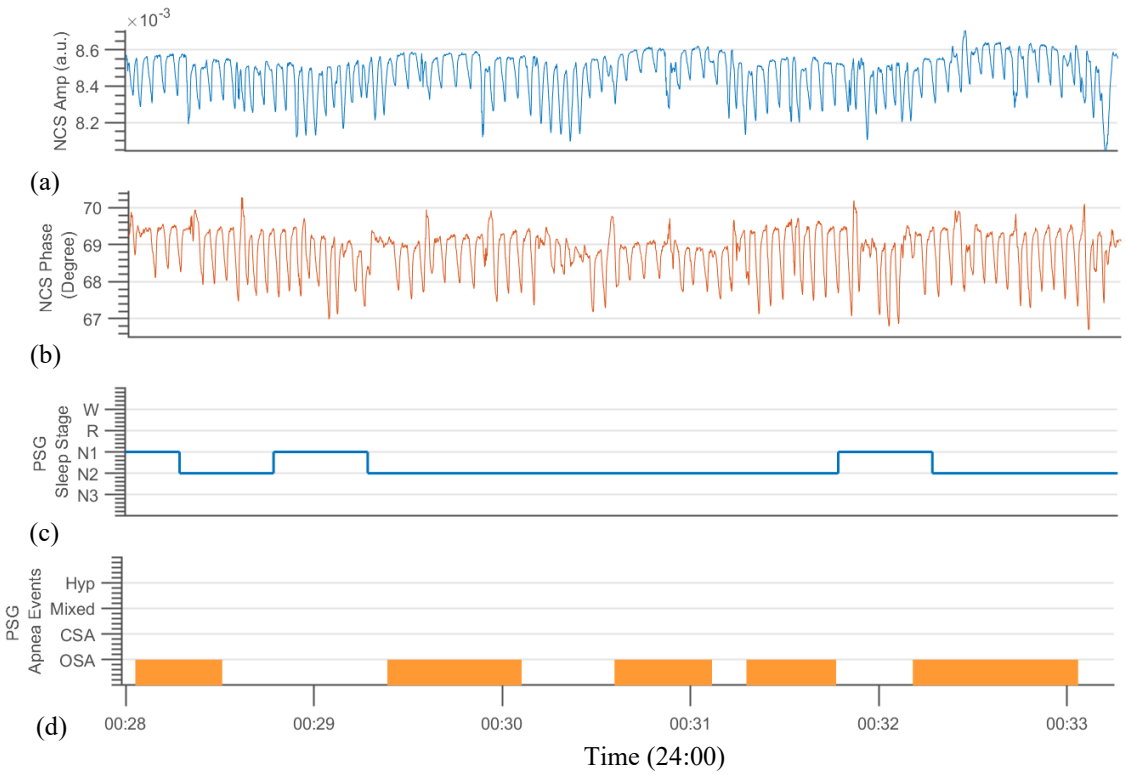


Figure 2.10 NCS and PSG readings similar to Fig. 2.9, showing cases of OSA. The amplitude (a) and phase (b) respiration waveforms are not identical and show complete (00:28:10) or little (00:29:20, 00:30:10) asynchronous motion.

three locations, requiring only different calibration coefficients. This active sensing setup is much more tolerant to ambient interferences and can be shielded against the motion of arms and legs as well. The detailed motion interference results are discussed

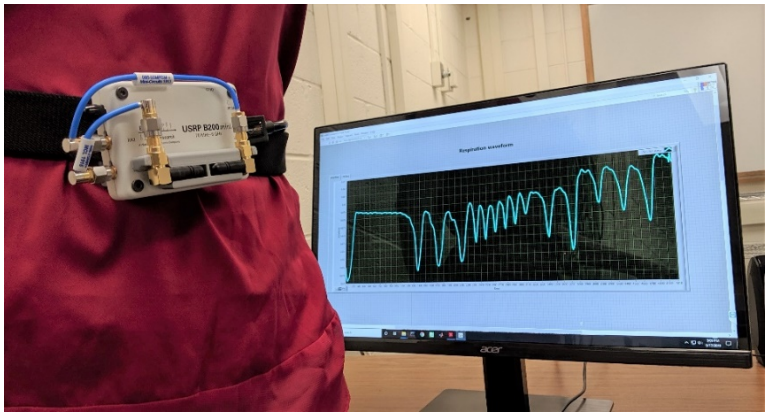


Figure 2.11 One active sensor implementation below xiphoid process.

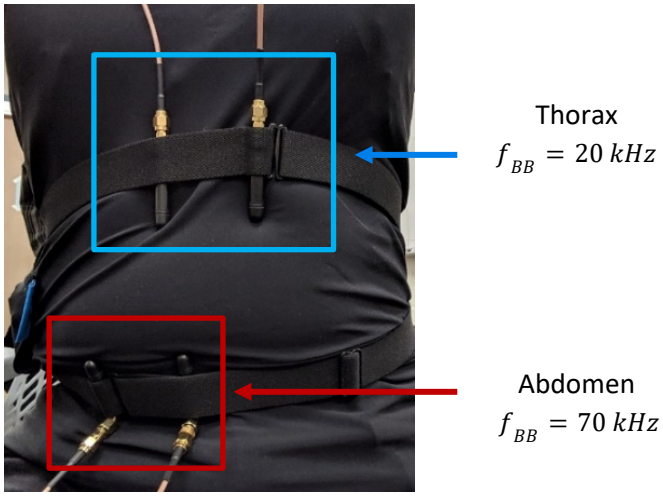


Figure 2.12 Two sensor implementation at thorax and abdomen.

in a later section and simulated respiratory disorder signals with this sensor are presented in detail in the next chapter.

Finally, a two-sensor placement is considered, with one sensor placed high at the thorax, and the other at the abdomen. This allows our sensor to accurately monitor the scenario of out-of-phase abdomen and thorax motion in OSA with paradoxical breathing [85]. This sensor is best suited to get both clear heartbeat and respiratory signal and provides increased tolerance to motion and other noise interference with two-point respiration measurement. An initial test setup is shown in Fig. 2.12, using one carrier frequency with two different intermediate baseband frequencies, f_{BB} with USRP B210. The setup is further refined and used for human studies in Chapter 4.

2.3 Signal model

The near-field motion is modulated onto the received data and can be retrieved by signal processing of the demodulated baseband data. The Tx quadrature baseband is presently a simple sinusoidal tone given by

$$I_{Tx}(t) = A\cos(2\pi f_{BB}t)$$

$$Q_{Tx}(t) = A\sin(2\pi f_{BB}t). \quad (2.1)$$

Here f_{BB} is the baseband frequency and A is the amplitude. Multiplexing is possible by setting different baseband frequencies or code divisions for different channels. B210 supports two channels that can be operated simultaneously. This digital baseband passes through the digital-to-analog converter (DAC) and is then mixed with the carrier frequency, f_{RF}. The Tx signal is modulated by the physical motion in the near-field region of the Tx antenna and the Rx signal is demodulated and sampled by the ADC to get the baseband quadrature, which needs to be bandpass-filtered around the channel tone f_{BB} to obtain data from different channels

$$I_{Rx}(t) = A(t)\cos(2\pi f_{BB}t + \theta(t))$$

$$Q_{Rx}(t) = A(t)\sin\big(2\pi f_{BB}t + \theta(t)\big). \quad (2.2)$$

The intended motion is modulated on the IQ amplitude $A(t)$ (NCS_{amp}) and phase $\theta(t)$ (NCS_{ph}), and can be extracted as,

$$NCS_{amp}(t) = \sqrt{I_{Rx}(t)^2 + Q_{Rx}(t)^2} \quad , \quad (2.3)$$

$$NCS_{ph}(t) = unwrap(\tan^{-1}\left(\frac{Q_{Rx}(t)}{I_{Rx}(t)}\right) - 2\pi f_{BB}t + \theta_0), \quad (2.4)$$

where θ_0 is a constant phase offset including the initial phase and accumulation from the Tx-Rx path and cables. While both amplitude and phase contain the modulated motion, their coupling strength varies with the sensor placement, which provides further resistance to noises when both magnitude and phase are included for further analysis. Additional studies are needed to clearly understand the coupling tradeoffs. For this work, when only one sensor is available, generally heart rate is extracted from amplitude

and respiration from the phase, as internal body coupling is usually better on amplitude. The two sensor setup is discussed in Chapter 4.

2.4 Physiological signal processing

The raw NCS signal needs to be correctly pre-processed to extract further information. This includes operations like filtering, peak detection, and signal artifact detection and correction. This section discusses both pre-processing and final HR, RR, and RV estimation algorithms in detail.

2.4.1 Peak detection

With the high SNR signal available, we can observe every heartbeat and respiratory cycle in the time-domain to support real-time analysis and variability in both the waveforms, without resorting to frequency-domain techniques. This work aims to use tuning-free or minimally tuned algorithms for both respiration and heartbeat peak detection. To that purpose, modified versions of two algorithms have been used that work with high accuracy without any tuning from person-to-person or varying respiratory patterns.

A modified algorithm of automated multiscale-based peak detection (AMPD) [86] is used, due to its effectiveness against noisy quasi-periodic waveforms and its simplicity in terms of no free parameter selection. The limitations of AMPD include 1) long processing time and large memory requirement for a long signal segment, and 2) the highest frequency (f_{max}) should not be larger than 4 times the lowest frequency (f_{min}) in one signal segment. As respiration signal can have a large frequency range for slow breathing at 6 breaths per minute (BPM) to panting at 60 BPM (0.1 – 1 Hz), a windowing was introduced with 50% overlap, as sudden frequency jump is not expected

over a small segment. The algorithm performs maxima and minima detection over multiple scales and keeps the peaks that are maximum over all the scales, and discards the rest of the maxima, as shown in the flowchart in Fig. 2.13. This repeats over each window, where the window length is selected to include 1 cycle at the least frequency, which is taken as 6 BPM for the breath. It can be further updated by first taking a FT over the next 30s and then updating the window to include few cycles corresponding to the peak RR. For this small window, it will satisfy the frequency range limit of [f_{min}, $4f_{min}$] while allowing the signal to have a broad frequency range over time. The processing speed and memory requirements are also improved. As the heart rate does not show sudden variation, it does not have any such strong window size dependence

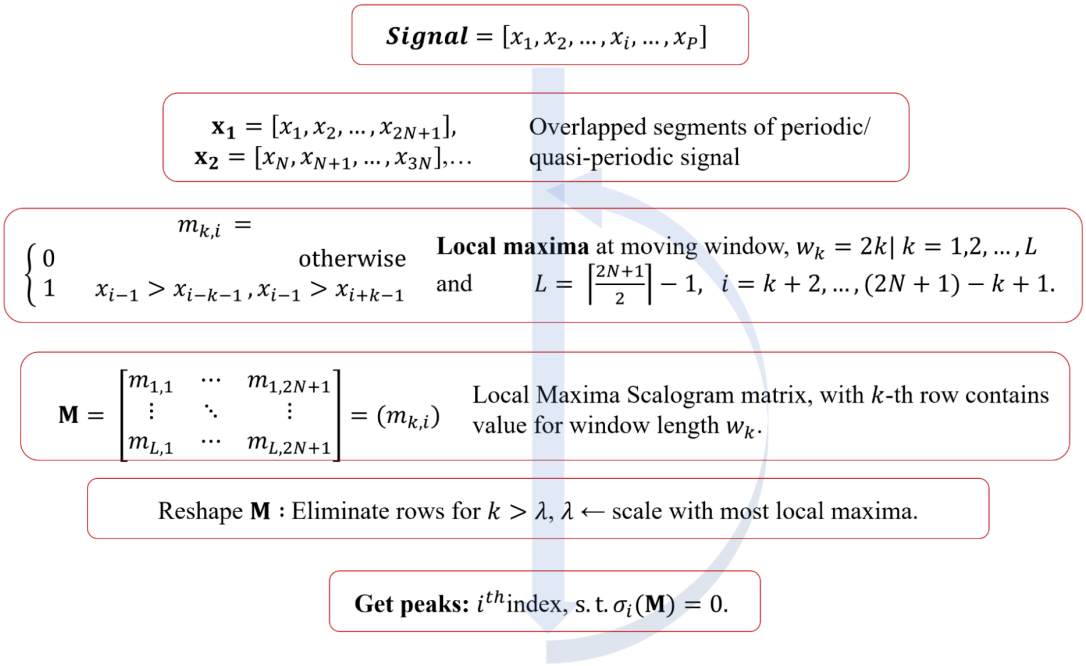


Figure 2.13 Flowchart showing implementation of modified AMPD algorithm using local maxima scalogram (LMS) matrix.

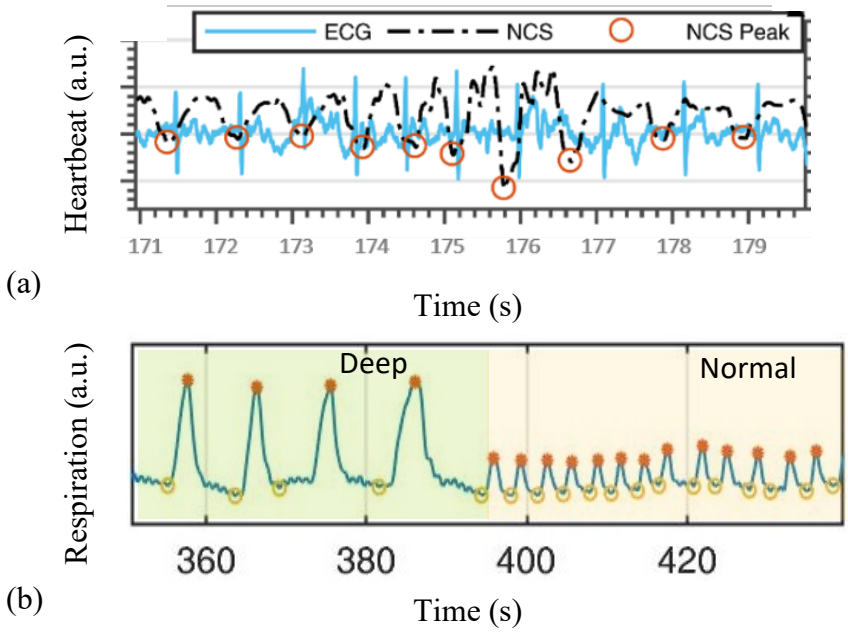


Figure 2.14 NCS heartbeat (a) and respiration (b) waveforms showing peaks detected by AMPD algorithm.

for this algorithm. Fig. 2.14 shows the sharp minima peaks for the heart rate and both minima and maxima peaks from the respiration using this algorithm.

While the above algorithm works well, it is prone to error at sudden peak jumps and very irregular breathing, which can occur in some of the respiratory disorders. Moreover, the computation requirements are huge, with a large local maxima scalogram (LMS) matrix M, shown in Fig. 2.13. An improved peak detection algorithm, called moving average-crossing algorithm [87] is implemented, which is effective for signals with varying amplitude and frequency characteristics, such as the respiration signal here with the RR in the range of 0 – 50 BPM, without any manual tuning. A moving average is estimated at every point using a given window length, resulting in a moving average curve (MAC). This window length is selected to have approximately one respiratory

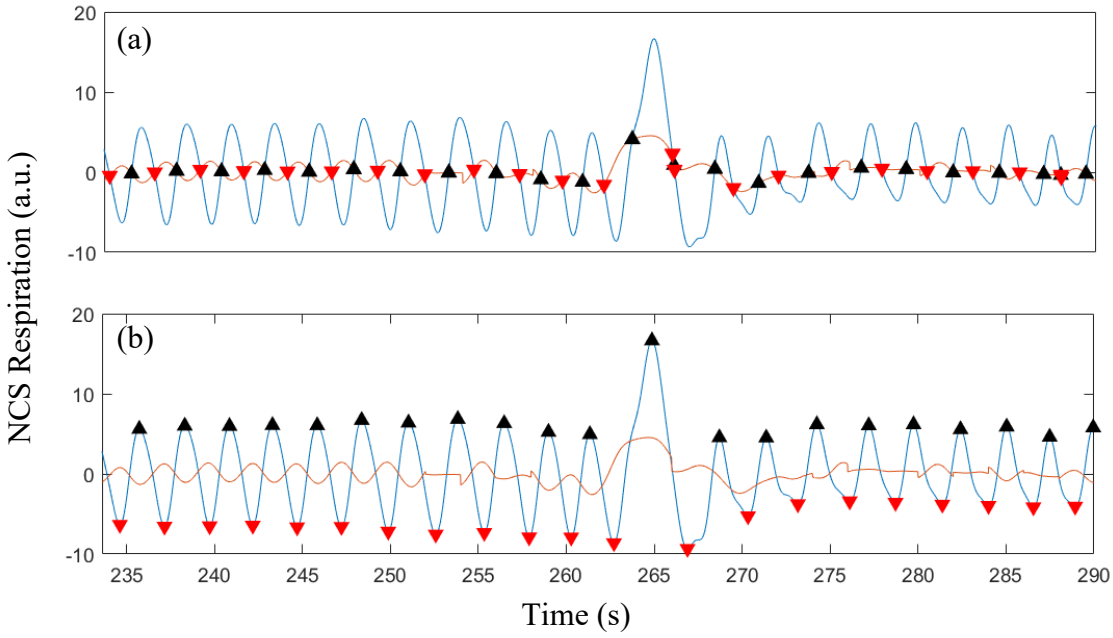


Figure 2.15 Normalized NCS respiration waveform and peak detection. (a) The MAC curve (orange) showing positive-slope intercepts and negative-slope intercepts in black and red colors, respectively. (b) Maxima (black) and minima (red) points detection showing end of inhalation and end of exhalation, respectively.

cycle in each window and is constantly updated by taking a Fourier transform over a fixed period to estimate the RR. The points where MAC crosses the original signal are labeled as intercepts and are classified as up or down intercepts for positive and negative slopes, respectively. Finally, a maximum is marked as the maximal point between two up-down intercepts and a minimum is also marked similarly. This algorithm is used in the estimation of all three parameters: RV, RR, and HR. Fig. 2.15 shows the original respiration signal, its MAC, and the final extracted minima and maxima peaks.

2.4.2 Signal conditioning

The NCS waveform needs to be pre-processed to remove baseline drift. Further, the respiration and heartbeat signal needs to be filtered properly. The entire processing

is performed using MATLAB. The respiratory and heartbeat signals are highpass filtered with a zero-phase Butterworth digital filter of order = 3 and 3-dB cutoff frequency, f_{3db} = 0.05 Hz and 0.7 Hz respectively. Both the signals are lowpass filtered with a Kaiser window filter with a passband ripple of 0.1 and stopband attenuation of 10 dB, with passband and stopband frequencies $[f_P, f_{st}]$ of [0.8,1.2] Hz and [9,10] Hz for respiration and heartbeat respectively.

Another observed phenomenon is the coupled signal's polarity reversal, i.e., increasing volume during inspiration is shown as a decreasing signal, which is dependent on sensor placement. While this information is not required directly for respiration rate or volume estimate, it can be helpful to distinguish end-of-inspiration and end-of-expiration peaks. A method based on respiration physiology is proposed to correct the waveform polarity, assuming end-of-inspiration is expected to be a peak. The normal respiration is expected to have a lesser inspiratory interval, t_{II} (time from end-of-expiration to end-of-inspiration) than expiratory interval, t_{EI} (time from end-of-

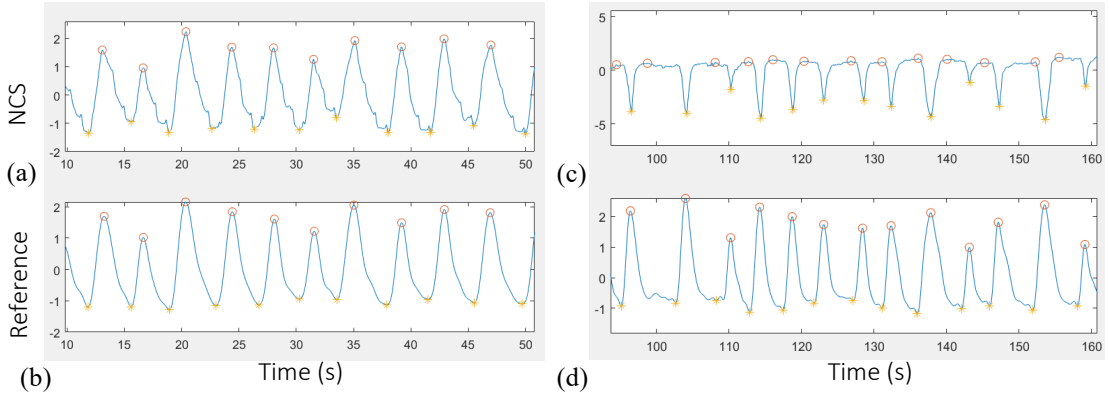


Figure 2.16 NCS waveform polarity correction. (a) and (b) Good case: Inspiration and expiration peaks are correctly detected in NCS leading to accurate sign correction, with NCS waveform in-phase with reference chest belt waveform. (c) and (d) Poor case: The NCS waveform respiration coupling has a flat characteristic during expiration, leading to poorer peak detection. Thus, waveform polarity correction fails.

inspiration to end-of- expiration), i.e., $t_{EI}/t_{II} > 1$. Following are the steps for respiration polarity correction: 1) maxima and minima detection on the filtered respiratory signal, 2) estimate t_{EI}/t_{II}; if <1, flip waveform by multiplying by -1, otherwise do nothing. This method resulted in accurate waveform polarity correction for 71/75 recorded normal breathing data segments with an accuracy of 94.67%. Fig. 2.16 shows a good and poor case using this approach.

Other signal distortions occur particularly during deep breathing, likely due to saturation. This can be taken care of by automatic Rx gain adjustment during data recording if the normalized received signal amplitude exceeds $1/\sqrt{2}$. Motion interference is another important consideration, that can occur due to sudden torso jerk or poor sensor stability on the body. Thus, another important step is signal quality detection to reject poor quality signal due to any of the above reasons. As the reference signals can also have motion interference, a signal with interference in either NCS or reference sensor is rejected.

- **Respiration Data Quality**: As thorax and abdomen sensors are independently placed, motion artifacts may be present differently, leading to a poor correlation between the two measurements. Also, data from NCS and reference chest belts can show dissimilar artifacts, with both or only either one of them showing the artifact. Thus, correlations between thorax and abdomen chest belts, and NCS and chest belts for both thorax and abdomen are calculated with an epoch

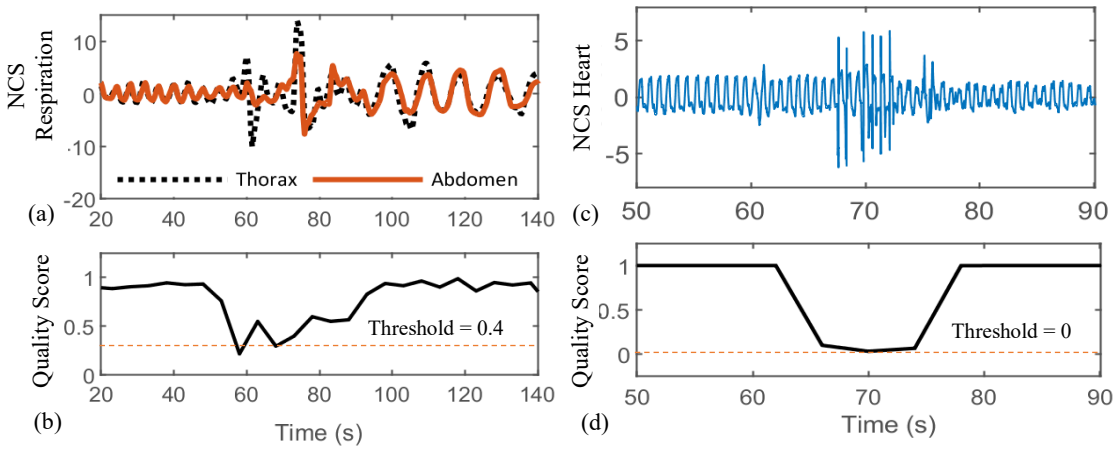


Figure 2.17 Quality score for respiration and heartbeat waveforms with motion artifact instance between 60 – 80 s. A fixed, empirically selected threshold is used to reject poor data, given by a low quality score. (a) Normalized NCS thorax and abdomen respiration waveforms. (b) Respiration quality score in the range [0,1], showing poor quality with score less than the threshold (0.4). (c) Normalized filtered NCS heartbeat waveform modulated on the thorax sensor, with most of the artifact filtered out. (d) Quality score normalized by tanh() in the range [-1, +1]. A low threshold of 0 is selected, as the second harmonic method used for heart rate estimate provides even higher motion tolerance.

duration of 5 s. Final data quality is taken as an absolute of the product of all these correlations, with higher values indicating better quality, as shown in Figs. 2.17 (a) and (b). A fixed threshold is empirically selected, below which the RR and RV estimates are discarded.

- **Heartbeat Data Quality**: As the NCS heartbeat waveform is only extracted from the thorax sensor, an outlier detection algorithm based on a one-class support vector machine (OCSVM) [88] is implemented to detect artifacts in the filtered heartbeat waveform. To speed up signal processing, the motion detection algorithm is implemented on a window of 4 s. The algorithm is trained on the entire routine and a window is marked as an outlier if the tanh() normalized score is less than 0. Figs. 2.17 (c) and (d) show the filtered heartbeat waveform

from the NCS thorax sensor and the corresponding score, respectively. Detailed algorithm and feature extraction are presented in the next chapter. For ECG waveforms, QRS complex peaks are less sensitive to noise, and thus no separate artifact detection is implemented. HR estimation is discarded if the window contained any poor-quality period.

2.4.3 Heart rate estimation

The heartbeat waveform is modulated strongly on the thorax sensor and can be separated from the respiration by proper filtering. However, during the fast breathing exercise, the RR can be around 40 BPM, close to the HR, resulting in some filtering ambiguity. In the time-domain waveform, beat-to-beat peak location may not be accurately detected due to smooth peaks with less sharp features. Therefore, we have proposed using the second harmonic of the heartbeat waveform (Fig. 2.18) to estimate the peak-to-peak heartbeat interval with reduced interferences from respiration and motion. The FT is used to first estimate the approximate HR, which guides the filtering to retrieve the second harmonic with a bandwidth of 0.8 Hz. The same peak detection method is then used to find peaks of the second-harmonic waveform. Instantaneous HR is estimated as an inverse of each heartbeat interval, taken as the sum of two neighboring peak-to-peak periods, as shown in Fig. 2.18 (Right). This HR is averaged over a window (10 s) to suppress any outliers. The reference HR is estimated from the ECG waveform with bandpass filtering between 0.4 – 20 Hz, which removes any baseline drift and suppresses the low-frequency P and T waves. The sharp QRS complex peak is then detected with a simple slope-based peak detection algorithm.

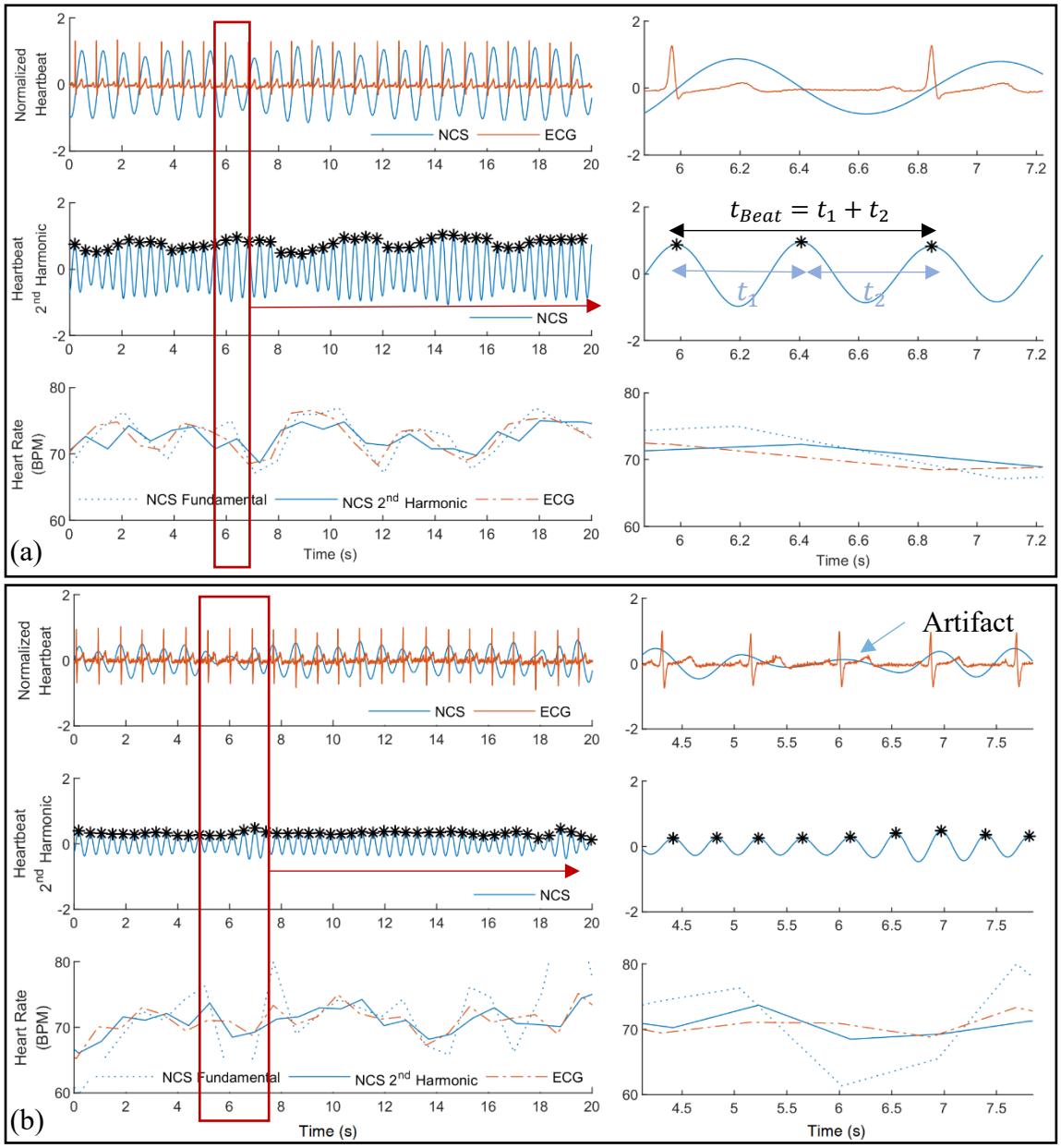


Figure 2.18 NCS heartbeat waveform and HR extraction. *Left:* Band-pass filtered NCS heartbeat waveforms and heart rate extraction from 2nd harmonic, showing instances without motion artifact (a) and with motion artifact (b). (a) Top figure shows ECG and filtered NCS heartbeat. The middle figure shows 2nd harmonic of the NCS heartbeat with detected peaks. The bottom figure shows the instantaneous HR as inverse of each beat-to-beat interval, t_{Beat}, from fundamental NCS (dotted blue), harmonic NCS (solid blue) and ECG (dash-dotted orange). (b) The same HR extraction, showing periods of artifact at $t =$ 6 s and 18 s, where the 2nd harmonic is more accurate. *Right:* Zoomed-in version of the left figures, showing t_{Beat} calculation from NCS 2nd harmonic as the sum of adjacent beat-to-beat intervals.

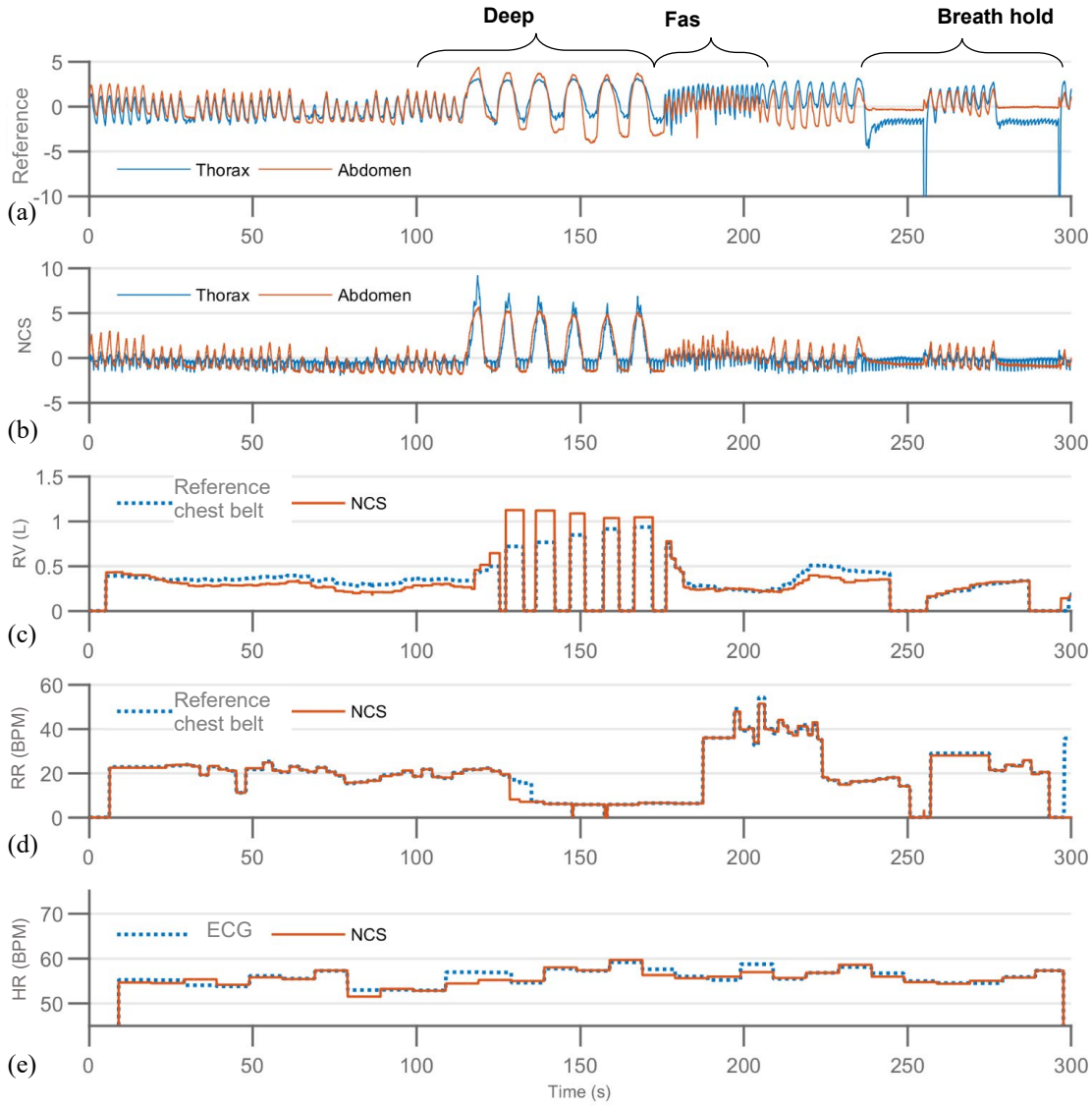


Figure 2.19 Examples of a 5-minute breathing protocol for one participant. The subject is in the supine posture while performing normal, deep, fast breathing and breath-hold. (a) Normalized reference strain chest-belt data at thorax and abdomen. During breath hold the thorax belt also shows a weak heartbeat motion. Different breathing periods are indicated here, with normal breathing in the unmarked sections. (b) Normalized NCS respiration data from thorax and abdomen sensors, as well as strong heartbeat on the thorax waveform. (c) RV during different breathing styles, based on the average volume exchanged in each inhalation and exhalation cycle over the past window containing at least two peaks, thus resulting in 0 estimates when the number of peaks is less than two (slow breathing). (d) RR estimation from NCS and chest belts clearly showing different breathing periods, with the normal RR around 20 BPM. (e) HR estimation from both sensors showing average resting HR in the range of 55 — 60 BPM.

Fig. 2.19 shows NCS and reference waveforms for a subject and the strong heartbeat modulated on the thorax NCS sensor while performing different breathing exercises, especially visible in the breath-hold periods (Fig. 2.19 (b)), with the corresponding average HR from NCS and reference ECG in Fig. 2.19 (e).

2.4.4 Respiratory rate estimation

The NCS signal is band-pass filtered with the cut-off frequencies of 0.05 Hz and 0.8 Hz to derive the respiration waveforms from both the sensors. As the diaphragm and resulting abdomen motion are usually larger, the respiration waveform from the abdomen NCS sensor alone is used to estimate RR. This sensor also has a weaker heartbeat coupling, and thus filtering requirements are less stringent during fast breathing and breath-hold. Similarly, the abdomen chest belt is also used to estimate reference RR with the same algorithm. RR is calculated as the number of detected breath cycles over the maximum window of the past 15 s. For each window, the number of cycles is calculated as the number of inhalation peaks minus 1, and the total time is the interval between first and last inhalation peaks. If no complete cycle is detected during the entire window, the RR is marked as 0. Here inspiration peaks are selected instead of expiration, as they can be more accurately detected due to sharper peaks. Fig. 2.19 (d) shows calculated RR from NCS and reference chest belts over varying normal, deep, and fast breathing and breath-holds, in the range of 0 — 45 BPM.

2.4.5 Respiratory volume estimation

In addition to HR and RR, this work presents an algorithm to extract RV estimation to get complete respiratory effort estimation. Both are related to each other

through a term called minute ventilation, $MV = RR \cdot RV$, that is related to blood carbon dioxide levels. As the peak-to-peak height of the respiration waveform is related to the amount of air inhaled and exhaled, it can be calibrated using a reference sensor to extract the RV. As the characteristic of the reference sensor change, so does the calibration approach as discussed below.

We first calibrated using pre-calibrated respiratory inductance plethysmography (RIP) chest belts, that showed slight waveform saturation for deep breathing. Thus, a quadratic model was proposed to calibrate NCS peak-to-peak height (h),

$$ah^2 + bh + c = RV_{RIP} \tag{2.5}$$

with RV_{RIP} as the pre-calculated respiratory volume derived from RIP chest belts. Results from this estimate are presented in Chapter 3.

To have a more accurate calibration, a pneumotachometer (PTM) was used for calibrating both strain chest belts and NCS. While PTM is a kind of spirometer, we have referred to it as such to state explicitly that the airflow is estimated from the pressure change. The facemask design includes a separator between nose and mouth and is tightly strapped on to minimize any air leakage from the mouth. Subjects are requested to inhale and exhale by mouth only, and they can feel rather uncomfortable during PTM measurements. Thus, PTM is removed from the participant except during the short initial calibration routine of 30 s, with 15 s of each normal and deep breathing. PTM measures the airflow rate waveform, from which the beginning of each inhalation and exhalation point is identified by a simple zero-cross detection algorithm, based on the change of sign before and after the zero-crossing points, and the slope at these points, shown in Fig. 2.20 (a). Integration is performed over each inhalation and exhalation

cycle to get the instantaneous air volume, Vol_{PTM}, without aggregating a bias over time. Least-square fitting is performed for solving the following equation by a trust-region algorithm with bounded constraints of $a, b > 0$

$$a \cdot Resp_{Th} + b \cdot Resp_{Abd} + c = Vol_{PTM}, \tag{2.6}$$

where $Resp_{Th}$ and $Resp_{Abd}$ are thorax and abdomen respiratory waveforms, respectively. Fig. 2.21 (a) shows PTM airflow waveform during the calibration, with the corresponding Vol_{PTM} and calibrated NCS and chest-belt volumes in Fig. 2.21 (b). RV is defined as the volume of the air exchanged during each inhalation and exhalation

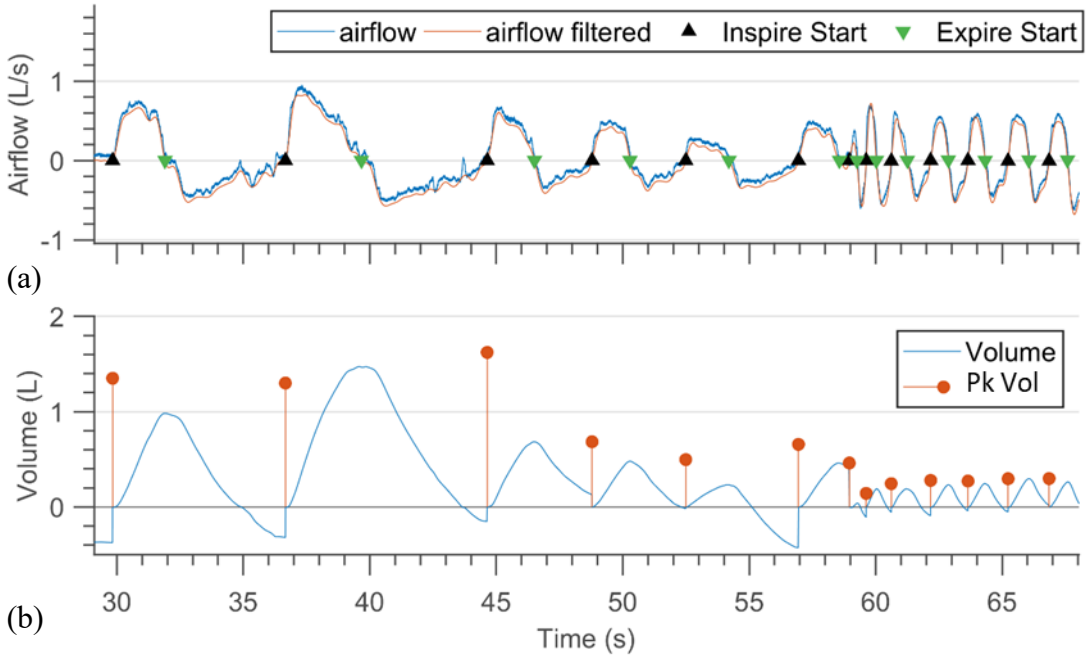


Figure 2.20 PTM airflow waveform and volume estimate. (a) PTM airflow (blue line) in L/s and its lowpass filtered signal (orange line). The detected inspiration and expiration start points are also marked. (b) The volume of air exchanged calculated by integrating over each cycle (between two inspire start points). In some cases, the volume goes negative, indicating either incorrect baseline shift, forced exhalation, or use of nose in addition to mouth. The peak-to-peak height is also calculated.

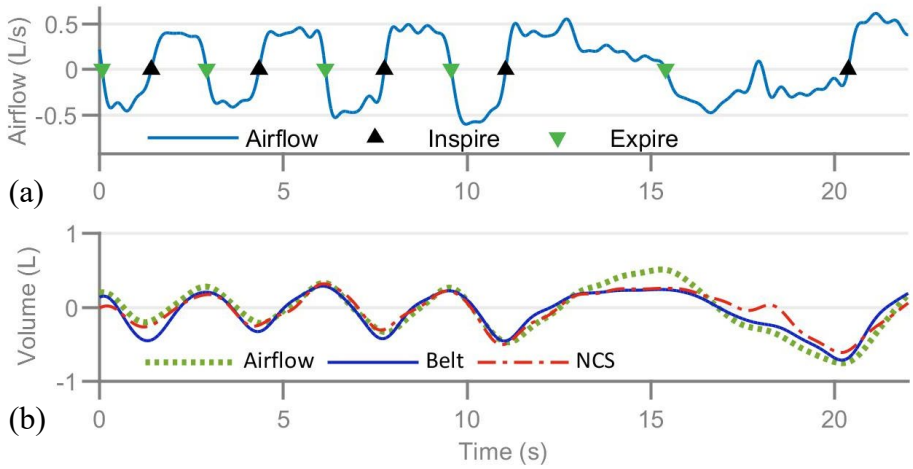


Figure 2.21 PTM airflow and calibrated chest belt and NCS respiration waveforms. (a) Baseline corrected PTM waveform. (b) Instantaneous volume of air exchanged Vol_{PTM} (dotted green) from the airflow, and calibrated chest belts (blue) and NCS (dashed-dotted red) respiration waveforms.

and is calculated as the peak height of each calibrated breath cycle as shown in Fig. 2.20 (b). Fig. 2.19 (c) showed the estimated RV from chest belts and NCS data over different breathing styles. Deep breathing clearly shows an increased volume of air exchange compared to normal breathing. RV is estimated as the average volume exchanged in each respiratory cycle over a window of 15 s, with at least two peaks in the window to get a robust estimate.

Due to sensor motion over time, mostly resulting from a posture change, the estimated RV may deviate from its original calibration. An additional corrective step is performed if the estimated RV during the normal breathing section of the breathing exercise routine $RV(Normal)_{Routine}$, deviates more than ± 5% from the calibrated $RV(Normal)_{Calibration}$, derived from the first 15 s of calibration. The NCS and reference RV is multiplied with a corresponding scaling factor (SF) given as

$RV(Normal)_{Calibration}/RV(Normal)_{Routine}$, assuming the normal breathing routine has nearly constant RV for the same person over time.

To validate this correction, an additional RV calibration consistency study was conducted on one subject over three consecutive days in the sitting posture. The subject performed the calibration, immediately followed by the voluntary breathing exercise routine. Table 2.1 shows the detailed results, with RV before scaling correction, and the required SF to improve the match when using day-1 calibration for all successive days. The $RV(Normal)_{Calibration}$ lies within a narrow range of 0.37 – 0.43 L, showing consistency in normal breathing volume for a healthy individual over time. Little scaling is required to correct RV when calibrated immediately before the routine, resulting in scaling factors close to 1 on day-1. Variations in sensor placement and coupling strength result in scaling factors other than 1 to correct the RV for the next two days. Figs. 2.22 (a) and 2.22 (b) show instantaneous volume and RV estimated on day-2 using day-1 calibration, with NCS underestimating and reference chest belt overestimating the expected $RV(Normal)_{Calibration}$ = 0.37 L, derived from day-1. The corrected scaled RV in Fig. 2.22 (c) shows improved agreement of NCS and reference to each other, as

Table 2.1 RV calibration consistency test over 3 consecutive days

1. Day $\boldsymbol{i}$	2. $\text{RV(Normal)}_{\text{Calibration}}$	3. $\text{RV(Normal)}_{\text{Routine}}$ (Calibration: Day 1)	4. Scaling Factor, SF (Calibration: Day 1)
$\boldsymbol{i = 1}$	$\text{RV}_{PTM} = 0.37\ L$	$\text{RV}_{NCS} = 0.39\ L$, $\text{RV}_{BIOPAC} = 0.35\ L$	$SF_{NCS} = 0.95, SF_{BIOPAC} = 1.06$
$\boldsymbol{i = 2}$	$\text{RV}_{PTM} = 0.43\ L$	$\text{RV}_{NCS} = 0.22\ L$, $\text{RV}_{BIOPAC} = 0.65\ L$	$SF_{NCS} = 1.68, SF_{BIOPAC} = 0.57$
$\boldsymbol{i = 3}$	$\text{RV}_{PTM} = 0.43\ L$	$\text{RV}_{NCS} = 0.24\ L$, $\text{RV}_{BIOPAC} = 0.60\ L$	$SF_{NCS} = 1.54, SF_{BIOPAC} = 0.62$

Table displaying the normal breathing RV calibration and scaling factor consistency over three days of testing on the same subject. Each day consisted of a calibration period immediately followed by a breathing routine, performed while sitting upright. **1.** The day of the measurements. **2.** The normal RV estimated from the PTM during the calibration routine. All three RV estimates are close to each other, within a range of 0.06 L. **3.** The normal breathing RV for NCS and BIOPAC during the breathing exercise routine, derived using calibration from day 1 for all three days. **4.** The calculated scaling factor to accurately map the routine's normal RV to normal RV of day 1 calibration. Little scaling is required to correct RV when calibrated immediately before the routine, resulting in scaling factors close to 1 on day-1. Variations in sensor placement and coupling strength result in scaling factors other than 1 to correct the normal RV on the next two days.

well as to the calibration value. As SF assumes accurate $RV(Normal)_{Calibration}$, some error may be introduced from inconsistencies in the calibration and correction process, including variation due to small calibration duration, difference in resistance when breathing through the mouth, and change in $RV(Normal)$ over time. The correction assumes knowledge of normal breathing period, that can be potentially identified based on the normal RR and HR range of an individual.

2.5 Motion, coughing, and speaking interferences

This section discusses how signal patterns from the body and ambient motion, cough, and speaking are observable in the respiratory and cardiac waveforms and their detection.

2.5.1 Motion detection during sleep with passive setup

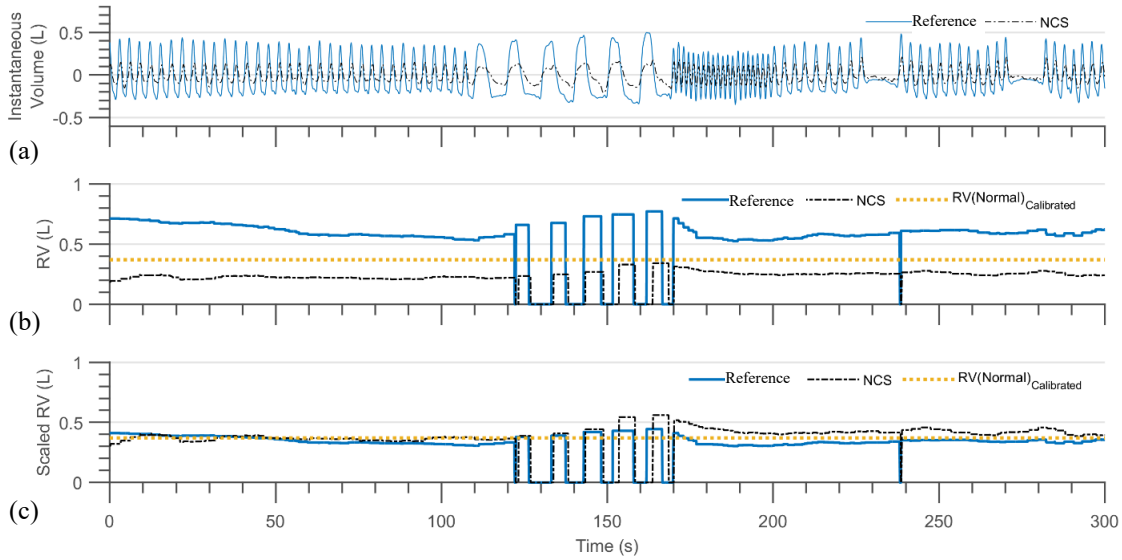


Figure 2.22 Results of calibration consistency test. Using RV calibration from day-1 on day-2 data, both collected in sitting posture. Subject is performing voluntary breathing exercises in the routine. (a) Calibrated instantaneous volume from reference chest belts and NCS. (b) Estimated RV from the top plot with chest belts overestimating and NCS underestimating the expected $\text{RV(Normal)}_{\text{Calibration}} = 0.37$ L. (c) Scaled RV estimation that corrects the error based on a normal breathing section.

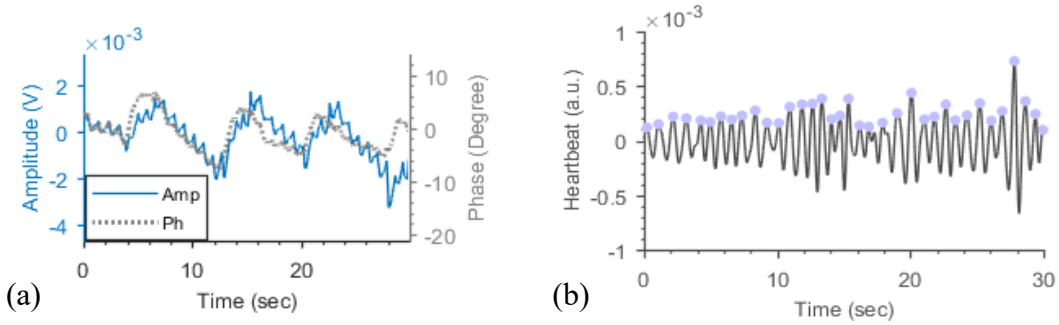


Figure 2.23 Heartbeat extraction from NCS using wavelet. (a) Raw NCS amplitude and phase waveforms, with heartbeat clearly indicated in amplitude. (b) Heartbeat detection from the wavelet d8 coefficient reconstruction.

Motion detection for sleep scoring with accurate heartbeat and respiration information cannot be directly achieved by simple filtering. Motion affects waveform features of both heartbeat and respiration. In this work, a beat-by-beat segmentation is performed to achieve finer time resolution of motion detection. Accurate peak detection is difficult in non-stationary heartbeat waveforms with multiple peaks depending on sensor placement, operating carrier frequency, and filtering range. As RR in sleep is limited, and not expected to be as high as panting, an approach based on multilevel 1-D wavelet decomposition using Daubechies db10 wavelet is used. Reconstructed waveforms by the detail coefficient at level 8 (NCS-d8), containing the main component of heartbeat waveforms are shown in Fig. 2.23. The wavelet coefficient level can be selected by adjusting the signal sampling rate, f_s and selecting the level containing 50 – 100 beats per minute (BPM) frequency, or simply by selecting the coefficient resulting in maximum correlation with the filtered heartbeat waveform. Peak detection can be performed at this stage to get individual heartbeats. This method is different than earlier HR estimation, as while that was designed to be motion tolerant, here the aim is two-fold: 1) to be sensitive to even the smallest motion, like a slight jerk, of arms or even legs which can be indicative of periodic limb movement disorder (PLMD) [89],

observed only during sleep and is an indicator of sleep quality, and 2) to identify motion affected beats and reject them.

A controlled data collection is performed imitating the following scenarios during sleep:

1) Stationary state: Subject intentionally remaining still.
2) Torso motion: Slight torso motion for 5 seconds.
3) Body jerk: A fast, high-energy motion of torso and arms for 0.5-2 seconds.
4) Turning: Subject turns to their left or right in 2-3 seconds.

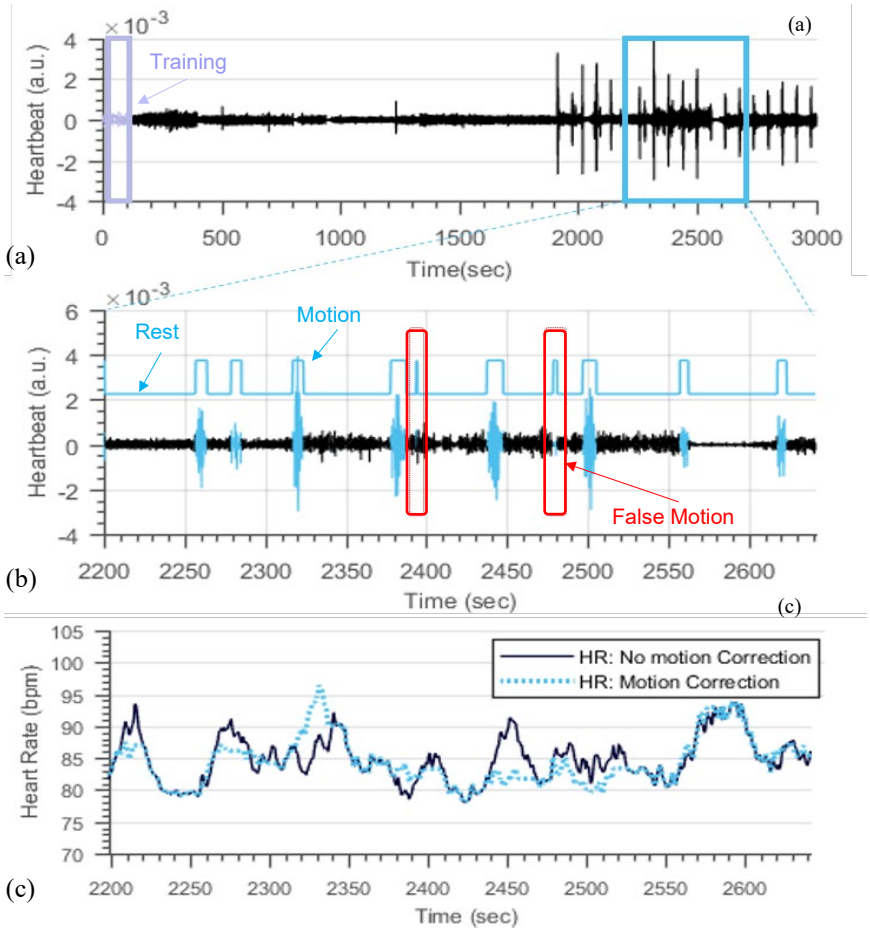


Figure 2.24 (a) NCS heartbeat waveforms with the training period shown in the initial two minutes. (b) Motion prediction for a section of NCS with false positive cases. (c) Motion corrected HR estimation.

Data were collected for 3 hours with the NCS signal sampling rate of 500 Sps. Simulated movements are performed at intervals of 1 minute.

Motion features are identified based on the difference between motion-affected waveforms and those obtained at rest. Also, features need to be robust to account for variation in RR, HR, and signal amplitude over time. Relative beat interval and relative beat root mean square (RMS) are features based on the assumption that beat interval and RMS are not expected to vary dramatically from beat to beat during rest and sleep. Statistical mean, variance, skewness, and kurtosis are calculated to capture major differences between waveforms at motion and rest, where a heuristic window of 5 beats is applied. Normalized spectral power in the range of 0.6-10 Hz over the same window is used as a feature to observe spectral changes resulting from motion.

SVM with the radial basis function kernel has been adopted to detect motion using the above 7 features. Semi-supervised learning is used for training the model, i.e., training is performed with data collected only at rest with spontaneous breathing. Exclusion of motion data for training allows adaptability of the classifier to other movements indicative of restlessness while avoiding overfitting, as well as reduces the inconvenience and subjective variation to perform movement routines during training.

Table 2.2 Motion Detection for Each Context

Context	**Performance Metric**	
	Detected	Missed
Slight Torso Motion	32	3
Body Jerk	48	1
Turning	32	7
Total	112	11

Table 2.3 Beat-by-Beat Motion Detection

		Actual	
		Motion	Rest
Predicted	Motion	648	252
	Rest	86	12985

Fig. 2.24 shows NCS heartbeat and corresponding motion detection. The performance of the motion detection algorithm was evaluated by manually annotated labels for each beat.

Table 2.2 shows the number of correctly detected and missed cases for each category of motion. This shows accurate motion detection in 91.06% of the cases. The finer, beat-by-beat motion classification results are arranged in a confusion matrix in Table 2.3. This shows the number of beats corresponding to each category and gives an accuracy of 97.58%, sensitivity of 88.28%, and specificity of 98.10% for the proposed algorithm. Fig. 2.24 (c) shows an example of the improved HR estimation after removing the motion artifact. The HR is estimated with a moving average of 30-beat window length, by discarding motion-affected beats.

Body jerk motion can be detected with good accuracy, while slow-turning can be more often misclassified as rest. The accuracy will depend on the training data and related algorithms. We have trained with spontaneous breathing data at rest, including regular breathing with occasional deep breaths. Training with a regular breathing pattern increases the sensitivity to motion but also results in increased false positive (FP) motion

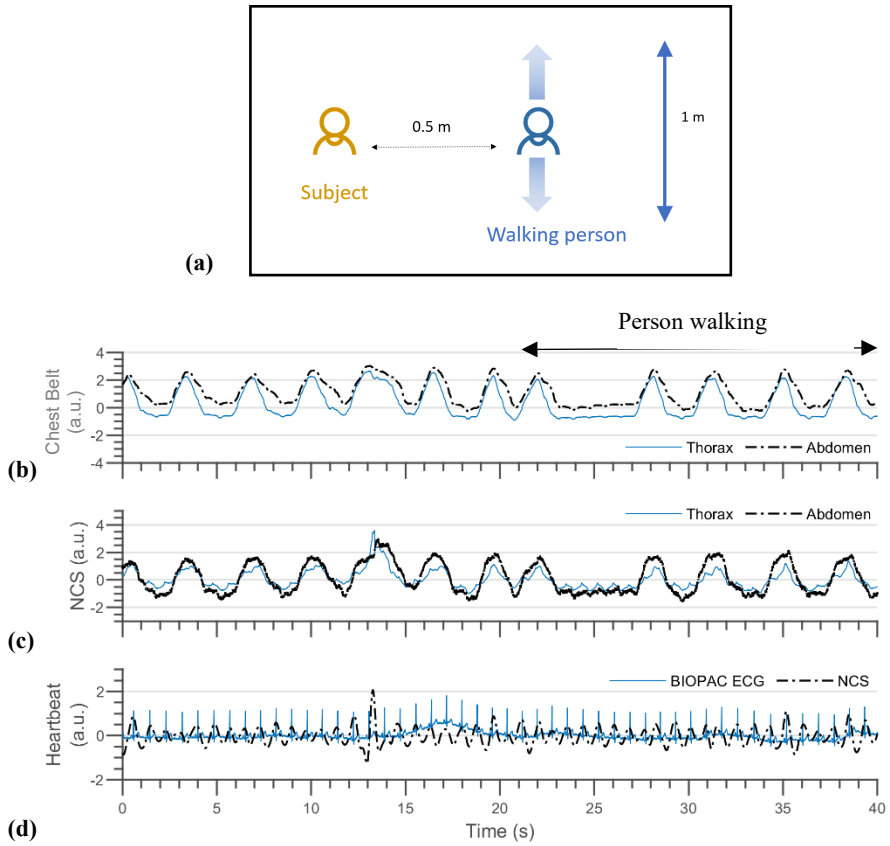


Figure 2.25 The effect of ambient motion on reference (chest belts and ECG) and NCS signals. Subject is seated at rest with a person standing nearby at a distance of 0.5 m. At $t = 20\ s$, the person starts walking forward and backward at a speed of 0.73 m/s as shown in (a). (b) Normalized chest belt waveforms showing clear respiratory signal with no motion interference. (c) Normalized NCS waveforms showing both respiration and clear heartbeat without any motion interference. (d) Clear heartbeat signal is observed from both ECG and NCS without any interference.

in case of irregular breathing. This also presents a further opportunity for respiratory disorder identification.

2.5.2 Motion interference with a wearable setup

Our wearable active NCS sensor setup has a high tolerance for ambient motion interference. As shown in Fig. 2.25, there is no interference from a nearby walking person on the NCS data, due to the strong near-field coupling. Additional tests are performed with the motion of different body parts during breath holds, as shown in Figs.

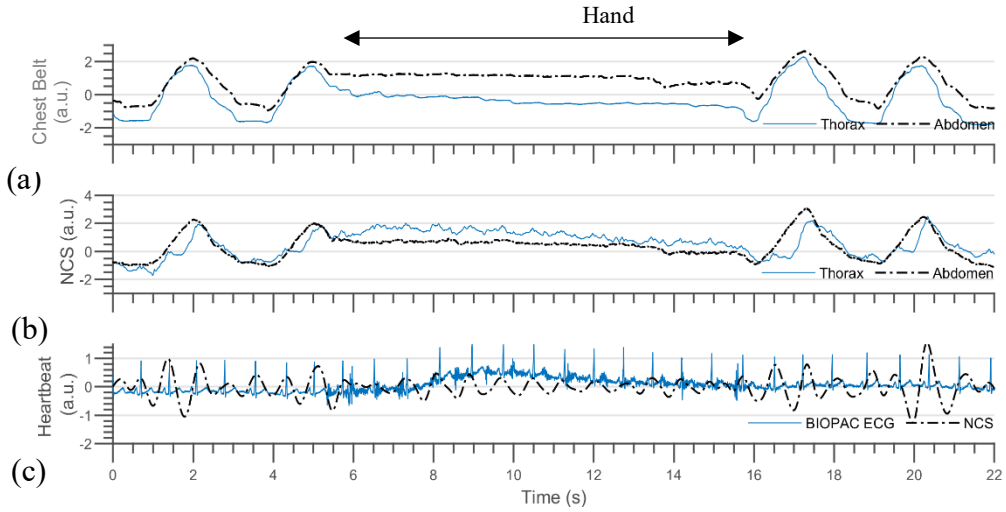


Figure 2.26 The effect of hand motion on reference (chest belts and ECG) and NCS signals. The subject is seated at rest with hands resting on thighs, performing normal breathing for the first 5 s. For the next 10 s, repeated instances of fist opening and closing are performed with the right hand, while holding breath. (a) Normalized chest-belt waveforms showing clear respiratory signals with no motion artifact. (b) Normalized NCS waveforms showing respiratory motion, with minimal motion artifact. (c) Heartbeat motion can be correctly extracted from NCS after further processing. Some interference is observed in the ECG waveform during the hand motion, but the R peaks can still be clearly seen.

2.26 and 2.27. There is no effect of hand motion on the chest belts, nor the NCS sensors unless it is very close to the antenna pair. ECG becomes noisy with the hand motion, likely due to electromyogram (EMG) interference. During arm motion, both thorax and abdomen belts show signal anomalies, while only the thorax NCS is affected. Additional studies are required to examine internal muscle motion interference resulting from near-field coupling and separate the external arm motion coupling, as abdomen NCS still shows a clear signal without any interference.

While external motion causes limited interference, large spurious torso motion results in motion artifacts. Respiratory measurements from chest belts can be interfered with belt displacement due to body movement, and an entire signal segment can be missed if the participant performs heavy or forced exhalation that leads to loss of

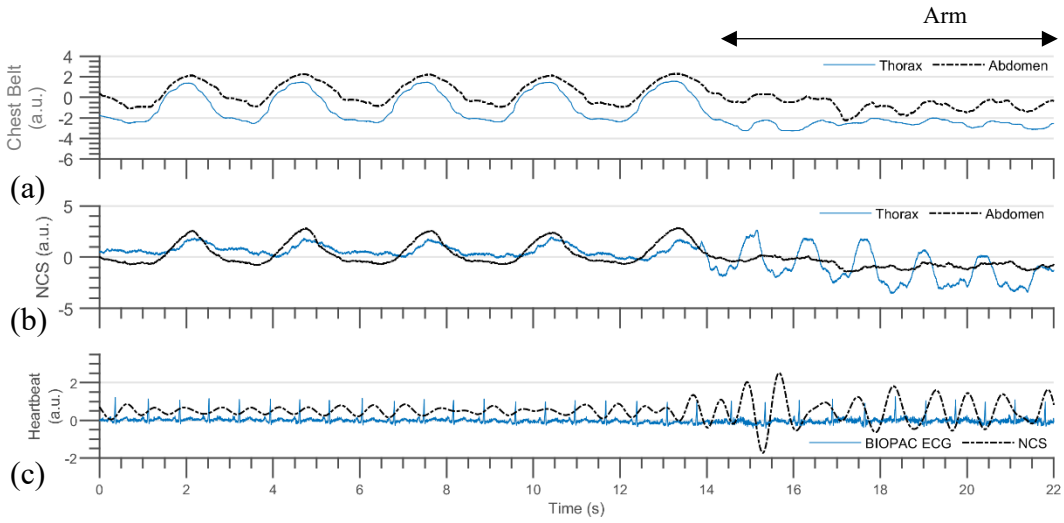


Figure 2.27 The effect of arm motion on reference (chest belts and ECG) and NCS signals. The subject is seated at rest with horizontal forearms in front of the body and upper arms in the vertical position. A forward arm swing of up to 30° is performed with both arms, starting at $t = 14\ s$, while holding breath and keeping forearms horizontal. (a) Normalized chest-belt waveforms showing clear respiratory signals and some motion interference during the breath hold period with arm motion. (b) Normalized NCS waveforms showing respiratory motion without any interference to the abdomen sensor, but arm motion is captured in the thorax sensor. (c) The heartbeat signal from NCS is interfered due to the arm swing and HR can only be possibly extracted from the harmonic.

tension. NCS waveforms are prone to motion artifact due to the current antenna pair packaging, which is loosely placed over the clothes with semi-isotropic radiation patterns. Free antenna movement relative to the body surface may result in signal artifacts, making NCS well-suited for static and quasi-static conditions, for example, when a person is typing, driving, or sleeping. Further, our active setup has slightly better coupling, which may lead to signal saturation issues as discussed in Sec. 2.4.2, and requires signal quality detection for a larger bandwidth including both respiration and heartbeat signals. This improved coupling is a result of sensor design and placement selection, which has a tradeoff of the possible saturation with improved heartbeat SNR, as well as a tradeoff of comfortable back sensing with improved coupling from the front.

2.5.3 Coughing and speaking

As the NCS sensor can capture the near-field motion at a high sampling rate, we can easily capture signatures of other common breathing conditions such as coughing and sneezing. We performed some preliminary data collection to observe voluntary coughing patterns with one sensor below the xiphoid process. Fig. 2.28 (a) shows the low-frequency (LF) NCS signal where regular breathing is mixed with irregular coughing signatures. Cough instances can be more clearly detected from the filtered high-frequency (HF) component, as shown in Fig. 2.28 (b), where the NCS waveform is filtered between 15 – 50 Hz. A cough was predicted if the waveform energy increased above a certain threshold, empirically selected to be 100 times the energy of the HF component during regular breathing. Results of this cough indicator are shown in Fig. 2.28 (b) as well. Multiple sensors and their positions can be optimized in the future to

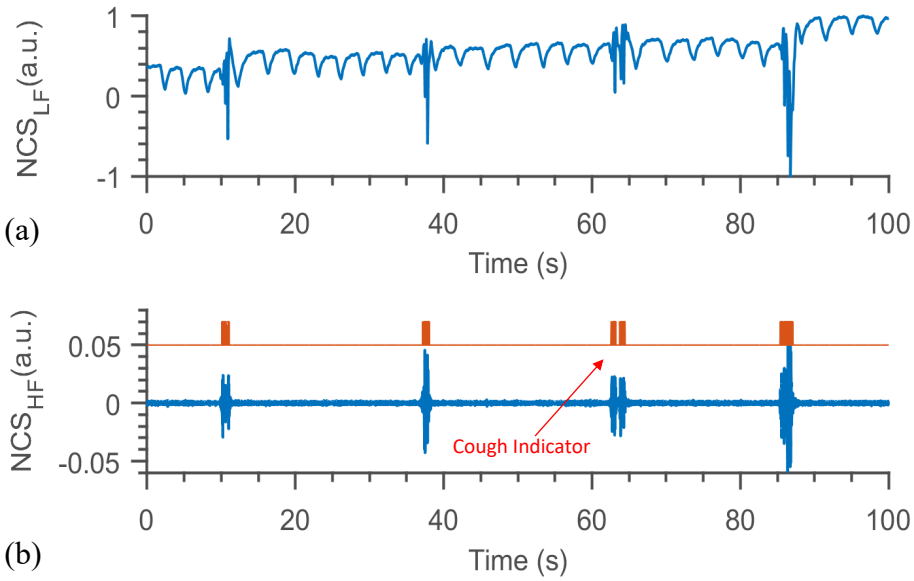


Figure 2.28 NCS waveform showing coughing instances at t = 20, 45, 68 and 85 s. (a) NCS LF component between 0.02 – 10 Hz, showing breathing and coughing movement. (b) Only HF component is shown with frequencies between 15 – 50 Hz, along with cough indicator based on detecting high energy instances.

possibly recognize coughing signatures at different respiratory tracts of the pharynx, larynx, trachea, and bronchi.

Speaking involves irregular controlled breathing patterns that can vary with culture, language, and individual gender and age differences [90], [91]. As it does not provide useful information in terms of RV, RR due to variation in patterns, speaking instances are detected as interference with poor signal quality indicator, and rejected for further analysis. Participants were asked to read the famous Gettysburg address. Fig. 2.29 shows a one-minute waveform of a participant reading the following excerpt:

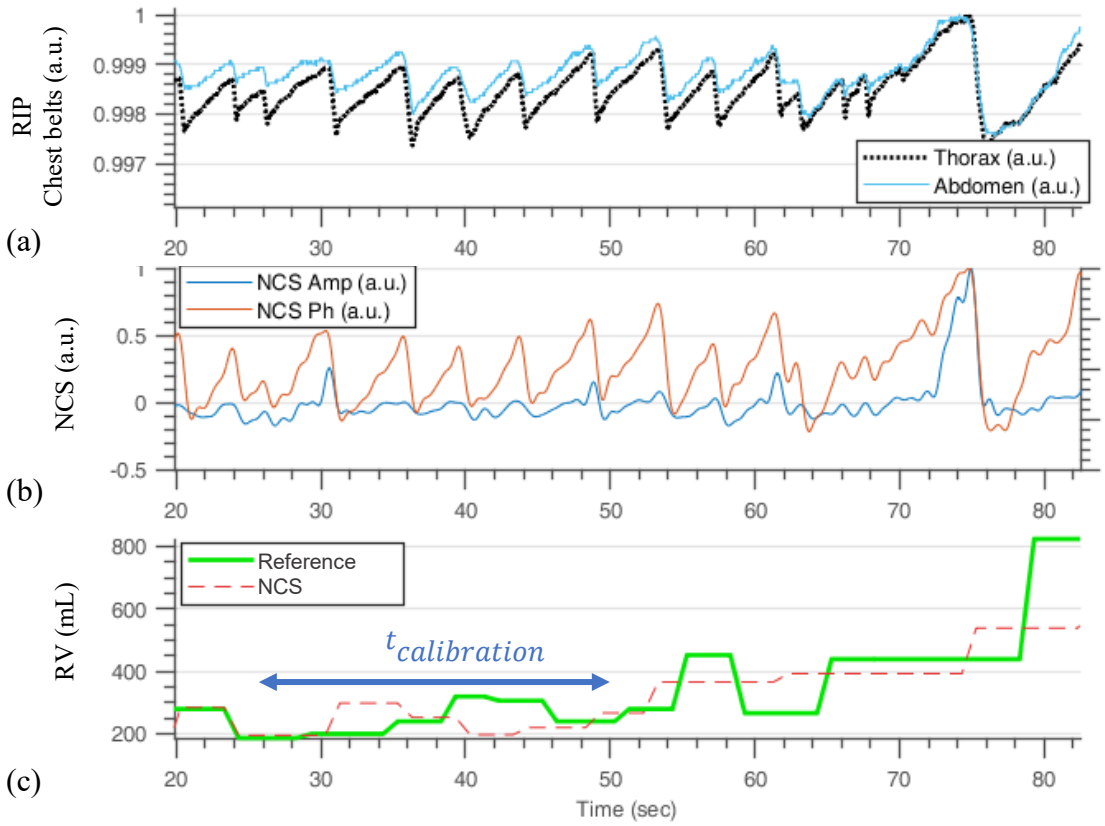


Figure 2.29 Waveform during speaking. (a) Reference RIP chest belt waveforms at thorax and abdomen, both normalized independently. (b) NCS amplitude and phase coupling, showing clear varying respiration patterns in phase. (c) The peak-to-peak RV estimated from both the reference chest belts and NCS showing varying patterns, and difficulty in accurate RV estimate for both the sensors. For example, reference misses the decreased RV around 67 s.

Four score and seven years ago our fathers brought forth on this continent, a new nation, conceived in Liberty, and dedicated to the proposition that all men are created equal. Now we are engaged in a great civil war, testing whether that nation, or any nation so conceived and so dedicated, can long endure. We are met on a great battle-field of that war. We have come to dedicate a portion of that field, as a final resting place for those who here gave their lives that that nation might live. It is altogether fitting and proper that we should do this. - But, in a larger sense, we cannot dedicate—we cannot consecrate—we cannot hallow—this ground.

2.6 Conclusion

In this chapter, we presented the sensor hardware setup and placement considerations. As signal can have various interferences, signal processing techniques including filtering and signal quality detection are also presented. Signal polarity correction algorithm, specific to NCS respiration signal is also presented. The HR, RR, and RV estimation algorithms are presented, with a detailed focus on RV calibration and its consistency over time. These parameters are useful for long-term health monitoring for both healthy users as well as patients. Finally, interference patterns due to motion, coughing, and speaking are discussed, which shows increased motion tolerance with the active wearable setup, and an approach for utilizing passive setup to give an indicator of sleep quality by detecting different motion instances. Later chapters present detailed results in simulated disorders compared to reference sensors and validation studies on a larger set of participants.

CHAPTER 3

VALIDATION OF NCS RESPIRATORY PATTERNS WITH HUMAN STUDIES

3.1 Breathing pattern variation study

Non-invasive respiration monitoring has received a lot of interest in recent years [92], which has led to many research works on a robust, accurate, comfortable, and affordable sensor that can provide the two key respiratory parameters of respiration rate and volume. As respiratory failure is often difficult to predict, such sensor can be even

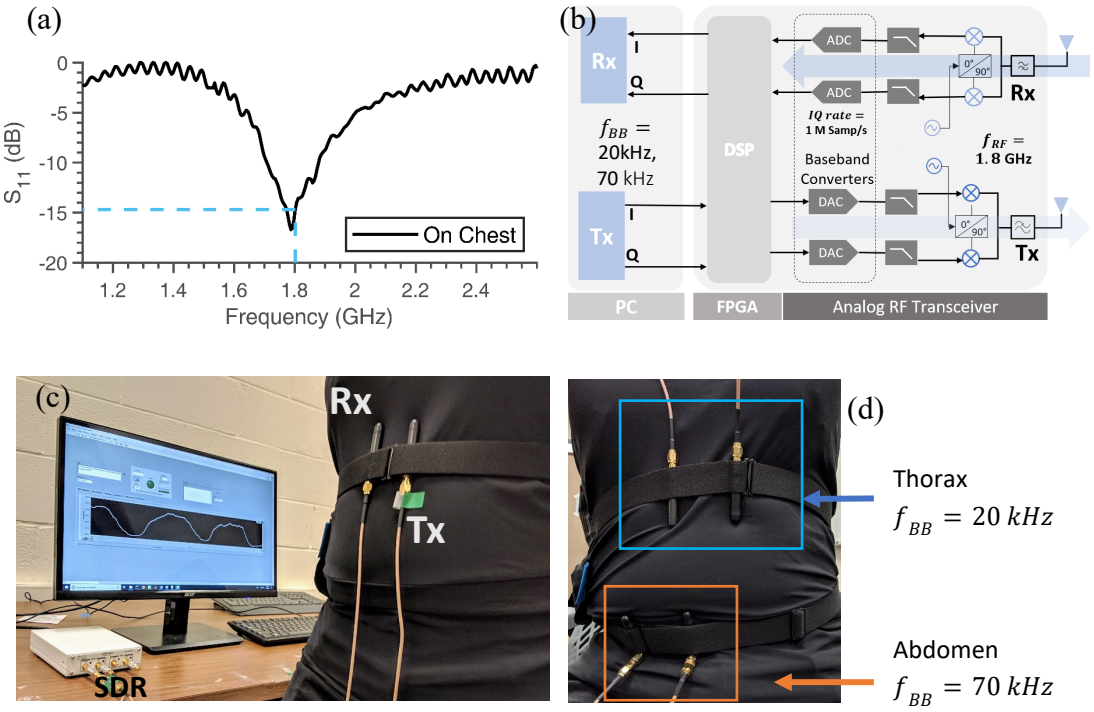


Figure 3.1 Setup for breathing pattern variation study. (a) Measured antenna S_{11} with chest placement showing $S_{11} = -15$ dB at 1.8 GHz. (b) SDR architecture showing one of the Tx – Rx paths. Two Tx and two Rx share an LO. (c) Placement of Tx and Rx antennas below the xiphoid, together with the chest-belt Hexoskin for comparison. SDR and LabVIEW for signal demodulation and analysis are also illustrated. (d) Setup showing two Tx – Rx antenna pairs to observe movement at thorax and abdomen. The person is wearing a Hexoskin smart shirt, and the sensors are placed close to the body belts at thorax and abdomen.

more useful for long-term continuous on-line measurements on patients pre-diagnosed with respiratory disorders, where the risk of mortality can be effectively reduced [93]. While many mobile sensors for exercise and wellness monitoring aim to estimate RR only, additional information on RV provides richer information for respiratory health diagnostics. In this study, an active NCS sensor is deployed to monitor respiratory patterns and estimate RR and RV for various simulated and real disorders [43].

3.1.1 Experiment setup and data collection

The sensor setup is implemented with the COTS component as described in the previous chapter. An Ettus USRP B210 with UHF monopole antennas is used with real-time visualization in NI LabVIEW. With both the antennas placed on the chest, the return loss, S_{11} of the Tx antenna is shown in Fig. 3.1 (a), with an operating frequency of 1.8 GHz, close to the minimum S_{11} point. Two baseband frequencies, f_{BB} of 20 kHz and 70 kHz are used for multiplexing signals from two sensing points. Fig. 3.1 (b) shows the SDR architecture with f_{BB} and f_{RF} selection. As a single sensor is more convenient, one active Tx-Rx pair is placed below the xiphoid process to monitor the respiratory patterns for most disorders. For OSA, a two-sensor placement is used to measure out-of-phase thorax and abdomen motion. From the quadrature baseband, the NCS signal here is taken as the modulated amplitude, with a bandpass filtering to remove the DC level and any small heartbeat component.

To calibrate our NCS measurements, we have used a Hexoskin smart shirt [29] that monitors respiration by RIP chest belts as a reference. Wires incorporated into fabric elastic tension bands are placed around the thorax and abdomen. The lungs expand and contract during respiration, which is observed as a change in the self-

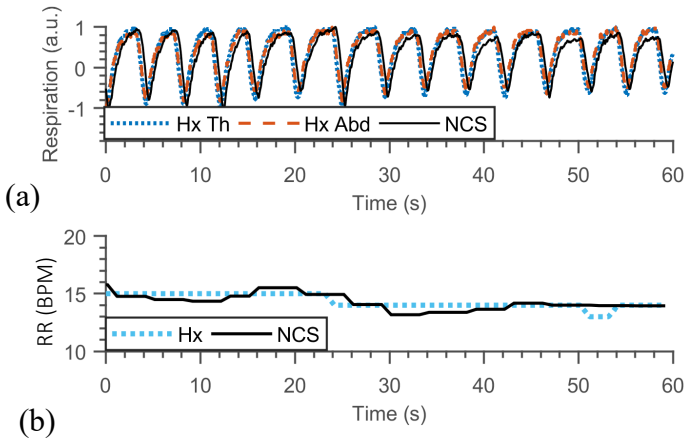


Figure 3.2 Respiration during normal breathing. (a) Normalized respiratory waveforms from Hexoskin thorax (Hx Th) and abdomen (Hx Abd) body belts and NCS. (b) Nearly constant RR estimated from Hexoskin and NCS in the normal breathing range of 12 – 20 BPM.

inductance of the coils used to sample the respiratory waveform at 128 Hz. RR and RV are then estimated at 1 Hz. The NCS waveforms are sampled at 500 Hz, and RR and RV are also calculated at 1 Hz for easy comparison. Fig. 3.2 (a) shows an example of the NCS respiration waveform with the sensor placed below the xiphoid process, and the thorax and abdomen waveforms retrieved from reference (Ref) during normal spontaneous breathing. The corresponding RR from NCS denoted as RR_{NCS}, and from reference Hexoskin as RR_{Ref} is shown in Fig. 3.2 (b) in BPM. RV under this normal breathing condition is often denoted as the tidal volume (TV).

We validated NCS respiration waveforms and parameters in simulated conditions of 1) normal, 2) deep, 3) slow-shallow and 4) fast-shallow breathing. Synchronous Hexoskin respiration data were collected. Additionally, we performed RR and RV analyses on several simulated abnormal patterns including, 1) CSA, 2) Cheyne-Stokes, 3) Biot's, and 4) ataxic breathing [94]. We also performed data collection with two NCS sensors, one at the abdomen and one at the thorax, for isovolumetric abdomen

movement. Repeated abdomen contraction and relaxation movements were executed with breath-holding, i.e., neither inhalation nor exhalation. As the total RV is conserved, this leads to the out-of-phase movement of the thorax and abdomen, similar to breathing efforts during airway collapse. Finally, we explored the coughing recognition from extracted NCS waveforms. All of the procedures involving human subjects were approved by the Cornell Institutional Review Board (IRB).

3.1.2 Normal, deep, and shallow respiration

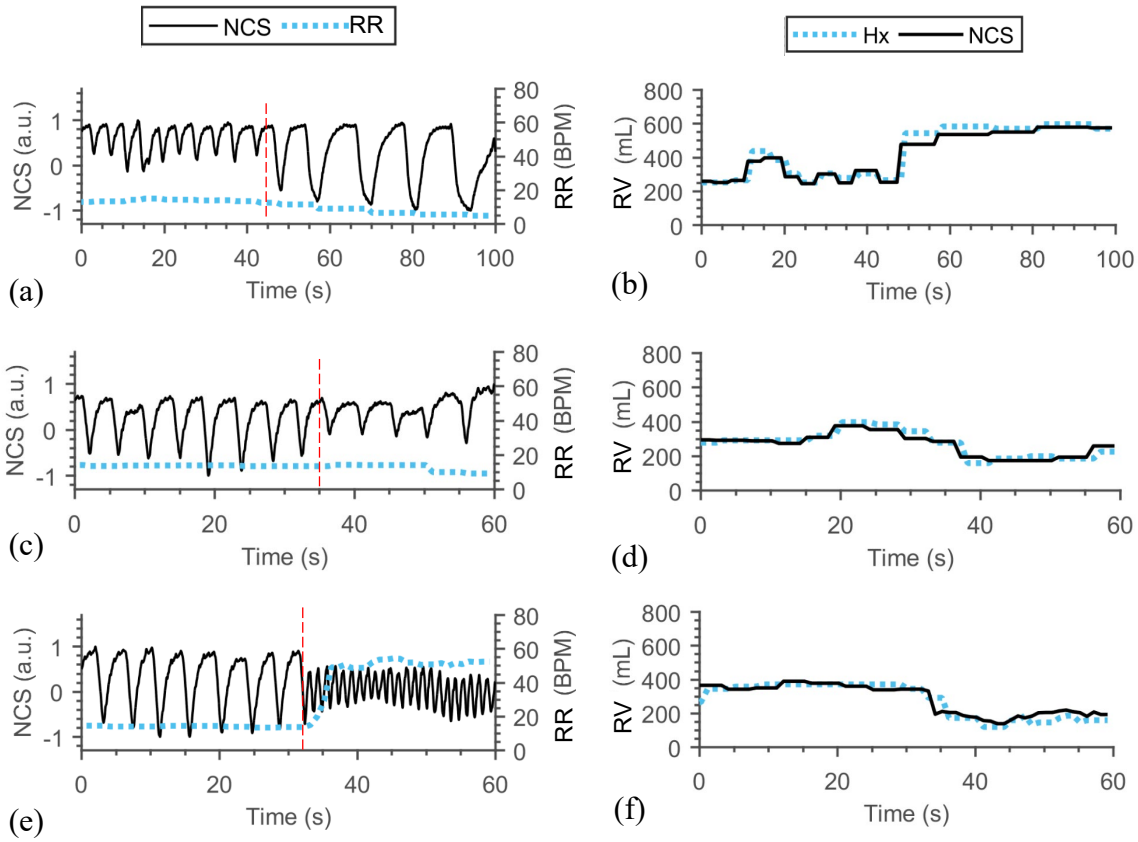


Figure 3.3 The top row shows normalized NCS respiration waveforms in solid black lines with transitions from normal to different breathing conditions, indicated by vertical dashed red lines. Estimated RR from NCS is also shown in BPM, as dotted blue lines. RR obtained from Hexoskin is very close in this set of experiments, and is hence not shown. The bottom row shows the corresponding RV estimate from Hexoskin (dotted blue lines) and NCS (solid black lines). (a) and (b) show transition from normal to deep breathing at $t = 45$ s. (c) and (d) show normal to slow-shallow breathing transition at $t = 35$ s, accompanied with decreased RV. (e) and (f) show transition to rapid shallow breathing at $t = 32$ s, marked with lower RV and higher RR.

We performed data collection showing transitions among normal, deep, and shallow breathing conditions. The normal RR for adults during rest is in the range of 12−20 BPM. The normal breathing patterns extracted from NCS and reference are shown in Fig. 3.2 (a). All respiration waveforms are normalized. Fig. 3.3 (a) shows the transition between normal and deep breathing conditions. RR estimated from NCS is also shown, and is very close to that from reference, which is not shown for graphical clarity. The corresponding RV estimates are shown in Fig. 3.3 (b). Two shallow breathing conditions were also considered, one with normal or slow RR, and the other with high RR. Slow-shallow breathing leads to low RV estimates and is one of the possible indicators of respiratory depression induced by opioid abuse. Fig. 3.3 (c) shows a transition from normal to shallow breathing conditions with similar RR, marked by a sharp decrease in RV. On the other hand, the breathing pattern of rapid shallow breathing in Fig. 3.3 (e) is marked with increased RR and reduction of RV in comparison with the normal breathing conditions. This pattern can be triggered by multiple scenarios, including anxiety attacks, asthma, and COPD. All transitions are marked with vertical dashed red lines.

3.1.3 Simulated respiratory disorder patterns

Central sleep apnea

CSA is a sleep disorder, marked by the cessation of airflow without any muscle activity [15]. This breathing disorder is frequently observed among people suffering from chronic heart failure and brain diseases, among other reasons, as well as can be induced by opiate uses.

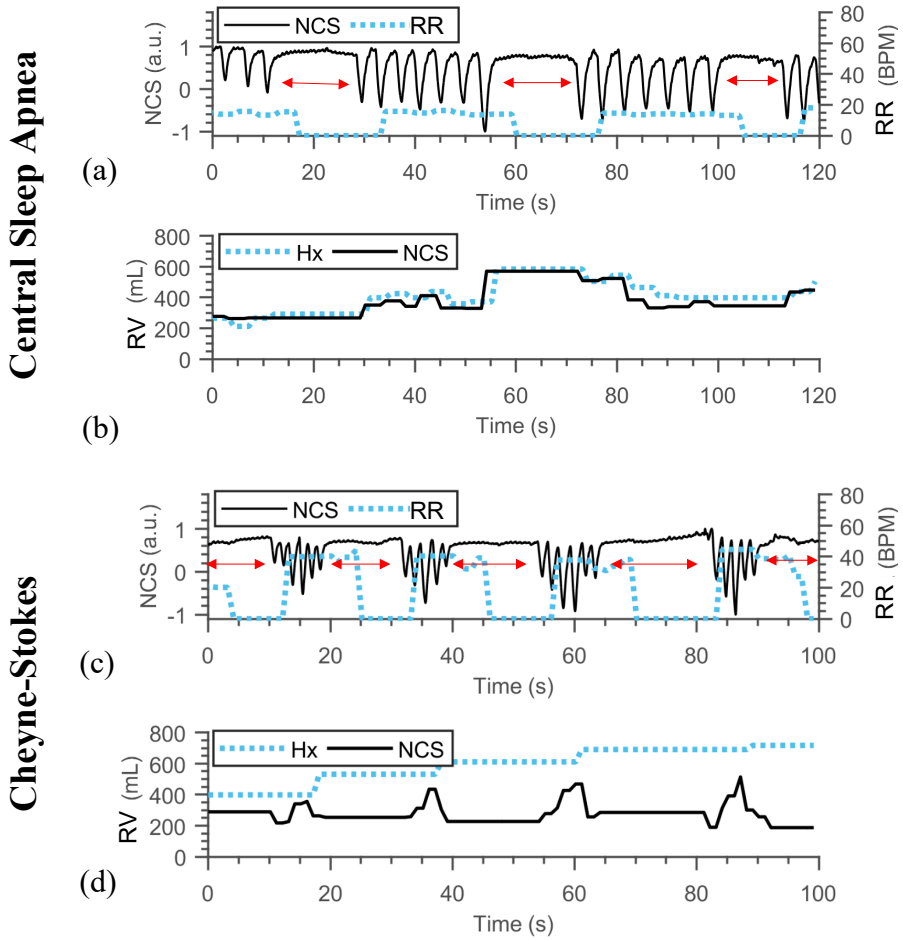


Figure 3.4 Respiration waveforms in CSA (a)-(b) and Cheyne Stokes (c)-(d). (a) NCS respiration waveforms with CSA, showing normal breathing with pauses between 12 – 28 s, 56 – 70 s and 102 – 112 s, indicated by the red arrows. Notice that although the NCS waveform indicates the breath pause timing instantaneously, RR drops to nearly zero only after a delay due the moving sampling window of 8 s. (b) RV estimates by calibrated NCS and Hexoskin. (c) The NCS respiration waveform and RR for the Cheyne-Stokes breathing pattern, separated by breath pauses marked with the red arrows. (d) The corresponding RV by Hexoskin and NCS, where Hexoskin cannot correctly capture the rise and fall pattern in RV.

Fig. 3.4 (a) shows the respiratory pattern with sections of normal breathing with in-between voluntary breath-holding marked by the red arrows, lasting around 10 – 15s. Although the NCS waveform shows the instantaneous transition to a nearly constant value, RR drops to zero only after a delay, as it is calculated by counting the number of complete breath cycles in the last 8s window. The seemingly delayed drop in RR is an

artifact from the window size selection; a large window will have a larger delay appearance while a small window may miss a slow cycle and report zero RR. NCS can output the accurate time registry for the endpoints of inhalation and exhalation and can thus be used for breathing cycle parameters other than RR. Nevertheless, the CSA pattern of interest here can still be derived from the present RR curve. The corresponding RV is shown in Fig. 3.4 (b), which is only updated when a breath motion is observed. As there is a complete cessation of breathing, RV is not updated until the next breath cycle is detected.

Cheyne-Stokes respiration

This is a variant of CSA, with the periodic shallow breathing pattern alternating with deep over-breathing and apnea, which leads to a crescendo-decrescendo waveform accompanied by the apnea state. Cheyne-Stokes breathing is found in the same patient group as those patients suffering from CSA induced by heart failure, kidney failure, and stroke. Fig. 3.4 (c) shows that we can capture the breath pauses, also indicated by the NCS waveform and RR. The shallow and deep over-breathing patterns can also be observed from the rise and fall of RV_{NCS} in Fig. 3.4 (d). In contrast, RV_{Ref} captures the increase in RV by deep over-breathing, but not the decrease during the shallow breathing period.

Biot's respiration

This breathing pattern is characterized by periods of apnea followed by fast respiration cycles with regular magnitudes. Different from Cheyne-Stokes, it lacks the crescendo-decrescendo cycles and is irregular with varying periods of apnea. This breathing pattern

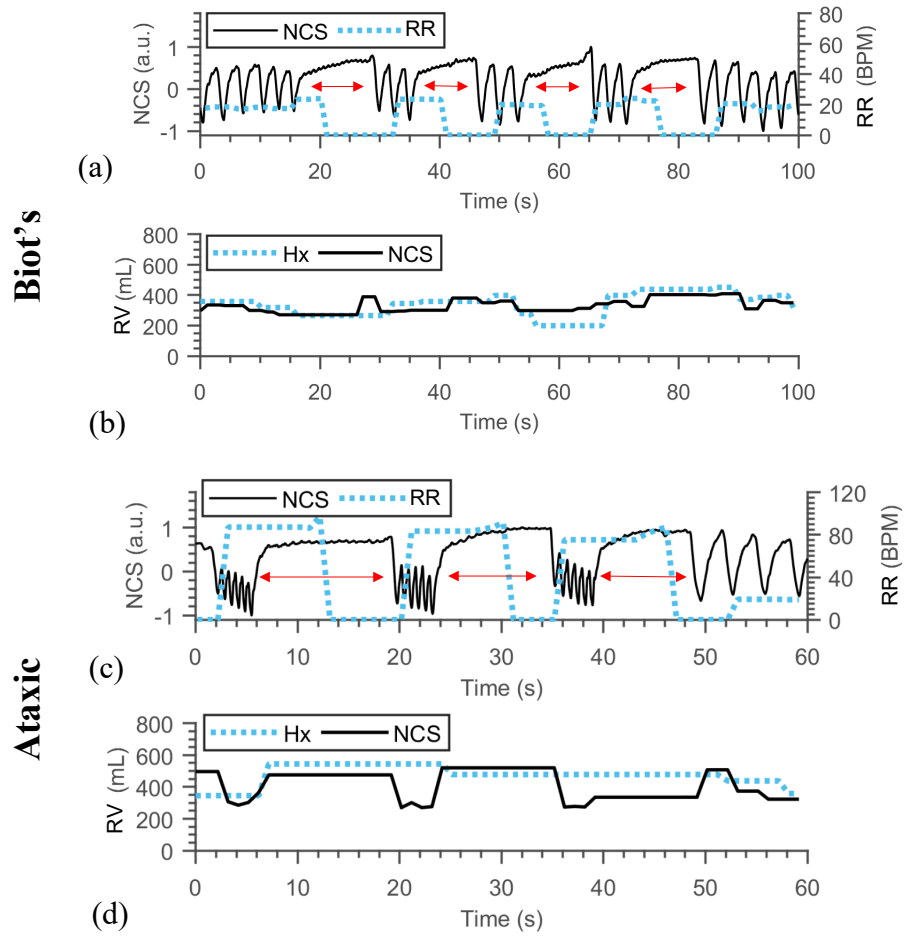


Figure 3.5 Respiration waveforms in Biot's (a)-(b) and Ataxic (c)-(d). (a) The NCS respiration waveform showing Biot's breathing with periods of apnea indicated by the red arrows, followed by rapid regular breathing. (b) RV estimates from Hexoskin and NCS. (c) The NCS waveform showing ataxic breathing with irregular rapid breaths followed by periods of apnea. (d) The correlated RV estimates from NCS and Hexoskin. NCS shows decrease in RV during the rapid irregular breathing.

is a result of damages to the central pattern generator in the medulla, possibly due to stroke, trauma, or injury. It is also observable in the case of opioid intoxication, which may also result in CSA and Cheyne-Stokes. Fig. 3.5 (a) shows sections of fast breathing with breath pauses, marked by an initial increase and then decrease in RR. The corresponding RV variation is shown in Fig. 3.5 (b) where the constant RV period corresponds to apnea.

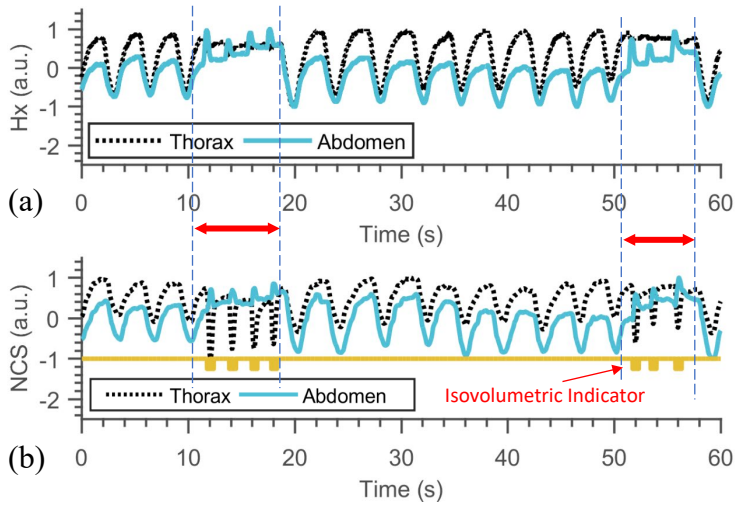


Figure 3.6 (a) Hexoskin thorax and abdomen waveforms with normal breathing and instances of isovolumetric abdomen contraction, marked by the red arrows. The breathing transitions are marked by the vertical dashed blue line. (b) Respiration waveforms from two NCS sensors showing asynchronous motion of thorax and abdomen during isovolumetric movement, similar to Hexoskin waveforms. Isovolumetric indicators as gold squares show instances of out-of-phase thorax and abdomen waveforms, detected from negative product of their slopes

Ataxic respiration

Ataxic breathing involves irregular breathing cycles, followed by irregular pauses and increasing apnea time. This pattern is observed when Biot's breathing deteriorates due to further damage to the brain center responsible for Biot's breathing. As shown in Fig. 3.5 (c), during the irregular rapid breathing, complete exhalation is not performed, resulting in lower RV estimates. Fig. 3.5 (d) shows NCS capturing decrease of RV and then increase at the last irregular breathing due to complete exhalation. No change is observed during breath-hold. The reference chest belts missed the decrease in RV during rapid breathing.

Obstructive sleep apnea

One of the major sleep apneas is OSA, caused by respiratory muscle weakness or air duct blockage. This apnea is characterized by the asynchronous movement of the

thorax and abdomen that can be observed as a time lag of the thoracoabdominal motion or complete out-of-phase motion of thorax and abdomen [95].

We performed isovolumetric abdomen movement by successively contracting and relaxing the abdomen muscles, without inhalation and exhalation. This resulted in a pure paradoxical breathing motion where the thorax and abdomen are moving in the opposite direction. With only one sensor placed below the thorax, it gets coupled to diaphragm, lungs, and some abdomen motion, and gives accurate RV estimations only when both abdomen and thorax are in sync. However, when the abdomen and thorax are moving asynchronously, this setup becomes sensitive to the sensor placement on the chest, with ambiguous results based on dominating thorax or abdomen movement. To separately observe both thorax and abdomen motion, we used B210 in the 2×2 MIMO mode with one pair of Tx – Rx on the thorax, and the other pair on the abdomen, with different baseband modulation to separate coupling of both movements.

Fig. 3.6 (a) shows the RIP chest belt data of regular breathing and intervals of the paradoxical abdomen and thorax movement. This strong voluntary movement is observable in the NCS data from thorax and abdomen sensors in Fig. 3.6 (b). Reference thorax belt does not clearly show opposite movement in all the cases, such as t = 55 s. A simple algorithm can detect time instances with isovolumetric breathing. As the slope of abdomen and thorax waveforms are opposite during isovolumetric inhalation and exhalation, we perform slope estimation at each sample point, using current and previous data points, and calculate the product of the two slope estimates at abdomen and thorax. The sign of product is negative during isovolumetric breathing, as indicated in Fig. 3.6 (b) by gold squares. The extracted NCS waveform from two sensors can

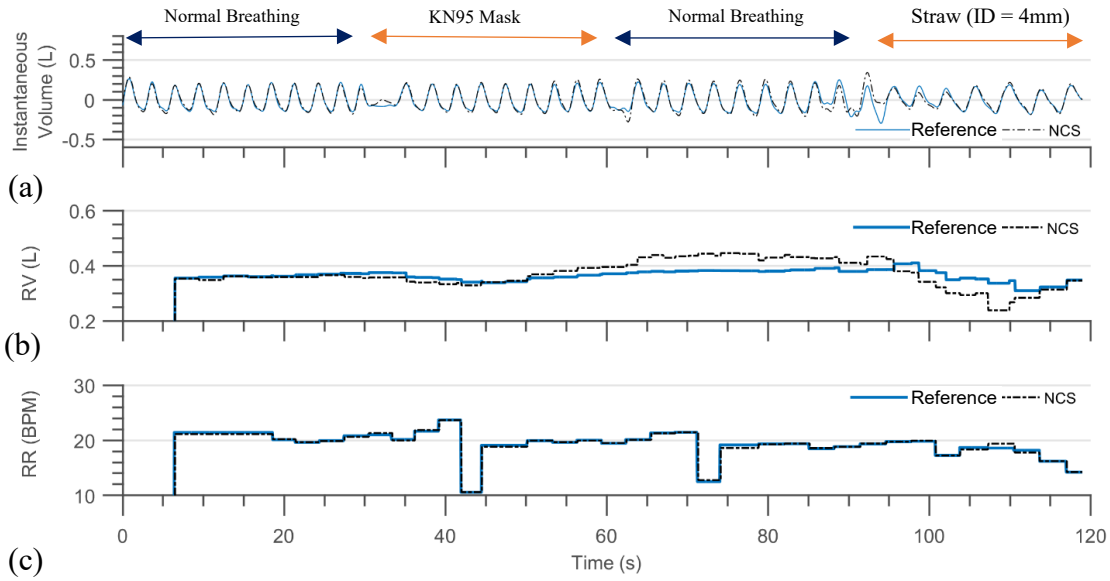


Figure 3.7 Studying the effect of simulated airway resistance variation. Subject is sitting upright while performing normal breathing. Two external resistances are introduced: 1) KN95 mask, while breathing normally through nose, and 2) straw with an inner diameter of 4 mm, only breathing through mouth. (a) & (b) plots show calibrated instantaneous volume and derived RV estimates respectively, (c) shows corresponding RR. The response to mask is nearly not distinguishable from the normal breathing from both RV and RR estimates, slight dips around $t = 45\ s$ and $70\ s$ are due to the breath holds during transition periods. Breathing through straw shows decreased RV in the beginning, while maintaining same RR, the following increase in RV is associated with decreased RR, showing the increased respiratory effort required due to the resistance offered by the narrow straw. While both NCS and reference chest belts can capture these variations, NCS appears to be more sensitive to the RV changes.

accurately capture the asynchronous thoracoabdominal movements when OSA is of high concern.

Airway resistance variation

We have also studied the influence of different simulated airway resistances on RV by introducing two resistances: 1) KN95 mask covering face and mouth while breathing normally through the nose, and 2) straw with a 4 mm inner diameter while only breathing through the mouth. The extracted instantaneous volume, RV, and RR from chest belts and NCS are shown in Fig. 3.7. RV calibration was performed before

the entire routine. While breathing through a mask does not show any noticeable change in RV and RR estimates, breathing through straw shows decreased RV towards the beginning with constant RR, followed by an increase in RV with decreased RR. This opposite RV and RR behavior indicate the increased respiratory effort required due to the resistance offered by the narrow straw. The consistency of RV between the NCS and reference chest belts suggests that the respective calibration remains reasonable within the tested range of varying simulated airway resistance.

3.1.4 Benchmarking results

To benchmark the overall NCS results against reference Hexoskin, linear regression, and the correlation between RR and RV are estimated from Hexoskin and NCS waveforms. The data set includes two sessions of 10 mins each, including normal, deep, slow-shallow, and fast-shallow breathing. Pearson's correlation coefficient (r) is very high for RR, $r(RR_{Ref}, RR_{NCS}) = 0.99$, and slightly lower for RV, $r(RV_{Ref}, RV_{NCS}) = 0.86$. Notice that RV_{Ref} has shown serious problems when RR changes suddenly. Figs. 3.8 (a) and 3.8 (b) show the scatter plot of RV_{NCS} vs RV_{Ref} and RR_{NCS} vs RR_{Ref}, respectively, together with the regression line. Instances of different breathing conditions are also marked, where the outlier cases are mainly at the transition point of two breathing conditions.

For comparison, we have also shown NCS vs Hexoskin scatter plots under various respiratory disorders in Figs. 3.8 (c) and 3.8 (d). These scatter plots summarize earlier observations in Figs. 3.3 – 3.5, where NCS and Hexoskin RV estimates were poorly correlated as RV_{Ref} often missed sudden changes in breathing patterns. RR

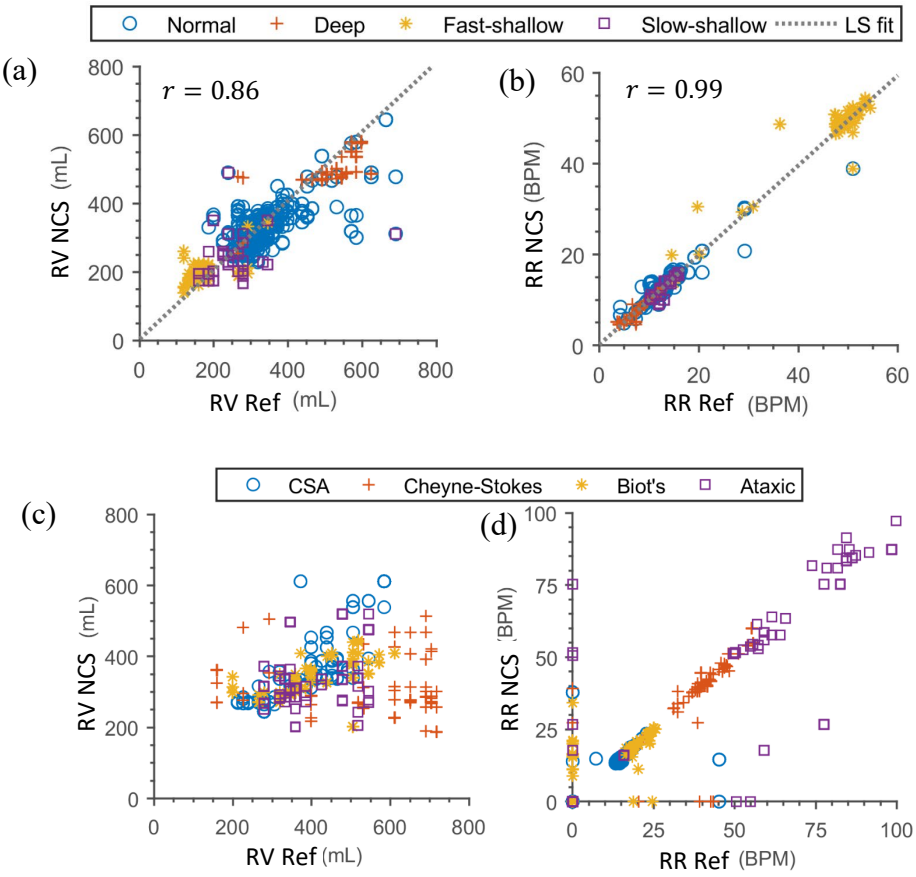


Figure 3.8 (a) Calibrated RV_{NCS} vs RV_{Ref} for normal, deep, fast-shallow and slow-shallow breathing conditions, showing the dotted least-squares fit line. Legend is shown at the top. (b) RR estimates, RR_{NCS} vs RR_{Ref} under the same breathing conditions. (c) and (d) show scatter plots, RV_{NCS} vs RV_{Ref} and RR_{NCS} vs RR_{Ref} respectively under various respiratory disorder conditions, only to illustrate low correlation between computed estimates from NCS and Hexoskin. Legend is shown at the top.

estimates are more correlated, with some outliers on the axes, again possibly due to differences at the transition points to apnea or missed breath peaks.

An additional subject was tested for normal, deep, fast shallow, and apnea breathing, which gave similar results. The root mean square deviation (RMSD) for both

Table 3.1 RMSD of RV and RR for normal, deep and shallow respiration

Subject	$\mathbf{RMSD}(x_{NCS}, x_{Ref}) = \sqrt{\frac{\sum_{i=1}^{N}(x_{NCS,i}-x_{Ref,i})^2}{N}}$	
	$x = RV\ (mL)$	$x = RR\ (BPM)$
1	53.3	1.1
2	114.5	5.8

subjects 1 and 2 are shown in Table 3.1. The larger difference in Subject 2 is likely due to the placement of the chest belts and NCS tags.

3.2 Gender, age, posture variation study

The previous section demonstrated preliminary NCS results where RV and RR can be estimated with high accuracy for one participant under various respiratory disorders. This section is an extension of the previous work with an over-clothing wearable RF NCS sensor study, conducted on 20 healthy participants with varying body mass index (BMI) and gender [42]. Two prototype sensors were placed on the participants, one close to the heart and the other below the xiphoid process to couple to the motion from heart, lungs, and diaphragm by the NCS principle and derive HR, RR, and RV parameters, as shown in Fig. 3.9 (a). Further, as RV is sensitive to body posture and breathing styles, we have presented analyses over various postures, as well as during conscious and spontaneous breathing exercises with a large breathing range of 0 — 45 BPM and a resting HR in the range of 50 — 90 BPM. A short calibration period by a gold-standard pneumotachometer (PTM) was performed for each subject and posture once, and the corresponding model was used on NCS and chest belts for further voluntary breathing exercises.

3.2.1 Experiment setup

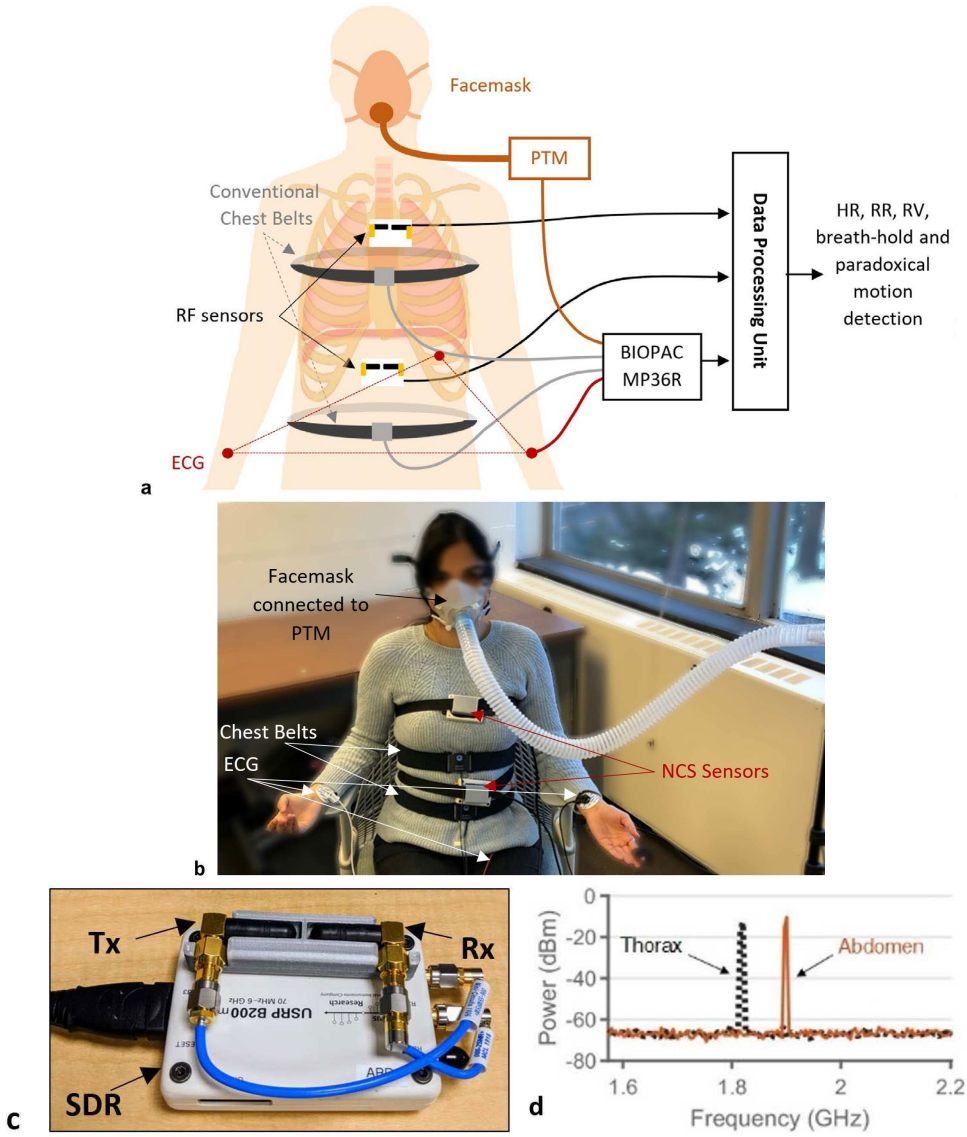


Figure 3.9 The large-scale human study experimental system. (a) Schematics of NCS and BIOPAC sensors and data flow. (b) Experimental setup with the participant wearing NCS and BIOPAC sensors in the sitting posture. Photo taken and published with written informed consent of the subject. (c) The NCS sensor consisting of SDR as well as the Tx and Rx antennas in a 3D-printed package. (d) Spectrogram of thorax and abdomen NCS sensors at their respective carrier frequencies of 1.82 GHz (-12.84 dBm) and 1.9 GHz (-10.42 dBm).

The NCS sensor prototype is implemented by an Ettus USRP B200mini [81] SDR, with monopole helical antennas (Taoglas TG.19.0112), packaged in a 3D-printed case, as shown in Fig. 3.9 (c). The relative antenna placement is designed to adjust the

direct Tx-Rx coupling and enhance the sensitivity to the reflected signal from internal organs. Two SDR units are multiplexed by frequency, with the carrier frequencies of 1.82 GHz and 1.90 GHz. This setup is a variation of the previous multiple-input MIMO implementation to improve sensor stability and participant comfort with a smaller SDR. The baseband tone is set at f_{BB} = 51 kHz, sampled at 2×10^6 samples per second (Sps). The NCS_{amp} and NCS_{ph} were derived in (2.3) and (2.4). For this work, we opted to use the best reference-correlated signal during the calibration phase for both thorax and abdomen sensors. For respiration, chest belt signals are taken as reference, and for the heartbeat, band-pass filtered (0.9 – 1.8 Hz) ECG signal is taken as the reference. The final demodulated data is sent to the control computer by a universal serial bus. The Tx powers are -12.84 dBm and -10.42 dBm for the thorax and abdomen, respectively, as shown in Fig. 2.9 (d), significantly below the allowable OSHA radiation exposure limit [80].

Reference measurements are performed by BIOPAC sensors, including a 3-lead ECG SS2LB, two torso belts SS5LB and PTM SS11LB, placed as shown in Figs. 3.9 (a) and (b). Unlike the previous smart-shirt, this unit is closer to hospital-grade instruments and allows using a more accurate RV measurement using PTM. The pre-calibrated PTM measures the airflow rate in L/s from the mouth using a facemask and is only placed on the subject for short calibration periods. For the remaining breathing exercises, reference respiration is only recorded by two belts placed at the thorax and abdomen that measure the change in local tension. During the experiment, male participants wore the thorax belt at 2 – 3 cm below the armpits, while female participants wore the belt below the breasts, close to the xiphoid process and costal margin,

considered as the dividing line between the rib cage and abdomen. This placement is selected for belt stability and user comfort. All sensors are connected to a 4-channel data

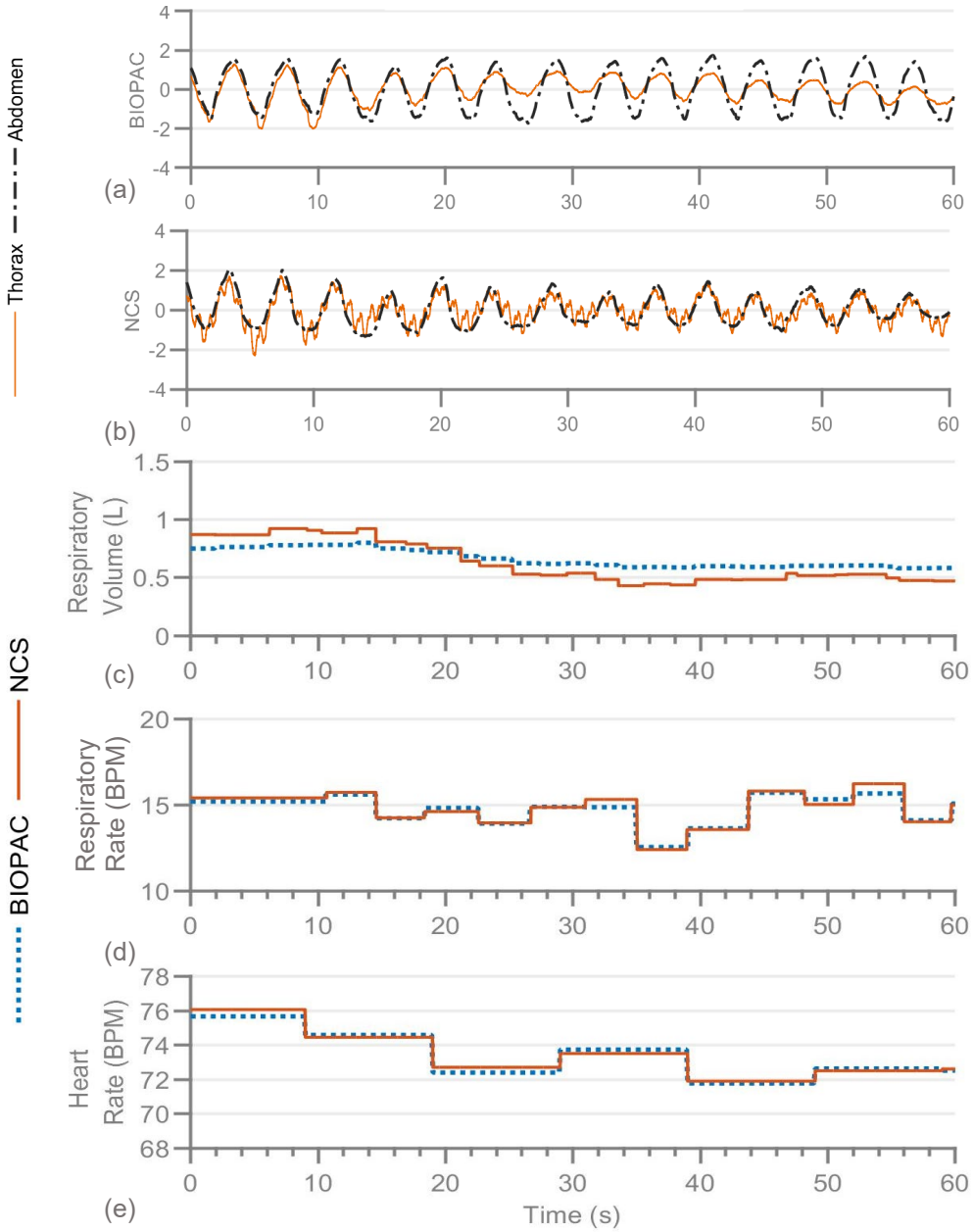


Figure 3.10 Examples of the sensor outputs and estimated RV, RR and HR. **(**a), (b) The normalized BIOPAC chest-belt signals and normalized unfiltered NCS signals, respectively, showing respiration waveforms during tidal (normal) breathing. The NCS thorax signal shows strong heartbeat motion as well. (c) – (e) The analysis of the data shown in **(**a) – (b); (c) and (d) show RV and RR respectively from NCS and BIOPAC chest belts; (e) shows the HR from NCS and BIOPAC ECG.

acquisition unit BIOPAC MP36R [96]. Fig. 3.10 shows example respiration waveforms from BIOPAC and NCS and estimated RV, RR, and HR.

3.2.2 Participants and protocol

The human study protocol was approved by the Cornell IRB, and participants provided written informed consent to take part in the study. Twenty-five healthy participants with no known history of cardiopulmonary diseases were recruited and instructed to follow breathing routines to the best of their abilities, without overexerting, in a sequence of three postures: supine left lateral recumbent and sitting upright in a chair. The instructions were provided to the participants in real-time using LabVIEW in both audio and visual formats. The data collection was carried out in a standard laboratory room with drywalls and supporting metal frames, without any radiation-absorbent material. The environment consisted of standard furniture including a bed, desks, and chairs, along with computers and various units of laboratory equipment. The participant attire was not controlled, and the NCS sensors were placed over their daily clothing, including shirts and loosely fitted hoodies made of different fabric materials. Five participants' data was rejected due to poor calibration data in any one of the three postures, which showed inconsistency due to either an inability to follow the mouth-only breath instruction or from the loose placement of the facemask that caused air leakage.

The 20 eligible participants included 14 females and 6 males with age from 18 to 34 years old ($\mu = 22.9, \sigma = 3.3$), weight from 49.8 to 79.5 kg ($\mu = 62.2, \sigma = 9.3$) and height from 158 to 183 cm ($\mu = 167, \sigma = 6.8$). The BMI of the participants fell

within the slightly underweight to slightly overweight range from 18.3 to 26.8 $\mathrm{kg/m^2}$ ($\mu = 22.2, \sigma = 2.2$).

A fixed breathing protocol of approximately 6 minutes was executed in each posture after calibration, which included:

- Normal breathing for 120 s.
- Deep breathing for 60 s.
- Fast breathing for 30 s.
- Normal breathing for 30 s.
- Breath-hold for maximum 20 s, followed by normal breathing for 20 s. Repeat once.

During normal or tidal breathing, no inhalation or exhalation instructions were provided, but the participants were asked not to take any deep breaths or long pauses. During deep breathing, real-time audio and visual instructions were given to start inhaling and exhaling with fixed durations of 4 s and 6 s respectively, which gave a RR of 6 BPM. Similarly, instructions were given for fast breathing at a rate of 40 BPM, with equal time for inhalation and exhalation. These instructions were provided as a guideline, and participants were advised to perform normal breathing if they felt uncomfortable during any of the routines, resulting in some variations around the expected rates. The procedure was clearly explained to the participants, with short examples for practice, before recording the calibration and main routines.

For observing separate thorax and abdomen motion, the participants were asked to perform the isovolumetric exercise with the following steps during breath-hold: contract the abdomen inwards, maintain the position for 3 s, then relax abdomen back to normal. Three such maneuvers were performed, separated by normal breathing intervals of 10 s to simulate paradoxical abdomen-thorax motion as observed in OSA.

As this motion is difficult to perform without extended practice, this exercise was only performed in the sitting posture, so that the participants could learn the motion, while looking at their real-time NCS/chest belt waveforms.

With the participants continued to be seated in a chair, an additional series of exercises were performed to test the breathing patterns in the relaxation state and under the given attention-engaging task. The participants were asked to close their eyes and relax for 5 minutes. This relaxed period allowed free-breathing over time without any instructions, thus increasing the likelihood of non-voluntary tidal breathing. The next routine was designed using PsyToolkit [97], [98] to render the participants under an attentive and cognitive task, where they watched and reacted to anomalous jumps of an on-screen rotating clock hand by pressing the space bar. Instantaneous feedback was given by flashing an on-screen indicator light of the correct keypress, to ensure continuous attention. This routine was performed for 6.5 minutes and rendered non-voluntary breathing patterns under fast temporal variations due to induced stress. Both routines utilized the RV calibration from the previous sitting posture.

3.2.3 RV, RR, and HR results

Overall statistics

With tests in multiple postures and under different breathing exercises, an analysis of 100 experiments, with over 590 minutes of total recorded data for 20 participants has been performed. Fig. 3.11 shows the results across all three postures, with different breathing styles. A high correlation of NCS and BIOPAC reference for RV ($r_{RV} = 0.84$), RR ($r_{RR} = 0.93$) and HR ($r_{HR} = 0.95$) has been achieved, as shown in the scatter plots in Figs. 3.11 (a) – (c). The Bland-Altman ($B\&A$) plot has been

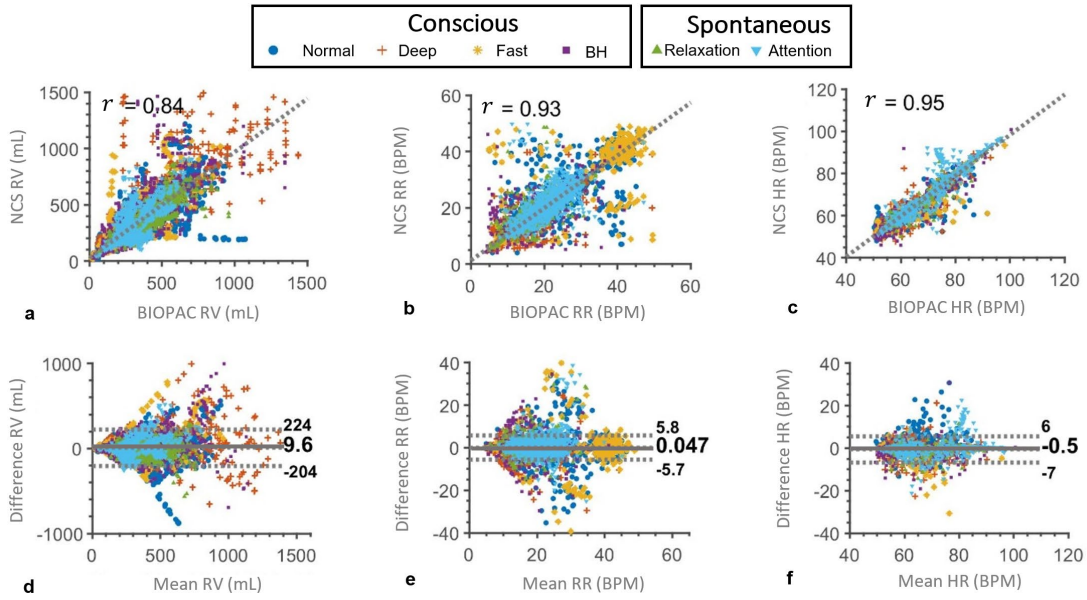


Figure 3.11 Correlation and agreement between NCS and BIOPAC estimate of RV, RR, and HR over the entire data. The label shows a marker for each breathing style, including conscious normal, deep, fast and breath-hold (BH), as well as spontaneous breathing in relaxation and attention states. (a) – (c): Scatter plots of NCS vs. BIOPAC RV, RR and HR, respectively, with denoted Pearson's correlation coefficient, r, showing high correlation between the two sensors. (d) – (f): $B\&A$ plots of NCS and BIOPAC showing the bias m at the center (solid line) and the corresponding LoA (dotted lines) given by $m \pm 1.96 \cdot \sigma$.

employed to quantify the agreement between NCS and BIOPAC, both of which may have errors. This agreement is estimated by the mean (m) and standard deviation (σ) of the measurement differences. $B\&A$ plots can also identify possible outliers visually by a XY scatter plot, with the Y axis as the pairwise difference, and the X axis as the mean of the two measurements. The systematic bias is estimated as the mean difference m, and limits of agreement (LoA) within which 95% of the differences are expected to lie, are estimated as LoA $= m \pm 1.96 \cdot \sigma$, assuming a normal distribution. Figs. 3.11 (d) – (f) show good agreement of both the sensors with low mean deviations ($m_{RV} =$ 9.6 mL, $m_{RR} = 0.05$ BPM, $m_{HR} = -0.5$ BPM) and narrow LoA, as denoted by the

dashed lines around the mean value. The results for each participant individually are shown in Table 3.2.

Table 3.2 Correlation and *B&A* statistics for each participant.

Measurement		Subjects																			
		1	2	3	4	5	6	8	9	10	11	12	13	14	15	16	17	18	19	20
RV (mL)	r	0.6	0.85	0.68	0.73	0.71	0.80	0.45	0.80	0.20	0.84	0.97	0.90	0.76	0.71	0.86	0.72	0.92	0.43	0.82
	m	40	4	210	12	-8	58	21	-57	88	-12	-29	-9	-41	14	21	-3	1	0	-46
	σ	125	41	195	39	102	93	96	123	179	84	36	83	64	55	105	188	28	99	140
RR (BPM)	r	0.85	0.91	0.87	0.92	0.90	1	0.93	0.90	0.97	0.91	0.91	0.91	0.97	0.78	0.82	0.91	0.90	1.00	0.93
	m	-0.07	0.20	0.88	-0.27	0.17	-0.02	-0.27	0.01	-0.16	0.32	-0.46	-0.13	-0.20	0.39	0.53	0.50	0.27	0.01	-0.29
	σ	3.08	2.50	5.38	2.86	2.64	0.68	2.45	3.36	2.09	2.80	3.85	2.49	1.99	4.26	4.84	2.63	2.87	0.45	3.43
HR (BPM)	r	0.97	0.45	0.36	0.57	0.98	0.87	0.33	0.92	0.81	0.87	0.61	0.91	0.92	0.81	0.74	0.72	0.80	0.86	0.92
	m	-0.20	-1.06	-3.96	-0.36	-0.20	-0.06	0.08	0.20	-0.41	-0.39	-0.90	-1.01	0.10	-0.13	-0.27	-1.06	-0.12	-0.33	-0.11
	σ	1.26	4.42	6.08	2.92	1.19	2.20	7.44	1.13	2.15	2.40	2.43	2.79	2.10	2.31	1.83	2.58	1.00	1.63	1.42

Table showing Pearson's correlation coefficient (r), and mean bias (m) and standard deviation (σ) of the difference of NCS and BIOPAC estimates (*B&A* statistics) for each subject across all the postures and breathing styles.
Subject 7 is not shown here as at least half of the data is discarded due to poor quality, resulting in a small data set for any statistical analysis.

Variation over breathing patterns

As breathing is a hybrid voluntary-autonomous process, data has been collected during both conscious and spontaneous states. During conscious state, participants performed guided normal, deep, fast breathing, and breath-holds. For the spontaneous tests, no breathing instructions were provided, and instead, participants were first asked to relax and then execute an attention-engaging task. As performing the task required the participants to be seated, spontaneous breathing tests were performed only in the sitting posture. Fig. 3.11 showed the results over all the experiments with breathing styles denoted by different markers. As normal breathing is similar to spontaneous breathing, most of these points are overlapped. The fast breathing pattern is centered around the instructed 40 BPM in the RR plots. Deep breathing is indicated by low RR around 6 BPM, as well as high RV. These points are also spread out with a low correlation between NCS and chest belts, possibly due to data distortion from the large motion, resulting in nonlinear strain gauge response, as well as NCS sensor motion relative to the body. Breath-hold periods have low RR at the beginning of the hold

Table 3.3 Correlation and B&A statistics with variation over breathing patterns

Measurement		**Breathing Patterns**					
		Normal	Deep	Fast	Hold	Relaxation	Attention
RV (mL)	r	0.88	0.80	0.76	0.84	0.93	0.76
	m	0	60	18	33	-10	11
	σ	89	186	144	138	65	97
RR (BPM)	r	0.95	0.91	0.89	0.77	0.93	0.87
	m	0.05	0.02	-0.49	-0.13	0.11	0.19
	σ	2.48	2.58	6.66	4.09	1.72	2.48
HR (BPM)	r	0.95	0.93	0.90	0.96	0.98	0.95
	m	-0.69	-0.98	-1.70	-0.80	-0.23	0.04
	σ	2.90	3.45	4.38	2.92	1.81	3.67

period and may also be accompanied with deep breaths after the long holding period, thus are seen closer to the deep-breathing points in both RV and RR plots. Zero RR is not marked, but these periods are treated separately during the breath-hold detection stage. Fig. 2.19 showed an example of one participant following the conscious breathing protocol in the supine posture, and the extracted RV, RR, and HR over time with consistent observations.

Table 3.3 compares the correlation and $B\&A$ statistics for different breathing patterns. The relaxation state with spontaneous breathing shows the highest correlation

Table 3.4 Average RV, RR and HR over 3 postures: supine, left lateral recumbent and sitting

Posture	RV (mL)	RR (BPM)	HR (BPM)
Supine	329 ± 200	17.3 ± 8.5	65.3 ± 9.9
Left Lateral Recumbent	263 ± 210	19.5 ± 9.5	63.8 ± 9.7
Sitting	317 ± 179	19 ± 9.7	67.6 ± 9.3

Table showing mean (m) and standard deviation (σ) of the RV, RR and HR statistics over different postures. The data is collected over all 20 participants following the same voluntary breathing exercise routine of normal, deep, and fast breathing, as well as breath holds. Average RV is observed to be highest in the supine posture, followed by sitting and lateral recumbent, while RR has the opposite trend. HR is highest in the sitting posture, followed by supine and lateral recumbent postures. The standard deviation likely demonstrates variation introduced from different breathing exercises.

Table 3.5 Correlation and B&A statistics with variation over postures

Measurement		Postures		
		Supine	Lateral Recumbent	Sitting
RV (mL)	r	0.81	0.89	0.81
	m	2	18	27
	σ	121	115	132
RR (BPM)	r	0.94	0.93	0.93
	m	-0.06	0.06	-0.08
	σ	2.84	3.60	3.52
HR (BPM)	r	0.96	0.93	0.94
	m	-0.70	-0.62	-1.22
	σ	2.72	3.54	3.20

of NCS and BIOPAC sensors for RV and HR estimates, close to conscious normal breathing results. This could be due to the regular breathing pattern with few motion artifacts. In comparison, participants tended to move in response to stress during the attention test, leading to motion artifacts.

Variation over postures

Adjusting body postures can lead to RV variations, as respiratory mechanisms are affected by different resistance or compliance of the lung and chest wall components. Thus, tests were performed with the guided breathing protocol in three postures: laying on a bed in supine and left lateral recumbent postures and sitting in a chair. The average reference RV, calculated over all 20 participants during the identical protocol of voluntary breathing exercises, is observed to be highest at 329 mL in the supine posture, followed by sitting and lateral recumbent at 317 and 263 mL, respectively. The RR has the opposite trend with an average RR of 17.3, 19, and 19.5 BPM in supine, sitting, and lateral recumbent postures, respectively. The average HR is highest during sitting at 67.6 BPM, compared to 65.3 and 63.8 in supine and lateral recumbent postures, respectively. These results are summarized in Table 3.4. The detailed statistics for each posture are shown in Table 3.5. Good correlation is observed across all the postures, with supine showing the least bias (m) for RV and RR estimates with narrow LoA (small σ), possibly due to the stable posture. No other clear trend is observed, indicating that the estimates are more sensitive to breathing types and patterns than to posture. Fig. 3.12 shows the correlation and $B\&A$ plots for different postures, both with and without the simulated breathing exercises.

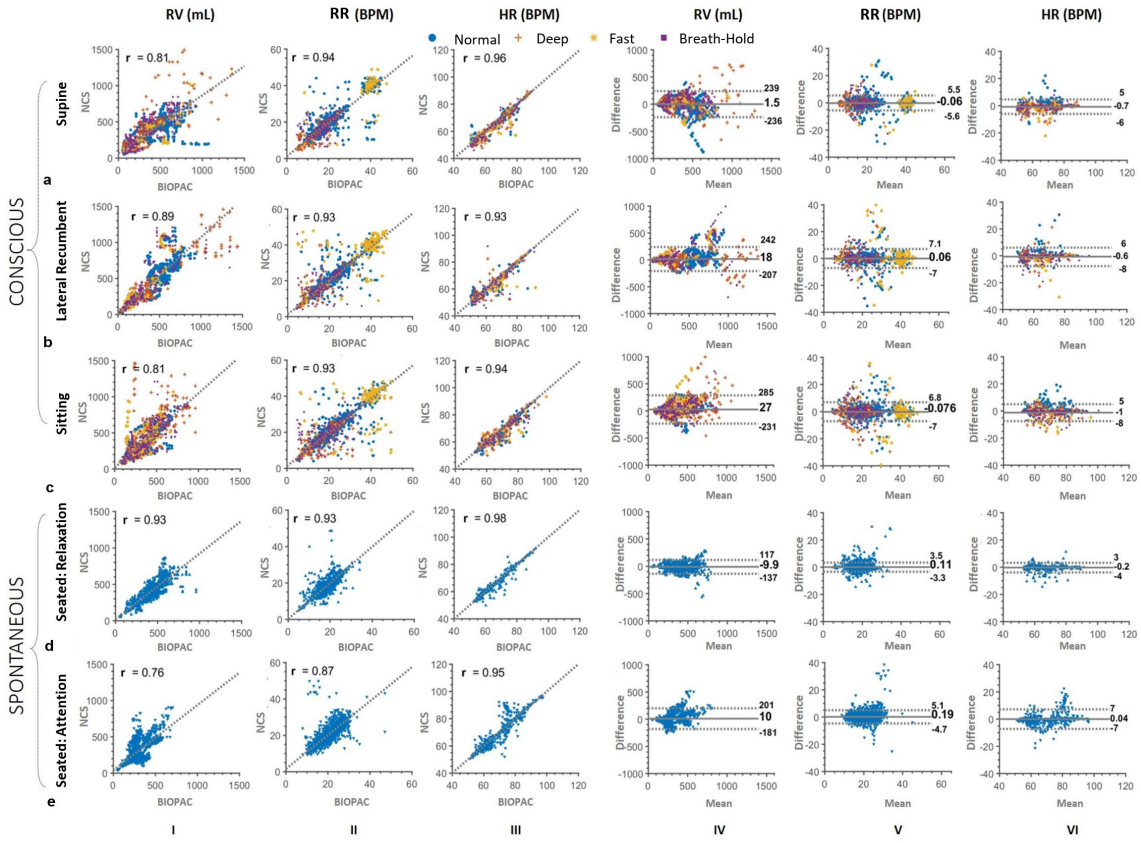


Figure 3.12 Comparison of NCS and BIOPAC data. Scatter (I – III) and *B&A* plots (IV – VI) between NCS and BIOPAC showing correlation coefficients (r), bias (m) and limits of agreement (LoA: $m \pm 1.96 \cdot \sigma$) for the RV, RR and HR across all routines. (a) – (c) The results across supine, lateral recumbent and sitting postures while following the breathing protocol of normal, deep, fast breathing and intermittent breath-hold. Different breathing protocol periods are indicated with a specific marker. (d), (e) The spontaneous breathing protocol results during the relaxation and attention states with the participant sitting upright.

Breath-hold detection

Each participant was instructed to perform two breath holds in both supine and lateral recumbent postures for a maximum duration of 20 s, simulating central sleep apnea (CSA), which is indicated by breath cessation for at least 10 s [15]. A simple detection algorithm was implemented based on the inspiration peak-to-peak interval. Overall, both NCS and BIOPAC performed well for breath-hold detection, as shown in

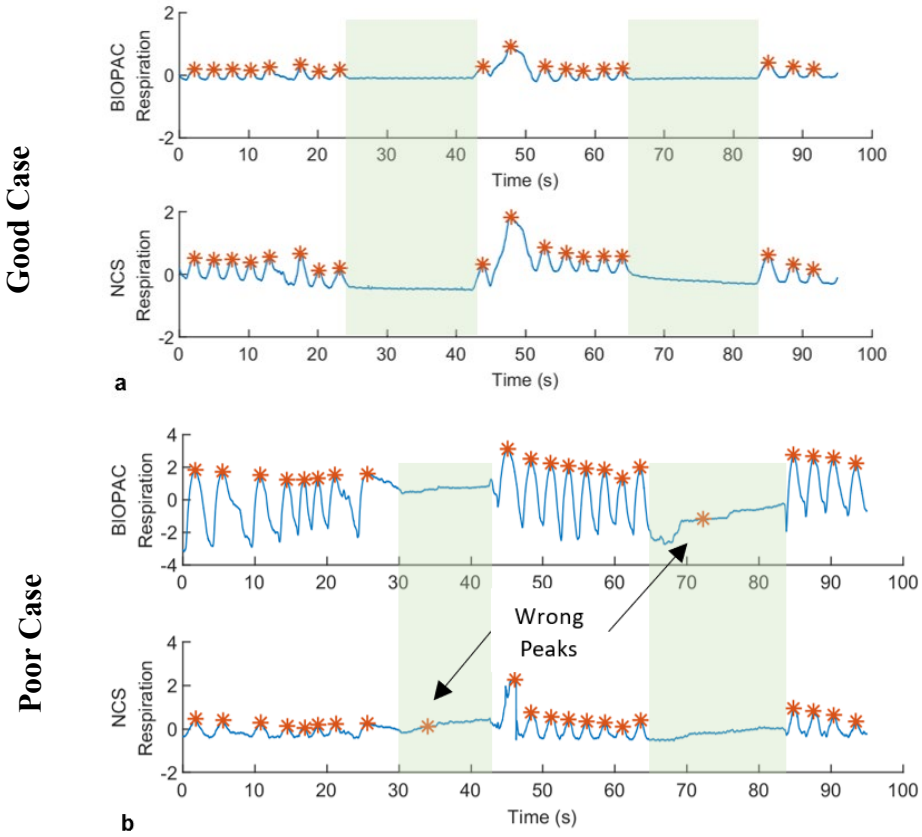


Figure 3.13 Breath-hold detection for two participants on instances of good and poor cases. Green shaded sections show manually annotated simulated apnea durations. The top and bottom figures show normalized BIOPAC and NCS abdomen respiration waveforms with detected peaks. The participant can perform the breath hold without any motion in (a), leading to accurate detection. In (b), wrong peaks are detected in both NCS and BIOPAC waveforms as shown, due to artifacts and peak detection limitations, as the participant is not able to maintain the breath hold without any motion. The motion coupling is different for the two sensors, as seen around $t = 68$ s, where BIOPAC shows some abdomen motion leading to wrong peak detection, but not NCS.

Table 3.6, which shows both sensors detected 74 cases out of 80 annotated cases, with each missing 3 non-overlapping instances. The errors possibly originate from the participant's noncompliance with the breath-hold protocol. Small torso motion is coupled differently to the two sensors, leading to incorrect peak detection during breath-hold. Fig. 3.13 shows the NCS and BIOPAC respiratory waveforms with detected peaks during breath-holds for representative good and poor cases.

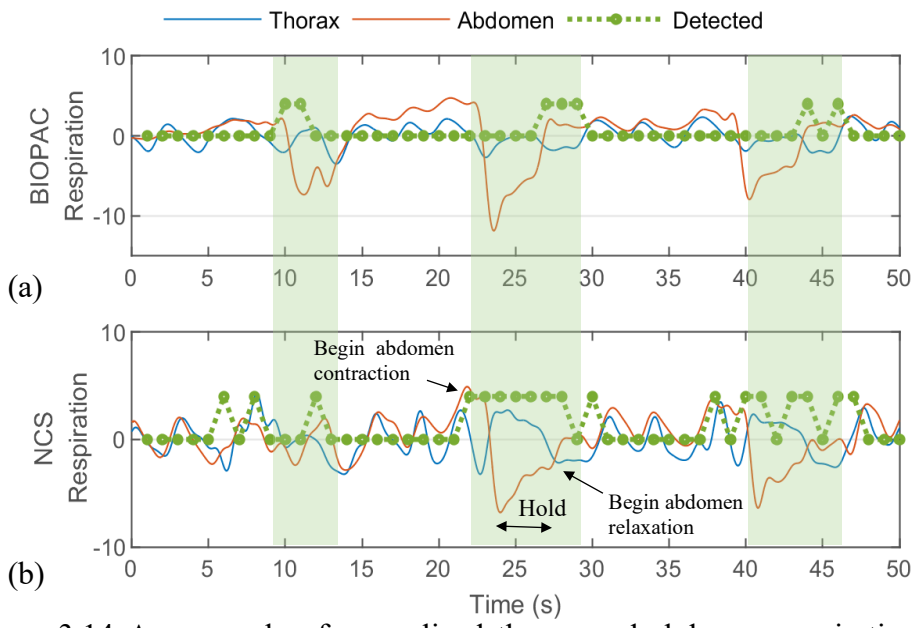


Figure 3.14 An example of normalized thorax and abdomen respiration waveforms during the isovolumetric maneuver by one participant. (a) Normalized BIOPAC chest belt waveforms, and (b) Filtered NCS respiration waveforms. The intended paradoxical motion windows are marked by green-shaded areas and detected instances are shown by positive value of the dotted green lines. Timing of abdomen contraction, hold and relaxation is denoted during the second cycle of the NCS waveform. Both BIOPAC and NCS can detect all three instances of paradoxical abdomen-thorax motion.

Paradoxical abdomen-thorax motion detection

To test the separate thorax and abdomen motion, the participants were asked to perform an isovolumetric abdomen exercise while holding their breath. With no airflow, the inward abdomen contraction results in outward motion of the thorax, as the total lung volume is conserved [78], simulating paradoxical abdomen-thorax motion similar to OSA with completer closure of the airway. We used the slope-product of thorax and abdomen respiration waveforms to detect paradoxical motion in BIOPAC and NCS waveforms, as shown in Fig. 3.14, where three instances of isovolumetric maneuver are successfully detected by both the sensors.

While some participants were able to successfully perform breath-hold and abdomen contraction, anomalous instances when paradoxical behavior was not observed can be attributed to the following possible reasons. 1) Placement sensitivity of sensors: the BIOPAC thorax belt placed near the xiphoid process may be coupled to the

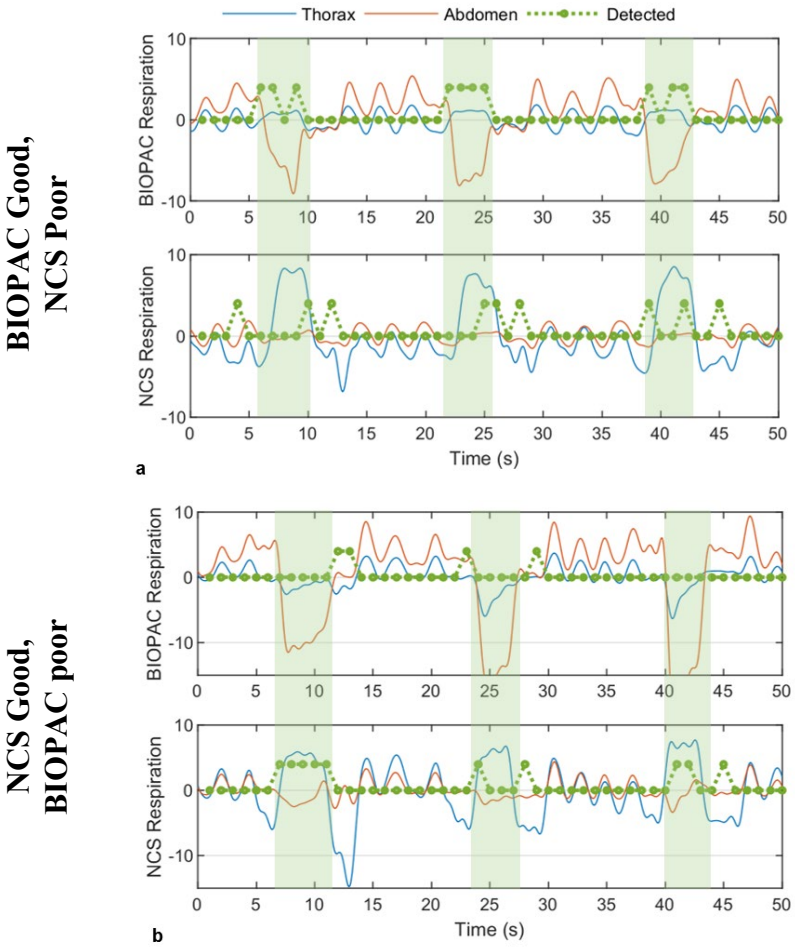


Figure 3.15 Two representative cases for paradoxical abdomen-thorax motion detection. Normalized respiratory waveforms from BIOPAC chest belts and filtered NCS are plotted in top and bottom figures, respectively. Intended instances are manually annotated by the green shaded areas. The detected periods are shown by positive value of the dotted green lines. True detection is marked if the annotated window overlaps with the observed instances. (a) The BIOPAC waveforms show clear paradoxical motion visually, as well as by the algorithm. NCS thorax and abdomen sensors do not show a complete paradoxical motion visually, but there are periods of opposite slope leading to detection of the second and third instances. Better abdomen sensor placement is required. (b) NCS waveforms show clear paradoxical motion visually, as well as by the algorithm. BIOPAC thorax belt do not show clear paradoxical motion in all three instances and requires better placement.

Table 3.6 Detection of breath hold (BH) and paradoxical abdomen-thorax (PAT) motion

BH ($n = 80$)		NCS	
		Detected	Missed
BIOPAC	Detected	74	3
	Missed	3	0
PAT ($n = 58$)		NCS	
		Detected	Missed
BIOPAC	Detected	42	9
	Missed	6	1

abdomen motion, and similarly the NCS sensor can be coupled to the accessory muscles. Cross-coupling between the two motions will reduce the paradoxical motion detectability. 2) Participants were unable to perform the isovolumetric exercise correctly during the intended period while following the breath-hold constraint. 3) The signal is lost due to sensor instability during the large chest circumference change resulting from the abdomen contraction. Overall, the algorithm was designed to be able to detect even slight paradoxical motion, resulting in the similar performance of both sensors as shown in Table 3.6. Both NCS and BIOPAC can detect 42 out of 58 instances, with fewer missed cases for BIOPAC (6) than for NCS (9). Figs. 3.15 (a) and (b) show two examples of poor detection for NCS and BIOPAC, respectively. While the algorithm performs well, as seen in these figures, it is more sensitive to baseline drift and other motion artifacts. Thus, sensor placement needs to be further investigated for robust paradoxical motion recognition.

3.3 Conclusion

This chapter first validated the one-sensor approach for monitoring various simulated respiratory disorder patterns and the associated RR and RV estimation compared to a reference sensor. The second part of the chapter validated these results over a broader population to test sensor performance with gender and BMI variation.

Overall, this RF NCS sensor and presented algorithm can achieve high accuracy under variation of user physical characteristics as well as their posture and breathing pattern variations. Thus, it promotes the applicability in real-life conditions for continuous monitoring of healthy subjects. A brief comparison with other noninvasive techniques focused on respiration is presented in Table 3.7. Different statistics are reported, including percentage accuracy and root mean square error (RMSE), apart from the ones used in this chapter, r, m, and σ. The proposed method provides a good

Table 3.7 Comparison with other noninvasive methods

Paper	Sensor	Statistics			Experimental Conditions
		Respiratory Rate	Respiratory Volume	Heart Rate	
Massagram [39]	Far-field Doppler RF; 1 m LoS	–	Sitting: r = 0.77; m = 39 mL; σ = 107 mL Supine: r = 0.72; m = 24 mL; σ = 129 mL	–	8 healthy spirometer
Nguyen [37]	Directional far-field CW RF	–	Bed accuracy: 95.4% (Error 58 mL)	–	6 healthy spirometer
Adib [38]	Far-field FMCW RF; 1 – 8 m LoS	Accuracy: 99.3% (Error 0.09 BPM)	–	Accuracy: 98.5% (Error 0.95 BPM)	14 healthy chest belt/oximeter
Reyes [32]	Smartphone camera	r = 1.0; RMSE = 0.4 BPM; $m = -0.02$ BPM; σ = 0.42 BPM	r = 0.98; RMSE = 182 mL; σ = 185 mL	–	15 healthy spirometer
Brüllmann [99]	RIP chest belts	–	Healthy: m = 0 mL; σ = 55 mL Patients: m = 20 mL; σ = 100 mL	–	5 healthy, 12 patients flow meter
Chu [19]	On-skin strain sensor	–	r = 0.96; m = –77 mL; σ = 152 mL	–	7 healthy spirometer
NCS [42]	**Wearable RF**	**r = 0.93; m = 0.05 BPM; σ = 2.93 BPM**	**r = 0.84; m = 10 mL; σ = 109 mL**	**r = 0.95; m = –0.5 BPM; σ = 3.32 BPM**	**20 healthy PTM**

A comparison of the proposed NCS method (across all positions and breathing styles) with existing technologies, focused on respiration. Gold-standard spirometer respiratory volume accuracy tolerance is $\pm 3\%$ [11].
Different performance metrics are used across papers including percentage accuracy, root mean square error (RMSE), correlation coefficient (r), mean (m) and standard deviation (σ) of the differences of measured and reference data ($B\&A$ statistics). Experimental conditions show number of participants and the reference measurement device for each work.
LoS line of sight, *FMCW* frequency modulated continuous wave, *RIP* respiratory inductance plethysmography.

estimation of both heart and respiratory motion characteristics with increased user comfort. The accuracy tolerance for a well-calibrated gold-standard spirometer is ± 3% [11]. For example, a quiet breath of 0.4 L will be measured as 0.4 ± 0.012 L, and a deep breath of 2 L will be measured as 2.0 ± 0.060 L. Table 3.3 showed the σ for different breathing patterns, with low values in spontaneous normal breathing conditions, $\sigma_{Relax} = 0.065$ L, $\sigma_{Attention} = 0.097$ L and $\sigma_{Normal} = 0.089$ L during the breathing protocol. While these values are higher than acceptable clinical values, they can be utilized for preliminary analysis where quantitative evaluation of respiratory efforts including both RR and RV along with HR is helpful, especially for geriatric patients with low cognitive function, where dyspnea information can only be currently determined from self-reporting or a caregiver's visual observation [100], [101].

The limitations of the current setup include data distortion due to antenna motion relative to the body, sensor placement sensitivity, and arduous calibration requirements. The former issue of antenna motion can be possibly resolved by improved antenna packaging and garment integration of the antenna. While sensor placement requirements are less stringent for most routines, it becomes more important for accurate paradoxical abdomen-thorax motion monitoring, as NCS sensor coupling to diaphragm and lungs is observed to be position sensitive and needs to be carefully deployed. The calibration protocol requires participants to breathe only through the mouth while wearing a facemask for a brief duration. This calibration is mainly to give an absolute scale in RV and can be omitted if the percentage volume change is sufficient.

Additionally, the NCS sensor with its small form-factor and simple transceiver architecture can be incorporated inconspicuously into the fabric with an improved

wireless design and can be worn over multiple layers of clothing without requiring any direct skin contact. We have used frequency division for multiplexing two sensors, which can be easily extended to place more sensors at different locations on the body, as well as on multiple people simultaneously. The Tx signal can be further modulated with a unique pseudo-noise code known to the corresponding Rx. These design options in NCS provide higher signal isolation against ambient interferences and inter-sensor collision than implementations based on the direct far-field RF and optical sensors. The detailed respiratory and heartbeat characteristics also open other areas of applications including cough monitoring, stress detection, and overall ambulatory healthcare monitoring.

CHAPTER 4

ATTENTION MONITORING USING NON-INVASIVE VITAL SIGN FEATURES

4.1 Introduction

The provision of individually tailored sensory response systems has been of increasing importance in recent years. Apart from sensing the surrounding context, it involves understanding the physical and physiological responses of the people. Using emotion recognition has been an important part of the human-machine interaction system by teaching a robot to give appropriate facial and speech responses. Smart buildings can be enabled to adjust lighting according to the emotional state and keep track of mental health, including depression and anxiety monitoring and alert family when required. With increasing work-from-home jobs, forced by the COVID-19 pandemic, such systems are even more important to monitor work fatigue. Keeping track of fatigue or the opposite, attention, is even more important for people working in critical conditions like soldiers in action, heavy machine operators, air traffic controllers, and drivers. With accurate knowledge of the attention state, actions in form of visual or sound stimuli can be provided if a person is slipping into drowsiness or a less attentive state.

The capability of an intelligent machine response requires an understanding of human emotional and cognitive reactions. There has been a debate starting from the early philosophers if reason and emotion are two opposite or related concepts, with recent research demonstrating a high level of interdependence [102]. When people perceive a vast amount of information, a subset processing is prioritized and extraneous

irrelevant information filtered out, to access the conscious awareness, which can be termed as attention [103]. Thus, attention is a basic function that controls focus, vigilance, and response simultaneously [104]. It can be classified as two types: exogenous attention is sensory-driven, involuntary response, while endogenous attention is a voluntary process with the conscious expectation of events [105]. There is another state called *emotional attention*, which is involuntary similar to exogenous attention but has been suggested to rely on special neural circuitry involving the amygdala [106]. Presently, psychological and neuroscientific research has extensively modeled emotion and its impact on attention and decision making. Brain studies to differentiate the amygdala response to emotion and attention have suggested a faster response to emotion over attention [107], [108], however, the overall results remain inconclusive [109]. In summary, emotional and attention states are highly intertwined, and the attention study can benefit from the vast research in the emotion monitoring domain to understand various physiological reactions.

In this work, we have focused primarily on attention versus relaxation classification. The attention under this study is endogenous, in response to certain events that the user interacts with during a specific task. The relaxation has been loosely defined as a state similar to drowsiness or fatigue but covers a broader aspect of user inattentiveness including the meditative state. This is an important factor, as the user needs to be attentive even though no harmful activity has been perceived for a long period, such as in driving or guarding. A study on 25 healthy participants sitting on the chair was performed while wearing NCS RF sensors on the thorax and abdomen with routines involving neutral, relaxation, and attention states while sitting in a chair.

Chapter 3 included data from these routines, which validated the sensor performance for respiratory effort and heart rate monitoring. The next section presents related works in the area of attention detection and emotion monitoring. The experiment setup and detailed data collection routine during attention are presented in Sec. 4.3. Sec. 4.4 presents the classification algorithm, including feature selection. Sec. 4.5 presents detailed results using NCS and reference sensors. Finally, discussion, including conclusion and future work are presented.

4.2 Background and related work

Long-term attentive tasks may result in mental fatigue [110] and require constant attention monitoring. Developing a safety attention monitoring setup ideally requires complete occupant centric knowledge, i.e.:

1) Sensing the surroundings.
2) Sensing the user's actions.
3) Sensing user's physiological responses like brain activity, heart and respiratory patterns, and motor expression including face, voice, and body.

The detailed sensing of the user environment is very difficult and also not suitable in most real-life applications. Autonomous or assisted driving fall in this category, but require a lot of research efforts that are difficult in all possible attention-demanding scenarios. Even with this sensing, increased drowsiness can result in the user's inability to take immediate action. Also, user action or response monitoring [111], [112] to record attention periodically can get very tiring and distracting by its action. Thus, monitoring

of physiological responses and motor expressions are the most suitable techniques for a generalized case.

Existing attention monitoring techniques mostly detect fatigue as an alternative. While fatigue can be both physical and mental, only the latter has been included here. This technique only focuses on cases arising out of tiredness or boredom leading to sleepiness. Cases, where the user is relaxed and not expecting any events to occur are not counted, which can be extremely dangerous, for example, in sudden combat attacks. Thus, a broader comparison of attention versus relaxation needs to be performed. This can be enabled by the interlink of attention and emotion. Emotions can be classified in different ways, where each feeling is an independent emotion of *joy, anger, sadness*, etc. [113] or they can be clustered in a smaller multidimensional space [114], [115], usually spanned by *valence* and *arousal* in addition to others. Valence is classified as

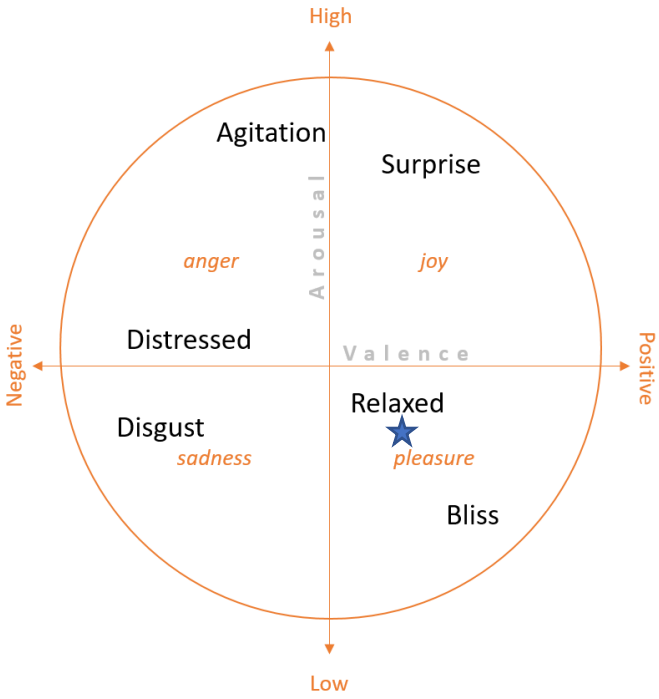


Figure 4.1 A two-dimensional arousal and valence model.

positive or happy emotions versus negative or sad emotions. Arousal refers to the activation level and ranges from low to high and is easier to differentiate than valence [115]. Fig. 4.1 shows the 2D arousal and valence model. Since the relaxed state is near the neutral valence and arousal states in the 2D model and attention may be accompanied with high agitation (high arousal – neutral valence) and distress (neutral arousal – negative valence), they are separable with this 2D model. In the following literature review, we will present the two areas of fatigue monitoring and emotion monitoring, due to connection of attention detection to both areas.

Fatigue monitoring: Previous research on attentiveness and fatigue detection comprises mainly of motor expression monitoring, including eye and mouth. The focus is mainly on fatigue detection, i.e., sleepiness which is characterized by change (increase or decrease) in blink rate and yawn, which may be an indicator of the arousal level [116]. As this approach requires facial expression monitoring, sensors in ambient are required. Thus, the system is often tailored to a specific setup like driving. PERCLOS (percentage of eye closure) has been widely used as an alertness measure but eye-only fatigue detection performance is inferior to EEG [117]. Computer-vision based eye detection systems and algorithms have been proposed [118]–[120], which may suffer from illumination conditions. Ji et al [121], [122] performed real-time driver fatigue prediction using near infra-red (IR) video cameras for eyelid, gaze, head movement, and facial expression tracking with a probabilistic model. The best accuracy can be achieved with physiological parameter sensing which may be inconvenient for driving with wearable sensors [121], [122]. Other alternatives include piezoelectric sensors placed

on the skin of eyelid for fatigue monitoring [123]. Electrooculogram (EOG) sensors have been used [124] to detect driver drowsiness.

Fatigue detection has also been performed by physiological measurements like changes in heart rate, brain activity, and thermoregulation, the last of which gives the best performance for monitoring physical fatigue in construction workers [125]. EEG has been extensively used for mental fatigue detection [126]–[130] with response time in 10s of seconds [130]. Skin conductance has also been used for mental workload [131].

Emotion monitoring: There has been a vast amount of research done in the area of emotion recognition from physiological signals [132]–[134]. The accuracy varies based on the recorded signals, sensing technology, emotion model, and classification algorithms. The real-life implementation heavily depends on the required comfort level of the sensor and accuracy tradeoff, for example, pulse oximetry (SpO_2) sensing from the wristwatch may achieve limited accuracy with high comfort.

The facial expression, speech, and other motor features have been extensively used for emotion recognition with improved results for multimodal approaches [135], [136]. A bimodal system used facial and acoustic information [135] to classify four emotions using k-nearest neighbor (kNN). A hidden Markov model (HMM) with maximum likelihood classifier was used for automatic segmentation and recognition from a continuous video into a sequence of emotions in [137], with a very low recognition rate for an unseen participant as the display of emotions varies with ethnicity and gender. Deep-learning based approaches have also been used in recent years [138], [139], including transfer learning [140] for emotion detection. The performance of these

methods relies on the expressiveness of the user, accurate baseline estimates, or large training data across different races, illumination, and background audio noises.

Galvanic skin response (GSR) or skin conductance (SC), blood volume pulse (BVP) and skin temperature (ST) from hands, and pupil diameter (PD) with an eye gaze tracking instrument were employed to monitor stress [141]. SVM classifier resulted in high cross-validation (CV) accuracy with PD as the most important feature. The existing literature has difficulty in detecting the intensity of emotion with only GSR and electromyography (EMG) [142], even with a relaxation baseline. Arousal versus valence classification has been performed using GSR and photoplethysmogram (PPG) [143] with leave-one-out subject accuracy of 70%, compared to 87% for a single-user model. While these parameters are important, they are associated with specific emotions. GSR and PD have been associated with arousal [144], [145]. EMG measurement of neck muscles could be an indicator of stress, and it is a single-dimensional valence measure [132]. Stress or anxiety may result in colder ST at the fingers. Thus, multiple sensors with different placements are required, all with direct skin contact, and are thus uncomfortable.

Another important category of work focuses particularly on heartbeat signals to measure heart rate variability (HRV) [146], [147]. Electrocardiogram (ECG) has been widely used for heartbeat based emotion recognition [148], [149]. A model based on empirical mode decomposition (EMD) with Hilbert-Huang transform (HHT) was used to derive features based on instantaneous frequency for arousal/valence study in [148]. As ECG requires skin contact, PPG has been seen as a desirable alternative, particularly with watch and fitness tracker integration [150], [151]. However, baseline drift, motion

artifacts, and signal quality need to be taken into account. As ECG has been established as a biometric tool [152], showing varying features across users, it should be used carefully for consistent emotion recognition across different subjects. Additionally, the emotional state has also been linked to the correlation of respiration and heartbeat in [115], indicating a correlation between different physiological channels. Only the respiration signal has been used with high accuracy in [153], [154] by a deep learning approach.

Noninvasive wireless RF sensing has also been used for emotion detection [155]–[157]. The traditional far-field sensing mostly measures respiratory signals only, which can be 10 times stronger than the heartbeat [155]. The heartbeat features were stated to be indirectly extracted from respiration in [155], achieving a high average leave-one-subject CV accuracy of 94.4% for stress detection. Raw RF data was used in [156] to train a convolutional neural network (CNN) followed by a long short-term memory (LSTM) deep learning architecture. This achieved 67 – 87% accuracies to classify 4 different emotions with user-independent learning. HRV was extracted with signal processing tools [157] with an FMCW radar, resulting in a 70 – 75% accuracy for a person-independent 4-emotion classification, using the valence/arousal model. This work required a baseline neutral state corresponding to the specific day and person for feature calibration. These wireless sensing techniques are restricted to indoor setups, supporting only a limited number of users in the room.

This work uses robust NCS wireless sensing, that can be placed over-the-clothes to record heartbeat and respiratory sensing. This allows us to monitor multiple people simultaneously without any interference from ambient motion, unlike existing

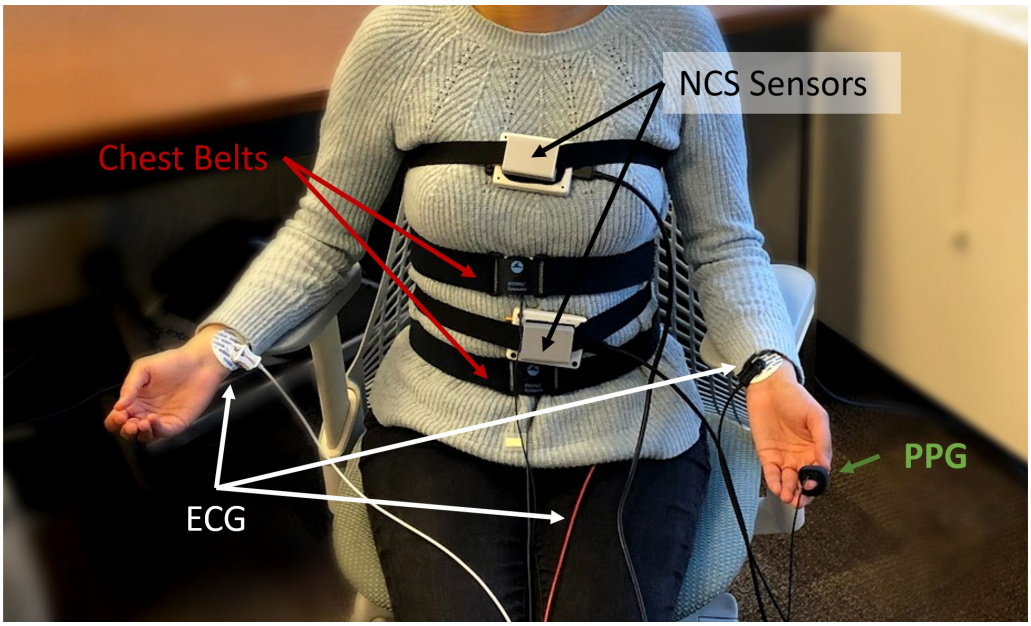


Figure 4.2 Setup for attention monitoring.

approaches. With near-field coupling, we can get a noticeably clear heartbeat signal, in addition to the respiration for feature extraction. We have established a relationship between emotion and attention and adopted similar tools as emotion recognition for a broader attention/relaxation classification, allowing adaptability to different applications. Further, with passive and active setups, NCS has high comfort levels in comparison with other approaches requiring direct skin contact.

4.3 Experimental setup

This experimental study was conducted as part of the previous study discussed in Sec. 3.2, approved by Cornell IRB, and performed with participants' informed written consent. The data collection was performed on 25 healthy subjects including 15 female subjects. The age range is 18 – 34 yrs., with BMI in the range 17.6 – 27 kg/m^2.

4.3.1 Vital-sign sensors

Two NCS sensors were placed at thorax and abdomen along with reference BIOPAC chest belts and ECG as shown in Fig. 3.9. No facemask calibration was required, and an additional finger PPG was placed as shown in Fig. 4.2. In this chapter, three routines in the sitting posture are considered:

R1 Breathing exercise routine (Initial 120 s only): Perform normal breathing without any deep breaths or breath holds.

R2 Relaxation routine (5 minutes): Close eyes, relax, and spontaneous breathing with no additional instructions.

R3 Attention routine (6.5 minutes): Play a vigilance-testing game with spontaneous breathing.

The attention monitoring setup uses heartbeat and respiration signals from NCS and compares them with the reference devices (ECG, chest belts). The information contained in these signals is discussed below.

Heartbeat: The gold standard sensor for heartbeat monitoring is ECG. The main features are heart rate, interbeat interval (IBI), and heart rate variability (HRV) [158] which can be measured in several different ways. HRV is the variation in IBI or RR interval (interval between two RR peaks in ECG) and reflects the effect of many physiological factors modulating normal heart rhythm.

Anatomically, the medulla located in the brainstem contains cell bodies for two main divisions of the autonomic nervous system (ANS), termed as sympathetic (SNS) and parasympathetic (PNS). Autonomic outflow from the medulla is divided into sympathetic and parasympathetic (vagal) branches, that modulate the activity of the

heart and blood vessels. These modulations are reflected on the HRV and are thus used extensively to monitor ANS functions. As the ANS, together with the hypothalamus gives response to emotional cues, the HRV is related to emotional responses [158].

The two branches, SNS and PNS, exhibit distinct functions. The fibers of SNS are present throughout the atria, particularly concentrated in the sino-atrial (SA) node and ventricles, and when stimulate the cardiac muscles to increase the heart rate (HR) when activated. This stimulation occurs in response to stress and emotional intensity and prepares the body for vigorous activity. PNS decreases the cardiac workload, primarily resulting from trauma and allergic reactions. The separate contributions from sympathetic and parasympathetic autonomic activity modulate the RR intervals of the QRS complex. The sympathetic activity exhibits a slow course of action, with effects observed 5 s after stimulation, returning to baseline after 20 – 30 s, and is thus associated with low frequency (LF) in the range 0.04 – 0.15 Hz of modulation frequencies of HR. Parasympathetic activity response appears rapidly with a response time of 400 ms and returns to baseline after 1s, associated with high frequency (HF) in the range 0.15 – 0.4 Hz of modulation frequencies of HR [147].

The HRV feature-based emotion detection has been explored in great detail in recent years, with time-domain (e.g. SDNN, RMSSD), frequency domain (LF, HF, 2D LF-HF), and non-linear (entropy and Poincare plot features) [158]–[160] analysis. However, care must be taken in signal processing, as even slight inaccuracy in IBI estimation can result in a huge variation. Thus, various standardized data processing tools have been established for the ECG data. As PPG data does not have a sharp peak and may have a small peak immediately following the main peak (dicrotic notch),

rigorous signal processing needs to be performed. For NCS heartbeat signal, 2nd harmonic based peak detection has been shown to perform better in previous chapters.

Respiration: Respiration is in general an unconscious, automatic, and involuntary process, controlled by the medulla and pons. Conscious factors can override the autonomous process for a limited period, for example, while speaking, smelling, or breath-hold. With the complexity of the respiration system, and in general favorability of heartbeat monitoring, the observations such as respiratory sinus arrhythmia [161] and rate variation relationship lack complete understanding [162]. In general, respiration is modulated by various emotional reactions including stress, anger, and relaxation. For instance, a decrease in respiratory rate (RR) is observed in a relaxed state. Tension, sudden fear, or happy surprise may result in a breath-hold, resulting in decreased RR, with increased variability. Deep and fast breathing can be anger or happiness. Further, while respiration is affected by emotions, voluntary respiration breathing techniques have been suggested to influence emotions and heartbeat [163]. Dyspnea or difficulty in breathing has been related to excitement or fatigue.

4.3.2 Attention routine

A vigilance based game, based on the Mackworth clock test [164] was developed to engage the participant's attention. This routine was designed in PsyToolkit [97], [98], where a graphical interface displayed rotating clock hand as shown in Fig. 4.3. Intermittently, the clock hand jumped more than normal, and the participant was expected to press the spacebar. The speed, normal rotation angle, and jump step were

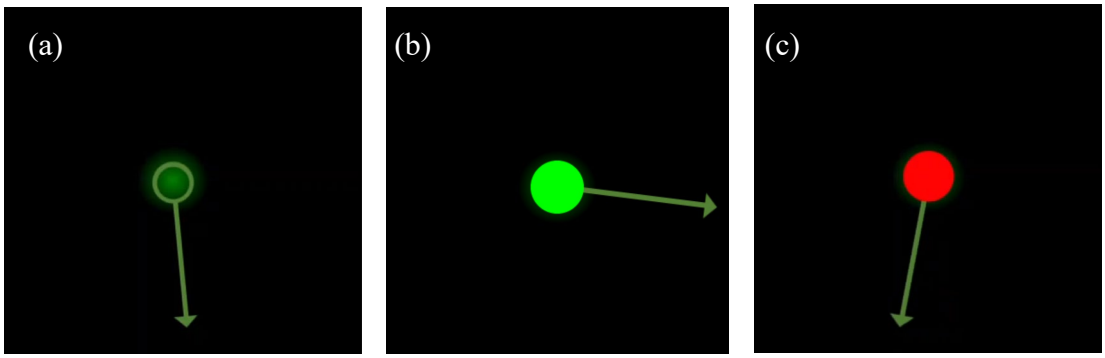


Figure 4.3 Attention test showing a clock hand. (a) Normal clock rotation. (b) Correctly detected jump. (c) Incorrect spacebar press or missed jump.

adjusted to keep the level easy. If the participant pressed the spacebar correctly, the middle bulb glowed green. If a jump was missed, or the spacebar was wrongly pressed, it would glow red. These two instances are shown in Figs 4.3 (b) and (c), respectively.

The attention routine of 6.5 min includes:

1. 30 s instructions as shown in Fig. 4.4,
2. 30 s trial round,
3. 10 s wait,
4. 5 min test round.

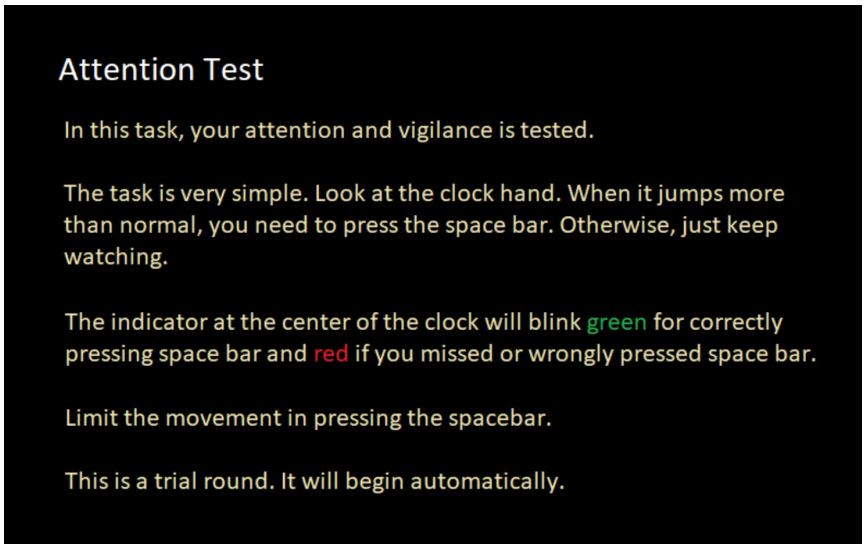


Figure 4.4 Instructions displayed at the beginning of the attention routine.

1	2	3	4	5	6	7	8	9	
32351	1000	3	126	10	0	0	1	trial	★ Missed jump
32892	500	1	162	4	0	1	0	trial	★ Incorrect press
34437	1000	3	198	6	0	0	0	trial	
35483	1000	3	234	5	0	0	0	trial	
36530	1000	3	270	3	0	0	0	trial	
37577	1000	3	306	7	0	0	0	trial	
38622	1000	3	342	6	0	0	0	trial	
39671	1000	3	378	1	0	0	0	trial	
40714	1000	3	414	8	0	0	0	trial	
41759	1000	3	450	4	0	0	0	trial	
42806	1000	3	486	9	0	0	0	trial	
43853	1000	3	522	6	0	0	0	trial	
44896	1000	3	558	7	0	0	0	trial	
45942	1000	3	594	6	0	0	0	trial	
46596	611	1	684	10	1	0	0	trial	★ Correct detection
48037	1000	3	720	5	0	0	0	trial	

Figure 4.5 The test generated dataset giving timing and reaction information, with column numbers shown on the top.

The normal clock hand rotation is 3.6°, and the abnormal jump is 5.4°. On each turn, a random number between 1 to 10 is generated, and an abnormal jump occurs if the number is 10. Thus, the jump probability with a uniform number generator is 0.1. A maximum reaction time of 1000 ms is allowed. Following the test, information about the game is returned, including 1) time in ms from experiment start corresponding to each rotation, 2) reaction time, 3) status (if a key was pressed), 4) the current position of the clock hand in the tenth of degrees, 5) generated random number, 6) value '1' if jump occurred and correctly detected 7) value '1' for the incorrect press, 8) value '1' if jump occurred and missed, and 9) 'trial' or 'test'. Fig. 4.5 shows the data arranged in

columns, with three star-marked instances showing correctly detected jump, incorrect press, and missed jump cases.

4.4 Algorithm

A feature-based machine learning (ML) classification algorithm has been used for attention and relaxation classification. The entire signal processing has been implemented in MATLAB.

4.4.1 Feature extraction and selection

The analysis is done at two different time scales. One uses the entire window of data collection for the routine, ranging from (2 min – 6.5 min) and another is an ultra-short window of 90 s with a 10s slide. Features are generated as various statistical measures estimated on the parameters extracted from waveforms:

1. Heartbeat: IBI
2. Respiration: Inter-respiratory interval (IRI), uncalibrated tidal volume (TV) or peak height, inspiratory interval (t_{II}), expiratory interval (t_{EI}), t_{EI}/t_{II}.

   Statistical analysis: mean, variance, approximate entropy (ApEn), successive differences at first-order (SD1), and second-order (SD2).

   Frequency domain: LF, HF

A subsection of these features is used for classification, to reduce noise and improve the classification. The features are ranked using minimum redundancy maximum relevance (MRMR) algorithm [165], using ground truth information of corresponding classes ($y = \{1: Normal, 2: Relaxed, 3: Attentive\}$). If a feature i, is differential for different classes, it should have large mutual information, $I(y, i)$. The mutual information is related to the joint probabilistic distribution of two variables

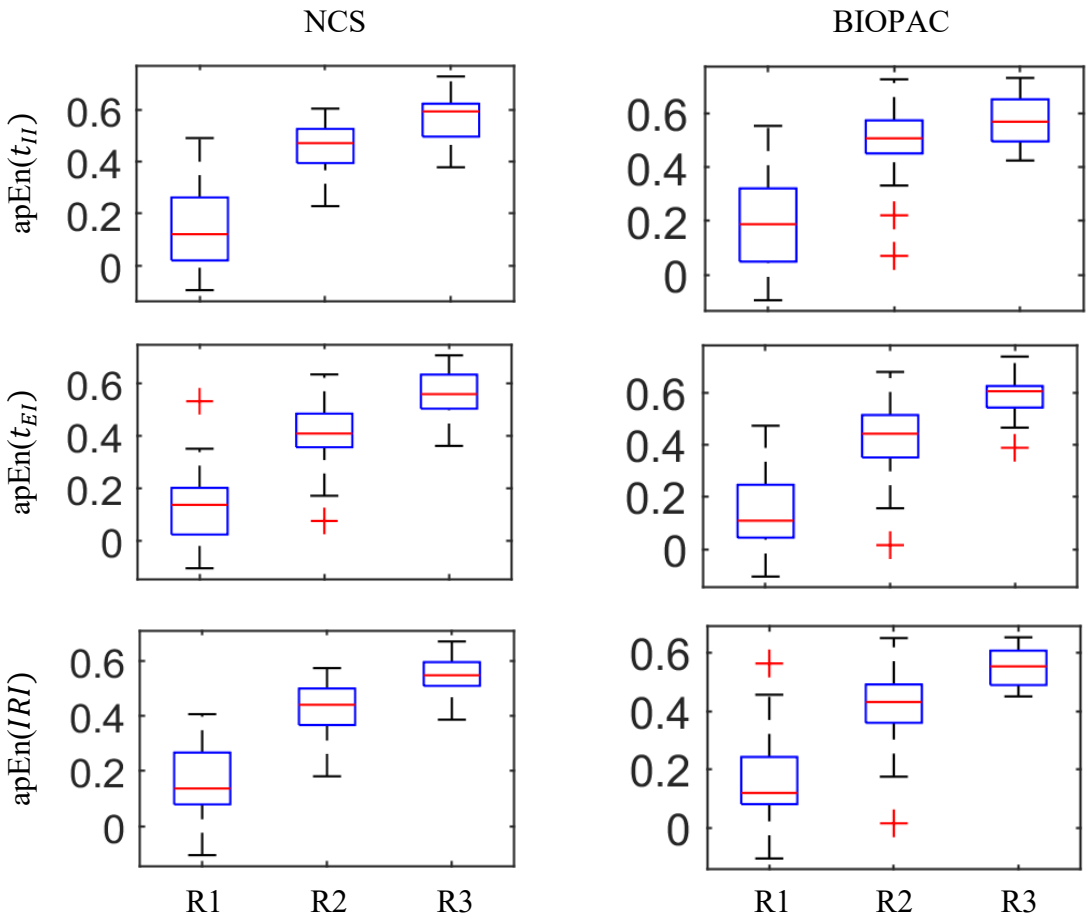
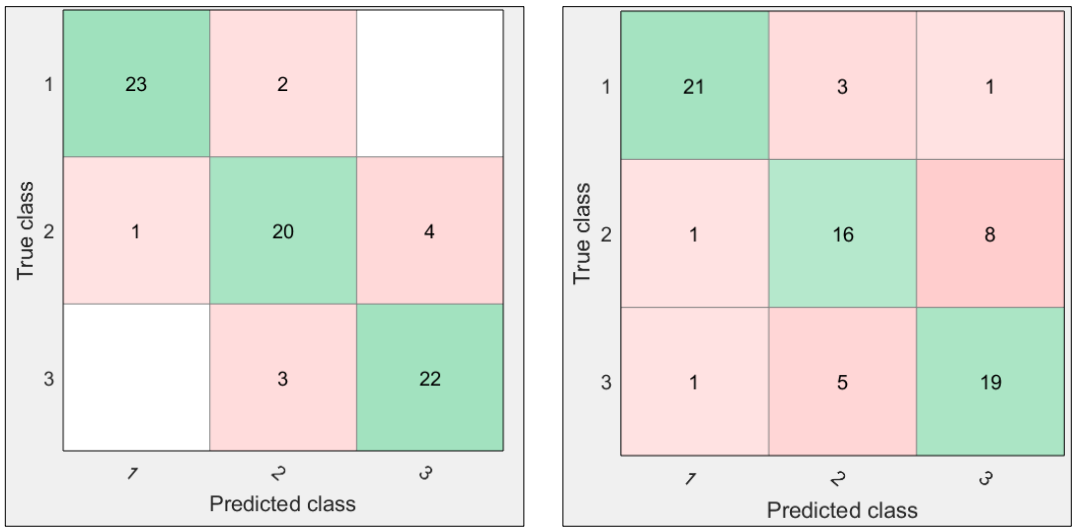


Figure 4.6 Approximate entropy (ApEn) feature values across different classes on the entire dataset for both NCS and BIOPAC. Large window sizes have been used.

(features) i & j, given as $I(i,j)$, and gives the level of similarity between the variables. The idea of minimum redundancy is to select the variables that are mutually maximally dissimilar. The two mutual information-based criteria need to be simultaneously optimized to generate results. The first feature is the one with the highest $I(y,i)$, then rest are determined incrementally by optimization using a heuristic algorithm.

4.4.2 Classification algorithm

Different classification algorithms have been implemented including support vector machines (SVM), kNN, trees, and their ensembles. The reported results are



(a) NCS (b) BIOPAC

Figure 4.7 5-fold CV performance on the entire dataset using long-window respiratory features, for (a) NCS and (b) BIOPAC. The three classes are R1, R2 and R3.

considered with the best accuracy. For longer window features, ensemble subspace discriminant achieves the best results. For ultra-short feature analysis, kNN works best. 5-fold cross-validation (CV) has been implemented for model training, and performance analysis has been done on an unseen holdout set. The results show high accuracy for both subject-dependent and independent results.

4.5 Results

4.5.1 Feature comparison and subject-dependent results

To find out information within respiration, we first restrict the analysis to the respiratory features. With a larger window size, only a subset (11) of all the features are used, with the approximate entry being the maximally contributing feature. The approximate entropy estimated on $t_{EI}, t_{II},$ and IRI are shown in Fig. 4.6, derived from both NCS respiratory waveforms and BIOPAC chest belt waveform at the abdomen.

Table 4.1 Features derived from respiration

$\text{mean}(t_{II})$	$\text{mean}(t_{EI})$	$\text{mean}(IRI)$	$\text{mean}(BR)$
$\text{mean}(TV)$	$\text{mean}(t_{EI}/t_{II})$	$\text{mean}(SD1_{II})$	$\text{mean}(SD1_{EI})$
$\text{mean}(SD1_{IRI})$	$\text{mean}(SD1_{TV})$	$\text{mean}(SD2_{II})$	$\text{mean}(SD2_{EI})$
$\text{mean}(SD2_{IRI})$	$\text{mean}(SD2_{TV})$	$\text{std}(t_{II})$	$\text{std}(t_{EI})$
$\text{std}(IRI)$	$\text{std}(BR)$	$\text{std}(TV)$	$\text{std}(t_{EI}/t_{II})$
$\text{std}(SD1_{II})$	$\text{std}(SD1_{EI})$	$\text{std}(SD1_{IRI})$	$\text{std}(SD1_{TV})$
$\text{std}(SD2_{II})$	$\text{std}(SD2_{EI})$	$\text{std}(SD2_{IRI})$	$\text{std}(SD2_{TV})$
$\frac{\text{mean}(SD1_{II})}{\text{std}(t_{II})}$	$\frac{\text{mean}(SD1_{EI})}{\text{std}(t_{EI})}$	$\frac{\text{mean}(SD1_{IRI})}{\text{std}(IRI)}$	$\frac{\text{mean}(SD1_{TV})}{\text{std}(t_{TV})}$
$\frac{\text{mean}(SD2_{II})}{\text{std}(t_{II})}$	$\frac{\text{mean}(SD2_{EI})}{\text{std}(t_{EI})}$	$\frac{\text{mean}(SD2_{IRI})}{\text{std}(IRI)}$	$\frac{\text{mean}(SD2_{TV})}{\text{std}(t_{TV})}$
$ApEn(t_{II})$	$ApEn(t_{EI})$	$ApEn(IRI)$	$ApEn(TV)$
$LF(t_{II})$	$LF(t_{EI})$	$LF(IRI)$	$HF(t_{II})$
$HF(t_{EI})$	$HF(IRI)$		

Fig. 4.7 shows the confusion matrix of average 5-fold CV performance on the entire NCS data, showing 86.7% accuracy for R1, R2, and R3 classification. The corresponding BIOPAC accuracy is 74.7%. For the ultra-short window, all 46 features are derived as shown in Table 4.1, and good results are seen when only the top 10 features selected from the MRMR algorithm are used. Fig. 4.8 shows the feature importance versus the rank plot with the top 6 features being: $\text{mean}(SD1_{II})$, $\text{mean}\left(\frac{t_{EI}}{t_{II}}\right)$, $\text{mean}(SD2_{IRI})/\text{std}(IRI)$, $\text{mean}(TV)$, $\text{mean}(SD2_{TV})/\text{std}(TV)$, $apEn(t_{EI})$. Here, $SD1_{t_{II}}$ is written as $SD1_{II}$ for simplification.

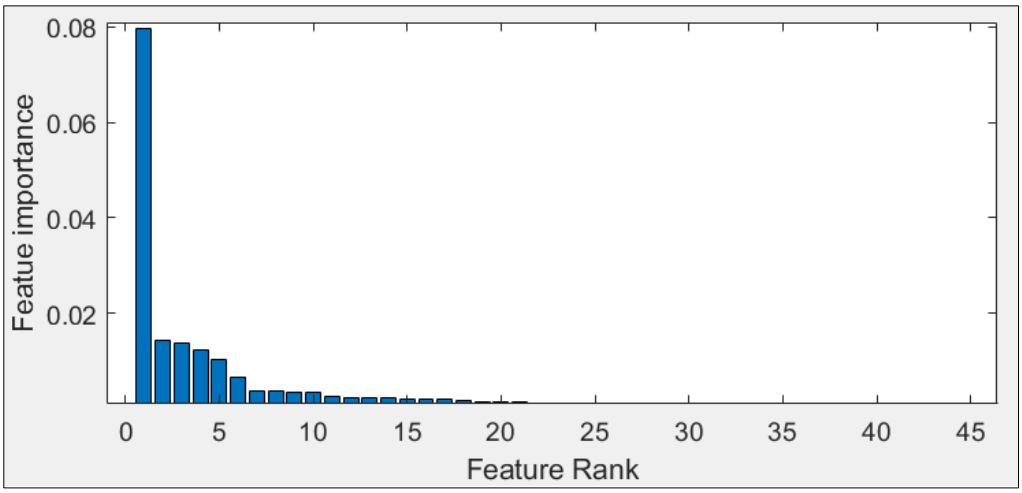
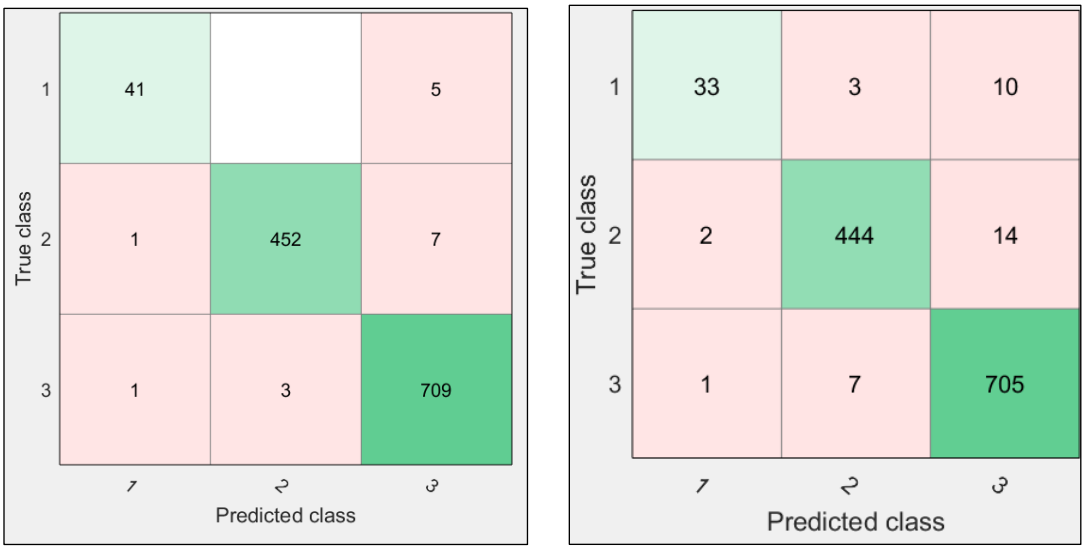


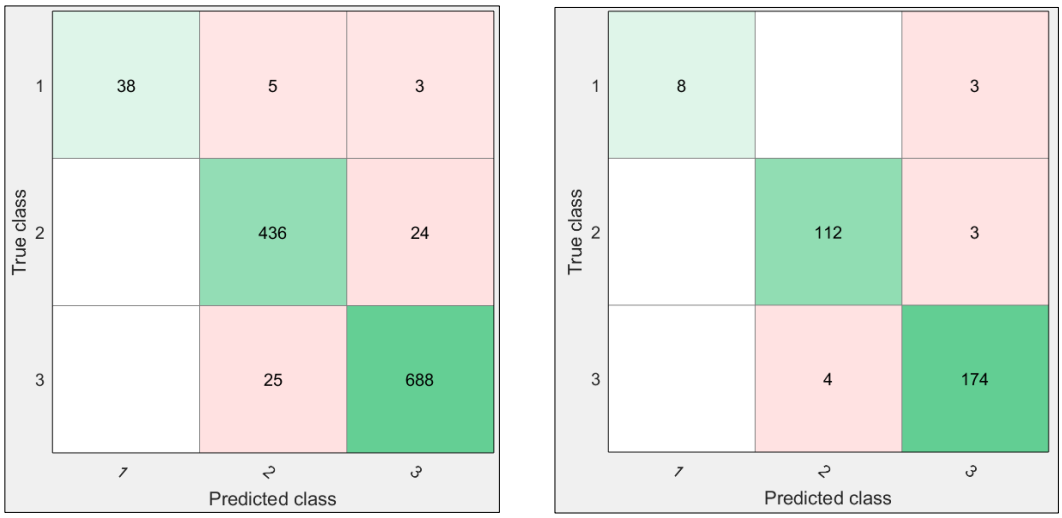
Figure 4.8 MRMR feature ranking results for ultra-short length features. Only top 6 ~ 10 have a high importance.

Fig. 4.9 shows the 5-fold CV accuracy performance with NCS and BIOPAC data using ultra-short windows, resulting in an accuracy of 98.6% and 97%, respectively. Fig. 4.10 shows the NCS performance with only the top 10 features on training data and 25% randomly selected holdout data for the same window size.



(a) NCS (b) BIOPAC

Figure 4.9 5-fold CV performance on the entire dataset using ultra-short respiratory features, for (a) NCS (accuracy: 98.6%) and (b) BIOPAC (accuracy: 97%).



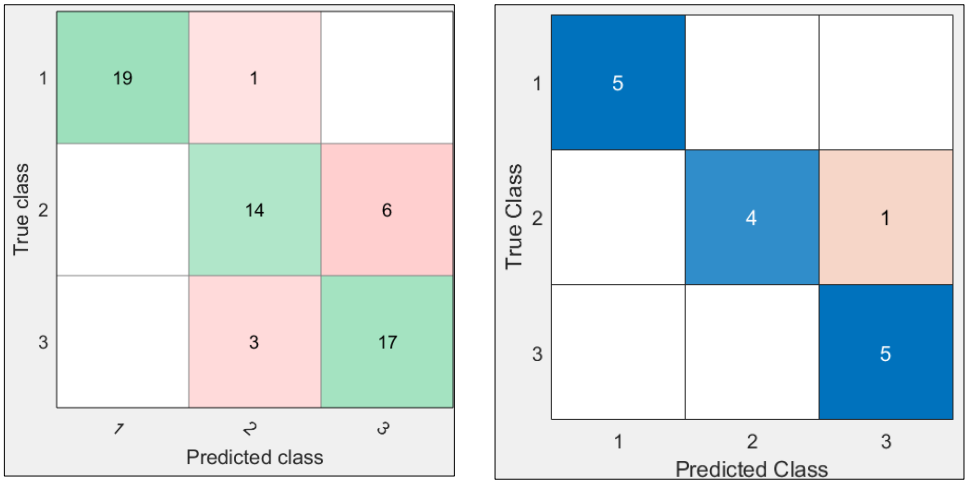
(a) NCS 5-fold CV (b) NCS 25% holdout data

Figure 4.10 Using top-10 features for training (a) and testing (b) on the NCS derived respiration features on ultra-short window.

In the next step, HRV features are also estimated considering only two classes R2 and R3. This provided similar training performance, but slightly poorer performance on the holdout set than the respiration features, as shown in Fig. 4.11.

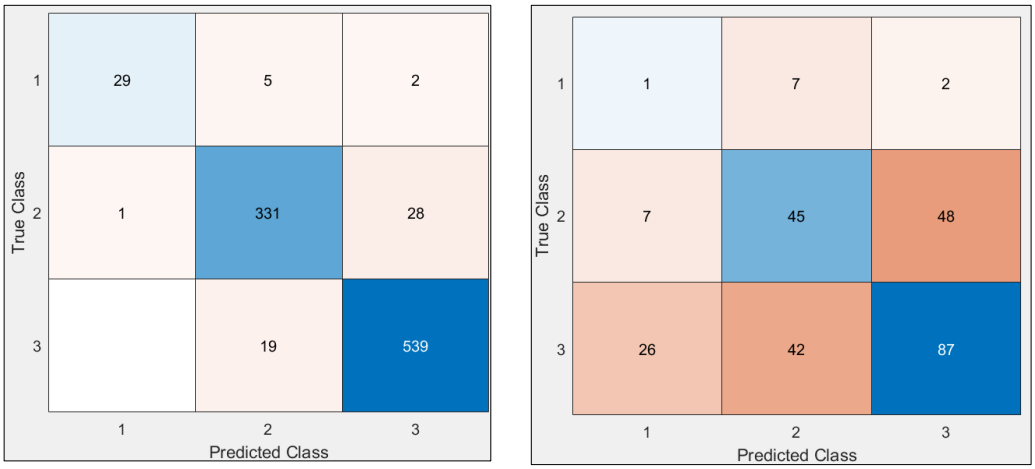
(a) NCS HRV (b) NCS Respiration

Figure 4.11 25% Holdout test dataset performance for R2 and R3 (relaxation and attention) classification using ultra-short features from (a) NCS HRV and (b) NCS respiration data.



(a) NCS CV training performance (b) Testing on unseen participants

Figure 4.12 Subject-independent train-test results on the NCS data with longer feature windows, showing very high accuracy.



(a) NCS CV training performance (b) Testing on unseen participants

Figure 4.13 Subject-independent train-test results on the NCS data with ultra-short feature windows.

4.5.2 Subject-independent results

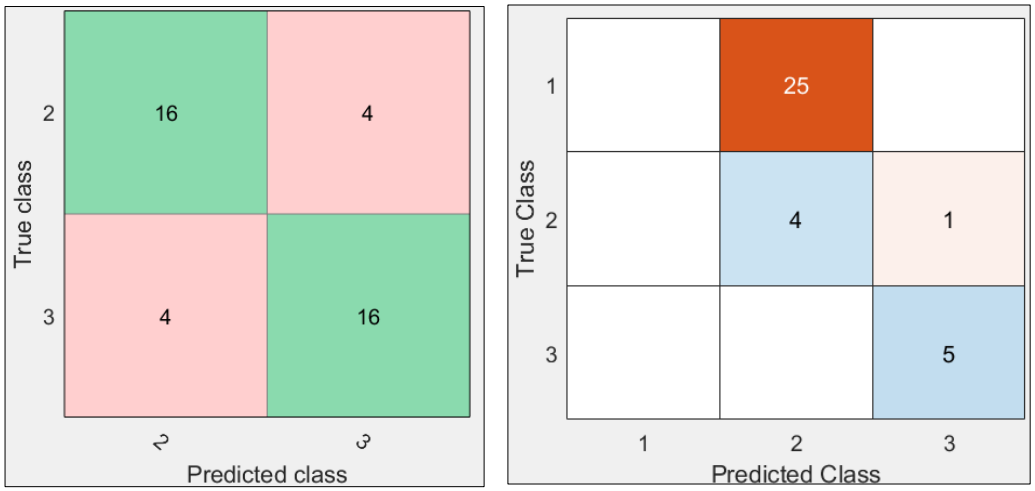
Subject independent tests were performed with both the feature lengths. 20 participants were used for training and the remaining for testing. In this case, results are still good for longer feature lengths as shown in Fig. 4.12. However, for ultra-short

features, performance degrades to 50% test accuracy which is better than by chance (33%) but still poor as shown in Fig. 4.13. This performance needs to be further investigated.

4.5.3 Effect of controlled respiratory pattern

While R1 is not part of the relaxation/attention test, the participants are asked to perform normal breathing (with no deep breaths or breath holds) for 2 min. Further normal breathing does not require excess voluntary attention and is not stress-inducing. This controlled breathing pattern is similar to a relaxation state and tests are performed to verify this intuition. Fig. 4.14 (a) shows the CV accuracy of 80% when only R2 and R3 data from 20 participants are used for training the classifier. Then, R1 of all the participants and the remaining 5 subjects is used as the holdout test set. Fig. 4.14 (b) shows the results, when all 25 cases, for seen and unseen participants are classified as R2, thus validating the hypothesis.

4.6 Discussion



(a) NCS CV training performance

(b) Testing on unseen participants and R1 data

Figure 4.14 Validating normal breathing routine (R1) classification as relaxation routine (R2) when only two classes are considered.

The ability to recognize attention and relaxation states is a major part of emotion detection, which can enhance safety measures in place for critical conditions and help reduce the overall stress level. This thesis has presented a general discussion of what is "attention" and its psychological perspective, as well as neuroscientific research linking emotion and attention. This has allowed us to generalize attention or vigilance, and relaxation and study them using a two-dimensional model, without limiting to stress or fatigue detection that are mostly treated as one-dimensional.

Further, the novel NCS sensor technology has been used to monitor both respiratory and heartbeat patterns. NCS presents a unique advantage that the sensor can operate in a passive mode with only a Tx antenna on the chest [45], thereby making the setup more comfortable for certain environments like driving, while still giving strong heartbeat and respiratory signals. Other furniture integrated versions can also be adopted with some tradeoff in relative sensor position and heartbeat sensing accuracy. Our results show significant accuracy with only respiration signals as well. While the presented HRV estimation algorithm performs well for some cases, with sensor placement variation, performance can degrade for some cases, resulting in lower accuracy than expected. The addition of noninvasive pulse oximetry can also be done in the future to possibly improve the overall estimate.

The feature window length and system response time to identify attention needs to be considered. The limitation of using respiratory features include1) learning slow or fast breathing requires a larger time window compared to HRV change, and 2) sudden inhalation or exhalation is a one-time event on a very small time scale. The feature identification on the respiration data can be further improved to incorporate these

variations. Traditionally, HRV features have been widely investigated for long-term (24 h) and short-term (> 5 min), and the reliability of ultra-short HRV (< 5 min) is not completely validated [166], but have been increasingly used for faster responses [167], [168]. The statistical features are also impacted by the window length and maybe highly biased with shorter lengths. ApEn has shown to be unreliable for excerpts of <3 min [169]. These could be the possible reasons for poorer performance on the shorter interval of 90 s compared to the entire window lengths. Another possible reason is mislabeling and misclassification of transition periods. When estimated on the larger window, it results in averaging out the transition period information. However, shorter windows might be incorrectly labeled at the transition period from relaxation to attention, leading to poor learning, as well as poor test results in these instances. This needs to be further studied and routines can be made longer for future studies to drop such transition periods.

The results presented in this work have been collected on 22 subjects and extensive validation has been performed on unseen test data, unseen routines, as well as entirely unseen participants. With smaller window sizes, this works well if baseline information is collected for each subject. Picard et al [170] considered this phenomenon in detail, with one subject selecting their own images to induce different emotions daily over multiple weeks. This showed more variation in day-to-day versus each emotion on the same day, showing huge baseline variation. Thus, future studies are required to consider daily variation in attention and relaxation classification.

The above baseline phenomenon needs to be compared with human perception during the experimental data collection. This can only be estimated by the user's self-

report. In this work, participants were asked to fill a survey indicating their relaxation and stress levels on a scale, as well as giving brief descriptions. The term stress was used instead of attention, as attention is more abstract for users. During the relaxation routine, most people responded to feeling relaxed or sleepy. Mixed responses were derived during the attention routine. Some subjects got increasingly frustrated as they received negative feedbacks, while some were calmer as they were good at the exercise. The game level was easy so that participants could quickly learn it, however, for future studies, games should become adaptively complex to demand continuous attention of skilled players and reduce mislabeling. For the presented results, as the survey feedback is not explicit, user data was not included to update labels. Another way to indicate real attention levels could be the response time from the game which is shown to be related to attention and fatigue [102]. Li et al [171] used another concept of group-based individual response specificity (IRS) for user-independent emotion recognition. In psychophysiology, the idea of IRS is of a consistent response of an individual for different stimuli, thereby introducing difficulties in subject-independent emotion detection. Thus, for every new user, IRS levels are estimated and associated with the closest cluster of people, before emotion labeling. The best performance is achieved when the population is classified into 10 clusters, however, testing is only done with 30 subjects [171]. Thus, more subjects are needed to study the expected number of clusters and corresponding performance improvement.

Overall, we have presented a novel sensor setup for user-independent attention – relaxation classification using respiratory and heartbeat features, resulting in a good accuracy for real-life implementation. While there has been a lot of research in the area

of emotion recognition, results must always be looked at carefully to understand the consideration of participant and baseline variations in the study. Our approach shows good results for unseen participants with sufficient window sizes. With the unique advantage of choices between passive and active NCS setups, the presented monitoring setup is suitable for a wide range of applications, from soldiers, to drivers and pilots, and even hospitals to detect accurate attention, vigilance, or stress versus relaxation states.

CHAPTER 5

INDOOR PASSIVE UHF RF-IMAGING

Recent years have seen a rapid increase in innovation focusing on smart buildings – from audio recording based smart interactive devices to camera-based security. Passive occupant monitoring is an important aspect of smart buildings that can be used to respond to user needs by learning accurate behavioral features, respond to emergencies, as well as save energy by smart HVAC control. This chapter focuses on using a UHF RF system for imaging indoor occupants. However, unlike the traditional outdoor radar applications, this setup suffers from nonideal near-field signal artifacts, strong multipath reflections and is even restricted in terms of bandwidth and spatial diversity of antennas – to limit the overall system costs.

This chapter discusses the traditional RSSI absorption/attenuation based signal model and both phase and RSSI comprising reflection-based model. The latter does not required occupant to be in LoS of a link, and allows higher resolution, but is also more prone to phase noises; thus, an improved calibration algorithm is proposed. Work has been done in combining both the models to get an improved result in high-noise environments. The experimental and simulation results are compared, and the real-life scenario and noises are studied in detail. The final image is generated by solving the linear inverse problem using proposed sparsity-approximation algorithms [205], which leverages the emptiness of the room to solve a highly underdetermined equation. Results show a performance comparison of traditional vs proposed solutions for both experimental and simulation cases.

5.1 Indoor RF propagation

In an indoor environment, the emitted electromagnetic (EM) waves have a complicated signal model from the transmitter (Tx) to the receiver (Rx), with reflection from surrounding walls and furniture resulting in multipath fading, and shadowing from obstacles resulting in attenuation or complete LoS blockage. Fig. 5.1a depicts a 2D

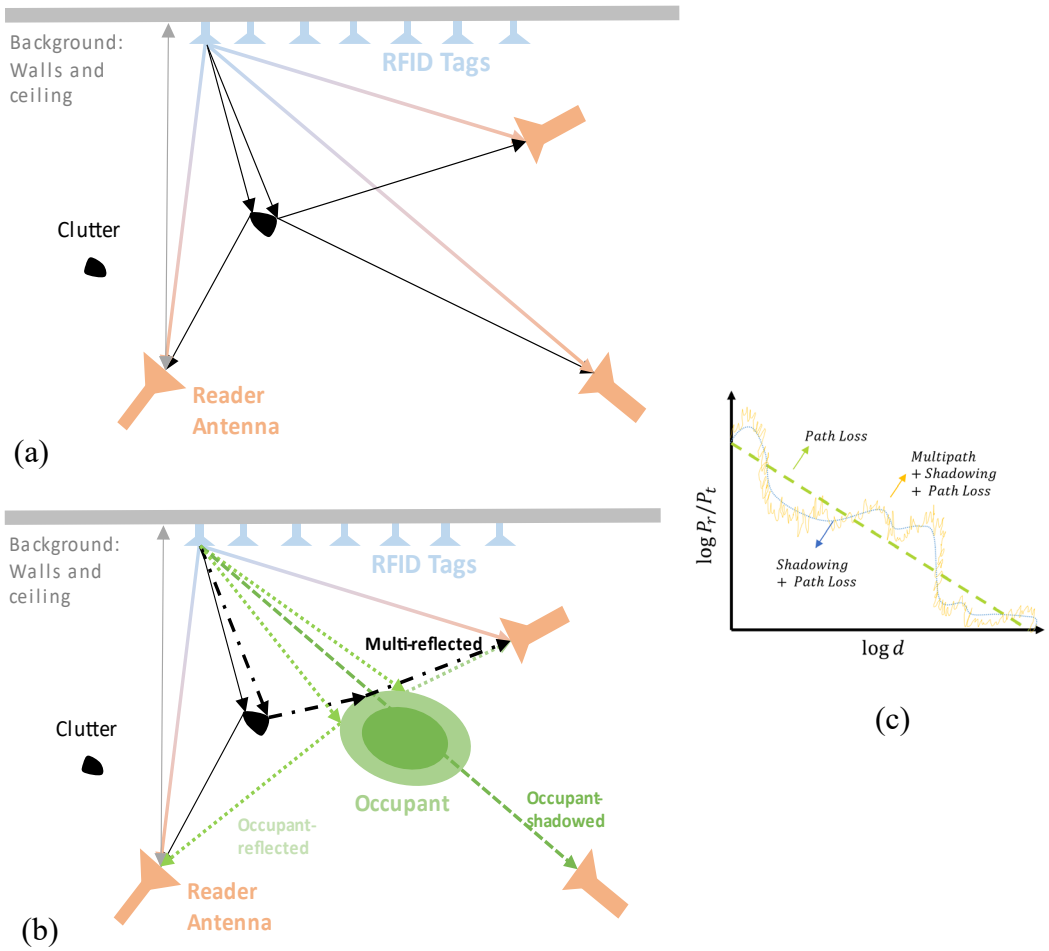


Figure 5.1 Illustration of RFID tags, reader antennas and backscattered signals from tags to reader antennas. (a) The unoccupied room with LoS from tag to reader and inherent multipath from wall and furniture (clutter). (b) Signal paths with an occupant, showing multipath: occupant reflection (light-green dotted), multi-reflection (black dash-dotted), and occupant shadowing (green dashed). (c) Effect of shadowing and multipath noise on the received power as a function of distance between Tx – Rx. Path loss is shows with a dashed green line. The effect of shadowing is shown with a dotted blue line, and with addition of multipath it gets even more noisy (yellow line).

diagram showing the backscattered uplink from tag to reader Rx antenna, including the LoS and multipath from walls and furnishing. The LoS path observes path loss with amplitude attenuation (α) given by Friis free-space transmission formula:

$$\alpha_{tr} \propto \left(\frac{\lambda}{4\pi d_{tr}}\right). \tag{5.1}$$

where λ is the wavelength and d_{tr} is the distance between the tag Tx and reader Rx antennas. The phase ϕ_{tr} accumulated on this path is:

$$\phi_{tr} = \frac{2\pi d_{tr}}{\lambda} \tag{5.2}$$

The multipath results in additional signals at Rx, often weaker than LoS, which can increase or decrease the overall RSSI, depending on the phase difference resulting from the additional distance traveled. This is often described by the conventional Rician fading channel [58], [172]. Occupant presence in the room can cause further RSSI variation or even complete blockage of the LoS paths and additional multipath, as shown in Fig. 5.1 (b). The indoor signal now follows more closely to the Rayleigh fading channel that assumes no dominant path, thus with multiple occupants, neither Rician nor Rayleigh model fits. Fig. 5.1 (c) plots the log of received power /transmit power (P_r/P_t) as a function of increasing distance d between Tx – Rx, which shows linear path loss, and received signal in presence of shadowing and multipath [173]. This shows serious RSSI deviation in the presence of shadowing and noise in the case of multipath.

5.2 Occupant imaging setup design

As shown in Fig. 5.1 above, we have selected a multistatic and multi-frequency system deployment using commercial RFID tags for occupant monitoring in the ISM frequency band of 902 – 928 MHz. As passive tags are low-cost sensing points, the

number of tags, $L \gg M$, the number of reader antennas. The previous work in our group [77] showed a $1m \times 1m$ 2D imaging setup for small less-reflective objects, using wired harmonic Tx-Rx setup and point-sized scatterers. Fig. 5.2 shows the generalized setup considered by that work, with LoS path from l^{th} Tx to m^{th} Rx, with l and m in the range $[1, L]$ and $[1, M]$, respectively. The P scatterers in the bounding box backscattered the RF signal transmitted at the frequency f_n with $n \in [1, N]$. The locations of Tx and Rx antennas and scatterers were defined with respect to the origin O with the vector r. The model ignored the attenuation with the point-sized scatterers, and the backscattered signal $y(r_l, r_m, r_p, k_n)$ was written as a function of location vectors (r_l, r_m, r_p) and the wavenumber $k_n = (2\pi f_n)/c$ as

$$y(r_l, r_m, r_p, k_n) = \iiint_{|r_l|<|r_{max}|} \rho(\mathrm{p})\, e^{-jk_n r(l,m,p)} dxdydz. \tag{5.3}$$

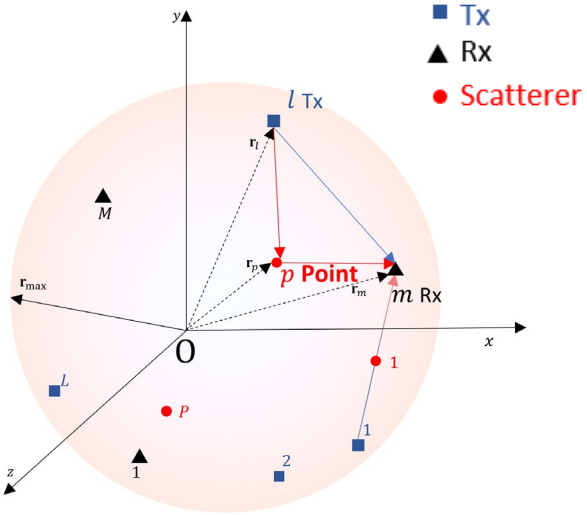


Figure 5.2 Generalized imaging system schematics showing the direct line-of-sight path from l^{th} Tx antenna to m^{th} Rx antenna, and the path reflected from p^{th} point scatterer. r_{max} is the radius of the imaging capture volume.

The integration was computed over the imaging volume, where (x, y, z) were bounded such that objects with reflectivity $\rho(p)$ were within a sphere of radius $|r_{max}|$. $\boldsymbol{r}(l, m, p)$ described the signal path from l^{th} Tx to p^{th}scatterer and then backscattered to m^{th} Rx.

However, the problem of occupant imaging is more complicated with large non-point scatterers as shown in Fig. 5.1. In this paper, the RF imaging setup is augmented to a wireless 3D imaging setup with commercial RFID tags, using improved noise-canceling calibration, considering cases of LoS attenuation and blockage, and sparsity-leveraging imaging algorithm. The prototype and feasibility studies are done extensively using a simulation environment in CST Microwave Studio to study the effects of large occupant (or any other object to be imaged) sizes. Two setup sizes are considered in both simulation and experiments – a $1/6^{th}$ scaled model, and a true room-scale model. While the former model is for demonstration, it can also be used for imaging high dielectric objects in closed containers, faults or crack detection, and gaming for hand-gesture recognition.

5.2.1 Simulation design

CST Microwave Studio is a 3D EM wave analysis software that uses methods such as the finite element method (FEM) for high-frequency problems with time or frequency domain solvers. In this work, a time-domain solver is used that performs broadband calculation of S-parameters from one single calculation run by applying DFT to time pulses.

The computation effort and the total calculation time in the time-domain solver is related to the total number of timesteps to be calculated and is defined by two factors: 1) The minimum timestep width Δt and 2) The ending time t_{end}. The former, Δt is

proportional to the smallest spatial mesh step and can be improved by avoiding unnecessarily small steps. However, to simulate complex antenna structures, such as the ones above with exceedingly small size along any one dimension, or different dielectric layers, the mesh size cannot be made arbitrarily large. The mesh size is also related to the smallest wavelength considered; thus, simulation upper frequency should be strictly limited to frequencies of interest. The t_{end} is related to excitation signal duration, which is smaller for a larger bandwidth. However, this setting does not seem to affect the overall time in test cases and is not tuned. Other stopping criteria include accuracy (the solver stops when the remaining energy decreases by this amount compared to the

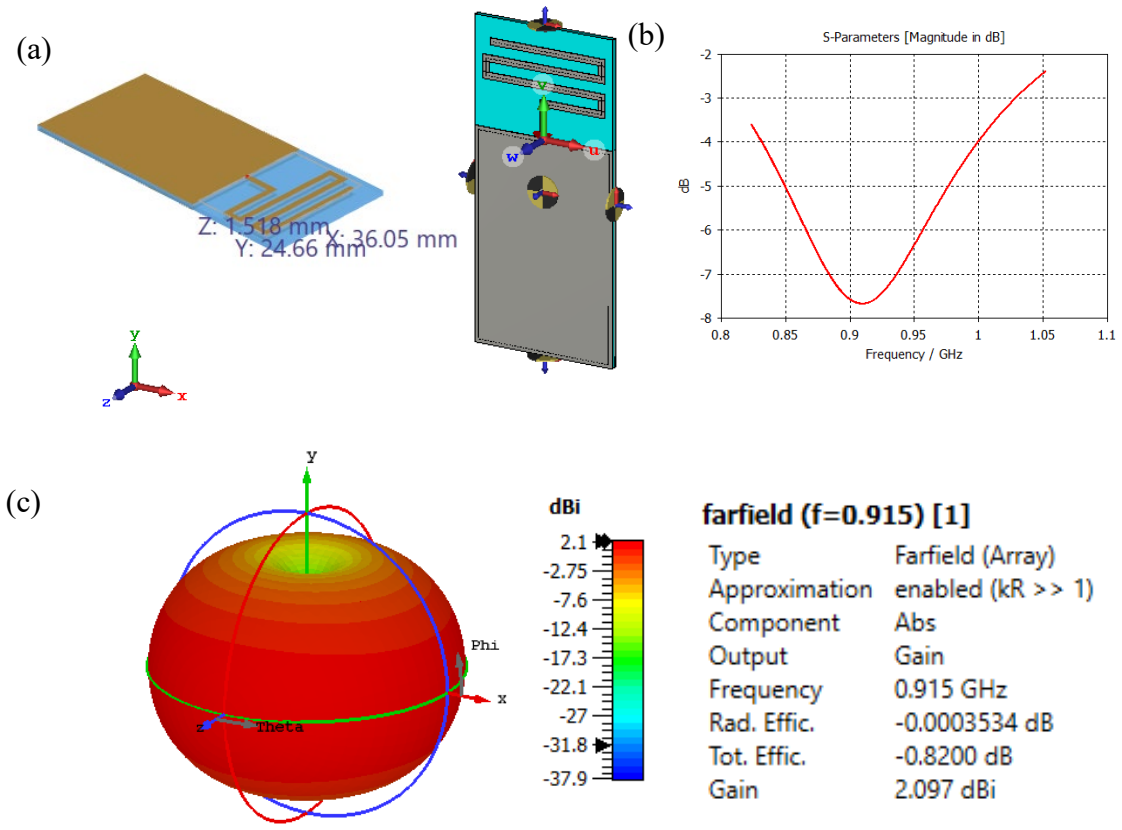


Figure 5.3 Printed meandering antenna. (a) Antenna design and specifications. (b) Antenna S_{11} characteristics with minimum S_{11} of -8 dB at 915 MHz. (c) Omnidirectional far-field radiation pattern at 915 MHz showing the antenna gain.

maximum energy), which is selected to be -40 dB to get reasonably low "truncation error."

Discrete ports are used as signal sources to estimate the S-parameters between different antenna pairs. These are modeled by a lumped element consisting of a current source with an inner impedance in parallel that absorbs power. A Gaussian pulse is used as an input signal for the time-domain simulation, that allows: 1) limited bandwidth, 2)

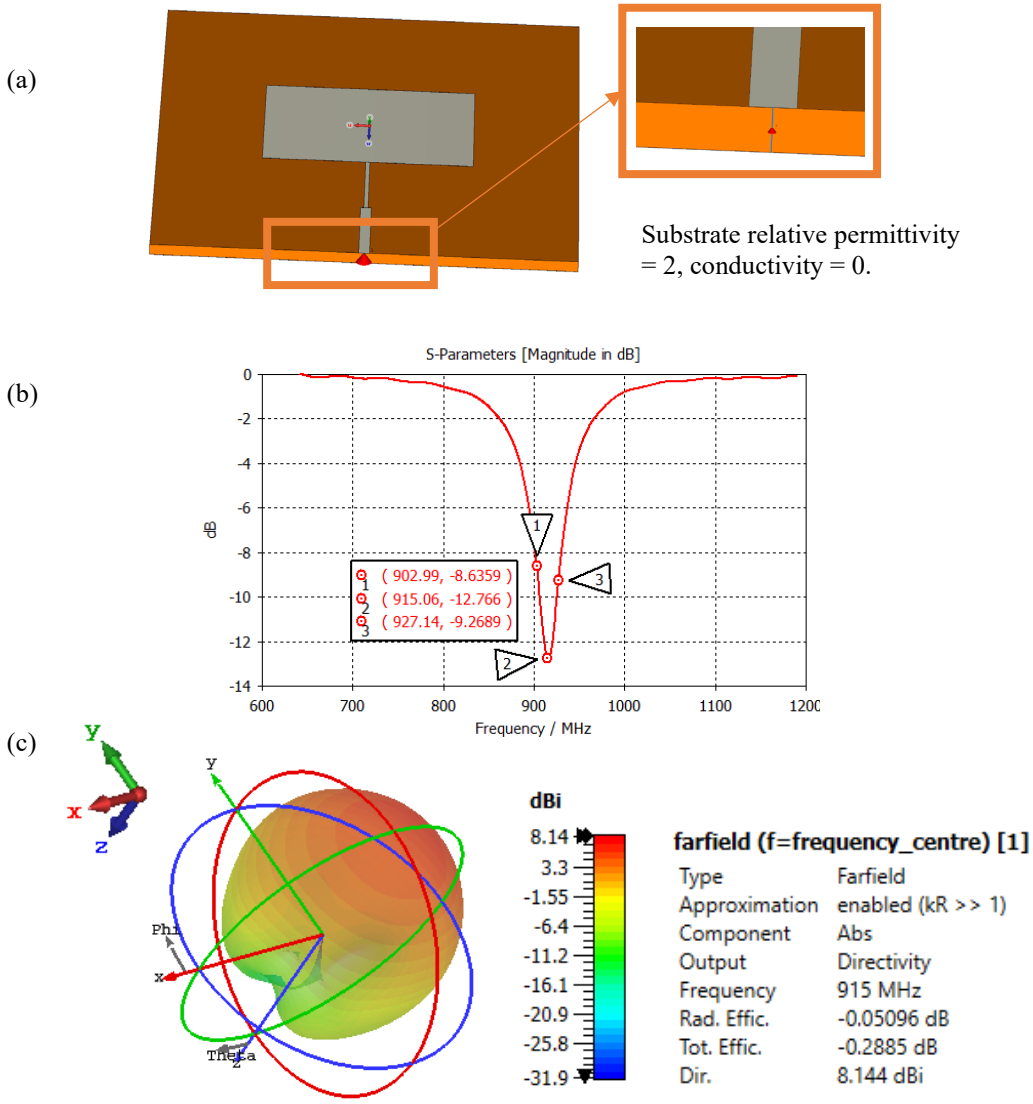


Figure 5.4 Patch antenna simulation. (a) Antenna design imported from Antenna Magus. (b) The S_{11} characteristics showing center frequency of 915 MHz. (c) Radiation pattern showing antenna directivity and gain of 8 dBi.

transformation to frequency domain maintaining Gaussian shape. As discussed earlier, bandwidth selection influences the simulation time – narrow bandwidth requires a longer signal.

To simulate the experimental system, the first step is antenna design. Fig. 5.3 (a) shows a printed meandering antenna structure from the Antenna Magus simulation tool, tuned to the desired frequency band. The corresponding S_{11} characteristics and radiation patterns are shown in Figs. 5.3 (b) and (c). This antenna is paper-thin, similar to RFID tags with a monopole radiation pattern. The reader antenna is a patch antenna of gain around 8 dBi, shown in Fig. 5.4.

Both antennas are very close to the ones used in the experimental system, however, the numerical efforts required to simulate setup with multiple of these antennas distributed in a room need to be considered carefully. Fig. 5.5 shows the smallest mesh sizes in red color for the rectangular patch antenna in Fig. 5.4, which is very small. To keep simulation time reasonably small to the order of a few hours for a

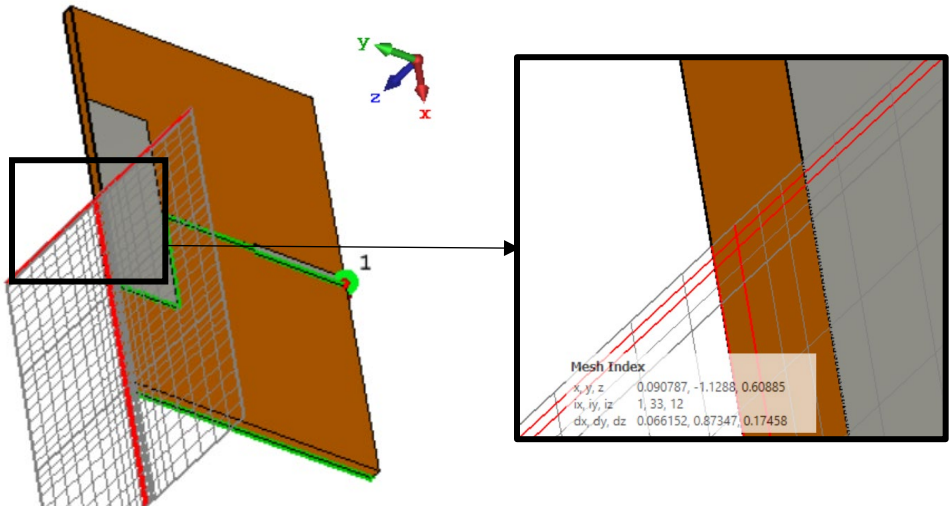
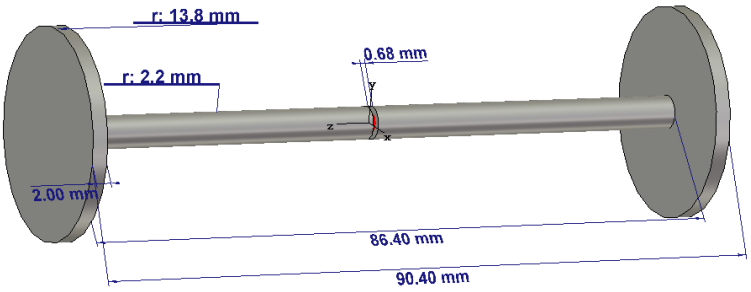


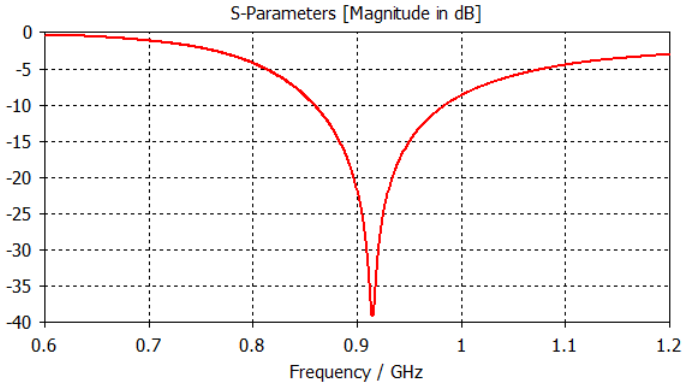
Figure 5.5 Smallest mesh size in CST simulation is directly related to the structure design, material and wavelength. The inset shows zoomed in view of the smallest mesh cell in red boundary, with dx, dy, dz showing its length along different dimensions in cm.

real-room simulation, the above antenna designs were replaced by loaded dipole antennas. Fig. 5.6 shows a loaded dipole antenna design made of a perfect electric conductor (PEC), with infinite electrical conductivity, and its S_{11} characteristics in a standalone setup. The antenna radiation pattern is similar to that of a printed meandering

(a)



(b)



(c)

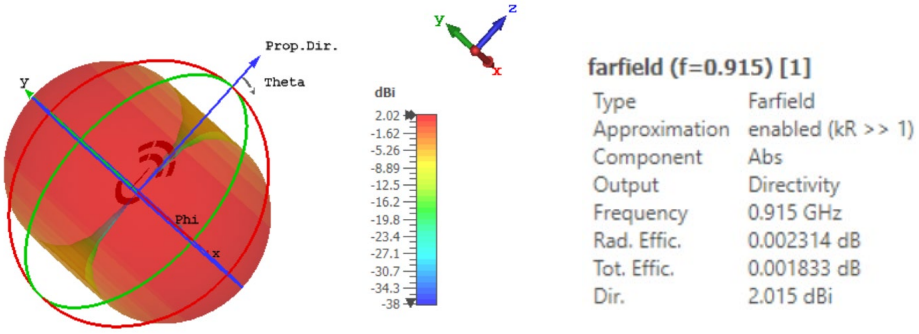


Figure 5.6 Dipole antenna design. (a) Antenna with all the dimension specifications. (b) S_{11} characteristics with center frequency at 915 MHz, designed to have a wide bandwidth, $S_{11} < -10$ dB between [0.86, 0.99] GHz. (c) Omnidirectional antenna radiation pattern, similar to printed meandering antenna.

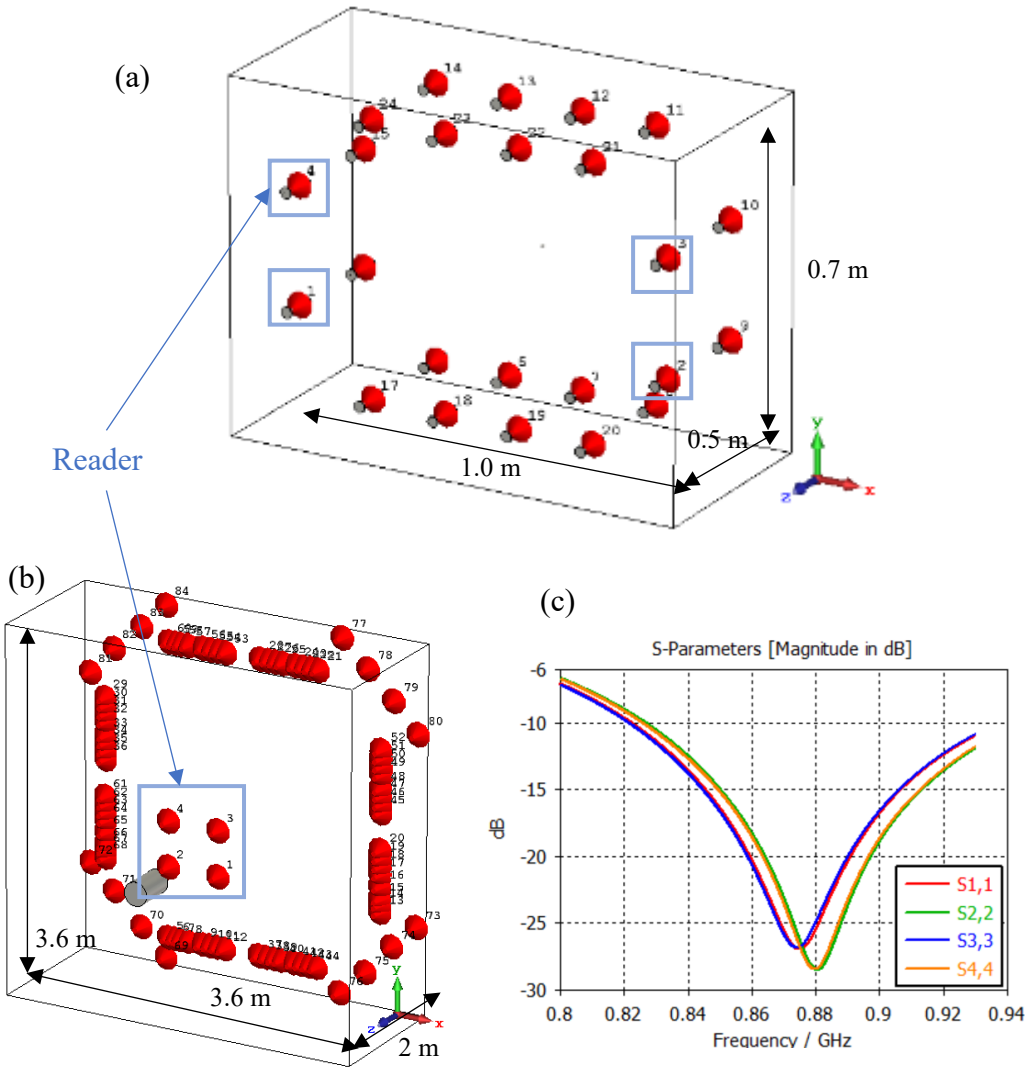


Figure 5.7 Simulation setups with dipole antenna. (a) $1/6^{th}$ scale room setup showing the area of size $1m \times 0.7m \times 0.5m$ with four reader antennas on the two opposite walls. Rest simulate RFID tags. (b) Real room of size $3.6m \times 3.6m \times 2m$, with reader antenna on the ceiling. Occupant's torso is simulated with a PEC cylinder of height 0.5 m. (c) The S_{ii} parameters with $i \in [1,4]$ for the four reader antenna corresponding to setup (b), with no PEC object.

antenna, so the substitution does not result in a drastic change. For the reader antenna, while the correct antenna behavior is not replicated, the overall general algorithm performance can be reasonably estimated.

Fig. 5.7 (a) shows a 1/6 scaled room setup surrounded by a dipole antenna, and Fig. 5.7 (b) shows a real-scale setup. The red cones indicate discrete antenna simulation

(a) (b)

(c)

Figure 5.8 The commercial RFID system. (a) Passive RFID tag. (b) A small patch antenna with metal ground. This small gain antenna was used for scaled room setup. (c) Impinj RFID reader.

ports, associated with each antenna, with reader antennas numbered 1 – 4. In the simulation, only a one-way path is simulated from Tx (tags) to four Rx (reader) antenna. As a human model is very complex, both in outline and material complexity, simple PEC cylindrical or ellipsoidal models are used in the simulation, that represent the reflective nature of the human body, without the attenuation. Keeping antennas close to each other and this PEC object might shift the antenna characteristics as shown in Fig. 5.7 (c) which shows S_{ii} parameters for the four reader antennas in a real room setup. Thus, interpretation of setup design from simulation to experiment needs to be carefully considered. Additionally, initial scaled setup tests were also done at the harmonic frequency band of 1.7-1.9 GHz, with dipole antenna modified accordingly, that

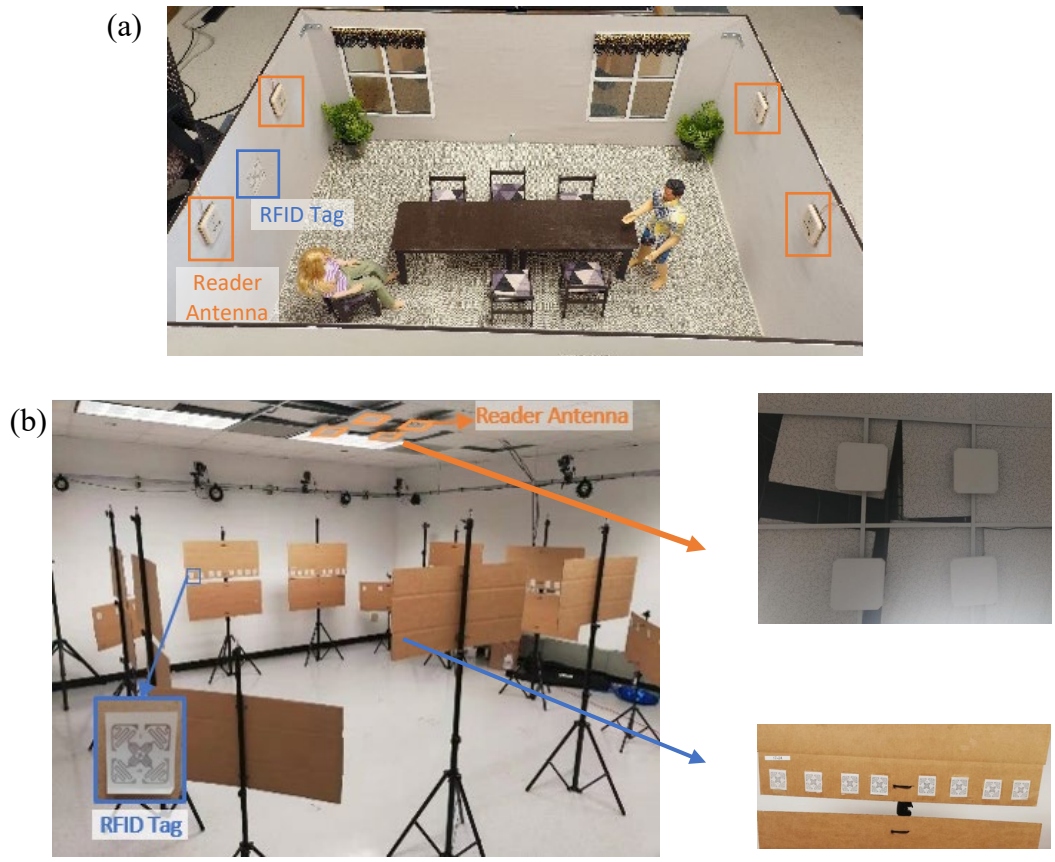


Figure 5.9 Experimental setups for occupancy monitoring with COTS RFID system. (a) A 1/6th scaled model of a conference room setup with four small patch reader antennas on the four walls and RFID tags under the wallpaper. (b) The room scale setup with RFID tags on the cardboard and reader antennas on the ceiling, also shown in the zoomed insets.

demonstrates the validity of the proposed algorithm and approaches at different UHF frequency bands.

5.2.2 Experimental setups

The experimental setups were designed after validating the simulation performance in both the scaled and real-size domain. A commercial RFID system has been used, with tag and reader as shown in Fig. 5.8. we employed ambient passive RFID tags, without requiring any markers or devices on occupants. The test setup included an Impinj Speedway R420 RFID reader that can be connected to four reader antennas as

shown in Fig. 5.8 (c), following time-division multiplexing. The reader acts as a transceiver and periodically emits RF signals in the 902 – 928 MHz range with 50 frequencies separated by 0.5 MHz. This EM energy is harvested by the tags to activate and backscatter the ID-modulated signal. The passive tags cost less than 10 cents each with a reading range of ~10 m, and thus, can increase the spatial RF link coverage without adding much to the overall cost in deployment and maintenance. The number of tags is only limited by the computation requirements and reader distance. Multiple readers can work collaboratively for larger spaces [174]. The reader collects tag IDs and the corresponding RSSI and phase values for each link, which is composed of a tag and a reader Rx antenna at a given frequency.

Fig. 5.9 (a) shows the $1/6^{th}$ scaled set up with a conference room design using wooden tables and chairs in a cardboard box of size $1m \times 0.7m \times 0.5m$. The human dolls had silicone exterior with metal joints and skeleton with high reflectivity. The position of the four reader antennas are indicated, tags were placed under the wallpaper, with one tag shown for reference. The real-sized lab test setup is shown in Fig. 5.9 (b) with 80 RFID tags arranged on cardboard. Except at the corners, tags are placed at a height of 1.24 m, with a minimum distance of 10 cm between the neighboring tags. The 16 corner tags are placed at a height of 1 m. Four reader antennas are placed on the ceiling. The next section discusses the signal model and calibration for this wireless passive setup.

5.3 Signal model and background calibration

While occupancy information is embedded in the phase and RSSI, additional signal components from the background furniture and walls remain, as well as the effects of cables and transceivers that can have frequency dependence. Also, the tag insertion loss and phase offset can vary with the incident power. As the phase has more serious uncertainties, it has been commonly ignored in most techniques. The traditional approach uses RSSI attenuation and is termed as radio tomography imaging (RTI), and suffers from poor RSSI resolution, and requires the object to be in LoS of the Tx – Rx path, that requires a dense network of antennas as shown in [65]. As humans also strongly reflect the RF signals, where the additional distance traveled is incorporated into the phase information, a reflection-based model can also be considered, such as for small-object imaging, without LoS blocking in a low-multipath environment [77]. This work has compared these two models and proposed improved calibration for the backscatter reflection model. Finally, preliminary work in the direction of combining both the models is also considered that gives superior results in high multipath cases.

5.3.1 RSSI attenuation based

Significant work in the area of RTI has been done in the past few decades [65], [76], [175], which shows RTI can work well for imaging objects of large sizes. To have sufficient LoS paths passing through the object of interest, a dense network is required

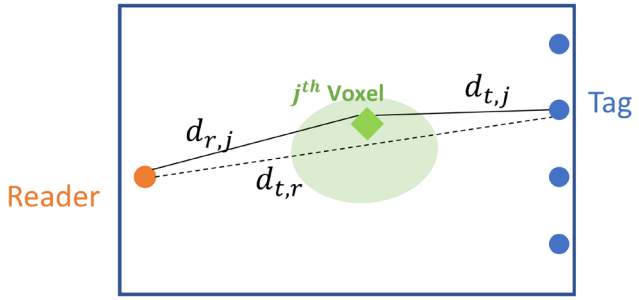


Figure 5.10 RSSI attenuation model in RTI.

which may be difficult without all active transponders such as RFID system with passive tags. This is demonstrated in [76], which shows increased imaging noise with RFID system for only one person.

The imaging problem is translated into a linear inverse problem. A simple measure of attenuation is the difference of RSSI (in dB) with and without the occupants' presence in the room, as shown below

$$z_{t,r}^{RTI} = z_{t,r}^{Occ} - z_{t,r}^{NoOcc} = Shadowing_{t,r} + \eta_{t,r} \tag{5.4}$$

where $z_{t,r}$ is the RSSI reading in dB for a tag – reader antenna pair, and the superscripts Occ and $NoOcc$ indicate the presence and absence of the occupants, respectively. And the attenuation or shadowing for each Tx – Rx pair is measured as the sum of attenuation in each voxel along the considered path. The weighting of each voxel is based on the linear ellipsoidal model in [65], as shown in Fig. 5.10, with d as the distance and x as the unknown attenuation at j^{th} voxel:

$$\sum_{j=1}^{N_{voxel}} w_{t,r,j} x_j, \qquad w_{t,r,j} = \frac{1}{\sqrt{d_{t,r}}} \begin{cases} 1 & d_{t,j} + d_{r,j} < d_{t,r} + \lambda \\ 0 & otherwise \end{cases} \tag{5.5}$$

where t, r correspond to tag and reader antennas. Thus, from (5.4) and (5.5), the model can be simplified as:

$$z^{RTI} = Wx + \eta \tag{5.6}$$

where z^{RTI} is a column vector of RSSI (dB) difference corresponding to each tag-reader antenna link of length $L \cdot M$, keeping earlier notations of the maximum number of tags and reader antennas corresponding to L and M respectively. x is also a column vector corresponding to an unknown attenuation value in each voxel in the area of interest, with length equal to the number of voxels, N_{voxel}. W is a rectangular matrix of

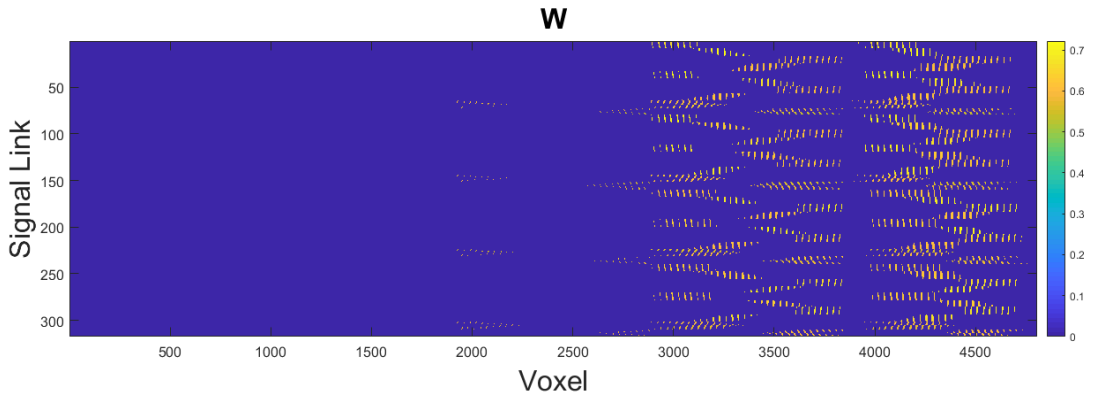


Figure 5.11 The W matrix for setup in Fig. 5.9 (b), with $L = 80, M = 4$, resulting in 320 (t, r) links. For voxel index < 3000 (x-axis in plot), or $height < 1.5\ m$, there are only a few links crossing the voxel, so image will be mostly 0s below that.

dimensions $L \cdot M \times N_{voxel}$ with columns corresponding to the weight of j^{th} voxel, $j \in [1, N_{voxel}]$ and rows corresponding to each (t, r) link, calculated according to (5.5).

Eq. (5.6) is a highly underdetermined problem and can be solved by various inverse problem solutions given in the next section. It should be noted that W matrix does not depend on the frequency, and thus matrix will have degeneracy if z_{RTI} include readings at multiple frequencies for the same tag – reader link, and may not help reduce ambiguity. Instead, z^{RTI} can be estimated as the mean across different frequencies to reduce some uncorrelated multipath noise. Further, RSSI readings from the reader are quantized and result in poor image resolution. Other practical concerns include relative tag – reader antenna placement constraints due to antenna beamwidth patterns and distance from the ground. With the reader antenna placed on the ceiling as shown in Fig. 5.9 (b), there are very few links with an occupant on the LoS, and mostly above height ~1.5 m, resulting in a highly sparse W matrix as shown in Fig. 5.11. This will result in a nearly 2D image.

5.3.2 Reflection based

This work focuses on reducing noise in the phase readings present in the considered experimental system due to 1) heavy multipath, 2) cable phase offsets, and 3) reader and tag circuitry phase. Here a background subtraction algorithm is proposed with additional corrections for these phase offsets from cables and reader circuitry lumped as ϕ_r, and tag circuitry, ϕ_t, in a heavy multipath environment. As LoS blockage and shadowing attenuation have a higher impact on RSSI, this calibration focuses on extracting the weaker signal reflected by the occupant with reflectivity γ, when LoS is not completely blocked. The LoS blocking can be detected when the signals at all frequencies experience attenuation in presence of the occupants. This derives from the fact that multipath fading can increase or decrease RSSI differently for uncorrelated frequency channels, but for LoS blockage most frequencies experience the attenuation or blockage. With significant LoS blockage and high multipath noise, the calibration proposed in [77] fails to work well.

In an unoccupied room, we first collect the complex signal y^{NoOcc} backscattered from L tags and received at M reader Rx antennas at N frequencies. Then, we collect the signal with occupants in the room, y^{Occ} , and perform background subtraction (BS) by:

$$y^{BS} = y^{Occ} - y^{NoOcc} \tag{5.7}$$

This gives us the occupant-reflected signal if there is no LoS blockage. The dotted light-green line in Fig. 5.1b shows the backscattered component involving: the reader (r_m), occupant (o_k), and tag (t_l), with m and l in the range $[1, M]$ and $[1, L]$, respectively. With $s(t)$ as the signal transmitted by r_m, the free space attenuation and phase accumulation (α, ϕ) at individual paths are indicated in the order of traversal as

$(\alpha_{or}, \phi_{or}): r \rightarrow o; (\alpha_{ot}, \phi_{ot}): o \rightarrow t; (\alpha_{ot}, \phi_{ot}): t \rightarrow o;\ (\alpha_{or}, \phi_{or}): o \rightarrow r$. The round-trip received signal can be expressed as:

$$y^{BS} = \beta_t^2 \cdot \gamma^2 \cdot \alpha_{or}^2 \cdot \alpha_{ot}^2 \cdot \exp\begin{pmatrix} -2j(\phi_{ot} + \phi_{or} + \\ \phi_t + \phi_r) \end{pmatrix} s(t) \tag{5.8}$$

Here β_t is the amplitude attenuation after tag modulation. We omitted the subscripts of $k, l,$ and m for convenience. The above equation shows unknown phase values ϕ_r and ϕ_t, estimated at the next steps using only y^{NoOcc}.

The round-trip LoS signal collected at the Rx of various reader antennas from the l^{th} tag, with multipath phase δ_{ml}, is given as:

$$\begin{aligned} y_{r_1 t_l}^{NoOcc} &= \beta_l^2 \alpha_{t_l r_1}^2 \cdot \exp\left(-2j\left(\phi_{r_1 t_l} + \phi_{r_1} + \phi_{t_l} + \delta_{1l}\right)\right) \\ y_{r_2 t_l}^{NoOcc} &= \beta_l^2 \alpha_{t_l r_2}^2 \cdot \exp\left(-2j\left(\phi_{r_2 t_l} + \phi_{r_2} + \phi_{t_l} + \delta_{2l}\right)\right) \\ &\vdots \\ y_{r_M t_l}^{NoOcc} &= \beta_l^2 \alpha_{t_l r_M}^2 \cdot \exp\left(-2j\left(\phi_{r_M t_l} + \phi_{r_M} + \phi_{t_l} + \delta_{Ml}\right)\right). \end{aligned} \tag{5.9}$$

In the above equations, $s(t)$ is common, and not mentioned explicitly. The phase ϕ_{t_l} is common across the receivers, and can be eliminated by taking $m = 1$ as the reference receiver to get the approximate estimate of the relative ϕ_m:

$$\eta_{r_{m1} t_l} = \left(y_{r_m t_l}^{NoOcc} \cdot \exp\left(2j\phi_{r_m t_l}\right)\right) \cdot \left(y_{r_1 t_l}^{NoOcc} \cdot \exp\left(2j\phi_{r_1 t_l}\right)\right)^* \tag{5.10}$$

where $(q)^*$ is the conjugate of q. For $m > 1, \eta_{r_{m1} t_l} \approx c_1 \cdot \exp\left(-2j\left(\phi_{r_m} - \phi_{r_1} + (\delta_{ml} - \delta_{1l})\right)\right)$, with c_1 as a constant. If the multipath signal is weak, the term $(\delta_{ml} - \delta_{1l})$ is expected to be randomly distributed and can be reduced by averaging over tags to give a better estimate. However, in our analysis, significant variation is observed over

some tags, invalidating the averaging process, and instead, η stays as a function of both ϕ_{r_m} and the multipath factor $(\delta_{ml} - \delta_{1l})$. The inclusion of this correction is one of the key factors in evaluating the final weak occupant-reflected signal and makes it different from the traditional calibration.

Next, to estimate ϕ_t, the unknown phase introduced by the tag circuitry, we substitute $\exp\left(-2j\left(\phi_{r_m}\right)\right)$in Eq. (5.9) by $\eta_{r_{m1}t_l}$ and average over all the reader Rx for each tag:

$$\rho_l = \sum_{m=1}^{M} y_{r_m t_l}^{NoOcc} \cdot \exp\left(2j\, \phi_{r_m t_l}\right) \cdot \eta_{r_{m1}t_l}^{*} \tag{5.11}$$

As the multipath variation is observed to be smaller over Rx, it is averaged out to get $\rho_l \approx c_2 \cdot \exp\left(-2j\phi_{t_l} + \phi_{r_1} + \delta_{1l} + \theta\right)$, with the averaged multipath effect lumped into θ.

In the final step, the calibrated signal y^{Cal} is derived as:

$$y_{r_m t_l}^{Cal} = y^{BS} \cdot \frac{\eta_{r_{m1}t_l}^{*}}{\rho_l}. \tag{5.12}$$

In the above model, we consider the signal at different frequencies to be independent with different initial phases and do not express explicitly the frequency subscript $n \in [1, N]$. Thus, the above steps are performed at each frequency for all the tag–Rx links.

Once the reflected signal is extracted, the linear signal model can be established as a function of the additional signal path traveled. Again only one-way tag to reader Rx path can be considered, considering the symmetricity to give the following phase accumulation on a path from l tag to m Rx, reflected from j^{th} voxel with a complex reflectivity of x_j at frequency f_n,

$$y_{r_m t_l f_n}^{Cal} = \sum_{j=1}^{N_{voxel}} e^{\frac{j2\pi f_n}{c}(d_{mj}+d_{lj})} \cdot x_j, \tag{5.13}$$

where d_{mj} and d_{lj} are the additional distances traveled. Only the phase variation is considered as the additional attenuation will be very small and limited by the RSSI resolution. The above equation can also be simplified to a linear matrix model,

$$y^{Cal} = Ax + \eta, \tag{5.14}$$

with y^{Cal} as a column vector of length $L \cdot M \cdot N \sim 1E4$, and x of length $N_{voxel} \sim 1E6$. A is a complex rectangular matrix of size $(L \cdot M \cdot N) \times N_{voxel}$ and η is the noise. This is also an underdetermined equation, but with more sensing points due to frequency diversity in the problem model. As x is a collection of reflectivity at each voxel, with only a limited part of the room being occupied, it is highly sparse. This thesis includes this additional information to propose a solution of (5.14) in the next section.

5.3.3 Combined reflection-attenuation model

While the reflection model gives a higher resolution 3D image compared to the RSSI attenuation model, it suffers strongly from the multipath noise, which is more prominent in the real-room setup with strong reflection from walls and ceiling, and also from people walking nearby. As the above two models measure two different

phenomena in occupants' presence, namely, reflection and attenuation, their combination is the next step to get improved results, especially with more than one occupant in the room, and high multipath noise. In such cases, if the noises are not completely correlated, some improvement is expected by the combined model, where x can be lumped as the same unknown variable, with low absolute value if a voxel does not attenuate or reflect, and high values otherwise.

The Eqs. (5.6) and (5.14) can be combined in two different ways, taking $u_1 = y^{Cal}$ and $u_2 = z^{RTI}$,

$$\text{Weighted Vector Augmentation: } [(1-\lambda)u_1; \lambda u_2] = [A; W] \cdot x + \eta \tag{5.15}$$

$$\text{Weighted Addition: } (1-\lambda)u_1 + \lambda u_2 = [(1-\lambda)A + \lambda W] \cdot x + \eta. \tag{5.16}$$

While the vector augmentation does not require u_1 and u_2 to be of the same lengths, addition requires the same length, thus, u_2 should be the vector at all frequencies.

Currently, λ is selected empirically – as the RSSI attenuation model has a higher tolerance to noise with poor resolution and lesser information content, it can be weighted higher in high noise conditions. An improved selection scheme is left for future works.

5.3.4 **Spatial and frequency diversity**

The spatial and frequency diversities help improve image reconstruction. This is expected as increased sensing points help reduce the underdetermined state of the problem and can help reach the unique solution without ambiguous solutions or aliasing. However, there is a limit on frequency range and the number of tags and reader antennas from a practical perspective. Thus, the problem comes down to selecting tag and reader Rx placement, and frequency selection with the given constraints. This thesis helps

provide the groundwork for this problem. The work in [77] introduced the relation of the inverse problem model in (5.14) to its Fourier transformed K-space domain, and thus solving the inverse problem can be considered as taking the inverse Fourier transform (FT) with non-uniformly sampled data points. From the knowledge of FT and the K-space domain, uniform coverage in the K-space domain is expected to give an improved performance. Thus, the following objective function is defined, that aims to minimize inverse of the distance between any two K-space sample points, K_i and K_j, and thus, reduce compactness and increase uniformity:

$$f(r_m, t_l, f_n): \min \sum_{1, i \neq j}^{L \cdot M \cdot N} \frac{1}{\left|\left|K_i - K_j\right|\right|_2}. \tag{5.17}$$

The global optimization is difficult to achieve, so a Simulated Annealing algorithm based on the Metropolis criterion is used. The initial temperature is set high (~1e10) to get closer to some local minima, starting from a random point. When some local minimum is reached, it is used as the starting point for a new search with a lower temperature (~1e6) and reannealing interval of ~45, to find the true minimum point.

Simulation validation is done by using this algorithm to select only 5 frequencies in the large bandwidth of [0.9, 3] GHz and compared with 100 equally distributed frequencies in the entire bandwidth. The number of tags and receivers are 16 and 4, respectively, and kept constant for both the cases and used to image a 2D ellipse shape as shown in Fig. 5.12 (a). Thus, increased K-space sampling points are observed in the second case, with 100 frequency points, as shown in Fig. 5.12 (b), resulting in image reconstruction shown in Fig. 5.12 (c), with one clear peak at the correct location. The z-dimension in the image shows the normalized reflectivity value at each pixel, and it can

be thresholded to give the object location and approximate shape. Simulated annealing was run to give the frequency set of: [1.7, 2.025, 2.35. 2.675, 3.0] GHz, and the results are shown in Figs. 5.12 (d), (e), giving a very similar performance as the previous case. Fig. 5.12 (f) shows an image generated by using 5 uniformly spaced frequencies in the entire bandwidth, resulting in many noisy side lobes. Thus, this approach led to better performance with limited bandwidth and a limited number of frequencies.

For our experiment setup, with practical concerns, reader Rx antennas are only placed on the ceiling and the corresponding RFID tags need to be placed in the reader's main beamwidth to be read, and thus can have limited spatial diversity. So, this problem

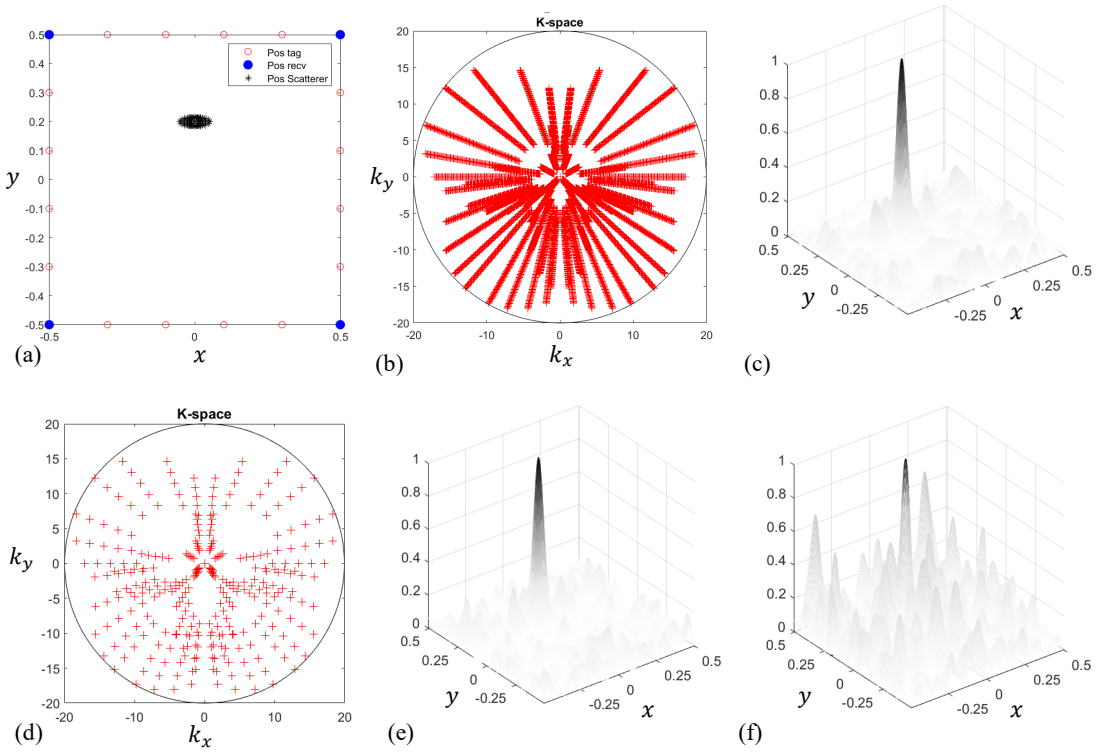


Figure 5.12 Optimal frequency selection using K-space optimization. (a) Simulated 2D imaging setup with 16 tags, 4 reader antennas and an elliptical object. (b) K-space samples with 100 uniformly selected frequencies in bandwidth [0.9,3] GHz. (c) Reconstructed image with 100 frequencies. (d) and (e) K-space coverage and image generated with 5 frequencies selected with proposed algorithm. (f) Image generated with 5 uniformly selected frequencies in range [0.9,3] GHz.

was only applied for frequency selection. However, the increased noise level in the experiment setup results in a high number of unread frequencies, and thus validation of improvement becomes difficult. Overall, this problem formulation can be very helpful in frequency and spatial diversity design with better hardware support.

5.4 Proposed inverse problem solution

The image is reconstructed by estimating the real or complex x in the inverse problems (5.6), (5.14) – (5.16) of the format

$$r = St + \eta, \tag{5.18}$$

with known r and S parameters. With a length of $len(r) \sim 1E4 \ll len(t) \sim 1E6$, this is an underdetermined set of equations with no unique solution. But with additional constraints or assumptions, some solution can be derived, that is validated against the ground truth of occupant location and size. This section presents the traditional approach, its variation, and proposed sparsity-based solutions and discusses their pros and cons.

As t usually denotes the occupant presence at each voxel in the room, with a non-zero real or complex value, the image is plotted as the magnitude square, $|t|^2$, and needs to be segmented to detect the number and location of the occupants. Traditional algorithms generate non-zero values at empty voxels and ghost images due to poor constraints and a high number of unknowns. Thus, the image first needs to be thresholded before object detection. All voxels below the threshold are taken as 0 and the image vector is converted to a 3D matrix corresponding to voxel arrangement in the room. A connected component clustering algorithm [176] is applied, that works by

scanning the thresholded image voxel-by-voxel (from top to bottom and left to right) and identifies the connected regions having all non-zero values. A voxel $V(i,j,k)$ is connected to $V(i-1,j,k), V(i+1,j,k), V(i,j-1,k), V(i,j+1,k), V(i,j,k-1), V(i,j,k+1)$, and so most pixels have 6 neighbors. All independent connect regions are taken as number of objects and their centroid as the object location.

5.4.1 Matched filtering

Matched-Filtering (MF) is the simplest traditional solution of (5.18), that gives an estimate of t as,

$$\widehat{t_{MF}} = S^T r \tag{5.19}$$

where S^T is the complex conjugate transpose or Hermitian transpose of S. MF minimizes the error component related to the noise power and maximizes the signal-to-noise ratio. Computing $\widehat{t_{MF}}$ involves the simple computation of S^T and is the fastest solution. However, without additional constraints, this does not reduce the entropy of the problem and is prone to artifacts and not well-resolved.

5.4.2 Conjugate gradient least squares

For underdetermined or rank-deficient case, the least-squares (LS) problem $\left||r - St|\right|_2$ has an infinite number of minimizers. MF, for instance, is a special case of LS estimation. The set of all minimizers is convex and has a unique element having a minimum length. Thus, the LS problem is set up to get the minimum norm solution of $r = St$, i.e., with the constraint minimum $||\hat{t}||$. This can be written as follows using the Lagrangian,

$$\widehat{t_{LS}} = \min_t \left||t|\right|^2 \; s.t. \; r \in \arg\min_x ||Sx - r|| \tag{5.20}$$

$$\mathcal{L}(t,\mu) = \left|\left|t\right|\right|_2^2 + \mu^T(r - St)$$

$$\widehat{t_{LS}} = S^T(SS^T)^{-1}r$$

where $S^T(SS^T)^{-1}$ is called a pseudo inverse. An extension to this is the weighted LS, corresponding to minimizing $\left|\left|W^{1/2}t\right|\right|_2^2$ where W is the diagonal matrix with w_{ii} weights, resulting in the solution $\widehat{t_{WLS}} = W^{-1}S^T(SW^{-1}S^T)^{-1}r$. This solution is less expensive than other direct approaches including singular value decomposition (SVD) and orthogonal factorization of S, but it is also less accurate, particularly for ill-conditioned problems.

Instead of the direct method such as given above, the damped LS problem

$$\min\left|\left|\begin{bmatrix} S \\ \lambda I \end{bmatrix} t - \begin{bmatrix} r \\ 0 \end{bmatrix}\right|\right|_2 \tag{5.21}$$

can be solved iteratively with a more stable conjugate gradient (CG) for minimizing. The above form is equivalent to Tikhonov regularization. The advantage of the iterative method is that solution can be controlled by the number of iterations, which provides additional regularization. At each step the residual $res_i = r - St_i$ is computed, and while residual may not decrease at each step in the direction of the correct solution, 0 residual is expected for the final result. The convergence rate of the CG-method is affected by the condition number of S^TS, and it remains unpredictable in our case, as the problem is not well-conditioned. In such cases, additional preconditioning can be applied to improve convergence [177]. This is less prone to artifacts than MF, however, the resolution is not much improved.

5.4.3 ℓ^1 and ℓ^0 sparsity approximations

The previous LS solution considered $\boldsymbol{\ell}^2$ norm regularization, which minimizes the norm of estimated $\hat{t}$, but does not utilize inherent sparsity in the image due to a large number of empty voxels in a room. In penalized form, the use of ℓ^0, ℓ^1 norms give rise to the problems:

$$\min_t \left|\left|r - St\right|\right|_2^2 \ subject\ to\ \left|\left|t\right|\right|_0 \leq k1 \qquad \text{(Best subset selection) (5.22)}$$

$$\min_t \left|\left|r - St\right|\right|_2^2 \ subject\ to\ \left|\left|t\right|\right|_1 \leq k2 \qquad \text{(Lasso regression) (5.23)}$$

where $\left|\left|t\right|\right|_0$ gives the number of nonzero elements in vector t and $\left|\left|t\right|\right|_1 = \sum_{i=1}^{N_{voxel}} |t_i|$. This section proposes image-reconstruction based on sparsity assumptions of ℓ^1 norm, and approximate solution of (5.22). Overall, sparse algorithms have two major advantages: 1) introduce a known sparsity constraint to make the problem more solvable and reducing the solution domain, and 2) remove thresholding dependence on the generated image for occupant counting and locating, which is an inherent requirement of both of the above algorithms.

FISTA: The fast iterative shrinkage thresholding algorithm (FISTA) [178] is a method to solve the non-differentiable convex optimization problem with high computational efficiency. This method utilizes proximal algorithms that are used for non-smooth constrained problems. The Eq. (5.23) is considered in the form

$$\min_t \frac{1}{2}\left|\left|St - r\right|\right|^2 + \gamma\left|\left|t\right|\right|_1, \qquad (5.24)$$

$$\min_t f(t) + g(t). \qquad (5.25)$$

Here, $f(t) = 1/2\left||St - r|\right|^2$ is a smooth continuously differentiable function, and $g(t) = \gamma\left||t|\right|_1$ is a non-smooth continuous convex function. As (5.25) is a complex problem, the solution t^* verifies:

$$0 \in \lambda\nabla f(t^*) + \lambda\partial g(t^*)$$
$$0 \in \lambda\nabla f(t^*) - t^* + t^* + \lambda\partial g(t^*) \tag{5.26}$$
$$(I + \lambda\partial g)(t^*) \in (I - \lambda\nabla f)(t^*)$$

Here, left is the backward Euler, which is the proximal step and right is the general forward Euler, i.e., gradient descent step. Multiplying by $(I + \lambda\partial g)^{-1}$ on both sides results in the following:

$$t^* = (I + \lambda\partial g)^{-1}(I - \lambda f)(t^*) \tag{5.27}$$

Thus, $k + 1^{th}$step in one FISTA iteration involves the gradient-based step with a shrinkage operator:

$$t_{k+1} = \text{prox}_{\lambda_k g}\big(t_k - \lambda_k\nabla f(t_k)\big). \tag{5.28}$$

The step size λ_k is in the range $(0,1/L_c)$, where L_c is the Lipschitz constant of $\nabla f(t)$. The proximal operator for the ℓ^1 norm of a real-valued vector is the shrinkage operator, which is the simple element-wise operator called soft-thresholding. With $g(t)$ as the ℓ^1 norm of a complex vector t, we derive the proximal operator of g in the next steps. The proximal operator seen in (5.28) can also be defined as

$$\text{prox}_{\lambda g}(v) = \arg\min_t g(t) + \frac{1}{2\lambda}\left||t - v|\right|_2^2 \tag{5.29}$$

$$\Rightarrow \sum_{i=1}^{N_{voxel}} \arg\min_{t} |t_i| + \frac{1}{2\lambda}|t_i - v_i|^2$$

Taking the derivative with respect to the real and imaginary part gives

$$\begin{gathered} \lambda t_{iR} + t_{iR}|t_i| - v_{iR}|t_i| = 0, \lambda t_{iI} + t_{iI}|t_i| - v_{iI}|t_i| = 0. \\ \Rightarrow t_{iR} = \frac{v_{iR}|t_i|}{\lambda + |t_i|}, t_{iI} = \frac{v_{iI}|t_i|}{\lambda + |t_i|} \end{gathered} \tag{5.30}$$

Now substituting these values in $t_{iI}^2 + t_{iR}^2 = |t_i|^2$,

$$\begin{gathered} |v_i|^2 = (\lambda + |t_i|)^2 \Longrightarrow |v_i| = \lambda + |t_i| \\ \Longrightarrow |t_i| = |v_i| - \lambda \geq 0. \end{gathered} \tag{5.31}$$

The proximal operator with phase-preservation is derived as

$$\text{prox}_{\lambda g}(v) = \begin{cases} (|v_i| - \lambda\gamma)e^{j\theta_{v_i}}, & |v_i| > \lambda\gamma \\ 0, & |v_i| \leq \lambda\gamma \end{cases}, \tag{5.32}$$

where θ_{v_i} is the angle of the complex v_i.

The initial t_0 is selected as an all-zero vector. A prior estimate of t_0 can improve convergence and computational cost. The regularization parameter γ and the tolerance limit δ are empirically selected and remain constant for a given imaging setup and signal-to-noise ratio (SNR). The step size λ is further fine-tuned by the backtracking FISTA implementation [178]. The stopping criteria are defined with the tolerance limit on the norm of Δt during two consecutive iterations.

For the gradient method on (5.24), the rate of convergence is no worse than $O(1/k)$, where k is the iteration counter. FISTA has an improved complexity result of $O(1/k^2)$, thus achieving faster convergence.

Orthogonal matching pursuit: The ℓ^0 "norm" is not a norm in a strict mathematical sense as it does not satisfy homogeneity, i.e., for p-norm, it should satisfy $\left|\left|\alpha t\right|\right|_p = \alpha\left|\left|t\right|\right|_p, \alpha \geq 0$. The $\left|\left|t\right|\right|_0$ is the ℓ^0 pseudo-norm which gives the number of nonzero elements of t. Thus, it is directly related to the sparsity, and solving the sparsity problem (5.22) exactly is NP-hard. Matching pursuit (MP) is an approximate

1) **a greedy** algorithm that solves problem optimally at each stage, and
2) **an iterative** algorithm that finds the elements of x in a step by step manner,

proposed by Mallat and Zhang [179]. Starting from an initial approximation of $t^0 = 0$, and residual estimate, it iteratively generates a sorted list of atoms (columns of over-sampled dictionary S) and weighting scalars (corresponding non-zero elements in t), which form the sub-optimal solution to the sparse problem. Orthogonal matching pursuit (OMP) is an extension of this approach, where at each step all extracted weights or coefficients are updated by computing the orthogonal projection of signal on the subspace spanned by the set of selected atoms.

To find the voxel reflectivity, the dictionary is just taken as S, but it can be anything. For a noiseless real-domain problem, the recovery of t can be proved based on the properties of S as follows:

Definition: For $t \in \mathbb{R}^n$, t is k-sparse if the number of non-zero elements in t is k, i.e., $\left|\left|t\right|\right|_0 = k$.

Definition: Given $r \in \mathbb{R}^m$, dictionary $S \in \mathbb{R}^{m \times n}$, $n \gg m$. The mutual coherence, μ of the i^{th}, j^{th} column vectors of S, denoted as s_i and s_j, is given as

$$\mu = \max_{i \neq j} \frac{\left| s_i' s_j \right|}{\left|\left| s_i \right|\right|_2 \left|\left| s_j \right|\right|_2} \tag{5.33}$$

and t can be recovered by OMP if the following inequality is satisfied:

$$\mu < \frac{1}{2k-1} \tag{5.34}$$

Existing literature has considered cases of the complex domain [180] and in noisy conditions with additional constraints. In our setup, the convergence is more difficult to be estimated because of no assumptions of Gaussian noise, and the non-randomness of S. However, it can be safely assumed that the sparser t is, the fewer coherence constraints are required on S. Thus, the solution will work well if a few people are

1. Initialize residual $\mathbf{res^{(0)}} = \mathrm{r}$.
   Support estimate $\mathcal{T}^0 = \phi$, iteration counter, $k = 1$.
2. Find column most correlated with the current residual $\mathbf{res^{(k-1)}}$, ie, $c_k = \mathrm{argmax}\, |S^T \mathbf{res^{(k-1)}}|$, $c \in \{1, \dots, n\}$.
   Check if that column has not been previously selected. Otherwise repeat 2.
3. Update support estimate $\mathcal{T}^k = \mathcal{T}^{k-1} \cup c_k$.
4. Estimate t using current support $\hat{t} = S_{\mathcal{T}^k}^{\dagger} r$.
5. Update residual $\mathbf{res^{(k)}} = r - S\hat{t} = \left(\mathbf{I}_m - \mathbf{P}_{\mathcal{T}^k}\right) r$.
6. Increment $k = k + 1$.
7. Repeat $2 - 6$, until Stopping Criteria (SC) is met.

Output: $\mathcal{T}^k$ and $\hat{t}$.
Notations:
Support $\mathcal{T}(t) = \{i : t_i \neq 0\}$
$S^{\dagger} = (S^T S)^{-1} S^T$, Moore-Penrose pseudo inverse.
$\mathbf{P}_S = SS^{\dagger}$. $\mathbf{P}_{S_{\mathcal{T}}}$ is denoted by $\mathbf{P}_{\mathcal{T}}$.

Figure 5.13 OMP iterative algorithm to solve (5.22).

present in a small room, considering the dictionary size is related to the number of voxels. The stopping condition can be an estimate of sparsity or condition on the residual. In this work, sparsity corresponding to the number of filled voxels for nearly two occupants is selected. This works well for both one and two occupant cases and only changes the size of the estimated clusters, which is acceptable for the application. The performance of this is superior to that of MF in most cases and will be discussed in the results section.

As OMP mainly selects the best dictionary columns (atoms), the same algorithm can be used with a different problem setup. If columns of S are collected data, i.e., r vectors corresponding to people at different grid points in the room, then the problem becomes equivalent to selecting the best possible location(s) out of the collected data. In this case, for $S \in \mathbb{R}^{m \times n}$, the dimension n corresponds to the number of data collected. This is termed as OMP2 for convenience. Here, the advantage is that the occupant location and number can be derived without going through image generation and segmentation steps. Thus, sparsity is directly related to the number of occupants, and stopping criteria cannot be an approximate number as earlier. An automated tuning free OMP approach [181], [182] is adopted that observes the residual ratio (RR) at each iteration. For the k^{th} iteration, $RR(k) = \left||\text{res}^k|\right|_2 / \left||\text{res}^{k-1}|\right|_2$, & $0 < RR(k) < 1$. The residual at k^{th} iteration is as shown in Fig. 5.13 (line 5). As our observations are non-linear, the noise model is non-Gaussian. Further, the sparse recovery mutual coherence criterion in (5.34) is not satisfied for $k > 2$. Thus, while this problem setup is advantageous, it may not work well consistently. Other limitations include prior data collection, and performance deterioration with the high noise level in real-world

experiments, that needs to be further studied. The preliminary results using automatic stopping conditions are presented in the next section.

Both the above problems can be solved by the same OMP algorithm, with different stopping conditions and dictionary S. The algorithm is presented in Fig. 5.13 for a generalized stopping criterion (SC).

5.5 Results

This section presents the results across simulation and experiment datasets for both scaled and true-size setups. The focus is on comparison and validation of the proposed 1) model, 2) calibration, and 3) algorithm approaches under these varying noise conditions. More importantly, these proposed approaches are aimed to improve image-generation at the true-scale setup. With very different multipath noise conditions and object sizes, the different setups need only one or more of the above approaches. Thus, the presented results are primarily grouped based on setup designs. The model used is the reflection model, unless mentioned otherwise.

For occupant or object counting and localization, further processing is required on the generated 3D image, including segmentation and clustering. As OMP and FISTA are sparse reconstruction algorithms, in most cases no thresholding is required, the image is already 0 where no occupants are detected. For FISTA, however, there are various tunable parameters, that may result in a noisy background, and some threshold is required, but a fixed threshold works well with a given set of parameters. MF always requires a threshold. As the image intensity changes with room size, the number of

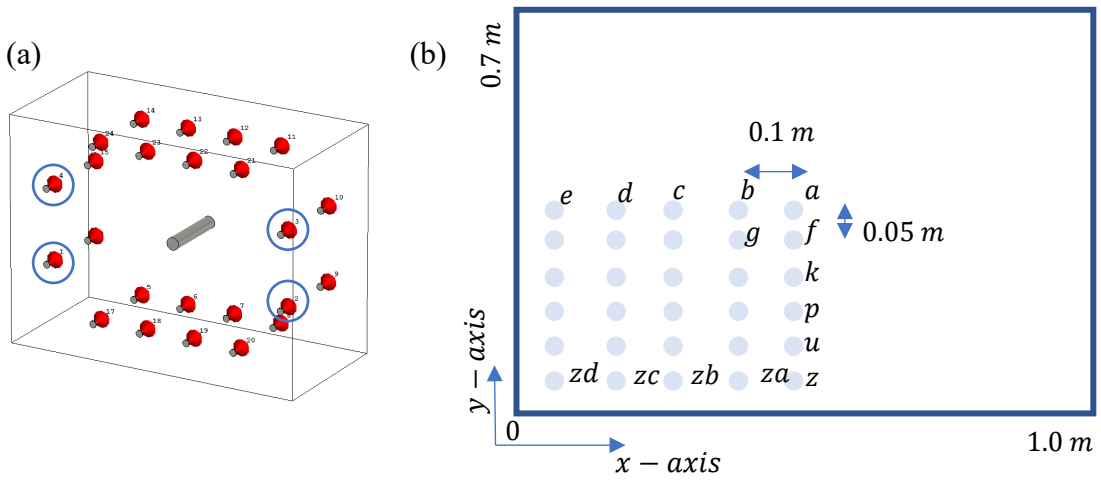


Figure 5.14 Scaled simulation setup with dipole antennas around the room. (a) Setup with cylindrical PEC object at the center and four reader antennas as highlighted. (b) Different object locations in the bottom-left quadrant.

occupants, and their locations, it is very difficult to come up with threshold selection strategies for MF and CGLS algorithms that do not generate sparse images.

5.5.1 1/6th scaled setup

Simulation setup: For the simulation setup shown in Fig. 5.7 (a), one cylindrical PEC object of radius 25 mm and height 270 mm was placed at 30 various locations in the bottom-left quadrant. As the simulation is noise-free, in this symmetrical room, all four quadrants are identical. Further, around 10 simulations were performed with 2 objects in the same quadrant. Fig. 5.14 shows the simulation setup and different object locations.

Fig. 5.15 shows images with objects at a few distinct locations with satisfactory results for all three algorithms. As it is a noise-free case and small PEC object, the reflection model is used with proposed calibration (denoted by "4.2"). The voxel's size is sufficiently small to not give a pixelated appearance. Six uniformly distributed frequencies in the range 900 – 925 MHz are used for reconstruction. The time

complexities of the algorithms are different, with MF being the fastest, only taking a few ms, compared to about 4 s for OMP and 6.5 s for FISTA. The calibration performance is similar for most locations, with exceptions such as corner-most

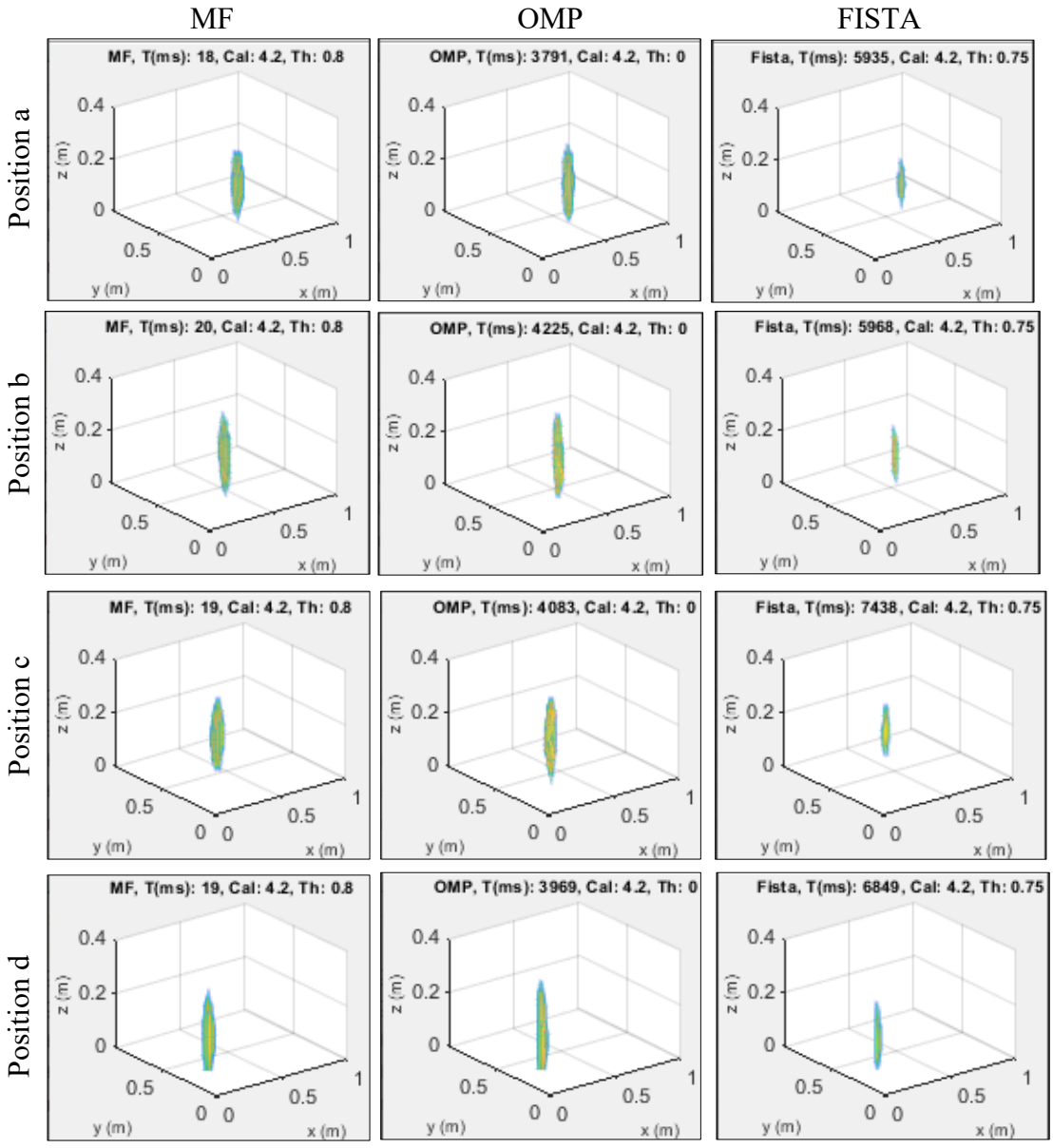


Figure 5.15 Generated images for scaled simulation with object at four different locations. Rows indicate object locations and columns correspond to different algorithms used. In all cases, correct location is estimated. OMP requires no threshold (Th) and gives accurate estimate. With one strong object, MF and FISTA thresholds are constant and empirically selected.

placement as shown in Fig. 5.16, where proposed calibration works better than earlier approach [77] (denoted by "1").

The two object cases are more difficult as one object may shadow another. Few good cases using proposed calibration are shown in Fig. 5.17 while comparing different

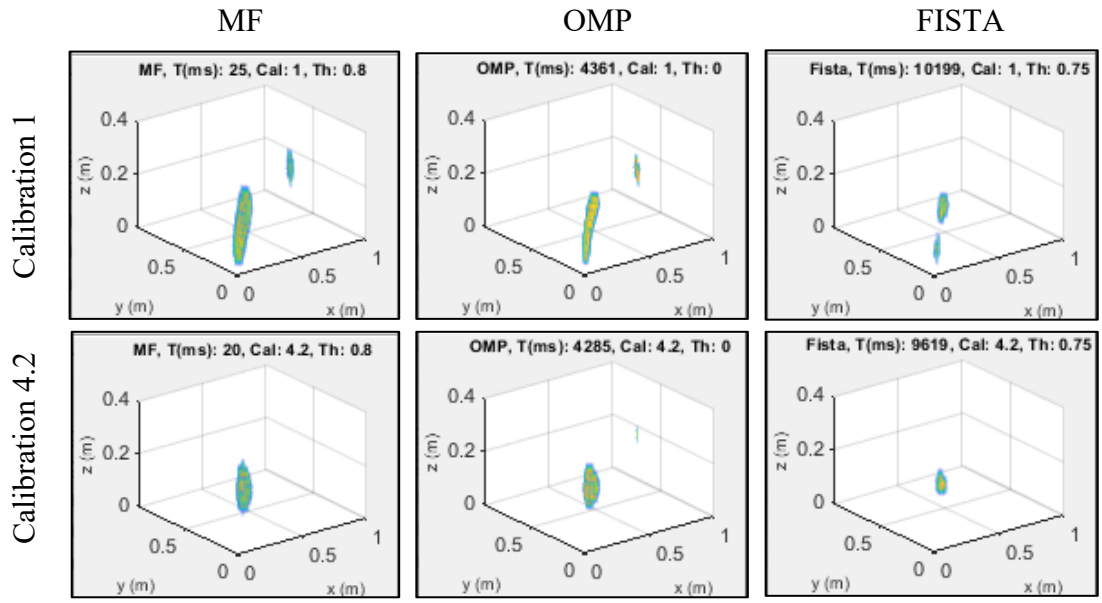


Figure 5.16 Comparing calibration for scaled simulation. Top row shows calibration 1 and bottom shows proposed 4.2, with constant thresholds.

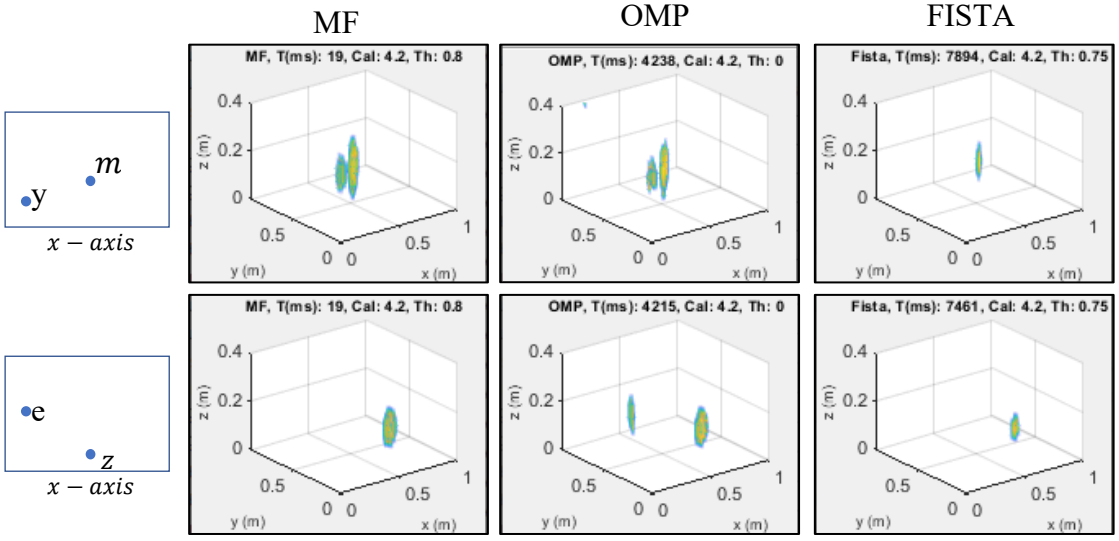


Figure 5.17 Scaled simulation results for two objects. OMP is consistently better performing without any threshold requirements.

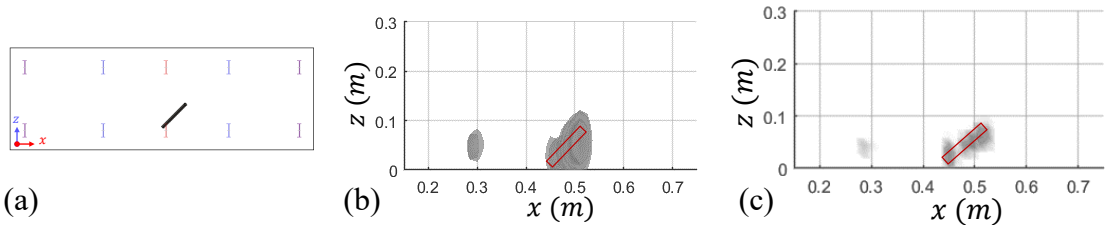


Figure 5.18 Improved FISTA performance for small objects. (a) Scaled simulation setup. (b) MF grayscale image with red rectangle showing true object. (c) FISTA image showing pixelated cluster giving accurate object size and angle.

algorithms. OMP can detect correct occupant count as well as locations for all the cases, with fixed stopping criteria. FISTA and MF can detect one object's location correctly. FISTA's performance is slightly poorer than MF, mostly because of non-linearities in the model with large object size. Fig. 5.18 compares the case of a smaller object at an angle, where FISTA is better at giving sparse reconstruction of an object showing the correct angle.

For some cases, like objects at $a\&e, b\&e$, the e location is difficult to be detected. With the symmetricity of the setup and $a\&e$ being on the same perpendicular line from the reader antenna plain, some shadowing effect may have taken place. However, with different frequency ranges, the less strong object can also be recovered,

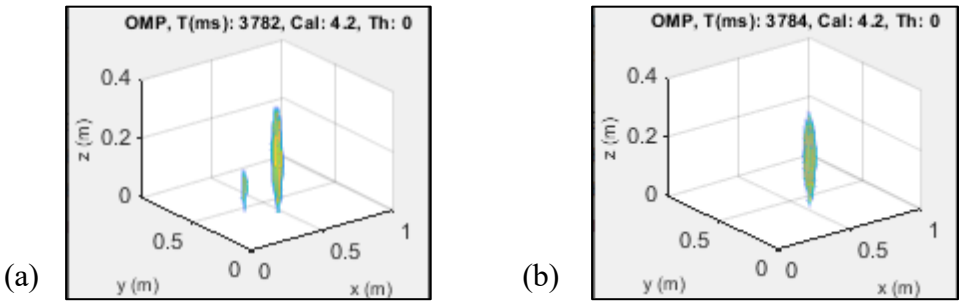


Figure 5.19 Poor cases with two objects. (a) Two objects at $a\&d$, with very small indication of d. (b) Two objects at $a\&e$ where only a is detected.

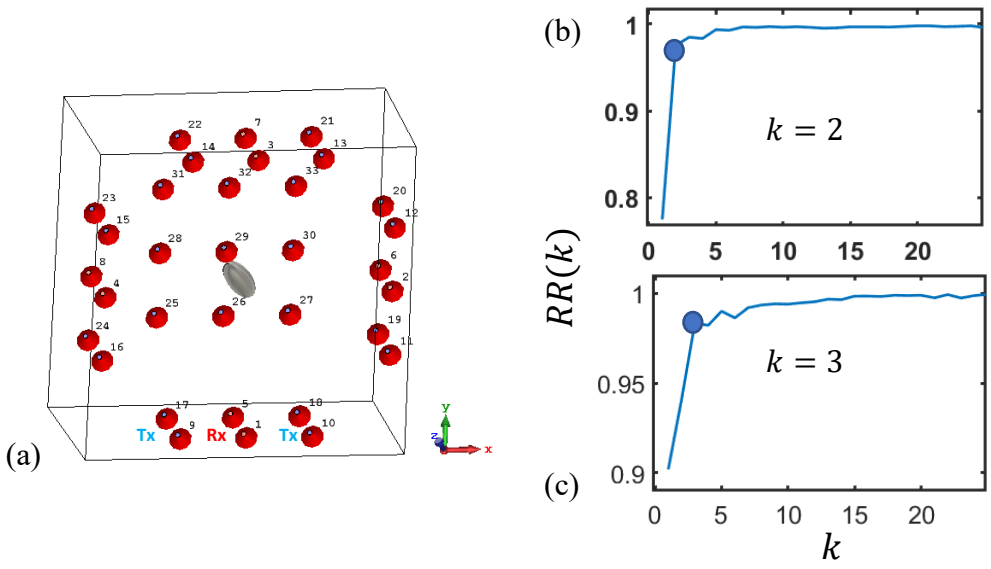


Figure 5.20 OMP2 results. (a) Simulation setup corresponding to dictionary generation. (b) & (c) $RR(k)$ vs k plots, with $k = \#iterations$ for one and two objects, respectively. Correct stopping criteria is predicted, equal to #objects + 1.

which will be presented in the next section. Fig. 5.19 shows the OMP reconstructed images for objects at $a\&d, a\&e$, where a very weak reconstruction is seen for the former case, and the latter shows no existence of the second object.

To test the performance of OMP2: OMP with dictionary made of collected data r, another setup was considered with eight reader antennas on the side walls (central columns are reader antennas, remaining all are tags) as shown in Fig. 5.20 (a) and an ellipsoidal object at different locations. The dictionary is composed of data collected with only one object placed at different grid points in the room. Test cases include data collected with one or two objects at these pre-selected grid points.

Corresponding $RR(k)$ vs k plots for one and two object cases are shown in Figs. 5.20 (b) and (c), where it can correctly get the knee point index to equal to $(1 + \#objects)$. For greater than 2 objects, this setup gives the wrong estimate more frequently. Thus, with improved spatial diversity, and low noise conditions, this method

Figure 5.21 The scaled experiment setups with different object count, locations and postures, and corresponding generated images on the monitor, marked by orange circles.

has the potential to detect correct occupant counting. Further, after the stopping condition is selected, the selected column of the dictionary also gives information about the object's location. More analysis needs to be done for cases when objects are randomly placed.

Experiment setup: Fig. 5.21 shows the images generated on the computer screen corresponding to different experiment setups using the MF algorithm. It has been further tested with different objects with higher reflectivity (cans) and lower reflectivity (a plastic object with metallic paint) as shown in Fig. 5.22. The corresponding results show that two high reflection objects placed closely together can be detected using the OMP

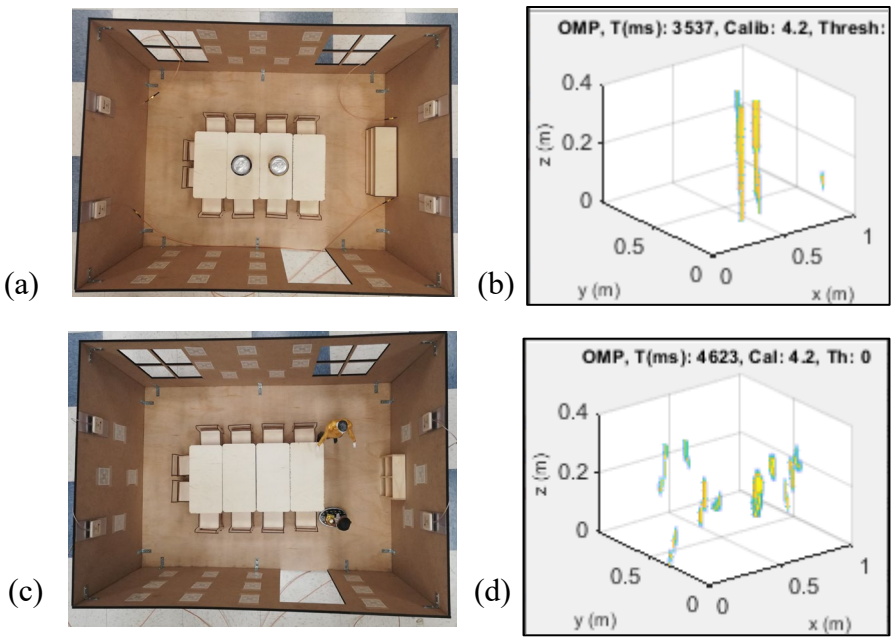


Figure 5.22 Scaled experiment results with proposed calibration and different objects. (a) & (b) Metal cans placed close by and OMP reconstructed image. (c) & (d) Two 3d printed plastic scaled human dolls with metal paint. Corresponding image is noisy.

algorithm (zero threshold), but smaller reflection objects are not that clear. Similar conclusions are observed with MF and FISTA.

5.5.2 True-scale setup

Simulation setup: The true scale simulation is as shown in Fig. 5.7 (b). The algorithm and calibration comparisons are similar to that in the scaled model. However, as the object size increases, the linearly approximated models have slightly poorer performance than earlier. This can be taken care of by putting additional tags on furniture and room dividers as shown in Fig. 5.23. Both one and two objects can be detected with additional center tags. As human muscle material requires a lot more

meshing with huge computation, only PEC materials can be simulated. Thus, the experimental setup provides real validation.

Experiment setup: The real-scale room size has more impact on the data quality compared to scaled setup due to 1) **stronger multipath** from walls and ceiling reflection, and people walking outside the room, that results in **phase noise** in tag readings over time, 2) **human** body material is not entirely reflective as considered in scaled model and simulation, 3) **large occupant size** resulting in loss of tag readings, and 4) **low tag read rate** likely as tags are not in the main beamwidth of the reader antenna.

Fig. 5.24 (a) shows image reconstruction result without any corrections, with one occupant standing at (0.6,3.0) m in a room of size $3.6m \times 3.6m$ as shown in Fig.

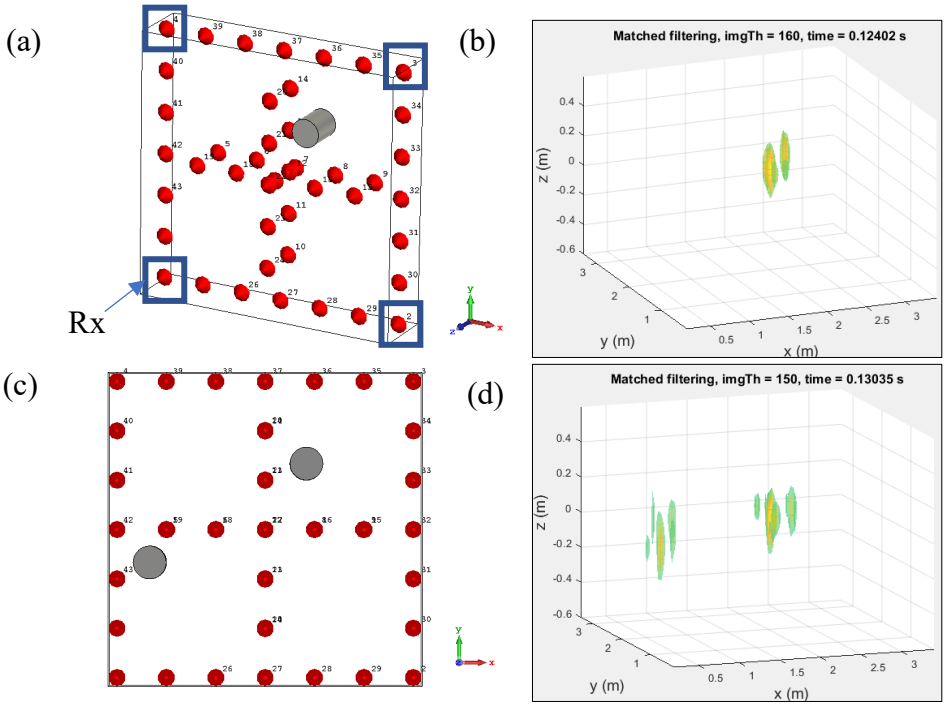
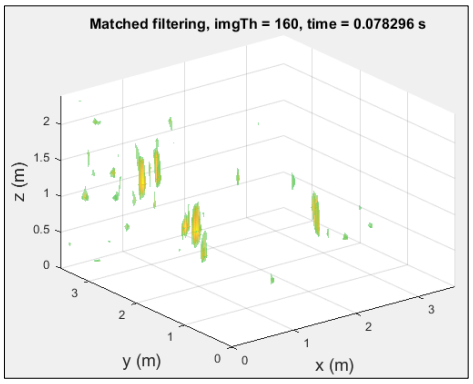


Figure 5.23 Tags on the furniture in center. (a) & (b) Simulation setup with four readers on the corner and 40 tags on the wall and middle dividers. MF results shows accurate location. (c) & (d) Same simulation setup with 2 objects. MF can detect both objects, with slightly different thresholds, that can take values between 0 to 255.

(a)



(b)

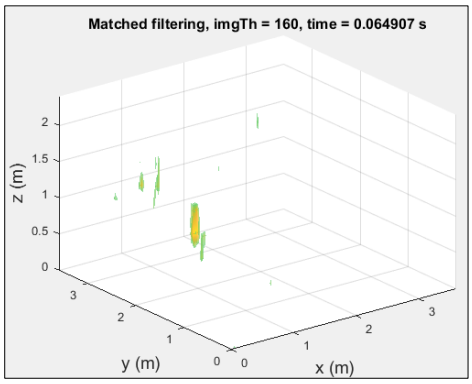
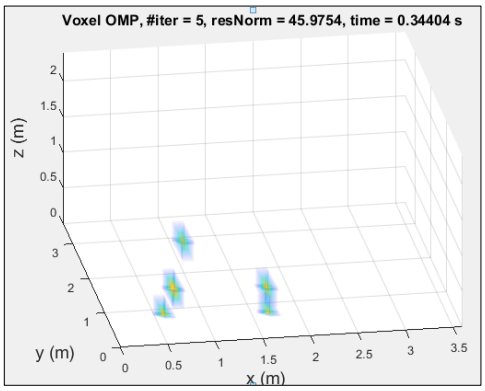


Figure 5.24 RFID tag phase error correction results with occupant standing at (0.6,3.0) m. (a) No correction, 24% missed tag-readings, multiple objects detected. (b) Rejecting tag-Rx pairs with phase standard deviation above 12°, 35% links not considered. One occupant detected at (0.8,1.8) m.

5.9 (b). Multiple ghost clusters are detected in this image. Fig. 5.24 (b) shows the reconstructed results by rejecting tag – reader Rx links with phase standard deviation over multiple readings to be greater than 12°. This substantially improves the results, showing only one occupant close to the true location. The calibration improvements can be clearly observed for the experiment setup as shown in Fig. 5.25. The left figure shows multiple clusters, while the right figure shows one clear occupant at the correct location.

The data collection was performed with one or two occupants standing or sitting at different grid points in the room. Fig. 5.26 shows the results of occupancy detection

(a)



(b)

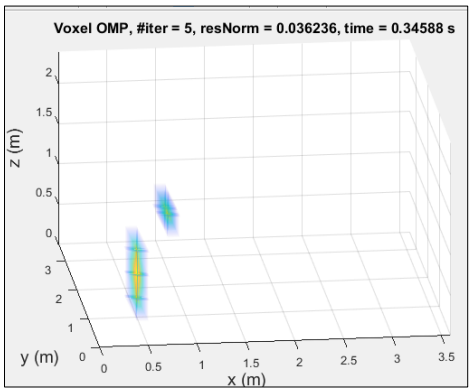


Figure 5.25 Calibration comparison. One occupant at (0.6,1.2) m, image reconstruction using OMP. (a) Earlier calibration. (b) Proposed calibration detecting one occupant.

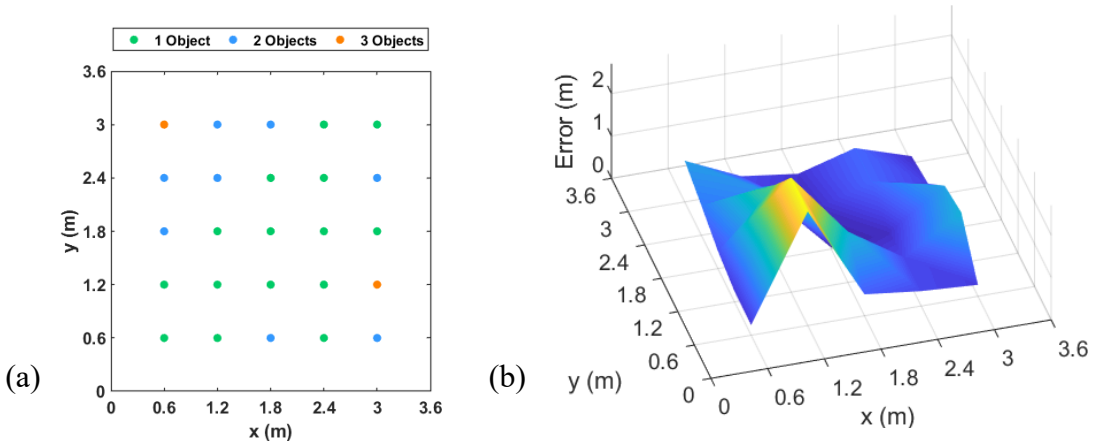


Figure 5.26 MF grid results for occupant counting and locating using earlier calibration and no phase correction. (a) Detected number of occupants indicated by circle color. Center locations are accurate. (b) Localization error in meters. For multiple occupant count, their center location is taken for result estimation.

for one occupant in the room using MF with no calibration and phase corrections. Only occupants at the center can be correctly counted. Fig. 5.27 shows the grid results using OMP with proposed calibration and phase correction. Correct occupant detection is performed at all locations, with low localization error in most cases, except 2 points.

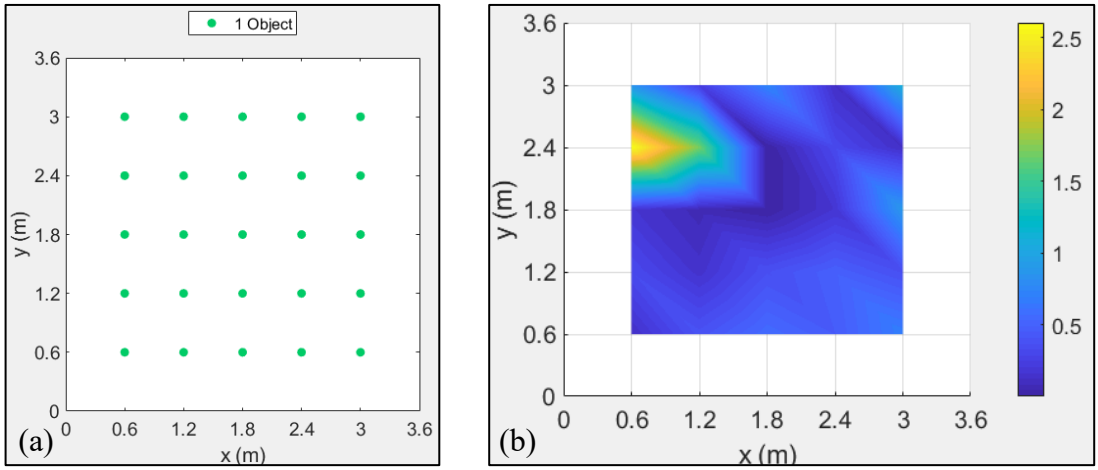


Figure 5.27 OMP grid results for occupant counting and locating using proposed calibration. (a) Detected number of occupants, green circles indicate correct count. (b) Location error at each grid point, zero at center, and very high for (x,y) = (0.6,2.4)m, which may be attributed to very high multipath noise.

Few cases with two occupants are shown in Fig. 5.28 with the same algorithm, however,

the performance is poor with only one occupant detection in most cases. Setup as shown in Fig. 5.23 is expected to perform better by improving tag readability with reduced phase noise if room tags are possible. For more than one occupant case, lower accuracy is also related to nearby standing or sitting occupants, which can be lumped as one.

The other improvement that is considered at this point is using the RSSI attenuation model, as occupants absorb the microwave radiation and attenuate the signal. With phase noise correction, this model generates better images for some cases, however with poor resolution and only in 2d. Fig. 5.29 shows the results of the reflection and RSSI attenuation model for two cases, with correct occupant counting using the latter approach for the second case. To reconstruct high-resolution 3d images, section 5.3 presented a weighted combination or *ensemble* of both as shown in (5.15) and (5.16). Fig. 5.30 shows the comparison of occupant counting using all four models: refection, RSSI attenuation, and weighted vector augmentation (WVA), and weighted addition

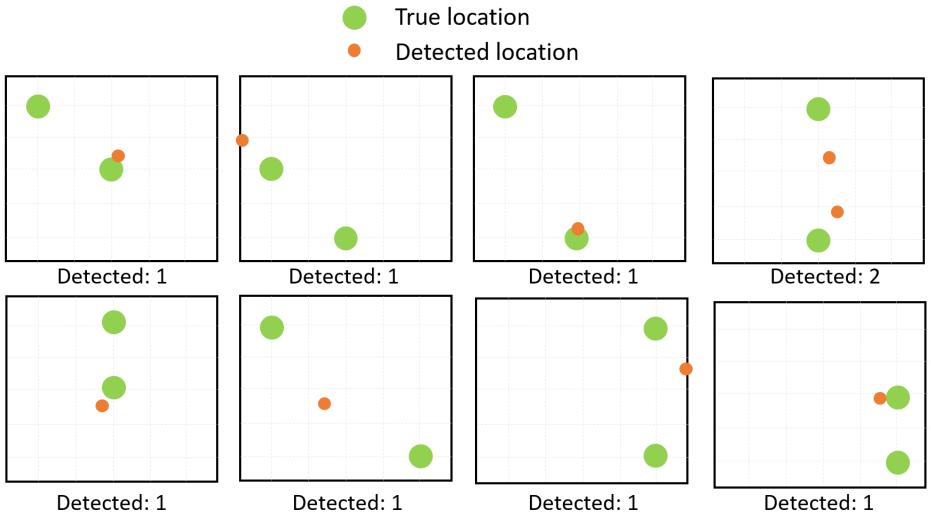


Figure 5.28 OMP grid results for two occupants showing true and detected locations. With two occupants, reflection model alone is very noisy.

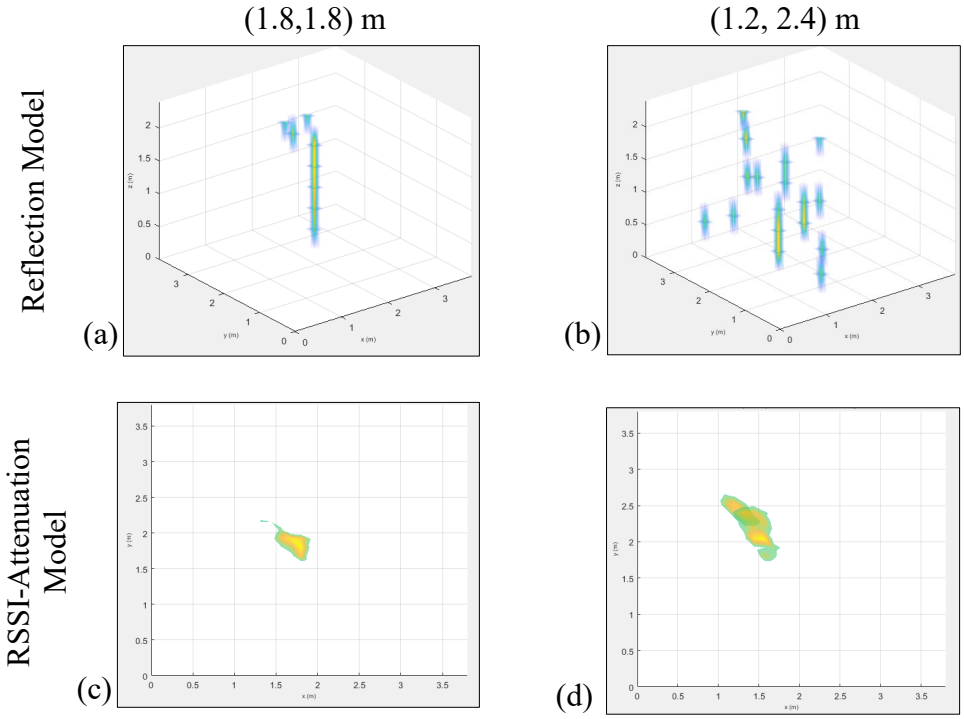


Figure 5.29 Model Comparison with two different occupants at different locations. (a) & (b) Show reflection model results where (b) result is noisy and results in multiple occupant detection. (c) & (d) Show results of RSSI attenuation model with one 2D cluster at nearly correct locations.

(WA) of both. The results are derived over the entire dataset where ≥ 2 occupants are taken as two. An empirical threshold derived from collected data is used for zero against all classification, resulting in most detections to be accurate. More details on this are presented in the next chapter. The weighted average (WA) model gives the best performance with an accuracy of 72.3%.

5.6 **Discussion**

This chapter discussed in detail RF imaging models, reconstruction algorithms, and real-life challenges with the noisy experimental data. In addition to this, other factors are affecting the overall results, that can be integrated to improve the results if

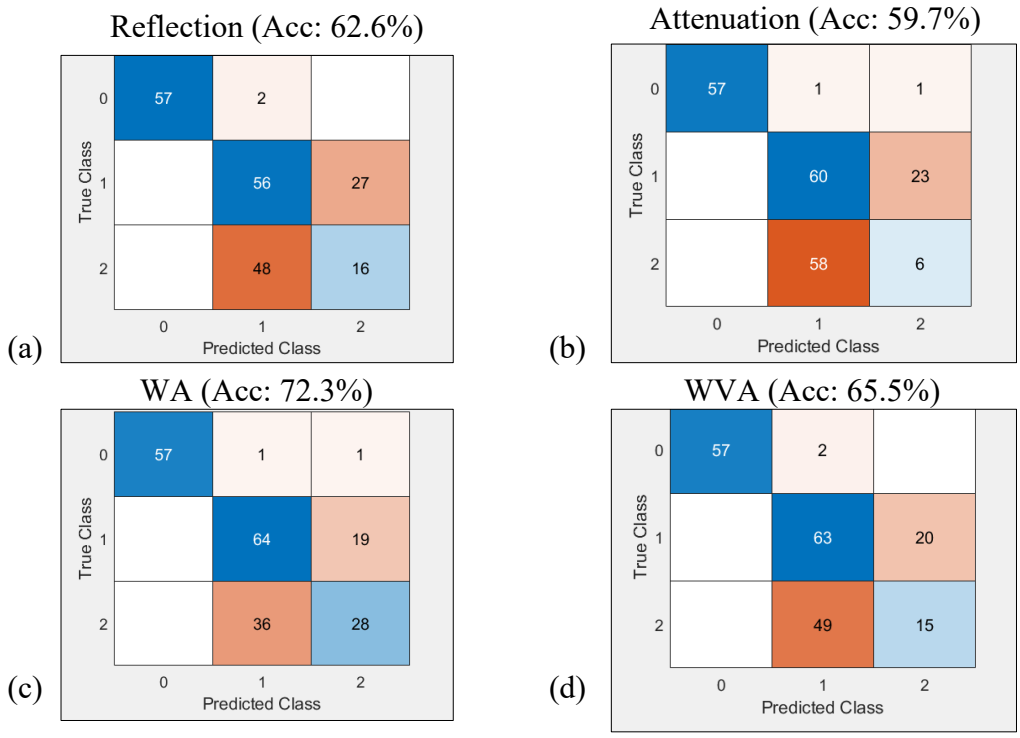


Figure 5.30 Confusion matrix showing predicted vs true occupant count for different models. Algorithm used is OMP, and proposed calibration is used for reflection model. The model ensemble approaches use $\lambda = 0.8$. (a) Reflection model. (b) RSSI attenuation model. (c) Weighted average of both models. (d) Weighted vector augmentation of both models.

supported by the hardware. Section 5.3.4 discussed such an approach based on uniform K-space sampling. Additionally, even limited bandwidth at a low-frequency range can be used to improve the results as shown in Fig. 5.31. The results show accurate image reconstruction with two objects in limited bandwidth of 860 – 890 MHz, which failed to reconstruct accurately in 900 – 925 MHz. This behavior is assumed to be attributed to a shift in antenna characteristics when placed close to each other and in presence of nearby PEC objects. However, this needs to be further studied.

The real-scale setup has increased noise, which needs to be studied carefully. Even after the removal of tag – Rx links with high phase deviation, the reconstructed results are noisier than the scaled model. One possible reason is an error in location and

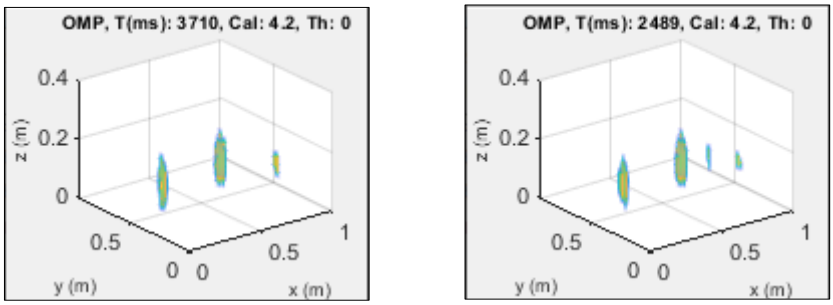


Figure 5.31 Frequency considerations: results of scaled simulation setup with objects at $a\&e$ locations, poorly detected in 900 – 925 MHz band, as shown in Fig. 5.20 (b). (a) 6 frequencies in range 850 – 900 MHz. (b) 3 frequencies in range 860 – 890 MHz. In both cases, two objects can be clearly detected at correct locations.

phase center estimates of tags and reader antennas. This is validated by adding noise of up to 5 cm to the x, y, z coordinates of the tag and reader antenna locations in the simulation dataset. Fig. 5.32 shows the effect of this error, which shows correct object count and locations when no error, and wrong location of one object in presence of error.

The ensemble is a technique to do a weighted combination of different "weak" estimators to generate a "strong" estimator. Additionally, a weighted combination of various "weak" estimated solutions with uncorrelated noise results in lower overall noise. Our problem setup proposes improved solutions at all the various stages of the problem: model, calibration, and algorithm. Thus, an ensemble of different models, calibrations, and algorithms intelligently can result in an overall improvement. Ensemble performance on the scaled setup data showed improvement over using one or the other method. However, the true scale image generation can be very noisy at times, and thus requires one-out-of-all selection instead of the ensemble.

5.7 **Conclusion**

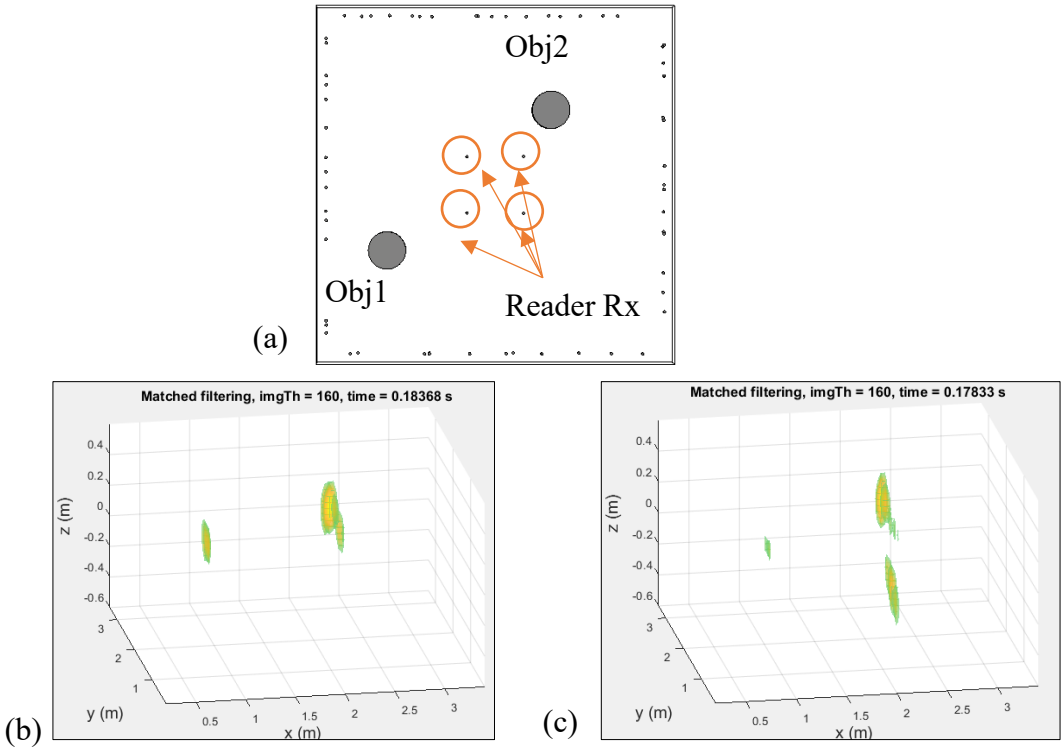


Figure 5.32 Phase and tag location noise tolerance. (a) True scale simulation setup with randomly placed tags and four reader antennas on the ceiling. (b) MF image results with accurate two object counting and location estimate. (c) MF image with introduction of random error between 0 to 5 cm in tags and receiver antennas.

In this work, a lot of effort has been put towards indoor RF imaging algorithm and signal processing, including problem model, algorithmic solution, and practical noise concerns in an indoor heavy multipath environment. This is a difficult problem, particularly with RFID system, with both inherent and added phase noise. But with increased RFID utilization and low-cost, this setup is of particular interest, and special attention has been paid to understand the noise characteristics of the collected dataset. All the presented methods are well applicable to any new hardware that performs multistatic, multifrequency imaging. The proposed sparsity-based OMP solution will further improve with a low-noise dataset. Also, this work has studied and compared two

different models: reflection and attenuation based. While the reflection model generates high-resolution 3d images, performance degrades in presence of high phase noise. The attenuation model requires many active reader units to generate a 3D image with uniform link density over the entire imaging volume. But the resolution is poorer. Thus, combination approaches have been proposed that show improved performance than any single method. Overall, the entire approach results in improved detection of one occupant, and their location without any fingerprinting, in near-real-time. This approach is ideal for an assisted living setup where one or at most two people are expected to live together and their approximate locations can help keep track of activity level without any privacy concerns.

CHAPTER 6

INDOOR OCCUPANT COUNTING USING DEEP-LEARNING

6.1 Introduction

With the ubiquity of the RFID tags, in the previous chapter, we have explored them for indoor occupant counting and location estimate of up to two occupants. The update is made in near real-time with data collection taking the majority of the time. With faster sampling, tracking algorithms can be included to improve the accuracy of the location estimate. This can be very helpful for assisted living applications. This chapter adds more capabilities to the RFID system, by using the same setup for accurate counting for a greater number of closely located occupants. This adds another functionality to the occupant-centered control (OCC) [183] where occupant monitoring can enable various automation functions like HVAC (heating, ventilation, and air-conditioning control) energy saving [184], security, and behavior analysis. Among these energy saving is a major topic and requires accurate counting of the occupants in a passive device-free manner. Key requirements for this passive sensing include low cost, low power, low computation, small training overhead, and high accuracy. Many indoor items will have RF links in the realization of the IoT, which can further facilitate the development of a smart environment.

In this work, occupant counting is performed using carrier phase and RSSI information as inputs to a convolutional neural network (CNN) based deep-learning (DL) model [204], without requiring detailed feature extraction. The proposed method can achieve high accuracy even with limited training data. The approach is verified by setting a testbed in a 150-ft^2 room with drywall and metal beams in the ceiling, without

any control on human traffic in the outside hallway or the nearby classroom. Data were collected with 0 – 5 participants either sitting or standing at different locations in the room, representing a typical home or office scenario. Further, we have proposed a calibration approach to nullify the effect of furniture, walls, and unknown transceiver and cable phase offsets to extract weaker occupant-reflected signals in presence of strong multipath, which was presented in the previous chapter. The calibrated RSSI and carrier phase information gives improved counting accuracy. In contrast to the traditional fingerprinting approach, we trained the CNN model with participants at a few locations and tested at unseen locations. The participants in training and testing cases were stationary, not moving randomly, which reduced the required data-collection time by allowing average over a smaller period. We have also explored tuning the same CNN model to perform accurate counting in different home settings with a single occupant, without requiring extensive new training data collection. High counting accuracy can still be maintained under placement variations of tags, receivers, and furnishing. The occupancy detection accuracy is further tolerant to the participant in the quasi-stationary or walking state, demonstrated in the home setting. The major contributions of this thesis include:

- A 2D CNN-based DL method that can learn features embedded in ambient RSSI and phase.
- A signal calibration model for RSSI and phase correction by estimating only the multipath from occupant-reflection.
- Use of passive UHF RFID tags to increase sensing points with training only at a limited number of occupant-locations without extensive fingerprinting.

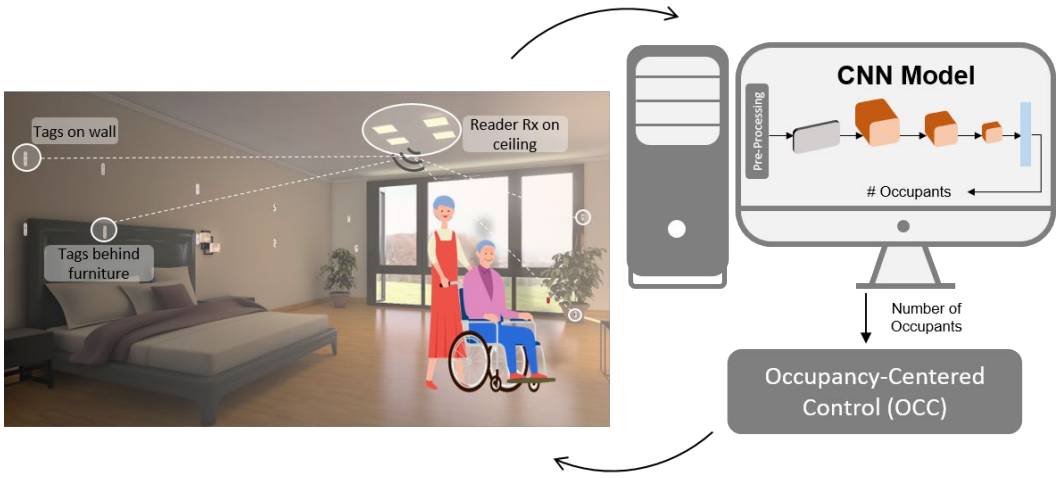


Figure 6.1 The occupancy counting setup using ambient passive RFID tags.

- Transferrable training from the test room to home setting with different furnishing and small validation data.
- Detection of the moving participant with training only from the stationary data.

Fig. 6.1 shows the setup summary and the entire processing pipeline. The next section presents the details of CNN architecture. Sections 3 and 4 show the experimental setup, data collection, and results.

6.2 Convolutional neural network (CNN) architecture

The learning process of CNN has been effectively demonstrated in the visual domain [185], [186]. For object recognition, the layers act as nonlinear filter banks with increasing complexity and details as the depth increases [187]. The shallow layers contain generalized information such as edge-detection filters, and the deeper layers learn the detailed object characteristics.

We have implemented an AlexNet-inspired [188] CNN model as shown in Fig. 6.2 (a). The model is implemented in Python with the open-source library PyTorch to learn information in the collected RSSI and phase. The data for each tag-Rx link is

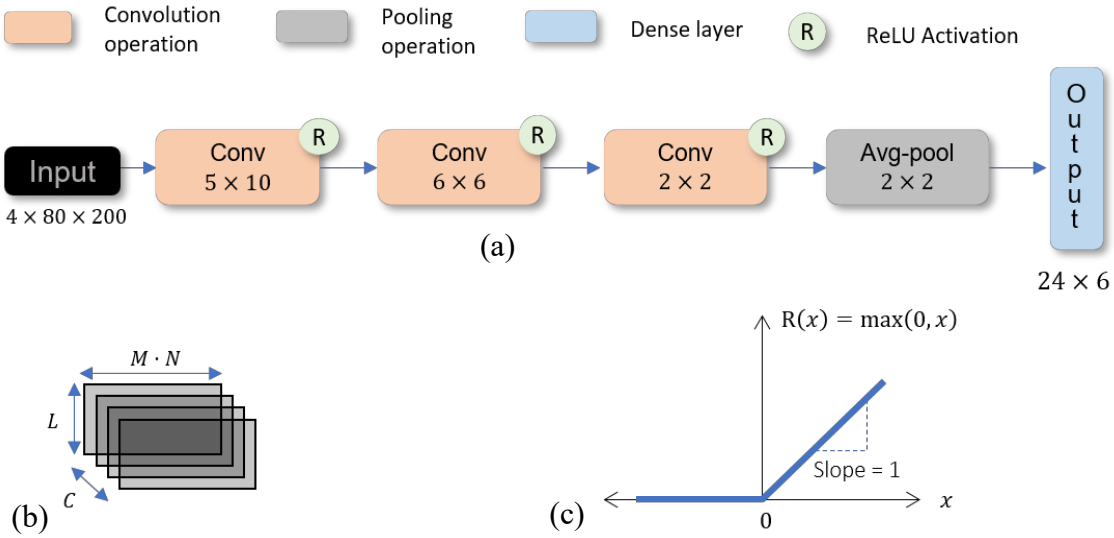


Figure 6.2 CNN architecture. (a) Network arrangement with three convolution layers, followed by a fully connected dense layer. The kernel sizes are shown in the boxes. Layer legend is shown at the top. (b) The arrangement of input dataset shown in 3D with $L = 80, M \cdot N = 200, \text{and } C = 4$. (c) ReLU function $R(x) = \max(0, x)$.

arranged in a 3-D format with RSSI and phase as the channels, C. We have explored using up to four channels: RSSI and phase collected in presence of occupants ($y^{Occ}_{RSSI}, y^{Occ}_{Phase}$) and the corresponding calibrated signal ($y^{Cal}_{RSSI}, y^{Cal}_{Phase}$). The rows contain reading from each tag, arranged over all the frequencies for each successive reader antenna in the columns. Thus, the data shape ($depth \times height \times length$) is given as $C \times L \times (M \cdot N)$, where $L, M \,\&\, N$ are the number of tags, reader antennas and frequencies respectively. With 80 tags, 4 reader antennas, 50 frequencies, and all 4 channels, the data are arranged as $4 \times 80 \times (4 \cdot 50)$, shown in Fig. 6.2 (b). This arrangement is designed to extract correlated information across the neighboring tags and frequencies and can learn effectively with increased spatial diversity of tags and Rx antennas, and the spectral diversity of frequencies. We started with a baseline

architecture consisting of three convolution layers and a fully-connected dense layer, and then progressively optimized the hyper-parameters to achieve the final model.

The convolution layers are defined by the *kernel* shape, i.e., the filter shape that performs convolution operation over the input data, *stride*, which indicates the filter sliding interval, and *padding*, the layer of zeros added to the data boundary that increases the data size over which convolution is performed. The *kernel* sizes for three layers are 5×10, 6×6, and 2×2 respectively, with large $strides$ of $[3,10]$, $[5,5]$, and $[1,1]$, respectively. For the first layer, as the nearby tags show a higher correlation, a small window size of 5 is selected in comparison with a window size of 10 across frequencies, which are more closely distributed in the spectral domain. The number of filters in the first layer was varied between 40 – 120, and we observed the performance initially improved and then dropped beyond 60. Thus, the first layer is selected to have 60 filters, then 30 in the next, and finally reduced to 12 in the last convolution layer.

A nonlinear activation function by the rectified linear unit (ReLU) [189] $\max(0, x)$, as shown in Fig. 6.2 (c), is applied at the output of the convolution layers. As network depth increases, ReLU allows more gradients to back-propagate through the model during training, updating each weight without the vanishing gradients problem commonly observed in other functions. In order to introduce regularization and avoid overfitting, we use the dropout layers [190] at the output of three convolution layers with probabilities of 0.4, 0.2, and 0.2, respectively. Dropout leads to sparse representations, even when no sparsity-based regularization is present, thus it prevents overfitting and acts as a regularization. At the output of the final convolution layer, we use an *average-pool* layer [191] that performs down-sampling by taking an average

across a 4×4 window, where the feature map decreases by $1/4^{th}$ along height and length dimensions and introduces the model's invariance to local transitions. Reduction in dimensionality in turn lowers the computation requirement with fewer parameters. While *max-pool* generally performs better for images, here *average-pool* performs better by smoothing the features. The flattened output of the pooling layer is provided as input to a fully connected layer of size 24×6, where 6 corresponds to the number of output classes for 0 – 5 occupants. The final output occupancy count is the node with the maximum value. We optimize a cross-entropy loss function that calculates a score summarizing the average difference between the actual and predicted probability distributions for all classes in the problem. If the probability of the correct class is low, the loss is high. If the correct class is predicted, a perfect zero loss is observed. An adaptive gradient algorithm, Adam, with decoupled weight decay regularization [192] of 0.2 and a learning rate of 0.001, is adopted for training.

Evaluation of different model hyper-parameters such as the number of layers, kernel size, etc., carries a risk of overfitting on the training dataset. Using test data for hyperparameter tuning introduces a bias, and the model may not be generalized as parameters can be tweaked to get maximum test accuracy. To solve this issue, we use K-fold cross-validation (CV) with $K = 5$. In this approach, training data are divided into K folds and K iterations are performed over the dataset. In each round, one part is used for validation, and the remaining $K - 1$ parts are merged into a training subset for learning. The CV accuracy is computed as the mean accuracy achieved over the K validation sets. As our data distribution across classes is biased, stratified CV is applied to maintain the same class distribution in each fold. The data are divided into non-

overlapping train and test sets. This holdout test set is not seen at all and CV is performed using the training data. The hyper-parameters are tuned to achieve the best CV accuracy. This model is then trained with the entire training data and gives good performance on the test set, as shown in later sections.

6.3 Experimental setup

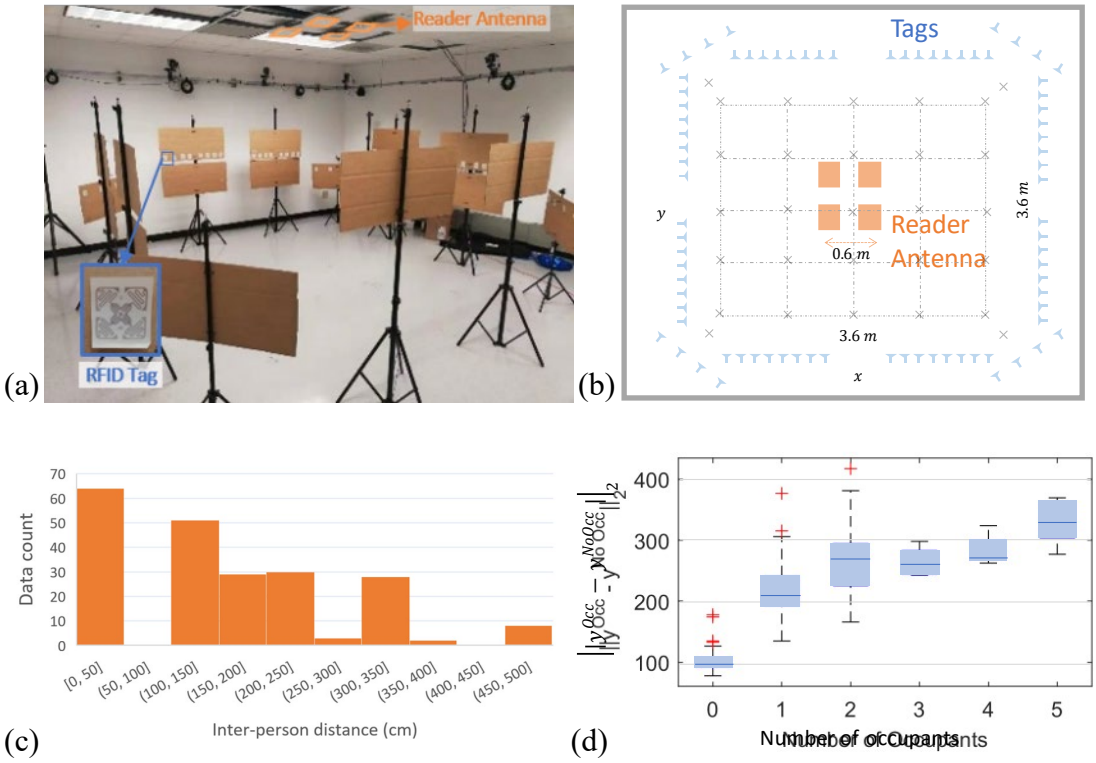


Figure 6.3 The lab test setup. (a) 80 RFID tags arranged on cardboard, with inset showing a magnified view of the passive paper tag. Except at the corners, tags are placed at a height of 1.24 m with a minimum distance of 10 cm between neighboring tags. The 16 corner tags are placed at a height of 1 m. Four reader antennas are placed on the ceiling. (b) The 2D top view and relative tag and reader antenna placement. The possible occupant locations are indicated by the cross symbol (×). (c) The histogram plot showing pairwise inter-person distance. (d) The box plot of the distribution of the norm $\left|\left|\boldsymbol{y^{Occ}} - \boldsymbol{y^{NoOcc}}\right|\right|_2$ as a function of number of occupants, with a clear increasing trend. The occupants are standing or sitting at different locations in the room. For 0 occupants, it is non-zero due to variations outside the test region.

In this work, we employed ambient passive RFID tags, without requiring any markers or devices on occupants. The test setup is the same real-scale room in the previous chapter, including an Impinj Speedway R420 RFID reader that can be connected to four reader antennas placed on the ceiling panels. 80 passive tags were uniformly distributed in a $3.6\,m \times 3.6\,m$ lab room as shown in Fig. 6.3 (a). The reader antennas were placed at four corners of an approximate square with an edge baseline of $0.6\,m$, as shown in Fig. 6.3 (b). The tag height of $1.24\,m$ was guided by antenna radiation pattern, as a tradeoff between area coverage and tag readability.

The reader acts as a transceiver and periodically emits RF signals in the 902 – 928 MHz range with 50 frequencies separated by 0.5 MHz. This EM energy is harvested by the tags to activate and backscatter the ID-modulated signal. The passive tags cost less than 10 cents each and can increase the RF link coverage without adding much to the overall cost in deployment and maintenance. The number of tags is only limited by the computation requirements and reader distance. The COTS UHF reader has a typical operating range of around 10 m, which is enough for a medium-sized room. Multiple readers can work collaboratively for larger spaces [174]. The reader collects tag IDs and the corresponding RSSI and phase values for each link, which is composed of a tag and a reader Rx antenna at a given frequency.

Our system works in three phases, data collection under different experimental conditions, pre-processing including calibration and data imputation of missed tag readings, and final counting by the DL model.

6.3.1 Data collection

The RFID reader follows EPCglobal C1G2 [193] [GS1] standards for tag inventory. A slotted ALOHA mechanism is used, under which each tag randomly selects a time slot for reporting. The reader employs a frequency hopping spread spectrum (FHSS) [194] across 50 carrier frequencies in the given bandwidth. Thus, RF links are read in a random order for different tags and frequencies for each reader Rx antenna, which is activated in a time-division manner. The impinged power on tags varies, depending on Rx-tag distance and ambient factors. Thus, our one-round sampling time is selected to accommodate FHSS, reader antenna time-division, and sufficient read rates for low-visibility tags. We implemented a two-minute data collection for an unoccupied room and took the average across the multiple readings for each link. With participants in the room, a one-minute data collection was performed. As a large number of closely distributed frequencies are not required, the sampling time can be significantly reduced with a custom reader.

Our initial test was in a lab-environment with people walking outside the hallway and high-occupancy neighboring classrooms. We performed experiments at different times of the day, on five days, distributed over three months. The room also had automated HVAC with fans on the ceiling. While the room had chairs, computers, and laboratory equipment, all the furniture in the active capture volume was removed during standing experiments, and chairs were provided for sitting participants.

The data collection protocol was approved by the Cornell institutional review board (IRB). The room was divided into 25 uniformly distributed grid points and 4 additional corner points as shown in Fig. 6.3 (b). We collected data on 9 participants with varying physical characteristics, with up to 5 occupants at a time in this mid-sized

room. The participants were instructed to stand or sit at different points. For 1 occupant, data were collected at the 25 uniform grid points. For ≥ 2 occupants, some location and posture combinations were selected based on the probability of occupancy in different parts of a room, with higher probability at the center, and the additional 4 corner points were selected as edge cases. The 0-occupant y^{Occ} data were essentially the data collected without any occupants in the room, taken at a different time than y^{NoOcc}.

The data collection included extreme cases – with all or selected occupants sitting or standing close at the same location, as well as when all were maximally separated. Fig. 6.3 (c) shows the histogram plot for pairwise inter-person distance when ≥2 occupants were present in the room. Table 6.1 shows the number of data points for each case and the percentage of sitting and standing occupants. A total of $N_{Data} = 206$ data points were collected, with more cases for ≤ 2 occupants. Fig. 6.3 (d) shows the 2-norm distance between complex signal vectors with and without occupants, $\left|\left|y^{Occ} - y^{NoOcc}\right|\right|_2$, which has an increasing trend for the number of occupants, caused by increasing multipath and body absorption. The norm with 0 occupants is lower than others but non-zero, showing a significant impact of surrounding changes on the multipath at different data collection instances.

Table 6.1 Total data distribution in lab room test

#Occupants (Class)	# Data	Posture	
		Sit	Stand
0	59	-	-
1	83	37%	63%
2	37	43%	57%
3	8	42%	58%
4	9	56%	44%
5	10	40%	60%

6.3.2 Calibration and data imputation

Before occupant detection, we extract the calibrated RSSI and phase as discussed in the previous section. No-occupant data, y^{NoOcc} were collected before each experiment with chairs for sitting participants. As the background subtraction introduces error for complete LoS blockage, we estimate the RSSI drop from y^{NoOcc} to y^{Occ} across all the frequencies. If more than 70% RSSI drop is observed across all the frequencies for a tag-Rx link, this link is considered blocked and thus exempted from the calibration process by assigning a zero reading.

In a room of moderate size, one reader can activate most of the tags in the empty room. However, with the multipath, angle-dependent antenna gain, and LoS blockage, a small number of tags may be unreadable. Also, when the frequency separation is more than the indoor coherence bandwidth, the multipath behavior is not correlated [195] and the same tag – Rx pair may not be read successfully at some of the frequencies. To take care of the missing information in y^{Occ}, average at other frequencies and neighboring tags are used for both RSSI and phase. Similarly, in the calibrated data y^{Cal}, non-overlapping unread instances of y^{Occ} and y^{NoOcc} may increase missing information. Furthermore, as stated previously, strong LoS blockage for some links might result in zero-readings, all of which are imputed using the following rules:

- If at least $1/3^{rd}$ of the frequencies are read, missed frequency readings for the same link are replaced by their average, as in the current setup the reader antennas were closely placed on the ceiling, observing similar multipath behavior.
- Otherwise, an average across the nearest 4 tags is taken where the phase variation is limited to $25°$ and the distance is limited to $0.25\ m$.

Both imputation rules are only approximations over the noisy measurements suffering heavy multipath and high channel loss due to poor visibility of the reader antenna. The phase and RSSI are further normalized as:

$$y_{Phase} = |y'_{Phase}|/\pi$$
$$y_{RSSI} = y'_{RSSI}/(0.05 \cdot \max y'_{RSSI}) - \text{mean}(y'_{RSSI}/(0.05 \cdot \max y'_{RSSI})). \quad (6.1)$$

In the above equations, the un-normalized data is indicated as y', and mean and max operations are taken across the entire dataset. As the RFID reader has a π phase ambiguity, the range is converted from $[-\pi, \pi]$ to $[0, \pi]$ before normalization, which also shows better algorithm convergence. The RSSI readings are mostly distributed in the range 0.02 – 10 nW (–77 to –50 dBm), with few readings close to the maximum value of 112 nW. Hence a factor of 0.05 is included so that most of the readings are normalized to the range $[-0.5, 0.5]$. The distributions of normalized RSSI and phase data before and after imputation are shown in Fig. 6.4, where the y-axes denote the number of instances.

6.4 Results from deep learning

In this section, we demonstrate effective occupant-count learning at untrained locations. We then investigate the effectiveness of calibration and data imputation. We also use a partial selection of tags to show that the number of tags can be significantly

reduced for occupant counting. Finally, we demonstrate learning transference to different setups in the case of single occupancy.

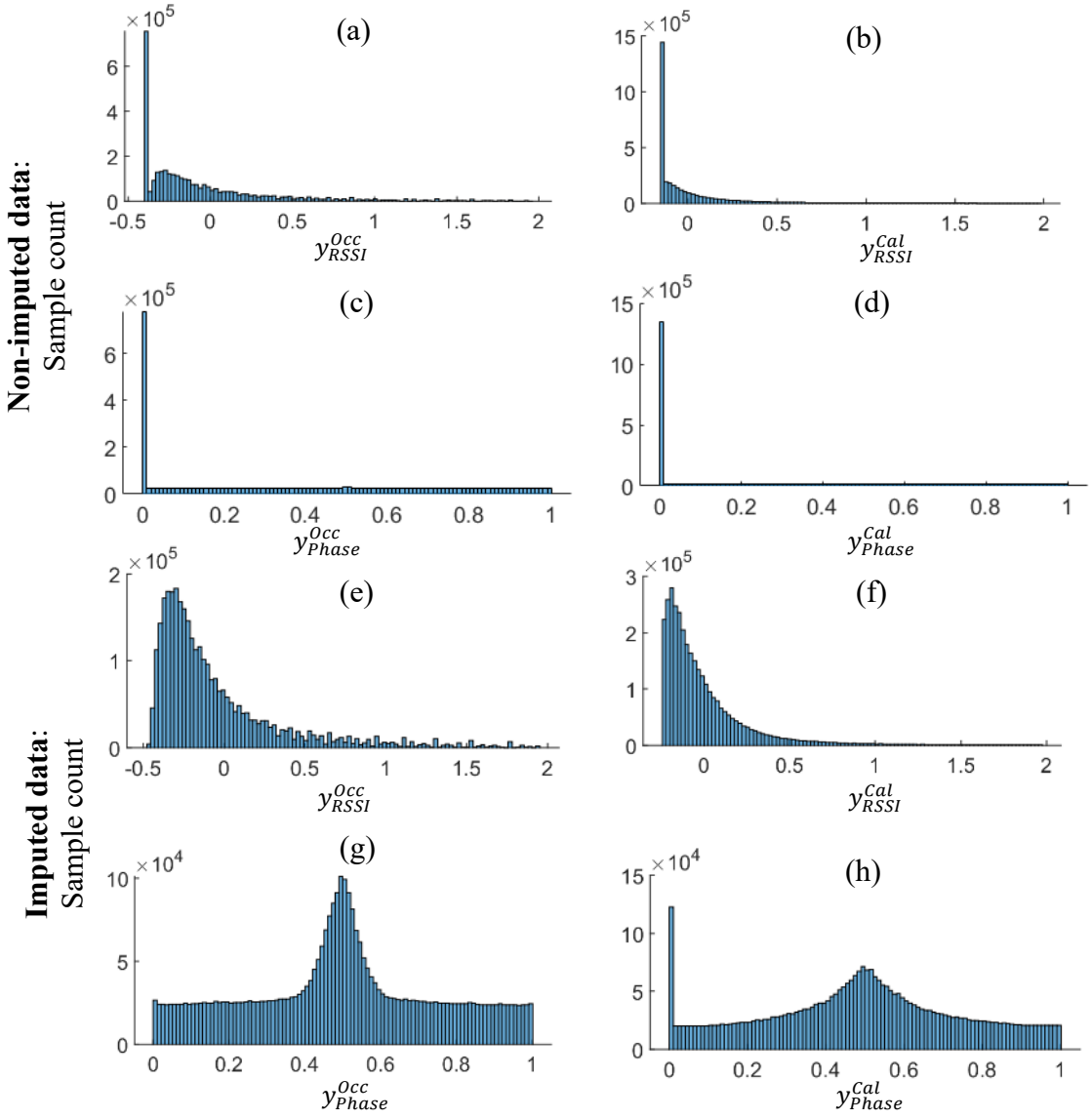


Figure 6.4 Data distribution without any imputation (a) – (d), and with imputation (e) – (h), showing RSSI and phase with occupants, and after calibration. While most zero-data are removed in y^{Occ}, some zero-readings remain in y^{Cal}_{RSSI} and y^{Cal}_{Phase}, when near-by tags also contain zero which may occur in case of large LoS blockage across all frequencies for a tag – Rx link.

The algorithm is implemented in Python using the PyTorch library for the CNN model on an Intel i7-8700 CPU with 16 Gb RAM and NVIDIA GeForce GTX 1050 Ti GPU.

6.4.1 Location-independent learning

In this paper, we have emphasized counting quasi-static and static occupants because moving participants have other ready solutions and might induce ambiguous

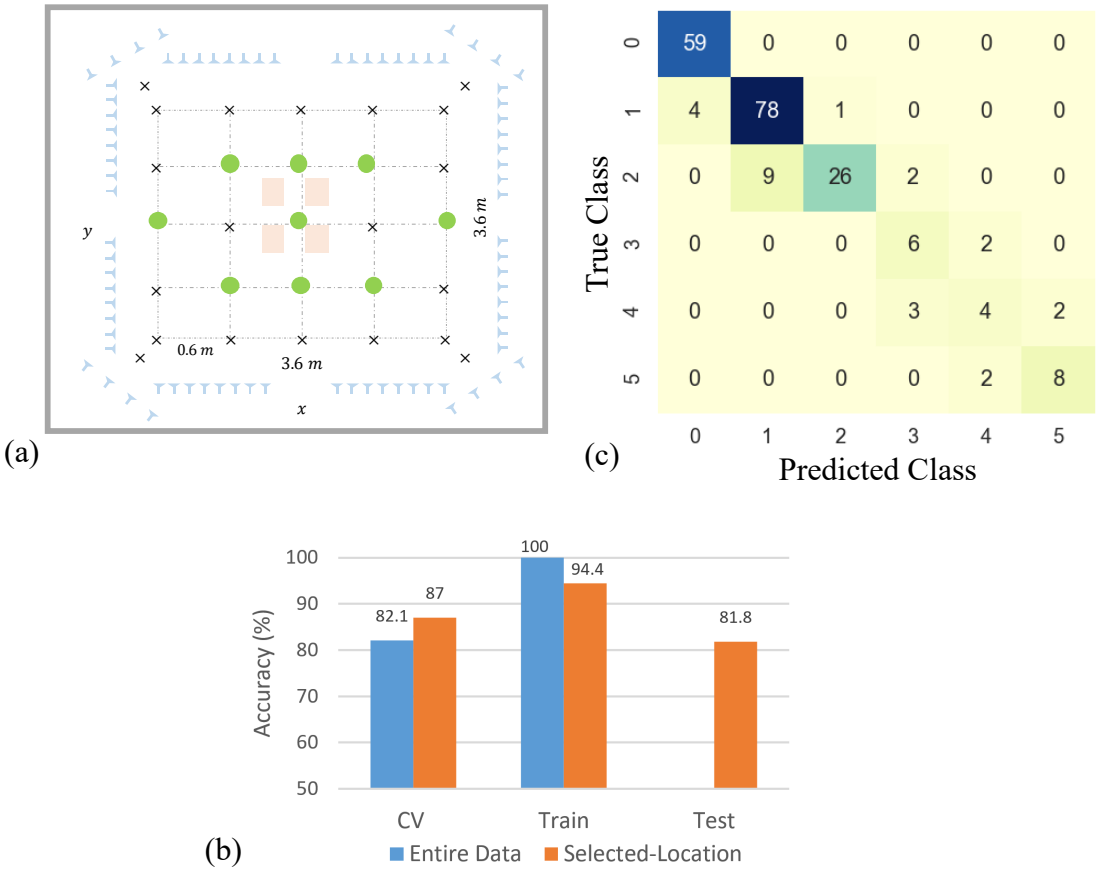


Figure 6.5 Selected-location training results, showing generalized location-independent learning with non-overlapping occupant locations in train and test data. (a) Top view with all possible occupant locations indicated by a cross (×). Selected training locations are shown with green dots. (b) CV, train, and test accuracies of the trained model, compared to when the entire data are used for training and CV. (c) The confusion matrix on the entire dataset with the location independent learning, showing high probability of the predicted number of occupants to be within ± 1 of the true number, $P(d \leq 1) = 1$.

Table 6.2 Train – test set distribution with selected – location training

#Occupants (Class)	Train set			Test set		
	Sit	Stand	#Data	Sit	Stand	#Data
0	-	-	29	-	-	30
1	13%	39%	43	24%	24%	40
2	30%	24%	20	14%	32%	17
3	42%	33%	6	0%	25%	2
4	28%	28%	5	28%	16%	4
5	36%	4%	4	4%	56%	6

interpretations in our initial trials. For one occupant, fingerprinting-based methods can be reasonably effective [196], but the increasing number of participants can readily overburden the training cost. The occupancy probability at different room locations is not homogeneous due to furnishing, lighting, and access to doors, which further complicates the selection of training cases. In our method, as the empty-room calibration is effective, we hence hypothesize that learning based only on selected grid points can be applied to detecting occupants at other locations not seen during the training. Fig. 6.5 (a) shows the grid locations selected for the training with green dots. For one occupant, there is no overlap of the training and testing locations. For ≥ 2 occupants, the training set included data with at least $\lceil \frac{\#occupants}{2} \rceil$ at the selected locations. Table 6.2 shows the number of train and test data points corresponding to each class based on this criterion.

With the data separated into training and location-holdout test sets, the CNN model is established using 5-fold CV for hyper-parameter tuning, without any bias of the holdout data. We compared the accuracy of this model with 5-fold CV training on the entire dataset, as shown in Fig. 6.5 (b). Similar CV accuracy is observed in both cases, with a higher accuracy of 87% in the case of selected-location training, also resulting in a high

Table 6.3 Probability $P(d \leq i)$ estimation under data imputation and channel variation

Data	Test Results		
	$P(d=0)$	$P(d \leq 1)$	$P(d \leq 2)$
Non-Imputed	0.77	0.98	1
Imputed ($C=4$)	0.82	1	1
y^{Occ} ($C=2$)	0.77	0.95	1
$y^{Cal}(C=2)$	0.78	0.94	0.98
y_{RSSI} ($C=2$)	0.79	0.97	1
$y_{Phase}(C=2)$	0.66	0.90	0.94
$y_{RSSI}^{Occ}(C=1)$	0.77	0.92	0.95
$y_{RSSI}^{Cal}(C=1)$	0.77	0.95	0.98

location-holdout test set accuracy of 81.8%. We define accuracy as the fraction of cases with a correct prediction of the number of occupants, i.e., the difference (d) between the ground truth and prediction is 0, denoted by the probability $P(d=0)=0.818$. Our model learns well, with all predictions within ± 1 of the true value, $P(d \leq 1)=1$. Fig. 6.5 (c) shows the confusion matrix on the entire dataset, including both train and test sets, for location-independent learning. For all the following comparisons, we will use this training setup as the baseline. The computation cost is low with only 2.7 s for training.

6.4.2 Impact of data imputation and calibration

Fig. 6.6 (a) shows the comparison of training on imputed and non-imputed data, using all four channels. Imputed data shows slightly higher CV and location-holdout test accuracy of 87% and 82%, compared to 83% and 77%, respectively. Table 6.3 shows the values $P(d \leq i)$ for $i = 0, 1, \text{and } 2$. The improvement from imputation is limited, as while we removed random reading defects, embedded information of

occupant shadowing may be partially lost. For all subsequent comparisons, imputed data are used.

We tested the use of different numbers of channels, $C = 1, 2, \text{or } 4$ in the model to quantify the contribution of both calibrated and un-calibrated RSSI and phase. Fig. 6.6 (a) shows the accuracies with different channels, and detailed probability values are

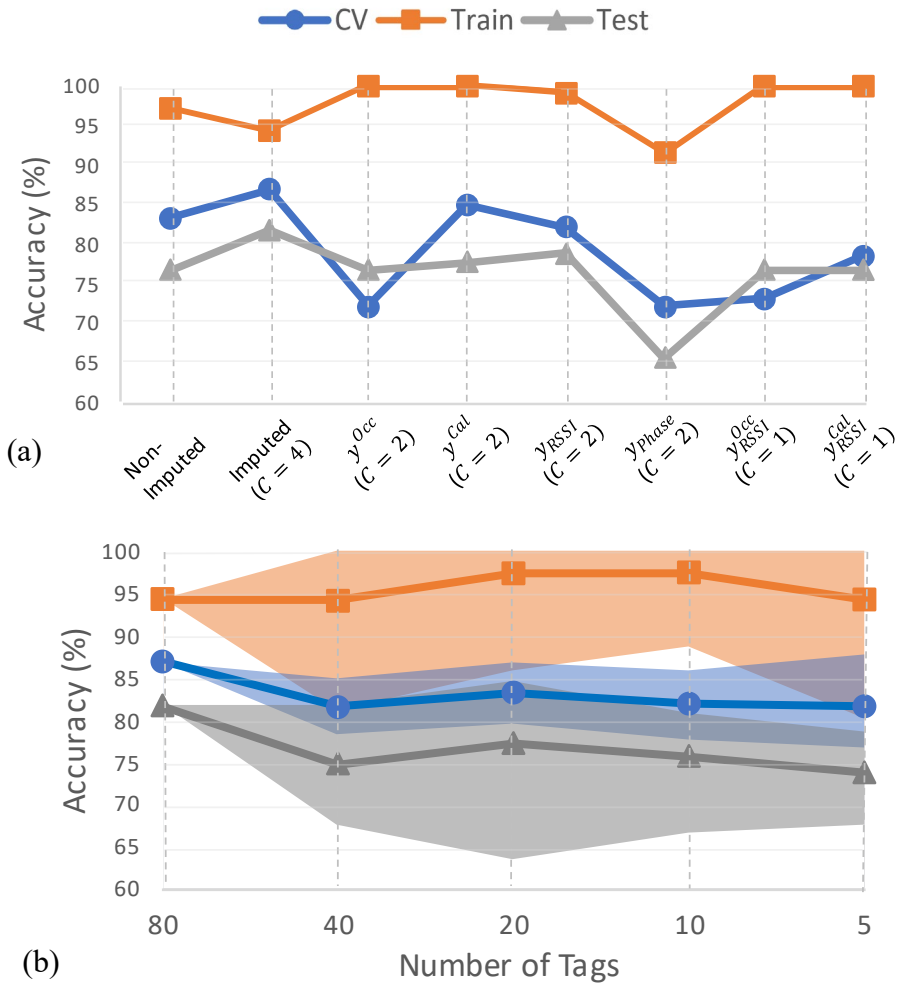


Figure 6.6 CV, train, and test accuracies with different variations of model and data. (a) First two points show comparison of non-imputed and imputed data with all four channels (C), y^{Occ}_{RSSI}, y^{Occ}_{Phase}, y^{Cal}_{RSSI}, and y^{Cal}_{Phase}. Fewer channels are tested, showing clear improvement with calibration. (b) Comparing results with fewer tags. Solid lines show average accuracy with different tag selections and the bands show minimum and maximum accuracies around the average. High accuracy is observed even with limited number of tags, but with higher probability of data loss.

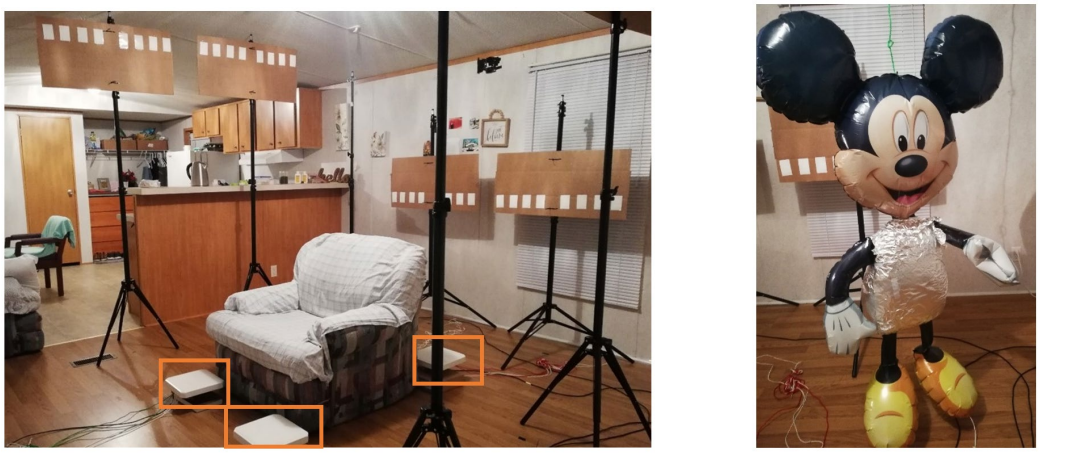
summarized in Table 6.3. Using two calibrated channels with y_{RSSI}^{Cal} and y_{Phase}^{Cal}, denoted as $y^{Cal}(C = 2)$, shows the highest CV accuracy next to using all 4 channels. As the calibration algorithm ignores multi-occupant RF reflection, scattering, and shadowing, the model performs best for the number of occupants ≤ 2. With only RSSI data, y_{RSSI} consisting of y_{RSSI}^{Cal} and y_{RSSI}^{Occ}, higher accuracy is achieved compared to only-phase based learning, suggesting RSSI has less ambiguous information. Comparison of using only RSSI data shows that y_{RSSI}^{Cal} performs better than y_{RSSI}^{Occ}, but poorer than both y_{RSSI}^{Cal} and y_{RSSI}^{Occ}. Overall, the best performance is achieved with all four channels.

6.4.3 Impact of tag density and placement

In our present lab room test setup, we distributed tags uniformly across the room. However, in real-life implementations, it is important to consider the use of fewer tags with random placement. Out of the 80 tags, we successively decreased the number of tags to 40, 20, 10, and 5 with random selection. We considered 25 different combinations of tags and averaged over all combinations to get the final CV, train, and test accuracies shown in Fig. 6.6 (b). The accuracy range is shown with colored bands around the solid line. When all 80 tags are considered, 81.8% test accuracy is achieved. However, with one of the randomly placed arrangements of 20 tags, results surpassed that of 80 tags, with 84.8% test accuracy. The average test accuracy with 20 tags is also the next best with 77.4%. Even the case of 10 tags shows good learning capabilities with >75% average test accuracy. However, with fewer tags, LoS blockage and tag distribution need to be carefully studied in the future.

6.4.4 Learning transference to the home setting

Our test design includes multi-occupant counting in a general mid-size office or home setup. As the train data collection is a significant overhead for every new setup, we used a home living room with a single occupant to investigate learning transference, as shown in Fig. 6.7 (a) (Home-I). The setup was constrained by the social-distance rule



(a) (b)

(c) (d)

Figure 6.7 Alternate home setups for occupancy detection and motion experiments. Four reader antennas (orange box) and 64 tags are placed in all the setup variations. (a) Home-I is nearly an upside-down version of the lab setup, with 16 tags on each wall, without the corner tags. Tags are placed at different heights on opposite wall pairs. A sofa was put in the middle. (b) A human-sized balloon of height 1.4 m pasted with aluminum foils to better mimic scattering from human body. This is used with setups (a) and (c) in standing and lying down poses. (c) Alternate Home-II, where the lower-height tags on two walls are rotated 45o as shown. (d) Home-III setup with arrangement of both reader antennas and tags different from the lab setup. A participant performs different motion routines in the monitored area of approximately 1.8 m × 1.8 m.

Table 6.4 Probability $P(d \leq i)$ estimation to show learning transference

Setup	Selected Channels (C)	Test Results	
		$P(d = 0)$	$P(d \leq 1)$
Lab	$C = 4$	0.82	1
Home-I	$C = 4$	1	1
Home-II	$C = 2: y_{RSSI}$	0.85	1
Home-III	$C = 2: y_{RSSI}$	1	1

during the COVID-19 pandemic and had the reader antennas on the floor in a larger rectangular arrangement of sides $0.75\ m$ and $1.3\ m$. A sofa was placed at the center of the room, and the room layout had an open kitchen, generating different multipath. The 64 tags were placed similarly, with different heights on the opposite wall pairs, without the 16 tags at the room corners. The room's occupancy was tested with a plastic phantom balloon of height 1.4 m, as shown in Fig. 6.7 (b), pasted with aluminum foils to increase RF reflectivity. The balloon was suspended in standing and lying-down positions at different grid points to collect 21 occupied and 20 unoccupied data points. The second home setup (Home-II) was created by rotating tag-holding boards on two walls by 45° as shown in Fig. 6.7 (c). In the final setup variation (Home-III), reader antennas were placed at the four corners with tags brought closer together in a two-tier arrangement in the z-direction, without any furniture, as shown in Fig. 6.7 (d). One participant performed various motions including arm movements, squats, and slow walking inside the area at a rate of approximately 0.7 m/s. The data pre-processing and normalization were performed as previously with a slightly faster data collection of about 48 s per data point.

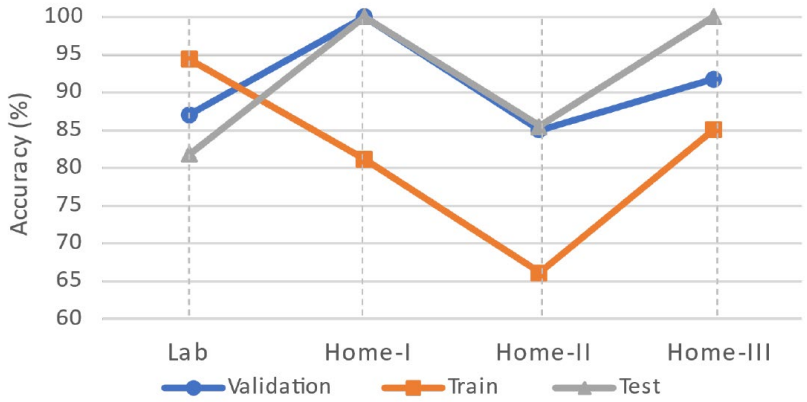


Figure 6.8 Validation, train, and test accuracy comparison of various setups with the balloon in Home-I and Home-II, and people motion in Home-III.

For the above setup variations with single occupancy, we use the entire data with 0 – 5 occupancy from the first lab set up to train and 1/4th of the new data for validation. An average of four different validation data sets is taken and reported as validation accuracy. Here, this is not denoted as CV, because the validation data is not part of the training data, with the main purpose of hyper-parameter tuning instead of training. The training is performed using 64 wall tags from the lab room setup. As the tag and reader positions varied a considerable amount from the lab setup to Home-II and Home-III, phase-based channels performed poorly during the validation and were not used for model learning. The results are shown in Fig. 6.8 and probabilities are summarized in Table 6.4. Good accuracies are reported for the balloon in different postures, as well as for the moving participant. It is to be noted that the output possible classes are still 0 – 5, and correct 0/1 detection is performed with no over-counting, and with 100% accuracies in Home-I and Home-III setups. Home-II setup shows a slightly poorer accuracy of 85%, with only up to ± 1 variation in the occupant count.

6.5 Simulation to experiment transfer learning

This section explores the possibility of entirely removing the training data collection process and perform real-scale simulations to generate training data. While the idea is attractive, performing real-scale simulation replicating the experiment scene is not possible with the limited available computation and memory power. Thus, some approximations need to be made, while understanding the repercussions of each assumption. Following approximations need to be performed:

1) Antenna: Fig. 5.3 showed that the tag antenna has a monopole pattern and can be easily substituted by the dipole antenna. Reader antenna, however, needs to be a patch antenna to correctly simulate antenna beamwidth. Thus, the antenna shown in Fig. 5.4 is selected.
2) Boundary conditions: Simulating concrete walls is computationally intractable for a large room, thus 'open add space' with maximum reflection level of 0.01 is selected.

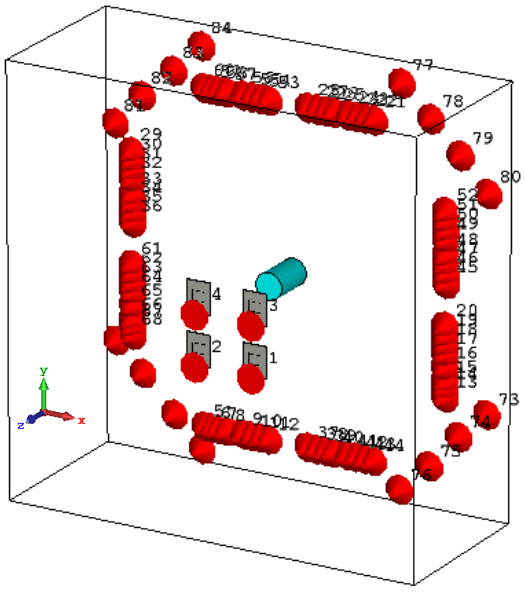


Figure 6.9 CST microwave setup replication the experiment setup.

3) Object material and size: Human muscle or even water results in a lot of meshing, but PEC is very reflective. Thus, a cylindrical torso of height 0.5 m, radius 0.15 m and material with relative permittivity, $\epsilon = 55$, and electrical conductivity of $\sigma = 0.948\ S/m$ is selected, with no material based mesh refinement or tangent loss.

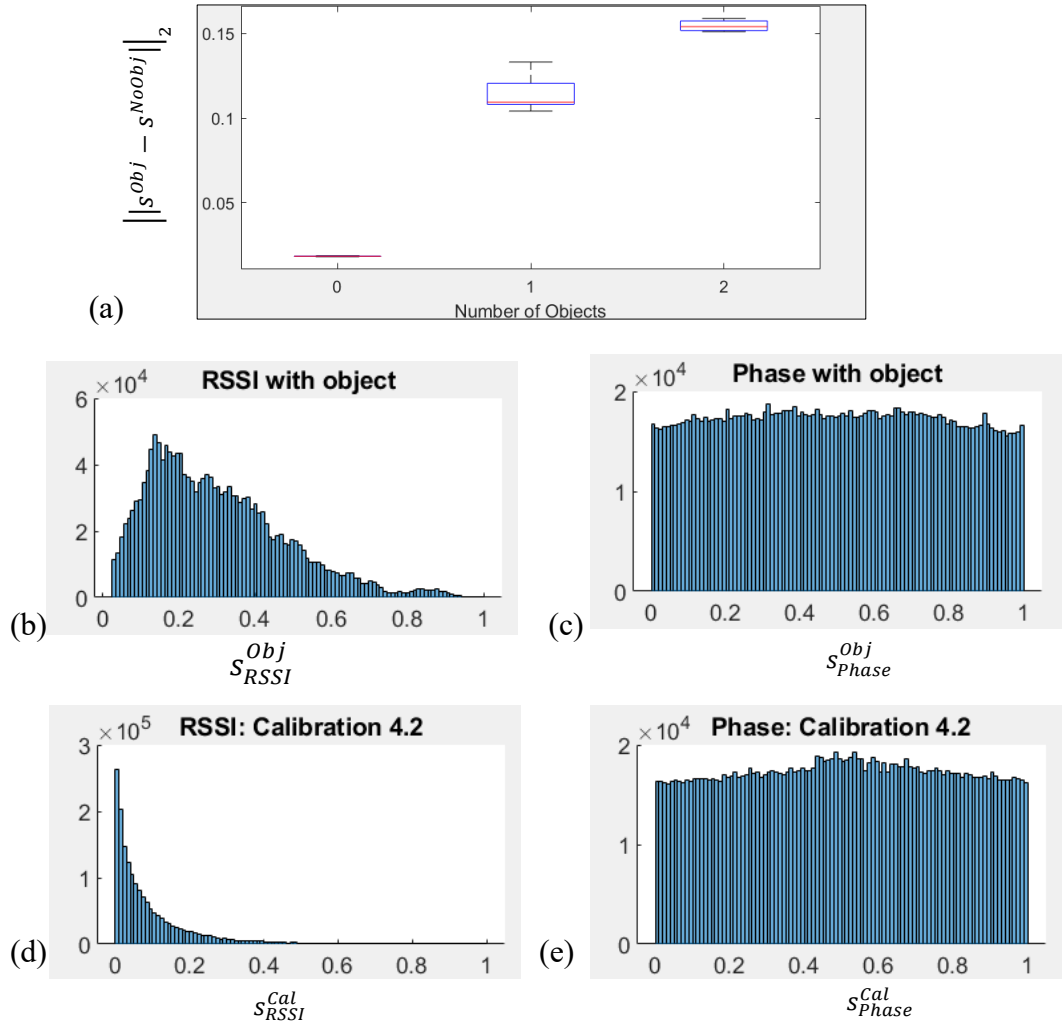


Figure 6.10 Simulation data characteristics. (a) Norm of data with and without object. (b) – (e) Histogram of RSSI and phase of the simulation data, similar to Fig. 6.4.

4) Amplitude and phase noise: These are added to have similar $\left|\left|y^{Occ} - y^{NoOcc}\right|\right|_2$ as seen with experiment data. With s as the simulated complex signal at the four reader antennas,

   a. $|s_{noisy}| = |s| + (|s| * \alpha * N(0,1))$, with $\alpha = 0.005$ as amplitude noise factor.

   b. $\angle s_{noisy} = \angle s + N(0,1) * \pi/180$.

Fig 6.9 shows the simulation setup with the object, and it takes about 9 – 10 hrs per simulation. Data is collected with no object (1 case), one object (6 non-symmetric locations), and two objects (4 data). Using the above amplitude and phase noise strategies, data can be augmented. Fig. 6.10 (a) shows the norm of simulated data, with 36 cases each of 0, 1, and 2 objects. Fig. 6.10 (b) – (e) show the corresponding RSSI and phase histograms with the object and calibrated. As the tag antenna transmitted signals are Gaussian pulses, the phase at different frequencies cannot be matched similar to the experiment, and hence only RSSI based features are transferable.

Fig. 6.11 (a) shows the confusion matrix of the test simulation data performance when the same CNN model is trained with the experiment data and testing done on the augmented simulation data using the two RSSI based features. It shows very good learning across 0 vs all object prediction, as well as for one object. 50% of the two object cases are predicted as one object, which needs further improvement. Fig. 6.11 (b) shows the confusion matrix for the testing experiment data when augmented simulation data is used for training. In this case, there is correct detection for ≥ 2 occupants as 2, with only 5 two-occupant cases predicted as 1. The 0 vs all prediction is also very accurate. The detection of one occupant however is confused with the two occupants, which needs

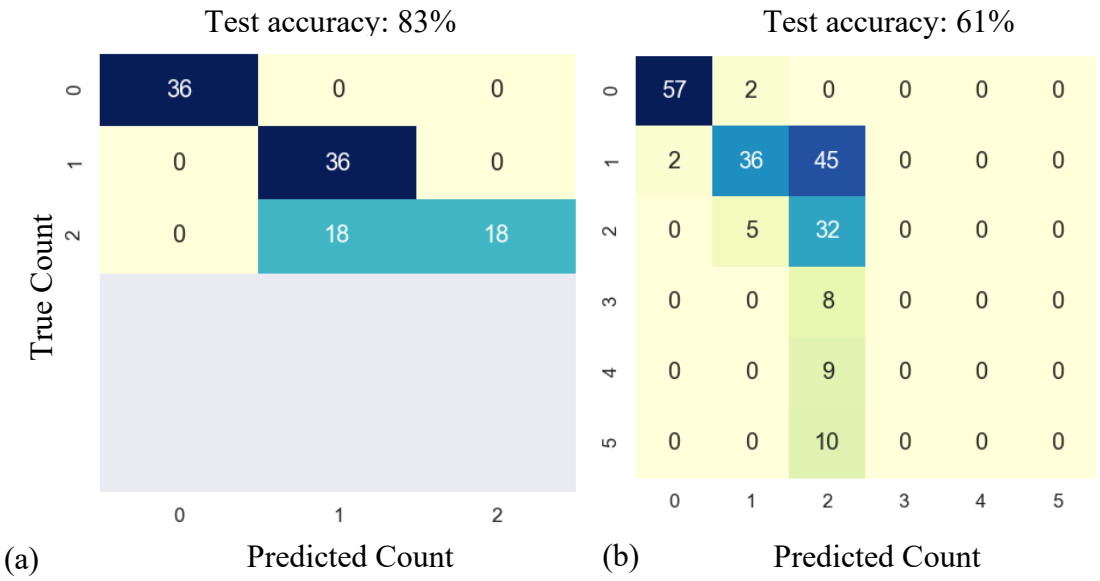


Figure 6.11 Simulation-experiment transfer learning confusion matrices. (a) Testing on simulation data. (b) Testing on experiment data.

to be further improved. In both cases, learning was stopped when training accuracy either stayed 100% for some iterations or when the growth was very small with each iteration. Validation with some of the test data can further help tune the hyperparameters. Overall, even with a very limited simulation dataset, it is possible to achieve satisfactory results. This simulation transfer learning is a great way to reduce the time and effort required in training data collection for future application purposes.

6.6 Conclusion

In this paper, we presented an indoor occupancy counting algorithm using a CNN-based DL architecture, which can learn features from RSSI and phase data. To the best of our knowledge, this is the first non-fingerprinting-based method to count ≤ 5 stationary occupants in an indoor environment. Further, comparable performance has been observed in camera-based systems under the same number of occupants in the

room [212], [213]. We have presented realistic test setups with comparisons of different channel selections, showing clear learning improvement with the calibration process. Also, good accuracy can be achieved with fewer tags, as small as 10.

Another major contribution of this work is the demonstration of occupancy learning transference to another similarly sized room, tolerating small variations in the placement of Rx and tag antennas. As multipath behavior can change significantly with the room layout and people outside the monitored region, the learning transference indicates generalized learning of the CNN model. Good performance is observed even with the participant in motion. Further tests need to be conducted with a larger number of occupants to identify learning transfer for more than one participant.

In continuation of learning transference, we have successfully demonstrated an alternative to experimental training data, by using only simulation data. Even with approximate simulations, similar data distribution can be achieved by normalization and appropriately added noise.

Future challenges and possible solutions include:

- Calibration: The present method requires calibration to nullify the effect of objects and furniture with high reflectivity. Thus, any change in large metallic furniture needs to be re-calibrated, which can be performed by frequent comparison of known no-occupancy data to updates. Also, as the RF channels are less affected by wood, cloth, and plastic, data are tolerant of furniture with small metallic content, as seen in the home setting.
- Biased dataset: Our dataset has a small number of samples for ≥ 3 occupants, introducing a bias in the learning model. As it is labor-intensive to collect data with

a large number of occupants across the grid, solutions like generative adversarial networks (GANs) [197] can be explored to formulate synthetic data mimicking the collected experimental data.

- Imputation: Our proposed imputation model is not effective in retaining shadowing or fading information; hence the approach needs to be further improved. Imputation is helpful with any defective tags over time, e.g., by a nail on the tag.
- Person size variation: Our CNN model learns RF perturbation features without identification of person size. Thus, a very large person may be counted as two, or some large pets may also be counted, as they can introduce sufficient phase fluctuation and RSSI attenuation. As one of the major applications is HVAC control, it is important to account for the overall heat load instead of the exact count.
- Noise characteristics: For simulation transfer learning, empirically defined amplitude and phase noises are added. This process should be further improved as this has drastic effects on learning performance.

Low-cost passive RFID tags have found increased usage towards keeping track of everyday objects such as pill-box and keys [70], [198]–[200]. Our presented approach can be readily augmented to benefit from the additional sensing points. Various methods have been proposed in the literature for tag locating [201]–[203] that can be used to locate these tags before the signal processing for occupant counting. Thus, our occupant counting system can be augmented to an existing IoT environment, enabling new functions in the applications of energy savings and assisted living.

CHAPTER 7

CONCLUSION

This thesis demonstrates the capabilities of RF sensors for two main assisted living necessities: *wellness* by respiratory and heartbeat monitoring, and *safety* and *accessibility* by occupant monitoring.

7.1 Contributions to respiratory monitoring

For respiratory monitoring, efforts have been put into sensor setup, design improvements, as well as understanding physiology, data collection, and processing. The summary of major contributions is as follows:

1. Proposed a two-point sensor measurement setup for obstructive sleep apnea monitoring by understanding the interplay of physiology and near-field RF sensing. Most noninvasive sensors do not focus on detecting apnea, due to complex paradoxical abdomen-thorax motion.
2. Developed an improved respiratory effort measuring sensor including not only the respiratory rate but also respiratory volume by a robust simple calibration scheme. Table 3.7 summarized the satisfactory performance of this sensor against other state-of-the-art noninvasive sensors including chest belts, strain sensor, camera, and far-field RF, for respiratory monitoring, along with good performance for heart rate estimation.

3. Validated accurate respiratory effort monitoring by a human study with 25 participants. This included protocol design, data collection, and developing robust signal processing tools for
    a. accurate peak estimation,
    b. signal quality detection,
    c. heart rate estimation, and
    d. motion detection.
4. Designed an on-the-bed sleep monitoring setup and installed it at the Cornell Weill sleep center, NYC, that performed reliable respiratory pattern monitoring under normal breathing and central sleep apnea. This remotely monitored setup provided a great understanding of one-sensor apnea measurement and its limitations.
5. Extended the respiratory effort monitoring setup to monitor the user's attention and relaxation states by understanding the relationship between attention and emotion recognition. Performed reliable feature extraction from respiration that is simpler than accurate HRV detection and subject independent.

7.2 Contributions to occupant monitoring

This thesis shows extensive work in the area of indoor RF occupant imaging and counting, focused on understanding the noisy signal model. For validation, CST simulations have been performed along with the analysis of the experimental data. Following are the major contributions:

1. Summarized and compared existing problem models for RF imaging and categorized them as attenuation and reflection models.
2. Proposed background calibration algorithm, with improved performance in the noisy multipath scenario.
3. Applied sparsity-based linear inverse problem solutions using ℓ_0 and ℓ_1 norm regularizations, with OMP giving superior performance. The knowledge of sparsity in the room is used to generate naturally sparse images without requiring threshold selection, which is difficult without knowing the true occupant count.
4. Developed a frequency selection algorithm based on uniform K-space sampling, that can help limit the required bandwidth, without degrading the image quality.
5. Proposed a novel CNN based deep-learning model for occupant counting, which is more tolerant to setup changes, allowing learning transference across different setups. This is further augmented by simulation learning, which can have huge potential by minimizing experimental requirements.

7.3 Future works

Both projects can have great real-life implications. Respiratory effort monitoring is not only useful for everyday health monitoring and sleep apnea diagnosis, but especially for geriatric patients with low cognitive function, where dyspnea information can only be currently determined from self-reporting or a caregiver's visual observation. With information about the respiratory pattern and comfortably placed on-the-bed sensor setup, NCS can be used to learn features associated with dyspnea-discomfort [100], [101]. A real-time sleep apnea detection algorithm would also be an important

work using the two-sensor approach. A deeper, theoretical understanding of RF interaction in the near-field is also critical.

The performance of the occupant monitoring setup is limited by the noise from the hardware setup. For occupant counting purposes, fewer tags are sufficient and thus, Bluetooth beacons can be used instead. For other purposes, custom hardware will greatly improve performance. With reduced noise, proposed sparsity algorithms will have much improved performances as well. A natural extension of this work is RFID tag integration in every object, with an improved tag localization algorithm. This will help identify when any furniture has been moved for calibration purposes, as well as learn user behavioral traits that can be helpful in smart occupancy-centered control.

BIBLIOGRAPHY

[1] World Health Organization, "Active ageing: A policy framework: World Health Organization," *Noncommunicable Disease Prevention and Health Promotion Department*, 2002.

[2] E. M. Matheson, D. E. King, and C. J. Everett, "Healthy lifestyle habits and mortality in overweight and obese individuals," *J. Am. Board Fam. Med.*, vol. 25, no. 1, pp. 9–15, Jan. 2012.

[3] F. Sanchis-Gomar *et al.*, "Physical inactivity and low fitness deserve more attention to alter cancer risk and prognosis," *Cancer Prev. Res.*, vol. 8, no. 2, pp. 105–110, Feb. 2015.

[4] K. M. Detering, A. D. Hancock, M. C. Reade, and W. Silvester, "The impact of advance care planning on end of life care in elderly patients: randomised controlled trial," *BMJ*, vol. 340, 2010.

[5] G. Mitchell, "Rapidly increasing end-of-life care needs: a timely warning," *BMC Med.*, vol. 15, no. 1, 2017.

[6] N. Rubio *et al.*, "Home monitoring of breathing rate in people with chronic obstructive pulmonary disease: observational study of feasibility, acceptability, and change after exacerbation," *Int. J. Chron. Obstruct. Pulmon. Dis.*, vol. 12, pp. 1221–1231, Apr. 2017.

[7] H. R. Colten and B. M. Altevogt, Eds., *Sleep Disorders and Sleep Deprivation: An Unmet Public Health Problem*. The National Academies Press, 2006.

[8] O. Oldenburg, B. Lamp, L. Faber, H. Teschler, D. Horstkotte, and V. Töpfer, "Sleep-disordered breathing in patients with symptomatic heart failure A contemporary study of prevalence in and characteristics of 700 patients," *Eur. J. Heart Fail.*, vol. 9, no. 3, pp. 251–257, 2007.

[9] D. Wang and H. Teichtahl, "Opioids, sleep architecture and sleep-disordered breathing," *Sleep Med. Rev.*, vol. 11, no. 1, pp. 35–46, 2007.

[10] J. Kim and E. André, "Emotion recognition based on physiological changes in music listening," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 12, pp. 2067–2083, 2008.

[11] B. L. Graham *et al.*, "Standardization of spirometry 2019 update an official American Thoracic Society and European Respiratory Society technical statement," *Am. J. Respir. Crit. Care Med.*, vol. 200, no. 8, pp. 70–88, 2019.

[12] W. J. DePaso, R. H. Winterbauer, J. A. Lusk, D. P. Dreis, and S. C. Springmeyer, "Chronic dyspnea unexplained by history, physical examination, chest roentgenogram, and spirometry: analysis of a seven-year experience," *Chest*, vol. 100, no. 5, pp. 1293–1299, 1991.

[13] J. Wanger *et al.*, "Standardisation of the measurement of lung volumes," *Eur. Respir. J.*, vol. 26, no. 3, pp. 511–522, 2005.

[14] L. M. Rofail, K. K. H. Wong, G. Unger, G. B. Marks, and R. R. Grunstein, "Comparison between a single-channel nasal airflow device and oximetry for the diagnosis of obstructive sleep apnea," *Sleep*, vol. 33, no. 8, pp. 1106–1114, Aug. 2010.

[15] R. T. Muza, "Central sleep apnoea-a clinical review," *J. Thorac. Dis.*, vol. 7, no. 5, pp. 930–937, May 2015.

[16] K. K. Tremper, "Pulse oximetry," *Chest*, vol. 95, no. 4, pp. 713–715, 1989.

[17] D. Biswas, N. Simoes-Capela, C. Van Hoof, and N. Van Helleputte, "Heart rate estimation from wrist-worn photoplethysmography: A review," *IEEE Sens. J.*, vol. 19, no. 16, pp. 6560–6570, 2019.

[18] Y. Zheng *et al.*, "Unobtrusive sensing and wearable devices for health informatics," *IEEE Trans. Biomed. Eng.*, vol. 61, no. 5, pp. 1538–1554, 2014.

[19] M. Chu *et al.*, "Respiration rate and volume measurements using wearable strain sensors," *NPJ Digit. Med.*, vol. 2, no. 1, p. 8, 2019.

[20] F. Q. AL-Khalidi, R. Saatchi, D. Burke, H. Elphick, and S. Tan, "Respiration rate monitoring methods: A review," *Pediatr. Pulmonol.*, vol. 46, no. 6, pp. 523–529, 2011.

[21] N. Netzer, A. H. Eliasson, C. Netzer, and D. A. Kristo, "Overnight pulse oximetry for sleep-disordered breathing in adults: a review," *Chest*, vol. 120, no. 2, pp. 625–633, 2001.

[22] O. T. Inan *et al.*, "Ballistocardiography and seismocardiography: A review of recent advances," *IEEE J. Biomed. Heal. Informatics*, vol. 19, no. 4, pp. 1414–1427, 2015.

[23] A. Taebi, B. Solar, A. Bomar, R. Sandler, and H. Mansy, "Recent advances in seismocardiography," *Vibration*, vol. 2, no. 1, pp. 64–86, 2019.

[24] Z. Xia, M. H. Shandhi, O. T. Inan, and Y. Zhang, "Non-contact sensing of seismocardiogram," *IEEE Sens. J.*, vol. 18, no. 14, pp. 5956–5964, 2018.

[25] E. Pinheiro, O. Postolache, and P. Girão, "Theory and developments in an unobtrusive cardiovascular system representation: Ballistocardiography," *Open Biomed. Eng. J.*, vol. 4, no. 1, pp. 201–216, 2010.

[26] C. S. Kim *et al.*, "Ballistocardiogram: Mechanism and potential for unobtrusive cardiovascular health monitoring," *Sci. Rep.*, vol. 6, pp. 1–6, 2016.

[27] I. Sadek, E. Seet, J. Biswas, B. Abdulrazak, and M. Mokhtari, "Nonintrusive vital signs monitoring for sleep apnea patients: a preliminary study," *IEEE Access*, vol. 6, pp. 2506–2514, 2018.

[28] E. Pinheiro, O. Postolache, and P. Girão, "Study on ballistocardiogram acquisition in a moving wheelchair with embedded sensors," *Metrol. Meas. Syst.*, vol. 19, no. 4, pp. 739–750, 2012.

[29] Carre Technologies inc (Hexoskin), "Hexoskin smart shirts - cardiac, respiratory, sleep & activity metrics." [Online]. Available: https://www.hexoskin.com/. [Accessed: 20-Feb-2020].

[30] T. Zeevi, A. Levy, N. Brauner, and A. Gefen, "Effects of ambient conditions on the risk of pressure injuries in bedridden patients—multi-physics modelling of microclimate," *Int. Wound J.*, vol. 15, no. 3, pp. 402–416, 2018.

[31] C. Bansal, R. Scott, D. Stewart, and C. J. Cockerell, "Decubitus ulcers: a review of the literature," *Int. J. Dermatol.*, vol. 44, no. 10, pp. 805–810, 2005.

[32] B. A. Reyes *et al.*, "Tidal volume and instantaneous respiration rate estimation using a volumetric surrogate signal acquired via a smartphone camera," *IEEE J. Biomed. Heal. Informatics*, vol. 21, no. 3, pp. 764–777, 2017.

[33] C. Li, V. M. Lubecke, O. Boric-Lubecke, and J. Lin, "A review on recent advances in Doppler radar sensors for noncontact healthcare monitoring," *IEEE Trans. Microw. Theory Tech.*, vol. 61, no. 5, pp. 2046–2060, 2013.

[34] K. Naishadham *et al.*, "Estimation of cardiopulmonary parameters from ultra wideband radar measurements using the state space method," *IEEE Trans. Biomed. Circuits Syst.*, vol. 10, no. 6, pp. 1037–1046, 2016.

[35] A. D. Droitcour, O. Boric-Lubecke, and G. T. A. Kovacs, "Signal-to-Noise ratio in Doppler radar system for heart and respiratory rate measurements," *IEEE Trans. Microw. Theory Tech.*, vol. 57, no. 10, pp. 2498–2507, 2009.

[36] D. Zito *et al.*, "SoC CMOS UWB pulse radar sensor for contactless respiratory rate monitoring," *IEEE Trans. Biomed. Circuits Syst.*, vol. 5, no. 6, pp. 503–510, 2011.

[37] P. Nguyen, X. Zhang, A. Halbower, and T. Vu, “Continuous and fine-grained breathing volume monitoring from afar using wireless signals,” in *IEEE INFOCOM*, pp. 1–9.

[38] F. Adib, H. Mao, Z. Kabelac, D. Katabi, and R. C. Miller, “Smart homes that monitor breathing and heart rate,” in *Proc. ACM CHI*, 2015, pp. 837–846.

[39] W. Massagram, N. Hafner, V. Lubecke, and O. Boric-Lubecke, “Tidal volume measurement through non-contact Doppler radar with DC reconstruction,” *IEEE Sens. J.*, vol. 13, no. 9, pp. 3397–3404, 2013.

[40] F. Pfanner, T. Allmendinger, T. Flohr, and M. Kachelrieß, “Modelling and simulation of a respiratory motion monitor using a continuous wave Doppler radar in near field,” in *Proc. SPIE Med. Imag.*, 2013, vol. 8668, pp. 37:1-37:12.

[41] D. Teichmann, A. Kuhn, S. Leonhardt, and M. Walter, “The main shirt:A textile-integrated magnetic induction sensor array,” *Sensors*, vol. 14, no. 1, pp. 1039–1056, 2014.

[42] P. Sharma, X. Hui, J. Zhou, T. B. Conroy, and E. C. Kan, “Wearable radio-frequency sensing of respiratory rate, respiratory volume, and heart rate,” *npj Digit. Med.*, vol. 3, no. 1, p. 98, 2020.

[43] P. Sharma, X. Hui, and E. C. Kan, “A wearable RF sensor for monitoring respiratory patterns,” in *IEEE EMBC*, 2019, pp. 1217–1223.

[44] X. Hui and E. C. Kan, “Monitoring vital signs over multiplexed radio by near-field coherent sensing,” *Nat. Electron.*, vol. 1, pp. 74–78, 2018.

[45] P. Sharma and E. C. Kan, “Sleep scoring with a UHF RFID tag by near field coherent sensing,” in *IEEE MTT-S Int. Microw. Symp. Dig.*, 2018, pp. 1419–1422.

[46] P. H. Gordon, R. Chen, H. Park, and E. C. Kan, “Embroidered antenna characterization for passive UHF RFID tags,” *IEEE RFID Conf.*, 2016. [Online]. Available: https://arxiv.org/abs/1710.02237.

[47] X. Hui, P. Sharma, and E. C. Kan, “Microwave stethoscope for heart sound by near-field coherent sensing,” *IEEE IMS*, vol. 2019-June, pp. 365–368, 2019.

[48] J. Y. Park *et al.*, “A critical review of field implementations of occupant-centric building controls,” *Build. Environ.*, vol. 165, p. 106351, 2019.

[49] H. Ghayvat, J. Liu, S. C. Mukhopadhyay, and X. Gui, “Wellness sensor networks: A proposal and implementation for smart home for assisted living,” *IEEE Sens. J.*, vol. 15, no. 12, pp. 7341–7348, Dec. 2015.

[50] K. Akkaya, I. Guvenc, R. Aygun, N. Pala, and A. Kadri, "IoT-based occupancy monitoring techniques for energy-efficient smart buildings," in *IEEE Wireless Communications and Networking Conference Workshops (WCNCW)*, 2015, pp. 58–63.

[51] Z. Chen, C. Jiang, and L. Xie, "Building occupancy estimation and detection: A review," *Energy Build.*, vol. 169, pp. 260–270, 2018.

[52] S. Yiu, M. Dashti, H. Claussen, and F. Perez-Cruz, "Wireless RSSI fingerprinting localization," *Signal Processing*, vol. 131, pp. 235–244, 2017.

[53] W. Wang, J. Chen, T. Hong, and N. Zhu, "Occupancy prediction through Markov based feedback recurrent neural network (M-FRNN) algorithm with WiFi probe technology," *Build. Environ.*, vol. 138, pp. 160–170, 2018.

[54] J. Yang, H. Zou, H. Jiang, and L. Xie, "Device-free occupant activity sensing using WiFi-enabled IoT devices for smart homes," *IEEE Internet Things J.*, vol. 5, no. 5, pp. 3991–4002, 2018.

[55] H. Zou, Y. Zhou, J. Yang, and C. J. Spanos, "Device-free occupancy detection and crowd counting in smart buildings with WiFi-enabled IoT," *Energy Build.*, vol. 174, pp. 309–322, 2018.

[56] W. Xi *et al.*, "Electronic frog eye: Counting crowd using WiFi," in *IEEE INFOCOM*, 2014, pp. 361–369.

[57] S. Di Domenico, G. Pecoraro, E. Cianca, and M. De Sanctis, "Trained-once device-free crowd counting and occupancy estimation using WiFi: A Doppler spectrum based approach," in *IEEE International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, 2016, pp. 1–8.

[58] S. Depatla, A. Muralidharan, and Y. Mostofi, "Occupancy estimation using only WiFi power measurements," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 7, pp. 1381–1393, 2015.

[59] Y. Yang, J. Cao, X. Liu, and X. Liu, "Wi-count: Passing people counting with COTS WiFi devices," in *International Conference on Computer Communications and Networks (ICCCN)*, 2018, pp. 1–9.

[60] P. E. Lopez-de-Teruel, F. J. Garcia, and O. Canovas, "Practical passive localization system based on wireless signals for fast deployment of occupancy services," *Futur. Gener. Comput. Syst.*, vol. 107, pp. 692–704, 2020.

[61] M. Youssef, M. Mah, and A. Agrawala, "Challenges: Device-free passive localization for wireless environments," in *International Conference on Mobile Computing and Networking (MobiCom)*, 2007, pp. 222–229.

[62] M. Seifeldin, A. Saeed, A. E. Kosba, A. El-Keyi, and M. Youssef, “Nuzzer: A large-scale device-free passive localization system for wireless environments,” *IEEE Trans. Mob. Comput.*, vol. 12, no. 7, pp. 1321–1334, 2013.

[63] C. Xu *et al.*, “SCPL: Indoor device-free multi-subject counting and localization using radio signal strength,” in *ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)*, 2013, pp. 79–90.

[64] Y. Yuan, J. Zhao, C. Qiu, and W. Xi, “Estimating crowd density in an RF-based dynamic environment,” *IEEE Sens. J.*, vol. 13, no. 10, pp. 3837–3845, Oct. 2013.

[65] J. Wilson and N. Patwari, “Radio tomographic imaging with wireless networks,” *IEEE Trans. Mob. Comput.*, vol. 9, no. 5, pp. 621–632, 2010.

[66] B. Abade, D. P. Abreu, and M. Curado, “A non-intrusive approach for indoor occupancy detection in smart environments,” *Sensors*, vol. 18, no. 11, p. 3953, 2018.

[67] O. Bin Tariq, M. T. Lazarescu, J. Iqbal, and L. Lavagno, “Performance of machine learning classifiers for indoor person localization with capacitive sensors,” *IEEE Access*, vol. 5, pp. 12913–12926, 2017.

[68] D. Ryan, S. Denman, S. Sridharan, and C. Fookes, “An evaluation of crowd counting methods, features and regression models,” *Comput. Vis. Image Underst.*, vol. 130, pp. 1–17, 2015.

[69] J. Zou, Q. Zhao, W. Yang, and F. Wang, “Occupancy detection in the office by analyzing surveillance videos and its application to building energy conservation,” *Energy Build.*, vol. 152, pp. 385–398, 2017.

[70] K. Jung and S. Lee, “A systematic review of RFID applications and diffusion: key areas and public policy issues,” *J. Open Innov. Technol. Mark. Complex.*, vol. 1, no. 1, p. 9, 2015.

[71] X. Zhu, S. K. Mukhopadhyay, and H. Kurata, “A review of RFID technology and its managerial applications in different industries,” *J. Eng. Technol. Manag.*, vol. 29, no. 1, pp. 152–167, 2012.

[72] I. Ehrenberg, C. Floerkemeier, and S. Sarma, “Inventory management with an RFID-equipped mobile robot,” in *IEEE International Conference on Automation Science and Engineering*, 2007, pp. 1020–1026.

[73] M. Haddara and A. Staaby, “RFID applications and adoptions in healthcare: A review on patient safety,” *Procedia Comput. Sci.*, vol. 138, pp. 80–88, 2018.

[74] S. F. Wamba, A. Anand, and L. Carter, “A literature review of RFID-enabled healthcare applications and issues,” *Int. J. Inf. Manage.*, vol. 33, no. 5, pp. 875–891, 2013.

[75] S. Palipana, B. Pietropaoli, and D. Pesch, “Recent advances in RF-based passive device-free localisation for indoor applications,” *Ad Hoc Networks*, vol. 64, pp. 80–98, 2017.

[76] B. Wagner, N. Patwari, and D. Timmermann, “Passive RFID tomographic imaging for device-free user localization,” *Work. Positioning, Navig. Commun.*, pp. 120–125, 2012.

[77] Y. Ma and E. C. Kan, “Ubiquitous tagless object locating with ambient harmonic tags,” in *IEEE International Conference on Computer Communications (INFOCOM)*, 2016.

[78] K. Konno and J. Mead, “Measurement of the separate volume changes of rib cage and abdomen during breathing,” *J. Appl. Physiol.*, vol. 22, no. 3, pp. 407–422, 1967.

[79] Ettus Research, “USRP B210 USB Software Defined Radio (SDR).” [Online]. Available: https://www.ettus.com/all-products/ub210-kit/. [Accessed: 28-Oct-2020].

[80] “OSHA safety and health topics: Radiofrequency and microwave radiation standards.” [Online]. Available: https://www.osha.gov/SLTC/radiofrequencyradiation/standards.html. [Accessed: 20-Feb-2020].

[81] Ettus Research, “USRP B200mini.” [Online]. Available: https://www.ettus.com/all-products/usrp-b200mini/. [Accessed: 28-Oct-2020].

[82] Y. Ma, X. Hui, and E. C. Kan, “Harmonic-WISP: A passive broadband harmonic RFID platform,” in *IEEE MTT-S Int. Microw. Symp. Dig.*, 2016, pp. 1–4.

[83] F. Yu, Y. Ma, and E. C. Kan, “Reflective nonlinear transmission lines for single-antenna non-self-jamming RFID,” in *IEEE MTT-S International Microwave Symposium*, 2011, pp. 1–4.

[84] Z. Zhang, P. Sharma, J. Zhou, X. Hui, and E. C. Kan, “Furniture-integrated respiration sensors by notched transmission lines,” *IEEE Sens. J.*, p. 1, 2020.

[85] W. T. McNicholas, “Diagnosis of obstructive sleep apnea in adults,” *Proc. Am. Thorac. Soc.*, vol. 5, no. 2, pp. 154–160, 2008.

[86] F. Scholkmann, J. Boss, and M. Wolf, “An efficient algorithm for automatic peak detection in noisy periodic and quasi-periodic signals,” *Algorithms*, vol. 5, no. 4, pp. 588–603, 2012.

[87] W. Lu *et al.*, “A semi-automatic method for peak and valley detection in free-breathing respiratory waveforms,” *Med. Phys.*, vol. 33, no. 10, pp. 3634–3636, 2006.

[88] B. Scholkopf, R. Williamson, A. Smola, J. Shawe-Taylor, and J. Platt, “Support vector method for novelty detection,” in *NIPS*, 1999, pp. 582–588.

[89] M. Hornyak, B. Feige, D. Riemann, and U. Voderholzer, “Periodic leg movements in sleep and periodic limb movement disorder: prevalence, clinical significance and treatment,” *Sleep Med. Rev.*, vol. 10, no. 3, pp. 169–177, 2006.

[90] H. J. E. and S. John, “Age-related changes to speech breathing with increased vocal loudness,” *J. speech, Lang. Hear. Res.*, vol. 51, no. 3, pp. 651–668, Jun. 2008.

[91] A. Maclarnon and G. Hewitt, “Increased breathing control: Another factor in the evolution of human language,” *Evol. Anthropol. Issues, News, Rev.*, vol. 13, no. 5, pp. 181–197, Jan. 2004.

[92] M. Folke, L. Cernerud, M. Ekström, and B. Hök, “Critical review of non-invasive respiratory monitoring in medical care,” *Med. Biol. Eng. Comput.*, vol. 41, no. 4, pp. 377–383, Jul. 2003.

[93] M. Á. Martínez-García *et al.*, “Continuous positive airway pressure treatment reduces mortality in patients with ischemic stroke and obstructive sleep apnea: A 5-year follow-up study,” *Am. J. Respir. Crit. Care Med.*, vol. 180, no. 1, pp. 36–41, 2009.

[94] G. Yuan, N. A. Drost, and R. A. McIvor, “Respiratory rate and breathing pattern,” *McMaster Univ. Med. J.*, vol. 10, no. 1, pp. 23–25, 2013.

[95] G. Jean-Louis, F. Zizi, L. T. Clark, C. D. Brown, and S. I. McFarlane, “Obstructive sleep apnea and cardiovascular disease: Role of the metabolic syndrome and its components,” *J. Clin. Sleep Med.*, vol. 4, no. 3, pp. 261–272, 2008.

[96] “Data acquisition and analysis system with AcqKnowledge for MP36R,” *BIOPAC Systems, Inc.* [Online]. Available: https://www.biopac.com/product/mp36r-systems/. [Accessed: 20-Feb-2020].

[97] G. Stoet, “PsyToolkit: a novel web-based method for running online questionnaires and reaction-time experiments,” *Teach. Psychol.*, vol. 44, no. 1, pp. 24–31, 2017.

[98] G. Stoet, “PsyToolkit: a software package for programming psychological experiments using Linux,” *Behav. Res. Methods*, vol. 42, no. 4, pp. 1096–1104, 2010.

[99] G. Brüllmann, R. Thurnheer, and E. Bloch, “Respiratory monitoring by inductive plethysmography in unrestrained subjects using position sensor-adjusted calibration,” *Respiration*, vol. 79, pp. 112–120, 2010.

[100] N. Ambrosino and G. Scano, “Dyspnoea and its measurement,” *Rev. Lit. Arts Am.*, vol. 1, no. 2, pp. 100–107, 2004.

[101] M. L. Campbell, T. Templin, and J. Walch, "A respiratory distress observation scale for patients unable to self-report dyspnea," *J. Palliat. Med.*, vol. 13, no. 3, pp. 285–290, Jan. 2010.

[102] T. Brosch, K. R. Scherer, D. M. Grandjean, and D. Sander, "The impact of emotion on perception, attention, memory, and decision-making," Swiss Med. Wkly., vol. 143, p. w13786, 2013.

[103] J. Driver, "A selective review of selective attention research from the past century," Br. J. Psychol., vol. 92, no. 1, pp. 53–78, 2001.

[104] M. M. Sohlberg and C. A. Mateer, Introduction to Cognitive Rehabilitation: Theory and Practice. Guilford Press, 1989.

[105] T. Brosch, G. Pourtois, D. Sander, and P. Vuilleumier, "Additive effects of emotional, endogenous, and exogenous attention: behavioral and electrophysiological evidence," Neuropsychologia, vol. 49, no. 7, pp. 1779–1787, 2011.

[106] P. Vuilleumier, "How brains beware: neural mechanisms of emotional attention," Trends Cogn. Sci., vol. 9, no. 12, pp. 585–594, 2005.

[107] Q. Luo, T. Holroyd, C. Majestic, X. Cheng, J. Schechter, and R. J. Blair, "Emotional automaticity is a matter of timing," J. Neurosci., vol. 30, no. 17, pp. 5825–5829, 2010.

[108] G. Pourtois, L. Spinelli, M. Seeck, and P. Vuilleumier, "Temporal precedence of emotion over attention modulations in the lateral amygdala: Intracranial ERP evidence from a patient with temporal lobe epilepsy," Cogn. Affect. Behav. Neurosci., vol. 10, no. 1, pp. 83–93, 2010.

[109] L. Pessoa, "Emotion and attention effects: Is it all a matter of timing? Not yet," Front. Hum. Neurosci., vol. 4, p. 172, 2010.

[110] M. Azarnoosh, A. L. I. M. Nasrabadi, M. R. Mohammadi, and M. Firoozabadi, "Evaluating nonlinear variability of mental fatigue behavioral indices during long-term attentive task," Complexity, vol. 17, no. 6, pp. 7–16, 2012.

[111] M. Basner and J. Rubinstein, "Fitness for duty: A 3 minute version of the Psychomotor Vigilance Test predicts fatigue related declines in luggage screening performance," *J. Occup. Environ. Med. Coll. Occup. Environ. Med.*, vol. 53, no. 10, p. 1146, 2011.

[112] J. Geiger-Brown, V. E. Rogers, A. M. Trinkoff, R. L. Kane, R. B. Bausell, and S. M. Scharf, "Sleep, sleepiness, fatigue, and performance of 12-hour-shift nurses," *Chronobiol. Int.*, vol. 29, no. 2, pp. 211–219, 2012.

[113] P. Ekman and W. V Friesen, “Constants across cultures in the face and emotion,” *J. Pers. Soc. Psychol.*, vol. 17, no. 2, p. 124, 1971.

[114] W. M. Wundt, *Principles of physiological psychology*, vol. 1. Sonnenschein, 1904.

[115] J. Kim and E. Andre, “Emotion recognition based on physiological changes in music listening,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 12, pp. 2067–2083, 2008.

[116] L. K. McIntire, R. A. McKinley, C. Goodyear, and J. P. McIntire, “Detection of vigilance performance using eye blinks.,” *Appl. Ergon.*, vol. 45, no. 2, pp. 354–362, Mar. 2014.

[117] U. Trutschel, B. Sirois, D. Sommer, M. Golz, and D. Edwards, “PERCLOS: An alertness measure of the past,” University of Iowa, 2011.

[118] B. Mandal, L. Li, G. S. Wang, and J. Lin, “Towards detection of bus driver fatigue based on robust visual analysis of eye state,” *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 3, pp. 545–557, 2016.

[119] W. Yimyam and M. Ketcham, “Eye region detection in fatigue monitoring for the military using AdaBoost algorithm,” in *International Symposium on Natural Language Processing*, 2016, pp. 151–161.

[120] Y.-W. Chen and K. Kubo, “A robust eye detection and tracking technique using gabor filters,” in *Third International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP 2007)*, 2007, vol. 1, pp. 109–112.

[121] Q. Ji, Z. Zhu, and P. Lan, “Real time non-intrusive monitoring and prediction of driver fatigue,” *IEEE Trans. Veh. Technol.*, vol. 53, no. 4, pp. 1052–1068, 2004.

[122] Q. Ji, P. Lan, and C. Looney, “A probabilistic framework for modeling and real-time monitoring human fatigue,” *IEEE Trans. Syst. Man, Cybern. Part A Syst. Humans*, vol. 36, no. 5, pp. 862–875, 2006.

[123] C. Lü, S. Wu, B. Lu, Y. Zhang, Y. Du, and X. Feng, “Ultrathin flexible piezoelectric sensors for monitoring eye fatigue,” *J. Micromechanics Microengineering*, vol. 28, no. 2, p. 25010, Jan. 2018.

[124] T. C. Chieh, M. M. Mustafa, A. Hussain, S. F. Hendi, and B. Y. Majlis, “Development of vehicle driver drowsiness detection system using electrooculogram (EOG),” in *International Conference on Computers, Communications, & Signal Processing with Special Track on Biomedical Engineering*, 2005, pp. 165–168.

[125] A. Aryal, A. Ghahramani, and B. Becerik-Gerber, “Monitoring fatigue in construction workers using physiological measurements,” *Autom. Constr.*, vol. 82, pp. 154–165, 2017.

[126] K. Q. Shen, X. P. Li, C. J. Ong, S. Y. Shao, and E. P. V. Wilder-Smith, “EEG-based mental fatigue measurement using multi-class support vector machines with confidence estimate,” *Clin. Neurophysiol.*, vol. 119, no. 7, pp. 1524–1533, 2008.

[127] K.-Q. Shen, C.-J. Ong, X.-P. Li, Z. Hui, and E. P. V Wilder-Smith, “A feature selection method for multilevel mental fatigue EEG classification,” *IEEE Trans. Biomed. Eng.*, vol. 54, no. 7, pp. 1231–1237, 2007.

[128] G. Borghini *et al.*, “Assessment of mental fatigue during car driving by using high resolution EEG activity and neurophysiologic indices,” in *IEEE Engineering in Medicine and Biology Society*, 2012, pp. 6442–6445.

[129] X. Zhou *et al.*, “Vigilance detection method for high-speed rail using wireless wearable EEG collection technology based on low-rank matrix decomposition,” *IET Intell. Transp. Syst.*, vol. 12, no. 8, pp. 819–825, 2018.

[130] L. J. Trejo, K. Kubitz, R. Rosipal, R. L. Kochavi, and L. D. Montgomery, “EEG-based estimation and classification of mental fatigue,” *Psychology*, vol. 6, no. 05, p. 572, 2015.

[131] Y. Shimomura, T. Yoda, K. Sugiura, A. Horiguchi, K. Iwanaga, and T. Katsuura, “Use of frequency domain analysis of skin conductance for evaluation of mental workload,” *J. Physiol. Anthropol.*, vol. 27, no. 4, pp. 173–177, 2008.

[132] M. Egger, M. Ley, and S. Hanke, “Emotion Recognition from Physiological Signal Analysis: A Review,” *Electron. Notes Theor. Comput. Sci.*, vol. 343, pp. 35–55, 2019.

[133] A. Dzedzickis, A. Kaklauskas, and V. Bucinskas, “Human emotion recognition: Review of sensors and methods,” *Sensors*, vol. 20, no. 3, 2020.

[134] L. Shu *et al.*, “A review of emotion recognition using physiological signals,” *Sensors*, vol. 18, no. 7, p. 2074, 2018.

[135] C. Busso *et al.*, “Analysis of emotion recognition using facial expressions, speech and multimodal information,” in *6th International Conference on Multimodal Interfaces*, 2004, pp. 205–211.

[136] G. Castellano, L. Kessous, and G. Caridakis, “Emotion recognition through multiple modalities: face, body gesture, speech,” in *Affect and Emotion in Human-Computer Interaction: From Theory to Applications*, C. Peter and R. Beale, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008, pp. 92–103.

[137] I. Cohen, A. Garg, and T. S. Huang, “Emotion recognition from facial expressions using multilevel HMM,” *Neural Inf. Process. Syst.*, vol. 2, 2000.

[138] B. Schuller, G. Rigoll, and M. Lang, “Speech emotion recognition combining acoustic features and linguistic information in a hybrid support vector machine-belief network architecture,” in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, 2004, vol. 1, pp. I–577.

[139] L. Chen, W. Su, Y. Feng, M. Wu, J. She, and K. Hirota, “Two-layer fuzzy multiple random forest for speech emotion recognition in human-robot interaction,” *Inf. Sci. (Ny).*, vol. 509, pp. 150–163, 2020.

[140] D. Cevher, S. Zepf, and R. Klinger, “Towards multimodal emotion recognition in German speech events in cars using transfer learning,” *arXiv Prepr. arXiv1909.02764*, 2019.

[141] J. Zhai and A. Barreto, “Stress detection in computer users based on digital signal processing of noninvasive physiological variables,” in *IEEE Engineering in Medicine and Biology*, 2006, pp. 1355–1358.

[142] A. Nakasone, H. Prendinger, and M. Ishizuka, “Emotion recognition from electromyography and skin conductance,” in *International workshop on biosignal interpretation*, 2005, pp. 219–222.

[143] G. Udovičić, J. Derek, M. Russo, and M. Sikora, “Wearable emotion recognition system based on GSR and PPG signals,” in *International Workshop on Multimedia for Personal Health and Health Care*, 2017, pp. 53–59.

[144] M. M. Bradley, L. Miccoli, M. A. Escrig, and P. J. Lang, “The pupil as a measure of emotional arousal and autonomic activation,” *Psychophysiology*, vol. 45, no. 4, pp. 602–607, 2008.

[145] E. Gatti, E. Calzolari, E. Maggioni, and M. Obrist, “Emotional ratings and skin conductance response to visual, auditory and haptic stimuli,” *Sci. Data*, vol. 5, no. 1, p. 180120, 2018.

[146] J. Zhu, L. Ji, and C. Liu, “Heart rate variability monitoring for emotion and disorders of emotion,” *Physiol. Meas.*, vol. 40, no. 6, p. 64004, Jul. 2019.

[147] H. G. Kim, E. J. Cheon, D. S. Bai, Y. H. Lee, and B. H. Koo, “Stress and heart rate variability: A meta-analysis and review of the literature,” *Psychiatry Investig.*, vol. 15, no. 3, pp. 235–245, 2018.

[148] F. Agrafioti, D. Hatzinakos, and A. K. Anderson, “ECG pattern analysis for emotion detection,” *IEEE Trans. Affect. Comput.*, vol. 3, no. 1, pp. 102–115, 2012.

[149] E.-. Jang, B.-. Park, S.-. Kim, Y. Eum, and J.-. Sohn, “Identification of the optimal emotion recognition algorithm using physiological signals,” in *International Conference on Engineering and Industries (ICEI)*, 2011, pp. 1–6.

[150] R. Harper and J. Southern, “End-to-end prediction of emotion from heartbeat data collected by a consumer fitness tracker,” in *International Conference on Affective Computing and Intelligent Interaction (ACII)*, 2019, pp. 1–7.

[151] T. Bhowmik, J. Dey, and V. N. Tiwari, “A novel method for accurate estimation of HRV from smartwatch PPG signals,” in *IEEE Engineering in Medicine and Biology Society (EMBC)*, 2017, pp. 109–112.

[152] M. O. Diab, A. Seif, M. Sabbah, M. El-Abed, and N. Aloulou, “A review on ECG-based biometric authentication systems,” in *Hidden Biometrics: When Biometric Security Meets Biomedical Engineering*, Singapore: Springer, 2020, pp. 17–44.

[153] Q. Zhang, X. Chen, Q. Zhan, T. Yang, and S. Xia, “Respiration-based emotion recognition with deep learning,” *Comput. Ind.*, vol. 92–93, pp. 84–90, 2017.

[154] S. Oh, J. Y. Lee, and D. K. Kim, “The design of CNN architectures for optimal six basic emotion classification using multiple physiological signals,” *Sensors*, vol. 20, no. 3, pp. 1–17, 2020.

[155] J. R. Machado Fernández and L. Anishchenko, “Mental stress detection using bioradar respiratory signals,” *Biomed. Signal Process. Control*, vol. 43, pp. 244–249, 2018.

[156] A. N. Khan, A. A. Ihalage, Y. Ma, B. Liu, Y. Liu, and Y. Hao, “Deep learning framework for subject-independent emotion detection using wireless signals,” pp. 1–13, 2020.

[157] M. Zhao, F. Adib, and D. Katabi, “Emotion recognition using wireless signals,” *Commun. ACM*, vol. 61, no. 9, pp. 91–100, 2018.

[158] U. R. Acharya, K. P. Joseph, N. Kannathal, C. M. Lim, and J. S. Suri, “Heart rate variability: A review,” *Med. Biol. Eng. Comput.*, vol. 44, no. 12, pp. 1031–1051, 2006.

[159] W. von Rosenberg, T. Chanwimalueang, T. Adjei, U. Jaffer, V. Goverdovsky, and D. P. Mandic, “Resolving ambiguities in the LF/HF ratio: LF-HF scatter plots for the categorization of mental and physical stress from HRV,” *Front. Physiol.*, vol. 8, no. JUN, pp. 1–12, 2017.

[160] T. Adjei, W. Von Rosenberg, T. Nakamura, T. Chanwimalueang, and D. P. Mandic, “The ClassA framework: HRV based assessment of SNS and PNS dynamics without LF-HF controversies,” *Front. Physiol.*, vol. 10, no. APR, pp. 1–15, 2019.

[161] G. Strauss-Blasche, M. Moser, M. Voica, D. R. McLeod, N. Klammer, and W. Marktl, "Relative timing of inspiration and expiration affects respiratory sinus arrhythmia," *Clin. Exp. Pharmacol. Physiol.*, vol. 27, no. 8, pp. 601–606, 2000.

[162] R. Jerath and C. Beveridge, "Respiratory rhythm, autonomic modulation, and the spectrum of emotions: The future of emotion recognition and modulation," *Front. Psychol.*, vol. 11, 2020.

[163] P. Philippot, G. Chapelle, and S. Blairy, "Respiratory feedback in the generation of emotion," *Cogn. Emot.*, vol. 16, no. 5, pp. 605–627, 2002.

[164] K. L. Lichstein, B. W. Riedel, and S. L. Richman, "The Mackworth clock test: a computerized version.," *J. Psychol.*, vol. 134, no. 2, pp. 153–161, Mar. 2000.

[165] C. Ding and H. Peng, "Minimum redundancy feature selection from microarray gene expression data," *J. Bioinform. Comput. Biol.*, vol. 3, no. 2, pp. 185–205, 2005.

[166] L. Pecchia, R. Castaldo, L. Montesinos, and P. Melillo, "Are ultra-short heart rate variability features good surrogates of short-term ones? State-of-the-art review and recommendations," *Healthc. Technol. Lett.*, vol. 5, no. 3, pp. 94–100, 2018.

[167] M. R. Esco and A. A. Flatt, "Ultra-short-term heart rate variability indexes at rest and post-exercise in athletes: evaluating the agreement with accepted recommendations," *J. Sports Sci. Med.*, vol. 13, no. 3, p. 535—541, Sep. 2014.

[168] L. Salahuddin, J. Cho, M. G. Jeong, and D. Kim, "Ultra Short Term Analysis of Heart Rate Variability for Monitoring Mental Stress in Mobile Settings," in *IEEE Engineering in Medicine and Biology*, 2007, pp. 4656–4659.

[169] J. M. Yentes, N. Hunt, K. K. Schmid, J. P. Kaipust, D. McGrath, and N. Stergiou, "The appropriate use of approximate entropy and sample entropy with short data sets.," *Ann. Biomed. Eng.*, vol. 41, no. 2, pp. 349–365, Feb. 2013.

[170] R. W. Picard, E. Vyzas, and J. Healey, "Toward machine emotional intelligence: Analysis of affective physiological state," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 10, pp. 1175–1191, 2001.

[171] C. Li, C. Xu, and Z. Feng, "Analysis of physiological for emotion recognition with the IRS model," *Neurocomputing*, vol. 178, pp. 103–111, 2016.

[172] H. Obeidat, A. Alabdullah, E. Elkhazmi, W. Suhaib, O. Obeidat, and M. Alkhambashi, "Indoor environment propagation review," *Comput. Sci. Rev.*, vol. 37, p. 100272, 2020.

[173] A. Goldsmith, *Wireless Communications*. New York, NY, USA: Cambridge University Press, 2005.

[174] X. Hui and E. C. Kan, “Collaborative reader code division multiple access in the harmonic RFID system,” *IEEE J. Radio Freq. Identif.*, vol. 2, no. 2, pp. 86–92, Jun. 2018.

[175] S. Xu, H. Liu, F. Gao, and Z. Wang, “Compressive sensing based radio tomographic imaging with spatial diversity,” *Sensors*, vol. 19, no. 3, 2019.

[176] L. di Stefano and A. Bulgarelli, “A simple an efficient connected components laboring alogothm,” in *10th Intl. Conf. Image Anal. Proc.*, 1999, pp. 322–327.

[177] X. Meng, M. A. Saunders, and M. W. Mahoney, “Lsrn: A parallel iterative solver for strongly over- or underdetermined systems,” *SIAM J. Sci. Comput.*, vol. 36, no. 2, 2014.

[178] A. Beck and M. Teboulle, “A fast iterative shrinkage-thresholding algorithm for linear inverse problem,” *SIAM J. Imaging Sci.*, vol. 2, no. 1, pp. 183–202, 2009.

[179] S. G. Mallat and Zhifeng Zhang, “Matching pursuits with time-frequency dictionaries,” *IEEE Trans. Signal Process.*, vol. 41, no. 12, pp. 3397–3415, 1993.

[180] R. Fan, Q. Wan, Y. Liu, H. Chen, and X. Zhang, “Complex orthogonal matching pursuit and its exact recovery conditions,” *arXiv Prepr. arXiv1206.2197.*, 2012.

[181] S. Kallummil and S. Kalyani, “Tuning free orthogonal matching pursuit,” *arXiv Prepr. arXiv1703.05080*, 2017.

[182] S. Kallummil and S. Kalyani, “Signal and noise statistics oblivious sparse reconstruction using omp/ols,” *arXiv Prepr. arXiv1707.08712*, 2017.

[183] M. Sofos *et al.*, “Innovations in sensors and controls for building energy management: Research and development opportunities report for emerging technologies,” 2020.

[184] J. Ahmad, H. Larijani, R. Emmanuel, M. Mannion, and A. Javed, “Occupancy detection in non-residential buildings – A survey and novel privacy preserved occupancy monitoring solution,” *Appl. Comput. Informatics*, no. xxxx, pp. 1–9, 2018.

[185] W. Rawat and Z. Wang, “Deep convolutional neural networks for image classification: A comprehensive review,” *Neural Comput.*, vol. 29, no. 9, pp. 2352–2449, 2017.

[186] W. Zhiqiang and L. Jun, “A review of object detection based on convolutional neural network,” in *Chinese Control Conference (CCC)*, 2017, pp. 11104–11109.

[187] L. A. Gatys, A. S. Ecker, and M. Bethge, “Image style transfer using convolutional neural networks,” in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 2414–2423.

[188] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems (NIPS)*, 2012, pp. 1097–1105.

[189] V. Nair and G. E. Hinton, "Rectified linear units improve restricted boltzmann machines," in *International Conference on Machine Learning (ICML)*, 2010, pp. 807–814.

[190] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," *J. Mach. Learn. Res.*, vol. 15, no. 56, pp. 1929–1958, Jan. 2014.

[191] B. Zhou, A. Khosla, A. Lapedriza, A. Oliva, and A. Torralba, "Learning deep features for discriminative localization," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 2921–2929.

[192] I. Loshchilov and F. Hutter, "Decoupled weight decay regularization," in *International Conference on Learning Representations*, 2019.

[193] EPCglobal, "EPC TM radio-frequency identity protocols generation-2 UHF RFID," 2013. [Online]. Available: https://www.gs1.org/sites/default/files/docs/epc/uhfc1g2_2_0_0_standard_20131101.pdf.

[194] D. Torrieri, *Principles of spread-spectrum communication systems.* Springer, 2005.

[195] H. Hashemi, "The indoor radio propagation channel," *Proc. IEEE*, vol. 81, no. 7, pp. 943–968, Jul. 1993.

[196] W. Ruan, L. Yao, Q. Z. Sheng, N. J. G. Falkner, and X. Li, "TagTrack: Device-free localization and tracking using passive RFID tags," *MobiQuitous 2014 - 11th Int. Conf. Mob. Ubiquitous Syst. Comput. Netw. Serv.*, pp. 80–89, 2014.

[197] A. Creswell, T. White, V. Dumoulin, K. Arulkumaran, B. Sengupta, and A. A. Bharath, "Generative adversarial networks: An overview," *IEEE Signal Process. Mag.*, vol. 35, no. 1, pp. 53–65, Jan. 2018.

[198] Y. Duroc and S. Tedjini, "RFID: A key technology for humanity," *Comptes Rendus Phys.*, vol. 19, no. 1–2, pp. 64–71, 2018.

[199] V. Stanford, "Pervasive computing goes the last hundred feet with RFID systems," *IEEE Pervasive Comput.*, vol. 2, no. 2, pp. 9–14, Apr. 2003.

[200] V. Mulloni and M. Donelli, "Chipless RFID sensors for the Internet of Things: Challenges and opportunities," *Sensors*, vol. 20, no. 7, p. 2135, 2020.

[201] J. Song, C. T. Haas, and C. H. Caldas, “A proximity-based method for locating RFID tagged objects,” *Adv. Eng. Informatics*, vol. 21, no. 4, pp. 367–376, 2007.

[202] Y. Ma, X. Hui, and E. C. Kan, “3D real-time indoor localization via broadband nonlinear backscatter in passive devices with centimeter precision,” in *International Conference on Mobile Computing and Networking (MobiCom)*, 2016, pp. 216–229.

[203] F. Xiao, Z. Wang, N. Ye, R. Wang, and X. Li, “One more tag enables fine-grained RFID localization and tracking,” *IEEE/ACM Trans. Netw.*, vol. 26, no. 1, pp. 161–174, Feb. 2018.

[204] P. Sharma, G. Xu, X. Hui, and E. C. Kan, “Deep-learning based occupant counting by ambient RF sensing, ” *IEEE Sensors J.,* 2020. (Under Review)

[205] D. L. Hysell, P. Sharma, M. Urco, and M. A. Milla, “Aperture-synthesis radar imaging with compressive sensing for ionospheric research, ”*Radio Science*, vol. 54, pp. 503–516, 2019.

[206] K. Sun, Q. Zhao, and J. Zou, “A review of building occupancy measurement systems,” Energy Build., vol. 216, p. 109965, 2020.

[207] V. L. Erickson, S. Achleitner, and A. E. Cerpa, “POEM: Power-efficient occupancy-based energy management system,” in International Conference on Information Processing in Sensor Networks, 2013, pp. 203–216.

[208] N. Cao, J. Ting, S. Sen, and A. Raychowdhury, “Smart sensing for HVAC control: collaborative intelligence in optical and IR cameras,” *IEEE Trans. Ind. Electron*., vol. 65, no. 12, pp. 9785–9794, 2018.

[209] M. Eldib et al., “A low resolution multi-camera system for person tracking,” in *IEEE International Conference on Image Processing (ICIP),* 2014, pp. 378–382.

[210] M. Upmanyu, A. M. Namboodiri, K. Srinathan, and C. V. Jawahar, “Efficient privacy preserving video surveillance,” in *IEEE International Conference on Computer Vision*, 2009, pp. 1639–1646.

[211] T. B. Moeslund and E. Granum, “A survey of computer vision-based human motion capture,” *Comput. Vis. Image Underst*., vol. 81, no. 3, pp. 231–268, Mar. 2001.

[212] C. Raghavachari, V. Aparna, S. Chithira, and V. Balasubramanian, “A comparative study of vision based human detection techniques in people counting applications,” *Procedia Comput. Sci.,* vol. 58, pp. 461–469, 2015.

[213] T. Teixeira and A. Savvides, “Lightweight people counting and localizing for easily deployable indoors WSNs,” *IEEE J. Sel. Top. Signal Process.*, vol. 2, no. 4, pp. 493–502, 2008.