Capturing Facial Expressions using COTS mmWave Radar

PROJECT - PART 1 (CS67101) report submitted to

Indian Institute of Technology Kharagpur

in partial fulfilment for the award of the degree of

Master of Technology

in

Computer Science and Engineering

by

Avijit Mandal

(18CS30010)

Under the supervision of

Professor Sandip Chakraborty

Computer Science and Engineering

Indian Institute of Technology Kharagpur

Autumn Semester, 2022-23

November, 2022

DECLARATION

I certify that

(a) The work contained in this report has been done by me under the guidance of my supervisor.

(b) The work has not been submitted to any other Institute for any degree or diploma.

(c) I have conformed to the norms and guidelines given in the Ethical Code of Conduct of the Institute.

(d) Whenever I have used materials (data, theoretical analysis, figures, and text) from other sources, I have given due credit to them by citing them in the text of the thesis and giving their details in the references. Further, I have taken permission from the copyright owners of the sources, whenever necessary.

Date: November, 2022 (Avijit Mandal)
Place: Kharagpur (18CS30010)

COMPUTER SCIENCE AND ENGINEERING

INDIAN INSTITUTE OF TECHNOLOGY KHARAGPUR

KHARAGPUR - 721302, INDIA

CERTIFICATE

This is to certify that the project report entitled "**Capturing Facial Expressions using COTS mmWave Radar**" submitted by **Avijit Mandal** (Roll No. 18CS30010) to Indian Institute of Technology Kharagpur towards partial fulfilment of requirements for the award of degree of Master of Technology in Computer Science and Engineering is a record of bona fide work carried out by him under my supervision and guidance during Autumn Semester, 2022-23.

Date: November, 2022
Place: Kharagpur

Professor Sandip Chakraborty
Computer Science and Engineering
Indian Institute of Technology Kharagpur
Kharagpur - 721302, India

Acknowledgements

I would like to thank my supervisor **Prof. Sandip Chakraborty** to give me an opportunity to work under his guidance, who gave me constructive advice throughout the course of my M.Tech project.

I would like to thank **Mr. Argha Sen**, Ph.D. scholar who helped me in this project whenever I was stuck. It would not have been possible to do this project without him.

I would like to thank every faculty member of the Department of Computer Science and Engineering, IIT Kharagpur, for guiding us in every way possible during the last year.

Finally, I would like to thank my family and friends for their constant love and support.

Contents

Abbreviations

ADC	ANALOG TO DIGITAL CONVERTER
COTS	COMMERCIAL OFF THE SHELF
FA	FACIAL AUTHENTICATION
mmWave	MILLIMETER WAVE

Chapter 1

Introduction

1.1 Importance of facial expression detection

In education, marketing, and advertising, measuring audience response and documenting user experience has shown to be extremely helpful and practical. These comments are typically gathered through crowd-sourcing services like IMDB, Nielsen TV, Rotten Tomatoes, etc. Even if such evaluations and ratings, considered overall, can typically reflect the public's general opinion, they are less effective in revealing

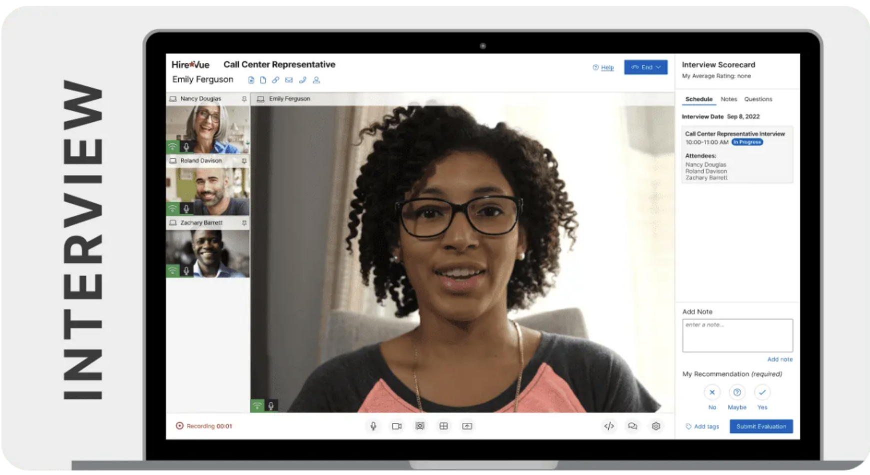


FIGURE 1.1: HireVue interview platform

specific, immediate, or fleeting emotional reactions that occur throughout the movie. Because first-hand experiences cannot be quickly recovered or replicated, gathering and assessing those fine-grained emotional responses necessitates active audience participation in real-time. Medical professionals are now utilizing facial expression recognition to help treat mental and behavioral illnesses. Micro-expressions naturally reveal a fleeting, unobtrusive mood, yet they frequently conceal vital information. The expression recognition system can collect and identify these minute changes. With the aid of this technology, medical professionals may quickly develop treatment plans by assessing the emotional and mental health of kids with autism and depression. With the increasing setup for virtual coding test and interviews during the covid-era, many hiring platforms such as HireVue [1] uses facial expression detection techniques to detect uninterested candidates, real-time feedback about coding test, etc.

1.2 Broad Idea

In this study, we contribute to developing a mmWave-based facial expression recognition system. Firstly, we generate mmWave signals and capture reflected signals (77 GHz - 81 GHz) using Texas Instruments (TI) mmWave frontend radar (IWR1642BOOST). The captured data gets transferred to the host machine via the DCA1000EVM board for further analysis to recognize the user's facial expression using several pre-processing libraries and, finally, ML/DL models.

[1] `https://www.hirevue.com/`

Chapter 2

Literature Survey

2.1 Modalities of Expression Recognition

There are different modalities for facial expression recognition available. Among them, Camera-based methods for facial expression recognition are widely used. Other popular recognition techniques include Signal based sensing techniques which can be further classified into two predominant techniques, namely, Acoustic-based and Radio Frequency based modalities. In the following section, we have discussed these modalities in detail.

2.1.1 Camera Based Recognition

Camera-based techniques are widely used due to their cheap hardware components availability and easy setup. Also, many people have already worked on image-based facial expression recognition techniques. Due to this, there are mainly available public datasets making the training and testing phase easier. But surveillance cameras may raise privacy concerns, especially in home and office environments. Moreover, camera-based sensing is highly dependent on environmental lighting conditions,

FIGURE 2.1: Camera-based Facial Expression Recognition

making it hard to work under conditions of low illumination, smoke, and opaque obstructions. Also, using masks, sunglasses, and scarves for protection is very common. These facial occlusions limit the capabilities of image-based recognition.

2.1.2 Signal Based

Comparatively speaking to a camera, signal-based sensing is less intrusive and employs fewer ways that violate privacy. Environmental variables such as dim lighting, obstructions (such as masks, clothing, accessories, etc.), smoke, etc., have nothing to do with signal transmission and reflection. Since cameras struggle in these circumstances, signal-based sensing performs better. Additionally, it can provide security in places like restrooms or locker rooms where video recording is prohibited. By identifying people's presence and showing them where they are moving, it can enhance automatic doors with several clever features. Using radar there simultaneously reduces energy consumption. The range of signal-based sensing is a few millimeters to a few hundred meters, and it doesn't require a straight line of sight to function (for example, through drywall or plywood). There are two widely used signal-based sensing, i.e., Acoustic sensing and RF-based sensing.

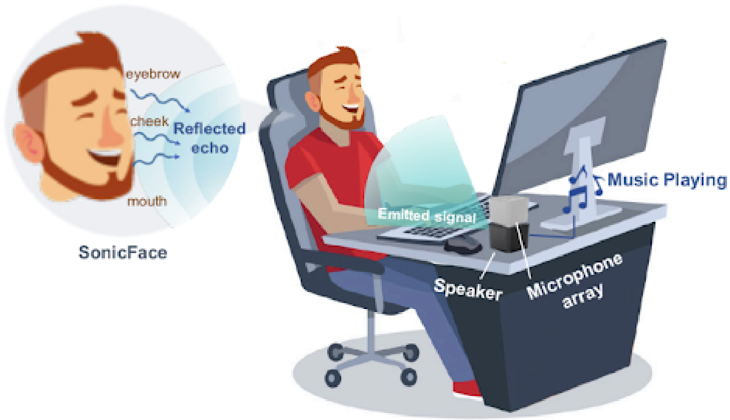


FIGURE 2.2: Signal-based Sensing for Facial Expression Recognition

2.1.2.1 Acoustic Based

Acoustic-based sensing sends a chirp signal in the audible/inaudible ranges and captures the reflected signal. The reflected sound signal is further analyzed to infer doppler shifts for the final sensing approach. Acoustic sensing setup is so accessible that even a commodity smartphone can also be used as a sensing device. Acoustic sensing has a Range Resolution (ΔR)of 4.3 cm, making it good for detecting motion detection.

$$\Delta R = \frac{c}{2B} \tag{2.1}$$

where c is speed of light and B is the bandwidth.

Also, acoustic-based sensing does not have privacy issues with images and is deployable in less favorable conditions. One of the limitations of acoustic sensing is we can not use the phase variation of the reflection signal much due to its high wavelength (λ), as phase variation is inversely proportional to λ. Previously I have worked on acoustic sensing for the detection of facial expressions during my BTP, and the accuracy we achieved keeping the environment controlled and noise-free is still 56.7%.

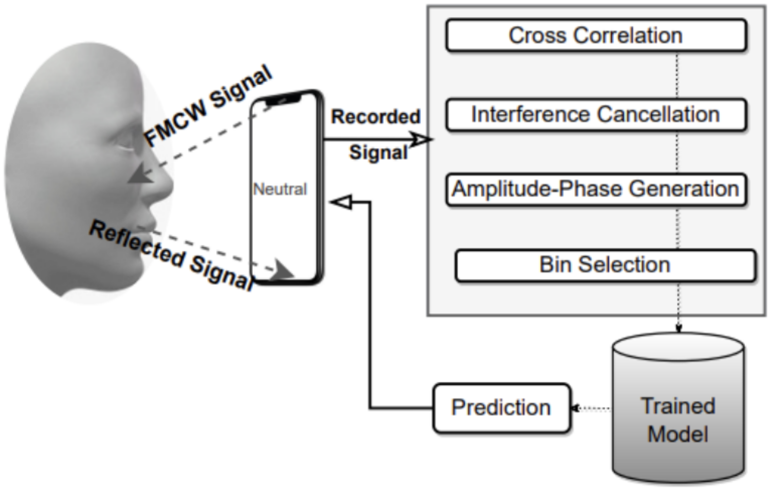


FIGURE 2.3: Acoustic-based Facial Expression Detection

2.1.2.2 RF Based

In RF-based sensing, we have two widely used techniques, i.e., WiFi Sensing and mmWave sensing. WiFi-based sensing suffers from low range resolution (1.8 m) as it has higher bandwidth (80MHz). mmWave sensing has a better range resolution (0.0375m) than Wifi-based sensing due to its lower bandwidth (4 GHz). According to the equation 2.2, the phase variations in mmWave can catch even the slightest differences in the reflector's velocity. The wavelength (λ) is in the range of millimeters (4mm), and the typical phase noise ($\Delta\phi$) is 0.057 deg. Therefore it is possible to track small changes in range (Δr 0.63 um) by taking advantage of the phase variations.

$$\Delta\phi = \frac{2\pi\Delta r}{\lambda} \tag{2.2}$$

2.2 Related Works

2.2.1 Acoustic Sensing and Blink Detection

Problem Statement Many real-life applications use eye blink detection, including Human-Computer Interaction (HCI), drowsy driving prevention, and eye disease detection. Traditional camera-based techniques are promising but have strict lighting conditions and line-of-sight requirements.

Solution Proposed The creators of BlinkListner [2] describes a method that uses acoustic waves to detect minor eye blink movements without requiring touch. In this work, the authors first model the connection between signal change and tiny motions induced by eye blink and interference. The researchers next offer a unique strategy for maximizing the modest signal fluctuation caused by eye blinks by using "harmful" interference.

Results BlinkListener gives a robust performance with an accuracy of 95%.

2.2.2 Robust Human Face Authentication on Smartphones Using Acoustic Sensing

Problem Statement Many essential daily apps, such as social networks, shopping, and banking, rely on smartphone user authentication. But most of the available solutions have their limitations. For example, the user can forget the PIN, while facial recognition can spoofed by video and images. The Iris Scan is the most effective one, but most commodity mobile phones don't have hardware for the same.

Solution Proposed This paper demonstrates the development of EchoPrint [5]. It has two phases; the registration phase, where the front camera detects the facial landmarks such as the nose and eyes while the speaker sends acoustic signals to capture acoustic features through the reflected signals using CNN and SVM models. Here both Visual and Acoustic features are used as a combination for the authentication phase.

Results EchoPrint obtains 93.75% balanced accuracy and 93.50% F-score in tests, while the average precision with fundamental visual elements is 98.05%. With advanced graphical features, the accuracy is increased to 99.96%.

2.2.3 Using a Commodity Microphone Array to Capture Facial Expressions

Problem Statement Tracking facial expressions is essential in entertainment and movies to capture audience responses. Available Solution based on Image processing leads to privacy concern. Also, taking answers at the end of a particular event doesn't capture transient audience responses between events.

Solution Proposed SonicFace [1] is an acoustic-based emotion recognition system that recognizes the user's facial expressions and movements while watching television. SonicFace bombards the face by emitting ultrasonic sound waves mixed with the speaker's background music. The microphone array captures the reflected echoes, which communicate information about facial emotions such as brow, cheek, and lip movements.

Results Six participants' accuracy is shown by SonicFace: the average accuracy is 78.6%, with a standard variation of 6.25%.

2.2.4 Face authentication against spoofing using mmWave without on-site registration

Problem Statement Schemes for face authentication are widely used. But because they primarily rely on cameras, today's face identification systems are subject to face occlusion (such as facial masks, sunglasses, and scarves) and spoofing assaults like 3D-printed masks.

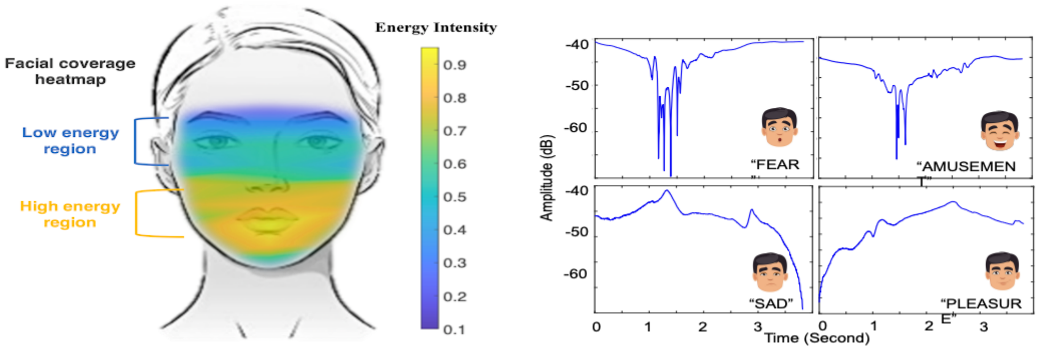
Solution Proposed To create the mmFace [4] anti-spoofing face authentication system, this article uses the penetrability, material sensitivity, and fine-grained sensing capacity of mmWave. By moving a commercial off-the-shelf (COTS) mmWave radar along a predetermined route, it scans the human face. The facial biometric attributes and structural attributes that are conveyed by the mmWave signals that are reflected off of a human face enable mmFace to perform facial authentication and liveness detection reliably.

Results The experiment's findings demonstrate that mmFace is reliable, secure, and accurate for user authentication.

Chapter 3

Methodology

3.1 Identification of Important Facial Regions



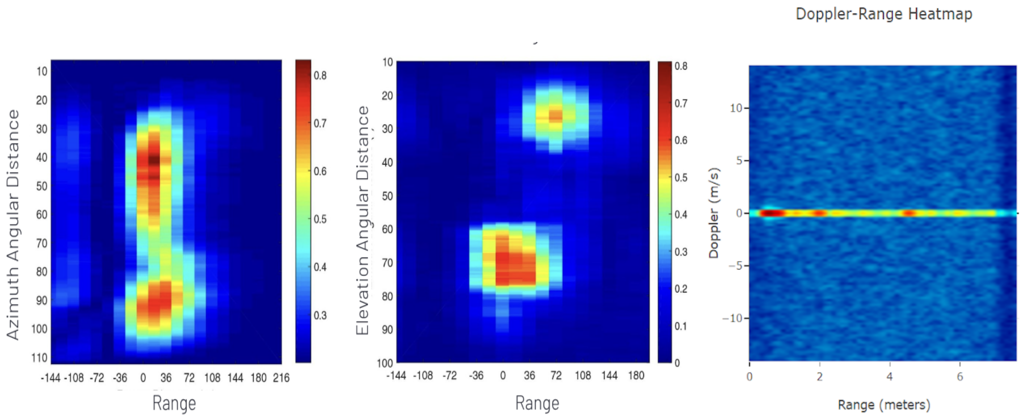
(a) Heatmap illustrates the intensity of the reflected echo signal

(b) Quadrature characteristics of the echo signals for various expressions

FIGURE 3.1: Facial coverage heatmap and Quadrature characteristics for various expressions

Different parts of the face respond to the incoming signal in various ways. For instance, Fig. 3.1(a) demonstrates that the energy intensity in the mouth and cheek regions is considerably higher than in the eye and brow regions, where weak muscle movements are seen. We illustrate the quadrature characteristics of the reflected echo signal for various expressions in Fig. 3.1(b). It has been noted that compared to "sad" and "joy," "fear" and "amusement" have substantially more comprehensive energy intensity ranges (high fluctuation). Because "fear" and "amusement" engage more mouth muscles, they are thought to reflect the echo signal with more significant multi-path modifications.

3.2 Definition



(a) Range Azimuthal Heatmap (b) Range Elevation Heatmap (c) Range Doppler Heatmap

FIGURE 3.2: Range-Azimuthal, Range-Elevation and Range-Doppler Heatmaps

The apparent position of an object in the sky concerning a particular observation point is defined by the angles of azimuth and elevation. Usually, but not always, the observer is situated on the earth's surface.

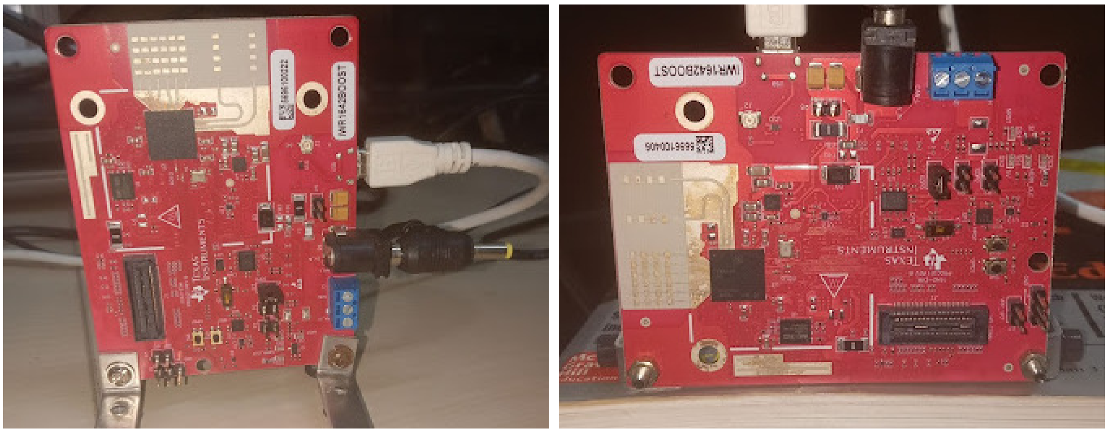
Azimuth A measurement of the horizon made between a fixed point (such as true north) and the vertical circle traversing the center of an object, generally used in astronomy and navigation clockwise through 360 degrees from the north point.

Elevation The elevation is a vertical angular measurement in a spherical coordinate system. Many take it as the vertical height of an object.

Range Doppler 2D heat-map whose ordinate is Doppler and abscissa is Range which gives information on user movement across different range bins.

3.3 Facial Expression Recognition using mmWave Radar (IWR1642BOOST)

Initially, we used TI's IWR1642BOOST board to send a mmWave chirp signal. We must keep the board in two different orientations to capture range-azimuth and range-elevation because two RX receiver antennas in IWR1642BOOST are in the same plane. So, to capture the azimuthal-range heatmap, we need to keep the board vertical, and to capture the elevation-range heatmap, we need to keep it horizontal. Fig. 3.3(a) and Fig. 3.3(b) show the board's vertical and horizontal orientation, respectively.

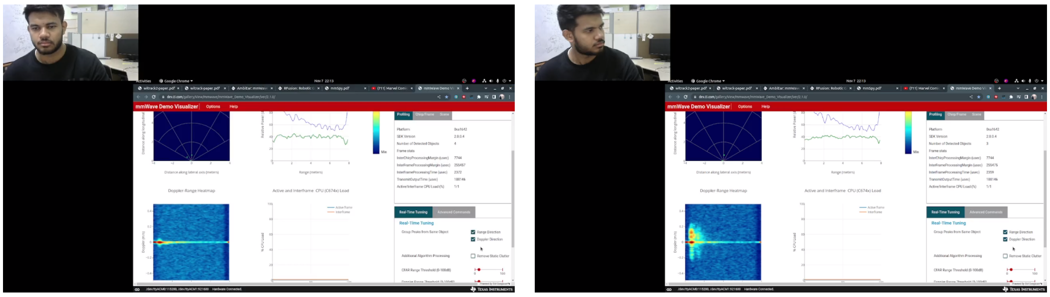


(a) Vertical Orientation for capturing range-azimuthal heatmap

(b) Horizontal Orientation for capturing range-elevation heatmap

Figure 3.3: Different Orientations of IWR1642BOOST board

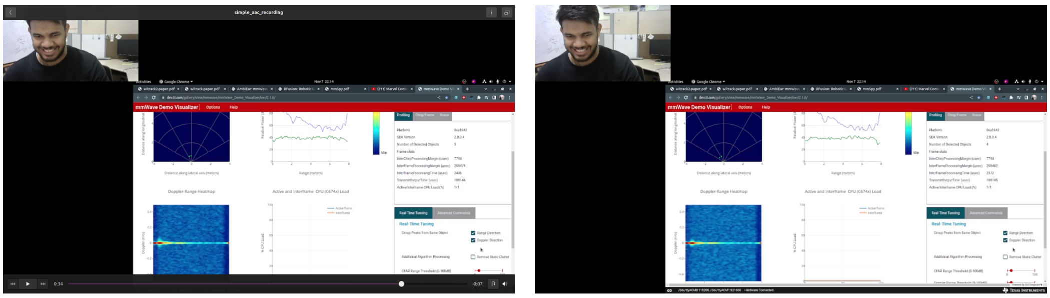
The following section demonstrates the change in different heatmaps due to the user's head and facial movements.



(a) No disturbance in range-doppler heatmap when the user is static

(b) Noticeable disturbance in range-doppler heatmap when the user is nodding his head

FIGURE 3.4: Behavioral change in range-doppler heatmap during static and continuous head movements



(a) Slight variation in range-doppler heatmap when user transits to smiling face from neutral state

(b) Range-doppler heatmap comes again back to equilibrium

FIGURE 3.5: An instantaneous change in range-doppler during the transition phase

3.3.1 Static User With No Facial Expression

Under neutral facial expression with no body part movement, i.e., the user is static, no variations are observed in the range-doppler. Fig. 3.4(a) depicts no change in range in the Doppler.

3.3.2 Capturing User's Facial Expression Using Range Doppler

During our pilot study, we found that when there is a macro movement like nodding of the head, there is a range change in range-Doppler, as shown in Fig. 3.4(b). Transition in the facial expression can be captured using the range-doppler patterns.

But a constant smile or facial expression can not be classified from the collected data as we can see in Fig.1 that initially when the user smiles, there is an instantaneous change in the Doppler pattern, which gradually comes to equilibrium.

3.4 Challenges with single mmWave Radar

We have already mentioned that the current IWR1642BOOST has all the RX antennae in the same plane. This limits the user to capture either range-azimuth or range-elevation as both need two different orientations of the radar. Also, our current IWR1642BOOST radar has two chips, one in front used for receiving the mmWave signal, and one in the back computes Point cloud, Range-FFT, and Range-Doppler computation. This chip does not give any functionalities for calculating range-azimuth and range-elevation. Also, these on-chip calculations add processing latency.

To address the above issues, we introduce two crucial changes in hardware. First, instead of IWR1642BOOST, we decided to use the AWR6843ISK. Fig. 3.6 depicts that out of three RX antennae, two are in the same plane and the other in a different plane. Thus, range-azimuth and range-elevation can be captured simultaneously.

Second, we captured raw data using DCA1000EVM to reduce on-chip computation overhead. Currently using single mmWave radar, we have 4 FPS as supported by the USB baud rate (250ms for each frame). DCA board enables capturing of High-Volume Raw ADC data using an Ethernet cable. We can further use raw data to calculate range-azimuth and range-elevation heatmaps.

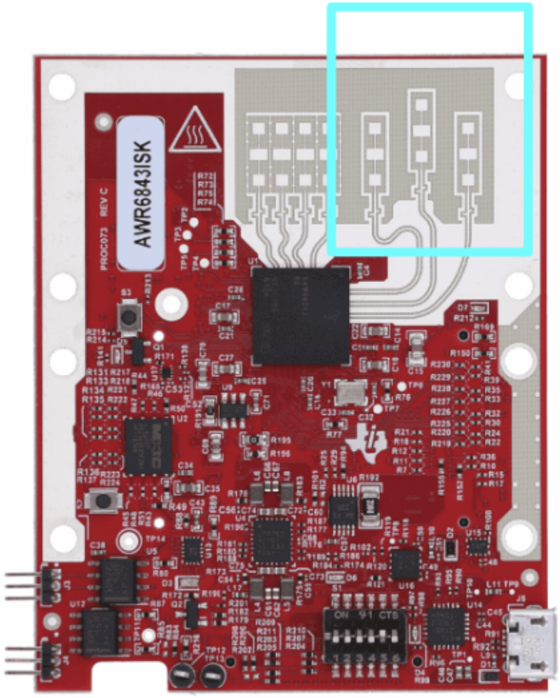


FIGURE 3.6: AWR6843ISK Board

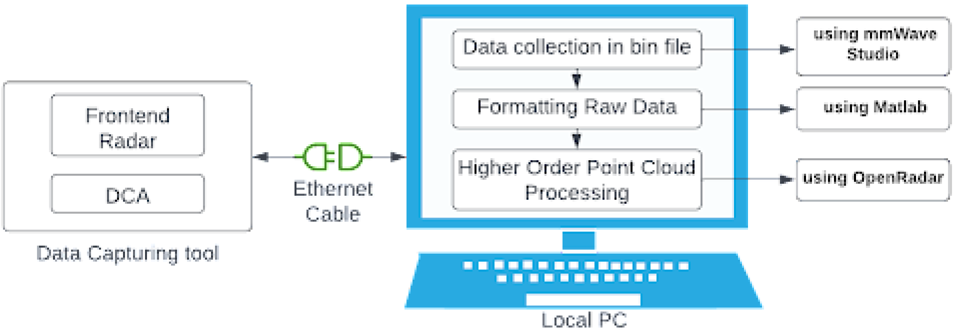


FIGURE 3.7: Processing Pipeline

3.5 Processing Pipeline

The overall processing pipeline is divided into four main parts. First, our Radar and DCA board jointly capture the data, which further send to the Local PC through an Ethernet cable. In Local PC, mmWave studio receives incoming data from the DCA board and dumps the raw ADC in the `.bin` file. This raw ADC data is further post-processed to format it into tabular form. After that, we computed the point cloud from the tabular data using OpenRadar [3]. Fig. 3.7 illustrates the overall processing pipeline.

3.5.1 Hardware Setup

For the hardware setup, we have the following components:

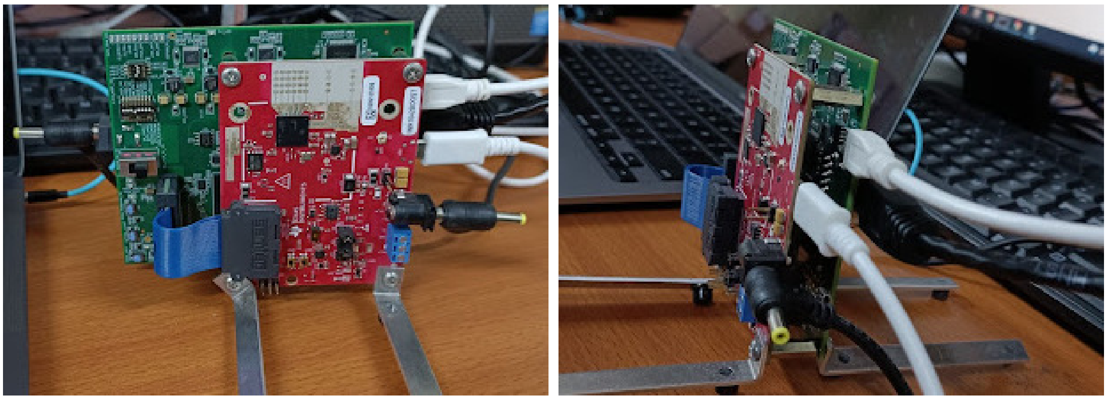
1. DCA1000EVM Data Capture Board [1],
2. IWR1642BOOST mmWave frontend Radar [2],
3. Ethernet Cable supporting 300 Mbps data transfer
4. Two 5V Power Adapter

The hardware components are shown in Fig. 3.8

3.5.2 mmWave Studio Setup

mmWave studio (mmWave Studio) [3] is a collection of tools that enhance the evaluation of TI mmWave sensors. These easy-to-use tools provide the capability to evaluate and prototype chirp designs and experiment with the out-of-the-box demo.

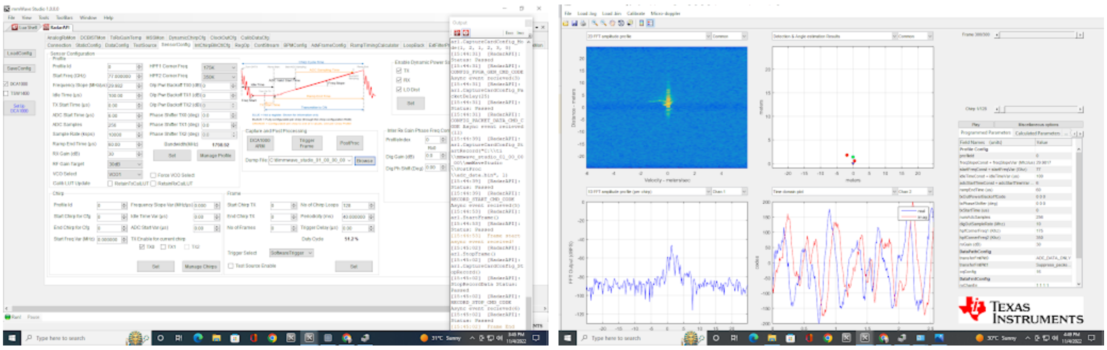
[1] `https://www.ti.com/tool/DCA1000EVM`
[2] https://www.ti.com/tool/IWR1642BOOST
[3] `https://www.ti.com/tool/MMWAVE-STUDIO`



(a) Front view

(b) Side view

FIGURE 3.8: Hardware setup



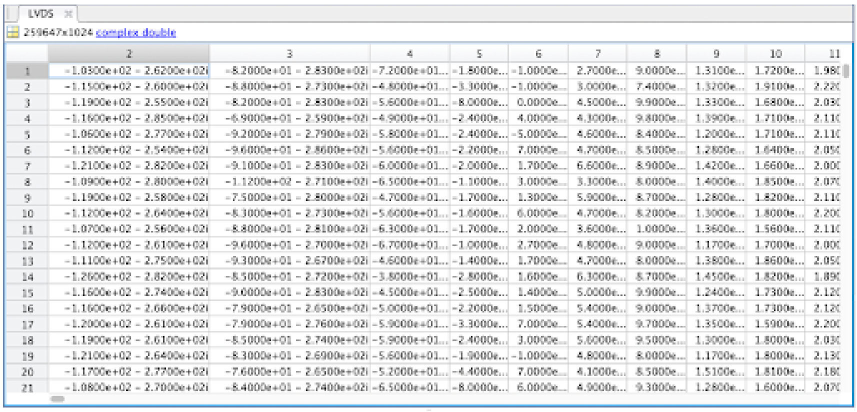
(a) mmWave studio sensor configuration

(b) Post processed data on mmWave studio

FIGURE 3.9: mmWave studio setup

The tools are hosted directly on TI.com and enable interaction with mmWave sensor evaluation modules (EVMs). The tools also can be saved for offline use.

mmWave Studio is a stand-alone Windows GUI that can configure and control mmWave sensor modules and collect analog-to-digital (ADC) data for offline analysis. ADC data capture is intended to enable the evaluation and characterization of radio-frequency (RF) performance and PC development of signal-processing algorithms.



LVDS

259647x1024 complex double

	2	3	4	5	6	7	8	9	10	11
1	-1.0300e+02 - 2.6200e+02i	-8.2000e+01 - 2.8300e+02i	-7.2000e+01...	-1.8000e...	-1.0000e...	2.7000e...	9.0000e...	1.3100e...	1.7200e...	1.98
2	-1.1500e+02 - 2.6000e+02i	-8.8000e+01 - 2.7300e+02i	-4.8000e+01...	-3.3000e...	-1.0000e...	3.0000e...	7.4000e...	1.3200e...	1.9100e...	2.22
3	-1.1900e+02 - 2.5500e+02i	-8.2000e+01 - 2.8300e+02i	-5.6000e+01...	-8.0000e...	0.0000e...	4.5000e...	9.9000e...	1.3300e...	1.6800e...	2.03
4	-1.1600e+02 - 2.8500e+02i	-6.9000e+01 - 2.5900e+02i	-4.9000e+01...	-2.4000e...	4.0000e...	4.3000e...	9.8000e...	1.3900e...	1.7100e...	2.11
5	-1.0600e+02 - 2.7700e+02i	-9.2000e+01 - 2.7900e+02i	-5.8000e+01...	-2.4000e...	-5.0000e...	4.6000e...	8.4000e...	1.2000e...	1.7100e...	2.11
6	-1.1200e+02 - 2.5400e+02i	-9.6000e+01 - 2.8600e+02i	-5.6000e+01...	-2.2000e...	7.0000e...	4.7000e...	8.5000e...	1.2800e...	1.6400e...	2.05
7	-1.2100e+02 - 2.8200e+02i	-9.1000e+01 - 2.8300e+02i	-6.0000e+01...	-2.0000e...	1.7000e...	6.6000e...	8.9000e...	1.4200e...	1.6600e...	2.00
8	-1.0900e+02 - 2.8000e+02i	-1.1200e+02 - 2.7100e+02i	-6.5000e+01...	-1.1000e...	3.0000e...	3.3000e...	8.0000e...	1.4000e...	1.8500e...	2.07
9	-1.1900e+02 - 2.5800e+02i	-7.5000e+01 - 2.8000e+02i	-4.7000e+01...	-1.7000e...	1.3000e...	5.9000e...	8.7000e...	1.2800e...	1.8200e...	2.11
10	-1.1200e+02 - 2.6400e+02i	-8.3000e+01 - 2.7300e+02i	-5.6000e+01...	-1.6000e...	6.0000e...	4.7000e...	8.2000e...	1.3000e...	1.8000e...	2.20
11	-1.0700e+02 - 2.5600e+02i	-8.8000e+01 - 2.8100e+02i	-6.3000e+01...	-1.7000e...	2.0000e...	3.6000e...	1.0000e...	1.3600e...	1.5600e...	2.11
12	-1.1200e+02 - 2.6100e+02i	-9.6000e+01 - 2.7000e+02i	-6.7000e+01...	-1.0000e...	2.7000e...	4.8000e...	9.0000e...	1.1700e...	1.7000e...	2.00
13	-1.1100e+02 - 2.7500e+02i	-9.3000e+01 - 2.6700e+02i	-4.6000e+01...	-1.4000e...	1.7000e...	4.7000e...	8.0000e...	1.3800e...	1.8600e...	2.05
14	-1.2600e+02 - 2.8200e+02i	-8.5000e+01 - 2.7200e+02i	-3.8000e+01...	-2.8000e...	1.6000e...	6.3000e...	8.7000e...	1.4500e...	1.8200e...	1.89
15	-1.1600e+02 - 2.7400e+02i	-9.0000e+01 - 2.8300e+02i	-4.5000e+01...	-2.5000e...	1.4000e...	5.0000e...	9.9000e...	1.2400e...	1.7300e...	2.12
16	-1.1600e+02 - 2.6600e+02i	-7.9000e+01 - 2.6500e+02i	-5.0000e+01...	-2.2000e...	1.5000e...	5.4000e...	9.0000e...	1.3700e...	1.7300e...	2.12
17	-1.2000e+02 - 2.6100e+02i	-7.9000e+01 - 2.7600e+02i	-5.9000e+01...	-3.3000e...	7.0000e...	5.4000e...	9.7000e...	1.3500e...	1.5900e...	2.20
18	-1.1900e+02 - 2.6100e+02i	-8.5000e+01 - 2.7400e+02i	-5.9000e+01...	-2.4000e...	3.0000e...	5.6000e...	9.5000e...	1.3000e...	1.8000e...	2.03
19	-1.2100e+02 - 2.6400e+02i	-8.3000e+01 - 2.6900e+02i	-5.6000e+01...	-1.9000e...	-1.0000e...	4.8000e...	8.0000e...	1.1700e...	1.8000e...	2.13
20	-1.1700e+02 - 2.7700e+02i	-7.6000e+01 - 2.6500e+02i	-5.2000e+01...	-4.4000e...	7.0000e...	4.1000e...	8.5000e...	1.5100e...	1.8100e...	2.18
21	-1.0800e+02 - 2.7000e+02i	-8.4000e+01 - 2.7400e+02i	-6.5000e+01...	-8.0000e...	6.0000e...	4.9000e...	9.3000e...	1.2800e...	1.6000e...	2.07

FIGURE 3.10: Post-processed matrix

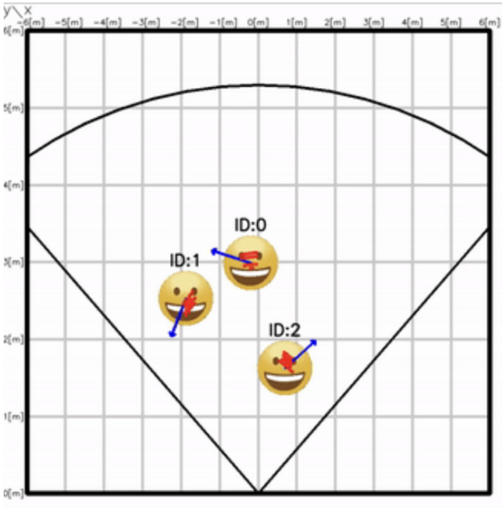


FIGURE 3.11: Point cloud plot using OpenRadar

3.5.3 Pre-processing Raw ADC Data Using MATLAB

MATLAB script is used to pre-process the raw ADC data. Hyperparameters such as the Number of ADC samples (`numADC`), Number of RX antennae (`numRX`), and Number of chirps per frame (`numCPF`) are collected from mmWave Studio. After pre-processing the data from the `.bin` file, we stored the data in the `.mat` file as an array of dimensions `[numChirps, (numADC * numRX)]`. This `.mat` file is further analyzed by python scripts (`openRadar`) for higher-order processing. An example post-processed matrix is shown in Fig. 3.10

3.5.4 Point Cloud information using OpenRadar

OpenRadar is used to process ADC data for higher-order calculation. It is open source and can be installed using the pip command: `pip install openradar`. It applies range and doppler processing on top of raw ADC data. We can use this platform for further signal processing to gain range-azimuth and range-elevation heatmap. The point cloud generated by openradar is illustrated in Fig. 3.11

Chapter 4

Implementation Challenges

We faced the following implementation challenges. Some of them are resolved, and some are yet to be resolved.

1. To capture high-volume raw ADC data, we had to connect mmWave radar to the DCA board, during which we faced many challenges to make a setup such that the DCA board gets proper input from radar to be able to transfer it using an ethernet cable to the host system.

2. mmWave studio has 2 versions, and we first tried using the latest version(2.0), but due to a lack of a proper setup guide, we could not use that.

3. We switched to version 1.0, and there also the guide broadly helped us with using the software interface but very little about pre-requisites and other system compatibility steps for which we had to spend hours trying out various configurations to find the perfect one with several queries posted on TI discussion forum.

4. Initially, the number of data frames captured on mmWave Studio was fixed; we modified the configuration to capture for infinite time.

5. We now have the data in the required format, but the data is still not accurate as the bandwidth utilized in the current setup is just 0.75GHz compared to the ideal bandwidth of 4GHz, resulting in low range resolution and hence low precision.

6. The data in bin files from mmWave studio required pre-processing to visualize the captured motion using publicly available tools, for which we had to write MATLAB Script to obtain data in a matrix form.

7. We modified and implemented a few publicly available repositories supporting pre-processing of raw ADC data to obtain required heatmaps, which failed as they were very specific to their use. We even created issues and mailed them, but all went unanswered.

8. We finally found the OpenRadar [3] library, which we are currently modifying to fit our data input and provide output to match our requirements.

9. We are the first to work on the recognition of facial expressions using mmWave. Also, as mmWave is a new tech, we don't have publicly available datasets.

Chapter 5

Future Work

As a part of future work, the following tasks will be performed:

1. Adjust hyperparameters in "Sensor Config → Profile" so that we utilize the entire 4 GHz bandwidth for data capturing with enhanced range resolution.

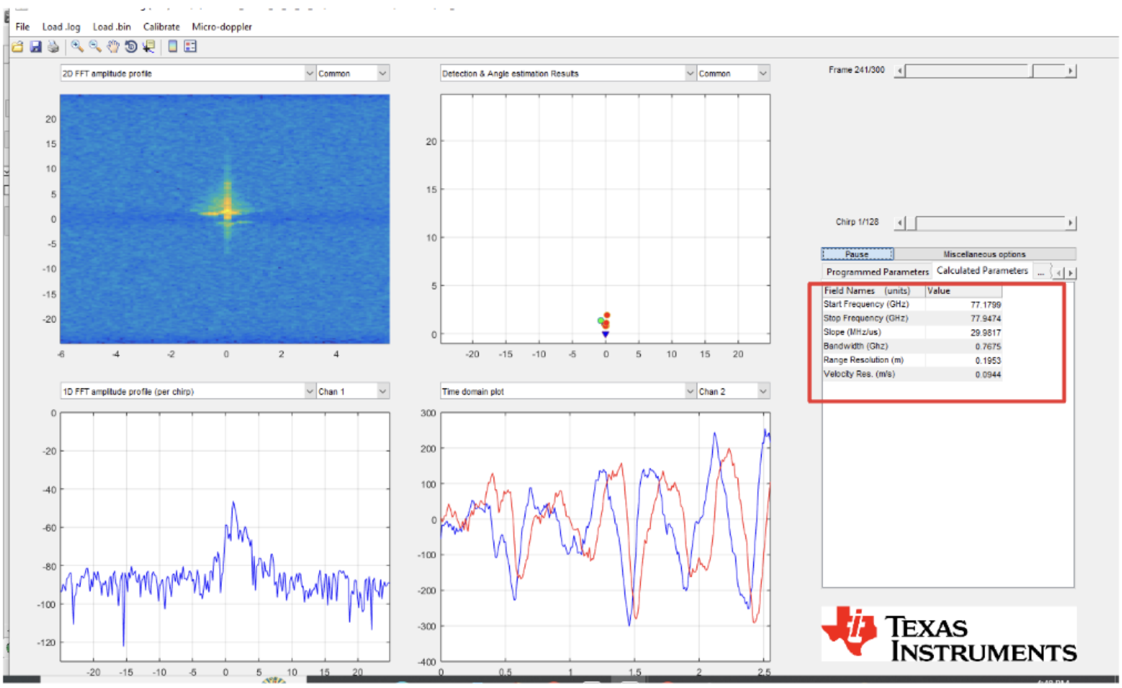


FIGURE 5.1: Bandwidth is not getting used completely

Fig. 5.1 illustrates the above-mentioned problem.

2. We are still unable to pre-process the complete data. Some tailing data we are discarding during pre-processing. We need to Fine-tune the pre-processing data pipeline.

3. We need to complete data collection of each type of facial expression, i.e., happy, amazed, angry, neutral, etc.

4. Training ML/DL model on Range doppler, Range Azimuth, and Range Elevation heatmap for facial expression detection.

Bibliography

[1] Gao, Y., Jin, Y., Choi, S., Li, J., Pan, J., Shu, L., Zhou, C., and Jin, Z. (2021). Sonicface: Tracking facial expressions using a commodity microphone array. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 5(4):1–33.

[2] Liu, J., Li, D., Wang, L., and Xiong, J. (2021). Blinklistener: " listen" to your eye blink using your smartphone. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 5(2):1–27.

[3] Pan, E., Tang, J., Kosaka, D., Yao, R., and Gupta, A. (2019). Openradar. `https://github.com/presenseradar/openradar`.

[4] Xu, W., Song, W., Liu, J., Liu, Y., Cui, X., Zheng, Y., Han, J., Wang, X., and Ren, K. (2022). Mask does not matter: anti-spoofing face authentication using mmwave without on-site registration. In *Proceedings of the 28th Annual International Conference on Mobile Computing And Networking*, pages 310–323.

[5] Zhou, B., Xie, Z., Zhang, Y., Lohokare, J., Gao, R., and Ye, F. (2021). Robust human face authentication leveraging acoustic sensing on smartphones. *IEEE Transactions on Mobile Computing*.