



University of Glasgow | School of  
Computing Science

# 3D Face Recognition from RGB Camera and Radar Sensor

Stergios Aji

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

MSci Interim Report

November 9, 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	Data Acquisition . . . . .	3
2.2	Multimodality of Data . . . . .	3
2.3	Data Fusion Techniques . . . . .	3
2.4	Deep Learning for Face Recognition . . . . .	3
<b>3</b>	<b>Proposed Approach</b>	<b>3</b>
<b>4</b>	<b>Work Plan</b>	<b>4</b>

# 1 Introduction

## 1.1 Motivation

Facial Recognition is a crucial area of research for its wide range of applications spanning security surveillance, forensics, human-computer interaction and healthcare. It's 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 [1](CITE) 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 FaceID [2](CITE). 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, millimeter 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 [3](CITE). 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 [(CITE)] camera to capture the RGB information of a subject's face. The Google Soli Radar sensor [(CITE)] will be used to capture depth information from reflected millimeter waves.

This RealSense camera also includes a LiDAR sensor which produces a more accurate dense 3D point cloud. 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

### 2.1 Data Acquisition

### 2.2 Multimodality of Data

### 2.3 Data Fusion Techniques

### 2.4 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**