Barts and The London
School of Medicine and Dentistry

Queen Mary
University of London

Three Dimensional (3D) Forensic Facial Reconstruction in an Egyptian Population using Computed Tomography Scanned Skulls and Average Facial Templates:

A Study Examining Subjective and Objective Assessment Methods of 3D Forensic Facial Reconstructions

Dalia Ahmed Selim Ali Abdou

Submitted in partial fulfillment of the requirements of the Degree of Doctor of Philosophy

Under the Supervision of:

Professor Peter Vanezis
Professor Atholl Johnston

December 2017

STATEMENT OF ORIGINALITY

I, Dalia Ahmed Selim Ali Abdou, confirm that the research included within this thesis is my own work or that where it has been carried out in collaboration with, or supported by others, that this is duly acknowledged below and my contribution indicated. Previously published material is also acknowledged below.

I attest that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge break any UK law, infringe any third party's copyright or other Intellectual Property Right, or contain any confidential material.

I accept that the College has the right to use plagiarism detection software to check the electronic version of the thesis.

I confirm that this thesis has not been previously submitted for the award of a degree by this or any other university.

Signature:

Date:

Details of collaboration and publications: A list of the study presentations is included in Appendix 21.

ACKNOWLEDGEMENTS

Throughout this thesis, I have been encountered with numerous difficulties on different levels and at different stages right from the beginning. However, if I could not find that amount of support and assistance with every obstacle I faced, the thesis would not have been completed. Therefore, I am extremely grateful to every person who helped me reach this point.

For a starter, I am thankful to my financial sponsors, the Egyptian Ministry of Higher Education and the Egyptian Cultural Affairs and the Mission Sector, as well as the Egyptian cultural Counsellor and the staff of the Egyptian Cultural Centre and Educational Bureau in London, UK. I am equally grateful to my home department, the Department of Forensic Medicine and Clinical Toxicology, Faculty of Medicine, Zagazig University, Egypt. In particular, I am forever grateful for Professor Fatma Farahat, the former head of the department, for nominating me for this PhD scholarship, and Professor Eman El-Zahed, the current head of the department who continually encourages me and all the junior staff to seize similar opportunities and work on improving and developing out professional careers. I have always been supported by Professor Wafaa Hussien.

I consider myself extremely lucky for having Professor Peter Vanzis and Professor Atholl Johnston, with their greatest expertise that I could ever dreamt of, as my PhD supervisors. Not just that, but I cannot express enough appreciation for their encouragement and belief in me from the very first day and in many difficult times. If it were not for their endless support and wide communications, this thesis would not have come to light. Their modesty, knowledge and relentless endeavour to give the best work, have always been admirable.

I would like to also thank Dr. Maria Vanezis, Director at the Academy of Forensic Medical Sciences and the coinvestigator of this study, for her substantial input to this thesis. Not only Dr. Vanezis has put the time and effort to transfer her practical experience in the field of 3D forensic facial reconstruction to me, but also she generously allowed me to use her own database in my thesis to conduct the pilot study. That generous action has saved this thesis and ensured its progress at the time due to the delay we have encountered to gain the ethics committee approval. My thanks are also to the William Harvey Research Institute, Queen Mary University of London committee members, Professor Nick Goulding, Dr. Matin Carrier and Professor Tim Warner, who assessed the progress of my thesis and provided me with constructive and sincere recommendations, and to Professor Arthur Tucker who generously helped me apply for the approval of the UK National Research Ethics Committee.

I would like to express my greatest gratitude to the efforts of my colleagues in the Faculty of Medicine, Zagazig University, Egypt, Dr. Mohamed Amin, Lecturer of Diagnostic Radiology, and Dr. Essam Youssef, Lecturer of Neurosurgery, who provided me with the main resources that I depended on for conducting the main part of this thesis. I would like to also thank Dr. Curtis Offiah, Consultant Radiologist, as well as the staff in the radiology department of St Bartholomew's Hospital, London, particularly Ms Olufunmilayo Soile, as well as Dr. Vito Cirielli, Forensic Pathologist at the University of Verona, Italy, for their assistance in the data collection for this thesis. I am also very grateful to Ms Jacqui Dokoupil, the former Administrator of the QMUL Clinical Pharmacology Department, for the time and effort she has put in obtaining the patients consents and taking their photographs.

I have faced a number of procedural and technical difficulties along this thesis for which I have consulted a number of experts in different fields. They have generously provided me with their expertise and professional advice which helped overcome many obstacles. Therefore, I highly appreciate the assistance of Dr. Claudio Capelli, Institute of Cardiovascular Science, UCL, UK, Dr. Tim Niblett from The Turing Institute, Glasgow, UK, Dr. Lars Ebert, Institute of Forensic Medicine, University of Zurich, Switzerland, Professor Niels Lynnerup, the Head of Forensic Medicine Department, University of Copenhagen, Denmark, and last but not least the late Dr. Allan McNeill, former Senior Lecturer of Cognitive Psychology Glasgow Caledonian University, UK. I am also very grateful for the generous assistance from the staff of iGene Digital Autopsy Facilities.

I am, also, in debt for the experts in different forensic fields who kindly agreed to assess my work in an invaluable contribution to the study. I am thankful to the facial identification/perception psychology experts; Dr Melanie Pitchford and Dr William Brown, University of Bedfordshire, UK, and Dr Faye Skelton, Edinburgh Napier University, UK. I am also grateful to the Forensic Medicine/Pathology experts in Zagazig, University, Egypt; Professor Mie Gomaa, Professor Amal Abdulkhalek, Assistant Professor Shereen El-Khateb, Dr. Aisha Abo-Hashem, Dr. Dena Mohamed, Dr. Marwa Amer, and Dr. Ghadeer Mohamed, in addition to Dr. Vito Cirielli, University of Verona, Italy, and Dr. Kunasilan Subramaniam, Dr. Hanif Mahmud, and Dr. Nur Liyana Johar, The Ministry of Health, Malaysia. My thanks are also to the Forensic Anthropology experts who kindly participated in this study; Dr. David Gonçalves, Melanie Henriques, Catarina Coelho, Maria Alejandra Acosta, Joanna Drath, and Calil Ibrahim Makhoul from the Department of Life Sciences, University of Coimbra, Portugal, in addition to Dr. Sara Casado Zapico, Department of Anthropology, NMNH, Smithsonian Institution, Washington, DC, USA, Dr. Joe Adserias Garriga, University of Girona, and University of Barcelona Spain, Mara Karell, University of Edinburgh, Janamarie Truesdell, University of Cambridge, as well as Professor Peter Vanezis, the primary supervisor of this thesis, and Dr. Maria Vanezis, Academy of Forensic Medical Sciences, UK. In addition, I would like to express my deepest gratitude for the wide participation from the students (particularly of the Forensic Medicine MSc and the Clinical Drug Development MSc courses) and the staff of William Harvey Research Institute, Queen Mary, University of London, as well as the Barts and The London School of Medicine and Dentistry and the NHS Health trust as well as from the public.

Last but not least, without the blessing and the prayers of my parents, brother, and sisters, I would have gone nowhere with my studies nor with my life. May my love and best wishes be always with you! To my dearest husband, Mohamed, and children, Omar (10), Layla (7) and Adam (1), I attribute any success in my life to you for your active support, encouragement and patience.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

List of Abbreviations

.hips	Bitmap Graphics
.lsm	Linux Software Map
.obj	Object
.stl	Stereolithography
.tiff	Tag Image File Format
.xml	Extensible Markup Language
2D	Two Dimensional
3D	Three Dimensional
BMI	Body Mass Index
CBCT	Cone-Beam Computed Tomography
CCTV	Closed-Circuit Television
CFT	Craniofacial Template
CT	Computed Tomography
DICOM	Digital Imaging and Communications in Medicine
DNA	Deoxyribonucleic acid
EDM	Euclidean Distance Matrix
FBI	Federal Bureau of Investigation
FFR	Forensic Facial Reconstruction
FOV	Field Of View
FR	Facial Reconstruction
GE	General Electric Company

GIMP	GNU Image Manipulation Program
HU	Hounsfield Units
MRI	Magnetic Resonance Imaging
MDCT	Multi Detector Computed Tomography
MSCT	Multi Slice Computed Tomography
PI	Proportion Indices
RE/FACE	Reality Enhancement Facial Approximation by Computational Estimation
RMS	Root Mean Square
ROI	Region Of Interest
sDT	Signed Distance Transform
SSD	Sum of Squared Differences

ABSTRACT

Introduction:

Forensic facial reconstruction can assist identification by reconstructing a face of the unknown person with the aim of its recognition by his/her family or friends. In the facial reconstruction approach adopted in this study, a 3D average face template was digitally warped onto a 3D scanned skull image. This study was carried out entirely on an Egyptian population, and was the first of its kind.

Aims:

This study aimed to demonstrate that 3D facial reconstructions using the novel methodology described could show significant resemblance to the faces corresponding to the persons in question when they were alive. Moreover, using techniques previously validated for facial reconstruction, the aim was to compare them to the method developed, and to assess approaches used to determine the accuracy of 3D facial reconstructions.

Methods:

Initially, a pilot study was conducted using a database of laser scanned skulls and faces. The faces were reconstructed using an average facial template generated by merging a number of faces of similar population, sex, and age. The applicability, as well as the main components of the facial reconstruction method, the single and average facial templates, and the facial soft tissue thickness measurements, were investigated. Furthermore, in the main study, the faces of computed tomography (CT) scanned heads of an Egyptian population were reconstructed using average facial templates. The accuracy of the reconstructed faces was assessed subjectively by face pool, and face resemblance tests, and objectively by measuring the surface distances between the real and reconstructed faces. In addition, a number of novel subjective and objective assessment methods were developed. These included assessment of individual facial regions using subjective resemblance scores, and objective surface distance comparisons. A new objective method, craniofacial anthropometry, was developed by taking and comparing direct measurements from the skull, and comparing the measurements from the real and reconstructed faces.

The studied cases were ranked according to all subjective, and objective, tests, and statistically correlated.

Results and Conclusions:

The average facial templates showed a higher identification rate than the single face templates. The approach of facial reconstruction used in this thesis showed a comparable accuracy to many other facial reconstruction methods, yet was superior in terms of its applicability, transferability, and ease of use. In the face pool tests, the younger assessors were able to correctly identify the reconstructed faces better than older assessors. Furthermore, the identification rate by the forensic anthropology experts was higher than the non-experts. The former group showed the highest agreement between the observers in giving the resemblance scores. Although there was a significant rank correlation between the subjective and objective assessment tests, the subjective tests are influenced by the assessors' subjective characteristics (e.g., age, professional experience), thus making objective assessment more reliable. However, in situations where subjective tests are used, it is better to use the face resemblance tests and consult forensic anthropologists. Also, Craniofacial Anthropometry, particularly the craniofacial angles, can successfully indicate the accuracy of the facial reconstructions. Importantly, this study shows that certain facial regions, particularly the cheek and the jaw, are more reliable than other areas in the subjective and objective assessment of the facial reconstruction.

What is Already Known about this Subject:

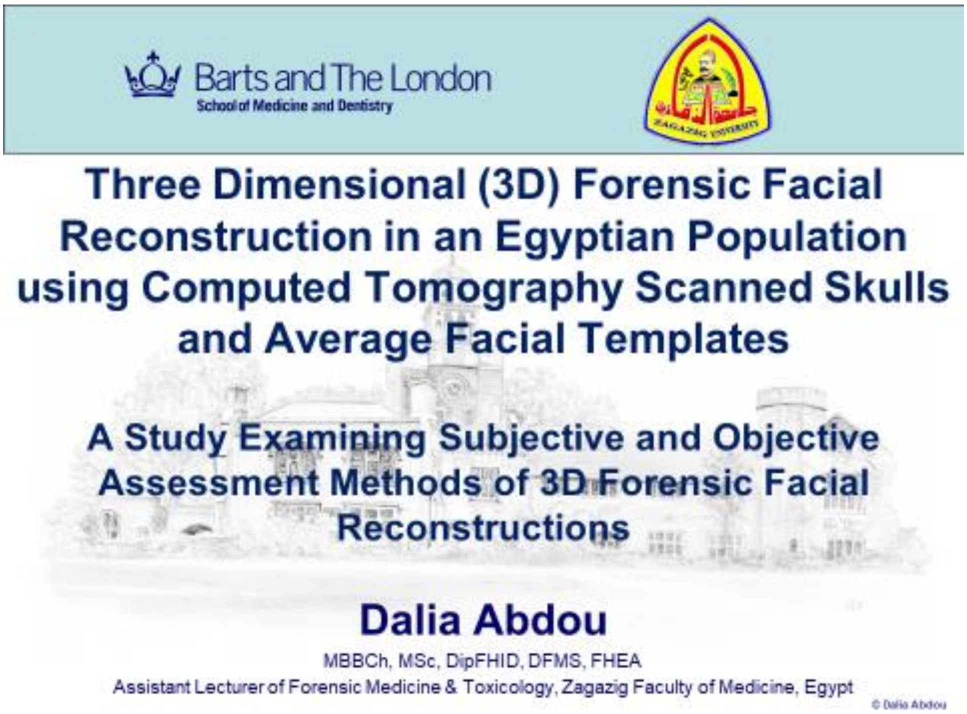
- Forensic facial reconstruction is not a primary method of identification, but an assisting method.
- Most methods of forensic facial reconstruction require specific experience in musculoskeletal head anatomy, and some techniques require additional input from mathematical and technological fields.
- Forensic facial reconstruction has been assessed via subjective and objective methods

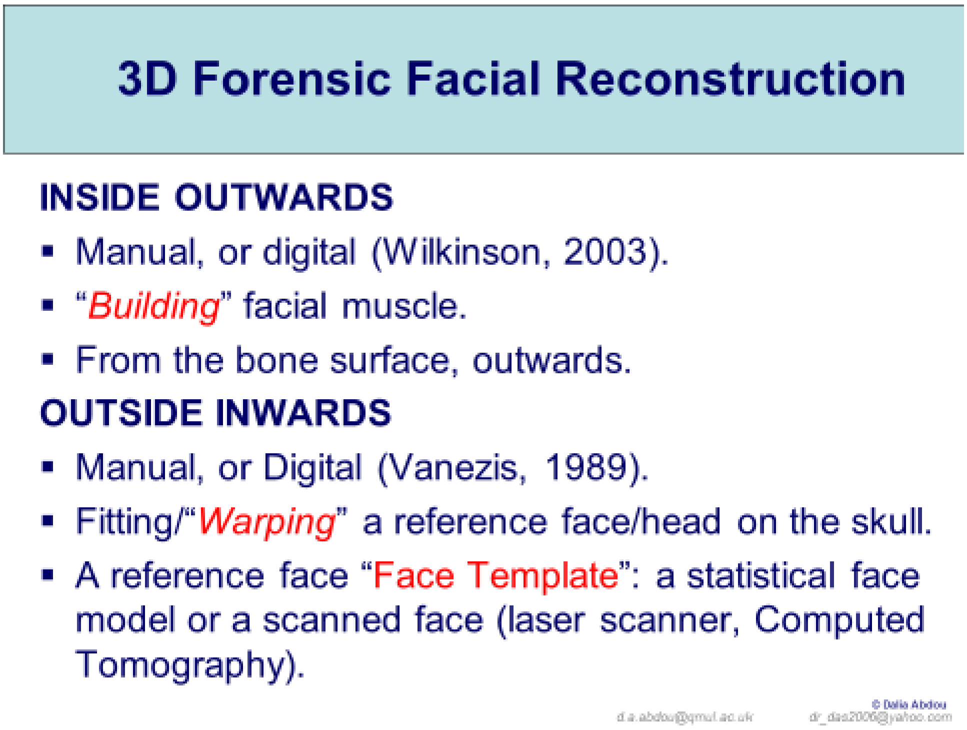
What this Study Adds:

- The first study of its kind to be applied to Egyptian population.
- The first study of its kind to present a comprehensive and detailed analysis of different subjective and objective method used for the assessment of the accuracy of facial reconstruction.
- The first study of its kind to recruit a number of experts in forensic face recognition psychology, forensic pathology and forensic anthropology with and without facial reconstruction experience.
- The first study of its kind to include 3D laser scanned and 3D CT scanned cases, as well as 2D photographs and 3D CT scanned faces for the assessment of facial reconstructions.
- The first study of its kind to conduct subjective analysis via 3D online testing.
- The "outside Inwards" approach to facial reconstruction adopted in this study was validated with comparable recognition rates and resemblance to the target, compared to other approaches.
- This approach to facial reconstruction, using scanned facial templates, proved to be a quick, easy to learn with no previous experience and cost-effective method.
- A detailed and illustrated user manual was developed for use in facial reconstruction using the "Outside Inwards" approach and employing the present facial reconstruction software. The manual was validated and tested by inexperienced users.

- Certain, single, scanned facial templates were better than others in obtaining identification and resemblance to the target.
- The average facial templates are better than many single scanned facial templates.
- A population-specific facial template and facial soft tissue depth are essential. However, the influence of the facial soft tissue depths on the resulting facial reconstruction is not strongly related to the method these depths were measured.
- Detailed analysis of the usual format of the face pool tests with evidence-based suggestions to improve the design of the tests for a more reliable assessment of the facial reconstructions.
- An evidence-based conclusion that objective methods are more reliable than subjective methods. However, the significant correlation between the test types allows the use of subjective tests when needed.
- An evidence-based conclusion that the subjective face resemblance tests are more accurate and reliable than the subjective face pool tests.
- Establishing the relationship between the subject’s age and professional experience in forensic anthropology and better performance in the subjective face pool tests.
- Developing and validating a novel method (Craniofacial Anthropometry) for objective assessment of the accuracy of facial reconstructions.
- Setting a guidance approach for designing any facial reconstruction study in general, starting from the facial reconstruction process to a step-wise way of selecting the method of assessment according to the circumstances and data availability to each study.

STUDY SUMMARY





The Facial Reconstruction Software

- Originally designed by Dr Maria Vanezis and Dr Tim Niblett, at Glasgow University.
- Digital reconstruction of a face from a skull.
- Certain objects; triangle meshes for both the skull, and the face template together with sets of skull and face landmarks.
- Facial soft tissue depths.
- Automatically warps the face onto the skull guided by the corresponding landmarks.

Study Structure

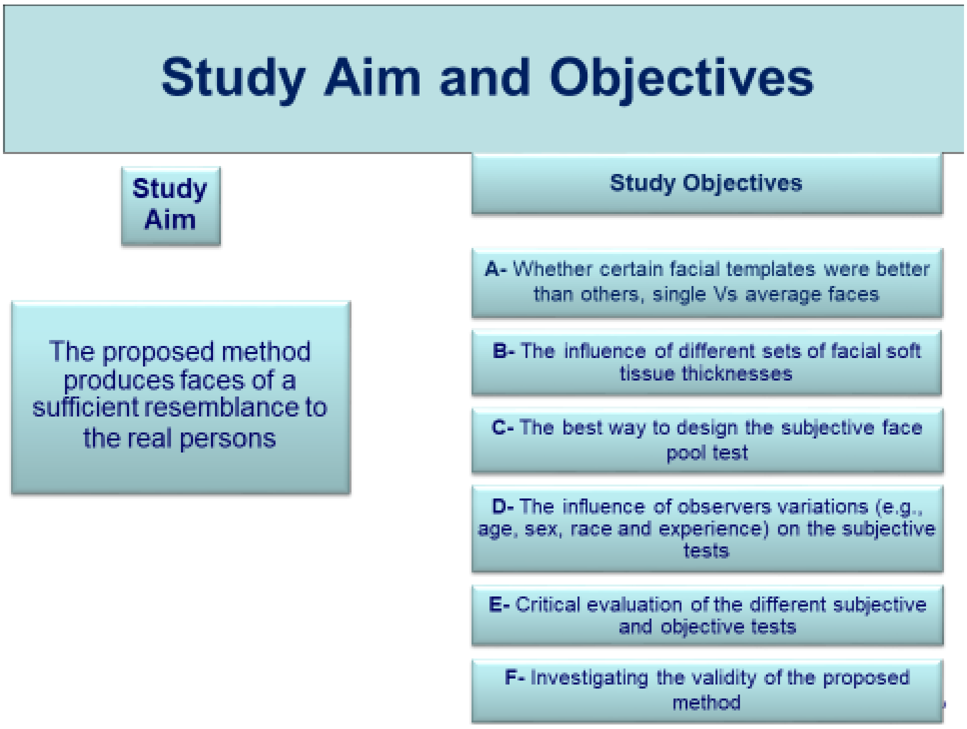
Pilot*

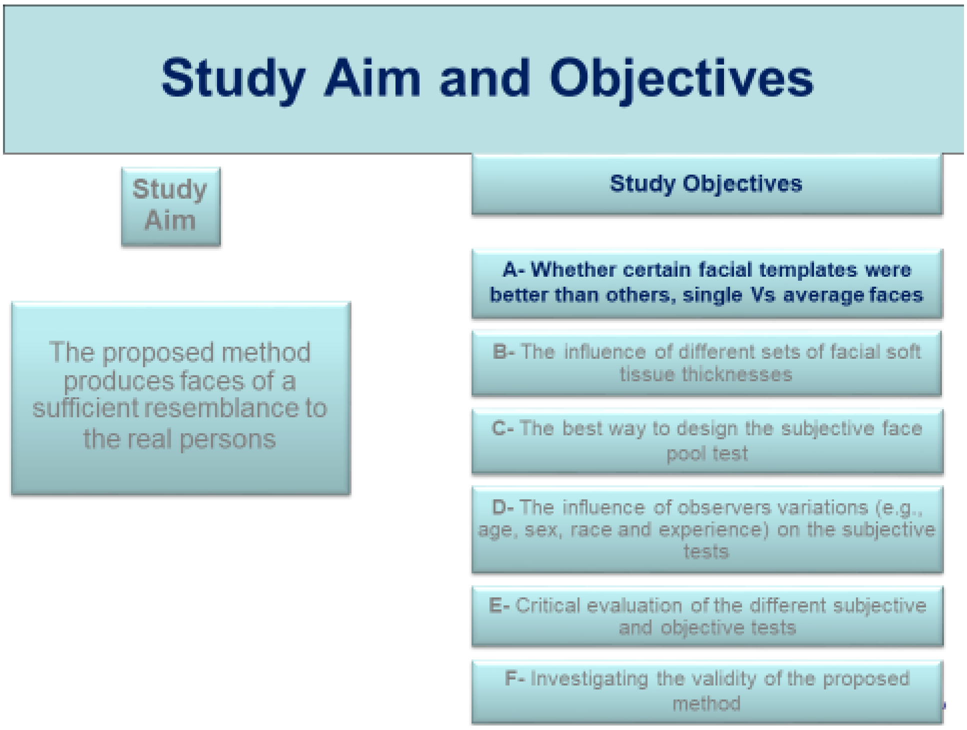
- 4 Caucasian laser scanned,
  15 Egyptian CT scanned skulls and faces.
- 2D Face pool tests: identification percentages.
  2D Resemblance tests: resemblance scores.
- Observers: 450
  Males and females
  Caucasian and non-Caucasian
  Non-experts
- Objective surface distance comparison between real and reconstructed faces.

Main

- 30 Egyptian CT scanned skulls and faces.
- 3D Face pool tests: identification percentages.
  3D Resemblance tests: resemblance scores.
- Observers: 80
  Males and females
  Egyptians and non-Egyptians
  Non-experts, and Experts (forensic pathology, facial identification psychology, forensic anthropology, and facial reconstruction)
- Objective surface distance comparison between real and reconstructed faces.

*To test the main components of the proposed method in this thesis (e.g. the facial templates, and the facial soft tissue depths).


Study Aim and Objectives
Study Aim
Study Objectives
The proposed method produces faces of a sufficient resemblance to the real persons
A- Whether certain facial templates were better than others, single Vs average faces
B- The influence of different sets of facial soft tissue thicknesses
C- The best way to design the subjective face pool test
D- The influence of observers variations (e.g., age, sex, race and experience) on the subjective tests
E- Critical evaluation of the different subjective and objective tests
F- Investigating the validity of the proposed method


Study Aim and Objectives
Study Aim
Study Objectives
The proposed method produces faces of a sufficient resemblance to the real persons
A- Whether certain facial templates were better than others, single Vs average faces
B- The influence of different sets of facial soft tissue thicknesses
C- The best way to design the subjective face pool test
D- The influence of observers variations (e.g., age, sex, race and experience) on the subjective tests
E- Critical evaluation of the different subjective and objective tests
F- Investigating the validity of the proposed method

A- Comparing Single and Average Human Faces as Templates for 3D Digital Forensic Facial Reconstruction

- 3D single and average faces.
- Certain faces can be more suitable as templates.
- Unpredictable!!
- Average faces: more accurate than most single faces.
- Safer choice.

Study Aim and Objectives

Study Aim

The proposed method produces faces of a sufficient resemblance to the real persons

Study Objectives

A- Whether certain facial templates were better than others, single Vs average faces

B- The influence of different sets of facial soft tissue thicknesses

C- The best way to design the subjective face pool test

D- The influence of observers variations (e.g., age, sex, race and experience) on the subjective tests

E- Critical evaluation of the different subjective and objective tests

F- Investigating the validity of the proposed method

B- The Influence of Facial Soft Tissue Depths on Facial Reconstructions-1

- Comparison between an old* and a recent** facial soft tissue depths data sets published approximately 30 years apart---------- No significant differences.

- Comparison between 4 different sets*** of facial soft tissue depths by modifications in the cheek region landmarks---------- No significant differences, but set (3) was the most accurate and set (4) was the least accurate.

* Rhine and Moore (1982) and Helmer (1984)
** Stephan (2014)
*** Rhine and Moore (1982) and Helmer (1984): Set (1): 40 landmarks, Set (2): 38 landmarks, Set (3): 36 landmarks, Set (4): 34 landmarks.

B- The Influence of Facial Soft Tissue Depths on Facial Reconstructions-2

- **Conclusions:**
  - Standardised set of facial soft tissue depths.
  - A previously validated set, even if old, is sufficient.
  - Standardised craniofacial landmarks set, with proper definitions and accurate descriptions of the locations and the directions (Brown et al., 2004).

Study Aim and Objectives

Study Aim

The proposed method produces faces of a sufficient resemblance to the real persons

Study Objectives

A- Whether certain facial templates were better than others, single Vs average faces

B- The influence of different sets of facial soft tissue thicknesses

C- The best way to design the subjective face pool test

D- The influence of observers variations (e.g., age, sex, race and experience) on the subjective tests

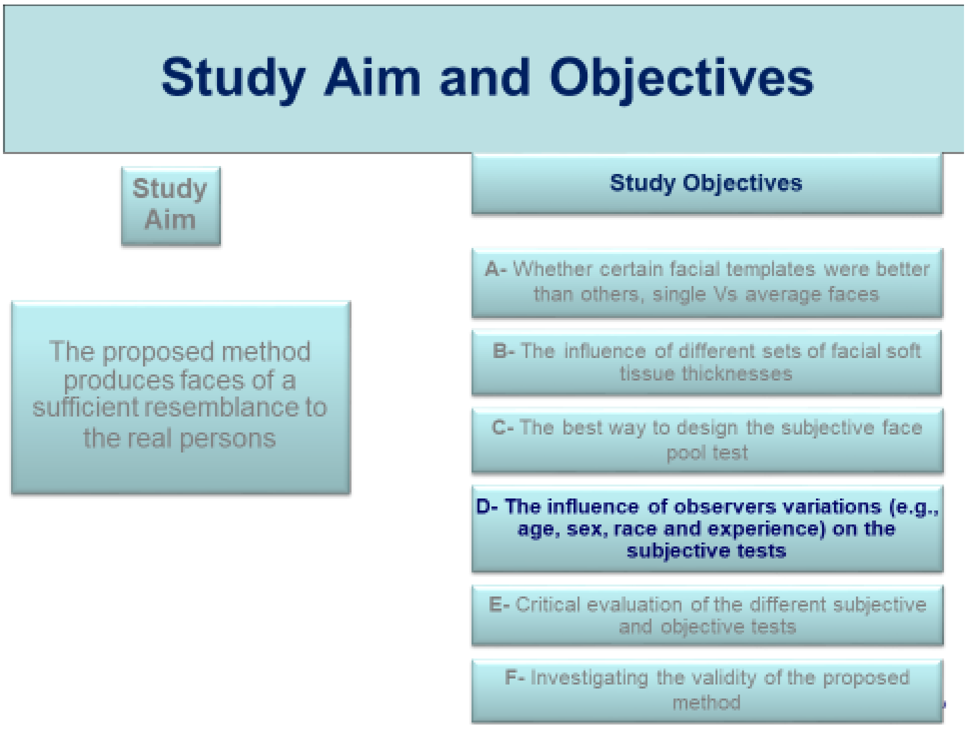
E- Critical evaluation of the different subjective and objective tests

F- Investigating the validity of the proposed method

C- Designing a Face Pool Test for Subjective Assessment of Forensic Facial Reconstruction

Higher identification rates with:

1. Faces of similar complexions and with no distracting facial features (computer generated faces) Vs photographs.
2. Neutral facial expressions Vs expressed smile.
3. Multiple facial views Vs one view only.



D- Observers' Characteristics Influence on the Subjective Tests for Facial Reconstruction Assessment

Observers: 450 in the pilot, 80 in the main.

- **Sex:** female participants better than male participants. But, not statistically significant.
- **Age:** young participants significantly better than older participants.
- **Race:** no indication that observers of a certain race would identify faces of the same race better.
- A **professional experience** in forensic anthropology, especially with FR experience, significantly improves the observers' performance in face pool and face resemblance tests.

Study Aim and Objectives

Study Aim

The proposed method produces faces of a sufficient resemblance to the real persons

Study Objectives

A- Whether certain facial templates were better than others, single Vs average faces

B- The influence of different sets of facial soft tissue thicknesses

C- The best way to design the subjective face pool test

D- The influence of observers variations (e.g., age, sex, race and experience) on the subjective tests

E- Critical evaluation of the different subjective and objective tests

F- Investigating the validity of the proposed method

E- Critical Evaluation of the Different Subjective and Objective Methods for Facial Reconstruction Assessment-1

Subjective Tests:

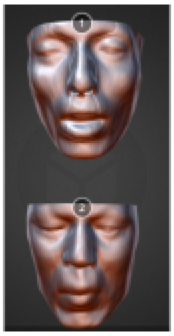
1. Face Pool Test: Identification percentage.

© Dalia Abdou

E- Critical Evaluation of the Different Subjective and Objective Methods for Facial Reconstruction Assessment-2

2. Face Resemblance Test: Overall Resemblance Score (0-10).

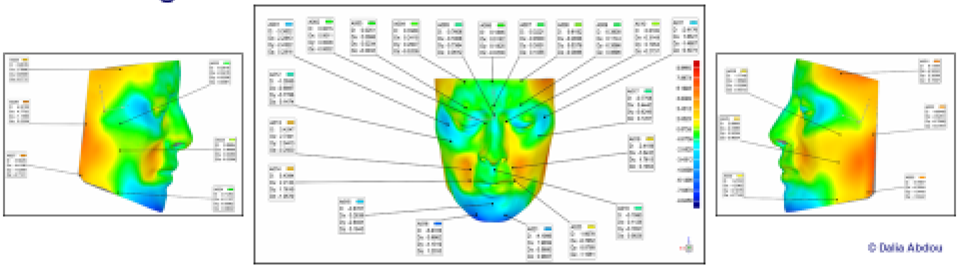
3. *Modified* Face Resemblance Test: Resemblance Score of Individual Facial Regions.



E- Critical Evaluation of the Different Subjective and Objective Methods for Facial Reconstruction Assessment-3

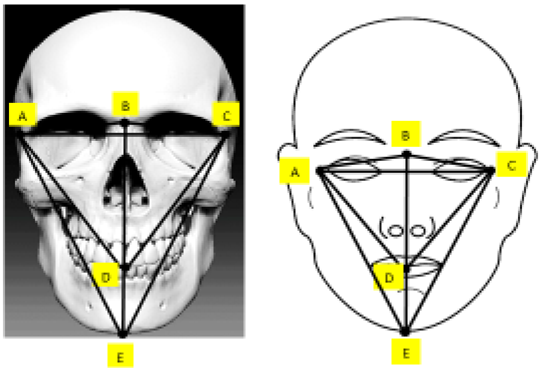
Objective Tests:

1. Objective Facial Surface Overall Distance Standard Deviation (SD).
2. *Modified* Objective Facial Surface Distance Standard Deviation (SD): at Individual Facial Regions.



E- Critical Evaluation of the Different Subjective and Objective Methods for Facial Reconstruction Assessment-4

3. *New* Objective Craniofacial Anthropometry (Linear Ratios and Angles differences).



E- Critical Evaluation of the Different Subjective and Objective Methods for Facial Reconstruction Assessment-5

TEST TYPE	TEST TITLE	TEST NAME	TEST DESCRIPTION
Subjective Assessment	Subjective Test (1)	Face Pool Test (n=20)	Identification Percentage/Rate Above Chance.
	Subjective Test (2)	Face Resemblance Test Version (1) (n=30)	Overall Facial Resemblance Score (0-10).
	Subjective Test (3)	Face Resemblance Test Version (2) (n=30)	▪ Individual Facial Regions Resemblance Score (0-10). ▪ Sum of Resemblance Score at All Facial Regions.
Objective Assessment	Objective Test (1)	Surface Distance Difference Test Version (1) (n=30)	Overall Facial Surface Distance Standard Deviation (SD) Differences.
	Objective Test (2)	Surface Distance Difference Test Version (2) (n=30)	▪ Individual Facial Regions Surface Distance Standard Deviation (SD) Differences. ▪ Sum of Surface Distance Standard Deviation (SD) Differences.at all Facial Regions.
	Objective Test (3)	Craniofacial Anthropometry (n=30)	Linear Ratios and Angles Differences.

E- Critical Evaluation of the Different Subjective and Objective Methods for Facial Reconstruction Assessment-6

- **Results:**

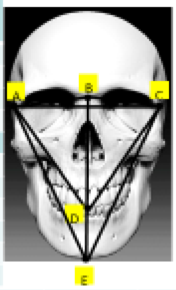
Valid

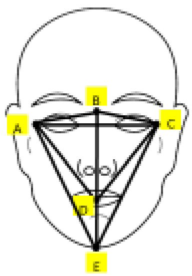
TEST	CALCULATIONS	Face Pool Test	Face Resemblance Test	Objective Surface Difference
Face Pool	ID%	N/A	0.003	0.028
Face Resemblance	Overall Resemblance Scores (V1)	0.003	N/A	0.000
Objective Surface Difference	Overall Obj. SD	0.028	0.000	N/A
Modified Face Resemblance	Sum of Facial Regions Resemblance Scores (V2)	0.012	0.000	0.017
	Orbital Bones Resemblance Score	0.023	0.013	
	Nasal Bones Resemblance Score	0.004	0.003	
	Cheek Bones Resemblance Score	0.026	0.000	0.027
	Chin Bones Resemblance Score		0.000	0.049
	Jaw Bones Resemblance Score	0.029	0.000	0.002
Modified Objective Surface Difference	Sum of Facial Regions Obj. SD		0.001	0.000
Craniofacial Anthropometry	Average Linear Ratios Differences		0.039	
	Average Angles Differences	0.008	0.008	0.008

E- Critical Evaluation of the Different Subjective and Objective Methods for Facial Reconstruction Assessment-7

- **Results (Continued):**

Linear Ratio	ID% (20 Cases)	Overall Resemblance Scores (30 Cases)	Obj. SD (30 Cases)
AB/AD	0.036	0.009	
AC/BE	0.018		
BC/BD			0.029
BC/CE			0.014
AEC	0.001	0.032	0.016
ADB	0.028	0.011	
CAE			0.038
ACE			0.002
CDB			0.039
CEB			0.018





Angles	Forehead	Orbital Bone	Nasal Bone	Cheek Bone	Chin Bone	Jaw Bone
AB/AD				0.022		0.012
AD/BD					0.030	
CE/BE		0.014				
AEC				0.048		
ADB				0.026		0.015
Objective Overall SD				0.027	0.049	0.002

E- Critical Evaluation of the Different Subjective and Objective Methods for Facial Reconstruction Assessment-8

Conclusions:

- The previously used subjective and objective tests are valid.
- The objective tests are more reliable than the subjective tests.
- The subjective face resemblance test is more reliable than the subjective face pool test.
- A step-wise approach.
- Certain facial regions (the cheek, and the jaw) can be more sensitive.

Study Aim and Objectives

Study Aim

The proposed method produces faces of a sufficient resemblance to the real persons

Study Objectives

- **A-** Whether certain facial templates were better than others, single Vs average faces
- **B-** The influence of different sets of facial soft tissue thicknesses
- **C-** The best way to design the subjective face pool test
- **D-** The influence of observers variations (e.g., age, sex, race and experience) on the subjective tests
- **E-** Critical evaluation of the different subjective and objective tests
- **F- Investigating the validity of the proposed method**

F- Investigating the Validity of the Proposed Method of Facial Reconstruction

Analysis of the subjective and objective test results:

- The subjective identification percentages in the face pool tests.
- The subjective mean resemblance scores in the face resemblance tests.
- The objective facial surface distance difference.
- These results were comparable to other facial reconstruction methods.
- The proposed method and the FR software were easily applied by a volunteer user, with no previous experience, under blind conditions---less subjective.

Superior to many other methods

Summary
What this Study Added

- Presented a quick and easy FR approach.
- First to be applied on an Egyptian population.
- Included both laser scanned and CT scanned skulls.
- Introduced subjective analysis via online 3D testing.
- Included 2D and 3D subjective assessment tests.
- Recruited a large number of forensic experts.
- Presented a comprehensive analysis of different subjective and objective methods.
- Presented an evidence based step-wise approach for FR assessment.
- Modified the current and develop new methods of assessment of the FR accuracy.

Broadly speaking, 3D forensic facial reconstruction can be performed by one of 2 approaches. The first one is the “Inside Outwards” approach which could be performed manually or digitally (Wilkinson, 2003). In this approach, the facial muscles are “built” starting from the bone surface and moving outwards towards the skin. The “Outside Inwards” approach could also be done manually, but in most cases it is performed digitally (Vanezis et al., 1989). In the latter, a reference face or head on the skull is digitally fitted/“warped” on the skull. A reference face is termed a “Face Template”, which could be a statistical face model or a scanned face (e.g. by a laser scanner, Computed Tomography). In the present study, the “Outside Inwards” approach using a reference scanned face was adopted.

The facial reconstruction software used in the present study was originally designed by Dr Maria Vanezis and Dr Tim Niblett, at Glasgow University. It adopts the concept of digital reconstruction of a face from a skull. This involves certain objects; triangle meshes for both the skull, and the face template together with sets of skull and face anatomical landmarks. Facial soft tissue depths define the distances between the skull and the face at these landmarks. The software then automatically warps the face onto the skull guided by the corresponding landmarks.

The present thesis consisted of 2 parts; a pilot and a main studies. The pilot study aimed at testing the main components of the proposed method in this thesis (e.g. the facial templates, and the facial soft tissue depths) before starting the main study. In the pilot study, 4 Caucasian laser scanned skulls and faces, and 15 Egyptian CT scanned skulls and faces were studied. Whereas in the main study, a total of 30 Egyptian CT scanned skulls and faces were studied. The reconstructed faces in both study parts were subjectively assessed via face pool and face resemblance tests. These tests were performed in 2D and in 3D formats in the pilot and the main studies respectively. The pilot study involved approximately 450 non-expert observers, including; males and females, as well as Caucasian and non-Caucasian observers. Whereas, the main study involved approximately 80 observers, including; males and females, Egyptians and non-Egyptians, as well as non-expert and experts (in forensic pathology, facial identification psychology, forensic anthropology, and facial reconstruction). Furthermore, the reconstructed faces in both study parts were objectively assessed via surface distance comparison between the real and reconstructed faces. The main aim of the present thesis was to investigate whether the proposed method in this thesis could produce faces of a sufficient resemblance to the real persons. To reach this aim, a number of objectives were sought.

A- Investigating whether certain faces were better than others as templates for the proposed method

For this purpose, a comparison was conducted between single faces and average faces, generated by digitally merging a number of single faces together, as templates. Faces reconstructed via both single and average templates were subjectively and objectively compared. The results showed that certain single faces can indeed be more suitable as templates than other single faces. However, the average faces were more accurate than most single faces tested in this study. Therefore, the average faces can be considered a safe, and accurate, choice as facial templates with no need to continue searching for the "best" single face.

B- Investigating the influence of facial soft tissue depths on facial reconstructions

For this objective, 2 comparisons were performed. The first comparison was conducted between an old and a recent facial soft tissue depths data sets published approximately 30 years apart. The first set was the combined set of Rhine and Moore (1982) and Helmer (1984) for Caucasian adults (Vanezis, 2008), and the second set was the non-population specific set of Stephan (2014). Faces reconstructed via both sets were objectively compared, and the results showed no significant differences between them.

The second comparison involved 4 different sets of facial soft tissue depths, starting with the full 40 landmarks set of Rhine and Moore (1982) and Helmer (1984) (Set 1). This set was then modified by removing a number of the cheek region landmarks, resulting in 3 other sets; Set (2): 38 landmarks, Set (3): 36 landmarks, Set (4): 34 landmarks. The results showed no significant differences between the 4 sets. However, set (3) was the most accurate and set (4) was the least accurate.

Based on the results of both comparisons, it was, concluded that it is important to have a standardised set of facial soft tissue depths, but a previously validated set, even if old, is sufficient. More importantly, a standardised craniofacial landmarks set, with proper definitions and accurate descriptions of the locations and the directions is more important for the accuracy of facial reconstruction (Brown et al., 2004), in order to standardise the facial reconstruction guidance without leaving that to the practitioner's judgement (Vanezis, 2008).

C- Designing a face pool test for subjective assessment of forensic facial reconstruction

For this purpose, a number of face pool test formats were tested. A face pool test aims at attempting to identify a target face from a pool of similar face (foils/fillers) (Moyers, 2007, Stephan and Henneberg, 2006, Wilkinson et al., 2006). Higher identification rates were observed when using faces of similar complexions that contained no distracting facial features (i.e., computer generated faces) than using photographs. In addition, a face with neutral facial expressions (i.e. passport-like photos) resulted in higher identification rates than a face of the same individual that showed an expressed smile. Moreover, multiple facial views resulted in higher identification rates than faces with one view only. Following these recommendations while designing a subjective face pool test can lead to a more reliable assessment results.

D- The influence of the observers' characteristics on the subjective tests for facial reconstruction assessment

The face resemblance test comprises direct comparison between a reconstructed face and a real face and assigning a score of resemblance to them (Stephan and Henneberg, 2006, Vanezis, 2008). The investigated subjective tests were; face pool test: by calculating the identification percentage of the target face, and face resemblance test: by calculating the overall resemblance score using a rising scale (0-10), where 0 = no resemblance, and 10 = the highest resemblance. The subjective assessment was conducted in the form of online surveys presenting 3D faces a more interactive, thus more reliable, assessment. The results of the face pool and face resemblance tests performed by the observers recruited in this thesis (approximately 450 in the pilot, and 80 in the main parts) were analysed to study the influence of the observers' characteristics (sex, age, race and professional experience) on these subjective tests. This showed that female participants were better than male participants, particularly in the pilot study. But, this was not statistically significant. In contrast, younger participants were significantly better than older participants. Moreover, the study showed no indication that observers of a certain race would identify faces of their own race better than others. Amongst the types of professional experience studied in this thesis, only a professional experience in forensic anthropology, especially with additional facial reconstruction experience, significantly improved the observers' performance in both face pool and face resemblance tests. Based on these results, it was concluded that these subjective tests are not only affected by the subjectivity in the way they are designed, but also by the observers performing them, which adds to the inherent subjectivity in, and the lower reliability of, these tests.

E- Critical evaluation of the different subjective and objective methods for facial reconstruction assessment

This was conducted in the light of the previous conclusion. In addition to calculating the identification percentage of the target face in the face pool test, and the face resemblance scores in the face resemblance test, the latter test was modified in the present thesis, for the first time in a similar research study. This modification involved using the same scale to assess the resemblance in individual facial regions (forehead, nasal bone, orbital bones, cheek bones, chin, and jaw bones).

Moreover, the investigated objective tests were; the overall facial surface distance standard deviation (SD) (Wilkinson et al., 2006, Lee et al., 2012, Short et al., 2014). This test was also modified in the present thesis, for the first time in a similar research study, by measuring the facial surface distance standard deviation (SD) at individual facial regions (similar to those assessed in the modified face resemblance test). Furthermore, a novel objective method (Craniofacial Anthropometry) was introduced and developed in the present thesis for assessing the accuracy of facial reconstructions. This was done by taking direct linear measurements from each skull as well as its respective real and the reconstructed faces. Linear ratios and angles were then calculated from these linear measurements, and the real and the reconstructed faces were compared via average linear ratios and angles differences.

All the examined cases were assessed via all the studied subjective and objective tests, and then ranked according to each test and the case ranks were then statistically correlated. In this experiment, 3 types of tests that were previously described in literature, the subjective face pool and face resemblance tests as well as the objective facial surface overall distance standard deviation (SD) were validated by statistical correlation. The results showed that these tests significantly correlated with each other, which confirm the validity of these tests as methods for assessing the accuracy of facial reconstructions.

To validate the newly modified subjective face resemblance test, the resemblance scores given to all the facial regions were summed and statistically correlated, with the validated tests, which was significant with the three previously validated methods. This shows the validity of this new method. Furthermore, it was observed that only two facial regions (the cheek, and the jaw) correlated significantly with the three previously validated methods.

In a similar way, to validate the newly modified objective test, the facial surface overall distance standard deviation (SD) at all the facial regions were summed and statistically correlated, with the previously validated tests. This was significant with two of the three previously validated methods, which shows a validity of the modified objective test.

When the real and the reconstructed faces were compared via cranial anthropometric average linear ratios and angles differences, only the angles differences showed validity as an objective method for assessing the accuracy of facial reconstructions as they significantly correlated with the three previously validated methods. Moreover, out of all the compared linear ratios and angles, only one angle correlated significantly with the three previously validated methods, which showed a higher sensitivity of this angle in reflecting the accuracy of facial reconstructions. This angle is related to the facial width and length, which can explain the higher value of this particular angle in objective assessment of the facial reconstructions. Furthermore, only two facial regions (the cheek, and the jaw) correlated significantly with the highest number of the compared linear rations and angles. Being consistent with the previous objective test results, this shows the higher sensitivity of these facial regions in comparison with the other facial regions in assessing the facial reconstructions.

The results of this experiment also showed that the subjective face resemblance test correlated significantly with all the other subjective and objective methods assessed in this study, whereas the subjective face pool test correlated significantly with a lower number of the studied tests. This showed that the former test is more reliable than the latter test.

F- Investigating the validity of the proposed method of facial reconstruction

Using the subjective identification percentages in the face pool tests, the subjective mean resemblance scores in the face resemblance tests, and the objective facial surface distance difference, the facial reconstruction method proposed in this thesis showed a comparable accuracy to other facial reconstruction methods presented in literature.

Moreover, the applicability of the proposed method and the ease of use of the present facial reconstruction software were investigated in this thesis. A number of faces were reconstructed by a volunteer user, with no previous experience, under blind conditions, helped only by a user manual prepared as a part of this study. The reconstructed faces were then compared with those reconstructed for the same cases by a more experienced user and under non-blind conditions. The results showed no significant differences between the faces reconstructed by both users. This shows the lower subjectivity of the proposed method as it is not user dependent. The presented facial reconstruction method is quick and easily transferrable from one user to another. Therefore, this method can be considered superior to many other facial reconstruction methods.

To summarise, the present thesis is the first to:

- Be applied on an Egyptian population,
- Recruit a large number of forensic experts,
- Include both laser scanned and CT scanned skulls in one study,
- Included 2D and 3D subjective assessment tests,

- Introduce subjective analysis via online 3D testing,
- Involve a comprehensive analysis of different subjective and objective methods with modifying the current and developing new methods for assessment of the facial reconstruction accuracy,
- Present an evidence-based step-wise approach for assessment of facial reconstructions in a research study as follows; the objective tests are more reliable than the subjective tests, as the latter tests are affected by the way they are designed and by the observers performing them. Thus, objective tests should be the first choice, followed by the subjective face resemblance test then the subjective face pool test as the former is more reliable. If the use of face pool tests is inevitable, the presented photos should have a similar complexion, with no distracting facial feature, neutral facial expression, and presented in multiple views.

CHAPTER 1: INTRODUCTION

The present introduction chapter includes a number of sections. The first section summarises the purpose of forensic facial reconstruction. The second section describes the history and developments of facial reconstruction. The third section describes the different methods of forensic facial reconstruction, with reference to the approach used in the present thesis. Whereas, the fourth and fifth sections describe the key aspects of the present approach of facial reconstruction, including; the acquisition of data using computed tomography and the facial soft tissue depths used for facial reconstruction.

1.1 PURPOSE OF FORENSIC FACIAL RECONSTRUCTION

In the forensic field, establishing the individual's identity is a requirement for both legal and civil (as in cases of refugees) situations. Therefore, no effort should be spared to confirm a positive identity. Forensic human identification is primarily a comparison process. In cases of the dead, this comparison is between antemortem and postmortem information to confirm, or exclude, the identity. Different identification methods carry different weights in confirming the identity and they vary depending on the case circumstances and the state of the victim. Primary identification methods are more conclusive in confirming the identity (e.g., DNA, dental records, fingerprints). However, it is not uncommon that an unknown corpse is recovered and no positive identification is reached by these primary methods (Stephan and Henneberg, 2006), most probably due to the lack of ante mortem data to complete the comparison. In that case, other less conclusive identification methods are attempted as screening or eliminating methods. Identification via the face of the deceased is considered one of the tertiary or assisting identification methods. This facial identification can be carried out directly from fresh bodies by the deceased's relatives and other acquaintances. However, this visual identification has many limitations (Saukko and Knight, 2004).

Facial identification can also be performed based on a reconstruction of the face from the skeletal remains. Anthropologically, the skeletal remains can provide circumstantial identification, via establishing the biological profile of the individual (age, sex, and race), as well as a positive identification via medical or dental records, unique ante mortem wounds or pathologies, and DNA analysis. The skull, in particular, is a good indicator of the general identification, through a number of skull traits that indicate sex, age, and race of the unknown remains (Saukko and Knight, 2004). It can also provide the basis for personal identification through facial superimposition or reconstruction (Vanezis, 2008).

Forensic facial reconstruction (FFR) is one of the tools used for human identification as a last resort when no other information is available. It is especially valuable when the remains are badly decomposed, mutilated, burnt or skeletonised, which make visual identification inapplicable. The aim of FFR is to create a facial image that bears an adequate resemblance to the deceased individual to contribute to their recognition. Facial reconstruction can be useful in assisting the search for missing persons by drawing the attention of the public, or the law enforcement authorities, to the reconstructed face to stimulate a response of a possible recognition. When a recognition is triggered with a suspected identity, the identification can then be confirmed by other methods of positive identification (Gupta et al., 2015, Phillips, 2001, Vanezis, 2008, Vanezis et al., 2000). The reconstructed face can also be superimposed onto an ante-mortem photograph of the deceased to help confirm or exclude the identity (Shahrom et al., 1996).

Phillips (2001) presented six forensic cases of victims of suspected unnatural death in South Africa, where identification could not be reached by other means. The cases were positively identified via facial reconstruction of the skeletal remains conducted by the author. Also, many cold forensic cases were solved with the aid of facial reconstruction (Chron News, 2008, Government Technology, 2013, KSBW 8 News, 2015, Woodstock Patch, 2016). In addition to forensic applications, facial reconstruction has archaeological applications such as; facial reconstruction of Egyptian mummies (Attardi et al., 1999, Baldock et al., 1994, Cesarani et al., 2004, Davy et al., 2005, Hughes, 1996, Wilkinson, 2003), as well as other mummies; such as, The Spirit Cave Man, in 2008 (Mathilda's Anthropology Blog, 2011), Ötzi the Iceman, in 1993 by Artist John Gurche, in 1998 by Professor Peter Vanezis, and in 2011 by paleontological artists, Alfons and Adrie Kennis (Mummy Tombs, 2011), King Tutankhamun, in 2005 (BBC News, 2005), and King Richard III, in 2013 (University of Leicester, 2013).

Turning to the question of admissibility of forensic facial reconstruction as expert witness evidence of identification, it is important to understand the principle of the Daubert standard. This was set by the Supreme Court, as Rule 702 of the Federal Rules of Evidence regarding the admissibility of expert witness testimony in 1993 following the legal proceedings of Daubert v. Merrell Dow Pharmaceuticals. It was then amended in 2001. Daubert standard involves a number of requirements to ensure the expert witness is sufficiently qualified by "knowledge, skill, experience, training, or education" to give expert testimony. In addition to having the appropriate scientific knowledge, the expert witness should be able to assist the court in understanding the issue of question that lies within their expertise, a role referred to as "gatekeeping". Thus, a number of factors were suggested to help judges evaluate the reliability of scientific evidence including; empirical testability, being based on sufficient facts or data; following reliable principles and methods, having an acceptable error rate, and being acceptable by publishing and peer reviewing, and by the scientific community in general.

Various scientists have employed different approaches of facial reconstruction with no single standardised and scientifically acceptable method of facial reconstructions. Thus, different facial reconstructions can be generated from the same skull (Davy et al., 2005), which carries a large degree of subjectivity. It is even claimed that both conventional and computerised facial reconstruction techniques share similar artistic subjectivity (Abate et al., 2004). So, although the facial reconstruction can reveal what the person might have looked like during life, the final appearance may vary because of the subjectivity involved in the facial reconstruction methods employed. Therefore, according to Daubert standard, facial reconstruction is not legally admissible evidence as a positive identification method, but rather an assisting identification method (Vanezis, 2008).

Thus, while it is generally agreed that facial identification by methods like image superimposition is more useful in eliminating or disproving than confirming the identity, superimposition might have different weights in court with different cases (Huete et al., 2015, Vanezis and Brierley, 1996). The 1935 Buck Ruxton murder case was one of the famous historical cases where image superimposition was accepted by the court as supportive to identification (Glaister and Brash, 1937). In that case, images of two partially macerated skulls were superimposed with ante mortem photographs of the suspected victims (Mrs. Ruxton and her maid Mary Rogerson) (Figure 1). Yet, the superimposition was not admissible as a method of positive identification standing alone, which confirms the ongoing debate among researchers (Vanezis, 2008).

a b c

Figure 1: The 1935 Buck Ruxton murder case: Portrait of Isabella Ruxton (The common law wife of and victim of hanged killer Buck Ruxton) used for superimposition of skull (a), Negative of skull in the portrait A position (b), Positive portrait and negative skull superimposition (c) (Glaister and Brash, 1937).

1.2 History and Developments of Forensic Facial Reconstruction

The 1952 discovery of ten defleshed human crania, from the Pre-pottery Neolithic B Levels (c.7500 – 5500 BC), by Kathleen Kenyon, the director of the British School of Archaeology in Jerusalem, refers to one of the earliest practices of facial reconstruction. With no clear purpose, the faces of these crania were reconstructed in plaster directly over the skull with shells replacing the eyeballs. Death masks were found in the Jordan Valley where clay was applied to the dry skulls of the dead ancestor as a symbol of worship. In 1884, Schaaffhausen reconstructed a woman's head using arbitrarily chosen soft tissue thicknesses (Gupta et al., 2015, Parks et al., 2013, Taylor, 2001, Vanezis, 2008, Verzé, 2009, Wilkinson, 2005).

Facial reconstructions from the skulls of alleged famous historical people was documented in the late nineteenth century when reconstructions from skulls found in their tombs were conducted by anatomists and then compared to portraits and death masks to verify the authenticity of these skulls. For example, Welcker reconstructed the faces of Schiller and Kant in 1883, and the face of Raphael in 1884, and Kollman and Büchly reconstructed the face of Dante in 1898. His reconstructed Bach's face in 1895, which was credited as the first scientific facial reconstruction (Gupta et al., 2015, Parks et al., 2013, Taylor, 2001, Vanezis, 2008, Verzé, 2009, Wilkinson, 2005).

In 1946, the anthropologist Dr. Wilton M. Krogman laid the real foundation for facial reconstruction, by conducting an experiment to investigate whether the shape of the face during life could be predicted from the skull. The reconstructed face carried a sufficient resemblance to the original face (Taylor, 2001). Krogman, then, suggested five principles to be followed for facial reconstruction; the relation of eyeball to orbit, the shape of nose tip, the ear location, the mouth width, and the ear length (Gupta et al., 2015).

Since the emerge of the facial reconstruction from the skull at the end of 19th century, it has greatly developed and improved by the evolving technological advancements, such as; scanning devices, computer software programs and 3D graphics. Traditionally, to manually reconstruct the face, the skull had to be defleshed or completely desiccated and clay, plasticine or another modelling material was then used to reconstruct the face directly onto the skull or its cast or replica (Vanezis et al., 1989). The corner stone of facial reconstruction development was the introduction of 3D computerised forensic facial reconstruction by Vanezis et al. (1989) by fitting a 3D image of a facial template onto a 3D image of the skull employing special computer software. As 3D computerised facial reconstruction substituted the conventional manual methods by many experts (Shahrom et al., 1996, Vanezis et al., 2000, Wilkinson et al., 2006, Vanezis, 2008), surface scanners were then used to obtain a 3D image of the skull and then import it into the 3D reconstruction software, a step that preserves the skull and allows its future examination if needed (Vanezis, 2008, Vanezis et al., 1989). More developments in the computer programming have expanded to generate a statistically calculated facial model or digital mesh and use it as a template fit it onto the skull (Kähler et al., 2003, Andersson and Valfridsson, 2005, Claes et al., 2006, Vandermeulen et al., 2006).

Other approaches of forensic facial reconstruction have been computerised to mimic the manual facial reconstruction by digitally applying pre-modeled individual facial muscles in addition to other facial features (ATOR, 2012, Lee et al., 2012, Wilkinson et al., 2006). Likewise, measuring the facial soft tissue depths has greatly developed from using knives or needles inserted directly into cadavers' faces, to non-invasive measuring tool by medical imaging devices. These include; cranial x-rays "craniography" (Aulsebrook et al., 1996, George, 1987, Smith and Throckmorton, 2006), ultrasound (Aulsebrook et al., 1996, De Greef et al., 2005, El-Mehallawi and Soliman, 2001, Smith and Throckmorton, 2006), computed tomography (Phillips and Smuts, 1996, Shimofusa et al., 2009), and magnetic resonance imaging (Sipahioğlu et al., 2012).

Another example of the continuing research in facial reconstruction lies in the different methods of subjective and objective assessment of the accuracy of the reconstructed faces. It was reported that, in 1913, H. Von Eggeling was the first to attempt validating the accuracy of a facial reconstruction produced from a cadaver's skull by comparing it to the cadaver's death mask (Verzé, 2009, Parks et al., 2013). Since then, numerous researchers have attempted to assess the accuracy of the facial reconstructions they produced to validate and compare different facial reconstruction techniques (Shahrom et al., 1996, Vanezis et al., 2000, Vanezis, 2008, Claes et al., 2006, Claes et al., 2010, Vandermeulen et al., 2006, Vandermeulen et al., 2012, Wilkinson et al., 2006, Lee et al., 2012, Quatrehomme et al., 1997, Kähler et al., 2003). Subjective assessment of forensic facial reconstructions has been performed mainly by two types of tests; a face recognition test referred to as "*face pool test*", and a face resemblance scoring test referred to as "*face resemblance test*" (Parks et al., 2013). Further developments introduced the objective comparison between the digital real and reconstructed faces, including image superimposition (Shahrom et al., 1996, Vanezis and Brierley, 1996, Curry et al., 2001, Jayaratne et al., 2012), as well as mathematical surface distance comparison (Claes et al., 2006, Vandermeulen et al., 2006, Wilkinson et al., 2006, Lee et al., 2012, Short et al., 2014, Decker et al., 2013).

1.3 METHODS OF FORENSIC FACIAL IDENTIFICATION

While a perfectly symmetrical face has been seen by artists as an indicator of idealised beauty and worth achieving by plastic surgeons for aesthetic reasons, forensic artists perceive things differently (Taylor, 2001). Forensic facial reconstruction is a way of recreating a face from the skull alone based on the influential relationship between the head soft tissues (i.e. muscles, fatty tissue, and skin) and the underlying hard tissue (i.e. the skull bone and the nasal cartilage), together with the colour and texture of the skin and hair, and the facial features (eyes, nose, and mouth) (Phillips, 2001, Vanezis, 2008). The face is "built" from the skull after anthropological examination of the unknown skeletal remains is done to establish the biological profile (i.e. age, sex, and race) (Phillips, 2001). Therefore, a thorough examination of the skull is required to identify any skull features that might affect the facial appearance, such as; the mandible and dentition, symmetry of the nasal bones, and wear of the occlusal surfaces, as well as bone pathologies, wounds or unusual landmarks (Vandermeulen et al., 2006, Vanezis, 2008, Vanezis et al., 1989). It is suggested that facial reconstruction research would be best delivered with the aid of a multi-disciplinary team (e.g., forensic pathologists, anthropologists, anatomists, radiologists and mathematicians) (Tilotta et al., 2009).

Numerous approaches developed by researchers over years in attempts to reach an optimum method. Therefore, it would be difficult to include all of them in this introduction. So, while describing the main approaches developed for facial reconstruction, only some of the commonly cited examples were mentioned.

1.3.1 Two-Dimensional Facial Reconstruction

Two-dimensional facial reconstructions started as manual hand drawn portraits onto 2D skull images, such as; skull photographs (Taylor, 2001), and radiographs (George, 1987). Taylor (2001) reported limited references of facial soft tissue depths for 2D facial reconstructions, compared to those available for 3D sculptures. Therefore, the author introduced a modified method for producing 2D reconstructions by combining an anatomical method with soft tissue depth estimates at certain markers positioned on the sides of the drawn skull. The markers were then joined creating a contour that is used as a reference for drawing the face. This approach included a number of successive phases starting with placing the depth markers on the skull, which is then photographed. The next steps involve developing the individual facial features, drawing the facial contour, and then placing the eyeball, the eyebrows, the nose and the mouth (in frontal and lateral views). Finally, the ears, hair, neck and other details were drawn (Taylor, 2001). George (1987) applied the technique suggested by Taylor (2001) for 2D facial reconstruction but with tracing skull lateral radiographic x-rays (craniographs), instead of photographs. The distances between the skulls and soft tissues traces were, then, measured to establish the midline soft tissue depths means and facial angles. A cephalometric analysis was conducted to determine the skull type from its tracing, and points were, then, plotted on the skull tracing and connected to outline the face. The average dimensions of the nose, lips, and chin were also determined. Finally, the face was "humanised" by adding additional features (e.g., tone, eye, hair patterns, and age lines) guided by the relevant anthropological data (sex, age, and race).

Although this manual 2D facial reconstruction technique led to the identification of numerous skulls, it required a larger degree of artistic ability compared to the anatomic and anthropological knowledge (Abate et al., 2004, Gupta et al., 2015, Taylor, 2001, Vanezis, 2008). This is because most cases required some degree of rendering to capture the subtle facial expression and “humanise” the drawn face, which made the final product rather artistic than accurate. However, the possibility to draw multiple variations from the same face (e.g., different nasal angles, lip positions, facial built, facial hair styles, etc.) provided alternatives to the same reconstruction that might increase the chance of recognition (Taylor, 2001). In addition, less artistic skills could be achieved using craniographs (George, 1987) or via computerised 2D facial reconstruction, which saves time, and produces multiple face varieties from the same skull. The latter method comprises using templates of facial components selected from a database, positioned onto a digital drawing of the skull (Miyasaka et al., 1995). Examples of these 2D facial reconstruction computer software programs include Computer Assisted Recovery Enhancement System (CARES) and Forensic Anthropology Computer Enhancement System (FACES). These systems work by digitising skull radiographs, photographs or images, and generating a digital version of them to allow the 2D facial reconstruction

(Gupta et al., 2015, Taylor, 2001, Vanezis, 2008). Further developments were introduced to 2D facial reconstruction, such as; Face Imaging Reconstructive Morphography (FIRM) technique for the production of objective facial composites based on cephalometric measurements, and Identi-Kit™ database of facial components (face contour, eyes, nose, lips, chin, etc.) (Miyasaka et al., 1995, Vanezis, 2008, Vanezis et al., 2000).

George (1987) pointed out the value of combining 2D and 3D formats by using the 2D reconstructions as "blueprints" for the 3D facial reconstructions, or photographing the 3D reconstructions in 2D from any angle and under varied lighting conditions to examine any cranial points. Moreover, Wilkinson (2008) suggested to increase the chance of recognition of the facial reconstructions presented to the public in 3D, and did not receive the required response, by producing 2D images of the reconstruction with different hairstyles, skin and eye colour, etc. This can be achieved using computer software (e.g., Adobe Photoshop™). However, caution is needed as this may entail adding additional and often uncertain facial details (Wilkinson, 2008).

1.3.2 Three-Dimensional Facial Reconstruction

Three-dimensional technology can be of utmost assistance to the forensic anthropologist, not only for performing the facial reconstruction process, but also for measuring the facial soft tissue depths, as well as for anthropological assessment of the remains (Attardi et al., 1999, Cesarani et al., 2003, Cesarani et al., 2004). The 3D devices can better examine the topography of the facial surface features. This helps distinguish between the different individual facial shapes from their contours as well as the underlying skeletal structures, and the tissue layers in between (Smith and Throckmorton, 2006). Three-dimensional facial reconstruction also started as a manual technique (i.e. sculptures) and then became computerised in the late 1980s by Vanezis et al. (1989) in the Medical Graphics Workstation at University College London. A laser scanner and a video camera were used to establish a database of facial templates of living subjects, and then these facial templates were warped onto the digitally scanned skulls for facial reconstruction. Compared to the manual facial reconstruction, the computerised 3D facial reconstruction is faster, more efficient, and more flexible with the ability to produce multiple variants of the same reconstructed face while preserving the original skull specimen. Additional facial features can also be added via computer programs to produce human-like faces (Gupta et al., 2015, Lee et al., 2012, Vanezis, 2008, Vanezis et al., 1989, Vanezis et al., 2000, Wilkinson, 2005, Wilkinson et al., 2006). This computer facilitated facial reconstruction can also be easily learnt in a short time (Shahrom et al., 1996). This is time and cost saving as the facial reconstruction can be conducted in more forensic facilities with the cost mostly directed to that of the software (Davy et al., 2005). Thus, with the possibility of recruiting more practitioners, 3D facial reconstruction is particularly useful in mass disaster situations for screening identification of the victims.

In many cases, however, the original skull specimen cannot be accessed, neither directly nor via 3D imaging, and only 2D skull are images (e.g., radiographs, photographs, and craniometrics) available. Therefore, 3D reconstruction can be attempted from the available 2D images. For example, multiple 2D views can be aligned and registered, using cranial points, to provide a template for the 3D skull reconstruction. Then, via computer modelling software, a 3D model is generated using the registered 2D images, with extrapolating (i.e. estimating) the surface morphology between the views. The generated 3D skull model can then be used for the facial reconstruction. This technique, however, entails a degree of assumptions and estimations with possible loss of the surface details. Therefore, for more accurate 3D model formation, the highest possible number of 2D views would be required (Wilkinson, 2007). Curry et al. (2001), developed a method of 3D craniofacial mapping from lateral and frontal stereo x-ray images of the cranium (Cephalograms), using tie points (radio-opaque markers) on the face and the teeth to compute the coordinates of the 3D model.

On the other hand, the physical skull can be scanned using a surface scanner. From the scanned skull, a 3D copy can be generated for a facial reconstruction in a location remote from the original specimen (Decker et al., 2013, Lynnerup, 2002). This can be helped by the possibility of rapid prototyping to generate a physical 3D skull replica from the digital copy via stereolithography (i.e. 3D printing) (Lynnerup, 2002), for example form laser scanners (Cesarani et al., 2004, Shahrom et al., 1996, Vanezis et al., 2000). Moreover, if the skull is fragmented, each piece can be scanned separately and then the skull can be digitally reassembled in 3D, with possible remodelling of any missing fragments. This is more efficient and time saving than physical reassembly of the fragmented pieces. However, some of the details of the digital fragment edges may be lost due to resolution problems, which would necessitate access to the original fragments (Wilkinson, 2007).

In general, the main facial reconstruction approaches adopted by researchers can be "inside outwards" conducted by building the facial muscles starting from the bone surface (Andersson and Valfridsson, 2005, ATOR, 2012, Lee et al., 2012, Wilkinson et al., 2006). In contrast, facial reconstruction can be conducted "outside inwards" using a face template placed from outside and "warped" onto the skull using landmarks (Moyers, 2007, Vanezis, 2008, Vanezis et al., 1989, Vanezis et al., 2000). Some researchers combined features of both approaches (Kähler et al., 2003). Only a number of the commonly cited studies of these different approaches were discussed in this section as examples of each approach.

1.3.2.1 The "Inside Outwards" Approach: Facial Reconstruction by Building the Facial Muscles

In this approach, the practitioner starts from the bone surface of the skull building the facial muscles one by one and moving outwards towards the skin, where a layer is added over the musculature representing the subcutaneous fat and skin. This is performed following the "Russian or Gerasimov" school, or the "Manchester or Combination" school. The former was first adopted by Gerasimov M. M. it is a morphoscopic approach, where the facial muscles, fat and skin are reconstructed from "inside outwards" in an anatomical pattern with no consideration to the soft tissue depths measurements between the skull and the face skin (Gupta et al., 2015, Vanezis, 2008, Verzé, 2009, Wilkinson, 2005). Whereas the "Manchester" technique was first adopted by Prag J., Neave R. A. H., and Wilkinson C. It takes into account both the facial muscles reconstruction and the soft tissue thicknesses that are dependent on age, sex, race, and body build, taken from various published literature, and guided by landmarks at certain anatomical locations (Wilkinson, 2005). Researchers following the "Manchester" school use soft tissue thickness measurements.

Different guidelines has been suggested by many researchers (Vanezis et al., 1989, Wilkinson, 2008) for reconstructing the different facial features (e.g., eyes, nose, ears, etc.). It is argued that these guidelines depend on artistic skills in addition to the subjectivity in deciding which guidelines to follow, which introduces unknown quantities of error to the facial reconstruction (Stephan and Henneberg, 2001). Moreover, these guidelines have limited predictive accuracy with no verified method for every feature of the face (Hayes, 2016). However, anthropological examination may reveal information that suggest the shape of these features. For example, if the skull is believed to be of an Asian race, facial features such as the eye folds characteristic should be considered (Wilkinson, 2007, Wilkinson, 2008). The facial reconstruction is then finished off by "fleshing" the face by adding clay until the tissue thickness markers are covered. External features are then added, such as; hair, skin colour, racial traits, glasses, and age-related facial details that are predicted from the skull morphology (e.g., eye bags, neck sagging, jaw line softness and eyelid drooping). Thus, unless indicated by associated evidence, estimation of the unknown facial features (e.g., too much ageing, wrong eye and hair colour, etc.) should absolutely be avoided as this will lead to false impressions, and discourage recognition (Wilkinson, 2007, Wilkinson, 2008).

Studies have emerged to digitise the "inside outwards" approach of facial reconstruction using computer software programs (ATOR, 2012, Lee et al., 2012, Wilkinson et al., 2006). Wilkinson et al. (2006) designed their own 3D modeling software (Freeform® Plus software) to mimic the manual "Manchester" method of facial reconstruction by employing a virtual sculpting technique with haptic touch-based feedback to be able to feel the skull surface during the analysis. This helped provide some important skeletal details for the facial reconstruction. Furthermore, the authors established a data bank of pre-modeled facial muscles. Each muscle is selected and rebuilt as accurately as possible following the anatomical guidelines of muscles' origins and insertions. This is done using pegs placed onto the skull surface at the corresponding anatomical sites, with their lengths derived from the facial soft tissue depths data at the respective anatomical sites. This software has been used for facial reconstruction in a number of studies (Wilkinson et al., 2006, Lee et al., 2012, Short et al., 2014).

On the other hand, researchers attempted the "inside outwards" approach of facial reconstruction using a free open source software program originally designed for visual effects, graphic designs, 3D modeling and animation. For example, Davy et al. (2005) performed a facial reconstruction of an Egyptian mummy using software (3ds max®) using frontal and lateral radiographs of the skull and statistically formed facial musculature. Then, FaceGen Modeller software (Face Gen®) package was used to add texturing of the face. In a similar way, a group of Brazilian archaeologists (archaeological research Arc-Team), led by a 3D designer Cícero Moraes, presented an all free process of facial reconstruction (ATOR, 2012). The process was entirely free starting from the 3D skull segmentation (i.e. extraction) from the medical computed tomography (CT) files using the free InVesalius software (InVesalius®). The 3D skull mesh extracted from CT was also edited using a free Meshlab software (Meshlab®). Then, the free Blender software (Blender®) was used to import the skull mesh and apply previously designed facial muscles guided by landmarks of certain facial soft tissue depths and 3D modeling. Finally, the free Gimp software (GIMP®) was used to add the textures and treat the image as required. This approach was presented in 2012 (Latinoware, 2012), and used for the archaeological facial reconstruction of St. Anthony of Padua (Blendernation, 2014), and the faces of hominids through evolution (Blendernation, 2013).

It should be noted that this "inside outwards" approach requires extensive knowledge of the facial muscle s anatomy (Gupta et al., 2015, Vanezis, 2008, Verzé, 2009, Wilkinson, 2005).

1.3.2.2 The "Outside Inwards" Approach: Reconstructing the Face from a Template

In this approach, the face is reconstructed as a unit without building the individual facial muscles. Out of the manual facial reconstruction techniques, the morphometric "American" school, adopted by Gatliff B. P., is the closest to the "outside inwards" approach. In this school the face is built with the facial soft tissue thicknesses as a bulk, without much regard to the details of the underlying musculature anatomy, and guided by pegs with lengths corresponding to the average facial depths at certain anatomical landmarks (Gupta et al., 2015, Vanezis, 2008, Verzé, 2009, Wilkinson, 2005).

Moreover, the “outside inwards” approach was adopted when the 3D facial reconstruction was first computerised by Vanezis et al. (1989). The face is reconstructed from the unknown skull from “outside inwards” by virtually “wearing” a reference face/head, like an elastic face mask, that is digitally fitted onto the studied skull. In general, this template could be a *craniofacial model* (i.e. involving a full head) (Attardi et al., 1999, Claes et al., 2010, Jones, 2001, Moyers, 2007, Nelson and Michael, 1998, Parks et al., 2013, Quatrehomme et al., 1997, Vandermeulen et al., 2013, Vandermeulen et al., 2006), or a *facial model* (i.e. involving a face only). The latter could be a scanned face (Shahrom et al., 1996, Vanezis, 2008, Vanezis and Vanezis, 2000, Vanezis et al., 1989), or a deformable facial model mathematically or statistically calculated from the skull (Andersson and Valfridsson, 2005, Claes et al., 2006). The craniofacial model or the scanned facial model should be anthropologically similar (i.e. of the same race, sex and age group) to the skull. The computerised “outside inwards” approach, the reconstruction can be performed either by a dense, or a sparse method.

(I) Facial Reconstruction via the Dense Approach: using Craniofacial Templates

This technique of the “outside inwards” approach takes into account the craniofacial relationship as a whole dense unit, and is not guided by certain facial depths tables. A craniofacial model in the form of a full head (i.e. a skull and a face together) is used as a reference. Face-specific deformations “volume deformations”, which are the geometric transformation from one volume (face) to another, are calculated by different algorithms from the reference head, and then applied to the unknown skull to predict the unknown face shape (Attardi et al., 1999, Claes et al., 2010, Jones, 2001, Moyers, 2007, Nelson and Michael, 1998, Parks et al., 2013, Quatrehomme et al., 1997, Vandermeulen et al., 2013, Vandermeulen et al., 2006). Studies using the dense approach can be classified into two main categories, landmark-independent and landmark-based, according to whether the reconstruction is guided by landmarks or not.

Examples of the *landmark-independent techniques* include Quatrehomme et al. (1997). In this study, the authors used salient "crest" lines on a CT scanned skull to calculate the volume deformation from one cadaver head "reference" to another cadaver head "test". This method was based on transforming a whole head model (skull and face) into another relying on the skull and face morphology, rather than a certain set of craniofacial landmarks. However, since this study involved one case only, a larger sample with more facial reconstructions are needed to validate this method. Moreover, the necessity of keeping accurate skull-face registrations, which was problematic in this study, should be taken into account. Similarly, Nelson and Michael (1998) used control points defined by key anatomical landmarks on a reference and a test skulls scanned by CT. These control points consisted of "disc fields", each consisting of 3 points (a centre and 2 vectors), for each studied volume (skull). In a database, reference heads were classified and selected according to sex and age group (e.g., young adult, mature adult and senile). The selected reference heads were further analysed, in a tree-shaped structure, to determine the head with the closest features matching the test head by comparing the spatial distribution of their control points. Jones (2001) used another algorithm by calculating volumetric "distance fields" data from one reference head selected from a database, to reconstruct the face of an unknown CT scanned skull, using corresponding points of similarity for registration. In addition, the author mapped all skull points to avoid possible inaccuracies that might result from interpolation between restricted points. However, the author acknowledged that this work had a number of limitations which needed further improvements.

Moreover, RE/FACE (Reality Enhancement Facial Approximation by Computational Estimation) is a software program designed by the Federal Bureau of Investigation (FBI), and the General Electric Company (GE). It has been used for the facial reconstruction of forensic cases following the landmark-independent techniques employing medical imaging and statistical techniques. From a large established database of human CT head scans, the software identifies a composite that represents an "average face" matching the skull, then derives a statistical facial template for the unknown skull (Turner et al., 2005). RE/FACE follows the dense approach using the "crest" lines for registering the unknown skull with the selected/known head (skull and face), then applies the deformation from the known head to the unknown skull for generating the warped face. The technique employs an automated and objective method to reduce the subjectivity as it did not require manual measurements or landmarks placing (Turner et al., 2005). Although this method is promising, the software was validated by FBI researchers only (Moyers, 2007, Parks et al., 2013, Turner et al., 2005), but it is not available to non-FBI researchers. Moreover, this software depends on a large database of head CT scans (Parks et al., 2013).

On the other hand, in the *landmark-based techniques*, the facial reconstruction is similar to the above discussed landmark-independent technique in that it also uses a facial/craniofacial model by calculating the volume deformation of a reference head and applying it to the unknown skull, without the use of facial depths tables (i.e. a dense technique). However, unlike the previous landmark-independent technique, it takes into account a number of craniofacial landmarks as a guide for the registration between the reference model and the unknown skull, which is a point of similarity to the sparse approach discussed below. For instance, Attardi et al. (1999) conducted a facial reconstruction of a CT scanned ancient Egyptian mummy, using a 3D European full head model as a reference that was registered with the mummy's head by a number of anatomical landmarks. Then, the cranial features were tracked by finding corresponding sets of characteristic points for each feature. The deformation of the reference volume to the mummy volume was then calculated via obtaining a "scattered motion field", which was then diffused (applied) to the whole head reference volume.

Vandermeulen et al. (2006) presented another example of the transformation algorithms applied for facial reconstruction using a full craniofacial model guided by a number of landmarks. The authors conducted a fully automated facial reconstruction using multiple CT scanned reference heads (skulls and facial surfaces). Each reference head was transformed into "signed distance transform (sDT)", representing the distance between the closest points on the skull and face surfaces. To reconstruct a face of a target skull, the sDT of the target skull was calculated and then warped onto the sDT of the reference skull. Then, the deformation between the two warps was applied to all the reference head sDT maps in the database. These deformed reference heads were then averaged and their arithmetic "average" was considered as the facial reconstruction of the target skull. A leave-one-out approach was adopted to validate the facial reconstructions by using each subject in the database in turn as the target, which was reconstructed from the remaining reference subjects. The sum of squared differences (SSD) between the corresponding elements in the compared facial surfaces was then calculated. The smaller the SSD value, the more similar two surfaces are. The reconstructed heads were then ranked according to these SSD. However, the limitations of this method were attempted to overcome in subsequent studies by the authors. For example, Claes et al. (2010) proposed a similar refined statistical craniofacial model, calculated from an extended facial database. The reconstruction was guided by anatomical craniofacial landmark positions associated with tissue depths, as well as age, BMI and sex values. Similarly, in order to increase the variabilities of the reconstructed faces, the reference heads database was continuously upgraded, extended and subcategorised according to the subject specific attributes (age, BMI, sex) (Vandermeulen et al., 2013).

(II) Facial Reconstruction via the Sparse Approach: using Facial Templates

In this technique, a face is used as a template that is warped onto the unknown skull, from outside inwards, guided by a number of craniofacial landmarks identified at certain anatomical locations, hence the name “sparse”. It starts with registering the selected facial template, represented by a 3D digital triangular mesh, with the target (unknown) skull, represented by another 3D digital triangular mesh. This registration ensures that the two meshes have the same placing, orientation, dimensions and resolution. The process is conducted via specific facial reconstruction software programs, which provide facilities to view the digitised skulls and facial templates as 3D scans, and give the user the ability to interactively manipulate the images as required (Vandermeulen et al., 2006, Vanezis, 2008). The following step is warping the face model onto the unknown skull like an elastic mask. The term “warping” refers to a statistical process that works by defining a warping or coordinate mathematical transformation function to minimise the distances between the corresponding points in the aligned (registered) images. Interpolating (i.e. estimating the values of the facial depths at the remaining parts of the registered meshes based on the known depths values of the landmarks), and extrapolating (i.e. filling any gaps) the warping then follow (Vandermeulen et al., 2006). The warping process is different from simple rigid transformation algorithms, although both start by identifying a set of points (landmarks). In the latter algorithm, an object can rotate and slide on a target object, preserving the source (first) object, and the transformation is simpler to calculate, whereas in the warping algorithm the source approaches the target. The selection of either method depends on the application. For example, the facial reconstruction of a skull requires warping, rather than rigid transformation, in order to specify a new position for each defined point on the reference skull due to its complex shape (Turner et al., 2005). The warping process is an automatic function of a number of software programs that are used for facial reconstruction (Claes et al., 2006, Claes et al., 2010, Vanezis, 2008).

The used facial template in this approach is either a specific scanned face anthropologically similar to the unknown skull (Shahrom et al., 1996, Vanezis, 2008, Vanezis and Vanezis, 2000, Vanezis et al., 1989), or a statistical face model calculated from the unknown skull (Andersson and Valfridsson, 2005, Claes et al., 2006). Andersson and Valfridsson (2005) developed a mathematical method of facial reconstruction using CT scanned skull images and the open access (3ds max®) software. This method involved deforming a cylinder mesh into a facial model calculated from the underlying skull, and then warping this facial model onto the skull using facial soft tissue depths at certain anatomical landmarks. Claes et al. (2006) used a statistically deformable facial model for facial reconstruction averaged from a database of faces of a large and diverse population. The generated statistical facial model was warped onto the skull guided by virtual dowels at certain craniofacial landmarks. Although this model relied on sparse craniofacial landmarks, it densely combined population-specific tissue depths in correlation with the skin surface shape at the defined landmarks. The authors validated their statistical models by a leave-one-out cross-validation procedure.

In contrast, a specific scanned partial or a whole face was used for facial reconstruction following the sparse approach. Partial faces (i.e. facial components/composites), as well as complete faces have been used for facial reconstruction. Nelson and Michael (1998) cited Evenhouse et al. (1991) who produced facial templates of average partial faces (facial features) from several 2D photographs. These features were then used as standard templates for different facial regions with the same position in every reconstructed face, changing their shapes according to the skull. Computerised 2D facial reconstruction (e.g. Face Imaging Reconstructive Morphography technique (FIRM)) was developed for the production of objective facial composites based on cephalometric measurements to prepare a database of these facial components (e.g. Identi-Kit™), then positioning them onto a digital drawing of the skull (Miyasaka et al., 1995, Vanezis, 2008, Vanezis et al., 2000). Miyasaka et al. (1995) constructed a computer imaging system for facial reconstruction consisting of an image processing unit for skull morphometry, and an image editing unit for placing the facial components, from a database, on the skull. Warping a full facial template digitally onto the skull was first described by Vanezis et al. (1989). It was further developed with designing of a special software program for facial reconstruction purposes (Shahrom et al., 1996, Vanezis et al., 2000, Vanezis, 2008). Furthermore, A database of scanned faces was established to be used as facial templates for the facial reconstruction (Vanezis, 2008). This approach was adopted, and the software and the database were used for the facial reconstructions in this thesis.

Kähler et al. (2003) performed facial reconstruction using an approach combining the features of the "inside outward" and the "outside inwards" approaches. The authors developed a physics-based animation system by placing anatomical landmarks on a 3D scanned skull, then designing an anatomy-based virtual head model. This model incorporated; (1) a triangular mesh representing the skin surface, (2) a group of virtual muscles representing the facial muscles of expression that can contract in linear and circular fashions, (3) a mass-spring system that pulls on the muscles allowing their animated deformation, and (4) skin landmarks, which correspond to the skull landmarks for better fitting. After the head model was designed, a space deformation of the skin landmarks was statistically set up, via mathematical equations, to fit the skin and muscle layout onto the skull. This animated head model provided different versions of the same face with different facial expressions, rather than just the neutral face produced by manual or other digital facial reconstruction techniques. This lively face appearance helps better recognition of the face by comparing the reconstructed face to multiple ante mortem photographs for more reliable identification. Moreover, this facial reconstruction produces a full head model, with relatively few landmarks, in a short time, and allows more variations of face shapes (e.g., slim, obese, etc.).

Table 1 and Table 2 summarise the main techniques of the "Inside Outwards" and the "Outside Inwards" approaches respectively.

Table 1: The main techniques of the "Inside Outwards" approach

The "Inside Outwards" Approach		
Manual	Russian/Gerasimov School (Muscles Only)	Gerasimov M. M.
	Manchester/Combined School (Muscles + Landmarks)	Prag J., Neave R. A. H., and Wilkinson C.
Computerised	Customised Software	(Wilkinson et al., 2006, Lee et al., 2012, Short et al., 2014)
	Free/Open Access Software	(Davy et al., 2005, ATOR, 2012)

Table 2: The main techniques of the "Outside Inwards" approach

The "Outside Inwards" Approach				
Manual	American School (Face as a bulk)			Gatliff B. P.
Computerised (Volume deformation/transformation)	**Dense** Craniofacial Model + No facial tissue depths tables	**Landmarks-Independent**		(Quatrehomme et al., 2007, Nelson and Michael, 1998, Jones, 2001, Turner et al., 2005)
		Landmarks-Based		(Attardi et al., 1999, Vandermeulen et al., 2006, Claes et al., 2010, Vandermeulen et al., 2013)
	Sparse Facial Model (Template) + Landmarks + Facial tissue depths tables	**Scanned Face**	**Partial**	(Evenhouse et al., 1991, Miyasaka et al., 1995)
			Whole	(Shahrom et al., 1996, Vanezis, 2008, Vanezis and Vanezis, 2000, Vanezis et al., 1989)
		Statistical		(Andersson and Valfridsson, 2005, Claes et al., 2006)

1.4 ACQUISITION OF 3D SKULL DATA USING COMPUTED TOMOGRAPHY

The imaging technologies used for obtaining a digital 3D image of an object are either surface or volumetric based. Laser scanners are surface based, and they acquire depth information by projecting laser rays for surface image scanning (Islam et al., 2015). Laser surface scanners were used in facial reconstruction studies for different purposes, such as; to establish a database of faces (Vanezis, 2008), and to obtain a 3D image of a completely skeletonised skull (Vanezis et al., 1989, Vanezis et al., 2000). In contrast, to examine tissues deeper than the skin surface, volumetric imaging modalities provide bony structure and soft-tissue facial information (Islam et al., 2015). Therefore, volumetric scanners, such as; ultrasound (De Greef et al., 2005, Vandermeulen et al., 2006, El-Mehallawi and Soliman, 2001), computed tomography (CT) (Phillips and Smuts, 1996, Shimofusa et al., 2009, Tilotta et al., 2009) and MRI (Sahni et al., 2008, Sipahioğlu et al., 2012) have been used to mesure the facial soft tissue depths for facial reconstruction. In addition, certain scanners, such as; a CT scanner can be employed to produce a 3D digital skull image (Attardi et al., 1999, Cavalcanti and Vannier, 1998, Cesarani et al., 2003, Cesarani et al., 2004, Jayaratne et al., 2012, Lee et al., 2012, Quatrehomme et al., 1997, Wilkinson et al., 2006). The latter medical imaging devices can provide a 3D image of the skull quickly and easily, especially if the skull specimen is not completely skeletonised (i.e. the soft tissue is still present) with no need for defleshing or damaging the skull (Lee et al., 2012, Wilkinson et al., 2006). In the present thesis, the facial soft tissue data used were obtained via ultrasound (El-Mehallawi and Soliman, 2001), while the skull and facial images were obtained using computed tomography (CT) scans of living patients.

With the introduction of CT in the early 1970s, it became possible to create radiological cross-sections of the entire body. In 1998, a new generation of spiral CT scanners was introduced which made it possible to produce numerous cross-sections of a complete body in less than one minute (Attardi et al., 1999, Poulsen and Simonsen, 2007). The overall spatial concept of CT imaging allows a better understanding of the tissue complexities than multiple 2D axial images, especially with the modern advances in image acquisition, and 3D images processing and display (Cavalcanti and Vannier, 1998).

Moreover, the introduction of Digital Imaging and Communications in Medicine (DICOM) allowed the raw slice images to be transferred from the CT scanner to a workstation with a simple and fast protocol. In addition, CT is capable of performing a large-volume examination, with thin scanning slices, and in a short time during a single-shot whole-body acquisition (Andersson and Valfridsson, 2005, Cavalcanti and Vannier, 1998, Lee et al., 2012, Tilotta et al., 2009, Vandermeulen et al., 2006).

Cavalcanti and Vannier (1998) conducted a study to compare the accuracy of 2D and 3D CT modalities. Craniometric measurements were taken from anatomical skull landmarks of cadaver heads, via 2D CT and 3D CT. they were, then, compared with direct measurements taken by an electromagnetic digitiser as the ground truth. The authors reported that the viewing was satisfactory in both image types. However, many landmarks were better identified in 3D CT than 2D CT, particularly of the points that involved the sutures, the midface and the mandible. Moreover, the 3D CT measurements were statistically different from those of the 2D CT, with lower mean differences observed within the former measurements. In 2D CT, 25% of the measurements were significantly different from the physical measurements, mostly in areas with skull trauma, in contrast to non-significant differences between the 3D CT and the direct methods in all the measurements. As well, Rocha et al. (2003) assessed the repeatability of certain craniometric anthropological linear measurements taken, by two practitioners, from 3D images that were reconstructed from 2D CT axial slices. According to the authors, the standard error percentage between the two practitioners was low, with adequate error factors in the bone and soft tissue measurements for this type of analysis.

Image visualisation, manipulation and analysis from CT scans start with the separation/extraction of a 3D image of a certain tissue (e.g., skull and face surface) from the 2D CT cut series, a process referred to as “segmentation”. There are various approaches to perform segmentation from the CT (e.g., intensity thresholding, histogram based selection of the threshold level, clustering of grey level of boundary, and Canny edge detection (Rathnayaka et al., 2011)). The most widely used method is the intensity thresholding-based technique, which is based on that different tissue types have different grayscale values. Each type of tissue can be separated into a separate volume via at a certain radio-density (i.e. thresholding) represented by CT numbers expressed in Hounsfield units (HU) scale (De Greef et al., 2005, Rathnayaka et al., 2011, Tilotta et al., 2009, Vandermeulen et al., 2006). This is helped by the rapidly growing 3D graphics features of the recent medical CT scanners, in particular the volume rendering technique that enables the user to reconstruct body parts from the CT scans in 3D according to their separate volumes (thresholds) (Cavalcanti and Vannier, 1998). To render a surface from 2D CT scans, the tissues are segmented, then a smooth surface is formed via interpolation between the slices, followed by surface illumination (Wilkinson, 2007).

Segmentation was performed manually by Attardi et al. (1999) (Attardi et al., 1999), where 2D slices were stacked up and interpolated to build a volume. Isosurfaces (i.e. surfaces with points of the same function value) were then generated from the stacked slices of the region of interest (ROI) according to their radio-densities (Attardi et al., 1999). Although, manual segmentation is simple, it is subjective and involves intra- and inter-personal variability. In addition, it is labour intensive and time consuming (Rathnayaka et al., 2011). On the other hand, 3D image reconstruction from 2D CT scans was achieved semi-automatically, where the user visually selects the threshold level of the ROI (Andersson and Valfridsson, 2005, Lee et al., 2012, Shweel et al., 2013, Vandermeulen et al., 2006, Wilkinson et al., 2006). This user interface, however, affects the accuracy and the repeatability of this method. further, the segmentation could be performed via fully automated techniques that invlove lower user intervention and thus, less inaccuracies (Rathnayaka et al., 2011). Although these segmentation software programs are complex and require comprehensive background in programming and/or mathematics, some of them are freely available (3D Slicer®, InVesalius®), thus preferred by many researchers (ATOR, 2012, Fedorov et al., 2012).

Sakuma et al. (2010) investigated the accuracy of using 3D images segmented from a mobile single helical CT for facial identification by superimposing the 3D reconstructed faces on both the skulls and the antemortem photographs of the victims. Jayaratne et al. (2012) also applied the superimposition between 3D images segmented from the CBCT and those reconstructed from stereophotography (3D photographs) to benefit from the advantages of the both methods. These advantages include low radiation dose, short scanning time, compact design, locating the bony landmarks, and viewing the natural surface colour and texture. The authors concluded that CT and 3D photographic data can be successfully fused with minimal errors. This helps accurately identify the anthropometric soft tissue landmarks and the face areas suitable for 3D image registration. In addition, the 3D reconstructed images have the advantages of image manipulation (translation, rotation, and segmentation) by computer graphics as well as interactive landmark identification (Cavalcanti and Vannier, 1998).

It should be noted, however, that even with improved visual presentation, a digital 3D image reconstructed from a scanning device is not an entirely exact copy, as the scanned copy is affected by the slice thickness, the scan plane, the spatial resolution, the filters applied, and the angle of rotation, in addition to the artefacts caused by dental filling and appliances (Claes et al., 2010, Wilkinson, 2007). One of the inherent limitations of the digital scanners in general, and the CT scanner in particular, is the limited resolution, which causes distortion of some of the skull details (e.g., apertures, fossae, and holes), and even loss of the finer details (e.g. nasal spine) (Claes et al., 2010). This is due to the computational processing as well as the manual editing required to reconstruct a 3D image from a number of viewpoints in a series of profiles (Kähler et al., 2003, Wilkinson, 2007). Furthermore, a skull replica produced from the 3D reconstructed skull image will suffer from the same limitations, in addition to other problems related to the replication procedure (Wilkinson, 2007). Therefore, referring to the original skull is always important as an adjunct to the scanned images (Kähler et al., 2003, Wilkinson, 2007).

The distinctive function of CT allows observing the cranial bones and any bone pathology with a very high image resolution, as well as confirming the anthropologist's observations on the skeletal remains, especially if it was covered by desiccated soft tissue (Attardi et al., 1999). Although other non-invasive scanners, such as; laser or ultrasonography, are more suitable to build up large databases of skull and face images, CT scans provide a better contrast to distinguish between different tissue types using the thresholding technique (De Greef et al., 2005, Tilotta et al., 2009, Vandermeulen et al., 2006). As well, as CT is a contactless scanner, it would be more suitable in many cases (e.g., burnt or decomposed remains) (Phillips and Smuts, 1996, Shimofusa et al., 2009). Further, despite that MRI does not carry a radiation hazard (Rathnayaka et al., 2012), this is not a problem when scanning cadavers for forensic identification purposes using a CT scanner. In addition, compared to MRI, CT is cheaper, quicker with better visualisation of the bone tissue, of a comparable accuracy (Rathnayaka et al., 2012), and more readily available than MRI in most hospitals and post mortem facilities.

Three dimensional image reconstruction from a CT scanner has been applied in human identification using various techniques, such as; photo superimposition (Jayaratne et al., 2012, Sakuma et al., 2010, Shahrom et al., 1996), and facial reconstruction (Andersson and Valfridsson, 2005, Lee et al., 2012, Quatrehomme et al., 1997, Wilkinson et al., 2006), in addition to obtaining facial scans to construct a large database of human faces for facial reconstruction (Claes et al., 2006, Vandermeulen et al., 2006, Turner et al., 2005, Moyers, 2007, Parks et al., 2013). The ability of CT to view soft tissue contrast as well as internal organs and structures helped archaeologists obtain better information from archaeological specimens. Cesarani et al. (2003) conducted virtual endoscopy with navigation in hollow structures filled with air using a dedicated software program. This fly-through endoscopy offered an inside view of the body that had previously been possible to attain only through surgical and invasive techniques (Cesarani et al., 2003). Cesarani et al. (2004) examined 13 wrapped well-preserved Egyptian mummies, and Friedrich et al. (2010) examined 12 Chachapoyan mummies using multidetector CT (MDCT) technology as a non-invasive investigative tool to obtain 3D reconstructions of the mummies' whole bodies with virtual removal of the bandages. Moreover, the faces of ancient mummies have been reconstructed with the aid of CT (Attardi et al., 1999, Baldock et al., 1994, BBC News, 2005, Cesarani et al., 2004, Hughes, 1996, Mathilda's Anthropology Blog, 2011, Mummy Tombs, 2011, University of Leicester, 2013, Wilkinson, 2003). CT was also used for the acquisition of 3D skull images for forensic facial reconstruction in a number of studies (Andersson and Valfridsson, 2005, Kähler et al., 2003, Kim et al., 2005, Lee et al., 2012, Quatrehomme et al., 1997, Rocha et al., 2003, Vandermeulen et al., 2006, Vanezis et al., 1989, Wilkinson et al., 2006). In addition, CT was useful to collect facial soft tissue thickness measurements for different populations (Phillips and Smuts, 1996, Shimofusa et al., 2009, Tilotta et al., 2009).

Different types of CT have been used in facial reconstruction research studies; such as, single-detector spiral/helical CT (Attardi et al., 1999, Kim et al., 2005), multislice/multidetector CT (MSCT/MDCT) (Cesarani et al., 2003, Cesarani et al., 2004, Friedrich et al., 2010) and cone-beam CT (CBCT) (Jayaratne et al., 2012, Lee et al., 2012). Spiral/helical CT, introduced in the 1980s, provides simultaneous continuous tube rotation and CT table advancement, while MDCT is equipped with multiple detector banks to shorten the scanning time and increases the resolution (Kim et al., 2005). Spiral CT can take slices with 2-mm thickness, while MDCT can take multiple slices with 0.5 mm in thickness with more information about hard tissue like bone. In addition, MDCT provides sufficient information on bone abnormalities and pathological changes via a fan-shaped beam that images the subject as sequential slices (Xi et al., 2013, Jayaratne et al., 2012). However, the inherent accuracy of spiral and MDCT can be jeopardized by metal artefacts, such as; amalgam teeth fillings, as these artefacts interfere mostly with the intensity thresholding procedures of the bone structures (Sakuma et al., 2010, Vandermeulen et al., 2006). They also involve exposure to high radiation dose, high cost, and difficult access (Jayaratne et al., 2012, Kim et al., 2012, Xi et al., 2013).

Cone-Beam CT (CBCT) is a CT model of real-size dataset that employs a cone-shaped beam of x-rays with shorter scanning time over a single low radiation dose scan (Xi et al., 2013, Jayaratne et al., 2012). It can produce slices down to 0.1 - 0.2 mm thickness, with more detailed measurement points (Sakuma et al., 2010). CBCT can reduce metal artefacts, thus can be applied in dentistry, odontology as well as in human identification. In addition, CBCT is cost-effective and has the capability of scanning the patient in an upright sitting position with a neutral and relaxed facial expression. This has the advantage of avoiding sagging down and subsequent distortion of the face by gravity that happens with other CT devices which take the subjects' images in supine position. The main drawbacks of the CBCT, however, are the low contrast resolution and the limited Field of view (FOV), which restrains the accurate 3D reconstruction process, so that a relatively large defect area, especially on the occipital region of the subject scan is usually present. Thus, in order to overcome this shortcoming, post-processing of the CBCT data is needed (Lee et al., 2012, Xi et al., 2013). In addition, CBCT has other limitations, such as; the lack of a soft-tissue window, and the precise Hounsfield Units (HU), as well as high image noise (Cha, 2013).

The main drawbacks of the CBCT, however, are the low contrast resolution and the limited Field of view (FOV), which restrains the accurate 3D reconstruction process, so that a relatively large defect area, especially on the occipital region of the subject scan is usually present. Thus, in order to overcome this shortcoming, post-processing of the CBCT data is needed (Lee et al., 2012, Xi et al., 2013). In addition, CBCT has other limitations, such as; the lack of a soft-tissue window, and the precise Hounsfield Units (HU), as well as high image noise (Cha, 2013).

Kim et al. (2012) compared the accuracy of linear measurements taken (1) directly from dry skulls by digital caliper, and from 3D images extracted by (2) MDCT and (3) CBCT. The results showed statistical differences within the repeated measurements of each of the three methods; with the least of them found in the CBCT 3D reconstruction images. However, when Shweel et al. (2013) compared the accuracy of MDCT and CBCT, the measurements taken by CBCT were slightly closer to, but not significantly different from, the real intraoperative measurements than those taken by MDCT. In addition, the 2 imaging modalities showed similar morphologic characteristics. This proves a comparable accuracy between the MDCT and the CBCT.

The accuracy of CT in measuring facial soft tissue depths has been assessed by Kim et al. (2005). Using a digital calliper perpendicular to the bone, direct physical soft tissue depths measurements were taken from punch holes consisting of the full soft tissue thickness from skin to bone at certain facial landmark. These direct measurements showed high agreements within and between users, therefore were considered as the ground truth. They were then compared with the CT measurements taken immediately after the physical measurements by the same users. The CT measurements were taken under 13 different CT protocols, which differed in the type of CT (conventional MDCT and spiral CT), slice thickness and pitch ratios within the same slice. The results showed high intraclass correlation within the physical measurements and the CT protocols. However, in a number of instances, the CT measurements were significantly different from the physical measurements, but with a small mean deviation in every instance. It was, also, observed that the image quality was better in the spiral than the conventional CT scanning. In addition, the image quality decreased with increasing the slice thickness, and the pitch ratio within the same slice thickness. Therefore, the authors recommended that CT scanning can be used to accurately measure the facial soft tissue thickness, and it is better to use slice thicknesses less than 5 mm, and a spiral/helical pitch less than 2:1 for more accurate measurements.

In spite of the advantages of CT in the field of forensic facial reconstruction, the use of imaging methods in research involving living persons is limited to clinical indications (Schuh et al., 2013). This is because these CT scanners carry an unnecessary radiation hazard when applied on a healthy living subject, which raises ethical and legal concerns (Quatrehomme et al., 1997, Shimofusa et al., 2009). Only a few studies used CT head images from live subjects either volunteering and consenting for the purpose of these studies (Andersson and Valfridsson, 2005, Lee et al., 2012, ATOR, 2012), or as a retrospective study using CT scans of live individuals provided by an investigative/authoritative institution (e.g., FBI) (Wilkinson et al., 2006).

Due to the advantages of CBCT over MDCT, it has become more popular in the clinical orthodontic field (Jayaratne et al., 2012, Kim et al., 2012, Lee et al., 2012, Xi et al., 2013). The low radiation dose associated with CBCT has led to its use in facial reconstruction studies. However, the accuracy of facial reconstructions by Lee et al. (2012) who used skulls scanned by CBCT were comparable to those of Wilkinson et al. (2006) who utilized spiral CT scanned heads. Moreover, some of the disadvantages of using other types of CT in research can be minimised by incorporating forensic facial reconstruction studies within the clinical field, where scans of patients' heads can be used for research without extra cost or radiation exposure with other diagnostic or medical indications for patients (Jayaratne et al., 2012, Kim et al., 2012, Lee et al., 2012, Shimofusa et al., 2009, Tilotta et al., 2009, Xi et al., 2013). Therefore, many facial reconstruction studies emphasise the importance of the multidisciplinary cooperation between the CT radiologists, computing specialists, anthropologists, and forensic artists in order to achieve the highest benefits from applying CT in facial reconstruction (Attardi et al., 1999, Cesarani et al., 2003, Cesarani et al., 2004, Vanezis et al., 2000).

1.5 Facial Soft Tissue Thickness Measurements

The depths of facial soft tissue constitute an integral part of forensic facial reconstruction techniques as they link the inner skull to the outer skin, hence they are thought to largely contribute the shape of the face (Vanezis, 2008).

1.5.1 Methods of measuring the facial soft tissue depths

Different methods were employed by different researchers for measuring the facial soft tissue thicknesses for facial reconstruction starting with taking these measurements directly from cadavers' faces. One of the earliest attempts of taking *direct measurements* was that by the German physiologist and anatomist Welcker (1883), where a thin small surgical knife blade was inserted into the faces of a number of cadavers at certain facial point corresponding to anthropometric cranial landmarks. The facial soft tissue thicknesses were marked as different lengths on this knife blade. Welcker used this facial technique in the facial reconstruction of Schiller, Kant and Dant. In 1895, the anatomist His modified Welcker's technique, to reconstruct Bach's face, using a thin sharp needle bearing a small piece of rubber pushed into the flesh at right angles to the bone at various locations until the needle struck bone. This piece of rubber was displaced upwards from the original point, and the soft tissue thickness at that particular site was measured as the distance from the point of contact to the point of displaced rubber (Brown et al., 2004, Gupta et al., 2015). More recently, Kim et al. (2005) took physical measurements by making a punch hole consisting of the full soft tissue thickness from skin to bone at certain facial landmark, leaving it to relax for one week, then measuring the lengths of the holes using a digital calliper perpendicular to the bone.

Furthermore, the facial soft tissue depths were measured using different imaging and scanning devices, such as; *craniographs/cephalographs* (i.e. cranial x-rays) (George, 1987, Pithon et al., 2014). Details from the cephalographs can help the anthropologist estimate the skull's sex, age, and race (George, 1987, Pithon et al., 2014). Also, as the craniographs are almost life-size, they require less calibration, which can be more precise with known distances from the camera. Furthermore, the visible critical sella point (S) in craniographs allows correct skull orientation and cephalometric analysis, which is essential to determine the skull type (George, 1987).

Facial soft tissue depths were also measured using *computed tomography (CT)* (Phillips and Smuts, 1996, Shimofusa et al., 2009, Tilotta et al., 2009), and *Magnetic Reasonance Imaging (MRI)* (Sahni et al., 2008, Sipahioğlu et al., 2012). CT was used by many researchers for measuring the facial soft tissue depths, not only for facial reconstruction purposes, but also for other applications (Cha, 2013). Kim et al. (2005) concluded that CT scanning can be used to measure the facial soft tissue thickness with acceptable accurately compared to physical measurements taken from cadaver heads. Facial soft tissue depths were collected using CT from cadavers (Shimofusa et al., 2009), as well as from living subjects (Phillips and Smuts, 1996, Tilotta et al., 2009, De Greef et al., 2005, El-Mehallawi and Soliman, 2001, Parks et al., 2014). Whereas CT gives better visualisation of the bone tissue, Magnetic Reasonance Imaging (MRI) allows a clear demarkation of the cortical bone and inherent soft tissue contrast, thus allows easy measuring of the facial soft tissue thickness with small observer errors (Sahni et al., 2008, Sipahioğlu et al., 2012). Likewise, *ultrasound* is useful in soft tissue visualisation. However, it is cheaper, and more widely available than both CT and MRI and unlike the CT, ultrasound presents no radiation hazard, thus it was usd for measuring facial soft tissue depths in many studies (De Greef et al., 2005, El-Mehallawi and Soliman, 2001).

A number of researchers combined craniographs with ultrasound for measuring the facial soft tissue depths (Aulsebrook et al., 1996, Smith and Throckmorton, 2006). Aulsebrook et al. (1996) tested the accuracy of radiography and ultrasound individually and combined. The authors stated that that dual system was found to be accurate and reliable. Smith and Throckmorton (2006) compared the facial soft tissue depths taken by lateral craniographs to those taken by ultrasound at 3 facial points (A: on the upper lip, B: on the chin indentation, and C: on the nose). The results revealed differences in the meseaurements taken by both modalities, being most in the chin, and least in the nose. This was explained by the different subject postures between the two modalities. The authors were aware of other limitations in their study that would account for the differences found between the 2 scanning methods, such as the absence of markers at landmark points on the radiographs, the lapse of time between the collection of the radiographs and the ultrasound scans, as well as the small sample size.

De Greef et al. (2006) used ultrasound to measure the facial soft tissue depths of Caucaian population, and then compared them with three former datasets for the same population. These datasets were collected by Rhine and Moore in 1984 (directly from cadavers), by Helmer in 1984 (using ultrasound), and by Manhein et al. in 2000 (using ultrasound). The authors conluded that the majority of their measurements were significantly different from other studies, with the least difference found between this study and that of Manhein. The authors attributed these differnces between their measurements and those of the other studies to a number of reasons realted to the other studies, including the smaller sample sizes, the supine posture of subjects compared to the upright position in their study, and the differences in defining the studied subcategories, particularly the BMI. To some degree, these differences were also thought to be due to the postmortem changes that could have influenced the facial soft tissue depths in the previous studies. De Greef et al. (2005) compared the facial soft tissue thickness measurements at 52 landmarks taken by ultrasound to those measured by CT. The results showed an insignificant difference between the ultrasound (US) and the CT measurements except at 6 out of the 52 landmarks (i.e. < 12%) only, which were located in the masseter region. This was thought to be due to the influence of gravity on the soft tissue thicknesses between the upright position during the ultrasound acquisition and the supine position during CT scanning (Phillips and Smuts, 1996, Shimofusa et al., 2009). Also, the results of the semi-automated ultrasound facial soft tissue depth registration method of De Greef et al. (2005) showed good repeatability of the ultrasound measurements.

On the other hand, Herrera et al. (2016) compared the recognition rates of faces reconstructed using four different facial soft tissue thicknesses datasets of Brazilian population from different methods (2 from cadavers, 1 from CT and 1 from MRI). The results showed that data from cadavers resulted in a higher recognition and more resemblance to the targets than those conducted with data from medical imaging (including CT and MRI). The authors attributed that to the inherent inaccuracies in medical imaging due to factors, such as; the measurement direction, the head positioning, the resolution of images, the segmentation algorithms of hard tissues and positioning of markers. However, when ultrasound was included in the comparison, Stephan and Simpson (2008a) observed that the average facial soft tissue depths obtained by needle puncture in the soft tissues of cadavers were similar to those measured by ultrasound, but thinner than those measured by CT data, especially at the cheeks.

It is very important that practitioners are cautious when applying facial soft tissue depths, collected using a certain scanning device, on skulls scanned by another scanning device with a different posture. For example, the subject is scanned in a supine position via radiography, MSCT and MRI, in contrast to the upright position in ultrasonography and CBCT (Aulsebrook et al., 1996, De Greef et al., 2005, Smith and Throckmorton, 2006). These body posture differences affect the facial soft tissue depths due to the influence of gravity on the face (e.g., dropping of the cheeks, and movements of the eyelids following movements of the cheeks), with minimal effect on the nose. These effects are especially important in obese and older subjects (Tilotta et al., 2009). Measurements taken in the upright posture are more realistic when reconstructing an unknown face for the identification purposes. Although this posture differences may not be reflected to a large degree on the final appearance and, hence, the recognisability of the reconstructed face (Smith and Throckmorton, 2006), many researchers preferred the use of ultrasound to collect facial soft tissue depths (De Greef et al., 2005, De Greef et al., 2006). This was shown in the review by Hayes (2014) of 15 facial reconstruction articles published between 2000-2013, which revealed that most of them used soft tissue depths of Helmer (1984) collected via ultrasound followed by the data collected directly from cadavers by Rhine and Moore (1984). Vanezis (2008) used combined data tables from the latter two studies for the facial reconstruction of adult Caucasian males and females.

It is, therefore, shown that facial soft tissue thicknesses are best measured using ultrasound, especially that it is better to measure the facial soft tissue depths in living subjects. However, , Smith and Throckmorton (2006) mentioned a number of technical limitations while using ultrasound, as well as some practical suggestions to overcome them for a more accurate application of ultrasound. For example, the tendency toward mediolateral (right-left) deviation in the collected scans, which usually requires correction can be overcome by using an automated, rather than a freehand, scanning process. Also, repeated scanning is usually needed to avoid image loss. Using a headrest allows a better control over head positioning. Moreover, better images could be obtained via more recent ultrasound machines, with improved video imaging.

1.5.2 Factors affecting the facial soft tissue depths

When performing the facial reconstruction, a number of factors that affect the facial soft tissue depths should be kept in mind as they may possess limitations to the success (i.e. recognition) of the reconstructed faces. For example, *population*-specific differences in the facial soft tissue depths have been long acknowledged by researchers, thus data tables were published for different anthropological races and, then, for their subcategories (i.e. mixed populations). Of the first attempts was that of Rhine and Moore for Caucasoids in 1984; Suzuki for Mongoloids in 1948, and Rhine and Campbell for Negroids in 1980 (Aulsebrook et al., 1996). Mixed or subpopulations related differences were noted to have unique facial features, and thus different facial depths from their counterparts of the original races. This called for careful anthropological examination and metric analysis of the skull before facial reconstruction and careful selection of the suitable population-specific tables (Phillips and Smuts, 1996). Moreover, researchers recommended that the accuracy of craniofacial reconstructions requires more data for refining the subject-specific attributes to help appreciate the relation between the physical properties of subject and their facial soft tissue depths (De Greef et al., 2006).

Therefore, various sets of facial soft tissue depths measurements were published for different populations as well as mixed or subpopulations. To name a few, American Whites and Blacks and South Africans of mixed racial origin (Phillips and Smuts, 1996), Caucasian population (De Greef et al., 2006), European population (Tilotta et al., 2009), Japanese population (Shimofusa et al., 2009), Turkish population (Bulut et al., 2014, Sipahioğlu et al., 2012), Korean population (Hwang et al., 2012), South Korean population (Cha, 2013), Chinese-Americans population (Chan, 2007), Taiwan population (Chung et al., 2015), Brazilian population (de Almeida et al., 2013, Tedeschi-Oliveira et al., 2009), Australian population (Domaracki and Stephan, 2006), Egyptian population

(El-Mehallawi and Soliman, 2001), French population (Guyomarc'h et al., 2013), Indian population (Sahni et al., 2008), Portuguese population (Codinha, 2009), Columbian population (Ruiz, 2013), and Northern Sudanese population (Sforza et al., 2013), etc.

The individual's *age* is one of the other factors that should be kept in mind as it affects the facial soft tissue depths. Aging has shown an influence on the facial soft tissue depths, where the whole tissue thickness is negatively correlated with advancing age due to skin wrinkling (Sahni et al., 2008). Thus, it should be kept in mind that children and young adults have different facial soft tissue depths compared to adults (Briers et al., 2015, Hodson et al., 1985, Wilkinson, 2002). In addition, there is a documented *sexual dimorphism* in the facial soft tissue depths, where male soft tissue depths are thicker than those of females (Cavanagh and Steyn, 2011, Cha, 2013, De Greef et al., 2006, De Greef et al., 2009, Dong et al., 2012, El-Mehallawi and Soliman, 2001, Utsuno et al., 2014, Wilkinson, 2002). Even in children aged 7 – 18 years old, sex differences in facial soft tissue depths increases as the age increases towards 18 years old (Utsuno et al., 2010).

On the other hand, Stephan et al. (2005a) argued that variations in facial soft tissue depths within each sex are far larger than those between sexes. Therefore, a number of researchers looked into the sexual dimorphism in the individual facial parts (De Greef et al., 2009, Sahni et al., 2008, Pithon et al., 2014, Utsuno et al., 2014). Significant statistical sexual differences were observed in the facial soft tissue at certain midline and bilateral landmarks. For example, in males, some facial areas appear thicker, such as; around the mouth (De Greef et al., 2009), particulalry chin lip fold and beneath the chin (Sahni et al., 2008). In contrast, it was observed that, in females, the cheeks (De Greef et al., 2009), particulalry the supra M2, which is located above the second upper molar tooth, appear thicker (Sahni et al., 2008). The brow, and the glabella tissue depths differences were negligible or slightly bigger in women (De Greef et al., 2009), while the supraglabella was thicker in men (Sahni et al., 2008). Moreover, the individual differences in the skeletal and facial types are thought to influence the facial soft tissue thickness. However, it was observed that the differences in the facial soft tissue thicknesses between skeletal classes were related to sexual differences in children (Pithon et al., 2014) as well as in adults (Utsuno et al., 2014), being more apparaent in females than in males. However, as the differences were mostly less than one mm, the effect of which on the facial reconstructions was not believed to be marked (Utsuno et al., 2014). The skeletal class is detemined by the ANB angle: (A) the deepest point on the premaxilla outline, (N), nasion, lateral view of the most anterior point of the frontonasal suture, and (B) deepest point on the anterior wall of the mandibular symphysis. The ANB angle indicates the position of the maxilla in relation to the mandible and categorises 3 skeletal classes. In class I, the ANB angle = 2 - 4º, in class II, the ANB angle is >4º; and in class III, the ANB angle is <2º. Accordingly, the facial profiles can then be classified into straight, convex, and concave, respectively (Pithon et al., 2014, Utsuno et al., 2014). In class III (the concave face), there is either an overgrowth of the mandible or a decreased growth in the maxilla region. The opposite is observed in class II (the convex face). Also, the soft tissue thickness values of skeletal class I range between the values of class II and III from the subnasale to the labiomentale (Utsuno et al., 2014).

This shows that the sexual dimorphism in facial soft tissue depths was apparent in individual facial parts rather than as an overall average thickness (De Greef et al., 2009, Pithon et al., 2014, Utsuno et al., 2014). However, the areas that showed sexual differences (e.g., the cheek, the mouth) contain a higher amount of subcutaneaus fat (De Greef et al., 2009). Accordingly, it is expected that the facial soft tissue depths change with the body weight. However, small changes in weight are not usually reflected on the face, rather are distributed throughout the body, in which case does not affect the likeness of the face to a large degree (De Greef et al., 2009, Utsuno et al., 2014). Starbuck and Ward (2007) measured the accuracy of three variants of the reconstructed faces (emaciated, normal and obese face) quantitatively/objectively, using the anthropometric craniofacial variability index, and qualitatively by subjectively assessing the resemblance of the three variants of the reconstructed faces to their photographs. The results of quantitative assessment were not consistent with the subjective assessment.

Similarly, the subject’s height does not affect the facial soft tissue thickness because the likeness depends on proportions more than on finite measurements (Aulsebrook et al., 1996). It can, therefore, be concluded that soft tissue depth differences among adults that are not attributed to sex and with no systematic unidirectional change with age might be attributed to the *height-to-weight ratio* and *Body Mass Index (BMI)* (De Greef et al., 2006). It is agreed that the body build (or the BMI) in particular has the most substantial effect on the facial soft tissue depths compared to the other factors (Aulsebrook et al., 1996, De Greef et al., 2009, Smith and Throckmorton, 2006, Vanezis, 2008).

However, it is important to be careful when excluding other influential factors (e.g., sex, or age), as their influence might not be equally applied to all facial parts (De Greef et al., 2009). Further, the BMI can sometimes explain the variations attributed to the other factors (e.g., sexual dimorphism) (Smith and Throckmorton, 2006). Therefore, care should be directed to categorising the individual properties of the subjects, especially when studying the impact of these influential factors on the facial soft tissue depth measurements (Smith and Throckmorton, 2006), or when selecting a certain facial soft tissue data table for an unknown skull with estimated anthropological characteristics (Vanezis, 2008). Particularly, the faces should be classified according to the body weight (or BMI) as; thin, very thin, well nourished, and very well nourished. Accordingly, the tissue thicknesses should be averaged, and then classified into average male (very thin and well nourished), and average female (thin and well nourished), as well as the maximum and minimum variations for both sexes in order to be able to produce more accurate facial reconstructions (Vanezis, 2008).

We should, also, bear in mind that factors related to the practitioner, the technique, and the condition of measuring the facial tissue depths would be expected to affect the measurements, thus the facial reconstruction. For instance, Shimofusa et al. (2009) used MDCT to measure facial depths from cadavers within 0 – 3 days after death without embalming. These measurements were compred to other studies measuring facial depths from cadavers and from living subjects of the same population. Although the cadaveric measurements taken by Shimofusa et al. (2009) were thicker than other previously published cadaveric data, they were thinner than data from live persons. This was attributed to the effects of postmortem changes on the facial soft tissue depths, such as tissues dehydration, muscle rigidity, and drooping due to loss of muscle activity and zero gravity after death (Aulsebrook et al., 1996, Phillips and Smuts, 1996, Tilotta et al., 2009). The dry cadaveric skin makes it hard to puncture and measure the soft tissue thickness directly perpendicular to the skin surface, due to the difficulty in finding the exact underlying bone sites using manual palpation (Kim et al., 2005). Even in fresh cadvers, postmortem dehydration is believed to be high particularly in the initial stages (Phillips and Smuts, 1996). Therefore, measuring the facial soft tissue thicknesses from the living subjects is thought to be more accurate than from cadavers (Phillips and Smuts, 1996, Tilotta et al., 2009).

Additionally, there are other less common factors that could affect the facial soft tissue depths. For instance, Smith and Throckmorton (2006) referred to that long time lapse between the repeated collections of facial soft tissue depths for the same subjects might be associated with a change in the soft tissue measurements. This may be, for example, due to the hormonal changes in women that could alter their tissue depths, or other seasonal differences affecting both sexes. Moreover, fractures, swellings, malformations, distortions and asymmetries, and even the administration of a local anaesthetic can distort the tissue through swelling and flaccidity (Aulsebrook et al., 1996). Even more, facial depths thickness can be affected by water retaining in the body (e.g. diet, menstruation, pregnancy, alcohol, etc.) (Mollov et al., 2012).

CHAPTER 2: AIMS & OBJECTIVES

This study aimed to demonstrate that 3D facial reconstructions using scanned facial templates digitally fitted onto 3D skulls showed a sufficient resemblance to the real faces of the persons in question when they were alive to allow the recognition of these reconstructed faces should they advertised. Moreover, the subjective and objective methods previously described in research studies to assess the accuracy of facial reconstructions were evaluated whilst developing the current methods and introducing new methods. Before starting the main study, a pilot study was conducted to test the main components of the 3D forensic facial reconstruction method proposed in this thesis (i.e. the facial templates, the facial soft tissue depths, and the accuracy assessment tests).

To achieve these aims, the following objectives were sought in the pilot study:

1- Investigating the validity of the proposed method of facial reconstruction in producing faces with a sufficient resemblance to the target.
2- Investigating whether certain facial templates were better than others in obtaining identification and resemblance to the target, together with comparing single and average faces as suitable facial templates for the proposed method of facial reconstruction.
3- Investigating the influence of using different sets of facial soft tissue thickness measurements on the accuracy of the facial reconstructions.
4- Investigating the applicability of the proposed method and the present facial reconstruction software with testing the possibility of transferring the process from one user to another.
5- Investigating the best way to design the subjective assessment tests of facial reconstruction (i.e. the face pool and face resemblance tests) and the best way to present the tested faces to assessors in order to suit a forensic facial reconstruction research study.

Also, the following objectives were sought in the main study:

1- Investigating the validity of the proposed method of facial reconstruction in producing faces with a sufficient resemblance to the target.
2- Assessing the reliability of different subjective assessment methods previously described in literature in determining the accuracy of facial reconstructions, taking into account the individual variations among the assessors (e.g., age, sex and professional experience in a related field).
3- Developing and validating Craniofacial Anthropometry as an objective assessment method for facial reconstruction.
4- Setting a guidance approach in setting a research study involving subjective and/or objective assessment of forensic facial reconstructions.

CHAPTER 3: A PILOT STUDY

The method of 3D facial reconstruction proposed in this thesis entails reconstructing the face of a certain skull using a template of a scanned face of an anthropologically similar skull (i.e. the "Outside Inwards" approach), and employing the present 3D computer graphic Facial Reconstruction (FR) software. This pilot study was conducted to test the main components of this method, including; the *scanned facial templates (single and averaged),* the *facial soft tissue thickness measurements,* and the *subjective and objective assessment tests*. In addition, the *applicability* of the proposed method and the facial reconstruction software were also tested. This pilot study lasted from April - December 2014.

3.1 Materials and Methods

The materials used in this pilot study, as well as the methods used for assessment of the resulting reconstructed faces, are described in this section.

3.1.1 Acquisition of the Pilot Study Materials

Case studies in this pilot study were obtained from two sources; a laser scanner and a computed tomography (CT) scanner. The former scans were obtained from a database collected by Dr. Maria Vanezis, the co-investigator of this thesis, as a part of her PhD thesis (Vanezis, 2008). This database included laser scanned skulls and ante mortem photographs of forensic cases of European Caucasian subjects, as well as laser scanned faces of European Caucasian healthy volunteers. Dr. Vanezis had obtained the required consents for using these materials in this research. Whereas the omputed tomography scanned data were provided from two Diagnostic Radiology Centres in Zagazig City, Sharkia Governorate, Egypt. The relevant ethical (e.g., taking consents) and governance issues were dealt with by the collaborating colleagues in the two centres according to the relevant Institutional Review Board (IRB). Patients admitted to these centres for head scanning indicated for medical and clinical purposes were approached and consented.

3.1.1.1 Laser Scanned Skulls and Faces

The 3D images of the laser scanned skulls and faces were obtained via the laser scanner (Facia Optical Surface Scanner™), which was developed by the Medical Physics Department of University College London (Vanezis, 2008). This pilot study included 4 out of the 5 cases studied by Vanezis (2008). For these cases, the ante mortem photographs were available as well as information about the skull's age, sex and race obtained from the forensic anthropological and pathological reports.

The studied cases were described as follows: Skull (I): a young Caucasian male with an estimated age of 18 – 30 years old, Skull (II): a young Caucasian male with an estimated age of 20 – 25 years old, Skull (III): a young Caucasian male with an estimated age of 24 – 32 years old, and Skull (IV): a Caucasian female with an estimated age of 30 – 40 years old. Appendix (1) shows the skulls and antemortem photographs of these cases. The laser scanned facial templates in the database were assessed prior to their use in this pilot study. As the studied skulls were adult Caucasian, non-Caucasian and children faces were initially excluded. Also, faces with artefacts due to scanning (e.g., due to very bushy dark eyebrows or beard) were excluded. As a result, a total of 86 scanned Caucasian faces (40 males and 46 females) were considered suitable. Moreover, the database included a number of "average faces" that were generated by mathematical merging of a number of faces of similar age, race and sex for each group. These faces were used for the facial reconstruction in this pilot study together with the single faces.

For each of the 3 male skull cases (I, II and III), 13 scanned male facial templates were selected from the database and each was separately used for the facial reconstruction of each skull. Of these 13 templates; 7 single templates aged < 30 years old, labelled (20Y-01, 02, 03, 04, 05, 06, 07), 3 single templates aged 30-39 years old, labelled (30Y-01, 02, 03), and 3 average faces representing 3 age groups; < 30 years old, 30-39 years old, and 40-49 years old were labelled (20Y-AV), (30Y-AV), and (40Y-AV) respectively. For the female case (IV), 2 average female facial templates representing 2 age groups; < 30 years old, and 30-39 years old were labelled (20Y-FAV) and (30Y-FAV) respectively.

Appendix (2) shows the single and average facial templates (n = 15) used for the facial reconstructions of the male and female cases.

3.1.1.2 Computed Tomography (CT) Scanned Skull and Faces

The group of Computed Tomography (CT) scanned skulls used in this study were obtained using GE (General Electric) Brightspeed model, 8 slice detectors MSCT (multislice CT) in the first centre, and using Siemens SOMATOM Emotion model, 16 slice detectors MSCT (multislice CT) in the second centre. In both centres, routine adult head scanning involved axial & coronal planes of the patient in supine position with the head in the head-holder. The indications for routine head scanning were; minor head injuries and cranial trauma, orbital lesion, facial bone injuries, and paranasal sinuses diseases, e.g. sinusitis, polyposis, tumour, etc. Scans were done through 5 mm slice thickness & 5 mm gap. Data were then transferred from the scanner to encrypted compact discs together with a viewer software program that enables the user to view the medical images on a personal computer (PC).

The retrieved CT images were stored in Digital Imaging and Communications in Medicine (DICOM) format as two-dimensional cut series. As 3D facial images (i.e. 3D polygon meshes) were required for the present study, initial processing of the retrieved images was performed before commencing the facial reconstruction process. A 3D mesh is formed of a collection of triangles that define the shape of a polyhedral object in 3D computer graphics. Each triangle is formed of vertices, edges and faces. The aim of this step was to separate the skulls from the head scans into separate volumes (i.e. 3D images), a process referred to as "*segmentation*". The segmentation process was performed via a "thresholding" technique, which is based on the difference of values among various tissues. In other words, each tissue of interest belong to a well-differentiated section determined in HU (Hounsfield units) in the CT grayscale. These HU, or CT numbers, represent the radio-densities of different structures scanned by a CT, according to which each type of tissue can be separated (De Greef et al., 2005, Vandermeulen et al., 2006, Tilotta et al., 2009).

For this segmentation process, an open source software package, (InVesalius®) version 3.0.0 Beta 5, was initially used to segment the 3D skull images. This software generates virtual three-dimensional models correspondent to the anatomical parts of the human body. The software can then generate 3D .stl (stereolithography) files (Herrera et al., 2016). To be compatible with the facial reconstruction software, the 3D images extracted from the CT scans data in (.stl) format were converted to another 3D file format (.obj) via a 3D mesh processing software system (Meshlab®) (ATOR, 2012).

Unfortunately, the segmented images were incompatible with and could not be imported into the present facial reconstruction (FR) software, and the reasons for that were not understood. Therefore, the FR software designer, Dr. Tim Niblett, Scotland, was consulted to diagnose and attempt to solve the problem.

The diagnosed problems were:

A) The triangle normals (i.e. the direction the mesh triangles point) were positioning in different directions so they don't make up a continuous surface.
B) The 3D meshes segmented from the CT scans by Invesalius software were formed of too many triangles, which created too large files (about 20 times the size of the 3D meshes obtained by a laser scanner). This led to inability to load the images segmented from the CT scans into the software or to a very poor and slow performance of the machine if could be loaded. This problem was not found while importing images taken from the laser scanner due to suitable size.
C) The segmented images were associated with non-human tissue noise (e.g., parts of the scanner table) which had the same grayscale as the segmented tissue. This associated noise increased the size of the files further, which increased the difficulty in loading these files into the facial reconstruction (FR) software.

The suggested solutions were to reduce the size of the segmented files and to correct the triangles' directions to create 3D meshes with continuous surfaces. Comprehensive investigation has been conducted; including visiting colleague from University College London who work in similar medical fields (e.g., 3D models designing and printing for medical and educational purposes and 3D imaging for facial surgery[1]), visiting colleagues from the UK well known iGene[2] digital autopsy facilities in Sheffield and Sandwell, UK who work in 3D medical images analysis from CT scanners for forensic pathology purposes, as well as contacting colleagues from the widely referenced Virtopsy project team, Institute of Forensic Medicine, University of Zurich, Switzerland[3].

This investigation has led to the identification of two types of open source software packages:

A) *Medical imaging segmenting software (3D Slicer®)*:

This software was designed for image analysis and scientific visualization. It is used to segment and reconstruct 3D skull image based on a sequence of 2D DICOM files acquired with CT equipment. This software generates virtual three-dimensional models correspondent to the anatomical parts of the human body. After reconstructing three-dimensionally Digital Imaging and Communications in Medicine (DICOM) images, the software allows the generation of .stl (stereolithography) files. Through this software, the bone and soft tissue of the head can be reconstructed. Furthermore, it was possible to crop the segmented images to remove the associated noise before being exported as 3D files. Moreover, the resulting images were of continuous surfaces which solved the first part of the initial problem, however, the sizes of the images were still too large.

B) *3D mesh processing software system (Meshlab®)*:

This software is widely used the technical fields of 3D development and data handling (ATOR, 2012). Using this software, a simplification process, known as mesh decimation, was performed to reduce the number of triangles forming the skull and faces meshes, hence to reduce the file size to be loaded into the FR software. Furthermore, the meshes segmented via the segmenting software (3D Slicer) in .stl file format were converted

[1] LIBRARY OF 3D ANATOMIES [ONLINE]. AVAILABLE AT: HTTP://WWW.UCL.AC.UK/CARDIAC-ENGINEERING/RESEARCH/LIBRARY-OF-3D-ANATOMIES [ACCESSED FEBRUARY 2016].
[2] Digital Autopsy by iGene [Online]. Available at: http://digitalautopsy.co.uk/ [Accessed July 2015].
[3] Virtopsy [Online]. Available at: http://www.virtopsy.com/index.php [Accessed July 2015].

to .obj file format via Meshlab software, which is the file format required for the FR software.

The previously described work summarises the main technical obstacles that were faced by the researcher during the course of the study and the attempts of resolving these issues. By March 2015, the researcher was able to solve the problem and to segment the target skull images from the CT scans images and import them into the facial reconstruction (FR) software in order to perform the FR process.

3.1.2 The Facial Reconstruction Software

The Facial Reconstruction (FR) software used in the present thesis was originally designed by Dr. Maria Vanezis and Dr Tim Niblett from the Turing Institute, Glasgow University for the purpose of facial reconstruction research (Vanezis, 2008). The software was later upgraded in 2009, and that updated version was used in this thesis. The design of this software adopted the concept of digital reconstruction of a face from a skull, which involves certain objects composing a session within the software. These objects include; triangle meshes for both the skull, and the face template together with sets of skull and face landmarks. These landmarks refer to certain anatomical locations on the face and the skull and can be moved interactively. Each landmark has a depth and an orientation, shown graphically by a small peg. The length of the peg corresponds to the facial soft tissue thickness at a given landmark, and the end of the peg away from the skull is where the point corresponding to the landmark on the face mesh should go. These objects can be viewed from three different vantage points at the same time (by default: left profile, anterior-posterior and right profile) to assist in the placement of landmarks, which are viewed in 3D to view and alter their direction. Furthermore, the alpha-blending (mixed view) allows the operator to see where the skull and skull landmarks are in relation to the reconstructed face. In addition, hiding the face or the skull images is possible while working on the other.

The FR software provides facilities to view the digitised skulls and facial templates as 3D scans. It is possible to import them from third party files in .lsm (Linux Software Map), .hips (Bitmap Graphics), and .obj (object) formats. In addition, images seen in the

3D window can be exported as a TIFF (Tag Image File Format) 2D image. The software gives the user the possibility to interactively manipulate the images as required. This could include; real time *Rotation* (to move the image in the main display window to rotate the view), *zooming -in and -out*, *translation* (to centre objects on the screen), *scaling* (to scale the objects in the window), and *identifying, adding*, *moving* or *removing* the landmarks.

3.1.3 The Proposed Method of Facial Reconstruction

The process of facial reconstruction is performed following the "Outside Inwards", Sparse Approach (Section 1.3.2.2). Blending the skull of one person and the face of another person aims at transforming or modelling the face to take the contour of the skull surface to take a similar shape to that of the original face for the purpose of human identification.

For each studied case in this thesis, the 3D skull image was imported into the software described in Section 3.1.2. The skull was then positioned in Frankfort horizontal plane, which can be reached when a horizontal line passes through the inferior border of the orbit and the anterior margin of the external auditory meatus (Figure 2). The cranium is in the anatomical position when the base line lies in the horizontal plane and right and left sides are level (Taylor, 2001, Wilkinson, 2007).

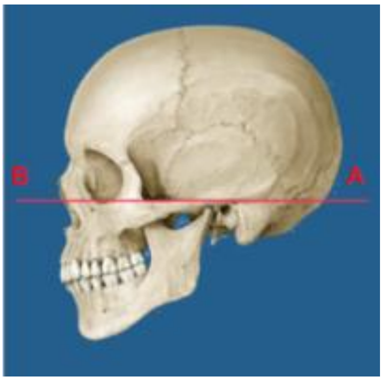


Figure 2: A skull positioned in the anatomical Frankfort horizontal plane

The skull landmarks, and the facial soft tissue depths on them, were adopted from the Rhine and Moore (1982) and Helmer (1984) combined landmarks set used by Vanezis (2008) (Appendix 12, Table 16). These landmarks were arranged in a set in which each landmark has a unique number and a name, which describes its anatomical location. The names and depths of these landmarks were saved in an .xml (Extensible Markup Language) file format, which was imported into the FR software. Different facial soft tissue thickness depths, for different populations, can be imported into the software as appropriate. Using a mouse cursor, the landmarks pegs were placed on their anatomical position on the skull. To ensure correct placing of the landmarks, the skull and the face images were moved, and rotated as required for better orientations, and were relocated and redirected to correct any error. To reconstruct the face of the skull, a facial template matching the skull's sex, and age group was selected from the database of the CT scanned faces. The selected facial template was then imported into the Facial Reconstruction (FR) software and positioned in Frankfort Horizontal position. The face landmarks, arranged in a similar set to that of the skull with corresponding numbers and names, were then placed on their corresponding anatomical positions on the face (Appendix 3). The facial image was fitted onto the skull image. The facial reconstruction is then completed by point-based fitting "warping" of the face mesh onto the skull mesh. The term "warp" is used generically to include linear transformations, and it is always the face that is warped. The software automatically "warps" the face onto the skull at the predefined anatomical landmarks on both meshes guided by the "pegs" that join the corresponding landmarks. This process of "warping" produces a new graphical object, which can be displayed in a variety of ways on a computer screen. This one-to-one mapping is used to calculate the mathematical transformation, which will produce the reconstructed face. A user manual for the software was prepared by the main researcher as a part of this thesis (Appendix 20).

3.1.4 Assessment of the Accuracy of the Facial Reconstructions

To assess the accuracy of the reconstructed faces in this pilot study, two forms of assessment tests were used; subjective (i.e. depending on the subjects' judgement) and objective (i.e. depending on computer software). The former included subjective face pool and face resemblance test types. For the subjective and objective assessment of the 3D reconstructed facial, images were exported from the FR software as 2D (.tiff), and 3D (.obj) file formats respectively. These images were then modified by removing the noise around the face, in addition to blackening of the top of the head, a process referred to as "burning". This was done in order to limit the possibility that the observers might assess the reconstructions based on these additional unreliable features (Stephan and Henneberg, 2006, Wilkinson et al., 2006). Image modification was performed using a free graphics manipulation software package GIMP (GNU Image Manipulation Program) Version 2.8.6 (GIMP®).

1- Subjective Assessment by Face Pool Tests:

The usual format of police line-up, used for confirming or excluding the identification of suspects by eyewitnesses, has been adopted in forensic facial reconstruction studies as a way of subjective assessment of the reconstructed faces. It is referred to as a *face pool test* and it aims at calculating the percentage or rate of correct identification of a target from a number of similar individuals.

In this pilot study, a comparison was conducted between two forms of the face pool tests. The first form, referred to in this study as form (A), involved a target's real face, which can be an antemortem photograph (Wilkinson et al., 2006, Vanezis, 2008), or a scanned face (Claes et al., 2006). The target's real face was compared to a pool of computer generated faces, including the target's reconstructed face and other computer generated or scanned faces) of the same complexion, appearance, and expression, referred to as "foils" or "fillers". The second form of the face pool tests, referred to in this study as form (B), involved comparing a target's facial reconstruction compared to a pool of photographs; including that of the target and other "foils" or "fillers" subjects similar to the target's age, sex, race, built, and general face morphology, but with eyes open and keeping head hair (Stephan and Henneberg, 2006, Moyers, 2007).

The foil photographs were selected from the freely available database of Glasgow Face Matching Test (GFMT) and Glasgow Unfamiliar Face Database (GUFD) databases (York Face Var Lab), as well as the Face Recognition Technology (FERET) database of the

National Institute of Standards and Technology (National Institute of Standards and Technology NIST).

2- Subjective Assessment by Face Resemblance Tests:

The resemblance ranking "*face resemblance test*" comprised a direct visual comparison using a rating/ranking scale the degree of similarity/resemblance between the reconstructed and the target faces and whether this would be sufficient to allow its correct identification had it been advertised in real life (Parks et al., 2013). In this pilot study, a certain form of the face resemblance tests, referred to in this study as form (A), was adopted. It involved subjectively assessing each facial reconstruction separately by assigning a score, from a rising numerical scale, to the reconstructed face according to its similarity to the target's real face (Moyers, 2007, Stephan and Henneberg, 2006). A numerical (e.g. 1 – 10) scale was used, where the (1) indicates no resemblance and (10) value indicate a high resemblance.

3- Objective Assessment:

In this pilot study, objective assessment was conducted via surface distance comparison between the real and the reconstructed faces of the same cases via Root Mean Square (RMS) surface distance difference.

Figure 3 summarises the subjective and objective tests involving the laser and the CT scanned skulls in a number of experiments involving the reconstructed faces of the studied cases to test the main components of the 3D facial reconstruction method proposed. The results of these experiments were analysed via Microsoft office Excel 2010 and Minitab® 17 statistical software package.

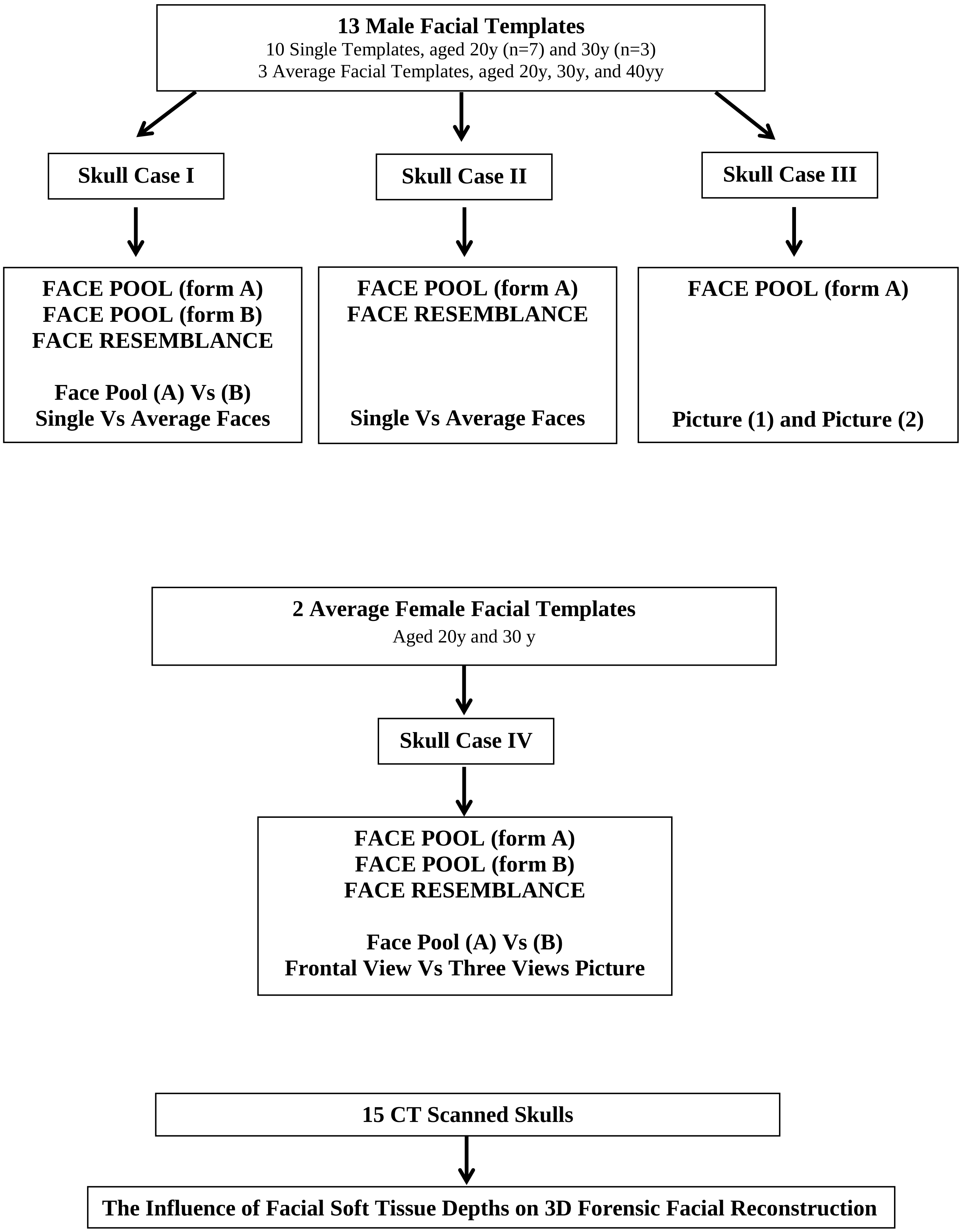


Figure 3: A flowchart showing the subjective and objective tests involving the laser and the CT scanned skulls

3.2 EXPERIMENT ONE: COMPARING SINGLE AND AVERAGE HUMAN FACES AS FACIAL TEMPLATES FOR 3D FORENSIC FACIAL RECONSTRUCTION

This experiment consisted of 2 parts aiming to compare between the influence of using single facial templates and that of the average facial templates on the resulting facial reconstructions. Comparison was conducted subjectively in Part One and objectively in Part Two.

3.2.1 Part One

- **Methods**

In this experiment, faces of two male Caucasian skulls, (I) and (II) (Section 3.1.1.1 and Appendix 1), were reconstructed one at a time using 13 male Caucasian facial templates (See Section 3.1.1.1 and Appendix 2). As a result, 6 faces reconstructed using average facial templates and 20 faces reconstructed using single facial templates (i.e. a total of 26 facial reconstructions) were generated and prepared for the assessment. The 2D facial reconstructions were compared to the targets' 2D ante mortem colour photographs via two types of subjective assessment tests.

(1) *Face Pool Test Form (A)* (Section 3.1.4):

The aim of face pool tests was to assess the ability of observers to identify a target subject from a face pool of faces, including the target and similar individuals. Images were presented in frontal views. Each test consisted of two rows of facial images; the upper row contained one colour photograph of the target individual, and the bottom row consisted of four computer generated facial images; including one facial reconstruction of the target. See Appendix (4-A) for an example. Each observer was asked to select only one image from the bottom row that they thought it best resembled the photograph in the upper row. A test instruction form was associated with each test (Appendix 4-B). For each case, tests were conducted 13 times to test the 13 facial reconstructions. Ten observers performed each test, with a total of 130 observers/case, hence 260 responses.

For each skull case (I & II), results from all tests were pooled and the percentage of correct identification was then calculated.

(2) *Face Resemblance Test Form (A)* (Section 3.1.4):

The aim of the face resemblance tests is to assess whether the degree of similarity/resemblance between the reconstructed and the target faces would be sufficient to allow its correct identification had it been advertised in real life. Images were presented in frontal views. Each test consisted of 2 facial reconstructions of the two skulls using the same facial template (i.e. a total of 13 tests). An example of the Resemblance Test Form (A) is shown in Appendix 5-A. Each observer was asked to give each facial reconstruction a score from 1 to 10 according to their similarity to the target face photograph, where 1 = no resemblance and 10 = the highest resemblance to the target. A test instruction form was associated with each test (Appendix 5-B). For each case, tests were conducted 13 times to test the 13 facial reconstructions. Ten observers have performed each test, with a total of 130 observers/case, hence 260 responses. For each skull case (I & II), results from all tests were pooled and the total resemblance scores of each of the 13 reconstructed faces was then calculated.

The correct identification percentage via face pool test form (A) and the resemblance score via face resemblance test form (A) of each facial reconstruction were also compared.

- **Results**

To analyse the results of a face pool test, the percentage of selecting any face from the face pool by random chance is first calculated. For example, the chance of selecting a face from 10 faces is 1:10 (i.e. 10%), and from 5 faces is 1:5 (i.e. 20%), and so on. So, to estimate the significance of a face pool test, the percentage of correct identification of the target's face is calculated, followed by deducting the chance rate initially calculated. Significant results are those which are above random chance (Wilkinson et al., 2006, Moyers, 2007).

- The chance of correct identification of any of the four faces in the designed face pool tests in this experiment (Appendix 4-A) is 1:4 (i.e. 25%). The identification rate above chance is considered significant.
- In both studies skulls, thirteen out of the twenty six reconstructed faces (13/26, i.e. 50 %) were correctly identified above chance (Figure 4 and Figure 5).
- In both studies skulls, five out of the six faces reconstructed using the 3 average facial templates (5/6, i.e. 83 %) were correctly identified above chance. In contrast, eight out of the twenty faces reconstructed using the 10 single facial templates (8/20, i.e. 35 %) were correctly identified above chance. These two proportions were statistically significantly different (P = 0.01).
- In both studies skulls, two out of the three average facial templates (2/3, i.e. 67%) resulted in correct identification above chance, in contrast to only one out of the ten single facial templates (i.e. 1/10, 10 %). These two proportions were statistically significantly different (P = 0.05).
- In skull (I), of an estimated age of 18 – 30y old, the identification percentages of the faces reconstructed with the 30y old, single and average, facial templates were higher than those with the 20y old facial templates. Moreover, the identification percentage of the face reconstructed with the 30 – 39y average facial template (30Y-AV) received the highest resemblance score and identification rate compared to other facial templates (Figure 4).
- In skull (II), of an estimated age of 20 – 25y old, the identification percentages of the faces reconstructed with the 20y old facial templates were higher than those with the 30y old facial template. Moreover, faces reconstructed with the 20 – 29y average facial template (20Y-AV) received the highest resemblance score and the second highest identification rate compared to other facial templates (Figure 5).

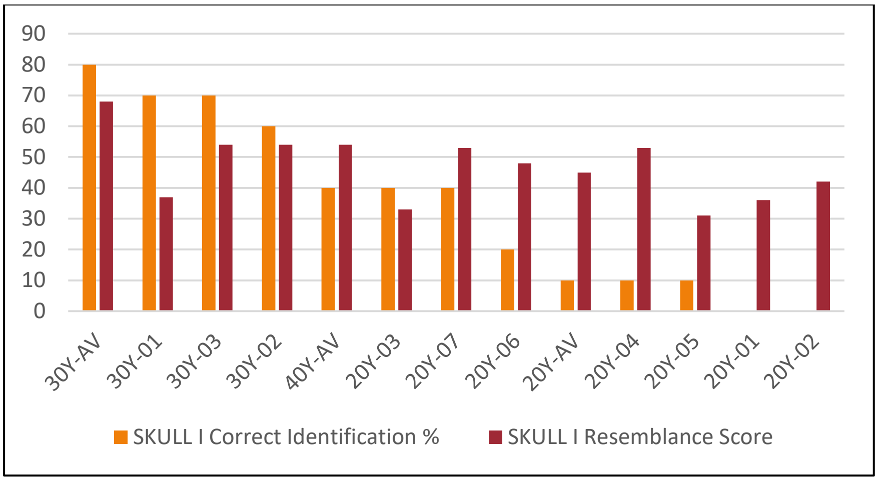


Figure 4: A chart showing the percentages of correct identification and the resemblance scores for faces reconstructed from skull case (I) (30y old) using the 13 facial templates.

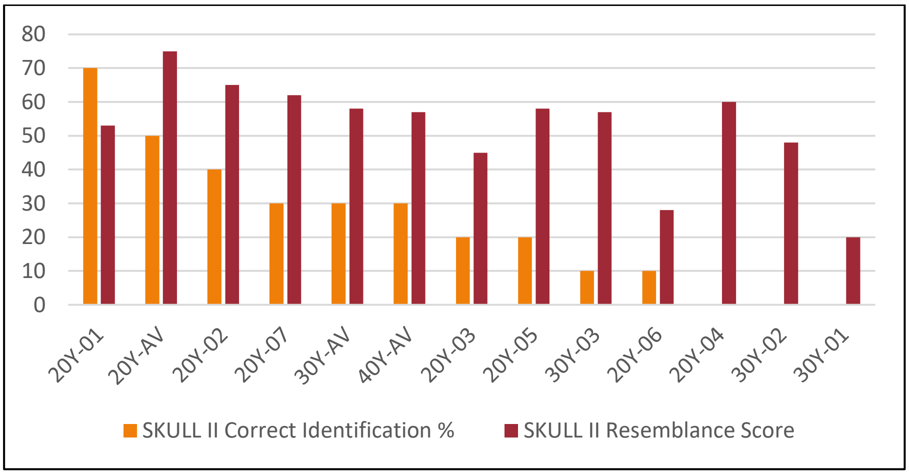


Figure 5: A Chart Showing the percentages of correct identification and resemblance scores for faces reconstructed from skull case (II) (20y old) using the 13 facial templates.

3.2.2 Part Two

- **Methods**

This part of the experiment included three male Egyptian cases aged 20 – 29 years old selected from the CT scanned database (Section 3.1.1.2). The 3D skulls (Appendix 6-A) and the facial templates (Appendix 6-B) were segmented as described in Section 3.1.1.2. From the CT scanned database (Section 3.1.1.2), three single Egyptian facial templates, that match the skulls' sex, race, and age groups were selected and labelled single (01), single (02), and single (03). Using a commercial software (Geomagic Wrap®), each 2 of the 3 segmented single faces were then digitally averaged into a new average facial template. The averaging process involved manual point-based aligning of a number of the single facial templates as 3D meshes, followed by automatic calculating and averaging the distances between them to generate one 3D mesh representing the average facial template.

This has resulted in 3 single (Appendix 6-B) and 3 average (Appendix 6-C) faces from the 3 cases. The 6 faces (3 single and 3 average) were used for the facial reconstructions of the 3 studies skulls. Facial reconstruction was performed using the facial reconstruction method and employing the FR software as described in Section 3.1.3, and using a combined set of Egyptian and European soft tissue depths to suit the studied Egyptian population (Section 4.2.2). The face of each skull was reconstructed 3 times; first using the single facial templates of the 2 other cases, then using the average face generated from these 2 faces, with a total of 9 facial reconstructions. For example, the single faces (single-02) and (single-03) and their averaged face (AV-02+03) were individually used for the facial reconstruction of skull case (01). Also, the single faces (single-01) and (single-03) and their averaged face (AV-01+03) were individually used for the facial reconstruction of skull case (02). Finally, the single faces (single-01) and (single-02) and their averaged face (AV-01+02) were individually used for the facial reconstruction of skull case (03).

The 3D reconstructed faces and the 3D real faces segmented from the CT scans were then aligned in (Meshlab®) and objectively compared by measuring the distance between the 2 faces as Root Mean Square (RMS) (in absolute units). The lower the RMS, the closer the fit between the real face and the facial reconstruction. In the second part, the comparison was conducted objectively via surface distance comparison between the real and the reconstructed faces after being objectively aligned via computer software (Section 3.2.2).

- **Results**

Figure 6 shows the measured Root Mean Square (RMS) distances (in absolute units) between:

- The real CT face of Case (01) and its facial reconstructions using the single facial templates (02 & 03) and the average facial template (02+03),
- The real CT face of Case (02) and its facial reconstructions using the single facial templates (01 & 03) and the average facial template (01+03), and
- The real CT face of Case (03) and its facial reconstructions using the single facial templates (01 & 02) and the average facial template (01+02).

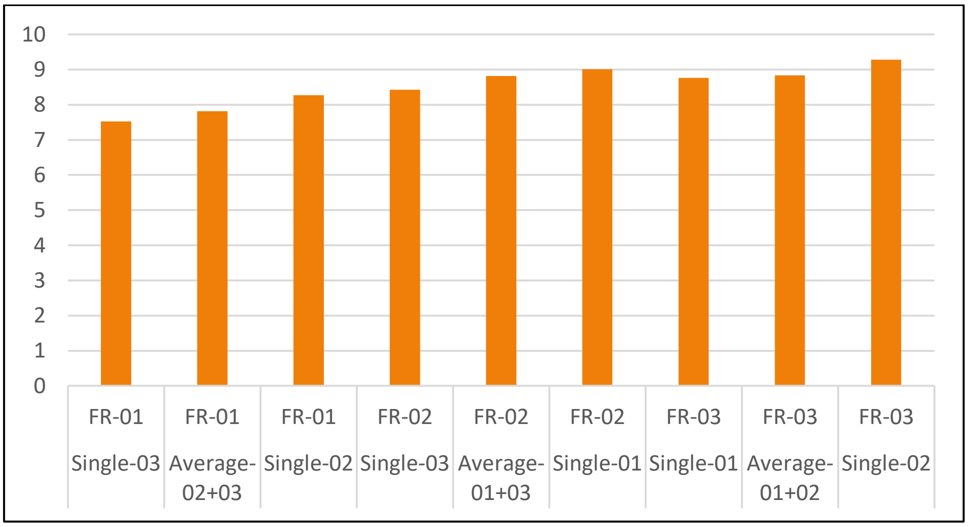


Figure 6: A chart showing the measured Root Mean Square (RMS) distances (in absolute units) between the real CT face of each case and its facial reconstructions using the single and the average facial templates.

When the single face (03) was used as a facial template to reconstruct the other 2 skulls, it showed the lowest RMS results (i.e. the closest fit between the real and the reconstructed faces), compared to the other 2 single facial templates (Figure 7). This indicates that certain single faces are better than others as facial templates for forensic facial reconstruction.

Similarly, when the average face (02+03) was used as a facial template to reconstruct the other 2 skulls, it showed the lowest RMS, compared to the other 2 average facial templates (Figure 7). However, in all cases, the mean of the combined RMS distances of the faces reconstructed using the three average faces was slightly lower (i.e. showed a closer fit between the real and the reconstructed faces) than that of the reconstructed faces using the three single faces.

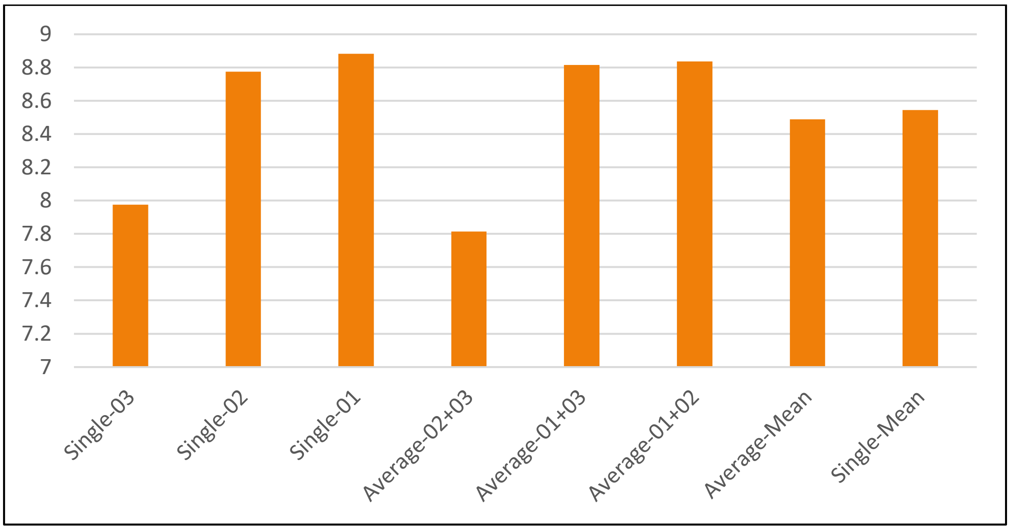


Figure 7: A chart showing mean of the Root Mean Square (RMS) distances (in absolute units) between the real CT faces and the reconstructed faces of the three cases separately and combined.

3.3 EXPERIMENT TWO: THE INFLUENCE OF FACIAL SOFT TISSUE DEPTHS ON 3D FORENSIC FACIAL RECONSTRUCTION

Tis experiment consisted of two parts designed to investigate how the changes in the facial soft tissue depths would influence the resulting facial reconstructions.

3.3.1 Part One

The aim of this part of the experiment was to explore whether modifications in the cheek region landmarks affects the facial reconstructions.

- **Methods**

The faces of fifteen (8 males and 7 females) Egyptian skulls from the Computed Tomography (CT) scanned database were reconstructed (Cases were described in Section 3.1.1.2). For each studied skull, an average facial template matching the skull's sex, race and age group was selected from the database of the laser scanned faces (Section 3.1.1.1). Facial reconstruction of the fifteen skulls was performed as described in Sections 3.1.2 and 3.1.3, and using the facial soft tissue depths of each of the following sets of landmarks (i.e. a total of sixty facial reconstructions):

(1) **Landmark Set (1):** including the full 40 landmarks of Rhine and Moore (1982) and Helmer (1984). This set incudes 40 landmarks; 10 in the midline and 14 on each side of the face with a total of 40 landmarks (Appendix 12, Table 17).

(2) **Landmark Set (2):** including 38 landmarks, after omitting the right and left Occlusal Line landmarks from the full set.

(3) **Landmark Set (3):** including 36 landmarks, after omitting the right and left Supra and Sub M2 landmarks from the full set.

(4) **Landmark Set (4):** including 34 landmarks, after omitting the right and left Occlusal Line, and right and left Supra and Sub M2 landmarks from the full set.

The 3D reconstructed faces were then objectively compared to the 3D real faces segmented from the CT scans using Root Mean Square (RMS) of surface distance differences (in absolute units).

- **Results**

Table 3 shows the Root Mean Square (RMS) distance differences (in absolute units) between the faces reconstructed of each case using the 4 sets of craniofacial landmarks and their respective CT faces. Using ANOVA, the RMS distances between the 4 sets were not significantly different (P-Value = 0.998). The highest and the lowest RMS median values were seen with the 36, and the 40 landmarks sets respectively.

Table 3: The Root Mean Square (RMS) distances (in absolute units) between the faces reconstructed of each case using the 4 sets of craniofacial landmarks and their respective CT faces

CASE	34LM*	36LM**	38LM***	40LM****
01	08.23	08.12	08.13	08.05
02	07.78	07.63	07.84	07.75
03	09.41	09.11	09.47	09.18
04	09.47	09.51	09.46	09.52
05	10.19	10.10	09.86	09.92
06	12.33	12.23	12.18	12.12
07	09.18	09.10	09.23	09.15
08	08.66	08.69	08.51	08.55
09	09.03	08.86	08.96	08.86
10	11.46	11.41	11.68	11.57
11	08.68	08.56	08.65	08.50
12	10.26	10.12	10.03	09.99
13	07.04	07.03	06.81	06.80
14	11.42	11.55	11.42	11.50
15	07.01	07.05	06.90	06.93
Median	**9.18**	**9.10**	**9.23**	**9.15**

***34LM:** 34 landmarks, after omitting the right and left Occlusal Line, and right and left Supra and Sub M2 landmarks from the full set.
****36LM:** 36 landmarks, after omitting the right and left Supra and Sub M2 landmarks from the full set.
*****38LM:** 38 landmarks, after omitting the right and left Occlusal Line landmarks from the full set.
******40LM:** The full 40 landmarks of Rhine and Moore (1982) and Helmer (1984).

3.3.2 Part Two

Over decades, various methods were employed to collect these measurements in different populations (Section 1.5). Rhine and Moore (1982) and Helmer (1984) data (Appendix 12, Table 17) were collected using needles from cadavers and ultrasound scans from living people respectively (Stephan and Henneberg, 2001, Vanezis, 2008). These data have been used by many researchers (Hayes, 2014). More recently, in 2008, Stephan and Simpsons started analysing and pooling the published facial soft tissue thickness data for adults (Stephan and Simpson, 2008a) and sub-adults (Stephan and Simpson, 2008b) that were collected by other researchers from both cadavers and living subjects using different methods (e.g., needle insertion, Ultrasound, CT, MRI). Updated data are published regularly in the researcher's website (Stephan), for other researchers to use in forensic facial reconstruction studies.

- **Methods**

In this part of the experiment, a comparison was conducted between two sets of facial soft tissue depths. The old set (Set A) was of the full 40 landmarks set defined by Rhine and Moore (1982) and Helmer (1984) (Appendix 12, Table 17). The new set was adopted mainly from the most recent data set presented by Stephan (2014) (Appendix, Table 18) at the time of performing this study. However, the latter data included 36 landmarks only; 14 in the midline and 11 on each side of the face. Of them only 26 landmarks were found to be common with Set (A) regarding their names and anatomical locations. As a result, following Parks et al. (2014), a modified was prepared (Set B). This second set was composed of the common 26 landmarks but with using the recent measure values from Stephan (2014) (Appendix 12, Table 19) in addition to the remaining 14 landmarks with using the old measures taken from old Rhine and Moore (1982) and Helmer (1984) set. This was done to ensure an equal comparison with Set (A) which was composed of 40 landmarks.

To compare between sets (A) and (B), 4 laser scanned skulls and 15 CT scanned skulls were reconstructed using facial soft tissue set (A) (Appendix 12, Table 17) and set (B) (Appendix 12, Table 19) separately. For each studied skull, an average facial template matching the skull's sex, race and age group was selected from the database of laser scanned faces (Section 3.1.1.1). This has resulted in 8 and 30 facial reconstructions from the laser and CT scanned skulls respectively.

Comparisons were then conducted subjectively, for the first 8 facial reconstructions, and objectively, for the following 30 facial reconstructions. For the *subjective* comparison, the reconstructed faces were compared with the real faces' photographs via face pool and face resemblance tests, with a total of 16 tests. Ten volunteer observers' responses were sought for each test (i.e. a total of 160 responses). The identification percentages of all cases in face pool tests and the total resemblance scores given by all observers in face resemblance tests to all cases were then calculated and compared between those using sets (A) and (B) using one-way ANOVA. For the *objective* comparison, the second 30 facial reconstructions were objectively compared via Root Mean Square (RMS) surface distance difference (in absolute units).

- **Results**

Table 4 shows the *subjective* identification percentages of all cases in face pool tests using the facial soft tissue depth sets (A) and (B). The identification rate median of faces reconstructed using the old data (set A) was higher than that of faces reconstructed using the new data (set B). However, using Paired-T Test, the difference between them was not statistically significant (P-Value = 0.628). Moreover, the *subjective* total resemblance scores of all cases in the face resemblance tests using the old (set A) and the new (set B) data were similar (Table 5).

Table 4: The identification percentages of all cases in face pool tests using the old (set A) and the new (set B) facial depths data

Case (Age-Sex)	Facial Depths (Set A)-ID %	Facial Depths (Set B)-ID %
I (30-M)	80	80
II (20-M)	0	40
III (20-M)	50	10
VI (30-F)	40	20
MEDIAN	**45**	**30**

Table 5: The total resemblance scores of all cases in face pool tests using the old (set A) and the new (set B) facial depths data

Case (Age-Sex)	Facial Depths (Set A)-RES. SCORE	Facial Depths (Set B)-RES. SCORE
I (30-M)	68	61
II (20-M)	54	61
III (20-M)	75	63
VI (30-F)	31	42
MEDIAN	**61**	**61**

The *objective* Root Mean Square (RMS) surface distance difference (in absolute units) between the old (set A) data and the new (set B) data (Table 6) showed no significant difference using Paired-T Test (P-Value = 0.244).

Table 6: The objective Root Mean Square (RMS) distance (in absolute units) using the old (set A) and the new (set B) facial depths data

CASE	Facial Depths (Set A)-RMS	Facial Depths (Set B)- RMS
01	08.05	08.21
02	07.75	07.64
03	09.18	09.12
04	09.52	09.53
05	09.92	09.95
06	12.12	12.23
07	09.15	09.06
08	08.55	08.63
09	08.86	08.88
10	11.57	11.46
11	08.50	08.60
12	09.99	10.01
13	06.80	06.95
14	11.450	11.57
15	06.93	06.97
MEDIAN	**9.15**	**9.06**

3.4 EXPERIMENT THREE: DESIGNING SUBJECTIVE ASSESSMENT TESTS FOR 3D FORENSIC FACIAL RECONSTRUCTION

The purpose of the three parts of this experiment was to validate the current design of police line-ups in face pool tests and explore the possibilities of implementing some modifications to better suit facial forensic reconstructions research.

3.4.1 Part One

This part was conducted to answer this research question: *Can the usual line-up format of a face pool test be improved for forensic facial reconstruction assessment?*

- **Methods**

To answer this question, the 2 face pool test forms (Form A and Form B), described in Section 3.1.4, were compared using facial reconstructions of one male skull (Skull I) and one female skull (Skull IV) selected from the laser scanned skull database (Section 3.1.1.1).

For the *male skull case (I)*, the identification rates of the 13 facial reconstructions assessed via a face pool test form (A) in Experiment One were also included in this experiment. The results showed that out of the 13 facial reconstructions of skull case (I), 7 facial reconstructions (7/13, i.e. 54 %) were correctly identified above chance (i.e. > 25 %) (Figure 4). In this experiment, these 7 facial reconstructions were re-assessed using a face pool tests form (B) consisting of a face pool of 8 coloured photographs including the target and 7 "foils" or "fillers" photographs. These photographs were compared against each of the 7 reconstructed faces of the target separately. The foils were selected to match the target's age, sex, race, and general face morphology. For each of the 7 facial reconstructions, 3 different forms of this Face Pool Test (B) were generated (See Appendix 7-A for an example) (i.e. a total of 21 tests). Six to ten different observers have performed each test, with a total of approximately 160 responses. Each observer was asked to select only one photograph that best resembled the target facial reconstruction. A test instruction form is shown in Appendix (7-B). The results from all tests were pooled and the percentages of correct identification obtained for the 7 facial reconstructions via face pool form (B) were compared to those obtained via face pool (A) in experiment one (Section 3.2.1). In the designed test, the chance of selecting any photograph was 1:8 (i.e. 13%). Correct identification rates above this were considered significant.

For the *female skull case (IV)*, aged 30-40 years old, 2 facial reconstructions were generated from the female Skull case (IV), using 2 averaged female faces, selected from the database of laser scanned faces; (20Y-FAV) and (30Y-FAV) (Appendix 8). The two facial reconstructions (Appendix 8) were then assessed first using the black and white ante mortem photograph of the target in frontal view only, then using 3 black and white ante mortem photographs of the target in 3 views (frontal, lateral and 3/4). Assessment was done via face pool test form (A), face pool test form (B) and face resemblance test form (A), with a total of 6 tests for the 2 reconstructed faces (2 face pool tests and 1 resemblance test each). Ten responses were sought for each test; with a total of 60 responses. Results from all tests were pooled and the percentage of correct identification using the face pool test forms (A) & (B) as well as the total resemblance scores given in the face resemblance test form (A) to each reconstructed face were calculated and compared.

The face pool tests form (A) were constructed using 3 computer generated faces, including the facial reconstruction of the target, and two foil similar female faces selected from the database of laser scanned faces. The face pool was compared to an antemortem photograph of the target. This form of the face pool tests was designed first with the faces in frontal view only (Appendix 9-A), then in 3 views (frontal, profile and 3/4) (Appendix 9-B). A test instruction form is shown in (Appendix 9-C). In contrast, face pool tests form (B) were constructed using 3 black and white photographs including the target and two similar female foil photographs. The face pool was compared to the computer generated facial reconstruction of the target. This form of the face pool tests was designed first with the faces in frontal view only (Appendix 10-A), then in 3 views (frontal, profile and 3/4) (Appendix 10-B). A test instruction form is shown in (Appendix 10-C). Resemblance Test Form (A) entailed direct comparison between the target's photograph and each of the reconstructed faces separately. The test was first designed with both images in frontal view only (Appendix 11-A), then in 3 views (frontal, profile and 3/4) (Appendix 11-B). A test instruction form is shown in Appendix (11-C).

The results of the female case were also used in the second part of this experiment.

- **Results**

For the male skull Case (I), the percentages of correct identification of the initial 13 facial reconstructions using the face pool test form (A) in Experiment One, Part One are shown in Figure 8.

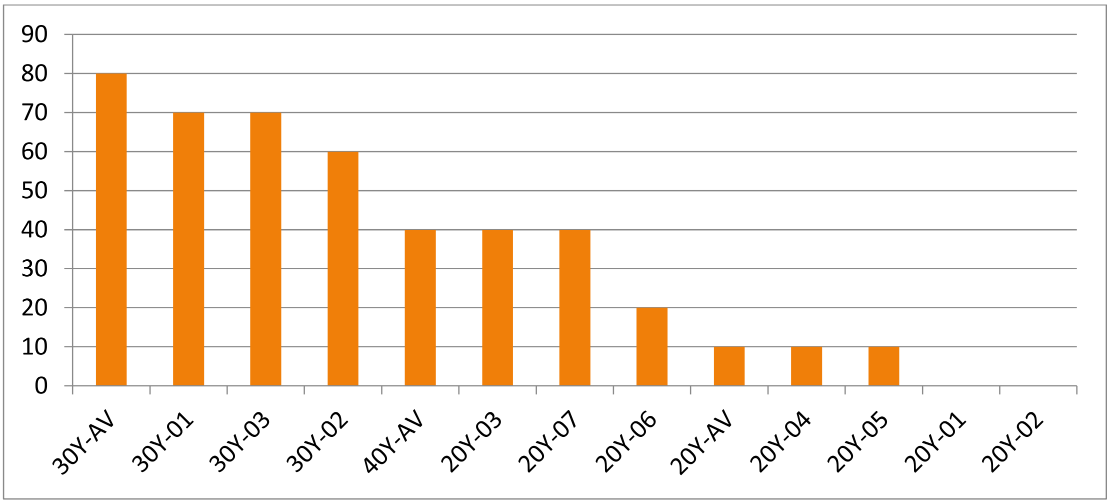


Figure 8: A chart showing the correct identification percentages of the initial 13 reconstructed faces of skull case (I) using the face pool test form (A) in Experiment One, Part One

Figure 8 shows that 7/13 facial reconstructions were identified above chance (25%) using Face Pool Form (A) (Appendix 4-A). Only these 7 facial reconstructions were re-tested using Form (B) (Appendix 7-A) and compared with their identification percentages using form (A). In 6 out of 7 (i.e. 86 %) of the tested facial reconstructions, the correct identification rates were higher using Face Pool Form (A) than Form (B) (Figure 9).

While in 4/4 (100 %) of the face pool tests of the female skull Case (IV), the correct identification rates were higher using Face Pool Form (A) than Form (B) (Figure 10 and Figure 11).

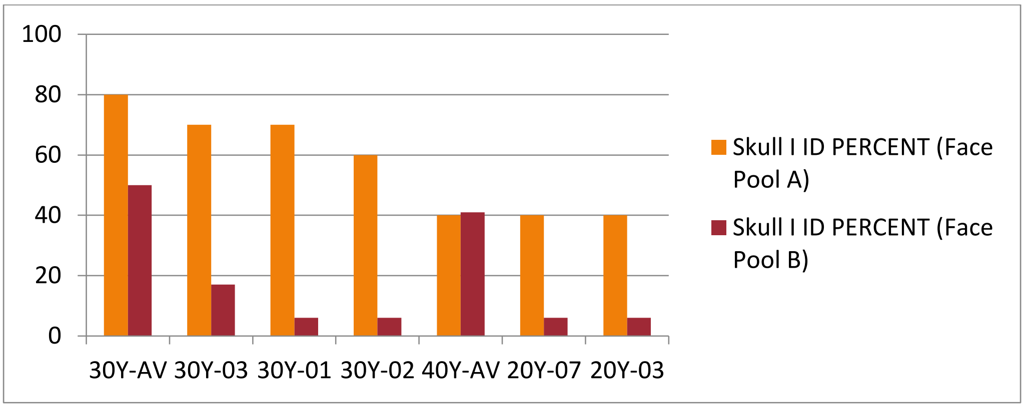


Figure 9: A chart showing the percentage of correct identification for skull case (I) obtained using face pool Form (A) and Form (B).

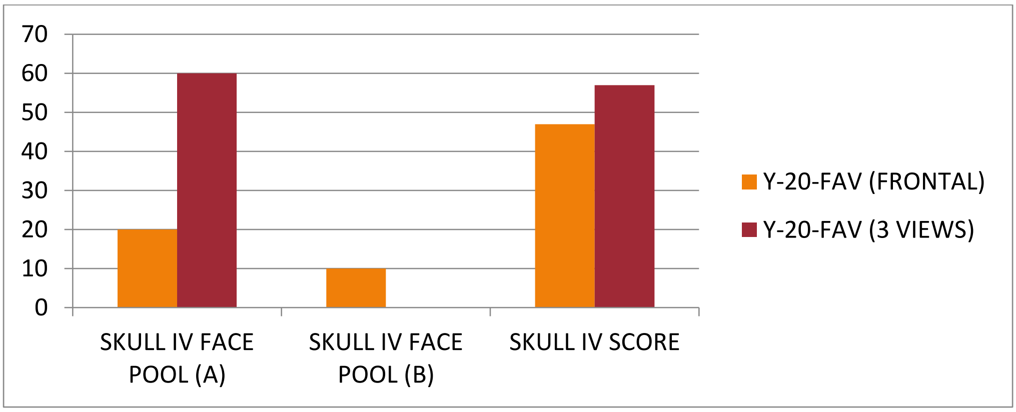


Figure 10: A chart showing the percentage of correct of the facial reconstruction of Skull case (IV) using the 20Y-FAV facial template.

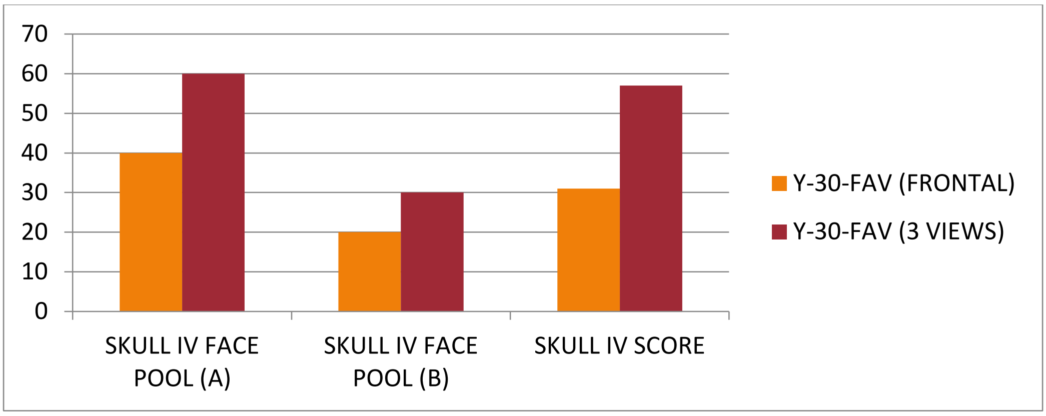


Figure 11: A chart showing the percentage of correct of the facial reconstruction of Skull case (IV) using the 30Y-FAV facial template.

3.4.2 Part Two

This part was conducted to answer this research question: *Can multiple orientations improve the identification rates?*

- **Methods**

Results obtained in Part One of this experiment, for the female skull case (IV), were re-analysed with the aim of exploring the differences between testing the target in frontal view only and compared to the 3 views (frontal, lateral and 3/4).

- **Results**

In 5/6 (i.e. 83 %) of the tests, those designed using photographs with multiple face views' showed higher identification rates and/or resemblance scores than with using photographs with frontal views only (Figure 12).

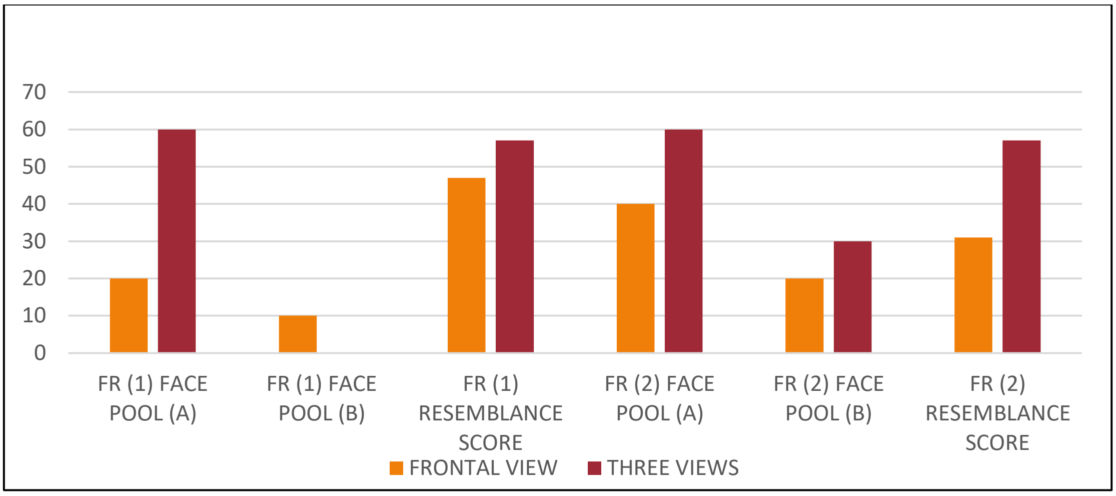


Figure 12: The correct identification percentages and resemblance scores of the 2 facial reconstructions of one skull using a frontal view and 3 views pictures

3.4.3 Part Three

This part was conducted to answer this research question: *Does the facial expression of the target in the photograph presented to assessors influence the identification rates?*

- **Methods**

To answer this question, 13 faces were reconstructed of the *male skull case (III)* using the 13 facial templates described in Section 3.1.1.1 and shown in Appendix 2. These faces were tested via face pool test form (A) using a colour photograph of the target with neutral expression (Picture 1) (Figure 13a), then retested also via the same face pool tests form, but with using another color photograph of the target showing an expressed smile (Picture 2) (Figure 13b).

a (Picture 1)

b (Picture 2)

Figure 13: The two pictures of skull case (III) used for facial expression experiment

Ten responses were sought for the 26 with a total of 260 responses. Results from all tests were pooled and the percentage of correct identification of each reconstructed face of skull case (III) using picture (1) and (2) were compared.

- **Results**

Using picture (1), with the neutral expression, 7/13 (i.e. 54%) of the facial reconstructions showed higher identification rates, 3/13 (i.e. 23%) of the facial reconstructions showed equal identification rates than using picture (2). In contrast, using picture (2), with an expressed smile, only 3/13 (i.e. 23%) of the facial reconstructions showed higher identification rates than picture (1) (Figure 14).

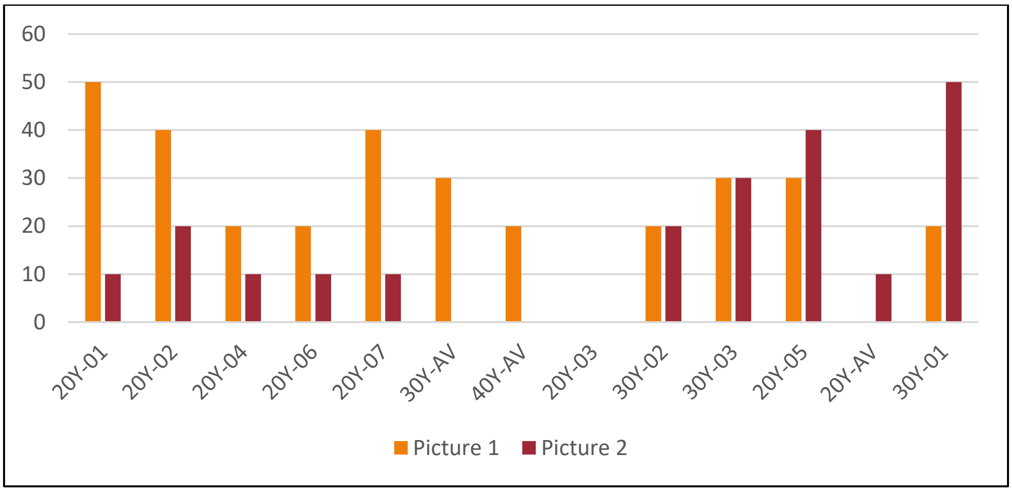


Figure 14: The correct identification percentages of skull case (III) using picture (1) and picture (2)

3.4.4 Part Four

This part of this experiment aimed to answer this research question: *Would the assessors' characteristics affect the identification rates in the face pool tests?*

In other words, whether the observers' own sex, race or age affects their correct identification rates of the facial reconstructions in face pool tests was investigated. If positive, this should be useful in designing the assessment tests for the next stage of the research project by selecting the group of observers that are more suitable for performing the tests.

- **Methods**

In this pilot study, 876 face pool tests were performed, which also involved information about the observers' age, sex and race (See Tests Instruction Forms in Appendices 4-B, 5-B, 7-B, 9-C, 10-C, and 11-C). This information was analysed to test the possibility of bias of the observer's selections by a certain sex, race or age group for higher identification rates. The percentages of the face pool tests in which the correct facial reconstruction was identified by male and female observers as well as by the Asian, African, Caucasian, and mixed races were calculated. In addition, variations among different age groups (< 30, 30 – 49 years, and 50 – 69 years) were also investigated.

- **Results**

The sex of the observer was available in 874 tests. Female observers answered 520 tests. Of them, 147 tests (i.e. 28%) were correctly answered (95% CI: 0.24-0.32). In contrast, male observers answered 354 tests. Of them, 83 tests (i.e. 23%) were correctly answered. Figure 15 shows all the identification percentages. There was no statistically significant difference between the performance of males and females (P = 0.118).

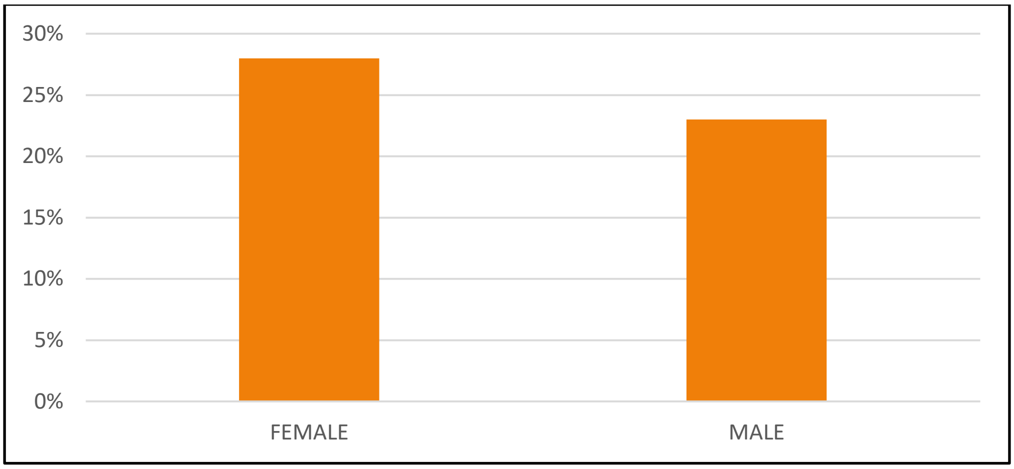


Figure 15: A chart showing the percentages of correct identification in the face pool tests by males and females

The race of the observer was available in 849 tests. Caucasian observers answered 452 tests. Of them, 95 tests (i.e. 21%) were correctly answered (95% CI: 0.17-0.25). Asian observers answered 296 tests. Of them, 79 tests (i.e. 27%) were correctly answered (95% CI: 0.22-0.32). African observers answered 30 tests. Of them, 8 tests (i.e. 27%) were correctly answered (95% CI: 0.11-0.43). Observers of mixed race answered 71 tests. Of them, 19 tests (i.e. 27%) were correctly answered (95% CI: 0.17-0.37). Figure 16 shows all the identification percentages.

Although the tested cases were of Caucasian race, Caucasian observers appear to be less able to identify the cases of their own race. However, the P-Values of the differences between the Caucasian race and the Asian, African and mixed races were 0.077, 0.491, and 0.281 respectively, hence this difference was not statistically significant. The small sample size of the African observers compared to other groups (3.5 %) should be taken into account.

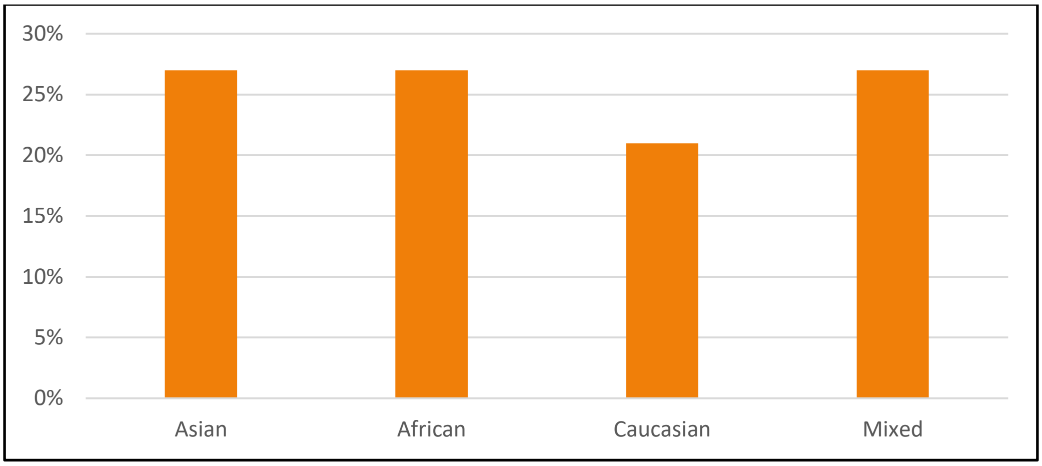


Figure 16: A chart showing the percentages of correct identification in the face pool tests by different races

The age of the observer was available in 874 tests, with age ranging between (17 – 61 years). The observers aged < 30 years answered 542 tests. Of them, 159 tests (i.e. 29%) were correctly answered (95% CI: 0.26-0.33). The observers aged 30 – 49 years answered 309 tests. Of them, 64 tests (i.e. 21%) were correctly answered (95% CI: 0.16-0.25). The observers aged 50 – 69 years answered 23 tests. Of them, 7 tests (i.e. 30%) were correctly answered (95% CI: 0.12-0.49). Figure 17 shows all the identification percentages.

The only statistically significant difference, however, was found between the < 30 years old and the 30 – 49 years old observes (P = 0.006). The small sample size of the 50 – 69 years old observers compared to other groups (2.6 %) should be taken into account.

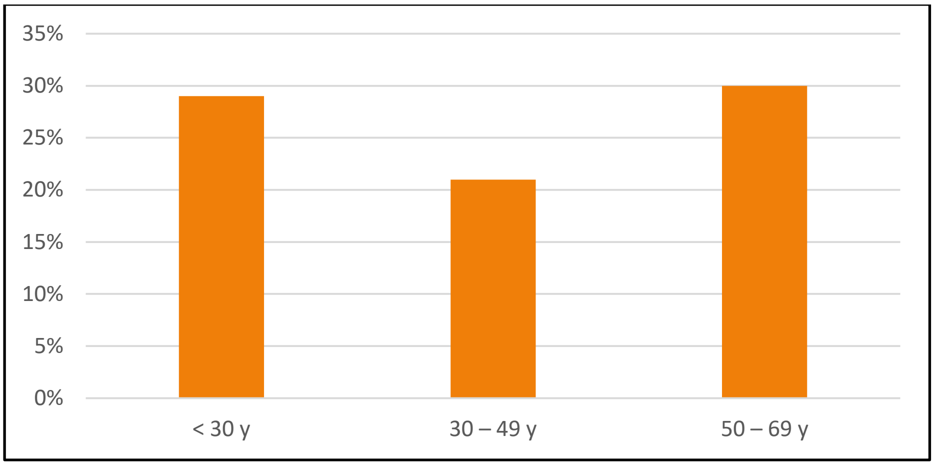


Figure 17: A chart showing the percentages of correct identification in the face pool tests by different age groups

3.5 EXPERIMENT FOUR: TESTING THE APPLICABILITY OF THE PROPOSED METHOD OF FACIAL RECONSTRUCTION

In this experiment, the variation between users in applying the facial reconstruction technique proposed in this study, using the present software, was investigated. The aim was to test the applicability and the subjectivity of the described method as well as the ease of use of the facial reconstruction (FR) software. This experiment was conducted over three stages. Initially, a user manual (V1) for the present software was prepared by the researcher. This first version was tested by a volunteer (user 1) with no previous contact with the facial reconstruction software nor experience in the forensic facial reconstruction field. The aim of this stage was to use the suggestions provided by user (1) to the main researcher to produce a fully developed manual that can be used by other users who perform the same facial reconstruction process and using the same facial reconstruction software. Accordingly, a second version (V2) was then prepared and a second volunteer (user 2), also with no previous contact with the facial reconstruction software nor experience in the forensic facial reconstruction field, was recruited. User (2) was asked to attempt to reconstruct four cases (2 males and 2 females) Egyptian skulls from the CT scanned database following the user manual updated version (V2) only, and under blind condition. The same four cases were also reconstructed again by the main researcher (user 3) under non-blind conditions. All faces were reconstructed using average facial templates matching the skull’s sex, race and age group was selected from the database of the laser scanned faces for the facial reconstruction. Faces reconstructed by users (2) and (3) were compared by being individually assessed against the respective targets’ CT segmented real faces via objective Root Mean Square (RMS) differences of the surface distances.

- **Results**

Figure 18 shows similar Root Mean Square (RMS) of the surface distance differences between the CT face of each target and the faces of the same target reconstructed by user (2) under blind conditions and user (3) under non-blind conditions. Comparison was statistically insignificantly different (P-Value = 0.981). In addition, user (2) also provided minor suggestions to further improve the manual. These suggestions have been taken into account and a 3rd and most updated version (V3) of the user manual has been prepared (Appendix 20).

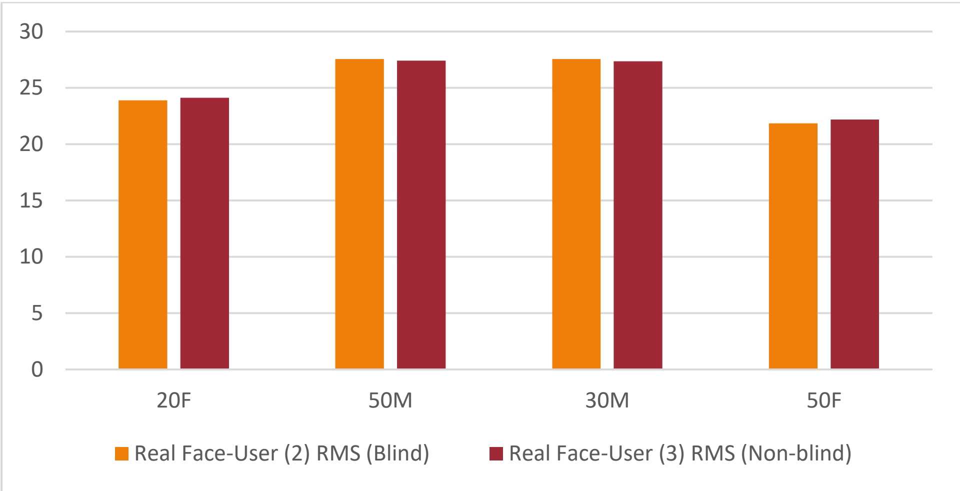


Figure 18: Surface comparison between the CT Face of each target and the faces of the same target reconstructed by non-expert user (2) under blind conditions and an experienced user (3) under non-blind conditions

3.6 The Pilot Study Summary and Conclusions

In the pilot study, the faces of 19 skulls were reconstructed with a total of 58 different facial reconstructions. In a number of preliminary experiments, 43 facial reconstructions were subjectively assessed via face pool and face resemblance tests and 15 facial reconstructions were objectively assessed via measuring the overall surface distance differences between each real and reconstructed face. The skulls and the faces of the cases studied in the pilot study were scanned by a laser scanner or a computed tomography scanner.

Initially, this pilot study showed that certain reconstructed faces received high and close ranks in both the face pool and face resemblance tests, as well as lower objective surface distance comparison between the real and the reconstructive faces than others.

Furthermore, the results of the pilot study showed that the averaged facial templates received higher identification rates and resemblance scores than most single faces used. It was, therefore, decided, for the main part of this thesis to use average facial templates for the facial reconstruction (Section 3.2).

In addition, this pilot showed an insignificant difference between an old and a new facial soft tissue depths data sets published approximately 30 years apart. Another experiment showed that although it is important to have a standardised set of population-specific facial depths for the facial reconstruction, a previously validated set, even if old, is sufficient. It is rather more important, to categorise the facial landmarks into more and less influential landmarks, without leaving this to the subjective practitioner's experience and judgment to add or omit facial landmarks (Vanezis, 2008). Therefore, efforts should be directed to standardising a craniofacial landmarks set, with proper definitions and accurate description of the locations and the directions between the corresponding cranial and facial landmarks with describing the best orientation for placing each landmark (Brown et al., 2004). Therefore, for the main part of this thesis, a previously published set of facial soft tissue depths for the Egyptian population (El-Mehallawi and Soliman, 2001) was used in combination with the midline landmarks of Rhine and Moore (1982) and Helmer (1984)'s set as being missing from the former set (Section 3.3).

This pilot study also showed that the usual format of the face pool test adopted in literature, could be improved for forensic facial reconstruction research with the aim of reaching a proper design that can more reliably reflect the accuracy of the tested facial reconstructions. It was concluded that it is better to use facial templates with closed eyes, no hair, neutral facial expression, and of similar complexion to allow the observers to base their selections on the shape of the skull itself with no distracting facial features. Furthermore, multiple views of the facial image of the same target allowed more reliable assessment of the facial reconstructions and thus a higher chance of correct identification as these views familiarised the observers with the target's face shape than facial images in frontal views only. Following on from that conclusion, rather than presenting 2D images with multiple views to the assessors, it was decided, for the main part of this thesis, to present the faces in a way that allows assessment of the faces from all angles by interactively rotating them in an online 3D view (Section 3.4).

Looking at the relationship between the assessors' characteristics (sex, age and race) and the correct identification rates in the face pool tests, the preliminary findings of this pilot study suggest that female observers aged < 30 years old represent an ideal group of observers for subjective assessment of forensic facial reconstructions by face pool tests. However, the sample size in certain observers groups were not large enough to restrict the recruited participants of the main part of the thesis to this group (female observers aged < 30 years old) and exclude others3.4.4).

CHAPTER 4: THE MAIN STUDY MATERIALS AND METHODS

In this section, the methodology used in the main study is described; starting from the acquisition of the materials (the skulls and facial templates), the preparation of the obtained materials, performing the facial reconstruction, and then assessing the resulting reconstructed faces.

4.1 ACQUISITION OF THE COMPUTED TOMOGRAPHY (CT) SCANS

The original protocol planned for this project was to obtain head CT scans of adult male and female Caucasian patients undergoing head CT scanning in the radiology department of St Bartholomew's Hospital, West Smithfield, Barts Health NHS trust, London, UK. This protocol was ethically approved by the Newcastle and North Tyneside National Research Ethics Committee (NREC) in May 2014.

4.1.1 Patient Recruiting and Consenting

Data required for this study included head CT scans as well as face photographs of the patients in five views (frontal, right profile, right three quarters, left profile, and left three quarters). It was planned to recruit a minimum of 10 case studies. Patients meeting the following inclusion criteria were aimed.

- **Inclusion Criteria**

1) Race: Caucasian. Caucasian ethnicity is the general physical type of some or all of the populations of Europe, North Africa, the Horn of Africa, Western Asia, Central Asia and South Asia.
2) Age: Adults (i.e. over 16 years of age).
3) Able to understand and willing to sign a written Informed Consent Form.
4) Able and willing to follow the protocol requirements.
5) Undergoing CT scanning of the head for medical or clinical purposes.
6) An intact complete head CT scan.

7) Complete patient record (including the age and sex of the patient).
8) It is **<u>Optional</u>** if any subject agrees to the access and use of their CT scans and/or face photographs for *teaching, research, or future research.*

Conversely, patients meeting the following exclusion criteria were excluded.

- **Exclusion Criteria**

1) Race: non-Caucasian.
2) Age: less than 16 years old.
3) Unable to understand or refusing to sign the written Informed Consent.
4) Unable or refusing to follow the protocol requirements.
5) Head trauma causing bone or soft tissue deformity or damage or scans with implantable devices (i.e. eye, head and neck, or cardiac).
6) Incomplete head scans.
7) Incomplete patient records.

As the facial reconstructions were planned to be performed under blind conditions, another research member was delegated, instructed and trained for consenting and photographing the patients. In addition, the main investigator and the delegate person, had a preliminary visit to the radiology department of St Bartholomew's Hospital to agree on a suitable private location to consent and photograph the patients. Preparatory meetings were held with the local co-investigator in the hospital, the consultant radiologist, to introduce the research team members to the radiology department staff and agree on the relevant arrangements. It was agreed that approaching the subjects will be conducted after they had scanned. A subsequent visit was conducted to set the photographing equipment and test the whole process prior to starting patients' recruitment.

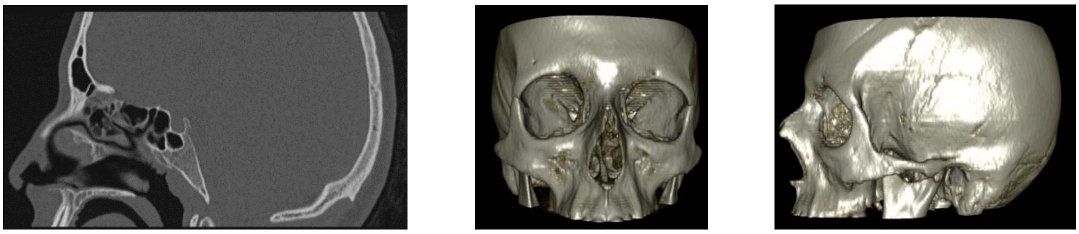
4.1.2 Patient Photographing

In July 2014, 7 patients scheduled for head CT scans were screened. The indication for scanning these patients was paranasal sinus disease. Two of these patients were excluded for not meeting the inclusion criteria (i.e. of African race, and having facial soft tissue pathology). The remaining 5 patients were approached and provided with a full explanation of the nature, the purpose and requirements of the study. Patients were given a concise patient information sheet (PIS) to read and keep, and asked to fill in the appropriate consent forms if they agreed to participate. Out of the 5 approached patients, 2 patients, a male aged 36 years old and a female aged 21 years old, agreed to participate. They have consented to; (1) use their head CT scan images that have already been taken as a part of their planned clinical care and acquired in accordance with the routine clinical scan protocols in the present study, and (2) take photographs of their face with a digital camera in 5 views (frontal, right profile, right three quarters, left profile, and left three quarters).

The face photographs of each patient were acquired by the delegate person using a Sony Cybershot DSC-RX100 Digital Camera (20.2MP 1.0-type Exmor R CMOS sensor, F1.8 lens, 3.6x optical zoom, Full HD 50p, 7.5cm) 3 inch LCD. The photographs were taken in colour, with the eyes open and clearly visible, with a neutral expression with the mouth closed (no grinning, frowning or raised eyebrows), and no hair across the eyes, similar to taking a passport photograph taken. In addition, the photographs were taken in sharp focus with a strong definition between the face and background with nothing covering the face. The photographs included the full head, neck and shoulders. A space around the full head, was included in the viewfinder or screen display of the camera. Furthermore, photographs were free from "redeye", airbrushing or similar enhancement, with no spectacles or sunglasses to avoid covering of the eyes by the frames as well as any reflection or glare on the glass. To reduce shadows on the background, the distance between the person being photographed and the background was minimised and free from patterns, objects or textures. Moreover, proper lighting and uniform illumination of the background was ensured to remove any shadows or other lighting effects that would otherwise interfere with clearly discerning the facial outline on the background. During photographing, the patient was sitting on a swivel chair at 90º with his/her head, neck and back at the same straight level. The camera was placed on a tripod adjusted at the eye level of the person being photographed. Five arrows, separated by 45º each, printed on an A1 poster fixed to the floor was placed in front of the patient to point to the required 5 views. Patients were asked to hold a scale and a label of his/her unique reference number. The scale and the label were obvious in each photograph of the 5 required views, yet not covering the face. Master and individual cases photo logs as well as screening and recruitment logs were filled as appropriate.

4.1.3 Acquiring and Processing the Head CT Scans

CT images were acquired by spiral multislice computed tomography (MSCT) scanner (SOMATOM Sensation 64) that is present in the imaging and X-ray department of St Bartholomew's Hospital, West Smithfield, Barts Health NHS trust, London, UK. The SOMATOM Sensation 64 CT-system is equipped with an x-ray tube that acquires the images in slice-by-slice imaging mode in which there is no table movement during data acquisition (SOMATOM Sensation 64 Application Guide, 2005). Scans were taken with a slice thickness of 0.6 mm and FOV (Field Of View) of 512 mm. CT images stored in Digital Imaging and Communications in Medicine (DICOM) format as two-dimensional cut series (Figure 19a), were then segmented and the 3D skulls were segmented via the (InVesalius®) software. Through this software, the 3D skull was reconstructed (Figure 19b, c).



a b c

Figure 19: A CT scan of one of the recruited patients, lateral view (a), A 3D reconstructed skull segmented from the CT scan in Figure (2) using InVesalius© software, frontal view (b), lateral view (c)

As seen in Figure 19, the protocol followed by the NHS for scanning the paranasal sinus patients was to scan only the part of the head where the sinuses were located (i.e., the forehead, orbits and maxilla). Thus, the obtained scanned images were incomplete, which is one of the exclusion criteria. On the other hand, patients admitted for full head CT scanning were those in the accidents and emergency department, thus incapable of being recruited for this study. As a result, the study was terminated in this trust and all the collected data (head CT scans and photographs) of the 2 recruited patients were discarded as per the ethics regulations.

The next step was to search for an alternative source of head CT scans. Colleagues in two different overseas locations were contacted:

1- The Department of Public Health and Community Medicine, Section of Forensic and Legal Medicine, University of Verona, Italy.
2- Two Diagnostic Radiology Centres, Zagazig City, Sharkia Governorate, Egypt.

They agreed to provide us with the appropriate data. Data from the first and the second sources were postmortem head CT scans of corpses, and head CT scans of living patients respectively. It was agreed that ten cases from each group (i.e. a total of 20 cases) would be studied. The ethical and governance issues applicable were dealt with by the collaborating colleagues in the two countries according to the relevant Institutional Review Board (IRB). All data were initially anonymised using an open source (DICOM anonymizer®) software. By reviewing he data from the first source, the faces of the cases had undergone postmortem changes which made them unidentifiable, which would make unreliable comparison between the reconstructed faces and the real faces segmented from the CT scans in their current status. It was, then, attempted to substitute the CT faces with antemortem photographs of the targets from their next of kin, which was not successful. Therefore, data obtained from the first source (Italy) were excluded from the study. Consequently, it was requested from colleagues in the second source (Egypt) to increase the number of cases collected from living patients undergoing head CT scans.

4.1.4 Head CT Scans of the Egyptian Population

A total of 85 head CT scans of Egyptian patients were provided to the research team (Section 3.1.1.2). However, only 61 scans (34 females, and 27 males) (Appendix 18, Table 20) were found suitable according to the inclusion and exclusion criteria mentioned in Section 4.1.1, except that photographs were no longer required. The 61 CT scans were then classified and labelled according to the case number, age and sex (Appendix 18, Table 21). These included the 15 cases used in the pilot study (Section 3.3.2).

4.2 Facial Reconstruction Using the Computed Tomography Scanned Data

In this section, processing of the CT scans, dealing with the technical problems faced, in addition to the selection process of the studied skulls and the facial templates are described.

4.2.1 CT Scans Processing and Technical Problems

The head CT scans were obtained as successive batches between December 2014 and August 2015, and every batch was processed once received as described in Section 3.1.1.2.

4.2.2 Performing the Facial Reconstruction of the Egyptian Skulls

The facial reconstruction process remains the same following the same method described in Section 3.1.3. However, as this study is the first to reconstruct faces of a modern Egyptian population, the process has been done over a number of successive steps depending on trials and errors in order to reach a protocol for this study, and similar studies in the future.

4.2.2.1 Facial Reconstruction of the Egyptian Skulls using Caucasian Facial Soft Tissue Thicknesses and Average European Facial Templates

At this stage, faces of only fifteen Egyptian skulls were studied as a part of the pilot study (Section 3.3). These 15 cases included 8 males aged; Av20y (n=4), Av30y (n=3), and Av50y (n=1), and 7 females aged; Av20y (n=2), Av30y (n=1), Av40y (n=2), and Av50y (n=2). The 15 cases were reconstructed using the proposed method of fitting an average facial template, matching the sex, and age group of the studied case, obtained from the European database of facial templates previously established by Dr. Maria Vanezis (Section 3.1.1.1), and using the Rhine and Moore (1982) and Helmer (1984) facial soft tissue depths for the white Caucasian population at 40 craniofacial landmarks (Appendix 12, Table 17).

The degree of resemblance between the reconstructed faces and their respective targets' real faces segmented from the CT scans was first visually assessed by the researcher. It was observed that the resemblance between the real and reconstructed faces was too weak to move to the assessment stage by volunteer observers. Therefore, further attempts were made to improve the results before starting the assessment stage. The first attempt was to modify the number of the used craniofacial landmarks, particularly at the cheek region, where the discrepancy was mostly noted between the reconstructed and the real faces. For that, 4 different sets of landmarks (Section 0) were used for facial reconstructions of each of the studied cases (i.e. a total of sixty facial reconstructions). This attempt constituted one of the experiments conducted in the pilot study (Section 0). The results of this attempt showed no satisfactory improvement in the resemblance between the facial reconstructions and their real CT faces using the original Rhine and Moore (1982) and Helmer (1984) 40-landmarks set or any of the other three modified sets.

The following attempt was to repeat the facial reconstruction of the 15 cases after replacing the Rhine and Moore (1982) and Helmer (1984) old facial soft tissue data with more recent data from Stephan (2014), but still with using the European facial templates. This also constituted one of the experiments conducted in the pilot study (Section 0).The results showed no difference between the old and the new data in the facial reconstructions.

4.2.2.2 Facial Reconstruction of the Egyptian Skulls using Egyptian Facial Soft Tissue Thicknesses and Average European Facial Templates

Another modification has been attempted by replacing the soft tissue data with Egyptian facial soft tissue thicknesses. Research for published facial soft tissue thicknesses for the Egyptian population revealed one study (El-Mehallawi and Soliman, 2001), where the soft tissue thicknesses at these landmarks were measured using ultrasound, and these landmarks were described in relation to the face only. Therefore, for accurate positioning of the landmarks on the face as well as the cranium, the definitions of the corresponding cranial landmarks were sought from additional studies (Brown et al., 2004, De Greef et al., 2005, Cha, 2013). A new .xml file containing the new set of landmarks' names and the Egyptian facial tissue depths (minimum, mean and maximum) has been prepared and imported into the FR software.

However, as the study by El-Mehallawi and Soliman (2001) presented bilateral landmarks only with no midline landmarks, fitting the facial templates on the skulls resulted in noticeable defects in the reconstructed face. Parts of the skull were bare from the overlying face, and parts of the face, mainly in the middle regions, were overstretched on the skull. Therefore, following Parks et al. (2014) to account for the lack of published midline landmarks by El-Mehallawi and Soliman (2001), a modified set of landmarks and their depths was prepared combining the bilateral landmarks from El-Mehallawi and Soliman (2001) and the midline landmarks from Rhine and Moore (1982) and Helmer (1984).

The directions between the corresponding facial and cranial landmarks were determined using trial and error as well as visual comparison between the reconstructed and the CT real faces via colour maps (histogram) generated by Meshlab software. These colour maps quantitatively represent the differences in depths between two aligned surfaces (i.e. faces). The reconstructions were repeated several times until the best fit possible was reached between the reconstructed faces and the real faces. The faces of the 15 Egyptian skulls were then reconstructed using the new modified set of the Egyptian facial tissue depths, but still with using the European average facial templates. However, visual assessment of the degree of the resemblance of the resulting reconstructed faces by the researcher showed no improvement over the previous attempts.

4.2.2.3 Facial Reconstruction of the Egyptian Skulls using Egyptian Facial Soft Tissue Thicknesses and Egyptian Facial Templates

The next attempt was to replace the European facial templates form the laser scanned database (Section 3.1.1.1) with Egyptian facial templates segmented form the CT scanned database (Section 3.1.1.2), in a similar way as skulls' segmentation, using the 3D Slicer software package. The faces of the 15 Egyptian skulls were then reconstructed using the new modified set of the Egyptian facial tissue depths, and with Egyptian facial templates of similar sex and age group. The reconstructed faces were visually assessed by the researcher and showed satisfactory resemblance to the target so they can be now be assessed objectively and subjectively by the observers. Table 7 shows the attempts followed to reach satisfactory facial reconstruction using the 15 CT scanned skulls as well as facial templates and facial soft tissue depths from different sources.

Table 7: The attempts followed to reach satisfactory facial reconstruction using the 15 CT scanned skulls as well as facial templates and facial soft tissue depths from different sources

ATTEMPT No.	SKULL SOURCE	FACIAL TEPLATES SOURCE	FACIAL SOFT TISSUE DEPTHS SOURCE	SUCCESSFULL
1	CT Scanned (Egyptian)	Laser Scanned (European)	Caucasian, Old ((Rhine and Moore (1982) and Helmer (1984))	N
2	CT Scanned (Egyptian)	Laser Scanned (European)	Caucasian, New Combined Set (Stephan, 2014) + ((Rhine and Moore (1982) and Helmer (1984))	N
3	CT Scanned (Egyptian)	Laser Scanned (European)	Egyptian (El-Mehallawi and Soliman, 2001)	N
4	CT Scanned (Egyptian)	CT Scanned (Egyptian)	Egyptian (El-Mehallawi and Soliman, 2001)	Y

The combined set of bilateral landmarks from El-Mehallawi and Soliman (2001) and the midline landmarks from Rhine and Moore (1982) and Helmer (1984) was included in the user manual of the facial reconstruction (FR) software with a special application on the Egyptian population, with descriptions of the landmarks' anatomical locations and directions, aided by figures and diagrams where appropriate (Appendix 20). This manual was validated as a part of the pilot study (Section 3.5).

4.2.2.4 Generation of the Average Egyptian Facial Templates

As the results of the pilot study showed that average facial templates were superior to single faces in the identification rates and resemblance scores, it was then decided by the researcher to create average faces from the obtained database of the Egyptian head CT scans and use them as facial templates for the facial reconstruction. For this purpose, research has then been conducted to find a suitable 3D image averaging software.

Rapidform 2004 PP2 – RF4 is one of the referenced software used for image analysis and quantitative comparison of the facial morphology in forensic facial reconstruction studies (Wilkinson et al., 2006, Lee et al., 2012, Short et al., 2014). This software was also used to generate average faces for clinical purposes (Kau et al., 2006). At the time of conducting this study, the software has been renamed as Geomagic Software, and commercialised by (3D Systems)TM company. The product of 3D Systems that is capable of combining multiple 3D objects and averaging them into one average 3D mesh is Geomagic Wrap®. The averaging process involves manual point-based aligning (image registration) of a number of single 3D meshes, then the software automatically calculates and averages the distances between the registered meshes and generates one average 3D mesh.

Out of the 61 cases used in this study, 30 scanned skulls (17 males, and 13 females) were selected to be the studied cases (Target Cases), whose faces were reconstructed. The remaining 31 cases were used to generate the average faces (Appendix 13). An average face representing each age group of both sexes has been generated to be used for facial reconstruction of the studied cases of the same group (Appendix 18, Table 22 and Table 23). However, in age groups (41-50 male) and (>50 female groups), there was no sufficient number of faces to generate an average face and reconstruct a face of the same age group. So, to include all age groups in the study, a face template was "borrowed" from a neighbouring age group and merged with a face from the studied age group to generate an average face and use it to reconstruct a face of the studied age group (Appendix 18, Table 22 and Table 23).

4.3 Preparation of the Real and Reconstructed Faces for Assessment

The resulting facial reconstruction (FR) images were exported from the FR software as 3D .obj files images. Similarly, the CT scanned real faces were also segmented and extracted from the 3D Slicer software as 3D .stl files images, which were then converted to 3D .obj files format via Meshlab software. Using the 3D mesh editing tools of Meshlab and Geomagic Wrap software programs, the 3D reconstructed faces and the CT segmented faces were then edited. The images were initially imported into Meshlab software to remove the unrequired parts of the segmented images, including; the non-face and the internal head parts that have been segmented from the CT scans because of having the same threshold range as that of the segmented face skin. Editing via Geomagic Wrap software included smoothing the facial surfaces and remove any spikes in the facial surface as well as cropping the facial templates artefacts caused by the CT scanner with refilling of the resulting holes in consistence with the contour of the face.

Further treatment included cropping and trimming of the part of the head above the forehead and behind the ears (i.e. the part of the head corresponding to the hair), a process referred to as "burning", to allow the observers to focus their assessment on the reconstructed faces only without any distracting features. This is in agreement with other studies (Nelson and Michael, 1998, Stephan and Henneberg, 2006, Wilkinson et al., 2006) as presenting the facial reconstruction without hair and as unidentified as possible is expected to favour positive recognition responses by triggering the memory. Absence of the other distracting, and wrongly estimated facial features, avoids confusion and allows the observer to use their imagination (Wilkinson, 2007). In addition, some authors (Lee et al., 2012) removed certain areas of the head (e.g., back of the head including ears and below of jaw line) as they were not believed to influence the reconstruction errors as the adult tissue depths at these areas shows a constant thickness regardless of age, sex, and ancestry. So, these areas were removed from the compared faces in the present study as well. Finally, to standardise the appearance of the compared faces, the reconstructed and the real faces of each case were aligned, by translation, rotation, and scaling, and edited in a similar manner. Table 8 shows the successive steps of the process of facial reconstruction starting with acquiring the CT scanned heads.

Table 8: The successive steps of the process of facial reconstruction starting with acquiring the CT scanned heads

DATA	PROCESSING	RESULTS	PROCESSING	RESULTS
30 CT scanned Heads (studied cases)	Segmentation (3D Slicer®)	30 faces	Editing (Meshlab®)	30 Real Faces
		30 skulls	Editing (Meshlab®) Facial Reconstruction (FR Software)	30 Reconstructed faces
31 CT scanned Heads	Segmentation (3D Slicer®)	31 faces	Generation of Average Facial Templates (Geomagic Control®) → Editing (Meshlab®) → Facial Reconstruction (FR Software)	
		31 skulls	N/A	N/A

4.4 Subjective Assessment of the Facial Reconstructions

This included subjective face pool and face resemblance tests (Section 3.1.4).

4.4.1 Recruitment of the Assessor Volunteers

Volunteers were recruited for this subjective assessment part of this experiment from inside and outside the Queen Mary University of London (QMUL), as well as from outside the UK. A full experiment for each observer was estimated to last for 6-8 weeks, subject to a timely. However the whole subjective assessment duration lasted from November 2015 to June 2016.

Participants were divided into two groups.

(I) **Non-Expert Group:** invited via emails or via the word of mouth. This group included; Forensic Medical Sciences MSc students at QMUL, Clinical Drug Development MSc students at QMUL, participants of Management of the Dead Course at QMUL, members of the Clinical Pharmacology Department, the Translational Medicine and Therapeutics Department, and the William Harvey Research Institute of QMUL, Forensic Science MSc students at King's College

London, colleagues from Institute of Cardiovascular Science, University College London as well as Egyptian participants living inside and outside the UK.

(II) **Expert Group:** invited via emails or via the word of mouth. They were divided into four main groups according to their experience, including; forensic face recognition psychology, forensic pathology, forensic anthropology and forensic facial reconstruction (Appendix 18, Table 24). Some participants had one or more of these experience types. The duration of experience ranged from 3 months per one type up to 28 years with more than one type of experience.

4.4.2 Presentation of the Subjective Tests to the Assessor Volunteers

In the pilot study (Chapter 3), it was concluded that the more views of the facial images presented to the assessors of the face pool test, the higher the identification rates of the target was found (Section 3.4). Following on from that conclusion, rather than presenting 2D images with multiple views to the assessors, it was decided to present the faces in an online 3D view that allows assessment of the faces from all angles by interactively rotating them. Online tests were in the form of a number of successive online surveys conducted via Google document forms (Google Docs), with restricted access only to participants who were provided with links to the tests, and passwords to maintain confidentiality. As each survey was completed by the participant, the response was automatically recorded and instantly returned to the researcher. Responses were then downloaded in Microsoft Office Excel spreadsheets. These online survey are shown in Appendices 14, 15 & 16. The face models were presented on a special website designed with the aim of online displaying and sharing 3D contents (Sketchfab). A private user account (Sketchfab Pro Account) has been set up on this website with the ability of publishing password-protected 3D models to keep the confidentiality of the studied subjects and providing the participants only with a password for each exercise.

In addition to a more reliable assessment in 3D view, the advantages of this online 3D test format are numerous. To name a few, the flexibility in performing the exercise at the convenient time of the participants, rather than inviting them for performing the tests on site, led to recruiting more off-site, and occasionally, overseas participants. In addition, off-site expert could be recruited easier sparing them from dealing with paper printed tests, with 2D photos, and having to repost the answers back to the research team, with possible material delay or loss, and thus saving time as well as cost. Moreover, with a higher degree of commitment to complete the whole experiment than expected, the number of the recruited participants was expanded more than originally planned. Moreover, the Sketchfab website, where 3D objects could be uploaded, provided a 3D model viewer tool to display the models on any mobile, desktop webpage or Virtual Reality headset. It also involved globalised and interactive viewing of the face models with the ability to move, rotate, and zoom in and out of the models, using an ordinary mouse, for more reliable assessment. In addition, as the objective assessment was done between the 3D real and reconstructed faces, performing the subjective assessment between the same 3D faces allowed a more reliable comparison between the subjective and objective tests. The main drawback, however, of this online testing was the inability to standardise the time spent in each test by all observers. This can be seen as appositive point, though, as setting a specific time for everyone wold not be reflective of the variations in human abilities in real life. The online surveys included both face pool and face resemblance tests, as described below.

4.4.3 Subjective Test (1): Face Pool Tests

The aim of face pool tests was to assess the ability of observers to identify a target subject from a face pool of faces, including the target and similar individuals (Section 3.1.4). In this study, 30 cases were reconstructed (Appendix 18, Table 22 and Table 23). According to the recommendations by Evidence and America (2003), five faces are required as "foils" for each target, with a total of 150 foil faces required to assess the 30 cases reconstructed in this study. However, due to the limited number of the available faces in our database (a total of 61), it was agreed that only 20 cases would be assessed via the face pool testing, while all 30 cases would be assessed via the face resemblance tests (Section 3.1.4). For the same reason, each face pool was designed consisting of one target face (facial reconstruction) and four test faces (CT scanned similar faces including the real face of the same case), instead of five foil faces as previously recommended. These test faces were selected to be matching the studies case's sex and age group and general face shape. However, in some situations where there was a limited availability, foil faces similar in general face morphology to the target face were selected from the same and the closest age group(s). Furthermore, the generated average face of a certain age group was used as a foil face for that group provided that it was not used for reconstructing the face in the same test (Appendix 18, Table 22, Table 23, and Table 25). To standardise the appearance of the compared faces, the test (foil) faces, forming each face pool, were respectively aligned with their target face by translation, rotation, and scaling, then edited in the same way as described in (Section 4.3).

Subjective assessment by face pool tests included 20 cases, which were performed over four successive exercises in four online surveys; each containing 5 different cases. Each exercise consisted of both male and female cases of different age groups (Appendix 18, Table 25). The cases were presented to observers in the order presented in Appendix 18, Table 25. Each face pool exercise consisted of a playlist of 3D models; One Target Face, representing the reconstructed face, labelled (Target Face), Four Test Faces, representing the foil faces, labelled (Test Faces A-D), and One collective model of all five faces together, labelled (Face Pool). The collective model is formed of the target face "the facial reconstruction" in the middle, surrounded by test face (A) in the top left corner, test face (B) in the top right corner, test face (C) in the bottom left corner, and test face (D) in the bottom right corner (Appendix 15-A). As Stephan and Henneberg (2006) found out that assessors tended to select a face from the face pool despite having the option not to choose any face, which appeared biased in this respect, the target face was always present in the face pool tests designed in the present study.

Prior to performing the tests, participants were instructed to make their assessments primarily based on the general shape of the face hard structures (e.g. forehead, orbits "eye bones", cheek bones, temple bone "the bone on the sides of the head", chin, etc...), rather than the face soft structures (e.g. nose, lips, etc...). Participants were asked to examine the faces carefully within their individual and collective models, then pick only one face from the 4 test faces that they thought it closely resembled the target face, then to insert their answers in the appropriate section in the assessment survey for each case. Each participant performed the tests separately and there was no time limit enforced. They were also asked not to share their answers with other participants before the end of the experiment. The test instructions are shown in Appendix (15-B).

Each observer was expected to perform 1-2 exercises per week. However, each observer was sent one exercise at a time. This face pool stage was estimated to last for 2-4 weeks for each observer, subject to a timely response.

The total number of participants in the face pool tests was 76 with 65-76 participants/case. Twenty-six (26) experts started this stage by performing the first face pool exercise. This group included 19 female (73%) and 7 male (27%) participants, 7 Egyptians (27%) and 19 non-Egyptians (73%), and ages ranging from 26 – 68 years old (mean age = 37.3 years old). The mean age of all participants was 34.2 years old. All experts performed all the 20 face pool tests, except one expert (Exp.19) who performed 5 tests only (Appendix 18, Table 26). In contrast, fifty (50) non-expert participants started this stage by performing the first face pool exercise. This group included 38 female (75%) and 12 male (25%) participants, 10 Egyptians (20%) and 40 non-Egyptians (80%), and ages ranging from 20 – 65 years old (mean age = 32.6 years old). However, not all the 50 non-expert participants completed the face pool stage (Exercises 1, 2, 3 and 4) (Appendix 18, Table 26).

The non-expert participants who completed the 20 face pool tests were asked to repeat the tests without knowing the correct answer. This was an additional and optional experiment and participants had the option to repeat as many tests as possible but in the order they have done the tests initially. Out of all participants, 3 participants repeated all the 4 face pool exercises (i.e. 20 cases), 1 participant repeated the first 2 face pool exercises (i.e. 10 cases), and 3 participants repeated the 1 face pool exercise (i.e. 5 cases) (Appendix 18, Table 25). Therefore, a total of 85 answers were obtained in the repeated tests. This experiment was done with the aim of testing if there was a learning curve in facial identification.

The results obtained in all face pool tests, and in the repeated tests, were statistically analysed, using Microsoft Excel 2010 and Minitab 17 statistical package, as follows:

- **The percentage of correct identification** of the real face from the face pool in a rate above chance. This was calculated based on the design of the face pool test in this study, which consisted of 4 test faces (including the real face). So, the chance of selecting any face was 1 in 4 (i.e. 25%). Identification rate of any face above 25% was considered significant.
- **Binary logistic regression statistical test** was carried out to estimate the probability, as a binary response, of the presence of a relationship between the correct identification rate of a participant and one or more independent characteristics of the participant. The binary logistic model is used to estimate the probability of a binary response based on one or more predictor (or independent) variables (features). It determines whether the presence of a risk factor increases the probability of a given outcome by a specific percentage which means a significant association between the variable and the outcome. In other words, it measures the probability/chance of whether a participant having a special characteristic would lead to higher identification rates of the real face out of a face pool. This also takes into account any overlap between participants' characteristics (Szumilas, 2010).

The participants' characteristics investigated were:

  - The participant's age,
  - The participant's sex,
  - The participant's nationality (Egyptian or not), and
  - The participant's professional experience in:
    - Forensic anthropology, with or without experience in forensic facial reconstruction,
    - Forensic face identification psychology, or
    - Forensic pathology.

- **The studied cases were ranked** according to the combined identification rates as well as the identification rate by each group of participants.
- **In the repeated face pool tests,** the proportions of correctly identified cases by the same participants in the second attempt were calculated and compared with those in the first attempt.

When a logistic regression is calculated, the regression coefficient is the estimated increase in the odds that the outcome increases in the presence of the variable. If an association is significant, an odds ratio (OR) quantifies how strongly that variable is associated with the outcome. The OR ratio is presented as follows: OR=1 (i.e. variable does not affect odds of outcome), OR>1 (i.e. variable associated with higher odds of outcome), and OR<1 (i.e. variable associated with lower odds of outcome). Therefore, the odds ratio is interpreted by first being deducted from a whole 1. The 95% confidence interval (CI) is used to estimate the precision of the OR. A large CI indicates a low level of precision of the OR, whereas a small CI indicates a higher precision of the OR (Szumilas, 2010).

4.4.4 Subjective Tests (2) and (3): Face Resemblance Tests

This face resemblance stage consisted of 30 direct "face-to-face" visual comparisons studying the 30 cases to assess the similarity between the facial reconstruction and the real face of a target using rating/ranking scale. For the present study, a resemblance test form (A) (Section 3.1.4) was designed and faces were assessed online in 3D on Sketchfab website. All the 30 cases reconstructed in the main part of this thesis were assessed via the face resemblance tests. If a participant had performed the face pool tests initially, the face resemblance tests started for each observer at least 2 weeks after they had finished the face pool tests. The 30 cases were assessed over 2 exercises (15 case/exercise), and the cases were presented to observers in the order presented in Appendix 18, Table 27. Exercises were in the form of online surveys as discussed in Section 4.4.2. For most participants, each exercise was performed in one session. However, it was convenient to some participants, particularly overseas, to send them the links to the 15 cases tested in each exercise in a word document to be spread over more than one session. After completing the whole exercise, those participants emailed the answers to the researcher. The second exercise was sent individually to each observer 2 weeks after the completion of the first exercise. This face resemblance stage was estimated to last for 2-4 weeks for each observer, subject to a timely response.

For each case, the face resemblance test consisted of one 3D model composed of 1 top face, representing the real face, and 1 bottom face, representing the reconstructed face

(Appendix 16-A). Each observer was asked to rate the target's reconstructed face, according to their similarity to the target's real face, using a rating scale from (0 to 10), where (0) represented no resemblance and (10) represented the highest resemblance to the target, then to insert their answers in the appropriate section of each case in the assessment surveys. Prior to performing the tests, participants were instructed to make their assessments primarily based on the general shape of the face hard structures (e.g. forehead, orbits "eye bones", cheek bones, temple bone "the bone on the sides of the head", chin, etc...), rather than the face soft structures (e.g. nose, lips, etc...). Participants were asked to examine the faces carefully within their individual and collective models, then pick only one face from the 4 test faces that they thought it closely resembled the target face, then to insert their answers in the appropriate section in the assessment survey for each case. Each participant performed the tests separately and there was no time limit enforced. They were also asked not to share their answers with other participants before the end of the experiment.

The results obtained in all tests were statistically analysed, using Microsoft Excel 2010 and Minitab 17 statistical package. Two versions of face resemblance tests were tested as follows:

- **Subjective Test (2): Face Resemblance Test Version (1); Overall Resemblance Scores**

This test was done by asking the participants to give an overall resemblance score to each reconstructed face using a numerical (0-10) scale. The test instructions are shown in appendix (16-B). Results were analysed as follows:

- The overall resemblance scores given by all participants (combined).
- The overall resemblance scores given by each participant group:
  - Experts in forensic anthropology with and without facial reconstruction experience. This was because results of the face pool test showed significant association between high identification rates and these professional characteristics only (Section 5.1).
  - Egyptian Participants.

  - Non-Expert, Non-Egyptian, Old Participants (NEX-NEG-OLD): who performed the face pool tests prior to the face resemblance tests.
  - Non-Expert, Non-Egyptian, New Participants (NEX-NEG-NEW): who did not perform the face pool tests prior to the face resemblance tests.
- The studied cases were ranked according to the combined resemblance scores as well as the resemblance scores by each group of participants.
- The between-observer agreement in the given overall resemblance scores between all participants (combined) using Kendall's Coefficient of Concordance (KCC) (a measure of the agreement among several observers that are rating/assessing a set of objects of interest).
- The between-observer agreement in the given the overall resemblance scores within the participant groups mentioned above using Kendall's Coefficient of Concordance (KCC).
- The statistical difference in the ranks of the assessed cases according to the overall resemblance scores given by participant groups using one way ANOVA and Spearman's rank correlation coefficient. Comparisons were particularly conducted between:
  - The NEX-NEG-OLD, and the NEX-NEG-NEW groups.
  - The NEX-NEG-OLD, and the non-experts, Egyptian groups.
  - The NEX-NEG-OLD, and the forensic anthropology experts.

- **Subjective Test (3): Face Resemblance Test Version (2); Individual Facial Regions Resemblance Scores**

This test was done by asking the participants to give a separate resemblance score to each facial region (1-Forehead, 2- Orbits, 3- Cheek Bone, 4- Chin, 5- Jaw), for each case using the numerical (0-10) scale for each facial part. The test instructions are shown in Appendix (16-C). Results were analysed as follows:

- The resemblance scores given by all participants (combined) to each facial region.
- The sum of the resemblance score for every case calculated from the resemblance scores given by all participants (combined) to each facial region.

- Ranking of the cases according to the resemblance scores given by all participants (combined) to each facial region.
- Ranking the cases according to the sum of the resemblance score for every case calculated from the resemblance scores given by all participants (combined) to each individual facial region.
- Using Spearman's rank correlation coefficient to correlate the case ranks according to the sum of the resemblance scores calculated from individual regions scores (subjective test 3) with the case ranks according to the objective overall surface distance standard deviation (SD) (objective test 1) (Section 4.5).
- Using Spearman's rank correlation coefficient to correlate the case ranks according to the sum of the resemblance scores calculated from individual regions scores (subjective test 3) with the case ranks according to the overall resemblance scores given in face resemblance test version (1) (subjective test 2) (Section 4.5).
- Using Spearman's rank correlation coefficient to correlate the case ranks according to the resemblance scores of individual facial regions in the face resemblance test version (2) (subjective test 3) with the case ranks according to the identification percentages of face pool tests (subjective test 1) (Section 4.5).
- Using Spearman's rank correlation coefficient to correlate the case ranks according to the resemblance scores of individual facial regions in the face resemblance test version (2) (subjective test 3) with the case ranks according to the objective overall surface distance standard deviation (SD) (objective test 1) (Section 4.5).
- Using Spearman's rank correlation coefficient to correlate the case ranks according to the resemblance scores of individual facial regions in the face resemblance test version (2) (subjective test 3) with the case ranks according to the objective surface distance standard deviation (SD) of the corresponding individual facial regions (objective test 2) (Section 5.4).

Some participants performed the face pool tests before starting this stage (Old participants) while others did not (New participants). Participants were divided into 2 groups according to which test version performed first (Appendix 18, Table 28):

- **Group (1)** performed exercise one (cases 01-15) in version (1), then exercise two (cases 16-30) in version (2).
- **Group (2)** performed exercise one (cases 01-15) in version (2), then exercise two (cases 16-30) in version (1).

The number of participants in this stage ranged from (65-76) per case. Fifty-one non-experts participants started this stage by performing the first face resemblance exercises versions 1 and 2 (Appendix 18, Table 28). This group included 34 female (67%) and 17 male (33%) participants, 8 Egyptians (16%) and 43 non-Egyptians (84%), and with ages ranging from 20 – 65 years old (mean age = 33.8 years old). However, not all the 51 participants completed the full experiment (Exercises 1 and 2) (Appendix 18, Table 27). Moreover, out of the 26 experts who performed the face pool tests (Appendix 18, Table 24), only 22 experts (all except Exp. 04, 07, 14 and 19) performed all the 30 face resemblance tests (Appendix 18, Table 28). The experts group included 15 female (68%) and 7 male (32%) participants, 6 Egyptians (27%) and 16 non-Egyptians (73%), and with ages ranging from 27 – 68 years old (mean age = 38.6 years old). The mean age of all participants was 35.2 years old.

4.5 OBJECTIVE ASSESSMENT OF THE FACIAL RECONSTRUCTIONS

This stage involved 3 types f objestive tests as described below. The results obtained these assessment tests were statistically analysed using Microsoft Excel 2010 and Minitab 17 statistical package.

4.5.1 Objective Test (1): Facial Surface Overall Distance Standard Deviation (SD)

This test was performed via 3D morphometric surface comparison using (Geomagic Control®) Software. This software was previously used, under the name of Geomagic Qualify©, by Lee et al. (2012) to assess the accuracy of faces reconstructed using the Manchester Method. In (Geomagic Control®) Software, surface comparison started by manually aligning the compared faces (the reconstructed and real face segmented from the CT scan) guided by the embedded skull and as well as a number of corresponding landmarks. Quantitative assessment of the surface morphology differences between the compared faces was then automatically calculated by the software and presented as average surface distance, root mean square (RMS) of surface deviations, Standard Deviation (SD) of the errors between the overlapped surfaces, and maximum and minimum ranges. The differences could also be shown via colour maps (histograms) (Appendix 17).

It was noticed that all these forms of surface distance differences changed with adjusting the maximum deviation settings of the software for the surface comparison. However, the Standard Deviation (SD) of the surface deviations showed the least changes. In addition, the RMS errors correlate well with the standard deviations (De Greef et al., 2005). Therefore, the latter was used as the indicator of the difference/fit between the compared faces, and the assessed cases were ranked according to their objective testing by that overall distance Standard Deviation (SD).

4.5.2 Objective Test (2): Facial Surface Distance at Individual Facial Regions

Using the same software program (Geomagic Control®), points were manually located on the combined facial model formed of registered real and reconstructed faces of each case. The distances between each 2 faces at each point were measured automatically via the software. Different facial regions were determined by different groups of points with a total of 33 points (Appendix 17) (some points are represented in more than one anatomical region) as follows:

- Forehead: 13 points.
- Orbit (right side): 4 points.
- Orbit (left side): 3 points.
- Nasal region: 3 points.
- Cheek (right side): 3 points.
- Cheek (left side): 3 points.
- Chin region: 3 points.
- Mouth: 5 points.
- Jaw (right side): 4 points.
- Jaw (left side): 4 points.

The absolute distances, including the average of bilateral regions points, at the points of the 6 regions (forehead, orbit, nasal region, cheek, chin and Jaw) were averaged and used as an objective measure of the surface distance at the respective regions. The average distances at each facial were then calculated.

The cases were separately ranked according to the objective absolute surface distance at each region and according to the sum of the absolute differences at all facial regions. Using Spearman's rank correlation coefficient, the following was correlated:

- The case ranks according to the sum of the absolute differences at all facial regions (objective tests 2) was correlated with the case ranks according to the overall surface distance SD (based on 30 cases) (objective tests 1).

- The case ranks according to the sum of the absolute differences at all facial regions (objective tests 2) was correlated with the case ranks according to the identification rate in face pool tests (based on 20 cases) (subjective tests 1).

- The case ranks according to the sum of the absolute differences at all facial regions (objective tests 2) was correlated with the case ranks according to the subjective overall resemblance score given to each case in the face resemblance test version 1 (subjective tests 2), based on 30 cases.

- The case ranks according to the objective absolute surface distance at each region was correlated with the case ranks according to the subjective resemblance score given to each region in the face resemblance test version 2 (subjective tests 3), based on 30 cases.

4.5.3 Objective Test (3): Craniofacial Anthropometry

Craniofacial Anthropometry, was applied as a method for objective assessment of facial reconstructions, by measuring a number of linear measurements (n=9) taken from the skull, real and reconstructed faces (Figure 20 and Table 9), from which linear ratios (n=13) and angles (n=11) were also calculated (Table 10). The set of linear measurements, ratios and angles used in this study was adopted, and developed, from (Kleinberg et al., 2007), where it was used for matching suspects' 2D faces to 2D faces from surveillance camera footages. In contrast, in the present study the facial comparison was used for the purpose of assessment of 3D forensic facial reconstructions.

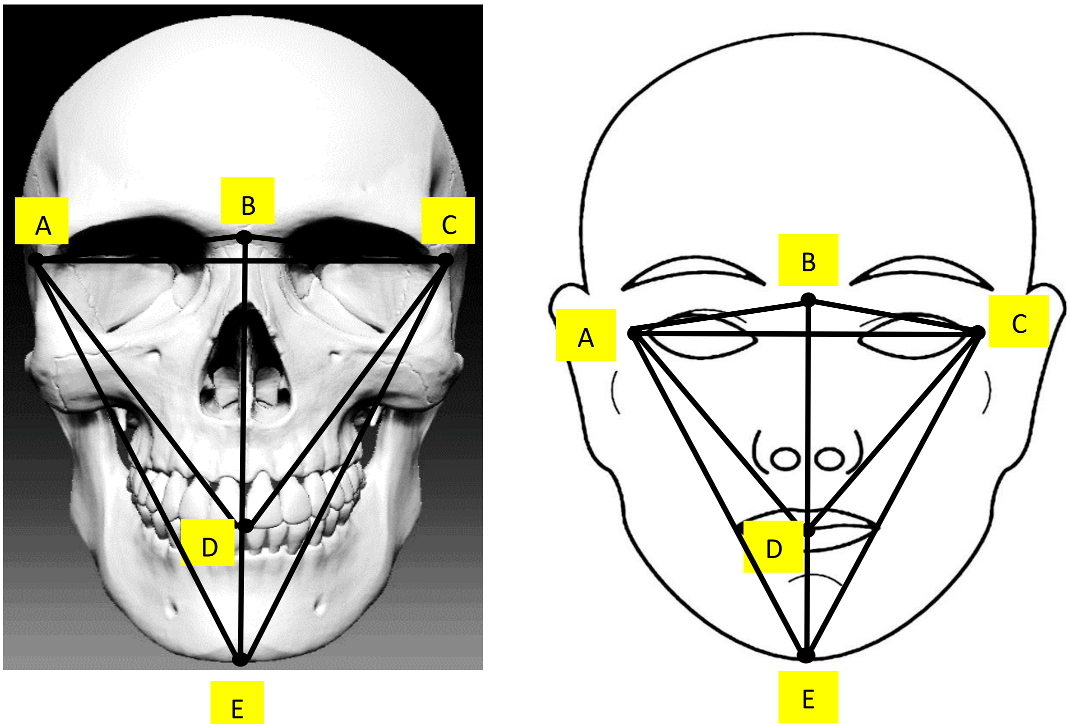


Figure 20: The used linear measurements, linear ratios, and angles

Table 9: The definitions of the anatomical points used for linear measurements

Anatomical Point	SKULL DEFINITION	FACE DEFINITION
(A) Left Ectocanthion	Bony projection of the ectocranial surface of the frontal bone, vertically centred on the orbit, next to the lateral orbital border	A point lateral to the outer canthus (angle) of the eye, vertically centred on the orbit, next to the lateral orbital border.
(B) Nasion	The midline of the Naso-frontal suture. A point at the top of the nasal bone, at the horizontal level of a line dividing the orbit into upper and lower halves.	
(C) Right Ectocanthion	Bony projection of the ectocranial surface of the frontal bone, vertically centred on the orbit, next to the lateral orbital border	A point lateral to the outer canthus (angle) of the eye, vertically centred on the orbit, next to the lateral orbital border.
(D)	A point in the midline halfway between the Supradentale (The jaw Centre, between the upper incisive teeth) and the Infradentale (The jaw Centre, between the lower incisive teeth).	A point in the midline halfway between the Labiale Superius (The midline point of the upper lip) and the Labiale Inferius (The midline point of the lower lip).
(E) Gnathion	Lowest point of the front of the chin in the midline	

Table 10: The used linear measurements, linear ratios, and angles

MEASUREMENT	RATIOS	ANGLE °
AC	AB/AD	AEC
AB	AB/BD	CAE
BC	BC/CD	CAD
AD	BC/BD	ACE
CD	AD/BD	ACD
BD	CD/BD	ABE
AE	AB/AE	CBE
BE	AB/BE	CDB
CE	BC/CE	ADB
	BC/BE	CEB
	AE/BE	AEB
	CE/BE	
	AC/BE	

The 3 parameters (linear measurements, linear ratios, and angles) were used for the objective comparison between the skull, the real and the reconstructed faces of each case. Linear ratios were calculated using Microsoft Excel 2010. Angles were calculated for each triangle separately using an online triangle calculator (Triangle Calculator) based on the mathematical rule "the angle of a triangle can be calculated from its sides". From the linear measurements, linear ratios, and angles, the following was calculated:

- The differences between the linear ratios of the skull and the real face were compared with the differences between the linear ratios of the skull and the reconstructed face.
- The differences between the angles of the skull and the real face were compared with the differences between the angles of the skull and the reconstructed face.
- The absolute differences between the linear ratios of the real and the reconstructed faces were calculated, and then averaged.
- The absolute differences between the angles of the real and the reconstructed faces were calculated, and then averaged.

To validate this method, the results were further analysed by ranking the assessed cases according to their:

- Objective testing by the individual linear ratios differences.
- Objective testing by the averaged linear ratios differences.
- Objective testing by the individual angles differences.
- Objective testing by the averaged angle differences.

4.6 COMBINED RESULTS ANALYSIS: CORRELATED TESTS

Table 11 shows all the subjective and objective tests, used for assessment of the reconstructed faces in the main part of the thesis, and their descriptions.

Table 11: The subjective and objective tests used for assessment of the reconstructed faces and their description

TEST TYPE	TEST TITLE	TEST NAME	TEST DESCRIPTION
Subjective Assessment	**Subjective Test (1)**	Face Pool Test (n=20)	Identification Percentage/Rate Above Chance
	Subjective Test (2)	Face Resemblance Test Version (1) (n=30)	Overall Facial Resemblance Score (0-10)
	Subjective Test (3)	Face Resemblance Test Version (2) (n=30)	Individual Facial Regions Resemblance Score (0-10)
Objective Assessment	**Objective Test (1)**	Surface Distance Difference Test Version (1) (n=30)	Overall Facial Surface Distance Standard Deviation (SD) Differences
	Objective Test (2)	Surface Distance Difference Test Version (2) (n=30)	Individual Facial Regions Surface Distance Standard Deviation (SD) Differences
	Objective Test (3)	Craniofacial Anthropometry (n=30)	Linear Ratios and Angles Differences

The ranks of all cases by all subjective and objective assessment methods were statistically correlated using Spearman's rank correlation coefficient. The ranks of the 20 cases (via face pool tests) and the 30 cases (via face resemblance tests and the objective tests) were verified using the non-parametric Freidman’s test to measure the difference between cases in the ranks received using different tests. Results of analysis are described in the Chapter 5 (Results).

CHAPTER 5: THE MAIN STUDY RESULTS

The faces reconstructed in this study were assessed both subjectively, by face pool and face resemblance tests, and objectively, by the overall surface distance comparison between the overall faces and the individual facial regions, as well as craniofacial anthropometry.

5.1 SUBJECTIVE TEST (1): FACE POOL TESTS

Twenty cases only were assessed via the subjective face pool tests (Section 4.4.3).

(A) The correct identification rates in face pool tests

Appendix 18, Table 29 shows the percentage of selection of each face in the face pool for each case, whether the target face was correctly identified above chance (25%), and the non-target faces that were identified above chance. The 20 face pools contained 20 target faces and 60 non-target faces. Out of the target faces, 13/20 faces (65%) were identified above chance. In 4/20 cases (25%), no non-target faces were identified above chance. Of them 4 cases Out of the 60 non-target faces, 20/60 faces (33%) were identified above chance, and further 12/60 faces (20%) were identified above that of their targets.

Appendix 18, Table 30 and Table 31 show the correct identification rates by all participants' groups:

- The mean identification rate of all participants combined (n=65-76) was 38.27% (13% above random chance), with 13/20 (65%) cases correctly identified.
- The mean identification rate of all the non-expert participants (n=40-50) was 38% (13% above random chance), with 13/20 cases correctly identified.
- The mean identification rate of the non-Egyptian non-expert participants (n=32-40) was 38% (13% above random chance), with 13/20 (65%) cases correctly identified.
- The mean identification rate of the Egyptian non-expert participants (n=8-10) was 36.49% (11% above random chance), with 11/20 (55%) cases correctly identified.

- The mean identification rate of all expert participants (n=25-26) was 36.37 (11% above random chance), with 11/20 (55%) cases correctly identified.
- The mean identification rate of the forensic pathology experts (n=13) was 36.56% (12% above random chance), with 12/20 (60%) cases correctly identified.
- The mean identification rate of the forensic anthropology experts (n=11-12) was 45.54 (21% above random chance), with 16/20 (80%) cases correctly identified.
- The mean identification rate of the facial identification psychology experts (n=3) was 40% (15% above random chance), with 16/20 (80%) cases correctly identified.

(B) Binary logistic regression statistical analysis of the odds ratio (OR) for participants' characteristics results:

- *Participants' age:*

Participants in this face pool tests were a total of 76 participants, with ages ranging from 20 – 68 years old (mean age = 34.2 years old). The binary logistic regression showed a *significant* association between the high identification rates and the participants' age. The calculated odds ratio (OR) for *participants' age* was 0.987 (95% CI: 0.976 - 0.998). This means that young participants have a 1-2 % higher chance of *significantly* higher correct identification rates than older participants (P = 0.025) (Figure 21).

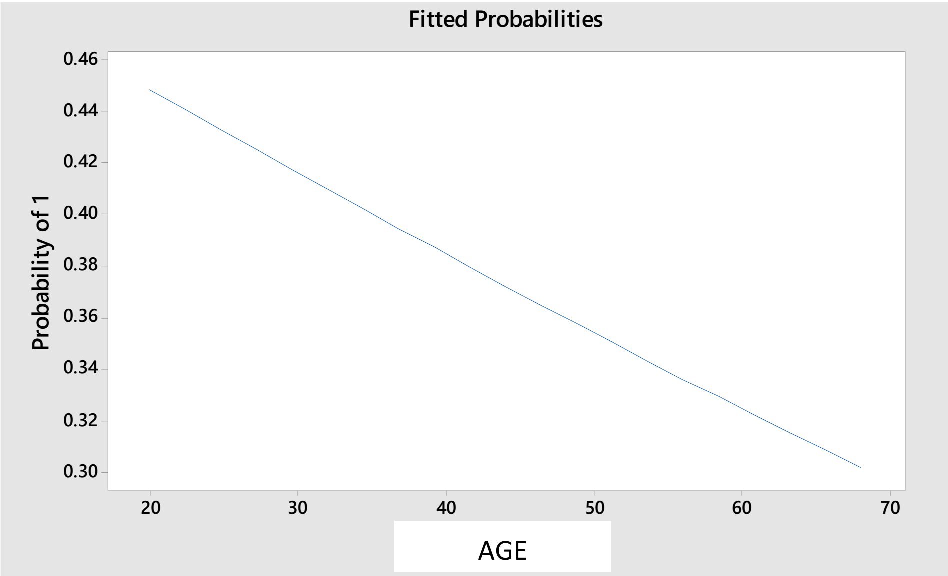


Figure 21: Plotted Odds Ratio (OR) of the correct identification rate of the participants' age

- *Participant's sex*

Participants in this face pool tests were a total of 76 participants, with a 57 females (75%) and 19 males (25%). The binary logistic regression showed a *non-significant* association between the high identification rates and the sex of the participant.

- *Egyptian Participant*

Participants in this face pool tests were a total of 76 participants, with 17 Egyptians (22%) and 59 non-Egyptians (78%). The binary logistic regression showed a *non-significant* association between the high identification rates and Egyptian participants compared to non-Egyptian participants.

- *Participant with professional experience:*

Participants in this face pool tests were a total of 26 participants. The types of experiences are summarised in Appendix 18, Table 24. The binary logistic regression showed a *non-significant* association between the high identification rates and participants with a professional experience in forensic medicine/pathology nor facial identification/perception psychology. However, the latter group includes only 3 participants

On the other hand, the binary logistic regression showed a *significant* association between the high identification rates and participants with a professional experience in forensic anthropology only. The calculated odds ratio (OR) was 1.594 (95% CI: 1.16 - 2.19). This means that participants with experience in forensic anthropology had 16 – 119 % (an average of 60%) higher chance of *significantly* higher correct identification than those with no such experience (P = 0.004) (Figure 22).

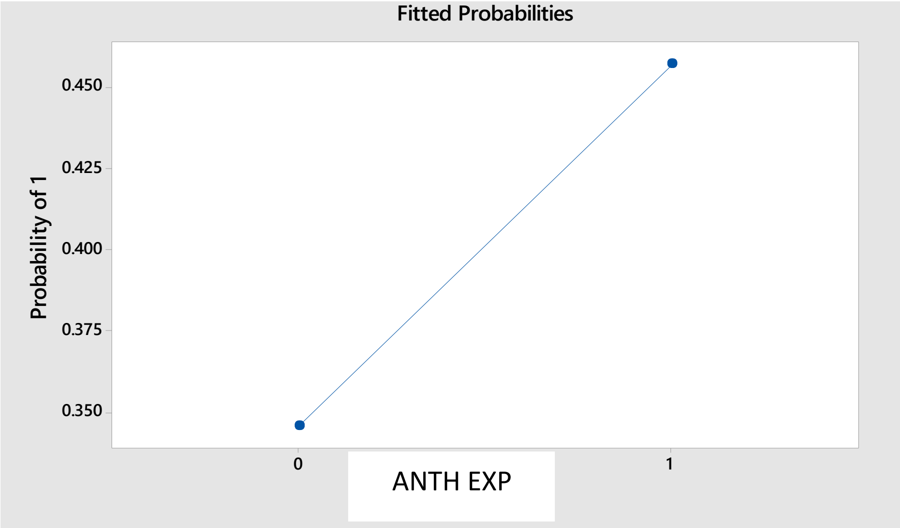


Figure 22: Plotted Odds Ratio (OR) of the correct identification rate of the participants with professional experience in forensic anthropology

Furthermore, it showed a *significant* association between the high identification rates and participants with a professional experience in forensic anthropology with forensic facial reconstruction. The calculated odds ratio (OR) was 2.996 (95% CI: 1.622 - 5.535). This means that participants with experience in forensic anthropology have 62 – 454 % (an average of 200%) higher chance of significantly higher correct identification than those with no such experience (P = 0.002) (Figure 23).

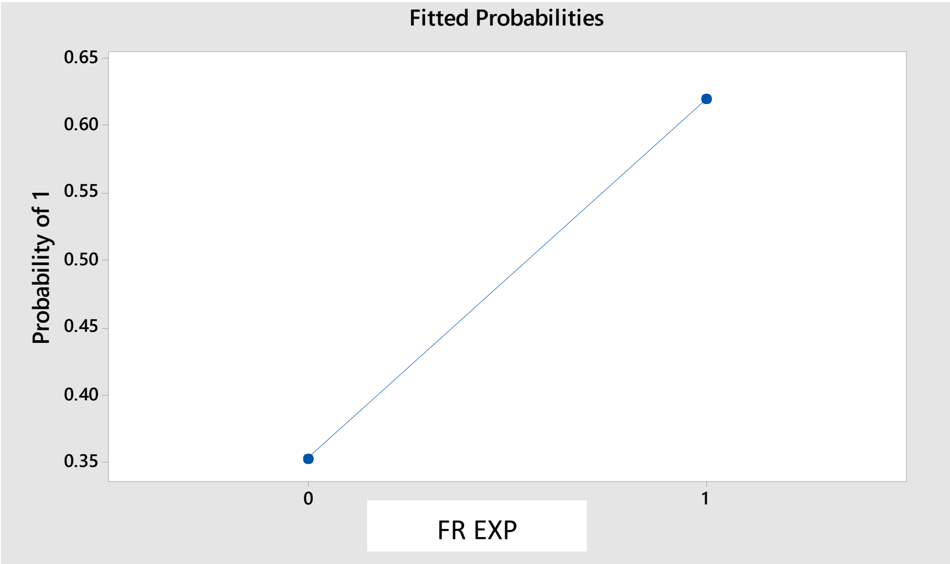


Figure 23: Plotted Odds Ratio (OR) of the correct identification rate of participants with professional experience in forensic anthropology and forensic facial reconstruction

(C) Ranking of the 20 cases according to the combined identification rates as well as the identification rate by each group of participants:

Results are shown in Appendix 18, Table 32 and Table 33. The higher the identification rate of a given facial reconstruction, the lower it was ranked. Results of statistical correlation of the case ranks according to each participants' groups' showed:

- A *significant* correlation between all experts and all non-experts (Spearman's rank correlation coefficient = 0.833) (P = 0.000).
- A *significant* correlation between non-expert Egyptians and non-expert non-Egyptians (Spearman's rank correlation coefficient = 0.754) (P = 0.000).
- A *significant* correlation between all non-experts and forensic anthropology experts (Spearman's rank correlation coefficient = 0.746) (P = 0.000). This is in spite of the *significant* difference in the identification rates between the 2 groups (P = 0.004).
- A *significant* correlation between all non-experts and the face identification psychology experts (Spearman's rank correlation coefficient = 0.729) (P = 0.000).
- A *significant* correlation between all non-experts and the forensic pathology experts (Spearman's rank correlation coefficient = 0.795) (P = 0.000).

The cases' ranks using this method were then statistically correlated with the case ranks using other subjective and objective methods (Section 5.6).

(D) Subjective Assessment by a Repeated Face Pool Test

Out of all the participants, 3 participants repeated all the 4 face pool exercises (i.e. 20 cases) (Appendix 18, Table 25), 1 participant repeated the first 2 face pool exercises (i.e. 10 cases), and 3 participants repeated the first 1 face pool exercise (i.e. 5 cases). Therefore, a total of 85 answers were obtained by 7 participants only on their second attempt of the face pool tests. The proportions of the correctly identified cases by the same participants in the first and second attempts were calculated and compared. The proportions of the correct answers in the 2^{nd} attempt were higher (Table 12). However, the statistical difference between the proportions of the correct answers in the 2 attempts was marginally *insignificant* (P = 0.086).

Table 12: Comparison between the proportions of correct identification in 1st and 2nd attempts of face pool tests

Attempt	Total Responses	Correct Responses	Proportion of Correct Identification
1st	85	31	36
2nd	85	42	49

5.2 SUBJECTIVE TESTS (2) AND (3): FACE RESEMBLANCE TESTS

All the 30 cases were assessed via the subjective face resemblance tests (i.e. scores from 0-10 representing the resemblance between the real and the reconstructed faces given by volunteering participants). As described in the previous chapter, the face resemblance tests were designed in 2 versions and the observers performing these tests were further distinguished into 2 groups (Section 4.4.4). Following the results of the face pool tests (Section0), only forensic anthropology, with or without facial reconstruction, experience showed a significant association between the high identification rates in face pool tests and these professional characteristics. Therefore, only the forensic anthropology experience was considered in this stage, while participants with other professional experience were merged with the non-expert group. Results of the face resemblance test versions (1) and (2) are described below.

(I) Subjective Tests (2): Face Resemblance Test Version (1); Overall Resemblance Scores

The first exercise (cases 01-15) was assessed by 37 participants (Group 1), while the second exercise (cases 16-30) was assessed by 33 participants (Group 2) (Appendix 18, Table 34). Results were analysed based on:

(A) The overall scores in face resemblance tests:

The average overall resemblance scores given by all participants (combined), and each of the participants groups to all the facial reconstructions are shown in Appendix (18), Table 36. The mean scores given by all participants was 45%, and by different groups were 49% for the forensic anthropology experts, 38% for the Non-Expert, Non-Egyptian, Old

(NEX-NEG-OLD) group (i.e. participants who performed the face pool tests prior to the face resemblance tests), 31% for the Non-Expert, Non-Egyptian, New (NEX-NEG-NEW) group (i.e. participants who did not perform the face pool tests prior to the face resemblance tests), and 62% for the non-expert y-Egyptian group.

(B) Ranking the assessed cases according to the combined resemblance scores as well as the resemblance scores by each participant group results:

The ranked cases according to the average overall resemblance scores given by all participants (combined), and each of the participants groups are shown in Appendix 18, Table 37. The higher the resemblance score of a given facial reconstruction, the lower it was ranked.

(C) Between-observer agreement in the ranking of the 30 cases, according to the overall resemblance scores, within all participants (combined) and within each participant group results:

This was performed using Kendall's Coefficient of Concordance (KCC) and the between-observers agreement was tested based on the ranks of the 30 cases by all participants (combined) as well as by each participants' group (Appendix 18, Table 37).

For exercise One (cases 01-15): Results showed:

- An overall *significant* agreement within all participants (combined) (n=37) (KCC = 0.23) (P = 0.000).
- A *significant* agreement within participants with a professional experience in forensic anthropology (n=5) (KCC = 0.48) (P = 0.003).
- A *significant* agreement within the non-expert non-Egyptian participants who performed the face pool tests before the face resemblance tests (NEX-NEG-OLD) (n=17) (KCC = 0.28) (P = 0.000).
- A *significant* agreement within the non-expert non-Egyptian participants who did not perform the face pool tests before the face resemblance tests (NEX-NEG-NEW) (n=9) (KCC = 0.24) (P = 0.008).
- A *disagreement* within the non-expert Egyptian participants (n=6) (P = 0.28).

For exercise Two (cases 16-30): Results showed:

- An overall *significant* agreement within all participants (combined) (n=33) (KCC = 0.33) (P = 0.000).
- A *significant* agreement within participants with a professional experience in forensic anthropology (n=5) (KCC = 0.48) (P = 0.002).
- A *significant* agreement within the (NEX-NEG-OLD) group (n=16) (KCC = 0.33) (P = 0.000).
- A *disagreement* within the (NEX-NEG-NEW) group (n=5) (P = 0.05).
- A *significant* agreement within the non-expert Egyptian participants (n=7) (KCC = 0.45) (P = 0.000).

(D) The statistical difference in the ranks of the assessed cases according to the overall resemblance scores given by each participant group was compared as follows:

1- Comparison (1): the NEX-NEG-OLD, and the NEX-NEG-NEW groups:

Appendix 18, Table 38 shows the average resemblance scores and the cases' ranks by both groups. One-way ANOVA showed that the average resemblance scores given by the NEX-NEG-OLD group were *significantly higher* (Figure 24) than those given by the NEX-NEG-OLD group (P = 0.027).

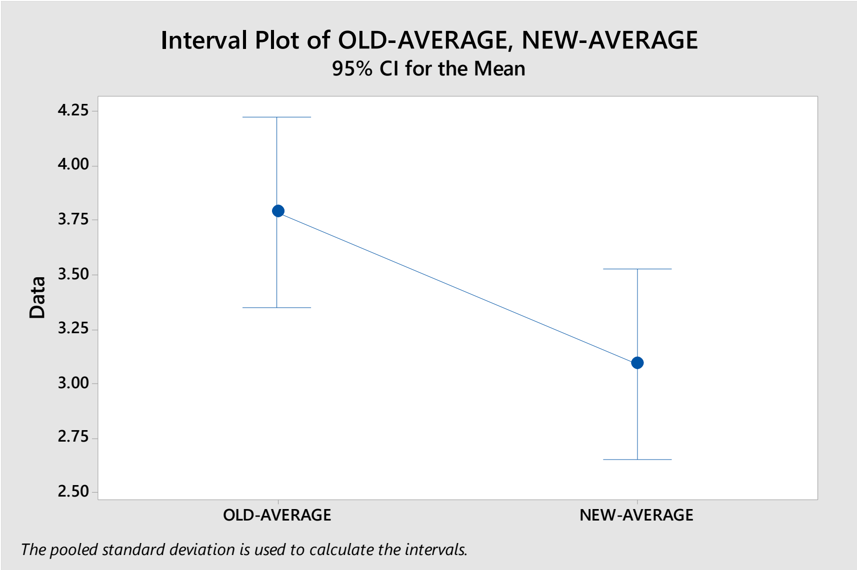


Figure 24: The mean average resemblance scores given by the NEX-NEG-OLD, and the NEX-NEG-NEW groups

However, the cases' ranks according to the average resemblance scores given by the NEX-NEG-OLD, and the NEX-NEG-NEW groups *correlated significantly* (Spearman's rank correlation coefficient 0.817) (P = 0.000).

2- Comparison (2): Non-expert non-Egyptian participants (NEX-NEG-OLD group), and non-expert Egyptian participants:

Appendix 18, Table 39 shows the average resemblance scores and the cases' ranks by both groups. One-way ANOVA showed that the average resemblance scores given by the non-expert Egyptian group were *significantly higher* (Figure 25) than those given by the NEX-NEG-OLD group (P = 0.000).

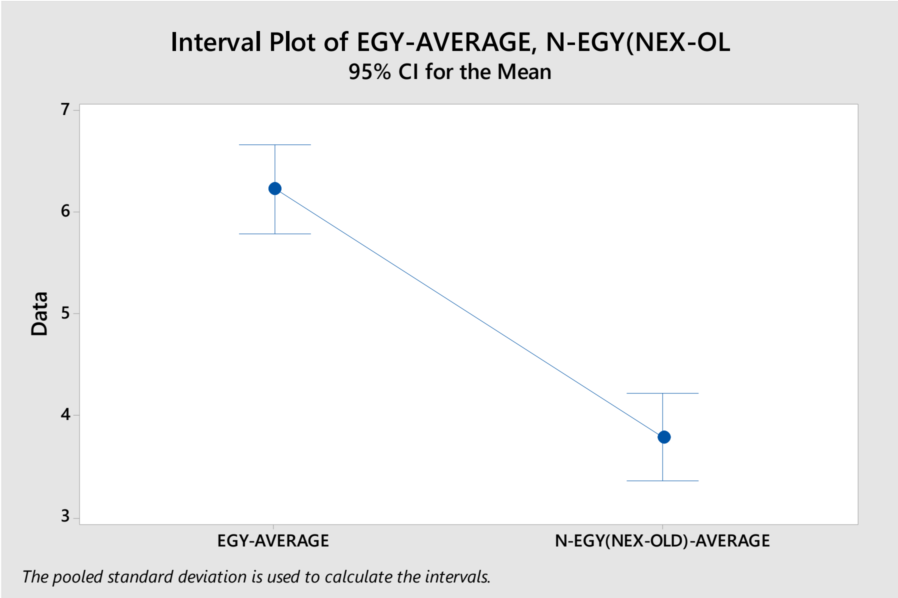


Figure 25: The mean average resemblance scores given by the NEX-NEG-OLD, and the non-expert Egyptian groups

However, the cases' ranks according to the average resemblance scores given by the NEX-NEG-OLD, and the non-expert Egyptian groups *correlated significantly* (Spearman's rank correlation coefficient = 0.368) (P = 0.045).

3- Comparison (3): Non-expert non-Egyptian participants (NEX-NEG-OLD group), and participants with professional experience in forensic anthropology:

Appendix 18, Table 40 shows the average resemblance scores and the cases' ranks by both groups. One-way ANOVA showed that the average resemblance scores given by the forensic anthropology experts group were *significantly higher* (Figure 26) than those given by the NEX-NEG-OLD group (P = 0.004).

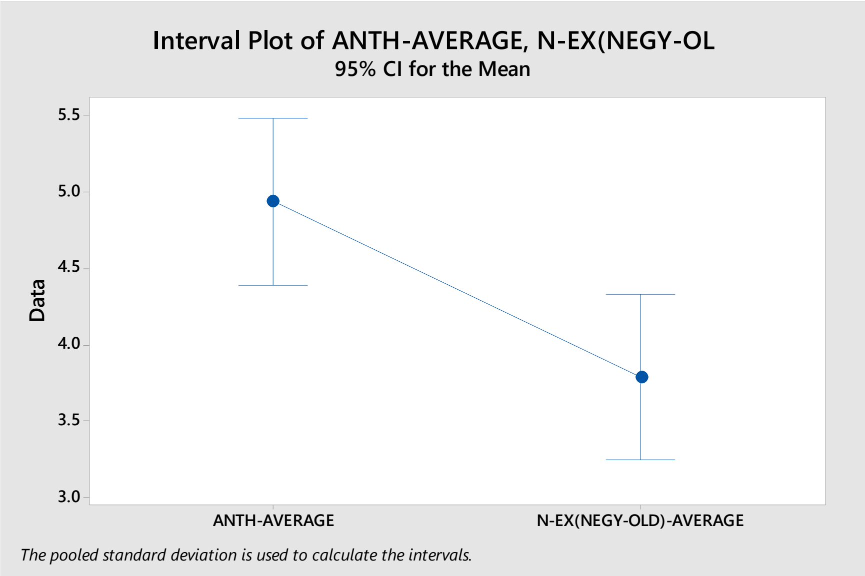


Figure 26: The mean average resemblance scores given by the NEX-NEG-OLD, and NEX-NEG-NEW groups

However, the cases' ranks according to the average resemblance scores given by the NEX-NEG-OLD, and the forensic anthropology experts groups *correlated significantly* (Spearman's rank correlation coefficient = 0.877) (P = 0.000).

(E) Different groups ranking of cases:

Results of statistical correlation of the case ranks according to each participants' groups showed:

- A *significant* correlation between the NEX-NEG-OLD and the NEX-NEG-NEW groups (Spearman's rank correlation coefficient = 0.817) (P = 0.000).
- A *significant* correlation between the NEX-NEG-OLD and the non-expert Egyptian groups (Spearman's rank correlation coefficient = 0.383) (P = 0.037).
- A *significant* correlation between the NEX-NEG-OLD and the forensic anthropology expert groups (Spearman's rank correlation coefficient = 0.877) (P = 0.000).

- A *significant* correlation between the NEX-NEG-NEW and the non-expert Egyptian groups (Spearman's rank correlation coefficient = 0.377) (P = 0.04).
- A *significant* correlation between the NEX-NEG-NEW and the forensic anthropology expert groups (Spearman's rank correlation coefficient = 0.673) (P = 0.000).
- An *insignificant* correlation between the non-expert Egyptian and the forensic anthropology expert group (Spearman's rank correlation coefficient = 0.215) (P = 0.254).

The cases' ranks using this method were then statistically correlated with the cases' ranks using other subjective and objective methods (Section 5.6).

(II) Subjective Tests (3): Face Resemblance Test Version (2); Individual Facial Regions Resemblance Scores

The first exercise (cases 01-15) has been performed by 36 participants (Group 2), while the second exercise (cases 16-30) has been performed by 34 participants (Group 1) Appendix 18, Table 35.

The average resemblance scores given by all participants (combined) to individual facial regions and the calculated overall resemblance scores from them for each case are shown in Appendix 18, Table 41. The ranks of the 30 assessed cases according to the average resemblance scores given by all participants (combined) to individual facial regions and the calculated overall resemblance scores from them are shown in Appendix 18, Table 42. The higher the resemblance score of a given facial reconstruction, the lower it was ranked.

The cases' ranks using this method were then statistically correlated with the case ranks using other subjective and objective methods (Section 5.6).

5.3 OBJECTIVE TEST (1): OBJECTIVE ASSESSMENT BY FACIAL SURFACE OVERALL DISTANCE STANDARD DEVIATION (SD)

All the 30 cases were objectively assessed via facial surface overall distance standard deviation (SD). Appendix 18, Table 43 shows the facial surface overall distance standard deviation (SD) and the cases' ranks accordingly. The lower the facial surface distance between a given facial reconstruction and its real face, the lower it was ranked. The cases' ranks using this method were then statistically correlated with the cases' ranks using the other objective and subjective methods (Section 5.6).

The results of objective assessment of the 30 cases via the overall surface distance standard deviation between the real and the reconstructive faces in this study showed that the surface differences ranged from 1.95 - 6.33 mm, with a mean difference of 3.39 mm. In addition, 25/30 cases (83%) showed a surface distance within a ± 5 mm (7/30 cases (23%) with a surface difference within a ± 2.5 mm, and 18/30 cases (60%) with a surface difference between 2.5 - 5 mm), and 5/30 cases (17%) showed a surface difference more than ± 5 mm.

5.4 OBJECTIVE TEST (2): OBJECTIVE ASSESSMENT BY FACIAL SURFACE DISTANCE AT INDIVIDUAL FACIAL REGIONS

All the 30 cases were objectively assessed via the facial surface distances at the individual facial regions. The ranks of the 30 cases according to the sum of the absolute objective surface differences at all facial regions are shown in Appendix 18, Table 44. The lower the facial surface distance between a given facial reconstruction and its real face, the lower it was ranked. The absolute difference of the surface distance at each facial regions are shown in Appendix 18, Table 45. The ranks of the 30 cases according to the absolute difference of the surface distance at each facial regions are shown in Appendix 18, Table 46. The cases' ranks using this method were then statistically correlated with the case ranks using the other objective and subjective methods (Section 5.6).

5.5 OBJECTIVE TEST (3): OBJECTIVE ASSESSMENT BY CRANIOFACIAL ANTHROPOMETRY

All the 30 cases were objectively assessed via Craniofacial Anthropometry using a number of linear measurements taken from the skull, real and reconstructed faces. From the linear measurements (n=9), linear ratios (n=13) and angles (n=11) were calculated and, then compared between the skull, the real and the reconstructed faces of each case as follows.

A. The correlation of the difference between the linear ratios of the skull and the real face with the differences between the linear ratios of the skull and the reconstructed face.

The differences between the linear ratios of the skull and the real face correlated *significantly* with the differences between the linear ratios of the skull and the reconstructed face (Spearman's rank correlation coefficient = 0.781) (P-Value = 0.000).

B. The correlation of the difference between the angles of the skull and the real face with the differences between the angles of the skull and the reconstructed face.

The differences between the angles of the skull and the real face correlated *significantly* with the differences between the angles of the skull and the reconstructed face (Spearman's rank correlation coefficient was 0.937) (P-Value = 0.000).

C. The absolute differences between the linear ratios of the real and the reconstructed faces

Appendix 18, Table 47 shows the absolute differences between the linear ratios of the real and the reconstructed faces and the averaged difference for each case. Appendix 18, Table 48 shows the ranked 30 cases according to individual linear ratios and their averages.

D. The absolute differences between the angles of the real and the reconstructed faces

Appendix 18, Table 49 shows the absolute differences between the angles of the real and the reconstructed faces and averaged differences for each case. Appendix 18, Table 50 shows the ranked 30 cases according to individual angles and their averages.

To validate the suggested Craniofacial Anthropometry method, the ranks of the assessed cases, according to their individual linear ratios differences, averaged linear ratios differences, individual angles differences, and averaged angle differences, were statistically correlated with the cases' ranks by other objective and subjective assessment methods (Section 5.6).

5.6 CORRELATED TESTS RESULTS

Appendix 18, Table 51 summarises the ranks of the 20 cases assessed via all subjective and objective tests. Appendix 18, Table 52 summarises the ranks of the 30 cases all objective tests and the subjective face resemblance tests.

5.6.1 Correlation between the Subjective Tests

1- Subjective assessment by the face pool versus subjective assessment by the face resemblance (version 1)

Analysis of the ranks of the 20 cases assessed via the 2 subjective tests (Appendix 18, Table 51) showed a *significant* correlation between them (Spearman rank correlation coefficient = 0.624) (P = 0.003).

2- Subjective assessment by the face pool versus the subjective assessment by face resemblance (version 2)

Analysis of the ranks of the 20 cases assessed via the overall resemblance scores calculated from the individual regions scores and the face pool identification rates (Appendix 18, Table 51) showed a *significant* correlation between them (Spearman rank correlation coefficient = 0.551) (P = 0.012).

Also, there was *significant* correlations between the cases ranks via the face pool identification rates and via the resemblance scores of the following individual facial regions:

- The orbital bone (Spearman rank correlation coefficient = 0.505) (P = 0.023).
- The nasal bone (Spearman rank correlation coefficient = 0.608) (P = 0.004).
- The cheek bone (Spearman rank correlation coefficient = 0.496) (P = 0.026).
- The jaw bone (Spearman rank correlation coefficient = 0.489) (P = 0.029).

3- Subjective assessment by the face resemblance (version 1) versus subjective assessment by the face resemblance (version 2)

Analysis of the ranks of the 30 cases assessed via the overall resemblance scores given in the face resemblance test version (1) and the overall resemblance scores calculated from individual regions scores in the face resemblance test version (2) (Appendix 18, Table 52) showed a *significant* correlation between them (Spearman rank correlation coefficient = 0.731) (P = 0.000).

Also, there was *significant* correlations between the cases' ranks via the overall resemblance scores given in face resemblance test version (1) and via the resemblance scores of the following individual facial regions:

- The orbital bone (Spearman rank correlation coefficient = 0.447) (P = 0.013).
- The nasal bone (Spearman rank correlation coefficient = 0.528) (P = 0.003).
- The cheek bone (Spearman rank correlation coefficient = 0.704) (P = 0.000).
- The chin bone (Spearman rank correlation coefficient = 0.623) (P = 0.000).
- The jaw bone (Spearman rank correlation coefficient = 0.753) (P = 0.000).

5.6.2 Correlation between the Subjective and the Objective Tests

1- Subjective assessment by the face pool versus objective assessment by the facial surface overall distance standard deviation (SD)

Analysis of the ranks of the 20 cases assessed via the 2 tests (Appendix 18, Table 51) showed a *significant* correlation between them (Spearman rank correlation coefficient = -0.492) (P = 0.028).

2- Subjective assessment by the face pool versus objective assessment by the facial surface distances at the individual facial regions

Analysis of the ranks of the 20 cases assessed via the 2 tests (Appendix 18, Table 51) showed an *insignificant* correlation between the cases ranks according to the identification rate in the face pool tests and their ranks according to the sum of the absolute objective surface differences at all facial regions (Spearman rank correlation coefficient = 0.05) (P = 0.835).

3- Subjective assessment by the face pool versus objective assessment by the average differences of the craniofacial anthropometric linear ratios

Analysis of the ranks of the 20 cases assessed via the 2 tests (Appendix 18, Table 51) showed an *insignificant* correlation between the 2 tests (Spearman rank correlation coefficient = -0.347) (P = 0.134).

4- Subjective assessment by the face pool versus objective assessment by the average differences of the craniofacial anthropometric angles

Analysis of the ranks of the 20 cases assessed via the 2 tests (Appendix 18, Table 51) showed a *significant* correlation between the 2 tests (Spearman rank correlation coefficient = -0.576) (P = 0.008).

5- Subjective assessment by the face resemblance (version 1) versus objective assessment by the facial surface overall distance standard deviation (SD)

Analysis of the ranks of the 30 cases assessed via the 2 tests (Appendix 18, Table 52) showed a *significant* correlation between them (Spearman rank correlation coefficient = -0.681) (P = 0.000).

6- Subjective assessment by the face resemblance test (version 1) versus objective assessment by the facial surface distances at the individual facial regions

Analysis of the ranks of the 30 cases assessed via the 2 tests (Appendix 18, Table 52) showed a *significant* correlation between the cases' ranks via the subjective overall resemblance score given to each case and via the sum of the absolute objective surface differences at all facial regions (face resemblance test version 1) (Spearman rank correlation coefficient = -0.555) (P = 0.001).

7- Subjective assessment by the face resemblance (version 1) versus objective assessment by the average differences of the craniofacial anthropometric linear ratios

Analysis of the ranks of the 30 cases assessed via the 2 tests (Appendix 18, Table 52) showed a *significant* correlation between them (Spearman rank correlation coefficient = -0.378) (P = 0.039).

8- Subjective assessment by the face resemblance (version 1) versus objective assessment by the average differences of the craniofacial anthropometric angles

Analysis of the ranks of the 30 cases assessed via the 2 tests (Appendix 18, Table 52) showed a *significant* correlation between them (Spearman rank correlation coefficient = -0.472) (P = 0.008).

9- Subjective assessment by the face resemblance (version 2) versus objective assessment by the facial surface overall distance standard deviation (SD)

Analysis of the ranks of the 30 cases assessed via the overall resemblance scores calculated from the individual regions' scores and via the objective facial surface overall distance standard deviation (SD) (Appendix 18, Table 52) showed a *significant* correlation between them (Spearman rank correlation coefficient = -0.431) (P = 0.017).

Also, there was *significant* correlations between the cases' ranks via the objective facial surface overall distance standard deviation (SD) and via the resemblance scores of the following individual facial regions:

- The cheek bone (Spearman rank correlation coefficient = -0.403) (P = 0.027).
- The chin bone (Spearman rank correlation coefficient = -0.363) (P = 0.049).
- The jaw bone (Spearman rank correlation coefficient = -0.553) (P = 0.002).

10- Subjective assessment by the face resemblance (version 2) versus objective assessment by the facial surface distances at the individual facial regions

Analysis of the 30 assessed cases showed an *insignificant* correlation between the cases' ranks via the subjective overall resemblance scores calculated from individual regions scores and their rank according to the sum of the absolute objective surface differences at all facial regions (Appendix 18, Table 52) (Spearman rank correlation coefficient = -0.247) (P = 0.188).

Also, the 30 cases' ranks via the sum of the absolute objective surface differences at all facial regions and via the subjective resemblance score given to each corresponding region (face resemblance test version 2) were correlated. This showed an *insignificant* correlation between the subjective and objective tests in the following individual facial regions:

- The forehead bone (Spearman rank correlation coefficient = 0.122) (P = 0.519).
- The orbital bone (Spearman rank correlation coefficient = -0.208) (P = 0.27).

- The nasal bone (Spearman rank correlation coefficient = 0.16) (P = 0.4).
- The chin bone (Spearman rank correlation coefficient = -0.31) (P = 0.096).

In contrast, this showed a *significant* correlation between the subjective and objective tests in the following individual facial regions:

- The cheek bone (Spearman rank correlation coefficient = -0.39) (P = 0.033).
- The jaw bone (Spearman rank correlation coefficient = -0.505) (P = 0.004).

11- Subjective assessment by the face resemblance (version 2) versus objective assessment by the average differences of the craniofacial anthropometric linear ratios

Analysis of the ranks of the 30 cases assessed via the overall resemblance scores calculated from the individual regions' scores and the objective assessment by average differences of the craniofacial anthropometric linear ratios (Appendix 18, Table 52) showed an *insignificant* correlation between them (Spearman rank correlation coefficient = -0.070) (P = 0.713). However, there was *significant* correlations between:

- The orbital bone's resemblance scores rank and linear ratio CE/BE rank: (Spearman rank correlation coefficient = -0.444) (P = 0.014).
- The cheek bone's resemblance scores rank and linear ratio AB/AD rank: (Spearman rank correlation coefficient = -0.418) (P = 0.022).
- The chin bone's resemblance scores rank and linear ratio AD/BD rank: (Spearman rank correlation coefficient = 0.397) (P = 0.030).
- The jaw bone's resemblance scores rank and linear ratio AB/AD rank: (Spearman rank correlation coefficient = -0.454) (P = 0.012).

12- Subjective assessment by the face resemblance (version 2) versus objective assessment by the average differences of the craniofacial anthropometric angles

Analysis of the ranks of the 30 assessed cases via the overall resemblance scores calculated from individual regions scores and via the objective assessment by the average differences in craniofacial anthropometric angles (Appendix 18, Table 52) showed an *insignificant* correlation between them (Spearman rank correlation coefficient = -0.186) (P = 0.326).

However, there was *significant* correlations between the cases' ranks via:

- The cheek bone's resemblance scores and the AEC angle: (Spearman rank correlation coefficient = -0.364) (P = 0.048).
- The cheek bone's resemblance scores and the ADB angle: (Spearman rank correlation coefficient = -0.406) (P = 0.026).
- The jaw bone's resemblance scores and the ADB angle: (Spearman rank correlation coefficient = -0.438) (P = 0.015).

13- Correlation between the Subjective Tests and the Individual Linear Ratios

There was *significant* correlations between:

- The face pool test and the linear ratio AB/AD (Spearman rank correlation coefficient = -0.471) (P = 0.036), based on 20 cases.
- The face pool test and the linear ratio AC/BE (Spearman rank correlation coefficient = -0.524) (P = 0.018), based on 20 cases.
- The face resemblance test and the linear ratio AB/AD (Spearman rank correlation coefficient = -0.568) (P = 0.009), based on the 30 cases.

14- Correlation between the Subjective Tests and the Individual Angles

There was *significant* correlations between:

- The face pool test and the AEC angle (Spearman rank correlation coefficient = -0.702) (P = 0.001), based on 20 cases.
- The face pool test and the ADB angle (Spearman rank correlation coefficient = -0.49) (P = 0.028), based on 20 cases.

- The face resemblance test and the AEC angle (Spearman rank correlation coefficient = -0.480) (P = 0.032), based on the 30 cases.
- The face resemblance test and ADB angle (Spearman rank correlation coefficient = -0.556) (P = 0.011), based on the 30 cases.

5.6.3 Correlation between the Objective Tests

1- Objective assessment by the facial surface overall distance standard deviation (SD) versus objective assessment by the facial surface distances at the individual facial regions

Analysis of the ranks of the 30 cases via the objective overall surface distance SD and via the sum of the absolute objective surface differences at all the facial regions (Appendix 18, Table 52) showed a *significant* correlation between (Spearman rank correlation coefficient = 0.883) (P = 0.000).

Also, the 30 cases' ranks via the absolute objective surface differences at each facial regions and via the objective overall surface distance SD (Appendix 18, Table 52) were correlated. This showed an *insignificant* correlation at the *chin* bone (Spearman rank correlation coefficient = 0.037) (P = 0.847). In contrast, it showed a *significant* correlation at the following individual facial regions:

- The forehead bone (Spearman rank correlation coefficient = 0.634) (P = 0.000).
- The orbit bone (Spearman rank correlation coefficient = 0.515) (P = 0.004).
- The nasal bone (Spearman rank correlation coefficient = 0.546) (P = 0.002).
- The cheek bone (Spearman rank correlation coefficient = 0.617) (P = 0.000).
- The jaw bone (Spearman rank correlation coefficient = 0.860) (P = 0.000).

2- Objective assessment by the facial surface overall distance standard deviation (SD) versus objective assessment by the average differences of the craniofacial anthropometric linear ratios

Analysis of the ranks of the 30 cases assessed via the 2 tests (Appendix 18, Table 52) showed an *insignificant* correlation between them (Spearman rank correlation coefficient = 0.328) (P = 0.077).

3- Objective assessment by the facial surface overall distance standard deviation (SD) versus objective assessment by the average difference of the craniofacial anthropometric angles

Analysis of the ranks of the 30 cases assessed via the 2 tests (Appendix 18, Table 52) showed a *significant* correlation between them (Spearman rank correlation coefficient = 0.475) (P = 0.008).

4- Objective assessment by the facial surface distance at individual facial regions versus objective assessment by the average differences of the craniofacial anthropometric linear ratios

Analysis of the ranks of the 30 cases assessed via the sum of the absolute objective surface differences at all the facial regions showed and the cases' ranks according to the objective average differences of the craniofacial anthropometric linear ratios (Appendix 18, Table 52) showed an *insignificant* correlation between them (Spearman rank correlation coefficient = 0.216) (P = 0.251).

5- Objective assessment by the facial surface distance at individual facial regions versus objective assessment by the average differences of the craniofacial anthropometric angles

Analysis of the ranks of the 30 cases assessed via the 2 tests (Appendix 18, Table 52) showed a *significant* correlation between them (Spearman rank correlation coefficient = 0.453) (P = 0.012).

6- **Correlation between the Objective Facial Surface Overall Distance Standard Deviation (SD) and the Individual Craniofacial Linear Ratios**

Results, based on 30 cases, showed *significant* correlations between facial surface overall distance standard deviation (SD) and the following linear ratios:

- The linear ratio BC/BD (Spearman rank correlation coefficient = 0.398) (P = 0.029).
- The linear ratio BC/CE (Spearman rank correlation coefficient = 0.443) (P = 0.014).

7- **Correlation between the Objective Facial Surface Overall Distance Standard Deviation (SD) and the Individual Craniofacial Angles**

Results, based on 30 cases, showed *significant* correlations between facial surface overall distance standard deviation (SD) and the following angles:

- The AEC angle (Spearman rank correlation coefficient = 0.437) (P = 0.016).
- The CAE angle (Spearman rank correlation coefficient = 0.380) (P = 0.038).
- The ACE angle (Spearman rank correlation coefficient = 0.546) (P = 0.002).
- The CDB angle (Spearman rank correlation coefficient = 0.378) (P = 0.039).
- The CEB angle (Spearman rank correlation coefficient = 0.430) (P = 0.018).

5.6.4 Summary of combined results

Table 13 shows a summary of the P-values of the significant correlations between different tests.

Table 14 shows a summary of the P-values of the significant correlations between different tests and the individual linear ratios and angles respectively.

Table 15 shows a summary of the P-values of the significant correlations between the individual linear ratios and angles and the objective facial surface overall distance standard deviation (SD) with the individual facial regions resemblance scores.

Table 13: The P-values representing the significant correlations between different tests

TEST	ID% (20 cases)	Overall Resemblance Scores (V1) (30 cases)	Overall Obj. SD (30 cases)
ID%	N/A	0.003	0.028
Overall Resemblance Scores (V1)	0.003	N/A	0.000
Overall Obj. SD	0.028	0.000	N/A
Overall Facial Regions Resemblance Scores (V2)	0.012	0.000	0.017
Orbital Bones Resemblance Score	0.023	0.013	
Nasal Bones Resemblance Score	0.004	0.003	
Cheek Bones Resemblance Score	0.026	0.000	0.027
Chin Bones Resemblance Score		0.000	0.049
Jaw Bones Resemblance Score	0.029	0.000	0.002
Overall Facial Regions Obj. SD		0.001	0.000
Average Linear Ratios		0.039	
Average Angles	0.008	0.008	0.008

Table 14: The P-values representing the significant correlations between the individual linear ratios and angles and other subjective and objective tests

Linear Ratio	ID% (20 Cases)	Overall Resemblance Scores (30 Cases)	Obj. SD (30 Cases)
AB/AD	0.036	0.009	
AC/BE	0.018		
BC/BD			0.029
BC/CE			0.014
AEC	0.001	0.032	0.016
ADB	0.028	0.011	
CAE			0.038

ACE			0.002
CDB			0.039
CEB			0.018

Table 15: The P-values representing the significant correlations between the individual linear ratios and angles and the objective facial surface overall distance standard deviation (SD) with the individual facial regions resemblance scores

	Forehead	**Orbital Bone**	**Nasal Bone**	**Cheek Bone**	**Chin Bone**	**Jaw Bone**
AB/AD				0.022		0.012
AD/BD					0.030	
CE/BE		0.014				
AEC				0.048		
ADB				0.026		0.015
Objective Overall SD				0.027	0.049	0.002

5.7 COMPARISON BETWEEN INDIVIDUAL CASES BASED ON DIFFERENT TESTS RANKS

To illustrate the individual differences between the 30 cases' ranks, it should be noted that the calculated ranks using the subjective tests correlate negatively with the calculated ranks using the objective tests. For example, a case receiving the first rank in a face resemblance test means that the reconstructed face received the highest resemblance scores (i.e. the highest accuracy). In contrast, a case receiving the last rank in the objective facial surface deviation means that it showed the least objective distance between the aligned real and the reconstructed faces (i.e. the most accurate fit between the real and the reconstructed face). Therefore, to account for this negative correlation in plotting the 30 cases ranks together, the objective tests' ranks were reversed by deducting them from 31 (not 30 to keep the rank between 1 and 30). The ranks of each case using each test were then summed, with exclusion of the cases' ranks via the average linear ratios' differences as they did not correlate with the majority of the other tests. The cases' ranks using the different subjective and objective tests, did not differ significantly when compared via the non-parametric Freidman's test in both the 20 cases (P- Values = 0.58) and the 30 cases (P- Values = 0.19). Figure 27 and Appendix 18, Table 53 show the ranks of the 30 cases ordered from the lowest (i.e. the most accurate facial reconstruction) to the highest (i.e. the least accurate facial reconstruction).

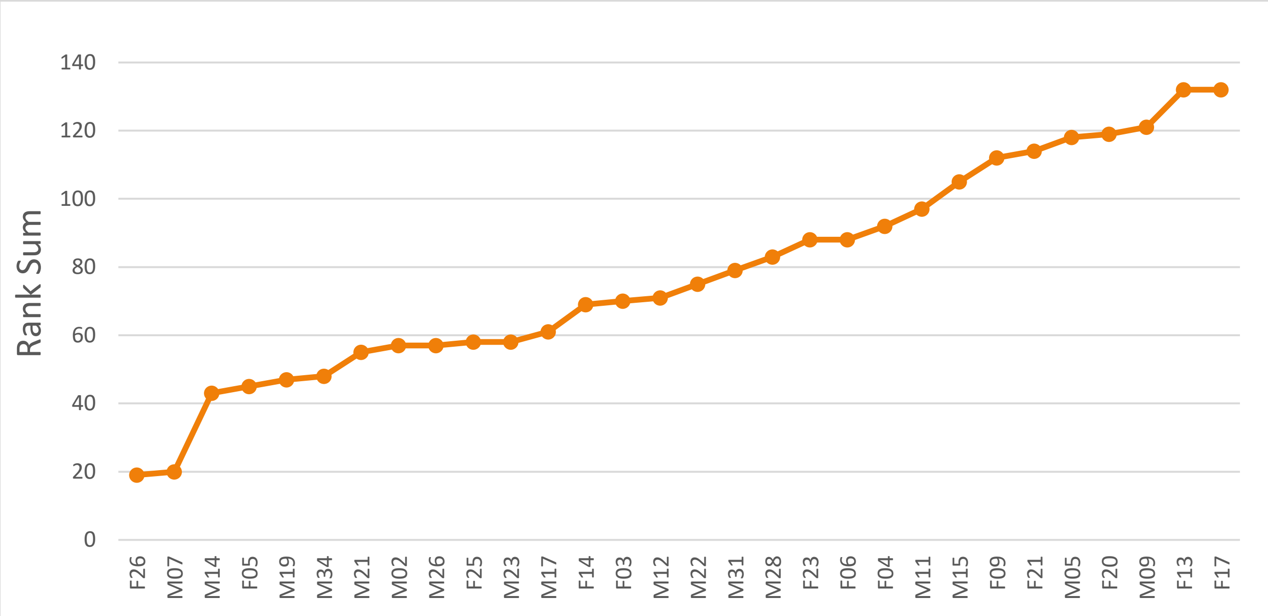


Figure 27: A graph showing the sum of the ranks of 30 assessed cases according to all subjective and objective tests (except Average differences of linear ratios), after adjusting for the negative ranks correlations

Appendix 19 shows the 30 studied skull cases, real and reconstructed faces, as well as the colour maps of the objective assessment, in addition to the results obtained in the objective assessment of the individual facial regions.

CHAPTER 6: DISCUSSION

This thesis consisted of two parts; a pilot study (Figure 3) and a main study (Table 8 and Table 11). The pilot study was conducted to validate the main components of the proposed facial reconstruction method (i.e. the facial templates, and the facial soft tissue depths). In addition, different formats of the subjective methods used for assessment of facial reconstructions were tested in to reach a proper design to be applied in the main part of this thesis in order to be more reliable in reflecting the accuracy of the tested facial reconstructions. In the main study of this thesis, the faces of 30 CT scanned skulls of an Egyptian population were reconstructed using average facial templates generated from CT scanned faces of the same population. The resulting facial reconstructions were subjectively and objectively assessed. Moreover, the reliability of different subjective assessment methods was investigated taking into account the subjective characteristics of the observers performing the subjective assessment while interpreting the results of these tests. Therefore, a more realistic evaluation of the proposed method was discussed. In addition, a new objective assessment method was developed and validated in comparison to with other methods previously published in literature.

6.1 The components of the proposed method

The forensic facial reconstruction method proposed in the present study adopts the “Outside Inwards” approach by digitally warping 3D scanned facial templates onto the skulls of the same population. The facial reconstruction was guided by a number of anatomical landmarks with pre-defined population-specific facial soft tissue depths. In the pilot study, the components of this proposed method (i.e. the facial templates and the facial soft tissue depths) were tested before being used in the main part of this thesis.

6.1.1 The facial templates

A specific face template was previously used as a template for facial reconstruction in other studies either as facial components/composites (i.e. partial) (Nelson and Michael, 1998), or in a complete form (i.e. as whole faces) (Shahrom et al., 1996, Vanezis, 2008, Vanezis and Vanezis, 2000, Vanezis et al., 1989). However, the selection of a face (or a face feature) from a database was claimed to be as subjective as manual sculpture techniques (Nelson and Michael, 1998). In addition, using a set of specific facial features of a certain face was questioned by some researchers as being reliant only on little information from the skull and imposing certain features on the reconstructed face (Andersson and Valfridsson, 2005, Wilkinson et al., 2006), especially when using a single face. This can be confusing not only to the acquaintances of the unknown individual, but also to those of the person whom facial template was used, which would result in false recognition (Claes et al., 2006, Nelson and Michael, 1998).

Moreover, problems rise when the facial template is disproportionate to the underlying skull. For example, a male facial template cannot be used on a female skull and vice versa due to sexual dimorphism in the skull features and dimensions (Salah et al., 2008). Even within the same sex, a long face template might be sagging when fitted onto a shorter/smaller skull, which can be encountered with a limited or restricted database. Similarly, a Caucasian face is disproportionate to a negroid skull (Jones, 2001). The outcome of such discrepancy is a more pronounced deformation, resulting in an unrealistic, or caricature-like facial reconstruction (Claes et al., 2006, Claes et al., 2010), or a face that is too stretched to a degree that a skin tear may be apparent (Jones, 2001). This might be more obvious when one or more of the biological profile features (race, age, sex) of the skull is uncertain.

On the other hand, the skull features can be strong enough or the facial template can be flexible enough that the facial contours would sufficiently change to take the shape of the underlying skull. Although a bigger database increases the chance that the reconstructed face resembles the deceased's face during life (Shahrom et al., 1996), the likelihood of finding that flexible single facial template can never be guaranteed (i.e. it can happen only by a mere chance). Additionally, there could be a "more suitable" facial template than that, supposedly, flexible facial template, for which the search could be endless. Furthermore, the "best" single face template may not always be suitable for all skulls, which again is difficult to be certain.

Therefore, the pilot study of this thesis investigated whether certain facial templates were better than others for the proposed method. The comparison included single and average facial templates. The results showed that certain reconstructed faces received high ranks in either the face pool or the face resemblance tests, and others received lower ranks in both tests. In contrast, certain reconstructed faces received similarly high ranks in both the face pool and the face resemblance tests. Therefore, these faces were considered more accurate than the faces which received high rank in one test type only. Hence, the facial templates used for reconstructing these accurate faces, were considered more useful or "better" than others as potential facial templates for forensic facial reconstruction.

To explain these individual differences between facial templates, we can describe fitting a template onto the target skull as a process which is performed by finding a set of parameters related to the template itself (i.e. the principal modes) (Claes et al., 2006), so that all the reference landmarks of the template fit the corresponding target skull landmarks. From a practical point of view, this might explain why a facial template is better than another (i.e. results in a higher recognition) when warped onto a skull of a third person, even if the skull and both facial templates are anthropologically similar (i.e. of the same, sex, race, and age group). In addition, the elasticity of the facial models in a database is determined by the database size, as well as the degree of freedom in the database (i.e. whether it contains enough facial models to represent each subject in the studied sample). This cannot be ensured with every studied/unknown case as the target faces (to be reconstructed) are unknown, there is no possible way of determining whether the database contains the appropriate facial model for each target skull or not. Therefore, it is difficult to attempt to select that "best" facial template from the database to guarantee a "successful" facial reconstruction. This highlights the problem mentioned above that certain facial templates are more convenient than others for facial reconstruction with the inherent difficulty in predicting this prior to the positive identification of the real face.

Suggestions to overcome the problems encountered with using a single face with a particular bio-mechanical form and genetic specificities have been presented by researchers. For example, in order to negate the specificities of a single face template, it was suggested to digitally merge (aggregate) a number of highly specific cranial forms (i.e. anthropologically similar single faces). Some researchers first deformed multiple face templates towards the given skull separately (i.e. deformable carniofacial models) and then averaging the resulting faces in to one face that is considered as the facial reconstruction (Claes et al., 2006, Claes et al., 2010, Vandermeulen et al., 2006). Other studies suggested to average the facial templates first to generate one 'average' face template that does not contain "actual" cranial data and then fit the resulting average face onto the skull (Nelson and Michael, 1998). As this latter technique involves directly merging the facial templates then warping the averaged face only onto the studied skull, it is less lengthy than the former technique which requires modifying all the facial models individually before merging them.

Furthermore, it is not uncommon that the anthropological profile of the skull is uncertain, or the database does not contain a full representative sample of different facial types. Thus, while it is recommended that the facial template belong to the same anthropological group as the target (Quatrehomme et al., 2007, Vanezis, 2008), the biological profile of the remains can be uncertain. In that case, it might be a good practice to use more than one facial template from different age, or race, groups for facial reconstruction to increase the chance of identification. This, therefore, adds to the advantages of average over single facial templates. In conclusion, an average facial template can be considered a "safer" and "better" selection, as the "best" selection may never be found. Moreover, Salah et al. (2008) suggested to extend the single face model into multiple category-specific alternative average face models based on gender and morphology.

While the term "average" face was used in some studies as a representative of a group of similar individuals but it was a single face (Shahrom et al., 1996, Attardi et al., 1999, Turner et al., 2005), in the present thesis, a digitally "averaged" face that was mathematically generated by digitally merging a number of single/specific faces was used. Facial image averaging is an active research area in a number of fields. For example, Salah et al. (2008) presented a method for facial image registration by registering a suspected face to an average face model (AFM), which automatically determined the correspondence to the faces in the database for face recognition. Moreover, Islam et al.

(2015) reported a number of applications of facial averages in orthodontics and facial surgeries.

The results of the pilot study of this thesis showed that the digitally averaged faces were better than single faces in reaching a good resemblance and a higher recognition of the target. This might be because “averaging” single faces generates a face with a higher robustness in fitting different skulls by different degrees. In addition, an average facial template reduces the subjectivity involved in the selection of single faces and leads to results that are less biased by a certain face, with a less chance of misidentification.

6.1.2 Facial soft tissue depths

Facial soft tissue thickness has been constituted as an integral component of forensic facial reconstruction that directly affects the appearance of the reconstructed face and, hence, its recognition. These facial soft tissue measurements are used as a guide for most 3D forensic facial reconstructions methods and are thought to be influenced by different individual factors (e.g. Body Mass Index (BMI), age, and sex).

The newer measuring methods and the increasingly published various population-specific data tables has raised the accuracy of the measured facial depths. However, this has increased the complexity and resulted in lack of standardised data sets (Brown et al., 2004). In 2008, Stephan and Simpsons started analysing and pooling the previously published facial soft tissue thickness data for adults (Stephan and Simpson, 2008a) and sub-adults (Stephan and Simpson, 2008b) that were collected by other researchers from both cadavers and living subjects using different methods (e.g., needle insertion, Ultrasound, CT, MRI). The authors pooled these data and presented Tallied Facial Soft Tissue Depth Data (T-Tables) of facial soft tissue thickness measurements. The aim of these tables was to provide a simple, standardised, and statistically validated facial soft tissue data, with low standard errors. Updated data are published regularly in the researcher's website (Stephan) to guide other researchers while performing the forensic facial reconstructions.

However, substantive differences were found when Parks et al. (2014) compared their facial soft tissue depths collected by CT scans of nearly 400 living subjects at the same facial landmarks and following the same guidelines as Stephan's Tallied Facial Soft Tissue Depth Data (T-Tables) published in 2012, and the same Stephan's non demographic-specific set (2012). Another comparison conducted between Parks et al. (2014)'s set and another demographically similar set by Rhine and Moore (1982) for European – American males of normal and obese BMI, also showed substantive differences. The authors explained these differences by a number of reasons, such as; the error propagation through measurements when taken by different users due to inherent human variability, in addition to different sample size, populations, collection period and method. For example; Parks et al. (2014) used CT scans of contemporary living Americans scanned in supine positions. In contrast, Stephan's non-population specific dataset was pooled based on variable collection methods, with different postures and over a large timeframes (1883–2012). Therefore, most of the Parks et al. (2014)'s measurements were larger than Stephan's data. In addition, Rhine and Moore collected data from cadavers using needles in supine position in the 1980s. Moreover, the recent secular weight change and increase in the prevalence of overweight in the United States of America compared to 1980s, might also explain the difference between the authors' and Rhine and Moore's data.

It should be noted, however, that statistically significant differences in facial depths measurements are not always significant to facial reconstructions from a practical point of view (Parks et al., 2014). For example, although the sexual dimorphism in facial soft tissue depths might have applications in the orthodontics field (Cha, 2013), it may have no practical influence on the likeness of the reconstructed face to the real face from a forensic identification point of view (Stephan et al., 2005a, Smith and Throckmorton, 2006). Similarly, De Greef et al. (2009) cited a number of studies which presented facial reconstructions using facial depths of non-matching ethnic groups.

In the pilot study of this thesis, a comparison was conducted, between two facial soft tissue depths. The first set was a combined set formed of the old but widely used population-specific Rhine and Moore (1982) and Helmer (1984) sets (Hayes, 2014). This combined set was previously used by Vanezis (2008). While, the second set was a modified set of depths produced by replacing the majority of the old depths in the first set with more recent and pooled depths adopted from Stephan (2014) that was aimed for all populations. However, the comparison involved the practical function of the facial soft tissue depths which is producing facial reconstructions. Thus, faces were reconstructed using the two sets separately and then objectively assessed by measuring the surface distances between the reconstructed and real faces. The results showed an insignificant difference between the faces reconstructed using the old and the new data. Thus, a previously validated data are enough even if they were published 30 years apart. Conversely, non-population specific measurements can work similarly to population specific data to a large degree. Therefore, precise measurement of the facial soft tissue depths is not critical to facial recognition, rather careful examination of the skull morphology and anatomy is what should be considered (De Greef et al., 2009, Smith and Throckmorton, 2006, Stephan et al., 2005a). This is particularly useful with unsure estimation of the skull traits (e.g., of mixed race, unsure sex).

Nonetheless, population specific facial soft tissue depths are still preferred by forensic facial reconstructors, even with the need to impose certain modifications to the used sets according to the researcher’s judgement (Vanezis, 2008). Even more, a review of 15 facial reconstruction articles published between 2000-2013 (Hayes, 2014) revealed that most of them used Helmer’s (1984) followed by Rhine and Moore’s (1982) soft tissue depths measured using ultrasound and direct needles of cadavers respectively. Vanezis (2008) used combined data tables from both studies for facial reconstruction of adult Caucasian males and females. However, when this combined set of Rhine and Moore (1982) and Helmer (1984) facial soft tissue depths for the white Caucasian population was used for reconstructing the faces of an Egyptian population in the main part of the present thesis, the reconstructed faces of the skulls using average European facial templates did not carry sufficient resemblance to the Egyptian targets as per the researcher’s judgement. Therefore, literature search have been carried out for an updated reference for facial soft tissue depths of Egyptian population studied in this thesis, which led to one reference (El-Mehallawi and Soliman, 2001), collected by ultrasound, which aimed at providing data tables for the Egyptian population, and showed the interpopulation difference between Egyptians and other populations by comparing their results to other studies. While, computed tomography is considered the gold standard for many forensic facial reconstruction studies in extracting 3D skull and facial surfaces (Claes et al., 2010), ultrasound has a comparable accuracy to CT in measuring facial soft tissue depths (De Greef et al., 2005). Therefore, El-Mehallawi and Soliman (2001)'s facial depths table was used as an acceptable reference for this thesis, which also could be a method for validating their data.

Furthermore, De Greef et al. (2006) suggested that the accuracy of the craniofacial reconstructions can be increased by including a high number of the landmarks used to allow better determination of the facial contours. In the pilot study of this thesis, an experiment was conducted in which four sets of facial soft tissue depths were compared by the objective assessment of the faces reconstructed using these sets; *Set (1): 40 landmarks* of Rhine and Moore (1982) and Helmer (1984), *Set (2): 38 landmarks*, removing only 1 pair of the cheek landmarks, *Set (3): 36 landmarks*, removing 2 pairs of the cheek landmarks, and *Set (4): 34 landmarks*, removing all 3 pairs of the cheek region. The differences between them were statistically insignificant when objectively compared via measuring the overall surface distance differences between the reconstructed and the real faces. However, the 36-landmarks set (Set 3), where 2 cheek landmarks on each side were removed, showed a closer fit between the real and the reconstructed faces, even more than the full set (1). These findings indicate that the cheek landmarks are very important for forensic facial reconstruction and cannot be omitted completely, as in set (4). This was consistent with Vanezis (2008) who recommended using the 36-landmarks set (Set 3), as the omitted landmarks were difficult to locate. Thus, within limits, some modifications can be made, subject to the judgement of the practitioner, for a better facial reconstruction.

In conclusion, once the population of the remains is confirmed, it would be better to use a population-specific set of facial depths. As only averaged soft tissue thickness measurements at a number of craniofacial landmarks are used, De Greef et al. (2006) suggested a large database for measuring the depths can increase the accuracy of the craniofacial reconstructions. So, although it is important to have a standardised set of population-specific facial depths for the accuracy of the facial reconstructions, standardising the facial soft tissue depths is difficult as it is as variable as the individuals within a population. Therefore, a previously validated set, even if old, is sufficient. This is because the differences between the sets are not always reflected on the resulting facial reconstructions in relation to the real faces (i.e. whether these differences would affect the recognition of the reconstructed face) (Stephan and Simpson, 2008a). However, as Stephan's T-Tables are not population specific, they might be useful if the population of the unknown remains is uncertain or in case of unavailable facial depths tables for the estimated population. In addition, continuous updating the facial depths for more modern population data is important for more accurate facial reconstructions.

As a result, it can be claimed that, the soft tissue depths have a less influential, although still necessary, role on the positive identification of the reconstructed face than that of the skull architecture, for example. This is simply because each skull has its own unique anthropometric measurements, thus, requires precise mathematical calculations, thus, cannot be standardised (Jedrzejowska, 2001). Therefore, precise measurement of the facial soft tissue depths is not critical to the facial recognition, but rather careful examination of the skull morphology and anatomy is what should be considered (De Greef et al., 2009, Smith and Throckmorton, 2006, Stephan et al., 2005a). This conclusion is particularly useful with unsure estimation of the skull traits (e.g., of mixed race, unsure sex). Moreover, it is rather more important, to categorise the facial landmarks into more and less influential landmarks. Otherwise, it would be left to the subjective practitioner's experience and judgment to add or omit facial landmarks (Vanezis, 2008), and as shown in the landmarks experiment of the pilot study. Furthermore, more efforts should be directed to standardising the set of craniofacial landmarks, with proper definitions and accurate descriptions of the locations and the directions between the corresponding cranial and facial landmarks with describing the best orientation for placing each landmark. This was attempted by Brown et al. (2004) who published a catalogue of a set of landmarks (located on the frontal, temporal, zygomatic, nasal, maxilla, mandible, occipital and parietal bones). The published catalogue, included definitions of these landmarks cross referenced to common definitions of landmarks by the bone after surveying a number of former literature. This catalogue also included; a table grouping the landmarks by approximate location, highlighted by a bone image, together with some encountered similarities and discontinuities between the tissue depth landmarks and the methods used for measuring the soft tissue depths. Therefore, as a part of the pilot study of this thesis, a user manual, including descriptions, with figures, of the locations of the recommended landmarks (Appendix 20). Population specific depths at these landmarks can then be used.

6.2 ASSESSMENT OF THE FACIAL RECONSTRUCTION ACCURACY

The accuracy of facial reconstructions can be seen, from a practical point of view as the final goal of a craniofacial reconstruction is the recognition or the identification success (Short et al., 2014). Thus, the success of the facial reconstruction can be measured as the ability of a facial reconstruction to generate purposeful and correct facial recognitions with no other identification methods available (Stephan and Cicolini, 2008). Moreover, a distinction should be made between a false positive (absolute) recognition of the facial reconstruction of a missing person, and an actionable recognition that will lead to an investigation to confirm the identity by conclusive identification methods (e.g., dental records, DNA) (Richard et al., 2014). Thus, the success rate of facial reconstruction can be defined as the number of forensic cases that were identified following the public recognition of the advertised reconstructed face of the unknown individual (Decker et al., 2013, Wilkinson, 2007). Wilkinson (2007) reported a success rate of 70% (16/23 cases) by the University of Manchester between 1982 and 2005, and 64% (7/11 cases) by herself between 1997 and 2005. Phillips (2001) presented a number of forensic cases that were solved aided by the facial reconstructions, where positive identification was confirmed by the victims' relatives.

On the other hand, in research studies, the accuracy of facial reconstructions is measured by whether they display true anatomical similarity to the target/unknown individual. The two types of comparisons are distinctive as high anatomical similarity, for example, does not guarantee correct recognition of the target individuals, and vice versa (Stephan and Henneberg, 2006). Therefore, the predictive evaluation of the facial reconstruction accuracy is a part of applied research to validate the presented methods of facial reconstruction, regardless of whether it would lead to recognition in real forensic cases or not (Hayes, 2016). While the recognisability is only subjectively assessed (i.e. depending on subjects' judgement), the anatomical similarity could be assessed via subjective, or objective (i.e. with the aid of computer programs) tests (Richard et al., 2014).

6.2.1 Subjective assessment methods

Broadly speaking, this includes two test types; the face pool and the face resemblance tests. A face pool test is similar to the eyewitness line-ups used in police investigation. In live eyewitness line-ups, the suspect, along with several "fillers" or "foils" (people of similar height, build, and complexion) stand side-by-side, both facing and in profile. Alternatively, photographs of the suspect and fillers can be shown to the identifier in what is called a "photo line-up", which might be presented to the eyewitness sequentially or simultaneously (Evidence and America, 2003). The concept of police line-ups is adopted when designing a face pool test in numerous forensic facial reconstruction studies (Moyers, 2007, Stephan and Henneberg, 2006, Wilkinson et al., 2006), but with the identification of a victim instead of a perpetrator (Stephan and Arthur, 2006, Stephan and Henneberg, 2006). In the face pool test, the percentage or rate of the correct identification of the target face is determined to assess the possibility of correctly identifying an unknown person by their family or acquaintances based their facial reconstruction. Other researchers (Richard et al., 2014) suggest the interpretation of the face pool test results in the form of test sensitivity [true positive responses/(true positive responses + false negative responses)], and test specificity [true negative responses/(true negative responses + false positive responses)], rather than absolute above chance identification rates. Based on that, the test sensitivity is considered as the primary benchmark for success, while the test specificity is more time saving (Richard et al., 2014).

In general, 2 forms of face pool tests were described in literature. These forms were referred to in the present thesis as face pool test forms (A) & (B). A face pool test form (A) involves comparing a target's real face with a pool of computer generated faces (the target's digital reconstructed face and other computer generated or scanned faces) of the same complexion, appearance, and expression as "foils" or "fillers". In this form, the target’s face can be an antemortem photograph (Wilkinson et al., 2006, Vanezis, 2008), or a 3D scanned face (Claes et al., 2006). In contrast, a face pool test form (B) involves comparing a target's facial reconstruction with a pool of photographs; including that of the target and other "foils" or "fillers" subjects similar to the target's age, sex, race, built, and general face morphology, but with eyes open and keeping head hair (Stephan and Henneberg, 2006, Moyers, 2007).

On the other hand, the second type of subjective tests is the face resemblance ranking “the face resemblance test”, which comprises a direct visual comparison between the reconstructed and the real faces. The aim of this test is to quantify the degree of similarity/resemblance between the reconstructed and the real faces and to test whether the target can be identified has the reconstructed face been advertised in real life (Parks et al., 2013). The face resemblance test has been also used in literature to assess the accuracy of the facial reconstructions (Snow et al., 1970, Stephan and Arthur, 2006, Stephan and Henneberg, 2006, Moyers, 2007, Stephan and Cicolini, 2008, Vanezis, 2008).

In general, two forms of resemblance tests were described in literature. These forms were referred to in the present thesis as face resemblance test forms (A) & (B). A face resemblance test form (A) involves subjectively assessing each facial reconstruction separately by assigning a score, from a rising scale, with lower and upper limits, to the reconstructed face according to its similarity to the target’s real face (Moyers, 2007, Stephan and Henneberg, 2006). The commonly employed resemblance scales are either a numerical scales only, or numerical and descriptive scales. In the former type, one end of the scale represents no resemblance and the other end represents a high resemblance and the observer assigns a score to the reconstructed face from this scale (Stephan and Arthur, 2006, Stephan and Henneberg, 2006). In contrast, in the numerical and descriptive scales, each number indicates a description of the degree of resemblance (e.g., no, slight, approximate, close, and strong resemblance) (Moyers, 2007, Stephan and Cicolini, 2008). The total scores given by all observers to each reconstructed face are then calculated and compared with those of other facial reconstructions. The facial reconstructions can then be ranked according to the scores given by the observers to each one of them. In contrast, a face resemblance test form (B) involves comparing a number of facial reconstructions of the same target individual at the same time. These reconstructions are directly ranked/rated by the observers against one another according to their similarity to the target (Vanezis, 2008). In this thesis, a face resemblance tests form (A) was used with a numerical scale with the lower and the upper ends representing no and a high resemblances respectively.

The face pool tests can be designed with foil individuals that match the general description of the target only (e.g., white male between 30 and 40 years of age) (Snow et al., 1970). Furthermore, instructions were set for designing a police line-up, and hence a face pool test, to minimise the chances of mistaken identification while still permitting witnesses to identify the suspect (Evidence and America, 2003). It was recommended that the face pool should be formed of one target and a minimum of five "foil" or "filler" (non-target) subjects, which should be similar to the target's sex, race, and age. In addition, the target should not be standing out from or too similar to the foils (Evidence and America, 2003). It is much harder to correctly identify the target from a face-pool of similar faces than dissimilar faces (Claes et al., 2006). Although a consistent appearance between the target and the foils should be ensured, the foils should have (or enhanced to have) similar, but not identical, features to the target (Evidence and America, 2003). Furthermore, it is important to select foil photos with sufficient similarity in quality and appearance to the target's photograph. Otherwise false positive identification could occur. For example, the target can be selected by the assessors even in the absence of the facial reconstruction in the face pool (Stephan and Henneberg, 2001).

It is acknowledged that "recognition" is what is sought by presenting a facial reconstruction to the public. Thus, researchers usually tend to prefer the face pool format to assess the facial reconstruction accuracy as it assesses the ability for the target individual to be "recognised" from a group of faces (Stephan and Henneberg, 2001, Stephan and Arthur, 2006, Stephan and Henneberg, 2006). However, in a real forensic case, what actually happens is that a friend or a relative of a missing person, who might have been keeping an eye on the media for such announcements, claims recognition. To confirm this "recognition", this person compares/assesses the "resemblance" between the advertised reconstructed face and the face of their missing one, either relying on their memory or by direct comparison with any ante mortem photos of the target. In essence, this scenario is similar to a face resemblance test rather than to a face pool test. Moreover, in other realistic situations, skeletal remains may be found and their face is reconstructed. The police first conducts a preliminary search among the photos of the missing individuals, and/or contacts their relatives aiming for a recognition and positive identification (Richard et al., 2014). This scenario also simulates a face resemblance test rather than a face pool test. It can therefore, be concluded that a face resemblance test format is more likely to occur in real forensic cases than a face pool test format. This was shown in Snow et al. (1970) who used the face resemblance scenarios to assess reconstructed faces subjectively assessed by the acquaintances, but used the face pool scenarios by unfamiliar volunteers.

However, it has been argued that in research studies the accuracy of the facial reconstructions is better measured in terms of recognisability rather than anatomical resemblance (i.e. by face pool rather than face resemblance tests) (Stephan and Arthur, 2006, Stephan and Henneberg, 2006). Testing the recognisability can be seen similar to the scenario of identification of a suspect by an eyewitness, but with identification of a victim instead of a perpetrator. In that sense, the facial resemblance ratings are described as an extreme degree of a show-up, where the assessors know the target, which makes them already biased, and do not have the option of identifying someone else (Stephan and Arthur, 2006, Stephan and Henneberg, 2006). Besides, as the assessors of the face resemblance tests, especially if non-experts, know this is the target, they tend to look for dissimilarities rather than similarities, unlike in the face pool tests (Wilkinson, 2008).

Stephan and Arthur (2006) conducted an experiment assessing two facial reconstructions of one female skull generated by 2 separate practitioners using manual clay modeling. The first facial reconstruction was performed by an experienced practitioner with direct access to the photograph of the target, while the second facial reconstruction was performed by an inexperienced practitioner under blind condition. The authors compared between the face pool test results (percentage of correct identification) and the face resemblance test results (scores from 0 – 5) of the 2 faces. The results showed that the first face was correctly recognised at a higher rate in face pool tests than the second face. Yet, when directly compared to the target, both reconstructed faces received close resemblance scores. Furthermore, Stephan and Henneberg (2006) compared between three subjective assessment methods (face resemblance test, simultaneous face pool test, and sequential face pool test) of 2 versions of a facial reconstruction; with and without head and facial hair. The results showed that both facial reconstruction versions received high, but similar, resemblance ratings, both by the assessors and by a facial reconstruction expert. Unexpectedly, in simultaneous and sequential face pool tests, the target identification rates by assessors with no previous knowledge of the target were higher than those with access to it.

In the aforementioned two studies, the authors explained the discrepancy between the face resemblance ranks and the face pool tests by that resemblance scores were not true indicators of the expected the recognisability of the reconstructed faces. They concluded that the resemblance ratings were insensitive measures of the accuracy of facial approximations and, thus, lent further weight to the use of recognition tests (i.e. face pool) tests in facial reconstruction assessment. However, in the former study (Stephan and Arthur, 2006), the non-standardised conditions under which the 2 faces were reconstructed falsely exaggerated a difference in their identification rate based on which the authors drew their conclusion based on one case only. Even with the second inexperienced practitioner who had no access to the ante mortem photograph of the target, the identification rate of the second face was also above random chance, which is still considered significant, even if low. Thus, although both faces were correctly identified above chance, the difference between their identification rates was attributed to the different reconstruction conditions. However, when it came to a direct comparison with the target, via the face resemblance tests, both faces received similar resemblance scores.

Moreover, as shown in Stephan and Henneberg (2006)'s study, a facial reconstruction expert stated that the high resemblance between the facial reconstructions and the target antemortem photograph would make them recongnisable. Yet, the face pool results did not reflect this conclusion even with the effort and thoughts put to design the face pool tests in the least biased and the more practical way. The authors explained that by the poorer resolution of the ante mortem target photograph than the foil photographs and still concluded that the face pool tests are more sensitive than the fact resemblance tests. Conversely, a simple opposite argument can be made as it can be suggested that the face pool test is less reliable than the face resemblance test, especially with its inherent limitations and subjectivity. Moreover, in contrast to the deductions of (Stephan and Arthur, 2006) and Stephan and Henneberg (2006), Wilkinson (2008) cited the work of Wilkinson and Whittaker (2002), who studied the reliability of the face resemblance ratings compared to the face pool tests. In the latter study, assessors performed face pool tests first and targets' identification rates were compared to the resemblance ratings given by the same assessors later. The results showed that all the tested facial reconstructions showed a comparable accuracy between the face pool and face resemblance test results. Even more, Richard et al. (2014) argued that although the face pool line-up format is appropriate for eyewitness identification forensic cases, it does not seem suitable for assessing facial reconstructions in research.

Based on the results of the pilot study regarding the best way to design and present the faces in the subjective assessment tests, additional advantages could be benefited from the recent technologies in the field of 3D forensic facial reconstruction to improve the outcome of these subjective tests. For example, 3D editing software programs may be used to enhance the lighting, texture, and presentation of the digital facial reconstructions. In addition, the presentation of virtual reconstruction images on the Internet using virtual reality modeling language (VRML) can also be implemented. This rapid, flexible and repeatable medium offered round-the-clock international accessibility to the images together with free and full interaction by the public. Scanners can be used to capture a 360° (or frontal 180°) image of a forensic craniofacial reconstruction, which can later be converted to VRML format. The VRML image format can then be loaded onto the Internet for assessors to assess them (or for the public to view them) in three dimensions. It can also be downloaded to videotape or used for stereolithographic 3D printing (Evison and Green, 1999). For these reasons and from the conclusions of the pilot study, facial reconstructions were presented to observers in the main study of this thesis an interactive 3D formats as online surveys.

In the main part of this study, the faces reconstructed were tested via a number of previously validated subjective assessment tests. These tests were the subjective face pool identification rates of 20 cases, and the subjective face resemblance scores of the overall face (test version 1) of 30 cases. In addition, all the 30 cases examined in the main study were tested via a number of newly introduced tests. These tests were the subjective face resemblance scores of individual facial regions (the forehead, the orbital bone, the nasal bone, the cheek bone, the chin bone, and the jaw bone) (test version 2), and the calculated sum of the subjective face resemblance scores from the scores given to each individual facial region in the face resemblance test version (2). To assess assessing the accuracy of the proposed method of facial reconstruction and to validate the newly introduced tests, all cases were ranked according to each test and the ranks were statistically correlated.

The results showed that the previously validated subjective and objective tests correlated significantly with each other, which means a more reliable assessment of the accuracy of the proposed method of facial reconstruction (Section 6.3). Moreover, compared to the face resemblance test version (1), the identification rates of the face pool test in this study correlated significantly with the resemblance scores of 4/6 facial regions (All except the forehead and the chin), while the face resemblance test version (1) correlated significantly with the resemblance scores of 5/6 facial regions (All except the forehead only), based on ranks of the initial 20 cases and the 30 cases respectively. Also, the correlation of the objective SD test was more significant with the face resemblance test ($P = 0.000$) than with the face pool test ($P = 0.028$).

Although various attempts and suggestions were made so that the face pool tests mimic forensic scenarios as much as possible, they remain inexact simulations to the real scenarios. Moreover, Richard et al. (2014) argued that although the face pool line-up format is appropriate for eyewitness identification forensic cases, it does not seem suitable for assessing facial reconstructions in research. Additionally, it has been claimed that the degree of subjectivity involved in selecting the foil faces is similar to that involved in manual clay sculpting (Nelson and Michael, 1998). As well, a face pool test might include two types of biases (I and II). Type I bias is the image bias (e.g., variations in resolution or pose) that causes the observers to tend to choose one image more than another, which could be more problematic with photographs (form B) than with computer generated faces (form A). Type II bias is related to the selected foil (distractor) faces (e.g., foils that are very similar or dissimilar to the target). Thus, designing a face pool test carries a high degree of subjectivity when selecting the foil faces according to the practitioner's judgement. Although recognised, it is difficult in many cases to avoid these biases, particularly type II (Stephan and Henneberg, 2006). In contrast, as the face resemblance test entails a direct comparison between the advertised reconstructed face and another face that is known to belong to the same person, it appears closer to the real forensic cases as explained, and, thus, designing such a test does not include any interference by the practitioner.

Therefore, based on this discussion, it can be concluded that the face resemblance test is more sensitive and thus more reliable than the face pool test in the subjective assessment of facial reconstructions, regarding their design and results.

It should be noted, however, that in forensic cases, the target, whether a missing person or a suspect, is usually recognised by individuals who are familiar with the target. In contrast, in forensic facial reconstruction research, in both types of subjective tests (face pool and face resemblance), the observers who assess the facial reconstructions are usually unfamiliar with the targets, even with simulated forensic scenarios. This is possibly because the brain response and neuronal activities for familiar faces differ from those for unfamiliar faces (Caharel et al., 2011, Eifuku et al., 2011). Some theories have been suggested to explain the mechanisms by which familiarity enhances the ability to match distinct face pictures of the same person. For example, recognising a familiar face is associated with specific person's information (e.g., occupation, name) with knowledge about face identity. This creates an indirect reinforced association between two facial images of the same person, with more confidence in matching or discriminating familiar faces, which makes recognising and matching familiar faces easier and faster than unfamiliar faces (Caharel et al., 2011). Therefore, such unfamiliar scenarios in facial reconstruction research studies are not true representations of forensic cases.

To overcome this, some researchers conducted studies using familiar situations in which the faces of healthy volunteers were reconstructed (Stephan et al., 2005b, Fernandes et al., 2012, Herrera et al., 2016). Stephan et al. (2005b) recruited 2 groups of assessors, unfamiliar and familiar with the target, to assess the facial reconstructions. However, both scenarios (unfamiliar and familiar scenarios) revealed a broad range of recognition success without a clear success of one scenario over the other. Moreover, Fernandes et al. (2012) used only familiar assessors to assess 3D facial reconstructions, and the identification rates did not exceed 24%. Similarly, when Herrera et al. (2016) tested the facial reconstruction of volunteers' scanned skulls by assessors who were familiar to them (students of the targets) did not seem to provide higher frequencies of correct recognitions or greater resemblance scores. Herrera et al. (2016)'s explanation was that being familiar with a person's face does not only depends on the facial appearance of the person, but also requires information about personal traits, emotions, etc. this is consistent with the psychological theories that explain face familiarity (Caharel et al., 2011). Therefore, familiarity has to be personal (i.e. related to close people like parents, children and friends) (Herrera et al., 2016).

However, in research studies, where the studied cases are either of cadavers or of living patients, it is difficult, from practical and ethical points of view, to recruit the deceased's or the patients' relatives for subjective assessment of the reconstructed faces of the studied subjects (Stephan and Henneberg, 2001, Herrera et al., 2016). Nonetheless, in the majority of research studies, subjective assessment is still performed with the aid of volunteers who are unfamiliar with the studied targets' faces. Various suggestions have been made by researchers to present the facial reconstructions in the subjective assessment tests in a way to ensure a reliable assessment by assessors who are unfamiliar with the targets. One of the objectives of the pilot study of this thesis was to investigate whether the usual format of the subjective face pool test adopted in literature, which is similar to the police line-up, can be improved for the purpose of forensic facial reconstruction research where assessors are unfamilia with the targets. Accordingly, different presentations of the facial reconstructions in these face pool tests were examined.

For example, Shahrom et al. (1996) advised to use facial templates with open eyes. Other researchers (George, 1987, Davy et al., 2005, Vanezis, 2008) suggested "humanising" the facial reconstruction with realistic facial features (e.g. from the electronic identikit system E-Fit™, FaceGen Modeller software (Face Gen®)). This was thought to improve the perceptual similarity in human observers as the human eye might be forced to a holistic (overall) view of the images rather than concentrating on isolated facial areas (Vanezis, 2008). Moreover, it was suggested that the used images should be normalised according to colour (i.e. should have the same texture on every facial surface), and pose (Claes et al., 2006, Stephan and Henneberg, 2001). Also, Vanezis (2008) referred to studies that showed that it was better not to use photographs in face pools to allow the observers to use their imagination and recognition skills while assessing the facial reconstructions by avoiding using images produced from "photographic" segments results in higher recognition. The viewer usually looks at the face in general, which may trigger recognition as a whole without specific characteristics. A photograph-like image of a person may not trigger such a response. Even more, using artistic sketches rather than Identikit composites can result in higher identification rates (Nelson and Michael, 1998, Vanezis, 2008).

In contrast, in the pilot study of this thesis, a comparison was conducted between face pool tests form (A) and (B). The results showed that using computer generated or scanned faces with closed eyes, no hair, and similar complexion (i.e. face pool form A) were better than the face pool tests formed of photographs with facial features (e.g., eye shape and colour, hair style and colour and skin colour) (i.e. face pool form B). These finding were consistent with Nelson and Michael (1998) who recommended to keep the facial reconstructions as 'undefined' as possible so that the memory is triggered by the form of the face, rather than by specific features. Likewise, Wilkinson et al. (2006) believed that the success of their facial reconstruction method over other methods was due to the way they presented the facial reconstructions in the face pool tests as neutral-colored, non-realistic models by avoiding the application of hair, and skin colors, and texture, so that the observers can concentrate on comparing the faces by relying on the shape of the skull underneath without being distracted by these additional features. Further, facial reconstructions without hair received slightly higher resemblance scores than those with hair in Stephan and Henneberg (2006)'s study.

It should be noted, however, that a database of photographs is easier to establish, from various websites or with a regular digital camera, than the databased of computer generated or scanned faces. Moreover, the collection of the scanned faces (e.g., by a laser scanner, CT) entails ethical considerations due to the possibility of associated radiation hazards. However, this can be overcome by a collaboration between different organisations (e.g., law enforcement authorities and medical institutions) to collect data already obtained for medical purposes to minimise the associated risks and be able to construct a larger database (Moyers, 2007, Shimofusa et al., 2009).

In addition, the pilot study of this thesis demonstrated that a face with a neutral expression was easier to assess than that with an expressed smile. This highlights the influence of poses and facial expressions in the photographs, presented to the assessors, on the identification rates (Stephan and Henneberg, 2001). So, the used photographs are better to be with neutral facial expressions (e.g., no grinning, frowning or raised eyebrows), and with no glasses, (i.e., a passport-like photograph). This is consistent with Snow et al. (1970)'s statement "…unretouched police photographs taken in a standard manner with little variation in lighting, expression, and pose, might make the comparison between faces easier and thus allows correct identification". Furthermore, it is better to use faces with no background, make up, or head or facial hair as this could not be determined from the skull (Claes et al., 2006, Stephan and Henneberg, 2001). Therefore, as the 3D computer generated or scanned faces usually have the same texture and pose and usually possess no facial expression, they constitute a more reliable alternative to photographs in research studies, where applicable, which add further to the advantages of these faces over photographs.

Richard et al. (2014) compared the recognition rates between 5 different presentations of facial reconstructions to the assessors. These included a "Basic" presentation (a single frontal image of the reconstruction without any adjustments for weight or age), a "Front and Profile" presentation, a "Weight Variation" presentation (including three frontal images of the reconstructions at thin, unadjusted average, and heavy weights), "Estimated Average Age" (for that demographic group), and "Estimated Age Range" presentation consisted of three reconstructions: mean age, 10 years younger, and 10 years older than mean age. The results showed that the "Front and Profile", "Weight Variation", and "Estimated Age Range" presentations had comparable high sensitivity. However, the "Front and Profile" presentation showed high values for both sensitivity and specificity, and it had the only statistically significant success rate. In addition, this latter presentation can be easily implemented in presenting both manual and computerised facial reconstructions. Moreover, the individual presentations were also useful when individuals' features were deviated from the "average". For example, the "Front and Profile" presentation helped recognising the target with an underbite. Furthermore, the "Weight Variation", and presentation was more useful when the target was thinner than the average weight for their group. Likewise, the "Estimated Age Range" presentation helped identifying individuals with 10 or 20 years difference from the average-aged facial reconstruction. This is because the sample of the studied individuals reflected a balanced physiological variation as is expected in a group of missing persons. Therefore, the authors concluded that more consistent recognition of the targets can be reached via na single presentation combining different presentations. Also, the authors recommended that it was more favourable to present facial reconstructions to the public in multiple images of different variations. Moreover, it was suggested that the correct recognition from a face pool test can be helped by using more than one view, particularly the "three-quarter" view, of the target and foil faces (e.g. frontal and profile views) (Wilkinson et al., 2006). Similarly, in the pilot study of this thesis, multiple views of the facial image of the same target allowed more reliable assessment of the facial reconstructions and thus a higher chance of correct identification as these views familiarised the observers with the target's face shape than the facial images in frontal views only. It was, then, concluded that multiple orientations of the target images improve the identification rates, which is consistent with Richard et al. (2014).

Moreover, there is an inherent limitation in the subjective assessment tests as they rely on the subjects who perform it. Variations related to the observers' individual characteristics (e.g. age, race, sex, related professional experience, etc.) were thought to influence their performance in these subjective tests, particularly in the face pool tests. Therefore, many researchers analysed the results of the subjective face pool tests taking into account the between-assessor/observer variations (Snow et al., 1970, Moyers, 2007, Herrera et al., 2016).

In the pilot study of this thesis, the relationship between the assessors' characteristics (sex, age and race) and the correct identification rates in the face pool tests showed some interesting observations. For example, although not statistically significant, female participants performed better than male participants in facial identification. This is consistent with some studies (Snow et al., 1970). However, in the main study, there was no significant association between the higher identification rates and the sex of the participant, which was consistent with other studies (Moyers, 2007, Herrera et al., 2016).

Furthermore, this pilot study showed that the Caucasian observers showed the least identification rates of the examined Caucasian faces. However, this was statistically different from the face pool identification rates by observers belonging to other races (Asian, African, and Mixed races). However, as the African participants in this pilot study represented only 3.5% of the study sample, their performance should be further investigated to confirm this conclusion. In addition, the identification of Asian and African reconstructed faces by observers from the same and other races should also be investigated. Moreover, in the main study, the Egyptian observers were not associated with significant high identification rates of the studied Egyptian cases. Even more, the identification rates by the Egyptian observers was lower than that of all participants combined. Furthermore, although the resemblance scores given by the Egyptian group of observers was higher than that of all participants group combined (62% compared to 49%), there was a disagreement within the Egyptian participants in the ranking of the studied cases, according to the overall resemblance scores. In contrast, the forensic anthropology expert group showed the highest inter-observer agreement, and their cases' ranks did not statistically correlate with that of the Egyptian participants. Therefore, the resemblance scores given by the Egyptian group can be considered less reliable than other groups, including the non-expert non-Egyptian participants. One explanation for this can be that the latter group were mostly recruited from London, UK. Having used to the diversity in the London population, this group might have been able to distinguish faces better than the Egyptian participants who are used to a more uniform population. It can, thus, be shown from the results of both the pilot and the main studies that there is no indication that a group of observers of a certain population would perform better in tests with targets' faces of the same race as the observers. Surprisingly, this contradicts the suggestion of a racial bias in face identification, where it is believed that the own-race faces are recognised more accurately than other-race faces (Goldinger et al., 2009).

On the other hand, in the pilot study, the distinction between the identification rates by observers of different age groups was more obvious. For instance, observers < 30 years old seemed to be able to match the faces of unknown individuals significantly better than those aging 30 – 49 years old. However, further investigation is needed by studying cases and observers of other age groups, especially that observers aged 50 – 69 years old represented 2.6 % of the total study sample. From these preliminary findings, it can, therefore, be suggested that observers < 30 years old represent an ideal group of observers for subjective assessment of forensic facial reconstructions by face pool tests. However, in Snow et al. (1970)'s study, the assessors' age was not found to be a significant factor in influencing correctness of choice among civilians. Interestingly, the main study results also showed that young participants have a 1-2 % higher chance of significantly higher correct identification rates than older participants. However, the age threshold for that could not be determined.

Furthermore, the assessors' professional experience is one of the individual characteristics, which was investigated in many studies and yielded controversial findings (Snow et al., 1970, Moyers, 2007, Herrera et al., 2016). For example, Herrera et al. (2016) conducted a study to investigate the performance of four facial soft tissue thicknesses (FSTT) datasets of Brazilian population. The authors assessed 16 faces reconstructed according to the manual American method by 120 participants using both types of subjective assessment tests (face pool and face resemblance tests). The influence of the assessors' sex and knowledge of Human Anatomy and Forensic Dentistry on recognising people was investigated. However, these features did not seem to play a determinant role to reach greater recognition rates.

On the other hand, Snow et al. (1970) classified the assessors in face pool tests into male and female groups, and into civilians and policemen groups. The policemen were generally more experienced in identification. In one case, the policemen and civilian females scored slightly better, but not statistically significant, than civilian males. However, in the second case, the policemen scored significantly better than the civilian males, and the civilian females scored better than their male counterparts, which is consistent with the results of this pilot study. However, there was no a significant difference in the number of years of experience among the policemen who correctly identified the reconstructions and those who did not (Snow et al., 1970). Further, it was observed that facial imagery (facial mapping) experts performed consistently better than members of the public when attempting to identify faces from CCTV footage (Wilkinson and Evans, 2009, Wilkinson and Evans, 2011). In these studies, the public showed high false acceptance rates (FAR) and low false rejection rates (FRR) of the target faces. Also, the error rate increased when the targets wore hats. This demonstrates the higher tendency of a jury (drawn from the public) to accept an innocent person than to reject a guilty person based on CCTV identification evidence. Therefore, Wilkinson and Evans (2009) and Wilkinson and Evans (2011) supported the conclusion that the skills, knowledge and abilities of CCTV facial imagery experts would be needed to assist the jury to reach reliable conclusions.

In a similar way, it might be suggested that observers with a professional experience in the "human face" identification and/or forensic facial reconstruction would be better performing in face pool tests than non-experts (Vanezis, 2008), although not yet investigated in previous facial reconstruction studies. Consequently, in the main part of this study, after completing the face pool tests of 20 cases, a number of participants from the non-expert group agreed to repeat a number of the face pool tests without revealing the targets' faces to them after the first attempts. The results showed that the proportions of the correct answers in the second attempt were higher than the first attempt. However, the statistical difference between the two proportions was marginally insignificant (P = 0.086). Additionally, the following stage (the face resemblance tests stage) included a number of non-expert observers who had initially participated in the face pool stage. This group was referred to as the "old group". Whereas participants who had not performed the face pool tests before participating in the face resemblance stage were referred to as the “new group”. Analysis of the face resemblance tests between the old and the new groups showed significantly higher resemblance scores in the old group compared to the new group, in addition to an inter-observer agreement in all cases assessed by the old group compared to an inter-observer disagreement in only half of the assessed cases by the new group.

Furthermore, all participants were classified according to their professional experience into non-experts, experts of facial identification psychology, forensic pathology, and forensic anthropology (with or without facial reconstruction experience). In the face pool tests, the ranking of the 20 cases according to the identification rates of all participants groups correlated significantly. However, comparing the performances of the expert groups with each other showed that the mean identification rate and the number of cases correctly identified by the forensic pathology experts were less than that of the non-expert group. In contrast, the mean identification rates and the number of cases correctly identified were higher than that of the non-experts and the forensic pathology experts. Moreover, there was no significant association between the high identification rates and a professional experience in forensic medicine/pathology nor facial identification/perception psychology. However, the small number of the recruited facial identification psychology experts (n=3) compared to other expert groups should be taken into consideration. On the other hand, a significant association was found between the high identification rates and a professional experience in forensic anthropology. This association markedly increased when the forensic anthropology experts also had experience forensic facial reconstruction. Furthermore, the mean resemblance rating given to the 30 cases by the forensic anthropology experts was higher than that given by all participants combined, with the highest agreement between the observers in the given resemblance scores found within the forensic anthropology experts compared to all other experts and no-experts groups.

It can, therefore, be concluded that professional and practical experiences in studying the human skeletal features (forensic anthropology) tend to improve the ability of an observer to correctly select the target face based on a facial reconstruction of the target only. On the other hand, this ability is not improved by an experience in the human anatomy (represented by the forensic pathologists), or the facial identification psychology. This might be because the forensic anthropology experts, particularly with facial reconstruction experience, can acknowledge the strengths and weaknesses of the facial reconstruction techniques. Thus, they look for the appropriate anatomical features that allow a reliable comparison between the real and reconstructed faces. In contrast, although instructed about what to look for in this study, the non-expert individuals would rely on less reliable features for selecting the target. However, as observed in this study, there is a learning curve in the subjective assessment of forensic facial reconstruction among the non-experts group which led to an improvement in the way the observers studied the faces shapes and hence higher identification rates when they repeated the face pool tests of the same cases. It might, therefore, be worthy to openly educate the public about what the capabilities and limitations of the facial reconstruction methods to achieve the best recognition rates possible (Stephan and Henneberg, 2006, Vanezis, 2008).

6.2.2 Objective assessment methods

Objective assessment of the facial reconstructions can follow a number of approaches, which are either landmark-independent or landmark-based. In the landmark-independent objective assessment, the assessment is made via defining points on the compared images, which is needed for registration and comparison between the reconstructed and the real faces. An example of this technique is image superimposition (Shahrom et al., 1996, Vanezis and Brierley, 1996, Curry et al., 2001, Jayaratne et al., 2012), which compares between the two faces at a number of corresponding anatomical landmarks, taking into account the tissue thicknesses and the general morphology. Corresponding features are then compared to find matches or differences. If the images are of the same individual, then the anatomical features should align accurately (Abate et al., 2004).

Moreover, the landmark-independent objective assessment of the 3D facial reconstructions can also be performed via mathematical surface distance comparison between the aligned 3D target's reconstructed and real faces (Claes et al., 2006, Vandermeulen et al., 2006, Wilkinson et al., 2006, Lee et al., 2012, Short et al., 2014, Decker et al., 2013). As well, this method requires initial registration between the compared surfaces performed via alignment, which is a point-based registration algorithm that defines a number of homologous points that sparsely selected on different facial regions of both surfaces (Wilkinson et al., 2006, Lee et al., 2012, Short et al., 2014, Decker et al., 2013). Then, the Euclidean Distance (i.e. the shortest distance between two points/landmarks irrespective of the direction) is calculated (Short et al., 2014). It is used as an objective indicator of the degree of closenessbetween the reconstructed and the target's real face. Not only the surface distance comparison provides an evaluation of the accuracy of the facial reconstructions, but also it provides a spatial map of the goodness of fit between the overall faces and at individual facial regions. The lower the surface difference, the higher the closeness, and, hence, the accuracy of the compared faces (Wilkinson et al., 2006). There is a number of mathematical examples for calculating the surface distance differences, including; Euclidean Distance Matrix (EDM) descriptors (Claes et al., 2006, Vandermeulen et al., 2006), Sum of Square Differences (SSD) (Vandermeulen et al., 2006), Surface Deviation (Wilkinson et al., 2006, Lee et al., 2012, Short et al., 2014, Decker et al., 2013), and Root Mean Square (RMS) (Jayaratne et al., 2012).

Software programs have been designed for surface distance differences assessment via surface deviation algorith. For example, Wilkinson et al. (2006) designed a software program (reverse modelling software Rapidform™ 2004 PP2 (@ INUS Technology Inc, Seoul, Korea)–RF4) for the purpose of digitising the combined "Manchester" facial reconstruction method, and for quantitative statistical comparison of the surface distance between the facial reconstructions and the real facial scans of the studied targets. The RF4 software allows the production of computed histogram plots and color maps for visual representation of the degree of discrepancies between the two surfaces. Each map consists of different colors indicating the distribution of errors across the face and numbers in mm corresponding to the surface deviation between the two surfaces, the "+" and the "-" numbers mean that the skin surface of the reconstruction was more and less prominent than the subject face respectively. Furthermore, visual assessment can be shown via the same software in a colour maps (histogram) that quantitatively represents the differences between the registered surfaces. Other studies presented similar methods involving similar sofware programs (Claes et al., 2006, Vandermeulen et al., 2006, ATOR, 2012).

On the other hand, the landmark-based objective comparison of the facial surfaces is dependent on the calculating the distances between surface landmarks. This approach includes a numer of techniques, such as; Procrustes Shape Analysis (Wilkinson, 2008, Short et al., 2014). Vanezis (2008) conducted an experiment assessing the mathematical significance of the facial reconstructions via full Ordinary Procrustes (Shape) Analysis. The aim of this experiment was to attempt to refine the choice of facial template(s) by excluding templates of extreme shapes by correlating this objective method results with the subjective resemblance ranking of the same faces. However, this correlation was statistically insignificant, and as the author explained, this could be because the mathematical Procrustes Analysis was based on a more holistic (overall) matching, rather than assessing isolated facial features, which might have been the basis of the subjective resemblance rating assessment.

Furthermore, geometrical morphometric analysis via craniofacial anthropometric measurements is another method of landmark-based objective assessment of the facial reconstructions (Jedrzejowska, 2001, Kleinberg and Vanezis, 2007, Kleinberg et al., 2007, Vanezis, 2008, Short et al., 2014, Hayes, 2016). It is a method of statistical shape analysis comparing the patterns of shape variance across a group of homologous landmarks (Hayes, 2016). Although the landmarks' coordinates retain all the geometrical information, including locations and orientations, this method requires reference images for comparison (Islam et al., 2015, Hayes, 2016). Cranial anthropometric mathematical calculations (craniometrics) was previously used for human identification purposes (Kleinberg and Vanezis, 2007, Jedrzejowska, 2001, Starbuck and Ward, 2007, Vanezis, 2008, Short et al., 2014, Hayes, 2016).

Kleinberg and Vanezis (2007) conducted an experiment of facial image comparison for facial identification from Closed-circuit television (CCTV) images. A number of facial landmarks were defined, and proportion indices (PI) were calculated for between-landmark lines as follows: PI = Numerator (lower value)/Denominator (higher value) x

100, in addition to angles between the landmarks. With inconsistent results, this study concluded that facial anthropometry, in the way it was applied, failed as a facial identification technique even when photos were taken from high quality surveillance footages. In addition, out of the measured landmarks, no one landmark made significantly better comparison than another. On the other hand, cranial measurements were used for the determination of a viscero-cranium profile from various skull segments and the different angles between them (Jedrzejowska, 2001). This allowed viscero-cranium reconstruction based on craniometry by identifying a significant mathematical correlation between the soft parts of the head (i.e. the face) and the cranial osseous structure, which is useful for human identification

Cranifacial anthropometry was used in different studies as an objective method for the assessment of the accuracy of facial reconstruction (Starbuck and Ward, 2007, Vanezis, 2008, Short et al., 2014, Hayes, 2016). With only a few exceptions (Starbuck and Ward, 2007, Vanezis, 2008), it was not validated (i.e. statistically correlated) with the subjective and other objective methods to confirm or decline its reliability. Starbuck and Ward (2007) used an anthropometric craniofacial assessment to quantitatively/objectively assess the facial reconstructions, and compared it to subjective assessment of the reconstructed faces by the face resemblance test using the targets' photographs. Similarly, Vanezis (2008) attempted to employ anthropometric comparison as an indication of the similarity between the reconstructed image and the photograph of each case, using distances between the landmark, as well as the corresponding differences in their proportion indices. The results of quantitative assessment in the above studies were not consistent with the subjective assessment.

Furthermore, Short et al. (2014) compared linear and angular measuremets taken from both the reconstructed and real faces as an indicator of the accuracy of facial reconstructions in comparison with the real faces. While there were no statistical differences in the linear and angular measurements between the reconstruction and the target, some anatomical regions of the reconstruction were smaller, or larger, than the taerget (i.e. underestimated or overestimated respectiely). The nose and the mouth were consistently statistically significantly overestimated. The authors, also, observed differences in the angular and linear measurements accuracy between class II and class III skeletal patterns (Pithon et al., 2014, Utsuno et al., 2014), especially at the naso-labial angle, which suggests that certain measuremens were less reliable than others as a measure of accuracy of the facial reconstructions. This suggests that further investigation and validation would be required.

Moreover, after the face of a young woman was reconstructed using predominantly verified methods (Hayes, 2014), and the identity was confirmed later by methods other than facial reconstruction (Hayes, 2016), applied geometric morphometrics was used to assess the accuracy of the reconstructed face. This was performed by comparing the reconstructed face with three antemortem photographs of the victim and a database of 64 images of matched sex, age, head pose and population. The geometric morphometrics showed a significant resemblance when most of the variance due to depicted head pose was removed from the analyses. However, this objective analysis highlighted the subjectively noticeable differences between the 2 faces, which could explain why the advertised reconstructed face could not lead to confimred identification. The authors attributed this discrepancy to the used facial reconstruction methods' limitations in their predictive accuracy or misapplication, the practitioner errors, as well as the photographic distortions. In addition, the authors recommended that different methods need to be verified to reconstruct individual facial features. However, Hayes (2016)'s study showed the usefulness of geometric morphometry in assessing the morphological accuracy of a forensic facial reconstruction. However, in addition to the limits of the 2D photographic distortion, the described technique required more than one antemortem image, as well as a comparative database compiled of sex, age, population and head pose matched images.

As discussed above, most of the previously used anthropometrics methods of assessing the facial reconstructions relied on 2D images (of the faial reconstruction or the antemortem photpgraphs. In 2D images it is difficult to obtain the exact same viewpoints and magnifications in the aligned images that will always prevent accurate metric assessment. So, matching the landmark lines between images was limited, which might explain the failure of anthropometry for the intended purpose in these studies. Thus, each individual's photograph should have the same head orientation and facial expression, with standardised camera angle and lens-subject distance (Kleinberg and Vanezis, 2007, Kleinberg et al., 2007). This might reduce, but not entirely remove, the photographic distortion. To the best of our knowledge, no studies have used anthropometry for objective assessment of facial reconstruction in three dimensions (3D).

In the main part of this study, the faces reconstructed were tested via the previously validated objective facial surface distance standard deviation (SD) of the overall face (test version 1) of 30 cases. In addition, all the 30 cases examined in the main study were tested via a number of newly introduced objective methods. The first method was the objective facial surface distance standard deviation (SD) of the individual facial regions (the forehead, the orbital bone, the nasal bone, the cheek bone, the chin bone, and the jaw bone) (test version 2). The second method was the calculated sum of the objective facial surface distance standard deviation (SD) of the individual facial regions in test version (2). The third method was the objective craniofacial anthropometry, in which linear measurements were taken from the skull and the real and reconstructed faces and linear ratios and angles were calculated. Then, the differences between the skulls and the real and the reconstructed faces, and then between the real and the reconstructed faces at these linear ratios and angles were used as an indicator of the accuracy of the reconstructed face.

Further, to assess assessing the accuracy of the proposed method of facial reconstruction and to validate the newly introduced tests, all cases were ranked according to each test and the ranks were statistically correlated. Based on the results, the newly introduced calculated sum of the subjective face resemblance scores from the scores given to each individual facial region in the face resemblance test version (2) correlated significantly with the previously validated subjective and objective tests. Moreover, the newly introduced calculated sum of the objective facial surface distance standard deviation (SD) of the individual facial regions in test version (2) correlated significantly only with the face resemblance test version (1), out of the previously validated tests.

In addition, the newly introduced craniofacial anthropometric linear ratios and angles differences between the skulls and the real faces significantly correlated with the differences between the skulls and reconstructed faces. This showed that the reconstructed faces were proportionate to the real faces, which was in favour of the facial reconstruction method used in this study (in terms of the landmarks placement, the used facial depths and the facial templates). Further, the average angles differences correlated significantly with all the previously validated subjective and objective methods. In contrast, the average linear ratios differences correlated significantly only with face resemblance test version (1).

Out of all linear ratios and angles, the linear ratio AB/AD and the angle (ADB) (Figure 20) correlated significantly with the previously validated subjective tests, as well as the resemblance scores of the cheek and jaw bones. Moreover, the linear ratio AB/BD (Figure 20) correlated significantly with the assessment by chin bones, as well as the objective overall SD test. However, only ***one*** angle (AEC) (Figure 20) correlated significantly with ***all*** the previously validated subjective and objective assessment tests, as well as the newly introduced subjective resemblance score of the cheek bones. This shows that this angle has a higher sensitivity in the assessment of accuracy of facial reconstruction, which can be explained by the fact that this specific angle indicates both facial length and width at the same time. This (AEC) angle points are; (A) the right angle of the orbit/eye, (E) the lowest point of the front of the chin in the midline, and (C) the left angle of the orbit/eye.

Furthermore, of the 6 assessed individual facial regions; 3 regions correlated significantly with the objective overall surface distance (SD), 4 regions correlated significantly with the identification rates by the face pool test, and 5 regions correlated significantly with the given overall face resemblance score (test version 1). In particular, the *cheek* and the *jaw* regions repeatedly provided significant correlations between the cases' ranks when assessed via all the previously validated subjective and objective assessment tests, as well as the newly introduced objective differences at the linear ratio AB/AD and at the angle ADB. In addition, the newly introduced individual resemblance scores at the *cheek*, *chin* and the *jaw* bones correlated significantly with the previously validated overall facial surface distance standard deviation (SD). Also, the resemblance scores at the *cheek*, and *jaw* bones correlated significantly with the facial surface distance standard deviation (SD) at these 2 regions. This shows a higher sensitivity of these facial areas in predicting the accuracy of the facial reconstructions.

This can be explained by that the face appearance is determined by the underlying skull as well as the distance between the skull and the face (i.e. the facial soft tissue depths). This distance is not the same at different craniofacial anatomical parts and is related to the hypodermic fat, which is thickest in the "malar fat pad" (i.e. cheeks), followed by the "premental fat pad" (i.e. the chin), and absent in the forehead and lip zones (De Greef et al., 2009). The areas with the thickest hypodermic fat are the most difficult to reconstruct based on information from the skull only as they are more distant from the bone. For example, the areas of the largest errors in the reconstructed faces were seen at minor portions of the lateral foreheads, the endocanthi, and the majority of both cheeks, in addition to the nose. In contrast, the most accurate areas were seen at the chin, orbits, upper cheeks, some of the forehead, cranium, and mouth (Wilkinson et al., 2006, Lee et al., 2012, Jayaratne et al., 2012, Short et al., 2014). It can, therefore, be concluded that the facial areas closer to the bone could be better indicators of the accuracy of the facial reconstruction in the subjective and objective assessment. As shown in this study, these areas were particularly the areas overlying the cheek bone, the jaw bones and the chin, which is consistent with most studies.

6.3 EVALUATION OF THE PROPOSED METHOD FOR FACIAL RECONSTRUCTION

Forensic facial reconstruction is considered as a method of recognition rather than identification (Wilkinson, 2007). Even more, Stephan and Henneberg (2001) conducted a study to objectively determine the accuracy of four commonly used methods of forensic facial reconstruction; a two dimensional (2D) drawing American method; a 2D computer "FACE" assisted American method; a 3D sculpting American method; and a 3D sculpting combination method. The authors assessed 16 facial reconstructions using the identification rates of face pool test. The results showed that only one case (conducted via the 3D sculpting American method) received a significant identification rate above chance, with an overall mean identification rate of 3% above chance, a result markedly lower than many similar studies. This suggested that the facial reconstructions were rarely accurate to allow identification of a target individual above chance nor to exclude individuals to whom skeletal remains may not belong. Therefore, the authors concluded that facial reconstructions were useful for estimations only and they were highly inaccurate and unreliable forensic techniques. It should be noted, however, that Stephan and Henneberg (2001)'s study compared manual and computerised 2D facial reconstruction methods only. Whereas 3D computerised techniques are now more widely implemented and believed to be more accurate than most of the methods tested in Stephan and Henneberg (2001)'s study (Davy et al., 2005). Therefore, the conclusion drawn by the authors cannot be generalised on the weight of facial reconstruction as an identification method, especially that the authors acknowledged that a much larger sample was needed to reach higher power.

In this study, the "outside inward" sparse approach, using scanned faces as templates, was adopted. This approach has previously produced successful facial reconstructions (Shahrom et al., 1996, Vanezis, 2008). The "outside inward" approach, however, was criticised by that these facial templates imposed certain facial features onto the resulting facial reconstructions. (Andersson and Valfridsson, 2005, Jones, 2001, Nelson and Michael, 1998, Vandermeulen et al., 2006, Wilkinson et al., 2006). In addition, the practitioners of this approach were criticised as not being experienced enough and did not perform adequate analysis of the skeletal remains, thus missed important clues that might lead to identification (Wilkinson et al., 2006). As a result, a number of researchers have developed other methods to overcome as many of these limitations as possible. For example, a completely different approach of facial reconstruction (the "inside outward" approach) was developed following the "Manchester" method of manual reconstruction, by digitally building the individual facial muscles rather than imposing certain facial features of a certain facial template (ATOR, 2012, Wilkinson et al., 2006). These studies, however, are still dependent on using craniofacial landmarks at certain anatomical locations and guided by facial soft tissue depths tables. Furthermore, other researchers combined the "outside inward" and the "inside outward" approaches using a mathematically calculated full head deformable model incorporating the skull and the face with facial muscles as well as anatomical landmarks in between (Kähler et al., 2003). Without a doubt, the digital environment made building the facial muscles as a separate layer in the "inside outwards" much easier, which is markedly helped by a database of pre-modeled muscles to save more time (Wilkinson et al., 2006). However, the external facial features are still required and may need an additional software program (ATOR, 2012, Davy et al., 2005). In addition, this technique creates only one face out of the skull (Wilkinson et al., 2006). Moreover, building the facial muscles requires a comprehensive experience of the musculature anatomy as well as the relationship between the facial hard and soft tissue.

Other than the sparse technique of the "outside inward" approach, used in the present thesis, a similar but dense technique was suggested by other researchers. The advantages of the dense approach using volume deformation is that it deals with the facial tissues as one unit with no regards to certain facial soft tissue depths. Additionally, it takes into account the anatomical 'guidelines' to determine the face shape according to the shape of the skull beneath, with any idiosyncrasies and asymmetries in the face shape. However, this method has similar limitations to the sparse approach. For example, using a reference head which also have an influential effect on the final appearance of the reconstructed face. Also, a database is required and the selection of the reference head is still a subjective process (Nelson and Michael, 1998). Moreover, the dense approach is quite complex in terms of computation with difficulty in reliably calculating the deformation (Jones, 2001). Also, some of the dense approach methods have their own inherent limitations (Claes et al., 2010, Quatrehomme et al., 1997).

Moreover, in the facial reconstruction techniques that do not rely on a real face as a template, certain limitations to skin modeling exist, including age-related features, such as wrinkles for which such considerations must be made (Davy et al., 2005). On the other hand, a facial template that is age-appropriate to the estimated age of the unknown remains would contain the natural facial features, such as the age related wrinkles. To overcome the criticised bias by a single facial template in the "outside inward" sparse approach, further modifications to the sparse approach were made. For instance, it was suggested to use a statistical facial model instead of a scanned face (Claes et al., 2006, Vandermeulen et al., 2006). However, this techniques require certain mathematical and statistical expertise). Furthermore, software programs that allow automatic selection of the facial template, thus reducing the subjectivity included, were developed, such as; the FBI's facial reconstruction software RE/FACE (Reality Enhanced Facial Approximation by Computational Estimation) (Turner et al., 2005, Moyers, 2007, Parks et al., 2013). It, however, still needs a database of faces as well as facial depths data.

In the present thesis, it was suggested to use average faces. These faces were automatically generated by digitally merging a number of specific facial templates via computer software program. Therefore, it was essential to test the proposed method in terms of whether reconstructing a face of a given skull using a scanned facial template can result in correct identification at a statistically significant rate and a sufficient resemblance to the target's face. This was done subjectively and objectively, in the pilot and main parts of the present study. Although the pilot study of this thesis was not designed to assess the accuracy of facial reconstructions, the results showed that the proposed approach of facial reconstruction was successful in generating a face of an unknown individual, with correct identification above chance in 20/43 (47%) face pool tests (of different formats as discussed before). Further, in the main part of this thesis, 30 cases were studied. Of them, 20 reconstructed cases were assessed by face pool tests, which showed that 13/20 (65%) cases were correctly identified above random chance by the participants. For those 13 cases, the mean identification rate was 49% (24% above random chance). Moreover, the forensic anthropology experts could identify 16/20 (80%) cases, with 53% (28% above random chance) as a mean identification rate of these 16 cases. Furthermore, all the 30 cases were assessed via the subjective face resemblance tests using a numerical (0 – 10) rating scale. This included giving an overall face resemblance scores (face resemblance test version 1), then, as scores to individual facial regions (1-Forehead, 2- Orbits, 3- Cheek Bone, 4- Chin, 5- Jaw) (face resemblance test version 2). The mean resemblance scores given to the 30 cases was 45% by all participants, rising to 49% by the forensic anthropology experts (with the highest between-observer agreement), and to 62% by the non-expert Egyptian group (but with between-observer agreement in half the cases only). Moreover, the results of the objective assessment of the 30 cases via the overall surface distance standard deviation between the real and the reconstructive faces in this study showed that the surface differences ranged from 1.95 - 6.33 mm, with a mean difference of 3.39 mm in all cases.

Subsequently, the proposed method in this thesis was evaluated in comparison with a number of the commonly cited studies adopting other approaches of forensic facial reconstruction in the light of the accuracy assessment as well as the advantages and disadvantages of each method. In an example of the *"Inside Outwards"* approach, Snow et al. (1970) presented four *manually* reconstructed faces; the resemblance of two of which was subjectively assessed by the acquaintances of each case resulting in a reasonable likeness in one case, and expressed reservations in the other. The other two facial reconstructions, of a young white male and an elderly white female respectively, were assessed via face pool tests by volunteer assessors. The identification rates were 54% above chance and 11% above chance of the male and female cases respectively. The latter results were thought to be due to the 25 years age difference between the individual at death and the available ante-mortem photograph.

Another example of the *"Inside Outwards"* approach was presented by Wilkinson et al. (2006), where *computerised* facial reconstruction resulted in a 50% correct identification percentage above chance. However, this was only conducted on two facial reconstructions (one male and one female), with the identification rate slightly higher in the female than in the male cases. However, when the authors validated the face pool design for both cases by another group of volunteers, 80% of the volunteers thought the target face from the face pool stood out from other faces in the female case, which may have accounted for the higher identification rate of that case. The same approach was further assessed objectively by surface distance comparison via surface deviation (Wilkinson et al., 2006, Lee et al., 2012, Short et al., 2014, Decker et al., 2013). In Wilkinson et al. (2006)'s study, the surface distance comparison showed that 54 - 62% of the reconstructed were within a ± 2.5 mm, and deviation error, and 75 - 90% were within a ±5.0 mm deviation error. In Lee et al. (2012)'s study, the surface distance comparison showed that 54 - 77% of the facial reconstruction surfaces had ± 2.5 mm deviation error, and 88 - 97% had deviation error of ± 5 mm. In Short et al. (2014)'s study, 56% - 90% of the compared faces lied within an error of ± 2.5 mm deviation. On comparison, in the main study of the present thesis, 25/30 cases (83%) showed a surface distance within a ± 5 mm, 7/30 cases (23%) showed a surface difference within a ± 2.5 mm, and 18/30 cases (60%) showed a surface difference between 2.5 - 5 mm, and 5/30 cases (17%) showed a surface difference more than ± 5 mm.

On the other hand, Claes et al. (2006) used face pool tests to compare the faces reconstructed via the *"outside inward"; sparse approach* using a *statistically deformable facial model*, in contrast to a *scanned facial template* in the present study. The overall average identification rate in the face pool tests was over 81%. However, the presented identification percentages of Claes et al. (2006)'s statistical facial model should be corrected to "above chance" as this sets the statistical significance of each percentage. With no indication about this percentage, it was not possible to compare Claes et al. (2006)'s study to the present study. Moreover, the authors calculated the Euclidean distance matrix (EDM) signatures of 118 facial reconstruction and correlated them with EDM signatures of the 118 corresponding original faces. The authors claimed that the face with the most similar EDM was found to correspond to the identification, which was used as an indicator of the quality of the reconstructions in terms of face recognition and identification. However, this validation was incomplete as there was no a 3D database of missing persons to compare the facial reconstructions with, as the authors acknowledged. Presented a similar statistical head models using average reference heads with calculation of the Sum of Square Distances (SSD) as an objective measure for the accuracy of the reconstructions. Although the above methods are automatic, the lack of anatomical guidance can produce errors, especially that the aligning of the target and reference skulls was non-linear (i.e. with no need to geometrically align the two surfaces).

Moreover, Moyers (2007) conducted a validation study for the FBI's facial reconstruction software RE/FACE (Reality Enhanced Facial Approximation by Computational Estimation), which adopts the *"Outside Inwards"; dense approach* of facial reconstruction using a *landmark-independent craniofacial template (CFT).* Although 9/10 (90%) cases were correctly identified above random chance, the mean identification rate for all non-expert subjects was 10% above random chance. In comparison, in the present study, the non-expert participants identified 13/20 (65%) cases but with an identification rate of 13% above random chance. However, the Moyers (2007) conducted a face resemblance test for the 9 subjects that received significant identification rates via a numerical and descriptive scale, in contrast to the numerical only scale used in this study, so it was not possible to compare between the 2 studies from the results of the face resemblance tests.

Another animated automatic approach that combine features of both the inside outwards and the outside inwards approaches was developed by Kähler et al. (2003). It involved a statistically deformation technique applied to a triangular mesh used as a template selected from a database of faces, in addition to building the facial muscles of expression only, guided by a number of landmarks at certain tissue thickness values. Thus, this animated head model provided various faces with different facial expressions, rather than just the neutral face produced by manual or other approaches of facial reconstruction. This facial animation would add a lively appearance to the face and helps better recognition of the face. This also allows comparing the reconstructed face to more ante mortem photographs with different facial expressions for a more reliable identification. Moreover, this facial reconstruction produced a full head model, with relatively few landmarks, in a short time, and allowed modifying the muscle mass if more variations of the face shapes were needed (e.g., slim, obese, etc.). The problem, however, with this approach was that it worked better with normal shaped skulls, while unusual skeletal features resulted in very sparse sampling of the unusual area. In addition, this approach involved a considerable amount of interpolation and heuristic additions to the reconstructed face. This was performed to overcome the drawbacks of using the facial soft tissue depths collected by Rhine and Campbell 1980 and Rhine and Moore 1984 (Kähler et al., 2003), which consisted of a number of landmarks that does not cover the full face and skull surfaces. This problem, however, made the technique reliant on the practitioner's experience and judgement, which does not make the approach fully automated as presented by the authors.

The present thesis showed that a facial template warped onto the skull could be modified to take the shape of the underlying skull and generate a face of sufficient similarity to the target's face that allows the identification of the target. The proposed method showed comparable accuracy to many of the other facial reconstruction techniques. While it is important to validate different facial reconstruction methods in comparison to each other, it is more important to place the accuracy assessment results into the correct perspective. This is taking into account the degree of the reliability of the accuracy assessment method (as discussed in the previous section) as well as interpretation of the results of these methods within the context of forensic scenarios. For example, Stephan and Henneberg (2006) stated that with "successful" facial reconstructions in real forensic cases, only one or a few individuals might come forwards claiming recognition, which may not be as a result of specific and purposeful facial identification. To incorporate this fact in research, the authors suggested that if a facial reconstruction method is able to offer more than 50% correct identification, it is guaranteed to provoke correct recognition responses. It could, therefore, be expected that the low rate of correct identification of a facial reconstruction does not mean that this facial reconstruction would not be successfully identified and recognised, and the chance of recognising of a facial reconstruction in a forensic environment could be much higher than was actually received in a research study. Even statistically insignificant recognition rates in certain studies may become significant if a larger sample size was used (Stephan and Henneberg, 2006).

Conversely, a face receiving high resemblance in a research study might have a weak ability to provoke recognition. Stephan and Henneberg (2006) presented an example of a forensic case where the advertised facial reconstruction was not successful in reaching identification although, later, was thought to carry a sufficient resemblance to the target when the identity was revealed by other means. Similarly, when the identity of the final 1987 King's Cross fire victim was confirmed 16 years later by means of medical records, his face that was previously reconstructed was believed to be of reasonable resemblance to the target (British Transport Police, The Guardian, 2004). Another example of the controversy around the reliability of facial reconstruction techniques was cited by Davy et al. (2005). The authors referred to the work of Haglund and Reay (1991), which tested the reproducibility of manual facial reconstruction techniques by multiple facial reconstruction practitioners who reconstructed the faces of several victims of the Green River serial killer in the 1980s and showed marked variability. In addition, Decker et al. (2013) showed a discrepancy between faces reconstructed from the same skull by different experienced practitioners applying manual and computerised facial reconstruction techniques (Decker et al., 2013). Furthermore, failed facial reconstructions are thought to be much more than those published in the literature as successful cases are usually given more attention, while many failures go unreported (Stephan and Henneberg, 2006, Stephan and Henneberg, 2001), attributed to the publication bias (Song et al., 2013).

As a result, it is believed that the success of the facial reconstructions in real forensic cases depends only to a small degree on the anatomical resemblance between the reconstructed and the real face (Phillips, 2001, Stephan and Henneberg, 2001, Stephan and Henneberg, 2006). Rather, it relies on the other factors, such as; the broadness and timing of the media coverage, who sees the advertisement, the presence of other assisting information advertised along with the facial reconstruction, etc. (Stephan and Henneberg, 2006). It can, therefore, be concluded that the identification rates presented by research studies do not truly reflect the recognisability of the facial reconstruction in real forensic scenarios. In particular, the identification rates of the face pool tests in research studies should not be considered as a reliable indicator of the accuracy of the facial reconstructions nor as a method of comparison between different facial reconstruction techniques. This is, also, because of the inevitable non-standardised design of the face pool tests, and the subjectivity included in the selection of the foil faces (Section 6.2.1).

Although the accuracy of the present method is comparable to many other methods and maybe higher than others, there are some limitations which might have led to lower accuracy than certain studies, which could be attributed to a number of factors. For example, many of the differences between the facial reconstructions and the real faces could be related to the different scanning positions and devices between the studied cases (supine position via a multidetector computed tomography) and the faces used to obtain the facial soft tissue depths in El-Mehallawi and Soliman (2001)'s study (upright position via ultrasound). The gravity has different effects on the face in both procedures. For example, it causes the cheek and mouth areas of the face to sag downward when the subject lies on his or her back, in contrast to fuller cheeks and a less stretched mouth while sitting upright (Wilkinson et al., 2006). Moreover, it was found by some researchers that facial soft tissue depths of modern populations may be different from those collected by former researchers possibly due to the increased rates of obesity in modern populations (Parks et al., 2014). Similarly, as (El-Mehallawi and Soliman, 2001) set was published in

2001 (14-15 years before being used in this study), changes in the facial soft tissue depths of a demographically similar population could, to some degree, have influenced the resulting facial reconstructions.

More importantly, as no similar studies were previously performed on the Egyptian population, the facial soft tissue depths set used in this study (El-Mehallawi and Soliman, 2001) was not previously validated. Thus, it had to be modified as appropriate, with a trial and error approach followed in this study to perform the reconstructions. So, while it would have been closer to forensic scenarios to perform the facial reconstructions under blind conditions, it was inevitable to do so in this study in privy (i.e. with access to the real faces of most cases). For starters, the landmarks in El-Mehallawi and Soliman (2001)'s study were described in relation to the face with no corresponding cranial definitions. Therefore, for accurate positioning of the landmarks on the face as well as the cranium, the definitions of the corresponding cranial landmarks were sought from additional studies (Brown et al., 2004, De Greef et al., 2005, Cha, 2013). Also, El-Mehallawi and Soliman (2001)'s defined landmarks were bilateral only with no midline landmarks. Therefore, fitting the facial templates on the skulls, using these bilateral landmarks only, resulted in noticeable defects in the reconstructed face. Parts of the skull were bare from the overlying face, and parts of the face, mainly in the middle regions, were overstretched on the skull. Therefore, following Parks et al. (2014), to account for the lack of the midline landmarks in El-Mehallawi and Soliman (2001)'s study, a modified set of landmarks and their depths was prepared. This modified set was formed of a combination between the bilateral El-Mehallawi and Soliman (2001)'s landmarks as well as the midline landmarks of the Caucasian population from Rhine and Moore (1982) and Helmer (1984)'s set. However, reconstructing the Egyptian faces using this modified set with an average European facial templates was not successful in achieving sufficient resemblance to the targets. Whereas, using average Egyptian facial templates resulted in a satisfactory resemblance to the targets. Moreover, it was essential to continually compare the resulting facial reconstructions with the target in order to achieve a satisfactory resemblance to the targets to proceed with the following stage of the study (assessment of the facial reconstructions).

From a practical point of view, the process of facial reconstruction is not expected to generate an exact replica of the unknown face, but to produce a face of a sufficient resemblance to draw the public attention for possible identification (Phillips, 2001, Andersson and Valfridsson, 2005). As a result, facial reconstruction can be rather seen as a screening or preliminary method of identification. Thus, one method can be advantageous over another, if, for instance, it is easily transferrable from one practitioner to another with accurate reproducibility (Shahrom et al., 1996), in addition to being simple, quickly applied, and as least subjective as possible. As described above, certain forensic facial reconstruction methods require a specialised expertise, which leads to that only knowledgeable experts can perform such methods, especially within a law enforcement agency. Many agencies, however, refrain from transporting the evidence (i.e. the remains), and restrict it to their local experts (Decker et al., 2013), which makes the facial reconstruction inapplicable in many potential cases.

Therefore, the pilot study of this thesis also investigated the applicability of the proposed method. With the help of a user manual, a user with no experience in the field of 3D facial reconstruction was able to perform the facial reconstruction process using the present facial reconstruction software. Even under blind conditions, the faces reconstructed by this user were closely similar in accuracy to those performed, for the same cases, by a more experienced practitioner under non-blind conditions. Not only these results show that the proposed method in this study could be easily learnt and implemented by a user with no previous experience in the field, but also prove the markedly reduced subjectivity in the proposed method as it is not user-dependent. Rather, it depends on to the availability and the applicability of the relevant software, which can be transferred from one source or facility to another. In contrast, years of experience would be needed before adequate skill and confidence can be developed to master the sculpting techniques, for example (Shahrom et al., 1996). In addition to having a comparable accuracy to many other methods, the proposed method in this thesis is time and cost effective. As well, with the possibility of being potentially accessible and available to other researchers and law enforcement authorities, the presence of an experienced forensic anthropologist as a consultant would be favourable to increase the chances of successful identification (Davy et al., 2005).

6.4 OUTSTANDING PROBLEMS WITH FACIAL RECONSTRUCTION

To reconstruct a face, three components/key aspects are included; a skull, a face and the facial soft tissue depths in between. The only sure component of facial reconstruction is the skull, while others are predicted from the skull. This is simply because each skull has its own unique anthropometric measurements, thus, requires precise mathematical calculations that cannot be standardised (Jedrzejowska, 2001). Consequently, more focus should be paid to obtaining as many clues as possible from the skull structure, and its relationship with the face, to guide more precise identification. For example, it is possible to determine, from the skull, the position and general shape of the main facial features which are in direct anatomical contact with the skull surface. In contrast, the facial parts that consist primarily of soft tissue or cartilage (e.g., lips, nose, and ears), and small details (e.g., hair colour and length, facial fatness, dimples, superficial scars, wrinkles, birthmarks, and skin folds) are difficult to be extrapolated solely from the bone because skeletal remains leave no evidence of their appearance, so they are usually speculative. Furthermore, hair style interpretation is highly subjective and possibly misleading (Shahrom et al., 1996).

Therefore, George (1987) stated that "the artist is technically limited by the "archetype approach", which is the determination of the average soft tissue dimensions that fit a given skull". In other words, despite the skull unique shape is a two-sided coin. On one side, it guides the shape of the face to a large degree, while on the other it is the only ground truth and the other compoenets are only estimates. Therefore, when a face is reconstructed, only a few modifications can be made to the face, and these are related to the soft tissue (e.g., nasal angulation and tip shape, lip position, chin form, and even nutritional variations) which are still attempts by researchers to set guidelines (Vanezis et al., 1989, Stephan and Henneberg, 2001, Wilkinson, 2007, Wilkinson, 2008, Hayes, 2016). It is argued, however, that although the approach of inserting facial features carries a subjective interference by the user, this is acceptable as long as the added features are neutral enough not to distract the public from the reconstructed model, unless supported by other evidence within the case (Andersson and Valfridsson, 2005). Thus, the presence of physical remains in association with the skull would be helpful in providing more information regarding the final appearance of the face. For example, the facial thickness could be approximately estimated from a part of the soft tissue attached to the skull. Other bodily evidence (e.g. jewellery, hair, glasses, etc.) are also useful in finishing the facial reconstruction by adding life like features that are known to belong to the unknown person which would increase the chance of his/her recognition (Gupta et al., 2015).

Moreover, according to the race estimated based on the anthropological examination of the skeletal remains, the facial depths measurements are usually selected. In addition, other factors, such as; diet and life style, also determine the face appearance especially the amount of facial fat, hence, the soft tissue thickness. As these factors are not directly influenced by the skull, they remain unpredictable. However, a problem rises if the skeletal remains belong to a mixed racial origin, which is difficult to identify from the skeletal remains and affects the accuracy of reconstructions due to the unavailability of the relevant soft tissue data. Therefore, the resemblance of the facial reconstructions to the real person may not be strong (Shahrom et al., 1996). It would be advisable that researchers should not rely only on forensic studies of facial soft tissue depths, but also on literature of orthodontics, cosmetic surgery, and cephalometric radiography (George, 1987, Smith and Throckmorton, 2006). This also emphasises the value of multidisciplinary cooperation in research.

One of the main difficulties encountered during using facial soft tissue depths in facial reconstruction is the lack of a standardised set of landmarks with concrete definitions to guide the forensic artist are. This leads to inaccurate positioning of the anatomical landmarks on the skull and the facial images, the degree of which is subject to the practitioner experience. Even with pre-defined anatomical locations, difficulties were faced due to the non-standardised descriptions of the locations and orientations of these anatomical landmarks (Brown et al., 2004, Stephan and Simpson, 2008a, Stephan and Simpson, 2008b, Stephan, 2014). As a result, some researchers had to apply modifications to previously validated sets of facial soft tissue depths (Vanezis, 2008), leading to an increased amount of subjectivity. To minimise this subjectivity, it is necessary to develop a standard set of landmarks and definitions. Also, an entirely objective and repeatable forensic facial reconstruction technique would be aided by an automatic tool for placing the landmarks on the skull (Brown et al., 2004). However, some researchers suggest it is better to adjust the orientation manually (Kähler et al., 2003).

In addition to the problem involved in the methods of assessing the accuracy of these reconstructions in research studies, there is a number of other problems and limitations associated with the key aspects of performing the facial reconstruction, the most obvious which is the subjectivity included. With the numerous approaches that have been developed by researchers and forensic investigators for facial reconstruction (Shahrom et al., 1996, Vanezis et al., 2000, Vanezis, 2008, Claes et al., 2006, Claes et al., 2010, Vandermeulen et al., 2006, Vandermeulen et al., 2012, Wilkinson et al., 2006, Lee et al., 2012, Quatrehomme et al., 1997, Kähler et al., 2003), there is no standardised validation protocol. Thus, the controversy questioning the value and accuracy of facial reconstruction as well as its acceptance by the legal community, as an inadmissible forensic evidence is increased (George, 1987, Vanezis and Brierley, 1996, Huete et al., 2015). It is agreed, however, that facial reconstruction is better used as a last resort only, to provide tentative identification, which is confirmed, or denied, by other tools (Stephan and Henneberg, 2001, Wilkinson, 2007).

Furthermore, in the main study of the present thesis, the ranks of all the cases by different tests were compared. It was observed that certain individual reconstructed faces showed higher accuracy than others using the facial reconstruction method described in this study, when compared to their respective real faces. There was no indication that a certain average facial template produced consistently higher (or lower) accuracy on all the skulls that it was warped onto. Similarly, there was no indication that a face of a certain sex or age group was reconstructed more accurately than that of another group. Rovultionary research, however, revealed the posibility of predicting the face appearnace based on the genetic information of the unknown subject (Claes et al., 2014, Kayser, 2015, Zbieć-Piekarska et al., 2015, Lippert et al., 2017). Ethical concerns were raised, however, because this discovery could breach the anonymity that was promised to volunteers who donated their DNA to research (The Times, 2017). Nonetheless, not only this will identify the remains of an unknown person, but also it can lead the search for a perpetrator based on material from the crime scene (Kayser, 2015). However, in order to explain why certain faces can be better reconstructed than others, further research in genetics should follow to predict the factors that cannot be predicted from the bone and affect the facial appearnace. We believe that these variations, however, could be largely attributed to variation in the body weight or more accurately the Body Mass Index (BMI) that would affect the facial soft tissue depths (De Greef et al., 2009). On the other hand, it was argued that the generated face based on DNA prediction (Lippert et al., 2017) was not entirley similar to the person it belonged, rather to a different person (MIT Technology Review).

Although this face prediction via DNA research can be promising, it requires long, widespread and comprehensive vaildation by fornsic anthropologiss in the first place, compared to the already established facial recontructino methods. The present study can be a guide of similar validation process as it presents an evidence based step-wise appraoch to the suggested methods of assessment, whether subjective or objective, acccording to the circumstances and data available of each research study. This approach should start with the objective asessment methods, wherever possible, as it was proved to more accurate than he subjective assessment. However, in research studies, when the objective assessment are not possible, for example, due to unavailable 3D digital images of the real faces, the subjective tests will be unavoidable. In these cases, the use of the face resemblance tests, as well as consulting forensic anthropologists (preferably with a practical experience in facial reconstruction) would be a relatively more reliable assessment of the accuracy of the facial reconstruction.

What is more agreed upon is that the human face is more complex than many of us think. Therefore, it should be kept in mind that many factors that affect the human face appearance are related to the individual human normal variations and the dietary life style of each individuals. Consequently, these factors are unavoidable limitation inehernt to the process of facial reconstruction in general. These unpredictable variations constitute a challenge to any facial reconstruction method used for forensic human identification as not directly drawn from the bone. However, some solutions have been suggested to increase the chances of recognition based on the facial reconstructions. For instance, Starbuck and Ward (2007) suggested to produce multiple facial variations based on the body weight (emaciated, normal and obese face) within the same ethnic group. Similarly, Richard et al. (2014) reported that presenting multiple images significantly increases the chance of correct recognition among unfamiliar participants. Interestingly, though, it was found that complete strangers could look identical (Daily Mail Online, 2016), which might result in a mistaken recognition. However, more positive identification methods will be conclusive or exclusive if this rare occasion happens.

CHAPTER 7: CONCLUSIONS AND RECOMMENDATIONS

To the best of our knowledge, no previous studies contained this number of previously validated and newly introduced subjective and objective tests. Hence, no extensive statistical correlation was previously performed between these tests. From the results of this study, the following conclusions and recommendations can be made.

The average facial templates were superior to single faces in the identification rates and resemblance scores, thus represent an ideal component of the facial reconstruction method suggested in this study. Also, it might be more convenient to average a number of faces from more than one anthropological group or of different facial types, which adds to the advantages of the average over single facial templates in case of uncertain biological profile of the skeletal remains. Also, regarding the facial soft tissue depths, it was concluded that a population-specific facial soft tissue depths is important for the facial reconstruction. More effort, however, should be put towards standardising the set of the landmarks locations, and definitions. As the proposed approach in this study was applied on a modern Egyptian population for the first time in this study, the accuracy of this method could be increased by using modern population-specific facial soft tissue depths. Moreover, it would be ideal to standardise the position in which the studied cases' heads are scanned with those used to collect the facial soft tissue depths to avoid any gravitation distortion of the facial appearance that discourages recognition of the target faces.

In addition, the present software allows a degree of flexibility by generating multiple facial reconstructions using multiple templates as well as multiple variants of the reconstructed faces according to the facial build (thin, average, and obese), which is expected to increase the chance of the recognition of unknown face (Starbuck and Ward, 2007, Richard et al., 2014). Moreover, the approach of facial reconstruction proposed in this thesis can be considered more superior than many other methods of forensic facial reconstruction in terms of its applicability, transferability from one user to another, and ease of use. Equally important, one of the objectives of this study was to present a cost effective approach to facial recontruction, and that included the selection of free open source software programs whenever possible.

The significant correlation between the previously validated subjective and objective tests made the assessment of the proposed method in this study more reliable by showing more consistent results. However, the inherent subjectivity in the subjective tests will lead to an inevitable and unpredictable error rate in these tests. Accordingly, the objective methods would be more reliable for assessing the accuracy of forensic facial reconstructions in an ideal research study. It should be noted, however, that, currently, there is no way to correlate the subjective recongnisability with the objective surface deviation (Richard et al., 2014), nor with the objective craniofacial anthropometry developed in this study. This would be an area for future research to follow on from this study.

In particular, the results of the subjective face pool assessment test is affected by the variability in their design, the selection of the foil faces, the presentations of the faces, as well as the psychology and characteristics of the participants, etc. The participants' age and professional experience in forensic anthropology, especially with additional experience in forensic facial reconstruction, seem to influence the observers' performance in the subjective face pool tests than other characteristics (e.g. sex, race, and other types of forensic professional experience). Therefore, the face pool tests can indicate whether a facial reconstruction method is successful or not, these tests cannot be a reliable way of drawing an accurate comparison between different facial reconstruction methods. In contrast, the subjective face resemblance assessment test is less affected by these factors, and are closer to the real forensic cases where acquaintances of the deceased make a comparison of resemblance between an advertised facial reconstruction and the face of their missing relative. Therefore, it is a reinforced conclusion to suggest that the face resemblance tests are favoured as a more reliable subjective tests than the face pool tests for assessing the accuracy of forensic facial reconstructions in research studies.

Therefore, in research studies, when the objective assessment are not possible, for example, due to unavailable 3D digital images of the real faces, the subjective tests will be unavoidable. In these cases, the use of the face resemblance tests, as well as consulting forensic anthropologists (preferably with a practical experience in facial reconstruction) would be a relatively more reliable assessment of the accuracy of the facial reconstruction.

Moreover, in the newly introduced subjective and objective assessment methods of the accuracy of the individual facial regions, the summed scores of all the individual regions were successfully correlated with the previously validated methods of assessing the face as a whole. Thus, these methods could give an idea about the goodness of fit in a certain facial feature. In addition, the newly introduced craniofacial anthropometry was validated and developed in this study. It could be successfully used as a method of objective assessment of forensic facial reconstructions by indicating the degree of proportion (i.e. fit) between the real and reconstructed faces. In addition, these findings suggested that the craniofacial angles differences are more sensitive, and thus more valid, as a cranial anthropometric objective test than the linear ratios differences.

Furthermore, it was shown that certain facial regions (particularly the cheek and the jaw regions) can be more reliable than others in the subjective and objective assessment of forensic facial reconstruction as they individually correlated with the other assessment methods. Thus, these areas are more sensitive than others in reflecting the accuracy of the reconstructed face as a whole. It can be recommended in future research to attempt to correlate between the individual facial regions subjective and objective assessment results.

REFERENCES

3D SLICER®. *3D Slicer Software* [Online]. Available: http://bit.ly/2k85eK0 [Accessed January 2017].

3D SYSTEMS. *Geomagic Software®* [Online]. Available: http://bit.ly/2j1mHYq [Accessed June 2015].

3D SYSTEMS. *Geomagic Wrap Software®* [Online]. Available: http://bit.ly/1olBCZ9 [Accessed June 2015].

3DS MAX®. *3Ds max Software* [Online]. Available: http://bit.ly/1A48R7T [Accessed January 2017].

ABATE, A. F., NAPPI, M., RICCIARDI, S. & TORTORA, G. 2004. FACES: 3D FAcial reConstruction from anciEnt Skulls using content based image retrieval. *Journal of Visual Languages & Computing,* 15**,** 373-389.

ANDERSSON, B. & VALFRIDSSON, M. 2005. *Digital 3D facial reconstruction based on computed tomography.*

ATOR. 2012. *Arc-Team Open Research. Forensic facial reconstruction with free software* [Online]. Available: http://bit.ly/2AYqDA6 [Accessed January 2017].

ATTARDI, G., BETRÒ, M., FORTE, M., GORI, R., GUIDAZZOLI, A., IMBODEN, S. & MALLEGNI, F. 1999. 3D facial reconstruction and visualization of ancient Egyptian mummies using spiral CT data: Soft tissues reconstruction and textures application.

AULSEBROOK, W., BECKER, P. & İŞCAN, M. Y. 1996. Facial soft-tissue thicknesses in the adult male Zulu. *Forensic science international,* 79**,** 83-102.

BALDOCK, C., HUGHES, S., WHITTAKER, D., TAYLOR, J., DAVIS, R., SPENCER, A., TONGE, K. & SOFAT, A. 1994. 3-D reconstruction of an ancient Egyptian mummy using X-ray computer tomography. *Journal of the Royal Society of Medicine,* 87**,** 806-808.

BBC NEWS. 2005. *Face of Tutankhamun reconstructed* [Online]. Available: http://bit.ly/2BIm49u [Accessed January 2017].

BLENDER®. *Blender software* [Online]. Available: http://bit.ly/1HVFvqR [Accessed January 2017].

BLENDERNATION. 2013. *Faces of Evolution - Exhibition of facial forensic reconstructions of hominids* [Online]. Available: http://bit.ly/2jpnnna [Accessed January 2017].

BLENDERNATION. 2014. *The face of St. Anthony finally revealed, and in 3D* [Online]. Available: http://bit.ly/2jrWC1m [Accessed January 2017].

BRIERS, N., BRIERS, T., BECKER, P. & STEYN, M. 2015. Soft tissue thickness values for black and coloured South African children aged 6–13 years. *Forensic science international,* 252**,** 188. e1-188. e10.

BRITISH TRANSPORT POLICE. *How the final King's Cross fire victim was identified* [Online]. Available: http://bit.ly/2jsdW6n [Accessed April 2017].

BROWN, R. E., KELLIHER, T. P., TU, P. H., TURNER, W. D., TAISTER, M. A. & MILLER, K. W. 2004. A survey of tissue-depth landmarks for facial approximation. *Forensic Sci. Commun,* 6.

BULUT, O., SIPAHIOGLU, S. & HEKIMOGLU, B. 2014. Facial soft tissue thickness database for craniofacial reconstruction in the Turkish adult population. *Forensic science international,* 242**,** 44-61.

CAHAREL, S., JACQUES, C., D'ARRIPE, O., RAMON, M. & ROSSION, B. 2011. Early electrophysiological correlates of adaptation to personally familiar and unfamiliar faces across viewpoint changes. *Brain research,* 1387**,** 85-98.

CAVALCANTI, M. & VANNIER, M. 1998. Quantitative analysis of spiral computed tomography for craniofacial clinical applications. *Dentomaxillofacial Radiology,* 27**,** 344-350.

CAVANAGH, D. & STEYN, M. 2011. Facial reconstruction: soft tissue thickness values for South African black females. *Forensic science international,* 206**,** 215. e1-215. e7.

CESARANI, F., MARTINA, M. C., FERRARIS, A., GRILLETTO, R., BOANO, R., MAROCHETTI, E. F., DONADONI, A. M. & GANDINI, G. 2003. Whole-body three-dimensional multidetector CT of 13 Egyptian human mummies. *American Journal of Roentgenology,* 180**,** 597-606.

CESARANI, F., MARTINA, M. C., GRILLETTO, R., BOANO, R., ROVERI, A. M. D., CAPUSSOTTO, V., GIULIANO, A., CELIA, M. & GANDINI, G. 2004. Facial reconstruction of a wrapped Egyptian mummy using MDCT. *American Journal of Roentgenology,* 183**,** 755-758.

CHA, K.-S. 2013. Soft-tissue thickness of South Korean adults with normal facial profiles. *The Korean Journal of Orthodontics,* 43**,** 178-185.

CHAN, W. N. J. 2007. *In vivo facial tissue depth study of Chinese-Americans in New York City.* Faculty of the Louisiana State University and Agricultural and Mechanical College in partial fulfillment of the requirements for the degree of Master of Arts in The Department of Geography and Anthropology by Wing Nam Joyce Chan BA, Macalester College.

CHRON NEWS. 2008. *Careful work solved one of Houston's oldest cold cases* [Online]. Available: http://bit.ly/2iBWf3u [Accessed January 2017].

CHUNG, J.-H., CHEN, H.-T., HSU, W.-Y., HUANG, G.-S. & SHAW, K.-P. 2015. A CT-scan database for the facial soft tissue thickness of Taiwan adults. *Forensic science international,* 253**,** 132. e1-132. e11.

CLAES, P., HILL, H. & SHRIVER, M. D. 2014. Toward DNA-based facial composites: preliminary results and validation. *Forensic Science International: Genetics,* 13**,** 208-216.

CLAES, P., VANDERMEULEN, D., DE GREEF, S., WILLEMS, G., CLEMENT, J. G. & SUETENS, P. 2010. Bayesian estimation of optimal craniofacial reconstructions. *Forensic science international,* 201**,** 146-152.

CLAES, P., VANDERMEULEN, D., DE GREEF, S., WILLEMS, G. & SUETENS, P. 2006. Craniofacial reconstruction using a combined statistical model of face shape and soft tissue depths: methodology and validation. *Forensic science international,* 159**,** S147-S158.

CODINHA, S. 2009. Facial soft tissue thicknesses for the Portuguese adult population. *Forensic science international,* 184**,** 80. e1-80. e7.

CURRY, S., BAUMRIND, S., CARLSON, S., BEERS, A. & BOYD, R. 2001. Integrated three-dimensional ciraniofacial mapping at the Craniofacial Research Instrumentation Laboratory/University of the Pacific. *Seminars in Orthodontics,* 7**,** 258-265.

DAILY MAIL ONLINE. 2016. Available: http://bit.ly/2BIm49u [Accessed August 2017].

DAVY, S., GILBERT, T., SCHOFIELD, D. & EVISON, M. 2005. Forensic facial reconstruction using computer modeling software. *Computer-Graphic Facial Reconstruction***,** 183-194.

DE ALMEIDA, N. H., MICHEL-CROSATO, E., DE PAIVA, L. A. S. & BIAZEVIC, M. G. H. 2013. Facial soft tissue thickness in the Brazilian population: new reference data and anatomical landmarks. *Forensic science international,* 231**,** 404. e1-404. e7.

DE GREEF, S., CLAES, P., MOLLEMANS, W., LOUBELE, M., VANDERMEULEN, D., SUETENS, P. & WILLEMS, G. 2005. Semi-automated ultrasound facial soft tissue depth registration: method and validation. *Journal of Forensic Science,* 50**,** JFS2004547-7.

DE GREEF, S., CLAES, P., VANDERMEULEN, D., MOLLEMANS, W., SUETENS, P. & WILLEMS, G. 2006. Large-scale in-vivo Caucasian facial soft tissue thickness database for craniofacial reconstruction. *Forensic science international,* 159**,** S126-S146.

DE GREEF, S., VANDERMEULEN, D., CLAES, P., SUETENS, P. & WILLEMS, G. 2009. The influence of sex, age and body mass index on facial soft tissue depths. *Forensic science, medicine, and pathology,* 5**,** 60-65.

DECKER, S., FORD, J., DAVY-JOW, S., FARAUT, P., NEVILLE, W. & HILBELINK, D. 2013. Who is this person? A comparison study of current three-dimensional facial approximation methods. *Forensic Sci Int,* 229**,** 161.e1-8.

DICOM ANONYMIZER®. Available: http://bit.ly/2AnMkKb [Accessed April 2015].

DOMARACKI, M. & STEPHAN, C. N. 2006. Facial soft tissue thicknesses in Australian adult cadavers. *Journal of forensic sciences,* 51**,** 5-10.

DONG, Y., HUANG, L., FENG, Z., BAI, S., WU, G. & ZHAO, Y. 2012. Influence of sex and body mass index on facial soft tissue thickness measurements of the northern Chinese adult population. *Forensic science international,* 222**,** 396. e1-396. e7.

EIFUKU, S., DE SOUZA, W. C., NAKATA, R., ONO, T. & TAMURA, R. 2011. Neural representations of personally familiar and unfamiliar faces in the anterior inferior temporal cortex of monkeys. *PLoS One,* 6**,** e18913.

EL-MEHALLAWI, I. H. & SOLIMAN, E. M. 2001. Ultrasonic assessment of facial soft tissue thicknesses in adult Egyptians. *Forensic Science International,* 117**,** 99-107.

EVENHOUSE, R. J., RASMUSSEN, M. & SADLER, L. L. Computer-aided forensic facial reconstruction. Boston-DL tentative, 1991. International Society for Optics and Photonics, 147-156.

EVIDENCE, T. W. G. F. E. & AMERICA, U. S. O. 2003. Eyewitness Evidence: A Trainer's Manual for Law Enforcement.

EVISON, M. P. & GREEN, M. A. 1999. Presenting three-dimensional forensic facial simulations on the Internet using VRML. *Journal of Forensic Science,* 44**,** 1219-1223.

FACE GEN®. *FaceGen Modeller software* [Online]. Available: http://bit.ly/2BRTyDg [Accessed April 2017].

FEDOROV, A., BEICHEL, R., KALPATHY-CRAMER, J., FINET, J., FILLION-ROBIN, J.-C., PUJOL, S., BAUER, C., JENNINGS, D., FENNESSY, F., SONKA, M., BUATTI, J., AYLWARD, S., MILLER, J. V., PIEPER, S. & KIKINIS, R. 2012. 3D Slicer as an image computing platform for the Quantitative Imaging Network. *Magnetic Resonance Imaging,* 30**,** 1323-1341.

FERNANDES, C. M. S., DA COSTA SERRA, M., DA SILVA, J. V. L., NORITOMI, P. Y., DE SENA PEREIRA, F. D. A. & MELANI, R. F. H. 2012. Tests of one Brazilian facial reconstruction method using three soft tissue depth sets and familiar assessors. *Forensic science international,* 214**,** 211. e1-211. e7.

FRIEDRICH, K. M., NEMEC, S., CZERNY, C., FISCHER, H., PLISCHKE, S., GAHLEITNER, A., VIOLA, T. B., IMHOF, H., SEIDLER, H. & GUILLEN, S. 2010. The story of 12 Chachapoyan mummies through multidetector computed tomography. *European journal of radiology,* 76**,** 143-150.

GEOMAGIC CONTROL®. *Geomagic Control Software* [Online]. Available: http://bit.ly/2BJWDoe [Accessed August 2016].

GEOMAGIC WRAP®. *Geomagic Wrap Software* [Online]. Available: http://bit.ly/1olBCZ9 [Accessed June 2015].

GEORGE, R. M. 1987. The lateral craniographic method of facial reconstruction. *Journal of Forensic Science,* 32**,** 1305-1330.

GIMP®. *GNU Image Manipulation Program* [Online]. Available: http://bit.ly/1I0gRf9 [Accessed January 2017].

GLAISTER, J. & BRASH, J. C. 1937. *Medico-legal aspects of the Ruxton case*, Livingstone.

GOLDINGER, S. D., HE, Y. & PAPESH, M. H. 2009. Deficits in cross-race face learning: insights from eye movements and pupillometry. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* 35**,** 1105.

GOOGLE DOCS. *Google document forms* [Online]. Available: http://bit.ly/1lEA3DK [Accessed February 2016].

GOVERNMENT TECHNOLOGY. 2013. *Facial Reconstruction Tech Helps Identify Skeletal Remains* [Online]. Available: http://bit.ly/2iC2zbp [Accessed January 2017].

GUPTA, S., GUPTA, V., VIJ, H., VIJ, R. & TYAGI, N. 2015. Forensic Facial Reconstruction: The Final Frontier. *Journal of clinical and diagnostic research: JCDR,* 9**,** ZE26.

GUYOMARC'H, P., SANTOS, F., DUTAILLY, B. & COQUEUGNIOT, H. 2013. Facial soft tissue depths in French adults: variability, specificity and estimation. *Forensic science international,* 231**,** 411. e1-411. e10.

HAGLUND, W. D. & REAY, D. T. 1991. Use of facial approximation techniques in identification of Green River serial murder victims. *The American journal of forensic medicine and pathology,* 12**,** 132-142.

HAYES, S. 2014. Facial approximation of 'Angel': Case specific methodological review. *Forensic science international,* 237**,** e30-e41.

HAYES, S. 2016. A geometric morphometric evaluation of the Belanglo 'Angel'facial approximation. *Forensic Science International,* 268**,** e1-e12.

HERRERA, L. M., STRAPASSON, R. A. P., DA SILVA, J. V. L. & MELANI, R. F. H. 2016. Forensic facial approximation assessment: can application of different average facial

tissue depth data facilitate recognition and establish acceptable level of resemblance? *Forensic Science International,* 266**,** 311-319.
HODSON, G., LIEBERMAN, L. S. & WRIGHT, P. 1985. In vivo measurements of facial tissue thicknesses in American Caucasoid children. *Journal of Forensic Science,* 30**,** 1100-1112.
HUETE, M. I., IBÁÑEZ, O., WILKINSON, C. & KAHANA, T. 2015. Past, present, and future of craniofacial superimposition: Literature and international surveys. *Legal Medicine,* 17**,** 267-278.
HUGHES, S. 1996. Three-dimensional reconstruction of an ancient Egyptian mummy. *Imaging the Past. Electronic Imaging and Computer Graphics in museums and archaeology,* 114**,** 211-225.
HWANG, H. S., PARK, M. K., LEE, W. J., CHO, J. H., KIM, B. K. & WILKINSON, C. M. 2012. Facial soft tissue thickness database for craniofacial reconstruction in Korean adults. *Journal of forensic sciences,* 57**,** 1442-1447.
INVESALIUS®. *InVesalius software* [Online]. Available: http://bit.ly/2BTYGqA [Accessed January 2017].
ISLAM, S. M., GOONEWARDENE, M. S. & FARELLA, M. 2015. A Review on Three Dimensional Facial Averaging for the Assessment of Orthodontic Disorders. *Innovations and Advances in Computing, Informatics, Systems Sciences, Networking and Engineering.* Springer.
JAYARATNE, Y. S., MCGRATH, C. P. & ZWAHLEN, R. A. 2012. How accurate are the fusion of cone-beam CT and 3-D stereophotographic images? *PloS one,* 7**,** e49585.
JEDRZEJOWSKA, Z. K. 2001. Craniometry and mathematical calculations as a method for viscero-cranium profile determination. *Forensic science international,* 117**,** 145-151.
JONES, M. W. Facial Reconstruction Using Volumetric Data. VMV, 2001. Citeseer, 135-150.
KÄHLER, K., HABER, J. & SEIDEL, H.-P. Reanimating the dead: reconstruction of expressive faces from skull data. ACM Transactions on Graphics (TOG), 2003. ACM, 554-561.
KAU, C. H., ZHUROV, A., RICHMOND, S., BIBB, R., SUGAR, A., KNOX, J. & HARTLES, F. 2006. The 3-dimensional construction of the average 11-year-old child face: a clinical evaluation and application. *Journal of oral and maxillofacial surgery,* 64**,** 1086-1092.
KAYSER, M. 2015. Forensic DNA phenotyping: predicting human appearance from crime scene material for investigative purposes. *Forensic Science International: Genetics,* 18**,** 33-48.
KIM, K.-D., RUPRECHT, A., WANG, G., LEE, J. B., DAWSON, D. V. & VANNIER, M. W. 2005. Accuracy of facial soft tissue thickness measurements in personal computer-based multiplanar reconstructed computed tomographic images. *Forensic science international,* 155**,** 28-34.
KIM, M., HUH, K.-H., YI, W.-J., HEO, M.-S., LEE, S.-S. & CHOI, S.-C. 2012. Evaluation of accuracy of 3D reconstruction images using multi-detector CT and cone-beam CT. *Imaging science in dentistry,* 42**,** 25-33.

KLEINBERG, K. F. & VANEZIS, P. 2007. Variation in proportion indices and angles between selected facial landmarks with rotation in the Frankfort plane. *Medicine, science and the law,* 47**,** 107-116.

KLEINBERG, K. F., VANEZIS, P. & BURTON, A. M. 2007. Failure of anthropometry as a facial identification technique using high - quality photographs. *Journal of forensic sciences,* 52**,** 779-783.

KSBW 8 NEWS. 2015. *Facial reconstruction provides new info for cold case* [Online]. Available: http://bit.ly/2j0xQIR [Accessed January 2017].

LATINOWARE. 2012. *The use of technology in police investigation* [Online]. Available: http://bit.ly/2AECn86 [Accessed January 2017].

LEE, W. J., WILKINSON, C. M. & HWANG, H. S. 2012. An Accuracy Assessment of Forensic Computerized Facial Reconstruction Employing Cone - Beam Computed Tomography from Live Subjects. *Journal of forensic sciences,* 57**,** 318-327.

LIPPERT, C., SABATINI, R., MAHER, M. C., KANG, E. Y., LEE, S., ARIKAN, O., HARLEY, A., BERNAL, A., GARST, P., LAVRENKO, V., YOCUM, K., WONG, T., ZHU, M., YANG, W.-Y., CHANG, C., LU, T., LEE, C. W. H., HICKS, B., RAMAKRISHNAN, S., TANG, H., XIE, C., PIPER, J., BREWERTON, S., TURPAZ, Y., TELENTI, A., ROBY, R. K., OCH, F. J. & VENTER, J. C. 2017. Identification of individuals by trait prediction using whole-genome sequencing data. *Proceedings of the National Academy of Sciences,* 114**,** 10166-10171.

LYNNERUP, N. 2002. 3-D CAT-scan: Anthropology, Archaeology and Virtual Reality. *BAR International Series,* 1016**,** 217-220.

MATHILDA'S ANTHROPOLOGY BLOG. 2011. *Spirit Cave Man* [Online]. Available: http://bit.ly/2zZqkB6 [Accessed January 2017].

MESHLAB®. *Meshlab software, Visual Computing Lab - ISTI - CNR* [Online]. Available: http://bit.ly/2AEdM3j [Accessed January 2017].

MIT TECHNOLOGY REVIEW. *Does Your Genome Predict Your Face? Not Quite Yet* [Online]. Available: http://bit.ly/2xS1anw [Accessed September 2017].

MIYASAKA, S., YOSHINO, M., IMAIZUMI, K. & SETA, S. 1995. The computer-aided facial reconstruction system. *Forensic science international,* 74**,** 155-165.

MOLLOV, N., BOSIO, J. A., PRUSZYNSKI, J. & WIRTZ, T. 2012. Intra-and inter-examiner reliability of direct facial soft tissue measurements using digital calipers. *Journal of the World Federation of Orthodontists,* 1**,** e157-e161.

MOYERS, D. K. 2007. Validation Study of ReFace (Reality Enhanced Facial Approximation by Computational Estimation).

MUMMY TOMBS. 2011. *Ötzi the Iceman; His Three Faces* [Online]. Available: http://bit.ly/2j2BGRC [Accessed January 2017].

NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY NIST. *Facial Recognition Technology (FERET)* [Online]. Available: http://bit.ly/2A24mgZ [Accessed May 2014].

NELSON, L. A. & MICHAEL, S. D. 1998. The application of volume deformation to three-dimensional facial reconstruction: A comparison with previous techniques. *Forensic Science International,* 94**,** 167-181.

PARKS, C. L., RICHARD, A. H. & MONSON, K. L. 2013. Preliminary performance assessment of computer automated facial approximations using computed tomography scans of living individuals. *Forensic science international,* 233**,** 133-139.

PARKS, C. L., RICHARD, A. H. & MONSON, K. L. 2014. Preliminary assessment of facial soft tissue thickness utilizing three-dimensional computed tomography models of living individuals. *Forensic science international,* 237**,** 146. e1-146. e10.

PHILLIPS, V. & SMUTS, N. 1996. Facial reconstruction: utilization of computerized tomography to measure facial tissue thickness in a mixed racial population. *Forensic Science International,* 83**,** 51-59.

PHILLIPS, V. M. 2001. Skeletal remains identification by facial reconstruction. *Forensic Science Communications,* 3**,** 4-15.

PITHON, M. M., RIBEIRO, D. L. R., DOS SANTOS, R. L., DE SANTANA, C. L. & CRUZ, J. P. P. 2014. Soft tissue thickness in young north eastern Brazilian individuals with different skeletal classes. *Journal of forensic and legal medicine,* 22**,** 115-120.

POULSEN, K. & SIMONSEN, J. 2007. Computed tomography as routine in connection with medico-legal autopsies. *Forensic science international,* 171**,** 190-197.

QUATREHOMME, G., BALAGUER, T., STACCINI, P. & ALUNNI-PERRET, V. 2007. Assessment of the accuracy of three-dimensional manual craniofacial reconstruction: a series of 25 controlled cases. *International Journal of Legal Medicine,* 121**,** 469-475.

QUATREHOMME, G., COTIN, S., SUBSOL, G., DELINGETTE, H., GARIDEL, Y., GREVIN, G., FIDRICH, M., BAILET, P. & OLLIER, A. 1997. A fully three-dimensional method for facial reconstruction based on deformable models. *J Forensic Sci,* 42**,** 649-52.

RATHNAYAKA, K., MOMOT, K. I., NOSER, H., VOLP, A., SCHUETZ, M. A., SAHAMA, T. & SCHMUTZ, B. 2012. Quantification of the accuracy of MRI generated 3D models of long bones compared to CT generated 3D models. *Medical engineering & physics,* 34**,** 357-363.

RATHNAYAKA, K., SAHAMA, T., SCHUETZ, M. A. & SCHMUTZ, B. 2011. Effects of CT image segmentation methods on the accuracy of long bone 3D reconstructions. *Medical engineering & physics,* 33**,** 226-233.

RICHARD, A. H., PARKS, C. L. & MONSON, K. L. 2014. Assessment of presentation methods for ReFace computerized facial approximations. *Forensic science international,* 242**,** 283-292.

ROCHA, S. D. S., RAMOS, D. L. D. P. & CAVALCANTI, M. D. G. P. 2003. Applicability of 3D-CT facial reconstruction for forensic individual identification. *Pesquisa Odontológica Brasileira,* 17**,** 24-28.

RUIZ, N. A. P. 2013. Facial soft tissue thickness of Colombian adults. *Forensic science international,* 229**,** 160. e1-160. e9.

SAHNI, D., SINGH, G., JIT, I. & SINGH, P. 2008. Facial soft tissue thickness in northwest Indian adults. *Forensic science international,* 176**,** 137-146.

SAKUMA, A., ISHII, M., YAMAMOTO, S., SHIMOFUSA, R., KOBAYASHI, K., MOTANI, H., HAYAKAWA, M., YAJIMA, D., TAKEICHI, H. & IWASE, H. 2010. Application of

Postmortem 3D‐CT Facial Reconstruction for Personal Identification. *Journal of forensic sciences,* 55**,** 1624-1629.
SALAH, A. A., ALYÜZ, N. & AKARUN, L. 2008. Registration of three-dimensional face scans with average face models. *Journal of Electronic Imaging,* 17**,** 011006-011006-14.
SAUKKO, P. & KNIGHT, B. 2004. The establishment of identity of human remains. *Knigth's Forensic Pathology,(3º ed). USA: Hodder Arnold.*
SCHUH, P., SCHEURER, E., FRITZ, K., PAVLIC, M., HASSLER, E., RIENMÜLLER, R. & YEN, K. 2013. Can clinical CT data improve forensic reconstruction? *International journal of legal medicine,* 127**,** 631-638.
SFORZA, C., ELAMIN, F., TOMMASI, D. G., DOLCI, C. & FERRARIO, V. F. 2013. Morphometry of the soft tissues of the orbital region in Northern Sudanese persons. *Forensic science international,* 228**,** 180. e1-180. e11.
SHAHROM, A., VANEZIS, P., CHAPMAN, R., GONZALES, A., BLENKINSOP, C. & ROSSI, M. 1996. Techniques in facial identification: computer-aided facial reconstruction using a laser scanner and video superimposition. *International journal of legal medicine,* 108**,** 194-200.
SHIMOFUSA, R., YAMAMOTO, S., HORIKOSHI, T., YOKOTA, H. & IWASE, H. 2009. Applicability of facial soft tissue thickness measurements in 3-dimensionally reconstructed multidetector-row CT images for forensic anthropological examination. *Legal Medicine,* 11**,** S256-S259.
SHORT, L. J., KHAMBAY, B., AYOUB, A., EROLIN, C., RYNN, C. & WILKINSON, C. 2014. Validation of a computer modelled forensic facial reconstruction technique using CT data from live subjects: a pilot study. *Forensic science international,* 237**,** 147. e1-147. e8.
SHWEEL, M., AMER, M. I. & EL-SHAMANHORY, A. F. 2013. A comparative study of cone-beam CT and multidetector CT in the preoperative assessment of odontogenic cysts and tumors. *The Egyptian Journal of Radiology and Nuclear Medicine,* 44**,** 23-32.
SIPAHIOĞLU, S., ULUBAY, H. & DIREN, H. B. 2012. Midline facial soft tissue thickness database of Turkish population: MRI study. *Forensic science international,* 219**,** 282. e1-282. e8.
SKETCHFAB. *3D models sharing website* [Online]. Available: http://bit.ly/1Itc05f [Accessed February 2016].
SKETCHFAB PRO ACCOUNT. Available: http://bit.ly/2nCU1Gm [Accessed February 2016].
SMITH, S. L. & THROCKMORTON, G. S. 2006. Comparability of Radiographic and 3D‐Ultrasound Measurements of Facial Midline Tissue Depths. *Journal of forensic sciences,* 51**,** 244-247.
SNOW, C. C., GATLIFF, B. P. & MCWILLIAMS, K. R. 1970. Reconstruction of facial features from the skull: an evaluation of its usefulness in forensic anthropology. *American Journal of Physical Anthropology,* 33**,** 221-227.
SONG, F., HOOPER, L. & LOKE, Y. 2013. Publication bias: what is it? How do we measure it? How do we avoid it? *Open Access Journal of Clinical Trials,* 2013**,** 71-81.
STARBUCK, J. M. & WARD, R. E. 2007. The affect of tissue depth variation on craniofacial reconstructions. *Forensic science international,* 172**,** 130-136.

STEPHAN, C. *CRANIOFACIALidentification.com* [Online]. Available: http://bit.ly/2BTh3Mo [Accessed January 2017].
STEPHAN, C. & ARTHUR, R. 2006. Assessing facial approximation accuracy: how do resemblance ratings of disparate faces compare to recognition tests? *Forensic science international,* 159**,** S159-S163.
STEPHAN, C. & HENNEBERG, M. 2006. Recognition by forensic facial approximation: case specific examples and empirical tests. *Forensic science international,* 156**,** 182-191.
STEPHAN, C. N. 2014. The Application of the Central Limit Theorem and the Law of Large Numbers to Facial Soft Tissue Depths: T - Table Robustness and Trends since 2008. *Journal of forensic sciences,* 59**,** 454-462.
STEPHAN, C. N. & CICOLINI, J. 2008. Measuring the accuracy of facial approximations: a comparative study of resemblance rating and face array methods. *Journal of forensic sciences,* 53**,** 58-64.
STEPHAN, C. N. & HENNEBERG, M. 2001. Building faces from dry skulls: are they recognized above chance rates? *Journal of Forensic Science,* 46**,** 432-440.
STEPHAN, C. N., NORRIS, R. M. & HENNEBERG, M. 2005a. Does sexual dimorphism in facial soft tissue depths justify sex distinction in craniofacial identification? *J Forensic Sci,* 50**,** 513-8.
STEPHAN, C. N., PENTON-VOAK, I. S., CLEMENT, J. G. & HENNEBERG, M. 2005b. Ceiling recognition limits of two-dimensional facial approximations constructed using averages. *In:* MARKS, J. G. C. A. M. K. (ed.) *Computer-graphic facial reconstruction.* Boston, United States: Academic Press.
STEPHAN, C. N. & SIMPSON, E. K. 2008a. Facial soft tissue depths in craniofacial identification (part I): an analytical review of the published adult data. *Journal of Forensic Sciences,* 53**,** 1257-1272.
STEPHAN, C. N. & SIMPSON, E. K. 2008b. Facial Soft Tissue Depths in Craniofacial Identification (Part II): An Analytical Review of the Published Sub - Adult Data. *Journal of forensic sciences,* 53**,** 1273-1279.
SZUMILAS, M. 2010. Explaining odds ratios. *Journal of the Canadian academy of child and adolescent psychiatry,* 19**,** 227.
TAYLOR, K. T. 2001. *Forensic art and illustration*, CRC Press.
TEDESCHI-OLIVEIRA, S. V., MELANI, R. F. H., DE ALMEIDA, N. H. & DE PAIVA, L. A. S. 2009. Facial soft tissue thickness of Brazilian adults. *Forensic science international,* 193**,** 127. e1-127. e7.
THE GUARDIAN. 2004. *The real story of body 115* [Online]. Available: http://bit.ly/2AZiLy9 [Accessed May 2017].
THE TIMES. 2017. *Scientists use genetic code to build a picture of your face* [Online]. Available: http://bit.ly/2AlUaE5 [Accessed September 2017].
TILOTTA, F., RICHARD, F., GLAUNÈS, J., BERAR, M., GEY, S., VERDEILLE, S., ROZENHOLC, Y. & GAUDY, J.-F. 2009. Construction and analysis of a head CT-scan database for craniofacial reconstruction. *Forensic science international,* 191**,** 112. e1-112. e12.
TRIANGLE CALCULATOR. Available: http://bit.ly/2Bceh7K [Accessed June 2016].

TURNER, W., BROWN, R., KELLIHER, T., TU, P., TAISTER, M. & MILLER, K. 2005. A novel method of automated skull registration for forensic facial approximation. *Forensic science international,* 154**,** 149-158.

UNIVERSITY OF LEICESTER. 2013. *Face and voice* [Online]. Available: http://bit.ly/2j4Ppav [Accessed January 2017].

UTSUNO, H., KAGEYAMA, T., UCHIDA, K. & KIBAYASHI, K. 2014. Facial soft tissue thickness differences among three skeletal classes in Japanese population. *Forensic science international,* 236**,** 175-180.

UTSUNO, H., KAGEYAMA, T., UCHIDA, K., YOSHINO, M., MIYAZAWA, H. & INOUE, K. 2010. Facial soft tissue thickness in Japanese children. *Forensic science international,* 199**,** 109. e1-109. e6.

VANDERMEULEN, D., CLAES, P., DE GREEF, S., WILLEMS, G., CLEMENT, J. & SUETENS, P. 2012. Automated facial reconstruction. *Craniofacial Identification,* 203.

VANDERMEULEN, D., CLAES, P., DEVELTER, W., DE GREEF, S., COUDYZER, W., WILLEMS, G. & SUETENS, P. 2013. A computational framework for CT-based craniofacial reconstruction. *Journal of Forensic Radiology and Imaging,* 1**,** 86-87.

VANDERMEULEN, D., CLAES, P., LOECKX, D., DE GREEF, S., WILLEMS, G. & SUETENS, P. 2006. Computerized craniofacial reconstruction using CT-derived implicit surface representations. *Forensic science international,* 159**,** S164-S174.

VANEZIS, M. 2008. *Forensic facial reconstruction using 3-D computer graphics: evaluation and improvement of its reliability in identification.* University of Glasgow.

VANEZIS, M. & VANEZIS, P. 2000. Cranio-facial reconstruction in forensic identification—historical development and a review of current practice. *Medicine, Science and the Law,* 40**,** 197-205.

VANEZIS, P., BLOWES, R. W., LINNEY, A. D., TAN, A. C., RICHARDS, R. & NEAVE, R. 1989. Application of 3-D computer graphics for facial reconstruction and comparison with sculpting techniques. *Forensic Sci Int,* 42**,** 69-84.

VANEZIS, P. & BRIERLEY, C. 1996. Facial image comparison of crime suspects using video superimposition. *Science & Justice,* 36**,** 27-33.

VANEZIS, P., VANEZIS, M., MCCOMBE, G. & NIBLETT, T. 2000. Facial reconstruction using 3-D computer graphics. *Forensic science international,* 108**,** 81-95.

VERZÉ, L. 2009. History of facial reconstruction. *Acta Biomed,* 80**,** 5-12.

WILKINSON, C. 2003. The facial reconstruction of the Marina el-Alamein mummy. *Polish Archaeology in the Mediterranean XIV Reports 2002. Warsaw University, 66,* 71.

WILKINSON, C. 2005. Computerized forensic facial reconstruction. *Forensic Science, Medicine, and Pathology,* 1**,** 173-177.

WILKINSON, C. 2007. Facial Anthropology and Reconstruction. *Forensic Human Identification.* CRC Press.

WILKINSON, C. 2008. *Forensic facial reconstruction*, Cambridge University Press.

WILKINSON, C. & EVANS, R. 2009. Are facial image analysis experts any better than the general public at identifying individuals from CCTV images? *Science & Justice,* 49**,** 191-196.

WILKINSON, C. & EVANS, R. 2011. Corrigendum to “Are facial image analysis experts any better than the general public at identifying individuals from CCTV images?”[Science and Justice 49 (2009) 191–196]. *Science & Justice,* 51**,** 218-221.

WILKINSON, C., RYNN, C., PETERS, H., TAISTER, M., KAU, C. H. & RICHMOND, S. 2006. A blind accuracy assessment of computer-modeled forensic facial reconstruction using computed tomography data from live subjects. *Forensic science, medicine, and pathology,* 2**,** 179-187.

WILKINSON, C. & WHITTAKER, D. Juvenile forensic facial reconstruction—a detailed accuracy study. Proceedings of the 10th conference of the International Association of Craniofacial Identification, 2002. 11-14.

WILKINSON, C. M. 2002. In vivo facial tissue depth measurements for white British children. *Journal of Forensic Science,* 47**,** 459-465.

WOODSTOCK PATCH. 2016. *Unidentified Human Remains: GBI Releases Facial Reconstruction in Cold Case* [Online]. Available: http://bit.ly/2j4q0hd [Accessed January 2017].

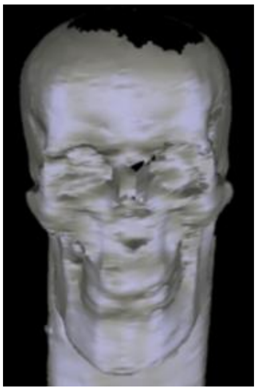
XI, T., VAN LOON, B., FUDALEJ, P., BERGE, S., SWENNEN, G. & MAAL, T. 2013. Validation of a novel semi-automated method for three-dimensional surface rendering of condyles using cone beam computed tomography data. *International journal of oral and maxillofacial surgery,* 42**,** 1023-1029.

YORK FACE VAR LAB. *Glasgow Face Matching Test (GFMT) and Glasgow Unfamiliar Face Database (GUFD)* [Online]. Available: http://bit.ly/2BIm49u [Accessed May 2014].

ZBIEĆ-PIEKARSKA, R., SPÓLNICKA, M., KUPIEC, T., PARYS-PROSZEK, A., MAKOWSKA, Ż., PAŁECZKA, A., KUCHARCZYK, K., PŁOSKI, R. & BRANICKI, W. 2015. Development of a forensically useful age prediction method based on DNA methylation analysis. *Forensic Science International: Genetics,* 17**,** 173-179.

APPENDICES

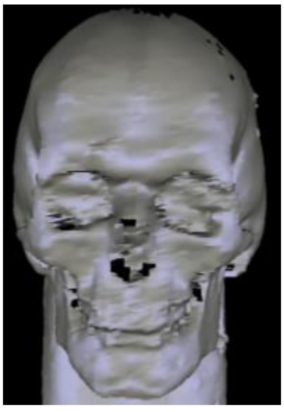
APPENDIX 1: THE SKULLS AND ANTEMORTEM PHOTOGRAPHS OF THE PILOT STUDY CASES



a

b

Skull (I) Case (a: Laser Scanned Skull, b: Ante Mortem Photograph)



a

b

Skull (II) Case (a: Laser Scanned Skull, b: Ante Mortem Photograph)

a b

Skull (III) Case (a: Laser Scanned Skull, b: Ante Mortem Photograph)

a b

Skull (IV) Case (a: Laser Scanned Skull, b: Ante Mortem Photographs)

APPENDIX 2: THE FACIAL TEMPLATES USED FOR THE FACIAL RECONSTRUCTIONS OF THE MALE AND FEMALE CASES


20Y-01


20Y-02


20Y-03


20Y-04


20Y-05


20Y-06


20Y-07

The 10 Single Facial Templates (aged 20 years old) used For the Facial Reconstruction of the 3 Male Cases


30Y-01


30Y-02


30Y-03

The 10 Single Facial Templates (aged 30 years old) used For the Facial Reconstruction of the 3 Male Cases

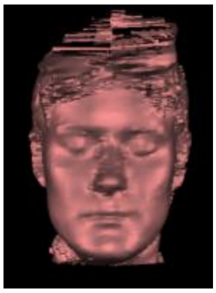
20Y-AV

30Y-AV

40Y-AV

The 3 Average Facial Templates (aged 20, 30, and 40 years old) used For the Facial Reconstruction of the 3 Male Cases

20Y-FAV



30Y-FAV

The 2 Average Facial Templates Used For the Facial Reconstruction of the Female Case

APPENDIX 3: THE WARPING PROCESS USING THE FACIAL RECONSTRUCTION SOFTWARE

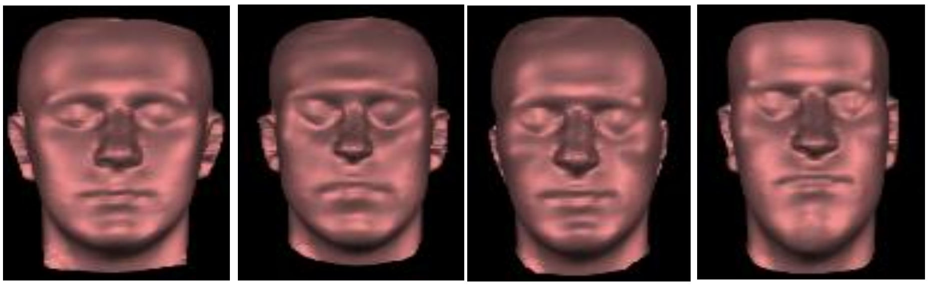
a b c

a: A 3D skull mesh in the facial reconstruction software after placing the anatomical landmarks

b: A 3D facial template mesh in the facial reconstruction software after placing the anatomical landmarks

c: The reconstructed face after automatic “warping” of the facial template (b) onto the skull (a) by the facial reconstruction software

APPENDIX 4-A: AN EXAMPLE OF THE FACE POOL TEST FORM (A) FOR SKULL CASE (I)



APPENDIX 4-B: THE 3D FORENSIC FACIAL RECONSTRUCTION EXPERIMENT (PAPER INSTRUCTION FORM)-FACE POOL TEST FORM (A)

3D Forensic Facial Reconstruction Experiment

(Instruction form)

Forensic Facial Reconstruction is a method used for identification of human skeletal remains (namely the skull). Three-dimensional reconstruction software is capable of producing faster and more flexible facial reconstructions.
Attached to this instruction form is a number of tests comprising computer generated facial reconstructions.

Each test consists of TWO rows of facial images

- The upper row consists of ONE photograph of a TARGET individual.
- The bottom row consists of FOUR computer generated facial images.

Please

- Tell us your:
  - Age:……………………………………………………………………………….
  - Sex:………………………………………………………………………………..
  - Ethnic Background:……………………………………………………………….

- Look carefully at the available images.
- Select **ONLY ONE** image from the bottom row (by circling the relevant image letter) that you think it closely matches the photograph in the upper row.
- Do not discuss your answers with anyone else.

Before you start read the following guidance:

Guidelines for your facial reconstruction choice

In assessing the facial images you should pay particular attention to the general shape of the face, chin, forehead, cheeks and the general overall impression of their combination. Individual features such as the nose, lips and eyes are of lesser importance because the underlying skull provides insufficient information to enable these structures to be accurately reconstructed.

APPENDIX 5-A: AN EXAMPLE OF THE FACE RESEMBLANCE TEST FORM (A) FOR SKULL CASE (I)

APPENDIX 5-B: THE 3D FORENSIC FACIAL RECONSTRUCTION EXPERIMENT (PAPER INSTRUCTION FORM)-FACE POOL TEST FORM (A)

3D Forensic Facial Reconstruction Experiment

(Instruction form)

Forensic Facial Reconstruction is a method used for identification of human skeletal remains (namely the skull). Three-dimensional reconstruction software is capable of producing faster and more flexible facial reconstructions.
Attached to this instruction form is a test comprising computer generated facial reconstructions.
Each test consists of TWO SEPARATE rows of facial images.
Please

- Tell us your:

  Age:……………………………………………………………………………………
  Sex: ……………………………………………………………………...……
  Ethnic Background:…………………………………………………………...……

- Look carefully at the available images.
- Please rate the resemblance between the TWO images in each row.
- Use a scale from 1 to 10, where **1** carries **NO** resemblance and **10** carries the **highest** resemblance.
- Write your answer in the relevant area at the RIGHT of each row.
- Do not discuss your answers with anyone else.

Before you start read the following guidance:

Guidelines for your facial reconstruction choice
In assessing the facial images you should pay particular attention to the general shape of the face, chin, forehead, cheeks and the general overall impression of their combination. Individual features such as the nose, lips and eyes are of lesser importance because the underlying skull provides insufficient information to enable these structures to be accurately reconstructed.

APPENDIX 6-A: THE SEGMENTED SKULLS OF THE STUDIED EGYPTIAN CASES IN PART TWO OF EXPERIMENT ONE

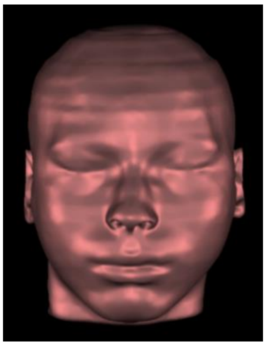

Egyptian Skull-01

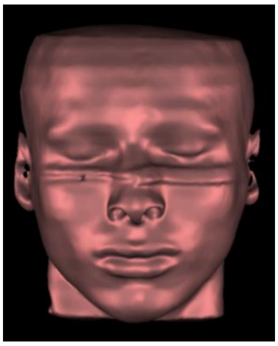

Egyptian Skull-02


Egyptian Skull-03

APPENDIX 6-B: THE SINGLE EGYPTIAN FACIAL TEMPLATES USED IN PART TWO OF EXPERIMENT ONE

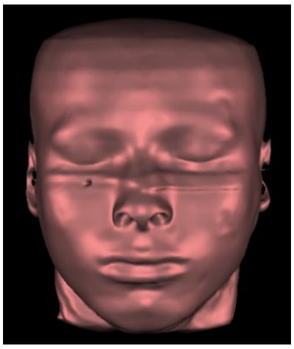

Egyptian Single Template-01

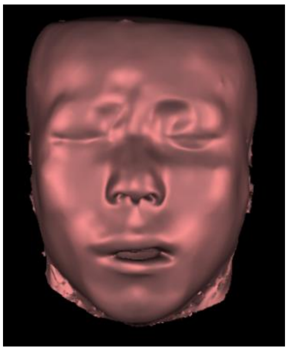

Egyptian Single Template-02


Egyptian Single Template-03

APPENDIX 6-C: THE AVERAGE EGYPTIAN FACIAL TEMPLATES USED IN PART TWO OF EXPERIMENT ONE


Average Template (01+02)


Average Template (02+03)


Average Template (01+03)

APPENDIX 7-A: AN EXAMPLE OF THE 3 FORMS OF THE FACE POOL TEST (B) USED FOR ONE OF THE 7 FACIAL RECONSTRUCTIONS OF SKULL CASE (I)

APPENDIX 7-B: THE 3D FORENSIC FACIAL RECONSTRUCTION EXPERIMENT (PAPER INSTRUCTION FORM)-FACE POOL TEST FORM (B)

3D Forensic Facial Reconstruction Experiment

(Instruction form)

Forensic Facial Reconstruction is a method used for identification of human skeletal remains (namely the skull). Three-dimensional reconstruction software is capable of producing faster and more flexible facial reconstructions.
Attached to this instruction form is a test comprising computer generated facial reconstructions.

Each test consists of:

- ONE facial reconstruction of a TARGET individual in the middle of the test.
- Eight photographs of different facial images.

Please

- Tell us your:
  - Age:……………………………………………………………………………….
  - Sex: ……………………………………………………………………………….
  - Ethnic Background:……………………………………………………………….

- Look carefully at the available images.
- Select **ONLY ONE** image form the 8 photographs (by circling the relevant image letter) that you think it closely matches the TARGET reconstructed face.
- Do not discuss your answers with anyone else.

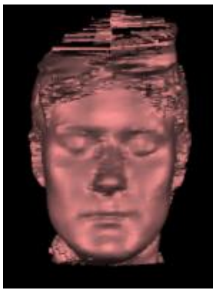
Before you start read the following guidance:

Guidelines for your facial reconstruction choice
In assessing the facial images you should pay particular attention to the general shape of the face, chin, forehead, cheeks and the general overall impression of their combination. Individual features such as the nose, lips and eyes are of lesser importance because the underlying skull provides insufficient information to enable these structures to be accurately reconstructed.

APPENDIX 8: THE FACIAL TEMPLATES AND THE FACIAL RECONSTRUCTION OF FEMALE CASE SKULL (IV) IN THE PILOT STUDY, EXPERIMENT THREE, PART ONE

20Y-FAV Facial Template

Facial Reconstruction using 20Y-FAV Facial Template



30Y-FAV Facial Template

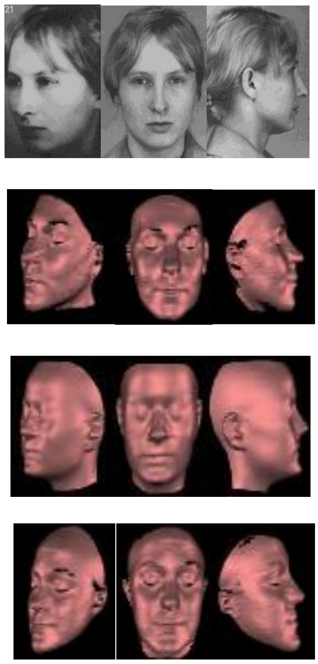
Facial Reconstruction using 30Y-FAV Facial Template

APPENDIX 9-A: FACE POOL TESTS FORM (A) (FRONTAL VIEW) FOR SKULL CASE (IV)

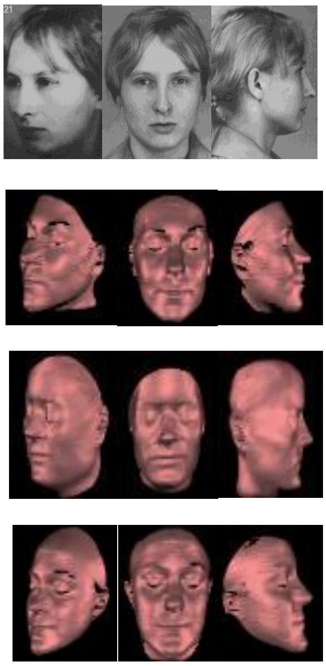
Testing the Facial Reconstruction of the target using the 20Y-FAV Facial Template

Testing the Facial Reconstruction of the target using the 30Y-FAV Facial Template

APPENDIX 9-B: FACE POOL TESTS FORM (A) (THREE VIEWS) FOR SKULL CASE (IV)



Testing Facial Reconstruction of the target using the 20Y-FAV Facial Template



Testing Facial Reconstruction of the target using the 30Y-FAV Facial Template

APPENDIX 9-C: 3D FORENSIC FACIAL RECONSTRUCTION EXPERIMENT (PAPER INSTRUCTION FORM)-FACE POOL TESTS FORM (A)

3D Forensic Facial Reconstruction Experiment

(Instruction form)

Forensic Facial Reconstruction is a method used for identification of human skeletal remains (namely the skull). Three-dimensional reconstruction software is capable of producing faster and more flexible facial reconstructions.
Attached to this instruction form is a test comprising computer generated facial reconstructions.

Each test consists of TWO rows of facial images.

- The TOP row consists of ONE facial photograph of a TARGET individual.
- The lower row(s) contains facial images of different individuals including a facial reconstruction of the TARGET.

Please

- Tell us your:
  - Age:…………………………………………………………………………….
  - Sex: …………………………………………………………………………….
  - Ethnic Background:…………………………………………………………….

- Look carefully at the available images.
- Select **ONLY ONE** image form the lower row(s) (by circling the relevant image letter) that you think it closely matches the TARGET reconstructed face.
- Do not discuss your answers with anyone else.

Before you start read the following guidance:

Guidelines for your facial reconstruction choice

In assessing the facial images you should pay particular attention to the general shape of the face, chin, forehead, cheeks and the general overall impression of their combination. Individual features such as the nose, lips and eyes are of lesser importance because the underlying skull provides insufficient information to enable these structures to be accurately reconstructed.

APPENDIX 10-A: FACE POOL TESTS FORM (B) (FRONTAL VIEW) FOR SKULL CASE (IV)-FACE POOL TESTS FORM (A) FOR SKULL CASE (IV)

Testing Facial Reconstruction of the target using the 20Y-FAV Facial Template

Testing Facial Reconstruction of the target using the 30Y-FAV Facial Template

APPENDIX 10-B: FACE POOL TESTS FORM (B) (THREE VIEWS) FOR SKULL CASE (IV)

Testing Facial Reconstruction of the target using the 20Y-FAV Facial Template

Testing Facial Reconstruction of the target using the 30Y-FAV Facial Template

APPENDIX 10-C: 3D FORENSIC FACIAL RECONSTRUCTION EXPERIMENT (PAPER INSTRUCTION FORM)-FACE POOL TESTS FORM (B)

3D Forensic Facial Reconstruction Experiment

(Instruction form)

Forensic Facial Reconstruction is a method used for identification of human skeletal remains (namely the skull). Three-dimensional reconstruction software is capable of producing faster and more flexible facial reconstructions.
Attached to this instruction form is a test comprising computer generated facial reconstructions.

Each test consists of TWO rows of facial images.

- The TOP row consists of ONE facial reconstruction of a TARGET individual.
- The lower row(s) contains photographs of different individuals including the TARGET.

Please

- Tell us your:
  - Age:…………………………………………………………………………….
  - Sex: ……………………………………………………………………………
  - Ethnic Background:……………………………………………………………
- Look carefully at the available images.
- Select **ONLY ONE** image form the lower row(s) (by circling the relevant image letter) that you think it closely matches the TARGET reconstructed face.
- Do not discuss your answers with anyone else.

Before you start read the following guidance:

Guidelines for your facial reconstruction choice

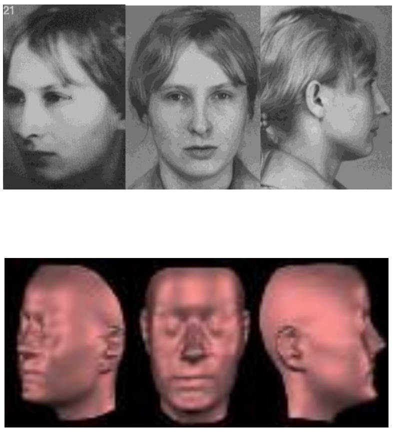
In assessing the facial images you should pay particular attention to the general shape of the face, chin, forehead, cheeks and the general overall impression of their combination. Individual features such as the nose, lips and eyes are of lesser importance because the underlying skull provides insufficient information to enable these structures to be accurately reconstructed.

APPENDIX 11-A: FACE RESEMBLANCE TESTS FORM (A) (FRONTAL VIEW) FOR SKULL CASE (IV)

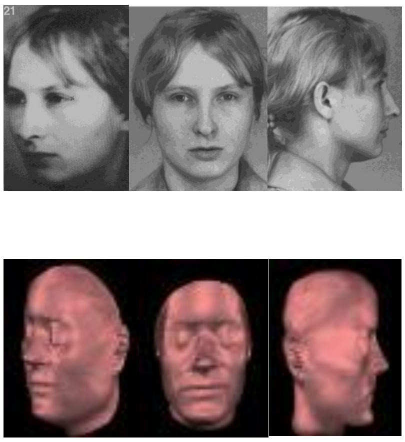
Testing Facial Reconstruction of the target using the 20Y-FAV Facial Template

Testing Facial Reconstruction of the target using the 30Y-FAV Facial Template

APPENDIX 11-B: THE FACE RESEMBLANCE TESTS FORM (A) (THREE VIEWS) FOR SKULL CASE (IV)



Testing Facial Reconstruction of the target using the 20Y-FAV Facial Template



Testing Facial Reconstruction of the target using the 30Y-FAV Facial Template

APPENDIX 11-C: THE 3D FORENSIC FACIAL RECONSTRUCTION EXPERIMENT (PAPER INSTRUCTION FORM)-FACE RESEMBLANCE TESTS FORM (A)

3D Forensic Facial Reconstruction Experiment

(Instruction form)

Forensic Facial Reconstruction is a method used for identification of human skeletal remains (namely the skull). Three-dimensional reconstruction software is capable of producing faster and more flexible facial reconstructions.
Attached to this instruction form is a test comprising computer generated facial reconstructions.

Please

- Tell us your:
  - Age:…………………………………………………………………………….
  - Sex: …………………………………………………………………………….
  - Ethnic Background:…………………………………………………………….

- Look carefully at the available images.
- Please rate the resemblance between the TWO images in each row.
- Use a scale from 1 to 10, where **1** carries **NO** resemblance and **10** carries the **highest** resemblance.
- Write your answer in the relevant area at the RIGHT of ach row.
- Do not discuss your answers with anyone else.

Before you start read the following guidance:

Guidelines for your facial reconstruction choice

In assessing the facial images you should pay particular attention to the general shape of the face, chin, forehead, cheeks and the general overall impression of their combination. Individual features such as the nose, lips and eyes are of lesser importance because the underlying skull provides insufficient information to enable these structures to be accurately reconstructed.

APPENDIX 12: PILOT STUDY TABLES

Table 16: Cranial landmarks' location description adopted from Rhine and Moore (1982) and Helmer (1984) (Vanezis, 2008)

Anatomical Landmarks	Description
Midline Landmarks	
Supraglabella	The foremost point in the midline, above glabella.
Glabella	The most forward projecting point of the forehead in the midline at the level of the Supraorbital ridges.
Nasion	The midline of the naso-frontal suture.
Rhinion	The end of the nasal bone at the junction between bone and cartilage of the nose.
Subspinale	The midline of the intranasal depression, below the nasal spine.
Supradentale	The jaw Centre, between the upper incisive teeth.
Infradentale	The jaw Centre, between the lower incisive teeth.
Supramentale	The most posterior midline point, above the chin in the jaw between the infradentale and the pogonion.
Pogonion (Mental Eminence)	The most prominent point of the chin.
Gnathion	Lowest point of the chin.
Bilateral Landmarks	
Frontal Eminence	Centered on pupil, most anterior point of the forehead at the level of the supraglabella.
Supraorbital Centre	The centre point of the upper margin of the orbit.
Suborbital Center	The centre point of the lower margin of the orbit.
Maxillo-Malar (Inferior Malar)	Centered on pupil, just interior to zygomatic Process.
Malar-Orbit Level	Lined up with the lateral border of the eye on the centre of the zygomatic process
Zygion	Most lateral curvature of the zygomatic bone
Supraglenoid	Root of the zygomatic arch just above and forward the acoustic meatus (ear canal).
Gonion	The outer margin of the angle of the mandible.
Supra M2	Above the second upper molar.
Occlusal Line	Point in the jaw in the plane of dental occlusion
SubM2	Below the second lower molar, horizontally lined up with supra-M2.
Ectoconchion	Bony projection of the ectocranial surface of the frontal bone, vertically centred on the orbit, next to the lateral orbital border.
Alare Level (Supracanine)	Vertically lined up with the cheilion, on the horizontal level of the chin–lip fold
Cheilion Level	On the midpoint of the Canine (1st Premolar) tooth.
Stephanion	The point on the side of the skull where the point where the coronal suture crosses the superior temporal line.

Table 17: Soft Tissue Depth Measurements Thickness (in mm) of an adult Caucasian male and female of average build at Cranial Landmarks Location, After Rhine and Moore (1982) and Helmer (1984) (Vanezis, 2008)

Anatomical Landmarks	**Number**	**Male Soft Tissue Depth (mm)**	**Female Soft Tissue Depth (mm)**
		Midline Landmarks	
Supraglabella (sg)	1	04.25	03.50
Glabella (g)	1	05.25	04.75
Nasion (n)	1	06.50	05.50
Rhinion (rhi)	1	3.00	02.75
Subspinale (sn)	1	10.00	08.50
Supradentale (ls)	1	9.75	09.00
Infradentale (li)	1	11.00	10.00
Supramentale (mls)	1	10.75	09.50
Pogonion (pg)	1	11.25	10.00
Gnathion (gn)	1	07.25	05.75
		Bilateral Landmarks	
*Frontal Eminence	2	04.25	03.50
Supraorbital (mso)	2	08.25	07.00
Suborbital (mio)	2	05.75	06.00
*Maxillo-Malar (InferiorMalar)	2	13.25	12.75
*Malar-Orbit Level	2	10.00	10.75
Zygion (zy)	2	07.25	07.50
*Supraglenoid	2	08.50	08.00
Gonion (go)	2	11.50	12.00
Supra M2 (sM^2)	2	19.50	19.25
Occlusal Line	2	18.25	17.00
SubM2 (iM^2)	2	16.00	15.50
*Ectoconchion	2	05.50	04.50
Alare Level (Supracanine) (acp)	2	12.35	11.22
*Cheilion Level (Canine-1stPM)	2	18.50	18.60
*Stephanion	2	04.00	04.00
TOTAL	**40**		

* Landmarks with no corresponding measurements in Stephan (2014)

Table 18: Soft tissue depth measurements (in mm), as weighted mean with SD, from adult T-tables of Stephan (2014)

Anatomical Landmarks	Soft Tissue Depth (mm)
Midline Landmarks	
Opisthocranion (op-op)	06.00
*Vertex (v)	05.00
Glabella (g)	05.50
Nasion (n)	06.50
*Midnasal (mn)	04.50
Rhinion (rhi)	03.00
Subnasale (sn)	13.00
*Mid-philtrum (mp)	11.00
Labrale superius (ls)	11.50
Labrale inferius (li)	13.00
Mentolabial sulcus (mls)	11.00
Pogonion (pg)	11.00
Gnathion (gn)	07.50
*Menton (m)	07.50
Bilateral Landmarks	
Mid-supraorbital (mso)	06.50
Suborbital (mio)	07.00
Zygion (zy)	06.50
Gonion (go)	11.50
Mid-ramus (mr)	18.50
*Mid-mandibular border (mmb)	11.50
Supra M2 (sM^2)	25.50
SubM2 (iM^2)	19.00
Alare curvature point (acp)	09.00
*Supra canine (sC)	10.00
*Infra canine (iC)	11.00

* Landmarks with no corresponding measurements in Rhine and Moore (1982) and Helmer (1984)

Table 19: Modified soft tissue depth measurements (in mm) at cranial landmarks location based on Rhine and Moore, 1982; and Helmer, 1984 and on Stephan (2014)

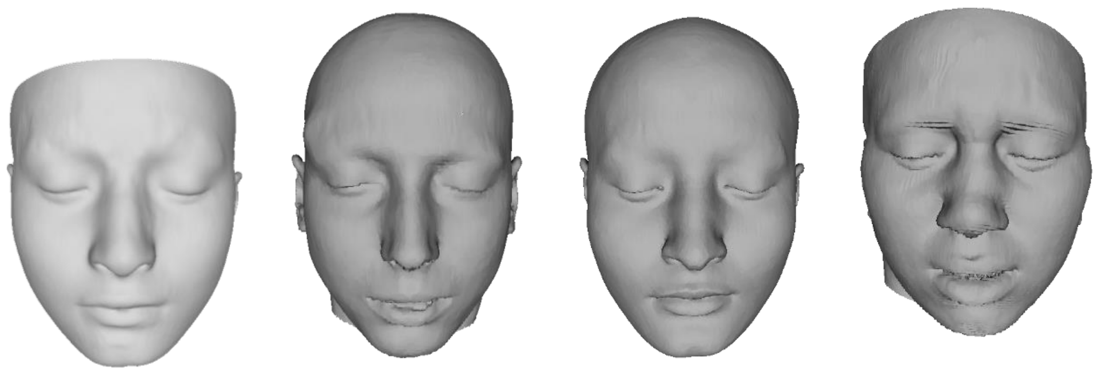
Landmarks names	**NEW (Modified) Soft Tissue Depths (mm)**
Midline Landmarks	
*Supraglabella = Opisthocranion	06.00
*Glabella	05.50
*Nasion	06.50
*Rhinion	03.00
*Subspinale = Subnasale	13.00
*Supradentale = Labrale superius	11.50
*Infradentale = Labrale inferius (li)	13.00
*Supramentale = Mentolabial sulcus	11.00
*Pogonion	11.00
*Gnathion	07.50
Bilateral Landmarks	
**Frontal Eminence	04.25
* Supraorbital (mso)	06.50
* Suborbital (mio)	07.00
**Maxillo-Malar(InferiorMalar)	13.25
**Malar-Orbit Level	10.00
* Zygion (zy)	06.50
**Supraglenoid	08.50
* Gonion (go)	11.50
**Ectoconchion	5.50
* Alare Level(Supracanine) (acp)	12.35
**Cheilion Level(Canine-1stPM)	18.50
**Stephanion	04.00

* Landmarks with measurements taken from Stephan (2014)

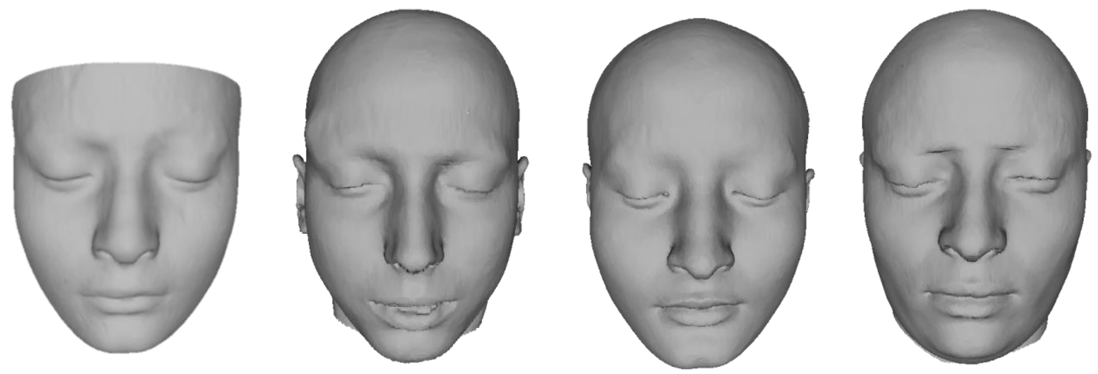
**Landmarks with measurements taken from Rhine and Moore (1982) and Helmer (1984)

APPENDIX 13: FORMATION OF THE AVERAGE FACIAL TEMPLATES

Female Average Faces



F-Av (16-20)Y-1 (F1+F02+F04) **F01** **F02** **F04**

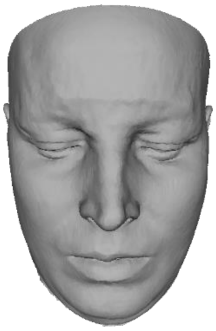


F-Av (16-20)Y-2 (F1+F02+F03) **F01** **F02** **F03**

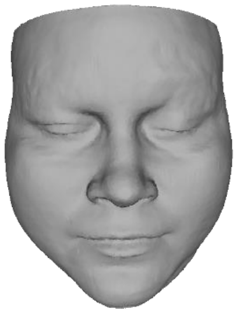
F-Av (21-30)Y (F07+F08+F10+F11) **F07** **F08** **F10** **F11**

F-Av (31-40)Y
(F12+F15+F17)

F12

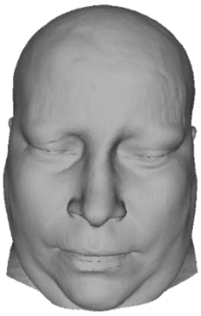


F15



F17

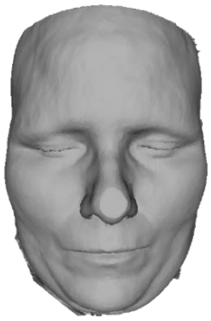
F-Av (41-50)Y
(F18+F19+F22+F24)



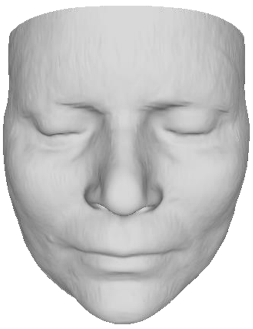
F18

F19

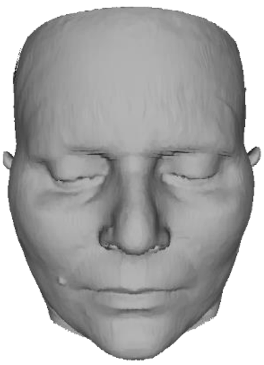
F22



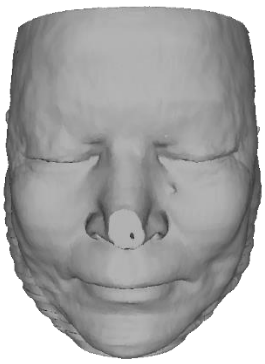
F24



F-Av (50+)Y
(F23+F27)



F23*



F27

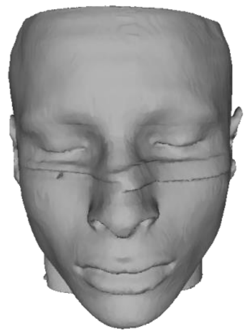
***(41-50)Y**

Male Average Faces

M-Av (16-20)Y
(M1+M03+M04)

M1

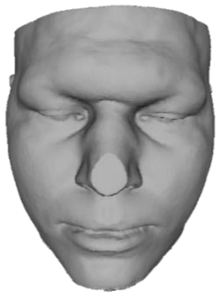
M03



M04

M-Av (21-30)Y
(M06+M08+M10+M13)

M06



M08

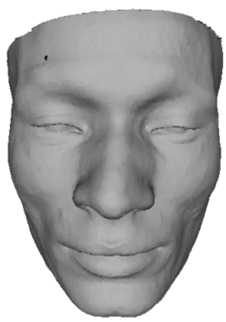
M10

M13

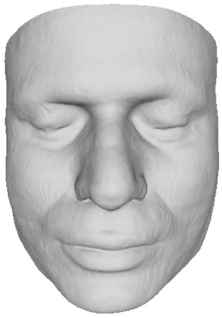
M-Av (31-40)Y
(M15+M16+M18)

M15

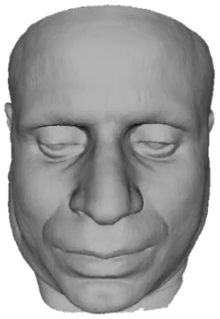
M16



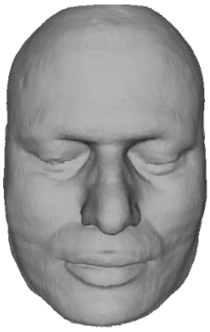
M18



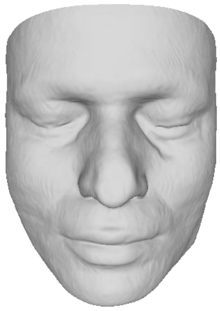
M-Av (41-50)Y
(M24+M25)



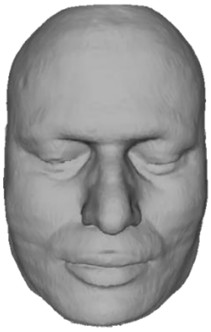
M24



M25

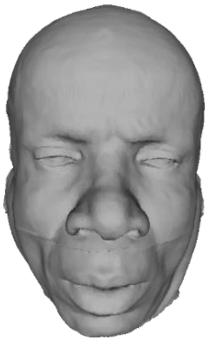


M-Av (50+)Y-1
(M25+M27+M29+M30)

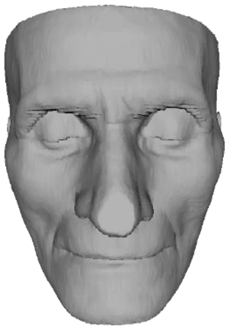


M25*

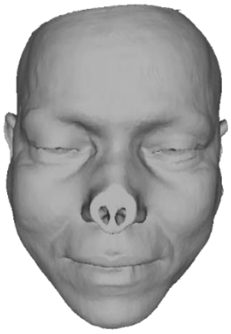
***(41-50)Y**



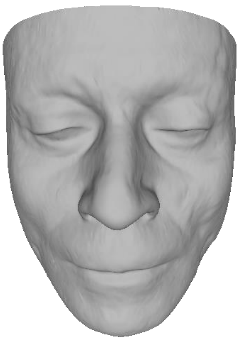
M27



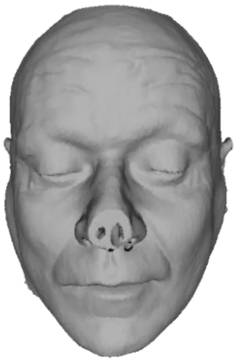
M29



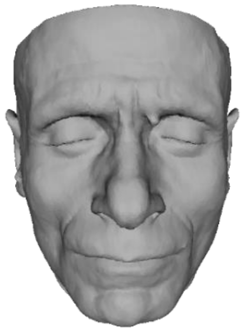
M30



M-Av (50+)Y-1
(M31+M32+M33)



M31



M32

M33

APPENDIX 14: THE VOLUNTEER CONFIDENTIALITY AGREEMENT FOR ONLINE TESTS

Volunteer Confidentiality Agreement

Description (optional)

This agreement applies to all volunteer assessors associated with and/or involved in the activities or affairs of FFR / Forensic Facial Reconstruction using CT scan research project. All data, materials, knowledge and information generated through, originating from, or having to do with FFR or persons associated with our activities, including patients, is to be considered privileged and confidential and is not to be disclosed to any third party. All pages, forms, information, photographs, documents, printed matter, policies and procedures, conversations, messages (received or transmitted), resources, contacts, e-mail lists, e-mail messages, client, staff or public information is confidential and the sole property of Queen Mary University London.

Please insert your name and email in the text box below. We are obligated not to share this information with any third party without your permission.

This agreement applies to all volunteer assessors associated with and/or involved in the activities or affairs of FFR / Forensic Facial Reconstruction using CT scan research project. All data, materials, knowledge and information generated through, originating from, or having to do with FFR or persons associated with our activities, including patients, is to be considered privileged and confidential and is not to be disclosed to any third party. All pages, forms, information, photographs, documents, printed matter, policies and procedures, conversations, messages (received or transmitted), resources, contacts, e-mail lists, e-mail messages, client, staff or public information is confidential and the sole property of Queen Mary University London.

Please insert your name and email in the text box below. We are obligated not to share this information with any third party without your permission.

APPENDIX 15-A: An Online Face Pool Assessment Test Example

18- Target Face

18- Test Face(A)

18- Test Face(B)

18- Test Face(C)

18- Test Face(D)

APPENDIX 15-B: THE ONLINE FACE POOL ASSESSMENT TEST INSTRUCTIONS

Face Pool Assessment Test Instructions

How the exercise works:

- Each session consists of 5 exercises, one for each case.
- Each exercise consists of a playlist of 3D models of:
  - One Target Face, labelled (Target Face).
  - Four Test Faces, labelled (Test Faces A-D).
  - One collective model of all five faces together, labelled (Face Pool).
- Please use the password provided to access each model.
- Using an ordinary mouse you can:
  - Rotate the model by pressing and holding the left mouse button.
  - Pan/move the model by pressing and holding the middle mouse button.
  - Zoom in and out using the middle mouse button.
- Please examine faces within their individual and collective.
- After careful assessment of each case, pick ONLY ONE face from the 4 test faces that you think it closely resembles the target face.
- Please complete the appropriate section of each case in this electronic answer/assessment survey.
- Please do not share your answers with other candidate(s) performing the same exercise before the end of the exercise.

Important Note for the exercise:

Please make your assessment primarily based on the general shape of the facial bones (e.g. forehead, orbits "eye bones", cheek bones, temple bone "the bone on the sides of the head", chin, etc...), rather than the face soft structures (e.g. nose, lips, etc...).

For further enquiries, please contact me on the email below.

Thank you for your participation in this study.

Dalia Abdou
d.a.abdou@qmul.ac.uk

How the exercise works:

- Each session consists of 5 exercises, one for each case.
- Each exercise consists of a playlist of 3D models of:
  - One Target Face, labelled (Target Face).
  - Four Test Faces, labelled (Test Faces A-D).
  - One collective model of all five faces together, labelled (Face Pool).
- Please use the password provided to access each model.
- Using an ordinary mouse you can:
  - Rotate the model by pressing and holding the left mouse button.
  - Pan/move the model by pressing and holding the middle mouse button.

- o Zoom in and out using the middle mouse button.
- • Please examine faces within their individual and collective.
- • After careful assessment of each case, pick ONLY ONE face from the 4 test faces that you think it closely resembles the target face.
- • Please complete the appropriate section of each case in this electronic answer/assessment survey.
- • Please do not share your answers with other candidate(s) performing the same exercise before the end of the exercise.

Important Note for the exercise:
Please make your assessment primarily based on the general shape of the facial bones (e.g. forehead, orbits "eye bones", cheek bones, temple bone "the bone on the sides of the head", chin, etc...), rather than the face soft structures (e.g. nose, lips, etc...).

For further enquiries, please contact me on the email below.

Thank you for your participation in this study.

Dalia Abdou
d.a.abdou@qmul.ac.uk

APPENDIX 16-A: AN ONLINE FACE RESEMBLANCE ASSESSMENT TEST EXAMPLE

APPENDIX 16-B: THE ONLINE FACE RESEMBLANCE ASSESSMENT TEST-V1 INSTRUCTIONS

Face Resemblance Assessment Test-V1 Instructions

How the exercise works:

• Each exercise consists of 15 tests, one for each case.

• Each test consists of a 3D model of 2 faces:
o Top Face, labelled (Real Face).
o Bottom Face, labelled (Reconstructed Face).

• Please use the password provided to access each model.

• Using an ordinary mouse you can:
o Rotate the model by pressing and holding the left mouse button.
o Pan/move the model by pressing and holding the middle mouse button.
o Zoom in and out using the middle mouse button.

• Please examine the two faces of each case.

• Give an overall score of resemblance (0-10) according to the degree of resemblance between the real and the reconstructed faces, where (0 = No Resemblance) and (10 = Highest Resemblance).

• Please complete the appropriate section of each case in this electronic answer/assessment survey.

• Please do not share your answers with other candidate(s) performing the same exercise before the end of the exercise.

Important Note for the exercise:

Please make your assessment primarily based on the general shape of the facial bones (e.g. forehead, orbits "eye bones", cheek bones, temple bone "the bone on the sides of the head", chin, etc...), rather than the face soft structures (e.g. nose, lips, etc...).

For further enquiries, please contact me on the email below.

Thank you for your participation in this study.

Dalia Abdou
d.a.abdou@qmul.ac.uk

How the exercise works:

• Each exercise consists of 15 tests, one for each case.
• Each test consists of a 3D model of 2 faces:
o Top Face, labelled (Real Face).
o Bottom Face, labelled (Reconstructed Face).
• Please use the password provided to access each model.
• Using an ordinary mouse you can:
o Rotate the model by pressing and holding the left mouse button.
o Pan/move the model by pressing and holding the middle mouse button.
o Zoom in and out using the middle mouse button.
• Please examine the two faces of each case.
• Give an overall score of resemblance (0-10) according to the degree of resemblance between the real and the reconstructed faces, where (0 = No Resemblance) and (10 = Highest Resemblance).

- Please complete the appropriate section of each case in this electronic answer/assessment survey.
- Please do not share your answers with other candidate(s) performing the same exercise before the end of the exercise.

Important Note for the exercise:

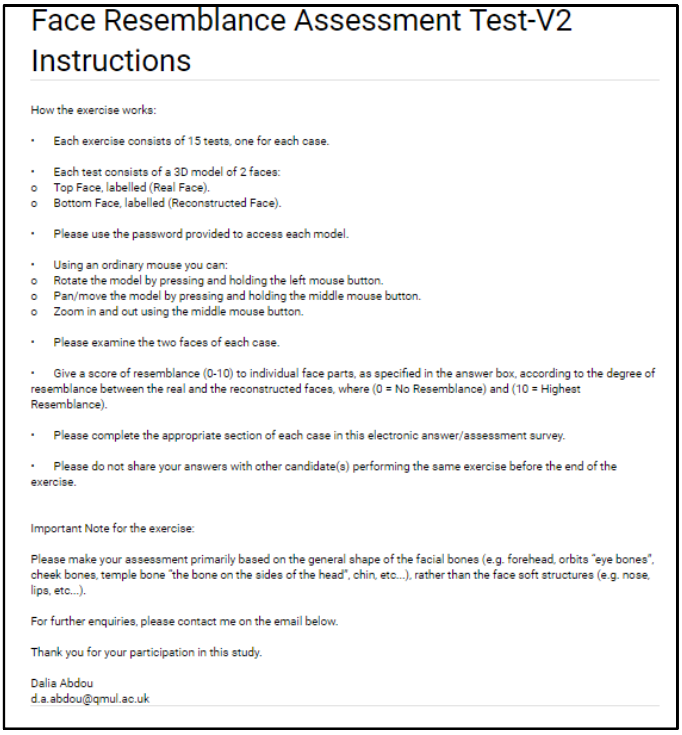
Please make your assessment primarily based on the general shape of the facial bones (e.g. forehead, orbits "eye bones", cheek bones, temple bone "the bone on the sides of the head", chin, etc...), rather than the face soft structures (e.g. nose, lips, etc...).

For further enquiries, please contact me on the email below.

Thank you for your participation in this study.

Dalia Abdou
d.a.abdou@qmul.ac.uk

APPENDIX 16-C: THE ONLINE FACE RESEMBLANCE ASSESSMENT TEST-V2 INSTRUCTIONS



Face Resemblance Assessment Test-V2 Instructions

How the exercise works:

• Each exercise consists of 15 tests, one for each case.

• Each test consists of a 3D model of 2 faces:
o Top Face, labelled (Real Face).
o Bottom Face, labelled (Reconstructed Face).

• Please use the password provided to access each model.

• Using an ordinary mouse you can:
o Rotate the model by pressing and holding the left mouse button.
o Pan/move the model by pressing and holding the middle mouse button.
o Zoom in and out using the middle mouse button.

• Please examine the two faces of each case.

• Give a score of resemblance (0-10) to individual face parts, as specified in the answer box, according to the degree of resemblance between the real and the reconstructed faces, where (0 = No Resemblance) and (10 = Highest Resemblance).

• Please complete the appropriate section of each case in this electronic answer/assessment survey.

• Please do not share your answers with other candidate(s) performing the same exercise before the end of the exercise.

Important Note for the exercise:

Please make your assessment primarily based on the general shape of the facial bones (e.g. forehead, orbits "eye bones", cheek bones, temple bone "the bone on the sides of the head", chin, etc…), rather than the face soft structures (e.g. nose, lips, etc…).

For further enquiries, please contact me on the email below.

Thank you for your participation in this study.

Dalia Abdou
d.a.abdou@qmul.ac.uk

How the exercise works:

• Each exercise consists of 15 tests, one for each case.
• Each test consists of a 3D model of 2 faces:
o Top Face, labelled (Real Face).
o Bottom Face, labelled (Reconstructed Face).
• Please use the password provided to access each model.
• Using an ordinary mouse you can:
o Rotate the model by pressing and holding the left mouse button.
o Pan/move the model by pressing and holding the middle mouse button.
o Zoom in and out using the middle mouse button.
• Please examine the two faces of each case.
• Give a score of resemblance (0-10) to individual face parts, as specified in the answer box, according to the degree of resemblance between the real and the reconstructed faces, where (0 = No Resemblance) and (10 = Highest Resemblance).
• Please complete the appropriate section of each case in this electronic answer/assessment survey.
• Please do not share your answers with other candidate(s) performing the same exercise before the end of the exercise.

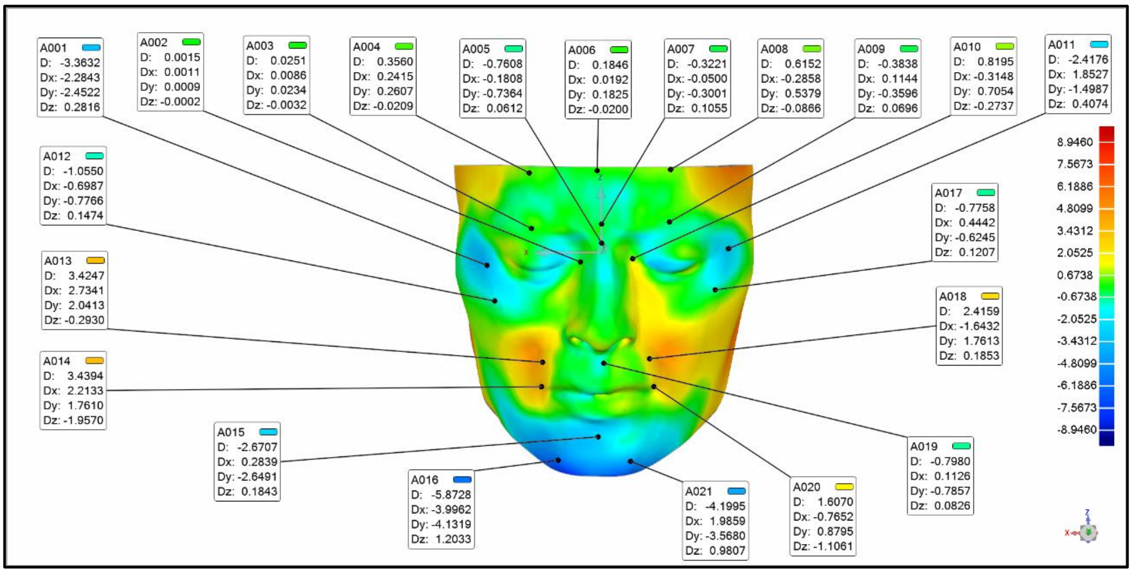
Important Note for the exercise:
Please make your assessment primarily based on the general shape of the facial bones (e.g. forehead, orbits "eye bones", cheek bones, temple bone "the bone on the sides of the head", chin, etc...), rather than the face soft structures (e.g. nose, lips, etc...).

For further enquiries, please contact me on the email below.

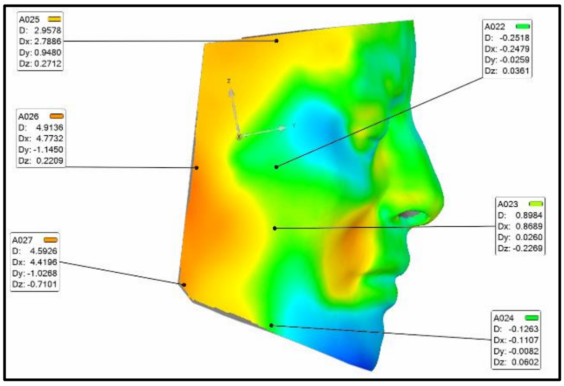
Thank you for your participation in this study.

Dalia Abdou
d.a.abdou@qmul.ac.uk

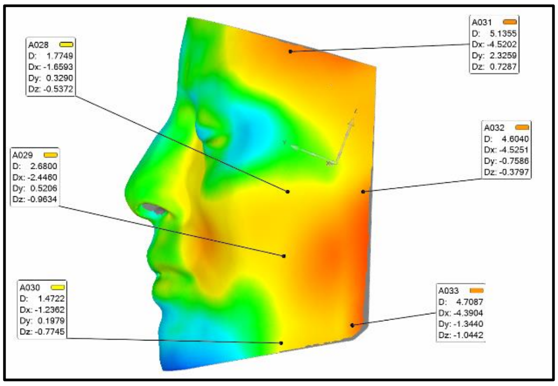
APPENDIX 17: SURFACE DISTANCE COMPARISON HISTOGRAM



A



b



c

A histogram (colour map) generated by (Geomagic Control®) software program, showing the difference in colour between aligned real and reconstructed faces at several points (a: frontal view, b: right view & c: left profile view).

The facial regions are represented by a number of points labelled on the figure above starting with (A00) as follows:

- Forehead: points A001- A0011, A0025 & A0031.
- Orbit (right side): points A001-3 & A0012.
- Orbit (left side): points A009-10 & A0017.
- Nasal region: points A002, A005 & A0010.
- Cheek (right side): points A001, A0012 & A0022.
- Cheek (left side): points A0011, A0017 & A0028.
- Chin region: points A0015, A0016 & A0021.
- Mouth: points A0013-14 & A0018-20.
- Jaw (right side): points A0023, A0024, A0026 & A0027.
- Jaw (left side): points A0029, A0030, A0032 & A0033.

APPENDIX 18: THE MAIN STUDY TABLES

Table 20: The 61 Egyptian male and female head CT scans used in the study with individual ages; males in the left table, and females in the right table

No.	Code	Sex	Age
1.	M1	M	17
2.	M2	M	17
3.	M3	M	18
4.	M4	M	18
5.	M5	M	21
6.	M6	M	21
7.	M7	M	22
8.	M8	M	25
9.	M9	M	27
10.	M10	M	28
11.	M11	M	30
12.	M12	M	30
13.	M13	M	30
14.	M14	M	34
15.	M15	M	35
16.	M16	M	35
17.	M17	M	35
18.	M18	M	36
19.	M19	M	37
20.	M20	M	38
21.	M21	M	38
22.	M22	M	38
23.	M23	M	48
24.	M24	M	48
25.	M25	M	52
26.	M26	M	53
27.	M27	M	54
28.	M28	M	55
29.	M29	M	56
30.	M30	M	57
31.	M31	M	57
32.	M32	M	60
33.	M33	M	63
34.	M34	M	67

No.	Code	Sex	Age
1.	F1	F	16
2.	F2	F	20
3.	F3	F	20
4.	F4	F	20
5.	F5	F	24
6.	F6	F	24
7.	F7	F	24
8.	F8	F	27
9.	F9	F	27
10.	F10	F	28
11.	F11	F	30
12.	F12	F	32
13.	F13	F	34
14.	F14	F	37
15.	F15	F	37
16.	F16	F	38
17.	F17	F	41
18.	F18	F	42
19.	F19	F	43
20.	F20	F	43
21.	F21	F	45
22.	F22	F	48
23.	F23	F	48
24.	F24	F	48
25.	F25	F	58
26.	F26	F	60
27.	F27	F	62

Table 21: The classified age groups of the Egyptian male and female head CT scans used in the study

Age Group (y)	Male	Assigned Number	Female	Assigned Number	Total
16 - 20	4	M1, M2, M3, M4	4	F1, F2, F3, F4	**8**
21 – 30	9	M5, M6, M7, M8, M9, M10, M11, M12, M13	7	F5, F6, F7, F8, F9, F10, F11	**16**
31 – 40	9	M14, M15, M16, M17, M18, M19, M20, M21, M22	5	F12, F13, F14, F15, F16	**14**
41 – 50	2	M23, M24	8	F17, F18, F19, F20, F21, F22, F23, F24	**10**
>50	10	M25, M26, M27, M28, M29, M30, M31, M32, M33, M34	3	F25, F26, F27	**13**
Total	**34**		**27**		**61**

Table 22: The male cases used for the facial reconstruction and for generating the average faces and as studied cases

Age Group (y)	Skulls' face reconstructed (Target Case)	Faces used for producing an averaged face	Average Face
16 - 20	M02	M01, M03, M04	Av-M(16 – 20)
21 – 30	M05, M07, M09, M11, M12	M06, M08, M10, M13	Av-M(21 – 30)
31 – 40	M14, M17, M19, M20, M21, M22	M15, M16, M18	Av-M(31 – 40)
41 – 50	M23	M24, M25*	Av-M(41 – 50)
> 50	M26, M28, M31	M25, M27, M29, M30	Av-M(> 50)-1
	M34	M31, M32, M33	Av-M(> 50)-1

* A face template was used "borrowed" from a neighbouring age group.

Table 23: The female cases used for the facial reconstruction and for generating the average faces and as studied cases

Age Group (y)	Skulls' face reconstructed (Target Case)	Faces used for producing an averaged face	Average Face
16 - 20	F03**	F01, F02, F04**	Av-F(16 – 20)
	F04**	F01, F02, F03**	Av-F(16 – 20)
21 – 30	F05, F06, F09	F07, F08, F10, F11	Av-F(21 – 30)
31 – 40	F13, F14	F12, F15, F16	Av-F(31 – 40)
41 – 50	F17, F20, F21, F23	F18, F19, F22, F24	Av-F(41 – 50)
> 50	F25, F26	F23*, F27	Av-F(> 50)

* A face template was used "borrowed" from a neighbouring age group.

** A case where face was reconstructed in one situation and used a facial template (to generate an average face) in another situation.

Table 24: The recruited experts, their type of experience, duration of experience and their institutions

Expert	Experience	Duration*	Institution
Experts Group One (Exp. 01 – 03): Experts in facial identification/perception psychology (n = 3)			
Exp.01	Cognitive Psychology Research and Teaching	8 years	Department of Psychology, University of Bedfordshire, UK
Exp.02	PhD Psychology	12 years	
Exp.03	PhD and research in face perception and recognition	16 years	Edinburgh Napier University, UK
Experts Group Two (Exp. 04 – 14): Experts in forensic medicine/pathology (n = 11) - **Egyptian Pathologists (n=7): (Exp. 04 – 10)** - **Non-Egyptian Pathologists (n=4): (Exp. 11 – 14)**			
Exp.04	Forensic Pathology	23 years	Forensic Medicine and Clinical Toxicology Department, Zagazig Faculty of Medicine, Egypt
Exp.05		23 years	
Exp.06		12 years	
Exp.07		8 years	
Exp.08		6 years	
Exp.09		6 years	
Exp.10		4 years	
Exp.11	Forensic Pathology	10 years	University of Verona, Italy
Exp.12	Forensic Pathology	3 years	UK DMJ (Diploma in Medical Jurisprudence) trainees, Malaysia
Exp.13		N/A	
Exp.14		N/A	
Experts Group Three (Exp. 15 – 23): Experts in forensic anthropology (n = 9)			
Exp.15	Forensic Anthropology	1.5 years	Department of Life Sciences, University of Coimbra, Coimbra
Exp.16	Biological, Anthropology, Forensic Anthropology, Human Osteology.	8 years	
Exp.17	Forensic Anthropology	2 years	
Exp.18	Forensic Anthropology	3 years	
Exp.19	Forensic Anthropology and Forensic Pathology	1 year	
Exp.20	MSc Forensic Anthropology, PhD, Postdoctoral experience	5 years	Department of Anthropology, NMNH, Smithsonian Institution, Washington, DC, USA
Exp.21	PhD, Forensic Anthropology. Forensic Odontology.	N/A	University of Girona, Spain. University of Barcelona, Spain.
Exp.22	MSc - Forensic Anthropology; PhD Candidate in Forensic Anthropology	N/A	PhD Candidate in the University of Edinburgh
Exp.23	MSc - Forensic Anthropology; PhD Candidate in Forensic Anthropology	6 years	PhD Candidate in the University of Cambridge
Experts Group Four (Exp. 24 – 26): Experts in forensic anthropology and forensic facial reconstruction (n = 3)			
Exp.24	Forensic Anthropology, Forensic Pathology, Forensic Facial Reconstruction.	1.5 years, 9 months, 3 months.	Department of Life Sciences, University of Coimbra, Coimbra
Exp.25	Forensic Anthropology, Forensic Facial Reconstruction	20 years	Cameron Forensic Medical Sciences, QMUL
Exp.26	Forensic Pathology, Forensic Anthropology, Forensic Facial Reconstruction	28 years	Barts and the London School of Medicine and Dentistry, QMUL

*Duration from the start of the experiment in November 2015

Table 25: The design of the four face pool exercise, including the studied target case (with its sex and age), and the used foil faces

Face Pool Exercise	Case (Code)	Sex	Age (y)	Face Pool Foil Faces (same age group)	Face Pool Foil Faces (different age group)
One	01 (F03)	F	16 - 20	F02, F04	F08 (21 – 30)
	02 (M07)	M	21 – 30	M05, M06, M08	
	03 (F13)	F	31 – 40	F12, F15, F16	
	04 (M23)	M	41 – 50	M24	M25, M29 (> 50)
	05 (F23)	F	41 – 50	F19, F21, F24	
Two	06 (M02)	M	16 - 20	M01, M03	M06
	07 (F14)	F	31 – 40	F13, F16	F19 (41 – 50)
	08 (M22)	M	31 – 40	M16, M20, M21	
	09 (F25)	F	> 50	F26, Av-F(> 50)	F23 (41 – 50)
	10 (M11)	M	21 – 30	M09, M10, M12	
Three	11 (F09)	F	21 – 30	F06, F07, F08	
	12 (M19)	M	31 – 40	M17, M18, M21	
	13 (F26)	F	> 50	F24, F25, Av-F(> 50)	
	14 (M26)	M	> 50	M28, M29, M30	
	15 (F04)	F	16 - 20	F03	F05, F09 (21 – 30)
Four	16 (M34)	M	> 50	M29, M30, M31	
	17 (F05)	F	21 – 30	F08, F09, F11	
	18 (M14)	M	31 – 40	M16, M20	M08 (21 – 30)
	19 (F17)	F	41 – 50	F18, F21, F23	
	20 (M31)	M	> 50	M26, M30, M33	

Table 26: The number of observers participated in each face pool exercise

Subjective Test (1): Face Pool Test	Non-Experts		Experts		TOTAL
	Egyptian	Non- Egyptian	Egyptian	Non- Egyptian	
Exercise One	10	40	7	19	76
Exercise Two	9	34	7	18	68
Exercise Three	9	32	7	18	66
Exercise Four	8	32	7	18	65

Table 27: The design of the two face resemblance exercise, including the studied target case (with its sex and age**)**

Face Resemblance Exercise	Case (Code)	Sex	Age (y)
One	01 (M11)	M	21 – 30
	02 (F09)	F	21 – 30
	03 (M17)	M	31 – 40
	04 (F03)	F	16 - 20
	05 (M28)	M	> 50
	06 (F20)	F	41 – 50
	07 (M09)	M	21 – 30
	08 (F25)	F	> 50
	09 (M15)	M	31 – 40
	10 (M23)	M	41 – 50
	11 (M05)	M	21 – 30
	12 (F06)	F	21 – 30
	13 (M21)	M	31 – 40
	14 (F13)	F	31 – 40
	15 (M26)	M	> 50
Two	16 (F17)	F	41 – 50
	17 (M07)	M	21 – 30
	18 (F26)	F	> 50
	19 (M14)	M	31 – 40
	20 (M34)	M	> 50
	21 (M12)	M	21 – 30
	22 (F05)	F	21 – 30
	23 (M19)	M	31 – 40
	24 (F21)	F	41 – 50
	25 (M02)	M	16 - 20
	26 (F23)	F	41 – 50
	27 (M31)	M	> 50
	28 (F14)	F	31 – 40
	29 (M22)	M	31 – 40
	30 (F04)	F	16 - 20

Table 28: The number of observers participated in each face resemblance exercise

Subjective Test (2): Face Resemblance Test		Non-Experts			Experts Continuing after Face Pool Test		TOTAL
		Egyptian	Non- Egyptian		Egyptian	Non-Egyptian	
		Continuing after Face Pool Test	Continuing after Face Pool Test	Starting from Face Resemblance Test			
Group 1	Exercise 1: Test Version 1*	3	14	9	3	8	37
	Exercise 2: Test Version 2**	3	12	8	3	8	34
Group 2	Exercise 1: Test Version 2	5	14	6	3	8	36
	Exercise 2: Test Version 1	4	13	5	3	8	33

* Face Resemblance Test Version 1 (i.e. Giving Overall Scores).

** Face Resemblance Test Version 2 (i.e. Giving Individual Facial Regions Scores).

Table 29: The percentages of selection of each face in the face pool for each case. The identification percentage of the target face of each case is marked in red colour.

Case	A	B	C	D	Target	Target face ID	Non-target faces ID above target
F03	30.3%	19.7%	44.7%	5.3%	**C**	**YES**	
M07	2.6%	81.6%	13.2%	1.3%	**B**	**YES**	
F13	50%	21.1%	14.5%	14.5%	**B**	**NO**	**YES (n =1)**
M23	27.6%	3.9%	61.8%	6.6%	**C**	**YES**	
F23	7.9%	5.3%	17.1%	68.4%	**C**	**NO**	**YES (n =1)**
M02	16.2%	13.2%	47.1%	23.5%	**C**	**YES**	
F14	33.8%	17.6%	5.9%	42.6%	**D**	**YES**	
M22	16.2%	44.1%	16.2%	23.5%	**D**	**NO**	**YES (n =1)**
F25	33.8%	29.4%	8.82%	27.9%	**B**	**YES**	**YES (n =1)**
M11	17.6%	19.1%	20.6%	41.2%	**A**	**NO**	
F09	28.8%	25.8%	28.8%	16.7%	**B**	**YES**	**YES (n =2)**
M19	36.4%	9.1%	34.8%	19.7%	**B**	**NO**	**YES (n =2)**
F26	3%	7.6%	68.2%	21.2%	**C**	**YES**	
M26	9.1%	12.1%	25.8%	53%	**D**	**YES**	
F04	19.7%	60.6%	7.6%	12.1%	**A**	**NO**	**YES (n =1)**
M34	3.1%	47.7%	13.8%	35.4%	**B**	**YES**	
F05	10.8%	24.6%	9.2%	55.4%	**D**	**YES**	
M14	9.2%	49.2%	35.4%	6.2%	**B**	**YES**	
F17	20%	16.9%	35.4%	27.7%	**A**	**NO**	**YES (n =2)**
M31	16.9%	18.5%	30.8%	33.8%	**C**	**YES**	**YES (n =1)**

Table 30: The correct identification rate of each of the 20 cases given by:

- All participants (combined) (n=65-76/case).
- All participants with a professional experience in (forensic anthropology, with or without an experience in forensic facial reconstruction, forensic face identification psychology, and forensic pathology). This group includes Egyptian and non-Egyptians participants (n=25-26/case).
- All the non-expert participants. This group includes Egyptian and non-Egyptians participants (n=40-50/case).
- Egyptian non-expert participants (n=8-10/case).
- Non-Egyptian non-expert participants (n=32-40/case).

Significant correct identification rates above chance (>25%) are marked in red colour.

Cases (n=20)	Correct Identification Percentage-Combined (n=65-76)	Correct Identification Percentage			
		All Experts (n=25-26)	All Non-Experts (n=40-50)	Egy-Non-Experts (n=8-10)	Non-Egy Non-Experts (n=32-40)
F03	44.7%	44.4%	46%	60%	42.5%
F04	19.7%	11.8%	22%	11.1%	25%
F05	55.4%	52.9%	52.5%	37.5%	56.3%
F09	25.8%	11.8%	31.7%	11.1	37.5%
F13	21.1%	16.7%	18%	0%	22.5%
F14	42.6%	41.2%	48.8%	55.6%	47.1%
F17	20%	23.5%	20%	0%	25%
F23	17.1%	16.7%	18%	10%	20%
F25	29.4%	35.3%	25.6%	22.2%	26.5%
F26	68.2%	52.9%	73.2%	77.8%	71.9%
M02	47.1%	35.3%	51.2%	33.3%	55.9%
M07	81.6%	77.8%	80%	90%	77.5%
M11	17.6%	17.6%	18.6%	22.2%	17.6%
M14	49.2%	35.3%	50%	37.5%	53.1%
M19	9.1%	11.8%	7.3%	0%	9.4%
M22	23.5%	17.6%	23.3%	22.2%	23.5%
M23	61.8%	77.8%	54%	60%	52.5%
M26	53%	70.6%	43.9%	66.7%	37.5%
M31	30.8%	23.5%	32.5%	37.5%	31.3%
M34	47.7%	52.9%	40%	75%	31.3%
AVERAGE	**38.27%**	**36.37%**	**37.83%**	**36.49%**	**38.2%**
No. of significantly identified cases	**13/20**	**11/20**	**13/20**	**11/20**	**13/20**

Table 31: The correct identification rate of each of the 20 cases given by:

- All non-expert participants. This group includes Egyptian and non-Egyptians participants (n=40-50/case).
- Participants with a professional experience in forensic anthropology (with or without an experience in forensic facial reconstruction) (n=25-26/case).
- Participants with a professional experience in forensic face identification psychology (n=3/case).
- Participants with a professional experience forensic pathology. This group includes Egyptian and non-Egyptians participants (n=13/case).

Significant correct identification rates above chance (>25%) are marked in red colour.

Cases (n=20)	Correct Identification Percentage			
	All Non-Experts (n=40-50)	Anthropology Experts (+/- FR) (n=11-12)	Facial Identification Psychology Experts (n=3)	Forensic Pathology Experts (n=13)
F03	46%	41.7%	33.3%	46.2%
F04	22%	27.3%	0%	15.4%
F05	52.5%	72.7%	66.7%	46.2%
F09	31.7%	18.2%	0%	15.4%
F13	18%	58.3%	0%	7.7%
F14	48.8%	27.3%	33.3%	30.8%
F17	20%	27.3%	33.3%	23.1%
F23	18%	8.3%	33.3%	15.4%
F25	25.6%	36.4%	33.3%	46.2%
F26	73.2%	63.6%	66.7%	46.2%
M02	51.2%	45.5%	66.7%	30.8%
M07	80%	100%	66.7%	76.9%
M11	18.6%	18.2%	0%	23.1%
M14	50%	63.6%	66.7%	38.5%
M19	7.3%	18.2%	33.3%	15.4%
M22	23.3%	36.4%	33.3%	15.4%
M23	54%	75%	66.7%	84.6%
M26	43.9%	72.7%	33.3%	76.9%
M31	32.5%	27.3%	33.3%	30.8%
M34	40%	72.7%	100%	46.2%
AVERAGE	**37.83%**	**45.54%**	**40%**	**36.56%**
No. of significantly identified cases	**13/20**	**16/20**	**16/20**	**12/20**

Table 32: The 20 ranked cases according to the correct identification rate of each case given by:

- All participants (combined).
- All participants with a professional experience in (forensic anthropology, with or without an experience in forensic facial reconstruction, forensic face identification psychology, and forensic pathology). This group includes Egyptian and non-Egyptians participants.
- All non-expert participants. This group includes Egyptian and non-Egyptians participants.
- Egyptian non-expert participants.
- Non-Egyptian non-expert participants.

Cases (n=20)	Ranked Cases According to Correct ID%-Combined	Ranked Cases According to Correct ID%			
		All Experts	All Non-Experts	Egy-Non-Experts	Non-Egy Non-Experts
F03	9	7	8	5	8
F04	17	18	15	15	14
F05	4	4	4	8	3
F09	13	18	12	15	9
F13	15	16	18	18	17
F14	10	8	7	7	7
F17	16	12	16	18	14
F23	19	16	18	17	18
F25	12	9	13	12	13
F26	2	4	2	2	2
M02	8	9	5	11	4
M07	1	1	1	1	1
M11	18	14	17	12	19
M14	6	9	6	8	5
M19	20	18	20	18	20
M22	14	14	14	12	16
M23	3	1	3	5	6
M26	5	3	9	4	9
M31	11	12	11	8	11
M34	7	4	10	3	11

Table 33: The 20 ranked cases according to the correct identification rate of each case given by:

- All non-expert participants. This group includes Egyptian and non-Egyptians participants (n=40-50/case).
- Participants with professional experience in forensic anthropology (with or without experience in forensic facial reconstruction) (n=25-26/case).
- Participants with professional experience in forensic face identification psychology (n=3/case).
- Participants with professional experience forensic pathology. This group includes Egyptian and non-Egyptians participants (n=13/case).

Cases (n=20)	Ranked Cases According to Correct ID%			
	All Non-Experts	Anthropology Experts	Facial Identification Psychology Experts	Forensic Pathology Experts
F03	8	10	8	4
F04	15	13	17	15
F05	4	3	2	4
F09	12	17	17	15
F13	18	8	17	20
F14	7	13	8	10
F17	16	13	8	13
F23	18	20	8	15
F25	13	11	8	4
F26	2	6	2	4
M02	5	9	2	10
M07	1	1	2	2
M11	17	17	17	13
M14	6	6	2	9
M19	20	17	8	15
M22	14	11	8	15
M23	3	2	2	1
M26	9	3	8	2
M31	11	13	8	10
M34	10	3	1	4

Table 34: The number of observers participated in face resemblance test version (1)

Subjective Face Resemblance Test: Version 1		**Non-Experts**			**Anthropology Experts**	**TOTAL**
		Egyptian	**Non- Egyptian**			
		Continuing after Face Pool Test	**Continuing after Face Pool Test**	**Starting from Face Resemblance Test**	**Continuing after Face Pool Test**	
Exercise 1	**Group 1**	6	17	9	5	37
Exercise 2	**Group 2**	7	16	5	5	33

Table 35: The number of observers participated in face resemblance test version (2)

Subjective Face Resemblance Test: Version 2		**Non-Experts**			**Anthropology Experts**	**TOTAL**
		Egyptian	**Non- Egyptian**			
		Continuing after Face Pool Test	**Continuing after Face Pool Test**	**Starting from Face Resemblance Test**	**Continuing after Face Pool Test**	
Exercise 1	**Group 2**	8	17	6	5	36
Exercise 2	**Group 1**	6	15	8	5	34

Table 36: The average overall resemblance scores of each of the 30 cases given by:

- All participants (combined).
- Participants with a professional experience in forensic anthropology.
- Non-expert non-Egyptian participants who performed the face pool tests before the face resemblance tests (NEX-NEG-OLD).
- Non-expert non-Egyptian participants who did not perform the face pool tests before the face resemblance tests (NEX-NEG-NEW).
- Non-expert Egyptian participants.

Cases (n=30)	COMBINED	EX-ANTH	NEX-NEG-OLD	NEX-NEG-NEW	NEX-EGY
F03	5.0	5.0	4.5	2.9	7.8
F04	4.0	5.2	3.0	2.2	5.4
F05	4.7	5.8	4.3	3.4	5.1
F09	2.6	1.8	1.4	1.0	6.0
F13	3.2	2.8	1.8	2.0	6.2
F14	4.7	6.6	4.1	3.0	5.0
F17	3.6	5.4	3.0	2.4	3.7
F23	4.7	6.0	3.8	2.8	6.0
F25	5.4	6.6	4.7	4.1	6.2
F26	7.3	8.2	7.0	6.2	7.9
M02	5.5	7.0	4.8	3.4	6.9
M07	5.8	6.2	5.6	4.2	7.3
M11	3.3	2.8	2.0	1.8	6.7
M14	6.5	7.8	6.1	4.4	7.7
M19	5.0	5.4	4.5	4.6	5.3
M22	5.5	6.2	4.8	3.8	7.0
M23	5.2	5.6	4.3	3.2	7.8
M26	4.9	4.8	4.1	3.3	7.5
M31	4.0	4.4	3.6	3.2	4.7
M34	5.1	5.8	4.7	4.2	5.6
F06	3.8	2.8	2.8	3.2	6.3
F20	2.7	1.8	2.2	1.1	5.8
F21	3.4	3.8	2.9	2.6	4.4
M05	3.3	2.2	2.4	2.8	5.8
M09	3.7	3.4	3.2	1.9	6.2
M12	4.3	5.8	3.4	2.4	5.4
M15	4.5	4.2	3.2	3.2	7.5
M17	4.9	5.8	4.1	2.6	7.2
M21	5.3	5.2	3.7	4.8	7.3
M28	3.7	3.8	3.6	2.0	5.2
Mean Scores	**4.5**	**4.9**	**3.8**	**3.1**	**6.2**

Table 37: The 30 ranked cases according to the average overall resemblance scores given by:

- All participants (combined).
- Participants with a professional experience in forensic anthropology.
- Non-expert non-Egyptian participants who performed the face pool tests before the face resemblance tests (NEX-NEG-OLD).
- Non-expert non-Egyptian participants who did not perform the face pool tests before the face resemblance tests (NEX-NEG-NEW).
- Non-expert Egyptian participants.

CASES (n=30)	COMBINED	EX-ANTH	NEX-NEG-OLD	NEX-NEG-NEW	NEX-EGY
F03	10	18	9	17	2
F04	20	16	22	24	22
F05	15	9	10	9	26
F09	30	29	30	30	17
F13	28	25	29	25	14
F14	14	4	12	16	27
F17	24	14	23	22	30
F23	16	8	15	18	17
F25	6	4	7	7	14
F26	1	1	1	1	1
M02	4	3	4	9	11
M07	3	6	3	5	8
M11	26	25	28	28	12
M14	2	2	2	4	4
M19	11	14	8	3	24
M22	5	6	4	8	10
M23	8	13	11	12	3
M26	13	19	14	11	5
M31	19	20	17	15	28
M34	9	9	6	5	21
F06	21	25	25	12	13
F20	29	29	27	29	19
F21	25	22	24	20	29
M05	27	28	26	19	19
M09	22	24	21	27	14
M12	18	9	19	22	22
M15	17	21	20	12	5
M17	12	9	13	21	9
M21	7	16	16	2	7
M28	22	22	18	25	25

Table 38: The average resemblance scores and case ranks of each of the 30 cases given by the NEX-NEG-OLD, and the NEX-NEG-NEW groups

CASES (n=30)	NEX-NEG-OLD-AVERAGE	NEX-NEG-NEW-AVERAGE	NEX-NEG-OLD-RANK	NEX-NEG-NEW-RANK
F03	4.5	2.9	9	17
F04	3.1	2.2	22	24
F05	4.3	3.4	10	9
F09	1.4	1.0	30	30
F13	1.8	2.0	29	25
F14	4.1	3.0	12	16
F17	3.0	2.4	23	22
F23	3.8	2.8	15	18
F25	4.7	4.1	7	7
F26	7.0	6.2	1	1
M02	4.8	3.4	4	9
M07	5.6	4.2	3	5
M11	2.0	1.8	28	28
M14	6.1	4.4	2	4
M19	4.5	4.6	8	3
M22	4.8	3.8	4	8
M23	4.3	3.2	11	12
M26	4.1	3.3	14	11
M31	3.6	3.2	17	15
M34	4.7	4.2	6	5
F06	2.8	3.2	25	12
F20	2.2	1.1	27	29
F21	2.9	2.6	24	20
M05	2.4	2.8	26	19
M09	3.2	1.9	21	27
M12	3.4	2.4	19	22
M15	3.2	3.2	20	12
M17	4.1	2.6	13	21
M21	3.7	4.8	16	2
M28	3.6	2.0	18	25

Table 39: The average resemblance scores and case ranks of each of the 30 cases given by the NEX-NEG-OLD, and the non-expert Egyptian groups

CASES (n=30)	NEX-NEG-OLD-AVERAGE	NEX-EGY-AVERAGE	NEX-NEG-OLD-RANK	NEX-EGY-RANK
F03	4.5	7.8	9	2
F04	3.1	5.4	22	22
F05	4.3	5.1	10	26
F09	1.4	6.0	30	17
F13	1.8	6.2	29	14
F14	4.1	5.0	12	27
F17	3.0	3.7	23	30
F23	3.8	6.0	15	17
F25	4.7	6.2	7	14
F26	7.0	7.9	1	1
M02	4.8	6.9	4	11
M07	5.6	7.3	3	8
M11	2.0	6.7	28	12
M14	6.1	7.7	2	4
M19	4.5	5.3	8	24
M22	4.8	7.0	4	10
M23	4.3	7.8	11	3
M26	4.1	7.5	14	5
M31	3.6	4.7	17	28
M34	4.7	5.6	6	21
F06	2.8	6.3	25	13
F20	2.2	5.8	27	19
F21	2.9	4.4	24	29
M05	2.4	5.8	26	19
M09	3.2	6.2	21	14
M12	3.4	5.4	19	22
M15	3.2	7.5	20	5
M17	4.1	7.3	13	9
M21	3.7	7.3	16	7
M28	3.6	5.3	18	25

Table 40: The average resemblance scores and case ranks of each of the 30 cases given by the NEX-NEG-OLD, and the forensic anthropology experts groups

CASES (n=30)	NEX-NEG-OLD-AVERAGE	ANTH-AVERAGE	NEX-NEG-OLD-RANK	ANTH-RANK
F03	2.0	2.8	28	25
F04	1.4	1.8	30	29
F05	3.4	5.8	19	9
F09	4.5	5.0	9	18
F13	3.6	3.8	18	22
F14	2.2	1.8	27	29
F17	3.2	3.4	21	24
F23	4.7	6.6	7	4
F25	3.2	4.2	20	21
F26	4.3	5.6	11	13
M02	2.4	2.2	26	28
M07	2.8	2.8	25	25
M11	3.7	5.2	16	16
M14	1.8	2.8	29	25
M19	4.1	4.8	14	19
M22	3.0	5.4	23	14
M23	5.6	6.2	3	6
M26	7.0	8.2	1	1
M31	6.1	7.8	2	2
M34	4.7	5.8	6	9
F06	4.1	5.8	13	9
F20	4.3	5.8	10	9
F21	4.5	5.4	8	14
M05	2.9	3.8	24	22
M09	4.8	7.0	4	3
M12	3.8	6.0	15	8
M15	3.6	4.4	17	20
M17	4.1	6.6	12	4
M21	4.8	6.2	4	6
M28	3.1	5.2	22	16

Table 41: The average resemblance scores of each of the 30 cases given by all participants (combined) to the individual facial regions

CASE	FOREHEAD	ORBIT	NASAL BONE	CHEEK BONE	CHIN BONE	JAW	Overall Face Regions Scores
F03	5.3	5.3	6.3	5.2	4.3	4.4	5.1
F04	5.0	4.4	2.5	4.3	4.1	3.9	4.0
F05	6.9	6.5	5.8	4.9	4.4	4.4	5.5
F09	5.2	3.6	4.8	3.9	3.6	3.4	4.1
F13	3.8	4.2	3.6	4.9	5.3	4.8	4.4
F14	4.1	3.4	3.0	4.2	4.6	4.2	3.9
F17	4.0	3.1	2.9	3.7	4.4	3.5	3.6
F23	5.5	4.5	4.5	4.7	4.1	3.7	4.5
F25	4.2	3.5	3.3	5.6	5.9	5.7	4.7
F26	5.9	6.3	5.9	6.2	6.7	5.9	6.1
M02	4.8	4.4	3.9	3.6	6.0	5.6	4.7
M07	5.2	5.2	5.7	6.0	3.5	4.7	5.0
M11	5.0	3.9	2.8	3.5	3.4	3.9	3.8
M14	5.8	6.1	5.5	6.1	6.6	6.7	6.1
M19	5.0	3.3	3.2	4.4	4.4	4.3	4.1
M22	5.5	5.3	4.3	5.2	6.0	5.6	5.3
M23	4.9	5.0	5.1	5.0	6.1	5.6	5.3
M26	2.9	3.7	2.9	4.1	4.7	3.8	3.7
M31	3.8	2.9	3.1	3.9	3.2	3.9	3.5
M34	4.3	4.2	5.6	6.4	5.2	5.8	5.2
F06	5.1	4.3	3.6	4.7	3.2	3.2	4.0
F20	5.8	4.5	3.1	3.7	2.9	3.3	3.9
F21	3.0	1.9	2.9	3.5	4.8	4.4	3.4
M05	4.6	3.1	3.2	3.0	3.8	3.5	3.5
M09	5.9	5.1	4.5	4.5	3.0	2.9	4.3
M12	3.0	4.7	3.2	4.3	3.8	3.7	3.8
M15	4.6	4.4	3.7	4.7	5.3	4.8	4.6
M17	5.9	4.9	6.5	5.8	4.5	5.2	5.4
M21	4.4	3.6	4.1	5.0	5.7	6.1	4.8
M28	4.1	3.6	3.1	3.6	4.9	4.7	4.0

Table 42: The ranks of the 30 assessed cases according to the average resemblance scores given by all participants (combined) to individual facial regions

CASE	FOREHEAD	ORBIT	NASAL BONE	CHEEK BONE	CHIN BONE	JAW	Overall Face Regions Ranks
F03	9	5	2	7	19	14	8
F04	15	14	30	18	21	20	19
F05	1	1	4	11	16	15	3
F09	11	21	9	22	24	27	18
F13	26	18	16	12	9	11	15
F14	24	25	25	20	14	18	22
F17	25	28	26	25	16	26	27
F23	7	12	10	14	20	24	14
F25	22	24	18	6	6	5	12
F26	4	2	3	2	1	3	1
M02	17	15	14	27	4	8	11
M07	10	6	5	4	25	12	9
M11	13	19	29	28	26	19	25
M14	6	3	7	3	2	1	2
M19	13	26	19	17	16	17	17
M22	8	4	12	8	5	6	5
M23	16	8	8	9	3	7	6
M26	30	20	28	21	13	22	26
M31	27	29	23	23	28	21	29
M34	21	17	6	1	10	4	7
F06	12	16	17	15	27	29	21
F20	5	11	22	24	30	28	23
F21	29	30	26	29	12	15	30
M05	19	27	19	30	22	25	28
M09	2	7	11	16	29	30	16
M12	28	10	21	18	23	23	24
M15	18	13	15	13	8	10	13
M17	3	9	1	5	15	9	4
M21	20	23	13	10	7	2	10
M28	23	21	23	26	11	13	20

Table 43: The facial surface overall distance standard deviation (SD) and the cases' ranks of each of the 30 cases

Cases (n=30)	SD Obj.	SD Obj. Rank
F03	2.19	27
F04	3.33	11
F05	2.20	26
F09	3.73	9
F13	5.44	5
F14	2.88	16
F17	3.92	8
F23	3.27	12
F25	1.95	30
F26	2.35	25
M02	2.83	19
M07	1.98	29
M11	3.21	13
M14	2.64	23
M19	2.06	28
M22	3.39	10
M23	2.85	17
M26	2.66	22
M31	2.74	21
M34	3.07	14
F06	4.21	6
F20	5.70	3
F21	2.98	15
M05	4.02	7
M09	6.24	2
M12	6.34	1
M15	5.45	4
M17	2.80	20
M21	2.38	24
M28	2.84	18

Table 44: The ranks of the 30 cases according to the sum of the absolute objective surface differences at all facial regions

Case	Sum of Regions Absolute Differences	Sum of Regions Absolute Differences Rank
F03	69.62	20
F04	84.42	15
F05	58.69	26
F09	103.76	10
F13	159.85	2
F14	106.39	9
F17	107.17	7
F23	102.27	11
F25	57.43	28
F26	69.24	21
M02	85.01	14
M07	61.46	25
M11	96.01	12
M14	40.51	30
M19	62.55	24
M22	106.63	8
M23	75.81	18
M26	57.75	27
M31	54.03	29
M34	74.68	19
F06	128.75	5
F20	171.54	1
F21	64.72	23
M05	89.71	13
M09	155.40	3
M12	121.19	6
M15	153.23	4
M17	82.20	16
M21	66.21	22
M28	80.52	17

Table 45: The absolute difference of the surface distance at each facial regions of each of the 30 cases

Case	Forehead	Orbit	Nasal Region	Cheek	Chin	Jaw
F03	17.34	8.84	1.58	9.64	12.74	24.00
F04	17.94	12.84	4.06	14.40	17.84	30.28
F05	15.13	8.55	1.41	9.53	8.71	19.03
F09	24.80	17.74	2.82	21.34	10.17	35.47
F13	23.02	14.99	7.42	22.71	2.23	98.90
F14	17.08	8.84	1.64	14.14	3.52	61.99
F17	27.29	17.54	2.72	24.67	2.04	51.61
F23	17.58	12.32	3.16	16.64	5.60	55.03
F25	15.35	13.87	2.67	15.44	3.90	19.61
F26	21.52	18.33	7.57	9.79	4.61	25.48
M02	30.21	14.57	4.29	14.36	5.80	29.65
M07	18.69	6.92	3.49	5.57	12.82	13.47
M11	23.05	14.56	4.87	7.38	5.97	46.44
M14	17.28	9.17	3.38	6.31	2.19	7.31
M19	19.84	10.11	2.13	11.03	2.92	27.32
M22	49.03	32.72	9.07	21.24	6.34	33.02
M23	15.41	8.30	4.46	6.71	1.47	31.61
M26	14.50	12.64	1.08	11.36	2.48	27.53
M31	13.96	15.99	1.39	16.67	3.16	21.65
M34	21.59	13.32	5.44	9.34	6.46	39.24
F06	19.74	10.34	4.61	20.06	4.44	77.74
F20	35.15	24.18	8.10	27.33	2.44	101.00
F21	15.81	15.89	2.24	18.78	3.28	23.75
M05	26.56	21.70	10.66	15.51	12.35	27.92
M09	30.32	19.26	4.93	25.97	7.85	88.33
M12	37.79	15.42	5.41	10.61	7.78	63.68
M15	38.43	25.71	4.00	31.86	7.13	72.74
M17	22.92	13.81	4.00	14.02	6.04	30.82
M21	23.30	20.42	4.34	18.21	7.65	16.75
M28	26.47	16.56	1.34	16.13	5.38	27.34

Table 46: The ranks of the 30 cases according to the absolute difference of the surface distance at each facial regions

Case	Forehead	Orbit	Nasal Region	Cheek	Chin	Jaw
F03	22	26	26	24	3	23
F04	20	20	14	16	1	16
F05	28	28	27	25	6	27
F09	10	8	20	6	5	12
F13	13	14	5	5	27	2
F14	24	27	25	18	21	7
F17	7	9	21	4	29	9
F23	21	22	19	12	16	8
F25	27	17	22	15	20	26
F26	16	7	4	23	18	22
M02	6	15	13	17	15	17
M07	19	30	17	30	2	29
M11	12	16	9	27	14	10
M14	23	25	18	29	28	30
M19	17	24	24	21	24	21
M22	1	1	2	7	12	13
M23	26	29	11	28	30	14
M26	29	21	30	20	25	19
M31	30	11	28	11	23	25
M34	15	19	6	26	11	11
F06	18	23	10	8	19	4
F20	4	3	3	2	26	1
F21	25	12	23	9	22	24
M05	8	4	1	14	4	18
M09	5	6	8	3	7	3
M12	3	13	7	22	8	6
M15	2	2	15	1	10	5
M17	14	18	16	19	13	15
M21	11	5	12	10	9	28
M28	9	10	29	13	17	20

Table 47: The absolute differences between the linear ratios of the real and the reconstructed faces and averaged differences of each of the 30 cases

Cases (n=30)	AB/ AD	AB/ BD	BC/ CD	BC/ BD	AD/ BD	CD/ BD	AB/ AE	AB/ BE	BC/ CE	BC/ BE	AE/ BE	CE/ BE	AC/ BE	Average
F03	0.059	0.112	0.013	0.063	0.050	0.067	0.057	0.080	0.037	0.048	0.034	0.015	0.063	**0.054**
F04	0.053	0.115	0.010	0.042	0.071	0.083	0.044	0.066	0.020	0.020	0.037	0.005	0.036	**0.046**
F05	0.023	0.092	0.042	0.001	0.085	0.075	0.035	0.038	0.011	0.020	0.000	0.016	0.007	**0.034**
F09	0.046	0.046	0.058	0.101	0.021	0.036	0.024	0.017	0.042	0.051	0.018	0.009	0.036	**0.039**
F13	0.068	0.173	0.095	0.258	0.103	0.171	0.033	0.040	0.057	0.094	0.004	0.051	0.095	**0.096**
F14	0.021	0.012	0.038	0.034	0.021	0.019	0.014	0.009	0.033	0.023	0.014	0.024	0.005	**0.021**
F17	0.090	0.166	0.053	0.130	0.058	0.078	0.075	0.111	0.059	0.088	0.048	0.036	0.122	**0.086**
F23	0.047	0.072	0.044	0.048	0.018	0.010	0.034	0.028	0.033	0.012	0.014	0.043	0.019	**0.032**
F25	0.004	0.044	0.030	0.034	0.073	0.114	0.002	0.031	0.002	0.025	0.058	0.057	0.032	**0.039**
F26	0.002	0.003	0.029	0.059	0.007	0.031	0.020	0.006	0.024	0.029	0.028	0.004	0.008	**0.019**
M02	0.037	0.015	0.008	0.002	0.046	0.017	0.036	0.031	0.027	0.018	0.014	0.021	0.037	**0.024**
M07	0.005	0.015	0.024	0.011	0.013	0.060	0.004	0.007	0.007	0.010	0.004	0.005	0.002	**0.013**
M11	0.027	0.044	0.008	0.027	0.009	0.053	0.043	0.032	0.003	0.014	0.027	0.018	0.029	**0.026**
M14	0.039	0.103	0.005	0.036	0.084	0.065	0.023	0.044	0.020	0.001	0.041	0.045	0.053	**0.043**
M19	0.043	0.046	0.001	0.042	0.010	0.061	0.006	0.016	0.002	0.013	0.017	0.026	0.039	**0.025**
M22	0.000	0.044	0.030	0.039	0.061	0.002	0.025	0.049	0.048	0.047	0.043	0.008	0.070	**0.036**
M23	0.033	0.113	0.004	0.048	0.106	0.065	0.034	0.031	0.020	0.011	0.009	0.020	0.011	**0.039**
M26	0.035	0.011	0.035	0.028	0.046	0.022	0.013	0.005	0.028	0.017	0.017	0.025	0.026	**0.024**
M31	0.031	0.039	0.004	0.021	0.002	0.037	0.026	0.051	0.006	0.039	0.042	0.062	0.082	**0.034**
M34	0.011	0.075	0.024	0.050	0.127	0.030	0.001	0.014	0.001	0.003	0.025	0.010	0.009	**0.029**
F06	0.018	0.087	0.012	0.068	0.088	0.073	0.026	0.024	0.003	0.011	0.010	0.014	0.017	**0.035**
F20	0.039	0.007	0.037	0.002	0.059	0.070	0.002	0.007	0.039	0.012	0.010	0.063	0.001	**0.027**
F21	0.059	0.039	0.065	0.069	0.051	0.021	0.051	0.038	0.068	0.058	0.030	0.028	0.015	**0.046**
M05	0.037	0.042	0.027	0.087	0.005	0.080	0.030	0.025	0.032	0.054	0.012	0.040	0.043	**0.04**
M09	0.060	0.096	0.032	0.077	0.019	0.049	0.066	0.057	0.049	0.045	0.032	0.020	0.054	**0.05**
M12	0.002	0.033	0.010	0.005	0.043	0.009	0.010	0.017	0.008	0.007	0.014	0.001	0.018	**0.014**
M15	0.015	0.074	0.057	0.126	0.081	0.077	0.028	0.046	0.053	0.079	0.031	0.040	0.096	**0.062**
M17	0.015	0.010	0.024	0.031	0.037	0.002	0.037	0.010	0.049	0.036	0.048	0.031	0.014	**0.026**
M21	0.020	0.003	0.033	0.010	0.032	0.047	0.018	0.017	0.017	0.011	0.068	0.060	0.052	**0.03**
M28	0.016	0.069	0.037	0.001	0.069	0.067	0.006	0.029	0.024	0.014	0.042	0.022	0.050	**0.034**

Table 48: The ranked 30 cases according to the individual linear ratios and their average

Cases (n=30)	AB/ AD	AB/ BD	BC/ CD	BC/ BD	AD/ BD	CD/ BD	AB/ AE	AB/ BE	BC/ CE	BC/ BE	AE/ BE	CE/ BE	AC/ BE	Average
F03	4	5	21	9	15	10	3	2	10	7	10	22	6	**4**
F04	6	3	23	14	9	3	5	3	18	15	9	27	14	**6**
F05	18	8	7	29	5	7	9	10	22	15	30	21	27	**15**
F09	8	14	3	4	21	20	18	20	8	6	17	25	14	**10**
F13	2	1	1	1	3	1	12	9	3	1	28	5	3	**1**
F14	19	25	8	18	21	25	22	26	11	14	20	15	28	**27**
F17	1	2	5	2	13	5	1	1	2	2	3	10	1	**2**
F23	7	12	6	12	24	27	10	17	11	22	20	7	19	**18**
F25	27	16	14	18	8	2	28	13	28	13	2	4	16	**10**
F26	28	29	16	10	28	21	20	29	16	12	14	29	26	**28**
M02	12	23	25	27	16	26	8	13	15	17	20	17	13	**25**
M07	26	23	18	24	25	15	27	27	24	27	28	27	29	**30**
M11	17	16	25	22	27	16	6	12	26	19	15	20	17	**22**
M14	10	6	27	17	6	12	19	8	18	30	8	6	8	**8**
M19	9	14	30	14	26	14	25	23	28	21	18	13	12	**24**
M22	30	16	14	16	11	29	17	6	7	8	5	26	5	**13**
M23	15	4	28	12	2	12	10	13	18	24	27	18	24	**10**
M26	14	26	11	21	16	23	23	30	14	18	18	14	18	**25**
M31	16	20	28	23	30	19	15	5	25	10	6	2	4	**15**
M34	25	10	18	11	1	22	30	24	30	29	16	24	25	**20**
F06	21	9	22	8	4	8	15	19	26	24	25	23	21	**14**
F20	10	28	9	27	12	9	28	27	9	22	25	1	30	**21**
F21	4	20	2	7	14	24	4	10	1	4	13	12	22	**6**
M05	12	19	17	5	29	4	13	18	13	5	24	8	11	**9**
M09	3	7	13	6	23	17	2	4	5	9	11	18	7	**5**
M12	28	22	23	26	18	28	24	20	23	28	20	30	20	**29**
M15	23	11	4	3	7	6	14	7	4	3	12	8	2	**3**
M17	23	27	18	20	19	29	7	25	5	11	3	11	23	**22**
M21	20	29	12	25	20	18	21	20	21	24	1	3	9	**19**
M28	22	13	9	29	10	10	25	16	16	19	6	16	10	**15**

Table 49: The absolute differences between the angles of the real and the reconstructed faces and averaged differences of each of the 30 cases

Cases (n=30)	AEC	CAE	CAD	ACE	ACD	ABE	CBE	CDB	ADB	CEB	AEB	Average
F03	2.464	2.238	0.325	0.226	0.669	0.461	0.385	1.228	4.677	2.421	3.938	**1.73**
F04	1.255	2.817	0.836	1.562	0.038	1.261	1.387	0.356	4.136	1.226	2.964	**1.622**
F05	0.872	1.209	0.637	0.337	1.203	1.657	0.799	2.823	2.153	0.810	2.225	**1.339**
F09	2.364	0.195	0.738	2.558	3.749	2.692	1.345	4.465	3.412	2.705	1.413	**2.331**
F13	3.930	0.050	2.779	3.979	6.168	1.512	0.578	7.958	5.713	3.986	2.169	**3.529**
F14	0.682	0.866	0.317	0.185	0.456	1.971	3.659	2.606	1.409	1.869	0.814	**1.349**
F17	4.690	3.120	1.927	1.570	2.840	0.393	0.918	4.301	7.239	4.003	5.092	**3.281**
F23	2.848	2.895	3.084	0.047	1.365	2.626	5.105	3.154	3.517	1.664	1.906	**2.565**
F25	1.050	0.375	3.234	0.675	0.541	5.549	5.805	1.764	0.159	0.157	0.670	**1.816**
F26	0.353	1.031	0.283	1.384	0.936	3.302	1.074	2.209	0.085	1.577	1.030	**1.206**
M02	3.197	1.946	0.349	1.252	2.012	2.684	3.033	0.492	2.610	1.535	2.107	**1.929**
M07	0.370	0.214	1.180	0.157	1.441	0.128	0.024	1.594	0.411	0.469	0.301	**0.572**
M11	3.096	1.088	3.254	2.007	0.075	4.094	1.294	0.421	2.162	0.358	2.567	**1.856**
M14	0.965	0.228	2.372	0.737	0.958	3.134	5.357	0.089	3.387	0.692	1.891	**1.801**
M19	1.154	0.147	0.063	1.006	3.772	1.169	2.169	0.146	3.038	0.143	0.553	**1.215**
M22	3.418	4.593	2.236	1.175	1.441	3.074	2.843	2.186	0.538	2.903	2.100	**2.41**
M23	0.331	1.376	2.153	1.707	0.347	2.161	2.836	0.428	2.723	1.086	1.964	**1.556**
M26	0.467	0.197	0.637	0.664	0.910	2.222	3.512	2.368	2.166	1.445	0.660	**1.386**
M31	1.987	0.111	0.093	1.876	2.088	2.250	4.797	0.188	2.294	0.690	1.923	**1.663**
M34	0.136	2.009	3.690	1.873	2.799	2.334	0.930	1.913	0.305	0.000	0.292	**1.48**
F06	0.831	0.685	0.862	1.516	0.072	2.208	0.839	1.020	1.578	0.255	1.634	**1.045**
F20	1.605	4.776	3.286	3.170	4.406	0.902	8.138	2.670	2.373	2.232	0.189	**3.068**
F21	2.388	1.061	0.210	1.327	1.730	5.005	5.471	4.725	3.973	3.952	2.826	**2.97**
M05	1.826	1.967	1.879	3.793	3.387	2.466	2.301	2.301	2.578	2.253	1.715	**2.406**
M09	4.454	1.611	0.561	2.843	2.231	6.248	4.481	2.550	4.749	3.115	4.099	**3.358**
M12	0.712	1.042	1.247	0.331	0.813	0.926	0.370	0.587	0.533	0.443	0.357	**0.669**
M15	3.778	1.695	2.085	2.083	1.324	1.687	0.692	4.400	1.445	3.564	2.003	**2.251**
M17	3.171	0.757	1.148	2.413	1.131	5.360	4.822	1.743	0.732	2.765	1.539	**2.326**
M21	0.381	0.157	0.887	0.538	0.020	6.439	6.520	2.228	1.420	0.750	0.708	**1.823**
M28	1.246	1.702	1.039	0.456	0.993	3.428	3.056	2.299	1.551	1.299	0.757	**1.621**

Table 50: The ranked 30 cases according to the individual angles and their averages

Cases (n=30)	AEC	CAE	CAD	ACE	ACD	ABE	CBE	CDB	ADB	CEB	AEB	Average
F03	10	6	25	27	23	28	28	21	4	9	3	**17**
F04	16	5	19	13	29	24	18	27	5	18	4	**19**
F05	21	14	21	25	16	22	25	7	18	20	7	**25**
F09	12	26	20	5	4	11	19	3	8	8	19	**9**
F13	3	30	6	1	1	23	27	1	2	2	8	**1**
F14	24	19	26	28	25	20	10	9	23	12	21	**24**
F17	1	3	11	12	6	29	23	5	1	1	1	**3**
F23	9	4	5	30	14	13	6	6	7	13	14	**6**
F25	19	22	4	21	24	3	3	18	29	28	24	**15**
F26	28	18	27	15	20	8	21	15	30	14	20	**27**
M02	6	9	24	17	10	12	13	24	12	15	9	**12**
M07	27	24	14	29	12	30	30	20	27	24	28	**30**
M11	8	15	3	8	27	6	20	26	17	26	6	**13**
M14	20	23	7	20	19	9	5	30	9	22	15	**16**
M19	18	28	30	19	3	25	17	29	10	29	26	**26**
M22	5	2	8	18	12	10	14	16	25	6	10	**7**
M23	29	13	9	11	26	19	15	25	11	19	12	**21**
M26	25	25	21	22	21	17	11	11	16	16	25	**23**
M31	13	29	29	9	9	16	8	28	15	23	13	**18**
M34	30	7	1	10	7	15	22	17	28	30	29	**22**
F06	22	21	18	14	28	18	24	22	19	27	17	**28**
F20	15	1	2	3	2	27	1	8	14	11	30	**4**
F21	11	16	28	16	11	5	4	2	6	3	5	**5**
M05	14	8	12	2	5	14	16	12	13	10	16	**8**
M09	2	12	23	4	8	2	9	10	3	5	2	**2**
M12	23	17	13	26	22	26	29	23	26	25	27	**29**
M15	4	11	10	7	15	21	26	4	21	4	11	**11**
M17	7	20	15	6	17	4	7	19	24	7	18	**10**
M21	26	27	17	23	30	1	2	14	22	21	23	**14**
M28	17	10	16	24	18	7	12	13	20	17	22	**20**

Table 51: The ranks of 20 assessed cases according to all subjective and objective tests

Case	Correct ID%	Overall resemblance scores	Overall face regions scores	Facial surface overall distance SD	Sum of absolute regions differences	Average differences of linear ratios	Average differences of angles
F03	9	9	7	17	9	3	10
F04	17	16	15	5	18	4	12
F05	4	13	3	16	7	10	17
F09	13	20	13	3	19	6	5
F13	15	19	12	1	17	1	1
F14	10	12	16	9	11	18	16
F17	16	17	19	2	20	2	2
F23	19	14	11	6	16	12	3
F25	12	6	10	20	3	6	8
F26	2	1	1	15	13	19	19
M02	8	4	9	11	1	16	6
M07	1	3	8	19	14	20	20
M11	18	18	17	7	5	14	7
M14	6	2	2	14	12	5	9
M19	20	10	13	18	4	15	18
M22	14	5	4	4	2	9	4
M23	3	7	5	10	15	6	13
M26	5	11	18	13	6	16	15
M31	11	15	20	12	8	10	11
M34	7	8	6	8	10	13	14

Table 52: The ranks of 30 assessed cases according to all objective tests and the subjective face resemblance tests

Case	Overall resemblance scores	Overall face regions scores	Facial surface overall distance standard deviation	Sum of absolute regions differences	Average differences of linear ratios	Average differences of angles
F03	10	8	27	20	4	17
F04	20	19	11	15	6	19
F05	15	3	26	26	15	25
F09	30	18	9	10	10	9
F13	28	15	5	2	1	1
F14	14	22	16	9	27	24
F17	24	27	8	7	2	3
F23	16	14	12	11	18	6
F25	6	12	30	28	10	15
F26	1	1	25	21	28	27
M02	4	11	19	14	25	12
M07	3	9	29	25	30	30
M11	26	25	13	12	22	13
M14	2	2	23	30	8	16
M19	11	17	28	24	24	26
M22	5	5	10	8	13	7
M23	8	6	17	18	10	21
M26	13	26	22	27	25	23
M31	19	29	21	29	15	18
M34	9	7	14	19	20	22
F06	21	21	6	5	14	28
F20	29	23	3	1	21	4
F21	25	30	15	23	6	5
M05	27	28	7	13	9	8
M09	22	16	2	3	5	2
M12	18	24	1	6	29	29
M15	17	13	4	4	3	11
M17	12	4	20	16	22	10
M21	7	10	24	22	19	14
M28	22	20	18	17	15	20

Table 53: The ranks of 30 assessed cases according to all subjective and objective tests (except the average differences of linear ratios), after adjusting for the negative ranks' correlations.

Cases	Overall resemblance scores	Overall face regions scores	Facial surface overall distance	Average differences of angles	Sum of absolute regions differences	Rank Sum
F26	1	1	3	4	10	**19**
M07	3	9	1	1	6	**20**
M14	2	2	23	15	1	**43**
F05	15	3	16	6	5	**45**
M19	11	17	7	5	7	**47**
M34	9	7	11	9	12	**48**
M21	7	10	12	17	9	**55**
M02	4	11	6	19	17	**57**
M26	13	26	6	8	4	**57**
F25	6	12	21	16	3	**58**
M23	8	6	21	10	13	**58**
M17	12	4	9	21	15	**61**
F14	14	22	4	7	22	**69**
F03	10	8	27	14	11	**70**
M12	18	24	2	2	25	**71**
M22	5	5	18	24	23	**75**
M31	19	29	16	13	2	**79**
M28	22	20	16	11	14	**83**
F23	16	14	13	25	20	**88**
F06	21	21	17	3	26	**88**
F04	20	19	25	12	16	**92**
M11	26	25	9	18	19	**97**
M15	17	13	28	20	27	**105**
F09	30	18	21	22	21	**112**
F21	25	30	25	26	8	**114**
M05	27	28	22	23	18	**118**
F20	29	23	10	27	30	**119**
M09	22	16	26	29	28	**121**
F13	28	15	30	30	29	**132**
F17	24	27	29	28	24	**132**

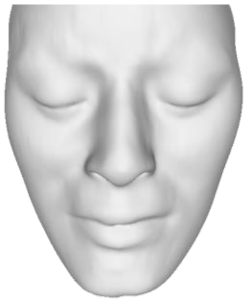
APPENDIX 19: INDIVIDUAL CASE STUDIES AND TEST RESULTS

1- Female Case- F03

Data

Facial Template
F-Av(16-20)Y-1

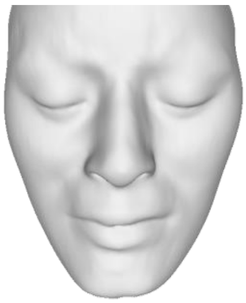
F03-Skull



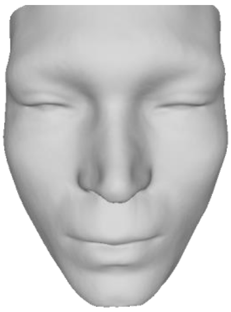
F03-Reconstructed Face

F03-Real Face

Face Pool Test



F03-Reconstructed Face



Test Face (A)
F08-Real Face

Test Face (B)
F02-Real Face

Test Face (C)
F03-Real Face

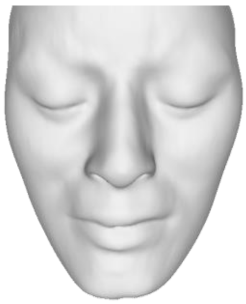
Test Face (D)
F04-Real Face

	Test Face (A)	Test Face (B)	Test Face (C)	Test Face (D)
ID %	**30.30**	**19.70**	**44.70***	**5.30**

*The correct face

Face Resemblance Test

F03-Real Face

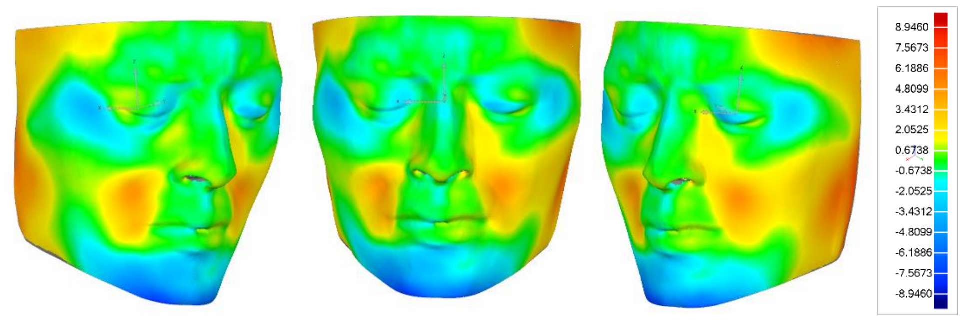


F03-Reconstructed Face

5.05

F03-Overall Resemblance Score

Objective Face Resemblance Test



F03-Objective Surface Difference Histogram (color Map)

2.188

F03-Surface Distance

2- Female Case- F04

Data

Facial Template
F-Av(16-20)Y-2

F04-Skull

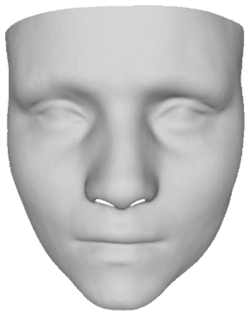
F04-Reconstructed Face

F04-Real Face

Face Pool Test

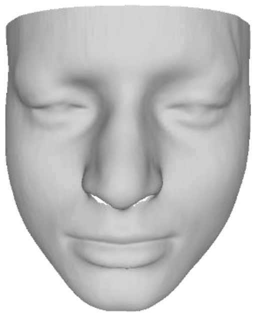
F04-Reconstructed Face

Test Face (A)
F04-Real Face



Test Face (B)
F09-Real Face

Test Face (C)
F03-Real Face



Test Face (D)
F05-Real Face

	Test Face (A)	Test Face (B)	Test Face (C)	Test Face (D)
ID %	**20.00***	**60.00**	**7.70**	**12.30**

*The correct face

Face Resemblance Test

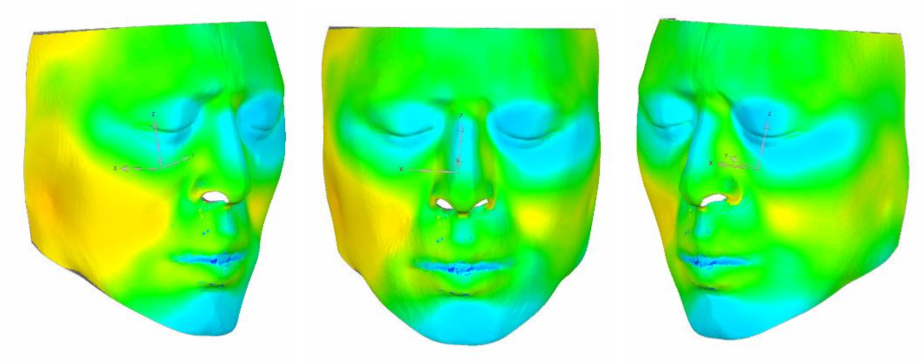
F04-Real Face

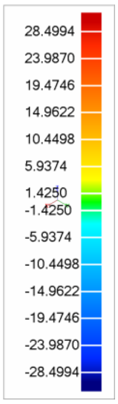
F04-Reconstructed Face

3.97

F04-Overall Resemblance Score

Objective Face Resemblance Test





F04-Objective Surface Difference Histogram (color Map)

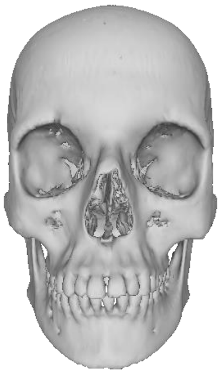
3.326

F04-Surface Distance

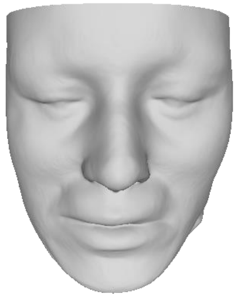
3- Female Case- F05

Data

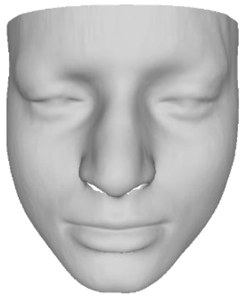
Facial Template
F-Av(21-30)Y



F05-Skull

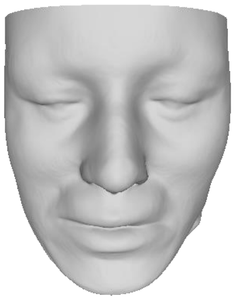


F05-Reconstructed Face

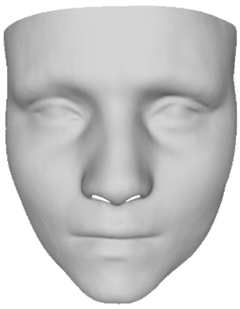


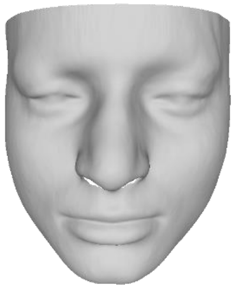
F05-Real Face

Face Pool Test



F05-Reconstructed Face

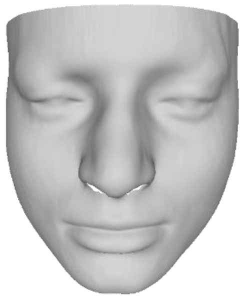


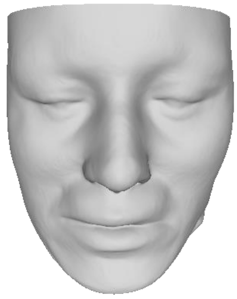


	Test Face (A) F11-Real Face	Test Face (B) F08-Real Face	Test Face (C) F09-Real Face	Test Face (D) F05-Real Face
ID %	**10.80**	**24.60**	**9.20**	**55.40***

*The correct face

Face Resemblance Test

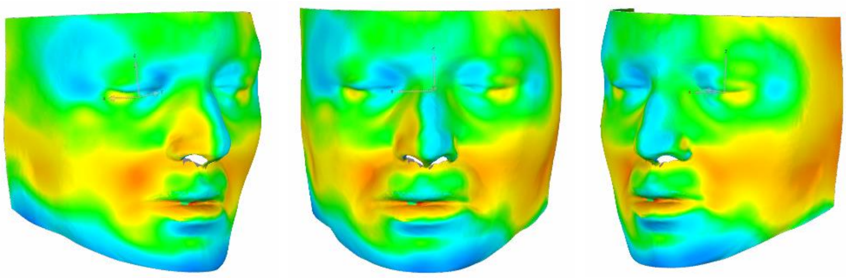


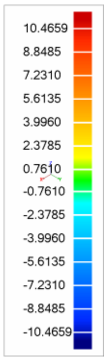


4.66

F05-Real Face F05-Reconstructed Face F05-Overall Resemblance Score

Objective Face Resemblance Test





2.199

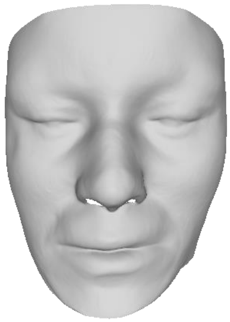
F05-Objective Surface Difference Histogram (color Map) F05-Surface Distance

4- Female Case- F06

Data

Facial Template
F-Av(21-30)Y

F06-Skull



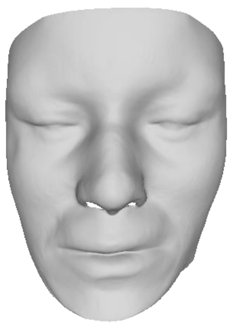
F06-Reconstructed Face

F06-Real Face

Face Pool Test
N/A

Face Resemblance Test

F06-Real Face

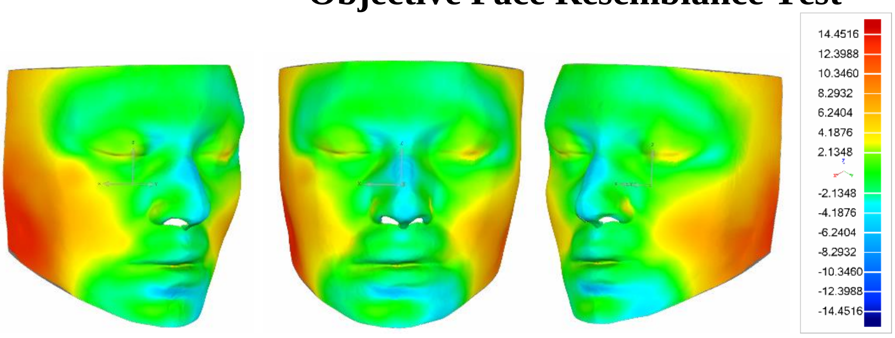


F06-Reconstructed Face

3.78

F06-Overall Resemblance Score

Objective Face Resemblance Test



F06-Objective Surface Difference Histogram (color Map)

4.209

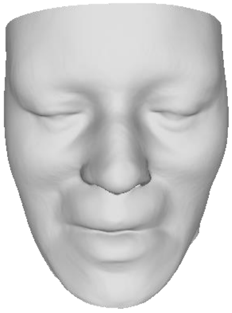
F06-Surface Distance

5- Female Case- F09

Data

Facial Template
F-Av(21-30)Y

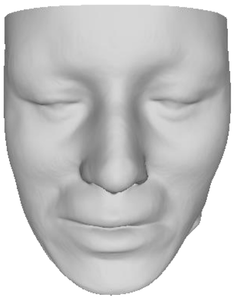
F09-Skull



F09-Reconstructed Face

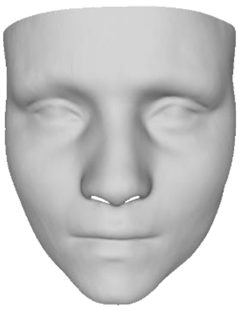
F09-Real Face

Face Pool Test



F09-Reconstructed Face

Test Face (A)
F08-Real Face



Test Face (B)
F09-Real Face

Test Face (C)
F06-Real Face

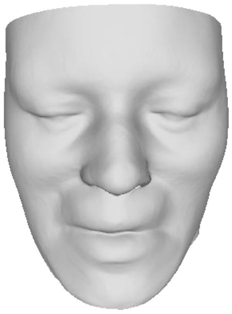
Test Face (D)
F07-Real Face

	Test Face (A)	Test Face (B)	Test Face (C)	Test Face (D)
ID %	**29.20%**	**26.20%***	**27.70%**	**16.90%**

*The correct face

Face Resemblance Test

F09-Real Face

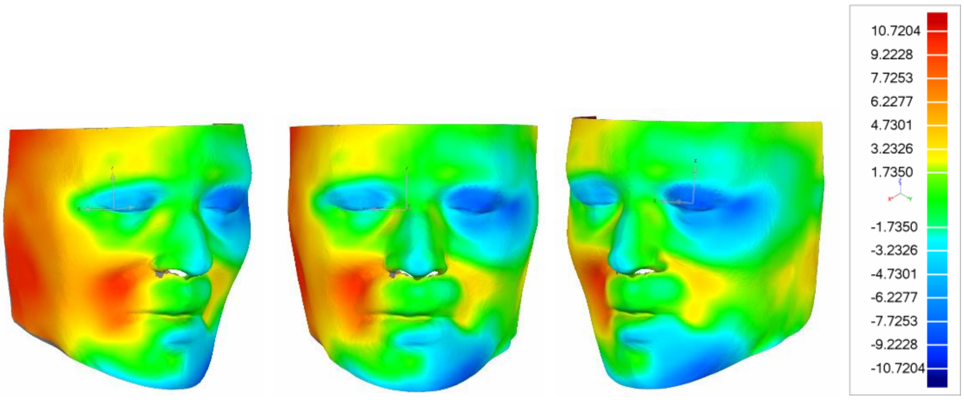


F09-Reconstructed Face

2.55

F09-Overall Resemblance Score

Objective Face Resemblance Test



F09-Objective Surface Difference Histogram (color Map)

3.726

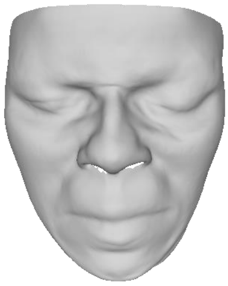
F09-Surface Distance

6- Female Case- F13

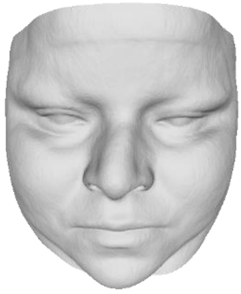
Data

Facial Template
F-Av(31-40)Y

F13-Skull

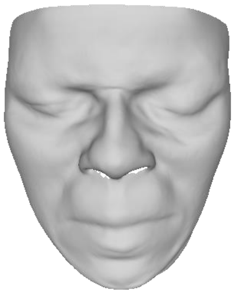


F13-Reconstructed Face



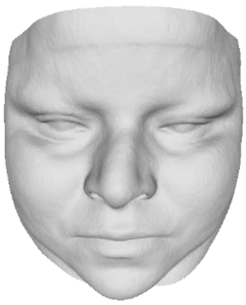
F13-Real Face

Face Pool Test

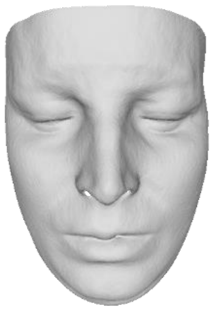


F13-Reconstructed Face

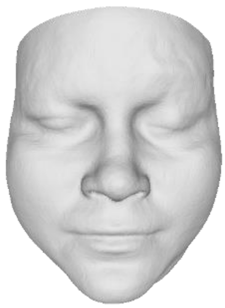
Test Face (A)
F15-Real Face



Test Face (B)
F13-Real Face



Test Face (C)
F12-Real Face

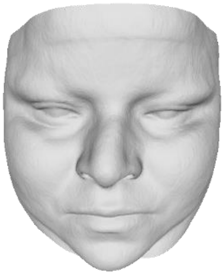


Test Face (D)
F16-Real Face

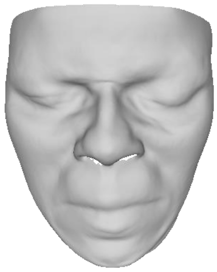
	Test Face (A)	Test Face (B)	Test Face (C)	Test Face (D)
ID %	**50.00%**	**21.10%***	**14.50%**	**14.50%**

*The correct face

Face Resemblance Test



F13-Real Face

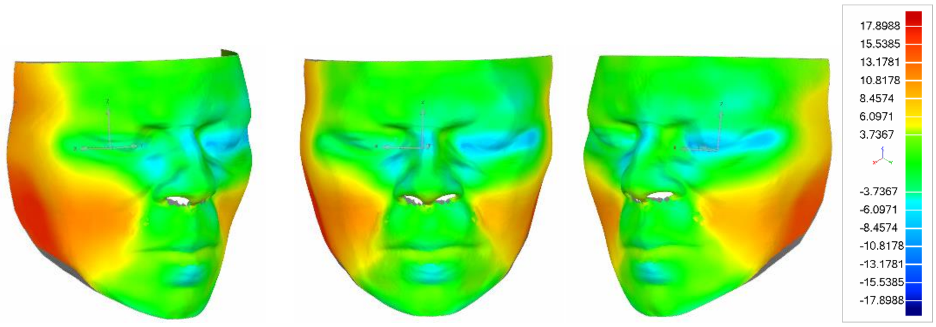


F13-Reconstructed Face

3.18

F13-Overall Resemblance Score

Objective Face Resemblance Test



F13-Objective Surface Difference Histogram (color Map)

5.436

F13-Surface Distance

7- Female Case- F14

Data

Facial Template
F-Av(31-40)Y

F14-Skull

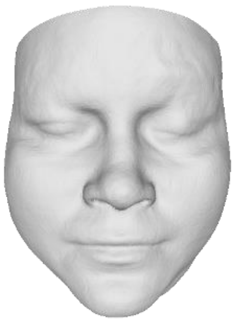
F14-Reconstructed Face

F14-Real Face

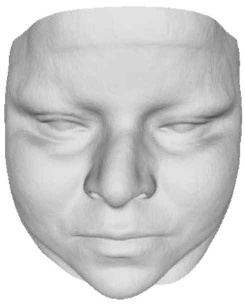
Face Pool Test

F14-Reconstructed Face

Test Face (A)
F19-Real Face



Test Face (B)
F16-Real Face



Test Face (C)
F13-Real Face

Test Face (D)
F14-Real Face

ID %	**34.30%**	**17.90%**	**6.00%**	**41.80%***

*The correct face

Face Resemblance Test

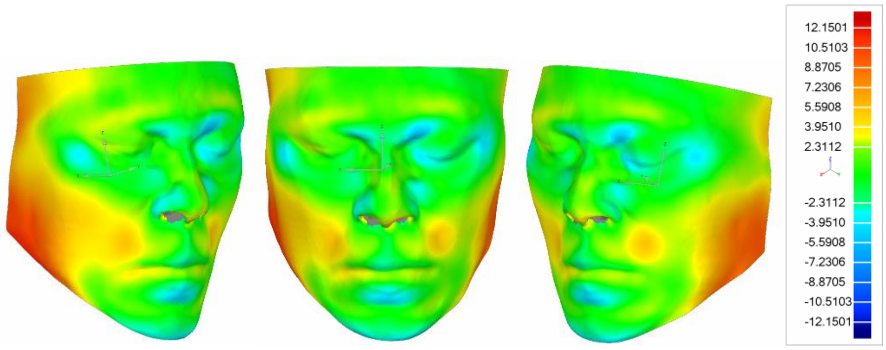
F14-Real Face

F14-Reconstructed Face

4.68

F14-Overall Resemblance Score

Objective Face Resemblance Test



F14-Objective Surface Difference Histogram (color Map)

2.881

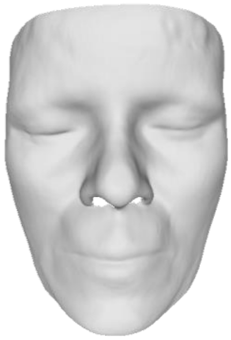
F14-Surface Distance

8- Female Case- F17

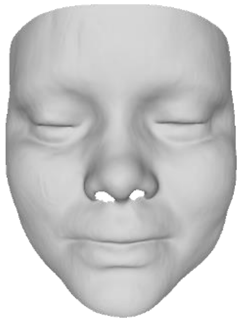
Data

Facial Template
F-Av(41-50)Y

F17-Skull

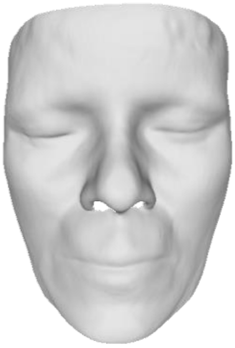


F17-Reconstructed Face

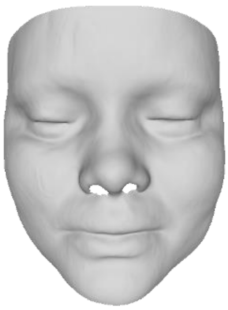


F17-Real Face

Face Pool Test



F17-Reconstructed Face



Test Face (A)
F17-Real Face

Test Face (B)
F21-Real Face

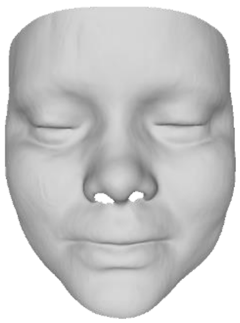
Test Face (C)
F18-Real Face

Test Face (D)
F23-Real Face

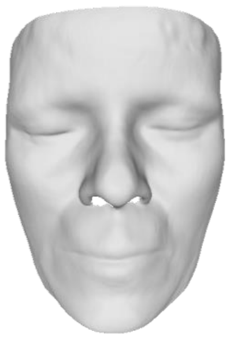
ID %	**20.00%***	**16.90%**	**35.40%**	**27.70%**

*The correct face

Face Resemblance Test



F17-Real Face

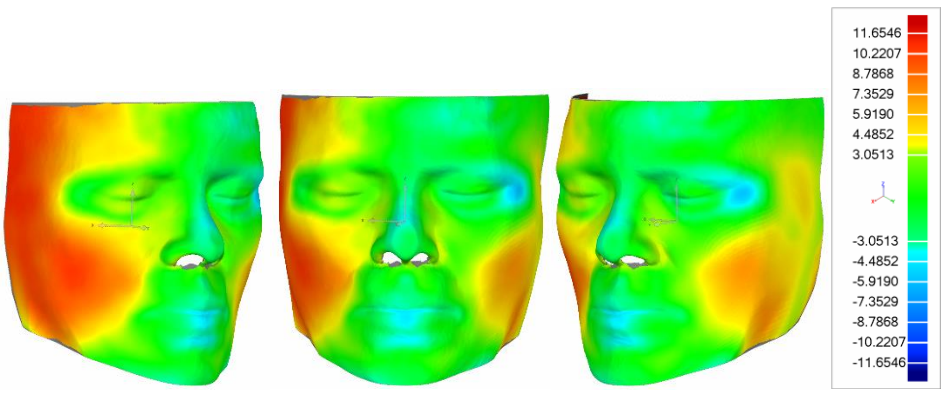


F17-Reconstructed Face

3.63

F17-Overall Resemblance Score

Objective Face Resemblance Test



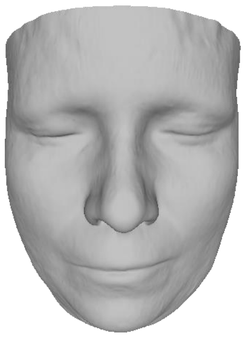
F17-Objective Surface Difference Histogram (color Map)

3.922

F17-Surface Distance

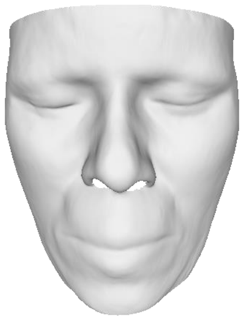
9- Female Case- F20

Data

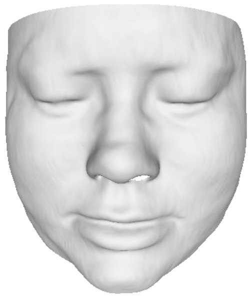


Facial Template
F-Av(41-50)Y

F20-Skull



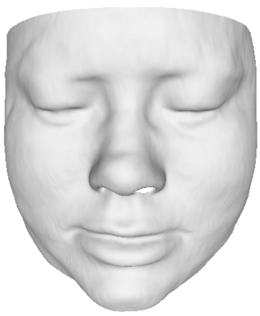
F20-Reconstructed Face



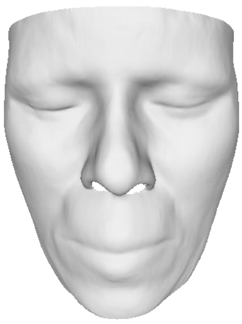
F20-Real Face

Face Pool Test
N/A

Face Resemblance Test



F20-Real Face

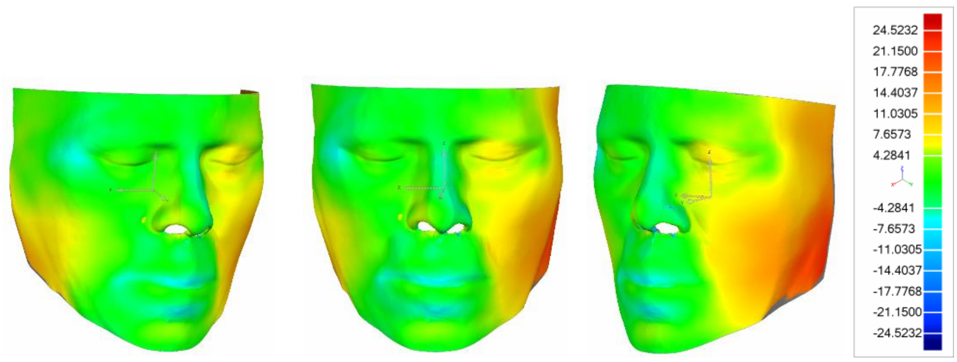


F20-Reconstructed Face

2.73

F20-Overall Resemblance Score

Objective Face Resemblance Test



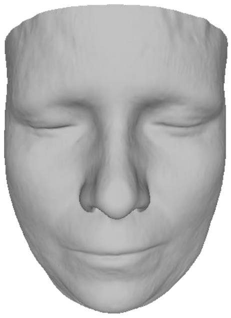
F20-Objective Surface Difference Histogram (color Map)

5.702

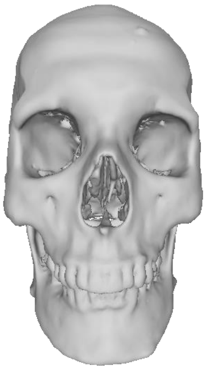
F20-Surface Distance

10- Female Case- F21

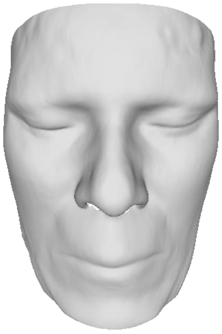
Data



Facial Template
F-Av(41-50)Y



F21-Skull



F21-Reconstructed Face

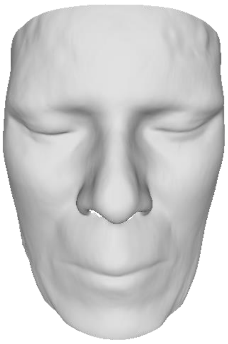
F21-Real Face

Face Pool Test

N/A

Face Resemblance Test

F21-Real Face

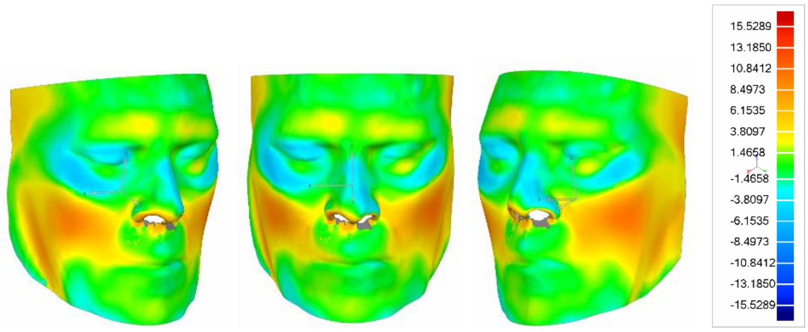


F21-Reconstructed Face

3.44

F21-Overall Resemblance Score

Objective Face Resemblance Test



F21-Objective Surface Difference Histogram (color Map)

2.984

F21-Surface Distance

11- Female Case- F23

Data

Facial Template
F-Av(41-50)Y

F23-Skull

F23-Reconstructed Face

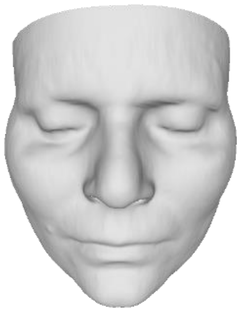
F23-Real Face

Face Pool Test

F23-Reconstructed Face

Test Face (A)
F21-Real Face

Test Face (B)
F19-Real Face



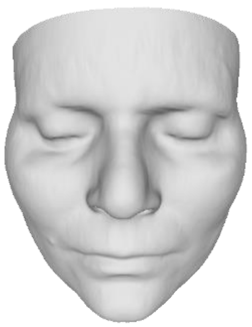
Test Face (C)
F23-Real Face

Test Face (D)
F24-Real Face

	Test Face (A)	Test Face (B)	Test Face (C)	Test Face (D)
ID %	**7.90%**	**5.30%**	**17.10%***	**68.40%**

*The correct face

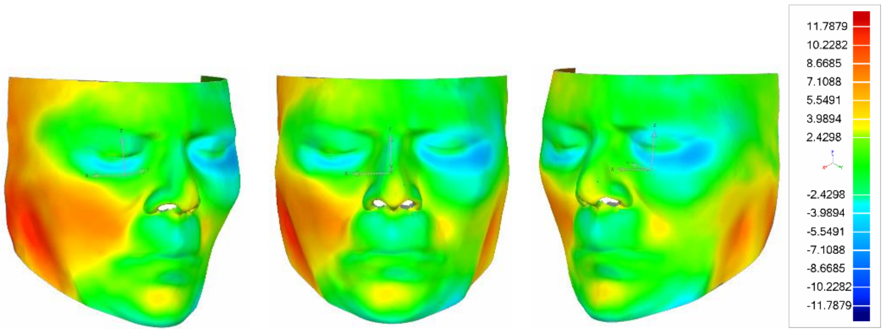
Face Resemblance Test



4.65

F23-Real Face F23-Reconstructed Face F23-Overall Resemblance Score

Objective Face Resemblance Test

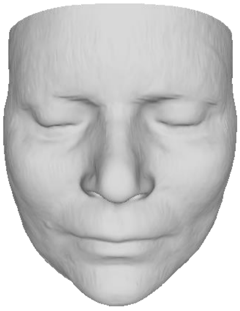


3.266

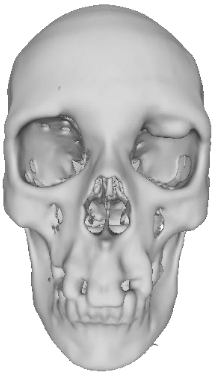
F23-Objective Surface Difference Histogram (color Map) F23-Surface Distance

12- Female Case- F25

Data

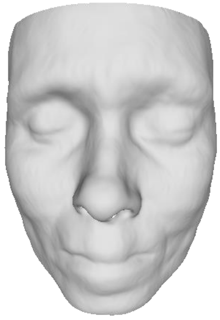


Facial Template
F-Av(>50)Y



F25-Skull

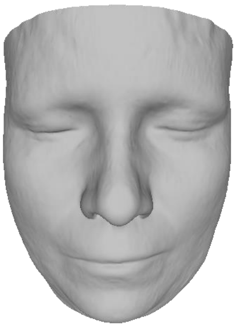
F25-Reconstructed Face



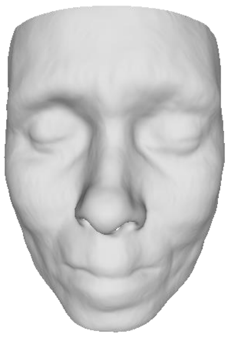
F25-Real Face

Face Pool Test

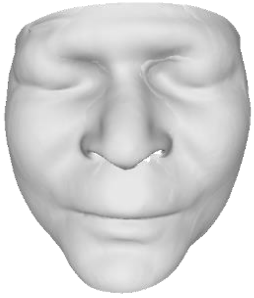
F25-Reconstructed Face



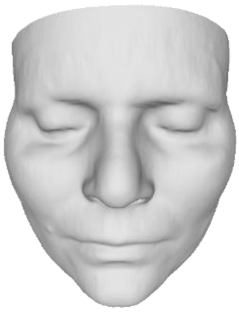
Test Face (A)
F-Av(41-50)Y



Test Face (B)
F25-Real Face



Test Face (C)
F26-Real Face



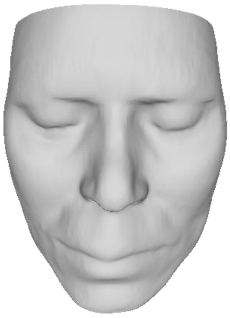
Test Face (D)
F23-Real Face

	Test Face (A)	Test Face (B)	Test Face (C)	Test Face (D)
ID %	**32.80%**	**29.90%***	**9.00%**	**28.40%**

*The correct face

Face Resemblance Test

F25-Real Face

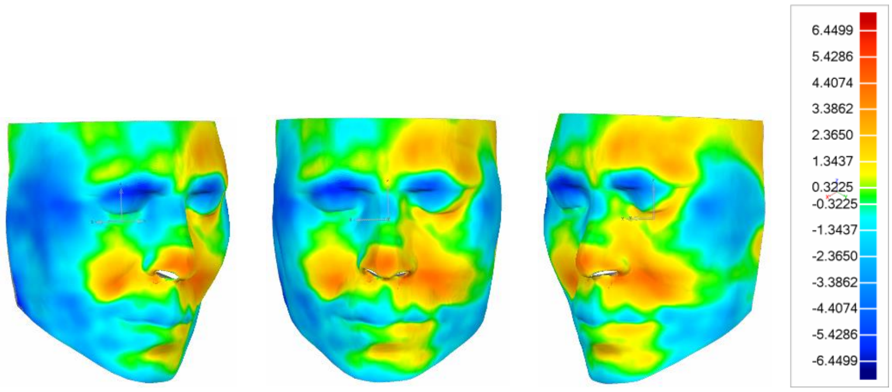


F25-Reconstructed Face

5.38

F25-Overall Resemblance Score

Objective Face Resemblance Test



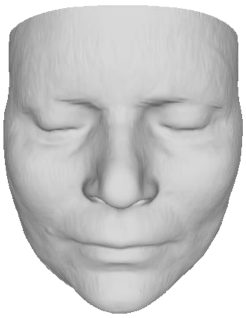
F25-Objective Surface Difference Histogram (color Map)

1.951

F25-Surface Distance

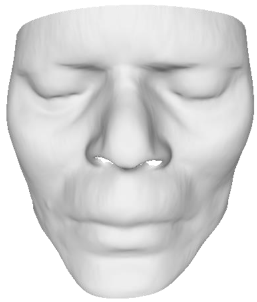
13- Female Case- F26

Data

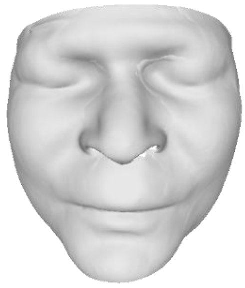


Facial Template
F-Av(>50)Y

F26-Skull

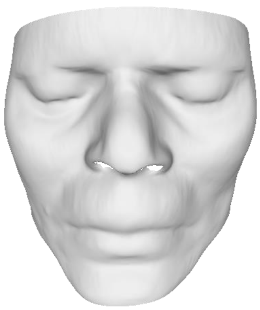


F26-Reconstructed Face



F26-Real Face

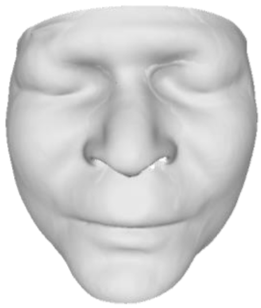
Face Pool Test



F26-Reconstructed Face

Test Face (A)
F-Av(41-50)Y

Test Face (B)
F25-Real Face



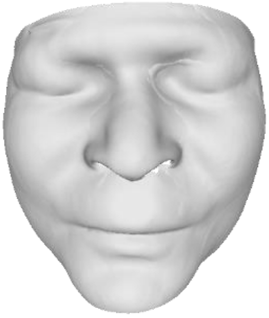
Test Face (C)
F26-Real Face

Test Face (D)
F24-Real Face

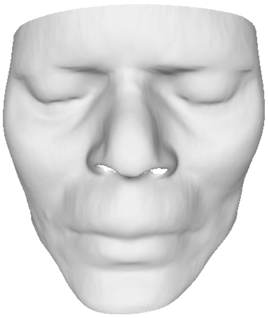
	Test Face (A)	Test Face (B)	Test Face (C)	Test Face (D)
ID %	**3.10%**	**7.70%**	**67.70%***	**21.50%**

*The correct face

Face Resemblance Test



F26-Real Face

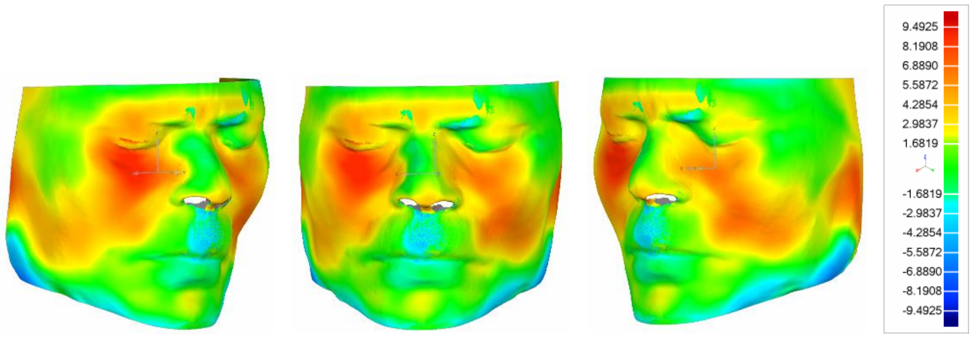


F26-Reconstructed Face

7.32

F26-Overall Resemblance Score

Objective Face Resemblance Test



F26-Objective Surface Difference Histogram (color Map)

2.355

F26-Surface Distance

14- Male Case- M02

Data

Facial Template
M-Av(16-20)Y

M02-Skull

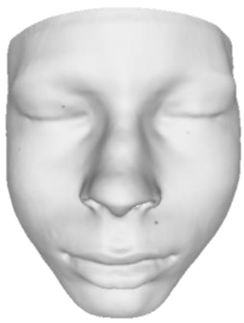
M02-Reconstructed Face

M02-Real Face

Face Pool Test

M02-Reconstructed Face

Test Face (A)
M01-Real Face



Test Face (B)
M03-Real Face

Test Face (C)
M02-Real Face

Test Face (D)
M06-Real Face

ID %	**16.40%**	**13.40%**	**47.80%***	**22.40%**

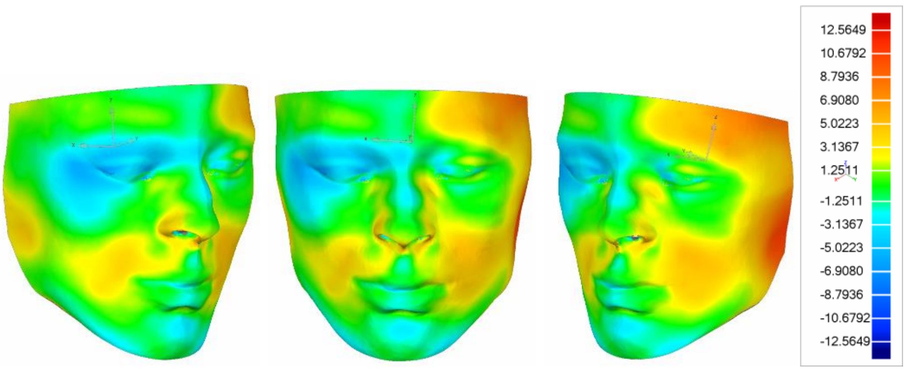
*The correct face

Face Resemblance Test

5.52

M02-Real Face M02-Reconstructed Face M02-Overall Resemblance Score

Objective Face Resemblance Test



2.825

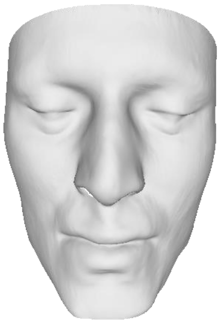
M02-Objective Surface Difference Histogram (color Map) M02-Surface Distance

15- Male Case- M05

Data

Facial Template
M-Av(21-30)Y

M05-Skull



M05-Reconstructed Face

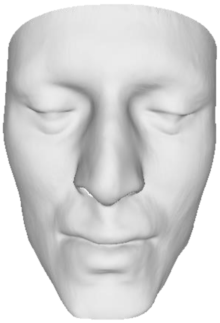
M05-Real Face

Face Pool Test

N/A

Face Resemblance Test

M05-Real Face

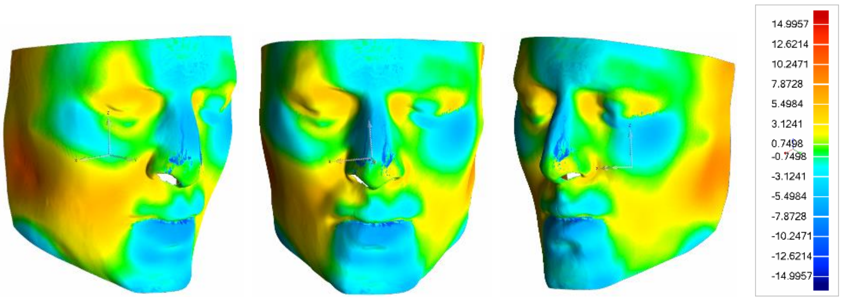


M05-Reconstructed Face

3.31

M05-Overall Resemblance Score

Objective Face Resemblance Test



M05-Objective Surface Difference Histogram (color Map)

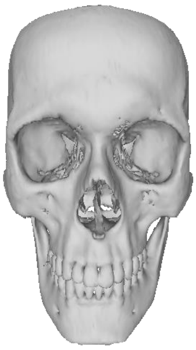
4.021

M05-Surface Distance

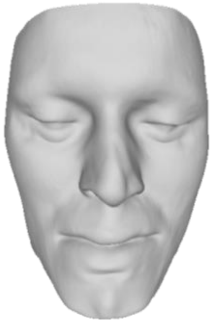
16- Male Case- M07

Data

Facial Template
M-Av(21-30)Y



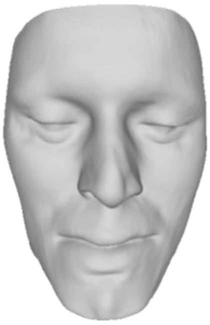
M07-Skull



M07-Reconstructed Face

M07-Real Face

Face Pool Test



M07-Reconstructed Face

Test Face (A)
M05-Real Face

Test Face (B)
M07-Real Face

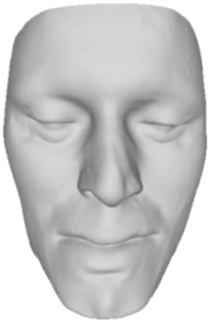
Test Face (C)
M06-Real Face

Test Face (D)
M08-Real Face

	Test Face (A)	Test Face (B)	Test Face (C)	Test Face (D)
ID %	**2.60%**	**81.60%***	**13.20%**	**1.30%**

*The correct face

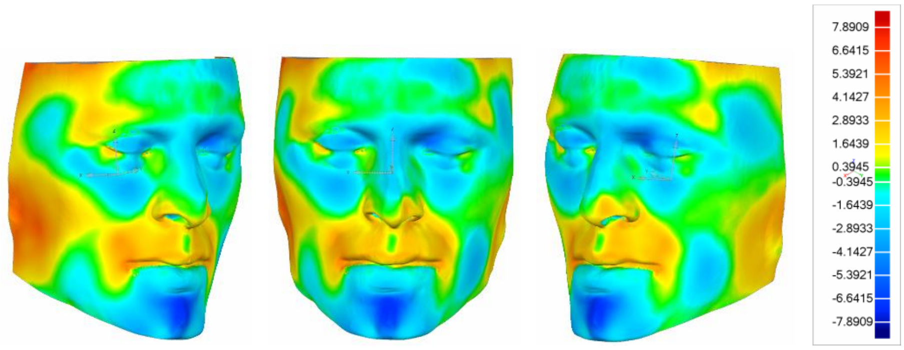
Face Resemblance Test



5.81

M07-Real Face M07-Reconstructed Face M07-Overall Resemblance Score

Objective Face Resemblance Test



1.980

M07-Objective Surface Difference Histogram (color Map) M07-Surface Distance

17- Male Case- M09

Data

Facial Template
M-Av(21-30)Y

M09-Skull

M09-Reconstructed Face

M09-Real Face

Face Pool Test

N/A

Face Resemblance Test

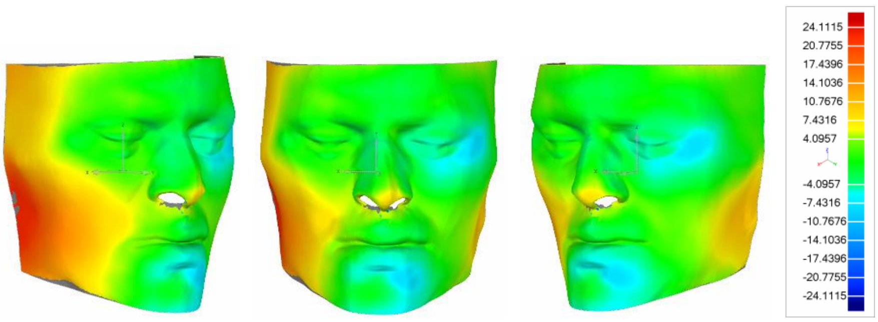
M09-Real Face

M09-Reconstructed Face

3.66

M09-Overall Resemblance Score

Objective Face Resemblance Test



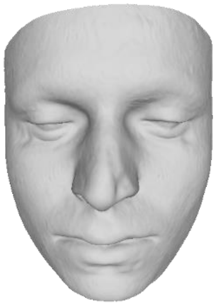
M09-Objective Surface Difference Histogram (color Map)

6.241

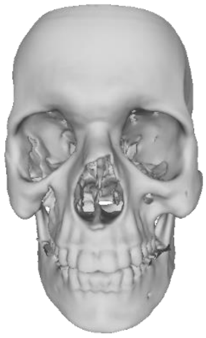
M09-Surface Distance

18- Male Case- M11

Data



Facial Template
M-Av(21-30)Y

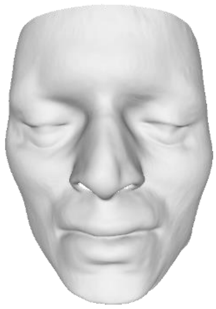


M11-Skull

M11-Reconstructed Face

M11-Real Face

Face Pool Test

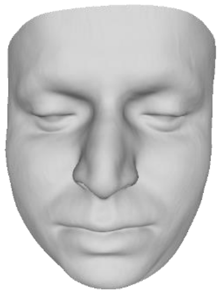


M11-Reconstructed Face

Test Face (A)
M11-Real Face

Test Face (B)
M12-Real Face

Test Face (C)
M09-Real Face



Test Face (D)
M10-Real Face

	Test Face (A)	Test Face (B)	Test Face (C)	Test Face (D)
ID %	**17.90%***	**19.40%**	**20.90%**	**40.30%**

*The correct face

Face Resemblance Test

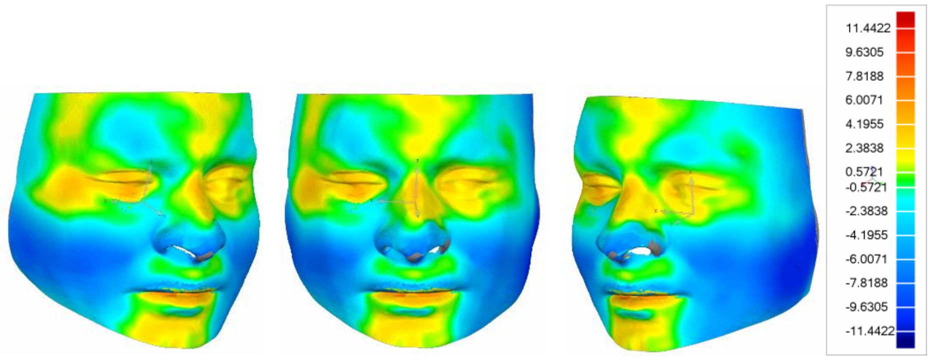
M11-Real Face

M11-Reconstructed Face

3.31

M11-Overall Resemblance Score

Objective Face Resemblance Test



M11-Objective Surface Difference Histogram (color Map)

3.206

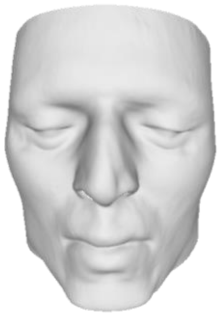
M11-Surface Distance

19- Male Case- M12

Data

Facial Template
M-Av(21-30)Y

M12-Skull



M12-Reconstructed Face

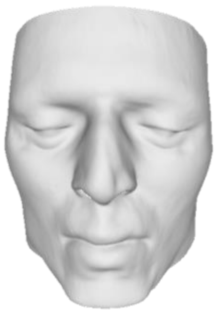
M12-Real Face

Face Pool Test

N/A

Face Resemblance Test

M12-Real Face

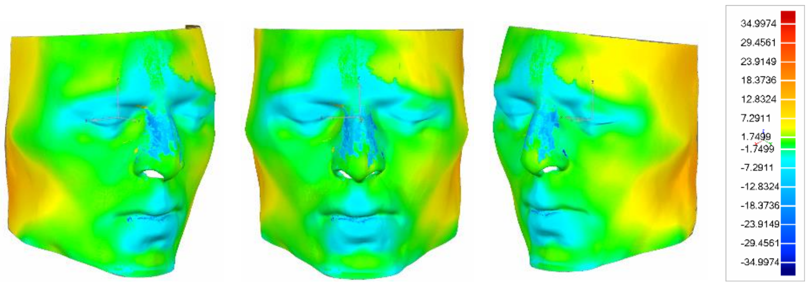


M12-Reconstructed Face

4.27

M12-Overall Resemblance Score

Objective Face Resemblance Test



M12-Objective Surface Difference Histogram (color Map)

6.337

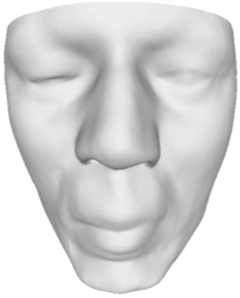
M12-Surface Distance

20- Male Case- M14

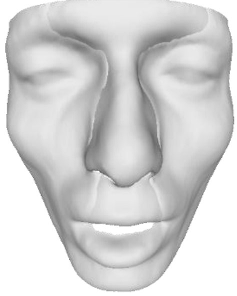
Data

Facial Template
M-Av(31-40)Y

M14-Skull

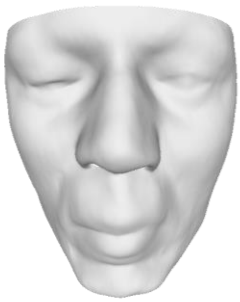


M14-Reconstructed Face



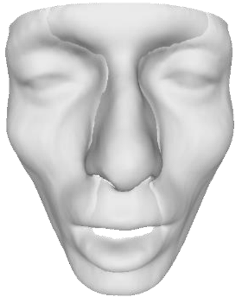
M14-Real Face

Face Pool Test



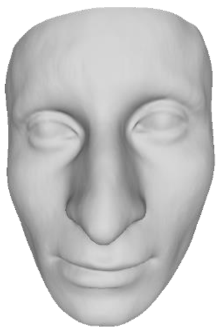
M14-Reconstructed Face

Test Face (A)
M20-Real Face



Test Face (B)
M14-Real Face

Test Face (C)
M16-Real Face



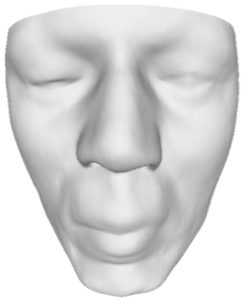
Test Face (D)
M08-Real Face

	Test Face (A)	Test Face (B)	Test Face (C)	Test Face (D)
ID %	**9.20%**	**49.20%***	**35.40%**	**6.20%**

*The correct face

Face Resemblance Test

M14-Real Face

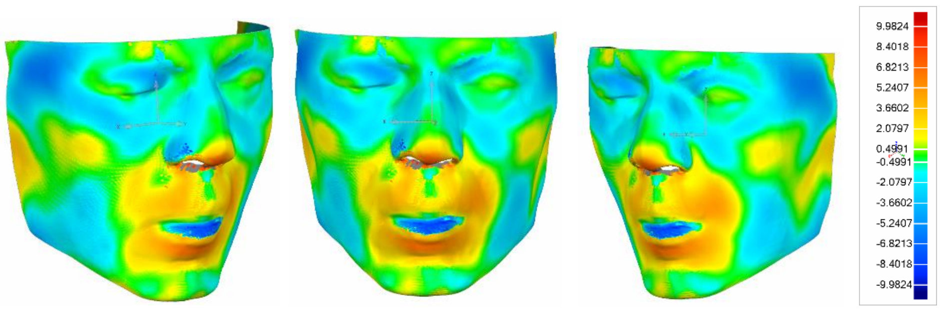


M14-Reconstructed Face

6.51

M14-Overall Resemblance Score

Objective Face Resemblance Test



M14-Objective Surface Difference Histogram (color Map)

2.642

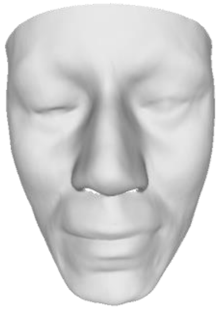
M14-Surface Distance

21- Male Case- M15

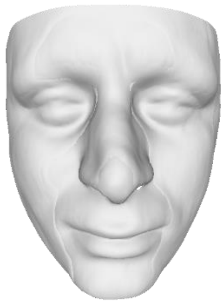
Data

Facial Template
M-Av(31-40)Y

M15-Skull



M15-Reconstructed Face

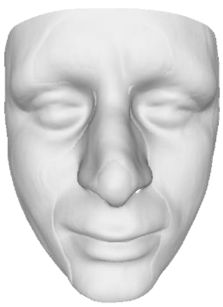


M15-Real Face

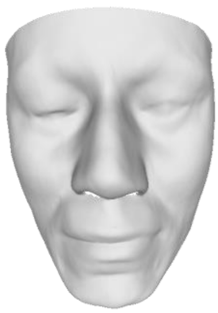
Face Pool Test

N/A

Face Resemblance Test



M15-Real Face

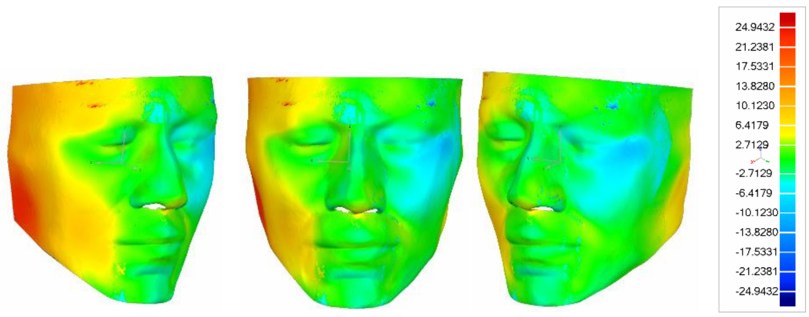


M15-Reconstructed Face

4.54

M15-Overall Resemblance Score

Objective Face Resemblance Test



M15-Objective Surface Difference Histogram (color Map)

5.449

M15-Surface Distance

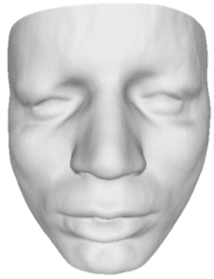
22- Male Case- M17

Data


Facial Template
M-Av(31-40)Y

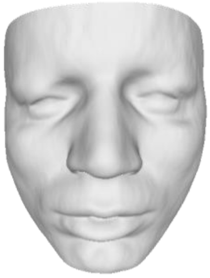

M17-Skull


M17-Reconstructed Face


M17-Real Face

Face Pool Test
N/A

Face Resemblance Test

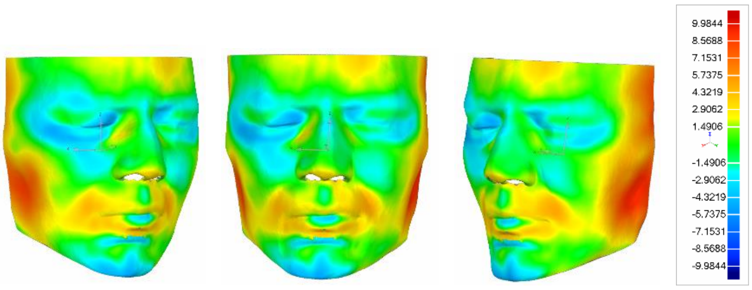

M17-Real Face


M17-Reconstructed Face

4.93

M17-Overall Resemblance Score

Objective Face Resemblance Test



M17-Objective Surface Difference Histogram (color Map)

2.804

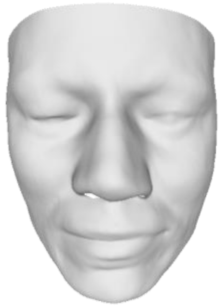
M17-Surface Distance

23- Male Case- M19

Data

Facial Template
M-Av(31-40)Y

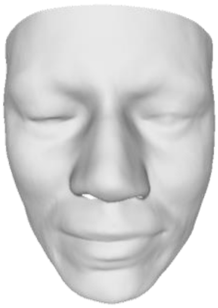
M19-Skull



M19-Reconstructed Face

M19-Real Face

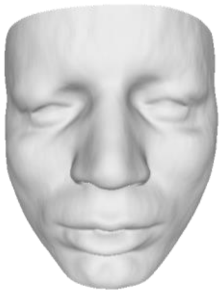
Face Pool Test



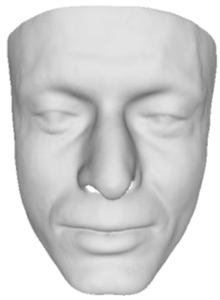
M19-Reconstructed Face

Test Face (A)
M18-Real Face

Test Face (B)
M19-Real Face



Test Face (C)
M17-Real Face

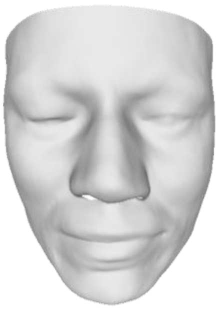


Test Face (D)
M21-Real Face

	Test Face (A)	Test Face (B)	Test Face (C)	Test Face (D)
ID %	**36.90%**	**9.20%***	**33.80%**	**20.00%**

*The correct face

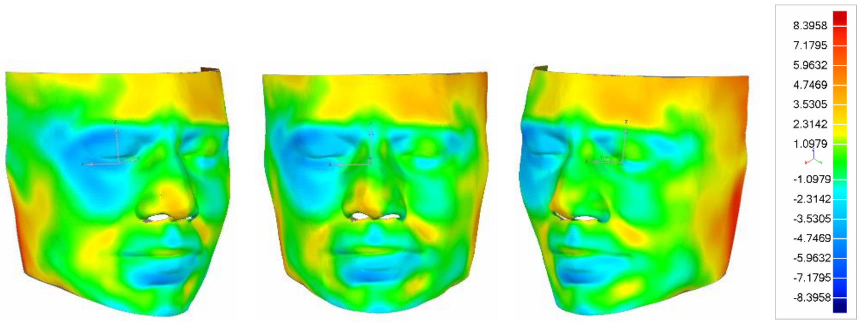
Face Resemblance Test



4.95

M19-Real Face M19-Reconstructed Face M19-Overall Resemblance Score

Objective Face Resemblance Test



2.060

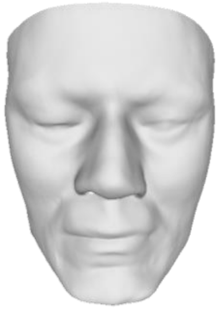
M19-Objective Surface Difference Histogram (color Map) M19-Surface Distance

24- Male Case- M21

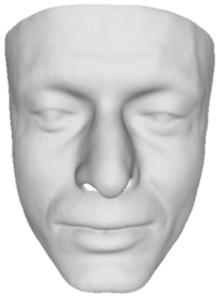
Data

Facial Template
M-Av(31-40)Y

M21-Skull



M21-Reconstructed Face

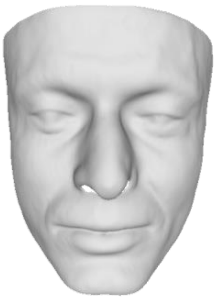


M21-Real Face

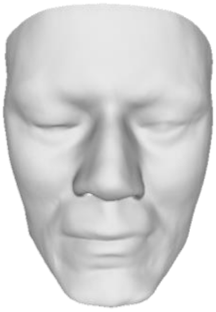
Face Pool Test

N/A

Face Resemblance Test



M21-Real Face

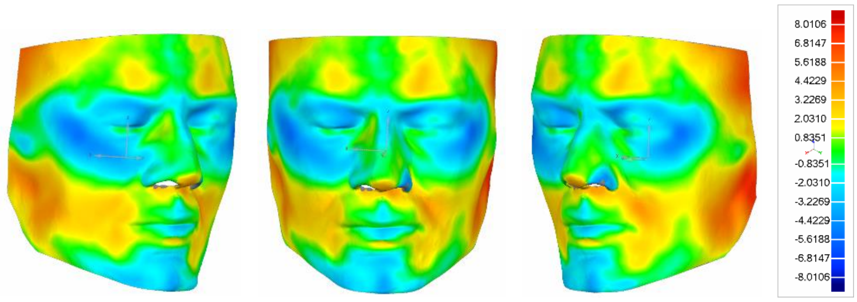


M21-Reconstructed Face

5.26

M21-Overall Resemblance Score

Objective Face Resemblance Test



M21-Objective Surface Difference Histogram (color Map)

2.378

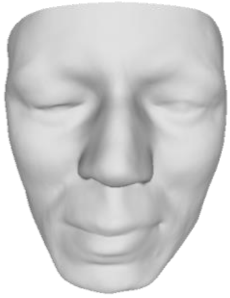
M21-Surface Distance

25- Male Case- M22

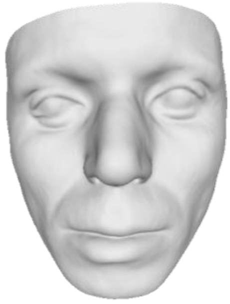
Data

Facial Template
M-Av(31-40)Y

M22-Skull

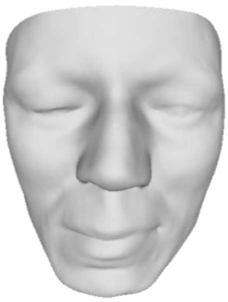


M22-Reconstructed Face

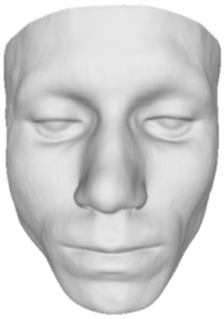


M22-Real Face

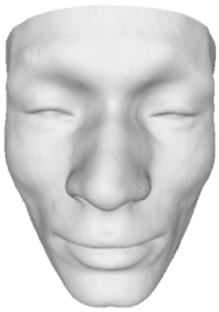
Face Pool Test



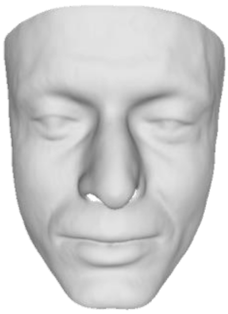
M22-Reconstructed Face



Test Face (A)
M20-Real Face



Test Face (B)
M16-Real Face



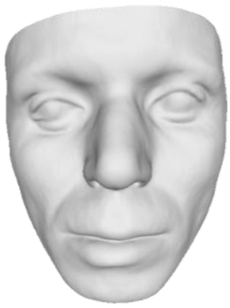
Test Face (C)
M21-Real Face

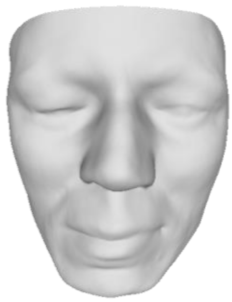
Test Face (D)
M22-Real Face

ID %	**16.40%**	**43.30%**	**16.40%**	**23.90%***

*The correct face

Face Resemblance Test

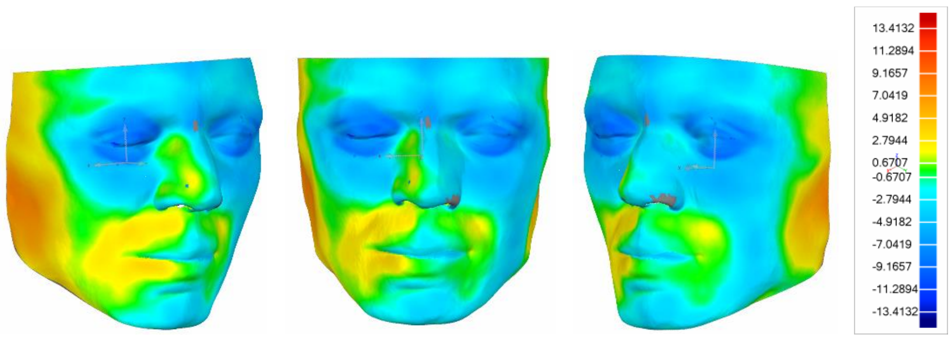




5.45

M22-Real Face   M22-Reconstructed Face   M22-Overall Resemblance Score

Objective Face Resemblance Test

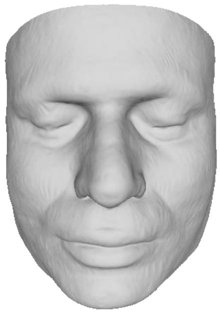


3.391

M22-Objective Surface Difference Histogram (color Map)   M22-Surface Distance

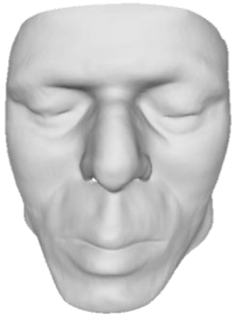
26- Male Case- M23

Data

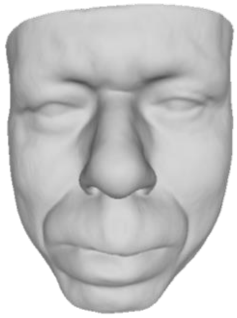


Facial Template
M-Av(41-50)Y

M23-Skull

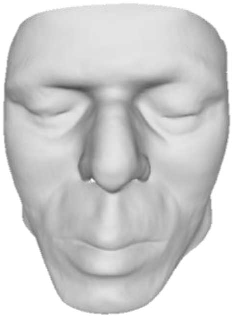


M23-Reconstructed Face

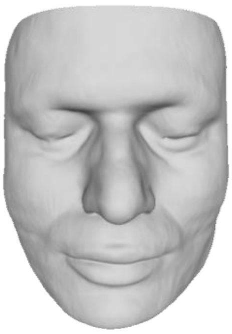


M23-Real Face

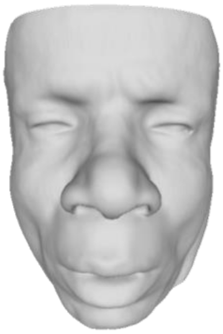
Face Pool Test



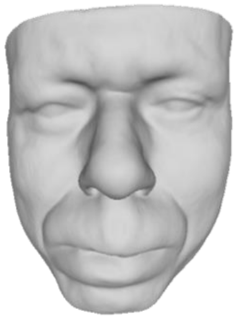
M23-Reconstructed Face



Test Face (A)
M25-Real Face



Test Face (B)
M29-Real Face



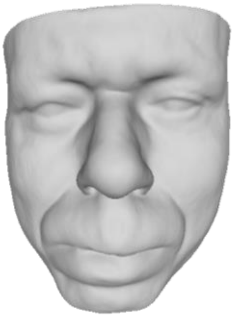
Test Face (C)
M23-Real Face

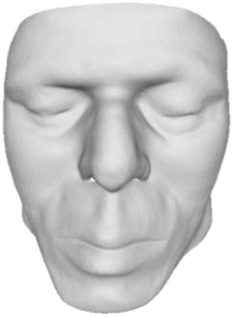
Test Face (D)
M24-Real Face

	Test Face (A)	Test Face (B)	Test Face (C)	Test Face (D)
ID %	**27.60%**	**3.90%**	**61.80%***	**6.60%**

*The correct face

Face Resemblance Test

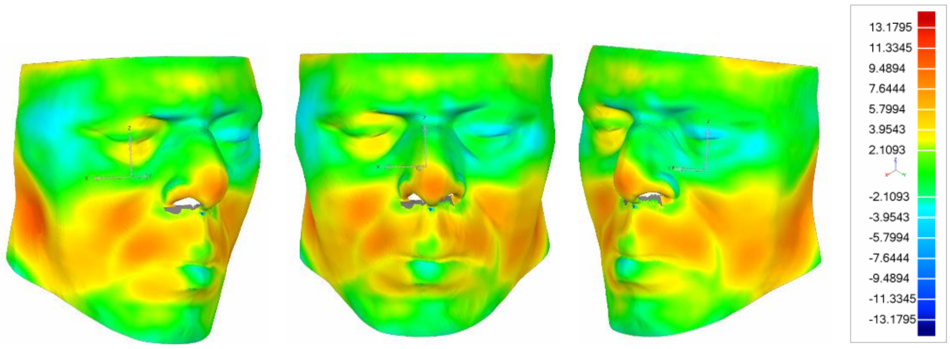




5.22

M23-Real Face M23-Reconstructed Face M23-Overall Resemblance Score

Objective Face Resemblance Test



2.852

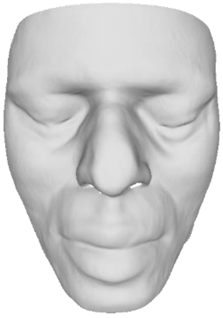
M23-Objective Surface Difference Histogram (color Map) M23-Surface Distance

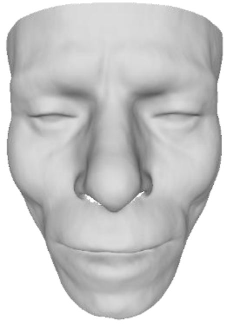
27- Male Case- M26

Data

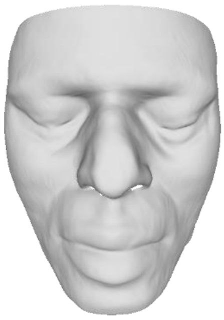

Facial Template
M-Av(>50)Y


M26-Skull

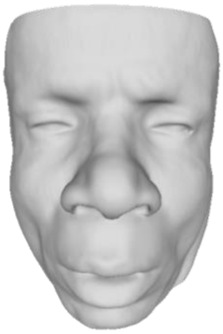

M26-Reconstructed Face

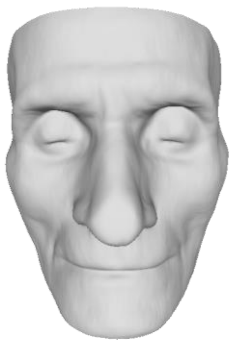

M26-Real Face

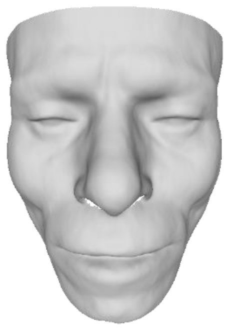
Face Pool Test


M26-Reconstructed Face


Test Face (A)
M28-Real Face


Test Face (B)
M29-Real Face

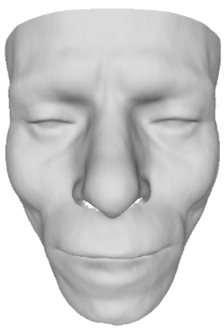

Test Face (C)
M30-Real Face


Test Face (D)
M26-Real Face

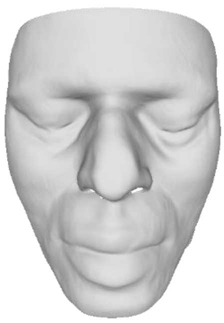
	Test Face (A)	Test Face (B)	Test Face (C)	Test Face (D)
ID %	**9.20%**	**12.30%**	**26.20%**	**52.30%***

*The correct face

Face Resemblance Test



M26-Real Face

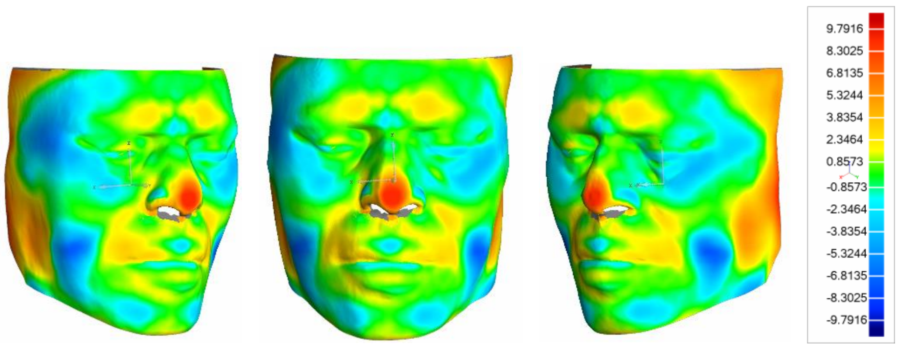


M26-Reconstructed Face

4.92

M26-Overall Resemblance Score

Objective Face Resemblance Test



M26-Objective Surface Difference Histogram (color Map)

2.655

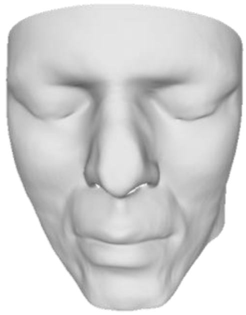
M26-Surface Distance

28- Male Case- M28

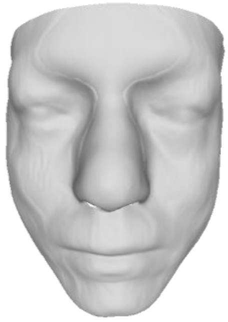
Data

Facial Template
M-Av(41-50)Y

M28-Skull



M28-Reconstructed Face

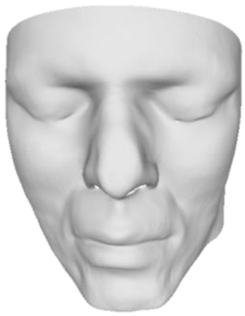


M28-Real Face

Face Pool Test
N/A

Face Resemblance Test

M28-Real Face

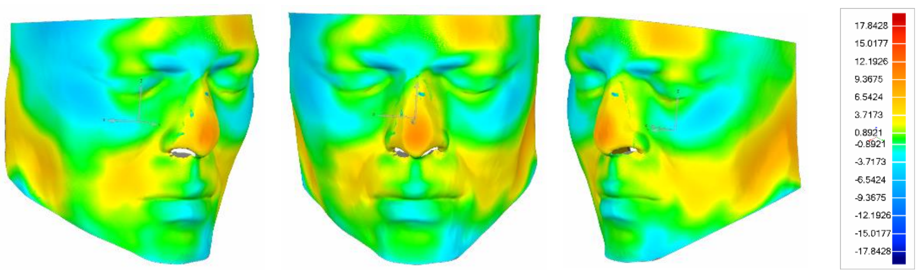


M28-Reconstructed Face

3.66

M28-Overall Resemblance Score

Objective Face Resemblance Test



M28-Objective Surface Difference Histogram (color Map)

2.844

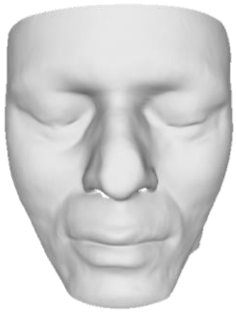
M28-Surface Distance

29- Male Case- M31

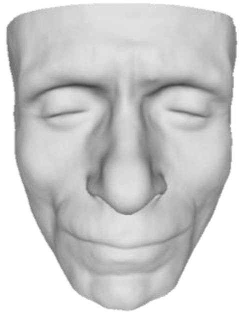
Data

Facial Template
M-Av(41-50)Y

M31-Skull

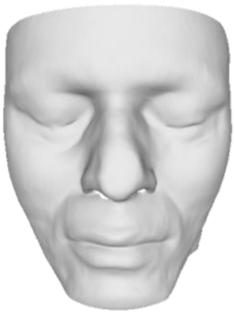


M31-Reconstructed Face

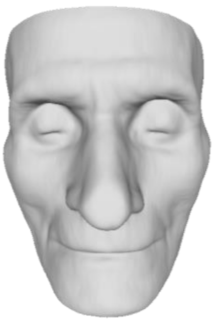


M31-Real Face

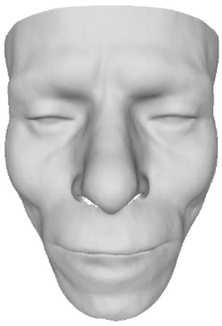
Face Pool Test



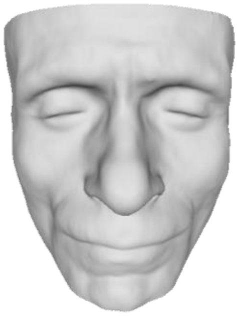
M31-Reconstructed Face



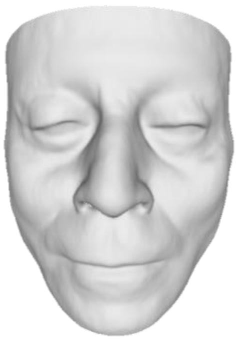
Test Face (A)
M30-Real Face



Test Face (B)
M26-Real Face



Test Face (C)
M31-Real Face

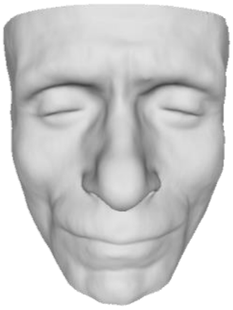


Test Face (D)
M33-Real Face

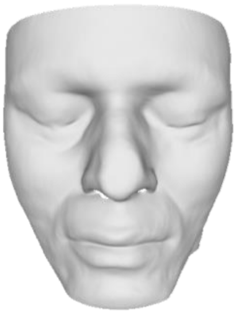
	Test Face (A)	Test Face (B)	Test Face (C)	Test Face (D)
ID %	**16.90%**	**18.50%**	**30.80%***	**33.80%**

*The correct face

Face Resemblance Test



M31-Real Face

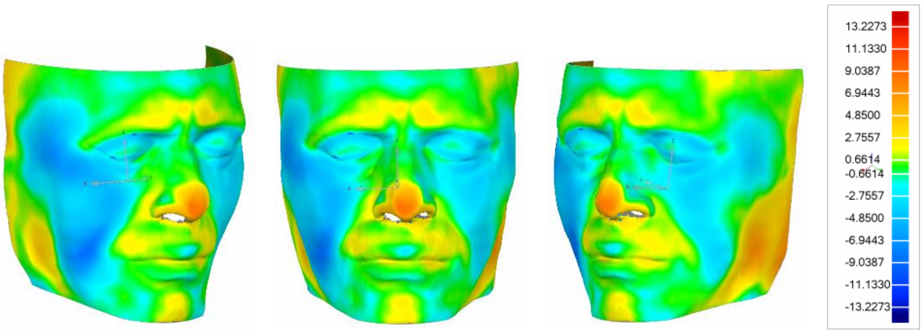


M31-Reconstructed Face

3.99

M31-Overall Resemblance Score

Objective Face Resemblance Test



M31-Objective Surface Difference Histogram (color Map)

2.742

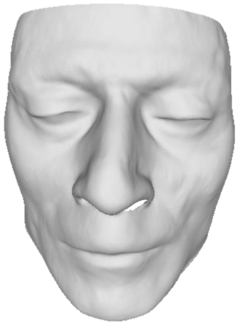
M31-Surface Distance

30- Male Case- M34

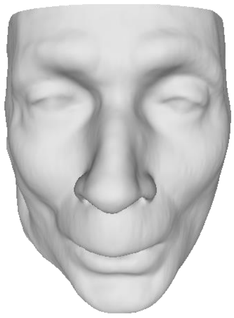
Data

Facial Template
M-Av(41-50)Y

M34-Skull

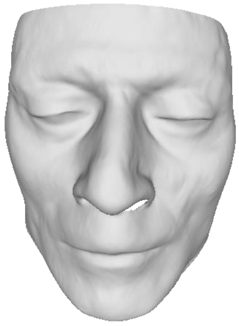


M34-Reconstructed Face

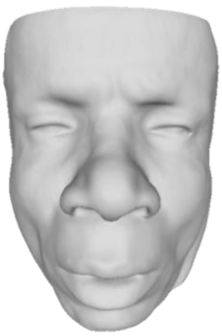


M34-Real Face

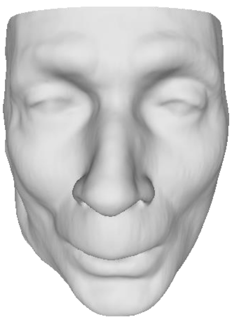
Face Pool Test



M34-Reconstructed Face

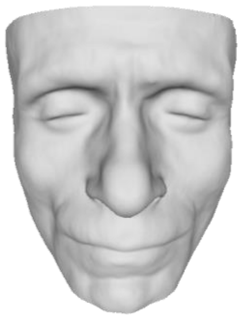


Test Face (A)
M29-Real Face



Test Face (B)
M34-Real Face

Test Face (C)
M30-Real Face

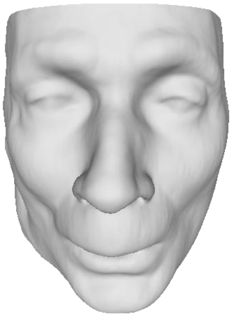


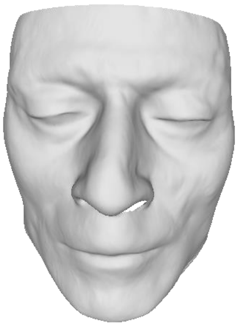
Test Face (D)
M31-Real Face

	Test Face (A)	Test Face (B)	Test Face (C)	Test Face (D)
ID %	**3.10%**	**47.70%***	**13.80%**	**35.40%**

*The correct face

Face Resemblance Test

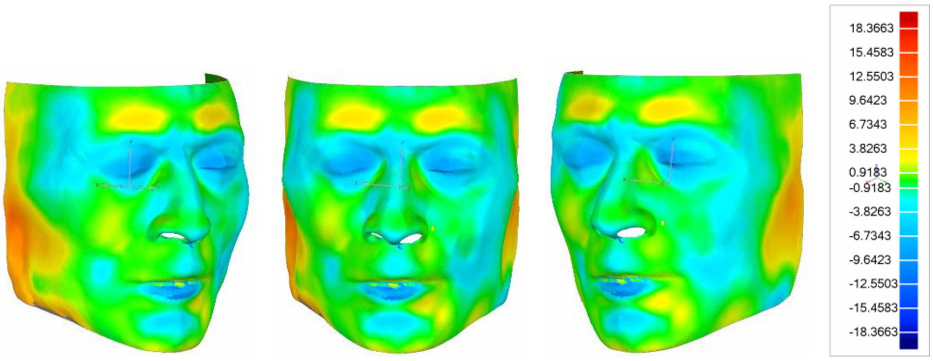




5.07

M34-Real Face M34-Reconstructed Face M34-Overall Resemblance Score

Objective Face Resemblance Test



3.070

M34-Objective Surface Difference Histogram (color Map) M34-Surface Distance

APPENDIX 20: "FACE FR" FACIAL RECONSTRUCTION SOFTWARE USER MANUAL

Software Designed By

Dr. Tim Niblett

Dr. Maria Vanezis

Manual Organised By

Dalia A. Abdou

Under the Supervision of

Professor Peter Vanezis

Professor Atholl Johnston

2015

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1 INTRODUCTION

For three-dimensional Forensic Facial Reconstruction (FFR), faces can be reconstructed from “inside outwards” by building facial muscles staring from the bone surface outwards towards the facial skin (Wilkinson, 2006; Lee et al., 2012), or from “outside inwards” by “wrapping” or “warping” a face template as a mask onto the skull (Quatrehomme et al., 1997; Vanezis, 2008).

“Face FR” Facial Reconstruction software was_designed by Dr Tim Niblett from the Turing Institute, Glasgow University in 1997 for the purpose of Forensic facial reconstruction research. The software was first used for Forensic facial reconstruction by Dr Maria Vanezis for her PhD thesis (Vanezis, 2008). The FR software provides facilities to view the digitised skulls and facial templates as 3-D scans. This software adopts the approach of facial reconstruction using facial templates.

Reconstruction is done by warping a 3D mesh of the face template onto a 3D skull mesh using a number of landmarks on both meshes. Each landmark on the face with a corresponding landmark on the skull (a landmark with the same name and side) is used to define the warp. The point on the skull landmark that is used is the top of the “peg”, so changing the orientation of the peg changes the reconstruction, as does changing the location of course.

There are several modes of warp available. All methods are point-based using the locations of corresponding landmarks. It is always the face which is warped. The term “warp” is used generically to include linear transformations, which is a general linear transform, including rotations, translations, scaling in each dimension.

The software gives the user the possibility to interactively manipulate the images as required. This could include; real time Rotation (to move the image in the main display window to rotate the view), and zooming -in and -out, translation (to centre objects on the screen), scaling (to scale the objects in the window), identifying, adding, moving or removing landmarks on its mesh to perform a reconstruction using a predefined set of tissue thicknesses, according to sex, ancestry, age and build (thin, medium or fat).

2 WINDOW LAYOUT

The software window layout is formed of the following windows.

2.1 THE MAIN WINDOW

Upon double clicking on the relative icon of the software application ".face", the configuration window (Figure 28) will appear followed by the main work window (Figure 29).

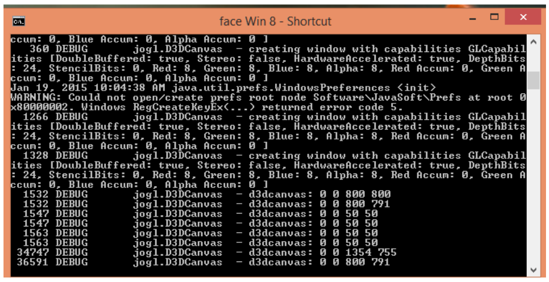


Figure 28: The configuration window of "Face FR" software

Figure 29: The main work window of "Face FR" software

The main window layout (Figure 30) consists of:

- The *3-D panel*, in which the 3-D skull and/or the face meshes can be loaded,

- A *slider* below the main 3-D panel allowing for zooming,
- A *menu bar* at the top, and
- *Aligning Buttons*; the top left 4 buttons which align the virtual camera to view the front, left, right, or back of the face & skull. This alignment is relative to the "default" coordinate system.

It also shows a side window (Section 2.2).

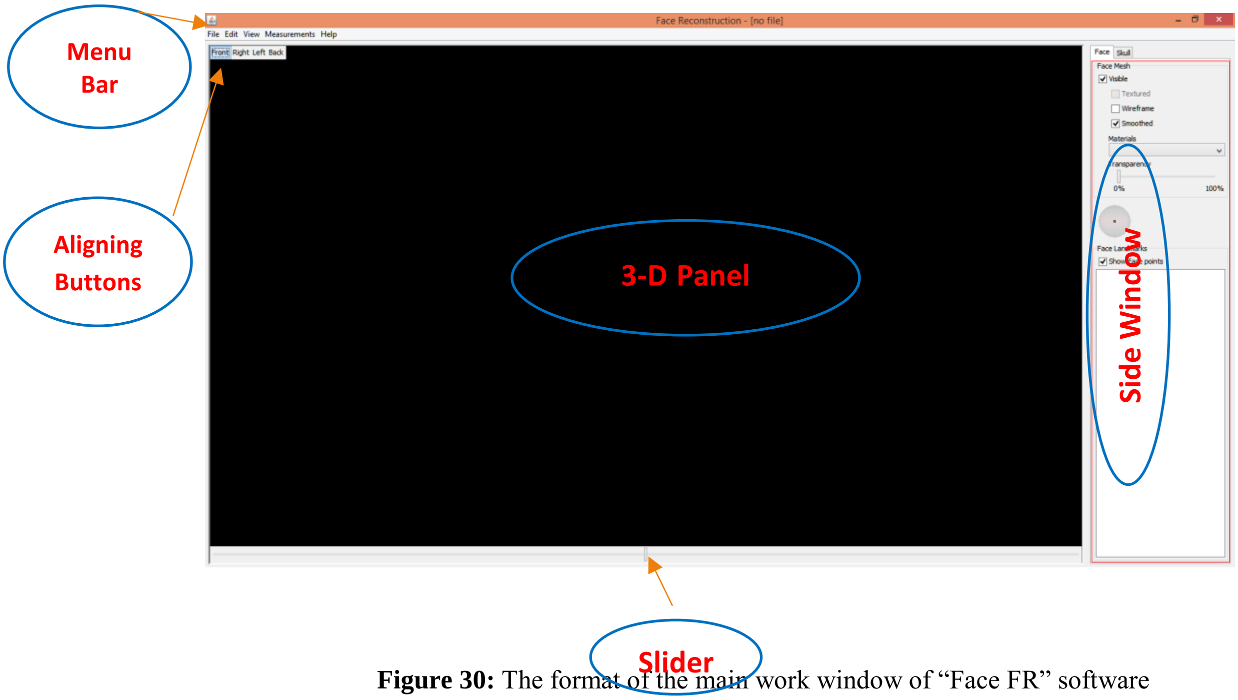


Figure 30: The format of the main work window of "Face FR" software

2.2 The Side Window

The side window (Figure 31) contains information about the face and skull meshes and their landmark sets. It contains two tabs; "Face" and "Skull". Pressing on these tabs brings either the face or the skull control panel to the front. Each tab shows 3 panels;

- The ***top*** panel contains mesh (face or skull) controls. It is possible to change the way the mesh is displayed and its transparency. The default is that the mesh is smoothed. In addition, hiding the face or the skull images is possible while working on the other by deactivating the "visible" button.

- The ***middle*** panel is a controller for the angle of the skull landmarks. Selecting a landmark and then moving this with the RIGHT button causes the angle of the landmark's "peg" to vary.
- The ***bottom*** panel contains the landmarks for the current tab (face or skull).

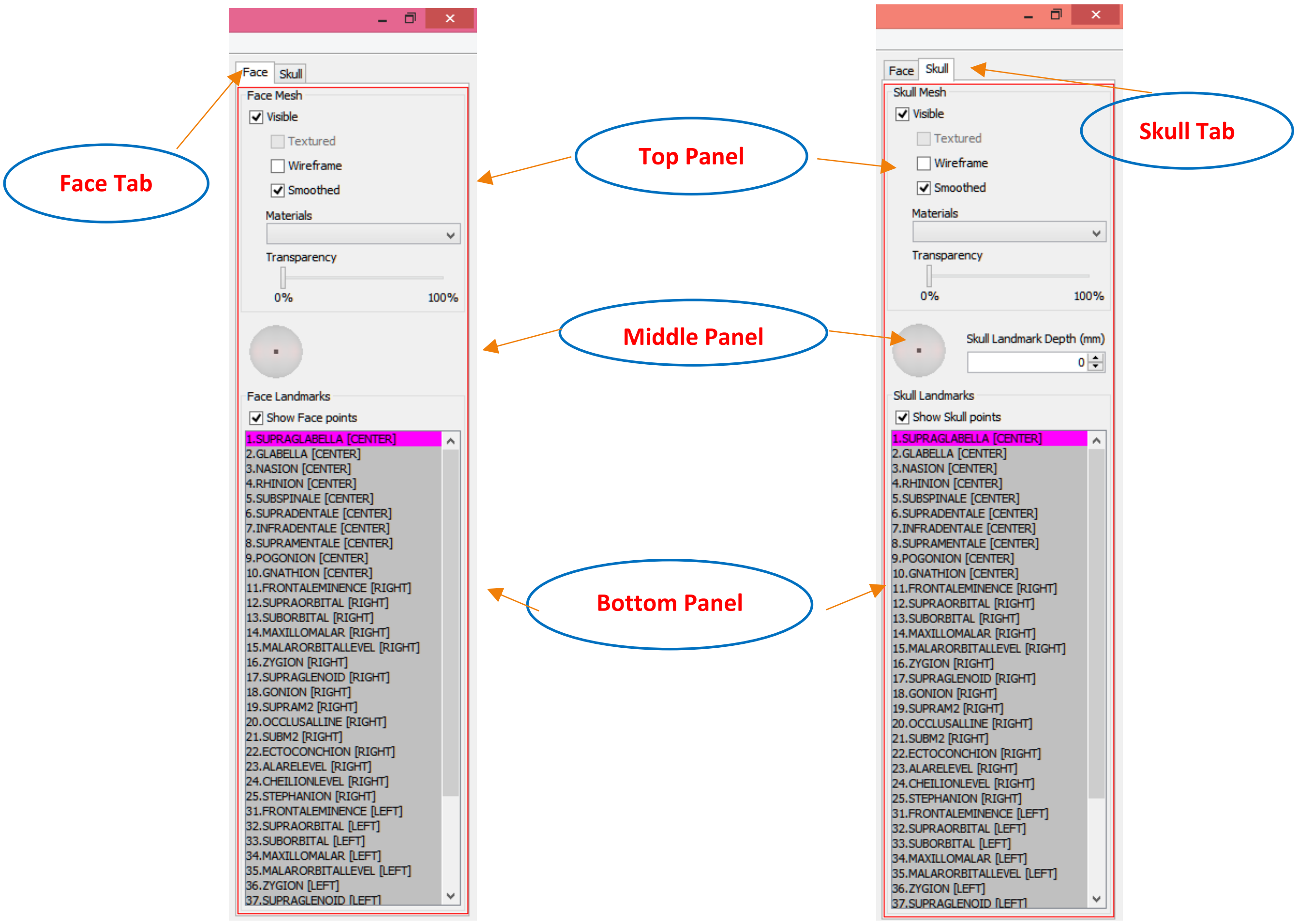


Figure 31: The format of the slide window

Landmarks are given locations when they are added to the face or skull. The location is on either the skull or face. Landmarks can be moved interactively.

Each landmark has a given name, and a side which should be one of “left”, “right”, or “centre”. The name and the side identify the landmark, and no two landmarks should have the same name and side.

Each landmark has a depth and orientation, shown graphically by a small peg with two ends, one at the skull and the other at the corresponding face point. The length of the peg corresponds to the facial soft tissue thickness at a given landmark.

The depths of the landmarks can be adjusted using the upper and lower arrows in the “Skull Landmarks Depths” window in the middle panel of the side window. Also, the direction of the landmarks can be adjusted using the Landmarks Rotating Cursor in the middle panel (Figure 32).

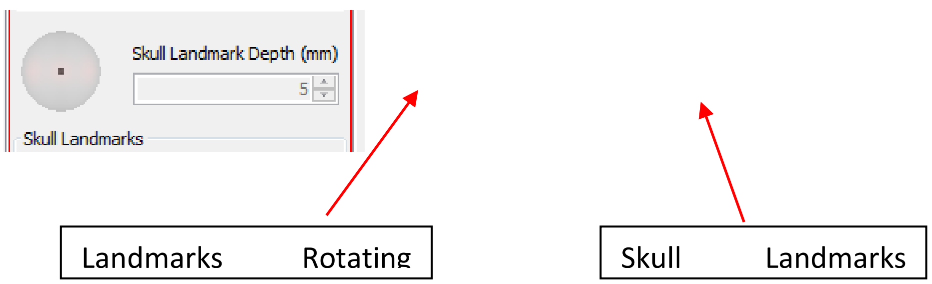
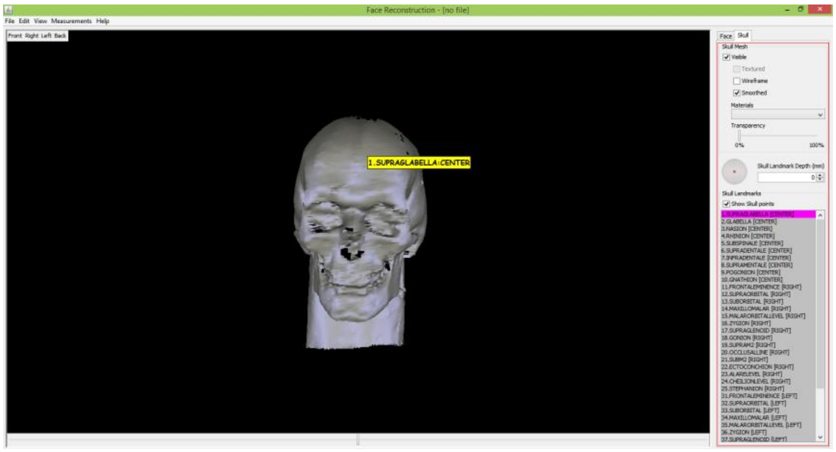


Figure 32: Skull landmarks depths and rotating cursor in the middle panel of the side window

A colour coding is used to describe the state of the landmark in the bottom panel of the side window as follows;

- **Grey**: The landmark has not been placed (Figure 33a).
- **Purple**: The landmark has not been placed & is the current landmark to be placed (Figure 33a). This means that if a landmark is placed with the relative mouse control, it will be this one (See Mouse Control: Section 2.3).
- **White**: The landmark has been placed (Figure 33b).
- **Blue**: The landmark has been placed & is currently selected (Figure 33b).
- **Red**: The landmark has not been placed yet (Figure 33b).



a

b

Figure 33 (a&b): Color coding for the placing the anatomical landmarks on the imported mesh

2.3 THE MENU BAR:

2.3.1 File Menu:

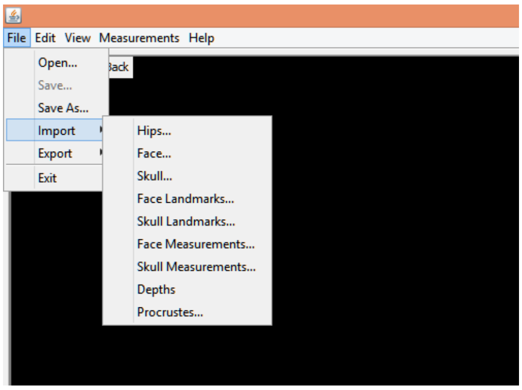


Figure 34: The file menu

Open	Open a session. This is a previously saved session with ".face" file extension. A session contains all or some of the data needed for a complete reconstruction. A file chooser dialog pops up.
Save	Save the current session (greyed out if there is no current session). The file containing the session is shown in the title bar.
Save As	Save the current session in a file which is selected in the file chooser dialog which pops up. This file will have the extension ".face".
Import	**Hips**: Import a face or a skull from a HIPS file.
	Face: Import a face mesh from a file. The file is either LSM or OBJ format.
	Skull: Import a skull mesh from a file. The file is either LSM or OBJ format.
	Face Landmarks: Import a set of landmark names for the face mesh. Only the names of the landmarks is imported, not their locations.

	Skull Landmarks: Import a set of landmark names for the skull mesh. Only the names of the landmarks is imported, not their locations.
	Face Measurements: Import a set of measurement specifications for the face.
	Skull Measurements: Import a set of measurement specifications for the skull.
	Depths: Import a file specifying depths for skull landmarks.
	Procrustes: Choose the best match of a series of face meshes using Procustes distance as the match criterion. A directory must be chosen which contains ".face" and/or ".hips" files. It is important that the file contain landmarks with positions as this is the data used to determine the Procrustes fits.
Export	**Image**: Export what can be seen in the 3-D window to a TIFF image. If only the face is wanted then the skull must be hidden, and the face points must be hidden.
Exit	Exit the application. A confirm window will pop up, just to make sure you want to exit.

2.3.2 **Edit menu:**

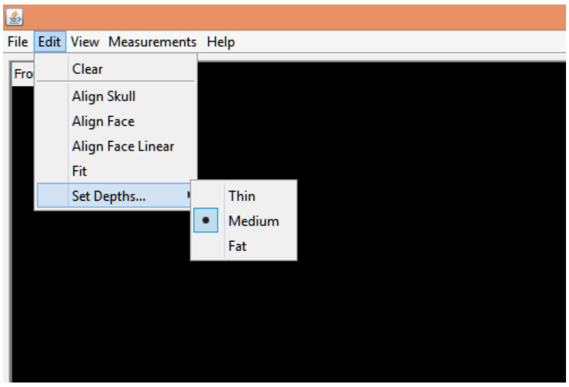


Figure 35: The edit menu

Clear	Clear any warps in the face mesh. This will return the face mesh to its original shape.
Align Skull	Use the skull landmarks to align the skull with respect to the standard axis coordinates. Information about which points are “left”, “right” and “center” is used here. The axes are OpenGL standard, with the X-axis increasing horizontally left to right, the Y-axis increasing vertically bottom to top, and the Z-axis decreasing near to far in the 3-D window.
Align Face	Aligns the face with the skull, using the common landmarks, and using a Procrustes transform to compute the best alignment. The alignment involves rotation, translation and scale.
Align Face Linear	Align the face with the skull using a full linear transform.
Fit	Perform a complete facial reconstruction, using the common landmarks between face and skull.
Set Depths	Set the depth of the points on the skull to be either “thin”, “medium”, or fat.

2.3.3 View Menu:

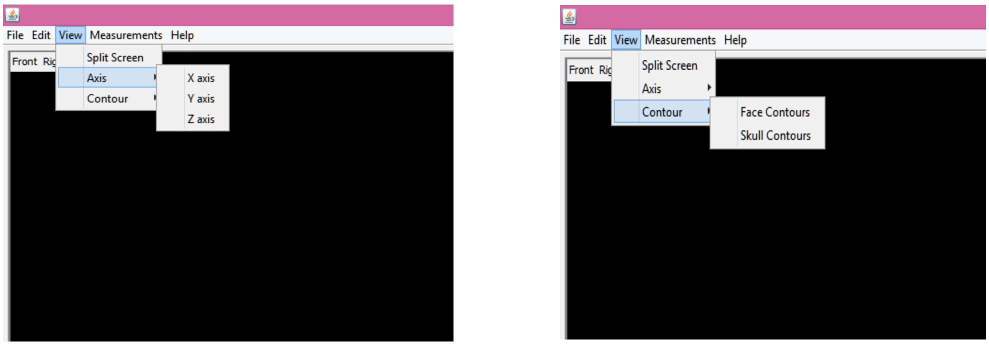


Figure 36: The view menu

Split Screen	A toggle to split the 3-D window into 2 independent parts (Figure 37). This is useful for precise location of landmarks.
Axis	Planes in the X, Y, and Z axes can be displayed. This helps with alignment of landmarks. Note that alignment of the skull and face is required before the axes can be relied upon. The X-axis is horizontal, with values increasing from left to right. The Y-axis is vertical, with values increasing from bottom to top. The Z-axis is the window depth, with values decreasing from near to far.
Contour	Show depth-based contours on skull or face in order to get a more precise view of the shapes when placing landmarks.

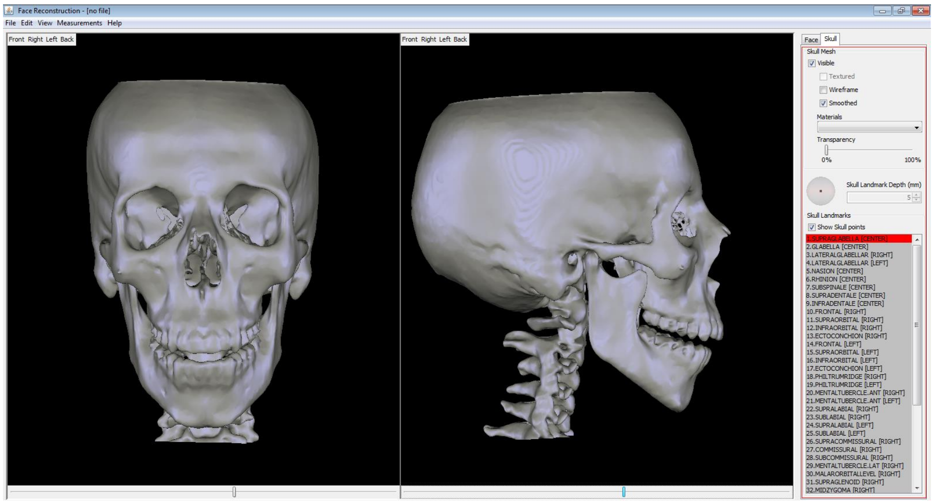


Figure 37: The main window in split screen view

2.3.4 **Measurements Menu:**

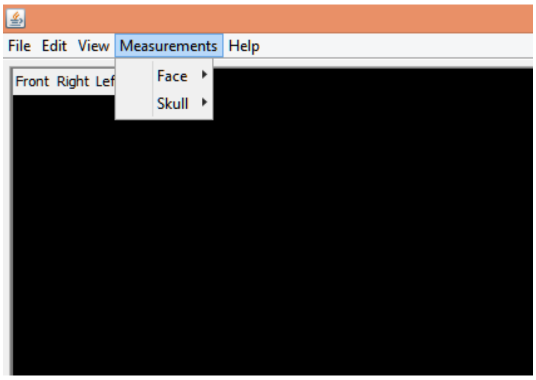


Figure 38: The measurements menu

Face	If face measurement definitions have been imported then these measurements can be selected and viewed on screen.
Skull	If skull measurement definitions have been imported then these measurements can be selected and viewed on screen.

2.3.5 **Help Menu:**

Figure 39: The help menu

Face Help	Describes the trajectory of the file imported as the facial template
Help	Not yet implemented. Will show the contents of this document.

3 THREE-DIMENSIONAL FORENSIC FACIAL RECONSTRUCTION USING "FR" SOFTWARE

The concept of digital reconstruction of a face from a skull via the software involves certain objects, which compose a session. These objects include; triangle meshes for both the skull (Figure 40), and the face template (Figure 41) together with sets of skull and face landmarks at given locations added to the face or skull and can be moved interactively.

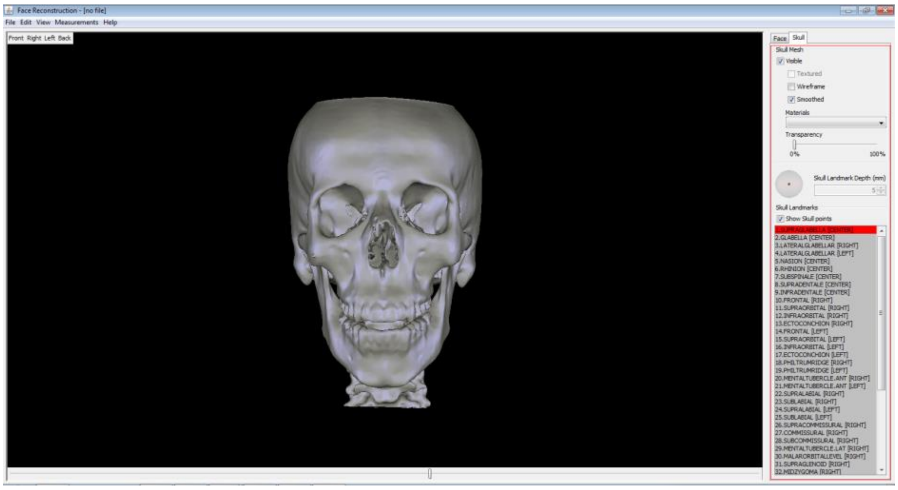


Figure 40: A three-dimensional skull mesh

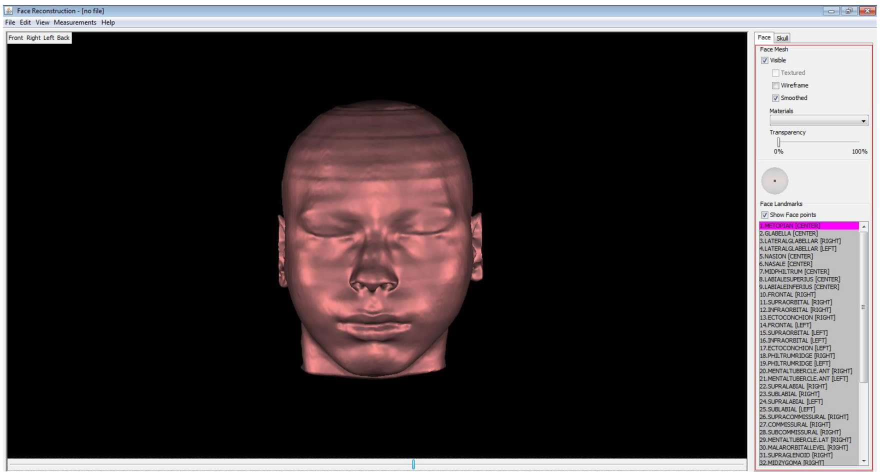


Figure 41: A three-dimensional face mesh

The skull and face meshes result from 3-D scans. It is possible to import them from third party files in .lsm, .hips, and .obj formats. In addition, images seen in the 3-D window can be exported to a .tiff image.

The objects can be viewed from three different vantage points at the same time (by default: left profile, anterior-posterior and right profile) to assist in the placement of landmarks, which are viewed in 3-D to view and alter their direction.

N.B. 01:

The default is that the mesh is smoothed.

N.B. 02:

Hiding the face or the skull images is possible while working on the other by deactivating the "visible" button.

N.B. 03:

The alpha-blending (mixed view) allows the operator to see where the skull and skull landmarks are in relation to the reconstructed face. It can be seen by adjusting the transparency of the meshes from the transparency slider in the top panel of the side window (Figure 42).

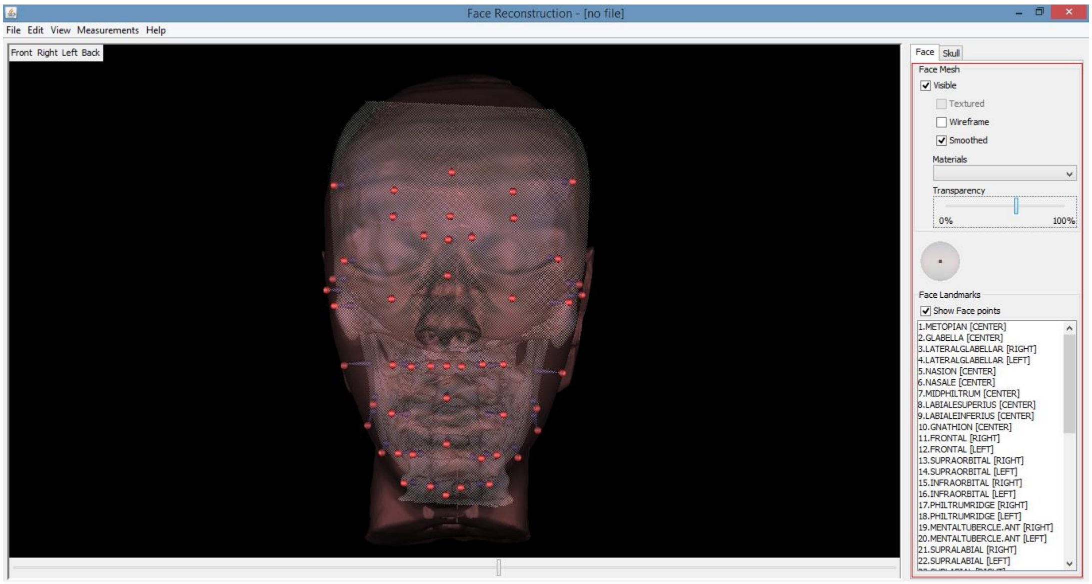


Figure 42: Mixed view showing the skull-face landmarks correlations

N.B. 04:

The following steps should be followed for the “Face FR” software to be used in Forensic Facial Reconstruction of any population. However, the files containing Skull and Face Landmarks and Depths are population specific and should be prepared prior to loading into the software.

N.B. 05:

The files containing the required information of skull landmarks’ names, depths, and face landmarks as well as the skull and the face meshes are loaded into the software separately from the file menu (Figure 43) in successive steps.

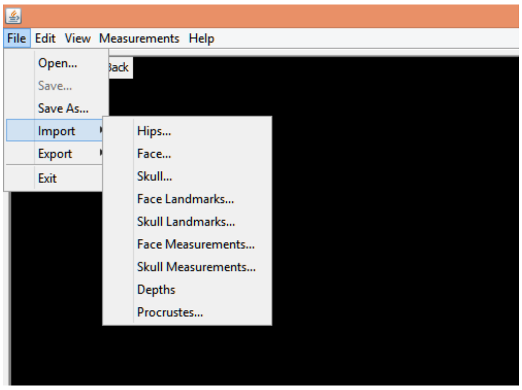


Figure 43: The import option of the file menu

3.1 STEP (1) IMPORTING SKULL LANDMARKS

To import the skull landmarks:

1) Make sure that the skull tab at top panel of the side window (on the right) is active (Figure 31).
2) Import the skull landmarks into the software by opening the file menu → import → skull landmarks (Figure 43).
3) A file browser dialog will pop up (Figure 44).

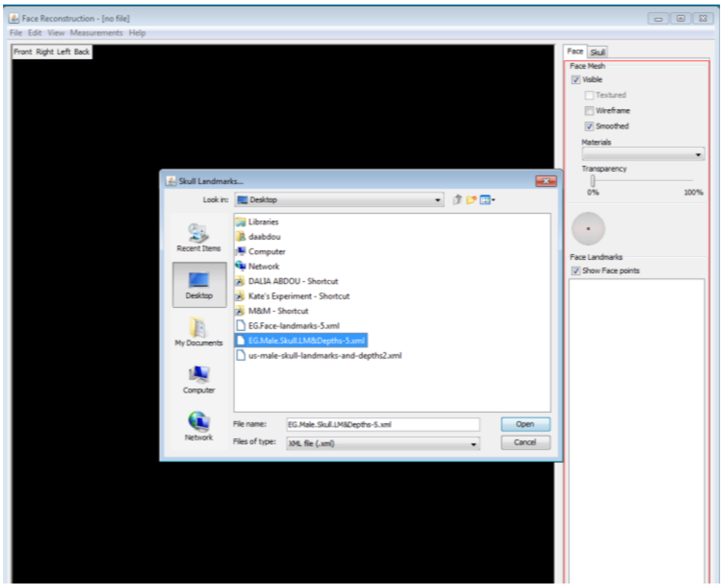


Figure 44: A file browser dialogue for skull landmarks

4) Browse to the location of the skull landmarks file. The file is in .xml format.
5) Click on the bottom panel of the side window (Figure 31) to see the landmarks.
6) The landmarks set can now be seen on the side window (Figure 45).

Figure 45: The main window after the skull landmarks are loaded

3.2 STEP (2): IMPORTING CRANIOFACIAL LANDMARKS DEPTHS

Although landmarks and depths are specified with the same file format (i.e. stored in the same file), at present depths are loaded separately, so that different sets of depths can be loaded into a running session. This will be useful if some depths need to be changed for example.

Internally, when saving sessions, the same format is also used to store the location of landmarks and the direction of skull landmarks. This format is not needed as an input.

To import the craniofacial landmarks depths:

1) Open the file menu → import → depths (Figure 43).
2) A file browser dialog will pop up (Figure 46).

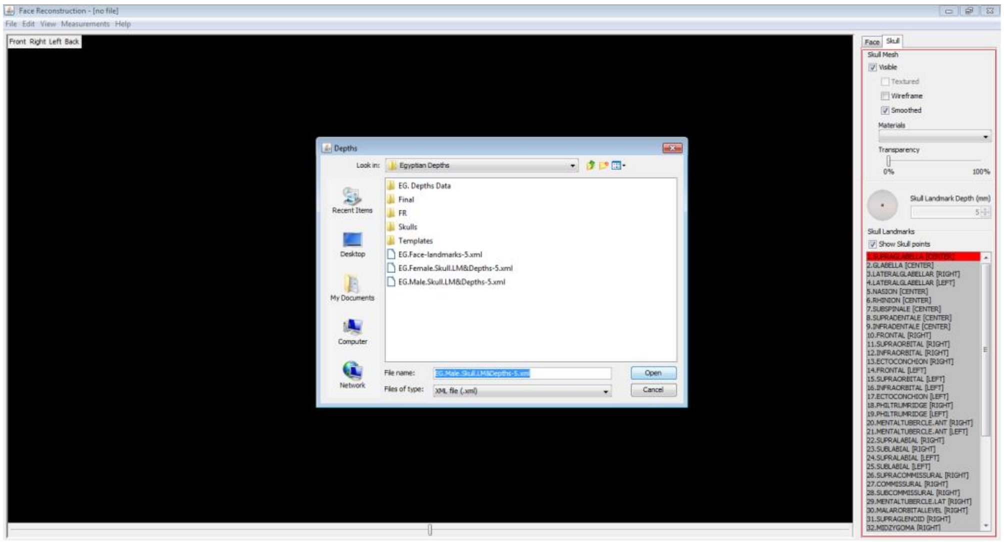


Figure 46: A file browser dialogue for skull landmarks depths

3) Browse to the location of the landmarks depths file. The file is in .xml format. Import the same file that was imported in step (1).
4) There will be no visible change, however after placing each landmark in step (4), the depths in (mm) will appear in the middle panel of the side window (Figure 32).

N.B. 06:

The upwards and downwards arrows in the middle panel of the side window (Figure 32) can be used to adjust the depths. However, as the landmarks' depths are loaded from the pre-prepared .xml file, depths measurements should NOT be changed after importing them unless recommended.

3.3 STEP (3): IMPORTING A SKULL MESH

To import the skull mesh:

1. Make sure that the skull tab at top panel of the side window (on the right) is active (Figure 31).
2. Import the 3-D skull mesh into the software by opening the file menu → import → skull (Figure 43).

3. A file browser dialog will pop up (
4. Figure **47**).

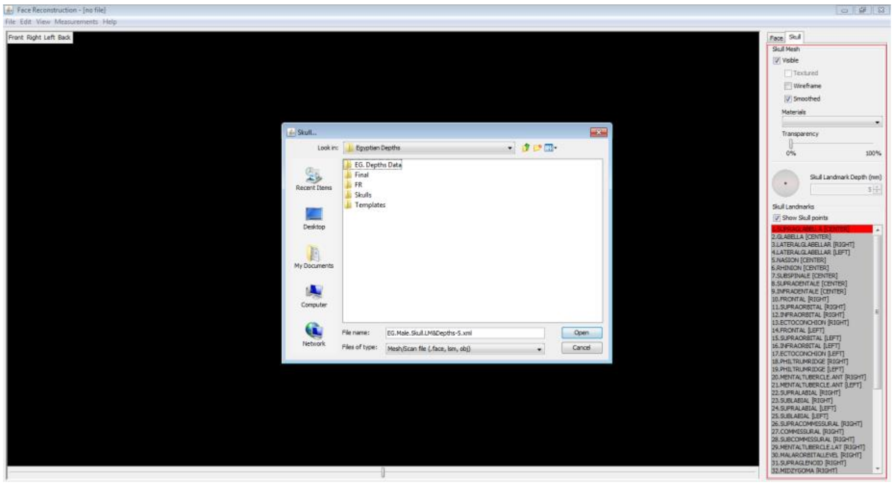


Figure 47: A file browser dialogue for the skull mesh

5. Browse to the location of the skull mesh file. The file is either .lsm or .obj format. The 3-D skull image will then appear in the main window (
6. Figure **48**).

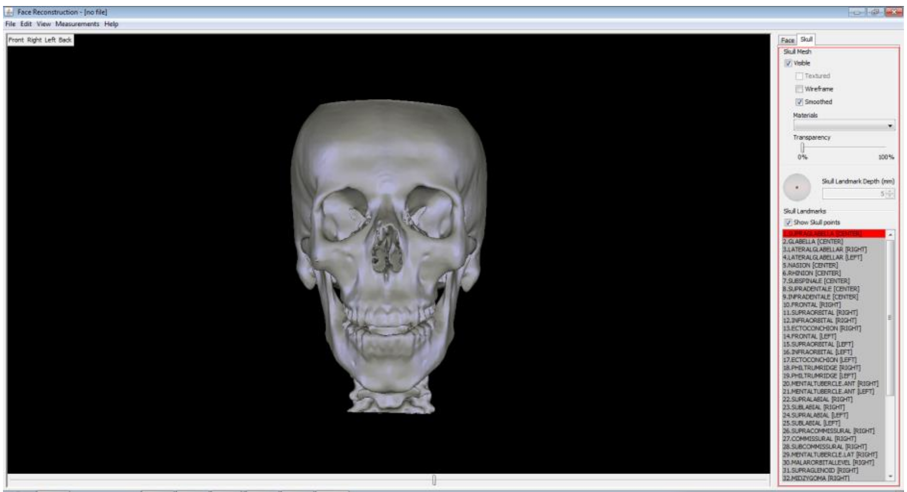


Figure 48: The main window after the skull is loaded

3.4 STEP (4): POSITIONING OF THE SKULL MESH

When a skull is loaded it could, in principle, be in any orientation. This depends on how the scanner capturing the skull works. So, it is centred on the origin, and then can be rotated appropriately.

The skull is ideally positioned in the anatomical Frankfort Horizontal position. Frankfort line is an imaginary line approximating the base of the cranium, passing from the infraorbital ridge (i.e. the lower border of the orbit) to the midline of the occiput (i.e. the back bone of the skull), intersecting the superior margin of the external auditory meatus (i.e. the upper border of the ear canal). The cranium is in the anatomic position when the base line lies in the horizontal plane and right and left sides are level (Figure 49).

The skull can be rotated as required for placing the landmarks, but should be kept in the anatomical Frankfort Horizontal position.

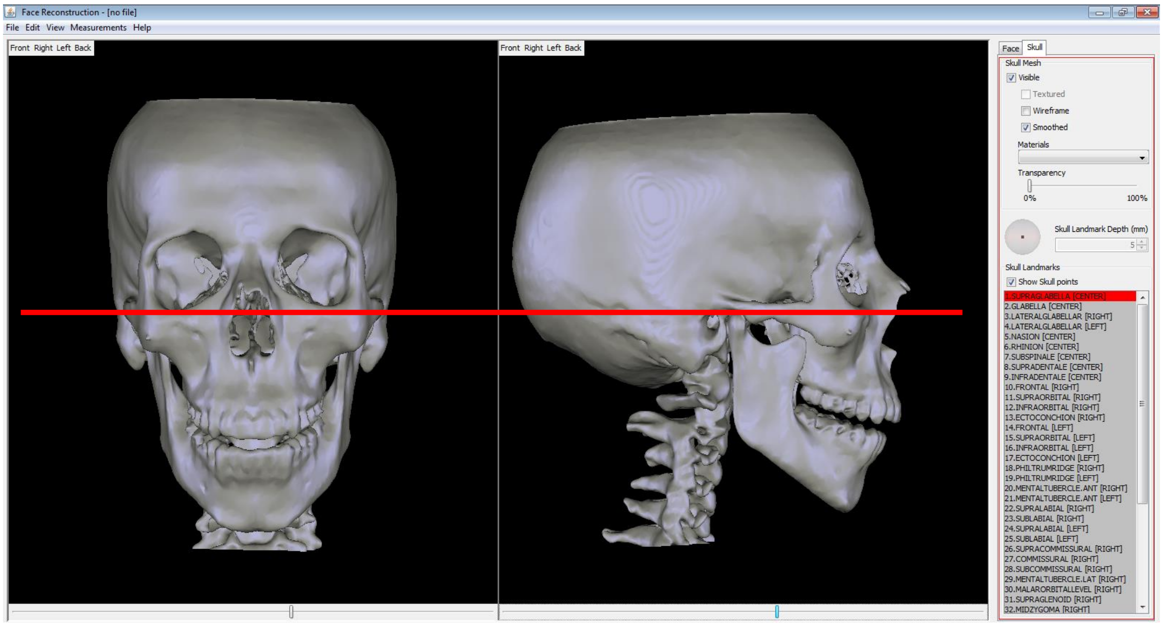


Figure 49: The imported skull positioned in the anatomical Frankfort horizontal plane

N.B. 07:

Mesh positioning:

1. To **Rotate** the mesh: press and hold the RIGHT mouse button and move the mesh in the main window to rotate the view.
2. To **Move** the mesh: press and hold the SHIFT + RIGHT mouse button and move the mesh in the main window.
3. To **Scale** the mesh: either press and hold the MIDDLE mouse button while moving the mouse up to zoom and down to shrink, or move the slider under the 3D panel of the main window (Figure 30).

3.5 STEP (5): PLACING THE LANDMARKS ON THE SKULL MESH

The anatomical landmarks represented by small projecting pegs, are then placed on their anatomical position on the skull using a mouse cursor. Each skull landmark is uniquely numbered and has a name, which describes its anatomical location.

N.B. 08:

For better orientation and more accurate placement of the landmarks, select the split screen view from the view menu (Figure 50) to view the skull from different views (e.g., Front and side) (Figure 49) at the same time while adjusting the same landmark.

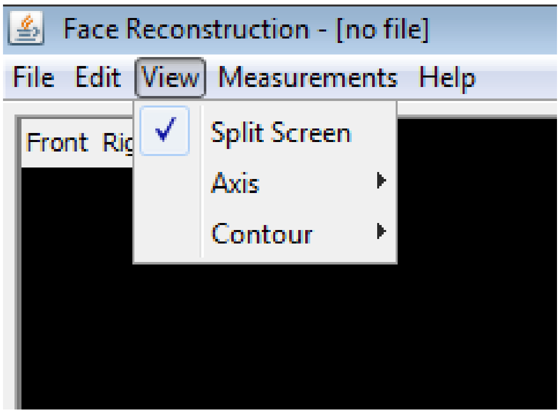


Figure 50: The split view

To add a landmark the mouse must be over a skull mesh at the expected anatomical site of the landmark. A set of landmarks must have been added for the mesh. The rightmost lower panel (the bottom panel of the side window) has the landmark names, and the mesh must be active by activating the relevant skull or the face tab from the uppermost in the right hand pane (the top panel of the side window).

N.B. 09:

To add a landmark:

- Click on the name of the required landmark in the side window bottom panel.
- Hold down the **CTRL** key and the name of the landmark to place will appear in a yellow rectangle where the mouse is located (Figure 33a).
- Click the *LEFT* mouse button to place the named landmark.
- If you want another landmark, release the **CTRL** key and select it in the landmark panel.

N.B. 10:

To move or remove a landmark:

- Click on the name of the required landmark in the side window bottom panel. Its color turns into green.

- Move the mouse over the landmark (until its color returns to blue).
- To **Move** the landmark: press and hold the LEFT mouse button and moving the mouse over the mesh.
- To **Remove** the landmark: click on it with the LEFT button with both **CTRL+SHIFT** keys pressed.

N.B. 11:

- Each landmarks should be oriented as perpendicular/vertical to the skull surface underneath.
- Adjusting the direction of any landmark can be done by rotating it via the landmarks rotating cursor (Figure 32) using the right mouse button.
- Zooming in, with the slider or with the middle mouse button, allows close orientation and adjusting of the landmarks.

N.B. 12:

- Some landmarks may be in line vertically or horizontally.
- So, it might be more practical to first position the mesh in one orientation view (e.g., frontal view), and then place the landmark or group of landmarks and adjust their direction to be perpendicular to the bone surface in that view.
- The mesh can then be rotated to the next orientation view (e.g., 3/4 view), so another landmark or group of landmarks can be placed and adjusted to the bone surface in that view.

To ensure placing the landmarks in the correct anatomical position on the skull, the skull mesh can be moved and rotated as required for better orientation of the location and direction of the landmarks (See Section 3.4, Positioning of the Skull Mesh). Then, each landmarks is rotated so it is vertical to the surface. Also, landmarks can be relocated and redirected to correct any error.

N.B. 13:

To rotate a landmark:

- Click on the name of the required landmark in the side window bottom panel. Its color turns into green.
- Move the mouse to the landmarks rotating cursor (Figure 32).
- Press and hold the RIGHT click mouse button on the cursor and rotate until correct.

3.6 STEP (6): SAVING THE SKULL MESH

After placing the landmarks on the skull mesh has completed, the file is then saved by selecting the "save as" option from the file menu (Figure 51). The file is saved in .face format. Close the working window after saving the work.

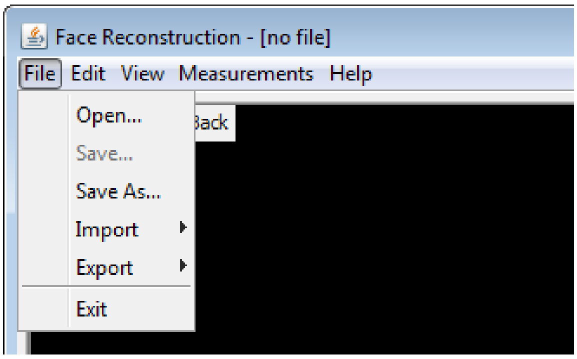


Figure 51: The file menu

N.B. 14:

Saving the work may take a few seconds or a minute, so it is important to wait until it is saved.

At any time, when the work has not completed, it can be saved by the same way. To resume working, start with repeating steps (1) and (2), then importing the saved skull file as in step (3), then repeating steps (4) and (5). The landmarks placement can then be resumed, and the file is saved after completion.

3.7 Step (7): Importing Face Landmarks

To import the face landmarks:

1. Start by opening a new working window.
2. Make sure that the face tab at top panel of the side window (on the right) is active (Figure 31).
3. Import the face landmarks into the software by opening the file menu → import → face landmarks (Figure 43).
4. A file browser dialog will pop up (Figure 52).

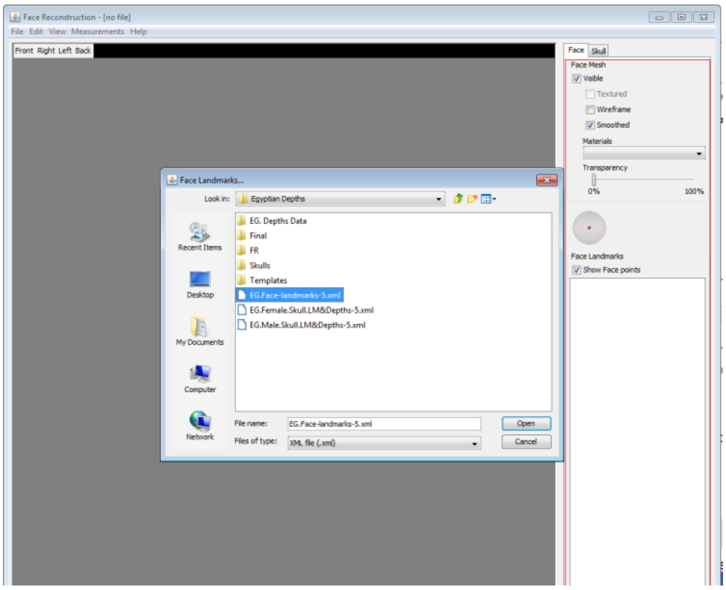


Figure 52: A file browser dialogue for face landmarks

5. Browse to the location of the face landmarks file. The file is in .xml format.
6. Click on the bottom panel of the side window (Figure 31) to activate the landmarks.
7. The landmarks set can now be seen on the side window (Figure 53).

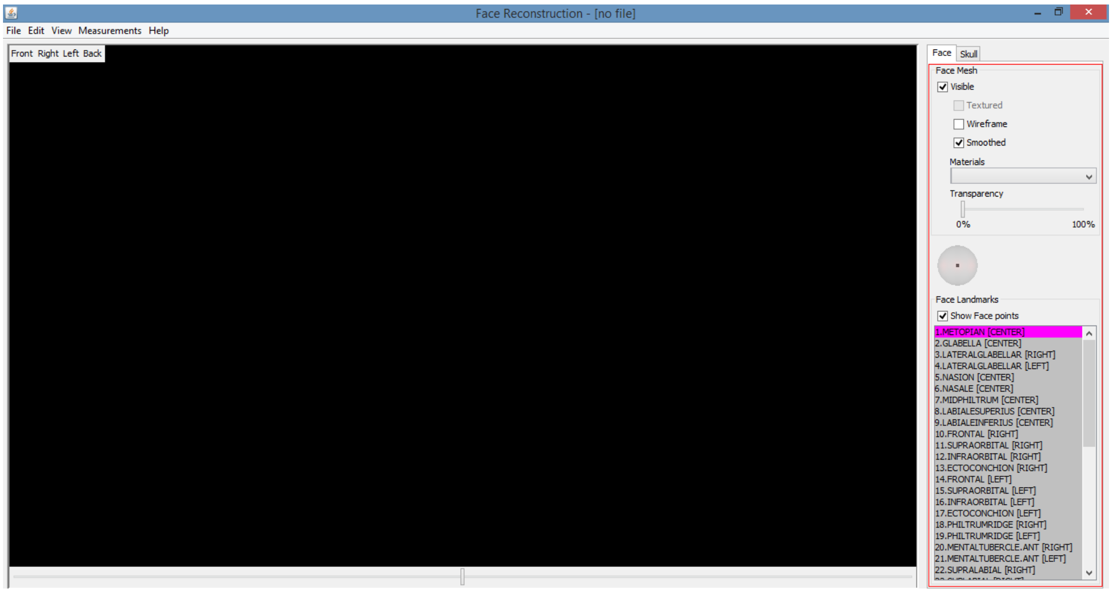


Figure 53: The main window after the face landmarks are loaded

N.B. 15:

No depths measurements for the face landmarks are needed. Only one landmarks depths file is loaded (Step 2).

3.8 STEP (8): IMPORTING A FACE MESH

To import the face mesh:

1- Make sure that the face tab at top panel of the side window (on the right) (Figure 31) is active.
2- Import the 3-D face mesh into the software by opening the file menu → import → face (Figure 43).
3- A file browser dialog will pop up (Figure 54).

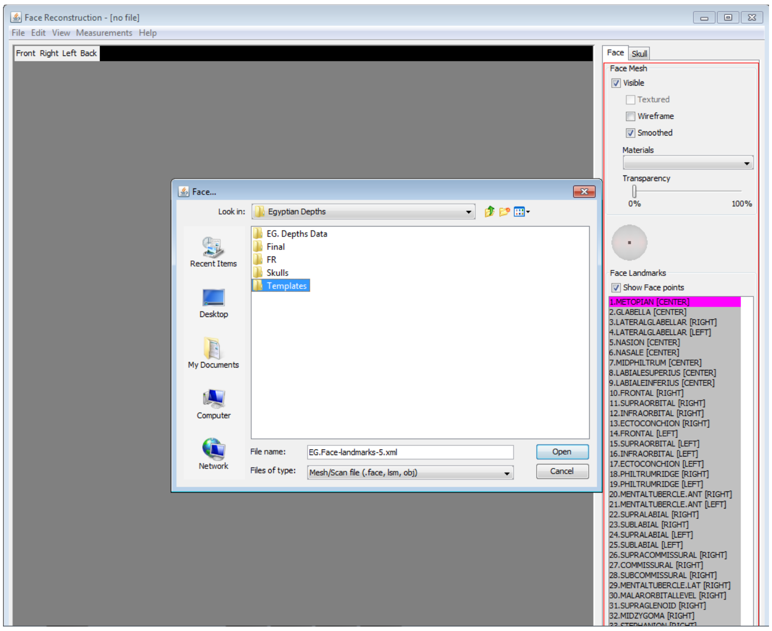


Figure 54: A file browser dialogue for the skull mesh

4- Browse to the location of the face mesh file. The file is either .lsm or .obj format. The 3-D face image will then appear in the main window (Figure 55).

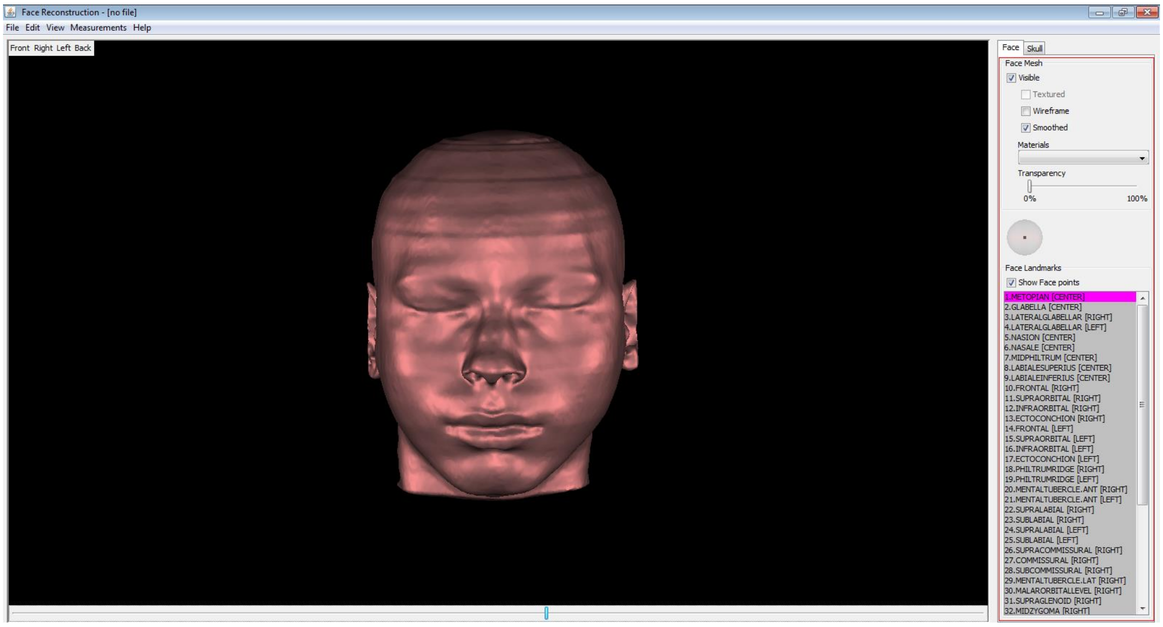


Figure 55: The main window after the skull is loaded

3.9 STEP (9): POSITIONING OF THE FACE MESH

When a face is loaded it could, in principle, be in any orientation. This depends on how the scanner capturing the face works. So, it is centred on the origin, and then can be rotated appropriately.

For better orientation and more accurate placement of the landmarks, select the split screen view from the view menu (Figure 50). The split view allows the user to view the face from different views (e.g., Front and side) at the same time (Figure 56).

Face positioning: Please refer to **N.B. 07**, Section 3.4.

The face is ideally positioned in the anatomical Frankfort Horizontal position (Figure 29). Frankfort line is an imaginary line approximating the base of the cranium, passing from the infraorbital ridge (i.e. the lower border of the orbit) to the midline of the occiput (i.e. the back bone of the skull), intersecting the superior margin of the external auditory meatus (i.e. the upper border of the ear canal). The head is in the anatomic position when the base line lies in the horizontal plane and right and left sides are level. The face can be rotated as required for placing the landmarks, but should be kept in the anatomical Frankfort Horizontal position.

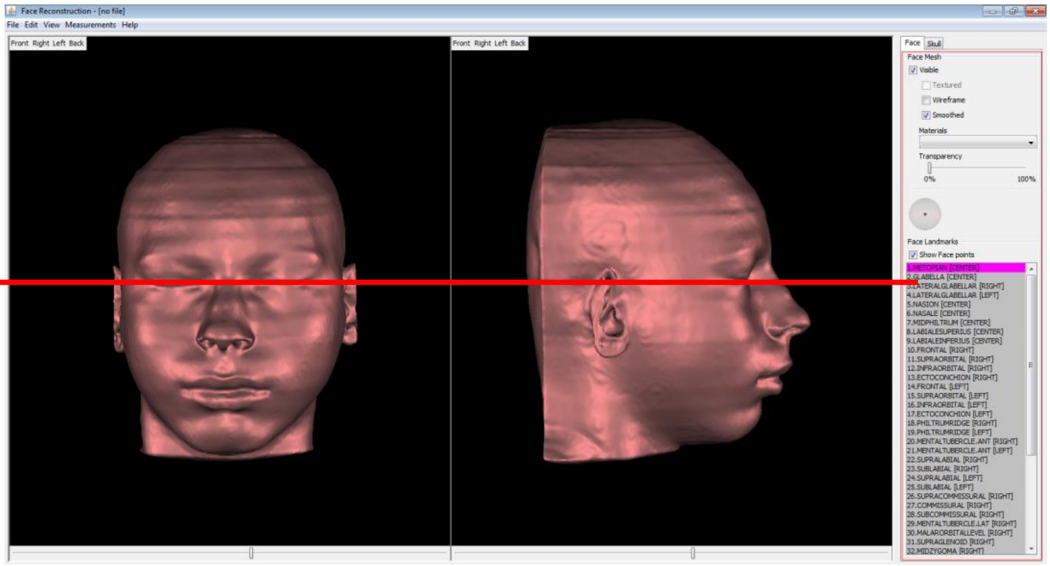


Figure 56: The imported skull positioned in the anatomical Frankfort horizontal plane

3.10 STEP (10): PLACING THE LANDMARKS ON THE FACE MESH

To add a landmark the mouse must be over a face mesh at the expected anatomical site of the landmark. A set of landmarks must have been added for the mesh. The rightmost lower panel (the bottom panel of the side window) has the landmark names, and the mesh must be active by activating the relevant face tab from the uppermost in the right hand pane (the top panel of the side window).

To add a landmark: Please refer to **N.B. 09**, Section 3.5.

To move or remove a landmark: Please refer to **N.B. 10**, Section 3.5.

3.11 STEP (11): SAVING THE FACE MESH

After placing the landmarks on the face mesh has completed, the file is then saved by selecting the “save as” option from the file menu (Figure 51). The file is saved in .face format. Close the work window after saving the work. Saving the work may take a few seconds or a minute, so it is important to wait until it is saved.

At any time, when the work has not completed, it can be saved by the same way. To resume working, start with repeating step (7), then importing the saved face file as in step (8), then repeating steps (9) and (10). The landmarks placement can then be resumed, and the file is saved after completion.

It is possible to save a session at any time and then load it back in again in exactly the same state.

3.12 Step (12): Completing the Warping Process and Saving the Work

To complete the facial reconstruction:

- Start by opening a new work window.
- Import the saved skull file.
- Import the saved face file.
- Select the "fit" function from the edit menu (Figure 35), so the two meshes are then automatically warped or fitted.
- To view the reconstructed face, make sure that the face tab at top panel of the side window (on the right) is active (Figure 31).
- To reverse the fitting process, select the "clear" function from the edit menu (Figure 35). This can be done to adjust any landmarks on the skull or the face meshes, then use the "fit" function again, and so on.
- Once happy with the reconstructed face, save the file by selecting the "save as" option from the file menu (Figure 51). The file is saved in .face format. Saving the work may take a few seconds or a minute, so it is important to wait until it is saved.

N.B. 16:

If changes are required, it is important to make them in the saved skull and/or the face mesh files separately. This can be done by importing (using the import function) the required file and saving it again for future reference.

N.B. 17:

To return to a saved file (face, skull or facial reconstruction), open a working window and select the open function from the file menu and select the required file.

N.B. 18:

The final appearance of the reconstructed serves as a better judge of the proper positioning of the landmarks. If the face is much distorted with this warp, for example, it usually indicates poor placement of landmarks or swapped landmarks. It always possible to move the landmarks in the way described above. Once the user is happy with the final appearance of the reconstruction, save the file as described.

N.B. 19:

A session, which when saved as a file has the extension ".face". It is a compressed file that contains all the information needed about a reconstruction, including the exact location of the landmarks and the state of the reconstruction.

N.B. 20:

To export the reconstructed face in .obj format for further analysis:

1. Copy the .face file.
2. Change the file extension to .zip.
3. Extract the file (face.obj) from the compressed file, and use it as required.

N.B. 21:

The skull, the face template or the output of a warp can also be exported in 2D as a .tiff (Tagged Image File Format) file for further use.

4 A CHART SUMMARY

Step (1): Import the skull landmarks (pre-prepared .xml file)

Section 3.1

Step (2): Import the skull landmarks depths (pre-prepared .xml file)

Section 3.2

Step (3): Import the skull mesh (activate the skull tab)

Section 3.3

Step (4): Position the skull mesh (Frankfurt position, split view)

Section 3.4

Step (5): Place the landmarks on the skull mesh

Section 3.5

Step (6): Save the skull mesh & close the work window

Section 3.6

Step (7): Import the face landmarks in a new work window

Section 3.7

Step (8): Import the face mesh (activate the face tab)

Section 3.8

Step (9): Position the face mesh (Frankfurt position, split View)

Section 3.9

Step (10): Place the landmarks on the face mesh

Section 3.10

Step (11): Save the face mesh & close the work window

Section 3.11

Step (12): Open a new window and fit the two meshes, save the FR & close the work window

Section 3.12

5 HINTS

N.B. 01: **Section 3**

The default is that the mesh is smoothed.

N.B. 02: **Section 3**

Hiding the face or the skull images is possible while working on the other by deactivating the "visible" button.

N.B. 03: **Section 3**

The alpha-blending (mixed view) allows the operator to see where the skull and skull landmarks are in relation to the reconstructed face. It can be seen by adjusting the transparency of the meshes from the transparency slider in the top panel of the side window (Figure 42).

N.B. 04: **Section 3**

The following steps should be followed for the "Face FR" software to be used in Forensic Facial Reconstruction of any population. However, the files containing Skull and Face Landmarks and Depths are population specific and should be pre-pared prior to loading into the software.

N.B. 05: **Section 3**

The files containing the required information of skull landmarks' names, depths, and face landmarks as well as the skull and the face meshes are loaded into the software separately from the file menu (Figure 43) in successive steps.

N.B. 06: **Section 3.2**

The upwards and downwards arrows in the middle panel of the side window (Figure 32) can be used to adjust the depths. However, as the landmarks' depths are loaded from the pre-prepared .xml file, depths measurements should NOT be changed after importing them unless recommended.

N.B. 07: **Section 3.4**

Mesh positioning:

- To **Rotate** the mesh: press and hold the RIGHT mouse button and move the mesh in the main window to rotate the view.
- To **Move** the mesh: press and hold the SHIFT + RIGHT mouse button and move the mesh in the main window.
- To **Scale** the mesh: either press and hold the MIDDLE mouse button while moving the mouse up to zoom and down to shrink, or move the slider under the 3D panel of the main window (Figure 30).

N.B. 08:

For better orientation and more accurate placement of the landmarks, select the split screen view from the view menu (Figure 50) to view the skull from different views (e.g., Front and side) (Figure 49) at the same time while adjusting the same landmark.

N.B. 09: **Section 3.5**

To add a landmark:

- Click on the name of the required landmark in the side window bottom panel.
- Hold down the CTRL key and the name of the landmark to place will appear in a yellow rectangle where the mouse is located (Figure 33a).
- Click the LEFT mouse button to place the named landmark.
- If you want another landmark, release the CTRL key and select it in the landmark panel.

N.B. 10: **Section 3.5**

To move or remove a landmark:

- Click on the name of the required landmark in the side window bottom panel. Its color turns into green.
- Move the mouse over the landmark (until its color returns to blue).

- To **Move** the landmark: press and hold the LEFT mouse button and moving the mouse over the mesh.
- To **Remove** the landmark: click on it with the LEFT button with both **CTRL**+**SHIFT** keys pressed.

N.B. 11: **Section 3.5**

- Each landmarks should be oriented as perpendicular/vertical to the skull surface underneath.
- Adjusting the direction of any landmark can be done by rotating it via the landmarks rotating cursor (Figure 32) using the right mouse button.
- Zooming in, with the slider or with the middle mouse button, allows close orientation and adjusting of the landmarks.

N.B. 12: **Section 3.5**

- Some landmarks may be in line vertically or horizontally.
- So, it might be more practical to first position the mesh in one orientation view (e.g., frontal view), and then place the landmark or group of landmarks and adjust their direction to be perpendicular to the bone surface in that view.
- The mesh can then be rotated to the next orientation view (e.g., 3/4 view), so another landmark or group of landmarks can be placed and adjusted to the bone surface in that view.
- Some landmarks may be in line vertically or horizontally.
- So, it might be more practical to first position the mesh in one orientation view (e.g., frontal view), and then place the landmark or group of landmarks and adjust their direction to be perpendicular to the bone surface in that view.
- The mesh can then be rotated to the next orientation view (e.g., 3/4 view), so another landmark or group of landmarks can be placed and adjusted to the bone surface in that view.

- Adjusting the direction of any landmark can be done by rotating it via the landmarks rotating cursor (Figure 32) using the right mouse button.

N.B. 13: **Section 3.5**

To rotate a landmark:

- Click on the name of the required landmark in the side window bottom panel. Its color turns into green.
- Move the mouse to the landmarks rotating cursor (Figure 32).
- Press and hold the RIGHT click mouse button on the cursor and rotate until correct.

N.B. 14: **Section 3.6**

Saving the work may take a few seconds or a minute, so it is important to wait until it is saved.

N.B. 15: **Section 3.7**

No depths measurements for the face landmarks are needed. Only one landmarks depths file is loaded (Step 2).

N.B. 16: **Section 3.12**

If changes are required, it is important to make them in the saved skull and/or the face mesh files separately. This can be done by importing (using the import function) the required file and saving it again for future reference.

N.B. 17: **Section 3.12**

To return to a saved file (face, skull or facial reconstruction), open a working window and select the open function from the file menu and select the required file.

N.B. 18: **Section 3.12**

The final appearance of the reconstructed serves as a better judge of the proper positioning of the landmarks. If the face is much distorted with this warp, for example, it usually indicates poor placement of landmarks or swapped landmarks. It always possible to move the landmarks in the way described above. Once the user is happy with the final appearance of the reconstruction, save the file as described.

N.B. 19: **Section 3.12**

A session, which when saved as a file has the extension ".face". It is a compressed file that contains all the information needed about a reconstruction, including the exact location of the landmarks and the state of the reconstruction.

N.B. 20: **Section 3.12**

To retrieve the reconstructed face in .obj format for further analysis:

- Copy the .face file.
- Change the file extension to .zip.
- Extract the file (face.obj) from the compressed file, and use it as required.

N.B. 21: **Section 3.12**

The skull, the face template or the output of a warp can also be exported in 2D as a .tiff (Tagged Image File Format) file for further use.

6 THREE-DIMENSIONAL FORENSIC FACIAL RECONSTRUCTION OF ADULT EGYPTIAN POPULATION USING “FR” SOFTWARE

6.1 STEP (1): IMPORTING EGYPTIAN SKULL LANDMARKS

6.2 STEP (2): IMPORTING EGYPTIAN CRANIOFACIAL LANDMARKS DEPTHS

Separate .xml files containing the cranial landmarks for adult Egyptian male and female populations have been prepared to be directly imported in the software as described in Sections 4.1 and 4.2.

6.3 STEP (3): IMPORTING A SKULL MESH

Import the studied skull mesh as described in section 3.3.

6.4 STEP (4): POSITIONING OF THE SKULL MESH

Position the studied skull mesh as described in section 3.4.

6.5 STEP (5): PLACING THE LANDMARKS ON THE SKULL MESH

The process of landmarks placement on the skull mesh in general has been described in details in section 3.5. However, the best working way to place the landmarks, the following hints are helpful.

6.5.1 **Landmarks Orientation:**

- **Points** 1, 2, 5, 6, 7, 8 & 9 are on the same vertical level (midline), in the frontal view.
- **Points** 2, 3, 4, 11 & 15 are on the same horizontal level (supraorbital line), in the frontal view.
- **Points** 10 & 14 are on the same horizontal level, in the frontal view.
- **Points** 10, 11 &12 are on the same vertical level, in the frontal view.
- **Points** 14, 15 &16 are on the same vertical level, in the frontal view.

- **Points** 7, 18, 19, 22, 24, 26 & 37 are on the same horizontal level.
- **Points** 23, 25, 28 & 39 are on the same horizontal level, in the frontal view.
- **Points** 20, 21, 29 & 40 are on the same horizontal level, in the frontal view.
- **Points** 26, 27, 28 & 29 are on the same vertical level, in the right 3/4 view.
- **Points** 37, 38, 39 & 40 are on the same vertical level, in the left 3/4 view.
- **Points** 13 & 30 are on the same vertical level, midway between the frontal and the right 3/4 views.
- **Points** 17 & 41 are on the same vertical level, midway between the frontal and the left 3/4 views.
- **Points** 30, 31 & 32 are on the same (slightly tilted) line, in the right lateral/profile view.
- **Points** 32, 33 & 34 are on the same vertical level, in the right lateral/profile view.
- **Points** 42, 43 & 44 are on the same horizontal level, in the left lateral/profile view.
- **Points** 43, 44 & 45 are on the same vertical level, in the left lateral/profile view.

6.5.2 **Landmarks direction:**

- **Points** 1-12, 14-16, 18-21 are vertical to the bone surface in frontal view.
- **Points** 22-25 are vertical to the bone surface in the view midway between the frontal view and the 3/4 view.
- **Points** 13, 17, 26-30, 37-41 are vertical to the bone surface in the 3/4 view.
- **Points** 31-36, 42-47 are vertical to the bone surface in the lateral/profile view.

6.5.3 **Landmarks Placing:**

For more practicality, a working model of landmark placing steps is described below.

Stage (1):

1- Place the skull in the frontal view.
2- Place points 1-25 (Table 54) and (

3-
4- Figure **57** and
5- Figure **58**).

Table 54: Description of the Skull Landmarks (1-25) for Adult Egyptian Population

LM No.	Cranial LM NAME	Cranial LM Definition
1 Midline	Supraglabella	The most anterior point in the midline, above the glabella (LM 2), **midway** between the frontal eminences (Fig. 31).
2 Midline	Glabella	The cross-point between midline and supraorbital line, a horizontal line at the upper border of the orbit (Fig. 31).
3 (Rt) 4 (Lt)	LateralGlabellar	A point at the junction between the inner border of the orbit and the supraorbital line, at the junction of the frontal, maxillary, and lacrimal bones.
5 Midline	Nasion	A point at the top of the nasal bone, in the midline of the Naso-frontal suture (Fig. 31), at the horizontal level of a line dividing the orbit into upper and lower halves.
6 Midline	Rhinion	The end of the nasal bone at the junction between bone and cartilage of the nose.
7 Midline	Subspinale	A point at the midline of the intranasal depression, below the nasal spine, **midway** between the nasal spine (Fig. 31) and the Supradentale (LM 8).
8 Midline	Supradentale	The jaw Centre, in the midline, between the upper incisive teeth (Fig. 31).
9 Midline	Infradentale	The jaw Centre, in the midline, between the lower incisive teeth (Fig. 31).
10 (Rt) 14 (Lt)	Frontal	A point on the forehead **midway** between the frontal eminence and the maximum curve of the supraorbital margin. It lies at the vertical level of a line passing through the centre point of the upper border of the orbit, and on the same

		horizontal level of the deepest point in the depression below the frontal eminence.
11 (Rt) **15 (Lt)**	SupraOrbital	The centre point of the upper orbital margin. On the horizontal level of the Glabella and Lateral Glabella.
12 (Rt) **16 (Lt)**	Infraorbital	The centre point of the lower orbital margin. On the horizontal level of the Glabella.
13 (Rt) **17 (Lt)**	Ectoconchion	Bony projection of the Ectocranial surface of the frontal bone, vertically centred on the orbit, next to the lateral orbital border.
18 (Rt) **19 (Lt)**	Philtrumridge	The prominence on the lateral ridge of the philtrum **midway** between the base of the nostril and the upper border of the jaw. On the horizontal level with the Subspinale.
20 (Rt) **21 (Lt)**	Mentaltubercle.ant	The most prominent point on the lateral bulge, and just above the lower border of the chin bone. On the vertical level with Philtrum Ridge.
22 (Rt) **24 (Lt)**	Supralabial	Over the maximum bulge of the canine eminence, on the horizontal level of the Philtrum Ridge, **midway** between the root of the nasal cartilage and the upper border of the jaw.
23 (Rt) **25 (Lt)**	Sublabial	A point on the depression below the teeth and above the chin prominence, on the vertical level of Supralabial.

6- Adjust points 1-21 to be vertical to the bone surface.

LM No.	LM NAME
1 (Midline)	Supraglabella
2 (Midline)	Glabella
3 (Rt), 4 (Lt)	LateralGlabellar
5 (Midline)	Nasion
6 (Midline)	Rhinion
7 (Midline)	Subspinale
8 (Midline)	Supradentale
9 (Midline)	Infradentale
10 (Rt), 14 (Lt)	Frontal
11 (Rt), 15 (Lt)	SupraOrbital
12 (Rt), 16 (Lt)	Infraorbital
13 (Rt), 17 (Lt)	Ectoconchion
18 (Rt), 19 (Lt)	Philtrumridge
20 (Rt), 21 (Lt)	Mentaltubercle.ant
22 (Rt), 24 (Lt)	Supralabial
23 (Rt), 25 (Lt)	Sublabial
26(Rt), 37 (Lt)	Supracommissural
27(Rt), 38 (Lt)	Commissural
28(Rt), 39 (Lt)	Subcommissural
29(Rt), 40 (Lt)	Mentaltubercle.lat
30(Rt), 41 (Lt)	MalarOrbitalLevel
31(Rt), 42 (Lt)	Supraglenoid
32(Rt), 43 (Lt)	Midzygoma
33(Rt), 44 (Lt)	Stephanion
34(Rt), 45 (Lt)	Midmasseteric
35(Rt), 46 (Lt)	Postero-masseteric
36(Rt), 47 (Lt)	Antero-masseteric

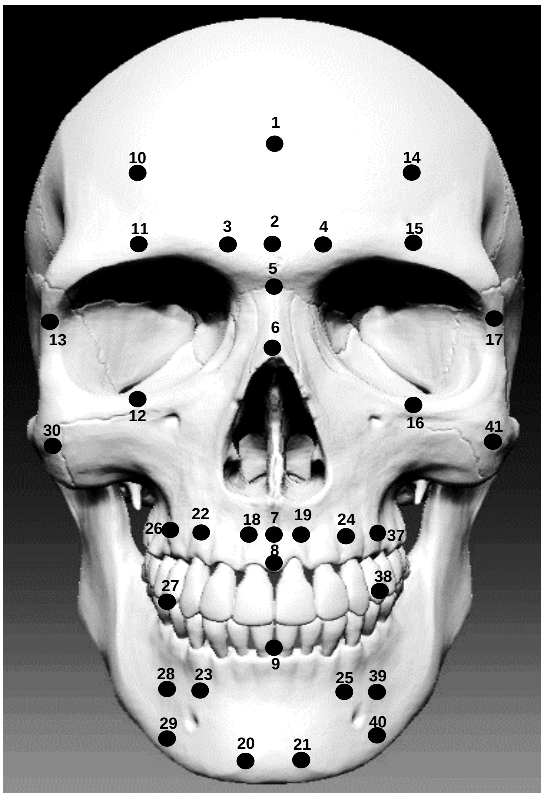


Figure 57: A skull diagram in frontal view showing the positions of the skull landmarks

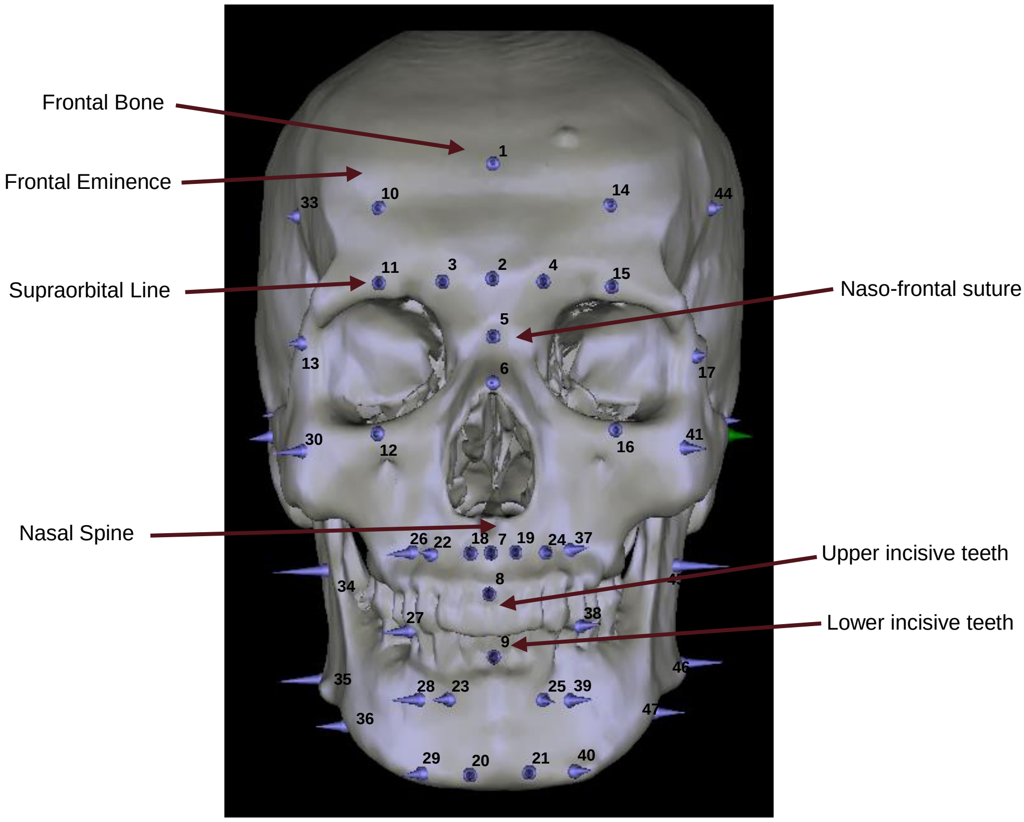


Figure 58: The skull mesh in the frontal view showing the positions and the directions of the skull landmarks

7- Turn the skull mesh to the right side of the skull to a view midway between the frontal view and the right 3/4 view.

8- Adjust the directions of landmarks 22 (Right Supralabial) & 23 (Right Sublabial) to be vertical to the bone surface (

9- Figure **59**).

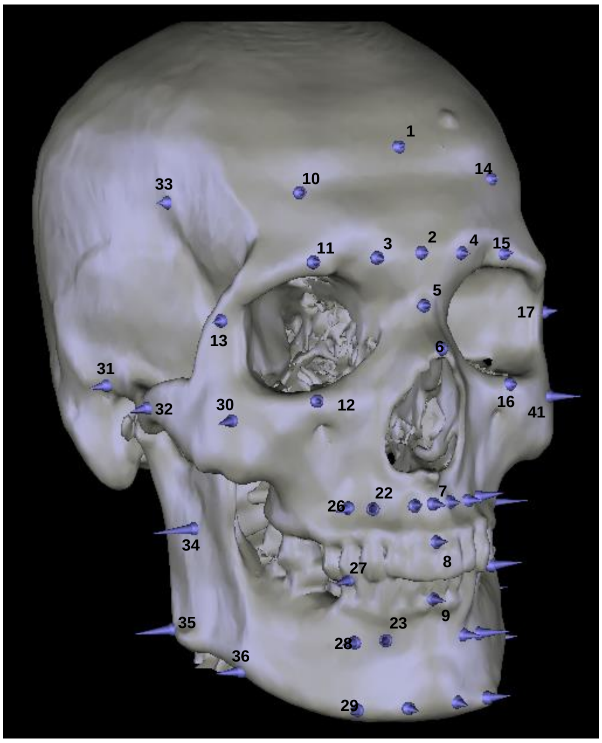


Figure 59: The skull mesh midway between the frontal and the right 3/4 views showing the positions and the directions of the skull landmarks

Stage (2):

10- Turn the skull mesh more to the right side of the skull to the right 3/4 view.

11- Adjust the direction of landmark 13 (Right Ectoconchion) to be vertical to the bone surface (

12- Figure **60**).

13- Place points 26-30 (Table 55).

Table 55: Description of the Skull Landmarks (26-30) for Adult Egyptian Population

LM No. (Rt)	Cranial LM NAME	Cranial LM Definition
26	Supracommissural	A point over the root of the first premolar (Figure 33). On the horizontal level of Supralabial.
27	Commissural	A point on the crown of the first premolar (Figure 33). On the vertical level of Supracommissural.
28	Subcommissural	A point lateral to the Sub-labial. On the horizontal level of Sublabial. On the vertical level of Supracommissural.
29	Mentaltubercle.lat	A point posterior to and on the horizontal level of the Mental tubercle Anterior. On the vertical level of the Supracommissural.
30	MalarOrbitalLevel	Lined up with the lateral border of the eye on the centre of the Zygomatic process (Figure 33). In vertical line with the Ectoconchion.

14- Adjust the directions of landmarks 26-30 to be vertical to the bone surface (
15- Figure **60**).

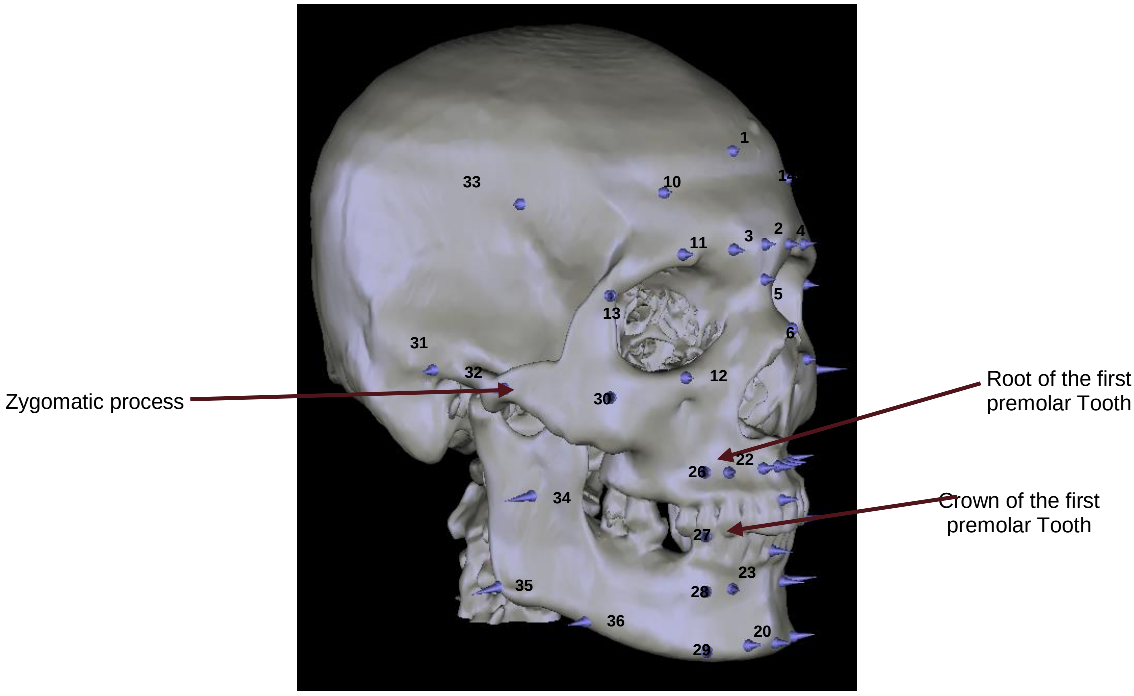


Figure 60: The Skull Mesh in the Right 3/4 View showing the positions and the directions of the skull landmarks

Stage (III):

16- Turn the skull mesh more to the right side of the skull to the right lateral/profile view.

17- Place points 31-36 (Table 56).

Table 56: Description of the Skull Landmarks (31-47) for Adult Egyptian Population

LM No. (Rt)	Cranial LM NAME	Cranial LM Definition
31	Supraglenoid	Root of Zygomatic arch (Figure 35), immediately above the mandibular condyle (Figure 35) and superficial to the posterior root of the Zygoma. Just at the upper border of the ear canal (Figure 36).
32	Midzygoma	A point overlying the maximum horizontal and vertical outer curvature of the Zygomatic arch.
33	Stephanion	The point on the side of the skull on the horizontal level of the frontal (LM 10, 14), and on the vertical level of the Mid-Zygoma and the Mid-Masseteric (LM 34, 45).
34	Midmasseteric	Middle of the masseter. A point at the centre of the outer surface of the mandibular ramus (Figure 35) **midway** between the zygomatic arch and the inferior border of the mandible. On the vertical level of the Sephanion and the Mid-Zygoma.
35	Postero-masseteric	This point lies at the lower and posterior edge of the mandible (Figure 35), just anterior to the angle of the mandible, and posterior to the Antero-Masseteric (LM 36, 47).
36	Antero-masseteric	This point lies at the inferior border of the mandible, halfway between the Postero-Masseteric and the lateral mental tubercle.

18- Adjust the directions of landmarks 31-36 to be vertical to the bone surface (
19- Figure **61** and Figure 62).

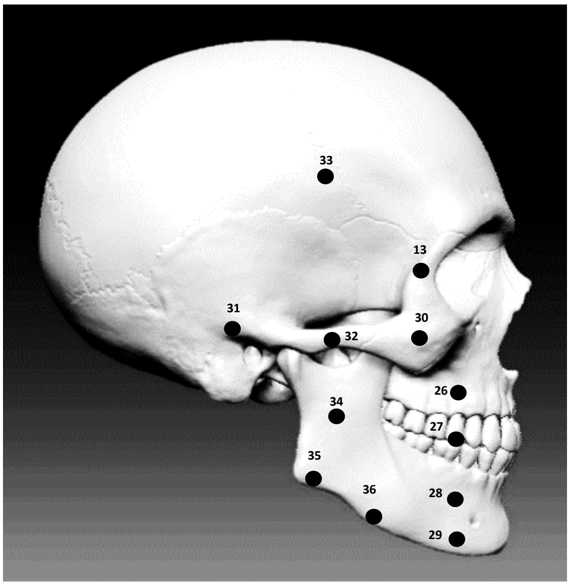


Figure 61: A skull diagram in frontal view showing the positions of the skull landmarks

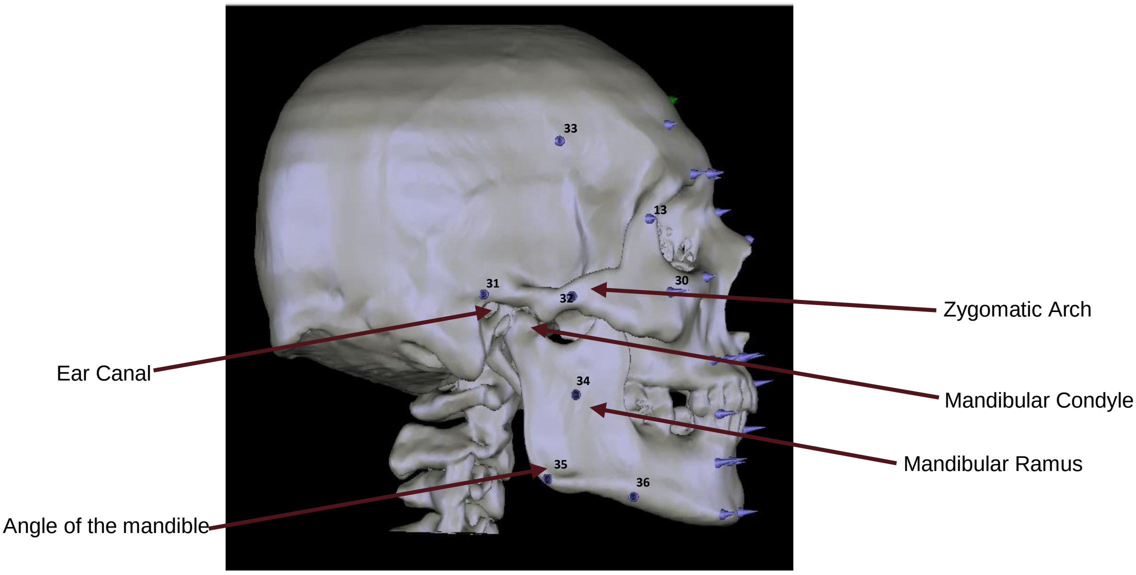


Figure 62: The skull mesh in the right lateral/profile view showing the positions and directions of the skull landmarks

Stage (IV):

20- Turn the skull mesh to the left side of the skull to a view midway between the frontal view and the left 3/4 view.

21- Adjust the directions of landmarks 24 (Left Supralabial) & 25 (Left Sublabial) to be vertical to the bone surface (Figure 63).

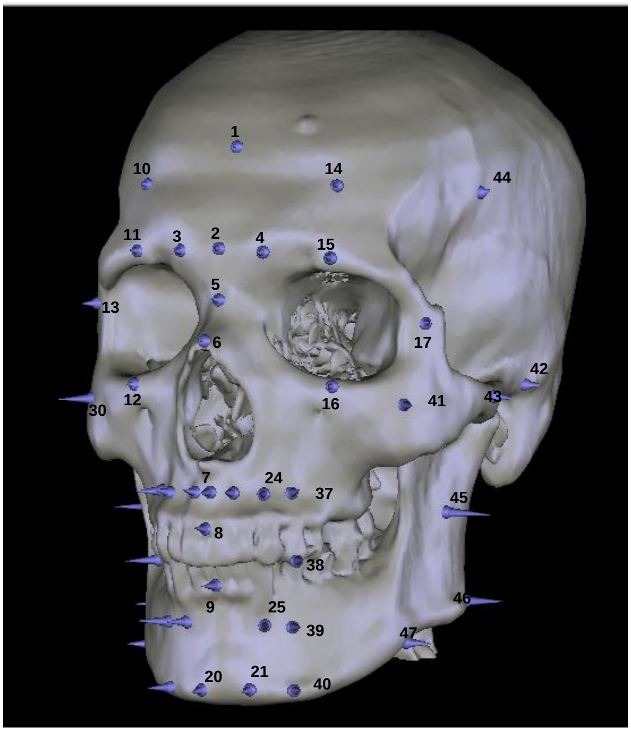


Figure 63: The skull mesh midway between the frontal and the left 3/4 view for showing the positions and the directions of the skull landmarks

Stage (V):

22- Turn the skull mesh more to the left side of the skull to the left 3/4 view.

23- Adjust the direction of landmark 17 to be vertical to the bone surface (

24- Figure **64**).

25- Place points 37-41 (Table 57).

Table 57: Description of the Skull Landmarks (37-41) for Adult Egyptian Population

LM No. **(Lt)**	**Cranial LM NAME**	**Cranial LM Definition**
37	Supracommissural	A point over the root of the first premolar (Fig. 33). On the horizontal level of Supralabial.
38	Commissural	A point on the crown of the first premolar (Fig. 33). On the vertical level of Supracommissural.
39	Subcommissural	A point lateral to the Sub-labial. On the horizontal level of Sublabial. On the vertical level of Supracommissural.
40	Mentaltubercle.lat	A point posterior to and on the horizontal level of the Mental tubercle Anterior. On the vertical level of the Supracommissural.
41	MalarOrbitalLevel	Lined up with the lateral border of the eye on the centre of the Zygomatic process (Fig. 33). In vertical line with the Ectoconchion.

Adjust the directions of landmarks 37-41 to be vertical to the bone surface (
26- **Figure 64**).

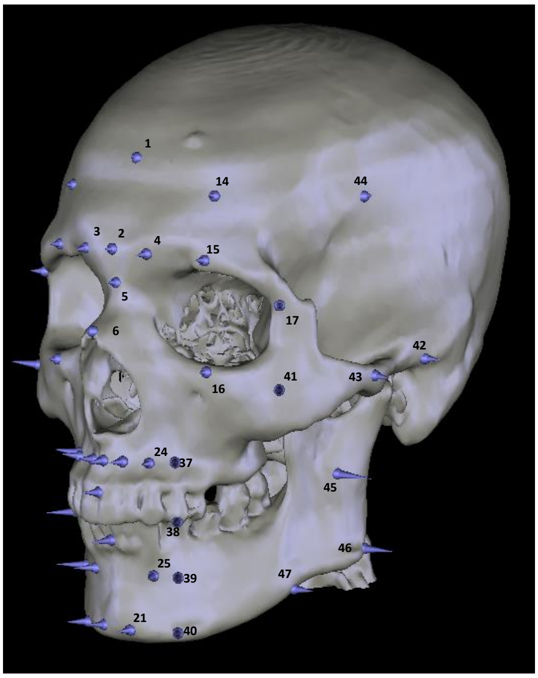


Figure 64: The skull mesh in the left 3/4 view showing the positions and directions of the skull landmarks

Stage (VI):

27- Turn the skull mesh more to the left side of the skull to the left lateral/profile view.
28- Place points 42-47 (Table 58).

Table 58: Description of the Skull Landmarks (42-47) for Adult Egyptian Population

LM No. (Lt)	**Cranial LM NAME**	**Cranial LM Definition**
42	Supraglenoid	Root of Zygomatic arch (Fig. 35), immediately above the mandibular condyle (Fig. 35) and superficial to the posterior root of the Zygoma. Just at the upper border of the ear canal (Fig. 35).
43	Midzygoma	A point overlying the maximum horizontal and vertical outer curvature of the Zygomatic arch.
44	Stephanion	The point on the side of the skull on the horizontal level of the frontal (LM 10, 14), and on the vertical level of the Mid-Zygoma and the Mid-Masseteric (LM 34, 45).
45	Midmasseteric	Middle of the masseter. A point at the centre of the outer surface of the mandibular ramus (Fig. 35) **midway** between the zygomatic arch and the inferior border of the mandible. On the vertical level of the Sephanion and the Mid-Zygoma.
46	Postero-masseteric	This point lies at the lower and posterior edge of the mandible (Figure 36), just anterior to the angle of the mandible, and posterior to the Antero-Masseteric (LM 36, 47).
47	Antero-masseteric	This point lies at the inferior border of the mandible, halfway between the Postero-Masseteric and the lateral mental tubercle.

29- Adjust the directions of landmarks 42-47 to be vertical to the bone surface (Figure 65 and

30- Figure **66**).

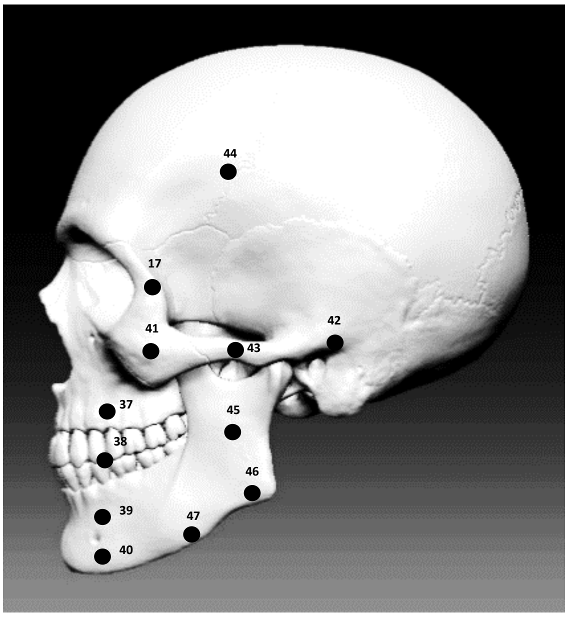


Figure 65: A skull diagram in the left lateral/profile view showing the positions of the skull landmarks

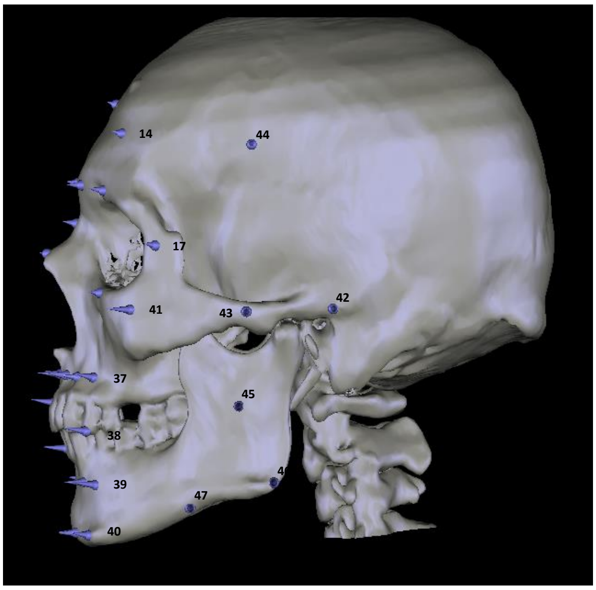


Figure 66: The skull mesh in the left lateral/profile view showing the positions and directions of the skull landmarks

6.6 STEP (6): SAVING THE SKULL MESH

As described in section 3.6.

6.7 STEP (7): IMPORTING FACE LANDMARKS

An .xml file containing the face landmarks for the adult Egyptian population has been prepared to be directly imported in the software as described in Section 3.7.

6.8 STEP (8): IMPORTING A FACE MESH

Import the studied face mesh as described in Section 3.8.

6.9 STEP (9): POSITIONING OF THE FACE MESH

Position the studied face mesh as described in Section 3.9.

6.10 STEP (10): PLACING THE LANDMARKS ON THE FACE MESH

The process of landmarks placement on the face mesh in general has been described in details in section 3.10. However, the best working way to place the landmarks, the following hints are helpful.

6.10.1 **Landmarks Orientation:**

- **Points** 1, 2, 5, 6, 7, 8 & 9 are on the same vertical level (midline), in the frontal view.
- **Points** 2, 3, 4, 11 & 15 are on the same horizontal level (supraorbital line), in the frontal view.
- **Points** 10 & 14 are on the same horizontal level, in the frontal view.
- **Points** 10, 11 &12 are on the same vertical level, in the frontal view.
- **Points** 14, 15 &16 are on the same vertical level, in the frontal view.
- **Points** 7, 18, 19, 22, 24, 26 & 37 are on the same horizontal level.
- **Points** 23, 25, 28 & 39 are on the same horizontal level, in the frontal view.
- **Points** 20, 21, 29 & 40 are on the same horizontal level, in the frontal view.
- **Points** 26, 27, 28 & 29 are on the same vertical level, in the right 3/4 view.
- **Points** 37, 38, 39 & 40 are on the same vertical level, in the left 3/4 view.

- **Points** 13 & 30 are on the same vertical level, midway between the frontal and the right 3/4 view.
- **Points** 17 & 41 are on the same vertical level, midway between the frontal and the left 3/4 view.
- **Points** 30, 31 & 32 are on the same slightly tilted line, in the right lateral/profile view.
- **Points** 32, 33 & 34 are on the same vertical level, in the right lateral/profile view.
- **Points** 42, 43 & 44 are on the same horizontal level, in the left lateral/profile view.
- **Points** 43, 44 & 45 are on the same vertical level, in the left lateral/profile view.

6.10.2 **Landmarks Placing:**

For more practicality, a working model of landmark placing steps is described below.

Stage (I):

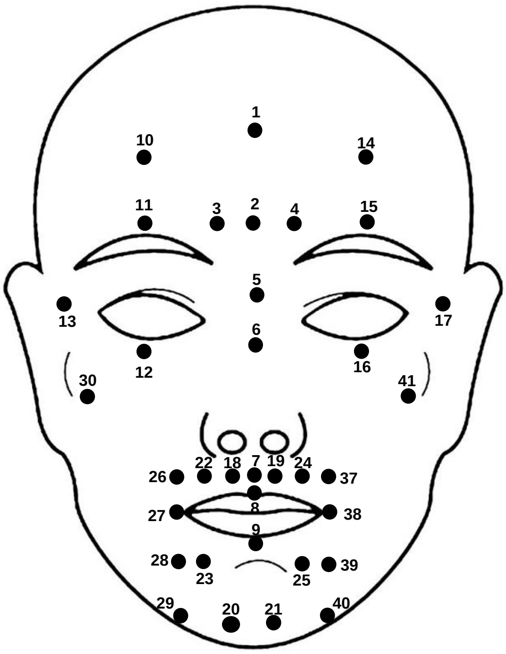
1. Position the face in the frontal view.
2. Place points 1-25 (Table 59 and
3. Figure **66**).

Table 59: Description of the Face Landmarks (1-25) for Adult Egyptian Population

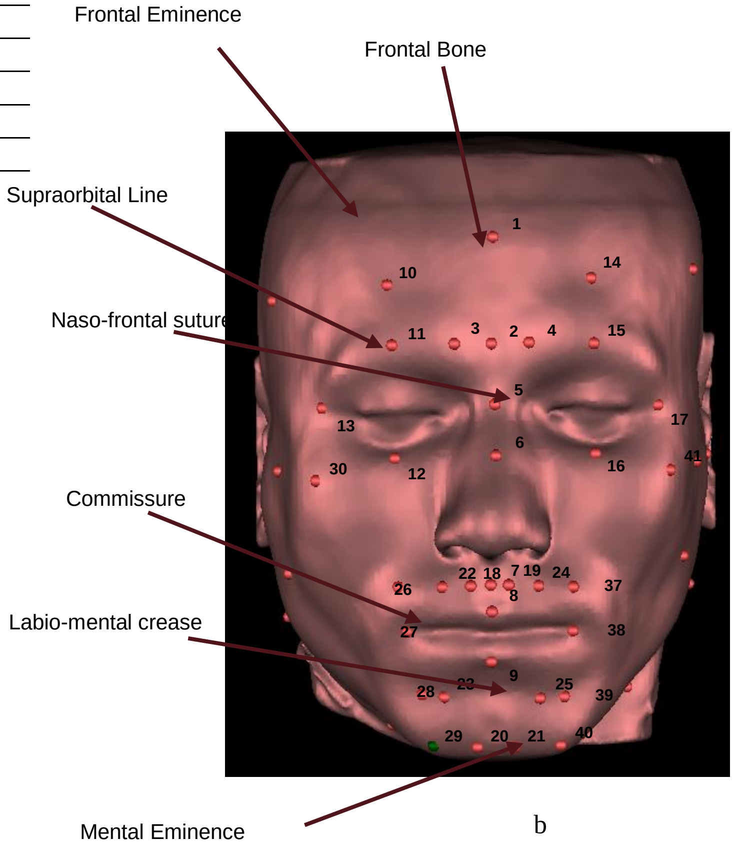
LM No.	Facial LM NAME	Facial LM Definition
1 Midline	Supraglabella	The most anterior point in the midline, above the glabella (LM 2), **midway** between the frontal eminences (Fig. 40).
2 Midline	Glabella	The cross-point between midline and supraorbital line (a horizontal line at the upper border of the orbit) (Fig. 40).
3 (Rt) 4 (Lt)	LateralGlabellar	A point on the soft tissue supraorbital ridge on a vertical line with the inner canthus of the eye, at the junction between the inner border of the orbit and the supraorbital line.
5 Midline	Nasion	A point at the top of the nasal bone, in the midline of the Naso-frontal suture (Fig. 40), at the horizontal level of a line dividing the orbit into upper and lower halves.
6 Midline	Nasal (The end of the nasal)	The end of the nasal bone at the junction between bone and cartilage of the nose.
7 Midline	Midphiltrum	The centre point on the midline **midway** between nose and mouth.
8 Midline	LabialeSuperius	The midline point of the upper lip.
9 Midline	LabialeInferius	The midline point of the lower lip.
10 (Rt) 14 (Lt)	Frontal	A point on the forehead **midway** between the frontal eminence and the maximum curve of the supraorbital margin. It lies at the vertical level of a line passing through the centre point of the upper border of the orbit, and on the same horizontal level of the deepest point in the depression below the frontal eminence.
11 (Rt) 15 (Lt)	SupraOrbital	The centre point of the upper orbital margin. On the horizontal level of the Glabella and Lateral Glabella.
12 (Rt) 16 (Lt)	Infraorbital	The centre point of the lower upper orbital margin. On the horizontal level of the Glabella.

13 (Rt) **17 (Lt)**	Ectoconchion	A point lateral to the outer canthus (angle) of the eye, vertically centred on the orbit, next to the lateral orbital border.
18 (Rt) **19 (Lt)**	Philtrumridge	The prominence on the lateral ridge of the philtrum **midway** between the base of the nostril and upper lip margin. On the horizontal level with the Subspinale.
20 (Rt) **21 (Lt)**	Mentaltubercle.ant	The most prominent point on the lateral bulge of the chin. On the vertical level with Philtrum Ridge, and the horizontal level with mental eminence (Fig. 40).
22 (Rt) **24 (Lt)**	Supralabial	Over the maximum bulge of the canine eminence **midway** between the angle of the mouth and the root of the nostril. On the vertical level with Philtrum Ridge.
23 (Rt) **25 (Lt)**	Sublabial	A point within the labio-mental crease (Fig. 40), on the vertical level of Supralabial.

LM No.	LM NAME
1 (Midline)	Supraglabella
2 (Midline)	Glabella
3 (Rt), 4 (Lt)	LateralGlabellar
5 (Midline)	Nasion
6 (Midline)	Nasal
7 (Midline)	Midphiltrum
8 (Midline)	LabialeSuperius
9 (Midline)	LabialeInferius
10 (Rt), 14 (Lt)	Frontal
11 (Rt), 15 (Lt)	SupraOrbital
12 (Rt), 16 (Lt)	Infraorbital
13 (Rt), 17 (Lt)	Ectoconchion
18 (Rt), 19 (Lt)	Philtrumridge
20 (Rt), 21 (Lt)	Mentaltubercle.ant
22 (Rt), 24 (Lt)	Supralabial
23 (Rt), 25 (Lt)	Sublabial
26(Rt), 37 (Lt)	Supracommissural
27(Rt), 38 (Lt)	Commissural
28(Rt), 39 (Lt)	Subcommissural
29(Rt), 40 (Lt)	Mentaltubercle.lat
30(Rt), 41 (Lt)	MalarOrbitalLevel
31(Rt), 42 (Lt)	Supraglenoid
32(Rt), 43 (Lt)	Midzygoma
33(Rt), 44 (Lt)	Stephanion
34(Rt), 45 (Lt)	Midmasseteric
35(Rt), 46 (Lt)	Postero-masseteric
36(Rt), 47 (Lt)	Antero-masseteric



a



b

Figure 67: Frontal view of a face diagram (a) and the face mesh (b) showing the positions of the face landmarks

Stage (II):

4. Turn the face mesh to the right side of the face to the right 3/4 view.
5. Place points 26-30 (Table 60 and
6. Figure **68**).

Table 60: Description of the Face Landmarks (26-30) for Adult Egyptian Population

LM No. (Rt)	Facial LM NAME	Facial LM Definition
26	Supracommissural	A point over the root of the first premolar. On the horizontal level of Supralabial. On the vertical level of Supracommissural.
27	Commissural	A point on the crown of the first premolar. Immediately posterior to the commissural bulge (the angle of the mouth) (Fig. 40).
28	Subcommissural	A point lateral to the Sub-labial. On the horizontal level of Sublabial. On the vertical level of Supracommissural.
29	Mentaltubercle.lat	A point posterior to and on the horizontal level of the Mental tubercle Anterior. On the vertical level of the Supracommissural.
30	MalarOrbitalLevel	Lined up with the lateral border of the eye on the centre of the Zygomatic process (Fig. 41), in vertical line with the Ectoconchion.

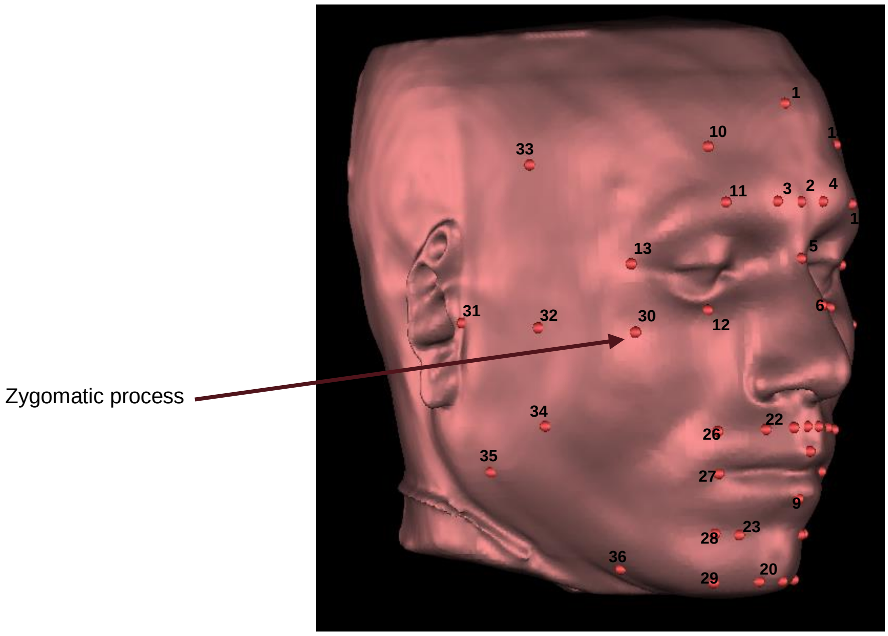


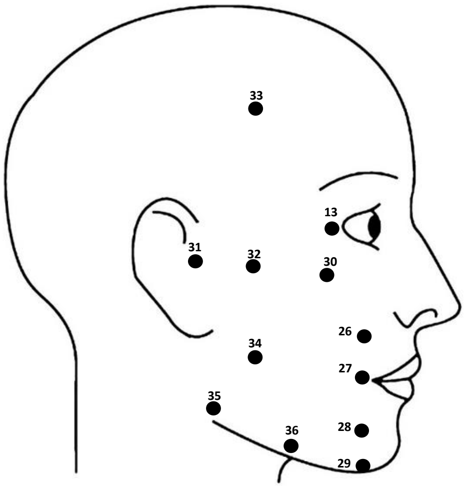
Figure 68: The face mesh in the right 3/4 view showing the positions of the face landmarks

Stage (III):

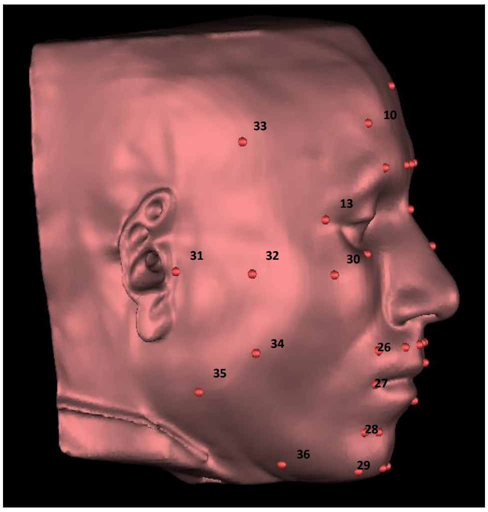
7. Turn the face mesh more to the right side of the face to the right lateral/profile view.
8. Place points 31-36 (Table 61 and Figure 69).

Table 61: Description of the Face Landmarks (31-36) for Adult Egyptian Population

LM No. (Rt)	Facial LM NAME	Facial LM Definition
31	Supraglenoid	A point on the skin surface just in front of the ear.
32	Midzygoma	A point overlying the maximum horizontal and vertical outer curvature of the Zygomatic arch.
33	Stephanion	The point on the side of the skull where the coronal suture crosses the superior temporal line. On the horizontal level of the frontal (LM 10, 14). On the vertical level of the Mid-Zygoma and the Mid-Masseteric (LM 34, 45).
34	Midmasseteric	Middle of the masseter. A point at the centre of the outer surface of the mandibular ramus halfway between the zygomatic arch and the inferior border of the mandible. On the vertical level of the Sephanion and the Mid-Zygoma.
35	Postero-masseteric	This point lies at the lower and posterior edge of the mandible, just anterior to the angle of the mandible, and posterior to the Antero-Masseteric (LM 36, 47).
36	Antero-masseteric	This point lies at the inferior border of the mandible, halfway between the Postero-Masseteric and the lateral mental tubercle.



a



b

Figure 69: Right lateral/profile view of a face diagram (a) and the face mesh (b) showing the positions of the face landmarks

Stage (IV):

9. Turn the face mesh to the left side of the face to the left 3/4 view.
10. Place points 37-41 (Table 62 and
11. Figure **70**).

Table 62: Description of the Face Landmarks (37-41) for Adult Egyptian Population

LM No. (Lt)	Facial LM NAME	Facial LM Definition
37	Supracommissural	A point over the root of the first premolar. On the horizontal level of Supralabial. On the vertical level of Supracommissural.
38	Commissural	A point on the crown of the first premolar. Immediately posterior to the commissural bulge (the angle of the mouth) (Fig. 40).
39	Subcommissural	A point lateral to the Sub-labial. On the horizontal level of Sublabial. On the vertical level of Supracommissural.
40	Mentaltubercle.lat	A point posterior to and on the horizontal level of the Mental tubercle Anterior. On the vertical level of the Supracommissural.
41	MalarOrbitalLevel	Lined up with the lateral border of the eye on the centre of the Zygomatic process (Fig. 41), in vertical line with the Ectoconchion.

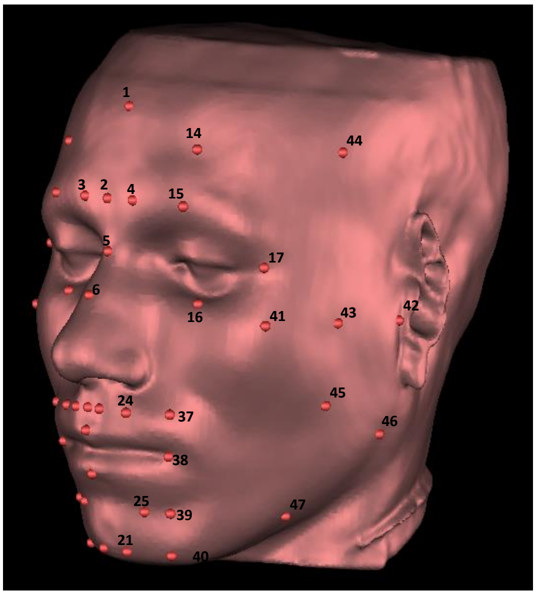


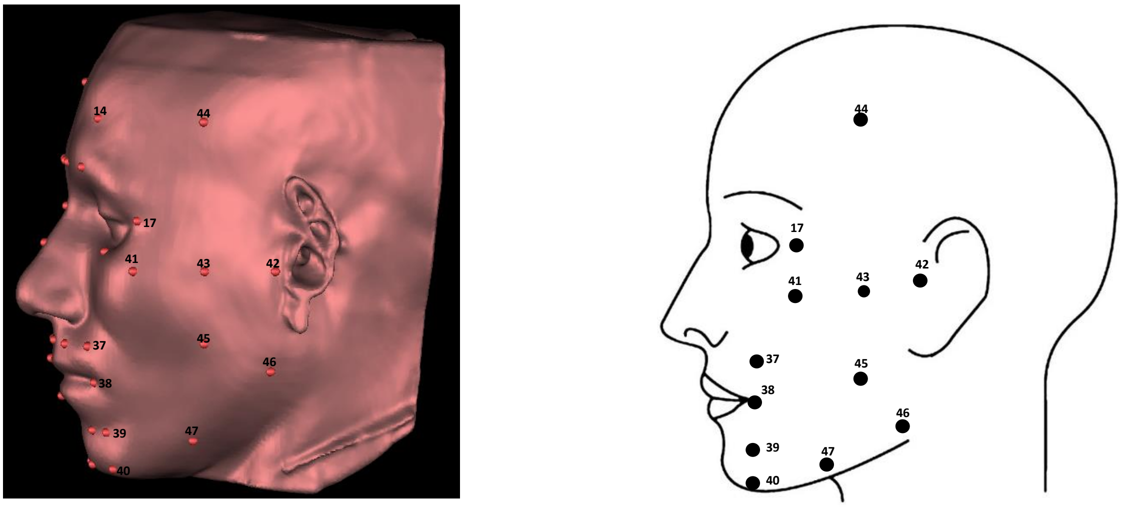
Figure 70: The face mesh in the left 3/4 view showing the positions of the face landmarks

Stage (V):

12. Turn the face mesh more to the left side of the face to the left lateral/profile view.
13. Place points 42-47 (Table 63 and Figure 71).

Table 63: Description of the Face Landmarks (42-47) for Adult Egyptian Population

LM No. (Lt)	Facial LM NAME	Facial LM Definition
42	Supraglenoid	A point on the skin surface just in front of the ear.
43	Midzygoma	A point overlying the maximum horizontal and vertical outer curvature of the Zygomatic arch.
44	Stephanion	The point on the side of the skull where the coronal suture crosses the superior temporal line. On the horizontal level of the frontal (LM 10, 14). On the vertical level of the Mid-Zygoma and the Mid-Masseteric (LM 34, 45).
45	Midmasseteric	Middle of the masseter. A point at the centre of the outer surface of the mandibular ramus halfway between the zygomatic arch and the inferior border of the mandible. On the vertical level of the Sephanion and the Mid-Zygoma.
46	Postero-masseteric	This point lies at the lower and posterior edge of the mandible, just anterior to the angle of the mandible, and posterior to the Antero-Masseteric (LM 36, 47).
47	Antero-masseteric	This point lies at the inferior border of the mandible, halfway between the Postero-Masseteric and the lateral mental tubercle.



a b

Figure 71: Left lateral/profile view of the face mesh (a) and a face diagram (b) showing the positions of the face landmarks

6.11 STEP (11): SAVING THE FACE MESH

As described in section 3.11.

6.12 STEP (12): COMPLETING THE WARPING PROCESS

As described in section 3.12.

APPENDIX 21: THE STUDY PRESENTATIONS

A) Queen Mary University of London:

1- Three-Minute Thesis Heats, School of Medicine and Dentistry Heats:

- June 11th, 2015.

2- William Harvey Research Institute Annual Review Day:

- June 30th, 2015.
- Oral presentation, titled "***Designing a Face Pool Test for the Subjective Assessment of 3D Digital Forensic Facial Reconstruction***".

3- William Harvey Research Institute 30 years Anniversary Celebration:

- June 23rd - 24th, 2016.
- Poster presentation, titled ***"Subjective and Objective Assessment of 3D Forensic Facial Reconstruction of Egyptian Population using Average Facial Templates"***.

4- William Harvey Day:

- October 18th, 2016.
- Poster presentation, titled ***"Subjective and Objective Assessment of 3D Forensic Facial Reconstruction of Egyptian Population using Average Facial Templates"***.

B) Forensic Sciences National Conferences:

1- FORREST (Forensic Research & Teaching) 2015 Conference:

- June 30th - July 2nd, 2015.
- Glasgow, Scotland, UK.
- Oral presentation, titled **"*Designing a Face Pool Test for the Subjective Assessment of 3D Digital Forensic Facial Reconstruction*"**.
- Three Poster presentations, titled:
  - ***"Comparing Single and Average Human Faces as Facial Templates for 3D Digital Forensic Facial Reconstruction"***.
  - ***"The Influence of Facial Soft Tissue Thickness Measures on the Accuracy of 3D Digital Forensic Facial Reconstruction"***.

- ***"The Relation between the Observer's Sex, Race and Age and the Performance in the Subjective Assessment Tests of 3D Digital Forensic Facial Reconstruction"***.

C) Forensic Sciences International Conferences:

1- FASE (Forensic Anthropology Society of Europe) one Day Symposium:

- September 5th, 2015.
- Montpellier, France.
- Oral presentation, titled:

  "Comparing Single and Average Human Faces as Facial Templates for 3D Digital Forensic Facial Reconstruction".
- Two Poster presentations, titled:
  - ***"Designing a Face Pool Test for the Subjective Assessment of 3D Digital Forensic Facial Reconstruction"***.
  - ***"The Influence of Observer's Sex, Ancestry and Age on the Correct Identification Rates of Forensic Facial Reconstructions"***.