Human Activity Recognition using Reconfigurable Intelligent Surfaces in controlled environments

Mariana Oliveira
Master in Engineering Physics
Department of Physics and Astronomy
Faculty of Sciences of University of Porto and Faculty of Engineering of University of Porto
2024

Human Activity Recognition using Reconfigurable Intelligent Surfaces in controlled environments

Mariana Oliveira

Project Report carried out as part of the Master in Engineering Physics

Department of Physics and Astronomy

2024

Supervisor
Dr. Luís Manuel Pessoa, Assistant Professor, FEUP

Co-supervisor
Dr. Nuno Miguel Paulino, Researcher, INESC TEC

INESCTEC

“ Success is not the key to happiness. Happiness is the key to success. If you love what you are doing, you will be successful. ”

Albert Schweitzer

Dedicated to my family, whose support made the tough times easier and the good times even better.

Acknowledgements

Em primeiro lugar, à família: à mãe, que tudo fez para que este percurso parecesse mais fácil, esteve presente nos momentos mais difíceis, apoiou-me nos momentos de maior vulnerabilidade e celebrou comigo as conquistas mais felizes; ao irmão, que tudo fez para me ajudar a manter focada e evitar que adiasse as tarefas, mesmo quando a tentação era grande; à tia, que tudo fez para me inspirar a cada dia com o seu exemplo de esforço e resiliência; aos avós, que tudo fizeram para gabar a neta, mas em especial à avó Laura, que tudo fez para que eu me sentisse feliz, amada e com a barriga cheia durante esta jornada. Este capítulo encerra-se para mim, mas também para vocês, que viveram esta aventura ao meu lado e me apoiaram incondicionalmente até ao fim.

Depois, aos amigos: àqueles que me apoiaram nas épocas mais complicadas do curso, mesmo sem saberem o porquê de ser complicado. E àqueles que conheci durante o curso, com quem partilhei e enfrentei essas mesmas épocas complicadas, sabendo exatamente o porquê de ser complicado. Houve alturas em que parecia impossível dar mais de mim, mas foram os amigos que estiveram lá para dar a força que eu precisava.

Ao Vasco, que esteve lá desde o dia 1 para apoiar, admirar, acalmar e amar. Que esteve lá para tornar este caminho mais divertido, bonito e fácil.

Por fim, seguem-se os agradecimentos especiais às pessoas do INESC TEC: ao Professor Luís Pessoa, um excelente orientador que esteve sempre presente para ajudar, para guiar e para discutir ideias, onde a partilha de conhecimento era constante. Ao Nuno Paulino, que me ajudou sempre nos momentos em que me sentia sem saída e com quem discutia ideias para que cada obstáculo se tornasse um desafio, e não um problema. Ao Francisco Ribeiro, que foi fundamental para o desenvolvimento deste trabalho, tanto a nível experimental como crítico. E, por fim, à Sofia Inácio, que muito me ajudou a nível de simulação e discussão de ideias. Obrigada por toda a ajuda que cada um de vocês deu; no INESC TEC, sempre senti muito apoio e sempre me senti muito feliz a desenvolver este trabalho, graças a vocês.

Resumo

O Reconhecimento de Atividades Humanas concentra-se na identificação e classificação de ações humanas através de diferentes tecnologias de *sensing*, incluindo a deteção por Radio frequência (RF). Esta dissertação investiga como as Superfícies Inteligentes Reconfiguráveis (RISs) podem ser integradas no reconhecimento de atividades humanas para melhorar a deteção por RF, em particular no Reconhecimento de Postura Humana e no Reconhecimento de Gestos das Mãos. As RISs, que consistem em elementos reflectivos programáveis, oferecem uma nova abordagem para controlar a propagação das ondas eletromagnéticas, ajudando a superar desafios como a atenuação por múltiplos trajetos, que frequentemente dificultam os sistemas de detecção baseados em RF.

Os métodos tradicionais de deteção por RF muitas vezes enfrentam dificuldades em ambientes complexos, onde obstáculos e interferências podem distorcer os sinais. Ao integrar a tecnologia RIS, este trabalho visa melhorar a precisão e a fiabilidade dos sistemas de HAR. As RISs permitem um melhor controlo das ondas de rádio,de modo a reduzir a interferência dos sinais e aprimorar o reconhecimento de atividades humanas.

O estudo analisa como a manipulação dos sinais de rádio em ambientes controlados pode criar ambientes de rádio inteligentes e utiliza parâmetros S_{21} para classificar atividades humanas. A dissertação contribui para os sistemas de reconhecimento de atividades humanas baseados em RIS ao avaliar os métodos de aquisição de dados, desenvolver modelos de aprendizagem supervisionada para a classificação de atividades e investigar a otimização dos elementos reflectivos das RISs. Estes avanços oferecem soluções práticas para melhorar a deteção de atividades humanas em cenários desafiantes.

No geral, esta dissertação avança o reconhecimento de atividades humanas ao integrar a tecnologia RIS, demonstrando como estas inovações podem melhorar significativamente a deteção por RF e otimizar o desempenho na identificação de atividades humanas.

Palavras-chave: HAR, RF Sensing, RIS, Wifi-6E, Redes Neuronais Convolucionais

Abstract

Human Activity Recognition (HAR) focuses on identifying and classifying human actions using various sensing technologies, such as Radio Frequency (RF) sensing. This dissertation explores how Reconfigurable Intelligent Surfaces (RISs) can be used to enhance HAR, specifically in Human Posture Recognition (HPR) and Hand Gesture Recognition (HGR). RISs, which consist of programmable reflective elements, offer a new approach to control how electromagnetic waves propagate. This can help address common issues like multipath fading that often make RF-based HAR systems challenging.

Traditional RF sensing methods often struggle in complex environments, where obstacles and interference distort signals. By integrating RIS technology, this work aims to improve the accuracy and reliability of HAR systems. RISs enable better control of radio waves, reducing signal interference and enhancing activity recognition.

The study examines how manipulating radio signals in controlled settings can create smart radio environments and uses S_{21} parameters to classify human-related tasks. It contributes to RIS-based HAR systems by analyzing data acquisition setups, developing supervised learning models for activity classification, and studying how to optimize RIS reflective elements. These efforts provide practical solutions for improving human activity detection in challenging scenarios.

Overall, this dissertation advances HAR by integrating RIS technology, showing how these innovations can significantly enhance RF sensing and improve the performance of human activity recognition.

Keywords: HAR, RF Sensing, RIS, Wifi-6E, Convolutional Neural Network (CNN)

Table of Contents

Acknowledgements iii

Resumo iv
Abstract v

Table of Contents vi
List of Figures viii
List of Abbreviations x

1. Introduction 1
   1.1. Context 1
   1.2. Motivation 2
   1.3. Objectives 3
   1.4. Contributions 3
   1.5. Document Structure 4
2. Related Work 5
   2.1. Background 5
      2.1.1. Human Activity Recognition Techniques 5
      2.1.2. Reconfigurable Intelligent Surfaces 8
   2.2. State-of-the-Art 9
      2.2.1. RIS-based HAR Approaches 10
      2.2.2. Summary 15
3. Proposed Approach 16
4. Human Posture Recognition with Metal Reflectors 18
   4.1. Design of the Setup 19
      4.1.1. Antenna Configuration 19
      4.1.2. Reflective surface (Metal Plate) 24
   4.2. Experimental Setup and Methodology 26
   4.3. Results 28
      4.3.1. Dataset Description and Analysis 28
      4.3.2. Posture Classification 30
   4.4. Summary 31
5. Hand Gesture Recognition with a RIS 33
   5.1. System Configuration 34
      5.1.1. Design of Set up 34
      5.1.2. FCAO Algorithm for RIS Optimization 35

5.2. RIS Validation 45
5.2.1. Initial Tests 46
5.3. Experimental Setup and Methodology 48
5.4. Results 50
5.4.1. Feasibility Study with Real Hand 50
5.4.2. Gesture Classification with Wooden Hand 52
5.5. Summary 59
6. Conclusions and Future Work 60

Bibliography 61
A. Transmitter Antenna for RIS-based set up 71
B. Preliminary Experiments with Metal Plate 76
B.1. Use of LNAs and Experimental Setup Adaptations 76
B.2. Transition to Anechoic Chamber 78
C. FCAO Algorithm 79
C.1. Initial Value of t_k using the *Pattern Search* Method 79
C.2. Gradient derivation for t_k minimization step 80
C.3. Code 81
C.4. Results from authors of FCAO Algorithm 86
D. Published Paper: SpecRF-Posture: Exploring Specular Reflections for Human Posture Recognition 87

List of Figures

2.1. Types of Human Activity Recognition. 5
2.2. VNA-based channel sounder system....7
2.3. RIS Capability: Reflector vs. Phase Shifts. 8
2.4. Signal Scattering in Conventional vs. Reconfigurable Environments. 9
2.5. Components of the RIS-based posture recognition system of [41] 11
2.6. Schematic configuration of the intelligent metasurface system....12
2.7. System configuration of the intelligent EM metasurface camera [44]. 13
2.8. Schematic map of the system [45]. 14

3.1. System architecture diagram of the general approach for HAR. 16

4.1. HPR System Architecture Diagram....18
4.2. E-plane Radiation Pattern of LB-20180 Antenna. 20
4.3. System Geometric Configuration. 21
4.4. Illustration of aperture losses. 22
4.5. Efficiency vs. H/D for $\theta = 45^{\circ}$. 23
4.6. Tx Antenna Emission and Reflection Distortion....25
4.7. Experimental Setup and Illumination Patterns. 26
4.8. HPR Experimental Setup in Anechoic Chamber....27
4.9. S_{21} Phase, Magnitude, and Posture Identification Analysis. 29
4.10CNN Architecture for Posture Classification. 30
4.11. Posture Recognition Accuracy: Magnitude vs. Magnitude & Phase....31

5.1. HGR System Architecture Diagram. 33
5.2. Efficiencies as functions of the design parameters of the setup. 34
5.3. Top-view of Setup Schematic. 36
5.4. Pattern search algorithm output. 42
5.5. Convergence of Pattern Search Algorithm 42
5.6. Convergence of the alternating minimization....43
5.7. Convergence of FCAO Algorithm 43
5.8. Mutual Coherence Matrix: Random vs. Optimized. 44
5.9. RIS details. 45
5.10RIS Configurations for Signal Steering. 46
5.11. RIS Steering Verification Setup. 47
5.12. Signal Magnitude vs. RIS Steering Angles at 5.91 GHz. 47
5.13. HGR Experimental Setup: Real vs. Wooden Hand. 49
5.14. S_{21} Magnitude for RIS Steering by 20 Degrees. 51
5.15. Frequency vs RIS Configuration: Gesture Identification. 52
5.16. Frequency vs RIS Random Configuration: Average Magnitude and Phase. 53
5.17. Frequency vs RIS Optimized Configuration: Average Magnitude and Phase. 54
5.18. Mean Magnitude Differences for Random vs. Optimized RIS Configurations. 55
5.19. Data augmentation example: white noise added to one image corresponding to measurements using random RIS configurations (top row), and another employing FCAO derived configurations (bottom row). 55
5.20. Convergence of Simple Model for Gesture Classification. 56

5.21. Simple model for Gesture Recognition Accuracy: Random vs. Optimized Config. ...57
5.22. CNN Architecture for Gesture Classification. ...58
5.23. CNN model for Gesture Recognition Accuracy: Random vs. Optimized Config. ...58

A.1. E-plane Radiation Patterns of LB-137 Antenna (10dB and 15dB, 6.5GHz) ...71
A.2. Normalized Radiation Patterns of LB-137 Antenna (15dB and 10dB). ...72
A.3. Efficiency vs. θ_0 ...73
A.4. Efficiency vs. H/D for $\theta = 35^\circ$ ...73
A.5. Efficiency vs. H/D for $\theta = 35^\circ$ and selected H's ...74
A.6. Intensity on Reflectarray (dB) ...74
A.7. Setup with Horn Antennas (10dB or 15dB Gain). ...75

B.1. First Experimental Setup in the laboratory. ...76
B.2. S_{21} parameters in the first experiment indoor. ...76
B.3. Second Experiment in the laboratory. ...77
B.4. S_{21} magnitude in the second experiment indoor. ...77
B.5. First Experiment in outdoor environment. ...78
B.6. S_{21} magnitude in the outdoor environment. ...78

C.1. [41] Results of T matrix and mutual coherence ...86
C.2. [41] Results of convergence ...86

List of Abbreviations

RF Radio Frequency

VNA Vector Network Analyser

HAR Human Activity Recognition

CNN Convolutional Neural Network

HPR Human Posture Recognition

HGR Hand Gesture Recognition

SoI Space-of-Interest

CS Compressive Sensing

LoS Line-of-Sight

RIS Reconfigurable Intelligent Surface

RSSI Received Signal Strength Indicator

CSI Channel State Information

RFID Radio Frequency Identification

LNA Low Noise Amplifier

FCAO Frame Configuration Alternating Optimization

ANN Artificial Neural Networks

1. Introduction

This chapter explains the motivation behind this dissertation, and the subsequent objectives. Specifically, Section 1.1 explains what Human Activity Recognition is and how Radio Frequency sensing with Reconfigurable Intelligent Surfaces can enhance it. The following Section 1.2 highlights why using Reconfigurable Intelligent Surfaces with Radio Frequency sensing is important and how it addresses current challenges. Section 1.3 outlines the established objectives. Section 1.4 summarizes the main achievements of the work. Finally, Section 1.5 gives a brief overview of the structure and contents of the remaining chapters.

1.1. Context

Human Activity Recognition is a growing field dedicated to identifying and classifying all forms of human activity, encompassing both static and dynamic motion [1]. The growth of HAR is driven by advances in sensor technology and machine learning algorithms, enabling more accurate recognition of activities. This technology has the potential to significantly impact various domains. In healthcare, HAR can assist in monitoring patient conditions and detecting early signs of medical issues [2]. In smart environments, HAR can enhance home assistance services to improve quality of life and autonomy [3]. In sports, it can contribute to improving personal health and well-being [4]. In security, HAR can be used for surveillance to identify unusual events [5]. Additionally, HAR improves human-computer interaction (HCI) by enabling more intuitive and responsive interfaces [6].

HAR can be accomplished using various types of data, such as video, motion sensors, and radio signals. As technology advances, traditional methods of human activity recognition are increasingly complemented by other approaches such as Radio Frequency (RF) Sensing.

RF Sensing is a privacy-preserving technique for HAR that addresses concerns about monitoring and privacy. By using radio frequency signals, RF Sensing detects human activities through the effects of body movements on the propagation of these signals. The attributes of the radio waves, such as amplitude or phase changes, as they travel from a transmitter to a receiver, reveal patterns of human activity. This method avoids direct capture of personal information, thereby preserving privacy while effectively recognizing and analyzing human activities [7].

While RF Sensing has many advantages, it faces some significant challenges because wireless environments are often complex and unpredictable. One important issue is multipath fading, which happens when signals reflect off different surfaces and reach the receiver through various paths [8]. This reflection can mix up the signals, making it harder to accurately detect and identify activities.

To overcome these challenges, Reconfigurable Intelligent Surfaces (RISs) are a promising solution with the ability to program and control electromagnetic wave propagation, enhancing RF Sensing and becoming a candidate technology for 6G wireless networks [9].

RISs are two-dimensional arrays composed of programmable reflective elements, typically made from advanced metamaterials. These surfaces can be tuned to alter the properties of electromagnetic waves, including their phase and amplitude, allowing for the creation of optimal signal paths even in complex or obstructed environments [10]. In addition to HAR, RF Sensing applications using RISs have been explored in other areas. For instance, they have been utilized in healthcare for heartbeat detection in complex environments [11], in tracking systems for real-time object positioning [12], and in localization systems to enhance communication and positioning accuracy [13–15]. By manipulating the propagation of radio waves, RISs can mitigate issues like multipath fading and signal interference, thus improving the accuracy of RF Sensing in HAR applications.

1.2. Motivation

Integrating RISs into Human Activity Recognition is driven by the need to tackle the limitations of current technology. As highlighted, traditional RF Sensing methods often struggle with challenges like multipath fading, which can interfere with accurate activity detection.

RIS technology offers a significant advantage by allowing us to actively control and shape the radio wave environment. This means we can improve signal quality and enhance activity recognition, even in difficult conditions where conventional methods might fail.

As we look ahead to the future of wireless technology, including the development of 6G networks, leveraging RIS for HAR not only solves existing problems but also positions us for new possibilities. This approach is set to play a crucial role in advancing how we recognize and analyze human activities in various scenarios.

1.3. Objectives

This work focused on both hardware devices and software techniques, including inference models, for human activity recognition using RF data, obtained using novel 6G radio devices, namely RISs. This approach aims to make sensing solutions more widely available and affordable through the use of low-cost hardware. Thus, the main objective was defined as follows:

- To advance the state of the art in human activity recognition by implementing novel smart radio environments for RF sensing.

The following intermediate goals were defined as milestones towards this main objective:

- Identify and implement methods to enhance the information retrieved from the RF channel.
- Model the Space-of-Interest (SoI), encompassing the RF channel and interactions with target objects.
- Configure experimental setups to build RF sensing datasets with ground truth.
- Train and evaluate the performance of machine learning models for human activity recognition.

1.4. Contributions

This work has made several important advances in the field of HAR by leveraging RF sensing and RISs. The key contributions are as follows:

- Methodologies for setup of RF sensing environments using RISs and data gathering.
- Implemented supervised learning models to classify human activity with RF data.
- Published open-source data sets for RF sensing, including: 1) dataset based on specular reflections for full body sensing*, and 2) dataset based on controllable RIS for hand gesture recognition†.
- Published a scientific paper [16] in 2024 IEEE International Mediterranean Conference on Communications and Networking (MeditCom) (see Chapter 4).

*M. Oliveira, F. Ribeiro, N. Paulino, O. Yurduseven and L. Pessoa, “SpecRF-Posture Dataset”. Zenodo, Apr. 03, 2024. doi: 10.5281/zenodo.10911412.

†M. S. F. de B. Oliveira, F. M. Ribeiro, N. Paulino and L. M. Pessoa, “RIS Based Hand Gesture Recognition Dataset”. Zenodo, Sep. 12, 2024. doi: 10.5281/zenodo.13754235.

1.5. Document Structure

This section presents the structure of the document, detailing the content and purpose of each chapter. Chapter 2 reviews related work, covering essential concepts needed to understand this work and providing a brief overview of the state of the art in comparable studies. Chapter 3 outlines the proposed approach in this dissertation, building on the state of the art. Chapter 4 focuses on human posture recognition using specular reflections with a metal plate, while Chapter 5 explores the replacement of the metal plate with a small RIS for human gesture recognition.

The appendices provide additional insights: Appendix A analyzes the positioning and material considerations for a setup using a larger RIS $(1\,\mathrm{m} \times 1\,\mathrm{m})$. Appendix B details the preliminary experiments with a metal plate, essential for the setup in Chapter 4. Appendix C includes details on a channel optimization algorithm, Frame Configuration Alternating Optimization, used in Chapter 5, and the Python code. Appendix D presents the published paper, relative to the work detailed in Chapter 4.

2. Related Work

In this chapter, we divide the related work into two main sections. The first section introduces essential concepts to help understand the methods and findings of this thesis. The second section reviews the latest advancements and approaches in the field RF sensing for human activity recognition using RISs.

2.1. Background

2.1.1 Human Activity Recognition Techniques

Figure 2.1 illustrates the three main categories of HAR: static activities, dynamic activities, and activities involving postural transitions [17]. Static activities include tasks like lying, sitting, and standing, which are generally easier to identify but can be confused with similar postures. Dynamic activities involve continuous movement, such as walking or running, which require distinguishing between different motion patterns. Activities with postural transitions, such as sitting to standing or walking to jogging, involve shifts between static and dynamic states and can be particularly complex due to overlapping motion patterns.

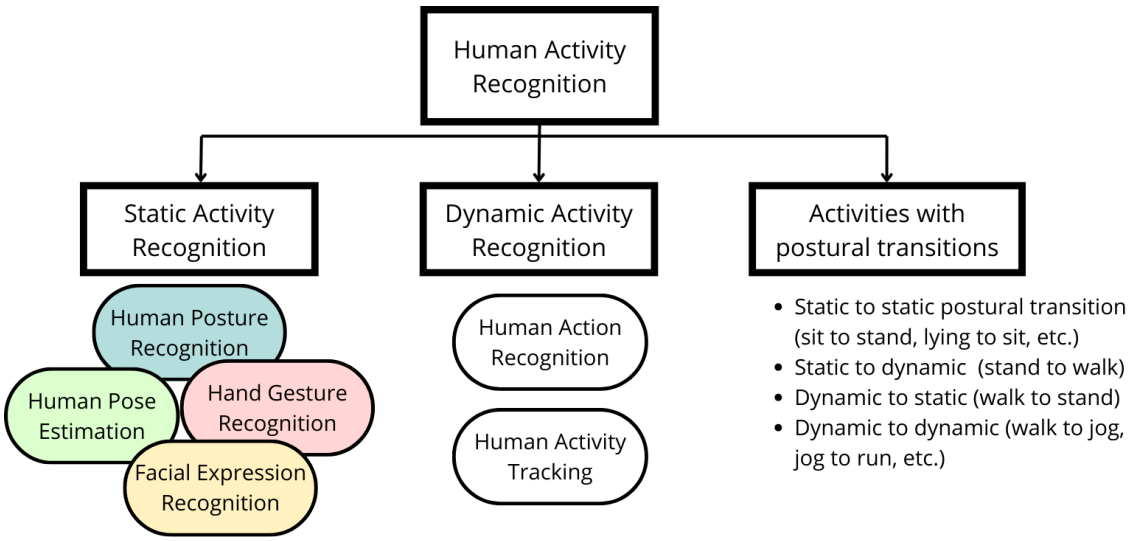


Figure 2.1: Types of Human Activity Recognition.

This work will focus on static methods, specifically Human Posture Recognition (HPR) and Hand Gesture Recognition (HGR). Human Posture Recognition involves classifying various static postures of the human body, such as lying, sitting, and standing. This technique is crucial in applications ranging from ergonomic monitoring to assistive technologies [18].

Hand Gesture Recognition focuses on detecting hand movements and static hand positions. This technique is relevant for applications like human-computer interaction, virtual reality, and sign language recognition [19].

Human activity recognition is a complex process that involves four key phases: selecting and deploying suitable sensors, collecting data from these sensors, pre-processing the data and extracting relevant features, and applying machine learning and deep learning algorithms to recognize activities [20]. The sensors used for sensing can include cameras, wearable devices, and sensors embedded in everyday objects or the environment. The techniques for HAR can be categorized based on the type of sensors used:

Vision-based Vision-based HAR utilizes optical sensors, such as cameras, to analyze images or videos. These optical sensors capture data in the form of 2D or 3D images and videos, which are then processed to identify and classify activities. Research in this area often distinguishes between RGB data, which captures color information, and RGB-D data, which also includes depth information to provide a more detailed understanding of the scene [21].

Sensor-based Sensor-based HAR employs various types of sensors to detect and analyze activities. Wearable sensors, such as accelerometers, magnetometers, and gyroscopes, are worn on the body or integrated into devices like smartphones and smartwatches to capture direct movement data [22]. Object sensors, attached to specific objects, help infer activities by monitoring movements related to those objects. Environmental sensors, placed in the surroundings, measure parameters like temperature and humidity, providing contextual information that helps in activity recognition [23].

Radio Frequency-based RF-based HAR utilizes radio signals and their reflections to detect human activities. Unlike vision-based systems, which depend on optical data, RF sensing is effective in low-light conditions and through obstacles. Additionally, wearable sensors, while useful, can be uncomfortable for prolonged usage, making RF techniques a noninvasive alternative. RF sensors capture the reflections of radio waves from objects and people, allowing for activity recognition even in non-line-of-sight (NLOS) scenarios [24]. This method is particularly useful for applications requiring through-wall detection [25, 26] and monitoring in complex environments, such as multi-person scenarios [27].

In recent years, some of the RF-based technologies that have become popular for human-related aplications are:

- RFID-based: Radio Frequency Identification (RFID) technology uses readers and tags where radio waves from the reader are backscattered by the tags. This allows the reader to capture information about the tags through the backscattered signals, typically using Received Signal Strength Indicator (RSSI) [28].
- Radar-based: Radar systems operate by emitting radio waves that reflect off objects, including humans [29]. By analyzing these reflections, radar systems can identify various attributes and recognize activities based on the differences between transmitted and received signals [30].
- WiFi-based: This method leverages disturbances in WiFi signals caused by human movement, analyzing Channel State Information (CSI) and variations across subcarriers. Human presence disrupts these subcarriers, altering signal attributes captured by the CSI [31].

These RF-based techniques demonstrate significant versatility in capturing human-related features but often require complex hardware configurations, limiting scalability and deployment. An alternative approach uses S_{21} parameters with a Vector Network Analyser (VNA) as a flexible solution for HAR. A VNA-based approach measures the complex frequency response of a channel by the S_{21} scattering parameters [32], which reveals how electromagnetic waves interact with the human body. This process involves wideband channel measurements conducted via stepped frequency sweeps across discrete points within a defined bandwidth [33]. Figure 2.2 illustrates a block diagram of the VNA-based channel sounder system.

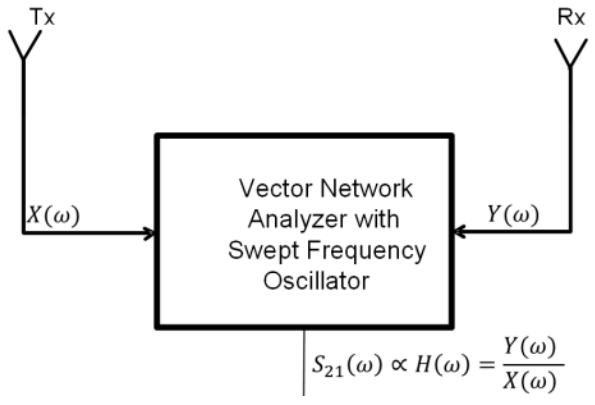
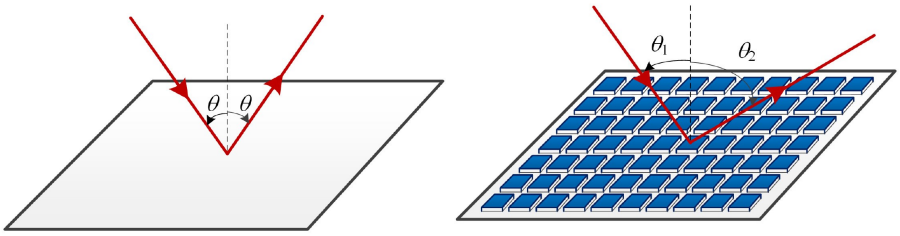


Figure 2.2: VNA-based channel sounder system. Adapted from [33].

2.1.2 Reconfigurable Intelligent Surfaces

RISs consists of several key components: an array of programmable elements that interact with incident signals, a copper layer to prevent signal leakage and a control circuit that adjusts the phase shifts of the elements [34]. These programmable elements manipulate electromagnetic waves by re-emitting them based on their material properties, such as permittivity, ϵ, and permeability, μ. RIS elements include PIN diodes that switch states to alter the signal response, precisely adjusting phase shifts and amplitudes. To illustrate the structure and function of RIS, it is useful to compare it with a regular reflector, such as a metal plate, where the angle of incidence equals the angle of reflection, as shown in Fig.2.3a. Unlike regular reflectors, RIS can adjust the surface impedance to produce specific phase shifts between incident and scattered waves, directing the reflected beam at any desired angle, as depicted in Fig. 2.3b.



(a) Regular Refletor

(b) Reconfigurable Intelligent Surface

Figure 2.3: RIS capability: Regular reflector vs. Programmable phase shifts. Adapted from [35].

This ability allows a RIS to act as a beamformer by programming the ON/OFF states of the diodes to achieve various beamforming patterns. This is achieved by controlling the reflection coefficients of each unit cell i.e., the single elements of the RIS [36]:

$$r_i = \Gamma e^{j\theta} \tag{2.1}$$

where $\Gamma \in [0,1]$ is the reflection amplitude and $\theta \in [0,2\pi]$ is the phase shift between incident and reflected signals.

Enhancing RF Sensing with Reconfigurable Intelligent Surfaces

Unlocking the full potential of RF sensing presents a challenge, particularly in the face of persistent obstacles such as multipath interference and radio environmental limitations that limit the number of transmission channels to carry information about targets to be sensed. To address these challenges, researchers have explored various solutions.

One approach involves integrating deep learning techniques to extract features directly from RF data without explicit modeling of complex physical processes, enhancing data processing capabilities [37]. Another promising strategy is incorporating RISs into RF sensing systems. RIS manipulates the propagation of electromagnetic waves, enabling dynamic adjustments of phase shifts and amplitude ratios along propagation paths. By reconfiguring these parameters, RIS creates more favorable channel conditions for RF sensing, optimizing signal transmission and reception [38].

Unlike conventional surfaces, which scatter radio waves based on their material properties and boundary conditions, RIS can dynamically alter the propagation of these waves. RIS achieves this by adjusting the phase of the reflected signals through programmable unit cells, allowing for precise beamforming and focusing. This capability enables RIS to direct reflected signals towards specific locations, enhancing the accuracy of RF sensing systems for human activity detection, as shown in the comparative example in Fig. 2.4.

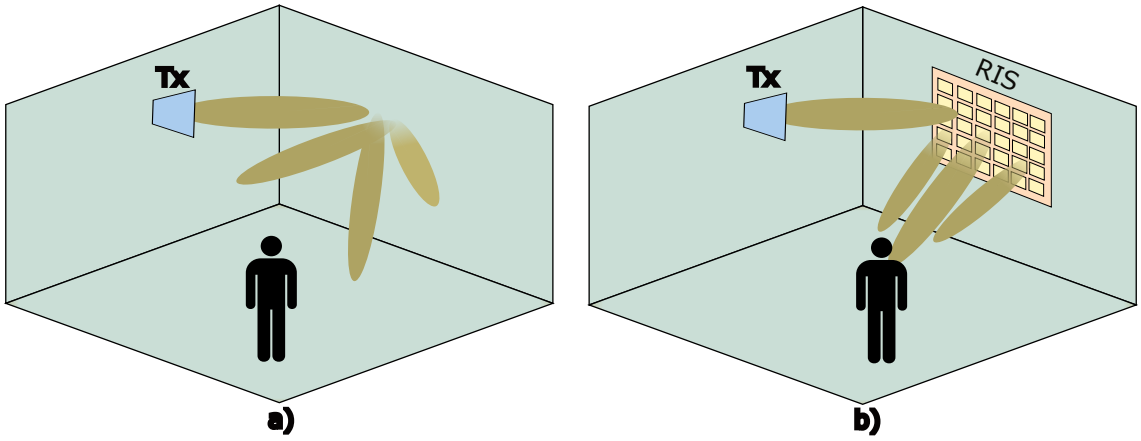


Figure 2.4: a) Conventional environment with diffuse signal scattering; b) Reconfigurable environment with RIS enabling targeted signal beamforming.

2.2. State-of-the-Art

Conventional HAR techniques that utilize RF signals have gained considerable attention due to their ability to detect activities without requiring direct line of sight. In [39], the authors explore a system that utilizes Wi-Fi and radar signals to read lip movements under face masks, enabling communication in challenging environments. The system employs machine learning and deep learning algorithms to analyze variations in RF signals for vowel identification. However, it is limited by its dependence on environmental conditions, which can adversely affect the quality of the received signal.

Similarly, in [40], the authors investigate a system that employs passive RFID signals for posture recognition in indoor environments, specifically focusing on monitoring the

activities of elderly individuals.This system analyzes variation patterns in RSSI values and employs machine learning techniques such as multinomial logistic regression and L1 regularization for classification. Despite its effectiveness, the system is sensitive to noise and interference, particularly during posture transitions, which can compromise detection accuracy.

These conventional RF-based systems often face challenges such as multipath interference, sensitivity to environmental changes, and reduced accuracy in complex scenarios. The reliance on environmental factors and the difficulties posed by signal distortion and interference underscore the need for more robust solutions for HAR.

RISs offer a promising advancement by actively shaping the radio environment. RIS-enhanced HAR systems can significantly improve signal quality, reduce interference, and focus RF energy more precisely on the areas of interest. This capability opens new possibilities for more accurate and robust human activity detection, even in challenging environments. Therefore, this section will explore the latest advancements in RIS-based HAR, examining how this technology is revolutionizing RF sensing and setting a new benchmark for future HAR systems.

2.2.1 RIS-based HAR Approaches

This section focuses on recent advancements in RF Sensing systems that incorporate RIS for applications in human activity recognition. These approaches have greatly influenced and shaped this research. Five specific works developed in the context of HAR based on the usage of RISs for manipulation of the radio frequency signals are reviewed below, highlighting their impact and contributions to the field.

Reconfigurable Intelligent Surface Based RF Sensing: Design, Optimization, and Implementation [41] This work explores the integration of RIS to enhance the accuracy and flexibility of posture recognition through the manipulation of RF signals.

The authors propose a novel RF sensing system that actively customizes the propagation environment by optimizing the configurations of the RIS. This approach addresses the limitations of conventional RF sensing techniques, which often struggle with multi-path fading and restricted transmission channels. By leveraging the capabilities of the RIS, the system is able to create multiple independent paths for signal transmission, thereby enriching the information available for accurate posture recognition.

A key contribution of this research is the formulation of a joint optimization problem that encompasses both the configuration of the RIS and the decision function for posture recognition. The authors employ the Frame Configuration Alternating Optimization (FCAO) algorithm, which minimizes the mutual coherence of the measurement matrix. This reduction in mutual coherence enhances the system's ability to distinguish between different postures and facilitates this classification task.

The experimental setup features a horn antenna that emits signals towards the RIS, while an omnidirectional antenna captures the reflections from the human body, utilizing S_{21} parameters as raw data to evaluate the system's performance (Fig. 2.5). The effectiveness of their approach is demonstrated through practical experiments, resulting in a remarkable improvement in recognition accuracy for 4 postures, achieving up to 96.7% with optimized configurations, compared to 82.1% in random configurations and 73.2% in non-configurable environments.

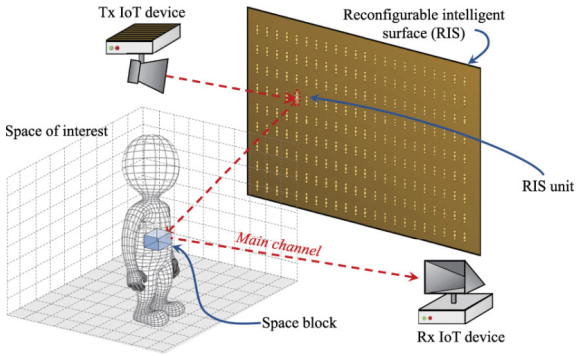


Figure 2.5: Components of the RIS-based posture recognition system of [41]

Intelligent metasurface imager and recognizer [42] Li et al. present a significant advancement in radio-frequency (RF) sensing systems with their intelligent metasurface imager and recognizer, which operates at approximately 2.4 GHz, one of the standard Wi-Fi frequency bands. This system enhances HAR by adaptively manipulating ambient RF signals to capture detailed images of human bodies and recognize gestures and vital signs in real-time.

A notable innovation in this study is the use of Artificial Neural Networkss (ANNs) to process RF signals, effectively transforming them into silhouette images of individuals by integrating a vision dataset. Rather than employing the FCAO algorithm to determine optimized configurations for various information retrieval tasks, the authors adapt the

configuration of the RIS in real-time, according to the processed RF data and the specific areas of interest identified by the ANNs. This capability allows the system to focus electromagnetic fields on selected body parts while minimizing interference from the surrounding environment and enables the system to distinguish between various hand signs and monitor vital signs, such as respiration, even when individuals are uncooperative or positioned behind obstacles.

The experimental setup includes a large-aperture programmable metasurface, transmitting and receiving antennas, and a VNA for data acquisition, along with a network of ANNs (Fig. 2.6). This configuration enables the system to distinguish between various hand signs and monitor vital signs, such as respiration, even when individuals are uncooperative or positioned behind obstacles.

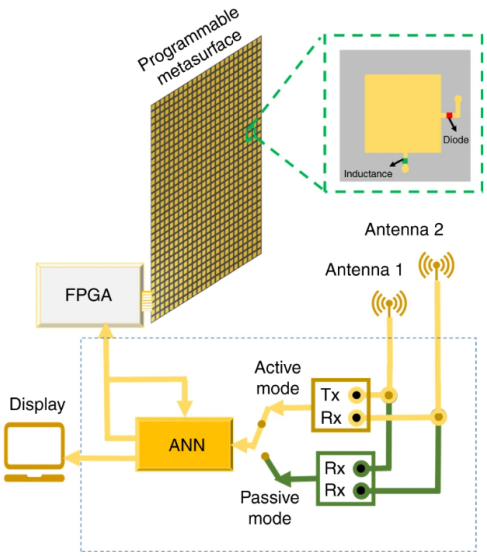


Figure 2.6: Schematic configuration of the intelligent metasurface system showing active and passive modes. [42]

Intelligent Electromagnetic Sensing with Learnable Data Acquisition and Processing [43] The authors present an approach of electromagnetic sensing utilizing a programmable metasurface for imaging and body gesture recognition tasks at 2.4 GHz. This innovative system shapes waves illuminating the scene and enhances sensing quality by employing multiple coding patterns of the metasurface, resulting in improved resolution in the image reconstruction of body postures.

A key contribution of this work is the integration of two deep ANNs: the measurement ANN (m-ANN) for adaptive data acquisition and the reconstruction ANN (r-ANN) for instant data

processing. This integrated framework allows for the simultaneous optimization of measurement strategies and data processing schemes, significantly enhancing the efficiency and accuracy of human activity recognition tasks.

The experimental setup employs a dual-antenna configuration using horn antennas, where one antenna serves as the transmitter and the other as the receiver. These antennas are connected to a VNA, which captures the S_{21} parameters as the data.

Intelligent electromagnetic metasurface camera: system design and experimental results [44] The authors introduce an innovative camera system that integrates programmable electromagnetic metasurfaces with deep learning techniques, establishing a Bayesian inference framework to enhance data acquisition and processing.

A key contribution of this work is the development of an intelligent EM metasurface camera capable of visualizing human behaviors behind a 60 cm-thick reinforced concrete wall. This capability is achieved through the use of a large-aperture programmable metasurface, which allows for adaptive data acquisition, and artificial neural networks that facilitate instant data processing. The system operates at a frequency of approximately 2.4 GHz, demonstrating robust performance in real-world settings, even in visually obstructed environments.

The experimental setup includes a low-cost commercial software-defined radio device (Ettus USRP X310), a transmitting antenna, and a three-antenna receiver, all synchronized with a personal computer (Fig. 2.7). The camera employs a compressive microwave measurement strategy, utilizing 18 control patterns per image to optimize the trade-off between imaging quality and efficiency. This innovative approach not only enhances the frame rate but also significantly improves the signal-to-noise ratio (SNR) of the acquired data, making it a promising tool for smart community applications and beyond.

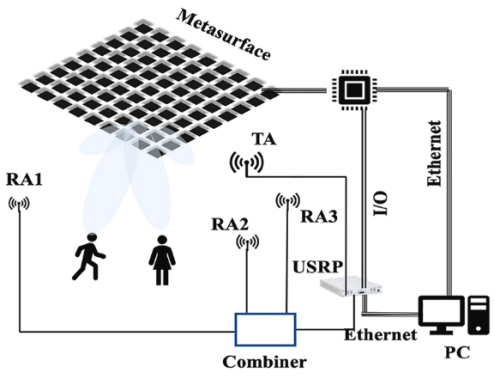


Figure 2.7: System configuration of the intelligent EM metasurface camera [44].

Machine-learning reprogrammable metasurface imager [45] This work investigates a digital imaging system based on a reprogrammable metasurface that utilizes machine-learning techniques to optimize real-time image capture. The system employs a two-step training strategy where desirable radiation patterns are first learned using machine learning algorithms, and then the corresponding coding patterns of the metasurface are designed to achieve the desired radiation configuration. This approach enables the generation of optimized measurement modes that facilitate high-quality object recognition and imaging while significantly reducing the number of required measurements. By leveraging a modified Gerchberg-Saxton algorithm [46] for discrete-valued optimization, the system effectively tailors the electromagnetic wave manipulation in real-time, enhancing the accuracy and efficiency of the imaging process.

A key contribution of this research is the implementation of an imager that can be electronically reprogrammed in real-time, allowing for the generation of measurement modes required for machine learning techniques, such as Principal Component Analysis (PCA). The authors demonstrate that by using a prototype of the imager, it is possible to achieve high-accuracy object classification and image reconstruction, even in dynamic scenarios. The experimental results validate the effectiveness of the system, showing that it can operate with a significantly reduced data acquisition time, representing an important advancement in the field of imaging and recognition.

The experimental setup consists of a $2\,\mathrm{m} \times 2\,\mathrm{m} \times 2\,\mathrm{m}$ anechoic chamber, equipped with a transmitting horn antenna and a waveguide probe receiver, along with a VNA for data acquisition (Fig. 2.8). The VNA measures the transmission coefficients S_{21} during the experiments to capture the system's response.

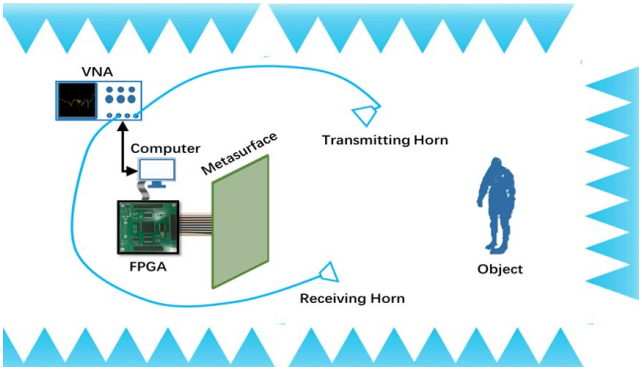


Figure 2.8: Schematic map of the system [45].

2.2.2 Summary

The comparison in Table 2.1 summarizes the key aspects of the RIS-based HAR systems presented in the state-of-the-art.

Approach	**Anechoic Chamber**	**Freq. (GHz)**	**Type of HAR**	**RIS Configuration**
Hu et al. [41]	✓	3.198	Posture Recognition	FCAO algorithm
Li et al. [42]	× (indoor)	2.4	Hand Gesture Recognition, Vital Signs	Gerchberg-Saxton algorithm
Li et al. [43]	× (indoor)	2.4	Posture Recognition	m-ANN
Wang et al. [44]	× (indoor)	2.4	Behavior Monitoring	Random patterns
Li et al. [45]	✓	2.4	Posture Recognition	Gerchberg-Saxton algorithm

Table 2.1: Comparison of RIS-based HAR Approaches

The reviewed works rely on highly controlled settings, such as anechoic chambers or closed laboratory environments, where external interference is minimized. This controlled environment allows researchers to fully explore and understand the contribution of RIS to signal manipulation and activity recognition. However, expanding the application of RIS-based HAR to more dynamic, uncontrolled environments with real-world interference remains an important goal.

All the systems analyzed operate within the standard WiFi frequency, particularly around $2.4\,\mathrm{GHz}$. Inspired by the enhanced network performance offered by WiFi-6E [47], our work will be developed within this novel frequency range, between $5.925\,\mathrm{GHz}$ and $7.1\,\mathrm{GHz}$.

Furthermore, the type of HAR addressed by these studies remains focused primarily on static recognition tasks, such as posture detection and simple gestures. Although these are foundational, future work should look toward more dynamic activities and a broader range of human behaviors.

The configuration of the RIS plays a crucial role in the performance of these systems, with various methods used across studies. Approaches range from optimized algorithms, such as the FCAO and Gerchberg-Saxton algorithms, to more adaptive methods like those based on ANNs. In our work, we will explore the potential of the FCAO algorithm for optimizing RIS configurations.

3. Proposed Approach

This chapter outlines the methodology and experimental strategy for assembling and evaluating a controllable smart radio environment for human activity recognition. The primary focus is on how controlled RF signals illuminate the space of interest, manipulating the emitted and reflected RF characteristics to enhance the effectiveness of activity recognition.

Our approach, presented in Fig. 3.1, begins with the creation of a controlled smart radio environment where the SoI, defined as the area where human activities are to be sensed, is illuminated with managed RF signals. The setup initially involves using a reflective metal plate to manipulate the RF signals. This metal plate configuration is discussed in detail in Chapter 4, where we explore how varying the reflection angles impacts activity recognition by altering the signal path within the environment. Following this, in Chapter 5, we introduce a RIS to achieve more dynamic control of the RF environment. A key element of our approach is recognizing that reflections from the RIS or metal plate are just one part of the process. The SoI itself also plays a crucial role by reflecting RF signals in unique ways, depending on the objects and activities present. This variation in reflection, combined with our control over the environment, allows us to gather comprehensive and relevant data for RF sensing and human activity recognition in smart radio environments.

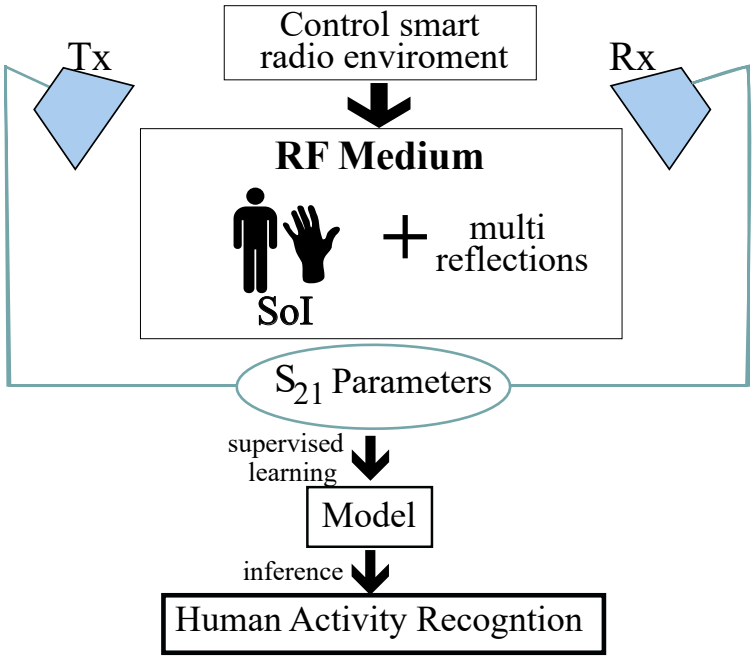


Figure 3.1: System architecture diagram of the general approach for HAR.

The experimental setup involves a transmitter (Tx) and receiver (Rx) of radio signals within this controlled environment, and then measuring the transmission coefficients, known as S_{21} parameters, between the Tx and Rx, which will serve as the data for HAR. By deploying both a metal plate (Chapter 4), and a RIS (Chapter Chapter 5), we were able to create distinct datasets, capturing signals reflected in various ways and configurations. The system was implemented in an anechoic chamber to ensure that external RF interference did not affect the results. This controlled environment allowed us to focus solely on the performance of the specific RF setup we created.

To optimize the RIS configurations, we adopted a strategy similar to the one presented by Hu et al. [41], utilizing the FCAO algorithm, which ensures optimal signal propagation and reflection for accurate activity recognition.

To evaluate the effectiveness of the retrieved data, we employed CNNs, which are well-suited for interpreting RF signal data represented as images. Our use of CNNs to interpret RF signals as images draws inspiration from the approach taken by Li et al. [42]. Firstly, we trained a CNN on a dataset capturing RF signals reflected from a metal plate, with varying angles of emission. Following this, an identical CNN was trained using a dataset that involved different configurations of the RIS. This approach allowed us to comprehensively assess the impact of both reflective surfaces on RF signal interpretation.

The performance of these models was then assessed by comparing their predictions with actual observations of human activities. This comparison enabled us to determine how well the data, including those for full-body postures and hand gestures, provided meaningful insights into activity patterns.

The following two chapters will present the respective implementations, experimental setups, and results.

4. Human Posture Recognition with Metal Reflectors

Recent studies have demonstrated the efficacy of S_{21} parameters for human body-related classification tasks [48–50]. Inspired by these advantages, this chapter explores the viability of using S_{21} parameters for posture recognition in semi-controlled environments.

Due to the initial unavailability of the RIS, a metal plate was employed as a reflective surface to manipulate RF signal propagation. The proposed system, named SpecRF Posture, relies on passive specular reflections as a baseline for future RIS-based solutions. It achieves beam scanning by rotating a steerable horn transmitter at regular intervals to capture a sweep of the S_{21} parameters from the SoI for various human postures.

Figure 4.1 illustrates the three main components of SpecRF Posture: Data Acquisition, Data Processing and Training, and Posture Classification. Initially, data is acquired by mechanically steering the Tx antenna to collect S_{21} parameters. Subsequently, the collected data is processed to extract relevant features, which are then used to train a CNN model. Finally, the trained model is used to classify five human postures.

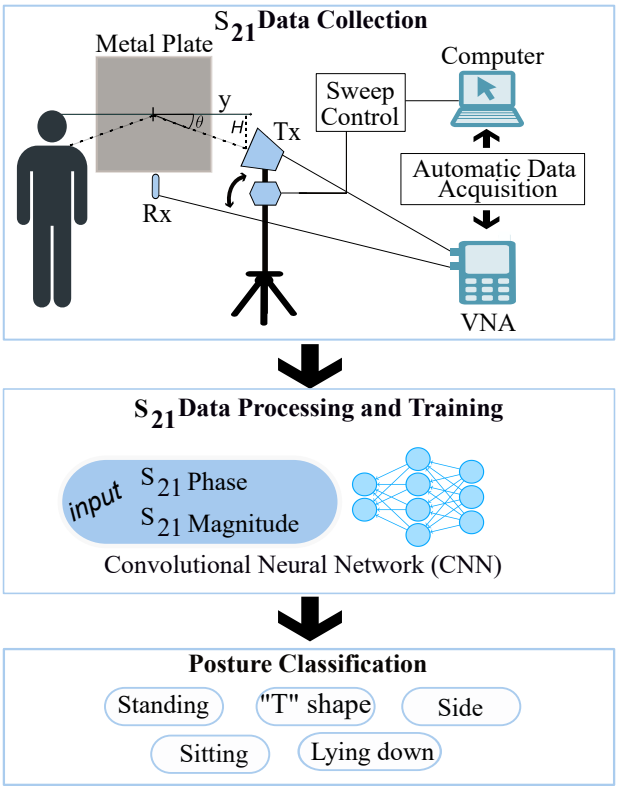


Figure 4.1: HPR System architecture diagram illustrating the three main components: Data Acquisition, Data Processing and Training, and Posture Classification.

The work presented in this chapter led to the publication of a conference paper [16], and the dataset created and used in this work has been made publicly available*.

This chapter is organized as follows: Section 4.1 details the design of the experimental setup, including the configuration of the transmitter and receiver antennas and the reflective surface. Section 4.2 describes the experimental setup within the anechoic chamber, the data acquisition process, and the overall methodology used for the experiments. Section 4.3 provides an overview and analysis of the collected data, including details on the data characteristics and posture classification results. Finally, Section 4.4 summarizes the key findings and conclusions from the experiments.

4.1. Design of the Setup

4.1.1 Antenna Configuration

This section describes the selection and positioning of the antennas used in the experimental system. Proper antenna positioning is crucial to ensure the quality of the transmitted and received signals, and to maximize the efficiency of signal emission and reception. Specifically, we were concerned with maximizing the metal plate area illuminated by the Tx, and that the metal plate would properly reflect towards the SoI.

Transmitter Antenna

Before analyzing the design of the setup, it was necessary to select a feed antenna. Various horn types were evaluated, including Conical, Standard Gain, and Broadband. Considering geometry, both conical and pyramidal horns were considered. The pyramidal horn antennas combine E-plane and H-plane sectoral horns for equal radiation patterns [51], which stood out as an ideal choice for our application. This choice was reinforced by the fact that the beamwidths in the principal planes E and H of the conical horn are generally unequal [52, p. 15-2], which do not meet the symmetrical characteristics required for our setup.

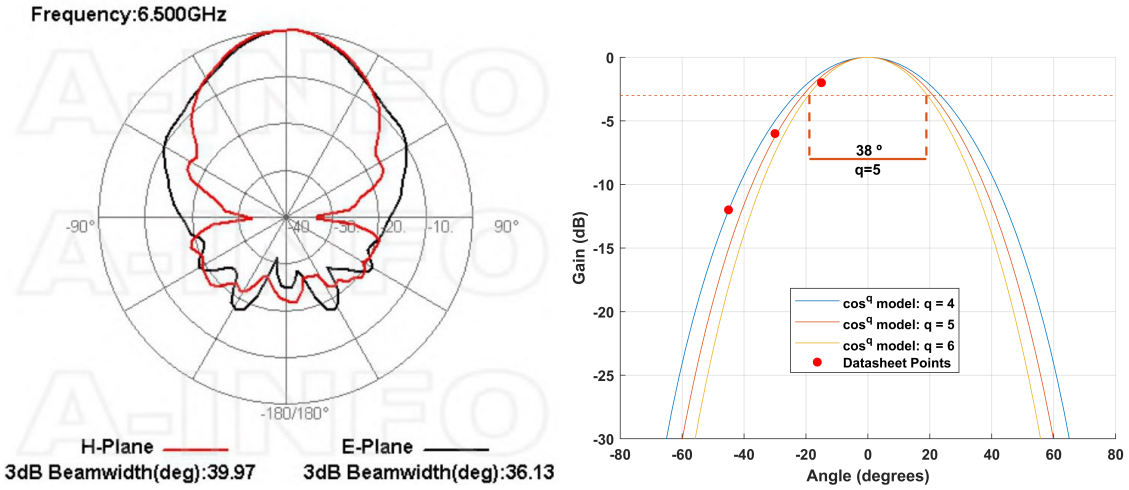
Among the pyramidal horn subtypes, Broadband and Standard Gain, the decision was influenced by the available equipment at the moment. Although Standard Gain horns are ideal for maintaining consistent gain across frequencies, a Broadband Horn Antenna, specifically the LB-20180-SF model with a 2-18 GHz frequency range and 12 dB gain, was selected due to its availability in the laboratory, and because it met our requirements.

*M. Oliveira, F. Ribeiro, N. Paulino, O. Yurduseven and L. Pessoa, “SpecRF-Posture Dataset”. Zenodo, Apr. 03, 2024. doi: 10.5281/zenodo.10911412.

Although the Broadband horn was used in this work, a Standard Gain horn is more suitable for studying signals across a frequency band. A study in Appendix A identified the LB-137-15 from A-INFO as the ideal choice for when a larger RIS (with $1\,\mathrm{m} \times 1\,\mathrm{m}$) is ready and therefore, we ordered this antenna for use in future experiments.

The position of the transmitter antenna was calculated based on a detailed analysis of the aperture efficiency as described in [53, Chapter 3]. This analysis aimed to achieve an edge taper of -10 dB to ensure that strong diffractions don't occur at the aperture edge. By setting an edge taper of -10 dB, the system minimizes unwanted radiation and diffraction effects that can degrade the quality of the transmitted signal.

To establish the ideal distances between components, a detailed analysis of the system's aperture efficiency is essential. The radiation of this antenna was modeled by a cos^q radiation pattern, where the power factor q is the same for both E and H planes ($q_E = q_H$). This balanced feed assumption leads to a rotationally symmetric pattern and negligible back radiation [53]. To determine the q value, a comparative analysis was conducted between the points of radiation pattern at a frequency of 6.5GHz using the antenna's datasheet and the cos^q model (Fig. 4.2). The choice of the 6.5 GHz frequency was made because it represents the central frequency of the Wi-Fi 6E band, which is the frequency range where our experiments will be conducted. This approach led to the choice of $q = 5$.



(a) Radiation Pattern from datasheet

(b) cos^q model patterns

Figure 4.2: E-plane radiation pattern points of LB-20180 datasheet and cos^q pattern. The 3 dB beamwidth of the E-plane from the datasheet and from the cos^5 model are very close.

The geometric configuration of the system is illustrated in Fig. 4.3. The metal plate aperture lies in the xy-plane, with its center positioned at the coordinate system's center. The feed antenna, denoted as Tx, is characterized by an offset angle θ_0 and situated at a distance H from the xy-plane *. It is oriented towards the Feed Beam Point.

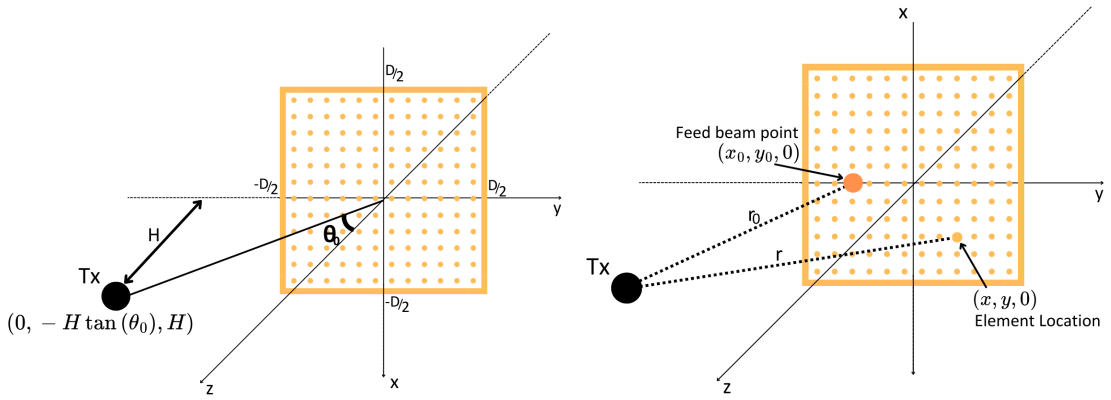


Figure 4.3: Geometric configuration of the system: offset feed towards aperture.

To facilitate the analysis, all the important geometric quantities involved are listed in Table 4.1, given the configuration of the Fig. 4.3, based on the equations provided in [54].

Parameter	**Formula**
Feed Location	$(0, -Htan\theta_0, H)$
Feed Beam Point (FBP)	$(x_0, y_0, 0)$
Element location	$(x, y, 0)$
Distance between feed and FBP	$r_0 = \sqrt{(x_0^2 + y_0^2 + H^2sec^2(\theta_0) + y_0(2Htan\theta_0)}$
Distance between feed and element	$r = \sqrt{(x^2 + y^2 + H^2sec^2(\theta_0) + y(2Htan\theta_0)}$
Distance between FBP and element	$s = \sqrt{(x - x_0)^2 + (y - y_0)^2}$

Table 4.1: Geometric quantities for the system configuration shown in Fig. 4.3.

To achieve the final goal of maximizing aperture efficiency and achieving -10dB of edge taper, the initial step involves a detailed analysis of the aperture's illumination efficiency (η_i) and spillover efficiency (η_s). Figure 4.4 illustrates these two efficiencies, showing how they impact the aperture loss based on the feed position relative to the aperture.

The aperture efficiency is defined as their product:

$$\eta_a = \eta_s \cdot \eta_i \tag{4.1}$$

*In Fig. 4.3, θ_0 represents the absolute value of $90^\circ - \theta$, where θ is the angle shown in Fig. 4.1. Specifically, θ_0 corresponds to the angle between the Tx antenna and the normal to the metal plate, while θ represents the angle between the Tx antenna and the y-axis in the plane of the metal plate.

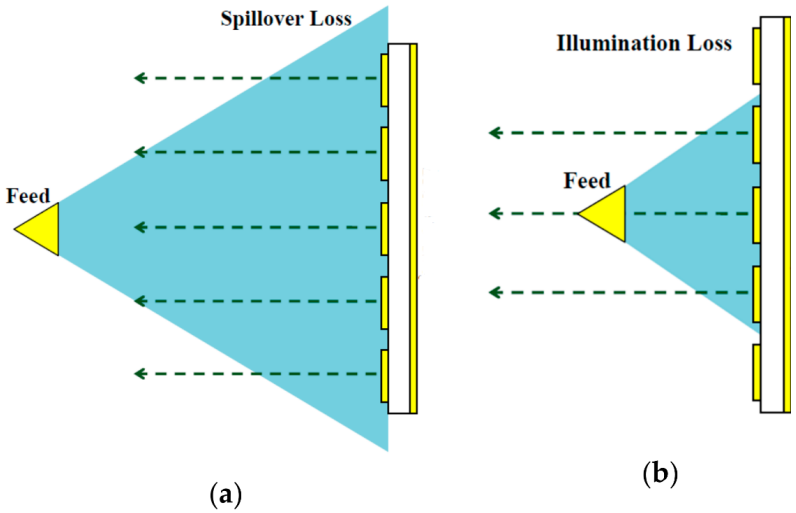


Figure 4.4: Illustration of (a) spillover and (b) illumination losses. Spillover loss occurs when the feed is too far from the aperture, while illumination loss occurs when the feed is too close. From [55].

The spillover efficiency (η_s) quantifies the fraction of power radiated by the feed that is captured by the reflecting aperture, compared to the total radiated power. It is defined as:

$$\eta_s = \frac{\iint_A P(r)\, dA}{\iint_{\text{sphere}} P(r)\, dA} \tag{4.2}$$

where both integrals represent fluxes of the Poynting vector, $P(r)$. The numerator calculates the power intercepted by the aperture, while the denominator represents the total radiated power integrated over a spherical surface centered at the feed. $P(r)$ can be expresed as $\hat{r}\frac{cos^{2q}\theta}{r^2}$ with $\theta \in [0; \frac{\pi}{2}]$ [55]. After performing the integration, the denominator simplifies to $\frac{2\pi}{2q+1}$ *. To evaluate the integral in the numerator, the Poynting vector is expressed in rectangular coordinates, leading to the following expression for the integrand:

$$P(r)dA = \frac{H}{r^3}\left(\frac{r_0^2 + r^2 - s^2}{2r_0 r}\right)^{2q} dxdy \tag{4.3}$$

The illumination efficiency (η_i) is defined by:

$$\eta_i = \frac{1}{A}\frac{|\iint_A I(x,y)\, dA|^2}{\iint_A |I(x,y)|^2\, dA} \tag{4.4}$$

Here, $I(x,y)$ is the amplitude distribution over the aperture, described by:

$$I(x,y) = \frac{H^{q_e}}{r^{1+q_e}}\left(\frac{r_0^2 + r^2 - s^2}{2r_0 r}\right)^{q} \tag{4.5}$$

where q_e is the element pattern power factor, conventionally defined as 1.

* $\int_0^{2\pi}\int_0^{\frac{\pi}{2}} \cos^{2q}(\theta)\sin(\theta)\, d\theta\, d\phi = \frac{2\pi}{2q+1}$.

Following this understanding of efficiencies, the study is used for the determination of key parameters such as θ_0, H and (x0, y0). By optimizing these parameters, the aim is to enhance the system's performance and, consequently, achieve the highest possible aperture efficiency.

The feed location is dependent on the parameters θ_0 and H. In this regard,a detailed analysis of the effect of these parameters on the variation of efficiencies was conducted, considering the values $q = 5$ and $q_e = 1$.

Since we are considering a metal plate here, where only specular reflections occur, the angle θ_0 is set to 45 degrees to avoid radiation distortion, as will be explored in the next section. However, in the case of the aperture being a RIS, this angle should be optimized, as done in the next chapter and in Appendix A.

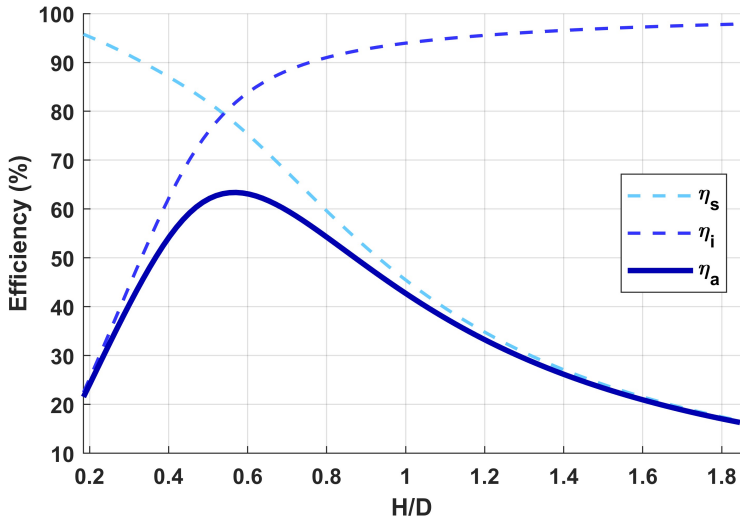


Figure 4.5: Efficiencies as a function of H/D with chosen $\theta = 45°$.

In Figure 4.5 the variation with H/D was explored, where D = 1m, and θ_0 is 45 degrees. As H increases, η_s decreases due to the increased distance between the metal plate and the feed source, resulting in a reduction of the aperture angle on the aperture plane. On the other hand, η_i increases as a more uniform field is generated in the aperture plane. Following some manipulation of values, the chosen H was determined to be 0.45m.

Finally, the positioning of the feed beam point (FBP) was investigated, specifically the values of x_0 and y_0. On the x-axis, the expected symmetric behavior was identified ($x_0 = 0\,\mathrm{m}$), while on the y-axis, an asymmetry was identified. However, since the experiment involves rotating the transmitter antenna along the y-axis on the metal plate, choosing an ideal y_0 is not necessary. Again, in the case of using a RIS, since the position of the Tx

antenna remains static, the FBP should be selected, as done in the next chapter and in Appendix A.

Receiver Antenna

The receiver antenna, referred to as the Rx antenna, plays a crucial role in capturing signals within the space of space. It must possess the capability to receive signals from all directions, allowing for effective detection of reflections from various locations within the SoI. Given these requirements, an omnidirectional vertical antenna was selected to ensure that the receiver has a wide field of view and can accurately capture reflection signals from all angles.

To optimize its performance, the Rx antenna is positioned in the same plane as the metal plate. This configuration ensures that the antenna primarily captures reflections that have interacted with the target space, rather than receiving direct reflections from the metal plate itself.

4.1.2 Reflective surface (Metal Plate)

As mentioned, a metal plate was utilized to illuminate the SoI, in the place of a RIS, which was still being developed in parallel by other members of our research group. Consequently, the approach in this Chapter evaluates the achievable human activity recognition performance relying on a low-cost architecture based solely on passive specular reflections occurring in the path between a steerable horn transmitter, a reflective surface, and the SoI.

Specular reflections are the primary type occurring in the path up to the SoI. However, upon reaching the SoI, reflections can occur in multiple directions depending on the target within the SoI. Therefore, it is essential to consider the illumination on the reflective surface and its subsequent reflection onto the SoI, to define the distances and arrangement of the components. As discussed in the previous section, the radiation emitted by the Tx can be modeled by a cos^q radiation pattern, where the power factor q is 5. Given that the radiation emitted by the Tx antenna is approximated by a cone, radiation distortion occurs after being reflected by the metal plate, as illustrated in Fig. 4.6. Consequently, the radiation incident on the central plane of the SoI exhibits a more dispersed distribution.

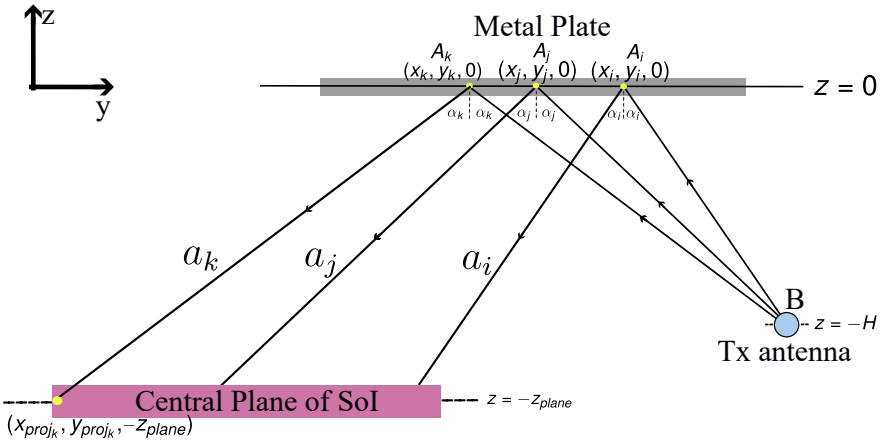


Figure 4.6: Top view of the Tx antenna emission towards the metal plate and distortion after reflection.

To quantify this phenomenon, the illumination pattern on the metal plate and its reflection on the central plane of the SoI were computed to visualize how the radiation varies with the emission angle of the Tx antenna, using Eq. (4.5).

To determine the subsequent illumination pattern in the SoI after reflection and distortion, the following equations were employed:

Geometric Projection: Projects a point $(x_i, y_i, 0)$ along the direction of the reflected vector $\vec{R}_i$ onto z_{plane}, where a_i is the scaling factor that adjusts the vector projection onto this plane:

$$(x_{proj_i}, y_{proj_i}, z_{plane}) = (x_i, y_i, 0) + a_i \cdot \vec{R}_i \tag{4.6}$$

$$a_i = \frac{z_{plane}}{\cos(\alpha_i)} \tag{4.7}$$

Vector Reflection: Calculates the direction of a vector after it has reflected off a surface:

$$\vec{R}_i = \vec{A_iB} - 2 \cdot (\vec{A_iB} \cdot \vec{n}) \cdot \vec{n} \tag{4.8}$$

where $\vec{A_iB}$ is the vector between the Tx antenna, represented by $B = (0, H \cdot \tan\theta_0, -H)$, and each point $A_i = (x_i, y_i, 0)$ of the metal plate. The normal vector to the surface is $\vec{n} = (0, 0, -1)$.

To illustrate these effects, Fig. 4.7 shows the illumination patterns when the Tx antenna is directed at three different points on the metal plate: $y_0 = -0.5\,\mathrm{m}$, $0\,\mathrm{m}$, and $0.5\,\mathrm{m}$. Additionally, it shows the resulting illumination pattern on the central plane of the SoI after the reflection and distortion. This schematic representation of the experimental setup illustrates the geometric relationships between the metal plate, Tx antenna, and the space occupied by the individual. The Tx antenna is characterized by $(\theta_0, H) = (45^\circ, 0.45\,\mathrm{m})$,

resulting in a $0.64\,\mathrm{m}$ $(= H/\cos(\theta_0))$ distance from the Tx to the center of the metal plate.

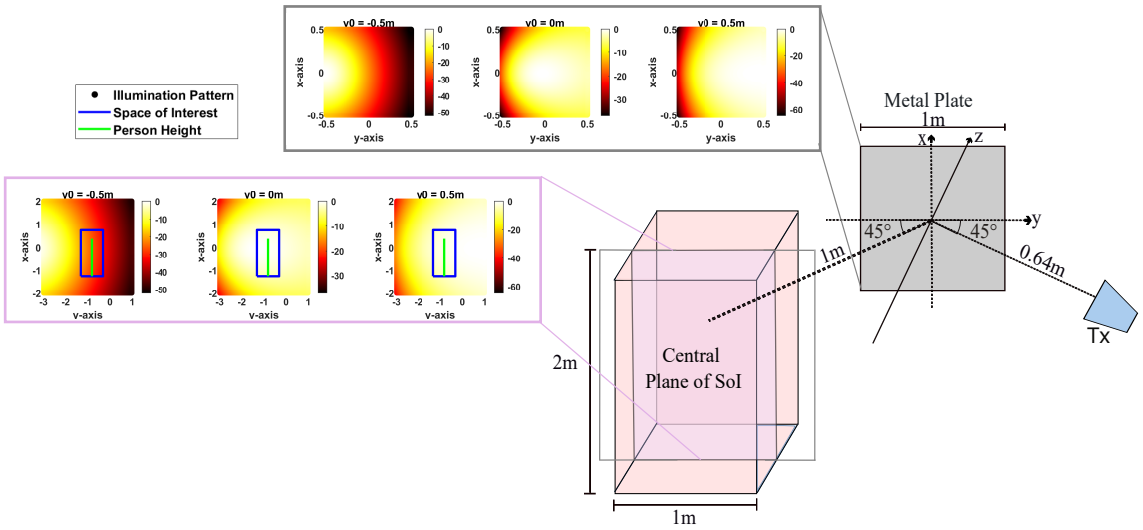


Figure 4.7: Schematic of the experimental setup illustrating the spatial configuration between setup elements. The figure also displays the illumination patterns for the Tx pointing towards different positions ($y_0 = -0.5\,\mathrm{m}$, $0\,\mathrm{m}$, and $0.5\,\mathrm{m}$) on the metal plate, as well as the corresponding distorted reflection patterns in the central plane of SoI.

4.2. Experimental Setup and Methodology

This section presents the final experimental setup and methodology used for human posture recognition with metal reflectors. Details of the preliminary experiments, which motivated the use of Low Noise Amplifiers (LNAs) and the selection of the anechoic chamber environment, are provided in Appendix B.

The experimental setup was assembled in an anechoic chamber at the Faculty of Engineering of the University of Porto (FEUP), to avoid interference from external signals and ensure accurate measurements. Fig. 4.8 provides a view of the setup from the perspective of the SoI. A metal plate with dimensions $1\,\mathrm{m} \times 1\,\mathrm{m}$ $(1\,\mathrm{m}^2)$ was used as the reflective surface.

The transmitter (Tx) antenna employed was the A-INFO LB-20180 horn antenna, selected in Section 4.1. It was positioned at an angle of 45 degrees to the metal plate. The omnidirectional antenna used as the receiver (Rx) was placed immediately below the metal plate, with a 60 cm offset from the center of the plate. This placement allows the Rx antenna to detect signals reflected from various directions within the SoI, facilitating the acquisition of more diverse data for posture recognition.

To mitigate the Line-of-Sight (LoS) component between the Tx and Rx antennas, a column of absorbers was strategically positioned, as illustrated in the right-hand side image

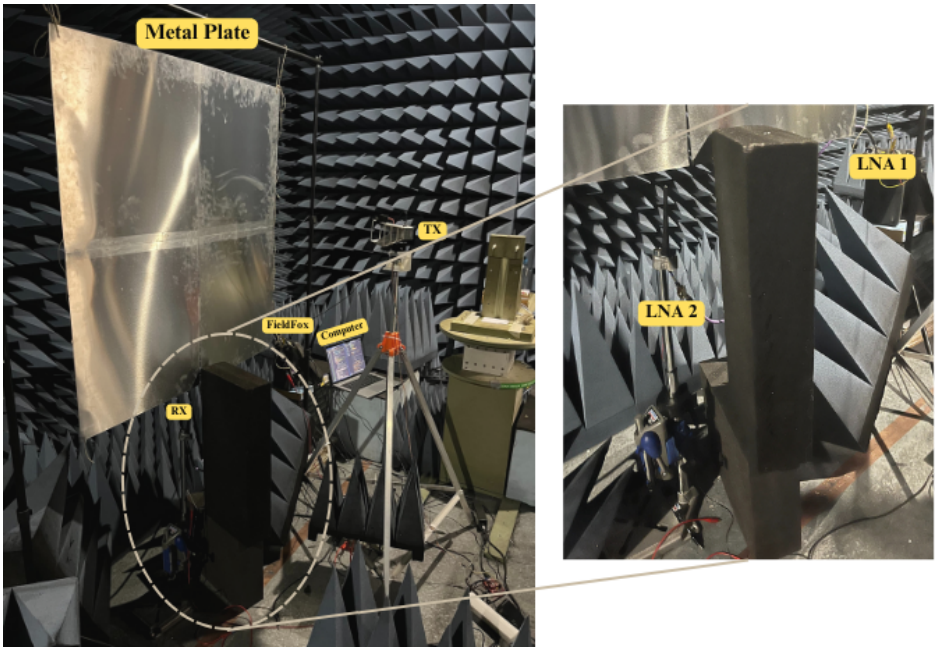


Figure 4.8: Experimental set up for HPR inside the anechoic chamber.

of Fig. 4.8. Additionally, two LNAs (ZX60-153LN-S+) were connected to the input and output ports of the Tx and Rx to amplify the transmitted and received RF signals, respectively, to compensate for any signal loss caused by absorption and reflections from the individual within the SoI. To manage data collection and control the antenna movement, a computer and a VNA (Keysight Fieldfox N9914A) were positioned outside the experimental area. The data acquisition system, developed by a member of our workgroup synchronized the rotation of the antenna and the collection of S_{21} measurements. This system automatically controlled the antenna's position through an Arduino Uno and a stepper motor, ensuring precise alignment. For each specified direction, the system triggered the VNA to capture the S_{21} parameters. The software and hardware setup for this system are publicly available *.

During each experiment, an individual assumed different postures within the SoI while the antenna systematically rotated to target various positions along the y-axis of the metal plate. The positions ranged from (0, $-0.65\,\mathrm{m}$, 0) to (0, $0.65\,\mathrm{m}$, 0), with intervals of $0.05\,\mathrm{m}$. This setup created a variety of illumination scenarios within the SoI, as illustrated by the illumination patterns shown in Fig. 4.7. The spacing between measurement points was determined by the sweep control unit's resolution, which is limited to 1.8°, and is also close to the wavelength at the highest frequency of $0.046\,\mathrm{m}$ in the chosen frequency band. Each y_0 position along the y-axis corresponds to a specific angle between the Tx antenna

*F. Ribeiro and M. Oliveira, "SpecRF-Posture". GitHub, Apr. 03, 2024. `https://github.com/franciscombr/SpecRF-Posture`

and the y-axis. At each of these angles, the S_{21} parameters were recorded across the frequency range of $5.925\,\mathrm{GHz}$ to $6.5\,\mathrm{GHz}$ using a VNA.

4.3. Results

This section presents the results of the study, divided into two main parts. First, the dataset created during the experiments is described and analyzed, providing insights into the characteristics of the collected data for each posture. Second, the performance of a CNN model trained for classifying human postures using the dataset is evaluated.

4.3.1 Dataset Description and Analysis

To create a dataset for posture recognition, data was collected for five distinct postures: Standing, T-shape, Side, Sitting, and Lying Down. A single volunteer assumed each posture, and 120 samples (referred to as *sweeps*) were collected for each one, resulting in a dataset of 600 samples $(1$ subject $\times\ 5$ postures $\times\ 120$ sweeps$)$.

Each sample is characterized by dimensions $(n_{\text{resolution}}, n_{\text{angles}})$, where $n_{\text{resolution}} = 151$ corresponds to the number of frequency points between $5.925\,\mathrm{GHz}$ and $6.5\,\mathrm{GHz}$, and $n_{\text{angles}} = 25$ represents the number of discrete angles through which the Tx antenna rotates.

To introduce variability and ensure that the dataset captures a range of realistic human movements, minor adjustments of the hands or head were incorporated while maintaining the overall posture. This approach helps in making the system more robust to different ways of assuming the same posture.

Each *sweep* required approximately 35 seconds for data collection, resulting in a total acquisition time of around 6 hours for the entire dataset. Due to the extensive time required for data acquisition, experiments were limited to a single volunteer. Spending 6 hours inside an anechoic chamber while maintaining static postures is not practical for multiple subjects. Therefore, this study focuses on demonstrating the feasibility of using S_{21} parameters for human posture recognition based on data from one individual.

Figure 4.9 illustrates an example of the variation in phase and magnitude of the S_{21} parameters with respect to the rotation angle and frequency within the WiFi-6E band for each posture.

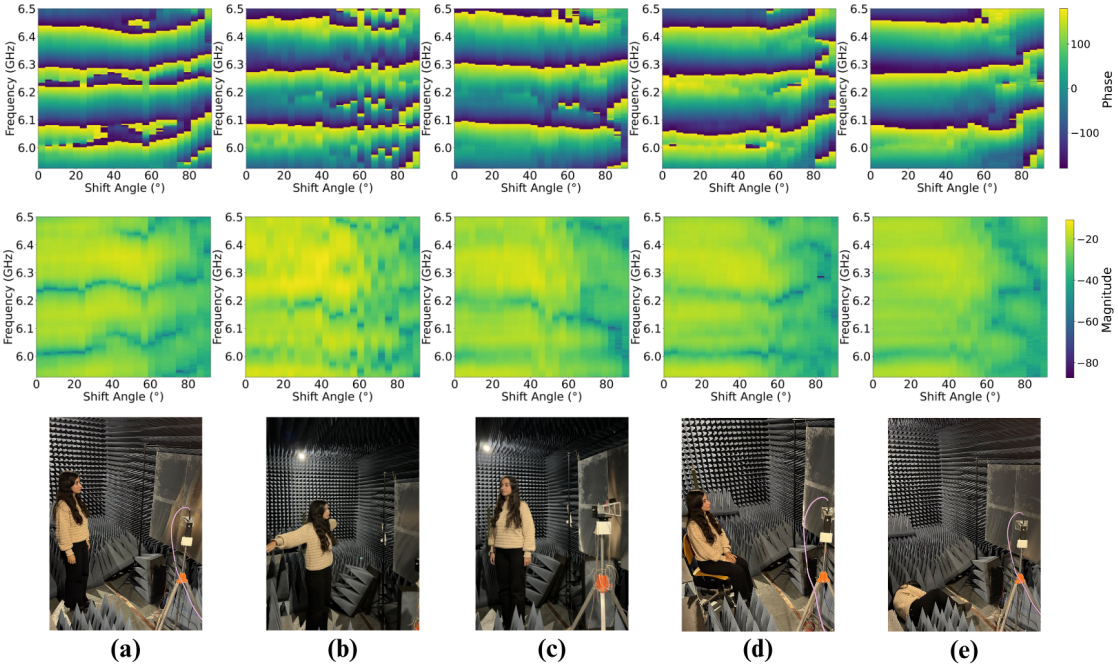


Figure 4.9: Frequency vs. Angle Analysis for S_{21} Phase (Top), Magnitude (Middle), and Posture Identification (Bottom) for: (a) Standing, (b) T-Shape, (c) Side, (d) Sitting, and (e) Lying Down.

The scanning range of the Tx antenna spanned from approximately $-0.65\,\mathrm{m}$ to $0.65\,\mathrm{m}$ along the y-axis, covering a broader area than the physical dimensions of the metal plate ($-0.50\,\mathrm{m}$ to $0.50\,\mathrm{m}$). This extensive scanning range ensured that the region of interest was adequately illuminated and captured in the dataset. By covering a wider area, we could determine the optimal illumination angles by analyzing the frequency-angle images of both magnitude and phase. The 0 degrees angle corresponds to the Tx antenna pointing at $-0.65\,\mathrm{m}$, while the maximum angle of 85 degrees corresponds to the antenna pointing at $0.65\,\mathrm{m}$.

The magnitude analysis revealed that at lower angles, higher intensity is observed due to the omnidirectional antenna capturing more information from the illuminated SoI. In contrast, higher angles showed less intensity as the antenna captured information from regions less effectively illuminated by the Tx antenna. Similarly, the phase analysis indicated greater variation at larger angles, attributed to inadequate illumination of the metal plate and irrelevant reflections captured by the Rx antenna.

In conclusion, we can accurately characterize the SoI up to approximately 50 degrees, corresponding to $y_0 = 0.45\,\mathrm{m}$. Within this range, distinctive profiles of magnitude and phase were observed for each posture, indicating that the received signal varies with different human body postures.

4.3.2 Posture Classification

The classification of human postures is performed using a CNN, which has proven effective in pattern recognition tasks for images and multidimensional signals [56]. Furthermore, this ability of CNNs to learn and extract relevant features from complex datasets has already been demonstrated to be applicable for radio signal classification tasks in recent research [57–61].

Fig. 4.10 illustrates the architecture of the CNN designed for posture classification. The network's input consists of angle-frequency images that represent the magnitude and phase of the S_{21} parameters, processed through two separate channels. The architecture includes multiple layers including convolutional layers, max-pooling layers, and dense layers. The output layer contains five nodes, each corresponding to one of the five postures: Standing, T-Shape, Side, Sitting, and Lying Down. The parameters used for the training of the CNN are detailed in Table 4.2.

Parameter	Value
Learning rate	0.00001
Training epochs	200
Optimizer	Adam
Loss Function	Categorical Crossentropy

Table 4.2: Hyperparameter configurations of CNN for SpecRF based Human Posture Recognition (HPR).

To assess the performance of the CNN model, the dataset was divided into training and testing subsets, with 80% of the data used for training and 20% reserved for testing. The raw data was processed to extract magnitude and phase information, which were then formatted into image-like structures suitable for CNN input.

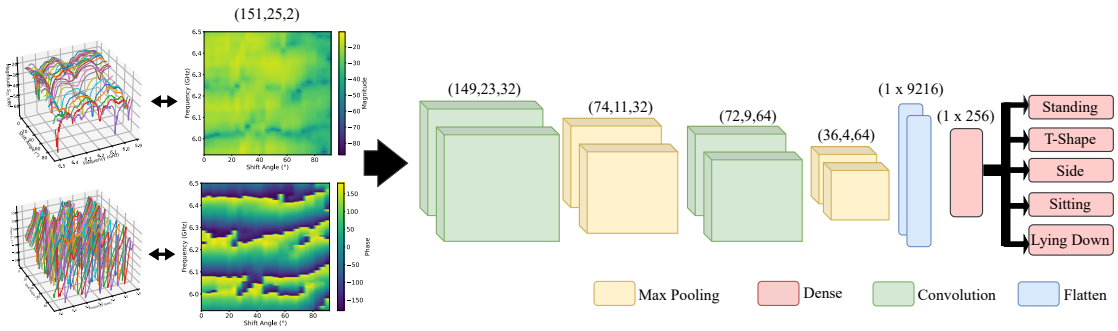


Figure 4.10: Architecture of the CNN developed for posture classification (Standing, T-Shape, Side, Sitting, Lying Down). The input to the network consists of the magnitude and phase of the S_{21} parameters, which can be represented both as a 3D plot and as a frequency-angle image, where the z-axis is represented as color data.

Fig. 4.11 compares two classification approaches. The first approach uses a two-channel input where one channel represents the magnitude and the other represents the phase of the S_{21} parameters. The second approach uses a single channel representing only the magnitude. The CNN achieved an impressive average accuracy of 99.17% on the validation dataset with the two-channel method. Accuracy rates were 100% for all postures except for the *Side* posture, which achieved a 96% accuracy.

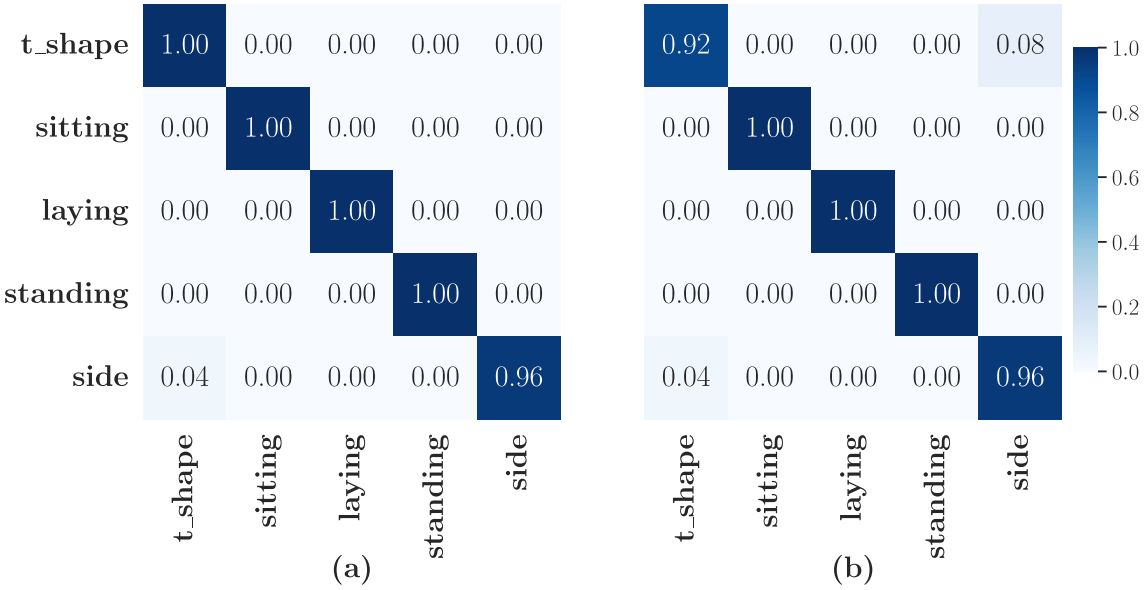


Figure 4.11: Posture Recognition accuracy utilizing two distinct approaches based on the S_{21} parameters: (a) incorporation of both magnitude and phase as input parameters for the CNN model, and (b) utilization of only magnitude as the input.

In comparison, the single-channel approach, which utilized only magnitude, showed a decrease in accuracy for the *T-Shape* posture, with accuracy dropping to 96%. This suggests that including phase information enhances the model's ability to distinguish between the *Side* and *T-Shape* postures, leading to more precise classification.

4.4. Summary

This chapter introduces an approach for human posture recognition using a simple hardware configuration. By leveraging a metal plate to passively manipulate RF signals through specular reflections, the system simplifies the HPR task. This approach eliminates the need for complex hardware setups, offering a cost-effective solution for creating semi-controllable signal paths using WiFi-6E signals.

Additionally, a design model for experimental setups involving the illumination of apertures, such as metal plates or RIS, is presented. This model, implemented in MATLAB,

allows for the manipulation of parameters such as H, θ_0, and the FBP to determine the optimal illumination characteristics of the aperture.

This chapter establishes a foundational approach for human posture recognition using a metal plate. The next chapter will extend this work by integrating a RIS in place of the metal plate, which will enable configurable signal paths, without the need of a steerable Tx antenna.

5. Hand Gesture Recognition with a RIS

In the previous chapter, we used a metal plate to study human posture recognition with RF signals. This chapter advances the work by incorporating a RIS, developed by members of our research team. Unlike the metal plate, the RIS enables dynamic and configurable signal paths, providing greater flexibility in RF signal manipulation [62–64].

Given the RIS's smaller dimensions ($18.46\,\mathrm{cm} \times 18.46\,\mathrm{cm}$) compared to the metal plate ($1\,\mathrm{m} \times 1\,\mathrm{m}$), the focus of the classification task shifts from human posture recognition to hand gesture recognition. The reduced size of the RIS requires a smaller area of interest, making hand gestures a more practical and appropriate application for this setup.

We now present an enhanced version based on the system proposed in Chapter 4. This new system uses the RIS's capability to modify the direction of the reflected beam. Instead of steering the Tx antenna to perform sweeps of the S_{21} parameters, the Tx antenna remains fixed and we vary the RIS configurations. Fig. 5.1 provides an overview of the main components in the new system, based on the Spec-RF Posture diagram (Fig. 4.1).

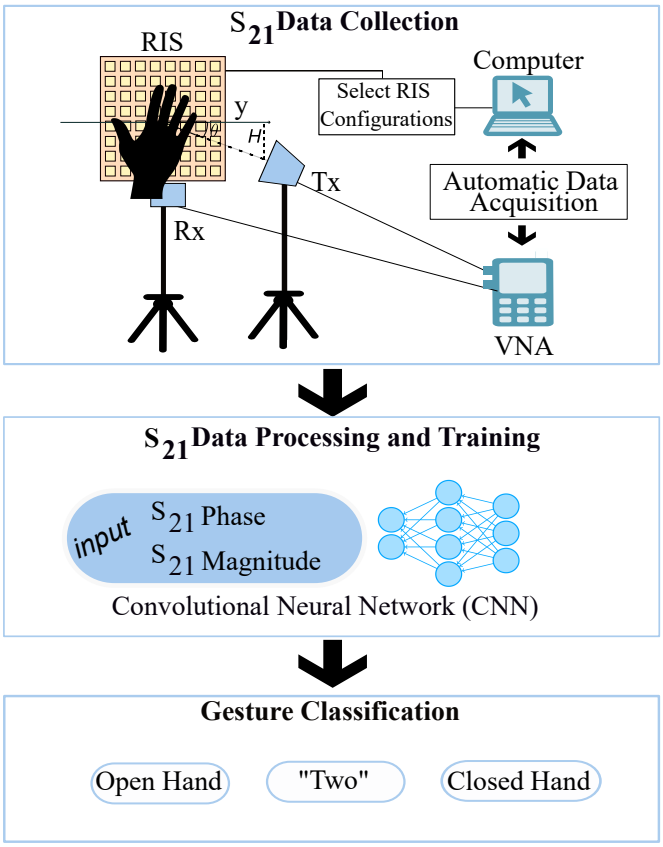


Figure 5.1: System architecture diagram illustrating the three main components: Data Acquisition, Data Processing and Training, and Gesture Classification.

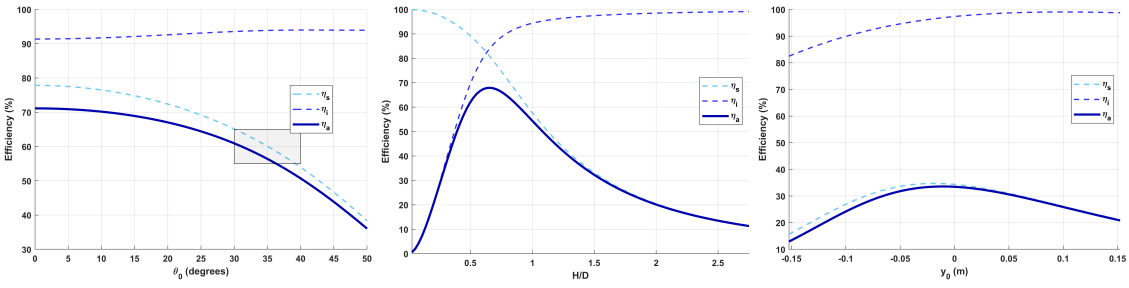
The dataset created and used in this work has been made publicly available*. This chapter is organized as follows: Section 5.1 covers the design of the experimental setup and the FCAO algorithm for defining the RIS configurations. Section 5.2 details the RIS and preliminary tests validating its functionality. Section 5.3 describes the experimental setup and methodology used for the hand gesture recognition. Section 5.4 provides the results for both a real hand and a wooden hand, evaluating the accuracy of gesture classification. Finally, Section 5.5 summarizes the conclusions from the experiments.

5.1. System Configuration

5.1.1 Design of Set up

Using the MATLAB model presented in the previous chapter, the parameters (H, θ_0) and the coordinates of the FBP were determined for the RIS-based setup. The LB-20180-SF antenna was used again for transmission, resulting in $q = 5$. As before, an offset feed was chosen to prevent the antenna from interfering with the SoI. However, since the system no longer relies on specular reflections, the SoI is placed directly in front of the RIS.

For our application, a highly inclined feed (high θ_0) would lead to inadequate illumination of the RIS, and being nearly centered on the RIS (small θ_0) would result in interference with the region of interest. Therefore, θ_0 was selected in the range $[30; 45]^\circ$ (Fig. 5.2a). Considering the aperture efficiency analysis and the RIS size ($18.46\,\mathrm{cm}^2$), the optimal values were determined to be $(H, \theta_0) = (33\,\mathrm{cm}, 35^\circ)$ and $FBP = (0, -0.02\,\mathrm{cm}, 0)$ (Fig. 5.2). This setup resulted in about 35% efficiency due to the minimum possible H in the anechoic chamber. Higher efficiency would have required $H \approx 16\,\mathrm{cm}$, but this wasn't feasible due to physical limitations and potential signal interference with the SoI.



(a) function of θ_0 (b) function of H/D (c) function of y_0

Figure 5.2: Efficiencies as functions of the design parameters of the setup.

*M. S. F. de B. Oliveira, F. M. Ribeiro, N. Paulino and L. M. Pessoa, "RIS Based Hand Gesture Recognition Dataset". Zenodo, Sep. 12, 2024. doi: 10.5281/zenodo.13754235.

Given the compact system and smaller SoI, the receiver antenna was chosen to be a horn antenna of the same model as the Tx to ensure that the Rx mainly captures signals from the SoI, minimizing the LoS component.

5.1.2 FCAO Algorithm for RIS Optimization

To determine the optimal configurations of the RIS to enhance the hand gesture recognition system, we draw inspiration from the work of [41], where the FCAO protocol is proposed for dynamic control of RIS elements. The FCAO protocol operates periodically, dividing time into frames during which each element of the RIS sequentially toggles between ON and OFF states. The key parameter to be optimized in this protocol is the duration each RIS element remains in its respective state within each frame.

First, the underlying theory of the RIS optimization process will be presented, followed by a detailed description of the algorithmic approach, and finally the results of the optimization process will be discussed.

Theory

The setup used for this task is shown in Fig. 5.3, including the distances between components and their respective angles. The RIS consists of $N \times N$ elements (where $N = 8$), divided into $L = 16$ square groups, each containing $N_l = 4$ elements. Each element of the RIS is controlled by a single PIN diode, which can be in two states, ON or OFF. Therefore, each element can switch between two states, s_i, with $i \in \{1, 2\}$ ($N_a = 2$). For simplicity of control, it is assumed that all elements within the same group l are in the same state s_i. The space of interest, located in front of the RIS, is divided into $M = 32$ equal cuboids. Each cuboid is characterized by a reflection coefficient η_m, which is zero when no object is present and non-zero otherwise.

In this RF medium, the electromagnetic waves emitted by the transmitter propagate in the environment, where they can be reflected, scattered, or diffracted by the walls and elements present in the setup. The most significant component in this scenario involves multiple reflections from the transmitter to RIS and the subsequent propagation of the signal to the space of interest, where further reflections occur off the object within the space of interest, ultimately captured by the Rx antenna. Thus, considering the states of

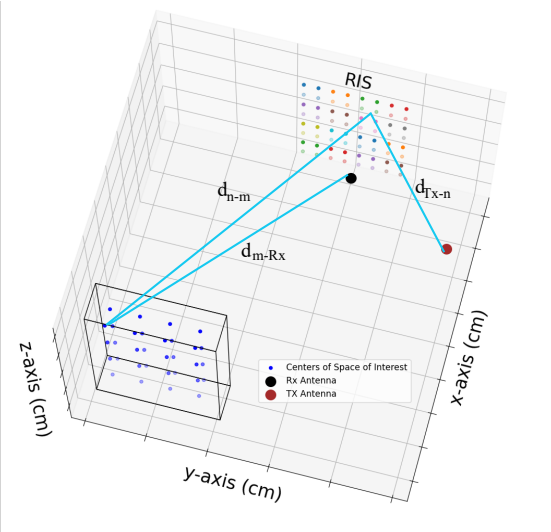


Figure 5.3: Top-view of Setup Schematic.

all groups as the vector $\mathbf{s}_L$, the signal at the receiver can be expressed as:

$$y_{reflection} = \sum_{m\in[1,M]} \sum_{l\in[1,L]} \sum_{n\in\mathcal{N}_l} h_{n,m}(s_l, \eta_m) \cdot P_t \cdot x \tag{5.1}$$

where P_t represents the Tx Power, x is a unit baseband with frequency f_c and h_r represents the channel gain that, based on [65], is represented by:

$$h_{n,m}(s_l, \eta_m) = \frac{\lambda \cdot r_{n,m}(s_i) \cdot \eta_m \cdot \sqrt{gT \cdot gR}}{4\pi \cdot d_{Tx-n} \cdot d_{n-m} \cdot d_{m-Rx}} \cdot e^{-j2\pi(d_{Tx-n}+d_{n-m}+d_{m-Rx})/\lambda} \tag{5.2}$$

where λ is the wavelength of the signal, gT and gR denote the gains of the transmitter and the receiver, respectively, and d_{Tx-n} denotes the distance from the Tx antenna to the n-th RIS element, d_{n-m} represents the distance from the n-th RIS element in the l-th group to the m-th spatial block and d_{m-Rx} is the distance from the m-th spatial block to the Rx antenna, as represented in Fig. 5.3. Also, $r_{n,m}$ are the reflection coefficients of the RIS in each state s_i, and η_m are the reflection coefficients of each cuboid m in the SoI.

Now let's define the dynamic protocol for changing the RIS configurations. The proposed protocol operates in a periodic time frame, where each frame has a duration δ. Instead of changing the configurations of all RIS elements simultaneously, each RIS element sequentially transitions between states s_1 and s_{Na} within each frame. The time that each element remains in a given state is a key parameter to be optimized.

For the l-th group of elements, the duration in each state is represented by the vector $\tilde{t}_l = (\tilde{t}_{l,1}, \ldots, \tilde{t}_{l,Na})^T$, with $\sum_{i=1}^{Na} \tilde{t}_{l,i} = \delta$. The complete matrix describing the time configuration

T, with dimensions $K \times (LN_a)$ where $K = 10$ frames, is defined as:

$$\begin{aligned} T &= \begin{pmatrix} \mathbf{t_1} & \mathbf{t_2} & \cdots & \mathbf{t_K} \end{pmatrix}^T \\ \text{where} \quad \mathbf{t_k} &= [t_{1,1} \;\; t_{1,Na} \;\; \ldots \;\; t_{L,1} \;\; t_{L,Na}] \end{aligned} \tag{5.3}$$

In conclusion, the received signal can be computed as:

$$y_{reflection} = P_t \cdot x \cdot T \cdot A \cdot \eta, \quad \text{with } \Gamma = T \cdot A \tag{5.4}$$

where matrix A represents the channel gain, with elements denoted by α. Each element represents the channel gain of the radio paths from Tx to Rx through the L RIS groups in the different states and the M spatial blocks, resulting in a dimension of $(LN_a) \times M$:

$$\begin{aligned} A &= [\boldsymbol{\alpha_1} \quad \boldsymbol{\alpha_2} \quad \ldots \quad \boldsymbol{\alpha_M}] \\ \boldsymbol{\alpha_m} &= [\alpha_{m,1,1} \quad \ldots \quad \alpha_{m,1,Na} \quad \ldots \quad \alpha_{m,L,1} \quad \ldots \quad \alpha_{m,L,Na}]^T \\ \alpha_{m,l,i} &= \sum_{n \in N_l} \frac{\lambda \cdot r_{n,m}(s_i) \cdot \eta_m \cdot \sqrt{gT \cdot gR}}{4\pi \cdot d_{Tx-n} \cdot d_{n-m} \cdot d_{m-Rx}} \cdot e^{-j2\pi \frac{d_{Tx-n} + d_{n-m} + d_{m-Rx}}{\lambda}} \end{aligned} \tag{5.5}$$

Given the measurement vector $y_{reflection}$, to effectively characterize the space of interest, it is crucial to be able to reconstruct the vector η. This reconstruction can be achieved through Compressive Sensing (CS). In this context, the signal model is given by:

$$y_{reflection} = \Gamma \cdot \eta + z \tag{5.6}$$

where $\Gamma \in \mathbb{C}^{K \times M}$ represents the measurement matrix, $z \in \mathbb{C}^K$ denotes additive noise, and $\eta \in \mathbb{C}^M$. Considering that the dimension of $y_{reflection} \in \mathbb{C}^K$ is much smaller than that of η (i.e., $K \ll M$), we deal with an underdetermined system. To accurately solve this underdetermined system using CS, two key conditions must be met [66]: 1) The vector η should be sparse, and 2) The matrix Γ needs to be appropriately designed.

Regarding the first condition, the space of interest is characterized by a cuboid region divided into M equal-sized cubes. Each cube m yields a reflection coefficient, resulting in a vector of reflection coefficients $\eta = [\eta_1, ..., \eta_M]$. As the object to be detected will only be present in a few of the blocks, most of the values η_i will be zero. Hence, η is expected to be sparse. Regarding the second condition, the matrix Γ is dependent on the system characteristics, such as distances between components and objects in the space of interest. The configuration of the RIS elements influences signal propagation and plays

a significant role in the design of Γ. Therefore, to evaluate whether the measurement matrix Γ is appropriately designed, the mutual coherence metric is considered, given by:

$$\mu(\Gamma) = \frac{1}{M(M-1)} \sum_{m,m' \in [1,M], m \neq m'} \frac{|\gamma_m^H \cdot \gamma_{m'}|}{||\gamma_m||_2 \cdot ||\gamma_{m'}||_2} \tag{5.7}$$

where γ_m with $m \in [1, M]$ represent the columns of Γ. The mutual coherence is as an important criterion to verify how well the columns of the measurement matrix are uncorrelated, and lower values of mutual coherence indicate a better design for signal recovery in compressive sensing applications. Thus, in the problem of minimizing mutual coherence, it is useful to define the Gram matrix $\hat{U} = \hat{\Gamma}^H \cdot \hat{\Gamma}$, where $\hat{\Gamma}$ corresponds to the matrix Θ normalized by columns. As $\hat{U}$ is Hermitian, it can be expressed in terms of a single vector of values $\mathbf{u} \in \mathbb{C}^{\frac{M \cdot (M+1)}{2}}$, represented by $[u_{1,2}, ..., u_{1,M}, u_{2,3}, ..., u_{M-1,M}]$ with $m, m' \in [1, M]$ and $m \neq m'$, which corresponds to the upper triangular sub-matrix of $\hat{U}$. This sub-matrix contains the necessary values to compute the mutual coherence [67].

$$\hat{U} = \begin{bmatrix} u_{1,1} & u_{1,2} & ... & u_{1,M-1} & u_{1,M} \\ u_{1,2}^* & u_{2,2} & ... & u_{M-2,M-1} & u_{M-2,M} \\ ... & ... & ... & .... & ... \\ u_{1,M-1}^* & u_{2,M-1}^* & ... & u_{M-1,M-1} & u_{M-1,M} \\ u_{1,M}^* & u_{2,M}^* & ... & u_{M-1,M}^* & u_{M,M} \end{bmatrix} \tag{5.8}$$

As previously mentioned, the matrix Γ depends on various parameters of the experiment. One adjustable parameter is the configuration of the RIS elements, which can be chosen in a way that minimizes mutual coherence. Therefore, the minimization of the mutual coherence can be defined as a optimization problem for the matrix T, maintaining A fixed:

$$(\text{P1}) \ \min_{\boldsymbol{T}} \quad \mu(\boldsymbol{\Gamma}), \tag{5.9}$$

$$with \quad \boldsymbol{\gamma}_m = \boldsymbol{t}_k \boldsymbol{\alpha}_m, \quad \forall m \in [1, M], \tag{5.10}$$

$$1^T \tilde{\boldsymbol{t}}_{k,l} = \delta, \quad \forall k \in [1, K], \ l \in [1, L], \tag{5.11}$$

$$\tilde{} \quad t_{k,l,i} \geq 0, \quad \forall k \in [1, K], \ l \in [1, L], \ i \in [1, N_a]. \tag{5.12}$$

To solve (P1), the authors of [41] present FCAO algortihm, where they suggest to solve (P2) in each iteration of the algortihm. For each iteration, it is necessary to minimize two quantities: the mutual coherence u and the times t_k.

$$(\text{P2}) \min_{\boldsymbol{t}_k, \boldsymbol{u}} \quad \|\boldsymbol{u}\|_1, \tag{5.13}$$

$$s.t. \quad u_{m,m'} = \frac{\boldsymbol{\gamma}_m^T \boldsymbol{\gamma}_{m'}}{\|\boldsymbol{\gamma}_m\|_2 \cdot \|\boldsymbol{\gamma}_{m'}\|_2}, \tag{5.14}$$

$$m,\ m' \in [1, M],\ m \neq m', \tag{5.15}$$

$$\mathbf{1}^T \tilde{\boldsymbol{t}}_{k,l} = \delta,\ \forall l \in [1, L], \tag{5.16}$$

$$\tilde{\boldsymbol{t}}_{k,l} \succeq 0,\ \forall l \in [1, L], \tag{5.17}$$

To solve (P2), the *augmented Lagrangian method* is employed to transform the constrained optimization problem into an unconstrained one. This is achieved using the indicator function $I(t_k)$ that incorporates the constraints, resulting in the following formulation:

$$\mathcal{L}(t_k, u, \beta, \alpha, T_{-k}) = ||u||_\infty + I(t_k) + \sum_{1 \leq m \neq m' \leq M} \beta_{m,m'} \left(u_{m,m'} - \frac{|f_m(p)^H \cdot f_{m'}(p)|}{||f_m||_2 \cdot ||f_{m'}||_2} \right) + \sum_{1 \leq m \neq m' \leq M} \frac{\rho}{2} \left| u_{m,m'} - \frac{|f_m(p)^H \cdot f_{m'}(p)|}{||f_m||_2 \cdot ||f_{m'}||_2} \right|^2 \tag{5.18}$$

$$I(t_k) = \begin{cases} 0\ , & \text{if } \mathbf{1}^T \cdot \mathbf{t_{k,l}} = \delta \text{ and } \mathbf{t_{k,l}} \succeq 0, \\ \infty\ , & \text{otherwise.} \end{cases} \tag{5.19}$$

where ρ is a positive scalar factor and β is a vector of Lagrange multipliers given by $\beta_0 = [\beta_{1,2}, ..., \beta_{1,M}, \beta_{2,3}, ..., \beta_{M-1,M}] \in \mathbb{C}^{M(M-1)/2}$ with random values between 0 and 1. This augmented Lagrangian function $\mathcal{L}(t_k, u, \beta, \alpha, T_{-k})$ can be rewritten in a simplified form as:

$$\mathcal{L}_{\mathcal{A}}(t_k, u, \beta; \rho) = \|u\|_{\ell_\infty} + I(t_k) + \sum_{1 \leq m \neq m' \leq M} \frac{\rho}{2} \left| u_{m,m'} - \frac{f_m(p)^H f_{m'}(p)}{\|f_m(p)\|_{\ell_2} \|f_{m'}(p)\|_{\ell_2}} + \frac{\beta_{m,m'}}{\rho} \right|^2. \tag{5.20}$$

The augmented Lagrangian method finds a local optimal solution for (P2) by solving a sequence of unconstrained subproblems according to Equation Eq. (5.20), where β is kept fixed. These subproblems can be solved using an alternating minimization procedure, where the value of **u** is updated while t_k is kept fixed, and vice versa.

Methodology and Implementation

This section details the FCAO Algorithm for problem (P1) and the Augmented Lagrangian Method for problem (P2).

Algorithm 1 FCAO Algorithm for (P1)

Obtain the initial value $\mu^{(0)}$ using equation (5.7) based on the matrix $C^{(0)}$
Set the number of iterations without changes $N_{non} = 0$ and the frame index $k = 0$ with $k \in [0, 9]$
for $i = 1, 2, 3, \ldots$ **do**
    Obtain the initial value t'_k using the pattern search method to find a good starting point for the augmented Lagrangian method
    Given t'_k, solve problem (P2) using the Augmented Lagrangian method (see Algorithm 2), obtaining a new minimum mutual coherence μ' and optimized time configuration t''_k
    If $\mu' < \mu^{(i)}$ then update $\mu^{(i)} = \mu'$, the matrix $T^{(i)}$ with ${t_k}^{(i)} = t''_k$, and reset $N_{non} = 0$; else $N_{non} = N_{non} + 1$
    If $N_{non} < K$, then $k = \text{mod}(k + 1, K)$; else **return** μ^* and T^*
end for

The FCAO Algorithm, detailed in Algorithm 1, iteratively optimizes the mutual coherence and time configurations. It begins with identifying an effective initial point using the pattern search method, implemented with the PyMOO library [68]. Details of the pattern seach algorithm are provided in Appendix C.1.

The problem (P2) is then solved using the Augmented Lagrangian Method. Optimization results are updated as better solutions are found, with the algorithm alternating between different frames until convergence is achieved.

Algorithm 2 Augmented Lagrangian Method for (P2)

Obtain the initial value $t_k^{(0)}$ using *pattern search* and set $\rho^{(1)}$
Calculate the values of $u_{m,m'}^{(0)}$ based on the values of $t_k^{(0)}$ and set $\beta_{m,m'}^{(1)} = 0$
for $j = 1, 2, 3, \ldots, NAL$ **do**
    Set ${t_k}_{(0)}^{(j)} = {t_k}^{(j-1)}$, $u_{(0)}^{(j)} = u^{(j-1)}$, $\beta^{(j)}$, and $\rho^{(j)}$
    for $i = 1, 2, 3, \ldots, NAM$ **do** ▷ alternating minimization
        update u: $u_{(i+1)}^{(j)} = \text{argmin}_u \mathcal{L}({t_k}_i^{(j)}, u, \beta, \alpha, T_{-k})$
        update t_k: ${t_k}_{(i+1)}^{(j)} = \text{argmin}_{t_k} \mathcal{L}(t_k, u_{(i+1)}^{(j)}, \beta, \alpha, T_{-k})$
    end for
    Update $\beta_{m,m'}^{(j+1)} = \beta_{m,m'}^{(j)} + \rho^{(j)} \cdot \left(u_{m,m'} - \frac{|f_m(p^{(j)})^H \cdot f_{m'}(p^{(j)})|}{\|f_m(p^{(j)})\|_2 \cdot \|f_{m'}(p^{(j)})\|_2} \right)$
    Update the scalar $\rho^{(j+1)}$
end for

Algorithm 2 presents the steps to solve the problem (P2) through alternating minimization. In each iteration j, the process involves minimizing two quantities: the mutual coherence

and the time configurations. The implementation of the two minimization strategies is based on the approach outlined in [67], which provides specific algorithms and procedures.

u minimization (Step 6 in Algorithm 2): consists of consists of solving the optimization of the vector **u** by introducing an auxiliary variable **z**, where $z_{m,m'} = \frac{\gamma_m^H \cdot \gamma_{m'}}{||\gamma_m||2 \cdot ||\gamma m'||2} - \frac{\beta m,m'}{\rho}$. This allows the use of the Proximal Gradient Method to minimize the problem P_u:

$$(P_u) \quad u^* = \min_u ||u||_1 + \frac{\rho}{2}||u - z||_2^2 \tag{5.21}$$

For this, the *scipy.optimize* package was used, with the *'L-BFGS-B'* method chosen for its suitability in minimizing this type of objective function *.

t_k **minimization** (Step 7 in Algorithm 2): consists of solving the optimization problem $P(t_k)$ formulated as follows, where $k_{m,m'} = u_{m,m'} + \frac{\beta_{m,m'}}{\rho}$:

$$(P_{t_k}) \quad t_k^* = \min_{t_k} I(t_k) + \sum_{1 \leq m \neq m' \leq M} \frac{\rho}{2} \left| \frac{f_m(t_k)^H \cdot f_{m'}(t_k)}{||f_m||_{l2} \cdot ||f_{m'}||_{l2}} - k_{m,m'} \right|^2 \tag{5.22}$$

Given the complexity and non-convex nature of the objective function, an approximate solution is obtained using the Proximal Gradient Descent method, a variant of Gradient Descent. This approach decomposes the objective function into a smooth, differentiable part and a non-differentiable indicator function $I(t_k)$. The update rule is:

$$x^+ = \underset{z \in I_{t_k}}{\text{argmin}} \frac{1}{2\lambda}||z - (x - \lambda \nabla g(x))||_2^2 \tag{5.23}$$

where λ is the step size, adjusted via inexact line search to ensure the objective function decreases with each iteration. The derivation of the gradient expression $\nabla g(x)$ used in the optimization process is provided in Appendix C.2.

The update of the ρ value in both (P_u) and (P_{t_k}) follows an adaptive approach, adjusting based on the feasibility of the solution [69]. After each iteration, ρ is updated to $\rho_{old} \cdot \omega$ if the absolute difference $|c(x)|$ between the new and old u vectors does not decrease significantly, and to ρ_{old}/ω otherwise.

Finally, the implementation of Algorithms 1 and 2 was done in Python, with relevant code snippets provided in Appendix C.3.

*L-BFGS-B Method Reference

Results

During the implementation of the code, several individual tests were conducted to validate its performance and determine the appropriate values for some variables. For instance, to determine the initial value t_{k_0} using pattern search, the optimization problem from eq. (C.2) was implemented in Python, and the algorithm's output was analyzed. As shown in Fig. 5.4, the algorithm first adjusts the solution to satisfy constraints (until "cv_avg" = 0), then minimizes the objective function while maintaining these constraints until stabilization.

n_gen	n_eval	cv_min	cv_avg	f_avg	f_min
1	20	3.8531908991	4.8595181387	-	-
2	148	1.8956274387	2.8744091689	-	-
3	261	0.2826198025	1.0891236206	-	-
...					
24	2910	0.0002174463	0.0002174463	-	-
25	3032	0.000000E+00	0.0001087231	1.1037146874	1.1037146874
26	3155	0.000000E+00	0.000000E+00	1.1036871604	1.1036596333
27	3277	0.000000E+00	0.000000E+00	1.1035944276	1.1035292219
...					
98	10879	0.000000E+00	0.000000E+00	1.0458581638	1.0457935879
99	10992	0.000000E+00	0.000000E+00	1.0457233380	1.0456530880
100	11072	0.000000E+00	0.000000E+00	1.0455652571	1.0454774262

Figure 5.4: Pattern search algorithm output.

A termination criterion was established based on the behavior of the objective function concerning the number of generations n_{gen} for various initial times t_k (Fig. 5.5). It was observed that from approximately $n_{\text{gen}} = 330$, the objective function stabilized across different frames, leading to this value being set as the stopping criterion.

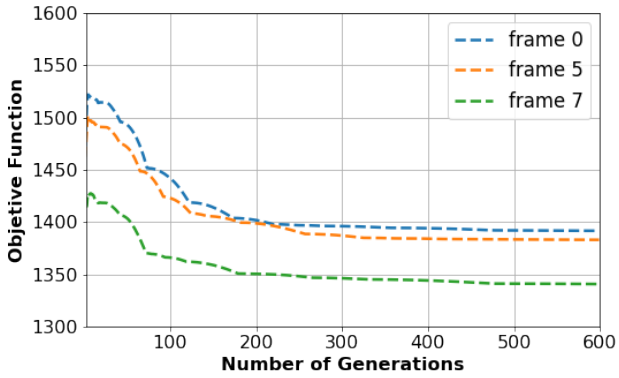


Figure 5.5: Convergence of Pattern Search Algorithm: Objective function vs n_{gen}

To define a convergence criterion for the FCAO algorithm, it was established that the algorithm would terminate if it could not find an initial value t_{k_0} using the pattern search method, as illustrated in Listing C.1. The parameter ρ was adaptively updated with $\omega = 1.3$, which was chosen to balance constraints and the objective function.

In the augmented Lagrangian method, choosing appropriate values for NAM and NAL presented in Algorithm 2 was necessary. After studying the algorithm's behavior for different iteration numbers NAM and NAL (Figure 5.6), NAM = 5 and NAL = 10 were selected. These values ensured that the norm of u converged to an optimal common value during the optimization of both t_k and u.

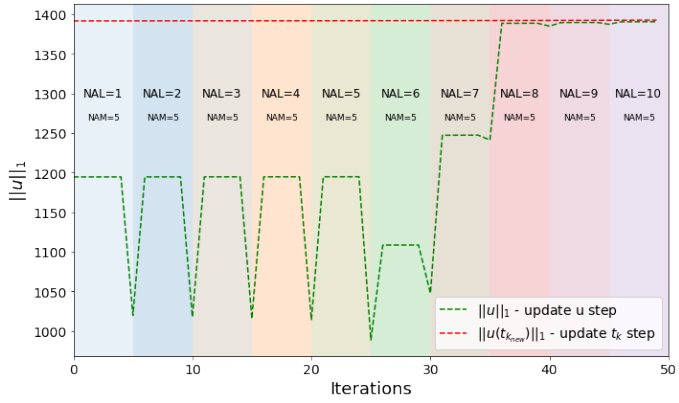


Figure 5.6: Convergence of the alternating minimization: u norm vs iterations for a certain frame k. The distinct colors in the graph represent different iterations of the alternating minimization process, with each color corresponding to an iteration NAL (line 3 in Algorithm 2) with NAM=5 iterations each, totaling 50 iterations (10×5).

It was also necessary to determine the value of the parameter λ presented in eq. (5.23) and the tolerance for the projected gradient descent in the augmented Lagrangian method. Figure 5.7 demonstrates the convergence behavior of the FCAO algorithm, with the average mutual coherence μ, defined in eq. (5.7), decreasing progressively across iterations. The values $\lambda = 1 \times 10^{-5}$ and tolerance 8×10^{-4} were selected after a trial-and-error process to identify the most effective parameters. With these settings, the average mutual coherence started at 0.7808, corresponding to a matrix Γ dependent on a random T, and decreased to 0.3817 after finding the optimized T that reduced the average mutual coherence of Γ's columns.

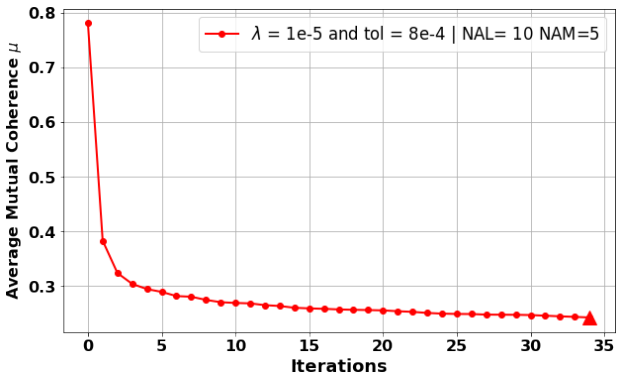


Figure 5.7: Convergence of FCAO Algorithm: μ vs iterations

Further insight into the optimization impact is provided by Figure 5.8, which compares mutual coherence matrices before and after optimization. These mutual coherence matrices correspond to the absolute value of each complex value u of the matrix $\hat{U}$ defined by eq. (5.8), providing insight into the relationship between different columns of Γ. In Figure 5.8a, the mutual coherence matrix for a random time configuration is presented, whereas Figure 5.8b shows the same matrix after the optimization process. It is observed that the average mutual coherence value decreases from 0.7808 to 0.3817, indicating that the FCAO algorithm effectively reduces the mutual coherence by 51.1%.

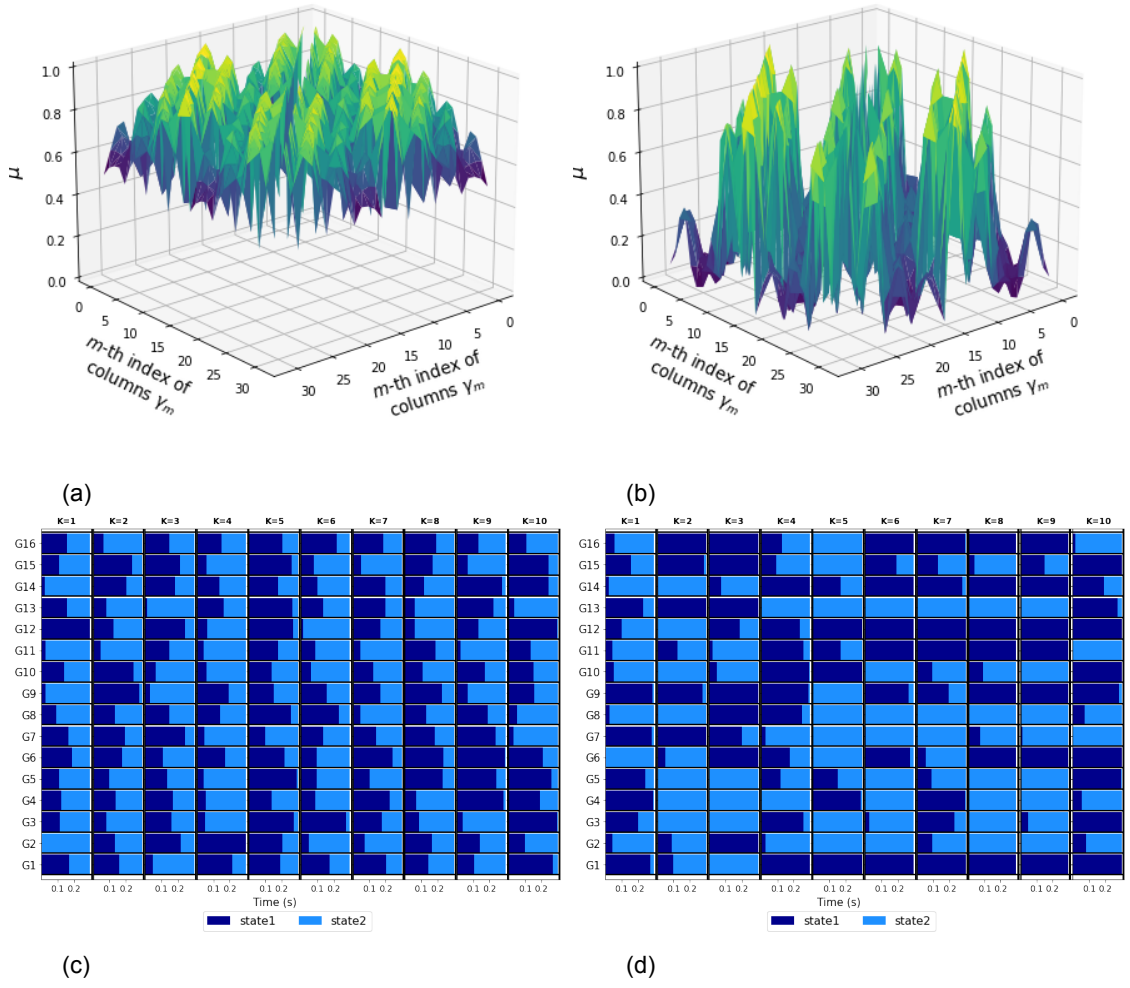
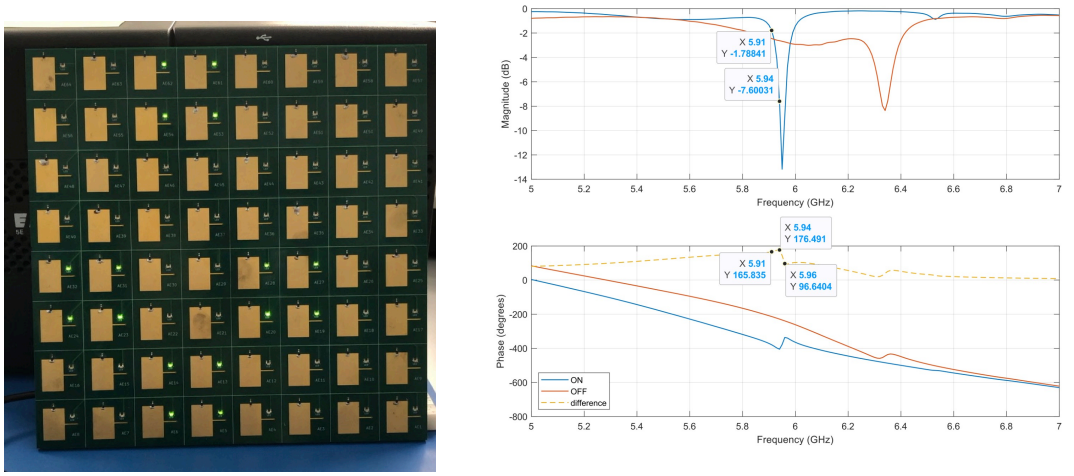


Figure 5.8: a) Mutual coherence matrix corresponding to the random T time configuration shown in c); b) Mutual coherence matrix corresponding to the optimized configuration and d) optimized T time configuration

To conclude, our FCAO implementation, adjusted for our experimental set up, successfully computes a sequence of 10 configurations frames and respective RIS element states in order to minimize the mutual coherence of the channel.

5.2. RIS Validation

The RIS used in this study consists of 64 unit cells, organized in an 8x8 grid. Each unit cell has dimensions of $23 \times 23\,\mathrm{m}^2$, resulting in a total size of $18.46 \times 18.46\,\mathrm{cm}^2$. Each unit cell incorporates one PIN diode, making it a 1-bit RIS capable of assuming two states. Each unit cell also has an LED that visually indicates whether it is in the ON or OFF state. The RIS was simulated and designed by members of our workgroup. An image of the RIS is shown in Fig. 5.9a, where each unit cell has an LED that visually indicates whether it is in the ON or OFF state.



(a) RIS with 64 unit cells arranged in an 8x8 grid.

(b) Magnitude (top) and phase (bottom) of the RIS vs. frequency for states ON and OFF.

Figure 5.9: RIS details.

The RIS was designed with a center frequency of 6.5 GHz, corresponding to the central frequency of the WiFi-6E band. A phase difference of 180 degrees between the ON and OFF states was expected at this frequency. However, initial tests conducted by our workgroup revealed that the RIS exhibited a phase difference of 180 degrees around 5.94 GHz. By combining this requirement with the region where the magnitude does not decay more than -3 dB, it was concluded that the RIS performs best at a frequency of 5.91 GHz, as shown in Fig. 5.9b, where the phase difference between the two configurations is 165º, lower than the desired 180º, but where the magnitude response for both states approximately equal. That is, although a greater phase difference is observed for 5.94GHz, the magnitude response of the RIS for the ON state greatly attenuates the reflected signal.

5.2.1 Initial Tests

To verify the functionality of the RIS, preliminary tests were conducted to ensure that the device effectively manipulates electromagnetic signals as intended. The primary objective was to confirm that the RIS could perform steering by altering the states of its unit cells (ON or OFF).

To achieve the desired steering effect, the RIS must be configured to direct the signal in specific directions by adjusting the phase of each unit cell. This involves changing the state of each unit cell, either ON or OFF, according to the following formulas:

$$\phi_{ij} = k\left(R - \sin(\theta)\left(x_i\cos(\phi) + y_j\sin(\phi)\right)\right) \tag{5.24}$$

where θ and phi define the desired beam direction, R is the euclidean distance between each unit cell and the Tx, k is the wave vector, and (x_i, y_i) are the coordinates of the center of each unit cell of the RIS. These formulas were implemented in a script developed by a workgroup member, inspired by previous work *. The phases ϕ_{ij}, which are continuous in the range [0º, 360º], are then quantified to two ranges, corresponding to an OFF state, or ON state, which are the phase setting possible with our RIS, where continuous phase configuration per unit cell is not supported.

My role involved providing the position of the transmitting antenna (Tx), as calculated in Section 5.1.1. The script then used this input to compute the RIS configurations necessary for achieving the desired steering patterns. The RIS was configured to steer the signal towards six different angles: 0°, 20°, 30°, 40°, 50°, and 60°. Figure 5.10 shows the corresponding RIS configurations for each steering angle.

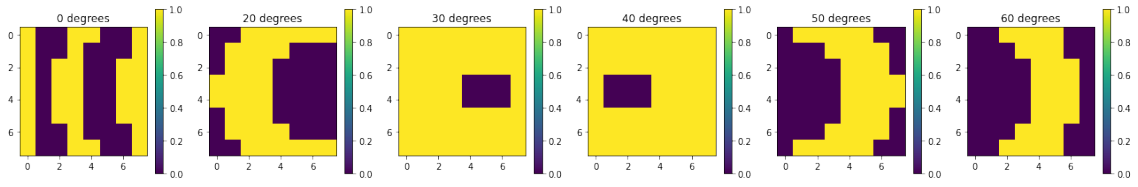
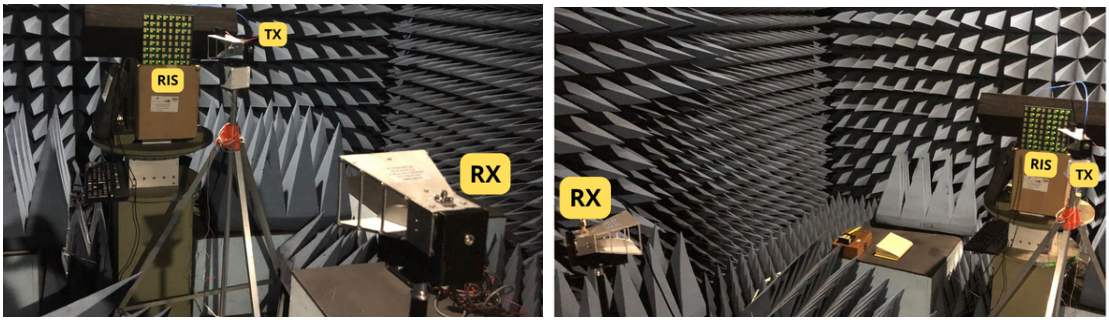


Figure 5.10: RIS configurations for steering the signal at different angles.

In the experimental setup, the transmitter remained fixed in its designated position, while the receiver was placed at various angles relative to the RIS—specifically, at 0°, 20°, 30°, 40°, 50°, and 60°. For each Rx position, S_{21} parameters were measured as a function of frequency, with six measurements taken at each position. This setup is illustrated in the photos in Figures 5.11a and 5.11b, showing the Rx positioned at 0° and 30°, respectively.

*Alexandros Pitilakis, "Calc-ScatPat". GitHub.`https://github.com/alexpiti/Calc-ScatPat`



(a) Rx positioned at 0° (b) Rx positioned at 30°

Figure 5.11: Experimental setup for verifying RIS steering capability.

Given that the central frequency for the RIS was chosen to be 5.91 GHz, the analysis focused on the signal magnitude at this frequency. For each angle of the Rx, it was expected that the magnitude of the signal would peak when the steering direction matched the angle of the Rx. To validate this, the normalized magnitude of the signal was measured across different Rx positions and RIS steering configurations.

The resulting data is presented as a matrix (Figure 5.12), where each column represents the results for a specific Rx angle, and each row corresponds to a RIS steering angle. The expectation was that the maximum magnitude would be observed when the steering angle aligns with the Rx angle. That is, when the RIS steers the reflection of the signal received from the Tx at an angle, relative to the normal of its plane, that directs it towards the Rx.

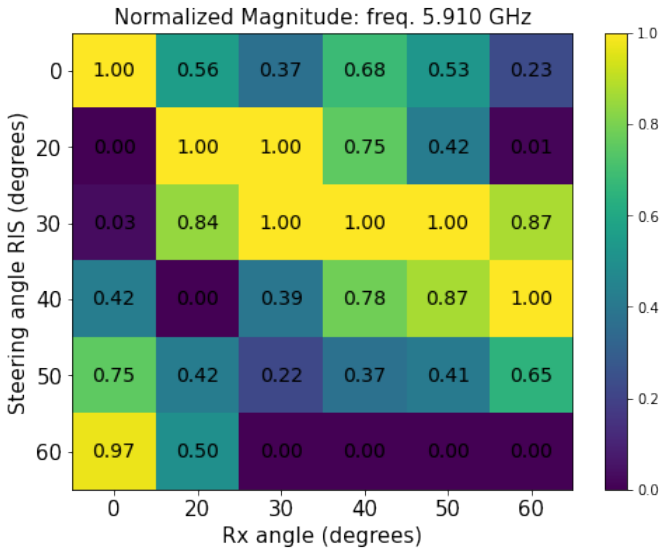


Figure 5.12: Normalized signal magnitude at 5.91 GHz for various Rx positions and RIS steering angles. The normalization is applied solely to help detect each peak, and we are not concerned with the radiation pattern for this functional validation.

For lower Rx angles (0° and 20°), the expected peak in signal magnitude was observed at the corresponding Rx angle. However, there was an unexpected increase in magnitude at larger steering angles, suggesting a potential capture of side lobes from the RIS. For mid-range Rx angles (30°), the magnitude peaked as expected around the steering angle and decreased beyond this point, aligning with the design expectations. For higher Rx angles (40°, 50°, and 60°), the peak signal magnitude was observed at angles different from the Rx angle, indicating possible side lobe effects or misalignment issues with the Rx antenna.

Overall, while side lobe effects and some setup inaccuracies influenced the results, the data confirms that the RIS is indeed capable of steering the signal as intended. This validates the capability of the RIS in controlling the direction of reflected signals.

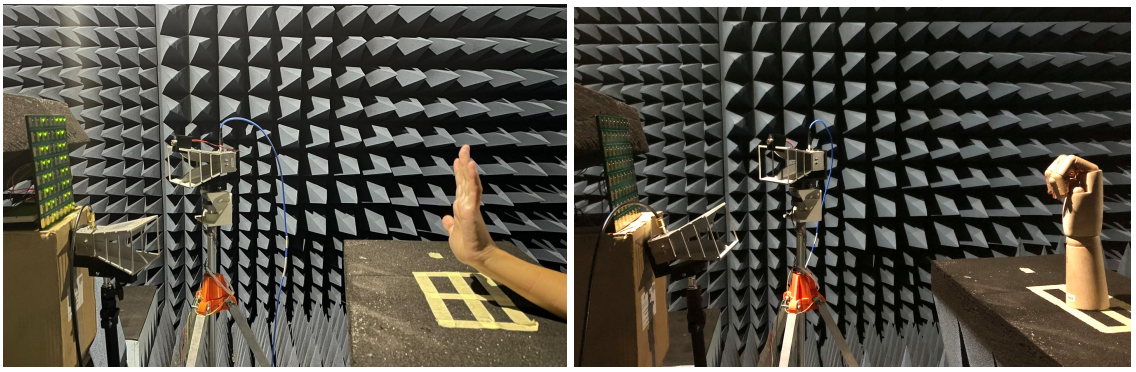
5.3. Experimental Setup and Methodology

This section describes the experimental setup and methodology for hand gesture recognition using the RIS. The methodology is divided into two main parts due to the distinct objectives of the experiments conducted (Table 5.1).

The experimental setup was conducted in an anechoic chamber at FEUP to minimize external signal interference. A-INFO LB-20180 horn antennas were used for both the transmitter and receiver. The Tx was positioned at a 35-degree angle relative to the normal of the RIS, while the SoI was shifted by 15 degrees to prevent interference from the Tx antenna. The Rx antenna was placed directly below the RIS, oriented towards the center of the SoI to capture reflections from hand gestures only. Additionally, a ZX60-153LN-S+ low-noise amplifier (LNA) was incorporated into the system to amplify the signal emitted by the Tx antenna.

Figure 5.13 illustrates the experimental setups used for the two different experiments. Figure 5.13a shows the setup with a real hand, while Figure 5.13b depicts the setup with a wooden hand. In the setup with the wooden hand, the Rx antenna is slightly tilted upwards to focus on the center of the wooden hand, thereby excluding the beginning of the arm from the measurements. This adjustment ensures that both setups, with the real hand and the wooden hand, exclude the initial part of the arm from the data, providing a consistent evaluation of hand gestures.

The objective of the initial experiment using a real hand was to validate the experimental



(a) (b)

Figure 5.13: Experimental setup for HGR with (a) real hand (b) wooden hand, inside the anechoic chamber.

setup and methodology for HGR. For this purpose, data was collected for three different hand gestures across distinct RIS configurations that directed the signal in specific directions: $5°, 10°, 15°, 20°$ and all elements OFF. For each of the three gestures, data acquisition involved manually saving the transmission coefficients S_{21} with the VNA and changing the RIS configurations using a computer. The computer communicates with the RIS via WiFi.

In the second experiment, the focus shifted toward evaluating the efficiency of the FCAO algorithm and training a Neural Network. The primary goal was to gather a significant amount of data to assess the performance of different RIS configurations, both randomized and optimized. To achieve this, each experimental run was divided into 10 frames, with each frame consisting of 39 distinct RIS configurations. These configurations were applied sequentially, resulting in a total of 390 measurements of the transmission coefficients S_{21} per run. The experiment was conducted twice: once with random configurations and once with the optimized configurations determined by the FCAO algorithm.

Inspired by the automatic data acquisition system described in Chapter 4, adjustments were made to adapt the system for the hand gesture recognition task. The new script synchronized the configuration changes in the RIS with the data collection process. For each configuration, the script triggered the VNA to record the corresponding S_{21} values, proceeding through all 39 configurations for each frame. This automation ensured consistent data capture across the 10 frames, resulting in 390 total measurements per run.

During each run, the wooden hand was positioned in three different postures within the SoI, with the RIS configurations changing systematically for all 10 frames. This process

was repeated for both the random and optimized configurations, ensuring a comprehensive dataset for evaluating the classification performance of the system and the effectiveness of the FCAO algorithm. During this experiment, it was observed that slightly varying the hand's orientation (e.g., tilting it left, right, forward, or backward) while maintaining the same gesture resulted in noticeable signal changes. To account for this, additional measurements were taken for each gesture, with the wooden hand placed in 8 different orientations. This was done for the random and optimized configurations, leading to a more diverse dataset.

Experiment	**Main Objective**	**Hand Used**	**RIS Configurations**
1º	Analyse feasibility of hand gesture recognition using S_{21} measurements.	Real Hand	Directional steering configurations: 5°, 10°, 15°, 20°
2º	Evaluate CNN accuracy for hand gesture recognition using S_{21} data under different RIS configurations.	Wooden Hand	Random vs Optimized configurations determined by the FCAO algorithm

Table 5.1: Summary of Experimental Objectives and Setups.

5.4. Results

5.4.1 Feasibility Study with Real Hand

The collected dataset for the three distinct gestures (Open Hand, Two Fingers, and Close Hand) included only one hand of a single volunteer. For each gesture, the hand remained static while the five configurations of the RIS changed. This process was repeated 60 times, resulting in 900 samples (1 hand's subject × 5 RIS configurations × 60 *runs*). Additionally, noise measurements were taken without any hand in the SoI. Ten measurements were taken for each of the five RIS configurations, resulting in 50 noise measurements.

Each sample is characterized by the number of frequency points between 5.0 GHz and 6.5 GHz, $n_{resolution}$, that was set to 201. Each sample was saved with a file name format "{date}_{angle}_{gesture}_{n_{med}}" , where n_{med} was the number of the measurement run.

During data capture, the signal varied slightly, which is expected given the difficulty in keeping the hand perfectly still. Considering that each *run* included changing the configuration and saving the data, each *run* took approximately 15 seconds, resulting in a total acquisition time of about 1 hour and 15 minutes (15 seconds × 5 RIS configurations × 60 *runs*), not including breaks to rest the arm. This data collection process was quite

exhausting, making it impractical to use a real hand for collecting a large dataset required for gesture classification with the FCAO configurations.

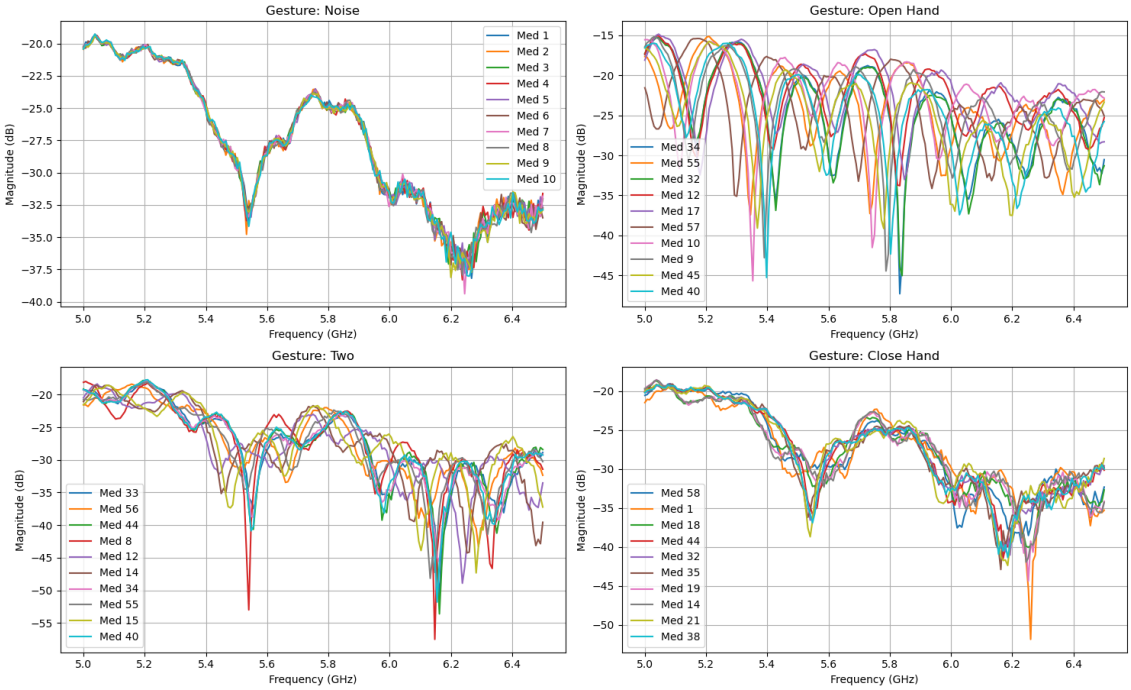


Figure 5.14: Magnitude of the S_{21} parameters for each gesture at a RIS configuration steering the signal by 20 degrees.

Thus, the obtained dataset was analyzed to observe if there were distinguishable differences between the hand gestures with this setup. Since the signal varied slightly due to the difficulty in maintaining a static hand position, averaging the signal would result in loss of information. Therefore, the raw signals were analyzed. First, we examined the magnitude of the S_{21} parameters for each gesture at a specific RIS configuration that steered the signal by 20 degrees. This approach allowed us to observe the variability between different *runs* for the same gesture. Figure 5.14 shows the magnitude for each of the gestures under this configuration, highlighting the variability present in the measurements. This variability is more pronounced for the 'Open Hand' and 'Two Fingers' gestures, while the 'Close Hand' gesture is closer to the noise level in terms of variability per measurement.

Figure 5.15 presents the frequency-RIS configuration images of the magnitude and phase of the S_{21} parameters. Each column shows 3 randomly selected runs (i.e., 3 pairs of magnitude and phase plots), for the case shown in the image below. That is, 3 examples from each of the 60 measurements per each gesture (5.15b), 5.15c), and 5.15d)), and 3 examples from the 10 measurements of noise (5.15a)). It is evident that each gesture has a distinct magnitude and phase image, leading to the conclusion that there are clear differences in the received signal at the Rx for each gesture.

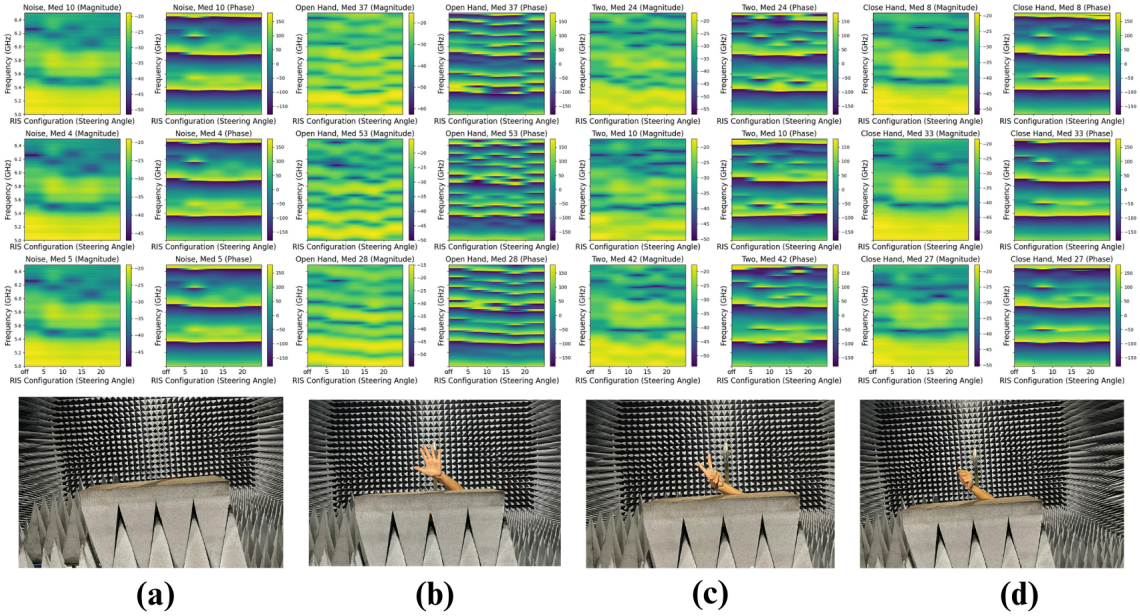


Figure 5.15: Frequency vs RIS configuration Analysis for S_{21} Magnitude, Phase and Gesture Identification for: (a) Noise, (b) Open Hand, (c) Two Fingers, and (d) Close Hand.

5.4.2 Gesture Classification with Wooden Hand

In this second experiment, a wooden hand was used for data collection. The collection process now applies the FCAO algorithm, and therefore, the concept of frames will be discussed again. Copper tape was applied to the fingers to enhance reflection, as wood is a material with low reflectivity [70]. The same three hand gestures from the first experiment—Open Hand, Two Fingers, and Close Hand—were studied. For each gesture, 115 runs were conducted using random RIS configurations and 115 runs with optimized configurations, both across 10 frames.

Each sample contained transmission coefficients S_{21} across the frequency range [5, 6.5] GHz, with 201 points per sample, as in the first experiment. The file naming format was "{frame}_{k}_{gesture}_{n_{med}" where k represented the frame number ($k \in [0, 9]$) and n_{med} the number of the measurement run ($n_{med} \in [0, 114]$), resulting in 10 files per measurement run. The data files were saved into two folders: "ris_optimized" and "ris_random", corresponding to the respective configuration strategies.

The time required to complete one run, which involved capturing data for all 390 configurations (39 configurations per frame $\times$ 10 frames), was about 3.5 minutes, for both the random and optimized configurations. The data acquisition time for the random configurations was 3.5 minutes$\times$ 115 runs = 6.7 hours per gesture, with the same duration for

the optimized configurations. Considering the three different gestures, this resulted in a total acquisition time of around 39 hours. This extended acquisition time demonstrated the impracticality of using a real hand for large datasets, as it demands prolonged data collection, which is physically exhausting for participants.

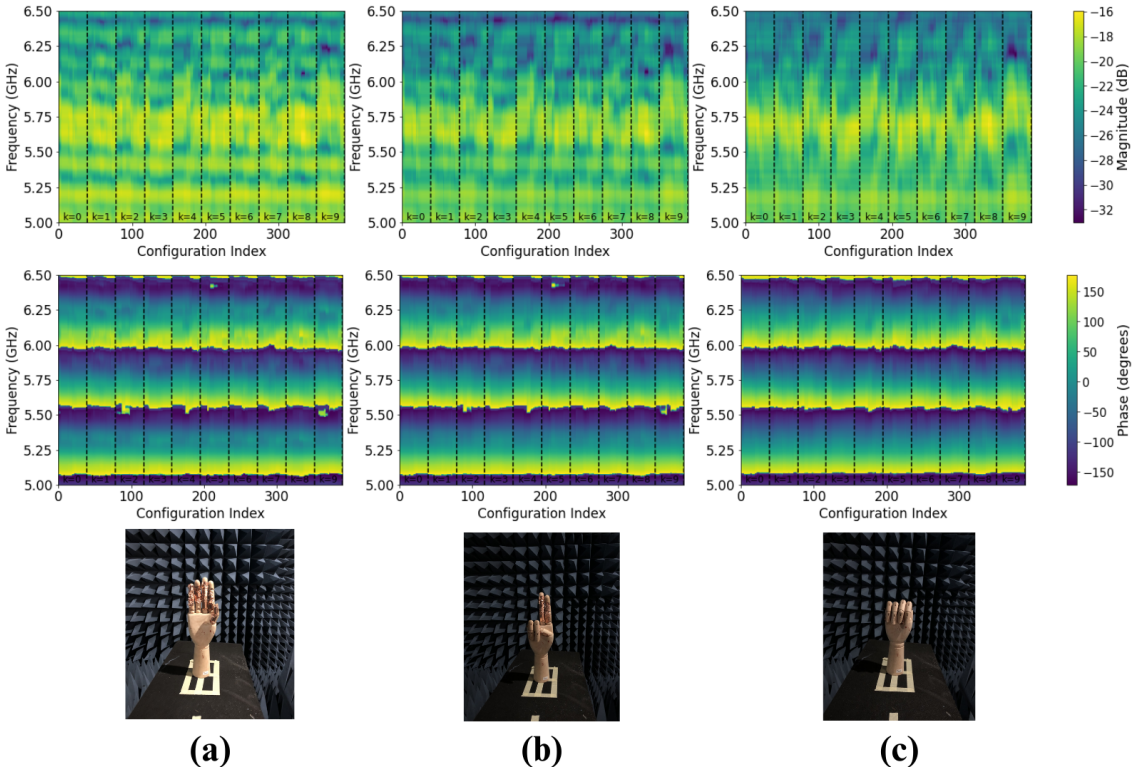


Figure 5.16: Frequency vs RIS Random configuration Analysis for S_{21} Average Magnitude, Average Phase and Gesture Identification for (a) Open Hand, (b) Two Fingers, and (c) Close Hand.

The collected data was again analyzed in the form of images, mapping frequency as a function of the RIS configurations over all frames. Fig. 5.16 and 5.17 illustrate the averaged phase and magnitude of the S_{21} parameters for each gesture, for the random and optimized RIS configurations respectively.

Fig. 5.16 shows little signal variability in magnitude across frames, except for frame 9, which is slightly different. The phase also in very equal between frames. This suggests that having multiple frames with random configurations may not add much value, as the information across them is redundant. However, there is variability between gestures, particularly in the 5.2–5.6 GHz range, where ”Close Hand” shows a weaker pattern and ”Open Hand” displays a more distinct feature.

Fig. 5.17 reveals noticeable variability in magnitude and phase across frames, indicating that more information is captured from the space of interest when using the configuration

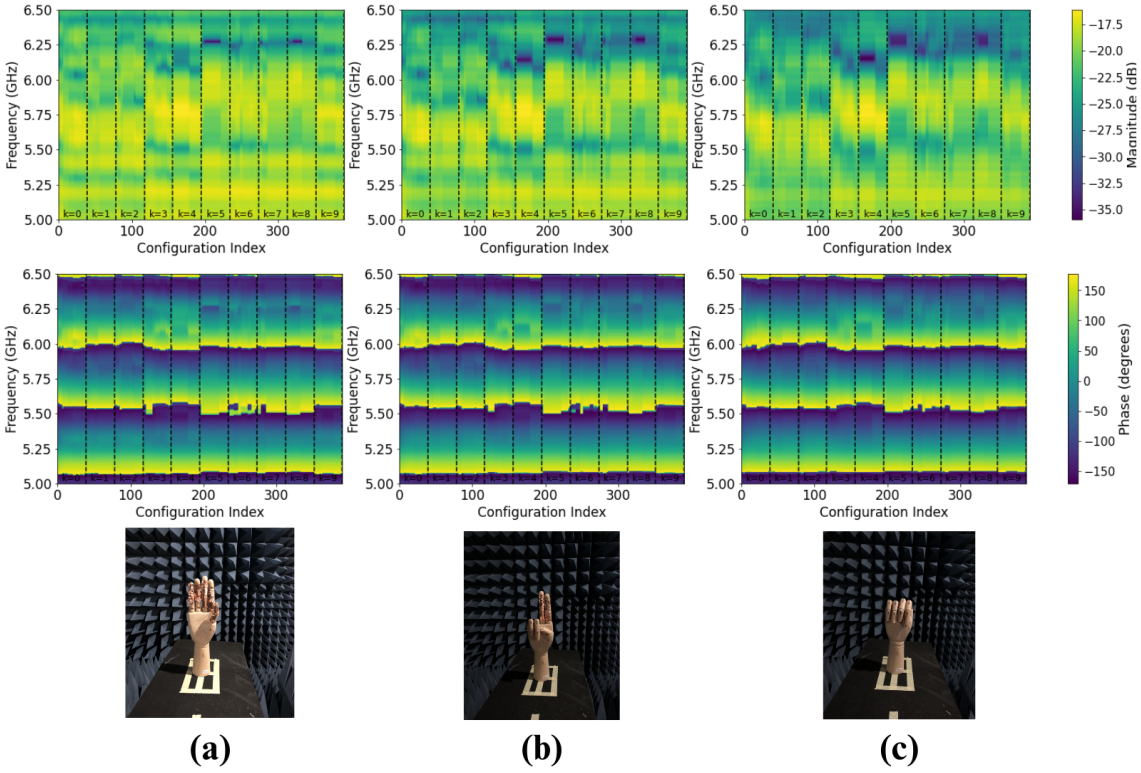
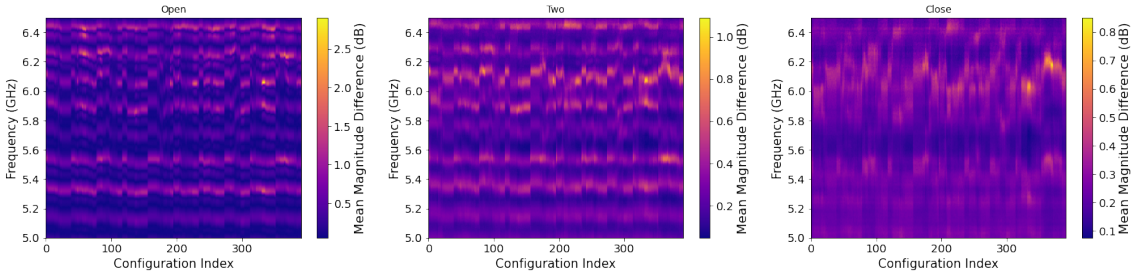


Figure 5.17: Frequency vs RIS Optimized configuration Analysis for S_{21} Average Magnitude, Average Phase and Gesture Identification for (a) Open Hand, (b) Two Fingers, and (c) Close Hand.

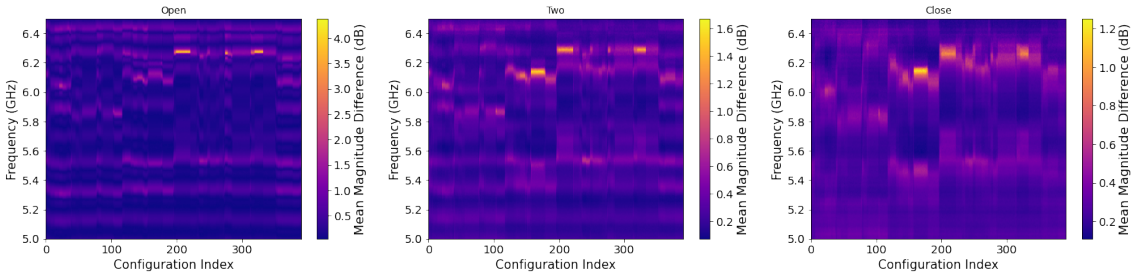
pattern computed by our implementation of the FCAO algorithm. Additionally, differences between gestures are more pronounced. For example, between 6 and 6.5 GHz, significant magnitude valleys are observed, with these dips being less pronounced for ”Open Hand” and more prominent for ”Close Hand”.

To justify presenting averaged magnitudes and phases, Fig. 5.18 displays the mean of the absolute differences in magnitude between the averaged signal and each of the 115 runs for each gesture, both for the random and optimized configurations. The first row shows the differences for the random configurations, and the second row for the optimized configurations. The differences range between 0 and 4 dB across gestures. The overall magnitude range spans approximately -40 dB to -14 dB. Despite some variability, particularly in the ”Open Hand” gesture, these differences are relatively minor, indicating that additional runs do not significantly impact the results.

In other words, there is little variability in data between different runs for the same gesture, likely due to the anechoic environment and use of an immobile hand. Therefore, data augmentation techniques were employed to ensure a more robust dataset for training a neural network for hand gesture recognition. First, additional measurements were



(a) Random Configurations



(b) Optimized Configurations

Figure 5.18: Mean magnitude differences of S_{21} across three gestures for random and optimized configurations.

taken with the wooden hand placed in 8 different orientations for each gesture. This approach aimed to capture slight variations in the signal caused by changes in the hand's positioning, providing a richer dataset for classification. Given the earlier analysis of mean differences in Fig. 5.18, it was determined that repeating measurements did not significantly contribute to new information. Therefore, only a single measurement was taken for each new orientation.

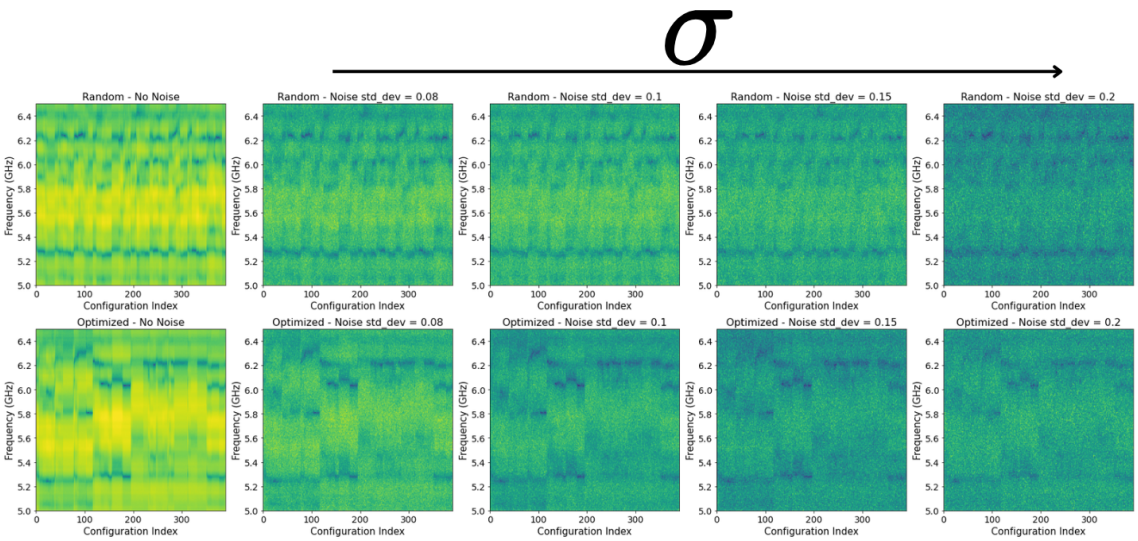


Figure 5.19: Data augmentation example: white noise added to one image corresponding to measurements using random RIS configurations (top row), and another employing FCAO derived configurations (bottom row).

Data augmentation techniques were applied to expand the dataset by replicating each

measurement 115 times and adding white noise with a mean of 0 and a standard deviation randomly chosen from [0.08, 0.1, 0.15, 0.2] as described in [71]. This approach ensured a more robust dataset, allowing the neural network to learn patterns beyond the limited original data and avoid overfitting. This process resulted in a total of 1035 samples per gesture for both random and optimized configurations (115 runs $\times$ 9 orientations). Fig. 5.19 provides an example of this process, where noise was added to create realistic variations while maintaining the core signal patterns. With three gestures, the final dataset comprised 3105 samples for random configurations and 3105 samples for optimized configurations, where each sample corresponds to an image.

For the classification stage, the first neural network approach was inspired by the one described by the authors that presented the FCAO algorithm [41]. In that method, only the information at the operating frequency of the RIS was used for classification. In our case, this frequency is f = 5.910 GHz. The data consisted of the averaged received signal over each of the 10 frames, resulting in a signal of dimension (1, 10) (i.e. f_c, K =10). The neural network consisted of three fully connected layers: an initial layer with 64 neurons, a hidden layer with 32 neurons, both using the ReLU activation function, followed by an output layer with softmax activation for gesture classification. The dataset was split into 80% for training, 10% for validation, and 10% for testing.

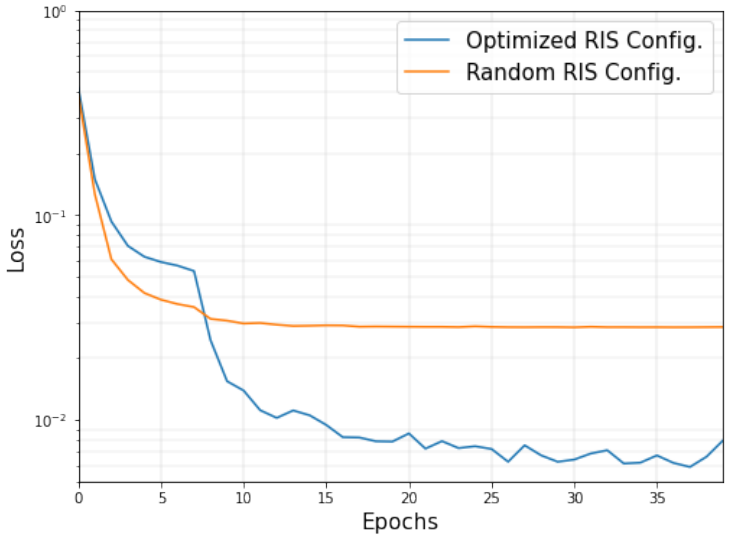


Figure 5.20: Convergence of Simple Model for Gesture Classification.

During training (as shown in Fig. 5.20), it was observed that in the initial epochs, the model performed better with the random configurations than with the optimized ones. This suggests that the optimized configuration contains more complex information, making it harder to learn at first. However, as training progressed, the optimized configuration

showed improved accuracy, surpassing the random configuration. This pattern was also observed in the study by Hu et al. [41] (see Fig. C.2).

Fig. 5.21 presents the confusion matrices for both configurations, demonstrating an advantage for the optimized setup. This confirms that the optimized configuration matrix, designed with low average mutual coherence, can achieve superior accuracy in gesture recognition. Specifically, the optimized configuration resulted in a 99.67% recognition accuracy, providing a 3.53% improvement compared to the random configuration, which achieved 96.14%.

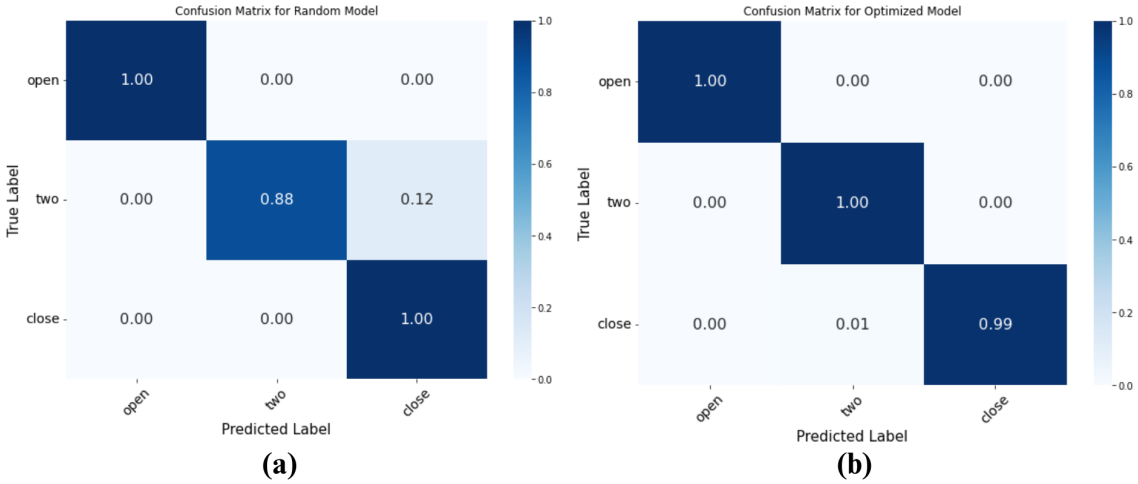


Figure 5.21: Gesture Recognition accuracy utilizing two distinct approaches based on the RIS configurations: (a) random (b) optimized,

The second classification approach employed a CNN with two branches, one for magnitude and one for phase, similar to the CNN described in Chapter 4. The architecture of this network is presented in Fig. 5.22, where the input layer consists of images of phase and magnitude with dimensions (201,390), and the output layer consists of three nodes for the hand gestures to be classified. This dual-branch architecture allows the network to learn features from both the magnitude and phase components of the RF signals, which can capture distinct and complementary information.

In this case, both the random and optimized configurations achieved high recognition accuracies, with the random configuration reaching 99.047% and the optimized configuration achieving 99.057% (Fig. 5.23). The improvement from the optimized configuration was minimal, only 0.01%, indicating that both methods were highly effective for gesture recognition.

The improvement in accuracy using an input size of (1, 10) in the first approach aligns with

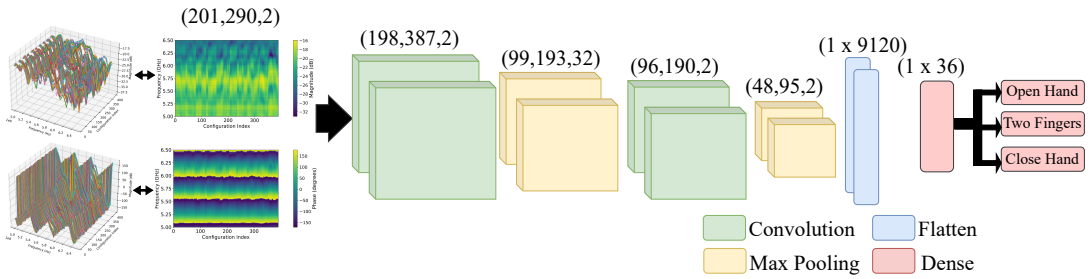


Figure 5.22: Architecture of the CNN developed for gesture classification (Open Hand, Two Fingers, Close Hand). The input to the network consists of the magnitude and phase of the S_{21} parameters, which can be represented both as a 3D plot and as a frequency-RIS configurations image.

the principles of the FCAO algorithm, which was inspired by compressive sensing theory. In this context, K (the number of measurements) is ideally much smaller than M, the number of distinct spatial regions of interest—in our case, M=32. In the first approach, where K=10, the network effectively captures gesture information as expected. However, in the second approach, by not averaging the signals across frames, we leverage all 390 distinct signals, which results in K=390, greatly exceeding M. This substantial increase in measurements makes the application of the FCAO algorithm less relevant, as the distinction between random and optimized configurations becomes negligible in terms of accuracy. Both configurations achieve similarly high performance, demonstrating that with such a large number of measurements, the choice of configuration has minimal impact on the classification results. In other words, while the use of an optimized configuration using FCAO increases the information captured from a narrowband signal, at the cost of the

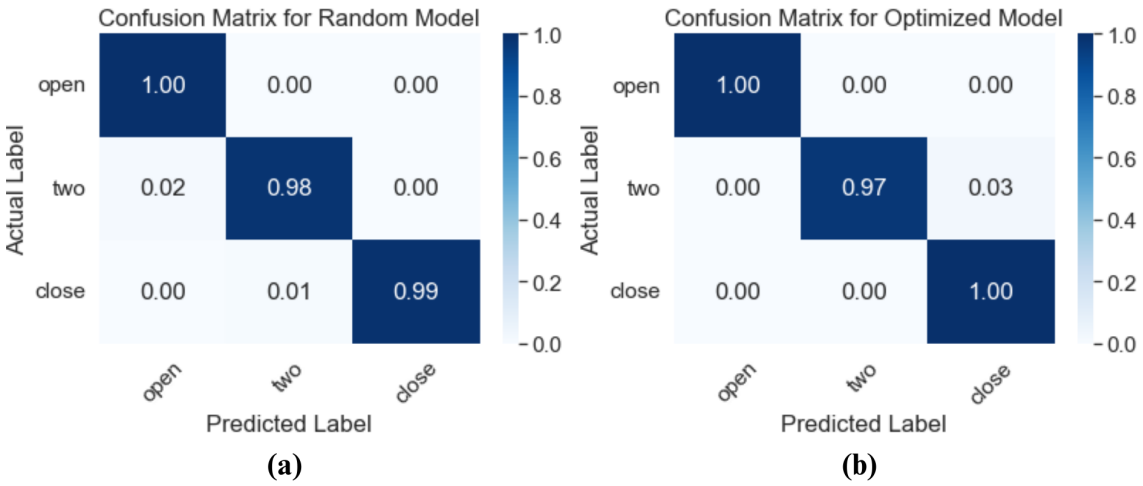


Figure 5.23: Gesture Recognition accuracy utilizing two distinct approaches based on the RIS configurations: (a) random (b) optimized,

computational effort of determining this configuration, we achieve a similar level of information by capturing magnitude and phase data for a broader frequency band, even using a random sequence of RIS configurations during data capture.

5.5. Summary

In this chapter, we successfully implement hand gesture recognition using a RIS with configurations optimized by the FCAO algorithm. The integration of the RIS, configured through the FCAO algorithm, resulted in a 3.53% improvement in classification accuracy compared to random configurations. The FCAO algorithm required significant coding efforts, detailed in Appendix C.3. Previous work by Hu et al. demonstrated a 14.6% accuracy improvement with a 2-bit RIS having four possible states. The greater accuracy in their study may be attributed to the increased number of possible configurations with a 2-bit RIS, which provides a richer set of options for optimization (see Fig. C.1) and can significantly affect performance. Additionally, their setup focused on human posture recognition with a larger RIS, which likely necessitates more precise control and optimization due to the larger size and distance between components. In contrast, our study with a 1-bit RIS for hand gesture recognition shows that even with fewer states and a smaller setup, optimized configurations can still improve performance. Future work could explore the benefits of using RIS with more states and larger dimensions in applications requiring extensive control and accuracy, such as human posture recognition.

This study confirms the potential of RIS-based systems for enhancing human-related classification tasks. By utilizing the RIS, we advanced the state of the art in 6G solutions for human activity recognition.

6. Conclusions and Future Work

In this work, we made significant progress in Human Activity Recognition using Reconfigurable Intelligent Surfaces for Radio Frequency sensing. By analyzing and comparing different studies in this area, we found that most existing work relies on highly controlled settings, like anechoic chambers or closed lab environments, which minimize external interference. Following this, our experiments also used an anechoic chamber to better understand the potential of RIS-based systems without the complexities of external noise, as we lacked any existing foundational implementation that allowed the exploration of real-world scenarios. Even so, some outdoor experiments were conducted (see Appendix B).

We explored two main approaches: human posture recognition using a metal plate, and hand gesture recognition with a RIS configured by the FCAO algorithm. The use of a metal plate simplified the hardware setup, making it a cost-effective solution for creating semi-controllable signal paths with WiFi-6E signals, achieving an accuracy of 99.17% in posture classification. Meanwhile, integrating a RIS optimized with the FCAO algorithm improved hand gesture recognition accuracy by 3.53% compared to using random configurations.

The objectives set out in the introduction have been successfully achieved. We advanced the state of the art in human activity recognition by implementing novel smart radio environments for RF sensing. This was accomplished by completing the intermediate goals defined in Section 1.3. Table 6.1 summarizes how each goal was achieved:

Intermediate Goal	**How It Was Achieved**
Identify and implement methods to enhance the information retrieved from the RF channel.	Identified RIS as a key component and used the FCAO algorithm to configure it; built a MATLAB model to determine the optimal Tx position for better illumination efficiency of the metal plate and RIS.
Model the Space-of-Interest (SoI), encompassing the RF channel and interactions with target objects.	Defined the SoI for human posture recognition (HPR) at the same angle as the Tx (specular reflections) and for hand gesture recognition (HGR) in front of the RIS.
Configure experimental setups to build RF sensing datasets with ground truth.	Created two public datasets of S_{21} parameters for HPR and HGR.
Train and evaluate the performance of machine learning models for human activity recognition.	Developed CNN models for classification and used confusion matrices to evaluate accuracy.

Table 6.1: Achievement of intermediate goals.

Looking ahead, there are several exciting directions for future work. One promising change is to integrate vision datasets, such as images and videos, to create a multi-modal sensing approach. This would enable the system to classify postures or gestures and also provide visual images of the person, a method referred to as *imaging* in recent studies [72–75]. Additionally, while controlled environment experiments are crucial for understanding the potential of RIS-based systems, expanding these applications to more dynamic, real-world environments is necessary to address interference challenges.

Another area for exploration is developing new methods for configuring RIS states. For example, a study of the Gerchberg-Saxton method, used in studies by Li et al. [42, 45] as described in the state-of-art, could provide valuable insights and allow for comparisons with the FCAO approach. Another direction is to expand the RIS-based setup for posture recognition explored in Chapter 4, using a larger RIS, like $1\,\mathrm{m} \times 1\,\mathrm{m}$. We have already laid the groundwork for this by analyzing and purchasing an ideal transmitter based on illumination efficiency, as detailed in Appendix A. Also, developing a 2-bit RIS with four different states could offer novel data and enable a re-evaluation of the studies from Chapter 5 using the FCAO algorithm to understand how the number of RIS states impacts classification accuracy.

Finally, shifting towards a WiFi-based system using CSI parameters instead of S_{21} in the WiFi-6 range could provide new insights and align with current trends in the field.

Overall, this work has made significant contributions to the field of human activity recognition and sets the stage for future developments that can further enhance the capabilities of RIS-based systems.

References

[1] M. H. Arshad, M. Bilal, and A. Gani, “Human activity recognition: Review, taxonomy and open challenges,” *Sensors*, vol. 22, no. 17, p. 6463, 2022. [Cited on page 1.]

[2] N. Dua, S. Singh, S. Challa, V. Semwal, and M. L. S. Kumar, *A Survey on Human Activity Recognition Using Deep Learning Techniques and Wearable Sensor Data*, 01 2023, pp. 52–71. [Cited on page 1.]

[3] D. Bouchabou, S. M. Nguyen, C. Lohr, B. LeDuc, and I. Kanellos, “A survey of human activity recognition in smart homes based on iot sensors algorithms: Taxonomies, challenges, and opportunities with deep learning,” *Sensors*, vol. 21, no. 18, 2021. [Online]. Available: https://www.mdpi.com/1424-8220/21/18/6037 [Cited on page 1.]

[4] Z. Zhuang and Y. Xue, “Sport-related human activity detection and recognition using a smartwatch,” *Sensors*, vol. 19, no. 22, 2019. [Online]. Available: https://www.mdpi.com/1424-8220/19/22/5001 [Cited on page 1.]

[5] A. Sunil, M. H. Sheth, S. E, and Mohana, “Usual and unusual human activity recognition in video using deep learning and artificial intelligence for security applications,” in *2021 Fourth International Conference on Electrical, Computer and Communication Technologies (ICECCT)*, 2021, pp. 1–6. [Cited on page 1.]

[6] P. Xu, “A real-time hand gesture recognition and human-computer interaction system,” *arXiv preprint arXiv:1704.07296*, 2017. [Cited on page 1.]

[7] Y. Yang, P. Hu, J. Shen, H. Cheng, Z. An, and X. Liu, “Privacy-preserving human activity sensing: A survey,” *High-Confidence Computing*, vol. 4, no. 1, p. 100204, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2667295224000072 [Cited on page 1.]

[8] L. Wang, C. Zhang, Q. Zhao, H. Zou, S. Lasaulce, G. Valenzise, Z. He, and M. Debbah, “Generative ai for rf sensing in iot systems,” *arXiv preprint arXiv:2407.07506*, 2024. [Cited on page 2.]

[9] G. C. Alexandropoulos, N. Shlezinger, I. Alamzadeh, M. F. Imani, H. Zhang, and Y. C. Eldar, “Hybrid reconfigurable intelligent metasurfaces: Enabling simultaneous

tunable reflections and sensing for 6g wireless communications," *IEEE Vehicular Technology Magazine*, 2023. [Cited on page 2.]

[10] Y. Liu, X. Liu, X. Mu, T. Hou, J. Xu, M. Di Renzo, and N. Al-Dhahir, "Reconfigurable intelligent surfaces: Principles and opportunities," *IEEE communications surveys & tutorials*, vol. 23, no. 3, pp. 1546–1577, 2021. [Cited on page 2.]

[11] J. ur Rehman Kazim, A. Tahir, J. Rains, T. J. Cui, A. Jabbar, M. A. Jamshed, M. Ur-Rehman, A. Alomainy, M. A. Imran, and Q. H. Abbasi, "In-home monitoring using wireless on the walls for future healthcare: Real-world demonstration," *Advanced Intelligent Systems*, vol. 5, 9 2023. [Cited on page 2.]

[12] W. Li, Q. Ma, C. Liu, Y. Zhang, X. Wu, J. Wang, S. Gao, T. Qiu, T. Liu, Q. Xiao, J. Wei, T. T. Gu, Z. Zhou, F. Li, Q. Cheng, L. Li, W. Tang, and T. J. Cui, "Intelligent metasurface system for automatic tracking of moving targets and wireless communications based on computer vision," *Nature Communications*, vol. 14, 12 2023. [Cited on page 2.]

[13] Z. Yang, H. Zhang, H. Zhang, and L. Song, "Reconfigurable intelligent surface-enabled integrated sensing and communication: a sensing-assisted communication framework," *Frontiers in Communications and Networks*, vol. 4, 2023. [Cited on page 2.]

[14] D. Dardari, N. Decarli, A. Guerra, and F. Guidi, "Los/nlos near-field localization with a large reconfigurable intelligent surface," *IEEE Transactions on Wireless Communications*, vol. 21, pp. 4282–4294, 6 2022.

[15] C. L. Nguyen, O. Georgiou, and G. Gradoni, "Reconfigurable intelligent surfaces and machine learning for wireless fingerprinting localization," 10 2020. [Online]. Available: http://arxiv.org/abs/2010.03251 [Cited on page 2.]

[16] M. Oliveira, F. M. Ribeiro, N. Paulino, O. Yurduseven, and L. M. Pessoa, "Specrf-posture: Exploring specular reflections for human posture recognition," in *2024 IEEE International Mediterranean Conference on Communications and Networking (MediterCom)*, 2024, pp. 227–232. [Cited on pages 3 and 19.]

[17] M. A. R. Ahad, A. D. Antar, and M. Ahmed, *Human Activity Recognition: Data Collection and Design Issues*. Cham: Springer International Publishing, 2021, pp. 63–75. [Online]. Available: https://doi.org/10.1007/978-3-030-51379-5_4 [Cited on page 5.]

[18] P. Patel, B. Bhatt, and B. Patel, “Human body posture recognition—a survey,” in *2017 International conference on innovative mechanisms for industry applications (ICIMIA)*. IEEE, 2017, pp. 473–477. [Cited on page 5.]

[19] N. Mohamed, M. B. Mustafa, and N. Jomhari, “A review of the hand gesture recognition system: Current progress and future directions,” *IEEE access*, vol. 9, pp. 157 422–157 436, 2021. [Cited on page 6.]

[20] Z. Hussain, Q. Z. Sheng, and W. E. Zhang, “A review and categorization of techniques on device-free human activity recognition,” *Journal of Network and Computer Applications*, vol. 167, p. 102738, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1084804520302125 [Cited on page 6.]

[21] A. Jalal, Y.-H. Kim, Y.-J. Kim, S. Kamal, and D. Kim, “Robust human activity recognition from depth video using spatiotemporal multi-fused features,” *Pattern Recognition*, vol. 61, pp. 295–308, 2017. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0031320316302126 [Cited on page 6.]

[22] S. K. Bashar, A. Al Fahim, and K. H. Chon, “Smartphone based human activity recognition with feature selection and dense neural network,” in *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. IEEE, 2020, pp. 5888–5891. [Cited on page 6.]

[23] N. Roy, A. Misra, and D. Cook, “Ambient and smartphone sensor assisted adl recognition in multi-inhabitant smart environments,” *Journal of ambient intelligence and humanized computing*, vol. 7, pp. 1–19, 2016. [Cited on page 6.]

[24] S. Khunteta, P. Saikrishna, A. Agrawal, A. Kumar, and A. K. R. Chavva, “Rf-sensing: A new way to observe surroundings,” *IEEE Access*, vol. 10, pp. 129 653–129 665, 2022. [Cited on page 6.]

[25] 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. [Cited on page 6.]

[26] 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, pp. 1–13, 2015. [Cited on page 6.]

[27] C. Xie, D. Zhang, Z. Wu, C. Yu, Y. Hu, and Y. Chen, "Rpm 2.0: Rf-based pose machines for multi-person 3d pose estimation," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 34, no. 1, pp. 490–503, 2023. [Cited on page 6.]

[28] L. Lubna, H. Hameed, S. Ansari, A. Zahid, A. Sharif, H. T. Abbas, F. Alqahtani, N. Mufti, S. Ullah, M. A. Imran *et al.*, "Radio frequency sensing and its innovative applications in diverse sectors: A comprehensive study," *Frontiers in Communications and Networks*, vol. 3, p. 1010228, 2022. [Cited on page 7.]

[29] R. Shahbazian and I. Trubitsyna, "Human sensing by using radio frequency signals: A survey on occupancy and activity detection," *IEEE Access*, vol. 11, pp. 40 878–40 904, 2023. [Cited on page 7.]

[30] S. A. Shah and F. Fioranelli, "Rf sensing technologies for assisted daily living in healthcare: A comprehensive review," *IEEE Aerospace and Electronic Systems Magazine*, vol. 34, no. 11, pp. 26–44, 2019. [Cited on page 7.]

[31] Y. Zhou, C. Xu, L. Zhao, A. Zhu, F. Hu, and Y. Li, "Csi-former: Pay more attention to pose estimation with wifi," *Entropy*, vol. 25, no. 1, p. 20, 2022. [Cited on page 7.]

[32] M. Mohtadifar, M. Cheffena, and A. Pourafzal, "Acoustic- and radio-frequency-based human activity recognition," *Sensors*, vol. 22, no. 9, 2022. [Online]. Available: https://www.mdpi.com/1424-8220/22/9/3125 [Cited on page 7.]

[33] A. Al-Saman, M. Cheffena, O. Elijah, Y. A. Al-Gumaei, S. K. Abdul Rahim, and T. Al-Hadhrami, "Survey of millimeter-wave propagation measurements and models in indoor environments," *Electronics*, vol. 10, no. 14, 2021. [Online]. Available: https://www.mdpi.com/2079-9292/10/14/1653 [Cited on page 7.]

[34] H. Zhang, B. Di, L. Song, and Z. Han, *Introductions and Basics*. Cham: Springer International Publishing, 2021, pp. 1–17. [Online]. Available: https://doi.org/10.1007/978-3-030-73499-2_1 [Cited on page 8.]

[35] M. Dajer, Z. Ma, L. Piazzi, N. Prasad, X.-F. Qi, B. Sheen, J. Yang, and G. Yue, "Reconfigurable intelligent surface: design the channel – a new opportunity for future wireless networks," *Digital Communications and Networks*, vol. 8, no. 2, pp. 87–104, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2352864821000912 [Cited on page 8.]

[36] Y. Liu, X. Liu, X. Mu, T. Hou, J. Xu, M. Di Renzo, and N. Al-Dhahir, "Reconfigurable intelligent surfaces: Principles and opportunities," *IEEE Communications Surveys & Tutorials*, vol. 23, no. 3, pp. 1546–1577, 2021. [Cited on page 8.]

[37] T. Zheng, Z. Chen, S. Ding, and J. Luo, "Enhancing rf sensing with deep learning: A layered approach," *IEEE Communications Magazine*, vol. 59, pp. 70–76, 2 2021. [Cited on page 9.]

[38] Z. Li, Y. Xie, L. Shangguan, R. I. Zelaya, J. Gummeson, W. Hu, and K. Jamieson, "Programmable radio environments with large arrays of inexpensive antennas," *GetMobile: Mobile Comp. and Comm.*, vol. 23, no. 3, p. 23–27, jan 2020. [Online]. Available: https://doi.org/10.1145/3379092.3379102 [Cited on page 9.]

[39] H. Hameed, M. Usman, A. Tahir, A. Hussain, H. Abbas, T. J. Cui, M. A. Imran, and Q. H. Abbasi, "Pushing the limits of remote rf sensing by reading lips under the face mask," *Nature communications*, vol. 13, no. 1, p. 5168, 2022. [Cited on page 9.]

[40] L. Yao, Q. Sheng, W. Ruan, T. Gu, X. Li, N. Falkner, and Z. Yang, "Rf-care: Device-free posture recognition for elderly people using a passive rfid tag array," 2015. [Cited on page 9.]

[41] J. Hu, H. Zhang, B. Di, L. Li, K. Bian, L. Song, Y. Li, Z. Han, and H. V. Poor, "Reconfigurable intelligent surface based rf sensing: Design, optimization, and implementation," *IEEE Journal on Selected Areas in Communications*, vol. 38, pp. 2700–2716, 11 2020. [Cited on pages viii, ix, 10, 11, 15, 17, 35, 38, 56, 57, 59, and 86.]

[42] L. Li, Y. Shuang, Q. Ma, H. Li, H. Zhao, M. Wei, C. Liu, C. Hao, C. W. Qiu, and T. J. Cui, "Intelligent metasurface imager and recognizer," *Light: Science and Applications*, vol. 8, 12 2019. [Cited on pages 11, 12, 15, 17, and 61.]

[43] H. Y. Li, H. T. Zhao, M. L. Wei, H. X. Ruan, Y. Shuang, T. J. Cui, P. del Hougne, and L. Li, "Intelligent electromagnetic sensing with learnable data acquisition and processing," *Patterns*, vol. 1, 4 2020. [Cited on pages 12 and 15.]

[44] Z. Wang, H. Zhang, H. Zhao, T. J. Cui, and L. Li, "Intelligent electromagnetic metasurface camera: System design and experimental results," *Nanophotonics*, vol. 11, pp. 2011–2024, 4 2022. [Cited on pages viii, 13, and 15.]

[45] L. Li, H. Ruan, C. Liu, Y. Li, Y. Shuang, A. Alù, C.-W. Qiu, and T. J. Cui, "Machine-learning reprogrammable metasurface imager," *Nature communications*, vol. 10, no. 1, p. 1082, 2019. [Cited on pages viii, 14, 15, and 61.]

[46] R. W. Gerchberg, "A practical algorithm for the determination of plane from image and diffraction pictures," *Optik*, vol. 35, no. 2, pp. 237–246, 1972. [Cited on page 14.]

[47] M. Eddy, "Rf interference and the role of filters in wi-fi 6e designs," Jul 2021. [Online]. Available: https://www.planetanalog.com/rf-interference-and-the-role-of-filters-in-wi-fi-6e-designs/ [Cited on page 15.]

[48] Y. Zhang, Y. Shao, L. Xiong, and J. Zhang, "Human activity recognition based on on-body antennas and semi-supervised learning," in *2023 17th European Conference on Antennas and Propagation (EuCAP)*. IEEE, 2023, pp. 1–5. [Cited on page 18.]

[49] Y. Kim and Y. Li, "Human activity classification with transmission and reflection coefficients of on-body antennas through deep convolutional neural networks," *IEEE Transactions on Antennas and Propagation*, vol. 65, no. 5, pp. 2764–2768, 2017.

[50] S. H. Gupta, A. Sharma, M. Mohta, and A. Rajawat, "Hand movement classification from measured scattering parameters using deep convolutional neural network," *Measurement*, vol. 151, p. 107258, 2020. [Cited on page 18.]

[51] S. Ikram and G. Ahmad, "Design & implementation of a standard gain horn antenna for optimized gain and radiation pattern using mathcad & hfss," in *2008 Second International Conference on Electrical Engineering*, 2008, pp. 1–5. [Cited on page 19.]

[52] R. Johnson and H. Jasik, *Antenna Engineering Handbook*, ser. Electronics Electrical Engineering. McGraw-Hill, 1993. [Online]. Available: https://books.google.pt/books?id=xTSNJhVlHGgC [Cited on page 19.]

[53] P. Nayeri, F. Yang, and A. Z. Elsherbeni, "Reflectarray antennas: theory, designs, and applications," 2018. [Cited on page 20.]

[54] A. Yu, F. Yang, A. Z. Elsherbeni, J. Huang, and Y. Rahmat-Samii, "Aperture efficiency analysis of reflectarray antennas," *Microwave and Optical Technology Letters*, vol. 52, no. 2, pp. 364–372, 2010. [Cited on page 21.]

[55] M. H. Dahri, M. H. Jamaluddin, F. C. Seman, M. I. Abbasi, N. F. Sallehuddin, A. Y. I. Ashyap, and M. R. Kamarudin, "Aspects of efficiency enhancement in reflectarrays

with analytical investigation and accurate measurement," *Electronics*, vol. 9, no. 11, 2020. [Online]. Available: https://www.mdpi.com/2079-9292/9/11/1887 [Cited on page 22.]

[56] I. Goodfellow, *Deep Learning*. MIT Press, 2016. [Cited on page 30.]

[57] R. Sharma, B. Deka, V. Fusco, and O. Yurduseven, "Integrated convolutional neural networks for joint super-resolution and classification of radar images," *Pattern Recognition*, vol. 150, p. 110351, 2024. [Cited on page 30.]

[58] H. Roh, S. Oh, H. Song, J. Han, and S. Lim, "Deep learning-based wireless signal classification in the iot environment." *Computers, Materials & Continua*, vol. 71, no. 3, 2022.

[59] Y. Zhou, C. Xu, L. Zhao, A. Zhu, F. Hu, and Y. Li, "Csi-former: Pay more attention to pose estimation with wifi," *Entropy*, vol. 25, no. 1, p. 20, 2022.

[60] X. Wang, X. Wang, and S. Mao, "Deep convolutional neural networks for indoor localization with csi images," *IEEE Transactions on Network Science and Engineering*, vol. 7, no. 1, pp. 316–327, 2018.

[61] M. Zhang, M. Diao, and L. Guo, "Convolutional neural networks for automatic cognitive radio waveform recognition," *IEEE access*, vol. 5, pp. 11 074–11 082, 2017. [Cited on page 30.]

[62] N. Rubab, S. Zeb, A. Mahmood, S. A. Hassan, M. I. Ashraf, and M. Gidlund, "Interference mitigation in ris-assisted 6g systems for indoor industrial iot networks," in *2022 IEEE 12th Sensor Array and Multichannel Signal Processing Workshop (SAM)*, 2022, pp. 211–215. [Cited on page 33.]

[63] V. K. Chapala and S. M. Zafaruddin, "Reconfigurable intelligent surface empowered multi-hop transmission over generalized fading," in *2022 IEEE 95th Vehicular Technology Conference: (VTC2022-Spring)*, 2022, pp. 1–5.

[64] R. Zhou, X. Chen, W. Tang, X. Li, S. Jin, E. Basar, Q. Cheng, and T. J. Cui, "Modeling and measurements for multi-path mitigation with reconfigurable intelligent surfaces," in *2022 16th European Conference on Antennas and Propagation (EuCAP)*, 2022, pp. 1–5. [Cited on page 33.]

[65] A. Goldsmith, *Wireless communications*. Cambridge university press, 2005. [Cited on page 36.]

[66] L. Carin, D. Liu, and B. Guo, “Coherence, compressive sensing, and random sensor arrays,” *IEEE Antennas and Propagation Magazine*, vol. 53, no. 4, pp. 28–39, 2011. [Cited on page 37.]

[67] R. Obermeier and J. A. Martinez-Lorenzo, “Sensing matrix design via mutual coherence minimization for electromagnetic compressive imaging applications,” *IEEE Transactions on Computational Imaging*, vol. 3, no. 2, pp. 217–229, 2017. [Cited on pages 38 and 41.]

[68] R. M. Lewis, A. Shepherd, and V. Torczon, “Implementing generating set search methods for linearly constrained minimization,” *SIAM Journal on Scientific Computing*, vol. 29, no. 6, pp. 2507–2530, 2007. [Cited on pages 40 and 79.]

[69] J. Nocedal and S. J. Wright, *Numerical optimization*. Springer, 1999. [Cited on page 41.]

[70] J. Gervacio and J. Cruz-Mandujano, “Measurement of light reflectance in different types of wood,” vol. 6046, 02 2006, pp. 604 617–604 617. [Cited on page 52.]

[71] M. Sheeny, A. Wallace, and S. Wang, “Radio: Parameterized generative radar data augmentation for small datasets,” *Applied Sciences*, vol. 10, no. 11, 2020. [Online]. Available: https://www.mdpi.com/2076-3417/10/11/3861 [Cited on page 56.]

[72] J. Gollub, O. Yurduseven, K. P. Trofatter, D. Arnitz, M. F. Imani, T. Sleasman, M. Boyarsky, A. Rose, A. Pedross-Engel, H. Odabasi *et al.*, “Large metasurface aperture for millimeter wave computational imaging at the human-scale,” *Scientific reports*, vol. 7, no. 1, p. 42650, 2017. [Cited on page 61.]

[73] J. Yao, Z. Zhang, X. Shao, C. Huang, C. Zhong, and X. Chen, “Concentrative intelligent reflecting surface aided computational imaging via fast block sparse bayesian learning,” in *2021 IEEE 93rd Vehicular Technology Conference (VTC2021-Spring)*. IEEE, 2021, pp. 1–6.

[74] J. Hu, H. Zhang, K. Bian, M. Di Renzo, Z. Han, and L. Song, “Metasensing: Intelligent metasurface assisted rf 3d sensing by deep reinforcement learning,” *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 7, pp. 2182–2197, 2021.

[75] 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, pp. 1–13, 2015. [Cited on page 61.]

[76] A-Info, *Datasheet LB-137-20 Standard Gain Horn Antenna*, https://reliantemc.com/download/A-Info/tr_LB-137-20.pdf. [Cited on page 71.]

[77] ——. Datasheet lb-137-25 standard gain horn antenna. https://reliantemc.com/download/A-Info/tr_LB-137-25.pdf. [Cited on page 71.]

[78] ——. Datasheet lb-137-15 standard gain horn antenna. https://reliantemc.com/download/A-Info/tr_LB-137-15.pdf. [Cited on page 75.]

A. Transmitter Antenna for RIS-based set up

This appendix provides a detailed analysis of the transmitter antenna selection and positioning, specifically for a RIS-based set up. While the main document discusses the configuration for a setup using a metal plate and a small RIS tile, this appendix focuses on the adjustments and considerations required when transitioning to an bigger RIS. The analysis follows similar principles but addresses the unique requirements and optimizations for using an RIS.

In transitioning to a RIS-based setup, the focus was on selecting the appropriate transmitter antenna from the available pyramidal horn subtypes: Broadband and Standard Gain. Broadband horns cover a wide frequency range, which is advantageous for applications requiring flexibility across various frequencies. However, for the specific case of operating within the WiFi-6E frequency range, consistent gain within this band is crucial. Therefore, the Standard Gain Horn Antenna was preferred. Therefore, for operation within the WiFi-6E frequency range, the selected transmitter antenna is the pyramidal standard gain horn antenna LB-137 from A-INFO.

The LB-137 antenna offers four different models featuring gains of [10, 15, 20, 25] dB. As mentioned earlier, symmetry in the radiation pattern in both planes is crucial. After analyzing the datasheets of the four models, it was observed that the 20 and 25 dB models exhibited a significantly less symmetrical pattern [76],[77], leading to their exclusion from consideration. Hence, the selection is limited to the 10 and 15 dB models.

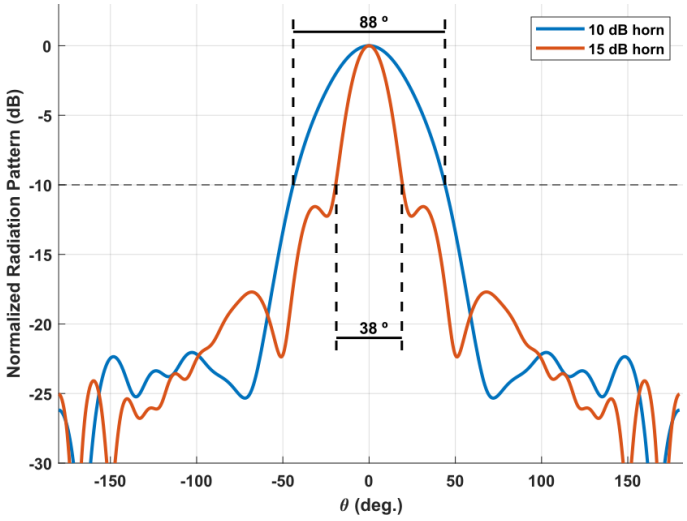


Figure A.1: E-plane radiation patterns of LB-137 horn antenna with gain 10dB and 15dB at 6.5GHz

Figure A.1 illustrates the E-plane radiation pattern at 6.5 GHz and corresponding $-10dB$ beamwidth for the 10 and 15dB horn antennas. It is observed that they exhibit a directive beam in their radiation, making them well-suited for illuminating a reflectarray aperture. Unlike the LB-20180 antenna, for which simulating the radiation pattern was not feasible due to its complexity and less comprehensive datasheet, the radiation patterns of the LB-137 antennas were successfully simulated by a member of our workgroup. The antennas were modeled using a cos^q radiation pattern, and the q values were determined through a comparative analysis between the simulated radiation patterns and the cos^q model. As shown in Figure A.2, the 2D radiation pattern closely aligns in both planes, extending up to approximately $\pm$ 50 degrees for the 10dB horn and $\pm$ 30 degrees for the 15dB horn. The chosen values are $q_{10dB} = 4$ and $q_{15dB} = 14$, where higher q values signify a more directional beam.

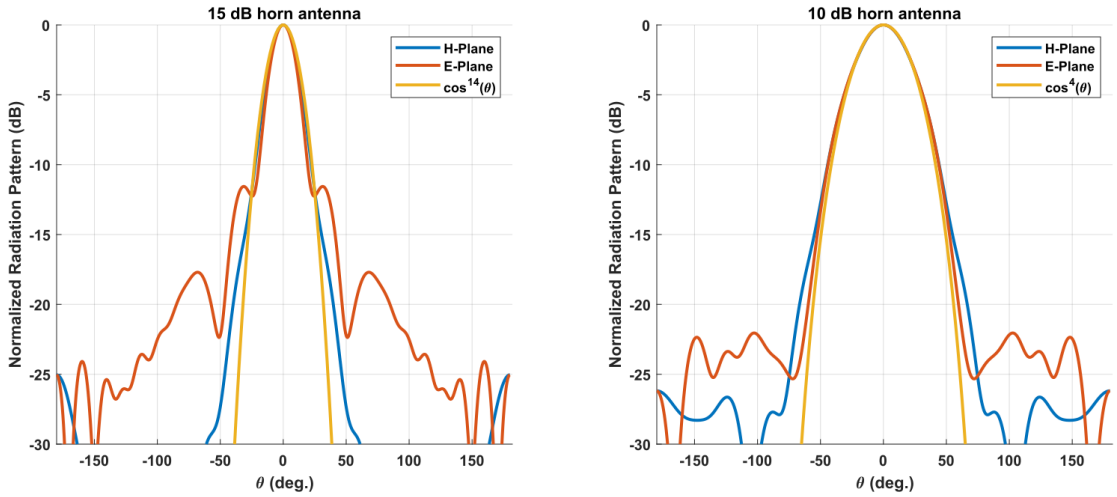


Figure A.2: Normalized pattern and cos^q model of radiation patterns of the a) LB-137-15 (15dB) b) LB-137-10 (10dB).

Based on the equations provided in Section 4.1.1, a new study of the aperture efficiency was conducted for a 1m $\times$ 1m RIS. This study aimed to determine the parameters θ_0, H, and the FBP to optimize the system's performance, considering the values $q_{10dB} = 4$ and $q_{15dB} = 14$ and $q_e = 1$.

Figure A.3 shows how the efficiencies vary with θ_0, revealing that as the angle increases, η_s decreases while η_i increases. Similar to the hand gesture recognition experiment (Chapter 5), θ_0 should be within the range [30;45] degrees to ensure proper illumination and avoid interference with the region of interest, as discussed in Section 5.1.1. Analysis of the graph suggests that within this range, the maximum aperture efficiency is achieved for both 10dB and 15dB. Consequently, the chosen θ_0 is 35 degrees.

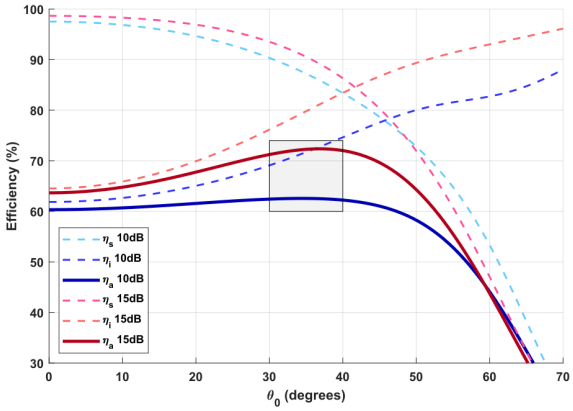


Figure A.3: Efficiencies as a function of θ_0

Figure A.4 shows the variation with H/D, where D = 1m, and θ_0 is 35 degrees.For the 10 dB horn antenna, the optimal distance H was determined to be 0.50 m, while for the 15 dB horn antenna, it was 0.95 m. These values were chosen after analyzing the trade-offs between aperture efficiency and the uniformity of the field generated across the RIS plane.

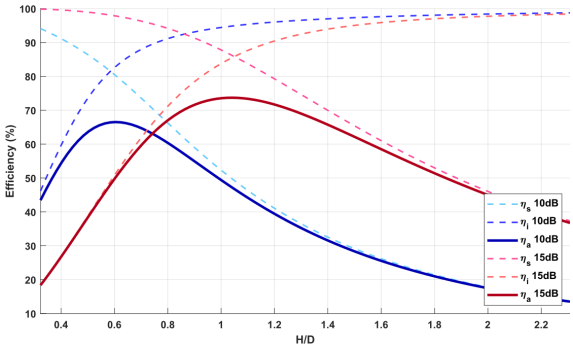


Figure A.4: Efficiencies as a function of H/D with chosen $\theta = 35°$

Finally, the ideal positioning of the focal point (FBP) was investigated, namely the values of x0 and y0. On the x-axis, the expected symmetric behavior was observed ($x_0 = 0$), while on the y-axis, an asymmetry was identified. By studying the optimal y0 value to achieve a beam with a greater amplitude at the center of the RIS, as depicted in Figure A.5, the selected values are $(x_0, y_0, z_0)_{10dB} = (0, -0.153m, 0)$ and $(x_0, y_0, z_0)_{15dB} = (0, -0.106m, 0)$.

Figure A.6 illustrates the radiation intensity distribution over the reflectarray for 10 and 15dB, respectly. An essencial metric is the relative field strength at the reflectarray's edges in comparison to its maximum, referred to as edge taper (ET). Conventionally, a typical ET value is around -10dB. It is very important to note that the chosen values H and y_0 were selected not only based on the range for maximum efficiency but also to achieve a -10 edge taper.

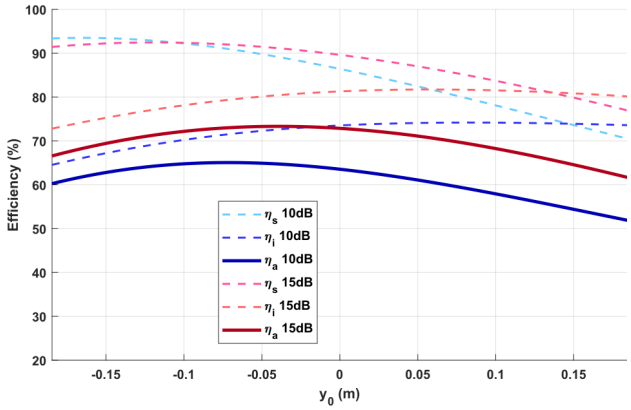
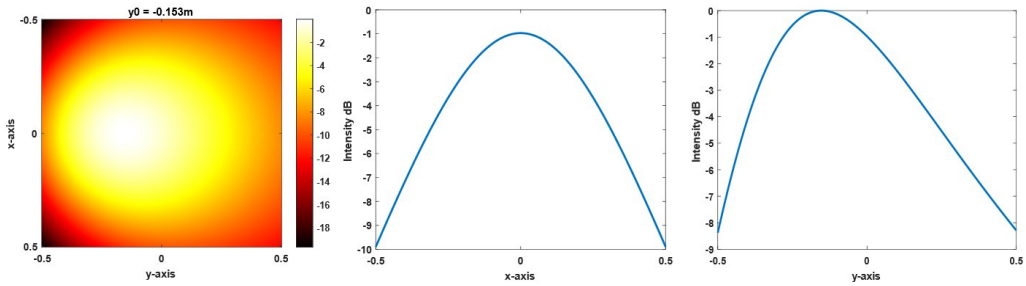


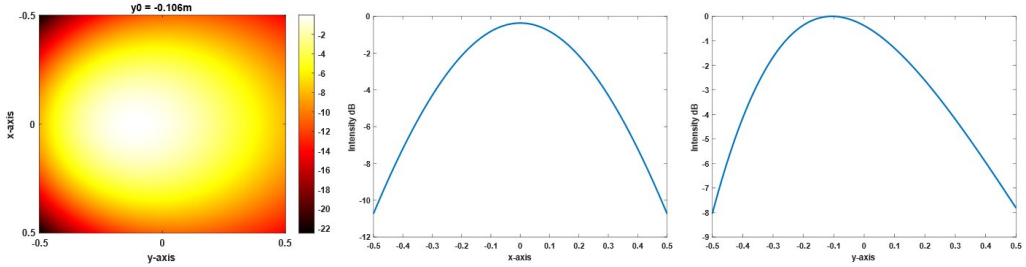
Figure A.5: Efficiencies as a function of y_0 with chosen $\theta = 35°$ and $H_{10dB} = 0.50m$, $H_{15dB} = 0.95m$

This analysis achieved an aperture efficiency value of $\eta_a = 63\%$ for 10dB gain and $\eta_a = 72\%$ for 15dB gains. Therefore, based solely on this analysis, choosing the 15dB antenna seems to be the most appropriate option. However, it is also required an analysis of which distance H is more suitable for our sensing application is necessary to make an appropriate choice, as illustrated in Figure A.7.

It is important to note that the distance from the antenna to the center of the reference frame is given by $H/cos(\theta)$, where $\theta = 35$ degrees. Consequently, the 15dB gain was chosen to ensure that the antenna's position does not interfere with the space of interest where the object to be sensed is located.



(a) 10dB horn antenna with feed beam point (0, -0.153,0)m



(b) 15dB horn antenna with feed beam point (0, -0.106,0)m

Figure A.6: Intensity in dB on the reflectarray (left), along x-axis (middle) and along y-axis (right).

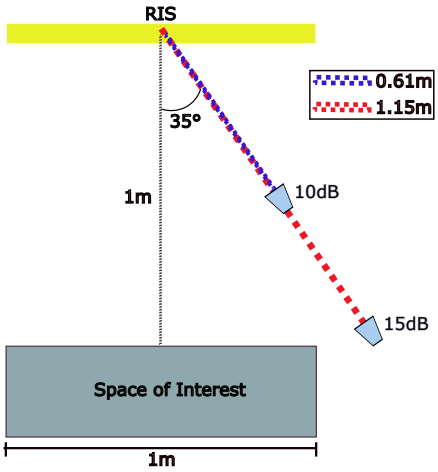


Figure A.7: Setup with Horn Antennas (10dB or 15dB Gain).

Therefore, the chosen antenna is the pyramidal horn antenna LB-137-15 from A-INFO, featuring an aperture opening of $113.20 \times 143.20 mm^2$ and a gain of 17.79dBi at the center design frequency of 6.5GHz [78].

B. Preliminary Experiments with Metal Plate

This appendix describes the preliminary experiments conducted to justify the use of LNAs and the selection of an anechoic chamber for human posture recognition with metal reflectors.

B.1. Use of LNAs and Experimental Setup Adaptations

Initial experiments were performed in the laboratory without LNAs, with the SoI positioned directly in front of the metal plate. The frequency range [3.17, 3.22] GHz was tested in these preliminary experiments. Boundary conditions of the SoI were tested: SoI empty, SoI with a person standing, and SoI with a person holding a metal plate to increase reflections (Fig. B.1). It was observed that the signal exhibited minimal variation (Fig. B.2).

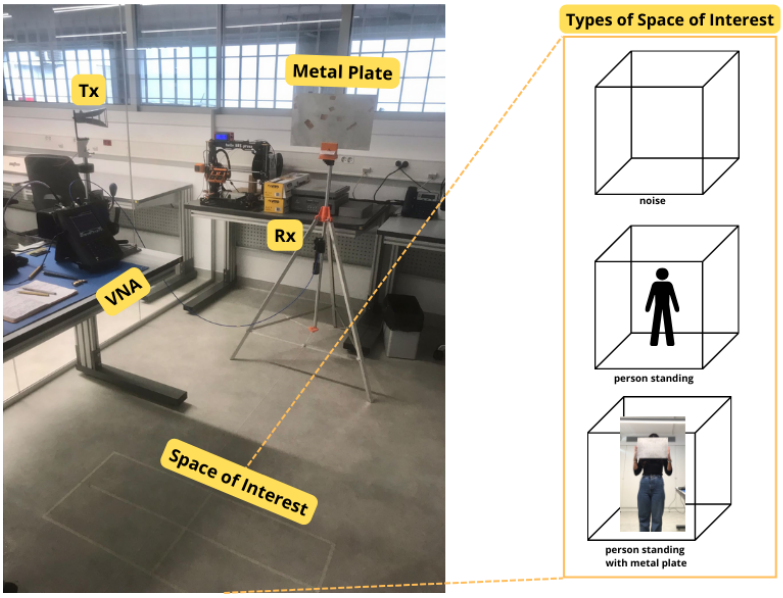
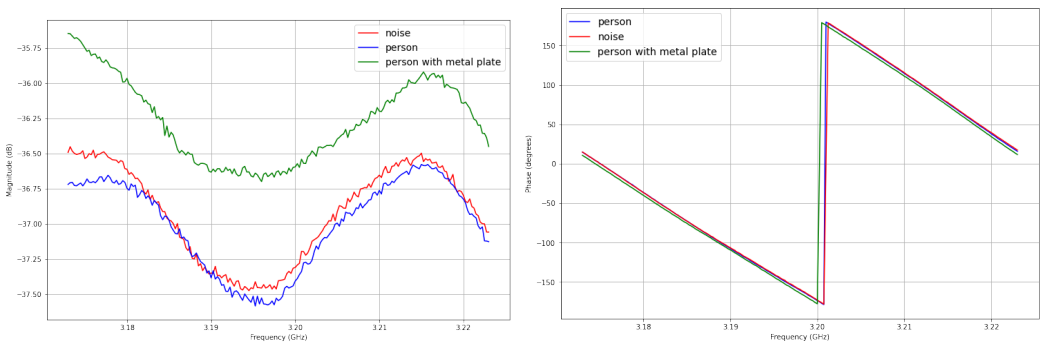


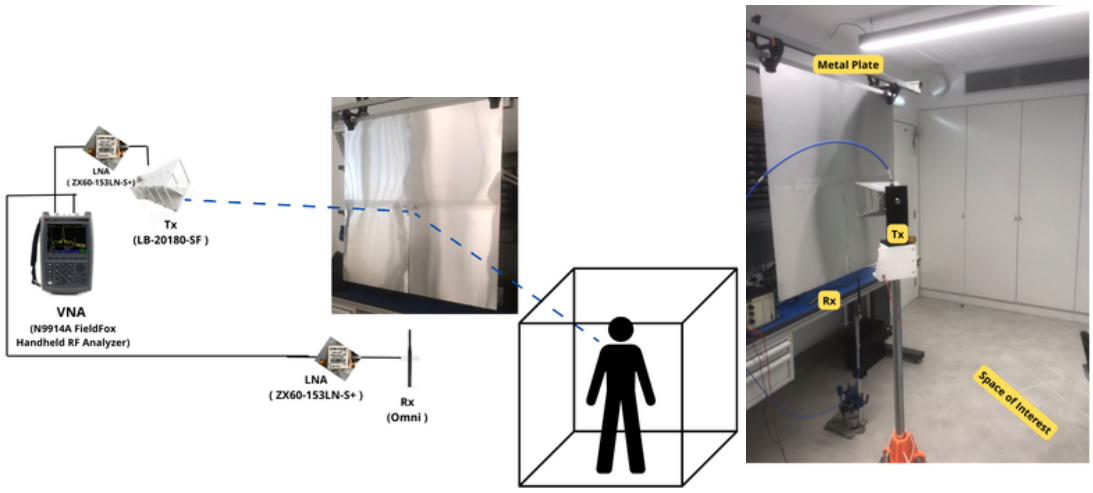
Figure B.1: First Experimental Setup in the laboratory.



(a) Magnitude (b) Phase

Figure B.2: S_{21} parameters in the first experiment indoor.

To address this, two changes were implemented: LNAs were introduced in the setup, connected to both the transmitter (Tx) and receiver (Rx), and the SoI was adjusted to align with the angle of the antenna emission. This adjustment was made to take advantage of the specular reflections from the metal plate for maximum signal reception (Fig. B.3a). Additionally, a larger metal plate measuring $1\,\mathrm{m} \times 1\,\mathrm{m}$ was used. The updated laboratory setup is shown in Fig. B.3b. The second set of experiments tested the frequency range [2.35, 2.45] GHz to identify the optimal frequency range for better signal performance.



(a) Schematic (b) Implementation

Figure B.3: Second Experiment in the laboratory.

With this revised setup, different postures: Standing, sitting, and Lying Down were tested. Fig. B.4 illustrates the new results, which revealed that the differences between curves were minimal, with only slight variations in the valley minimum.

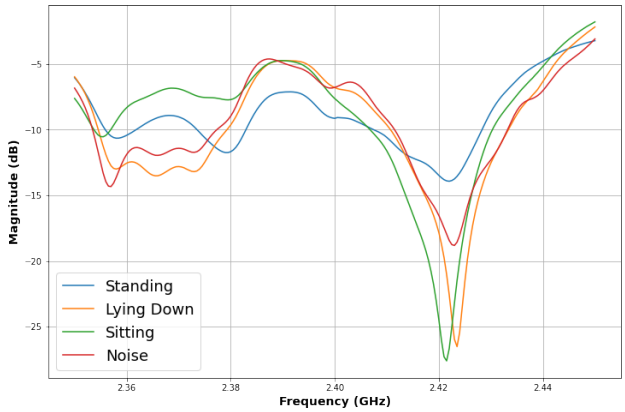


Figure B.4: S_{21} magnitude in the second experiment indoor.

This suggested possible interference from reflections off other laboratory components and signal capture from routers or other indoor signal sources. Consequently, it was concluded that the experiments should be conducted in an outdoor environment.

B.2. Transition to Anechoic Chamber

The outdoor experiments were conducted at the INESC TEC in a open area to avoid reflections from other surfaces. Fig. B.5a shows the setup without absorbent material between Tx and Rx, while Fig. B.5b includes the absorbent to block the Line-of-Sight. The postures tested are shown in Fig. B.5c. The frequency range used was [5.9, 6.5] GHz, aligning with the WiFi-6E frequency range, which was the focus of the final study.



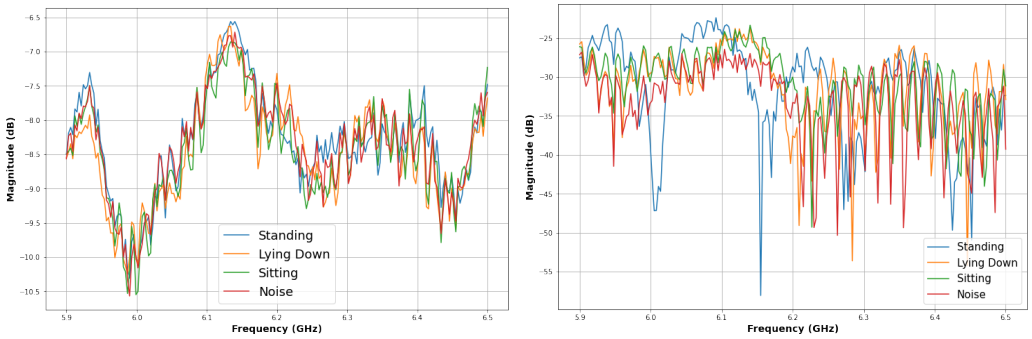
(a) Implementation of the set up

(b) Absorbent between path Tx-Rx.

(c) Type of postures.

Figure B.5: First Experiment in outdoor environment.

Fig. B.6 illustrates the magnitude of S_{21} measurements with and without the absorbent. Without it, the direct Tx-Rx path dominated the signal, masking posture differences. By blocking the direct Tx-Rx path with the absorber, distinct behavioral patterns for each posture became observable.



(a) without absorbent

(b) with absorbent

Figure B.6: S_{21} magnitude in the outdoor environment.

Although better results were achieved, differences between postures remained minimal, likely due to outdoor interference. Therefore, the setup was moved to an anechoic chamber.

C. FCAO Algorithm

C.1. Initial Value of t_k using the *Pattern Search* Method

The Augmented Lagrangian Method may converge to a local minimum far from the global minimum if the starting point is chosen poorly. Therefore, the *pattern search* algorithm was employed to determine the initial value of t_k [68]. This method was used to solve the problem (P5) to find the initial value $t_k^{(0)}$ that guarantees the minimum mutual coherence μ with the other time configurations of matrix T fixed, i.e., with T_{-k} fixed.

In general, the *pattern search* method is used to solve an optimization problem expressed by the mathematical formulation described in equation C.1.

$$\begin{aligned} \min \quad & f_m(x) && m = 1, .., M \\ \text{s.t.} \quad & g_j(x) \leq 0 && j = 1, .., J \\ & h_k(x) = 0 && k = 1, .., K \\ & x_i^L \leq x_i \leq x_i^U && i = 1, .., N \\ & x \in \Omega \end{aligned} \tag{C.1}$$

where x_i is the i-th variable to optimize in a vector **x**, bounded by limits x_i^L and x_i^U; f_m is the objective function—there can be multiple in a multi-objective optimization case; the function g represents an inequality constraint; and the function h represents an equality constraint.

This method is then used to minimize f_m while adhering to the three conditions mentioned. Specifically, in the context discussed in previous sections, the following concrete formulations can be applied to solve problem (P5):

$$\begin{aligned} \min \quad & ||\mathbf{u}||_1 && (m = 1) \\ \text{s.t.} \quad & \mathbf{1}^T \cdot \ t_{k,l} \ = \delta && l = 1, .., L \\ & 0 \leq t_{k_i} \leq \delta && i = 1, .., L \cdot Na \end{aligned} \tag{C.2}$$

where $\delta = 0.3s$; $u_{m,m'} = \frac{\gamma_m^H \cdot \gamma_{m'}}{||\gamma_m||_2 \cdot ||\gamma_{m'}||_2}$; and $\gamma_m = \gamma_m(t_k)$.

C.2. Gradient derivation for t_k minimization step

In the gradient calculation for the function g, we assume $||\gamma_m(t_k)||_{l_2} > 0$ for all columns of Γ. The gradient vector's l-th component is given by:

$$\nabla g_l = \sum_{1 \leq m \neq m' \leq M} \rho \left(\frac{\partial}{\partial p_l} \left(\hat{f}_m^H(t_k) \hat{f}_{m'}(t_k) \right) \right) \left(\frac{f_m(t_k)^H f_{m'}(t_k)}{\|f_m(t_k)\|_{l_2} \|f_{m'}(t_k)\|_{l_2}} - z_{m,m'} \right) \tag{C.3}$$

where:

$$\frac{\partial \hat{f}_i(p)}{\partial p_l} = \frac{\partial}{\partial p_l} \left(\frac{f_i(p)}{\|f_i(p)\|_{\ell_2}} \right) = \|f_i(p)\|_{l_2}^{-1} \left(I - \frac{f_i(p) f_i^H(p)}{\|f_i(p)\|_{l_2}^2} \right) \frac{\partial f_i(p)}{\partial p_l} \tag{C.4}$$

$$\frac{\partial}{\partial p_l} \left(\hat{f}_i^H(p) \hat{f}_j(p) \right) = \frac{\partial \hat{f}_i^H(p)}{\partial p_l} \hat{f}_j(p) + \hat{f}_i^H(p) \frac{\partial \hat{f}_j(p)}{\partial p_l} \tag{C.5}$$

For the specific case where $f = \gamma_m$, a column of the matrix Γ, and given $\gamma_m = t_k \cdot \alpha_m$, the partial derivative with respect to t_k is:

$$\frac{\partial \gamma_m(t_k)}{\partial t_{kl}} = \alpha_{m_l} \tag{C.6}$$

Given that $t_k = (t_{k,1,1} \ \ldots \ t_{k,1,Na} \ \ldots \ t_{k,L,1} \ \ldots \ t_{k,L,Na})$ and $\alpha_m = (\alpha_{m,1,1} \ \ldots \ \alpha_{m,1,Na} \ \ldots \ \alpha_{m,L,1} \ \ldots \ \alpha_{m,L,Na})$, the gradient's partial derivatives vector becomes:

$$\frac{\partial \gamma_m(t_k)}{\partial t_k} = \left[\frac{\partial \gamma_m(tk)}{\partial t_{k1,1}} \ \ldots\ldots \ \frac{\partial \gamma_m(tk)}{\partial t_{kL,Na}} \right] = \left[\alpha_{m_{1,1}}, \ldots, \alpha_{m_{L,Na}} \right]$$

Substituting this into the gradient expressions gives:

$$\frac{\partial \hat{\gamma}_m(t_k)}{\partial t_{kl}} = \|\gamma_m(t_k)\|_{l_2}^{-1} \left(I - \frac{\gamma_m(t_k) \gamma_m^H(t_k)}{\|\gamma_m(t_k)\|_{l_2}^2} \right) \cdot \alpha_{m_l} \tag{C.7}$$

$$\frac{\partial}{\partial t_{kl}} \left(\hat{\gamma}_m^H(t_k) \hat{\gamma}_{m'}(t_k) \right) = \frac{\partial \hat{\gamma}_m^H(t_k)}{\partial t_{kl}} \hat{\gamma}_{m'}(t_k) + \hat{\gamma}_m^H(t_k) \frac{\partial \hat{\gamma}_{m'}(t_k)}{\partial t_{kl}} \tag{C.8}$$

C.3. Code

This section provides Python code snippets relevant to the implementation of the FCAO Algorithm and the Augmented Lagrangian Method.

```
def FCAO_optimization(T_,A_, K, L_,Na_, M_, NAL, NAM,tolerance, step_size,filename):
    algorithm = PatternSearch()
    mus = np.empty(200, dtype=np.float64)
    T_upt = T_.copy()
    mu_ = calc_mu(T_,A_,M_)
    u_max_norm_ = 20000 #for the fisrt value with k=0
    original_stdout = sys.stdout
    with open(filename, 'wt') as file:
        sys.stdout = file
        i =0
        while True:
            print(i)
            k = i%K
            #pattern search for frame k
            problemk = MutualCoherenceProblem3(Na_,L_,M_,T_upt, A_, k,u_max_norm_)
            tk0 = T_upt[k, :]
            resk = minimize(problemk, algorithm, verbose=True, seed=1,
            x0=tk0,termination=('n_gen', 330), save_history=True)
            tk0_novo = resk.X

            try:
                u_final = resk.opt.get("u_final")
                parte_real, parte_imaginaria = u_final[:, :, 0],  u_final[:, :, 1]
                u_complex = parte_real + 1j * parte_imaginaria
                u_final = u_complex[0]

            except AttributeError:  #convergence! (pattern search didnt found a solution)
                break
            i+=1
            #augmented lagrangian
            u_2, tk2, u_tk2, u_update_, u_de_tk_update_, u_max_norm_=aug_lagrag(u_final,
            tk0_novo, T_upt, A_,k ,L_ ,Na_, M_, NAL, NAM, tolerance, step_size)
            tk_linha = tk2
            uu = np.abs(u_tk2)
            mu_linha = np.sum(2*uu)/(M*(M-1))
            mus[i] = mu_linha

            if (mu_linha <= mu_):
                mu_ = mu_linha
                T_upt[k,:] = tk_linha

        sys.stdout = original_stdout
    return T_upt, mu_, mus
```

Listing C.1: FCAO Algorithm function.

```
from pymoo.algorithms.soo.nonconvex.pattern import PatternSearch
from pymoo.optimize import minimize
from pymoo.core.problem import ElementwiseProblem

class MutualCoherenceProblem3(ElementwiseProblem):

def __init__(self,Na,L,M, Tx, Ax, k):
    self.Na = Na
    self.L = L
    self.M = M
    self.m_m = combine_mm(M)
    self.Tx = Tx
    self.Ax = Ax
    self.k = k
    xl = np.zeros(Na * L)  #lower limit
    xu = np.ones(Na * L) * 0.3  #upper limit
    super().__init__(n_var=Na * L, n_obj=1, n_ieq_constr=1, n_eq_constr=L,
    xl=xl, xu=xu)
    self.Tx = Tx
    self.Ax = Ax
    self.k = k
    self.max_u_norm = max_u_norm

def _evaluate(self, x, out, *args, **kwargs):
    Tk_opt = x
    Tx_opt = self.Tx.copy()
    Tx_opt[self.k, :] = Tk_opt

    #Recalculate Gamma and u with the new values of Tk_opt
    Gama = np.dot(Tx_opt, self.Ax)
    u = vetor_u2(Gama, self.M)
    out["F"] = norm(u, 1)
    out["u_final"] = np.column_stack((u.real, u.imag))

     # Check the constraint: the sum of every Na terms should be 0.3
    ttk_reshaped = Tk_opt.reshape(self.L, self.Na)   # Rearrange groups into columns
    group_sums = np.sum(tk_reshaped, axis=1) # Sum each group of Na terms
    constraint_violation = (group_sums - 0.3)
    out["H"] = constraint_violation

    u_norm_constraint = np.linalg.norm(u, 1) - self.max_u_norm
    out["G"] = u_norm_constraint
```

Listing C.2: Pattern search problem with Pymoo library.

```
def aug_lagrag(u_final_, tk0_novo_, T_, A_,k,L_,Na_, M, NAL, NAM, tolerance, step_size):
    dim = int(M * (M - 1) / 2)
    betas = np.random.rand(dim) + 1j * np.random.rand(dim)
    rho = 1
    mm = combine_mm(M)
    u0 = u_final_
    tk0 = tk0_novo_
    u_update, u_de_tk_update = [], []
    for i in range(NAL):
        print('IT', i)
        u_j0, tk_j0 = u0, tk0
        T0 = T_.copy()
        T0[k, :] = tk_j0
        Gama_f = np.dot(T0, A_)
        u_beta = vetor_u2(Gama_f, M,mm)
        #atualization of betas e rho
        betas += rho * (u_j0 - u_beta)
        lim_conv = 1e-4
        convergence_criterion = np.linalg.norm(u_j0 - u_beta)
        if convergence_criterion < lim_conv:
            rho /= 1.3 # Increase rho
        else:
            rho *= 1.3  # Decrease rho
        z_ = u_j0 - (betas / rho)
        u_initial_flattened = np.concatenate([u_j0.real, u_j0.imag])

        for j in range(NAM):
            u_upt = optt.minimize(obj_func, u_initial_flattened,
            args=(z_, rho), method='L-BFGS-B')
            u_optimized = u_upt.x[:dim] + 1j * u_upt.x[dim:]
            u_update.append(np.linalg.norm(u_optimized,1))
            tk_upt = projected_gradient_descent2(tk_j0, Gama_f, A_,T0, u_optimized
            ,tolerance,k,rho,
            betas, L_,Na_, step_size)
            T0[k, :] = tk_upt
            Gama_f = np.dot(T0, A_)
            u_new = vetor_u2(Gama_f, M,mm)
            z_ = u_new - (betas / rho)
            u_initial_flattened = np.concatenate([u_new.real, u_new.imag])
            u_de_tk_update.append(np.linalg.norm(u_new,1))
            tk_j0 = tk_upt
            print('iteração',j)
        u0 = u_optimized
        tk0 = tk_upt
    max_u_norm = np.linalg.norm(u_new,1)
    return u_optimized, tk_upt, u_new, u_update, u_de_tk_update, max_u_norm
```

Listing C.3: Augmented Lagrangian function.

```
def obj_func(x, z, rho):
    u_real = x[:len(x)//2]
    u_imag = x[len(x)//2:]
    u = u_real + 1j * u_imag
    return np.linalg.norm(u, 1) + (rho/2) * np.linalg.norm(u - z, 2)**2
```

Listing C.4: Objective function for u minimization with Gradient Descendent Method using L-BFGS-B.

```
def grad_g3(Gama, A,k,u_optimized_,rho,betas_,L,Na):

    z = u_optimized_ - (betas_/rho)
    K_, M_ = Gama.shape
    m_m_prime = combine_mm(M_)

    grad = np.zeros(L*Na,dtype = 'complex_')

    f_m,f_m_prime = np.zeros(len(m_m_prime), dtype = 'complex'), np.zeros(len(m_m_prime),
    dtype = 'complex')
    f_norm_m, f_norm_m_prime = np.zeros(len(m_m_prime), dtype = 'complex'),
    np.zeros(len(m_m_prime), dtype = 'complex')

    m = 0
    for i in m_m_prime:
        f_m[m],f_m_prime[m]  = np.conj(Gama[k, i[0]]), Gama[k,  i[1]]
        f_norm_m[m], f_norm_m_prime[m] = np.linalg.norm(f_m[m]) , np.linalg.norm(f_m_prime[m])
        m+= 1

    for j in range(L * Na):
        n = 0
        for i in m_m_prime:
            z_value = z[n]
            deriv_f_m, deriv_f_m_prime = A[j, i[0]],  A[j,i[1]]

            deriv_f_m_norm = (1/f_norm_m[n]) * deriv_f_m  * (np.eye(1) -
            ((f_m[n]*np.conj(f_m[n]))/f_norm_m[n]**2))
            deriv_f_m_prime_norm = (1/f_norm_m_prime[n]) * deriv_f_m_prime *
            (np.eye(1) -((f_m_prime[n]*np.conj(f_m_prime[n]))/f_norm_m_prime[n]**2))

            deriv = (deriv_f_m * f_m_prime[n]) + (f_m[n] * deriv_f_m_prime)

            calc_2 = ((f_m[n] *f_m_prime[n])/(f_norm_m[n] *f_norm_m_prime[n])) - z_value
            grad[j] += rho* deriv * calc_2
            n+=1

    return grad.real
```

```
# Projection onto the feasible set using Itk constraints
def project_to_set_Itk(z, Na=2,L=16, target_sum=0.3):
    z_clipped = np.maximum(z, 0)
    z_sum = np.sum(z_clipped)
    # Calculate the sum of each group of L terms
    group_sums = np.array([np.sum(z_clipped[i:i+Na]) for i in range(0, len(z_clipped), Na)])

    # Calculate the scale factor for each group
    scale_factors = target_sum / group_sums

    # Distribute the scaled z values over each group of L terms
    z_normalized = np.zeros_like(z)
    for i, scale_factor in enumerate(scale_factors):
        z_normalized[i*Na: (i+1)*Na] = z_clipped[i*Na: (i+1)*Na] * scale_factor

    return z_normalized

def projected_gradient_descent2(initial_x, Gama, A,T_, u_optimized_,
convergence_threshold, k,rho,betas_, L, Na, step_size):
    'Algorithm with a given tolerance'
    x = initial_x.copy()
    prev_x = x.copy()
    convergence = False
    T_new = T_.copy()
    u_norm_target =np.linalg.norm(u_optimized_,1)
    M=32
    mm = combine_mm(M)
    i=0
    while not convergence:
        T_new[k, :] = x
        Gama_ = np.dot(T_new, A)
        gradient_g = grad_g3(Gama_, A,k,u_optimized_,
        rho,betas_,L=16,Na=2)  # Recalculate the gradient
        #tirei a linha que tinha grad_g2
        z = x - step_size * gradient_g
        z_real = z.real
        x = project_to_set_Itk(z_real, Na=2,L=16)  #Project norm of z onto feasible set

        # Check convergence criterion
        change_in_x = np.linalg.norm(x - prev_x)
        print(change_in_x, np.linalg.norm(vetor_u2(Gama_,M,mm),1))
        if change_in_x < convergence_threshold:
            convergence = True
        prev_x = x.copy()
    return x
```

Listing C.5: t_k minimization with Project Gradient Descendent Mehtod.

C.4. Results from authors of FCAO Algorithm

Below are the main results from the original FCAO algorithm study [41], used as a benchmark for comparing with our findings.

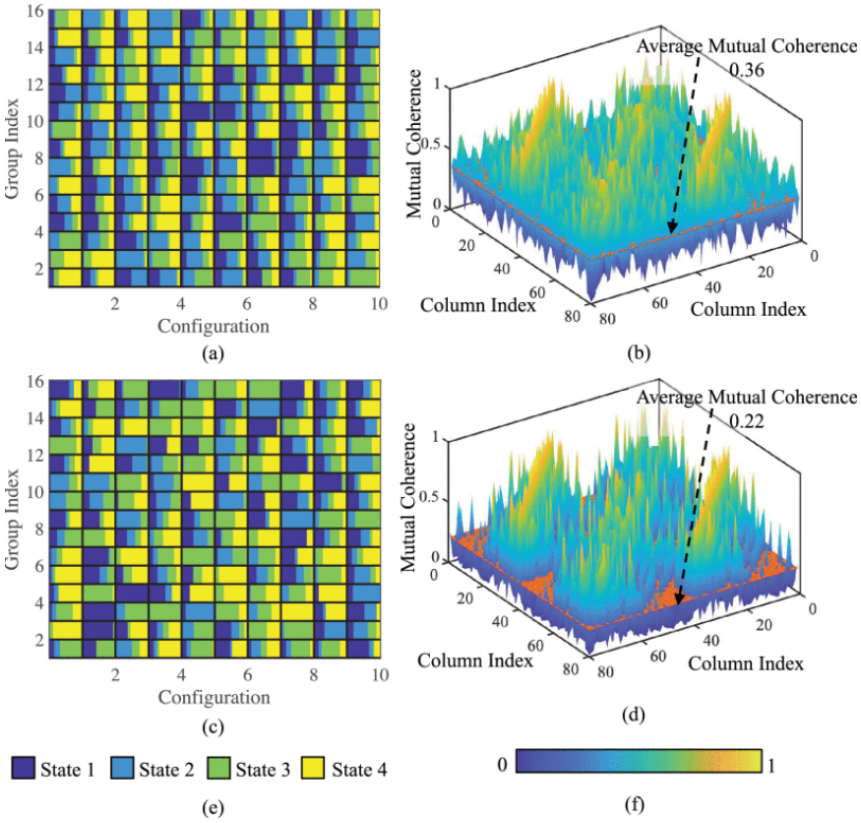


Figure C.1: From [41]: (a) Illustration of a random configuration; (b) Mutual coherence of the measurement matrix given the random configuration; (c) Illustration of the optimized configuration obtained by Algorithm 1; (d) Mutual coherence of the measurement matrix given the optimized configuration; (e) Colors that represents the duration of different states; (f) Color scaling representing the coherence to be from 0 to 1.

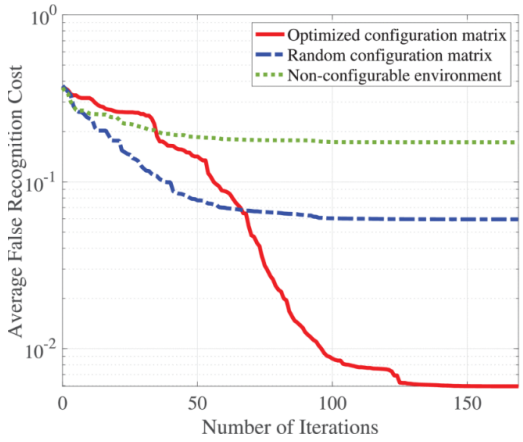


Figure C.2: From [41]: Average false recognition cost vs. the number of iterations of Algorithm 2 given optimized configuration matrix and random configuration matrix. K=10 .

D. Published Paper: SpecRF-Posture: Exploring Specular Reflections for Human Posture Recognition

This paper was presented at the 2024 IEEE International Mediterranean Conference on Communications and Networking (MeditCom), held from July 8 to 11, 2024, in Madrid, Spain. It was added to IEEE Xplore on August 12, 2024, and can be accessed via the doi: 10.1109/MeditCom61057.2024.10621437. The paper was published by IEEE.

Titled "SpecRF-Posture: Exploring Specular Reflections for Human Posture Recognition", the paper presents a low-cost method for Human Posture Recognition using RF signals. By manipulating the environment with specular reflections and analyzing S_{21} parameters in the WiFi-6E range, the approach achieves 99.17% accuracy. The paper also provides publicly available datasets* and software for data acquisition†.

Authors: Mariana Oliveira, Francisco M. Ribeiro, Nuno Paulino, Okan Yurduseven, Luís M. Pessoa

*M. Oliveira, F. Ribeiro, N. Paulino, O. Yurduseven and L. Pessoa, "SpecRF-Posture Dataset". Zenodo, Apr. 03, 2024. doi: 10.5281/zenodo.10911412.

†F. Ribeiro and M. Oliveira, "SpecRF-Posture". GitHub, Apr. 03, 2024. `https://github.com/franciscombr/SpecRF-Posture`

2024 IEEE International Mediterranean Conference on Communications and Networking (MeditCom) | 979-8-3503-0948-5/24/$31.00 ©2024 IEEE | DOI: 10.1109/MediCom61057.2024.10621437

SpecRF-Posture: Exploring Specular Reflections for Human Posture Recognition

Mariana Oliveira*, Francisco M. Ribeiro*, Nuno Paulino*†, Okan Yurduseven‡, and Luís M. Pessoa*†
*INESC TEC, Porto, Portugal
†Faculdade de Engenharia, Universidade do Porto, Portugal
‡Queen's University Belfast, United Kingdom
{mariana.s.fonseca, francisco.m.ribeiro, luis.m.pessoa, nuno.m.paulino}@inesctec.pt, okan.yurduseven@qub.ac.uk

***Abstract*—This paper presents SpecRF-Posture, a novel low-cost approach for accurate Human Posture Recognition (HPR) using Radio Frequency (RF) signals. SpecRF-Posture leverages S21 parameters within the WiFi-6E frequency range for classification. We obtain a dataset of S21 parameters for different postures by performing beamscanning through mechanical rotation of a horn transmitter aimed at a reflective surface that illuminates the space of interest. We determine the S21 parameters of the signals that are then reflected back from the space onto an omni-directional receiver. Thus for each posture we attain the S21 parameters of each possible illumination direction of the space. Experimental results demonstrate that SpecRF-Posture achieves an accuracy of $99.17\,\%$ in posture classification, highlighting its effectiveness. Additionally, an RF dataset was acquired using a software package for automatic data acquisition within the WiFi-6E frequency range, and both the dataset and the software package have been made publicly available.**

***Index Terms*—Posture Recognition, RF Sensing, Wifi-6E, Convolutional Neural Network (CNN)**

I. Introduction

Human Posture Recognition (HPR) is widely used in various applications such as smart home environments, surveillance, and sports analysis [1]–[3]. However, traditional vision-based systems are often susceptible to performance degradation under poor lighting conditions, often require significant processing power, and raise significant privacy concerns. These limitations have motivated the exploration of alternative methodologies. Namely, in recent years, the utilization of Radio Frequency (RF) signals for HPR has emerged as a promising approach in wireless sensing technology, as an alternative to computer vision. HPR using radio signals is based on the influence of human body disposition on the propagation of wireless signals, and on the analysis of signal characteristics to detect and classify human posture [4].

Recent studies with RF systems have demonstrated the effectiveness of S_{21} parameters for human body-related classification tasks [9]–[11]. Inspired by these advancements and the enhanced network performance offered by WiFi-6E [12], our system leverages S_{21} parameters within this frequency range for posture recognition. By analyzing the S_{21} parameters, we introduce SpecRF-Posture, a novel system designed for accurate classification of human posture.

In this work, we explore the use of a low cost architecture based only on passive specular reflections that occur in the path between a steerable horn transmitter, a reflective surface, the Space-of-Interest (SoI), and an omni-directional receiver. That is, our system achieves beam scanning by mechanically rotating the transmitter at regular intervals, capturing a sweep of the S_{21} parameters of the SoI for several human postures. With this, we aim to evaluate the viability of performing HPR or similar tasks using passive scatters within the propagation medium, using data retrieved from a real-word experimental setup. Moreover, we provide an evaluation based on specular reflections as baseline for a future Reconfigurable Intelligent Surface (RIS) based system we envision as future work.

The contributions of this paper are as follows:

- we propose low-cost RF system for posture recognition that utilizes specular reflections, achieving an accuracy of $99.17\,\%$ based on a simple CNN model;
- we implement and distribute the experimental method for automated gathering of an RF dataset in the WiFi-6E frequency range, by controlling a portable Vector Network Analyser (VNA) and steerable TX horn antenna;
- we present an RF dataset obtained in an anechoic environment, capturing the resulting signals from human reflections for five different postures.

To the best of our knowledge, this is the first work to explore HPR within the WiFi-6E range, and the dataset has been made publicly available[1].

The rest of this paper is organized as follows. Section II briefly presents related work on HPR. Section III introduces the SpecRF-Posture system, covering the software package for automatic VNA runs, the system architecture and the CNN based posture classification. Section IV presents experimental details, including the experimental setup, dataset analysis, and posture classification evaluation. Section V concludes the paper.

The work was supported by the European Union's Horizon Europe Research and Innovation Program (Project Converge—Telecommunications and Computer Vision Convergence Tools for Research Infrastructures) under Grant 101094831.

[1]M. Oliveira, F. Ribeiro, N. Paulino, O. Yurduseven and L. Pessoa, "SpecRF-Posture Dataset". Zenodo, Apr. 03, 2024. doi: 10.5281/zenodo.10911412.

Authorized licensed use limited to: b-on: UNIVERSIDADE DO PORTO. Downloaded on August 22,2024 at 09:30:04 UTC from IEEE Xplore. Restrictions apply.

II. Related Work

Several innovative approaches leveraging RF signals have been proposed in the literature for wireless posture recognition, each offering distinctive methodologies. For instance, in [5], the use of Frequency-Modulated (FM) signals is explored to detect the posture of an individual. Additionally, [6] explores the utilization of Received Signal Strength Indicator (RSSI) collected by RFID tags to identify six sleeping postures. Furthermore, [7] describes a system using microwave Multiple-Input-Multiple-Output (MIMO) arrays to detect whether a person is standing, sitting, or lying on their back. Finally, [8] discusses a WiFi-based posture recognition system leveraging Channel State Information (CSI). These methods demonstrate the versatility of wireless technologies in capturing human posture, but often require specific and complex hardware setups. This motivates exploring a simpler alternative approach utilizing S_{21} parameters with a Vector Network Analyser (VNA), providing flexibility to conduct research in this topic. Previous works, such as [9] employing semi-supervised learning and [10] utilizing deep CNN, have used these parameters for human activity recognition tasks, achieving high accuracies of $95\,\%$ and $98.8\,\%$, respectively. Similarly, [11] achieved $98.67\,\%$ accuracy for hand movement classification using a deep CNN approach. Remarkably, our approach adopts a simple CNN model and achieves a competitive accuracy of $99.17\,\%$, comparable to existing methods in the literature.

III. SpecRF-Posture

Fig. 1 illustrates our process of data acquisition, training, and HPR. The setup consists of a volunteer within the SoI, the steerable horn antenna, and the reflective surface, i.e, a metal plate. The process of acquiring specular reflections from the SoI through beamsteering is as follows:

*a) **Data Collection**:* Data is collected in the WiFi-6E band, between $5.925\,\mathrm{GHz}$ and $6.5\,\mathrm{GHz}$. A computer drives the steering of the horn antenna along the y-axis, specifically within the range of $\pm 0.65\,\mathrm{m}$ for each frequency. For each steering direction, the computer triggers acquisition of the S_{21} parameters via a Keysight Fieldfox N9914A portable VNA. This process is further explained in Section III-A.

*b) **Data Processing and Training**:* The collected data undergoes processing to extract phase and magnitude information from the parameters. Subsequently, these parameters serve as input for a CNN that is trained to classify the posture of the individual. The architecture of the CNN model is explained in detail in Section III-C.

*c) **Posture Classification**:* Five different postures were considered to test the robustness and sensitivity of the system for detecting differences in the human posture.

A. Automatic Data Acquisition

To streamline data acquisition, a system was developed to synchronously rotate the antenna to a set angular position, and trigger the VNA to get the corresponding S_{21} measurements. The system includes the antenna position controller and the other being the VNA interface. The respective design files and control software are publicly available[2].

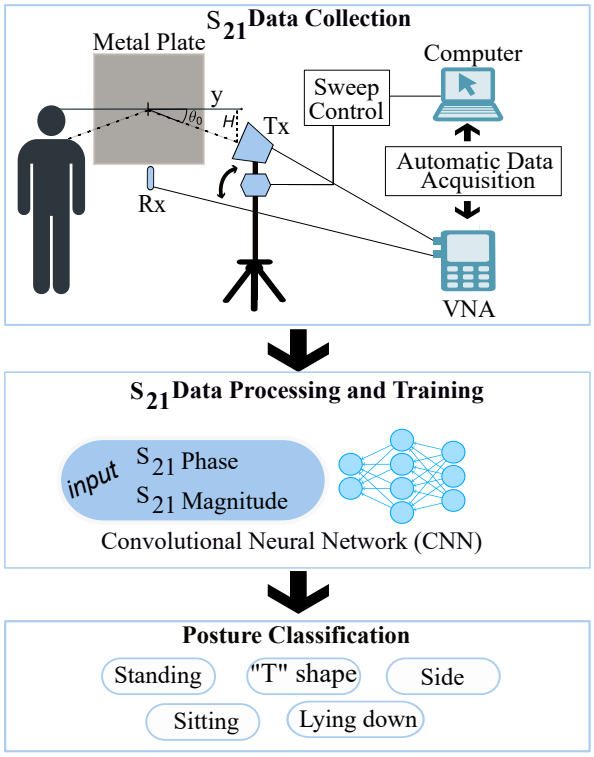


Fig. 1. System architecture diagram illustrating the three main components: Data Acquisition, Data Processing and Training, and Posture Classification.

The antenna position is controlled through an Arduino Uno and a stepper motor. The computer sends a command to the Arduino serial port containing the reference angle at which the motor should be placed and waits for the movement to be completed which is signaled by a command sent from the Arduino. The VNA is controlled through a Python class extending the Virtual Instrument Software Architecture (VISA). This control includes establishing a connection, selecting the parameters to measure, setting the frequency sweep range, controlling the pre-processing of data, and setting the data storage format.

A Python script that performs the acquisition connects the computer to the VNA and Arduino. The user is queried to obtain the posture that will be observed in the following run and how many angular sweeps of the antenna to perform. The main loop is responsible for iterating over a list of angles that result from the analysis in Section III-B. At each step, the TX is steered to the next angle, a measurement request is sent to the VNA, and the data is received.

B. Modeling the Experimental Setup

To better configure the experimental setup, we modeled the radiation emitted by the Tx antenna and its interaction with the environment. This section presents the modeling approach,

[2]F. Ribeiro and M. Oliveira, "SpecRF-Posture". GitHub, Apr. 03, 2024. https://github.com/franciscombr/SpecRF-Posture

Authorized licensed use limited to: b-on: UNIVERSIDADE DO PORTO. Downloaded on August 22,2024 at 09:30:04 UTC from IEEE Xplore. Restrictions apply.

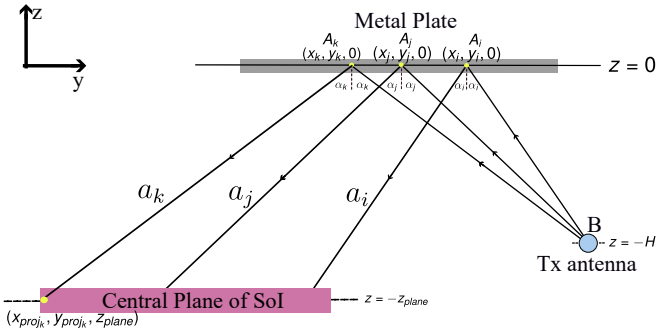


Fig. 2. Top view of the Tx antenna emission towards the metal plate and distortion after reflection.

considering especially the resulting illumination on the plate, and its reflection onto the SoI. This insight allows for a better interpretation of the gathered dataset (Section IV-B).

The radiation emitted by the Tx antenna was modeled by a cos^q radiation pattern, where the power factor q is the same for both E and H planes ($q_E = q_H$). This balanced feed assumption leads to a rotationally symmetric pattern and negligible back radiation [13]. After analyzing the radiation pattern available in the antenna datasheet and comparing it with the cos^q model, the chosen value for q was $q = 5$. Given that the radiation emitted by the Tx antenna is approximated by a cone, radiation distortion occurs after being reflected by the metal plate (Fig. 2).

Consequently, the radiation incident on the central plane of the SoI exhibits a more dispersed distribution. To quantify this phenomenon, the illumination pattern on the metal plate and its reflection on the central plane of the SoI were computed. Equation (1) represents the illumination pattern, denoted as I, at each point (x, y) on the metal plate, calculated using the equations provided in [14].

$$I(x,y) = 20 \cdot \log_{10}\left(\left(\frac{r_{01}^2 + r_1^2 - s_1^2}{2 \cdot r_{01} \cdot r_1}\right)^q\right) \quad (1)$$

$$r_{01} = \sqrt{x_0^2 + y_0^2 + H^2 \cdot \sec^2(\theta_0) + 2 \cdot H \cdot \tan(\theta_0) \cdot y_0} \quad (2)$$

$$r_1 = \sqrt{x^2 + y^2 + H^2 \cdot \sec^2(\theta_0) + 2 \cdot H \cdot \tan(\theta_0) \cdot y} \quad (3)$$

$$s_1 = \sqrt{(x - x_0)^2 + (y - y_0)^2} \quad (4)$$

In (1), $(x_0, y_0, 0)$ are the coordinates of the point where the Tx antenna is emitting towards, H is the distance of the Tx antenna to the y-axis and θ_0 is the angle between the metal plate and the Tx antenna.

The following calculations determine the subsequent illumination pattern in the SoI after reflection with distortion:

$$(x_{proj_i}, y_{proj_i}, z_{plane}) = (x_i, y_i, 0) + a_i \cdot \vec{R_i} \quad (5)$$

$$\vec{R_i} = \vec{A_iB} - 2 \cdot (\vec{A_iB} \cdot \vec{n}) \cdot \vec{n} \quad (6)$$

$$a_i = \frac{z_{plane}}{\cos(\alpha_i)} \quad (7)$$

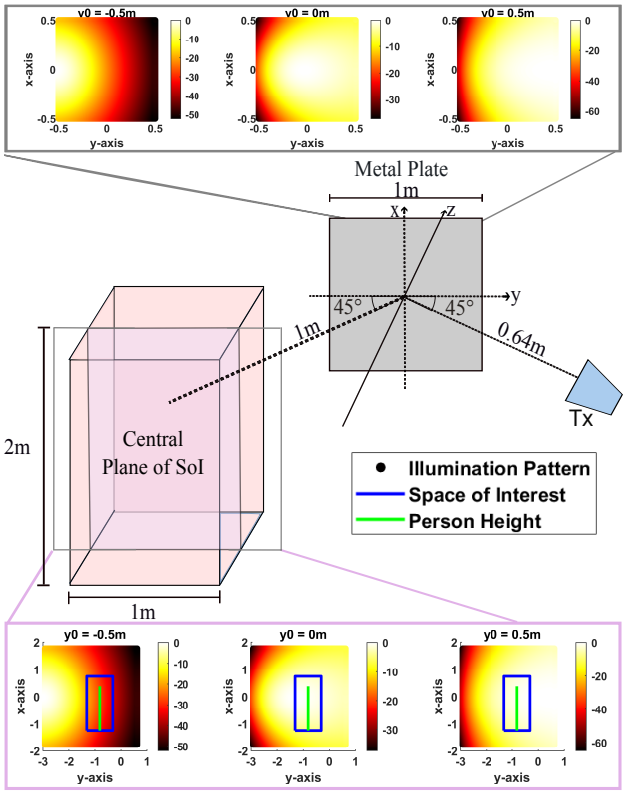


Fig. 3. Schematic representation of the experimental setup illustrating the geometric relationship between setup elements. The figure also displays the illumination patterns for the Tx antenna at different positions ($y_0 = -0.5\,\mathrm{m}$, $0\,\mathrm{m}$, and $0.5\,\mathrm{m}$) on the metal plate, along with the corresponding distorted reflection pattern in the central plane of SoI.

In (6), $\vec{A_iB}$ corresponds to the vector between the Tx antenna represented by $B = (0, H \cdot \tan\theta_0, -H)$ and each point $A_i = (x_i, y_i, 0)$ of the metal plate, with $\vec{n} = (0, 0, -1)$.

Fig. 3 shows the illumination patterns when the Tx antenna points towards $y_0 = -0.5\,\mathrm{m}$, $0\,\mathrm{m}$, and $0.5\,\mathrm{m}$ on the metal plate, and the illumination pattern in the central plane of the SoI after the reflection, with distortion. This schematic representation of the experimental setup illustrates the geometric relationships between the metal plate, Tx antenna, and the space occupied by the individual, aiming to achieve optimal illumination conditions. The angle θ_0 that corresponds to minimal radiation distortion is $45°$, and a distance H of $0.45\,\mathrm{m}$ ($= 0.64/\cos(\theta_0)\mathrm{m}$) provides an edge taper close to -10 dB when the antenna points towards the center of the plate ($y_0 = 0\,\mathrm{m}$). This minimizes diffraction effects on the reflector edges, maintaining optimal radiation performance [13].

C. CNN-Based Posture Classification

Convolutional Neural Networks (CNNs) have been widely recognized as an effective choice for pattern recognition tasks in images and multidimensional signals [15]. Therefore, the adoption of a CNN for human posture classification was driven by its capability to effectively learn and extract meaningful features from the data, demonstrated by various recent works

Authorized licensed use limited to: b-on: UNIVERSIDADE DO PORTO. Downloaded on August 22,2024 at 09:30:04 UTC from IEEE Xplore. Restrictions apply.

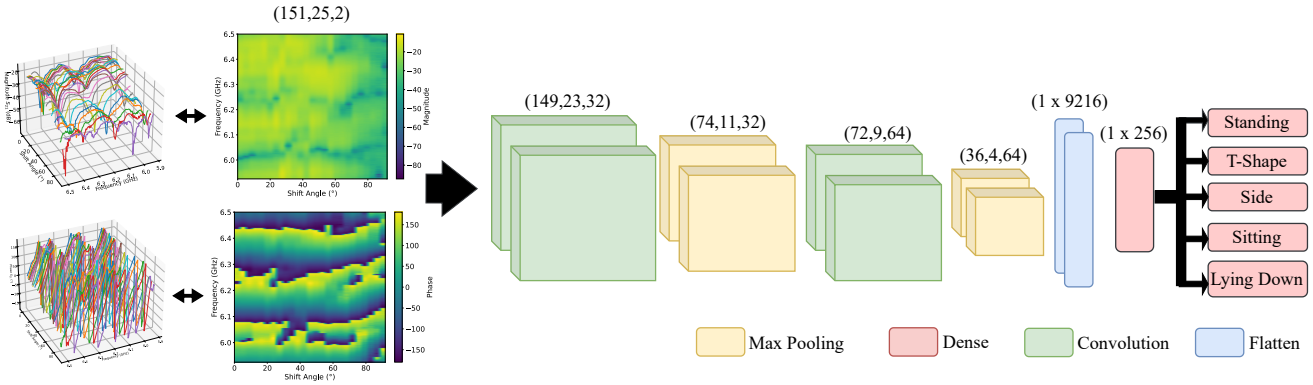


Fig. 4. Architecture of the CNN developed for posture classification (Standing, T-Shape, Side, Sitting, Lying Down). The input to the network consists of the magnitude and phase of the S_{21} parameters, which can be represented both as a 3D plot and as a frequency-angle image.

that utilize CNNs for different classification tasks using radio signals as images [16]–[20].

The architecture of the developed CNN, depicted in Fig. 4, comprises multiple layers, including convolutional layers, max-pooling layers, and dense layers. The input layer of the network receives a 2D feature, specifically angle-frequency images representing the magnitude and phase of the S_{21} parameters in two separate channels, while the output layer consists of five nodes, each representing one of the five postures. The parameters used for the training of the CNN are detailed in Table I.

TABLE I
HYPERPARAMETER CONFIGURATIONS OF CNN FOR SPECRF BASED HUMAN POSTURE RECOGNITION (HPR)

Parameter	Value
Learning rate	0.00001
Training epochs	200
Optimizer	Adam
Loss Function	Categorical Crossentropy

IV. EXPERIMENTAL EVALUATION

A. Experimental Setup

The experimental setup was assembled in an anechoic chamber, with a metal plate measuring $1\,\mathrm{m} \times 1\,\mathrm{m}$ ($1\,\mathrm{m}^2$). Fig. 5 illustrates the experimental setup from the perspective of the SoI. A horn antenna (A-INFO LB-20180) was employed as the transmitter (Tx), while an omnidirectional antenna served as the receiver (Rx), allowing for signal detection from various directions within the SoI.

A column of absorbers was placed to mitigate the Line-of-Sight (LoS) component between the Tx and Rx antennas, as shown in Fig. 5. The receiving antenna is positioned immediately below the metal plate, with a $10\,\mathrm{cm}$ offset from the center of the plate to capture richer signals from the SoI. Additionally, two Low Noise Amplifiers (ZX60-153LN-S+) were connected to the input and output ports of the Tx and

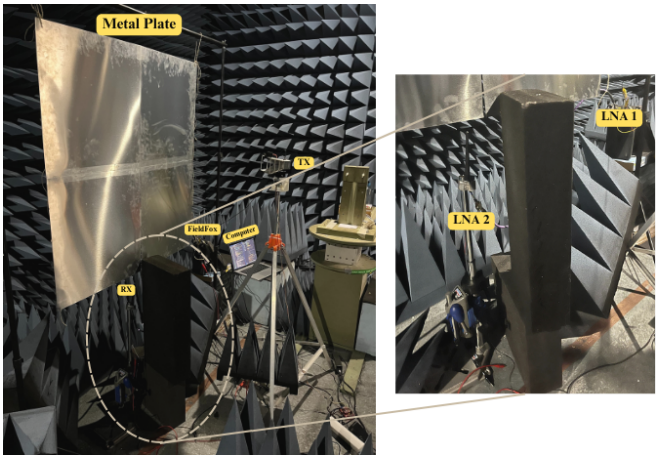


Fig. 5. Experimental set up inside the anechoic chamber.

Rx to amplify the transmitted and received RF signals, respectively, to compensate signal attenuation due to absorption and reflections from the person in the SoI. The computer and VNA were located outside the experimental area and were used to acquire the S_{21} parameters and perform the Tx sweep.

During the experiment, a single individual assumed various postures within the SoI while the antenna rotated in order to be directed towards points along the y-axis on the metal plate between $(0, -0.65\,\mathrm{m}, 0)$ to $(0, 0.65\,\mathrm{m}, 0)$, with a spacing of $0.05\,\mathrm{m}$. This results in diverse illumination scenarios within the SoI, as depicted in the illumination pattern plots shown in Fig. 3. This spacing is constrained by the minimum resolution of the sweep control unit ($1.8°$) but is also in proximity to the minimum wavelength (at the maximum frequency) of $0.046\,\mathrm{m}$ for the operating frequency band. Each y_0 position along the y-axis corresponds to a specific angle between the Tx antenna and this axis. Notably, due to the offset feeding of the antenna, the resulting illumination pattern exhibits a distorted shape resembling an ellipse. At each of these angles, the S_{21} parameters within the frequency range of $5.925\,\mathrm{GHz}$ to $6.5\,\mathrm{GHz}$ were captured using a VNA.

Authorized licensed use limited to: b-on: UNIVERSIDADE DO PORTO. Downloaded on August 22,2024 at 09:30:04 UTC from IEEE Xplore. Restrictions apply.

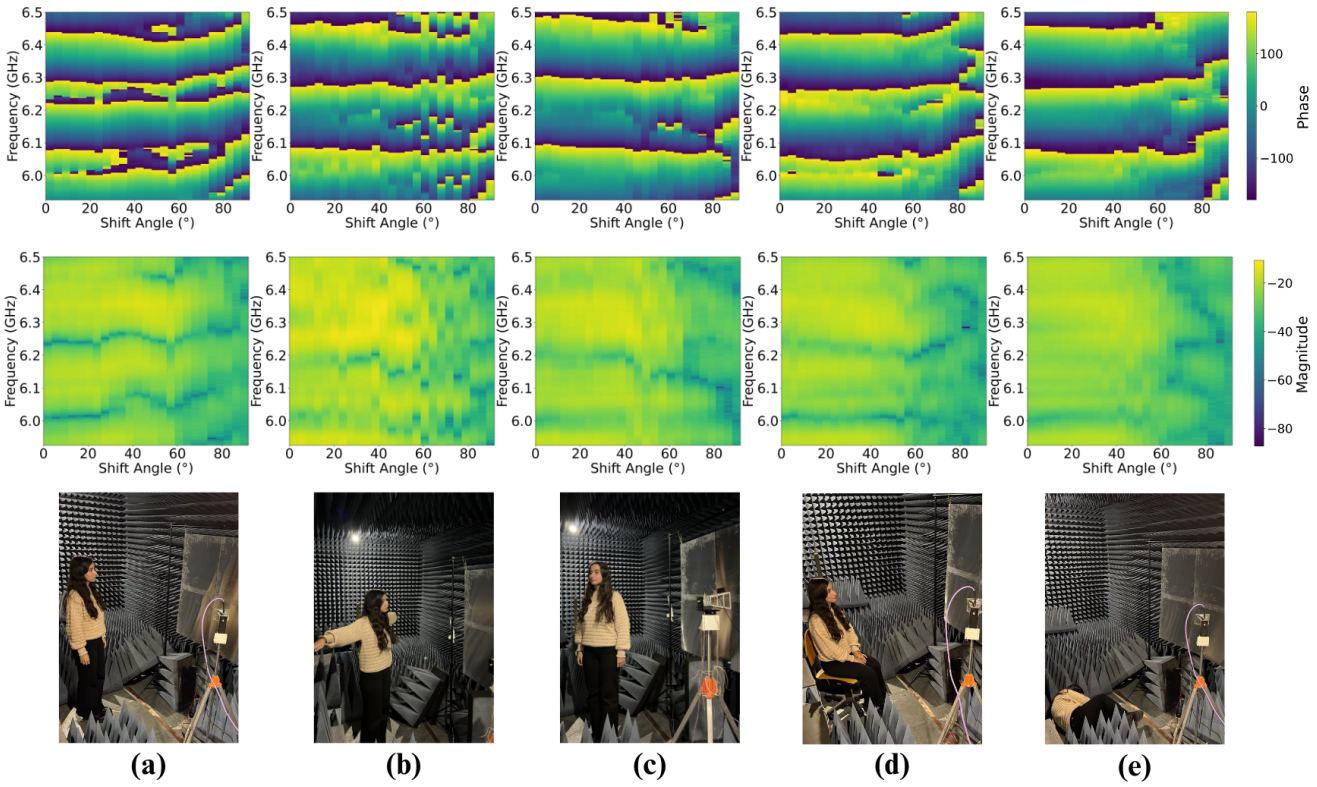


Fig. 6. Frequency vs. Angle Analysis for S_{21} Phase (Top), Magnitude (Middle), and Posture Identification (Bottom) for: (a) Standing, (b) T-Shape, (c) Side, (d) Sitting, and (e) Lying Down.

B. Dataset Description and Analysis

For each posture (Standing, T-shape, Side, Sitting, Lying down), we collected 120 samples (referred to as *sweeps*) from one volunteer, resulting in a dataset of 600 samples (1 subject $\times$ 5 postures $\times$ 120 sweeps). Each sample is characterized by dimensions $(n_{\text{resolution}}, n_{\text{angles}})$, where $n_{\text{resolution}} = 151$ represents the number of frequencies between $5.925\,\text{GHz}$ and $6.5\,\text{GHz}$, defining the resolution of the frequency range, and $n_{\text{angles}} = 25$ denotes the number of angles through which the antenna rotates. This spectrum of postures included minor movements of the hands or head to introduce variability into the data and make the model robust to different ways of assuming these postures. Finally, each *sweep* required 35 seconds for data collection, resulting in a total time of approximately 6 hours to create the dataset.

Fig. 6 presents an example illustrating the variation in phase and magnitude of the S_{21} parameters concerning rotation angle and frequency within the WiFi-6E range for each posture to be classified. The experiment involved a scanning range broader than the physical dimensions of the metal plate, extending from approximately $-0.65\,\text{m}$ to $0.65\,\text{m}$, to cover areas where the SoI was not adequately illuminated. Consequently, in the magnitude analysis, lower angles exhibit higher intensity due to the omnidirectional antenna capturing information from the illuminated space of interest by the Tx antenna, while higher angles result in less intensity as the antenna captures information from regions not adequately illuminated by the Tx antenna. Similarly, in the phase analysis, larger angles show greater variation because of inadequate illumination of the metal plate, resulting in irrelevant reflections captured by the Rx antenna. In conclusion, this broader scanning range enabled observation of the phase and magnitude data to determine the angle up to which the region where the person was located could be accurately characterized. As a result, the SoI is accurately characterized up to an angle of approximately 50 degrees, corresponding to $y_0 = 0.45\text{m}$.

Observations within the region that effectively characterize the SoI reveal distinctive profiles of magnitude and phase for each posture. This indicates that the received signal captured by the Rx exhibits variations corresponding to the human body's posture. Consequently, identical spaces of interest demonstrate similar profiles of magnitude and phase in the S_{21} parameters, with the reciprocity principle applying as well.

C. Posture Classification Evaluation

In this subsection, we evaluate the accuracy of the CNN model described in Section III-C. The dataset was initially split into $80\,\%$ for training and $20\,\%$ for testing. Then, the data was processed by extracting magnitude and phase information

Authorized licensed use limited to: b-on: UNIVERSIDADE DO PORTO. Downloaded on August 22,2024 at 09:30:04 UTC from IEEE Xplore. Restrictions apply.

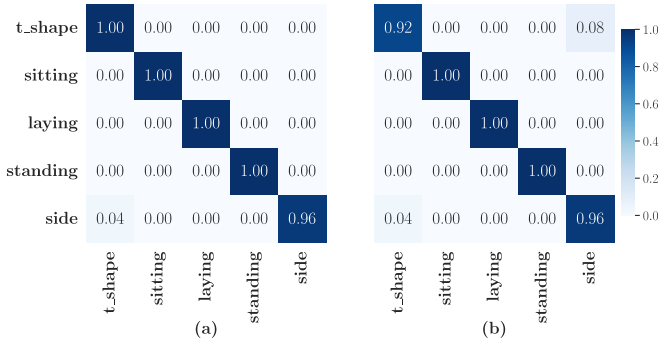


Fig. 7. Posture Recognition accuracy utilizing two distinct approaches based on the S_{21} parameters: (a) incorporation of both magnitude and phase as input parameters for the CNN model, and (b) utilization of only magnitude as the input.

from the raw data, which were then combined to form image-like data for input into the CNN model.

We compared two approaches: the first utilizes two channels, where one channel represents the magnitude and the other represents the phase of the S_{21} parameters, while the second approach uses only one channel representing magnitude. The classification results are presented in Fig. 7. SpecRF-Posture achieves an average accuracy in the validation data of $99.17\,\%$ for the first method. The individual accuracy rates are $100\,\%$ for all postures, with exception of *Side*, for which it is $96\,\%$. In contrast, the second approach yielded a slight decrease in accuracy specifically for the *T-shape* posture, with the accuracy declining to $96\,\%$. Hence, the use of phase information results in a more precise differentiation between the *Side* and *T-shape* postures.

V. Conclusions

In this paper, we present a system with a simple hardware architecture, leveraging a metal plate to passively manipulate Radio Frequency (RF) signals through specular reflections. This passive manipulation of the WiFi-6E signal simplifies the Human Posture Recognition (HPR) process by providing semi-controllable paths for RF signals without the need for complex hardware setups. This enables the use of passive scatters within the medium for similar tasks at a low hardware cost.

We illustrate systematic data gathering, driven by our hardware setup and control software. By covering the entire metal plate at regular intervals, our system captures significant data for precise posture recognition cost-effectively, demonstrating the viability of HPR via S_{21} parameters of signals reflected in the SoI, with potential for broader applications in wireless sensing and monitoring.

Given the classification results, we will gather richer datasets of specular reflections in future work, by including additional volunteers and postures, as well as address the use of non-specular reflections by resorting to a Reconfigurable Intelligent Surface (RIS), and measuring the potential performance increase versus increased system complexity and cost.

References

[1] Shu, Yangxia & Hu, Lei. (2023). A Vision-based Human Posture Detection Approach for Smart Home Applications. International Journal of Advanced Computer Science and Applications. 14. 10.14569/IJACSA.2023.0141023.

[2] Lee, Min-Fan & Chen, Yen-Chun & Tsai, Cheng-Yo. (2022). Deep Learning-Based Human Body Posture Recognition and Tracking for Unmanned Aerial Vehicles. Processes. 10. 2295. 10.3390/pr10112295.

[3] Cheng, Lingxiao. (2024). Athletes' Action Recognition and Posture Estimation Algorithm Based on Image Processing Technology. Journal of Electrical Systems. 20. 2059-2069. 10.52783/jes.3121.

[4] R. Shahbazian and I. Trubitsyna, "Human Sensing by Using Radio Frequency Signals: A Survey on Occupancy and Activity Detection," in IEEE Access, vol. 11, pp. 40878-40904, 2023, doi: 10.1109/ACCESS.2023.3269843.

[5] Pereira, João & Casmin, Eugene & Oliveira, Rodolfo. (2023). A Passive RF Testbed for Human Posture Classification in FM Radio Bands. Sensors. 23. 9563. 10.3390/s23239563.

[6] Liu, Jia & Chen, Xingyu & Chen, Shigang & Liu, Xiulong & Wang, Yanyan & Chen, Lijun. (2019). TagSheet: Sleeping Posture Recognition with an Unobtrusive Passive Tag Matrix. 874-882. 10.1109/INFOCOM.2019.8737599.

[7] Sasakawa, Dai & Honma, Naoki & Nakayama, Takeshi & Iizuka, Shoichi. (2018). Human Posture Identification Using a MIMO Array. Electronics. 7. 37. 10.3390/electronics7030037.

[8] Zuo, Jiancun & Zhu, Xianxun & Peng, Yue & Zhao, Zhiyang & Wei, Xiong & Wang, Xun. (2021). A New Method of Posture Recognition Based on WiFi Signal. IEEE Communications Letters. PP. 1-1. 10.1109/LCOMM.2021.3081135.

[9] Y. Zhang, Y. Shao, L. Xiong and J. Zhang, "Human Activity Recognition Based on On-Body Antennas and Semi-Supervised Learning," 2023 17th European Conference on Antennas and Propagation (EuCAP), Florence, Italy, 2023, pp. 1-5, doi: 10.23919/EuCAP57121.2023.10133047.

[10] Y. Kim and Y. Li, "Human Activity Classification With Transmission and Reflection Coefficients of On-Body Antennas Through Deep Convolutional Neural Networks," in IEEE Transactions on Antennas and Propagation, vol. 65, no. 5, pp. 2764-2768, May 2017, doi: 10.1109/TAP.2017.2677918.

[11] hak gupta, Sindhu & Sharma, Aayush & Mohta, Mohit & Rajawat, Asmita. (2019). Hand Movement Classification from Measured Scattering Parameters using Deep Convolutional Neural Network. Measurement. 151. 107258. 10.1016/j.measurement.2019.107258.

[12] M. Eddy. RF interference and the role of filters in Wi-Fi 6E designs. planetanalog.com. Accessed: April. 3, 2024. [Online.] https://www.planetanalog.com/rf-interference-and-the-role-of-filters-in-wi-fi-6e-designs/

[13] Nayeri, Payam & Elsherbeni, Atef. (2018). Reflectarray Antennas: Theory, Designs, and Applications. 10.1002/9781118846728.

[14] Yu, Ang & Elsherbeni, Atef& Huang, John & Rahmat-Samii, Yahya. (2010). Aperture efficiency analysis of reflectarray antennas. Microwave and Optical Technology Letters. 52. 364 - 372. 10.1002/mop.24949.

[15] Deep Learning (Ian J. Goodfellow, Yoshua Bengio and Aaron Courville), MIT Press, 2016, [Online.] http://www.deeplearningbook.org

[16] R. Sharma, B. Deka, V. Fusco, and O. Yurduseven, "Integrated convolutional neural networks for joint super-resolution and classification of radar images," Pattern Recognition, Elsevier, vol. 150, pp. 110351. 2024, doi: 10.1016/j.patcog.2024.110351.

[17] Roh, Hyeji & Oh, Sheungmin & Song, Hajun & Han, Jinseo & Lim, Sangsoon. (2022). Deep Learning-based Wireless Signal Classification in the IoT Environment. Computers, Materials & Continua. 71. 5717-5732. 10.32604/cmc.2022.024135.

[18] Zhou Y, Xu C, Zhao L, Zhu A, Hu F, Li Y. CSI-Former: Pay More Attention to Pose Estimation with WiFi. Entropy (Basel). 2022 Dec 22;25(1):20. doi: 10.3390/e25010020.

[19] X. Wang, X. Wang and S. Mao, "Deep Convolutional Neural Networks for Indoor Localization with CSI Images," in IEEE Transactions on Network Science and Engineering, vol. 7, no. 1, pp. 316-327, 1 Jan.-March 2020, doi: 10.1109/TNSE.2018.2871165.

[20] M. Zhang, M. Diao and L. Guo, "Convolutional Neural Networks for Automatic Cognitive Radio Waveform Recognition," in IEEE Access, vol. 5, pp. 11074-11082, 2017, doi: 10.1109/ACCESS.2017.2716191.

Authorized licensed use limited to: b-on: UNIVERSIDADE DO PORTO. Downloaded on August 22,2024 at 09:30:04 UTC from IEEE Xplore. Restrictions apply.