Kent Academic Repository

Robertson, Joshua (2017) ***The Application of the Human-Biometric Sensor Interaction Method to Automated Border Control Systems.*** **Doctor of Philosophy (PhD) thesis, University of Kent,.**

The Application of the
Human-Biometric Sensor Interaction Method
to Automated Border Control Systems

A Thesis submitted to the University of Kent for the Degree of Doctor of Philosophy
in Electronic Engineering

By Joshua Robertson

October 2017

ABSTRACT

Biometrics components are used in many different systems and technologies to verify that the user is whom they say they are. In Automated Border Control systems, biometrics components used in conjunction with a traveller's documents to make sure the user is whom they say they are so that they can cross into a countries borders. The systems are expected to verify the identity with a higher degree than officers who manually check travellers.

Each year the number of travellers crossing through a country borders increases and so systems are expected to handle bigger demands; through improving the user experience to ensuring accuracy and performance standards increase.

While the system does bring its benefits through increased speed and higher security, there are drawbacks. One of the main issues with the systems is a lack of standardisation across implementations. Passing through an automated process at Heathrow may be different to Hong Kong. The infrastructure, information, environment and guidance given during the transaction will all greatly differ for the user. Furthermore, the individual components and subsequent processing will be evaluated using a different methodology too.

This thesis reports on the contrasts between implementations, looking at solutions which utilise different biometric modalities and travel documents. Several models are devised to establish a process map which can be applied to all systems. Investigating further, a framework is described for a novel assessment method to evaluate the performance of a system. An RGB-D sensor is implemented, to track and locate the user within an interactive environment. By doing so, the user's interaction is assessed in real-time. Studies then report on the effectiveness of the solution within a replicated border control scenario. Several relationships are studied to improve the technologies used within the scenario. Successful implementation of the automated assessment method may improve the user's experience with systems, improving information and guidance, increasing the likelihood of successful interaction while maintaining a high level of security and quicker processing times.

ACKNOWLEDGEMENTS

I would like to thank my supervisor Dr Richard Guest for his extensive guidance and support over the last four years. Your feedback throughout this process has been invaluable, and I am extremely grateful for your support throughout this journey. Additionally, I would like to extend my thanks to Dr Stephen Elliott for his support and feedback throughout the early stages of my PhD.

I also need to extend my gratitude to my colleagues in the department. Chiara Lunerti and Christina Komliki, you both were there for me throughout my studies, and I will always cherish the times we had together in the lab.

Finally, I would like to thank Aodhán, Emma and my dad for their support and encouragement throughout my studies.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ABC – Automated Border Control

APC – Automated Passport Control

BSI – British Standards Institute

CI – Concealed Interaction

DHS – (United States) Department of Homeland Security

DI – Defective Interaction

DCT - Discrete Cosine Transform

FC – False Claim

FAR - False Acceptance Rate

FI – False Interaction

FMNR – False Non-Match Rate

FMR – False Match Rate

FRR – False Recognition Rate

FTA – Failure to Acquire

FTD – Failure to Detect

FTE – Failure to Enrol

FTP – Failure to Process

GM – Generic Model (ABC)

HBSI – Human-Biometric Sensor Interaction

HCI – Human-Computer Interaction

ICAO - International Civil Aviation Organization

ICP – Identity Claim Process

IEC - International Electrotechnical Commission

IRIS - Iris Recognition Immigration System

IOM – Iris on the Move

ISO - International Organization for Standardization

MRZ – Machine Readable Zone

MRTD – Machine Readable Travel Document

NIST - National Institute of Standards and Technology

RFID – Radio-frequency identification

RGB – Red Green Blue

SDK – Software Development Kit

SPS – Successfully Processed Sample

UI – User Interface

CHAPTER 1. INTRODUCTION

The work in this thesis is concerned with user interaction performance with the biometric components of Automated Border Control systems. Biometric modalities provide an additional step to the verification process of a traveller crossing through a countries border. The interaction between a human and sensor plays a significant role which will impact the system's decision and the outcome of the process. If a ‘correct’ interaction is presented with little to no errors, the biometric sample is likely to be accepted for matching. If an ‘incorrect’ presentation is made, then it increases the chance that the sample will not be accepted. If the algorithm decides that the sample is poor and is therefore rejected due to not meeting the threshold or by a processing error, the user may either try again or queue for manual processing.

Common biometric applications do not distinguish whether an error is either system based (e.g. the system failed to acquire a satisfactory sample) or if an error is user caused (e.g. the user presented a sample in such a way that it was difficult for the system to obtain a satisfactory sample) and merely wrap the two together.

Performance assessment of interactions is a challenging area and requires an in-depth understanding of the subtle difference between system generated and user interaction errors. This thesis suggests novel techniques in the assessment of human-biometric based interactions with two common biometric modalities used in border control systems. Unlike existing works described in the literature, this study focuses on how assessing user interaction in real time will not only enable improved methods of assessment but also how the captured data may enhance the guidance and information provided throughout the process.

In the rest of this chapter, the motivation for this research is further expanded upon. Section 1.1 will present a brief introduction to biometric systems along with a summary of the potential vulnerability in current performance evaluation techniques. Section 1.2 will introduce the motivation for this research. Aims and objectives are listed in Section 1.3. A scope of the work will be explored in Section 1.4. Finally, Section 1.5 will outline the structure of the thesis.

1.1 Biometric Systems

Automated Border Control (ABC) systems consist of several components which all contribute to the main aim of increasing the level of security in verifying travellers to cross a country’s borders. One of the key components of the system, the biometric module, aims to reliably recognise or verify a person attempting to pass through the border. Many questions are raised during the process, for example, “Is she/he really whom she/he claims to be?” or “Is this person using this system authorised to access this process?” Moreover, the biometric element is processed alongside other internal functions which run in the background; comparing the captured sample against a previously enrolled image, checking the traveller against databases, watch lists and so forth.

Research into techniques that can improve the performance of an automated recognition system is fundamental. Most applications (e.g. unlocking a phone, accessing a website) traditionally verifies a user based on what “she/he remembers or knows”, for example, through remembering a password or PIN. For ABC Systems, such recognition is also used in conjunction with biometric verification and are typically based on tokens possessed by the traveller, such as electronic passports and identity cards. However, if the ID card is stolen or a password is known to an unauthorised user, security can be breached, especially if the system is unattended. Recognition based on what a person is or does can address the problems related to these traditional methods. Combining the two is a powerful tool in identifying and subsequently verifying a user’s claim of identity.

Biometric-based systems verify a user based on various parts of the human body or human behaviour. Behavioural biometrics usually consists of a user acting such as speaking (voice), writing a signature and through walking. Physiological biometrics commonly uses face, fingerprint, iris and hand geometry. In general, biometrics modalities are based on the following criteria:

- Universality – every person should have it
- Uniqueness – it cannot be replicated in two people
- Permanence – it cannot alter greatly with time
- Measurability – it can be measured in some form

Combining biometric-based systems with traditional methods such as PIN codes, passwords and keys provide an extra level of security. In ABC systems, it is a common scenario for a user to present their passport (token) which contains a facial reference image. The token has been pre-validated by the country of issue and is a document that verifies the user's identity. The image is stored locally on the electronic chip embedded within the passport, and a live image is captured from the user which is then compared by a matching algorithm. If the system accepts the match and the documents pass other security checks, the user is authorised by the system and therefore can cross the country's borders.

Enabling a biometric component within a system has its advantages over using token only based methods. For example, standard security methods which may require the user to remember a PIN or password can easily be forgotten. Also, carrying cumbersome bunches of keys, tokens or cards can be easily lost or stolen. Biometrics guarantee that the user who accesses facilities cannot deny using it (non-repudiation) and in several scenarios, do not require the possession of any physical tokens, nor rely on the uncertainty of the human memory. However, passports may by-pass this rule; the traveller must carry his or her passport when travelling to another country. Despite obvious advantages to using biometric components in security systems, biometric systems do have some disadvantages and can be vulnerable to several issues; imposters, attacks and erroneous interaction with a sensor. Incorrect interactions occur due to a number of reasons but can range from when a user does not know what they're doing (confusion), tiredness or simply not knowing they have performed an unwanted action.

1.2 Motivation

ABC systems largely verify a user based on passport interaction and a facial presentation to a camera. Some systems rely on other biometric modalities such as fingerprints while others rely on different types of tokens such as electronic national identity cards and pre-registered cards, however other combinations do exist.

Systems that use a combination of an electronic passport (sometimes referred to as an ePassport) and face biometrics account for most ABC systems. As the face is usually visible, it is the easiest biometric sample to capture.

ABC recognition systems are highly influenced by the physical designs and topologies of the implementation. Solutions using facial modalities, for example, are directly affected by certain components of the system such as the placement of the camera and the use of feet symbols, whereby icons are placed on the ground to ensure that a user is standing within range of the camera. Users not standing correctly may be out of position, possibly resulting in a longer transaction or an erroneous result.

A major drawback to these automated processes is that they are reliant on compliant reference images stored on the token or database. Not only must the enrolled image meet certain standards, but the captured image from the biometric component of the system must also be meet similar specifications. While this is to be expected, the way in which these images are captured, compared and assessed differ from country to country and from airport to airport.

Several factors affect the captured image quality, which can be classified according to the relation with the users (physical and behavioural), user-sensor interactions (environmental and operational), acquisition sensors and processing systems. Several factors, however, should not contribute directly to the quality of the image:

- Operational environmental (e.g. no interference from daylight)
- Background and object occlusion
- Temperature and humidity
- Illumination and light reflection
- User’s age, gender, ethnic origin, skin conditions

Also, performance assessment within these systems typically report on a small number of measurements, such as throughput and standard biometric error ratings. The result is a simplified measurement that will not necessarily identify why an error occurred. Several groups are working on improving analysis of the performance of biometric implementations from a user’s perspective but there is little research to identify both the task interaction between user and sensor. Typical methodologies focus on the user, looking at the user’s (subjective) satisfaction of a process and how efficient and effective the system may be perceived.

Furthermore, the introduction of new technology in the Action Recognition field has enabled human recognition in a wide range of applications. One area looks at skeletal tracking using an RGB-D camera, using depth data to map a human's movement by classifying a behaviour. Sensors that can measure the user's location within a specified area can be employed for a variety of scenarios such as for gaming and physiotherapy purposes. Many affordable devices are available, increasing the popularity of using these tools for research purposes. In addition, there are plenty of applications that could benefit from identifying user movements but have not yet been studied in any detail.

Border control scenarios are designed to monitor a single user in a straightforward process; the user moves forward, enters a token, submits a biometric sample and exits the system. The placement of the camera and screen displaying the information is positioned to face the user and capture a frontal image. If the RGB-D camera can distinguish the position, body movement and pose for a presentation, specific feedback could be presented to guide the user through the process. For example, if the user is looking up or down the image may not meet the required standards. It would also impact transaction time and add to the already lengthy queues at busy airports as the user may delay the system's ability to capture. Additionally, if the behaviour continued without being corrected, the supervising border guard might have to stop and redirect the user to either restart the process or divert to manual control.

The motivation for this work, therefore, seeks to improve the overall transaction process when considering the user through their interaction with a system. Exploring the application of skeletal tracking within border control scenarios, a sensor that can monitor the user's movement is investigated. The proposed application can then also automate the performance assessment, whereby an unsuccessful interaction can be tracked and analysed, and possibly identifying key bottlenecks within a system where common mistakes occur.

1.3 Aims and Objectives

The general purpose of this research is to make recommendations for improving the border control process using a robust and efficient tracking system to enhance throughput, sample quality and in the reduction of user interaction errors. This work aims to explore the effectiveness of such systems to not only utilise tracking to improve performance assessment but to assess if conveying specific information back to the user when an erroneous presentation is detected benefits the process. The specific objectives of the research are to:

- Review the literature on biometric and ABC systems as well as user interaction assessment methodologies
- Propose an automated performance evaluation framework to facilitate the use of an activity tracking device
- To explore wherever the information obtained by the device will benefit in the performance assessment
- To explore the effect of information and guidance on user-interaction
- Make recommendations based on the literature, simulations of a border control scenario and results from surveys and questionnaires

1.4 Scope of the Project

The list of the work that will be carried out in this study is summarised below and explains the areas that will be covered in this research. The areas which will not be referred to in this research are also listed.

This study will only explore user interaction and tracking processing methods within a biometric interaction context. The research will look at the effect of user interaction in two biometric modalities; fingerprint and facial recognition. In this study, skeletal tracking will be explored within a self-service biometric scenario, seeking to replicate as many variables as possible when comparing to a live implementation.

This work will not explore passport interaction. While passport interaction is a fundamental component of ABC systems, a high-end passport reader which can read both the MRZ and RFID elements of the passport was not available for this study.

The research will investigate features and relationships based on user interaction performance and biometric modality. Through tracking methods whereby certain behaviours can be classified, the effect of feedback on erroneous presentations will also be studied.

1.5 Structure of the Thesis

The organisation of the thesis is given below:

- Chapter 2 explores the General Biometric Model, the Human-Biometric Sensor Interaction (HBSI) framework and an overview of Automated Border Control systems. The ABC section reports on biometric modalities, the use of tokens, border control performance assessment methodologies and current obstacles
- Chapter 3 proposes a framework to assess automated border control performance based on identifying general steps in systems and analysing further steps to break down the identity claim process
- Chapter 4 looks at the application of the Kinect sensor, a tracking device commonly used for Action Recognition in gaming and physiotherapy activities. Specifically, the accuracy and robustness of the device are investigated to ensure that the data captured is an accurate representation of the user. The success of the results will directly affect the data collections introduced in the next chapters
- Chapters 5 and 6 explore fingerprint and facial interaction. The use of the tracking system is explored further, and the automated method of performance assessment is reviewed for both modalities. In addition, variables relating to the border control process are investigated and studied
- Chapter 7 presents recommendations based on the findings of the previous chapters to improve current processes. This chapter explores information flow, biometric modalities and opportunities for tracking and image processing elements for not only automated border control systems but biometric processes in general. Conclusions, a summary of the contributions of this work and suggestions for future work are also provided in Chapter 7

CHAPTER 2. LITERATURE REVIEW

2.1 Introduction

The following review of literature is composed of three main sections. The first section discusses the General Biometric Model and introduces the process flow through the biometric system, highlighting the capture, matching and decision processes. Traditional system performance measurements are then reported followed by presenting the concept of user interaction and the importance it holds in biometric systems.

The second sections discuss The Human-Biometric Sensor Interaction (HBSI) framework and the history behind the model. The third section discusses the history of border control solutions and the use of the travel documents, investigating specific systems and the associated biometric modalities. Furthermore, the design, other contributing factors and a discussion of the user's acceptance of biometrics are discussed.

2.2 Generic Biometric Model

The General Biometric Model outlines the general process for capturing, matching and deciding wherever a biometric is accepted for verification or recognition.

This section is broken down to the process flow of the model, traditional assessment methods of system performance and key insights into user interaction.

2.2.1 Process Flow

The process flow of the General Biometric Model outlines the process for capturing, matching and deciding the output of the sample.

Figure 1 demonstrates the model. Typically, each subsystem contributes to the recognition process carrying out a task.

The Data capture subsystem is composed of the biometric capture device and will change based on the modality being used. This process requires a presentation to be made by the user. The user's characteristics are presented to a given sensor, which yields the system's input data based on the biometric measure and the technical features of the sensor [1].

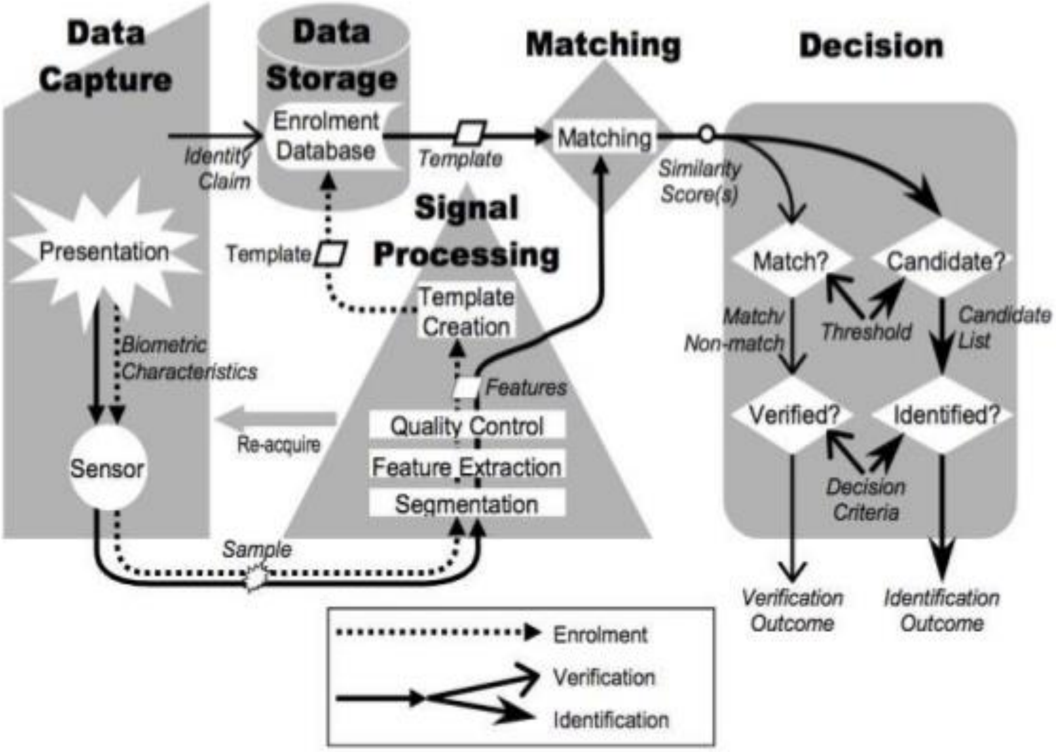

Figure 1: Generic Biometric Model

The Signal Processing subsystem generates a vector from the biometric sample. There are several stages involved:

- Quality Control – checks that the sample meets a predefined set of quality specifications with a goal of ensuring that feature extraction and segmentation will be successful
- Feature extraction – essential features (depending on the biometric used) is extracted and localised
- Segmentation – other information is localised such as detection, alignment, sample segmentation, normalisation and enhancement

Once the features have been successfully obtained, the template is created and sent to different parts of the system depending on the required function. In enrolment processes, the signal processing subsystem creates a biometric reference from the features and is sent to the enrolment database in the data store subsystem. This subsystem changes depending on the specifications for that system; data may be stored in a distributed or centralised database.

Verification or Identification functions are sent to the matching and decision subsystem. The matching subsystem compares a feature vector to a single biometric reference for verification purposes, or for identification purposes, several biometric references. A similarity score is produced for verification and similarity scores for a list of potential users for identification.

The decision subsystem, therefore, decides wherever the sample can be verified or identified based on set threshold and criteria. For verification systems, the result will either accept or reject the user who claims his/her identity. Identification systems will produce an output of candidate lists which contains the user's identifiers for those whom biometric references match the sample list. This could either be an empty list or a list with a fixed number of users identified.

2.2.2 System Performance

Biometric performance refers to the recognition accuracy and speed, the resources required to achieve that desired recognition and speed, as well as the operation and environmental factors that affect accuracy and speed of a biometric system [2]. Performance metrics are captured throughout the entire Generic Biometric Model process.

The International Organization for Standardization defines ISO/IEC 19795, a document which describes global standards for biometric performance testing and reporting [3] define two types of mandatory metrics that all systems must be able to report: error and throughput rates.

2.*2.2*.1 Error Rates

Recording error rates are useful for quantifying the accuracy of the system. These rates measure the number of errors that occur during biometric sample acquisition; it’s processing and the comparison with the biometric template and further decision attempts. There are two error rates which are reported during acquisition and signal processing processes:

- Failure to Acquire (FTA), a measurement of samples which the device failed to acquire. It can indicate issues of user performance [4]
- Failure to Enrol (FTE), the rate at which attempts to create a template from an input are unsuccessful, will point to the success of individuals to interact with a system [4]

There are two error rates for the comparison and decision subsystem process:

- False Non-Match Rate (FNMR) – the rate of samples, acquired through genuine attempts, which are falsely declared not to match a biometric reference of the same characteristic from the same user who provided the original sample
- False Match Rate (FMR) – the rate of samples, acquired from zero-effort imposter attempts, which are falsely declared to match the compared non-self biometric reference

Conventionally the overall performance of all biometric systems, including ABC implementations, monitor two key rates:

- False Rejection Rate (FRR) - the percentage of false rejections made by a system
- False Acceptance Rate (FAR) - the measure of performance that a biometric system will incorrectly accept an access attempt by a non-authorised user

2.*2.2*.2 Throughput Rates

Throughput rates measure the speed of use for a system, reporting on the number of users that can be processed per unit time based on computational speed and human-machine interaction. While ISO/IEC 19759 does not define a specific metric, nearly all systems report on the measurement of time spent on user interaction and processing speed. Several measures are typically recorded as evidenced by multiple studies [5] [6] [7] these are usually: enrolment (time taken for an image sample to be captured) and recognition duration time (time taken to perform matching) which is typically expressed in seconds, as well as a measurement of speed of the human-machine interaction, which should indicate when a user starts an interaction to ending it.

2.2.2.3 User Interaction

The term ‘User Interaction’ can be defined by how the user acts on the system and how the system acts on the user [8].

All biometric systems require the user to present a sample in some shape or form to a sensor. Depending on the modality and the design of the system, the process may be intrusive or obtrusive, but for all systems, will require a certain biometric to be presented in an accurate and timely manner.

Depending on the modality chosen, a successful capture for a sample will either require a movement requiring the user to act in a certain manner (e.g. the flow of the arm/hand for a signature) or a physiological template (e.g. presenting their iris or fingerprint in a certain way to a sensor). Both require the user to exhibit the desired behaviour. While systems can relay information and instructions to the user during this interaction to aid the user, the success of this process will *usually* rely entirely on correct user input.

Attempts made to a sensor, therefore, can be observed and categorised as either ‘correct’ or ‘incorrect’. Common incorrect interactions could be where users are presenting the wrong finger to a sensor, closing an eye during an iris scan or looking away from the camera during facial recognition. Correct presentations, on the other hand, could be classified when a required behaviour or sample is presented, e.g. both eyes are open for face recognition; the right finger is captured successfully in fingerprint verification.

Improvements in technology have enabled systems to improve the capture of a presentation even though an incorrect ‘behaviour’ has been performed. For example, in border control systems, multiple images are captured and the highest quality image is selected for matching purposes [9] [10].

It is common to assume the term ‘User Interaction’ relates specifically to Human-Computer Interaction (HCI), the field which explores how human beings interact with computational devices. In most cases, HCI specifically investigates interaction with User Interfaces (UI) and HCI-based systems [11].

HCI usually involves the study, planning and design of the interaction between users and computers. It stems from Interaction Design [12] [13], which is defined as the practice of:

- Understanding user's requirements and goals
- Designing tools for users to achieve those goals
- Envisioning all states and transitions of the system
- Considering the limitations of the user’s environment and technology

Although there is plenty of research available in the area within its field, HCI is not yet fully reported when regarding biometric systems, especially when considering the interaction in self-service systems. HCI might specifically relate to biometric systems when the HCI interaction is based on a biometric, e.g. identifying a user through analysing their use of an input device such as a keyboard or mouse. HCI based biometrics can be divided into two main categories, Direct and Indirect, according to Saaed [14] and Yampolskiy [15]. Direct interactions consider human interaction with input devices such as keyboards and mice; Indirect collects information from system calls, audit logs and GUI interaction.

There is, then, a difference between assessing how the user interacts with the system and assessing a user through the interaction. The former is reported through Usability Testing, which focuses on the user and aims to assess user satisfaction with a system. While HCI does influence biometric systems, there are no standardised definitions or methodologies that apply. Belen Fernandez-Saavedra reports [16] that HCI influence of performance is focused on system performance and is not widely applied to the overall area. Several institutions work in both areas, the NIST group, the HBSI project and the University of Carlos III Madrid are considered the main contributors to the area. Belen’s research primarily focuses on the Human-Biometric Interaction, considering the many factors that influence performance.

Research into defining User Interaction specifically within a biometric context has largely been led by the Human-Biometric Sensor Interaction (HBSI) project.

2.3 The Human-Biometric Sensor Interaction (HBSI) Model

The Human-Biometric Sensor Interaction (HBSI) Model illustrates how metrics measured from biometrics sensors (such as sample quality and system performance) can be tied to ergonomics (physical and cognitive) and usability (effectiveness, efficiency and satisfaction) to evaluate the overall performance of a biometric system. Applying this framework to a system often provides a better understanding of what affects a biometric systems performance.

Over the past nine years, the initial team of Kukula and Elliott developed one of the first models that linked usability and biometrics. The model (Figure 2 below) has its origins at the intersection of usability, human factors, and image quality/performance [17] [18] [19].

Initial work discussed the issue of hand placement in hand geometry systems [20] [21], based on evidence collected during a biometric feasibility study. The first HBSI model then was shortly introduced in 2005 and continues to build on previous research in the area of human-biometric device interaction.

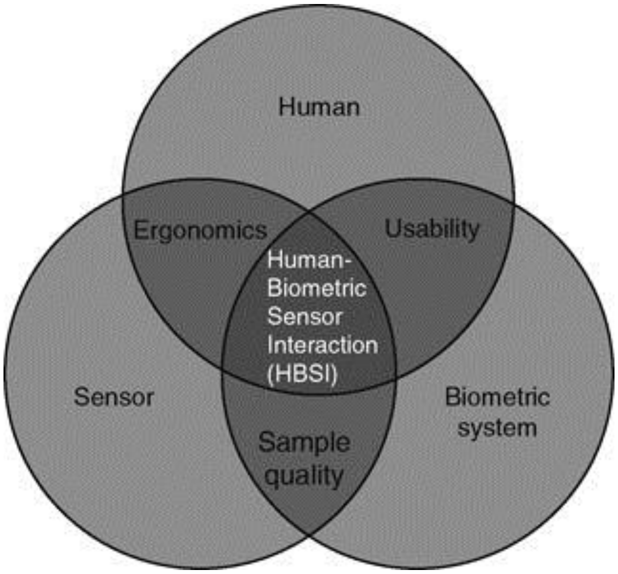

Figure 2: The Original HBSI Model [22]

The model has been validated against several modalities over the next decade such as Fingerprints [22] [23], Iris [24] and Signatures [25]. The next generation of the model has adapted to consider intelligent sensors [26] [27] – those that have some signal processing, sample quality and feature extraction intelligence during detection and acquisition. Recently the model has begun to consider the introduction of other authentication methods, such as the ePassport at an ABC gate. S. Elliott reports on the latest research and process model in a recent report [26].

Six different types of metrics were developed based on the HBSI model. The HBSI Interaction Framework [23] (Figure 3 below) comprises of; Defective interaction (DI), Concealed Interaction (CI), and False Interaction (FI), which are all based on incorrect presentations. For example, if an individual interacts with the sensor incorrectly, and the sensor does not “see” this interaction, then the framework defines this as a Defective Interaction. In this case, it is not the sensor’s “fault”, but further action must take place to consider why this happened. Separating a DI from a traditional Failure to Detect (FTD) is crucial to understand if it was a user or system generated an error.

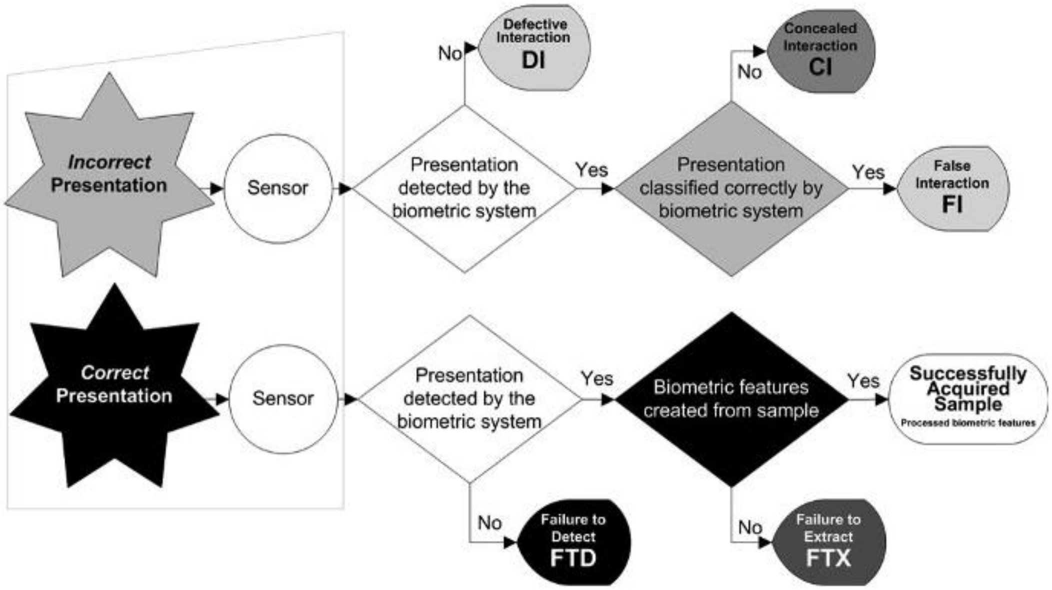

Figure 3: HBSI Interaction Framework

Concealed Interactions occur when the subject presents an incorrect biometric sample and is accepted by the system as a correct sample. For example, in a single fingerprint data collection, this could be when the system is expecting a middle finger, and the subject presents a ring finger. The system accepts the presentation and subsequently, processes it and stores it as such. Concealed Interactions are incredibly difficult to

categories without confirmation that the presentation was incorrect. If a user interacts with a device in a real-life scenario without human oversight, who is to say that the error was due to an incorrect or that the features were not processed?
False Interactions is when the system provides feedback to the user of an incorrect presentation and is the ideal result in an erroneous scenario. The system correctly handles the sample as an error and displays information to the user. Typically, this will require the user to restart the process, meaning more time and effort from the user.

Overall defining these metrics in the evaluation performance of the system will enable a deeper understanding of the reported metrics such as sample quality and throughput. Although the framework provides a benefit in this regard, it does lack, however, the ability to automatically detect an incorrect or correct presentation during the interaction. The HBSI Interaction Framework is discussed further in Section 2.3.2.

2.3.1 HBSI Evaluation Framework

The Evaluation Framework was developed shortly after the conception of the HBSI model and outlines the measurements for each intersection introduced in the model. The framework is developed through Kukula's thesis [28] and combines several disciplines that have been well researched and documented.

HBSI Evaluation, then, as previously discussed, can be used as an extension to analyse biometric system performance in a much wider sense than that proposed in the Generic Biometric Model.

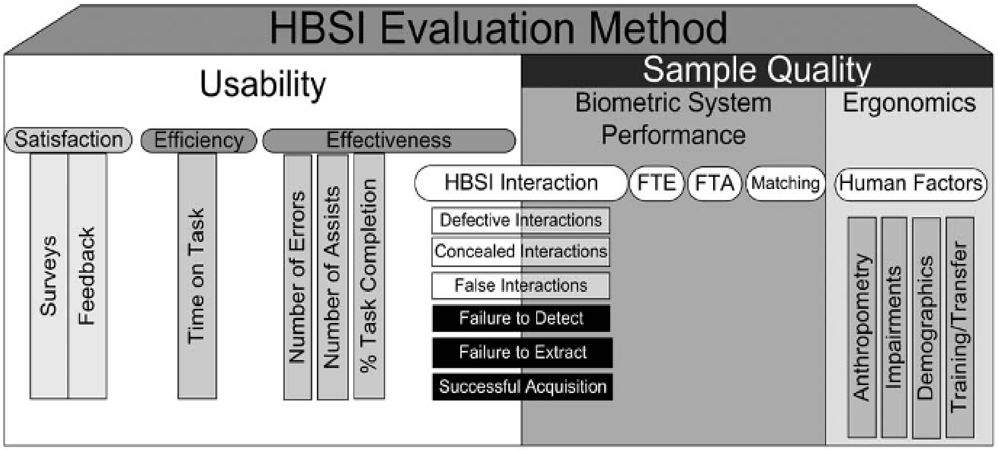

Figure 4: HBSI Evaluation Method Version 1

The Evaluation Method (Figure 4 above) considers a range of measurements that are obtained through presentation and acquisition. Usability assessment is completed by analysing user satisfaction, efficiency and effectiveness. Sample quality considers traditional metrics such as FTE, FTA and Matching Scores. Alongside system performance, human factors may also affect sample quality, such as demographic information and anthropometry factors. At the heart of the Evaluation Method, lies the HBSI Interaction Framework that considers the categorisations of a user presentation to a single sensor.

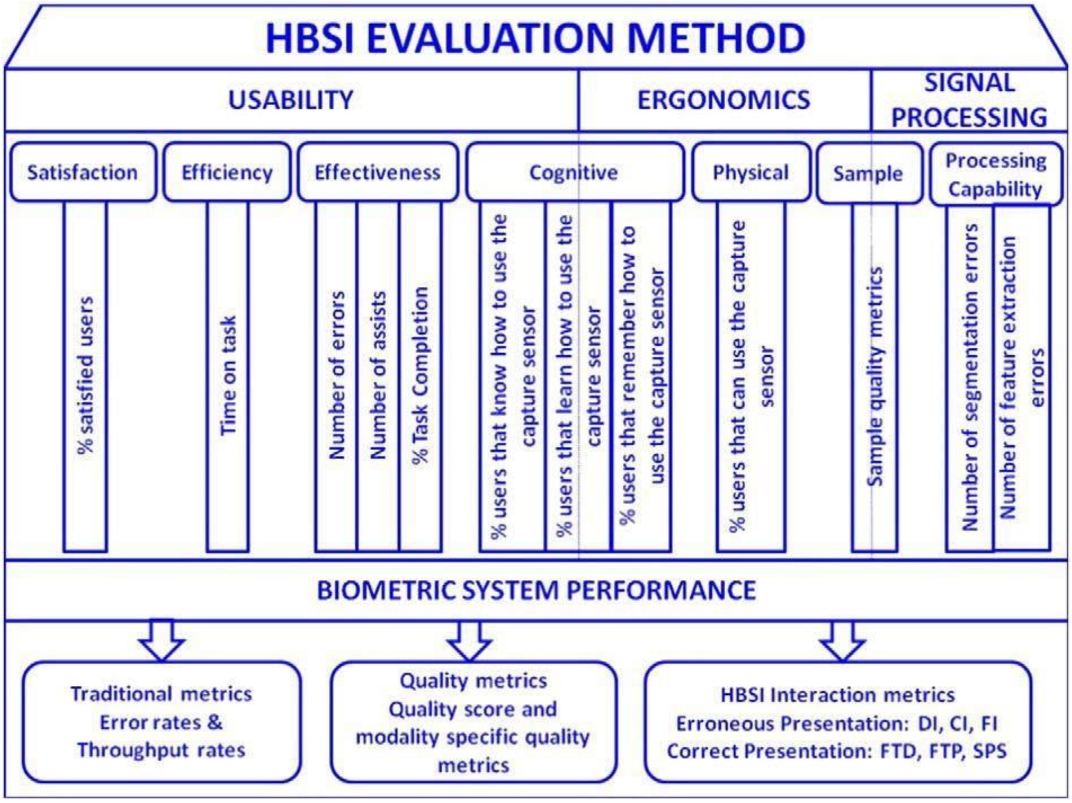

Figure 5: HBSI Evaluation Method Version 2

The framework then supports a broad range of metrics that are used throughout the development and implementation stages of a biometric system. The level of detail that this assessment can report could indicate and support deeper, and more dangerously, overlooked issues. For example, in ABC systems throughput rate is one of the most important factors, directly reporting on the efficiency of a process. If a user presents their finger to a fingerprint sensor a total of five times, but only one presentation is detected, then the system would traditionally report four FTDs. However, if the system was being evaluated further and was also looking at the time on task, the evaluator may question why one attempt took so long. In this scenario, there would be no data to analyse and determine the cause.

The model has evolved since 2010 and has been updated slightly to consider several other variables. Figure 5 above demonstrates the latest version of the HBSI Evaluation Method [27].

Version two decomposes usability, ergonomic and sample quality into further specific metrics, considering modality specific as well as HBSI and traditional metrics. Sample quality metrics here has been combined with signal processing, where the process capability of the system can also be measured.

2.3.1.1 Sample Quality

Different biometric modalities comprise of their own quality metrics and scores. A range of biometric standards published by the International Organisation for Standardization (ISO) and by the Electrotechnical Commission (IEC) define characteristics that all biometric systems must adhere to ABC systems or otherwise. It is important to note that while these standards do state-specific requirements, vendors are free to design and implement their quality assessment methods into systems if they adhere to the standards for that biometric modality.

For the quality of a face image, there have been many proposed quality assessment methods [29] [30] [31] [32]. An advantage in ABC systems is that it is usually possible to capture multiple face images from each subject and select the face image with the highest quality [33]. Most ABC systems must, however, adhere to several ISO standards which denote several quality metrics.

The ISO/IEC standard 19794-5 “Standards for Biometric Data Interchange Formats (Face Image Data)” [34] reference detailed instructions for lighting, facial pose, focus and so on for taking face photos in biometric systems. These standards must be adhered to when taking photos for both enrolment and verification stages of the process. Requirements are displayed below in Table 1. A further discussion of these standards is reported in Chapter 6, Facial Recognition.

Non-standard lighting or pose and out of focus are among the main reasons responsible for the performance degradation in face capture systems [30] [35]. Research in improvements into algorithms has investigated the robustness of several implementations including ABC systems [36].

Table 1. Face Image Requirements in ISO/IEC 19794-5

Clause	Attribute	Constraint
Scene	Posture	Control on deviation from frontal
	Illumination	Uniformly illuminated with no shadow
	Background	Plain light coloured
	Eyes	Open and visible
	Glasses	No flash reflections, dark tint or heavy frames
	Mouth	Closed and visible
Photographic	Head position	Placed in the centre
	Distance to camera	Moderated head size
	Colour	Colour neutral and no red eye
	Exposure	Appropriate brightness
Digital	Focus	No out-of-focus and in excellent sharpness
	Resolution	Width constraint of the head

The ISO/IEC 29794-5 Face Image Data standard [37] refers to specified methodologies for computation of objective and quantitative quality scores for facial images that are utilised in ABC systems. The document details approach to determine certain characteristics, such as facial symmetry, resolution and size.

The standard also suggests that facial quality can be categorised into the static subject and dynamic subject characteristics as demonstrated in Table 2 below. Different factors affect the quality of the image; static characteristics relate to anatomical features of the subject (head dimensions, eye positions) while dynamic characteristics consider subject related behaviours during the acquisition process (eyes open, pose).

Also, other static and dynamic characteristics are considered but mainly relate to properties to do with the build and environment of the system; background, the influence of lighting and camera characteristics (resolution).

ISO/IEC 19794-5 and 29794-5 are reviewed in more detail by J. Sang et al. [35]. The research reviews methods to tackle both static and dynamic features. Several algorithms are introduced for face image quality assessment including Gabor-Based Facial Symmetry, to evaluate changing illumination and improper posture, and DCT-Based Sharpness, to discern out-of-focus.

Table 2: Static and Dynamic Features considered in ISO/IEC 29794-5

Static Features	Dynamic Features
Biological: Anatomical characteristics (e.g. head dimensions, eye positions) Injuries and scars Ethnic Group Impairment Other factors: Heavy facial wears, thick or dark glasses Makeup Jewellery	Subject Behaviours: Closed Eyes Expression (exaggerated, smiling etc.) Hair across eyes Head Pose Subject Posing (frontal/non-frontal to camera)

While there are many strides to improve these systems from an algorithm point of view, ultimately the capture process differs between systems. Additionally, the matching process between a stored image on a passport or a token between a captured image also changes, making it extremely difficult for standardisation in the performance assessment of these systems.

Fingerprint scanners must produce images that exhibit good geometric fidelity, sharpness, detail rendition, grey-level uniformity and grey-scale dynamic range, with low noise characteristics as reported by M. Carmen at al. [38]. The required sample quality of fingerprint images is defined through several ISO/IEC standards. Following on the 19794 series, part 1 [39] defines that the fingerprint scanner produces of a certain standard based on image resolution, size, grey level colour range, sample rate, light intensity and signal to noise ratio. Parts 2-4 [40] [41] [42] specifies minutiae data, pattern spectral data and image data standards. Several studies investigate the accuracy and potential issues based on these formats [43] [44] [45]. The sample quality of fingerprints is discussed further in Chapter 5.

Matching algorithms of fingerprints largely used minutiae-based features, particularly restricted to two types of minutiae points; bifurcation and ridge endings. Several studies have investigated sample quality measures using these characteristics [46] [47] [48].

2.3.1.2 Usability

Performance assessment concerning the interaction with devices (including biometric systems) is assessed from either a user perspective or by the effect on system performance through interaction. The usability community, in general, is concerned with the assessment of a system through reporting efficiency, effectiveness and user satisfaction [49] [50] from a users point of view.

The term usability is defined by ISO 92411-11 [51] by the extent to which a product, biometric or otherwise, can be used by subjects to achieve their goals. It can be assessed according to three criteria: efficiency, effectiveness and user satisfaction. Regarding an ABC system, it is possible to define task performance as effective when an interaction supports users who can achieve their goal of successfully crossing a border (including the sub-tasks of token reading and biometric verification). The interaction with the system is considered efficient if the traveller can pass through the process promptly, which is subjective to an individual user but averages at around 15-20 seconds for European ABC configurations [52] [53] [7] [54]. A user's (subjective) satisfaction can depend on the level of the physical or mental workload that they may encounter throughout the process.

Research in usability evaluation has been largely led by the National Institute of Standards and Technology (NIST), who have contributed significantly to studying the usability of a wide range of systems [50] [55] [56] [5] [57] [58] [59]. Other studies in the area [60] [16] [61] [62] have investigated the influence of usability factors that affect biometric performance and user experience in some similar applications.

Several NIST studies have investigated the impact of many variables on performance in biometric systems. Choong et al. reported on several studies on ten-print fingerprint capture within a US manual border crossing scenario [55] [5] [63] [64]. Variables included: the height of the kiosk, angle of the sensor and impact of information.

Usability metrics are reported throughout all four studies on efficiency (time on task), effectiveness (task success and quality of the fingerprint) and user satisfaction (post-task questionnaires looking at binary answers as well as comments). The first study [63] investigated the impact of information on user performance using a poster, video, and verbal instructions. Participants who received verbal and video instructions outperformed users who were shown a poster, resulting in fewer errors and a quicker transaction. Another study from the NIST group report on the use of face overlay in facial verification systems [65] which indicated the use of an overlay image improved the quality of capture face images.

Usability is also closely linked to other issues which have been defined throughout literature. Acceptability testing or user acceptance testing, analyses how users can accept the use of a specific biometric characteristic, method or system for biometric recognition [66]. The ergonomic design focuses on the area of interaction between the user and the biometric system; analysing tasks, movements, and user behaviours [67].

There are many tools used in usability testing. NIST has provided a handbook on 'Usability and Biometrics' [68], outlining a user-centred design approach to aid the design and development of biometric technology systems. Common usability evaluation methods that are detailed include; cognitive walkthroughs, contextual inquiries, requirements analysis, user and task analysis and user evaluation.

Distinguishing the difference between usability and user interaction is important. Whereas usability defines how usable a system may be; whether that is overall or for a process or task – user interaction is defined as a combination of movements that result in either a successful or unsuccessful presentation made to a sensor [69].

2.3.1.3 Ergonomics

The Ergonomic or 'Human' Factors in HBSI is the study to achieve an optimal relationship between human and machines in an environment. HBSI has previously looked at ergonomic design to adopt a system to a user, rather than adapt the user to the system [70] [67] which is a common design concern for many implementations.

HBSI research has defined several relationships between the user, environment and the outcome of the algorithm [18] [70]. Figure 6 below demonstrates.

Inter-relationships between these groups may impact biometric performance. The user-environment relationship will depend on variables such as clothing (e.g. protective equipment used in some scenarios may affect the ability to capture a sample) and temperature and humidity, which may impact the skin also affect the acquisition for some modalities. The environment to algorithm relationship could be influenced by external factors such as noise, illumination or busy backgrounds. Finally, the user to algorithm relationship may be affected by physiological factors such as skin age, colour and moisture or behavioural factors such as finger preference which in part can affect recognition. Also, social factors such as hair length or wearing head coverings can impact facial and iris recognition due to the occlusion of necessary features.

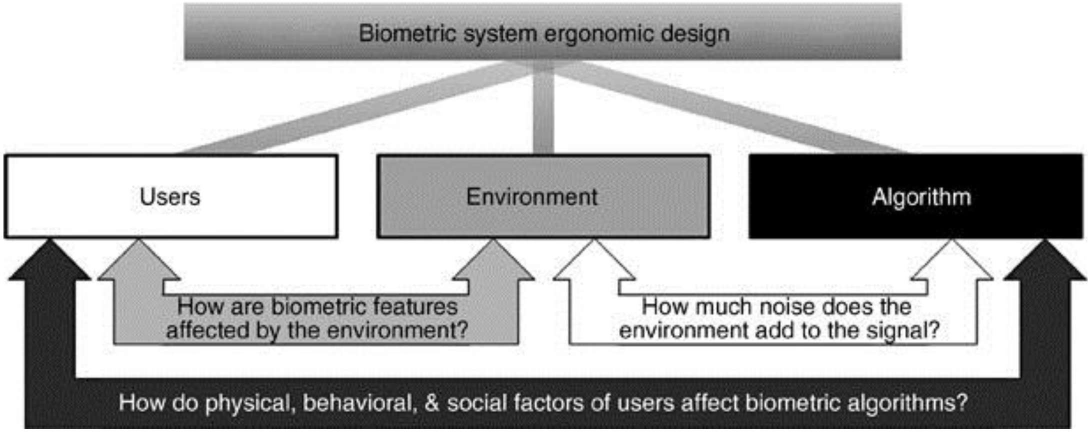

Figure 6: HBSI Ergonomic Design Factors

Biometric systems are heavily dependent on the sensors ability to acquire a sample, segment and extract features from (multiple) samples to determine the correct output. Common design concerns for any biometric systems include attributing accuracy, scale (size of the user base) and usability [2] [71]. In general, ABC systems are designed and implemented to overcome many of these issues raised. However, when assessing the deployment of these systems, it is important to analyse both ergonomic and usability factors.

2.3.2 HBSI Interaction Framework

The HBSI Evaluation Framework considers the individual interaction made with a biometric sensor. The process allows for an understanding of correct and incorrect behaviours typically occurring within a biometric system. Correct presentations for interaction can be categorised as either:

- Failure to Detects (FTD) are correct presentations that are not detected by the system
- Failure to Process (FTP) within biometric systems can occur due to reasons such as problems in segmentation, feature extraction or quality control and is a system error generated by the biometric system
- Successfully Processed Sample (SPS) is the ‘correct’ transaction which results from a correct presentation and successful processing

There are three possible categorisations of incorrect presentations:

- Defective Interactions (DI) which occur when a biometric sample is incorrectly presented and is not detected by the system
- Concealed Interactions (CI) occur when an incorrect presentation is detected by the system but is not handled correctly as an error. An example could be in fingerprint recognition where a user, for whatever reason, uses a different finger from that of the enrolled one but is still accepted by the system
- False Interactions (FI) occur when a user erroneously presents their biometric, and the system correctly identifies the error as an incorrect presentation

While the Interaction Framework provides a full range of categorisations, its drawbacks lie in requiring manual confirmation of errors. During data collections, the interactions are coded and recorded by the researcher as the study progresses. The framework is designed to be used for generic purposes and in its current state, for a single modality only. Multiple interactions may require a claim of identity, adding an element of the framework that is yet to be explored.

Recent work on the HBSI Model has investigated token presentations made to a sensor, creating a process chart that allows the categorisation of False Claims and Potential Attacks [26].

2.3.3 The Full HBSI Model

To provide practitioners and researchers with components that allow the assessment of operational times, false claims, attack, and token presentations, new sub-models of the HBSI presentation framework have been developed [69]. These new models have been integrated comfortably within the HBSI Interaction Framework to produce the full HBSI model, allowing a wide range of categorisations to be made within an identity claim scenario.

The Operational Times Model (Figure 7), reported in previous HBSI research [72] defines the transaction time that is required to use a biometric system and segments the presentation process into individual tasks, demonstrating the token and biometric presentation. The research presented by Brockly lays the foundations in automating transaction time's posthoc without the need for a human operator. Although the Operational Times Model was developed, it has not been applied in line with the Full HBSI framework.

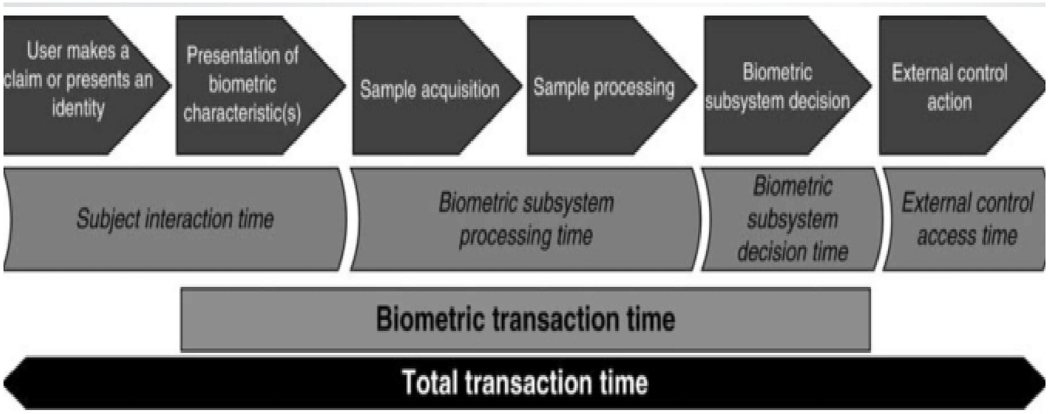

Figure 7: Operational Times Model [72]

The Full HBSI Model (Figure 8 below) accounts for systems that allow for one or more factors of authentication. An example of a one-factor authentication system would be a single biometric interaction process, while a multi-factor system may be any combination of a token, password, and biometric sample(s). This version of the model works to include token, attack and false presentations, looping to the start of the process (if necessary) once a method of authentication has been completed.

Trending technologies that implement anti-spoofing or liveness detection components and the ability to flag potential attacks to a figure of authority for subsequent processing were originally not considered in the original implementation of the HBSI Model. Therefore, an advantage of using the Full HBSI Model allows the categorisations of potential false claims and attack presentations, which are both introduced as separate models below.

The False Claim HBSI Model (Figure 9) occurs when an identity claim is made that does not belong to the user, and the system no longer requires an additional factor of authentication. This is needed in the event of an "accidental impostor presentation". For example, in a scenario of individuals travelling together, they could accidentally swap passports and therefore present an incorrect identity claim to the system.

Although this is not defined as a malicious attack, it must be classified as an invalid claim of identity, and this is when the False Claim Model is used. The model uses the same decisions as the HBSI Interaction Framework, but specific to identity claims made by the incorrect user. False Claims error metrics are denoted by the subscript FC.

In the case of an ABC system, personnel are employed and trained to supervise multiple transactions from different users and are expected to handle exceptions where applicable. For example, if a False Claim is made (e.g. an accidental swapping of the passport) and the system can detect and subsequently flag the claim to the border guard, then personnel will intervene and action the sample as either a Refused Sample or Forwarded Sample. It will be important for systems to be able to classify false claims as this could lead to breaches of security.

Systems involving some form of anti-spoofing or liveness detection will leverage the Attack Presentation HBSI Model (Figure 10). The HBSI Attack Presentation Model confirms that the biometric sample is detected, attempts to classify it as an attack sample, and determines if the presentation is suitable for matching to save the sample. If the biometric subsystem classifies the presentation as an attack, it either flags and forwards the sample to the respective authority or simply flags the sample and refuses it. If the presentation is not classified as an attack, it can achieve one of three attack HBSI error metrics.

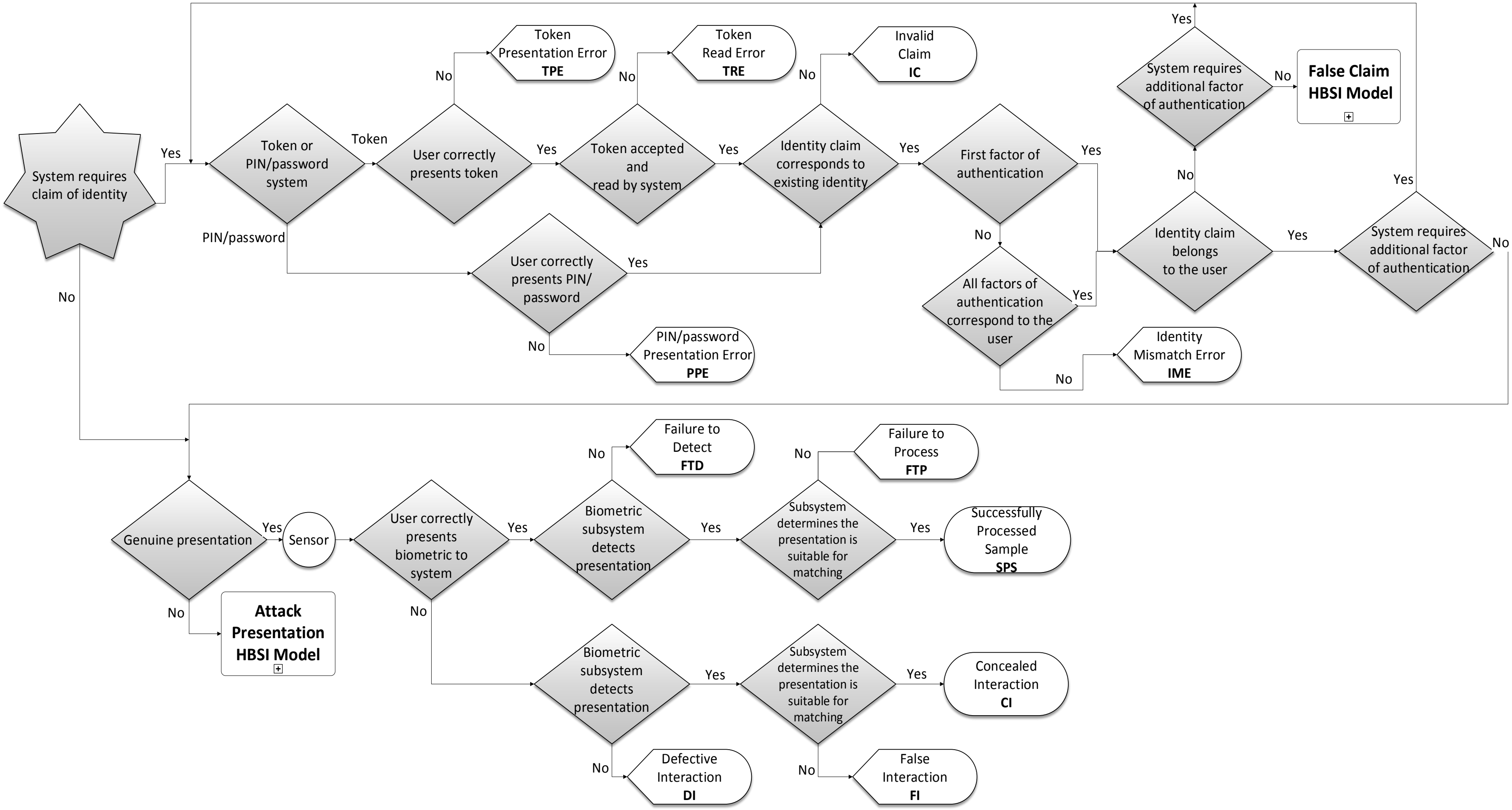

Figure 8: The Full HBSI Model

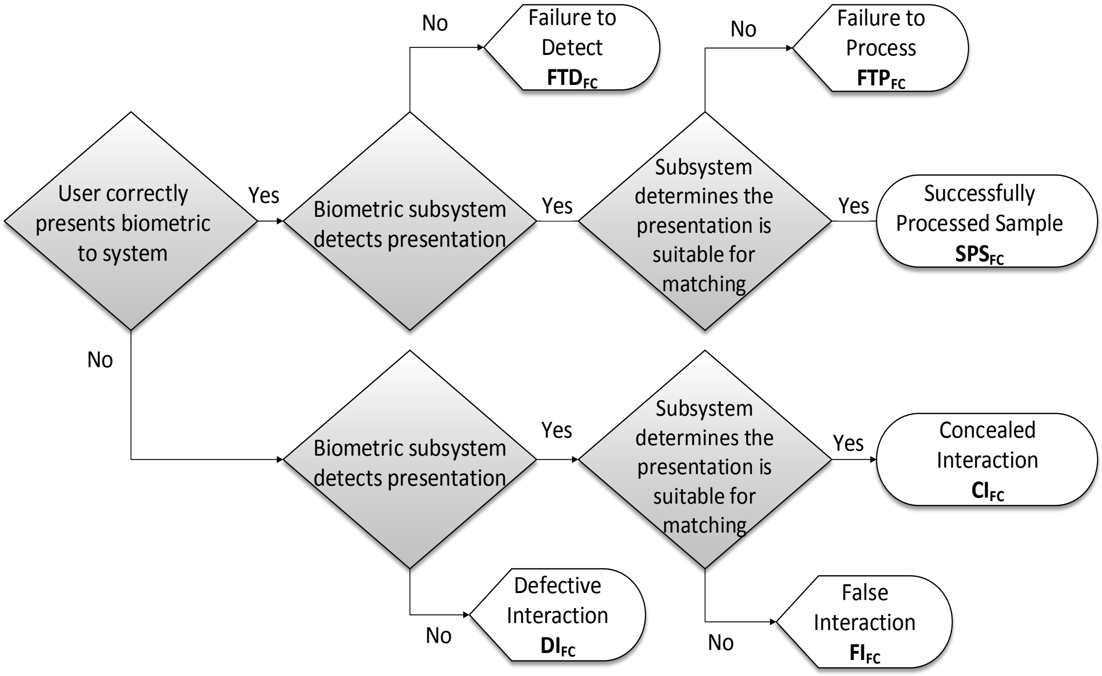

Figure 9: The HBSI False Claim Model

While improvements to hardware and software are continuously being developed to counter specific threats and certain types of attacks in large-scale biometrics systems; there is the underlying issue of the possibility of identity attacks.

Where an identity claim is required, the possible outcome of allowing an attack sample through could have devastating consequences in these systems. The HBSI Attack Model demonstrates how an attack presentation could be presented to the system and if successfully recognised, the output to be flagged to the appropriate operator who can intercept the attack.

There is an obvious case of ABC systems (controlling our borders and preventing security/terrorist threats) but how would this be controlled for the case of banking? Presentations made here are often not supervised, and so verification attempts are usually unattended. For example, if the user was to make an authentication attempt using the sensor at work or on the move this opens the device (and perhaps the associated token) to a greater risk of an attack. Implementing the ability to enable anti-spoofing or liveliness detection components in these scenarios will allow the model to be implemented, but this would be difficult to achieve in mobile applications where there is no additional oversight.

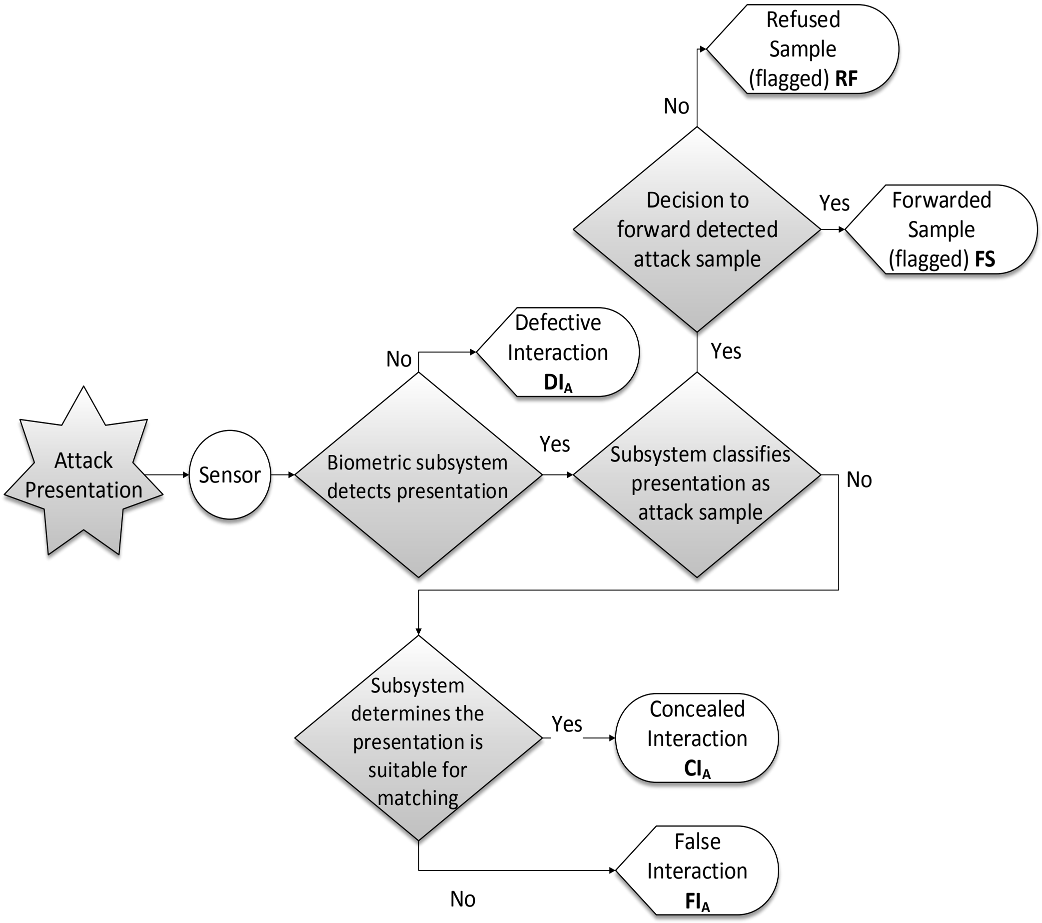

Figure 10: The HBSI Attack Presentation Model

Future work outside of the scope of this thesis could potentially investigate the effects of the environment and training within the HBSI models which may alter the user's behaviour when interacting with a system. It will be important to investigate how the design of the sensor can be altered to counter-attack presentations too. However, the application of the HBSI model will be able to provide a clearer picture in all cases. The work proposed in this thesis focuses on automating the original Interaction Framework model, applying the metrics to ABC scenarios and removing the need for coding errors by the evaluator during the trial.

2.4 Automated Border Control

Arriving at any international airport will require travellers to process through immigration control, whereby upon inspection of the traveller's documents and in some cases, biometric data, the border guard or the system will authorise the traveller access into the country.

Travellers who wish to travel internationality must hold a valid passport; a travel document which certifies the identity and nationality of its holder [73] [74]. Standard passports contain information such as the holder's name, place and date of birth, photograph, signature and other identifying information. In the last decade, passports have started to move towards including biometric information on a microchip which is embedded in the document, making them machine-readable and increasing the difficulty of counterfeiting [54].

With an increase in international flights and the availability of sophisticated biometric solutions, border control systems have adapted to new technology and security demands over the last several years.

In general, an ABC system consists of several components (See Figure 11 for example) which include, but is not limited to:

- Physical barriers (single-door or double-door)
- Monitoring and control station and equipment for the operator
- A document reader (optical devices including a radio frequency reader module)
- A biometric capture device (fingerprint reader, camera)
- User interfaces (LED signals, audio devices, monitors)
- Processing units and network drives
- Cameras/Sensors to monitor queues

The general process requires the traveller to verify a document at the first stage, and if the documents are verified successfully, move to a second stage where biometric verification is carried out. A general process flow for ABC processes can be seen in Figure 12.

Although the architecture and design of the system may have changed over the years, the core functionality of the system has remained the same.

Figure 11: An example of eGates in Heathrow (London, UK)

One of the earliest border control systems which used biometrics for verification of the traveller was Ben Gurion Airport [75] in Israel in 1985. The system, which is used by Israeli citizens only, used hand geometry to validate the traveller’s identity. This system is still in use today but has adapted to modern technology to utilise secure travel documents such as ePassports and a biometric card, which is given to travellers during enrolment. Hand Geometry recognition required physical contacts between the users and the capture device, which lead to user interaction issues. The hand needs to be placed correctly around the guidance pegs to trigger the capture. An incorrect placement would not trigger capture and cause inconvenience and often required supervised training to reduce FTA and FTPs.

Another early adopter of biometric technology for registered travellers was CANPASS, a Canadian programme released in 1995 [75]. At the core, INPASS was a standalone kiosk which utilised hand geometry and fingerprint biometrics to verify a traveller's identity. The program was eventually replaced in 2004 by the DHS Registered

Traveller Programmes (Global Entry, NEXUS), which requires fingerprints and Iris scans respectively.

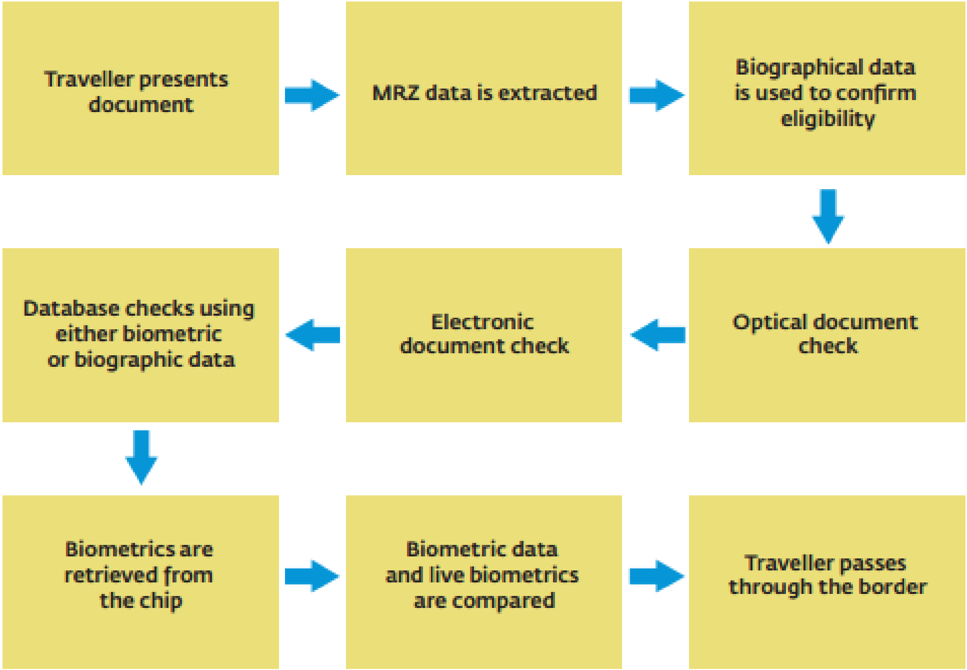

Figure 12: General ABC Process Flow [7]

J. Wayman evaluated the INPASS geometry system [76] which details the system concept, implementation and summary of the program. The main findings from the report disclosed that the INSPASS, in concept, “has the potential to be a cost-effective means of reducing processing time for frequent travellers by automating the primary inspection process without sacrificing security.” The report details some weaknesses, however; the kiosk design needed improvements. The evaluation disclosed that the system monitor and the hand geometry reader were not logically arranged and components were placed in awkward places that confused travellers who were using the system. Some components were configured to be difficult for left-handed users to interact with the machine. User feedback commented that instructions displayed were not clear and that there were a high number of false rejects. Another complaint from users was that the INSPASS inspection took as long as the manual primary inspection. Comments from some travellers noted that they would only use the program if the manual inspection queues were long and if the INPASS queue was shorter.

In the mid-2000's border control solutions started to move towards mass verification systems, which required modalities that were accessible by the general public. The introduction of electronic passports enabled facial verification. However, some pre-registered traveller programs are still in use and utilise a combination of modalities and documents.

The UK Iris Recognition Immigration System (IRIS) for example, was designed to allow enrolled passengers to cross through the UK border controls using automated barriers [77] [78]. The system was completely phased out by 2013 and replaced by the eGate (face recognition) system. The system would compare live iris images, captured by the system, against the iris images stored in the central database. A. Palmer and C. Hurrey [78] analysed problems with the system and the events to what led it to be retired in favour of eGates. The system suffered from many problems, but there were common themes; travellers had trouble lining up their eyes to the camera, resulting in a longer transaction time. Some passengers were not recognised at all and would have to be sent to the manual control. In the end, the system was not a match against the ePassport gate (or eGate) system which was easier to use and more accessible.

2.4.1 Biometric Modalities

A wide range of biometric modalities can be found in ABC systems across the globe. The most common modality found in both ABC and non-automated systems is facial verification, largely due to the access to a reference image stored on an electronic passport [7].

Several systems use fingerprint and iris biometrics. Fingerprints are typically used for immigration purposes and are usually found in semi-automated or manual systems as opposed to ABC solutions. Fingerprints can also be stored in the second generation of ePassports, which led to an increase in the number of systems using fingerprint modalities. Iris modalities are commonly used in registered traveller programmes and enrolled images are either stored directly onto a token or on a database.

Surveying systems across the globe, Table 3 below demonstrates some current examples and the respective modality used.

Table 3: Examples of ABC Systems

Modality	System and Country
Face	eGates (UK), APC (USA), No-Q (Netherlands), easyPass (Germany), RAPID (Portugal), APC (USA), Smart Gate (Australia/NZ)
Face & Fingerprint	ABC (Spain)
Fingerprint	PARAFES (France), USVISIT (USA), e-Channel (Hong Kong)
Iris	Privium (Netherlands), ABG (Germany), IRIS (UK – Retired), CANPASS (Canada)
Hand Geometry	Ben Gurion Airport (Israel), INPASS (US – Retired)

Face recognition is considered socially-accepted, nonintrusive and does not require any special training which is some of the reasons why it is favoured as the leading biometric modality for border crossing [52]. Facial verification is typically completed by comparing a live image to a stored, reference image on an electronic ID (passport or identity card). The token is read, and the image is extracted and stored temporarily in the system. The camera within the system then captures an image of the traveller and makes the comparison.

According to the BIOPASS II study [7], some ABC systems can capture a sequence of images over the course of capture. The system analyses the images from the camera in real time and the recognition software processes images to see if they meet certain quality requirements (as explained in Section 2.3.1) focus or face orientation).

In some cases, the camera within the system will automatically adjust to the user's height (e.g. UK eGates in Heathrow) or are in a fixed position (e.g. USA APC Kiosks).

In general, the biometric face verification system in a common eGate scenario must complete six steps:

1. The system chooses the camera position based on the traveller's height (if installed)
2. Information is then displayed on a monitor, instructing the user to look at the camera
3. Illumination is automatically adjusted based on environment lights
4. An image of the face is captured

5. Quality assessment is performed as well as determining if the image meets ISO standard requirements [34] [37]
6. Perform matching between the live images captured and the referenced image extracted from the document

Fingerprint recognition features high recognition performance and good social acceptance. Fingerprint verification typically consists of four steps:

1. Information is displayed to instruct the traveller how to position the finger
2. The fingerprint image is captured
3. Quality assessment is performed to determine if image(s) meet ISO standard requirements
4. Performing matching between the live images captured and the referenced images extracted in the document

Iris recognition features very high recognition performances and is considered to be highly intrusive.

1. Information is displayed to instruct the traveller how to position their face/eyes to the camera
2. A near-infrared light pulse is used to illuminate the eye, as well as control direction and dilution of the pupil
3. Iris is captured
4. The live image and the sample contained in the document/database are matched

Research trends in the design of innovative ABC systems typically investigate the use of multi-biometrics and less-constrained recognition. Multi-biometrics can increase biometric recognition accuracy, usability, and robustness to spoofing attacks, by combining multiple biometric sources [79] [80] [81] [60] [82]. Several studies demonstrate the increase of accuracy fusing face and fingerprint biometrics together in one system [83] [84] [31].

There are some multi-modal systems currently in action, such as the ABC eGate system used in Spain [7]. The face image is used as the main biometric modality, and fingerprint interaction occurs in three different scenarios:

1. In the segregated two-step process, fingerprints are used as a token to grant exit
2. During the identity verification process of the Spanish nationals, the live captured fingerprint image is used in a Match-on-Card operation, matching to the template stored on the Spanish e-ID Card
3. During the identity verification process of Spanish nationals, the live captured fingerprint image is compared against the reference data stored in the chip when the travel documents are the Spanish second generation ePassport

The decision matrix in the original setup of the Spanish ABC system required that the result of both biometric modalities comparisons be satisfactory to authorise the travellers crossing. If either the facial or fingerprint verification failed, the system considered the traveller identity verification process as unsuccessful.

In a report on the three-month study of the original implementation in 2012 [85] results indicated that 96.61% of the Spanish citizens (67,508 travellers in total) who used fingerprint-enabled travel documents could cross the border after successful fingerprint verification. However, up to 13.34% (FRR) were rejected because of the face verification result. The multi-modal implementation trialled in 2013 is based on ISO/IEC TR 24722 [86] fusion. In the original scenario, only 85.45% of the travellers could use the system. After the introduction of multi-modal biometric verification as displayed in Figure 13 below, the biometric overall error rates lowered to 4.78%, allowing 95.22% of the travellers to successfully use the system.

At the core of this fusion, two thresholds were selected for the facial verification component; a lower threshold level will reject travellers based on a given score, while an upper score will allow the border crossing to travellers who exceed the score. Travellers whose facial verification score falls between the lower and upper threshold will be required to present their fingerprint for additional verification. If at least one of the modality verification fails, the system considers the traveller identity verification process as unsuccessful.

When using biometrics in border control systems, an inevitable trade-off decision between FRR and FAR must be made. Lowering the FRR to increase the throughput in ABC systems since fewer passengers are erroneously rejected inevitably results in an increased FAR and vice versa.

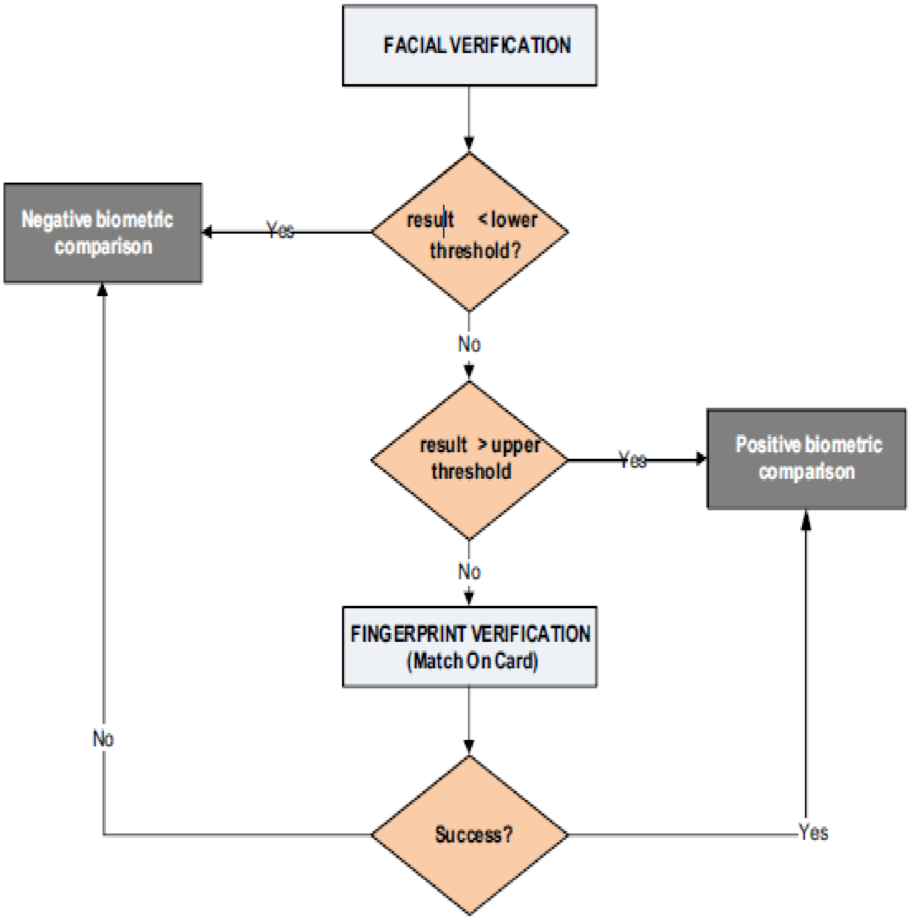

Figure 13: Multi-Modal Fusion in Spanish ABC eGate Trial [90]

Other research trends investigate less-constrained recognition, which is linked to an increase in usability and social acceptance of biometric systems [57] [79]. Contactless recognition has been researched by several institutes looking at vein [87] and fingerprint [88] [89] and iris [92] modalities.

Contactless fingerprints have yet to be implemented into ABC systems. However, research is pointing in favour of utilising the devices. Several studies have investigated the ability to capture fingerprints while the traveller is moving. A report by R. Donida Labita et al. [91] performed an analysis of user acceptability between a touch and touchless fingerprint scanner. Results indicated that 96.7% of the participants preferred the proposed touchless system over the touch-based system, with 100% considering the system more hygienic. Also, participants considered that the proposed system was more privacy compliant. This was due to the perception that no latent fingerprint could be left on the system.

On the move iris recognition acquires images in less constrained environments, capturing images as travellers walk past a sensor. A report by J. Matey et al. [92] describes the methodology behind at a distance iris capture, which is suitable for matching against a database of images. A similar implementation has recently been announced for deployment in 2017 in Dubai airports [93]. Dubbed “Iris on the Move” (IOM), systems will be able to capture images of the iris up to 3 away. The Dubai smart gates will capture both facial and iris biometrics of registered users only. As users must be registered, it will allow travellers to use either electronic passports, Emirates ID, e-gate card or a smartphone as a registered token.

2.4.2 Tokens and Travel Documents

Travel documents are used to certify the identity of a traveller when crossing a country’s border. Passports are the most common travel documents and are usually issued by a country’s government to verify the traveller and their nationality. From 2006, a new generation of passports introduced throughout the globe included a biometric element; a microchip embedded in the documents making them machine-readable and difficult to counterfeit [73].

These biometric passports, or as commonly referred to as ‘ePassport’, contain a contactless smart card which includes a microprocessor chip and an antenna embedded in the front, back or centre page depending on the country’s design. Documents and chip characteristics are well documented in several standards reported by ISO and the International Civil Aviation Organization (ICAO) [94] [95] [73]. Biometric passports may also be referred to as an e-MRTD (A machine-readable travel documents equipped with an electronic chip).

ePassports contain a digitally signed biometric file and various communication protocols (as appropriate for the country that issues the passport) as stated in ICAO specifications [96]. The digital image of the traveller (the same as the one on the passport page) is stored on the chip in a JPEG or JPEG2000 format. The latter format being a newer standard that offers better compression rates for comparable image quality. In addition to the position (i.e. coordinates) of certain facial features (e.g. eyes) can also be stored on the passport [96].

The quality of the reference biometric data stored on the passport is extremely important and is likely to have a major influence on the systems ability to match a live image. Several studies have investigated performance issues within border control systems about the quality of the reference image [97] [98] [7] [99]. ISO/IEC 19794-5 [34] defines requirements for facial images that are to be stored on the passport. The specifications include requirements in areas of pose, expression, backgrounds, shadows, glasses. Technical requirements are also detailed for focus, colours, radial distortion and colour space. See Figure 14 an example of an acceptable image.

Figure 14: Examples of Non-Accepted Images and an Acceptable Image [34]

Images of iris or fingerprints can also be stored on the electronic passport. However, these are optional and are only enrolled if required by the issuing state. These reference images can only be used for identity-verification (1:1 matching). Fingerprints are typically stored as WSQ (Wavelet Scalar Quantisation – a lossless compression format optimised for fingerprints). Two images are usually stored, the image of the left and right index finger, however, the passport can hold all ten fingerprint images if required (e.g. for certain scenarios such as visa permits). ePassports which contain fingerprints may take longer to read than passports without. The chip authentication and terminal authentication protocols require transmission of cryptographic keys and various certificates which are required to be performed by the chip. Additionally, the fingerprints must also be read, which typically add 25kilobytes of data to the transmission. The latest generation and chips, however, are significantly faster and can finish the inspection procedure in less than 3 seconds.

All passports contain a Machine-Readable Zone (MRZ) which is made up of two lines at the bottom edge of the document on the data page. When interacting with a passport reader, the device scans the relevant region first and then by using optical character recognition, recognise individual characters and obtain the digital form of the printed data.

There are two main stages to border control systems; enrolment and verification. During verification (or border crossing), a comparison of the biometric features is performed outside of the passport by either through a manual or an automated process. In either case, authorised bodies can read biometric and other data stored on the chip off the passport and compare the stored photos and images of the fingerprints/iris to those taken to at the checkpoint/system.

During verification, tokens/passports are typically authenticated and checked for fraudulence [100] [73]. Passports must be inspected for the following during reading:

- Systems must be able to read the MRZ (via optical character recognition – OCR) to be able to perform basic access control authentication
- Inspection systems must be equipped with the list of country signing certification authorities (CSCA) certificates of all countries, whose electronic passports are to be validated
- Check the physical security features of the passport under ultraviolet (UV) and infrared (IR) illumination

Although ePassports are the most common token used in identity claim scenarios, several other forms or permits and visas also contain biometric elements. Biometric visas are becoming increasingly popular throughout the world. The Biometric Resident Permit (BRP) [101] in the UK is given to those who apply to come to the UK for longer than six months. This visa contains personal details such as name, date and place of birth, fingerprints, a photo of the face and immigration status.

Permits/visas such as these are typically not used in ABC systems due the design and build of a system (the system may not be built to check immigration status etc.). However, in some cases, permits/visas can be used in the same way as a passport and act as a token for automated border control. Similarities point to the personal detail page of a passport, where personal information is stored as well as an MRZ zone. For example, German nationals can hold an electronic ID (e-ID) card which is usable by their EasyPass program (A sample can be seen Figure 15) [102].

BUNDESREPUBLIK DEUTSCHLAND
PERSONALAUSWEIS
T22000129
MUSTERMANN
GEB. GABLER
ERIKA
12.08.1964 DEUTSCH
BERLIN
31.10.2020 938568

SAMPLE

GRÜN
160 cm
01.11.10
STADT KÖLN
KÖLN
HEIDESTR. 17

```
IDD<<T220001293<<<<<<<<<<<<<<<
6408125<2010315D<<<<<<<<<<<<<4
MUSTERMANN<<ERIKA<<<<<<<<<<<<<
```

Figure 15: An example of an Electronic ID [102]

For systems that do allow additional e-MRTDs to be used other than ePassports, the ABC system may need to connect externally to a specific national system which allows the validation of the documents for access to its protected data areas.

Other systems may use registered tokens or pre-registered traveller cards as part of a frequent flier program. For example, PRIVIUM is Amsterdam's Airport Schiphol's service programme for frequent fliers which uses iris verification [53].

During enrolment, the images of the iris are directly stored onto the PRIVIUM membership card. At the border crossing, the details on the card are compared to live images of the eye. The design of these systems will alter slightly from ABC systems which use an e-MRTD as a token. The design and topology of the system will change based on the sensor (Privium cards are swiped), the biometric modality, and how the information is extracted from the chip and compared.

Several border control systems use a combination of tokens. The hand geometry system at Ben Gurion airport now uses a combination of an ePassport and a smart card [103]. The smart card code stores an array of encrypted personal information, from criminal histories to measurements of fingers, knuckle shapes and distances between joints in hand. Upon arrival, travellers go to a kiosk and swipe their smart cards through a reader and then place their hands on a biometric scanner. Once the scanner verifies a passenger's identity, a coupon is printed that allows the traveller to continue to the next part of the process. In this second stage, the passport is presented to a kiosk, and the document is then verified to authorise the traveller border crossing.

Table 4 demonstrates several examples of ABC systems and the respective biometric modality and token used. A further report of global systems is conducted in Chapter 3.

Table 4: Examples of ABC systems, the modality used and the required Token

Modality	Token Required	Examples
Face	Passport	eGates (UK), eGates (EU), APC (USA), No-Q (Netherlands)
Face	Resident Card	easyPass (Germany), RAPID (Portugal)
Face & Fingerprint	Passport or e-ID	ABC (Spain)

Although Table 4 presents a few examples, other combinations do exist. The most common modality used in ABC systems would be the face verification while using the ePassport as the required token. For this study, the work is concerned with automated systems only. The design and topologies of these systems differ, however, there are various formats that the build and implementation of the solution should follow. Section 2.4.3 also considers the role of the border guard within an ABC system.

2.4.3 Design and Other Factors

Regarding topology, Frontex [104] classifies current ABC systems into three categories:

- One-Step Process (Figure 16, below): when the token verification, identity verification and the border crossing happens in one single process

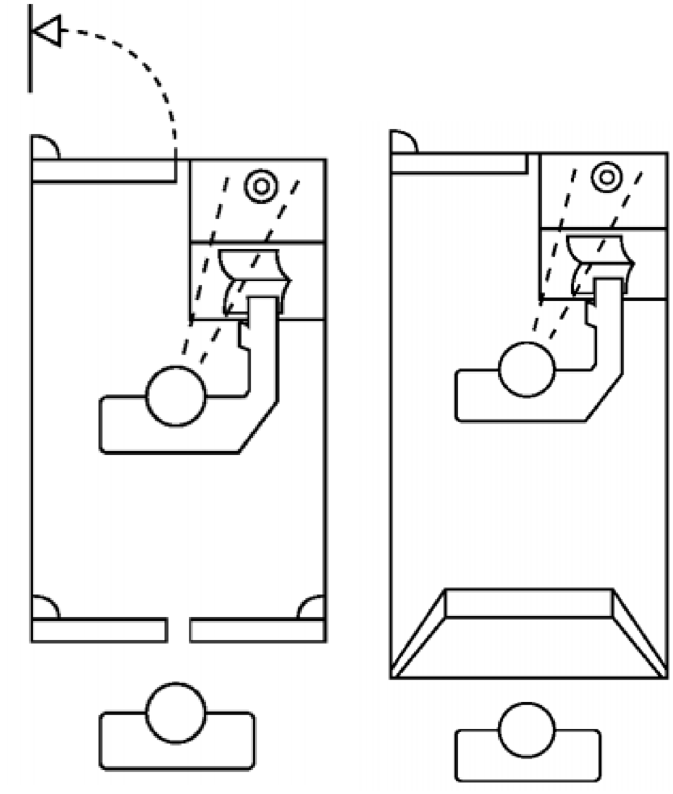

Figure 16: One Step Process [104]

- Integrated Two-Step Process (Figure 17): when the token verification and eligibility to use the system is performed in advance and, if successful, the identity verification process is conducted at a different stage in the same physical location

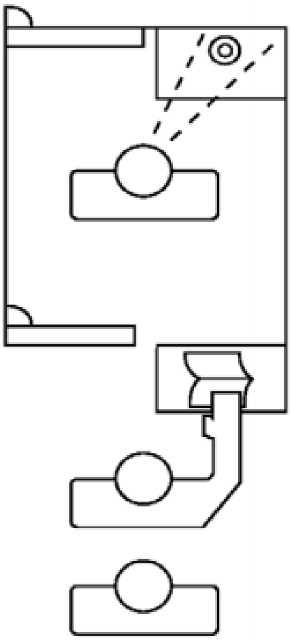

Figure 17: Two-Step Process [104]

- Segregated Two-Step Process (Figure 18): when the process of traveller verification and the border crossing are completely separated. A further token is sometimes required to link both processes, sometimes in the form of a biometric sample or ticket

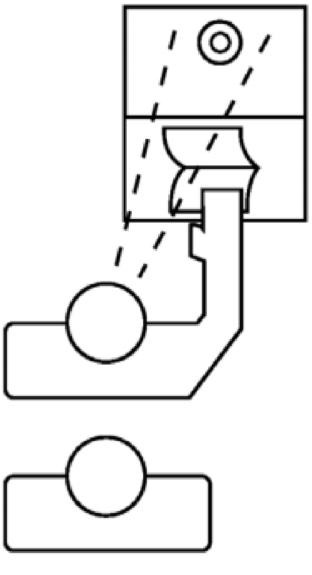

Figure 18: Segregated Two-Step Process [104]

Systems typically use a combination of physical barriers, full-page token readers, visual displays for instructions, biometric capture devices and system management hardware and software. The systems may also include uniqueness and liveness detection technologies [7] [53] [77] [104].

Several environmental factors must be considered to for the placement of ABC systems. Several constraints are to be expected during early implementation (existing infrastructure, cabling, lighting, etc.) but ultimately where systems are placed will determine how many travellers use it, how successful it is and what level of performance can be achieved [35] [104].

Security personnel oversee all aspects of ABC systems at some level. There are two main roles in the operation of an ABC system; the one of the operators and that of assisting personnel. The operator is responsible for the remote monitoring and control of the ABC system. Their most important task is to bring the necessary human factor into automated tasks. This is done by:

- Overseeing the user interface of the application
- Reacting to any notification given by the system (warnings, errors)
- Managing exceptions and providing a decision
- Monitoring and profiling travellers queuing in the ABC line and using the eGates to look for suspicious behaviour in travellers

Assisting personnel are the border guard(s) whose tasks are to handle the exceptions that take place at the ABC system, redirect travellers as needed, and assist travellers in specific situations (e.g. how to interact with the system if they are unsure).

2.4.4 Border Control Performance

It is a particularly difficult task to evaluate an ABC system. Performance is assessed based on individual components; technical performance of physical components, matching performance, timings, observations of the interaction process and traveller's perception of the system. Evaluating the biometric matching performance will determine the overall accuracy of the system, but the calculation may change between airports. Many countries include legalisation that limits the collection and storage of biometric data, and in most cases, the thresholds and matching scores are not commonly available to the public [105].

Several reports assess the Operational Reject Rate (ORR), the overall ratio of people sent to manual examination from an ABC system [106] [77] [107] [108], and through standard biometric throughput and error performance ratings. While standards require all systems to report on the biometric component; there is very limited information on other systems ORR rate. M.Nuppeny [108] states the target goal for the future operation of the German EasyPass system in 2012 was to reach an ORR <10%. From this number, a fraction of the ORR from biometrics should aim for <5% and the remaining <5% from watch list and documents checks. However, there has not been an update whether EasyPass is currently hitting these targets.

When measuring transaction time, most vendors typically measure the transaction time from the point a traveller places the passport on the reader until he/she exits the ABC systems, which typically does not exceed the 30 seconds [97] [109] [53] [7]. However, there are no current standards to define this measure, and so there is no uniform method for measuring transaction time across global implementations. Also, systems with multiple biometrics or are part of a segregated two-step process may occur longer transaction times or have multiple transaction times associated with the process.

A study on the automatic face recognition for ABC systems based on real data recorded of travellers at Schipol Airport [31] investigated the cause of performance errors and described how performance tests were concluded for a typical eGate scenario. Error rates are discussed based on the matching algorithm used to compare images to the enrolled digital passport image. The Receiver Operating Characteristic (ROC) test enables the operator to choose a threshold based on a suitable FAR rating. To complete the test, each comparison results in a score and is then compared to a threshold. The threshold differs between scenarios but is typically set at 30-40%~ [7]. If the score is above the assigned threshold and the comparison is an imposter pair (e.g. the live image and the digital passport are from different travellers), this results in FAR. Likewise, the number of genuine comparisons (both images are from the same subject) where the score is above the threshold divided by the total number of genuine gives a Verification Rate (VR). The number of genuine comparisons where the score is below a threshold divided by the total number of genuine comparisons gives FRR. By the studies definition (Equation 1):

$$VR = 1 - \mathrm{FRR} \tag{1}$$

The study notes that a common choice for FAR is 0.001 (0.1%) in ABC systems, i.e. 1 out of 1000 imposters are allowed. This is backed up by several other studies that have also noted that this a standard set across the board [110] [111] [112]. With an increase of FRR however, of more than a few percent, this could lead to delays and longer queues. This is due to an increase in manual checking by a border guard who will have to handle a small number out of each 100 genuine travellers, which may not be recognised by the system due to constraints outside of their control.

Cantarero et al. report [85] on ABC performance for some traveller's groups. FRR was reported for Portuguese travellers at 7.42% while FRR for Danish citizens was at a huge 21.59%. The average FRR across 31 countries was reported at 12.32% for over 92,000 interactions. Frontex reports [7] that Vision-box systems used in Portugal have a theoretical FRR of 4.25%. However, a study at Algarve University reported 5.2% FRR. After the study, the design of the light source was improved and led to a lower rate, although the number was not given. Frontex also suggests that there is an estimation of 17% of FRR across many systems can be attributed to the use of glasses, wearing hats, or occluding the face with hair.

Transaction time, on the other hand, can be particularly difficult to compare between systems. The process can differ greatly between scenarios, resulting in an unclear definition of time on task. One system could report on time on task from the moment a traveller steps into the eGate system until the system is completed, yet, another system may not have barriers and report on the time spent interacting with a biometric sensor only. Usability studies on non-border control-biometric systems report on efficiency and effectiveness in self-service scenarios and individual component times [55] [60]. Research is not widely available on measured total transaction time when investigating live scenarios [110] [31].

2.4.5 User Acceptance of Biometrics in Border Control Systems

User acceptance can be defined from the user's perceived need for a system and the utility it provides [113] [66]. The system then must be both convenient and usable to remain reliable and trustworthy for passengers to use. There are many considerations in biometric usability; information, guidance, ergonomics and more. This section reports on the issues surrounding user acceptance and usability issues within a border control system context.

Biometric components in border control solutions can cause problems. In some cases, travellers found a modality awkward and time-consuming to use (as documented by a user experience study on the now retired UK IRIS programme [114] and the challenges of iris recognition in UAE [115].

A study on multiple verification systems conducted by the UK Passport Service (UKPS) [116] also revealed some usability issues which affected system performance at an interaction level. More than 10,000 users participated in the study, with some 750 users who had some form of disability. Results suggested that fingerprint recognition was preferred but that some groups were more comfortable with iris recognition. Users, who identified as disabled found iris recognition very challenging, which was mostly due to the design and setup of the system being tested.

If the user has previously experienced 'slow' system performance or has erroneously been denied access, these negative experiences may cause the traveller to avoid the process in the future [114]. How the system experience is conveyed through publicity documentation and to the public through the news media can also affect the user presentation [117]. A positive user experience is typically based on convenience, confidence that the system is functioning correctly, and its perceived utility [66].

There is general conception from users that there is little trust on the use of the technology [118] [119]. Biometrics can be considered as sometimes as intrusive through both interaction and the subsequent storage of personal data. The UPKS report [116] found although most participants rated four systems they tested either satisfactory or positive; many raised several usability and acceptance issues. For example, within a fingerprint system, subjects commented on hygiene and the visible dirt which was highlighted due to illumination on the sensor.

Current global consortiums such as the FastPass [120] and ABC4EU [121] projects have noted the need to find, standardise and counter non-technical factors. Both projects project covers the broader area of border control solutions and considers algorithms improvements, benchmarking, queue analysis and more. Relevant papers are referenced to throughout this thesis.

Ylikauppila et al. [119] report on factors affecting UX and technology acceptance within ABC systems. Points of view are reported both from a traveller and border guard perspective. Data was collected from expert evaluations, passenger observations and interviews from border guards. Conducted in 2014, the results detail that passengers were still not aware of the overall process, which led to inactivity during the use of the ABC system. Passengers reported that they often do not know if they can use an ABC system based on what travel documents they are holding. According to the observations reported, many travellers who did not have the correct travel documents tried multiple times to interact with the system. For passengers using the system, observations witnessed struggles with individual components of the system, i.e. when to enter or exit the system and where to insert their passport. The researchers noted that the traveller’s restless actions caused disruption with the capture process, cancelling a transaction which in turn increased time, non-matches and rejections as well as retries. Border Guards emphasised the importance of the first time experience that travellers have. Positive experiences will influence the attitude towards the concept and their willingness to try the system in the future.

Pirelli [113] notes the importance of usability in border control when considering users with a disability. The paper reports on the scale of disability and the associated challenges for users interacting with automated systems. The report also suggests that users tend to miss key pieces of information during the process and so there should be careful consideration of the environment and situation where the system is integrated.

An important topic for travellers using ABC would be the system’s ability to be able to communicate with people regardless of native language. Implementations that utilise a Segregated Two-Step Process with an interactive kiosk have an easier task of deploying (limited) language options [104], while one-step solutions offer little to no choice. These configurations often rely on icons or simple pictorial instructions.

There is a clear need for consistent presentation and communication of biometric processes to maintain successful performance. While there is not a current standardised approach to how this is achieved in worldwide systems, some attempts have tried to improve usability and acceptance of some systems. In the UK, a number of organisations such as the British Standards Institute (BSI) and NIST have collaborated on the FaceSymbol project [122], which collected some graphical symbols representing facial biometrics. The workgroup aimed to establish a core set of icons to be used in the UK border control systems with the aim to form several ISO/IEC reports such as for systems which use face applications [123].

The idea to standardise information will make the process across the globe easily recognisable, comprehensible and consistent and therefore improve performance. However, this is for many, what makes the process confusing. Arriving in a foreign country already poses language problems but for those who are new to the airport and indeed the country, may have difficulty deciding how to proceed through border control. Yee-Yin Choon et al. conducted case studies on Biometric Symbol Design [124] prior to the FaceSymbol project. The NIST group evaluated a set of symbols intended for use in biometric systems to help users better understand biometric operations. There were six studies, with a total of 186 participants from the United States and four Asian countries. The survey reported on the matching of a symbol to its meaning. Seven symbols were determined to show ‘great promise’, with four symbols that were not well received and nine that needed further examination. The symbols were later assessed in the FaceSymbol project. See Figure 19 below for examples of the proposed symbols from the project.

To enhance acceptance and to improve the user experience, an ABC implementation needs to accommodate: a population with different demographics, language barriers and travellers from a variety of cultural backgrounds through the standardisation of signage and instructions. Also, to travellers whose interactions may be affected by stress, fatigue and a reaction to unfamiliar surroundings. Furthermore, a system must exhibit an ability to convey errors and to offer solutions leading to a more efficient process for all travellers. Moreover, this must be able to accommodate user performance and acceptance concerns, accounting for confidence, and physical or mental workload.

Figure 19: Examples of Icons gathered by the FaceSymbol project. From Top-Left across Facial Recognition, Look at Camera, Manual Passport Control, Move Hair. From Bottom-Left across: Do Not Smile, Open Passport, Remove Hats, Wait.

CHAPTER 3. BORDER CONTROL INTERACTION

3.1 Introduction

There are various approaches to evaluating the performance of Automated Border Control systems. As previously discussed in Section 2.4.4, all implementations that use a biometric component must report standardised biometric rates such as FAR, FRR and FMNR and FMR [104]. Although assessments are not readily available for all systems and are typically analysed internally, several studies have reported on the performance of the biometric sensor from a range of implementations [7] [53] [61] [52].

Reporting aspects of throughput of any system is crucial if systems are to improve. Time spent on a transaction or interaction will indicate where improvements may need to be made, for example, are there bottlenecks due to users not responding correctly to or not understanding the information displayed? There is currently not a standard definition for reporting transaction times, and therefore metrics differ from vendor to vendor. In a common scenario, one system may report on a transaction time measured from the point when a user enters their passport into the reader to when he or she exits the gate. Another system may define a transaction from the moment a user enters the first gate [77]. Different builds and topologies can also cause discrepancies in the recorded rates.

Systems may also report an Operational Reject Rate (ORR), defined as the overall rate of travellers rejected from the entire system [108]. This measurement does not consider wherever the traveller made a genuine or false interaction, or help identify where and why an error was generated. It also does not establish a difference between biometric and token interactions or possible system errors, for example, was this was due to not establishing a connection to an internal database? Or through a failure to read passport chip?

While these traditional performance evaluation techniques cover the basic operational and deployment scenarios of ABC systems, there is further work to be done. To understand the behaviours, and the system responses to these, the performance scenarios must be measured and analysed in higher detail. In this chapter, three models are proposed to identify and unify the performance assessment process against components in implementations across the globe.

From the survey of multiple ABC systems, a general process is mapped for all common implementations. To apply HSBI metrics and to perform a deeper analysis of performance and interaction, the Generic Model (GM) is proposed. The model outlines a process flow for both the enrolment and verification stages of the border crossing process, mapping points where the HBSI model can be enabled to establish metrics.

Further analysis of the verification stage of the GM reveals the Identity Claim Process (ICP), a definition of the formal stages that require the user to submit a token, present a biometric which upon successful verification, enables border crossing. These four user-focused tasks are later described in more detail in the Behavioural Framework.

The research proposed in this thesis investigates the applicability of a tracking sensor to analyse user movements and behaviours in real time. Upon identifying the common steps involved in a transaction, a breakdown of the desirable and undesirable behaviours describes what movements or actions may contribute to performance. Further research through data collections in this thesis builds on these established behaviours to assess the impact of these actions on performance.

3.2 Survey of Systems

ABC systems across the globe use a broad range of biometric devices combining either single or multiple sensors. Typically, the process will also require the presence of a token, a travel document which aims establish the identity whereas the biometric aims to verify. Requirements differ from country to country and have different usage implications for travellers depending on the configuration. To facilitate the application of the HBSI Presentation Framework to ABC, a Generic Model (GM) was developed which is based on existing systems, encapsulating key points and stages across implementations.

To faciliate the development of the GM, 23 global ABC implementations were assessed. Systems were selected based on the information available, largely through online brochures, reports and in some cases, performance assessments.

Table 5 below surveys systems across the globe providing the system name, country, token and biometric used as well as the design of the build.

Table 5: Survey of Global ABC Implementations

System Name	Country	Token	Biometric	Design
eGate	UK, IRE, FR, DE, PRT, NOR, ESP, ITL, DNK, EGY, POL	Passport	Face	One Step
SmartGate	Australia/New Zealand	Passport	Face	Segregated Two Step
eChannel/AutoGate	Hong Kong/Malaysia	Identity Card (eMRTD)	Finger	Integrated Two Step
Automated Passport Control (APC)	US	Passport	Face	Segregated Two Step
PRIVIUM	Netherlands (Schiphol Airport)	Membership Card	Iris	Integrated Two Step
PARAFRE	France	Passport or Identity Card	Fingerprint	Integrated Two Step
CANPASS	Canada	Membership Card	Iris	Integrated Two Step
Smart Gate	Dubai UEA	Identity Card, eGate card or Passport	Face	One Step
e-Gate	Taiwan	Passport	Face & Fingerprint	Integrated Two Step
Automated Gate	Japan	Passport	Face & Fingerprint	Integrated Two Step
Passport Control	Israel	Passport	Hand Geometry	Integrated Two Step

Out of the 23 systems, three configurations were a segregated two-step system, whereby the process is split between a kiosk and gate elements. In both instances, the user approaches a kiosk and enters their passport to be read. The user is then required to answer a series of questions based on immigration control on a touchscreen. Upon successful completion, a receipt/ticket is produced for the traveller who will use it as a temporary token for the next step. The traveller will then enter a typical eGate setup using the ticket produced from the kiosk. The ticket is then issued back to the traveller who must present the ticket to an immigration officer after collecting their baggage. Twelve of the surveyed systems used a one-step solution which all used facial verification as the biometric modality. The passport is the most common token, used in 18 of the 23 systems.

Although deviations do exist across implementations, a general process flow can be seen throughout the enrolment and verification stages. Looking at both sides to the process, a Generic Model is devised based on common steps of the implementations surveyed. These systems are composed of both automated (using technologies that typically do not require intervention by human operation) and manual elements.

3.3 Generic ABC Model

A Generic Model (GM) for ABC systems is devised outlining both the enrolment and verification stages of ABC systems. Both stages of the model are built based on the systems surveyed in Section 3.2 and through discussion in Section 2.4. The main purpose of outlining systems via this method is to permit the identification of automated steps within each system, and where possible, highlight areas in a system where it may applicable to identify HBSI errors which may occur during a presentation.

The Enrolment Stage (Figure 20) and Verification Stage (Figure 21) outline the general process flow that a traveller must complete for successful ABC crossing. While the enrolment stage is typically completed once per token, the verification stage will be completed each time a user wishes to cross a country via an ABC system.

For both parts of the GM, grey sections refer to areas where a form of manual intervention is required (e.g. this part must be completed by a border guard). Processes were donated by a blue section are automated (e.g. a biometric capture is algorithmically assessed, or a component can automatically detect movements via sensors within a gate). The white node indicates the starting point for interaction. Exit points within the GM, where travellers may be rejected from the system, are shown in red, while green nodes denote possible success or approval through a process. Orange sections refer to processes where border guard personnel may need to assist if the traveller is having difficulty with a certain action (e.g. struggling to complete the passport interaction stage).

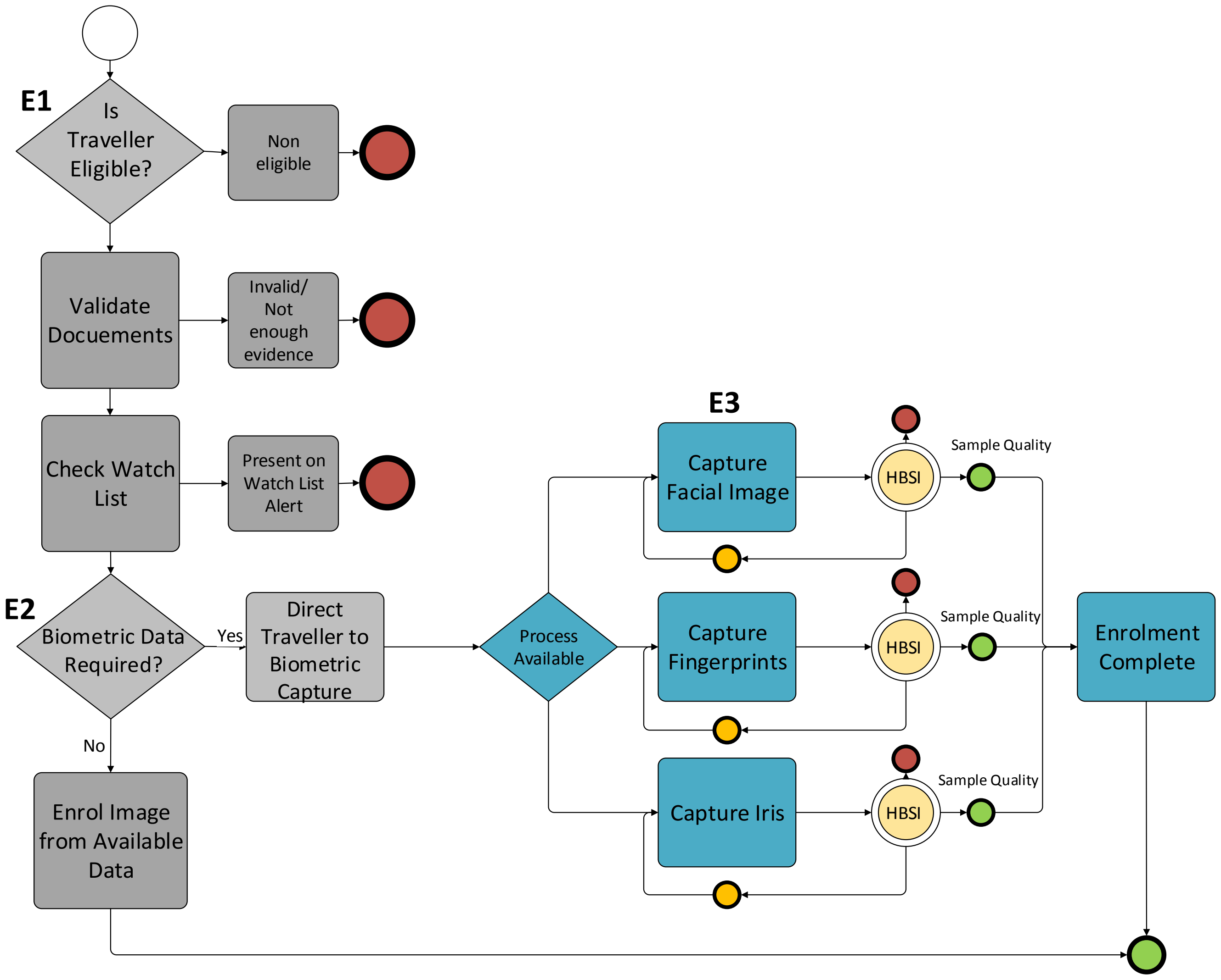

Figure 20: Enrolment Stage of the Generic Model

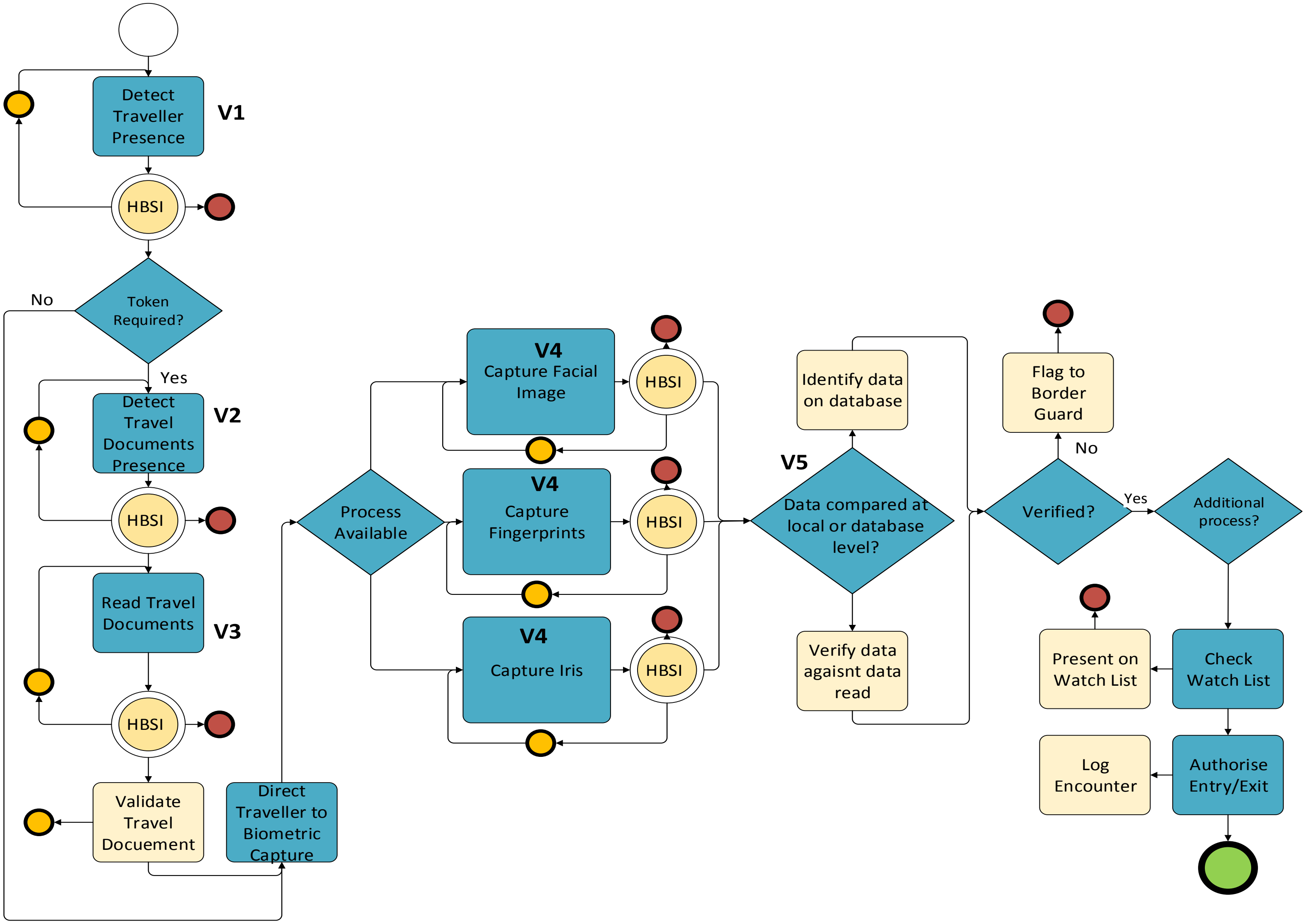

Figure 21: Verification Stage of the Generic Model

To facilitate the work conducted in this thesis, additional nodes have been added to points throughout both processes where the HBSI framework can be applied and adapted to identify specific behaviours. These are shown in yellow and are used for performance interaction assessment purposes. For example, identifying HBSI categorisations at biometric capture will allow further classifications to be made for that modality. For this first version of the GM, the route map is restricted to the major biometric modalities found in ABC systems which are facial, fingerprint and iris.

The first stage of the process will require the enrolment of both biometric and personal data to generate a token. After being approved onto a border access programme by assessing eligibility (E1), the traveller may be asked to provide biometric data at an enrolment centre or via self-captured facial images (E2). As discussed previously, passport tokens are typically generated from a passport photo, while other systems may require the traveller to physically present their sample to a system usually at an airport or a country's designated embassy or enrolment centre. Enrolment differs from each configuration reflecting the specific requirements for that country and the modalities used. HBSI assessment can be applied to all outcomes of point evaluation E3 ('Biometric Data Capture.') After successful biometric capture and processing, the enrolment stage, from a user point of view, is completed.

For the verification stage (Figure 21), the model proposes an outline for assessing the transaction for crossing a country's borders via an automated system. This stage of the GM includes a provision to detect the traveller's document (V2), enabling the HBSI Full Model as discussed in Section 2.4. If the user successfully enters their token, reading (V3) is performed, and the result is displayed back to the user. Upon successful validation of the token (if appropriate), subsequent biometric capture (V4) and verification at a local (token) or non-local (database) level (V5) of the traveller will lead to either authorisation or rejection to cross the border. Some configurations may contain liveness detection components which can identify a passenger's presence (V1). Systems may have built-in sensors that can detect a traveller entering the gate, therefore starting the transaction process when a traveller enters and displays instructions on the screen. The component that detects a travellers presence could be potentially assessed through an adapted version of the HBSI Presentation Framework.

Whereby identifying possible conditions such as: if the user has entered the system correctly, too quickly (e.g. straight after another user has exited), if another traveller is detected, or if the traveller is using the system already. Detection of such conditions are vital in the first stages of the process as this may alter how the system proceeds.

Table 6: Evaluation Points in the GM

Evaluation Point	Definition	Possible Outcome	HBSI
V1 Traveller Presence	Is traveller standing in the required area?	Yes (A2), No (Reject/Assist)	FTD/DI
V2 Token Presence	Is the token detected? Can the MRZ and other components be read?	Yes (A3), No (Reject/Assist)	FTD/DI
V3 Token Read	Was the token successfully processed?	Yes (A4), No (Reject/Assist)	SPS, FTP/CI
V4 Biometric Capture	Which biometric is required? Can the sample be captured?	Identify Modality (Iris, Finger, Face)	All
V5 Data Verification	Matching biometric data against information on the token	Performance at Database or Local Level – separate metrics	N/A

Table 6 shows the evaluation points throughout the Verification Stage of the GM, highlighting possible outcomes and HBSI categorisations. There are five stages or evaluation points; the user must enter the system (V1 traveller presence), insert the token (V2 token presence) and wait for the reading process (V3). After, the user must present a biometric (V4) and wait for the result (V5). After successful verification, the traveller can cross the border/leave the system. Although there are obvious points at which HBSI can be applied to improve the categorisation of errors, such as for biometric capture, evaluation points V1, V2 and V3 can also be assessed using automated processes. Detecting movements and behaviours throughout each of these tasks will enable an efficient and effective process, improving the ability to identify why and when an error may occur.

The verification stage of the GM can be decomposed further into individual stages which detail the process that the user must complete for a successful crossing. The proposed Identity Claim Process (ICP) framework discussed in Section 3.4 details steps of the verification stage, whereby a breakdown of the steps involved can outline expected or unexpected behaviours (system or user) for a sub-system component.

3.4 Identity Claim Process

The Identity Claim Process (ICP) model outlines the eight required steps for a system that requires a claim of identity alongside a biometric sample. The ICP can only occur during the Verification Stage in the GM and runs parallel to the evaluation points (GM Node). In Table 7, a definition of each proposed step and the related evaluation points from the GM are detailed.

Table 7: Identity Claim Process

Step	Title	Description	GM
1	System Requires a Claim of Identity	The system may or may not require the user to make an identity claim	V1
2	User Makes Identity Claim	The user either presents their token or submit their travel documents to the reader. The user must submit their token in such a way that the system should be expected to accept it	V2
3	Identity Claim Accepted by System	If the token can be read then, it should be accepted by the system. If this step fails, it is a failure of the token or the system, not the user	V3
4A	Identity Claim Corresponds to Valid Identity	The token exists in the database, or the token has a valid enrolment sample, digital signature, expiry date. The token has not been revoked	V3
4B	Claimed Identity belongs to a different user	The user may be using a false identity; for example, the token may have been (accidentally) swapped with a travel companion. If the intent was malicious, then this counts as an attack	V3
5	User Correctly Presents Biometric to System	A correct presentation can be defined when the user presents their biometric corresponding to the requirements of the system	V4
6	Biometric Subsystem Detects Presentation	The biometric system correctly detects the biometric data and can perform subsequent processing.	V4
7	Biometric Subsystem determines that presentation is suitable for matching	Biometric subsystem determines that the quality of the biometric sample be sufficient and can extract features to enable biometric matching to take place	V4
8	Biometric matching validates user against claimed identity	If the system is an identification system, then this means that the user is determined to be an enrolled user. If it is a verification system, then the identity claim of the user is verified	V5

Each step of the ICP must be completed to achieve successful border crossing. By breaking down the process into these tasks, a performance interaction assessment can be made at an individual step to identify where an error may possibly occur.

For example, Step 1 (System Requires a Claim of Identity) can only occur during the evaluation point V1 as defined in the GM. When a traveller enters the ABC system, the 'interaction' can be categorised in several ways; the system should be able to detect movement and initiate the next step, by displaying information on the screen. However, this may not process correctly by failing to detect any movement, which will either be the user not moving forward or from the sensor failing to detect movement. This stage requires both the user and system to work together – the user making the 'desired' or 'correct' behaviour of walking to the system, whilst the system correctly identifies human presence.

Another example is at Step 3 (Identity Claim Accepted by System), which can only occur after Token Read (V3) has been successful, will rely on the system's ability to complete the required verification processes for the presented token.

To facilitate the model, it is necessary to outline which ICP steps are either directly attributed to the user's behaviour or to a system process. Table 8 reports below.

Table 8: ICP Steps and Attributions

ICP Step	ICP Description	Attribution
1	System Requires a Claim of Identity	System/User
2	User Makes Identity Claim	User
3	Identity Claim Accepted by System	System
4A	Identity Claim Corresponds to Valid Identity	System
4B	Claimed Identity belongs to a different user	System
5	User Correctly Presents Biometric to System	User
6	Biometric Subsystem Detects Presentation	System
7	Biometric Subsystem determines that presentation is suitable for biometric matching	System
8	Biometric matching validates user against claimed identity	System

Further work using the model can outline various scenarios for assessment. Table 9 illustrates several potential system responses for a scenario where the user has already entered the ABC system and has successfully had his or her token read (Steps 3 and 4A). The traveller has failed the step where they are required to present their biometric correctly to the sensor due to incorrectly presenting their biometric (Step 5).

Table 9: Performance Assessment using ICP and HBSI

ICP Step									
1	2	3	4	5	6	7	8	**Potential System Response**	**HBSI**
Y	Y	Y	Y	N	N	N	Y	True Match	CI
Y	Y	Y	Y	N	N	Y	Y	False Non-Match	CI
Y	Y	Y	Y	N	N	Y	N	Failure to Process	FI
Y	Y	Y	Y	N	N	N	N	Biometric Not Presented	DI

In this illustrative scenario, only the potential system responses would be in effect in conventional assessment metrics. The ideal system response in this example would 'Biometric Not Presented' whereas a 'True Match' would be the worst scenario, shadowing a potential security risk. The inclusion of the HBSI framework can help to establish cases where the system was correct in the identification of the scenario as erroneous (False Interaction) and that the user incorrectly presented a biometric sample.

Therefore, HBSI works both ways, reporting on correct or 'good' system performance (FI's) and highlighting potential security threats where the biometric is not detected but the system grants access (Concealed Interaction). If the incorrect interaction was categorised in real time, feedback could be presented to the user with an option to recapture the required biometric sample or if necessary, to restart the process (for example, if a biometric sample was only deemed acceptable over the required threshold or it failed certain feature checks).

The ability to track and analyse certain user movements for stages that require an action from the user will enable a range of new features that could be adopted for biometric systems. Step 1 (System Requires a Claim of Identity), Step 2 (User Makes Identity Claim) and Step 5 (User Correctly Presents Biometric Sample) could be analysed through several technologies such as image processing elements or by using Action Recognition methodologies using a tracking component in the system.

A system with this capability may enable biometric solutions to produce a higher level of quality in guidance and feedback. For example, if movement/object tracking could detect non-movements then specific feedback could be produced to alert the user in a different format (e.g. produce a sound or display textual information based on the language in the passport).

Upon identifying scenarios with automated processes for detecting movement and behaviours, the performance of recognition algorithms, human-computer interfaces and the ergonomics of the systems can be analysed further, revealing possible areas in the system which could be improved.

For all forms of testing and analysis, it is important for systems, regardless of tracking abilities, to be able to handle certain unwanted actions. Early design stages usually address potential ergonomic issues, and usability engineers are commonly tasked with making the system accessible, intuitive and user-friendly [68].

During the initial system design, there must always be careful considerations for unexpected user and system responses. The research proposed in this thesis is concerned with feedback and performance assessment when unwanted actions are performed by the user. A typical scenario that is likely to be common in most facial recognition implementations is how the system can handle unwanted interaction errors, such as looking away from the camera or not fitting in the region of interest for capture. For the eGates in Heathrow, these solutions can adjust to the height of the user which helps reduce problems of travellers not looking at the camera. However, not all systems are purposely built to tackle this hurdle. In addition, nearly all ABC systems are limited by the information (through standardisation of information such as icons) that is displayed on the monitor. Indeed, systems can be restricted in many factors due to the design and placement of the physical build as well.

In summary, the ICP has outlined eight steps of the verification procedure that must be successfully completed for the user to complete the ABC process. The Behaviour Framework, proposed in this thesis through the next section, outlines ‘desirable’ and ‘undesirable’ actions for steps of the ICP which require a user movement or action. The framework will then lay the foundation in highlighting areas which can be assessed in real time by using the proposed methods to improve system interaction performance.

3.5 Behaviour Framework

For steps of the ICP where a movement or interaction from the user is required, the Behavioural Framework proposed in this section outlines some examples of desired, undesired and cautious behaviours that may occur during that specific task. The behaviours are defined based on observations made in several usability studies of biometric systems [7] [116] [125] [62] [15] and possible categorical errors observed in previous studies collected by the HBSI team.

The impact on these behaviours within a purpose-built system simulating the self-service environment is compared to performance assessment metrics, to identify if the ability to monitor this information sheds any light on user-interaction errors.

Before identifying behaviours, the specific tasks related to all self-service biometric systems need to be identified. Table 10 below outlines four tasks where user input is required and its related ICP step.

Table 10: Overview of the Behaviour Framework

Task	ICP Step
Entry – Movement from the queue to the designated feet symbols in the system. Task requires the user to move in front of the system and place baggage in an appropriate setting. The user will then follow instructions on the screen	1
Token Read – The physical movement of the token to the sensor. Requires the user to locate their token and place it on the sensor according to instructions on the screen/near the sensor. Most sensors will require the token to be placed for several seconds to ensure successful capture. Information will likely be displayed confirming the result	2
Biometric Read - A physiological sample is required to a sensor. Information on a monitor will typically instruct the user to move the required body part. The movement or action will differ based on the modality used in the scenario	5
Exit - Movement from the feet symbols through the gate/to the point of exit. Task requires the user to move from the system and remove any baggage	8

Although human behaviour can at times be unpredictable in an operational scenario, common responses can be mapped with the potential to be tracked and categorised by systems. The following section breaks down each task and identifies typical behaviours that can be observed from both an operational and data collection point of view.

Certain behaviours can, therefore, lead to ‘desired’ outcomes in the system, which in most cases, will result in correct sub-system processing and increase the likelihood of successful border crossing for the traveller. These should be the behaviours that users should aim to perform to successfully complete each task.

Table 11 outlines potential behaviours for the Entry task.

Table 11: Task 1: Entry

Behaviour	Categorisation	Impact
User moves forward in a timely manner and steps on the feet symbols, placing luggage in an appropriate position (e.g. does not interfere with equipment)	Desired	Increased efficiency
User moves forward but does not step on the feet symbols in the desired location	Cautious	Potential impact on the ability to capture in subsequent tasks
User moves forward and places luggage in an obstructive manner (e.g. in front of equipment)	Cautious	Potential impact on interaction for both token and biometric
User does not move forward	Undesired	Decreases throughput and efficiency, might require assistance

The desired behaviour for entry would be for the user to move in a timely manner to the system and be ready for the following instructions typically displayed on a monitor. The impact is likely to improve efficiency as the user demonstrates they are aware of the process and what is required from them. Cautious behaviours are identifiable through categorisations that could either lead to a ‘good/desired’ or ‘bad/undesired’ result, however, the categorisation should be made after an elapsed period.

In the example of the user not moving forward, this behaviour could potentially demonstrate that the user is not aware of the process and they may be confused on how to proceed. Not moving forward could have ramifications for other travellers, for example, the queue can build, and therefore new arrivals will be sent to the manual control. A border guard will usually be stationed close by to assist with the traveller if they do not move forward.

Table 12 outlines the expected behaviours for Token Read.

Table 12: Task 2: Token Read

Behaviour	Categorisation	Impact
User presents token in a timely manner with appropriate pressure and little movement	Desired	Increased efficiency
User presents token but makes small movements making token difficult to be read	Cautious	Depends on systems ability to capture. Small movements maybe tolerated
User presents but does not allow sufficient time for reading (either through movement or removing token to early)	Undesired	Increased likelihood of unsuccessful capture
User does not present token to the system	Undesired	Unable to continue without assistance/until behaviour is corrected

Token Read requires the user to present travel documents in a timely manner with precision. A successful interaction is likely to improve the rate at which the token is read and therefore processed. For several passport scanners, the token must be held for several seconds for the information on the chip to be read [104] and so the user should not move the documents during reading. The build and design of token readers may change between systems and so how the user interacts with the sensor will also change. Typical scanners require the passport to be pressed down whilst users interacting with some registered traveller programmes may be required to swipe a card instead. In this case, the movement is controlled, and therefore there is less chance of the document being moved around. However, some sensors do give users more range of motion, possibly leading to a higher chance of erroneous presentation, especially when juggling luggage and other accessories. In some cases, users may remove their passport too early (as assessed by information on the screen) which will typically require the process to start from the beginning [126] [110].

Another element to Token Read is the process that follows, Biometric Read. During this time users may be watching for more information or making a movement to suggest they are preparing for the next task or exiting early.

Table 13 below details the next task in the Behaviour Framework, the required Biometric Read.

Table 13: Task 3: Biometric Read

Behaviour	Categorisation	Impact
User presents biometric in a timely manner within the systems specifications and limits	Desired	Increased efficiency
User presents biometric but with slight/small movements during interaction (e.g. small face movements, partially moving finger that is not severe enough to affect the outcome of interaction)	Cautious	Most sensors will take multiple images and process a sample that meets a quality threshold. The impact will depend on systems ability to capture the image
User presents biometric and makes significant/erroneous movements during interaction (e.g. looking away from the camera for an extended amount of time, placing the wrong finger on a camera)	Undesired	Erroneous movements will likely increase chances of unsuccessful presentations. Depends on systems ability to capture a sample
User does not present biometric to the system	Undesired	Usually, results in a failure and rejection from the system

Possible interaction errors with a biometric sensor have been well covered in years of research as discussed throughout Chapter 2. A correct presentation will require the user to submit the correct biometric to the sensor within a timely manner. Some sensors can adapt slightly to ‘cautious’ behaviours by capturing multiple images and choosing the best quality sample through the systems processing algorithm. In some cases, users may make an undesired behaviour through incorrect interactions, for example, providing the wrong biometric, looking away or making too much movement so that the sensor is unable to capture. Interaction errors will change with the biometric modality used in the system and this may mean that systems that use more intrusive biometrics, such as fingerprints, may be open to a higher number of errors and undesirable behaviours. Some fingerprint errors could include: not applying enough pressure to the sensor, by not having the full finger on the sensor, or applying at an awkward angle.

Biometric Read is also accompanied by the subsequent processing, where users will typically be looking for further information on wherever their sample was accepted or not.

Table 14 below details the behaviours expected for users exiting the system.

Table 14: Task 4: Exit

Behaviour	Categorisation	Impact
User exits system in a timely manner	Desired	Quicker queue
User is gathering belongings, putting items away etc	Cautious	Not as efficient in reducing queue
User does not move forward	Undesired	Will slow throughput and may warrant assistance

The task of exiting a system is relatively straightforward; once a decision has been made and displayed on the screen, the user should exit the system by the designated process, e.g. passing through gates. The impact then is a successful transaction that helps to bring down the busy queue, one of the main goal of ABC systems. Cautious behaviours could see travellers getting their belongings ready and undesirable behaviours where travellers do not move forward, potentially not understanding their result or what to do next.

In general, most systems can account for small movements during the human-sensor interaction. Token and biometric sensors will be reading on a ‘loop’ – looking for the required information until an allotted amount of time has passed. Some facial recognition systems such as the UK eGates take multiple images and select the best image based on a quality score. However, consistent or severe errors in user movement are difficult to control and ultimately will have a higher impact on the result, likely contributing to increased time on task and possibly a higher chance of a reject from the system. In real scenarios, confusion or undesirable behaviours will usually end with some form of assistance from a border guard.

3.6 Research Goals

Performance assessment in ABC systems is achieved by reporting on the different components of the system, such as the speed of the algorithm, error rates captured from the biometric sensor and checks against the database. Not all this data is made publicly available. However, all systems with a biometric component must report on the standardised biometric performance rates such as FAR and FRR.

There has yet to be a system or study that assesses the entire transaction, by measuring the user movement to the system, their biometric and token sample, and if the user has exited the process.

The ability to gauge the movement of the user throughout the transaction in real time will offer significant advantages over other already proposed methods. One leading advantage for this proposed solution would allow the system to adapt to the user based on their behaviour to prevent potential errors within an interaction. Tailored feedback may reduce the likelihood of mistakes and improve various usability metrics such as user satisfaction and efficiency.

There have been several standards produced by ISO/IEC that establish a range of icons for use in implementations [127] [128] [129] [130]. See the FaceSymbol project for a scope of the icons used in ABC systems which use facial recognition [122]. While there are little studies to evaluate the effectiveness of this work, some research has investigated other variables that may affect performance; for example, through forms of instructions [63], the use of overlays [65] and the impact of the placement of sensors on a self-service environment [55].

Moving forward, research should consider the entire transaction and not just the output of individual components or sensors. It is important to consider where the user is and if they respond appropriately to a given task. Additionally, the use of adaptive information may also improve the process and reduce the likelihood of errors.

Data collections outlined in this thesis highlight the applicability of tracking systems in a self-service environment. Enabling the ability to track, record and measure actions during an interaction will allow a deeper understanding of where and why errors may occur. Furthermore, the data collected will be cross-referenced with other studies in the field, comparing the results from the studies to those that have previously been obtained for similar systems.

There are many tools available to enable tracking of the user throughout a process. A popular research field which seeks to evaluate body movements is Human Action Recognition, which aims to recognise human behaviours within a scenario. Studies in this area may use devices equipped with a depth sensor, a component to enable computer-aided vision. One sensor that has recently gained favourable attention is the Microsoft Kinect sensor, a motion sensing input device originally made for the Microsoft Xbox 360 games console. The Kinect has been used for real-time calculations on body movements and skeleton detection. Chapter 4. Kinect Stability, investigates the application of the Kinect further, discussing research in the field and the stability of the data captured.

CHAPTER 4. KINECT ANALYSIS

4.1 Introduction

Action Recognition, by definition, aims to recognise the actions and goals of one or more users from a series of observations, either through manual or automated means [131]. Recent technological advances have enabled Activity Recognition using a range of sensors and components, whereby devices are used to model a broad range of human activities using computer-vision methodologies.

Research into Action Recognition is highly active due to its extensive applications such as human-computer interfaces [132], human-robot interaction [133] and video surveillance [134]. With advances in the last decade and the availability of low-cost sensors, several systems have been introduced for recognising specific human actions in certain scenarios. Some examples include human recognition through shape analysis [135] and body motion skeleton tracking [136] [137]. There are two distinct areas within the field; Sensor-Based Activity Recognition which uses sensors (such as an accelerometer, GPS, microphone) to establish either a single user, multiple users or group activity [138] [139] [140].

The other leading field is Vision-Based Activity Recognition, which aims to track and understand the behaviour of users using computer-aided vision, usually through video sequences or digitised video data [141]. Zhang et al. [142] review several methodologies within the area and highlight a relatively new trend of research using depth-based sensors.

One device that is commonly used for motion sensing based recognition is the Microsoft Kinect, a device that enables users to control and interact with a console or computer without the need for a controller [143]. Typically, users interact with the device through gestures and spoken commands for gaming purposes. The Microsoft Kinect device is from the RGB-D (Red-Green-Blue & Depth) camera family that are increasingly utilised in the detection of human activities. The depth camera enables 3D capture, allowing the ability to generate real-time skeleton models of people with different body positions. The Kinect is one of the most affordable, whole-body markerless motion capture technology that is appropriate for both home use [144] and a range of other applications thanks to the support of the Microsoft Software Development Kit (SDK) [145].

There have been two iterations of the Kinect device. Microsoft introduced the Kinect Version 1 (V1) in 2010 to expand the service offered by the Microsoft Xbox 360 games console and therefore enable motion tracking in some games. The device was adopted by PC systems in 2012, with Microsoft introducing the first version of the SDK Kit to allow developers to create applications for specific scenarios. The Kinect Version 2 (V2) was introduced as part of the Xbox One console in 2013 and later released as Kinect for Windows in 2014.

There are several differences between the Kinect V1 and V2 devices. The Kinect V2 can quantify body motion by tracking the 3D coordinates of 25 anatomic joint centroids (Figure 22) compared to the 20 joints captured on the older Kinect V1.

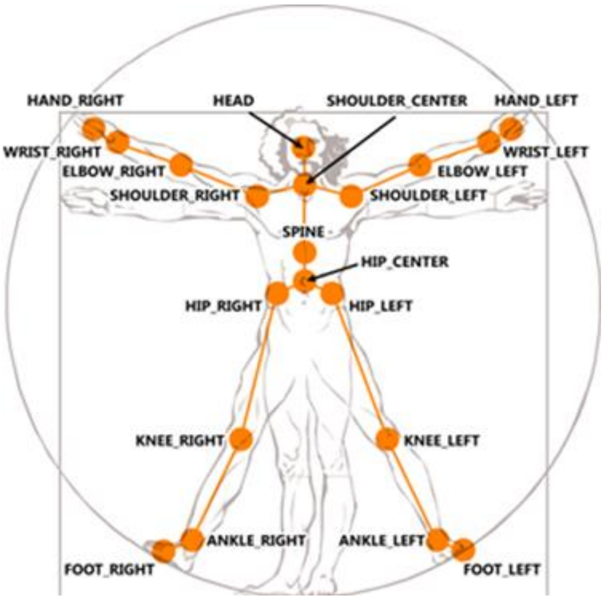

Figure 22: Skeletal Tracking Map for Kinect V1 [146]

Other technical differences between the devices are displayed in Table 15 below.

Table 15: Feature Comparisons of the Kinect V1 and V2

Feature	Kinect V1	Kinect V2
Colour Camera	620 x 480, 30FPS	1920 x 1080, 30FPS
Depth Camera	320 x 240	512 x 424
Max Depth Distance	~4.5m	~4.5m
Min Depth Distance	40cm (near mode)	50cm
Horizontal Field of View	57 degrees	70 degrees
Vertical Field of View	43 degrees	60 degrees
Tilt Motor	Yes	No
Skeleton Joints Defined	20	25
Full Skeletons Tracked	2	6
USB Standard	2.0	3.0
Supported OS	Win 7-10	Win 8-10

The Kinect device uses skeletal tracking features to recognise users and follow their actions. Using the infrared (IR) camera, the Kinect V2 can recognise up to six people in the field of the view of a sensor. The application can locate the joints of the tracked users in 3D space and track their movements over a transaction. The sensor is designed to recognise users standing or sitting, and facing the Kinect instrument. Sideway poses provide information-processing challenges regarding the part of the user that is not visible to the sensor, so the device works best when viewing the user from a frontal view. The field of view of users is determined by the settings of the IR camera. In the default range mode, which is used in this study, the Kinect can recognise people standing between 0.8m and 4.0m away from the device, however for optimal results, Microsoft recommends users be between 1.2m and 3.5m away from the apparatus. The device also works best when the tracked user is within 43.5 degrees of the sensor [143]. Figure 23 below provides the Vertical Field of View in the default range.

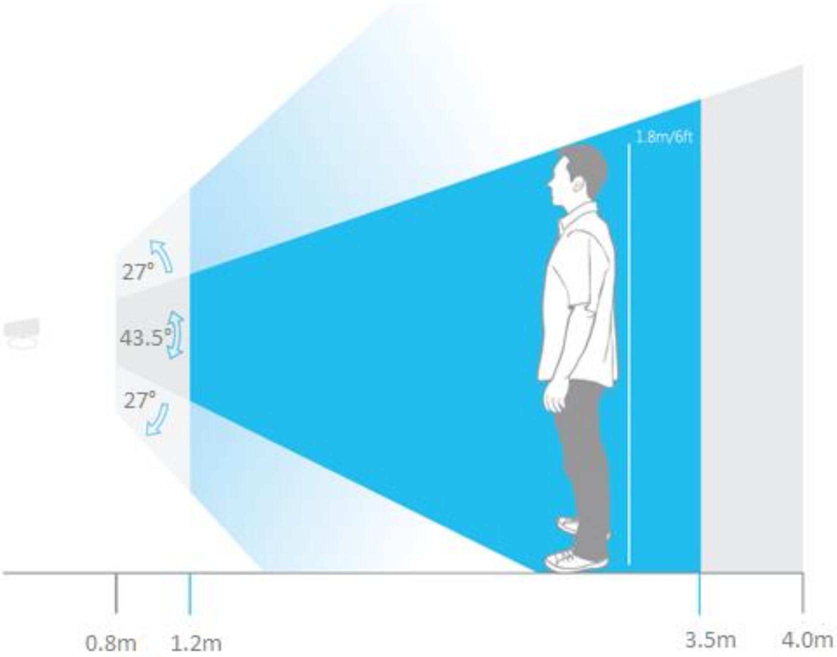

Figure 23: Vertical Field of View for Kinect V2 [143]

The distance of objects within the camera's range of view is calculated from time-of-flight analysis of reflected light beams, which yields a depth model of surrounding structures [147]. Based on machine learning techniques, the SDK application detects human shapes. It further provides an artificial skeleton based on 25 artificial anatomical landmarks (‘Kinect Joints’) projected into these shapes based on depth data.

3D coordinates define each joint position; X, Y and Z and is presented in a Cartesian coordinate system. The (0, 0, 0) point represents the sensor, and all other points are measured regarding the position of the sensor. Refer to Figure 24 below for more information.

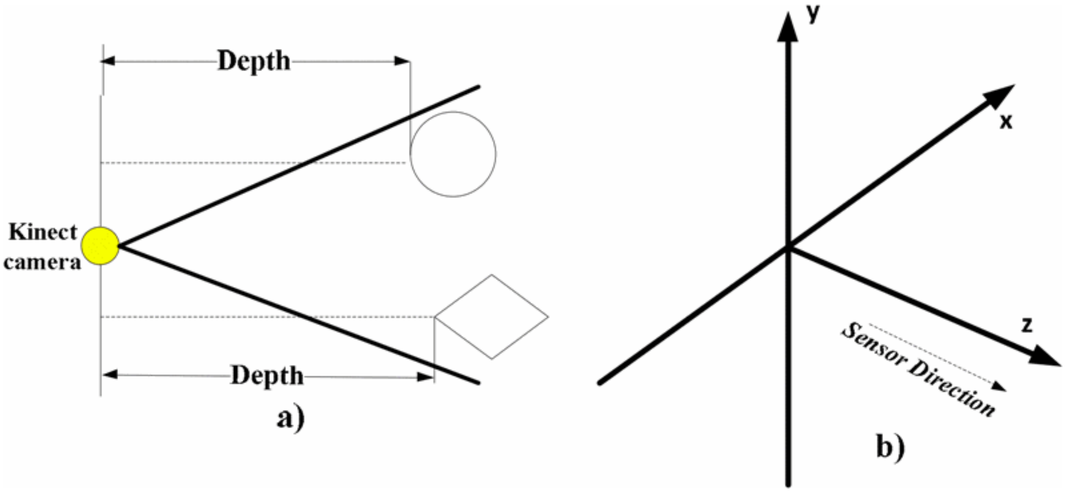

Figure 24: (a) Kinect Depth space and (b) skeleton space: X is the horizontal axis; Y vertical and Z is the position of the depth axis [143]

To summarise, the X is the distance in the horizontal axis; increases to the left decreases to the right. Y vertical increases are going up and decreases going down. The Z-axis distance; by standard, is measured from the Kinect X Plane (Figure 24) rather than the distance from the sensor.

Additionally, all values are reported in metres. So, for example, a standard CameraSpacePoint (or joint) [-1.5, 2.0, 2.5] is located 1.5m to the left 2.0m metres above and 2.5 metres from the sensor. The CameraSpacePoint is then reported for each of the 25 joints for each frame. Both Kinect sensors output 30 frames per second.

The Kinect cannot ‘see’ a single joint directly; it calculates joint position based on multiple body parts (or CameraSpacePoints), and the orientations of those joints, which produces the skeletal map of the user. There lies one of the sensors’ shortcomings, results are reported from the sensors point of view and will require adjustment for accurate reporting. For example, if the sensor is tilted slightly downwards, joints position will be slightly warped due to the sensors ‘view’. Take the foot joint, the sensor will consider the depth to be closer than the head. Therefore some joints will appear closer to the sensor than the actual field. The easy method to correct the vision would be to make sure the sensor is adjusted to be on a level with the person being recorded. However, this is not always possible in certain situations such as those in self-service environments, where users will be off different heights and will require the user to move closer or further from the sensor.

The Kinect can be programmed to calculate the map of skeletal points by distinguishing the body from the floor using the FloorClipPlane [148] [149]. Each skeleton frame contains a floor-clipping-plane vector, which defines the plane the user is standing on and returns a 4-floating point value in the Hessian normal form (Equation 2 below)

$$Ax + By + Cz + D = 0 \tag{2}$$

The Kinect sensor returns equation 2 as a vector (X,Y,Z,W), where (X,Y,Z) denotes the camera origin to the floor plane such that it is a perpendicular intersection with the floor plane. The W value is the magnitude and indicates the height of where the Kinect sensor is positioned in relation to the floor plane. If W is equal to 0, it means the field of view is limited, meaning there are likely to be objects blocking the field of view. The accuracy of the data reported by the Kinect is improved if the sensor has information of the FloorPlane [148] [149].

To measure the distance between the floor and a joint, the Point-Plane Distance Formula is used (Equation 3) [150]. Where (A,B,C,D), represents the FloorClipPlane formula (X,Y,Z,W) and (x,y,z) the joint location from the sensor. X returns the height.f

$$X = \frac{|Ax+By+Cz+D|}{\sqrt{A^2+B^2+C^2}} \tag{3}$$

The result of using the above calculation allows the 3D position of any joint to be calculated from the distance from the floor. The information can then provide data on the height of a certain joint from the floor, which is useful for certain measurements such as comparing the exact user's height against calculated height. There is a potential disadvantage for calculating height with this method, especially in a self-service environment. If the sensor is too close to a user and cannot get an accurate reading of the FloorClipPlane, there is likely to be higher variance in the reporting in the subject's height as inference will occur.

Regarding tracking, Microsoft makes recommendations to operators to ensure more accuracy when tracking users within the optimum range [146]. For example, interference with the IR sensor through wearing baggy or reflecting clothing or contrasting colours may affect the devices ability to locate a joint. Also, the environment in which the sensor operates will also affect results - low-lighting areas work better than brightly lit scenarios. Initial trials in this study looked at the lighting conditions before data collection, ensuring adequate lighting and space for the sensor to track accurately. Section 4.3 details more information on the effect of these variables on data collected throughout this work.

The Kinect SDK can identify the tracking state through the Joint.TrackingState property. A joint can either be ‘tracked’, ‘inferred’ or ‘not-tracked’. Each state is reported alongside the positional data for each frame. Tracked data signifies that the Kinect could fully determine the joints position and therefore confidence in the accuracy of the data is very high. Non-tracked means that there is no joint data available and inferred states are determined when the Kinect must rely on tracking data to report on the joints location, often signifying it couldn’t fully track the joint. Since inferred states are calculated, the confidence level in the data is low [146].

The Kinect SDK enables joint information to be adjusted (smoothed) across frames to minimise jittering and stabilise the joint positions over time. There are five smoothing parameters; Smoothing, Correction, Prediction, JitterRadius and MaxDeviationRadius [151]. The smoothing filter is based on Holt Double Exponential method. For this study, JitterRadius and MaxDeviationRadius were adjusted to 0.05 and 0.04 to filter any positions that deviate from raw data. The values were based on the C# example provided by Microsoft [151] which was recommended for scenarios such as gesture recognition in games. The filter provided some smoothing while retaining minimal latency, which was ideal when exploring self-service scenarios.

Kinect Accuracy, then, focuses on the 3D skeletal joint accuracy of the V2 sensor when used in a self-service biometric interaction environment. Although the Kinect has many uses through the ability to track a user, at this stage, the positional data and identifying regions in 3D space is key to enhancing the information displayed to the user throughout the transaction process. Research in this chapter investigates capturing joint data in multiple scenarios and seeks to report on the accuracy, identifying if the captured transaction is a truthful representation of the user’s movements. If successful, the outcome of the study will enable automated assessment methods within self-service scenarios such as border control implementations, and by doing so, enable biometric systems to enhance the ability to categorise performance metrics as well as relay high-quality feedback and information to the user.

Investigating previous work with both versions of the Kinect will establish its capabilities and in its stability and accuracy within multiple testing scenarios. A small data collection is introduced to verify the sensor's abilities to track a user within a border control scenario accurately.

4.2 Previous Work

Research into both versions of the Kinect device has typically investigated human Activity Recognition, using state of the art machine learning paradigms for applications such as posture and gesture analysis.

The ability of both devices to capture stable and reliable data is characteristically the primary concern for these studies through investigating the feasibility to track movements within specific scenarios. However, with the introduction of the Windows for Kinect V2 in 2013, there are only a handful of studies to date that has begun to utilise the upgraded device’s capabilities. The Kinect V1, on the other hand, has been extensively investigated, particularly in its ability to accurately report 3D depth data. A study by Khoselman and Elbernik [137] inquire into the geometric quality of depth data obtained by the Kinect V1 sensor. Through calibration and error analysis, the report found the accuracy of the depth component to be in error ranges of a few millimetres up to a maximum range of 400mm. The study recommended that depth data for mapping applications should be acquired within a 1 to the 3-metre range. The Kinect pose estimate is compared to more established motion detection techniques that used LED marker technology. Results found that in controlled body posture (standing, raising arms), the accuracy of the joint estimation was comparable to motion capture. However, where there was occlusion through sitting down, non-distinguishing depth (limbs were close to the body) or clutter (other objects in the scene) then the depth information was not as reliable. Variation between methods, for the V1 device, was found to be about 10cm.

Several studies use the Kinect device to investigate posture in a range of users. Obdržálek et al. [152] performed accuracy and robustness analysis of the Kinect V1 skeletal tracking in six exercises for the elderly population reporting on a high degree of precision for classified movements. Patients were asked to perform a variety of tasks such as sitting down, swinging a leg forward and lifting the knee up. Hoai-An et al. [153] report on a classification algorithm for human fall recognition based on Kinect V1 skeletal data. The Support Vector Machine (SMV) algorithm is programmed to recognise several fall and non-fall activities in different scenarios. Three experiments are conducted with a database from the human skeleton captured by the Kinect. Results conclude that the SMV algorithm could recognise falling scenarios with up to a 91.3% accuracy in the classification of a movement. In summary, the study concluded that future work would need to consider more scenarios to establish posture accuracy.

Clark et al. [154] report on the feasibility of the Kinect for postural control assessment. Twenty healthy subjects were asked to perform three tests; forward reach, lateral reach and single-leg eyes-closed standing balance. The Kinect data was compared to a 3D motion analysis system which had ‘comparable’ high inter-trial reliability and concurrent validity for most of the measurements. However, some biases were reported for some measures when looking at the sternum and pelvis evaluations. Several other papers have examined the body tracking accuracy for specific applications in physical therapy, such as upper extremity function evaluation [155], assessment of balance disorders [156], full-body functional assessment [157] and movement analysis in Parkinson's disease [158].

Improving the degree of accuracy in the reporting of joint position data obtained by Kinect has also attracted research into using multiple cameras. Asteriadis et al. [159] and Tong et al. [160] investigated using multiple Kinect V1 cameras for capturing joint position and depth data. A major problem for capturing any information from a single sensor setup is the occlusion of certain body parts, which is typically due to the placement of the sensor and the ranges it operates within. Both studies introduced several multiple camera setup methods from different perspectives to exhibit advantages, such as increased accuracy, and reduced shortcomings, as well as to reveal potential methodologies for image acquisition in Activity Recognition.

Yang et al. [161] report on early accuracy evaluations of the depth component of the Kinect V2. The study investigates accuracy distribution, depth resolution, depth entropy, edge noise and structural noise to assess the performance of the depth camera. The study reports ‘good accuracy’ if the object is within the optimal range as previously defined by Microsoft (47.5-degree angle and between 0.5m -3m distance), averaging a depth accuracy error of up to 2mm in some scenarios. The report concluded that some variables might affect the Kinect’s V2 performance, like the V1 sensor. Reflected objects and light-absorbing material (like carbon black) could cause issues with the IR light emitted by the Kinect sensor. However, there was no reported data establishing these differences in the depth components accuracy.

Dehbandi et al. [162] use depth data from the Kinect V2 to quantify upper limb behaviour and reports on the stability of the results. The sensor was placed on a tripod which stood at 0.92 metres from the floor; subjects were required to sit on a table placed 2.7m away from the sensor and consequently were within optimal tracking range of the Kinect V2. The study used the Wolf Motor Function Test (WMFT) protocol, an automated algorithm to collect data and classify behaviour. The score the WMFT produces details subject assessment of task performance in three areas; time, functional ability and strength. Each user was required to perform 15 functional tasks, with a maximum of 120s allocated for each. Movements required the subject to be in a seated position, and therefore only used 16 of the 25 available joints for analysis, not using any joints below the Hips. The WMFT calculated a classification algorithm and demonstrated up to a 91.7% classification accuracy with only six classification errors across the experimental conditions.

A feasibility trial of using the data from the Kinect V2 to determine postural stability in ‘healthy’ subjects was reported in early 2017 [144]. Twelve subjects were recruited and instructed to perform a sequence of postural stability tasks while on top of a force platform. The data were compared to the force platform and the Kinect V2 to quantify the degree to which the Kinect V2 was returning reliable data. An evaluation of the results showed a strong agreement between the Kinect and force platform classifiers, reporting a task classification of 87.8% accuracy in predicting which of the tasks were being performed.

Gonzalez-Jorge et al. [164] reported on the results of accuracy and precision tests on both the Kinect V1 and V2 sensors. The results were performed for different ranges and changing the inclination angle of the sensor. Results at a 1m range show similar precision for both sensors, Kinect 1 accuracy values ranging between 2mm and 12mm for a 1m range and between 4mm and 25mm for a 2m range. The Kinect 2 shows higher accuracy values between 0.1mm and 7.5mm for 1m range and 5m to 7mm for a 2m range. For precision, both sensors showed a decreasing result with an increase of range, although it was more prominent with the Kinect 1. For a 1m range, both sensors reported a standard deviation between 1.5mm and 6mm. At a 2m range, however, the Kinect 2 could provide a higher precision with values lower than 8mm while the Kinect 1 was over 10mm in most scenarios.

While most studies have reported on the Kinect’s ability to report on depth data or the accuracy of classifying a movement based on an algorithm, there is little research into analysing the accuracy of the joint positions obtained by the Kinect V2. Moreover, there is little to no research of using the sensor in a self-service environment whereby positional data is used to identify where the location of certain joints and relay information based on those conditions.

4.3 Data Collection

Three scenarios were designed to test the stability of the joint positional data captured within limits of a self-service environment. The main goal of this data collection was to ensure that the V2 sensor could track multiple joints accurately for stationary and movement based tasks. Joint location data is analysed to assess consistency across the transaction, confirming if movement measured by the sensor is made by the user and not what the sensor infers. The three scenarios were captured with the user facing the sensor and within the optimal limits of the sensor.

The first task required users to stand still on feet markers at three different locations from the sensor. The first position was at 1.0m, the second 1.3m and the third 1.6m away from the centre of the sensor. The markers on the floor indicated where each user was required to stand. The purpose of this experiment was to examine the variance in joint data over the transaction period. Each of the five participants was asked to stand still for a total of twenty seconds at each marker. During the task, the user was requested to keep joint movement to a minimum.

The second task requested users to move their right arm from their side to moving the arm forwards so that the arm was perpendicular to the floor. This was to mimic the token/fingerprint interaction movement. The goal of this task was to check the stability on the other joints not in motion when there is movement from other body parts. Each user was required to repeat this movement twice throughout the interaction, five seconds apart, again, on the 1.6m, 1.3m and 1.0m markers.

The final task required the users to walk between an additional marker 1.9m from the sensor. Users were asked to walk to the 1.3 marker, and then back to the start. The aim of the task is to establish stability in tracked joints across the movement in the interaction. The results from this task will outline the Kinect V2's ability to track a new user both enter and exit a self-service system. Results for this task largely assessed the Z measurement for all joints.

4.3.1 Data Capture

The system used for this data capture was built in collaboration with Purdue University (Indiana, US). The software was developed to further the work in HBSI automation, e.g. to detect and categorise errors in real time rather than using manual methods during post-processing. Z. Moore reports on classifying Human-Biometric Sensor Interaction errors in real time using an iris recognition system [165].

The Kinect system developed could export each frame with a timestamp, skeletal joint and X, Y and Z plane values from the 3D body point to an external file for analysis. The distance between the head joint and the floor plane was also calculated, reporting on the W metric for Head Joint only. All coordinates are displayed in metres. Additionally, the tracking state was obtained for each joint, identifying wherever the point was tracked, inferred or not tracked.

4.3.2 Users

Five users were recruited for this data collection. The only selecting criteria were that participants must be able to walk without the use of equipment. Participants were asked not to wear any obtrusive clothing (e.g. baggy hoodies, hats, bags) that may affect the stability tests. The participant's height and shoulder width were recorded for each trial. Height was measured from the floor to the tip of the head, and shoulder width was measured from each shoulder across the chest.

4.3.3 Scenarios

Scenarios are based on the three tasks as previously described. Each participant was required to complete each task once but for each required distance. Task 3 was only performed one time.

4.3.4 Guidance and Training

Participants were given guidance on where to stand for this data collection to comply with the goals for the data collection. Tasks were described to participants before capture and information were given during task indicating participants when to move (e.g. for Task 2 participants were asked to raise their hand within a five-second interval).

4.3.5 Experimental Setup

The Kinect V2 device was positioned on a professional tripod. The sensor was placed at the height of 1.8m from the ground and initially positioned 1.0m away from the feet symbols put on the floor. The camera was positioned so that all users would be within 43.5 degrees in the range of view (the recommended optimal range). However, the distance between user and device changed based on what task was being performed. Initial testing of the Kinect device confirmed the presence of the FloorClipPlane value.

4.3.6 Recording

Only the positional data from the Kinect device was recorded for this trial. No personal information or video footage of the participant’s interaction was recorded.

The data captured comprised of coordinate information for all 25 skeletal joints over the course of the transaction. Data was captured when subjects were in position, and each transaction was timed for a total of 20 seconds.

4.3.7 Data Storage

No sensitive information was captured from any of the participants. Skeletal information was stored on a spreadsheet and given a unique code based on the user, task and distance from the camera. The Kinect Data was stored on a secure hard drive backed up regularly at the University of Kent. No other information were obtained regarding the identity of participants other than the participant’s height and shoulder width.

4.4 Results

The results for each task will report on the percentage of fully tracked joints, the measured participant's height (defined from the distance of the head joint to the floor plane (W) and the measured distance between participant’s shoulders, through calculating Euclidean distance between the Shoulder Left and Right joint locations. Descriptive statistics and visual graphs will identify any abnormalities in the accuracy of obtaining information from the user. Results are separated by each task. Table 16 below details each participant’s measured height and shoulder width. Both measurements were conducted with a tape measure, reporting on the measurement in metres.

Table 16: Participants Height and Shoulder Width for Kinect Interaction

Participant	Height	Shoulder Width
1	1.82m	0.325m
2	1.68m	0.290m
3	1.75m	0.315m
4	1.66m	0.340m
5	1.88m	0.330m

4.4.1 Task 1

Users were asked to stand stationary on three different markers on the floor; 1.6m, 1.3m and 1.0m away from the centre of the Kinect Device. Table 17 below details some joints that were Fully Tracked for each of the three markers.

Table 17: Percentage of each Joint that were Fully Tracked for Task 1, Kinect Interaction (%)

Joint	1.0m	1.3m	1.6m
Head	100.00	100.00	100.00
Neck	100.00	100.00	100.00
ShoulderLeft	100.00	100.00	100.00
ShoulderRight	100.00	100.00	100.00
ElbowLeft	88.58	100.00	100.00
ElbowRight	88.98	100.00	100.00
WristLeft	78.69	100.00	100.00
WristRight	75.25	100.00	100.00
HandLeft	55.25	100.00	100.00
HandRight	49.64	100.00	100.00
SpineShoulder	100.00	100.00	100.00
SpineMid	100.00	100.00	100.00
SpineBase	98.85	100.00	100.00
HipLeft	95.50	100.00	100.00
HipRight	94.56	100.00	100.00

Overall, it can be observed that for stationary users at 1.3m and 1.6m distance from the sensor that there is a high degree of reported full tracking. Users 1.0m away from the sensor were not able to be tracked as accurately due to some body parts were not in range of the sensor and therefore marked as inferred.

The average height of the user is compared to the calculated height taken from the head joint to floor measurement for each task. The value measured by the Kinect was calculated as an average over the entire transaction.

Table 18: Participant Height: Measured against (Median) Calculated for Task 1: Kinect Interaction

Participant	Height	1.0M	1.3M	1.6M
1	1.82	1.90	1.86	1.83
2	1.68	1.72	1.69	1.67
3	1.75	1.74	1.75	1.77
4	1.66	1.74	1.65	1.65
5	1.88	1.72	1.82	1.84

As Table 18 demonstrates, the calculated height for each participant was accurate for tasks where users were standing further away from the camera. Higher deviations were recorded when calculating height for users standing 1.0m away from the system. Figure 25 demonstrates below.

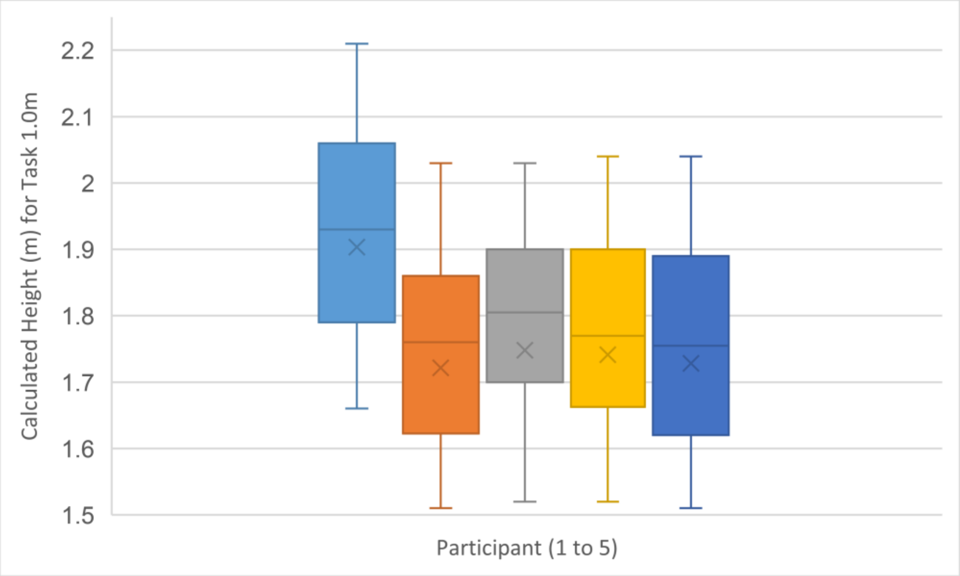

Figure 25: Box Plot Diagram for Calculated Participant Height for 1.0m (Kinect Interaction)

Calculated height varied highly for participants standing 1.0m away from the sensor. In some cases, a participant's calculated height varied up to 0.25m from the physical height of the user. While the Kinect could fully track the head joint for all distances, the most likely reason for the high variance was possibly due to occlusion of the floor plane by the positioning of the user. Users who were closer toward the camera may inhibit the sensors ability to detect the floor plane. As results for the lower body (Hip Joints and below) were either reported as inferred or non-tracked for 1.0m, this is the most likely reason for the result.

Table 19 below shows a comparison of the measured and Euclidean (calculated) distance between Shoulder Left and Shoulder Right joints measured by the researcher and by the sensor respectively. The Euclidean distance calculated based on an average measurement across the joints across all frames for all three interactions.

Table 19: Standard Deviation between Measured and Euclidean Calculated Shoulder Distances for Distance for Task 1

User	Measured	1.0m	st.dev	1.3m	st.dev	1.6m	st.dev
1	0.325m	0.338m	0.0842	0.327m	0.01568	0.384m	0.01568
2	0.290m	0.315m	0.0931	0.359m	0.01431	0.323m	0.02431
3	0.315m	0.312m	0.0954	0.311m	0.01546	0.333m	0.03546
4	0.340m	0.345m	0.0530	0.344m	0.05830	0.325m	0.02830
5	0.330m	0.327m	0.0841	0.338m	0.06431	0.366m	0.01431

There was a minimal difference between the physical and sensors measurement of the length between the shoulders in all tasks. Standard Deviation closes in for participants standing further away, indicating more accuracy for measuring the distance between shoulder joints. Although this was for users who were standing completely still, it is a result that defines the Kinect’s ability to track accurately within an optimal view range.

In addition to looking at the accuracy of the calculated participant height and distance between shoulder Joints, Table 20 below reports on the average variance in the movement for all joints across the three distances. With the inclusion of the smoothing filter which was applied , an ideal result is a smaller variance between data for each frame.

For stationary capture, 1.3m or higher for this scenario results in more tracking with the Kinect Sensor V2. For tracking the lower body, a lower variance is measured for users at the 1.9 markers. However, this distance may be impractical because on facial capture and with the inclusion of additional luggage that may be placed by the user's side.

Table 20: Mean-Variance for Z Joints across Task 1. ('+' variance <0.005, '++' variance <0.0005) for Kinect Interaction

Joint	1.0m	1.3m	1.6m
Head	+	++	++
Neck	+	++	++
ShoulderLeft	+	++	++
ShoulderRight	+	++	++
ElbowLeft	+	++	++
ElbowRight	+	++	++
WristLeft	+	++	++
WristRight	+	++	++
HandLeft	+	++	++
HandRight	+	++	++
SpineShoulder	+	++	++
SpineMid	+	++	++
SpineBase	+	+	++
HipLeft	+	+	++
HipRight	+	+	++

Although there are no formal requirements for how far the camera should be located from the user, facial capture must be at a 45-degree angle and able to capture photos from users who are between 1.4 and 2.0m tall (Kinect V2's Field of View) [104]. Standing too far away from the camera may result in lower quality images, which may lead to lower verification rates.

4.4.2 Task 2

The second scenario introduced a movement with the right arm while users were standing still on one of the distance markers. Users were asked to imagine there was a fingerprint sensor in front of them and to move to presenting their finger to a sensor while keeping still and looking forward. The movement was repeated twice with a five-second delay between each movement. Users will not make the same movement (i.e. move their wrist to an exact location each time) when repeating an action. The goal of this task was to first, check that joints were able to be tracked as accurately when compared to stationary movement in Task 1 and secondly, ensure that the body position is represented in the data. The results also detail the accuracy of the Kinect by again comparing the height of the user and the measurement across the shoulders.

Most joints will be stationary, and there should only be movement in joints relating to the right arm. These results will be shown through investigating variances in data over the course of the transaction. Firstly, Table 21 below details the percentage of joints that were classed as fully tracked throughout the interaction for all users.

Table 21: Percentage of Joints that were Fully Tracked for Task 2, Kinect Interaction (%)

Joint	1.0m	1.3m	1.6m
Head	100.00	100.00	100.00
Neck	100.00	100.00	100.00
ShoulderLeft	100.00	100.00	100.00
ShoulderRight	100.00	100.00	100.00
ElbowLeft	92.06	100.00	100.00
ElbowRight	91.47	100.00	100.00
WristLeft	87.29	100.00	100.00
WristRight	85.25	100.00	100.00
HandLeft	74.25	100.00	100.00
HandRight	76.11	100.00	100.00
SpineShoulder	100.00	100.00	100.00
SpineMid	99.86	100.00	100.00
SpineBase	97.54	100.00	100.00
HipLeft	95.50	100.00	100.00
HipRight	94.56	100.00	100.00

Table 21 details a high percentage of fully tracked joints for users standing 1.3m and 1.6m away for the sensor. The sensor had some difficulty with tracking joints for users 1.0m away from the sensor, with less than ideal results for following the movement in the specific joints required for the task.

The measured and calculated distance between shoulder joints is compared in Table 22 below. Calculated joints are taken as an average over the transaction including the movement of extending the right arm.

Table 22: Participant Shoulder Width: Measured against (Mean) Calculated for Task 2, Kinect Interaction

User	Measured	1.0m	st.dev	1.3m	st.dev	1.6m	st.dev
1	0.325m	0.352	0.0957	0.330	0.01218	0.332	0.01461
2	0.290m	0.325	0.0846	0.285	0.01651	0.285	0.01253
3	0.315m	0.294	0.0995	0.317	0.01366	0.318	0.01574
4	0.340m	0.368	0.0794	0.338	0.04590	0.345	0.01689
5	0.330m	0.287	0.0942	0.335	0.03235	0.376	0.01720

Even with the arm extension, the Kinect could distinguish the distance between shoulder joints accurately for users 1.3m and 1.6m away for the sensor. Users were asked not to make any major movements using their shoulder and by extending the right arm forward only, which enabled the shoulder joint to remain consistent across the scenario. The results indicate a slightly larger standard deviation when compared to Task 1. However, this was to be expected as there was a small movement in both position and rotation in the movement.

To gain an insight into the Kinect's vision, Figure 26 below shows some joints involved in the right arm extension which was tracked over the length of the transaction.

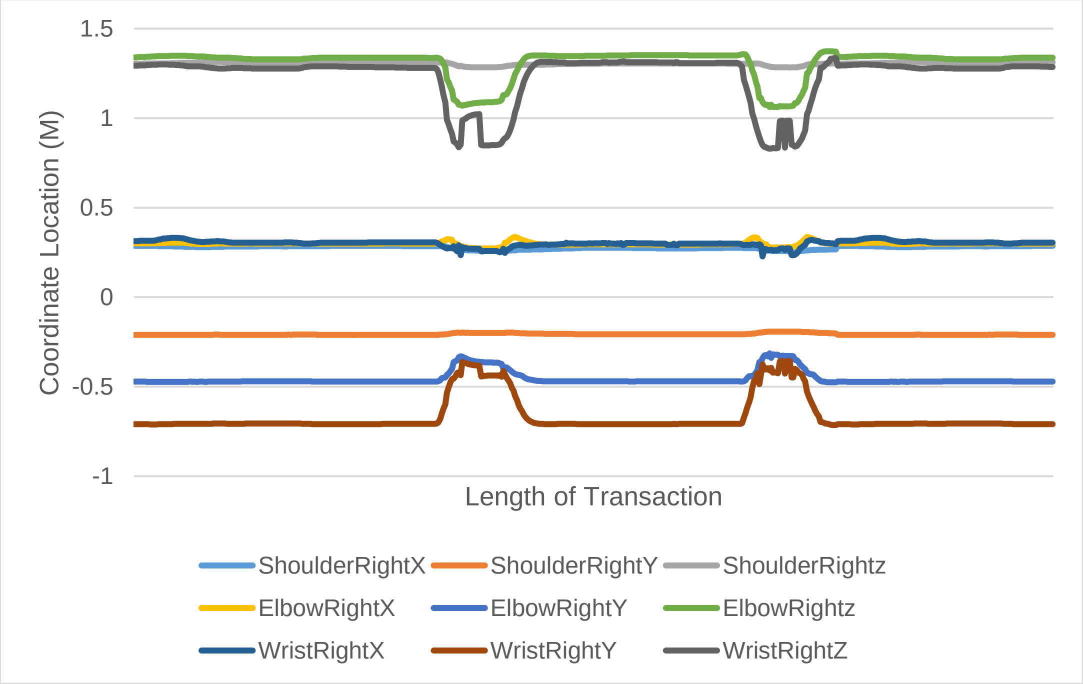

Figure 26: Critical Joint Co-ordinates for an Individual Participant for Task 2, 1.6m for Kinect Interaction

The data reported for one participant's interaction details the movement of the right arm moving forward. The ElbowRight joint remains consistent across the scenario, highlighting only precise movement for ElbowRight and WristRight joints.

Participant Height was again investigated to compare the stability of the measurement of the distance between the Head joint to the floor plane. Table 23 below reports.

The accuracy of the Kinect sensor again increased for subjects further away at 1.3m and 1.6m. A larger inaccuracy was reported for users at 1.0m away, with an average difference of up to 0.14m.

Table 23: Participant Height: Measured against (Mean) Calculated for Task 3, Kinect Interaction

Participant	Height	1.0M	1.3M	1.6M
1	1.82	1.96	1.80	1.85
2	1.68	1.71	1.67	1.67
3	1.75	1.69	1.75	1.77
4	1.66	1.70	1.64	1.68
5	1.88	1.83	1.86	1.86

To conclude, Table 24 below details the average variance in depth joint data across the interaction, stating different thresholds for levels of differences. While a lower variance in stationary joints are favourable (joints should not be moving as much compared to the joints that are used in the movement), it is expected that there should be a higher variance for joints associated with moving the right arm forward (marked with an ‘m’).

Table 24: Mean-Variance for Z Joints across Task 2 (‘m’ variance >0.005, ‘+’ variance <0.005, ‘++’ variance <0.0005) for Kinect Interaction

Joint	1.3m	1.6m	1.9m
Head	+	++	++
Neck	+	++	++
ShoulderLeft	+	++	++
ShoulderRight	+	++	++
ElbowLeft	+	+	++
ElbowRight	m	m	m
WristLeft	+	++	++
WristRight	m	m	m
HandLeft	+	++	++
HandRight	m	m	m
SpineShoulder	+	++	++
SpineMid	+	++	++
SpineBase	+	+	++
HipLeft	+	+	++
HipRight	+	+	++

As Table 24 demonstrates, a higher variance in joint data was found for the critical and associated joints for the right arm movement. A larger variance was obtained during the movement of the right arm for all critical joints. Although this is to be expected with a movement, it shows that those other joints during the interaction remain very stable and that there was minimal movement recorded during the process. Of course, this was a controlled scenario and results will differ in live testing; however, this task has proven that the Kinect can accurately track the user within this type of scenario within the setup reported.

4.4.3 Task 3

For the final task, users were asked to simply move forwards from the 1.9m marker to the 1.3m and back to the start by walking backwards. The data recorded in this study will have very high variances in the joint co-ordinates due to the movement, but the main idea of this task was to assess the ability to measure the skeleton across the scenario. The results then look at the ability for the Kinect Sensor to establish the user at and between the 1.9 and 1.3 markers. The user stopped movement at both markers indicating the start; middle and end of the transaction respectively. To begin, however, Table 25 below details the percentage of joints that were fully tracked in this scenario.

Table 25: Percentage of Fully Tracked Joints over Task 3 for Kinect Interaction

Joint	Task 3
Head	100.00
Neck	100.00
ShoulderLeft	100.00
ShoulderRight	100.00
ElbowLeft	98.59
ElbowRight	97.85
WristLeft	89.25
WristRight	88.44
HandLeft	81.58
HandRight	83.37
SpineShoulder	100.00
SpineMid	100.00
SpineBase	100.00
HipLeft	98.70
HipRight	99.46

The sensor again could track most joints to a high degree of accuracy. While this may seem lower than previous studies, it is important to remember that users were walking towards the camera and therefore as they got closer, the Kinect tracking state may turn from tracked to infer for some joints, especially those in the lower body.

To determine the Kinect's ability to track movement within the task, Figure 27 below detail several joints tracked distance over the course of the transaction. The joints selected in this task follow the spine from the base to the head.

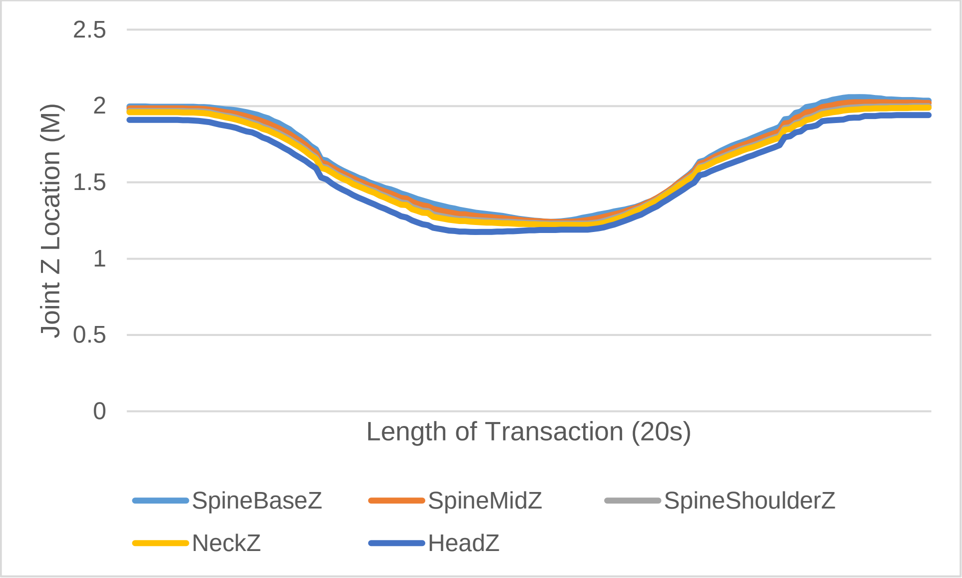

Figure 27: Spine Joints Location for an Individual Participant over Transaction Period for Kinect Interaction

Figure 27 above displays the depth data for many joints for an individual interaction. The smooth curve represents the movements of many joints associated with the spine. This figure outlines the smooth transaction of a user walking forwards (towards the sensor) and then back to the sensor.

Table 26 below details the calculated distance between the shoulder joints for users moving towards and backwards from the sensor. The user was facing towards the sensor for the entire transaction, so there were no significant movements in the joint position.

Table 26: Standard Deviation between Measured and Euclidean Calculated Shoulder Distances for Task 3

User	Measured	Task 3	Difference	st.dev
1	0.325m	0.345m	0.02	0.1357
2	0.290m	0.281m	-0.009	0.1251
3	0.315m	0.295m	-0.02	0.0972
4	0.340m	0.364m	-0.024	0.0855
5	0.330m	0.348m	0.018	0.0994

Results indicate that the movement did impact the sensors ability to measure the distance between the two joints accurately. When compared to previous scenarios, the standard deviation and average calculated position differ from the ground truth data. The deviance in these results indicates that there were discrepancies in the reported joint position across the task. As explained in Section 4.1, the sensor ‘sees’ a joint from the point of origin and so as the subject moves closer or away from the sensor, the calculated distance is likely to change due to the new position of the studied joints. However, while the standard deviation was relatively high, the calculated average for users 1-3 was less than +/- 10cm. An average difference of 24cm and 18cm was reported for users 4 and 5.

Looking at the accuracy of calculating participant height, Table 27 reports on the calculated descriptive statistics for Task 3.

Table 27: Participant Height: Measured against (Average) Calculated for Task 3: Kinect Interaction

Participant	Height	Task 3	st.dev
1	1.82	1.84	0.0113
2	1.68	1.70	0.0195
3	1.75	1.74	0.0200
4	1.66	1.63	0.0114
5	1.88	1.86	0.0214

Although this was a straightforward task, the results once again indicate the sensors ability to map a user in motion to a high degree of stability. The next part of this chapter discusses the outcomes of this experiment in further detail.

4.5 Conclusions

The results confirm the Kinect V2 sensor’s ability to track user movements within controlled scenarios accurately. Introducing the Kinect in Section 4.1, the capacity to measure accurate positional data by correct positioning of the sensor and smoothing techniques is outlined. Additionally, a comparison of the improvements over the original sensor is made while discussing the ability to use the FloorClipPlane to improve the likelihood of precise information and to also obtain data such as the distance of a joint from the floor.

Previous research described in this chapter outlines the devices features and limitations through various studies. Studies using both Kinect devices typically report on the accuracy of depth data and the ability to classify a movement using a predetermined classification algorithm. Additional sensors introduced increase accuracy but require extra room to set-up, a luxury ABC system may not have. The most common conclusion from the literature review demonstrates the optimal position for a sensor is within Microsoft’s recommended guidelines for setup. Studies on the accuracy of the positional data obtained by both devices are limited.

There is very little to no work available on the Kinect's ability to distinguish body positional data to track the user in real time.

In this chapter, three tasks were presented to ascertain the Kinect's ability to track movements over a 20-second transaction. The goal was to establish the Kinect's ability to track a single user within in a self-service scenario.

Trials with the Kinect sensor throughout the state of the art research agree that the device can accurately track the human skeleton when used in ideal conditions, e.g. through the optimal positioning of the sensor, by making certain movements and wearing non-interfering clothing. The latter may be a factor when investigating border control scenarios, as travellers may be wearing baggy clothing, accessories and backpacks. However, if implemented, a recommendation may be only to use fully visible joints that can be fully tracked throughout the interaction; upper body joints may provide enough information for facial verification scenarios. The benefit of the Kinect is that each joint reveal information about the user's position. For instance, if they are too far from the sensor or are in movement. By comparing the position of multiple joints, a representation of the user's body can be obtained. For face verification, it may only be necessary to determine the head position, as this is the only 'object' and that is physically required. The only time other joints maybe required is for token and fingerprint interactions; assessing if a movement has been made to a sensor and whether there are errors within that interaction.

The inclusion of bags, loose items of clothing and luggage may cause unreliable results in tracking the joints in the body, leading to inferred and non-tracked data. This will have an impact on the ability to capture within an ABC verification process, mainly due to the restrictive build and design of these systems. The results outlined in this Chapter indicate that the optimal position for a stable set of joints is for sensors placed 1.3m to 1.6m away from the user. While the Kinect can fully track several joints at 1.0m away, there were larger discrepancies in the accuracy of measuring distance between joints. The trade-off then is the design of the ABC system; the build of the eGate is usually between 2-3m in length but this could vary between systems. The token and biometric sensor is then usually positioned toward the back end of the ABC systems, but close enough to the user that no additional effort should be needed. This then limits where a sensor such as the Kinect can be placed, as another 1.3-1.6m from a sensor may require larger builds in already restricted spaces.

If the sensor were ever to be trialled in deployment scenarios, caution would be needed to make sure users place luggage to the side and to assist with tracking as much as possible. Removal of hats will aid tracking of the face and will be required for face capture regardless. Fingerprint Interactions will be even trickier, the placement of the Kinect, gate and fingerprint sensor may inhibit the ability to capture joint movements within that scenario.

Nevertheless, the sensor has its advantages within controlled self-service scenarios. Capacity to measure and obtain certain metrics such as joint positional data may enable systems to develop dynamic feedback, identifying correct or incorrect presentations in real time and enhancing the quality of the guidance given to the user. Additionally, measuring users by looking for everyday movements, the height of the traveller and so forth has scope beyond live deployments of a system. In operational testing, flaws may be highlighted by analysing captured data. By coupling positional data with performance assessment may reveal anomalies during an interaction, answering questions such as how long does the transaction take to complete and why? Are users struggling to interact with a sensor? Are there any obstacles the user is having the difficulty that may affect time on task or sample quality? The ability to track and analyse movements then may highlight problematic areas within the system.

In summary, the key contributions from this chapter are as follows:

- A review of literature in the human Activity Recognition field, specifically looking at the applicability of the Kinect Sensor within a self-service environment. Differences between both versions of the Kinect are explored. The research proposes different techniques and methodologies in improving the accuracy of the data reported by the device
- Investigation of the Kinect’s ability to accurately report on the joints position and location in three scenarios; standing still, moving an arm forward and walking forwards and then backwards
- The initial study with the Kinect has outline optimal position of the sensor for the next data collections

Chapter 5 and 6 introduce the Kinect V2 sensor to facial and fingerprint interactions, defining the applicability of using the sensor within a self-service environment using the two biometric modalities.

CHAPTER 5. FINGERPRINT INTERACTION

5.1 Introduction

The main aim of this study was to explore the ability to track and analyse the impact of erroneous behaviours on fingerprint interaction. Also, multiple sensors are compared between a flat and a raised surface to compare HBSI metrics. Fingerprint Interaction in ABC systems has been previously reported in Chapter 2.

A system was developed to enable users to interact with a self-service fingerprint system. The implementation uses the Kinect V2 sensor to track the user's position by collecting and analysing the user's actions involved throughout the transaction.

In addition to exploring the tracking accuracy in a self-service fingerprint system, the Kinect program also took the first steps into HBSI automation, categorising interactions in real time using positional data. The program is therefore referred to as the HBSI Automation program for this study.

Fingerprint Interaction can be analysed via several measures in biometric systems; biometric transaction times (throughput), quality of the sample and the matching score against previously enrolled images. A statistical analysis is provided to establish wherever the variables introduced in this study have a significant impact on throughput and sample quality metrics.

5.2 Data Collection

The experimental procedure details the setup, the information recorded and users recruited to the study. Additionally, this section details what guidance and training are given to participants.

5.2.1 Experimental Setup

Feet symbols were placed centrally in front of a kiosk to indicate to participants where to stand. Although the use of feet symbols is typically to provide visual instructions to users where to stand during facial verification [7], the purpose of including these in this study was to make sure participants were within the range for both Kinect recording and fingerprint interaction. There were no physical gates or barriers used for this collection due to the size of the room.

Five fingerprint sensors were stationed on a pedestal kiosk. The kiosk was 100cm tall and 105cm wide. The pedestal used in this collection was an ideal height for the experiment, matching the height of most systems that use a fingerprint modality. The NIST study on user interaction with fingerprint devices for kiosks of different height [55] previously determined that from the analysis, a height of 0.91m yielded faster performance results and was the most comfortable for travellers in the US-VISIT scenario.

The sensors were positioned on an adjustable slope platform (Figure 28 below) which could be positioned at a flat angle or a gradient of 20 degrees. The angle was measured using a protractor. The slope mechanism was securely positioned in the centre of the kiosk to limit the movement of sensors throughout the study.

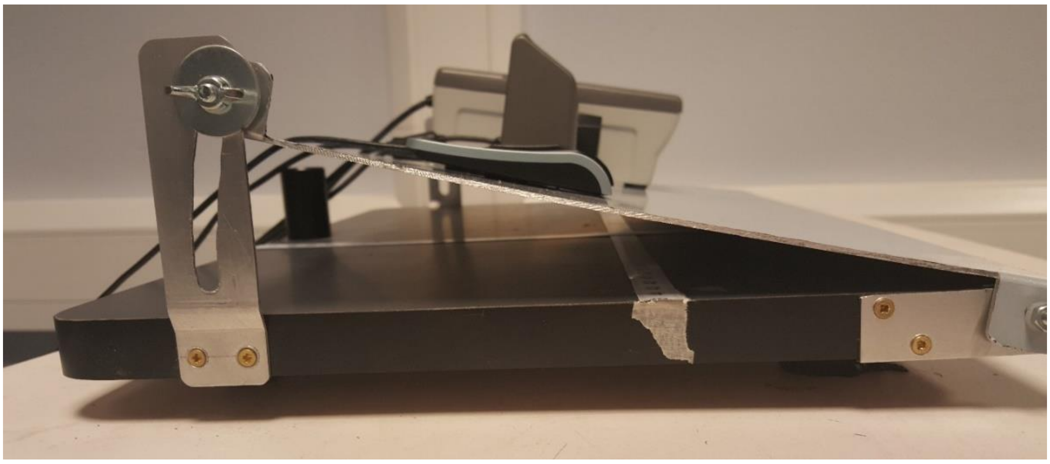

Figure 28: Slope Adjustment Mechanism with Fingerprint Sensors

The fingerprint devices that were used in this data collection were collected from several resources. Two of the fingerprint sensors, the UPEK touch and swipe sensors, were purchased from a UK distributor. The other three sensors had previously been purchased and stored for research use within the School of Engineering and Digital Arts at the University of Kent. Table 28 below details the sensor type and original manufacturer. All sensors were connected via USB 2.0.

Each sensor was separated on the kiosk by 5 cm (Figure 29). These distances were checked after each experiment to ensure consistency across scenarios. The front of the sensor was measured to be around 1.0m away from the Kinect sensor.

Participants viewed instructions from the system on an 18” BENQ monitor, which was placed on a free-standing custom-made iron stand placed 45 cm behind the kiosk.

Table 28: Sensors in Fingerprint Interaction

Sensor	Type	Resolution	Image Size	Manufacture
UPEK Eikon Solo Swipe	Capacitive	508ppi	144 pixels (width)	Digital Persona, Inc. Previously by UPEK, Inc.
UPEK EikonTouch 700	Capacitive	508ppi	256 x 360 pixels	Digital Persona, Inc. Previously by UPEK, Inc.
Hamster IV	Optical	508ppi	258 x 336 pixels	SecuGen
Hamster Pro 20	Optical	500ppi	300 x 400 pixels	SecuGen
CrossMatch Verifier 300 LC 2.0	Optical	500ppi	31 x 31 mm capture area	Cross Match Technologies Inc.

Figure 29: Sensor Setup for Fingerprint Interaction (From Left to Right: Eikon Swipe, Eikon Touch, Hamster IV, Hamster Pro and Crossmatch LC)

The biometric acquisition and processing component was based on the Neurotechnology MegaMatcher SDK (Version 8.1) [166]. The MegaMatcher SDK is designed for large-scale Automated Fingerprint Identification Systems (AFIS) and multi-biometric systems developers. The fingerprint engine is based on the NIST Minutiae Interoperability Exchange (MINEX) [167] which is a fingerprint template standard, which creates the possibility of a fully interoperable between tokens and systems. The standard is capable of matching rolled and flat fingerprints. Enrolled images captured during the initial stages were matched using the algorithm.

Two computers were used for data capture, both controlled by the investigator. One computer ran the Neurotechnology MegaMatcher program, on a Windows 7 System. The computer was placed on the pedestal, positioned below the monitor. Two monitors connected to the PC where the display was extended over both monitors. The other computer operated the Kinect HBSI Automation and ISpy Surveillance Software [168] on a Windows 10 based system. Additional effort was made to separate both the kiosk area from any computers, cables and other devices that would not be found in a typical border control setting.

The Kinect V2 sensor was placed 1.3m away from the sensors, 1.8m off the floor. The tilt was angled centrally toward the users. Video footage was captured by four Logitech HD webcameras using the ISpy Surveillance Program; one camera (C2) was set up to record an overview of movements on the feet icons so the positioning of users could be categorised. Two cameras recorded footage of the fingerprint interactions, one from a top-down view (C1) and one (C3) from a side view. See Figure 30 for the room layout configured for Fingerprint Interaction.

Figure 30 details the experimental room. Data capture is initiated when the user is standing on the ‘Start Position’, which is marked by tape on the floor which was 2.4m away from the sensor. Users walked toward the feet symbols, which were 25cm in length. The front tip of the symbols (toes) was 1.15m away from the sensor, and the back of the icons (heel) was 1.4m away from the Kinect V2 Sensor.

All video footage was compared against the Kinect categorisations and evaluated in Section 5.3.1.

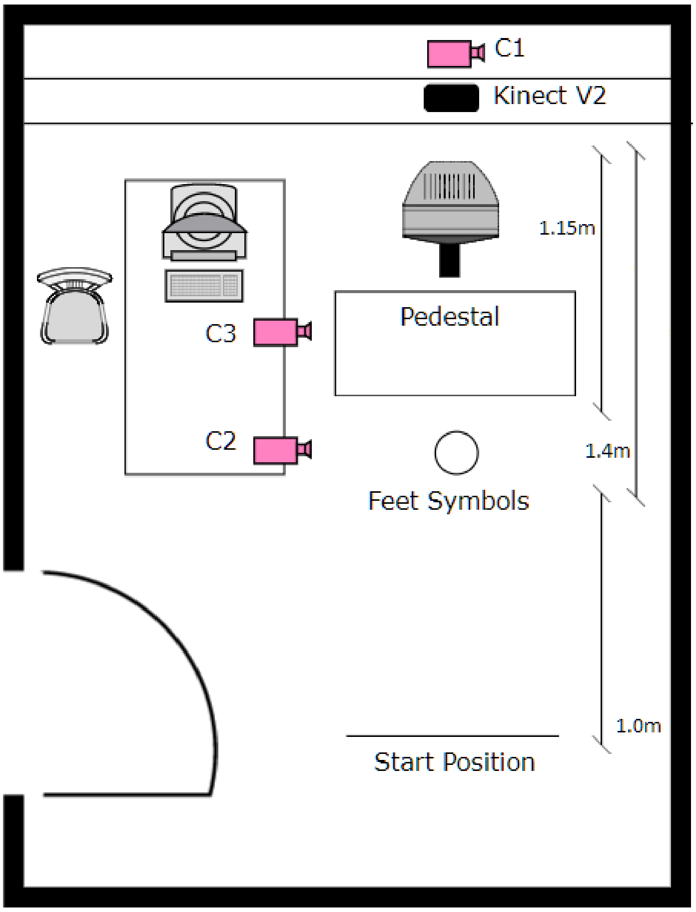

Figure 30: Experimental Room Setup (All measurements provided are in metres) for Fingerprint Interaction

5.2.2 Recording

Data was collected from several resources. The Neurotechnology MegaMatcher SDK captured biometric samples with timestamp information, which was used to calculate timings (time taken to complete a task, overall time etc.)

Fingerprint images and their respective NFIQ score were captured and stored on a secure hard drive. NFIQ is an image quality algorithm that was engineered by the NIST Biometric Image Software (NBIS) package [169]. The NFIQ algorithm reports quality scores on a nominal scale from one to five, with one being perceived as the best quality and five being the lowest. However, in the Neurotechnology program used in this system, the NFIQ algorithm is reported on a reverse scale, where images are rated a star quality; five for best quality and one for being the lowest.

An advantage of the NFIQ algorithm is that it is an open source tool for measuring the quality of fingerprint images independent of the fingerprint verification software used. Although not currently used during verification at border control, NFIQ ratings can provide an insight into the quality of captured images across a range of scenarios. Popular among previous work, acquired fingerprint images often have set enrolment thresholds for prints from the thumb, index finger, middle finger and ring finger at a NFIQ of one or two. In this data collection, all enrolled index finger images were captured at a NFIQ of four or five (reversed scale).

Matching Score was determined by VeriLook 5.1 SDK provided by Neurotechnology [166]. The Finger Matcher performs fingerprint template matching in 1:1 verification, matching the enrolled image against the captured images. The matching threshold is linked to FRR, the higher the threshold, the lower is FAR and higher FRR. Table 29 details the matching thresholds.

Table 29: Matching Threshold to FAR (as reported by Neurotechnology [166])

FAR	Matching Threshold (Score)
1%	24
0.1%	36
0.01%	48
0.001%	60
0.0001%	72
0.00001%	84
0.000001%	96

Three web cameras were set up to record video footage of the experiment, defining ground truth actions throughout the tasks. Video footage was reviewed after the experiment, and the actions/behaviours were recorded on a spreadsheet, defining each a timestamp for each task and notes on the images. A personal profile (not identifying the user) was collected and established post-data collection. See Table 30 below for further information on the data gathered.

As discussed previously in our Behavioural Framework defined in Section 3.5, there were four behavioural led tasks to this data collection experiment. Aligning these critical movements to our Behavioural Framework allows the breakdown of performance measurement at a task level. Table 30 below explains the behaviour led tasks and their associated variables for this experiment.

Table 30: Recorded Information for Fingerprint Interaction

Information	Description	Resource
Timestamp – HH: MM: SSS.SS	Printed and Exported to individual log file identified by a unique code and scenario number – processed when a form is first displayed	Neurotechnology MegaMatcher SDK
Fingerprint Image	Jpeg file	Neurotechnology MegaMatcher SDK
Fingerprint NFIQ Score	Quality Score (1-5)	Neurotechnology MegaMatcher SDK
Behaviours/Movements	Video Footage (.avi file)	Cameras/ISpy
Personal Profile	See Section 5.2.3	Questionnaires
Task Evaluation	See Section 5.2.3	Task Evaluation Form
Skeletal Joint Data	See Section 5.3.1	HBSI Automation Kinect Program

Table 31: Fingerprint Interaction Tasks and Associated Variables

Task	Related Variables
Entry (1) – Movement from starting position to feet symbols. Defined between FeetForm and ReadyForm	FeetForm (Timestamp) Feet_1 (Behaviour) ReadyForm (Timestamp) Time_1 (Time taken to complete)
Biometric Read (2) – The movement of placing the finger onto the sensor Defined as interaction between ReadyForm And CapturedForm	ReadyForm (Timestamp) Face_2 (Behaviour) Finger_2(Behaviour) CapturedForm(Timestamp) Time_2 (Time taken to complete)
Biometric Accept (3) – Response from System Defined between CapturedForm and TrialCompleteForm	CapturedForm (Timestamp) Face_3(Behaviour) Time_3 (Time taken to complete) TrialCompleteForm (Timestamp)
Exit (4) – Movement from system to starting position	Feet_2 (Behaviour) Overall Time (Total Time)

In addition, some face behaviours were captured to see wherever the participant was looking at the screen or their fingerprint. See Table 31 for reference to the forms which were displayed.

In total, four forms were displayed. Each form was based on icons adapted from ISO standards [104] and was shown on the freestanding monitor. Each form was displayed at the start of the four critical tasks involved in the interaction process (Table 32).

Table 32: Form Description for Fingerprint Interaction

Form	Description
Feet Form	Form displays instructions requesting the user to move to the feet symbols
Ready Form	Displays instruction to the user to present finger to the sensor
Captured Form	Informs user capture is successful
Trial Complete Form	Informs user process is over and to exit system

Behaviours were encoded a numerical value to assist in statistical analysis and are assessed in the Kinect analysis Section 5.3.1. Table 33 below describes the information collected from the Personal Profile, ABC Questionnaire and Task Evaluation forms. Results are further discussed in Section 6.3.2.5 and Section 7.2.

Table 33: Collected Information for Fingerprint Interaction

Personal Profile	ABC Questionnaire	Task Evaluation
Gender, Ethnicity, Height, Age, Handedness	Terminology (Yes/No)	Information Evaluation (1-5)
Accessibility	Experience (Yes/No)	Fingerprints (1-5)
Fingerprints Captured Before, Fingerprint Training	Training (Yes/No)	Results (1-5)
ABC Systems Before, If yes, which modality	Knowledge of Biometric Systems (Yes/No)	Conclusion (Descriptive Feedback)
Times Travelled, Travel Alone/Companions Illness, Hours of Sleep Accessories Temporal Illness (Burns, Cuts)	Descriptive Feedback (On ABC Systems, airports, atmosphere, queues)	

Table 34 summarises each system used in the study and the various devices and output associated.

Table 34: Various systems, connected devices and output files for Fingerprint Interaction

System	Device(s)	Output
Windows 10 PC – HBSI and Kinect Automation Program	Kinect Sensor	BodyValues spreadsheet, containing timestamps and processed x, y and z 3D spatial coordinates. Records 30 frames per second
Windows 10 PC – ISpy Surveillance System	Cameras 1-3	Video clips recorded at 720p from each camera. All recordings are initiated on movement detection and end after 10 seconds on non-movement within the detection zone
Windows 7 PC – Customised Neurotechnology Program	Fingerprint Scanners (1,2,3,4 & 5), Monitor	Timestamps when the process was started, each form was displayed, the fingerprint was captured. Fingerprint Samples with associated NFIQ ratings

Additional information such as which HBSI categorisations and the number of assists were noted during on paper during data collection.

The General Model of the System can, therefore, be defined through Table 35. The associated steps from the GM are identified, and the information displayed to the user to initiate an action. The system used in this experiment did not contain a token reader.

Table 35: General Model of the system for Fingerprint Interaction

GM	Step	Form	Assessment
V1	User enters system	Welcome Form	Kinect
V4	Presentation of biometric characteristic	Fingerprint Form	Kinect, HBSI
V4	Sample Acquisition and Processing	Processing Form	Kinect, HBSI
V5	Biometric Subsystem Decision	Matched Form	HBSI

5.2.3 Users

For this pilot study, participation was open to anyone over the age of 18. Individuals were recruited based on availability and the ability to communicate in English. There was no specific criteria for participants based upon their previous use. The following demographic information examined the participants:

- Age: 18-24, 25-34, 35-44, 45-54, 55-64, 65+
- Gender: Female, Male, Other
- Handedness: Left, Right or Ambidextrous
- Height: In centimetres
- Country of Origin
- Accessibility: participants were asked to state any attributes regarding Accessibility

Also, several more questions included in the demographic questionnaire helped to build a personal profile of subjects. Information collected beyond the demographic information above is discussed further in Chapter 7 Recommendations.

- Have you had your fingerprints electronically captured before? Yes/No
- Have you used an Automated Border Control system before? Yes/No
- Have you had any training to use a fingerprint of facial biometrics before this data collection? Yes/No
- Have you travelled by air in the last three months? Yes/No
- How many times do you travel abroad in a year?
- Do you follow the latest technological updates?
- Do you typically travel alone, with companions or with both when travelling abroad?
- Have you experienced any symptoms of illness in the last few days? (Cold, Flu, etc.)?
- How many hours of sleep did you receive last night?
- Please indicate wherever you are suffering from any temporary illnesses and if so, where are they located? (e.g. cuts, bites, burns, allergies)

When an individual arrived for their visit, they were instructed to read the Participant Information Sheet (PIS), ask any questions about the study and if they were comfortable to sign the consent form.

Participants were informed that if they wished to withdraw from the study that they could do so at any time. However, no participants withdrew from the data collection.

To participate in this study, each participant provided consent to the collection of his or her profile (demographic information), fingerprints and video recordings of their interaction with the fingerprint sensors. Each participant was assigned a unique identification number that was linked to the video, fingerprint images and

Kinect data. Names were not associated with the data collected. However, video recordings did include footage of the participant's face, which was used to analyse individual behaviours.

Subjects were required to enrol their fingerprints before interaction. Ten fingerprints of all fingers were collected on a CrossMatch LC Guardian sensor, a duplicate of the fifth sensor used in this data collection.

All fingerprints captured from participants were deleted immediately after obtaining both NFIQ quality and matching scores. All data was analysed offline and saved securely on a computer hard drive that was protected by a password.

The data collection had minimal risk of stress or discomfort for participants throughout the experiment. All participants could withdraw from the study at any time. Participants were also given plenty of time to study the PIS and ask questions before the experiment. Ample rest time was given between attempts to maintain an acceptable level of comfort during the study. Participants were not subjected to any level of additional risk when using the fingerprint sensors then they would do so when interacting with conventional devices such as phones, laptops, and USB drivers.

Seven participants were female and 13 males. Nineteen were right-handed and one participant was left handed. All 20 subjects stated that they travelled abroad by aeroplane in the previous year. However, only five identified that they used an ABC system upon returning to the country. Nineteen participants typically travelled with companions while only one subject travelled alone. All 20 participants held a current electronic passport. On average, participants slept for 7.4 hours the previous night before starting the data collection. Some participants stated that they were wearing additional accessories, such as glasses, a hat and contact lenses. Although contact lenses would have no effect on this study, accessories and items of clothing may influence the Kinect's ability to track skeletal data by interfering with skeletal tracking points as discussed in Chapter 4. See Figure 31 for a breakdown of descriptive demographics.

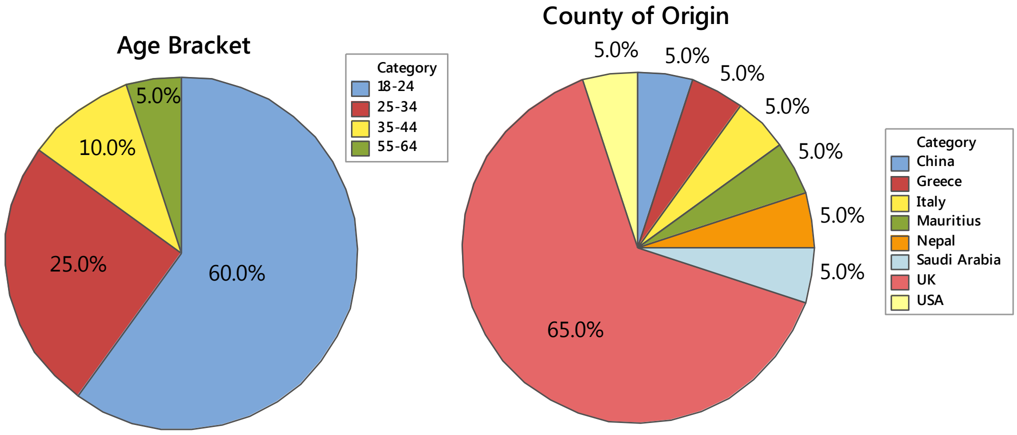

Figure 31: Gender and Handedness for Fingerprint Interaction

None of the participants identified themselves as having any accessibility problems, however one person identified as having a temporary illness, where the researchers noted that the subject's right index fingers had very slight burns.

Prior use of biometric systems was self-reported to understand participants and their previous experience with ABC systems further. 50% of participants stated that they had previously used an ABC system before. However, all twenty subjects expressed that they were aware of the process at airports. Of the 50% who had used an ABC system before, 100% previously used an ABC system which uses facial images. None of the subjects had previously interacted with a border control system which implements fingerprint verification.

Fifteen subjects stated they had previously used an electronic fingerprint system before and of those five had received a form of training before using the system. All five subjects said that this was for immigration purposes.

One participant out of the ten that had previously used an ABC system before identified that they received some form of training before using an ABC system for the first time (a UK eGate system) by interacting with a border guard officer.

Figure 32 below details the frequency of the height of participants for the study.

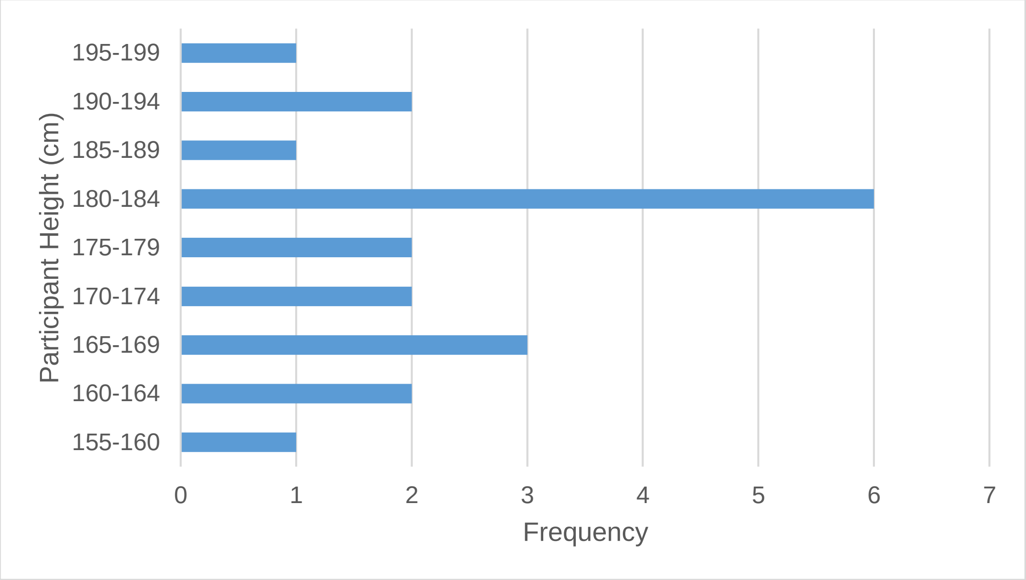

Figure 32: Height Distribution for Participants for Fingerprint Interaction

Questions asked for the Task Evaluation questionnaire were based on a 1-5 Likert Scale with one being strongly disagreed, and five being strongly agreed. Users were asked eleven questions regarding their use of the system.

1. I was given enough information to complete the task
2. I found the process easy to complete
3. I completed the task without any difficulty
4. The information provided clearly described what to do during the process
5. It was clear when the process begun
6. The prompts given on the screen were clear
7. I was confused by the entire process
8. The order of the capture process was clear
9. It was clear when the process had been completed
10. This experiment will benefit me when I use ABC systems in airports

5.2.4 Scenarios

Each participant was required to attempt fingerprint recognition for all five sensors, on both the slope and non-slope setting, for a total of ten interactions. Users were allocated a random order to complete the scenarios. This was done to minimise the effect of order in performance as well as in user's habituation. Each Scenario represented a fingerprint sensor and slope setting; Table 36 below demonstrates the possible combinations.

Table 36: Fingerprint Device and Slope Setting Combinations for Fingerprint Interaction

Scenario	Fingerprint Device	Slope Setting
1	UPEK Swipe	No Slope
2	UPEK Touch	No Slope
3	Hamster IV	No Slope
4	Hamster Pro 20	No Slope
5	Cross Match Guardian LC	No Slope
6	UPEK Swipe	Slope
7	UPEK Touch	Slope
8	Hamster IV	Slope
9	Hamster Pro 20	Slope
10	Cross Match Guardian LC	Slope

5.2.5 Guidance and Training

Participants were asked to read the Participant Information Sheet (PIS), which explained the purpose of the study also, to explain confidentiality issues such as where their fingerprints were stored and how long each sample would be saved for. Upon the completion of the consent form paperwork, the researcher started the video and Kinect recording.

Participants were asked to complete three surveys during the trial. One survey collected demographic and information to compile a personal profile, while the other surveys were for Task Evaluation and ABC Knowledge. Task Evaluation is discussed in Section 5.3.5 and results from the ABC Knowledge questionnaire are reviewed in Chapter 7.

Each participant was asked to follow instructions on the monitor without any further information from the researcher. Users were instructed that they might only ask questions if they were completely stuck and unsure how to proceed.

The researcher communicated which sensor the participation should interact with before each trial. Each sensor was labelled 1-5. Instructions prompted the user to interact with a fingerprint sensor. After a result was shown on the monitor, the process was repeated but for a different sensor for ten times. The researcher maintained a record log of any assistance that was required throughout the experiment.

5.2.6 Data Storage

At the end of each trial, the image of the fingerprint was saved to a secure, local drive and the timestamp of when each form was displayed was saved to a log file in a text document. Each fingerprint was given a unique identifier to keep anonymity.

The database was of a sensitive and personal nature. Hence it was stored on a secure server where access to the database was limited to the investigator. The size of the database after the images and video footage was deleted was around 400mb.

5.3 Performance Analysis

Performance analysis will be split into two categories; Kinect and HBSI Assessment. The main objective of the performance analysis is to identify the impact of the various variables introduced into this study affect on the interaction with the system and to determine how well the system performs compared to analysis from other similar implementations.

An investigation of the Kinect data will serve to ensure that data collected throughout the data collection is an accurate representation of the movement and to ensure the data captured can be used for behavioural analysis. The Kinect Analysis section reports on the tracking states, stability and the critical movements that were tracked throughout the experiment. The application and analysis of this data, if viable, will seek to enable the successful use of skeletal tracking within a range of biometric systems.

The HBSI assessment section details presentations and an evaluation of the system including reporting on the sample quality, usability and ergonomic factors. The presentation framework outlines correct and incorrect interactions throughout the study, investigating where errors may lie. In some scenarios, results are compared to previous studies conducted by NIST.

Performance Analysis ultimately investigates four main behavioural led tasks, based on the Behavioural Framework defined in Chapter 3. Table 37 below details each task, a definition and the expected behaviour.

Table 37: Task Analysis for Fingerprint Interaction

Task	Definition	Expected Behaviour
1	Start of the transaction (entry) – user moves towards feet symbols on the floor and stands within the designated area	User stands on feet symbols, looks at screen awaiting further instruction
2	Information is displayed on the screen requesting the user to place their finger on the sensor (biometric presentation)	User moves right arm forward, placing right index finger on sensor
3	Information on the screen confirms successful capture, processes and displays result (biometric read)	User moves right arm away from the sensor
4	Systems display information to confirm trial is over and to move forward – end of the transaction (exit)	User moves way from feet symbols back to the starting area

As there was not a token sensor present in this task, Task 2 and 3 seek to monitor the actions of the human-biometric interaction and the response to the system capture process. The four tasks will be subject to analysis throughout this chapter.

5.3.1 Kinect Analysis

Following from Chapter 4, Kinect Analysis, the concept of tracking states and performance analysis have previously been defined. Results from this chapter have proven the Kinect's ability to accurately track a single user within a self-service environment when the sensor is placed 1.3m away from the user. For this data collection, critical and associated tracking joints are defined for the behaviours tracked throughout the four tasks. Furthermore, tracking states and accuracy tests are reported and compared to the results obtained in Chapter 4. The ability to track specific movements successfully within a self-service environment will change the way we capture metrics, improve data collection scenarios and enable higher levels of quality feedback.

5.3.1.1 Definitions

To achieve a Successfully Processed Sample, the ideal HBSI result, the four critical tasks are required to be completed which enable the system to classify the presentation correctly. In summary; the user must move towards the feet symbols, place his or her finger on the device, remove the finger, and upon instruction, leave the designated area. Table 38 below details the definition of tracking for each task, identifying critical and associated joints that are tracked and analysed in real time.

The required 3D space coordinates are also stated. Successful completion of all four tasks should lead to a Successfully Processed Sample, having an overall positive impact on the quality and matching score.

Table 38: Kinect Task Definitions for Fingerprint Interaction

Task	Definition	3D Space	Critical	Associated
1	The user should be standing still on feet symbols	Z	Hip Left, Hip Right	Spine, Shoulders, Neck and Head
2	User should move their right arm towards the sensor, placing their finger on the device	X, Z	Right Shoulder, Elbow, Wrist, Hand	Shoulders, Neck, Head
3	User should move right arm away from sensor	X, Z	Right Shoulder, Elbow, Wrist, Hand	Shoulders, Neck, Head
4	User should leave the designated area	Z	Hip Left, Hip Right	Spine, Shoulders, Neck and Head

It is imperative that the critical joints for each task be considered fully tracked by the Kinect Sensor to ensure a high level of confidence that the data is an accurate representation of the user within the scenario. As discussed in Chapter 4, the sensor will only mark a joint as tracked if the sensor can fully establish the joints location based on the devices mapping algorithm. Critical and associated joints for this data capture will require the specified joints to be fully tracked so that the mapping of the skeleton is a highly accurate as possible.

5.3.1.2 Tracking States

Before analysing the Kinect data in detail, the first step was to ensure that only the fully tracked joint data was used. The main issue described in this data collection was the inclusion of the pedestal, which would remove the Kinect's ability to track joints in the lower body. To ensure that only fully tracked joints are used, the next process involved removing inferred and non-tracked information. Table 39 below describes the overall data tracked for all joints per scenario.

Inferred tracking information was devised from the Kinect sensor guessing the position of skeletal points that were out of range of the camera such as the ankle and knee. The system could infer these positions based on the available data from the hip joints and above. However as previously described from Section 4 an inferred joint is not an accurate presentation and is an estimate given by the sensor.

Table 39: Overview of Tracking States for all users in Fingerprint Interaction

Scenario	% Tracked	% Inferred	% Not Tracked
1	82.13%	17.85%	0.02%
2	81.52%	18.48%	0.00%
3	83.83%	16.17%	0.00%
4	83.49%	16.51%	0.00%
5	80.61%	19.37%	0.02%
6	84.68%	15.32%	0.00%
7	84.18%	15.82%	0.00%
8	84.08%	15.92%	0.00%
9	84.26%	15.74%	0.00%
10	84.14%	15.86%	0.00%

To investigate tracking states further, Table 40 below describes the percentage of joints ‘fully’ tracked for all users on each task.

Table 40: Fully Tracked Joints per Task (Dark Grey highlights critical joints for a task, Light Grey Associated joints) for Fingerprint Interaction

Joint	Task 1	Task 2	Task 3	Task 4
Head	100%	100%	100%	100%
Neck	100%	100%	100%	100%
ShoulderLeft	100%	100%	100%	100%
ShoulderRight	100%	100%	100%	100%
ElbowLeft	100%	100%	98.74%	100%
ElbowRight	100%	100%	100%	100%
WristLeft	96.58%	77.11%	73.57%	96.25%
WristRight	100%	100%	100%	100%
SpineTop	100%	100%	100%	100%
SpineMid	100%	100%	100%	100%
SpineBase	88.89%	100%	98.77%	97.58%
HipLeft	100%	100%	100%	100%
HipRight	100%	100%	100%	100%
KneeLeft	34.11%	32.58%	30.33%	41.58%
KneeRight	33.25%	28.58%	31.13%	40.25%
HandLeft	74.58%	75.28%	71.58%	68.56%
HandRight	70.22%	100%	100%	64.68%
FootLeft	10.58%	5.58%	6.11%	13.21%
FootRight	11.25%	4.58%	5.58%	12.55%

As Table 40 indicates, all critical joints (dark grey) for each associated task achieved a 100% tracking state and therefore based on previous stability tests in Chapter 4, were considered as usable for the following analysis.

5.3.1.3 Critical Tasks

The Behavioural Framework, as described in Section 3.5, outlines four tasks involved in the interaction process at a level where identifiable behaviour can be observed and defined in tracking by the Kinect. Through specifying a task, an analysis of the behaviours performed will provide insight into the movements users make through interaction, and how these may impact the performance of a system.

Based on both preliminary testing on a range of subjects and through Kinect stability testing in Section 4, the Kinect system was configured to categorise different behavioural codes for certain skeletal joint locations at the end of the four tasks. Location points were defined based on the experimental setup procedure. For example, by physical measurement, the user is standing on the feet symbols when they are standing between 1.15 to 1.40m away from the sensor. The sensor was aligned centrally to the symbols on a level with the user at 1.75m off the ground.

A definition was recorded by the Kinect HBSI Automation program to the log file depending on the joint location at the end of each task. For example, for Task 1, if the Hip Left and Hip Right joint met the requirements of >=1.15m && <=1.40m for the Z coordinate measurement, a GFEET01 metric was recorded. GFEET01 is defined as when the user is standing on the provided feet symbols, based on the physical measurements discussed in Section 5.2. Likewise, if there was no movement recorded (Hip Left and Hip Right are >=2.4m, the starting position), then a RFEET04 has been registered to the log file. Each task described below details the behaviour and the requirements for a Kinect definition.

Task 1: Entry

Participants were required to move and stand on feet symbols at the start of each transaction. Symbols were directly placed in front of the kiosk centrally to the Kinect sensor. Instructions based on ISO standards were displayed on the monitor informing a subject to move towards the system.

The feet symbols, which are traditionally used in ABC systems, have many advantages in this data collection such as limiting the range of the captured data and identifying a region of interest for where users should be standing for interactions to begin. The main purpose for feet symbols, however, serves as a standpoint for face capture in operational scenarios, where standing in the required area will likely have an impact on accurate face capture.

Table 41 below details the behaviour recorded and the Kinect definition for Task 1 entry.

Table 41: Behavioural Framework for Task 1: Entry (Fingerprint Interaction)

Code	Behaviour	Kinect Definition
GFEET01	The user approaches feet symbols and aligns feet to stand on the symbols correctly	HL & HR Z <= 1.40m && >= 1.15m
GFEET02	The user approaches feet symbols and feet are very slightly (2-5cm) off the feet symbols	HL & HR Z <= 1.45m && >= 1.10m
OFEET01	User is slightly off centre but standing on the symbols	HL & HR X>= ±>.01m && Z <=1.40m && >= 1.10m
RFEET01	User is stood in front of the feet symbols	HL & HR Z<= 1.15m
RFEET02	User is stood behind the feet symbols	HL & HR Z>= 1.45m
RFEET03	The user has not moved to feet symbols and is standing still	HL & HR Z >= 2.4m && Minimal variance

Kinect definitions were determined by the placement of feet symbols described in the experimental room layout in Section 5.2.3. As any joints below the knee were not fully tracked, behaviours were based on Hip Left and Hip Right joints (on the assumption that these are directly above the feet and therefore have the same depth distance from the sensor). Larger variances such as GFEET02 allow users to be on symbols but not necessarily within the full area.

Although it is realised that these behaviours will have a larger impact on face recognition (which was not captured in this study), it is necessary to identify if the Kinect sensor can capture the movement of the user before initiating the next task. Identifying these behaviours in real time would be beneficial for real-life application. If a RFEET03 behaviour is detected for example, and categorised after the allotted time, the system could attempt to alert the user or to change perhaps the method it communicates (e.g. icon instructions to text/audio). It would also enable the system to categorise the error in such a way that it would indicate a user fault, not a system generated an error.

Table 42: Observed and Tracked Behaviours for Task 1: Entry for Fingerprint Interaction

Task 1 (Entry)				
	Observed		**Kinect**	
Behaviour	N	%	N	%
GFEET01	157	77.7	161	79.7
GFEET02	44	21.8	41	20.3
RFEET01	1	0.5	0	0

Table 42 above presents the observed and Kinect measured Task 1 entry behaviours for the Fingerprint Interaction. 95.5% of the interactions resulted in ideal behaviours, where participants followed instructions and made correct use of the feet symbols. The Kinect categorised that 100% of the interactions were positioned correctly on the feet symbols, incorrectly identifying that one interaction was slightly too far forward. The significance of the RFEET01 behaviour has implications for biometric interaction. If this behaviour was repeated within a facial biometric modality system, it is likely that the user would be too close to the camera and therefore would not meet the requirements of the capture process.

It could be a possibility that the fingerprint scanners were placed in such a location that the user may have had to adjust themselves before interaction. A larger frequency of RFEET01 classifications could reveal insights into the design and build of a system, maybe resulting in an adjustment of the placement of symbols.

Task 2: Biometric Read

Previous work (as discussed in Section 4.2) with the Kinect V1 device has demonstrated it is unsuitable for determining finer activities such as finger movements. Kinect V2 is also unable to distinguish individual fingers and can only refer to Hand Tip Joint as an extension of the hand; through identifying either one or all five fingers. Therefore Task 2 for Fingerprint Interaction focuses on hand and wrist placement, with the assumption that users are making a correct presentation with their finger.

To begin this task, the user must be standing within the feet symbols (GFEET01, GFEET02 or OFEET01) which is a pre-requirement from Task 1. Any movement was recorded, for instance, if the user took a step back. However, all participants remained within the required area. Table 43 below details the different behaviours for the data collection and the Kinect’s tracking requirements for those movements.

Table 43: Behavioural Framework for Task 2: Biometric Read for Fingerprint Interaction

Code	Behaviour	Kinect Definition
GFING01	The user places finger on the sensor, wrist and hand joints are inside detection zone	RightWrist and RightHand Joints Z <= 1.0m
RFING01	Vigorously adjusts finger	High Variance in X or Y for Wrist/Hand Joints (>0.02m every 30frames)
RFING02	User presents incorrect finger	Absence of Right Wrist/Hand. Left Wrist/Hand movement detected instead
RFING03	Not swiping (Swipe sensor only)	No movement detected in X or Z for Wrist/Hand Joints

Table 44 below demonstrates the observed and tracked behaviours for Task 2; Biometric Read. The results indicate that the Kinect can track full movements of the user's right arm for this self-service scenario, but without the finesse of tracking an individual finger. The Kinect was unable to detect finer movements such as adjusting the finger or removing the finger too early. A setback for this system design was from only tracking the wrist and hand joints for interaction purposes. If the wrist and hand were stationary and within the required area, there was no following methodology to track the movement of the finger (e.g. the finger may be raised above the sensor at an angle, or only partially on the sensor). From ground truth data, this did not occur during this data collection, but it is possible that this undesired movement may take place in possible future implementations.

Table 44: Observed and Tracked Behaviours for Task 2: Biometric Read in Fingerprint Interaction

Task 2 (Finger)				
	Observed		**Kinect**	
Behaviour	N	%	N	%
GFING01	188	94.5	189	95.0
RFING01	1	0.5	0	0.0
RFING02	4	2.0	4	2.0
RFING03	6	2.0	6	3.0

The Kinect could distinguish all the correct presentations made to the sensor. The individual occurrence of RFING02 was observed both during the data collection and through video footage. The HBSI automation program classified the movement as a correct interaction. The sensor, in this case, was unable to detect the dynamic movement as observed in the video footage. The presentation was still categorised as a ‘correct’ presentation as data shows the hand and wrist joints were within the required region.

Task 3: Biometric Accept

This task proceeds Biometric Read, and for most systems, information is displayed to inform the user that their biometric sample has either been accepted or rejected after processing sub-system elements have been completed. The desired action for this task would be for the user to remove their sample from the sensor and to proceed to the next stage, Exit. There should be a very short delay between Tasks 2, 3 and 4.

Although this task signifies the end of the interaction, it is an important stage to assess when attempting to understand the impact of the information to the user.

For this system and step, the classification is determined by the user who should remove their hand from the sensor – which is an indication that they have perceived the information as confirming the biometric sample has been read and processed. The implementation can either display the ‘captured’ or ‘rejected’ form – in both cases; these are simple pictorial images usually depicting a tick or a cross. Table 45 below details the behaviour and Kinect measurements.

Table 45: Behavioural Framework for Task 3: Biometric Accept

Code	Observed Behaviour	Kinect
GHAND01	User removes hand from the sensor	Right Wrist and Hand Joints Z > 1.0m
FHAND02	User does not remove hand	Right Wrist and Hand Joints Z <= 1.0m

Table 46 below details observed and tracked behaviours for all scenarios in this data collection. The Kinect sensor correctly classified all hand movements for this task. The sensor could correctly detect that the hand had been removed from the sensor.

Table 46: Observed and Tracked Behaviours for Task 3: Biometric Accept for Fingerprint Interaction

Task 3 (Face)				
	Observed		Kinect	
Behaviour	N	%	N	%
GHAND01	195	100	195	100
FHAND02	0	0	0	0

Task 4: Exit

This task occurred after the interaction process (Task 2-3) had been completed and information had been displayed to the user relaying successful or unsuccessful capture. In either case, the user is expected to leave the station and return to the starting position. Table 47 details the behaviours below.

Table 47: Behavioural Framework for Task 4: Exit for Fingerprint Interaction

	Observed Behaviour	Kinect
GLE01	The user leaves the station	Hip Joints >1.45m
RLE01	The user does not leave the station	Hip Joints <=1.45m

Table 48 details the observed and tracked behaviours for Task 4: Exit. Again, all behaviours were classified correctly for this short study.

Table 48: Observed and Tracked Behaviours for Task 4: Biometric Accept for Fingerprint Interaction

Task 4 (Exit)				
	Observed		Kinect	
Behaviour	N	%	N	%
GLE01	200	99.0	200	99.0
REL01	2	1.0	2	1.0

The sensor could correctly detect the behaviours as confirmed by manual observations.

5.3.2 HBSI Assessment

The HBSI model can be used to evaluate both system performance and individual transactions. This section uses the Kinect V2 device to analyse both the presentation framework and evaluation method in a real-time scenario, analysing movements to identify errors within a transaction.

5.3.2.1 Presentation Framework

The application of the skeletal tracking system with the Kinect device enabled the first steps into HBSI automation, the ability to categorise presentations in real time based on body movements and positions. Figure 33 below details the adapted HBSI Presentation Framework, using the Kinect sensor to identify correct or incorrect interactions.

In addition to the main functions of the Neurotechnology MegaMatcher SDK, logical assessments were added to enable the automated categorisation of HBSI presentations.

The HBSI automation program was configured to assess movements for each of the four tasks as previously described in Section 5.3.1. Each task was specified certain conditions that the Kinect was searching for, e.g. the position of a hip or wrist joint within certain fixed parameters of the Kinect device. The specification for these limits was based on the knowledge that joint locations should be within certain regions at certain points throughout the transaction. For example, for Task 1, users should be within a certain distance that is the fixed location of the feet symbols on the floor. See Section 5.2.1 for the room setup and distances the sensor and the kiosk and feet symbols.

The main 3D coordinate that was tracked in this study was the Z distance, the distance of the joint from the sensor. As this was preliminary work with the Kinect to investigate the capability of the device, other coordinates were only used when looking for variances in data (e.g. no movement on the X or Y plane).

Table 49 demonstrates the scores for the ‘good’ behaviours as previously described in Section 5.3.1. When a good behaviour was met, a ‘score’ was given based on successful completion of the task.

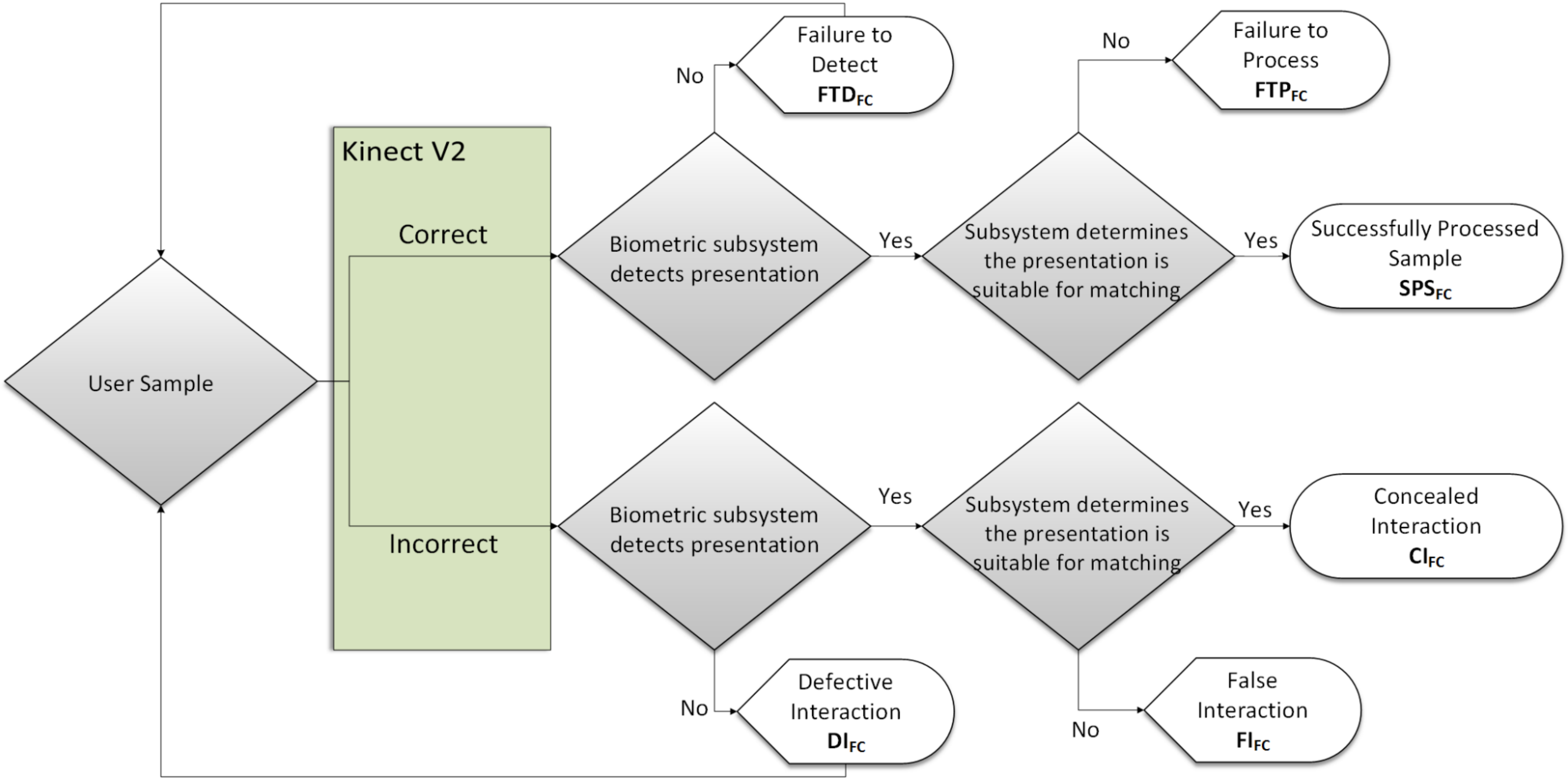

Figure 33: HBSI with the Kinect V2

Table 49: Kinect assessment limits and associated score for Fingerprint Interaction

Task	Control Limit	Score
1	Z-Distance for Hip Left/Right < 1.45m & >1.10m	1
2	Z-Distance Wrist is < .01m away from sensor X-Variance Wrist is <.0005m for swipe sensor	3
3	Z-Distance Wrist is > .01m away from sensor X-Variance Wrist is <.0005m for swipe sensor	1
4	Z-Distance for Hip Left/Right >1.35m	1

When a behaviour led task was documented by the Kinect, a counter incremented to a total of six points. One point was awarded for completing task one, three and four; three points have been granted for task 2. Task 2 was deemed the ‘critical’ task of the experiment and therefore rewarded a higher score. If the counter reached 4 points or above, then the system determined a ‘correct’ presentation, while a score of three or lower determine an ‘incorrect’ presentation.

To categorise a HBSI metric, the system required two key components; a score obtained from the behavioural tasks and either the presence or absence of a fingerprint sample. For this initial study, the researchers had to confirm the presence of a sample to the HBSI program by the end of the transaction. This was because the two systems (MegaMatcher and HBSI automation) were run on two separate computers. Therefore, while the process automated the classification process, the system was not fully automated due to a form of input required by the observer.

Table 50 below demonstrated the logical assessment required for a HBSI categorisation.

Table 50: HBSI Assessment for Fingerprint Interaction

HBSI	Behaviour	Completed Tasks	Processed Sample
False Interaction	Incorrect	Yes	Yes
Concealed Interaction	Incorrect	Yes	No
Defective Interaction	Incorrect	No	No
Successfully Processed Sample	Correct	Yes	Yes
Failure to Detect	Correct	Yes	No
Failure to Process	Correct	No	No

Table 51 below displays the frequency of categorisations, comparing automated versus manual HBSI categorisations.

Table 51: Frequency of HBSI Categorisations for Fingerprint Interaction

HBSI Categorisation	Automated	Manual
False Interaction	0	0
Concealed Interaction	10	4
Defective Interaction	4	7
Successfully Processed Sample	185	188
Failure to Acquire	3	3
Failure to Detect	0	0

All manual observations were recorded by the researcher during data collection and further verified by the camera footage post-data collection.

Of the total 202 presentations made to each sensor, 185 were classified as correct presentations and were successfully processed by the system.

All three Failure to Acquires occurred due to a system crash during sample acquisition. These were processed as correct presentations through a Kinect score and an absence of a processed sample.

The system categorised Fourteen incorrect presentations. Four of these incorrect interactions were categorised when a user presented their left finger instead of their right. In these cases, the presentations should be classified by the system as Concealed Interactions. The system determined six more CI's from a split of categorisations that through manual observation, should have been defined as DI's and SPS's.

There were four instances of users not swiping correctly for the first time when interacting with a swipe sensor. Three occurrences were due to no swiping movement and one instance of minimal movement during the reading process. These were classified correctly by the Kinect as a Defective Interactions through a lack of deviation in movement as set by the control limits.

There were recorded Failure to Detect (FTD) errors for this data collection.

5.3.2.2 Evaluation Framework

This next section reports on sample quality, user satisfaction, efficiency and effectiveness metrics as defined in the HBSI Evaluation Framework. Identifying and analysing usability in any scenario testing is critical to understand how and why people use a system or product.

5.3.2.2.1 Efficiency

The efficiency component of the HBSI Model reports the speed in which users can complete the tasks for which they use an ABC system.

In this section, following efficiency metrics are reported:

- Successful Task Completion
- Time on Task
- Number of Assists

A total of 202 interactions were recorded for Fingerprint Interaction. Of the 202, 195 were completed successfully.

Time on Task, for this experiment, was defined from the point in time when instructions were displayed asking the user to present their finger, to the end of the task where the fingerprint image had been captured and processed. For each response variable for time on task, the factor of angle and participant height against the sensor was considered. The timing data was not normally distributed therefore we used the non-parametric Kruskal-Wallis test. Table 52 below details significance for Time on Task.

Table 52: Significance of Variables on Time on Task ("+": p<0.05) for Fingerprint Interaction

Sensor	Sensor Angle	Participant Height	Mean Time
1	+	-	15.78s
2	-	-	12.35s
3	-	-	9.30s
4	-	-	8.90s
5	-	-	9.60s

The effect of angle was the only variable that had a significant effect on time on task for the swipe sensor only. The effect of subject height or angle on other sensors did not have an impact on time on task. However, it can be noted that the swipe sensor took longer to complete whereas non-swipe sensors have a lower average time on task.

Figure 34 below details the difference between the average Time on Task for sensors both on a flat and raised surface.

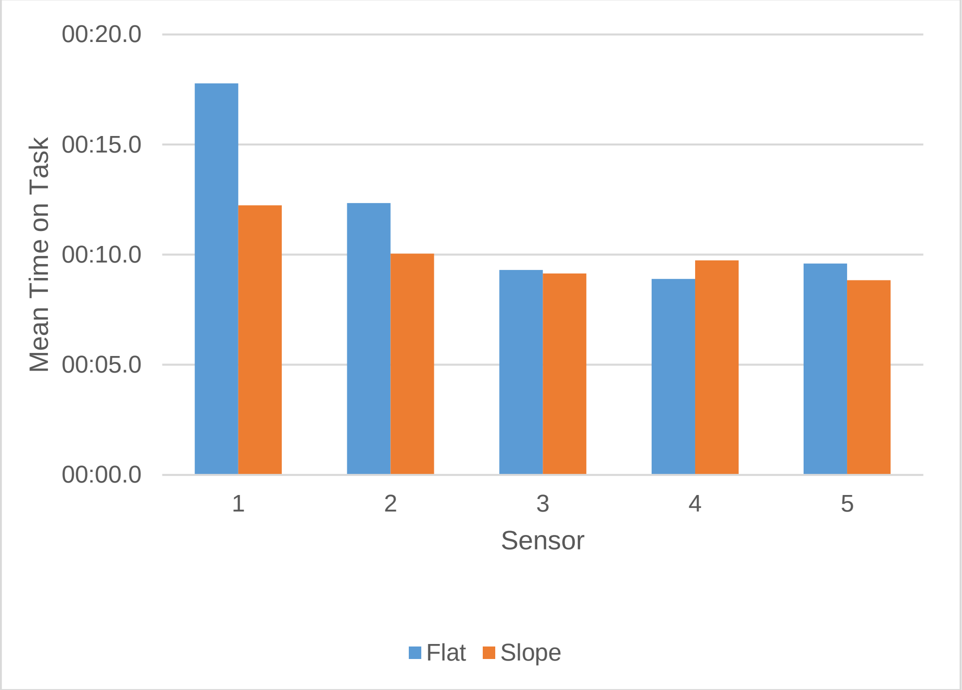

Figure 34: Mean Time on Task for Fingerprint Interaction

When comparing results to previous studies on different sensors and angles, the findings in this study show similarities. The Usability Testing of Height and Angles of Ten-Print Fingerprint Capture from NIST [55] found an impact on time on task when looking at the table height. The height of the table had a significant effect (p = 0.01) for a left thumb sample on time on task while participant height had an impact for right slap, left slap, and both thumbs for one (of two) sensor studied. In summary, the NIST study reported no significant effect due to angle, table height or subject height for the time required to complete a fingerprint task for both scanners except for the tasks described previously.

Fifteen interactions required assistance from the researcher. The researcher communicated each assist verbally. Of the 15 interactions, seven of the assists were verbal commands to swipe the finger for Sensor 1. Feedback from participants suggested that the information provided was unclear and left subjects confused initially. All subjects who received a verbal communication completed the task. The remaining eight assists were from a combination of prompting participants to walk to the sensor (4), remove finger (2) and step away from the system (2).

5.3.2.2.2 Effectiveness

In this section, following efficiency metrics are studied:

- Total time is taken to complete transaction
- Number of attempts

Each user was only allowed one attempt per scenario unless there was a system fault (freezing, crash) which occurred three times.

Time Taken to Complete (or Transaction Time) was defined between the point the Kinect started tracking the subject from the start position and to the point in time the final piece of information had been shown indicating that the fingerprint had been successfully processed. The time taken to exit after the last instruction was displayed was not included in this analysis. Time Taken to Complete is an important aspect of studying as it directly relates to throughput rates and the rate at which passengers can be processed through ABC systems. Average Time Taken to Complete is reported in Figure 35 below.

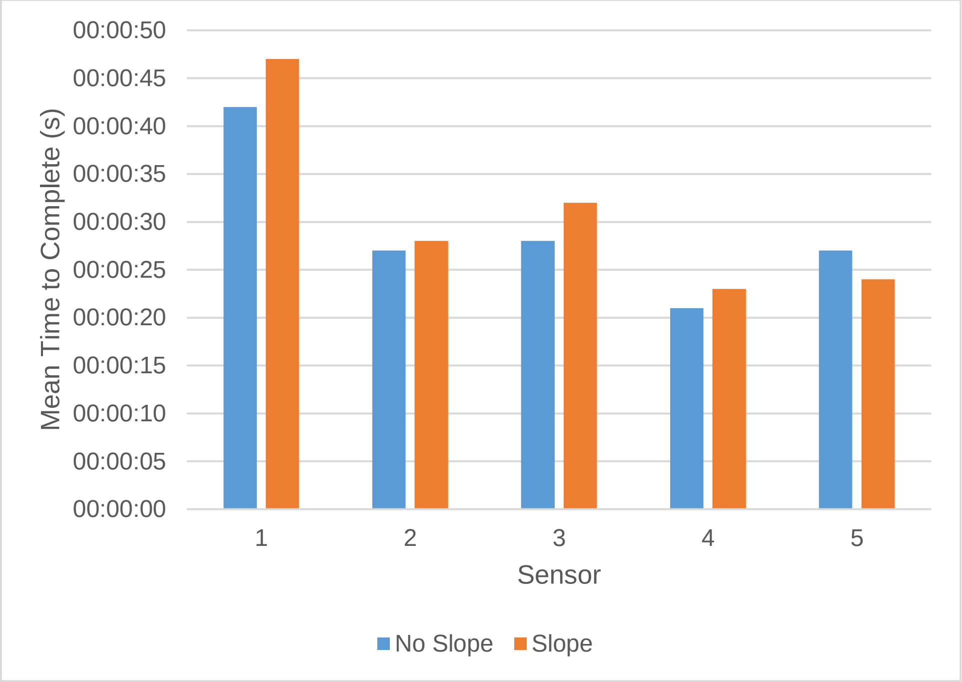

Figure 35: Mean Time to Complete for each sensor for Fingerprint Interaction

Figure 35 demonstrates the average time taken to complete the transaction for each sensor. Swipe sensors took on average much longer to complete than its non-swipe counterparts. Sloped positioned sensors 1 to 4 took slightly longer to complete than its non-sloped surfaces.

A Kruskal-Wallis test is used to see if there was a statistically significant difference for factors of sensor, angle and subject height on Time Taken to Complete. The distributions of NFIQ were statistically significant between groups, H (4) = 76.805 p = < .005. Table 53 identifies if the variables were statistically significant.

Table 53: Statistical Significance for Average Time to Complete ("+": p<0.05) for Fingerprint Interaction

Sensor	Sensor Angle	Participant Height	Median Time Taken to Complete
1	+	+	38.31s
2	+	-	24.32s
3	-	-	25.26s
4	-	-	20.32s
5	-	-	23.11s

Time Taken to Complete is typically a difficult factor to compare for system performance due to a nonstandard practice for reporting the measurement. However, most studies seem to follow the route of measuring the time from the start of the transaction until the last piece of information has been conveyed [97] [7].

A significant relationship was found for the angle of the slope for sensors one and two as well as the height of the participant for sensor one.

Although the data provides interesting results for the effect of the angle and participant height on Time Taken to Complete the interaction, these results may not necessarily outline errors in performance. In a standard performance report, throughput metrics will be declared, and feedback from users may be reported to distinguish if there were any reasons behind a higher transaction time.

5.3.2.2.3 Sample Quality

This part of the study reports on the quality of captured images and investigates which variable (angle, sensor, subject height) impacts the quality of the samples obtained. The analysis of sample quality is based on two factors; the NIST Fingerprint Image Quality (NFIQ) and the matching score against the enrolled image. The NFIQ quality rating was recorded for all fingerprint interactions during both the enrolment and verification components of the experiment.

Seven fingerprints were not captured successfully due to either system errors or erroneous user presentations as discussed in previously in Section 5.3.2 HBSI Assessment.

NFIQ ratings were totalled for the ten individual scenarios in Figure 36 below. As it can be observed from the chart, Sensor 1 (sloped and non-sloped surface) had a higher frequency of lower rated NFIQ samples. A higher frequency of the best quality samples was collected from Sensor 5 (Sloped and sloped surface).

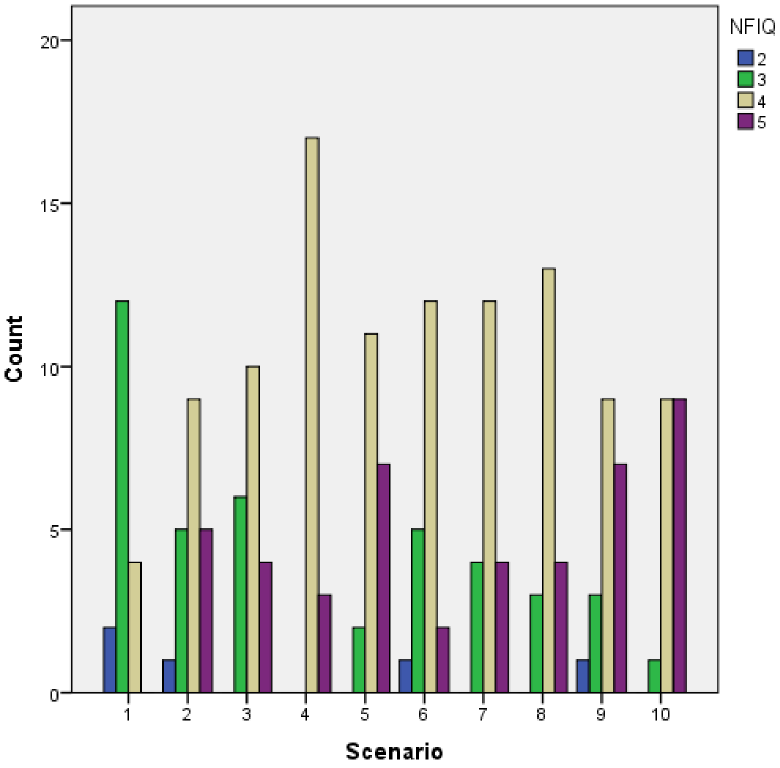

Figure 36: NFIQ Score in Fingerprint Interaction

To investigate the significance of the angle of the slope on NFIQ ratings for sensor pairs, a Sign test was conducted to study the difference between the slope vs non-slope sensors. Results are displayed in Table 54 below.

Table 54: Significance of Sensor Angle for NFIQ ("+": p<0.05) for Fingerprint Interaction

Sensor	Median	Significance
1	3	+
2	4	-
3	4	-
4	4	-
5	4	-

A significant result was found for the swipe sensor only when comparing NFIQ results on slope vs non-slope tasks. Interestingly, a lower mean NFIQ of 2.62 was found for the swipe sensor on a flat surface, compared to a 3.64 NFIQ when captured on a slope. There were no other statistically significance between NFIQ for an angle on other sensor pairs. Figure 37 below compares mean NFIQ between sensors.

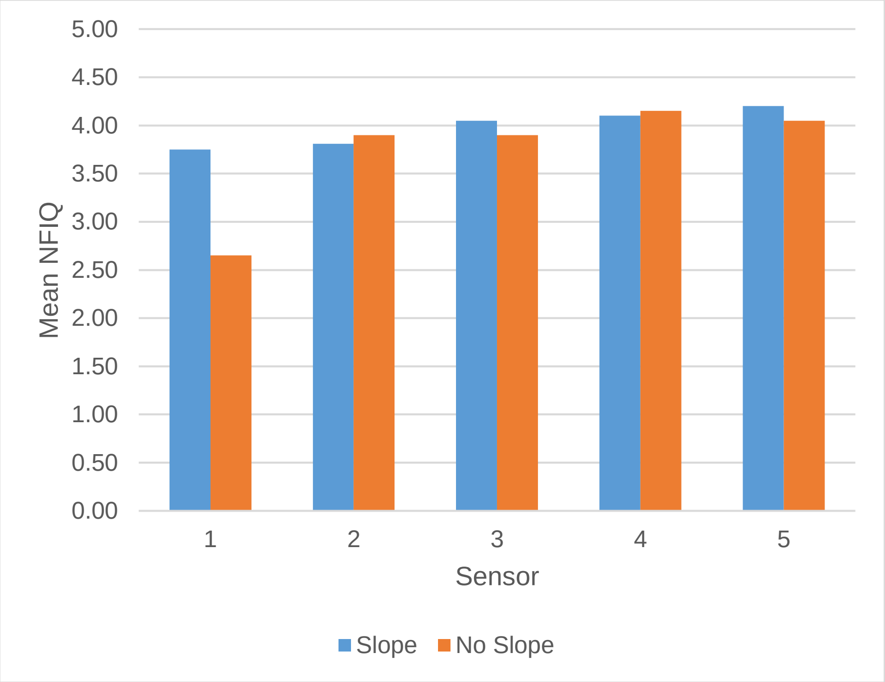

Figure 37: Mean NFIQ between Slope Settings for Fingerprint Interaction

To investigate sample quality further, a Kruskal-Wallis test is used to see if there was a statistically significant difference for factors of the sensor, angle height and subject height on NFIQ. The distributions of NFIQ were statistically significant between groups, H (9) = 40.479, p = < 0.005. Table 55 identifies if the variables were significant.

Table 55: Statistical Significance for Angle and Participant on NFIQ ("+": p<0.05) for Fingerprint Interaction

Sensor	Angle	Participant Height	Median NFIQ
1	+	-	3
2	-	-	4
3	-	-	4
4	-	-	4
5	-	-	4

The results demonstrate similarities with the NIST studies on fingerprint interaction. Although one relationship was found between the two possible angles for the swipe sensor, there were no other significant relationships between sensor pairs for both participant height and angle of the surface. The study on usability testing of height and angles of a fingerprint captured by two sensors from the NIST group [55] also demonstrate no significant results for sensor pairs on different angled surfaces. No other studies have investigated an effect on NFIQ for swipe sensorsc at an angle so results can only be reasonably compared to NIST's work. Swipe sensors, however, are not used in ABC scenarios, and therefore the outcome of the research may have little impact on this area.

Determining the matching score for captured images is a crucial element to ABC scenarios as samples must meet a quality threshold value to enable traveller verification for border crossing. In this study, the captured image in the verification element of the task is compared against the enrolled index finger sample that matched ISO token standards. The average Matching Scores for all sensors are reported in Figure 38 below.

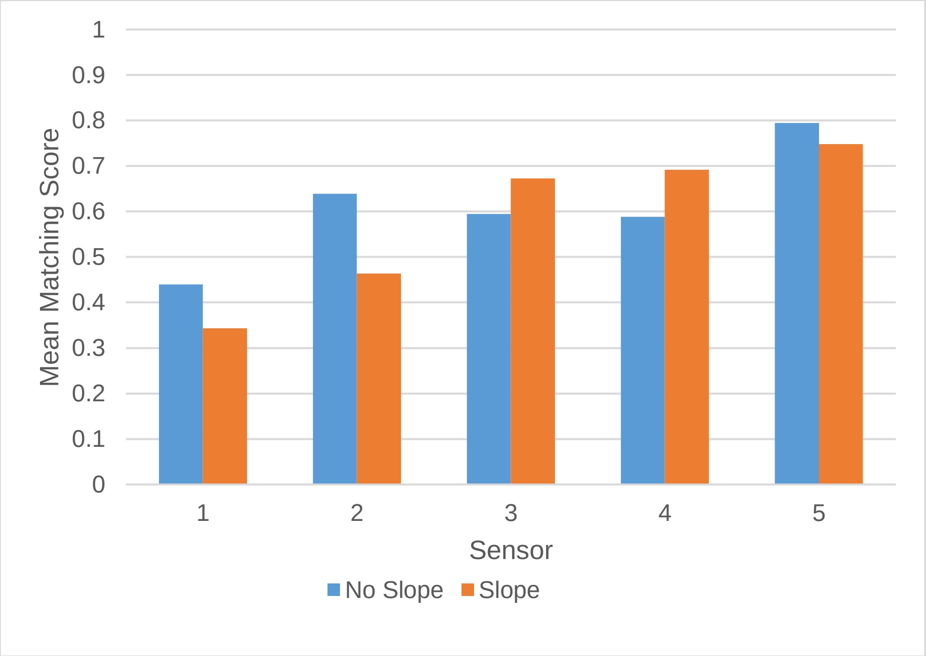

Figure 38: Mean Matching Score for Slope Settings in Fingerprint Interaction

A Kruskal-Wallis test is used to see if there was a statistically significant difference between factors of the sensor, angle height and subject height on Matching Score. Table 56 identifies if there were significances between the variables described.

Table 56: Statistical Significance for Angle and Participant Height on Matching Score ("+": p<0.05) for Fingerprint Interaction

Sensor	Angle	Participant Height	Median Matching Score
1	+	-	0.45
2	-	-	0.53
3	-	-	0.65
4	-	-	0.66
5	-	-	0.75

Table 56 reveals one statistically significant relationship for matching score for sensor one between sloped and non-sloped variables, as did Table 54 for NFIQ samples. Matching scores did improve for each sensor but this not due to either angle or participant height. No other studies are looking at matching score against usability aspects of a system (e.g. sensor placement, types of information, participant variables).

If variables are having an impact on throughput or other areas, then analysing tracked skeletal data could highlight why these factors may be having a significant effect on the performance.

Section 5.4 Data Analysis begins to investigate these relationships further with these preliminary findings.

5.3.2.2.4 User Satisfaction

A task evaluation was conducted to assess user satisfaction with the system. The main goal was to assess wherever the user perceived the system to be efficient. Answers were collected after all ten trials and so Figure 39 details the response to several questions directly relating to the performance of the system. Further results are discussed in Chapter 7, Recommendations.

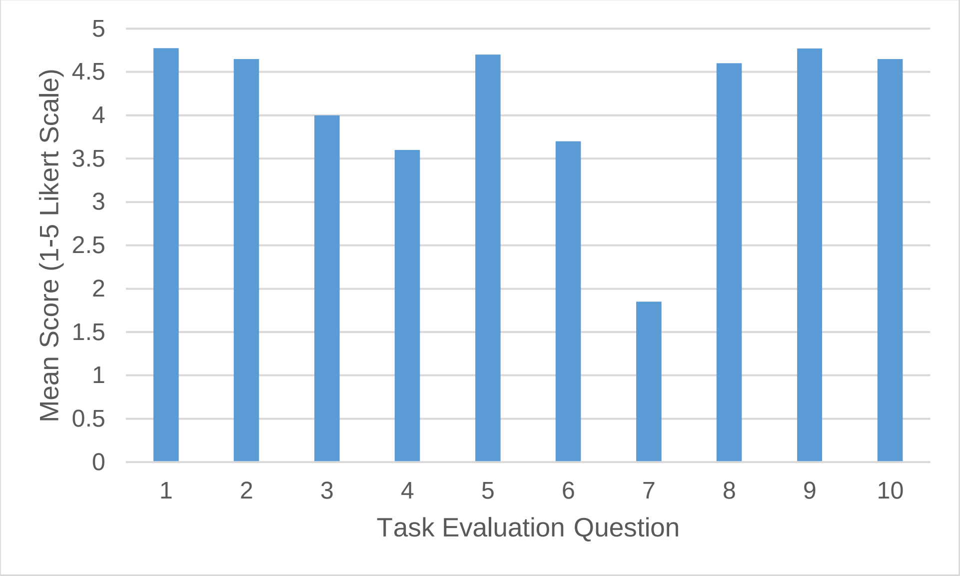

Figure 39: Mean Score for Task Evaluation Questions for Fingerprint Interaction

Users indicated they were given enough information to complete the task (Q1) and found the process relatively easy to complete (Q2). Information provided to the user was seeming clear (Q3-6). Users also indicated that the process was not entirely confusing (Q7) and that capture process was clear (Q8) and very clear the process had completed (Q9). Users revealed that they were likely to use an ABC system in the future based on their experience with the system in this study (Q10).

Although User Satisfaction is an important aspect to understand in any system, the feedback detailed for Fingerprint Interaction is somewhat flawed, the user concluded the metrics based on their entire process. Therefore, their answers may be skewed based on their whole experience rather than their interaction with an individual sensor.

5.4 Data Analysis

This section differs from performance analysis in respect to seeking answers to defined research questions. The main objective of this part of the research was to compare positional and movement data from the Kinect against the impact on performance. However, the first two questions seek to answer some unexplored relationships in the literature.

Question 1

The hypothesis was that users who used ABC systems before provided better fingerprints samples regardless of previous fingerprint usage. The theory was that previous users had a better understanding of information and the process involved. However, because the sample size for this collection was relatively small compared to large-scale scenario testing, a null hypothesis was assumed. A chi-squared test was conducted to reveal wherever there was any significance between the two groups (Previous ABC users and non-ABC users). Figure 40 demonstrates the difference between the groups for the total number of NFIQ samples.

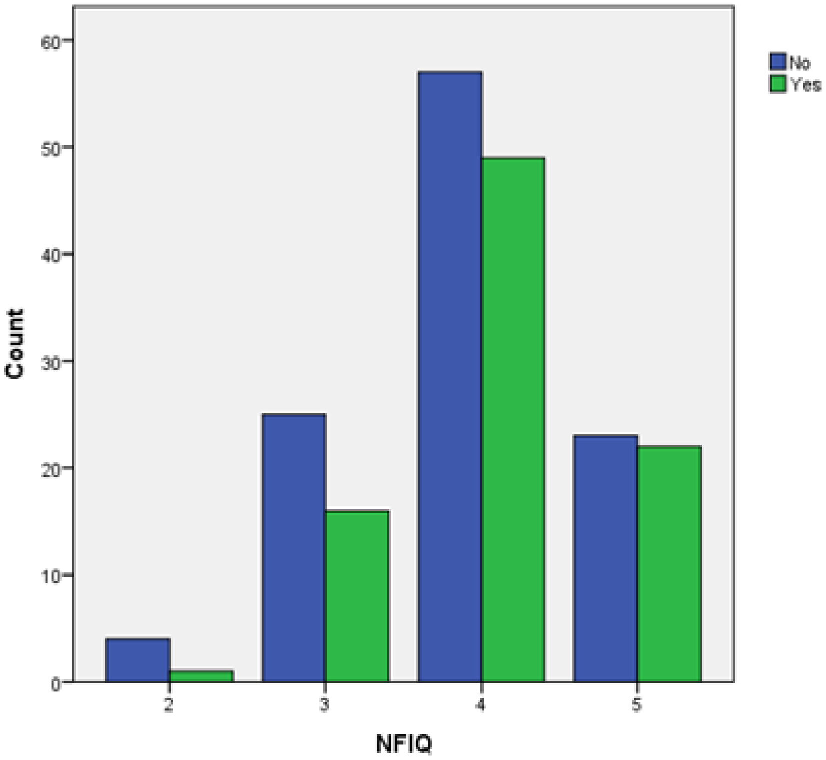

Figure 40: A comparison of NFIQ scores between users who previously used any form of ABC System Before (Fingerprint Interaction)

A chi-square statistical test was performed to examine the relationship between previous ABC users and NFIQ quality score. There was no statistical significance found, X^2 (3, N = 197) = 2.18, p =.53. As stated previously, zero subjects had previously used an ABC system with a fingerprint modality before the experiment but had used systems which used facial recognition.

Question 2

The next research question investigated the relationship between the user and sample quality and stated that people who used electronic fingerprint capture devices before provided higher quality fingerprint samples. Figure 41 below demonstrates the relationship between the two groups.

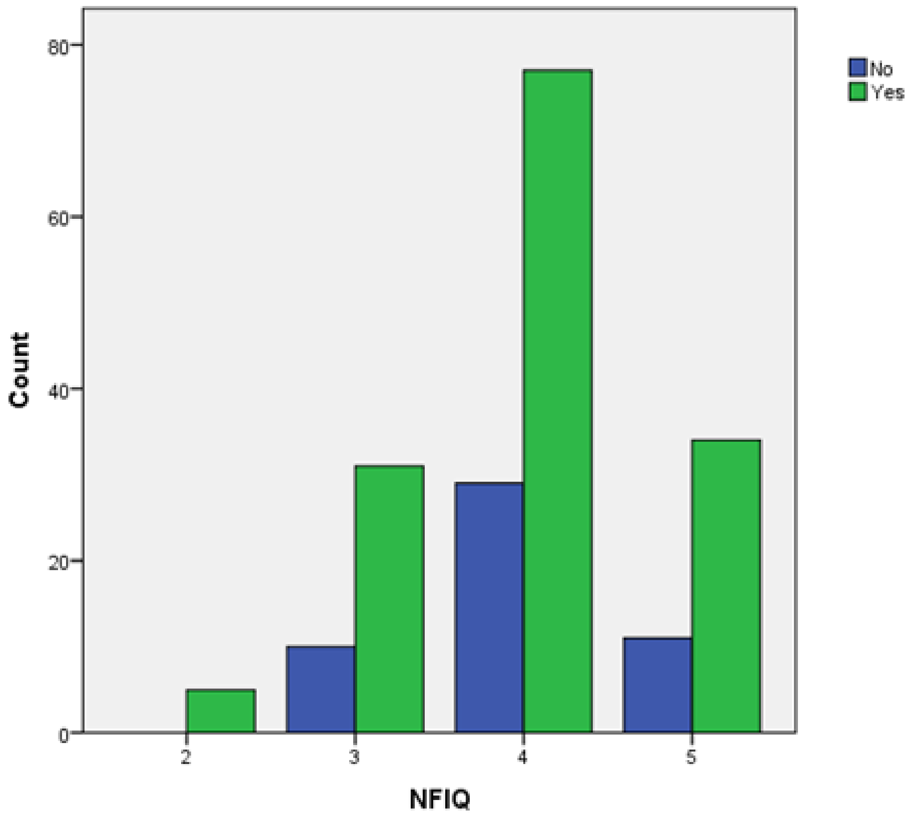

Figure 41: A Comparison of NFIQ scores between users who have submitted their fingerprints to an electronic system before (Fingerprint Interaction)

A chi-square statistical test was performed, and no relationship was found between previous electronic fingerprint system users and the NFIQ quality score, X^2 (3, N = 197) = 1.962, p =.580. Only five subjects had no previous experience using an electronic fingerprint capture system before the experiment.

Question 3

In this section, the relationship between performance and natural movement during an interaction is studied. Specifically reporting on the impact of wrist movement during fingerprint interaction. It is expected that there are naturally tiny variances recorded for wrist and hand joints during fingerprint interaction due to movement of the finger. However, there should only be any substantial movements in data for the swiping movement for the swipe sensor (Sensor 1). Skeletal data for the Right Wrist Joints were given the fully tracked status during Tasks 2 and 3, and so data captured throughout the study obtained a high degree of accuracy. However, as with all tracking systems, there will always remain a minimal variance in the data collected. Therefore, the analysis will be based on the average variance for all users in both X and Z axis data for the Right Wrist joint over the course of Task 2.

The ability to capture this data within testing scenarios will reveal potential usability issues with the system before deployment. For example, does the placement of sensors provide discomfort or awkwardness in interaction? It could require a user to reposition them for a better sample. If this is the case, is this measurable? From user feedback, there was no reported awkwardness of using any of the sensors.

In addition to the benefit of analysing these movements through testing before deployment, the ability to detect larger than average (or spike) variances in movement could also enable enhanced information to be relayed in an attempt to control the unwanted action (e.g. icons/instructions to keep your hand/finger still).

To compare this relationship, the first step is to investigate the variance in movement captured for the Right Wrist Joint during physical interaction (Task 2: Biometric Capture). This is the period where the arm is within the detection zone, and the participant has placed their finger on the sensor for reading. Figures 42 and 43 below report on the average variance for both X and Z 3D space coordinates (measured in m) over the period participants were interacting with a fingerprint sensor.

As results indicate, the average variance for both X and Z coordinates for the Right Wrist Joint was mostly consistent throughout each scenario except for sensor three on a non-slope platform. For variance in the X coordinate, it would be expected that there is very little to no movements during biometric reading. There was an expected variance for Sensor 1 for both the slope and flat surface; however, there was a higher variance for the X coordinate for the non-sloped version, suggesting users moved their wrist joint more vigorously during the interaction.

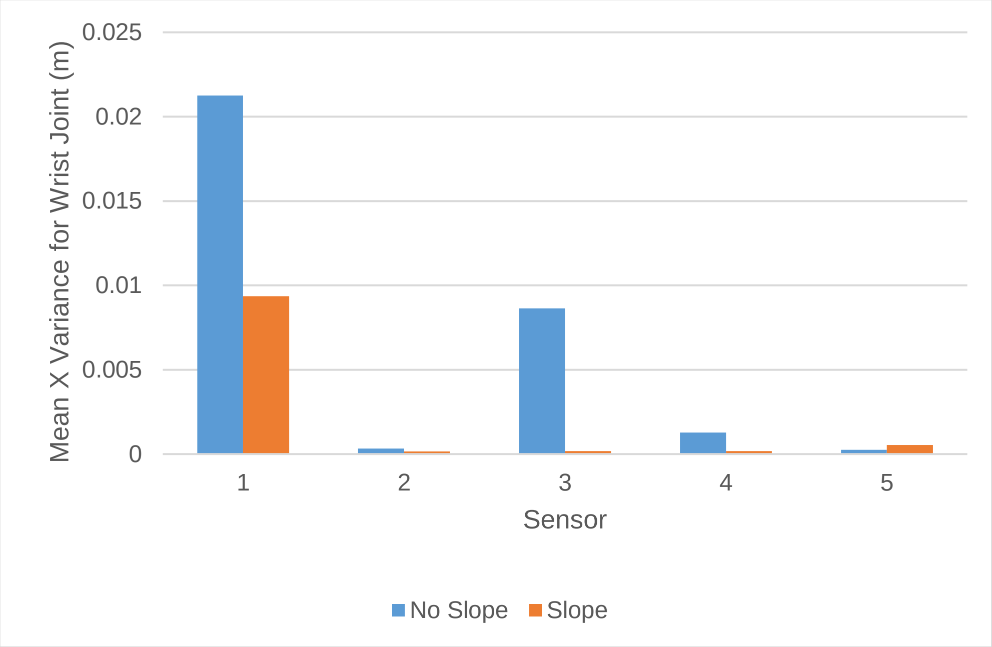

Figure 42: Mean X Variance for Task 2 Right Wrist Joint for Fingerprint Interaction

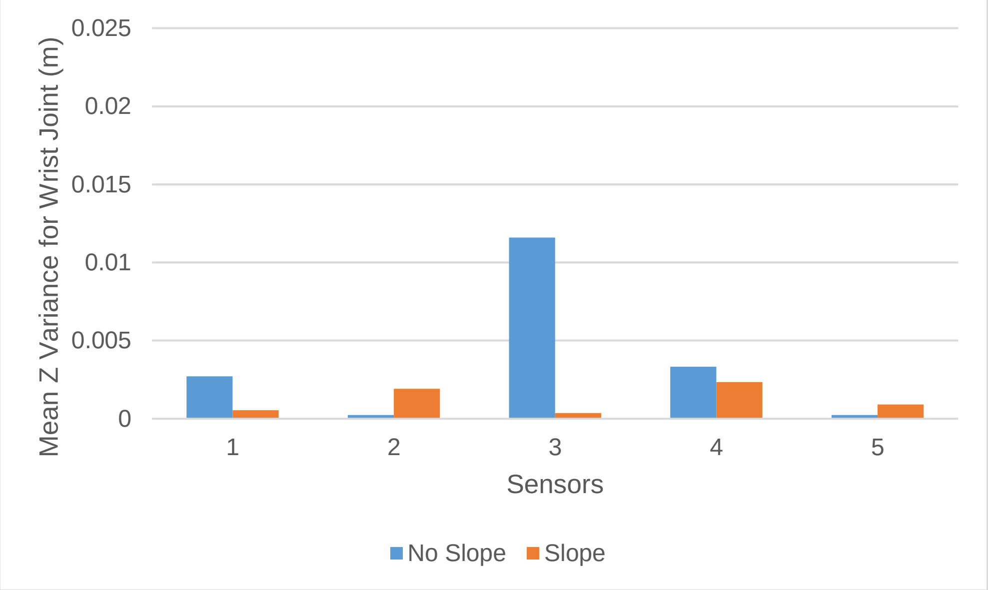

Figure 43: Mean Z Variance for Task 2 Right Wrist Joint for Fingerprint Interaction

Sensor three on a sloped surface reported in a high variance in both x and z data on a flat surface. Studying the form factor of sensor 3 (see Section 5.2.2), the design was different to other sensors. The housing was much taller and had a smaller size for the reader. The result suggests that users had difficulty to keep their wrist still during the interaction. Observations and video footage support this case; participants seem to adjust their wrist to place their finger comfortably on the sensor; however, through performance assessment, there were no outliers with NFIQ or throughput metrics. There was also no reported awkwardness of using the sensor as the post-questionnaire reports.

Table 57: Statistical Significance of X Variance for Right Wrist Joint on NFIQ and Throughput during Task 2 (“+”: p<0.05) for Fingerprint Interaction

Scenario	NFIQ		Throughput	
	Spearman	Sig	Spearman	Sig
1	.187	-	-.025	-
2	.346	-	.028	-
3	-.293	-	-.087	-
4	.514	+	.010	-
5	.202	-	.013	-
6	.091	-	.357	-
7	.018	-	-.217	-
8	.032	-	0.97	-
9	-.003	-	-.057	-
10	.330	-	-.274	-

Table 58: Statistical Significance of Z Variance for Right Wrist Joint on NFIQ and Throughput) during Task 2 (“+”: p<0.05) for Fingerprint Interaction

Scenario	NFIQ		Throughput	
	Spearman	Sig	Spearman	Sig
1	.319	-	.505	+
2	.062	-	-.048	-
3	.156	-	-.107	-
4	.056	-	-.283	-
5	.089	-	-.066	-
6	-.013	-	.013	-
7	-.034	-	.053	-
8	.264	-	.187	-
9	-.103	-	-.121	-
10	.030	-	-.138	-

The Z co-ordinate variance also provides insight into the interaction, users, on average, made more frequent movements for the non-sloped version of the sensor three compared to a sloped version, further supporting the difficulty in using this sensor. To investigate the impact of these variances on sample quality and transaction time, a Spearman's Correlation was performed. Tables 57 and 58 above provide insight into the interaction of multiple fingerprint sensors.

From the data collected, it is evident that the wrist variance measured through this task did not have a significant effect on sample quality or throughput metrics. Two relationships were identified, however; a High X variance in wrist movement for the fourth sensor on a flat slope did influence NFIQ. Additionally, there was also a relationship found between Time on Task and the sloped swipe sensor, stating that a higher variance increased time; however, this was to be expected for a swiping scenario. Therefore, as there were spikes in variances in the data for sensor three, this did not have an impact on the performance metrics recorded in this study.

In summary, while it is a useful tool to be able to measure the wrist movement and location from the sensor, from these results the conclusion is that there is not a meaningful impact from wrist variance on performance.

Question 4

Studying positional data such as where users are standing during interaction and the implications of that position on performance, has yet to be assessed in academic literature. Traditionally, feet symbols are placed within self-service systems to guide the user to stand within an optimal camera view to capture a face image. Additionally, the vendor sets where the symbols are placed, and as such, this will differ between systems and airports. There are currently no standards to determine where symbols should be put unlike many other forms of information within a border control environment. It is assumed that feet symbols are placed based on testing throughout the development cycle, but this is not largely discussed. It is possible that the placement of these feet symbols may affect more than just the quality of the face biometric presented to a camera. Also, multi-biometric systems, which may combine both face and fingerprint modalities, may require users to stand in the same area for both interactions. Therefore, it is a possibility that where users stand may influence their performance during an interaction with different modalities.

For this data capture, feet symbols were placed directly in front of the pedestal kiosk, centrally placed 1.15m and 1.35m from the Kinect. Participants were not given specific instructions to stand on the feet symbols; however, information displayed on the screen did suggest to users to stand on the symbols. Therefore, the data captured by the Kinect sensor can report the stationary position of the user during interaction. The analysis will then identify where the user is standing for a scenario and if there is a statistically significant effect on performance. Table 59 below provides an insight into the average position for users across the ten scenarios during the interaction process (Task 2: Biometric Read).

Table 59: Hip Position during Task 2 Biometric Read for Fingerprint Interaction

	Hip Left			Hip Right		
Scenario	**X**	**Y**	**Z**	**X**	**Y**	**Z**
1	-.056	-.141	1.251	.079	-.144	1.241
2	-.046	-.140	1.348	.086	-.144	1.338
3	-.030	-.133	1.347	.105	-.136	1.339
4	-.014	-.137	1.341	.120	-.140	1.337
5	-.010	-.142	1.347	.127	-.145	1.338
6	-.070	-.127	1.269	.143	-.125	1.252
7	-.049	-.152	1.337	.162	-.149	1.365
8	-.029	-.146	1.348	.110	-.148	1.341
9	-.006	-.162	1.569	.134	-.163	1.350
10	-.010	-.148	1.356	.128	-.150	1.356

Table 59 reveals that users were positioned on the feet symbols between 1.2-1.4m (the feet symbols were 20cm long). It seems subjects stood slightly closer for the swipe sensors (Scenario 1 and 6) than other sensors.

As discussed previously, Hip Left and Hip Right joints were 100% fully tracked for Task 2 and therefore retained a high degree of confidence in the accuracy of the recorded data.

Figures 44 and 45 below provide an insight into the Z coordinate variance for both Hip Left and Hip Right joint during the interaction (Task 2) process.

Both figures show the average Z variance for Hip Position during fingerprint interaction. Results indicate that users are making moving closer towards the sensor four on a sloped surface and slightly further away on a flat surface. This data differs from our previous research question, which did not highlight any significant movement in the wrist movement for Scenarios 4 and 9. The information here suggests that users must alter the way they stand to interact with the sensor.

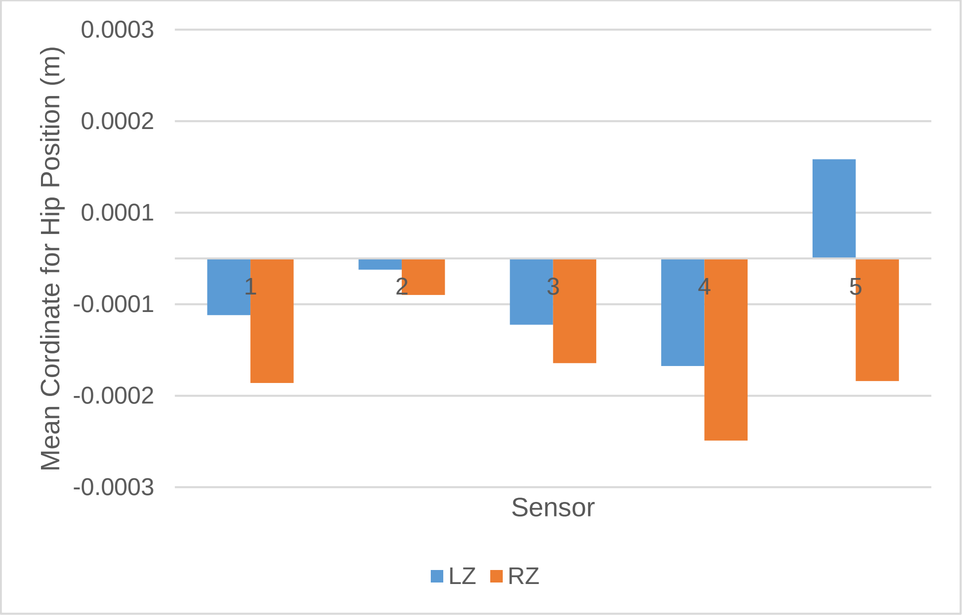

Figure 44: Mean Z Variance for Flat Surface sensors for Fingerprint Interaction

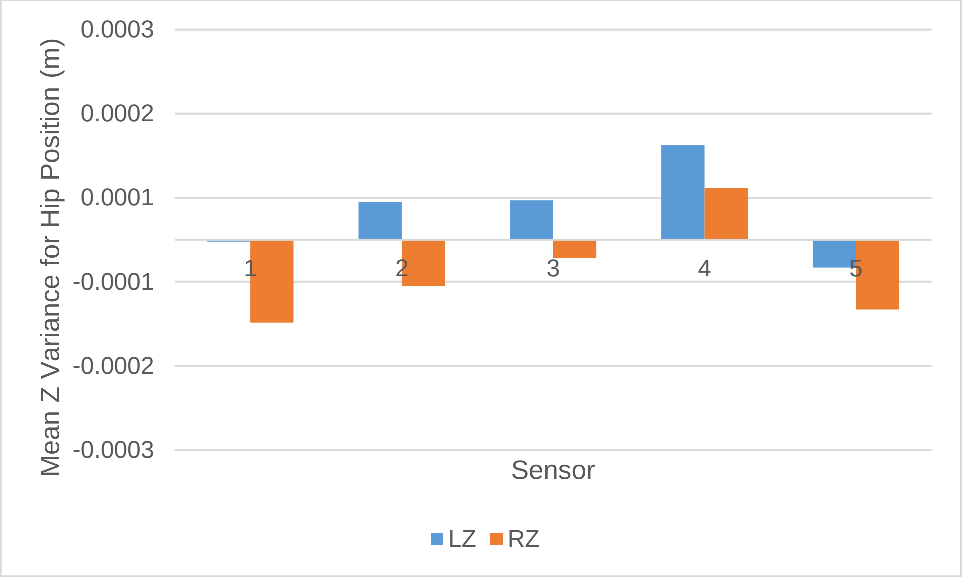

Figure 45: Mean Z Variance for Sloped Surface sensors Fingerprint Interaction

Otherwise, results remain relatively stable between surfaces. If users are positioning themselves further or forward during the interaction, this could lead to new questions. It could be based on sensor placement, the design of the system or perhaps other external factors. Ultimately, from this ability to report this information, the main goal will be to assess wherever there is an impact on performance. To assess if there was a relationship, a Pearson’s correlation was performed. Table 60 demonstrates the results.

Table 60: Statistical Significance of HipRight Z Variance on NFIQ and Throughput on Task 2 ("+": $p<0.05$) for Fingerprint Interaction

Scenario	NFIQ		Throughput	
	Pearson	Sig	Pearson	Sig
1	-.271	-	.311	-
2	.031	-	-.177	-
3	.146	-	-.008	-
4	.090	-	.531	+
5	.265	+	.010	-
6	.190	-	-.112	-
7	.254	-	-.010	-
8	.115	-	-.038	-
9	-.025	-	-.054	-
10	.069	-	-.255	-

For both Hip Left and Hip Right Joints, a statistical relationship was found between variance (how much the user is moving) and throughput for the fourth sensor on a flat slope (Scenario 4). Although there was not a significant impact on NFIQ, it seems users had to adjust themselves during fingerprint interaction. Looking for feedback and observations for Scenario 4, our findings did not reveal any difficulties during data collection.

Looking back further at Table 57 for Question 3, a significant relationship was found between Right Wrist Joint X variance and NFIQ for the fourth sensor on a flat surface. These relationships show through tracking that there may be underlying problems with the sensor possibly due to the size, shape or placement. Without these tracking methods, this may not be a result likely to be found during regular testing of biometric systems.

5.5 Summary

Self-service biometric systems are increasingly implemented across the globe for verification and identification solutions. Although there has been a range of studies investigating usability and performance assessment, the work undergone in this chapter has provided initial steps into studying the applicability of using movement tracking in this type of environment.

This chapter has identified the novelty of applying a tracking system to data collection methods. The results have demonstrated the ability to detect HBSI categorisations based on skeletal positional data within a biometric testing scenario. These first steps using the Kinect sensor has enabled the transition from a manual to an automated method for performance assessment. Alongside HBSI categorisation, the ability to categorise different joint positions during the interaction has also been introduced. Through this approach, the system can determine if a certain joint is within the desired area before proceeding to the next task (e.g. a user’s wrist must be detected near the sensor for fingerprint interaction to begin) which may reduce the likelihood of an error occurring further in the process.

Investigating skeletal data, results have highlighted unexpected variance in some joints throughout multiple scenarios. Further analysis has highlighted a significant effect on some variables on performance metrics such as NFIQ and throughput results. This initial work has revealed a potential unique ability to dwell deeper into operational design and testing before biometric systems are fully deployed.

Future testing should consider looking and classifying the impact of erroneous presentations only, determining the sensors ability to classify unlike incorrect behaviours in more detail in addition to assessing the consequences of those actions on both sample quality and throughput. Users in this study were asked to complete the scenario based on the information provided on the screen, revealing a high frequency of successful captures.

While this study was to investigate the applicability of gathering positional data initially, the next study will begin to consider improving information feedback based on this novel assessment method. Tracking movements and actions of a user within a controlled environment have the potential to increase throughput and the quality of the biometric sample obtained. Conducting a usability assessment of this initial system has identified similarities to other systems while introducing a new methodology for performance evaluation.

In summary, the chapter has contributed:

- Foundations for the HBSI Automation Program, a system to automatically categorise a presentation made to a sensor
- Results outline the effect on performance based on multiple variables such as user’s height, the sensor used and angle of the slope
- A methodology to assess users position within a controlled environment, aiding the system’s ability to detect movements and presentations to a sensor before relaying task-specific instructions

CHAPTER 6. FACIAL INTERACTION

6.1 Introduction

Facial interaction is the leading biometric used in border control scenarios and is, therefore, the main modality employed in implementations across the globe. The design, build, and the information displayed throughout the interaction process for these systems differ between vendor and country and maybe a confusing process for users. There are a number of recommendations and standards participating countries must adhere to; such as the quality of the image stored on the biometric passport and certain symbols used throughout the process. It is not known how much of an impact this information has on the interaction process. For instance, is there a difference in performance between variations of the system based on the information displayed during facial interaction? Some systems, such as the eGates in Heathrow, UK, display a mirror image of the user on the screen during capture. Other systems, such as the APC kiosks in the US offer simple pictorial information to 'look at the camera', pointing upwards to the camera built into the system. This chapter focuses primarily on building upon the success of the previous data collections using skeletal tracking methods to assess human-biometric performance within the facial biometric modality.

One of the research goals of this study is to investigate the impact of information in facial biometric systems and uses skeletal tracking to enhance performance assessment further. There are three main categorical groups for information displayed within a system; icons by itself, icons with added text, and text by itself. Textual information is rarely used within border control scenarios due to the language barriers between countries, but for countries to the east, textual information may be more common.

Information in this study is relayed as 'dynamic' feedback to the user, whereby their presentation is processed in real time by the Kinect Sensor and feedback is displayed based on individual elements of correct or erroneous interactions. For example, if the users head is tilted, outside of the acceptable region, then feedback will be produced on the screen to correct the user's presentation. The analysis of the data captured will consider several additional variables such as; joint data, yaw, pitch, roll of the face presentation, wherever the eyes closed and the subject's facial expression.

This implementation consisted of using skeletal tracking to automate the HBSI process further enabling a fully automated assessment based on user tracking. The overall design and build of this system were similar to the previous data collection, mimicking the size and layout of a self-service e-Gate without the use of barriers or gates.

Facial Interaction can be assessed through many means; reporting on sample quality, time on task and usability assessments, which look at efficiency, effectiveness and user satisfaction for any given system.

The biggest problem with non-standardisation between systems is the effect on acquisition. Face images captured by a system must meet requirements set by various ISO publications (as discussed in Chapter 2), but matching is widely led by the country and the designer of the system, which will choose a threshold for a quality that samples must meet.

ISO/IEC JTC 1/SC 37 [128] [127] is the technical committee who develop the standardisation of biometric technologies. To date, there are 121 published ISO standards from the workgroup with 30 currently under development. The committee is made up of six working groups (WGs) that carry out standardisation in specific areas within biometrics. Table 61 reports on the current areas covered.

Table 61: Working Groups within the ISO/IEC JTC 1/SC 37 Technical Committee

WG	**Area**
WG1	Harmonised Biometric Vocabulary
WG2	Biometric Technical Interfaces
WG3	Biometric Data Interchange Formats
WG4	Technical Implementations of Biometric Systems
WG5	Biometric Testing and Reporting
WG6	Cross-Jurisdictional and Societal Aspects of Biometrics

The image captured during the border crossing process must match an enrolment image within a certain threshold, typically stored in an electronic document such as a passport. Algorithms are used to match the captured image against the enrolment image while completing other jobs: checking for fraudulence in the reference image, checking the image against a watch list and internal databases. The outcome of the matching process is determined by the various factors such; the quality of the enrolled passport, which the ABC process cannot control, and the threshold for FRR and FAR rates and quality of the stored image. However, both factors differ from country to country and even between implementations used within those countries. The protocols for storing a passport image in one country may differ from another, and the image stored on the travel documents may have different static and dynamic properties from another countries passport. The focus for ABC systems then is to ensure that the captured image meets a required standard.

The general approach to the face image format for border control systems requires developers to specify the required format of the image, compression and the best practices for taking images within that system.

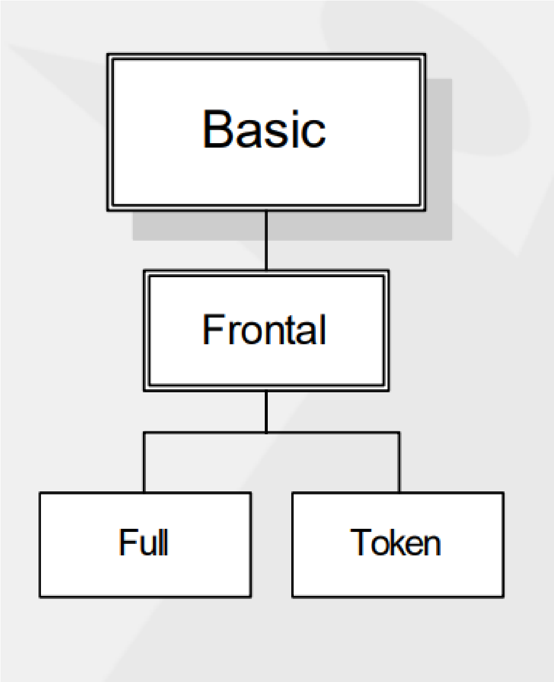

Figure 46: ISO 19794-5 Flowchart of Interchange Formats [34]

ISO/IEC 19794-5 Biometric data interchange formats – Part 5: Face Image data [34] from WG3 describes the interchange formats for facial recognition systems. Figure 46 above details a flowchart for interchange formats as defined by 19794-5. In brief, the requirements for face images used are as follows:

- Basic: Specifies a record format and what image data to use. It does not detail a mandatory scene (environment), photographic or digital requirements
- Frontal: A face record that adheres to additional requirements to the Basic standard which is appropriate for frontal face recognition (automated) and human examination (manual). Frontal photos then, must either be full or token:
    - Full: Includes the full head with all hair and in most cases, neck and shoulders. Full Frontal is the standard used for ePassports
    - Token Frontal: Specifies frontal images with a specific geometric size and eye positioning based on the width and height of the image. The images may be used in manual situations or for specific identification scenarios. However, this standard is typically not used in automated border control scenarios due to the specific geometric requirements that cannot always be replicated in different passports

All frontal images, however, must adhere to specific requirements, namely:

- Pose
    - Full-face frontal pose should be used. Rotation of the head should be less than +/- 5 degrees from frontal in every direction (yaw, pitch and roll)
- Expression
    - Should be neutral (non-smiling) with both eyes open (not wide open) and mouth closed
- Background
    - The background should be plain and contain no texture containing lines or curves that could cause problems when matching. Background should be a uniform colour or a single colour pattern
- Lighting
    - No shadows or point light source (Flashes). Should be equally distributed across the face. Multiple or diffused balanced sources or other lighting methods can be used

The enrolled image stored on the passport must also adhere to scene constraints, such as no hair covering the front of the face, no shadows, no sunglasses or glare on glasses. Glasses, in general, are accepted as they if they do not obstruct the eyes. Hats and other accessories must also be removed.

With these standards in mind, this data collection focuses on capturing images that adhere to addressing these standards in real time. Rather than through the traditional practice of capturing several pictures and assessing the quality of the image based on the best sample, the system proposed in this chapter will require the user to be presenting the required pose and expressional requirements for the capture process to begin. Background and lighting will be controlled and tested before trials, and all images will be assessed against the full-frontal standard to ensure conformity.

Also, the way guidance and information during the transaction is displayed to the user will also be examined. There have been various research before on the impact of instruction and guidance within the interaction process [65] [63] [8] but largely from a non-automated environment. Although studies conclude that different types of feedback do have an impact on performance, to date there is no work on specific or ‘dynamic’ feedback that is adapted based on the user’s biometric interaction. As the information displayed on the screen changes between trials, results may indicate differences between scenarios and may offer an insight into future best practices.

This chapter then, introduces the second data collection, using the Kinect sensor to assess, capture and process a user’s face image. As discussed in Chapter 5, the Kinect sensor is an accurate, low-cost, RGB-D camera that can be used for a variety of scenarios. So far, the device has been used to measure the user’s position, pose and movement in both simple movements and through fingerprint interaction. In this chapter, the ability to assess the face presentation is introduced, enabling the system to assess face pose and expression.

HBSI Interaction assessment is fully automated, and data from the Kinect is cross-referenced alongside traditional HBSI evaluation metrics. The impact of information through feedback is analysed and discussed.

6.2 Data Collection

A border control system using facial interaction is typically designed to match a captured image during the transaction process against a previously captured image enrolled on a token such as an electronic passport. Both photos must meet ISO defined standards and meet a threshold when comparing the two images. Border control solutions may be automated or semi-automated, but in most situations, the system will involve the user walking to a camera, presenting their (face) image and waiting for a result before continuing through a gate or barriers.

The data collection conducted in this chapter uses the HBSI assessment method to categorise performance of the system, looking at efficiency, effectiveness and to a degree, user satisfaction.

The system used in this study is an enhanced version of the HBSI automation program used in Chapter 5. However, instead of collecting fingerprints through a secondary program, the HBSI program was configured to capture images by using a combination of skeletal tracking and image processing elements. Through this method, the face pose and body position can be tracked and analysed in real time.

Section 6.2 presents the methodology used for this data collection, detailing the experimental setup of the system, the scenarios the users face and the settings for recording. The demographics of the users are discussed, and guidance, training and how the data is stored is also considered.

6.2.1 Experimental Setup

The hardware setup for this data collection consists of a Kinect V2 device, multiple web cameras, a PC and a display monitor. The scenario and layout were configured similarly to the Fingerprint Data Collection in Chapter 5. However, there were several differences, the first being the removal of the pedestal and the fingerprint sensors. The Kinect sensor was placed directly on top of the monitor instead of behind as it was in the previous data collection.

Feet symbols were placed centrally in front of the kiosk, 1.6m away from the monitor, to indicate participants where to stand. The feet symbols measured 25cm in length. The use of feet symbols during facial interaction is crucial to capturing the biometric of a user as it dictates the range of the camera. There were no physical gates or barriers used for this collection due to the size of the room used. Figure 47 below demonstrates the layout of the room used for data capture.

A Logitech HD camera was placed on a tripod on the table for enrolment photos. Further information on enrolment images is discussed in Section 6.2.3.

Two web cameras were configured to capture video footage of each trial to enable ground truth comparisons. Cameras were set up in the experiment (overview, C1) and above the feet symbols (feet view, C2). All video was captured in 720p.

A white backdrop and a lamp were placed behind and in front of the Kinect, Sensor V2 to adhere to the environmental conditions for full frontal images. More information on the effect of this setup is discussed in Chapter 6.2.3.

The Kinect V2 sensor was fixed on the top of the monitor which was measured to be 1.6m from the ground. The camera component of the Kinect V2 was positioned in the centre of the monitor. The sensor was angled downwards slightly at 10 degrees to account for users of a slightly smaller height but remained within the recommended 43.5-degree optimal range. The tilt also enabled the sensor to detect the floor at all distances.

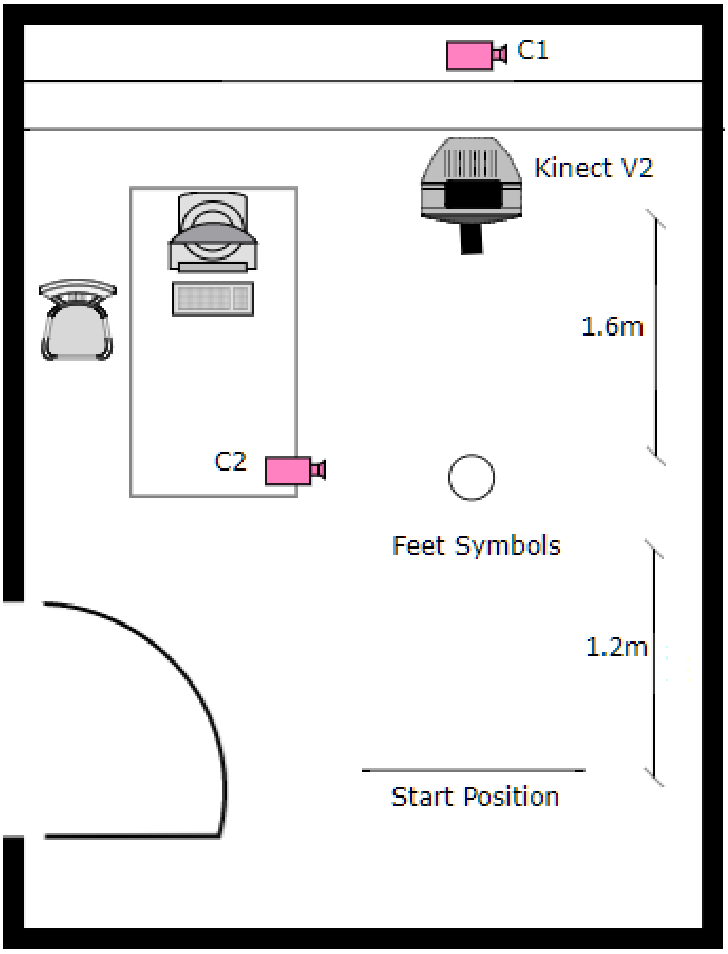

Figure 47: Experimental Setup of the Face Interaction System

The camera component of the Kinect V2 was configured this way to comply with full frontal standard [34]. A range of images at different angles and monitor heights were captured and compared with a range of participants with different heights before selecting a final position. The position of the sensor also closely resembles many other systems where the camera is placed directly atop of the screen.

6.2.2 Scenarios

Each participant was required to attempt facial recognition a total of ten times. Users were allocated a group at random. Each group was designated a type of information that would be displayed on the screen. The three groups were; Text, Text and Icons or Icons by itself. Each user was required to complete five interactions from their allocated group in a random order as well five baseline scenarios. The baseline scenarios were based on current live implementations. Users were then allocated a random order to complete all ten scenarios to minimise the effect of order in performance as well as in user's habituation. The user was not made aware of which order or what information they would be attempting on screen. The information displayed is detailed further in Table 62 below.

Table 62: Scenarios and Information for Face Interaction System

Scenario	Information	Specific Information
BASE01	Baseline	Look at Camera Image (UK)
BASE02		Look at Camera (US)
BASE03		Mirror Image (UK)
BASE04		Mirror Image with Text (US eGate)
BASE05		Camera Icon with arrow pointing towards camera (EU eGates)
ICONS01	Icons	Large ISO Icons (centred)
ICONS02		Small ISO Icons (centred)
ICONS03		Large ISO Icons w/ Live Image
ICONS04		Medium ISO Icons (Top)
ICONS05		Medium ISO Icons (Bottom)
TEXT01	Text	English
TEXT02		French
TEXT03		German
TEXT04		Japanese
TEXT05		Czech
ICONS&TEXT01	Icons & Text	English & Small Icons
ICONS&TEXT02		English & Large Icons
ICONS&TEXT03		On Top
ICONS&TEXT04		On Bottom
ICONS&TEXT05		Large Icons/Small Text

The information presented to the user was considered ‘dynamic’, during the capture process feedback was displayed that adapted to the user’s presentation. The baseline group, however, displayed traditional ‘static’ information, where there was no assistance on the screen to correct erroneous presentations.

Dynamic feedback, for this study, is defined as a reactive information through which the skeletal data and face pose is analysed in real time, and appropriate feedback is given. Dynamic feedback could include presenting guidance to the user in an attempt to address the following issues:

- Tilt head left or right to correct Roll
- Lift head up or down to correct Pitch
- Face camera straight to correct Yaw
- Remove glasses
    - Only if glasses were obstructing the cameras ability to take a picture
- Remove Hat/Accessories
    - If detected (the system could not obtain 100% tracking for the joint due to occlusion)
- Stop Smiling
    - To provide a neutral expression

If all the conditions were met according to the Kinect based evaluation, then a 3-second countdown was displayed. The system would then capture a series of quick photos (on average 5-6 images) over a two second period. Users were then instructed that the process was over and that they should exit the system.

Non-baseline scenarios in this study displayed dynamic feedback in different ways, such as in; different languages (text), simple large or small icons, either on the bottom of the screen or on the top (icons) or a mixture of both (icons and text). Several scenarios included a live image with a combination of icons and text.

The Baseline scenarios (BASE01-05) were based on the information displayed in real life implementations across the globe (e.g. EU eGates in the UK, Germany, Japan) and did not provide any form of dynamic feedback.

Apart from the type of information and feedback displayed during the interaction, all other variables were kept the same throughout the transaction.

Figures 48 below displays some of the visual differences between several of the baseline scenarios.

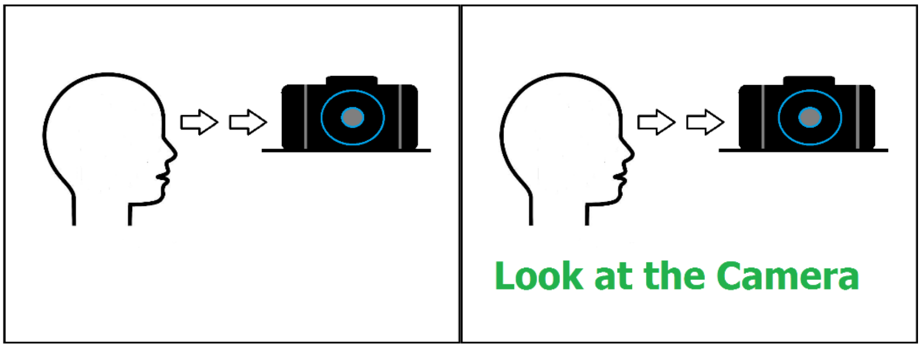

Figure 48: Difference in Scenarios for Face Interaction. Top Left: BASELINE01. Top Right: BASE0LINE2. Based on information in eGates (UK)

6.2.3 Recording

Most of the data was collected from the HBSI Automation package. The Neurotechnology MegaMatcher biometric SDK (from Chapter 5) was not used during data capture and was only used for post-processing purposes to check the quality of the image.

The (HBSI Automation) Kinect package was developed by the Purdue University and University of Kent team. The SDK was updated to capture additional data using the Face Basics Package introduced by Microsoft in 2015 [146]. Face Basics can detect additional features by using image processing techniques such as; expression, wherever the user is wearing glasses, face pose rotation (yaw, pitch and roll) and wherever the eyes were open or not for each frame. The results were also logged in addition to the joint data as described in previous studies.

Timing information, collected through time stamps, was previously collected from the Neurotechnology SDK in previous data collections but was captured from the HBSI Kinect program. Timestamps were printed for each frame and in a separate log file when each form was displayed. Timestamps are used to calculate timings (time taken to complete a task, overall time).

Two Logitech HD C920 cameras were configured to record video footage of the experiment, allowing ground truth actions to be compared against automated HBSI categorisations. Specific behaviours and any other relevant information, such as if assistance was given, is registered as observation notes throughout the interaction for each user (see Section 6.4) and compared to video footage after the experiment. A personal profile (not identifying the user) was collected and established for post-data capture analysis. See Table 63 below for further information on the data gathered.

Table 63: Recorded Information for Face Interaction System

Information	Description	Resource
Timestamp – HH: MM: SSS.SS	Printed and Exported to individual log file identified by a unique code and scenario number – processed when a form is first displayed	HBSI Automation Kinect Package
Face Image	Jpeg file	HBSI Automation
Behaviours/Movements	720p Video (.avi file)	Logitech HD C920
Personal Profile	See Section 6.2.4	Questionnaires
Kinect Data	See Section 6.3.1	HBSI Automation

Subjects were allocated to a group at random (Text, Text and Icons or Icons). All guidance was presented on the same standing monitor as in Chapter 5. Users were presented a total of five forms/screens during the interaction. Each form represented a stage in the system (e.g. move to feet, look at screen, capture and process complete). Scenarios adapted icons and other symbols based on the ISO/IEC 24779-1 standards [127]. Each form was displayed at the start of the four critical tasks involved in the interaction process (Table 64).

Table 64: Forms displayed in Face Interaction System

Form	Description
Feet Form	The form displays instructions requesting the user to move to the feet symbols.
Ready Form	Displays instruction to the user to present their face to the camera
'Dynamic' Processing Form	Instructions that either 1) gave feedback to correct presentation or if presentation correct 2) to countdown to capture
Captured Form	Informs user capture is successful
Trial Complete Form	Informs user process is over and to exit system

As discussed previously in the Behavioural Framework defined in Section 3.5, there were four behavioural led tasks to this data collection experiment. Token Interaction was not used during this study and so the third and fourth task investigated the presentation and response to the result. Aligning these critical movements to our Behavioural Framework allows the breakdown of performance measurement at a task level. Table 65 below explains the behaviour led tasks and their associated variables for this experiment.

Kinect Categorisations or Behaviours were recorded by the Kinect HBSI program and encoded with a numerical value to assist in statistical analysis. Results are assessed in the Kinect analysis section, 6.3.1. For Task 2, Biometric Read, specific errors were categorised, and a tally of the number of errors presented was collected. See Section 6.3.1 for further analysis.

Table 65: Facial Interaction Tasks and Associated Variables

Task	Related Variables
Entry (1) – Movement from starting position to feet symbols Defined between FeetForm and ReadyForm	Timestamp 1 Kinect Categorisation
Biometric Read (2) – The movement of presenting the face to the sensor Defined as the interaction between ReadyForm And CapturedForm. ** Dynamic Form appears during this process relaying feedback to the user	Timestamp 2 Kinect Categorisation Kinect Specific Error
Biometric Accept (3) – Response from System Defined between CapturedForm and TrialCompleteForm	Timestamp 3 Kinect Categorisation
Exit (4) – Movement from system to starting position	Timestamp 4 Kinect Categorisation

Table 66: Personal Profile and Task Evaluation for Face Interaction System

Personal Profile	Task Evaluation
Gender, Ethnicity Height, Age, Handedness Accessibility Facial Images Captured Before ABC Systems Before, If yes, which modality Times Travelled, Travel Alone/Companions Illness, Hours of Sleep Accessories Temporal Illness (Burns, Cuts)	Information Evaluation (1-5) Results (1-5) Conclusion (Descriptive Feedback)

In addition to the performance analysis performed, a personal profile and task evaluation form were collected from users prior and after the data collection respectively. Personal profiles help to establish the user base and provide a scope on the data captured. Table 66 below describes the information collected from the Personal Profile, ABC Questionnaire and Task Evaluation forms. The information collected may provide an insight into differences in results between scenarios.

The main goal of collecting data from both the personal profile and task evaluation forms was to build a ‘profile’ of users based on their demographic data as well as looking at some non-technical factors. The non-demographic information collected is discussed in Chapter 7, Future Recommendations. Section 6.2.4 discusses the demographics further while Section 6.3.2 HBSI assessment looks at the results from task evaluation questionnaire.

Questions asked for Task Evaluation were based on a 1-5 Likert Scale with one being strongly disagreed, and five being strongly agreed. Users were asked ten questions regarding their use of the system:

1. I was given enough information to complete the task
2. I found the process easy to complete
3. I completed the task without any difficulty
4. The information provided clearly described what to do during the process
5. It was clear when the process begun
6. The prompts given on the screen were clear
7. I was confused by the entire process
8. The order of the capture process was clear
9. It was clear when the process had been completed
10. This experiment will benefit me when I use ABC systems in airports

Face pose detected is achieved by the infrared component of the Kinect V2 sensor. The image is analysed through the camera in real-time, the head pose is deduced, and facial expressions can be collected. Embedded in the HBSI Automation program, data is collected on face rotations throughout the transaction. To enable a capture of the face image in Task 2, the rotation of the head must be less than ±5-degree rotation in each direction. The rotations must stay within these limits for three seconds which is then followed by a series of quick captures. Once the countdown has ended, and the capture process has started, which takes roughly 1-2s to complete, multiple images are taken of the face. There is a small possibility that the pose may change if the user is distracted during this period, so the best image is chosen for matching.

For analytical purposes, poses are ranked base on the rotations. Giving a ‘ranking’ to a pose enables a range of statistical tests to be conducted for data analysis and enables categorisation to be made. For example, a presentation may have a 1-degree rotation for both yaw and pitch, but roll rotation may be measured as 3 degrees. While this is nearly a perfect presentation (the user is looking straight on with a slight tilt to the left or right), for this study, this presentation was given a ‘good’ ranking, to account for the roll rotation. Table 67 below demonstrates ranked associations for face pose for this data collection. ISO specifications indicate any presentation within ≤ ± 5 degrees should be an acceptable sample for face matching. In addition, for the face presentation to be a rank 3 presentation or above; tracking elements required the eyes to be open, the mouth closed and the body skeleton within the required range of the camera. During initial testing, however, the mouth closed/no smiling expression had a poor accuracy of 68% (of ten users tested) and so was not used for analysis in this study.

Table 67: Yaw, Pitch and Roll Rankings for Facial Interaction (Dark grey rows refer to accepted presentations)

Rank	Association	Yaw	Pitch	Roll
1	Perfect	< ± 1	< ± 1	< ± 1
2	Good	< ± 3	< ± 3.5	< ± 3.5
3	Acceptable	≤ ± 5	≤ ± 5	≤ ±5
4	Unacceptable	> ± 5	> ± 5	> ± 5
5	Poor	> ± 7.5	> ± 7.5	> ± 7.5
6	Very Poor	>± 10	>± 10	>± 10

Throughout the data collection, users were asked seven questions relating to the information they had just received during the scenario. The questions asked were:

1. Was it clear when the task was begun
2. Was clear when to face camera for capture
3. Were the on-screen instructions clear
4. Was it clear how to position to the camera
5. Was it clear when the capture process was complete
6. Was it clear what the result was
7. Are you confident you completed the face capture process as intended

Each user answered either yes or no before continuing. Results are discussed in Section 6.3.2.1.

6.2.4 Users

Sixty participants (28 men and 32 women) were recruited without any special requirements. The only condition for taking part was that users must be over the age of 18 (a requirement to use ABC systems) and able to speak English.

Forty users were aged between 18-24 years old; eleven are 25-34, three 35-44, five are 45 to 54 and one person is between 55-64 years old.

Looking at the diversity of the participants, thirty-four of the participant's primary language was English. Four users were Italian, and three were Spanish. Nineteen users were from different parts of the globe such as Hong Kong, Iran, France and Slovakia.

None of the users had any issues regarding accessibility. Eleven participants were wearing glasses before the experiment began. Four participants were wearing daily contact lenses, and five users were wearing scarves. Users wearing glasses were asked to remove them for the enrolment but not for the verification stage of the trial.

Ten participants had reported that they were suffering from a temporary illness such as cold and headaches before the trial started. However, none of these factors were likely to have a significant effect on a presentation.

Users were also asked if they had used an ABC system previously and wherever they had a passport style photo taken within the last year. Twenty-two participants stated they had not had a passport style photo taken while 38 users did. Regarding wherever participants had used an ABC system before, 39 users had while 21 stated they had not.

On average, users had slept 7.24 hours the night before which might signify that the users may be well rested and are alert; a possible contrast to real users in ABC systems.

6.2.5 Guidance and Training

Participants were asked to read the Participant Information Sheet (PIS), which explained the purpose of the study in addition to confidentiality issues such as where their facial images were stored and how long each sample would be saved for on storage.

Upon the completion of the consent form paperwork, the researcher enrolled a passport style photo on a Logitech HD camera.

During the experiment, each participant was asked to follow instructions on the monitor and was instructed not to ask for any further information from the researcher. Users were instructed that they might only ask questions if they were completely stuck.

The researcher maintained a record log of any assistance that was required throughout the experiment.

Users were only allowed one attempt per scenario. As the system was built to correct erroneous presentations, through HBSI assessments the only errors that should be categorised are Failures to Acquire and Defective Interactions.

6.2.6 Data Storage

At the end of each trial, the facial image was saved to a secure, local drive and the timestamp of when each form was displayed was saved to a log file in a text document. Each face image was given a unique identifier to keep anonymity. The database was of a sensitive and personal nature. Hence, it was stored on a secure server where access to the database was limited to the investigator. The size of the database after the images and video footage was deleted after the data collection.

6.3 Performance Analysis

Performance analysis will be split into two categories; Kinect and HBSI Assessment. The main objective of the analysis is to determine the impact of information given through guidance on the screen on the transaction and if the introduction of face tracking throughout the interaction can highlight any performance issues.

As before, performance analysis will focus on the four critical tasks. The only tasks that have a change in the required behaviours are Task 2 and 3, Biometric Presentation and Read, which focused on a facial interaction rather than the fingerprint presentation. Table 68 below details more information on the task, its definition and an expected behaviour from a user.

Table 68: Task Breakdown and Expected Behaviours for Face Interaction System

Task	Definition	Expected Behaviour
1	Start of the transaction (entry) – user moves towards feet symbols on the floor and stands within the designated area	User stands on feet symbols, looks at screen awaiting further instruction
2	Information is displayed on the screen requesting the user to present their face to the camera (biometric presentation).	User stands still and looks at the camera, remaining still
3	Information on the screen confirms successful capture, processes the sample and displays result (biometric read).	The user should continue to watch screen looking for information. In real scenarios, may grab bags etc.
4	Systems display information to confirm trial is over and to move forward – end of the transaction (exit)	User moves way from feet symbols back to the starting area

6.3.1 Kinect Analysis

Following from Chapter 4 Kinect Analysis and Chapter 5 Fingerprint Interaction, the concept of tracking states and performance analysis using the Kinect sensor has previously been defined and studied. Results from these chapters have proven the Kinect's ability to track a single user within a self-service environment with a high-level of confidence that the movements pertain to the user. The data collection introduced in this study will focus on facial interaction and will further explore critical and associated tracking joints for behaviours tracked throughout the four critical tasks. Furthermore, face tracking elements such as the yaw, pitch, and roll rotation will be analysed in further detail, analysing how users present themselves to the sensor during the interaction. Tracking states and accuracy tests will be compared against the benchmark set in previous data collections.

6.3.1.1 Definitions

To achieve a Successfully Processed Sample, the ideal HBSI result, the four critical tasks are required to be completed successfully in order. In summary; the user must move towards the feet symbols, follow instructions on screen, look at the camera and upon instruction, and leave the designated area. Table 69 below details the definition of tracking for each task, identifying critical and associated joints that are tracked and analysed in real time.

Although this task did not require any specific movements of certain limbs for biometric interaction, such as the right-hand movement tracked in fingerprint interaction, it did require users to look at the camera for face capture. In addition to these joints, face tracking elements; yaw, pitch, roll and expression were considered. These variables, while not technically joints, will be referred to as 'critical' points for this study. The required 3D space coordinates are also stated. See Table 69 for more information.

Table 69: Behaviour Definitions for each task for the Face Interaction System

Task	Definition	Coordinate	Critical Joint	Associated Joint
1	The user should be standing still on feet symbols	Z	Hip Left, Hip Right	Spine, Shoulders, Neck and Head
2	User is following instructions on screen (dynamic feedback) and is facing camera	X, Z	Head, Neck Yaw-Pitch-Roll* Expression	ShoulderLeft, ShoulderRight, SpineTop, SpineMid
3	Feedback informs user of successful or unsuccessful capture	X, Z	Head, Neck	ShoulderLeft, ShoulderRight, SpineTop, SpineMid
4	User should the designated area	Z	Hip Left, Hip Right	Spine, Shoulders, Neck and Head

*Not joints but are critical to this step

Associated joints should have a fully tracked status to ensure a higher degree of accuracy in critical joints. However, there are no set requirements on where these should be positioned in the scenario.

6.3.1.2 Tracking States

Looking at individual joints further, Table 70 details the percentage of fully tracked joints per task for all users across all scenarios.

Observing the data, all critical and associated joints achieved fully tracked status for each of their respective tasks. Fully tracked critical joints testifies that the Kinect sensor was accurately tracking users within this scenario with a 'high' confidence level in the data

Table 70: Percentage of Fully Tracked Joints across all Tasks for Face Interaction. (Dark grey refers to critical joints and lighter grey associated)

Joint	Task 1	Task 2	Task 3	Task 4
Head	100%	100%	100%	100%
Neck	100%	100%	100%	100%
ShoulderLeft	100%	100%	100%	100%
ShoulderRight	100%	100%	100%	100%
ElbowLeft	100%	100%	98.74%	100%
ElbowRight	100%	98.75%	97.69%	100%
WristLeft	96.58%	89.87%	92.55%	96.25%
WristRight	100%	91.11%	93.57%	100%
SpineTop	100%	100%	100%	100%
SpineMid	100%	100%	100%	100%
SpineBase	88.89%	100%	98.77%	97.58%
HipLeft	100%	100%	100%	100%
HipRight	100%	100%	100%	100%
KneeLeft	34.11%	32.58%	30.33%	41.58%
KneeRight	33.25%	28.58%	31.13%	40.25%
HandLeft	74.58%	100%	91.58%	68.56%
HandRight	70.22%	100%	100%	64.68%
FootLeft	10.58%	5.58%	6.11%	13.21%
FootRight	11.25%	4.58%	5.58%	12.55%

Furthermore, the analysis in this section focuses heavily on the face pose, through analysing the yaw, pitch and roll of a presentation. While face movements may change drastically throughout a presentation; the sensor does not measure these elements the same way as joints. Face tracking is only enabled when the head joint is fully tracked, reporting on a continuous number referring to the degree of rotation. Figure 49 below details the mean yaw, pitch and roll recorded for all users and scenarios across the four tasks

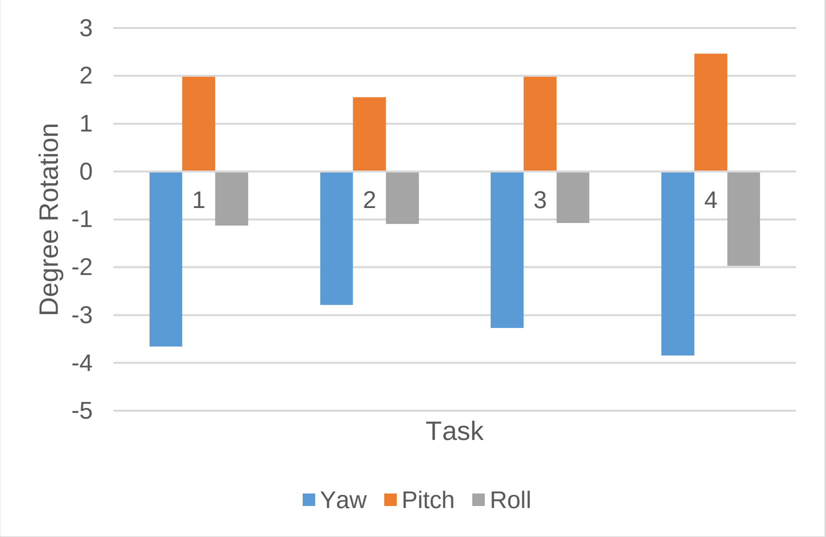

Figure 49: Mean Yaw, Pitch and Roll Rotations across all the four tasks for Face Interaction

As Figure 49 demonstrates; yaw, pitch and roll remained relatively consistent across all tasks. The average results demonstrate that users were typically exhibiting a good presentation and were looking towards the camera and monitor for information. For Task 2, biometric interaction, all three rotations improved during facial interaction before falling significantly for Task 4: exit, which was to be expected when users are turning and walking away. Although this data was an average of all users who are of different heights and sizes, a further analysis throughout this chapter will determine if other variables affected the captured face pose rotations.

6.3.1.3 Critical Tasks

The Behavioural Framework, as described in Section 3.5, breaks down the critical tasks involved in the interaction process at a level where identifiable behaviour can be tracked by the Kinect device. An analysis of the actions recorded will provide insight into the movements users naturally make throughout an interaction and wherever these have a significant impact on performance.

Like the Fingerprint Collection, the Kinect was configured to detect certain behaviours based on joint locations and image processing elements for each task as described in Section 6.2.3. The definition of the behaviour was logged to a record file with a timestamp for each task. Task 1 and 4 were almost identical to Task 1 and 4 from the fingerprint study concerning the movement the user makes towards the system. However, the Z-depth location changed slightly to account for the new system setup (feet symbols were placed slightly further away to achieve the requirements for a face image). Task 2 and Task 3 behaviours changed from a focus on right arm movement to facial position, with emphasis on the face pose.

Task 1: Entry

Participants were required to move forward and stand on feet symbols which were directly placed in front of the monitor. Instructions based on ISO standards were displayed on the monitor. See Figure 50 for an example of an icon displayed on the screen.

Figure 50: Feet Symbols used for Task 1: Entry

Table 71 below details the behaviour recorded and the Kinect definition for Task 1: Entry. Each OFEET or RFEET code (a cautious or undesired behaviour) was logged for each frame until a GFEET code was achieved. Dynamic feedback instructed users to either move either backwards, forwards or to the left or right if the desired position had not been achieved.

Table 71: Behavioural Framework for Task 1: Entry for Face Interaction

Code	Behaviour	Kinect Definition
GFEET01	The user approaches feet symbols and aligns feet correctly on the symbols.	HL & HR Z <= 1.6m && >= 1.35m
GFEET02	The user approaches feet symbols and feet are very slightly (2-5cm) off the feet symbols.	HL & HR Z <= 1.65m && >= 1.3m
OFEET01	User is slightly off centre	HL & HR X>= ±.05 && Z <=1.65m && >= 1.3m
RFEET01	User is in front of the feet symbols	HL & HR Z <= 1.35m
RFEET02	User is behind the feet symbols	HL & HR Z > 1.65m && <=2.0m
RFEET03	The user has not moved to feet symbols and is standing still.	HL & HR Z >= 2.8m && No variance

Where users are standing will have a larger impact on facial interaction. Standing too close or too far from the camera will make capture difficult. Although this system can detect the body in real time and relay information to the user by either asking them to step forwards or backwards (RFEET01 and RFEET02), problems can still occur. The most problematic error could be RFEET03, which may indicate that the user is distracted or confused on what to do. Identifying this in a real-time scenario could prompt extra assistance for a border guard through flagging procedures.

Table 72 below details the categorisation of the final behaviour recorded at Task 1: Entry.

Table 72: Observed and Tracked Behaviours for Task 1: Entry in the Facial Interaction System

Task 1 (Entry)				
	Observed		**Kinect**	
Behaviour	N	%	N	%
GFEET01	554	92.3	552	92.0
GFEET02	44	7.3	48	8.0
RFEET01	2	0.3	0	0.0

Of the 600 transactions, 37 transactions (6.1%) relayed dynamic information, requesting the user to either make an additional movement backwards, forwards, left or right. Baseline scenarios did not relay dynamic information to the user. All scenarios resulted in the user being in the correct position as determined by the Kinect.

Task 2: Biometric Capture

The capture process for Facial Interaction relies on two parts; 1) where the user is standing and their posture. Users should be standing upright looking forward towards the camera and 2) the head presentation, measured through the yaw, pitch and roll rotation. The eyes should be open and users should have a neutral expression. To begin this task, the user must be standing within the feet symbols, if any movements were detected, information requested the user to correct their position before continuing.

Table 73 below details the classification of the face presentation that was tracked for Task 2. The requirements for those movements are also detailed and require the head joint to be in range as well as a required rank presentation to form the behaviour code. See Section 6.2.3 for rankings in face rotation.

A very high percentage (93%) of presentations for this task was classified as 'good' interactions. Immediately recognising a good presentation to the camera should indicate that users are receiving the information correctly. Once an interaction was considered as a GFACE behaviour, the capture process could begin. A countdown began on screen, counting down from 3 seconds.

A successful countdown indicates the start of Task 3.

Table 73: Behavioural Framework for Task 2: Biometric Capture for Facial Interaction

Code	Behaviour	Kinect Definition
GFACE01	Ideal presentation. Straight face pose and little to no movement	Rank 1 Presentation Head Z <1.60m && >= 1.35m
GFACE02	Good presentation. Straight face pose, with little to no movements	Rank 2 Presentation Head Z <1.65m && >= 1.30m
OFACE01	Looking to the side/up/down very slightly that might affect capture	Rank 3 Presentation Head Z <1.65m && >= 1.30m
OFACE02	Minor movements in rotation that might affect capture	Rank 4 Presentation
RFACE01	Some noticeable movements during capture (e.g. looking to side slightly)	Rank 5 Presentation
RFACE02	Severe noticeable movements (e.g. looking down, looking away)	Rank 6+ Presentation

Table 74 reports on the frequency of tracked behaviours for the Facial Interaction System.

Table 74: Tracked Behaviours for Task 2: Biometric Read in the Facial Interaction System

Behaviour	N	Percent
GFACE01	523	87.3
GFACE02	34	5.7
OFACE01	17	2.8
OFACE02	0	0
RFACE01	20	3.3
RFACE02	5	0.8

Twenty-five interactions (4.1%) were considered as an ‘incorrect’ interactions (RFACE). However, these were corrected using the dynamic feedback system. To advance to Task 3, presentations were required to be considered a ‘correct’ interaction, so appropriate information was given to users to help them adjust their presentation. Through this system, there were no issues at this stage, and all users (100%) advanced to Task 3. Specific errors captured during this stage using the Kinect tracking is assessed in Section 6.3.2 HBSI Assessment.

Face Presentations are broken down further by information group. Table 75 details the frequency of classifications per group for Task 2.

Table 75: Face Classifications for Task 2 by Information Group

Information	Behaviour	N	Percent
Baseline	GFACE01	256	88.9
	GFACE02	14	4.9
	OFACE01	8	2.8
	RFACE01	10	3.4
Icons	GFACE01	93	92.0
	GFACE02	5	5.0
	OFACE01	1	1.0
	RFACE01	2	2.0
Icons & Text	GFACE01	81	81.0
	GFACE02	10	10.0
	OFACE01	4	4.0
	RFACE02	5	5.0
Text	GFACE01	93	84.5
	GFACE02	5	4.5
	OFACE01	4	3.6
	RFACE01	8	7.4

Task 3: Biometric Accept

Biometric Accept, for this system, takes place during the sub-system capture process. Successful capture presented a green tick on the screen after the countdown process displayed in Task 2. Unsuccessful captures were presented with a red cross. This task is primarily assessed on the overall quality of the image captured and the time taken to capture. However, Kinect tracking was still enabled for this task and the data captured measured if the user was standing still, to measure if participants were waiting for the next instruction. This mimics a border control system which will display information after successful matching to indicate that the user is free to pass through the barrier/exit the system.

This task will then focus on user position through looking at Hip Placement (where users are standing) and Head Position (are users still looking at the screen during post-capture) like the two previous tasks. Table 76 below detail the requirements for tracking for Task 3.

Table 77 details the behaviours tracked by the Kinect sensor and observed by the researcher. Twenty-four of the interactions failed at this stage due to Failure to Detect and Failure to Process errors (discussed in Section 6.3.2) and therefore were not processed by the Kinect sensor. No users moved before the system displayed the next piece of information.

Table 76: Behavioural Framework for Task 3: Biometric Accept for Facial Interaction System

Code	Observed Behaviour	Kinect Definition
GPOS01	User is standing still and looking at camera waiting for further instructions	Head, Hip L and Hip R Z <1.65m && >= 1.35m Rank 1-3 Presentation
OPOS01	The user is getting ready to move, but still standing in the same area. Not necessary looking at the camera	Head, Hip L and Hip R <1.65m && >= 1.30m Rank 3+ Presentation
RPOS01	User is moving away from system before process has completed and information has changed on screen	Head, Hip Left and Hip R >1.65m

Table 77: Observed and Tracked Behaviours for Task 3: Biometric Accept for the Facial Interaction System

Task 3 (Face)				
	Observed		**Kinect**	
Behaviour	**N**	**%**	**N**	**%**
GPOS01	575	95.80	555	96.68
OPOS01	25	4.16	19	3.30
RPOS01	0	0	0	0

Task 4: Exit

This task occurred after the interaction process (Task 2-3) had been completed and information had been displayed to the user depicting successful or unsuccessful capture. In either case, the user is expected to leave the station and return to the starting position. Table 78 details the behaviours below.

Table 78: Behavioural Framework for Task 4: Exit for Facial Interaction

Code	Observed Behaviour	Kinect Definition
GLE01	The user leaves the station	Hip Joints Z >=1.60m
RLE01	The user does not leave the station/does nothing	Hip Joints Z < 1.60m

Table 79 Observed and Tracked Behaviours for Task 4: Exit for the Facial Interaction System. Again, all behaviours were classified correctly for this short study.

Table 79: Observed and Tracked Behaviours for Task 4: Biometric Accept in the Phase 1 Data Collection

Task 4 (Exit)				
	Observed		**Kinect**	
Behaviour	N	%	N	%
GLE01	600	100	576	100
REL01	0	0	0	0

The ability to track critical tasks throughout a transaction has advantages in tracking specific behaviours or actions. Desired actions will lead the user through to the next task or stage of a system, and can positively impact system performance through on-the-spot training and improve presentation errors and thereby increase the likelihood of successful capture and subsequent verification or identification.

6.3.2 HBSI Assessment

The HBSI model is used to evaluate both system performance and individual transactions by looking at a range of correct and incorrect presentations made to a sensor. The Kinect V2 sensor analyses the interaction based upon the four defined tasks and allocates a weighting to the behaviour based on the impact on a potential presentation.

6.3.2.1 Presentation Framework

The application of the skeletal tracking system with the Kinect device enabled HBSI presentations to categorise interactions in real time based on body movements and face positions.

The HBSI Automation system was configured to assess movements for each of the four tasks like the system used in Chapter 5 Fingerprint Interaction, but changing requirements for Task 2 and 3 from a right arm movement to a face presentation. Each task specified certain conditions that the Kinect was searching for, e.g. the position of the hip joint and elements of a facial presentation within fixed parameters. The parameters and the score allocated to each task is defined in Table 80 below.

Table 80: HBSI Scores for each task for Facial Interaction

Task	Control Limit	Score
1	Z-Distance for Hip Left/Right < 1.65m & >1.35m	1
2	Head & Neck Joint = Z < 1.6m Yaw-Pitch-Roll = ± 5-degree rotation (Rank 3 or better)	3
3	Head & Neck Joint = Z < 1.6m Yaw-Pitch-Roll = ± 5-degree rotation (Rank 3 or better)	1
4	Z-Distance for Hip Left/Right >1.6m	1

HBSI categorisations follow the same formula introduced in Chapter 5, Section 5.3.2. The system requires two components for a classification; a score obtained from the behavioural tasks and either the presence or absence of a facial image. A combination of an image and score of 4 or above resulted in an SPS. Failures were reported for each task (Task 2 was a FTD presentation, Task 3 FTP). Table 81 details the manual and automated classifications for Facial Interactions.

Table 81: Manual and Tracked HBSI Categorisations for Facial Interaction

	Observed		Kinect	
HBSI	**N**	**Percent**	**N**	**Percent**
CI	0	0.0%	0	0.0%
DI	2	0.3%	0	0.0%
FTD	22	3.6%	22	3.7%
FTP	2	0.3%	2	0.3%
SPS	574	95.6%	576	96.0%

Automated HBSI categorisations for this data collection resulted in a very high amount of correct interactions through identifying a high number of Successfully Processed Samples (96.0%) with just 24 interactions resulted in either through a failure to acquire or process. There were two manual observations of a Concealed Interaction (Incorrect presentation but accepted as a correct) where users were not looking directly at the camera during capture. Through post-analysis, the users were slightly looking away in both cases.

However, the system accepted these two presentations as correct and did not record Defective Interactions or Concealed Interactions. Reviewing video footage both users gave an incorrect presentation but looked away during capture. The absence of many other CI and DI classifications was because the system required users to present a correct presentation to the system. Incorrect presentations were analysed in real time, and appropriate feedback was given to ensure correct presentations.

6.3.2.2 Evaluation Framework

This section reports on the HBSI evaluation framework component, identifying metrics in efficiency, effectiveness, user satisfaction and sample quality.

6.3.2.2.1 Efficiency

Through HBSI Classifications, 574 presentations (95.6%) were completed and marked as a Successfully Processed Sample (SPS) on a first attempt. Two presentations were accepted as completed automatically by the system but should have resulted in a reject and requested the user to re-attempt their interaction. However, these presentations were altered during the capture sequence, which the system automatically began after assessing correct presentation. Twenty-four presentations failed to process correctly, which all occurred due to a system failure during Task 3. This was a mixture of Failure to Detects and Failure to Process through system generated errors. As stated previously, users were only given one attempt per scenario and were not allowed to repeat even if the system crashed.

A major advantage of using ABC systems is its ability to increase throughput for travellers crossing a countries borders. The average time taken to cross the border via manual control is on average 32 seconds [104] while automated systems take on average 17 seconds to complete a transaction [31] [52]. Although systems differ in the method time taken to capture is measured, typically a transaction time is reported between the moment a user steps inside the gate to the point in time the exit gate opens. Time Taken to Capture then, is for the time taken to complete an interaction with a sensor. This measurement may change between system due to design and vendor but in all cases, the time taken to complete an interaction is an indication of how quickly users are proceeding through the process.

Timestamps were captured for the following events throughout the data collection:

- 1) System Started
- 2) Kinect Recording
- 3) Information Shown (Task 1) Feet Symbols
- 4) Information Shown (Task 2) Dynamic Feedback
- 5) Information Shown (Task 2) Capture Sequence
- 6) Image Captured (Task 3) Captured

- 7) Thank You (Task 4) Completed

Time Taken to Capture is measured between Task 2 and Task 3, more specifically between Dynamic Feedback (4) and Image Captured (6) events. Table 82 below details the descriptive statistics for Time Taken to Capture based on the information displayed to the user.

Table 82: Time Taken to Capture for Facial Interaction

Type of Information Displayed	Mean	Median	Minimum	Maximum
Baseline	7.47s	6.22s	4.85s	12.34s
Icons	6.92s	6.23s	5.12s	22.14s
Icons & Text	8.82s	6.22s	5.15s	15.65s
Text	7.85s	6.23s	5.22s	13.02s

Results indicate that information with icons only provided the best results. Users who completed the Icon scenarios took an average of 6.92s to complete the task, saving almost two seconds of Icons & Text which took the longest to complete at 8.82s. Baseline information, which is based on the information provided in live scenarios, performed somewhat in the middle, close to results from other systems [77] [7] [104]. It was expected that the addition of dynamic feedback for non-baseline groups would decrease the time taken to capture, but it is possible that presentations took longer to capture because of the time taken by users to understand the displayed errors. Feedback from users did indicate that the scenarios that used a form of text took extra time due to the additional time required to read and understand the information that was displayed. Indeed, scenarios in the text group were in different languages, so a larger time on task was to be expected.

A Welch T-Test was conducted to investigate the impact of the type of information against the baseline group. Each participant was allocated to a ‘group’ based on either icons, icons and text, or text. Participants were compared to a different group to ensure that the assumptions of the test were met. Results indicate a difference with a between-subjects design and are displayed in Table 83 below.

Table 83: Welch T-Test results for Time Taken to Capture Differences between Information Groups

Group	t	df	p
Icons	1.921	329.79	.167
Icons & Text	2.298	113.13	.132
Text	0.58	208.37	.809

Table 83 details the results from the test which indicates that there was no statistical significance between the adaptive information to the baseline groups for time on task. However, this simply accounts for all users and scenarios and not for any specific information displayed.

The focus of this study was to investigate the impact of different types of information on the transaction process. The four groups consisted of different scenarios within, forming subgroups for information. For example, three groups each contained a single scenario which displayed a live mirror feed to the user. Other groups contained variances of larger or smaller icons and text or relayed information towards the top or bottom of the screen. Therefore, a statistical analysis is conducted in the next section which investigates if there were any significant differences between these types of scenarios.

Several set-ups include a live image or ‘feed’ of the user during the transaction. The general purpose of displaying the feed is to provide a visual basis for users to make their presentation. While there is no research to back this hypothesis up, the survey presented in Chapter 2 demonstrated that 24% of configurations displayed a mirror or live image feed during the transaction.

Three groups each contained one scenario which contained a live feed during the transaction. Table 84 below demonstrates the overview of results for time on task for scenarios that used a live-image against scenarios with non-live images.

Table 84: Descriptive Statistics for Time taken to Capture for Live and Non-Live Scenarios for Facial Interaction

Type of Information	N	Mean	Median	Minimum	Maximum
Live Image	168	7.39s	6.24s	4.85s	36.14s
Non-Live Image	408	7.82s	6.22s	5.15s	56.21s

Studying the results further, a paired samples T-test was conducted to determine if there was a statistically significant difference between the inclusion of a live image compared to a non-live image.

As there were only three scenarios that included a live image, paired comparisons were made against the closest matching scenario within that group. For example, if the other information such as text or icons and text were in a similar position. The baseline scenario was compared to its UK counterpart. Table 85 reveals the results below.

Table 85: Paired Sample T-Test Results for Live and Non-Live Image Scenarios

Live Image Scenario	Non-Live Image Scenario	t	df	p
BASE03	BASE01	.639	112	.524
ICONS03	ICONS01	-2.37	36	.023
ICONS&TEXT03	ICONS&TEXT01	-4.78	41	.636

Results indicate a significant difference in the means for time to capture for the system displaying a live feed with icons and against a non-live feed version with just icons. Both scenarios displayed icons along the top side of the screen, which was directly below the camera. The non-live scenario had larger icons to compensate for space. However, the size of the live feed image filled roughly 80% of the screen size. Icons displayed at the top of the screen is more likely to benefit users who are making incorrect presentations while focusing on the camera. The average difference between these scenarios was 1.34s. However, other similar comparisons did not draw significant results which suggest a live feed does not necessarily benefit users within a transaction.

Similar dependent T-tests were conducted for other group comparisons to be made. Table 86 details descriptive statistics in time to capture between groups with large icons/text and small icons/text.

Table 86: Descriptive Statistics for Time to Capture for large icons/text and small icons/text for Facial Interaction

Type of Information	N	Mean	Median	Minimum	Maximum
Large Icons/Text	90	6.58s	5.95s	4.88s	8.56s
Small Icons/Text	90	7.74s	6.23s	5.21s	10.32s

Large icons/text resulted in a slightly faster average when recording time is taken to capture. Comparing similar scenarios, Table 87 details results for information with bigger icons/text against their smaller counterparts.

Table 87: Paired Sample T-Test Results for Large and Small Information for Face Interaction

Large Information	Small Information	t	df	p
ICONS01	ICONS02	-2.54	36	.002
ICONS & TEXT 02	ICONS & TEXT 01	-3.78	38	.432

Table 88 details results for information displayed on the bottom of the screen against the top.

Table 88: Descriptive Statistics for Time taken to Capture for Large Icons/Text and Small Icons/Text for Facial Interaction

Type of Information	N	Mean	Median	Minimum	Maximum
Top Icons/Text	90	6.58s	5.95s	4.88s	8.56s
Bottom Icons/Text	90	7.74s	6.23s	5.21s	10.32s

Information provided at the top of the screen performs slightly better by almost a full second. A paired sample T-Test reveals if this was a significant effect or not between matching scenarios within the same information groups. Table 89 reports.

Table 89: Paired Sample T-Test Results for Information displayed at the Top and Bottom of the screen for Face Interaction

Information Top	Information Bottom	t	df	p
ICONS04	ICONS05	-1.94	36	.001
ICONS&TEXT03	ICONS&TEXT04	-4.28	38	.325

Results indicate that there was a significant difference between information with icons on the top against icons along the bottom of the screen. Adaptive information was displayed just below the camera, in the direct eyesight with the camera so did not require the users to adjust their face pose.

6.3.2.2.2 Effectiveness

The Kinect system can adapt to the user's presence and therefore this had a significant effect on completion rate. Completion rate is calculated as a percentage of tasks completed out of the possible 600. As the HBSI assessment stated, 574 interactions were completed which resulted in a 95.6% task completion rate.

As this system was fully automated, there were no recorded assists from the observer during the data collection. Errors were recorded through the HBSI system and automatically processed based on detecting joint and face movements outside of the desired regions.

While this research has explored the ability to automate errors, it is crucial to understand if the impact of being able to classify these errors in real time can improve system performance. Understanding the relationship between errors and time taken to complete a task will be key in going forward to developing new implementations.

The total number of errors observed for each transaction was tracked by the sensor and confirmed through observations and video footage. Like the data captured in Chapter 5, Fingerprint Interaction, errors were defined as when an undesirable behaviour was performed during the interaction. These errors could have an impact on the user's interaction and be therefore tracked and processed through the HBSI system. As stated in Kinect Analysis, Section 6.3.1, the overall presentation was allocated a code based on the final presentation (GFACE01 etc.) made to the sensor, however, the number and what type of errors were also tracked.

Table 90 below defines all possible errors that were captured during Task 2: Biometric Read. When an error was identified, dynamic feedback was displayed to correct the incorrect presentation. Some errors may have a larger impact on the performance than others. For example, displaying feedback for users who are not looking toward the camera or turning their head away will likely cause a greater time taken to complete the task.

To investigate further, several relationships are assessed. The quantity of errors against time on task is investigated, with a hypothesis that more errors totalled will have a greater impact on time. Table 91 below details the overall number of errors that occurred during a single task for all scenarios.

Table 90: Possible Errors and their associated sources for Facial Interaction

Possible Error	Kinect Tracking	Confirmation	Potential Impact
Wearing Glasses	Image Processing	Final Image	Low-Medium. Depends on glass frame
Not standing on feet symbols	Body Tracking	Video Feed	Low-High
Rapidly turning head to the side	Image Processing (Face, yaw, pitch)	Video Feed	High
Not looking directly into camera during capture (Face Presentation 4+)	Image Processing (Face, yaw, pitch)	Final Image	Medium-High
Looking away, up, or down	Image Processing (Face, yaw, pitch)	Final Image	High
Raising head up or down	Image Processing (Face, yaw, pitch)	Video Feed/Final Image	High

Table 91: Descriptive Statistics for Total Number of Errors and Time to Capture for Facial Interaction

Total Number of Errors	N	Mean	Median	Minimum	Maximum
0	404	7.49s	6.21s	4.85s	55.65s
1	168	8.00s	6.24s	5.18s	60.32s
2	20	8.49s	6.21s	5.20s	40.10s
3	7	8.55s	6.22s	5.95s	18.17s

From looking briefly at Table 91 above, a larger number of errors seemed to have a small effect on time to capture. Table 92 breaks down the totals the number of errors based on the type of information shown.

A two-way ANOVA was performed to test for an interaction effect between the total number of errors and type of information on time is taken to capture. A statistically significant interaction was found between type of information and the number of errors on time taken to capture, F (7, 561) = 17.615, $p = <0.05$, partial $\eta 2 = .180$.

Table 93 below details the post hoc analysis through pairwise comparisons, revealing a statistically significant difference between the total number of errors within each information group.

Table 92: Number of Errors per Information Group for Facial Interaction

Type of Information	No of Errors	N	Percent	Average Time to Capture	st.dev
Baseline	0	215	78.75	6.24s	0.50s
	1	37	13.55	8.02s	3.52s
	2	14	5.12	13.12s	6.17s
	3	7	2.56	17.56s	8.34s
Icons	0	91	92.85	6.34s	0.48s
	1	4	4.08	12.09s	2.79s
	2	3	3.06	16.08s	5.97s
	3	0	0.00	-	-
Text	0	87	89.69	7.06s	5.37s
	1	5	5.10	12.39s	6.19s
	2	4	4.12	16.84s	6.03s
	3	1	1.03	18.24s	-
Icons & Text	0	96	88.88	6.84s	3.75s
	1	8	7.47	8.38s	2.85s
	2	3	2.80	17.89s	3.92s
	3	0	0	-	-

Table 93: Two-way ANOVA results between Number of Errors and Information Groups ('+' indicates p < .005) for Facial interaction

Information	Number of Errors			
	0	1	2	3
Baseline	+	+	+	+
Icons	+	+	-	+
Text	+	+	+	+
Icons & Text	+	-	+	+

In nearly all cases, the total number of errors had a significant impact on the time taken to capture. There was no evidence of a significant relationship between transaction time for users who made two errors for the icons group and one error for the icons and text group. Overwhelmingly, a larger number of errors did have an impact on transaction time throughout the four groups.

Identifying which of these errors occur in real time and relaying information to the user will benefit in correcting erroneous presentations and provide a form of guidance to users. As a statistically significant relationship was found between the number of errors and time on task, the next step would investigate specific errors which may have a higher impact on the time taken to capture. Establishing that there is a significant relationship then enables a full analysis of interaction to take place. For example, consider the scenario where the user is presenting their face image to a system which displays a live feed image with icons on the top. Results have indicated that there is a difference in time taken to capture between live and non-live scenarios in the first instance, showing us that live feed scenarios were slightly quicker to complete. Data has also shown that the more errors that are performed during the transaction, the higher the impact on transaction time.

6.3.2.2.3 User Satisfaction

Like Chapter 5, user satisfaction is reported on a minimal level due to the nature of the data collection.

A task evaluation was conducted to assess user satisfaction with the system. The main goal was to assess wherever the user perceived the system to be efficient. Answers were collected after all ten trials and so Figure 51 details the response to several questions directly relating to the performance of the system. Further results from the survey are discussed in Chapter 7, recommendations.

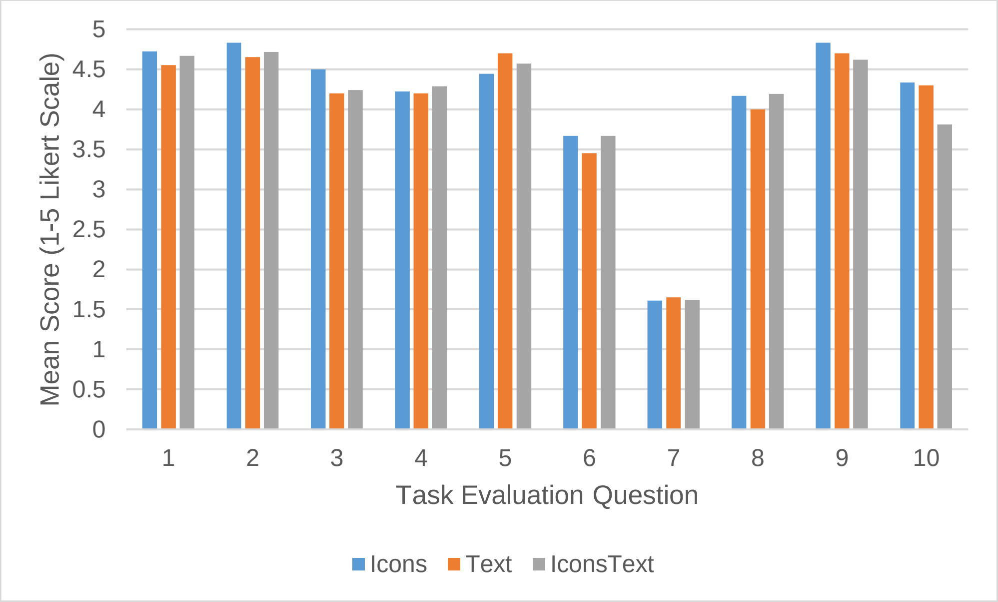

Figure 51: Mean Score for Task Evaluation Questions for Face Interaction

Results remained consistent for each information group. Users indicated they were given enough information to complete the task (Q1) and found the process relatively easy to complete (Q2) with the icon group leading with a result of 4.83 out of 5. Information provided to the user seemed to do the job, describing what to do during the process (Q4-5) although users did find some of the prompts confusing (Q6). Users did indicate that the entire process was not entirely confusing (Q7) and that capture process was clear (Q8) and very clear that the process had completed (Q9). Question Ten, although not directly relating to user satisfaction, did ask if the user would use ABC systems in the future based on this system, only the Icons & Text group did not state that they strongly agreed when compared to the other two groups.

Although this information is useful, the three groups did perform a total of ten scenarios, which combined baseline scenarios in addition to scenarios from the adaptive groups. Therefore, the results are not an indicator on the scenarios themselves but rather the general overall system that users interacted with.

However, after completing each scenario, the user was asked a series of questions relating to the information they had just seen. Due to time constraints, answers were recorded as a yes or a no. Figure 52 below details the percent of each group that answered yes to each question. The Questions are previously defined in Section 6.2.3.

Throughout all scenarios users indicated that it was often clear when the process began (83.19%) and when the process was over (86.59%). For many, it was unclear what the result of the system was (49.40%), but this was likely because the system did not have physical gate opening to indicate success. Information displayed in stand-alone kiosks must make the result clear and visible, so travellers are not confused.

Many were clear about the capture process. Most users knew when it was clear to the face the camera (78.27%) and to position themselves accordingly (66.84%). Scenarios differed greatly in responses on how to position themselves to the camera, with an expected lower result for the text-based scenarios. Language-based scenarios are not recommended for border control scenarios, but several systems do use some text (such as the eChannels in Hong Kong). Baseline scenarios also performed poorly, indicating that users may not understand fully how to present themselves to the camera in systems that relay simple information. Information with icons improved, however, reporting up to 89.5% of satisfied users. Overall, it seemed that users reported a higher level of clarity when engaging with the system that used enhanced icons to adapt feedback during the presentation.

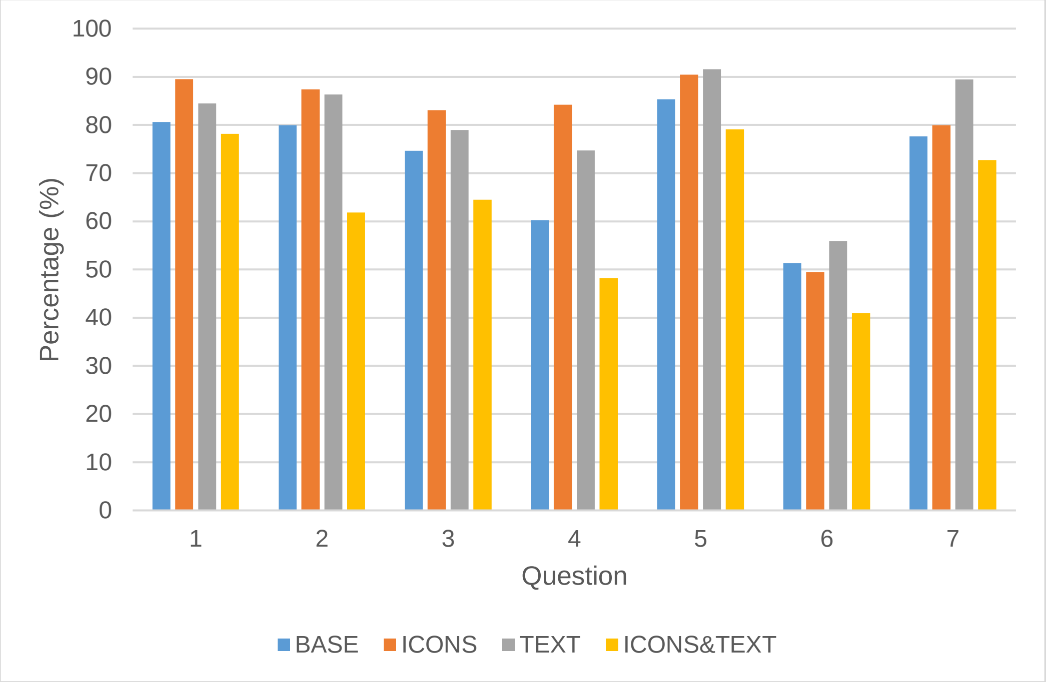

Figure 52: Percentage of Users who answered ‘Yes’ to Task Evaluation Questions in Face Interaction System

6.3.2.2.4 Sample Quality

Separate from traditional usability measurements, sample quality relates to the image captured during the transaction. This section investigates:

- Compliant Images
- Identification Matching

Images were only captured once they met the Kinect’s specifications as part of the system’s objective was to test the ability to capture images based on skeletal tracking and image processing. The Kinect program did not verify users against an enrolled image during the transaction, however during post-processing images were assessed based on the ISO full frontal standard and compared to the enrolled image and given a matching score.

The Aware PreFace program was used to determine if an image met ISO Full Front image standards. Successful images are marked as compliant and non-successful as noncompliant. Figure 53 below details the frequency of images captured across all scenarios and wherever they were ISO compliant.

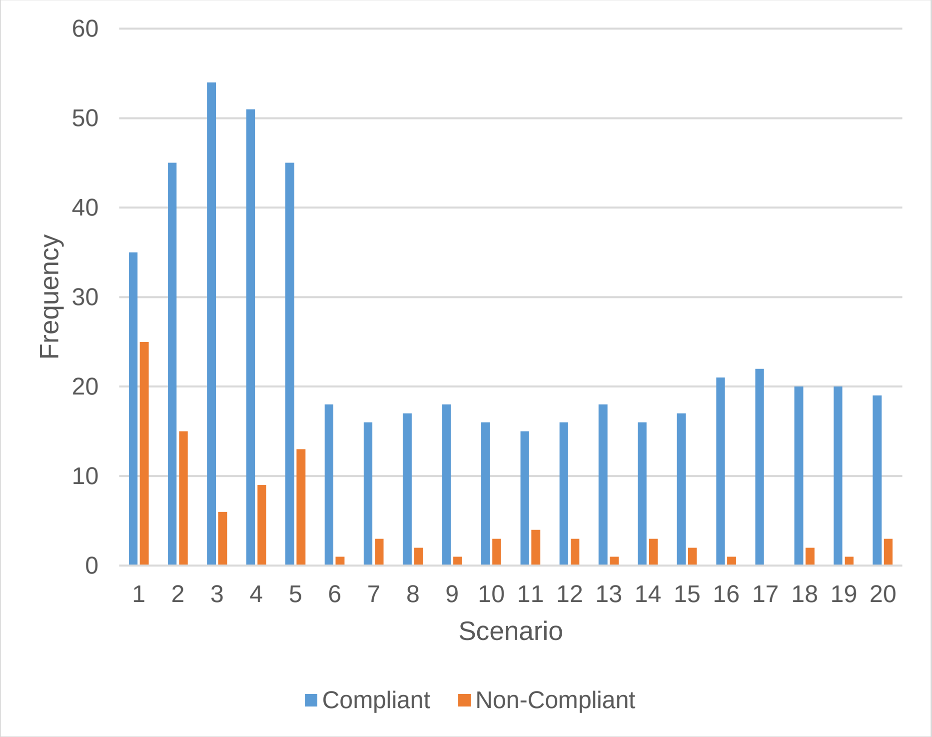

Figure 53: Frequency of Compliant and Non-Compliant Images for all scenarios in Facial Interaction

As it can be seen in Figure 53, many images conformed to ISO full frontal standards across all scenarios. BASE01 and BASE02 had a lot of non-compliant images with 35 complying with the standard and 25 that did not. For Scenario 17, Text and Icons 02, there were no reported non-compliant images. Grouping the images to information only, Table 95 below displays the frequency of compliant and non-compliant images against information shown.

Table 94: Number of Compliant and Non-ISO compliant images for Information in the Facial Interaction System

Compliance	Baseline	Icons	Text	Icons&Text
Non-ISO	41	11	6	17
ISO	248	90	93	93
Total	289	101	99	110

Furthermore, compliant and non-compliant images are compared against the behaviour code determined at the end of Task 2 by the Kinect sensor. Table 96 reports.

Table 95: Compliant and Non-Compliant Images based on Task 2 Presentation

	GFACE01	GFACE02	OFACE01	RFACE01	RFACE02
Non-ISO	74	6	9	12	2
ISO	449	28	8	8	3
Total	523	34	17	20	5

A total of 80 images (13.42% of all images) were analysed as non-compliant ISO images. These images, during the interaction process, were determined by the system as presenting a ‘good’ facial presentation with little to no errors. Several users who provided an incorrect presentation were accepted as an ISO-compliant image. This could be because a system takes several images during capture and the user-adjusted their presentation during the process.

A Chi-Square test revealed that there was a statistically significant association between the classifications for the face presentations (Task 2) and wherever an image was ISO compliant or not (Table 95). The test revealed a significant result $\chi 2$ (4) = 25.403, $p < .001$. Considering this result further, the focus was to investigate the difference between groups.

A Chi-Square test was conducted to see if there was a statistically significant association between the types of information received on compliant images. The baseline group was not included in this test as it would violate the assumption of independence of observations. The test revealed a significant result $\chi 2$ (2) = 13.064, $p = .001$.

Post hoc analysis involved pairwise comparisons using the Z-test of two proportions with a Bonferroni correction. The proportion of users who used information with icons only was statistically significantly higher than other groups, $p < .05$. The proportion of ISO compliant images for users in the text and icons and text group was not statistically significant differences, $p < .005$.

Matching Score was determined by VeriLook 5.1 SDK provided by Neurotechnology [166]. The Face Matcher performs facial template matching in 1:1 verification, matching the enrolled image against the captured images. Images must be near frontal face standards, with a rotation deviation of up to 15 degrees in any direction, meaning erroneous images captured were matched against the enrolled image. The matching threshold is linked to FRR, the higher the threshold, the lower is FAR and higher FRR. See Section 5.2.4 for the link between thresholds and FAR.

Looking at matching score between the enrolled and captured image, Table 96 reports on the scores across scenario.

Table 96: Descriptive Statistics for Matching Score for Information Groups in Face Interaction

Information	Mean	Median	st.dev
Baseline	61.63	63	8.45
Icons	71.36	77	14.78
Icons & Text	60.89	59	10.08
Text	64.18	68	12.23

Again, Icons performed slightly better than the other groups with a lead of an average score of 71.36, demonstrating that icons provided higher quality images in both matching score and the number of images conforming to standards.

A Kruskal-Wallis test was conducted between the three main groups to see which had a statistically significant effect on the matching score. There were no outliers in this data, as assessed by inspection of a boxplot. Median scores were statistically significantly different between groups, $\chi2$ (2) = 34.75, p = <.001.

Pairwise comparisons were performed using Dunn's 1964 procedure with a Bonferroni correction for multiple comparisons. Adjusted p-values were presented. The post hoc analysis revealed statistically significant different scores between comparisons between icons and the other two groups ($p < .001$). There was no significant relationship between text and text and icons.

Looking at the Icon group, Table 98 below details the descriptive statistics between the individual scenarios within the scenario, investigating if there were any significant differences between the five scenarios.

Table 97: Descriptive Statistics for Matching Score for Icon Scenarios in Face Interaction

Information	Mean	Median	st.dev
ICONS01	72.44	77	13.85
ICONS02	72.16	76	15.98
ICONS03	73.00	79	16.49
ICONS04	67.94	68	14.02
ICONS05	70.78	77	14.60

Matching score between the scenarios was very similar, with only ICONS04 performing slightly with a score of 67.94.

6.4 Data Analysis

While Performance Analysis seeks to investigate the relationship between variables obtained only through results HBSI and Kinect methods, Data Analysis explores the relationship between several variables further, seeking insight to questions that have yet to be explored in recent research.

Question 1

Wherever a user has previously used an ABC system that used the face modality was also considered against the impact on performance. It is theorised that increasing the use of a system will increase the overall rate of performance on subsequent systems. However, this is not necessarily true within border control systems. Systems change by design, requirements and through information displayed between different vendors and countries.

A two-way ANOVA was conducted to examine the effects of previous use with ABC systems and type of group on time is taken to capture. Outliers were assessed by inspection of a box plot, and normality was assessed using Shapiro-Wilk's normality test for each cell of design and homogeneity of variances was assessed by Levene’s Test. There were no outliers, residuals were normally distributed ($p > 0.5$), and there was the homogeneity of variances ($p = .072$).

The interaction effect was not a statistically significant, $F(2, 292) = .662$, $p = .517$, partial $\eta 2 = .004$. Pairwise comparisons were run, and p-values were Bonferroni-adjusted, again, there were no statistically significant differences found.

Question 2

An important variable in ABC systems that may impact the interaction is the subject height. Some systems can adapt to the user's height by automatically moving the camera to the eye level of the user. In some cases, such as the eGate scenarios in Heathrow, the monitor will also travel with the camera to match the user's height.

For this data collection, height is considered against performance through assessing a relationship between sample quality and time spent on the task. Users are grouped based on their height. Table 98 reports the average time to capture for each user across all scenarios.

Table 98: User Height and Descriptive Statistics for Time to Capture in Face Interaction

User Height	N	Mean	Median	Minimum	Maximum
150-155	10	7.86s	6.24s	5.66s	6.45s
155-160	90	7.91s	6.22s	5.18s	8.43s
161-165	110	8.79s	6.20s	5.12s	7.25s
166-170	100	7.68s	6.29s	5.17s	5.15s
171-175	69	7.81s	6.21s	5.16s	8.11s
176-180	80	7.19s	6.23s	5.51s	12.38s
181-185	60	7.18s	6.21s	3.44s	9.15s
186-190	40	6.34s	6.20s	3.44s	8.60s
191-195	40	7.03s	6.27s	3.58s	7.25s

To determine wherever that was an impact on the time taken to capture based on height, gender, type of information shown (baseline, icons etc.) and group. A multiple regression analysis was performed. See Table 99 below.

Table 99: Summary of Multiple Regression Analysis for Time to Capture in Face Interaction (* p <.05; β = unstandardised regression coefficient, SE = standardised error of the coefficient; B = standardised coefficient)

	β	SE	B
Intercept	6.753	5.11	
Gender	4.15	.20	0.48
Height	-.005	.049	-.005
Type	.042	.070	.025
Group	-.038	.105	-.015

Table 99 reports on the multiple regression results. All four variables were not found to impact time on task significantly.

6.5 Summary

In this chapter, an adaptive facial interaction system has been explored and tested with over sixty users. The participants were recruited mostly consisted of students from the university, but most had indicated that they had previous experience of using biometric systems such as ABC gates. Each user was requested to complete ten scenarios, always forming of at least five ‘baseline’ scenarios, systems that are based on current ABC systems and five ‘dynamic’ scenarios.

Dynamic scenarios (formally split into three groups; Icons, Text and Icons and Text) analysed user interactions in real time using the Kinect Sensor. The adaptive system required users to perform a correct movement before succeeding to the next task; move to the system, look at the camera, stay still for capture and finally exit the system. If the user did not present an ideal demonstration to the camera, specific information was provided to correct the presentation.

The Kinect sensor was not only used to capture face images but to automate the HBSI process by using skeletal tracking. The results were compared against manual observations and enabled a vast range of performance metrics at the end of each transaction. The system could automate the assessment of interactions and log the results for further analysis. Improving the system performance assessment this way allows vendors to investigate a range of information such as behaviours, movements, system errors as well as traditional metrics such as biometric verification rates.

Ground truth comparisons revealed that the system had an extremely high ratio of automating the HSBI categorisation process, in most cases, resulting in over 90% of correct automated classifications. Although used in a controlled environment, testing self-service biometric systems before deployment in this way may enable a deeper revelation into the flaws of a system.

In general, the icon group performed slightly better than the other information groups provided. Users who were presented adaptive icons provided a higher throughput result (6.92s) and were more likely to provide an ISO compliant image (91%) and a higher matching score against the enrolled image (71.36 at >0.001% FAR). Statistical tests revealed in many cases that there were significant differences between the information groups and where the placement of guidance was given.

In summary, this chapter provided the following:

- A report on the investigation into the effect of information displayed on the screen, looking at the impact of different sizes and layouts of icons, language and a combination of icons and text
- A performance assessment based on HBSI with an interest in throughput and sample quality, detailing how the inclusion of the Kinect sensor can enable a deeper picture of user movements and interactions
- Combining the HBSI framework with Facial Interaction capture to enable automated categorisations of an interaction, using face tracking ability of the Kinect sensor to detect, categorise and capture a user’s face based on specific requirements

CHAPTER 7. RECOMMENDATIONS AND CONCLUSIONS

7.1 Introduction

This section reports on the recommendations and findings based on the research presented in this thesis. The analysis conducted throughout this work has highlighted several novel approaches to the assessment and implementation of biometric systems. A highlight of the work was the introduction an automated method of assessing user interaction errors within self-service biometric systems, harnessing skeletal tracking methods to detect and analyse a user's position.

In this chapter, the key findings and contribution of the research are reported in a summary of the work covered in this thesis. This will be followed by a discussion of the significant research findings of this work, making recommendations for both biometric and border control processes. Also, the future of border control systems is discussed, ending with a summary and a closing statement.

7.2 Key Findings and Contributions

Airports across the globe are progressively installing ABC systems to improve security, streamline the travelling process and working towards facilitating a better passenger experience. In this work, performance assessment through user interaction was studied in two standard biometric modalities utilised in ABC systems. Each study attempted to replicate an automated eGate system, a solution to automatically verify travellers identify and allow border crossing.

Biometric performance is traditionally assessed through two common areas; error and throughput metrics. Standard measures such as FTA, FTE and FRR, can sometimes mask the true reason behind why an error occurred. Harnessing the HBSI method allows for a full categorisation of a range of metrics that will benefit in analysing system performance. Evaluating user behaviour for each task and mapping out all possible scenarios within the system will be crucial to configuring and adapting system responses. By applying the HBSI framework, this may lead to enhancing areas of system performance, helping to reduce errors and enhance overall usability of ABC systems for travellers worldwide.

The Generic Model can be used to standardise the mapping of ABC configurations to identify where variations and similarities lie. Having defined both the enrolment and verification stages of any given system through distinguishing components used, individual interactions with each sensor can be analysed using the HBSI method. Looking at the verification stage of the GM, a breakdown of tasks from both a user and system point of view are reported in Identity Claim Process model.

ICP demonstrates a framework for each step of the verification stage and identifies user and system responsibilities. The methodology can be used to identify conditions where the HBSI Presentation Framework can be implemented or adapted to specific situations and sensors.

For example, a system categorisation would be in effect in conventional assessment metrics for users who may present an incorrect finger for fingerprint verification. The ideal system output would be 'Biometric Not Presented' whereas if the system concluded a 'True Match', then there could be a potential security risk in play. Current performance assessment on biometric modalities would not distinguish a system error from a user-generated one. The inclusion of HBSI can help to establish cases where the system was correct in the identification of the scenario as erroneous (False Interaction) or that the user incorrectly presented a biometric sample. If a HBSI categorisation could be automated during this process, the likelihood of a security risk would be reduced as the incorrect presentation would be detected before the system handled the sample.

Human Activity Recognition is a large, growing field that aims to study tools to assess and classify movements for many scenarios. The Kinect V2 sensor is a tool used for tracking the user using skeletal tracking. The device is being adopted quickly by the Activity Recognition community, using the sensors depth component for a broad range of activities. The device reports on 25 joint locations in 3D space, providing the opportunity to track user movements within a certain range of the sensor. Research in the area has already established the Kinect's ability to provide accurate results but has not yet been adopted for assessing and assisting user interactions within a biometric system scenario.

Chapter 4 introduced three tasks to test the Kinect's ability to measure and follow a user's movement, assessing if the outcome would enable the applicability of the device to a full biometric interaction. Data indicated that the Kinect was highly accurate in measuring user's height and the distance between shoulder blades.

Movements were precisely identified and reported in relation to each of the five users. Again, the sensor was able to fully track the user within the controlled environment. In addition to a report on the findings, a wide range of literature using the Kinect was also introduced earlier in the chapter. Research indicates the accuracy of the different sensors within the Kinect camera.

Implementing the Kinect camera or a similar sensor could enable automated assessment of traveller interactions. Introducing methods to assess the user through the introduction of new tools such as face tracking or image-processing elements will provide many benefits. For instance, for the tired or stressed traveller, introducing an automated feedback system to relay information to the user on how to correct their presentation, e.g. look up, open eyes will begin the first steps into offsetting incorrect behaviours. The information can only be presented if the sensor was not able to detect certain features or joints within a required location. Further work will be needed to identify common presentations and appropriate methods in responding to users making incorrect interactions.

The two data collections in this study evaluated the sensors ability to track a user within a controlled, self-service environment. Chapter 5, Fingerprint Interaction, introduced a program to automate the process of classifying HBSI in real time, analysing an incorrect or correct presentation based on simple joint tracking locations. For example, if the user was standing in the correct place and their right hand/wrist joint was in the required zone, the sensor would determine that a correct presentation had taken place. The procedure has only enabled the first step of identifying actions within this type of scenario. This initial activity was somewhat crude, as classifying this interaction does not reveal detail in the finer fingerprint movements. However, the results indicate that the system may be better suited for face recognition systems.

On Fingerprint Interaction, the results summarised from the study confirm findings from previous reports conducted by the NIST group. There was little effect on performance, through sample quality and throughput, based on user’s height, the sensor used and angle of the slope. Several key relationships were found for a swipe sensor, but unfortunately, the swipe-based sensor has little applicability in ABC systems.

The additional benefit of studying fingerprint interaction and the various variables that were introduced into the study was to not only investigate the impact on performance but also to display how the Kinect sensor can reveal anomalies in the user's interaction. By cross-referencing performance data with skeletal information, the performance assessment provides a much deeper insight into the user's interaction. Facial Interaction improves the automation process further, combining biometric capture, processing and evaluation with skeletal tracking elements. The HBSI automation system utilises the face-tracking package from the Kinect SDK to provide real-time information on the face by; reporting on both face pose and expression. The system was better suited for facial interaction, as the location and position of the user would be more important for these scenarios.

The study introduced three different types of information through feedback and instruction to the user; simple pictorial icons, language (text) and a combination of the two. All icons were based on pictorial instructions that have been previously defined by various ISO working groups. Language, although not commonly employed in these types of systems, sometimes do appear and may throw the user off the process. A combination of the two is rare, but the goal of the study was to assess if any information yielded better performance results than the others.

In conclusion, this research may contribute to an improvement in the detail and accuracy of reporting of system performance in ABC systems. The application of the HBSI framework will allow a range of metrics, defining a set of interaction measurements which must be a priority (while adhering to the systems intended use) in the design and implementation of these public systems. Reporting on the six HBSI presentation metrics will allow a deeper understanding of where problems lie within a system. The models proposed will enable the breakdown of the process so that each stage can be assessed beyond the traditional reporting of a system level error. In defining a process map, user and system handlings are measured at each key component.

The main contributions in this thesis are the following:

- A study and review of the literature in Biometrics, HBSI and Automated Border Control systems. Additionally the full HBSI model is proposed, considering metrics beyond the single traditional human-sensor interaction.
- The design and production of two models that can be used to identify and evaluate ABC systems. The Generic Model (ABC)

identifies a general enrolment and verification process where the HBSI evaluation framework can be applied throughout significant steps of the process. Building on this, the ICP framework decomposes the verification stages of the Generic Model (ABC) to highlight which steps can be attributed to a user and/or system based input or response. The Behaviour Framework identifies several key actions that can be taken at specific user-interaction steps throughout the verification process

- The development of a tracking system to improve performance assessment and assist with user interaction errors in both facial and fingerprint systems. A study using the Kinect device is documented to establish the stability and reliability of the data returned by the sensor for use within a self-service scenario.
- Analysis of two ABC systems based on common biometric modalities found in border control processes.
- A report on the recommendations to improve current processes, exploring the results obtained from the fingerprint and facial chapters and suggesting suitable considerations for future design and implementation of ABC systems.

7.3 Limitations & Recommendations

Based on the findings and contributions presented in this study, the recommendations are as follows:

- Implementing the HBSI Framework will enable vendors to provide a methodology for assessing system performance by considering metrics beyond traditional error ratings. The HBSI model is continually updating and expanding; recent versions have been introduced to include the categorisation of token presentations, looking at false claims, attacks and the full interaction beyond a single modality interaction
- Introduce an additional sensor such as an RGB-D camera that can track the user’s skeleton within an ABC scenario. The results may be beneficial in ensuring that the correct position, posture and movements of the user are detected before starting a user-based task. The main advantage of implementing this system is it would enable the ability to relay dynamic feedback that can tailor instructions to the user based on their presentation

- Tackling standardisation, the Generic Model and Identity Claim Process enable process flows that outline common themes between scenarios. Identifying user and system processes early on within the design stage of an ABC solution will allow implementations to assess the performance of a transaction beyond individual elements of the system
- Improve process flow for travellers; simplifying information will make progress efficiency and effectiveness. The results from Facial Interaction suggest that systems that present dynamic feedback using icons will yield a higher rate of performance over current versions of the system

In addition to the main conclusions, participants were presented with several questionnaires throughout the studies that were intended to identify their knowledge on border control and biometric solutions. Out of 80 users, 69 stated that they followed the latest technological updates and understood the basic idea of a biometric system. Sixty percent of users had used a form of biometric verification processes for everyday purposes such as interacting with their phone or laptop. Seventy percent of the users had used an ABC system in the last year, using an eGate type system within the last year, 20% of those used a system that required an eMRTD or national ID card instead of a passport. All users indicated that they had used a facial interaction type system.

The Task evaluation surveys assessed the capability of the tested system through providing clear instructions (as discussed in Section 5.3.2 and 6.3.2). The questionnaire also gave the opportunity for participants to share their thoughts as well as their recommendations based on their interactions.

Out of eighty participants, at least 70% of users identified that, at some level, they did not trust the storage and use of biometrics, 15% suggested that they do not trust the storage of their data within government systems. Several users suggested that systems should detail a higher level of transparency with how biometric systems use and store their sensitive data. 85% of users did not know that their fingerprints could be kept on an electronic passport whereas 50% of users understood how the facial image captured in an ABC system was compared against their passport photo. Several comments suggested that they thought the UK Border Control held a national database of travellers with their details and photos on. Just 25% of the participants understood how biometric systems work, through capture, processing and matching against a previously captured sample.

45% of participants did, however, agree that use of biometrics provided a speedier process of crossing the border.

One of the most common items of feedback from participants refers to the use of similar systems and comparing the process between the study and real implementations. 82% of subjects suggested that while the system was alike in many ways, the lack of feedback for facial interaction on some of the baseline scenarios was consistent with their own experiences. The common conception from users was that during the process they did not know if they performed the interaction correctly or incorrectly, and in some cases reported that during the live scenario, some processes during the interaction would fail without providing any feedback. One of the most certain recommendations then is to make the information available simple, informative and above all, transparent in the process it is showing. Further work with the feedback displayed within any system should consider the user's interaction and their response to processes that may fail. Also, providing feedback through assessing a 'good' interactions and providing a result to the user may increase the likelihood of desired behaviours. This may cause other travellers who are in the queue to repeat the actions, as it is often the case that those who are waiting to use the system will tend to repeat the actions of those who went before them.

Further work will be required to attribute HBSI to the use of token presentations and other processes such as a user entering or exiting the system. A limitation of this study was not using the Kinect to assess a token interaction or evaluate the full interaction, considering the combination of multiple presentations using a token and a biometric. The design of travel documents through its dimensions, materials and quality of electronic and data components changes from country to country and will inevitably have an impact on the speed and accuracy of the token reading process. Passport readers are not readily available and if purchased from a supplier, are extremely expensive. There are several programs available to interact with basic features of a passport but firstly, the device typically will interact with the MRZ field only (ABC systems read the RFID chip), and due to the security features contained within, it would be difficult to combine the process with a biometric matching algorithm. One idea was to recreate the dimensions of a token reader and to produce a 'mock-up' of a passport, to enable a study to collect Kinect and HBSI data. However, the conditions in which a token is presented to a system would be difficult to replicate, as the design of any travel document needs to be carefully considered. In addition, in live scenarios, the token needs to be held for several seconds to a sensor while internal processes take place which would not be possible to replicate to a full extent in a research environment. Gschwandtner et al. [170] report further on the difficulty of simulating the piece of hardware.

Further data will be required from a wider range of participants and live implementations to validate the contributions of this work further. Also, more work is needed in categorising user behaviours and the effects these have on the system. In any case, when performing data collections that aim to replicate a border control scenario, care must also be taken to make sure that certain conditions that are prevalent in a live scenario are replicated as closely as possible. In a controlled environment, non-technical factors such as stress or tiredness will not be able to be easily replicated, and thus, the significance of these factors will be difficult to assess.

7.4 The Future of Border Control Processes

Border control solutions have changed dramatically throughout the last decade with thanks to the introduction of the ePassport and improvements in capture technology. Looking at the verification process, the future of border control systems looks promising in making the traveller experience more intuitive and user-friendly.

Considering recent news, a new trend in the implementation of systems is using biometrics on-the-move, facilitating a non-stop and contactless verification process for travellers [93] [171]. Recently unveiled, the systems which are described as a “new multi-context facial recognition technology”, can be used for a range of applications in the airport: check-in, boarding and at border control.

Research with contactless fingerprint scanners is becoming increasingly popular due to the development in the imaging technology field. Introducing methods to capture fingerprints without touching the sensor introduces some benefits such as distortion-free fingerprint acquisition (no pressure needed) and free from hygienic issues. Contactless technology is somewhat limited by the environment they operate in. If the sensor is capturing multiple fingerprints, then the sensor must be able to scale and capture fingerprints from subjects who may present at different distances. The sensors ability to capture at various ranges is limited, and so there are likely decreases in the capacity to focus and capture a high-quality image. Another major implication will be how the users interact with the sensor, users by instinct may touch the surface regardless, and so the surface will need to be regularly cleared to ensure there is no interference with the sensors ability to capture. Attention will be required to make sure that users are educated to use the system property. The limitations may outweigh the benefits, especially when considering an ABC scenario. Contactless sensors are more apt to appear in semi-automated border control solutions where a border officer may be present and may perhaps to be used for enrolment purposes only.

The ABC process is complex, and many factors contribute to the performance. Some uncontrollable factors such as social, psychological and ethnic factors will with no doubt, be one area which systems should be targeting to attempt to at least monitor and oversee in the future. Accounting for these non-technical factors may increase system performance, but to do so will require special attention to the design and development of both the build and user interface implemented into a system.

There are three key components to systems that should always be considered; the user, the environment and the system itself. The system should consider both the environment it operates within and the profile of the individual traveller. The passenger's experience will determine the longevity of a system and therefore the overall operation. The traveller’s demographic information will influence their ability to use a given system. For example, elderly users may be more likely to avoid using a system due to impairments. Users who are likely to have a bad first impression of a system are unlikely to use the system again and may prefer manual control methods. These are just some of the many challenges that the ABC system faces.

The design and build of the system should then consider the traveller's previous experience, training and attitude towards biometric systems. A first step should be to improve the information available in both before and during the interaction process. Information before queuing for the system will increase the general awareness of the process while displaying clear and precise information during the interaction will guide users to a successful transaction. The system status should be visible, and indications of progress should be given to enhancing the user’s experience. The research proposed in this thesis, therefore, will provide ABC systems with a foundation to improve the process flow and guidance provided throughout the transaction process.

REFERENCES

[1] A. Leniski, R. Skinner, S. McGann and S. Elliott, “Securing the Biometric Model,” in *Proc. IEEE 37th Annual Security Technology International Carnahan Conference*, 2004.

[2] A. Jain, A. Ross and S. Prabhakar, “An Introduction to Biometric Recognition,” *IEEE Transactions on Circuits and Systems for Video Technology,* Vol. 14, No. 1, pp. 4-20, 2004.

[3] Information technology -- ISO/IEC 19795-1:2006 Biometric performance testing and reporting -- Part 1: Principles and framework”, “International Organization for Standardization, 2003.

[4] K. Singla and S. Kumar, “A Review of Data Acquisition and Difficulties in Sensor Module of Biometric Systems,” *Songklankarin Journal of Science and Technology*, Vol. 5, No. 35, pp. 589-597, 2013.

[5] Y. Choong and M. Theofanos, “Ten-print fingerprint self-captures: Graphics-only user guidance without language,” in *Proc. IEEE International Joint Conference on Biometrics (IJCB)*, 2014.

[6] B. Fernandez-Saavedra, R. Alonso-Moreno, A. Mendaza-Ormaza and R. Sanchez-Reillo, “Usability Evaluation of Fingerprint-Based Access Control Systems,” in *Proc. Sixth International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP*), Darmstadt, 2010.

[7] Frontex Europa, “BIOPASS II Automated Biometric Border Crossing Systems Based on electronic passports and facial recognition,” Frontex, Warsaw, 2010.

[8] C. Riley, G. Johnson, H. McCracken and A. Al-Saffar, “Instruction, Feedback and Biometrics: The User Interface for Fingerprint Authentication Systems,” in *Proc. Human-Computer Interaction INTERACT*, 2009.

[9] K. Lai, S. Samoli and S. Yanushkevich, “Multi-spectral facial biometrics in access control in *Proc. IEEE Symposium on Computational Intelligence in Biometrics and Identity Management*, Orlando, USA 2016.

[10] R. Donida-Labati, A. Genovese, E. Muñoz, V. Piuri, F. Scott and G. Sforza, “Emerging biometric technologies for Automated Border Control gates,” in *Proc. 13th Int. Conf. on Pattern Recognition and Information Processing (PRIP 2016)*, Minsk, Belarus, 2016.

[11] Y. Yongzhi, Z. Peng and X. Yun, “Automatic User Interface Generating for Simple Interaction in Pervasive Computing,” in *Proc. IEEE International Conference on Computational Science and Engineering (CSE) and Embedded and Ubiquitous Computing (EUC)*, Guangzhou, China, 2017.

[12] The Encyclopaedia of Human-Computer Interaction, “The Interaction Design Foundation,” [Online]. Available: https://www.interaction-design.org/literature/book/the-encyclopedia-of-human-computer-interaction-2nd-ed/human-computer-interaction-brief-intro. [Accessed September 2017].

[13] J. Valacich, R. Wright and A. Dimoka, “Introduction to Human-Computer Interaction (HCI) Minitrack: Informing Design Choices Utilizing Behavioural, Neurophysiological, and Design Science Methods,” in *Proc. 46th Hawaii International Conference on System Sciences (HICSS),* 2013, Wailea, Mau, USA, 2013.

[14] K. Saeed, *New Directions in Behavioural Biometrics*, CRC Press, 2016.

[15] R. Yampolskiy and V. Govidaraju, “Direct and Indirect Human-Computer Interaction Based Biometrics,” *Journal of Computers*, Vol. 2, No. 10, pp. 76-88, 2007.

[16] B. Fernandez-Saavedra, “Evaluation Methodologies for Security Testing of Biometric Systems beyond Technological Evaluation”. Ph.D dissertation, University of Universidad Carlos III de Madrid, Madrid, Spain, 2013.

[17] E. Kukula, “Framework for Human, System, and Administrative Errors in Biometric Systems,” in *INCITS*, Washington, DC, 2009.

[18] E. Kukula and S. Elliott, “A Definitional Framework for the Human-Biometric Sensor Interaction Model,” in *Proc. SPIE Defence, Security and Sensing,* Orlando, FL, 2010.

[19] E. Kukula, S. Elliot, M. Sutton, N. Latif and V. Duffy, “Design and Evaluation of the Human-Biometric Sensor Interaction,” in *Proc. 9th Annual Information Security Symposium*, West Lafayette, Indiana, 2008.

[20] E. Kukula and S. Elliott, “Critical Anthropometric and Ergonomic Elements for Reliable Hand Placement in Hand Geometry Based Authentication System,” 2006. [Online]. Available: https://www.slideshare.net/bspalabs/2006-hbsi-and-hand-geometry. [Accessed September 2017].

[21] E. Kukula and S. Elliott, “Implementation of hand geometry,” *IEEE Aerosp. Electron. Syst. Mag*, Vol. 21, No. 3, 2006.

[22] E. Kukula, C. Blomeke, S. Modi and S. Elliott, “Effect of human-biometric sensor interaction on fingerprint matching performance, image quality and minutiae count,” *International Journal of Computer Applications in Technology*, Vol. 34, No. 4, 2009.

[23] E. Kukula, M. Sutton and S. Elliott, “The Human–Biometric-Sensor Interaction Evaluation Method: Biometric Performance and Usability Measurements,” in *IEEE Trans. Instrum. Meas,* Vol. 59, No. 4, pp 784-791, 2010.

[24] S. Elliott, “Accuracy, Throughput and Usability of an Aoptix InSight Iris Recognition System,” *Biometrics Consortium*, 2010.

[25] M. Brockly, R. Guest, S. Elliott and J. Scott, “Dynamic Signature Verification and the Human Biometric Sensor Interaction Model,” in *Proc. IEEE International Carnahan Conference on Security Technology*, Barcelona, 2011.

[26] S. Elliott, “Human Biometric Sensor Interaction: Latest Research and Process Model,” [Online]. Available: https://www.nist.gov/sites/default/files/documents/2016/12/06/08_elliott_nist_presentation.pdf. [Accessed September 2017].

[27] S. Elliott, “Evolution of the HBSI Model,” West Lafayette, Indiana, 2012.

[28] E. Kukula, “Design and Evaluation of the Human-biometric Sensor Interaction,” Ph.D dissertation, West Lafayette, Indiana, 2008.

[29] J. Chen, Y. Deng, G. Bai and G. Su, “Face Image Quality Assessment Based on Learning to Rank,” *IEEE Signal Processing Letters*, Vol. 2, No. 1, pp. 90-94, 2015.

[30] X. Gao, S. Li, R. Liu and P. Zhang, “Standardization of Face Image Sample Quality,” in *Proc. Second International Conference on Biometrics (ICB)*, Seoul, Korea, 2007.

[31] L. Spreeuwers, A. Hendrikse and K. Gerritsen, “Evaluation of automatic face recognition for automatic border control on actual data recorded of travellers at Schiphol Airport” in *Biometrics Special Interest Group (BIOSIG)*, 2012.

[32] R. Raghavendra and C. Busch, “Improved face recognition by combining information from multiple cameras in Automatic Border Control system,” in *Proc. 12th IEEE International Conference Advanced Video and Signal Based Surveillance (AVSS),* 2015.

[33] M. Kosmerlj, T. Fladsrud, E. Hjelmas and E. Snekkenes, “Face recognition issues in a border control environment,” in *Proc. International Conference on Advances in Biometrics*, Hong Kong, China, 2006.

[34] International Organization for Standardization, ISO/IEC, “ISO/IEC 19794-5:2011 Information Technology -- Biometric Data Interchange Formats -- Part 5: Face Image Data, 2011.

[35] J. Sang, Z. Lei and L. Stan, “Face Image Quality Evaluation for ISO/IEC Standards 19794-5 and 29794-5,” in *Proc. Advances in Biometrics*, 2009.

[36] A. Optiz and A. Zabini, “Evaluation of face recognition technologies for identity verification in an eGate based on operational data of an airport,” in *Proc. IEEE International 12th Conference on Advanced Video and Signal Based Surveillance (AVSS)*, August 2015.

[37] International Organization for Standardization, ISO/IEC, “Information technology — Biometric sample quality — Part 5: Face image data”, 2010.

[38] M. Carmen, L. Toscano-Medina, G. Aguilar-Torres and G. Sanchez-Perez, “Fingerprint Scanners Comparative Analysis Based on International Biometric,” in *Proc. International Conference on Internet Monitoring and Protection (ICIMP)*, 2012.

[39] International Organization for Standardization. ISO/IEC, “ISO/IEC 19794-1 2011 Information Technology -- Biometric data interchange formats -- Part 1: Framework”, 2011.

[40] International Organization for Standardization. ISO/IEC, "ISO/IEC 19794-2:2011 Information technology -- Biometric data interchange formats -- Part 2: Finger minutiae data", 2011.

[41] International Organization for Standardization. ISO/IEC, "ISO/IEC 19794-3:2006 Information Technology -- Biometric data interchange formats -- Part 3: Finger pattern spectral data,", 2006.

[42] International Organization for Standardization. ISO/IEC, "ISO/IEC 19794-4:2011 Information Technology -- Biometric Data Interchange Formats -- Part 4: Finger Image Data,", 2011.

[43] A. Bazin and T. Mansfield, "An Investigation of Minutiae Template Interoperability," in *Proc. IEEE Workshop on Automatic Identification Advanced Technologies*, Alghero, Italy, 2007.

[44] O. Bausiner and E. Tabassi, "Fingerprint Sample Quality Metric NFIQ 2.0," in *Proc. Special Interest Group on Biometrics and Electronic Signatures*, Darmstadt, Germany, 2010.

[45] N. Abdal-Ghafour, A. Abdel-Hamid, M. Nasir and S. Khamis, "Authentication enhancement techniques for BAC in 2G E-passport", in *Proc. Innovations in Information Technology (ITT)*, 2016.

[46] Z. Yao, J. Bars, C. Charrier and C. Roenberger, "A Literature Review of Fingerprint Quality Assessment and Its Evaluation," *IET Biometrics*, Vol. 5, No. 3, pp. 243-251, 2016.

[47] J. Ryu, J. Jang and H. Kim, "Analysis of Effect of Fingerprint," in *Proc. NIST Biometric Quality Workshop II*, 2007.

[48] C. Jin, H. Kim, X. Cui, E. Park, J. Kim, J. Hwang and S. Elliott, "Comparative Assessment of Fingerprint Sample Quality Measures Based on Minutiae-Based Matching Performance," in *Proc. Second International Symposium on Electronic Commerce and Security*, Nachang City, China, 2009.

[49] R. J. Michaels, B. Stanton, M. Theofanos and S. Orandi, "A taxonomy of Definitions for Usability Studies in Biometrics," NIST Interagency/Internal Report (NISTIR) – 7378, US Department of Commerce, National Institute of Standards and Technology, 2006.

[50] M. Theofanos, R. J. Michaels and B. C. Stanton, "Biometric Systems Include Users", *IEEE Systems Journal*, Vol. 3, No. 4, pp. 461-468, 2009.

[51] International Organization for Standardization. ISO/IEC, "92411-11 Ergonomic Requirements for Office Work With Visual Display Terminals (VDTs) - Part 11 Guidance on Usability," ISO/IEC 9241-11, 1998.

[52] C.Gohringer, "The application of face recognition in airports", *Biometric Technology Today*, Vol. 2012. No. 7, pp. 5-9, 2012.

[53] Frontex Europa, "BIOPASS Study on Automated Biometric Border Crossing Systems for Registered Passengers," European Agency for the Management of Operational and Cooperation at the External Borders, Warsaw, 2007.

[54] A. Drygajlo, “Multimodal Biometrics for Identity Documents and Smart Cards: European challenge,” in *Proc. 15th European Signal Processing Conference*, Poznan, Poland, 2015.

[55] M. Theofanos, B. Stanton, C. Sheppard, R. Michaels, N. Zhan, J. Wydler, L. Nadel and W. Rubin, “Usability Testing of Height and Angles of Ten-Print Fingerprint Capture,” NIST Interagency/Internal Report (NISTIR) – 7504, NIST, 2008.

[56] S. Furman, C. Stanton, M. Theofanos, J. Libert and J. Grantham, “Contactless Fingerprint Devices Usability Test,” NIST Interagency/Internal Report (NISTIR) – 8171, NIST, 2017.

[57] Y. Y. Choony, M. Theofanos and H. Guan, “Fingerprint Self-Captures: Usability of a fingerprint system with real-time feedback,” In *Proc. IEEE Fifth International Conference on Biometrics: Theory, Applications and Systems*, 2012.

[58] M. Theofanos, B. Stanton, C. Sheppard and R. Michaels, “Usability Testing of Face Images Capture for US Ports of Entry,” in *Proc. IEEE International Conference on Biometrics: Theory, Applications and Systems*, 2008.

[59] H. Guan, M. Theofanos, Y. Choong and B. Stanton, “Real-time Feedback for Usable Fingerprint Systems,” in *Proc. IEEE Fifth International Conference on Biometrics: Theory, Applications and Systems (BTAS)*, Arlington, VA, USA, 2012.

[60] E. Schiavone, A. Ceccareli, A. Bondavalli and A. Carvalho, “Usability Assessment in a Multi-Biometric Continuous Authentication System,” in *Proc. Seventh Latin-American Symposium on Dependable Computing (LADC)*, Cali, Colombia, 2016.

[61] B. R. Sanchez-Reillo, J. Liu-Jimenez and O. Miguel-Hurtado, “Evaluation of biometric system performance in the context of Common Criteria,” *Information Sciences*, Vol. 245, No. 1, pp. 240-254, 2013.

[62] B. Fernandez-Saavedra, R. Alonso-Morendo, J. Uriart-Antoniou and R. Sanchez-Reillo, “Evaluation methodology for analyzing usability factors in biometrics,” in *Proc. International Carnahan Conference on Security Technology*, 2009.

[63] B. Stanton, M. Theofanos, S. Orandi, R. Michaels and N. Zhang, “Usability Testing of Ten-Print Fingerprint Capture,” NIST Interagency/Internal Report (NISTIR) – 7403, NIST 2007.

[64] B. Stanton, M. Theofanos, S. Furman and P. Grother, “Usability Testing of a Contactless Fingerprint Device,” NISTIR 8159, NIST, 2016.

[65] M. Theofanos, B. Stanton, C. Sheppard, R. Michaels and Y. Choong, “Assessing Face Overlay,” NIST Interagency/Internal Report (NISTIR) - 7578, NIST, 2008.

[66] M. El-Abed, R. Giot, B. Hemery and C. Rosenberger, “A study of user's acceptance and satisfaction of biometric systems,” in *IEEE International Carnahan Conference on Security Technology (ICCST)*, San Francisco, 2010.

[67] E. Kukula and S. Elliott, “Ergonomic Design for Biometric Systems,” in *Encyclopaedia of Biometrics*, Springer US, pp. 274-280, 2009.

[68] NIST, “Usability & Biometrics,” National Institute of Standards and Technology”, 2008.

[69] S. Eliott, K. O'Connor, J. Robertson and R. Guest, “Expanding the Human-Biometric Sensor Interaction Model to Identity Claim Scenarios,” in *Proc. IEEE International Conference on Identity, Security and Behaviour Analysis (ISBA)*, Hong Kong, 2015.

[70] E. Kukula and S. Elliott, “Implementing Ergonomic Principles in a Biometric System: A Look at the Human Biometric Sensor Interaction (HBSI),” in *Proc. IEEE International Carnahan Conferences Security Technology*, 2006.

[71] A. Jain, S. Pankanti, S. Prabhaka, L. Hong, A. Ross and J. Wayman, “Biometrics: A Grand Challenge,” in *Proc. International Conference on Pattern Recognition*, Cambridge, UK, 2006.

[72] M. Brockly and S. Elliott, “Automatic Detection of Biometric Transaction Times,” in *Proc. The 8th International Conference on Information Technology and Applications (ICITA)*, 2013.

[73] International Civil Aviation Organization, “Machine Readable Travel Documents. Part 3: Specifications Common to all MRTDs,” ICAO, 2015.

[74] International Civil Aviation Organization, “Machine Readable Travel Documents, Doc 9303 (Sixth Ed),” ICAO, 2006.

[75] United States General Accounting Office (GAO), “Border Control Applications Piloted and Deployed,” Technology Assessment using Biometrics for Border Security,” GAO Report GAO-03-174, 2002, pp. 164-166.

[76] J. L. Wayman, “Report on the evaluation of the inpass hand geometry system,” Technical Report, National Biometric Test Centre Collected Works 1997-200, San Jose, 2000.

[77] D. Gorodnichy, S. Yanuskevich and V. Shmerko, “Automated border control: Problem formalization,” in *Proc. IEEE Symposium on Computational Intelligence in Biometrics and Identity Management (CIBIM)*, 2014.

[78] A. Palmer and C. Hurrey, “Ten Reasons Why IRIS Needed 20:20 Foresight: Some Lessons for Introducing Biometric Border Control Systems,” in *Proc. 2012 European Intelligence and Security Informatics Conference (EISIC)*, Odense, 2012.

[79] A. Ross, K. Nandakumar and A.Jain, “Handbook of Multibiometrics. International Series on Biometrics,” International Series on Biometrics. Vol 6., Springer, 2006.

[80] F. Scotti, “Computational intelligence techniques for reflections identification in iris biometric images,” in *Proc. IEEE International Conference on Computational Intelligence for Measurement Systems and Applications (CIMSA)*, Ostuni, Italy, 2007.

[81] F. Scotti, “Adaptive reflection detection and location in iris biometric images by using computational intelligence techniques,” *IEEE Trans. Instrum. Measur*, Vol. 59, No. 7, pp. 1825–1833, 2010.

[82] R. Snelick, M. Indovina, J. Yen and A. Mink, “Multimodal biometrics: issues in design and testing,” In *Proc. the ACM International Conference on Multimodal Interfaces (ICMI)*, Vancouver BC, Canada, 2003.

[83] P. Wild, P. Radu, L. Chen and J. Ferryman, “Robust multimodal face and fingerprint fusion in the presence of spoofing attacks,” *Pattern Recognition*, Vol. 50, pp. 17-25, February 2016.

[84] H. Wei, L. Chen and I. Ferryman, “Biometrics in ABC: counter-spoofing research,” *In Proc. FRONTEX 2nd Global Conference on Future Developments of Automated Border Control*, Warsaw, Poland, 2013.

[85] D. Cantaerero, D. Herrero and F. Mendez, “A Multi-modal Biometric Fusion Implementation for ABC Systems,” in *Proc. Intelligence and Security Informatics Conference (EISIC)*, Uppsala, Sweden, 2013.

[86] International Organization for Standardization. 24722:2007, “Information technology -- Biometrics--Multimodal and other multibiometric fusion,” ISO/IEC, 2007.

[87] G. Michael, T. Connie and A. Beng Jin, “A preliminary acclimatization study of a contactless biometrics using palm vein feature,” in *Proc. 6th IEEE Conference on Industrial Electronics and Applications (ICIEA)*, 2011.

[88] X. Liu, M. Pedersen, C. Charrier, F. Cheik and P. Bours, “An improved 3-step contactless fingerprint image enhancement approach for minutiae detection,” in *Proc. 6th European Workshop on Visual Information Processing (EUVIP),* Marseille, France, 2016.

[89] A. Kumar and C. Kwong, “Towards Contactless, Low-Cost and Accurate 3D Fingerprint Identification,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 37, No. 3, pp. 681-696, 2015.

[90] A. Anand, R. Donida-Labati, A. Genovese, E. Munoz, V. Piuri, F. Scotti and G. Sforza, “Enhancing the Performance of Multimodal Automated Border Control Systems,” in *Proc. International Conference of the Biometrics Special Interest Group (BIOSIG)*, 2016.

[91] R. Donida Labati, A. Genovese, V. Piuri and F. Scotti, “Toward Unconstrained Fingerprint Recognition: A Fully Touchless 3-D System Based on Two Views on the Move,” *IEEE Transactions on Systems, Man, and Cybernetic Systems*, Vol. 46, No. 2, pp. 202-219, 2015.

[92] J. Matey, O. Naroditsky, K. Hanna, R. Kolcynski, D. Lolcano, S. Mangru and M. Tinker, "Iris on the Move: Acquisition of Images for Iris Recognition in Less Constrained Environments," In *Proc. IEEE*, Vol. 94, No. 11, pp. 1936-1947, 2006.

[93] planetbiometrics (2015 November), "*Dubai Airport readies for 'iris on-the-move*'," Available: http://www.planetbiometrics.com/article-details/i/3666/desc/dubai-airport-readies-for-lsquoiris-on-the-move/

[94] International Organization for Standardization ISO/IEC, "Identification cards -- Machine readable travel documents -- Part 3: Machine readable official travel documents,", 2005.

[95] International Organization for Standardization ISO/IEC, "Identification cards -- Machine readable travel documents -- Part 1: Machine readable passport," 2008.

[96] International Civil Aviation Organization, "Doc 9303 Machine Readable Travel Documents. Part 9: Deployment of Biometric Identification and Electronic Storage of Data in eMRTDs," ICAO, 2015.

[97] J. Sanchez del Rio, D. Moctezuma, C. Conde, I. Martin de Diego and E. Cabello, "Automated border control e-gates and facial recognition systems," *Computers and Security*, Vol. 62, pp. 49-72, September 2016.

[98] P. Wild, S. Stolc, K. Valentin, F. Daubner and M. Ciabian, "Compression Effects on Security Document Images," in *Proc. Intelligence and Security Informatics Conference (EISIC)* Uppsala, Sweden, 2017.

[99] BSI, "An investigation into the performance of facial recognition systems relative to their planned use in photo identification documents - BioP I," secunet, 2004.

[100] Y. Choi, Y. Jeon and S. Park, "A study on secure protocol using the public key infrastructure approach in an e-passport," in *Proc. 12th International Conference on Advanced Communication Technology (ICACT)*, 2010.

[101] Gov.uk, "Biometric residence permits (BRPs)," [Online]. Available: https://www.gov.uk/biometric-residence-permits. [Accessed September 2017].

[102] B. Marzahn, "ePassport Based Identity Check," *ICAO Regional Seminar on MRTDs*, 2011.

[103] iaa.gov, "Arriving at Ben Guan Airport," [Online]. Available: http://www.iaa.gov.il/en-US/airports/bengurion/Pages/ArrivingPassengers.aspx. [Accessed September 2017].

[104] Frontex Europa, "Best Practice Operational Guidelines for Automated Border Control (ABC) systems," 31 August 2012. [Online]. Available: http://frontex.europa.eu/assets/Publications/Research/Best_Practice_Operational_Guidelines_for_Automated_Border_Control.pdf. [Accessed 25 July 2014].

[105] R. Labati, A. Genovese, E. Munoz, V. Piuri and F. Scotti, “Advanced design of Automated Border Control gates: biometric system techniques and research trends,” in *Proc. IEEE International Symposium on Systems Engineering (ISSE),* 2015.

[106] B. Marzahn, “E-gate case study: The German EasyPass Project,” *12th Conference Technical Advisory Group on Machine Readable Travel Documents*, ICAO, Montreal, 2011.

[107] N. Nuppeney, “Automated Border Control - state of play and latest developments,” in *Proc. NIST International Biometric Performance Conference*, 2014.

[108] M. Nuppeney, “Automated Border Control based on (ICAO compliant) eMRTDs,” in *Proc. NIST International Biometric Performance Conference*, 2012.

[109] Öffentlicher Abschlussbericht, “Untersuchung der Leistungsfähigkeit von biometrischen Verifikationssystemen – BioP II (A Study of the Peformance of Biometric Systems),” Secunet, Essen, 2005.

[110] A. Oostveen, M. Kaufmann, E. Krempel and G. Grasemann, “Automated Border Control: A Comparative Usability Study at Two European Airports,” in *Proc. 8th International Conference on Interfaces and Human Computer Interaction (IHCI 2014),* Lisbon, Portugal, 2014.

[111] R. Donida Labita, A. Genovese, E. Munoz, V. Pluri, F. Scotti and G. Sforza, “Biometric Recognition in Automated Border Control: A Survey,” *ACM Computing Surveys (CSUR)*, Vol. 49, No. 2, 2016.

[112] J. Sanchez, C. Conde, A. Tsitirdis, J. Gomez, I. Diego and E. Cabello, “Face-based recognition systems in the ABC e-gates,” in *Proc. 9th Annual IEEE International Systems Conference (SysCon)*, Vancouver, BC, Canada, 2015.

[113] G. Pirelli, “Usability in Public Services and Border Control,” in *Proc. Symposium of the Austrian HCI and Usability Engineering group (USAB),* 2009.

[114] A. Sasse, “Red-Eye Blink, Bendy Shuffle and the Yuck Factor. A User Experience of Biometric Airport Systems,” *IEEE Security & Privacy*, Vol. 5, No. 3, pp. 78-81, 2007.

[115] A. Al-Raisi and A. Al-Khouri, “Iris Recognition and the Challenge of Homeland and Border Control Security in UAE,” *Telematics and Informatics*, Vol. 25, No. 2, pp. 117-132, 2008.

[116] UKPS, “UK Passport Service Biometrics Enrolment Trial,” Atos Origin, London, 2005.

[117] B. Shackel, “Usability - Context, Framework, Definition, Design and Evaluation,” *Interacting with Computers*, Vol. 21, No. 5-6, pp. 339-346, 2009.

[118] M. Thoefanos, B. Stanton, R. Michaels and S. Orandi, “Biometric Systematic Uncertainty and the User,” in *Proc. IEEE International Conference on Biometrics: Theory, Applications, and Systems*, Crystal City, VA, 2007.

[119] M. Ylikauppila, S. Toivonen, M. Kulju and M. Jokela, “Understanding the Factors Affecting UX and Technology Acceptance in the Context of Automated Border Controls,” in *Proc. IEEE Joint Intelligence and Security Informatics Conference (JISIC)*, 2014.

[120] Fastpass EU, “FastPass Overview,” [Online]. Available: https://www.fastpass-project.eu/. [Accessed September 2017].

[121] European Commission, “ABC4EU Project Overview,” June 2014. [Online]. Available: http://abc4eu.com/about/. [Accessed September 2017].

[122] British Standards Institute, “FaceSymbol,” Report: FaceSymbol Design, Development & Testing, 2011.

[123] International Organization for Standardization. “Information technology -- ISO/IEC 24779-4:2017 Information technology -- Cross-jurisdictional and societal aspects of implementation of biometric technologies -- Pictograms, icons and symbols for use with biometric systems – Part 5: Face applications.

[124] Y. Choong, B. Stanton and M. Theofanos, “Biometric symbol design for the public - case studies in the United States and four Asian countries,” In *Proc. International Conference on Applied Human Factors and Ergonomics (AHFE2010)*, 2010.

[125] M.Drahansky, “Liveness Detection in Biometrics,” *Advanced Biometric Technologies*, InTech, pp. 439-444, 2011.

[126] A. Oostveen, “Non-use of Automated Border Control Systems: Identifying Reasons and Solutions” in *Proc. International BCS Human Computer Interaction Conference*, 2014.

[127] International Organization for Standardization ISO/IEC 1/SC 37 N4274, “ISO/IEC 24779-1, Biometrics — Cross-jurisdictional and societal aspects of the implementation of biometric technologies - Pictograms, Icons and Symbols for Use with Biometric Systems - Part 1: General Principles,” 2016.

[128] International Organization for Standardization. “Information technology -- ISO/IEC 24779-4:2016 Information technology -- Cross-jurisdictional and societal aspects of implementation of biometric technologies -- Pictograms, icons and symbols for use with biometric systems -- Part 4: Fingerprint applications.” 2016.

[129] International Organization for Standardization. “Information technology -- ISO/IEC 19795-1:2006 Biometric performance testing and reporting -- Part 4: Interoperability Testing”, “International Organization for Standardization, 2003.

[130] International Organization for Standardization. “Information technology -- ISO/IEC 24779-4:2016 Information technology -- Cross-jurisdictional and societal aspects of implementation of biometric technologies -- Pictograms, icons and symbols for use with biometric systems -- Part 9: Vascular applications, 2016.

[131] T. Gu, Z. Wu, L. Wang, X. Tao and J. Lu, “Mining Emerging Patterns for recognizing activities of multiple users in pervasive computing,” in *Proc. IEEE Mobile and Ubiquitous Systems: Networking & Services, MobiQuitous*, 2009.

[132] J. Aggarwal and M. Ryoo, “Human Activity Analysis: A Review,” *ACM Computing Surveys (CSUR),* Vol. 43, No. 3, 2011.

[133] A. Saxena, “Anticipating human activities using object affordances for reactive robotic response,” *IEEE Trans. Pattern Anal. Mac. Intell*, Vol. 1, pp. 1-14, 2015.

[134] W. Lin, M. Sun, R. Poovandran and Z. Zhang, “Human activity recognition for video surveillance,” in *Proc. IEEE International Symposium on Circuits and Systems (ISCAS)*, Seattle, WA, USA, 2008.

[135] L. Shao, L. Ji, Y. Liu and J. Zhang, “Human action segmentation and recognition via motion and shape analysis,” *Pattern Recognition Letters*, Vol. 33, No. 4, pp. 438-445, 2012.

[136] B. Wang, Z. Chen and J. Chen, “Gesture Recognition by Using Kinect Skeleton Tracking System,” in *Proc. Intelligent Human-Machine Systems and Cybernetics (IHMSC)*, 2013.

[137] K. Khoshelhman and S. Elberink, “Accuracy and resolution of Kinect depth data for indoor mapping applications,” *Sensors*, Vol. 12, No. 2, pp. 1437-1454, 2012.

[138] L. Wang, X. Gu, X. Tao and J. Lu, “Sensor-Based Human Activity Recognition in a Multi-user Scenario,” in *Proc. European Conference on Ambient Intelligence*, 2009.

[139] B. Logan, J. Healey, M. Philipose, E. Mungai-Tapia and S. Intille, “A long-term evaluation of sensing modalities for activity recognition,” in *Proc. International Conference on Ubiquitous computing (UbiComp),* 2007.

[140] L. Chen, J. Hoey, C. Nugent, D. Cook and Z. Yu, “Sensor Based Activity Recognition,” *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, Vol. 42, No. 6, pp. 790-808, 2012.

[141] I. Jang and K. Lee, “Depth Video based Human Model Reconstruction resolving Self-Occlusion,” *IEEE Transactions on Consumer Electronics*, Vol. 56, No. 3, pp. 1933-1941, 2010.

[142] S. Zhang, Z. Wei, J. Nie, L. Huang, S. Wang and Z. Li, “A Review on Human Activity Recognition Using Vision-Based Method,” *Journal of Healthcare Engineering*, 2017.

[143] Microsoft (2016), “Skeletal Tracking Microsoft Kinect” [Online]. Available: https://msdn.microsoft.com/en-us/library/hh973074.aspx. [Accessed 27 September 2017].

[144] B. Dehbandi, A. Barachant, A. Smeragliuolo, J. Long, S. Bumanlag, V. He, A. Lampe and D. Putrino, “Using data from the Microsoft Kinect 2 to determine postural stability in healthy subjects; A feasibility trial,” *PLoS ONE*, 2017.

[145] Microsoft (2017), “Kinect for Windows (SDK),” [Online]. Available: https://developer.microsoft.com/en-us/windows/kinect. [Accessed September 2017].

[146] Microsoft (2017), “Tracking Users with Kinect Skeletal Tracking,” [Online]. Available: https://msdn.microsoft.com/en-us/library/jj131025.aspx. [Accessed September 2017].

[147] K. Otte, B. Kayser and S. Mansow-Model, “Accuracy and Reliability of the Kinect Version 2 for Clinical Measurement of Motor Function,” *PLoS One*, 2016.

[148] Microsoft, “BodyFrame.FloorClipPlane Property,” [Online]. Available: https://msdn.microsoft.com/en-us/library/windowspreview.kinect.bodyframe.floorclipplane.aspx. [Accessed September 2017].

[149] Microsoft, “FloorClipPlane Defintion from Kinect SDK,” [Online]. Available: https://msdn.microsoft.com/en-us/library/windowspreview.kinect.bodyframe.floorclipplane.aspx. [Accessed September 2017].

[150] Wolfram MathWorld, “Point-Plane Distance,” [Online]. Available: http://mathworld.wolfram.com/Point-PlaneDistance.html. [Accessed September 2017].

[151] Microsoft, “Joint Filtering,” [Online]. Available: https://msdn.microsoft.com/en-us/library/jj131024.aspx. [Accessed September 2017].

[152] S. Obdržálek, G. Kurillo, F. Ofli, R. Bajcsy, E. Seto, H. Jimison and M. Pavel, “Accuracy and robustness of Kinect pose estimation in the context of coaching of the elderly population,” in *Proc. IEEE Eng. Med. Biol. Soc.*, 2012.

[153] T. Hoai-An, N. Hai, T. Phuc and T. Mai, “Support vector machine algorithm for human fall recognition Kinect-based skeletal data,” in *Proc. National Foundation for Science and Technology Development Conference on Information and Computer Science (NICS),* 2015.

[154] R. Clark, Y. Pua, K. Fortin, C. Ritchie, E. Webster, L. Dehehy and A. Byrant, “Validity of the Microsoft Kinect for assessment of postural control*,” Gait & Posture*, Vol. 36, No. 3, pp. 372-377, 2012.

[155] G. Kurillo, A. Chen, R. Bajcsy and J. Han, “Evaluation of upper extremity reachable workspace using Kinect camera",” *Technology and Health Care: Official Journal of the European Society for Engineering and Medicine*, Vol. 21, 2013.

[156] H. Funaya, T. Shibata, Y. Wada and T. Yamanaka, “Accuracy Assessment of Kinect Body Tracker in Instant Posturography for Balance Disorders,” in *Proc. International Symposium on Medical Information and Communication Technology (ISMICT),* 2013.

[157] B. Bonnechere, B. Jansen, P. Salvia, L. Omelna, F. Moiseev, V. Sholukha, J. Cornelis, M. Rooze and S. Van Sint Jan, “Validity and reliability of the Kinect within functional assessment activities: Comparison with standard stereophotogrammetry,” *Gait & Posture*, Vol. 39, No. 1, pp. 593-598, 2014.

[158] B. Galna, G. Barry, D. Jackson, D. Mhiripiri, P. Oliver and L. Rochester, “Accuracy of the Microsoft Kinect sensor for measuring movement in people with Parkinson’s disease,” *Gait & Posture*, Vol. 39, No. 4, pp. 1062-1068, 2014.

[159] S. Asteriadis, A. Chatzitofis, D. Zarpalas, D. Alexiadis and P. Daras, “Estimating human motion from multiple Kinect Sensors,” in *Proc. International Conference on Computer Vision / Computer Graphics Collaboration Techniques and Applications*, 2013.

[160] J. Tong, J. Zhou, L. Liu, Z. Pan and H. Yan, “Scanning 3D Human Bodies Using Kinects,” *IEEE Transactions on Visualization and Computer Graphics*, Vol. 18, No. 4, pp. 643-650, 2012.

[161] L. Yang, L. Zhang, H. Dong, A. Alelaiwi and A. El Saddik, “Evaluating and Improving the Depth Accuracy of Kinect for Windows v2,” *IEEE Sensors Journal*, Vol. 15, No. 8, pp. 4275-4285, 2015.

[162] B. Dehbandi, A. Barachant, D. Haray, J. Long, K. Tsgaris, S. Bumanlag, V. He and D. Putrino, “Using data from the Microsoft Kinect 2 to quantify upper limb behaviour: a feasibility study*,” IEEE Journal of Biomedical and Health Information*, No. 99, 2016.

[163] J. See, L. Dodakian, C. Chou, V. Chan, A. Mckenzie, D. Reinkesmeyer and S. Cramer, “A Standardized Approach to the Fughl-Meyer Assessment and its Implications for Clinical Trials,” *Neuroehabilitation and Neural Repair*, Vol. 27, No. 8, pp. 732-741, 2013.

[164] H. Gonzalez-Jorge, P. Rodriguez-Gonzalvez, J. Martinez-Sanchez, D. Gonzalez-Aguilera, P. Arias, M. Gesto and L. Diaz-Vilarino, “Metrological comparison between Kinect 1 and Kinect 2 sensors,” *Measurement*, Vol. 70, pp. 21-26, 2015.

[165] Z. Moore, “Human-Biometric Sensor Interaction Automation Using the Kinect 2.”, M.S Thesis, Purdue University, West Lafayette, USA, 2016.

[166] Neurotechnology, “Neurotechnology MegaMatcher SDK,” [Online]. Available: http://www.neurotechnology.com/megamatcher.html. [Accessed September 2017].

[167] NIST, “Minutiae Interoperability Exchange (MINEX) Overview,” [Online]. Available: https://www.nist.gov/programs-projects/minutiae-interoperability-exchange-minex-overview. [Accessed September 2017].

[168] ISpy Open Source Video Surveillance Software, [Online]. Available: https://www.ispyconnect.com/. [Accessed September 2017].

[169] National Institute of Standards and Technology “Development of NFIQ 2.0,” NIST, 2011.

[170] M. Gschwandtner, S. Tolc and F. Daubner, “Optical Security Document Simulator for Black-Box Testing of ABC Systems,” in *Proc. Intelligence and Security Informatics Conference (JISIC)*, The Hague, Netherlands, 2017.

[171] Biometricupdate.com, [Online]. Available: http://www.biometricupdate.com/201709/vision-box-unveils-new-facial-recognition-solution-for-border-control. [Accessed September 2017].