|  |
| --- |
| University of reading |
| A Face Authentication System |
| Final Year Project |
|  |

**Project Supervisor: Hong Wei**

**Project ID: A-FACE**

**Module Code: SE3IP11**

**Student Name: Tom Bedford**

**Student Number: 21009807**

**Submitted: TBD**

|  |
| --- |
|  |

**Abstract**

Computerised Face authentication is vastly being integrated into today’s technologies and systems. As an efficient and effortless method of character recognition this technology aspires to be prevalent in every-day activities in the aims of managing data and location access and delivering customized user experiences. This report illustrates the design, implementation and effectiveness of a proof of concept facial authentication system. The system utilises Viola-Jones methodology of feature extraction and PCA (Principal Component Analysis) to derive a fast and efficient means of digital face detection and verification. It features a configurable threshold acceptance value in order to handle a range of illumination levels which provide a challenging constraint in image analysis. The developed system can potentially complement an existing user authentication layer or operate as an independent authentication system. Its application can be easily extended to provide some form of greeting message or profile to future students attending a university open day

**Acknowledgments**

Hong Wei – project supervisor

Contents

[Glossary of Terms and Abbreviations 2](#_Toc448872262)

[Introduction 3](#_Toc448872263)

[Problem Articulation / Technical Specification 3](#_Toc448872264)

[The Solution 5](#_Toc448872265)

[Literature Reviews 6](#_Toc448872266)

[Implementation 8](#_Toc448872267)

[Testing: Verification and Validation 9](#_Toc448872268)

[Discussion 9](#_Toc448872269)

[Conclusion 9](#_Toc448872270)

[Project Commentary 9](#_Toc448872271)

[Social, Legal, Health & Safety and Ethical Issues 9](#_Toc448872272)

[Reflection 10](#_Toc448872273)

[References 10](#_Toc448872274)

[Appendices 11](#_Toc448872275)

## Glossary of Terms and Abbreviations

OPENCV – Open computer vision

PCA – Principal Component Analysis

Eigen Faces – definition and ref

Haar-like feature – geometrical feature extraction

## Introduction

Biometric analysis has been a focused area of research for many years due to its distinctive features, minimally invasive recognition methods and potential automation. Technological advancements have provided more powerful and affordable hardware and open sourced software libraries allowing the technology to be researched more openly. Although biometric evaluation such as finger print recognition has delivered more reliable results than face recognition it requires additional expensive hardware to install and requires focused interaction from its user. With the least overhead for system instalment, least invasive recognition technique and fastest functionality facial recognition has become the most popular choice of biometric analysis.

Facial recognition is the identification and verification of noticeable characteristics of a human face. In the field of image analysis it is leading the race in research as its speed and versatile application out weights previous methods. It is less invasive than other biometrical analysis methods like retinal and finger print recognition with less hardware requirements. Human to computer authentication is an integral functionality of many software systems as it manages data and/or location security. Facial recognition is a task performed effortlessly by humans on a daily basis. The sheer complexity of the human brain can truly be admired whilst replicating this complex functionality in machine software. Distinctive features of the human face such as the nose, mouth and eyes are algorithmically recognised by object detection and association to geometrical shapes.

This project explores the application of face detection and verification and its effectiveness as an authentication based system using the open vision library developed by IBM.

## Problem Articulation / Technical Specification

The fundamental goals of the developed system in this study is the successful detection of a user’s face and the association of the face image and the target databases facial images within a set threshold to grant a user access through the login interface. The threshold of acceptance will be adaptable to suit the needs of the systems environment which substantially dictates acceptance rates in the recognition process. Illumination and user positioning are the greatest defining factors in the algorithm.

It is assumed that users will present there face at an offset no greater than 40 degrees and the camera will be of sufficient quality to capture images to an adequate resolution. The user will not be wearing items that obstruct large portions of the face. To provide a reliably secure authentication system recognition results must fall within a refined acceptance threshold that ensures access only to known users of the system. These assumptions and constraints where originally defined in the PID [] report referenced at the end of this document which was submitted at the beginning of the project.

To satisfy the project deliverables derived in the PID each deliverable must meet its acceptance criteria.

The proposed solutions are based on addressing three main modules of the system’s functionality, the GUI (Graphical User Interface), face detection and verification functionality.

The GUI must be developed to provide a simple interface for the user to detect and capture an image of their face to be submitted as part of the authentication process. The GUI will also need to host a training screen where the user can select, load and train a face database prior to face recognition.

**Figure. 1** face authentication graphical user Interface design.

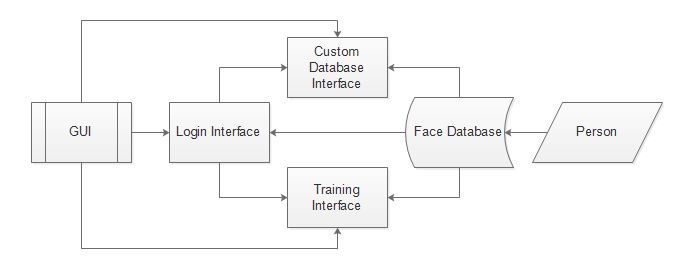
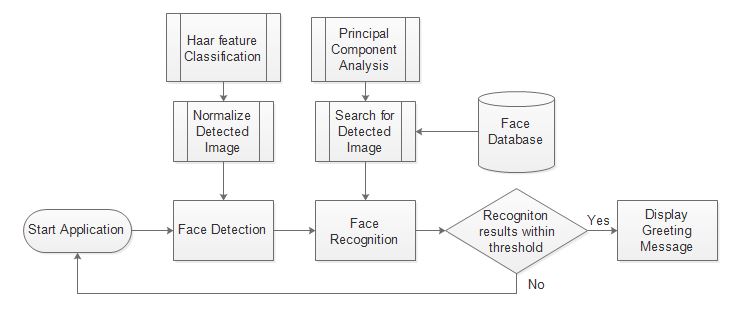


Figure 1 illustrates the work flow of the proposed GUI. It can be seen that the programs user interface is built up of three main displays, the login, training and custom database screens. The login and training screens are essential to the systems functionality for loading and training data and the custom database screen is to provide further customisation for creating and adapting databases. Developing the GUI in Java language will be less complex and time consuming than C++. This solution will be developed in a ported version of the OpenCV library natively written in C++ too Java [Java CV ref]. The OpenCV java library essentially wraps the C++ functionality in Java so that C++ code can be utilised in Java code. JavaFX will be used to develop the systems GUI allowing it be implemented as an MVC (Model View Controller) structure. This will allow the main algorithms of the system to be used by other systems easily.

**Figure 2.** Recognition process of the facial authentication system.



.

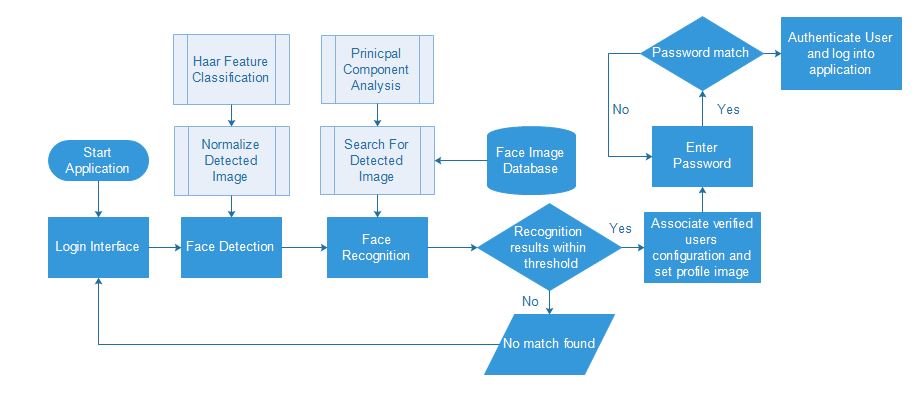
The face detection functionality will be capable of acknowledging a human face by recognising geometrical features of a human face with the use of trained Haar feature cascade classifiers provided by the Open CV library [Haar]. The classifier iterates through image pixel data to locate any know geometrical shapes whilst optimising the search results using an Adaboost algorithm [Adaboost]. To satisfy the systems acceptance criteria detection must successfully capture a frontal facing image within an offset of 40 degree angle at a success rate of 90%.

Detected images then need to be normalised to conform to standardized image dimensions of the system and reduce computation. This is achieved by resizing the detected image and then converting the colour channels to an average greyscale channel.

Principal component analysis will be used reduce and optimise data representation. It is the most popular and efficient LDA’s (Linear Detection Algorithm) in developing facial recognition technologies today. The PCA approach

## The Solution

**Figure. N** Proposed system solution design



The proposed solution as displayed in figure.N will be developed using a Java wrapper of the OpenCV library to enable simpler development of the program user interface so that efforts can be directed towards the main algorithms of the system. The OpenCV library is implemented in C++ programming language. Over the years wrappers and translated versions of the library have been developed to provide the ability for developers to utilise the libraries functionality from other programming languages such as Java. JavaFX library [JavaFX] will be used to implement an MVC (Model View Controller) structure which will allow the main functionality of the system to be utilised or integrated into other systems easily.

The open vison library provides a variety of classifiers that have been developed and trained to detect geometrical shapes. The classifier ‘haarClassifier\_frontalFace\_alt’ that will be used to perform face detection is based on viola jones methodology which incorporates an Adaboost algorithm. Detection of faces can be performed at around a 92% success rate of detection of frontal face images within an offset of roughly 40 degrees. The detected image is then normalized by rescaling and greyscale conversion to reduce computation and conform to dimensional standards of the recognition algorithm.

The prepared image is then added to the targeted database and re-trained to determine eigen value decomposition.

This solution looks to associate the user as a known user at the login stage. A password is then entered to ensure a secure authentication system. It was decided not to use the facial recognition results alone to determine system access as the face detection carried out by the haar feature classifier could potentially be spoofed to gain access.

**PCA (Principal Component Analysis) to refine and optimise data representation**

Acceptance testing through thresholding comparison of trained data sets. The principal components represented as Eigen values and vectors are computed on the database image sets. The user image Is then captured using face detection and added to the relevant database image set. PCA analysis is then run on the updated image set which delivers a deviation percentage from the original analysis. The new image set is accepted if the value deviation is still in the given threshold of acceptance.

Recognition Acceptance:

Recognition acceptance is tested by performing PCA on the original database image set. The new user image is then added to the images set where PCA analysis is performed again. If the new image set PCA analysis falls within an acceptable threshold derived from the original PCA analysis the image is verified and validated. This image is then associated to the relevant user profile and image set.

Eigen Value Decomposition

## Literature Reviews

Methods of human authentication have previously and still are performed using unique credentials such as personal ID cards, passports, unique passwords and phrases. These methods of authentication often require some form of physical interaction or focus from its user. All methods of computerised human identification operates on the basis that a set of uniquely identifiable data is submitted by a user to the system where it is compared and associated to its relevant data within a determined threshold of acceptance. The form of this uniquely identifiable data has largely been an extension of a person such as a bank card or password. Biometrics offer the convenience of no extra materials, assets or memory cognition as they are physical features that individually identify a person.

The face is the most exposed definitive characteristic of a human and therefore can be analysed with little demand of interaction from the person if any. A person need only look towards the scanning device which in this case is a camera for seconds and recognition can be successfully performed in real time or from a captured image frame. Face recognition is a popular method of identification in government surveillance systems [government surveillance]. It can be carried out on multiples entities simultaneously with speed which is extremely effective in crowds of people.

In the infantile stages of facial analysis feature extraction was performed manually. The automation of this process in real time was introduced by Paul Viola and Michael Jones in 2001 [voila jones ref]. Not only does the algorithm detect facial features in real time but performs this with a simpler complexity of computation allowing computers with low processing power to utilise the software. The Viola-Jones method of Haar-like feature classification has revolutionised face detection techniques.

In December last year the company bioID [bioID] released a facial recognition software solution for apple hardware. The software allows the use of face authentication to authorize access to devices such as the IPad and IPhone. It features a patented liveness detection algorithm [liveness detection] to determine whether the user is fact a real person or not. Liveness detection is become a common practice in face detection in the efforts to prevent spoofing and unauthorized access. The software also host a rich development package that can be easily integrated into software development to encourage developers to incorporate facial authentication into future software applications. Other programming language versions of this software are slowly becoming available.

Again last year there was the release of an advanced facial recognition system called DDFD (Deep Dense Facial Detection) [DDFD]. The detection algorithm has been extended from the foundations of Viola-Jones methodology of face detection [viola ref] and uses a convolution neural network to learn a composition of facial features. Impressively this algorithm is trained on databases containing over 21k images. This is one of the most robust face detection systems on the market today as it can successfully detect partial faces, upside down oriented faces and largely obstructed faces.

Cognitec is another company offering multipurpose face recognition system solutions for use in security authentication in fields such as law enforcement, border control and ID fraud protection. The company has been researching facial recognition technology in collaboration with the university of Surrey and Technical University of Dresden since 2010 [cognitec].

In a few months from now Microsoft windows 10 will be introducing a new component dubbed ‘Windows Hello’ [windows hello] to their popular operating system. This new functionality will allow windows users to login using face recognition technologies. Microsoft claim this will be more secure than older authentication methods like passwords and will strengthen computer security as a whole.

There are many organisations investing largely in face recognition development as the vision of effortless human computer interaction becomes a reality. The United States National Intelligence Agency launched a development program in early 2014 named Janis []. The aims of Janis is to compile a broad database of facial morphology derived from online digital media and sources to understand and develop more sophisticated and optimal facial analysis techniques. This rich database will provide an intensive testing plane for facial recognition algorithms and will accelerate their development.

Well renowned security conferences such as the Black Hat [Back Hat] and Def Con [Def Con] conferences attract some of the world’s leading specialists in information security under one roof to diverge and discuss security practices and technologies. An article posted shortly after the Black Hat conference back in 2009 in Washington DC voiced concerns over facial biometric use for authentication. The salient features of the report demonstrated how early versions of face authentication integrated into Lenovo, Asus and Toshiba machines could be hacked and did not fulfil the necessary requirements. Although facial analysis methods are becoming more sophisticated and reliable for security there are still concerns over the handling and management of biometric data due to its permanent nature.

A group of biometric computation experts demonstrated at the 2015 Black Hat conference how retinal analysis could be re-engineered and potentially hacked to falsify authentication [retinal hack black hat]. This form of re-engineering firstly required gaining access to the stored biometrical data but illustrated that even the most complex of biometric analysis could be exploited. Amongst the concerns of biometrical data management their also worries that the integration of facial recognition systems increase surveillance coverage and reducing the public privacy.

PCA is one of the most successful face detection algorithms to date. PCA is the method of reducing dimensionality of a data space also referred to as observational variables) and extracting data from the feature space (Independent Variables) which are needed to represent the data economically. The feature space in this case is the Eigen Space in which derived principal components are projected to deduce Eigen values.

* PCA is one of the most popular methods for face image analysis. It essentially optimizes the data needed to represent a defined set of data.

APPLICATIONS OF FACE RECOGNITION

Infrared technologies are currently being used in carparks to simulate a daytime like environment to extract car number plates in poor lighting conditions. This allows image recognition methods to extract the characters of the number plate. Infrared scanners used in this capacity are expensive to install and maintain but would provide more mobility of the system and more potential recognition environments.

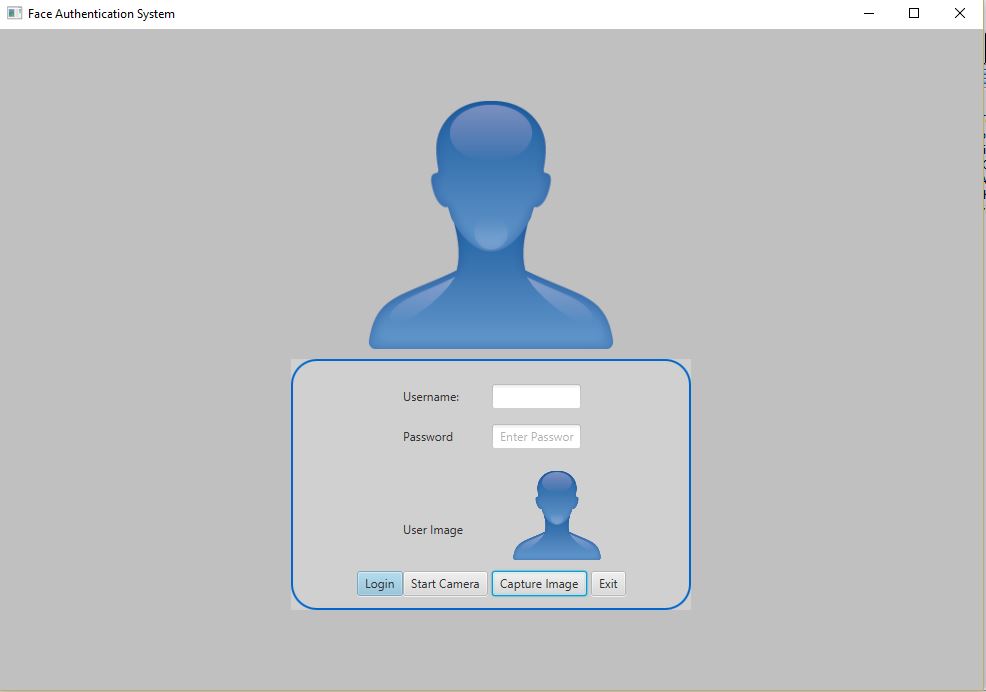
As the environments lighting conditions are a CCP (Critical Control Point) of the systems functionality its placement must be considered careful to cater for its needs.

* Face recognition has been inherited by many high street retailers to recognise VIP guests/customers so that a platinum service can be delivered.
* Most commonly face recognition is being used in surveillance and Security. Crowd control. Lost device services.
* Extended face recodnition research investigates the use of thermal cameras to collect image data. This allows a better representation of a 3D face model as it does not detect facial items such as glasses and records only heat signature data. <https://en.wikipedia.org/wiki/Facial_recognition_system>

## Implementation

A basic GUI framework was initially implemented and later enhanced to satisfy the needs of interaction of the system. The JavaFX library used to develop the GUI allowed the creation of an MVC based model isolating the API and GUI allowing modular reuse of the system. As seen in figure.1 illustration of the login screen the login interface hosts camera feed functionality

**Figure.n** Screenshot of login screen graphical user interface



The login screen provides an area where the user can test user authentication. On first time login of the system access is granted using an administrative username and password so that an initial face image data set can be loaded and trained. After the initial trained data has been produced a known user of the data set can then login using the facial recognition functionality.

The screen hosts a video capture stream activated by the ‘Start/Stop Camera’ button. The video capture objects implemented throughout the system are sourced from the ‘opencv.videoio’ library. Video stream data is handled in memory using OpenCV’s Matrices object format ‘Mat’. To comply with Java.awt image display standards ‘Mat’ objects are converted to a Java.awt Image using OpenCV’s image codecs in order to display the images in the GUI.

|  |
| --- |
| **protected** Image Mat2Image(Mat frame) {  // temporary buffer  MatOfByte buffer = **new** MatOfByte();  // encode image frame into PNG format  Imgcodecs.*imencode*(".PNG", frame, buffer);  // build image from encoded buffered data  **return** **new** Image(**new** ByteArrayInputStream(buffer.toArray()));  } |

Figure. N OpenCV Matrix to Java awt conversion function

Each grabbed frame from the video stream is processed using the FaceDetector.class developed for this application. The face detector class loads into memory a trained cascade classifier ‘haarcascade\_FrontalFace\_alt.xml’ supplied by the OpenCV library. The grabbed frame is converted to greyscale to reduce colour channels and effectively data to process. Greyscale conversion functionality is incorporated in OpenCV Imgproc module. During system runtime there are multiple areas of the application where image type conversion is computed. The class AppTools.class has been developed to manage these conversions and has been referenced in the Appendices of this report.

The face detection algorithm is performed on each grabbed frame of the video sequence to determine a human face through haar-like feature cascade classification. This involves loading the classifier into the Cascade Classifier object supported by the Open CV library. The frame is then prepares for detection by converting the frame to greyscale and calculating the minimum face size. The following code seen in figure.n illustrates this preparation.

|  |
| --- |
| //create cascade classifier  CascadeClassifier faceDetector = **new** CascadeClassifier();  //absolute path to classifiers  String classifierPath = "C:\\Users\\user\\workspace\\AFaceAuthenticationSystem\\src\\application\\Resources\\HaarCascades\\haarcascade\_frontalface\_alt.xml";  //load classifiers  faceDetector.load(classifierPath);    //Create Mat canvas to store detections  MatOfRect faceDetections = **new** MatOfRect();  Mat greyScaleImg = **new** Mat();    // convert the frame in gray scale  Imgproc.*cvtColor*(imageMat, greyScaleImg, Imgproc.***COLOR\_BGR2GRAY***);  // compute minimum face size (20% of the frame height)  **if** (absoluteFaceSize == 0)  {  **int** height = greyScaleImg.rows();  **if** (Math.round(height \* 0.2f) > 0)  {  absoluteFaceSize = Math.round(height \* 0.2f);  }  } |

**Figure.n classification preparation**

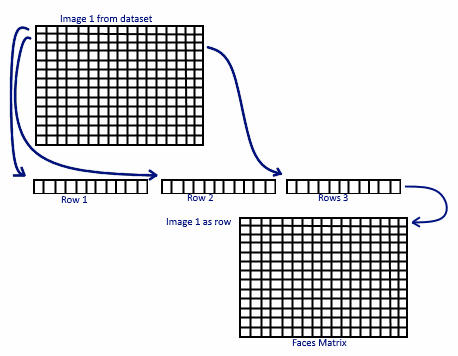
Detection is then performed on the prepared frame. This process iterates through the frames pixel data to distinguished known features of the classifier.

|  |
| --- |
| // detect faces  faceDetector.detectMultiScale(greyScaleImg, faceDetections, 1.1, 2, 0 | Objdetect.***CASCADE\_SCALE\_IMAGE***, **new** Size(  absoluteFaceSize, absoluteFaceSize), **new** Size());  //create array of face detections computed  Rect[] faceArray = faceDetections.toArray();  //iterate through the image  **for** (**int** i=0; i<faceArray.length; i++)  {  //find rectangle contours of faces  Imgproc.*rectangle*(imageMat, faceArray[i].tl(), faceArray[i].br(), **new** Scalar(0, 250, 0, 255), 3);  //crop image of face  Mat crop = imageMat.submat(faceArray[i]); |

**Figure. 2** Face detection code snippet

Limitations of the ported OpenCV library JavaCV where discovered here. The native C++ OpencV library supports a FaceRecognizer class that features PCA (Principal Component Analysis ) comutation functionality. As this module has not currently been ported to the JavaCV library a custom PCA class was implemented in Java with the aid of third party apache commons Math [] and Jama Matrices libraries [].

PCA concatenates each row of pixel data from an image of the dataset into a one dimensional array. Once each image is represented as a row a Matrix is built up of each image as a row. The logic to method can be seen in figure.n below.



**Figure.n** Face matrix preparation

Once the faces matrix has been constructed pixel values of the data are normalised between zero and one. During the greyscale averaging computation the RGB values are combined to create large decimal numbers.

|  |
| --- |
| **private** **double**[][] normalizeImageData(**int**[][] image) {  **double**[][] returnData = **new** **double**[image.length][image[0].length];  // Normalise data between 0 and 1  **for** (**int** faceInd = 0; faceInd < image.length; ++faceInd) {  **double** min = (**double**) getMinValue(image[faceInd]);  **double** max = (**double**) getMaxValue(image[faceInd]);  **for** (**int** j = 0; j < image[faceInd].length; ++j)  returnData[faceInd][j] = ((((**double**) image[faceInd][j])) - min) / (max - min);  }  **return** returnData;  } |

The strength of a stand-alone face recognition system used as a main authentication tool doesn’t appear to be the safest method of computer recognition as the face detection can be spoofed.

It is common practice to select a sub set of the computed Eigen values. The highest variations is distance of the Eigen values are removed as they represent high frequencies of data.

## Testing: Verification and Validation

Inanimate images/photos can be held in front of the camera and a face is successfully detected. Open CV’s trained classifiers are robust enough to detect images with very little features. This poses a security issue as user recognition can be spoofed without the necessary precautions being implemented to counter such forgery.

Show recognition results and thresholds. Discuss false negative matches, why and how can be countered.

## Discussion

When opencv was ported too javaCV the contrib module was not part of the implementation. This module contained useful functionlkity for FaceRecognize class. A JNI (Java Native Library) could be used to access opencv functionality in c++ although this would be more resource intensive

## Conclusion

This system can be fooled by holding an image which is then successfully detected. So if an unknown user is attempting to login they would simply have to find a reasonable high resolution image of the user they wish to impersonate and present this to the face detection process whilst producing a login image.

## Project Commentary

## Social, Legal, Health & Safety and Ethical Issues

There are many ethical arguments to the analysis and storage of biometrical data. Biometrics uniquesly identify a person with biometric data and can’t be changed easily. If your authentication password or forms of identity such as bank card become compromised they can easily be changed although human like features such as the face are much harder to modify.

The ethical and social use of face recognition has been questioned much already as the public feel that their civil liberties and privacy are being invaded by the continuous surveillance and gathering of public data. With the ever growing technical strive to innovate methods of human computer interaction biometric analysis.

## Reflection

Naturally the originally determined success criteria in the PID document are very general and needed further definition.

To ensure a reliable and secure authentication system scenarios such as user image spoofing where the unknown user attempting to log in holds up an image of a known user must be consider.

The project deliverable defined in the PID where unrealistic. Which ones why, what took up time.

JNI (Java Native Interface) can be used to utlise C++ code. The native C++ openCV library offers a very robust FaceRecognizer class that can determine the age and gender of a user. It also features live detection functionality that can differentiatre between a live and inanimate subject. This functionality cleverly exploits the movement of eyes and skin tone changes to conclude whether the user is a real person [ref to live detctor]

## References

[1] University of Reading, PID reference. Available at:

[] Network World, US intelligence wants to radically enhance facial recognition software [online]. Available at : <http://www.networkworld.com/article/2225788/applications/us-intelligence-wants-to-radically-advance-facial-recognition-software.html>

[viola ref]

[Haar feature classification]

[Java CV ref] <http://docs.opencv.org/>

[DDFD]

[conigtec] <http://www.cognitec.com/research.html>

[JavaFX]

[MVC] Model View Controller

[bioID]

[Windows Hello]

[Black Hat]

[Def Con]

[retinal hack black hat]

[Apache commons library]

[Jama Library]

## Appendices