Radar Based Classification/Regression Head Position Estimation

Presentation

Marcus Fong, Joseph Scott

20962575@student.curtin.edu.au
20558042@student.curtin.edu.au

School of EECMS
Curtin University

March 18, 2025

Table of Contents

1 Introduction
- Background

2 Hardware Outline
- Experimental Setup

3 Data Pre-Processing

4 Machine Learning

5 Future Improvements

Acknowledgement to Country (Boodja)

Acknowledgement

I would like to acknowledge the Wadjuk Noongar people as the traditional custodians of the land on which the Bentley campus is located, and would like to pay respect to elders past, present and emerging.

Curtin University

Project Goals

- Optimising head pose estimation of a subject in controlled indoor spaces
  - A sensing/imaging method that transmits sufficient information without raising privacy concerns will have to be chosen
  - Extracting head pose information (e.g head movement direction and angles) will have to be done using deep learning and transfer learning techniques

Background

[1] used a single XeThru UWB Radar Module to record head movements. The radar data was inputted into a classification ML model which was able to classify movements into six different positions with an accuracy of around 80%.

- The goal of this project will be to optimise the method used in [1] to accurately estimate the pitch, roll, and yaw of head movements recorded using UWB Radar.

Hardware Outline

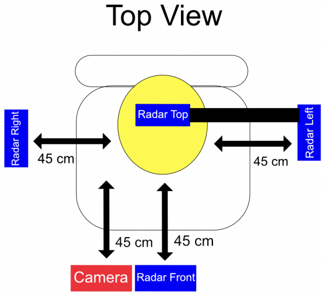
4x XeThru X4M03 UWB Radar Modules

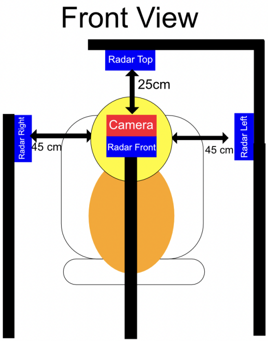
- Each radar module was purchased from an online seller at $550 USD per piece
- Four radars are planned to be used in the final regression model setup

1x ZED 2 Stereo Camera

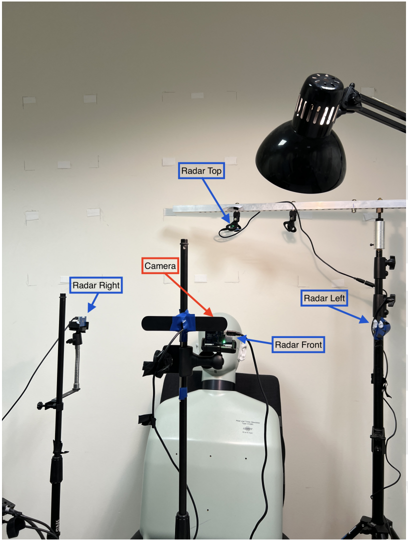
- Camera provided from Curtin Acoustics Lab building (retail price $500 USD per piece
- Video and radar data will be simultaneously recorded, with the video used as ground truth data to train the regression model
- Could potentially be placed with a cheaper/higher resolution standalone digital camera to improve ground truth data

Setup Schematic





Setup Photo (Front View)



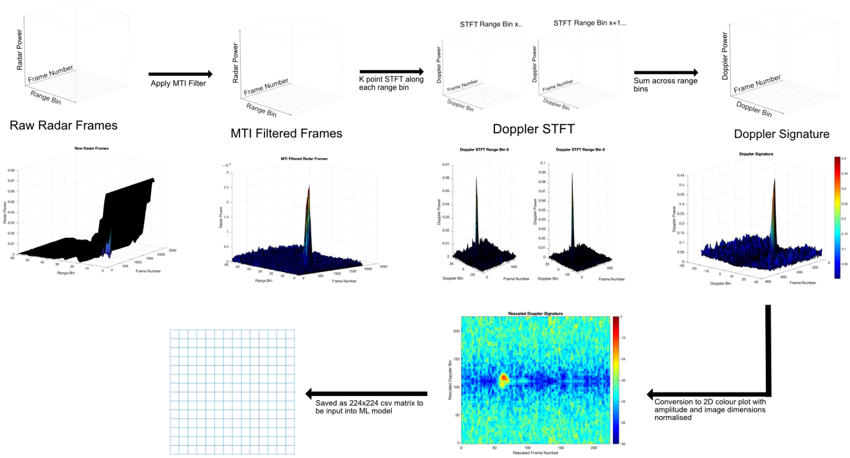
Raw Radar Frames and Specifications

Using MATLAB, radar data is recorded and saved as in-phase quadrature raw radar frames. Radar is recorded at 60 fps for two seconds. Four radars are placed in specific positions:

- in front of the subject
- to the left and right of the subject
- above the subject

The front, left, and right radars are positioned 45cm away from the subject, and the top radar is positioned 25cm away from the subject.

Radar Data Processing Diagram



Radar Data Processing

Raw data is processed into Doppler Signature to measure how the distance of the subject from the radar changes over time. The Doppler frequency shift caused by a target moving away from the radar with velocity v and angle θ is given by: $\Delta f = \frac{2 \cdot f_c \cdot v \cdot cos\theta}{c}$, where f_c is the centre frequency of the radar ($f_c = 7.92GHz$), and c is the speed of light.

- Radar data has been processed into Doppler information to highlight specific features from head movements:
  - if a target is moving towards or away from the radar ($\pm$ Doppler)
  - The speed at which the target is moving (Frequency)
  - The strength at which the target is moving (Power)

Radar Data Processing

The maximum Doppler frequency that can be detected by the radar is $= \pm\frac{fps}{2}Hz$, where fps is the radar recording fps.

- For a radar recording at 60fps, the maximum doppler frequency that can be detected is $\pm 30Hz$, which corresponds to a maximum movement speed of $v_{max} = \frac{30 \cdot 299792458}{2 \cdot 7.29 \cdot 10^9} = 61.6cm/s$.

The resolution of Doppler frequency measurements is given by $\frac{fps}{K}$, where K is the number of points in the Doppler STFT.

- For a K=56 point STFT, the resolution of each doppler-bin is $\frac{60}{56} =$ 1.07Hz. A resolution of 1.07Hz corresponds to a minimum measurable movement speed of $v_{min} = \frac{1.07 \cdot 299792458}{2 \cdot 7.29 \cdot 10^9} = 2.2cm/s$.

Radar Data Processing

From [2], the predominant frequency for pitch, yaw and roll head movements is 2.6-3.2Hz, with signficant harmonics occurring up to 15-20Hz. The min. frequency resolution ($1.07Hz$) and max. frequency ($\pm 30Hz$) of our setup allows us to measure the full spectrum of head movement speed frequencies. Headroom has also been given for the measurement of extreme slow or fast head movement cases.

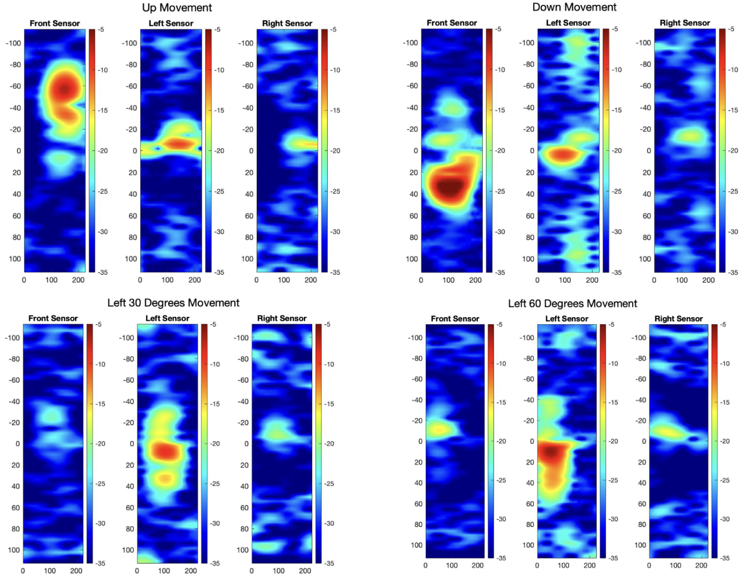
Curtin University

Radar Data Processing

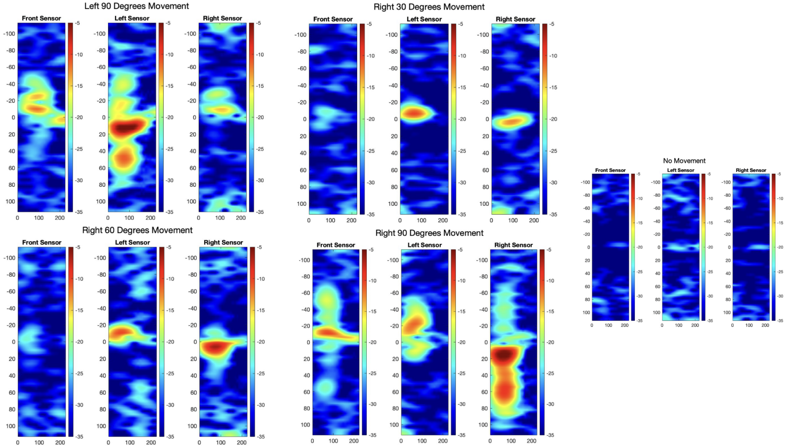
Head position movements can be correlated to different levels of Doppler frequency shift:

- A positive Doppler frequency shift indicates the subject is moving towards the sensor (e.g if a radar is placed in front of you and you move your head down, the centre of mass of your head moves towards the radar and causes a positive Doppler shift)
- A negative Doppler frequency shift indicates the subject is moving towards the sensor (e.g if a radar is placed in front of you and you move your head up, the centre of mass of your head moves away from the radar and causes a negative Doppler shift)

Radar Data Spectrograms



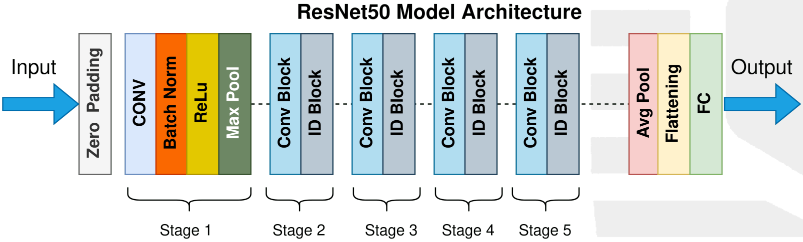
Radar Data Spectrograms



Curtin University

Model Overview

A deep neural network (ResNet) consisting of 50 layers, with other features like skip connections. This is used as the backbone, then some small fully connected layers are used as the head. Although this model is very deep, the inference time (time to classify) is nearly instant, even without a GPU.



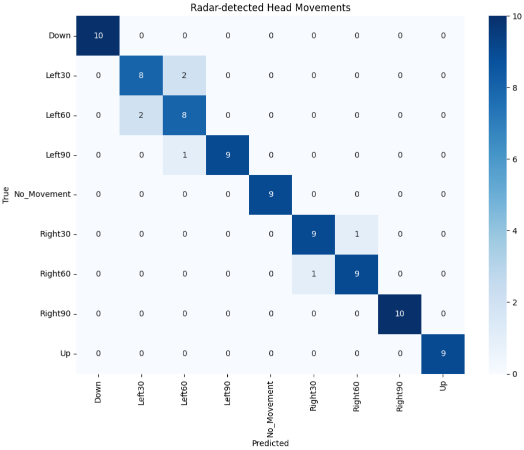
Method

1. Stack spectrograms from the 3 sensors together, similar to how RGB images are made up of 3 matrices of red, green and blue channels.
2. Pass this 3D spectrogram into our image an classification model (ResNet50).
3. Get probabilities that the spectograms fall into each class, with each class being a head movement.

The total dataset consists of 700 3D spectograms

Confusion Matrix

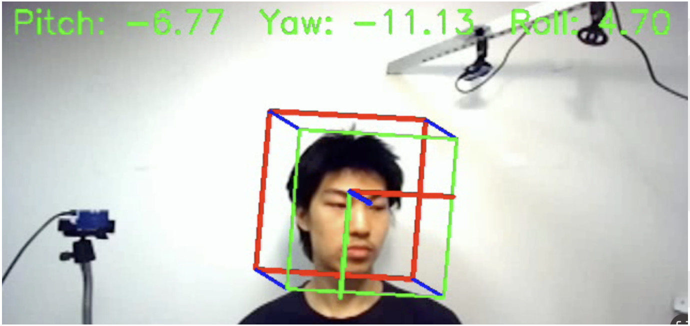
Final Test Accuracy: 92%



.

Regression Model

Using LSTM and Transformer regression model to estimate continuous pitch, yaw, and roll angles from radar. We'll use video footage like the one below as our ground truth for training.



Curtin University

Future Improvements

Working on simultaneously recording radar data and video recordings so the precise pitch, yaw, and roll head movements picked up by the video recording can be used to train the radar

References

[1] H Hameed et al. "Wi-Fi and radar fusion for head movement sensing through walls leveraging deep learning". In: **IEEE Sensors Journal** 24.9 (2024), pp. 14952–14961.

[2] G E Grossman et al. "Frequency and velocity of rotational head perturbations during locomotion". In: **Experimental Brain Research** 70.3 (1988), pp. 470–476.

[3] Kalyanaraman Avinash et al. "FormaTrack: Tracking People based on Body Shape". In: **Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.** 1.3 (2017), pp. 1–21.

[4] Gon Woo Kim et al. "A Study on 3D Human Pose Estimation Using Through-Wall IR-UWB Radar and Transformer". In: **IEEE Access** 11.99 (2023), pp. 15082–15095.

[5] Xiaolong Zhou et al. "Three-Dimensional Human Pose Estimation from Micro-Doppler Signature Based on SISO UWB Radar". In: **Remote Sensing** 16.7 (2024), pp. 1295–1316.

[6] Abolfazl Farahani et al. "A Brief Review of Domain Adaptation". In: **Advances in Data Science and Information Engineering**. Cham: Springer International Publishing, 2021, pp. 877–894.

Thank you!

Contact:

20962575@student.curtin.edu.au
20558042@student.curtin.edu.au

Curtin University