

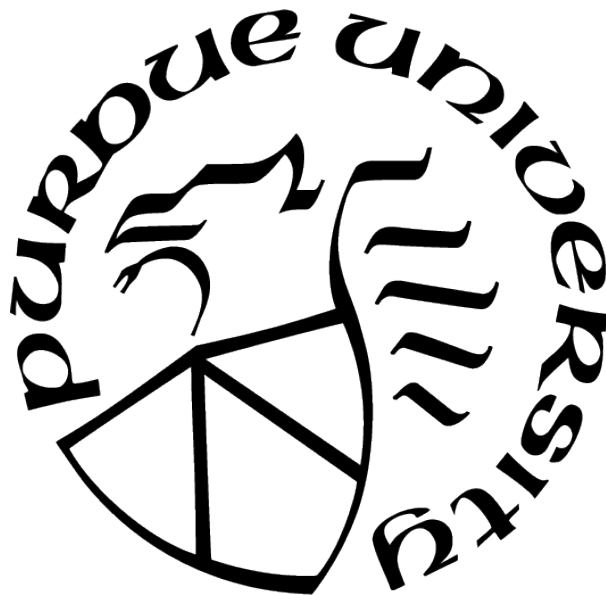
# VEHICLE OCCUPANT MONITORING WITH MMWAVE WIDEBAND PLANAR ARRAY RADAR

by  
Qiwen Deng

A Dissertation

*Submitted to the Faculty of Purdue University  
In Partial Fulfillment of the Requirements for the degree of*

Doctor of Philosophy



School of Electrical and Computer Engineering

West Lafayette, Indiana

May 2021

**THE PURDUE UNIVERSITY GRADUATE SCHOOL  
STATEMENT OF COMMITTEE APPROVAL**

**Dr. Yaobin Chen, Co-Chair**

School of Electrical and Computer Engineering

**Dr. Jianghai Hu, Co-Chair**

School of Electrical and Computer Engineering

**Dr. Lingxi Li**

School of Electrical and Computer Engineering

**Dr. Shreyas Sundaram**

School of Electrical and Computer Engineering

**Approved by:**

Dr. Dimitri Peroulis

To my beloved Ruchao and my parents for their love and support

## ACKNOWLEDGMENTS

Firstly, I would like to express my sincere gratitude to my advisors Dr. Yaobin Chen and Dr. Jianghai Hu, for the continuous support of my Ph.D. study and related research, for their patience, motivation, and immense knowledge. Thanks to Dr. Chen for his guidance to help me in all the research and writing of this dissertation. I could not have imagined having a better advisor and mentor for my Ph.D. study. Besides my advisors, I would like to thank the rest of my committee: Dr. Lingxi Li, and Dr. Shreyas Sundaram, for their insightful comments and encouragement for this research.

Secondly, my dissertation research is a subset of the research project sponsored by Ford Motor Company. Therefore, I had an opportunity to work at Ford Research and Innovation Center when the project kicked-off. My sincere and special thanks also go to Mr. Jialiang Le, a wonderful mentor and technical advisor at Ford Research and Advanced Engineering who directed my research and helped me accumulate academic and industrial experiences. Without his precious advice, it would not be possible to complete this research. Thanks for his inspiring and challenging questions, which helped me widen my research from various perspectives.

I am grateful for the help I obtained through my Ph.D. study. Dr. Renran Tian at Transportation and Autonomous Systems Institute (TASI) deserves special thanks for enlightening me on the first glance of research and also recommended and supported me to join this research project. Besides, I would like to thank other professors at TASI who provide their help and support during my Ph.D. study, such as Dr. Stanley Chien. I have been fortunate to have my friends and lab mates Wensen Niu, Dan Shen, and Keyu Ruan, for their help and support. I have been fortunate, too, to have my friend and colleague, Jun Lin, for the constructive discussions in my dissertation research and other projects we were working on together at Ford.

Last, but certainly not least, are my parents, who supported me unconditionally and encouraged me throughout the writing of this dissertation and to my life in general. Another precious support is from my girlfriend Ruchao Shan, who always provides me the courage to challenge myself for a better life. I could not have achieved this without her support.



# TABLE OF CONTENTS

LIST OF TABLES . . . . .	10
LIST OF FIGURES . . . . .	11
ABBREVIATIONS . . . . .	14
ABSTRACT . . . . .	16
1 INTRODUCTION . . . . .	17
1.1 Importance of vehicle in-cabin sensing . . . . .	17
1.2 Related works . . . . .	18
1.3 Research objectives and structure . . . . .	23
2 MATHEMATIC MODEL FOR HUMAN SUBJECT DETECTION . . . . .	24
2.1 Frequency domain model . . . . .	24
2.1.1 Mathematical model . . . . .	24
2.1.2 Derivation . . . . .	25
2.2 Time-domain model . . . . .	29
2.2.1 Mathematical model . . . . .	29
2.2.2 Derivation . . . . .	33
2.3 SINR enhancement . . . . .	37
3 BASIS OF VEHICLE IN-CABIN RADAR SYSTEM DESIGN . . . . .	40
3.1 Design constraint and requirement . . . . .	40
3.1.1 Vehicle size and layout . . . . .	40

3.1.2	Safety requirements . . . . .	41
3.2	Radar parameter design . . . . .	41
3.2.1	Central frequency . . . . .	41
3.2.2	Attenuation and penetration . . . . .	42
3.2.3	Waveform transmission . . . . .	45
3.2.4	Range resolution . . . . .	47
3.2.5	Antenna size . . . . .	50
3.2.6	Patch antenna . . . . .	50
3.2.7	Single pair antenna . . . . .	52
3.2.8	Antenna array . . . . .	55
3.2.9	Antenna subarray . . . . .	56
3.3	The Angle of Arrival (AoA) approaches . . . . .	61
3.3.1	Angle of Arrival (AoA): Beamforming . . . . .	61
3.3.2	Angle of Arrival (AoA): Subspace-based approach - MUSIC . . . . .	63
3.3.3	Angle of Arrival (AoA): Adaptive beamforming - MVDR . . . . .	66
3.4	Radiation regulations and compliances . . . . .	66
3.4.1	Radiation regulations . . . . .	66
3.4.2	Radiation compliance . . . . .	68
4	VEHICLE IN-CABIN RADAR SYSTEM SIMULATION . . . . .	71
4.1	Simulation platform . . . . .	71

4.1.1	System design . . . . .	71
4.1.2	Parameter selection for waveform . . . . .	72
4.1.3	Antenna array design . . . . .	74
4.1.4	Planer antenna array configuration . . . . .	75
4.1.5	Steering matrix . . . . .	76
4.2	Signal and image processing . . . . .	77
4.2.1	TDM-based approach and algorithm . . . . .	77
4.2.2	Vital sign estimation . . . . .	78
4.2.3	AoA estimation and heatmap generation . . . . .	79
4.2.4	CFAR thresholds . . . . .	81
4.2.5	Coordinate transformation . . . . .	82
4.2.6	Point cloud approach . . . . .	83
4.2.7	Processing flow . . . . .	84
4.3	System validation with human subject . . . . .	86
4.3.1	Design of the reflection-based human subject model . . . . .	86
4.3.2	Experiments . . . . .	87
4.3.3	Reflection-based human subject model . . . . .	89
5	HUMAN SUBJECT CLASSIFICATION . . . . .	92
5.1	Overview of human subject classification . . . . .	92
5.2	Imaging dataset preparation for classification . . . . .	93

5.2.1	Vision-based in-cabin synthetic imaging dataset . . . . .	93
5.2.2	Camera-radar alignment . . . . .	94
5.2.3	Radar-based in-cabin imaging dataset generation . . . . .	96
5.2.4	Point cloud 2-D projection . . . . .	98
5.3	Deep learning framework . . . . .	99
5.3.1	Network structure . . . . .	99
5.3.2	Loss function . . . . .	100
5.3.3	Training . . . . .	100
5.3.4	Testing . . . . .	102
5.4	Results and analysis . . . . .	104
5.4.1	Evaluation metrics . . . . .	104
5.4.2	Accuracy . . . . .	104
5.4.3	Failure case and discussion . . . . .	107
6	HUMAN SUBJECT POSTURE DETECTION . . . . .	109
6.1	Introduction to posture detection . . . . .	109
6.2	Keypoint-based model and imaging method . . . . .	109
6.2.1	Design of the Keypoint-based human subject model . . . . .	109
6.2.2	Imaging method for posture detection . . . . .	111
6.2.3	Design of vital sign signal . . . . .	112
6.3	Simulation results . . . . .	113

6.3.1	Planar radar array in different sizes . . . . .	113
6.3.2	Keypoint skeleton models and their postures . . . . .	115
6.3.3	Accuracy and analysis . . . . .	120
7	CONCLUSION AND FUTURE WORK . . . . .	124
7.1	Conclusion . . . . .	124
7.2	Future work . . . . .	125
	REFERENCES . . . . .	126
	VITA . . . . .	137

## LIST OF TABLES

2.1	Selected relative permittivity for adults and children . . . . .	32
3.1	The typical size of Ford vehicle . . . . .	40
3.2	Dielectric constant and loss tangent of thorax related tissue under 7.29 GHz and 79 GHz radars . . . . .	44
3.3	Antenna size and minimum antenna distance . . . . .	50
3.4	RF antenna specs of AWR1843BOOST . . . . .	52
3.5	Basic parameters for a simulation system . . . . .	69
4.1	Parameters for chirp design . . . . .	73
5.1	Intrinsic parameters . . . . .	95
5.2	Selected Keypoints for human model . . . . .	97
5.3	Accuracy comparison between radar and camera based occupant classification .	107
6.1	Detection and MSE in polar coordinate with different size of antenna array . . .	115
6.2	Performance summary of the simulation tests in Cartesian coordinate . . . . .	122

## LIST OF FIGURES

1.1	The vehicle occupant monitoring with the wideband antenna array . . . . .	17
3.1	One-way attenuation of radar wave under difference frequency [90] . . . . .	43
3.2	Section of thorax of a male [91] . . . . .	44
3.3	UWB radar power spectrum density distribution [92] . . . . .	46
3.4	Impulse radio waveform . . . . .	47
3.5	FMCW chirp waveform . . . . .	48
3.6	Micro-strip antenna configuration . . . . .	51
3.7	System setup for two subjects . . . . .	53
3.8	Weak signal detection method flowchart . . . . .	53
3.9	Signal amplification for system setup . . . . .	54
3.10	Real aperture radar antenna array . . . . .	56
3.11	Subarray based occupant monitoring method . . . . .	57
3.12	Partition of back row in vehicle cabin . . . . .	58
3.13	Subarray based two-step scan . . . . .	58
3.14	(a) reduction of beamforming number with different value of repetition interval (b) number of beamforming for different number of occupants and size of subarray . . . . .	61
4.1	The simulation system . . . . .	71
4.2	Signal power level change by propagation stage . . . . .	73
4.3	The spectrum of designed FMCW chirp . . . . .	74
4.4	T-shape planar antenna array (a) actual antenna array with normal direction and (b) virtual planar antenna array . . . . .	75
4.5	Demonstration of virtual planar antenna array . . . . .	76
4.6	Distance between each antenna pair . . . . .	77
4.7	Vital sign detection method . . . . .	79
4.8	Comparison of the signal (a) before and (b) after the unwrapped algorithm applied . . . . .	79
4.9	Point spread function . . . . .	80
4.10	A heatmap measurement of a standing human in real world . . . . .	81

4.11	2D CA-CFAR threshold method . . . . .	82
4.12	Range and cross-range resolution in Cartesian coordinate . . . . .	83
4.13	Point cloud plot of a standing human . . . . .	84
4.14	Imaging and vital sign detection pipeline . . . . .	85
4.15	Multi-path effect inside vehicle cabin . . . . .	86
4.16	Procedure to find the generalized reflection-based model for human subject .	87
4.17	Tests of real human subject and corresponding design of reflection-based human subject model in simulation . . . . .	88
4.18	Tests of real human subject and corresponding design of reflection-based human subject model in simulation with offset . . . . .	90
4.19	Reflection-based human subject simulation model . . . . .	91
5.1	Overview of the deep-learning-based method for human subject classification	93
5.2	Comparison of real world data with SVIRO [114] . . . . .	94
5.3	Generated ground truth of camera data . . . . .	94
5.4	(a) RCS pattern of occupant model and (b) point cloud of occupants . . . .	98
5.5	Point cloud projection overlayed with segmentation ground truth . . . . .	98
5.6	Revised Unet for occupant classification [75] . . . . .	99
5.7	Layout of (a) Ford Escape and (b) Tesla Model 3 . . . . .	101
5.8	Statistics of occupant inside Ford Escape by (a) occupant category and (b) seat position and occupant inside Tesla Model 3 by (c) occupant category and (d) seat position . . . . .	102
5.9	Layout of Hyundai Tucson . . . . .	103
5.10	Statistics of occupants inside Hyundai Tuscon by (a) occupant category and (b) seat position . . . . .	103
5.11	(a) mIoU and (b) cross entropy loss of training . . . . .	104
5.12	Output mask of data from the training dataset . . . . .	105
5.13	Pre-processed data and results from the testing data set . . . . .	106
5.14	(a) extreme posture and (b) lack of radar point cloud points . . . . .	107
6.1	Keypoint model and ratio of a standing human . . . . .	110
6.2	Keypoint adult models in 3D space (a) standing (b) sitting . . . . .	111
6.3	Comparison of (a) imaging with single range bin and (b) imaging with improved algorithm . . . . .	111



6.4	The designed displacements of (a) breath signal and (b) heart signal . . . . .	113
6.5	The imaging of a standing human with different numbers of radar antenna (a) $4 \times 4$ (b) $8 \times 8$ (c) $16 \times 16$ and (d) $24 \times 24$ . . . . .	114
6.6	Standing adult posture: (a) Keypoint imaging (b) Keypoint estimation . . . .	116
6.7	(a) breath rate and (b) heart rate detected for a standing adult . . . . .	116
6.8	Sitting posture of an adult: (a) Keypoint imaging (b) Keypoint estimation .	117
6.9	Sitting adult Keypoints detected with DBSCAN . . . . .	118
6.10	(a) breath rate and (b) heart rate detected for the sitting adult . . . . .	118
6.11	Sitting child Keypoints detected with DBSCAN . . . . .	119
6.12	Unfiltered (a) breath signal and (b) heart signal of a sitting child . . . . .	120
6.13	(a) breath rate and (b) heart rate for the sitting child . . . . .	120
6.14	2D error of the standing posture case . . . . .	121
6.15	3D error of (a) the sitting adult and (b) the sitting child . . . . .	122

## ABBREVIATIONS

CW	continuous wave
IR	impulse radio
FMCW	frequency modulated continuous wave
SFCW	stepped frequency continuous wave
FFT	fast fourier transform
AoA	angle of arrival
MUSIC	multiple signal classification
VMD	variational mode decomposition
PCA	principle component analysis
ICA	independent component analysis
RF	radio frequency
RCS	radar cross section
CTFT	continuous time fourier transform
TWB	total body water
ROI	range of interest
WSS	wide sense stationary
SINR	signal to interference plus noise ratio
ADAS	advanced driver assistance systems
IF	instantaneous frequency
MIMO	multiple input multiple outputs
SIMO	single input multiple output
SISO	single input single output
MVDR	minimum variance distortionless response
EIRP	effective (or equivalent) isotropic radiated power
SAR	specific absorption rate
RMS	root mean square
TDM	time division modulation
SUT	subject under test

ADC	analog-to-digital
CNN	convolutional neural network
mIoU	mean Intersection over Union
MSE	mean square error
DBSCAN	density-based spatial clustering of applications with noise
OTS	off-the-shelf
CFAR	constant false alarm detection

## ABSTRACT

This dissertation proposes a vehicle occupant monitoring method using a mmWave wide-band planar radar array to monitor multiple occupants' status. The radar array provides high range resolution with a wide field of view in both azimuth and elevation domain, making multiple occupant detection possible. Several methods are developed for posture detection, vital sign estimation, and classification of multiple occupants inside the vehicle cabin. Firstly, a mathematical model is proposed to describe the occupant reflection in the radio frequency environment. A signal processing pipeline is proposed based on the mathematical model. Next, a simulation framework is developed for the occupant's posture detection and vital signs estimation. A reflection-based model is created to include both the size of each part of the human body and its reflection pattern for various sizes of the occupant simulation. A deep-learning-based method is then proposed based on the radar images reflected from the model for the occupant classification. This method utilizes the image information from the aligned camera as supervision to translate the radar point cloud to semantic segmentation masks. The designed network uses a sparse projected radar point cloud in 2D to generate occupants' segmentation masks in different categories. The overall prediction accuracy of the designed method is acceptable and compatible with the accuracy of the camera-based image segmentation using the same network. For posture detection, a Keypoint-based model is proposed containing both posture and vital signs. Various sizes of the planer antennas are investigated. The optimal size of the antenna is selected to evaluate various human subjects in the simulation system. The results of the detection capability and accuracy are sufficient to distinguish the given sizes of the occupants as well as their liveness estimation in the simulation.

# 1. INTRODUCTION

## 1.1 Importance of vehicle in-cabin sensing

Vehicle in-cabin sensing is becoming a crucial technology and hot research topic for both vehicle safety and comfort of conventional and future intelligent vehicles. The future vehicle will have flexible interior configurations, which is very challenging to the sensing technology development for occupancy monitoring and classification. Various state-of-the-art sensing technologies are being proposed and developed for the challenge. Compared to other sensors, radar has two additional advantages - privacy and subject detection through non-metallic materials. A typical overhead setup of a radar system for in-cabin occupant monitoring is shown in Fig. 1.1 for the occupant detection and classification. Besides safety applications, these advantages can be used for a better ride experience by continuously monitoring their state change by measuring both the posture and vital signs. Air conditioner control, gesture control for human-machine interface, and health monitoring could be possible when a more accurate radar-based occupant sensing technology is available.



**Figure 1.1.** The vehicle occupant monitoring with the wideband antenna array

On the other hand, Euro NCAP is discussing and proposing child-presence detection requirements for helping save child life due to heatstroke [1]. According to statistics, 792 children have died from vehicular heatstroke since 1998 in the U.S. and more than half of the death was caused by being left in vehicles unattended [2]. Detection of a child left behind in vehicle is difficult for most of the state-of-the-art sensing technologies, especially when using a camera-based system. The camera-based system fails when an occlusion occurs, as some small children may be covered by blankets or infants be blocked by the canopy of the rearward-facing infant seat. A better approach is to design a system that is capable of detecting a certain level of the vital-sign, e.g., live signature. However, the live signature of a child, especially an infant, is a subtle micro-movement and hard to be detected by most of the remote sensors, not to mention infants could be in sleep or cocooned by the heavy clothes during the winter season. Radar-based sensing technology can be a good fit to extract the live signature via subtle movement.

## 1.2 Related works

Current research using radar for human subject monitoring is by measuring the vital sign signals. The continuous wave (CW) based radar measures small physiological movement by calculating the Doppler shift of the received signal [3]. By measuring the vital sign signal using the Doppler shift, occupant inside cabin can be distinguished from static object. Most of the processing algorithm for received signal involves both the in-phase (I) channel and quadrature (Q) channel [4], because of the I/Q channel phase imbalance and to avoid null-point problem [5]. The central frequency of CW radar used in previous work is either located in low frequency band (1-7 GHz) [6][7][8][9][10] or high frequency band ( $\geq 10GHz$ ) [4][11][12][13][14][15][16]. The advantage of using CW based radar is that it is energy efficient and has a simple structure due to the simple transmission signal type. However, it has limitations on multiple human subject localization and clutter differentiation because of the narrow bandwidth.

Different types of wideband radar have been implemented for vital sign monitoring in recent years [17]. They can be categorized into three different types as impulse-radio (IR)

based wideband radar, frequency modulated continuous wave (FMCW) based wideband radar, and stepped frequency continuous wave (SFCW) with wide bandwidth.

Most of the previous works using either IR, FMCW, and SFCW based radar for vital sign detection for the commercial purpose focus on the patient-related vital sign warning [18][19][20], smart home elder person monitoring [21] and rescue and discovery of trapped person underground [22][23]. For clinical applications with wideband radar, they are used for detecting both the breath rate and the heart rate of patients. This non-contact method requires the radar kits to be placed under the bed to have a better measurement of the patient's signal. Because the distance between the patient and radar is relatively short, it could provide the displacement of the chest or belly of that patient with rich details. Wideband radar can also be used to detect the motion of patients so as to protect them from falling and continuously monitor their sleeping stage [24]. Because of the purpose of this scenario, one pair of radar transmitter will be sufficient. The smart home monitoring scenario will, on the other hand, require the penetration ability of radar and need it to have the full coverage of the room. The radar used in this case is to detect the posture change of a person, or more specifically, to detect falling of an elder person [24]. Another approach is to detect emotional change with radar kits [25]. A similar method for smart home application is to use Wi-Fi [26] [27] with a portable device to detect people inside a room. The fact that certain requirement needs to be reached to guarantee the penetration ability limits the choice of radar and makes a radar device bulkier compared to radars with higher central frequency. Similar to smart home vital sign detection, to detect trapped people underground also requires that radar has the ability to penetrate through wall and concrete. Therefore, this type of radar will also have a relatively larger size because of lower central frequency. In recent years, wideband radar with high penetration ability is proposed for driver's state monitoring. One pair of antenna is used for the vital sign detection for driver and other occupants. It can be placed either in front of driver [28] [29] or inside driver's seatback [30]. The research work of driver's state monitoring and vital sign detection is only demonstrated in static scenarios and non-automotive environment. Further research is required to improve the robustness of state monitoring when the vehicle is in operation.

The signal processing for vital sign detection improves over time to increase the respiration rate and heartbeat rate estimation accuracy. Using wideband radar to locate occupants needs data from both fast and slow time domains. Fast time is the time delay of a signal received by receiver and transmitter. Each bin of the fast time axis represents a distance from an object to radar antenna. Slow time, on the other hand, indicates the fast time changing of each detected object. Therefore, localization of living subjects needs to use both information from the slow time axis and the fast time axis for the detection. Single person localization can be based on either fast fourier transform (FFT) [31] [32] based method, variance calculation [30], or entropy-based method [33] over a certain amount of a signal in the slow time. For most of the localization methods in current publications, variance or FFT of certain slow time data has been evaluated by applying the method to every range bin along the fast time axis [34]. Another approach introduced by Schires et al. [30] and Shen et al. [35] calculated the vital signs by collecting a set of the maximum phase change data along the slow time axis. The cross-correlation method is applied in the calculation by feeding two sets of signals from  $\tau_i$  to  $\tau_{i+n}$  along the fast time axis. The first set is the base on  $t_0$  in slow time. The second set is selected from  $t_j$  to  $t_{j+m}$  in sequence along the slow time axis. Walterscheid et al. [36] proposed another approach to locate a person using distributed radars. It helps to mitigate the random movement impact on the vital signs but may have difficulty placing the radars in a limited space such as a vehicle cabin.

Multiple person detection requires radar to estimate the angle of arrival (AoA) of occupants in the cabin. In this case, a radar array will be suitable for estimation and localization, [37][38]. Vehicle cabin is a rather small space with different types of occupants, and it is still challenging to separate two people sitting next to each other. One method proposed using capon-based method as an optimization problem to separate the live signature from two peoples sitting side by side [39]. Another method is to use a deep learning and blind source separation based method to separate multiple people sitting shoulder by shoulder [40]. For multi-person detection in a noisy environment, a clean algorithm is proposed to differentiate the effective power peak from the clutter subtracted received signal peaks [41]. Angle estimation is also essential to distinguish person in the cabin. It can be measured by associating data in data cube [42] such as beamforming or other AoA method such as



multiple signal classification (MUSIC) [43] and its derivative [44]. Single pair radar based multiple occupants detection method is also possible in vehicle cabin [29]. A radar kit is placed on A-pillar to differentiate people in different seats. However, it may not be able to detect weak signals of occupants sitting in rear seats, especially vital sign signals from infants or children.

Vital sign detection is an effective way for occupant localization and detection [45]. A basic method for vital sign detection is using FFT and filter bank to separate breathing signal and heartbeat signal [46]. Other methods that can be used to extract sinusoidal signal from a noisy data are wavelet-based method [47], Hilbert Huang transform based method [48] and variational mode decomposition (VMD) [29]. Based on the Hilbert Huang method, there are also plenty of derivatives proposed to improve the accuracy of detection for sinusoidal data estimation with noise [49] [50] [51][52][53] [54]. However, Hilbert Huang based methods are not robust in noisy environment and data need to be pre-processed before feeding into the method. For noisy signal with low signal to noise ratio, nonlinear Levenberg-Marquardt [55] can be used to effectively extract vital sign signal from the background. Kalman filter [34], principle component analysis (PCA) [56], and independent component analysis (ICA) [57] can also be used for vital sign filtering.

Motion can sometimes interfere with detection of vital signs of occupants in vehicle. The detection of occupants' motion, energy, power spectrum density, [58], entropy [59] and autocorrelation [60] can be used to effectively indicate their movements. The acceleration method [59] can also be applied to distinguish occupants' certain movements by calculating their accelerations. Another method proposed for random movement cancellation is to increase the amount of radar. The subject's movement can be effectively canceled by either placing two radar kits in the front and back [61][62] or placing four radar kits surrounding the subject [63].

Other than monitoring the vital sign for occupant detection, two other approaches are widely used for human subject detection based on radio frequency (RF) reflection pattern. The first is the Keypoint-based method. The majority of researchers select a deep-learning related approach. The heatmap data feedings into learning models are projections of power reflections in azimuth and elevation domains. Power reflections are estimated by conducting

a coherence summation of each radar channel to find the strongest power value targets. Adib et al. [64] proposed the method first, and then Zhao et al. [65] improved it by using a low central frequency radar. The other method is the point cloud approach. It collected a bulk of points in which voxel's power intensity is above threshold in 3D space. Zhao et al. [66] developed a point cloud-based human tracking algorithm for walking subjects inside a room. Both approaches extracted the test subject from the background. Due to the nature of the low central frequency radar, the antenna array package is too large to use limited space, e.g., inside a small car cabin. The large antenna array with higher central frequency radar could provide enough resolution with smaller package sizes [67]. Above are all about subject's shape detection. With knowing the vital signs, the radar can distinguish the living subject from other test subjects accurately.

To understand the reflection pattern of human body, simulation based on human body in RF environment is required. Only a few researchers have discussed the human model for RF system simulation. Van Dorp et al. [68] introduced their approach to simulate the approximate human body. They defined 12 body parts, with head as a sphere, and the rest as either cylinders or ellipsoids, based on different radar cross section (RCS) pattern of each body part. David Winter [69] first proposed a height normalized human body model with 17 key-points based on the human skeleton with consideration of the RCS pattern and later adopted by Chen et al. in [70]. Then, Ram et al. [71] improved Winter's model with added-on RCS gain calculation. The data used for the model came from animation and relative permittivity of tissue. The latter is for the estimation of body reflection. The relative permittivity is calculated according to the radar with a 7.5 GHz central frequency. The use-case of the above research is for walking human subject detection outside or through a wall.

Classification of different types of occupants is usually based on the segmentation of 2D image. Previous works on segmentation are based on different deep learning networks such as FCN [72], Unet [73], Mask R-CNN [74] and Pix2Pix [75]. The majority of posture estimation also uses deep learning architectures to find the Keypoint of subject from 2D images. The posture estimation on 2D images is to find 2D Keypoint from images; popular architectures include Openpose [76], AlphaPose [77], and HRNet [78]. The pose estimation in 3D uses

deep learning networks in higher dimensions, such as V2V-Posenet [79]. Although the 3D posture estimation gives location of each Keypoint in three dimensions, it usually takes more computational power compared to 2D scenarios.

### 1.3 Research objectives and structure

The objectives of the research are to establish the simulation framework and methodology of mmWave wideband planer array radar system for the presence detection, location detection, classification, and liveness signature estimation of in-cabin human subjects, through the application of various cutting-edge radar signal and computer vision imaging processing technologies.

The main contributions of this dissertation are as follows:

- Derived mathematical radar model for occupant imaging and vital sign estimation
- Established a FMCW-based radar simulation system for 3D imaging, posture, and limited vital sign estimation
- Developed reflection-based and Keypoint-based human subject models for occupant classification and posture/limited vital-sign estimation
- Introduced a deep learning method for occupant classification using the radar data from planer antenna array radar

The dissertation is organized as follows. Chapter two proposes a mathematical model for a vehicle in-cabin radar sensing system. Chapter three introduces the basis of in-cabin radar system design. Chapter four introduces the simulation of vehicle in-cabin radar system and design of a generalized human subject model. Chapter five introduces a deep-learning-based method for occupant classification using synthetic data and the designed human subject model. Chapter six presents the posture detection of human subjects. Chapter seven discuss the conclusion and future work.

## 2. MATHEMATIC MODEL FOR HUMAN SUBJECT DETECTION

### 2.1 Frequency domain model

#### 2.1.1 Mathematical model

Inspired by research in [80] using single input single output (SISO) radar for vital sign detection. A Mathematical model for vital sign monitoring using a radar antenna array is derived from the time delay model of a single pulse. It can be described as the following form, as in Eq. (2.1). In this equation,  $X(f)$  is the augmented received signal by radar antenna array in the frequency domain.  $I$  is the number of antenna pairs in the antenna array. Sum of  $\mathbb{K}$  related to both the breathing signal and the heartbeat signal is described in Eq. (2.2).  $A_n$  is the amplitude of the received signal and  $G(\mathbb{K}, f, \phi, \theta, i)$  is described in Eq. (2.4).  $f_c$  in Eq. (2.1) stands for the weighted summation of breathing frequency  $f_b$ , the frequency of heartbeat  $f_h$  and repetition frequency  $f_r$ .  $\delta(f - f_c)$  is the time delay Dirac delta function describing the displacement of a single person.

$$X(f) = \sum_{u=-\infty}^{\infty} \sum_{i=1}^I \sum_{\mathbb{K}=-\infty}^{\infty} A_n G(\mathbb{K}, f, \phi, \theta, i) \delta(f - f_c) \quad (2.1)$$

Where

$$\begin{aligned} \sum_{\mathbb{K}=-\infty}^{\infty} &= \sum_{K_1=-\infty}^{\infty} \sum_{K_2=-\infty}^{\infty} \dots \sum_{K_{N_b}=-\infty}^{\infty} \sum_{K_1=-\infty}^{\infty} \sum_{K_2=-\infty}^{\infty} \dots \sum_{K_{N_b}=-\infty}^{\infty} \\ &\quad \sum_{H_1=-\infty}^{\infty} \sum_{H_2=-\infty}^{\infty} \dots \sum_{H_{N_h}=-\infty}^{\infty} \sum_{H_1=-\infty}^{\infty} \sum_{H_2=-\infty}^{\infty} \dots \sum_{H_{N_h}=-\infty}^{\infty} \end{aligned} \quad (2.2)$$

$$f_c = \left[ \sum_{p=1}^{N_b} p(K_p + K_p) \right] f_b + \left[ \sum_{q=1}^{N_h} q(H_p + H_p) \right] f_h + u f_r \quad (2.3)$$

$$G(\mathbb{K}, f, \phi, \theta, i) = f_r P(f) e^{-j2\pi f(\tau_a(\phi, \theta, i) + \tau_0)} (-1)^{\sum_{p=1}^{N_b} K_p + \sum_{q=1}^{N_h} H_q} j^{\left(\sum_{p=1}^{N_b} K_p + \sum_{q=1}^{N_h} H_q\right)} \prod_{p=1}^{N_b} J_{K_p}(2\pi f A_p^b) \prod_{p=1}^{N_b} J_{K_p}(2\pi f B_p^b) \prod_{q=1}^{N_h} J_{H_q}(2\pi f A_q^h) \prod_{q=1}^{N_h} J_{H_q}(2\pi f B_q^h) \quad (2.4)$$

### 2.1.2 Derivation

A detailed proof of this model can be shown step by step as follow. The live signature of a living subject can be measured by getting the time delay  $\tau_d$  of the transmitted signal as in Eq. (2.5). Time delay here is assumed without any phase delay. The time delay contains two parts. One is the constant time delay  $\tau_0$ , and the other part is the time delay of breathing and heartbeat in Fourier expansion form where  $f_b$  is the frequency of respiration and  $f_h$  is the frequency for the heartbeat.

$$\begin{aligned} \tau_d(t_s) = \tau_0 &+ \sum_{p=1}^{N_b} \left[ A_p^b \sin(2\pi p f_b t_s) + B_p^b \cos(2\pi p f_b t_s) \right] \\ &+ \sum_{q=1}^{N_h} \left[ A_q^h \sin(2\pi q f_h t_s) + B_q^h \cos(2\pi q f_h t_s) \right] \end{aligned} \quad (2.5)$$

Since reflected signal received by each receiver antenna pair of radar antenna array has a time delay  $\tau_a$  between each other, it needs to be corrected to increase the received gain. The time delay  $\tau_a$  depends on azimuth incident angle  $\phi$ , elevation incident angle  $\theta$ , and distance of the antenna to the array origin.

Therefore, the total reflected power from each location in 2D-domain can be shown in the FMCW waveform's convolution with the impulse signal in Eq. (2.6).

$$x(t_s) = p(t_s) * h(t_s) \quad (2.6)$$

where  $p(t_s)$  is the FMCW burst waveform in time domain that responds to the form of  $A_n e^{j\theta}$ , and  $h(t_s) = \sum_{n=-\infty}^{\infty} \sum_{i=1}^I \delta(t_s - nT_r - \tau_d(t_s) - \tau_a(\phi, \theta, i))$  is for impulse waveform time response.  $T_r$  is the repetition time of each transmitted signal.

By applying continuous time Fourier transform (*CTFT*) to Eq. (2.5), we can express the received signal in the frequency domain as in Eq. (2.8)

$$X(f) = \int_{-\infty}^{\infty} x(t_s) e^{-2\pi f t_s} dt_s \quad (2.7)$$

$$= \sum_{n=-\infty}^{\infty} \sum_{i=1}^I P(f) A_n e^{-j2\pi f(nT_r + \tau_d(t_s) + \tau_a(\phi, \theta, i))} \quad (2.8)$$

For simplicity, assuming  $\beta = 2\pi f$ ,  $\beta_b = 2\pi f_b$ ,  $\beta_h = 2\pi f_h$  for each frequency, we can get the following Eq. (2.10).

$$X(f) = \sum_{n=-\infty}^{\infty} \sum_{i=1}^I P(f) A_n e^{-j2\pi f(nT_r + \tau_a(\phi, \theta, i))} e^{-j2\pi f(\tau_d(t_s))} \quad (2.9)$$

$$= \sum_{n=-\infty}^{\infty} \sum_{i=1}^I P(f) A_n e^{-j\beta(nT_r + \tau_a(\phi, \theta, i))} e^{-j\beta(\tau_d(t_s))} \quad (2.10)$$

Substitute  $t_s = nT_r$  and Eq. (2.5) into Eq. (2.10), we get

$$X(f) = \sum_{n=-\infty}^{\infty} \sum_{i=1}^I P(f) A_n e^{-j\beta(nT_r + \tau_a(\phi, \theta, i) + \tau_0)} \cdot e^{-j\beta \left( \sum_{p=1}^{N_b} [A_p^b \sin(p\beta_b nT_r) + B_p^b \cos(p\beta_b nT_r)] + \sum_{q=1}^{N_h} [A_q^h \sin(q\beta_h nT_r) + B_q^h \cos(q\beta_h nT_r)] \right)} \quad (2.11)$$

Since  $e^{\sum_k \alpha_k} = \prod_k e^{\alpha_k}$ , second exponential term in Eq. (2.11) could be expressed as in Eq. (2.12).

$$\prod_{p=1}^{N_b} e^{-j\beta A_p^b \sin(p\beta_b nT_r)} \prod_{p=1}^{N_b} e^{-j\beta B_p^b \cos(p\beta_b nT_r)} \prod_{q=1}^{N_h} e^{-j\beta A_q^h \sin(q\beta_h nT_r)} \prod_{q=1}^{N_h} e^{-j\beta B_q^h \cos(q\beta_h nT_r)} \quad (2.12)$$

With Jacobi-Anger expansion, terms in Eq. (2.12) will be further expanded in the form of Eq. (2.13) with the Bessel function of the first kind  $J_\alpha(x)$ .

$$\begin{aligned} & \prod_{p=1}^{N_b} \sum_{K_p=-\infty}^{\infty} J_{K_p}(\beta A_p^b) e^{-jK_p p \beta_b n T_r} \prod_{p=1}^{N_b} \sum_{K_p=-\infty}^{\infty} J_{K_p}(\beta B_p^b) j^{-K_p} e^{-jK_p p \beta_b n T_r} \\ & \prod_{q=1}^{N_h} \sum_{H_q=-\infty}^{\infty} J_{H_q}(\beta A_q^h) e^{-jH_q q \beta_h n T_r} \prod_{q=1}^{N_h} \sum_{H_q=-\infty}^{\infty} J_{H_q}(\beta B_q^h) j^{-H_q} e^{-jH_q q \beta_h n T_r} \end{aligned} \quad (2.13)$$

Applying the distributive law across the summations in (2.13), Eq. (2.8) can be reorganized as in Eq.(2.14).

$$\begin{aligned} X(f) &= \sum_{n=-\infty}^{\infty} \sum_{i=1}^I P(f) A_n e^{-j\beta(nT_r + \tau_a(\phi, \theta, i) + \tau_0)} \sum_{\mathbb{K}=-\infty}^{\infty} \mathbb{J}(\alpha, \mathbb{K}, \beta) e^{-j(\cdot)nT_r} \\ &= \sum_{n=-\infty}^{\infty} \sum_{i=1}^I \sum_{\mathbb{K}=-\infty}^{\infty} P(f) A_n \mathbb{J}(\alpha, \mathbb{K}, \beta) e^{-j(\beta + \cdot)(nT_r)} e^{-j\beta(\tau_a(\phi, \theta, i) + \tau_0)} \end{aligned} \quad (2.14)$$

Where  $\mathbb{K}$  is defined as the product of summations in the form of Eq. (2.15)

$$\begin{aligned} \sum_{\mathbb{K}=-\infty}^{\infty} &= \sum_{K_1=-\infty}^{\infty} \sum_{K_2=-\infty}^{\infty} \cdots \sum_{K_{N_b}=-\infty}^{\infty} \sum_{K_1=-\infty}^{\infty} \sum_{K_2=-\infty}^{\infty} \cdots \sum_{K_{N_b}=-\infty}^{\infty} \\ &\quad \sum_{H_1=-\infty}^{\infty} \sum_{H_2=-\infty}^{\infty} \cdots \sum_{H_{N_h}=-\infty}^{\infty} \sum_{H_1=-\infty}^{\infty} \sum_{H_2=-\infty}^{\infty} \cdots \sum_{H_{N_h}=-\infty}^{\infty} \end{aligned} \quad (2.15)$$

$\mathbb{J}(\alpha, \mathbb{K}, \beta)$  is the product of the Bessel function of the first kind defined as follows.  $\alpha$  is defined as the scale factor of Bessel function and  $\alpha \in \{-1, 1\}$ .

$$\begin{aligned} \mathbb{J}(\alpha, \mathbb{K}, \beta) &= J_{\alpha K_1}(\beta A_1^b) J_{\alpha K_2}(\beta A_2^b) \cdots J_{\alpha K_{N_b}}(\beta A_{N_b}^b) \\ &\quad J_{\alpha K_1}(\beta B_1^b) J_{\alpha K_2}(\beta B_2^b) \cdots J_{\alpha K_{N_b}}(\beta B_{N_b}^b) \\ &\quad J_{\alpha H_1}(\beta A_1^h) J_{\alpha H_2}(\beta A_2^h) \cdots J_{\alpha H_{N_h}}(\beta A_{N_h}^h) \\ &\quad J_{\alpha H_1}(\beta B_1^h) J_{\alpha H_2}(\beta B_2^h) \cdots J_{\alpha H_{N_h}}(\beta B_{N_h}^h) \\ &\quad j^{-\alpha(K_1 + K_2 + \cdots + K_{N_b} + H_1 + H_2 + \cdots + H_{N_h})} \end{aligned} \quad (2.16)$$

Where  $\cdot$  is defined as in Eq. (2.14). Here it can be shown that the frequency of breathing and heartbeat can be expressed linearly in the form of  $kf_b + lf_h$ .

$$\begin{aligned}
\cdot &= \left[ (K_1 + K_1) + 2(K_2 + K_2) \dots + N_b(K_{N_b} + K_{N_b}) \right] f_b \\
&+ \left[ (H_1 + H_1) + 2(H_2 + H_2) \dots + N_h(H_{N_h} + H_{N_h}) \right] f_h \\
&= \underbrace{\left[ \sum_{p=1}^{N_b} p(K_p + K_p) \right]}_k f_b + \underbrace{\left[ \sum_{q=1}^{N_h} q(H_q + H_q) \right]}_l f_h
\end{aligned} \tag{2.17}$$

According to Dirac delta function's property and applying Eq. (2.18) from the definition of Dirac comb.

$$\sum_{n=-\infty}^{\infty} e^{-j2\pi f n T_r} = \frac{1}{T_r} \sum_{m=-\infty}^{\infty} \delta\left(f - \frac{m}{T_r}\right) \tag{2.18}$$

to Eq. (2.14) and define  $f_r = \frac{1}{T_r}$  as the repetition frequency.

Since each  $K$  and  $H$  is within the range of  $(-\infty, \infty)$  and the Bessel function is symmetric, the sign of  $K$  and  $H$  can be flipped. Furthermore, from the property of Bessel function, we have  $J_{-n}(x) = (-1)^n J_n(x)$ . Finally, we can have the final expression for vital sign estimation using a radar antenna array as Eq. (2.19)

$$\begin{aligned}
X(f) &= \sum_{u=-\infty}^{\infty} \sum_{i=1}^I \sum_{\mathbb{K}=-\infty}^{\infty} A_n f_r e^{-j\beta(\tau_a(\phi, \theta, i) + \tau_0)} \mathbb{J}(1, \mathbb{K}, \beta) P(f) \delta(f + \cdot - u f_r) \\
&= \sum_{u=-\infty}^{\infty} \sum_{i=1}^I \sum_{\mathbb{K}=-\infty}^{\infty} A_n f_r e^{-j\beta(\tau_a(\phi, \theta, i) + \tau_0)} \mathbb{J}(-1, \mathbb{K}, \beta) P(f) \delta(f - \cdot - u f_r) \\
&= \sum_{u=-\infty}^{\infty} \sum_{i=1}^I \sum_{\mathbb{K}=-\infty}^{\infty} A_n f_r e^{-j\beta(\tau_a(\phi, \theta, i) + \tau_0)} \underbrace{(-1)^{\sum_{p=1}^{N_b} K_p + \sum_{q=1}^{N_h} H_q}}_{G(\mathbb{K}, f, \phi, \theta, i)} \mathbb{J}(1, \mathbb{K}, \beta) P(f) \dots \tag{2.19} \\
&\quad \delta(f - \underbrace{k f_b + l f_h - u f_r}_{f_c}) \\
&= \sum_{u=-\infty}^{\infty} \sum_{i=1}^I \sum_{\mathbb{K}=-\infty}^{\infty} A_n G(\mathbb{K}, f, \phi, \theta, i) \delta(f - f_c)
\end{aligned}$$



## 2.2 Time-domain model

### 2.2.1 Mathematical model

A time-domain mathematical model can be derived from the FMCW chirp equation for a single human body. The transmitted signal hits the human body and reflects the antenna array. The back-scattered signal is an attenuated and phase-shifted version of the transmitted signal and follows humans' RCS pattern. The received signal is then mixed with the transmitted signal to convert it to the based band. A planar radar antenna array model can detect a human body in both azimuth and elevation domain in this model.

The model for a single human body can be extended to multiple occupants detection with random movement concerning slow time  $t_s$  and fast time  $t_f$ , and it can be described as in Eq. (2.20).

$$y(t_f, t_s) = \sum_{r=R_s}^{R_e} \sum_{l=1}^L \sum_{i=1}^I [A_{R_l} \sqrt{G(\sigma_e)} e^{j(\phi_{rl} + \phi_{ml} + \phi_a)} + w_i] \quad (2.20)$$

with

$$\begin{aligned} \phi_{r_l}(t_f(r_l), t_s) = & 2\pi f_b(R_{mi}(t_s))t_f(r_l) + \frac{4\pi}{\lambda_m} [A_l^B \sin(2\pi f_{br}t_s) \\ & + A_l^H \sin(2\pi f_{hr}t_s)] \end{aligned} \quad (2.21)$$

$$\phi_{m_l}(t_f(r_l), t_s) = 2\pi f_{ib}(R_{ma}(t_s))t_f(r_l) + \frac{4\pi}{\lambda_m} R_{ma}(t_s) \quad (2.22)$$

$$\phi_a(\phi(t_s), \theta(t_s), i) = \frac{2\pi}{\lambda_m} \tau_a(\phi(t_s), \theta(t_s), i) \quad (2.23)$$

where

$t_f, t_s :$	Fast time and slow time
$I :$	Number of antenna pair
$L :$	Number of targets
$A_{R_l} :$	Received power magnitude by certain target
$D_s, D_e :$	Start and end radial distance
$G(\sigma_e) :$	Gain factor with respect to equivalent RCS
$f_{b_l} :$	Beat frequency of certain target
$f_{bi_l} :$	Incremental beat frequency of certain target
$A^B, A^H :$	Amplitude of breathing and heart beat
$f_b, f_h :$	Frequency of breathing and heart beat
$R_{ma}(t_s) :$	Change of range for macro-movement
$\lambda_m :$	Maximum wavelength
$\tau_a :$	radar antenna array related delay
$w_i :$	white noise

In this model, the change of amplitude consists of two parts. The first part  $A_{R_l}$  is two-way power loss during transmission in free space. The second part is the reflected gain with respect to RCS  $G(\sigma_e)$  in form of Eq. (2.24) [81]

$$G(\sigma_e) = \frac{4\pi\sigma_e}{\lambda^2} \quad (2.24)$$

Where the equivalent RCS  $\sigma_e$  is the summation of each RCS value of N body parts as in Eq. (2.25) [82]

$$\sigma_e = \Gamma_h \left| \sum_{n=1}^N \sqrt{\sigma_n} e^{j\phi_n} \right|^2 \quad (2.25)$$

Because the human body is not a metallic object, the relative permittivity of humans needs to be considered in Eq. (2.25) as a reflection coefficient  $\Gamma_h$  [71]. The  $\Gamma_h$  is derived from the Maxwell equation. It is known that the speed of light is given by relative permeability  $\mu_0$  and relative permittivity  $\epsilon_0$  in air. For human tissues with relative permeability  $\mu_r$  and the average relative permittivity of the human body  $\epsilon_{r_{avg}}$ , the relative speed of light is calculated by

$$c_r = \frac{1}{\sqrt{\mu_0 \mu_r \epsilon_0 \epsilon_{r_{avg}}}} \quad (2.26)$$

Then the index of refraction can be given by

$$n = \frac{c}{c_r} = \sqrt{\mu_r \epsilon_{r_{avg}}} \quad (2.27)$$

So  $\Gamma_h$  is calculated by the Fresnel reflection coefficient at normal incidence as in Eq. (2.28)

$$\Gamma_h = \left| \frac{1 - n}{1 + n} \right|^2 \quad (2.28)$$

All the human body's permeability  $\mu_r$  in this dissertation is assumed to be constant and equal to  $\mu_0$  for simplicity. The value of  $\Gamma_h$  can vary for humans of different ages. This is because the total body water content (TBW) can affect the value of relative permittivity  $\epsilon_r$ . In [83], they introduced an empirical equation to calculate  $\epsilon_r$  for a male child aged from 3 months to 13 years old, as shown in Eq. (2.29)

$$\epsilon_{r_{ch}} = 2.616^{(X-6.63)} \epsilon_{r_A}^{2.4813(1-0.09X)} \quad (2.29)$$

Where  $\epsilon_{r_{ch}}$  is the relative permittivity for child and  $\epsilon_{r_A}$  is the relative permittivity for adult.  $X$  is a parameter related to the height  $H$  and weight  $W$  of child yield Eq. (2.30)

$$X = H^{0.65} W^{0.65} \quad (2.30)$$

An average relative permittivity of a complex object with multiple layers of tissues can be obtained by combining each tissue layer's relative permittivity weighted by its thickness. The organs and bones in the thorax are surrounded by skin and fat. The electromagnetic wave reaches the skin and fat first, then a part of the wave goes through the body following a path that depends on where the wave reaches first. Thus, the average relative permittivity for each body may vary in terms of incident angle of radar waveform.

Table 2.1 shows the relative permittivity of some tissues for adults and children using radar with a central frequency of 79 GHz. The former is from Gabriel's report, [84] while the latter is calculated by Eq. (2.29). In this work, the chest and belly are the areas of interest because they are where the lungs and heart are located.

**Table 2.1.** Selected relative permittivity for adults and children

Tissue name	Relative permittivity (F/m)	
	Adult	Child
Aorta	7.882	10.181
Blood	9.826	12.381
Blood Vessel	7.882	10.181
Body Fluid	8.679	11.090
Bone Cancellous	4.491	6.181
Bone Cortical	3.503	4.958
Bone Marrow	2.989	4.307
Breast Fat	2.619	3.831
Fat	2.988	4.306
Heart	9.719	12.261
Lung Deflated	9.05	11.509
Lung Inflated	4.532	6.231
Muscle	10.27	12.876
Nerve	6.948	9.104
Oesophagus	10.66	13.309
Skin Dry	6.45	8.522
Skin Wet	8.391	10.763
Spinal Cord	6.948	9.104
Thymus	10.1	12.686

Based on the previous discussion, the average relative permittivity  $\epsilon_{r_{avg}}$  is 9  $F/m$  for an adult and 11  $F/m$  for a child. The Fresnel reflection coefficient is then 0.25 for an adult and 0.29 for a child.

Besides the change of amplitude, there are also phase changes in the time domain mathematical model. The first part  $\phi_{rl}^r$  is due to the radial movement of the human body that stayed without big movement for the radar antenna array. It depends on both changes in a slow time and fast time. Vital signs such as respiration rate and heart rate are of interest for automotive applications among all radial movements. For automotive radar with 79 GHz central frequency, the range resolution is about 4cm. With tracking of phase changes, the vital sign signal with mm-level displacement can be detected. By detecting a human's respiration and heart rate, we could distinguish them from other static objects.

The second part of phase change  $\phi_m^r$  is the Doppler shift that could describe the macro movement of the human body. It depends on the change in a slow time. It can be used to track significant movements of human body.

The third part of the phase difference is by the radar antenna array. For a planar array in the horizontal and vertical dimension, the phase corresponds to the time delay  $\tau_a$ , which depends on the azimuth  $\theta$  and elevation angle  $\phi$  of each virtual antenna towards the target.

Last but not least, some noise factor is also considered in this model. A white noise  $w_i$  is introduced in this model for each channel to mimic the noise during transmission. Also, radar's mutual coupling effect can be substantial because the equivalent aperture of the radar antenna array is small due to high center frequency. To mitigate the noise effect, the range of interest is selected while neglecting signals from other distances.

### 2.2.2 Derivation

The proof of Eq. (2.20) starts with an equation for a transmitting signal given by Brooker [85]

$$X_T(t_f) = A_T e^{j(2\pi f_s t_f + \pi \frac{B}{T_c} t_f^2 + \phi_0(t_f))} \quad (2.31)$$

with the real part of the signal expressed by Euler's formula

$$X_T(t_f) = A_T \cos(2\pi f_s t_f + \pi \frac{B}{T_c} t_f^2 + \phi_0(t_f)) \quad (2.32)$$

where  $t_f$  is fast time,  $A_T$  stands for the amplitude transmitting power,  $f_s$  is the starting frequency of the chirp,  $B$  represents the bandwidth of the chirp,  $T_c$  is the chirp duration,  $\phi_0$  is the initial phase of the transmitted signal.

Power loss and traveling phase delay of electromagnetic waves during the transmission makes the receiving signal a scaled and shifted version of the transmitting signal. The receiving signal has certain scale factors of attenuation and time delay, including the RCS's reflected gain. The phase delay due to the projection of the radial axis's movements is derived from the time delay. Both the receiving power and phase delay are expressed as,

$$X_R(t_f) = \alpha A_T \sqrt{G(\sigma_e)} \cos(2\pi f_s(t_f - t_d) + \pi \frac{B}{T_c}(t_f - t_d)^2 + \phi_0(t_f - t_d)) \quad (2.33)$$

where  $t_d$  is the time delay.

The transmitting signal is converted to I and Q channels and then mixed with the receiving signal represented by

$$y(t) = X_T(t) \times X_R(t) \quad (2.34)$$

Substituting Eq. (2.32) and Eq. (2.33) with Eq. (2.34) and after some derivations, we get

$$y(t_f, t_s) = A_R \sqrt{G(\sigma_e)} \cos(j(2\pi f_s t_d + 2\pi \frac{B}{T_c} t_d t_f - \pi \frac{B}{T_c} t_d^2 + \Delta\phi_0(t))) \quad (2.35)$$

Eq. (2.35) can be expressed as

$$y(t_f, t_s) = A_R \sqrt{G(\sigma_e)} e^{j(2\pi f_s t_d + 2\pi \frac{B}{T_c} t_d t_f - \pi \frac{B}{T_c} t_d^2 + \Delta\phi_0(t))} \quad (2.36)$$

where  $A_R$  is the power magnitude of receiver defined by  $\frac{\alpha A_T^2}{2}$ .  $\Delta\phi_0(t_d)$  represents the residual phase noise as

$$\Delta\phi_0(t_d) = \phi_0(t_f) - \phi_0(t_f - t_d) \quad (2.37)$$

For random movement targets, the time delay consists of three range changes  $R_{mb}$ ,  $R_{mi}$ , and  $R_{ma}$

$$t_d = \frac{2(R_{mb}(t_s) + R_{ma}(t_s) + R_{mi}(t_s))}{c} \quad (2.38)$$

$R_{mb}$  is the base-movement. It is the distance from test subject to radar.  $R_{ma}$  is the macro-movement representing the head, torso, and limb activities of subject.  $R_{mi}$  is a micro-movement that contains chest and heart signals, both of which can be expressed by sinusoidal waves

$$R_{mi}(t_s) = A^B \sin(2\pi f_{br} t_s) + A^H \sin(2\pi f_{hr} t_s) \quad (2.39)$$

where  $A^B$  and  $A^H$  are the amplitude of human subject's chest and heart signals, and  $f_{br}$  and  $f_{hr}$  are the corresponding frequencies.

Based on [67], the third term of exponential in Eq. (2.36) can be neglected because it is negligibly small in practice. The phase difference can also be neglected due to the range-correlation effect.

Therefore, Eq.(2.36) can be simplified to

$$y(t_f, t_s) = A_R \sqrt{G(\sigma_e)} e^{j(2\pi f_s t_d + 2\pi \frac{B}{T_c} t_d t_f)} \quad (2.40)$$

It is known that the beat frequency  $f_b$  is defined by chirp slope  $K_c$  and travel time of signal  $T_r$  and can be divided to three parts,  $f_{bmb}$ ,  $f_{bmi}$  and  $f_{bma}$ , correspondingly

$$f_b = K_c T_r = f_{bmb} + f_{bma} + f_{bmi} \quad (2.41)$$

where

$$f_{bmb} = \frac{2BR_{mb}(t_s)}{cT_c} \quad (2.42)$$

$$f_{bma} = \frac{2BR_{ma}(t_s)}{cT_c} \quad (2.43)$$

$$f_{bmi} = \frac{2BR_{mi}(t_s)}{cT_c} \quad (2.44)$$

After a few steps of derivation by substituting Eq. (2.38), (2.41)-(2.44), Eq. (2.40) can be expressed as

$$\begin{aligned} y(t_f, t_s) &= A_R \sqrt{G(\sigma_e)} e^{j(\frac{4\pi}{\lambda_m}(R_{mb}(t_s)+R_{ma}(t_s)+R_{mi}(t_s)))} e^{j(2\pi(f_{bmb}+f_{bma}+f_{bmi})t_f)} \\ &= A_R \sqrt{G(\sigma_e)} e^{j(2\pi f_{bmb}t_f + \frac{4\pi}{\lambda_m}R_{mb}(t_s))} e^{j(2\pi f_{bma}t_f + \frac{4\pi}{\lambda_m}R_{ma}(t_s))} e^{j(2\pi f_{bmi}t_f + \frac{4\pi}{\lambda_m}R_{mi}(t_s))} \\ &= A_R \sqrt{G(\sigma_e)} e^{j(\phi^{mb}(t_f, t_s) + \phi^{ma}(t_f, t_s) + \phi^{mi}(t_f, t_s))} \end{aligned} \quad (2.45)$$

For a radar with antenna array, the time delay of each antenna pair varies. An additional phase delay term  $\phi^{ant}(\phi(t_s), \theta(t_s), i)$  is included, which is a function of the incident angle  $(\phi(t_s), \theta(t_s))$  with argument  $t_s$  and position of  $i^{th}$  antenna pair. Therefore, the model with the antenna array is expressed by

$$y(t_f, t_s, i) = A_R \sqrt{G(\sigma_e)} e^{j(\phi^{mb}(t_f, t_s) + \phi^{ma}(t_f, t_s) + \phi^{mi}(t_f, t_s))} e^{j(\phi^{ant}(\phi(t_s), \theta(t_s), i))} \quad (2.46)$$

Next is to add multiple-targets detection and range of interest (ROI) features into the model. The features are used to detect all the meaningful targets within the interest range, increase detection accuracy, mitigate the coupling effect, and reduce computation resources. Thus, the model is rearranged as

$$y(t_f, t_s) = \sum_{d=D_s}^{D_e} \sum_{l=1}^L \sum_{i=1}^I A_{R_l} \sqrt{G(\sigma_e)} e^{j(\phi_l^{mb}(t_f, t_s) + \phi_l^{ma}(t_f, t_s) + \phi_l^{mi}(t_f, t_s) + \phi_i^{ant}(\phi(t_s), \theta(t_s)))} \quad (2.47)$$



where  $D_s$  and  $D_e$  are start and end radial ranges.  $L$  is the total number of detected targets.

Lastly, a white noise function represents the noises that come from the environment, electronic hardware, system package, and others. The final model is given by

$$y(t_f, t_s) = \sum_{d=D_s}^{D_e} \sum_{l=1}^L \sum_{i=1}^I [A_{Rl} \sqrt{G(\sigma_e)} e^{j(\phi_l^{mb}(t_f, t_s) + \phi_l^{ma}(t_f, t_s) + \phi_l^{mi}(t_f, t_s) + \phi_i^{ant}(\phi(t_s), \theta(t_s)))} + w_i(t_s)] \quad (2.48)$$

where  $w_i(t_s)$  is the white noise.

### 2.3 SINR enhancement

From Eq. (2.48), the output amplitude of received signal using radar antenna array is related to the amplitude  $\sum_{i=1}^L A_R$ . Assuming  $A_R$  of the received signal were all equal, then the summation of the received signal would be

$$A_R = L A_R \quad (2.49)$$

The phase difference of each received can be corrected by coherent integration using either summer or FFT.

Assuming output power from SISO radar were  $P_{signal} = A_R^2$ . We could assume the output power of the signal from the radar antenna array as

$$P_{signal} = (A_R)^2 = L^2 A_R^2 = L^2 P_{signal} \quad (2.50)$$

The noise signal's overall noise level can be estimated, including clutter, side lobe, multi-path, and jamming signal. All of the signals are in the I and Q channel. Assuming amplitude of the noise signal were  $A_n$ , and amplitude of jamming signal were  $A_j$ , for SISO radar, the amplitude and power could be shown as

$$A_{noise} = \frac{1}{\sqrt{2}}((A_{nI} + jA_{nQ}) + (A_{jI} + jA_{jQ})) \quad (2.51)$$

$$P_{noise} = A_{nI}^2 + A_{nQ}^2 + A_{jI}^2 + A_{jQ}^2 \quad (2.52)$$

When using a radar antenna array, the noise is a summation of the noise of  $n$  antennas shown below

$$\begin{aligned} A_{noise} &= \sum_n A_{noise}^n \\ &= \frac{1}{\sqrt{2}}((\sum_n A_{nI}^n + \sum_n A_{jI}^n) + j(\sum_n A_{nQ}^n + \sum_n A_{jQ}^n)) \\ &= \frac{1}{\sqrt{2}}((A_{nI} + A_{jI}) + j(A_{nQ} + A_{jQ})) \end{aligned} \quad (2.53)$$

So the output noise power of received radar is given by

$$\begin{aligned} P_{noise} &= E(A_{noise} A_{noise}^*) \\ &= \frac{1}{2}E((A_{nI} + A_{jI})^2) + \frac{1}{2}E((A_{nQ} + A_{jQ})^2) \end{aligned} \quad (2.54)$$

It can be further expanded as

$$\begin{aligned} P_{noise} &= \frac{1}{2}[E((A_{nI})^2) + E((A_{jI})^2) + E(2A_{nI}A_{jI}) \\ &\quad + E((A_{nQ})^2) + E((A_{jQ})^2) + E(2A_{nQ}A_{jQ})] \end{aligned} \quad (2.55)$$

The clutter signal and jammer signal received by each antenna were assumed uncorrelated, then each of the cross expectation would equal to zero. So the amplitude of noise power can be simplified as

$$P_{noise} = \frac{1}{2}[E((A_{nI})^2) + E((A_{jI})^2) + E((A_{nQ})^2) + E((A_{jQ})^2)] \quad (2.56)$$

where each of expectation of Eq. (2.56) with noise level  $NL = \{noise, jammer\}$  and  $i = \{I, Q\}$  calculated by

$$\begin{aligned} E((A_{NLi})^2) &= E\left(\sum_{j=1}^n A_{NLi}^j \sum_{k=1}^n A_{NLi}^k\right) \\ &= \sum_n E((A_{NLi}^n)^2) + \sum_{j,k \in [1,n], j \neq k} E(A_{NLi}^j A_{NLi}^k) \end{aligned} \quad (2.57)$$

Assuming each receiver's noise signal were not correlated, then the second term in Eq. (2.57) equaled zero.

$$E(A_{NLi}^j A_{NLi}^k) = 0, \forall j \neq k \quad (2.58)$$

The first term in Eq. (2.57) is wide sense stationary (WSS) and with zero-mean, so we can get Eq. (2.59) with  $\sigma$  standing for the standard deviation of the noise power.

$$E((A_{NLi})^2) = \sum_n E((A_{NLi}^n)^2) = L\sigma^2 \quad (2.59)$$

Therefore, the final noise power is as in Eq. (2.60) where  $\sigma^2$  is the noise power of each receiver

$$P_{noise} = 2L\sigma^2 \triangleq 2LP_{noise} \quad (2.60)$$

The signal-to-interference-plus-noise ratio (SINR) can be finally calculated by the ratio of output signal power versus the output noise power shown in Eq. (2.60). By using radar antenna array, the SINR of the received signal will be  $\frac{L}{2}$  times larger than the original SINR as in Eq. (2.61).

$$\begin{aligned} SINR &= \frac{P_{signal}}{P_{noise}} \\ &= \frac{L^2 P_{signal}}{2LP_{noise}} \\ &= \frac{L}{2} * SINR \end{aligned} \quad (2.61)$$

### 3. BASIS OF VEHICLE IN-CABIN RADAR SYSTEM DESIGN

#### 3.1 Design constraint and requirement

Compared to other applications using wideband radar for localization and occupancy detection, the vehicle cabin space is relatively smaller, leading to challenges for detection, classification, and liveness estimation.

##### 3.1.1 Vehicle size and layout

Vehicle cabin size varies according to its usage. A radar system needs to be designed to cover the whole in-cabin arena, from sports cars having small cabin space with lower ceiling height to full-size vans having large cabin space with higher ceiling. The table 3.1 listed below is the cabin space of the typical Ford vehicles.

Five seats with two rows are chosen for the evaluation of vehicle occupant detection. The Sedan car and sport utility vehicle (SUV) are the main focus in this research for occupant monitoring. Therefore, the typical cabin length from windshield to trunk's back for this vehicle type is about three meters. It limits the maximum detection range to three meters.

**Table 3.1.** The typical size of Ford vehicle

Model	Size (L×W×H mm)
Fiesta	4060 ×1756 ×1498
Focus	4668 ×1825 ×1481
Fusion	4872 ×1852 ×1478
EcoSport	4273 ×1765 ×1653
Escape	4524 ×1839 ×1684
Edge	4808 ×1928 ×1692
Explorer	5047 ×2004 ×1778
Expedition	5636 ×2029 ×1941
F-150	5890-6190 ×2029 ×1920
Mustang	4784 ×1916 ×1381

For vehicle cabin layout, most vehicles have two front seats and three back seats. There are five zones in-cabin that need to be evaluated to detect movement and occupants' vital signs. Driver and one passenger usually sit in the front while other passengers such as adults,

children, infants, or pets sit in the back; children and infants in the rear seat should be inside special seats hooked to back seats.

### **3.1.2 Safety requirements**

According to requirements of different states in the U.S. [86], children and infants need to be protected with certain types of child seats or infant seats. The requirement varies from state to state and can be categorized into four types: rearward-facing seats, forward-facing seats, belt-positioning booster seats, and lap and shoulder belts [87].

The rearward-facing seats are for children aged zero to two, or for those whose weight is within the limit of rearward-facing seats. Once children surpass the rearward-facing seat's weight limit, they should be switched to forward-facing seats with a harness. They should switch to belt-positioning booster seats before they reach 4'9" and are aged between eight and twelve. If children fit the belt properly and are under thirteen years old, they must ride in the back seat using the lap and shoulder belts.

Seats for infants and children have different heights of back and cover. For rearward-facing infant seats, they have a canopy to cover infants' sunshade; during winter, infants are covered to keep them warm. The convertible seats have convertible back, which can be adjusted to comply with children's height. Booster seats also have either a low back or a high back according to children's size.

## **3.2 Radar parameter design**

Low central frequency UWB radar has been studied and used for health monitoring inside hospital or house. Its design and parameter setup do not fit the automotive applications. The mmWave radar is selected, and all the parameters are redesigned.

### **3.2.1 Central frequency**

For currently available radar for advanced driver assistance systems (ADAS) in vehicle, both central frequency 24 GHz and 79 GHz are popular choices in most countries. In Japan, 60 GHz is selected for automotive radar in its ADAS system. Most of radars currently

mounted on vehicle are for the detection of outside objects [88]. FCC has just recently approved 60 GHz central frequency radar for vehicle in-cabin application [89]. This central frequency radar system will be the direction for vehicle in-cabin occupant detection and classification.

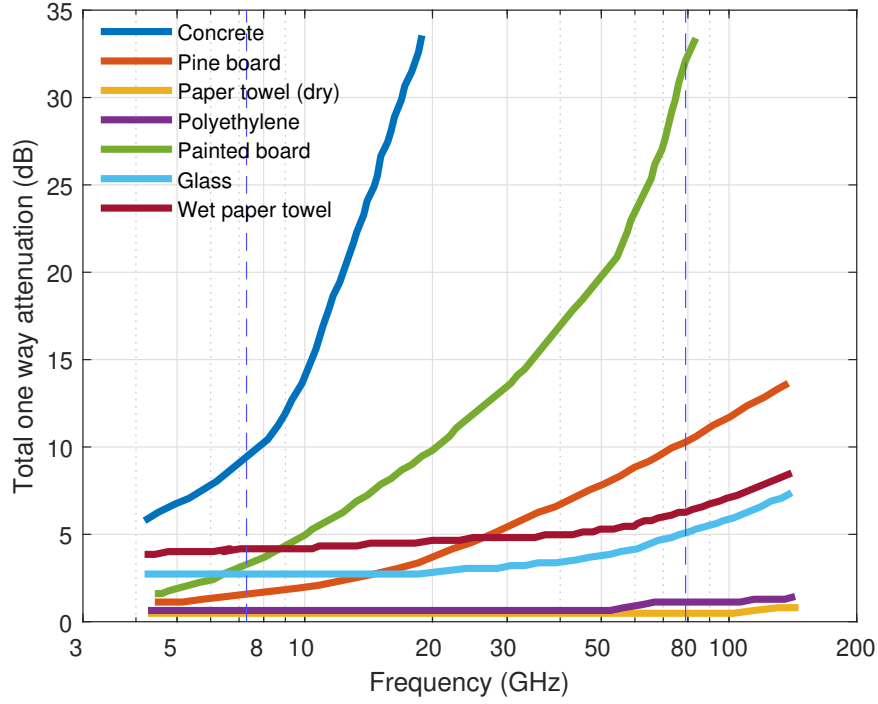
The algorithm for 60 GHz in-cabin occupant monitoring is the same as the 79 GHz based system, and also, the 60 GHz radar hardware is not available in the market now, therefore, the 79 GHz central frequency radar is selected for the system design in this research.

### 3.2.2 Attenuation and penetration

The reflection and penetration properties of materials are affected by the radar waveform. When radar signal reaches a test subject, one part of the signal is reflected towards radar as a backscattered signal to the receiver, and another part of the signal penetrates the subject. The amount of reflection and penetration varies under different materials of subject.

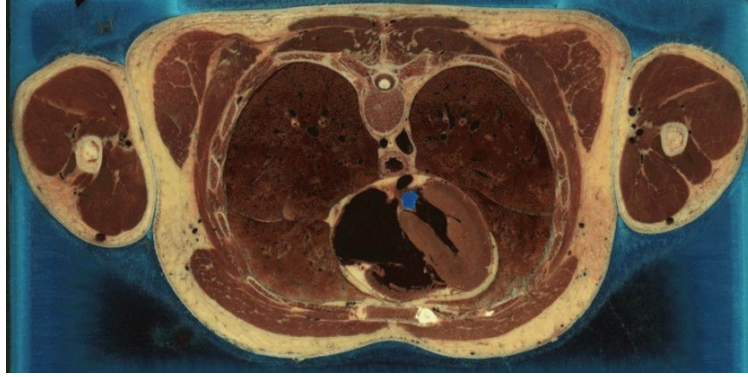
The following Fig. 3.1 is the one-way attenuation of different materials under different frequency bands. The two blue dash lines indicate the attenuation of materials in the frequency bands of interest. As one can tell from the figure, the C-band (4 to 8 GHz) and X-band (8 to 12 GHz) have good penetration ability, shown in the first blue dash line. It could penetrate through even thick material to detect moving objects behind. The V-band (40 to 75 GHz) and W-band (75 to 110 GHz), on the other hand, cannot penetrate thick material as C-band and X-band radar do. However, for materials such as glass and fabric, the V-band and W-band show relatively low attenuation, making them appropriate choice of in-cabin occupant detection radar.

Human body has multiple layers, and chest displacement is a popular use for vital sign detection; thus, thorax area of human body is then the area of interest. One example of the thorax section of a human male is shown in Fig. 3.2. The heart, lungs, spinal, blood vessels can be seen clearly. Organs in the thorax are surrounded by skin and fat, where the radar signal reaches first after being transmitted by a radar transmitter. The radar signal could have a different path through the human body when placed in a different location inside vehicle cabin.



**Figure 3.1.** One-way attenuation of radar wave under difference frequency [90]

Permittivity and loss tangent of human tissue for radars with central frequency 7.29 GHz and 79 GHz are shown in Table 3.2 [84]. Both values of permittivity and loss tangent are shown in the table. Relative permittivity, or dielectric constant, of tissue and organs of individuals reflects the ability to store electrical energy in the electric field; huge contrast between two tissues will induce more reflections at the interface. On the other hand, the loss tangent is correlated to signal attenuation [30]. A larger loss tangent will induce higher signal attenuation.



**Figure 3.2.** Section of thorax of a male [91]

**Table 3.2.** Dielectric constant and loss tangent of thorax related tissue under 7.29 GHz and 79 GHz radars

	7.29 GHz		79 GHz	
Tissue name	Dielectric constant	Loss tangent	Dielectric constant	Loss tangent
Aorta	36.23	0.4067	7.882	1.172
Blood	49.88	0.4325	9.826	1.382
Blood Vessel	36.23	0.4067	7.882	1.172
Body Fluid	62.57	0.3694	8.679	1.975
Bone Cancellous	14.3	0.4779	4.491	0.7455
Bone Cortical	9.057	0.4117	3.503	0.5192
Bone Marrow	4.829	0.1969	2.989	0.2449
Breast Fat	4.248	0.3162	2.619	0.1536
Fat	4.822	0.2012	2.988	0.2453
Heart	46.52	0.4176	9.719	1.317
Lung Deflated	41.67	0.3907	9.05	1.248
Lung Inflated	17.66	0.3903	4.532	1.008
Muscle	46.47	0.363	10.27	1.293

*continued on next page*



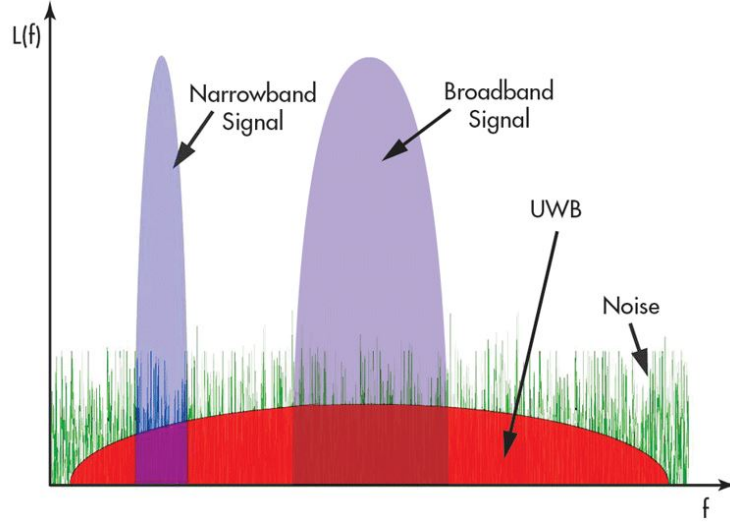
**Table 3.2.** *continued*

	7.29 GHz		79 GHz	
Tissue name	Dielectric constant	Loss tangent	Dielectric constant	Loss tangent
Nerve	25.98	0.3781	6.948	0.9501
Oesophagus	53.79	0.3972	10.66	1.402
Skin Dry	33.83	0.3719	6.45	1.352
Skin Wet	36.79	0.3951	8.391	1.174
Spinal Cord	25.98	0.3781	6.948	0.9501
Thymus	49.6	0.3915	10.1	1.355

Table 3.2 shows that the permittivity of tissues for 79 GHz radar is smaller than that of the 7.29 GHz radar while the loss tangent value of tissue for 79 GHz is higher than that for the 7.29 GHz radar. It means that radar with a central frequency of 79 GHz absorbs less energy and has higher attenuation than 7.29 GHz, which is in compliance with the tendency shown in Fig. 3.1. The heartbeat signal measured by radar with a central frequency of 79 GHz is weaker than that measured by radar with a central frequency of 7.29 GHz, when the two radars have the same output power level. In general, using radar with either central frequency 7.29 GHz or 79 GHz, the radar signal path will try to avoid tissue with high loss tangents such as blood, body fluid, esophagus, and Thymus to reach the heart.

### 3.2.3 Waveform transmission

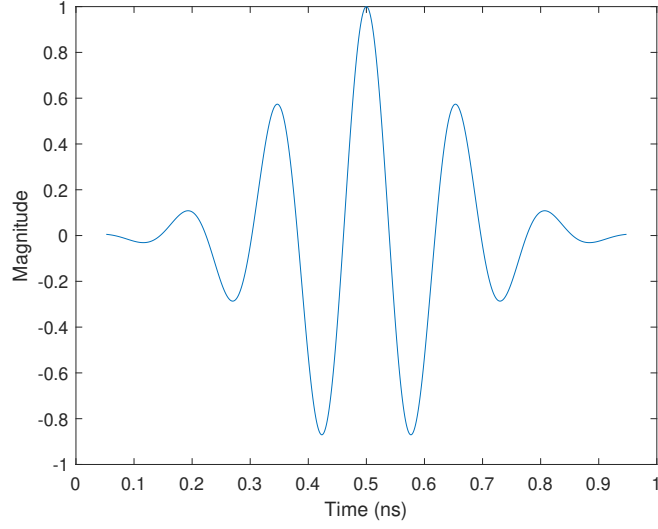
One advantage of UWB radar is its waveform leads to a low average transmit power over the frequency span. Because the transmitted radar waveform has a wide bandwidth, the output power is distributed over the whole frequency band, making the average power of the entire frequency band much lower than the radar waveform with narrowband signal and broadband signal. As shown in Fig. 3.3. Y-axis stands for the power spectrum density, and X-axis represents the frequency bandwidth.



**Figure 3.3.** UWB radar power spectrum density distribution [92]

For radar with their frequency band in C-band, X-band, or W-band, two waveform types are often used. Impulse radio-based radar transmits out pulse (often used Gaussian modulated pulse) signal continuously as in Fig. 3.4 and measures the reflected signal for object detection and movement detection. Impulse signal is often used for C-band, X-band for indoor sensing, and through-wall sensing. However, some research or device also chooses to use the FMCW radar as in Fig. 3.5 [21]. For W-band radar-equipped on the vehicle, most of the research chooses FMCW as the signal waveform.

The benefits of using impulse radio based waveform are twofold. Because the pulse length of impulse radio-based radar is less than 1 ns, and will wait for around 40 ns to collect all the signals. The average power between pulses is low. On the other hand, more signals will be received from the detected object due to the wide frequency band. When the transmitted signal arrives at the surface of the test subject, a signal with different frequencies will have different reflection ability. It is related to the signal wavelength and the smoothness of the surface. If the radar waveform wavelength is larger than the concave region of the surface, this region will not act as a corner reflector. The signal will be scattered away and no longer be detected by the receiver antenna; if the waveform wavelength is less or equal to the concave region, the transmitted signal will be backscattered to the radar antenna.



**Figure 3.4.** Impulse radio waveform

FMCW waveform is shown in Fig. 3.5. FMCW based radar detects the object by comparing the difference between the received frequency with the reference signal to cover the whole detection range. On the other hand, the impulse radio-based radar has a “blind spot” based on Eq. (3.1). It must wait until the end of the impulse transmission to start measuring the test subject.

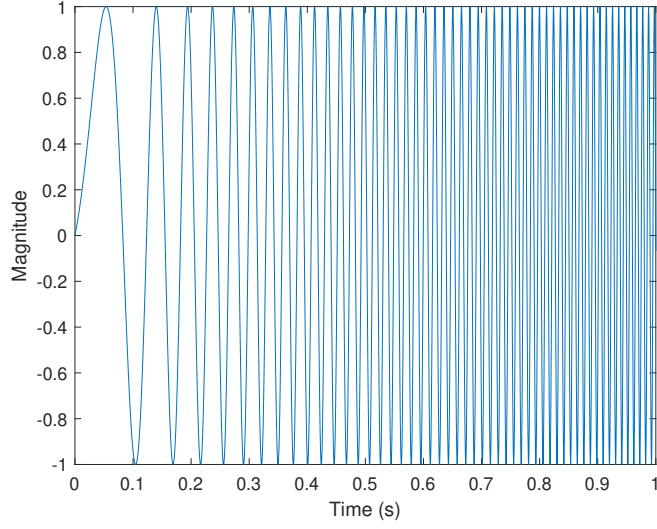
$$d = c \frac{t}{2} \quad (3.1)$$

where  $d$  is the distance from the subject to the radar antenna,  $c$  the speed of light,  $t$  is the time delay from the beginning of impulse signal transmission.

#### 3.2.4 Range resolution

Radar waveform with an ultra-wide frequency band will increase the range resolution for distance detection. The fractional bandwidth defines the ultra-wideband as in Eq. (3.2) [93].

$$BW_f = \frac{2(f_h - f_l)}{f_h + f_l} \geq BW_{lb} \quad (3.2)$$



**Figure 3.5.** FMCW chirp waveform

where  $BW_f$  is the fractional bandwidth and  $f_h$  and  $f_l$  represent the highest and the lowest frequencies of the signal, respectively.  $BW_{lb}$  is the lower bound of the bandwidth defined as 0.2 with a cut-off level of  $L_{cf} = -10dB$  [93]. Also, there is another definition by limiting the minimum absolute bandwidth to 500MHz.

The IR radar with wide bandwidth has finer range resolution compared to the radar with narrow bandwidth. The range resolution of an impulse-based radar depends on the signal scattered from the subject of interest. Impulse radar with a wide bandwidth signal enhances the range resolution by backscattering more signals than impulse radar with narrow bandwidth. The surface of a test subject, such as a human subject, contains lots of wrinkles. Therefore, as discussed previously, by increasing the frequency span, more signals with different frequencies can be received by the receiver antenna, leading to high resolution eventually.

For FMCW radar, the range resolution is directly related to the length of bandwidth. Since FMCW-based radar detects the subject by comparing the frequency difference between the received signal with the transmitted one, the longer the frequency span is for the FMCW chirp signal, the better the range resolution of radar can be achieved.

Micro-movement detection, such as human subject vital sign detection, requires accurate measurement of chest displacement. Typically, the displacement of the chest is in mm level. Therefore, to achieve this accuracy, a specific range of resolution needs to be achieved.

For chest displacement, the movement is in a range from 1mm to 12 mm; heart displacement, on the other hand, is in a range of 0.2 mm to 0.5 mm. For IR-based radar, samples collected for the range profile calculation are essential. If the range resolution is set to be 6 mm, then the sampling rate for 3 meters of cabin length should be at least 50 Gs/s to have enough samples as the following equations.

$$samples = \frac{3}{0.006} = 500samples \quad (3.3)$$

$$time = \frac{3}{3 \times 10^8} = 10^{-8}s \quad (3.4)$$

$$sampling\_rate = \frac{samples}{time} = 50Gs/s \quad (3.5)$$

For FMCW waveform-based radar, the micro-movement detection differs from that of the impulse-based waveform. It could be calculated by Eq. (3.6) that the range resolution is only 4 cm.

$$d_r = \frac{c}{2B} \quad (3.6)$$

where  $d_r$  represents the range resolution,  $c$  is the speed of light, and  $B$  is the bandwidth.

The micro-movement is detected by evaluating the phase difference between transmitted and received signals. This is valid for the micro-movement estimation. Assume the subject did not move during the whole assessment period, and the range bin in the range profile remained the same, then the instantaneous frequency (IF) of the FMCW chirp would be a constant value. Since FMCW radiates a sinusoidal waveform, and the detected subject's initial phase value could be estimated, then the phase change of that range bin can be continuously monitored. By recording the phase change of a subject unmoved, the micro-movement of subject is known.

### 3.2.5 Antenna size

Single antenna size is related to the central frequency. Typically the size of a radar antenna is one-quarter of the wavelength of the central frequency [94][95]. For an antenna array, each receiver's distance is typically half-length of the central frequency wavelength [96]. The size of a single radar antenna and the minimum antenna distance between receivers for antenna array are estimated with multiple central frequencies shown in Table 3.3.

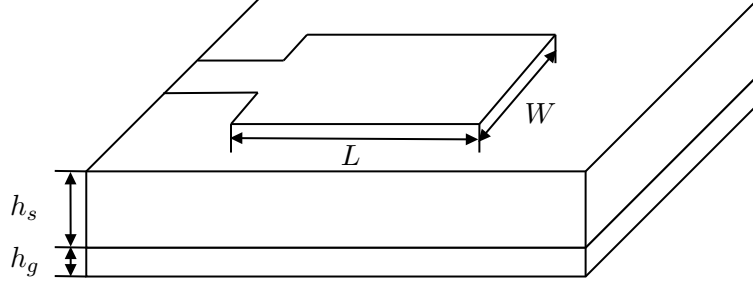
**Table 3.3.** Antenna size and minimum antenna distance

Central frequency (GHz)	Wavelength (m)	Antenna size (m)	Min antenna distance (m)
10	0.03	0.0075	0.015
24	0.0125	0.00312	0.00625
60	0.005	0.00125	0.0025
79	0.00379	0.00094	0.00189

For occupant detection in vehicle cabin, a radar antenna array with high central frequency requires a smaller space inside vehicle and thus provide more flexibility for deployment. As can be seen from Table 3.3, both the antenna size and minimum distance between the receiving antennas decrease with increasing of central frequency. Specifically, both the antenna size and minimum antenna distance of radar with 79 GHz central frequency is reduced by 87% compared to that for radar with 10 GHz central frequency. Therefore, the total size of radar with antenna array will be reduced by 87% if with the same number of the antenna array, which provides more choices of mounting position for the antenna array inside the cabin.

### 3.2.6 Patch antenna

For 79 GHz radar with MIMO configuration, the antenna can be designed using the patched micro-strip antenna, shown in Fig. 3.6. In this 3D figure of a single microstrip antenna,  $L$  stands for the antenna's length,  $W$  stands for the patch antenna's width,  $h_s$  stands for the height of the antenna substrate, and  $h_g$  is the height of the ground plane.



**Figure 3.6.** Micro-strip antenna configuration

To determine the length  $L$  and the width  $W$  of each microstrip antenna, both the effective dielectric constant  $\epsilon_{eff}$ , extended length  $\Delta L$  need to be evaluated according to substrate height  $h_s$  and central frequency. Balanis gives the calculation of effective dielectric constant [97] as Eq. (3.7)

$$\epsilon_{eff} = \frac{\epsilon_r + 1}{2} + \frac{\epsilon_r - 1}{2} \left[ 1 + 12 \frac{h_s}{W} \right]^{-\frac{1}{2}} \quad (3.7)$$

where  $\epsilon_{eff}$  is the effective dielectric constant,  $\epsilon_r$  is the dielectric constant of the substrate.

The extended length can be referred to empirically by Hammerstad [98] in Eq. (3.8)

$$\Delta L = 0.412h \frac{(\epsilon_{eff} + 0.3) \left( \frac{W}{h_s} + 0.264 \right)}{(\epsilon_{eff} - 0.258) \left( \frac{W}{h_s} + 0.8 \right)} \quad (3.8)$$

The length of one patch can then be calculated by Eq. (3.9)

$$L = L_{eff} - 2\Delta L \quad (3.9)$$

The effective length of one patch antenna  $L_{eff}$  is given by [99] Eq. (3.10). In this equation,  $c$  is the speed of light, and  $f_0$  is the central frequency.

$$L_{eff} = \frac{c}{2f_0\sqrt{\epsilon_{eff}}} \quad (3.10)$$

The width is calculated [100] as in Eq. (3.11)

$$W = \frac{c}{2f_0\sqrt{\frac{(\epsilon_r+1)}{2}}} \quad (3.11)$$

A patch antenna design example using parameters from Texas Instrument's AWR1843 BOOST can be shown in Table 3.4.

**Table 3.4.** RF antenna specs of AWR1843BOOST

Parameter	Symbol	Value	Unit
Dielectric constant	$\epsilon_r$	3.66	F/m
Substrate height	$h_s$	0.1016	mm

The AWR1843BOOST has a central frequency 79 GHz with dielectric constant (also called relative permittivity) of substrate  $\epsilon_r$  and height of substrate  $h_s$ . By applying the aforementioned Eq. (3.7) - (3.11), the length L and width W can be calculated as 0.95 mm and 1.24 mm, which match the measurement from the hardware.

### 3.2.7 Single pair antenna

For a single antenna approach, ultra-wideband impulse radar (UWB-IR) is often selected. The advantage of UWB-IR is its high range accuracy and short pulse repetition time (PRT). The radar transmits very short impulse wave signals. Gaussian impulse or modulated Gaussian impulse is typically used as the impulse signal to increase bandwidth to achieve a better range resolution. The duration of the impulse signal is much shorter compared to FMCW. Compared to the duration of chirp signals of FMCW - several microseconds, the signal duration of IR is shorter - within one nanosecond. Using the radar to locate object needs to use data from both the fast time and the slow time.

Sensor is often located in front of test subjects, or faces them. A single-pair antenna radar is typically placed away from boresight to achieve better subject detection. It can be placed at a vehicle's A-pillar to make sure that the distance from each subject to the antenna is different. In most of the published research for this scenario, algorithms are applied to

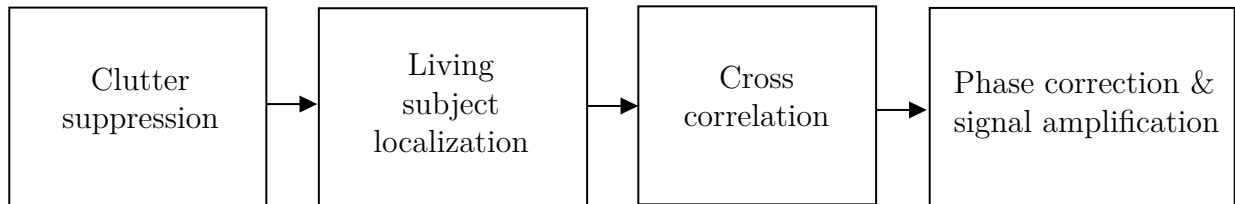


data collected from ideal scenarios and/or short distances where SNR is large. This makes the position and vital signs of the human subject relatively easy to be located. However, in a practical environment, noise signals can be much larger in some scenarios, such as metal reflection. Occlusion can also increase difficulties in vital sign signal detection of human subjects sitting behind when most signals are reflected and attenuated by subjects in the front. One alternative way to sense vital signs is to embed radar into seats in this strong reflection scenario [30]. However, it requires one radar system per seat. If radar systems are too far away from occupants, reflected signals can be weak and noisy signals may be buried in background noises. The concept system setup for single pair antenna based occupant detection is shown in Fig. 3.7.



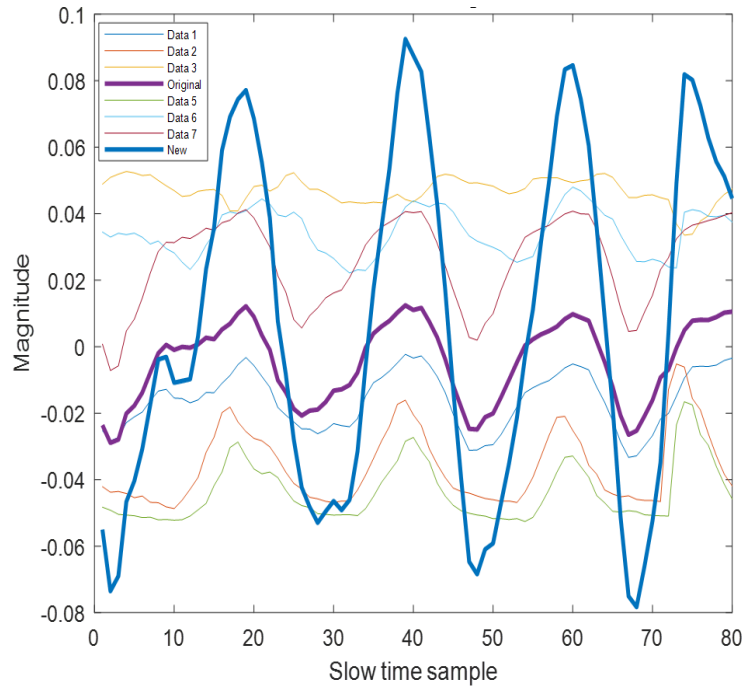
**Figure 3.7.** System setup for two subjects

One approach is to apply a cross-correlated based method, which consists of four different parts [101]. This approach includes four steps: clutter suppression, human subject localization, cross-correlation, and phase correction and signal amplification. The flowchart of this method is shown in Fig. 3.8.



**Figure 3.8.** Weak signal detection method flowchart

By applying the setup in Fig. 3.7 and the method mentioned above, the vital sign signal can be extracted with high SNR as shown in Fig. 3.9. In Fig. 3.9, a signal with legend named Original is the signal before enhancement; Data 1-3 and 5-7 are signals in adjacent bins of the original signal; a signal with legend New is the enhanced output signal. From this figure, it can be observed that all sinusoidal waves have been lined up and added up for amplification. The magnitude of each sinusoidal signal has been added up while the signal's frequency has been maintained. The SNR increases by 46.7% from the original signal. Since the cross-correlated signal also has its oscillation around zero, no DC tendency can affect FFT computation in the following steps. The reason is that the waveform that has been transmitted out has a Gaussian-like oscillation pattern. In contrast to the original signal, the adjacent signal's DC component will cancel each other out once they are added up together. It can also be seen that there will be a rather flat signal, such as Data 3 in Fig. 3.9 added to the original one, which may result in more noise in the original signal. This can be fixed by applying a bandpass filter.



**Figure 3.9.** Signal amplification for system setup

### 3.2.8 Antenna array

The proper antenna array design increases the gain and angular resolution to locate the test subject in a high dimension. With more receiver antennas in an antenna array, the radar receives more signals with a certain time delay. By synchronizing and summing up all received signals, the total amplitude of the signal will increase. The antenna array requires at least one transmitter and two receiver antennas to estimate the AoA for subjects in the same range of antenna. In principle, the AoA is estimated by calculating the phase delay from the same transmitter. It is calculated as in Eq. (3.12) [96].

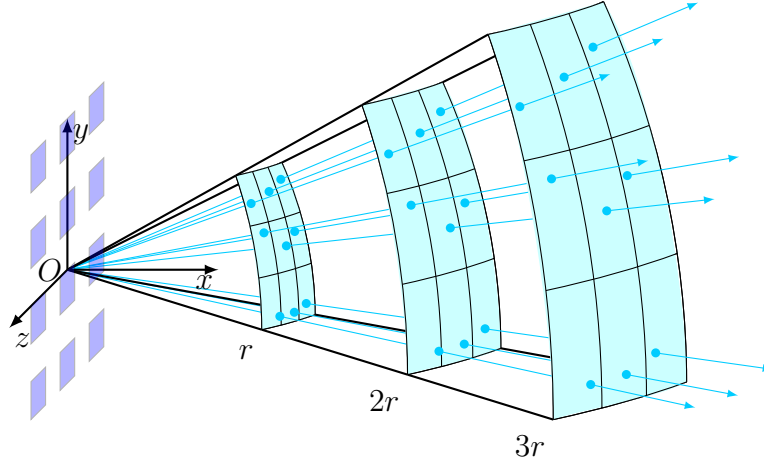
$$\theta_{res} = 0.89 \frac{\lambda}{D} \approx \frac{2}{N_{Tx} N_{Rx}} \quad (3.12)$$

where  $\theta_{res}$  is the angular resolution,  $\lambda$  is the wavelength,  $D$  is the equivalent aperture, and  $N_{Tx}$ ,  $N_{Rx}$  are numbers of transmitter and receiver, respectively.

A multiple input multiple output (MIMO) structure-based antenna array can effectively reduce size of antenna array and thus further reduce space occupied inside a vehicle. Compared to synthetic aperture radar, MIMO radar requires less data acquisition and processing time achieve the same cross-range resolution [102]. The MIMO radar equivalents the single input multiple output (SIMO) structure by increasing the number of transmitters. The transmitters are divided by a fixed distance depending on the central frequency of signal to represent the difference of phase delay in antenna array. Transmitters radiate signals in sequence while multiple receivers receive the signal as a group and repeat per transmitter. Nevertheless, using the MIMO radar array for in-cabin occupant detection takes a long time and needs more memory for processing; it will need to process a data cube instead of a matrix.

Although the angular resolution increased with the size of aperture, the angular resolution is also affected by the distance between the target and the origin of the antenna array. A sketch of a real aperture of the antenna array radar is shown in Fig. 3.10. The real aperture of the antenna array angular resolution decreases along the radial axis, meaning the size of the angular pixel shown in the color of cyan in Fig. 3.10, increases with range. Specifically, for a set of point targets in range  $r$ ,  $2r$ ,  $3r$  as shown in Fig. 3.10, it can be seen that the

angular resolution decrease with the increase of distance between the target and the array origin  $O$ . As a result, the size of point target from detection increase with distance between target and origin.



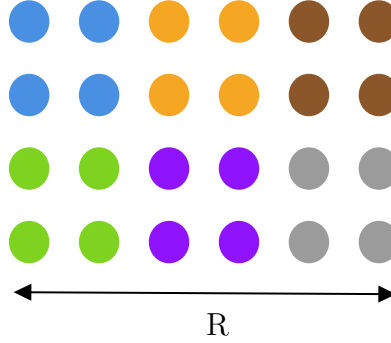
**Figure 3.10.** Real aperture radar antenna array

### 3.2.9 Antenna subarray

For high-resolution monitoring, the projection of the angular pixel onto the area of interest should be fine enough to separate two or more adjacent targets. It is achievable by adding adequate antennas in the array. However, more antennas would cost computational power. Also, since subjects in cabin do not change seats frequently, it is not necessary to scan all antennas in the array each time if a vehicle is not fully occupied. Instead, we should only focus on the area of occupied seats.

An approach based on a subarray-based concept is proposed and shown in Fig. 3.11. The dots in this figure are the equivalent antenna in the antenna array. The key to this approach is to divide the full antenna array into several subarrays. Each subarray corresponds to a zone inside a vehicle. Each subarray in the plot is marked in a different color. This figure is modeled on a five-seat vehicle. The antenna array is divided into two rows to cover both the front seats and the rear seats. For simplicity, each row is designed to have three different subarrays to cover three different zones inside the cabin. Each zone is assumed square, so subarray needs an equivalent number of antennas on horizontal and vertical axes. The

method consists of two parts, occupancy detection and beamforming. The first part of this method is to process the received data cube by using defined subarrays. If there is no signal detected in the defined zone, the zone will be excluded in the second step. An additional step will be applied for those zones with detected signals to generate high-resolution reflections by using beamforming with all antennas in the antenna array.



**Figure 3.11.** Subarray based occupant monitoring method

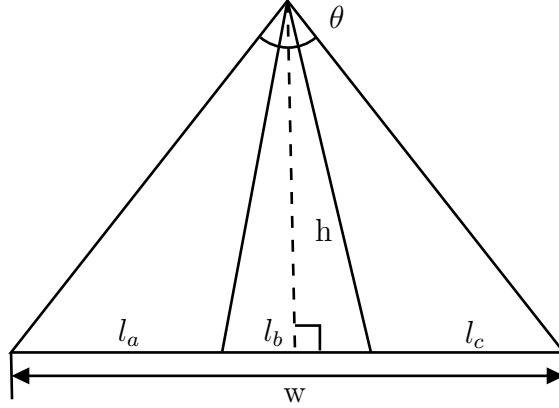
It is assumed that the total number of antennas in horizontal dimension in Fig. 3.11 as  $R$ , the number of subarray is  $M$ , and each of them has  $N$  antennas. For example, in Fig. 3.11,  $M$  is three and  $N$  is two. These three parameters yield

$$R = M \times N \quad (3.13)$$

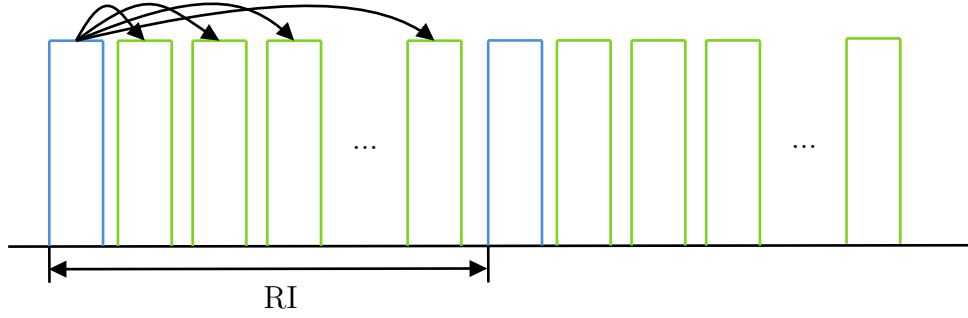
Two conditions need to be considered in selecting an optimal numbers of antennas for each subarray, i.e., the maximum length of partition and total times of the computation. The maximum length of partition is determined by the width or the height of the detection field. Fig. 3.12 presents a simplified 1-D model for the back row for example.

where  $\theta$  is the field of view (FoV) of the radar,  $w$  is the total width of the rear row,  $h$  is the vehicle cabin's height. The three seats divided the total width of the back seat into a, b, c three parts with lengths  $l_a$ ,  $l_b$  and  $l_c$ . By small angle approximation for each partition, the division of FoV can be expressed by

$$\theta_p = \frac{l_i}{w} \theta, i \in \{a, b, c\} \quad (3.14)$$



**Figure 3.12.** Partition of back row in vehicle cabin



**Figure 3.13.** Subarray based two-step scan

To detect the subject located in this partition, the value of the subarray's angular resolution should be less than  $\theta_p$ . By Eq. (3.12), we get Eq.(3.15)

$$\frac{2}{N} \frac{180}{\pi} \leq \theta_p \quad (3.15)$$

Then the computational time is assessed. Here it is assumed that the memory resource of hardware were sufficient for processing. Therefore, the total times of beamforming angles is the only focus. The traditional beamforming approach scans all the receiving antenna cells. The time of the traditional beamforming costs are,

$$T_{trad} = \frac{\theta}{\frac{2}{R}} = \frac{\theta R}{2} \quad (3.16)$$

For a subarray-based two-step scan, a mechanism of processing is shown in Fig. 3.13. The block in blue represents the time spent for occupancy detection. Since the occupant is relatively static inside the vehicle cabin, it is unnecessary to detect occupancy in each iteration. So a repetition interval  $RI$  is used to represent the repetition time of this step. After finding the occupied zone in the first step, the next is to generate high-resolution reflection signals in each zone. The blocks in green show the procedure of beamforming. Each green block represents time spent for beamforming, and the number of beamforming is determined by the zones found in the first step. For the convenience of comparison with the traditional method, the time consumed in the repetition interval is amortized into the following beamforming showed in Fig. 3.13.  $P$  out of  $M$  partition is selected after applying the first step. The next step is to use all antennas in this partition for the highest angular resolution. Thus, the total computational times are

$$\begin{aligned}
T_{sub} &= \frac{\frac{\theta}{M}}{\frac{2}{N}} M \frac{1}{RI} + \frac{\frac{\theta}{M}}{\frac{2}{R}} P \\
&= \frac{\theta}{2} \frac{N}{RI} + \frac{\theta}{2} \frac{P}{M} R \\
&= \frac{\theta}{2} \left( \frac{N}{RI} + \frac{P}{M} R \right)
\end{aligned} \tag{3.17}$$

To save computational power, the number of beamforming for subarray-based method  $T_{sub}$  should be less than the number of beamforming for the traditional based method  $T_{trad}$ , meaning the relationship between these two numbers should yield  $T_{sub} < T_{trad}$ . To get maximum upper bound for this inequality, the value of  $RI$  is set to one. Then, Eq. (3.18) can be obtained after some simple derivations.

$$\begin{aligned}
\frac{\theta}{2} \left( N + \frac{P}{M} R \right) &< \frac{\theta}{2} R \\
N &< \left( 1 - \frac{P}{M} \right) R
\end{aligned} \tag{3.18}$$

Eq. (3.18) provides the upper limit for the number of antennas in a subarray. By reorganizing Eq. (3.15), the lower bound of  $N$  is given by Eq. (3.19)

$$N \geq \frac{360}{\theta(\frac{l_i}{w})\pi} \quad (3.19)$$

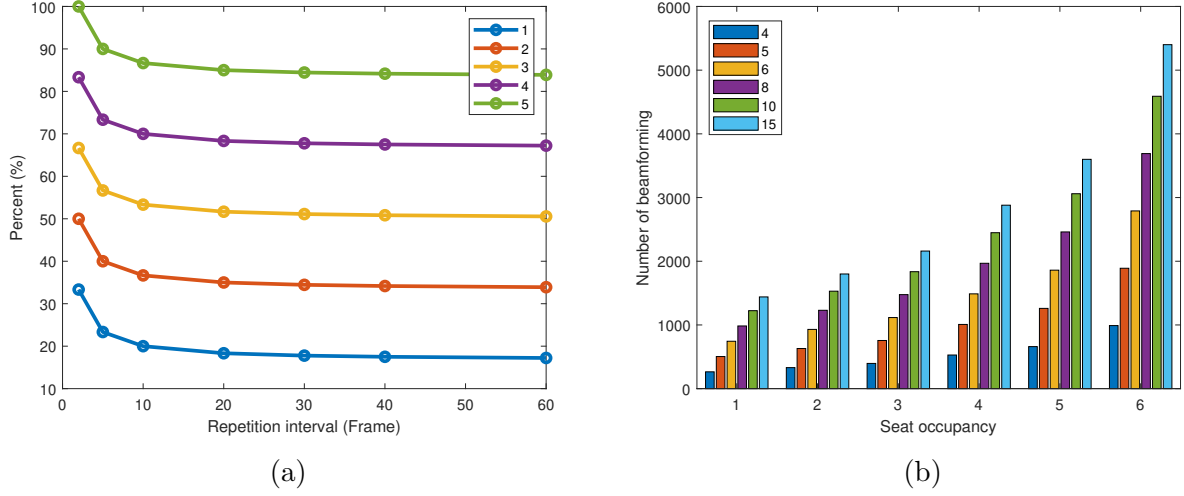
To sum up, the number of antenna in each subarray should yield Eq. (3.20). If this equation is valid,  $N_{opt}$  equals the lower bound value.

$$\frac{360}{\theta(\frac{l_i}{w})\pi} \leq N < (1 - \frac{P}{M})R \quad (3.20)$$

In this research, the goal is to determine the optimal size of a planar array that is able to cover a vehicle with five seats in two rows based on the subarray-based approach. For simplicity, the width of all five seats is set to 50 cm. The planar array is divided into six square sub-arrays, with five of them responsible for detecting the occupancy of 5 seats. After some straightforward calculations based on Eq. (3.20), the minimum number of  $R$  can be found as five, and the optimal number of  $N$  is three. Therefore, the minimum number of antenna needed for the sub-array-based occupancy detection is  $9 \times 6$ .

The selection of repetition interval is key in this proposed method to save computational power. In Fig. 3.14a, different repetition intervals are tested, and the percentage of the subarray-based beamforming numbers can be found decrease as the repetition interval increases. Lines in different colors represent different numbers of occupant inside a vehicle cabin. The decreasing tendency is similar in different occupancy scenarios. If the number of repetition interval is fixed at 20, the number of beamforming is found to change along with the sub-array size and the number of occupied seats, as shown in Fig. 3.14b. In the bar chart, the first five sets of bars show the number of beamforming using the subarray-based method; the last set of bars is the number of beamforming using the traditional method and set as reference. It can be found that the subarray-based method helps decrease the number of beamforming with different sizes of the antenna array and thus saves computational power.





**Figure 3.14.** (a) reduction of beamforming number with different value of repetition interval (b) number of beamforming for different number of occupants and size of subarray

### 3.3 The Angle of Arrival (AoA) approaches

AoA approach is widely used in determining the direction of propagation of the radio-frequency wave and estimating its magnitude [103]. The results can be used for test subject localization, tracking, and recognition in 3D space.

#### 3.3.1 Angle of Arrival (AoA): Beamforming

The beamforming collects signals from each antenna pair of the antenna array to generate optimized weights and to determine where their locations as expressed in Eq. (3.21).

$$y(t_f, t_s, i) = \sum_{l=1}^L [A_{Rl} \sqrt{G(\sigma_e)} e^{j(\phi_{rl}(t_s, t_f) + \phi_a(\phi(t_s), \theta(t_s), i))}] \quad (3.21)$$

The mechanism of the generated weights can be described using a signal model for an antenna array. Assume that there were  $L$  uncorrelated or partially correlated point target sources,  $s_l(t_s)$ . The receiving data,  $X_R(t_s)$  is comprised of the backscattered signals  $A_R s_l(t_s)$  and noise  $w(t_s)$ .  $A$  is the steering matrix for the search of target of interest. In this research,

it is assumed that the AoA in the steering matrix were known and can be expressed in the form of  $\theta_l$  for simplicity;  $\theta_l$  represents the AoA of the  $l$ -th point target source.

$$\begin{aligned} X_R(t_s) &= A_R s_l(t_s) + w(t_s) \\ s_l(t_s) &= [s_1(t_s), s_2(t_s), \dots, s_L(t_s)] \\ A &= [a(\theta_1), a(\theta_2), \dots, a(\theta_L)] \end{aligned} \quad (3.22)$$

The size of parameters mentioned in Eq. (3.22) can be described as follows. Both  $X_R(t_s)$  and  $w(t_s)$  are vectors and they have same length that equals to the number of antenna pairs  $I$ .  $s_l(t_s)$  is a  $L \times 1$  vector representing signal values from  $L$  point sources. Therefore, the size of  $X_R(t_s)$  is  $L \times I$ .  $A$  is an  $I \times L$  steering matrix calculated by Eq. (4.5). The steering matrix consists of the relative phase shifts arrived at different receiver in planar array from one target source. Each column of  $A$  represents a steering vector from one of the sources and depends on the AoA in both azimuth and elevation.

The principle of beamforming is to assign weights  $\omega_a$  to each antenna pair in the receiving signal  $X_R$ . Therefore, it has a size of  $I \times L$ . The weights could enhance the power of the main lobe and suppress the power of sidelobe in the direction of AoA. Thus, the optimized received power is described in Eq. (3.23).

$$X'_R = \omega_a X_R \quad (3.23)$$

Then the output power is calculated by Eq. (3.24).

$$P_{out} = X'_R (X'_R)^H = \omega_a^H (X_R X_R^H) \omega_a = \omega_a^H R_x \omega_a \quad (3.24)$$

where the superscript  $H$  represents the Hermitian matrix of the original matrix.  $R_x$  is the covariance matrix of the received signal matrix  $X_R$ .

The key idea of the beamforming is to determine the distribution of weights to the receiving signals. The most straightforward approach is to use the matrix  $A$  as the weights  $\omega_a$  directly. It also named conventional or Bartlett beamforming [103]. However, this method

provides a relatively low resolution compared to other methods. To achieve better angular resolution, a more adaptive AoA approach is needed.

### 3.3.2 Angle of Arrival (AoA): Subspace-based approach - MUSIC

The MUSIC is a specific algorithm in subspace-based approaches based on the eigenvalue decomposition of the covariance matrix observed at an array [104]. It decomposes the radar covariance matrix  $R_x$  into two subspaces, a signal subspace and a noise subspace.

$$R_x = E\{X_R X_R^H\} = A R_s A^H + \sigma_n^2 I \quad (3.25)$$

where the first item on the right of Eq. (3.25) is the signal subspace and the second item on the right is the noise subspace.  $R_s$  is the source covariance matrix. The source covariance matrix's diagonal elements represent source power, and the off-diagonal elements represent source correlations.  $R_s$  can be expressed by signal vector  $s$  as

$$R_s = E\{s s^H\} \quad (3.26)$$

The algorithm assumes that the noise powers were equal at all sensors and uncorrelated between sensors. Thus, the noise covariance matrix became an I-by-I diagonal matrix with equal values along the diagonal.

It is noted that for uncorrelated sources or even partially correlated sources,  $R_s$  is a positive-definite Hermitian matrix and has full rank  $L$ , which is equal to the number of target sources. The signal covariance matrix,  $A R_s A^H$ , is an I-by-I matrix, also with rank  $L < I$ .

Because the true sensor covariance matrix is unknown in the real-world scenarios, the MUSIC algorithm can take the sampled sensor covariance matrix  $R_x$  for the calculation. The sampled sensor covariance matrix is an average of multiple reflected chirps from the sensor data.  $K$  is the number of reflected chirps.

$$R_x = \frac{1}{K} \sum_{k=1}^K X_R(t_s) X_R(t_s)^H \quad (3.27)$$

If the number of target sources is known, the MUSIC algorithm is able to provide better spatial resolution than other algorithms of AoA. When the number of sources specified is incorrect, MVDR and Beamscan may show insignificant peaks from the correct spatial spectrum. In contrast, the MUSIC spatial spectrum may be inaccurate when the number of sources is not specified correctly. Also, the amplitudes of MUSIC spectral peaks cannot be interpreted as the power of the sources [105].

The derivation of the algorithm starts with equation (3.25). Assume  $A$  and  $R_s$  had rank of  $L$ , both  $A$  and  $R_s$  would then have  $L$  positive real eigenvalues and  $I - L$  zero eigenvalues. Next, these eigenvalues were reorganized in descent sequence with  $\mu_1 \geq \mu_2 \geq \dots \geq \mu_L > 0$  in signal subspace. When noise is introduced, a noise power  $\sigma_0^2$  needs to be added to each of the eigenvalue in both subspaces. To sum up, the eigenvalue for signal and noise subspaces are given by

$$\lambda_k = \begin{cases} \mu_d + \sigma_0^2 & l = 1, 2, \dots, L \\ \sigma_0^2 & l = l + 1, l + 2, \dots, I \end{cases} \quad (3.28)$$

The corresponding eigenvectors can also be expressed as

$$e_l = \{\underbrace{e_1, e_2, \dots, e_L}_{\text{signal}}, \underbrace{e_{L+1}, \dots, e_I}_{\text{noise}}\}$$

The eigenvectors in the first bracket correspond to the positive eigenvalues that span the signal subspace, while those in the second bracket correspond to the zero eigenvalues that span the null subspace and are orthogonal to the signal space. Thus,  $R_s$  can be presented as

$$R_s = \sum_{l=1}^L \lambda_l e_l e_l^H + \sum_{l=L+1}^I \lambda_l e_l e_l^H \quad (3.29)$$

When  $l \geq L$ , the first item on the right of the Eq. (3.29) goes to zero. Multiply  $e_l$  to both sides, we get

$$R_s e_l = \lambda_l e_l = \sigma_0^2 e_l = 0 \quad (3.30)$$

By further multiplying  $e_l$  to both side of Eq. (3.25), we get

$$AR_s A^H e_l = 0, \quad \forall l > L \quad (3.31)$$

If multiply  $e_l$  to the above equation, we further get

$$\begin{aligned} e_l^H AR_s A^H e_l &= 0 \\ (A^H e_l)^H R_s (A^H e_l) &= 0 \\ A^H e_l &= 0 \end{aligned} \quad (3.32)$$

From the above derivation, those arrival vectors are orthogonal to the null space, which is a key foundation of MUSIC.

The mechanism of MUSIC is to search all vectors in the steering matrix that are orthogonal to the noise subspace. Therefore, MUSIC constructs an AoA dependent pseudo-power expression, called the MUSIC pseudo-spectrum  $P_{music}$  as in Eq. (3.33).

$$\begin{aligned} P_{music}(\theta) &= \frac{1}{a^H(\theta) (\sum_{d=L+1}^I e_l e_l^H) a(\theta)} \\ &= \frac{1}{\sum_{d=L+1}^I |a^H(\theta) e_l|^2} \end{aligned} \quad (3.33)$$

If there is a vector orthogonal to the noise subspace, the pseudo spectrum peaks are infinite. However, because there is noise in  $X_R$  and the true covariance matrix is replaced by the sampled covariance matrix, the arrival vectors are not exactly orthogonal to the noise subspace. Then, the angles at which  $P_{music}$  have finite peaks are the desired arrival directions. Because the pseudo-spectrum can have more peaks than the actual target sources, the algorithm requires specifying the number of desired target sources,  $L$ , as a parameter. The algorithm is responsible for picking the  $L$  largest peaks by searching the space of interest. For the planar array, the search space is a two-dimensional grid of azimuth and elevation angles.

### 3.3.3 Angle of Arrival (AoA): Adaptive beamforming - MVDR

In the real-world, the number of signal sources is usually unknown because the number of signal subspaces and noise subspaces can not always be separated correctly. Extra targets will appear in the set of detected points; those extra points are also called ghost targets. In some situations, the covariance estimation of  $R_x$  is singular or near-singular, meaning the matrix is not positive definite. It will cause numeric stability issues in doing the inversion for AoA spectral estimation. Thus, to avoid ghost points and singularity, the minimum variance distortionless response (MVDR), an adaptive beamforming algorithm is introduced to the real-world scenario and its equation is shown as follow,

$$\hat{P}_{MVDR} = \frac{1}{a(\theta)^H [R_x + \alpha_n I]^{-1} a(\theta)} \quad (3.34)$$

The diagonally loading factor  $\alpha_n$  may be determined using the noise variance estimation as follow:

$$\alpha_n = \beta P_n \quad (3.35)$$

where  $\beta$  is small constant number and  $P_n$  is defined as  $\frac{1}{N_\alpha} \text{trace}(R_x)$  where  $N_\alpha$  is the number of virtual antenna pair under test.

## 3.4 Radiation regulations and compliances

Related radiation regulations will be presented. Then, the system performance will be assessed through theoretical analysis using parameters from the simulation system.

### 3.4.1 Radiation regulations

There are two standards currently available to assess the effect of the radiation power on health. One is called effective (or equivalent) isotropic radiated power (EIRP), and the other is called specific absorption rate (SAR) in the electric field. They measure the safety of radar from two sides.

EIRP is an indicator in evaluating the power transmitted from the antenna side. It calculates the effective radiation power in the maximum radiated signal direction [106]. The direction of the maximum radiated pattern is from the direction of the main lobe of the radar antenna array. The EIRP is determined by two factors, the radiated power from the transmitter and the antenna gain of the radar. It follows the Eq. (3.36) to calculate EIRP in Watt.

$$EIRP = P_T * G_T \quad (3.36)$$

$P_T$  is the transmitter's radiation power, and  $G_T$  stands for the transmitter's antenna gain.

SAR is another indicator in assessing the radiation power received by human body. It measures energy absorption rate of materials when they are exposed in a radio frequency electromagnetic field radiated by radar [107]. It is widely used in the consumer electronics industry as an indicator of radiation power effect on health. There is a mandatory requirement of SAR from the FCC and International Commission on Non-Ionizing Radiation Protection (ICNIRP) for protecting the human body's health. SAR is calculated as in Eq. (3.37).

$$SAR = \frac{\delta |E|^2}{m_d} \quad (3.37)$$

where SAR is in W/Kg,  $\delta$  is the electrical conductivity of human tissue in S/m, and  $m_d$  is the mass density of the human body.  $E$  stands for the electrical field root mean square (RMS), it measuring the effective magnitude of the electrical field. The RMS value of the electrical field is calculated by,

$$E = \sqrt{\frac{G_T P_T Z_0}{4\pi d^2}} \quad (3.38)$$

$G_T$  and  $P_T$  are the antenna's gain and radiation power, as in the previous equation.  $d$  is the distance from the antenna to the object.  $Z_0$  is the free space impedance, which is a product of magnetic constant and speed of light in free space, and it has a constant value of  $120\pi$  in  $\Omega$ .

The upper limits, 55 dBm peak radiation power and 50 dBm/MHz average radiation power, are for the radar from 76 GHz up to 81 GHz [108], Nevertheless, it did not specify the use scenarios for radar within this bandwidth. On the other hand, the upper limit of the EIRP density mask is -51.3 dBm/MHz for 79 GHz radar [109]. Therefore, -51.3 dBm/MHz is selected as the desired upper limit for this research.

FCC requires that the SAR in an uncontrolled environment is less than 1.6 W/kg for spatial peak and 0.08 W/kg for the whole body in a six-minute time domain [110]. ICNIRP guidance requires that power density of the body surface for the general public should be less than  $10W/m^2$  [111]. In Europe, the International Electrotechnical Commission (IEC) sets an average SAR limit of 2 W/kg over the 10 g of tissue, which absorbs most of the energy inside the radar signal field. Noted that this rule is for RF signal with a central frequency under 6 GHz. As discussed in the previous part, the radar signal with a central frequency under 10 GHz and above 10 GHz affects the human body differently. It is known from the World Health Organization (WHO) that, for radar with a central frequency above 10 GHz, power density above  $1000W/m^2$  will produce adverse health effects for human tissues. Therefore, to ensure the overall safety for 24-hours occupant monitoring using a wideband radar with the central frequency of 79 GHz, SAR should be less than 0.08 W/kg, or the power density of body surface is less than  $10W/m^2$ .

### 3.4.2 Radiation compliance

To build a radar sensing simulation system, the parameters of the simulation system need to be set similar to an actual system, which should comply with regulations. Some basic parameters for the radar sensing system are listed in Table 3.5.

As discussed in the previous section, most radiation energy absorption is stored in the skin for radar with central frequency of 79GHz. RMS electrical field can be calculated according to the value of Table 3.5 and Eq. (3.38), where the peak power of Tx has been convert from  $dBm$  to  $W$ .

$$E = \sqrt{\frac{9 \times 0.00355 \times 120\pi}{4\pi \times 3^2}} = 0.326V/m \quad (3.39)$$



**Table 3.5.** Basic parameters for a simulation system

Parameter	Value	Unit
Center frequency (f)	$7.90 \times 10^{10}$	Hz
Bandwidth (B)	$4.00 \times 10^9$	Hz
Tx number ( $N_{Tx}$ )	16	
Rx number ( $N_{Rx}$ )	16	
Tx gain ( $G_T$ )	9	dBi
Rx gain ( $G_f$ )	9	dBi
Tx peak power ( $P_T$ )	5.5	dBm
Distance (d)	3	m
Chirp duration( $T_c$ )	50	us
Frame time ( $T_f$ )	50	ms

Therefore, the assessed result of SAR is as following according to Eq. (3.37)

$$S = \frac{38.18 \times 5.48^2}{1000} = 0.013W/kg \quad (3.40)$$

where, The electrical conductivity  $\delta$  is set as 38.18 S/m for dry skin while the human density is set as  $1000kg/m^3$  [112].

According to the result, SAR is below the spatial peak requirement and satisfies the requirement of whole body, which is 0.08 W/kg mentioned in the previous part. This means that the design in this work meets the requirement of SAR to ensure the safety of the subject inside the vehicle cabin.

If radiation level is evaluated by power density on the surface, it is estimated as,

$$S = \frac{EIRP}{4\pi \times d^2} = \frac{28.2}{113.04} = 0.25mW/m^2 \quad (3.41)$$

It can be seen that the calculated power density value on the surface  $S$  is much lower than the standard set by ICNIRP.

For EIRP, it depends on both the transmitter's radiation power and the frame time set to calculate one frame. EIRP can be estimated by Eq. (3.36) and data from Table 3.5 as 28.2 mW, and the average power over the frame needs to be calculated as in Eq. (3.42).

$$P_{avg} = EIRP * \frac{T_c}{T_f} = 0.0028mW \quad (3.42)$$

Finally, by dividing the average power over the bandwidth, the power density power of the frequency band can be obtained as in Eq. (3.43)

$$EIRP_{density} = \frac{P_{avg}}{B} = -61.55dBm/mHz \quad (3.43)$$

By comparing the density value in Eq. (3.43) with the requirement for indoor sensing as -51.3 dBm/MHz, the density is found to be lower than the requirement. It can be seen from the previous calculation that the EIRP density is affected by the frame time. In this research, the frame time should be shorter than a specific upper limit, according to the requirement of vital sign signal detection. To detect occupants' vital signs, the sampling rate needs to be at least more than 8 Hz. Apparently, increasing the frame time for this case will impact the sensing capability of the occupant's micro-movement in-cabin. In this example, the frame time is 50 ms, which is 20Hz, much higher than the minimum sampling frequency.

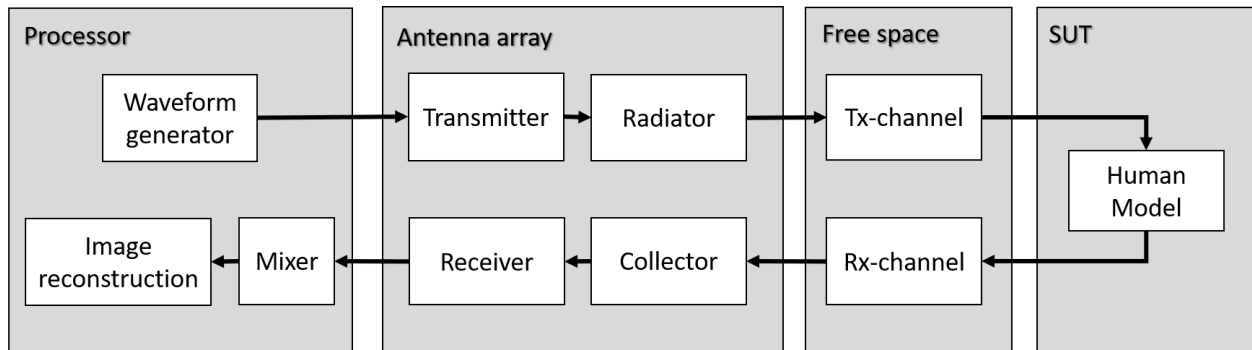
## 4. VEHICLE IN-CABIN RADAR SYSTEM SIMULATION

### 4.1 Simulation platform

The simulation system is built on the MATLAB environment. In this chapter, the radar simulation system design will be introduced and the flow and methods of radar signal processing will be discussed. The system will be used for vehicle in-cabin human subject detection and classification.

#### 4.1.1 System design

Fig. 4.1 shows the designed simulation system which is superposed on a simplified hardware system block diagram for illustration. It starts with an electromagnetic waveform generator. A baseband signal is generated according to a selected signal configuration and sent to the radiator with predefined gains, transmitting power, and carrier frequency. The signal is radiated by one transmitting antenna after another through all the transmitting antennas with a defined sequence based on the time-division modulation (TDM) based approach. Tx-channel simulates the signal propagation in the free space and can be further extended to imitate multi-path signals. The human model used for the subject under test (SUT) module is designed to mimic the reflected signal off human subjects according to the RCS pattern of each body part and the incident angle. After reflected from the target, the signals will travel the same path back to the radar antenna array. Similar to Tx-channel, the Rx-channel is established for propagation simulation of reflecting signals in the free space.



**Figure 4.1.** The simulation system

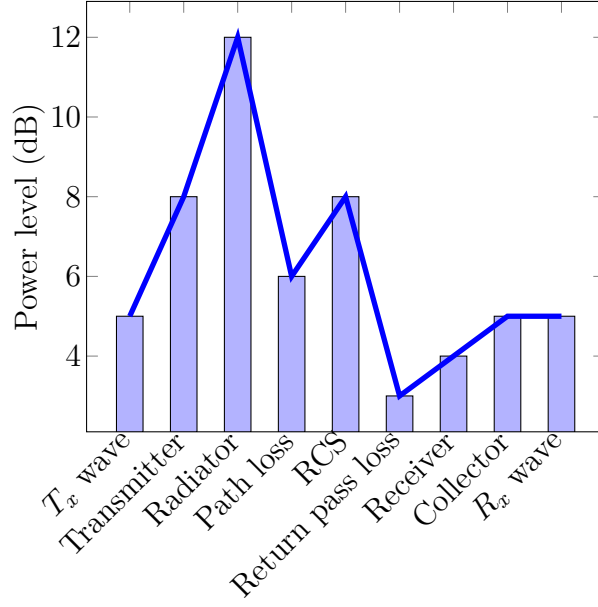
The reflecting signals are collected through the receiving antennas and then amplified with predefined gains. I and Q signals are generated by mixing and receiving signals for further signal processing and image reconstruction. For simplicity, the noise factors are not considered in this system.

During the transmission of signals, the power level of each signal changes when the signal travels through each module of the designed system. Based on the Two-Way equation and the example discussed in [81], an example is shown in Fig. 4.2 to represent the variation of power level. The radiation power off the radar antenna to the free space is assumed to be 12 dBm. The Tx waveform is generated from the hardware and typically it has a relatively low power level which is around 5 dBm in our example. After passing through a two-stage amplification, all the generated signals will see their power increase to 12 dBm when reaching the radiator. Then, when the signals are transmitted through the free space, their power level lowers when they reach the reflectable subjects, as most of the power is lost during transmission. Once signals are reflected by subjects and their power is increased by the RCS gain of subjects, the power will be further reduced when they travel back through the free space. When the receiver and collector receive the backscattered signals, the powers of the received signals will be increased through another two-stage amplifier before using them for image reconstruction.

#### **4.1.2 Parameter selection for waveform**

Nowadays, only a few 3D imaging radar technologies are currently available in the market. To make the mathematical-model-based simulation system compatible with the actual radar system, several radars are assessed and reviewed and a set of typical parameters of automotive-related FMCW radars from Texas Instruments are selected. The parameters are listed in Table 4.1.

Due to the limitations of the design, material, manufacturing processing, and so forth, it is challenging for an actual radar system to perform as well as a theoretical designed



**Figure 4.2.** Signal power level change by propagation stage

**Table 4.1.** Parameters for chirp design

Parameter	Value	Unit
Center frequency (f)	79	GHz
Bandwidth ( $B_{eff}$ )	3.74	GHz
Max distance ( $d_{max}$ )	3	m
Chirp duration( $T_c$ )	33	$\mu s$
Frame time ( $T_f$ )	30	ms
Sweep slope ( $S$ )	112	$MHz/\mu s$

simulation system. One of the examples is the limitation of sweep slope  $S$ . Theoretically, chirp duration  $T_c$  depends on the maximum distance  $d_{max}$  and speed of light  $c$  by

$$T_c = \frac{2d_{max}}{c} \quad (4.1)$$

$S$  is an effective bandwidth  $B_{eff}$  divided by  $T_c$  as

$$S = \frac{B_{eff}}{T_c} \quad (4.2)$$

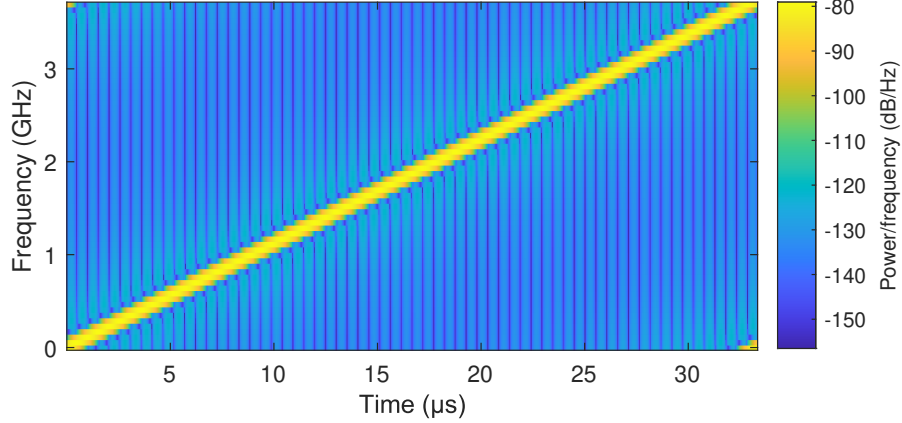
Then,  $B_{eff}$  is calculated by

$$B_{eff} = \frac{c}{2d_{max}} \quad (4.3)$$

If the maximum distance from Table 4.1 is selected, the chirp duration  $T_c$  becomes 33.3 ns, then the sweeping slope will be  $112 \text{ GHz}/\mu\text{s}$ , and the maximum beat frequency  $f_{b_{max}}$  is increased accordingly at a mixer as

$$f_{b_{max}} = \frac{2Sd_{max}}{c} \quad (4.4)$$

The current radar SOC technology could hardly help to reach this sweeping slope, which has to be reduced to an achievable level. A practical approach is to extend the chirp duration of  $T_c$ . Therefore,  $T_c$  is extended it by 1000 times for this simulation, as shown in Fig. 4.3 and extend the total chirp time to  $33.3 \mu\text{s}$ . The frequency span in baseband will be upconverted by central frequency 79 GHz, to increase the radiated wideband chirping signal from 77 GHz to 81 GHz.

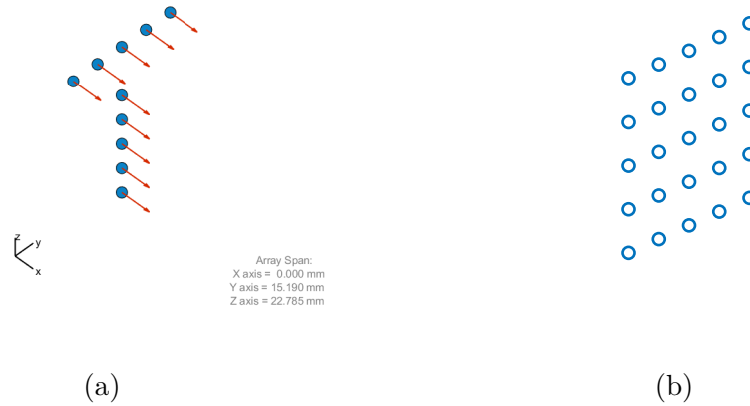


**Figure 4.3.** The spectrum of designed FMCW chirp

#### 4.1.3 Antenna array design

A typical design of a planer antenna array is to place a transmitting array and a receiving array on a different axis. The shape of the array design can vary, e.g., L-shape, T-shape, or others. To demonstrate the realization process of the MIMO antenna array in the simulation

system, a  $5 \times 5$  T-shape design is selected as an example shown in Fig. 4.4a. The red arrows indicate the normal direction of electromagnetic waves. The transmitting antenna array is located along the vertical axis while the receiving antenna array situates the horizontal axis. The virtual antenna array, i.e., a completed planar array, is formed, as shown in Fig. 4.4b. Therefore, a  $5 \times 5$  virtual antenna array is generated and the location of virtual antennas is shown in blue circles.

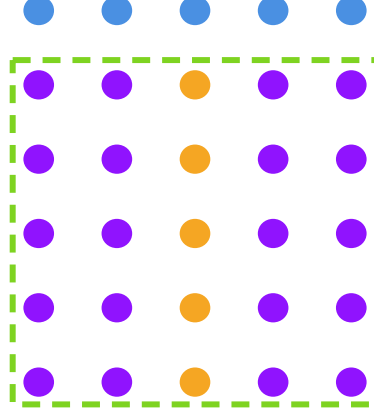


**Figure 4.4.** T-shape planar antenna array (a) actual antenna array with normal direction and (b) virtual planar antenna array

#### 4.1.4 Planer antenna array configuration

A TDM-based MIMO configuration is to create a planer array from the two actual antenna arrays, as shown in Fig. 4.5. The horizontal physical antenna array shown in blue dots are assumed as the transmitting antennas. The vertical physical antenna array shown in orange dots are assumed as receiving antennas. The transmitting antennas radiate signals from left to right in sequence; the time between each radiation is equivalent and predefined. Between each radiation, the receiving antennas receive the backscattered signals at the same time. With this modulation, the virtual antennas are established as the purple dots in Fig. 4.5. These antennas do not exist physically but are generated by the TDM-MIMO modulation. The planer antenna array is created using geometric equivalent by combining all the

actual and virtual antennas. The equivalent planar antenna array is the dots in the green dash-line box in Fig. 4.5.



**Figure 4.5.** Demonstration of virtual planar antenna array

#### 4.1.5 Steering matrix

The origin of the planar antenna array is defined as the center of the equivalent antenna array. The geometric position of each antenna pair in the antenna array is defined as the horizontal and vertical distance between the origin and each equivalent antenna. The receiving signal of each antenna pair in the antenna array can be expressed as Eq. (4.5) below,

$$S_a(i) = e^{j\frac{2\pi}{\lambda}\tau_a(\phi,\theta,i)} \quad (4.5)$$

where  $\tau_a(\phi, \theta, i)$  is time delay function of radar antenna array with respect to its geometric position  $pos$ , a function of distance between antenna pairs  $i$ , and incident direction, a function of incident angle with both elevation angle  $\phi$  and azimuth angle  $\theta$ .

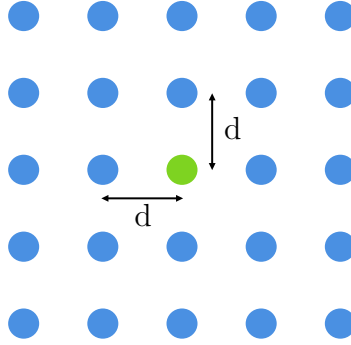
The antenna pair positions are the same as the planar antenna array in the dash-line box of Fig. 4.5. The distance between adjacent antenna pairs is a half wavelength,  $d = \frac{\lambda}{2}$ , as introduced previously, and the origin is set as the color of green. The distance matrix can



be generated accordingly, as shown in Fig. 4.6. Each element in the matrix should follow Eq. (4.6)

$$pos(i) = (md, nd) \quad (4.6)$$

where  $m, n$  are the antenna's geometric distance to the origin in horizontal and vertical axes, respectively.



**Figure 4.6.** Distance between each antenna pair

The estimation of incident direction  $dir(\phi, \theta)$  is based on Eq. (4.9) should follow

$$\begin{aligned} dir_x(\phi, \theta) &= \cos\phi\cos\theta \\ dir_y(\phi, \theta) &= \cos\phi\sin\theta \\ dir_z(\phi, \theta) &= \sin\phi \end{aligned} \quad (4.7)$$

## 4.2 Signal and image processing

### 4.2.1 TDM-based approach and algorithm

The TDM approach for a MIMO-based radar simulation system can be realized through a predefined one-hot weight vector which controls the sequence of signal transmission. The pseudo-code of generating the vector is shown in Algorithm 1. Each chirp loop, i.e., completing the transmitting and receiving process once, collects all the backscattering signals through the receiving antennas. The signals are reflected from the targets and environment

radiated by all the transmitting antennas. The receiving signals are processed and matched to the corresponding pair of transmitting-receiving virtual planar antennas. They are then fed into the matrix's relevant cell to match the virtual planar antenna array. Assuming the number of range bin were set to 100, then the radar cube's dimensions would become  $100 \times 5 \times 5$ . The data cube could also be rearranged to a  $100 \times 25 \times 1$  matrix. The three dimensions of the matrix are the number of range bins, the number of virtual channels, and the number of chirp loops.

---

**Algorithm 1** One-hot weight vector definition

---

```

1: function ONEHOTWEIGHT( $N_{chirploop}, N_{tx}$ )
2:    $w0 = \text{zeros}(N_{tx}, 1)$ 
3:   for 1 to  $N_{chirploop}$  do
4:      $k = \text{mod}(N_{chirploop} + N_{tx}, N_{tx})$ 
5:     if  $k \neq 0$  then
6:        $w0(k, 1) = 1$ 
7:     else
8:        $w0(\text{end}, 1) = 1$ 
9:     end if
10:  end for
11:  return  $w0$ 
12: end function

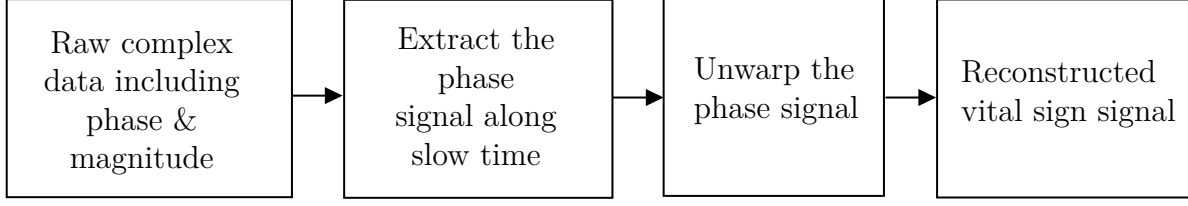
```

---

#### 4.2.2 Vital sign estimation

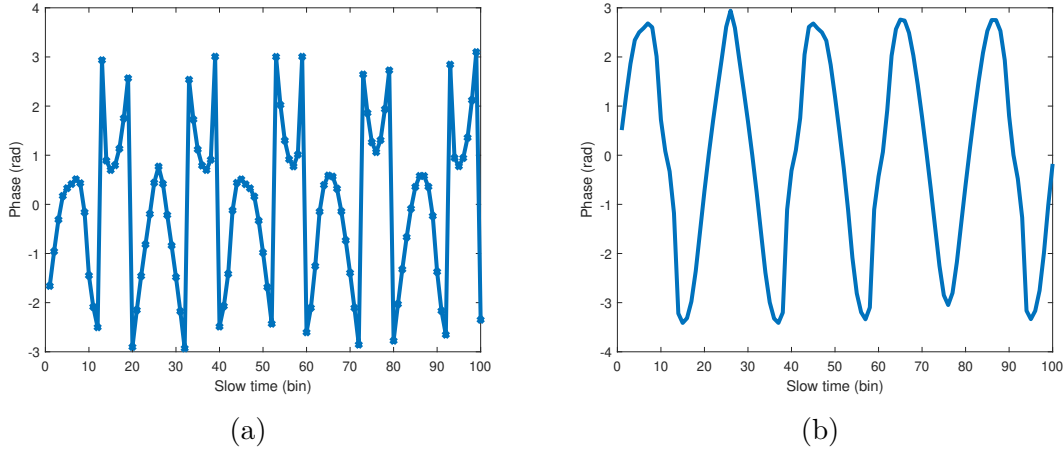
For a radar system equipped with nearly 4 GHz bandwidth, the macro-movement, i.e., head, upper-torso, arm, and leg movements, can be detected by signals along with the radiation direction. However, the radar range resolution may not be fine enough to catch the amplitude changes of micro-movement signals, i.e., breath and heart movements. In this system, a vital sign detection method is included to find the vital sign signal of a human subject. The approach used for the method is shown in Fig. 4.7.

This approach is to measure the phase change of signals that can be used to sense micro-movement. Once the location of the SUT is found, the raw complex data, including phase and magnitude, is recorded. Then, the phase is extracted from signals along the slow time. Nonetheless, the recorded change of phase includes inconsistency at the connection of two



**Figure 4.7.** Vital sign detection method

consecutive phase samples around  $\pi$ . To correct the inconsistency, an unwrapped algorithm is needed for vital sign signal reconstruction, as shown in the third block in Fig. 4.7. Fig. 4.8a shows the signal before the unwrap algorithm applied. After unwrapping the phase signal, the reconstructed vital sign signals can be obtained in Fig. 4.8b.

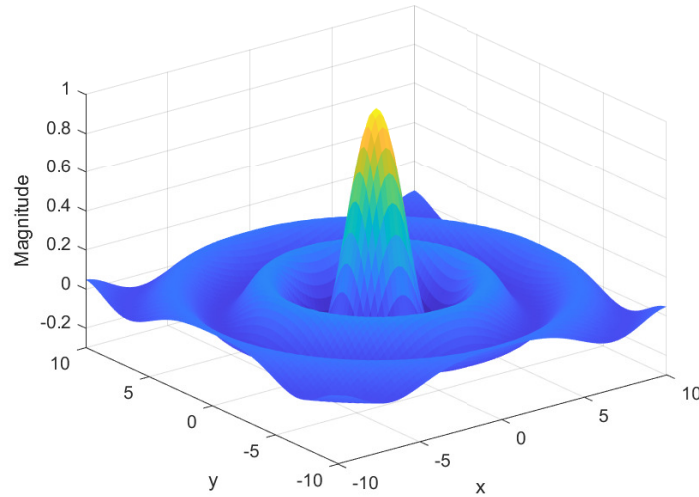


**Figure 4.8.** Comparison of the signal (a) before and (b) after the unwrapped algorithm applied

### 4.2.3 AoA estimation and heatmap generation

To obtain high-resolution images of the SUT, AoA estimation method is applied to the received data to find the power reflection level of subjects within the range of interest. The resolution of the image depends on the number of equivalent antennas in both horizontal and vertical axes. With more equivalent antennas in both axes, targets located in the close angular distance can be distinguished. The point spread function (PSF) for a single target imaging is typically in the form of 2D sinc function as in Fig. 4.9. The heatmap image of

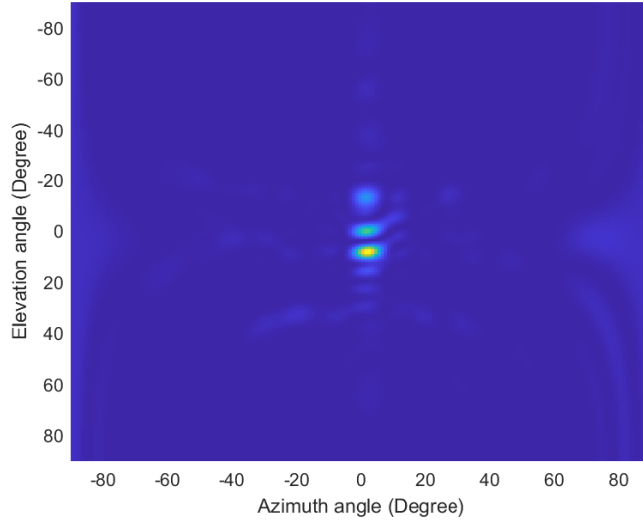
a target in azimuth and elevation includes the power reflection from the target of interest and noise with low power level. For a subject such as human body, the reflection pattern in the generated heatmap is complicated. The power reflection from each body part of human body is mixed with noises, making some parts of the human body not as explicit as others.



**Figure 4.9.** Point spread function

Fig. 4.10 shows a heatmap measurement of a human facing a radar antenna array. The center of radar is facing the belly. The human is lifting both of the arms horizontally. The radar array scan both azimuth and elevation domain from  $-90$  to  $90$  degree using the AoA approach mentioned in the previous chapter. The brighter the color is in the heatmap, the stronger the power level will be. As can be seen from the heatmap, only a few points are explicit in the heatmap, making the detection of other body parts difficult.

In this simulation system, the Bartlett beamforming is chosen when the angular resolution is adequate. Although it provides a worse resolution than other advanced methods, it consumes less computational power and saves processing time. If higher resolution is required and the number of targets is known, MUSIC can be a better fit, such as for posture estimation of human subjects.



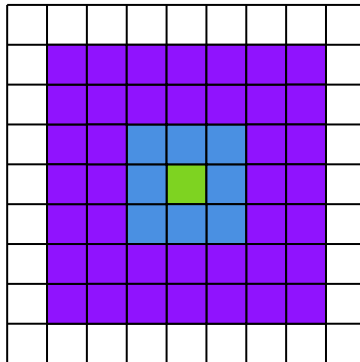
**Figure 4.10.** A heatmap measurement of a standing human in real world

#### 4.2.4 CFAR thresholds

To find the signal of interest and eliminate noise and interference signals, an adaptive threshold method needs to be applied to the generated heatmap from AoA. In this work, 2D cell-averaging constant false alarm detection (CA-CFAR) is selected to find the signal of interest from noise. In principle, CFAR uses cells around the cell containing the signal under test to calculate the noise level and then check if the signal under test has a higher value than the calculated noise level. If its value is higher than the noise level, the signal in the cell will be selected as a signal of interest. Fig. 4.11 shows how this method works. The grid in Fig. 4.11 is assumed to be the heatmap generated from AoA. The signal under test is in green color. First, the guard cells are defined and will not be included in the noise level calculation. In Fig. 4.11, the guard cells with size one are shown in blue. For the calculation of noise level, a window is defined. The window used for noise calculation in Fig. 4.11 is shown in purple and its size is defined as 2. Assume the magnitude of the signal in the cell

is  $P_s$  and the total number of cell in the window for noise estimation is  $N_w$ , the noise level  $NL$  can be estimated by Eq. (4.8)

$$NL = \frac{1}{N_w} \sum_{s=1}^{N_w} P_s \quad (4.8)$$

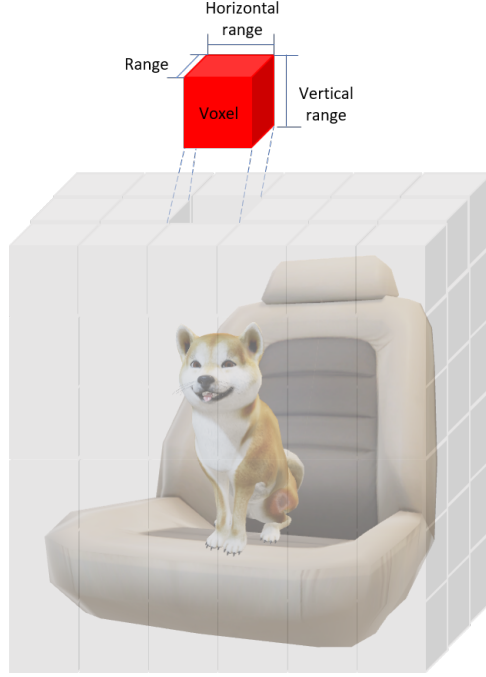


**Figure 4.11.** 2D CA-CFAR threshold method

#### 4.2.5 Coordinate transformation

As mentioned previously, the radar signal is measured in a Polar coordinate as in Fig. 3.10. It measures reflections with respect to distance, azimuth angle and elevation, which is different from the measurement in the Cartesian coordinate.

The concept using Cartesian-coordinate-based measure of range resolution and cross-range resolution in 3D space is shown in Fig. 4.12. The 3D space is divided into Voxels  $(\Delta x, \Delta y, \Delta z)$  with each dimension standing for range, horizontal range, and vertical range. The power level of each Voxel stands for the magnitude of reflections of the incident radar signal. Each Voxel is represented in a Cartesian coordinate.



**Figure 4.12.** Range and cross-range resolution in Cartesian coordinate

All the detected reflection signals from the last step need to be mapped to the Cartesian coordinate for post-processing with Eq. (4.9).  $R$  represents the distance between the radar and pixel of interest;  $\phi$  and  $\theta$  are the elevation angle and azimuth angle, respectively.

$$\begin{cases} x = R \cdot \cos\phi \cos\theta \\ y = R \cdot \cos\phi \sin\theta \\ z = R \cdot \sin\phi \end{cases} \quad (4.9)$$

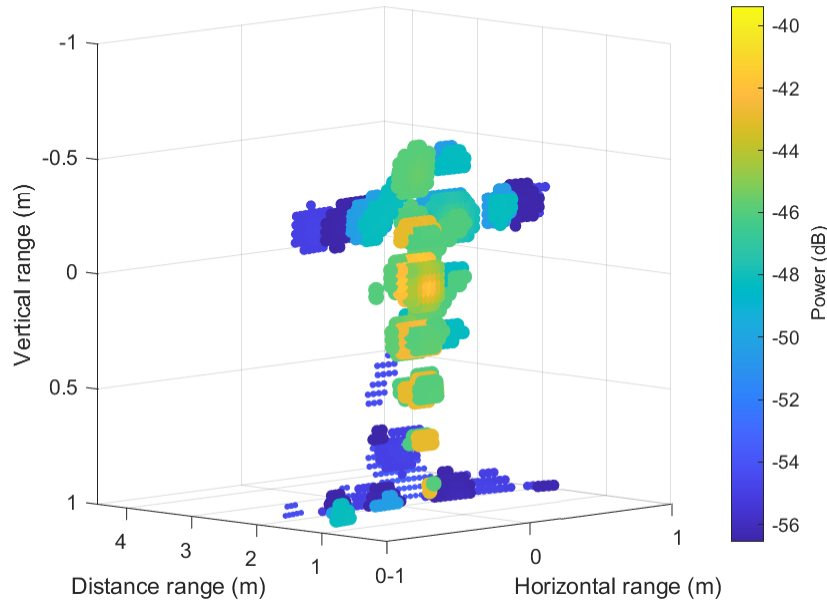
#### 4.2.6 Point cloud approach

The radar signals through the several processing steps discussed above can be placed into the corresponding Voxel of the 3-D point cloud space. The magnitude of each Voxel shows the power reflection level of the subject under test in different ranges. However, because the point cloud is estimated along the range axis in the polar coordinate, the generated point cloud could contain few ghost signals with a low power level. Those signals can be eliminated

by applying a power-based threshold  $T_{power}$  as in Eq. (4.10). If the magnitude of Voxel is above the threshold, the point cloud will be kept; otherwise, the magnitude of the Voxel will be set to zero.

$$\begin{cases} y & \text{if } y > T_{power} \\ 0 & \text{Otherwise} \end{cases} \quad (4.10)$$

By applying the signal and image processing method mentioned above, the point cloud of the standing human mentioned previously is shown in Fig. 4.13. All body parts are visible in 3D space with each detected body part in the form of a filled circle. The magnitude of each Voxel shows the reflective power level of the corresponding body part of a human. It is obvious that belly has the strongest power reflection compared to other parts.



**Figure 4.13.** Point cloud plot of a standing human

#### 4.2.7 Processing flow

A 3-D imaging radar processing flow is developed as illustrated in Fig. 4.14. First, the raw complex signals are received and stored into a data cube, following analog-to-digital



(ADC) samples, chirps, and antennas, by defined orders. The ADC samples which reflect the time delay of signals, are converted to a range profile. If the phase change of the interested range profiles, or called range bin, is recorded, the vital signs from the test subjects can be obtained accordingly. The next is to calculate the Doppler profile by applying FFT along the chirps axis. It tracks the velocity changes of the test subjects. This research focuses on the detection and classification of human subjects in-cabin. Thus, we are only interested in the data from one bin, the doppler-zero bin. The green slice shown in the data cube represents the signals with zero doppler velocity. Then, the beamforming approach is applied on the data of doppler-zero bin, which is to apply one range bin (the orange slice) after another along with the range profile with CFAR threshold and get the heatmaps accordingly. Lastly, these heatmaps are stacked together before being transformed into Cartesian space; thus, a 3D image of the subject in  $(x, y, z)$  space is obtained. The quality of 3D imaging depends on the number of antennas in both the azimuth and elevation axes.

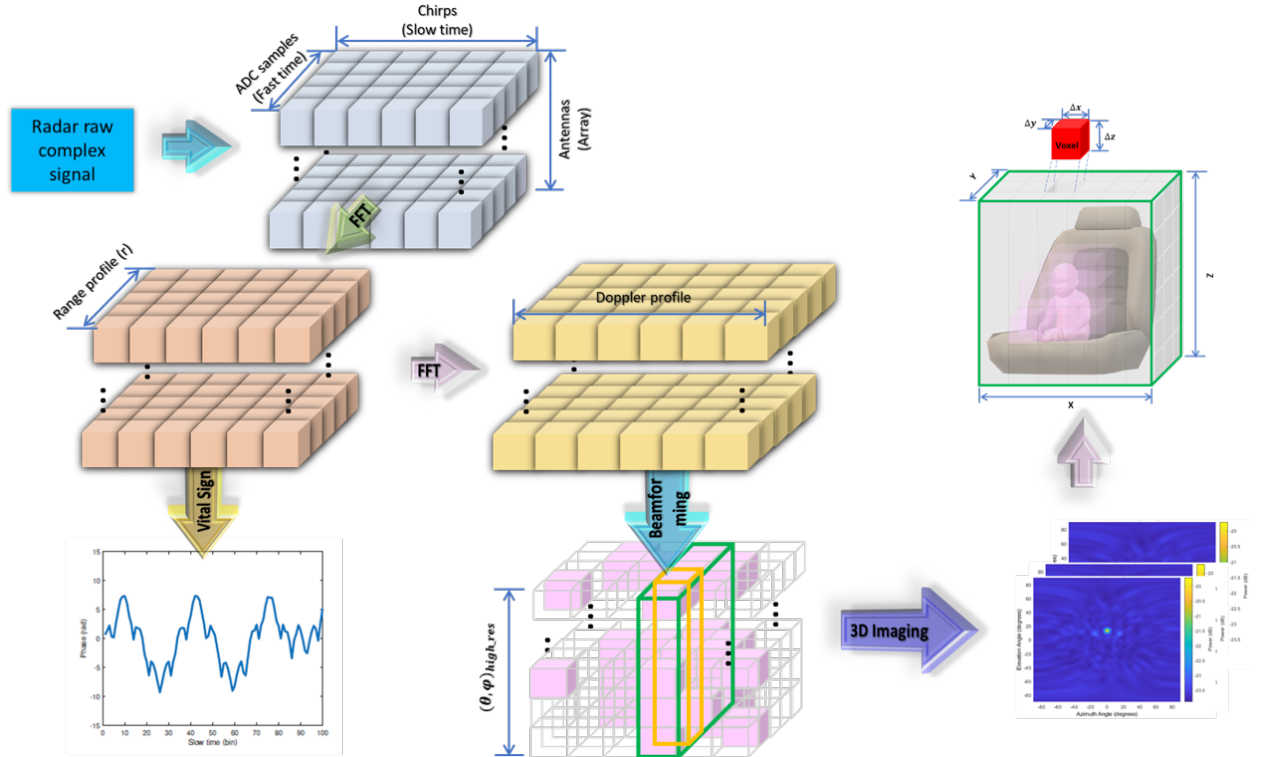
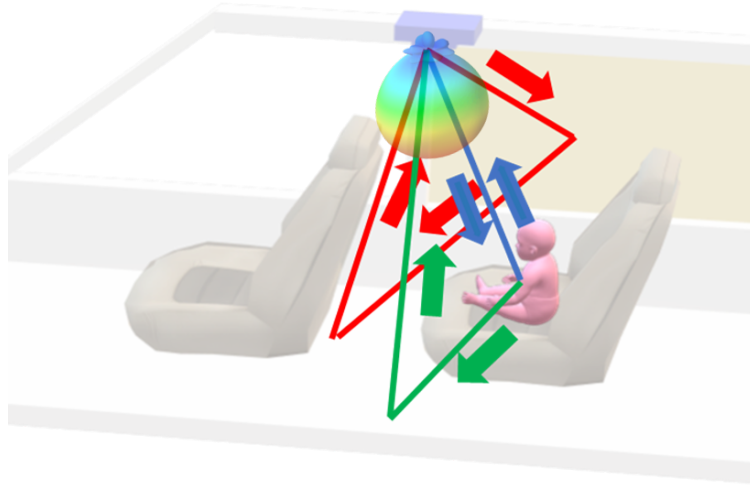


Figure 4.14. Imaging and vital sign detection pipeline

For radar-based vehicle cabin monitoring system, there are multi-path due to the small space inside the vehicle cabin and side lobe related to the radar design. The side lobe can be suppressed by digital beamforming. Fig. 4.15 shows the direct path and multi-path reflections inside the vehicle cabin. The blue line shows the direct reflection of the radiated radar signal; the red and green line is the multi-path of signals that are not directly reflected back to the antennas. Due to the randomness of multi-path signals, they could be mitigated by accumulating the detected results across the slow time and filtering out themselves with low temporal frequency in a certain Voxel.

For the reflection signals from other noises or static subjects, such as seats or other interior structure, inside cabin, they can be estimated separately to reduce their effects. These signals can be treated as clutter during the second data cube calculation for range estimation and can be removed.



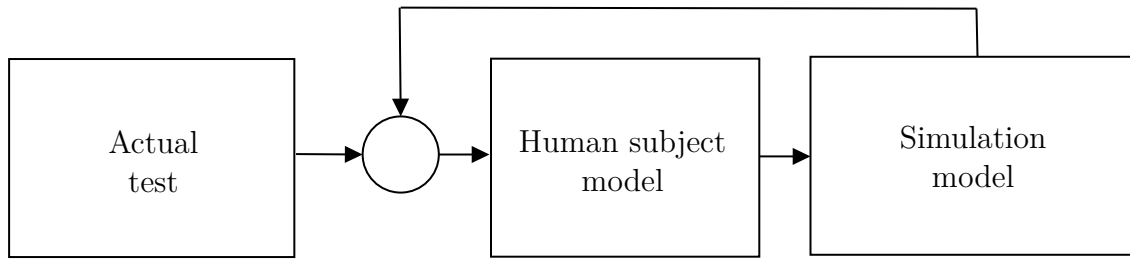
**Figure 4.15.** Multi-path effect inside vehicle cabin

### 4.3 System validation with human subject

#### 4.3.1 Design of the reflection-based human subject model

As discussed previously, the generated point cloud of human subjects is comprised of a bunch of filled circles. The size and magnitude of those circles are affected by the posture and

orientation of the human body towards radar. To design a simulation human subject model for radar sensing system applications, a procedure is designed in Fig. 4.16. Firstly, the point cloud of the actual test is generated using the aforementioned signal and image processing methods as reference. Then, a human subject model is proposed based on observation from the actual test to mimic the power reflection of the actual test. The model is then used in the radar simulation system to generate a point cloud of the simulation human subject using the same parameters as those of the actual test. The two results are then compared to find the difference. The procedure will stop until the simulated point cloud has a similar pattern to the actual test. The same procedure is repeated under other test scenarios to update human subject model design and find the simulation model that can be used for different postures and orientations.

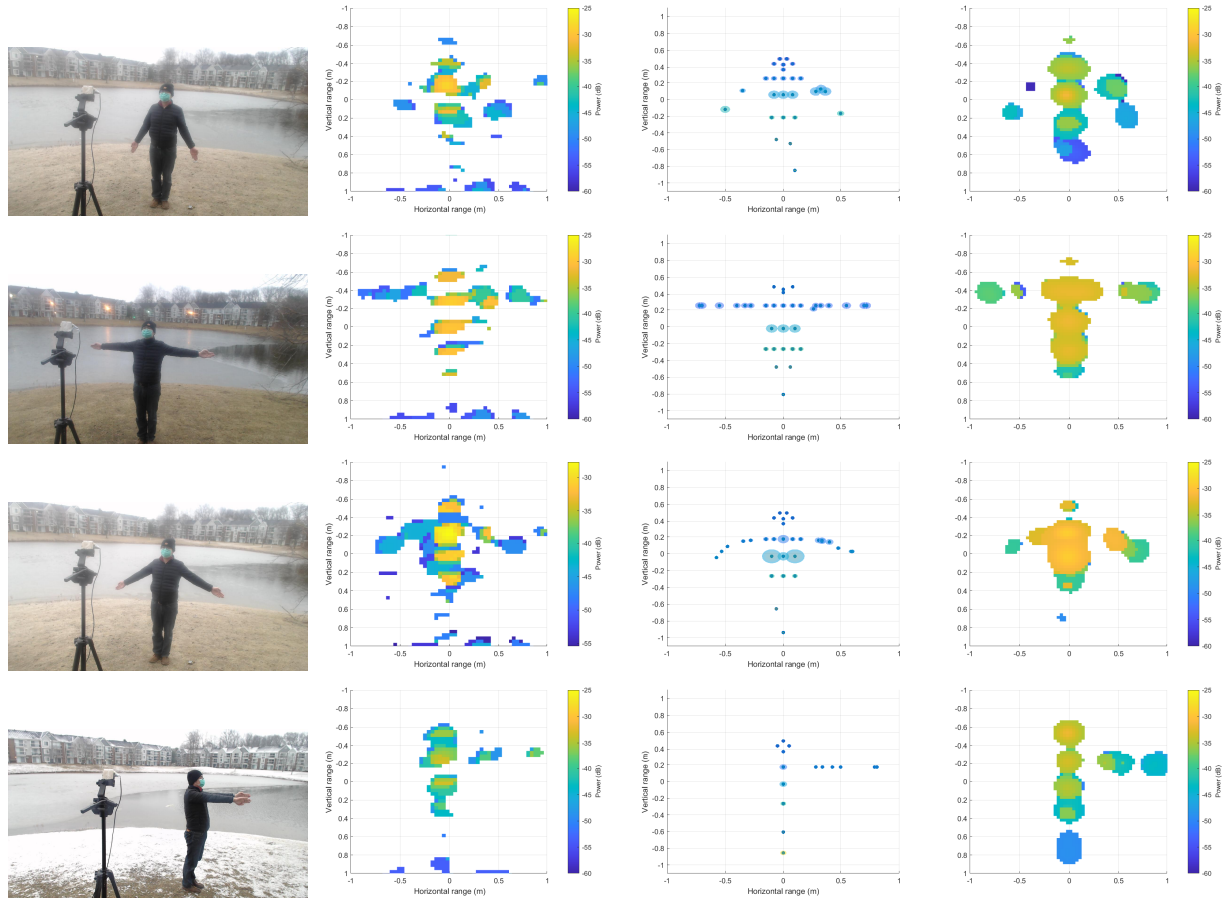


**Figure 4.16.** Procedure to find the generalized reflection-based model for human subject

### 4.3.2 Experiments

Four tests are designed in an open field to mimic the power reflection of human body inside vehicle. The radar is placed above a person's head and is tilted to an angle of 15 degrees with respect to the vertical line on the ground. The various of the four tests are 1) standing human subject with both arms close to torso 2) standing human subject with both arms lifting horizontally 3) standing human subject with both arms forming a 45-degree angle with torso 4) standing human subject lifting both arms to the front of body and has the side of torso facing radar. The results are shown in Fig. 4.17. The first column in Fig. 4.17 is the ground truth image of a human subject and the radar mounting position. The second column shows the point cloud result of the actual test. The third column represents

the designed human subject model for the corresponding test. The fourth column in the figure is the power reflection level of the human subject in simulation.



**Figure 4.17.** Tests of real human subject and corresponding design of reflection-based human subject model in simulation

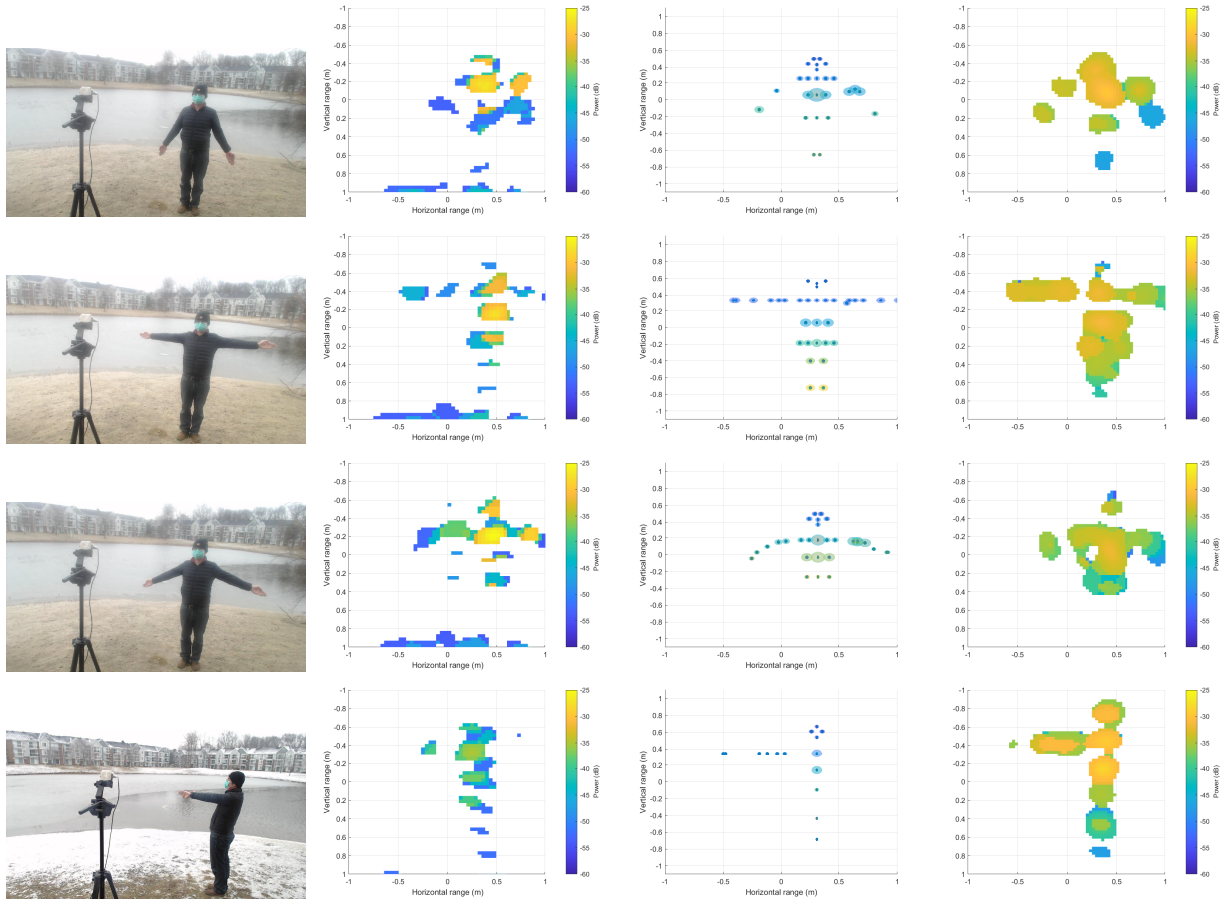
It can be seen that the image pattern of these four test scenarios in the fourth column are similar to the actual test in the second column. Nevertheless, the magnitudes each body part in these two plots are not similar to each other. The discrepancy occurs because the actual test under the designed test scenarios uses the off-the-shelf (OTS) radar antenna array with a different waveform. There is no FMCW-based antenna array with a sufficient number of antennas in the market to meet the requirement of human subject imaging in this research; only a SFCW-based planar antenna array radar with similar central frequency and MIMO configuration is available. Both the OTS radar antenna array system and the simulation radar array system can provide similar image patterns for the same test subject,

because the two systems have similar array layout. Therefore, only the array configuration and some of the parameters of the simulation system are adjusted to comply with the known configurations of the off-the-self radar antenna array.

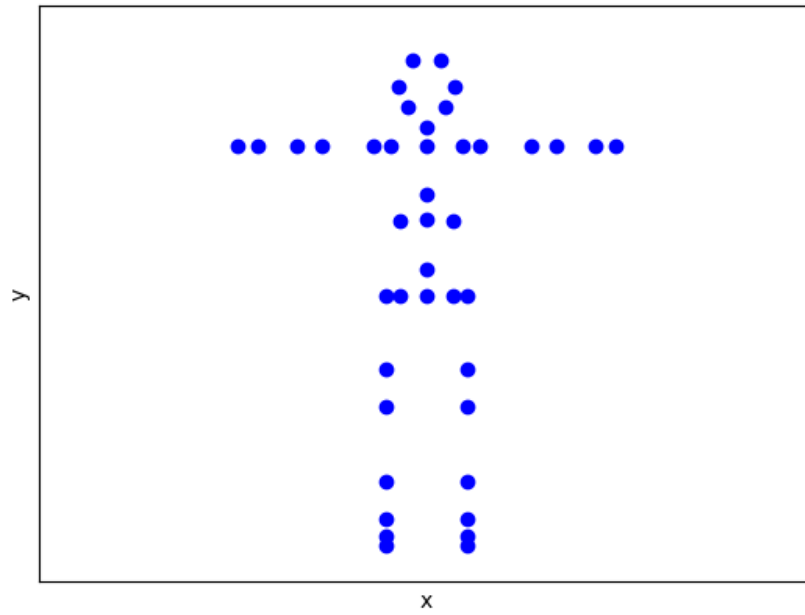
Fig. 4.18 shows the results of another four tests with offset. These four tests using the corresponding human subject model from Fig. 4.17 to check the robustness of the designed model. Both the point locations and RCS patterns are either slightly modified or unchanged according to the counterparts in the first four tests; a 0.3 meters displacement is apply to each of the four cases for comparison with the reflection patterns in the first four tests. It can be seen that the designed model provides a similar reflection pattern in the second column even with a distance offset; the reflection patterns are also close to the patterns in the first four tests. This shows that this type of human body model and corresponding simulation system can provide adequate reflection information for occupancy monitoring occupant in different locations.

### 4.3.3 Reflection-based human subject model

Based on the designed model for the eight tests mentioned above, a human subject simulation model is proposed below in Fig. 4.19. The model considers the symmetry of human body by including more points for legs and feet. The head is designed with points to form the shape of an ellipse. For each body part, two closely located points are used to generate an explicit reflection pattern of the corresponding body parts. Two more points are added horizontally at the torso area to represent its width. Based on the results from the eight designed test scenarios, this model is adequate to represent the reflection of a human test subject for the selected mmWave radar and will be used in the later chapter to generate reflections radar signals of human body with various sizes, postures, and orientations.



**Figure 4.18.** Tests of real human subject and corresponding design of reflection-based human subject model in simulation with offset



**Figure 4.19.** Reflection-based human subject simulation model

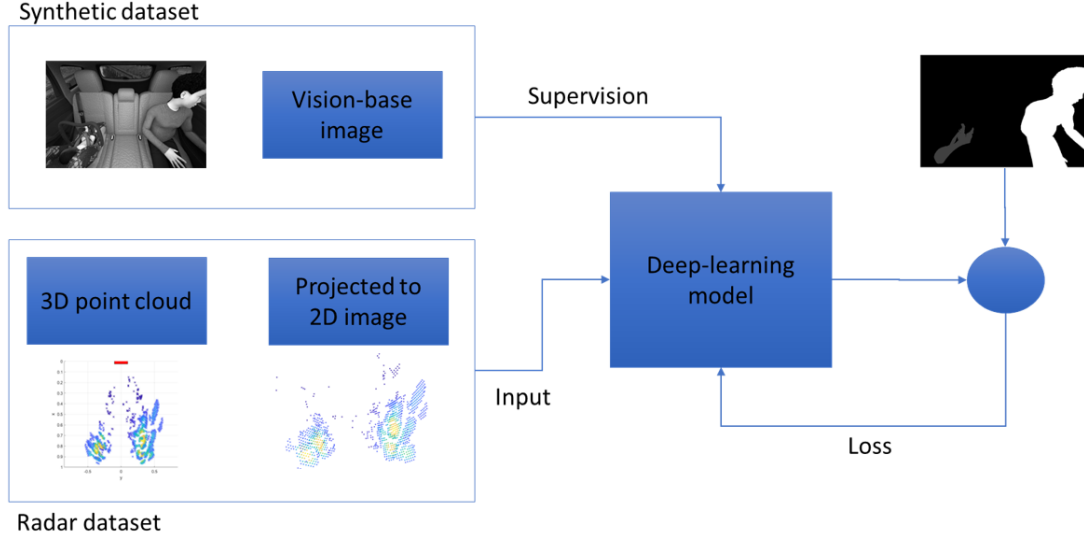
## 5. HUMAN SUBJECT CLASSIFICATION

### 5.1 Overview of human subject classification

As discussed in previous chapters, the radar-based image is a sparse set of point clouds. The power level of each point indicates the reflection of backscattered signals from a test subject. The characteristic of the radar image is totally different comparing with the image from the camera taken from the same test subject. Various radar signal processing methods were proposed to identify or categorize the test subject. A 3D zone-based method is proposed for occupant classification [113]. However, this method is not flexible because occupants could move between different zones in 3D space. A 3D convolutional-neural-network (CNN)-based segmentation method is another approach for the classification, but it requires large computation resources and is not practicable for real-time automotive applications. In this chapter, a 2D point-cloud-projection-based method is proposed. It reduces data dimension comparing with the 3D approach. A deep-learning-based method is used to transfer the point cloud reflection pattern to vision-based image segmentation. The vision-based image data are required as ground truth for the training.

The overview of the approach is shown in Fig. 5.1. The top branch uses vision-based image from synthetic dataset as supervision in the deep-learning model. The branch on the bottom uses data from radar-based images as input to the deep-learning model. The radar dataset is generated from the defined human subject simulation model to obtain a 3D point cloud and then project the point cloud in 3D space to a 2D image by computer-vision-based method. The input data to the deep-learning model are projected images in 2D. The output of the deep-learning model is the segmentation of the input images. The output segmentation masks are compared with the ground truth data to compute the loss. The loss is then feedback to the deep-learning model to update the weights and biases.





**Figure 5.1.** Overview of the deep-learning-based method for human subject classification

## 5.2 Imaging dataset preparation for classification

### 5.2.1 Vision-based in-cabin synthetic imaging dataset

The data set used in this research is a synthetic data set called Synthetic Vehicle Interior Rear Seat Occupancy (SVIRO) [114]. This data set provides different types of occupants, including infants, children, and adults. These occupants have different postures between frames and they are designed for indifferent vehicle types, including Ford Escape, Hyundai Tucson, and Tesla Model 3. The authors in [114] verified the synthetic data with data collected from a vehicle in the real world, as in Fig. 5.2. They also test some machine learning and deep learning algorithm which verifies that this data set can reflect occupant posture and occupancy in real world. This data set provides RGB and greyscale images, depth images, segmentation masks, Keypoints for pose estimation, and a bounding box.

The camera data needs to be included to solve the radar point cloud's ambiguity caused by the size of the antenna array so that occupants inside cabin can be classified precisely. The segmentation ground truth data from the SVIRO dataset have seven categories. The ground truth used in this chapter only has three categories, child, adult, and others. The camera ground truth of Fig. 5.4a and Fig. 5.4b can be shown in Fig. 5.3. Different categories



**Figure 5.2.** Comparison of real world data with SVIRO [114]

are displayed in different colors. Specifically, the adult in Fig. 5.3 is depicted as white, the child in this figure is assigned with grey color, and others are shown as black.



**Figure 5.3.** Generated ground truth of camera data

### 5.2.2 Camera-radar alignment

Although the radar antenna array and camera have been allied to the same location, their poses are different. It requires rotating and projecting the radar image to a 2D camera map to align with the camera ground truth data. Intrinsic and extrinsic matrices are created for the alignment. The extrinsic matrix maps the point cloud data from radar to camera, and the intrinsic matrix aligns the rotated radar signal with the image plane.

In SVIRO, the intrinsic parameters are shown in Table 5.1. It is noted that the focal length to the pixel is for both the x-axis and y-axis of images. These parameters in Table 5.1 can help map the points in the world coordinate to the camera coordinate when no rotation or translation exists.

**Table 5.1.** Intrinsic parameters

Parameter	Value
focal length	3.4 mm
sensor width	8.5 mm
fnumber	2.5
skew coefficient	0
focal length (pixel)	514.4208
principle point	(640, 480)

The intrinsic matrix can be represented as in Eq. (5.1).

$$K = \begin{bmatrix} f & 0 & p_x \\ 0 & f & p_y \\ 0 & 0 & 1 \end{bmatrix} \quad (5.1)$$

where  $f$  is the focal length,  $p_x$  and  $p_y$  are focal length with respect to pixel.

The extrinsic matrix represents the rotation and translation from the world coordinate to the camera coordinate. The extrinsic matrix  $Z$  can be described in Eq. (5.2).  $R$  represents the rotation matrix by multiplying rotation matrices along (x, y, z) axes in this equation.  $C$  is the translation matrix between two coordinate systems, which, in this case, is zero.

$$Z = R[I] - C = \begin{bmatrix} r_1 & r_2 & r_3 \\ r_4 & r_5 & r_6 \\ r_7 & r_8 & r_9 \end{bmatrix} \quad (5.2)$$

Using these parameters, the intrinsic matrix and extrinsic matrix can be built up by Eq. (5.3). By mapping the radar point cloud data to the camera coordinate system, the 3D point cloud could be aligned with the image plane.

$$\begin{bmatrix} x_p \\ y_p \\ z_p \end{bmatrix} = \underbrace{\begin{bmatrix} f & 0 & p_x \\ 0 & f & p_y \\ 0 & 0 & 1 \end{bmatrix}}_K \underbrace{\begin{bmatrix} r_1 & r_2 & r_3 \\ r_4 & r_5 & r_6 \\ r_7 & r_8 & r_9 \end{bmatrix}}_Z \begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad (5.3)$$

### 5.2.3 Radar-based in-cabin imaging dataset generation

As discussed in the previous Chapter, the human subject simulation model consists of 52 key points for the classification. The 18 points below are selected from those 52 points that a similar Keypoints model could be obtained as described in Fig. 6.1. The location of heart is neglected since only the reflection image of occupant is considered. The selected points are shown in Table 5.2.

The ground truth dataset contains the 3D coordination information for the 18 Keypoint model. Based on these 3D Keypoints, additional points are inserted to generate a reflection-based simulation model discussed previously. The model is used to generate radar image dataset for the training and testing.

In this model, each point's RCS pattern is defined as a sphere of the same size, so each point of the same body's reflection pattern is the same in terms of incident angle. One way to distinguish the power reflection patterns of children and adults is that the radius of the point's RCS pattern of children is usually designed smaller than that of the adult.

The point cloud generation follows the imaging method discussed in the previous section. The point cloud data consists of two types of information, the coordinate location and the power reflection level. The power reflection pattern of occupants is different from each other in terms of posture, size, and position to radar sensor. Since the reflection-based model has more points, a more dense point cloud will be generated from the radar simulation system. To reduce the total number of points, a method is proposed to reduce the point cloud in 3D space. This method selects the point with maximum power reflection in each

**Table 5.2.** Selected Keypoints for human model

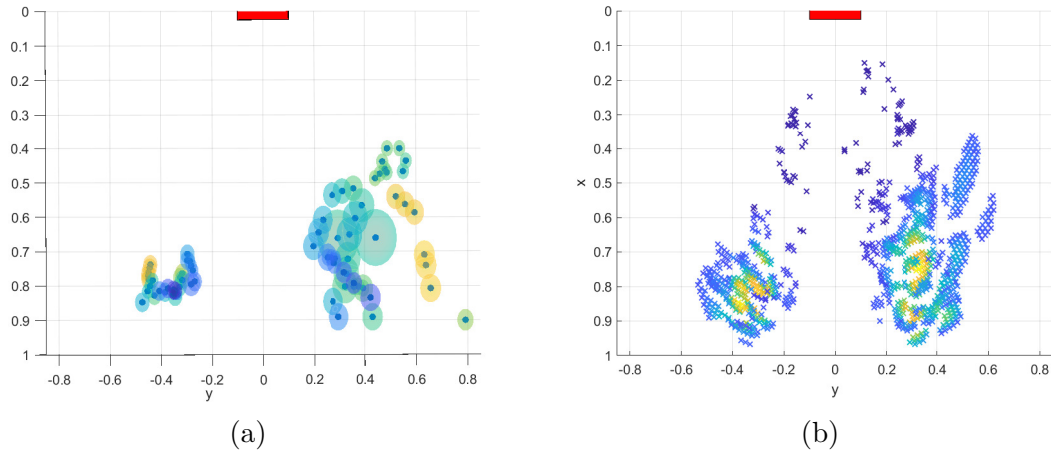
Name	Keypoint
head	"head"
neck	"neck_01"
chest	"spine_03"
pelvis	"pelvis"
left shoulder	"upperarm_l"
right shoulder	"upperarm_r"
left elbow	"lowerarm_l"
right elbow	"lowerarm_r"
left hand	"hand_l"
right hand	"hand_r"
left hip	"thigh_l"
right hip	"thigh_r"
left knee	"calf_l"
right knee	"calf_r"
left ankle	"foot_l"
right ankle	"foot_r"
left toe	"ball_l"
right toe	"ball_r"

azimuth-elevation angular pixel along the range axis. The equation can be described as Eq. (5.4).

$$P(\theta, \phi, r) = \max_r P(\theta, \phi, R) \quad (5.4)$$

where  $\theta$  is the azimuth angle,  $\phi$  is the elevation angle, and  $R$  is the range.

An example of RCS pattern and generated point cloud design can be shown in Fig. 5.4a and Fig. 5.4b respectively. In both plots, the red box is the radar sensor's location. It is mounted overhead on the roof of a vehicle. In this scenario, a child sits on the left and an adult sits on the right. The size of the child's RCS pattern is smaller than that of the adult. Fig. 5.4b, the radar image, is generated by the model with RCS feature added as shown in Fig. 5.4a. The lighter portion in this plot represents the body part that reflects more incident radar signal. The  $30 \times 30$  MIMO antenna array is selected to strike a balance between optimizing computational time and achieving a high-qualified point cloud. Some

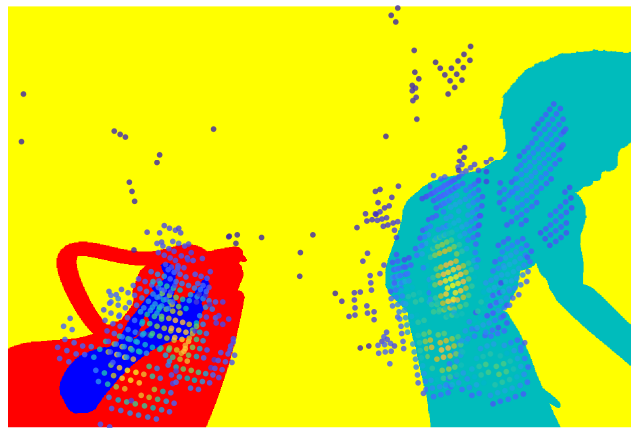


**Figure 5.4.** (a) RCS pattern of occupant model and (b) point cloud of occupants

body portion of the simulation model may be still unclear or missing due to the angular resolution of the radar antenna array.

#### 5.2.4 Point cloud 2-D projection

The last step is to project the 3D point cloud data to the plane of the 2D image for both reflected power pattern of the test subjects. The overlay of the point cloud image with the segmentation ground truth is shown in Fig. 5.5.

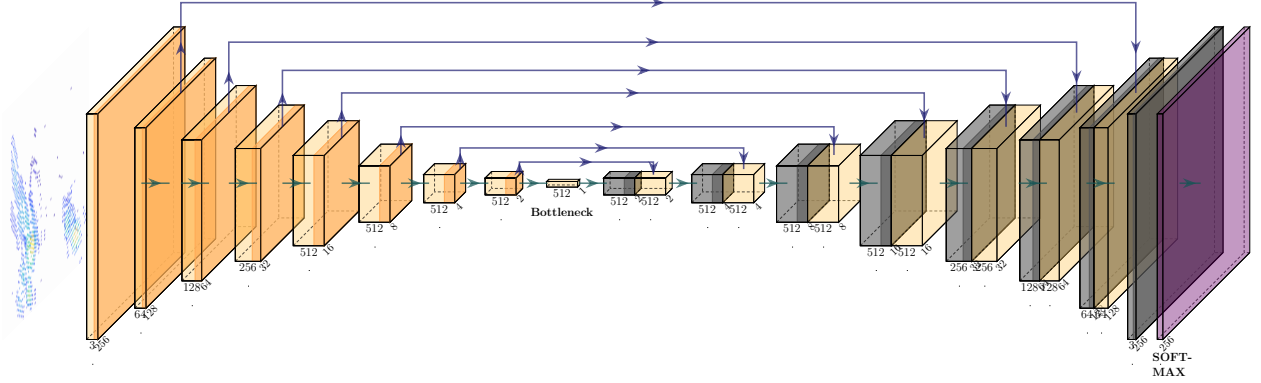


**Figure 5.5.** Point cloud projection overlaid with segmentation ground truth

## 5.3 Deep learning framework

### 5.3.1 Network structure

A typical deep learning model for image segmentation or image denoise, Encoder-Decoder based network, is used for prediction of vision-based segmentation from radar image data for human test subject classification.



**Figure 5.6.** Revised Unet for occupant classification [75]

The model, a revised Unet network, is comprised of several parts. The first part is an encoder to extract the features of input image and compress it into a  $1 \times 1$  feature map. The input image is first resized to a  $256 \times 256$  image to fit the structure of this network. Each block consists of a convolutional layer and leaky Relu layer. The first three blocks reduce the image size by half and increase the feature map's size, while the last four blocks keep the size of the feature map and reduce the size of the image to extract more features from the image. After seven blocks of encoding, a  $1 \times 1$  feature map is obtained as the bottleneck of this network.

After this network's bottleneck, corresponding seven decoder blocks are used to translate the point cloud projection image into image masks. Each decoder block consists of a transpose convolutional layer and a Relu layer for upsampling the image. Besides the transpose convolutional layer and the Relu layer, a skip connection from the encoder block is also used. The skip connection is concatenated with its corresponding decode block to add more features.

The last layer is a  $256 \times 256$  image with three classes as previously defined for the occupant classification. To output each pixel category in the image, a softmax layer is used to obtain the mask and classification of occupants.

### 5.3.2 Loss function

The loss function used in this chapter is Lovász-Softmax loss [115] for multi-class classification with each class  $c \in \mathcal{C}$ . It is an enhanced softmax starting with Eq. (5.5), where  $i$  is the  $i$ -th pixel in the image.

$$\mathbf{f}(c) = \frac{e^{F_i(c)}}{\sum_{c' \in \mathcal{C}} e^{F_i(c')}} \quad (5.5)$$

The vector of pixel errors  $\mathbf{m}_i(c)$  is proposed in [115] as Eq. (5.6) with ground truth  $y^*$ .

$$\mathbf{m}_i(c) = \begin{cases} 1 - f_i(c) & c = y_i^* \\ f_i(c) & otherwise \end{cases} \quad (5.6)$$

The loss is shown as Eq. (5.7)

$$loss(\mathbf{f}(c)) = \overline{\Delta}_{J_c}(\mathbf{m}(c)) \quad (5.7)$$

where  $\Delta_{J_c}$  is defined as the set of mispredictions  $M_c$  in Eq. (5.8) and  $\overline{\Delta}_{J_c}$  is the Lovász extension [115] of the misprediction set.

$$\Delta_{J_c} : M_c \in \{0, 1\}^p \mapsto \frac{|M_c|}{|\{y^* = c\} \cup M_c|} \quad (5.8)$$

### 5.3.3 Training

In SVIRO dataset, each vehicle has only 2000 images in the training set. Each vehicle contains eight adults and six children serve as test subjects. Posture of each human subject is randomly generated between images according to [114]. In the training set, some images only have non-human subjects, so the total number of images containing human subject is



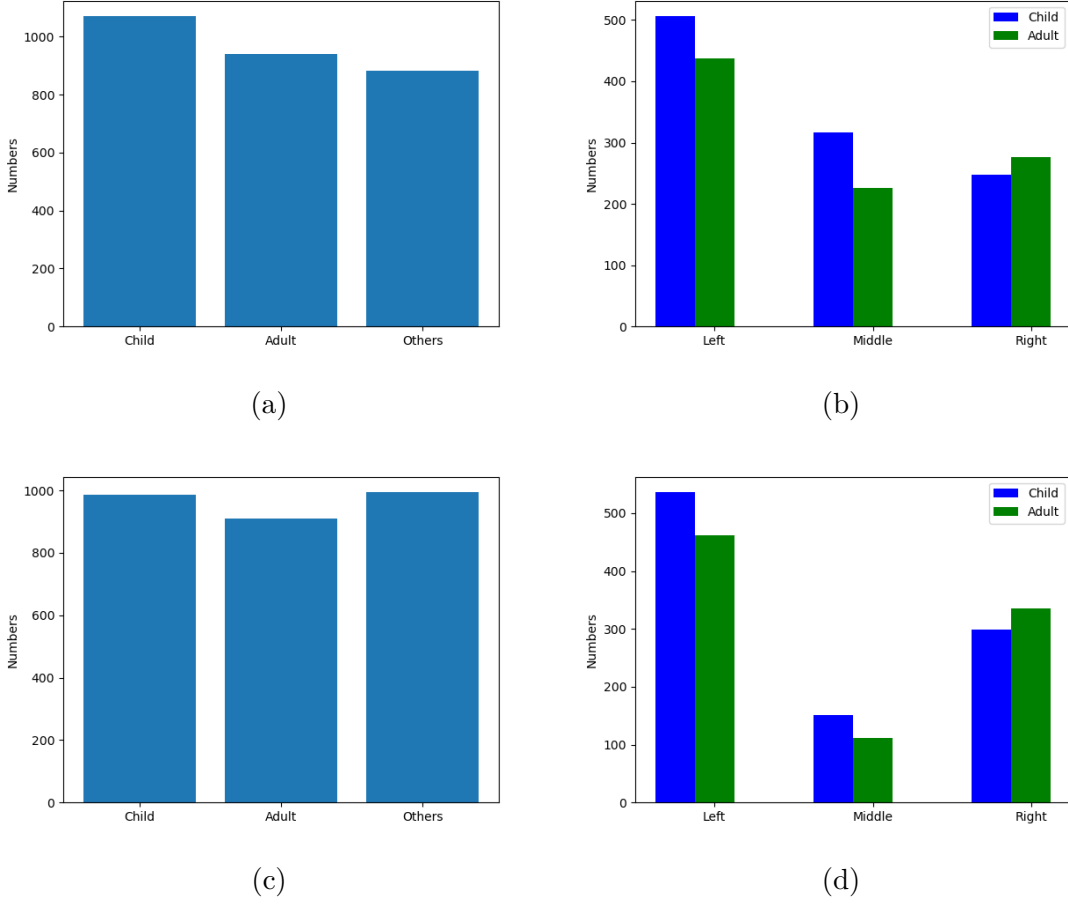
fewer than 2000. Therefore, the number of the image dataset is doubled by flipping images along the horizontal axis. In order to obtain adequate amount of the images for the training, the dataset from two different vehicle as shown in Fig. 5.7a and Fig. 5.7b are used. One is Ford escape and the other is Tesla model 3.



**Figure 5.7.** Layout of (a) Ford Escape and (b) Tesla Model 3

For the training set, 80% of the data is used for training while the rest of data is used as a validation set. Each mini-batch used for training has images in it. The total number of epoch for the training is 50.

The distribution of occupants inside Ford Escape is summarized in Fig. 5.8a and Fig. 5.8b. Fig. 5.8a shows the number of occupants in each category. The dataset is a relatively balanced, i.e. children, adults and others, is more or less the same - around 1,000. If we delve into the portion of the data set where there is at least one occupant inside, as shown in Fig. 5.8b, the figure shows that the distribution of different occupants is imbalanced. Therefore, data augmentation is applied in this research to balance the number of each category. Every image in the data set is flipped horizontally and concatenated to the original data set to keep the number of children and adults of different seats at the same level. Distribution of occupant inside Tesla Model 3 is slightly different from that of Ford Escape. As shown in Fig. 5.8d, the number of occupants sitting in the middle seat is fewer than the number of occupants in the left and right seats, as the middle seat is smaller in size.



**Figure 5.8.** Statistics of occupant inside Ford Escape by (a) occupant category and (b) seat position and occupant inside Tesla Model 3 by (c) occupant category and (d) seat position

In this study, *RMSPProp* is used as the optimizer, the learning rate is set at 0.0001. The batch size used for the training is 2, and the threshold for classification is 0.5. The weights and bias of this network are initialized every step by He initialization [116].

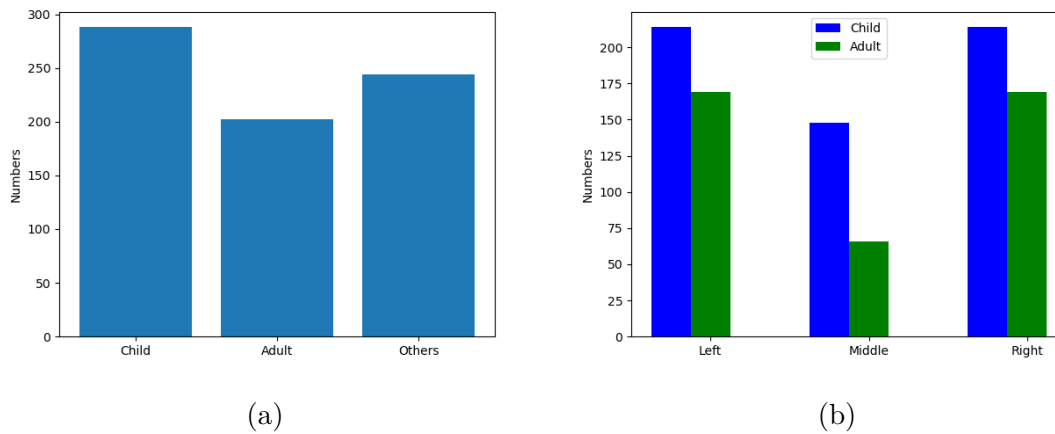
### 5.3.4 Testing

The image dataset from another vehicle, Hyundai Tucson, is selected for testing. Tucson's layout is as Fig. 5.9. The data is a randomly picked from the dataset and its statistics are shown in Fig. 5.10a and Fig. 5.10b.



**Figure 5.9.** Layout of Hyundai Tucson

Distributions of occupant inside Hyundai Tucson are shown in Fig. 5.10a and Fig. 5.10b. The total number of adults and children are balanced, and their number is at a similar level. Based on the distribution of occupant to seat position in Fig. 5.10b, the number of occupants in the middle seat is slightly smaller than the number in the left and right seats.



**Figure 5.10.** Statistics of occupants inside Hyundai Tucson by (a) occupant category and (b) seat position

## 5.4 Results and analysis

### 5.4.1 Evaluation metrics

The evaluation metrics for the study are mean intersection over union (mIoU) and cross-entropy loss. The first is to assess accuracy of masks and category for each occupant. The second is to measure the loss change of the algorithm. The mIoU is defined as Eq. (5.9).  $M$  is the number of classes. TP represents the true positive. FP is the false positive. FN is the false negative.

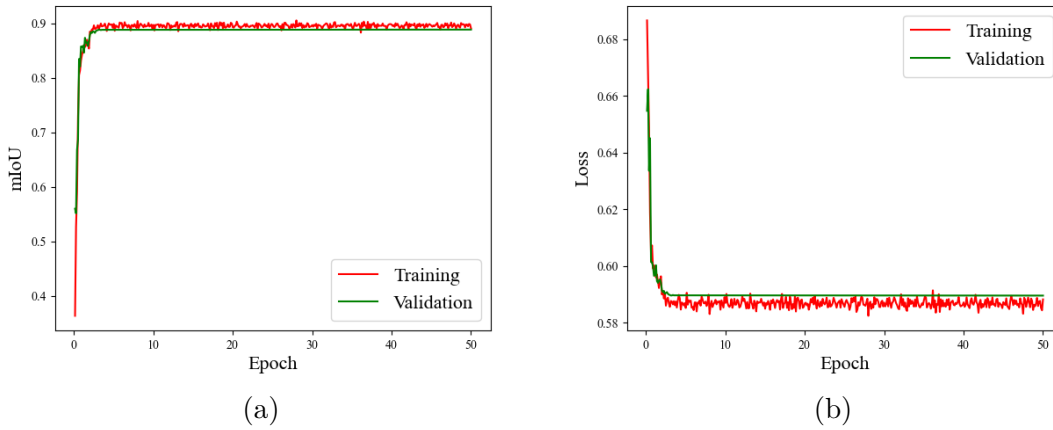
$$mIoU = \frac{1}{M} \sum_{c=1}^M \frac{TP}{TP + FP + FN} \quad (5.9)$$

The multi-class cross entropy loss is defined as Eq. (5.10) where  $y_c$  is binary indicator,  $p_c$  is the predicted probability of class  $c$ .

$$crossentropy = - \sum_{c=1}^M y_c \log(p_c) \quad (5.10)$$

### 5.4.2 Accuracy

The change of mIoU and cross-entropy loss of this training along epochs are shown in Fig. 5.11a and Fig. 5.11b. The red line indicates the change of mIoU and Loss in the training set, and the green line stands for a change of these two metrics in the validation set.



**Figure 5.11.** (a) mIoU and (b) cross entropy loss of training

The mIoU accuracy of the training set calculates the IoU of both the child and adult. As from Fig. 5.11a, the final mIoU at epoch 50 is around 89.9%, and the mIoU at epoch 50 for the validation data set is 88.9%. In Fig. 5.11b, both the training and validation loss is reduced rapidly to 0.586 and 0.589, respectively, and holds at a fixed value. The gap between the training and validation lines is acceptable in both figures, and the validation plot is kept flat over epochs. Fig. 5.12 shows an example of the testing result where both child and adult are detected and classified as purple and cyan color accordingly.

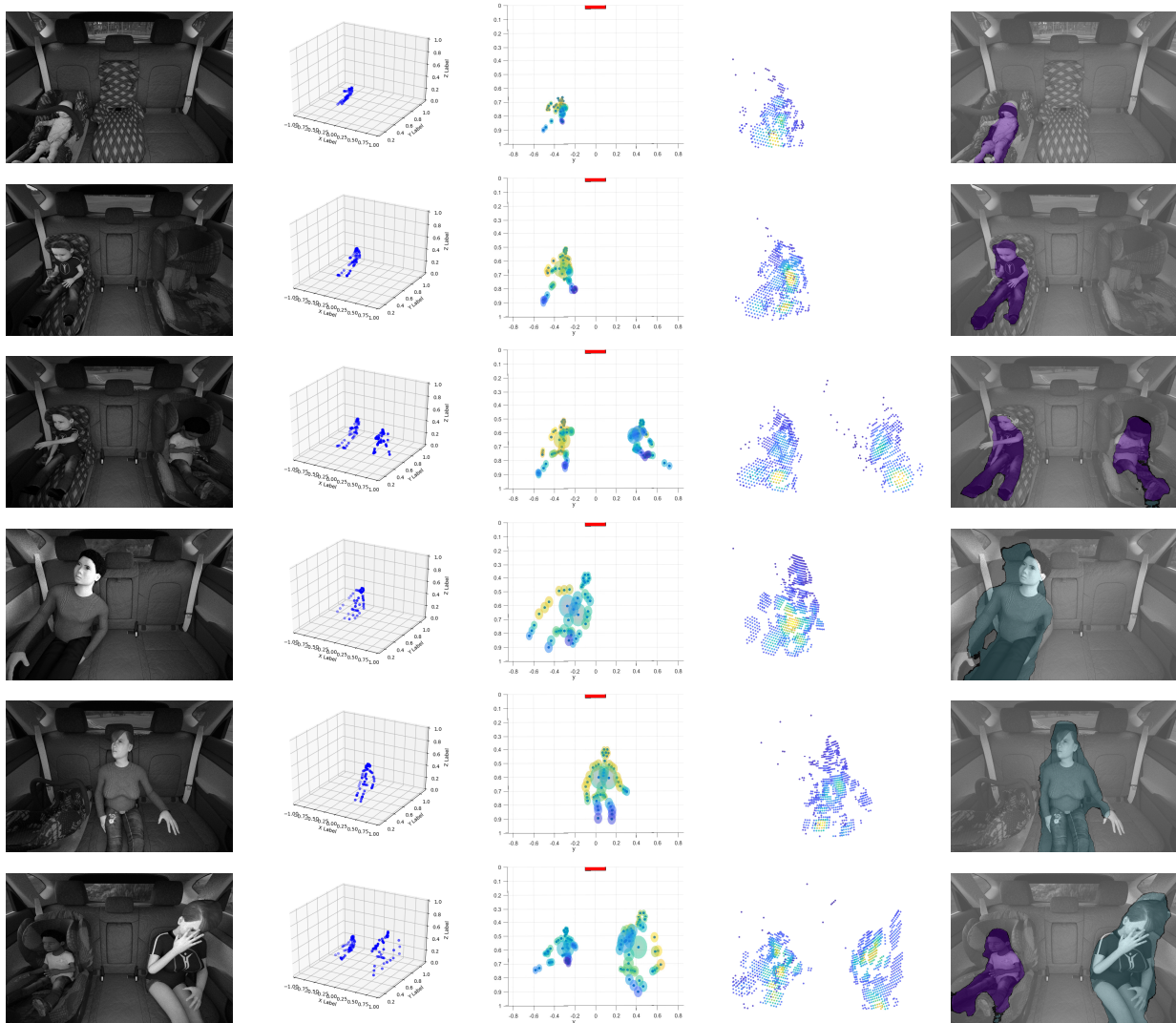


**Figure 5.12.** Output mask of data from the training dataset

The additional testing results are shown in Fig. 5.13. The first column shows the original camera image. The second column shows the proposed 3D reflection-based model based on ground truth. The third column is the radar RCS pattern of the reflection-based model. The fourth column is the radar reflection data using the proposed simulation model. The last column is the output mask of classification from the network.

The mIoU result of the testing dataset is 85.7%. This is because both the occupant's posture and the relationship position between the camera/radar posture and occupant have changed. When the model is applied to the new dataset, the prediction accuracy is decreased.

This model with the same hyper-parameters designed for the radar-based segmentation also used to assess the accuracy of camera-based semantic segmentation. The accuracy of training is around 97.6%, and the validation accuracy is approximately 97.1%. Both the



**Figure 5.13.** Pre-processed data and results from the testing data set

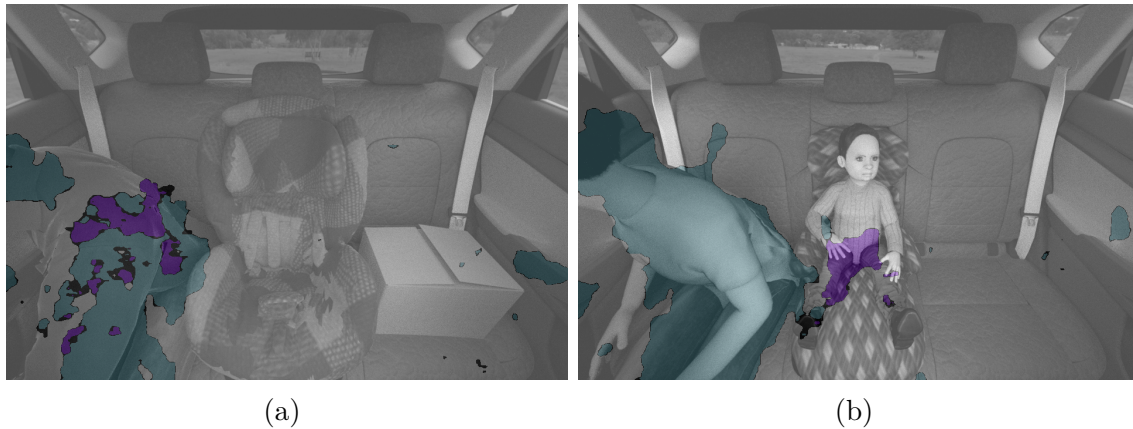
training and validation accuracy of radar-based segmentation are about 8% lower than the camera-based method. However, the overall accuracy is still acceptable and at the same level with the camera-based method. The testing accuracy of the radar-based method, on the other hand, outperforms the camera-based method. This is because the camera-based method needs to consider the texture of occupants, such as the color of clothes and hair, while the radar-based method only considers the power level of the detected point cloud. Therefore, the radar-based method is adaptable.

**Table 5.3.** Accuracy comparison between radar and camera based occupant classification

	Radar based	Camera based
Training	89.9%	<b>97.6%</b>
Validation	88.9%	<b>97.1%</b>
Testing	<b>85.7%</b>	72.8%

### 5.4.3 Failure case and discussion

As discussed previously, there are cases where the model does not work as predicted. The reason are divided into two categories. The first is the extreme posture of human body. As described in the previous section, the camera has a limited field of view for in-cabin sensing; if the occupant is leaning forward to the front, the camera often can't capture the human's whole body. If this happened, the Keypoint model and 3D reconstruction would be inaccurate. Fig. 5.14a shows an example of inaccurate detection due to the extreme posture of occupants. In this case, the body part belonging to the adult is either out of the image or occluded by other parts of the body. As in the image, the head is out of the image, and the arm covers the torso area. The model can not detect the pattern of this adult. Thus, both the classes of adult and child mixed in the image. It is difficult for algorithm to distinguish occupant types in this region.



**Figure 5.14.** (a) extreme posture and (b) lack of radar point cloud points



The second is due to a lack of radar point cloud data. As known from previous chapters, the radar imaging system highly depends on the power reflection level related to the RCS pattern and range. For the case shown in Fig. 5.14b, it is possible that the power reflection level of child is too low to be detected. Although the class is correct for this child, it is difficult for the algorithm to understand the posture according to the shape of the mask.

Some potential solutions could be used to solve these aforementioned problems. A multiview ground truth camera system might be used to obtain the complete 3D model of occupant. The radar system will get a better estimation of the power reflected from the occupant. To solve the problem of insufficient point cloud data, the power reflections of radar signals should be accumulated across multiple frames to better image the occupant and increase detection accuracy.



## 6. HUMAN SUBJECT POSTURE DETECTION

### 6.1 Introduction to posture detection

Previous research has examined using the Keypoint-based model for pedestrian motion detection. However, in this research, the main focus is to detect human subjects inside vehicle cabin. Human subjects often remain in the seat with subtle movements. Besides detecting postures, a vital-sign-based detection is added to find human subjects. In this chapter, the design of the Keypoint-based model is presented, and simulations of different postures with vital signs are shown and discussed.

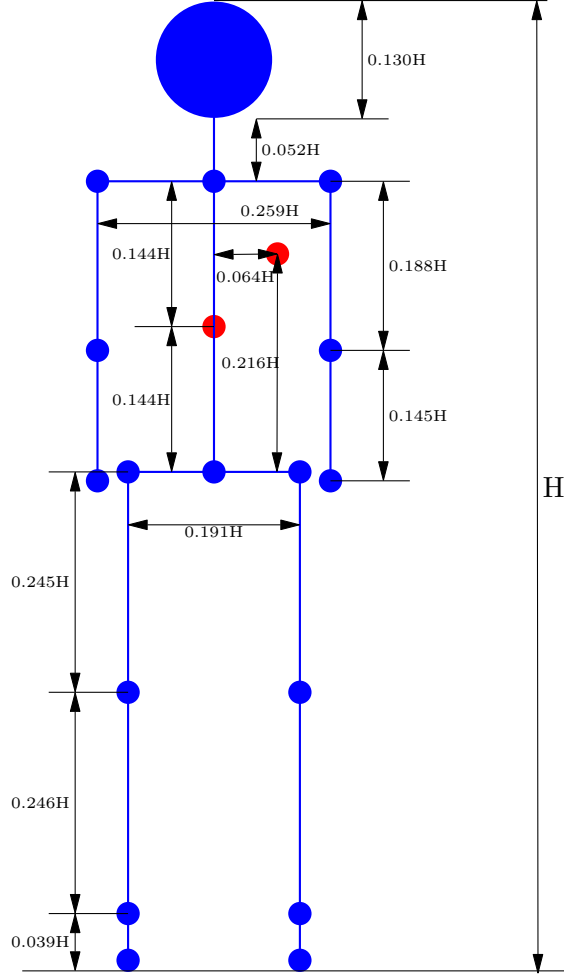
### 6.2 Keypoint-based model and imaging method

#### 6.2.1 Design of the Keypoint-based human subject model

The new 19 Keypoint skeleton model is designed based on the 17 point model [68]. The difference is that two more points are added to the torso area representing the chest and heart for sources of the vital signs as shown in Fig. 6.1. The two red circles are the designed locations for chest and heart. The location of chest is designed to be at the center of the torso and the location of heart is on the upper right of the location of chest. The sizes of the body parts are defined based on anthropometry [69][117] and projected to the total height proportionally.

Fig. 6.2a shows a standing adult model in Cartesian 3D space. The red solid square indicates radar. Each point of the model represents a reflection pattern from the corresponding body part. A sphere models the head while the ellipsoids model all other parts. The RCS pattern determines reflecting intensity according to the sizes of each body parts [118]. Orientation of each equivalent shape is defined to better approximate corresponding parts of human subjects. Besides, the model can be placed on a single 2D plane towards the radar.

Fig. 6.2b shows a sitting adult model in front of the radar. Compared with the standing model's placement, the sitting model is placed on multiple 2D planes at the axis towards



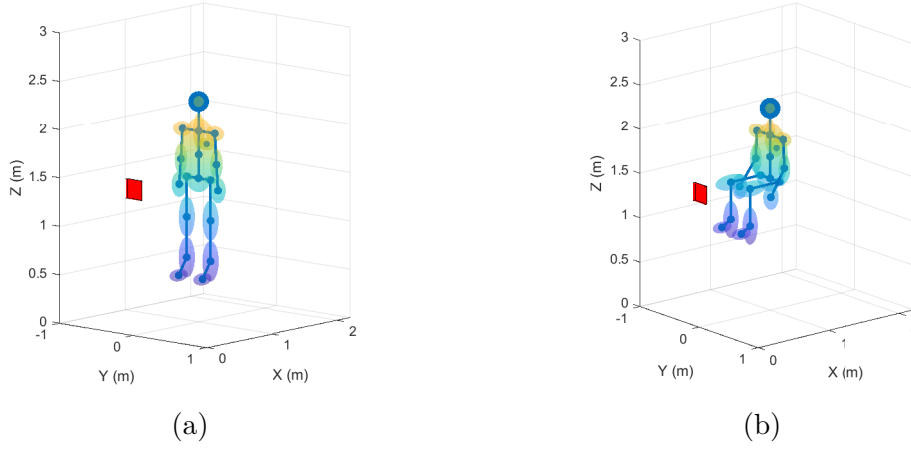
**Figure 6.1.** Keypoint model and ratio of a standing human

the radar in the 3D space. The number of the 2D planes are based on the radar resolution converted from 3D polar space.

The model can be scaled down according to the anthropometric database for a child or an infant simulation. If we want to detect where breath and heart rates come from, e.g., chest and heart, the angular resolutions for both azimuth and elevation domains should be adequate enough to distinguish a selected target from others. the angular resolution is calculated by

$$\theta_{res} = \frac{L_{arc}}{R} \quad (6.1)$$

where  $L_{arc}$  is arc length and  $R$  is the range in Polar space

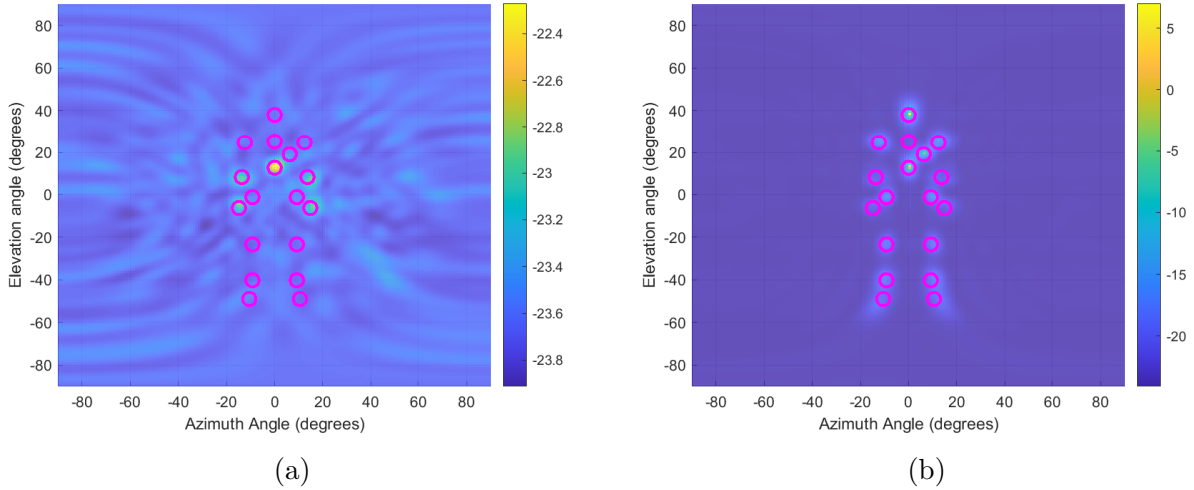


**Figure 6.2.** Keypoint adult models in 3D space (a) standing (b) sitting

Based on Eq. (6.1), the angular resolutions can be improved by increasing the number of receiving antennas.

### 6.2.2 Imaging method for posture detection

For posture estimation, the MUSIC-based method is used because the number of sources is known, and the number of sources is fixed.



**Figure 6.3.** Comparison of (a) imaging with single range bin and (b) imaging with improved algorithm

It is known that radar detection is based on a Polar coordinate where estimation is along the radial axis. The estimation of angle-of-arrival in polar coordinate will result in some inaccurate detection for objects in the same range in Cartesian space, as shown in Fig. 6.3a. To improve the estimation of Keypoints, a simple 3D high-resolution imaging method is introduced to improve the detection of human Keypoints. The basic idea is to feed the data with both the range of interest and its adjacent bins. Then we can use the data to calculate the covariance matrix to itself. Points in each radial range are calculated following a rolling buffer. This helps to distinguish Keypoints located in the different distances in 3D space. The proposed method is shown in Fig. 6.3b. The pseudo-code of this algorithm can be shown below in Algorithm 2 below.

---

**Algorithm 2** Improved key point imaging method

---

```

1: function IMPROVCROSSCORR( $X, n$ )           ▷ Where  $X$  - set of signals,  $n$  - number of
   adjacent bin
2:   for  $i = d_{start}$  to  $d_{end}$  do
3:      $X_{corr}(i) = xcorr(X(i - n : i + n), X(i - n : i + n))$ 
4:   end for
5:   return  $X_{corr}$ 
6: end function

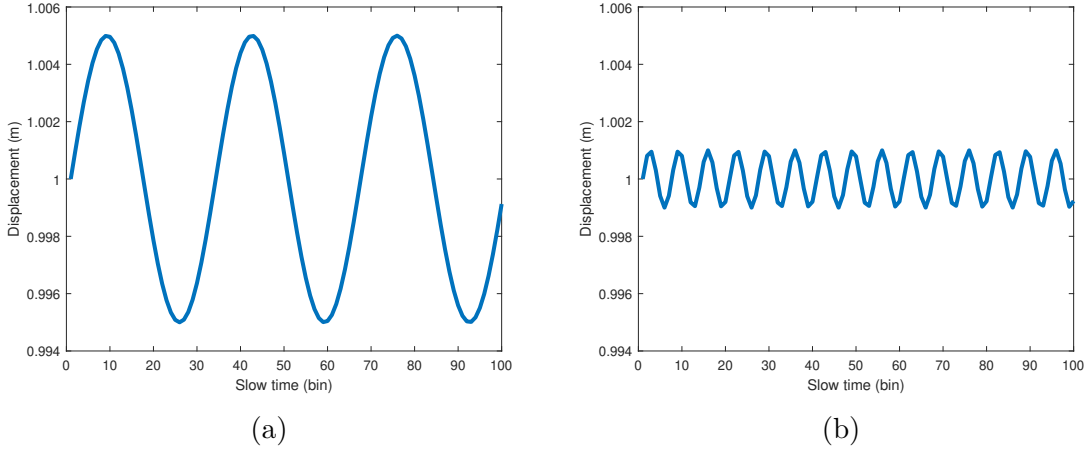
```

---

### 6.2.3 Design of vital sign signal

The vital signs indicate breath and heart rates for this study. Based on empirical data [119], the maximum breathing displacement is 5 mm, and the maximum heart displacement is 1 mm. A breath rate of 0.3 Hz and a heart rate of 1.5 Hz for the human subject model are applied. Both signals are modulated as a sinusoidal waveform as shown in Fig. 6.4a and Fig. 6.4b, respectively.

Inspired by Ahmad et al. [67], the MVDR is used toward the predefined angles where vital signs are located to calculate the corresponding weights. The raw data, i.e., processed radar signal from a selected pair of transmitting-receiving antennas, are multiplied by the weights. A range FFT is then applied to data along the slow-time axis to detect breath and heartbeat displacements.



**Figure 6.4.** The designed displacements of (a) breath signal and (b) heart signal

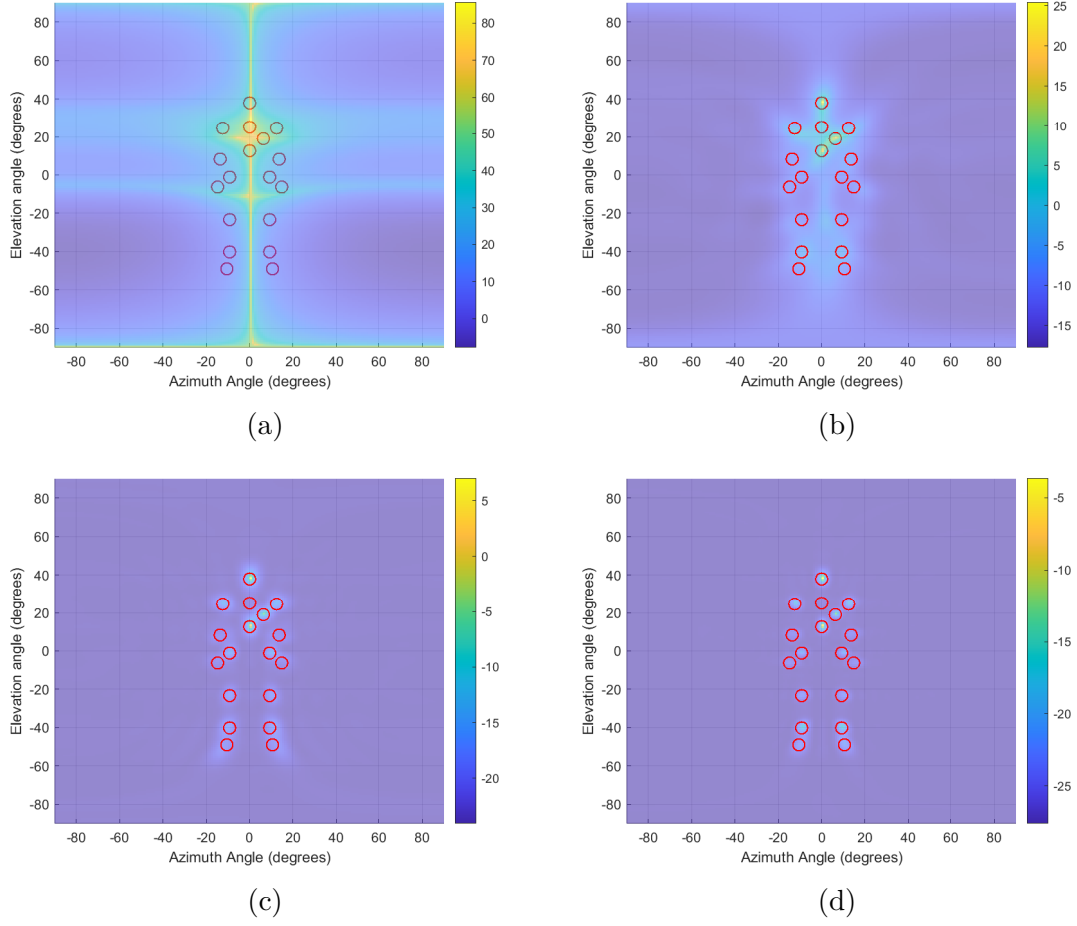
### 6.3 Simulation results

With all the designs and parameter selection, the proposed simulated 3D imaging radar system can detect the subject represented by the Keypoint-based skeleton model and its breath and heart rates. In the simulation, the center of the hips is defined as the origin of the human subject model and is neglected in the simulation as a target. Therefore, the total number of targets in this simulation is 18. This work focuses on the simulation system building for the imaging and vital signs assessment within a selected ROI. All the designed experiments are in an ideal situation without any noise considered.

#### 6.3.1 Planar radar array in different sizes

Four different array sizes,  $4 \times 4$ ,  $8 \times 8$ ,  $16 \times 16$ , and  $24 \times 24$ , with the standing adult model, are selected for the testing. The model is 1.7 meters in height and is placed at one meter away toward the antenna array, as shown in Fig. 6.2a. The noise level of these tests is assumed to be the same.

Fig. 6.5a – 6.5d show the imaging with four different numbers of antennas in the array. The sensing resolution increases while the array size grows. Only a few key points with stronger reflected intensity from the model are detected with the  $4 \times 4$  array in Fig. 6.5a. For the  $8 \times 8$  array, a rough shape of the model can be found in Fig. 6.5b. However, the



**Figure 6.5.** The imaging of a standing human with different numbers of radar antenna (a)  $4 \times 4$  (b)  $8 \times 8$  (c)  $16 \times 16$  and (d)  $24 \times 24$

lower part of the model's key points are still blurry as knees and feet can hardly be found. All 18 targets emerge clearly in Fig. 6.5c after the array size is doubled. The resolution can be further increased by adding more antennas in the array, as shown in Fig. 6.5d.

Table 6.1 shows the difference in Keypoint detection and mean square error (MSE) with different antenna array size. From the result, it can be found that the number of Keypoint detected increases with the size of the array. When the antenna size reaches  $16 \times 16$  and above, the antenna array can detect all defined Keypoints in the heatmap. In terms of the error of detection, the value of MSE can be high if the size of the antenna array is not large enough to provide enough resolution, such as the  $8 \times 8$  case in these tests. The MSE value of  $16 \times 16$  and  $24 \times 24$  are kept in a small value, and both values are acceptable.

**Table 6.1.** Detection and MSE in polar coordinate with different size of antenna array

	Point detected	MSE in degree
$4 \times 4$	3	15.19
$8 \times 8$	10	333.01
$16 \times 16$	18	0.88
$24 \times 24$	18	0.20

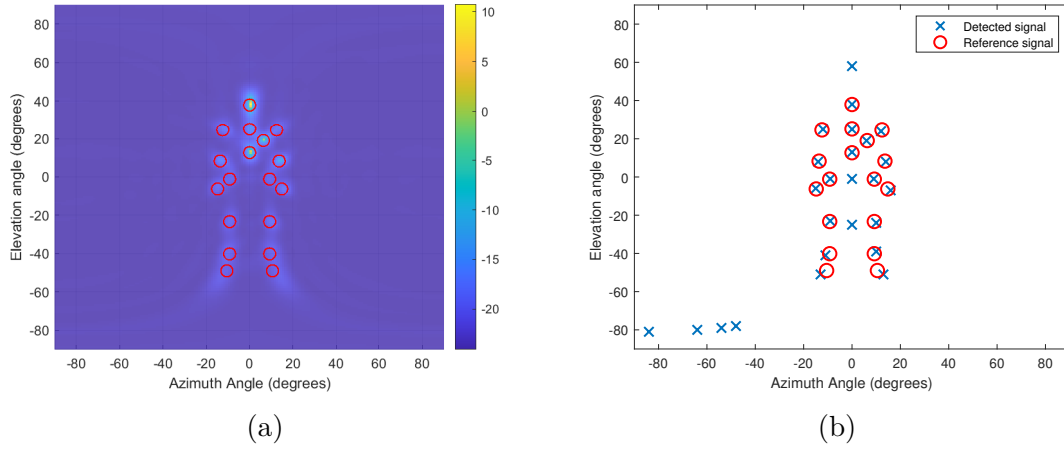
Compared with low central frequency radar, a high central frequency radar has a smaller aperture of antennas. It is feasible to pack a larger antenna array into a smaller package with a 79 GHz central frequency radar. On the other hand, increasing the number of antennas requires increasing data cube memory and computational power. Therefore, the application cases determine the resolution required for radar antenna design and system selection. The number of antennas can be further reduced by diminishing side-lobes and multi-path effects with other advanced methods, which are not included in this work.

### 6.3.2 Keypoint skeleton models and their postures

The system performance is assessed by using three testing cases with the  $16 \times 16$  antenna array. The system performance includes 3D imaging and vital signs. Two types of datasets are collected -one is for 3D imaging evaluation, and the other is for the vital sign assessment.

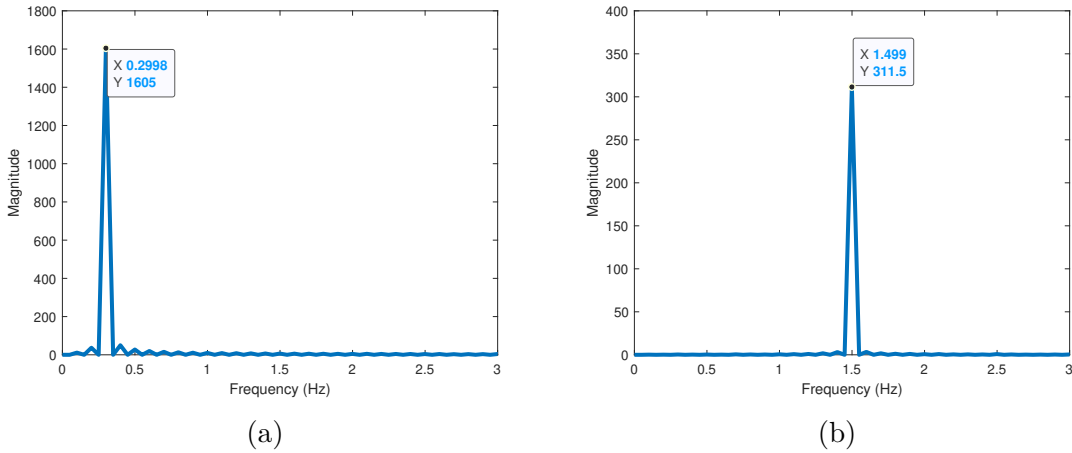
The first case is to use the adult model with the standing posture. The model is placed at one meter away and faces the radar, as shown in Fig. 6.2a. By applying the 3D imaging method, the model's reflected points are constructed in the form of 3D imaging and key points accordingly. In Fig. 6.6a, reflected points of chest and head are stronger while the others are slightly weaker. The points for feet are still detectable, even though their size is small, according to research done by Adib et al. [64]. Fig. 6.6b shows that reflected points marked as blue crosses are sensed after a simple threshold is applied. Their locations are more accurate compared with the reference points marked as red circles. Several unrelated points detected can be eliminated if certain thresholds are further fine tuned.

Signals of vital signs, i.e., breath and heart rates, can be extracted by tracking magnitude changes or phase changes along the slow-time axis. In this research, the phase changes are



**Figure 6.6.** Standing adult posture: (a) Keypoint imaging (b) Keypoint estimation

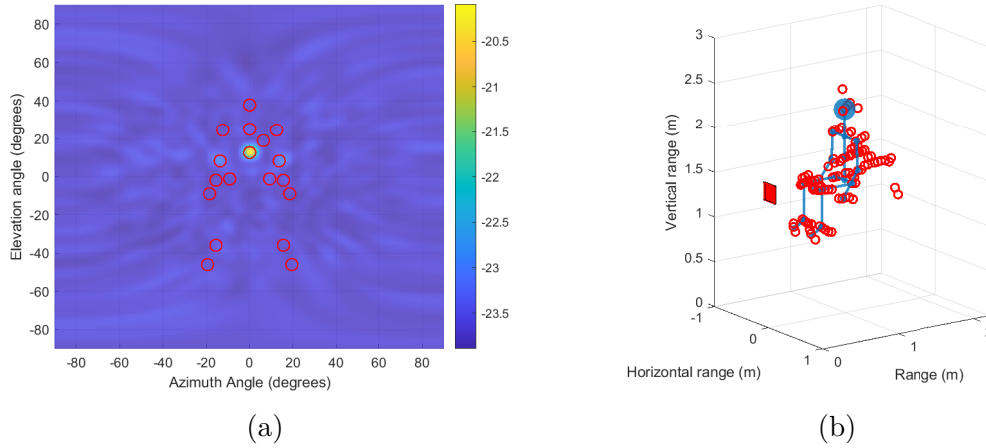
tracked by computing the phase change, unwrapping, and converting the signals to the frequency domain. Then, breath and heart rates are extracted through the beamforming of signals towards the chest and heart locations. The results are shown in Fig. 6.7a – 6.7b. The frequency of breath and heart rates are 0.3 Hz and 1.5 Hz, respectively, as previously discussed. Other signals can be barely found in the two plots since the radar angular resolution is fine enough to separate the two reflected points representing chest and heart.



**Figure 6.7.** (a) breath rate and (b) heart rate detected for a standing adult



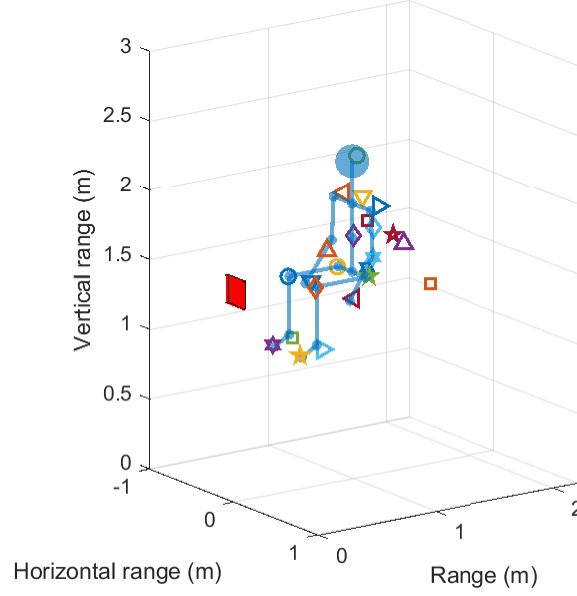
The next case is to use the adult model with sitting posture. Again, the model's central hip point is placed at one meter away, and the model faces radar, as shown in Fig. 6.2b. Unlike the standing posture, the sitting posture results of reflected points in the model spanning a bulk of the range bins. If using the same approach, i.e., estimating reflected points in a plane, on the standing model, many reflected points of the model will be lost, as shown in Fig. 6.8a. The reference key points (red circles) of the model are superimposed to 3D imaging. Thus, the MUSIC method and Algorithm 2, which have been discussed in a previous section, are applied in the searching and detecting of the reflected points in 3D space. These algorithms improve efficiency and save computation power by avoiding calculating data in all range bins. Fig. 6.8b. shows all the reflected points detected in 3D space.



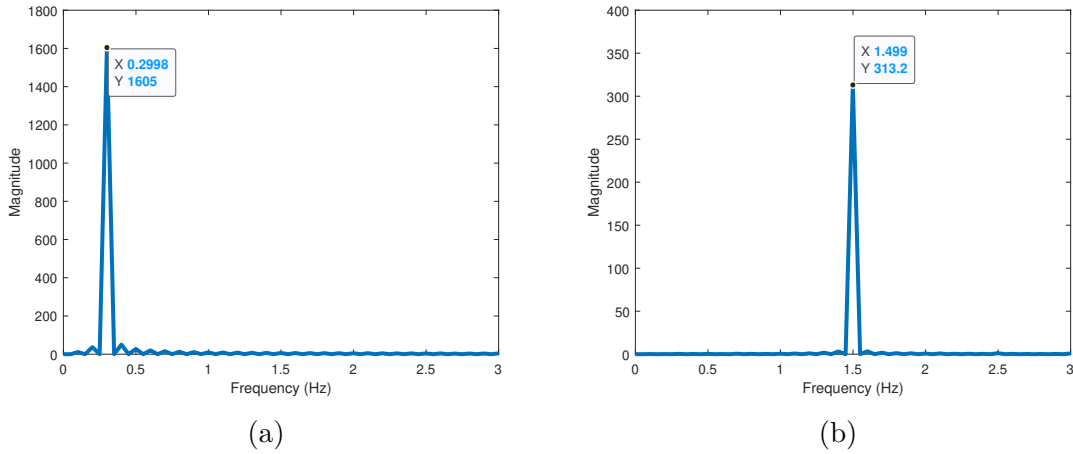
**Figure 6.8.** Sitting posture of an adult: (a) Keypoint imaging (b) Keypoint estimation

To extract points potentially belonging to the model, the density-based spatial clustering of applications with noise (DBSCAN) algorithm is used, which is an unsupervised machine-learning-based method to cluster the detected points belonging to different categories in 3D space. The relative distance between each point is calculated as inputs of the method.  $\epsilon$  is set at 0.1, and the minimum points for one category are set at 2 because the detected point cloud is sparse. The center point of each category computed by the algorithm can be used for the Keypoint-based skeleton model estimation, as shown in Fig. 6.9. Further improvement can be achieved by applying a supervised machine-learning-based method. The results of

breath and heart rate calculations are shown in Fig. 6.10a – 6.10b, similar to the standing adult results.



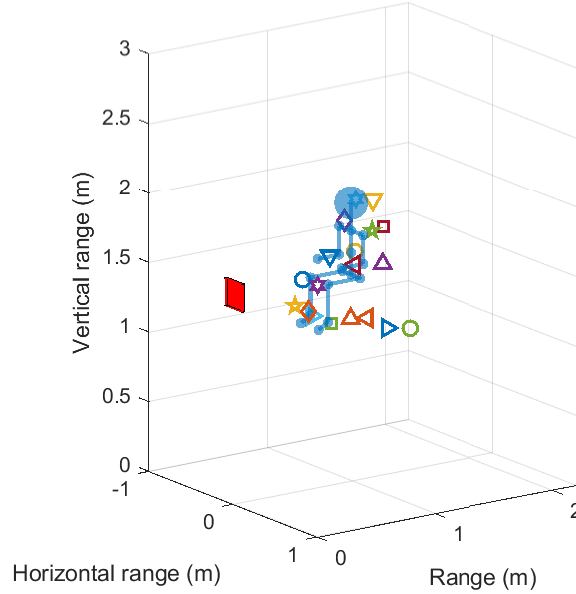
**Figure 6.9.** Sitting adult Keypoints detected with DBSCAN



**Figure 6.10.** (a) breath rate and (b) heart rate detected for the sitting adult

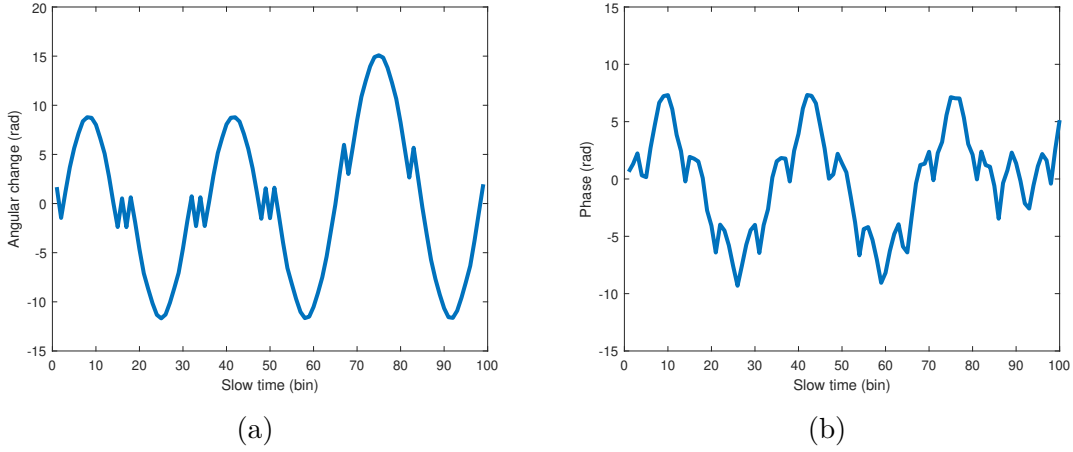
The last case is to use the child model in the sitting posture. The setup is the same as the previous cases. As discussed in the previous section, the lengths and widths of the skeleton model's Keypoints are designed in proportion to the total height. The child model's

total height is 1.1 meters, about 65% of the adult model height. Thus, the relative distances between each point is much shorter than the adult. The sensor resolutions for the approach thus have difficulty locating some of the key points accurately, as shown in Fig. 6.11.

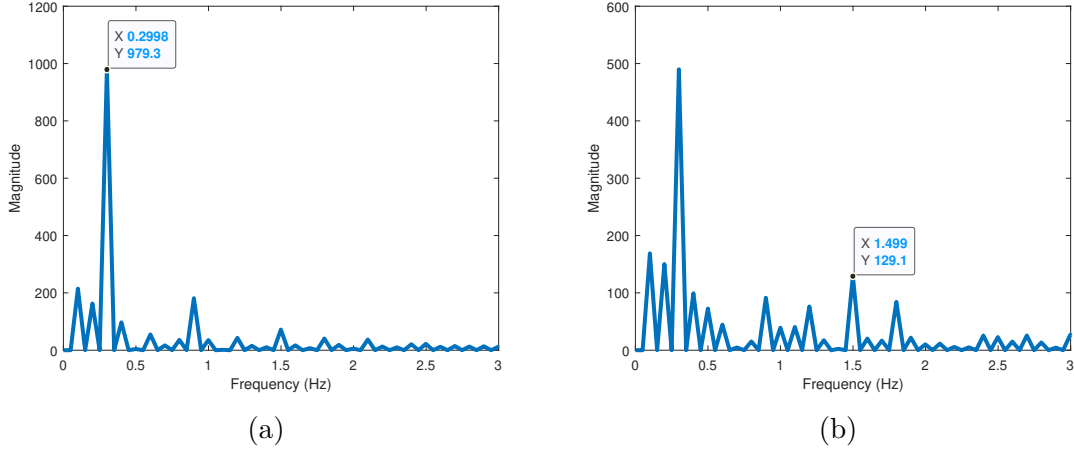


**Figure 6.11.** Sitting child Keypoints detected with DBSCAN

Different from the two previous cases, reflected signals received from either chest or heart location contain both breath and heart rate information. The radar angular resolution is not fine enough to separate these two reflected points representing chest and heart for the child model, as shown in Fig. 6.12a – 6.12b. In this case, appropriate bandpass filters and thresholds can be applied to capture breath and heart rates in a frequency domain as in Fig. 6.13a – 6.13b.



**Figure 6.12.** Unfiltered (a) breath signal and (b) heart signal of a sitting child



**Figure 6.13.** (a) breath rate and (b) heart rate for the sitting child

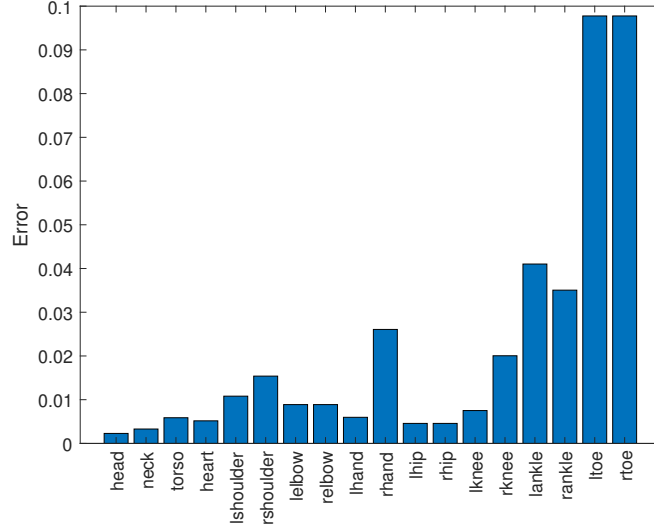
### 6.3.3 Accuracy and analysis

The accuracy of estimation is achieved through measuring the Euclidean distance between the predicted Keypoint and the ground truth. For the 2D case, the error is calculated by Eq. (6.2), where the distance between each point on z-axis and the radar is the same.

$$2D\_error = \sqrt{(x - x_0)^2 + (y - y_0)^2} \quad (6.2)$$

In Fig. 6.6a, a standing posture is described, and the human model's Keypoints are estimated as in Fig. 6.6b. The errors of all detected Keypoint signal and the reference signal

is calculated by Eq. (6.2). The result is shown in Fig. 6.14, where errors of all points are in low level except the errors of both the left and right toe. This is because the two points are located at the edge of the main lobe, so the AoA estimation is not accurate.



**Figure 6.14.** 2D error of the standing posture case

The Euclidean distance for 3D point cloud is calculated by Eq. (6.3), where  $x, y, z$  are the predicted points and  $x_0, y_0, z_0$  are the ground truth points.

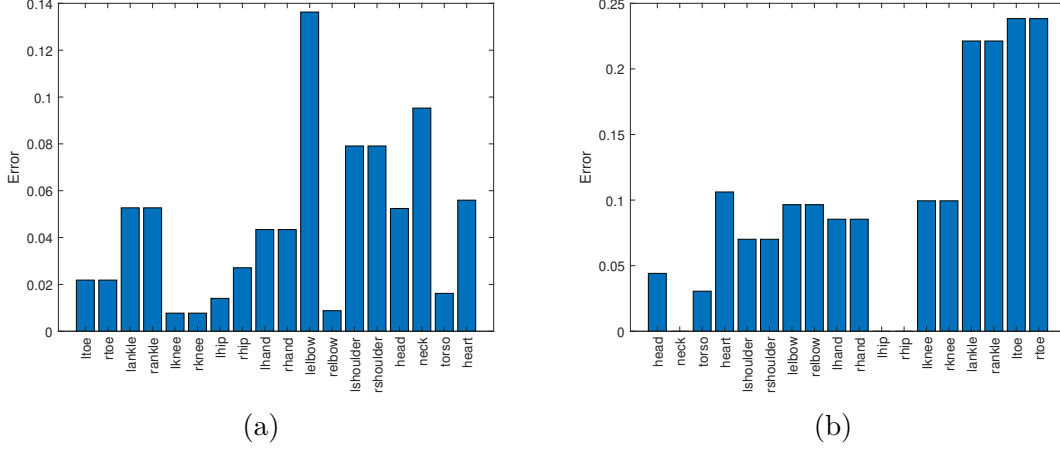
$$3D\_error = \sqrt{(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2} \quad (6.3)$$

The overall performance of the three tests is concluded in Table 6.2. In all three tests, most of the Keypoints are detected. The MSE values of all three tests are based on the detection in Cartesian coordinate. The MSE values in the Table 6.2 of all three tests are low, meaning the detected points are close to the reference points. Therefore, the simulation system successfully simulates postures of occupants based on the designed Keypoint model.

Based on the result of the first 3D test with a sitting adult in Fig. 6.9, accuracy of all Keypoints, which is kept at a low level, is listed in Fig. 6.15a. There is only one point with a relatively higher error. This indicates that the unsupervised learning of the Keypoints may not be robust in some estimations.

**Table 6.2.** Performance summary of the simulation tests in Cartesian coordinate

Test	Point detected	MSE in distance
2D case	18	$6.62 \times 10^{-4}$
3D sitting adult	18	$8.66 \times 10^{-4}$
3D sitting child	15	$4.5 \times 10^{-3}$

**Figure 6.15.** 3D error of (a) the sitting adult and (b) the sitting child

Compared to the estimation of adult, the average estimation error of Keypoints of a sitting child in Fig. 6.11 is higher. From results in both Fig. 6.15b and Fig. 6.11, some Keypoints of the sitting child are not detected. The value of those undetected points is set to zero in Fig. 6.11. Particularly for this test, the Keypoints of neck, left hip, and right hip are not detected. These points are undetected because the body size is smaller, and the angular resolution is not fine enough to distinguish different Keypoints. For other points, the detection error is also high because of the insufficient angular resolution of the antenna array.

The accuracy of the vital sign is based on the frequency. For vehicle occupant monitoring, the purpose of using the vital sign is to distinguish living subjects from non-living subjects. Therefore, it is adequate to only measure the vital sign frequency rather than the shape of the vital sign signals. In both cases, the accuracy of breath rate and heart rate is around 99.9%. However, the sitting child's frequency domain result shows that the heart rate signal can be affected by other signals since the amplitude of adjacent bins is high enough to affect

estimation results. In other words, noises in vehicle can affect the estimation accuracy of vital sign signals. A more advanced algorithm needs to be designed to track and filter out other noise signals to solve this problem.

## 7. CONCLUSION AND FUTURE WORK

### 7.1 Conclusion

In this dissertation, a planar array radar system for occupant imaging and vital sign detection has been built for the purpose of in-cabin occupant monitoring.

To simulate the power reflection pattern of occupants, a mathematical model is derived. A signal processing pipeline is then proposed based on the mathematical model, which includes the processor, the antenna array, the free space propagation, and the subject under test. A simulation system is then built based on the pipeline to simulate the radar system. Both the posture of occupants and their vital sign signals are estimated. The vital sign signal is to evaluate the displacement of chest or belly of occupants.

For the classification, a human subject simulation model is developed. Based on this model, a supervised deep learning method, which combines camera image information with radar image information, is introduced for occupant segmentation and classification. The accuracy of test and validation is around 90%, and the testing accuracy is approximately 87%, which are compatible to the performance of the camera-based method for classification.

The posture detection and vital sign estimation of occupant monitoring are based on a proposed Keypoint-based model. Two more points are added to the previous 17 points Keypoint body model to reflect the movement of these two areas. Occupant posture can be sensed by either a machine-learning-based method or a deep-learning-based method. For the Keypoint-based method, a machine-learning-based method is used to cluster the detected points and combine detected points in different categories for posture detection. In the simulation system, an antenna array with a size of  $16 \times 16$  is selected for human body posture imaging. Then a sitting adult and a sitting child are tested using a different human model. The results in the simulation system show that most of the Keypoints belonging to occupants are detected. The accuracy of vital sign estimation is acceptable.



## 7.2 Future work

The future work of this research is threefold. The first is to introduce more models and sitting positions into the system to simulate their reflections, such as models for pets and people in front seats. Posture and vital signs of occupants in front seats need to be considered. The only difference between occupants at front and in the rear is their direction towards radar. In this dissertation, a top-mounted radar antenna array system is discussed, in which occupants in front seats have their back to the radar antenna array. It is possible to measure vital signs and provide a point cloud image of occupants in front seats. However, it is hard to use both information collected by radar antenna array in post-processing, as vital sign signals measured from the back are weak and mixed with noises, and the point cloud image of adult is not complete because many signals are blocked by vehicle seats.

The second is to improve the algorithm performance for the classification enhancement. In this dissertation, children and infants are not distinguished because they have similar sizes. Nonetheless, both posture and facing direction of children and infants are different. It is difficult to distinguish them because their power reflection patterns are similar. For example, power reflection patterns of forward-facing children and infants are similar. A 3D model-based model to detect occupants' sitting direction may be able to solve this problem.

Another work that can be done in the future is to develop a real-time Keypoint-based posture detection algorithm to monitor posture and vital signs of different occupants. This system can be used to improve safety and ride experience of the future intelligent vehicle system.

## REFERENCES

- [1] E. NCAP. (). “2025 roadmap,” [Online]. Available: <http://www.euroncap.com/en/for-engineers/technical-papers/>. (Last accessed: 02.03.2019).
- [2] J. Null. (). “Heatstroke deaths of children in vehicles,” [Online]. Available: <https://www.noheatstroke.org/>. (Last accessed: 01.20.2019).
- [3] 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 Transactions on microwave theory and techniques*, vol. 61, no. 5, pp. 2046–2060, 2013.
- [4] M. Zakrzewski, H. Raittinen, and J. Vanhala, “Comparison of center estimation algorithms for heart and respiration monitoring with microwave doppler radar,” *IEEE Sensors Journal*, vol. 12, no. 3, pp. 627–634, 2011.
- [5] Y. Xiao, J. Lin, O. Boric-Lubecke, and M. Lubecke, “Frequency-tuning technique for remote detection of heartbeat and respiration using low-power double-sideband transmission in the ka-band,” *IEEE Transactions on Microwave Theory and Techniques*, vol. 54, no. 5, pp. 2023–2032, 2006.
- [6] A. D. Droitcour, O. Boric-Lubecke, and G. T. Kovacs, “Signal-to-noise ratio in doppler radar system for heart and respiratory rate measurements,” *IEEE transactions on microwave theory and techniques*, vol. 57, no. 10, pp. 2498–2507, 2009.
- [7] W. Massagram, V. M. Lubecke, A. HØst-Madsen, and O. Boric-Lubecke, “Assessment of heart rate variability and respiratory sinus arrhythmia via doppler radar,” *IEEE Transactions on microwave theory and techniques*, vol. 57, no. 10, pp. 2542–2549, 2009.
- [8] H. Avagyan, A. Hakhoumian, H. Hayrapetyan, N. Pogosyan, and T. Zakaryan, “Portable non-contact microwave doppler radar for respiration and heartbeat sensing,” *Armenian Journal of Physics*, vol. 5, no. 1, pp. 8–14, 2012.
- [9] J. E. Kiriazi, O. Boric-Lubecke, and V. M. Lubecke, “Dual-frequency technique for assessment of cardiopulmonary effective rcs and displacement,” *IEEE Sensors Journal*, vol. 12, no. 3, pp. 574–582, 2011.
- [10] W. Hu, Z. Zhao, Y. Wang, H. Zhang, and F. Lin, “Noncontact accurate measurement of cardiopulmonary activity using a compact quadrature doppler radar sensor,” *IEEE Transactions on Biomedical Engineering*, vol. 61, no. 3, pp. 725–735, 2013.

- [11] J.-Q. Wang, C.-X. Zheng, X.-J. Jin, G.-H. Lu, H.-B. Wang, and A.-S. Ni, "Study on a non-contact life parameter detection system using millimeter wave.," *Hang tian yi xue yu yi xue gong cheng= Space medicine & medical engineering*, vol. 17, no. 3, pp. 157–161, 2004.
- [12] C. Li, Y. Xiao, and J. Lin, "Experiment and spectral analysis of a low-power *ka*-band heartbeat detector measuring from four sides of a human body," *IEEE Transactions on Microwave Theory and Techniques*, vol. 54, no. 12, pp. 4464–4471, 2006.
- [13] Y. Xiao, J. Lin, O. Boric-Lubecke, and M. Lubecke, "Frequency-tuning technique for remote detection of heartbeat and respiration using low-power double-sideband transmission in the *ka*-band," *IEEE Transactions on Microwave Theory and Techniques*, vol. 54, no. 5, pp. 2023–2032, 2006.
- [14] J.-H. Lee, J. M. Hwang, D. H. Choi, and S.-O. Park, "Noninvasive biosignal detection radar system using circular polarization," *IEEE Transactions on Information Technology in Biomedicine*, vol. 13, no. 3, pp. 400–404, 2009.
- [15] S. Suzuki, T. Matsui, M. Kagawa, T. Asao, and K. Kotani, "An approach to a non-contact vital sign monitoring using dual-frequency microwave radars for elderly care," *Journal of Biomedical Science and Engineering*, vol. 6, no. 07, p. 704, 2013.
- [16] J. Wang, X. Wang, Z. Zhu, J. Huangfu, C. Li, and L. Ran, "1-d microwave imaging of human cardiac motion: An ab-initio investigation," *IEEE Transactions on Microwave Theory and Techniques*, vol. 61, no. 5, pp. 2101–2107, 2013.
- [17] J. Muñoz-Ferreras, Z. Peng, R. Gómez-García, and C. Li, "Review on advanced short-range multimode continuous-wave radar architectures for healthcare applications," *IEEE Journal of Electromagnetics, RF and Microwaves in Medicine and Biology*, vol. 1, no. 1, pp. 14–25, 2017, ISSN: 2469-7249.
- [18] V. Nguyen, "Non-contact physiologically related motion sensing with ultra-wideband impulse radar," Thesis, Georgia Institute of Technology, 2016.
- [19] K. Mostov, E. Liptsen, and R. Boutchko, "Medical applications of shortwave fm radar: Remote monitoring of cardiac and respiratory motion," *Medical physics*, vol. 37, no. 3, pp. 1332–1338, 2010.
- [20] L. Ren, "Noncontact vital signs detection," *PhD dissertation*, 2017.
- [21] F. Adib, H. Mao, Z. Kabelac, D. Katabi, and R. C. Miller, "Smart homes that monitor breathing and heart rate," in *Proceedings of the 33rd annual ACM conference on human factors in computing systems*, ACM, 2015, pp. 837–846, ISBN: 1450331459.

- [22] A. Nezirović, “Trapped-victim detection in post-disaster scenarios using ultra-wideband radar,” Thesis, TU Delft, Delft University of Technology, 2010.
- [23] N. Maaref, P. Millot, C. Pichot, and O. Picon, “A study of uwb fm-cw radar for the detection of human beings in motion inside a building,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 5, pp. 1297–1300, 2009.
- [24] Y. Tian, G.-H. Lee, H. He, C.-Y. Hsu, and D. Katabi, “Rf-based fall monitoring using convolutional neural networks,” *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 2, no. 3, pp. 1–24, 2018.
- [25] M. Zhao, F. Adib, and D. Katabi, “Emotion recognition using wireless signals,” in *Proceedings of the 22nd Annual International Conference on Mobile Computing and Networking*, ACM, 2016, pp. 95–108.
- [26] Y. Zhu, Z. Xiao, Y. Chen, Z. Li, M. Liu, B. Y. Zhao, and H. Zheng, “Adversarial wifi sensing,” *arXiv preprint arXiv:1810.10109*, 2018.
- [27] Z. Li, Z. Xiao, Y. Zhu, I. Pattarachanyakul, B. Y. Zhao, and H. Zheng, “Adversarial localization against wireless cameras,” in *Proceedings of the 19th International Workshop on Mobile Computing Systems & Applications*, ACM, 2018, pp. 87–92, ISBN: 1450356303.
- [28] Z. Yang, M. Bocca, V. Jain, and P. Mohapatra, “Contactless breathing rate monitoring in vehicle using uwb radar,” in *Proceedings of the 7th International Workshop on Real-World Embedded Wireless Systems and Networks*, ACM, 2018, pp. 13–18.
- [29] M. Yang, X. Yang, L. Li, and L. Zhang, “In-car multiple targets vital sign monitoring using location-based vmd algorithm,” in *2018 10th International Conference on Wireless Communications and Signal Processing (WCSP)*, IEEE, 2018, pp. 1–6, ISBN: 1538661195.
- [30] E. Schires, P. Georgiou, and T. S. Lande, “Vital sign monitoring through the back using an uwb impulse radar with body coupled antennas,” *IEEE Transactions on Biomedical Circuits and Systems*, vol. 12, no. 2, pp. 292–302, 2018, ISSN: 1932-4545. DOI: [10.1109/TBCAS.2018.2799322](https://doi.org/10.1109/TBCAS.2018.2799322).
- [31] S. A. T. Hosseini and H. Amindavar, “Uwb radar signal processing in measurement of heartbeat features,” in *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, 2017, pp. 1004–1007.
- [32] H.-S. Cho and Y.-J. Park, “Accurate heartbeat frequency extraction method using uwb impulse radar,” *IEIE Transactions on Smart Processing & Computing*, vol. 6, no. 4, pp. 246–252, 2017.

- [33] X. Hu, L. Qiu, D. Zhao, Q. Chen, R. Qian, and T. Jin, “Acceleration-based algorithm for long monitoring the micro motions on a stationary subject using uwb radar,” in *2017 International Conference on Circuits, Devices and Systems (ICCCDS)*, IEEE, 2017, pp. 205–208.
- [34] F. Khan and S. Cho, “A detailed algorithm for vital sign monitoring of a stationary/non-stationary human through ir-uwb radar,” *Sensors*, vol. 17, no. 2, p. 290, 2017.
- [35] H. Shen, C. Xu, Y. Yang, L. Sun, Z. Cai, L. Bai, E. Clancy, and X. Huang, “Respiration and heartbeat rates measurement based on autocorrelation using ir-uwb radar,” *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 65, no. 10, pp. 1470–1474, 2018.
- [36] I. Walterscheid and G. E. Smith, “Respiration and heartbeat monitoring using a distributed pulsed mimo radar,” in *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2017, pp. 3449–3452. DOI: [10.1109/EMBC.2017.8037598](https://doi.org/10.1109/EMBC.2017.8037598).
- [37] I. Bekkerman and J. Tabrikian, “Target detection and localization using mimo radars and sonars,” *IEEE Transactions on Signal Processing*, vol. 54, no. 10, pp. 3873–3883, 2006.
- [38] D. Kocur, D. Novák, and J. Demčák, “A joint localization and breathing rate estimation of static persons using uwb radar,” in *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 2017, pp. 1728–1733. DOI: [10.1109/SMC.2017.8122865](https://doi.org/10.1109/SMC.2017.8122865).
- [39] M. Muragaki, S. Okumura, K. Maehara, T. Sakamoto, M. Yoshioka, K. Inoue, T. Fukuda, H. Sakai, and T. Sato, “Noncontact respiration monitoring of multiple closely positioned patients using ultra-wideband array radar with adaptive beamforming technique,” in *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, 2017, pp. 1118–1122.
- [40] S. Yue, H. He, H. Wang, H. Rahul, and D. Katabi, “Extracting multi-person respiration from entangled rf signals,” *Proceedings of the ACM on Interactive, Mobile, Wearable and Technologies, Ubiquitous*, vol. 2, no. 2, p. 86, 2018.
- [41] J. W. Choi and S. H. Cho, “A new multi-human detection algorithm using an ir-uwb radar system,” in *Third International Conference on Innovative Computing Technology (INTECH 2013)*, IEEE, 2013, pp. 467–472.

- [42] J. Rovňáková and D. Kocur, “Toa estimation and data association for through-wall tracking of moving targets,” *EURASIP Journal on Wireless Communications and Networking*, vol. 2010, no. 1, p. 420 767, 2010, ISSN: 1687-1499. DOI: [10.1155/2010/420767](https://doi.org/10.1155/2010/420767).
- [43] X. Zhang, L. Xu, L. Xu, and D. Xu, “Direction of departure (dod) and direction of arrival (doa) estimation in mimo radar with reduced-dimension music,” *IEEE communications letters*, vol. 14, no. 12, pp. 1161–1163, 2010.
- [44] C. Qian, L. Huang, and H.-C. So, “Improved unitary root-music for doa estimation based on pseudo-noise resampling,” *IEEE Signal Processing Letters*, vol. 21, no. 2, pp. 140–144, 2014.
- [45] D. Novák and D. Kocur, “Multiple static person localization based on respiratory motion detection by uwb radar,” in *2016 26th International Conference Radioelektronika (RADIOELEKTRONIKA)*, 2016, pp. 252–257. DOI: [10.1109/RADIOELEK.2016.7477386](https://doi.org/10.1109/RADIOELEK.2016.7477386).
- [46] X. Hu and T. Jin, “Preliminary results of noncontact respiration and heartbeat detection using ir-uwb radar,” in *2016 First IEEE International Conference on Computer Communication and the Internet (ICCCI)*, 2016, pp. 320–323. DOI: [10.1109/CCI.2016.7778934](https://doi.org/10.1109/CCI.2016.7778934).
- [47] M. Baboli, O. Boric-Lubecke, and V. Lubecke, “A new algorithm for detection of heart and respiration rate with uwb signals,” in *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, IEEE, 2012, pp. 3947–3950.
- [48] M. Adjrad, S. Dudley, and M. Ghavami, “Experimental vital signs estimation using commercially available ir-uwb radar,” in *2014 International Radar Conference*, IEEE, 2014, pp. 1–4.
- [49] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu, “The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis,” *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, vol. 454, no. 1971, pp. 903–995, 1998. DOI: [10.1098/rspa.1998.0193](https://doi.org/10.1098/rspa.1998.0193).
- [50] Z. WU and N. E. HUANG, “Ensemble empirical mode decomposition: A noise-assisted data analysis method,” *Advances in adaptive data analysis*, vol. 01, no. 01, pp. 1–41, 2009. DOI: [10.1142/s1793536909000047](https://doi.org/10.1142/s1793536909000047).

- [51] M. E. Torres, M. A. Colominas, G. Schlotthauer, and P. Flandrin, "A complete ensemble empirical mode decomposition with adaptive noise," in *2011 IEEE international conference on acoustics, speech and signal processing (ICASSP)*, IEEE, 2011, pp. 4144–4147.
- [52] X. Hu and T. Jin, "Short-range vital signs sensing based on eemd and cwt using ir-uwb radar," *Sensors (Basel, Switzerland)*, vol. 16, no. 12, p. 2025, 2016, ISSN: 1424-8220. DOI: [10.3390/s16122025](https://doi.org/10.3390/s16122025).
- [53] M. A. Colominas, G. Schlotthauer, and M. E. Torres, "Improved complete ensemble emd: A suitable tool for biomedical signal processing," *Biomedical Signal Processing and Control*, vol. 14, pp. 19–29, 2014, ISSN: 1746-8094. DOI: <https://doi.org/10.1016/j.bspc.2014.06.009>.
- [54] K. Shyu, L. Chiu, P. Lee, T. Tung, and S. Yang, "Detection of breathing and heart rates in uwb radar sensor data using fvpief based two-layer eemd," *IEEE Sensors Journal*, pp. 1–1, 2018, ISSN: 1530-437X. DOI: [10.1109/JSEN.2018.2878607](https://doi.org/10.1109/JSEN.2018.2878607).
- [55] S. Bakhtiari, S. Liao, T. Elmer, A. Raptis, *et al.*, "A real-time heart rate analysis for a remote millimeter wave iq sensor," *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 6, pp. 1839–1845, 2011.
- [56] E. Pittella, A. Bottiglieri, S. Pisa, and M. Cavagnaro, "Cardiorespiratory frequency monitoring using the principal component analysis technique on uwb radar signal," *International Journal of Antennas and Propagation*, vol. 2017, 2017.
- [57] M. Donelli, "A rescue radar system for the detection of victims trapped under rubble based on the independent component analysis algorithm," *Progress In Electromagnetics Research*, vol. 19, pp. 173–181, 2011.
- [58] Y. Uomoto and A. Kajiwarra, "Heartbeat monitoring uwb sensor robust to body movement," in *2018 IEEE 4th World Forum on Internet of Things (WF-IoT)*, 2018, pp. 280–285. DOI: [10.1109/WF-IoT.2018.8355194](https://doi.org/10.1109/WF-IoT.2018.8355194).
- [59] X. Hu, L. Qiu, D. Zhao, Q. Chen, R. Qian, and T. Jin, "Acceleration-based algorithm for long monitoring the micro motions on a stationary subject using uwb radar," in *2017 International Conference on Circuits, Devices and Systems (ICCCDS)*, 2017, pp. 205–208. DOI: [10.1109/ICCCDS.2017.8120479](https://doi.org/10.1109/ICCCDS.2017.8120479).
- [60] F. Khan and S. H. Cho, "A detailed algorithm for vital sign monitoring of a stationary/non-stationary human through ir-uwb radar," *Sensors*, vol. 17, no. 2, p. 290, 2017.

- [61] J. Muñoz-Ferreras, Z. Peng, R. Gómez-García, and C. Li, "Random body movement mitigation for fmcw-radar-based vital-sign monitoring," in *2016 IEEE Topical Conference on Biomedical Wireless Technologies, Networks, and Sensing Systems (BioWireless)*, 2016, pp. 22–24. DOI: [10.1109/BIOWIRELESS.2016.7445551](https://doi.org/10.1109/BIOWIRELESS.2016.7445551).
- [62] F.-K. Wang, T.-S. Horng, K.-C. Peng, J.-K. Jau, J.-Y. Li, and C.-C. Chen, "Single-antenna doppler radars using self and mutual injection locking for vital sign detection with random body movement cancellation," *IEEE Transactions on Microwave Theory and Techniques*, vol. 59, no. 12, pp. 3577–3587, 2011.
- [63] X. Yu, C. Li, and J. Lin, "Two-dimensional noncontact vital sign detection using doppler radar array approach," in *2011 IEEE MTT-S International Microwave Symposium*, IEEE, 2011, pp. 1–4.
- [64] F. Adib, C.-Y. Hsu, H. Mao, D. Katabi, and F. Durand, "Capturing the human figure through a wall," *ACM Transactions on Graphics (TOG)*, vol. 34, no. 6, p. 219, 2015.
- [65] M. Zhao, T. Li, M. Abu Alsheikh, Y. Tian, H. Zhao, A. Torralba, and D. Katabi, "Through-wall human pose estimation using radio signals," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 7356–7365.
- [66] P. Zhao, C. X. Lu, J. Wang, C. Chen, W. Wang, N. Trigoni, and A. Markham, "Mid: Tracking and identifying people with millimeter wave radar," in *2019 15th International Conference on Distributed Computing in Sensor Systems (DCOSS)*, IEEE, 2019, pp. 33–40.
- [67] A. Ahmad, J. C. Roh, D. Wang, and A. Dubey, "Vital signs monitoring of multiple people using a fmcw millimeter-wave sensor," in *2018 IEEE Radar Conference (RadarConf18)*, IEEE, 2018, pp. 1450–1455.
- [68] P. van Dorp and F. Groen, "Human walking estimation with radar," *IEE Proceedings-Radar, Sonar and Navigation*, vol. 150, no. 5, pp. 356–365, 2003.
- [69] D. A. Winter, *Biomechanics and motor control of human movement*. John Wiley & Sons, 2009.
- [70] V. C. Chen, F. Li, S.-S. Ho, and H. Wechsler, "Micro-doppler effect in radar: Phenomenon, model, and simulation study," *IEEE Transactions on Aerospace and electronic systems*, vol. 42, no. 1, pp. 157–158, 2006.
- [71] S. S. Ram, C. Christianson, Y. Kim, and H. Ling, "Simulation and analysis of human micro-dopplers in through-wall environments," *IEEE Transactions on Geoscience and remote sensing*, vol. 48, no. 4, pp. 2015–2023, 2010.



- [72] J. Long, E. Shelhamer, and T. Darrell, “Fully convolutional networks for semantic segmentation,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 3431–3440.
- [73] O. Ronneberger, P. Fischer, and T. Brox, “U-net: Convolutional networks for biomedical image segmentation,” in *International Conference on Medical image computing and computer-assisted intervention*, Springer, 2015, pp. 234–241.
- [74] K. He, G. Gkioxari, P. Dollár, and R. Girshick, “Mask r-cnn,” in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 2961–2969.
- [75] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, “Image-to-image translation with conditional adversarial networks,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 1125–1134.
- [76] Z. Cao, G. Hidalgo, T. Simon, S.-E. Wei, and Y. Sheikh, “Openpose: Realtime multi-person 2d pose estimation using part affinity fields,” *IEEE transactions on pattern analysis and machine intelligence*, vol. 43, no. 1, pp. 172–186, 2019.
- [77] H.-S. Fang, S. Xie, Y.-W. Tai, and C. Lu, “Rmpe: Regional multi-person pose estimation,” in *Proceedings of the IEEE International Conference on Computer Vision*, 2017, pp. 2334–2343.
- [78] J. Wang, K. Sun, T. Cheng, B. Jiang, C. Deng, Y. Zhao, D. Liu, Y. Mu, M. Tan, X. Wang, *et al.*, “Deep high-resolution representation learning for visual recognition,” *IEEE transactions on pattern analysis and machine intelligence*, 2020.
- [79] G. Moon, J. Yong Chang, and K. Mu Lee, “V2v-posenet: Voxel-to-voxel prediction network for accurate 3d hand and human pose estimation from a single depth map,” in *Proceedings of the IEEE conference on computer vision and pattern Recognition*, 2018, pp. 5079–5088.
- [80] V. Nguyen and M. A. Weitnauer, “Uwb impulse radar for vital signs sensing — a modeling framework for arbitrary periodic heart and lung motion,” in *2015 IEEE Biomedical Circuits and Systems Conference (BioCAS)*, 2015, pp. 1–4. DOI: [10.1109/BioCAS.2015.7348371](https://doi.org/10.1109/BioCAS.2015.7348371).
- [81] N. A. W. Center, “Electronic warfare and radar systems engineering handbook,” *Electronic Warfare Division, Pont Mugu, CA*, 1997.
- [82] M. I. Skolnik, “Radar handbook second edition,” *McGrawHill*, 1990.

- [83] M. Ibrani, L. Ahma, and E. Hamiti, "The age-dependence of microwave dielectric parameters of biological tissues," *Microwave Materials Characterization*, vol. 10, p. 51 400, 2012.
- [84] C. Gabriel, "Compilation of the dielectric properties of body tissues at rf and microwave frequencies.," KING'S COLL LONDON (UNITED KINGDOM) DEPT OF PHYSICS, Tech. Rep., 1996.
- [85] G. M. Brooker, "Understanding millimetre wave fmcw radars," in *1st International Conference on Sensing Technology*, 2005, pp. 152–157.
- [86] A. A. Association. (). "Child passenger safety," [Online]. Available: <https://drivinglaws.aaa.com/tag/child-passenger-safety/>. (Last accessed: 01.20.2019).
- [87] A. A. Association. (). "Car seat guide," [Online]. Available: <https://safeseats4kids.aaa.com/car-seat-guide/>. (Last accessed: 01.20.2019).
- [88] I. Gresham, A. Jenkins, R. Egri, C. Eswarappa, N. Kinayman, N. Jain, R. Anderson, F. Kolak, R. Wohler, S. P. Bawell, *et al.*, "Ultra-wideband radar sensors for short-range vehicular applications," *IEEE Transactions on Microwave Theory and Techniques*, vol. 52, no. 9, pp. 2105–2122, 2004.
- [89] FCC. (). "Da 18-1308," [Online]. Available: <https://docs.fcc.gov/public/attachments/DA-18-1308A1.pdf>. (Last accessed: 10.20.2020).
- [90] L. M. Frazier, "Radar surveillance through solid materials," in *Command, Control, Communications, and Intelligence Systems for Law Enforcement*, International Society for Optics and Photonics, vol. 2938, 1997, pp. 139–146.
- [91] M. J. Ackerman, "The visible human project," *Proceedings of the IEEE*, vol. 86, no. 3, pp. 504–511, 1998.
- [92] C. Wolff. (). "Radar basics," [Online]. Available: <http://www.radartutorial.eu/02.basics/UWB%5C%20Radar.en.html>. (Last accessed: 03.20.2019).
- [93] J. Sachs, *Handbook of ultra-wideband short-range sensing: theory, sensors, applications*. John Wiley & Sons, 2013, ISBN: 3527651837.
- [94] W. L. Stutzman and G. A. Thiele, *Antenna theory and design*. John Wiley & Sons, 2012.
- [95] S. K. Das, *Antenna and wave propagation*. Tata McGraw-Hill Education, 2016.

- [96] M. A. Richards, *Fundamentals of radar signal processing*. Tata McGraw-Hill Education, 2014.
- [97] C. A. Balanis, *Advanced engineering electromagnetics*. John Wiley & Sons, 1999.
- [98] E. O. Hammerstad, "Equations for microstrip circuit design," in *1975 5th European Microwave Conference*, IEEE, 1975, pp. 268–272.
- [99] G. Kumar and K. P. Ray, *Broadband microstrip antennas*. Artech house, 2003.
- [100] I. J. Bahl and P. Bhartia, *Microstrip antennas*. Artech house, 1980.
- [101] Q. Deng, J. Le, S. Barbat, R. Tian, and Y. Chen, "Efficient living subject localization and weak vital-sign signal enhancement using impulse radio based uwb radar," in *2019 IEEE Intelligent Vehicles Symposium (IV)*, IEEE, 2019, pp. 777–782.
- [102] T. G. Savelyev, X. Zhuge, A. G. Yarovoy, L. P. Ligthart, and B. Levitas, "Comparison of uwb sar and mimo-based short-range imaging radars," in *2009 European Radar Conference (EuRAD)*, 2009, pp. 109–112.
- [103] H. L. Van Trees, *Optimum array processing: Part IV of detection, estimation, and modulation theory*. John Wiley & Sons, 2004.
- [104] MATLAB, *Music super-resolution doa estimation*, <https://www.mathworks.com/help/phased/ug/music-super-resolution-doa-estimation.html>, (Last accessed Sep, 2020).
- [105] MATLAB, *Direction of arrival estimation with beamscan, mvdr, and music*, <https://www.mathworks.com/help/phased/examples/direction-of-arrival-estimation-with-beamscan-mvdr-and-music.html>, (Last accessed Nov, 2020).
- [106] Y. Huang and K. Boyle, *Antennas: from theory to practice*. John Wiley & Sons, 2008.
- [107] J. Jin, *Electromagnetic analysis and design in magnetic resonance imaging*. CRC press, 1998, vol. 1.
- [108] FCC. (). "17-94," [Online]. Available: <https://docs.fcc.gov/public/attachments/FCC-17-94A1.pdf>. (Last accessed: 01.20.2019).
- [109] F. C. Commission *et al.*, *Fcc rules section 15.517*, 2003.
- [110] R. E. Fields, "Evaluating compliance with fcc guidelines for human exposure to radiofrequency electromagnetic fields," *OET Bull* 65, 2001.

- [111] I. Guideline, “Guidelines for limiting exposure to time-varying electric, magnetic, and electromagnetic fields (up to 300 ghz),” *Health phys*, vol. 74, no. 4, pp. 494–522, 1998.
- [112] H. J. Krzywicki and K. S. Chinn, “Human body density and fat of an adult male population as measured by water displacement,” *The American journal of clinical nutrition*, vol. 20, no. 4, pp. 305–310, 1967.
- [113] MATLAB, *Overhead mount vehicle occupancy detection*, [https://dev.ti.com/tirex/explore/node?node=ALRSW97rqR7Y.z46NeSpEQ\\_\\_\\_AocYeEd\\_\\_\\_LATEST](https://dev.ti.com/tirex/explore/node?node=ALRSW97rqR7Y.z46NeSpEQ___AocYeEd___LATEST), (Last accessed Nov, 2020).
- [114] S. D. D. Cruz, O. Wasenmuller, H.-P. Beise, T. Stifter, and D. Stricker, “Sviro: Synthetic vehicle interior rear seat occupancy dataset and benchmark,” in *The IEEE Winter Conference on Applications of Computer Vision*, 2020, pp. 973–982.
- [115] M. Berman, A. Rannen Triki, and M. B. Blaschko, “The lovász-softmax loss: A tractable surrogate for the optimization of the intersection-over-union measure in neural networks,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 4413–4421.
- [116] K. He, X. Zhang, S. Ren, and J. Sun, “Delving deep into rectifiers: Surpassing human-level performance on imagenet classification,” in *Proceedings of the IEEE international conference on computer vision*, 2015, pp. 1026–1034.
- [117] P. Malmivuo, J. Malmivuo, and R. Plonsey, *Bioelectromagnetism: principles and applications of bioelectric and biomagnetic fields*. Oxford University Press, USA, 1995.
- [118] B. R. Mahafza, *Radar systems analysis and design using MATLAB*. CRC press, 2002.
- [119] T. Instrument, “Vital signs 68xx developer’s guide,” *Industrial Toolbox*, 2019.

## VITA

Mr. Qiwen Deng graduated from Wuhan University, China in 2013 with the B.S. degree in Mechanical Engineering and B.A. degree in Finance. He attended Purdue University in Fall 2013 for his graduate study in Electrical and Computer Engineering and received his MSECE in Spring 2015. He worked briefly at Jing-Jin Electric Technologies Co. in China as vehicle system software engineer before returning to Purdue University in spring 2016 to pursue his doctoral degree in Electrical and Computer Engineering. From February 2018 to December 2020, he worked on collaborative research projects at Ford Research and Innovation Center funded by Ford University Research Program.