UNIVERSITY
of
GLASGOW

Facial Anthropometry as an Evidential Tool in Forensic Image Comparison

Submitted to the University of Glasgow for the degree of
Doctor of Philosophy
By
Krista F. Kleinberg, RPh
Forensic Medicine and Science
Division of Cancer Sciences and Molecular Pathology

March 2008

Acknowledgements

I would like to express my sincere appreciation to my supervisors, Dr. Robert A. Anderson, Dr. J Paul Siebert, Professor Peter Vanezis, and Professor A. Mike Burton, for their time, guidance, encouragement and patience in the development of my research project. I would like to give an extra special thank you to Dr. Anderson and Dr. Siebert. Their help during my time at the University of Glasgow cannot be over stated. The end result is not only my completed thesis but also my love of forensic science and the discovery of new ideas and thoughts.

I wish to extend special thanks to my University Advisor, Dr. Mike Edward and colleagues Maria Vanezis and Sarah Sholl, and to Fran Ackermann who was instrumental in guiding me through the PhD process.

I want also to extend my thanks to John C. Hay and to the volunteers who graciously lent their free time to help with the research process. For their continued help and support I wish to thank Elizabeth Stein and Yvonne Spence and lastly to my friends and family for their never-ending encouragement.

To my Dad and Mom, Michael and Marcylyn, I owe my deepest gratitude for supporting my desire to return to school and to study an area that has meaning and importance in today's world. Thank you so much for your love and support.

List of Contents

List of Tables

List of Figures

Glossary of Terms

Anthropometry: n. the comparative study of sizes and proportions of the human body [1].

Frankfort horizontal plane: n. a plane used in craniometry that is determined the highest point on the upper margin of the opening of each external auditory canal and the low point on the lower margin of the left orbit and that is used to orient a human skull or head usually so that the plane is horizontal [2].

Morphology: n. the form and structure of anything [1].

Nasal alae: relating to the spread of the nostrils.

Palpebral slit: the line drawn from the tear duct towards the opposite corer of the eye.

Philtrum: the vertical groove in the median portion of the upper lip [3].

Vector: column or row of elements.

Element: a member of a set. In Chapter 6, the term is used interchangeably with proportions.

Summary

Anthropometry can be used in certain circumstances to facilitate comparison of a photograph of a suspect with the potential offender portrayed in video surveillance crime footage. Anthropometry does not have the same success rate in identification as DNA or fingerprinting. However, these types of evidence are not always left at crime scenes. Sometimes the only evidence available relating to an offence is from surveillance videos and research was needed to lend credence to anthropometry as a viable method of identification. An alternative method of detecting individuals from surveillance video, morphology, was also investigated to determine its accuracy in confirming the identity of individuals based on facial descriptions and for use as a comparison tool in forensic identification.

Pilot Laboratory Study: A number of different techniques are employed in facial image comparison of living persons. In this study, the effect of rotation on angles and proportions between selected facial landmarks is evaluated as a first step to assess whether facial anthropometry could be usefully applied to facial image comparison. The faces of five volunteers were photographed in the Frankfort plane at different angles of rotation from 0º (frontal) to 90º (side view), rotating every 10º both clockwise and counter-clockwise. Four landmarks were used: right and left ectocanthions, nasion, and stomion. The proportions of the measurements between these landmarks were calculated as well as the angles created by the lines connecting the same landmarks. The results show a consistent and predictable variation between the five subjects. With rotation, the greatest variation is seen where horizontal landmark connecting lines are combined with the ectocanthion/stomion or nasion/stomion lines. There is less variation in the proportions for vertical and diagonally orientated landmark connecting lines. In principle, the data from these empirical measurements could also be used to develop a photogrammetric model of the face which, if calibrated, could be used to correct anthropometric measurements for distortions caused by a camera angle which differs from the one specified in a protocol for facial comparison. The purpose of developing such a model would lie in its use to calculate correction factors to convert observed proportions and angles back to the full-face orientation values, which could then, for example, be used to search a database of the proportions.

Investigation of Uncertainty of Anthropometric Measurements: The objective of this study was to estimate the uncertainty in the measurements of the chosen facial proportions caused by landmark placement and by operators taking photographs, including the uncertainty contributions resulting from different people performing these tasks. The aim of this was to simulate effects found in the real world, as there would be different operators both placing landmarks and taking suspect photographs in various police departments. In addition, this study was completed in order to address variables encountered in the Pilot Laboratory Study that occurred as a result of the experimental set up. The first section of the study reviewed the errors involved in measuring facial proportions as a result of variations in landmark placement. Intra and inter-operator studies in landmark placement were conducted and as expected the average and range of coefficients of variation for the set of proportions were larger in the inter-operator error than that obtained in the intra-operator error. The second section of the study reviewed the errors in measuring facial proportions as a result of the process of taking photographs. The lowest variation in facial measurements was seen in the series of photographs taken of a single subject by a single operator and in general, the lowest variation in facial measurements was seen at 45° and the highest variation at 20°. The contributions of errors from landmark placement and photography were determined to produce an overall estimated uncertainty of 5%. When a comparison of 2D images is conducted in this manner this estimation of uncertainty should be taken into account.

Anthropometry Study: An existing database of video and photograhic images was examined, which had previously been used in a psychological research project with the aim to test the hypothesis: "Using a comparison of anthropometric facial proportions, it is possible to discriminate between individuals of two samples." Material avaliable consisted of 80 video (Sample 1) and 119 photograhic (Sample 2) images and were of high resolution, though taken with different cameras. A set of 37 anthropometric landmarks were placed measuring 59 proportions to conduct within sample and between sample comparisons using the following calculations; mean absolute value between proportions, Eulcidean distance and Cosine θ distance between proportions. First, the statistics of the two samples were examined to determine which calculation best ascertained if there were any differences between faces which fall under the same conditions. Subsequent to a between sample, the removal of up to 50% of the lowest variant proportions along with the determination of a subsample of faces requiring human verification were tested. Relative frequency distribution histograms were created from the data and the normal histogram curves of true positive and true negative faces were superimposed to determine their separation rate and how likley it may be to mix up the two categories of faces. Presented results showed that the Cosine θ distance equation using Z-normalized values was the preferred equation because it achieved the largest separation between true positive and true negative faces. Results also indicated that there was no benefit to removing up to 50% of the lowest variant proportions in the comparison of Sample 1 against Sample 2. Finally, applying the Cosine θ distance equation allowed a decrease to five database images to be verified by a human in approximately 75% of the cases tested.

Morphology study: A morphological analysis was conducted on high resolution images and although highly relevant to the process of facial identification did not contribute to the continuity of the thesis and thus was included as an Appendix. The morphological analysis was performed on a total of 199 images: 119 photographs and 80 images from video using a checklist of 20 facial characteristics. Each facial characteristic had numerous choices in which it could be described. Once the analysis on all 199 images was carried out, a comparison was conducted between each video (unknown) image and the database of 199 (known) images. In the research conducted, only 2.5% of the comparisons showed a true positive match between video and photograph with zero false positives in the group. Subsequent to analysis it was determined not possible to differentiate between individuals, however, when looking directly at the individuals' photographs, it is clear that there were differentiating characteristics amongst them.

Conclusions: After embarking upon a series of anthropometrical investigations using high resolution images to compare video images with photographic images, it was concluded that anthropometry, when used as a comparison method of identification, does not generate the results necessary for use as evidence in a court of law. Identifying individuals based on a morphological analysis of a check list of features alone also did not result in clear consistent identifications. If descriptions of facial characteristics are to be fully utilized, a side by side comparison is likely to be less subjective. This outcome was as expected and provides additional insights into forensic morphological research.

1 Forensic Science and Human Identification

By definition, forensic science is the application of one or more branches of science to the investigation of questions of interest to the legal system. This incorporates a large number of professionals in a variety of disciplines. The personnel involved in the law enforcement and legal side include police, procurators fiscal/coroners and lawyers, all of whom are imperative to the criminal investigation and judicial processes. The scientists, including pathologists, are imperative to the collection, processing and exposition of the evidence.

The identification of individuals, either living or dead, is essential for legal and medical reasons as well as bringing peace of mind to relatives and friends. Living persons may not remember who they are due to amnesia [4] and identification of the deceased is the one of the first steps in the investigation of suspicious deaths. Positive identification requires comparison and matching of at least one unique feature, either using dental or medical records or, in recent years, using the DNA profile [5]. Examining the dead to determine identity in cases where the individual is visually unrecognizable can be undertaken in a collaborative effort by forensic pathologists, forensic odontologists, forensic anthropologists and other forensic science experts such as fingerprint and DNA specialists.

The process of identifying individuals is typically based on a comparison of an unknown and a known and is used when there is someone to compare the evidence against or it is used to create the potential for comparing the evidence against a suspect in the future. The methods of comparison of DNA, fingerprint and dental records all rely on taking details from the unidentified individual and comparing them against the same details from a known individual. Facial identification methods based on comparisons include anthropometry, morphology, and superimposition. Facial reconstruction is one method of identification that does not initially use comparison in a traditional sense but is used as a method of last resort when there is no other possible way of identifying the individual. A more detailed explanation of the different types of facial identification follows in the next chapter.

Due to recent scientific advances and reports in the media, DNA analysis is probably the most widely recognised method of identifying individuals and linking suspects to crime scenes. DNA has also been used as a way to identify positively human remains which have been extremely damaged or aged, including victims of war crimes [6, 7] and, for example, the victims of the September 11th terrorist attacks [8]. A person can even be identified through DNA found on partially eaten food [9] that has been found at the scene or from DNA obtained from fixed and mounted histology slides [10]. DNA analysis and comparison is used both to exonerate innocent prisoners and to strengthen cases for prosecution [11, 12].

Using fingerprints as a means of identification dates back even further, to the 1820’s, when Johann Purkinje stated that everyone’s fingerprints were different [13]. The first person to use fingerprints to identify criminals was Henry Faulds in 1880 [13]. He did this after the police showed him some sooty fingerprints found on a wall after a suspected robbery. Faulds took the fingerprints of the suspect they had in custody, and found they were very different from the ones on the wall. A few days later, another suspect was arrested whose fingerprints matched the ones on the wall.

Dental records are an effective way to identify or corroborate identification of an unknown fatality, although this can be difficult at times because of a lack of antemortem patient records. However, even in the absence of antemortem records, it may still be possible to obtain general information on the sex, age, and ethnic origin of the individual [4]. Forensic odontologists have two main roles: they help to confirm identification of a deceased person based on a comparison of postmortem and antemortem dental records [14], and they examine and compare bite marks [4]. Forensic odontologists are especially useful during the investigation of mass deaths and disasters when visual identification may be impossible.

The most easily recognizable way to identify someone is from their face and therefore the methods of identification that cater to the face are all very important to forensic science. Each method of identification has its purpose depending on which type of evidence is available. Whether the evidence available is from a video recording or eye witness, the application of facial identification techniques is vital to the investigation of crime and crucial to the well being of family, friends and the victims themselves.

1.1 Aims and Scope of Present Study

The aims of the research undertaken were to determine the operational characteristics of forensic anthropometry when used in identification of individuals and so establish if forensic anthropometry is sufficiently accurate and consistent to be used as evidence in court. Forensic anthropometry has traditionally been considered to be a controversial and unproven method of identification, The general feeling within the forensic science community is that there are too many factors which make this method subjective and that even when high quality photographs taken in a controlled setting are available, factors such as lighting, head position, camera position, and operator experience may all contribute to the inaccuracy of this technique for identification purposes.

DNA, if left at the crime scene, can be the most reliable way to identify an individual. However, the question remains as to what happens when DNA is not found at the scene. DNA does not solve all identification problems: DNA cannot be obtained from a surveillance video image. The field of forensic research is always looking for new methods of identification. All ideas are speculative in the beginning of the research process but new ideas cannot be dismissed. They must be evaluated and they must be tested. Disciplines such as gait recognition, ear print identification and speaker identification are all in the research phase and identification from footwear impressions, which now has its own database, was once thought of as obscure.

The potential for using anthropometry as an identification method has been created by the widespread use of surveillance cameras as well as mobile phone cameras. There are few places, especially in the urban environment, that cannot be photographed or monitored. Cameras play a predominant role in present day society and anthropometry, if of evidential value, would benefit from this. Interest in anthropometry is present amongst forensic practitioners, for example, during the course of the study presented here a collaborative research project was published which aimed to assess facial variability and was funded by the US Government on behalf of the Federal Bureau of Investigation [15, 16].

The aims of the present study were as follows:

(a) To conduct literature reviews on:

- different facial identification methods, including anthropometry, morphology, superimposition and facial reconstruction;
- identifications made by eyewitnesses and biometric systems;
- systems for video surveillance and image enhancement;

(b) To determine the overall uncertainty that can occur as a result of the measurement process;

(c) To evaluate anthropometry as a method for facial identification and to determine the degree of accuracy with which identification can be made of individuals based on anthropometric measurements;

(d) To evaluate facial morphology to obtain an insight into facial image comparison. This analysis although highly relevant to the process of facial identification did not contribute to the continuity of the thesis and thus was included as an Appendix.

These aims were assessed to disprove the null hypothesis “It is not possible by utilizing a series of facial anthropometric measurements to be able to satisfactorily discriminate between individual subjects when comparing images of a known subject with those of a subject whose identity is not known in order to assess whether they are the same individual.”

No other methods were to be employed in the research and techniques such as biometrics, although available and useful, were taken out of the scope of this study.

2 Craniofacial Identification

2.1 Introduction

The following chapter is an introduction to craniofacial identification. It reviews different facial identification methods, including anthropometry, morphology, superimposition and facial reconstruction. The chapter also includes information on identifications made by eyewitnesses and biometric systems, and concludes with a general description of surveillance video systems and image enhancement techniques.

2.2 Anthropometry

Alphonse Bertillon transformed the French criminal justice system in 1882 when he implemented the first modern system of criminal identification. He achieved this through anthropometry, which became known as *Bertillonage* [17]. Moenssens states that "Bertillon's system of anthropometrical measurements was based on three ideas: the fixed condition of the bone system from the age of twenty until death; the extreme diversity of dimensions present in the skeleton of one individual compared to those in another; the ease and relative precision with which certain dimensions of the bone structure of a living person can be measured using simply constructed caliphers" [18]. Bertillon formulated a classification system based on eleven measurements including height, length and width of head, arm span, sitting height, length of left middle and little fingers, left foot length, left forearm length, right ear length and cheek width [17]. He insisted on precision when obtaining measurements, and was adamant that the prisoner being measured should stand or sit in a very specific way (Figure 2.1).

Bertillon recorded the person's measurements on file cards, along with a description of any unusual characteristics and photographs from the frontal and profile viewpoints. His superiors at the Préfecture of Police were sceptical about his system, but eventually gave him a three-month trial period in which to test it. From then on, Bertillon acquired measurements from every prisoner that was brought in. In February 1883, during the third month of this trial period, he succeeded in making his first anthropometric identification. Bertillon and his system gained credibility during the next year by producing 241 identifications [17].

Figure 2.1 has been removed due to Copyright restrictions

Bertillonage soon spread throughout the world. However, without Bertillon himself to oversee it, the system became sloppy and careless [17]. Foreign officials also took it upon themselves to change certain measurements or modify the instruments. The effect this had was that, as the system moved further away from France, the identification success rate decreased. The early 1900's saw the transformation in identification from anthropometry, to be replaced by dactlyoscopy (fingerprint identification) as the dominant identification method in the United States. This change was hastened by the Will West case. When Will West, an African American, was admitted into the United States Penitentiary at Leavenworth, he was mistaken for another inmate named William West, who looked almost identical to Will West and was indistinguishable based on his anthropometric measurements (Figures 2.2 – 2.3) [17].

Figure 2.2 has been removed due to Copyright restrictions

Figure 2.3 has been removed due to Copyright restrictions

Subsequent fingerprinting proved they were two different inmates and confidence in anthropometry was lost. However, this need not have been the case had Bertillon's rules been followed. Bertillon allowed for a 3mm margin of error, but the left foot measurements of the two individuals clearly exceeded this (Figure 2.2). The story contains many flaws and was most likely created to prove a point. One of these flaws was that, as groundbreaking a case as this was, it went unreported in both the press and the scientific literature. The earliest mention of the case was in 1918, fifteen years after the supposed incident. As the use of fingerprinting began to increase, some departments used it in conjunction with anthropometry. Anthropometry was still looked upon as the more scientific method, but fingerprinting was the future.

Alphonse Bertillon paved the way for anthropometry to be used in forensic cases as a means of identifying individuals. However, aside from anthropometry, other measurement techniques were used for classification purposes. Craniometry, which is the measurement of the skull and its contents, was used to categorize people in the early 1900's [19]. Brains were weighed and measured in research carried out by Robert Bennett Bean to determine the differences between the brains of African Americans and those of Caucasians. Bean noticed many differences and used these observations to support his ideology that African Americans were inferior to white Americans. Measurements were also used in phrenology, now considered a pseudoscience, which is the study of the structure of the skull to determine a person's character and mental capacity and was used along with anthropometry by the German National Socialist (Nazi) Party during World War II. The group, Bureau for Enlightenment on Population Policy and Racial Welfare, recommended that Aryans and non-Aryans should be classified on the basis of measurements of their skulls and other physical features [20]. One of the exhibitions at the United States Holocaust Memorial Museum called "Deadly Medicine: Creating the Master Race," contains medical memorabilia and includes amongst other things "…a wide array of callipers for racial anthropometry" [21].

While society today may be disturbed by the use of anthropometry to separate individuals during the Nazi era, the technique has nevertheless been investigated for use in a similar manner in the present day, in a study to establish the differences between paedophiles and rapists [22]. Before undertaking the study, Taylor et al. observed through their work with sexual offenders that rapists were more athletic and muscular, with low percentages of body fat, while paedophiles were less muscular and athletic and had higher percentages of body fat. For their study, they took ten body measurements, including height and weight, of 23 paedophiles and 13 rapists. The results showed a lack of significant anthropometric differences for most measurements and therefore separating the subjects into somatotypes was not useful. The findings did show that the rapists in the study were more muscular and had less body fat then the paedophiles. However, the results could not be generalized due to the small sample size and the fact that the sample was taken from a prison population.

Anthropometry was one of the methods employed in the identification of John Paul Jones, who was an American 18th Century naval hero [23]. John Paul Jones, originally born in Scotland, emigrated to the North American Colonies and fought in the Revolutionary War. He passed away in Paris in the year 1792 and was buried without a marker in the St. Louis Cemetery for Protestants of foreign birth which was subsequently sold and covered over by landfill. In 1899, the U.S.

Ambassador to France wanted to find the body of John Paul Jones and return him to the United States. Six years of research led him to the abandoned cemetery and in February of 1905 excavations began. Identification was based on comparison of the corpse with written descriptions and artwork made during his life. Several methods were used to attempt identification and anthropometry was included, along with photographic superimposition and a morphological analysis. Two busts and a medal portrait were used for comparison: one of the busts was unsigned and the other, by Houdon, was a result of a 1780 sitting with the artist. Photographs were taken of the corpse and superimposed against the bust by Houdon. Morphology of the shape of the ear was also noted.

Critiques of the identification question the authenticity of Houdon's bust. It is said that Houdon frequently used life masks to help better replicate his subjects, although there is no proof that he used one when he created Paul Jones. This was in doubt because as well as aging 12 years from when the bust was made until his death, Paul Jones suffered through many illnesses. To refute this, it is said that the measurements were taken before the photographs were taken, while the corpse was still moist from preservative. It was concluded that the measurements taken of the bust and corpse support the positive identification of John Paul Jones (Table 2.1).

Table 2.1 has been removed due to Copyright restrictions

The use of facial measurements as a form of evidence is still new and its reliability in a court of law is debatable. Identification from photographs using facial anthropometry has been used in present day court cases, both in the United Kingdom and in the United States [24]. An expert witness report will usually contain multiple comparisons [24] and explain how the images were compared. Any similarities or differences are emphasized and distinguishing marks, such as moles or deformities, are used to support a conclusion that the individuals are the same or different.

"Facial mapping" is a term used to describe the process of identifying an individual through the facial features and can include video/photo superimposition, anthropometric measurements and morphological comparisons [24]. Anthropometric measurements use landmarks on the face as a way to identify specific areas. Although it is a relatively new form of evidence, it is being used more frequently in court and with greater confidence. The use of photographs in anthropometry is further described in Section 2.3.

Anthropometry played a role in the case of *R (on the application of Taj) v Chief Immigration Officer, Midlands Enforcement Unit, Queen's Bench Division (Administrative Court), CO/1084/99, 29 January 2001* [25]. The question considered was whether the claimant was the person identified in the passport or whether he was in fact the brother of the claimant. Two photographs were compared, taken from a passport application made in 1974 and from a British passport issued in 1996. One defence expert carried out a visual analysis, comparing specific facial features and shapes. His report explained that there were slight differences but that he felt, when first looking at the photographs that the faces were similar, indicating the two people could be related. Superimposition was carried out and particular attention was given to analysis of the facial measurements with the conclusion that, although the measurements did not provide absolute proof, they added support to the initial assessment. However, a second expert witness concluded that the differences between the faces in the photographs could have been caused by a range of factors, including weight gain, hair loss, head position, different facial expression, and that facial measurements were unreliable because of the different head positions in the photographs and because the method had not been tested enough to use in court.

In another case the following year [26], facial measurements were carried out for the prosecution on a suspected bank robber. An expert witness compared images taken from a videotape of the robbery with a photograph of the suspect, which was obtained without his knowledge or consent and declared a successful match.

In a study by Farkas et al. based on measurements found reliable in one of his previous studies, landmarks that were able to be seen clearly on a photograph were used to create age progression photographs for missing children using anthropometry [27]. The necessary materials to carry out this process were photographs of the missing child from the front and side, photographs of parents, siblings, and relatives, and the age of the missing child, as well as race and nationality. They concluded that the main role of anthropometry was to convert a small picture into a life sized photograph that could be used for gathering further information from the face. Proportions between anthropometric landmarks were useful in showing inconsistencies or uniformities of the face. They also found that measurements alone were not enough to carry out the task, but that an artist's interpretation was an essential tool as well.

2.3 Photogrammetry

The term used to describe the taking of precise measurements from photographs is *photogrammetry*. When anthropometric measurements are taken from photographs rather than from bodies or the faces of living and deceased persons, this is "photo-anthropometry". Ïşcan describes photo-anthropometry as "*the analysis of anthropometric landmarks, dimensions, and angles to quantify facial characteristics and proportions from a photograph*" [28].

2.3.1 Introduction

Photogrammetry can be traced back to Leonardo da Vinci's exploration of geometry, optics, mechanics, and geophysics and to Albrecht Dürer's outline of the laws of perspective [29]. Now mostly used to make maps from aerial photographs, photogrammetry has evolved from its inception in 1715 by Brook Taylor who published a book on linear perspective, one hundred years before the invention of photography [30]. In 1759, J.H. Lambert suggested that perspective could be valuable to mapping, but it was not until 1840 that Aimé Laussedate applied the rules of perspective to map making. Photogrammetry continued to develop when in 1909 the German scientist Pulfrich, conducted experiments in the use of stereoscopic pairs of photographs for mapping purposes.

Cameras used in present day society can be traced back to the pinhole camera (Figure 2.4). The materials that make up a pinhole camera may vary widely but basically include a light tight box, a piece of aluminum foil for the pinhole aperture and a piece of cut film taped to the opposite end of the box. Light surrounding the object filters in through the pinhole and the image is burned onto the film as a mirror image from the original subject. If the pinhole is either too small or too big, the result is a blurry image. The principal aim of photogrammetry is to obtain information concerning the real objects depicted in a photograph from this two-dimensional image.

Figure 2.4 has been removed due to Copyright restrictions

As photogrammetry involves the use of two dimensional images, it is constructive to understand the concept of perspective. Perspective is a method of depicting three dimensional objects within a two dimensional plane, giving the illusion of volume and space [32]. Linear perspective dominated Western painting from the Renaissance until the end of the 19th century. The effect linear perspective has on a two dimensional plane is that close objects appear to be larger than those further away. To create this illusion in an image with perspective, lines which are parallel in the object space appear to converge to a single point [33]. These points are called vanishing points and may be as numerous as there are sets of parallel lines in the object space.

Photogrammetry has many applications apart from mapping. These include: geology, forestry, agriculture, design and construction, planning of cities and highways, cadastre, environmental studies, exploration, military intelligence, medicine and surgery and miscellaneous applications, such as tailors using customers' measurements to make individually tailored suits [29].

2.3.2 Forensic Photogrammetry

Photogrammetrists may apply their knowledge and expertise in aerial and terrestrial photogrammetry to aid in law enforcement investigations [34]. Forensic aerial photogrammetry has been used to settle land disputes such as a dispute over extensive landfill work on one property which affected drainage on an adjacent property [34]. Forensic terrestrial photogrammetry has been used, for example, to determine responsibility for a death during a power boat race and in a case in which personal injury resulted from an accidental fire and explosion.

Often for these types of investigations, the photogrammetrist is asked to perform their analysis on a photograph of an unknown scene taken with an unknown camera [34]. As stated in the 5th edition of the Manual of Photogrammetry [34], in these types of cases "…the major reduction issue is the recovery of the interior orientation parameters of the camera," whereas in cases where forensic photogrammetry is used to do site surveys or object measurements with a known, metric camera, the methods used are the same as in industrial and architectural close-range photogrammetry.

Photogrammetry is often used in the analysis of traffic accidents, major disasters, and major crime scenes. In certain circumstances it may be a useful tool to facilitate comparison of a photograph of a suspect with the offender observed in surveillance crime footage, although identification by this means cannot be used in all cases. Images captured with Closed Circuit Television (CCTV) cameras are frequently difficult to evaluate because most images are of poor quality as well as being from significantly differing viewpoints in relation to the camera. For maximum identification accuracy, the photographs being compared should be from the same angle and direction [35] and clear enough to be able to locate the facial landmarks. It is also helpful to compare photographs taken in similar lighting conditions and similar time periods [28]. Identifications were found to be less accurate when measuring only four facial features [36], therefore it is advantageous to compare as many landmarks as are available on the photograph.

When comparing photographs, the images must be large enough to enable landmarks to be identified. In cases in which it is necessary for one or both of the photographs to be enlarged, Porter and Doran measured interpupillary distance as a way to ensure that the correct magnification was produced [37]. In their research, they took three photographs of the same subject who had different hair lengths and amounts of facial hair and used measurements, such as the horizontal face width between ear roots, to determine if the person in the three different photographs was the same. Their study outcome was successful in concluding that there was a high probability of the three individuals being the same person. Identification from photographs using facial anthropometry has been used in present day court cases with mixed results [38]. Halberstein used head length and height along with other landmarks in three cases in which he compared photographs of a perpetrator with the actual suspect [38]. Anthropometric proportions, along with a morphological comparison, were used to assist in identification resulting in convictions in two of the three cases mentioned.

Measurements taken from photographs play an important role when analysing the usefulness of photographs as evidence [38]. Photographic measurements are used to determine the height of the subject portrayed in video surveillance images [39], or, when taken from the face, are used for direct identification. Factors such as weight gain, facial expression, hair loss, and head position can all affect the outcome of the measurements. At present, facial measurements may be better used to exclude someone, rather than to make a positive identification. When the facial measurements of two people are similar, it can be stated that the two people *may* be the same but it cannot be stated that there is a positive match.

2.4 Morphology

Morphology is defined as the form and structure of something [1] and can be applied to a wide variety of uses, including identification.

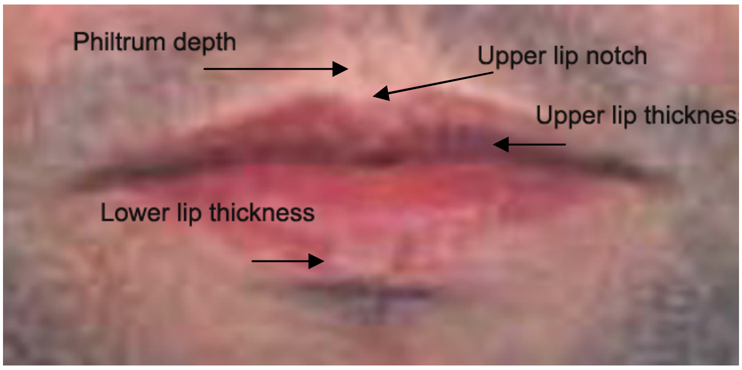
Morphological characteristics alone are not enough to identify someone, either living or deceased, but instead a comparison of at least two individuals must be carried out. Specific features distinctive to an individual will have a profound influence on the outcome of a presumed identification of that individual [4]. When the morphological characteristics of an individual are recorded, the height, weight, and physique are the features which are first noted. Also important to observe are the hair colour and length and the presence of any facial hair or other body hair. Racial and ethnic facial appearances, the colour of the individual's skin, and any unique characteristics such as tattoos, surgical scars, congenital deformities, tribal scars or markings, circumcision, moles, warts or other skin blemishes are photographed if possible. Attention is paid to the clothing worn as well as jewelry as they may provide information on the sex, race, occupation, or social status of the individual. However, it should be remembered that there are many men who dress in women's clothing and vice versa, so clothing may not be indicative alone. The age of the person may be approximated using the general appearance of the individual. Any loss of skin elasticity, hair colour, and changes in the joints due to arthritis are observed.

Presently, morphological analyses are conducted on numerous facets of the human body, including facial images [40], fingerprints [41], teeth [42], and ears [43]. Morphology has also been utilized to determine the sex of human remains, based on individual features of the skull, such as the nasal aperture, supraorbital ridge, and malar size [44]. In the comparison of facial images, morphological evaluations are conducted frequently as corroborative evidence to video superimposition and other techniques of facial identification. A Morphological analysis can be especially helpful when the individual in question has obscured their face in some way. A study by Yoshino et al (2002), attempted to find a reliable computer-based comparison method to aid in the identification of facially disguised individuals [45]. This study resulted in a system which proved reliable enough to identify disguised individuals and, just as importantly, the authors believed that the method and results would be easily understood in court.

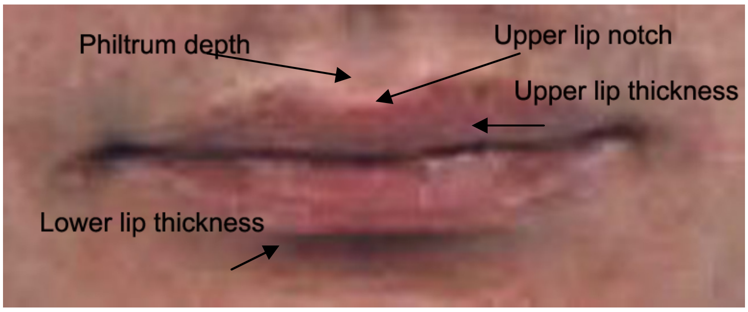
Morphological comparisons can be carried out on photographs or video images of suspects as a way of looking for similarities as well as differences. Lighting, camera angles, and camera distance can all affect the appearance of the suspect and make identification more difficult. A morphological comparison should be a part of any facial identification scheme, whenever possible, helping to supply additional evidence to support or disprove an identification. Useful features that help with identification are those which are present in only a small percentage of the population, for example, highly pronounced ears or a dimple on the chin. Identifying and distinguishing marks such as warts, moles, cauliflower ears, or scars are also especially valuable in the identification process.

A morphological description can contribute to identification of individuals; however, there are drawbacks to this type of method. Ïşcan notes that emotions such as anger, fear, happiness and surprise can significantly change facial features [28]. Changes resulting from disease, aging and exposure to the elements also affect morphology. Fat distribution also plays a contributing role in the changing morphology of the face [46].

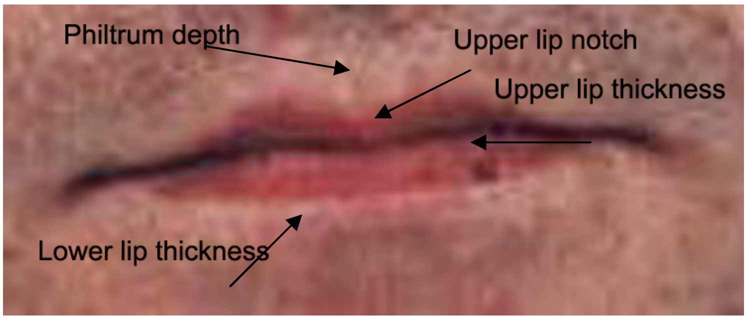
The photographs in Figures 2.5a-c demonstrate how varied a facial feature such as the mouth can be. From an examination of these photographs several aspects can be described, including the size of the top and bottom lips, the depth of the philtrum and the shape of the upper lip notch. Morphological analyses can be very subjective and there can be a fine line between what one person would consider a deep or shallow philtrum compared to another. It can be slightly less subjective if two facial features are directly compared but may become a grey area when facial features are described on their own. The description of the mouth in photograph (a) may indicate that the person has a deep philtrum, a wavy upper lip notch, a thin upper lip thickness and an average lower lip thickness. Obviously there are different degrees of thickness and this is where an individual's own perception plays a part in the description process. The mouth in photograph (b) could be described as having a deep philtrum, a wavy upper lip notch, a thick upper lip thickness and a thin lower lip thickness. Photograph (c) shows a much different mouth which could be described as having a shallow philtrum, an angular upper lip notch and thin upper and lower lip thicknesses.



(a)



(b)



(c)

Figure 2.1a-c Series of photographs showing morphological characteristics of mouths on three different Caucasian men

Once a facial analysis has been completed, with or without a successful identification, a report will need to be prepared which outlines the process as well as the outcome of the analysis. Any materials used, such as images, must be included in the report to provide a visual record of what was completed. Any similarities or differences are noted [24], including any distinguishing marks which cause the suspect to stand out from other individuals. The final conclusion is usually couched in terms of, “is definitely not the same person,” “could be the same person,” or “is likely to be the same person.” The condition and quality of the material should also be addressed enabling the reader to know if the evidence was recorded under poor conditions or if indeed the conditions provided crystal clear material. In the case of *R v Stubbs* [47], facial mapping was conducted, comparing photographs from a robbery with photographs of the defendant taken at the time of arrest. The expert witness concluded that there were no significant differences but many similarities in the nose, hair, lips and chin. A ruler was also held between two points on both photographs, which also showed similarities. In another case, a morphological comparison was completed on images of a suspected arsonist and a man caught on film trying to start a fire at the Manningham Ward Labour Club [48]. Facial features such as nose, eyebrows, and mouth were compared and the complete analysis included video superimposition but not anthropometry. When cross examined, the expert witness involved admitted that the quality of the video was poor and that the distances between the cameraman and subject as well as the lighting could also pose problems.

Lynnerup and Vedel studied suspects in a bank robbery and compared video footage from the bank with video of one of the suspects walking in a corridor and walking around outside in a yard [49]. They looked at the morphology and gait of the suspects and carried out a photogrammetry analysis. The authors realized that because the suspects wore helmets and loose fitting clothing only general body features could be compared, but knew that if even the general body features did not match there would be no point in continuing with the analysis.

Morphology, in conjunction with a computer program, can also assist in the creation of line-ups in an attempt to cut down on the number of police identification photographs that a witness may have to view as researched by Lee and Whalen [50]. For example, their computer retrieval system pulls photographs from the database based on physical descriptions given by the witness. Positive results were achieved by the researchers who chose ninety features (e.g. straight vs. curly hair, thin vs. thick neck) for their studies based on five criteria: “maximally informative; mutually orthogonal; observable and memorable; consistently coded by different people; and the number of system features minimized while still discriminating among police photographs [51].”

2.5 Superimposition

The techniques of photographic and video superimposition can be used for identification of both deceased and living persons. With the deceased it is sometimes used when no medical or dental records are available at the time, but there is an idea of who the person may be. A skull and a photograph can be superimposed as well as two photographs. Once the person’s identity has been determined, DNA or dental records, if available, should be examined for confirmation. In some cases, video superimposition on deceased persons is better used to exclude someone rather than to provide positive identification because it can definitely be stated that a skull and photograph may not be of the same person [52]. However, if they do appear similar it can only be stated that they could be the same person, due to the possibility that another skull of the same size may coincide with the photograph.

Although the comparison of skulls and photographs has been used for many years to identify historical and archaeological figures [53, 54], it was first introduced into the legal system in 1935 by Glaister and Brash in the Ruxton case [55]. Two dismembered bodies found in Moffat, Scotland, were thought after examination to be one younger woman and one older woman. Visual identification was not possible, as both bodies were far too decomposed, but they were thought to be the remains of Isabella Ruxton (Skull No. 2) and Mary Rogerson (Skull No. 1). Measurements of the crania and face were taken and compared. This revealed the face parts were different sizes but the size and shape of the crania were similar. Photographs of each woman were compared against the skulls. The photographs made available were enlarged to life size (Figure 2.6), using an object in the photograph to keep the correct scale. The skulls were positioned in four poses as close to the position of the heads in the photographs as possible and photographed at natural size (Figure 2.7). Outlines of the skulls were compared with outlines of the portraits. Finally, the skulls were photographically superimposed on the corresponding photographs in which the poses matched (Figure 2.8). This technique was not used alone for positive identification of the two individuals but was used in conjunction with other evidence. Because the superimposition of photographs on skulls was so new and the technique not rigorously tested, a positive identification was not concluded. It was known that however close a match was obtained between photographs and skulls, the Defence would question the identification in court. Therefore, the final conclusion given in court was that Skull No. 1 could not possibly be Mrs. Ruxton but could be Mary Rogerson and Skull No. 2 could not possibly be Mary Rogerson but could in fact be Mrs. Ruxton.

Figure 2.6 has been removed due to Copyright restrictions

Figure 2.7 has been removed due to Copyright restrictions

Figure 2.8 has been removed due to Copyright restrictions

Video superimposition, published by Helmer and Grűner in 1977 [56], further improved on photo superimposition. One advantage of this method is that the images can be compared quickly [57] and can easily be seen on a screen. Video superimposition has been used to identify such individuals as Dr. Eugenio Berríos Sagredo [58] and Joseph Merrick (The Elephant Man) [59].

Although there are different ways to set up the necessary equipment to conduct a video superimposition [60-63], materials usually include one video camera and one monitor for the original skull or photograph [52], a second video camera and monitor for the photograph being compared, a mixing unit allowing variable mixing of the two pictures, and a third monitor to view the mixed picture [54]. Video superimposition is particularly helpful if it is intended to match a surveillance video with a photograph of a known suspect because perpetrators often try to hide their appearance [45, 62]. With the mixing device, different sections, such as the ear or cheekbone, can be directly compared and viewed on the screen [62]. In cases where the video tape is of poor quality, video superimposition can be used for exclusion of possible suspects [64].

In order to establish identification, the two images in question should be from the same viewpoint or at least a very similar position [62]. When the original image is either a photograph or a skull, a digital camera can be used to obtain the image and display it on the monitor. Images taken from a surveillance tape are transferred to a video tape recorder. To start the process of identification, the two images are superimposed together. This confirms that the images are in alignment. The images are then faded out to further check alignment and line up facial landmarks. When both images are from living subjects, as opposed to one being of a skull, landmarks vary according to the position from which the image was obtained [63]. Landmarks on a skull which are usually aligned may include the nasal cavity, bite line and eye sockets [65]; eyebrows in the photograph are matched to the supraciliary arches of the skull and the uppermost part of the philtrum in the photograph is matched with the anterior nasal spine (subnasal point) of the skull [66]. When comparing a skull to a lateral photograph, according to Austin-Smith and Maples [65], special attention is paid to the curves of the forehead, the depth of the nasal bridge, the shape and projection of the nasal bones, the lower face and chin prominence, and the height of the vault. To compare frontal photographs Austin-Smith and Maples advise that notice is taken of orbital size and shape, the breadth of the nasal bridge, the width of the nasal aperture, the total facial length and width, the ratio of mid-face to upper or lower face length, and mandibular shape [65]. Once the images are in alignment, one image is wiped electronically over the other. This allows the viewer to focus on specific facial features. The image can be wiped horizontally, vertically, or diagonally [62, 65, 66]. The most difficult part of the procedure, and one of the most important, is matching the positions of the two items being compared [62]. Equally important is enlarging the antemortem photograph to the correct size [67].

There is some controversy concerning the reliability of video superimposition in making accurate identifications. While some institutions have used it in court as a means of obtaining a positive identification [62, 67], others have used it to corroborate additional evidence [55, 67, 68]. A false positive identification can easily be made if the image available is unclear or the facial position is not similar. Austin-Smith and Maples carried out a study to determine the reliability of superimposition [65]. Their goal was to determine if superimposition could be used alone as a means of positive identification. In the study, they compared 3 skulls to 97 lateral view and 98 frontal view photographs. None of these skulls belonged to people in the photographs and the study did not use the front teeth (anterior dentition). It is very useful when anterior dentition is available to compare against a smiling photograph. However, this is not always possible and comparison of the skull alone is often required. After comparison using video superimposition, their results showed 9.6% of the lateral comparisons were consistent fits and 8.5% of the frontal comparisons were consistent fits. When both lateral and frontal view comparisons were used only 0.6% of the photographs matched a skull. They, along with others, came to the conclusion that, whenever possible, several photographs should be compared against the original as this increased the chance for correct identification [64, 66, 69]. If this is not possible, corroborating evidence should be used along with the comparison to further validate the identification [63, 67, 68].

One study aimed to show that similar-looking crania have unique differences. This study involved an adult black female skeleton which had been found in Ohio. Measurements of the Ohio skull were compared to measurements of skulls in the National Museum of Natural History of the Smithsonian Institution in Washington, D.C to assess the likelihood of another cranium and mandible matching [70]. The Smithsonian collection contains 30,000 skeletons from all over the world. Of these, 52 matched the anthropological characterises of the Ohio case. To compare the skulls, the upper facial height (nasion to prosthion) and lower facial height (prosthion to gnathion) were measured and the ratio of lower height divided by upper height was calculated. This decreased the number of similar skulls to eighteen. To further reduce the number, the width of face (bizygomic breadth, zygion to zygion) and upper facial length (nasion to prosthion) were measured. The ratio of upper facial length to facial width was then calculated and once again this was compared to the Ohio case. The four closest metrical matches were then superimposed with a photograph of the skull found in Ohio. Although the ratios were all very similar, the skulls did not match the Ohio case photograph when superimposed. This demonstrated that even though skulls may be very similar, there are still differences that distinguish them, validating the use of video superimposition in this case and in general.

Photographs can be presented as persuasive evidence to jurors as providing a record of the events of a case. What jurors may not realize is that photographs do not explain the full story but give only a snapshot of the account. Photographs can also be manipulated into displaying information that is untrue, for example by using filters to change colours or adding or removing objects [71].

Although controversial, video superimposition, an aspect of facial mapping, has been used in UK courts of law both to convict criminals and exonerate the accused as supportive evidence as well as being the sole evidence against the defendant in some cases. In the case of *R v Stockwell* [72], facial mapping was used to compare photographs of the suspect in two robberies. This case admitted expert witness testimony provided the jury was told it was not bound by the evidence and was free to draw its own conclusions. The suspect was found guilty and a later appeal dismissed. Another case, *R v Clarke* [73], was one of the first court cases to include video superimposition as evidence. The judge allowed video superimposition as a new technique to be used in court as evidence. Closed circuit television (CCTV) images from a robbery were compared with police identification photographs of the suspect. One photograph was electronically wiped on top of the other, both vertically and horizontally, and it was concluded that they were the same person. The defendant was convicted and the verdict was subsequently upheld at an appeal.

The subjective nature of video superimposition as a form of identification was demonstrated in the case of *R v Anderson* [74], in which both the prosecution and the defence submitted expert witness testimony. The expert for the prosecution compared video images from a robbery with photographs of the defendant and concluded that it was likely both images were of the same person but that he could not be certain. The expert for the defence said that he could not fully exclude the defendant as the perpetrator but that the images were unlikely to be of the same person.

Video superimposition has also been used to aid in the identification of historical figures such as Dr. Josef Mengele [75]. The remains of the former Auschwitz concentration camp doctor were believed to have been buried in a cemetery in Embu near Sao Paulo, Brazil, under the name Wolfgang Gerhard, having passed away on February 7, 1979. Anthropological characteristics, such as sex, height, age, and any skeletal deformities were determined. Video superimposition was carried out on the exhumed body with the help of the complete skull and several photographs. The photographs available included two from the age of 27 in the frontal and lateral positions and three from an age of older than 60. Subsequent to both the anthropological analysis, which was consistent with information available on Mengele, and the video superimposition, it was determined that there was "…no room for doubts that exhumed skeletal parts are remains of Josef Mengele."

2.6 Facial Reconstruction

When a skeleton or skull is found, the primary goal is to find the identity of the person. An anthropologist, looking at the skull or skeleton, can identify characteristics such as race, sex and age. It is up to the police, identification specialists, and occasionally the public to do the rest. The police gather information on any missing persons fitting the anthropologic profile. If there is an idea of identity and antemortem images are available, video superimposition can be done. If there are no leads as to who the person is, facial reconstruction may be used [26, 47]. Facial reconstruction is not meant to provide an exact replica of the person, but is used to create enough of a similarity that when shown to the public it may jog their memory [26, 47, 76]. The method of identification employed depends on the evidence available. Facial reconstruction is used as a last resort when identifying factors such as DNA analysis or dental records are not available at the time.

2.6.1 Traditional methods

There are several ways in which facial reconstruction is carried out, but before beginning any method, as much information about the individual as possible is gathered from the skull, i.e. age, sex, and race. In the traditional method, a cast of the skull is made, or in some cases the actual skull is used, and clay is moulded on to it to recreate the face. Individual scientists will have a slightly different procedural order, but generally methods are similar. Karen T. Taylor, author of the book *Forensic Art and Illustration*, glues the mandible to the cranium as part of her skull preparation technique [77]. She places the skull on an adjustable stand and orientates the skull into the Frankfort Horizontal [2] position. She then places tissue depth markers in appropriate places. An important but often difficult task of completing a facial reconstruction is matching the cranial skeletal landmarks with their corresponding soft tissue landmarks [78]. Finally, the prosthetic eyes are secured and clay is used to build muscle groups and subcutaneous tissue and skin [77], never exceeding the tissue markers. It is at this stage that artistic merits come into play. Creating the mouth and nose utilizes measurements along with some artistic license. The width of the mouth is determined by measuring the front six teeth. To estimate the size of a Caucasian nose, the nasal aperture is measured at its widest point and then 10mm is added to obtain the total width. Finally, the ears and neck are also created by estimation [47].

In 1997, a reconstruction was completed at the University of Manchester on the skull of a garage fire victim [79]. The Staffordshire Police had an idea of who the individual was but identification was unconfirmed due to a lack of dental records. The skull needed reassembly and missing parts were remodelled based on a mirror image. The reconstruction was completed using the clay method and along with a DNA analysis, a positive identification transpired. To test their hypothesis that it was possible to achieve a correct identification even when parts of the skull were missing, a photograph of this reconstruction was posed in the same position and had the same hair as a photograph of the individual. Ten members of the public were shown the photographs, and 90% agreed that the reconstruction was a great or close resemblance.

Forensic drawings are used to re-create an image of an individual and can be completed from viewing the body or drawing from crime scene photographs [77]. This would be done when the body is in too bad a condition to be shown to the public. Two-dimensional facial reconstruction from a skull is also used to build an identity from scratch and is developed through studying a radiograph of the skull or studying the actual skull [77].

2.6.2 3-D Facial Reconstruction Using Computer Graphics

Facial reconstruction using 3-D computer graphics is another and more contemporary technique. A laser is used to replicate an image of the skull in digital form in a computer. The advantage of this is that once the original skull has been scanned with the laser it can then be stored safely away [76]. Another advantage is that it only takes minutes to make additional faces fitting the skull but with different soft tissue facial features [47]. Using a computer program to create a facial reconstruction takes about one to two hours, which is a considerably shorter time period than that of the sculpting method.

The Iceman, the oldest and best preserved mummy ever found at the time of writing (Figure 2.9), is estimated to be 5,300 years old and was discovered in 1991 in the Otzaler Alps of the Southern Tyrol. The body was found wearing goatskin leggings and a grass cape with a copper-headed axe and arrows nearby [80]. Ownership of the mummy was disputed between Italy and Austria but was eventually taken by Italy. A 3-D facial reconstruction was completed by the Department of Forensic Medicine and Science at the University of Glasgow using a stereolithograph of the skull taken from a CT scan [76]. A beard and moustache were added to the face afterwards using the computer program CD-fit™.

Figure 2.9 has been removed due to Copyright restrictions

The process of facial reconstruction from 3-D computer graphics is outlined as follows [76]. The skull is prepared by placing cotton into the eye sockets and additional holes are filled in with BLUE-TAC™. This is to prevent the laser going through the skull and loss of data. The skull is placed on a holder and rotated 360° horizontally under computer control. A laser is scanned over the skull and viewed by a video camera. The image is digitised and analysed by the computer producing a 3-D image. The operator selects a facial thickness from a choice of thin, normal, or fat. Forty landmarks are available to be placed at specific points on the skull image. Given the anthropological information of race, age, and sex, an average face is chosen from the database (example of a female shown in Figure 2.10). The same 40 landmarks are available to be placed on the image of the face. The operator aligns the skull with the face and the markers are rechecked. The 3-D image of the skull is superimposed on the face. The peg markers are matched and moved manually on the average face. The computer then reconstructs an image of the face on the skull. The final image can be printed out or photographed directly from the monitor and can be saved for future reference. This is especially helpful when identifying characteristics are known at a later date [26, 76]. The image can then be manipulated using a program such as CD-fit™, to add characteristics such as hair that make it more public-friendly.

Figure 2.10 has been removed due to Copyright restrictions

One problem facing artists using either method is estimating the size and shape of soft tissue parts such as the eyes, nose, lips, hair colour and style, eyebrow shape, and ears [26]. Additional clues that are helpful for recognition but not available on a skull are tattoos, scars, or wrinkles. Not only are there no clues obtained from the skull as to what these features should look like, they also drastically alter the appearance of the face and therefore influence the chances of it being recognised by the public. It is especially helpful if accessories are found with the body, such as glasses, or jewellery. Features that are unavailable from the skull but that have a great impact on identification, like skin colouring or hair style, should be made as ambiguous as possible in the reconstruction. Often, to the lay person, the finished reconstruction bears no resemblance to the actual individual, but to someone who knew the deceased, the reconstruction can be instantly recognizable. Completing and displaying multiple reconstructions, with varying hair styles and/or colouring may seem like a good idea initially, but in the end has been shown to cause confusion [77]. Research conducted by Helmer, Roohricht, Peterson and Mohr [82] found that skulls that had pronounced characteristics in relation to age, constitution or illness, created more agreement on resemblance, whereas agreed resemblance results were worse in cases of female and younger individuals.

Facial reconstruction was useful in a case occurring in 1991. An exhumation was carried out of a woman who had previously been thought to have passed away from natural causes, but in fact had been strangled [76]. Her identity was unknown as she had been homeless and an alcoholic. Her hair was available but the only description on hand was that she wore glasses. A 3-D computer facial reconstruction was developed and within a few days of circulation, a tentative identification was made. It was confirmed by comparison of mitochondrial DNA from the skull with DNA from her mother and sister.

Facial reconstruction, while in some cases justified and useful, was found not to be very helpful in cases such as the identification of victims of the Green River Serial Murderer [83]. The “Green River Killer” had been linked to 41 sets of remains since 1982. Most of the remains which have been found have been skeletonised, leading to problems in identification. Twenty four reconstructions or approximations were produced to help identify these individuals. Multiple reconstructions were created including clay reconstructions and drawings made from radiographs. None of them contributed towards identification of the deceased. This could have been because they were not seen by the necessary public or they did not display enough of a likeness and therefore were unrecognizable. A successful reconstruction is one that has enough information to lead to identification, no matter how closely the reconstruction resembles the individual before death. Research comparing different methods of facial reconstruction conducted by Stephan and Henneberg [84] found that only the sculpting method of facial reconstruction gave identification rates above chance at a statistically significant rate. They also concluded that facial reconstruction is not a method that could be used to exclude someone.

Facial reconstruction, part science and part art, has its believers and its sceptics. But the end result everyone is looking for is that a family gains peace of mind when their relative is recognized as a result of a reconstruction, no matter which method is used.

2.7 Eye Witness Identification

Subsequent to a crime being committed, identifications are made either by forensic evidence or with the help of eyewitnesses [24]. Numerous studies have been carried out to demonstrate that eyewitness accounts can be unreliable, unless the witness knows the perpetrator [85-87]. With CCTV as an increasingly common method of crime documentation, scientists must be prepared to use an identification technique that is accurate and reliable enough to stand up in court.

2.7.1 Police Line-ups

Most of the population is familiar with the idea of a typical police line-up, as it is a common feature of many films and television programmes. Traditional police line-ups are conducted with several individuals in a room, lined up side by side. The witness is located in a separate room behind a one-way mirror, along with a police officer investigating the case and possibly the defence solicitor. The line-up may or may not include the suspect. Each part of the line-up process has issues that may affect identification, including the choice of distracters, or known innocent parties, in the line-up. Luus and Wells have recommended choosing the distracters to match the description given by the eyewitness over choosing distracters that look similar to the suspect [88]. They also believe that all members of a line-up should have both similar physical features as well as differing features. The similar features among members would be features that the eyewitness had recalled while the differing features would be characteristics that the witness did not recall.

Another feature of the line-up process that may determine the reliability of the identification is whether or not the witness was provided with instructions at the start of the process stating that the suspect may or may not appear in the line-up [89]. In research conducted by Malpass and Devine, eyewitnesses to a staged crime were given various instructions for a line-up where the offender was either present or absent. Biased instructions implied that the witnesses were to choose someone, and the unbiased instructions allowed the witness to choose no one if necessary. When the offender was not included in the line-up, there was generally a high incidence of choosing someone, but the lowest rate of choosing someone came in the unbiased group. Wrongful identifications occurred most often when the witnesses were given biased instructions and when the offender was not present in the line-up.

The first idea for an alternative line-up was researched by Wells in 1984 [90]. His theory was to have the suspect absent from the first line-up in order to see if the witness would choose anyone. If the witness chose someone then their identification would not count and they would be discredited as a witness, but if they did not choose anyone, then they would be shown the real line-up with the suspect included. Another alternative idea for a line-up involved showing the witness each individual one at a time, instead of all at once [91]. The eyewitness is shown only one individual at a time and must make a decision of whether or not that is the person they are looking for before being shown the next individual. This helps to minimize the chance that the witness will pick one of the individuals just because he looks more like the suspect than any of the others [92]. However, Pozzulo and Lindsay showed that adults, more than children, were more likely to correctly reject a line up which did not include a suspect and that for sequential line-ups, the gap between child and adult correct rejections was increased compared with simultaneous line-ups [93].

An alternative to the traditional line-up, either sequential or simultaneous, is a video identification. A video identification may include short video clips of each person that are shown to the witness one at a time. The persons in the videos may turn both to the right and left in order to show the witness their profiles, much as in the traditional police line-up. Valentine and Heaton found in their experiment that even though 25 percent of the witnesses chose the suspect from a 25 photograph line-up, only 15 percent chose the suspect in a video identification [94]. However, they concluded that video identifications were fairer than line-ups and that wider use of video identifications has the potential to improve the reliability of eyewitness identification evidence [94]. Using virtual heads instead of photographs is another way of creating a line-up and research has been done using virtual heads that were three-dimensional and photogrammetrically-generated [95]. Subjects were shown video clips of the suspects' faces and then were shown line-ups consisting of virtual versions of the suspects' faces, to determine which allowed a more effective recognition tool for persons seen previously. Results from the study showed that although virtual heads show promise as an effective identification method, they are not as effective as photographs when it comes to recognition from video clips. However, it is believed that with a better computer program and more time spent generating the virtual heads, it is possible to increase the quality of the photogrammetrically-generated heads.

2.7.2 Mistaken Identification

Research has shown that it is not possible to rely on witness identification alone, even from a video let alone from memory. Mistaken identification played a significant part in the convictions of nine innocent men in the state of Virginia as noted from serious felony cases since 1980 [96]. This is why it is important to have a scientific method of identification to back up what a witness may report. Even victims who have experienced a violent crime close up have misidentified their attackers. Research has been conducted over the years to determine which factors may affect a victim's ability to make a correct identification. Highly attractive and unattractive women's faces were found to be more recognizable than the neutral faces that they were mixed with [97]. Along the same lines, distinctive faces were shown to be more recognizable than faces considered to be similar to a prototype [98]. Research into child versus adult identifications show that children are less likely than adults to make a correct identification from a line-up [93]. Individuals have also been shown to be better at recognizing their own race compared to people from another race [99], and this is described as own-race bias (OCB). Individuals were found to be less recognizable after examining photographs taken two years apart. A study conducted by Read, Vokey and Hammersley [100] demonstrated that when the two photographs appeared similar to each other, the duration of exposure to the photograph was positively related to the rate of recognition. However, when the two photographs were not similar to each other, the increase in exposure time either had no effect, or had an increased or significantly decreased recognition rate. Whether or not a suspect wears a disguise can also have an effect on the ability of a witness to give a correct identification. Research by Yarmey [101], in which some of the witnesses were informed beforehand that they would be asked to describe a suspect, showed that the witnesses had better recall of eye colour, hair colour, and hair length in a non-disguised suspect over the disguised suspect wearing a baseball cap and sunglasses. Yarmey also concluded that the accuracy of descriptions such as recalling the estimated age of the suspect were high with 97 percent, while height and weight estimations were less accurate with 60 percent and 44 percent respectively. When the witnesses attempted to recall the suspect after a delay of four hours, hair colour and colour of footwear were significantly better recalled than when describing the suspect immediately. However, when the witnesses attempted to recall the suspect immediately, complexion and height were better recalled than on the delayed test and Yarmey could offer no explanation for this outcome.

2.7.3 Psychological Facial Recognition

Eyewitness testimony can often be the principal evidence contributing to a conviction in a court case but in many cases this can condemn an innocent person [102]. Wells et al. state that eyewitness identification evidence is among the least reliable forms of evidence and yet persuasive to juries. Both the processing [103, 104] and the psychology of facial recognition have been studied for years by researchers and have shown accurate recognition is more likely to result when the person is familiar [105], even from low quality images [85, 86]. However, there is a high error rate when attempting to recognize unfamiliar faces [85]. This was further demonstrated in experiments conducted by Kemp, Towell and Pike [106], who showed that cashiers had a difficult time determining if customers matched the photograph on their credit card. The cashiers knew the experiment was happening and reported that they had spent a longer time checking the cards than normal; however, still slightly more than 50 percent of the fraudulent cards were accepted by the cashiers. These people were standing directly in front of the cashiers, illustrating the difficulty of trying to recognize a customer from a photograph.

The exactness of facial matching is also decreased by changes in viewpoint, but less so with changes in facial expression, and the use of colour photographs versus a grey scale image has had no effect on outcome [85]. Gait, body shape, and clothing had little to do with recognizing people as shown in experiments by

Burton et al [85]. This could have been due to the degree of familiarity that the subjects had with the individuals on the video. The experiment was conducted by university staff and the subjects consisted of university students who knew the experiment was to take place in the psychology department, and that the participants would probably be connected to the university.

2.8 Identification through Biometrics

Biometric based automated facial recognition systems are another approach used to compare offender and suspect images. Such systems are playing an increasingly important identification role and are used in airports, banks, and retail shops. Biometrics can be applied by three distinct groups: commercial, government, and forensic [107]. A biometric system can identify or verify the identity of a person by using human physiological or behavioural traits [108]. Fingerprint recognition is perhaps the best known biometric technology used today [109], but other common biometric characteristics that may be used to identify individuals are iris patterns, facial characteristics, speech, hand geometry, ear shape and body odour [110]. However, all vary with regard to usefulness and accuracy in performing the identifications.

Biometric systems for identification have been used previously on the public, such as in 2001 during Super Bowl XXXV in Tampa, Florida [111], but with the terrorist attacks in the United States on September 11, 2001 and the introduction of the 2001 US Patriot Act, biometric identification has been pushed to the forefront. The United States requires that any foreign passport issued after October 2005 should contain a biometric chip with the owner's facial characteristics, fingerprints, or iris scans in order to speed up border controls and make it harder to falsify passports [112]. The European Union has had difficulty meeting this deadline for various reasons, but until it does, travellers will need a visa to visit the United States. Biometric systems are used on personal computers and laptops to act as security devices [113] and awareness of biometric identification has become more common amongst the general public by being a regular feature in cinema films and television programmes.

An ideal biometric system according to Phillips, Martin, Wilson and Przybocki [114] displays four qualities:

1. "All members of the population possess the characteristic that the biometric

identifies, such as irises or fingerprints;

2. each biometric signature differs from all others in the controlled population;
3. the biometric signatures don't vary under the conditions in which they are collected, and
4. the system resists countermeasures."

A biometric system can be broken down into two components: the enrollment module and the identification module [109]. The enrollment module is the part of the system that trains it to identify a specific person and the identification module is the part of the system that recognizes the person. For example, in the enrollment phase of a facial recognition system, the person's face would be scanned and a template would be created from a digital representation. Included in the template may be the size and positions of the nose, mouth, and eyes. Templates of all individuals included would be stored on a database. During the identification phase, the biometric sensor scans the person and obtains the biometric used for identification, in this case the face, and the data is transferred into the same type of template used in the previous phase. This template is compared against each of the previously stored templates to determine if there is a match.

One concern facing many private citizens is the effect of biometric identifications on an individual's personal privacy. Hadley's report in 2004 [110] quotes a document on airport security in which the American Civil Liberties Union (ACLU) states that the organization "does not oppose using biometric identification techniques with a proven record of accuracy, such as iris scans or digital fingerprints, to identify and authenticate persons working in secured areas of airports." However, the ACLU is against using biometrics on all airline passengers because they feel this would be a violation of the Fourth Amendment of the United States Constitution, protecting the right of personal privacy. The ACLU believes it is possible that biometric data will eventually be abused in the same way that social security cards, used for identification purposes in the United States, are used in identity theft [110].

An advantage of using a biometric system is that the information necessary cannot be lost like a key or forgotten like a password [110]. Disadvantages include accuracy, cost, ease of use and a fear that information such as fingerprints will be stored elsewhere [115]. With any type of security system, there will be criminals who will try to circumvent it, necessitating the continuous overhaul of the system. In Kuala Lumpur, Malaysia, when a gang of carjackers realized that they would need the fingerprint of the car owner every time to start the car, they chopped off the driver's finger and took it with them [116]. This type of crime would indicate that biometrics alone might not be enough of a security deterrent; but that the biometric would need to be foolproof and it perhaps would be necessary for a live specimen to be scanned. Biometric security may be increased by having the person use a password in addition to the biometric [107] or using a system that requires multiple biometric characteristics for identification.

2.8.1 Security Biometric System: Facial Recognition

From a forensic perspective, facial recognition can be an important and powerful tool in the following scenarios [117]:

1. "Searching a crowd scene to pick out faces drawn from a given database of candidate images;
2. Picking out the best face match to a query subject from an array of previously acquired face images
3. Given an image of a criminal involved in a crime, producing corroborative evidence to support or reject the hypothesis that the person in the image is the suspect in custody."

Facial recognition, by use of a computer, is accomplished using measurements, location of facial features, or analysing the facial image as a whole. Facial recognition systems currently have a high error rate and at this stage should be used to decrease the number of possible positive identifications so that they can then be analysed by an expert human eye. Inaccurate results from appearance-based facial recognition systems can occur due to changes in lighting, the angle at which the image captured or a changing facial expression [118, 119]. The precision of facial recognition systems is being improved upon by introducing sampling techniques to account for these types of changes. Ideally, facial recognition systems should produce radically different images of different people, but similar images of the same person even in environmentally different conditions [120]. It is often the case that different faces appear similar when placed under the same conditions of lighting, expression or position; whereas the same face under these conditions appears different.

Advantages of facial recognition systems are that they are relatively easy to use and they work quickly. However, the conditions under which they are used must be ‘fast, inexpensive and right’ [121]. They are also unobtrusive, compared to other biometric identification systems, do not require the user to stand still or move a certain way and the systems are becoming less expensive [122].

One problem concerning facial recognition systems and indeed all biometric identification systems is the possibility of obtaining a false match or a false non-match. Typically, with a biometric identification system, the computer will assign a matching score between the input and the sample located in the database [107]. A higher score means a more likely match. The score must be equal to or greater than a pre-determined threshold score to be considered a match. To make the system more secure, the threshold score number would have to be increased; however, this causes the number of false non-matches to increase. To allow for an increased number of input variables, the set score would need to be lowered, but then the number of false matches would increase.

In theory, biometric identification systems seem like they would be very effective but the following are two examples in which results were not as hoped. A two year facial recognition project, in which 36 surveillance cameras were set up around Ybor city in Tampa, Florida, was deemed unsuccessful [123]. The cameras, along with facial recognition software, were supposed to recognize facial characteristics of felons and runaway children from a database of 24,000 police and family photographs. The project failed but the cameras were still staffed on Thursday, Friday and Saturday nights leading to arrests for fighting and illegal drug sales. In 1998, a facial recognition system was tested in Hertfordshire, England, at Watford Football Club and while initial results were not as positive as they wanted, officers were willing to use the system again [124]. The software used concentrated on the area of the face between the top of the eyebrows and the bottom of the chin and from one side of the temple to the other.

Not only is research being conducted to obtain an accurate facial recognition system, but research is also being done to protect the anonymity of innocent individuals who may inadvertently find themselves on a video tape that may need to be used to show other illicit activities [125]. Newton, Sweeney and Malin accomplished this by digitally changing the facial images of the innocent so that the facial recognition software cannot match people to their images captured on video. Initial research found that covering the eyes and nose, reducing the number of pixels in the face, or changing grey scale pixels to black or white values were ineffective in blocking the facial recognition software from doing its job. However, their '*k*-Same' algorithm averaged similar looking faces together in such a way that many characteristics remain yet the resulting image is not reliably recognized by face recognition software [125].

Research conducted in 2002 used a computer system to create facial composites using an interactive genetic algorithm to improve upon the e-fit technique of obtaining a likeness of an individual [126]. Among an initial sample of faces, the user would assign a "fitness" score to each one to establish how closely it resembled the target. Based on the score, the computer randomly selects additional members from the database to mutate and crossover, creating additional faces for comparison.

2.8.2 Pose Invariant Facial Recognition Systems

One of the most difficult challenges of image comparison occurs when two images with differing facial positions are compared. The chances of two photographs being in the same exact positioning are limited. Heads can rotate horizontally, vertically and diagonally, and it is unlikely that two images, especially if one is in motion as would be in a surveillance video, will ever be in the same position relative to the camera, which is necessary for the most accurate image comparison. Extensive research has been carried out and continues to be done into finding ways to address this problem successfully.

Facial recognition is a large and important part of the field of biometrics, but has met with limited success, in part due to the fundamental problem of facial position [118, 119]. Research has been conducted into developing pose invariant facial recognition methods in an attempt to address this. Two such methods are described below and both report greater recognition rates than when used without the pose transformations.

The first method uses eigenface analysis to extract features in order to recognize faces from a viewing range of 30º out of plane rotation from both left and right sides and then performs the facial recognition [127]. The researchers restricted their rotation range to ±30º because this allows both eyes to remain in the image as they use the location of the eyes to position the faces. In order to estimate the facial pose θ, they use $\theta = \arcsin(a / r)$, where a is the distance between the projection of the mid-point of the two eyes and the centre of the face and r is the radius of the head. From their results, they reiterate the fact that pose estimation is an essential component of the process and that the result of the pose estimation directly affects the rate results of the recognition. Although the researchers report success with this method, it may not be a favourable long term method to employ as it is limited to a rotational change in facial positions of ±30º which will only occur in a partial number of cases.

The second method of pose invariant facial recognition computes 3D images to estimate pose by initially distinguishing points on the ear to detect facial features [128]. In this method, the rough face contour and locations of nine facial features are then used to detect the ear points. As the rough face contour will change if the face changes due to a different hairline or weight gain or loss, it may also limit this method for comparison to be used within a relatively short time elapse. General anthropometry has shown the ear to be an ineffective feature for accurate measurement [27], most likely due to the points being on different planes. Chen and Cham acknowledge that ear points are not used in most face recognition algorithms but are important in their algorithm for estimating the seesaw rotation [128]. According to photographs in their article which show examples of detected ear points, the mark, placed by an algorithm, included the joint point of both the ear boundary and the face boundary, but did not appear to be consistently placed across examples, although this may be an effect of the photographs shown. This approach allows the inclusion of more people than if the super aurale had been used but still limits the number of people included if their hair covers the entire length of the ear. Positive attributes of this method are that it allows the recognition of faces at a large rotation and can effectively estimate both the seesaw and horizontal rotations of the test face.

2.9 Video Surveillance

There are no reliable estimates of the number of surveillance cameras watching over the UK. However, as stated by the BBC, there are up to 4.2 million cameras in Britain [129]. According to the Glasgow City Council website [130], accessed in 2007, there are over 400 CCTV cameras placed in the city centre and residential areas of the city of Glasgow which are monitored 24-hours a day, 365 days a year.

Society is not only monitored by video cameras but by automatic number plate recognition systems, shop RFID tags, mobile phone triangulation, store loyalty cards, credit card transactions, oyster cards, satellites, the electoral roll, NHS patient records, personal video recorders, phone-tapping, electronic listening devices and hidden cameras, worker call monitoring, and cookies.

2.9.1 The Role of Video Surveillance in Crime Detection/Prevention

Video surveillance is often the last or possibly only “witness” to a crime. An example of this was when surveillance video was responsible for recording the abduction of eleven year old murder victim Carlie Brucia in Sarasota, Florida, leading to the arrest of a suspect when a police informant identified the man in the video [131]. In 1999, over a million surveillance cameras watched over the UK [132] and research has been conducted specifically in Liverpool city centre to determine the effects of CCTV [133]. It is also increasingly common for mobile phones to be used by the public to capture a crime being committed and in some cases the criminals will video themselves committing offences.

Cameras are becoming more sophisticated in parallel with improvements in technology. This is evident in an example involving a failed attempt to disable a camera that was placed in New Orleans. A man shot the camera with a paintball rifle, hoping to rid the neighbourhood of police surveillance, but instead was arrested when his image was saved on the camera’s hard drive [134]. As it turned out the man was wanted on a murder warrant.

However, although rapidly improving, the majority of video surveillance equipment does not produce the crystal clear images needed to provide irrefutable identifications. It is in these kinds of cases where identification techniques such as anthropometry would be helpful. Information obtained from surveillance tapes can be important supportive evidence because it may show a crime being committed, although it is not always easy to recognize and therefore convict a criminal using CCTV evidence alone. Video surveillance may be more reliable than eyewitness testimony because the story it tells is consistent and also corroborates with what the eyewitness saw [135]. However, a more comprehensive analysis is necessary because even when video images are clear, two people can look very similar to each other. Furthermore it is also easy and common for individuals to disguise themselves.

It can be difficult, if not impossible, for surveillance operators to keep constant attention on cameras when they have to keep track of such large numbers of them. New Scientist reported about software that has been developed which analyzes CCTV footage and determines if further investigation is needed [136]. This system has the potential to find suspicious packages or spot a possible suicide attempt. The system has been tested in London Underground stations and works by comparing images from the CCTV cameras against images from the same empty station. The images are analysed to determine what objects differ between the two. The system does not know what the images signify, but it can alert an operator who can investigate further, thus allowing selected cameras to be actively monitored.

2.9.2 Positive and Negative Aspects of Video Surveillance

The amount of publicity given to CCTV has led the public to believe crimes can easily be solved from video alone when, in reality, the video produced is often blurred and precise image details usually nonexistent. Several factors can influence the quality of images produced from these surveillance systems [137] and of greatest assistance to law enforcement include placement of cameras and lenses, recorders, storage space, and compression schemes [138]. The visible area is greatest when a camera is placed at a high level in relation to the subject. However, close detail of an individual person may be impossible to see, rendering the footage useless. Cameras placed at low levels may show more detail of a person, but the camera view has a greater chance of being obstructed and the camera has a greater chance of being vandalised. In order to get the most constructive information for surveillance, businesses must decide what their objectives are when setting up a surveillance system. Options are to concentrate on employee theft prevention or to protect their employees and customers from unknown individuals and placement of the cameras depends on the intention. Obviously, if more cameras are placed, a greater area is covered and the more likely the chances of identification, keeping in mind that if a camera is placed too far away, there will not be enough detail to determine identity. The type of lens on the video camera can also affect what is captured on videotape. A wide-angle lens may allow a large area to be seen, but it can also distort the picture. A telephoto lens, while able to obtain images close-up, may also affect the image outcome.

In an effort to improve the quality of an image, some systems allow the operator to sharpen the edges of an image or to enhance all edges except where the camera detects flesh tone. This may result in misleading images, which can be questioned by the defence in court. Analysts who enhance images must be able to explain to the court exactly what was carried out in a manner that is understandable to everyone [71]. Their skills must also include the ability to reproduce any changes made. The more images captured on videotape, the greater chance of obtaining a sequence of events useful to the investigator [137], as pertinent information can be lost using time lapse recordings [85]. A further problem which may impede identification is that frequently, due to the high cost of storage and tape replacement, most images are discarded before they are viewed [137].

Introduction of digital recording has helped to overcome some of the problems caused by video surveillance systems. Businesses using VHS taped video surveillance systems may constantly reuse the tapes to save money, which over time causes the images to degrade. Analog video systems are limited to a certain resolution and enlargement of a portion of the image will not add any additional detail [135]. Presently, more systems are switching to digital, which have distinct advantages and disadvantages. Advantages include a better quality of image that is produced and the ease of using the system. A much faster retrieval process is present to look for images on a digital system [139] and when viewing long periods, can save valuable time. An important disadvantage is that when the images are exported, they are compressed and essential details may be removed.

2.10 Forensic Image Analysis

Forensic image analysis includes photogrammetry, photographic comparison, content analysis and image authentication [39]. The process of analyzing images for forensic purposes falls into three categorical tasks: interpretation, examination and technical preparation [39]. The first, interpretation, involves deriving conclusions based upon what is seen in the images. The second, examination, uses image analysis techniques to extract information from the images. The third, technical preparation is the preparation of evidence or images for examination.

2.10.1 Resolution

The resolution of an image refers to the number of pixels in that image [140]. Resolution can also refer to the height and width of the image as well as the number of pixels in the whole image. In digital cameras, the term megapixel refers to the number of pixels in the whole image. Digital cameras with a larger number of megapixels produce larger images.

Image size can be changed by adjusting the number of pixels and the quality of an image can be changed by compressing it [140]. This reduces the size of the file making it able to load faster. Uncompressed images are saved, for example, as a bitmap (BMP) or Tagged Image File Format (TIFF) file, whereas compressed images are saved, for example, as Joint Photographic Experts Group (JPEG) files. A compressed image loses detail and decreases the quality of the image. If images are going to be compressed to make the file smaller, the uncompressed image should always be saved as a master copy [140]. Also, the process of compression is iterative, so that opening and editing a JPEG file and then re-saving it in JPEG format results in the progressive loss of information and image detail.

2.10.2 Digital Image Enhancement Guidelines

Digital images are commonly used in forensic practice to record evidence at crime scenes and record evidence of victims. Often it is necessary to enhance the images, such as those obtained from a poor quality video surveillance system. However, it is easy to manipulate images using software such as Adobe Photoshop® and it is important that when digital evidence is used in court that everyone involved in the justice system realizes this and knows the process and methods that are used to prevent any changes to the image. In 2002 the Scientific Working Group on Imaging Technologies (SWGIT) developed guidelines which should be followed when using digital images [141]. Key features include:

- "The original image is preserved;
- The processing steps are logged when they include techniques other than those used in a traditional photographic darkroom;
- The end result is presented as an enhanced image, which may be reproduced by applying the logged steps to the original image;
- The recommendations of this [SWGIT Version 1.2, June 2002] document are followed."

Images are segregated into two categories depending on their intended uses [141]. Category One images express what was recorded but is not analyzed by experts. Examples of these are general crime scene images, surveillance images and arrest photographs. Category Two images are analysed by experts. Examples of these include latent prints, questioned documents and Category One images to be analyzed.

Before the evidence is photographed, careful consideration is given to positioning of the evidence, camera, or measurement reference scale which can all affect the image produced and if done incorrectly can create a perspective and parallax distortion [142]. Accepted enhancement techniques include brightness and contrast adjustment, colour balancing, cropping, and dodging and burning [141]. The use of any enhancement techniques performed on Category Two images should be documented [143] as well as the sequence in which they were carried out. However, any exploratory steps that were taken and not used in the final analysis do not need to be documented. Requirements for documentation include noting the software and techniques used, along with their settings and parameters. The process of spotting, which removes blemishes due to dust or scratches on traditional negatives is not an accepted enhancement technique used on forensic images. The final report, based on recommendations by SWGIT, written by the operator who performed the analysis, includes a review of the materials received, the request, the methods used, the results obtained, an estimate of accuracy and precision, the basis for the conclusion, and the conclusion [39].

Evidence that has been photographed incorrectly, causing distortion of the image, may lead to the evidence being analyzed incorrectly. The detrimental effect this can have to a case is obvious. Bowers and Johansen conducted research into repairing certain types of distortion with Adobe® Photoshop® [142]. They found that using Adobe® Photoshop®, they could reliably repair Type I distortion, which occurs when the reference scale in the photograph is in the same plane as the evidence sample (such as a bite mark) but the camera angle is not perpendicular to that plane.

2.11 Police Photographs ("Mug Shots") in the United Kingdom

2.11.1 Trends from 1989

"Mug shot" photographs are taken at each of the 36 local police divisions in Scotland, depending on where the individual is arrested [144]. There is no standard way that the photographs are taken and they are taken by photographically untrained police officers. Once brought into the police station, suspect photographs are taken as part of a series of identification procedures including DNA sample, fingerprints, footwear impressions, hair combing results, height, weight, distinguishing marks and anything else that can be used to identify the person, and all information obtained is entered into the Strathclyde Police database. Each person is assigned a number that will be used for the current and any future arrests.

Until 1989, the process of taking mug shot photographs was carried out in the traditional way, that is, with a front facing view and a profile facing view. In 1989, this was changed to taking a single ¾ facial shot as this direction was shown to be more beneficial for identifications because it is more common to view someone that way. The photographs are taken with film, processed and then scanned digitally into the computer.

When the photographs are taken, the person is always sitting down to avoid any issues with height. As their height is taken separately, it does not need to be recorded in the photograph. The camera is not set at a specific distance from the subject but the lens is sealed and the f-stop setting and shutter speed of the camera are fixed so that nothing can be changed. The lens used has a focal length of 52 mm because this is what the eye sees and therefore the photo is not distorted. The individual is asked to sit down and is instructed to look over at something in the room in order to the give the camera a ¾ view of their face. One photograph is taken and because film is used, it is not possible to retake poor photographs. Also, because of the nature of the situation, the individuals are not always compliant. Therefore, the photographs produced are not uniform and can be of variable quality, for example, the eyes may be closed or the facial orientation may differ from the ¾ view requested.

2.11.2 Forensic Identification National Database (FIND)

A pilot study [145], is currently being carried out by the Police Forces of Lancashire, West Yorkshire and Merseyside, which supply data and images for a national database. Other participating police forces in Devon, Cornwall, British Transport Police (BTP) North Eastern Region Greater, Manchester (GMP) and North Wales Police have read only access [145]. The project is called *FIND* which stands for "Forensic Identification National Database", which allows the police to search, retrieve, store and transmit facial images and/or video images. These images of arrested individuals can be referenced against criminal data held on the Police National Computer (PNC). The images and their descriptive data can be easily removed or re-classified as deemed necessary by the Police. Each force obtains their own images, stores them centrally and can send them to the central National repository held at the Hendon Data Centre (HDC). Forces can access an image immediately which is beneficial when investigating the high level of cross border crime.

FIND will initially be used by the police forces of England and Wales, the British Transport Police, the Channel Islands and the Isle of Man Forces and police forces in Scotland will be included at a later date. The database will also be available to other accredited agencies on a view only basis once they are granted permission.

What makes this database different from previous databases in the UK is that the photographs included are taken in a standardized way. The photographs are taken with a standard backdrop, standard lighting (two lamps) and photographs taken are frontal, profile and a ¾ view which are to be taken simultaneously. The Police standards for Still Digital Image Capture and Data Interchange of Facial/Mugshot and Scar, Mark and Tattoo Images were created and endorsed by the Association of Chief Police Officers (ACPO). Version 2.0 of the standards was released in May 2007 [146].

The standards fall into the following categories [146]:

1. "Scene Requirements
2. Photographic Requirements
3. Digital Requirements
4. Image compression and Interchange Format Requirements
5. Image Data Guidelines"

One example of the way this database may be used is that subsequent to an eyewitness giving a description of an individual to the police, a randomly selected group of photographs based on this description can be chosen from the database to be shown to the witness, which may save time and additional stress to the witness from having to look at books full of mug shot photographs. It is expected that the FIND project will come into service in 2009.

In conclusion, there are many aspects of forensic identification that are put into practice today and more to be researched. From the first methods, such as anthropometry and fingerprinting to the recent advances in biometrics, experts have been trying to find ways to find missing persons and identify the victims and perpetrators of crime. Anthropometry, first used in the 1800's, has since been replaced with more discriminating methods of identification. It is still used today but the varying results it achieves indicate that more research is necessary to ensure a reliable forensic identification method.

3 Facial Landmarks

3.1 Introduction

Eleven unilateral and 13 bilateral, totalling 37 landmarks were chosen for inclusion in subsequent anthropometric studies. Bilateral landmarks are located on both sides of the face. From these landmarks, a total of 73 linear measurements and 59 proportions (also unilateral and bilateral) were selected for comparison of images. This chapter is separated into three sections covering individual landmarks, linear measurements and proportions, and explains the choices made and the reasoning behind those choices.

3.2 Landmarks

Figure 3.1 Facial landmarks and their location

The definitions of landmarks given below are adapted from the literature and are given in Table 3.1. Their locations are shown on 45° and frontal view photographs in Figure 3.1.

Table 3.1 Definitions of Landmarks Used in this Study

1. Glabella (g): the most prominent midline point between the eyebrows.

2. Nasion (n): the point in the midline of both the nasal root and the nasofrontal suture. This point is always above the line that connects the two inner canthi. A canthus is the angle at either end of the fissure between the eyelids.

3. Exocanthion (ex): the point at the outer commissure of the eye fissure. A commissure is the site of union of corresponding parts and a fissure is any cleft or groove, in this case of the eye [bilateral].

4. Endocanthion (en): the point at the inner commissure of the eye fissure [bilateral].

5. Palpebrale superius (ps): highest point in the midportion of the free margin of each upper eyelid. The free margin portion of the eyelid is the unattached edge. [bilateral]

6. Palpebrale inferius (pi): the lowest point in the midportion of the free margin of each lower eyelid [bilateral].

7. Orbitale (or): the lowest point on the margin of the orbit. The orbit is the bony cavity that contains the eyeball [bilateral].

8. Superaurle (sa): the highest point of the free margin of the auricle. The auricle is the portion of the external ear that is not contained within the head [bilateral].

9. Subaurale (sba): the lowest point on the free margin of the ear lobe [bilateral].

10. Postaurale (pa): the most posterior point on the free margin of the ear helix. The helix refers to the coiled structure of the ear. [bilateral].

11. Otobasion inferius (obi): the lowest point of attachment of the external ear to the head [bilateral].

12. Alare (al): the most lateral point on each nostril contour [bilateral].

13. Subnasale (sn): the midpoint of the angle at the columella (fleshy, lower

margin) base where the lower border of the nasal septum and the surface of the upper lip meet.

14. Pronasale (prn): the most protruded point of the nasal tip.

15. Subalare (sbal): the point on the lower margin of the base of the nasal ala where the ala disappears into the upper lip skin [bilateral].

16. Stomion (sto): the imaginary point at the crossing of the vertical facial midline and the horizontal labial (lip) fissure between gently closed lips, with teeth shut in the natural position.

17. Crista philtri landmark (cph): the point on the elevated margin of the philtrum just above the vermilion line. The philtrum is the vertical groove in the median portion of the upper lip and vermilion refers to the exposed red portion of the upper or lower lip [bilateral].

18. Cheilion (ch): the point located at each labial commissure [bilateral].

19. Labiale inferius (li): the midpoint of the vermilion border of the lower lip.

20. Labiale superius (ls): the midpoint of the vermilion border of the upper lip.

21. Gonion (go): the most lateral point at the angle of the mandible. The mandible is the bone of the lower jaw [bilateral].

22. Sublabiale (sl): determines the lower border of the lower lip or the upper border of the chin.

23. Pogonion (pg): the most anterior midpoint of the chin.

24. Gnathion (gn): the lowest point in the midline on the lower border of the chin.

Careful consideration was given to the selection of landmarks that were to be used in the comparison of faces. When choosing a landmark it was important that it was one that could be placed consistently. It had to be a point where an operator performing the comparison would be able to locate it in the same place within an acceptable error. According to Fieller [147], the criteria used to determine a successful/reliable landmark are:

- observer knowledge;
- consistency of landmark placement;
- discriminatory power;
- landmark visible in majority of cases.

Initially, studies by Farkas [27] and Purkait [148] were consulted in choosing the landmarks that were selected for this research. In both of these studies, the reliability of landmarks was tested by comparing direct measurements on the face to indirect measurements made using photographs. In both studies, the researchers marked the landmarks on the face with black dots before taking the photographs and used these to make the measurements. The advantage afforded by this approach was that it is much easier to locate landmarks on a person rather than on a photograph because it is possible to feel for bony indications of the landmarks. In a real case scenario in which two or more photographs are being compared, as in the present study, this approach could not be used.

A more detailed consideration of the studies by Farkas [27] and Purkait [148] is pertinent to the present research study.

Farkas studied anthropometric measurements from photographs (indirect) vs. anthropometric measurements taken from a person (direct) [27]. Landmarks can be more difficult to find when they are being located on a 2-D surface, rather than on a live person [45]. Farkas used 36 North American Caucasians in his study, comprising 18 men and 18 women, and 100 direct linear measurements between two landmarks were taken from the head, face and ears [27]. In addition, 60 measurements were obtained from photographs taken from frontal and lateral views; landmarks were marked previously on the skin with a pen. Direct measurements were compared against indirect measurements and any differences were noted. Measurements were regarded as reliable if the differences between them were no more than 1mm or 2 degrees. Forty of the direct measurements involved landmarks that were not visible because of the two dimensional nature of photographs, or were covered by hair, which prevented them from being obtained from the photographs. Of the indirect measurements, 20 were reliable, of which 7 involved the lips and mouth, none used the ear, and 9 were inclinations, small areas of the face joined at angles.

Farkas observed that the glabella can be blocked by bushy eyebrows and the trichion can by hidden by hair [27]. The trichion is the point at which the median plane of the forehead intersects with the hairline. He also noticed that the greatest distortion was the measurement between the subnasale and the pronasale because of the difference between the planes that the landmarks are located on. However, the measurement between the root of the nose (nasion) and the labial fissure (stomion) was found to be precise because the two landmarks were closely related to the plane of focus. Farkas also noted that frontal view prints supplied the most accurate measurements of orbits, lips, and mouth but that profile prints were more useful because they provided a number of vertical measurements that when compared to direct measurements fell within the accepted error as well as exhibiting accurate inclinations of the facial profile. The opportunity to place landmarks on the skin for subsequent measurement before photographs are taken will not likely have a place in forensic cases. Placement of landmarks will most probably occur on photographs.

The study conducted by Purkait consisted of measurements, i.e. distances between landmarks, taken from 17 landmarks in total. The chosen landmarks are listed in the table below. All of these landmarks, with the exception of the tragion and zygion, were used for the research conducted in this thesis. Reasons why the tragion (the point in the notch just above the front of the external opening of the ear) and zygion (most lateral point of the cheekbone) were not used are given section 3.2.

Table 3.2 Landmarks Used in the Study by Purkait [148]

• Exocanthion (ex)	• Postaurale (pa)	• Labrale superior (ls)
• Endocanthion (en)	• Superaurale (sa)	• Labrale inferior (li)
• Orbitale (or)	• Subaurale (sba)	• Gonion (go)
• Zygion (zy)	• Gnathion (gn)	• Subalare (sbal)
• Subnasale (sn)	• Cheilion (ch)	• Otobasion inferior (obi)
• Tragion (t)	• Stomion (sto)	

Purkait chose measurements in which the landmarks were located on the same plane, with the exception of the zygion, which was used to check the reliability of the measurements related to the landmark. Purkait's study compared direct to indirect measurements and the camera used to take the photographs was placed at a fixed distance of 80cm from the subject and the horizontal axial plane of the camera lens in frontal image was focused to the nasion. It was found that indirect measurement values were consistently larger than the same dimensions measured directly on the face and this was attributed to a flattening of features when projected on a two dimensional surface (the photograph) and therefore the landmarks are 'pushed out' laterally. However, distances that exhibited larger values obtained from direct measurements include those measured directly around the eyes (en-ex) and mouth (sto-ch). Purkait explains this as result of the soft tissue of these two areas being manipulated during the measurement process.

In Purkait's study of direct measurements, landmarks were considered to be variable or to change significantly with minor alterations in facial expression if they produced measurements with a percentage error above 2% in intra-observer study or above 4% in inter-observer study. Examples of these measurements included en-en (R&L), sto-ch (R&L), t-pa (R&L), t-go (R&L), and obi-go (R&L). A more detailed inter-observer study was then carried out on two male subjects to compare the variation in data when the same measurements were repeated 10 times. The mean and standard deviation were calculated for each measurement. The distances between landmarks that showed statistically significant differences between means obtained by the two measuring operators included the orbital, gnathion, cheilion and gonion landmarks.

A collaborative study funded by the FBI, Chief Investigator Evison, was also used as a guide in the choice of landmarks [16, 147]. In this study, 61 landmarks were originally chosen for consideration but a final selection of 30 landmarks was used. For Evison's research, two operators placed landmarks three times on 35 faces. The landmarks were also tested to determine if certain ones were more difficult to place for operators who were new to the field by having two experienced and three new operators measure five faces three times. Facial landmarks used for the research in this thesis that were eliminated from Evison's original 61 landmarks were the gonion, gnathion, subnasale and orbitale.

Research on facial recognition was consulted when considering which landmarks would be included in the present study. Craw, Costen, Kato and Akamatsu chose 34 landmarks for their investigation into automatic face recognition [149]. They termed their landmarks as 'true' or 'deficient'. An example of a deficient landmark was described as, "the edge of the chin 'half way' between two true landmarks." Craw, Costen, Kato and Akamatsu avoided the use of landmarks whose location might change over the long term, such as around the hair or hairline. They instead focused on points on the smaller part of the face where changes are more unvarying. Okada, von der Malsburg and Akamatsu documented the use of 20 facial landmarks in their face recognition system research [150]. Sixteen of these landmarks, apart from four located on the eyebrows, were used for the research in this thesis. Reasons for not using the eyebrows are mentioned further on in this section.

Although small numbers of landmarks do not provide enough data to carry out a facial identification, there are four landmarks in particular which may be more beneficial than others. The four landmarks include the right and left ectocanthions, the nasion and the stomion. They are located in the middle of the face and, because of this, remain relatively fixed and do not alter with a changing facial expression. The exception to this is the stomion, which may change with expression or conversation. These landmarks are visible whilst wearing a hat and are not affected by changing hairstyles or hairlines. They are also relatively easy to locate on a photograph. At the beginning of this work, these four landmarks were used for pilot studies designed to evaluate the potential of anthropometry for use in facial comparisons.

Landmarks that were excluded for the present study were eliminated on the basis of their ability to be located on photographs. For example, the tragus (t), located on the ear, would be very difficult to spot on a low resolution video image that may have been enlarged, since simply increasing the size of the image does not improve its quality. Landmarks on the head such as the euryon (eu), frontotemporale (ft), frontozygomaticus and (fz) and condylion laterale (cdl) are identified by palpation, which would not be possible when placing the landmarks on two dimensional images. Landmarks that rely on hair, such as the superciliare (sci), trichion (tr) and vertex (v) are of little value because hairlines change with time and the age of the subject and so they may not produce reproducible results when images taken at different times are compared. A good example is the superciliare, the highest point of the eyebrow, which cannot be identified reliably in subjects with plucked eyebrows. It would also be difficult to locate any point on the eyebrow on a low resolution video image or one that is produced from a camera located far away, especially if the subject has light coloured hair. The vertex, i.e. the crown of the head, may change depending on what hairstyle the subject is wearing. In addition, the opisthocranion (op) is located on the back of the head and is therefore of no use when trying to identify someone based on their facial features. Lastly, the zygion (zy), according to Kolar and Salter [151], is the most lateral point on the zygomatic arch, or cheekbone, and is not a fixed point but is identified by the maximum facial breadth as measured by the spreading calliper. This is also not practicable when comparing two dimensional images.

3.3 Linear Measurements

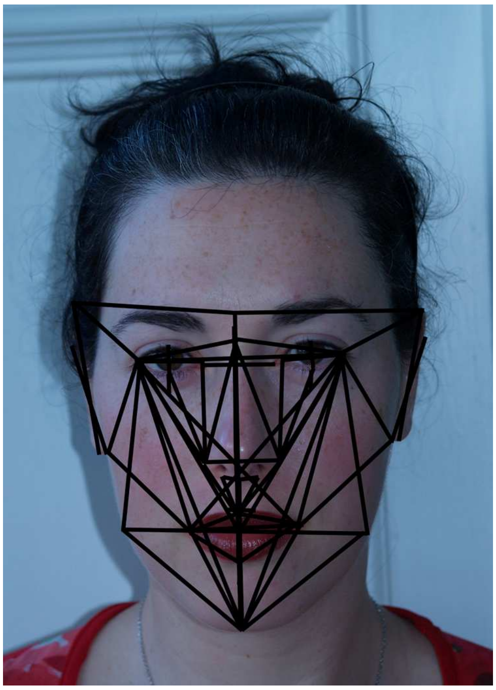


Figure 3.2 Linear Measurements

Linear measurements are the distances between facial landmarks and were chosen in order to create the proportions used to compare images. In principle, it is possible to produce 37 x 36 (i.e. 1332) distances using the 37 landmarks selected for the present study. While this is computationally possible, practical requirements dictate that a subset of these should be selected, especially if each linear measurement in turn becomes one of a pair used to determine proportions or angles, when 1332 x 1331 (i.e. 1,772,892) combinations would be possible. The 73 linear measurements selected are shown in Figure 3.2 and are listed below. They are unilateral unless noted as bilateral.

Table 3.3 Linear Measurements used in the Present Study.

1. go-go: Mandible breadth	25. li-sl: Cutaneous lower lip height
2. gn-go: Mandibular body length [bilateral]	26. sto-sl: Lower lip height
3. n-gn: Morphological face height	27. sa-sba: Ear length [bilateral]
4. n-sto: Upper face height	28. g-pg
5. sn-gn: Lower face height	29. pa-obi: [bilateral]
6. sto-gn: Anterior mandible height	30. ex-n: [bilateral]
7. sl-gn: Chin height	31. ex-sto: [bilateral]
8. ex-go: Lateral facial height [bilateral]	32. Sa-ex [bilateral]
9. en-en: Intercanthal width	33. Ex-obi [bilateral]
10. ex-ex: Binocular width	34. Obi-ch [bilateral]
11. en-ex: Eye fissure length [bilateral]	35. Ex-ch [bilateral]
12. ps-pi: Eye fissure height [bilateral]	36. Ex-al [bilateral]
13. pi-or: Lower eyelid height [bilateral]	37. En-al[bilateral]
14. al-al: Nose width	38. Ch-gn [bilateral]
15. sbal-sn: Nostril floor width [bilateral]	39. Pi-al [bilateral]
16. sn-prn: Nasal tip protrusion	40. Ch-ls [bilateral]
17. n-sn: Nose height	41. Ch-li [bilateral]
18. n-prn: Nasal bridge length	42. Al-ls [bilateral]
19. cph-cph: Philtrum width	43. n-al [bilateral]
20. ch-ch: Labial fissure width	44. ex-gn [bilateral]
21. sn-sto: Upper lip height	45. sbal-ls [bilateral]

22. sn-ls: Philtrum length
23. ls-sto: Upper vermilion height
24. sto-li: Lower vermilion height
46. Rt ex-lt ch [bilateral]
47. g-sa [bilateral]

Although the number of possible linear measurements increases exponentially with the number of landmarks, not all are reliable or pertinent to the research undertaken for this thesis. The list of linear measurements compiled by Kolar and Salter includes measurements used by earlier authors including Broca (1879) and Farkas (1981) [151].

Measurements 1 through 28 in Table 3.3 include all measurements from Kolar and Salter which are based on landmarks chosen for this research as listed in Table 3.1.

Measurements numbered 32 through 44 in Table 3.3 were obtained from Craw et al [149] and measurement number 45 was found to be reliable by Farkas [27]. Measurements 29-31 and 46-47, all bilateral, were unique to the present study. Two of these measurements, ex-n and ex-sto, were chosen because they utilize landmarks that were considered to be less affected by facial expression than others and also because they would be visible even if the subject was wearing a hat. Measurement, pa-obi was chosen in order to get some semblance of the width of the ear. Normally, the width of the ear is measured from the preaurale to the postaurale; however, the preaurale is a landmark that was considered to be too difficult to locate on a video image and was therefore not used. Instead, the otobasion inferius was the landmark chosen to be included in the linear measurement with the postaurale.

3.4 Proportions

A total of 59 proportions were selected from the 73 x 72 (i.e. 5256) which were possible using the linear measurements in Table 3.3 and are listed below (Table 3.4). Twenty two of the proportions are bilateral and are so noted.

Table 3.4 Proportions selected for the Present Study

1. go-go/n-gn	2. sn-sto/sto-sl
3. sn-gn/n-sto	4. li-sl/sn-ls
5. n-prn/g-pg	6. sbal-sn/sn-prn [bilateral]
7. gn-go/n-gn [bilateral]	8. sl-gn/sto-gn
9. al-al/ex-ex	10. ex-go/go-go [bilateral]
11. al-al/n-sn	12. n-sn/n-sto
13. sa-sba/n-gn [bilateral]	14. n-gn/n-sto
15. n-sn/sa-sba [bilateral]	16. en-al/ex-ch [bilateral]
17. ex-ex/go-go	18. obi-ch/g-sa [bilateral]
19. ex-n/ex-sto [bilateral]	20. sbal-ls/n-al [bilateral]
21. ex-n/n-sto [bilateral]	22. pi-al/sa-ex [bilateral]
23. ex-sto/n-sto [bilateral]	24. ex-obi/ex-ch [bilateral]
25. en-ex/ps-pi [bilateral]	26. ex-al/ch-gn [bilateral]
27. en-en/ex-ex	28. ch-ls/n-prn [bilateral]
29. pi-or/en-ex [bilateral]	30. al-ls/ch-gn [bilateral]
31. sa-sba/pa-obi [bilateral]	32. ch-li/ex-ch [bilateral]
33. cph-cph/sn-ls	34. ex-sto/rt ex-lt ch [bilateral]
35. ls-sto/ch-ch	36. sn-gn/ex-gn [bilateral]
37. sto-li/ch-ch	

As it is more common to use absolute measurements in anthropometric comparisons [27, 37, 61, 148] rather than proportions, there was less guidance with respect to which proportions would be more reliable or more relevant than others in the present study. Halberstein used a combination of up to twelve face and body proportions when comparing a photograph to a live subject, and three of these proportions were used [38]. These proportions were ear length/facial height (sa-sba/n-gn), nasal height/ear length (n-sn/sa-sba) and nasal width/nasal height (al-al/n-sn). The remainder of the proportions that were used by Halberstein were not incorporated into this research because they either included facial landmarks that were not chosen for the present study or because they were body proportions, such as shoulder width, leg or shoe lengths. Two proportions (n-sn/n-sto, n-gn/n-sto) were used by Catterick for his research [36]. The remainder of the proportions chosen were unique to this study.

Figure 3.3 Photograph of a face divided into three sections

During the selection process, the face was considered to consist of three sections (Figure 3.3): the top section includes the eyes and above (1), the middle section contains the nose (2), and the lower section would include the mouth and below (3). Intuitively, it is expected that longer lines between landmarks located in two different sections of the face would make a more reliable proportion than two short lines in the same section of the face. This is because small variations in landmark placement making up short lines would result in large changes in proportions, which may not accurately portray true variations between individuals. The proportions utilized in this research were deliberately chosen to include linear measurements between landmarks in different sections of the face and others that covered a small section of the face, such as the length versus the width of the eye.

The majority of anthropometric literature discusses the use of distances [23, 27, 37, 148, 151] and to a lesser extent, proportions [36, 38] between landmarks. However, it should be noted that another way of utilizing facial landmarks, is to measure the angles formed between landmarks. Angles, along with proportions, were examined in the initial Pilot Study discussed in Chapter 4.

4 Pilot Anthropometry Laboratory Study of Still Images

4.1 Introduction and Aims

A pilot study was initially conducted to determine if there was any basis for using anthropometric measurements in facial comparisons. The study considered the effects of camera angle on anthropometric measurements, consisting of proportions and angles, made from facial photographs of five individuals. The pilot study was an effective means of gaining experience in locating the landmarks and using the equipment and software to place these landmarks and, in addition, was a good basis on which to consider anthropometry as a means for forensic identification and whether further investigation was warranted using the same facial landmarks.

In principle, the data from these empirical measurements could also be used to develop a photogrammetric model of the face which, if calibrated, could be used to correct anthropometric measurements for distortions caused by a camera angle which differs from the one specified in a protocol for facial comparison. The purpose of developing such a model would lie in its use to calculate correction factors to convert observed proportions and angles back to the full-face orientation values, which could then, for example, be used to search a database of the proportions.

The methodology, results, discussion and conclusions of the study are presented in this chapter.

4.2 Experimental Section

4.2.1 Subjects

Facial photographs were taken of five subjects. The first five people who volunteered for the study were included, paying no attention to race, gender, or age. No one was compensated financially or otherwise for time spent in the study. The photographs used in the study were of three females, one English, one Irish, and one Malaysian, and two males, one from Sri Lanka and the other from India (Table 4.1). Permission was granted by the volunteers to use their photographs in the research and no additional ethical approval was required.

Table 4.1 List of five subjects and their nationalities

Subject 1	English female
Subject 2	Irish female
Subject 3	Malaysian female
Subject 4	Indian male
Subject 5	Sri Lankan male

4.2.2 Equipment and Photography

The photographs required for the pilot study were taken with a conventional 2.1 megapixel digital camera with automatic settings for film and shutter speed and aperture, in a laboratory setting with the lens axis in the horizontal plane. Facial photographs were taken at 10° intervals up to 90° both clockwise and counter-clockwise from the 0° (full face) position. The subjects were seated on a swivel chair placed on a marked base sheet to show angle position in relation to the camera (Figure 4.1). A compass was used to draw a circle around the centre of the base sheet and a protractor was used to mark 10 degree intervals. From the centre of the circle, straight lines were drawn outwards. The lines were marked with tape for better visualisation.

Figure 4.1 Experimental design

A swivel office chair was then placed on the base sheet in the centre of the circle and rotated as required. A ruler was fixed vertically to the front of the chair and the chair was turned incrementally such that the ruler lined up sequentially with the tapes on the base sheet, every 10º up to 90º in both directions from the full-face position.

To ensure the photographs were all taken at the same vertical level and at the same distance from the subject, the camera was attached to a tripod which was placed at a fixed distance (approximately one metre) from the rotational axis of the chair. The subjects were initially placed in the full-face position. To standardise the position of the head, the volunteers were instructed to look straight ahead and to keep as still as possible. A ruler was held vertically down the centre of their face to ensure that their head was not tilted to either side. The ruler was then held to the side of their face, lining their tragus with the lateral side of their cheekbone to avoid the head from being tilted too far up or down. The subject was instructed to hold that position with eyes open and mouth closed, but relaxed, whilst the photographs were being taken. This method was repeated for each 10º of rotation until the chair was in the 90º lateral position both in clockwise and counter-clockwise directions. Thus, a total of nineteen photographs were taken for each subject (Figure 4.2).

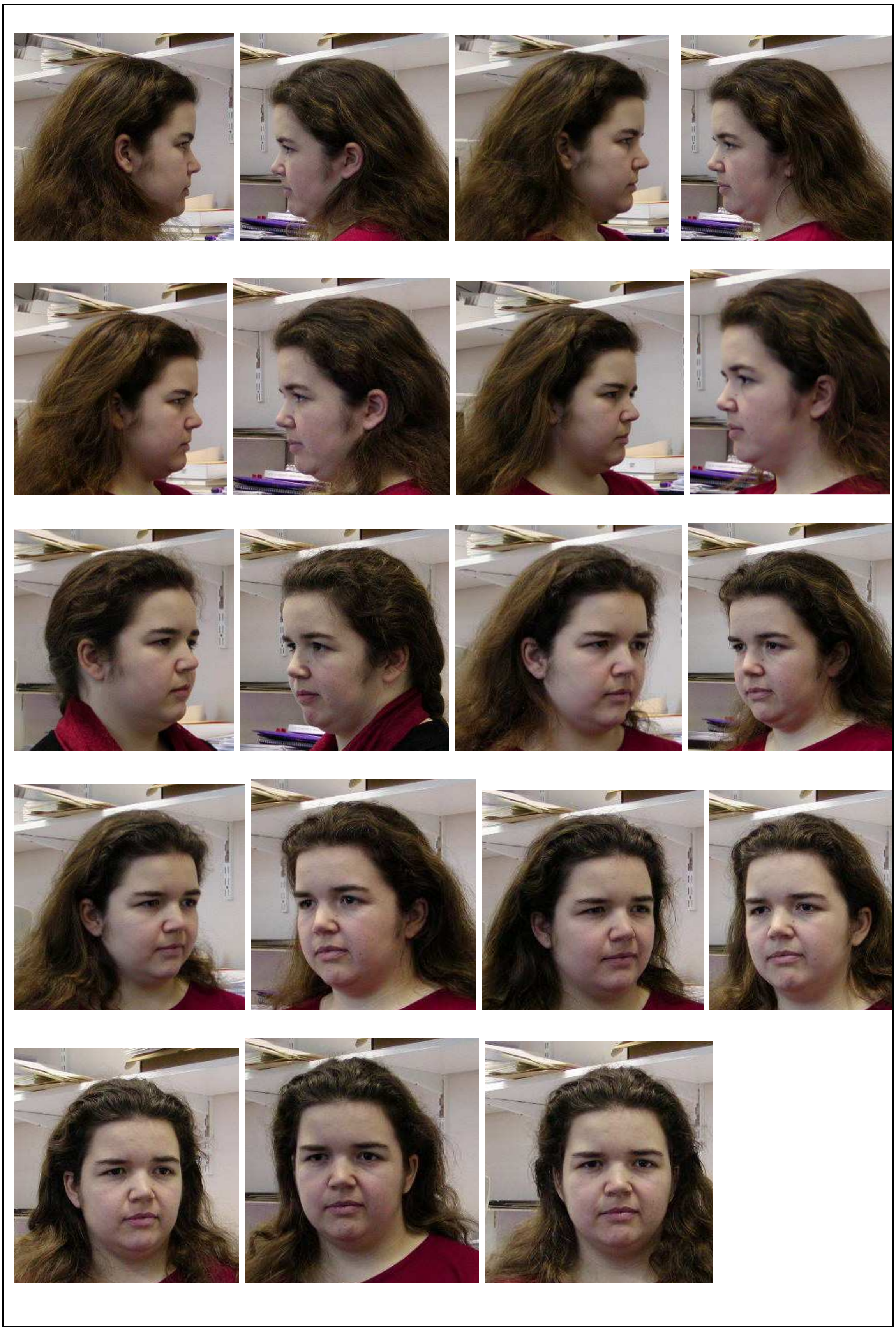


Figure 4.2 Photographs taken at camera angles between 0º and 90º, in 10º increments

4.3 Derivation of Anthropometric Measurements

The photographs were downloaded as Jpeg files and a measurement programme produced in-house, Facial Identification Centre Version 0.1 © Forensic Medicine and Science Glasgow University, was used to place landmarks on the image and carry out linear and angular measurements.

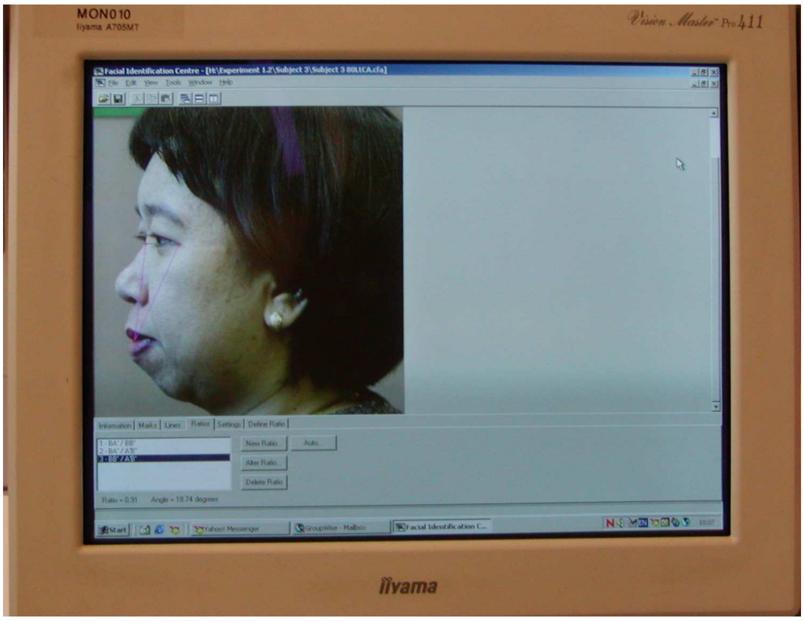


Figure 4.3 Screen shot of the software used to place landmarks and calculate proportions and angles

The software employed (Figure 4.3) was simple to use and included a tool that allowed the operator to place a small circle on each landmark chosen. It was possible to place a total of 61 landmarks if so desired. The landmarks were placed initially with the mouse and their positions could be fine tuned using the keyboard to the precise locations required by the operator. Additional tools in the programme allowed lines to be drawn between specific landmarks and ratios and angles between lines to be calculated automatically. The colours of circles representing landmarks and lines could be changed and the operator could decide whether lines should be displayed or not.

For the pilot study, four facial landmarks were chosen as discussed in Section 3.2, consisting of the right and left exocanthions, the stomion and nasion as illustrated in Figure 4.4. Linear measurements were made between the landmarks and proportions calculated between them. Similarly the angles between the lines were calculated.

Each landmark was assigned the following notation:

A = right exocanthion

A' = left exocanthion

B = nasion

B' = stomion

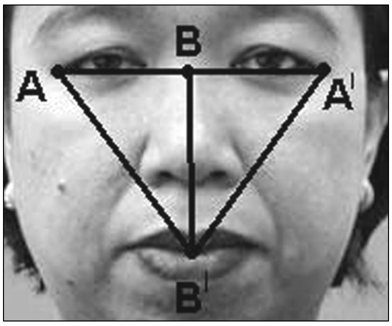


Figure 4.4 Landmarks selected for this study

4.3.1 Data Processing

Lines were created between each pair of landmarks, giving 5 linear measurements in total. Proportions and angles between the lines were calculated as follows. A proportion was derived using the equation:

$$P = \frac{\text{Numerator (smaller linear measurement)}}{\text{Denominator (larger linear measurement)}} \times 100$$

An angle value consisted of the angle between any two specific lines. The proportions measured were BB' / A'B', BA' / A'B', BA' / BB', AB / AB', AB / BB', BB' / AB' and AB' / A'B'. Angle values recorded were ∠ BB'A', ∠ BA'B', ∠ BAB' and ∠ BB'A. Both sides of the face were included since faces are usually asymmetrical to some degree. Also, two angles in each of the two triangles sharing BB' as a common side were included, since the angles ABB' and A'BB' were not always 90° angles (Figure 4.5 indices.)

In cases where either exocanthion could not be seen, the landmark was not used and only the data available using the nasion and the visible exocanthion were included (Figure 4.5). This usually occurred with camera angles beginning at 20° on either side of the full-face position.

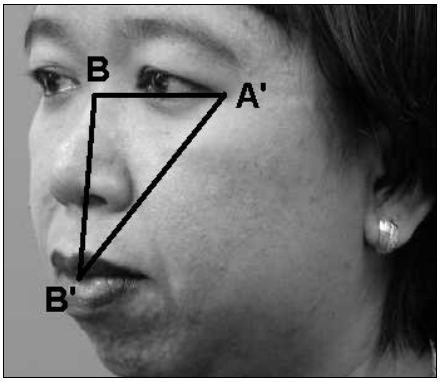


Figure 4.5 Example of subject rotated 50º clockwise

4.3.2 Tests Used

The calculated values of the proportions indices and angles were recorded in an Excel® spreadsheet. The numerical data for each proportion index and each angle value were then plotted on a line graph against camera angle to illustrate the variation of proportions and angular measurements with angle of rotation of the subject relative to the camera. The average values and relative standard deviations of the proportions and angle measurements for the five subjects were also determined and placed in Table 4.2.

4.4 Results

Tables 4.2a and 4.2b show the mean values and relative standard deviations of the proportions and angle measurements and Figures 4.6a-k illustrate the variation of proportions and angles with angle of rotation of the subject for each of the five subjects, both in an absolute sense and as a proportionate change compared to the 0° value. In positions where landmarks were not visible, measurements were not carried out, usually, for example, for rotations of 20° and greater. This was because, as the subject was rotated sideways, one exocanthion was no longer visible.

The experimental error in the camera angle and errors resulting from placement of the landmarks were not taken into account when analyzing the facial proportions and angles but were assumed to be constant for the five subjects.

Tables 4.2a-b Means and Relative Standard Deviations for (a) Proportions and (b) Angles in five subjects

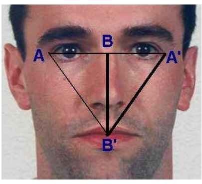
(a)

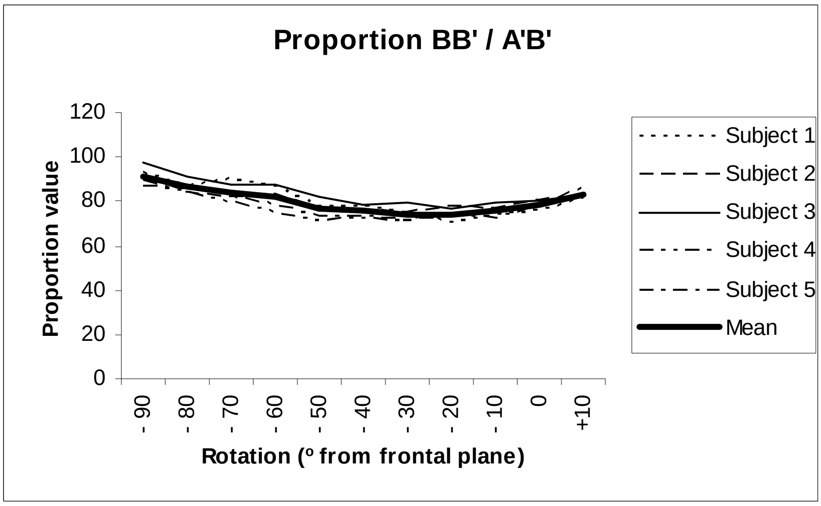
	Means and % Relative Standard Deviations (in parenthesis) for Proportions						
Rotation (degrees)	**BB' / A'B'**	**BA' / A'B'**	**BA' / BB'**	**AB / AB'**	**AB / BB'**	**BB' / AB'**	**AB' / A'B'**
-90	91.3 (4.5)	38.7 (27.4)	42.8 (29.5)				
-80	86.3 (3.6)	49.4 (21)	57.6 (23.4)				
-70	84.2 (5.5)	54.1 (11.6)	64.7 (16.1)				
-60	82.1 (6.6)	57.5 (13.6)	70.7 (19.8)				
-50	76.4 (5.6)	65.1 (10.3)	85.6 (14.1)				
-40	75.4 (3.7)	67.4 (8.2)	89.5 (10.4)				
-30	74.4 (4.3)	67.1 (10.6)	93.1 (10.8)				
-20	74.1 (3.9)	69.1 (6.2)	93.5 (9.7)				
-10	75.7 (3.3)	67.4 (4.7)	89.3 (7.4)	54.3 (5.9)	65.9 (9.4)	82.6 (3.8)	91.7 (4.5)
0	**78.6 (2.4)**	**63.1 (5.2)**	**80.3 (8.0)**	**62.4 (4.5)**	**81.3 (7.7)**	**76.9 (3.4)**	**102.3 (1.2)**
+10	83.1 (1.8)	56.2 (6.8)	67.7 (8.4)	65.4 (5.8)	87.4 (9.4)	75.1 (3.9)	110.7 (3.1)
+20				67.8 (6.5)	92.9 (9.9)	73.2 (3.7)	
+30				68.5 (7.9)	92.4 (8.6)	74.2 (1.8)	
+40				67.3 (8.6)	90.7 (10.4)	74.3 (2.6)	
+50				64.0 (8.3)	85.2 (12.2)	75.5 (4.6)	
+60				57.9 (10)	72.8 (13.5)	79.9 (4.4)	
+70				54.0 (12.8)	64.5 (16)	84.2 (5.1)	
+80				49.3 (20.1)	68.0 (23.4)	85.8 (5.2)	
+90				38.1 (30.7)	42.4 (34)	91.1 (4.6)	

(b)

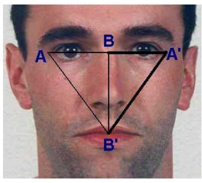
Means and % Relative Standard Deviations (in parenthesis) for Angles				
Rotation (degrees)	**Angle: BB'A'**	**Angle: BA'B'**	**Angle: BAB'**	**Angle: BB'A**
-90	22.8 (28.5)	66.4 (9.6)		
-80	29.7 (22.5)	59.7 (5.9)		
-70	32.8 (13.1)	57.4 (8.4)		
-60	35.2 (15.3)	55.4 (9.6)		
-50	40.7 (12.1)	49.9 (7.6)		
-40	42.4 (9.7)	48.9 (4.9)		
-30	43.9 (10)	48.1 (5.8)		
-20	43.8 (7.8)	47.9 (5.2)		
-10	42.4 (5.9)	49.2 (4.5)	55.7 (5.6)	32.9 (6.7)
0	**39.1 (6.4)**	**51.9 (3.5)**	**50.3 (4.6)**	**38.7 (5.4)**
+10	34.2 (7.6)	56.2 (2.8)	48.7 (5.1)	40.9 (7.1)
+20			47.0 (4.9)	42.7 (8.2)
+30			47.8 (2.3)	43.3 (9.7)
+40			47.9 (3.3)	42.4 (10.4)
+50			49.1 (6.3)	39.9 (9.8)
+60			53.0 (6.4)	35.4 (11.3)
+70			57.4 (7.8)	32.7 (14.4)
+80			59.4 (8.9)	29.7 (21.2)
+90			66.2 (10.4)	22.5 (32)

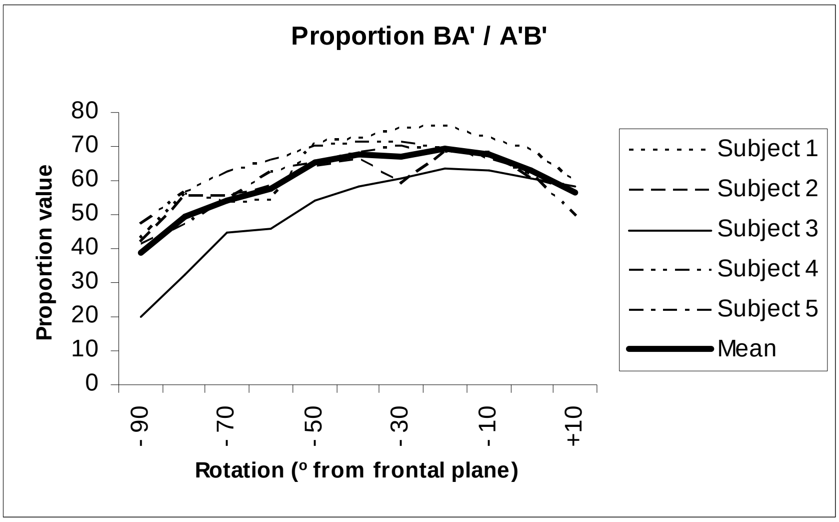
(a)



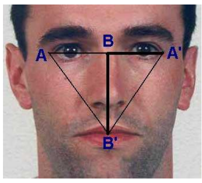


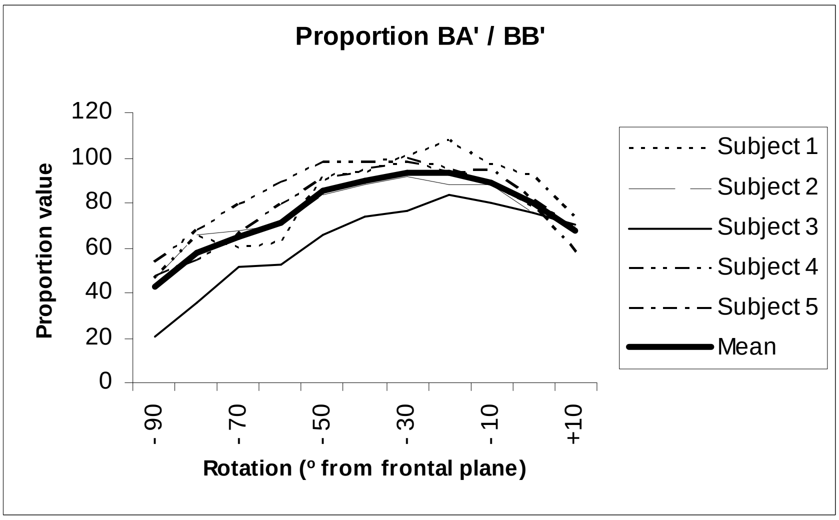
(b)


A
B
A'
B'

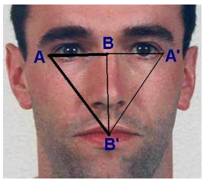

Proportion BA' / A'B'
Proportion value
Rotation (° from frontal plane)
- 90
- 70
- 50
- 30
- 10
+10
Subject 1
Subject 2
Subject 3
Subject 4
Subject 5
Mean

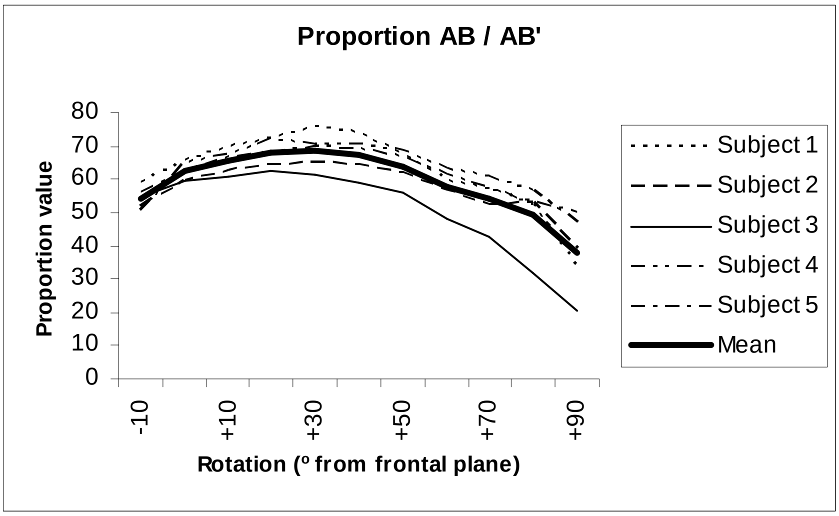
(c)


A
B
A'
B'

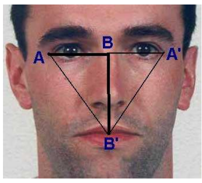

Proportion BA' / BB'
Proportion value
Rotation (° from frontal plane)
- 90
- 70
- 50
- 30
- 10
+10
Subject 1
Subject 2
Subject 3
Subject 4
Subject 5
Mean

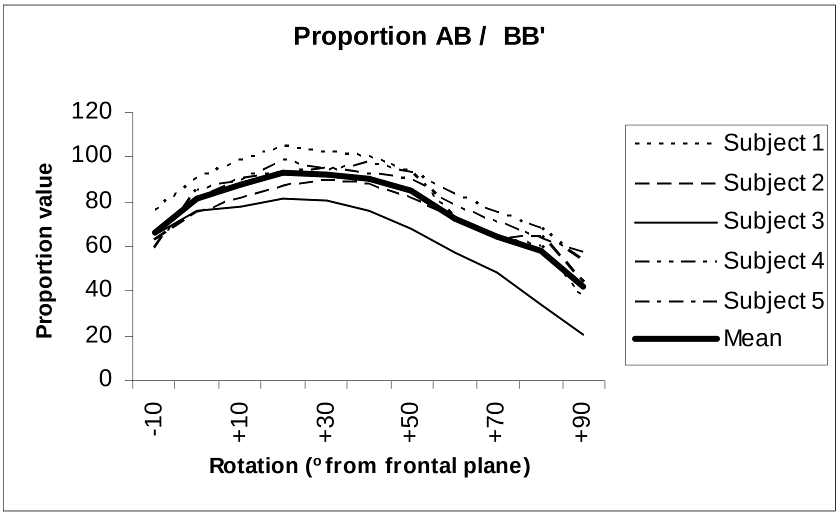
(d)


A
B
A'
B'

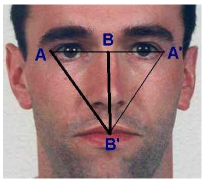

Proportion AB / AB'
Proportion value
Rotation (° from frontal plane)
-10
+10
+30
+50
+70
+90
Subject 1
Subject 2
Subject 3
Subject 4
Subject 5
Mean

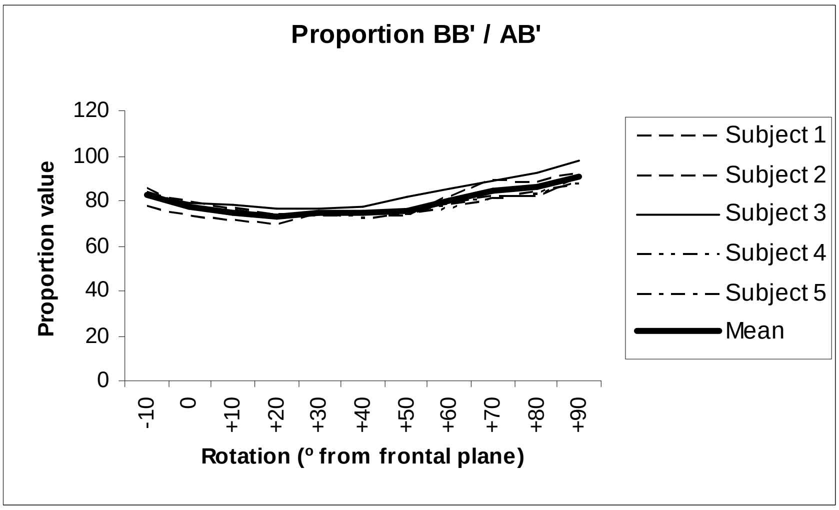
(e)


B
A
A'
B'

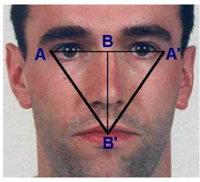

Proportion AB / BB'
Proportion value
120
100
80
60
40
20
0
-10
+10
+30
+50
+70
+90
Rotation (° from frontal plane)
Subject 1
Subject 2
Subject 3
Subject 4
Subject 5
Mean

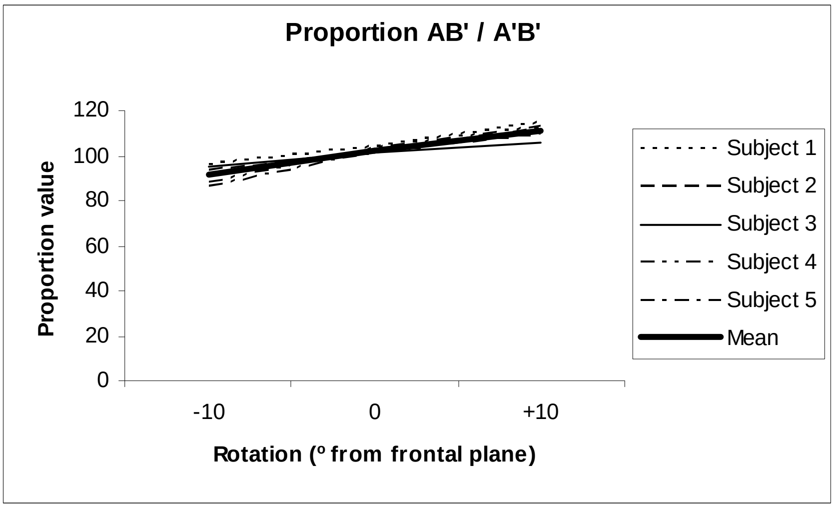
(f)


B
A
A'
B'

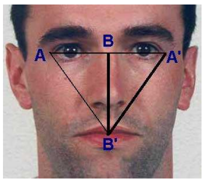

Proportion BB' / AB'
Proportion value
120
100
80
60
40
20
0
-10
0
+10
+20
+30
+40
+50
+60
+70
+80
+90
Rotation (° from frontal plane)
Subject 1
Subject 2
Subject 3
Subject 4
Subject 5
Mean

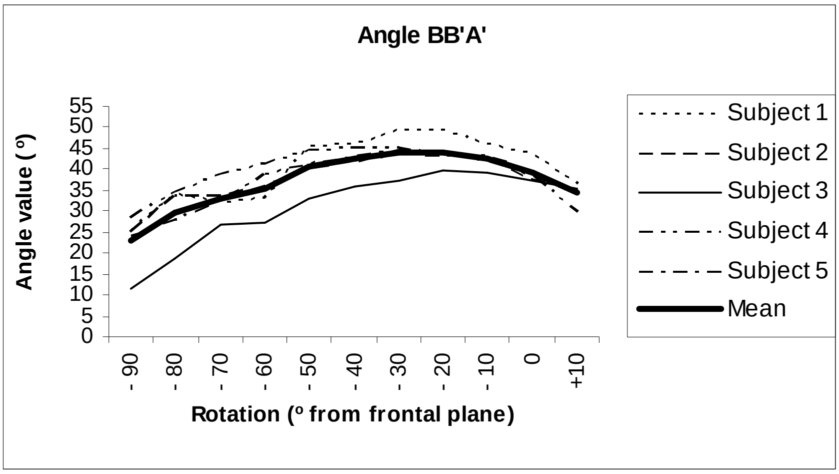
(g)


B
A
A'
B'

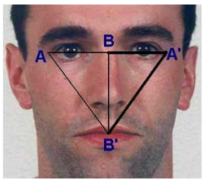

Proportion AB' / A'B'
Proportion value
120
100
80
60
40
20
0
-10
0
+10
Rotation (° from frontal plane)
Subject 1
Subject 2
Subject 3
Subject 4
Subject 5
Mean

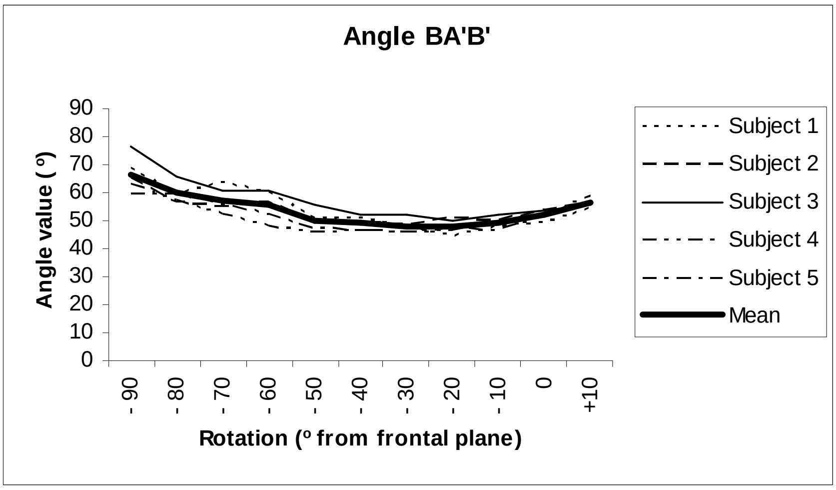
(h)


A
B
A'
B'

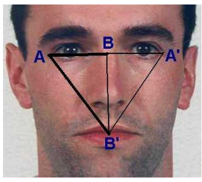

Angle BB'A'
Angle value (°)
55
50
45
40
35
30
25
20
15
10
5
0
- 90
- 80
- 70
- 60
- 50
- 40
- 30
- 20
- 10
0
+10
Rotation (° from frontal plane)
Subject 1
Subject 2
Subject 3
Subject 4
Subject 5
Mean

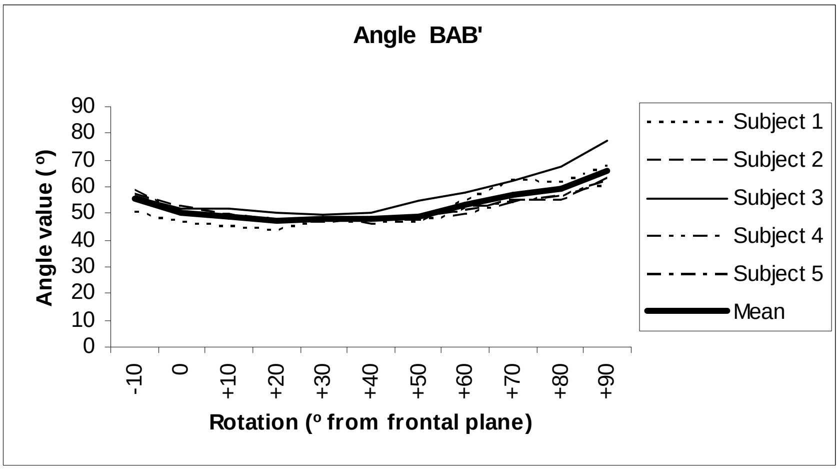
(i)


A
B
A'
B'

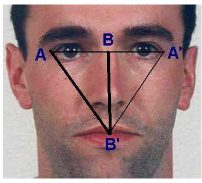

Angle BA'B'
Angle value (°)
90
80
70
60
50
40
30
20
10
0
- 90
- 80
- 70
- 60
- 50
- 40
- 30
- 20
- 10
0
+10
Rotation (° from frontal plane)
Subject 1
Subject 2
Subject 3
Subject 4
Subject 5
Mean

(j)


A
B
A'
B'


Angle BAB'
Angle value (°)
90
80
70
60
50
40
30
20
10
0
-10
0
+10
+20
+30
+40
+50
+60
+70
+80
+90
Rotation (° from frontal plane)
Subject 1
Subject 2
Subject 3
Subject 4
Subject 5
Mean

(k)



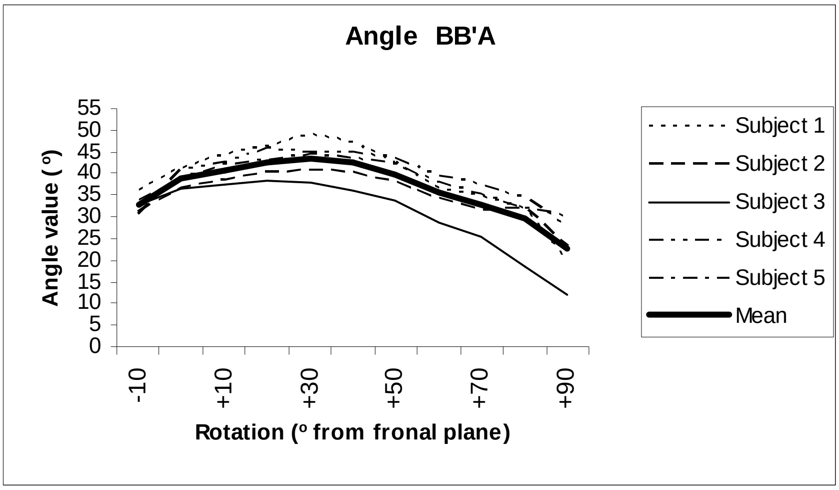


Figure 4.6a-k Proportions and angle measurements in five subjects

4.5 Discussion

4.5.1 Experimental Set-up

In this study, it was found that a significant limitation during the process of taking photographs was in positioning the head of the subject accurately. It was a challenge first for the photographer to position the head so that both left and right sides of the corresponding degree of rotation matched exactly. It was also difficult for the volunteer to keep their head in the position posed. It is natural for the head to move slightly, even in the short time it takes for the photographer to position and then walk around to the camera to take the picture, and any deviation in head position would affect subsequent measurements made on the photographs. The chair used might also have affected how level the head appeared. If the chair was not perfectly level, i.e. the axis of rotation was not vertical, then the angle of the head would change with the angle of rotation. The subject's posture, when seated can also affect the orientation of their head. All of these variables contribute to subsequent measurement errors. Good results not only depend on the skill of the photographer and his/her ability to position the head, but also rely on the volunteer's ability to hold his/her head still.

During the process of taking the photographs, other problems were encountered. The first time the photographs were taken, the subjects started in the left lateral position and continued around every 10 degrees until reaching the right lateral position. The subjects were told only to look straight ahead and to keep a neutral face. These photographs could not be used in the study because when the left and right corresponding degrees were compared, they were found to be inconsistent with each other. The right and left corresponding degrees should be in a similar position because this provides a more accurate comparison.

One volunteer had trouble sitting still. Once his head was positioned correctly, the photographer would return to the camera and tell him to look straight ahead, keeping his mouth closed. He would then proceed to nod his head in agreement, effectively ruining his head position. Once analysed, this subject's group of pictures was deemed unusable due to lack of angle continuity.

If this study was to be repeated, many of these problems could be resolved by using a purpose-built rig in which the subject's head is guided into the correct position and restrained from moving, for example by using a cradle as in a computed tomography scan. In addition, a rig of this type could ensure that the axis of rotation of the head was at a selected location within the cross-section of the head. In the chair set-up used in this pilot study, the axis of rotation was set by the base of the chair and there was no accurate method of positioning the subject such that, for example the centre of the head was positioned on this axis. Subsequent work attempted to address these problems by keeping the subject fixed in position and moving the camera (Chapter 5).

4.5.2 Landmarks

Landmark placement during analysis of images may also affect the results. Some level of experience at this point is required because placement of the landmarks can be highly subjective. For example, one person's interpretation of the location of the exocanthion may not be exactly the same as another's. Standardised marks also may not be easily located because the majority of images used for comparison in forensic casework are of poor quality. For this reason, in the study, the exocanthion point was placed on the larger reference point where the upper eyelid overlapped with the lower lid. Each of the landmarks in this study was located on photographs rather than by using the 3D face of the subject, which was the approach used in some published studies. It was not feasible to feel the 3D face for the exact position of landmarks because all CCTV or other offender footage and suspect images are compared in 2D and location of landmarks on such images is more appropriate.

The points selected for the study were ones that were relatively fixed on the face and, except for the stomion do not alter with a changing facial expression. One approach to achieving placement continuity would be to create a new set of standards for placing landmarks on photographs, keeping the variability of forensic cases in mind. A starting point may be to divide faces into small sections with the same number of gridlines in which to focus on each landmark. Or, as mentioned in the previous paragraph, the exocanthion could be placed where the upper eyelid overlapped with the lower lid instead of the lateral corner of the eye. It would be important not to increase the area in which a point could be placed, but instead, choose larger reference areas on the face. If this were to be undertaken, extensive studies would need to be carried out on possible alternative landmark locations that were both easily found on standard lower quality images as well as being easily and reproducibly located by different operators. Landmarks that have been found to be less reliable could be examined first and, if studies show promise, a move towards creating a whole new set of landmark locations specifically for the forensic science field could be constructed.

Not all landmarks were visible from every camera angle, since many were hidden by the rotation of the head. This result does not indicate that these landmarks are of no value in facial comparisons but that, from any given camera angle, a subset of measurements should be used. The empirical approach used in this study could be extended to include a wider range of measurements in order to characterise their useful angular fields. This would also include angle of tilt as well as rotation of the camera relative to the subject. Since the new database being constructed for use by law enforcement in the United Kingdom (FIND) has selected the 0° and 45° angles of rotation in the horizontal plane, then the number of landmarks accessible is necessarily restricted but a list of those which are accessible could be prepared.

4.5.3 Proportions and Angle Measurements

As shown in Table 4.2a, all six proportions were available only on photographs taken from the frontal view or taken when the subject was rotated 10° to the left and/or right.

Proportions which have small relative standard deviations indicate a small variation between the proportions or angles of the subjects in the study group. An example of this is proportion index AB'/A'B' at 0°, where this proportion was very similar in all subjects and the corresponding relative standard deviation was only 1.2%. Subsequent work (Chapter 5) showed that this degree of variation could be explained by measurement error. By contrast, proportion BA'/BB' has a larger relative standard deviation of 8%, indicating a wider degree of variation between individuals. Variations in angle measurements were also observed between the individuals in the small study group and, as for the proportions, the RSD increased with camera angle.

The implication of these results for forensic science is that proportions and angles are needed which vary significantly between individuals, otherwise anthropometry cannot be used to distinguish between them. The proportion AB'/A'B' would be of little value since it does not change much between people. At the 0° angle the two proportions BA'/BB' and AB/BB' have the largest RSDs. These might be considered to be giving a measure of the ratio of the half-face width to the central face length.

The distances between landmarks measured on the photographs decreased at oblique camera angles, but the errors in the measurements remained the same, and this is reflected by increases in the relative standard deviations. The relative standard deviations of the ratios did not increase in proportion when any of the measured distances became small, as occurs by foreshortening due to an oblique camera angle. This is most likely caused by the error in setting the camera angle. The relative standard deviations of angles did follow the expected pattern for the counter clockwise rotations, but did not follow the expected pattern for the clockwise rotations. Again, this is likely due to an error caused by rotating the subject.

One of the aims of this pilot project was to determine the effects of the camera angle on proportions and angles. The effects can be deduced from the individual curves in Figures 4.5a-k, which show the absolute and proportionate changes in the measurements according to camera angle. Flatter curves indicate that the camera angle is less significant whereas pronounced curves indicate a significant effect. A change in camera angle of even 10 degrees produces a change in proportions and angles which is bigger than the measurement error (as indicated by subsequent work described in Chapter 5). Clearly, the empirical data support the intuitive expectation that anthropometric measurements from photographs are significantly affected by the camera angle relative to the subject. If these measurements are used to compare faces depicted in images such as photographs or else if they are used to identify an unknown individual using an image database, then the camera angles used in the two images under comparison must be closely similar. Alternatively, an allowance must be made for the camera angle and the measurements corrected to those obtained from a standard view (such as zero degrees) before they are used.

4.5.4 Modelling the Data

Given the opportunity to compare a suspect photograph on file with an image taken from a crime scene surveillance video, it is more than likely that the facial orientations in the two images with respect to the camera will differ from each other. If it was then possible to calculate a correction factor, it would aid in solving one of the fundamental problems of forensic anthropometry, that is, the comparison of two photographs taken from different camera angles. This would be very helpful to use in the type of situation where, as part of a suspect profile when they are arrested, a record of their landmark locations at the frontal view were kept to be used for comparison. Any image obtained during a subsequent arrest that was to be compared to the original police photograph could then be corrected back to the frontal view as a way of negating facial rotation.

In principle, the changes in the measured parameters could be calculated using a mathematical model of the head using geometry and trigonometry. However, this type of model would require three-dimensional data relating to the spatial arrangement of the landmarks, and this cannot be obtained from a single two-dimensional image although more can be learned from a sequence of images [152, 153]. In practice it would also be difficult to construct the model by direct measurements of the head. The purpose of developing such a model would, as mentioned above, lie in its use to calculate correction factors to convert observed proportions and angles back to the full-face orientation values, which could then, for example, be used to search a database of the proportions and angles. Also, although the experimental set-up used was as controlled as possible with the equipment available, there were unknown dimensions introduced such as the exact distance of the camera from the axis of rotation and the position of the subject's head relative to the axis of rotation. A three dimensional scanner, as used in facial reconstruction, would be necessary for the development of a model which would allow proportions to be calculated for any camera angle.

The chosen landmarks for this study form triangles and, although it is a basic principle of geometry that given two sides or angles of a triangle, information for the third can be easily found, it was important in this study to consider data obtained from all sides and angles of the triangles. This was because the landmarks are in three dimensional space and lie on different vertical planes with respect to the line of sight of the camera lens, whereas the photographs are two dimensional images which show the projections of the points on a single plane. They were therefore expected to generate data different from that of a triangle created from points on the same vertical plane. Empirical measurements of all sides and angles were therefore required to determine which would give the most representative and usable data.

The results indicated that proportions and angles varied according to the angle of rotation of the subject relative to the camera and that the graphs of indices and angles mostly formed curves with maxima or minima. Considering a right angled triangle similar to those shown in Figure 4.4, rules of perspective dictate that, as the camera angle rotates horizontally, the horizontal side of the triangle will foreshorten and the vertical will stay constant. An expected result of this is that a proportion involving the horizontal side should be maximum at 0° and decrease to zero at 90° and a graph of the proportion versus angle would show a turning point at 0°. However, this was not the case and the presence of turning points in the graphs other than at 0° clearly indicates that the landmark points do not lie in a single plane at right angles to the full face position.

A simple illustration of this type of model is given in Figures 4.7a-c, in which a horizontal cross-section of the head through the nasion is approximated to a circle. BA' represents the line between the nasion and the right exocanthion, assumed to lie in the plane of the cross-section. If the camera is sufficiently far from the subject (assumed to be true), light from the subject reaching the camera can be approximated to parallel light, and the observed length of BA' is ba', which is the projection of the line BA' on the image plane, i.e. equivalent to a photograph taken by the camera. The angle θ is the angle between the 0° line and the line BA' on the face. Figure 4.7a represents the observed length of BA' in the full-face view. Note that the observed length ba' is smaller than BA'.

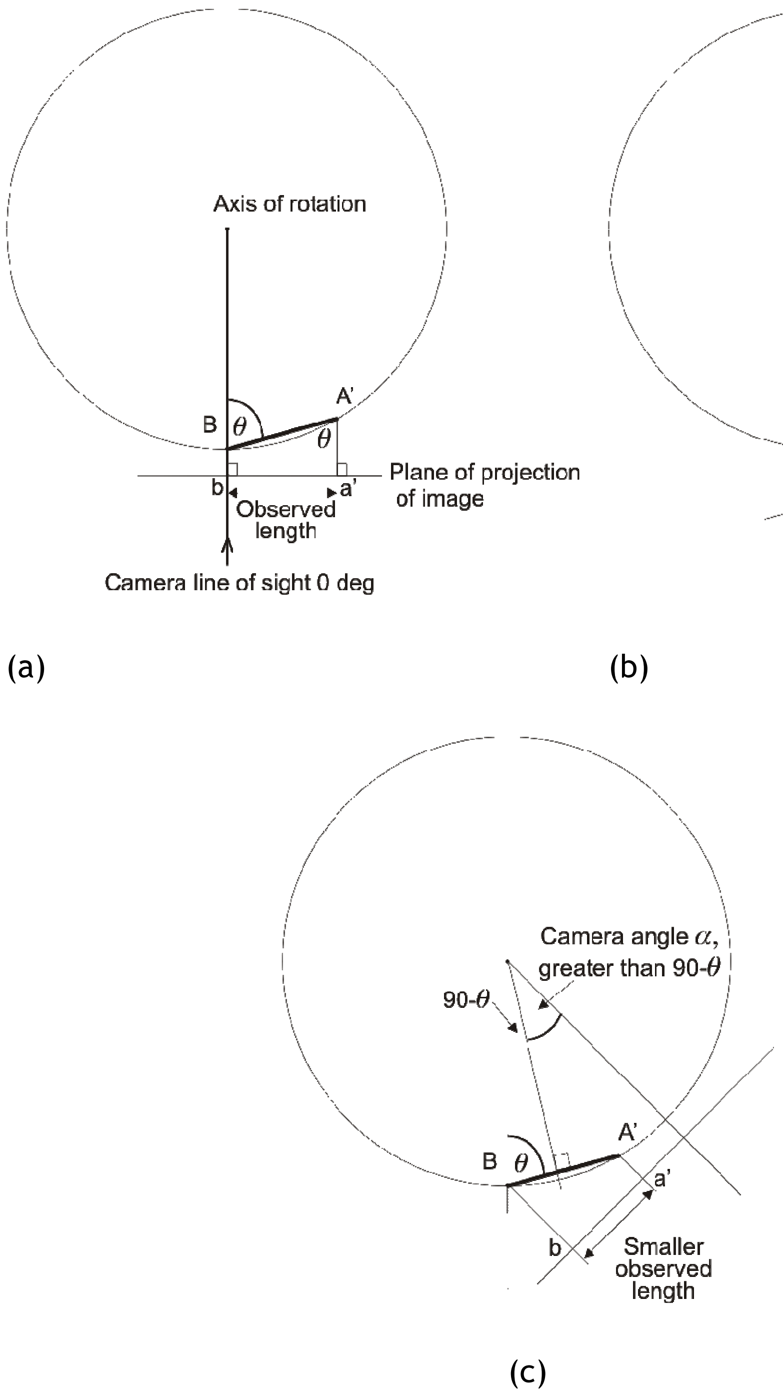


Figure 4.7a-c Simplified model of the effect of camera angle on observed length of a facial measurement.

The model indicates that ba' will be maximum when the camera angle is equal to 90°-θ (Figure 4.7b). At angles other than 90°-θ (Figure 4.7c), ba' will decrease in proportion to the sine of angle α between the line of sight of the camera and 90°-θ. Clearly there will be a maximum in the observed length of ba' and a graph of ba' versus camera angle will also have a maximum. Similarly, a proportion involving ba’ as the numerator will tend to maximise also, depending on the linear measurement used for the denominator of the index.

The model is simplified insofar as it does not include a tilt angle for the camera or the facial line, but it illustrates that, in principle, empirical measurements such as those made in this study could be used to determine facial angles such as $90-\theta$, which is the camera angle at which ba’ is maximum. Angles such as θ would be difficult to measure directly on the face but could be obtained using a three-dimension laser scanner, if available. The model could be extended to include proportions and camera angles above or below the horizontal plane and in principle could provide an estimate of the anthropometric measurements which would be obtained when measuring photographs taken at different camera angles. This includes the full-face view, thereby allowing correction factors to be introduced when comparing photographs taken at different camera angles.

In a laboratory setting, camera parameters and angles of rotation will be known. When faced with a photograph in which these parameters are unknown, it may be possible to determine them using photogrammetry. Common mathematical techniques are the Direct Linear Transform (DLT), the 8-parameter transformation, the bundle adjustment, and other vanishing point methods (based on manual line construction or analytically constrained coplanarity equations) [34]. In cases where angles of rotation are unknown, a system was developed for use with facial recognition software that, using 30 facial landmarks, estimates the angle of the face [154]. Forensically, this type of system may be contributive to cases where multiple images of a suspect (of unknown angles) were acquired from a video surveillance camera; it may be possible to choose the image with the same angle of rotation as the one located in the database. This system could be constructive to use in conjunction with the model to first determine the angle rotation of the face in order to correct the facial measurements back to frontal view.

Equally of interest, if a model head of known dimensions were to be photographed from an unspecified angle, the camera angle relative to the head could be estimated by comparing the observed proportions with the true values (full face view), which would be known for the model head. This is one way of calibrating a mathematical model of the locus of a crime which has produced images of a suspect for analysis. After the model has been calibrated, correction factors might be derived for anthropometric measurements made from the images of the suspect, allowing them to be submitted for a search of a database.

The pose invariant systems discussed previously in Section 2.8.2 of Chapter 2 have their own specific equations for estimating pose to then correct images back to frontal view. The difference between those methods and the theoretical model discussed here is that using a more complex approach to correct images back to a frontal view, they then apply their results in a facial recognition program. They are creating a whole new frontal view image, while the theoretical model developed focuses on specific measurements between landmarks. It is possible that the theoretical model in this study can be applied to the method of comparison of anthropometric measurements undertaken in this thesis.

This approach to address the problem of comparing images of differing facial rotations may be a simpler concept compared to the more complex methods acknowledged in Section 2.8.2. However, a simpler approach may be readily available to a more widespread audience, such as those organizations with smaller budgets, i.e. police departments. Given the choice in an ideal world, it may be desirable to use one of the more complex methods; however, they are not without their problems as well.

At present, there may not be a guaranteed method of facial comparison that is able to be used for all cases, covering every angle of rotation at different ages. Methods that require specialized variables, such as a certain angle rotation, may be better used for those circumstances where that variable can be controlled, such as for business security. The pose invariant facial comparison methods described only deal with facial rotation, and other constraints such as facial expression and lighting are not dealt with. This demonstrates the complexity of the human face and how difficult it is to develop software that can be used for all types of scenarios such as needed for forensic work.

Despite problems encountered while obtaining the data, which are further discussed below, the results demonstrated a consistent change in values between each 10° step rotation for each person. The graphs illustrate that each subject follows a similar curve, both with respect to angles and proportions, indicating a predictable change between each 10° rotation. Given this information, it may be possible to estimate from which angle a photograph is taken. However, from this study it is clear that in order to attempt to predict facial position with an acceptable degree of accuracy, a much larger reference group of subjects categorized by gender, ethnicity, and facial build would need to be studied. Although all subjects followed a similar curve, Subject 3 in particular had the greatest deviation from the mean compared to the other four subjects. This may have been a result of variations in photographic angles but can most likely be attributed to the race of the subject, again indicating the need of a larger categorized study.

4.6 Conclusions

This initial pilot study identified problems regarding the uncertainty of the angle of the subject and the chair which were attempted to be remedied in subsequent work discussed in Chapter 5. As the location of the head of the subject was not held consistent with no known axis of rotation, the next study attempted to rectify this by identifying a specific axis of rotation and although there was not a ‘rig’ available to hold the head, an attempt was made to keep the head position constant by placing a stand in which the subject could place their chin, resting their head and moving the camera instead of the head to minimize movement. The distance from the camera to the axis of rotation was recorded and held constant for all photographs and lastly, during the process of taking the new series of photographs, the focal point of the camera was centred on the nasion.

The information gained from this study cannot be applied to present day casework, although it does show the degree of change of proportions and angles in relation to landmark distances and therefore the way facial features change with rotation. A predictable variability is demonstrated in measurement outcomes (angles and proportions) with facial position using a limited number of landmarks. However, in addition to demonstrating such predictability, the fundamental question, which is addressed in the major part of the thesis, is whether or not anthropometry is sufficiently discriminatory between individuals, especially between subjects with a similar physiognomy, and therefore whether such a technique might be useful as corroborative evidence or even for a positive identification in facial image comparisons. It was concluded that there was evidence of sufficient potential discrimination to warrant further investigations.

5 Investigation of Uncertainty of Anthropometric Measurements

5.1 Introduction and Aims

The objective of this chapter was to estimate the uncertainty in the measurements of the chosen facial proportions caused by the processes involved in landmark placement and taking photographs, including the uncertainty contributions resulting from different people performing these tasks. The aim of these studies was to simulate effects found in the real world, as there would be different operators both placing landmarks and taking suspect photographs in police departments across the world. In addition, this study was completed in order to address variables encountered in the Pilot Study that occurred as a result of the experimental set up.

Within the field of forensic science, the estimation of uncertainty has become a focus of attention in recent years, to a large extent because of the use of forensic science results in court and the consequent need of the judiciary and jury to have an indication of the reliability of these results. It is important to note the distinction between uncertainty and error, which is clearly expressed by Rowley [155]. “An error in measurement is a consequence of a mistake and so can be corrected by eliminating the error and then repeating the measurement correctly. Uncertainty may be able to be reduced and estimated but it is an intrinsic property of the measurement system and cannot be completely eliminated.” Estimation of uncertainty requires firstly the estimation of random measurement errors (Type A uncertainty contributions) since these cannot be eliminated and largely determine the net uncertainty of the measurements [155]. Other sources of uncertainty (Type B contributions) are those which can be estimated without recourse to empirical studies, for example, calibration data of an instrument, manufacturer’s specifications or other published sources [155]. These can then be combined to obtain an overall estimate of uncertainty. Sources of contributions to the overall uncertainty relative to this study are given in Figure 5.1. Only some of these sources of error were amenable to study, for example, different measurement systems could not be evaluated and only the computer-based system which was available was used.

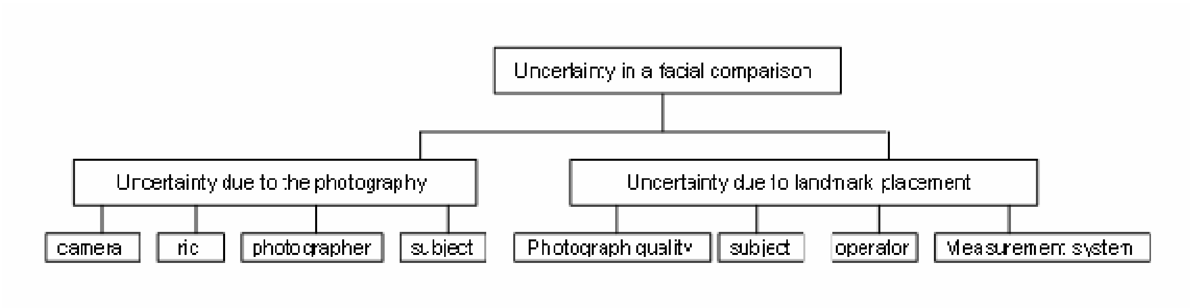


Figure 5.1 Sources of uncertainty

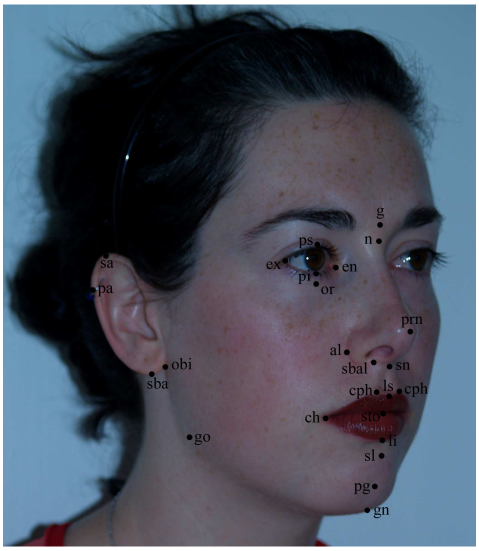
This chapter is divided into two sections. The first section reviews the errors involved in measuring facial proportions as a result of variations in landmark placement. The second section reviews the errors in measuring facial proportions as a result of the process of taking photographs. This chapter reviews the methodology, results and discussion of these studies.

5.2 Errors in Facial Proportions Resulting from Variations in Landmark Placement

Two analyses using repeatability data were carried out to determine the errors which occurred in placing landmarks. These analyses were necessary because the placement of landmark points is subjective to a certain extent and is subject to variation. By measuring the degree of variation which arises from the measurement process itself, robust criteria might be established to test whether images matched or were different.

5.2.1 Materials and methodology

A full description of the research material used for Section 5.2 can be found in Chapter 6, although a brief description is provided here. Video and database images provided were of male Caucasian policemen. Faces were displayed from the frontal viewpoint, showing features from the neck up, in the format of police identification photographs. Figure 5.2 is included as a reminder of the location of the facial landmarks used to derive proportions measured for the study.



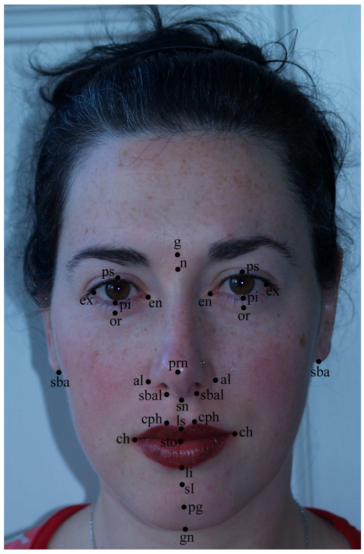


Figure 5.2 Facial landmarks and their location

5.2.1.1 Methodology of intra-operator study

This test was conducted on high resolution research material and included samples of both video images and photographs. In total, the video and photographic images of a single subject were measured 10 times. Landmark placement was carried out by the same operator using the same equipment on the same day (i.e. under repeatable conditions). The operator participating in this study had previous experience both of using the equipment and placing the chosen landmarks. For this study, a total of four representative landmarks and six proportions were used as in the Pilot Study. The landmarks chosen were located on the central portion of the face and were found to be easily located.

A second intra-operator analysis was then conducted on images of ten different subjects, which were each measured once a day for seven days. The same ten subjects were used for both video images and database photographs. Subjects were chosen at random, ensuring that they covered a range of faces and a range of difficulty in locating landmarks. The landmark placement was carried out on the images by the same operator on the same equipment at the same time but only once daily. However, each day the landmark placement was performed on the subjects in a different order. These conditions do not constitute the widest set of variations (for example, different operators, different equipment) but were relevant to the present study.

The coefficient of variation for each proportion was found and results are reported in Section 5.2.2.1.

5.2.1.2 Methodology of inter-operator study

This analysis was conducted on high resolution research material and included samples of video images chosen at random. The inter-operator study developed from the previous intra-operator study and was carried out on the full set of 37 landmarks and 59 proportions. Facial landmarks were placed on a total of six video images, six times each by five different operators. One operator had previous experience in using the equipment and knowledge of the landmarks. The remaining operators had no experience in using the equipment and no previous knowledge of anthropometric landmarks. The same equipment was used by all operators and each operator conducted their landmark placement of images in a single day. The coefficient of variation for each proportion was found and results for this analysis are reported in Section 5.2.2.2.

5.2.2 Results

The analyses were carried out to measure the errors in proportions as a result of landmark placement and were completed in two steps: intra-operator and inter-operator.

5.2.2.1 Analysis One: intra-operator study

The average and range of coefficients of variation (CV) for the proportions of both the video images and database photographs is summarized in Table 5.1.

Table 5.1 Average and range of coefficients of variation for measurements of proportions in video images and database photographs for intra-operator study.

Video images				
	Average CV % (n = 10)		Range of CV %	
	Within Day	Between Day	Within Day	Between Day
Proportions	1.6	2.1	0.6-2.4	0.7-3.7
Database photographs				
	Average CV % (n = 10)		Range of CV %	
	Within Day	Between Day	Within Day	Between Day
Proportions	1.3	1.7	0.5-1.7	0.6-3.0

The summarized results follow the expected pattern in that there was less variation in proportions due to landmark placement on the database photographs compared to that of the video images as well as within day compared to between day. It is expected that the database photographs would produce a smaller variation than the video images because although the video images were high resolution, they were still slightly more pixelated and thus compared to the photographs it was more difficult to place the landmarks on these images.

5.2.2.2 Analysis Two: inter-operator study

Table 5.2 displays the summarized results for each video image that was landmarked by the five operators. The average and range of coefficients of variation (CV) for the 59 proportions are listed along with the average and range of CV's once the outliers were removed. An outlier was defined as a CV that was above 10.0%.

Table 5.2 Each image was measured 6 times by each operator. The table lists the average (n=6) and range of coefficients of variation obtained by each operator for measurements of proportions of each video image for inter-operator study.

Video image 1					
n=6	Operator 1	Operator 2	Operator 3	Operator 4	Operator 5
Average CV %	3.2	8.0	4.8	6.7	10.0
Range of CV %	0.3-17.2	0.9-48.8	0.0-28.7	0.6-39.3	0.7-107.7
Average and range after removal of outliers (CV% above 10%)					
Average CV %	2.7	4.2	2.7	4.3	3.3
Range of CV %	0.3-9.0	0.9-10.0	0.0-9.7	0.6-9.8	0.7-10.0
Video image 2					
n=6	Operator 1	Operator 2	Operator 3	Operator 4	Operator 5
Average CV %	5.7	3.0	5.0	4.7	4.1
Range of CV %	0.7-24.0	0.5-11.1	0.8-21.6	0.5-16.7	0.0-21.5
Average and range after removal of outliers (CV% above 10%)					
Average CV %	4.2	2.7	3.4	3.5	3.1
Range of CV %	0.7-9.1	0.5-8.7	0.8-8.8	0.5-10.0	0.0-9.0
Video image 3					
n=6	Operator 1	Operator 2	Operator 3	Operator 4	Operator 5
Average CV %	4.5	3.2	7.0	4.4	6.5

Range of CV %	0.5-14.6	0.5-11.9	0.4-47.6	0.6-15.4	0.0-92.7
Average and range after removal of outliers (CV% above 10%)					
Average CV %	3.7	2.7	3.3	3.4	3.3
Range of CV %	0.5-9.6	0.5-9.1	0.4-8.4	0.6-10.0	0.0-8.8
Video image 4					
n=6	Operator 1	Operator 2	Operator 3	Operator 4	Operator 5
Average CV %	6.1	3.4	4.4	5.8	3.8
Range of CV %	1.2-19.1	0.6-15.3	0.5-21.0	0.6-24.1	0.8-17.6
Average and range after removal of outliers (CV% above 10%)					
Average CV %	4.8	3.0	3.3	4.3	3.1
Range of CV %	1.2-9.9	0.6-7.7	0.5-9.6	0.6-9.5	0.8-7.6
Video image 5					
n=6	Operator 1	Operator 2	Operator 3	Operator 4	Operator 5
Average CV %	3.2	9.0	2.9	4.7	4.2
Range of CV %	0.5-25.4	0.7-59.3	0.0-12.1	0.6-27.4	0.7-16.9
Average and range after removal of outliers (CV% above 10%)					
Average CV %	2.6	3.3	2.3	3.1	3.1
Range of CV %	0.5-10.0	0.7-8.9	0.0-9.3	0.6-8.1	0.7-9.7
Video image 6					
n=6	Operator 1	Operator 2	Operator 3	Operator 4	Operator 5
Average CV %	3.3	5.5	3.9	4.7	3.9
Range of CV %	0.5-15.4	0.5-34.7	0.6-14.0	1.0-20.8	0.5-16.5
Average and range after removal of outliers (CV% above 10%)					
Average CV %	2.8	3.6	3.4	3.5	3.4
Range of CV %	0.5-8.5	0.5-9.9	0.6-10.0	1.0-8.4	0.5-9.5
Overall average % after removal of outliers (CV% above 10%) for five operators and six images (n=30)					
3.3					

Operator 1 was the only operator of the group to have previous experience in the process, yet only had the smallest average CV of the group of operators in 20% of the video images tested. This was not expected but does not necessarily mean that the experienced operator placed the landmarks incorrectly, only that there was more variation in relation to other operators to where they did place the landmarks.

The average and range of CV's3 are larger than that obtained in the intra-operator error and this is expected due to the variability introduced with additional operators. The number of proportions included in the inter-operator study was much larger, 59 compared to 6 in the intra-operator study. Some of these proportions use landmarks which are difficult to place reproducibly and because of this the average coefficients of variation are larger. In practice, proportion measurements which show very large variations would be excluded from any system of image comparison or database searching. In the present study, these were termed "outliers" and were defined as proportions giving CV% values greater than 10%. Once the outliers were removed from the average CV, the data still showed larger average CV's than occurred as a result of the intra-operator study.

5.3 Errors in Facial Proportions Resulting from the Process of Taking Photographs

5.3.1 Materials

5.3.1.1 Subjects

The study group consisted of 14 volunteers (5 male, 9 female) from amongst the students and staff of the Forensic Medicine and Science Section, University of Glasgow, who were selected randomly, avoiding bias with respect to race, gender, or age. Photographs were taken of three men and 7 women in the study cohort and in addition 4 men and 6 women from the cohort participated in taking of photographs. No one was compensated for time spent in the study and permission was granted by the volunteers for their photographs to be used in this research.

5.3.1.2 Equipment

The photographs required for this study were taken with a SLR Olympus E-500 8.0 megapixel digital camera with automatic settings and an Olympus 40-150mm lens set on a tripod in a laboratory setting with the lens axis in the horizontal plane (Figure 5.1). The lens was set to a focal length of 52mm (the focal length specified by the police in the creation of suspect identity photographs) and a piece of adhesive tape was placed on the lens so that it could not be changed. Facial photographs were taken from the frontal position at 0º and then the tripod was moved in an arc and photographs were taken of the right side of the face at +20 degrees and +45 degrees. The subjects were seated in a stationary chair with a stand located in front of the chair for the subject to rest their chin in order to keep their head stationary. A protractor printed on transparency paper and a metre stick was used to mark the three angles on the floor. The locations of the tripod feet for each angle were marked with tape.

Figure 5.3 Experimental set up for taking four series of photographs

To ensure the photographs were all taken at the same vertical level, the camera was attached to a tripod and placed approximately 1.5 metres from the chair. The subjects were placed in the frontal position. To standardise the position of the head, the volunteers were instructed to only rest their chin on the stand but not to put any weight on it. A ruler was held to the side of their face, lining up their tragus with the lowest part of the orbit in the Frankfort horizontal position. The subject was instructed to hold that position with eyes open and mouth closed, but relaxed, whilst the photographs were being taken.

The legs of the tripod were set but height of the camera could be adjusted so that the focal point in the view finder lined up with the nasion and, if needed, the arm on the tripod could be used to swivel the camera to find this point. The subject kept their head still while the tripod was moved to the left at 20 and 45 degree rotations in order to get photographs of the right side of the volunteers' face. To keep the subjects from following the camera with their eyes while it was being rotated, a small stand was placed on a table behind the camera and the subject was asked to focus their attention on that.

5.3.2 Methodology

5.3.2.1 Procedure

Four series of photographs were taken.

1. Ten volunteers took photographs of the same subject at three different angles (0°, 20°, 45°). Two examples are shown in Figure 5.4.

Figure 5.4 Photographs taken by two different operators at angles 0°, 20° and 45°. The photographs were not cropped in order to show the differing placement of the subject in the viewfinder.

2. One volunteer took photographs of ten different subjects at three different angles (0°, 20°, 45°). Two examples are shown in Figure 5.5.

Figure 5.5 Photographs of two different subjects taken by one volunteer

3. One volunteer took photographs of one subject (a) at three different angles (0°, 20°, 45°) six times. The tripod was not adjusted.

4. One volunteer took photographs of one subject (b) at three different angles (0°, 20°, 45°) six times. The height of the tripod was adjusted to account for the sloping floor. An example is shown in Figure 5.6.

Figure 5.6 Photographs taken by one volunteer at angles 0°, 20° and 45°.

The first series of photographs were taken to assess the variation in facial measurements resulting from the use of different operators/photographers. This simulated the real-life situation in which suspect photographs are taken by different operators in police stations across Britain. Ten different volunteers were asked to take three photographs (0°, 20° and 45°) of the same subject resulting in a total of 30 photographs. The volunteers who had agreed to take part were given a set of instructions, listed below, to read and follow (Figure

5.7). They were not allowed to move any equipment apart from adjusting the height and angle of the camera on the tripod to line up the focal point with the nasion.

Standard Operating Procedure

1. Line up tripod legs on the three dots representing the frontal view point of at an angle of 0º.

2. Ensure that the camera has remained in the horizontal position by checking the spirit level bubble.

3. Sit the subject on the chair so they are facing the camera and so that their chin is resting on the stand.

4. Using a ruler for help with alignment, guide their head in the Frankfort horizontal plane. This is a horizontal line from the tragus on the ear to the bottom of the orbit. A reference photograph is available for guidance.

5. Instruct them that they are to remain still in this position with a relaxed face, their mouth closed, and looking straight ahead.

6. The focal length of the camera is set and should not be touched.

7. Keeping the legs of the tripod set, adjust the height of the camera so that the centre dot in the view finder lines up with the Nasion. The Nasion is located at the midpoint of the nasofrontal suture. A reference photograph is available for guidance. Do not change any settings of the camera or tripod. If you need to swivel the camera to centre the viewpoint, use the arm on the tripod only.

8. The camera is set on auto focus. Press the shutter button half way down to focus and then fully to take the photograph.

9. Move and line up the tripod legs to the blue dots which represent a camera angle of 20º and repeat steps 2-8.

10. Repeat steps 2-8 using the green dots, for a camera angle of 45º.

Figure 5.7 Standard operating procedure given to each volunteer before they took any photographs

The second series of photographs was taken to assess the variation in facial measurements resulting from the fact that faces differ from each other. One operator took photographs of ten different volunteers at three angles (0°, 20°, 45°) to give a total of 30 photographs. The standard operating procedure shown in Figure 5.8 was followed and only the height and angle of the camera were adjusted to centre the viewpoint on the nasion.

The third series of photographs was taken by one volunteer of one subject at three different angles (0°, 20°, 45°). The subject then stood up and left the rig. The subject was then relocated in the rig prior to the next set of photographs. This sequence was repeated six times to give a total of 18 photographs. This series was designed to assess the variation in facial measurements resulting from the actual process of setting up the rig, even if every effort was made to avoid variation. In this series, neither the camera nor tripod was adjusted. Due to a sloping floor, the focal point in the viewfinder did not always align with the nasion at angles 20° and 45°.

The fourth and final series of photographs was taken by one volunteer of one subject at three different angles six times for at total of 18 photographs. This differed from the third series insofar as the tripod height and angle were adjusted to ensure that the focal point in the viewfinder was aligned with the nasion at all three angles. The subject used in this series was different from that used in the third series.

5.3.2.2 Data Collection

The photographs were downloaded as Jpeg files and a measurement programme produced in-house, Facial Identification Centre Version 0.32 © Forensic Medicine and Science, Glasgow University, was used to place landmarks on the image, carry out linear measurements and calculate proportions. A single experienced operator carried out all landmark placements on these photographs.

The software was simple to use and very basic. Initially a set of metrics is programmed to be used as a template. This includes any landmark that the operator wishes to be placed and any line and proportion that will be determined from these landmarks. When a new analysis is started all that is needed is the image and then the saved metrics file is applied. Multiple images can be added to each new analysis. Landmarks are placed by the operator with the mouse in the location the operator wishes. Once landmark placement on all images is finished, the analysis can then be exported into an Excel™ spreadsheet and includes the x, y coordinates, lengths of the lines between landmarks and the calculated proportions to two decimal places.

A total of 37 landmarks (as described in Chapter 3) were placed and 59 proportions calculated. In this study, angles were not calculated.

5.3.3 Results

5.3.3.1 General

Once the four series of photographs were taken, landmark placement was carried out and the statistical analyses described below were performed. Blank cells in the tables, for which the number of measurements was zero (n = 0), occurred when landmarks could not be placed on photographs due to the camera angle. When only one measurement was available (n=1), it was not meaningful to calculate statistics such as the mean and standard deviation. All landmarks were easily located on frontal view (0°) photographs, with the exception of the gnathion, because the chinrest used by the subjects covered or changed the position of the landmark. For the same reason, the gnathion could not be located on photographs taken at a camera angle of 20°. In addition, landmarks on the left side of the face apart from those around the left eye were hidden with the change in camera angle. At 45°, all landmarks on the left side of the face were hidden although the gnathion could be located.

5.3.3.2 Analysis One: variation in facial measurements resulting from use of different operators

A single subject was photographed by ten different operators at each of the three camera angles. Fifty nine facial proportions were measured according to the visibility of the relevant landmarks and the coefficient of variation (%) was calculated for each proportion at each camera angle (Table 5.3). The largest possible ‘n’ was 10.

Table 5.3 The coefficient variation (CV %) of 59 proportions obtained from photographs of a subject taken by 10 operators at three different camera angles. Proportions with “ ’ ” indicate the left side of the face.

Multiple operators taking photographs of one subject

		0°		20°		45°	
		n	Average CV %	n	Average CV %	n	Average CV %
1.	go-go'/n-gn	0		0		0	
2.	sn-gn/n-sto	0		0		9	2.3
3.	n-prn/g-pg	10	4.5	9	3.1	10	2.6
4.	gn-go/n-gn	0		0		9	2.6
5.	gn-go'/n-gn	0		0		0	
6.	al-al'/ex-ex'	10	1.6	0		0	
7.	al-al'/n-sn	10	2.8	0		0	
8.	sa-sba/n-gn	0		0		8	2.1
9.	sa'-sba'/n-gn	0		0		0	
10.	n-sn/sa-sba	10	3.6	10	1.5	9	2.1
11.	n-sn/sa'-sba'	5	4.1	0		0	
12.	ex-ex'/go-go'	8	33.5	1		0	
13.	ex-n/ex-sto	10	2.2	10	1.4	10	1.8
14.	ex'-n/ex'-sto	10	1.6	10	35.4	0	
15.	ex-n/n-sto	10	2.2	10	1.9	10	2.5
16.	ex'-n/n-sto	10	2.4	10	38.2	0	
17.	ex-sto/n-sto	10	2.3	10	1.4	10	1.7
18.	ex'-sto/n-sto	10	2.4	10	6.9	0	
19.	en-ex/ps-pi	10	3.5	10	4.2	10	4.5
20.	en'-ex'/ps'-pi'	10	3.2	10	5.2	0	
21.	en-en'/ex-ex'	10	0.9	10	50.9	0	
22.	pi-or/en-ex	10	8.5	10	13.5	10	12.6
23.	pi'-or'/en'-ex'	10	7.7	10	9.4	0	
24.	sa-sba/pa-obi	1		10	1.8	9	5.0
25.	sa'-sba'/pa'-obi'	0		0		0	
26.	cph-cph'/sn-ls	10	8.4	10	10.8	10	11.3
27.	ls-sto/ch-ch'	10	5.1	10	4.6	0	

28. sto-li/ch-ch'	10	7.0	10	6.6	0	
29. sn-sto/sto-sl	10	4.3	10	14.5	10	7.6
30. li-sl/sn-ls	10	12.2	10	44.7	10	13.1
31. sbal-sn/sn-prn	10	8.0	10	9.9	10	18.6
32. sbal'-sn/sn-prn	10	13.7	1		0	
33. sl-gn/sto-gn	0		0		9	3.0
34. ex-go/go-go'	8	7.6	1		0	
35. ex'-go'/go-go'	8	7.8	1		0	
36. n-sn/n-sto	10	1.4	10	1.6	10	1.4
37. n-gn/n-sto	0		0		9	1.3
38. en-al/ex-ch	10	1.3	10	3.8	10	1.6
39. en'-al'/ex'-ch'	10	3.1	0		0	
40. obi-ch/g-sa	10	5.4	10	1.7	9	1.3
41. obi'-ch'/g-sa'	6	3.4	0		0	
42. sbal-ls/n-al	10	5.5	10	4.8	10	4.0
43. sbal'-ls/n-al'	10	5.9	0		0	
44. pi-al/sa-ex	10	6.5	10	4.0	9	2.0
45. pi'-al'/sa'-ex'	6	4.2	0		0	
46. ex-obi/ex-ch	10	7.3	10	2.7	10	1.9
47. ex'-obi'/ex'-ch'	10	6.6	0		0	
48. ex-al/ch-gn	0		0		9	2.3
49. ex'-al'/ch'-gn	0		0		0	
50. ch-ls/n-prn	10	6.5	10	2.4	10	3.7
51. ch'-ls/n-prn	10	6.9	10	7.0	0	
52. al-ls/ch-gn	0		0		9	3.7
53. al'-ls/ch'-gn	0		0		0	
54. ch-li/ex-ch	10	3.4	10	4.3	10	4.8
55. ch'-li/ex'-ch'	10	3.0	10	8.4	0	
56. ex-sto/ex-ch'	10	1.3	10	1.0	0	
57. ex'-sto/ex'-ch	10	0.9	10	3.5	0	
58. sn-gn/ex-gn	0		0		9	1.9
59. sn-gn/ex'-gn	0		0		0	
Average CV%		5.4		9.7		4.4
Range				1-50.9		

of CV%	0.9-33.5		1.3-18.6
Average and range after removal of outliers (CV% above 10%)			
Average CV%	4.4	4.1	2.0
Range of CV%	0.9-8.5	1-9.9	1.3-7.6

In this set of measurements, the same individual was photographed, the photographer/operators were instructed to follow closely the set procedure and all photographic measurements were carried out by a single experienced operator, but nevertheless a large range of variation in proportions was obtained. These variations can be attributed to the errors introduced by the camera operators when implementing the photography procedure, for example, differences in where they placed the camera relative to the subject, plus any uncertainty associated with landmark placement. The average CV% for proportions obtained from 0° photographs was 5.4% whereas the separate study made of errors due to landmark placement (Section 5.2.1.1), showed that an experienced operator placing landmarks produces proportion measurements which vary by 1.6% on average (under repeatability conditions). This indicates that the photography process may introduce significant variation compared to the measurement process. If the CV% values in Table 5.3 are examined, it can be seen that a few proportions give much bigger errors than the others, for example, ex-ex'/go-go' has a CV of 33.5%. If a system of facial comparison were ever to be introduced, these unreliable proportion measurements would be excluded. A parallel type of data manipulation is to consider them as outliers and exclude them from the data analysis. By doing this the average CV is reduced to 4.4%, which is still significantly larger than the error produced by the landmark placement process.

As mentioned, proportion ex-ex’/go-go’ at frontal view, had the highest CV at 33.5%, which can be attributed to the problem of obtaining a consistent location of the gonion, which has shown to be a difficult landmark to place. Any error acquired from placing that landmark will be propagated when using the landmark to calculate a proportion, particularly if the landmark is used more than once in calculating the proportion.

In general, most variation was observed for proportions obtained from photographs taken at a camera angle of 20º relative to the frontal view.

There were a few proportions at a rotational angle of 20º that had unusually high coefficients of variation compared to the other proportions at the same angle. These were ex'-n/ex'-sto (35.4%), ex'-n/n-sto (38.2%), en-en'/ex-ex' (50.9%), and li-sl/sn-ls (44.7%). As all ten photographs were included in the calculation, the greatest contribution to the variation was most likely to be an error caused by landmark placement.

5.3.3.3 Analysis Two: variation in facial measurements resulting from different subjects

Table 5.4 contains the coefficients of variation for facial proportions resulting from different subjects, when one volunteer took photographs of ten different subjects.

Of the four series of photographs, the highest coefficients of variation occurred for proportions in which a single volunteer took photographs of ten different subjects. This is to be expected, and indeed if this was not the case then it would negate any chance of determining any differences between individuals. Ideal proportions would have a large CV because it shows that the proportion has a high variation and therefore a good indicator of making a distinction between two individuals. A proportion with a low CV may not be as good as indicator if there is little variation among different faces.

In this analysis, the majority of any variation will be caused as a result of faces being different from each other; however, there will still be contributions resulting from the landmark placement and the photographic process. There were no trends with respect to the effect of camera angle on the range of variation.

Table 5.4 The coefficient variation (CV %) of 59 proportions for the series of photographs in which a single volunteer took photographs of 10 different subjects. Proportions with " ' " indicate the left side of the face.

One operator taking photographs of 10 different subjects

	0°		20°		45°	
	n	Average CV %	n	Average CV %	n	Average CV %
1. go-go'/n-gn	0		0		0	
2. sn-gn/n-sto	0		0		4	7.6
3. n-prn/g-pg	8	9.4	7	7.1	9	8.0
4. gn-go/n-gn	0		0		4	3.6
5. gn-go'/n-gn	0		0		0	
6. al-al'/ex-ex'	10	10.2	2	30.6	0	
7. al-al'/n-sn	9	35.9	2	21.8	1	
8. sa-sba/n-gn	0		0		1	
9. sa'-sba'/n-gn	0		0		0	
10. n-sn/sa-sba	5	7.1	7	9.6	6	6.2
11. n-sn/sa'-sba'	5	3.5	0		0	
12. ex-ex'/go-go'	4	39.8	1		0	
13. ex-n/ex-sto	9	4.3	9	4.0	9	4.2
14. ex'-n/ex'-sto	9	4.4	9	11.2	0	
15. ex-n/n-sto	9	7.0	9	6.9	9	9.5
16. ex'-n/n-sto	9	7.0	9	12.5	0	
17. ex-sto/n-sto	9	4.0	9	5.4	9	7.2
18. ex'-sto/n-sto	9	3.2	9	3.1	0	
19. en-ex/ps-pi	10	10.8	10	11.3	10	13.9
20. en'-ex'/ps'-pi'	10	12.4	10	12.3	0	
21. en-en'/ex-ex'	10	48.5	10	42.6	0	
22. pi-or/en-ex	10	18.5	10	19.4	10	16.3
23. pi'-or'/en'-ex'	10	21.8	10	20.5	0	
24. sa-sba/pa-obi	4	6.2	6	3.8	6	7.8
25. sa'-sba'/pa'-obi'	4	5.9	0		0	
26. cph-cph'/sn-ls	10	32.2	10	39.3	10	25.1
27. ls-sto/ch-ch'	10	35.5	10	37.6	0	

28. sto-li/ch-ch'	10	13.3	9	14.2	0	
29. sn-sto/sto-sl	10	19.4	10	23.5	10	12.1
30. li-sl/sn-ls	10	66.2	9	47.5	10	21.6
31. sbal-sn/sn-prn	10	22.3	9	29.3	10	36.8
32. sbal'-sn/sn-prn	10	25.5	3	12.0	1	
33. sl-gn/sto-gn	0		0		4	9.4
34. ex-go/go-go'	4	56.1	1		0	
35. ex'-go'/go-go'	4	64.3	1		0	
36. n-sn/n-sto	9	4.6	9	8.3	9	4.8
37. n-gn/n-sto	0		0		4	1.2
38. en-al/ex-ch	10	8.9	10	16.2	10	10.9
39. en'-al'/ex'-ch'	10	14.8	2	59.1	0	
40. obi-ch/g-sa	5	10.2	7	11.0	6	8.1
41. obi'-ch'/g-sa'	5	10.4	0		0	
42. sbal-ls/n-al	9	33.0	9	51.6	9	34.6
43. sbal'-ls/n-al'	9	32.8	1		0	
44. pi-al/sa-ex	5	8.9	7	24.8	6	13.4
45. pi'-al'/sa'-ex'	5	13.7	0		0	
46. ex-obi/ex-ch	7	9.4	9	32.2	9	5.5
47. ex'-obi'/ex'-ch'	8	15.3	0		0	
48. ex-al/ch-gn	0		0		4	7.0
49. ex'-al'/ch'-gn	0		0		0	
50. ch-ls/n-prn	9	36.2	8	14.9	9	12.5
51. ch'-ls/n-prn	9	36.4	8	26.2	0	
52. al-ls/ch-gn	0		0		4	17.0
53. al'-ls/ch'-gn	0		0		0	
54. ch-li/ex-ch	10	11.7	9	7.7	10	7.0
55. ch'-li/ex'-ch'	10	12.2	9	14.7	0	
56. ex-sto/ex-ch'	10	2.4	10	2.4	0	
57. ex'-sto/ex'-ch	10	2.5	10	2.4	0	
58. sn-gn/ex-gn	0		0		4	4.6
59. sn-gn/ex'-gn	0		0		0	
Average CV%		19.1		19.4		11.7
Range of CV%		2.4-66.2		2.4-59.1		1.2-36.8

5.3.3.4 Analysis Three: variation in facial measurements resulting from same subject

Table 5.5 contains the data when one volunteer took photographs of one subject (A) at three different angles making no adjustments to the tripod and Table 5.6 contains the data when a volunteer took photographs of one subject (B) adjusting the height of the tripod to accommodate the sloping floor.

Table 5.5 The coefficient variation (CV %) of 59 proportions for the series of photographs in which a single volunteer took photographs of a single subject. The height of the camera was not adjusted to accommodate the sloping floor. Proportions with " ' " indicate the left side of the face.

One operator taking photographs of one subject: multiple times (A)

		0°		20°		45°	
		n	Average CV %	n	Average CV %	n	Average CV %
1.	go-go'/n-gn	0		0		0	
2.	sn-gn/n-sto	0		0		0	
3.	n-prn/g-pg	6	2.3	5	2.5	5	1.3
4.	gn-go/n-gn	0		0		0	
5.	gn-go'/n-gn	0		0		0	
6.	al-al'/ex-ex'	6	0.0	0		0	
7.	al-al'/n-sn	6	2.3	0		0	
8.	sa-sba/n-gn	0		0		0	
9.	sa'-sba'/n-gn	0		0		0	
10.	n-sn/sa-sba	6	3.6	6	1.6	6	2.2
11.	n-sn/sa'-sba'	6	2.5	0		0	
12.	ex-ex'/go-go'	0		0		0	
13.	ex-n/ex-sto	6	1.6	6	1.5	6	1.1
14.	ex'-n/ex'-sto	6	2.0	6	2.9	0	
15.	ex-n/n-sto	6	1.6	6	1.1	6	1.0
16.	ex'-n/n-sto	6	1.5	6	2.2	0	
17.	ex-sto/n-sto	6	2.1	6	1.4	6	1.3
18.	ex'-sto/n-sto	6	1.5	6	1.3	0	
19.	en-ex/ps-pi	6	4.4	6	3.4	6	3.5
20.	en'-ex'/ps'-pi'	6	1.8	6	1.7	0	
21.	en-en'/ex-ex'	6	1.5	6	1.8	0	

22. pi-or/en-ex	6	8.9	6	6.7	6	7.5
23. pi'-or'/en'-ex'	6	5.7	6	7.4	0	
24. sa-sba/pa-obi	6	2.1	6	2.4	6	2.9
25. sa'-sba'/pa'-obi'	6	1.0	0		0	
26. cph-cph'/sn-ls	6	6.7	6	4.9	6	4.8
27. ls-sto/ch-ch'	6	4.0	6	11.0	0	
28. sto-li/ch-ch'	6	6.9	6	5.8	0	
29. sn-sto/sto-sl	6	4.4	6	21.7	6	21.9
30. li-sl/sn-ls	6	14.1	6	66.9	6	58.4
31. sbal-sn/sn-prn	6	11.6	6	13.0	6	12.4
32. sbal'-sn/sn-prn	6	2.6	0		0	
33. sl-gn/sto-gn	0		0		0	
34. ex-go/go-go'	0		0		0	
35. ex'-go'/go-go'	0		0		0	
36. n-sn/n-sto	6	1.4	6	0.8	6	1.2
37. n-gn/n-sto	0		0		0	
38. en-al/ex-ch	6	0.9	6	1.3	6	1.4
39. en'-al'/ex'-ch'	6	0.9	0		0	
40. obi-ch/g-sa	6	2.9	6	2.4	6	2.6
41. obi'-ch'/g-sa'	6	4.3	0		0	
42. sbal-ls/n-al	6	2.4	6	4.0	6	2.0
43. sbal'-ls/n-al'	6	2.6	0		0	
44. pi-al/sa-ex	6	3.4	6	3.7	6	3.4
45. pi'-al'/sa'-ex'	6	5.1	0		0	
46. ex-obi/ex-ch	6	2.2	6	1.9	6	2.1
47. ex'-obi'/ex'-ch'	6	4.0	0		0	
48. ex-al/ch-gn	0		0		0	
49. ex'-al'/ch'-gn	0		0		0	
50. ch-ls/n-prn	6	4.5	6	2.8	6	2.8
51. ch'-ls/n-prn	6	5.5	6	4.5	0	
52. al-ls/ch-gn	0		0		0	
53. al'-ls/ch'-gn	0		0		0	
54. ch-li/ex-ch	6	3.1	6	1.8	6	3.0

55. ch'-li/ex'-ch'	6	3.7	6	4.6	0	
56. ex-sto/ex-ch'	6	0.9	6	1.1	0	
57. ex'-sto/ex'-ch	6	0.7	6	1.1	0	
58. sn-gn/ex-gn	0		0		0	
59. sn-gn/ex'-gn	0		0		0	
Average CV%		3.5		6.0		6.8
Range of CV%		0.0-14.1		0.8-66.9		1.0-58.4
Average and range after removal of outliers (CV% above 10%)						
Average CV%		3.0		2.8		2.6
Range of CV%		0.0-8.9		0.8-7.4		1.0-7.5

Lower levels of variation accompany this series of photographs of photographs that were taken of a single subject by a single volunteer. The camera was not adjusted to account for the sloping floor and therefore there will still be a level of uncertainty. This low level of variation is to be expected because as the face in the photographs is the same, it would not be expected to get the same variation as between different people. The photographer is also held constant therefore eliminating the error acquired from using different photographers.

In this series of photographs, the proportion with the highest percentage of variation across the board is li-sl/sn-ls. While the frontal view had a much lower CV percentage at 14.1% compared to the other rotations, both 20º (66.9%) and 45º (58.4%) had high levels of variation and was the highest in this series for all proportions. The two linear measurements involved in this ratio are relatively short distances. It is known that a small error in landmark placement between two points that are a short distance apart from each other will yield a large error. Therefore, because both lines are short, this ratio may not be discriminating enough to use.

Table 5.6 The coefficient variation (CV %) of 59 proportions for the series of photographs in which a single volunteer took photographs of a single subject. The height of the camera was adjusted to accommodate the sloping floor. Proportions with “ ’ ” indicate the left side of the face.

One operator taking photographs of one subject: multiple times (B)

		0°		20°		45°	
		n	Average CV %	n	Average CV %	n	Average CV %
1.	go-go'/n-gn	0		0		0	
2.	sn-gn/n-sto	0		0		6	2.3
3.	n-prn/g-pg	6	4.5	6	2.6	6	3.0
4.	gn-go/n-gn	0		0		6	3.1
5.	gn-go'/n-gn	0		0		0	
6.	al-al'/ex-ex'	6	2.0	0		0	
7.	al-al'/n-sn	6	1.7	0		0	
8.	sa-sba/n-gn	0		0		6	1.6
9.	sa'-sba'/n-gn	0		0		0	
10.	n-sn/sa-sba	5	5.1	6	1.6	6	1.5
11.	n-sn/sa'-sba'	0		0		0	
12.	ex-ex'/go-go'	5	4.2	0		0	
13.	ex-n/ex-sto	6	1.8	6	1.5	6	2.0
14.	ex'-n/ex'-sto	6	1.5	6	1.9	0	
15.	ex-n/n-sto	6	2.5	6	1.1	6	2.2
16.	ex'-n/n-sto	6	2.1	6	2.0	0	
17.	ex-sto/n-sto	6	2.2	6	0.8	6	1.0
18.	ex'-sto/n-sto	6	2.6	6	1.8	0	
19.	en-ex/ps-pi	6	7.1	6	2.5	6	4.5
20.	en'-ex'/ps'-pi'	6	4.5	6	5.0	0	
21.	en-en'/ex-ex'	6	0.0	6	3.0	0	
22.	pi-or/en-ex	6	11.0	6	7.1	6	8.0
23.	pi'-or'/en'-ex'	6	8.8	6	7.5	0	
24.	sa-sba/pa-obi	1		6	3.2	6	1.8
25.	sa'-sba'/pa'-obi'	0		0		0	
26.	cph-cph'/sn-	6	6.6	6	8.1	5	7.8

ls						
27. ls-sto/ch-ch'	6	4.8	6	9.5	0	
28. sto-li/ch-ch'	6	4.1	6	5.7	0	
29. sn-sto/sto-sl	6	4.4	6	7.3	6	2.6
30. li-sl/sn-ls	6	6.1	6	8.1	6	5.6
31. sbal-sn/sn-prn	6	13.6	6	10.5	6	14.2
32. sbal'-sn/sn-prn	6	13.4	1	40.8	0	
33. sl-gn/sto-gn	0		0		6	2.9
34. ex-go/go-go'	5	8.5	0		0	
35. ex'-go'/go-go'	5	9.7	0		0	
36. n-sn/n-sto	6	1.2	6	1.1	6	1.1
37. n-gn/n-sto	0		0		6	1.0
38. en-al/ex-ch	6	2.4	6	1.5	6	2.3
39. en'-al'/ex'-ch'	6	2.8	0		0	
40. obi-ch/g-sa	5	5.8	6	1.8	6	2.4
41. obi'-ch'/g-sa'	0		0		0	
42. sbal-ls/n-al	6	2.7	6	3.2	6	2.6
43. sbal'-ls/n-al'	6	3.0	0		0	
44. pi-al/sa-ex	5	12.0	6	5.5	6	2.9
45. pi'-al'/sa'-ex'	0		0		0	
46. ex-obi/ex-ch	6	9.2	6	4.5	6	1.5
47. ex'-obi'/ex'-ch'	0		0		0	
48. ex-al/ch-gn	0		0		6	1.7
49. ex'-al'/ch'-gn	0		0		0	
50. ch-ls/n-prn	6	6.5	6	5.0	6	3.5
51. ch'-ls/n-prn	6	7.4	6	5.8	0	
52. al-ls/ch-gn	0		0		6	3.7
53. al'-ls/ch'-gn	0		0		0	
54. ch-li/ex-ch	6	3.5	6	3.4	6	3.4
55. ch'-li/ex'-ch'	6	4.1	6	8.3	0	
56. ex-sto/ex-ch'	6	0.6	6	0.5	0	
57. ex'-sto/ex'-ch	6	1.7	6	1.7	0	
58. sn-gn/ex-gn	0		0		6	2.1
59. sn-gn/ex'-gn	0		0		0	
Average				5.3		3.3

CV%	5.0		
Range of CV%	0.0-13.6	0.5-40.8	1.0-14.2
Average and range after removal of outliers (CV% above 10%)			
Average CV%	4.2	4.0	2.9
Range of CV%	0.0-9.7	0.5-9.5	1.0-8.0

This series of photographs produced a higher CV for the frontal view than those used for Analysis Three. In this series of photographs, the camera was adjusted to account for the sloping floor and the focal point remained focused on the nasion, introducing an additional operation with an additional contribution to the overall variation. The CV for photographs at 45° was marginally smaller that for frontal photographs, perhaps because corrections were made for the slope in the floor, even though an additional operation was introduced.

Overall, the last two series of photographs gave the lowest variations. This low level of variation is to be expected because a single individual was the subject of the photographs, a single volunteer took the photographs, and the focal point remained constant. There will still be an uncertainty caused by landmark placement and minimal uncertainty caused by angle rotation, but not to the degree that was found earlier and that is shown in the tables.

The practical implications of this for a usable system are that operator involvement should be reduced to a minimum by using a fixed rig when taking photographs, perhaps with separate cameras permanently mounted for each angle required.

5.4 Estimation of Uncertainty

The following contributions to the uncertainty of anthropometric measurements are summarised from Figure 5.1 and placed into Table 5.7. The uncertainty assessment must include all contributions and not just those which were observed in one particular laboratory on a given occasion.

Table 5.7 Contributions to the uncertainty of anthropometric measurements

	Description	Uncertainty Estimated?
1	Camera focal length fixed at 52 mm	fixed at 52 mm - effect on facial measurements assumed to be negligible
2	Rig used for taking photographs: involves:- distance from subject axis of rotation position height of camera relative to subject	 Fixed Fixed Fixed on nasion
3	Photographer (different operators will interpret and implement the method differently)	Table 5.3
4	Subject (orientation of head when photographs are being taken)	Table 5.5 B
5	Photograph quality (resolution, contrast)	Not known
6	Subject (some landmarks are less distinct and more difficult to place depending on the physiognomy of the subject)	Table 5.2
7	Operator (experience)	Table 5.1
8	Measurement system (direct measurement of photographs, computer analysis of images on a monitor)	Not known

Combined uncertainty [155] (single experienced operator) includes contributions from photography and landmark placement:

$$\text{At } 0^\circ\text{, Uncertainty} = \surd\ [(4.2)^2 + (1.6)^2] = \surd 20.2 = 4.5\ \%$$

$$\text{At } 45^\circ\text{, Uncertainty} = \surd\ [(2.9)^2 + (1.6)^2] = \surd 10.97 = 3.3\ \%$$

Combined uncertainty (multiple operators) includes contributions from photography and landmark placement:

$$\text{At } 0^\circ\text{, Uncertainty} = \surd\ [(4.4)^2 + (3.3)^2] = \surd 30.25 = 5.5\ \%$$

$$\text{At } 45^\circ\text{, Uncertainty} = \surd\ [(2.0)^2 + (3.3)^2] = \surd 14.89 = 3.9\ \%$$

The resulting combined uncertainty observes two things:

1. Variation in facial proportions is less at a facial pose of 45° than at full frontal, 0°.

2. Variation in facial proportions is less with a single (experienced) operator than with multiple operators.

These observations have clear implications for use in the real world. Although not all of the contributions to uncertainty were known and therefore able to be included in the calculation of the overall estimation of uncertainty, the contributions that were included are relevant to this study. Police departments take their suspect identity photographs with a ¾ facial shot because this rotation was found to be more beneficial for identifications as it is more common to view someone that way [144] and the combined uncertainty determined in this study supports this. The uncertainty of facial proportions calculated at 45° was less than at the full frontal pose of 0° when using either a single or multiple operators.

The uncertainty of facial proportions calculated when a single experienced operator was used to take the photographs and place the landmarks were less at both facial rotations than when multiple operators were involved. This supports both the need for the involvement of a small number of operators and for those operators to be experienced.

The conclusion made from the two observations stated above is that to procure the least amount of uncertainty, the best case scenario for a 2D image comparison of this nature would entail a single operator with experience placing landmarks on a facial image rotated at 45°. At present this scenario is unlikely and although shown to produce a higher uncertainty, it is presumable that multiple operators will be involved at both the landmark placement and the photography stages. Therefore, it can be stated that at least a 5% estimation of uncertainty is expected due to the combined errors that occur and when a comparison of 2D images is conducted in this manner this estimation of uncertainty should be taken into account.

5.5 Discussion

5.5.1 Introduction

It is important that the overall uncertainty of any identification method includes all sources of uncertainty. This section of will first consider the uncertainty of facial proportions caused by landmark placement and the photography process and then examine the limitations uncovered while conducting this study. It will also review the significance of the uncertainty associated with anthropometric measurements and the implications for facial comparisons and for future studies in this field.

5.5.2 Landmark placement

The intra and inter-operator studies conducted were repeatability measurements and, although the conditions did not constitute the widest possible set of variations (for example, different equipment), they were relevant to the present study. It would be advantageous in future to expand the inter-operator study to include landmark placement on database photographs as well as on video images, which would allow an assessment to be made of the effect of image quality on landmark placement. Testing the uncertainty of landmark placement on high resolution video images, as in this study, probably results in an under-estimate of uncertainty, because the material does not represent the majority of video images that would be available in real world situations - most real world images would be of low resolution - but is relevant to the present study.

The Intra-operator study was designed to answer the question: “Will the same operator always produce the same anthropometric measurements from the same photograph?” Although the inter-operator study may be more relevant to possible real world situations, there is still a certain amount of variation of landmark placement expected with a single operator and it was therefore important to know how big this variation is. When two faces are compared using anthropometric measurements, the window allowed for a match should be at least as big as the error involved in making the measurements.

For the six proportions tested, the average CV was under 2% and this was smaller than for any single operator in the inter-operator study. The four landmarks used, and resulting proportions, were clearly a factor and a number of the 59 proportions used in the inter-operator study had a large degree of variation, raising the average CV. Nevertheless, the intra-operator study gives an indication of what is possible if these unreliable landmarks and proportions are excluded and if a single experience operator is used.

The inter-operator study was designed to answer the question: “Will different operators produce the same anthropometric measurements from the same photograph?” The inter-operator data was arranged to show the variation in anthropometric measurements obtained for the different images by different operators. The study therefore showed the variation that may occur for a given image if different operators in different places (e.g. different police offices) placed the landmarks. This is in contrast to assessing the performance of each operator by examining the extent of variation which results from placing landmarks on a series of different images. It is more relevant for this study to show the variation among facial proportions of the photographs rather than variation among operators because of the inexperience of the operators. An inter-operator study of this type should ideally use operators with similar levels of experience, in which case it would be prudent to establish their levels of expertise by determining the CVs of measurements they produce.

The inter-operator study therefore examines the variation in landmark placement by the operators for each video image tested, but does not test the accuracy of location of the landmarks in an absolute sense. To do this, it would be necessary for operators involved in anthropometric measurements around the world, with previous anthropometric knowledge and experience, to measure the images and produce a consensus value for each of the measurements. This mirrors a general problem in proficiency testing, since there is no way of determining the “true” value but only different ways of estimating what the true value might be.

It appeared as though the majority of proportions that achieved the highest variation were concentrated around the eyes and involved the landmarks of endocathion, exocanthion, orbital, palpebrale superius and palpebrale inferius. This could result from problems in placing these particular landmarks accurately, but may also be due to the short linear distances between the landmarks, since propagation of errors in landmark placement will lead to larger relative variations in the calculated proportions. Using a digital sliding calliper to measure photographs, intra and inter-observer studies were carried out to test reliability of measurements [156]. The intra-observer study resulted in low reliability measurements of ls-sto and n-sn. Their inter-observer study produced low reliability in measurements around the eye (en-en, ex-ex) and mouth (ch-ch), consistent with results encountered here, and height of the face (gn-n).

The primary goal of repeating the process of taking photographs at different angles was to address the problems discovered in the initial pilot study. These were primarily related to converting as many variables as possible to constants, all related to the axis of rotation. First it was necessary to set up a specific axis of rotation of the camera around the head of the subject and keep it constant. From this axis of rotation it was possible to record the distance to the camera as well as to define a set of camera angles relative to the head that originated from the axis of rotation. Finally, it was necessary to develop a course of action to keep the head located consistently on the axis of rotation in a way that was comfortable for the subject. This was perhaps the most difficult variable to accommodate. In the pilot study, the camera was fixed and the head rotated, but, in the absence of a rig built to keep the head in the Frankfort horizontal plane in a comfortable way, it was easier to keep the head immobile and to move the camera around it to change the angle. The methods in which these variables were taken into account were mentioned in the methodology section of this chapter and will not be repeated here.

As everything was carefully planned beforehand, there were few limitations while taking the photographs. The most significant limitation was to position the head in the Frankfort horizontal plane. It was simple to align the tragus with the low point of the lower part of the orbit; however, once the photographs were taken and examined, it was evident that there was some diagonal rotation that would preferably have been absent. The pose invariant facial recognition system by Chen and Cham, takes this type of rotation into account with a correction for what they call the ‘seesaw’ rotation [128]. Rotating the camera instead of the head was also easier on the subject and they did not have any trouble keeping their head stationary as there was a stand used as a chin rest. It was straightforward to adjust the camera rotation based on the marks made previously on the floor. Only three photographs were taken, which did not consume a lot of time or contribute towards making the subject tired. As there was no rig to hold the subject’s head in an exact place, when it came to multiple volunteers taking photographs of a single subject, there was still the possibility of the subject placing their head in different positions on the axis of rotation or holding it at different angles diagonally.

As the proportion values were used to judge the uncertainty contributions of other factors, it was necessary to look at any limitations in placing the landmarks. The photographs were clear and large enough to cause no difficulties in locating the majority of landmarks. The major exception to this was not due to the quality of the photograph but instead was caused by interference from the chin rest used to help keep the subject's head in one location. This was taken into consideration when setting up the chin rest, and it was thought that if subjects were instructed not to put their full head weight on the stand, any problems associated with landmark placement would be avoided. However, when the large photographs were examined on the computer monitor, the gnathion could not be located from either the frontal or 20º angle views.

In this study, the variations in proportions provide a quantitative analysis of how different variables influence the measurements. In order to minimize the amount of proportional variation, three specific parameters were kept constant. These include the distance the camera was placed from the axis of rotation, camera focal length and the use of a single operator to place landmarks. Keeping parameters constant makes it more likely to differentiate between proportional variation caused by rotation rather than the fact that the landmarks involved are on different planes.

In this study, the variation from four different variables was taken into account. These included four different series of photographs taken in an attempt to account for different variables that are true to life. The one type of variable not taken into account was the error shown in the proportions if different types of cameras were used. The procedure for taking photographs followed by police departments in the UK was followed as closely as possible. Although the police use a different kind of camera (film) than was used in this experiment (digital), the actual camera used would be standardized across the country. It seems unlikely that photographic images from a digital camera would produce different anthropometric measurements than those from a traditional film camera, although the latter would have to be printed and presumably digitised to enable a computer system to be used, thereby introducing more contributions to uncertainty. The question also arises concerning whether different digital cameras would ultimately result in different anthropometric measurements: although the focal length of the lens may be the same on each camera, the lens configuration may produce different degrees of image distortion in the horizontal and vertical planes, resulting in different images.

Another type of uncertainty not taken into account in these studies, which was mentioned by Fieller in his lecture notes [147] was the variation within the same face caused by the effects of ageing or different facial expressions. The effects of aging would affect facial comparisons using a long-established database of images, rather than the measurement process itself. One way to combat the effects of ageing, lighting, different camera angles, or different camera parameters is by creating a face average of an individual using as many images as possible. Although not tested using the anthropometric comparison method described in this thesis, image-averaging was met with success in artificial, as opposed to human, facial recognition [157, 158].

As could perhaps have been predicted, the majority of variability in this study was between the facial proportions of different subjects and between proportions for the same subject when photographs were taken by different volunteers. A lower degree of proportional variation within the same subject was obtained when photographs were taken by a single volunteer. The variables examined in this study, combined with the variables involved in placing landmarks, can be combined to assess the uncertainty in a real life scenario and to give an indication of the reliability of this kind of identification process.

The variables addressed in this study are important as the study mimics the present protocol for taking suspect photographs in police departments around the country. Attempts have been made to address the widespread variations in the suspect photographs that are currently obtained with the development of the FIND program to standardize the method in which photographs are taken. In general, the average CV was higher at the 20° rotation and lowest at the 45° rotation. This result is important because it shows that the angle rotations used in the FIND system would achieve the least amount of variation in anthropometric measurements. Having a country-wide database that is standardized with the same types of photographs and the same identifying information would be of substantial benefit to any type of facial identification system, especially for investigations that cross jurisdictions.

The next chapter focuses on the effectiveness that the landmarks and proportions have as the basis of a comparison method for frontal view database (known) and video (unknown) images.

6 Anthropometry Study of Paired Still and Video Images

6.1 Introduction

This chapter on anthropometry covers the materials used for the experiments, followed by the methodology, results and discussion of a research study undertaken on high resolution images. The aim was to test the hypothesis: "Using a comparison of anthropometric facial proportions, it is possible to discriminate between individuals of two samples." The objective was to derive measurements between specific landmarks on the face in both print and video media and incorporate them into a feature vector to use in statistical analysis to determine to statistical accuracy if identifications of an individual can be made based on these measurements.

6.2 Materials and Methodology

6.2.1 Subjects

A total of 199 images of Caucasian male police volunteers were available which had been used previously in research conducted by Bruce et al [85]. However, the ideal sample size for this study would be equal to the number of vector elements tested (59 proportions), multiplied by six, as determined by the statistical power calculation. The 199 images were made up of 80 video still images and 119 photographic images. The 80 video images consisted of 80 different faces and the 119 photographic images consisted of 119 different faces. The photographic images included the same 80 faces that were video images in addition to 39 extra faces not included as video images. The majority of images provided are unable to be shown in this thesis because of the sensitive nature of the material but an example of each type of image is provided in Figure 6.1, with the video image located on top and the photographs on the bottom. The photographs were of policemen, both retired and presently working and except for photographs which have already been published elsewhere, are not included in this work. Both sets of images were displayed from the frontal viewpoint, showing features from the neck up, in what appeared to be the format of police identification photographs and were taken on the same day. In this study the identity of the subjects in the video images was known and could be correlated with the corresponding photographic images. This means that identifications made on the basis of facial anthropometry could be designated as true or false identifications.

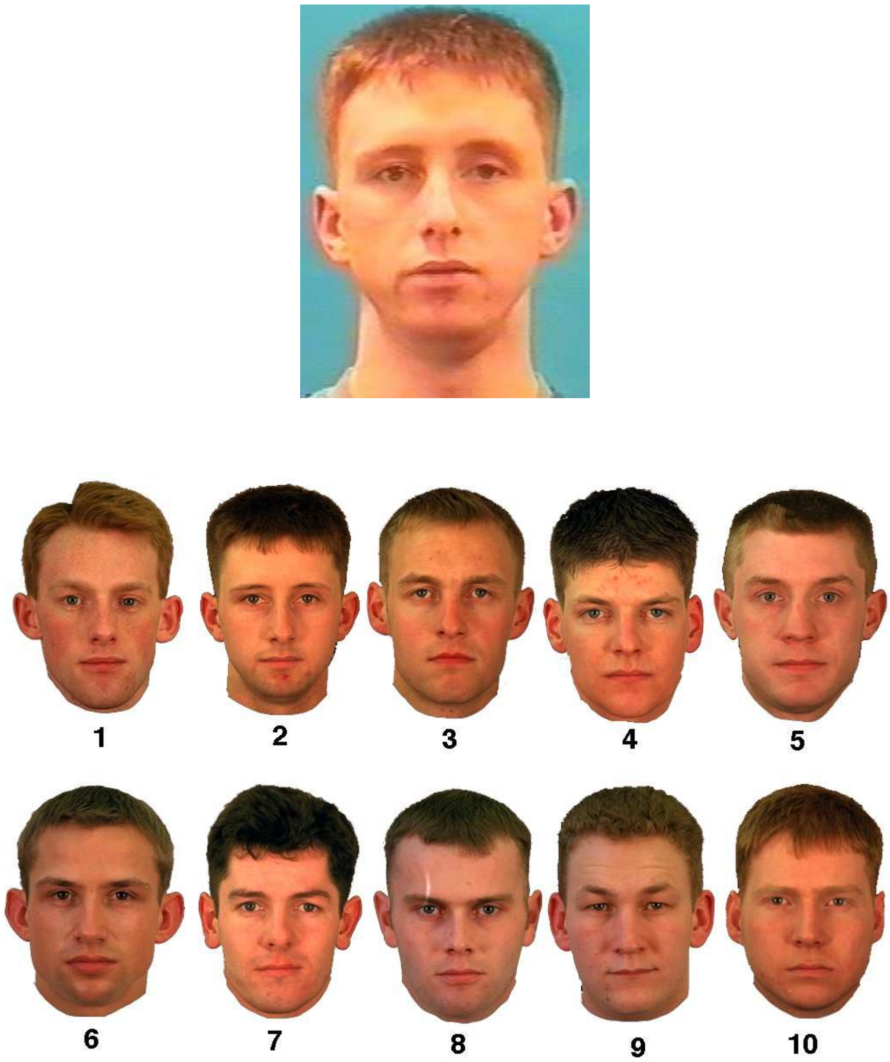


Figure 6.1 High resolution video image and selection of ten database photographs

The images obtained for this study were originally used for the purpose of testing a person's ability to recognize faces; therefore the angle of rotation of the subject relative to the camera was not a primary concern when the images were recorded and not controlled accordingly. One limitation resulting from this was that some of the images were rotated to the left or right by an estimated angle of 10 degrees from the frontal position, therefore comparing this to a full frontal image could not yield ideal results. Some of the subjects had their mouths open, again resulting in a comparison that could not yield a completely ideal result, even though some of the landmarks chosen are less subject to the effects of facial expression than others. Photographs with variables such as these however, more closely mimicked a “real life” situation. One positive feature of these video images was that they were recorded on the same day as the photographs. This meant that the study images did not have any of the possible facial changes which can occur due to time factors such as weight loss/gain or increase in age.

In forensic science, a sample is a subset of a population. The images provided are samples from the population of male Caucasians and will be referred as such. For the investigations carried out in this study, the following connotations were used. Video images are hereafter called Sample 1 and photographic images are hereafter called Sample 2. Sample 1 contained 80 video images and Sample 2 consisted of 119 database photographs. The only difference between the two samples was the Media in which the images were captured. When the two samples were compared against each other, Sample 2 was further broken down into a subset of the 80 faces that matched the faces in Sample 1.

Table 6.1 Definitions of two Samples used in the present research study

Sample 1 = 80 Video images
Sample 2 = 119 Photographic images

6.2.2 Data Collection

One volunteer, with previous experience in placing landmarks on print images, placed the 37 landmarks on all 199 images using the measurement programme produced in-house, Facial Identification Centre Version 0.32 © Forensic Medicine and Science Glasgow University. The measurement programme carried out the same linear measurements and calculated the same 59 proportions as was described previously in Chapter 3.

6.2.3 Experimental Design

The following list of questions along with the application of several tests was undertaken with the aim to test the hypothesis; “*Using a comparison of anthropometric facial proportions, it is possible to discriminate between individuals of two samples.”*

1. Can similar faces be separated from dissimilar faces within a single sample using simple dot product comparisons and are the statistics for the two samples stationary?

2. Using a small sample of re-landmarked images, how significant is the error contribution in re-landmarked images and what is the operator spread under ideal conditions?

3. How distinguishable are individual faces in the samples? Is it possible to distinguish true positive faces from true negative faces using simple dot product comparisons where statistics from two samples are known?

4. Is there a subset of vector elements (proportions) that give a better classification than others? Does the removal of the lowest variant elements improve the separation rate between true positive faces and true negative faces?

5. Given the classification accuracy of the system, what subsample requires verification by a human observer for a given degree of matches between a suspect and a database of subjects?

6.2.3.1 Terminology and Equations

The group of 59 proportions was treated as a multi dimensional vector and evaluated as a means of comparing all faces. A vector is a column or row of elements. In this study, the feature vector is the series of 59 proportions derived from chosen linear measurements between facial landmarks. The following tests were conducted on within sample and between sample comparisons: the mean absolute differences between proportions, the Euclidean differences between proportions, the Cosine θ (raw values) distance between feature vectors, Cosine θ (mean subtracted) distance between feature vectors, and the Cosine θ (Z-normalized) [159] distance between feature vectors. These equations are listed below and relative frequency histograms were used to illustrate the distribution of results. For these tests, identical elements must be included in each feature vector. Table 6.2 lists the terminology adopted for same faces and different faces within a comparison.

Table 6.2 Definitions given to correct and incorrect face matches

True Positive (TP): a match which is also a correct match between the video image and photograph of the same subject. **True Negative (TN):** a match which is excluded and which is a correct exclusion because it involves a video image and a photograph of two different subjects. **False Positive (FP):**a match which is an incorrect match between a video image and a photograph of two different subjects. **False Negative (FN):** a match which is excluded but which is an incorrect exclusion because it involves the video image and photograph of the same subject.

The mean absolute difference between proportions

This equation, determines the distance that separates one face from another by taking the absolute value of one face proportion subtracted from the same proportion of a second face. This is carried out for each proportion in the feature vector. The summation of this feature vector is then divided by the total number of elements (59 proportions in this case). A difference of 0 between two faces establishes that those two faces have identical facial proportion vectors. The smaller the difference in facial proportions is indicative of a smaller difference between faces. A disadvantage of using this equation is that the maximum difference between faces is not bonded.

$$\text{Mean abs diff} = \frac{\sum_{n=1}^{n=N} \left| F_1(n) - F_2(n) \right|}{N}$$

The Euclidean distance between proportions

The Euclidean distance is another method of measuring the distance between two points. This is the square root of the sum of the squares of the differences between corresponding elements, in this case proportions, and is often simply noted as $\|F_1 - F_2\|$. A difference of 0 between two faces establishes that those two faces have identical facial proportion vectors. The smaller the difference in facial proportions is indicative of a smaller difference between faces. A disadvantage of using this equation is that the maximum difference between faces is not bonded.

$$\text{Euclidean distance} = \sqrt{\sum_{n=1}^{n=N} \left(F_1(n) - F_2(n)\right)^2}$$

The Cosine *θ* distance

The Cosine θ distance equation is a similarity measurement and is used to measure the angle between two vectors. This equation was completed using the application of three different proportion values; raw, mean subtracted and in order to standardize the measure, Z-normalized. The first equation used the raw values from each facial proportion. The second equation (mean subtracted), was completed in order to create a more standardized vector and was determined by finding the mean for each proportion across the sample and then subtracting the number from the raw values. The third and final Cosine θ distance equation applied a full Z-normalization to each proportion. This was accomplished by dividing the mean subtracted element by the sample standard deviation for the particular proportion.

A difference of 1.0 between two faces establishes that those two faces have identical vectors of facial proportions. An advantage of using this equation is that the range of values is bound from -1.0 to +1.0 and useful comparisons are ranged from zero to one. A difference of zero is indicative of a face that shows no correlation whereas a result of 0.5 is achieved by random chance. Any negative result shows the face comparison has an inverse correlation.

(a)

Cosine θ (raw values)

$$Cos\,\theta = \frac{\sum_{n=1}^{n=N} F_1(n) * F_2(n)}{\|F_1\|\,\|F_2\|}$$

(b)

Cosine θ (mean subtracted)

$$Mean\ subtracted\ element = R_{raw} - R_{pop\,\mu}$$

$$Cos\,\theta = \frac{\sum_{n=1}^{n=N} F_1(n) * F_2(n)}{\|F_1\|\,\|F_2\|}$$

(c)

Cosine θ (Z-normalized)

$$Z - normalized\ element = \frac{R_{raw} - R_{pop\,\mu}}{\sigma_{pop}}$$

$$Cos\,\theta = \frac{\sum_{n=1}^{n=N} F_1(n) * F_2(n)}{\|F_1\|\,\|F_2\|}$$

This equation is also called statistical correlation

6.2.3.2 Experimental Design

> *"Can similar faces be separated from dissimilar faces within a single sample using simple dot product comparisons and are the statistics for the two samples stationary?"*

To answer these questions, the equations in Section 6.2.3.1 were applied so that every face in a single sample was compared to itself and every other face in the same sample. Each sample contained one image of each face and for this reason all that can be determined is the true negativity of two faces, but not if the faces are true positive. The same tests were carried out on Sample 1 and then separately on Sample 2. Testing all combinations of pairs of faces within a single sample is important because it tests faces which fall under the same conditions, including the same camera parameters, allowing the tests to ascertain if there are any differences between faces. This means that the facial proportions will be the only changeable variable between faces as all other variables remain constant; same media, same operator placing landmarks and same facial pose. If a face can not be distinguished from other faces in the same sample using facial proportions, then it will be difficult to use those facial proportions to discriminate a face from a different sample, leading to poor discrimination.

Showing the statistical (mean and standard deviations) similarity in the two samples would reveal that any difference between faces would be attributed to the face and not the media in which the image was obtained. The comparison of distribution means and standard deviations of the two samples was illustrated with a graph exhibiting the normal histogram curves of the two samples superimposed and completed for each equation tested. It is predicted that subsequent to the application of the equations, the data will produce means and standard deviations of two samples which will differ from each other. This may be due to the increased difficulty in landmark placement on faces in Sample 1 due to the image media or because the sizes of the two samples are different.

Comparing faces within a sample is an efficient first step in determining if facial proportions and in turn, feature vectors are an effective method of facial image comparison. It is predicted that because the dot product tests take all proportions into account for the comparison, as a multi dimensional feature vector, there will be enough of a difference in facial proportions to determine similarity or dissimilarity of faces within a single sample.

> *"Using a small sample of re-landmarked images, how significant is the error contribution in re-landmarked images and what is the operator spread under ideal conditions?"*

In order to test the error of landmark placement and witness the effect, if any, on distribution, a small subset of six images from Sample 1 was chosen randomly and each image was re-landmarked six times on the same day by the same operator, creating a subset sample of 36 images. Those same six images were also re-landmarked six times on the same day by multiple operators with no previous landmarking or anthropometry experience. Although, the error of landmarking images by inexperienced operators was discussed in Chapter 5, it remains to be seen how much of an influence the additional non-experienced operators will have on the error when the tests are conducted. It is predicted that the addition of operators with little experience will have a strong effect on the outcome of the error.

"How distinguishable are the faces in the two samples? Is it possible to distinguish true positive faces from true negative faces using dot product comparisons where statistics from two samples are known?"

To answer the above question, a between sample comparison was carried out; every face in Sample 1 was compared to every face in Sample 2 that was found in Sample 1. The remaining 39 faces in Sample 2 that were not in Sample 1 were not included for this test. The two samples contained a finite number of 80 images and it was possible to determine the statistics (mean and standard deviation) of each proportion necessary to carry out the Cosine θ mean subtracted and Z-normalized distance equations. The normal histogram distribution curves of true positive faces and true negative faces were superimposed to determine if it was possible to distinguish between faces in the two groups. The amount of overlap shows the possibility of achieving either a false positive or false negative face match, also known as the rate of misclassification.

The two samples contained the same faces posed in the same position, taken on the same day with facial landmarks placed by the same operator. It was however, the camera parameters (Media) of the two samples which was what differentiated them. Results obtained from comparing faces within-sample can be used to predict the outcome of how separable faces in the two samples are. If the within-sample discriminability is poor then the between-sample discriminability will also be poor.

"Is there a subset of vector elements (proportions) that give better classification than others? Does the removal of the lowest variant elements improve the separation rate between true positive faces and true negative faces?"

Low variant vector elements (proportions) can be indicative of low discriminablity between faces and therefore little information is being added to assist the classification. To determine if removing the lowest variant elements had any effect on the separation rate between true positive and true negative faces, three steps were performed. First, to determine which elements had the lowest variation in each sample, the mean and standard deviation of each element was computed and placed in descending order. Second, using the standard deviations, increasing amounts of the lowest variant elements in amounts of 5%, 10%, 25% and 50% were removed. The third and final step was to apply the Cosine θ (Z-normalized) distance equation to conduct the comparisons between the two samples. The increasing amounts of low variant elements were first removed in a within sample comparison to determine any influence it had on the statistics of each sample. Once this was completed, the same degrees of low variant elements were removed and used in a between sample comparison. Relative frequency histograms were used to illustrate the analyses.

Hypothetically, if elements with low variances were to be removed from a summation vector, it can be predicted that the comparison will bear more weight because elements are only useful if they can be used to successfully distinguish between individuals. However, removing too many parameters can also have a detrimental effect on a comparison, which is why eliminating varying amounts of the lowest variant elements were tested.

> *"Given the classification accuracy of the system, what subsample requires verification by a human observer for a given degree of matches between a suspect and a database of subjects?"*

Given the task of comparing a suspect's image to a large database of identity photographs, the ability to decrease the number of possible face matches could potentially save countless investigation hours. The accuracy of the classification system was shown in the application of the Cosine θ (Z-normalized) distance equation in the between sample comparison and was illustrated by overlapping the normal histogram curves of true positive and true negative faces. Subsequent to using this same equation, a record was kept of the number of faces in Sample 2 that were closer to or equal in distance to the true positive face in Sample 1. Judging the accuracy rate in this way should give similar results to the overlapped normal histogram curves of true positive and true negative face matches.

6.3 Results

The results section is further subdivided into five sections;

1. Comparisons within a single sample
2. Re-landmarked images in Sample 1
3. Between Sample Comparisons
4. Removal of the lowest variant elements
5. Subsample of images requiring human verification.

The differences in faces, after completion of the above mentioned equations, were compiled and illustrated as relative frequency distribution histograms. In cases where faces were compared within a single sample, duplicate comparisons and each identical face comparison was removed from the distributions and the resulting histogram depicted a true negative distribution of faces. In addition, a graph showing the normal histogram curves of Sample 1 superimposed with Sample 2 was provided to illustrate the statistical difference in distributions of the two samples. In the between sample comparisons, four histograms were made to illustrate the results; true positive faces, true negative faces, true positive and true negative faces together and the superimposed normal histogram curves of true positive and true negative faces. This final histogram was included to illustrate the separation rate of true positive and true negative faces and the amount of overlap indicated the possibility of misclassification. Below each histogram is a box containing the samples’ size, and mean and standard deviation of the distribution.

6.3.1 Comparisons within a Single Sample

The following equations were executed within a single sample. Every face within a single sample was compared against itself and every other face in the same sample.

- Mean absolute difference between proportions
- Euclidean distance between proportions
- Cosine θ distance equations
    - Raw values

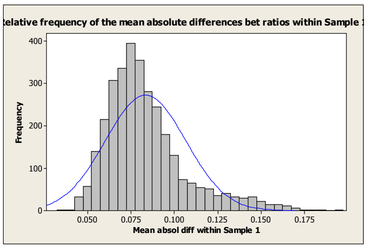
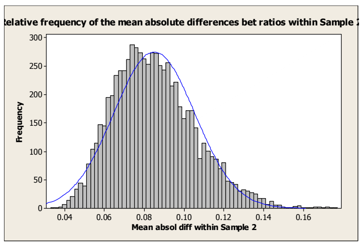
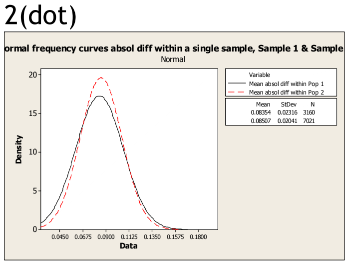
- Mean subtracted values
- Z-normalized values

Equations were first administered on Sample 1 and then separately on Sample 2. These equations were conducted to determine their ability to discriminate between similar faces and dissimilar faces within a single sample using dot product comparisons and if the statistics for the two samples were stationary. As only one image of each face in each sample was available from the research material, it is only possible to determine if two faces were different, rather than if two faces were the same.

6.3.1.1 Mean absolute difference between proportions

Table 6.3 shows the relative frequency histograms, with normal curves, of the mean absolute differences between proportions within Sample 1 (a) and the mean absolute differences between proportions within Sample 2 (b). Section (c) of Table 6.3 shows the normal curves of both samples superimposed. The solid line represents Sample 1, and the dotted line represents Sample 2. This analysis assesses the differences in proportions between faces and a result of “0” between two faces will be indicative of those faces with identical facial proportion vectors.

Table 6.3 Relative frequency histograms illustrating comparisons of the mean absolute differences between proportions in true negative faces within a single sample.

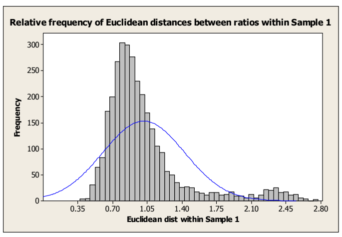
Mean absolute differences between proportions within Sample 1 and within Sample 2		
(a) Sample 1	(b) Sample 2	(c) Normal curve of Sample 1 (solid) superimposed with the normal curve of Sample 2(dot)
		
n= 3160	n= 7021	
Mean 0.0835	Mean 0.0851	
St. Dev 0.0232	St. Dev 0.0204	

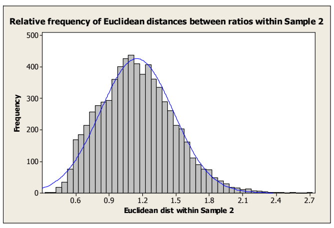
The distributions of the two samples both lean towards the left, or 0.0. Sample 2 shows a much denser histogram caused by the larger sample size. The mean and standard deviations of the distributions in the two samples are very similar which can be seen in the superimposed histogram curves, with Sample 2 showing a higher density most likely due to the larger sample size. Sample 1 has a slightly larger standard deviation and this can be expected because of the slightly more difficulty in placing landmarks on the media that Sample 1 was captured. Due to the difference in mean and standard deviations between the two samples, once this equation is applied to the comparison of Sample 1 against Sample 2 it will not be known if differences between faces are a result of a difference in Media or a difference in faces.

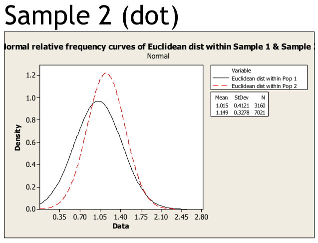
6.3.1.2 Euclidean distances between proportions

Relative frequency histograms with normal curves are shown below in Table 6.4 describing the product of finding the Euclidean distance between feature vectors of true negative faces within a single sample carried out separately for Sample 1 (a) and Sample 2 (b). Section (c) of Table 6.4 shows the normal curves of both samples superimposed. The solid line represents Sample 1 and the dotted line represents Sample 2. This analysis tests the distances between proportions for faces, and therefore a “0” indicates that two faces have identical facial proportion vectors. The distributions from each sample closely resemble the mean absolute difference distributions shown above in which the distributions lean towards the left, closer to 0.0. Again, Sample 2 shows a denser histogram most likely caused by the larger sample size.

Table 6.4 Relative frequency histograms illustrating comparisons of the Euclidean distances between proportions in true negative faces within a single sample.

<table>
<tr><th colspan="3">Euclidean distances between proportions within Sample 1 and within Sample 2</th></tr>
<tr>
<td>(a) Sample 1


<table><tr><td colspan="2">n= 3160</td></tr><tr><td>Mean</td><td>1.0149</td></tr><tr><td>St. Dev</td><td>0.4121</td></tr></table></td>
<td>(b) Sample 2


<table><tr><td colspan="2">n= 7021</td></tr><tr><td>Mean</td><td>1.1493</td></tr><tr><td>St. Dev</td><td>0.3278</td></tr></table></td>
<td>(c) Normal curve of Sample 1 (solid) superimposed with the normal curve of Sample 2 (dot)

</td>
</tr>
</table>

The superimposed normal histogram curves illustrate the difference between the distribution mean and standard deviations of the two samples and this difference was greater than the mean absolute difference equation (Table 6.3), with again Sample 1 having a larger standard deviation than Sample 2. As a result of this difference in mean and standard deviations between the two samples, it will not be certain if the differences between faces result from the Media of the images or an true difference in faces when this equation is applied in the comparison of Sample 1 against Sample 2.

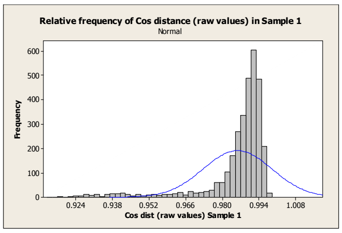
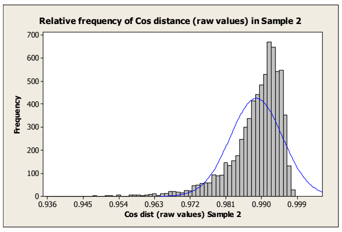
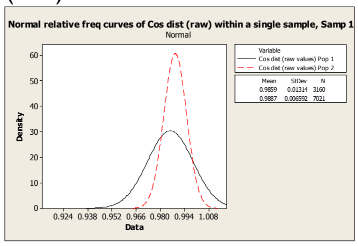
6.3.1.3 Cosine θ distance

Tables 6.5, 6.6, and 6.7 illustrate results obtained from applying the Cosine θ distance equations to faces in Sample 1 (a) and separately to the faces in Sample 2 (b). Section (c) of the tables show the normal curves of both samples superimposed. The solid line represents Sample 1 and the dotted line represents Sample 2. Three different forms of the equation were applied to determine if subtracting the mean (Table 6.6) and dividing by the variance (Table 6.7) made any difference to the distributions and statistics of the two samples. For all three equations, a result of a “1.0” subsequent to a comparison indicates two faces with identical facial proportion vectors.

Table 6.5 shows the relative frequency histograms with normal curves for each sample as a result of using the proportions’ raw value. The distribution from this equation leans heavily toward 1.0. This indicates that there were not many discriminatory factors between faces in the sample to distinguish between true negative faces. In both samples, the means are similar, but the variation within Sample 1 was greater than that of Sample 2 which can again be attributed to the difference in sample size or image media.

When the normal curves of the two histograms were superimposed, this difference in statistics is apparent. Sample 2 has a higher density which can be attributed to the larger size of Sample 2. The superimposed curves also lend evidence to the poor likelihood of this equation as an effective tool for distinguishing similar faces from dissimilar faces because of the strong difference in mean and standard deviations between the two samples. It will not be known if the difference in feature vectors is a result of a difference in Media or a difference in faces.

Table 6.5 Relative frequency histograms illustrating comparisons of Cosine θ distance equations (raw values) in true negative faces within a single sample for Sample 1 (a) and Sample 2 (b). Normal curves are superimposed in (c).

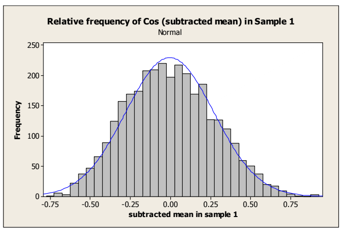
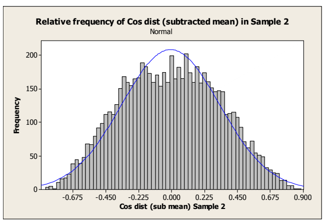
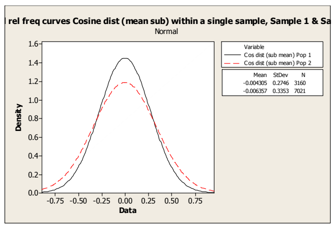
Cosine θ (*raw values*) distance within Sample 1 and within Sample 2		
(a) Sample 1	(b) Sample 2	(c) Normal curve of Sample 1(solid) superimposed with the normal curve of Sample 2 (dot)
		

n=3160	
Mean	0.9859
St. Dev	0.0131

n=7021	
Mean	0.9887
St. Dev	0.0066

Table 6.6 illustrates the histogram distributions and normal curves resulting from applying the Cosine θ distance equation using mean subtracted proportion values. This normalized the distribution and although there is a difference of mean and standard deviation within the two samples which can be seen when the normal curves are superimposed (c), this difference is smaller than the equations previously tested.

Table 6.6 Relative frequency histograms illustrating comparisons of Cosine θ distance equations (mean subtracted) in true negative faces within a single sample for Sample 1 (a) and Sample 2 (b). Normal curves are superimposed in (c).

Cosine θ (*mean subtracted*) distance within Sample 1 and within Sample 2		
(a) Sample 1	(b) Sample 2	(c) Normal curve of Sample 1 (solid) superimposed with the normal curve of Sample 2 (dot)
		

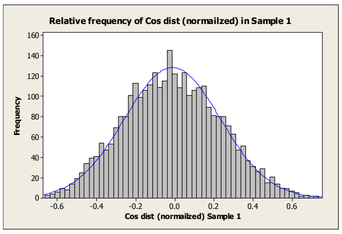
n=3160	
Mean	-0.0043
St. Dev	0.2746

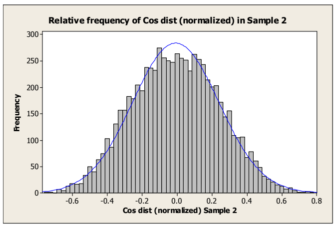
n=7021	
Mean	-0.0064
St. Dev	0.3353

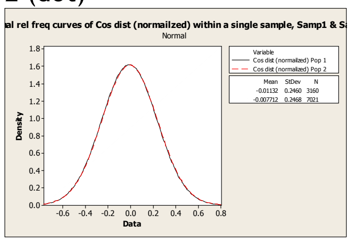
Table 6.7 illustrates the histogram distributions and normal curves obtained from applying the Z-normalized proportion values to the Cosine θ distance equation which resulted in the greatest degree of normal distribution with virtually indistinguishable standard deviations. Note that even though the sample sizes are different, the distributions shown in the superimposed curves remain

superimposed and this is because Z-normalization creates a more standardized vector. Thus far, Cosine θ distance equation using Z-normalized values has been the best representation of establishing the stationary statistics of the two samples. The lack of overlap observed in the normal curves assumes that any difference in faces is a result of the true difference in faces rather than a difference in media.

Table 6.7 Relative frequency histograms illustrating comparisons of Cosine *θ* distance equations (Z-normalized) in true negative faces within a single sample for Sample 1 (a) and Sample 2 (b). Normal curves are superimposed in (c).

<table>
<tr><th colspan="3">Cosine θ distance (Z-normalized) within Sample 1 and within Sample 2</th></tr>
<tr>
<td>(a) Sample 1


<table>
<tr><td colspan="2">n=3160</td></tr>
<tr><td>Mean</td><td>-0.0113</td></tr>
<tr><td>St. Dev</td><td>0.2460</td></tr>
</table>
</td>
<td>(b) Sample 2


<table>
<tr><td colspan="2">n=7021</td></tr>
<tr><td>Mean</td><td>-0.0077</td></tr>
<tr><td>St. Dev</td><td>0.2468</td></tr>
</table>
</td>
<td>(c) Normal curve of Sample 1 (solid) superimposed with the normal curve of Sample 2 (dot)


</td>
</tr>
</table>

6.3.2 Re-landmarked Images in Sample 1

In order to test the error of landmark placement on images and the effect on distribution under ideal conditions, six images in Sample 1 were landmarked six times each on the same day and superimposed normal curves of true positive and true negative faces were created from the data. The six images were re-landmarked by a single experienced operator and four inexperienced operators. These were the same re-landmarked images examined and discussed previously in Chapter 5. Landmark placement by multiple operators is to be expected if this technique were to be used in forensic science laboratories. Comparisons of re-landmarked images were analyzed first from the experienced operator and second, with all operators. The sample size of re-landmarked images by the experienced operator was 36 and the sample size of re-landmarked images by all operators was 180. The following equations were applied to the comparison of re-landmarked images of a single experienced operator:

- Cosine θ distance equations
  - o Raw values
  - o Mean subtracted values
  - o Z-normalized values derived from statistics of sample of re-landmarked images
  - o Z-normalized values derived from statistics of Sample 1

The following equations were applied to the comparison of re-landmarked images of all operators:

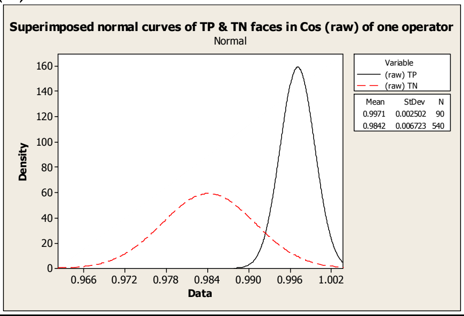
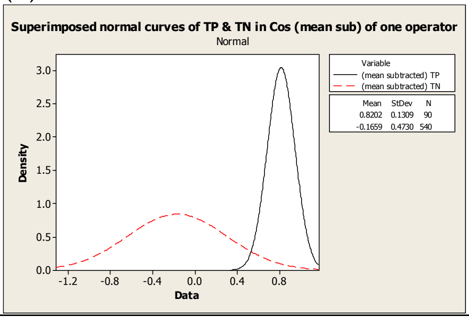
- Cosine θ distance equations
  - o Z-normalized values using statistics derived from the sample of re-landmarked images
  - o Z-normalized values using statistics derived from Sample 1

Only the Cosine θ distance equations were conducted on this sample of images as a result of the stationary statistics produced by the equation in Section 6.3.1.3. Only the Cosine θ (Z-normalized) distance equation was applied in the comparison of re-landmarked images by all operators because of the strong separation rate achieved from the single trained operator. Each image was compared against itself and every other image and duplicate comparisons along with identical face comparisons were removed from the distribution. Results are illustrated as superimposed normal curves separating true positive and true negative faces. True positive faces are represented by the solid curve and true negative faces by the dotted curve. The amount that the curves overlap is equal to the number of true positive and true negative faces that could be confused in the comparison. The smaller the area, the smaller the chances of obtaining a false positive or false negative face match.

6.3.2.1 Cosine θ distance

Table 6.8 illustrates the superimposed normal histogram curves of true positive and true negative faces from the Cosine θ distance equations using raw values (a) and mean subtracted values (b). The amount of overlap is smaller when the proportion values used in the Cosine θ equation was computed by subtracting the mean of the proportion. These amounts of overlap of true positive and true negative faces appear to be relatively small indicating that that these equations would be very useful to separate the true positive from the true negative faces.

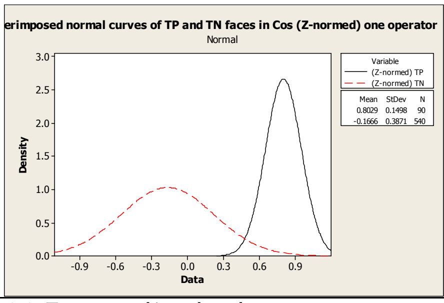
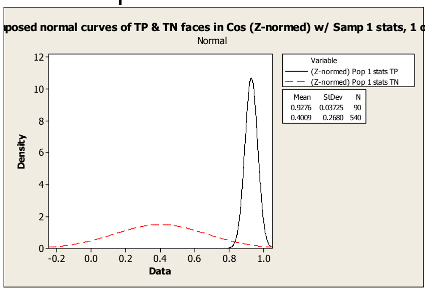
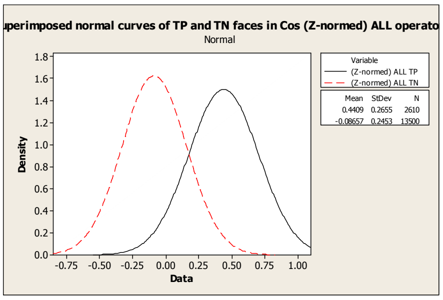
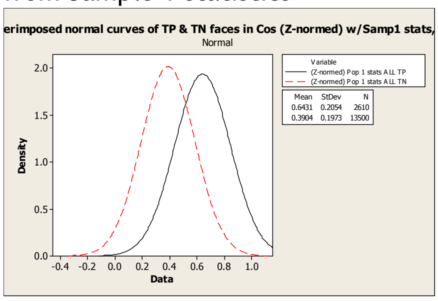
Table 6.8 Superimposed normal curve histograms illustrating true positive and true negative face comparisons of the Cosine *θ* (raw and mean subtracted) distance equations in six re-landmarked images from Sample 1 using one experienced operator.

Superimposed normal curves of true positive and true negative faces using Cosine *θ* distance equations for six re-landmarked video images in Sample 1 using one experienced operator	
(a) Raw values	(b) Mean subtracted values
	

However, the smallest amount of overlap was seen when Z-normalization was applied, which is illustrated in Table 6.9. Results obtained from the experienced operator are shown on top row of the table and results from all operators are located on the bottom row of the table. Z-normalization was derived from both the subsamples statistics and, to witness the effect, from the statistics of the entire Sample 1. The Cosine *θ* distance equation using Z-normalized values derived from the subset sample's own statistics is in sections (a, c) and the Cosine *θ* distance using Z-normalized values derived from the statistics from the entire Sample 1 is in sections (b, d). The experienced operator and all operators were placed in the same table in order to directly compare the effect that inexperienced operators had on the separation rate of true positive and true negative faces.

When compared to the experienced operator, the effect of inexperience operators can clearly be seen in the separation rates of true positive and true negative faces in both cases of Z-normalization. Results from the experienced operator, using the re-landmarked subset samples statistics, achieved a smaller separation rate then when deriving Z-normalization from the statistics of the entire Sample 1. This indicates that for this case a larger sample size in which to derive statistics could be influential. In contrast, results from all operators achieved a noticeably smaller separation rate using the re-landmarked subset samples statistics then when deriving Z-normalization from the statistics of the entire Sample 1. Although this contrast occurred, the possible reasoning is the same. In this case, to derive the statistics for Z-normalization, the larger sample size was the re-landmarked images by all operators. In addition, even though the re-landmarked image sample was larger, the number of different faces in the sample was smaller.

Table 6.9 Superimposed normal curve histograms illustrating true positive and true negative face comparisons of the Cosine θ (Z-normalized) distance equations in six re-landmarked images from Sample 1 using one experienced operator and all operators.

<table>
<tr><th colspan="3">Cosine θ (Z-normalized) distance equations for six re-landmarked video images in Sample 1</th></tr>
<tr><td>One operator (experienced)</td><td>(a) Z-normalized values

</td><td>(b) Z-normalized values derived from Sample 1 statistics

</td></tr>
<tr><td>Multiple operators (experienced and inexperienced)</td><td>(c) Z-normalized values

</td><td>(d) Z-normalized values derived from Sample 1 statistics

</td></tr>
</table>

The most important point witnessed in Table 6.9 from this comparison of the experienced operator and multiple operators is not necessarily what size of sample to derive statistics for Z-normalization, but to observe the strong effect that inexperienced operators have in separating true positive and true negative face cases. As multiple experienced operators were not tested, it cannot be stated that this difference in separation rates between the experienced operator and all operators was due to the experience of the operators or instead the effect of that will naturally occur with multiple operators. It is expected that multiple experienced operators would also detrimentally effect separation rates compared to a single experienced operator, although it is assumed that this effect would be considerably less than with multiple inexperienced operators and is something to be tested in future inter-operator studies.

6.3.3 Between Sample Comparisons

The following equations were applied to the between sample comparisons of Sample 1 against Sample 2.

- Mean absolute difference between proportions
  - Raw values
  - Z-normalized values
- Euclidean distance between proportions using Z-normalized values
- Cosine θ distance equations
  - Z-normalized values
  - Log of the raw proportional values
    - Raw values
    - Mean subtracted values
    - Z-normalized values
  - Z-normalized values derived from a single sample’s statistics
    - statistics derived from Sample 1
    - statistics derived from Sample 2
    - statistics derived from summation of Sample 1 and Sample 2

Only faces from Sample 2 that had a matching face in Sample 1 were included in the comparisons and the 39 extra faces in Sample 2 were excluded. Every face in Sample 1 was compared against every one of those matching faces in Sample 2 and the true positive match excluded from the distribution. The distributions were separated into three histograms showing relative frequency; true positive face comparisons, true negative face comparisons, and true positive and true negative faces together. The overlapped normal histogram curves from true positive faces and true negative faces were also shown in order to determine the separation rate between true positive and true negative face comparisons. The true positive normal curve is represented by the dotted line and the true negative normal curve is represented by the solid line. The smaller the amount of overlap illustrates a larger degree of separation between the two categories which in turn demonstrates the potential of the equation to locate a correct match between samples. This degree of overlap can also be referred to as the quantity of false positive and false negative face comparisons, or the degree of misclassification.

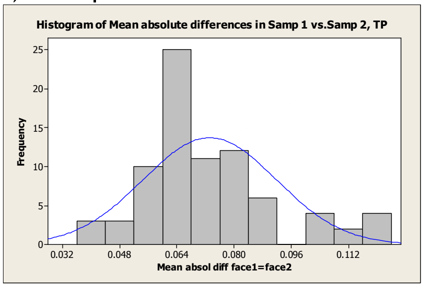
6.3.3.1 Mean absolute difference between proportions

To begin, Sample 1 was compared to the matched faces from Sample 2 by applying the mean absolute difference between proportions equation. This was completed first using the raw proportional values (Table 6.10) and then using Z-normalized values (Table 6.11) derived from each sample's own statistics. The Z-normalized values were tested using this equation because when the Z-normalized values were applied to the Cosine θ distance equation in re-landmarked images of Section 6.3.2.1, the greatest separation between the group of true positive faces and true negative faces was produced. It was not known if using Z-normalized values would have the same effect on this equation, however, it was thought separation rates may be improved and so was tested.

Table 6.10 Relative frequency histograms illustrating comparisons of the mean absolute difference between proportions of Sample 1 vs. the matched faces in Sample 2 showing true positive cases (a), true negative cases (b), both cases together (c), and both cases superimposed (d).

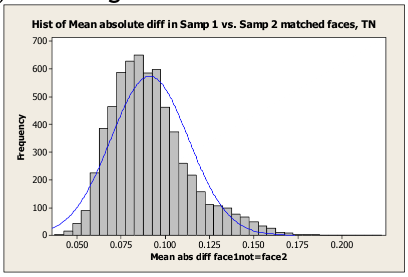
Mean absolute difference between proportions in Sample 1 vs. Sample 2 matched faces using *raw values*

(a) True positive faces



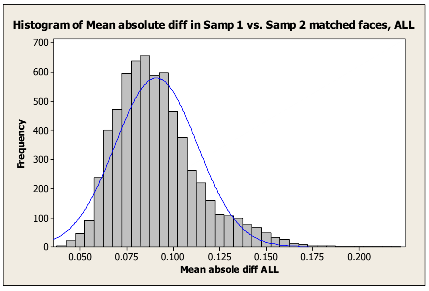
n= 80	
Mean	0.0732
St. Dev	0.0186

b) True negative faces



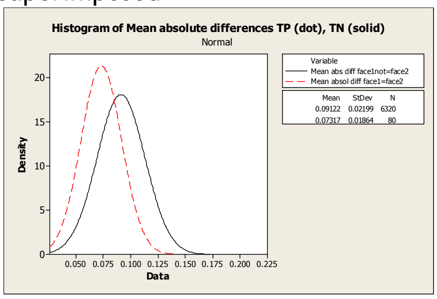
n= 6320	
Mean	0.0912
St. Dev	0.0220

(c)True positive faces + True negative faces



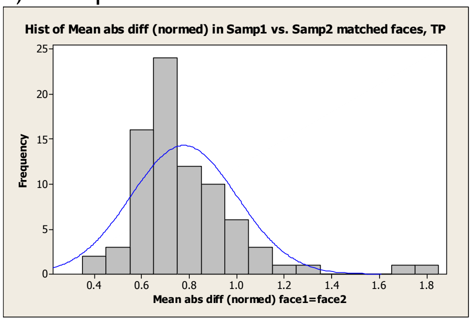
n= 6400	
Mean	0.0910
St. Dev	0.0220

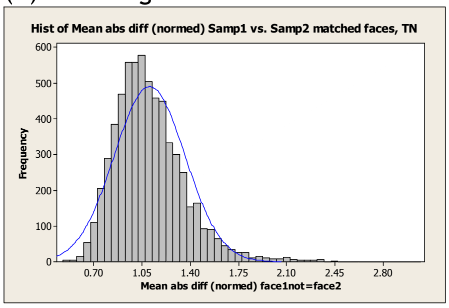
(d) Normal curves of true positive (dot) and true negative (solid) faces superimposed

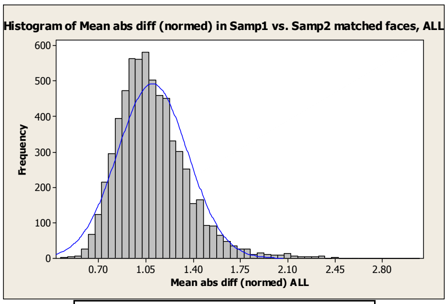


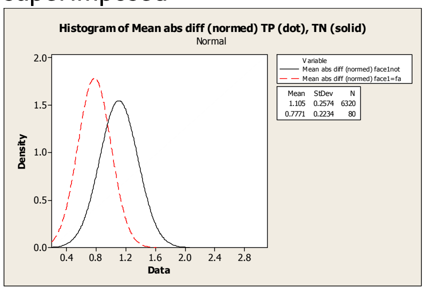
The histogram distributions showing raw values (Table 6.10) and Z-normalized values (Table 6.11) looked very similar to each other. Both groups showed distributions that leaned towards the left or 0.0 mark. A slightly smaller overlap of true positive faces and true negative faces occurred when using the Z-normalized values, indicating that the Z-normalized values showed a greater promise for making a facial comparison when this equation was applied to the feature vectors. However, the amount of overlap is still more than would be desired of an identification method.

Table 6.11 Relative frequency histograms illustrating comparisons of the mean absolute differences between proportions of Sample 1 vs. the matched faces in Sample 2 using Z-normalized values showing true positive cases (a), true negative cases (b), both cases together (c), and both cases superimposed (d).

<table>
<tr><th colspan="2">Mean absolute differences between proportions in Sample 1 vs. Sample 2 matched faces using Z-normalized values</th></tr>
<tr><td>(a) True positive faces


<table><tr><td colspan="2">n= 80</td></tr><tr><td>Mean</td><td>0.7771</td></tr><tr><td>St. Dev</td><td>0.2234</td></tr></table></td>
<td>(b) True negative faces


<table><tr><td colspan="2">n= 6320</td></tr><tr><td>Mean</td><td>1.1053</td></tr><tr><td>St. Dev</td><td>0.2574</td></tr></table></td></tr>
<tr><td>(c) True positive faces + true negative faces


<table><tr><td colspan="2">n= 6400</td></tr><tr><td>Mean</td><td>1.1012</td></tr><tr><td>St. Dev</td><td>0.2596</td></tr></table></td>
<td>(d) Normal curves of true positive (dot) and true negative (solid) faces superimposed

</td></tr>
</table>

6.3.3.2 Euclidean distance between proportions

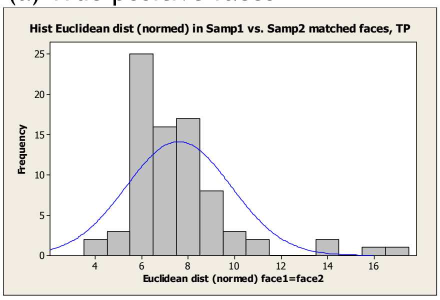
The Euclidean distance between proportions was found between each face in Sample 1 vs. each matched face in Sample 2. The relative frequency histograms in Table 6.12 measure the distribution of the Euclidean distance using Z-normalized values derived from each samples own statistics. The Z-normalized values were used in this comparison because they showed a slightly smaller overlap as opposed to the raw values when using the mean absolute difference between proportions (Section 6.3.3.1). It was not known if using Z-normalized values would have the same effect on this equation as for the mean absolute differences between proportions equation, however, as separation rates were slightly greater, it was thought the same effect may be witnessed here and so was the sole equation tested.

The histogram distributions again leaned towards the left. The amount of overlap of true positive faces to true negative faces appeared to be similar to that of the mean absolute differences between proportions test when using Z-normalized values and would indicate that this equation did not offer a more advantageous value when distinguishing between individuals with the same or different faces from two different samples.

Table 6.12 Relative frequency histograms illustrating comparisons using the Euclidean distance between proportions of Sample 1 vs. the matched faces in Sample 2 using Z-normalized values showing true positive cases (a), true negative cases (b), both cases together (c), and both cases superimposed (d).

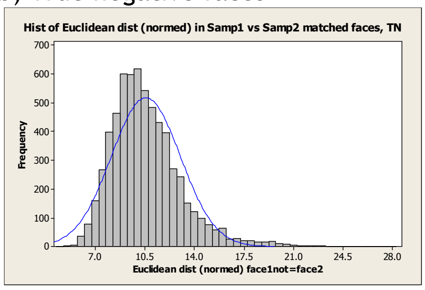
Euclidean distance between proportions in Sample 1 vs. Sample 2 matched faces using *Z-normalized values*

(a) True positive faces



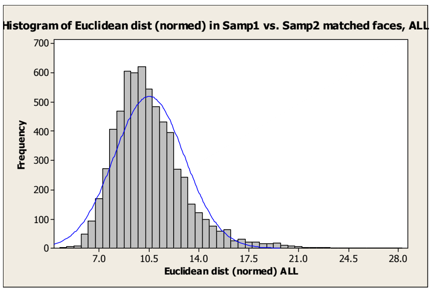
n= 80	
Mean	7.594
St. Dev	2.258

(b) True negative faces



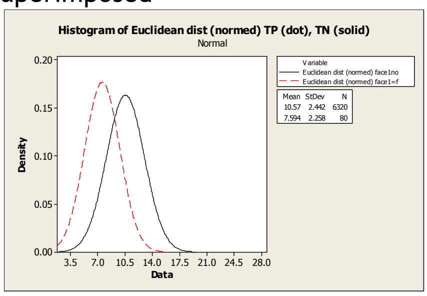
n= 6320	
Mean	10.572
St. Dev	2.442

(c) True positive faces + true negative faces



n= 6400	
Mean	10.534
St. Dev	2.462

(d) Normal curves of true positive (dot) and true negative (solid) faces superimposed



6.3.3.3 Cosine θ distance

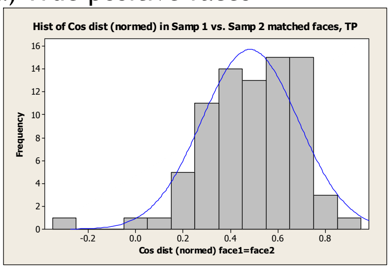
The histograms in Table 6.13 were derived from the Cosine θ (Z-normalized) distance equation and used to determine if it was possible to distinguish same faces from different faces in two samples where statistics from those two samples are known.

The data used was Z-normalized using each sample's own statistics. Table 6.13 shows the distributions of true positive faces (a), true negative faces (b), true positive and true negative faces together (c) and the normal curves of true positive and true negatives overlapped (d). Of the superimposed curves, the dotted line represents true positive faces and the solid line represents true negative faces. The histogram of true negative faces (b) illustrates a distribution more normal to that seen in the histogram of true positive faces (a). This could be a result of the larger sample size of true negative faces or because the frequency of true positive faces should have the majority of their values situated towards the 1.0 mark as the closer a result to 1.0 equals a closer representation between faces. The Cosine θ distance (Z-normalized) equation has shown the least amount of overlap between the true positive and true negative faces. This represents a smaller rate of image misclassification (approximately 25%) and indicates this equation is the best predictor of face discrimination that has been tested during this research thus far.

Table 6.13 Relative frequency histograms illustrating comparisons using the Cosine θ distance (raw) between proportions of Sample 1 vs. the matched faces in Sample 2 using Z-normalized values showing true positive cases (a), true negative cases (b), both cases together (c), and both cases superimposed (d).

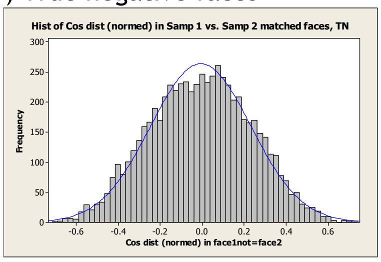
Cosine θ distance between proportions in Sample 1 vs. Sample 2 matched faces using *Z-normalized* values

(a) True positive faces



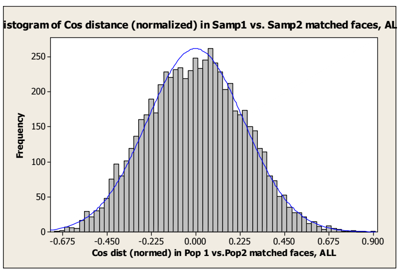
N=80	
Mean	0.4822
St. Dev	0.2035

(b) True negative faces



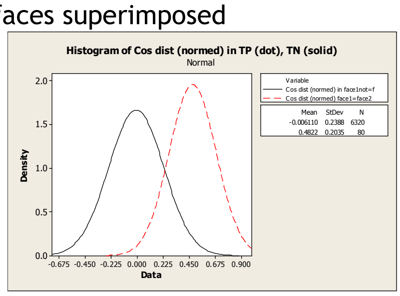
n=6320	
Mean	-0.0061
St. Dev	0.2389

(c) True positive faces + true negative faces



n=6400	
Mean	-0.00001
St. Dev	0.2445

(d) Normal curves of true positive (dot) and true negative (solid) faces superimposed



The tests conducted in the remainder of this section, were carried out in an attempt to further improve upon results achieved from the Cosine θ (Z-normalized) distance equation. All tests were conducted using the Cosine θ distance equations and the only differences between the tests were the values used in the equation. The log of the proportional values was first applied to the raw, mean subtracted and Z-normalized equations and then Z-normalized values were derived from using a single sample's statistics; Sample 1, Sample 2 and statistics derived from a summation of both samples.

Cosine θ distance using the log of the proportion values

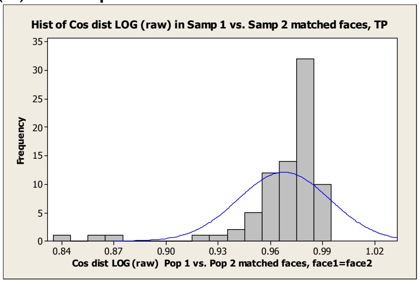
The log of the proportion values was taken and inputted into the three Cosine θ distance equations; raw, mean subtracted, and Z-normalized. Distribution results are illustrated in the histograms below. Raw value distributions are located in Table 6.14, mean subtracted distributions are located in Table 6.15, and Z-normalized distributions are located in Table 6.16. Each table is subdivided into true positive faces (a), true negative faces (b), true positive and true negative faces together (c) and the superimposed normal curves of true positive faces and true negative faces (d). The dotted line of the superimposed normal curves represents true positive faces and the solid line represents true negative faces.

The log of the raw values showed the largest amount of overlap between true positive and true negative faces indicating a minimal discrimination factor between the true positive and true negative face groups. The distributions using the mean subtracted values showed considerably less overlap but the least amount of normal curve overlap of true positive and true negative faces was again found in the histogram for the Cosine θ distance equation using Z-normalized values. Results mirrored that of Table 6.13 (Cosine θ (Z-normalized) distance), indicating there is no advantage to using the log of the proportion values.

Table 6.14 Relative frequency histograms illustrating comparisons using the log of proportional values to derive the Cosine θ (raw) distance of Sample 1 vs. the matched faces in Sample 2 showing true positive cases (a), true negative cases (b), both cases together (c), and both cases superimposed (d).

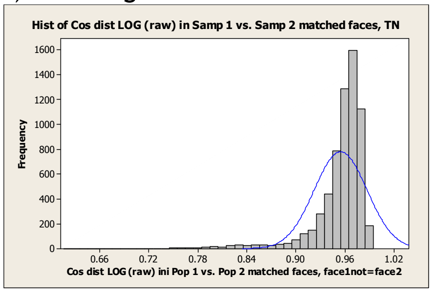
Cosine θ distance between Sample 1 vs. Sample 2 matched faces *raw* values derived from the log of proportional values

(a) True positive faces



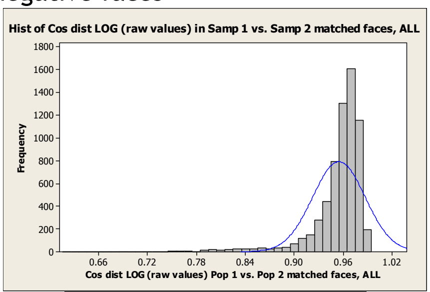
n= 80	
Mean	0.9677
St. Dev	0.0264

(a) True negative faces



n= 6320	
Mean	0.9548
St. Dev	0.0324

(a) True positive faces + true negative faces



n= 6400	
Mean	0.9550
St. Dev	0.0324

(d) Normal curves of true positive and true negative faces superimposed

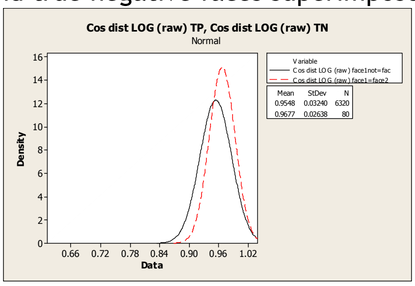
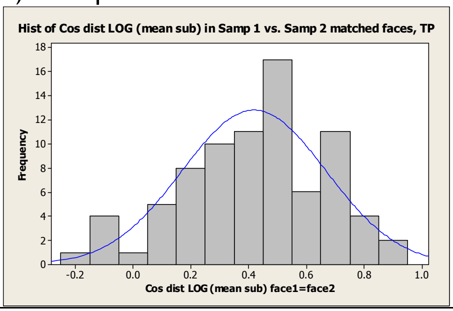


Table 6.15 Relative frequency histograms illustrating comparisons using the log of proportional values to derive the Cosine θ (mean subtracted) distance of Sample 1 vs. the matched faces in Sample 2 showing true positive cases (a), true negative cases (b), both cases together (c), and both cases superimposed (d).

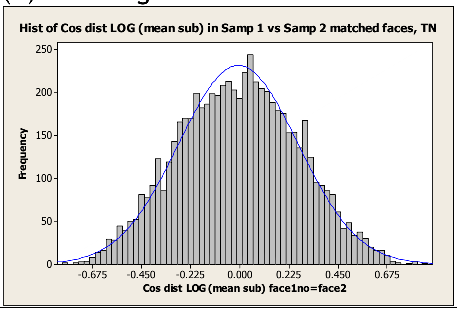
Cosine θ distance between Sample 1 vs. Sample 2 matched faces using *mean subtracted* values derived from the log of proportional values

(a) True positive faces



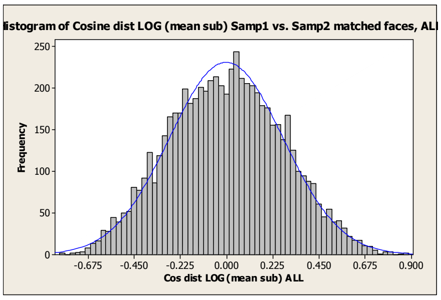
n= 80	
Mean	0.4185
St. Dev	0.2492

(b) True negative faces



n= 6320	
Mean	-0.0063
St. Dev	0.2734

(b) True positive faces + true negative faces



n= 6400	
Mean	-0.0010
St. Dev	0.2771

(d) Normal curves of true positive (dot) and true negative (solid) faces superimposed

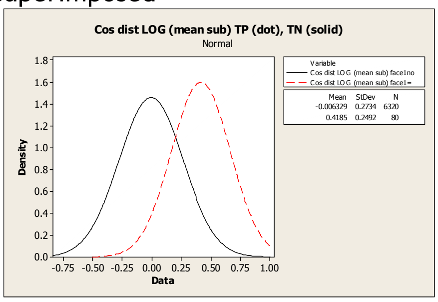
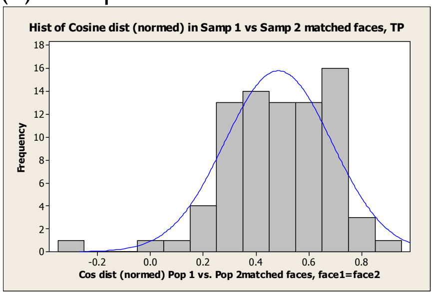


Table 6.16 Relative frequency histograms illustrating comparisons using the log of proportional values to derive the Cosine θ (Z-normalized) distance of Sample 1 vs. the matched faces in Sample 2 showing true positive cases (a), true negative cases (b), both cases together (c), and both cases superimposed (d).

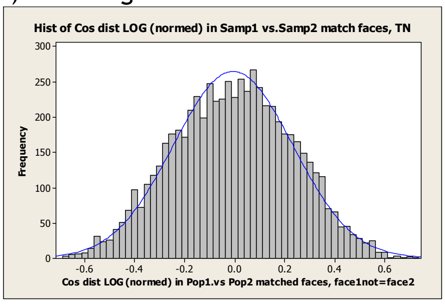
Cosine θ distance between Sample 1 vs. Sample 2 matched faces using *Z-normalized* values derived from the log of proportional values

(a) True positive faces



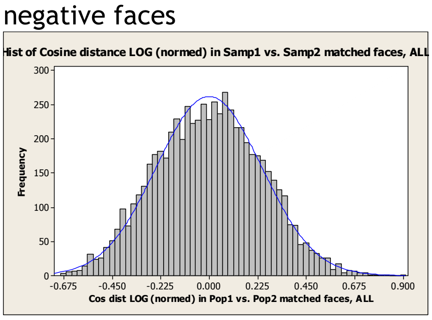
n= 80	
Mean	0.4840
St. Dev	0.2024

(b) True negative faces



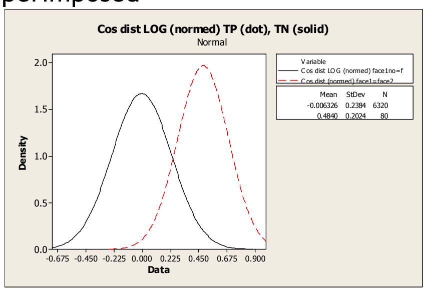
n= 6320	
Mean	-0.0063
St. Dev	0.2384

(c) True positive faces + true negative faces



n= 6400	
Mean	-0.0002
St. Dev	0.2442

(d) Normal curves of true positive (dot) and true negative (solid) faces superimposed



Cosine θ (Z-normalized) distance using a single samples' statistics

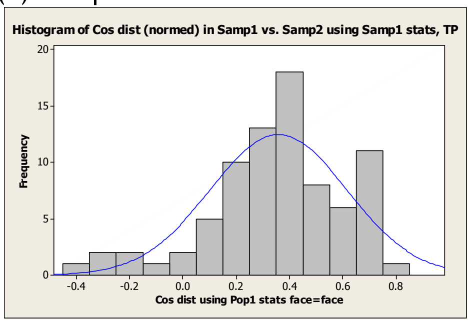
It may often be the case where the statistics of only one sample or the other are known which therefore must be applied to both samples and these cases should be assessed. Results from this analysis are illustrated as relative frequency histograms. Table 6.17 illustrates the Cosine θ (Z-normalized) distance equation where Z-normalization is derived from statistics of Sample 1. Table 6.18 illustrates results when Z-normalization was derived using statistics from Sample 2 and Table 6.19 illustrated results when Z-normalization was derived from statistics of a summation of Samples 1 and 2. Following the format of previous tables, histograms are distributed as true positive faces (a), true negative faces (b), true positive and true negative faces together (c) and the superimposed normal curves of true positive faces and true negative faces (d). The dotted line of the superimposed normal curves represents true positive faces and the solid line represents true negative faces.

Frequency distributions as a result of Z-normalization derived from statistics from Sample 1 were similar to that when using statistics from Sample 2. There was a slightly smaller overlap of normal curves in true positive faces and true negative face using statistics derived from Sample 2, but not enough to be significant and was not improved from the situation where each samples own statistics were used to derive Z-normalization (Table 6.13).

Table 6.17 Relative frequency histograms illustrating comparisons using the Cosine θ (Z-normalized) distance of Sample 1 vs. the matched faces in Sample 2 showing true positive cases (a), true negative cases (b), both cases together (c), and both cases superimposed (d). Statistics to determine Z-normalized values were derived from Sample 1.

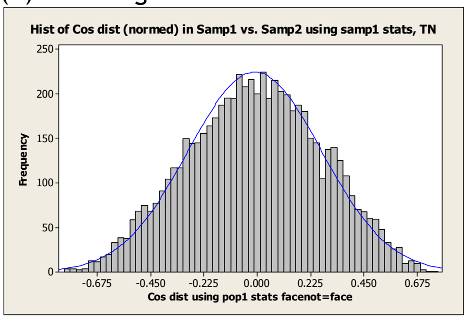
Cosine θ distance (Z-normalized) for Sample 1 vs. Sample 2 using statistics derived from Sample 1

(a)True positive faces



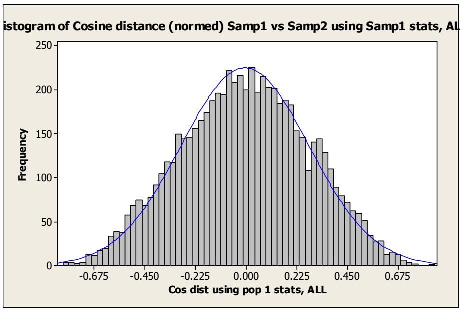
n= 80	
Mean	0.3584
St. Dev	0.2565

(b) True negative faces



n= 6320	
Mean	-0.0097
St. Dev	0.2816

(c) True positive faces + true negative faces



n= 6400	
Mean	-0.0051
St. Dev	0.2842

(d) Normal curves of true positive (dot) and true negative (solid) faces superimposed

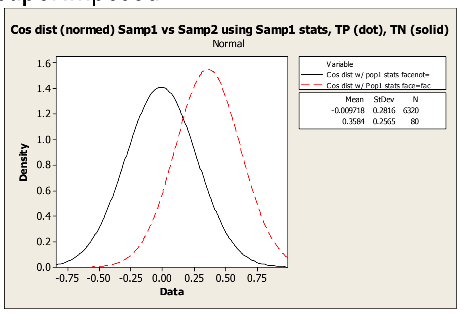
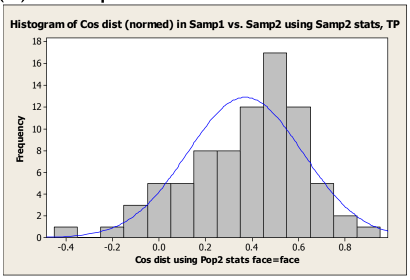


Table 6.18 Relative frequency histograms illustrating comparisons using the Cosine θ (Z-normalized) distance of Sample 1 vs. the matched faces in Sample 2 showing true positive cases (a), true negative cases (b), both cases together (c), and both cases superimposed (d). Statistics to determine Z-normalized values were derived from Sample 2.

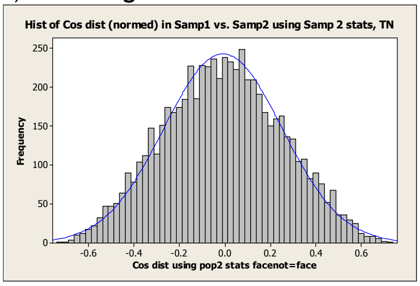
Cosine θ distance (Z-normalized) for Sample 1 vs. Sample 2 using statistics derived from Sample 2

(a) True positive faces



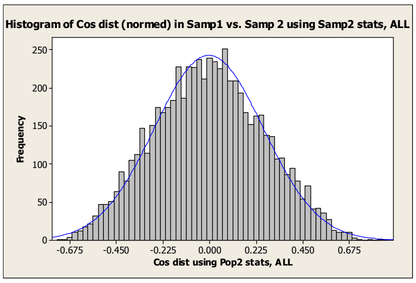
n= 80	
Mean	0.3706
St. Dev	0.2477

(b) True negative faces



n= 6320	
Mean	-0.0083
St. Dev	0.2599

(c) True positive faces + true negative faces



n= 6400	
Mean	-0.0036
St. Dev	0.2631

(d) Normal curves of true positive and true negative faces superimposed

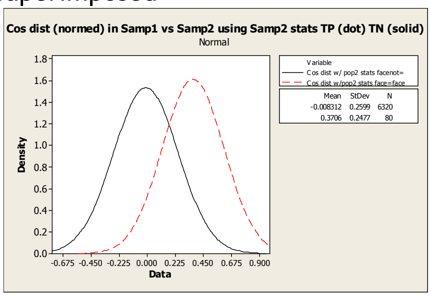


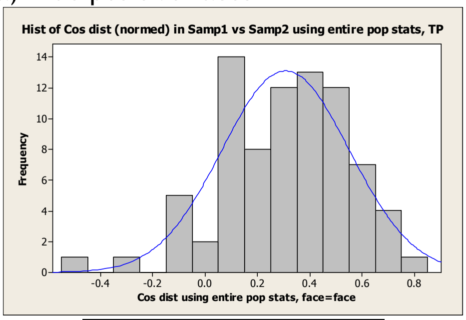
Table 6.19 illustrates the frequency distribution data when applying the Cosine θ distance (Z-normalized) equation to the between sample comparison, using statistics derived from a summation of the two samples in order to calculate the Z-normalized values. The relative frequency histogram distributions for true positive faces (a), true negative faces (b), and true positive with true negative faces (c) appeared to be similar to the histogram distributions above using either Sample 1 (Table 6.17) or Sample 2 (Table 6.18) statistics to calculate Z-normalized values. The superimposed normal histogram curves of true positive faces and true negative faces histogram for this comparison appeared to have a

smaller overlap than when using a single samples' statistics, however, the overlap was not an improvement from deriving Z-normalized values from each samples own statistics (Table 6.13).

Table 6.19 Relative frequency histograms illustrating comparisons using the Cosine *θ* (Z-normalized) distance of Sample 1 vs. the matched faces in Sample 2 showing true positive cases (a), true negative cases (b), both cases together (c), and both cases superimposed (d).Statistics to determine Z-normalized values were derived from a summation of the two samples.

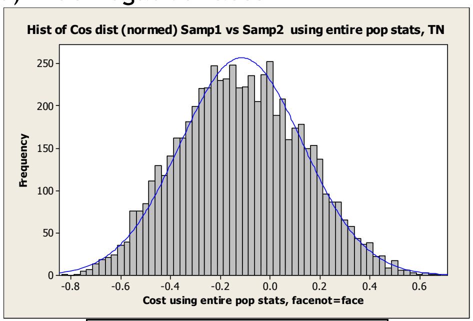
Cosine *θ* distance (Z-normalized) for Sample 1 vs. Sample 2 using statistics derived from a summation of both samples

(a) True positive faces



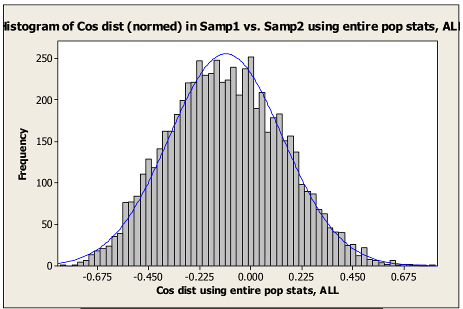
n= 80	
Mean	0.3069
St. Dev	0.2436

(b) True negative faces



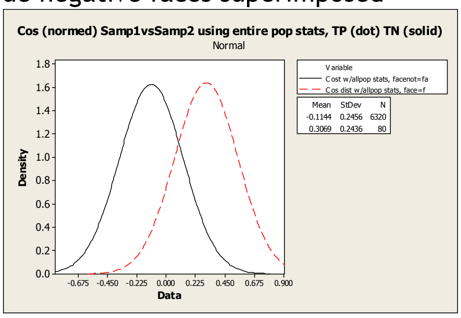
n= 6320	
Mean	-0.1144
St. Dev	0.2456

(c) True positive faces + true negative faces



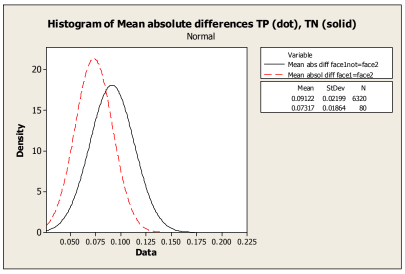
n= 6400	
Mean	-0.1091
St. Dev	0.2500

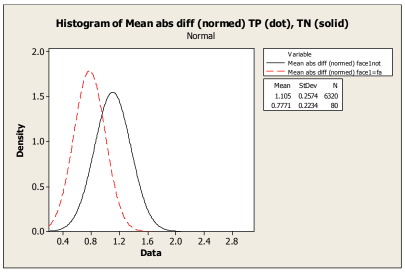
(d) Normal curves of true positive and true negative faces superimposed

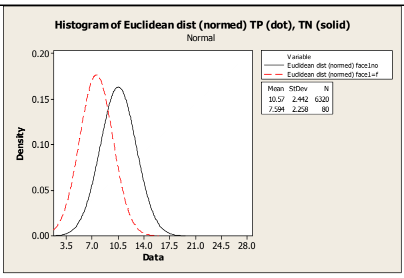


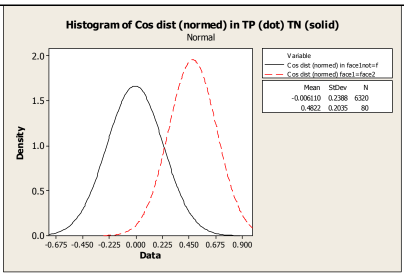
To briefly summarize Section 6.3.3, the between sample comparisons, the greatest separation of true positive and true negative faces was achieved with the Cosine θ (Z-normalized) equation and a larger separation rate was not fulfilled using either the log of proportions or when applying several different Z-normalization values. Table 6.20 was created to illustrate this summarization and used the same histograms showing the degree of overlap in true positive and true negative faces as in (d) of each table in Section 6.3.3. The next stage in the investigation of feature vectors to discriminate between faces of two samples was to remove the lowest variant elements from the feature vector.

Table 6.20 Summary of the conditions imposed and results achieved in the comparison of faces in Sample 1 and Sample 2. Results are illustrated by the superimposed normal histogram curves showing the amount of overlap in true positive (dot) and true negative (solid) faces indicating the chance of misclassification.

<table>
<tr><th colspan="2">Between Sample Comparisons</th></tr>
<tr><th>Condition</th><th>Result</th></tr>
<tr><td colspan="2">Mean absolute difference between proportions</td></tr>
<tr><td>• Raw values</td><td>


</td></tr>
<tr><td>• Z-normalized values</td><td>


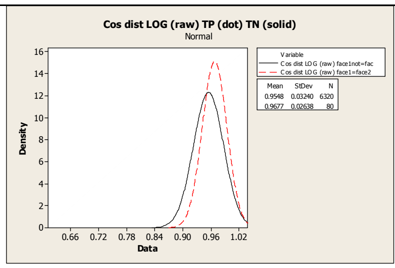
</td></tr>
<tr><td colspan="2">Euclidean distance between proportions</td></tr>
<tr><td>• Z-normalized values</td><td>


</td></tr>
<tr><td colspan="2">Cosine θ distance equations</td></tr>
<tr><td>• Z-normalized values</td><td>


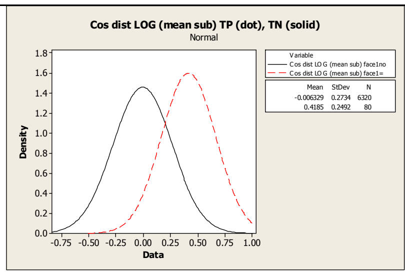
</td></tr>
</table>

Cosine θ distance equations using the log of the raw proportional values

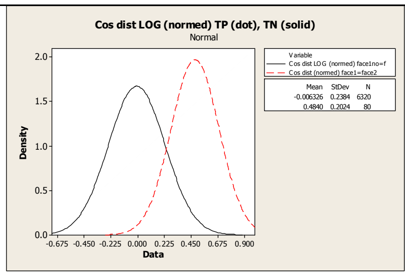
- Raw values



- Mean subtracted values

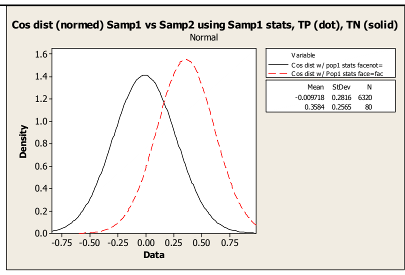


- Z-normalized values

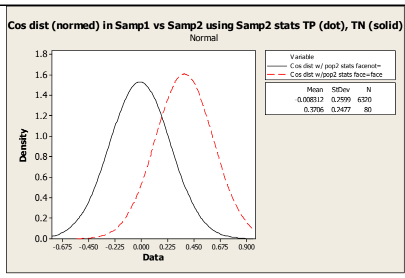


Cosine θ distance equations Z-normalized with values derived from a single sample's statistics

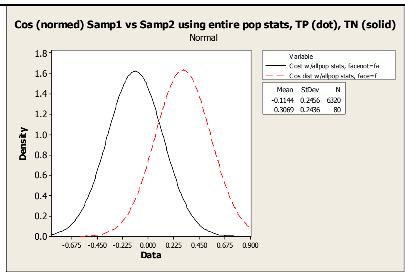
- statistics derived from Sample 1



- statistics derived from Sample 2



- statistics derived from a summation of Sample 1 and Sample 2



6.3.4 *Removal of Lowest Variant Elements*

The greatest separation between true positive and true negative faces was achieved with the application of the Cosine θ (Z-normalized) equation using each samples own statistics. To determine if this separation rate could be further improved, the elements (proportions in this study), which had a low variance, and therefore may be of little use or interfere with the comparison were removed from the feature vector. Feature vector elements which have a low standard deviation between individuals are of little use when trying to discriminate between the individuals. Forensically, an element that proves to be similar amongst the majority of the sample is of little benefit as a discriminating factor.

To determine which elements would be removed from the feature vector, the mean and standard variation for each proportion was found in each sample. For this process the sample sizes were different in each sample; Sample 1 contained 80 images and Sample 2 consisted of 119 images. Determination of the lowest variant parameters can be summarized in 3 steps:

1. Find the mean and standard deviation for each element (proportion) in Sample 1 and in Sample 2.

2. Rank the element standard deviations in descending order.

3. Remove the chosen elements in increasing amounts between 5-50%.

Increasing amounts of lowest variant elements were first removed in a within sample comparison to determine the effect it had on the statistics of the samples. Once this was completed, the same degrees of low variant elements were removed and applied in a between sample comparison. With the exception of one element, the 30 lowest variant elements were found to be the same in both Sample 1 and Sample 2 and they were removed in the order as predetermined by Sample 1. Relative frequency histograms were used to illustrate the distributions of data in the same format as used previously.

The analysis is first illustrated showing the relative frequency distributions of the mean (Table 6.21) and standard deviation (Table 6.22) of the elements in each sample and are shown along with their respective superimposed normal curve distributions. The superimposed normal curves in (c) of Table 6.21 show a nearly identical distribution mean for both samples. The standard deviation normal curve overlap is slightly greater in Table 6.22 which shows a larger variation between samples. This slight variation between samples could be a result from the error that occurs as a result of landmark placement on different media; the standard deviation of elements in Sample 2 was smaller than Sample 1 and the resolution of images in Sample 2 was clearer to images in Sample 1.

Table 6.21 Relative frequency histograms of the element mean of Sample 1 and the element mean of Sample 2

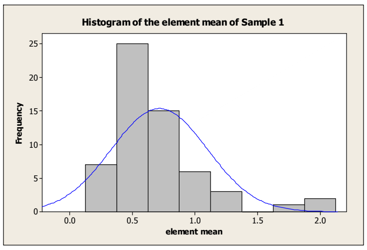
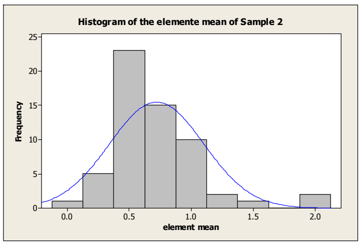
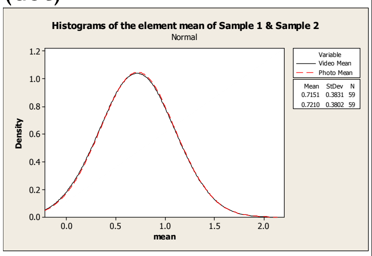
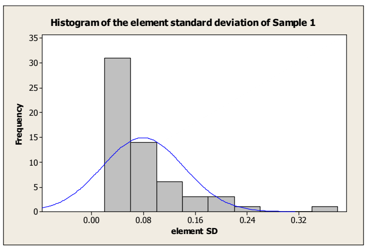
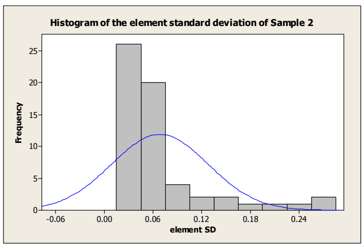
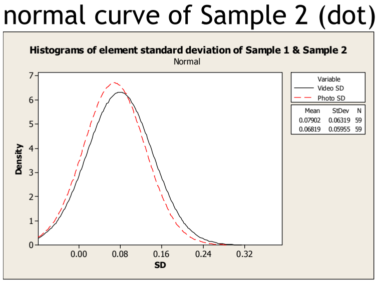
Element mean of Sample 1 and element mean of Sample 2		
(a) Sample 1	(b) Sample 2	(c) Normal curve of Sample 1(solid) superimposed with the normal curve of Sample 2 (dot)
		
n=59 Mean 0.7151 St. Dev 0.3831	n=59 Mean 0.7210 St. Dev 0.3802	

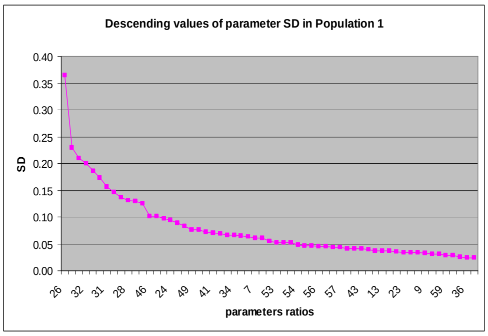
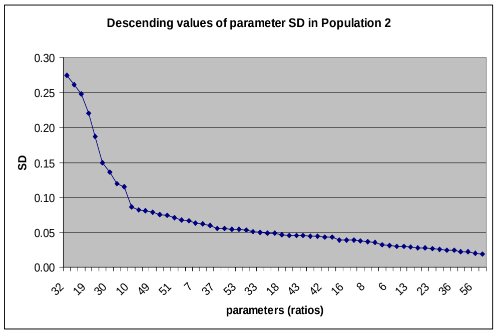
Table 6.22 Relative frequency histograms of the element standard deviation for Sample 1 and the element standard deviation for Sample 2

Element standard deviation of Sample 1 and element standard deviation of Sample 2		
(a) Sample 1	(b) Sample 2	(c) Normal curve of Sample 1 (solid) superimposed with the normal curve of Sample 2 (dot)
		

n=59		n=59	
Mean	0.0790	Mean	0.0682
St. Dev	0.0632	St. Dev	0.0596

The standard deviations of elements in each sample were placed in descending order and shown in Table 6.23. Increasing amounts, starting at 5% and ending at 50% of the lowest variant elements were removed to determine what effect if any this would have on the degree of separation between true positive and true negative faces in the comparison of Sample 1 and Sample 2.

Table 6.23 Descending order of element standard deviations in each sample place on a line graph.

Sample 1	Sample 2
	

Up to 50% of the lowest variant elements were removed from the feature vector and Figure 6.2 below shows in (a) the linear measurements that make up the 59 proportions used thus far in the anthropometric analysis, and in (b) the linear measurement that comprise the *remaining* 29 proportions after 50% of the lowest variant proportions were removed.

(a) (b)

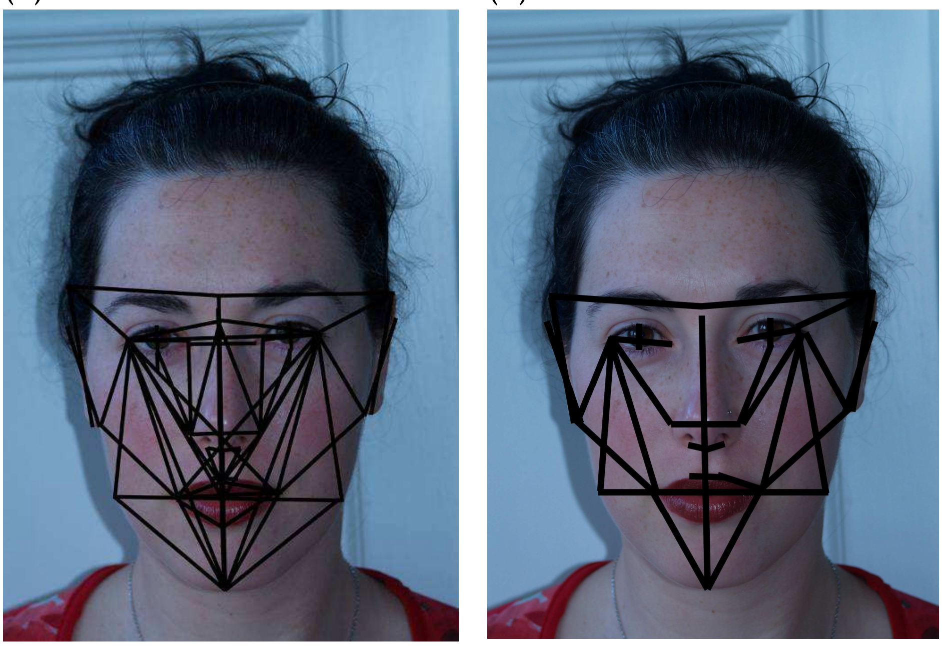


Figure 6.2 Figure (a) represents the linear measurements of all proportions that were tested. Figure (b) represents the linear measurements of the 29 proportions retained after 50% of the lowest variant proportions were removed.

The 30 proportions that were eliminated from the comparison process in varying degrees are listed in the appendix and Figure 6.3 below illustrates the linear measurements that made up the lowest variant proportions in increasing stages; 5%, 10%, 25% and 50%. The ideal proportion will depict a high level of variance and considered to be distinguishable among individuals. However, proportions can also show a high level of variance when the linear measurements involve short distances between landmarks due to the error in landmark placement. It is possible that many of the high variant proportions retained were done so because they involved landmarks with short linear distances and not because they were necessarily distinguishable amongst individuals.

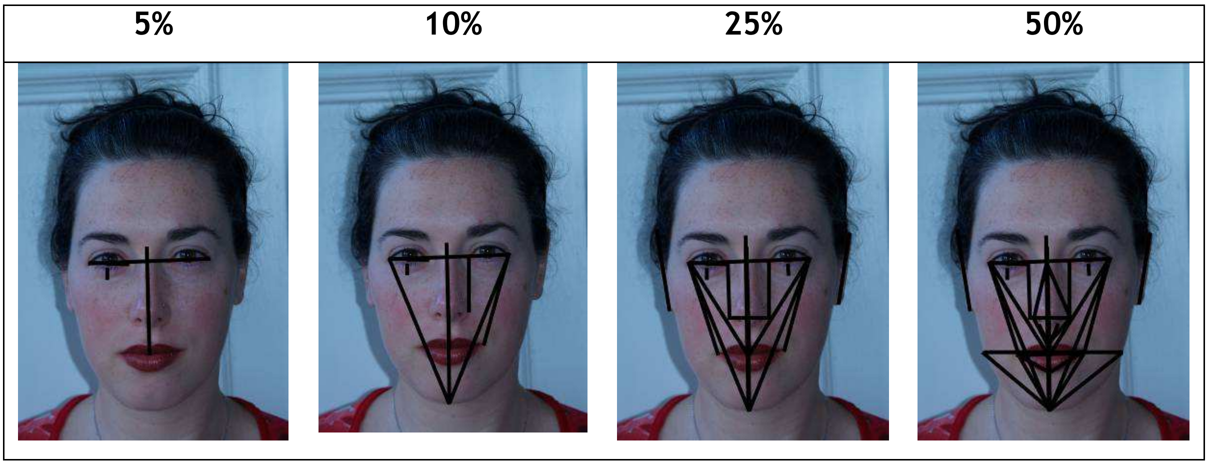


Figure 6.3 The increasing percentages of low variant elements that were removed from the feature vector for comparison shown in four photographs illustrating the linear measurements making up those proportions.

The first comparisons completed after removing the chosen lowest variant elements were conducted within sample and distribution results are shown as relative frequency histograms. For these comparisons, the sizes of the two samples were held consistent, using the same 80 faces that appeared in Sample 1 and Sample 2. The percentages of lowest variant elements were removed in four stages and each is represented by a different table (Tables 6.24-6.27). Only the Cosine θ distance (Z-normalized) equation was applied to this analysis because of the close statistics the two samples achieved when compared within sample using this equation. Sections (a) and (b) of the tables illustrate the relative frequency distribution achieved when each face in a sample was compared against itself and every other face in the same sample; duplicates and true positive faces removed. Section (c) of the tables shows the normal curves of the two samples superimposed. In these four tables, the superimposed normal curves of Sample 1 are represented by the solid line and the dotted line represents Sample 2.

The histograms of Sample 1 and Sample 2 in each table showed a normal distribution and the superimposed normal curves of Sample 1 and Sample 2 in Tables 6.24 (5%), 6.25 (10%), and 6.26 (25%), were virtually superimposed. This is similar to what happened when this equation was applied to the full feature vector in Table 6.7. However, when 50% of the elements were removed, there was a slight difference in standard deviation between samples shown in the distribution at the height of the curve, which shows that removing 50% of the 59 elements has a slight but not necessarily significant effect. The superimposed normal curves are useful for the next analysis of Sample 1 against Sample 2 because it demonstrates that the statistics are stationary and any separation rate achieved between true positive and true negative faces can be attributed to a variable other than the media of the images. The superimposed curves also establish that the removing 5%, 10% or 25% of the 59 elements had little effect in changing the statistics of the two samples.

Table 6.24 Relative frequency histograms illustrating comparisons of Cosine θ distance equations (Z-normalized) within a single sample for Sample 1 (a) and Sample 2 (b) with 5% of lowest variant elements removed. Normal curves are superimposed in (c).

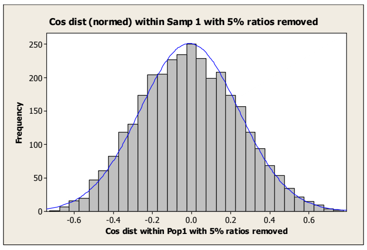
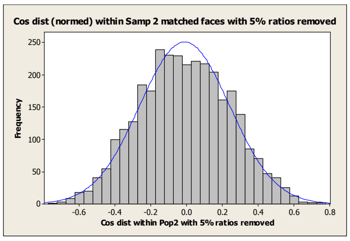
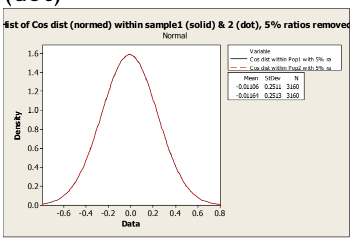
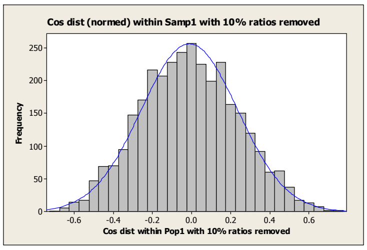
Cosine θ distance (Z-normalized) within Sample 1 and within matched faces Sample 2 with 5% of lowest variant elements removed		
(a) Sample 1	(b) Sample 2 matched faces	(c) Normal curve of Sample 1 (solid) superimposed with the normal curve of Sample 2 (dot)
		
n= 3160 Mean -0.0111 St. Dev 0.2511	n= 3160 Mean -0.0116 St. Dev 0.2513	

Table 6.25 Relative frequency histograms illustrating comparisons of Cosine θ distance equations (Z-normalized) within a single sample for Sample 1 (a) and Sample 2 (b) with 10% of lowest variant elements removed. Normal curves are superimposed in (c).

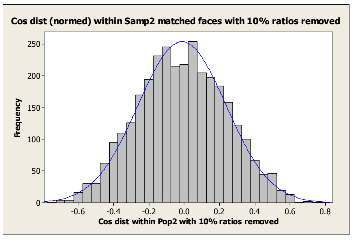
Cosine θ distance (Z-normalized) within Sample 1 and within matched faces of Sample 2 with 10% of lowest variant elements removed

(a) Sample 1



n= 3160	
Mean	-0.0109
St. Dev	0.2466

(b) Sample 2 matched faces



n= 3160	
Mean	-0.0117
St. Dev	0.2489

(c) Normal curve of Sample 1 (solid) superimposed with the normal curve of Sample 2 (dot)

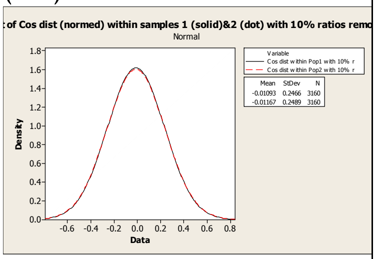
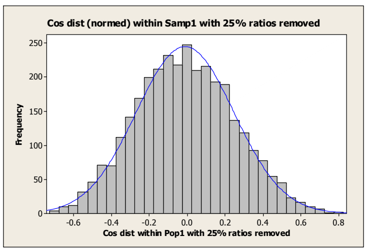


Table 6.26 Relative frequency histograms illustrating comparisons of Cosine θ distance equations (Z-normalized) within a single sample for Sample 1 (a) and Sample 2 (b) with 25% of lowest variant elements removed. Normal curves are superimposed in (c).

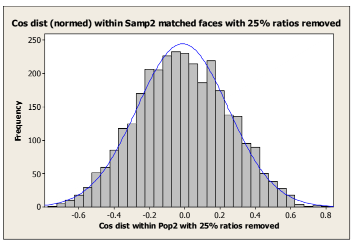
Cosine θ distance (Z-normalized) within Sample 1 and within matched faces of Sample 2 with 25% of lowest variant elements removed

(a) Sample 1



n= 3160	
Mean	-0.0104
St. Dev	0.2580

(b) Sample 2 matched faces



n= 3160	
Mean	-0.0116
St. Dev	0.2579

(c) Normal curve of Sample 1 (solid) superimposed with the normal curve of Sample 2 (dot)

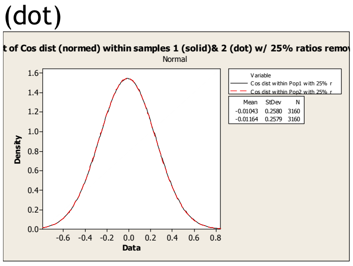
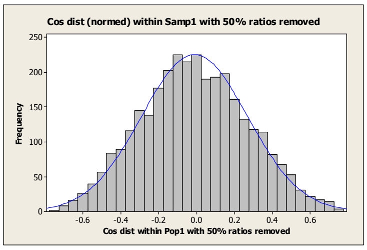
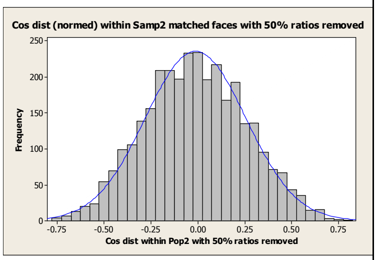
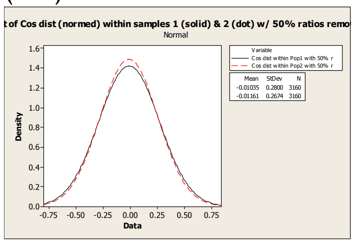


Table 6.27 Relative frequency histograms illustrating comparisons of Cosine θ distance equations (Z-normalized) within a single sample for Sample 1 (a) and Sample 2 (b) with 50% of lowest variant elements removed. Normal curves are superimposed in (c).

<table>
<tr><th colspan="3">Cosine θ distance (Z-normalized) within Sample 1 and within Sample 2 with 50% of lowest variant elements removed</th></tr>
<tr>
<td>(a) Sample 1


<table><tr><td colspan="2">n= 3160</td></tr><tr><td>Mean</td><td>-0.0104</td></tr><tr><td>St. Dev</td><td>0.2801</td></tr></table></td>
<td>(b) Sample 2 matched faces


<table><tr><td colspan="2">n=3160</td></tr><tr><td>Mean</td><td>-0.0116</td></tr><tr><td>St. Dev</td><td>0.2674</td></tr></table></td>
<td>(c) Normal curve of Sample 1 (solid) superimposed with the normal curve of Sample 2 (dot)

</td>
</tr>
</table>

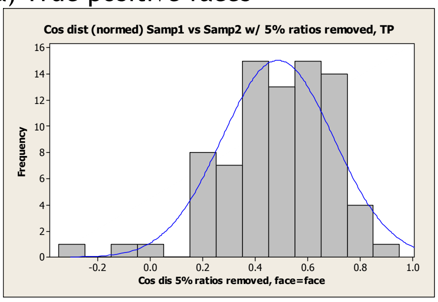
Tables 6.28-6.31 illustrate the distributions achieved in the comparison of Sample 1 against the matched faces of Sample 2 when the Cosine θ distance (Z-normalized) equation was applied and up to 50% of the lowest variant elements were removed. Only this equation was used for the analysis because it achieved the highest separation rate of true positive and true negative faces in the between sample comparisons (Table 6.13.) This analysis compared Sample 1 against Sample 2 and the distributions in the tables were shown as true positive cases (a), true negative cases (b) and the combination of the true positive and true negative cases (c). The superimposed normal curves of the true positive and true negative faces were shown in (d). The true positive normal curve is represented by the dotted line and the true negative normal curve represented by the solid line. Following Table 6.31 is Figure 6.4 which displays section (d) of Tables 6.28-6.31 and was created to illustrate the relationship between of each degree of element removal.

In the four stages, 5%-50%, of low variant element removal from the feature vector, the relative frequency histograms display very similar distributions of true positive and true negative faces both to each other and to the previous comparison utilizing the full feature vector shown in Table 6.13. A summarization of this relationship is shown in Table 6.32. Of the superimposed normal curves of section (c) located in Tables 6.28-6.31, the dotted curves (true positive faces) display a higher density and this is because as these are true positive faces, the majority of the data is concentrated towards the 1.0 mark which will be reflected in the distribution. According to Figure 6.4, neither percentage group of the lowest variant elements removed from the feature vector had and any more influence than another in separation rates. This indicates that no constructive effect was achieved with varying stages of removal of low variant elements to the separation rates of true positive and true negative faces in this between sample comparison and with the application of the Cosine θ distance (Z-normalized) equation using the full feature vector. The separation rate was not improved with the removal of certain elements in the feature vector, but neither was it harmed.

Table 6.28 Relative frequency histograms illustrating comparisons using the Cosine θ (Z-normalized) distance of Sample 1 vs. the matched faces in Sample 2 showing true positive cases (a), true negative cases (b), both cases together (c), and both cases superimposed (d) with 5% of lowest variant elements removed from vector.

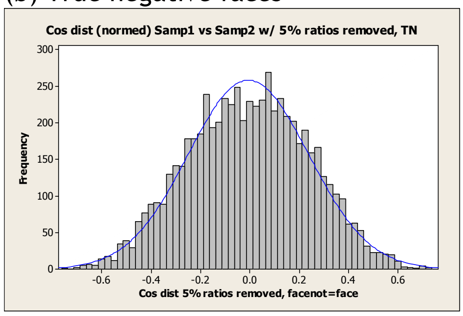
Cosine θ distance (Z-normalized) for Sample 1 vs. Sample 2 matched faces with 5% of lowest variant elements removed

(a) True positive faces



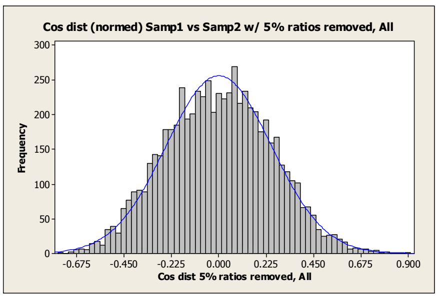
N=80	
Mean	0.4863
St. Dev	0.2123

(b) True negative faces



n=6320	
Mean	-0.0054
St. Dev	0.2443

(c) True positive faces + true negative faces



n=6400	
Mean	0.0007
St. Dev	0.2500

(d) Normal curves of true positive (dot) and true negative (solid) faces superimposed

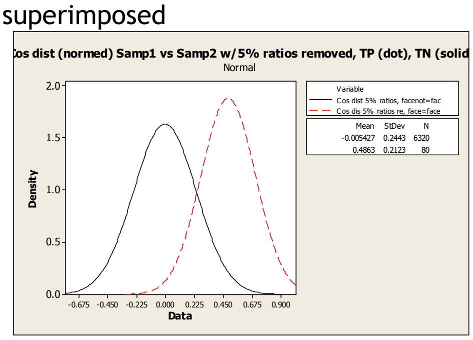
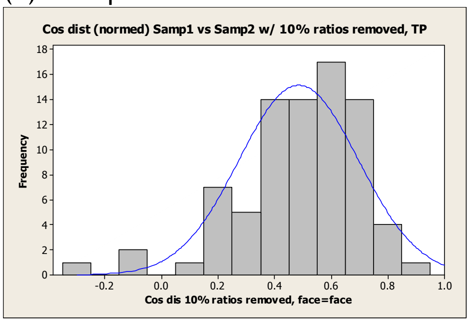


Table 6.29 Relative frequency histograms illustrating comparisons using the Cosine θ (Z-normalized) distance of Sample 1 vs. the matched faces in Sample 2 showing true positive cases (a), true negative cases (b), both cases together (c), and both cases superimposed (d) with10% of lowest variant elements removed from vector.

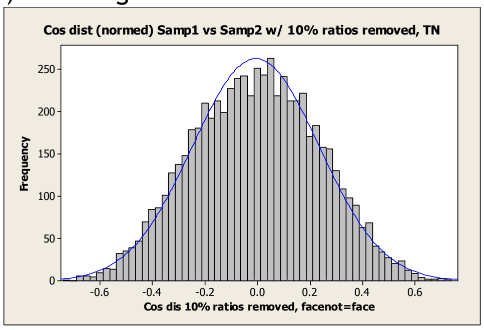
Cosine θ distance (Z-normalized) for Sample 1 vs. Sample 2 matched faces with 10% of lowest variant elements removed

(a) True positive faces



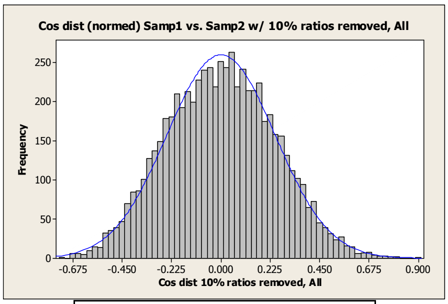
N=80	
Mean	0.4850
St. Dev	0.2108

(b) True negative faces



n=6320	
Mean	-0.0055
St. Dev	0.2406

(c) True positive faces + true negative faces



n=6400	
Mean	0.0007
St. Dev	0.2463

(d) Normal curves of true positive (dot) and true negative (solid) faces superimposed

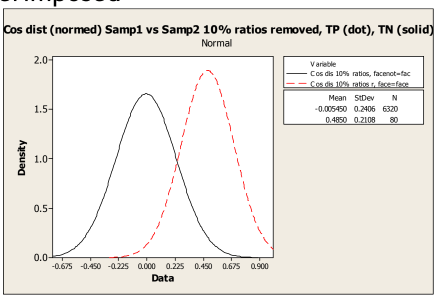
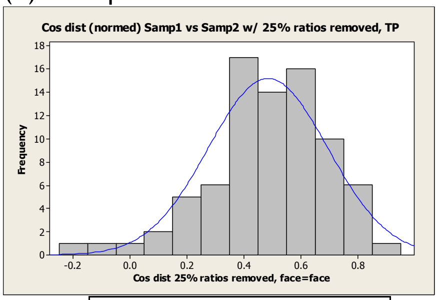


Table 6.30 Relative frequency histograms illustrating comparisons using the Cosine θ (Z-normalized) distance of Sample 1 vs. the matched faces in Sample 2 showing true positive cases (a), true negative cases (b), both cases together (c), and both cases superimposed (d) with 25% of lowest variant elements removed from vector.

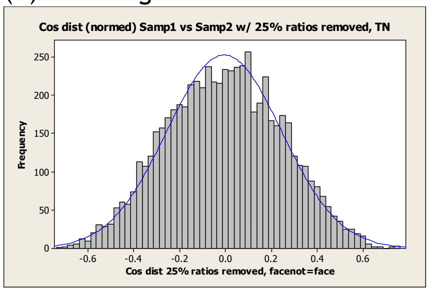
Cosine θ distance (Z-normalized) for Sample 1 vs. Sample 2 matched faces with 25% of lowest variant elements removed

(a) True positive faces



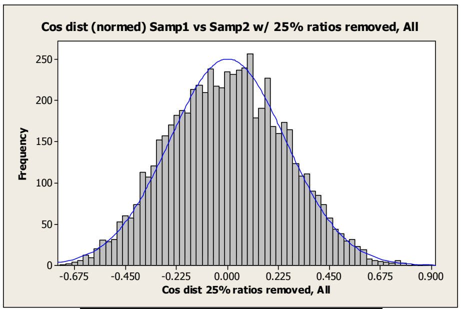
N=80	
Mean	0.4843
St. Dev	0.2104

(b) True negative faces



n=6320	
Mean	-0.0052
St. Dev	0.2502

(c) True positive faces + true negative faces



n=6400	
Mean	0.0009
St. Dev	0.2555

(d) Normal curves of true positive (dot) and true negative (solid) faces superimposed

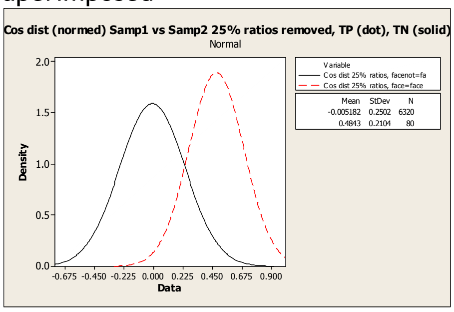
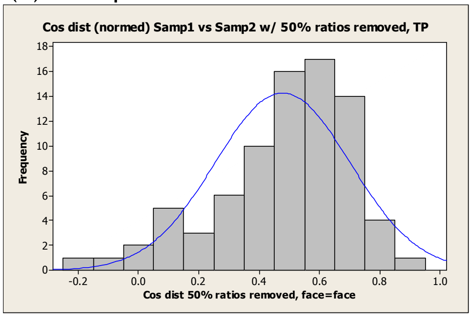


Table 6.31 Relative frequency histograms illustrating comparisons using the Cosine θ (Z-normalized) distance of Sample 1 vs. the matched faces in Sample 2 showing true positive cases (a), true negative cases (b), both cases together (c), and both cases superimposed (d) with 50% of lowest variant elements removed from vector.

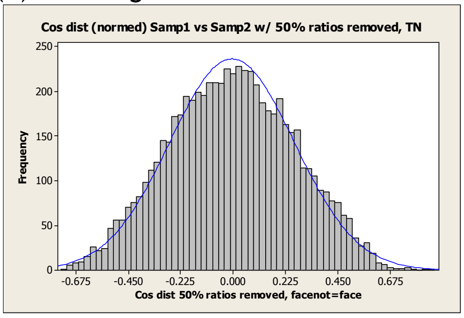
Cosine θ distance (Z-normalized) for Sample 1 vs. Sample 2 matched faces with 50% of lowest variant elements removed

(a) True positive faces



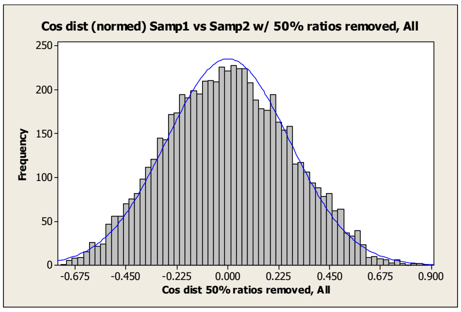
N=80	
Mean	0.4775
St. Dev	0.2236

(b) True negative faces



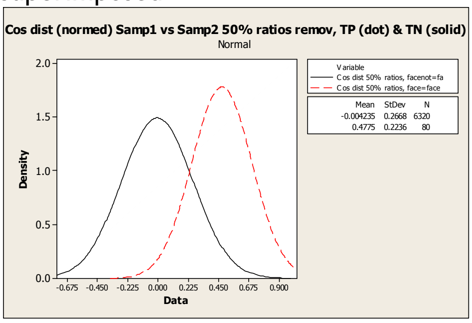
n=6320	
Mean	-0.0042
St. Dev	0.2668

(c) True positive faces + true negative faces



n=6400	
Mean	0.0018
St. Dev	0.2716

(d) Normal curves of true positive (solid) and true negative (dot) faces superimposed



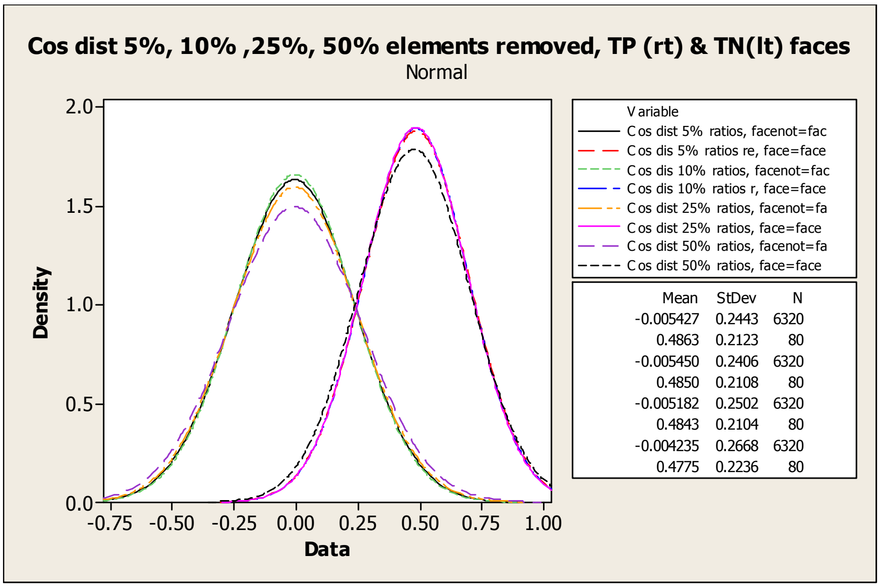
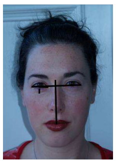
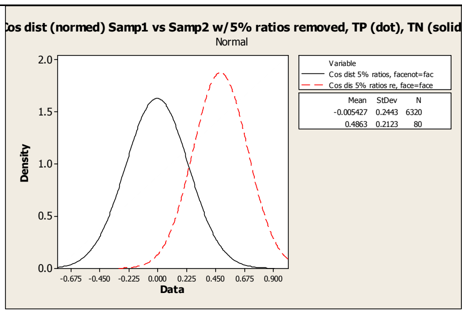
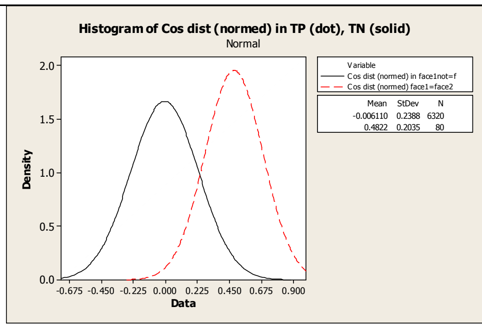
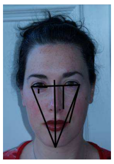
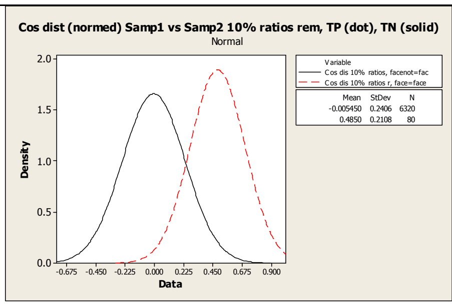
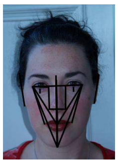
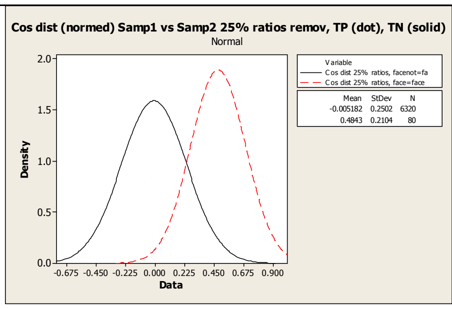
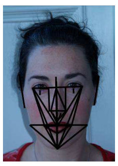
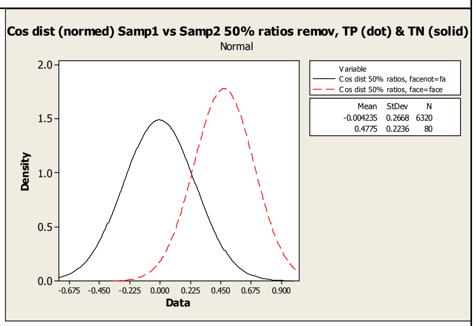


Figure 6.4 Normal curves of true positive and true negative faces superimposed with 5, 10, 25 and 50% of lowest variant elements removed from feature vector.

Table 6.32 Summarization of the overlapped normal curves of true positive (dot) and true negative (solid) faces. The feature vector with the increasing amounts of lowest variant proportions removed is alongside the full feature vector using the Cosine θ (Z-normalized) distance equation in the comparison of Sample 1 vs. Sample 2.

	Feature vector with increasing amounts of lowest variant proportions removed		Full feature vector
5% 			
10% 			
25% 			
50% 			

6.3.5 Subsample of Images Requiring Human Verification

Although not named as such, the subsample of images that would require human verification was monitored previously as separation rates of true positive and true negative faces. The overlapped area under the normal distribution curves defines statistically the chance of mixing up true positive and true negative faces, and in turn describes the chance of achieving false positive or false negative face matches which is referred as the misclassification rate in this analysis. To investigate this subsample of images that would require human verification, the Cosine θ (Z-normalized) distance equation was applied to the comparison of Sample 1 against the entire 119 images of Sample 2. As the Cosine θ (Z-normalized) distance equation using the full feature vector generated the most successful separation rate of true positive and true negative faces, it was used to further analyze the misclassification rate. The 119 photographic images of Sample 2 were included to more mimic real life situations where the database photographs in Sample 2 will most likely have extraneous images to that of the suspect images in Sample 1.

To investigate the misclassification rate, the following 3 steps were carried out:

1. The Cosine θ (Z-normalized) distance equation was found between every face in Sample 1 against every face in Sample 2 and all duplicate cases were removed giving a total of 80 comparisons. Other than using the entire Sample 2, this was completed no differently than in previous between sample comparisons. Z-normalized values were derived from each samples own statistics.

2. The true positive face match between Sample 1 and Sample 2 was noted and the value was established as the cut off point for each of the 80 comparisons.

3. For this test, a match value of 1.0 signifies an identical face and therefore any result that was greater than or equal to that cut off point was counted and tallied minus the true positive case.

The minimum count possible was 0 and the maximum count possible was 118. The count tallied in this range was equal to the number of faces in Sample 2 that were closer to or equal in Cosine angle distance to the true positive face of Sample 1.

Results are illustrated in the form of a pie chart located below in Figure 6.5. The range of faces in Sample 2 that were closer to or equal in distance to the true positive face of Sample 1 was from 0 to 109 and is arranged in order in a clockwise rotation in the pie chart. The distribution shows that 49% of counts were at zero signifying that for almost half of the comparisons there were no other faces in Sample 2 that were closer in Cosine angle distance than the true positive match in Sample 1. About 75% of the comparisons achieved five false positives or less and the maximum count was 109 which only occurred in 1% of the 80 comparisons meaning that there were 109 false positive cases for that comparison.

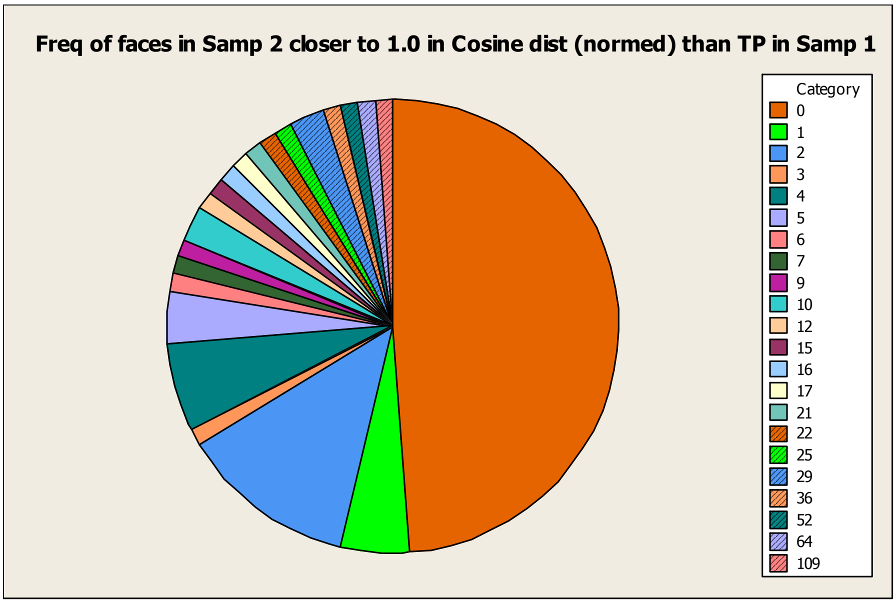


Figure 6.5 Pie chart illustrating the subsample of images requiring human verification when the Cosine θ (Z-normalized) distance equation was applied in the between sample comparison.

6.4 Discussion

The remainder of this chapter is a summarization and discussion of the results obtained after the completion of the study undertaken. A generalized discussion of anthropometry and its relation to this study is examined in Chapter 7.

6.4.1 Overview of the Study

Using high resolution photographic research material, the object of the study was to assess whether if a facial anthropometric feature vector could be utilized to distinguish between individuals of a similar age group, ancestry and sex. A total of 59 proportions derived from 37 landmarks were chosen to achieve a balance of the horizontal and vertical regions of the face, taking into account angle rotation. It becomes more difficult to manage landmarks which are difficult to place and which show large variations because of slight changes in the camera angle. Knowledge of the type of information gathered in this study may help in future to narrow down the number of possible suspects in an investigation. In order to establish that two faces were the same and use this identification method to positively identify rather than eliminate suspects, it would be required to show that the probability of this happening in the rest of the population at random was in effect zero [40].

Attempts to test the hypothesis: "Using a comparison of anthropometric facial proportions, it is possible to discriminate between individuals of two samples," was pursued in this chapter by conducting several analyses. The 59 proportions were treated as a multi dimensional feature vector, effectively treating the 59 proportions as a whole and simple dot product and distance comparisons were first performed within a single sample and then performed in between sample comparisons; video images against photographic images. The feature vector was utilized in three types of equations testing the differences between faces in the samples. The mean absolute difference, the Euclidean distance and the Cosine θ distance. Normalization was applied to the proportional values as a way to equalize the feature vector values in each sample and account for the statistics that different camera parameters would produce. Z-normalization enhances any differences between means and makes the interpretation of the data more straightforward allowing small differences in the data can to be more simply seen.

The first step to answering the objectives laid forth in the experimental design of the study was to evaluate each sample of images to determine if once the equations were applied, any differences could be seen between the two samples. As only one image of each face in each sample was available from the research material, it was only possible to determine if two faces were different, rather than if two faces were the same. For each sample, every face in the respective sample was compared to itself and every other face in the same sample. Testing faces against those found in the same sample is important because it allows the equations to ascertain if there are any differences between faces which fall under the same conditions. This means that other than the possibility of slight changes in facial expression the facial proportions will be the only changeable variable between faces as all other variables remain constant; same media, same operator placing landmarks and same facial pose. The similarity or dissimilarity between the groups of faces in a single sample is presented when the distribution statistics, in this case mean and standard deviation, of the two samples are compared. A comparison of the statistics produced from comparing faces within its own sample can be used to predict the outcome of how separable faces in the two samples are. When the statistics of two samples are similar, it is possible to show if two faces are distinguishable. The likelihood of faces still being distinguishable from each other does not diminish with two samples whose statistics are dissimilar but the possibility may be more remote.

In each sample the equations of mean absolute difference and Euclidean distance between proportions resulted in a one sided distribution, especially with the smaller sample sized Sample 1. However, the application of the Cosine θ distance (mean subtracted) and (Z-normalized) equations produced normal distributions in each sample. The means and standard deviations for the two samples were similar and this was supported by the illustration of the superimposed normal distribution curves of both samples. The Cosine θ (Z-normalized) distance equation provided the best representation of stationary statistics between the two samples as it was the only equation that resulted in a direct superimposition of curves even taking into account the different sample sizes and allowing for landmark placement error. The lack of overlap as a result of the application of this equation shows that any difference in faces is a result of the true difference in faces rather than a result of the different media of the two samples of images.

Before commencing the between sample comparison of the faces using the chosen equations, a small inter-operator study was carried out, to assess the influence of landmark placement conducted by multiple operators. This is important to test because although landmark placement on all 199 images used in the comparative process of this chapter was conducted by a single operator, this would not likely be the case in the real world. Landmark placement has been tested by other researchers on 3D images in a clinical setting and it was suggested that average operator error varies widely [160]. The current study was conducted with one experienced operator but the remaining operators were inexperienced. It would be beneficial to analyze this data further in an inter-operator study using experienced operators located in different graphical regions because this scenario would be more likely as a police procedure, but this option was not available. Experience was shown to be a benefiting factor when the inter-operator variation in taking standard skeletal measurements was tested with a panel of experienced forensic anthropologists and found to be minimal [161]. For evidence interpretation in the court of law, any variation that occurs in the data as a result of multiple operators placing landmarks should be small.

The Cosine θ distance (Z-normalized) equation was used to compare the re-landmarked images because when applied in the comparison of faces within a single sample it was found to be the equation in which the statistics of the two samples were the most stationary. Normalization for the Cosine θ distance (Z-normalized) equation was carried out both using the statistics from the subset sample and also using the statistics derived from the entire sample. Both scenarios were tested in case a larger pool of faces had an effect on the statistics of the sample and in turn the comparisons. Data resulting from the application of these equations were illustrated as superimposed normal histogram curves of true positive and true negative faces. Each face in the subset sample was compared to every other face in the subset sample and all duplicates and identical face matches were removed from the distribution. It was hypothesized conducting an inter-operator test, using high resolution research material but completed by inexperienced operators, would produce a greater amount of variation than from an experienced operator and this theory was tested and found to hold.

A strong separation rate of true positive and true negative faces was seen with the experienced operator; however, once the data from the inexperienced operators were included, the distribution no longer depicted the strong separation. Although variation of landmark placement caused by each operator was investigated and discussed in Chapter 5, a further study analyzing the distribution achieved from the re-landmarked images of each operator after applying the Cosine θ distance (Z-normalized) equation could determine if any of the inexperienced operators also achieved the same strong separation rate as the experienced operator. An inexperienced operator producing a similar degree of separation to the experienced operator would signify that the small separation rate produced from all operators was caused by the inclusion of multiple operators rather than their experience. However, from the literature [161], it can be predicted that the spread from a single inexperienced operator would be larger than an experienced operator.

Once samples were looked at individually, a between sample comparison was conducted to determine how distinguishable the faces were in the two samples. Every face in Sample 1 was compared to every face in Sample 2 using only the faces in Sample 2 that had a true positive match to a face in Sample 1. Although the two samples contained the same faces posed the same, taken on the same day with facial landmarks placed by the same operator, the Media in which the images were obtained made the images of the two samples distinguishable.

The mean absolute difference, the Euclidean distance and the Cosine θ distance equations were applied in the between sample comparison. The two samples contained a finite number of 80 images and it was possible to determine the sample statistics (mean and standard deviation) necessary to carry out the Z-normalization of the equation. Further attempts to distinguish faces of two samples were accomplished by calculating the Cosine θ distance equations using the log of raw proportions to determine if a larger separation rate would be yielded. Tests were conducted assuming that the statistics from each sample were known but there is a strong possibility that this may not be the case in a real world setting. Therefore additional analyses were conducted applying the Cosine θ (Z-normalized) distance equation and Z-normalization was carried out as if the statistics from either of the samples were known and statistics of the summation of the two samples. The final between sample comparison conducted to test separation rates of true positive and true negative faces was to remove the lowest variant proportions of the feature vector.

Once the respective equations were conducted, superimposed normal histogram distribution curves of true positive faces and true negative faces were used to illustrate the discrimination of the two groups. In general, a narrower distribution was seen for the true positive faces. This was because as the distribution contained only true positive matches, the data should be centred on a smaller range of values. The amount of overlap correlated to the possibility of achieving either a false positive or false negative face match.

Superimposing the normal curves to demonstrate the separation between true positive and true negative faces, the Cosine θ distance (Z-normalized) equation produced the least amount of overlap between true positive faces and true negative faces when statistics of the two samples were known. The match values of true negative faces in the superimposed histogram normal curves begin to trail off at 0.68, indicating that although it is still possible to achieve a true negative identification above this value, it is likely that a returned match score of below 0.68 will result in a true negative face after closer examination. Although this result occurred in this study, it may not be replicated with a larger test database. Examination of the superimposed curves showed approximately a 25% chance of achieving an incorrect identification. The conclusion made from this investigation was that this equation was the best predictor of face discrimination tested on the available images.

Attempts to improve this separation rate saw the application of the log of the raw proportion values to the Cosine θ distance equations. Of the three equations applied, the Cosine θ (Z-normalized) distance equation resulted in the greatest separation rate of true positive and true negative faces. This supplies additional evidence that applying the Cosine θ distance equation using Z-normalized values was the most beneficial way to distinguish between individuals for this research. The overlap of true positive and true negative faces was not any greater than that of using the raw Z-normalized proportional values and therefore did not provide any incentive for use of log values.

Not all circumstances will provide the sample statistics needed to derive the required Z-normalized values. For cases where a single sample's statistics are known, those statistics were applied to both samples to determine their Z-normalized values. There was a slightly smaller amount of overlap in normal curves of true positive face and true negative faces using statistics from Sample 2 compared to using statistics derived from Sample 1. This means that using the statistics from Sample 2 provided a smaller chance of misclassification, however, the overlap when applying the statistics from Sample 2 to both samples still supersedes that of using each samples own statistics to Z-normalize values.

Testing different cases of a single samples' statistics demonstrates the need for knowing each sample's own statistics for comparing images using the Cosine θ (Z-normalized) distance equation.

In practice, statistics for the sample of database photographs will most likely be known, although will change constantly with any new addition to the sample. The statistics will be known because as the photographs are posed at standardized angles, the landmarks and consequent proportions will be consistent throughout the entire sample. Tracking the statistics of a group of video images will be more complicated as the images with each case will contain different positions and successive landmarks and proportions. Sample statistics would need to be generated by subdividing video images into the facial poses they are positioned.

Subsequent to the conclusion that the Cosine θ (Z-normalized) distance equation using each samples own statistics was the most effective way to discriminate between faces of two samples in this study, further attempts were made to improve the separation rate of true positive and true negative face matches by removing the lowest variant elements (proportions) of the feature vector. Increasing amounts of the lowest variant elements in amounts of 5%, 10%, 25% and 50% were removed from the feature vector of each face. Removing either too many or not enough elements can have a detrimental effect on a comparison, which is why eliminating varying amounts of the lowest variant elements were tested. Separation rates of true positive and true negative faces in the between sample comparison were found by applying the Cosine θ (Z-normalized) distance equation. This equation was chosen over any other because it garnered the largest separation rate when applied to the full feature vector.

The increasing amounts of low variant elements were first removed and a within sample comparison conducted to determine the effect on the statistics of each sample. Once this was completed, the same degrees of low variant elements were removed in between sample comparisons. Relative frequency histograms were used to illustrate the analyses. Removing the elements had no effect of the statistics of the two samples, and except for a slight variation with 50% removal, the statistics were held stationary. Following the between sample comparisons, no improvement was seen in the separation rate of true positive faces and true negative faces subsequent to removal of any of the varying percentages of lowest variant proportions. The conclusion drawn from this is that there was an inconsequential difference when removing increasing numbers of the chosen low variant elements in this study; neither helping nor hindering the comparison.

The majority of the elements with a high variation that remained for the comparison appeared to be proportions in which either there was a small distance between landmarks, such as around the eyes, or included landmarks that the literature has shown to be of little comparative worth, such as the ears. This indicates that the high variance that certain elements produced could have been due to an error that occurs in landmark placement and not necessarily a difference in individuals' facial features. It is possible that in the course of choosing which proportions should be removed; first normalizing the proportions by dividing the mean by the standard deviation would produce significantly different results.

Although highly individual, the complexity of the human face indicates that comparing 2D images does not offer the same rate of individuality assessment that DNA or fingerprints does. However, if comparing 2D images can narrow down the possible matches to be further verified by a human operator, the time saved can be extensive. Once all comparisons were completed and the true positive and true negative separation rates were found between Sample 1 and Sample 2, a method of determining the misclassification rate, or rate of obtaining a false positive or false negative face match in a between sample comparison was analyzed. This can also be referred to the number of images in a subsample that would be required to be verified by a human with further identification methods, including a morphological analysis.

In order to address the rate of misclassification between faces in Sample 1 and faces in Sample 2, the Cosine θ (Z-normalized) distance equation was applied. The true positive face match between Sample 1 and Sample 2 was noted, the value was established as the cut off point for each of the 80 comparisons and a count was recorded of how many other faces in Sample 2 produced a value closer to or equal in distance to the true positive face in Sample 1. In practice, the method in which this experiment was conducted would not be available, as the true positive face would not be known; however, it was necessary to initially conduct the experiment in this way in order to detect the limitations of this method.

The Cosine θ (Z-normalized) distance equation was used because it achieved the smallest amount of overlap between true positive and true negative faces in the between sample comparison. Results of tallying the count of false positives achieved in this analysis, agree with previous true positive and true negative separation rates. The distribution of 49% of counts was at zero signifying that for almost half of the comparisons there were no misclassifications. About 75% of the comparisons achieved five false positives or less. A larger sample of images should be tested to determine if consistent results would be produced.

6.4.2 Limitations of the Study and Further Work

One of the limitations encountered over the course of this study was that the number of images for comparison did not provide the ideal sample size in relation to the number of elements tested. The method was tested on images of a similar physiognomy, as occurs in practice, and these were the images available that suited the criteria. Another limitation was that number of proportions tested was considerably smaller than the possibilities based on the number of landmarks. The number of landmarks chosen was not in question as it was believed that enough landmarks were selected to gain an acceptable representation of the face. Given the equipment available, in addition to time restrictions, it was not feasible to include the entire number of possible proportions. Incorporating the 1,772,892 possibilities into a feature vector is something that could be done in further study. It would then be necessary to determine which of those would be beneficial and which would hinder the comparison process. Depending on the pose of the image, it would be suitable to know which proportions were valuable. The final limitation of the study was that a more comprehensive statistical analysis could have further improved the data analysis. Possible alternatives are discussed below.

When comparing a suspect image to a database of photographs using the chosen landmarks and anthropometric proportions, the best-case scenario would produce one true positive, zero false negatives and zero false positives. In other words there would not be a single other face in Sample 2 that would be close in distance than its true positive face in Sample 1. The rate of misclassification tested thus far takes into account the true positive match between Sample 1 and Sample 2 plus any matches deemed closer to 1.0. However, in cases where the true positive face was the best match, it does not inform the researcher how close the next best match was to the true positive match. A true positive match will hold more weight as evidence when the next best match proves to be of significant distance away. Future work would determine the distance between the true positive match and the next best match by finding the log likelihood ratio between matches. This is the log of the ratio of the second best match to the best match. The full analysis would include finding the log likelihood ratio between each match based upon a descending order of matches which would show their relation to each other. A poor log likelihood ratio could be indicative of the necessity for more images to be tested.

A portion of the false positive and false negative faces under the superimposed curves of true positive and true negative faces must be included in the selection of faces for verification by a human. The question arises as to how wide would the margin need to be to include more, or all, of the population: two standard deviations from the distribution mean, three standard deviations? A wider range will result in more false positives, so in order to exclude all false positives a narrower range would need to be employed. It is not beneficial to have too wide a range that would include everyone. Ideally, as many people as possible need to be excluded correctly in order to focus upon a decreased number of suspects. One possible way this method could be used in the future to narrow down possible suspects is to operate within a margin of two standard deviations.

In order for anthropometry to be used with confidence to confirm a comparison of two or more individuals, research results would have to show some level of reliability in regard to the number of true identifications made. The investigations undertaken in this chapter to determine if it is possible to discriminate between individuals of two samples using a multi dimensional facial feature vector found that the Cosine θ distance was the best discriminator this but could further be improved upon if time allowed by administering a more comprehensive statistical analysis. A common precursor to further statistical investigation is a multivariate technique called a Procrustes analysis. A Procrustes analysis would match landmarks or shapes from two sets of data removing the variation of translation, rotation and scaling in the data so that the data becomes a single frame of reference. Statistical tests such as Hotelling's or Students t test could be applied to the data investigated in this study; however, the results achieved thus far can be interpreted and determined from looking at the data directly.

Halberstein used Spearman and Kendall Tau statistical tests to determine if there was a statistical significance between the proportions of the two individuals being compared [38]. However, the article does not report if any tests were conducted to assess the possibility of the suspect's proportions achieving a statistical significance with another member of the population and this category of data analysis should not be ignored.

6.4.3 Conclusions

At the moment, the practical implications for the method investigated are limited. Using the method to eliminate possibilities from the database could happen in two ways. As people can easily change their appearance, facial measurements from a wide pool of suspects could be compared and the method used to narrow down suspects from the entire database. With large databases the complexity of this increases due to long lists of suspect possibilities. In the absence of DNA or fingerprints, it may be more beneficial to reduce the database first with a general description of the suspect based on hair colour, sex, or any other identifying feature. A facial measurement comparison could then be implemented to further reduce the list of suspects and additional evidence, such as locality would decrease the list even more.

The general conclusion derived from the investigations undertaken in this study was that these tests do not offer a significant and infallible method of discriminating between individuals of two samples. At best they may offer corroborating evidence and could be used to narrow down a list of suspects as long as other evidence was available. The between sample comparisons conducted in this chapter did not account for any type of error in the method. The effect on facial measurements caused by the error was described as intra and inter-operator studies in Chapter 5 and Section 6.3.2. Conducting the comparisons again taking into account the error in facial measurements as a result of landmark placement by different operators would most likely modify the misclassification rate of face matches and possibly negate any substantive effect of facial matching that was shown. The cases tested in this research were done as a best case scenario; facial poses in both samples were faced frontal, images from both samples were taken on the same day, landmarks on images from both samples were placed by the same operator and the quality of video was high. These 'best case' samples would most likely not be available to forensic scientists but the advantages for further testing on worst case scenarios are non-existent if the discriminatory power is not sufficient with the best cases.

7 Discussion of Anthropometry

Eyewitness testimony can often be the principal evidence resulting in a conviction in court cases, many times condemning an innocent person [102]. Wells et al. [102] state that "…eyewitness identification evidence is among the least reliable forms of evidence and yet persuasive to juries." The psychology of facial recognition has been studied for years by researchers and has shown that accurate recognition is more likely to result when the person is familiar, even from low quality images [85, 86]. However, there is a high error rate when trying to recognize unfamiliar faces [106]. Experiments conducted by Kemp, Towell and Pike [106], show that cashiers performed poorly in determining whether customers, who were standing right in front of them, matched the photograph on their credit card. The exactness of psychological facial matching is also decreased with changes in viewpoint, but is less affected by changes in facial expression [85].

Because mistakes can easily be made, eyewitness testimony should not be the sole evidence provided in court. A more scientific method of identification is needed to prove guilt or innocence in the eyes of the judicial system. A comparison of two dimensional images using facial anthropometry, one such method, has not been thoroughly tested but as 2D images are the most readily available, priority should be given to this. Proportions, rather than absolute measurements, were utilized in this research allowing the subject in the photographs to be of different sizes. However, the two photographs being compared must be taken from the same angle and viewpoint. The advantage of measuring photographs to that of a living person is that it is then possible to return to the photographs if there is any question about them in the future. It is also feasible for several operators to complete the analysis in order to get differing opinions or to confirm/refute identification.

Although rapidly improving, in terms of spatial resolution, the majority of video surveillance equipment does not produce the clear images needed to provide irrefutable identifications, when other more conclusive evidence such as DNA or fingerprints is not available. It is in these kinds of cases that anthropometry would be a useful identification technique. Surveillance video tape can be important supportive evidence to have because it may show a crime being committed, although, it is not always easy to recognize and therefore convict a criminal caught on CCTV. Video surveillance may be more reliable than eyewitness testimony because the story told is consistent as well as corroborating what the eyewitness reported [135]. However, a more comprehensive analysis is necessary because even when video images are clear, it is possible that two people may look similar to each other in addition to the ease in which individuals often disguise themselves. To aid in the identification of disguised individuals, research was conducted in the comparison of 2D with 3D images [61]. Anthropometric landmarks found to be significant in that study were superaurale and otobasion inferius, as well as the outline of the ear helix.

7.1 Probability

Once all evidence is gathered and a case goes to trial, the ultimate question the jury is asked to answer is regarding the guilt or innocence of the defendant. It is up to both the prosecution and the defence to present evidence to the jury to help them answer this question. For the forensic scientists involved in the interpretation of evidence, there is a degree of uncertainty associated. According to Aitken and Stoney [162], "How can uncertainty be measured? By probability."

Concepts of evidence and probability are nothing new and were incorporated during the period *Bertillonage* was in use. The very basis of *Bertillonage* was that, over the entire range of measurements taken, it would be unlikely that two people would have the same ones [163]. However, there were several problems involved in the rationalities and taking of measurements during this time. One of these was that as the database grew, the probability of two (or more) people having the same set of measurements was unknown. The greatest restriction of the system at the time was that it was for comparative use only and therefore could only tell if it were possible that the person could be the same as the one who had been arrested at an earlier time but could not prove that a specific person was at a particular place.

There are three laws of probability [162-165]. The first law of probability concerns one event, the second law is used for events that are mutually exclusive and the third law of probability examines the concurrence of two events. For an event which is known to be impossible, the probability is zero. Conversely, for an event that is absolute, the probability is one.

The following are questions to consider when evaluating crime scene surveillance video evidence and probability.

> What is the probability that the suspect is at the scene of the crime at the time the crime happened?
>
> What is the probability that the suspect was wearing the clothes that can be seen on the surveillance video?
>
> What is the probability that the suspect will be matched by other forensic evidence found at the crime scene?
>
> What is the probability that the suspect has the same facial proportions as the person shown on the surveillance video? Or reversely, what is the probability that the suspect does not have the same facial proportions as anyone else in the population?

For these questions to be answered, the likelihood of an event must be measured by referring to the relevant population in which the event is likely to have occurred [162]. An important interpretation of forensic evidence is rendered by the Bayes' Theorem. "Bayes Theorem always takes the same basic form: prior probabilities are converted to posterior probabilities by multiplication in proportion to likelihoods [166]" and states that for two events, P and R,

$$\Pr(S|R) = \frac{\Pr(R|S) * \Pr(S)}{\Pr(R)}$$

Understanding Bayes' Theorem is key to the interpretation of forensic evidence and in this case anthropometric facial image comparison because it defines what the odds are that the facial proportions of a suspect correlates to the facial proportions of anyone else in the population. To determine these odds, extensive studies would need to be conducted into the differences in facial proportions among those of a similar and dissimilar physiognomy, and those of different genders and races. The case of Jean Charles De Menezes depicts a good example Bayes' Theorem. Jean Charles De Menezes was shot and killed by police based upon a description of a suspected suicide bomber after the London bombings in 2005 [167]. Had the police considered Bayes' Theorem, they may have realized that the probability of someone else in the population fitting the description of the person they were looking for (false positive) were greater than the probability of Jean Charles De Menezes being that person (true positive).

7.2 Landmark Placement

Once the selected images for comparison have been chosen, the first step in attempting to make an identification though a facial anthropometric comparison is to place the facial landmarks in their appropriate locations. Therefore it is imperative to use a set of landmarks that are easily and accurately placed, discriminate between individuals and that will be available on the majority of photographs. A landmark that is accurately placed and discriminating between individuals is of not much use if it is located in a place that is commonly disguised within the course of a crime.

Fieller reports in his lecture notes that an initial study in 2002 showed that the variation between faces to be sufficiently greater than the variability between different operators measuring shapes from the same photo and that of the variability within the same operator making measurements on the same photograph [147]. He also reported that the corner of the jaw was not easily identifiable but was variable between people, while the corner of the mouth was easily located but less variable between people as well as depended on the expression given [147]. While, research conducted by Purkait, who took direct measurements of the face, cautioned use of the exocanthion in comparisons, because any minor change in facial expression resulted in inconsistent locations [148]. Although in the research conducted for this thesis, facial expression was not a factor, as both the video images and photographs had the same neutral expression, movement of the ectocanthion in multiple 2D images is one of the least likely landmarks to be affected by facial expression unless the person is squinting and was found to be a factor in Purkait's study because she was measuring live people who may become weary during the measurement process.

To place the landmarks necessary to conduct the analyses in this thesis, the second version of the computer programme, Facial Identification Centre Version 0.32, was greatly improved from the original version, 0.01 used in the pilot study. The ability to import a standardized list of metrics (landmarks, lines, proportions) for measurement and then export the results into an Excel® file saved countless hours of inputting the data manually. It also created a greater accuracy of the data by removing the possibility of errors caused by an inaccurate transcription of data. It is recognized that this programme is very basic but it could be improved by allowing the adjustment of certain image parameters in order to test the size and quality of the image with the landmarks once it is loaded into the computer programme. As it stands now, no changes are able to be made to the image once it is imported into the program file.

Justification of the chosen landmarks for this research was explained in Chapter 3 and will not be repeated here. The remainder of this section in the discussion will focus on the experience and observations of the landmark placement process on 2D images experienced by the author. In this study, the biggest influence on the variation in landmark placement was a result of the experience of the operators. By the nature of the process, variation of landmark location will increase with the inclusion of multiple operators to place landmarks rather than a single operator but this is necessary to replicate the effect that different operators in multiple police departments would have on the error. The variation of landmark placement was also increased when using inexperienced operators to perform the task. Any variation detected in landmark location amongst different operators will become magnified when involving images where it is difficult to detect the location of the landmarks. Landmarks are difficult to locate on photographs that are small, blurry, or grainy. Landmarks located on bony surfaces, such as the gonion, are difficult to place on 2D images in the majority of people whose faces are not chiselled.

A hypothetical solution to the problem of large variations within landmark placement would be to establish a single centre where police departments across Britain would send their photographs to a handful of trained operators who would place the landmarks and record the data in a centralized database alongside other identifying information. Initially the system could be conducted on police identification photographs taken by the FIND program. Therefore, the photographs would be standardized and as a small group of operators would be placing landmarks, the error in facial measurements caused by both landmark placement and rotational variation would be minimized. Although variation would still occur with the use of multiple operators, the operators would be highly experienced following a standardized training program, ideally keeping the error to a minimum. In cases where an anthropometric identification method may be beneficial, any police department across the country could then send their available crime footage images to this centralized centre for their expertise. This would ensure that the video image analysis, including landmark placement, was completed by the same group of individuals, lowering the variation between landmark placement in the original database photographs and then any suspect images submitted. Of course before any of that could be implemented, anthropometry as a method of identification would need to be validated as obtaining accurate comparisons between images.

Variability in landmark placement could depend on the operator's mood and degree of tiredness or on how many comparisons had been done that day. Variability in measurements was noted in *Bertillonage* when measurements of one individual taken by several police officers varied because of the difficulty experienced in reading minute graduations on the callipers [18]. Although a computer program was used for landmark placement in the research undertaken in this project, the choice of where to place the individual landmarks was made by the operator. In order to remove the subjectivity and variability caused by humans placing landmarks, a more scientific and computer orientated method of placing landmarks would be advantageous. A three-dimensional digitizer has been tested on its ability to detect landmarks on the faces of living people and a stone cast [168]. It was shown that the digitizer could determine landmark positions with sufficient precision. While this system may not be useful to compare photographic images it does illustrate that there has been research into different ways of detecting landmarks for measurement purposes.

7.3 Other Research into Anthropometric Analysis

A summary of the anthropometric comparison method carried out in this thesis compared to that of other research found in the literature is laid out in Table 7.1. With the exception of Halberstein, the majority of the researchers use absolute measurements as opposed to that of proportions. One advantage of using direct linear measurements between landmarks is that fewer landmarks are required to make up a line rather than a proportion, which can be a consequence if considering camera angles or disguises. Also as fewer landmarks are required, the error produced when locating the landmarks is less. The disadvantage to using absolute measurements is that the size of the two or more images being compared must be equalized which can be difficult if camera parameters are unknown. If direct measurements were to be utilized, it would be necessary to implement a Procrustes analysis to aid in the comparison.

For images on print or digital media, the availability of a computer to directly calculate distances or proportions is advantageous to record a more precise measurement product, although this is not an option when measuring a living person and high quality callipers should be used as was reportedly done by Halberstein, Farkas and Purkait. In cases where the suspect offers to be anthropometrically measured in person, 3D imaging of the suspect should be employed to further combine the advantage of digital media with the presentation of a living person.

All researchers listed in Table 7.1 utilized landmarks located on the face and in addition, Halberstein used proportions derived from parts of the body. The addition of multiple areas from the body can potentially strengthen the identification but should also be tested for their statistical difference amongst the population. No other identification methods were used in combination with the anthropometric comparisons in this research; however, a morphological analysis and superimposition were used by some of the researchers in addition to their anthropometric technique. Any available methods used in conjunction with an anthropometrical analysis should be used as supplementary evidence as it strengthens the conclusion of the identification.

High resolution images were used as research material because it was thought that if the technique did not work on high quality images, then it would have little chance of working on typical surveillance camera images, which can be blurry or grainy. The facial positions of these images were taken in a controlled setting, with both video images and photographs positioned as close as possible to a frontal view. In order to achieve the best possible results, video images compared against photographs must have facial positions and facial expressions that resemble each other. In an attempt to compare images with differing positions, Yoshino et al [61] developed a 3D physiognomic range finder which adjusted a 3D image to match the orientation and size of the 2D image. This incorporated a morphological comparison, an anthropometric analysis, and morphometric matching combined to achieve facial identification.

Table 7.1 Summary of author's facial anthropometric comparison research compared to that of other research reported in the literature.**

	Proportions or absolute measurements	Method	2D/3D/ Live person	Landmarks used	Scaling	Additional ID methods used	Findings
Kleinberg [169]**	**Proportions**	**Computer program, operator places landmarks with mouse**	**2D / 2D**	**Face**	**Used proportions so no need to enlarge photo**	**None**	**Not applied to casework, poor identification results**
Halberstein [38]	Proportions	Callipers	2D / Live	Head/face/ body	Used proportions so no need to enlarge photo	morphology	Used in 3 cases. 2 guilty, 1 non-guilty
Yoshino [61]	Absolute measurements	Computer	2D / 3D	Face	Converted original 3-D measurement data into the number of pixels in display	Superimposition to help line up 2-D and 3-D images	3-D image rotated to match facial position of 2-D image
Porter [37]	Absolute measurements	Digital callipers	2D / 2D	Face	Enlarged both photographs	morphology	Avoid measurements on the vertical axis due to distortion
Farkas [27]	Absolute measurements	Callipers. Marked landmarks on person before taking photos	2D / Live	Face	Enlarged photo to life size	None	Nasion and stomion good points b/c on same plane Frontal view supplied best results for orbits, lips, mouth
Purkait [148]	Absolute measurements	Callipers. Marked landmarks on person before taking photos	2D / Live	Face. Approx same plane	Enlarged photo to life size	None	Indirect measurements consistently larger due to flattening of photo with exception of measurements around eyes and mouth consistently larger in direct meas.
Rogers [23]	Absolute measurements	Article does not say	Bust / corpse	Face	Assumed bust was life size due to artist reputation	Morphology, superimposition	Identified corpse of John Paul Jones

A study published by Weinberg et al [170] compared the anthropometric precision and accuracy of two different digital 3D photogrammetry systems (Genex and 3dMD), as well as comparing the two systems against direct anthropometry. The results indicated that the mean differences of the three methods were too small to be of importance and therefore 3D image systems would be of use for landmark placement. This has important implications on future research and may help in overcoming the problem of facial orientation. One reason that their direct anthropometric results were at least equally precise to the 3D image measurements was that they had conducted their direct measurements on inanimate objects and there is a certain amount of measurement error to be expected when measuring living subjects. This would not be an issue when comparing photographs to video images as they are both print media.

As well as comparing photographs positioned at the same angle, it is also beneficial to compare two photographs in which only a short amount of time has elapsed between them, because age and weight changes can affect landmark position and therefore, outcome [28]. Other conditions that can affect the comparison are photographs taken under different lighting conditions or at different distances between the individual and the camera [28].

7.4 Conclusions

On completion of the comparisons and analysis of the results in Chapter 6, it was concluded that facial anthropometry has little value in producing positive identifications between video and photographic images. Results from this project show that the use of the selected landmarks and proportions in this type of comparison is not accurate enough to gather the results necessary for a court of law. It appears from this research that distances between facial features alone are not enough to distinguish between individuals. It is possible that there is a small measurable difference between what someone looks like and the variables of the situation when comparing two different photographs. In other words, the proportion values between different individuals is not a big enough difference to offset the variables and errors resulting from comparing two photographs.

An integral question to be asked before conducting a comparison between surveillance video and photographs was whether or not photographs taken from similar angles can be compared accurately. Based on the data from this analysis, the answer is no. The concept of using anthropometry as a tool to compare 2D images in forensic facial identification is one with intuitive appeal. However, in reality there are too many variables involved to guarantee consistent results. Other researchers have been able to apply various methods of anthropometric comparison to cases resulting in convictions [37, 38]. Nevertheless in this study, the proportions derived from linear measurements between the chosen landmarks have shown not to be useful in securing identifications. It appears from this research as though distances between facial features alone are not enough to distinguish between individuals. Several factors contribute to unease about the future use of comparing 2D images using facial anthropometry, including the quality of video images produced from surveillance systems. Lighting on the video may effect the images produced and may create shadows which obscure facial landmarks. Blurriness from motion also plays a role to generating poor quality images. The position of the head in both images being compared must be the same, and often the position of the surveillance camera does not allow for this. It is difficult to precisely match the angle in which the head is tilted in the photographs and video images because the human head can move in so many different ways, both horizontally, vertically and diagonally. When the whole body is in movement, as in a video, different angles are going to be recorded compared to when an individual is photographed sitting or resting. It is difficult to compare images with different camera parameters especially if the parameters of one or both cameras are unknown. And finally, even though there are guidelines for exact landmark positions, there is still a level of subjectivity to placing the landmarks. Operator experience and judgement is a contributing factor to the success of landmark placement.

A significant problem with distinguishing between two 2D images is that the face is a complex 3D structure and any comparison should therefore keep that in mind. Future research should concentrate on creating a 3D reconstruction of the suspect from their police identity photograph and this 3D image could then be matched to the facial position of the individual in the video image. Creating a 3D image should demonstrate a unique fit factor and superimposing a 2D image (individual on video) onto a 3D image (suspect photograph) will clearly exhibit the distribution of the fitting error vs. the pose angle and individuality of the person. It is possible to extract 3D landmarks from video by matching two or more frames and applying photogrammetry for a fixed camera position [171].

A sizeable amount of research has been done into creating a 3D facial image [152, 153, 172, 173] for many purposes. One of these is to compare a 3D image with a 2D image [147] of the known individual in order to manipulate the head positioning to mirror that of the unknown individual [174]. Researchers in the Netherlands compared a 2D image with that of a 3D laser scan image with discouraging results [174]. To prepare for the comparison, the camera position and orientation were estimated while keeping the focal length and pixel ratio fixed and the 3D image was then projected onto the 2D image and manipulated so that the head positioning matched. The locations of the X and Y coordinates of a maximum of seven landmarks were used and compared in both images. Their set of comparisons was small and combined with the fact that they only used seven landmarks could have influenced their poor outcome. In reality, it may not be feasible to take a 3D image scan of a possible suspect as the subject will probably not agree, but using photographs taken of the suspect at several different angles to create a 3D image of the suspect would be possible [152, 153].

A system also exists that automatically creates a 3D face using a database of the anthropometric measurement statistics of a population. Once a group of random measurements have been chosen, the system constructs the face by finding the best surface that fits into the geometric parameters set by the measurements [172]. The end product, as shown in the published article, is a face that looks cartoon-like but if a more lifelike version were created, theoretically the system could be applied to forensic work.

The effect of distortion on photographs may affect the ability to compare images taken at different times with different camera parameters. The focal length of the camera lens and the subject distance from the camera are factors which contribute to distortion, however, lens distortion can be corrected by calibration. Farkas found that the greatest effect from distortion was shortening of the upper third of the face and that one reason the nasion and stomion landmarks proved to be accurate was because they were on the same focusing plane [27].

The camera parameters for both sets of images in which the present study was conducted were unknown. The worst case scenario in a photogrammetric analysis is when nothing is known about the cameras which captured the images. In this type of circumstance there is a strong danger of generating any comparison to ‘fit’, whereas compelling comparisons can be made from calibrated images. In practice it is likely that camera parameters will be known for at least one set of images. Lee et al. found it necessary to be aware of the distortion parameter of a camera lens in addition to obtaining a sufficient number of calibration points in order to effectively measure the height of a person standing in a fixed location [175]. The advantage given to the comparison of images taken for the FIND program is that the suspect identity photographs taken by police departments will be standardized and camera parameters clearly stated. Camera parameters of video images obtained from surveillance cameras can be acquired using basic photogrammetry skills. In an ideal comparison, nothing would deviate between the two cameras parameters, or at most only the focal length would be varied.

A significant difference between the facial proportion feature vector of the unknown individual and facial proportion feature vector from the database photograph could lead to the belief that either the two people were different or that the angle position of the faces in the two photographs was different. If further research led to the adoption of this technique for police cases, it would be important for the operator to understand that not all scenarios would be appropriate material to work with and it would be important to be able to tell the difference between a situation where it is possible to compare photographs and one in which it is not. Inappropriate material would include images that are too degraded or of too low a quality, or would involve the comparison of two images where the camera angles are too varied with no means to correct one image position to the other.

Any thoroughly tested and validated method of identification could only improve the administration of justice. However the current method of utilizing anthropometric proportions to compare faces from 2D images tested in this thesis is not yet at the stage where it would benefit the judicial system. It is preferable to error on the side of caution and potentially miss a potential suspect based upon facial measurements than begin to commence a miscarriage of justice against the individual. In cases where someone is eliminated falsely based upon the anthropometric analysis carried out, ideally additional evidence found at the crime scene would provide the intelligence necessary to include the individual in the suspect list. The basic principle of matching evidence from an unknown suspect to a database of known individuals, as the methods in Chapter 6 were designed, could be used for police casework as investigations are likely to be conducted in the same manner. Using this identification method in the elimination of a wide profusion of suspects may also provide benefits to police cases. In cases where additional evidence has already eliminated a pool of suspects, attempts could be made to use this identification method to provide additional supporting evidence in the defence or prosecution of the individual.

8 Final Conclusions and Suggestions for Further Work

The possibility of using anthropometry for criminal identification has been studied and analyzed since the late 19th century, starting with the contributions of Bertillon. There are individuals who recognize the important contributions that Bertillon made and hope endures that any old anthropometric files that were created based on Bertillon's method were not destroyed and may be made available for further use [176].

The studies that were completed for this thesis were intended to add to the research that Bertillon started with the intention of creating an accurate form of identification for use in forensic cases. The most extensive study involved anthropometry and was carried out to validate a method to discriminate between subjects when the identity of one was known and the other was unknown and so to determine if they were the same individual.

In addition to, or in lieu of, completing an anthropological analysis, a morphological analysis should be carried out whenever possible. However, when images are compared on the basis of descriptions obtained from a checklist, they may be too subjective to achieve accurate identifications.

On the basis of the work carried out it has not been possible to disprove the null hypothesis as set out in my objectives, i.e. "It is not possible by utilizing a series of facial anthropometric measurements to be able to satisfactorily discriminate between individual subjects when comparing images of a known subject with those of a subject whose identity is not known in order to assess whether they are the same individual."

8.1 A Retrospective Overview of the Project

A pool of facial landmarks was selected after consulting the literature to address the problem of the low discriminatory power obtained using the initial four landmarks. A total of 37 landmarks and 59 proportions were chosen to be used for further research. The increased number of landmarks and proportions provided a more thorough representation of the face than was obtained in the initial pilot study.

Despite the problems encountered in the initial pilot study, the results demonstrated a consistent change in values between each 10° step rotation for each person. The graphs illustrated that each subject followed the same type of curve, with respect to both angles and proportions, indicating a predictable change between each 10º rotation. Given this information, it may be possible to estimate from which angle a photograph is taken and, as a corollary, it may be possible to calculate correction factors to normalise proportions to their full-face values. If it is possible computationally to correct the distorted images from the Hubble space telescope, it should be possible to improve anthropometric measurements by application of mathematics.

A second laboratory study was undertaken to address specific experimental design problems encountered in the initial pilot laboratory study. What was different about this study was that the photographs were taken in a fixed and closely monitored setting in an attempt to minimize variables which could be controlled, such as the head pose. The photographs were also taken by multiple operators, much in the way that photographs are obtained by the police, to determine the effect this had on proportion values.

Intra and inter-operator error studies were conduced on high resolution images to assess the process of landmark placement and quantify the effect caused by inexperienced operators. The combination of errors in facial measurements from landmark placement and photography contributes to the overall uncertainty of the method.

Utilizing the 59 proportions as a multivariate feature vector, a study was conducted to compare two samples of 2D images - video images and database photographs. Research material consisting of high resolution still video images positioned in the frontal view was made available, which proved to be easily comparable against a database of frontal view photographs, as the faces were in the same pose. Results demonstrated that using the combination of ‘best case scenario’ images in addition to a single experienced operator placing landmarks on all images achieved at best a decrease to five database images to be verified by a human in approximately 75% of the cases tested.

Any discrimination power the method has is greatly diminished when variables are factored in that occur with images taken at different times with different cameras capturing different poses and expressions in addition to different operators conducting landmarks placement. When the operator is required to place the landmarks manually, as with the computer program used in these experiments, it is a subjective placement based on the individual's own assessment of where the landmark should be placed. Variability among different operators is to be expected but if the operator is not thoroughly trained, the potential is created for inaccurate placements, directly affecting any results obtained. One possible solution to the problem of variability in landmark placement amongst operators is to create a centralized department in which a small group of thoroughly trained operators are responsible for placing landmarks on images for the entire United Kingdom, keeping the data in a centralized biometric database. The United States of America Federal Bureau of Investigation is currently in the process of developing their own vast biometric database which could be used as model [177].

The research materials used were the best possible images available, offering the operator comparable images with the same facial position. These images were also useful because, as the video images and photographs were taken on the same day, there were no possible facial changes due to age or weight. However, even with these high quality images taken in a controlled environment, it still proved impossible to make identifications from the measurements and proportions obtained.

When conducting an anthropometrical image analysis, the images chosen for comparison are just as important as using a trained operator. The position of the head is important when determining which images are to be compared. Both facial proportions and morphological descriptions of facial features could change based on the position of the head so it is important that the position of the head is the same in both the image from the surveillance video and that of the photograph. Promising research into image averaging [178] could be the answer to the fundamental issues of differing poses, lighting, expressions and camera parameters that are problematic for achieving accurate facial recognition.

8.2 Future Research

From the initial pilot study in this thesis, it is clear that in order to attempt to predict facial position with an acceptable degree of accuracy, a much larger reference group of subjects categorized by sex, ethnicity, and facial build would need to be studied. The small study demonstrated a predictable variability in measurement outcomes (angles and proportions) with facial position using a limited number of landmarks. The fundamental question which needs to be addressed is whether anthropometry is sufficiently discriminatory between individuals. This is particularly important when analysing data of subjects with a similar physiognomy, to enable such a technique to be used as corroborative evidence or even for a positive identification. Further studies need to examine to what extent accurate comparisons can be made of photographs that are from similar viewpoints in relation to a camera, and to find what the acceptable range of positional differences between images is which allows useful conclusions to be drawn.

Further research into creating 3D images taken from surveillance cameras may help to correct potential problems with orientation. Before any two images can be compared, they must be from similar viewpoints and it can be quite difficult to accomplish this. However, if there was a way to manipulate the image so that is was in the same position as the image it is being compared to, there might be a greater chance for accurate identification.

When both laboratory studies in this thesis had been completed and further research using anthropometry as a type of image comparison in human identification had been carried out, several avenues that future research might explore became apparent. Now that this study has compared video images to the photographs of similar looking individuals, much in the way a traditional police line-up operates, the method could be tested against a random sample to determine the differences in measurements and what if any effect this has on identification.

Anthropometry has been used successfully in some instances [37, 38] and, although different methods were used, there may be some benefit to conducting further research. Research could delve into specific reasons affecting the accuracy of the technique. Experiments to test camera distance or lighting distortion, and individual variables dealing with surveillance systems could be explored to determine if any one affects results more or less than others. Findings from this could influence the way security systems are set up in the future.

The data obtained from the anthropometry study demonstrated that an identification based on a two dimensional image is weak, even given best case scenario materials. Given additional time to further investigate the comparison of anthropometric facial proportions, a detailed consideration of photogrammetry would be advisable. Using a desktop photogrammetry package, a 3D representation of the face could be developed and used to find the homographies between two images. A 3D image will also allow a fitting error to be determined, allowing the question “Is there enough information available in the image to allow a comparison to be made?” to be answered. A comparison with a 3D image as its foundation launches the process with stronger data than if just two 2D images were available.

Instead of using anthropometry to attempt probable identification, research into using the technique to decrease the number of individuals (or as a process of elimination) from a larger pool of suspects could be helpful. As the anthropometric identifications have been shown to be of limited value, and the small morphological research undertaken also showed limited value, it may be useful to investigate how it would be to use the two methods in conjunction.

In contrast, it may be more beneficial to the field of facial identification to put on hold research into anthropometric comparisons of 2D images and instead focus on related fields. There has been encouraging work done in the field of biometrics and although not yet a perfect method of identification, it is a very promising one. Studying the facial morphology of two images and conducting a direct comparison is being done currently in real life identifications and should not be ignored. Finally, and perhaps most importantly, research into improving the images produced from surveillance camera security systems would be of great benefit to the field of facial image comparison.

8.3 Final Words

Although the findings of this research were not the expected or hoped for outcomes, nonetheless they provide useful information for the research community. If anthropometry is to be taken seriously as a method of identifying individuals and used as evidence in court then it needs to have some serious research to support it. The purpose of these studies was to test the science behind a method of identification that has been used with varying results and is not yet commonly accepted in the field of facial identification. Hopefully this contribution will help convince researchers interested in photographic anthropometry of 2D images that although the studies tested show there may be potential to narrow down a list of suspects from a database, it is not at the stage where it can be used to make a direct identification with the aim to convict in court.

References

[1] Hanks P, ed. The Collins English Dictionary. Second ed. Glasgow: William Collins Sons & Co. Ltd 1986.

[2] Frankfort horizontal plane. (n.d.). Merriam-Webster's Medical Dictionary Retrieved, from Dictionarycom website [cited December 19, 2006]; Available from: http://dictionary.reference.com/browse/frankforthorizontalplane

[3] Friel JP, ed. Dorland's Illustrated Medical Dictionary. 25th ed. Philadelphia: W.B. Saunders Company 1974.

[4] Shepherd R. Simpson's forensic medicine. 12th Edition ed. London: Arnold 2003:49-51.

[5] Ubelaker DH. Human skeletal remains: Excavation, analysis, interpretation. 2nd Edition ed. Washington D.C.: Taraxacum 1989:118-9,28.

[6] Primorac D, Andelinovic S, Definis-Gojanovic M, Drmic I, Rezic B, Baden MM, et al. Identification of war victims from mass graves in Croatia, Bosnia, and Herzegovina by the use of standard forensic methods and DNA typing. J Forensic Sci. 1996;41(5):891-4.

[7] Randerson J. Srebrenica victims named at last. New Scientist. 2003;177(2385):6.

[8] Vastag B. Out of tragedy, identification innovation. JAMA. 2002;288(10):1221-23.

[9] Hollon T. Reforming criminal law, exposing junk forensic science. The Scientist. 2001;15(17):12.

[10] Elliott K, Hill DS, Lambert C, Burroughes TR, Gill P. Use of laser microdissection greatly improves the recovery of DNA from sperm on microscope slides. Forensic Sci Int. 2003;137:28-36.

[11] Balding DJ, Donally P. Evaluating DNA profile evidence when the suspect is identified through a database search. J Forensic Sci. 1996;41(4):603-07.

[12] Williamson R, Duncan R. DNA testing for all. Nature. 2002;418:585-6.

[13] Wilson C. Written in blood: A history of forensic detection. Wellingborough, North Hamptonshire: Equation 1989.

[14] Rudnick SA. The identification of a murder victim using a comparison of the postmortem and antemortem dental records. J Forensic Sci, JFSCA. 1984 Jan 29(1):349-54.

[15] News. Computer-assisted facial identification. Significance. 2004;1(1):2-5.

[16] Evison MP, Bruegge RWV. The magna database: A database of three-dimensional facial images for research in human identification and recognition. *Forensic Sci Com* April 2008.

[17] Cole SA. Suspect identities: A history of fingerprinting and criminal identification. Cambridge, Massachusetts: Harvard University Press 2002:32-59, 140-6.

[18] Moenssens AA. Fingerprint techniques. Philadelphia: Chilton Book Company; 1971.

[19] Gould SJ. The mismeasure of man. Toronto: George J. McLeod 1981:77-9.

[20] Sabbatini RME. Phrenology, the history of brain localization. *Brain & Mind* 1997.

[21] Engber D. Deadly medicine: Creating the master race. Br J Med. 2004 August 28;329(7464):517.

[22] Taylor D, Myers WC, Robbins L, Barnard GW. An anthropometric study of pedophiles and rapists. J Forensic Sci. 1993;38(4):765-8.

[23] Rogers NL, Field K, Froede RC, Towne B. The belated autopsy and identification of an eighteenth century naval hero-The saga of John Paul Jones. J Forensic Sci. 2004;49(5):1036-49.

[24] Bromby MC. At face value? New Law J Expert Witness Supplement. 2003 28 February;153(7069):302-4.

[25] Crim. L.R. 1999, SEP. 1999:750-1.

[26] Vanezis M, Vanezis P. Cranio-facial reconstruction in forensic identification- Historical development and a review of current practice. Med Sci Law. 2000;40(3):197-205.

[27] Farkas LG. Anthropometry of the head and face. Second edition ed. New York: Raven Press, Ltd. 1994.

[28] Ïşcan M. Introduction of techniques for photographic comparison: potential and problems. In: Ïşcan M, Helmer R, eds. *Forensic analysis of the skull*. New York: Wiley-Liss 1993:57-70.

[29] Slama CC, ed. Manual of Photogrammetry. 4th ed. Falls Church: American Society of Photogrammetry 1980.

[30] Sharp HO. Practical Photogrammetry. New York: The Macmillan Company 1951.

[31] Forsyth DA. Powerpoint slide sets for "Computer Vision: A Modern Approach". [cited 2007 23 August]; Available from: http://www.cs.berkeley.edu/~daf/bookpages/slides.html

[32] Perspective. 2007 Encyclopædia Britannica [cited 2007 18 September]; Available from: http://www.search.eb.com/eb/article-9059357

[33] Mikhail EM, Bethel JS, McGlone JC. Introduction to modern photogrammetry. New York: John Wiley & Sons, Inc. 2001.

[34] McGlone JC, Mikhail EM, Bethel J, eds. Manual of Photogrammetry 5th ed. Bethesda: American Society for Photogrammetry and Remote Sensing 2004.

[35] Vanezis P, Lu D, Cockburn J, Gonzalez A, McCombe G, Trujillo O, et al. Morphological classification of facial feature in adult Caucasian males based on an assessment of photographs of 50 subjects. J Forensic Sci. 1996;41(5):786-91.

[36] Catterick T. Facial measurements as an aid to recognition. Forensic Sci Int. 1992;56:23-7.

[37] Porter G, Doran G. An anatomical and photographic technique for forensic facial identification. Forensic Sci Int. 2000;114:97-105.

[38] Halberstein RA. The application of anthropometric indices in forensic photography: Three case studies. J Forensic Sci. 2001;46(6):1438-41.

[39] SWGIT. Best practices for forensic image analysis. *Forensic Sci Com* October 2005.

[40] Mardia KV, Coombes A, Kirkbride J, Linney A, Bowie JL. On statistical problems with face identification from photographs. J Appl Stat. 1996;23(6):655-75.

[41] Alessandrini F, Cecati M, Pesaresi M, Turchi C, Carle F, Tagliabracci A. Fingerprints as evidence for a genetic profile: Morphological study on fingerprints and analysis of exogenous and individual factors affecting DNA typing. J Forensic Sci. 2003;48(3):586-92.

[42] Kolltveit KM, Solheim T, Kvaal SI. Methods of measuring morphological parameters in dental radiographs Comparison between image analysis and manual measurements. Forensic Sci Int. 1998;94:87-95.

[43] Swift B, Rutty GN. The human ear: It's role in forensic practice. J Forensic Sci. 2003;48(1):153-60.

[44] Rogers TL. Determining the sex of human remains through cranial morphology. J Forensic Sci. 2005;50(3):493-500.

[45] Yoshino M, Noguchi K, Atsuchi M, Kubota S, Imaizumi K, Thomas CD, et al. Individual identification of disguised faces by morphometrical matching. Forensic Sci Int. 2002;127:97-103.

[46] Donofrio LM. Fat distribution : A morphologic study of the aging face. Dermatol Surg. 2000;26(12):1107-12.

[47] Vanezis P, Blowes RW, Linney AD, Tan AC, Richards R, Neave R. Application of 3-D computer graphics for facial reconstruction and comparison with sculpting techniques. Forensic Sci Int. 1989;42:69-84.

[48] Jury hears of 'similarities'. *UK Newsquest Regional Press-This is Bradford* 2003.

[49] Lynnerup N, Vedel J. Person identification by gait analysis and photogrammetry. J Forensic Sci. 2005;50(1):112-8.

[50] Lee E, Whalen T. Computer image retrieval by features: Selecting the best facial features for suspect identification systems. Proceedings of the third international conference on Information and knowledge management; 1994; Gaithersburg, Maryland: ACM Press; 1994. p. 105-11.

[51] Lee E, Whalen T. Computer image retrieval by features: suspect identification. Proceedings of the SIGCHI conference on Human factors in computing systems 1993; Amsterdam, The Netherlands: ACM Press; 1993. p. 494-9.

[52] Reddy KSN. Identification of dismembered parts: The medicolegal aspects of the Nagaraju Case. Forensic Sci. 1973;2:351-74.

[53] Sognnaes RF. Hitler and Borman identifications compared by postmortem craniofacial and dental characteristics. Am J Forensic Med Pathol. 1980;1(2):105-15.

[54] Iten PX. Identification of skulls by video superimposition. J Forensic Sci. 1987;32(1):173-88.

[55] Glaister J, Brash JC. Medico-legal aspects of the Ruxton case. Edinburgh: Livingston 1937:144-70.

[56] Helmer R, Grűner O. Vereinfachte Scha¨delidentifizierung nach dem Superprojektions verfahren mit Hilfe einer Video-Anlage. Zeitschrift fur Rechtsmedizin. 1977;80(3):183-7.

[57] Ghosh AK, Sinha P. An economised craniofacial identification system. Forensic Sci Int. 2001;117:109-19.

[58] Solla HE, Ïşcan MY. Case report: Skeletal remains of Dr. Eugenio Antonio Berríos Sagredo. Forensic Sci Int. 2001 116:201-11.

[59] Goza WM. William R. Maples, forensic historian: Four men, four centuries, four countries. J Forensic Sci. 1999;44(4):692-4.

[60] Yoshino M, Matsuda H, Kubota S, Imaizumi K, Miyasaka S, Seta S. Computer-assisted skull identification system using video superimposition. Forensic Sci Int. 1997;90:231-44.

[61] Yoshino M, Matsuda H, Kubota S, Imaizumi K, Miyasaka S. Computer-assisted facial image identification system using a 3-D physiognomic range finder. Forensic Sci Int. 2000;109:225-37.

[62] Vanezis P, Brierley C. Facial image comparison of crime suspects using video superimposition. Sci Justice. 1996;36:27-34.

[63] Jayaprakash PT, Srinivasan GJ, Amravaneswaran MG. Cranio-facial morphanalysis: a new method for enhancing reliability while identifying skulls by photo superimposition. Forensic Sci Int. 2001;117:121-43.

[64] Austin D. Video superimposition at the C.A. Pound Laboratory 1987-1992. J Forensic Sci. 1999;44(4):695-9.

[65] Austin-Smith D, Maples WR. The reliability of skull/photograph superimposition in individual identification. J Forensic Sci. 1994;39(2):446-55.

[66] Shahrom AW, Vanezis P, Chapman RC, Gonzales A, Blenkinsop C, Rossi ML. Techniques in facial identification: computer-aided facial reconstruction using a laser scanner and video superimposition. Int J Legal Med. 1996;108:194-200.

[67] Bastiaan RJ, Dalitz GD, Woodward C. Video superimposition of skulls and photographic portraits- A new aid to identification. J Forensic Sci. 1986;31(4):1373-9.

[68] Koelmeyer TD. Videocamera superimposition and facial reconstruction as an aid to identification. Am J Forensic Med Pathol. 1982;3(1):45-8.

[69] Yoshino M, Imaizumi K, Miyasaka S, Seta S. Evaluation of anatomical consistency in cranio-facial superimposition images. Forensic Sci Int. 1995;74:125-34.

[70] Ubelaker DH, Bubniak E, O'Donnell G. Computer-assisted photographic superimposition. J Forensic Sci. 1992;37(3):750-62.

[71] Oxlee G. Evidence from imagery. New Law Journal: Expert Witness Supplement. 25 June 1993 June 25;143(6606):915.

[72] R (on the application of Taj) v Chief Immigration Officer, Midlands Enforcement Unit, Queen's Bench Division (Administration Court), CO/1084/99, 29 January 2001. 2001.

[73] R v Stubbs, Court of Appeal (Criminal Division), [2002] EWCA Crim 2254, 3 October 2002. 2002.

[74] In the Matter of an Application by Paul McFadden for Judicial Review, High Court of Justice in Northern Ireland, 12 March 2002. 2002.

[75] Helmer R. Identification of the cadaver remains of Josef Mengele J Forensic Sci. 1987;32(6):1622-44.

[76] Vanezis P, Vanezis M, McCombe G, Niblett T. Facial reconstruction using 3-D computer graphics. Forensic Sci Int. 2000;108:81-95.

[77] Taylor KT. Forensic art and illustration. Boca Raton: CRC Press LLC 2001.

[78] George RM. Anatomical and artistic guidelines for forensic facial reconstruction. In: İşcan MY, Helmer RP, eds. *Forensic Analysis of the Skull*. New York: Wiley-Liss, Inc. 1993:215-27.

[79] Wilkinson C, Neave R. Skull re-assembly and the implication for forensic facial reconstruction. Sci Justice. 2001;41(3):233-4.

[80] Iceman keeps scientists guessing. BBC News 2000 1 March 2008 [cited; Available from: http://news.bbc.co.uk/1/hi/sci/tech/1077816.stm

[81] Curse of the mummy? 2005 [cited 2006 December 7]; Available from: http://images.google.co.uk/imgres?imgurl=http://www.smh.com.au/ffximage/2005/04/21/mummy_wideweb__430x336.jpg&imgrefurl=http://www.smh.com.au/news/World/Curse-of-the-mummy/2005/04/21/1114028454815.html&h=336&w=430&sz=28&hl=en&start=3&tbnid=CoxIcmMclsO0IM:&tbnh=98&tbnw=126&prev=/images%3Fq%3DOetzi,%2Bthe%2BIceman%2Bphotographs%26svnum%3D10%26hl%3Den%26lr%3D%26rls%3DGGLD,GGLD:2004-13,GGLD:en%26sa%3DN

[82] Helmer RP, Roohricht S, Petersen D, Mohr F. Assessment of the reliability of facial reconstruction. In: İşcan MY, Helmer RP, eds. *Forensic Analysis of the Skull*. New York: Wiley-Liss, Inc. 1993:229-46.

[83] Haglund WD, Reay DT. Use of facial approximation techniques in identification of green river serial murder victims. Am J Forensic Med Pathol. 1991;12(2):132-42.

[84] Stephan CN, Henneberg M. Building faces from dry skulls: Are they recognized above chance rates? J Forensic Sci. 2001;43(3):432-40.

[85] Burton AM, Wilson S, Cowan M, Bruce V. Face recognition in poor-quality video: Evidence from security surveillance. Psychol Sci. 1999 May;10(3):243-8.

[86] Davies G, Thasen S. Closed-circuit television: How effective an identification aid? B Journal Psychol. 2000;91:411-26.

[87] Burton A, Miller P, Bruce V, Hancock P, Henderson Z. Human and automatic face recognition: a comparison across image formats. Vision Res. 2001;41:3185-95.

[88] Luus CAE, Wells GL. Eyewitness identification and the selection of distracters for lineups. Law Hum Behav. 1991;15(1):43-57.

[89] Malpass RS, Devine PG. Eyewitness identification: Lineup instructions and the absence of the offender. J Appl Psychol. 1981;66(4):482-9.

[90] Wells GL. The psychology of lineup identifications. J Appl Social Psychol. 1984a;14(2):89-103.

[91] Lindsay RCL, Wells GL. Improving eyewitness identifications from lineups: Simultaneous versus sequential lineup presentation. J Appl Psychol. 1985;70(3):556-64.

[92] Kloberdanz K. Building a better lineup. TIME. 2005 March 21:19.

[93] Pozzulo JD, Lindsay RCL. Identification accuracy of children versus adults: A meta-analysis. Law Hum Behav. 1998;22(5):549-70.

[94] Valentine T, Heaton P. An evaluation of the fairness of police line-ups and video identifications. App Cog Psychol. 1999;13:S59-S72.

[95] Bailenson JN. Using virtual heads for person identification: An empirical study comparing photographs to photogrammetrically-generated models. J Forensic Ident. 2003;53(6):722-8.

[96] Green F. Preventing wrongful convictions. Richmond Times-Dispatch. 2005 March 30;Sect. A1.

[97] Fleishman JJ, Buckley M, Klosinsky M, Smith N, Tuck B. Judged attractiveness in recognition memory of women's faces. Percept Mot Skills. 1976;43(3):709-10.

[98] Light LL, Kayra-Stuart F, Hollander S. Recognition memory for typical and unusual faces. J Exp Psychol Hum Learn Mem. 1979;5(3):221-8.

[99] Wright DB, Boyd CE, Tredoux CG. A field study of own-race bias in South Africa and England. Psychol Public Policy Law. 2001;7(1):119-33.

[100] Read JD, Vokey JR, Hammersley R. Changing photos of faces: Effects of exposure duration and photo similarity on recognition and the accuracy - confidence relationship. J Exp Psychol Learn Mem Cogn. 1990;16(5):870-82.

[101] Yarmey AD. Eyewitness recall and photo identification: A field experiment. Psychol Crime Law. 2004;10(1):53-68.

[102] Wells GL, Small M, Penrod S, Malpass RS, Fulero SM, Brinacombe CAE. Eyewitness identification procedures: Recommendations for lineups and photospreads. Law Hum Behav. 1998;22(6):1-39.

[103] Ellis HD. Recognizing Faces. Br J Psychol. 1975;66(4):409-26.

[104] Bruce V, Young A. Understanding face recognition. Br J Psychol. 1986;77:305-27.

[105] Bruce V. Influences of familiarity on the processing of faces. Perception. 1986;15:387-97.

[106] Kemp R, Towell N, Pike G. When seeing should not be believing: Photographs, credit cards and fraud. App Cog Psychol. 1997;11:211-22.

[107] Prabhakar S, Pankanti S, Jain AK. Biometric recognition: Security and privacy concerns. IEEE Security & Privacy. 2003 March/April:33-42.

[108] Jain AK, Ross A, Prabhakar S. An introduction to biometric recognition. IEEE Trans on Circuits and Systems for Video Technology, Special Issue on Image- and Video- Based Biometrics. August 2003.

[109] Pankanti S, Bolle RM, Jain A. Biometrics: The future of identification. Computer. 2000 February;33(2):46-9.

[110] Hadley C. Your personal passport. EMBO reports. 2004;5(2):124-6.

[111] Woodward JD. And now, the good side of facial profiling. The Washington Post. 2001 February 4;Sect. B.04.

[112] Graham-Rowe D. ID row bad news for transatlantic travellers. New Scientist. 2005 April 16(2495):23.

[113] Grotta SW. Bio-Keys - Fingerprint readers, retinal scanners, and facial recognition cameras are being used increasingly by businesses to keep intruders out of corporate networks. Is the future finally here? PC Magazine. 2001 June 5:163.

[114] Phillips PJ, Martin A, Wilson CL, Przybocki M. An introduction to evaluating biometric systems. Computer. 2000 February;33(2):56-63.

[115] Briggs B. Biometrics: Can we have a show of hands? Health Data Management. 2002;10(12):48.

[116] Finger chopped off to beat car security. New Scientist. 2005 April 9(2494):4.

[117] Peacock C, Goode A. Automatic forensic face recognition from digital images. Sci Justice. 2004;44(1):29-34.

[118] Lu X, Jain AK. Resampling for face recognition. Proc of 4th Int'l Conf on Audio- and Video- Based Biometric Person Authentication (AVBPA); 2003 June 9-11; Guildford, UK; 2003. p. 869-77.

[119] Lu X, Wang Y, Jain AK. Combining classifiers for face recognition. IEEE International Conference on Mulitmedia & Expo; 2003 July 6-9; Baltimore, MD; 2003. p. 13-6.

[120] Daugman J. Face and gesture recognition: Overview. IEEE Trans Pattern Analysis and Machine Intelligence. 1997;19(7):675-6.

[121] Biometrics use growing. R & D. July 2002;44(7):11.

[122] Pentland AS, Choudury T. Face recognition for smart environments. Computer. 2000;33(2):50-5.

[123] Tampa's facial recognition project a flop. United Press International. 2003 August 20;Sect. 1008232w1871.

[124] Hook P. Face to face identification. Police Magazine. 1998 October.

[125] Newton E, Sweeney L, Malin B. Preserving privacy by de-identifying facial images. Pittsburgh: Carnegie Mellon University, School of Computer Sciences, Technical Report, CMU-CS-03-119; 2003 March.

[126] Solomon CJ, Gibson SJ, Pallares-Bejarano A. EigenFIT - the generation of photographic-quality facial composites. Advancing Biometric Technologies: One Day British Machine Vision Association Symposium at the Royal Statistical Society; 2002 March 6; London, England; 2002.

[127] Huang FJ, Zhang H-J, Chen T, Zhou Z. Pose invariant face recognition. Fourth IEEE International Conference on Automatic Face and Gesture Recognition 2000 March 26-30; Grenoble, France; 2000.

[128] Chen Q, Cham W-k. 3D model based pose invariant face recognition from a single frontal view. Electronic Letters on Comp Vis Img Ana. 2007;6(1):13-26.

[129] Britain is 'surveillance society'. 2006 [cited 2007 18 September]; Available from: http://news.bbc.co.uk/1/hi/uk/6108496.stm

[130] Public Space CCTV. 2006 [cited 2007 18 September]; Available from: http://www.glasgow.gov.uk/en/Residents/YourCommunity/CommunityServices/publicspacecctv.htm

[131] Seattetimes.com. Man charged after missing girl's body found. [newspaper] 2004 [cited 2005 May 17]; Available from: http://archives.seattletimes.nwsource.com/cgi-bin/texis.cgi/web/vortex/display?slug=girl07&date=20040207&query=surveillance+video

[132] Edwards R. The camera lies. New Scientist. 1999 March 27(2179):27.

[133] Coleman R, Sim J. "You'll never walk alone": CCTV surveillance, order and neo-liberal rule in Liverpool city centre. Br J Socio. 2000;51(4):623-39.

[134] Foster M. New Orleans tries cameras to watch criminals. The Seattle Times. 2005 March 9.

[135] Lewis DL. Surveillance video in law enforcement. J Forensic Ident. 2004;54(5):547-59.

[136] Hogan J. Your every move will be analysed. New Scientist. 2003 12 July(2403):4.

[137] Aldridge J. Effective CCTV security and safety systems. CCTV Today. 1994;Feb:12-7.

[138] SWGIT. Recommendations and guidelines for using closed-circuit television security systems in commercial institutions: Part 1. *Forensic Sci Com* January 2005.

[139] Chmielewski T. Digital video surveillance: New visual technology means better protection for you. *Inside Self-Storage Magazine* 2001.

[140] Image Resolution, size and compression: What does it really mean? [cited 2007 May 8]; resolution]. Available from: http://www.microscope-microscope.org/imaging/image-resolution.htm

[141] SWGIT. Recommendations and guidelines for the use of digital image processing in the criminal justice system. *Forensic Sci Com* January 2003.

[142] Bowers CM, Johansen RJ. Photographic evidence protocol: The use of digital imaging methods to rectify angular distortion and create life size reproductions of bite mark evidence. J Forensic Sci. 2002;47(1):178-85.

[143] SWGIT. Best practices for documenting image enhancement. *Forensic Sci Com* July 2005.

[144] Hogg I. Chief Inspector, Strathclyde Police. In: Kleinberg K, ed. Glasgow: Personal communication 2007.

[145] Facial Images National Database (FIND). 2007 [cited 2007 July 11]; Available from: http://www.npia.police.uk/en/5967.htm

[146] Islam R. Police standard for still digital image capture and data interchange of facial/mugshot and scar, mark and tattoo images. 2007 [cited; Available from: http://www.npia.police.uk/en/docs/Capture_interchange_standard_Facial_SMT_images.pdf

[147] Fieller N. Statistical facial identification. 2006 [cited 2007 12 September]; Available from: http://nickfieller.staff.shef.ac.uk/seminars/faces04-10-06.pdf

[148] Purkait R. Anthropometric landmarks: How reliable are they? Anthropometric landmarks. Med Leg Update. 2004;4(4):133-40.

[149] Craw I, Costen N, Kato T, Akamatsu S. How should we represent faces for automatic recognition. IEEE Trans Pattern Analysis and Machine Intelligence. 1999 Aug 1999;21(8):725-35.

[150] Okada K, Malsburg Cvd, Akamatsu S. A pose-invariant face recognition system using linear pcmap model. In proceedings of IEICE Workshop of Human Information Processing; 1999; Okinawa; 1999. p. 7-12.

[151] Kolar JC, Salter EM. Craniofacial anthropometry Practical measurement of the head and face for clinical, surgical and research use. Springfield, Illinois: Charles C Thomas 1997.

[152] Siebert JP, Marshall SJ. Human body 3D imaging by speckle texture projection photogrammetry. Sensor Review. 2000;20(3):218-26.

[153] Urquhart CW, Siebert JP. Towards real-time dynamic close range photogrammetry. *SPIE Videometrics II*. Boston, USA 1993.

[154] Hirayama T, Iwai Y, Yachida M. Integration of facial position estimation and person identification for face authentication. Syst Comp Jpn. 2007;38(5):43-58.

[155] Rowley AG. Evaluating uncertainty for laboratories: a practical guide and handbook; 2001 January.

[156] Roelofse MM, Steyn M, Becker PJ. Photo identification: Facial metrical and morphological features in South African males. Forensic Sci Int. 2008:doi:10.1016/j.forsciint.2007.12.003

[157] Jenkins R, Burton AM, White D. Face recognition from unconstrained images: progress with prototypes. Seventh IEEE International Conference on Automatic Face and Gesture Recognition (FG'06); 2006; 2006. p. 25-30.

[158] Burton AM, Jenkins R, Hancock PJB, White D. Robust representations for face recognition: The power of averages. Cog Psych. 2005;51:256–84.

[159] Miller S. Experimental Design and Statistics. Second ed. Philadelphia: Brummer-Routledge 1974.

[160] Ayoub A, Garrachy A, Hood C, White J, Siebert JP, Spencer R, et al. Validation of a vision-based, three-dimensional facial imaging system. Cleft Palate Craniofac J. 2003;40(5):523-9.

[161] Adams BJ, Byrd JE. Interobserver variation of selected postcranial skeletal measurements. J Forensic Sci. 2002;47(6):1193-202.

[162] Aitken CGG, Stoney DA. The use of statistics in forensic science. West Sussex: Ellis Horwood Limited 1991.

[163] Robertson B, Vignaux GA. Interpreting evidence evaluating forensic science in the courtroom. West Sussex: John Wiley & Sons Ltd. 1995.

[164] Aitken CGG. Statistics and the evaluation of evidence for forensic scientists. West Sussex: John Wiley & Sons Ltd 1995.

[165] Lucy D. Introduction to statistics for forensic scientists. West Sussex: John Wiley & Sons Ltd. 2005.

[166] Armitage P, Berry G, Matthews JNS. Statistical Methods in Medical Research. 4th ed. Malden, Massachusetts: Blackwell Science Ltd 2002.

[167] Holden M. Errors led London police to kill Brazilian. Reuters, UK 2007 [cited 1 February 2008]; Available from: http://uk.reuters.com/article/idUKL0188687220071001

[168] Ferrario V, Sforza C, Poggio C, Cova M, Tartaglia G. Preliminary evaluation of an electromagnetic three-dimensional digitizer in facial anthropometry. Cleft Palate Craniofac J. 1998;35:9-15.

[169] Kleinberg KF, Vanezis P, Burton AM. Failure of anthropometry as a facial identification technique using high-quality photographs. J Forensic Sci. 2007;52(4):779-83.

[170] Weinberg SM, Naidoo S, Govier DP, Martin RA, Kane AA, Marazita ML. Anthropometric precision and accuracy of digital three-dimensional photogrammetry: Systems with one another and with direct anthropometry. J Craniofac Surg. 2006;17(3):477-83.

[171] Longuet-Higgins HC. A computer algorithm for reconstructing a scene from two projections. Nature. 1981;293(10):133-5.

[172] DeCarlo D, Metaxas D, Stone M. An anthropometric face model using variational techniques. Proceedings of the 25th annual conference on Computer Graphics and Interactive techniques; 1998: ACM Press; 1998. p. 67-74.

[173] Zhang C, Cohen FS. 3-D face structure extraction and recognition from images using 3-D morphing and distance mapping. IEEE Trans Image Proc. 2002;11(11):1249-59.

[174] Goos MIM, Alberink IB, Ruifrok ACC. 2D/3D image (facial) comparison using camera matching Forensic Sci Int. 2006;163:10-7.

[175] Lee J, Lee E-D, Tark H-O, Hwang J-W, Yoon D-Y. Efficient height measurement method of surveillance camera image. Forensic Sci Int. 2007;doi:10.1016/j.forsciint.2007.10.008

[176] Gloor PA. Bertillon's method and anthropological research; a new use for old anthropometric files. J Forensic Sci Society. 1980;20:99-101.

[177] Nakashima E. FBI prepares vast biometric database. The Washington Post. 2007 22 December;Sect. A01.

[178] Jenkins R, Burton AM. 100% accuracy in automatic face recognition. Science. 2008;319:435.

[179] Dougherty J. Poisoning diagnosis 'rock solid'. CNN 2004 [cited 2005 May 5]; Available from: http://www.cnn.com/2004/WORLD/europe/12/12/yushchenko/

[180] Walsh NP. Doctors prove Yuschenko was victim of poisoning. The Observer. 2004 December 12.

Appendix- Low Variant Proportions

Appendix Table 1 Descending proportion values and rank of element standard deviations in Sample 1 and Sample 2

Sample 1 (video)			Sample 2 (photos)	
Proportion No.	**Proportion SD**	**Rank**	**Proportion No.**	**Proportion SD**
cph-cph'/sn-ls	0.3649	**1.**	sbal'-sn/sn-prn	0.2746
en'-ex'/ps'-pi'	0.2295	**2.**	en'-ex'/ps'-pi'	0.2614
sn-sto/sto-sl	0.2093	**3.**	sbal-sn/sn-prn	0.2479
sbal'-sn/sn-prn	0.2004	**4.**	en-ex/ps-pi	0.2201
en-ex/ps-pi	0.1857	**5.**	sn-sto/sto-sl	0.1870
pi'-al'/sa'-ex'	0.1737	**6.**	cph-cph'/sn-ls	0.1493
sbal-sn/sn-prn	0.1557	**7.**	li-sl/sn-ls	0.1362
pi-al/sa-ex	0.1463	**8.**	pi-al/sa-ex	0.1197
ch-ls/n-prn	0.1367	**9.**	pi'-al'/sa'-ex'	0.1150
sto-li/ch-ch'	0.1310	**10.**	n-sn/sa-sba	0.0863
ch'-ls/n-prn	0.1302	**11.**	sa-sba/pa-obi	0.0824
li-sl/sn-ls	0.1247	**12.**	n-sn/sa'sba'	0.0813
ex-obi/ex-ch	0.1020	**13.**	ex'-al'/ch'-gn	0.0789
ex'-obi'/ex'-ch'	0.1009	**14.**	ex-al/ch-gn	0.0749
sa'-sba'/pa'-obi'	0.0966	**15.**	go-go'/n-gn	0.0742
sa-sba/pa-obi	0.0937	**16.**	ch'-ls/n-prn	0.0712
n-sn/sa-sba	0.0890	**17.**	sa'-sba'/pa'-obi'	0.0675
n-sn/sa'sba'	0.0829	**18.**	ch-ls/n-prn	0.0667
ex'-al'/ch'-gn	0.0758	**19.**	al-al'/n-sn	0.0626
ex-al/ch-gn	0.0755	**20.**	sn-gn/n-sto	0.0617
sn-gn/n-sto	0.0713	**21.**	ex-obi/ex-ch	0.0593
obi'-ch'/g-sa'	0.0705	**22.**	n-gn/n-sto	0.0551
go-go'/n-gn	0.0687	**23.**	ex'-obi'/ex'-ch'	0.0550
obi-ch/g-sa	0.0665	**24.**	sto-li/ch-ch'	0.0545
ex-go/go-go'	0.0660	**25.**	al'-ls/ch'-gn	0.0540
ex'-go'/go-go'	0.0653	**26.**	obi'-ch'/g-sa'	0.0530
n-gn/n-sto	0.0636	**27.**	ex-go/go-go'	0.0513
al-al'/n-sn	0.0612	**28.**	sl-gn/sto-gn	0.0500
sl-gn/sto-gn	0.0600	**29.**	ex'-go'/go-go'	0.0490
ls-sto/ch-ch'	0.0554	**30.**	al-ls/ch-gn	0.0483
al'-ls/ch'-gn	0.0527	**31.**	ex'-sto/n-sto	0.0461
gn-go'/n-gn	0.0516	**32.**	ch-li/ex-ch	0.0457
al-ls/ch-gn	0.0516	**33.**	obi-ch/g-sa	0.0456
ch-li/ex-ch	0.0473	**34.**	sbal'-ls/n-al'	0.0449
gn-go/n-gn	0.0467	**35.**	gn-go'/n-gn	0.0448
ch'-li/ex'-ch'	0.0458	**36.**	ex-ex'/go-go'	0.0441
ex-sto/ex-ch'	0.0457	**37.**	sbal-ls/n-al	0.0436
ex-ex'/go-go'	0.0451	**38.**	ex-sto/n-sto	0.0432
ex-n/n-sto	0.0442	**39.**	ch'-li/ex'-ch'	0.0392
ex'-sto/ex'-ch	0.0434	**40.**	ex'-n/n-sto	0.0387
sbal-ls/n-al	0.0409	**41.**	ex-n/n-sto	0.0384
ex-sto/n-sto	0.0406	**42.**	sa'-sba'/n-gn	0.0372
sbal'-ls/n-al'	0.0402	**43.**	sa-sba/n-gn	0.0363
ex'-n/n-sto	0.0400	**44.**	gn-go/n-gn	0.0355
ex'-sto/n-sto	0.0372	**45.**	n-prn/g-pg	0.0318
ex-n/ex-sto	0.0369	**46.**	al-al'/ex-ex'	0.0307
sa-sba/n-gn	0.0366	**47.**	pi-or/en-ex	0.0303
ex'-n/ex'-sto	0.0357	**48.**	en-al/ex-ch	0.0302
pi'-or'/en'-ex'	0.0340	**49.**	ex-n/ex-sto	0.0288
al-al'/ex-ex'	0.0335	**50.**	en'-al'/ex'-ch'	0.0277
en-al/ex-ch	0.0334	**51.**	ls-sto/ch-ch'	0.0276
sa'-sba'/n-gn	0.0328	**52.**	pi'-or'/en'-ex'	0.0270
n-prn/g-pg	0.0314	**53.**	ex'-n/ex'-sto	0.0257
en'-al'/ex'-ch'	0.0310	**54.**	en-en'/ex-ex'	0.0241

sn-gn/ex'-gn	0.0286	**55.**	n-sn/n-sto	0.0240
sn-gn/ex-gn	0.0277	**56.**	sn-gn/ex'-gn	0.0227
en-en'/ex-ex'	0.0257	**57.**	sn-gn/ex-gn	0.0222
n-sn/n-sto	0.0245	**58.**	ex-sto/ex-ch'	0.0197
pi-or/en-ex	0.0245	**59.**	ex'-sto/ex'-ch	0.0189

Appendix Table 2 Lowest variance proportions that were removed from comparison of Sample 1 vs. matched faces from Sample 2 for each percentage category

5% of lowest variance proportions removed	10% of lowest variance proportions removed	25% of lowest variance proportions removed	50% of lowest variance proportions removed
22. pi-or/en-ex	22. pi-or/en-ex	22. pi-or/en-ex	22. pi-or/en-ex
36. n-sn/n-sto	36. n-sn/n-sto	36. n-sn/n-sto	36. n-sn/n-sto
21. en-en'/ex-ex'	21. en-en'/ex-ex'	21. en-en'/ex-ex'	21. en-en'/ex-ex'
	58. sn-gn/ex-gn	58. sn-gn/ex-gn	58. sn-gn/ex-gn
	59. sn-gn/ex'-gn	59. sn-gn/ex'-gn	59. sn-gn/ex'-gn
	39. en'-al'/ex'-ch'	39. en'-al'/ex'-ch'	39. en'-al'/ex'-ch'
		3. n-prn/g-pg	3. n-prn/g-pg
		9. sa'-sba'/n-gn	9. sa'-sba'/n-gn
		38. en-al/ex-ch	38. en-al/ex-ch
		6. al-al'/ex-ex'	6. al-al'/ex-ex'
		23. pi'-or'/en'-ex'	23. pi'-or'/en'-ex'
		14. ex'-n/ex'-sto	14. ex'-n/ex'-sto
		8. sa-sba/n-gn	8. sa-sba/n-gn
		13. ex-n/ex-sto	13. ex-n/ex-sto
		18. ex'-sto/n-sto	18. ex'-sto/n-sto
			16. ex'-n/n-sto
			43. sbal'-ls/n-al'
			17. ex-sto/n-sto
			42. sbal-ls/n-al
			57. ex'-sto/ex'-ch
			15. ex-n/n-sto
			12. ex-ex'/go-go'
			56. ex-sto/ex-ch'
			55. ch'-li/ex'-ch'
			4. gn-go/n-gn
			54. ch-li/ex-ch
			52. al-ls/ch-gn
			5. gn-go'/n-gn
			53. al'-ls/ch'-gn
			27. ls-sto/ch-ch'

Appendix- Morphological Analysis

1 Morphology Study of Paired Still and Video Images

The anthropometric identification system developed by Alphonse Bertillon was devised to measure parts of the body, which also incorporated a morphological vocabulary that described different aspects of both the face and body [17]. At that point in time, individuals were being described in general terminology, contributing to what Bertillon believed was poor identification by the police. His descriptions were detailed and it was his goal that when given an account of an individual using this specialized vocabulary, a replication could be created of what they looked like.

This chapter consists of the methodology, results and discussion of a study undertaken on high resolution still and video images. The aim of this study was to conduct a morphological analysis and determine the possibility of confirming the identity of individuals based on a set of facial characteristic descriptions and facial feature comparisons. The results show that performing a morphological analysis in the manner would not be used in situations where there is already a suspect photograph to compare to a crime scene video image, but perhaps would help narrow down the possible suspects an eye witness would have to look at.

1.1 Materials and Methods

1.1.1 Subjects

The same set of high resolution photographs and video images taken in a controlled setting were used as for the anthropometry research which included a total of 199 images, comprising 119 photographs and 80 video images.

1.1.2 Data Collection

Using the checklist in Figure 0.1, a morphological analysis was performed on the 199 images using characteristics determined from a paper written by Vanezis et al. (1996) [35]. Vanezis et al. selected their list after conducting research from criteria listed in Ïşcan's book, Forensic Analysis of the Skull [28].

Figure 0.1 has been removed due to Copyright restrictions

The characteristics from Vanezis et al., which apply only to adult Caucasian males, were chosen based on research carried out by them and the criteria for inclusion were [35]:

- "Ease of discrimination among subset features,
- Good agreement among assessors,
- Non-reliance on anthropometric data (linear measurements and proportions),
- Permanence of feature, and
- Feature, part of normal morphological anatomical variation"

This study used twenty facial characteristics for comparison. As only one frontal view photograph was provided, categories that utilized criteria observed from the side, such as ear lobes, septum tilt and chin projection were not used. Each category had three to seven descriptions to choose from, with facial form as the category with the most choices and philtrum depth and shape with the least. All sections had "undecided" as one of their choices, allowing an "out" for the operator if needed.

Once descriptions of the 199 images had been recorded, the following analyses were carried out.

1.2 Results

The data obtained from this study is qualitative containing nominal variables. Comparisons of video images to database photographs were based on subjective observations of facial features. The chosen facial features were done so assisted by science.

1.2.1 Analysis One: Match rates of facial features between video images and database photographs

To determine the number of morphological descriptions from the database photographs that coincided with each video image, a comparison was set up consisting of one video image evaluated against 119 database photographs. Each video image had a true positive photograph which was always included in the database of photographs. As there were 80 video images, there were a total of 80 comparisons. Match rate analysis is demonstrated by first exhibiting match rates for each facial feature across all 80 comparisons and then illustrating match rates between each video image and its true positive photograph.

To establish match rates (definitions given in Table 0.1), if the same description was given for the video image and its true positive photograph within the database, it was deemed a True Positive and if differing descriptions were given it was deemed to be a False Negative. All other database photographs that matched descriptions with the video image were regarded to be False Positives and if differing descriptions were given, they were regarded as True Negatives. Results are summarized for the 80 comparisons in Table 0.2 separating facial features giving the sensitivity, specificity, positive predictive value, and negative predictive value and accuracy percentages.

Table 0.1 Definitions given to correct and incorrect match rates

True Positive (TP): a match which is also a correct match between the video image and photograph of the same subject.

False Positive (FP):a match which is an incorrect match between a video image and a photograph of two different subjects.

True Negative (TN): a match which is excluded and which is a correct exclusion because it involves a video image and a photograph of two different subjects.

False Negative (FN): a match which is excluded but which is an incorrect exclusion because it involves the video image and photograph of the same subject.

Sensitivity: The proportion of true matches that were correctly identified, which is calculated as:

TPx100/ (TP+FN)

Specificity: The proportion of false matches, which is calculated as:

TNx100/ (TN+FP)

Positive Predictive Value (PPV): The proportion of true matches out of all the comparisons, which is calculated as:

TPx100/ (TP+FP)

Negative Predictive Value (NPV): The proportion of true exclusions out of all the comparisons, which is calculated as:

TNx100/ (FN+TN)

Accuracy: The proportion of video images and photographs that were correctly matched in the study, which is calculated as:

(TP+TN)x100/(TP+TN+FP+FN)

Table 0.2 Summary of match rates between video images and database photographs for N=80 separated by facial features

	TN	TP	FP	FN	Sensitivity %	Specificity %	PPV %	NPV %	Accuracy %
Facial form	6184	51	3256	29	63.8	65.5	1.5	99.5	65.5
Facial fatness	5565	55	3875	25	68.8	59	1.4	99.6	59
Chin feature	5956	48	3484	32	60	63.1	1.4	99.5	63.1
Chin shape	6134	46	3306	34	57.5	65	1.4	99.4	64.9
Malars	5088	53	4352	27	66.3	53.9	1.2	99.5	54
Eyebrow shape	3787	55	5653	25	68.8	40.1	1	99.3	40.4
External eyebrow ends	6783	32	2657	48	40	71.9	1.2	99.3	71.6
Eyebrow density	5641	59	3799	21	73.8	59.8	1.5	99.6	59.9
Eye shape	6573	38	2867	42	47.5	69.6	1.3	99.4	69.4
Palpebral slit	6914	36	2526	44	45	73.2	1.4	99.4	73
Eye bag	5630	45	3810	35	56.3	59.6	1.2	99.4	59.6
Nose tip shape	7229	28	2211	52	35	76.6	1.3	99.3	76.2
Nostril	5342	45	4098	35	56.3	56.6	1.1	99.3	56.6

visibility									
Nasal alae	7577	25	1863	55	31.3	80.3	1.3	99.3	79.9
Philtrum depth	5024	49	4416	31	61.3	53.2	1.1	99.4	53.3
Philtrum shape	5227	36	4213	44	45	55.4	0.8	99.2	55.3
Upper lip notch	6227	39	3213	41	48.8	66	1.2	99.3	65.8
Upper lip thickness	5554	50	3886	30	62.5	58.8	1.3	99.5	58.9
Lower lip thickness	5832	48	3608	32	60	61.8	1.3	99.5	61.8
Ear projection	7207	35	2233	45	43.8	76.3	1.5	99.4	76.1
Total	**119474**	**873**	**69326**	**727**	**54.6**	**63.3**	**1.2**	**99.4**	**63.2**

Across 80 comparisons, the correct match rates of facial features from video images to their true positive database photographs can be seen in the number of true positives. The sensitivity of the test is the percentage of true positive results out of all comparisons and is a way to show the discriminatory value of each facial feature. The range was from 31% for nasal alae to 74% for eyebrow density. The low percentage from the sensitivity test shows that nasal alae may not have been a feature that was easily able to differentiate on an image but that eyebrow density was a more distinctive facial characteristic to describe on both video images and photographs because it had a much higher sensitivity and therefore number of true positives in proportion to the total number of comparisons.

The match rates of facial features of a video image to incorrect database photographs can be seen in the number of false positives. A high number of photographs from the database that match the descriptions of the video image could indicate a common facial feature while the lower numbers may signify a feature description not as widespread. This regularity is important to note because it demonstrates that the greater the frequency of the facial feature amongst the database photographs then the less effective the comparison is with that facial feature because if may be considered to have a low discriminatory factor.

The graph in Figure 0.2 illustrates correct and incorrect match rates for each facial feature between the video images and their true positive photographs. Series one (bottom), shows the percentage of true positives each facial feature had amongst the 80 comparisons, while series two (top), shows the percentage of false negatives that each facial feature had from the 80 comparisons, equalling 100%. Among the entire group of facial features, there are no noticeable trends between the percentages of true positives and false negatives.

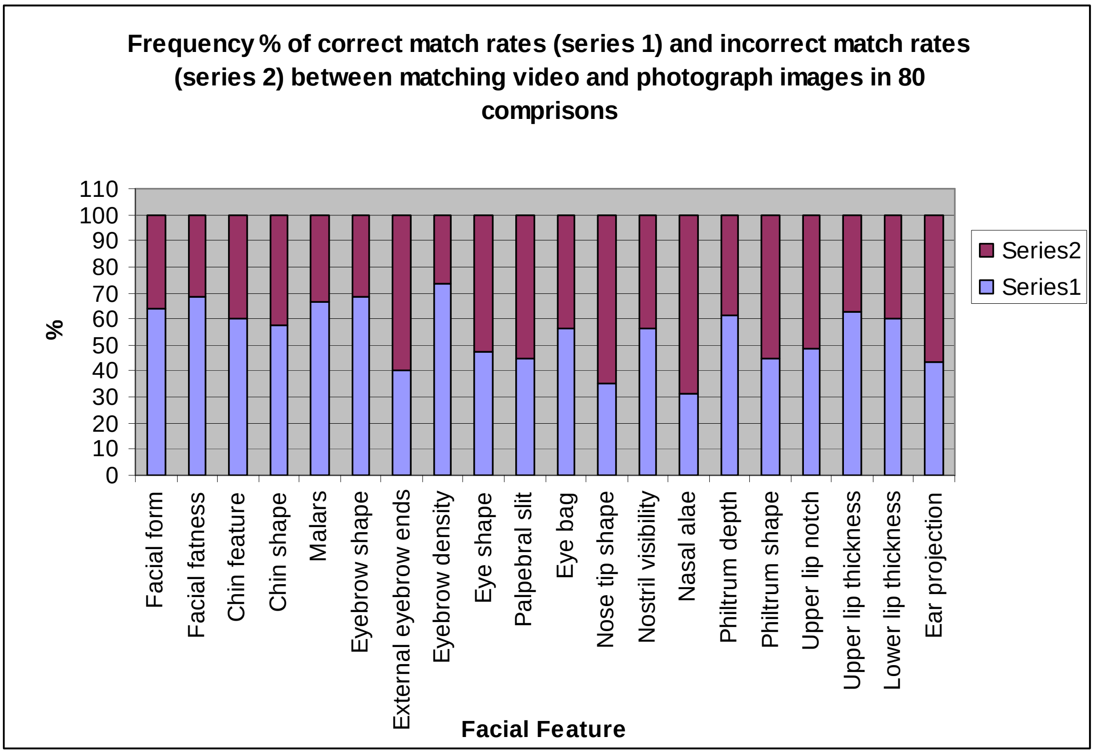


Figure 0.1 Match rates (Series 1) vs. incorrect match rates (Series 2) for each facial feature of video images with their true positive photograph in 80 comparisons

1.2.2 Analysis two: Match rates between video image and their true positive database photograph in 80 comparisons

After an examination of the match rates of the total number of facial features between each video image and its true positive database photograph, the descriptive statistics in this analysis display results from correct matches shared from the 20 total facial features. In practice, to label a video and photograph match successful enough to warrant further investigation into the suspect, a high rate of true positives would be needed in addition to a low number of false positives within the comparison.

Table 0.3 contains descriptive statistics for correct matches between the video images and the database of photographs in 80 comparisons. Part (a) describes the true positives achieved between the video images and their true positive photograph and part (b) describes the true negatives achieved between the video images and the remaining photographs in the database. The greatest number of true positive matches between a video image and photograph could have been 20 but the mean was only 11, which is not much over 50% and the range was 5 to 16. The mean of true negatives in the comparisons between a video image and the remaining photographs in the database was 1493 but the highest possible number could have been 2360 (75%) in each video image comparison. A comparison that contains a large amount of true negatives along with the will help to narrow down a suspect list but only if the amount of true positives in the comparison is large as well.

Table 0.3 Correct matches: descriptive statistics for true positive match rates (a) between video image and its true positive database photograph and true negative match rates (b) between the video images and database photographs in 80 comparisons.

	N	Mean	SE Mean	SD	CV	Min	Max	Median	Mode
(a)	80	11	0.27	2	22	5	16	11	11
(b)	80	1493	10	91	6	1287	1718	1490	1372,1444, 1449, 1483*

*The data contain at least five mode values. Only the smallest four are shown.

Table 0.4 contains descriptive statistics for incorrect matches between the video images and the database of photographs in 80 comparisons. Part (a) describes the false negatives achieved between the video images and their true positive photograph and part (b) describes the false positives achieved between the video images and the remaining photographs in the database. The range of false negative matches between a video image and its true positive photograph was 4 to 15 and while the greatest number could have been 20, the mean was as high as 9. The mean of false positives in the comparisons between a video image and the remaining photographs in the database was 867 out of a possible 2360 (37%) in each video image comparison. A low number of false positives are necessary in a comparison to avoid implicating the wrong person, but obtaining a small amount of false negatives are essential as well to avoid missing the correct individual.

Table 0.4 Incorrect matches: descriptive statistics for false negative match rates (a) between a video image and its true positive database photograph and false positive match rates (b) between video images and database photographs and in 80 comparisons.

	N	Mean	SE Mean	SD	CV	Min	Max	Median	Mode
(a)	80	9	0.27	2	27	4	15	9	9

(b)	80	867	10	91	11	642	1073	871	763, 787, 832, 833*

*The data contain at least five mode values. Only the smallest four are shown.

Figure 0.3 illustrates the numbers of true positive in (ascending order) and false negatives matches between each video image and its true positive photograph in the total of 80 comparisons. The x-axis labels do not indicate which video image is attributed to which bar but instead show that 80 comparisons were completed.

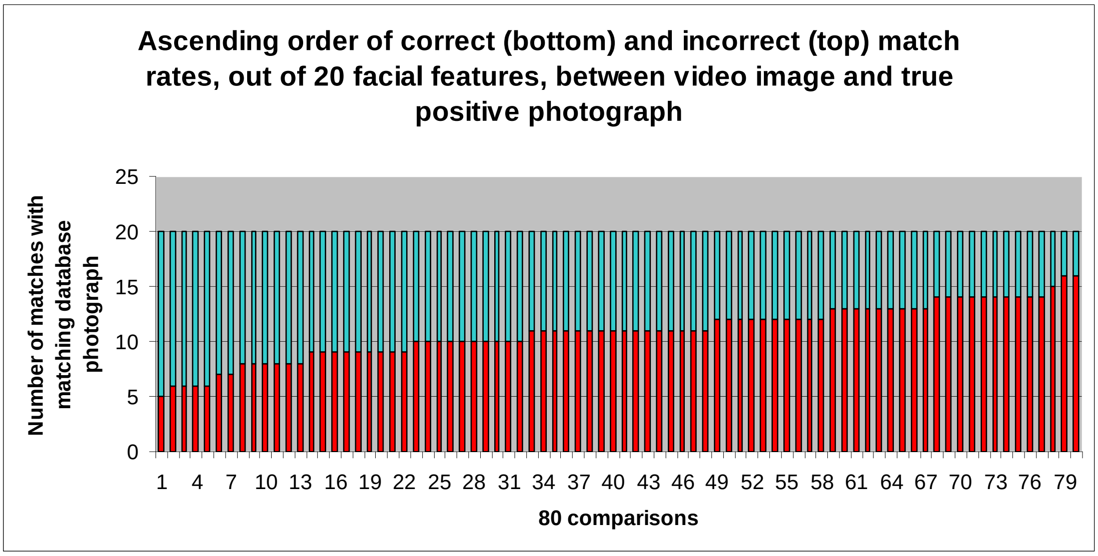


Figure 0.2 Match rates (correct (bottom) and incorrect (top)) out of 20 facial features between the video image and its true positive photograph grouped in ascending order

The histogram in Figure 0.4 shows the frequency distribution in 80 comparisons of the number of matches with a video image and its true positive photograph. The data show a normal distribution of true positive matches between the video images and their true positive database photographs.

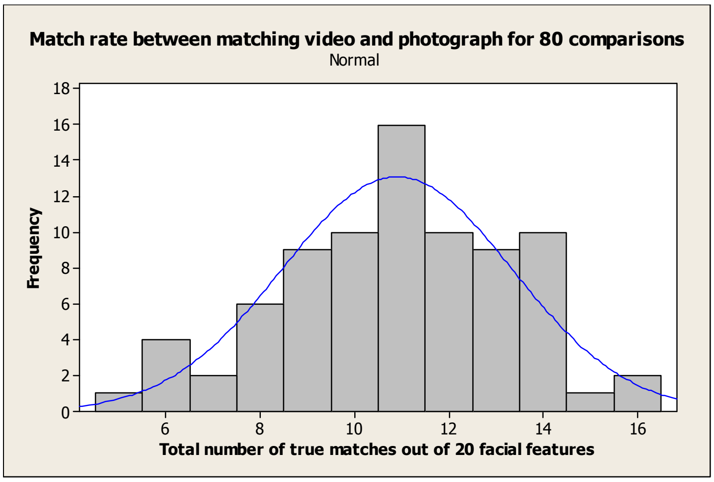


Figure 0.3 Normal distribution histogram showing the number of correct matches between video image and its true positive photograph in 80 comparisons

1.2.3 Analysis two: Chi square

The majority of the facial zones contained multiple groupings (i.e. external eyebrow ends, eyebrow shape) to describe the images that were compared against each other. In this analysis, the statistical test chi-square (x^2), shown in Table 0.5, was used to illustrate the comparison of the number of true positive matches for each facial feature and therefore the significance of the different categories for each section of the face. Significance levels show how likely a result is due to chance. A p value of .05 equals a 5% chance of being not true or a 95% of being true. Just over half of the findings were observed to be significant. Subsequent to completing the calculations, it was observed that when determining x^2 in a two by two table, the larger the difference between observed true positive numbers in each category, the larger the x^2, which equals a significant *p*. Therefore, the more significant the two features are in determining which one is better for comparison. In general, when the observed and expected frequencies are close together, the two variables are not associated.

Table 0.5 Chi square analysis comparing different facial features completed on true positive results.

Facial feature comparison	*Degree of freedom (df)*	x^2	*P*
Facial form vs. Facial fatness	1	0.45	0.504**
Chin feature vs. chin shape	1	0.10	0.748**

Nose tip shape vs. nostril visibility vs. nasal alae	2	12.04	0.002*
Nose tip shape vs. nostril visibility	1	7.28	0.007*
Nostril visibility vs. nasal alae	1	10.16	0.001*
Nose tip shape vs. nasal alae	1	0.25	0.614**
Eyebrow shape vs. eyebrow density vs. external eyebrow ends	2	20.21	0.000*
Eye shape vs. palpebral slit	1	0.10	0.751**
Eye shape vs. eyebrow shape	1	7.42	0.006*
Philtrum shape vs. philtrum depth	1	4.24	0.039*
Upper lip thickness vs. lower lip thickness	1	0.11	0.746**

*** Not significant*
** Significant*

1.3 Discussion

Difficulties will arise when conducting comparisons in any method of facial identification, especially when working with photographs of individuals and a morphological analysis is no exception.

As with an anthropometric comparison, careful attention should be paid to the position of the head in photographs [77]. If the head is tilted to either side or to the back or front then a different description may occur. For example, the base of the nose in relation to the position of the ears should be noted. In a head that is tilted backward, the nose appears shorter and the nostrils more visible. Appearances can vary with even the slightest change in expression or photographic angle [28]. Weight changes, facial hair, dentition, illness, medication, disguises, hair colour, makeup, tattoos, accessories (hat, glasses), plastic surgery all affect appearance and therefore affect the description [77]. Lipstick may hide an upper lip notch or change the appearance of lip thickness. Eyebrow thickness can be made thinner by plucking eyebrows or thicker by adding makeup. Aging affects the morphology of the face and is especially hard to define or predict because as a whole everyone ages differently and various features on different people age differently. Tissue above the eyes drops, ears appear to get longer, the tip of the nose may drop and lips may thin [77]. Dental changes may affect the shape of the jaw and mouth. Medications, such as Prednisone, may change facial shape as can poisons such as dioxin. This was displayed recently with the poisoning of Ukrainian opposition presidential candidate Viktor Yushchenko in September 2004 [179]. This resulted in the now president's face becoming disfigured with a severe form of acne called chloracne [180].

The position of the head may affect how much the ears appear to be projected and this also shows in a photograph that is not completely facing forward. It may be difficult to establish the differences between slight and pronounced ear projections because of the angle or blurriness of the image. Also, if the head is not fully facing the front, nostril visibility may be compromised and some features may appear to be asymmetrical when they are in fact symmetrical. Head position also affects the palpebral slit because if the head is tilted in one direction or another, it is likely that the direction of the palpebral slit may change.

The results achieved from a morphological analysis are subjective. The individual who conducts the analysis should have a strong background in facial anatomy. In addition, a background in art may be useful too. Data may be less subjective when the individual has features that stand out or are to the extreme. Facial features have many degrees of portrayal and the description will reflect this. For example, some noses may be very pointed but some slightly pointed. Most operators would be in agreement when describing a very pointed nose, but if the nose is only slightly pointed, some operators may describe it as pointed while others may see it as rounded. For this reason, extreme features will be described more consistently. More often than not, the feature was an uncommon one, such as a dimple in the chin, straight eyebrows, asymmetrical nostrils, a fat face, or extended or flared nasal alae. This may indicate either the subjective nature of the technique or that it is difficult to describe the human face with so few descriptive words.

For this experiment, a morphological analysis was conducted on a total of 199 high resolution images. Eighty of the images were from video and each one was compared against the remaining 119 photographic images in a database. Previous research recorded these same individual's facial anthropometric measurements. Subsequent to that anthropometric analysis it was determined not possible to differentiate between individuals on the basis of anthropometry alone and when looking at the different persons' photographs, it is clear that there are differentiating characteristics amongst them.

A side by side comparison was carried out in all of the literature found to date in which a morphological analysis was conducted in addition to an anthropometric analysis [23, 37, 38]. The experiment undertaken for this thesis diverges from this method and was carried out with a different approach. The reason for this was to show that if there was a list of morphological descriptions on file, such as in Bertillon's time, of past and current offenders, could descriptions made at the time of arrest or descriptions based on eye witness accounts be matched to an offender database? This could effectively narrow down the number of potential suspects saving valuable police time. In this study, all descriptions are subjective with no scientific way of gaining descriptions. This could be one reason why identifications are not carried out in this manner. There is no way of scientifically proving the benefit of the test or the truth of the identification.

This was not a study to determine which facial features are useful to describe a person's face in a photograph. However, it is helpful to note which facial features accumulated higher match rates because it may indicate features that are easier to describe from video images. From this research, when examining morphology from a still video image, according to the number of true positives obtained, the easiest features to establish were the face form (64%) and facial fatness (69%). The facial shape is the largest area to observe and does not require a minute knowledge of the facial features. It may however, be more difficult to distinguish between oval and angular down faces in the video images due to blurriness. Facial fatness is also relatively easy to observe, however, the face may tend to appear fatter than it normally does, again due to image fuzziness. Another feature which was relatively easy to determine was the presence of malars (66%). This may be because of the position of the head in relation to the angle of the video camera. Careful consideration should be paid to this area because the appearance of malars may depend on where the light hits the photograph.

A noticeable limitation of the morphological analysis as a comparison tool occurred when looking at the video images. Video images were obviously not as sharp as the photographs, leading to the question of "Is the facial feature a result of pixels or a shadow or the actual feature?" Out of eighty video image analysed, 63% had chins that appeared to be featureless. Was this because their chin was featureless, or is it just that not enough detail could be seen? In this experiment, there seemed to be a large percentage of targets with rounded noses. This could have been due to a lack of photographic sharpness rather than an accurate description. The video images used as research material were of a high standard and therefore the match rates obtained are a best case scenario for this type of study. Video images of a lower standard can be hypothesised to compile much lower true positive percentages, although a study would need to be completed in order to prove this.

Facial landmarks such as the mouth, nose, and chin are difficult features to visualize clearly when a video produces a fuzzy image. Examples of these features are: nose tip shape, chin feature, upper lip notch, palpebral slit, septum tilt, upper and lower lip thickness, nostril visibility, external eyebrow ends, and nasal alae. Shadows, in addition to fuzziness may make it difficult to determine eye shape. Attention should be paid when observing narrow vs. oval eyes because the person could be slightly squinting in one of the photos compared to the other, giving the appearance of two different shaped eyes. Eyebrow density can be a difficult feature to make out in a video image because if the image is of low quality, the fuzziness renders it difficult to distinguish between thick, sparse, or normal eyebrows. For video images, eyebrow density and shape and external eyebrow ends were easier to determine when the hair was dark and more difficult when the hair was light brown or blond. These variables show the difficulty when assessing images from a 2D photograph, rather than a living person. The photographs below aid in demonstrating the difficulty that can occur when trying to assign facial features minute physical descriptions.

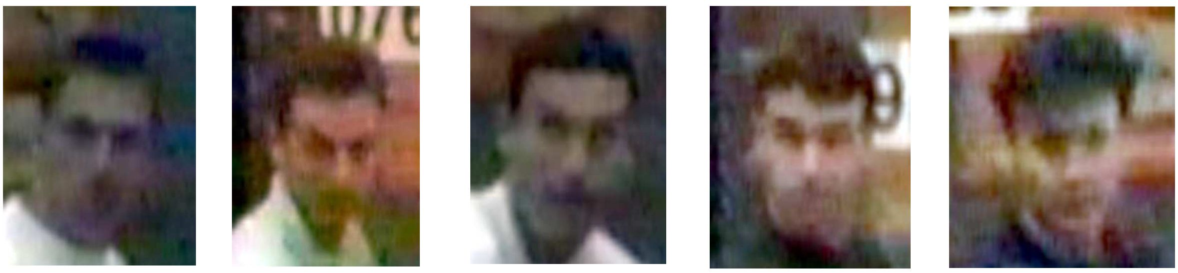


Figure 0.4 Examples of still images recorded on a VHS video tape taken from a surveillance camera

Categories where a majority of the subjects had the same feature description (example: eyebrow shaped---curved) would not indicate features of the best distinguishing nature. On the other hand, features that are not as common (straight eyebrow shape) would make a good comparison. Direct comparison of specific features such as is carried out during video superimposition might be better for this reason than a comparison by way of terminology alone. However, there is no denying that the more ways of corroborating evidence that proves or disproves identification the better. A morphological analysis would be of benefit as corroborating information. Experience of the operator is also important to achieve consistent results. The results were obtained by comparing one type of video but differing conclusions may be obtained from observing higher as well as lower quality video images as well as be obtained by different analysts.

Some features are useful to differentiate between individuals but are only easily established on quality images taken reasonably close. These for example are chin features, such as a cleft or double chin, nose tip shape, upper lip notch, and philtrum depth. However, these same features would not be easily established on poor quality image. Malars would tend not to be as noticeable in fuller faced individuals. If able to be seen clearly, eyebrow shape can be a good distinguisher between men because they usually do not change the shape of their eyebrows as women may do and therefore would likely be less varying. The presence of eye bags may not be a good feature to use for comparison due to the possible changeable nature of the feature. Circumstances such as stress or lack of sleep may affect the feature and may not be useful when comparing images taken from two different time periods.

As a morphological comparison is subjective, a written explanation of this and the limitations that occur because of this should be included with any analysis. Similar to conducting an anthropometric analysis or any facet of facial mapping, a more experienced operator will achieve greater accuracy. Caution should be followed when conducting the analysis especially when stating the final conclusion. Terms such as “more than likely the same person” or “could be the same person” are better to use than stating that there has been a positive identification.

The final conclusion which emerges from the results of this experiment is that morphology is useful when facial features are not too rigidly categorized in check lists, but instead are looked at as shapes with small, describable, contour differences that can be recorded and used by the trained operator to confirm or refute identification.