A Photometric Stereo Approach to Face Recognition

Roger Woodman
[www.razorrobotics.com/roger-woodman]

A dissertation submitted in partial fulfilment of the requirements of the University of the West of England, Bristol for the Degree of Master of Science

Faculty of Computing, Engineering and Mathematical Sciences

November 2007

Abstract

Face recognition has received much interest in the last decade, as the need for reliable personal identification security has become ever more critical. At present for face recognition to be a viable personal identification method an accurate low cost solution is required. Many two-dimensional (2D) face recognition systems have been implemented, which demonstrate the potential of using face recognition, although at present these systems are often unreliable. To produce more reliable face recognition systems, a recent trend has been to use three-dimensional (3D) data, which as research has shown is more accurate and robust than traditional 2D techniques. 3D face recognition systems, however, are often expensive relative to 2D alternatives, as precise capture equipment is required. This research project has identified photometric stereo as a low cost 3D capture system, which has been little researched in the area of face recognition. An investigation is presented, which evaluates the capabilities of photometric stereo for use in the area of face recognition by means of a number of experiments. These experiments are conducted using a photometric stereo system, designed and implemented for this research project.

Acknowledgments

To all those that encouraged, supported and guided me, I am eternally grateful.

Contents

List of Tables

List of Figures

1. Introduction

During recent times traditional personal identification methods have come under much scrutiny. This has mainly been due to increase fraud and attention surrounding terrorist activities. To meet the modern day needs of personal identification, a number of more robust methods have been researched. These methods, which aim to uniquely identify individuals using information about a person's physical make up, are known as biometrics. An investigation into fourteen such biometric techniques has been conducted by Jain et al. (2004). Their research identifies face recognition as the only method with the potential of providing a non-intrusive system that could be publicly accepted in today's society.

At present, two main factors restrain the widespread use of face recognition technology; these factors are system reliability and cost. The issue of system reliability has been addressed in many research projects, with recent advances in 3D imaging technology, reported by Zhao and Chellappa (2006), demonstrating increased levels of recognition accuracy compared to traditional 2D systems. The use of 3D data to produce 3D face models has been shown as an effective approach for handling a number of key difficulties in face recognition, which can degrade system performance. These issues are associated with illumination variance, head pose orientation and changes in facial expression. However, although 3D imaging systems have been shown to outperform equivalent 2D systems, more accurate equipment is often required at a greater expense.

The aim of this research project is to investigate a low-cost face recognition system that uses 3D data to perform face matching. The chosen 3D capture technique is photometric stereo, which has been proven to produce accurate 3D models of different objects, while remaining computationally efficient (Vogiatzis et al., 2006). Although photometric stereo has been shown as a reliable method of modelling inanimate objects, little research has been conducted in the area of face recognition. Recent research by Zhou et al. (2007), has presented evidence that photometric stereo can accurately capture

images of a face in order to construct a 3D model. As little research has been performed in the area of photometric stereo for face recognition, it will be the further aim of this research project to investigate the capabilities of photometric stereo and how captured data can be applied to face matching.

To fully investigate photometric stereo for low-cost face recognition, it is necessary to implement a working face recognition system from which experiments can be conducted and results analysed. Therefore, for this research project, a photometric stereo capture system and face matching software will be designed and constructed. To test the system and produce results of face matching, an image database will be created consisting of a number of volunteer faces. To achieve the aims of this research project a number of objectives have been created which are used to organise the investigation process. These objectives and the scope of the overall project are presented in the following chapter.

2. Objectives

It has been argued that 3D face recognition systems have the potential to be more robust and reliable than current 2D systems. This argument is supported by Zhao and Chellappa (2006), whose research has shown that 3D face recognition systems are able to handle much larger variations in subject pose, illumination and facial occlusion. The work by Tangelder (2005) provides slightly different evidence, where 3D methods are shown to enhance 2D face recognition capabilities. Further results from Bronstein et al. (2004a) have shown that 2D face recognition methods can be greatly improved with accurate feature location, which can be achieved more reliably using 3D models.

For face recognition to be a viable personal identification technology, both a reliable and low cost solution is required. As described by Akarun et al. (2005) and Tangelder (2005), currently face recognition technology is not accurate enough to be relied upon alone in the majority of security applications. In addition, the cost of 3D face recognition systems, such as laser scanning and coded light can be expensive, although the evidence suggests that these technologies are the most accurate (Li and Jain, 2004). An interesting 3D modelling technique, which has been little researched in the area of face recognition, is photometric stereo. This method is a low cost alternative to many 3D modelling techniques, which has been shown to produce accurate 3D models of a number of static objects. The work by both Miyazaki et al. (2007) and Vogiatzis et al. (2006) have provided results, which demonstrate the accuracy of the 3D models compared with the actual objects being modelled.

This research project aims to investigate photometric stereo as a low cost face recognition solution. In order to achieve this, an evaluation must be made of the reliability and robustness of the photometric stereo method and face recognition evidence provided from a working low cost photometric stereo system. Furthermore, evidence should be provided detailing the impact that

results from this 3D system could have on traditional 2D recognition techniques.

Based on the aims and evidence presented thus far, the objectives of this research project are to:

- Investigate the latest techniques used in both face recognition and photometric stereo object modelling.
- Implement apparatus capable of acquiring accurate photometric stereo images.
- Identify processes and procedures for implementing an automated photometric stereo face recognition system, using practical experiments to support investigation.
- Evaluate the photometric stereo method for use in face recognition, using practical experiments to support conclusions.

2.1 Scope of Research

Face recognition, as a research area, covers many different disciplines, such as a computer vision, psychology and biology. The research in these areas, although relevant, cannot be covered entirely in this research project. Therefore, it is essential at this juncture to define the project scope.

A 3D face recognition system will be designed, implemented and tested. The implemented system will aim to address a number of face recognition issues, including illumination, pose and distance from camera. The issues of facial expressions, however, will not be addressed in the system, as this covers a huge field of psychology, biology, and 3D modelling. Although this is beyond the scope of the actual system, a full explanation of expression handling techniques will be given in the literature review that follows.

Results will be presented in order to evaluate the face recognition system created for this research project. However, there is no scope at this time for comparing the developed system directly with results of other similar systems, although the relative merit of this is fully understood.

3. Literature Review

Personal identification has seen a number of advances in the last two decades. The term biometrics is used to encompass all methods which identify an individual based on physical attributes that can be both perceived and measured based on actions. Traditional biometric methods such as fingerprints and handwriting signatures have been shown as not reliable when operated on large data sets (Zhang et al., 2004).

In order to meet the demands of today's security conscious society, a number of biometric methods have been proposed that in theory offer greater levels of reliability than traditional methods. A summary of what is required from a modern biometric system is given by Jain et al. (2004) "A practical biometric system should meet the specified recognition accuracy, speed, and resource requirements, be harmless to the users, be accepted by the intended population, and be sufficiently robust to various fraudulent methods and attacks to the system". A number of biometric methods which meet these criteria have been investigated by Akarun et al. (2005). Their research has identified the following as viable personal identification methods:

- Iris
- Retina
- Hand
- 3D face

Of these biometric methods, 3D face has the potential to be the most cost effective and require the least cooperation from the individual being identified. "Although 2D and 3D face recognition are not as accurate as iris scans, their ease of use and lower costs make them a preferable choice for some scenarios." (Akarun et al., 2005). The study by Zhao and Chellappa (2006) has reviewed the current research in biometric identification and the possibility of implementing each in real world situations. Their work has described the benefits of face recognition as being socially accepted, relatively

low cost and has the potential to be integrated into current CCTV security networks, which exist in nearly all cities and secure locations.

The following chapter sections provide a review of the most established and up-to-date technologies in face recognition. An assessment of 2D and 3D recognition methods is given by means of compare and contrast. From these recognition methods a low cost solution for producing 3D face models is identified and investigated in more detail. Finally, techniques for evaluating different face recognition systems based on standardised tests are presented.

3.1 Face Recognition in Humans

Much of the generated interest surrounding automated face recognition is based on the ability of humans to effortlessly recognise other human faces, and the desire to recreate it using computer vision techniques (Hallinan et al., 1999). However, although it is apparent that humans in general are excellent at recognising familiar faces, many studies have shown that humans do not fare well when recognising unfamiliar faces from a single image. Research conducted by Sinha et al. (2006) shows that humans are able to recognise familiar faces of celebrities from blurred or partially occluded images. However, evidence is also presented, which shows that the ability to recognise faces is significantly reduced when presented with an unfamiliar face and asked to recall it from a collection of unfamiliar faces.

A number of studies have been conducted which compare human face recognition with the latest automated techniques. A study by Lund (2001) demonstrates that human's process faces holistically. This means that the face is processed without analysing each feature individually. Further studies have shown that the individual measurements of facial features are not relied upon. An investigation by Zhao and Chellappa (2006) uses a single model of a face in order to align and resize the features of a number of images of different faces to a standard configuration. Their results show that there is little difference in the ability to recognise a face when altered in this way. Their investigation concludes that a human's ability to recognise a face is heavily

reliant on the face texture and shape of facial features and the shape of the face as a whole.

3.2 Face Recognition Using Computer Vision

Computer based face recognition is the process of recognising individual faces from a series of digital images, using computer driven algorithms to aid identification. Whether a face recognition system is semi-automated or fully automated, a similar process flow is used to identify a subject's face from single or multiple images. This general process flow, as described by Zhao et al. (2003), involves; face detection, face quantifying and face recognition. These processes have been implemented in a variety of diverse ways, from using only the 'raw' 2D image data to 3D reconstructions of the entire face. However, these diverse methods all share a common aim, to quantify each individual face in order to prove its uniqueness (Hallinan et al., 1999). "To remember a face as an individual, and to distinguish it from other known and unknown faces, we must encode information that makes the face unique." (Zhao and Chellappa, 2006).

For personal identification, face recognition can be used to authenticate an individual using one of three modes. These modes, as described by Lu (2003) are verification, identification and the watch list task. Verification is concerned with matching a query face with the enrolled image of a claimed individual's identity using a one-to-one matching process. Identification uses a one-to-many scheme and involves matching a query face with an image in a face database to determine the individual's identity. The watch list authentication task has been introduced recently and involves an open test where the query image may or may not represent an individual contained in the database. The query image is compared with all enrolled images of the database with a similarity score computed for each. If the similarity score is equal to or higher than a predefined alarm value then a match is made. This authentication task is designed for security situations, such as locating missing persons or for identifying prohibited individuals at border controls.

Many terms are often associated with how images are processed in face recognition. The term used to define the database of faces, used for recognition matching, is often called the enrolment database. This database is required in order to match a current face image, known as the query image, with a number of enrolled faces in the database (Li and Jain, 2004). The data contained in the enrolment database can be made up of 2D images and/or face information extracted from 3D models. The difference between a 2D image and a 3D model is that the 2D image generally provides no height information, thus the image can only be manipulated on the x and y axes. 3D models are made up of data defining the structure of an object; therefore the model can be rotated on the z axis, which allows for geometric data to be calculated, providing information about the actual distance on the face and the extent of protrusions for facial features. (Bronstein et al., 2004a)

The remainder of this chapter will explore both automated and semi-automated face recognition approaches using computers. A review of face recognition systems by Zhao et al. (2003) shows that the majority of face recognition research use only greyscale images, due to the difficulties associated with colour images with respect to illumination conditions. As stated by Li and Jain (2004), the use of colour information is typically used only as a pre-processing step in order to separate a human face from the image background. Therefore, greyscale images will be assumed from here on unless otherwise stated.

3.3 Image Acquisition

Image acquisition in face recognition is the method of capturing a number of digital images for use in the recognition process. The techniques and equipment used to acquire images vary considerably depending on which face recognition approach is adopted. In general 2D techniques require much less equipment than 3D alternatives. The review of 2D systems by Zhao et al. (2003) states that for 2D systems all that is generally required is a single camera and a well-lit location. An exception, which has received little attention, is the use of thermal imaging cameras (Wolff et al., 2001). These

cameras capture 2D images of the face surface temperature, with each pixel of the image representing a specific temperature.

It is necessary at this juncture to discuss the difference between a 3D model of the head and a 3D model of only the face. A 3D model of the head assumes height information for the entire head, thus the head model can be rotated 360°. A 3D model of the face assumes height information for only the facial features and the surrounding face surface. This is often known as a 2.5 Dimensional (2.5D) surface reconstruction (Vogiatzis et al., 2006). As face recognition is clearly only interested in the face and to avoid confusion, the term 3D will be used to address all models of the head whether present in 2.5D or 3D. The remainder of this chapter will discuss image acquisition for 3D systems.

Stereo imaging is a capture method that uses two cameras in order to simultaneously acquire two images of an object from slightly different viewpoints. Triangulation is applied to a number of points identified in both images in order to calculate the geometry of the presented object. Correspondence is a technique used to locate precise points of an object that can be readily identified in both images. An imaging process, which also relies on correspondence, uses orthogonal profile images of the face (Yin and Basu, 1997). The process can use either two cameras to capture the object in a single shot or a single camera held stationary and the object rotated between shots.

Photometric stereo is a modelling technique that use multiple images captured under different illumination directions to recover the surface gradient (Agrawal et al., 2005). The surface gradient defines the surface orientation for each point of the face, known as the surface normals. These surface normals can be integrated using a surface reconstruction algorithm to produce a model of the face surface. The technique is traditionally used for modelling static objects such as museum artifacts, however, recently evidence has been presented by Zhou et al. (2007), which demonstrates that acceptable face models can be produced using this technique. Photometric stereo will be covered in more depth in section 3.5.

The final type of 3D imaging system is that which projects a light onto an object in order to measure how the object interacts with the light compared to a flat surface. Two such techniques are laser scanners and coded light. Laser-based techniques use a laser line to scan the entire surface; triangulation is then applied to each laser line to calculate the surface geometry. The coded light technique, projects a specially coded light pattern onto an object. Images are captured of the object and the light pattern, which are then processed to reveal depth information based on distortion of the project patterns.

In order to compare the 3D modelling techniques, which have been identified, for use in the area of face recognition, a number of key considerations need to be made. Zhao et al. (2003) defines a number of aspects which can be used to compare face recognition systems, these are:

- Quality of captured data
- Capture and recognition time
- Equipment cost
- Requirements of the user
- Intrusiveness

Of the capture methods identified here, laser scanning systems provide the most accurate results, and are often used to test the precision of results from systems using different capture methods. However, laser sensors are relatively expensive and much time is required in order to completely scan an object. "The acquisition of a single 3D head scan can take more than 30 seconds, a restricting factor for the deployment of laser-based systems." (Akarun et al., 2005). Due to the time taken and visible laser projected onto the face, much cooperation is required from the user. The use of coded light is a relatively fast process. However, as with laser scanning a visible light pattern is projected onto the users face. Two methods, which require no projected light, are stereo and orthogonal imaging. These techniques use standard cameras, with stereo imaging requiring two cameras and orthogonal needing only one. Stereo imaging, as described by Li and Jain (2004), can

produce accurate 3D models using little equipment. The main issue with these methods is the use of correspondence, which can be computationally expensive and produce errors in the results. 'Often failure to identify a stereo correspondence point can result in a 3D surface, which contains 'holes' due to absent surface information' (Zhao and Chellappa, 2006). A relatively fast 3D capture method, which does not require correspondence, is photometric stereo. This method requires a single camera and three lights. Further details of photometric stereo as a face recognition method will be presented in section 3.5.

3.4 Prominent Face Recognition Methods

All computer aided face recognition systems use either 2D or 3D data in one form or another. The form of this data can differ between 2D and 3D systems, with 2D methods usually using colour or intensity values and 3D methods using geometric based values, range data or surface gradient data. According to Zhao et al. (2003) face recognition methods, using either 2D, 3D or both techniques together, can be organised into one of three groups:

- Feature-based methods
- Appearance-based methods
- Mixed methods

The first group, feature–based methods, segment the face into a number of components which can be analysed individually. This usually involves the identification of local features such as the eyes, nose and mouth. (Axnick and Ng, 2005). These local features can be analysed in a number of ways, using simple measurement techniques or more sophisticated template matching. Appearance-based methods use the whole face area as the input to the face recognition system. These methods typically operate on 2D images, using the raw image data to produce comparisons of other face images, from which a match can be made. The final mixed methods group combines both feature- and appearance-based methods to produce different results. It is argued that mixed methods should produce more reliable results, albeit with a greater

computational cost. “Combination of feature- and appearance-based methods may obtain the best results.” Tangelder (2005).

The following sections describe a number of face recognition methods in more detail. Each method has relative merits and demerits and it is these properties that are used when choosing a method for any given face recognition task. ”Appropriate schemes should be chosen based on the specific requirements of a given task.” (Zhao et al., 2003)

3.4.1 Feature-based Methods

Early feature-based methods used a set of geometric feature points in order to measure and compare distances and angles of the face. Research by Starovoitov et al. (2002) uses 28 face points, manually selected on a number of 2D images, in order to describe the face. Their results show that this method is accurate in recognising faces, which are all of the same pose orientation. However, with varying pose the ability to recognise a face is significantly reduced. This process of locating geometric feature points has recently been applied to 3D face models. The evidence presented by Zhao et al. (2003), demonstrates that 3D feature location is more robust than 2D methods, allowing for large variation in subject pose and changes in facial expression.

Elastic bunch graph matching (EBGM) is a recently introduced technique, where a sparse grid is overlaid on the face and a number of grid points are adjusted to pre-defined feature locations. These grid points are used to calculate local feature vectors, which can then be used in face comparisons using a graph matching algorithm (Aguerrebere et al., 2007). A critical part of this method is accurate location of grid points, which can be difficult to achieve reliably. “The difficulty with this method is the requirement of accurate landmark localization.” (Bolme et al., 2003).

3.4.2 Appearance-Based Methods

The most prominent appearance-based method, which is also one of the earliest, is the Eigenface method introduced by Turk and Pentland. The Eigenface algorithm, as described by Turk and Pentland (1991), uses principal component analysis (PCA) to reduce the number of dimensions that make up the face to its principal components. These principal components are defined as the parts of the face image that exhibit the largest variations when compared to the training set. PCA is used to avoid direct correlation with all components of the face image, which as well as being computationally inefficient, also can result in unreliable data being processed. In contrast to feature-based methods, which rely on the location of facial points, the Eigenface method processes the face using a holistic approach, thus negating the requirement of any feature locating. For recognition the Eigenface method relies on the assumption that "the difference between two instances of the same face will be less than that between two different faces." Hallinan et al. (1999)

Evidence of the Eigenface method, has shown it to produce accurate, reliable recognition results when presented with frontal face images under constant illumination conditions (Hallinan et al., 1999). However, the method is sensitive to variations in pose and illumination, which can have a significant effect on the reliability to match faces. To address these issues a recent method known as Fisherface has been conceived (Shakhnarovich and Moghaddam, 2004). The Fisherface method adds another dimension to the Eigenface technique, by using a larger training set containing multiple varying images of each subject face. The advantage of this is that more data is available to describe illumination conditions and pose orientation, which as a result makes the method more robust as more training images are presented. The method is inherently more complex than the Eigenface method and thus requires more computation time to both train the system and generate a face match.

3.4.3 Mixed Methods

Mixed methods combine aspects of both feature- and appearance based methods with the aim to produce more accurate and robust recognition results. The most widely used mixed method is Eigenfeatures (Torres et al., 1999). Eigenfeatures is a direct extension to Eigenface that uses feature locations in an aim to improve face matching. The process involves locating and separating individual face features, these face features are then processed individually using the Eigenface technique. Evidence is presented by (Torres et al., 1999), that this method provides more accurate recognition than the appearance-based method, Eigenface.

The following section provides a detailed explanation of photometric stereo and presents a review of the latest research conducted and technical developments made. The photometric stereo method can be used with mixed method face recognition, as both 2D images and 3D reconstructions can be produced.

3.5 Photometric Stereo

Photometric stereo is a technique used to recover the shape of an object from a number of images taken under different lighting conditions. The shape of the recovered object is defined by a gradient map, which is made up of an array of surface normals (Zhao and Chellappa, 2006). Each surface normal is a vector, which dictates the orientation for each of the facets of a surface. The diagram of Figure 3.1 illustrates a single surface normal. It can be seen from this diagram that the normal orientation is defined by an x, y and z component.

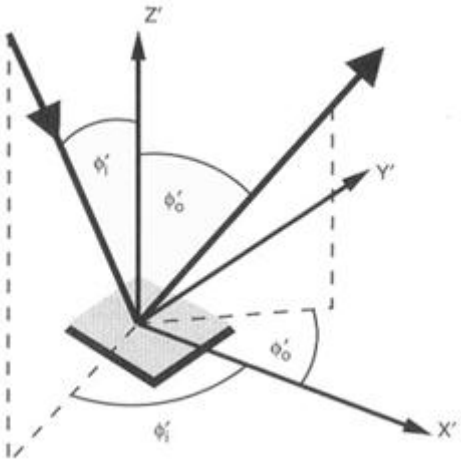


Figure 3.1 – Surface normal diagram (Zhao and Chellappa, 2006)

As photometric stereo uses computer images, each surface normal is attributed to the smallest unit of an image, which is known as a pixel. Therefore, the number of surface normals depends on the number of pixels contained in the image. To recover a surface normal at least three images are needed, each with a different light source direction (Kee et al., 2000). The requirement of three images under different light sources, adds minimal constraints on any photometric stereo system, these constraints are the use of three lights and the acquisition of three images. It has been shown by Vogiatzis et al. (2006) that it is possible to use multiple camera views to capture a larger surface of an object, however, the simplest and most accurate method is to use a constant camera view point. In addition to the gradient map, photometric stereo can also be used to produce an albedo image (Kee et al, 2000). This image is illumination invariant and describes the surface texture of the object without shape.

Although it has been shown that the minimum requirements of a photometric stereo system are three images taken under three different light sources, this in practice has associated with it two main issues, these issues are cast shadows and ambient light. The issue of cast shadows occur for objects where not all light sources reach the entire object surface, often caused by protrusions or overhangs of the object surface. This can result in missing data as surface normal calculations require at least three different intensity values to calculate a surface normal. To prevent missing data, a number of methods

have been developed which capture more data of the object surface. The research by Raskar et al. (2004) and Chandraker et al. (2007) use four lights to capture four differently illuminated images of an object. The algorithm implemented by Rasker et al. (2004) uses all four images to calculate every surface normal of the object. An Alternative method, described by Chandraker et al. (2007) presents an algorithm where cast shadows are removed as a pre-processing step. Their approach is to select the three pixels of highest intensity from the four images and use these to calculate each surface normal. The second issue of using a minimal photometric stereo system is due to ambient light, which can cause incorrect calculations of the surface normals. This issue of ambient light, as highlighted by Bronstein et al. (2004a), can be 'compensated for by subtraction of the darkness image'. The darkness image, which is referred to, is simply an image captured under ambient illumination alone.

The process of photometric stereo for 3D modelling is a well-documented technique. Much research has been shown that inanimate objects such as museum artifacts can be accurately modelled (Miyazaki et al., 2007) and (Vogiatzis et al., 2006). However, the research into photometric stereo as a face recognition technique is relatively under researched. Recent research into using photometric stereo for face recognition has been conducted by Zhou et al. (2007) and Bronstein et al. (2004a). Much of their research has shown that photometric stereo has the capability to produce accurate 3D models of human faces, which has the potential to be used in a face matching system.

A recent paper published by Miyazaki et al. (2007), describes a photometric stereo system which is able to model museum artifacts thought glass screens. An algorithm is presented, which removes the reflective effects of a glass surface, by selecting three of the five captured images, which contain no saturation due to direct reflection of light from the surface to the camera lens. This paper highlights the potential of using photometric stereo in face recognition. Although no mention of face recognition is made in the research by Miyazaki et al. (2007), it could be argued that this algorithm demonstrates

a method that could be extended to the area of face recognition, in order to remove the effects of subjects wearing glasses or to capture face images through glass screens.

3.5.1 Surface types

Photometric stereo can be used to model the majority of surfaces, however depending on the surface type different surface normal calculations must be used. Different surfaces can be divided into one of two types, Lambertian (after Johann Heinrich Lambert) and specular (Li and Jain, 2004). The diagram of Figure 3.2 illustrates the different properties of the two surfaces, when illuminated from a point source.

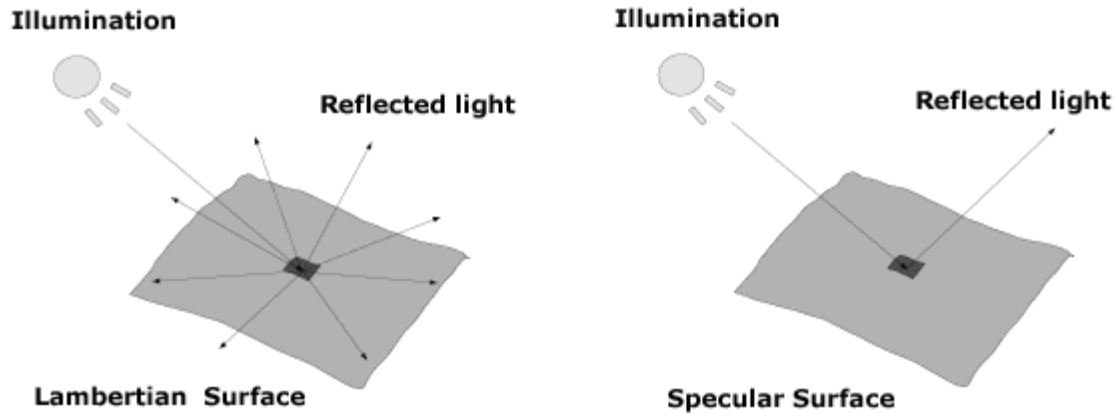


Figure 3.2 – Surface types: (a) Lambertian surface (b) Specular surface

As illustrated in the diagram, Lambertian surfaces (Figure 3.2 (a)) reflect the incident light in all directions and appear equally bright when observed from any view point (Kee et al, 2000). Lambertian surfaces are found in materials such as cotton cloth, brick and matt paper and paints. Specular surfaces (Figure 3.2 (b)) reflect light at the same angle as the incident light, which results in different brightness being observed dependent on the view point. Specular surfaces are present on materials which reflect light, such as mirrors and polished metals.

For face recognition Lambertian surfaces are assumed, as the face skin is mostly a Lambertian surface (Kee et al, 2000). However, as Hallinan et al. (1999) states, the face as a whole can contain specular surfaces in the hair and eyes, which can cause specularities in the face data. These specularities are produced, when a specular surface is treated as Lambertian, producing an error value. The error values are often removed by averaging the surrounding values to reconstruct the surface (Zhao and Chellappa, 2006).

3.5.2 Surface Reconstruction

To produce 3D models from photometric stereo data it is necessary to integrate the gradient map to produce a height map. This height map, also known as a range map, defines the relative heights for each part of the surface, with the highest values representing those surfaces closest to the camera at time of capture (Lenglet, 2003)

Two integration techniques exist that can be used to reconstruct the gradient map, to produce a 3D model of the actual objects surface, these are local and global techniques. These techniques as described by Schlüns and Klette (1997) vary considerably in their performance and the reconstruction results. Local integration techniques start at a location within the gradient data, often the centre, and perform integration on neighbouring values, moving around the gradient map until a complete integration is achieved. Global integration techniques integrate the gradient map as a whole, using multiple scans in order to reconstruct the surface. Both local and global techniques have relative merits and demerits, which determine their use. Local techniques are relatively easy to implement and computationally fast, although dependent on data accuracy. Global techniques are computationally more expensive, however, are significantly more robust to noisy data (Schlüns and Klette, 1997). For photometric stereo, global integration is usually always used due to the noise present in computer images captured using this technique (Kovesi, 2005).

3.6 Recognition Issues

A number of issues exist in face recognition, which can impair the system's ability to recognise faces. These issues can be grouped into two classes, internal factors and external factors (Bronstein et al., 2004a). Internal factors are associated with issues concerning the physical aspect of the face, such as facial expression, eyes open or closed and ageing. External factors are those that affect face recognition independent to the actual face; these are issues such as illumination conditions, head pose relative to the camera, facial occlusion and low quality images. As described by Li and Jain (2004) face recognition systems in general aim to handle these internal and external factors in one of three ways. The first approach is to use a standard capture system, where lighting and image quality is regulated, and to enforce rules, which the subject face must adhere to, such as central pose and a neutral facial expression. The second approach is to use a pre-processing step, which normalises the captured faces to the appropriate size, illumination level, pose and expression, based on an ideal face template. The final approach is to use an invariant representation of the face, which is insensitive to all changes in both internal and external factors. It is important to note that an invariant representation of the face has not yet been achieved and as stated by Bronstein et al. (2004a) "The ultimate goal of face recognition is to find some invariant representation of the face, which would be insensitive to all these changes". The remainder of this section addresses the principal issues that presently exist in face recognition.

Illumination variations due to ambient light or unknown lighting directions can have the largest effect on the ability to recognise a face. "The changes induced by illumination are often larger than the differences between individuals" (Zhao and Chellappa, 2006). To manage variations in illumination a number of techniques have been developed, one of which has already been described for photometric stereo, where an image is taken under ambient lighting and subtracted from all other images. This technique can be used for any photometric systems. For 2D systems, the two main techniques

used, are histogram equalisation and edge detection (Li and Jain, 2004). Histogram equalisation uses a graph of all image intensity values to determine information about the lighting conditions and offset their effects. Edge detection, also used for low quality images, is a technique used to highlight features within images based on steep transitions of intensity values. For face images this technique can be used to define the edge of the face and the majority of facial features.

The issue of head pose is a major concern for 2D face recognition developers, as slight variations in pose can affect the image results significantly. Methods which aim to handle pose variations typically use 3D models (Akarun et al. 2005). Face recognition systems can either use 3D models exclusively or as a pre-processing step to rotate the head to a central position from which 2D images can be produced. The later method is described by Bronstein et al. (2004a), where in their research a traditional Eigenface method is enhanced by fusing 2D and 3D data to produce a more robust method which is invariant to illumination and pose.

Variations in facial expression can affect the location of facial features and the overall shape of the face. This can affect both feature- and appearance-based face recognition methods. As with head pose variations, 3D models have been used in mathematical modelling to determine the facial expression and how the face has changed as a result (Zhao and Chellappa, 2006). These mathematical models can be adjusted to produce a neutral expression, which can be applied to 3D recognition or used to produce neutral 2D images.

The final prominent issue in face recognition is that of facial occlusion. Facial occlusion techniques are concerned with any articles that cover the face and subsequently affect recognition. These include items such as glasses, hats, jewellery, hair and cosmetics. The majority of techniques used to identify facial occlusions are based on 2D images (Zhao et al., 2003). The traditional method is to use template images of items, which could be present within an image, in order to search for and exclude their locations from the matching process. This process of excluding occluded areas, as described by Li and Jain

(2004), is an often used technique for handling large facial occlusions, although it is shown that for smaller facial occlusions the face surface can sometimes be reconstructed using information from other parts of the face, however, this approach is purely subjective to the type and degree of occlusion. A final approach worth citing is the use of thermal imaging cameras to capture images of the face which highlight different areas of temperature (Wolff et al., 2001). This technique as examined by Zhao and Chellappa (2006) can reliably identify a number of facial occlusions such as glasses, hats and jewellery based on their temperature difference from the face surface. The identified items can subsequently be handled using one of the previously described methods.

3.7 Performance Measures

In order to compare face recognition systems and algorithms it is essential to use standard performance measures. The two most common performance measures are false accept rate (FAR) and false reject rate (FRR) (Li and Jain, 2004). The FAR measure denotes the number of unknown individuals identified as belonging to the enrolment database, whereas the FRR measure denotes the number of individuals, who appear in the database, falsely rejected. The importance of these performance measures is highlighted by Zhao and Chellappa (2006) where it is shown that the threshold used to determine if a face match has been found must be set to compromise between the number of query images falsely rejected and the number falsely accepted. As their research shows, the compromise between the two measures is often application specific, with more secure applications favouring much lower FAR's with the associated effect of higher FRR's.

As of 2000 the face recognition vendor test (FRVT) was established to meet the demands of researchers and companies who wanted to compare the abilities of their system with that of others (Li and Jain, 2004). The FRVT provides the means to compare different system performance using a number of standardised face databases. According to Zhao and Chellappa (2006), the

databases created for the FRVT have become the benchmark for many face recognition research projects. The FRVT has been conducted in 2000, 2002 and again recently in 2006, the report by Phillips et al. (2007) presents the results of the 2006 FRVT. In their report they identify an improvement in the overall performance of the tested systems, which is described as being "due to advancement in algorithm design, sensors, and understanding of the importance of correcting for varying illumination across images." (Phillips et al., 2007).

The bench mark used for the FRVT is a FRR of 0.01 (1 in 100) at a FAR of 0.001 (1 in 1000), with any system achieving these results, receiving an accreditation. As a comparison the ideal system would produce the results of 100% correct detection with zero false acceptances. The FRVT benchmark figure was first achieved in the 2006 test by two different systems, 'the first by Neven Vision using high resolution still images and the second using a 3D imaging system produced by Viisage.' (Phillips et al., 2007).

The remainder of this project aims to build on the research conducted in this chapter. The Photometric stereo method, identified as a low cost 3D modelling technique, will be investigated by means of a system implementation. An evaluation of the system and findings will be presented and conclusions made on the capabilities of photometric stereo for face recognition.

4. Methods of Investigation

The methods used in this research project involve an equal share of background research and practical experimentation. In this way it is believed that a balanced and evident evaluation can be made of photometric stereo as a face recognition technique and of face recognition as a practical method for personal identification. A thorough examination of the available research literature is to be first conducted in order to determine the scope of the project based on the current level and focus of research that has already been achieved in this area.

A number of experiments will be designed and conducted in order to investigate the aims of this project. These experiments will test photometric stereo approaches to face recognition. The experiments will require the design and construction of photometric stereo apparatus and software to capture and process face images. A database of different face images will be used to verify the effectiveness of the face recognition system. The database will be populated with face images of a number of volunteers. The literature research will continue at key stages of the project to further reinforce the context of the investigation and the experiments performed.

A summary of the methods of investigation for this research project are given here:

- A survey of the relevant literature will be conducted throughout the research project, to ensure a well-rounded review of the literature is presented.
- Regular discussions with project supervisor to provide guidance and ensure research aims are feasible.
- A number of experiments will be conducted to evaluate the effectiveness of photometric stereo as an accurate and robust method for face recognition.

- The investigation of the experiments will require image acquisition apparatus and image processing software to be designed and implemented.

- A database of subject images will be created in order to conduct recognition experiments. This will require the cooperation of a number of individuals.

5. Image Acquisition and Representation

Photometric stereo, which is being exclusively used in this project, is a technique employed to capture images of an object, be it a chair, a face, or even an entire landscape. The technique involves using three or more images taken from a stationary camera with light sources projected from different directions. The group of images that are taken of an object with each of the lights is known as the image set. Thus for each face an image set has to be acquired in order to extract the necessary photometric stereo data.

As previously stated, the aims of this research project are to investigate the effectiveness of photometric stereo for use in a face recognition system. The actual photometric procedure itself is not under direct scrutiny and therefore the methods used to capture the image sets will not be dealing with issues such as ambient lighting or subject movement. A detailed description of how to handle these issues is given by Bronstein et al. (2004a). For this research project, the image set will be made up of three images of a human face, taken under three unique lighting conditions. The images will be taken as fast as possible to avoid movement of the face between captures. All of the images will be taken in a dark room, which will avoid any ambient light affecting the images.

5.1 Image Intensity Equations

This section defines the equations used in this research project, to calculate both the surface albedo and surface normals. Both sets of equations are applied to each pixel in all three images. The result of each equation is a single matrix the same dimension as the input image size. As greyscale images are being used, the pixel value between 0 and 255 defines the intensity.

The equation used to calculate the surface albedo can be expressed as

$$albedo(x, y) = \sqrt{I1^2 + I2^2 + I3^2} * 0.58$$

Where x and y are the current pixel position and I1, I2, I3 are the intensity values for each of the three images. The equation has been adapted from a calculation presented by Kee et al. (2000).

The surface normal equations are based on those described by Lenglet (2003). To make the equation more applicable to this research project, the x, y and z components of the surface normal have been separated into three separate equations. The surface normal equations are defined as:

$$normalX(x, y) = (I2 - I3)/(\text{distance} * 0.8660254)$$
$$normalY(x, y) = -(((2 * I1) - I2 - I3)/(3 * \text{distance}))$$
$$normalZ(x, y) = (I1 + I2 + I3)/(3 * \text{height})$$

Where distance is equal to the distance apart for two lamps on the photometric stereo apparatus and the height is the distance from a lamp to the face. These two values are expressed as a ratio of each other, thus any measurement unit can be used.
Before discussing the intended experiments and the design of the photometric apparatus, it is first important to understand what kind of data will be utilised. All the figures presented in the subsequent sections have been created during this project using the tools and techniques discussed in the following chapters.

5.2 Surface Gradient

As explained in chapter 3, the surface gradient describes the surface normals as a whole, calculated using the image set. Each surface normal has x, y, and z components. Using the x and y component the surface gradient can be plotted in 2D, to form an image illustrating how each surface normal is orientated. The image in Figure 5.1 (a) shows an albedo image and (b) surface gradient map of a face.

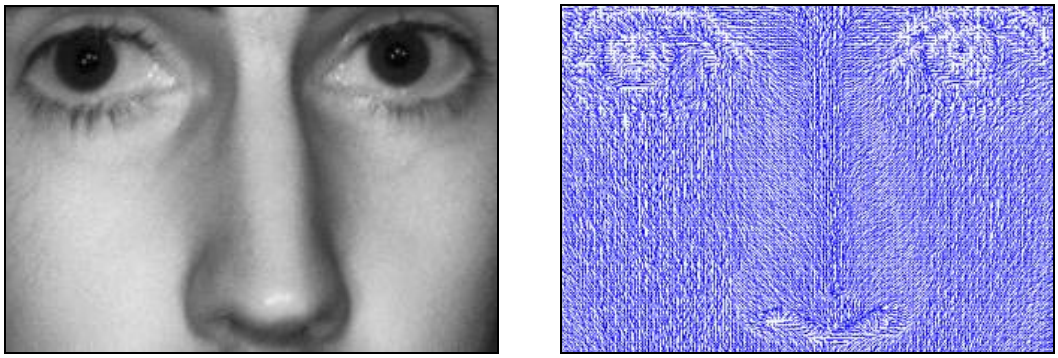


Figure 5.1 – (a) Albedo face image (b) surface gradient map

The surface gradient can be considered as the 'raw' data obtained from the photometric stereo process. This data is often integrated to produce a 3D surface, which can be analysed using the height data.

It was found during the initial testing stages that the quality of the surface gradient data had a direct effect on the ability to extract accurate height information. The measure of quality of the surface gradient is simply how clearly the facial features can be seen. This means that the more visible the facial features appear the more accurate the surface gradient.

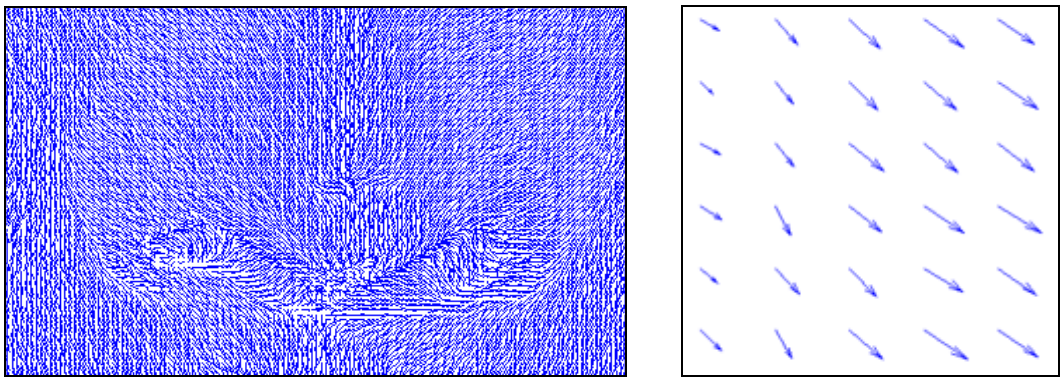


Figure 5.2 – (a) Surface gradient map of nose (b) Surface normal sample

The images shown in Figure 5.2, illustrate the gradient information contained in a section of the lower part of the nose, taken from Figure 5.1. The images show how gradient information can be represented by equally spaced lines directed according to the orientation of each surface normal.

5.3 Surface Reconstruction

As stated in chapter 3, surface reconstruction involves transforming the gradient data into height data. Both local and global techniques have been reviewed and evidence presented that global techniques are more suited for photometric stereo, due to their ability to handle noisy images.

In order to produce a 3D model of the face, surface gradient data has to be integrated to create a smooth reconstruction from one surface gradient to the next. The surface reconstruction method chosen for this research project is the Frankot-Chellappa algorithm. This algorithm is used in many of the research papers, and is described by Kovesi (2005) as producing accurate reconstructions as well as being resilient to noisy images and computationally fast. "This is a quick one-step algorithm that is highly robust to noise. Indeed, it possibly remains one of the most noise tolerant algorithms to date." (Kovesi, 2005).

The Frankot-Chellappa algorithm uses Fourier transforms to integrate each surface normal of the gradient map to produce a reconstructed surface. The Frankot-Chellappa surface integration equation (Schlüns and Klette, 1997) can be expressed as:

$$Z(x,y) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} Z^{(F)}(u,v) \cdot e^{-j \cdot (u \cdot x + v \cdot y)} du\, dv$$

To illustrate the surface reconstruction, the remainder of this chapter presents a series of images produced using the Frankot-Chellappa surface integration algorithm. An albedo image, displayed in Figure 5.3, shows a section of a face containing the eyes and nose, which is used in the subsequent reconstructed images.

Figure 5.3 – Albedo image used for surface reconstruction

Using the Matlab '*surf()*' function, reconstructed surfaces can be simply transformed to produce a 3D model. The image of Figure 5.4 shows the modelled height data of the reconstructed surface created using the Frankot-Chellappa algorithm. Further reconstructed faces, produced during this research project, are provided in appendix A.

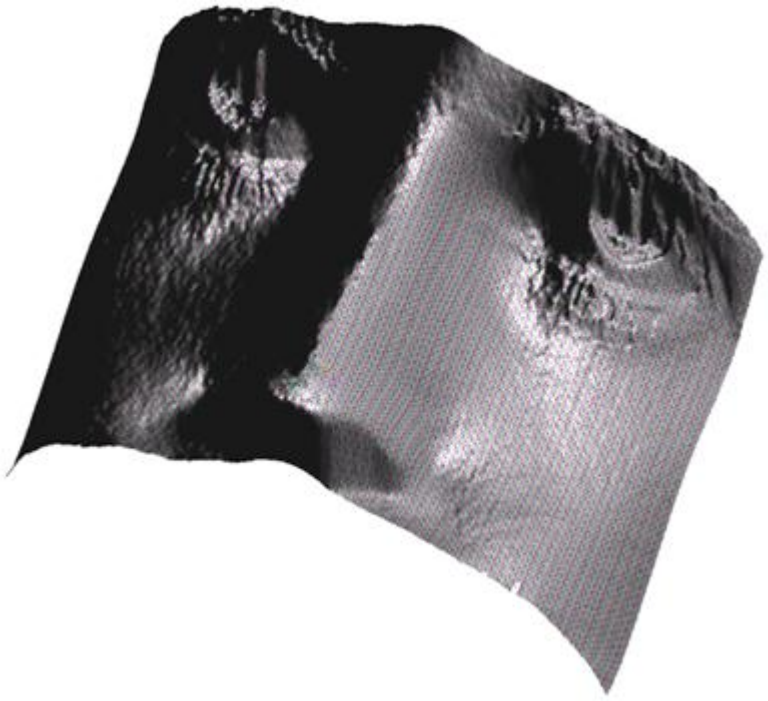


Figure 5.4 – Reconstructed surface: Raw height data

This reconstructed surface is displayed as 'raw' height data, which is equivalent to the texture of the face surface devoid of any actual surface colour (intensity) information. In addition to displaying the 'raw' height data, it is possible, using the Matlab *'surf()'* function, to overlay the albedo. This

produces a 3D face model with the surface texture and colour (intensity). The images presented in

Figure 5.5 show multiple views of a reconstructed face surface overlaid with the albedo face image. As it can be seen from these images, it is possible to produce 'new' 2D images of a face by adjusting the orientation of the 3D model. These 2D images could be exported from Matlab and have the potential to be processed using 2D face recognition techniques. However, this is outside the scope of this research project.

Figure 5.5 – Multiple views of a reconstructed surface: Albedo overlay

It is important to note that the 3D reconstruction images produced in Matlab show height information, thus any details that can be viewed must represent a difference in height. Facial features such as the nose represent large height values, whereas wrinkles, bumps and furrows represent small fluctuations in height data. It has been shown that 3D reconstructed face surfaces overlaid with the albedo have the potential to be used for 2D face recognition. However, as stated by Zhao and Chellappa (2006) it is critical when measuring the surface height that only the reconstructed 'raw' height data is used and not the additional albedo image, as this can distort the surface and conceal errors in the height data. "Evaluation of 3D shape should only be done when the colour texture is not displayed." (Zhao and Chellappa, 2006).

The following two chapters detail the design of the experiments and experimental apparatus constructed for this research project. It is important to note for verification of the apparatus realisation that the images presented in this chapter have been produced during testing stages of the photometric stereo apparatus.

6. Design of Experiments

To achieve the research goals of this project, it is necessary to conduct a number of practical experiments. These experiments are required to demonstrate the photometric stereo technique for face recognition. The remainder of this chapter describes the overall design of the experimental work and the rationale behind each individual experiment.

The overall aim of the experimental work is to investigate the theory behind 3D face recognition in practical terms using photometric stereo. Both height and gradient data will be used in the experiments in order to perform a balanced investigation. The experiments are separated into the following three groups:

- Feature extraction
- Face normalisation
- Face matching

These three experimental groups have been designed to follow the process used by the majority of face recognition systems as identified in the literature review. The natural progression of the experiments explains how the face recognition process is to be achieved, from the initial image acquisition to the final face matching process. The ordering of the experiments is critical, as each experiment will rely on results and research conducted in the previous experiment.

An important consideration when designing the experiments, which will later form the basis of the photometric stereo system, is to achieve automated processing while avoiding spending time re-creating procedures and techniques that have already been well researched and documented. The main process that will be done manually is face isolation. The process of isolating the face in 2D images is a well-researched topic, with numerous algorithms existing which can extract a face from the majority of background images. A detailed review of face isolation algorithms is given by Li and Jain

(2004) and Zhao et al. (2003). In order to focus on the prevalent issues of 3D face recognition in these experiments, the subjects face will be manually isolated by the user. The isolation of the image will be around the forehead, cheeks and chin, avoiding any hair where possible. It is important to note that the isolation will only be done approximately to ensure that if automatic isolating is implemented in the future that there is no reliance of the system for precise facial isolation.

The design of the first experiment group, feature extraction, involved first determining which features were to be extracted and then, what techniques would be used for extraction. This decision was made based on the reviewed literature of similar feature based systems. Based mainly on the work by (Hallinan et al., 1999), the nose was chosen as the initial face feature for extraction. Using the nose location as a point of reference, the eyes will be then located. "The most striking aspect of the face in terms of range data is the nose. If the nose can be located reliably, it can be used to put constraints on the location of eyes and other features." (Hallinan et al., 1999). The rationale behind choosing the nose as the reference feature of the face is it is the only feature which exhibits the following properties:

- Protrudes outward from the rest of the face
- Characteristic roof-like ridge which is vertically orientated
- Ridge falls on the approximate symmetry plane of the face

An experiment will be conducted using the symmetry of the face to locate the eyes. The design of this experiment is similar to that proposed by Gupta et al. (2004). Using the nose bridge and nose tip, an approximate vertical line of symmetry can be formed. This reference line can be used to facilitate the location of the eyes by limiting their location and adding the constraint that they must appear either side of the nose. A second eye locating experiment will be conducted using gradient maps without height data, in order to investigate the potential of non-integrated gradient data for feature location.

Height data will be processed by analysing cross sections of the face to identify patterns. This procedure is used in many 3D modelling systems and is described by Mian et al. (2005), as an efficient and robust way of identifying features that exhibit noticeable height differences to the surrounding area. The diagram in Figure 6.1, illustrates the intended method for presenting height data.

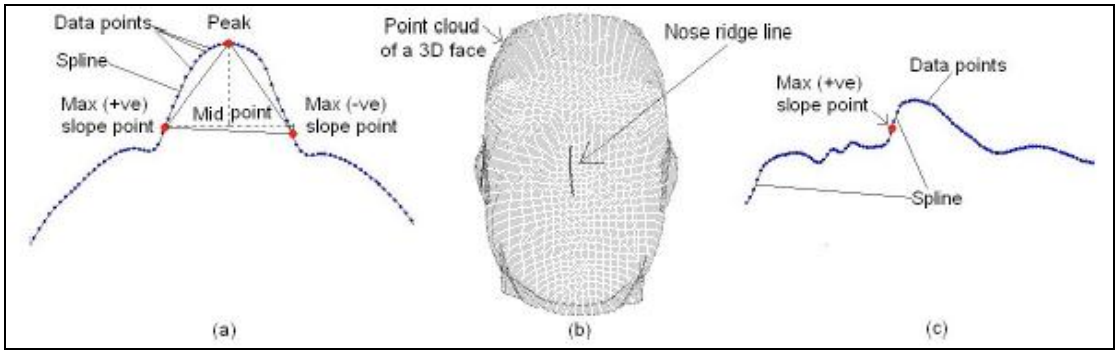


Figure 6.1 – Face profiles of height data (Mian et al., 2005)

The design of the normalisation experiment group is such that each experiment investigates the process of manipulating the 3D face, in order to standardise the size and positioning prior to recognition. Each of the normalisation experiments will use the feature information determined in the feature extraction experiments. The final experimental group will use the normalised face information from the previous experiments, and will investigate methods of matching a query face image with a database of enrolled face images. Different face matching procedures will be investigated, using both feature size and location information, with aim of determining which feature metrics produce the most reliable matching results.

6.1 Test Data

For the purpose of the experimental work, a small database of 35 images will be used containing seven different subjects. The size of the database is small compared to those used in the majority of face recognition research projects. This is due to the unavailability of any public face databases containing photometric stereo images, which are often employed in face recognition projects. Also the capture of photometric stereo images is often more time

consuming relative to images captured for 2D recognition tasks. The test data will be made up of a number of image sets, taken of a number of subjects. The image sets consist of three images taken of a subject's face under three different lighting directions. All images will be captured in a dark room to avoid ambient lighting. The face database is provided in appendix B and is divided into sections, one for each subject, identified by the letters 'A' to 'G'. Each section consists of all the image sets of the subject with details of facial feature locations.

The database of face images will be used for both the experimental work and verification stages of the project. The success of the experimental work will depend heavily on the quality of the images captured. As stated in chapter 5, the quality of the captured images will be determined by how accurately gradient data can be extracted. The verification stage will measure the effectiveness of the face recognition system in identifying a correct face match, when presented with a new image of a face contained in the database. From these results it will be possible to quantify the success of the face recognition system in terms of both accuracy and robustness.

7. Design of Experimental Apparatus

The method used for image acquisition is photometric stereo. This method enforces a number of requirements for how the images are captured and therefore adds a number of constraints on the overall design.

The following sections detail the lighting rig, the circuit used to control the lighting and the software used to process the captured images. It is important to note that all of the apparatus, including the lighting rig, the synchronisation circuit and the test software was created for this research project solely by the researcher.

7.1 Photometric Stereo Arrangement

The lighting arrangement used for this research project is based on that employed by Bronstein et al. (2004a) and Kee et al. (2000). These systems use three illumination sources equally distributed around a centralised camera and is described as providing the largest illumination coverage, which is desirable for angular objects. The configuration used for this research project is illustrated in Figure 7.1.

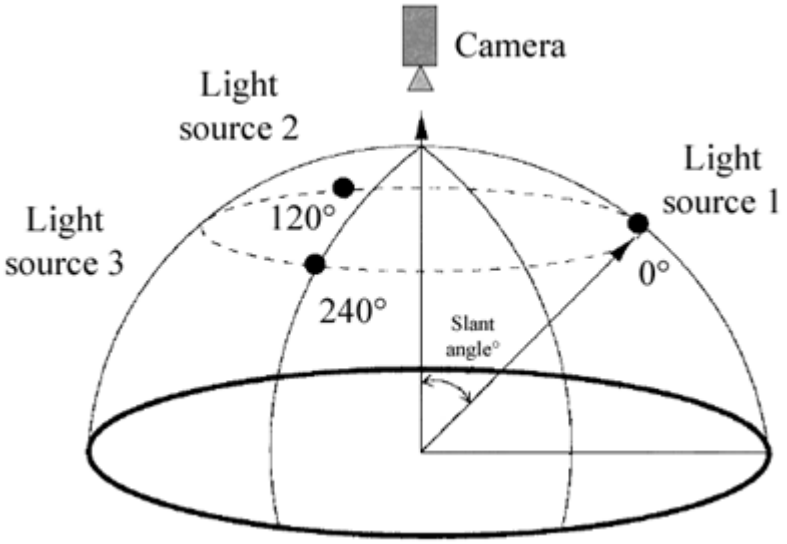


Figure 7.1 – Photometric stereo configuration (adapted from Kee et al., 2004)

The distances of the object from the lights dictate the slant angle of each light. The desired angle points the light to the centre of the object. As the only available lens was a 25mm F/1.4, this set the distance of the face to the camera at 195 cm and the slant angle to 50°.

The implemented photometric stereo capture apparatus is shown in Figure 7.2. Each of the three lights used, are in the form of a self-contained array of 50 light emitting diodes. These lights were chosen over other lights, such as incandescent, fluorescent and halogen, as the switch on/off times are much more rapid and significantly less power is used. The specification details of the lights are provided in appendix C.

Figure 7.2 – Photometric stereo rig: Focus on the lights

A side view of the capture apparatus is shown in Figure 7.3. This image highlights the positioning of the camera in relation to the lights. The camera used to acquire the image sets was chosen to be greyscale, as this would capture images with only the data that is needed. An equivalent colour camera could have been used, however, the resultant images would contain

three channels of colour, which is inherently slower and would take more time to process. The specification of the camera used is provided in appendix C.

Figure 7.3 - Photometric stereo rig: Focus on the camera

Finally a full view of the capture apparatus is provided in Figure 7.4. As this images shows the entire capture system is supported on an adjustable tripod, which by design is extremely stable and allows the system to be raised and lowered to suit the need of the user.

Figure 7.4 - Photometric stereo rig: Full view

This section has presented details of the photometric stereo capture apparatus. In order to produce a complete photometric stereo face recognition system, both a circuit and a piece of software are required to synchronise the lights and camera and process the captured image sets, this will be the subject for the remainder of this chapter.

7.1.1 Lighting Synchronisation Circuit

As previously stated, photometric stereo has been traditionally used to model inanimate objects such as museum artifacts. It clear that these objects do not move, thus images can be captured easily using manual switches to turn lights on and off and to take each photo. As the images captured in this research project are faces, a fast automatic capture process is required. To avoid any movement of the face between images it is essential that the three images are captured in quick succession. To achieve this, an electronic circuit is used to synchronise the switching on and off of the three lights and the capture of each of the three images.

The circuit designed and constructed for this research project is illustrated in Figure 7.5. The design of this circuit is inspired by a collection of circuits created by Braga (2002). This collection of circuits is based around a 4017 integrated circuit (IC), which produces a sequence of ten high/low outputs.

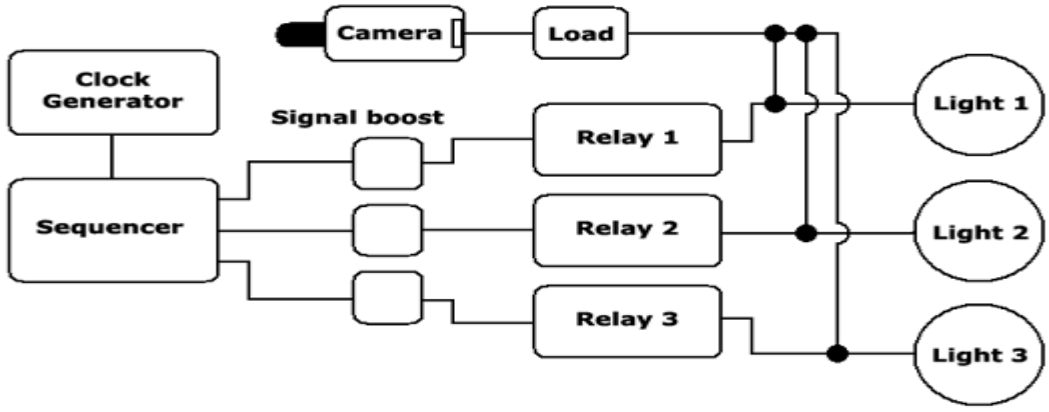


Figure 7.5 – Lighting and camera trigger circuit summary

The main design requirement of the synchronisation circuit was the ability to vary the switching speed between the three lights. This was essential in order to capture the three images in the shortest period possible. A 555 IC timer coupled with a potentiometer was used to produce the variable clock cycle that was passed into the 4017 IC, which in turn varied the light sequence and image capture speed. The finished circuit diagram is provided in appendix D.

It was found that the camera trigger is extremely sensitive to even the smallest fluctuation in current. Due to this issue, it was found that transistors could not be used to trigger the camera and lights simultaneously as the current leakage was too high. To cope with this, relays were used coupled with a diode to prevent any residual current triggering the camera inadvertently when switching off.

7.2 Face Recognition Test Software

An essential component of any automated face recognition system is the software used to both enrol subjects and process face queries. It has been

found during research, that no publicly available photometric stereo software exists. For this reason, in order to perform the experiments of this research project, software had to be designed and created.

Although many of the experiments and theories that were being investigated were chosen prior to the software design, it was supposed that during the initial experimentation stage other similar experiments would be conceived. Therefore, the software was designed as a test bed, providing various tools for testing many aspects of the photometric stereo data. These tools were designed for extracting, presenting and correlating all data from any set of three photometric stereo images.

The main design ethos was to create a piece of software that was highly functional, while remaining flexible and easily upgradeable. This was realised by selecting the key parts of C#, a high level programming language and Matlab, a matrix processing application. The reasons behind using both were that each provided certain technologies, either pre-built or that is simple to implement. The C# language provides a wealth of libraries, which make implementing graphical user interfaces (GUI's) straightforward and that allows for efficient code execution and debugging. The Matlab tool set also provides a wealth of libraries, but unlike C# these are specific to manipulating matrices. As the information of an image is encoded in a matrix, Matlab can be considered as an ideal processing tool to manipulate the data contained within these images. In addition to Matlab's matrix processing abilities, it has also the facility to interpret and run new code, without needing to be restarted. This provides the ideal functionality to allow users of the software to change code and manipulate any data present in Matlab's workspace, without the need to restart or compile and test code independently.

The overall design of the software is shown in appendix E. This design shows that the system is separated into a number of classes, each class encompassing one of the main elements of the software. The main software elements can be summarised as:

- Image acquisition
- Surface normal calculation
- Albedo calculation
- Gradient visualisation
- Surface reconstruction
- Feature identification
- Feature alignment
- Feature scaling
- Feature comparison

The finished software is composed of approximately 9,000 lines of code and is provided on the attached CD-ROM. All of the main functions of the code are commented, with an explanation of what each does. Pseudo code for the key processing functions is provided in appendix F. As stated in chapter 5, the Frankot-Chellappa integration technique is used for surface reconstruction. The algorithm is implemented in Matlab code (appendix G) and was obtained from the website,
www.csse.uwa.edu.au/~pk/research/matlabfns. The code is provided freely and is cited for public use in the authors associated publication, "Shapelets Correlated with Surface Normals Produce Surfaces" Kovesi (2005).

8. Feature Extraction

In this chapter a series of experiments are presented, which are designed to investigate finding facial features using both surface gradient and reconstructed range data.

A number of experiments are conducted which aim to locate the nose, the eyes and the mouth of all of the faces in the database. As stated previously the faces will be cropped by the user to remove the surrounding background and the subject's hair prior to feature processing. Figure 8.1 shows the position of each of the features used in these experiments and is based on the feature location diagram of Hallinan et al. (1999).

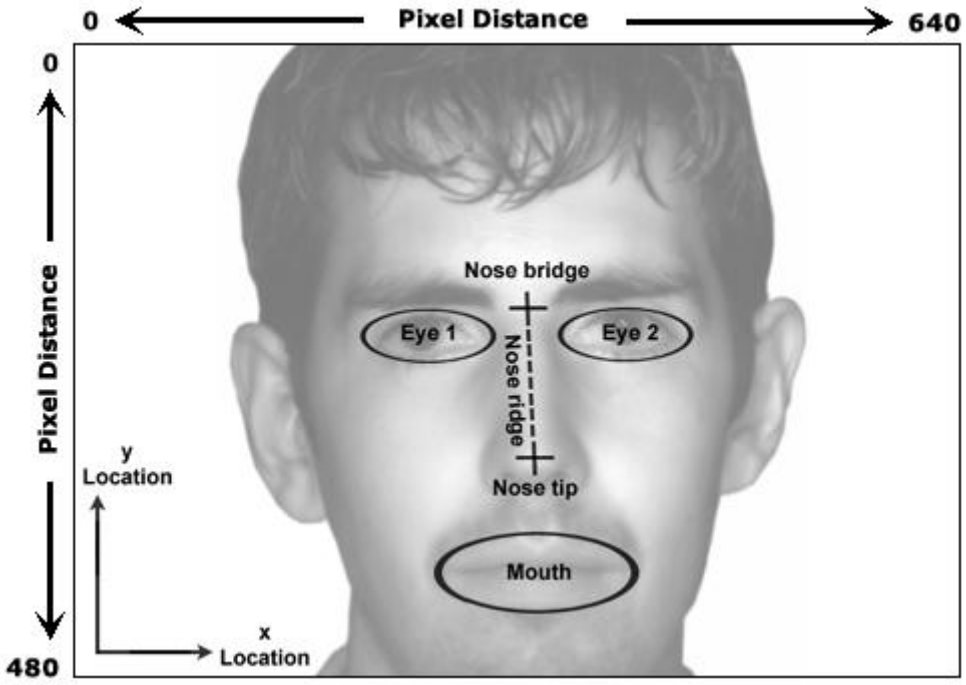


Figure 8.1 – Facial feature locations

As previously stated the organisation of the experiments conducted in this chapter is based on the feature localisation process described by Gupta et al. (2004 and Hallinan et al. (1999). Their procedures identify the nose as being a reliable starting position in order to put constraints on the position of the eyes and mouth, based on the symmetry of the face.

The following feature extraction experiments mainly use height data to identify facial features. The height data of the face surface can be plotted onto a graph in order to visualise a cross section of the face profile. The image of Figure 8.2 illustrates how height data plotted on a graph can be used to represent the horizontal profile of a face through the nose (Ng and Schlüns, 1998).

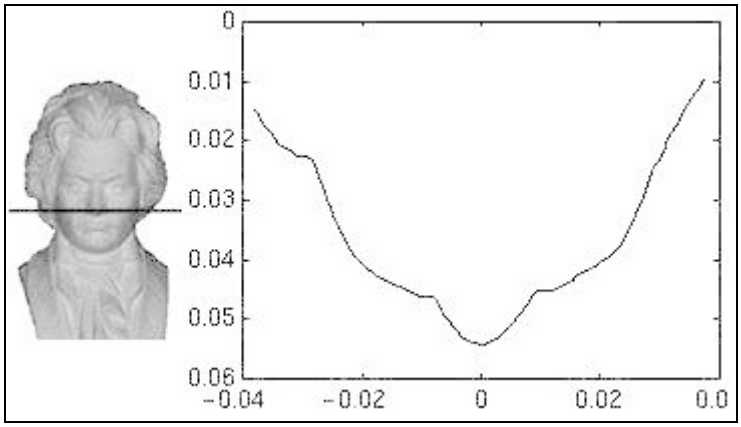


Figure 8.2 - A graph of horizontal height data (Ng and Schlüns, 1998)

It is necessary to understand the concept of plotting the height data in this way, as this representation is drawn upon in the experiments to explain the feature searching algorithms. In addition it is important to note that the research software provides the tools for plotting height data graphs. Consequently, all of the height data graphs of the subsequent chapter sections have been produced using the research software and actual images from the face database.

The experiments conducted in this chapter are successively organised. In order to avoid reiteration and to introduce research findings progressively, each experiment uses results and findings of the one previous.

8.1 Finding the Nose Tip

Using the approach described by Hallinan et al. (1999), the first feature to be located is the nose. The location of the nose will subsequently facilitate the location of other facial features. To handle different shaped noses, the tip of

the nose, in this research project, is defined as the most outwardly pointing section of the nose. Therefore the tip of the nose, in general, can be considered as the most protruding facial feature, exhibiting the largest height difference from the surrounding face surface.

In order to locate the nose tip a technique was conceived that segmented each face profile into a number of sections. As described previously, for each face a horizontal profile can be produced using the height values of the face surface. Therefore for each pixel on the y axis, a horizontal profile can be plotted. The images of Figure 8.3 illustrate (a) a horizontal profile line and (b) the location on the face where the profile line was formed from.

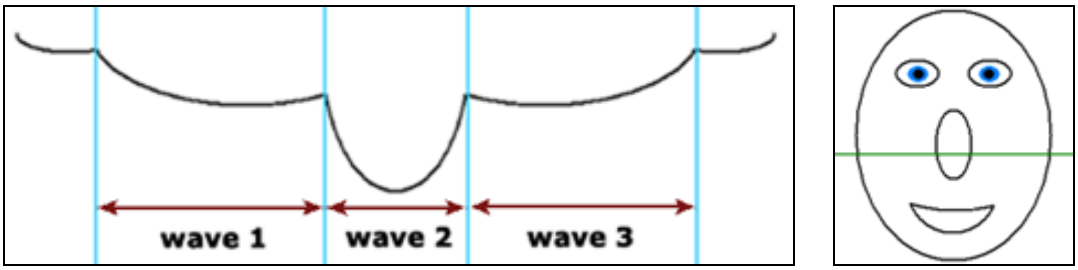


Figure 8.3 – (a) Horizontal profile through nose (b) Profile Location

In order to measure and extract information from the height data, the idea of segmenting the profile into a series of waves was considered. The technique, which all evidence suggests has not been used before, was conceived by inspecting the shapes of horizontal profile lines and observing that the height data of the face is represented by a number of smooth height variations from one distance to another. As can be seen in Figure 8.3 (a) a wave on the profile line is defined as a transition from a low-point to a high-point back to a low-point. This technique is used broadly in all experiments of this chapter.

The nose tip search algorithm is explained in the pseudo code which follows. As stated, this algorithm is based on the preposition that the nose tip is the most protruding feature of the face. Therefore, a simple search is made of the height map in order to locate the x and y location of the highest value. This highest value defines the height of the nose tip and the x and y coordinates define its location.

```
SET wave = 0

/* Get largest height values and set global variables */
FindLargestHeightValue()

/* Set the profile line to */
STORE horizontal profile_Line at largest_Height_Value_y

/* Loop through all the waves in the profile line */
FOR ALL waves in horizontal profile_Line
   /* Store the wave, which contains the largest height value x */
   STORE wave containing largest_Height_Value_x
END-FOR

/* Set facial feature locations */
noseTip_x =  wave_Largest_Middle_Value_Index
noseTip_y =  largest_Height_Value_y
noseTip_z =  largest_Height_Value
```

The nose tip algorithm implements the wave profile technique. However, at this point of the proceedings the location of the wave is not essential. The inclusion of the wave searching component will become apparent in further experiments, where the wave location is used to facilitate the extraction of other face features.

8.1.1 Results

The results of the nose tip search experiment are presented as a table of facial feature location coordinates and corresponding facial images with facial feature positions indicated. The facial images have been taken directly from the research project software. All of the results taken from the face database show that the highest point for each cropped face are located at the tip of the nose. The following diagrams illustrate examples taken from the experiment results. Further results showing evidence of nose tip locating, for each of the subject faces in the database, is provided in appendix B.

The images of Figure 8.4 show an example of the nose tip location and have been taken directly from the research project software. The two intersecting lines of Figure 8.4 (a) identify the largest height location of the face. The vertical line through the nose is illustrated in Figure 8.4 (b). The line is a

plotted using the height values through the y axis of the nose, which effectively represents a segment of the face profile one pixel wide.

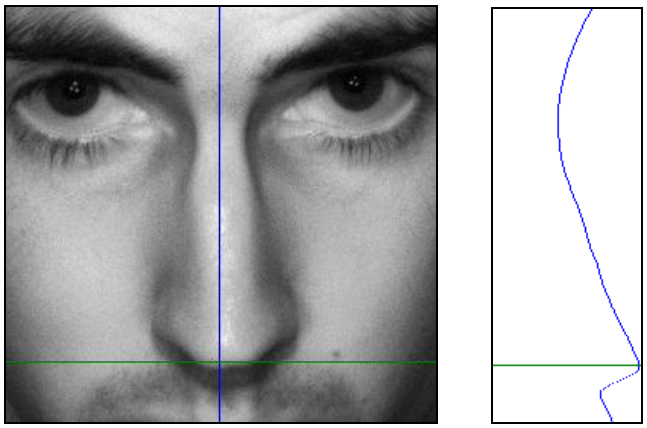


Figure 8.4 – (a) Nose tip (b) Vertical profile

The horizontal line through the nose tip is illustrated in Figure 8.5. Figure 8.5 (a) shows two vertical lines for the nose edge and a middle vertical line, which is the position of the highest point of the nose tip. It was found during the initial testing stage, that occasionally the highest point of the horizontal profile (x axis) was not located on the centre of the nose. This is illustrated in Figure 8.5 (a). To handle this issue and to correct the nose tip position to the centre of the nose, the horizontal profile line was clipped by a small percentage. The resulting clipped profile is then used to find the middle highest value, which identifies the nose tip x coordinate.

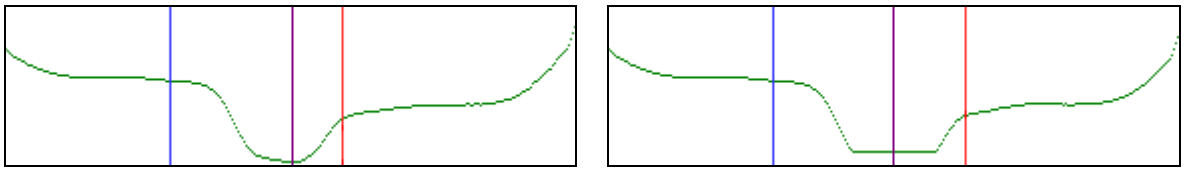


Figure 8.5 – (a) Horizontal profile (b) Height restricted to remove errors

The following pseudo code shows the addition to the nose tip search function which was used to clip the maximum height of the profile line.

```
/* Loop through all the height values in wave and remove top 10% */
FOR ALL values in wave
   /* If a value is within 10% of the largest value */
   IF wave_Height_Value >= (largest_Height_Value - 10%) THEN
      /* Cap the wave value to 10% off the largest value */
      wave_Value = largest_Height_Value - 10%
   END-IF
END-FOR
```

As can be seen from the algorithm, 10% of the height value is clipped, which through testing, was found to be an effective clip value.

8.1.2 Discussion

The results presented for this experiment has shown that the nose tip algorithm is capable of accurately locating the tip of the nose for each face in the database. Initial testing identified that height data alone does not always locate the nose tip precisely. To handle this inaccuracy an improvement of the original algorithm is presented, which is shown to provide a more robust positioning. Although the cause of the anomalies in the height data is presently unknown, the height information suggests that it is due to reflectance at the tip of the nose. This is suggested, as the height data fluctuates at the points of the nose where high image intensity values can be seen on the image. This identification cannot be presented more clearly than the profile line shown in Figure 8.5 (a), as the variation in intensity value is only slight.

This experiment has introduced a novel technique for obtaining information from the height map, by separating the data into a series of waves. This elucidation allows horizontal profiles of the face to be described in higher terms than single height values alone. The technique developed in this experiment, will be extended and applied to each of the experiments that follow.

8.2 Finding the Nose Bridge

The bridge of the nose is the lowest part of the nose and is situated between the nose tip and the forehead. No stipulation is made on exactly where the

nose bridge should be located, only that the location is the same for all images of a single subject face. As with the nose tip algorithm, height profiles will be utilised in order to locate the x and y coordinates of the nose bridge. The aim of locating the nose bridge is that in conjunction with the nose tip location, the length of the nose should be determinable by subtraction of the nose tip and nose bridge y coordinates.

The nose bridge algorithm uses the nose tip as a reference point to locate the bridge. The following pseudo code outlines the nose bridge algorithm.

```
SET temp_Wave            = 0
SET temp_Highest_Value   = 0
/* While loop count index */
SET current_y            = noseTip_y

/* Set the profile line */
STORE horizontal profile_Line at largest_Height_Value_y
/* Get largest height values and set global variables */
FindLargestHeightValue()
/* Store the wave, which contains the largest height value x */
STORE wave containing largest_Height_Value_x

/* Loop: Start from noseTip_y and decrement until 0 is reached */
WHILE current_y >= 0
   /* Set the profile line to the current_y index */
   STORE horizontal profile_Line at current_y
   /* Get the wave on the current profile line at the noseTip x */
   temp_Wave = wave containing noseTip_x

   /* If previous is less than the highest then check for a trend */
   IF temp_Highest_Value  < temp_Wave THEN
        /* Set facial feature locations */
        noseBridge_x = middle index of temp_Highest_Value
        noseBridge_y = current_y
        noseBridge_z = temp_Highest_Value
        /* Exit the WHILE loop */
        BREAK
      END-IF
   /* Set the highest value to that of the current wave */
   temp_Highest_Value = temp_Wave
   /* Decrement the while current_y index */
   current_y --
END-WHILE
```

The concept of the face profile as a series of waves is used to a greater extent in this experiment than the previous. As the pseudo code shows, first the wave containing the nose tip location is identified. A loop is then used, starting from the nose tip y location and terminating at the top of the image. The loop first stores the wave containing the nose tip x location for the current

profile line on the y axis and then compares the highest value of the current wave with the one of the previous. With this algorithm the nose bridge is declared as found when the current highest value is greater than the one previous.

8.2.1 Results

The results of the nose bridge search experiment are provided in appendix B. The results are presented as a table of facial feature location coordinates and corresponding facial images with facial feature positions indicated. As with the nose tip experiment, all results have been produced using the subject images from the face database. The images of Figure 8.6 show an example of the nose bridge location and have been taken directly from the research project software. The two intersecting lines of Figure 8.6 (a) identify the location of the nose bridge for this face. As with the nose tip experiment, Figure 8.6 (b) illustrates a vertical cross section through the centre of the nose. It can be seen in this image that the horizontal line intersects the nose profile at the point where the nose height is starting to increase.

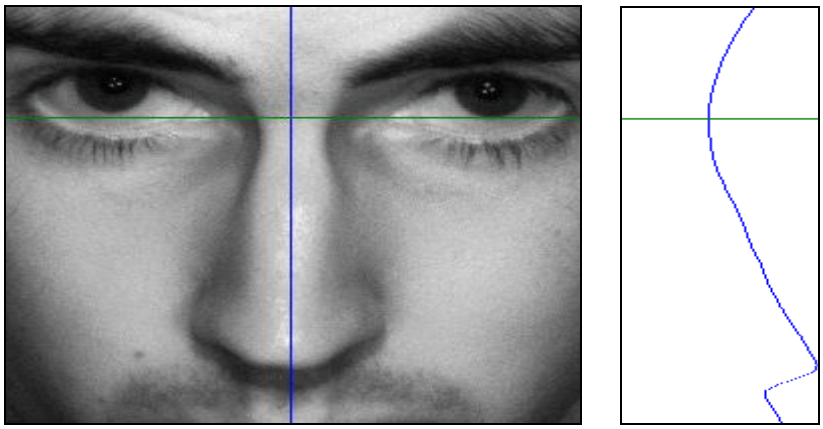


Figure 8.6 - (a) Nose tip (b) Vertical profile

It was found during the initial testing stage that often the nose bridge was being located in the middle of the nose, when it appeared on the profile graph to be at a higher position. This was found to be due to small anomalies in the height data caused by specularities, much like those identified in the nose tip

experiment. To handle this problem, the following pseudo code was added to the original algorithm.

```
SET TREND_MIN = 5

/* If trend iteration has reached the trend constant then
   save the location of the nose bridge/mid and exit */
IF trend_Iteration >= TREND_MIN THEN
/* Set facial feature locations */
  noseBridge_x = middle index of temp_Highest_Value
  noseBridge_y = current_y - TREND_MIN
  noseBridge_z = temp_Highest_Value
  /* Exit the WHILE loop */
  BREAK
END-IF
/* Increment the decreasing height trend count */
trend_Iteration ++

/* If trend has stopped then reset the trend iteration count */
ELSE IF trend_Iteration > 0 THEN
  trend_Iteration = 0
END-IF
```

The addition of this code allows for a number of 'error' pixels to be ignored. As this code demonstrates, a 'trend' counter is used, in order to enforce a minimum number of pixels to all be greater than the previous, before the nose bridge location is confirmed. The code addition uses a constant value for the minimum increasing 'trend' condition. This value is set to an arbitrary 5 value, which worked for all images of the face database. If necessary this value could be adjusted if further testing with a larger data set proves it to be too low.

8.2.2 Discussion

The experiment presented here has demonstrated that it is possible to locate the nose bridge using the nose tip as a reference point. Results have been presented which verify the location of the nose bridge for each of the subject faces in the face database. An improvement to the initial searching algorithm has been described, which aims to handle height anomalies. The improved algorithm uses height trends to enforce a minimum number of height changes that are required before location can be confirmed.

With the location of the nose bridge and previous location of the nose tip, a number of further feature points can be extracted and facial approximations made. Using the distance from the nose bridge to the nose tip, the mid-point can be calculated as shown here:

```
noseMid_x    = noseBridge_x
noseMid_y    = noseBridge_y + ((noseBridge_y - noseTip_y) / 2)
noseMid_z    = height_Map(noseMid_x, noseMid_y)
```

In addition to the location of the mid-point of the nose, the height is also found. This further nose location point provides more information which can be used to describe the nose as a separate entity to the face. The illustration in Figure 8.7, demonstrates how the nose mid-point and nose height can be used to describe the shape of the nose.

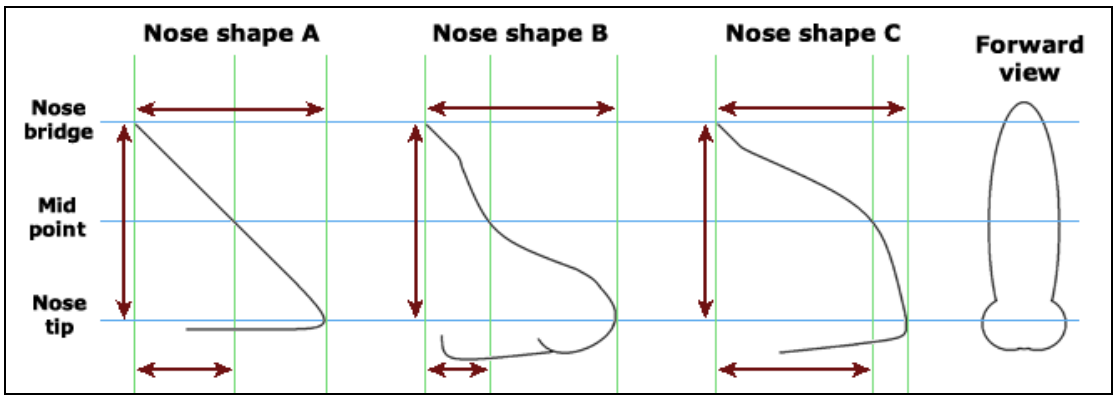


Figure 8.7– Nose Shape examples: Demonstrating use of mid-point

The addition of the nose bridge location, allows for a line of symmetry to be placed on the face, using the x coordinates of the two nose points. This line of symmetry has been previously identified by Hallinan et al. (1999), where the face division is used as a constraint in approximating the location of the eyes. In addition the results also identify that the eyes of all of the subject faces are approximately on the same horizontal line as the nose bridge. This observation could prove useful as a reference point for locating the eyes and will be the subject of the next experiment.

8.3 Locating the Eyes Using Height Data

The final feature extraction experiment involves locating the position of the eyes. Using the eyes and nose locations it is believed that a description of the face in terms of length and width can be realised. The technique, which will be employed for this experiment, is based on the findings from the initial testing stage of the photometric stereo software. It was observed during initial trials of the software that a large variation in the height data existed around the pupils of the eye. These observed height variation can be seen in the example of Figure 8.8.

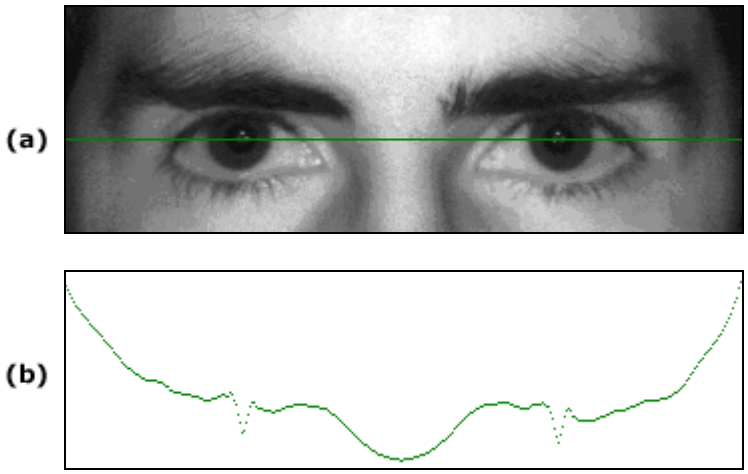


Figure 8.8 – (a) Profile Location (b) Horizontal profile through eyes

The image of Figure 8.8 (a) shows a section of the face, with a horizontal line through the eyes, highlighting the position where the height data has been extracted from. The face image has been rotated in order to show the height data of both eyes in a single profile line. As it can be seen in Figure 8.8 (b), two anomalies of approximately the same shape appear in the height data at the location of the eye pupils. This observation is also made in the work by Hallinan et al. (1999), "specularities in the pupil are more difficult because their intensities are typically extremely different from those of the rest of the iris". In their work they describe how specular surfaces of the face can be controlled by smoothing the surface where they appear. Eye specularities are also observed by Zhao and Chellappa (2006), where these anomalies or 'spikes' are identified as artifacts in the 3D data. This research evidence shows that a number of researchers have identified these 'artifacts' in the pupils of the eyes, thus it can be seen that the anomalies are not due to the capture system or processing software used for this project. Therefore, it will be the

aim of this experiment to investigate the potential of using height variations in the eyes in order to find their location.

As with the previous experiment, the eye location technique will involve segmenting the height profile line into a series of waves. The wave at the first eye pupil will be found manually and used to search for the second eye. In order to match the waves that are present in both eyes, it is necessary to extend the profile wave concept further. Therefore, in addition to separating the profile into a series of waves and using only middle and highest positions, ten other wave attributes have been identified. These twelve attributes, which describe the overall shape of each wave, are shown in
Figure 8.9. This illustration shows twelve dimension arrows, each arrow is associated with a single attribute of the wave.

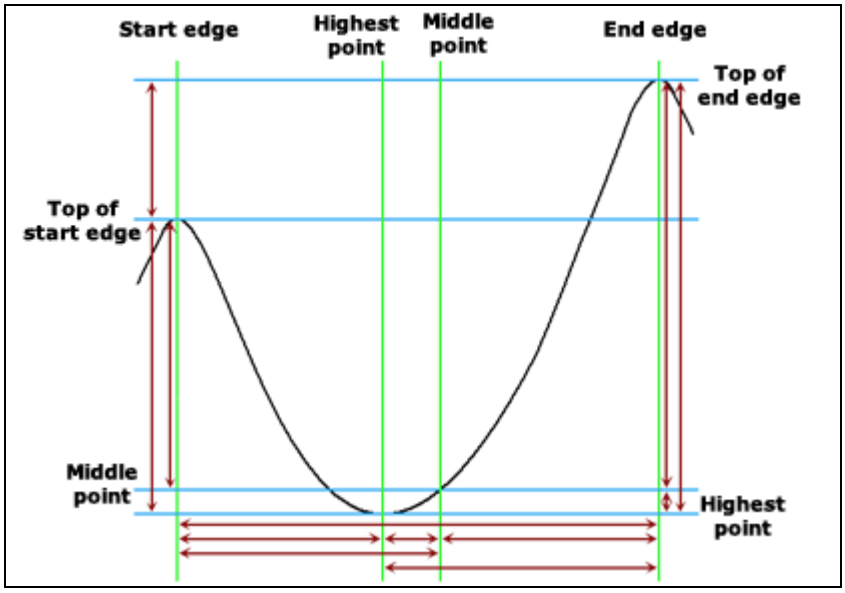


Figure 8.9 – The twelve wave attributes

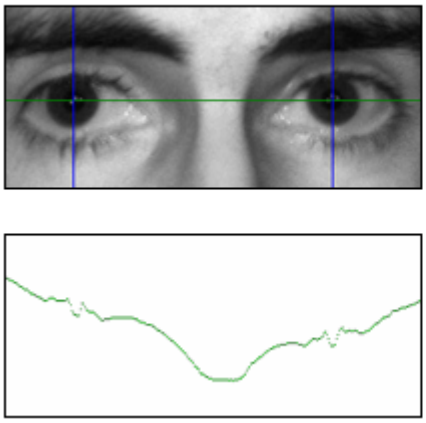
The algorithm used for the eye localisation and matching is provided as pseudo code in appendix F. The algorithm first calculates the twelve attributes of the manually selected wave of the query eye pupil. A search is then performed by plotting a height line for each y location of the face, from the nose x location on the opposite side of the query eye to the edge of the face. The height line is then segmented into a number of waves and each wave

compared to the query wave. The wave comparisons are made using the percentage difference of the attributes for both the current wave and the query wave. The wave comparison which produces the highest percentage match is selected as the second eye. The results of the eye locating algorithm performed on images from the face database are given in the following section.

8.3.1 Results

The results presented in this section have been taken from the complete eye location experiment results, provided in appendix H.

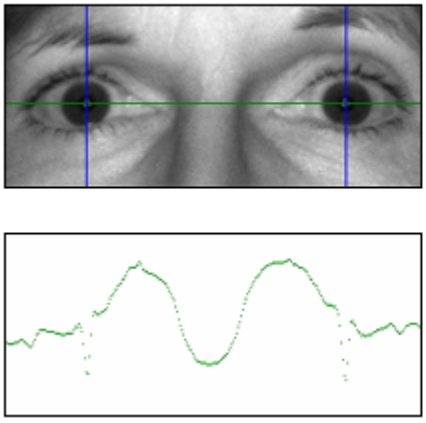
Subject A: Set 1



Wave attribute	Eye 1	Eye 2	Match %
Edge difference	-7.589	-10.07	75.385
Start to highest value	111.9	74.67	66.717
Middle to highest value	12.76	14.37	88.737
End to highest value	104.3	64.61	61.922
Start to middle value	99.17	60.3	60.803
End to middle value	91.58	50.23	54.849
Width start to end	12	11	91.666
Width start to middle	6	5	83.333
Width start to highest	7	4	57.143
Width middle to highest	1	1	100
Width middle to end	6	6	100
Width highest to end	5	7	71.429

Table 8.1 – Eye matching results for subject A

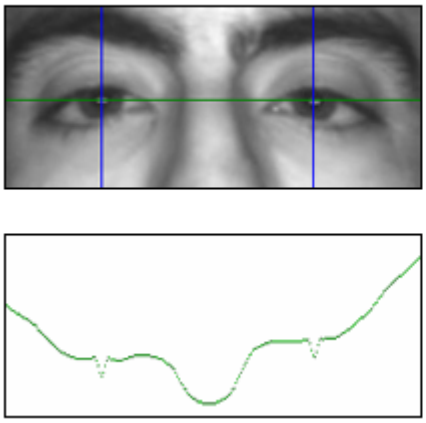
Subject B: Set 1



Wave attribute	Eye 1	Eye 2	Match %
Edge difference	28.06	-36.64	0
Start to highest value	113.7	88.0	77.407
Middle to highest value	25.94	9.545	36.796
End to highest value	141.8	51.36	36.23
Start to middle value	87.75	78.46	89.41
End to middle value	115.8	41.81	36.103
Width start to end	10	11	90.909
Width start to middle	5	5	100
Width start to highest	4	4	100
Width middle to highest	1	1	100
Width middle to end	5	6	83.333
Width highest to end	6	7	85.714

Table 8.2 – Eye matching results for subject B

Subject C: Set 1



Wave attribute	Eye 1	Eye 2	Match %
Edge difference	12.61	28.08	44.9
Start to highest value	92.39	82.41	89.20
Middle to highest value	0	0	100
End to highest value	105	110.5	95.03
Start to middle value	92.39	82.41	89.20
End to middle value	104.9	110.5	95.028
Width start to end	12	10	83.333
Width start to middle	4	5	80
Width start to highest	4	5	80
Width middle to highest	0	0	100
Width middle to end	4	5	80
Width highest to end	4	5	80

Table 8.3 – Eye matching results for subject C

The results presented in Table 8.1, 8.2 and 8.3 details the eye matching evidence for the first three subject faces taken from the complete results of appendix H. Each result is made up of a table of wave attributes and match percentages between the two eyes, the actual found eye locations and the height profile line plotted at the y position of the eyes.

8.3.2 Discussion

This experiment has investigated a novel technique for locating the position of the eyes. The results presented demonstrate that the eyes can be accurately found in all image sets, where both pupils are visible. The method presented is currently semi-automatic and requires the location of one eye in order to automatically locate the second. As stated, specularities, which occur in the eyes due to specular reflection, are used to locate the corresponding position in both eyes. It has been found that the height data variations in the pupils of the eyes can be represented as a wave pattern and twelve dimensions calculated. These dimensions are used to match a corresponding wave on the opposite side of the face defined by the nose line symmetry. For all open eye face images of the database, it has been found that for each query eye wave the best match wave is the location of the second eye pupil.

It can be seen from the eye localisation experiment that a potential refinement of the wave matching algorithm can be made. The current algorithm searches half of the face, which is on the opposite side of the nose from the query eye. This algorithm could be further refined by reducing the potential location of the matching eye. From the feature locations provided in appendix B, it can be observed that the eyes are located approximately on the same horizontal line as the nose bridge. To be more precise, the eyes relative to the nose bridge vary from -15 pixels higher to +48 pixels lower. Therefore, the nose bridge location could be used to further constrain the initial likely location of the matching eye. An excepted percentage match would be required with this method in order to perform a wider search if a matching eye is not found.

It could be argued that using specular reflection of the eye pupils is not an appropriate technique, as the height variations are essentially errors created as a Lambertian surface is being assumed. However, these variations are observed in at least two other independent research projects and it has been shown that this technique can achieve accurate eye locations with the potential to be fully automated.

8.4 Locating the Eyes Using Gradient Data

The aim of this experiment is to investigate the potential of using gradient data for locating the position of the eyes. This differs from the previous experiments, as those used reconstructed gradient data, which provides height information for each pixel in place of surface normal data. Using gradient data for identifying and quantifying shapes of the face is a relatively untouched research area, with only the research by Bronstein et al. (2004b) identified during the review of literature. Their research investigates the use of surface gradient data in order to determine local properties of the face surface. "Avoiding the explicit surface reconstruction stage saves computational time and reduces numerical errors." (Bronstein et al., 2004b). The further aim of the experiment is to investigate whether closed eyes can be located using this method. "most eye-localization techniques assume some geometric and

textural models and do not work if the eye is closed." (Zhao and Chellappa, 2006).

In order to extract information from the gradient map a graph will be used to plot each surface normal based on its x and y value. As the research currently achieved in the area of shape matching using gradient data is relatively small, the aim of this experiment is to simply investigate the potential of using gradient graphs in order to locate the eyes. This will involve producing a number of gradient graphs from specific, manually selected, sections of the face and analysing the results.

The algorithm used to create the gradient graph is shown in pseudo code below. The steps used, is to first create a graph with the x and y axes corresponding to the minimum and maximum surface normal values. The next step is to place a mark on the graph for each of the surface normals. The x and y components of each surface normal is used as the location for the graph mark. Finally the graph is displayed with a central cross overlaid and the minimum and maximum surface normal values at each point

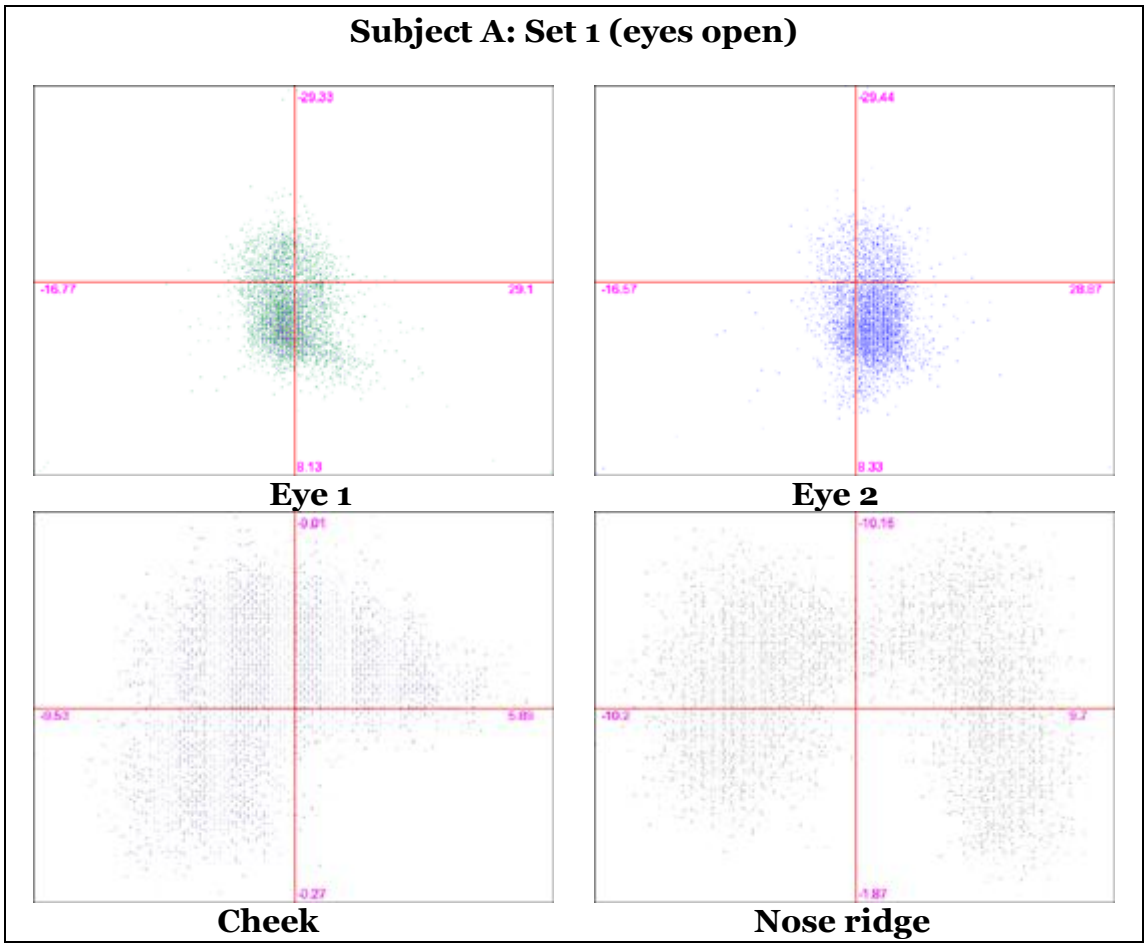
```
/* Create graph with the x and y scale to fit the surface normals */
CREATE Graph to accommodate min/max surface_Normals

/* Loop through all the surface normals from the current image set */
FOR ALL surface_Normals
   /* Store the surface normal x and y components */
   normal_x = GetXSurfaceNormal(x_Index, y_Index)
   normal_y = GetYSurfaceNormal(x_Index, y_Index)

   /* Put a mark on the graph at the normal x and y position */
   MarkOnGraph(normal_x, normal_y)
END-FOR

/* Show the graph on screen */
Display Graph
```

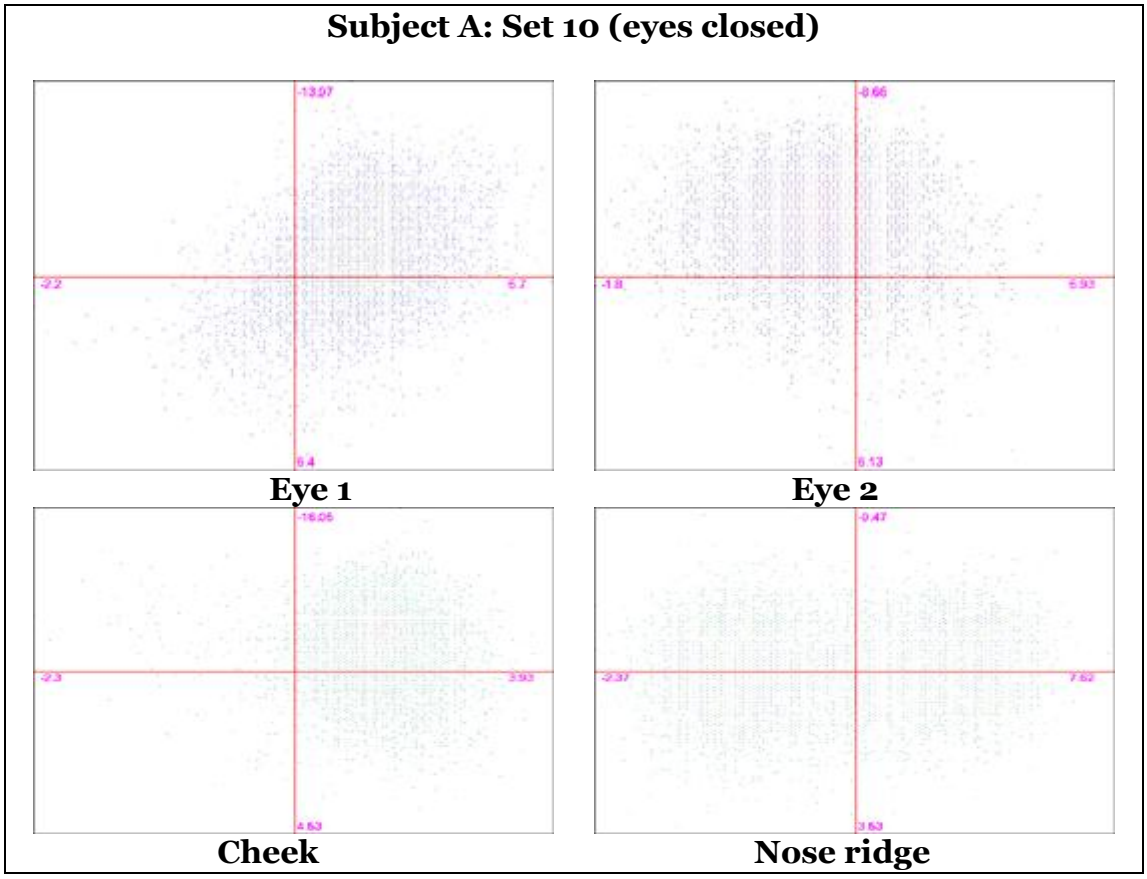
In order examine the gradient graphs for the eyes it will be necessary to provide a controlled comparison from other sections of the face. Results will be presented for a number of faces, consisting of a number of gradient graphs produced from different areas of the face. The faces that will be examined

have been selected from the face database based on image sets where both eyes are open and closed.

8.4.1 Results

The results presented in this section have been taken from the complete gradient plotting results, provided in appendix I.





These two sets of results show four gradient graphs for four sections of the face. Both graphs have been produced from the face of subject A. The first set of graphs the eyes are open and the second set of graphs the eyes are closed.

8.4.2 Discussion

The aim of this experiment has been to demonstrate the use of gradient data as a method for locating the eyes. A technique has been developed to display the gradient data by plotting each surface normal orientation onto a graph. A number of gradient graphs have been plotted for each face, to compare the gradient distribution for both the eyes and two other control face locations. The control locations were chosen to be the cheek and the ridge of the nose, as these areas exhibit both a flat surface and a large curved area. Results have been presented for a number of faces with both eyes open and eyes closed. This has been done to investigate if gradient data can be used to identify the location of the eyes both open and closed. This has been done as most eye

location algorithms, including the previous experiment, are unable to locate the position of closed eyes.

The results for the open eye face images show that the gradient data is distributed around a single point, close to the centre of the graph, whereas the gradient data for the control face areas are widely distributed. The results for the closed eye face images show that for both the eyes and the control areas equal distribution occurs. These results have shown that with the eyes open it is possible to observe localised gradient data on the gradient graphs, which is not present in other areas of the face. This evidence shows that it could be possible to search for these localised gradient distributions in order to locate the area of the face where the eyes are located. The results, where the eyes are closed have not shown any noticeable difference in distribution compared to the control locations. Therefore, from this evidence alone it cannot be determined if gradient data can be used to locate the position of closed eyes.

It is not possible to directly compare the results of this experiment with that of the previous height data based eye location experiment. This is due to the results presented in this experiment being based on manual selection and observed areas. However, this experiment has identified that gradient data has the potential to be used for eye location based on localised distribution of the surface gradient orientations. It has further been shown that compared to the height based experiment, no surface integration is required and only a simple algorithm is needed to produce the gradient graphs. It could be argued therefore that this technique could provide a computationally faster method for eye localisation compared to the previous height based method. Further investigation is required to identify a technique for measuring local distribution in the gradient graphs.

9. Face Normalisation

In this chapter, a number of experiments are conducted which aim to produce a normalised face that can be used in direct comparisons with other normalised faces. As stated earlier, it is essential in any face recognition system to enforce standards that all face images must adhere to prior to the recognition process. It is the aim of these experiments to make the recognition system more robust to variations in head orientation and distance from the camera.

The common normalisation criteria, as noted by (Li and Jain, 2004), are size, pose orientation and illumination. Systems which are able to accurately normalise variances of these criteria can be considered invariant to their effects. In many 3D modelling techniques, the issue of illumination is either irrelevant or, as in the case of photometric stereo, a much higher concern which must be dealt with before all other processing. It is important to reiterate that the face albedo image, which is created, after the surface normals are calculated, is a 2D image of the face normalised against illumination.

In these experiments the face database will be used to test the normalisation methods and measure the results. Methods for both face size and pose orientation will be presented. Face orientation can be defined by three rotations of the x, y and z planes. Heseltine et al. (2002), defines these orientations as roll, tilt and pan. The face images in Figure 9.1 illustrate these rotations. It can be observed that roll is a 2D rotation around the centre point; tilt is an up-down nodding rotation; pan is left-right rotation.

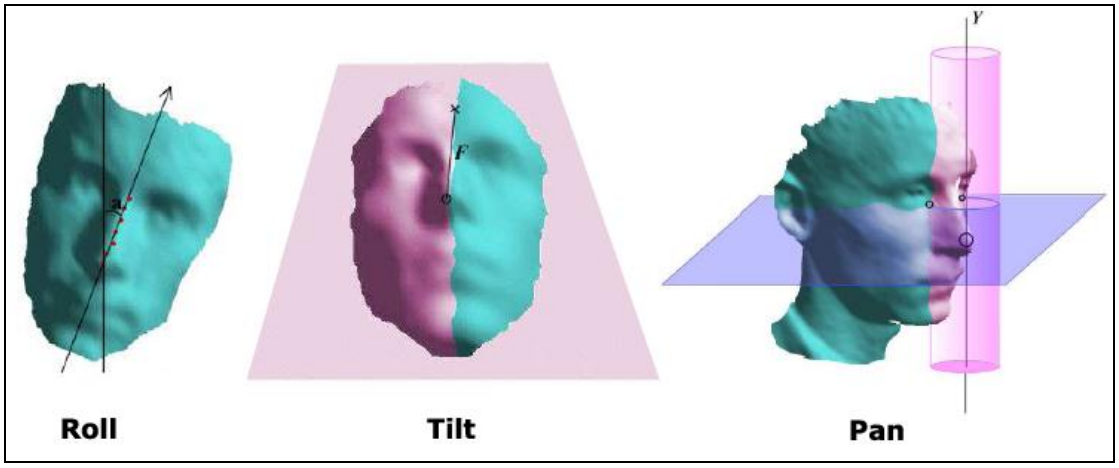


Figure 9.1 – Head orientations – images from Heseltine et al. (2002)

Of the three types of orientation, only roll and tilt will be considered in the experiments which follow. Normalising pan orientations will not be attempted in the experiments of this research project. This is due to the extreme complexity which is associated with rotating 3D models produced with photometric stereo. As stated earlier the models created in photometric stereo are not 3D models of the entire head, but of only the facial features which are in view of the fixed camera. This can result in missing data when the face is normalised. A number of research papers have addressed this issue; in the work by Heseltine et al. (2002) correspondence is used to reconstruct the missing data based on data of the opposite side of the face. This technique, although proven effective, shares many of the difficulties associated with stereo correspondence imaging.

9.1 Correcting Face Roll

In order to rotate a 2D object around a fixed centre location, two or more points of the object must be known (Zhao and Chellappa, 2006). Of the face features found in the previous experiment chapter, the eyes appear to be the apparent choice for re-aligning the face. The face would be considered normalised against roll if the eye 1 and eye 2 y location values are equal. The nose bridge and nose tip x values could also be used for alignment in the same way. However, if the pan orientation of the face is not centred then the nose bridge and nose tip would not be in line when considering 2D rotation.

The initial rotation algorithm used the eye locations to calculate a rotation angle, which could be used to rotate the image set. Figure 9.2 shows two images of a face exhibiting roll and a normalised face where both eyes are located on the same vertical plane.

Figure 9.2 – (a) Rotation angle (b) Normalised roll orientation

The image of Figure 9.2 (a) illustrates the calculation used to produce the rotation angle. The following sample code demonstrates the actual calculation used in the software.

```
/* adjacent = 65, opposite = 30, angle = 24.78 */
double opposite = g_Current_Face.eye1_y - g_Current_Face.eye2_y;
double adjacent = g_Current_Face.eye2_x - g_Current_Face.eye1_x;
double angle    = Math.Atan(opposite / adjacent);
```

This rotation algorithm produces a rotation angle, which is used to rotate the image set. The algorithm requires the location of the eyes, which in turn requires the location of the nose. Following the rotation of the image set, the location of the nose is lost and therefore relocation would be required.

A more efficient normalisation method, which would only require feature location to be performed once, is to calculate the normalised locations for just the facial features. The diagram of Figure 9.3 illustrates how the facial features can be divided into a series of four triangles labelled from t1 to t4.

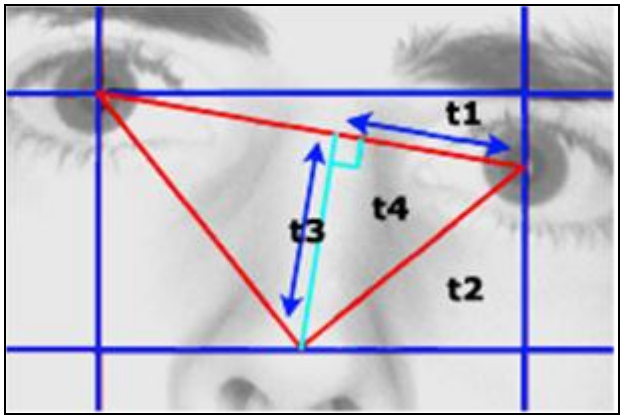


Figure 9.3 – Feature rotation triangle

The improved rotation algorithm is an extension of the previous algorithm, where the normalised eye locations were identified. This algorithm uses feature locations in order to calculate the dimensions and angles of four internal triangles. The triangle t4, shown in Figure 9.3, is the main triangle used to normalise the nose tip and nose bridge locations. The following pseudo code explains the improved feature rotation algorithm.

```
/* Triangle 1 Calculations */
t1_Opposite   = eye2_x - eye1_x
t1_Adjacent   = eye1_y - eye2_y
t1_Hypotenuse = GetHypotenuse(t1_Opposite, t1_Adjacent)
t1_Angle      = GetAngle(t1_Opposite, t1_Adjacent)

/* Triangle 2 Calculations */
t2_Opposite   = eye2_x    - noseTip_x
t2_Adjacent   = noseTip_y - eye2_y
t2_Hypotenuse = GetHypotenuse(t2_Opposite, t2_Adjacent)
t2_Angle      = GetAngle(t2_Opposite, t2_Adjacent)

/* Triangle 3 Calculations */
t3_Angle1     = 180 - (t1_Angle + t2_Angle);
t3_Adjacent   = t1_Hypotenuse
t3_Opposite   = t2_Hypotenuse
t3_Hypotenuse = GetSideUsingCosine(t3_Adjacent,t3_Opposite,t3_Angle1)

/* Triangle 4 Calculations */
t4_Angle1     = t3_Angle1
t4_Angle2     = 180 - (t4_Angle1 + 90)
t4_Angle3     = 90
t4_Hypotenuse = t3_Opposite
t4_Opposite   = GetSideUsingSine(t4_Angle3, t4_Angle2, t4_Hypotenuse)
t4_Adjacent   = GetAdjacent(t4_Opposite, t4_Hypotenuse)
```

As this code shows, each triangle is used to calculate the sides and angles of other triangles, finishing with the calculation of the fourth centre triangle. The opposite and adjacent sides of the fourth triangle provide the distance of the nose tip to the eyes y and eye 2 x locations respectively. The final piece of

pseudo code presents the method used to adjust the current feature locations to the new normalised locations.

```
/* As eye_1 is used for aligning its location remains */
eye2_y = eye1_y
eye2_x = eye1_x + t1_Hypotenuse

old_noseTip_y = noseTip_y

noseTip_x     = eye2_x - t4_Adjacent
noseTip_y     = eye1_y + t4_Opposite
noseBridge_x  = eye2_x - t4_Adjacent
noseBridge_y  = noseBridge_y + (noseTip_y - old_noseTip_y)
```

Both the face roll normalisation algorithms presented in this section have been implemented in the research software. The following section presents results of the face roll implementations carried out using the face database.

9.1.1 Results

Presented in this section is evidence of face roll normalisation using the developed feature triangle solving algorithm.

Subject	Rotated (yes/no)	Rotation angle (°)	Eyes				Nose			
			1 x	1 y	2 x	2 y	Bridge X	Bridge Y	Tip X	Tip Y
A (Set 1)	no		219	135	404	133	330	115	330	271
A (Set 1)	yes	0.691	219	135	404	135	328	116	328	272
B (Set 1)	no		324	230	489	236	399	218	399	324
B (Set 1)	yes	-2.083	323	236	489	236	402	221	402	327
C (Set 1)	no		254	181	377	166	324	139	324	240
C (Set 1)	yes	6.953	254	181	377	181	316	147	316	248
D (Set 1)	no		302	246	438	241	379	220	379	311
D (Set 1)	yes	2.106	302	246	438	246	376	222	376	313
E (Set 1)	no		213	144	415	140	316	135	316	281
E (Set 1)	yes	1.912	213	144	415	144	312	136	312	282
F (Set 1)	no		263	176	421	185	333	146	333	252
F (Set 1)	yes	-3.260	262	185	421	185	336	150	336	265
G (Set 1)	no		276	215	438	209	368	208	368	336
G (Set 1)	yes	2.121	276	215	438	215	363	211	363	339

Table 9.1 – Face roll normalisation results

The results presented in Table 9.1 show the normalisation results for each subject of the database. The eyes and nose feature locations before and after normalisation are given in addition to the angle of face rotation.

9.1.2 Discussion

This experiment has achieved its aim of normalising the face against roll. Results have been presented, which demonstrate the normalisation algorithm processing images from the face database. The method used is based on solving a number of triangles, using the eye locations to produce a rotation angle. The use of triangles means the calculations are simple to understand and implement compared to other roll normalising techniques. As the eyes are being exclusively used in the normalising algorithm, the accuracy of the rotation angle is solely based on the accurate location of the eyes. This would seem to suggest that a more robust algorithm could be achieved if more facial features were used to ensure accurate reference location points. In general this method should not be affected by variances in subject gaze, as eyes typically move in unison which indicates that lining up the pupils will still normalise against subject face roll.

The initial algorithm demonstrates that a rotation angle can be calculated based on the eye locations, which can be used to rotate the image set. The rotated image set consists of 2D normalised images, which potentially could be used in conventional 2D face recognition systems. This finding demonstrates how 3D height data can be used, although indirectly by aiding eye location, to implement 2D image processing techniques.

9.2 Correcting Face Tilt

According to Zhao et al. (2003) the issue of differing face tilt in 2D face recognition has one of the largest effects on reliable recognition after varying illumination. The issues of face tilt is also described by Hallinan et al. (1999), who identifies that the perceived size of facial features in 2D images can vary considerably with differing face tilt orientation. Therefore, the aim of this experiment is to normalise the face to a known repeatable tilt orientation.

The images shown in Figure 9.4, demonstrate how the length of the nose and distance from the eyes to the nose tip can vary dramatically with increasing

head tilt. The three images are of the same subject, with each image overlaid with a feature triangle and vertical line, identifying the eyes, nose tip and nose bridge respectively.

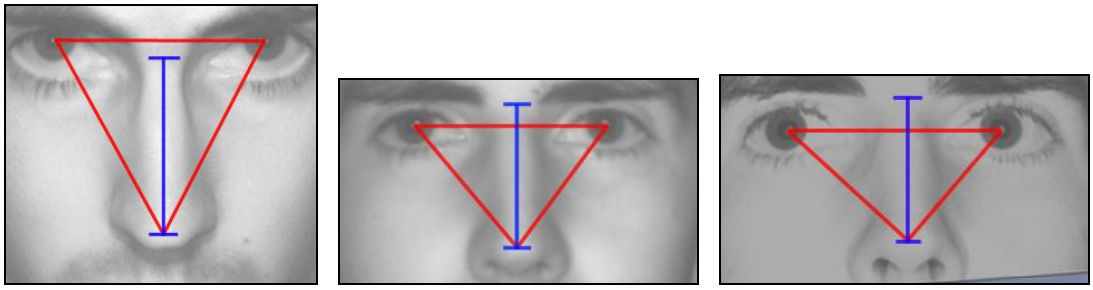


Figure 9.4 – Increasing negative face tilt with feature locations

To achieve normalisation of the face images, an algorithm has been conceived, which uses both pixel distance and height information of the nose. The two images of Figure 9.5, (a) illustrate the tilt normalisation calculation and (b) a representation of a normalised face viewed from the side.

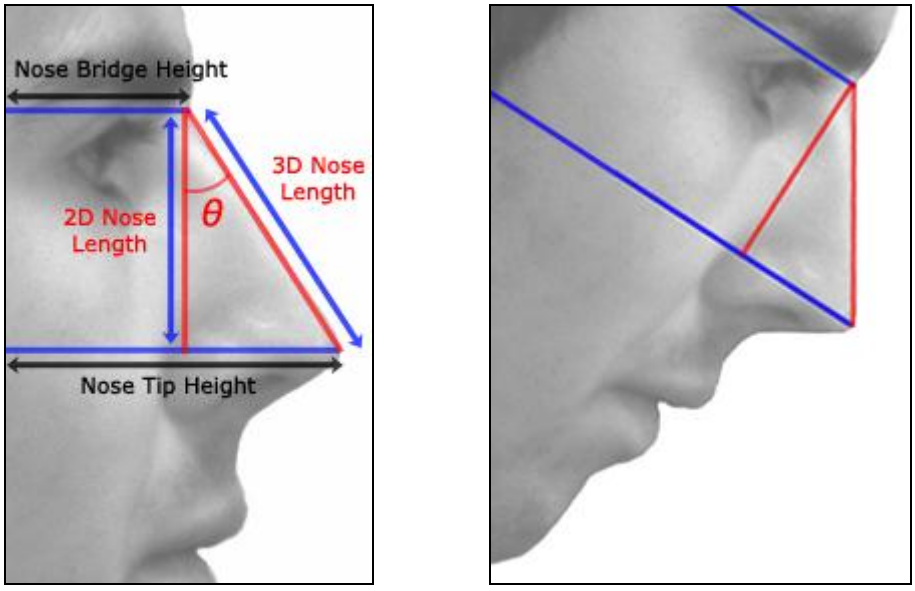


Figure 9.5 – (a) Normalisation tilt calculation (b) Normalised face

The following pseudo code details the calculation used to normalise the facial features which contain associated height information. The nose length, in this code, is described as 2D and 3D. 2D refers to the length of the nose in pixels as perceived in the 2D face images. Whereas, the 3D nose length is the actual

nose length as it would be perceived if the image had been captured from the side profile.

```
/* Calculate the nose protrusion values */
noseBridgeToTip = noseTip_z - noseBridge_z
noseBridgeToMid = noseMid_z - noseBridge_z

/* Calculate the length of the nose in 2D and 3D */
noseLength2D    = noseTip_y - noseBridge_y
noseLength3D    = GetHypotenuse(noseBridgeToTip, noseLength2D)

/* Calculate the rotation angle to normalise the face */
rotation_Angle  = GetAngle(noseBridgeToTip, noseLength2D)
```

The described calculations explain how the rotation angle is found. This angle defines the number of degrees of rotation required to normalise the face against tilt. This rotation angle could be used to manipulate the orientation of the face images, the height map, or the individual face features. In addition to calculating the rotation angle, the nose ridge length and nose height protrusion have also been found.

9.2.1 Results

Presented in this section is evidence of face tilt normalisation performed using the nose alignment algorithm.

Subject	Rotation angle (°)	2D nose length (pixels)	Nose tip protrusion (height values)	3D nose Length (height values)
A (Set 1)	66.9	156	366.5	398
B (Set 1)	62.4	106	203.01	229
C (Set 1)	67.1	101	238.8	259
D (Set 1)	75.6	91	354.6	366
E (Set 1)	61	146	262.96	301
F (Set 1)	55.3	106	152.98	186
G (Set 1)	63.6	128	257.81	288

Table 9.2 - Face tilt normalisation results

The results presented in

Table 9.2 show the normalisation results for each subject of the database. It is important to note that the 2D data is provided in pixels and the 3D data in height values. The comparison of pixels to height data is unknown. Therefore

rotation angles and 3D nose lengths are not equivalent to the 2D measurements.

9.2.2 Discussion

This experiment has found that it is possible to normalise the face against tilt, using both 2D and 3D data of the face. The results show that in addition to calculating the angle of rotation, the actual length and height protrusion of the nose can be found using height information and the perceived length of the nose in 2D. This demonstrates that the face as a whole does not need to be normalised against tilt in order to obtain the actual length of the nose if 3D data is available. Although the angle of rotation will not be required, it could prove useful in future experiments, as this would allow for the rotation of facial features other than the nose such as the forehead and chin.

The results of this experiment will be validated in the final experimental chapter, where a comparison will be made of the 3D nose protrusion lengths for different images of the same subject.

9.3 Face Scaling

The final normalisation experiment is to scale the size of the face to a dimension that all other faces can be compared with, while maintaining the aspect ratio. To accomplish this, first the maximum face dimensions in pixels must be determined. In this way normalisation of scale will only require enlarging the face image, thus there will be no possibility of lost information by reducing the size. As the captures image dimension is 640 x 480 pixels, all face images must be of this size or less. Therefore, each face image will be scaled to 640 x 480 pixels, while maintaining aspect ratios.

An algorithm has been designed and implemented that manipulates the positioning of the feature locations. By extending the feature locations, the relative positions of the eyes and nose can be transformed to fit the boundaries of the image. As the overall size of the face is unknown, the

outermost feature locations are used to define the face width and height. The following pseudo code describes the first step of the scaling algorithm. This step involves calculating the scale value that will be used to enlarge all the face features.

```
/* Use face width as the scale */
scaleValue = image_Width / faceWidth

IF (faceHeight * scaleValue) > image_Height THEN
  /* Use face height as the scale */
  scaleValue = image_Height / faceHeight
END-IF
```

As this code demonstrates, first the scale size is calculated for the width which is then compared with the scale size for the height. The largest scale is used, as this will allow the features to be enlarged to the full dimensions of the image, while keeping the aspect ratio. Following the calculation of the scale value, all the face feature locations are moved relative to eye1 x and y = 0. This shift allows for the face features to be enlarged by the scale value, while remaining within the dimensions of the image. The pseudo code which follows demonstrates how the features are enlarged by multiplying the x and y locations with the scale value.

```
noseBridge_y = noseBridge_y * scaleValue
eye2_x       = eye2_x       * scaleValue
noseBridge_x = noseBridge_x * scaleValue
noseTip_x    = noseTip_x    * scaleValue
noseTip_y    = noseTip_y    * scaleValue
```

The complete face scaling algorithm is provided in appendix F. The algorithm has been implemented in the research software and the results of face scaling, for the face database, are presented in the following section.

9.3.1 Results

Presented in this section is evidence of face scaling for handling face distance from camera.

Subject	Scaled (yes/no)	Eye1 x	Eye1 y	Eye2 x	Eye2 y	Nose BridgeX	Nose BridgeY	Nose TipX	Nose TipY
A (Set 1)	no	219	135	404	133	330	115	330	271
A (Set 1)	yes	0	58.5	569.2	58.5	335.4	0	335.4	480
B (Set 1)	no	324	230	489	236	399	218	399	324
B (Set 1)	yes	0	57.8	640	57.8	304.6	293.41	304.6	408.7
C (Set 1)	no	254	181	377	166	324	139	324	240
C (Set 1)	yes	0	161.6	584.6	161.6	294.7	0	294.7	480
D (Set 1)	no	302	246	438	241	379	220	379	311
D (Set 1)	yes	0	112.9	640	112.9	348.2	0	348.2	428.2
E (Set 1)	no	213	144	415	144	316	135	316	281
E (Set 1)	yes	0	10	609.2	10	329.2	0	329.2	480
F (Set 1)	no	263	176	421	176	333	146	333	252
F (Set 1)	yes	0	140.9	640	140.9	297.9	0	297.9	426.7
G (Set 1)	no	276	215	438	215	368	208	368	336
G (Set 1)	yes	0	15	607.5	15	326.3	0	326.3	480

Table 9.3 – Feature scaling results

The results presented in Table 9.3 show the feature scaling results for each subject of the database. The eyes and nose feature locations before and after scaling are presented.

9.3.2 Discussion

The results of this experiment have shown that by using the image dimensions, the facial feature location and sizes can be scaled appropriately. The scaling process used here, avoids the traditional approach of scaling the entire image. Currently four features are used in normalisation of the face, thus only four locations of the face are manipulated. The algorithm used for scaling has been flexibly designed to incorporate more feature locations in the future. It can be argued that this approach would always computationally outperform image scaling techniques, as the number of feature locations would always be less than the number of image pixels.

The aim of this experiment has been to scale each of the subject faces of the face database, in order to manage faces captured at different distances from the camera. The following chapter will investigate whether these scaled images and the normalised images in general, can improve the capability of matching two different images of a single subject.

10. Face Matching

This final experiment is a culmination off all of the previous experiments and research. The aim of this experiment is to investigate how the photometric stereo data can be used to match two different image sets of the same face. The matching methods that used will apply the same technique for both encoding the enrolled images in the database and extracting the information from the query image. This means that if the same image is used for a query face that exists in the enrolment database then a 100% match will be made. This differs to techniques which use training to gradually change the internal matching processes based on 'experience', resulting in systems which are difficult to predict and test (Li and Jain, 2004).

This experiment will use the measurements and localities of the facial features to determine a match between a subject query image set and the individuals enrolled on the database. Feature measurements are compared in both 2D and 3D perspectives in order to evaluate the disparity between the two methods. A number of different combinations of facial feature measurements will be tested and the results analysed in order to determine the most accurate recognition technique. A further experiment will be conducted using a double weighting for 3D comparisons. The idea of this experiment is to test if a bias towards 3D comparisons has an effect on the recognition results.

The 2D and 3D measurements that will be used for face matching are:

- 2D features:
    - Nose length (nose bridge to nose tip)
    - Eye distance (eye 1 to eye 2)
    - Eye 1 to nose tip
    - Eye 2 to nose tip
    - Eyes centre to nose tip

- 3D features:
    - Nose length (nose bridge to nose tip)

- Nose mid-point length
- Nose middle height
- Nose tip height

The 2D distance measurements are indicated on the image of Figure 10.1(a) and the 3D distance measurements on the image of Figure 10.1(b).

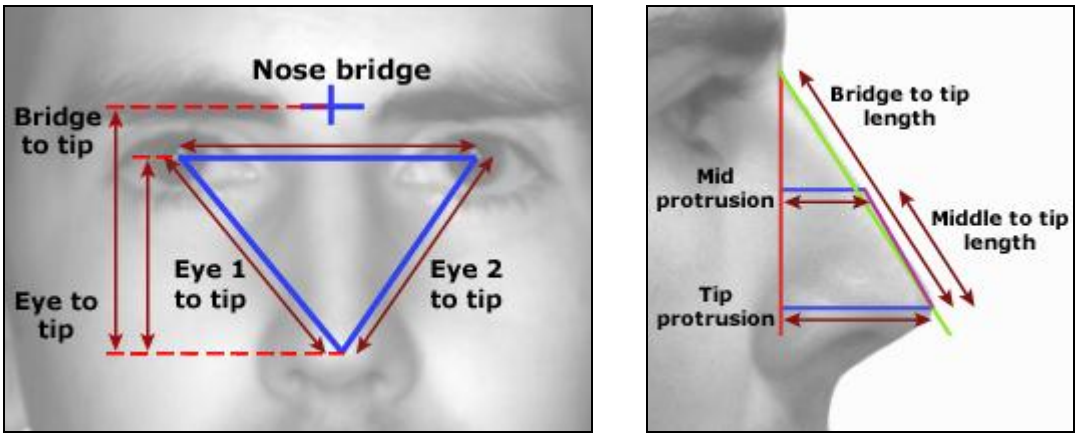


Figure 10.1 – (a) 2D feature points **(b) 3D feature points**

A query face will be selected from the database and a number of feature measurements will be compared with each of the remaining faces. The comparison will be made by calculating the percentage difference of each feature measurement. A face match percentage value will then be calculated by averaging the total difference of all feature percentages, with the number of features used for the match. The face comparison that produces the highest face match value will be presented as the matching face. The following section provides the results of the face matching experiments.

10.1 Results

This section presents the face matching results produced from 27 face images from the face database. The face database images of subjects wearing glasses and subjects with closed eyes have been excluded from the search as the eye locations have been chosen manually, which could affect the reliability of the results. For each subject the number of face images varies, which means that for each query face there is a varying number of possible correct matches. It

has been reasoned that this should not affect the results for two reasons, the first being that the same database of images will be used for each matching experiment. The second reason being that the increased probability that a correct image will be found is balanced by the number of additional searchers that are performed.

A total of four matching experiments were conducted, which use different combinations of the 2D and 3D data. The fourth experiment used a weighting for the 3D measurement comparisons, which results in the 3D measurements having twice the effect on the matching process. The results presented in Table 10.1 have been compiled from the complete set of experiment results provided in appendix J. The table of results are sectioned to show the face matching experiment process and the percentage of correct matches found using the specific feature data combination. The correct match percentage is calculated by averaging the number of correct matches with the number of face images used in the search.

Face matching experiment	Correct matches (%)
2D feature comparison	48.148
3D feature comparison	74.074
2D + 3D feature comparison	77.778
2D + 3D (double weighted) feature comparison	81.481

Table 10.1 – Face matching results

The following chapter provides a discussion of these results and identifies the face matching experiment which has produced the highest matching rate.

10.2 Discussion

This final experiment has brought together all of the research and findings of the previous experiments to produce the basis of a complete face recognition

system. The method presented in this chapter uses normalised measurements of the face in order to conduct face matching. A number of matching techniques have been tested and results collected which identify the combination of facial measurements that result in the most reliable face matching. The results show that using 3D measurements in place of 2D measurements result in a 25.9% increase in recognition. Further results show that the combination of 2D and 3D measurements increases recognition by 3.7%. The final results show that using a double weighting for 3D comparisons combined with 2D measurements, results in a recognition value of 81.48%, which is a 33.3% increase compared to using 2D measurements alone. This evidence demonstrates that 3D face measurements can be used to produce better face recognition results than 2D measurements. It is further shown that a combination of 2D and 3D measurements produce better recognition than either type separately.

As identified in the literature review, performance measures can be used to compare different face recognition systems. The report compiled by Phillips et al. (2007), presents the results of the latest face recognition vendor test (FRVT). In these tests, performance is compared based on the false acceptance rate (FAR) and the false reject rate (FRR). The FRR value is used to measure the system performance and the FAR value is used to enforce a minimum rate of 0.001 for each test. For this experiment, the best matching technique FAR can be calculated as 0.1852 (18.52 in 100) based on 5 false matches from 27 faces. From the results the FRR cannot be calculated as no reject threshold was used in the algorithm as the threshold was unknown. Using the results of this experiment it is possible to calculate a reject threshold based on the lowest match percentage of the best matching technique. This value is 85.138%, which could be used in future experiments in order to calculate the FRR for the matching algorithm. As the FRR value is unavailable it is not possible to directly compare the results of this experiment with those of the FRVT. However, the face matching results demonstrate that it is possible to use data captured from a photometric stereo system to match different images of a single individual from a database of multiple face images.

The following chapter presents a summary of the conducted experiments and identifies what has been achieved overall.

11. Overall Experimental Work Discussion

The experiments conducted in this research project have provided evidence which demonstrates the use of photometric stereo for face recognition. The experimental work has been divided into the three groups; feature extraction, face normalisation and face matching. Each group has been designed to examine the use of photometric stereo data and how it can be applied to each stage of face recognition. To perform the experiments a photometric stereo capture system and processing software were designed and implemented. The capture system was required to compile a database of different face images for testing and the processing software used to conduct the actual experiments. The design of the processing software was critical as it was required to perform all the experiments and produce testing results. The experimental stage has demonstrated that the capture system and processing software were able to produce and process accurate results for all experiments.

The first experimental group involved extracting information about a number of local features of the face. The face features identified were the eyes and nose and were selected as both are little affected by changes in facial expression and are visible even with moderate variations in pose. To locate the eyes and nose a novel technique was conceived and implemented that segmented the height profile line into a series of waves, from which each one could be measured separately. In order to measure the waves and derive information about the shape, twelve properties based on the dimensions of each wave were identified. These wave properties were used to identify waves of similar size and shape for locating both the nose bridge and the eyes. This technique of segmenting and measuring the height profile line has not been identified in any of the reviewed literature. Therefore, it seems this is a new technique for processing height data, although further research would be required to make certain. Nevertheless, for the experiments conducted in this research project, the use of height line waves has proved to be a reliable method for locating specific features such as the eyes and for following the height path of the nose tip along the nose ridge to locate the nose bridge.

The localisation of the eyes was conducted using two experiments both using different search data and location techniques. The first eye locating experiment used height data in order to locate specular reflections of the eye. This novel method used the height line wave matching technique, to locate matching waves that appear in both eyes. As with the wave technique itself, the use of measuring and matching the specular reflections of the face has not been found in any of the reviewed literature. The second eye locating experiment used gradient data plotted onto a graph in order to observe the difference in gradient distribution for the eyes and other areas of the face. This experiment demonstrated that for faces where the eyes are open, the gradient distribution is localised compared to other areas of the face. However, for faces where the eyes are closed, no correlation was found to suggest that the gradient distribution differed from the control areas of the face. This experiment has identified the potential of using gradient data for recognising a feature of the face without using reconstructed height data. This could prove useful in future research for locating other features of the face without reconstructing the height data, which is both computationally intensive and as Bronstein et al., (2004b) identifies, can introduce errors in the data.

The normalisation experiments have been designed to handle pose variations and face distance from the camera. Each experiment has used both 2D and 3D data to solve feature triangles from which distance and angles were applied to scaling and rotation. The normalisation functions implemented in the software are automatic processes that require no input from the user. The method used was designed to only normalise the five feature locations and height values, instead of the traditional approach which performs normalisation on each pixel of the face image. For the 640 x 480 images used in these experiments, this would involve the manipulation of 307,200 pixels for every change in size and orientation. Therefore, the normalisation processes developed in these experiments are computationally more efficient compared to similar image manipulation techniques.

The final experiment brought together all of the techniques developed in the previous experiment to produce an automated face matching process. The process used different images from the database as a query face, normalised the features and then compared the face with each of the database faces to locate the best match. A number of matching techniques were trialled, each using different combinations of 2D and 3D feature properties. The results of these trials identified that the least reliable matching method used only 2D feature properties and the most reliable method used both 2D and 3D feature properties. As it has been shown, it is not possible to directly compare the face matching performance with other face recognition systems. However, the evidence presented overall in the experimental work, demonstrates that it is possible to create a face recognition system that uses photometric stereo data to match different images of the same face.

12. Conclusions

It has been identified that for face recognition to be a viable solution for personal identification, it needs to be cost effective, fast and non-intrusive. It has been shown from the literature review and experimental work that photometric stereo compared to other 3D modelling techniques, requires less equipment, is computationally fast, requires little cooperation from the subject and as the results show, is able to produce accurate 3D models of the face, which can be used in face matching. The aim of this research project has been to investigate photometric stereo and its potential as a method for a low cost face recognition system. In order to ascertain the capabilities of photometric stereo, both as a 3D capture technique and as a method for face recognition, an investigation was made based on a number of objectives. These objectives, in summary, were to perform a review of the relevant face recognition and photometric stereo literature, implement a photometric stereo capture system, identify face recognition processes using practical experiments and finally to evaluate photometric stereo for use in face recognition.

This research project has achieved all of its initial aims and objectives. A literature review has been conducted, which discusses the latest research in face recognition, comparing different methods and techniques with photometric stereo. A complete photometric stereo capture system has been designed and implemented, to satisfy the low cost project aim only a simple lighting circuit and a standard imaging camera was used. To investigate the capabilities of photometric stereo for face recognition, a number of experiments were conceived that aimed to test and provide results for different aspects of the face recognition process. These experiments used a piece of software designed for this research project; the software provided a number of generalised tools that could be used to complete all necessary experiments. The experimental work as a whole resulted in an implementation for a complete face recognition system.

A discussion has been made of the overall experimental work, which analyses the photometric stereo implementation. Results have been presented that demonstrate more reliable face recognition is made when using both 2D and 3D data compared to either data types individually. Further evidence demonstrates that it is possible to use photometric stereo to correctly match a new image of a face, with a face that exists in the database. However, these results only show successful matches on a small scale. Further testing with many more individual faces would be required to further determine the system reliability. To conclude, the literature review and experimental work has provided evidence that photometric stereo has all of the properties to be considered as a low cost face recognition system.

12.1 Future Research

The photometric stereo system presented in this research project implements the fundamental components required to produce 3D face models. Subsequently, there is much scope in improving the system robustness. Two potential improvements, identified in the literature review, are using extra lights to handle shadows produced by extreme poses, and capturing an additional image with no lights on to remove the effects of ambient lighting. To test the effect of these additions and to assess the overall performance of the system it would be necessary to use a larger and more diverse enrolment database. The diversity of the database would involve a range of subject ethnicities, genders and ages coupled with multiple images of the same individual exhibiting different poses and facial expressions.

The 3D face reconstructions, produced using the research software, demonstrate that fine details of the face surface can be captured using photometric stereo. It seems plausible therefore, that the surface texture of the face, for example moles, furrows on the forehead, creases around the nose and wrinkles around the eyes, could be compared between faces much like fingerprint identification. This could be the subject of further research.

Another area of research touched upon in this project, is that of using surface gradients in order to extract shape information from the face. Experimentation has demonstrated that surface gradient has the potential to be used for locating the eyes, based on surface gradient distribution. A natural progression to the conducted gradient experiment would be to compare gradient results of the face with that of simple 3D geometric shapes. These comparisons could be used to define features of the face, such as the nose, in terms of gradient shape, which could subsequently be used as a face matching technique. Research has shown that for photometric stereo, gradient data has the potential to produce more accurate results than height data created using integration. Therefore, future research could be conducted to investigate the potential of using gradient data, without integration, to produce a more accurate face recognition system.

References

Zhao, W., Chellappa, R., 2006. Face processing: Advanced modeling and methods. Academic Press, Elsevier.

Li, S., Jain, A. K., 2005. Handbook of face recognition. Springer-Verlag, New York.

Hallinan, P. L., Gordon, G. G., Yuille, A. L., Giblin, P., Mumford, D., 1999. Two- and three-dimensional patterns of the face, A K Peters Limited, USA.

Braga, N. C., 2002. Robotics, mechatronics, and artificial intelligence. Newnes, USA.

Vogiatzis, G., Hernández, C., Cipolla, R., 2006. Reconstruction in the round using photometric normals and silhouettes. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2, pp 1847-1854.

Chandraker, M., Agarwal, S., Kriegman, D., 2007. ShadowCuts: Photometric stereo with shadows. Computer Vision and Pattern Recognition (CVPR '07), pp 1-8.

Raskar, R., Tan, R., Feris, J., Turk, M., 2004. Non-photorealistic camera: Depth edge detection and stylized rendering using multi-flash imaging. ACM Transactions on Graphics (TOG), 23(3), pp 679-688.

Ng, A. K., Schlüns, K., 1998. Towards 3D model reconstruction from photometric stereo. IVCNZ 1998, pp 309-314

Bronstein, A. M., Bronstein, M. M., Gordon, E., Kimmel, R., 2004a. Fusion of 2D and 3D data in three-dimensional face recognition. International Conference on Image Processing (ICIP), 1, pp 87-90.

Bronstein, A. M., Bronstein, M. M., Kimmel, R., Spira, A., 2004b. Face recognition from facial surface metric. Proceedings of the European Conference on Computer Vision (ECCV), pp 225-237.

Miyazaki, D., Hara, K., Ikeuchi, K., 2007. Photometric stereo beyond glass: Active separation of transparent layer and five-light photometric stereo with M-estimator using laplace distribution for a virtual museum. Proceedings of the IEEE Workshop on Photometric Analysis For Computer Vision (PACV'07, in conjunction with ICCV 2007).

Zhao, W., Chellappa, R., Phillips, P. J., Rosenfeld, A., 2003. Face recognition: A literature survey. ACM Computer Survey, 35(4), pp 399-458.

Kee, S. C., Lee, K. M., Lee, S. U., 2000. Illumination invariant face recognition using photometric stereo. IEICE, 83(7), pp 1466-1474.

Gupta, H., Roy-Chowdhury, A. K., Chellappa, R., 2004. Contour-based 3D face modeling from a monocular video. Proceedings of the British Machine Vision Conference 2004, 1, pp 136-146.

Heseltine, T., Pears, N., Austin, J., 2002. Evaluation of image pre-processing techniques for Eigenface based face recognition. Proceedings of the Second International Conference on Image and Graphics, 4875, pp 677-685.

Kovesi, P., 2005. Shapelets correlated with surface normals produce surfaces. Tenth IEEE International Conference on Computer Vision, 2, pp 994-1001.

Lenglet, C., 2003. 3D surface reconstruction by shape from shading and photometric stereo methods. Proceedings of the ninth IEEE International Conference on Computer Vision (ICCV2003), pp 1-20.

Schlüns, K., Klette, R., 1997. Local and global integration of discrete vector fields. Advances in Theoretical Computer Vision, pp 149-158.

Tangelder, J., 2005. Survey of 2D face recognition methods for robust identification by smart devices. Technical report PNA-E0504, CWI, Amsterdam.

Axnick, K., Ng, A. K., 2005. Fast face recognition. Image and Vision Computing New Zealand (IVCNZ 2005).

Bolme, D. S., Beveridge, R., Teixeira, M., Draper, B. A., 2003. The CSU face identification evaluation system: Its purpose, features, and structure. Springer-Verlag, Berlin.

Aguerrebere, C., Capdehourat, G., Delbracio, M., Mateu, M., Fernandez, A., Lecumberry, F., 2007. Aguara: An improved face recognition algorithm through Gabor filter adaptation. Automatic Identification Advanced Technologies 2007 IEEE Workshop, pp 74-79.

Mian, A. S., Bennamoun, M., Owens, R., 2005. Region-based matching for robust 3D face recognition. British Machine Vision Conference (BMVC), 1, pp 199-208.

School of computer science and software engineering at the University of Western Australia, 2006. Frankot-Chellappa surface integration algorithm implemented in Matlab [online]. Available from: http://www.csse.uwa.edu.au/~pk/research/matlabfns [Accessed 20 July 2007].

Akarun, L., Gökberk, B., Salah, A. A., 2005. 3D Face Recognition for Biometric Applications. Proceedings of the European Signal Processing Conference.

Jain, K., Ross, A., Prabhakar, S., 2004. An introduction to biometric recognition. IEEE Transactions on Circuits and Systems for Video technology, 14, pp 4-20.

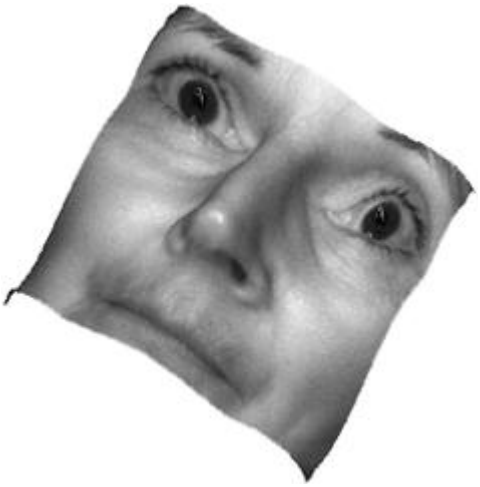
Wolff, L. B., Socolinsky, D. A., Eveland, C. K., 2001. Quantitative measurement of illumination invariance for face recognition using thermal infrared imagery. In Proceedings of the IEEE Workshop on Computer Vision Beyond the Visible Spectrum.

Agrawal, A., Chellappa, R., Raskar, R., 2005. An Algebraic Approach to Surface Reconstruction from Gradient Fields. IEEE International Conference on Computer Vision (ICCV), 1, pp 174-181.

Appendix A: Reconstructed Faces

Subject A

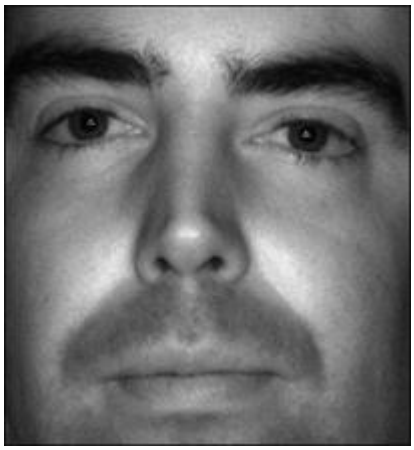
Subject B



Subject C

Subject D

Subject F



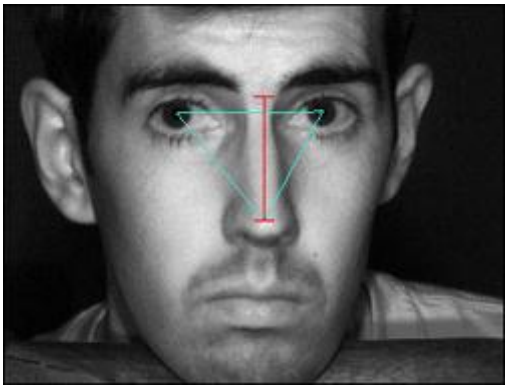
Subject G

Appendix B: Face Database

This document presents the face images and details contained in the enrolment database, made up of 35 image sets, consisting of 105 images, of 7 different individual faces.

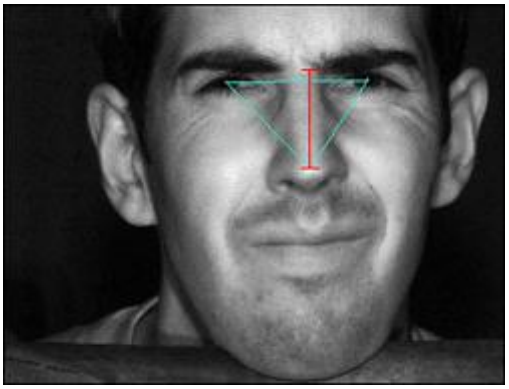
Note: For images where eyes are closed or glasses are worn eye locations have been selected manually.

Subject A: Set 1



eye1 x	219	**nose bridge x**	330
eye1 y	135	**nose bridge y**	115
eye2 x	404	**nose bridge z**	1191.1
eye2 y	133	**nose mid x**	330
		nose mid y	193
		nose mid z	1309.3
		nose tip x	330
		nose tip y	271
		nose tip z	1557.6

Subject A: Set 2



eye1 x	282	**nose bridge x**	389
eye1 y	96	**nose bridge y**	81
eye2 x	462	**nose bridge z**	768.32
eye2 y	93	**nose mid x**	389
		nose mid y	143
		nose mid z	849.62
		nose tip x	389
		nose tip y	205
		nose tip z	1007.5

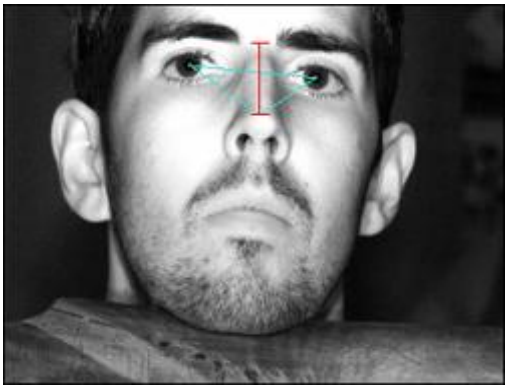
Subject A: Set 3

eye1 x	339	**nose bridge x**	442
eye1 y	176	**nose bridge y**	199
eye2 x	525	**nose bridge z**	842.93
eye2 y	172	**nose mid x**	442
		nose mid y	279
		nose mid z	1029.6
		nose tip x	442
		nose tip y	359
		nose tip z	1241.5

Subject A: Set 4

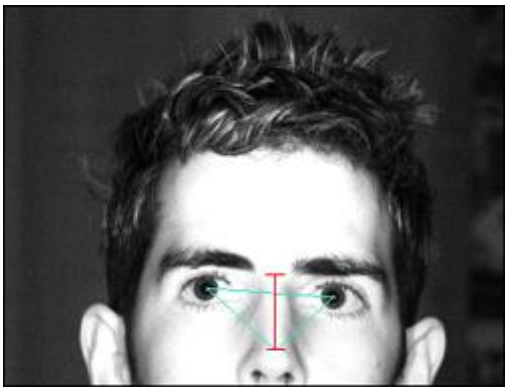
eye1 x	337	**nose bridge x**	436
eye1 y	174	**nose bridge y**	188
eye2 x	522	**nose bridge z**	742.83
eye2 y	167	**nose mid x**	436
		nose mid y	271
		nose mid z	999.52
		nose tip x	436
		nose tip y	354
		nose tip z	1380.7

Subject A: Set 5



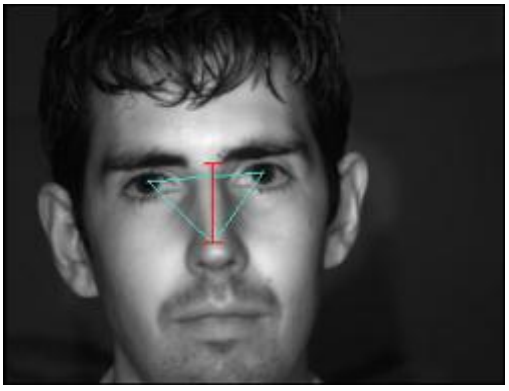
eye1 x	236	**nose bridge x**	325
eye1 y	75	**nose bridge y**	47
eye2 x	396	**nose bridge z**	228.39
eye2 y	90	**nose mid x**	325
		nose mid y	92
		nose mid z	385.64
		nose tip x	325
		nose tip y	137
		nose tip z	477.28

Subject A: Set 6



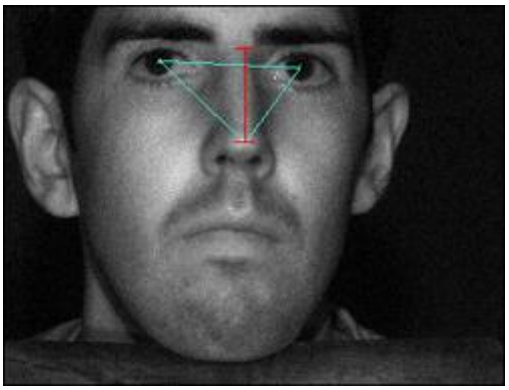
eye1 x	257	**nose bridge x**	345
eye1 y	353	**nose bridge y**	336
eye2 x	418	**nose bridge z**	207.38
eye2 y	365	**nose mid x**	345
		nose mid y	383
		nose mid z	279.89
		nose tip x	345
		nose tip y	431
		nose tip z	349.2

Subject A: Set 7



eye1 x	182	**nose bridge x**	265
eye1 y	221	**nose bridge y**	199
eye2 x	329	**nose bridge z**	818
eye2 y	209	**nose mid x**	265
		nose mid y	249
		nose mid z	978.92
		nose tip x	265
		nose tip y	299
		nose tip z	1057.9

Subject A: Set 8

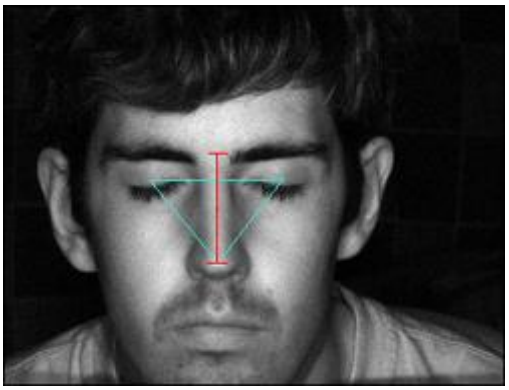


eye1 x	197	**nose bridge x**	306
eye1 y	68	**nose bridge y**	52
eye2 x	376	**nose bridge z**	1014.4
eye2 y	76	**nose mid x**	306
		nose mid y	111
		nose mid z	1064.3
		nose tip x	306
		nose tip y	171
		nose tip z	1251.2

Subject A: Set 9

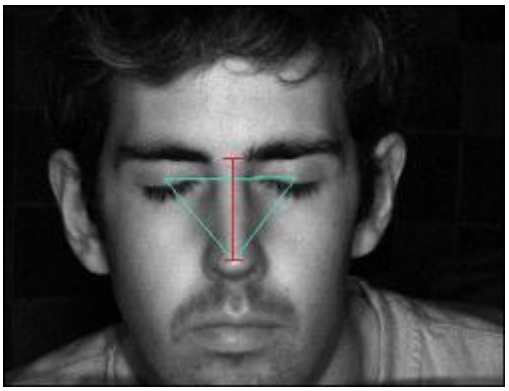
eye1 x	279	**nose bridge x**	336
eye1 y	206	**nose bridge y**	193
eye2 x	384	**nose bridge z**	1669.5
eye2 y	210	**nose mid x**	336
		nose mid y	238
		nose mid z	1888.1
		nose tip x	336
		nose tip y	284
		nose tip z	2168.1

Subject A: Set 10



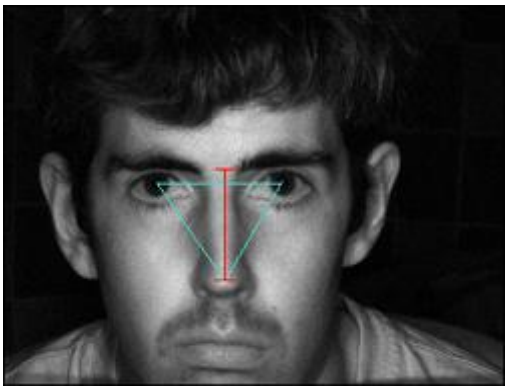
eye1 x	185	**nose bridge x**	271
eye1 y	219	**nose bridge y**	185
eye2 x	354	**nose bridge z**	703.15
eye2 y	217	**nose mid x**	271
		nose mid y	253
		nose mid z	832.97
		nose tip x	271
		nose tip y	322
		nose tip z	992.78

Subject A: Set 11



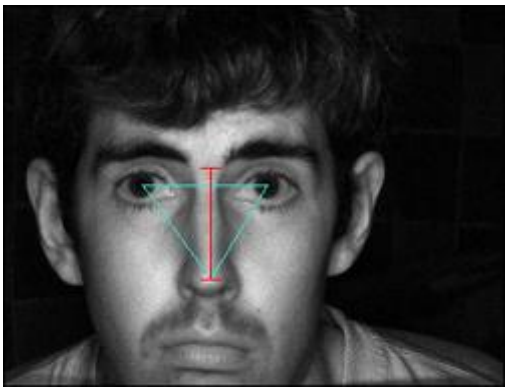
eye1 x	204	**nose bridge x**	291
eye1 y	217	**nose bridge y**	190
eye2 x	371	**nose bridge z**	851.67
eye2 y	214	**nose mid x**	291
		nose mid y	254
		nose mid z	985.38
		nose tip x	291
		nose tip y	319
		nose tip z	1070.5

Subject A: Set 12



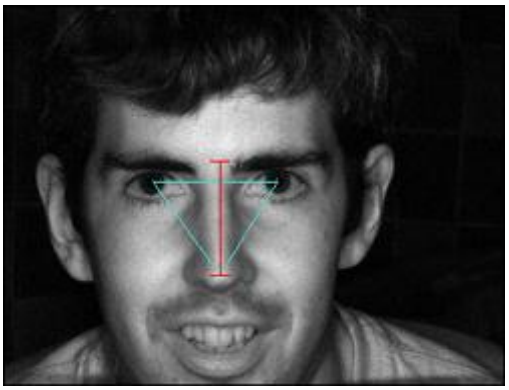
eye1 x	194	**nose bridge x**	280
eye1 y	222	**nose bridge y**	204
eye2 x	352	**nose bridge z**	1036.6
eye2 y	222	**nose mid x**	280
		nose mid y	274
		nose mid z	1194.8
		nose tip x	280
		nose tip y	345
		nose tip z	1329.3

Subject A: Set 13



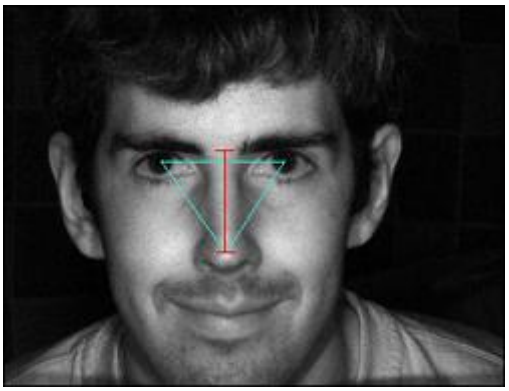
eye1 x	176	**nose bridge x**	262
eye1 y	224	**nose bridge y**	203
eye2 x	334	**nose bridge z**	771.09
eye2 y	224	**nose mid x**	262
		nose mid y	272
		nose mid z	886.8
		nose tip x	262
		nose tip y	342
		nose tip z	997.53

Subject A: Set 14



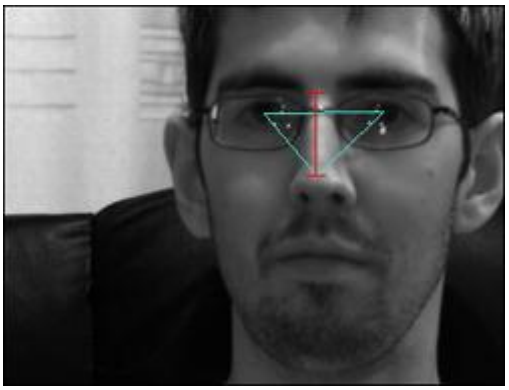
eye1 x	189	**nose bridge x**	274
eye1 y	221	**nose bridge y**	195
eye2 x	347	**nose bridge z**	678.32
eye2 y	220	**nose mid x**	274
		nose mid y	266
		nose mid z	802.11
		nose tip x	274
		nose tip y	338
		nose tip z	916.44

Subject A: Set 15



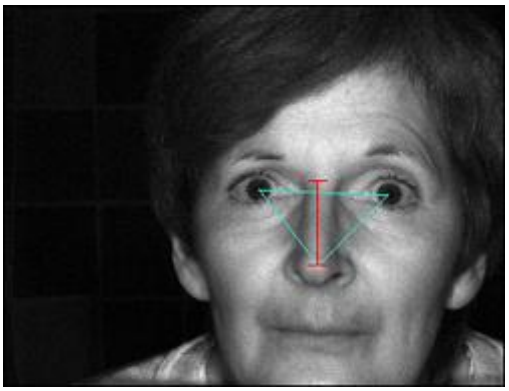
eye1 x	200	**nose bridge x**	281
eye1 y	195	**nose bridge y**	180
eye2 x	356	**nose bridge z**	273.87
eye2 y	194	**nose mid x**	281
		nose mid y	244
		nose mid z	399.75
		nose tip x	281
		nose tip y	309
		nose tip z	460.36

Subject A: Set 16



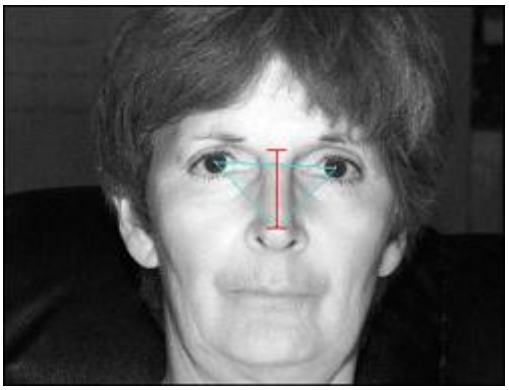
eye1 x	330	**nose bridge x**	397
eye1 y	134	**nose bridge y**	108
eye2 x	483	**nose bridge z**	158.95
eye2 y	132	**nose mid x**	397
		nose mid y	160
		nose mid z	167.29
		nose tip x	397
		nose tip y	212
		nose tip z	156.53

Subject B: Set 1



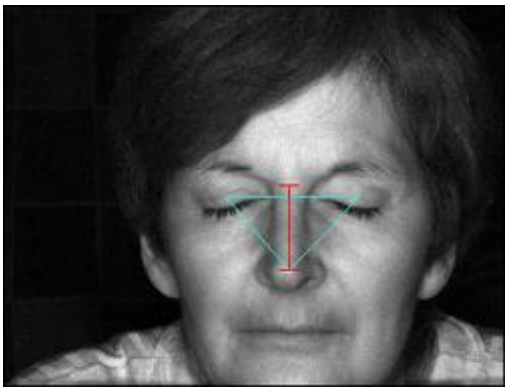
eye1 x	324	**nose bridge x**	399
eye1 y	230	**nose bridge y**	218
eye2 x	489	**nose bridge z**	293.41
eye2 y	236	**nose mid x**	399
		nose mid y	271
		nose mid z	407.66
		nose tip x	399
		nose tip y	324
		nose tip z	496.42

Subject B: Set 2



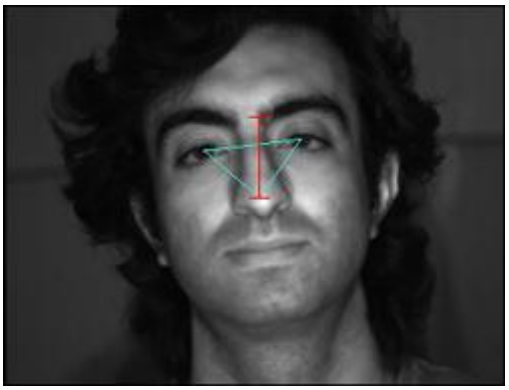
eye1 x	268	**nose bridge x**	346
eye1 y	195	**nose bridge y**	181
eye2 x	421	**nose bridge z**	397.03
eye2 y	202	**nose mid x**	346
		nose mid y	230
		nose mid z	509.13
		nose tip x	346
		nose tip y	279
		nose tip z	602.12

Subject B: Set 3



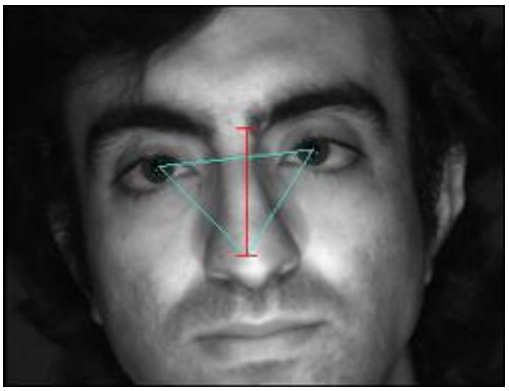
eye1 x	282	**nose bridge x**	362
eye1 y	236	**nose bridge y**	224
eye2 x	452	**nose bridge z**	756.42
eye2 y	238	**nose mid x**	362
		nose mid y	277
		nose mid z	873.15
		nose tip x	362
		nose tip y	331
		nose tip z	969.45

Subject C: Set 1



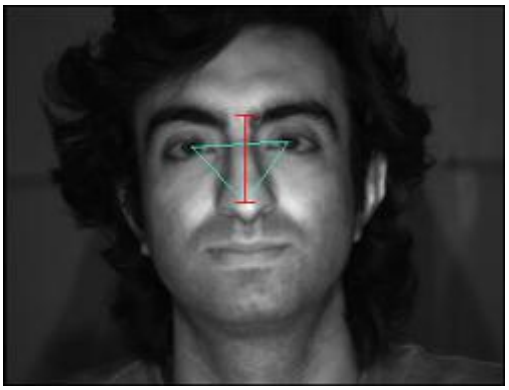
eye1 x	254	**nose bridge x**	324
eye1 y	181	**nose bridge y**	139
eye2 x	377	**nose bridge z**	1654.2
eye2 y	166	**nose mid x**	324
		nose mid y	189
		nose mid z	1800.5
		nose tip x	324
		nose tip y	240
		nose tip z	1893

Subject C: Set 2



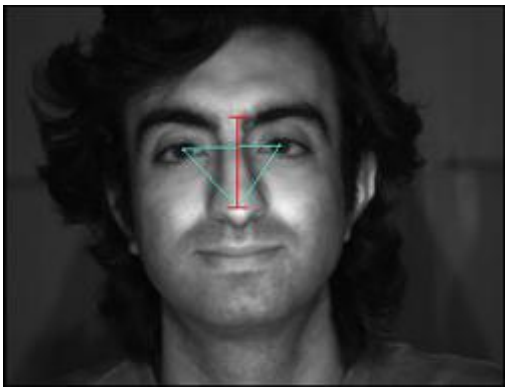
eye1 x	196	**nose bridge x**	308
eye1 y	200	**nose bridge y**	152
eye2 x	392	**nose bridge z**	2347.9
eye2 y	179	**nose mid x**	308
		nose mid y	232
		nose mid z	2993.9
		nose tip x	308
		nose tip y	312
		nose tip z	3332.6

Subject C: Set 3



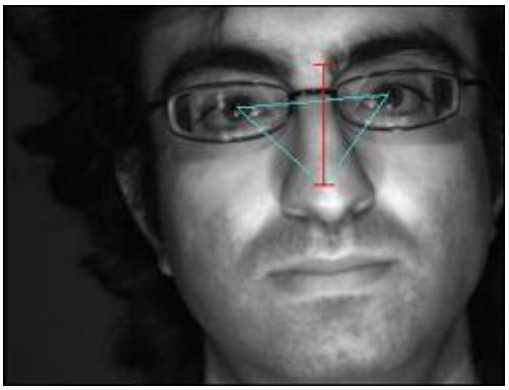
eye1 x	239	**nose bridge x**	306
eye1 y	176	**nose bridge y**	137
eye2 x	361	**nose bridge z**	700.86
eye2 y	169	**nose mid x**	306
		nose mid y	192
		nose mid z	861.16
		nose tip x	306
		nose tip y	247
		nose tip z	966.03

Subject C: Set 4



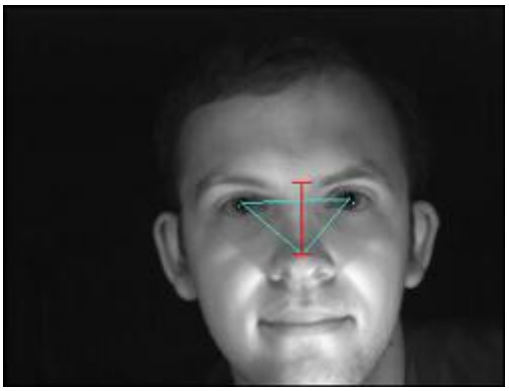
eye1 x	227	**nose bridge x**	297
eye1 y	178	**nose bridge y**	139
eye2 x	350	**nose bridge z**	886.69
eye2 y	173	**nose mid x**	297
		nose mid y	195
		nose mid z	1061.3
		nose tip x	297
		nose tip y	252
		nose tip z	1188.5

Subject C: Set 5



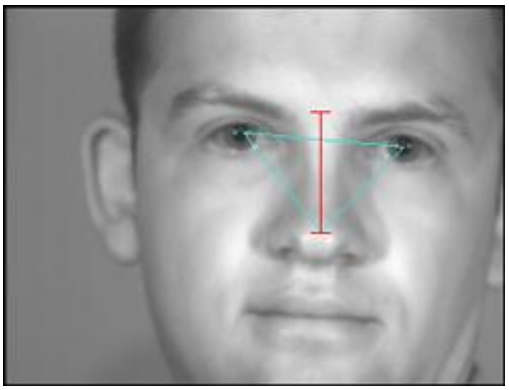
eye1 x	296	**nose bridge x**	406
eye1 y	127	**nose bridge y**	72
eye2 x	488	**nose bridge z**	1876
eye2 y	110	**nose mid x**	406
		nose mid y	148
		nose mid z	2333.6
		nose tip x	406
		nose tip y	225
		nose tip z	2698.2

Subject D: Set 1



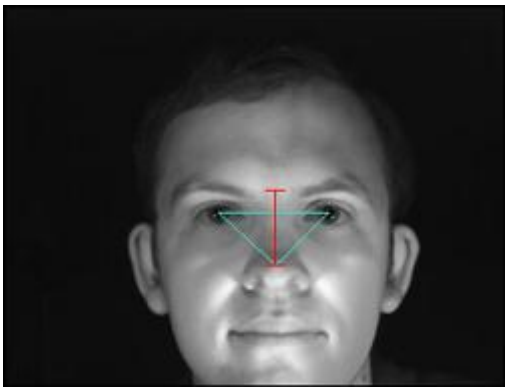
eye1 x	302	**nose bridge x**	379
eye1 y	246	**nose bridge y**	220
eye2 x	438	**nose bridge z**	1466.3
eye2 y	241	**nose mid x**	379
		nose mid y	265
		nose mid z	1691.4
		nose tip x	379
		nose tip y	311
		nose tip z	1820.9

Subject D: Set 2



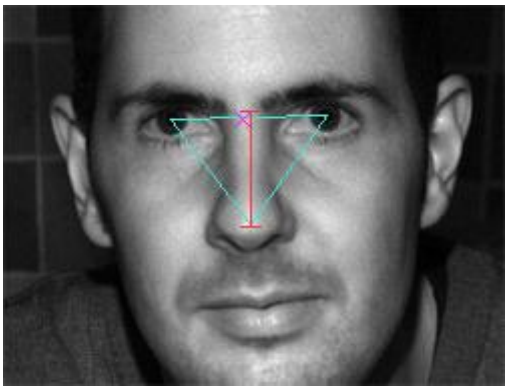
eye1 x	304	**nose bridge x**	403
eye1 y	158	**nose bridge y**	132
eye2 x	508	**nose bridge z**	1360.3
eye2 y	176	**nose mid x**	403
		nose mid y	208
		nose mid z	1635.5
		nose tip x	403
		nose tip y	284
		nose tip z	1676.9

Subject D: Set 3



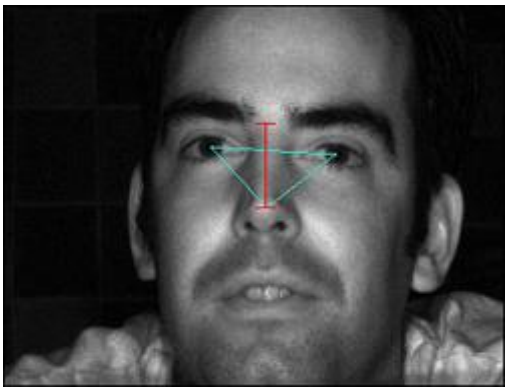
eye1 x	272	**nose bridge x**	345
eye1 y	258	**nose bridge y**	230
eye2 x	413	**nose bridge z**	1386.5
eye2 y	258	**nose mid x**	345
		nose mid y	277
		nose mid z	1622.2
		nose tip x	345
		nose tip y	324
		nose tip z	1689.2

Subject E: Set 1



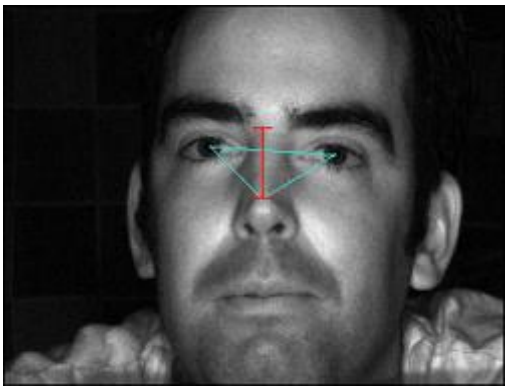
eye1 x	213	**nose bridge x**	316
eye1 y	144	**nose bridge y**	135
eye2 x	415	**nose bridge z**	846.34
eye2 y	140	**nose mid x**	316
		nose mid y	208
		nose mid z	985.19
		nose tip x	316
		nose tip y	281
		nose tip z	1109.3

Subject F: Set 1



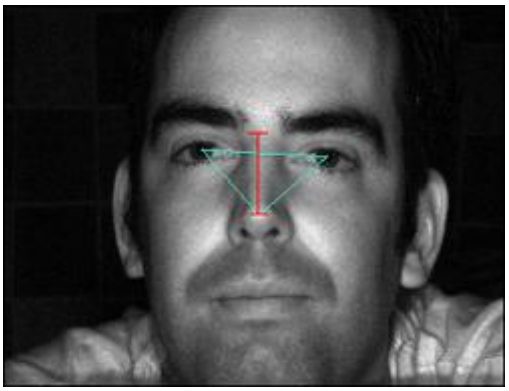
eye1 x	263	**nose bridge x**	333
eye1 y	176	**nose bridge y**	146
eye2 x	421	**nose bridge z**	517.25
eye2 y	185	**nose mid x**	333
		nose mid y	199
		nose mid z	603.05
		nose tip x	333
		nose tip y	252
		nose tip z	670.23

Subject F: Set 2



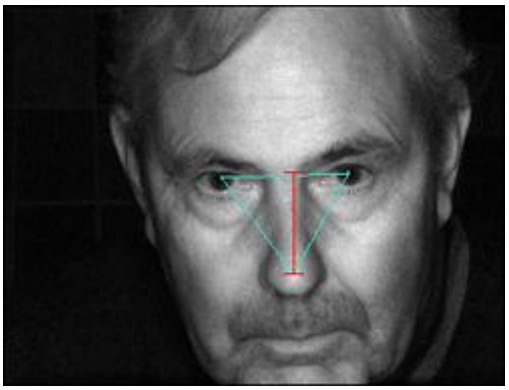
eye1 x	262	**nose bridge x**	328
eye1 y	176	**nose bridge y**	152
eye2 x	420	**nose bridge z**	454.20
eye2 y	186	**nose mid x**	328
		nose mid y	196
		nose mid z	507.30
		nose tip x	328
		nose tip y	241
		nose tip z	599.65

Subject F: Set 3



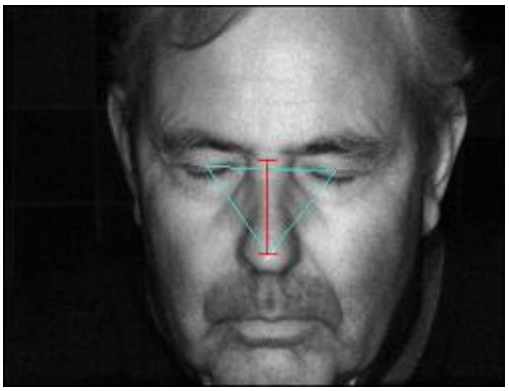
eye1 x	251	**nose bridge x**	323
eye1 y	179	**nose bridge y**	158
eye2 x	410	**nose bridge z**	910.57
eye2 y	188	**nose mid x**	323
		nose mid y	209
		nose mid z	996.01
		nose tip x	323
		nose tip y	261
		nose tip z	1066.7

Subject G: Set 1



eye1 x	276	**nose bridge x**	368
eye1 y	215	**nose bridge y**	208
eye2 x	438	**nose bridge z**	328.95
eye2 y	209	**nose mid x**	368
		nose mid y	272
		nose mid z	433.40
		nose tip x	368
		nose tip y	336
		nose tip z	586.76

Subject G: Set 2



eye1 x	257	**nose bridge x**	335
eye1 y	200	**nose bridge y**	192
eye2 x	421	**nose bridge z**	677.45
eye2 y	206	**nose mid x**	335
		nose mid y	251
		nose mid z	788.37
		nose tip x	335
		nose tip y	311
		nose tip z	885.45

Subject Z: Set 1

eye1 x	188	**nose bridge x**	223
eye1 y	216	**nose bridge y**	206
eye2 x	331	**nose bridge z**	2119.9
eye2 y	223	**nose mid x**	223
		nose mid y	256
		nose mid z	2018.2
		nose tip x	223
		nose tip y	306
		nose tip z	1970.9

Subject Z: Set 2

eye1 x	252	**nose bridge x**	314
eye1 y	261	**nose bridge y**	259
eye2 x	368	**nose bridge z**	1525.8
eye2 y	265	**nose mid x**	314
		nose mid y	301
		nose mid z	1868.5
		nose tip x	314
		nose tip y	343
		nose tip z	2483.7

Appendix C: Hardware Specification

The ImagingSource

DMK 21BF04 - Specification

General behaviour:	
Sensitivity	0.5 lx at 1/30s, gain 20 dB
Interface (optical):	
Type	progressive scan
Format	1/4 "
Resolution	H: 640, V: 480
Pixel size	H: 5.6 µm, V: 5.6 µm
Lens mount	C/CS mount
Adjustments (man):	
Shutter	1/10000 to 30 s
Gain	0 to 36 dB

http://theimagingsource.com/en/products/cameras/firewire_mono

Kingbright

52mm LED CLUSTER

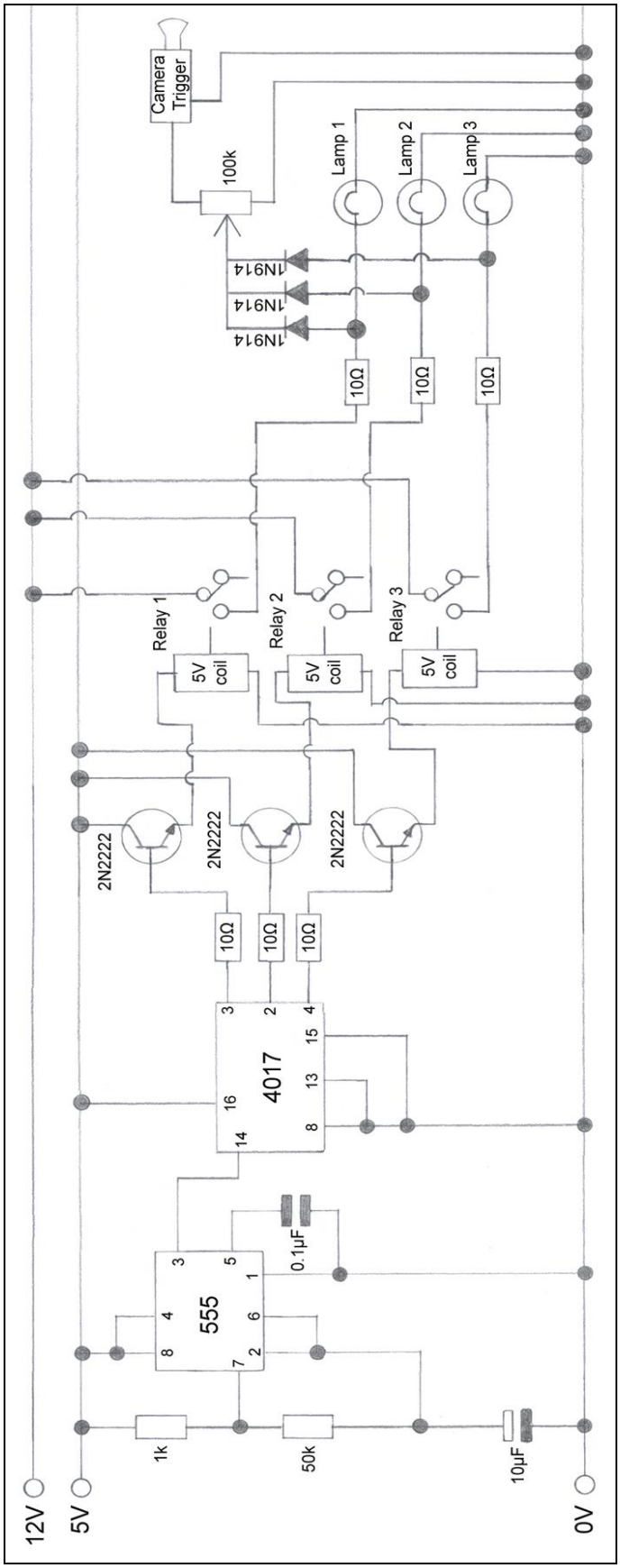
Selection Guide

Parameter	Super Bright Red	Units
Power dissipation	2.5	W
DC Forward Current	250	mA

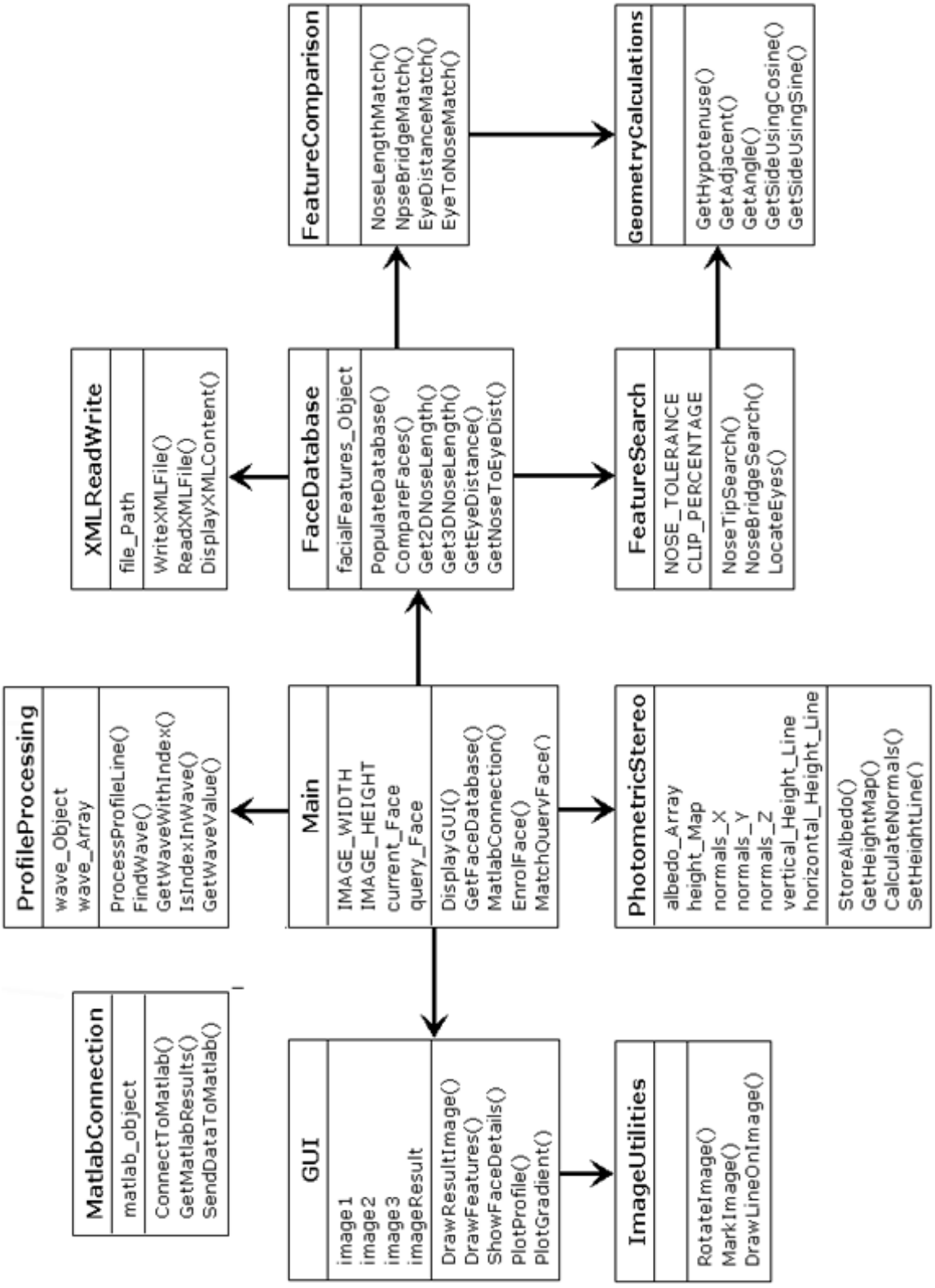
Symbol	Parameter	Device	Min.	Typ.	Max.	Units	Test Conditions
Iv	Luminous Intensity	BL0307-50-44	16000	20000	-	mcd	IF=200mA
θ	Viewing Angle		-	40	-	Deg	-
V_F	Forword Voltage		-	9.25	12.5	V	IF=200mA
λpeak	Peak Wavelength		-	660	-	nm	IF=200mA
λD	Dominate Wavelength		-	640	-	-	IF=200mA
Δλ1/2	Spectral Line Halfwidth		-	20	-	nm	IF=200mA
I_R	Reverse Current		-	-	100	uA	VR = 5V

http://www.kingbright.com/product_main.php

Appendix D: Capturing Circuit Diagram



Appendix E: Software Class Diagram



Appendix F: Software Function Pseudo Code

FeatureMatching(*query_Face*)

```
SET best_Match    =   0

/* Calculate facial features location and sizes (query_Face) */
NoseTipSearch()
NoseBridgeSearch()
LocateEyes(query_Face_eye1)

/* Normalise face */
PoseNormalisation(query_Face)
ScaleNormalisation(query_Face)

/* Search for and store the best face match in the database */
FOR ALL entries in the face_Database
    STORE best_Match(current_Face, query_Face)
END-FOR

/* Display the best match */
Display best_Match_Image
Display best_Match_Percentage
```

NoseTipSearch()

```
SET wave = 0

/* Get largest height values and set global variables */
FindLargestHeightValue()
/* Set the profile line to */
STORE horizontal profile_Line at largest_Height_Value_Y

/* Loop through all the waves in the profile line */
FOR ALL waves in horizontal profile_Line
    /* Store the wave, which contains the largest height value x */
    STORE wave containing largest_Height_Value_X
END-FOR

/* Loop through all the height values in wave and remove top 10% */
FOR ALL values in wave
    /* If a value is within 10% of the largest value */
    IF wave_Height_Value >= (largest_Height_Value – 10%) THEN
        /* Cap the wave value to 10% off the largest value */
        wave_Value = largest_Height_Value – 10%
    END-IF
END-FOR

/* Set facial feature locations */
noseTip_X   =   wave_Largest_Middle_Value_Index
noseTip_Y   =   largest_Height_Value_Y
noseTip_Z   =   largest_Height_Value
```

NoseBridgeSearch()

```
SET temp_Wave                    = 0
SET temp_Highest_Value    = 0
/* Trend iteration, used to skip height anomalies */
SET trend_Iteration        = 0
SET TREND_MIN                    = 5
/* While loop count index */
SET current_Y                  = noseTip_Y

/* Set the profile line */
STORE horizontal profile_Line at largest_Height_Value_Y

/* Get largest height values and set global variables */
FindLargestHeightValue()

/* Loop through all the waves in the profile line */
FOR ALL waves in horizontal profile_Line
    /* Store the wave, which contains the largest height value x */
    STORE wave containing largest_Height_Value_X
END-FOR

/* Loop: Start from noseTip_Y and decrement until 0 is reached */
WHILE current_Y >= 0
    /* Set the profile line to the current_Y index */
    STORE horizontal profile_Line at current_Y
    /* Get the wave on the current profile line at the noseTip x */
    temp_Wave = wave containing noseTip_X

    /* If previous is less than the highest then check for a trend */
    IF temp_Highest_Value  < temp_Wave THEN

        /* If trend iteration has reached the trend constant then
             save the location of the nose bridge/mid and exit */
        IF trend_Iteration >= TREND_MIN THEN
            /* Set facial feature locations */
            noseBridge_X =     middle index of temp_Highest_Value
            noseBridge_Y =     trend_StartIndex
            noseBridge_Z =     current_Y - TREND_MIN
            noseMid_X    =  noseBridge_X
            noseMid_Y    = noseBridge_Y + ((noseBridge_Y – noseTip_Y) / 2)
            noseMid_Z    = height_Map(noseMid_X, noseMid_Y)
            /* Exit the WHILE loop */
            BREAK
   END-IF
        /* Increment the decreasing height trend count */
        trend_Iteration ++

    /* If trend has stopped then reset the trend iteration count */
    ELSE IF trend_Iteration > 0 THEN
        trend_Iteration = 0
    END-IF

    /* Set the highest value to that of the current wave */
    temp_Highest_Value = temp_Wave
    /* Decrement the while current_Y index */
    current_Y --
END-WHILE
```

LocateEyes(*eye_Query, noseBridge_tolerance*)

```
/* Set search Y as a small amount above and below the nose bridge */
SET startSearch_Y =  noseBridge_Y – noseBridge_tolerance /* 60 */
SET endSearch_Y     =  noseBridge_Y + noseBridge_tolerance /* 60 */
SET startSearch_X=  0
SET endSearch_X     =  0

/* Determine if the query eye is the left or right */
IF eye_Query_X < noseBridge_X THEN
    /* Query eye1: set search right of nose bridge */
    startSearch_X      =  noseBridge_X
    endSearch_X          =  IMAGE_WIDTH
ELSE
    /* Query eye2: set search left of nose bridge */
    endSearch_X          =  noseBridge_X
END-IF

/* Loop through all waves between startSearch and endSearch */
FOR ALL waves between startSearch(x,y) AND endSearch(x,y)
    /* Store the wave, which best matches the eye_Query */
    STORE best_Match(current_Wave, eye_Query)
END-FOR

/* Set facial feature locations */
eye1_X   =  eye_Query_X
eye1_Y   =  eye_Query_Y
eye2_X   =  best_Match_X
eye2_Y   =  best_Match_Y
```

FindLargestHeightValue()

```
SET largest_Height_Value       = 0
SET largest_Height_Value_X   = 0
SET largest_Height_Value_Y   = 0

/* Loop through all the values in the height map */
FOR ALL values in height_Map

    /* If the current value is larger then store it */
    IF current_Height_Value > largest_Height_Value THEN
        largest_Height_Value          = current_Height_Value
        largest_Height_Value_X = current_X
        largest_Height_Value_Y = current_Y
    END-IF
END-FOR
```

PlotGradient()

```
SET normal_X = 0
SET normal_Y = 0

/* Create graph with the x and y scale to fit the surface normals */
CREATE Graph to accommodate min/max surface_Normals

/* Loop through all the surface normals from the current image set */
FOR ALL surface_Normals
```

```
    /* Store the surface normal x and y components */
    normal_X = GetXSurfaceNormal(x_Index, y_Index)
    normal_Y = GetYSurfaceNormal(x_Index, y_Index)

    /* Put a mark on the graph at the normal x and y position */
    MarkOnGraph(normal_X, normal_Y)
END-FOR

/* Show the graph on screen */
Display Graph
```

ScaleNormalisation(*face*)

```
SET faceWidth      = eye2_X – eye1_X
SET faceHeight     = 0
SET scaleValue     = 0

/* Determine if the eyes are above or below the nose bridge */
IF noseBridge_Y > eye1_Y THEN
    faceHeight = noseTip_Y – eye1_Y
ELSE
    faceHeight = noseTip_Y – noseBridge_Y
END-IF

/* Use face width as the scale */
scaleValue = image_Width / faceWidth

IF (faceHeight * scaleValue) > image_Height THEN
    /* Use face height as the scale */
    scaleValue = image_Height / faceHeight
END-IF

/* Move face features to the left */
eye2_X             = eye2_X             – eye1_X
noseBridge_X   = noseBridge_X   – eye1_X
noseTip_X      = noseTip_X          – eye1_X

/* Set the left most feature to 0 */
eye1_X = 0

IF eyeToNoseTip > noseLength THEN
    /* Move face features to the top */
    noseTip_Y      = noseTip_Y      – eye1_Y
    noseBridge_Y   = noseBridge_Y   – eye1_Y
    eye1_Y             = 0
    eye2_Y             = 0
    noseBridge_Y = noseBridge_Y * scaleValue
ELSE
    /* Move face features to the top */
    eye1_Y             = eye1_y     – noseBridge_Y
    eye2_Y             = eye2_y         – noseBridge_Y
    noseTip_Y      = noseTip_Y – noseBridge_Y
    noseBridge_Y   = 0
    eye1_Y             = eye1_Y * scaleValue
    eye2_Y             = eye2_Y * scaleValue
END-IF

/* Set facial feature locations */
eye2_X             = eye2_X * scaleValue
```

```
noseBridge_X   = noseBridge_X * scaleValue
noseTip_X      = noseTip_X * scaleValue
noseTip_Y      = noseTip_Y * scaleValue
```

Appendix G: Frankot-Chellappa Surface Integration Algorithm

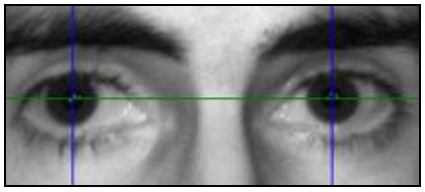
```
% FRANKOTCHELLAPPA  - Generates integrable surface from gradients
%
% An implementation of Frankot and Chellappa'a algorithm for
% constructing an integrable surface from gradient information.

% Usage:     z = frankotchellappa(dzdx,dzdy,pz)
%
% Arguments:  dzdx,  -  2D matrices specifying a grid of gradients
%           dzdy, pz  with respect to x, y and z.
% Returns:    z     - Inferred surface heights.

% Copyright (c) 2004 Peter Kovesi
% School of Computer Science & Software Engineering
% The University of Western Australia
% http://www.csse.uwa.edu.au/

% Permission is hereby granted, free of charge, to any person
% obtaining a copy of this software and associated documentation
% files (the "Software"), to deal in the Software without
% restriction, subject to the following conditions:
% The above copyright notice and this permission notice shall be
% included in all copies or substantial portions of the Software.
%
% The Software is provided "as is", without warranty of any kind.
% October 2004

function z = fcint(dzdx,dzdy,pz)
% Z = fcint(Px,Py,Pz)
% Integrate surface nmormals, Px, Py, Pz, using the Frankot-
% Chellappa surface integration algorithm to recover height, Z

if ~all(size(dzdx) == size(dzdy))
   error('Gradient matrices must match');
end

[rows,cols] = size(pz);

% scale normal vectors to derivative format
for n = 1:rows
  for x = 1:cols
    % Handle zero values
    if (pz(n,x) == 0)
          dzdx(n,x)     = -0.00000000001;
          dzdy(n,x)     = -0.00000000001;
    else
      dzdx(n,x)      = -dzdx(n,x) / pz(n,x);
      dzdy(n,x)= -dzdy(n,x) / pz(n,x);
    end
  end
end

[wx, wy] = meshgrid(-pi/2:pi/(cols-1):pi/2,-pi/2:pi/(rows-1):pi/2);

% Quadrant shift to put zero frequency at the appropriate edge
```

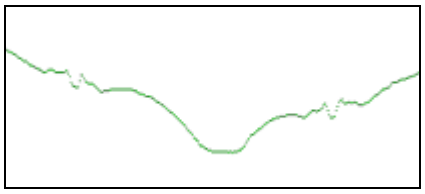
```
wx = ifftshift(wx); wy = ifftshift(wy);

% Fourier transforms of gradients
DZDX = fft2(dzdx);
DZDY = fft2(dzdy);

% Integrate in the frequency domain by phase shifting by pi/2 and
% weighting the Fourier coefficients by their frequencies in x and y
% and then dividing by the squared frequency.  eps is added to the
% denominator to avoid division by 0.
j = sqrt(-1);

dd = wx.^2 + wy.^2;
Z = (-j*wx.*DZDX -j*wy.*DZDY)./(wx.^2 + wy.^2 + eps);

z = real(ifft2(Z));  % Reconstruction

minNumber = GetMinNumber(z);
minNumber = minNumber * -1;

for n = 1:rows
  for x = 1:cols
    z(n,x) = z(n,x) + minNumber;
  end
end

z = z/2;
```

Appendix H: Eye Location Results

This document presents the results of the eye locating experiment, produced using the face database and the research project software. The results show the position of the eyes, the profile height line and a comparison of each eye wave. The image of the eyes has been rotated using the software in order to display both eye artifacts in a single profile line. The left eye of the image is identified as 'eye 1' and the right eye as 'eye 2'.

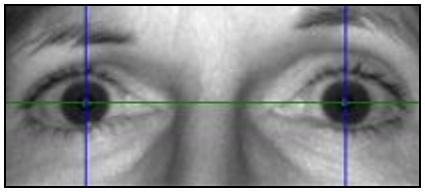
Subject A: Set 1

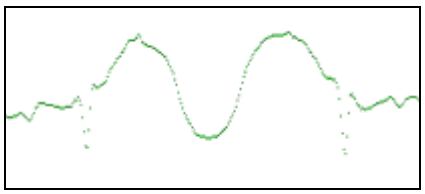




Wave attribute	Eye 1	Eye 2	Match %
Edge difference	-7.589	-10.07	75.385
Start to highest value	111.9	74.67	66.717
Middle to highest value	12.76	14.37	88.737
End to highest value	104.3	64.61	61.922
Start to middle value	99.17	60.3	60.803
End to middle value	91.58	50.23	54.849
Width start to end	12	11	91.666
Width start to middle	6	5	83.333
Width start to highest	7	4	57.143
Width middle to highest	1	1	100
Width middle to end	6	6	100
Width highest to end	5	7	71.429

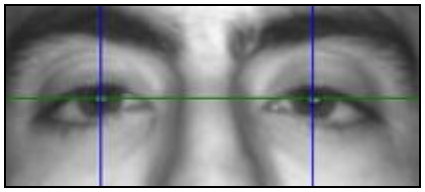
Subject B: Set 1

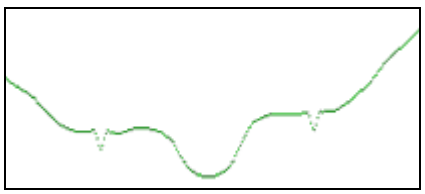




Wave attribute	Eye 1	Eye 2	Match %
Edge difference	28.06	-36.64	0
Start to highest value	113.7	88.0	77.407
Middle to highest value	25.94	9.545	36.796
End to highest value	141.8	51.36	36.23
Start to middle value	87.75	78.46	89.41
End to middle value	115.8	41.81	36.103
Width start to end	10	11	90.909
Width start to middle	5	5	100
Width start to highest	4	4	100
Width middle to highest	1	1	100
Width middle to end	5	6	83.333
Width highest to end	6	7	85.714

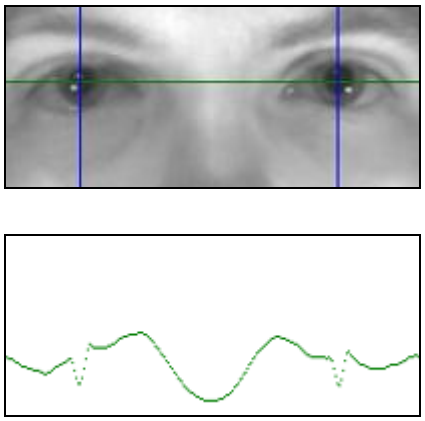
Subject C: Set 1





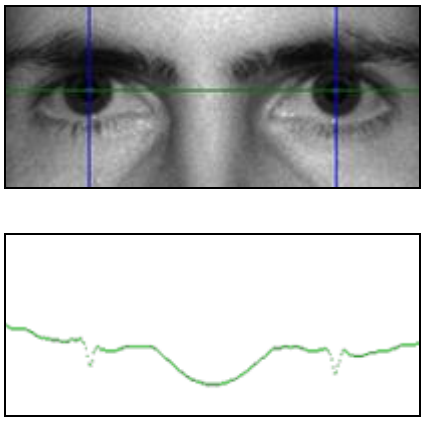
Wave attribute	Eye 1	Eye 2	Match %
Edge difference	12.61	28.08	44.9
Start to highest value	92.39	82.41	89.20
Middle to highest value	0	0	100
End to highest value	105	110.5	95.03
Start to middle value	92.39	82.41	89.20
End to middle value	104.9	110.5	95.028
Width start to end	12	10	83.333
Width start to middle	4	5	80
Width start to highest	4	5	80
Width middle to highest	0	0	100
Width middle to end	4	5	80
Width highest to end	4	5	80

Subject D: Set 1



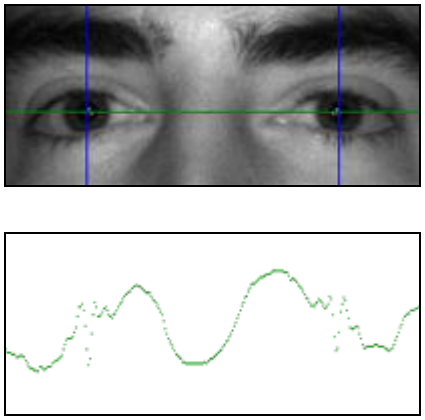
Wave attribute	Eye 1	Eye 2	Match %
Edge difference	58.16	19.39	33.34
Start to highest value	118.8	93.04	78.34
Middle to highest value	21.99	27.12	81.11
End to highest value	176.9	112.4	63.546
Start to middle value	96.77	65.92	68.122
End to middle value	154.9	85.31	55.063
Width start to end	10	10	100
Width start to middle	5	5	100
Width start to highest	4	4	100
Width middle to highest	1	1	100
Width middle to end	5	5	100
Width highest to end	6	6	100

Subject E: Set 1



Wave attribute	Eye 1	Eye 2	Match %
Edge difference	-44.25	-56.84	77.86
Start to highest value	173.4	144.1	83.089
Middle to highest value	0	0	100
End to highest value	129.1	87.21	67.557
Start to middle value	173.4	144.1	83.088
End to middle value	129.1	87.21	67.547
Width start to end	12	12	100
Width start to middle	6	6	100
Width start to highest	6	6	100
Width middle to highest	0	0	100
Width middle to end	6	6	100
Width highest to end	6	6	100

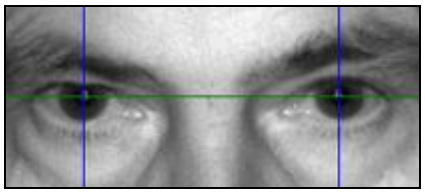
Subject F: Set 1

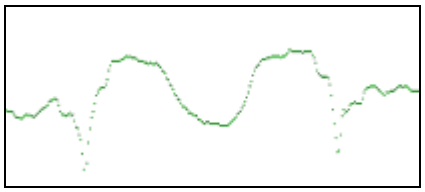


Wave attribute	Eye 1	Eye 2	Match %
Edge difference	-14.75	-12.64	85.74
Start to highest value	134.2	130.7	97.42
Middle to highest value	0	17.36	0
End to highest value	119.4	118.1	98.87
Start to middle value	134.2	113.4	84.49
End to middle value	119.4	100.7	84.33
Width start to end	15	17	88.235
Width start to middle	4	4	100
Width start to highest	4	3	75
Width middle to highest	0	1	0
Width middle to end	5	5	100
Width highest to end	5	6	83.333

Wave attribute	Eye 1	Eye 2	Match %

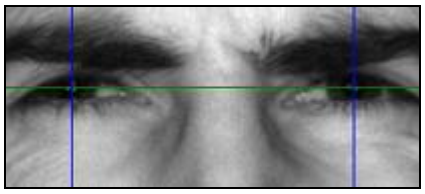
Subject G: Set 1

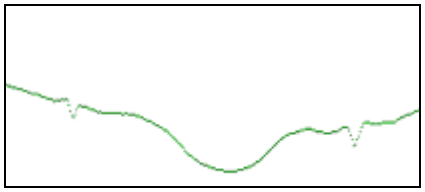




Edge difference	61.04	-88.95	0
Start to highest value	139.7	186.3	74.956
Middle to highest value	69.62	66.57	95.607
End to highest value	200.7	97.36	48.511
Start to middle value	70.02	119.7	58.478
End to middle value	131.1	30.79	23.429
Width start to end	18	14	77.777
Width start to middle	9	7	77.777
Width start to highest	7	9	77.777
Width middle to highest	2	2	100
Width middle to end	9	7	77.777
Width highest to end	11	5	45.455

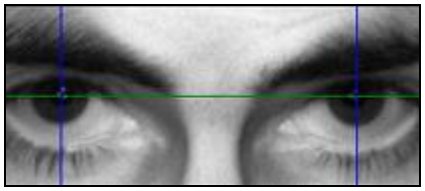
Subject A: Set 2

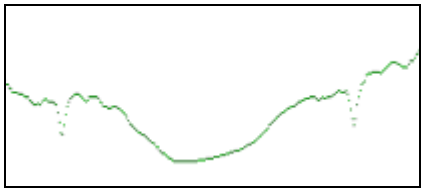




Wave attribute	Eye 1	Eye 2	Match %
Edge difference	-37.50	9.165	0
Start to highest value	126.1	102.7	81.445
Middle to highest value	3.667	2.092	57.043
End to highest value	88.55	111.8	79.181
Start to middle value	122.4	100.6	82.176
End to middle value	84.88	109.7	77.349
Width start to end	12	11	91.666
Width start to middle	6	4	66.666
Width start to highest	7	5	71.429
Width middle to highest	1	1	100
Width middle to end	6	5	83.333
Width highest to end	5	4	80

Subject A: Set 3

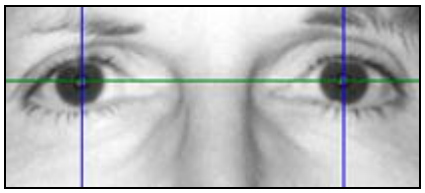


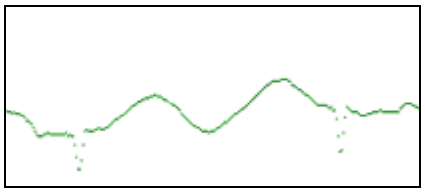


Wave attribute	Eye 1	Eye 2	Match %
Edge difference	28.58	57.88	49.376
Start to highest value	151.4	117.2	77.396
Middle to highest value	49.54	88.22	56.159
End to highest value	180	175.1	97.269
Start to middle value	101.8	28.95	28.428
End to middle value	130.4	86.84	66.579
Width start to end	15	13	86.666
Width start to middle	7	6	85.714
Width start to highest	5	3	60
Width middle to highest	2	3	66.666
Width middle to end	8	7	87.5
Width highest to end	10	10	100

Wave attribute	Eye 1	Eye 2	Match %

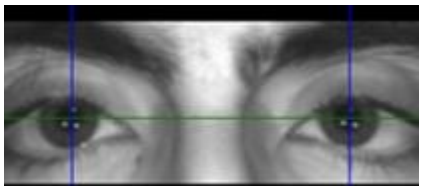
Subject B: Set 2

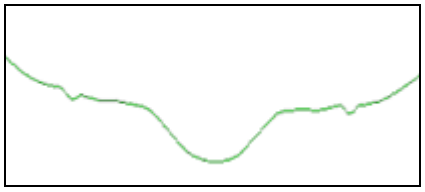




Edge difference	6.615	-4.81	0
Start to highest value	67.92	95.22	71.325
Middle to highest value	0	0	100
End to highest value	74.53	90.41	82.436
Start to middle value	67.92	95.22	71.324
End to middle value	74.53	90.41	82.44
Width start to end	10	15	66.666
Width start to middle	5	4	80
Width start to highest	5	4	80
Width middle to highest	0	0	100
Width middle to end	5	5	100
Width highest to end	5	5	100

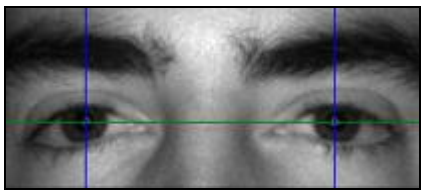
Subject C: Set 2

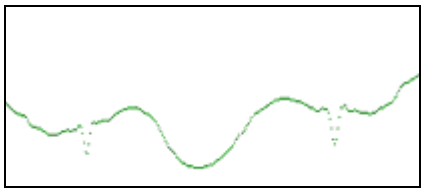




Wave attribute	Eye 1	Eye 2	Match %
Edge difference	20.1	-75.42	0
Start to highest value	133.8	170.8	78.34
Middle to highest value	7.531	7.051	93.631
End to highest value	153.9	95.33	61.954
Start to middle value	126.2	163.7	77.114
End to middle value	146.3	88.29	60.324
Width start to end	16	13	81.25
Width start to middle	8	6	75
Width start to highest	7	7	100
Width middle to highest	1	1	100
Width middle to end	8	7	87.5
Width highest to end	9	6	66.666

Subject F: Set 3

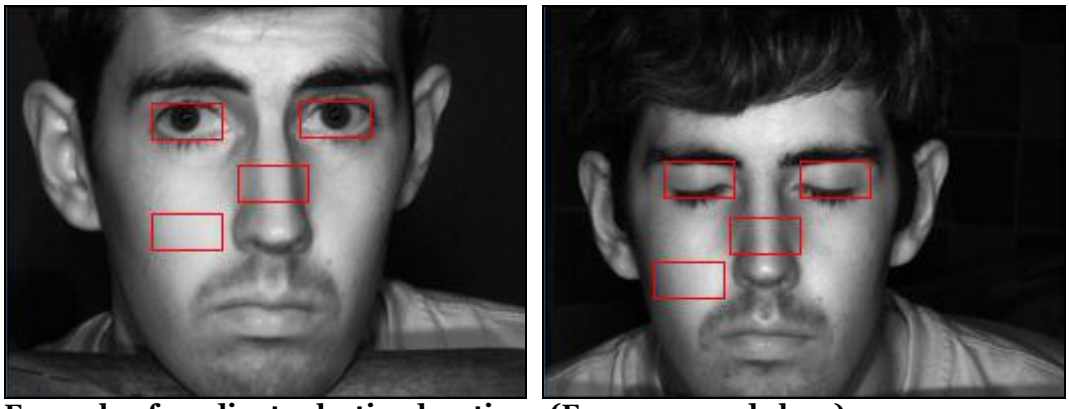




Wave attribute	Eye 1	Eye 2	Match %
Edge difference	23.27	-5.325	0
Start to highest value	100.6	142.9	70.391
Middle to highest value	6.711	8.754	76.662
End to highest value	123.9	137.6	90.021
Start to middle value	93.91	134.2	69.982
End to middle value	117.2	128.9	90.929
Width start to end	13	11	84.615
Width start to middle	4	5	80
Width start to highest	3	6	50
Width middle to highest	1	1	100
Width middle to end	5	6	83.333
Width highest to end	6	5	83.333

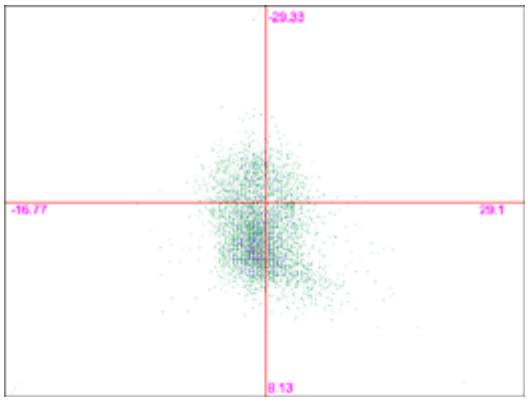
Appendix I: Gradient Plotting Results

This document provides the results of the eye location experiment using gradient graphs. Three subjects have been selected from the face database, with results presented for eyes open and eyes closed images of each subject. The different colours on the graph are used to represent the number of surface gradients that are present at each location. The colours range from black, for few identical locations, to dark blue for many.

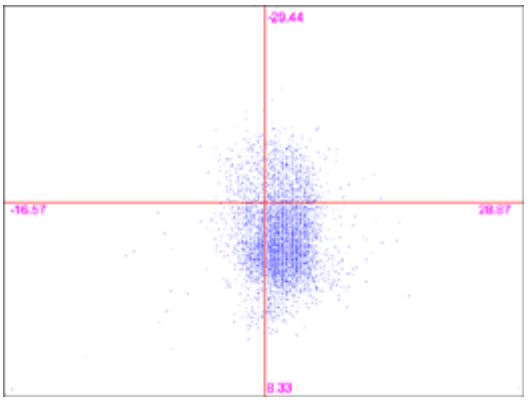


Example of gradient selection locations (Eyes open and close)

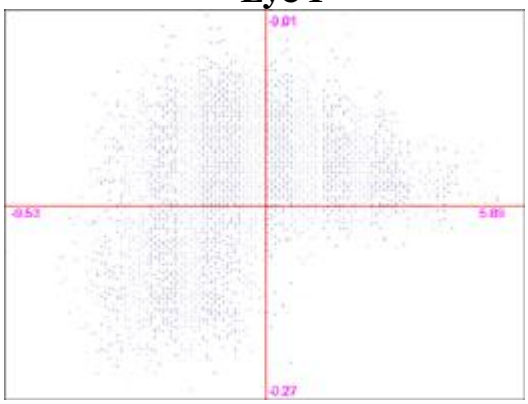
Subject A: Set 1 (eyes open)



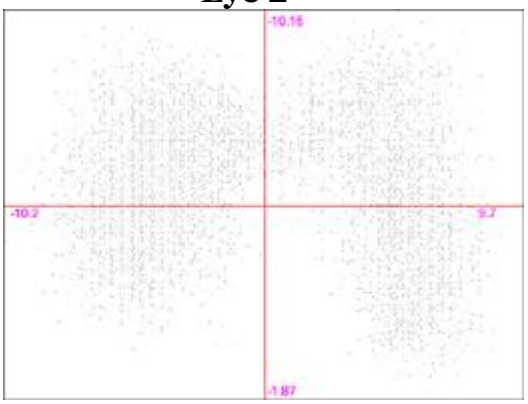
Eye 1



Eye 2

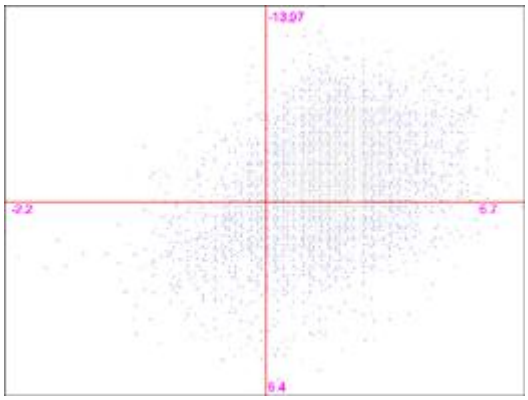


Cheek

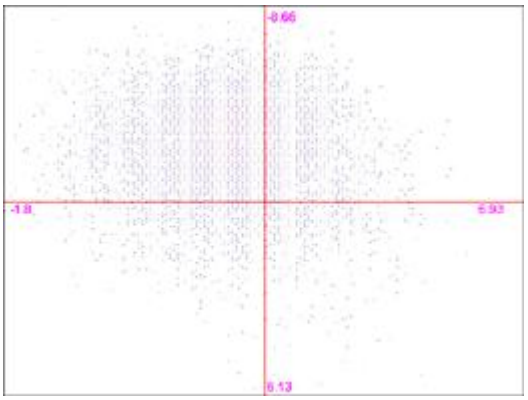


Nose Ridge

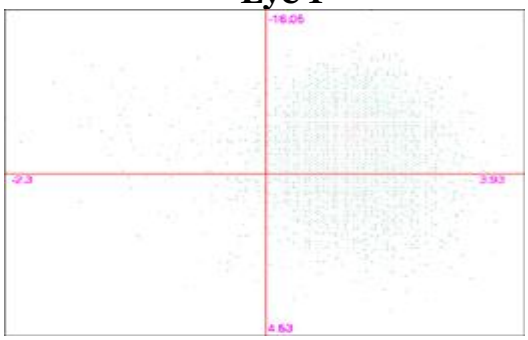
Subject A: Set 10 (eyes closed)



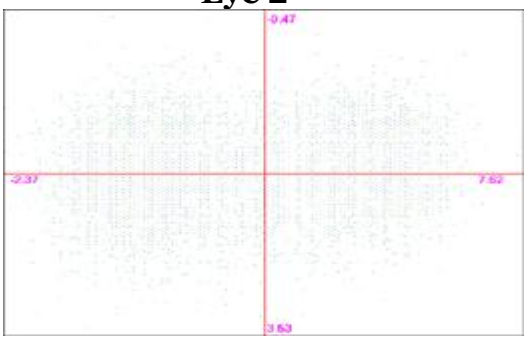
Eye 1



Eye 2

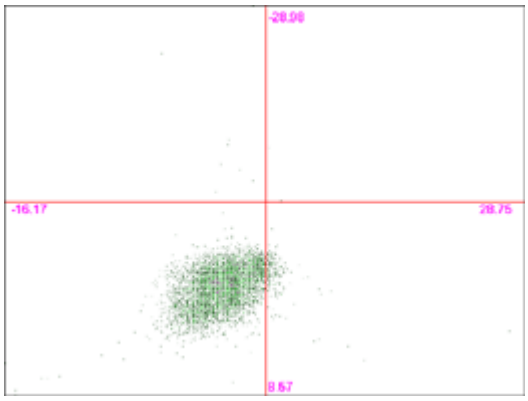


Cheek

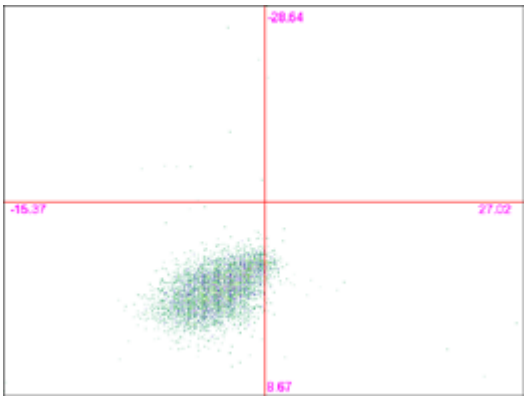


Nose Ridge

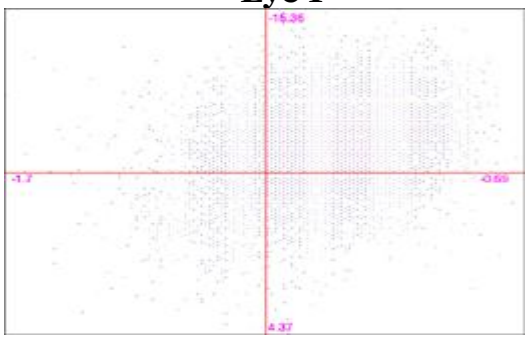
Subject B: Set 1 (eyes open)



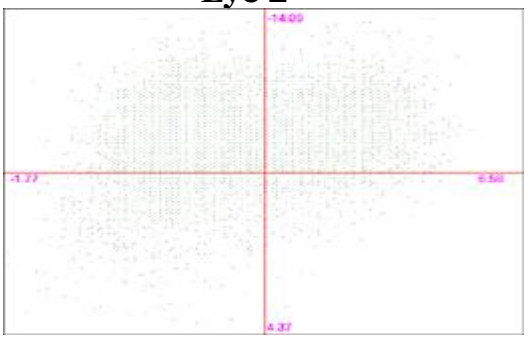
Eye 1



Eye 2

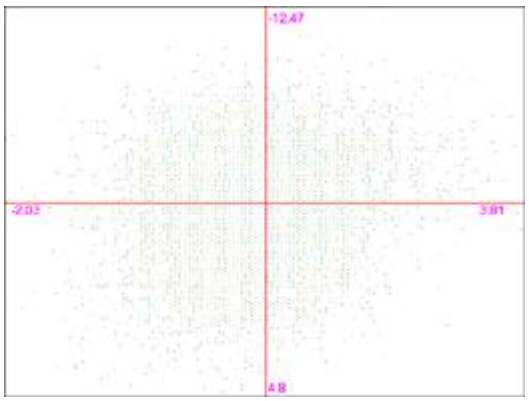


Cheek

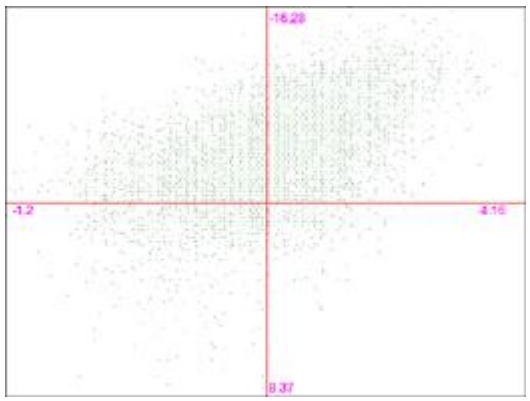


Nose Ridge

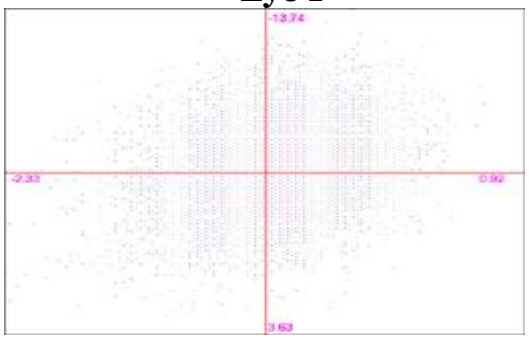
Subject B: Set 3 (eyes closed)



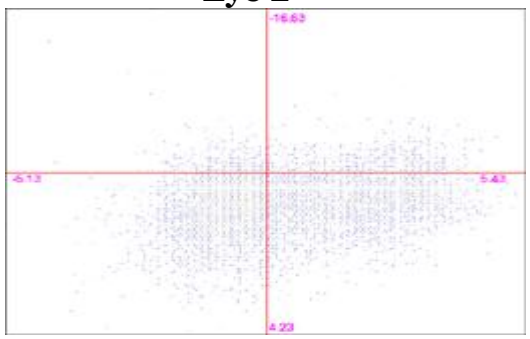
Eye 1



Eye 2

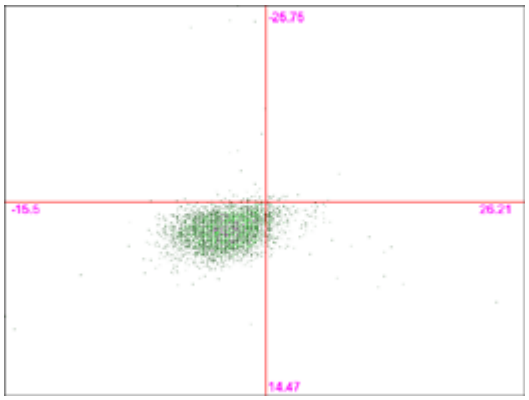


Cheek

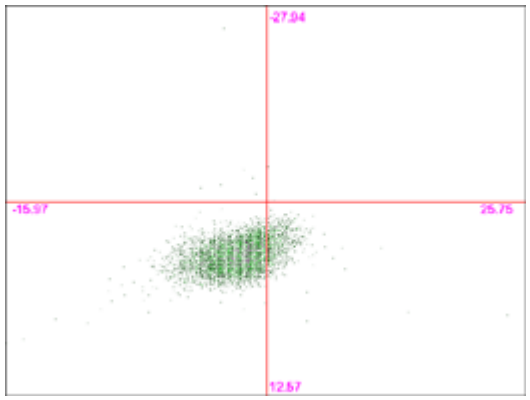


Nose Ridge

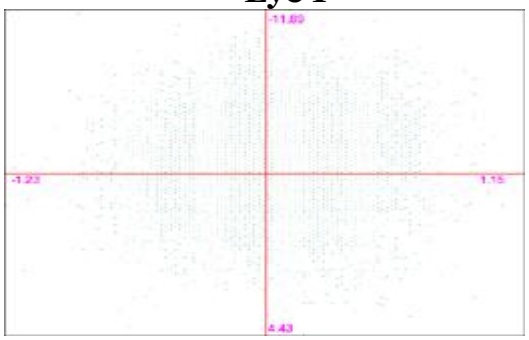
Subject G: Set 1 (eyes open)



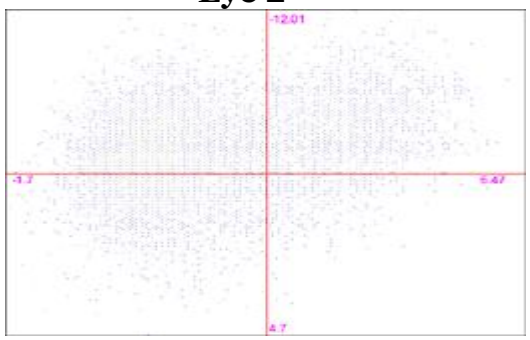
Eye 1



Eye 2

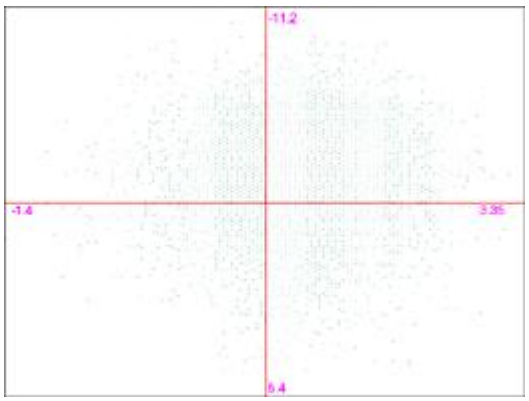


Cheek

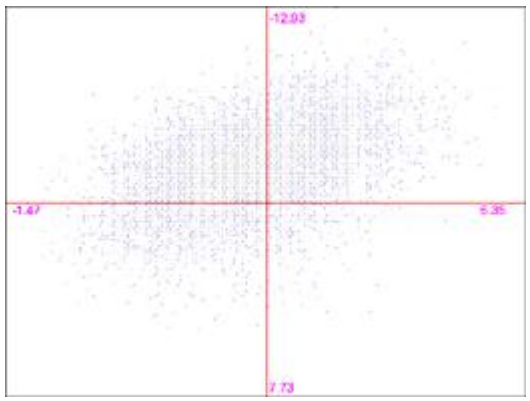


Nose Ridge

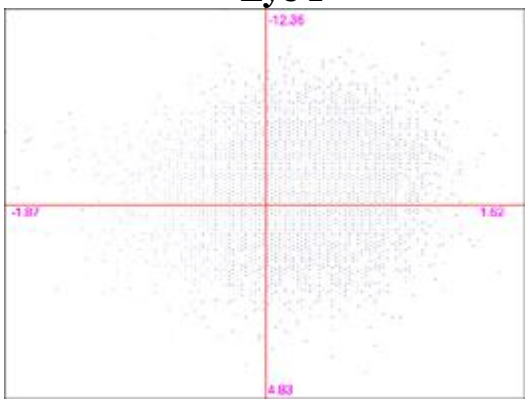
Subject G: Set 2 (eyes closed)



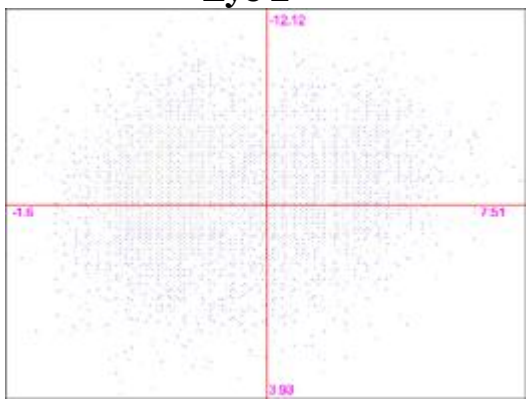
Eye 1



Eye 2



Cheek



Nose Ridge

Appendix J: Face Matching Results

These results consist of 27 face images from the face database (Faces wear glasses are worn or eyes are closed have been excluded from the test although are used in the search)

2D Feature comparison

Query face	Correct Match (Yes/No)	Matched Face	Match Value %
Subject A: Set 1	Yes	Subject A: Set 2	94.432
Subject B: Set 1	Yes	Subject B: Set 2	93.062
Subject C: Set 1	No	Subject D: Set 3	89.629
Subject D: Set 1	No	Subject B: Set 1	93.040
Subject F: Set 1	No	Subject B: Set 1	93.607
Subject G: Set 1	No	Subject A: Set 12	92.038
Subject A: Set 2	Yes	Subject A: Set 1	94.432
Subject A: Set 3	Yes	Subject A: Set 4	98.742
Subject A: Set 4	Yes	Subject A: Set 3	94.742
Subject A: Set 5	Yes	Subject A: Set 6	90.958
Subject A: Set 6	Yes	Subject A: Set 5	90.958
Subject A: Set 7	No	Subject D: Set 3	92.705
Subject A: Set 8	No	Subject B: Set 2	89.195
Subject A: Set 9	No	Subject C: Set 2	92.814
Subject A: Set 12	Yes	Subject A: Set 13	98.591
Subject A: Set 13	Yes	Subject A: Set 12	98.591
Subject A: Set 14	Yes	Subject A: Set 13	92.707
Subject A: Set 15	Yes	Subject A: Set 7	87.514
Subject B: Set 2	No	Subject D: Set 2	94.374
Subject C: Set 2	Yes	Subject C: Set 1	88.531
Subject C: Set 3	Yes	Subject C: Set 1	89.630
Subject C: Set 4	No	Subject A: Set 14	90.633
Subject D: Set 2	No	Subject B: Set 2	94.374
Subject D: Set 3	No	Subject F: Set 3	96.813
Subject F: Set 2	No	Subject A: Set 6	87.607
Subject F: Set 3	No	Subject D: Set 2	96.813
Subject G: Set 2	No	Subject A: Set 3	89.159

3D Feature comparison

Query face	Correct Match (Yes/No)	Matched Face	Match Value %
Subject A: Set 1	Yes	Subject A: Set 3	89.043
Subject B: Set 1	Yes	Subject B: Set 2	94.932
Subject C: Set 1	No	Subject A: Set 2	95.424
Subject D: Set 1	Yes	Subject D: Set 2	95.584
Subject F: Set 1	Yes	Subject F: Set 3	83.782
Subject G: Set 1	Yes	Subject G: Set 2	94.532
Subject A: Set 2	Yes	Subject A: Set 8	96.856
Subject A: Set 3	Yes	Subject A: Set 1	89.043
Subject A: Set 4	Yes	Subject A: Set 9	94.186
Subject A: Set 5	No	Subject C: Set 1	91.867
Subject A: Set 6	No	Subject F: Set 2	82.182
Subject A: Set 7	Yes	Subject A: Set 14	93.684
Subject A: Set 8	Yes	Subject A: Set 2	96.857
Subject A: Set 9	Yes	Subject A: Set 4	94.186
Subject A: Set 12	No	Subject C: Set 4	92.993
Subject A: Set 13	Yes	Subject A: Set 14	96.592
Subject A: Set 14	Yes	Subject A: Set 13	96.592
Subject A: Set 15	No	Subject G: Set 2	93.199
Subject B: Set 2	Yes	Subject B: Set 1	94.932
Subject C: Set 2	No	Subject D: Set 1	69.9137
Subject C: Set 3	Yes	Subject C: Set 1	95.005
Subject C: Set 4	No	Subject A: Set 12	92.993
Subject D: Set 2	Yes	Subject D: Set 1	95.584
Subject D: Set 3	Yes	Subject D: Set 2	89.170
Subject F: Set 2	Yes	Subject F: Set 3	90.237
Subject F: Set 3	Yes	Subject F: Set 2	90.237
Subject G: Set 2	Yes	Subject G: Set 1	94.532

2D + 3D Feature comparison

Query face	Correct Match (Yes/No)	Matched Face	Match Value %
Subject A: Set 1	Yes	Subject A: Set 9	88.700
Subject B: Set 1	Yes	Subject B: Set 2	93.810
Subject C: Set 1	Yes	Subject C: Set 3	91.780
Subject D: Set 1	Yes	Subject D: Set 2	91.743
Subject F: Set 1	Yes	Subject F: Set 3	87.189
Subject G: Set 1	Yes	Subject G: Set 2	91.934
Subject A: Set 2	Yes	Subject A: Set 8	92.259
Subject A: Set 3	Yes	Subject A: Set 12	84.286
Subject A: Set 4	Yes	Subject A: Set 3	84.263
Subject A: Set 5	Yes	Subject A: Set 2	84.793
Subject A: Set 6	No	Subject F: Set 2	85.437
Subject A: Set 7	Yes	Subject A: Set 14	90.135
Subject A: Set 8	Yes	Subject A: Set 2	92.259
Subject A: Set 9	Yes	Subject A: Set 1	88.691
Subject A: Set 12	Yes	Subject A: Set 13	93.646
Subject A: Set 13	Yes	Subject A: Set 14	94.261
Subject A: Set 14	Yes	Subject A: Set 13	94.261
Subject A: Set 15	No	Subject G: Set 1	87.949
Subject B: Set 2	Yes	Subject B: Set 1	93.810
Subject C: Set 2	No	Subject D: Set 2	75.906
Subject C: Set 3	Yes	Subject C: Set 1	91.780
Subject C: Set 4	No	Subject A: Set 12	89.328
Subject D: Set 2	Yes	Subject D: Set 1	91.743
Subject D: Set 3	Yes	Subject D: Set 1	87.001
Subject F: Set 2	No	Subject A: Set 6	85.437
Subject F: Set 3	No	Subject B: Set 2	88.149
Subject G: Set 2	Yes	Subject G: Set 1	91.934

2D + 3D Feature comparison
(3D features double weighted)

Query face	Correct Match (Yes/No)	Matched Face	Match Value %
Subject A: Set 1	Yes	Subject A: Set 9	87.052
Subject B: Set 1	Yes	Subject B: Set 2	94.131
Subject C: Set 1	Yes	Subject C: Set 3	92.702
Subject D: Set 1	Yes	Subject D: Set 2	92.840
Subject F: Set 1	Yes	Subject F: Set 3	86.215
Subject G: Set 1	Yes	Subject G: Set 2	92.057
Subject A: Set 2	Yes	Subject A: Set 8	93.573
Subject A: Set 3	Yes	Subject A: Set 1	85.138
Subject A: Set 4	Yes	Subject A: Set 9	85.532
Subject A: Set 5	Yes	Subject A: Set 2	86.456
Subject A: Set 6	No	Subject F: Set 2	84.507
Subject A: Set 7	Yes	Subject A: Set 14	91.149
Subject A: Set 8	Yes	Subject A: Set 2	93.573
Subject A: Set 9	Yes	Subject A: Set 1	87.053
Subject A: Set 12	Yes	Subject A: Set 13	92.057
Subject A: Set 13	Yes	Subject A: Set 14	94.927
Subject A: Set 14	Yes	Subject A: Set 13	94.927
Subject A: Set 15	No	Subject G: Set 2	89.4491
Subject B: Set 2	Yes	Subject B: Set 1	94.131
Subject C: Set 2	No	Subject D: Set 2	73.317
Subject C: Set 3	Yes	Subject C: Set 1	92.702
Subject C: Set 4	No	Subject A: Set 12	90.375
Subject D: Set 2	Yes	Subject D: Set 1	92.840
Subject D: Set 3	Yes	Subject D: Set 2	87.379
Subject F: Set 2	Yes	Subject F: Set 3	85.660
Subject F: Set 3	No	Subject B: Set 2	87.390
Subject G: Set 2	Yes	Subject G: Set 1	92.057