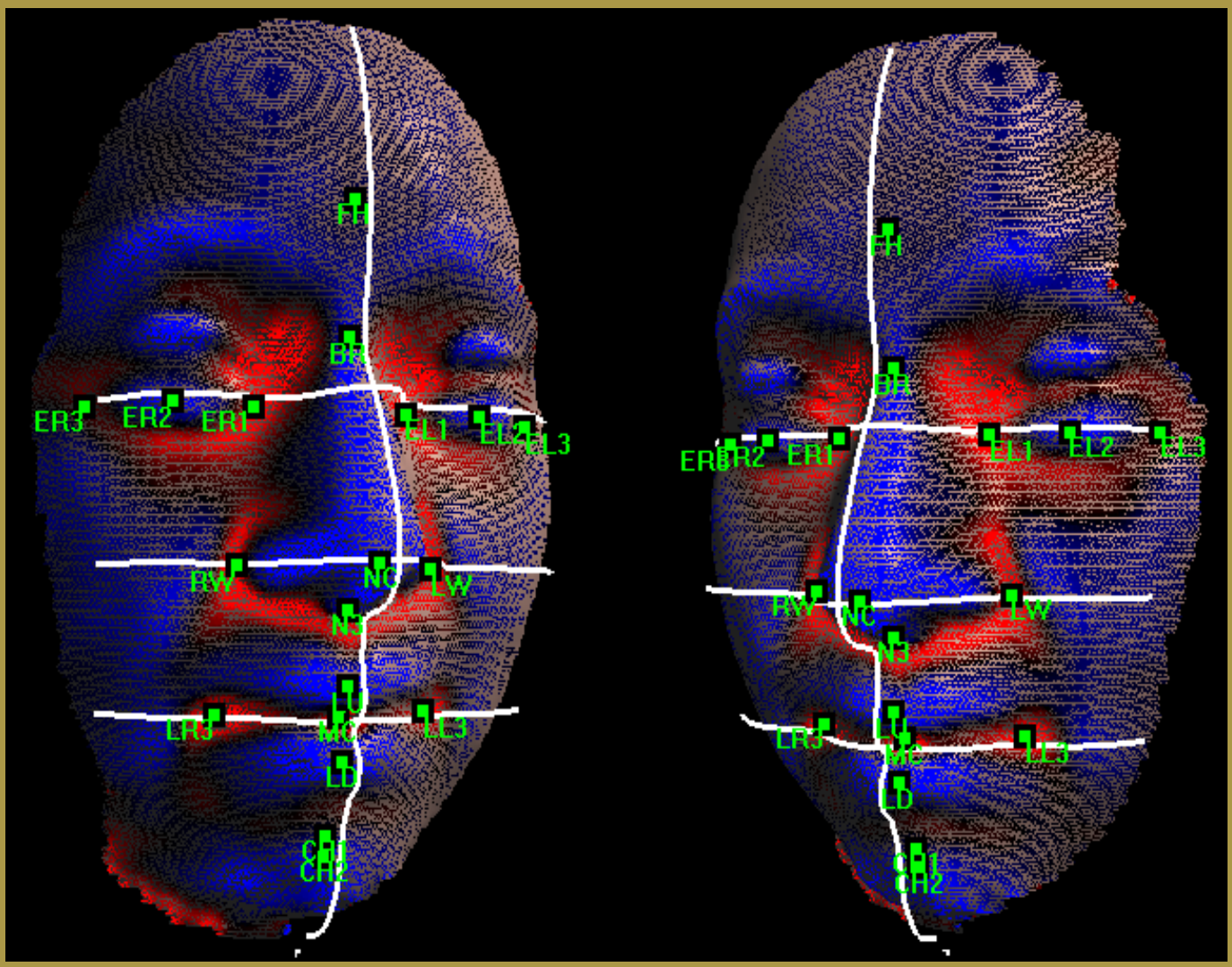
NPIA
National Policing Improvement Agency

Automated Face Recognition

Applications within Law Enforcement



Market and Technology Review

October 2006

Version 1.1
Ambika Suman
Biometric Technologies Manager
Biometrics Team – Identification

This report has been produced by NPIA's Biometrics Team as part of an ongoing programme of work to investigate and understand the current capabilities of automated Face Recognition (FR) technology, and to identify opportunities for its deployment within the police service.

This report was originally produced under PITO (Police Information Technology Organisation) in the preceding months of transition to the NPIA. The version that is now distributed has been re-branded for publication under the NPIA; however, its content remains the same as references are made to both PITO and the NPIA.

The report and the recommendations that are made still hold true and the NPIA Biometrics Team continue its programme of work in this area.

There are many operational scenarios where automated FR could have an important role to play; a number of these are considered, along with the significant challenges they all present to current state-of-the-art FR systems. However, whilst such examples are useful to highlight specific issues, it is important to point out that this report does not attempt to establish a business case for the use of FR technology in policing in any of these areas. Rather, it is designed to inform stakeholders about how the technology works, why its use in law enforcement presents unique, challenges and how these may be overcome in the future; through academic research and ongoing R&D by companies specialising in this field.

The diverse nature of policing means that the user requirements for FR will also be very varied. For example, the requirements of a Special Branch officer at an airport will be very different to that of a Senior Investigating Officer looking to use FR in a post-incident scenario.

An exercise is already underway within NPIA to identify all the business areas that could benefit from FR and to capture the different user requirements in each area.

As part of this work, NPIA's Biometrics Team is working closely with the NPIA FIND project; this project will provide an initial national database of standard mugshot images that in the future could be used in conjunction with FR technology.

The role of this report is therefore to inform discussions and stimulate debate on how the police service might make use of automated FR technology both now and in the medium to long term as new developments in this area increase performance and open up new

opportunities for its deployment. The report is also designed to inform academia and the vendor community of the challenges and issues that need to be addressed if such technology is to ever be fully effective in a law enforcement environment.

It should be regarded as a supporting document to assist with the timely development of business case(s) for deploying automated FR technology, and to help in shaping future Police strategy for coherent use of identification systems.

Colin Patton
Director, Identification
NPIA

DCC Stephen Long
ACPO Chair, Science & Technology
Senior Responsible Officer for the FIND Project

Product Control Page

Authors:	Ambika Suman	Biometric Technologies Manager
Address:	National Policing Improvement Agency New Kings Beam House 22 Upper Ground London SE1 9QY	☎ 020 8358 5446 020 8358 5533 ambika.suman@npia.pnn.police.uk

Distribution List:	
Mark Blake (MB)	Central Customer (CC)
Simon Moore (SM)	Head ID Projects
Colin Patton (CP)	Director
Geoff Whitaker (GW)	Head of Biometrics
Deviani Pitrola (DP)	Systems Engineer
Aidan Littlewood (AL)	Systems Engineer
Prof. Farzin Deravi	University of Kent
Prof. Mike Fairhurst	University of Kent
Josef Kittler	University of Surrey
Richard Garner	IPS (UKPS)
Mike Franklin	IPS (UKPS)

Sign Off Authorised By:			
NAME	**Responsibility**	**Signature**	**Date**
Colin Patton	Director		
Geoff Whitaker	Head of Biometrics		

Issue Control:		
Version	**Date**	**DESCRIPTION**
0.1	1st June 2006	First draft for review
0.2	20th June 2006	Peer review – Update of Executive Summary
0.3	8th August 2006	Inclusion of comments from GW and CP Completion of Glossary and Appendix Addition of diagrams and list of figures

Issue Control:		
Version	**Date**	**DESCRIPTION**
0.4	10th August 2006	Distribution to CC
0.5	28th September	Incorporation of comments from CC, Kent and Sign off, Removal of "Restricted" markings.
Version 0.6	4th October	Final comments and sign off

CONTENTS

LIST OF FIGURES

Executive Summary

Automated face recognition is increasingly becoming a mature technology and has clear applications for policing; many of these are highlighted in the latest version of PITO's Identification Roadmap (2005 – 2020). However, the mass market drive in the development of the technology is from the civilian and commercial sectors; applications can already be seen in areas such as biometric based physical access control, passport control and immigration.

It is important to appreciate that the use of face recognition technology in operational policing environments is fundamentally different to its use in the majority of civilian application scenarios. Firstly, a major source of police data will be from CCTV or surveillance images, captured in uncontrolled conditions, of varying poor quality, but also because the technology will normally be required to operate in identification mode ('one to many' searches) rather than merely performing a 'one to one' verification against a claimed identity. Thus, the performance and accuracy requirements of the systems will also be diverse, having to cater for both real time searching as well as non-real time (*forensic or post event analysis*) applications.

PITO's Biometrics team has investigated a variety of technologies, using both 2D and 3D images as well as techniques to manipulate and combine multiple views. It has become clear that some of these are far more suited than others to the law enforcement domain, and in particular to use in forensic investigations. The reasons for this are summarised below but are explained in more detail in the main body of this report.

In spite of its shortcomings, there are some clear advantages to deploying face recognition technology in law enforcement applications, particularly in forensic or 'post event analysis' scenarios. Rapid probing of police mugshot collections, identification of offenders on 'watch-lists' in mobile scenarios, compilation of witness albums, video based ID parades, or bail reporting, and other diverse uses within the wider CJS could all deliver significant benefit to policing. The resulting decrease in the manual overhead of performing such tasks, coupled with more efficient and reliable identification of individuals and detection of offenders, has many parallels with the benefits originally identified from the introduction of automated fingerprint searching in England and Wales in the 1990s. National systems such as NAFIS / IDENT1 have evolved to become essential tools in the fight against crime and have revolutionised the way in which the fingerprint service operates.

This report considers a number of scenarios whereby face recognition could similarly impact operational policing if deployed either now or in the future, highlighting the pros and cons, as well as the opportunities and challenges presented by each. However, whilst a centralised national search capability may be appropriate in some circumstances, it seems likely that a nationally co-ordinated rollout of targeted local applications may be a better approach. One of the recommendations of this report is to explore such opportunities to better understand police requirements for identification using facial image data.

The arrival of a UK-wide mugshot database in the near future (as a result of PITO's FIND project) is both a key enabler and a major driver for the introduction of some form of automated face recognition capability, since without it the full benefits to policing of creating such a national database cannot be fully realised. It is important to understand however, that such technologies have somewhat different requirements of facial image data than a database such as FIND will initially deliver. FIND will provide only image sharing and retrieval (generally of just a frontal or 'full face' image) and a limited search capability, based on demographic data captured at the same time as the images. These images may be perfectly adequate for human examination and comparison, but there is no guarantee that such images are best suited for use with automated systems, particularly when used in conjunction with poor quality, uncontrolled CCTV images.

The use of 3D facial image data, as well as the ability to 'fuse' multiple images to obtain 3D facial representations, is still relatively new and is the subject of much ongoing research. However, there are numerous advantages in moving towards 3D mugshot capture in terms of improved matching accuracy

and more opportunity to make identifications using such images with CCTV. PITO's Biometrics team believe that in order to make effective use of automated face recognition in a forensic environment, it will be necessary to employ either 3D images, or multiple 2D images taken from different angles.

Research and development into the design of effective user interface functions and image enhancement tools for forensic image comparison, in conjunction with a better understanding of how human beings analyse and compare faces, may yield improved performance from the systems when operating with police data. This area has been largely ignored by face recognition vendors as the technology has developed to cater for the civilian need of fully automated, high throughput matching systems.

Much important work is going on regarding standards for 3D imaging and MPEG (particularly MPGE7) Standards looking at 'Advanced Face Descriptors' and 'Advanced Image Coding and Searching', and ISO 15378 Part 8 will apply MPEG7 to photo management, allowing searching of databases for duplicate images, changed images, or for specific scene elements, such as people and vehicles. Whilst some of the technology to do this is still at the research stage, adoption of such standards will be essential in enabling large (possibly distributed) databases of images to be efficiently managed, manipulated and searched, whether using face recognition technology or other search descriptors. It is essential to maintain a close watch on developments in this area.

Recommendations and Opportunities

Below is a summarised list of recommendations within the areas of technological improvement and potential opportunities of benefit to policing. These are relevant to stakeholders such as ACPO, industry, the police service, wider CJS stakeholders and academia, and can also be used to inform the development of the business case for national roll out of face recognition (i.e. the Autoface project to be run by PITO)

1. Explore the benefits of 3D mugshot capture

Opportunity: To pilot 3D mugshot capture and 3D to 3D mugshot searching at point of entry in custody

Benefits: Allows one to explore the benefits of 3D mugshot capture in terms of image quality, speed and the value of the 3D data for identifying persons caught on CCTV. It also provides a 'non contact' form of identification in custody.

2. Combine Automated Video Analysis and Forensic Identification technologies

Opportunity: To pilot Face Recognition with CCTV, based on an integrated model of intelligent video surveillance and analysis with a forensic identification interface.

Benefits: To gather hard evidence on the performance of Face Recognition with CCTV data in an operational context, using real police data as well as 'post event' or forensic identification applications that may be acquired to complement such applications.

3. Combine cognitive approaches with the development of User Interfaces for Face Recognition

Opportunity: To pilot systems that explore the role of human computer assisted tools (for example, Identix facial search interface) to evaluate the performance of face recognition searching with operator assisted tools in comparison to fully automated searching.

Benefits: This allows one to better understand the user requirements of facial identification in terms of the operator's need to analyse such information. This can be used to inform the specification of future user requirements for deployment of face recognition technology.

4. Explore applications that are targeted for specific user scenarios i.e. serious and organised crime

Opportunity: To pilot the use of face recognition with CCTV images whereby cameras are strategically placed at choke points in public areas (i.e. stairwells at stations, near posters or signs, or at entry points to grounds) in order to ensure a good frontal image of a face can be captured, and subsequently be searched against a watchlist of known suspects and offenders.

Opportunity: To pilot 2D mugshot to 2D mugshot searching with the roll out of FIND to detect duplicates in the FIND collection.

Opportunity: To pilot Face Recognition in support of Anti terrorist Intelligence Operations.

Benefits: One would have a better understanding of the value of face recognition for different aspects of policing which can then help to inform the strategy of nationally rolling out Face Recognition – whether it should be a centralised search capability or a centrally co-ordinated series of local deployments.

5. Deployment of High Definition CCTV with a view to capture facial image data

Opportunity: To pilot or perform evaluative benchmarks of Face Recognition technology used with High Definition CCTV for mugshot to CCTV and CCTV to CCTV searching.

Benefits: One would obtain a better understanding of what is potentially achievable with high quality data which may lead to further improvements in CCTV data capture and perhaps justify the need for funding to enable this.

6. Explore the potential of Face Recognition at range (using LIDAR)

Opportunity: To evaluate Face Recognition at range using LIDAR 3D technology, perhaps specifically working in areas of covert surveillance such as anti terrorist operations or operations to combat serious and organised crime.

Benefits: Fast, accurate and covert identification of individuals at range which could hugely impact many areas of secure surveillance and intelligence operations**.**

7. Close monitoring of MPEG standards (MPEG7) and Biometric standards for 3D imaging for face

Opportunity: To monitor and influence development of future MPEG and biometric standards and to understand how they may be applied to large scale image databases (such as FIND)

Benefits: This work could have implications for future large scale image database design as well as enabling greater interoperability between systems.

Background

The Biometrics team within PITO's Identification Directorate has been investigating facial recognition as part of its Facial Recognition Evaluation and Demonstration Strategy (FRED). As part of the first phase of this work, the team has acquired a number of state of the art systems for the purpose of gaining familiarity with face recognition applications representative of those that are currently, or will soon be, commercially available.

The overall aim of this work area is to assess the viability of facial recognition technologies in meeting current and future requirements of the police service. Over recent years, the team has been extensively collaborating with universities, both in the UK and internationally, as part of its Academic Collaboration Programme. Many universities are focussing their research into automated methods of facial identification, and the team has been working with academic partners to gain a better understanding of current research in this field. The team has also been working closely with members of the police user community, wider Home Office (HO) and commercial sector stakeholders in order to better understand their requirements for facial identification, and the value and limitations of current technology and face recognition research in meeting them.

The FIND database will shortly deliver the capability for police forces in England, Wales and Scotland to access, for the first time, a national collection of mugshot images. The ability to retrieve and share such images, and to search the data using demographic descriptors, will have an enormous impact on many areas of everyday police business. An inevitable outcome of delivering such a capability is an increased interest in the use of automated Face Recognition technology in conjunction with these images. PITO received a mandate from ACPO towards the end of 2005, to start work on developing a business case for using automated Face Recognition technology and this is currently being progressed by PITO's Central Customer.

A key objective of PITO's Biometrics team's work on Face Recognition over the past 12 months, and the main purpose of this report, is to provide stakeholders in the 'Auto Face' project with a better understanding of the technology, its uses and limitations, and to consider possible applications for its future deployment. It is designed to help with the development of the Strategic Business Case and future requirements definition for automated Face Recognition systems within policing.

Report Purpose

The purpose of this report is to provide guidance to ACPO, the user community, industry and academia on the potential use of face recognition for the police service in order to maximise the benefits to operational policing where face recognition is employed. The report includes recommendations for future work in this area which PITO's Biometrics Team feels is necessary if the opportunities provided by FIND are to be fully exploited.

The report provides a brief analysis of the current market and technological trends in face recognition applications. It is aimed at those who are unfamiliar with face recognition technology and its possible applications, and wish to obtain an overall understanding of the subject and its application to policing. Technical aspects that are presented are at a high level, to provide a simple overview of the various algorithmic approaches that are typically used by vendors. Note that although the report is heavily focussed on police applications of the technology, it also provides sufficiently in depth information for those interested in other broader aspects of the technology's application.

A number of application areas are examined and the pros and cons of each are considered. In addition, an overview of ongoing research and existing implementations around the world is provided.

NOTE: The area of automated Face Recognition is a rapidly evolving field. This report endeavours to provide an accurate picture of the market place, but it is inevitable that some of what is reported here may have changed during the course of publishing the report and its wider circulation. PITO would welcome feedback from those reading this report on any errors or omissions in order that they can be addressed and corrected in future reports by PITO's Biometrics Team.

Report Structure

For ease of reference, the following provides a breakdown of the report and descriptions of each section.

Section 1: Introduction
This section provides a brief overview of: where face recognition technology is currently used, the difference in the requirements of commercial and civilian applications in comparison to law enforcement applications, the challenges with the current performance of the technology, and why PITO has been exploring the potential of face recognition for the police service.

Section 2: Market Overview
This section very briefly describes the current trends in the face recognition market and how it may be impacted by the rise in take up of the technology, the availability of low cost cameras integrated with mobile technology, and legacy infrastructures.

Section 3: Face Recognition Technology
This section explains in very basic terms how face recognition systems work and describes the processes involved, from capturing a facial image right the way through to declaring a match. 2D, 3D and multiple view based approaches are considered and the pros and cons of each are explored.

Section 4: Face Recognition Application Areas
This part of the report explores, using a SWOT analysis, the deployment of face recognition in four generic application areas based on the use of images captured in 'controlled' (mugshot/portrait) and 'uncontrolled' (CCTV) environments.

Section 5: Government Initiatives for Face Recognition
This section illustrates the level of interest that there currently is for face recognition from governments across the world and the range of application areas where the technology has been implemented. The information in this section is by no means exhaustive, and it is likely that new developments will have arisen during the preparation of this report, and following its publication.

Section 6: Standards
This section provides information on standards activities with respects to face recognition and other (related) technologies that may impact future developments of applications.

Section 7: Vendors
A number of vendors providing face recognition applications are listed here. Though this is accurate at the time of print, the market is rapidly evolving, thus the information here is subject to change.

Section 8: Research
The section provides details of research activities (past, present and future) in the area of face recognition.

Section 9: Report Summary
The report is summarised in this section

Section 10: Recommendations and Appendix I: Recommendations to PITO
The findings of this report provide two sets of recommendations; the first set is overviewed in this Section and is aimed at external stakeholders to PITO, namely, ACPO, the police service, academia and industry. The second set of recommendations informs the future business planning for PITO with respects to continued effort into exploring the potential of face recognition. These are detailed in Appendix I of this report.

1 Introduction

Humans have a remarkable ability to recognise faces and to use them to distinguish between people. Research into automating this capability dates back over 30 years. More recently Face Recognition as a "biometric" technology (whereby the face is the physiological trait that uniquely identifies an individual) has become a hot topic of modern day research as a result of the growing pressure to exploit faces as a means of identification from both the commercial and law enforcement sectors. This is evidenced by the increasing number of companies that are offering face recognition solutions and by the considerable amount of new research that is being funded.

1.1 Where is there a need for automated face recognition?

The need to identify people can be split into two distinct areas based on the diversity of the requirements that are typical to each area: these are **Law Enforcement** and **Civilian** application areas.

In all cases, speed of identification is a key business driver, but the main advantages that face has over other biometrics are "universality" and "acceptance". Everyone has a face and it is widely used and accepted as a means of identification. Facial images are used in identity documents such as passports, driving licences etc, but images are also available from CCTV surveillance cameras, and are used in witness albums, video parading, composite identification, etc. As a result, there are a number of both covert and overt sources from which facial images are readily available and can subsequently be used to determine identity.

Finally, face recognition is a task that we all perform naturally in our social environments; and so any part of an identification process that requires human intervention is likely to require little or no operator expertise (*compared, for example, with fingerprints where a qualified fingerprint expert is needed to confirm an identification*).

1.2 Where is face recognition typically used?

Automated face recognition technologies are already in use in both the civilian and law enforcement areas. Civilian applications are typically focussed on improving the current areas where faces are used as a means of proving or verifying the identity that a person claims to be. (This is typically a one-to-one comparison otherwise referred to as *Verification*). The objective is to either authorise, or deny access to a certain privilege. Thus, civilian based applications have very apparent incentives for the end user, and are typically required to identify lots of people, in real time, with minimal effort, and high accuracy. For example, presenting a face on a passport enables verification of an individual when entering a country or crossing a border. However, (automated) face recognition helps to speed this process up and provides a more accurate method of identity verification (*often superior to humans in such circumstances*). In this case, the benefits to the end user would be entry into the country and fast passage through immigration. Most applications in the civilian domain are for security and access control for airports and buildings, and also in logical access applications where the face is used either in conjunction with, or in place of a pin code or password.

In all of these cases, the reference image of the face that is stored with the identity information is a full face image of the person facing straight ahead to the camera, taken in a controlled or indoor environment. The image that it is compared to will often be another image, taken in very similar conditions, possibly with the same type of camera, or it may be captured 'live', while the individual is present. Overall the images will generally be of good quality and acquired under optimal conditions.

Within policing there are some needs that are similar to the civilian requirements to identification. Thus some police requirements for automated face recognition are met by applications catering for civilian

users, but there are also some uses of the technology that are fundamentally different for policing. Within this report these are referred to as the "Law Enforcement" type applications of face recognition. They are forensically motivated, involving image evidence from crime scenes, CCTV footage, surveillance images, and other sources of facial identification data, that are used in non real time (*post event analysis*) mode to provide intelligence that may further investigations. The facial images used in these cases are typically of very poor quality, acquired in variable conditions, with off-centre views of the face, often containing expressions, or even occlusion within regions of the face. The applications are often required to function in an *identification* mode, whereby one image is searched against a background dataset of images as a one-to-many comparison. One can draw analogies to the use of face recognition (or facial images per se) with that of fingerprints. Ten-print images of the fingers are captured in a controlled custody environment, scene of crimes marks (latent fingerprint impressions) are often randomly deposited and thus of varied quality and appearance. However, employing similar approaches to the treatment of both types of forensic data may benefit police capabilities, as illustrated by the following table:

Fingerprints searching on AFIS[1]	Face Recognition Equivalents	Capability
Ten print[2] to Ten print searching	Mugshot to Mugshot Live Image to Mugshot	Identification/Verification of an individual in custody
Ten print to Mark searching	Mugshot to CCTV image	Linking known individuals to unsolved crimes
Mark to Ten print searching	CCTV image to Mugshot	Identifying suspects from forensic evidence at crime scenes
Mark to Mark	CCTV images from multiple cameras	Linking unsolved crimes together

Figure 1: Analogous Forensic value of Fingerprint and Face

1.3 Why is automated face recognition difficult?

In either application area, (and in biometric applications generally) the quality of the image is the key factor that impacts the performance of the technology. 'Quality', in this context, does not just mean the quality of the captured image, but includes the distinctiveness of features on the face, the impact of the environment and equipment on the captured image, and any pre or post processing of the image which may impact its suitability for automated searching. The quality of the image captured has a knock on effect on all subsequent recognition processes. For Law Enforcement applications of the technology there will often be little or no control over the acquisition or source of image data that may potentially be used for identification, whereas in the civilian domain the technology and capture process can usually be configured to optimise the quality of the images. Even so, facial image data capture for face recognition is still largely non-standardised. It was only very recently in 2002 that work on defining image capture standards for facial identification was commenced by international standards organisations such as ISO[3],

[1] AFIS Automated Fingerprint Identification System traditionally used within law enforcement for identifying fingerprints captured in custody or for searching latent fingerprint impressions left at crime scenes.
[2] Ten rolled fingerprints captured using ink or on Livescan from arrestee in police custody
[3] National Institute of Standards and Technology International Standards Organisation

BSI[4] and ICAO[5]. The Police Information Technology Organisation (PITO) produced the FIND image standard in 2005 (very similar to the ICAO standard) as the common method to be adopted for mugshot image capture for police forces in the UK. Typically face recognition vendors will now cite compliance with such standards, however, prior to these being established, acceptable performance of the technologies was heavily reliant on the use of imaging equipment specified by vendors.

Most facial recognition applications are highly sensitive to changes in lighting, pose, age, population, etc. In most real world scenarios these factors are difficult to control and thus have an adverse effect on the performance of the technologies. Most face recognition technologies that are currently on the market or being researched cater primarily to those applications where the environment is controllable. Uncontrolled face recognition such as CCTV surveillance, etc. is generally considered to be too challenging, offering little return on investment in comparison to those operating in controlled environments, as typically seen in the civilian domain.

The table below shows a comparison between the use of fingerprints and faces for identification. It clearly shows why the accuracy of face recognition will never match fingerprints.

Fingerprints	Face
There are normally ten unique fingerprints per individual	There is only one face per individual
Fingerprints differ considerably between individuals	Faces are fundamentally all very similar
Fingerprints do not change over time	Faces change considerably with age
Fingerprints can be accurately captured and represented in 2D	Faces are 3D objects and are thus affected by changes in viewing angle and lighting
There are well established processes for fingerprint comparison	There is currently no general consensus regarding the use of facial images for identification
Fingerprint recognition technology developed from a well defined forensic science and human expertise, traditionally applied to assist experts in the examination of fingerprints from crime scenes and individuals.	Face Recognition has been designed to be fully automated, that is, with limited expert involvement and motivated by civilian requirements. It has not been designed as forensic tool as human expertise for forensic facial identification is yet in the process of being established.

Figure 2: Fingerprint Identification Vs Facial Identification within Law Enforcement

Face recognition in a forensic environment is fundamentally different to its uses in conventional civilian applications. The vast majority of subjects will be non compliant/non co-operative, enquiry data will typically be of low quality, varying in lighting, pose angle, expression, occlusion, etc. The technology is more often required to work in an identification mode, searching very large databases (FIND may hold in

[4] British Standards Institute
[5] International Civil Aviation Organisation

excess of 25 million records. Frequently even human operators will be unable to recognise or confirm matches output from a system due to the quality of images data.

1.4 Why does PITO have an interest in Face Recognition?

The second iteration of the PITO Identification Roadmap (2005) highlighted a number of areas that PITO and other stakeholders across the police and wider CJS need to act on in order to realise the full potential of identification capabilities that meet the requirements of the police service. The study specifically identified Face Recognition technologies as potentially having a huge impact on many areas of policing.

The UK currently has the most widely deployed CCTV coverage in the world, with a ratio of approximately 14 people to each camera. Ongoing work to develop a national strategy for CCTV, led by the National CCTV User Group, could result in a doubling of this figure to around 1 camera per 7 individuals. The most abundant source of facial image data comes from CCTV footage. However, when a crime is recorded on camera, the process of actually locating all of the relevant footage and identifying the possible suspects who may have committed it is a manual, and extremely time consuming, process. As a result, this abundant data source remains widely redundant. It would be of great operational advantage to the police if this process was assisted by automation – to more efficiently trawl through hours of footage, and ultimately, to identify the individuals within the data. For this reason, the combination of Face Recognition with CCTV has clear applications for advancing operational policing. **If successful, Face Recognition will be a revolutionary improvement in policing**, much in the way that automatic fingerprint identification enabled capabilities to automatically search, **on a national scale**, fingerprints from custody and crime scenes and **return matches with minutes or even seconds.**

Whilst PITO is not currently involved in any operational projects with respects to CCTV, it is very actively engaged in monitoring and driving forward developments of the CCTV technology with a view to exploiting its potential use Face Recognition (and other identification technologies, such as voice and gait) The Biometrics team within PITO has very rapidly developed its knowledge base and expertise in this area with a view to informing future capabilities that be may be delivered through national services such as IDENT1 and FIND, etc.

Facial Recognition is rapidly becoming a mature technology which many forces are eager to take up. Some forces such as Greater Manchester Police, West Yorkshire, and others have already procured and implemented Facial Recognition systems and a “Futures Working Group” on Automated Facial Identification has also been established by ACPO (2002). The key issue identified by all is that although there are many different systems and suppliers, there is little independent data on accuracy, in particular with respect to UK police operational data. In effect, the choice of a vendor is difficult and users often have unrealistic expectations of what the technology is capable of doing and where it can be used.

By 2007 PITO will have delivered “FIND” to police forces in England, Wales and Scotland enabling them to capture, store, and exchange mugshot images on a national scale. (However, by this time PITO will have ceased to exist and the FIND project will have been delivered under the new National Policing Improvement Agency). Inevitably, the FIND project has also generated a huge amount of interest in face recognition technologies both within forces and across the wider CJS. Searching on faces may be expected to be a key requirement for expanding the FIND project in the near future.

The IDENT1 “multimodal” platform was specifically procured in anticipation of such a requirement emerging. It is envisaged that FIND will either sit on, or will interface to, IDENT1 thereby extending the range of identification feature data (*range of biometric modalities*) from fingers and palms to include mugshots, scars marks and tattoos. In addition, IDENT1 will look to fuse the outputs of different classifier systems, perhaps by combining algorithms for different biometrics (*e.g. face and finger*) to improve the accuracy of identifications being made and output from the system.

The Lantern project is one example of where such an approach could prove beneficial. The aim of the current project is to demonstrate the capability of performing a mobile identity check on a person, without the need for them to be taken to a Livescan unit in a custody suite. The demonstrator uses a mobile application that is able to capture and search a person's fingerprints against IDENT1 and return this information to a police officer at the roadside. Two of the main challenges highlighted by the Lantern demonstrator are the accuracy of searches when using only a limited number of fingers and the quality of fingerprints that can be captured from a mobile device. The use of a second feature, such as face, to search on can readily improve the level of confidence in the returned results whilst also providing an alternative method of identification when fingerprints cannot be captured.

Another PITO project, that is not detailed here, is NVIS (National Video Identification Strategy). This will also have a major impact on the future use of facial image data for identification and investigative purposes within policing.

In November 2005, PITO received a mandate from ACPO to develop a Business Case for the national deployment of a face recognition capability to forces. This report is expected to be the first in a series that forms part of ongoing investigations by PITO's Biometrics Team into the feasibility of face recognition technology being able to meet policing requirements, both now, and in the near future. This will feed directly into the development of this Business Case, and will inform requirements for future national face recognition applications, should the Business Case be accepted by ACPO and the Home Office.

2 Market Overview

The accelerated growth of face recognition systems is impressive. In 1998, revenue forecasts for the face recognition technology market were just under $7 Million and grew exponentially to approximately $139 Million by 2003. It is likely that now (2006) the face recognition market revenue exceeds $170 million worldwide. (Based on figures from Biometrics Market Report, 2003)

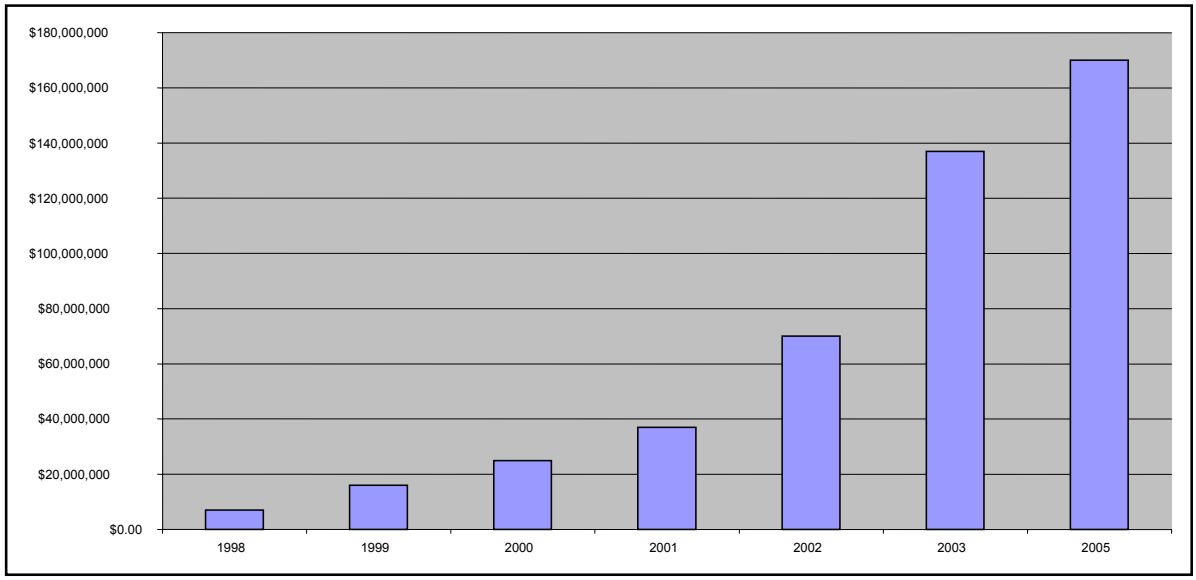


Figure 3: Revenue forecasts for the face recognition market 1998 - 2005

Face Recognition systems are typically based on having full frontal portrait (i.e. passport) images, taken at the time of enrolment against an uncluttered background, and later comparing these to similar photographic images acquired from still or video cameras. Such images are currently used for passports, driving licences, physical access control, and photo cards as a means of identification by close to 95% of all professions. Therefore, the development of the face recognition market is strengthened by the fact that a supporting infrastructure is already in place that makes the technology well aligned for being widely adopted.

Historically, face recognition has been out run by the enormous market growth in fingerprints. The fingerprint market is not only one of the largest in terms of revenue, but also in the broad spectrum of applications in both the civilian and law enforcement domains. Fingerprint identification is very well established in policing. However, despite the various societal and privacy issues surrounding the biometric being associated with criminality, uptake of fingerprint technology (ten-print identification searching or single finger verification) is strengthened by accuracy figures, and by existing national fingerprints systems that enable the fingerprint data and applications to be used in a wider sense. For example, roadside identification, ID cards, passports, visa applications are all currently looking to exploit the accuracy and availability of fingerprint data and low cost technologies, both in the UK and internationally.

Face recognition has a clear application in the civilian domain that is primarily driven by public acceptability of people using their face as an identifier and the ease of use of the technology (people are used to looking at a camera). However, the mass market in face recognition technology is currently limited and more suited to civilian applications, where the images captured will generally be of a standard quality, the majority of users will be compliant and the method of capture and environment will be controllable. As such the majority of face recognition applications are fully automated to facilitate the needs in the civilian

domain where operators are likely to be non-expert, have a large throughput (to be able to process large numbers of people quickly) and need to have a very low false accept rate (FAR)[6].

However, the technology must develop further to cope with unconstrained environments where lighting, background or the user's pose may vary considerably. Only then will vendors be able to take full advantage of the widespread availability of cheap digital cameras, integrated with desktop computers, mobile phones and visually enabled PDAs that have the potential to give facial biometrics a significant edge over alternative technologies.

Within law enforcement, the forensic use of such technology is very different in that a degree of manual intervention is acceptable in most scenarios. Images may be marked up by experts, the technology might be complimented by visual aid tools, and the return of false matches is acceptable to a much larger extent than for civilian applications. In short, Face Recognition technology has more potential to be used as an aid to filter information, or provide final lists of individuals with facial images, which can then be manually verified by expert operators. Currently there are very few vendors with face recognition applications deployed within law enforcement agencies where the technology has been specifically developed, or is able to function, in this manner.

There is also a need for research into the cognitive side of facial comparison and identification. Very few vendors of face recognition technology currently appear to be putting serious effort into this aspect of their system. Certainly more focus on cognitive approaches to the way faces are processed by humans may help to develop visual aids and UI tools etc. that may assist operators with analysis of faces, both for forensic purposes, or when required in real time (e.g. at immigration). This will inevitably impact the implementation of the technology and the performance required/expected by end users.

Until the technology is further developed to cope with unconstrained environments where lighting, background and pose vary considerably, the market growth and investment for facial recognition applications will be limited for law enforcement. However, some immediate gains could potentially be sought if vendors and researchers focussed on the development of applications that assist "facial identification", whereby the technology is forensically motivated in its functionality, operation and user base, in much the same way that AFIS applications are designed to support the work of fingerprint experts, rather than replacing them with a fully automated system.

It is reassuring that, although very few in number, there are some interesting developments where companies have begun to explore this area. CyberExtruder, Animetrics and XID, originally marketed at the video graphic and computer games industry, have recently developed applications exploiting 3D imaging processes that are aimed to help examiners visualise facial image data. L1 Identity Solutions (which has now taken over Identix, Viisage, Iridian and Securimetrics) recently launched a system specifically geared at forensic examination of facial image data. Unlike many other distributors or integrators, Identix is able to take full advantage of having direct access to the algorithm source code such that the input provided by operators on the Identix application interface can directly feed into the algorithm, thus potentially improving its accuracy.

[6] False Accept Rate (FAR) is term used to measure the number of instances that a system incorrectly returns a match for two biometric samples that are in fact from different individuals. (See Glossary)

3 Face Recognition: Technology

3.1 A Generic Application: How does it typically work?

The generic principle behind most face recognition applications is to capture a facial image from either a still or video camera and launch it as either:

- an automated one-to-many search against a database of similarly captured images (*identification*); or
- perform a one-to-one automated comparison of the captured image (a 'claimed' identity) against an image previously registered on the application's database (*verification*).

The overall functions of a generic application can be split into the following main processes: image capture, specification of eye co-ordinates, normalisation, feature extraction/encoding (templification), matching and decision. The diagram below shows the flow of data within an application.

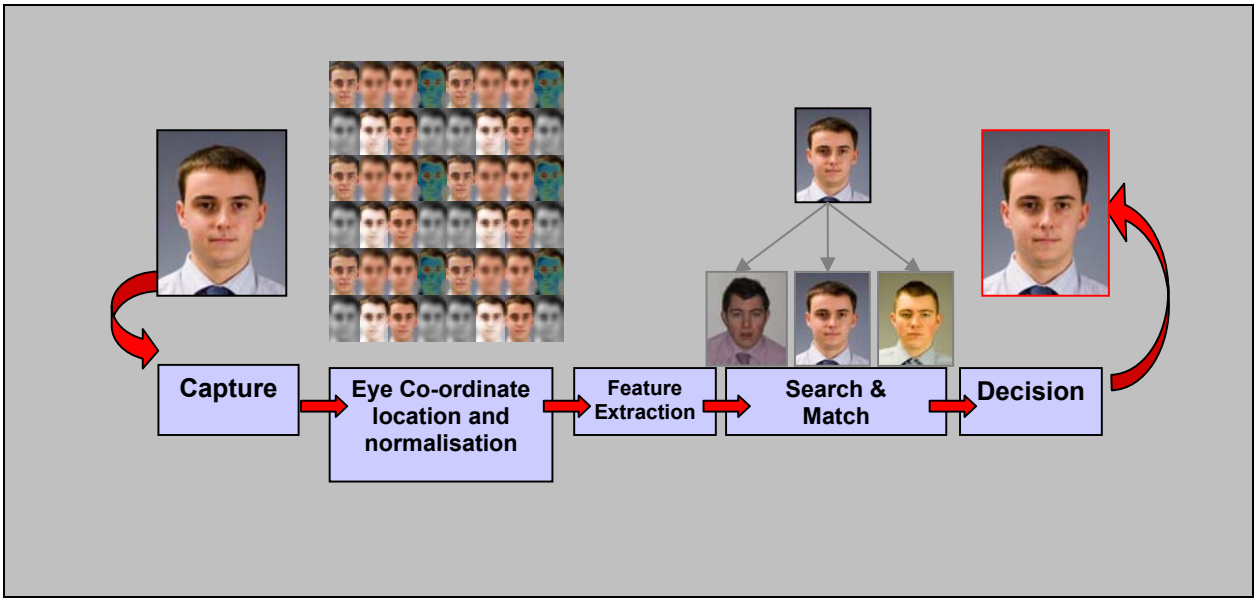


Figure 4 Generic Application Process

3.1.1 Capture

Most face recognition applications are designed to work with standard digital cameras, web cams or still images captured and extracted from live video feeds. The person using the application is typically required to pose for a frontal image to be captured, generally standing two to three feet away from the camera. Needless to say, the better the quality of camera the more likely it is that the image will be of good quality, with greater robustness to variations in indoor lighting conditions. This helps to maximise the performance of the face recognition application. Most face recognition applications will also accept scanned or other facial images captured from other media into electronic format. (e.g. printed passport photographs) However, such images will typically require manual mark up of the eyes and alignment in order for the application to be able to use them.

The use of Live/streaming video for capturing images is becoming increasingly popular amongst face recognition vendors. There are several performance advantages in using images acquired from such sources:

1) Multiple still images of an individual can be captured and an operator or the system can then choose the best quality images to perform the search or comparison

2) The fusion of multiple images to produce a single image which is then used for comparison can provide significant improvements in the accuracy of matches

3) Use of live video to obtain images can often improve usability of the application when acquiring image data and speed up this process too. It helps to minimise common data capture errors that arise from individuals either moving at the time a photograph is taken or for those that find it difficult to keep a neutral expression.

4) It is also provides the opportunity to incorporate "Liveness Detection"[7] measures – for example, by aid of voice, expression, gesture analysis, etc.

Figure 5: Example of face recognition with live capture – NEC's Neoface application

3.1.2 Eye co-ordinate location and normalisation

Most applications will require the eye co-ordinates to be detected and, for some, the tip of the nose too. This is normally an automated process but may require manual intervention if the software cannot accurately locate the features (perhaps as a result of glasses, heavy eyebrows, moustache, beard etc). During this process the image can also be cropped, rotated for alignment and scaled. Normalisation of the image across pixels to dispel the effects of sharp illumination will also be performed. Inevitably images captured in very controlled conditions with fixed pose and lighting help to make this process easier. In particular, controlling pose can help automate things like eye co-ordinate detection, thereby minimising the need for operator intervention.

[7] Liveness Detection is a process to verify the authenticity of the sensor is detecting an image or biometric sample from a live, human being. This process is employed at the time of capturing an image.

3.1.3 Feature extraction

This part of the process is generally fully automated and dependent on the methodology of the matching algorithm. The purpose is to extract key information (features) that can be used for automated comparison and matching. For face recognition these are more than just eye, nose or facial landmarks; they can also be pixel level information about a region in the facial image. Indeed, it would be wrong to think of "features" as physical entities at all in terms of the way Face Recognition systems typically work. Some applications create holistic representations of the face based on the high frequency pixel data; others use local patterns around certain salient features (fiducial points) of the face. In short the method of feature detection and the "feature" that is detected is reliant on the approach used for matching. These are referred to below and described in detail in the Appendix.

3.1.4 Matching

The matching algorithm will typically work in either verification (one-to-one) and/or identification (one-to-many) mode – the difference being that in identification, the image acquired is always searched against a database of pre-enrolled images that are stored for comparison, whereas with verification, a comparison is made to a 'claimed' identity, perhaps using an image that is stored locally on a card.[8] All of these images will ideally have been captured as specified above in optimal conditions and encoded as described previously.

It is in this part of the process that the source and "quality" of the image data used with the application has the greatest effect on the performance of the system, and hence the most significant impact on where and how the technology can be successfully deployed.

3.1.5 Decision

The output of the matcher will yield a result whereby pairs of images are classified as matches or non matches based on their similarity / difference. Typically, these decisions are dependent on thresholds which can be configured to suit certain population biases of data, tuned for certain individuals and for specific operational scenarios. The output of the match will usually have a score provided with it. However, this score is application specific and heavily reliant on the threshold setting of the application. Analysis of the matcher score may help inform the operator in the use of an application, but it is not a means by which to compare the performance of two applications.

3.2 Algorithms

Face (or local feature) detection, classification, and matching can be considered as being three main processes in face recognition matching. There are many different approaches that have been researched and employed in developing applications; each one has its strengths and weaknesses for each of the three aspects. It is typically found that most mainstream application algorithms are based on similar underlying (*mathematical and information theoretic*) concepts and can also be seen employed in other areas of Digital Imaging and Pattern Recognition. However, each one will vary in the approach and the process for which different algorithm principles are adopted. Most commonly used methods are Principal Component Analysis (PCA), Local Feature Analysis (LFA), Elastic Graph Matching and Eigenface Bases, to name a few. The Appendix of this report provides more information on these and many other algorithmic methods used by modern day Face Recognition. All of these techniques are extensively documented in academic literature and many textbooks on Pattern Recognition.

[8] This is commonly referred to as a Match on Card application where a biometric (in this case facial image) template is stored locally on a smartcard.

There are some generic challenges in face recognition that affect the computation and complexity of algorithms; two of which are the levels of variation between and within faces and the high dimensional nature of the facial feature space. It is important to understand these concepts in order to appreciate the pros and cons of the difference algorithmic approaches overviewed in this report.

3.2.1 Intra-class Vs Inter-class Variation

A set of faces of different individuals will have very many similarities or features in common. For example, all human faces will have eyes, ears, a nose and a mouth as well as other parts of the face with high levels of similarity. This is referred to as low 'interclass' (*or between-class*) variation for faces; this is the level of differences between two different classes of face. However, faces belonging to the same individual can vary significantly too; this is referred to as 'intra-class' (*within–class*) variation. As a result classifying faces correctly (or in other words identifying the correct class of individual the face image belongs to) can be difficult and prone to high degrees of error. Algorithms for face recognition require this classification error to be minimised. The variant nature of faces is exploited by using principles that focus on the difference in faces as opposed to similarities. Many Face Recognition algorithms exploit the most discriminating features as an approach to classification (PCA, Eigenface, LFA etc).

3.2.2 High dimensional feature space

It is important to understand that "features" in this context do not necessarily map to physical features on the face; rather they are better considered as mathematical entities / constructs derived from processing pixel level information in the image.

A face can be represented in a feature space of multiple dimensions, with each dimension being a form of variation for each instance an image is captured. Consider for example a digital image of a face captured as 430 x 640 pixels. Each pixel can be considered to be a varying feature of the face as a result there are 430x 640 dimensions to consider. The high levels of similar or generic attributes between differing faces means that there are numerous variables and directions in which attributes can vary, (e.g. as a result of changes in angle, pose, lighting, etc). Thus face recognition is a high dimensional problem both at the individual pixel level and also at the level of intrinsic information within the face. In this form it is an extremely complex operation to perform. Algorithmic approaches typically, therefore, first look to reduce the number of dimensions and hence reduce the complexity of the algorithm.

3.3 3D Face Recognition

Until very recently 3D face recognition was considered to be too complex a problem to solve. The problems mainly arose (and still do) in the capture of a 3D model of the face, as this process is extremely sensitive to misalignment or any slight movement by an individual during acquisition of the 3D image.

However, it is acknowledged that 3D recognition potentially offers several advantages over traditional 2D techniques. Firstly, in a 3D geometric representation of the face there is a significant amount of depth information about the face which is absent from the flat, 2D images of the face; second, the geometry of rigid features of the face (i.e. nose, chin etc) may be used to make recognition techniques more robust to common issues such as lighting, facial expressions, make up, and head orientation that hamper the performance of 2D.
It is important to understand that in this context the term 3D refers only to the ability to capture and store depth (or range) information about a face. Many 3D applications do not capture or use a conventional 2D image at all, and if viewed on screen will merely produce a 'mask', or 3D 'mesh' which is of very little use for human aided identification. In some applications a 2D image is overlaid on this 3D model to create a

true 3D representation of the face, which can then be viewed from multiple angles. This is often referred to as 5D, and this approach is the one that is most likely to be of use in forensic applications.
It may also be possible to use a 3D model to improve the accuracy of traditional 2D image based recognition. One might be able to extract a 2D image from a 3D model of the face from a specific view in order to compare it to another image of a face taken from the same view point. (Dynamic Vision). Indeed, many "3D" systems merely use the additional information to assist in 'normalising' the face (i.e. rotate to frontal image, correct lighting etc.) in order that a conventional 2D matching algorithm can then be employed. A common argument for adopting this approach is that there are numerous 'legacy' databases of 2D images already in existence against which searches could be launched.

Other applications attempt to overlay a 2D image onto a generic 3D model of a face, in order to create a pseudo 3D image, which can then be manipulated (either manually or automatically) and normalised to correct for pose, lighting etc. This is sometimes referred to as 2.5D, and may well have application in forensic face recognition, although the validity of this approach as a means of improving search accuracy has not yet been proven.

The relatively recent development of commercial 3D applications and the current lack of agreed standards for capturing and storing the data, means that there is no market at the present time for developing applications that can search large 3D databases. However, work is now going on within ISO to draw up international standards for 3D image data capture and storage, and this, coupled with improvements in capture techniques, could result in many more 3D applications coming on to the market in the next few years.

From the above, it can be seen that currently 3D face recognition is still an open research field, though several vendors already offer commercial solutions. Most methods of 3D model representations are based on reconstructing the structure of the face using structured light and/or stereo cameras, and these approaches are explained below.

3.3.1 Structured Light based approaches

The use of structured light to illuminate the face with a known intensity pattern has proven to be very successful in generating face models. This approach allows conventional cameras to be used to measure a reflected pattern of a light signal plane/pattern (e.g. lines, circles or sinusoidal) that is projected onto the surface to be modelled. The distortion of the pattern provides the range of depth over the surface.

Infra red (IR) light is cost efficient and it has proven to be very robust to variations in illumination for indoor capture environments. However, there is still a serious limitation to its use outdoors, due to the higher level of ambient light which then requires a much higher intensity IR projector to be used.

3.3.2 Use of stereo cameras

Stereo visual or video methods are similar in concept to human binocular vision, in that two cameras or two video channels are used to obtain two images from which distance information can be obtained. The way that machine stereo vision generates the third dimension is achieved by finding the same features in each of the two images, and then deriving the distances to these features by triangulation; that is, by intersecting the lines of sight from each camera to the object.

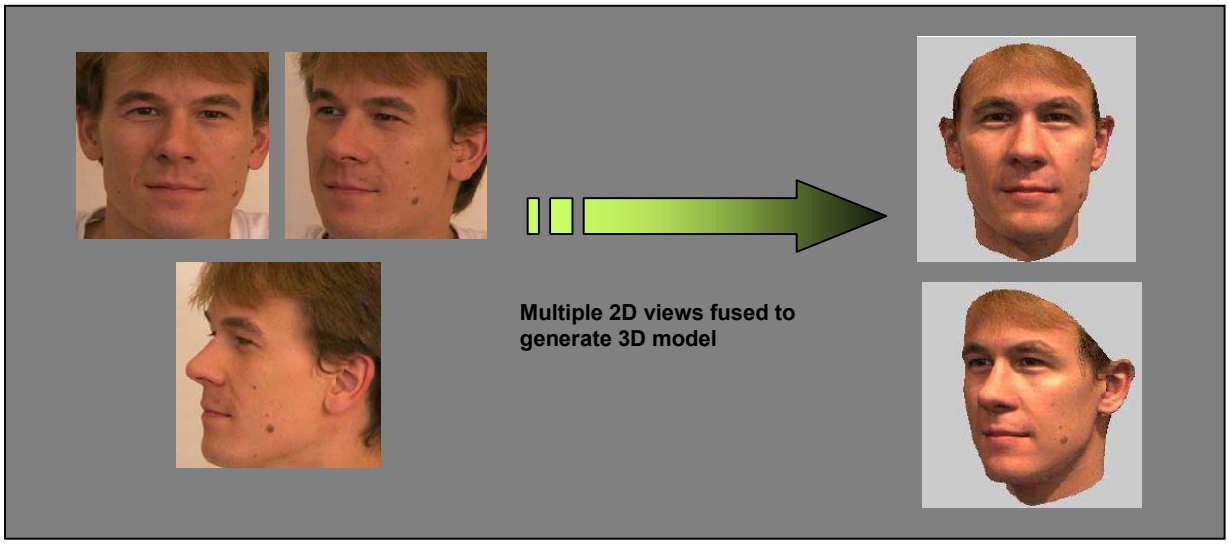


Figure 6: Multiple 2D views to generate 3D model

Although this can produce a dense depth map of a scene, the computed depth values may not be very accurate. Structured light on the other hand produces more accurate depth values, but it is very restrictive. However, with recent advances in video, calibrated stereo-visual methods for 3D images, commercial solutions have been able to implement depth perception by combining binocular stereo and structured light. Projecting a grid onto the face and integrating video capture of it into a high resolution 3D model allows for better capture and thus good recognition accuracy with low cost off-the-shelf components. (Dynamic Vision).

Another approach to 3D is to avoid the need to explicitly reconstruct the 3D model but to instead make use of multiple two dimensional views together with dense correspondence maps between them. These multiple views are then used to reconstruct a full 3D view of the face. However, faces are non rigid objects and suffer a degree of deformation which can result in errors in the representation.

4 Face Recognition Applications Areas

This section investigates the strengths and weaknesses of Face Recognition technologies in meeting police requirements. Four broad categories typifying the various scenarios in which face recognition technologies might be used are considered. These are:

1. Controlled Facial Image to Controlled Facial Image
2. Uncontrolled Facial Image to Controlled Facial Image
3. Controlled Facial Image to Uncontrolled Facial Image
4. Uncontrolled Facial Image to Uncontrolled Facial Image

Terminology:

- A "***Probe***" image refers to an electronic facial image that is captured and presented for matching. This is the image that will either be searched against a database collection of facial images to find a match, or verified against a given image thought to be a match to the probe.
- A "***Gallery***" is a database collection of electronically stored facial images that is referenced by an application. This collection is searched to find possible matches to a probe image or to verify a claimed identity. An image that is stored in the gallery collection is referred to as a "*gallery image*".
- A **"Controlled"** image is a facial image captured in a controlled environment (e.g. passport)
- An **"Uncontrolled"** image refers to ad hoc or random capture of facial images in an uncontrolled environment (e.g. CCTV)

Note: For each of the following categories the potential uses of face recognition refer to 2D face recognition only. 3D and multiple view based face recognition approaches are described in a later section.

4.1 Category 1: Controlled Facial Image to Controlled Facial Image (Mugshot to Mugshot)

This category represents the conventional approach to face recognition and includes the vast majority of civilian applications. Facial images for both the probe and gallery are captured in a controlled manner. The purpose of controlling the capture is to ensure that only the best possible quality images are used with the application. Such images often require capture to a certain specified standard, whereby the lighting, pose, expression, distance to camera, etc. are defined (e.g. ICAO standards for travel documents). This helps to minimise the effects of variance that these factors have on the matcher performance.

4.1.1 Scenario

Applications aimed at matching controlled portrait image to controlled portrait image are particularly suited for access control and passport applications. Civilian applications such as these will typically operate in a verification mode. They are also well suited for law enforcement in the case of mugshot to mugshot image matching, both in verification and identification modes, provided that the mugshot images are of sufficient quality. (See diagram in Figure 7)

4.1.2 Strengths

- ***Availability of Technology***

  The face recognition technology market is very well established in this area, particularly for civilian applications. Recent benchmarks have shown that the accuracy of face recognition technologies when operating in a verification mode is very high and can be on par with that of fingerprints when high quality images are used.

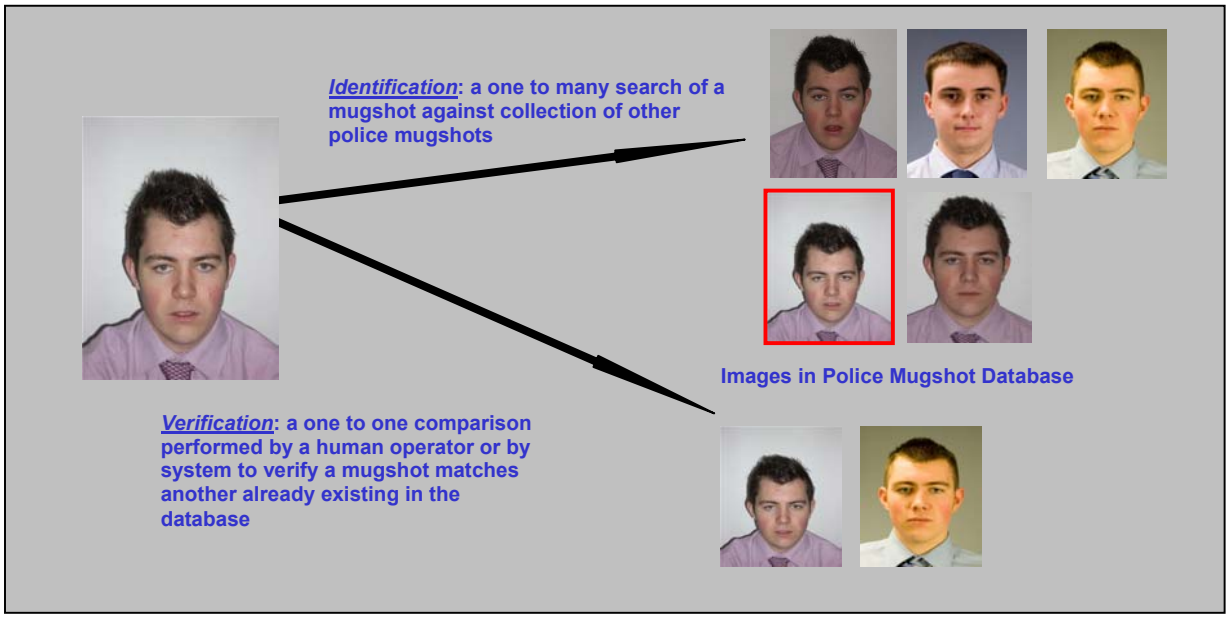


Figure 7: Scenario diagram of Mugshot to Mugshot searching

- ***Documented Accuracy***

  Some implementations and trials of face recognition technology have reported the technology to have superior accuracy to that of humans. Although we are very good at recognising persons known to us, we often fail to accurately distinguish unfamiliar faces. In situations such as immigration/passport control where officials need to compare large numbers of unfamiliar individuals to their passport photographs, the technology can offer clear benefits; it does not suffer from preconceptions, bias or fatigue as do humans.

4.1.3 Weaknesses

- ***Background Noise***

  Coloured or cluttered backgrounds adversely affect the matcher. Thus probe and gallery images have to be acquired, or electronically edited, to have plain (typically 18% grey scale) backgrounds. Algorithms that use global image features are particularly sensitive to background noise. Often features may be present in the background which will not be present in the reference image which will cause algorithms to mismatch images. Although local feature based algorithms offer a better degree of robustness they are still adversely affected. Some vendors have got around this by closely cropping images around the face to eliminate as much background data from the image as possible.

- ***Algorithmic Robustness***

  Very few algorithms are robust to the effects of variations in age, gender, ethnicity, appearance, occlusion or expression. For optimised performance these variances have to be controlled at the time of image capture, and yet this is only possible for the latter three that are mentioned. Global feature based algorithms are sensitive to variations; however, more localised approaches can offer some performance improvement. Those that use elastic bunch graph approaches that filter variations at nodes (fiducial points) on the face are shown to provide some robustness to expression and pose. Some Eigenface/PCA based technologies have also demonstrated invariance to significant differences in age. Moreover, the face is often similar in pose, expression and there is little occlusion or variation in appearance.

- ***Normalisation and Manual Editing***

  Accurate location of the eye co-ordinates is fundamental to the performance of the system and often has to be performed manually. For more standardised images automated eye co-ordinate location is generally found to be quite accurate but in most cases still requires manual checking (for example, thick eyebrows or spectacles can confuse the software). For probe images this is simple as it is quick and requires no operator expertise, however, for batch processing of facial images in the gallery this is a weakness and is most often application specific.

- ***Circumvention***

  For typical civilian access control scenarios the use of the technology is simple and well suited. It is relatively difficult to spoof as it is difficult to occlude the face when live capture of the probe image is required. (Note, however, that the technology is generally not robust to the use of photos or disguises such as make up, beards, wigs, etc.) The majority of users in the civilian domain are compliant and co-operative as it is in their interest for the system to work well for them. On the other hand, for a small subset of the population in the civilian application domain, users will go to extremes to circumvent the technology. Dramatic changes in appearance such as plastic surgery present considerable performance challenges for the system. There is no data on the performance of systems in this case.

- ***Intra-class Separation***

  Facial recognition is not robust to distinguishing between identical twins. However, there is an argument that for the performance of the technology this is not an issue. Systems that could differentiate between twins may be too sensitive to "intra-class" image variations and thus suffer from high false rejection rates, whereby two images are in fact of the same person.

4.1.4 Opportunities

- ***Watchlists***

  Searching in this context in an identification mode may be suitable for both the use of police mugshots or civilian portrait images. For the civilian domain, capabilities such as searching an applicant's passport image against a gallery of known fraudulent applicants in a watch list will provide significant operational gains in securing borders/ points of entry into a country.

  Searching mugshots against a watchlist can assist officers in many ways. For example, early and rapid identification of prolific/repeat offenders, violent offenders when brought into custody could be readily identified by using face as a crosscheck against a watchlist, held either locally or shared from a central facility.

- ***Point of Entry Access Control***

  A gallery of mugshots may provide valuable intelligence for closed operations such as football matches where certain individuals have been previously identified as troublemakers. Face recognition deployed on entry to the ground may assist both the management of the point of entry and also provide intelligence to the officers securing the grounds.

- ***Continuous on line verification***

  Continuous monitoring of a face may provide certain opportunities. Many facial recognition technologies that are used for access control (logical/ physical) currently operate with video cameras that continuously capture facial images. When used on a workstation or mobile platform, having first established the user's identity, the system can continue to monitor the video stream for changes. This can be used to ensure that the same user is present for the entire on line session which may be particularly useful where the user's identity must be authenticated throughout it. (e.g. court appearance, real time video evidence, computer logon, ATM access, etc.)

- ***Rapid probe of other mugshot galleries***

  Mugshot to mugshot recognition will allow forces to rapidly probe each others' mugshot galleries. Although this might be achievable through searches based on the Criminal Record Office (CRO) number, for images which have no available fingerprints or information on PNC, forces currently have no way finding mugshots that might belong to the same person other than by manually checking individual records. Therefore, mugshot to mugshot searching could enable duplicates within force collections to be identified quickly and efficiently, returning a small number of likely matches for human examination.

  Note that this is largely reliant of the quality of data. For example, for new images on FIND, rapid probing of forces' collections will enable one to determine if the person is previously known to the police, as long as the legacy data is compliant to the minimum image quality requirements as specified by FIND. However, in many cases the force's legacy data, predating FIND, is of poor quality and unsuitable for use with face recognition. Thus, in these situations, individuals who may have had

their mugshots captured in the past may not be identifiable in this way when brought into custody at a later date.

- ***Accuracy Requirements for Law Enforcement***

  One can argue that even at low levels of accuracy, automatic Face Recognition systems are of considerable value if a search yields a result on a case where no other intelligence leads are available, particularly so in the case of serious crimes. Thus, the accuracy requirement of mugshot to mugshot searching is very different for policing than for the civilian domain.

- ***Witness Identification***

  Mugshot to mugshot searching may be useful for generating witness albums or for obtaining facial image data to assist a witness in building a composite image. The technology may be used to automatically generate a list of similar looking candidates to add integrity and eliminate bias when compiling line ups or video identification parades.

4.1.5 Threats

- ***Legacy Data***

  Prior to implementation of FIND and the FIND Image Standard, mugshot image quality across the UK police service has been extremely varied and often poor. One cannot guarantee the performance of face recognition when used for mugshot to mugshot searching with such images and so for some considerable time to come forces will have a large volume of legacy images that may not be searchable. Therefore the full benefits of this capability may not be available/ realisable to the police service until after the FIND database is sufficiently well established.

- ***Managing Expectation***

  Face recognition operating in this manner can vary in performance based on the mode of operation: verification or identification. The source of image data will also impact on the performance of the technology. Note also that as explained above the accuracy achievable in policing will not be as high as in the civilian domain. Whilst this does not negate its potential benefits, it may result in a false understanding of what accuracy to expect and thus poor perception of the technology. This could undermine the value of introducing this technology into the police service. Therefore the expectation of accuracy has to be monitored and managed very carefully amongst users and the media. Vendors have historically made extremely optimistic and often unjustifiable claims regarding the accuracy of their products. (This is true of the biometrics market generally, not just face recognition). PITO's Biometrics Team therefore has an important role to play here to investigate and understand the potential and the realistic accuracy that can be expected from the technology for various police requirements. This knowledge has to be communicated as widely as possible to the police service, wider CJS and industry in order to manage the expectation of this technology and to drive its future development. The current FR Evaluation and Demonstration project is a key area in which the team is already influencing the development and implementation of face recognition technology for the police service.

4.1.6 Suitability for Law Enforcement

- ***Non co-operative Subjects and Variation in Head Position***

  Often, within law enforcement, users will not be fully co-operative. In a custody suite capturing a mugshot (that is also FIND Compliant) from a non co-operative individual can very be difficult, and thus there is little guarantee that the acquired image will be of an optimal quality. To get around this some vendors have deployed systems with PTZ (pan, tilt and zoom) cameras that can automatically zoom in on an individual to capture a suitable image. However, some individuals may be under the influence of drugs, alcohol, or simply be in a condition whereby they are physically unable to position themselves in front of a camera for a mugshot to be captured. In these instances, the effectiveness of a PTZ camera will be limited as such images will still contain variation in the position and orientation of the person's head.

  Work by PITO's Biometrics Team has suggested that 'Head Pitch' (or the vertical displacement of the individuals head) has an effect on face recognition performance. Recent reports on Face Recognition from other Home Office pilot evaluations have also reported that very small perturbations in the vertical alignment of the head appears to have a **significant** affect on face recognition performance, more so than variations in head pose from right to left. This has a huge implication on the use of face recognition for mugshot to mugshot searching and the accuracy that can be expected in such circumstances.

- ***Typical Policing Environments***

  Other than in a custody situation, the police rarely operate in controlled environments. The potential for using Face Recognition technology under these circumstances is therefore limited, and in many cases other means of identification, such as fingerprints, may be more appropriate. However, there is value for the police to be able to automatically probe/search against collections from other agencies such as UKPS, Asylum seekers, and DVLA where fingerprint data may not be available. The adoption of international standards for facial image capture by these agencies is an important pre-requisite for such searches to be carried out effectively.

- ***Composites***

  There is considerable interest from the user community in using automated face recognition to search composite images against mugshot databases such as FIND. Indeed some forces have already tried this and claim some degree of success, in either identifying a perpetrator, or providing valuable investigative leads.

  However, current approaches to composites such as EFit, ProFit, and sketches, are not particularly well suited to use with face recognition, due to the manner by which they are produced.

  New approaches to composites such as Evo Fit and EigenFit, are based on similar principles to those employed by Face Recognition algorithms (namely PCA) and therefore it may well be that such images are more compatible for use with the technology, and may thus produce better results. This is an area that requires further investigation, but such research is inherently difficult to do due to the method by which such composite images are created (i.e. images recalled from memory).

4.2 Category 2 & 3: Controlled Facial Image and Uncontrolled Facial Image (Mugshot ↔ CCTV)

The largest source of facial image data available to the police for investigative purposes is from CCTV cameras. Furthermore, the civilian domain also has widespread access to such data. Categories 2 and 3 exploit these sources of data for use with mugshot or controlled portraits as defined for category 1. The issues surrounding these categories are very similar thus these shall be addresses together.

For category 2, facial images acquired under controlled conditions are used as probe images for comparison against a gallery of images that have been sourced from uncontrolled environments or ad hoc situations. The use of automated face recognition to link known suspects to images captured on CCTV would be extremely advantageous for both law enforcement and security applications in the civilian domain. For policing this capability will enable one to link known individuals to unsolved crimes.

For category 3, probe facial images are acquired or extracted from video data typically in conditions that cannot be controlled. These are used for comparison against a gallery of images acquired by controlled methods. These applications are most likely required to operate in an identification mode and allow use of multiple sources of facial image data. For policing such a capability will allow one to identify suspects from crime scene footage.

4.2.1 Scenario

Face recognition applications provide the most significant advantage when used forensically to identify individuals thought to have been captured on CCTV. (See diagram in Figure 7)

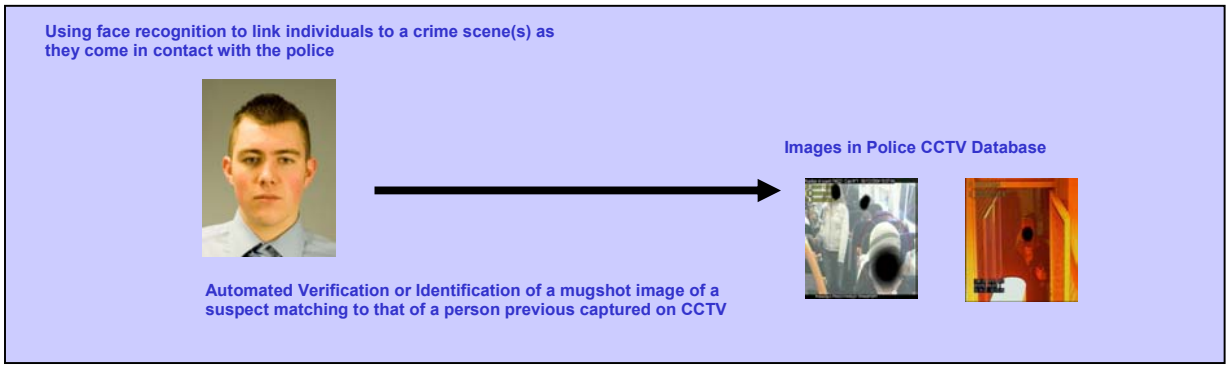


Figure 8: Diagram of Category 2 Scenario Mugshot to CCTV searching

Similarly, being able to search facial images from CCTV may result in technologies for real time surveillance and identification of individuals. Typical application areas include airport/border surveillance, monitoring of crowds (sports stadiums), or post event analysis applications for the forensic identification of offenders that have been caught on camera. (See diagram in Figure 8)

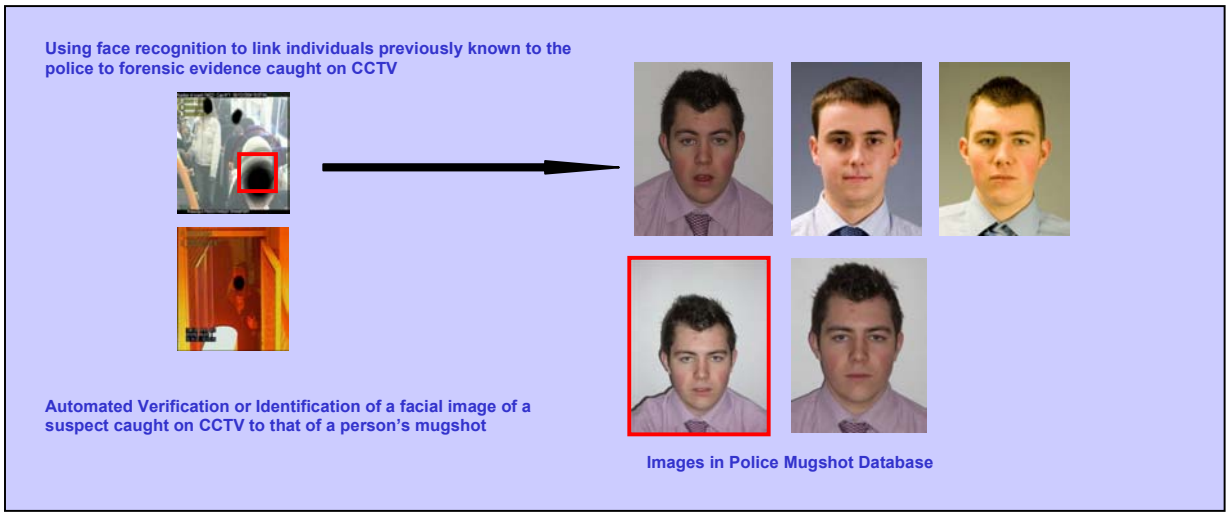


Figure 9: Diagram of Category 3 Scenario CCTV to mugshot searching

4.2.2 Strengths

- ***Abundance of Data***

  The UK currently has the most widely deployed CCTV coverage in the world with approximately a ratio of 14 people to each camera, and ongoing work to develop a national strategy for CCTV could result in a doubling of this figure to 1 camera per 7 individuals. Thus, there is an abundant source of facial image data – a resource that is currently undermined by the lack of available technology that would allow effective and efficient means of utilising this source of information – both in terms of storage, indexing and analysis and searching. The ACPO CCTV national strategy now being developed is likely to result not only in more cameras but also in increased image quality. The strategy is being developed in conjunction with the police, industry and academia and the potential for automated face recognition is being taken into account.

- ***Technological functions that are currently available***

  Technologies for automated face detection and tracking from video are currently available and found to be generally effective, however, the face must be visible and (ideally) directly facing the camera.

- ***Legacy infrastructure or data***

  There is an existing infrastructure of portrait or mugshot data available for probing CCTV sources or galleries. Furthermore, with increasing use of face recognition applications for access control and mugshot to mugshot searching such images are widely available in electronic format.

4.2.3 Weaknesses

- ***Randomness of CCTV data***

  Typically, facial images from CCTV are of poor quality with variations in pose, illumination, expression, angle and distance to camera (to name a few). The backgrounds are often cluttered and the subjects clothing can also present image recognition issues. These factors impact the performance and

applicability of face recognition technologies being used for comparison of facial image data sourced from CCTV/video footage.

- ***Current usage of Facial Image Sensors***

  Mostly all 2D face recognition applications are able to take facial images from any video source. To extract a facial image from CCTV may cause problems in additional degradation in quality from the conversion of video format to still image format, and also in enlarging the image for facial recognition and subsequent operator analysis.

- ***Accuracy and performance***

  The current performance of the technology has been very poor when used with such images. Some vendors such as Aurora (currently using an Identix G3 search engine) have reported robustness to some of these factors such as illumination, lighting and pose to a certain extent (up to 15º off centre). However, it is likely that the images that were used to determine this would generally have been compared to very similar images in the gallery set or have little variation in other respects. CCTV images will more often contain multiple factors that pose a challenge to the performance of the technology. Moreover the images with which they are compared (e.g. controlled portrait/mugshot) will be very different. For this reason local feature based techniques often provide slight improvements over global image based techniques to face recognition. However, most face recognition algorithms are fundamentally based on image matching. Therefore, whilst some technologies may display robustness to certain types of images which have a high degree of similarity to the image in the database, the performance with the global data remains poor.

  Recent work by UKPS and others has also identified a problem with performance when the head is tilted. Since the vast majority of CCTV cameras are positioned above head height this will have a very detrimental effect on the search accuracy for mugshot to CCTV searching and vice versa.

- ***Confidence of matches***

  The current state of the technology is poorly suited to civilian domain applications where the accuracy requirement is high with a small rate of false identifications and there is a need to operate in real time. Contrary to the face recognition scenarios described in category 1 above; for comparing CCTV/ other uncontrolled images, the operator will not be able to rely on the match given by the face recognition application to inform their judgement. Where a match is found it might be returned in a very low position in a 'respondent list' in which case (depending on the circumstances) it may not even be viewed by the operator. Furthermore, when a genuine match is returned by the system, it might not be apparent to the operator that they are the same person. With the accuracy of the technology being so poor it is likely that it will be ignored when making a final decision on confirming a match. Forensic use of the technology may differ from civilian applications in that low ranking matches may be analysed off line. Improvements in accuracy may be achievable if the operator spends more time in extracting and marking up an image prior to searching. However, most face recognition applications currently have little by way of added functionality to assist with enhancing images to improve matching or verification, nor do they currently support significant input by the operator, other than in marking the eye locations.

- ***Moving towards Skin Texture Analysis***

  Many face recognition application providers are looking to exploit skin texture detail to improve the performance of their systems. However, these approaches are reliant on good quality images for mugshot or portrait capture and are more suited to access control applications. Evaluations of face recognition have shown that it performs better on images that are like for like in quality; that is, a low quality image is more likely to match against an image that is of similar poor quality. Matching a poor

quality (i.e. low resolution) CCTV image to a good quality portrait/ mugshot may suffer from redundant information or noise which adversely affects the performance of the matcher.

- ***Incompatible images for comparison***

  One of the main obstacles in the use of face recognition for uncontrolled images is the varying angle or pose at which the facial image is acquired. Both eyes have to be clearly visible as current matching techniques require the eye co-ordinates to be located in order to normalise the image. Moreover these images must be matched against a previously enrolled portrait or mugshot image which further constrains the use of images from CCTV footage. There have been some developments in using multiple pose generated mugshots/ portraits (commonly referred to as 2.5D) to overcome pose/angle constraints. However, these methods suffer from poor alignment between the probe and gallery images.

- ***Occlusion***

  The most obvious constraint with uncontrolled capture of facial images is occlusion. If people do not want to be recognised on camera then they can simply wear a mask or ensure that regions of the face are occluded by sunglasses, hair, beards and other dramatic changes in appearance. More simply, knowledge of where a camera is located may enable one to simply position oneself facing away from the camera. On the other hand, whilst most people are aware of the use of fingerprints in forensic investigations they are still commonly left at crime scenes, providing valuable information to investigating officers. It seems likely that with so many CCTV cameras now in operation, with suitable searching tools, a usable image (without occlusion) could be found in many cases, with its value mainly being limited by other factors such as resolution, lighting, angle etc.

4.2.4 Opportunities

- ***Mobile Applications***

  Lantern will shortly provide the capability for mobile fingerprint checks in support of Automatic Number Plate Recognition (ANPR) technology. The fundamental purpose is the “identification of individuals at the point of contact” and the “rapid identification of persistent offenders”. This capability can be readily enhanced by the return of a mugshot to act as a manual check of the returned respondent. Furthermore, inclusion of a camera on a Lantern unit may allow for remote or local automated face recognition to act as a secondary identifier where fingerprints cannot be captured, or for searching against locally stored watch-lists of known/prolific offenders. If successful this could have clear business efficiency improvements for roadside identification. LAPD have already deployed an application like this for identifying gang members using local watch-lists held on PDA’s equipped with cameras and Face Recognition software.

- ***Improvements in Video/CCTV enhancement***

  Video enhancement technology needs to develop to provide improved quality video footage to retrieve higher quality facial images from CCTV. This is an ongoing research area, looking at ways of fusing multiple low quality images to provide a single high quality image, or to infer 3D information from multiple 2D views of an individual from varying angles.

- ***Efficient Processing of CCTV/Video evidence***

  Typically, operators spend hours trawling through video footage before finding any information of value. Furthermore, in order to use Face Recognition with CCTV images the faces must first be detected and extracted from the footage. Therefore, if more efficient methods of automated processing, indexing and archiving of such data were available and used, it would provide an opportunity whereby Face Recognition may be used to identify the persons caught on camera. Furthermore, face detection technology itself may assist in this process by detecting faces within footage (as opposed to matching) such that all faces can be archived to a separate gallery for subsequent (automated/ manual) analysis.

- ***Improvements in Video/CCTV Image Capture and Storage***

  The single most contributory factor to poor Face Recognition performance with CCTV images is quality. Raw or non extracted footage caught on camera is frequently of sufficiently high resolution for use with face recognition, so long as a 'near' full frontal face image is available. However, typical business procedures involve applying high levels of compression when archiving the footage onto tape or to digital media such that the images are then virtually unusable, even for human analysis. Moving towards the use of high definition or good quality CCTV capture, with minimal compression, could yield sufficiently good quality data that may be usable for automated searching.

- ***Forensic evidence***

  If significant performance improvements are made, the potential to identify persons from data at a crime scene will be more widely extendable to linking faces to known individuals, and to other crimes. CCTV to mugshot, and mugshot to CCTV searching will provide significant leads for anti-terrorist intelligence and surveillance investigations too. Even with limited accuracy, Face Recognition, in these contexts, could potentially reduce the amount of CCTV needed to be referred for manual analysis, and moreover, witnesses will also be able to spot individuals from footage of the subject in a more natural context (i.e. pose and environment) than mugshot images, which often lack expression.

- ***Manual Expertise and R&D into image enhancement***

  Development and provision of automated tools that assist operators in manually marking up features or points of interest may complement algorithm performance, particularly so for local feature based or graph matching approaches. Other image enhancement functions such as lighting correction, enlargement, rotation, mirror imaging, etc. may be used by operators to prepare the image for searching in order to improve its quality and minimise the effects of variance that arise as result of capture or environment, etc. At present, Face Recognition applications are fully automated (with the exception of manually marking the eye co-ordinates). Very few vendors provide such functionality and there is little research and development into HCI operations such as these. This is reflective of the market currently being geared towards civilian applications which require speed, high throughput and minimal expert intervention, which is in many ways the opposite of automated forensic analysis of facial images that may be employed in law enforcement.

4.2.5 Threats

- ***Negative impact of past implementations***

  Following the US attacks on 9/11, there have been numerous implementations of facial recognition with CCTV (mainly mugshot to CCTV within policing and for civilian security environments such as spotting terrorists at airports). Many of these have received a vast amount of media attention and subsequent criticism of the technology as it has proved to be immature for implementation. There has also been huge debate and serious criticism of the technology in terms of the privacy implications of

face recognition combined with other surveillance technology. As a result many vendors and end user stakeholders have decided against continued use of the technology in this way citing the possibility of false matches and other technological shortcomings of face recognition systems. Research and development is primarily aimed at other mass market opportunities such as access control or those described for category 1. As a result, unless this research is directly funded by government in some way, there is a real risk that this potential area of application will remain unrealised to the detriment of future policing capabilities.

- ***Lack of test data sources***

  One of the main obstacles to research and development is the lack of (operationally representative) test data available for CCTV facial images. There is currently no suitable test data with CCTV probe or gallery images with matching mugshots that can be readily provided to academia or industry to develop the face recognition technology in this area.

4.3 Category 4: Uncontrolled Facial Image to Uncontrolled Facial Image (CCTV to CCTV)

In this category both the probe and gallery facial images are acquired or extracted from video data, typically captured in conditions that cannot be controlled (i.e. CCTV). The application is most likely to require operation in an identification mode and allow use of multiple sources of facial image data. This category will enable the linking of unsolved crimes together if it is forensically used. Searching of CCTV images from multiple cameras taken at different instances would be a pivotal change to the use of CCTV data and covert forensic identification. The potential to link unsolved crimes together would provide valuable intelligence. However, the current limitations surrounding the use of CCTV/video images with face recognition, as described above, are too great for the technology to be used in this manner. As a result, there is currently little, or no, research and development addressing face recognition technology for application in this category.

It is possible that accuracy in this area could be considerably improved by making use of other identification data (clothing, hats, bags, build and gait) – i.e. whole body recognition rather than face recognition. Some research in this area is already being done, looking primarily at crowd behaviour and detection of abnormal events, but this work is closely related to techniques for linking individuals across multiple CCTV sources.

As described in section 4.2.4, camera placement, quality, and the use of automated CCTV may provide opportunities to obtain good quality data from such image sources, and hence, there may be some occasions whereby face recognition proves to be a useful analytical tool that can sit alongside a CCTV system. Tracking persons in scenes, reducing the time taken to trawl through hours of footage, and closed covert surveillance operations are some examples where this may prove beneficial.

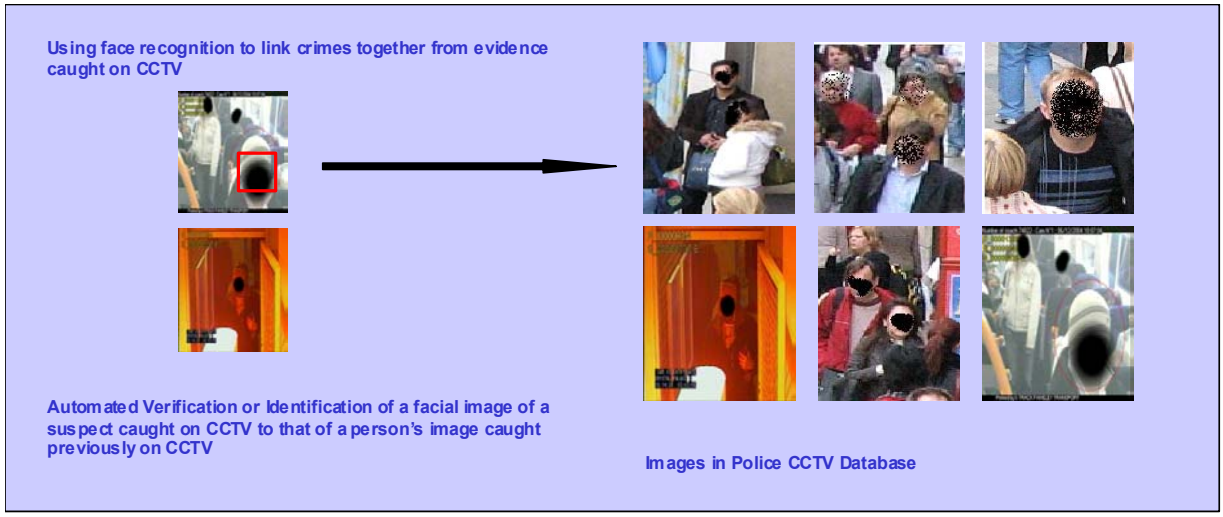


Figure 10: Diagram of Scenario CCTV to CCTV searching

4.4 3D Face or Multiple View Based Approaches

From the scenarios above it is apparent that the majority of police enquiry data will be of 2D images of faces taken from multiple angles, illumination and environments, captured from 2D CCTV data. Nonetheless the emergence of 3D facial image capture and recognition can still be used for many of the policing purposes as described in the scenarios above and is in many ways more suitable than the use of 2D face recognition.

The following are types of 3D face recognition that may be applied to the policing scenarios as described above.

4.4.1 3D – 3D Searching

The capture of 3D mugshots in custody suites as opposed to 2D photographs offers many advantages. Firstly, 3D would not require the subject to face any single camera. The time taken to capture a 3D image is becoming increasingly faster and may soon be more suited than 2D for operation with non co-operative subjects in custody. Furthermore, 3D images would capture far more information on subjects. Not only would an individual be viewable or identifiable from any angle, but it also enables the simultaneous capture of a second biometric, namely the ear. The skin texture detail on many 3D systems is very good, enabling scars, marks, tattoos and other distinguishing characteristics to be easily discerned and viewed from any angle.

The capability to search 3D mugshots against 3D mugshots is already provided by a number of suppliers (3D mugshot, in this case, implies capturing a 3D frontal posed image of the face). The difference from 2D capture is that often more than one camera is used, and as a result the face model is viewable from any profile of the face).

Figure 11: An example of a 3D facial image viewable from any angle[9]

In general, such applications have been shown to be significantly more robust in comparison to 2D recognition and noticeably faster. The capture and search of 3D mugshots can help to overcome a number of issues associated with 2D such as variations in illumination, pose angle, expression. As a result, depending on the application, they can offer far higher accuracy under sub-optimal conditions than a 2D system is normally able to provide.

[9] The image in Figure 11 is taken from the CyberExtruder 3D viewer application.

Note, however, that 3D capture for police mugshots will be reliant on the 2D skin and texture information being captured and viewable by operators as well as the 3D shape models. There are some 3D applications that can only display or capture the structural detail of the face (i.e. those that only use structured light to produce facial images without texture or shape data). (See Figure 12) These images will not be usable for police purposes, at least in the scenarios being considered here, as they will be meaningless for human visualisation or verification of the face.

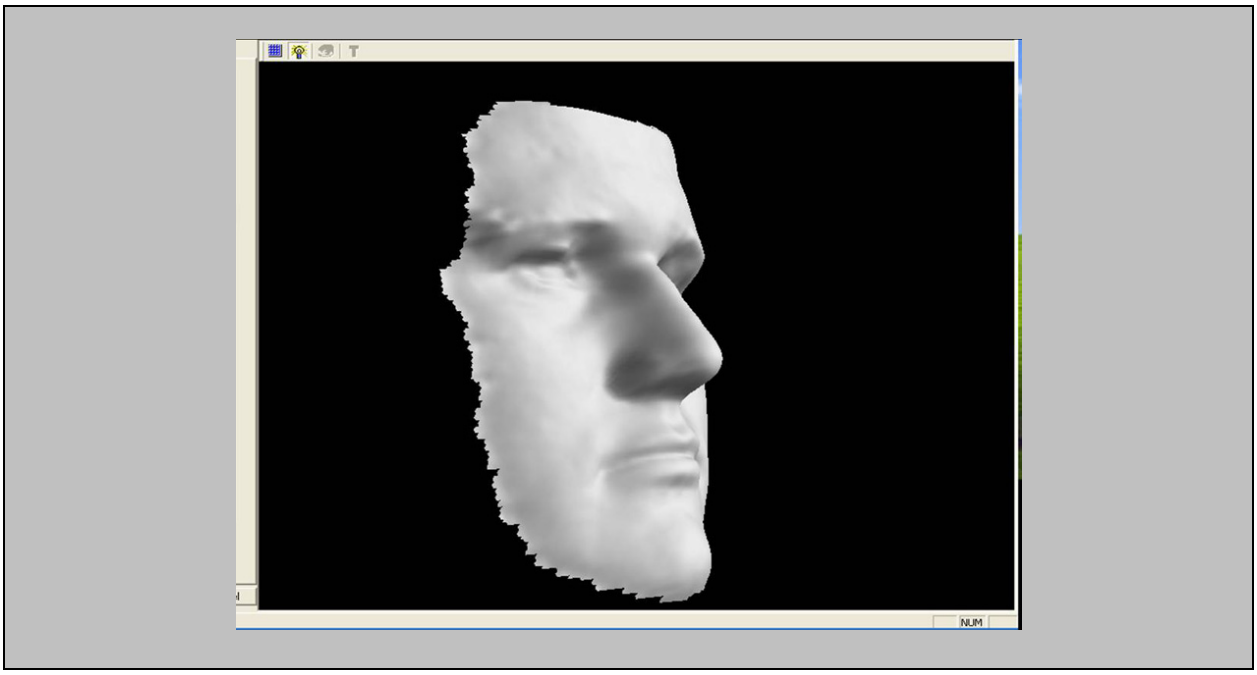


Figure 12: 3D Face image captured using infra red light. Structural image absent of skin texture

4.4.2 3D mugshot to 2D images from CCTV data

The main advantage of using 3D images for recognition is so that one can potentially identify a person from any viewpoint. However, with current 3D approaches there are numerous issues that must be overcome in order to use this data in this manner.

Firstly, current 3D recognition algorithms are designed to do 3D-to-3D image searches which is well suited for verification of an individual (1 to 1 matching) for controlling access to a room, or 3D mugshot to 3D mugshot searching. Some applications on the market (for example, Geometrix, Active ID[10]) will capture both a 3D and 2D mugshot images, simultaneously.

However, the problem still remains in that the bulk source of police data is from 2D facial images from CCTV, and it is here that 3D image capture could be of benefit in that, potentially, faces could be identified from any pose/angle captured on CCTV. However, this can only be achieved when 3D face recognition applications are designed to perform 3D mugshot to 2D CCTV facial image comparisons and vice versa. Furthermore, identifying faces caught on CCTV with 3D mugshots will inevitably require an operator to

[10] Note that during the time this report was being produced, Geometrix were bought out by a company called, ALIVE. The technology product, however, is still marketed under the original name, 'Active ID' and is currently used in some law enforcement applications in the USA.

perform a manual comparison of the two image types which may only be achievable if vendors design their applications such that they provide operators the tools to perform such tasks.

Theoretically, automated 3D mugshot to 2D CCTV image searching is a much more complex task than searching 2D images from CCTV against a 3D gallery. The reason for this is as follows: consider a 3D mugshot as the probe image whereby one is searching an image of an individual to detect if they are linked to crime previously caught on CCTV (as previously described in category 2 for 2D images). The images in the CCTV gallery will be of faces caught at random profiles and angles. In theory, the 3D mugshot probe would have to be realigned to match the profile angle for every image it is compared against in the CCTV gallery. Therefore, this realignment step would have to be performed repeatedly, for each and every image that the 3D mugshot is compared to as part of the 1-to-many (identification) search.

However, launching a 2D CCTV facial image against a collection of 3D mugshots one could specify the orientation of the face or a region of a 3D model that the algorithm should focus on as a matching region. (See Figure 13)

Though few in number, there are some vendors and academics that are exploring the potential of 3D facial image data, by developing automated comparison techniques of 3D image data with 2D data.

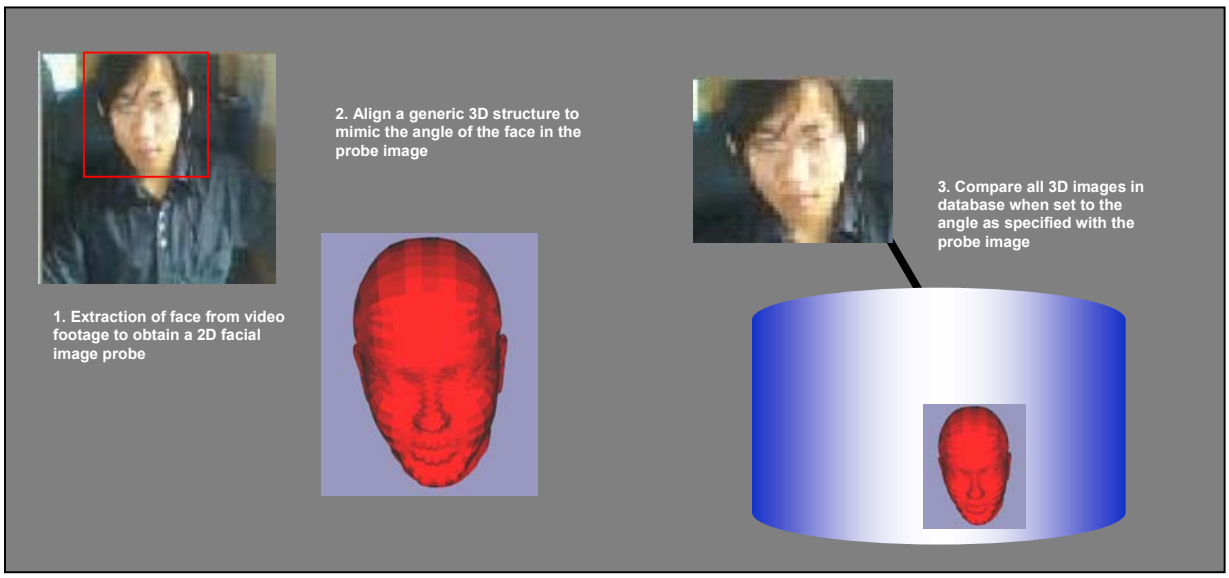


Figure 13: 2D CCTV to 3D mugshot searching by specifying the orientation of the face profile in the CCTV image

4.4.3 Multiple view based models

Multiple view based approaches have the potential to be employed in all four applications areas described. In fact many face recognition vendors are currently looking to exploit multiple images from different views to perform face recognition – creating fused models for near (synthetic) 3D representations of a face. For mugshot images multiple profiles can be fused together or a single mugshot image may be processed to synthesize a 3D representation of the face. Alternatively by fusing together multiple frames of a person from CCTV footage one might be able to generate a representation of a 3D model of a face for searching against either a 2D or 3D collection.

These are novel approaches to face recognition that are still largely at the R&D stage and have yet to be investigated and fully understood, although products are now starting to appear on the market. (See

section 7 Vendors, Animetrics) However, should such an approach be successful they would offer significant benefits to policing, improving both human and automated matching performance using conventional CCTV images without the need to deploy special 3D cameras.

4.4.4 3D scanning with Long Range Multispectral Laser Imaging (LIDAR)

The ability to scan and classify objects in scenes using information from multiple light signal channels has been explored to the extent that a select number of vendors have developed products specifically for use in military surveillance and intelligence operations. More recently, this technology has been applied to acquiring 3D facial image scans at a distance, or '*at range*'. The technology, commonly called 'LIDAR' (**LI**ght **D**etection **A**nd **R**anging), is based on using lasers in multiple bands of the light spectrum measuring the reflectance of signal as it comes into contact with objects within the range of the sensor. If successful this technology has the potential to be used for 3D to 3D face detection and tracking for targeted, covert surveillance operations and may prove to be more robust than traditional CCTV detection and tracking applications based on 2D recognition.

Perhaps in the future such applications may be more widely deployed in public environments. However, this will only be the case once issues over public acceptability and privacy have been addressed.

5 Government Initiatives for Face Recognition

This section provides information on a number of government initiatives where the potential for face recognition is currently being explored and where the technology has already been implemented. The list below is by no means exhaustive, but it aims to demonstrate the level of interest that there currently is for face recognition from governments across the world and the range of application areas where the technology has been implemented.

5.1 France

5.1.1 3D recognition for employee authentication at Lyon Airport

The French Civil Aviation Authority has been looking into the use of face recognition technologies. Since the beginning of the year (2006), Lyon airport has been using a 3D facial imaging and recognition system supplied by A4 Vision to create security badges for 500 pilots, mechanics, and other employees with access to the airport's highly secure tarmac. The airport hopes to issue as many as 5000 badges to its employees by June. The CAA are currently working in plans to roll out 3D biometric systems at airports in Bordeaux, Lille, Nice, Paris, and Toulouse.

5.2 Germany and Australia

5.2.1 Authentication of frequent passengers and visa applicants

The German government recently awarded a major contract to deliver facial recognition to automate the validation of visa applications. The contract was awarded to Cognitec Systems. The company has also recently won a contract from the Australian Department of Foreign Affairs and Trade (DFAT) for face recognition software to be used with its passport issuance and renewal process. This follows the success of the physical access control installation, SmartGate, at Sydney International Airport (2002) whereby frequent travellers and Qantas airline staff were enrolled for automatic verification by the system.

5.3 Malaysia

5.3.1 Biometric Passport

The Malaysian government has been issuing biometric passports for sometime and now the majority of domestic passport holders have these. It is based on the use of three biometrics; face, finger and iris. The user is able to use a choice of biometric modality, the most commonly used when required to verify being face.

5.4 Pakistan

5.4.1 Duplicate checking for passports

The National Database and Registration Authority of Pakistan (NADRA) is currently using face recognition in conjunction with its passport images. The system, supplied by Viisage, has been used to identify duplicates from 34 million images. The database is expected to grow to 50 million records once enrolment is complete.

5.5 United Kingdom

5.5.1 UK Passport Service (UKPS): New Biometric Passport

UKPS (now part of the Identity and Passport Service - IPS) has been evaluating face recognition for the capability to electronically match facial images stored on, or presented to UKPS systems, with the stored passport records of individuals or watch lists. The purpose is to improve counter fraud capabilities by recognising fraudsters and high-risk individuals previously identified and photographed when they access UKPS services. The UKPS performed scanning trials to determine solution options. Images were scanned and enrolled using an application provided by Imagis (now known as Visiphor). The trial included a study into identification methods used by operators to perform visual verifications, which led to the development of bespoke automated tools to assist an operator with on screen verification.

By early next year the IPS will have begun roll out of e-passports equipped with a facial biometric. There are plans to include fingerprints by 2009. The passports include a microchip that holds a digitised facial image but has space to hold additional biometrics in the future if needed. Phased roll out of e-Passports began in February 2006.

5.5.2 National Identity Cards (UK) and the Wider Criminal Justice: Public Entitlements and cross reference checking

Under the UK's national ID proposal (the details of which are still subject to change) face, finger and possibly iris scans will be used to identify people. Although finger is the primary biometric, face may provide other opportunities for establishing a person's identity that are socially more acceptable than fingerprints, which the general public still associate with criminality. This may open the potential for stakeholders of the ID card scheme to look towards the benefits of biometric authentication. The national identity register is primarily advocated as a more reliable method of providing entitlements to those that have a right to them. The very nature of the capability will undoubtedly result in the card holders or the system being used to cross check information when required. In particular, the Criminal Records Bureau has considered the use of the National Identity Register to perform checks on individuals which require screening because of the nature of their job i.e. working with children or vulnerable adults. The present system is limited in that criminal history checks are searched on name, gender, and date of birth - all three of which do not necessarily relate back to a single individual. Matching the fingerprints of a CRB applicant and searching them against the national ten-print collection stored on IDENT1 would provide a far more reliable and efficient process for checking criminal history information on PNC. However, it was felt that applicants would oppose having their fingerprints taken because of they would feel they were being treated as criminals. However, the adoption of automated Face Recognition technology, in schemes such as the ID cards and IDENT1, may provide an opportunity for improving the reliability and efficiency of criminal records bureau disclosure process and possible methods of establishing identity in wider areas of the CJS.

5.5.3 London borough of Newham: CCTV monitoring and watchlists

In 1998 the London Borough of Newham deployed a face recognition system (supplied by Visionics) to accompany its CCTV operation. This implementation of the technology received a vast amount of media and industry attention. However, the trial was widely reported to have failed as the immaturity of the technology was highlighted, even though it resulted in an immediate drop in crime. Evidently, the use of Face Recognition technology regardless of its accuracy was successful in deterring crime in the area.

5.5.4 NCIS – ChildBase: Detecting faces on the Internet

The National Criminal Intelligence Service, in 2003, launched the ChildBase system: a bespoke facial recognition software based on technology developed by the Canadian company Imagis Technologies (now known as Visiphor). The system's database relies on some 850,000 images taken from past investigations, such as Operation Cathedral, which cracked the 'Wonderland Club' of internet paedophiles in 1998. Paedophiles often download thousands of images of abuse but presently each one must be manually retrieved and categorised - a mammoth task that takes up huge amounts of police time and resources so that much of the evidence is left without being retrieved. The system is designed to log images and cross-reference them automatically, matching up new images which have no clues to their history with images that have previously been analysed.

This particular deployment of face recognition technology is regarded by its end users and by industry as a success. It has significantly reduced the need for investigators to trawl through thousands of disturbing images and has resulted in faster and more efficient development of investigative leads.

5.5.5 West Yorkshire Police: Matching against mugshot data collection

West Yorkshire Police's Imaging Unit has been very active in furthering facial identification methods. It was one of the first forces to acquire a face recognition system (from Aurora) and now use it to search CCTV images against a database of mugshots. Though the force has acknowledged that the technology is still somewhat immature for this type of application they have reported that it has proven to be a useful investigative tool.

5.6 United States

5.6.1 Tampa Florida Police Department: Scanning crowds

In January 2001, the city of Tampa, Florida used face recognition technology to scan the faces of people in crowds at the Super Bowl, comparing them with images in a database of digital mug shots. Cameras equipped with face recognition technology were subsequently installed in the Ybor City nightlife district. Both implementations sparked huge controversy over the use of the technology being used to spy against people and citing it as a breach of a civilian's right to privacy. The city encountered opposition from people wearing masks and making obscene gestures at the cameras. Subsequently, in August of 2003, the Tampa Police Department scrapped Ybor City's facial-recognition system, citing the system's ineffectiveness as bearing heavily on their decision.

5.6.2 Logan Airport, Boston: Airport Surveillance

In 2002 face recognition trials commenced at Logan Airport in Boston whereby the Identix system (Visionics at the time) was used to pilot the technology for surveillance purposes. During the test, photographs of 40 employees who volunteered for the program were scanned into a database. Cameras at two checkpoints at the airport relayed the images of everyone (passengers and employees) passing through to a computer, which compared them to the pictures stored in its memory. It used the facial recognition technology to come up with a match. The technology was reported to have successfully identified enrolled individuals 153 times but failed to match them on 95 occasions. This implementation was also heavily criticised by the media and privacy activists.

5.6.3 Pinellas County Sheriffs Department: Initial check on arrival at custody

The Sheriff's office in Pinellas County Florida has rolled out a system provided by Viisage both at the county Jail house (custody) and at the nearby court house. The jail house system enables an individual bought into custody to have an initial identity check using a mugshot photograph prior to their formal booking. The image is searched against the force's own database collections of mugshots to see if an individual has an existing record, but can also be searched against data from neighbouring counties. Fingerprints and a second mugshot are taken later during the formal booking process and a final search at the time of release confirms that the individual has not swapped identities with someone else. The Viisage application was also installed as an access control application, performing real time watchlist searching to secure access to the county court house.

5.7 Canada

5.7.1 Bluebear: Rapid probing of forces' mugshot collections

In 2003 the pilot project "Bluebear" was launched with three forces in Canada (Chatham-Kent, York and Windsor regional forces). The project allows the participants to quickly and simultaneously search each other's mugshot and text databases, in a highly secure environment over the Internet. The technology was supplied by Sun Microsystems and used Visionsphere's face recognition engine. Bluebear is now a company in its own right when Visionsphere became Bluebear Network International providing automated tools for facial identification and text searching of demographic data.

6 Face Recognition Standards

Within the last four years there has been significant activity in the development of a wide range of Biometric standards. More specifically, however, standards for face recognition have been pushed forward in order to facilitate take up and development of the technology as a result of huge interest in deploying applications post 9/11 (e.g. the introduction of e-passports worldwide).

The US Visit programme, whereby all visa applicants to the US are now required to provide a biometric (fingerprints) to enter the US at all ports and borders of entry, resulted in the International Civil Aviation Organisation (ICAO) mandating that all visas and passports must contain a biometric identifier, with face as the primary biometric, with fingerprint and iris as optional secondary biometrics. As a result, the face recognition market has boomed, but there was a global call from industry and government for standardisation of facial image capture, storage and transmission, etc. in order to ensure the face recognition community and systems integrators will be able to overcome large scale implementation issues such as interoperability, and legacy datasets and infrastructure.

The following is an overview of current relevant standards published so far by the International Standards Organisation's (ISO) committee for Biometrics.

6.1 International Standards Organisation (ISO)

6.1.1 19794-1 Biometric Data Interchange Format: Framework

This Standard sets the context for the standardisation of biometric data and its use in other biometrics data structures. The Standard discusses the issues involved in the capture, feature extraction, and use of biometric data at a "Biometric Data Block (BDB)" level, including the distinction between a BDB containing image data and one based on feature extraction. It also discusses requirements for a sensor, some of the terminology used in multi-modal work (multiple BDBs, possibly using different biometrics), and the BDB format identifier registration mechanism. This standard has a number of sections for different biometrics one of which is for face - *19794-5 Biometric Data Interchange Format: Face Image Data.*

6.1.2 19794-5 Biometric Data Interchange Format: Face Image Data

This Standard defines a data structure (called a Biometric Data Block format) that contains a digital record of the image of a face. It specifies how the image is to be acquired (including lighting, pose of the subject, smiling, head-dress, etc.) and how it is to be converted to a digital representation, with a full specification of the digital format. This Standard enables equipment from one vendor to produce a face image data block format that can be compared directly with a face image data block produced by equipment from a different vendor without any collaboration between the two vendors (open 'inter-working').

Associated matching algorithms are not standardised, and are generally company confidential. Normally a Biometric Data Block would be "captured" when a person is "enrolled", and stored with some additional metadata about the time of capture, the equipment used, and so forth. It might be stored on a smart-card to be carried by the person, or on a central database, or both. These options for storing data are subject to privacy concerns that might be expressed by the individual or in national legislation, and to the need to maintain backups. The diagram below shows the structure of the facial image data, as described in the standard.

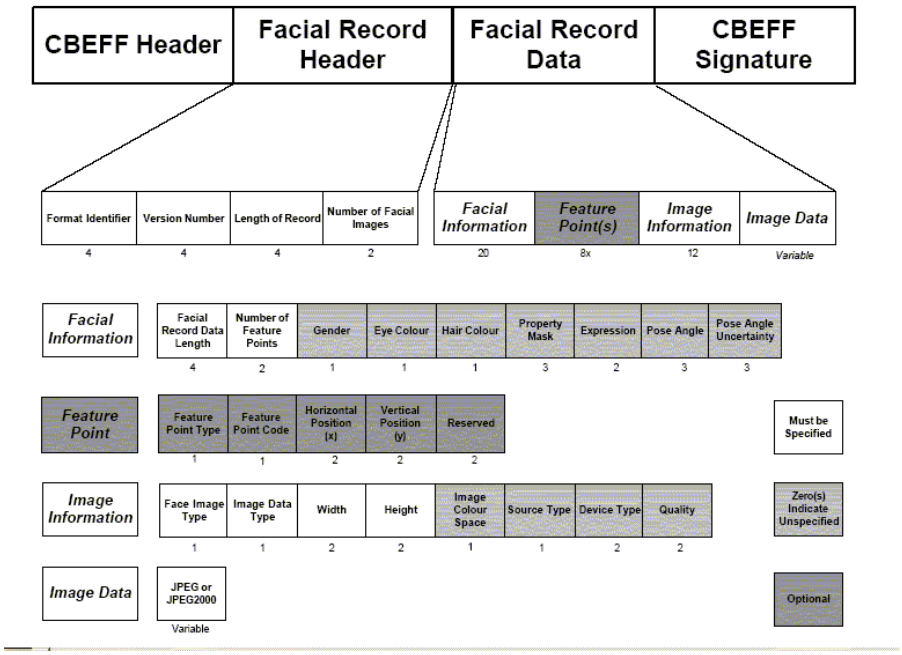


Figure 14: Diagram of Facial Image Data format for ISO interoperability standard

Note that the standard is currently being amended to include specifications for a 3D Face Image Format. It is currently in a working draft. The changes will include definition of the range of the information block, type of 3D images, and the capture.

In addition, ISO working groups are currently working on further standards and technical reports, addressing related issues such as Best Practice' guidelines for capturing compliant facial images, an extension to 19794-5 to incorporate storage of 3D data, and developing standardised methods for assessing the 'quality' of a facial image. Storing an image quality metric alongside an image can provide useful information about the likely accuracy of results obtained from a database search.

Outside of the SC37 Biometrics Standards body, MPEG are also actively working on standards which allow for Advanced Face Descriptors, and Advanced Image Coding information to be stored with images, facilitating image searching and retrieval based on common descriptors. If the proposed new work item on Face Recognition using Advanced Face Descriptors is ever taken forward, ISO 19794-12 could one day provide a standard approach to face recognition using such data.

6.2 FIND Police Standard for Mugshot Capture

This Standard was produced by PITO in 2005, in preparation for the delivery of FIND. The purpose of the Standard is to specify the capture of still digital images and Data Interchange of facial mugshot and scar, mark and tattoo (SMT) images for the police. This Standard draws on the following pre-existing documents:

- ANSI/NIST-ITL 1-2000, Data Format for the Interchange of Fingerprint, Facial, & Scar Mark & Tattoo (SMT) Information.
- NIST Best Practice Recommendation for the Capture Of Mugshots,
- Annex D of the Machine Readable Travel standard (Facial Image Format for Interoperable Data Interchange) by International Civil Aviation Organisation
- IS0/IEC FCD 19794-5, Biometric Data Interchange Formats (as above)

- VIPER Remote Site Design Specification (see page 30 of previous section)
- UKPS Photo Guidelines for Passport Applications.

It is important to note that the FIND standard is a minimum standard for mugshot capture and due to the lack of prior work in face recognition standards it draws heavily on earlier standards that were not written specifically with a view to using such images with face recognition technology. For example, the FIND standard specifies an 18% grey background, but recent evaluations of Face Recognition applications have shown that this may be too dark. A lighter shade may be required for face recognition as otherwise the technology can have problems in deciphering the background from the outline of the head. The specification of 18% grey was inherited from predefined standards prior to face recognition and has been specified in FIND for the purpose of having a uniform background and format of image. This is important for when human operators or witnesses are viewing such images in order to prevent any one image standing out due to extraneous factors such as background etc.

Ownership of the Standard has been accepted by ACPO, although PITO closely monitor it and will be responsible for updating it as and when necessary.

6.3 The VIPER Standard

The VIPER (Video Identification Parade Electronic Records) Standard was produced by West Yorkshire Police for the VIPER system. It specifies how video mugshots need to be taken in order to be used with the electronic video ID parade application. Both the VIPER system and standard are being adopted by PITO in the development of (NVIS) National Video Identification Strategy whereby a gallery of video images of volunteers and offenders will be established for police forces.

7 Vendors

In terms of the algorithms providers that have been most consistently used by companies winning large face recognition contracts, the top four, in alphabetical order, are A4 vision, Cognitec, Identix, and Viisage. (BTT, May 2006). Note, however, that recently a number of companies have merged, including those with their own successful algorithms - namely, Identix and Viisage. The Face Recognition market is rapidly changing and such mergers are frequent. The table below provides a snap shot of the vendors' offerings which is accurate at the time of writing this report, but due to the nature of the market it is likely that the information presented here will change. The list is by no means exhaustive but is intended more for the purpose of presenting a range of companies adopting different approaches to Face Recognition.

Company	Technology	Comments
AcSys Biometrics	2D video capture based with face tracking • FRS Discovery • VeraShield • VeraPort	AcSys Biometrics is one of the few on the market that uses a Neural Network based approach for its algorithm. FRS discovery is a functional product designed for evaluations and giving customers a feel for what FR is. VeraShield and Port are access controlled based applications whereas FRS is geared towards surveillance. AcSys's Holographic/Quantum Neural Technology learns changes in facial features which allows the software to be robust to aging processes and other cosmetic variations. Previous benchmarks have shown that although it does not perform as well as other algorithms on the market, the technology is able to improve over time as it learns data. Furthermore the technology has claimed to be robust to up to 60 degrees off centre.
A4 Vision Primary focus in physical access control, border control and registered traveller programs.	(Face Reader) 3D Face recognition using structured light projection	The technology is based on projecting structured light on to the face of a person to capture a geometric model of the facial structure. It is largely dependent on the nose being correctly located but provides a vast amount of structural information about the face. The main application on the market only acquires structural images without skin texture data. Thus for policing, and in particular for the purpose of mugshot to mugshot recognition, this technology is not readily applicable. However, the company has recently released a version of their product that will overlay a 2D face texture image over the structural image for use. The US Department of Homeland Security's Federal Protective service has recently implemented A4 Visions Access 3D Face reader at its region 10 headquarters. PITO also uses the system to control access to its Biometrics Laboratory at its London offices.

Animetrics Although heavily aimed at forensic applications the company also caters to the civilian market.	FaceEngine (Forensica) 3D face recognition using 2D photographic models	This technology enables the conversion of a 2-dimensional photo into a 3-dimensional avatar, which can then be rotated and seen from any view. It is a technology that is able to take a portrait captured at any angle and return it to a frontal pose or any other desired orientation. Unlike face detection it the technology uses knowledge about the head shape and change in an image to locate and track the head in a picture. Although heavily aimed at forensic applications, the company also caters to the civilian market. No information is currently published on its performance, but such an approach clearly has potential benefits for forensic face identification (automated and manual).
Aurora Police mugshot recognition	e-Gallery Face recognition technology distributer (Identix Engine)	Aurora are very well known in the UK police arena with West Yorkshire police having a system installed, and also chairing the Aurora user group for facial identification. The company's technology is currently based on the Identix G3 engine, which Aurora claim provides optimum performance with police mugshot and CCTV images, but can easily be changed should new algorithms prove to be more suitable. The application has been developed specifically for Law enforcement use of mugshots. Aurora and their sister company Advantage also provided the system that is currently being used by UKPS (now IPS) to evaluate face recognition technology for identifying and preventing passport fraud.
Cognitec Physical access control, border control and registered traveller programs.	FaceVACS 2D face recognition	The Cognitec FaceVACS application is based on portrait matching for ID cards, travel documents etc. Cognitec have received extensive recognition by independent evaluations of their technology and was one of the top the performers in DARPA's last round of tests. The technology is 2D based, combining feature recognition with texture and light intensity analysis, and Neural Network approaches. The technology is verification based however there is also the FaceVACS – alert which is designed more specifically for identification. With recent developments in 3D sensors Cognitec are also looking to enhance the capability of the FaceVACS product to include 3D data capture and recognition.
Identix	FaceIt G3 and G6	The Identix face recognition technology works optimally

Note - Identix merged with Viisage (Jan 2006) and now both companies have been taken over by L1 Identity Solutions.	2D face recognition, Forensic Analysis User Interface Application	when matching frontal images. It is able to detect faces as long as both eyes are visible. Recognition is not significantly affected by variations in pose up to 15 degrees. The company claims robustness at 15 to 35 degrees with more significant loss of matching for variation beyond 35 degrees off centre. The FaceIt G6 engine, the most recent release of the Identix technology is unique in that it also uses skin texture information. It is based on fusing 3 algorithms Vector Feature Analysis (VFA): Local Feature Analysis and Surface Texture Analysis (STA). However, it is reliant on high quality images. UKPS (now IPS) are currently using the G6 engine in trials to help prevent passport fraud by detecting duplicate applications. Recently the company has released a new user interface application which is forensically motivated to assist investigative officers to process image data to launch and verify images.
Viisage Viisage merged with their prime competitor Identix (Jan 2006) and now both companies have been taken over by L1 Identity Solutions. Civil ID. Criminal ID. Border and Area Security	2D face recognition	The algorithm behind the Viisage technology is an Eigenface based approach. The combined company will potentially be able to provide multimodal finger, face and skin imaging identity solutions. Viisage's deployment of a custody system in Pinellas County Sheriff's Department (Florida) was high publicised and considered to be successful. The deployment involved immediate mugshot capture and searching to identify and track arrestees previously registered with the Police department. This included a web enabled search and retrieval interface and mobile identification capability.
Neven Vision now taken over by Google.	2D face recognition	Prior to being taken over by Google, Neven Vision were quite unique in their use of face recognition technology with mobile devices such as mobile phones, PDA's, etc. based on local watchlist searching. The company had their own algorithm, originally from Eyematic which the company was previously known as. However, now under Google, the technology has been moved away from being applied to this area of face recognition. Instead, Google aim to use the technology in conjunction with consumer online photo collection applications.
Geometrix (Now taken over by Alive)	3D face recognition	The Geometrix ActiveID system performs 3D recognition but the 3D model representation is generated from stereo

	Active ID	camera stills and does not use projected light as nearly all other 3D based face cameras do. Two images are obtained with dense correspondence maps to fuse them together to produce a 3D representation together with skin texture data. As a result the system can capture a 3D model and 2D portrait that is ICAO (or even FIND compliant) from a single stage process.
OmniPerception	2D face recognition	OmniPerception are a spin off company from the University of Surrey's Computer Vision department. The technology uses an Independent Component Analysis based approach to face recognition and has reportedly been able to achieve a level of accuracy on par with some of the major face recognition vendors. The company has not yet identified its niche market (civilian or law enforcement) and primarily supply their technology as an SDK or as a bespoke service. OmniPerception have had some commercial success with a related product (Magellan) which analyses video to detect particular images or patterns within the frames. Its primary application is for detecting advertising within broadcast sporting events, but the underlying technology may also be well suited to law enforcement requirements for video analysis.
CyberExtruder	3D model generator	CyberExtruder's technology uses reconstruction algorithms to create, from 2D representations of a face, a 3D model of a persons head potentially viewable from all angles. The technology works by projecting the 2D image data onto a generic mesh model of a head.
XID	Face Scan Access Control Generates a synthesised 3D face model	In addition to an Access Control technology based on Face Scans, XID also have a face synthesiser application based on using a 2D image of a face and generating thousands of representation of the image distorted to simulate variation in illumination, pose etc and fusing then together to generate a 3D model representation. The tool may then be used to manually mark up features or co-ordinates on the face by an operator that may assist is automated or human matching of the facial image.
3dMD	Qlonerator – 3D head model imaging system	Though not a face recognition company 3dMD is worth mentioning here for its product designed to capture 3D face models, the Qlonerator. The company claim to have sold up

		to 125 3D systems worldwide, mainly within the medical arena. The technology uses multiple high definition cameras to (near instantaneously) capture a person's image from multiple views to then generate a 3D representation within seconds.

Though not listed here, other companies that are worth monitoring for new applications in the near future are Sagem, Visiphor (both systems integrators with in house Face Recognition algorithms) and Northrop Grumman and Intuvision who have developed intelligent CCTV data mining systems that are specifically designed to be coupled with Face Recognition technology. Clearly many other companies can be expected to be developing similar applications, although since much of this work is targeted at military and covert surveillance scenarios, reliable information is difficult to obtain.

8 Research

Face Recognition has been a research area in computer imaging and cognitive psychology for over 30 years and continues to generate immense academic interest. In the early days, the objectives of developing face recognition algorithms were for the purpose of applying them to improving methods of facial identification in law enforcement. However, over the years, the complexity of dynamic, real life environments has posed some very difficult research challenges for academia to address. As more and more technologies acquire poor publicity as a result of incorrect expectations being set and innapropriate deployment of the technology, fewer vendors and researchers are focussing their attention on the law enforcement market. As a result improved methods of face recognition have favoured the civilian domain where there is a mass market for which the technology is currently well suited (access control applications, drivers licences, passport photographs, visas etc).

This section provides details of past research activities that have been pivotal to developments in this area. Some key papers that typify present research are summarised to present a snapshot of current academic interests too.

Finally, recommendations are given for future areas of research that have been identified as specifically addressing approaches that may result in face recognition technology being viable for application for law enforcement as described earlier in this report.

8.1 Past Activities

8.1.1 The FERET Protocol – National Institute of Standards and Technology

Prior to FERET there had been no formal evaluations of face recognition technologies. The FERET protocol was a series of evaluations that were aimed to determine the current range and state of the art performance of the technology and to advance the technology by providing test and evaluation data captured from a large population. The FERET database was divided into two partitions: one was made available to the algorithm providers (vendors and academia) to assist with development of the technology and the second was retained to perform independent evaluations and testing of algorithms.

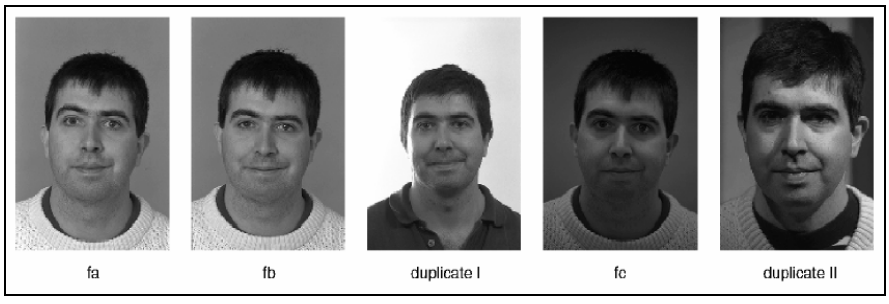


Figure 15: Examples of typical FERET images

Evaluations were performed in August 1994, March 1995, September 1996 and March 1997 all of which are reported in publicly available documents for reference.

The FERET database is available without charge and consists of 1564 sets of images (1199 original sets and 365 duplicate sets) – a total of 14,126 images. (See Figure 15) The availability of the FERET database and the evaluation protocols has had significant impact on the progress of developing face recognition algorithms and the manner in which they are evaluated. Conclusions of the FERET program

resulted in academia addressing the problem of recognition from images collected months or years apart and recognition under pose changes.

8.2 Present Interest

8.2.1 Face Recognition Grand Challenge (FRGC 2005)

The Face Recognition Grand Challenge (FRGC) was set up by NIST for the purpose of driving the accuracy of face recognition technology up by an order of magnitude from the FRVT2002. As with the FERET protocol, FRGC was designed to achieve this performance goal by providing researchers with a corpus of 50,000 images and presenting them with six research challenge problems to overcome. These were controlled indoor still versus indoor still, indoor multiple still versus indoor multiple still, 3D versus 3D, controlled indoor still versus uncontrolled indoor still (full frontal pose), 3D versus controlled single still, 3D versus uncontrolled single still.

Figure 16: FRGC Images from left to right these consist of: indoor still, outdoor still, multiple still, 3D single view and 3D full face

The FRGC scores showed that for comparison of controlled/indoor still the mean recognition rate claimed was 91%, with the highest performing algorithm achieving 99.9% and employing multiple still images – a significant improvement over single still image matching. The maximum score for 3D to 3D was found to be 97% and showed robustness to variations in expression, thus demonstrating the potential for 3D facial imagery. [11]

8.2.2 Face Recognition Vendor Test (FRVT)

The Face Recognition Vendor Test was preceded by FERET and most recently performed in 2002. Evaluations of face recognition technologies are now being conducted under FRVT2006 which is directly linked to FRGC2005. The tests will independently validate the vendors' reported accuracies from the FRGC. The corpus of images used in FRVT2002 was taken indoors, under controlled lighting conditions. The top three performing participants had an average verification rate of 80%, an error rate of 20% with 0.1% being falsely accepted.

[11] The accuracy figures reported in this section are based on reported figures claimed by FRGC participants. Note that there are numerous definitions of the term accuracy and therefore many interpretations of what these figures actually present.

8.2.3 IEEE Automated Face and Gesture Recognition 2006

This conference is held annually across the world and is one of the premier forums for presenting research in face recognition. This year over 186 papers were reviewed for publication which is evidence of the intense academic attention in face recognition from across the world. Face recognition papers focussed heavily on analysis of expression and the performance improvements of 3D recognition and multiple view based approaches. Very few papers directly addressed the issues surrounding the of uncontrolled CCTV/video images with face recognition technology.

The following two papers that were presented at this year's conference are summarised as examples of current research.

- Face Recognition from Unconstrained Images – Progress and Prototypes: *Rob Jenkins (Cambridge), Mike Burton, David White (Glasgow)*

  This paper looked at using a face averaging technique to provide a facial image representation suitable for storage in a face recognition gallery set. This involved using multiple images of the same person and constructing an average image (averaged over pixel values). The approach was shown to offer robustness to variations of faces themselves, environment and the capture device. This method allowed the average image representation to be updated with each new instance of a person's face. For the purpose of mugshot capture this approach may be useful for the case of recidivist arrestees, where with each new mugshot taken, their average face image stored in a police database may be updated. Note that this research was conducted on pose invariant images (full front images only).

- Subspace based age group classification using facial images under various lighting conditions – *Kazuya Ueki et al. (Waseda University Okubo, Japan & NEC Soft Ltd, Japan)*

  Research into gender classification has been very popular for over 20 years, but this paper looked at the challenging area of age classification. This is challenging as the assessment of age by humans is inaccurate. Enormous time and effort is required to acquire data with age variations of subjects. This research involved developing the Waseda HCI Technology Database (WITDB) which consists of facial images under a variety of lighting conditions and age groups. As with gender classification that adopts appearance based approaches to classify and learn from training data (PCA, NN and SVM's), the algorithmic approach for age classification used 2DLDA. Testing of the algorithm showed promising results with accuracies of 46.3% for 30 – 34 year olds, 67.8% for 55-59 year olds and 78.1% for 60-85 year olds.

8.2.4 Research initiatives supported by PITO

In the past four years the Biometrics team has extensively engaged with academia and has established a very strong and active programme of collaboration with universities throughout the UK. There are a number of academic institutions that PITO is currently supporting which have a direct connection to face recognition research.

- **Sheffield University** has obtained funding (brokered by PITO) from the US government for the development of a large scale facial image test set for use in assessing the performance of face recognition systems. Sheffield University has now gathered over 3,000 3D head images on behalf of the FBI. The primary purpose is to look at the variation of features across the population. It is hoped that this raw data will eventually be a made available for research purposes.
- **Kingston University** is heading up two EPSRC funded network collaborations. The aim is to produce innovative developments in the use of CCTV technology and in the area of motion prediction. The Vitab network (Video Identification and Threat Assessment using Biometrics) is focussed on

identifying threats from abnormal gestures or behaviour and providing more robust approaches to tracking vehicles and people on CCTV.

- **University of Surrey** is working on new facial identification techniques and the 3D modelling of faces. As part of a separate research proposal Surrey has recently also teamed with Glasgow University and Queen Mary and Westfield College on research into CCTV image enhancement specifically for facial identification.

- **University of Kent** has a long standing reputation in the area of biometrics research and for collaboration with PITO. More recently, PITO has supported an EPSRC proposal from the university looking at the use of automated face recognition in conjunction with poor quality CCTV.

8.2.5 Face Image Quality

NIST is currently looking at developing methods for obtaining automated image quality measures for faces following the success of the NFIQ (NIST Fingerprint Image Quality) checker. This image quality measure may provide an indication to an operator or algorithm of the quality of a face image. This information can then be used to provide a confidence measure to results returned from a matching algorithm.

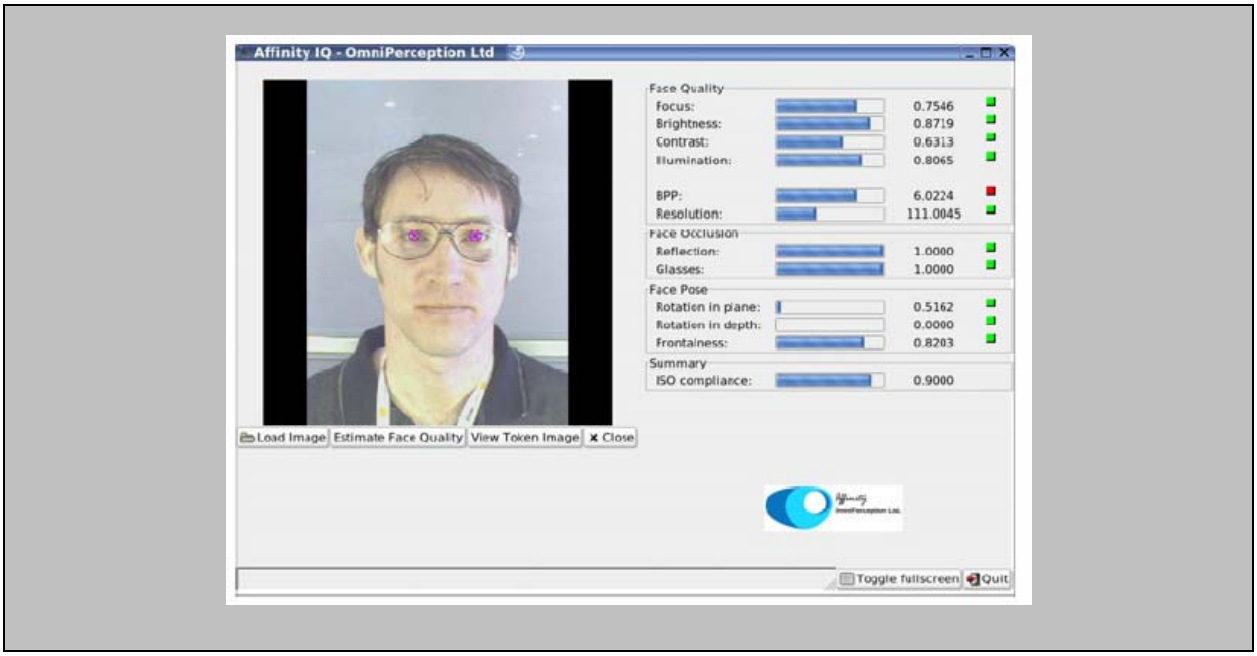


Figure 17: Image Quality Checking Screen by OmniPerception. Used to assess compliancy to ICAO facial image standard

Many face recognition vendors such as Omniperception already provide such tools with their application, and others are developing similar software. However, the image quality measures are tuned to the vendor's specification as opposed to a generic standard. With the exception of Aware's Preface SDK these image quality tools are currently not marketed as isolated applications, rather they are provided with the Face Recognition software.

ISO has plans to develop a quality standard measure for faces but as yet there has been little progress on this. A similar standard for fingerprint data is currently at 2nd working draft but it is proving difficult to do

this in a way that does not favour any particular quality algorithm or matching technique. At the heart of the problem is the challenge of defining exactly what 'quality' means in this context.

PITO's Biometrics team has recently been looking closely at a number of such applications for the purpose of investigating their potential use within custody suites to facilitate operators to capture a FIND compliant mugshot and for automated quality control of images being loaded onto the FIND system. The team's work to assess the performance of such software suggests that there is still some considerable way to go before these can be used for fully automated checks of an image.

8.3 "Facing the Future"

8.3.1 Advances in HCI for face recognition

The matching of fingerprints on automated fingerprint identification systems (AFIS) for law enforcement is analogous to the need to automatically match mugshots and facial images from CCTV or other forensic sources. An AFIS operates in an identification mode and verification is performed manually by fingerprint experts using the system. Moreover, identifying marks left at crime scenes is often made difficult in that they are deposited randomly and often of varying poor quality. For this reason searching of marks is manually assisted by the use of automated tools that help an expert to mark up a fingerprint image with features that can be used by the matching algorithm. Furthermore onscreen viewing tools are also present to assist experts to manually verify the list of respondents returned by the system.

Forensic searching and matching of facial images should also be treated in a similar manner to that employed with fingerprints. Viewing tools should be developed to assist on screen analysis; editing tools should enable marking up of images for local features or regions of interest. Finally, capture and quality should be tightly controlled and the entire process should be logged to maintain integrity of the data for the purpose of presenting it in court. There are currently no standards or guidance that stipulate how electronic facial image data from CCTV should be handled in order to ensure that it remains eligible for submission as evidence in court.

8.3.2 Advances in video data capture and efficient surveillance monitoring

It is clear that the ability to use automated face recognition for searching faces obtained from CCTV would provide a huge operational advantage; furthering criminal investigations and preventing and detecting crime. However, this capability will always be limited in terms of the performance of the Face Recognition technology due to the constraints in the quality of images that can be captured from CCTV. Therefore, in addition to research into improving the robustness of face recognition when used with such images, research and development should also be focussed on improving capture methods and the efficiency with which surveillance/video footage can be automatically processed. At present operators may spend hours viewing video footage to find evidence of the crime (or even to discover that a crime (or event of interest) has occurred). More efficient methods of viewing, archiving, indexing and searching CCTV data may provide opportunities whereby face recognition may be usefully employed.

Some surveillance vendors are looking into incorporating into their systems intelligent agents, crowd analysis, motion flow etc. that can detect abnormal occurrences or events on camera by using behavioural analysis of scenes, people, and events. Recent research in this area has resulted in the development of visual tools and automated functions such as cues or triggers to assist operators viewing CCTV footage. This coupled with automated face recognition has the potential to produce a powerful surveillance and intelligence tool. However, to date research and development in this area has not adequately been focussed at addressing the use of CCTV data with face recognition. Of the companies that the Biometrics team has come across, Northrop Grumman with Alert Video, Intuvision with ViLO, and 3VR are the few

that have already attempted to incorporate both intelligent agent technology and face recognition into their products. (See list of vendors in Section 7).

PITO's Biometrics team has (for the past four years) been very actively engaged with academia and industry stakeholders that are focussed in the area of video surveillance and monitoring in order to keep apprised of new developments in the technology that may provide an opportunity to make use of CCTV data. For example, the team has research collaborations with Kingston, Surrey, and Kent universities under EPSRC and has installed an intelligent surveillance application, supplied by NEC, for in-house evaluation and demonstration purposes. The Biometrics team has, through many such opportunities, been able to contribute its knowledge and expertise to push forward development of new approaches to video and surveillance monitoring with a view to its future use with face recognition within policing.

8.3.3 Future Research Possibilities

It is apparent that 3D image capture and recognition has several advantages over 2D which will always be limited in terms of its use in law enforcement applications. Within the last few years 3D face recognition has advanced dramatically. More and more research will be performed into improving the ease and efficiency of 3D capture using stereo camera and multiple view based approaches as these have the largest potential for building on the existing infrastructure of technology and data that is available. An interesting avenue of future research would be to look into head (as opposed to just face) recognition as this would be more compatible with the type of images that will be available forensically from CCTV. There is also the question of whether full face images (required for 2D automated recognition) are the most appropriate for human recognition. FIND will store full face images as the norm, but there is considerable evidence from research into the cognitive aspects of face recognition, that ¾ views of heads are easier for humans to recognise. The use of 3D mugshots might deliver considerable improvements in recognition rates, especially when comparing faces from CCTV.

8.3.4 Cognitive approaches and issues to recognising faces

Historically, there have been some very significant pieces of research into the cognitive psychology of how humans recognise people. Work has been performed into how humans recognise one another. It is known that individuals can identify people from their own ethnic group better than those from different groups. In addition, people are very good at recognising known faces but find it is more difficult to match images of unfamiliar faces. This is also the case for expression; classifying facial expressions is much easier for known individuals.

Much of the research into cognitive psychology has been applied to the development of police witness applications such as '*Efit*' and '*Photofit*' which are composite applications used to generate pictorial representations of a witness's description of an offender. Research efforts include the development of visual aids, interview methods to stimulate memory and recollection, and analysis of the terms and diction commonly used by children or other vulnerable witnesses in describing offenders. For example, children will often use terms such as oval or square to describe angular faces or point to shapes or objects that are similar to what they are trying to describe.

Further research into cognitive approaches to recognising faces applied to the development of Face Recognition would be very beneficial. It would be useful to understand how the performance of face recognition technology may potentially be impacted by different populations. How important is the ethnicity of the viewer / investigator when comparing images of faces of different ethic origin? How do children compare with adults in these respects? How do male / female compare and what can it provide by means of informing development of automated gender classification algorithms for searching faces.

One key area to look into would be to determine how many images a person is able to examine before their ability to correctly classify faces as matches / non matches deteriorates, (e.g. due to false memory,

tiredness etc.) How long a break is required before performance improves again? Answers to questions such as these would be very informative for establishing best practice guidelines and processes for viewing facial images. For example fingerprint experts traditionally look at the top 15 returned respondents from a search of crime scene marks for fingers and are encouraged to take breaks at least every 2 hours to avoid fatigue. It seems likely that much longer respondent lists of facial images could be viewed without strain (as there is less detailed examination of the images) but it is not clear how much benefit there would be in terms of the numbers of identifications made. It is also not clear at this stage that manually examining respondent lists of facial images is the best approach. It may be more appropriate to simply use the top 50 returned images (for example) as a shortlist of individuals for further investigation, rather than attempting to make a match / no match decision based on a single full face image.

PITO's Biometrics team has already raised these issues with a number of leading academics in the field (Peter Hancock, Mike Burton, to name but two) in order to stimulate further research into these issues. As a result, recent EPSRC submissions for research into face recognition have included an element of cognitive research, with PITO as prime collaborators/ stakeholders.

There may be an opportunity for PITO to contribute its knowledge in this area to similar efforts across governments overseas. For example the US has recently set up a focussed group to investigate face recognition for police application whereby these issues will be addressed. The Biometrics team has an extensive understanding in this area and should take advantage of this opportunity to both share its knowledge and learn of developments by its overseas counterparts.

8.3.5 US Research interests in Face Recognition

Although the US does not have the same quantity of CCTV cameras as the UK, the NIJ (National Institute of Justice) has an active research program into biometrics and recently announced several new awards related to face recognition applications. These include:

- Automatic normalisation of uncooperative subjects
- Extracting 3D face data from video surveillance footage
- High definition 3D face from video
- Development of 3D surveillance technology
- High definition CCTV to enable 'in car' face recognition of drivers
- Creation of a test database of images from different quality CCTV cameras

It is interesting to note the emphasis on the use of 3D techniques which closely mirrors current UK academic research and is also in line with the views of PITO's Biometrics Team on how face recognition from uncontrolled subjects such as those captured on CCTV will develop in the future.

9 Report Summary

Face Recognition, though it has been a research area for over 30 years has only recently started to mature, sparked by growth of the biometrics market in general after the US terrorist attacks of 9/11. The last few years have seen considerable progress in the development of the technology, particularly when operating with controlled images for one to one comparison (verification). Recent evaluations and research such as FRGC have indicated that the technology fares very well for verification, reaching accuracy on a par with fingerprint recognition. However, this level of performance from the technology is dependent on it being used in conjunction with good quality, high resolution, and controlled portrait images for both the probe and gallery sets. When operating with controlled, full frontal portraits in an identification mode (or one to many matching) automated face recognition technology is considered as 'promising' at the present time, although more development and understanding is required on how age, varying populations, appearance (or facial artefacts such as glasses), ethnicity and expressions impact on performance.

There is a clear distinction between the civilian and law enforcement areas of application in terms of the requirements, implementation and expectations of the technology. The civilian domain is primarily geared to using the application for verification or access control; whereby the images used with the technology will generally be captured as good quality, full face portraits, in a controlled environment and will accommodate high volumes of searches. Subjects will also be co-operative as there will be clear incentives or benefits to the end users, such as ease of access to a facility or fast passage through immigration. Though the technology has sparked huge debate amongst privacy advocates this has mainly been in the surveillance domain as opposed to civilian type application of the technology. Face Recognition is being widely adopted in the civilian domain and applications can be seen in airports, passport applications, and access control, to name a few.

The Face Recognition market has clearly developed in the direction of civilian applications. However, the technology must overcome issues such as variance in expression, pose, illumination and low image quality (or resolution) in order for it to take advantage of growth opportunities provided by the availability of supporting technologies such as low cost cameras/sensors, which can be embedded or integrated into mobile phones, PDA's, and desktop computers. These technologies will enable the Face Recognition market to grow in areas that are directed towards personal use of access control for such devices but also potentially in mobile identification application areas.

The application of Face Recognition has very clear benefits in Law Enforcement and could potentially revolutionise the use of facial images and CCTV data in policing. Much of this data is readily available, but redundant, as it is too difficult and time consuming to process. However, searching police image data, using Face Recognition, remains a significant challenge and will always result in lower performance figures than are seen in the civilian domain, as the majority of police data is from poor quality CCTV, captured in uncontrolled conditions, with a need to search it in identification mode; though this is not to say that it is of no value for Law Enforcement purposes. In spite of its various shortcomings, there are some clear benefits that can be made from deploying Face Recognition in the Law Enforcement arena, and particularly so in the case of non real-time or forensic analysis of police data.

The use of the technology for searching 2D mugshots (i.e. 'controlled to controlled' searching) can provide significant leads for intelligence purposes and may allow for fast probing of multiple mugshot collections between forces. Incorporating such a capability in conjunction with FIND would enable the detection of duplicates in a force's legacy mugshot collections whilst also providing an opportunity to evaluate the technology. The creation of mugshot watchlists on PDA's or other mobile media may provide useful opportunities where face recognition may help to identify threats and/or capture known or prolific offenders such as football hooligans, sex offenders, and other targeted individuals thus preventing crimes. In terms of forensic application of the technology the NCIS's ChildBase system is cited as a successful implementation of face recognition technology whereby it is being used as a tool to aid searching and comparison of internet images of child pornography. It has already provided clear benefits to intelligence operations and investigators by significantly reducing the manual overhead in these operations. Another

example of where this may be achievable is in the viewing and compiling of witness albums and assisting with the creation of facial composites; adding reliability and confidence in the overall process and resulting identifications.

The use of 3D recognition has many advantages to 2D applications. This is already apparent in the type of access control applications available. However, police data will always be based on 2D images caught from CCTV or mugshot and thus regardless of how well 2D approaches may develop to overcome the issues of uncontrolled capture it will always be hampered by variations in pose. Arguably, the most significant edge offered by 3D is in its suitability for use in policing. 3D approaches provide robustness to pose, illumination and to some extent expression. There could be significant benefit for the police in moving towards the capture of 3D mugshots as this will potentially allow for images to be identified from any angle caught on CCTV whilst still allowing for 2D/3D mugshots to be searched too. Furthermore, with 3D capture one might also be able to capture a second biometric, namely ears which are widely used in manual facial identification and available as forensic evidence (*though not currently used*). Note also that the difference in cost between 2D and 3D systems is decreasing and the publication of 3D biometric data standards for face by ISO in the near future will provide further stimulus for the market.

PITO's Biometrics team has a key role to play in managing stakeholder expectations of face recognition technology However, the team must also ensure that the police community fully appreciate the enormous potential of the technology as an investigative tool to aid forensic or post event analysis of police data.

Albeit through the development of alternative approaches to image capture, or the development of forensic tools for verification or image enhancements, face recognition technology can be developed so that it is ready for use in policing. Even so, the police user community has to understand where and when novel approaches are of real benefit when procuring systems. For example, the high performance improvements that skin texture analysis provides to face recognition are based on the use of high resolution images, thus would not apply to police data which is generally from poor quality CCTV. It is clear then that PITO must maintain its expertise in this area to closely monitor developments in Face Recognition in order to be able to provide informed advice on when and where new advances in the technology are able to offer real benefit to policing.

10 Recommendations

A number of recommendations have arisen through the investigation of face recognition technology and the current state of the market, as described in this report. Many of these recommendations are aimed at informing both industry and the police community, as well as informing PITO's own future business planning in the area of face recognition.

Therefore, the recommendations that follow from this report are separated as follows:

- Section 10.1 below focuses on the areas of technological improvement and identifies opportunities for policing. It is intended for stakeholders such as ACPO, industry, the police service, wider CJS stakeholders and academia, and it will also be used to inform the development of the business case for national roll out of face recognition (i.e. PITO's Autoface project)

- Specific 'Recommendations for PITO' are described separately, in Appendix I of this report, for the purpose of informing future business planning within PITO / NPIA with regards to continuing efforts in this area.

10.1 Technology Development and Opportunities

The following recommendations are based on areas that have been identified as most likely to enable key, future capabilities, of benefit to the police service, to be achieved as a result of current and future technological development.

Recommendation 1: Explore the benefits of 3D mugshot capture

Continuing to only capture 2D mugshots in custody will limit the use of this data for identifying persons caught on CCTV. Acquiring 3D mugshots in custody offers the potential benefit of making it easier to identify a person from any profile or angle from CCTV, simultaneously capturing a person's ear image, and potentially future proofing force legacy data in anticipation of one day searching both 2D and 3D facial images. Note also that 3D capture is rapidly becoming not only cheaper, but faster too. This has particular advantages in that, capturing data from non–cooperative subjects in custody may be far easier using 3D than capturing 2D FIND compliant mugshots. Finally, 3D capture systems can simultaneously capture a 2D mugshot and will almost always result in an image that is of sufficient quality for use with face recognition.

Opportunity: **To pilot 3D mugshot capture and 3D to 3D mugshot searching at point of entry in custody**

Benefits: **Fast capture of image in custody, more opportunity to identify people from CCTV, a non contact form of identification in custody**

Recommendation 2: Automated Video Analysis and Forensic Identification

Though much development in both intelligent automated CCTV and Face Recognition is required before the two can be used together successfully there is much that the two technologies can do to improve each other's capabilities. Face Recognition may provide automated means for real time or offline face detection and tracking to help process video data. Intelligent automated processing of video footage, high quality CCTV and informed placement of cameras may provide data from such sources that may be usable by face recognition systems. Developments in both CCTV should be made with a view to couple the technologies. Moreover systems that gather information from these technologies should be designed in a way that allows effective management and analysis of the data in real time (on line) and offline, forensic examination and management.

Opportunity: **To pilot Face Recognition with CCTV, based on an integrated model of Intelligent Video surveillance and analysis with a Forensic Identification Interface.**

Benefits: **Real operational evidence of the merits of Face Recognition with CCTV in policing, investigative leads found faster and thus making better use of this abundant source of forensic evidence**

Recommendation 3: Combining cognitive approaches with User Interface Development

Facial identification and searching of facial images from CCTV will undoubtedly require operator intervention when working with police data. The technology should be developed to incorporate human operator expertise, for example; to extract faces from the image data, 'encoding' or marking up' the data prior to launching a search, and in providing viewing tools that assist an operator to view and verify matches that are returned.
Face Recognition is currently fully automated thus in order to improve its use with CCTV data, development of tools that exploit human cognitive visual approaches to facial identification may prove to be of value in improving the performance of face recognition, much in the way that fingerprint matching technology is assisted by trained fingerprint experts.

Opportunity: **To trial systems that explore the use of innovative HCI tools (for example Identix facial search interface) in order to evaluate the performance improvements they can provide in comparison to fully automated searching.**

Benefits: **Development of HCI tools and Face Recognition systems such that the chances of detecting individuals and making identifications is optimised**

Recommendation 4: Targeted applications for specific user scenarios i.e. serious and organised crime

Whilst a centralised national search capability may be appropriate in some circumstances, it is more likely that a nationally co-ordinated rollout of targeted local applications may be a better approach to meeting police requirements. From the scenarios described in Section 4, it is clear that there is no one method of implementing face recognition that would benefit operational policing. Rather face recognition has the potential to provide numerous capabilities which should be explored separately. Mobile identification using PDA's or other image capture device, combined with watchlist searching, crowd surveillance at open public environments such as shopping centres and stations, are examples of where face recognition may help to combat serious and prolific crime, whereas mugshot to mugshot searching may provide opportunities to identify people when detained in custody.

Opportunity:

To pilot the use of a face recognition with CCTV images whereby cameras are strategically placed at choke points in public areas (i.e. stairwells at stations, near posters or signs, or at entry points to grounds) in order to ensure a good frontal image of a face can be captured, and subsequently be used to search against a watchlist (of similar images) of known suspects and offenders.

To pilot 2D mugshot to 2D mugshot searching with the roll out of FIND to detect duplicates in the FIND collection.

To pilot Face Recognition in support of Anti terrorist Intelligence Operations.

Benefits: **Centralised, small scale, targeted deployments of the technology thus more suited to a specific business need and more likely to operate at a higher level of accuracy that on a national scale resulting in better ROI and more evident efficiency gains.**

Recommendation 5: Deployment of High Definition CCTV with a view to capture facial image data

As part of the National CCTV Strategy, it is possible that the number of cameras being deployed within the UK will increase significantly and this will include movement towards high quality CCTV camera capture. However, it is not only the camera quality that matters, but more importantly the high levels of compression typically applied that degrades CCTV data such that it is virtually unusable.

There are little, or no, independent figures available on the performance of face recognition with high quality CCTV data. With a better understanding of what is achievable with high quality capture and storage, more and more (police and public) stakeholders may start to invest in such technologies. This raises te question of whether reliable, automated searching of CCTV against CCTV data will be viable in the future;. The answers to such questions are needed in order to manage the expectation of what can be achieved with face recognition for policing and to justify investment in this area.

Opportunity:

To pilot or perform evaluative benchmarks of Face Recognition technology when used with High Definition CCTV for mugshot to CCTV and CCTV to CCTV searching.

Benefits:

Ascertain state of the art and set the benchmark for what is highest level achievable. It will inform whether Face Recognition with CCTV is worthwhile exploring at all is the accuracy with even the best quality of images from CCTV is poor.

Recommendation 6: Explore the potential of Face Recognition at range (using LIDAR)

Long range acquisition of facial image data could be extremely beneficial for covert identification. Furthermore, to acquire 3D facial image data with LIDAR suggests that identification of individuals is not only possible when coupled with 3D mugshot data, but also it will allow for robust and accurate linking of suspects in multiple samples of surveillance footage and over long periods. This technology may result in significantly innovative approaches for future face recognition surveillance technology.

Opportunity:

To evaluate the potential of Face Recognition at range using LIDAR technology, perhaps specifically working in areas of covert surveillance such as anti terrorist operations or operations to combat serious and organised crime to pilot such technologies with 3D face recognition.

Benefits: **Fast, accurate and covert identification of individuals at range and widening the opportunity to identify individuals of interest that may not be identified otherwise.**

Recommendation 7: Close monitoring of emerging standards in biometrics for 3D face and MPEG (MPEG7)

Another important issue, which is extremely relevant to both future automated face recognition and to FIND is that of Standards. A revision of the ANSI NIST ITL-2000 standard[12] is currently underway, and a large number of ISO biometric standards are either published, or soon will be. These include, of particular relevance here, 2D and 3D face image data standards (ISO 19794 – 5). However, important related work is also going on in regards to MPEG standards (particularly MPGE7) looking at 'Advanced Face Descriptors' and 'Advanced Image Coding and Searching', and ISO 15378 Part 8 will apply MPEG7 to photo management, allowing searching of databases for duplicate images, changed images, or for specific scene elements, such as people and vehicles. Whilst some of the technology to do this is still at the research stage, adoption of such standards will be essential in enabling large (possibly distributed) databases of images to be efficiently managed, manipulated and searched, whether using face recognition technology or other search descriptors. It is essential to maintain a close watch on developments in this area.

[12] ANSI/NIST-ITL 1-2000, Data Format for the Interchange of Fingerprint, Facial, & Scar Mark & Tattoo (SMT) Information

Appendix I: Recommendations to PITO

This section details the recommendations of how PITO and the Biometrics team should proceed with regards to face recognition in light of the findings of this report.

1. **The Biometrics team should continue to work with the FIND team to ensure that future requirements for automated face recognition are taken into consideration at the design stage. Moreover, ACPO should be encouraged to consider the merits of moving away from traditional 2D methods of capturing mugshots and instead advance to 3D or multiview based capture of facial images that will better facilitate the use of custody images for identifying images caught on CCTV and will aid witness identification.**

It is clear that face recognition with 2D mugshots will always be limited in its application for law enforcement whereas 3D capture will potentially provide very significant advantages over traditional 2D methods and opportunities for matching faces with face recognition that would otherwise be unachievable. PITO should therefore look to build on the FIND standard for mugshot capture and influence ACPO to push for the capture of 3D mugshots in custody suites. The FIND collection is still in the early stages of delivery and thus it is economically and strategically advantageous to move to 3D capture of mugshots as soon as possible. In doing so a collection of operational 3D data will be available for use with face recognition albeit 2D, multiple views based or 3D data, and thus future proof the database against improvements in the technology as it develops.

2. **PITO should maintain and expand its links with academia to drive research in directions that are likely to deliver real benefits to the police service.**

Over the past four years the Biometrics team has extensively engaged with academia and has established a very strong and active programme of collaboration with academia. The success of research proposals such as those supported by EPSRC and the range of disciplines that they cover are reflective of the team's awareness and understanding of the wider context of how these applications must function in reality.

The police requirements of face recognition (and biometrics in general) are fundamentally different to the civilian domain and these diverse challenges are now being acknowledged and appreciated by the industry and academia. The team has been successful in promoting research into areas to develop the technology to meet the police services diverse needs. PITO should build on its collaboration with academia; perhaps by directly funding research and/or by taking on students under placements to work on the team.

One of the key obstacles in furthering face recognition is the availability of test data sets. The team is planning a large scale data capture exercise of CCTV and mugshot data in collaboration with Kent University, in order to use this data for in house evaluations of face recognition as part of the FRED strategy and in conjunction with its support of Kent's EPSRC proposal of research into the *"Recovery of video forensic identification data for use with face recognition"*. The lack of such data has hampered progress in the development of the technology thus this data set will be an invaluable resource for academic research and should ideally be made widely available to academia and algorithm developers.

The team's collaboration with Kent is an opportunity to share resources with the university and further develop its in house expertise in face recognition. The team should thus seriously consider defining

work areas or establishing roles and responsibilities that sandwich or MSc students from the university may be able to fulfil on the team.

3. PITO should maintain its links with industry to ensure that PITO, and in particular the Biometrics Team, is fully aware of new developments in this area, and ensure (e.g. via attendance / presentations at conferences) that vendors appreciate the challenges of automated face recognition in the law enforcement arena.

Since the July attacks in London and as identified in the Roadmap increasing amounts of attention is being made on the need for more efficient processes by which to use CCTV evidence for intelligence and surveillance. Face Recognition could have a dramatic impact on this capability and thus is an important area for PITO to monitor despite its current limitations. The team must therefore maintain its expertise in face recognition technology in order to influence the development of the technology and be in a position to advise on the suitability of the technology as it matures. Face recognition is rapidly maturing for the civilian domain but similar levels of performance may not be achievable for law enforcement applications. The team must look to manage the police community's expectations and ensure that they are realistic, and understand the potentials and limitations that face recognition can offer.

4. PITO should take every opportunity to speak to the user community, at all levels, in order to ensure they have realistic expectations of what automated face recognition will be able to deliver.

In developing the business case for face recognition it should be made clear that whilst the technology is currently immature for use with CCTV there are some benefits that can be achieved in the short term by deploying face recognition. The business case should look at the potential for wider use of the technology in applications such as the Lantern project whereby face recognition used with watchlists may facilitate mobile identification, the automatic generation of witness albums, and mugshot to mugshot probes of force collections as a means of establishing identity in where fingerprints are not appropriate or available. Wider CJS stakeholders such as the CRB should also be consulted to determine whether face recognition may be suitable as an identifier, and as a means of tracking individuals throughout the CJS cycle, from arrest, through to bail, and release. Use of one's face may be a more socially acceptable means by which to check for criminal history (as opposed to the use of fingerprints) as well as a more secure check than that currently performed by CRB. Checks are currently performed on name, gender and date of birth which can be easily falsified when claiming an identity. Access control applications based on face recognition may be suitable for use in prisons, courts and areas such as bail reporting if not now then in the near future. In these applications it is less the reliability of the face recognition than it is the lack of supporting IT or data infrastructure that constrains the use of the technology.

5. It is vital to continue to work closely with other government agencies (both within the UK and internationally) with an interest in automated face recognition, such as HOSDB's Biometrics Centre of Expertise and UKPS (now IPS), in order to ensure that lessons are learned and duplication of effort is avoided.

It is vital that PITO, through the Biometrics team, maintain close links with other stakeholders currently looking at face recognition. For instance, HOSDB has extensive knowledge in the area of CCTV and video data capture which PITO's team must draw on and IPS has performed trials of face recognition

and is one of the first to have explored the development of HCI tools or automated viewing and enhancement techniques that may be used to facilitate human face recognition. The Biometrics team can thus benefit from their experience in these areas to inform its own trials and inform further R & D into the development of tools to assist forensic analysis and searching of facial image data.

PITO in turn has a considerable expertise in testing and evaluation of biometrics technology in general and more importantly in the procurement of large scale identification systems. There is a great deal that PITO can offer by way of guidance and expertise in these areas relating to face recognition to stakeholders across government(s).

6. The Biometrics team should continue to participate in relevant standards bodies (ISO, ANSI NIST) to ensure Standards meet the needs of the police service. Consider expanding participation beyond SC37 to include related standards working groups such as MPEG;

Another important issue, which is extremely relevant to both future automated face recognition and to FIND is that of Standards. It is essential that PITO not only maintains a close watch on developments in this area, but also provides input where appropriate through its continued participation in IST44 / SC37. A revision of the ANSI NIST ITL-2000 standard is currently underway, and a large number of ISO biometric standards are either published, or soon will be. These include, of particular relevance here, 2D and 3D face image data standards (ISO 19794 – 5). However, important related work is also going on in regards to MPEG standards (particularly MPGE7) looking at 'Advanced Face Descriptors' and 'Advanced Image Coding and Searching', and ISO 15378 Part 8 will apply MPEG7 to photo management, allowing searching of databases for duplicate images, changed images, or for specific scene elements, such as people and vehicles. Whilst some of the technology to do this is still at the research stage, such standards will be essential in enabling large (possibly distributed) databases of images to be efficiently managed, manipulated and searched, whether using face recognition technology or other search descriptors. Consideration should be given to providing additional resources in order to enable PITO to actively participate in the MPEG working groups.

7. The Biometrics team should continue with the in-house face recognition evaluation and testing project, as laid out in the FRED strategy. (Ref. 11) and ensure the results of such tests are documented and fed back not only within PITO but, where appropriate, also made available to industry and academia. Every effort must be made to acquire operationally representative data for testing, and where possible to also make such data widely available to researchers and algorithm developers;

The FRED project has already provided a great deal of useful information on a wide variety of issues surrounding the use of automated Face Recognition. This has fed directly into projects such as FIND and the development of the business case for 'Autoface'. The Biometrics team has investigated a variety of technologies, using both 2D and 3D images as well as techniques to manipulate and combine multiple 2D views, and it has become clear that some of these are far more suited than others to the law enforcement domain, and in particular to use in forensic investigations. Through the FRED project the team has been able to establish an good level of in house knowledge and expertise in the area of Face Recognition.. Partly as a result of the team's work in this area, it has become apparent that some companies are now starting to address the issues specific to law enforcement. . The team should continue its work in this area.

Appendix II: Face and Pattern Recognition Algorithms

The following is a list of algorithms that are commonly associated with Face Recognition technology. These methods are extensively documented within literature on Pattern Recognition as well as in papers on Face Recognition. For ease of reference, many of the definitions below have been taken from a single source, in this case the Face Recognition Home Page. To view the reports directly follow the link www.face-rec.org

Principle Component Analysis (PCA)

This is a very common approach used in face recognition that is derived from general pattern recognition techniques. This method identifies the distinct characteristics of an image (albeit a facial image) and sorts them to pick out "features" of the image/ face which vary the most from the rest of the image. This methods thus selects the most distinct features with which to represent the facial image by. These are used as the base characteristics for matching the image thus reducing the complexity of matching by having fewer features that have to be looked.

See Reference:

- *M.A. Turk, A.P. Pentland, Face Recognition Using Eigenfaces, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 3-6 June 1991, Maui, Hawaii, USA, pp. 586-591*
- *M. Turk, A. Pentland, Eigenfaces for Recognition, Journal of Cognitive Neurosicence, Vol. 3, No. 1, 1991, pp. 71-86*

Linear Discriminant Analysis

Linear Discriminant Analysis is another principle derived from matrix algebra and used in pattern recognition. It is essentially a form of linear classification. It allows points in a dataset (for example faces in set of facial images) to be separated by maximising the variance between faces and minimising variance within faces. Linear discriminant analysis finds the underlying vectors in the facial feature space that best separates the faces of different subjects within a dataset of facial images and minimising the difference within a group of points that represent that same subject in the set.However, LDA can only separate out the different classes of features where there is a linear relationship between the features. But faces are non linear; thus there will always be an error of classification where there exists an overlapping of features (data points in feature space) that cannot be clearly separated by a line between the two classes. In such cases other methods are looked at such as nearest neighbour or other more complex non linear classification techniques.

See Reference:

- *K. Etemad, R. Chellappa, Discriminant Analysis for Recognition of Human Face Images, Journal of the Optical Society of America A, Vol. 14, No. 8, August 1997, pp. 1724-1733*

Independent Component Analysis (ICA)

Independent Component Analysis (ICA) is a principle from pattern recognition similar to PCA whereby instead of pulling out the maximum variations ICA pulls out features that vary in an image that are statistically independent (or have no correlation) and separated from variations in an image that are correlated.

See Reference:

- *M.S. Bartlett, J.R. Movellan, T.J. Sejnowski, Face Recognition by Independent Component Analysis, IEEE Trans. on Neural Networks, Vol. 13, No. 6, November 2002, pp. 1450-1464*

Evolutionary Pursuit (EP)

An adaptive approach that searches for the best set of projection axes in order to maximize a fitness function, measuring at the same time the classification accuracy and generalization ability of the system. Because the dimension of the solution space of this problem is too big, it is solved using a specific kind of genetic algorithm called Evolutionary Pursuit (EP). In short having applied PCA the EP algorithm adaptively selects the best set of projection axes that describes the feature set whilst simultaneously reducing the chance of misclassifying the features.

See Reference:

- C. Liu, H. Wechsler, Evolutionary Pursuit and Its Application to Face Recognition, IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 22, No. 6, June 2000, pp. 570-582

Evolutionary Computing

A population training based technique that uses the probability densities of the population entities to find the best combination of variables.

Elastic Bunch Graph Matching (EBGM)

Represents faces as a graph and exploits similarities in the topological structure of human faces. Faces are represented as graphs, with nodes positioned at fiducial points. (exes, nose...) and edges labelled with 2-D distance vectors. Each node contains a set of complex Gabor wavelet coefficients at different scales and orientations (phase, amplitude). Recognition is based on the labelled graphs looking at the connection of by edges, and distances.

See Reference:

- *L. Wiskott, J.-M. Fellous, N. Krueuger, C. von der Malsburg, Face Recognition by Elastic Bunch Graph Matching, Chapter 11 in Intelligent Biometric Techniques in Fingerprint and Face Recognition, eds. L.C. Jain et al., CRC Press, 1999, pp. 355-396*

3-D Morphable Model

A 3-D morphable face model encodes shape and texture in terms of model parameters, and exploits an algorithm that recovers these parameters from a single image of a face.

See Reference:

- *V. Blanz, T. Vetter, A Morphable Model for the Synthesis of 3D Faces, Proc. of the SIGGRAPH'99, 08-13 August 1999, Los Angeles, USA, pp. 187-194*

Support Vector Machine (SVM)

Given a set of points belonging to two classes, a Support Vector Machine (SVM) finds the hyperplane that separates the largest possible fraction of points of the same class on the same side, while maximizing the distance from either class to the hyperplane. PCA is first used to extract features of face images and then discrimination functions between each pair of images are learned by SVMs.

See Reference:

- *G. Guo, S.Z. Li, K. Chan, Face Recognition by Support Vector Machines, Proc. of the IEEE International Conference on Automatic Face and Gesture Recognition, 26-30 March 2000, Grenoble, France, pp. 196-201*

Artificial Neural Networks

These are computational methods that model neural network processes in the brain whereby networks algorithms are trained on training samples of data to learn and predict outputs.

See Reference:

- *Principal Component Analysis and Neural Network Based Face Recognition, Qing Jiang http://people.cs.uchicago.edu/~qingj/ThesisHtml/*

Glossary

Access Control
Prevention of unauthorised access or use of a resource.

Acceptability
Indicates how a user or public feels about a system.

Acquisition
Acquiring a facial image.

AFIS
Automated Fingerprint Identification Systems.

Algorithm
A mathematical principle applied to form a computational process.

ANPR
Automatic Number Plate Recognition.

ANSI
American National Standards Institute.

Appearance Based Models
A computerised methodology for recognition based on extracting low level features in an image for recognition.

Arrestee
A person arrested by the police.

Artificial Neural Networks
A computing method that emulates neural processes in the human brain; learning by past experience.

ATM
Automated Telling Machine.

Authentication
The process whereby an individual is checked to ensure that he/she is a valid person.

Automation
A process that is performed computationally either in part or full.

Biometrics
The automated process of establishing identity based on physical or behavioural characteristics of a person.

BSI
British Standards Institute.

BWG
Biometric Working Group.

CCTV
Closed Circuit TV.

Controlled Image
A facial image captured in controlled environment to meet specified standard. For example a passport photo or mugshot.

CRB
Criminal Records Bureau.

Decision
The output of a matcher declaring two facial images as a match or non match.

DFAT
Department of Foreign Affairs and Trade

Elastic Bunch Graph Matching (EBGM)
See Appendix.

Enrolee
A person who has a template on file with application.

Enrolment
A process of collecting the biometric sample (face image) and storing it with an identity of the application database.

EPSRC
Engineering and Physical Sciences Research Council

False Accept Rate (FAR)
The number instances or the rate that an application incorrectly matches two images as belonging to the same individual.

False Reject Rate
The number of instances or the rate that an application fails to match two images belonging to the same individual.

Face Recognition (FR)
A biometric that uses a face as an identifier.

Feature
A particular detail pertaining to the image or biometric data within it. A feature may not necessarily be a physical entity.

Fiducial Points
Local feature points on the face.

FIND
Facial Images National Database – PITO project to deliver national mugshot collection to forces across the UK.

Fingerprint
The pattern of ridges and valleys that is deposited as a finger comes into contact with a material or surface.

FRGC

Face Recognition Grand Challenge

FRVT
Face Recognition Vendor Test

Full face (or, 'full frontal' face)
This refers to the profile of a person when facing straight ahead at a camera.

Gallery
A collection of electronically stored images that is referenced by an application to find a possible match. Also called the background database.

Gallery Image
An image stored within the Gallery collection.

Head Pitch
Displacement of the face/head position in the vertical direction. See also Head Yaw

HOSDB
Home Office Scientific Development Branch

ICAO
International Civil Aviation Organisation

IDENT1
National Strategic Identification Services Platform (SISP) delivered by PITO providing national search capabilities of multiple biometrics, namely fingerprints, palms, possible extending to faces, DNA, earprints, footprints etc in future. (IDENT1 replaced NAFIS in December 2004).

Identification
A one-to-many search and comparison of an image for the purpose of finding a match to an individual's image

IPS
Identity and Passport Service (See also UKPS)

ISO
International Standards Organisation

JPEG
Stands for "Joint Photographic Experts Group". JPEG is standard format for storage compression of static digital images. JPEG2000 is a recent release of the JPEG standard providing Region of Interest (ROI) compression and addresses the problems of visual degradation such as graininess in poor quality or low resolution images.

Lantern
PITO project, piloting mobile identification of individuals at the roadside using fingerprints captured on a ruggedised PDA with embedded optical fingerprint sensor.

Liveness Detection
Liveness Detection is a process embedded in the system at the sensor to authenticate that the person using the system and presenting there biometric to the sensor is indeed genuine - that is from a living, human being.

Multiples
More than one image of the same individual stored in a gallery collection. Also referred to as “duplicates”.

Match
When two images are classed as belonging to the same individual.

MPEG
Moving Pictures Expert Group. Standard for video/dynamic image compression.

NAFIS (National Automated Fingerprints Identification System)
NAFIS was replaced by the IDENT1 system. It provided fingerprint bureaux of England and Wales with the capability to search fingerprints from crime scenes, and ten prints from custody, against the national collection of unidentified crime scene marks and the national fingerprint collection.

NIST
National Institute of Standards and Technology

NVIS
National Video Identification Strategy

PNC (Police National Computer)
Stores all criminal history data of persons arrested or convicted of a crime in UK. Managed and developed by PITO.

Post Event Analysis
This term is used to describe analyses or operations in non real time after an event such as a crime or serious incident has occurred. For example forensic analysis is a post event analysis function.

Probe
An image that is presented for automated matching against one or many images from a collection of electronically stored images.

UKPS
United Kingdom Passport Service. Now known as Identity and Passport Service (IPS)

Uncontrolled Image
A facial image captured under ad hoc or uncontrolled conditions. For example CCTV.

Verification
A one to one comparison of two images for the purpose of validating a that two images do indeed belong to the same subject.

References

1. Intelligent Biometric Techniques in Fingerprint and Face Recognition – *L.C Jain et al., CRC Press 1999.*

2. Eigenfaces for Recognition – *M Turk, A Pentland, Journal of Cognitive Neuroscience Vol.3 1991.*

3. Evolutionary Pursuit and its Application to Face Recognition – *C. Lui and H.Wechsler, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 22, June 2000.*

4. Dynamic Vision: From Images to Face Recognition – *Shaogang Gong et al. Imperial College Press, 2000.*

5. Face Recognition Using LDA based Algorithms – *J Lu et al. IEEE Transaction on Neural Networks, Vol. 14, 2003.*

6. The Biometric Industry Report – Market and Technology Forecasts to 2003, *1st Edition, Elsevier 2003.*

7. Biometric Technology Today: Face Recognition Part 11 – *May 2006, Elsevier (www.compseconline.com)*

8. Face Recognition: A Literature Survey – *Zhao et al. NIST 2000*

9. Identification Roadmap 2005 – 2020 Part 1: *UK Police Information Technology Organisation, 2005.*

10. ISO/IEC JTC1/ SC37 – N1511 – Face Image Data Standards on Conditions for Taking Pictures – *March 2006*

11. PITO Face Recognition Test and Evaluation (FRED) Strategy – *July 2005, Ambika Suman.*

12. FREDS Data Capture Requirements – *November 2005, Ambika Suman, PITO.*

13. Face Recognition Grand Challenge – *Face and Gesture, Jonathon Philips, NIST, 2006.*

14. Face Recognition Vendor Test 2002 – *Jonathan Philips, NIST, 2003*

15. FRS1 Project Report 2006 – Identity and Passport Service (IPS)♦

16. FRS1 Pilot Evaluation Report 2006 – Identity and Passport Service (IPS)♦

17. G. Guo, S.Z. Li, K. Chan, Face Recognition by Support Vector Machines, Proc. of the IEEE International Conference on Automatic Face and Gesture Recognition, 26-30 March 2000, Grenoble, France, pp. 196-201

♦ Report circulated as "RESTRICTED"

18. A Morphable Model for the Synthesis of 3D Faces, *V. Blanz, T. Vetter, Proc. of the SIGGRAPH'99, 08-13 August 1999, Los Angeles, USA, pp. 187-194*

19. Face Recognition by Elastic Bunch Graph Matching, *L. Wiskott, J.-M. Fellous, N. Krueuger, C. von der Malsburg, Chapter 11 in Intelligent Biometric Techniques in Fingerprint and Face Recognition, eds. L.C. Jain et al., CRC Press, 1999, pp. 355-396*
20. Evolutionary Pursuit and Its Application to Face Recognition, *C. Liu, H. Wechsler, IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 22, No. 6, June 2000, pp. 570-582*

21. Face Recognition by Independent Component Analysis, *M.S. Bartlett, J.R. Movellan, T.J. Sejnowski, IEEE Trans. on Neural Networks, Vol. 13, No. 6, November 2002, pp. 1450-1464*

22. Discriminant Analysis for Recognition of Human Face Images, *K. Etemad, R. Chellappa, Journal of the Optical Society of America A, Vol. 14, No. 8, August 1997, pp. 1724-1733*

23. Face Recognition Using Eigenfaces, *M.A. Turk, A.P. Pentland, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 3-6 June 1991, Maui, Hawaii, USA, pp. 586-591*

24. Eigenfaces for Recognition, *M. Turk, A. Pentland, Journal of Cognitive Neurosicence, Vol. 3, No. 1, 1991, pp. 71-86*