



University of Glasgow | School of  
Computing Science

# 3D Face Recognition from RGB Camera and Radar Sensor

Stergious Aji

School of Computing Science  
Sir Alwyn Williams Building  
University of Glasgow  
G12 8QQ

MSci Interim Report

November 26, 2023

# Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
1.1	Motivation . . . . .	2
1.2	Aims . . . . .	3
<b>2</b>	<b>Background Survey</b>	<b>3</b>
2.1	Radar Technology . . . . .	3
2.2	Data Acquisition . . . . .	4
2.3	Prior Work on Radar-based Face Recognition . . . . .	5
2.4	Multimodality of Data . . . . .	6
2.5	Data Fusion Techniques . . . . .	6
2.6	Deep Learning for Face Recognition . . . . .	6
<b>3</b>	<b>Proposed Approach</b>	<b>6</b>
<b>4</b>	<b>Work Plan</b>	<b>6</b>

# 1 Introduction

## 1.1 Motivation

PRIVACY

RADAR PENETRATE HAIR, CLOTH [\[1\]](#)

3D FACE RECOGNITION SURVEY [\[2\]](#)

Facial Recognition is a crucial area of research for its wide range of applications spanning security surveillance, forensics, human-computer interaction and healthcare. Its most popular application being access control using biometric authentication. This removes the need to remember passwords and provides a non-invasive, hands-free approach to human verification. Facial metrics are naturally more accessible in comparison to other biometrics like fingerprint, iris or palm print.

Facial recognition systems have come a long way since its dawn in the 1960s. The earliest work by Bledsoe [\[3\]](#) distinguished faces by comparing distances of manually annotated landmark features such as the nose, eyes, ears and mouth. In more recent years, the advent of Deep Learning techniques has greatly improved face recognition performance with help of the sheer number of images of faces online. However, these systems primarily rely on 2D imagery from RGB cameras which are vulnerable to lighting changes and pose variations. To compensate for this, depth information of facial attributes are required. Additionally, moving to 3D facial recognition increases the security of biometric authentication systems.

3D face recognition systems are becoming more popular with the likes of many smartphone companies integrating a type of face unlock such as Apple's Face ID [\[4\]](#). Furthermore, this demand has pushed this depth sensing technology to smaller form factors and requiring little power and computation to work efficiently on mobile devices on the fly.

Most depth cameras used for this purpose use a form of active face acquisition where non-visible light is emitted and reflected back from a person's face which is subsequently captured by sensors and measured to estimate facial features. The most popular approach uses LiDAR which emit waves in the near-infrared spectrum. The main disadvantage to this is that it is usually too weak to penetrate clothing or hair. In contrast, millimetre waves used in Radar can penetrate thin objects to directly reach the dermal layer of the skin meaning that it may perform better against occlusion like obstacles or even rain or fog.

Very little research has been done in the effectiveness of using Radar waves for 3D face recognition but what has been done show positive results [\[5, 6, 7, 8, 9\]](#). Radar technology is often less expensive in terms of acquisition cost and computational cost since it requires little energy to power compared to LiDAR cameras. However, Radar has its drawbacks since it is less accurate and sparse which may hinder its performance for facial recognition.

A solution is to combine RGB information with the depth captured by the Radar sensors to effectively learn facial features and identify them.

## 1.2 Aims

This project aims to investigate the effectiveness of using RGB cameras in conjunction with mmWave Radar sensors for 3D facial recognition. Since there are no easily accessible datasets online for this purpose, we will require acquiring this data ourselves. We will be using an Intel RealSense L515 [10] camera to capture the RGB information of a subject's face. The Google Soli Radar sensor [11] will be used to capture depth information from reflected millimetre waves.

This RealSense camera also includes a LiDAR sensor which produces a more accurate dense depth image. As a backup, a separate model can be trained to transform the sparse Soli data into a more dense representation before inputting into the facial recognition model. However, if the Radar data works well this may not be needed. (MAY NOT INCLUDE)

Since data must be collected we aim to collect faces of around 50 participants which will be enough to train and evaluate our proposed model given the tight timeframe. We aim to acquire face data of different poses, lighting conditions and with multiple occlusion scenarios.

Next we aim to produce our very own face recognition model using a Deep Concurrent Neural Network to learn facial features using both the RGB and depth information acquired from the Soli sensor. We will investigate different data fusion techniques in order to find the best approach to combine RGB features with the Radar data.

## 2 Background Survey

Before any work is conducted, a review of relevant literature in the field is conducted in order to corroborate best practices and gain a deeper understanding of the strengths and weaknesses of using a Radar sensor, specifically for face identification. Firstly, a look into the datasets available for our purpose and the subsequent reasoning behind the requirement to gather our own dataset will be explored. This section also provides recent study of using machine learning techniques to classify radar signatures of human faces.

### 2.1 Radar Technology

Radio Detection and Ranging, or Radar, has been around for decades and is used in many important applications such as space exploration, military and commercial aircraft and ship navigation, as well as, weather forecasting. However, only recently with the miniaturisation

of Radar sensors to the millimetre wave band (mmWave) has research expanded to other small scale domains [12]. Most notably, Google integrated their own Soli sensor for face detection and motion gesture recognition into their Pixel 4 smartphone [13]. Furthermore, mmWave sensing is being explored in the domain of autonomous driving for instance, the application of collision warning and adaptive cruise control systems [14]. mmWaves offer an advantage over the traditional near-infrared waves used by Light Detection and Ranging (Lidar) cameras, with its robustness to atmospheric conditions such as dust, smoke, fog and rain as well as extreme lighting conditions [15]. This is counter-balanced by the lower accuracy of mmWaves in comparison.

## 2.2 Data Acquisition

There are various techniques available in capturing 3D face data, but they can be categorised into two main types: *Active* and *Passive*. Active acquisition systems emit non-visible light on to the target and use the reflected light information to estimate a 3D representation of the object. Contrastingly, passive systems use the available light in the scene to capture facial feature information. For instance, stereoscopic cameras use two or more cameras to capture images of the target from different perspectives enabling depth perception. In addition, 3D facial features can be inferred using shape-from-shading by analysing the luminance values of a 2D grayscale image.

While passive systems offer advantages of not requiring emission of light and being able to work in real-time, it is highly sensitive to the lighting conditions in the environment. Active systems are robust against this since they use their own light to illuminate the target being able to work in extremely dim conditions.

Active systems often combine the use of structured light and triangulation based methods to capture depth information. In the case of Lidar cameras such as the Intel RealSense, the time-of-flight of emitted light is measured to infer the distances of points on the target. The Radar sensor is also a form of active acquisition using 3 receiving antennas to capture reflected mmWaves from the target, measuring the phase difference and Doppler shift to estimate the distance and velocity of the target.

When collecting data of human faces for 3D face recognition, it is often ideal to ensure the data encompasses a wide range of facial poses and expressions, lighting conditions and occlusion scenarios. Unlike 2D face images, the pure geometric information from 3D face scans ensure the models are insensitive to pose and lighting changes. [16] found that the maximum angle that their Local Binary Pattern (LBP) based model is robust against is 60°. A wide range of genders, ages and ethnicities are also vital for the model to be robust enough against real-world scenarios.

In recent years, the number of 3D face databases available has grown using a variety of different acquisition techniques and devices. Most notable datasets include the BU-3DFE [17] and the FRGC [18] database, widely accepted as the standard reference database for performance evaluation of 3D face recognition systems. The BU-3DFE dataset primarily

focusses on expression-invariance containing 6 types of expressions from 100 individuals using stereo photography. While these datasets are unsuitable for this project, it is useful to survey the data collection process used to amass them. From the extensive research done, there is currently only one public database including Radar signatures of 206 human faces available [19]. This dataset was captured with a Qualcomm 60GHz mmWave Radar, however, does not include any RGB face images of the participants of the study. These factors motivate building our own dataset with both RGB face images and mmWave Radar face signatures. This allows us to investigate the model’s effectiveness against varying conditions such as lighting, pose variation and occlusion.

## 2.3 Prior Work on Radar-based Face Recognition

The use of mmWaves for human face identification is a relatively new research field driven by the miniaturisation of Radar sensor technology to the millimetre wave band. The earliest paper found to investigate human identification using mmWaves dates back to 2019 [20]. While this paper focusses on simultaneous classification of people by their gait and body shape rather than facial features, it displayed the ability for mmWaves to capture subtle idiosyncracies between individuals for machine learning models to achieve accurate classification accuracies.

Following this, Hof et. al. [5] proposed a Deep Neural Network (DNN) based Autoencoder that is able to distinguish human faces captured by a 802.11ad/y networking chipset operating at a centre frequency of 60GHz. The Autoencoder is able to encode mmWave face signatures of over 200 individuals with enough separation to distinguish positive and negative instances by measuring their Mean Squared Error (MSE) against a reference encoding of a face. The study involved a decently-sized data acquisition procedure capturing mmWave signatures of 206 individuals of varying genders and ages with 5 poses each: frontal, as well as, 15° and 25° left and right. This dataset was subsequently made available as a IEEE Data Port [19], however, does not include common occlusion scenarios which our study hopes to investigate. Additionally, the chipset used contained a large sensor containing over 1024 transmitting and receiving antenna pairs. This is in contrast to the compact mobile Soli chip with a single transmit and three receiver antennas. The study also simulated the effect of reducing the number of antennas to 10 which made a significant reduction in the distinctivity of the faces. Promisingly, increasing the number of neurons and an additional hidden layer to their Neural Network was able to maintain high accuracy.

Lim et. al. [6]

Kim et. al. [7]

Pho et. al. [8]

Challa et. al. [9]

## 2.4 Multimodality of Data

## 2.5 Data Fusion Techniques

## 2.6 Deep Learning for Face Recognition

# 3 Proposed Approach

state how you propose to solve the software development problem. Show that your proposed approach is feasible, but identify any risks.

# 4 Work Plan

show how you plan to organize your work, identifying intermediate deliverables and dates.

# References

- [1] David R Vizard and R Doyle. Advances in millimeter wave imaging and radar systems for civil applications. In *2006 IEEE MTT-S International Microwave Symposium Digest*, pages 94–97. IEEE, 2006.
- [2] Song Zhou and Sheng Xiao. 3d face recognition: a survey. *Human-centric Computing and Information Sciences*, 8(1):1–27, 2018.
- [3] Woodrow Wilson Bledsoe. The model method in facial recognition. *Panoramic Research Inc., Palo Alto, CA, Rep. PR1*, 15(47):2, 1966.
- [4] Apple Inc. About Face ID advanced technology, 2023. Accessed: 2023-11-19 <https://support.apple.com/en-gb/102381>.
- [5] Eran Hof, Amichai Sanderovich, Mohammad Salama, and Evyatar Hemo. Face verification using mmwave radar sensor. In *2020 International Conference on Artificial Intelligence in Information and Communication (ICAIIIC)*, pages 320–324, 2020.
- [6] Hae-Seung Lim, Jaehoon Jung, Jae-Eun Lee, Hyung-Min Park, and Seongwook Lee. Dnn-based human face classification using 61 ghz fmcw radar sensor. *IEEE Sensors Journal*, 20(20):12217–12224, 2020.
- [7] J Kim, J-E Lee, H-S Lim, and S Lee. Face identification using millimetre-wave radar sensor data. *Electronics Letters*, 56(20):1077–1079, 2020.

- [8] Ha-Anh Pho, Seongwook Lee, Vo-Nguyen Tuyet-Doan, and Yong-Hwa Kim. Radar-based face recognition: One-shot learning approach. *IEEE Sensors Journal*, 21(5):6335–6341, 2021.
- [9] Muralidhar Reddy Challa, Abhinav Kumar, and Linga Reddy Cenkeramaddi. Face recognition using mmwave radar imaging. In *2021 IEEE International Symposium on Smart Electronic Systems (iSES)*, pages 319–322, 2021.
- [10] Intel Corporation. Intel RealSense LiDAR Camera L515, 2023. Accessed: 2023-11-19 <https://www.intelrealsense.com/lidar-camera-l515/>.
- [11] Jaime Lien, Nicholas Gillian, M Emre Karagozler, Patrick Amihood, Carsten Schwesig, Erik Olson, Hakim Raja, and Ivan Poupyrev. Soli: Ubiquitous gesture sensing with millimeter wave radar. *ACM Transactions on Graphics (TOG)*, 35(4):1–19, 2016.
- [12] A Soumya, C Krishna Mohan, and Linga Reddy Cenkeramaddi. Recent advances in mmwave-radar-based sensing, its applications, and machine learning techniques: A review. *Sensors*, 23(21):8901, 2023.
- [13] Nicholas Gillian Jaime Lien. Soli: Radar-based perception and interaction, 2020. Accessed: 2023-11-25 <https://blog.research.google/2020/03/soli-radar-based-perception-and.html>.
- [14] DF Robot. Eight Practical Applications of mmWave Radar Technology, 2023. Accessed: 2023-11-19 <https://www.dfrobot.com/blog-1650.html>.
- [15] Cadence Design Systems. mmwave radar applications and advantages, 2022. Accessed: 2023-11-25 <https://resources.system-analysis.cadence.com/blog/msa2022-mmwave-radar-applications-and-advantages>.
- [16] Utsav Prabhu, Jingu Heo, and Marios Savvides. Unconstrained pose-invariant face recognition using 3d generic elastic models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(10):1952–1961, 2011.
- [17] Lijun Yin, Xiaozhou Wei, Yi Sun, Jun Wang, and Matthew J Rosato. A 3d facial expression database for facial behavior research. In *7th international conference on automatic face and gesture recognition (FGRO6)*, pages 211–216. IEEE, 2006.
- [18] P Jonathon Phillips, Patrick J Flynn, Todd Scruggs, Kevin W Bowyer, Jin Chang, Kevin Hoffman, Joe Marques, Jaesik Min, and William Worek. Overview of the face recognition grand challenge. In *2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05)*, volume 1, pages 947–954. IEEE, 2005.
- [19] Evyatar Hemo, Amichai Sanderovich, and Eran Hof. mmwave radar face signatures, 2018. <https://dx.doi.org/10.21227/wr67-kx23>.
- [20] Peijun Zhao, Chris Xiaoxuan Lu, Jianan Wang, Changhao Chen, Wei Wang, Niki Trigoni, and Andrew Markham. mid: Tracking and identifying people with millimeter wave radar. In *2019 15th International Conference on Distributed Computing in Sensor Systems (DCOSS)*, pages 33–40. IEEE, 2019.