

**Statistical Analysis and Visualisation of
3D Anthropological and Archaeological
Materials**

Wuyang Shui

PhD

University of York

Archaeology

January 2023

Abstract

The expanding scale of 3D hominid fossils and archaeological artefacts have encouraged a growing use of rigorous scientific study to reveal mechanisms of human evolution and understand human activities and behaviour. Often statistical analysis and visualisation methods, e.g. geometric morphometrics, are applied to obtain what may appear to be reliable findings. However, owing to the diversity and complexity of research materials and questions, off-the-shelf methods may not work well. Thus, there are some important considerations that need to be borne in mind regarding the interdisciplinary study, and innovative approaches are required to enhance performance.

There are multiple challenges that I have addressed. First, three different semilandmarking approaches on morphometric analyses and visualisation of mean and allometrically scaled surfaces are assessed. These approaches produce different semilandmark locations, which in turn lead to different results, although non-rigid approaches are broadly consistent. My second concern is about virtual restoration by means of reflection. A landmark-free method is presented to quantify gross and regional surface asymmetry, and then a landmark-based deformation method is developed to compute geometric models of missing data with the predicted boundary curve as a constraint. Finally, I explore craniofacial relationships between hard and soft tissues among modern humans, and develop computerised methods to recreate probable faces of archaic humans and *Homo sapiens* based on the learnt relationships. Results suggest average dense facial soft tissue thickness depths contribute to enhancing the approximation accuracy. Although nasal (and oral) hard tissues have an effect on the corresponding soft tissues, some caution is needed when approximating soft tissue structures.

Overall, this thesis makes contributions within the field of virtual anthropology and archaeology. I examine the effect of semilandmarking approaches on geometric morphometrics, and develop new methods for asymmetry detection, virtual restoration and facial approximation. These proposed methods have been applied to different case studies.

Table of Contents

Abstract	2
Table of Contents	3
List of Tables	10
List of Figures	14
Acknowledgements	19
Declaration	20
Chapter 1 Introduction	22
1.1 General background and problem statement.....	23
1.1.1 Geometric morphometrics	23
1.1.2 Different applications	25
1.2 Motivation and objectives.....	28
1.3 Research questions.....	29
1.4 Structure of thesis	30
1.4.1 Chapter 2	30
1.4.2 Chapter 3	31
1.4.3 Chapter 4	31
1.4.4 Chapter 5	32
1.4.5 Chapter 6	33
1.4.6 Chapter 7	33
1.4.7 Chapter 8	34
1.5 References.....	35
Chapter 2 A comparison of semilandmarking approaches in the analysis of size and shape	43
2.1 Introduction.....	44
2.2 Materials and Methods.....	51
2.2.1 Materials	51
2.2.1.1 Adult human heads	51
2.2.1.2 Ape crania	51

2.2.2	Methods	52
2.2.2.1	Three semilandmarking approaches.....	52
2.2.2.1.1	Generation of template.....	52
2.2.2.1.2	Semilandmarking approaches	53
2.2.2.2	Comparison of three semilandmarking approaches	55
2.2.2.2.1	The locations of semilandmarks	55
2.2.2.2.2	Comparisons of mean landmark and semilandmark configurations	55
2.2.2.2.3	Procrustes distances among specimens obtained using different semilandmarking approaches and densities.....	55
2.2.2.2.4	PCA and allometry.....	56
2.3	Results.....	57
2.3.1	Human head scans	57
2.3.1.1	The locations of semilandmarks	57
2.3.1.2	Comparison of landmark and semilandmark configurations among different semilandmarking approaches.....	59
2.3.1.3	Comparison of centroid sizes and Procrustes distance matrices.....	61
2.3.1.3.1	Differences between semilandmarking approaches	61
2.3.1.3.2	Differences densities of semilandmarks	63
2.3.1.4	PCA and allometry.....	64
2.3.1.4.1	Correlations between PC scores from different semilandmarking methods and densities.....	64
2.3.1.4.2	Comparison of allometric vectors.....	67
2.3.1.4.3	Comparison of allometric scaling of landmark and semilandmark configurations.....	67
2.3.2	Ape crania.....	68
2.3.2.1	The locations of semilandmarks	68
2.3.2.2	Differences among Mean Landmark and Semilandmark Locations.....	70
2.3.2.3	Comparison of centroid sizes and Procrustes distance matrices.....	70
2.3.2.4	PCA and allometry.....	72
2.4	Discussion.....	74
2.4.1	Tests of hypotheses.....	74

2.4.1.1	Hypothesis i	74
2.4.1.2	Hypothesis ii	75
2.4.1.3	Hypothesis iii	75
2.4.1.4	Hypothesis iv	76
2.4.1.5	Hypothesis v.....	78
2.4.1.6	Hypothesis vi	80
2.4.2	Significance and implications of findings	80
2.4.3	Limitations and future work	85
2.5	Conclusion	86
2.6	Acknowledgments.....	86
2.7	References.....	87
	Supplementary material	93
	Chapter 3 A comparison of semilandmarking approaches in the visualisation of shape differences.....	94
3.1	Introduction.....	95
3.2	Materials and Methods.....	98
3.2.1	Materials	98
3.2.1.1	Datasets and templates	98
3.2.1.2	Semilandmarks.....	99
3.2.2	Methods	101
3.2.2.1	Comparisons of mean surface meshes between different approaches	101
3.2.2.2	Comparisons of allometrically scaled surface meshes.....	104
3.2.2.3	Comparisons of surface meshes resulting from different semilandmarking densities.....	104
3.2.2.4	Comparisons of mean and allometrically scaled surface meshes resulting from landmarks alone	105
3.3	Results.....	106
3.3.1	Comparison of estimates of mean surfaces	106
3.3.1.1	Different semilandmarking approaches	106
3.3.1.2	Different densities of semilandmarks.....	110

3.3.2	Comparison of estimates of allometrically scaled surfaces.....	113
3.3.2.1	Different semilandmarking approaches	114
3.3.2.2	Different densities of semilandmarks.....	118
3.3.3	Comparisons of mean and allometrically scaled surface resulting from landmarks alone.....	122
3.4	Discussion.....	127
3.5	Conclusion	133
3.6	Acknowledgments.....	134
3.7	References.....	135
	Supplementary material	139
Chapter 4 A Landmark-free Approach for Surface Asymmetry Detection and Profile Drawings from Bilaterally Symmetrical Geometry		140
4.1	Introduction.....	141
4.2	Related work	142
4.3	Materials and methods	144
4.3.1	Materials.....	144
4.3.2	Methods	145
4.3.2.1	An approximation of the plane of symmetry	146
4.3.2.2	Asymmetry detection.....	151
4.3.2.3	Profile drawings.....	152
4.4	Results.....	152
4.4.1	The approximation of the plane of symmetry	152
4.4.1.1	The plane of symmetry in synthetic data	152
4.4.1.2	The plane of symmetry in real objects.....	153
4.4.2	Asymmetry detection.....	156
4.4.3	Profile drawings.....	157
4.5	Discussion.....	158
4.6	Conclusion	160
4.7	Acknowledgements.....	160
4.8	References.....	161

Supplementary material	166
Chapter 5 A geometric completion and shape analysis method for damaged bilaterally symmetrical artefacts.....	167
5.1 Introduction.....	168
5.2 Related work	170
5.3 Materials and methods	171
5.3.1 Materials	171
5.3.2 Methods	172
5.3.2.1 The restoration of the damage mask	173
5.3.2.2 Shape analysis of the mask	177
5.4 Results.....	178
5.4.1 The restoration of the damaged mask.....	178
5.4.2 Shape analysis of the mask.....	183
5.5 Discussion.....	184
5.6 Conclusion	186
5.7 Acknowledgements.....	187
5.8 References.....	188
Chapter 6 A computerized facial approximation method for archaic humans based on dense facial soft tissue thickness depths.....	192
6.1 Introduction.....	193
6.2 Materials and methods	195
6.2.1 Materials	195
6.2.1.1 The archaic human fossil	195
6.2.1.2 Skull and face datasets of modern living humans.....	197
6.2.1.3 Anatomical landmark definitions.....	198
6.2.2 Methods.....	200
6.2.2.1 The restoration of JNS 1	201
6.2.2.2 Computerized facial approximation.....	201
6.2.2.3 Morphological shape variations of facial approximation.....	205
6.2.2.4 The effects of FSTDs and skull morphology on the approximated face.....	206

6.3	Results.....	206
6.3.1	Facial approximation of JNS 1	206
6.3.2	Facial approximation of modern humans	210
6.3.3	The effects of FSTDs and skull morphology on facial approximation.....	212
6.3.4	Shape analysis of the approximated faces	213
6.4	Discussion	214
6.5	Conclusion	217
6.6	Acknowledgements.....	218
6.7	References.....	219
	Supplementary material	224
	Chapter 7 A computerized facial approximation method for <i>Homo sapiens</i> based on facial soft tissue thickness depths and geometric morphometrics	225
7.1	Introduction.....	226
7.2	Materials and Methods.....	228
7.2.1	Materials	228
7.2.2	Methods	231
7.3	Results.....	237
7.3.1	Quantification of craniofacial relationships among modern humans	237
7.3.2	Evaluation of the proposed method	242
7.3.3	FA of the UC 101 skull	244
7.4	Discussion	245
7.4.1	The quantification of craniofacial relationships	245
7.4.2	The computerized FA method	247
7.4.3	Facial approximation of UC 101	248
7.5	Conclusion	249
7.6	Acknowledgements.....	250
7.7	References.....	251
	Chapter 8 Conclusion and future research.....	257
8.1	General summary	257
8.2	Key findings and discussion	259

8.2.1	Different semilandmarking approaches and densities	259
8.2.2	Asymmetry detection and virtual restoration	261
8.2.3	Facial approximation	262
8.3	Future research.....	265
8.4	References.....	267

List of Tables

Table 2.1 Comparison of semilandmarks from sliding TPS and TPS&NICP	57
Table 2.2 Comparison of semilandmarks from sliding TPS and LS&ICP	57
Table 2.3 Comparison of semilandmarks from TPS&NICP and LS&ICP	58
Table 2.4 The centroid sizes of the mean landmark and semilandmark configurations generated by different semilandmarking approaches and different densities of semilandmarks	59
Table 2.5 Procrustes distances computed between mean landmark and semilandmark configurations	60
Table 2.6 Procrustes distances computed between mean landmark and semilandmark configurations in the face and scalp.....	61
Table 2.7 Pearson correlations among vectors of Procrustes distances between each individual and the mean and Mantel tests of association between the Procrustes distance matrices derived using different semilandmarking approaches.....	63
Table 2.8 Pearson correlations among vectors of Procrustes distances and Mantel tests comparing Procrustes distance matrices between each density of semilandmarking and the maximum density.	64
Table 2.9 Pearson and Mantel correlations between vectors and matrices of Procrustes distances from each semilandmarking approach and density and those from landmarks alone.....	64
Table 2.10 Pearson correlations between PC1 and PC 2 scores derived using different semilandmarking approaches.....	65
Table 2.11 Pearson correlations of PC1 and PC 2 scores between each semilandmark density and the maximum (1000) density	65
Table 2.12 Pearson correlations of PC1 and PC 2 scores from landmarks alone and those from each semilandmarking approach and density.....	66
Table 2.13 Pearson correlations between PC1 and PC 2 scores based on facial landmarks and semilandmarks derived using different semilandmarking approaches	66
Table 2.14 Correlations of PC1 and PC 2 scores between each facial landmark and semilandmark configuration at lower densities of semilandmarks and the configuration with the maximum density.	66
Table 2.15 Correlations of PC1 and PC 2 scores derived from the landmarks alone and each facial landmark and semilandmark configuration generated by different approaches and densities....	67
Table 2.16 The angles (°) between allometric vectors from different semilandmarking approaches	

and densities.....	67
Table 2.17 Comparison of Procrustes distances between the predicted shapes corresponding to the maximum and minimum centroid size derived using semilandmarking approaches and densities	68
Table 2.18 Comparison of semilandmarks from sliding TPS and TPS&NICP approaches	69
Table 2.19 Comparison of semilandmarks from sliding TPS and LS&ICP approaches	69
Table 2.20 Procrustes distance (dist) between mean landmark and semilandmark configurations derived at varying densities from sliding TPS and TPS&NICP approaches.....	70
Table 2.21 Pearson correlations among the vectors of Procrustes distances between each ape cranium and the mean and Mantel tests among Procrustes distance matrices.....	71
Table 2.22 Pearson and Mantel correlations between vectors and matrices of Procrustes distances from each semilandmarking approach and density and those from the landmarks alone.....	71
Table 2.23 Comparison of Pearson correlations of PC1 and PC 2 between Sliding TPS and TPS&NICP	72
Table 2.24 Pearson correlations among PC scores from each semilandmarking density and from the landmarks alone	73
Table 2.25 Pearson correlations among PC scores from each semilandmarking density and the maximum density.....	73
Table 2.26 The angles (°) between allometric vectors derived using sliding TPS and TPS&NICP approaches.....	73
Table 2.27 Procrustes distances between the predicted landmark and semilandmark configurations from sliding TPS and TPS&NICP corresponding to the maximum and minimum centroid sizes.....	74
Table 2.S1 Ape cranium specimen	93
Table 3.1 Procrustes distances computed between the mean surfaces of ape crania generated by Sliding TPS and TPS&NICP after re-semilandmarking and re-warping the template mesh....	108
Table 3.2 Procrustes distances computed between the mean landmarks and semilandmarks of ape crania generated by Sliding TPS and TPS&NICP after re-semilandmarking the warped mesh.	109
Table 3.3 Percentages of total variance explained by PC 1 and PC 2 of mean surface shape...	110
Table 3.4 Procrustes distances between the vertices of the estimated mean human head surfaces using 1000 semilandmarks and those using increasing numbers of semilandmarks from each approach after re-semilandmarking and re-warping the template mesh	111
Table 3.5 Procrustes distances between the vertices of mean ape cranial surfaces estimated by	

each approach, using 800 semilandmarks and those estimated using increasing numbers of semilandmarks after re-semilandmarking and re-warping the template mesh.....	111
Table 3.6 Procrustes distances computed between vertices of ape cranial surfaces, allometrically scaled to the maximum and minimum centroid size, from sliding TPS and TPS&NICP semilandmarking approaches after re-semilandmarking and re-warping the template mesh	116
Table 3.7 Procrustes distances between vertices of the allometrically scaled surfaces of heads at the maximum and minimum centroid sizes using the landmarks and highest density of semilandmarks and surfaces estimated using the landmarks and lower densities of semilandmarks, after re-semilandmarking and re-warping.....	118
Table 3.8 Procrustes distances between vertices of the estimated predictions of ape cranial surfaces at the maximum and minimum centroid sizes derived from the maximum density of semilandmarks and those from lower densities of semilandmarks, after re-semilandmarking and re-warping	120
Table 3.S1 Ape cranium specimen	139
Table 4.1 The approximated planes of symmetry of different artefacts	155
Table 4.S1 Archaeological and anthropological objects	166
Table 5.1 Eighteen Landmarks	172
Table 5.2 Geometric deviation between the damaged mask and the resulting deformation using different coefficients	181
Table 5.3 The comparison of anthropometric measurements	184
Table 6.1 Anatomical landmarks	198
Table 6.2 Eight patches on left side of JNS 1	200
Table 6.3 FSTDs deviation (mm) and the percentage distribution between three approximations and the template	209
Table 6.4 FSTDs deviation (mm) and percentage distribution between the approximated and actual faces.....	211
Table 6.5 Geometric deviation (mm) and percentage distribution between the approximated and actual faces.....	212
Table 7.1 Anatomical facial landmarks	230
Table 7.2 The 2B-PLS analysis between nasal hard and soft tissue shapes	238
Table 7.3 A comparison of the approximated and actual nasal soft tissues based on Procrustes distance	240
Table 7.4 2B-PLS analysis between mouth hard and soft tissue shapes	241

Table 7.5 Comparison of approximated and actual mouth soft tissue shapes based on Procrustes distance 242

List of Figures

Figure 2.1 Two datasets with anatomical landmarks.....	52
Figure 2.2 Workflow of comparison of three semilandmarking landmarks.....	52
Figure 2.3 Locations of the average coordinates of 1000 semilandmarks generated by sliding TPS (black points), LS&ICP (red points) and TPS&NICP (green points) approaches	58
Figure 2.4 Average differences (Euclidean distance in location in mm) between 1000 semilandmarks generated by different approaches	59
Figure 2.5 The mean semilandmarks generated by sliding TPS over the face (yellow points) and scalp (red points).....	60
Figure 2.6 Comparison of the centroid sizes of landmarks and semilandmark configurations computed by different approaches	61
Figure 2.7 Vectors of Procrustes distances between each individual and the mean computed for each semilandmarking approach using different densities of semilandmarks	62
Figure 2.8 Comparison of the vector of Procrustes distances between every specimen and the mean among different approaches	62
Figure 2.9 800 semilandmarks generated by different semilandmarking approaches on the mean surface generated by Sliding TPS	69
Figure 2.10 The average differences in location (mm) between 800 semilandmarks generated by different approaches.....	70
Figure 2.11 Comparisons of centroid sizes estimated by different semilandmarking approaches.	72
Figure 2.12 Comparison of the vector of Procrustes distances from the mean between semilandmarking approaches.....	72
Figure 2.13 Visualisation of superimposed scatterplots of PC1 and PC 2 from analyses of 20 ape crania using landmarks and semilandmarks from sliding TPS and TPS&NICP approaches with varying densities	73
Figure 3.1 Two datasets.....	99
Figure 3.2 Different densities of semilandmarks generated by the sliding TPS (black points), LS&ICP (red points) and TPS&NICP (green points) approaches	101
Figure 3.3 Visualisation of the differences in mesh triangle surface areas among mean surface shapes generated using different semilandmarking approaches	102
Figure 3.4 Procrustes distances computed between vertices of the mean surfaces of human heads	

generated by warping the template mesh to the semilandmarks obtained by different approaches.	103
Figure 3.5 Visualisation of the head mesh surfaces generated using different semilandmarking approaches after re-semilandmarking and re-warping.....	106
Figure 3.6 Procrustes distances computed between the mean surfaces of human heads obtained by different approaches after re-semilandmarking and re-warping the template mesh	107
Figure 3.7 Visualisation of the differences in mesh triangle surface areas among mean surface shapes generated using different semilandmarking approaches after re-semilandmarking and re-warping	108
Figure 3.8 Surface meshes of the estimated mean ape cranium generated using different semilandmarking approaches after re-semilandmarking and re-warping	109
Figure 3.9 Visualisation of the regional differences in local surface areas of mean ape cranial shapes from sliding TPS (reference) and TPS&NICP (target) approaches.....	110
Figure 3.10 Superimposed scatterplots of PC 1 and PC 2 of mean shape surfaces using sliding TPS (red points), LS&ICP (green points) and TPS&NICP (cyan points) approaches.....	111
Figure 3.11 Procrustes distances of estimated mean surfaces between every density and the maximum density from each approach to semilandmarking	112
Figure 3.12 Colour maps of local surface mesh area differences among mean surfaces of human heads derived using lower densities (target) and that from 1000 semilandmarks (reference) using different semilandmarking approaches	113
Figure 3.13 Colour maps of local surface mesh area differences among mean surfaces of ape crania derived using lower densities (target) and 800 semilandmarks (reference) based on different semilandmarking approaches.....	113
Figure 3.14 Procrustes distances computed between the vertices of human head surfaces, allometrically scaled to the maximum and minimum centroid sizes, based on different semilandmarking approaches after re-semilandmarking and re-warping the template mesh	114
Figure 3.15 Visualisation of the differences in mesh triangle surface areas among predicted allometrically scaled surfaces of human heads representing the maximum and minimum centroid size generated by different semilandmarking approaches.....	115
Figure 3.16 Visualisation of the differences in mesh triangle surface areas among fitted surface shapes of ape crania generated by the sliding TPS (reference) and TPS&NICP (target) approaches after re-semilandmarking and re-warping of the template.....	116
Figure 3.17 PC1 (92.4% of total variance) vs PC2 (4.7%) from PCA of the mean and allometrically scaled head surfaces derived using varying densities of semilandmarks and each semilandmarking approach	117

Figure 3.18 PC1 (99.5% of total variance) vs PC2 (0.15%) from PCA of the mean and allometrically scaled head surfaces derived using varying densities of semilandmarks and the sliding TPS and TPS&NICP semilandmarking approaches.....	117
Figure 3.19 Procrustes distances, after re-semilandmarking and re-warping, between the allometrically scaled head surfaces derived from the maximum density and those from lower densities of semilandmarks	119
Figure 3.20 Colour map of local area differences computed between the re-semilandmarked and re-warped allometric predictions of surfaces of human heads at the maximum and minimum sample centroid sizes, computed between lower densities (reference) and the maximum density (target) of semilandmarking.....	120
Figure 3.21 Procrustes distances, after re-semilandmarking and re-warping, between the allometrically scaled ape cranial surfaces derived from the maximum density and those from lower densities of semilandmarks	120
Figure 3.22 Colour map of local area differences computed between the re-semilandmarking and re-warped allometric prediction of surfaces of ape crania at the maximum and minimum sample centroid sizes, computed between lower densities (reference) and the maximum density (target) of semilandmarking	121
Figure 3.23 The effects of different template surfaces and semilandmarks on the visualisations of mean head	122
Figure 3.24 The effects of different template surfaces and semilandmarks on the visualisations of mean ape cranium	123
Figure 3.25 The effects of different template surfaces and semilandmarks on the visualisations of allometric prediction of surface at the maximum centroid size	124
Figure 3.26 The effects of template surface and semilandmarks on the visualisation of allometric prediction of surface at the minimum centroid size	124
Figure 3.27 Scatter plots of the mean and allometrically scaled head surfaces	126
Figure 3.28 Scatter plots of the mean and allometrically scaled ape surfaces	126
Figure 4.1 The workflow of asymmetry detection and profile drawings based on the identification of the plane of symmetry	146
Figure 4.2 The WKS values of two models at different four energy frequencies, including $k=1$, $k=10$, $k=50$ and $k=100$	147
Figure 4.3 The workflow of the extraction of symmetrical regions from segmented regions.	148
Figure 4.4 The extraction of symmetrical regions	150
Figure 4.5 The qualification of surface asymmetry of the mandible	151

Figure 4.6 The comparison of the approximated and actual MSPs in synthetic skulls. (a) A nearly perfect skull; (b) An asymmetrical skull. (c) A damaged skull with missing data	153
Figure 4.7 The comparison of the approximated and landmark-based MSPs.....	154
Figure 4.8 The identification of MSPs in three damaged skulls	154
Figure 4.9 The comparison of the approximated planes of symmetry generated by the proposed (red colour) and the landmark-free morphometric (green colour) methods.....	156
Figure 4.10 The identification of the plane of symmetry in a terracotta warriors hand.....	156
Figure 4.11 Surface asymmetry detection of thirty skulls.....	157
Figure 4.12 Surface asymmetry detection of archaeological artefacts.....	157
Figure 4.13 Profile drawings of symmetrical objects.....	158
Figure 5.1 The metal mask.....	169
Figure 5.2 The digital mask as repaired through the reassembled fragments	172
Figure 5.3 The pipeline of the proposed method.....	173
Figure 5.4 The repaired boundary curve of the missing area	175
Figure 5.5 The reflection structure of the intact side of the mask.....	179
Figure 5.6 The deformation of the template model using our method.....	180
Figure 5.7 The evaluation of the reliability of the restoration.....	181
Figure 5.8 The effect of every weight on the deformation.....	182
Figure 5.9 The restoration of synthetic data.....	183
Figure 5.10 Shape variations between human faces and the masks in shape space.....	184
Figure 6.1 JNS 1 cranium (peach colour) and two different mandibles (grey colour).....	197
Figure 6.2 The pipeline of the computerized facial approximation method	201
Figure 6.3 The restoration of JNS 1 using Tabun 2 that comprised the anatomical landmark and semilandmark configuration.	201
Figure 6.4 Segmentation of template skull.....	204
Figure 6.5 The deformation of the template skull and face using TPS	207
Figure 6.6 The deformation of the template skull using the proposed method.....	207
Figure 6.7 The facial approximation of JNS 1 using the proposed method.....	208
Figure 6.8 Comparisons of the facial approximations generated by different methods.....	209
Figure 6.9 The coarsely approximated face using the normal vectors of overall the vertices.	210

Figure 6.10 The coarsely approximated face using the FSTDs along the cylindrical sampling vectors.....	210
Figure 6.11 Four approximation examples of modern humans.....	211
Figure 6.12 Facial approximation based on the average FSTDs of the males	212
Figure 6.13 Facial approximation of JNS 1 using Mauer 1	213
Figure 6.14 Three approximated faces using linear interpolation.....	213
Figure 6.15 Comparison of the approximated faces and modern human faces in the shape space.	214
Figure 6.S1 Different approximated faces of JNS 1 based on different FSTDs and mandibles	224
Figure 7.1 Average skull and face including the landmark and semilandmark configuration.	230
Figure 7.2 Semilandmarks of UC 101.....	231
Figure 7.3 Workflow of quantification of craniofacial relationships and FA method	232
Figure 7.4 The coarsely approximated face	235
Figure 7.5 Average FSTDs at landmarks and semilandmarks.....	238
Figure 7.6 Covariation of PLS1 between nasal hard and soft tissue shapes	239
Figure 7.7 Geometric discrepancy of approximated and actual nasal soft tissue shapes	240
Figure 7.8 Covariation of PLS1 between mouth hard and soft tissue shapes	241
Figure 7.9 Geometric discrepancy of the approximated and mouth soft tissue shapes.....	242
Figure 7.10 A comparison of the approximated and actual faces.....	243
Figure 7.11 Colour maps of the approximation difference between the approximated and actual faces	243
Figure 7.12 The approximated and actual faces of eight test samples	244
Figure 7.13 Facial approximation of UC 101	245
Figure 8.1 The structure of the thesis.....	258

Acknowledgements

This thesis would not have been possible without the support and encouragement of many people. I am very delighted to express my gratitude to them.

First and foremost, I would like to thank my supervisors, Professor Julian Richards and Professor Paul O'Higgins, who gave me valuable support, comments and mentoring, and encouraged me to write a thesis on this topic. Their efforts in editing and proofreading my reports, papers and thesis, and inspirational ideas from our supervision and thesis advisory panel (TAP) meetings were really very helpful and encouraging. Their positive outlook and confidence in my research inspired me and gave me tremendous assurance. Additionally, I want to thank them to let me to begin my new academic journey in the archaeology and anthropology fields.

I am grateful to all the people and institutions who provided access to modern human head datasets, fossil specimens, and archaeological artefacts. Additionally, I would like to extend my thanks to all my friends and colleagues who gave me valuable discussion, comments and suggestions on my papers and thesis: Professor Nick Pears, Dr. Christian Duncan, Dr. Antonio Profico, Dr. Olivia Smith, Professor Xiujie Wu, Professor Mingquan Zhou, Professor Zhongke Wu, Professor Shengling Geng, Professor Wu Liu, Mr. Yuan Ji, Mr. Fei Gao, Mr. Taiping He, Dr. Yameng Zhang, Dr. Qingqiong Deng, Dr. Xia Zheng, Dr. Pianpian Wei, Dr. Dan Zhang, and Dr. Haijun Li.

I want to express my gratitude to the Department of Archaeology at the University of York. Thanks to Dr. Colleen Morgan and Dr. Jessica Hendy for their comments and suggestions in my TAP and progression meetings. My thanks go to all the members of PalaeoHub who expand my horizons and knowledge in anthropology and paleoanthropology during annual academic meeting, especially thank to Dr. Samuel Cobb, Dr. Laura Fitton, Dr. Philip Cox, Dr. Andrew Cuff. I am fortunate to be accompanied by a group of PhD students and Postdoctoral researchers supervised by Professor Paul O'Higgins: Dr. Antonio Profico, Dr. Lou Albessard-Ball, and Lisa Genochio. Additionally, my thanks go to all the members of Archaeology Data Service (ADS), especially thank to Dr. Holly wright and Dr. Tim Evans. I would like to thank anonymous reviewers and editors for their feedback and comments during the review process of my papers submitted to different journals. I extend my sincere thanks to Dr. Laura Fitton for being my internal examiner and Professor Bastir Markus for being my external examiner.

Last but not least, this thesis is dedicated to my wife, Dr. Wenxuan Li, and my daughter Miya. My deepest gratitude goes to my wife for her love, companionship and everlasting encouragement. Thanks our parents for supporting my choice to fulfil my potential and aspirations.

Declaration

I declare that this thesis is a presentation of original work and I am the sole author. This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as References. For the co-authored papers presented as Chapters 2-7 in Section 1.4, I provide details of the respective contributions of each co-author to every paper.

At the time of submission, the following papers as part of this thesis have been published:

- Chapter 3
Wuyang Shui, Antonio Profico and Paul O’Higgins. A comparison of semilandmarking approaches in the visualisation of shape differences. *Animals*. 2023, 13(3): 385.
- Chapter 5
Wuyang Shui and Fei Gao. A geometric completion and shape analysis method for damaged bilaterally symmetrical artefacts. *Journal of Cultural Heritage*. 2021, 52: 118-127.
- Chapter 6
Wuyang Shui, Yameng Zhang, Xiujie Wu and Minquan Zhou. A computerized facial approximation method for archaic humans based on dense facial soft tissue thickness depths. *Archaeological and Anthropological Science*. 2021, 13:186.

At the time of submission, the following paper as part of this thesis has been accepted:

- Chapter 4
Wuyang Shui, Pianpian Wei, Xia Zheng, Shengling Geng. A landmark-free approach for surface asymmetry detection and profile drawings from bilaterally symmetrical geometry. *ACM Journal on Computing and Cultural Heritage*. 2022. (accepted)

At the time of submission, the following papers as part of this thesis have been submitted for publication:

- Chapter 2

Wuyang Shui, Antonio Profico and Paul O’Higgins. A comparison of semilandmarking approaches in the analysis of size and shape. *Animals*. 2023. (under review)

- Chapter 7

Wuyang Shui, Xiujie Wu and Minquan Zhou. A computerized facial approximation method for *Homo sapiens* based on facial soft tissue thickness depths and geometric morphometrics. *Journal of Anatomy*. 2022. (under review)

Chapter 1 Introduction

Physical anthropologists and archaeologists try to make sense of human evolution and history by observing and analysing its tangible remains (e.g. fossil specimens and artefacts). Examples of particular interest include the examination of morphological variation and taxonomic classification of hominid fossils (Hershkovitz et al., 2021) (Lacruz et al., 2019), and modelling and simulation of human activities and behaviour (Brooks et al., 2018; Dolbunova et al., 2022). The expanding scale of fossils and artefacts has encouraged a growing use of a computational methods to understand variation and the factors that contribute to it, minimising subjective interpretation and misunderstanding. The underpinning idea is to translate the research question at hand into a geometric or statistical question. For example, traditional morphometric methods are applied to make quantitative comparisons of organismal form in the study of morphological variations based on the statistical analysis of a set of measurements (e.g. distances, angles and ratios) (Harvati et al., 2011; Roth & Mercer, 2000). However, such measurements tend to capture limited detailed information and statistical methods alone cannot provide the graphical representation to improve understanding of shape variations. A more powerful statistical analysis method in conjunction with visualisation is desirable.

The past thirty years witnessed the fast development of the acquisition of digital (or virtual) specimens and artefacts through non-contact measurement techniques, such as laser scanning, photogrammetry and medical imaging techniques (Davies et al., 2017; Falkingham et al., 2018; Ruiz et al., 2021). Such techniques can create a large collection of digital models to access, reuse and analyse the archaeological and anthropological materials flexibly and easily, thereby revolutionising conventional studies. One of particular interest is virtual anthropology (or virtual morphology) that provides innovative tools as well as software and methods to quantify shape and size variations and study morphology (Recheis et al., 1999; Weber, 2015). Generally, it includes six main areas: exposure of internal structures, comparison of shapes and forms, reconstruction, materialization of digital specimens, and sharing data (Weber, 2015). Often the related concepts, techniques and tools can also contribute to the research and application in an archaeological context when artefacts are used instead of fossil specimens (Okumura & Araujo, 2019; Weber, 2014).

Digital resources, tools and quantitative methods facilitate academic research and practical applications in both Archaeology and Anthropology, thereby promoting an understanding of the materials. However, owing to the diversity and complexity of data and research questions, generalised computational methods may not work well. Inappropriate methods probably reduce the scientific rigour and might hinder understanding of the analytical results. Thus, innovative scientific techniques and

methods must be developed and applied to discover new knowledge and get new research findings. This chapter first introduces the general background of the most commonly used methods and some of their applications in Anthropology and Archaeology, and then introduces the research aims and questions, the structure of the thesis and its scientific contribution.

1.1 General background and problem statement

1.1.1 Geometric morphometrics

Landmarks refer to discrete points which are in some sense equivalent among biological structures, taking into account anatomical, developmental, biomechanical or evolutionary knowledge (Bookstein, 1991). Using landmark coordinates reflecting the shape and form (shape plus size) of physical specimens, geometric morphometric methods are regularly employed to quantify shape and size variations and visually interpret morphological variations, revolutionising conventional morphometrical studies (Adams et al., 2004; Bastir et al., 2019; Mitteroecker & Gunz, 2009; O'Higgins, 2000; Theska et al., 2020; Viscosi & Cardini, 2011). However, few homologous landmarks can be identified from biological structures, especially in the smooth regions, such as the cranium vault and tooth crown. To address this problem, several equivalent point correspondences (called semilandmarks) are generated to describe the detailed and rich information. It is important to note that the establishment of semilandmarks relies primarily on mathematical mapping algorithms based on topographic features rather than developmental or evolutionary equivalences. In practice, sliding approaches that minimise the bending energy of Thin plate-splines (TPS) or Procrustes distance, a measure of the absolute magnitude of shape difference, are carried out to yield sliding semilandmarks over curves or surface regions between landmarks (Gunz & Mitteroecker, 2013; Perez et al., 2006). Recently, a practical guide and a range of solutions have been provided to help researchers place sliding semilandmarks during data collection (Bardua et al., 2019). It should be noted that landmarks are reliable and respect homologies as far as they can be identified, while semilandmarks may add uncertainties or errors in subsequent morphometric analysis (Cardini, 2020).

After obtaining the digital models of specimens, the first step of a geometric morphometric analysis is to select and design a landmark (and semilandmark) configuration with regard to the research question at hand (Oxnard & O'Higgins, 2009). Because landmarks are frequently scant and cannot capture rich shape information, high-density semilandmarks are often used to promote the quantification of morphological shape (Bardua et al., 2019). Subsequently, General Procrustes Analysis (GPA) is applied to the landmark and semilandmark configurations among specimens, removing the non-shape variation (e.g. location, orientation and scaling) (Mitteroecker et al., 2013). After completing the superimposition, Procrustes shape coordinates of every specimen are

projected into the Kendall's shape space, which is a multidimensional Riemannian space with points in the space representing the shapes of specimens (Mitteroecker & Gunz, 2009). Afterwards, Procrustes distance matrices can be analysed to assess the shape differences. Next, multivariate statistical methods, such as principal component analysis (PCA), two-block partial least squares analysis (PLS), multivariate regression, multivariate analysis of variance (MANOVA), and clustering, are used to perform statistical analysis with regard to the research question (Mitteroecker & Schaefer, 2022). Finally, TPS transformation grids or surface warps facilitate visual interpretation of the patterns of shape variation of mean and allometrically scaled shapes and quantifying the shape changes (Bookstein, 1989; Klingenberg, 2013). Alternatively, other visualisation techniques (e.g. colour maps between the reference and target surfaces, lollipops, etc.) are implemented to understand the underlying shape variations. It is worth mentioning that the locations of semilandmarks cannot be singly interpreted (Bastir et al., 2019; Mitteroecker & Schaefer, 2022; Oxnard & O'Higgins, 2009).

Archaeological artefacts as a living witness of significant cultures and human values are an important form of tangible cultural heritage. They provide reliable historic evidence for examining the past organization of human societies and historical manufacturing capacities (McNabb et al., 2018; Saragusti et al., 1998). For example, artefact tool typologies facilitate understanding of human and early human cognitive development, and cultural transmission processes (Herzlinger et al., 2017). In similar applications in biology and anthropology, several previous studies have often employed the concept and general process of geometric morphometrics to assess archaeological artefact groups (Thulman, 2012) and quantify morphological difference between tools (Hashemi et al., 2021; Lycett et al., 2010). However, the definition and design of landmark and semilandmarks are problematic because the concept of developmental or evolutionary homology does not apply to artefacts, unlike landmarks among biological fossil specimens (Okumura & Araujo, 2019). Often equally spaced semilandmarks are placed over artefacts to capture geometric information (García-Medrano et al., 2020; Herzlinger & Grosman, 2018).

Geometric morphometrics is a set of powerful visual statistical tools to quantify shape variations, regularly applied among human fossils (Cui & Wu, 2015; Mori et al., 2020; Torres-Tamayo et al., 2020) and artefacts (Archer et al., 2018; Serwatka, 2015). Several software packages are available, such as morphologika (O'Higgins & Jones, 1998), IDAV Landmark Editor (Wiley et al., 2005), EVAN Toolbox (www.evan-society.org), MorphoJ (Klingenberg, 2011), the Thin-plate spline (TPS) series (Rohlf, 2015), SlicerMorph (Rolfe et al., 2021), and R packages, e.g. geomorph and Morpho (Adams & Otárola-Castillo, 2013; Schlager, 2017) to carry out geometric morphometric analyses. However, the definition of an appropriate landmark and semilandmark

configuration with regard to the research question is the prerequisite step. It is an important task to examine the effects of using different semilandmarking approaches and densities of semilandmarks on morphometric analyses and visualisations. Understanding these effects can assist researchers to design reliable and appropriate configurations with regard to a specific question.

1.1.2 Different applications

The geometric morphometric toolkit has wide applications in Anthropology and Archaeology (Archer et al., 2018; Bastir et al., 2019; Mitteroecker & Schaefer, 2022). Examples include asymmetry surface detection (Melchionna et al., 2021), virtual restoration of damaged fossils and artefacts (Gunz et al., 2009; Lautenschlager, 2016), and examination of the relationships between hard and soft tissues in conjunction with facial approximation (Guyomarc'h et al., 2014; Malá et al., 2018).

1.1.2.1 Asymmetry detection

Bilaterally symmetrical objects represent a large and important proportion of archaeological artefacts and biological objects. Asymmetry detection provides considerable guidance and reliable evidence to allow experts to examine historical manufacturing capacities (McNabb et al., 2018), repair damaged artefacts (Jo et al., 2020; Profico et al., 2019), evaluate human developmental variability and instability (Hou & Fagan, 2021), and perceive the evolution of human cognition (Wynn, 2002). Geometric morphometrics is the preferred way of quantifying asymmetry among a collection of samples (McGrath et al., 2022; Neubauer et al., 2020). In this process, GPA and PCA are applied to symmetric components comprising the original and mirrored landmark configurations, and asymmetric components including differences of the original configuration from the symmetric averages, separately. Then, asymmetry can be detected and visualised using geometric morphometric and multivariate statistical analyses.

To quantify surface asymmetry of every specimen or artefact, the current method involves the comparison of geometric differences between the original and registered mirrored landmark configurations with regard to the plane of symmetry determined by the landmark configuration (Damstra et al., 2012). However, the locations of landmarks are primarily dependent upon the experts' knowledge and visual perception, especially in the featureless and smooth regions. Thus, placement errors are inevitable and the choice of landmark configuration affects the subsequent results. To avoid landmark placements, a landmark-free morphometric approach is proposed to identify the plane of symmetry based on the mid-points of a full set of vertices between original and aligned mirrored meshes (Di Angelo et al., 2019; Noori et al., 2020). Then, the gross asymmetry is assessed by computing the Euclidean distance between the original and registered surfaces and the

colour map of distance of every vertex is used to recognise the regional asymmetry (Claes et al., 2011). However, the existing morphometric approach is influenced by missing data and shape distortion. New methods and tools are desirable to deal with these issues and to promote accuracy.

1.1.2.2 Virtual restoration

Human fossils are often discovered broken and exhibit missing anatomy, hence virtual restoration is an important and inevitable step towards subsequent analysis, e.g. to compare morphology (Davis et al., 2021) and in facial approximation of the dry skull (Hayes et al., 2013). Additionally, virtual restoration methods can facilitate computation of missing structures of the damaged artefacts and in producing physical models via 3D printing, avoiding physical intervention and secondary damage. Several factors, e.g. biological constraints and morphological integration (Gunz et al., 2009), require consideration when dealing with biological and anthropological specimens. Similarly, when recreating the geometry and texture of damaged artefacts, various aspects, such as geometrical constraints, decorative and cultural information, and artistic style, should be considered to ensure the accuracy of the restoration (Jo et al., 2020; Lanitis et al., 2012).

In a situation where the damaged object exhibits bilateral symmetry and more than half of each object is perfectly preserved, reflection is a straightforward and reliable means for restoration through the prediction of the locations of landmarks or point correspondences over the missing regions. Subsequently, the non-rigid registration approaches, e.g. TPS, can be carried out to warp the reflection of an intact region to a damaged region, thereby generating the missing geometry surface (Gunz et al., 2009; O'Higgins et al., 2019). Notably, the prediction of landmark and semilandmark locations, and the choice of deformation (or registration) approaches will affect the ultimate results. Additionally, when both sides of e.g. a fossil are incomplete, a surface from a well-preserved specimen is often used as a template; it is commonly warped to the damaged specimen to achieve virtual restoration (Amano et al., 2015; Benazzi et al., 2009). However, the choice of the template will affect the restoration result, because the morphology of the deformed template is used to complete the damaged specimen.

Recently, statistical shape models (SSM) (Ebert et al., 2022; Schlager & Rüdell, 2017) have been shown to be a promising way to reconstruct the complete model from a partial fossil. They incorporate prior knowledge from within the same group to capture shape variability (Brunton et al., 2014). Generally, PCA is applied to the aligned dense point correspondences among specimens to generate the SSM, and hence shape variability can be parametrized by the mean surface and principal components (PCs). Subsequently, the damaged fossil is rigidly fit to the SSM, and the parameters of the SSM are optimized to allow the reconstructed surface derived from the SSM to match the damaged fossil. When there are great fitting errors, the reconstructed surface is further warped to the damaged

fossils using TPS (Fuessinger et al., 2018). This method is said to provide a more reliable reconstruction of fossil specimens with missing geometry, however it requires a suitable sample of specimens.

1.1.2.3 Facial approximation

Individual human faces are unique due to the differences in bony morphology, facial soft tissues, and locations, sizes and shapes of facial features (e.g., eyes, nose, mouths, ears) (Richmond et al., 2018). Facial approximation (sometimes known as facial reconstruction) aims to recreate a probable soft tissue face based on a dry skull alone. It is a last resort for human identification of a seriously decomposed cadaver when there are no other clues in forensic science. It also has provided new insights, facilitating the visual depiction of archaic humans and *Homo sapiens* (Baldasso et al., 2021; Guyomarc'h et al., 2018; Hamre et al., 2017; Hayes, 2016; Lee et al., 2014). Over the past century, there are two different viewpoints about the reliability of facial approximation (Stephan, 2015; Wilkinson, 2005). Some researchers argue that the detailed facial surface can be recreated accurately based on skull morphology. They suggest failed prediction is attributable to lack of use of detailed skull morphology and technical knowledge in modelling. By contrast, others point out there are varying degrees of facial approximation errors and uncertainties, and hence multiple faces of the same person can be recreated with similar degrees of confidence.

The conventional manual method was first proposed by anthropologists and artists who sculptured clays over a replica of the dry skull based primarily upon facial soft tissue thickness depths (FSTDs) at a limited number of landmarks and the predicted muscle structures (Wilkinson, 2010). Therefore, a subjective and artistic interpretation is inevitable, resulting in potentially biased and inconsistent approximations (Campbell et al., 2021; Stephan, 2015). This process is very time consuming and requires sculpturing skill. To address these limitations, interactive virtual sculpturing (Wilkinson et al., 2006) and automated computerised methods based on algorithms (Claes et al., 2010; Stephan et al., 2019a) have been developed and applied to the dry skull. These computerized methods are faster, more objective and efficient. The quantification of craniofacial (anatomical modelling) relationships between hard and soft tissues is the fundamental basis of facial approximation. Average FSTDs (Stephan et al., 2019b) or regressed FSTDs (De Greef et al., 2009; Guyomarc'h et al., 2013) at landmarks associated with personal information (e.g. sex and age) is used to predict the facial points of the approximated result. Additionally, geometric morphometric methods are used to explore the craniofacial relationships between facial features and bony structures (Guyomarc'h et al., 2014; Kustár et al., 2013; Ridet et al., 2020) and relationships between the facial profile and bony structures (Malá et al., 2018). These learnt relationships can further be used to recreate facial features of the dry skull, thereby improving approximation accuracy.

Many studies have assessed performance in facial approximation through a quantitative assessment of similarity between the approximated and actual faces or a recognition rate tested by a face pool (Claes et al., 2010; Gietzen et al., 2019; Lee et al., 2012). However, there is considerable controversy regarding approximated facial features (Guyomarc'h et al., 2014; Wilkinson, 2010). Therefore, a degree of artistic interpretation and modelling is required, especially in archaeological applications. New methods need to be developed to quantify craniofacial relationships and perform facial approximation, minimising the approximation errors and subjective interpretation.

1.2 Motivation and objectives

This study is mainly motivated by my own experience in interdisciplinary collaborations, which illustrated the need to implement geometric morphometric approaches in archaeological contexts. Given recent progress in the accessibility of 3D archaeological and anthropological datasets, it is common for archaeologists and anthropologists to collaborate with computer scientists to undertake research using statistical analysis and visualisation techniques. However, each side in such collaborations has their own knowledge, preference and experience, and hence there have sometimes been some controversial suggestions and concerns in addressing the same questions. For example, if it is of interest to place landmarks and semilandmarks, computer scientists are concerned with the speed, computation cost, geometric accuracy, and the degree of human intervention, however, they always ignore biological concepts of homology, which are important in studies of variation, evolution and development. Thus, they need to know the principles and methods of anthropology and archaeology, while anthropologists require understanding of limitations of the computerised methods (Greener et al., 2022). As a computer scientist working in an archaeology department, I have attempted to bridge the gaps in interdisciplinary study and explore advanced and novel algorithms and methods for paleoanthropological and archaeological data.

Owing to the diversity and complexity of 3D data research questions, existing computational methods may not be directly applicable. Novel methods are necessary to provide a better interpretation of human fossils and artefacts, and to offer an opportunity for virtual simulation of the past. This thesis aims to advance methods by pursuing two objectives. These are as follows:

- Assess the extent to which morphometric analyses and visual descriptions of shapes of interest generated by different semilandmarking approaches and densities of semilandmarks differ and how they differ, allowing archaeologists and anthropologists to carefully consider whether semilandmarks are necessary to answer the question at hand and balance this need against the statistical and biological downsides.

- Develop a complete workflow and novel computerised method for each special application, including semilandmarking, asymmetry detection, virtual restoration, and facial approximation, thereby minimising human intervention and enhancing the performance.

1.3 Research questions

The main research questions address topics in anthropological and archaeological research using statistical analysis and visualisation methods. These questions may broadly be formulated as follows:

- 1) Semilandmarking methods and densities in geometric morphometrics
 - To what extent do different semilandmarking methods and densities lead to different morphometric results, concerning semilandmark locations, mean landmark and semilandmark configurations, centroid sizes, Procrustes distance matrices, PCs of shape variation, and allometrically scaled shapes? (Chapter 2)
 - To what extent do the visualisations of mean and allometrically scaled shapes differ when using different semilandmarking approaches and densities of semilandmarks? (Chapter 3)
 - To what extent does the use of landmarks and semilandmarks affect morphometric results and visualisation in comparison to those generated by using landmarks alone? (Chapters 2 and 3)
- 2) Asymmetry detection and virtual restoration of bilaterally symmetrical objects
 - Can a landmark-free method identify a suitable plane of symmetry and quantify the surface asymmetry from objects with little geometric distortion or simple missing geometry? If so, are there differences between the planes of symmetry and asymmetry generated by landmark-free and landmark-based morphometric methods? (Chapter 4)
 - Can computerised methods generate a digital model of missing geometry that tightly matches the boundary curves of missing geometry, thereby supporting virtual and physical restoration? (Chapter 5)
- 3) Facial approximation of archaic humans and *Homo sapiens*
 - Can average dense FSTDs be used to quantify overall craniofacial relationships in order to achieve reliable facial approximation (Chapters 6 and 7)?
 - To what extent do skull morphology and the choice of FSTDs affect the approximated faces (Chapter 6)?
 - To what extent do nasal (and oral) hard tissues affect soft tissue shapes? Can high-density semilandmarks be used to improve the accuracy of the approximated

nose (and mouth) soft tissue shapes (Chapter 7)?

1.4 Structure of thesis

This thesis is composed of six individual articles (Chapters 2-7), supplemented by the introductory Chapter 1 and concluding Chapter 8. Each article covers a theme within the overall topic. Each is intended for separate journal publication and can be read independently or as part of the whole. From a theoretical perspective, Chapters 2 and 3 are related to the geometric morphometric methods. Chapter 2 has been submitted to a special issue “Geometric Morphometrics Applied to Biological Structures” of *Animals* and Chapter 3 has been published in *Animals*. Based on the fundamental principles of geometric morphometrics, Chapters 4-7 give different applications. Chapters 4 and 5 seek to quantify asymmetry detection and achieve virtual restoration of artefacts and fossils exhibiting near bilateral symmetry. Chapter 4 has been accepted by *ACM Journal on Computing and Cultural Heritage* and Chapter 5 has been published in *Journal of Cultural Heritage*. Chapters 6 and 7 aim to recreate probable facial appearances of archaic humans and *Homo sapiens*, respectively. Chapter 6 has been published in *Archaeological and Anthropological Sciences* and Chapter 7 has been submitted to *Journal of Anatomy*. Because different journals have different formats, each article has been re-formatted according to a standard style to provide a coherent thesis. The main revision is to heading numbering and reference formatting.

The study has been approved by the Ethics Review Committee of Department of Archaeology, University of York (20 April 2021).

1.4.1 Chapter 2

Different semilandmarking approaches and densities of semilandmarks result in different locations of semilandmarks, therefore they can be expected to yield different analytical results. The article presented in Chapter 2 assesses the performance of three landmark driven semilandmarking approaches and different densities of semilandmarks using two different datasets (adult human head and ape cranial surfaces) with different degrees of variation and complexity. Six null hypotheses are tested, of no difference in semilandmark locations, mean landmark and semilandmark configurations, centroid sizes, Procrustes distance matrices, PCs of shape variation, and allometrically scaled shape. Additionally, this article examines the effect of semilandmarks on morphometric analyses by computing the correlations between the results calculated using the landmarks alone and those based on the landmarks and semilandmarks. These analytical results facilitate assessment of the performance of different semilandmarking approaches and densities.

This article has been submitted as follows:

Wuyang Shui, Antonio Profico and Paul O’Higgins. A comparison of semilandmarking approaches in the analysis of size and shape. *Animals*. 2023. (under

review)

Author contributions: Conceptualization, Shui, O’Higgins and Profico; methodology, Shui, O’Higgins and Profico; software, Shui, Profico; validation, Shui and O’Higgins; formal analysis, Shui; investigation, Shui, O’Higgins and Profico; resources, Shui, O’Higgins and Profico; data curation, Shui, O’Higgins and Profico; writing—original draft preparation, Shui; writing—review and editing, O’Higgins, Shui and Profico; visualisation, Shui; supervision, O’Higgins; project administration, O’Higgins, Shui.

1.4.2 Chapter 3

Different semilandmarking approaches can yield different semilandmark locations, thereby generating different estimates of mean and allometrically scaled shapes by warping the template surface. It is important to note that semilandmarks cannot be singly interpreted, but rather be analysed as a whole. The article presented in Chapter 3 assesses how different semilandmarking approaches and densities of semilandmarks affect the visualisations of the estimates of mean and allometrically scaled shapes using two different datasets (adult human head and ape cranial surfaces) with different degrees of variation and complexity datasets. Additionally, this article examines the effect of the choice of template surface and semilandmarks on the visualisation, respectively. These visualisation results facilitate assessment of the performance of different semilandmarking approaches and determining whether semilandmarks are necessary for the visualisation.

This article has been published as follows:

Wuyang Shui, Antonio Profico and Paul O’Higgins. A comparison of semilandmarking approaches in the visualisation of shape differences. *Animals*. 2023, 13(3): 385.

Author contributions: Conceptualization, Shui, O’Higgins and Profico; methodology, Shui, O’Higgins and Profico; software, Shui, Profico; validation, Shui and O’Higgins; formal analysis, Shui; investigation, Shui, O’Higgins and Profico; resources, Shui, O’Higgins and Profico; data curation, Shui, O’Higgins and Profico; writing—original draft preparation, Shui; writing—review and editing, O’Higgins, Shui and Profico; visualisation, Shui; supervision, O’Higgins; project administration, O’Higgins, Shui.

1.4.3 Chapter 4

Landmark-based and landmark-free morphometric methods for detecting planes of symmetry have both been proposed. However, the landmark-based approach requires manual identification of landmark locations (time consuming and prone to error) and landmark independent morphometric method is always influenced by missing data. The article presented in Chapter 4 explores a novel landmark-free approach to identify the

best-fitted plane of symmetry from nearly bilaterally symmetrical objects by finding the plane with the minimum geometric differences between the original and mirrored meshes. The potentially symmetrical regions are extracted through the examination of the extent of asymmetry and plane clustering, and hence this method is robust to small geometric distortion and simple missing geometry. Based on the recognised plane, this article quantifies gross and regional surface asymmetry and produces the profile drawings. The proposed method is evaluated and applied to both synthetic and real anthropological (e.g. skull, cranium, endocast and femur) and archaeological (terracotta warriors head, bronze mask, ceramic plate and pottery spoon) data.

This article has been accepted as follows:

Wuyang Shui, Pianpian Wei, Xia Zheng, Shengling Geng. A landmark-free approach for surface asymmetry detection and profile drawings from bilaterally symmetrical geometry. *ACM Journal on Computing and Cultural Heritage*. 2022. (accepted)

Author contributions: Conceptualization, Shui, Wei, Zheng and Geng; methodology, Shui; software, Shui; validation, Shui; formal analysis, Shui; investigation, Shui, Wei, Zheng and Geng; resources, Shui, Wei, Zheng and Geng; data curation, Shui, Wei, Zheng and Geng; writing—original draft preparation, Shui; writing—review and editing, Shui, Wei, Zheng and Geng; visualisation, Shui; project administration, Shui.

1.4.4 Chapter 5

Symmetry may be useful in reconstructing missing geometry by deforming the intact side to the damaged area. However, it remains challenging to improve the quality of deformation when artefacts are heavily corroded and have irregular boundaries. The article presented in Chapter 5 develops a non-rigid deformation approach to allow the reflection of the intact region to match the damaged artefact, thereby generating a digital model of missing geometry. This model can be converted to a physical one to achieve physical restoration of artefacts via 3D printing technique. Additionally, this article employs geometric morphometrics and anthropometric measurements to capture characteristic features based on landmarks. The proposed method is evaluated and applied to a metal mask of the Liao dynasty from the Palace Museum, China.

This article has been published as follows:

Wuyang Shui and Fei Gao. A geometric completion and shape analysis method for damaged bilaterally symmetrical artefacts. *Journal of Cultural Heritage*. 2021, 52, 118-127.

Author contributions: Conceptualization, Shui and Gao; methodology, Shui; software, Shui; validation, Shui; formal analysis, Shui and Gao; investigation, Shui and Gao; resources, Shui and Gao; data curation, Shui and Gao; writing—original draft preparation,

Shui; writing—review and editing, Shui and Gao; visualisation, Shui; project administration, Shui.

1.4.5 Chapter 6

The conventional manual facial approximation method is often used to recreate a probable face of archaic humans. However, this process consumes time and the approximate results rely primarily upon the practitioner's experience and interpretation. The article presented in Chapter 6 develops a computerized facial approximation method for archaic humans based on the assumption that the distributions of FSTDs of modern living humans are similar to those of archaic humans. The deformation method is used to fit the warped template face to the coarsely facial appearance recreated by assigning the average dense FSTDs to the dry skull, thereby filling in the missing geometry and adding facial features. Additionally, this article assesses the extent to which skull morphology and the choice of FSTDs affect the approximated faces and then employs the geometric morphometric methods to examine shape variations among modern human faces and approximated faces. The difference in the distributions of FSTDs is used to evaluate the proposed method. Finally, the proposed method is applied to the Jinniushan 1 archaic human skull which is one of the most important fossils of the Middle Pleistocene from China, dating back to approximately 260,000 BP.

This article has been published as follows:

Wuyang Shui, Yameng Zhang, Xiujie Wu and Minquan Zhou. A computerized facial approximation method for archaic humans based on dense facial soft tissue thickness depths. *Archaeological and Anthropological Science*. 2021, 13:186.

Author contributions: Conceptualization, Shui, Wu, Zhang and Zhou; methodology, Shui; software, Shui; validation, Shui; formal analysis, Shui and Zhang; investigation, Shui, Zhang, Wu and Zhou; resources, Shui, Zhang, Wu and Zhou; data curation, Shui, Zhang, Wu and Zhou; writing—original draft preparation, Shui; writing—review and editing, Shui, Zhang, Wu and Zhou; visualisation, Shui; project administration, Shui.

1.4.6 Chapter 7

The article presented in Chapter 7 quantifies the covariations of nasal (and oral) hard and soft tissue morphologies and presents a computerised method facial approximation method for *Homo sapiens*. The facial statistical shape model is used to fit the coarsely approximated face recreated by assigning the average dense FSTDs to the dry skull and the predicted nose and mouth soft tissues, thereby filling in the missing geometry and adding facial features. Resemblance comparison and recognition rate are used to test the accuracy of the proposed method. Finally, the proposed method is applied to approximate the facial appearance of the Upper Cave (UC) 101 skull, an important 30,000-year-old human skull excavated from the village of Zhoukoudian in northern China.

This article has been submitted as follows:

Wuyang Shui, Xiujie Wu and Minquan Zhou. A computerized facial approximation method for *Homo sapiens* based on facial soft tissue thickness depths and geometric morphometrics. *Journal of Anatomy*. (under review)

Author contributions: Conceptualization, Shui, Wu and Zhou; methodology, Shui; software, Shui; validation, Shui; formal analysis, Shui; investigation, Shui, Wu and Zhou; resources, Shui, Wu and Zhou; data curation, Shui, Wu and Zhou; writing—original draft preparation, Shui; writing—review and editing, Shui, Wu and Zhou; visualisation, Shui; project administration, Shui.

1.4.7 Chapter 8

Finally, Chapter 8 provides conclusions and ideas for future research.

1.5 References

- Adams, D, Rohlf, J, & Slice, D. (2004). Geometric morphometrics: ten years of progress following the 'revolution'. *Italian Journal of Zoology*, 71(1), 5-16.
- Adams, DC, & Otárola-Castillo, E. (2013). geomorph: an R package for the collection and analysis of geometric morphometric shape data. *Methods in Ecology and Evolution*, 4(4), 393-399.
- Amano, H, Kikuchi, T, Morita, Y, Kondo, O, Suzuki, H, Ponce de León, MS, Zollikofer, CP, et al. (2015). Virtual reconstruction of the Neanderthal Amud 1 cranium. *American Journal of Physical Anthropology*, 158(2), 185-197.
- Archer, W, Pop, CM, Rezek, Z, Schlager, S, Lin, SC, Weiss, M, Dogandžić, T, et al. (2018). A geometric morphometric relationship predicts stone flake shape and size variability. *Archaeological Anthropological Sciences*, 10(8), 1991-2003.
- Baldasso, RP, Moraes, C, Gallardo, E, Stumvoll, MB, Crespo, KC, Strapasson, RAP, & de Oliveira, RN. (2021). 3D forensic facial approximation: Implementation protocol in a forensic activity. *Journal of Forensic Sciences*, 66(1), 383-388.
- Bardua, C, Felice, R, Watanabe, A, Fabre, A, & Goswami, A. (2019). A practical guide to sliding and surface semilandmarks in morphometric analyses. *Integrative Organismal Biology*, 1(1), obz016.
- Bastir, M, García-Martínez, D, Torres-Tamayo, N, Palancar, CA, Fernández-Pérez, FJ, Riesco-López, A, Osborne-Márquez, P, et al. (2019). Workflows in a Virtual Morphology Lab: 3D scanning, measuring, and printing. *Journal of Anthropological Sciences*, 97, 107-134.
- Benazzi, S, Stansfield, E, Milani, C, & Gruppioni, G. (2009). Geometric morphometric methods for three-dimensional virtual reconstruction of a fragmented cranium: the case of Angelo Poliziano. *International journal of legal medicine*, 123(4), 333-344.
- Bookstein, FL. (1989). Principal warps: Thin-plate splines and the decomposition of deformations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(6), 567-585.
- Bookstein, FL. (1991). *Morphometric tools for landmark data: geometry and biology*: Cambridge University Press.
- Brooks, AS, Yellen, JE, Potts, R, Behrensmeier, AK, Deino, AL, Leslie, DE, Ambrose, SH, et al. (2018). Long-distance stone transport and pigment use in the earliest Middle Stone Age. *Science*, 360(6384), 90-94.
- Brunton, A, Salazar, A, Bolkart, T, & Wuhler, S. (2014). Review of Statistical Shape Spaces for 3D Data with Comparative Analysis for Human Faces. *Computer Vision and Image Understanding*, 128(11), 1-17.

- Campbell, RM, Vinas, G, Henneberg, M, Diogo, R, & Evolution. (2021). Visual depictions of our evolutionary past: A broad case study concerning the need for quantitative methods of soft tissue reconstruction and art-science collaborations. *Frontiers in Ecology and Evolution*, 9, 639048.
- Cardini, A. (2020). Less tautology, more biology? A comment on “high-density” morphometrics. *Zoomorphology*, 139(4), 513-529.
- Claes, P, Vandermeulen, D, De Greef, S, Willems, G, Clement, JG, & Suetens, P. (2010). Computerized craniofacial reconstruction: Conceptual framework and review. *Forensic Science International*, 201(1-3), 138-145.
- Claes, P, Walters, M, Vandermeulen, D, & Clement, G. (2011). Spatially-dense 3D facial asymmetry assessment in both typical and disordered growth. *Journal of anatomy*, 219(4), 444-455.
- Cui, Y, & Wu, X. (2015). A geometric morphometric study of a Middle Pleistocene cranium from Hexian, China. *Journal of Human Evolution*, 88, 54-69.
- Damstra, J, Fourie, Z, De Wit, M, & Ren, Y. (2012). A three-dimensional comparison of a morphometric and conventional cephalometric midsagittal planes for craniofacial asymmetry. *Clinical Oral Investigations*, 16(1), 285-294.
- Davies, TG, Rahman, IA, Lautenschlager, S, Cunningham, JA, Asher, RJ, Barrett, PM, Bates, KT, et al. (2017). Open data and digital morphology. *Proceedings of the Royal Society B: Biological Sciences*, 284(1852), 20170194.
- Davis, CA, Profico, A, & Kappelman, J. (2021). Digital restoration of the Wilson-Leonard 2 Paleoindian skull (~ 10,000 BP) from central Texas with comparison to other early American and modern crania. *American Journal of Physical Anthropology*, 176(3), 486-503.
- 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(2), 60-65.
- Di Angelo, L, Di Stefano, P, Governi, L, Marzola, A, & Volpe, Y. (2019). A robust and automatic method for the best symmetry plane detection of craniofacial skeletons. *Symmetry*, 11(2), 245.
- Dolbunova, E, Lucquin, A, McLaughlin, TR, Bondetti, M, Courel, B, Oras, E, Piezonka, H, et al. (2022). The transmission of pottery technology among prehistoric European hunter-gatherers. *Nature Human Behaviour*, 1-13.
- Ebert, LC, Rahbani, D, Lüthi, M, Thali, MJ, Christensen, AM, & Fliss, B. (2022). Reconstruction of full femora from partial bone fragments for anthropological analyses using statistical shape modeling. *Forensic Science International*, 332, 111196.

Falkingham, PL, Bates, KT, Avanzini, M, Bennett, M, Bordy, EM, Breithaupt, BH, Castanera, D, et al. (2018). A standard protocol for documenting modern and fossil ichnological data. *Palaeontology*, 61(4), 469-480.

Fuessinger, MA, Schwarz, S, Cornelius, CP, Metzger, MC, Ellis, E, Probst, F, Semper-Hogg, W, et al. (2018). Planning of skull reconstruction based on a statistical shape model combined with geometric morphometrics. *International Journal of Computer Assisted Radiology Surgery*, 13(4), 519-529.

García-Medrano, P, Ashton, N, Moncel, MH, & Ollé, A. (2020). The WEAP method: A new age in the analysis of the Acheulean Handaxes. *Journal of Paleolithic Archaeology*, 3(4), 756-793.

Gietzen, T, Brylka, R, Achenbach, J, zum Hebel, K, Schömer, E, Botsch, M, Schwanecke, U, et al. (2019). A method for automatic forensic facial reconstruction based on dense statistics of soft tissue thickness. *Plos one*, 14(1), e0210257.

Greener, JG, Kandathil, SM, Moffat, L, & Jones, DT. (2022). A guide to machine learning for biologists. *Nature Reviews Molecular Cell Biology*, 23, 40-55.

Gunz, P, & Mitteroecker, P. (2013). Semilandmarks: a method for quantifying curves and surfaces. *Hystrix, the Italian Journal of Mammalogy*, 24(1), 103-109.

Gunz, P, Mitteroecker, P, Neubauer, S, Weber, GW, & Bookstein, FL. (2009). Principles for the virtual reconstruction of hominin crania. *Journal of Human Evolution*, 57(1), 48-62.

Guyomarc'h, P, Dutailly, B, Charton, J, Santos, F, Desbarats, P, & Coqueugniot, H. (2014). Anthropological Facial Approximation in Three Dimensions (AFA3D): Computer-Assisted Estimation of the Facial Morphology Using Geometric Morphometrics. *Journal of Forensic Sciences*, 59(6), 1502-1516.

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(1), 411. e411-411. e410.

Guyomarc'h, P, Velemínský, P, Brůžek, J, Lynnerup, N, Horák, M, Kučera, J, Rasmussen, KL, et al. (2018). Facial approximation of Tycho Brahe's partial skull based on estimated data with TIVMI-AFA3D. *Forensic science international*, 292, 131-137.

Hamre, SS, Erstrand, GA, Daux, V, Parson, W, & Wilkinson, C. (2017). Three individuals, three stories, three burials from medieval Trondheim, Norway. *Plos one*, 12(7), e0180277.

Harvati, K, Stringer, C, & Karkanas, P. (2011). Multivariate analysis and classification of the Apidima 2 cranium from Mani, Southern Greece. *Journal of Human Evolution*, 60(2), 246.

Hashemi, SM, Nasab, HV, Berillon, G, & Oryat, M. (2021). An investigation of the flake-

based lithic tool morphology using 3D geometric morphometrics: A case study from the Mirak Paleolithic Site, Iran. *Journal of Archaeological Science: Reports*, 37, 102948.

Hayes, S. (2016). Faces in the museum: revising the methods of facial reconstructions. *Museum Management and Curatorship*, 31(3), 218-245.

Hayes, S, Sutikna, T, & Morwood, M. (2013). Faces of Homo floresiensis (LB1). *Journal of Archaeological Science*, 40(12), 4400-4410.

Hershkovitz, I, May, H, Sarig, R, Pokhojaev, A, Grimaud-Hervé, D, Bruner, E, Fornai, C, et al. (2021). A middle pleistocene homo from Nesher Ramla, Israel. 372(6549), 1424-1428.

Herzlinger, G, Goren-Inbar, N, & Grosman, L. (2017). A new method for 3D geometric morphometric shape analysis: The case study of handaxe knapping skill. *Journal of Archaeological Science: Reports*, 14, 163-173.

Herzlinger, G, & Grosman, L. (2018). AGMT3-D: A software for 3-D landmarks-based geometric morphometric shape analysis of archaeological artifacts. *Plos One*, 13(11), e0207890.

Hou, M, & Fagan, M. (2021). Assessments of bilateral asymmetry with application in human skull analysis. *Plos one*, 16(10), e0258146.

Jo, YH, Hong, S, Jo, SY, & Kwon, YM. (2020). Noncontact restoration of missing parts of stone Buddha statue based on three-dimensional virtual modeling and assembly simulation. *Heritage Science*, 8, 103.

Klingenberg, CP. (2011). MorphoJ: an integrated software package for geometric morphometrics. *Molecular Ecology Resources*, 11(2), 353-357.

Klingenberg, CP. (2013). Visualizations in geometric morphometrics: how to read and how to make graphs showing shape changes. *Hystrix, the Italian Journal of Mammalogy*, 24(1), 15-24.

Kustár, A, Forro, L, Kalina, I, Fazekas, F, Honti, S, Makra, S, & Friess, M. (2013). FACE-R—a 3D database of 400 living individuals' full head CT- and face scans and preliminary GMM analysis for craniofacial reconstruction. *Journal of Forensic Sciences*, 58(6), 1420-1428.

Lacruz, RS, Stringer, CB, Kimbel, WH, Wood, B, Harvati, K, O'Higgins, P, Bromage, TG, et al. (2019). The evolutionary history of the human face. *Nature ecology & evolution*, 3(5), 726-736.

Lanitis, A, Stylianou, G, & Voutounos, C. (2012). Virtual restoration of faces appearing in byzantine icons. *Journal of Cultural Heritage*, 13(4), 404-412.

Lautenschlager, S. (2016). Reconstructing the past: methods and techniques for the digital

restoration of fossils. *Royal Society Open Science*, 3(10), 160342.

Lee, WJ, Wilkinson, CM, & Hwang, HS. (2012). An accuracy assessment of forensic computerized facial reconstruction employing cone-beam computed tomography from live subjects. *Journal of Forensic Sciences*, 57(2), 318-327.

Lee, WJ, Yoon, AY, Mi, KS, Wilkinson, CM, & Dong, HS. (2014). The archaeological contribution of forensic craniofacial reconstruction to a portrait drawing of a Korean historical figure. *Journal of Archaeological Science*, 49, 228-236.

Lycett, SJ, von Cramon-Taubadel, N, & Gowlett, JA. (2010). A comparative 3D geometric morphometric analysis of Victoria West cores: implications for the origins of Levallois technology. *Journal of Archaeological Science*, 37(5), 1110-1117.

Malá, PZ, Krajíček, V, & Velemínská, J. (2018). How tight is the relationship between the skeletal and soft-tissue facial profile: A geometric morphometric analysis of the facial outline. *Forensic Science International*, 292, 212-223.

McGrath, K, Eriksen, AB, García-Martínez, D, Galbany, J, Gómez-Robles, A, Massey, JS, Fatica, LM, et al. (2022). Facial asymmetry tracks genetic diversity among Gorilla subspecies. *Proceedings of the Royal Society B: Biological Sciences*, 289(1969), 20212564.

McNabb, J, Cole, J, & Hoggard, CS. (2018). From side to side: Symmetry in handaxes in the British Lower and Middle Palaeolithic. *Journal of Archaeological Science: Reports*, 17, 293-310.

Melchionna, M, Profico, A, Buzi, C, Castiglione, S, Mondanaro, A, Del Bove, A, Sansalone, G, et al. (2021). A New Integrated Tool to Calculate and Map Bilateral Asymmetry on Three-Dimensional Digital Models. *Symmetry*, 13(9), 1644.

Mitteroecker, P, & Gunz, P. (2009). Advances in geometric morphometrics. *Evolutionary Biology*, 36(2), 235-247.

Mitteroecker, P, Gunz, P, Windhager, S, & Schaefer, K. (2013). A brief review of shape, form, and allometry in geometric morphometrics, with applications to human facial morphology. *Hystrix, the Italian Journal of Mammalogy*, 24(1), 59-66.

Mitteroecker, P, & Schaefer, K. (2022). Thirty years of geometric morphometrics: Achievements, challenges, and the ongoing quest for biological meaningfulness. *American Journal of Biological Anthropology*, 178(S 74), 181-210.

Mori, T, Profico, A, Reyes-Centeno, H, & Harvati, K. (2020). Frontal bone virtual reconstruction and geometric morphometric analysis of the mid-Pleistocene hominin KNM-OG 45500 (Olorgesailie, Kenya). *Journal of Anthropological Sciences*, 98, 49-72.

Neubauer, S, Gunz, P, Scott, NA, Hublin, JJ, & Mitteroecker, P. (2020). Evolution of brain lateralization: A shared hominid pattern of endocranial asymmetry is much more variable

in humans than in great apes. *Science Advances*, 6(7), eaax9935.

Noori, SMR, Farnia, P, Bayat, M, Bahrami, N, Shakourirad, A, & Ahmadian, A. (2020). Automatic detection of symmetry plane for computer-aided surgical simulation in craniomaxillofacial surgery. *Physical Engineering Sciences in Medicine*, 43(3), 1087-1099.

O'Higgins, P. (2000). The study of morphological variation in the hominid fossil record: biology, landmarks and geometry. *Journal of Anatomy*, 197(1), 103-120.

O'Higgins, P, Fitton, LC, & Godinho, RM. (2019). Geometric morphometrics and finite elements analysis: Assessing the functional implications of differences in craniofacial form in the hominin fossil record. *Journal of Archaeological Science*, 101, 159-168.

O'Higgins, P, & Jones, N. (1998). Facial growth in *Cercocebus torquatus*: an application of three-dimensional geometric morphometric techniques to the study of morphological variation. *Journal of Anatomy*, 193(2), 251-272.

Okumura, M, & Araujo, AG. (2019). Archaeology, biology, and borrowing: A critical examination of Geometric Morphometrics in Archaeology. *Journal of Archaeological Science*, 101, 149-158.

Oxnard, C, & O'Higgins, P. (2009). Biology clearly needs morphometrics. Does morphometrics need biology? *Biological Theory*, 4(1), 84-97.

Perez, SI, Bernal, V, & Gonzalez, PN. (2006). Differences between sliding semi-landmark methods in geometric morphometrics, with an application to human craniofacial and dental variation. *Journal of Anatomy*, 208(6), 769-784.

Profico, A, Bellucci, L, Buzi, C, Di Vincenzo, F, Micarelli, I, Strani, F, Tafuri, MA, et al. (2019). Virtual anthropology and its application in cultural heritage studies. *Studies in Conservation*, 64(6), 323-336.

Recheis, W, Weber, GW, Schäfer, K, Knapp, R, Seidler, H, & Zur Nedden, D. (1999). Virtual reality and anthropology. *European journal of radiology*, 31(2), 88-96.

Ridel, A, Demeter, F, L'abbé, E, & Oettlé, A. (2020). Nose approximation among South African groups from cone-beam computed tomography (CBCT) using a new computer-assisted method based on automatic landmarking. *Forensic Science International*, 313, 110357.

Rohlf, FJ. (2015). The tps series of software. *Hystrix, the Italian Journal of Mammalogy*, 26(1), 9-12.

Rolfe, S, Pieper, S, Porto, A, Diamond, K, Winchester, J, Shan, S, Kirveslahti, H, et al. (2021). SlicerMorph: An open and extensible platform to retrieve, visualize and analyse 3D morphology. *Methods in Ecology and Evolutionary*, 12(10), 1816-1825.

- Roth, VL, & Mercer, JM. (2000). Morphometrics in development and evolution. *American Zoologist*, 40(5), 801-810.
- Ruiz, RM, Torres, MTM, & Allegue, PS. (2021). Comparative Analysis Between the Main 3D Scanning Techniques: Photogrammetry, Terrestrial Laser Scanner, and Structured Light Scanner in Religious Imagery: The Case of The Holy Christ of the Blood. *Journal on Computing Cultural Heritage*, 15(1), 1-23.
- Saragusti, I, Sharon, I, Katzenelson, O, & Avnir, D. (1998). Quantitative analysis of the symmetry of artefacts: Lower Paleolithic handaxes. *Journal of Archaeological Science*, 25, 817-825.
- Schlager, S. (2017). Morpho and Rvcg–Shape Analysis in R: R-Packages for geometric morphometrics, shape analysis and surface manipulations. In *Statistical shape and deformation analysis*, 217-256.
- Schlager, S, & Rüdell, A. (2017). Sexual Dimorphism and population affinity in the human zygomatic structure—comparing surface to outline data. *The Anatomical Record*, 300(1), 226-237.
- Serwatka, KJLT. (2015). Shape variation of Middle Palaeolithic bifacial tools from southern Poland: a geometric morphometric approach to Keilmessergruppen handaxes and backed knives. *Lithics*(35), 18-32.
- Stephan, CN. (2015). Facial approximation—from facial reconstruction synonym to face prediction paradigm. *Journal of Forensic Sciences*, 60(3), 566-571.
- Stephan, CN, Caple, JM, Guyomarc'h, P, & Claes, P. (2019a). An overview of the latest developments in facial imaging. *Forensic Sciences Research*, 4(1), 10-28.
- Stephan, CN, Meikle, B, Freudenstein, N, Taylor, R, & Claes, P. (2019b). Facial soft tissue thicknesses in craniofacial identification: Data collection protocols and associated measurement errors. *Forensic Science International*, 304, 109965.
- Theska, T, Sieriebriennikov, B, Wighard, SS, Werner, MS, & Sommer, RJ. (2020). Geometric morphometrics of microscopic animals as exemplified by model nematodes. *Nature Protocols*, 15(8), 2611-2644.
- Thulman, DK. (2012). Discriminating Paleoindian point types from Florida using landmark geometric morphometrics. *Journal of Archaeological Science*, 39(5), 1599-1607.
- Torres-Tamayo, N, Schlager, S, García-Martínez, D, Sanchis-Gimeno, JA, Nalla, S, Ogihara, N, Oishi, M, et al. (2020). Three-dimensional geometric morphometrics of thorax-pelvis covariation and its potential for predicting the thorax morphology: A case study on Kebara 2 Neandertal. *Journal of Human Evolution*, 147, 102854.
- Viscosi, V, & Cardini, A. (2011). Leaf morphology, taxonomy and geometric

morphometrics: a simplified protocol for beginners. *PloS one*, 6(10), e25630.

Weber, GW. (2014). Another link between archaeology and anthropology: virtual anthropology. *Digital applications in archaeology and cultural heritage*, 1(1), 3-11.

Weber, GW. (2015). Virtual anthropology. *American Journal of Physical Anthropology*, 156, 22-42.

Wiley, DF, Amenta, N, Alcantara, DA, Ghosh, D, Kil, YJ, Delson, E, Harcourt-Smith, W, et al. (2005). Evolutionary morphing. The *IEEE Visualization Conference (VIS 2005)*, 431-438.

Wilkinson, C. (2005). Computerized forensic facial reconstruction. *Forensic Science, Medicine, and Pathology*, 1(3), 173-177.

Wilkinson, C. (2010). Facial reconstruction--anatomical art or artistic anatomy? *Journal of Anatomy* 216(2), 235-250.

Wilkinson, C, Rynn, C, Peters, H, Taister, M, Kau, CH, & 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(2), 179-187.

Wynn, T. (2002). Archaeology and cognitive evolution. *Behavioral Brain Sciences*, 25(3), 389-402.

Chapter 2 A comparison of semilandmarking approaches in the analysis of size and shape

Abstract: Often few landmarks can be reliably identified in analyses of form variation and covariation. Thus, semilandmarking algorithms have increasingly been applied to surfaces and curves. However, the locations of semilandmarks depend on the choice of algorithm and their density, by the investigator. In consequence, to the extent that different semilandmarking approaches and densities result in different locations of semilandmarks, they can be expected to yield different results concerning patterns of and variation and co-variation. The extent of such differences due to methodology, rather than sampling is as yet unclear and often ignored. In this study, the performance of three landmark driven semilandmarking approaches is assessed, using two different surface mesh data sets with different degrees of variation and complexity, by comparing the results of morphometric analyses. These approaches produce different semilandmark locations, which in turn lead to different statistical results, although the non-rigid semilandmarking approaches are broadly consistent. Thus, the results of morphometric analyses using semilandmarks must be interpreted with due caution, recognising that error is inevitable and that, in consequence, statistical descriptions of transformations based on any semilandmarking approach are approximations of reality.

Article details:

Wuyang Shui, Antonio Profico and Paul O’Higgins. A comparison of semilandmarking approaches in the analysis of size and shape. *Animals*. 2023. (under review)

At the time of submission, this chapter was pending. It was published by *Animals* on March 28, 2023. See the link: <https://www.mdpi.com/2076-2615/13/7/1179>.

2.1 Introduction

Geometric morphometric (GM) methods are regularly applied in the analysis of size and shape variation among landmark configurations taken from biological structures (Adams et al., 2004; Mitteroecker & Gunz, 2009; Mitteroecker & Schaefer, 2022; O'Higgins, 2000; Viscosi & Cardini, 2011). Landmarks are matched points among objects that define a map of point equivalences among samples. In biology, the selection and identification of landmarks is a critical first step because it translates the question at hand into a geometric one, amenable to statistical analysis. As such it has been argued that the choice of landmarks and the basis upon which they are considered equivalent should derive from the specific question at hand (Oxnard & O'Higgins, 2009). However, probably more common in practice is 'question free' landmarking, in the sense that the landmarks do not specifically aim to represent a particular hypothesis to be tested, but rather to 'capture overall form' (Bardua et al., 2019). As such, as many landmarks as possible are placed on the specimens, with the basis of equivalence among them being defined according to different criteria (i.e. with mixing of landmarks deemed to be developmentally or evolutionarily equivalent or homologous, with points that are equivalent in terms of function, topography or geometry). Analyses then address specific biological questions by examining vectors of covariation of landmark configurations with extrinsic variables of interest such as size, age, sex, ecology, etc. (Chakravarty et al., 2011).

The choice of landmarks inevitably impacts the results (e.g. distance matrices) obtained from subsequent analyses. In studies where the aim is simply to discriminate groups, or classify specimens as belonging to one or other predetermined group, the effects of landmark choice on the distance matrix among specimens may matter little. What does matter is the degree of discrimination and the accuracy of classification or identification. As such, landmarks (and semilandmarks, see below) may be chosen and marked up, based only on eventual discrimination and classification accuracy.

However, in studies of biological transformations of form it is argued (Cardini, 2020; Oxnard & O'Higgins, 2009) that the basis of matching of points between specimens is important and requires prior knowledge of developmental or evolutionary homology. Thus, in studies of developmental or evolutionary transformations or of variation resulting from these processes, the process of landmarking relates to developmental or evolutionary equivalences ('this point, becomes that point over time'), while in studies of function, different criteria for matching are more appropriate. Uncertainties in identifying such corresponding points, differences in these criteria and choice of landmarks will result in different landmarking schemes (Oxnard & O'Higgins, 2009), which in turn will lead to different distance matrices among specimens and so to different descriptions of transformation or variation.

Beyond these issues, further problems arise, in that that few, if any reliably equivalent

landmarks, defined in terms of homology, can be identified in most studies. This is less of an issue where the equivalence of landmarks required for a particular study is functional. In this case these points represent biomechanical equivalences (e.g. end points, fulcrum, etc. of levers (Oxnard & O'Higgins, 2009)). However, the claim of equivalence between landmarks in terms of homology, is problematic, because as Macleod (MacLeod, 1999) has noted, 'The methods for inferring biological homology classically apply to entire structures; not to infinitesimally small geometric points (e.g., landmarks) within those structures.' This is echoed by Oxnard and O'Higgins (2009) who note that 'There are many cases where landmarks are poorly defined, not repeatable, and uncertain with regard to equivalence'. The extent to which landmarks can be considered homologous in a developmental or evolutionary sense is debatable. Some features such as the tips of tooth cusps might be argued to be homologous, and to present readily identified landmarks, while others such as the surface of the parietal bone in mammals may also be homologous but do not present readily identifiable landmarks. Morphometricians often refer to 'homologous landmarks', but this is shorthand for 'equivalent points on homologous structures'. We adopt this convention here, as shorthand, but recognize that homology of points is at best uncertain and that this uncertainty relates to the both identifiability (how readily the location of a landmark can be identified) and a hypothesis of homology (shared developmental or evolutionary origin). Landmarks may be readily identifiable (e.g. from topography) but not biologically homologous. The extent to which identifiability and homology coincide is largely unknowable, but clearly varies according to anatomical (topographic) features and the extent to which the material on which they are located is the product of shared developmental or evolutionary processes. This is recognized by Bookstein's (1991) classification of landmarks and is considered in greater detail by Oxnard and O'Higgins (2009). While recognizing these difficulties in identifying homologous landmarks, in this study, we treat landmarks as if they are truly homologous, because the focus of this work is on the effects of choosing different algorithms for locating semilandmarks.

Landmarks cannot be readily identified over smooth surfaces such as the cranial vault. It has therefore become increasingly common to use algorithms to place densely matched points over such regions in every specimen. In these cases, equivalence of placement is determined by algorithms that often, but not always, use the locations of identifiable landmarks as a guide. The aim may be more explicitly to capture 'overall form', e.g. by 'increasing the density of the shape information' (Marshall et al., 2019) rather than to identify and mark up homologous points among the sample. Algorithmic landmarking methods that use 'known' point homologies (landmarks) to estimate 'dense point correspondences' (a term commonly used in computer science (Blanz & Vetter, 1999)) among surfaces between landmarks are commonly termed semilandmarking methods or approaches and such dense point correspondences are known as semilandmarks

(Bookstein, 1997). This terminology distinguishes them from landmarks that are considered to be equivalent among specimens in the sense of developmental or evolutionary homology (Oxnard & O’Higgins, 2009).

While identification and equivalence of landmarks are critical issues to resolve, the design of a landmark configuration is also of central importance. In practice, landmark configurations in any one study should be determined by the question at hand, the biologists’ knowledge and experience, prior work, preservation of material and available time for digitization. This is an unsatisfactory situation because different researchers approaching the same question may use different landmarks and obtain different analytical results. Inevitably, this has prompted debate over how to define an appropriate configuration of landmarks for each study (Cardini, 2020; Oxnard & O’Higgins, 2009; Viscosi & Cardini, 2011). This goes beyond the question of whether or not the landmark configuration as a whole captures enough ‘information’ (Bardua et al., 2019; Goswami et al., 2019), because the ‘information’ required will vary according to the hypothesis. A simple configuration may be perfectly adequate and entirely appropriate, depending on the question.

Where landmarks are sparse and data on regions between landmarks is deemed necessary especially in smooth regions, several approaches to creating dense point correspondences have been proposed. One approach is to identify points on each specimen based on some simple mathematical rules applied to each separately (e.g. evenly spaced over a curve (2D) or surface (3D) using sampling algorithms (Bardua et al., 2019; Boyer et al., 2015)). The correspondence of points among specimens is specified in terms of the algorithmic rule used to place them. While terminology is inconsistent, such point correspondences have been referred to as pseudo-landmarks, rather than semilandmarks (e.g. (Goswami et al., 2019)), reflecting the fact that they are placed without reference to other specimens (based on a prior model of equivalence such as developmental equivalence; developmental homology). Semilandmarks, on the other hand, are defined on curves or surfaces in such a way that their locations are controlled by true landmarks and depend on the locations of landmarks in a template or in the sample as a whole. Both pseudo- and semilandmarks are treated as true “landmarks” in subsequent analyses (Bookstein, 1997; Gunz et al., 2009).

Since different landmarking philosophies and approaches lead to different landmarks, and so, are used to ‘ask’ different questions and produce different analytical results, the question of how to automatically yield landmarks has generated great interest. In studies of transformations and variation arising as a result of development or evolution, homologies inevitably underpin landmarking (Bookstein, 1989; Oxnard & O’Higgins, 2009). Within computer science where the focus has been more on ‘general’ similarity or geometric difference (discrimination) and identification/classification, biological

homology, is usually not explicitly considered, rather the focus is on matching of points taken on similar topographical features. At least four approaches are commonly used. In the first, an approach used both in biological morphometrics and computer vision, landmarks are recognized (usually on the basis of equivalence between specimens in terms of topographical features, e.g. curvature, known homology or by defined rules and relationships) and marked up manually or semi-automatically (Negrillo-Cárdenas et al., 2022). In the second approach, landmarks are marked up or automatically identified on a template and these are then projected onto the target specimens using rigid and non-rigid alignment algorithms that are very similar to those used in the majority of semilandmarking approaches (Li et al., 2017; Porto et al., 2021). That landmarks need to be transferred from a template might be considered a drawback. This may be avoided in the third approach, which predicts landmarks based on a training model derived from a statistical shape model describing a priori knowledge of 3D shape variation (Canavan et al., 2015; Cootes et al., 1995). Finally, deep learning has been used to mark up the locations of landmarks (Chen et al., 2021; Le et al., 2020). Despite the considerable effort that has been devoted to them, there has been limited interaction between biological morphometricians and computer scientists with respect to the theoretical and philosophical underpinnings of these last two approaches. In consequence different landmarking strategies are routinely applied in each field.

Further, landmarks are frequently scant and this has driven the development of approaches to marking up dense correspondences (= semilandmarks) between specimens over surface regions between landmarks. Different strategies are commonly applied to semilandmarking in biological morphometrics and computer vision.

In morphometrics as applied in biology, methods for semilandmarking have been developed that use equivalent landmarks (based on prior knowledge) as control points to estimate the locations of semilandmarks by projection followed by sliding. Under this procedure a template specimen is manually landmarked and then semilandmarks are manually or semiautomatically placed on curves and surfaces. Subsequently, the semilandmarks on the template specimen are transferred to each specimen (e.g. by selecting the nearest point between the template and target specimen (Rolfe et al., 2021)). This is followed by sliding of semilandmarks, usually to minimize either the bending energy of a triplet of thin-plate splines (TPS) or Procrustes distances among specimens (Gunz et al., 2005; Perez et al., 2006). Sliding is achieved iteratively, replacing and refitting the template with the mean for the first iteration, and with the recomputed mean for subsequent ones. Of the two approaches, sliding TPS through the minimization of bending energy is most commonly applied (Gunz & Mitteroecker, 2013). It is argued that in Procrustes distance minimization, all landmarks and semilandmarks influence the sliding, even if very distant from the semilandmarks being slid, while minimization of

bending energy gives greater weight to landmarks and semilandmarks that are local to the semilandmarks (note, however, that landmarks may be close to semilandmarks, but on different surfaces). However, in both cases a set of landmarks is necessary to guide the sliding approach, and it cannot be applied when none are present.

Different strategies for marking up dense point correspondences (=semilandmarks) have been pursued in the field of computer vision. These have been applied to biological material as well as to non-biological objects where they rely on mathematical mappings based on topographic features, rather than developmental or evolutionary equivalences. While the use of topographic features to identify point correspondences is different in principle to how landmarks are said to be identified in biology, in practice biologists often rely on anatomical features defined topographically rather than through detailed developmental or evolutionary analysis. This is for the simple reason that in closely related species and within species, similar structures in similar locations are usually developmentally and evolutionarily homologous. However, in regions with simple topography the locations of semilandmarks depend more on the algorithm used to place them. Algorithms used in computer science for mapping include optic flow (Banz & Vetter, 1999), Generalized Procrustes analysis (GPA) and TPS (Mydlová et al., 2015; Velemínská et al., 2012), GPA and coherent point drift (CPD) (Musilová et al., 2016), non-rigid Iterative closest point (NICP) (Amberg et al., 2007; Booth et al., 2018; Shui et al., 2020), scaled rigid Iterative closest point (ICP) and visco-elastic models (White et al., 2019) and a 3D registration method integrating ICP, CPD and the Laplace-Beltrami operator (Dai et al., 2020). Notably, in each of these, a set of point correspondences (determined algorithmically or visually) provides an initial map of equivalences among specimens that is used to guide subsequent algorithmic marking up of dense correspondences among the surfaces, between landmarks.

To avoid manually placing landmarks, several landmark-free algorithms have also been proposed for marking up dense correspondences (Van Kaick et al., 2011). The fitting of a template (reference) specimen surface to each target via registration or alignment algorithms underlies the most common approaches, such as the ICP algorithm (Besl & McKay, 1992). This comprises two steps, iterated until the sum of squared distances among point correspondences between the template and target specimens reaches a minimum. First, correspondences are updated by searching for the nearest points to the registered template points in the target. Second, point clouds of template and target surfaces are rigidly registered by minimizing the squared Euclidean distances among candidate pairwise corresponding points. Registration and identification of correspondences are then iterated until a minimum is reached. Many variants of the ICP algorithm have been proposed with the aim of improving the accuracy of registration (minimisation of template-target distances) by e.g. using different distance metrics or

assigning different weights to vertices and rejecting outliers (Rusinkiewicz & Levoy, 2001). For example, an improved ICP-based approach has been proposed to register the surfaces of specimens by minimising the symmetric point-to-plane distances (along the surface normal vector) instead of point-to-point distances (Pomidor et al., 2016). Additionally, the point correspondences found by the ICP algorithm rely on the initial alignment of two surfaces, by for example, least squares. An alternative is to use principal component analysis (PCA) to find the principal axes of the template and target point sets to provide a sensible initial position. Different approaches lead to different maps of point correspondences (semilandmarks).

Another landmark-free algorithm, available as an auto3dgm package (Boyer et al., 2015) based on the ICP framework, has been proposed to yield semilandmarks among specimens. In this, a set of points on the template specimen are projected onto the target. However, the choice of the template, the degree of complexity, and the density and locations of points affect the results (Gao et al., 2018; Vitek et al., 2017). To mitigate this, a template is chosen that has the greatest overall geometric similarity to the members of the sample. Then semilandmarks of the template specimen are projected to each specimen. Vitek et al. (2017) have indicated that the choice of initial alignment influences the resulting estimates of point correspondences when using auto3dgm and the lack of true landmarks as control points impacts registration. Moreover, this approach involves rigid registration and so can result in equivalent points on the template and target specimens being placed on different anatomical features. This is most pronounced among specimens with large differences in shape and size, e.g. points around the zygomatic process of a rigidly registered temporal bone might be projected from the reference onto the condyle of the target (Vitek et al., 2017).

One landmark-free approach (Gu et al., 2004; Wang et al., 2007) uses conformal geometry to establish point equivalences among 3D meshes, because any genus zero surface can be mapped conformably onto a sphere and surface with a single boundary can be mapped onto a unit disk. In practice, the conformal transformation is applied to the 3D surface and then the correspondences between two surfaces are found in the 2D domain. Examples are provided by the work of (Boyer et al., 2011) and (Koehl & Hass, 2015), however, these conformal methods are sensitive to the quality of surfaces and the complexity of topologies.

Landmark-free algorithms, e.g. ICP-based method, for marking up point correspondences between surfaces can result in mappings that are quite different from the map of point equivalences based on postulated homologies. This effect can be large, with semilandmarks from the template projected to different anatomical features in the target. In any case, even when ‘appropriately located’, such equivalences have no implicit biological basis, they might, or might not, be good approximations of homology. In

consequence, statistical (e.g. PCA) and/or visual (e.g. warping between surfaces) descriptions of (developmental or evolutionary) transformation might or might not properly describe them and analyses of variation based on them may or may not reflect the developmental or evolutionary basis of variation. By their nature, point correspondences identified without paying attention to homology have an uncertain relationship to the underlying processes responsible for differences in form. They may however be useful in discrimination, identification or classification of specimens to prior groups. These are a different, yet important and common application of landmark data in computer vision, but very different tasks to that of describing developmental and evolutionary transformations.

It is worth noting that all semilandmarking approaches are dependent on mathematical models of matching. As such, homology is only respected to the extent that the mathematical model uses the homologies of true landmarks to estimate semilandmarks, and to the extent that the landmarks actually represent biologically homologous points. Semilandmarks can be intended as estimates of, rather than true, homologous point matchings (landmarks). Thus, different semilandmarking approaches will yield different semilandmarks. The reasons for preferring one approach over another cannot rest entirely on arguments of developmental or evolutionary equivalence of the resulting semilandmarks because the definition and identification of homologous points relies on prior developmental and evolutionary knowledge (which is often lacking). No algorithm without a knowledge driven model of homology can properly determine or interpolate homology from surface or texture features. Instead, assessment of algorithms has focused on their ‘performance’ defined in various ways.

A few previous studies have attempted to assess the performance of different semilandmarking approaches. Evaluated criteria for comparing different semilandmarking approaches include: the Euclidean distances between semilandmarks (or landmarks) from each approach or with manually placed ones (Porto et al., 2021), comparison between methods of: the resulting distributions of groups (Boyer et al., 2015; Gonzalez et al., 2016; Rolfe et al., 2021), the geometric deviation between template and transformed meshes (Rolfe et al., 2021; Shui et al., 2020), the first two principal components (PCs) (Boyer et al., 2015; Gonzalez et al., 2016; Harper et al., 2022), distance matrices to quantify shape variations (Boyer et al., 2011; Boyer et al., 2015; Harper et al., 2022; Pomidor et al., 2016), and comparison of estimates of centroid size of resulting configurations (Porto et al., 2021). These criteria may indicate how different semilandmarking strategies perform in matching surfaces, distinguishing groups or identifying unknown specimens, but they do not relate to how well the homology map is represented by the resulting semilandmarks. All suffer from the fact that semilandmarks are not point homologies, they may be estimates of such homologies, but each estimate

is different. As such, the extent to which they correctly describe biologically homologous anatomical differences and transformations is limited by the extent to which knowledge of homology is embedded in their construction.

The consequences of choosing alternative semilandmarking approaches in studies of biological form variation need to be further investigated. In this study, the degree to which they generate different results is investigated and the significance of the findings for future studies is discussed. To these ends, we employed three landmark driven semilandmarking methods, sliding TPS, an example of a rigid, and an example of a non-rigid registration approach to yield semilandmarks for two datasets comprising surfaces with different degrees of complexity and distributions of identifiable landmarks to guide semilandmarking. These semilandmarking approaches are compared by empirically testing six hypotheses using surface scans of human faces and ape crania based on the same template of landmarks and semilandmarks. These are: (i) that for the same density of semilandmarking there are no differences in semilandmark locations generated by different approaches and (ii) in mean landmark and semilandmark configurations. Further, between densities of semilandmarking using any one approach and between semilandmarking approaches, there are no differences in resulting estimates of (iii) centroid sizes (iv) distance matrices (v) PCs of shape variation (vi) allometrically scaled shapes.

2.2 Materials and Methods

2.2.1 Materials

Two high resolution datasets were used in this study. These comprise surface meshes with different degrees of surface complexity and variation in form. Each human head surface scan comprises more than 146,000 vertices and 290,000 triangles and each ape cranial surface model comprises more than 196,000 vertices and 391,000 triangle meshes.

2.2.1.1 Adult human heads

The sample comprises 100 surface scans of adult human male heads from the Liverpool-York 3D ‘headspace’ dataset, consisting of 1519 subjects. Geometry and texture were captured using a 3dMD five-camera system (Dai et al., 2020). 16 anatomical (homologous) landmarks had already been manually marked up on each using the EVAN toolbox, in the course of a previous study (Smith et al., 2021) and these are shown in Figure 2.1a. These cover much of the facial region but landmarks are not identifiable over the scalp.

2.2.1.2 Ape crania

This sample included 20 surface meshes of ape crania captured by CT scanning, including 5 *Gorilla*, 5 *Hylobates lar*, 5 *Pan troglodytes*, and 5 *Pongo abelii* (Table 2.S1).

These present more complex surfaces and a far greater degree of variation in size and shape than the head surface dataset. As shown in Figure 2.1b, 41 anatomical landmarks had already been manually placed over the entire cranium, in the course of a previous study (Profico et al., 2017).

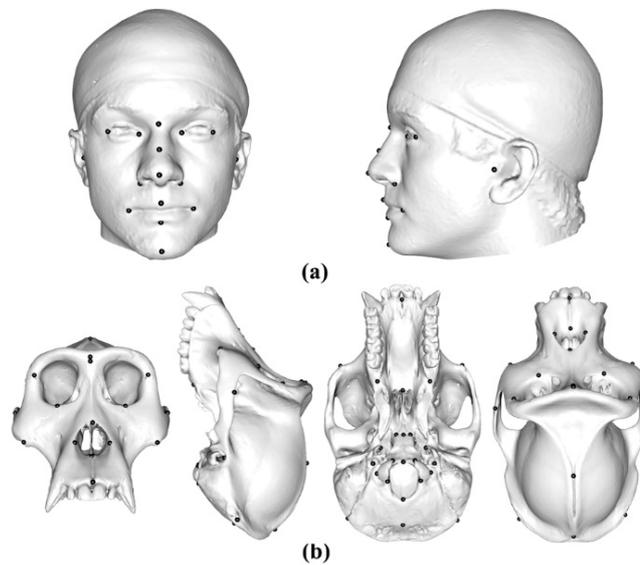


Figure 2.1 Two datasets with anatomical landmarks. (a) An adult human male head with 16 landmarks. (b) An ape cranium with 41 landmarks.

2.2.2 Methods

2.2.2.1 Three semilandmarking approaches

Based on the fixed landmarks, we used three different methods to mark up semilandmark sets of varying density among specimens using a common template and then tested six hypotheses (Figure 2.2).

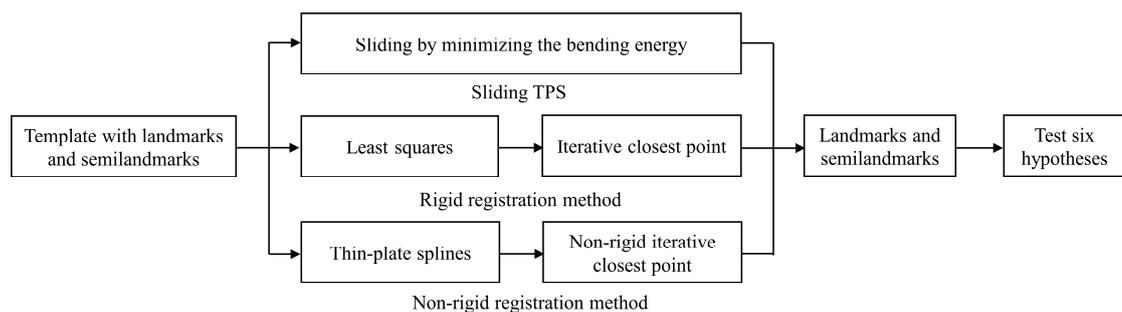


Figure 2.2 Workflow of comparison of three semilandmarking landmarks

2.2.2.1.1 Generation of template

A landmark and semilandmark template was created for each dataset to be used as the basis of semilandmarking using three different approaches. To create this template for the head data, we used the NICP algorithm (Amberg et al., 2007) developed in Matlab to align all of the human heads using landmarks as a ‘soft constraint’, and from this alignment computed a mean template form (size and shape) surface, by averaging the coordinates of every vertex of all the heads. Then, we used the k-means clustering

algorithm to evenly distribute five different densities (200, 400, 600, 800, and 1000) of surface semilandmarks over the template, ignoring the locations of true landmarks and avoiding the ears, which have a complex surface.

For the ape crania, the external surface of every specimen was extracted by computing the intersection points of each cross-section plane and the 3D meshes (Shui et al., 2020), to avoid the internal surface interfering with the sliding and projection of semilandmarks. Next, we used the `fastKmeans` function in the `Morpho` R package to perform the k-means clustering algorithm to evenly sample 800 surface semilandmarks over an arbitrary specimen (a male gorilla, the Procrustes distance between it and the template, based on the landmarks and semilandmarks was 0.0988, which is very similar to the average difference between individuals and the mean, estimated using sliding TPS semilandmarking), ignoring symmetry, and employed the sliding TPS approach to yield semilandmarks among specimens. Then, the mean form of landmarks and semilandmarks was calculated and the arbitrary specimen surface was deformed to approximate the mean surface form. This surface was used as the template. Finally, we utilized the k-means clustering algorithm to evenly sample five different densities of surface semilandmarks (50, 100, 200, 400, 800) on the template, avoiding the cranial base and teeth.

2.2.2.1.2 Semilandmarking approaches

Once the templates were created, the three commonly used methods of semilandmarking described above were applied as follows:

a) Sliding TPS

Semilandmarks were projected from the template surface onto each target surface and then iteratively slid over the target surface to minimize the bending energy of the TPS between each specimen and template. This is the classic approach first described by (Bookstein, 1997) and developed by (Gunz & Mitteroecker, 2013; Gunz et al., 2005). It is advised not to locate semilandmarks beyond the bounds of the true landmarks (Oxnard & O'Higgins, 2009) where the TPS, whose bending energies are minimised during subsequent sliding, is progressively less constrained by the landmarks. Consequently, after sliding the scalp semilandmarks may be located on non-homologous features (Gunz & Mitteroecker, 2013). However, the other two semilandmarking approaches (below) are expected to function adequately over the scalp, depending on landmarks only for the initial registration. Thus, for consistency and to be able to compare results, the sliding TPS algorithm was applied to scalp semilandmarks, despite known issues. Subsequent analyses compare results from the scalp and face among semilandmarking approaches, and so provide insights into the severity of this issue in this dataset. We used the patching (`placePatch`) and sliding (`slider3d`) procedures in the `Morpho` R package to slide the semilandmarks based on the templates (Schlager, 2017).

b) Rigid registration

Rigid registration aims to find the linear transformation of the template to each specimen surface (translation and rotation) that aligns two surfaces without scaling in such a way that the sum of squared Euclidean distance between landmarks (and, if present, semilandmarks) is minimised. Note, this method does not deform a surface to optimise fitting. So, it is prone to error due to the difference in size between the template and every specimen.

Semilandmarking of the target was achieved iteratively using a hybrid rigid registration combining LS and point-to-point ICP algorithms (LS&ICP). First, the LS algorithm was used to fit the template to each specimen, minimizing the distances between landmarks on the template and each target specimen by translating and rotating the template to best fit the target. Subsequently, the ICP algorithm iteratively rigidly refitted the template to the target, minimising the sum of the squared distances between the landmarks and current estimate of semilandmarks, found by searching for the nearest points on the target surface from the registered template semilandmarks. The initial rigid alignment based on landmarks speeds up convergence during the ICP phase. The C++ programming language was used to code the algorithm to generate semilandmarks using this LS&ICP approach.

c) Non-rigid registration

We presented a hybrid non-rigid registration approach (Shui et al., 2021) to deform the template specimen to fit each target specimen and then projected the semilandmarks from the warped template onto each specimen by searching for the nearest points on its surface, to yield semilandmarks across all specimens. More details of the algorithm can be seen in Chapter 6. Unlike rigid registration, in fitting, each vertex of the template can be moved freely with stretching based on a non-rigid transformation and landmarks acting as constraints. This comprised two steps: first, the TPS algorithm was used to warp the template to every specimen based on the fixed landmarks. This removed size and shape differences between the template and each target set of landmarks and provided a reasonable initial alignment of surfaces. Second, the NICP algorithm (Amberg et al., 2007) was applied to warp the deformed template surface to each specimen as rigidly as possible, optimizing the cost function by assigning an affine transformation to each vertex, rather than an interpolation function as used in TPS. For this procedure, the cost function comprised a landmark term, a local affine regularization term and a stiffness term. Registration loops were performed by decreasing stiffness weights iteratively and deforming the template incrementally. This resulted in the warped template surface matching the target closely. Here, this approach is referred to as TPS&NICP. The C++ programming language and Matlab were used to apply this approach. The software prototype used to generate the semilandmarks using three different approaches can be

downloaded from the Internet (https://github.com/sissun/Geometric_morphometrics.git).

2.2.2.2 Comparison of three semilandmarking approaches

We compared the different semilandmarking approaches by testing the null hypotheses.

2.2.2.2.1 The locations of semilandmarks

Differences between methods in the placement of semilandmarks were assessed by visualizing them and computing the Euclidean distances between each semilandmark (that share the same initial template position), computed using each semilandmarking approach. These were used to compute the average semilandmarking ‘error’ between approaches and to examine their distributions. Note these are ‘errors’ between algorithmic results and not in homology mapping per se, which cannot be evaluated because truly homologous dense point correspondences are unknowable.

2.2.2.2.2 Comparisons of mean landmark and semilandmark configurations

Generalised Procrustes analysis (GPA) was applied to the landmark and semilandmark configurations estimated for the sample and then the mean centroid sizes were compared among different semilandmarking methods and densities. Subsequently, the Procrustes distances among estimates of the mean shape were computed to quantify the differences between them arising from different semilandmarking approaches. To contextualise the extent to which Procrustes distances between means differ, these were compared to the average distances between individuals and the mean for each density of semilandmarks.

2.2.2.2.3 Procrustes distances among specimens obtained using different semilandmarking approaches and densities

We examined the effect of different semilandmarking approaches and different densities of semilandmarks on Procrustes distance matrices.

a) The effect of different semilandmarking approaches

Generalised Procrustes analysis was applied to each of the landmark and semilandmark sets generated by the different semilandmarking approaches. Then, Procrustes shape coordinates and their sample means and centroid sizes were obtained. Procrustes distance matrices among all individuals were calculated and a Mantel Test (Dutilleul et al., 2000) performed to compare distance matrices obtained by the different semilandmarking approaches. Additionally, vectors of Procrustes distances between each individual and the mean were compared among semilandmarking methods by plotting bivariate graphs and computing Pearson correlations.

b) The effect of different densities of semilandmarks

The results obtained from analyses of the landmarks and different densities of semilandmarks were compared with those obtained using the landmarks and maximum density of semilandmarks from each method. This was achieved by computing the Pearson correlations among vectors of Procrustes distances to the mean and by performing a Mantel Test between the Procrustes distance matrices derived from each density and that from the maximum density. Additionally, the matrix of Procrustes distances among specimens based on the landmarks alone was computed in order to compare these distances with those obtained by different semilandmarking approaches and densities.

2.2.2.2.4 PCA and allometry

For the landmarks alone, a GPA and PCA of the covariance matrix was carried out in order to compare PCs with those from the semilandmarking methods. Then, for the landmarks and each density of semilandmarking, a separate GPA was carried out of the landmark and semilandmark configurations derived from each of the three semilandmarking methods. A PCA was then carried out on the resulting shape coordinates from each GPA at each semilandmarking density. To assess how the major vectors of variation (PCs) differ between approaches and semilandmarking densities, we compared the distributions of specimens along the first two PCs by computing the Pearson correlations among the PC 1 and among the PC 2 scores arising from each semilandmarking approach and from landmarks alone.

Next, a joint GPA and PCA was carried out combining landmarks and semilandmark sets of the same density from each semilandmarking approach. The full set of PC scores (which completely account for the Procrustes distances among specimens) was used to estimate allometry (the relationship between size and shape) based on each semilandmarking method (Klingenberg, 2016). Allometry was estimated for the whole sample and each semilandmarking density by multivariate regression of PC scores on the natural logarithm of centroid size. These estimates of allometry were then compared between different semilandmarking approaches based on the angles between allometric vectors (Gonzalez et al., 2010). Small angles indicate that semilandmarks generated by different approaches are similar and large angles indicate that they are more different.

Subsequently, the predicted shapes of landmarks and semilandmarks representing the extreme limits (smallest-largest) of the allometric vectors derived using each semilandmarking method and density were compared. This comprised two steps: first, the predicted shapes (landmark and semilandmark configurations) corresponding to the upper and lower limits of centroid sizes estimated by each approach were estimated from the multivariate regression (O'Higgins, 2000). Second, Procrustes distances were computed

between the predicted shapes representing the maximum and minimum centroid sizes from each of the landmark and semilandmark sets generated by the different semilandmarking approaches.

2.3 Results

We first compared semilandmarking approaches using the adult human head surfaces and then repeated key analyses using the ape cranial dataset to compare findings between surfaces that differ in their complexity, distribution of identifiable landmarks, and degree of variation in size and shape.

2.3.1 Human head scans

2.3.1.1 The locations of semilandmarks

The resulting differences in placement of semilandmarks ('errors') from the three semilandmarking approaches are presented in Tables 2.1-2.3 for the comparisons of sliding TPS and TPS&NICP, sliding TPS and LS&ICP, and TPS&NICP and LS&ICP approaches, respectively. The tables list differences in location (diff, in mm), the average deviation (dev in mm) and the % of semilandmarks that differ in location by 0.0-1.0 mm, 1-2.5 mm, 2.5-5 mm and ≥ 5.0 mm.

Table 2.1 Comparison of semilandmarks from sliding TPS and TPS&NICP.

diff mm	200		400		600		800		1000	
	dev	%								
[0.0-1.0)	0.68	16.00	0.71	15.75	0.72	15.17	0.74	15.13	0.75	15.00
[1.0-2.5)	1.63	18.00	1.68	21.75	1.69	23.67	1.68	27.12	1.72	29.40
[2.5-5.0)	3.81	40.00	3.83	45.25	3.74	50.67	3.73	52.25	3.67	55.50
≥ 5.00	5.75	26.00	5.64	17.25	5.35	10.50	5.12	5.50	5.02	0.10
Total	3.42	100.00	3.18	100.00	2.96	100.00	2.80	100.00	2.66	100.00

Table 2.2 Comparison of semilandmarks from sliding TPS and LS&ICP.

diff. mm	200		400		600		800		1000	
	dev	%								
[0.0-1.0)	-	-	-	-	-	-	-	-	-	-
[1.0-2.5)	-	-	-	-	-	-	-	-	-	-
[2.5-5.0)	4.21	37.50	4.26	47.25	4.28	51.50	4.34	58.63	4.36	52.60
≥ 5.00	5.83	62.50	5.60	52.75	5.54	48.50	5.51	41.37	5.52	47.40
Total	5.22	100.00	4.97	100.00	4.89	100.00	4.82	100.00	4.91	100.00

Table 2.3 Comparison of semilandmarks from TPS&NICP and LS&ICP.

diff. mm	200		400		600		800		1000	
	dev	%								
[0.0-1.0)	-	-	-	-	-	-	-	-	-	-
[1.0-2.5)	-	-	-	-	-	-	-	-	-	-
[2.5-5.0)	4.35	12.50	4.37	15.00	4.33	15.83	4.38	20.25	4.38	17.90
≥5.00	6.84	87.50	6.56	85.00	6.45	84.17	6.43	79.75	6.40	82.10
Total	6.53	100.00	6.23	100.00	6.12	100.00	6.02	100.00	6.04	100.00

The mean locations of 1000 semilandmarks generated by sliding TPS (black points), LS&ICP (red points) and TPS&NICP (green points) approaches are illustrated in Figure 2.3 on the template surface warped to the mean of the sliding TPS generated landmark and semilandmarks. It is evident that their locations differ among semilandmarking approaches. The mean configurations from both TPS&NICP and LS&ICP lie close to, or on the surface defined by the mean configuration from sliding TPS. Similar results are found for all semilandmarking densities, with sliding TPS and TPS&NICP approaches producing the most similar semilandmark locations.

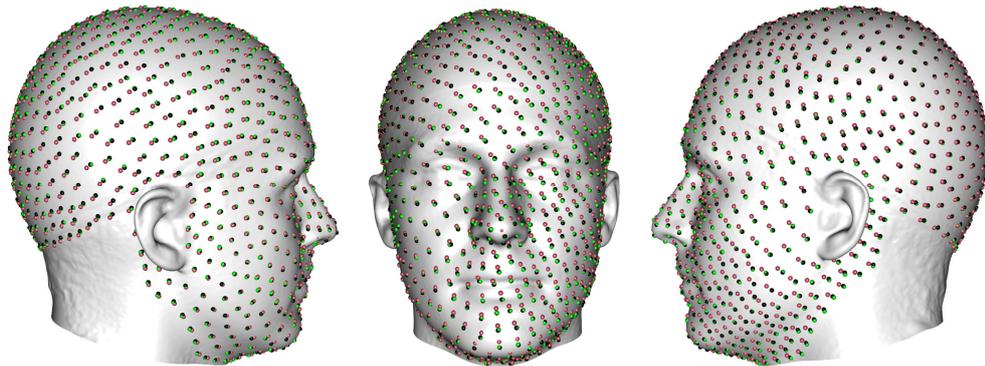


Figure 2.3 Locations of the average coordinates of 1000 semilandmarks generated by sliding TPS (black points), LS&ICP (red points) and TPS&NICP (green points) approaches.

Focusing on the highest semilandmarking density (but note that these findings apply equally to all higher densities), Figure 2.4a shows the differences in location among 1000 semilandmarks generated by sliding TPS and TPS&NICP approaches. 15% (Table 2.1) of the full set of semilandmarks (yellow points) differ in location by <1.0 mm and nearly all of these are located on the face (Figure 2.4a), where the density of landmarks is high. As the density of landmarks in a region falls, the differences in semilandmark placement increase. Thus, 29.4% (Table 2.1) of the semilandmarks (Figure 2.4a, blue points) present a difference of 1.0-2.5 mm, principally around the forehead and cheeks and 55.6% of the semilandmarks (red and black points) present differences greater than 2.5 mm in location among semilandmarking approaches, principally these larger differences are located over the scalp. Figure 2.4b shows the differences in placement of 1000 semilandmarks generated by sliding TPS and LS&ICP approaches. Compared to Figure 2.4a, the deviations between equivalent semilandmarks are larger. This reflects the differences

presented in Table 2.2. For 1000 semilandmarks, all deviations between sliding TPS and LS&ICP are >2.5mm, while sliding TPS and TPS&NICP, locate 44.4% of semilandmarks with difference in location of less than 2.5mm. These greater differences are most evident in the face (Figure 2.4b) where increased asymmetry of differences in the locations of semilandmarks is also evident.

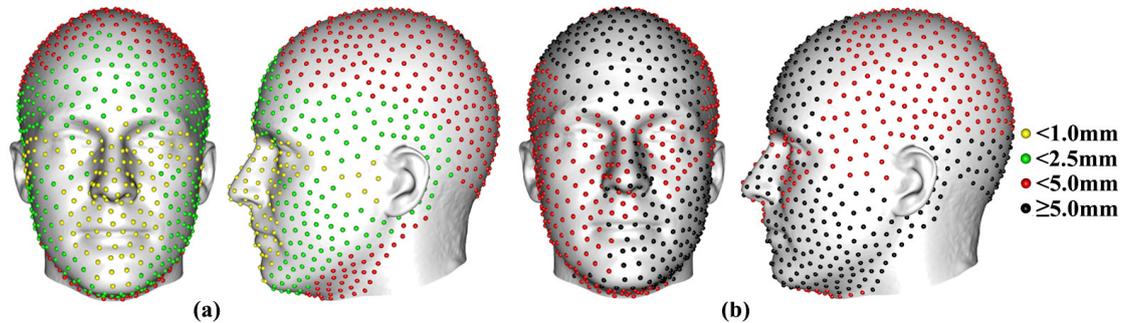


Figure 2.4 Average differences (Euclidean distance in location in mm) between 1000 semilandmarks generated by different approaches. (a) differences between sliding TPS and TPS&NICP. (b) differences between sliding TPS and LS&ICP. Differences between TPS&NICP and LS&ICP approaches are not shown because they are very similar to those in b.

2.3.1.2 Comparison of landmark and semilandmark configurations among different semilandmarking approaches

For adult human male heads, the centroid sizes of the estimated mean configurations derived from different semilandmarking approaches, are presented in Table 2.4. They indicate that the mean centroid sizes estimated using each approach are almost the same and increase similarly with increasing semilandmark density, as expected.

Table 2.4 The centroid sizes of the mean landmark and semilandmark configurations generated by different semilandmarking approaches and different densities of semilandmarks.

	200	400	600	800	1000
Sliding TPS	1514	2126	2599	2990	3334
LS&ICP	1513	2123	2595	2986	3330
TPS&NICP	1513	2124	2598	2989	3333

Additionally, Procrustes distances were computed among estimates of the mean landmark and semilandmark configurations from different semilandmarking approaches (Table 2.5). Comparison of these distances among semilandmarking methods and densities indicates that the estimates of the mean landmark and semilandmark configurations generated from sliding TPS are more similar to TPS&NICP than to LS&ICP. With increasing numbers of semilandmarks, the LS&ICP approach converges on the results obtained using sliding TPS and TPS&NICP, as indicated by the progressively reducing Procrustes distances with increasing semilandmarking density in rows 1 and 3 of Table 2.5. In contrast, the estimated mean configurations from sliding TPS and TPS&NICP present a very similar Procrustes distance (0.0049+/-1) at all densities.

Table 2.5 Procrustes distances computed between mean landmark and semilandmark configurations.

	200	400	600	800	1000
Sliding TPS LS&ICP	0.0112	0.0093	0.0086	0.0083	0.0082
Sliding TPS TPS&NICP	0.0048	0.0049	0.0049	0.0050	0.0049
LS&ICP TPS&NICP	0.0122	0.0103	0.0096	0.0093	0.0091

At each semilandmarking density the estimates of mean shape arising from the different semilandmarking approaches differ by between 14.15% and 37.43% of the average difference between individuals and the mean, estimated using sliding TPS semilandmarking. Between sliding TPS and TPS&NICP this relative distance remains stable (~14.5%) with increasing semilandmark density, whereas it decreases from ~35% to 25% for the comparisons between the mean derived by LS&ICP and those from both sliding TPS and TPS&NICP.

To further explore why estimates of mean landmark and semilandmark configurations from different semilandmarking approaches show the patterns of difference presented in Table 2.5, the semilandmarks generated by sliding TPS were separated into two regions, face (yellow points) and scalp (red points), as in Figure 2.5. Then, Procrustes distances were computed between the landmark and semilandmark sets from these regions, obtained using different approaches and semilandmarking densities (Table 2.6). At all densities, the Procrustes distances between landmarks and semilandmarks in the face and scalp, computed from sliding TPS and TPS&NICP are the smallest, with distances among face semilandmarks being the largest, and these are consistent among semilandmarking densities. By contrast, the Procrustes distances from LS&ICP and both sliding TPS and TPS&NICP (Table 2.6 rows 1 and 3) are larger. Similar to the comparison of sliding TPS and TPS&NICP, distances are greater among facial landmarks and semilandmarks, than among scalp semilandmarks. Further, the distances among face landmarks and semilandmarks from LS&ICP and both sliding TPS and TPS&NICP decrease with increasing semilandmark density. This likely accounts for the convergence with increasing numbers of semilandmarks between the LS&ICP approach and both sliding TPS and TPS&NICP mean shapes seen in Table 2.5.

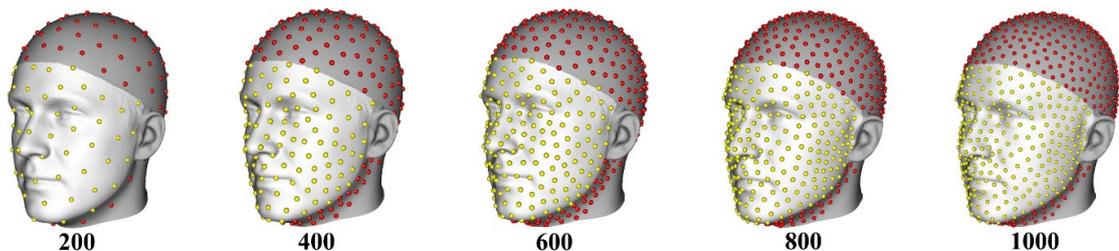


Figure 2.5 The mean semilandmarks generated by sliding TPS over the face (yellow points) and scalp (red points).

Table 2.6 Procrustes distances computed between mean landmark and semilandmark configurations in the face and scalp. Sliding TPS vs LS&ICP (first row), Sliding TPS vs TPS&NICP (middle row), and LS&ICP vs TPS&NICP (bottom row).

200		400		600		800		1000	
Face	Scalp								
0.0229	0.0067	0.0195	0.0058	0.0179	0.0058	0.0168	0.0060	0.0166	0.0062
0.0056	0.0055	0.0057	0.0055	0.0061	0.0054	0.0061	0.0056	0.0061	0.0054
0.0238	0.0081	0.0205	0.0074	0.0190	0.0072	0.0181	0.0071	0.0175	0.0072

2.3.1.3 Comparison of centroid sizes and Procrustes distance matrices

2.3.1.3.1 Differences between semilandmarking approaches

Centroid sizes of individuals, computed using the landmarks and semilandmarks generated by alternative semilandmarking approaches are compared between sliding TPS and the other two semilandmarking approaches in Figure 2.6 (the comparison of TPS&NICP and LS&ICP approaches is not shown because it is very similar to those in Figure 2.6a). Figure 2.6a compares centroid sizes between sliding TPS (horizontal axis) and LS&ICP (vertical axis). The dashed line denotes the expected relationship if centroid sizes are identical, and the red line represents the fitted line from a regression of centroid sizes of individuals from the landmark and semilandmark configurations computed by LS&ICP (dependent variable) on those computed by sliding TPS (independent variable). In each case, the fitted red line shows a smaller gradient than the expected linear relationship indicated by the dashed line. Likewise, Figure 2.6b compares centroid sizes from sliding TPS (independent variable) and TPS&NICP (dependent variable). The plots indicate that the centroid sizes from sliding TPS (red lines) were very similar to those of TPS&NICP at all assessed semilandmarking densities.

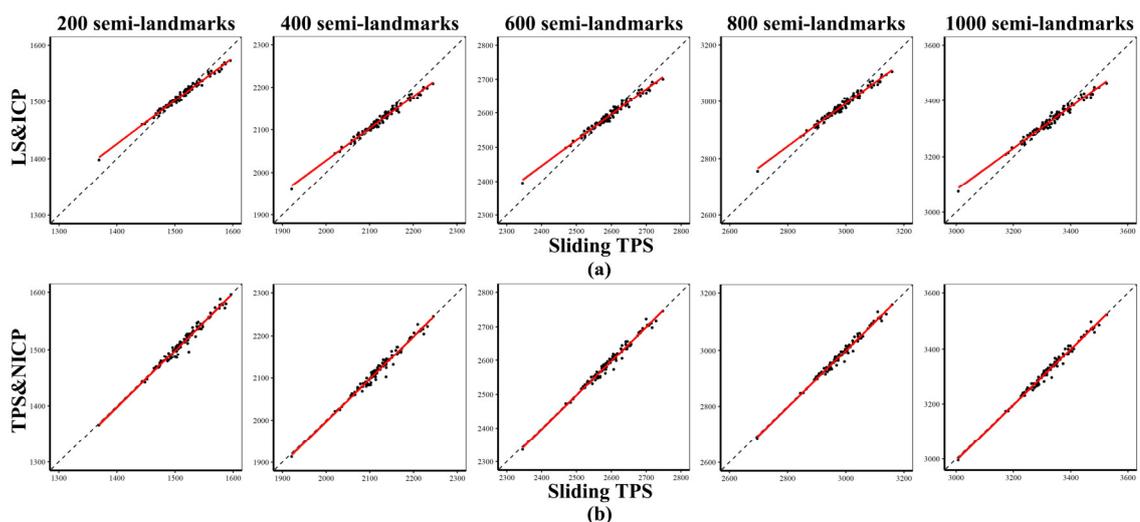


Figure 2.6 Comparison of the centroid sizes of landmarks and semilandmark configurations computed by different approaches. (a) Comparison of TPS sliding and LS&ICP. (b) Comparison of TPS sliding and TPS&NICP. TPS&NICP and LS&ICP approaches are not compared in this figure because the results are similar to those in a.

Figure 2.7 summarizes the vectors of Procrustes distances between each individual and the mean calculated using landmarks and semilandmarks from the sliding TPS, LS&ICP, and TPS&NICP approaches. In Figure 2.7, the horizontal axis represents different densities of semilandmarks and the vertical axis represents Procrustes distances. Consistently across semilandmarking densities within each method, the smallest distances arise from the LS&ICP approach and the largest from TPS&NICP, with sliding TPS intermediate.

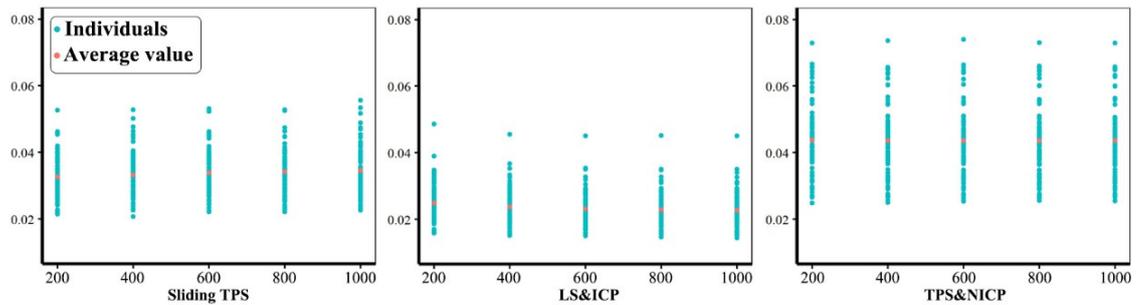


Figure 2.7 Vectors of Procrustes distances between each individual and the mean computed for each semilandmarking approach using different densities of semilandmarks. Cyan points represent Procrustes distance between every specimen and the Procrustes mean shape and red points represent the average value of the Procrustes distance vector.

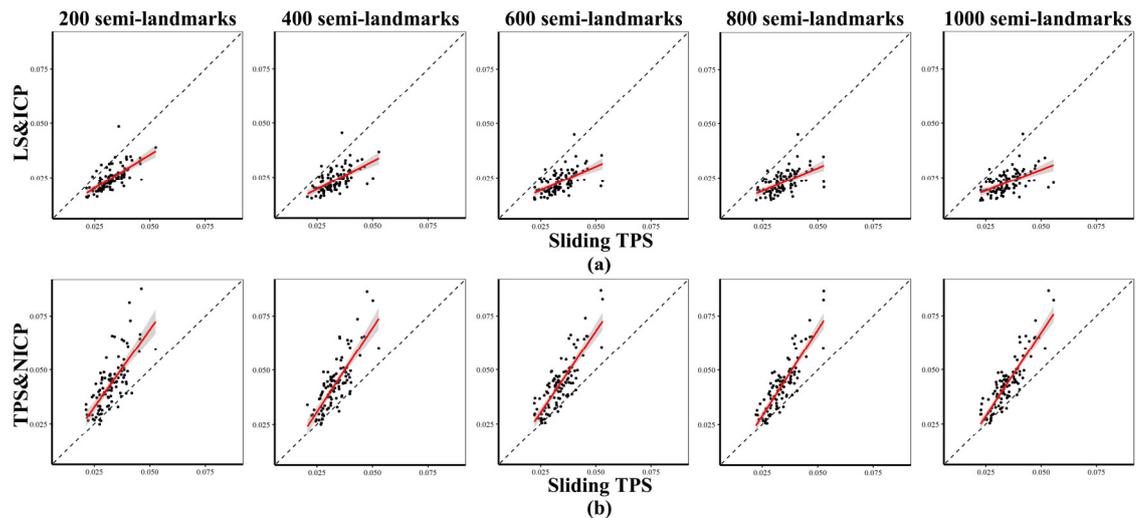


Figure 2.8 Comparison of the vector of Procrustes distances between every specimen and the mean among different approaches. (a) Comparison of TPS sliding and LS&ICP approaches. (b) Comparison of TPS sliding and TPS&NICP approaches. TPS&NICP and LS&ICP approaches are not compared in this figure because the results are very similar to those in a.

Figure 2.8 directly compares the distances from the sliding TPS approach with those from the other two approaches. The results comparing TPS&NICP and LS&ICP are not shown, because they are similar to those in Figure 2.8a comparing vectors of Procrustes distances between each individual and the mean produced by sliding TPS (horizontal axis) and LS&ICP (vertical axis). The dashed line denotes the expected relationship if Procrustes distance vectors are identical, and the red line represents the linearly fitted regression of the vector derived from LS&ICP (dependent variable) on the vector from sliding TPS (independent variable). For each density of semilandmarking, the fitted red

line has a smaller gradient and lies below the expected linear relationship indicated by the dashed line, indicating that Procrustes distances from sliding TPS are greater than those from LS&ICP (supported by Figure 2.7) and this difference increases as Procrustes distance increases. Likewise, Figure 2.8b compares Procrustes distance vectors between sliding TPS and TPS&NICP. The plots indicate that the Procrustes distance vectors from sliding TPS (red lines) are less than those from TPS&NICP at all semilandmarking densities.

Table 2.7 compares the Pearson correlations among the vectors of Procrustes distances between every individual and the mean as well as the Mantel correlation between the Procrustes distance matrices. At all semilandmarking densities, the largest Pearson correlations (r) are between Procrustes distance vectors from sliding TPS and TPS&NICP and these correlations increase with increasing numbers of semilandmarks (second row in Table 2.7). In contrast, the weakest correlations are consistently between LS&ICP and TPS&NICP and these decrease with increasing semilandmark density.

Table 2.7 Pearson correlations among vectors of Procrustes distances between each individual and the mean and Mantel tests of association between the Procrustes distance matrices derived using different semilandmarking approaches. Sliding TPS vs LS&ICP (first row), Sliding TPS vs TPS&NICP (middle row), and LS&ICP vs TPS&NICP (bottom row).

200		400		600		800		1000	
Pearson	Mantel								
0.6830	0.6841	0.6068	0.6260	0.5790	0.5978	0.5555	0.5790	0.5323	0.5528
0.7284	0.7017	0.7901	0.7576	0.8304	0.8009	0.8623	0.8311	0.8829	0.8564
0.3955	0.3916	0.3615	0.3646	0.3478	0.3504	0.3438	0.3454	0.3368	0.3382

Using Mantel tests, Table 2.7 also compares the matrices of Procrustes distances among all individuals calculated using landmarks and semilandmarks generated by the different semilandmarking approaches. The largest matrix correlations were found between sliding TPS and TPS&NICP and the association becomes stronger with increasing numbers of semilandmarks (second row in Table 2.7). The weakest association, as assessed by Mantel tests, is between matrices calculated from semilandmarks derived from LS&ICP and TPS&NICP. These become weaker with increasing numbers of semilandmarks. Additionally, the matrix (Mantel) correlations for this comparison are consistently low relative to the other comparisons (bottom row in Table 2.7 vs rows 1 and 2).

2.3.1.3.2 Differences densities of semilandmarks

Consistency within semilandmarking methods, of relative Procrustes distances among specimens computed using varying densities of semilandmarks, was assessed by tests of association. Pearson correlations were computed among vectors of Procrustes distances between each individual and the mean, as were Mantel correlations among the

matrices of Procrustes distances. These, presented in Table 2.8, compared the distances from the landmark and semilandmark configuration comprising 1000 semilandmarks from the head surface data with those from configurations comprising 200-800 semilandmarks. Within each semilandmarking approach, these correlations are generally large (>0.90) and increase with increasing numbers of semilandmarks. The largest correlations are found across semilandmarking densities arising from TPS&NICP and the smallest, from sliding TPS, especially at lower semilandmarking densities.

Table 2.8 Pearson correlations among vectors of Procrustes distances and Mantel tests comparing Procrustes distance matrices between each density of semilandmarking and the maximum density.

	200		400		600		800	
	Pearson	Mantel	Pearson	Mantel	Pearson	Mantel	Pearson	Mantel
Sliding TPS	0.9029	0.9075	0.9580	0.9597	0.9829	0.9832	0.9886	0.9898
LS&ICP	0.9919	0.9893	0.9975	0.9969	0.9981	0.9977	0.9991	0.9990
TPS&NICP	0.9978	0.9976	0.9994	0.9993	0.9994	0.9994	0.9998	0.9997

Additionally, the Pearson correlations among vectors of Procrustes distances between each individual and the mean and the Mantel correlations among distance matrices were computed between distances derived from each semilandmarking approach and density and the distances based on landmarks alone are computed (Table 2.9). The TPS&NICP approach results in distances that are most highly correlated with those from landmarks and the sliding TPS approach results in distances that are a little less strongly correlated while the LS&ICP approach results in distances that are poorly correlated with those from landmarks alone. For LS&ICP and TPS&NICP approaches, but not sliding TPS, the correlations become smaller with increasing semilandmarking density.

Table 2.9 Pearson and Mantel correlations between vectors and matrices of Procrustes distances from each semilandmarking approach and density and those from landmarks alone. Sliding TPS (first row), LS&ICP (middle row), and TPS&NICP (bottom row).

	200		400		600		800		1000	
	Pearson	Mantel								
Sliding TPS	0.6424	0.5718	0.6604	0.5826	0.6729	0.5910	0.6928	0.6084	0.6848	0.6115
LS&ICP	0.3234	0.3164	0.2677	0.2634	0.2549	0.2472	0.2469	0.2378	0.2342	0.2260
TPS&NICP	0.7315	0.7035	0.7237	0.6976	0.7221	0.6949	0.7222	0.6952	0.7208	0.6944

2.3.1.4 PCA and allometry

2.3.1.4.1 Correlations between PC scores from different semilandmarking methods and densities

To broadly compare distributions of specimens arising from different semilandmarking approaches and densities, correlations (ignoring sign, and so arbitrary reflections on PCs) were computed between scores of individuals on the first two PCs of shape derived from separate GPA/PCA of the landmark and semilandmark configurations (Table 2.10). These first two PCs account for greater than 30% of the total

shape variance in each analysis of landmarks and semilandmarks from sliding TPS and LS&ICP, and 49% for those from TPS&NICP. Correlations are generally moderate to weak. They are low for comparisons of PC1 and PC2 scores between LS&ICP and TPS&NICP and become smaller for comparisons of PC 1 scores between LS&ICP and sliding TPS with increasing density. In contrast, PC1 and PC2 correlations from sliding TPS and TPS&NICP are low for lower densities, rising for higher ones.

Table 2.10 Pearson correlations between PC1 and PC 2 scores derived using different semilandmarking approaches. Sliding TPS vs LS&ICP (first row), Sliding TPS vs TPS&NICP (middle row), and LS&ICP vs TPS&NICP (bottom row).

200		400		600		800		1000	
PC1	PC2								
0.6246	0.0788	0.5441	0.6581	0.4280	0.8798	0.3951	0.7776	0.3804	0.5059
0.2829	0.4818	0.4947	0.0921	0.7287	0.2434	0.8004	0.5523	0.7954	0.6903
0.1436	0.1030	0.1621	0.0848	0.1727	0.0713	0.1718	0.0411	0.1735	0.0085

Within each semilandmarking approach, the correlations among PC1 and PC2 scores from every density and the maximum density of each semilandmarking approach are presented in Table 2.11. The correlations from LS&ICP and TPS&NICP were consistently large, whereas the correlations from sliding TPS increase with increasing semilandmark densities. Table 2.12 details the extent to which PC scores for the analyses of Table 2.11 correlate with those from PCA of landmarks alone. For TPS&NICP landmarks and semilandmarks, at all densities, PC1 scores show a consistent correlation of approximately 0.89, and PC2 scores, ~0.49. In contrast, PC scores based on landmarks and semilandmarks from sliding TPS, show a lower correlation of PC1 ~0.55, except at the lowest density of semilandmarking where the correlation is 0.2437. PC2 correlations are uniformly low. LS&ICP semilandmarking at all densities results in PC1 and PC2 scores with uniformly low correlations.

Table 2.11 Pearson correlations of PC1 and PC 2 scores between each semilandmark density and the maximum (1000) density.

	200		400		600		800	
	PC1	PC2	PC1	PC2	PC1	PC2	PC1	PC2
Sliding TPS	0.4697	0.5826	0.7051	0.1701	0.9532	0.6054	0.9815	0.9131
LS&ICP	0.8661	0.8934	0.9469	0.9534	0.9799	0.9805	0.9942	0.9938
TPS&NICP	0.9989	0.9988	0.9994	0.9992	0.9996	0.9995	0.9998	0.9998

To assess the extent to which scalp semilandmarks, which lack nearby fixed landmarks to guide their placement, affect the relative performance of semilandmarking approaches, a set of PCAs was carried out using facial landmarks and semilandmarks alone. Table 2.13 presents the resulting correlations among PC scores extracted using each semilandmarking approach at each density of semilandmarking. These first two PCs

account for almost 40% of the total shape variance in each analysis. Those between sliding TPS and TPS&NICP are consistently high (>0.9), while correlations between LS&ICP and both sliding TPS and TPS&NICP are consistently low.

Table 2.12 Pearson correlations of PC1 and PC 2 scores from landmarks alone and those from each semilandmarking approach and density. Sliding TPS (first row), LS&ICP (middle row), and TPS&NICP (bottom row).

200		400		600		800		1000	
PC1	PC2								
0.2437	0.2256	0.4119	0.0538	0.5586	0.0485	0.6165	0.1081	0.5782	0.2153
0.0385	0.1486	0.0422	0.1751	0.0342	0.1821	0.0222	0.1929	0.0240	0.2085
0.8897	0.4997	0.8844	0.4937	0.8847	0.4913	0.8877	0.4906	0.8932	0.4931

Table 2.13 Pearson correlations between PC1 and PC 2 scores based on facial landmarks and semilandmarks derived using different semilandmarking approaches. Sliding TPS vs LS&ICP (first row), Sliding TPS vs TPS&NICP (middle row), and LS&ICP vs TPS&NICP (bottom row).

200		400		600		800		1000	
PC1	PC2								
0.2112	0.5431	0.2130	0.4862	0.2095	0.4510	0.2245	0.3525	0.2445	0.2261
0.9617	0.9388	0.9520	0.9258	0.9525	0.9338	0.9603	0.9351	0.9654	0.9337
0.0859	0.4849	0.0626	0.4541	0.0565	0.4235	0.0912	0.3150	0.1310	0.1869

In Table 2.14, for each semilandmarking approach, the correlations are presented between scores on PC1 and PC2 calculated using maximum semilandmarks and each lower semilandmarking density on the face. These correlations are consistently high for the sliding TPS and TPS&NICP approaches, while they are consistently low for the LS&ICP approach. Table 2.15 presents the correlations between PC1 and PC2 scores computed using the facial landmarks alone and those from the set of landmarks and semilandmarks over the face derived from each semilandmarking approach and density. TPS&NICP derived PC1 scores are consistently highly correlated with those from the landmarks and PC2 scores are less strongly, but still highly correlated. The PC scores from the sliding TPS approach show intermediate levels of correlation with those from landmarks and the LS&ICP approach present the lowest correlation.

Table 2.14 Correlations of PC1 and PC 2 scores between each facial landmark and semilandmark configuration at lower densities of semilandmarks and the configuration with the maximum density.

	200		400		600		800	
	PC1	PC2	PC1	PC2	PC1	PC2	PC1	PC2
Sliding TPS	0.9861	0.9855	0.9941	0.9943	0.9944	0.9931	0.9982	0.9976
LS&ICP	0.8549	0.8663	0.9149	0.9196	0.9429	0.9447	0.9848	0.9846
TPS&NICP	0.9971	0.9915	0.9996	0.9991	0.9998	0.9994	0.9998	0.9993

Table 2.15 Correlations of PC1 and PC 2 scores derived from the landmarks alone and each facial landmark and semilandmark configuration generated by different approaches and densities. Sliding TPS (first row), LS&ICP (middle row), and TPS&NICP (bottom row).

200		400		600		800		1000	
PC1	PC2								
0.8371	0.5124	0.7841	0.3921	0.7775	0.4043	0.8021	0.3888	0.8159	0.4043
0.1129	0.2667	0.1775	0.2187	0.2052	0.1938	0.1921	0.1522	0.1755	0.1097
0.9193	0.7084	0.9023	0.6413	0.8955	0.6351	0.8997	0.6233	0.8971	0.6407

2.3.1.4.2 Comparison of allometric vectors

To assess the effects of semilandmarking approach on allometric vectors at each density, the landmark and semilandmark configurations from each approach were submitted to a joint GPA followed by PCA. Angles, as a measure of dissimilarity, were then computed among allometric vectors estimated by multivariate regression of shape (scores on all PCs) on the natural logarithm of centroid size (Table 2.16). Angles exceed 90° for some comparisons because the allometric vectors were compared respecting their polarity (i.e. from small to large). At all semilandmarking densities, the angles between sliding TPS and TPS&NICP derived vectors were smaller than those in comparisons involving LS&ICP. This indicates that the allometric vector resulting from semilandmarking using the LS&ICP approach is rather different from those derived from sliding TPS and TPS&NICP. Meanwhile, the angles between allometric vectors from sliding TPS and TPS&NICP become smaller with increasing semilandmark densities, whereas they increase in the comparisons with LS&ICP.

Table 2.16 The angles (°) between allometric vectors from different semilandmarking approaches and densities.

	200	400	600	800	1000
Sliding TPS LS&ICP	87.39	90.21	92.37	93.44	93.68
Sliding TPS TPS&NICP	32.47	29.57	26.29	24.00	23.12
LS&ICP TPS&NICP	92.76	97.60	98.64	100.01	100.32

2.3.1.4.3 Comparison of allometric scaling of landmark and semilandmark configurations

We compared the predicted landmark and semilandmark configurations representing the extreme limits (maximum and minimum centroid sizes) of each allometric vector derived from each semilandmarking method by computing Procrustes distances between them. The results are presented in Table 2.17. The Procrustes distances among predictions are smallest and converge with increasing density for comparisons between predicted shapes from sliding TPS and TPS&NICP. For these comparisons, the distances among

predicted shapes at minimum size (~0.024) are approximately double those at maximum size (~0.012) because the distribution of centroid sizes is skewed towards the maximum. These distances are approximately 70% of the average Procrustes distance to the mean for the comparisons of predicted shapes at minimum centroid size and 35% at maximum centroid size. Distances between predicted shapes from LS&ICP and those from the other semilandmarking approaches are generally greater.

Table 2.17 Comparison of Procrustes distances between the predicted shapes corresponding to the maximum (Max) and minimum (Min) centroid size derived using semilandmarking approaches and densities. Sliding TPS vs LS&ICP (first row), Sliding TPS vs TPS&NICP (middle row), and LS&ICP vs TPS&NICP (bottom row).

200		400		600		800		1000	
Max	Min								
0.0298	0.0494	0.0288	0.0495	0.0281	0.0504	0.0278	0.0520	0.0283	0.0529
0.0128	0.0260	0.0131	0.0257	0.0125	0.0231	0.0118	0.0222	0.0117	0.0217
0.0352	0.0628	0.0355	0.0653	0.0347	0.0646	0.0347	0.0659	0.0350	0.0668

Taken together, these findings indicate that predicted landmark and semilandmark configurations from allometric analysis are most similar between the sliding TPS and TPS&NICP approaches and least between LS&ICP and the other approaches.

2.3.2 Ape crania

2.3.2.1 The locations of semilandmarks

We repeated a subset of the foregoing analyses using ape cranial surface meshes. Ape crania vary more in form (shape and size) and have a more complex surface than the head scans used in the previous analyses. They present identifiable landmarks over the whole surface, unlike the head surface data which offer no identifiable scalp landmarks (Figure 2.1). Figure 2.9 shows the average locations of 800 semilandmarks generated by sliding TPS (red points), LS&ICP (red points), and TPS&NICP (green points) approaches on the mean surface generated by deforming the template cranium to the mean landmarks and semilandmarks derived from the sliding TPS approach. Semilandmarks generated by LS&ICP tend to be located in different positions to those from the other approaches. Additionally, the mean configurations from the TPS&NICP and particularly LS&ICP do not exactly lie on the sliding TPS mean surface. These differences are particularly evident for semilandmarks around and in the orbits, temporal fossae, over the brow ridges, zygomatic arch and maxilla.

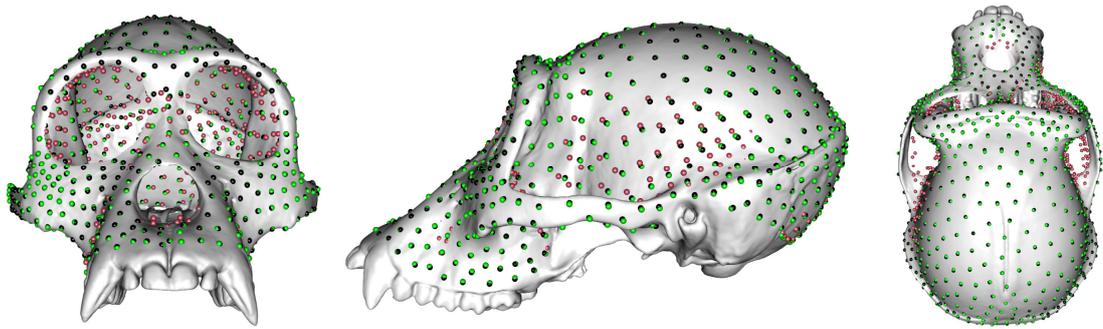


Figure 2.9 800 semilandmarks generated by different semilandmarking approaches on the mean surface generated by sliding TPS.

Table 2.18 shows the differences in locations of semilandmarks generated by sliding TPS and TPS&NICP approaches. It lists average differences in location (dev in mm) and the percentage of semilandmarks (%) that differ in location by 0.0-1.0 mm, 1-2.5 mm, 2.5-5.0 mm and ≥ 5.0 mm. Figure 2.10a, illustrates these differences. Between sliding TPS and TPS&NICP approaches, differences are all less than 5.0 mm, with the majority (>99%) less than 2.5 mm (Figure 2.10a) and the proportion of semilandmarks from sliding TPS and TPS&NICP located within 1.0 mm of each other tends to decrease with increasing semilandmarking density. By contrast, the semilandmark locations derived from LS&ICP are more different from those derived by both sliding TPS and TPS&NICP. For brevity, only the differences in semilandmarks between sliding TPS and LS&ICP (Table 2.19 and Figure 2.10b; 98.62% ≥ 5.0 mm) are presented, but the results are similar for the comparison of TPS&NICP with LS&ICP.

Table 2.18 Comparison of semilandmarks from sliding TPS and TPS&NICP approaches.

diff. mm	50		100		200		400		800	
	dev	%								
[0.0-1.0)	0.56	52.00	0.64	46.00	0.70	35.50	0.68	30.25	0.70	33.75
[1.0-2.5)	1.34	48.00	1.44	54.00	1.47	64.00	1.49	69.50	1.50	66.13
[2.5-5.0)	-	-	-	-	2.99	0.5	2.57	0.25	2.73	0.12
≥ 5.00	-	-	-	-	-	-	-	-	-	-
Total	0.94	100.00	1.08	100.00	1.21	100.00	1.25	100.00	1.23	100.00

Table 2.19 Comparison of semilandmarks from sliding TPS and LS&ICP approaches.

diff. mm	50		100		200		400		800	
	dev	%	dev	%	dev	%	dev	%	dev	%
[0.0-1.0)	-	-	-	-	-	-	-	-	-	-
[1.0-2.5)	-	-	-	-	-	-	-	-	-	-
[2.5-5.0)	4.93	2.00	4.54	4.00	4.58	5.50	4.82	1.25	4.80	1.38
≥ 5.00	9.04	98.00	9.89	96.00	10.41	94.50	11.05	98.75	11.52	98.62
Total	8.96	100.00	9.68	100.00	10.09	100.00	10.97	100.00	11.37	100.00

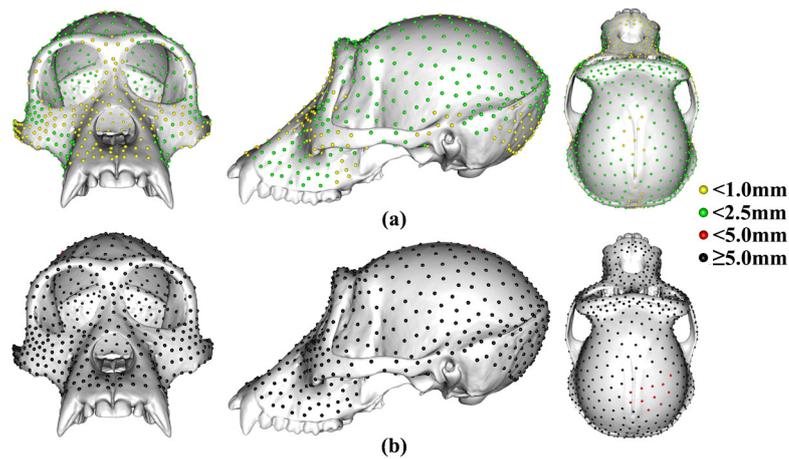


Figure 2.10 The average differences in location (mm) between 800 semilandmarks generated by different approaches. (a) Differences between sliding TPS and TPS&NICP approaches. (b) Differences between sliding TPS and LS&ICP approaches. Differences between TPS&NICP and LS&ICP approaches are not shown because they are very similar to those in b.

2.3.2.2 Differences among Mean Landmark and Semilandmark Locations

To compare the estimates of mean landmark and semilandmark configurations, Procrustes distances were computed among the mean landmarks and semilandmarks generated by different approaches (Table 2.20). The Procrustes distances between mean shapes from sliding TPS and TPS&NICP increase with increasing density, unlike the results using the head surface data (middle row, Table 2.5). These distances range from 0.0051 with 50 semilandmarks to 0.0072 with 800 semilandmarks, which are 4.87% and 7.62% of the average distance (computed using semilandmarks derived from sliding TPS plus landmarks) of specimens from the mean at each density.

These Procrustes distances among estimates of the mean are similar to those from the equivalent analyses of the head surface data (Table 2.5, middle row), while they are a much smaller proportion of the average distance of specimens from the mean compared to the equivalent percentages for the head surface data (~14.5%), reflecting the greater degree of variation in shape of the ape crania.

Table 2.20 Procrustes distance (dist) between mean landmark and semilandmark configurations derived at varying densities from sliding TPS and TPS&NICP approaches.

	50	100	200	400	800
dist	0.0051	0.0061	0.0072	0.0067	0.0072

2.3.2.3 Comparison of centroid sizes and Procrustes distance matrices

The differences in semilandmark locations derived using LS&ICP from those derived by the other two approaches are emphasised by the analyses of Table 2.21. This table presents the Pearson correlations among the vectors of Procrustes distances between every individual and the mean as well as the correlations, from Mantel tests, comparing the matrices of Procrustes distances among all individuals, calculated using landmarks and semilandmarks generated by different semilandmarking approaches and densities.

The largest (all >0.99) Pearson and Mantel correlations were found between sliding TPS and TPS&NICP. LS&ICP clearly produces quite different results to those obtained using the other semilandmarking approaches, especially at higher semilandmarking densities. For this reason, we focus on comparison of the results from analyses of landmarks and semilandmarks from sliding TPS and TPS&ICP.

Table 2.21 Pearson correlations among the vectors of Procrustes distances between each ape cranium and the mean and Mantel tests among Procrustes distance matrices. Sliding TPS vs LS&ICP (first row), Sliding TPS vs TPS&NICP (middle row), and LS&ICP vs TPS&NICP (bottom row).

50		100		200		400		800	
Pearson	Mantel								
0.7630	0.8024	0.6789	0.7403	0.5954	0.6711	0.5066	0.5993	0.4185	0.5241
0.9988	0.9986	0.9962	0.9970	0.9959	0.9961	0.9951	0.9947	0.9948	0.9944
0.7540	0.7929	0.6643	0.7268	0.5806	0.6561	0.4739	0.5761	0.3881	0.5050

Table 2.22 presents Pearson correlations between the distance vectors and matrices calculated using the full set of landmarks alone, and those from the landmarks and semilandmarks generated by each semilandmarking approach and density. Distances from sliding TPS and TPS&NICP approaches are similarly correlated with those from landmarks alone. Thus, these correlations are ~0.96 for the lowest density of semilandmarking and fall gradually to ~0.9 for the highest. In contrast correlations for the LS&ICP approach are moderate (~0.5-0.75), but follow the same trend by becoming smaller with increasing semilandmarking density.

Table 2.22 Pearson and Mantel correlations between vectors and matrices of Procrustes distances from each semilandmarking approach and density and those from the landmarks alone. Sliding TPS (first row), LS&ICP (middle row), and TPS&NICP (bottom row).

50		100		200		400		800	
Pearson	Mantel								
0.9619	0.9579	0.9460	0.9401	0.9303	0.9244	0.9160	0.9079	0.9105	0.8995
0.7424	0.7951	0.6627	0.7532	0.6081	0.7153	0.5413	0.6742	0.4916	0.6402
0.9602	0.9551	0.9473	0.9391	0.9350	0.9241	0.9260	0.9135	0.9221	0.9076

Centroid sizes and Procrustes distances between each cranium and the mean were very similar between landmark and semilandmark configurations from sliding TPS and TPS&NICP, as indicated in Figures 2.11 and 2.12, where the fitted lines are almost coincident with the dashed lines, thereby indicating identity. However, LS&ICP tends to produce landmarks and semilandmark configurations with larger centroid sizes than those from sliding TPS for small crania, and smaller for large crania (Figure 2.11). Comparisons of Procrustes distances from the mean derived from LS&ICP and sliding TPS (Figure 2.12b) at varying semilandmark densities show marked differences, unlike comparisons between the same distances from TPS&NICP and sliding TPS.

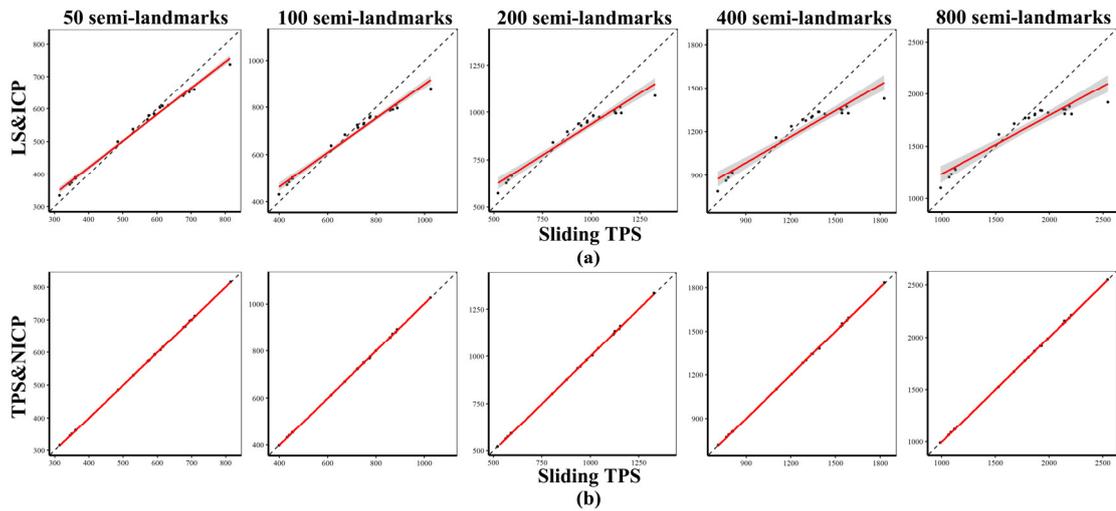


Figure 2.11 Comparisons of centroid sizes estimated by different semilandmarking approaches. (a) Sliding TPS vs LS&ICP. (b) Sliding TPS vs TPS&NICP.

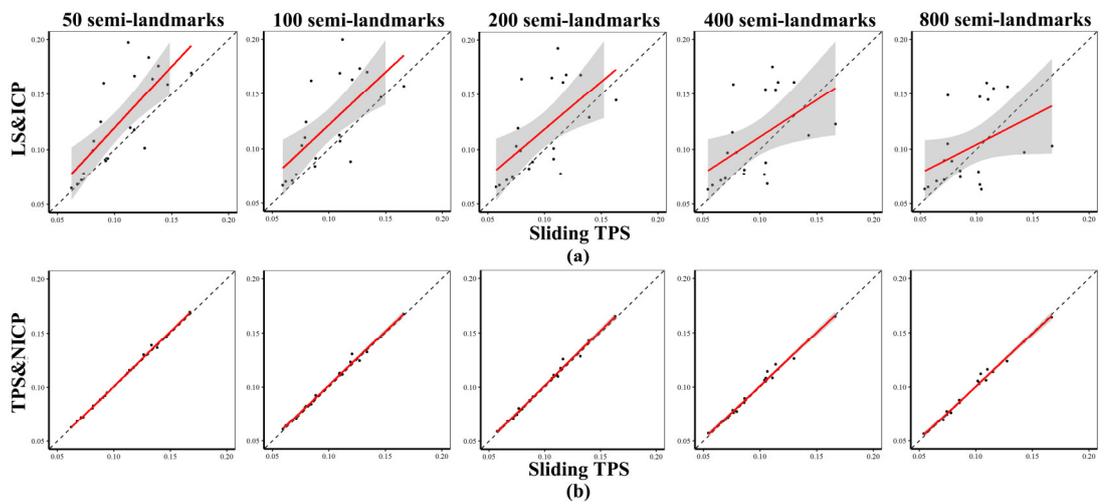


Figure 2.12 Comparison of the vector of Procrustes distances from the mean between semilandmarking approaches. (a) Sliding TPS vs LS&ICP. (b) Sliding TPS vs TPS&NICP.

2.3.2.4 PCA and allometry

Additionally, we calculated the correlations of scores on the first two PCs of shape variation resulting from the separate GPA and then PCA of each semilandmark configuration and density (Table 2.23). These are all greater than 0.99, considerably greater than the values obtained for the same comparisons using the head surface data (middle row, Table 2.10). The scatterplots of the first two PCs reflect this and account for ~60% of the total shape variance. Examples are superimposed in Figure 2.13.

Table 2.23 Comparison of Pearson correlations of PC1 and PC 2 between Sliding TPS and TPS&NICP.

50		100		200		400		800	
PC1	PC2								
0.9993	0.9989	0.9983	0.9992	0.9978	0.9985	0.9974	0.9977	0.9972	0.9967

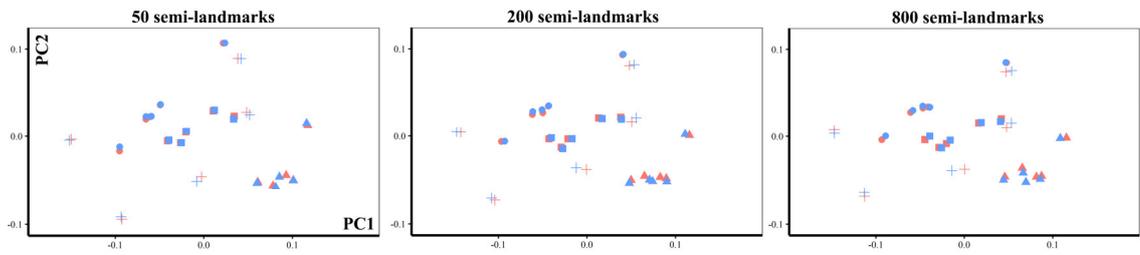


Figure 2.13 Visualisation of superimposed scatterplots of PC1 and PC 2 from analyses of 20 ape crania using landmarks and semilandmarks from sliding TPS and TPS&NICP approaches with varying densities. Horizontal axis represents PC 1 and vertical represents PC 2. Cross: *Pongo abeli*. Circle: *Gorilla*. Rectangle: *Pan troglodytes*; Triangle: *Hylobates lar*. Red: Sliding TPS. Cyan: TPS&NICP.

Table 2.24 presents the correlations between PC scores on the first PCs of shape from each landmark and semilandmark configuration generated by semilandmarking approaches and densities and those from GPA and PCA of the landmarks alone. These are very similar for both semilandmarking approaches, being large to moderate and slightly greater for PC1 than PC2 scores, decreasing ~5-7.5% with increasing semilandmarking density. Additionally, the correlations of PC1 and PC2 scores between each semilandmarking density and the maximum (800) density are presented in Table 2.25. This shows that the correlations within each approach are high, and increase with increasing density.

Table 2.24 Pearson correlations among PC scores from each semilandmarking density and from the landmarks alone. Sliding TPS (first row), and TPS&NICP (bottom row).

50		100		200		400		800	
PC1	PC2								
0.9487	0.8634	0.9238	0.8435	0.9104	0.8188	0.9021	0.8155	0.8963	0.8077
0.9430	0.8732	0.9186	0.8434	0.8999	0.8020	0.8880	0.8108	0.8833	0.8012

Table 2.25 Pearson correlations among PC scores from each semilandmarking density and the maximum density.

	50		100		200		400	
	PC1	PC2	PC1	PC2	PC1	PC2	PC1	PC2
Sliding TPS	0.9888	0.9769	0.9971	0.9900	0.9987	0.9963	0.9996	0.9984
TPS&NICP	0.9876	0.9765	0.9962	0.9908	0.9989	0.9965	0.9998	0.9993

The angles between allometric vectors derived by multivariate regression of shape (the full set of non-zero PC scores) on the natural logarithm of centroid size using landmarks and semilandmarks from sliding TPS and TPS&NICP (Table 2.26) are generally small (< 9°) and increase moderately with semilandmarking density.

Table 2.26 The angles (°) between allometric vectors derived using sliding TPS and TPS&NICP approaches.

	50	100	200	400	800
Sliding TPS					
TPS&NICP	6.52	7.67	8.68	8.97	8.73

The Procrustes distances between predicted landmark and semilandmark

configurations at the extreme limits of the allometric vector are shown in Table 2.27. The Procrustes distances between allometric predictions of cranial shape at the maximum centroid size are between a half and two thirds of those between predictions at the minimum centroid size. As with the head surface data, this is explained by the distribution of centroid sizes being skewed towards the maximum. These distances increase with semilandmarking density. They range between 9% and 14% of the average distance of specimens from the mean, for the predictions of cranial shape at the maximum centroid size and between 16% and 21% for predictions at the minimum centroid size.

Table 2.27 Procrustes distances between the predicted landmark and semilandmark configurations from sliding TPS and TPS&NICP corresponding to the maximum (Max) and minimum (Min) centroid sizes.

	50	100	200	400	800
Max	0.0095	0.0113	0.0122	0.0124	0.0134
Min	0.0168	0.0189	0.0211	0.0214	0.0202

2.4 Discussion

This study compares alternative strategies for marking up dense point correspondences (semilandmarks) among biological structures for subsequent statistical analyses. We compared these semilandmarking approaches by empirically testing six hypotheses (i-vi, below) using surface scans of human faces and ape crania.

2.4.1 Tests of hypotheses

2.4.1.1 Hypothesis i

Taking each hypothesis in turn, i) regarding semilandmark locations, Tables 2.1-2.3 show that for the head surfaces, at all densities of semilandmarking, sliding TPS and TPS&NICP result in smaller differences in semilandmark locations ($\sim 34\text{-}45\% < 2.5$ mm, Table 2.1) than between these methods and LS&ICP, where all differences are greater than 2.5 mm (Tables 2.2-2.3). The least differences between sliding TPS and TPS&NICP are found at higher semilandmarking densities and where true landmarks are nearby (Figure 2.4). For the ape cranial surfaces, sliding TPS and TPS&NICP, produce similar semilandmark locations, with more than 99% lying within 2.5 mm of their corresponding one (Table 2.18). Larger differences in semilandmark locations are found between LS&ICP and sliding TPS (94%-98% ≥ 5.0 mm apart; Table 2.19).

These findings falsify the hypothesis of no difference. However, beginning with the same template, sliding TPS and TPS&NICP produce the most similar semilandmark locations among the three methods, especially at higher semilandmarking densities and where landmarks are nearby. LS&ICP produces semilandmark locations that differ more from those derived from sliding TPS and TPS&NICP, than they differ from each other, especially for the more complex ape cranial surfaces. It should be noted that differences in semilandmark locations do not in themselves indicate that one or more methods fails to identify homologous semilandmarks, rather they could be identifying different

homologous semilandmarks. However, it is unlikely that any approach identifies semilandmarks that are homologous, because they are constructed by algorithms that do not take account of evolutionary or developmental history, except in the sense that homologous landmarks may be used to guide their placement and only to the extent that the landmarks are indeed homologous.

2.4.1.2 Hypothesis ii

With regard to hypothesis ii), that there is no difference between mean landmark and semilandmark configurations estimated using different semilandmarking approaches, from Table 2.4 it is evident that within each density of semilandmarking, there are negligible differences in estimates of the centroid size of the mean configuration of landmarks and semilandmarks from the head data. However, configuration shapes do differ, as shown in Table 2.5, where the Procrustes distances between the means from sliding TPS and TPS&NICP are consistently in the range 0.0048 to 0.005 irrespective of semilandmark density. In contrast means from LS&ICP become more similar to the means estimated by the other approaches with increasing semilandmarking density. Differences in estimates of the mean shape not inconsiderable, particularly where the comparisons involve means estimated by LS&ICP.

Table 2.20 presents Procrustes distances between estimates of the mean ape cranial landmark and semilandmark configuration derived from sliding TPS and TPS&NICP. These increase with increasing density, unlike the results based the head surface data (middle row, Table 2.5). The reason for this difference between analyses of the two datasets requires further investigation, but likely relates to differences in surface complexity among the two datasets. They range from 4.87% to 7.62% of the average distance of specimens from the mean at each density, and are smaller than the equivalent percentages for the head surfaces, likely reflecting the greater variance (and so average distances from the mean) of the ape cranial surfaces.

The differences between estimates of the mean are the result of differences in semilandmark locations achieved by each method.

2.4.1.3 Hypothesis iii

With regard to hypothesis iii), Figure 2.6 for the head surfaces and Figure 2.11 for the ape cranial surfaces plot the centroid sizes of configurations from sliding TPS against those from TPS&NICP as well as those from sliding TPS against those from LS&ICP. For both sets of surfaces (Figures 2.6 and 2.11), the plots indicate that the centroid sizes from sliding TPS are very similar to those of TPS&NICP at all assessed semilandmarking densities, while those from LS&ICP are larger than those from sliding TPS at low semilandmark densities and smaller at higher densities. The plot of TPS&NICP vs LS&ICP (not shown) is similar to that of sliding TPS vs LS&ICP. This

falsifies the hypothesis that centroid sizes of configurations do not differ between semilandmarking approaches, because the centroid sizes produced by LS&ICP differ somewhat from those generated by the other semilandmarking approaches. However, sliding TPS and TPS&NICP approaches result in configurations with very similar centroid sizes at all semilandmarking densities.

2.4.1.4 Hypothesis iv

Considering the impact of different semilandmarking approaches on the distribution of specimens, hypothesis iv) was tested by quantifying the extent to which distance matrices differ. The relative Procrustes distances among specimens were compared using Pearson correlations of vectors of Procrustes distances to the mean and Mantel correlations, computed among entire Procrustes distance matrices derived using different semilandmarking approaches and densities. Additionally, the vectors of distances from the mean were plotted against each other to visualise their associations.

For the head surfaces, the vectors of distances to the mean show differences between semilandmarking approaches. Thus, in Figure 2.7, at every semilandmarking density, estimates of Procrustes distances are largest for configurations using semilandmarks computed by TPS&NICP, and least for those computed using LS&ICP. The correlations among distance vectors and matrices from sliding TPS and TPS&NICP approximately range from 0.7 to 0.85 and become greater with increasing numbers of semilandmarks (Table 2.7). The vectors and matrices from both sliding TPS and TPS&NICP show smaller correlations with those from LS&ICP and these become smaller with increasing numbers of semilandmarks (Table 2.7). The plots of Figure 2.8 support these findings, and show that at all semilandmarking densities, TPS&NICP and sliding TPS result in vectors of distances from the mean which are highly correlated, although distances are larger when computed from semilandmarks derived by TPS&NICP than those from sliding TPS. As Procrustes distances between specimens increase, those from TPS&NICP increase at a faster rate than those from sliding TPS. From Figure 2.8 it is also apparent that correlations between vectors from sliding TPS and LS&ICP are smaller and that as Procrustes distances from the mean increase, those from LS&ICP do so at a lower rate than sliding TPS.

For the ape crania (Table 2.21) all distance vector and matrix correlations among semilandmarking approaches are larger than those for the heads (Table 2.7). Correlations among distance vectors and matrices from sliding TPS and TPS&NICP become slightly smaller with increasing numbers of semilandmarks, while the correlations between these vectors and matrices with those from LS&ICP reduce more markedly with increasing numbers of semilandmarks (Table 2.21). These findings are reflected by the plots of Figure 2.12, which compares vectors of Procrustes distances computed using different semilandmarking densities. These indicate that, at all densities of semilandmarking, the

vectors of distances are almost identical between TPS&NICP and sliding TPS, reflecting the very high correlations in Table 2.21, while between sliding TPS and LS&ICP there are greater differences in both the relative distances (bigger scatters) and in how the vectors differ between semilandmarking methods with increasing semilandmarking density (varying slopes and positions of regression lines).

These results falsify the hypothesis of no differences among distance matrices derived using different semilandmarking approaches. The most similar vectors of distances from the mean and distance matrices are found between sliding TPS and TPS&NICP, with those from LS&ICP being most distinctive. Further, they indicate that the extent to which different semilandmarking approaches result in different distance matrices depends on the density of semilandmarks and on the landmark distribution over surfaces, as well as the complexity and degree of variation among surfaces and the distribution.

Beyond this, analyses were conducted using the head surface data to assess, within each semilandmarking approach, if distance matrices converge with increasing semilandmark density. Table 2.8 presents, for the head surfaces, the results of comparisons of the distance matrices using 1000 semilandmarks with those with fewer semilandmarks for each method of semilandmarking. All correlations are high, with distance matrices from all densities below 1000 semilandmarks correlating >0.99 , and increasing slightly from low to high densities, for semilandmarks derived from TPS&NICP and LS&ICP. For semilandmarks derived from sliding TPS, the distance vector and matrix from 200 semilandmarks show correlations >0.90 , rising progressively with increasing density to >0.988 with 800 semilandmarks. Thus, slight converge of distance matrices, from a high initial level of similarity, occurs with both TPS&NICP and LS&ICP. Sliding TPS also shows convergence of distance matrices and vectors, but from an initially lower correlation. Thus, convergence does occur, most markedly for sliding TPS, while the other two approaches are more consistent across semilandmarking densities.

For the head surface data, Table 2.9 tabulates the correlations between the vectors from the mean, the distance matrices calculated using the landmarks alone and those based on the landmarks and semilandmarks generated by each semilandmarking approach. The same calculations for the ape crania are presented in Table 2.22. Correlations are much higher for the ape crania, indicating that the semilandmarks and landmarks more closely reflect the landmark data. This may be because in the head surface data a substantial region, the scalp, is not landmarked, and so variations in configuration shape over the scalp are unaccounted for by the landmark data alone, while landmarks are distributed over the whole surface of the ape crania, as are semilandmarks. Additionally, it may be because the ape crania represent several species and differ much more than the

human heads. Correlations with landmark data fall with increasing semilandmarking density in all cases, for both surface datasets, except for sliding TPS in the head surface data (Table 2.9, row 1). The fall in strength of correlation may indicate that increased density of semilandmarking results in additional aspects of shape variation, not described by landmarks alone, being captured by the semilandmarks. However, the decrease in strength of correlation is slight to moderate for TPS&NICP for both head and ape data and for sliding TPS with the ape data (<5%), while for LS&ICP it is relatively much greater (~40% for the head surface data and ~15% for the ape crania). It is not clear if the decreases in strength of correlation occur because higher semilandmarking densities describe additional shape variation or additional error. However, in the case of LS&ICP, the anatomical placement of semilandmarks is clearly very different from the template and from those arising from the other semilandmarking approaches in the ape crania (Figures 2.9 and 2.10) and to a lesser extent, the head surface data (Figures 2.3 and 2.4). For sliding TPS, applied to the head surface data, correlations with distances from the landmark data increase slightly (by ~6%) between the lowest and highest semilandmarking densities, in contrast to the other semilandmarking approaches. Landmarks are not present over the scalp in this dataset and so, scalp variations are not accounted for by the landmark data. It is likely that this, and the consequent inadequate control of sliding over the scalp have led to the difference in result with sliding TPS.

2.4.1.5 Hypothesis v

Further analyses tested the hypothesis (v) that PCs of shape variation arising from analyses of landmarks and semilandmarks derived by different methods and at different densities do not differ. PCA is an exploratory technique whose results are prone to change with changes in included specimens or data. However, it is commonly used to assess relationships among specimens based on plots of PCs. For this reason, PC scores from separate GPA and PCA of each landmark and semilandmark set were compared among landmark and semilandmark sets generated by different methods and densities of semilandmarking.

For the head surfaces, Table 2.10 presents the correlations among PC1 and PC2 scores from analyses of landmark and semilandmark configurations from different semilandmarking approaches and with different densities of semilandmarks. These are low to moderate and tend to increase for the comparisons between sliding TPS and TPS&NICP, but vary less predictably for the comparisons between LS&ICP and the other approaches. Additionally, within each semilandmarking approach PC1 and PC2 scores converge on the PC scores obtained using 1000 semilandmarks as semilandmarking density increases (Table 2.11). This convergence is most marked for sliding TPS and least for TPS&NICP, which shows very high correlations for all densities (>0.99). These findings indicate that PCA results are sensitive to both semilandmarking approach and

density, with TPS&NICP producing the most stable results with varying semilandmark density.

Table 2.12 compares PC scores derived from the landmarks alone with those from the different semilandmarking approaches and densities. The results indicate that correlations are moderate or low for all comparisons except for those between PC1 scores from TPS&NICP and landmarks alone. TPS&NICP therefore most consistently replicates the distances that arise from analysis of the landmarks alone, while LS&ICP results in the biggest differences. Sliding TPS results in differences that are intermediate. This is likely because scalp semilandmark sliding is inadequately controlled. This was investigated by carrying out further analyses omitting scalp semilandmarks.

The results of analyses of facial landmarks and semilandmarks in Tables 2.13-2.15 differ from those using the full set of facial landmarks and neurocranial semilandmarks presented in Tables 2.10-2.12, in particular the sliding TPS approach shows much higher correlations with results from TPS&NICP (Table 2.13) and landmarks alone (Table 2.15). Additionally, the PC scores from sliding TPS at lower semilandmarking densities are consistently highly correlated (~ 0.99) with those from landmarks and the maximum semilandmarks located on the face (Table 2.14), unlike when scalp semilandmarks were included (Table 2.11). These findings likely reflect the lack of fixed landmarks to control sliding over the neurocranium, which inevitably impacts the functioning of the sliding TPS algorithm. In the face, where landmarks are present, it achieves very similar results to the TPS&NICP approach. The TPS&NICP approach appears to be much less sensitive to the presence or absence of landmarks, reflecting that fact that this method does not rely on landmarks to control the eventual positions of semilandmarks (except in the initial registration and projection steps).

In contrast similar analyses using the ape crania show consistently very high (>0.99 in Table 2.23) correlations among PC1 and PC2 scores between landmark and semilandmark sets of varying density from sliding TPS and TPS&NICP. This is reflected in the superimposed PC plots of Figure 2.13, which show close correspondence between semilandmarking approaches at varying densities. For both of these semilandmarking methods, PC1 and PC2 scores correlate strongly (>0.8) with those from analyses of landmarks alone (Table 2.24). Additionally, analyses of surfaces from lower densities of semilandmarks are very highly correlated with those from 1000 semilandmarks (>0.97 in Table 2.25).

These findings falsify hypothesis (v), PC scores do change depending on the semilandmarking algorithm, and the density of semilandmarking. Greater consistency among sliding TPS and TPS&NICP is found in the analyses of ape crania than in those of the heads. This is, part because landmarks are scattered over the whole surface of the ape crania but are not present over the scalp in the head surface data, providing inadequate

control of the sliding TPS algorithm. Additionally, the ape crania are more variable and so the PCs are more strongly determined and less prone to change because of small changes in the semilandmark configurations. The correlations between PC scores from landmarks alone and those obtained using different semilandmarking approaches are very high for the ape crania (Table 2.24), which suggests that the semilandmarks add little additional information, but they are lower for the head surface data, plausibly because the landmarks are not located over the scalp. Scalp semilandmarks reflect additional aspects of variance that result in lower correlations with the landmark data. The results of Table 2.15, concerning facial landmarks and semilandmarks alone separately, lend support to this explanation.

2.4.1.6 Hypothesis vi

Finally, hypothesis (vi) that there are no differences in resulting estimates of allometric scaling of shape was tested. With the head surface data, as expected from the differences in semilandmark locations produced by each of semilandmarking approaches, the angles among allometric trajectories (Table 2.16) are very large ($\sim 90^\circ - 100^\circ$, respecting polarity of the vectors) for comparisons of LS&ICP with the other approaches and somewhat smaller between sliding TPS and TPS&NICP ($\sim 32^\circ - 24^\circ$). For the ape crania only the angles between the vectors from sliding TPS and TPS&NICP were computed and these are smaller than those found with the head surface data ($\sim 6^\circ - 9^\circ$), becoming greater with increasing semilandmark density (Table 2.26), whereas they decline among human heads (Table 2.16).

With regard to the differences in shape of the predicted individuals with the minimum and maximum centroid sizes from each semilandmarking approach, for both data sets, the Procrustes distance between them is greater for the comparisons of predicted shapes with minimum centroid size. This is because in both data sets the distribution of centroid size is skewed towards the maximum.

It would be of interest to more widely explore the consequences of different semilandmarking approaches in comparing allometric trajectories between sexes or species. This was not possible in the present study, which was limited to comparing predicted allometrically scaled mean shapes and the angles between allometric vectors derived using different semilandmarking approaches. These angles indicate that the semilandmarking approaches generate different results (because semilandmarks are in different locations), but they do not inform us about the effects of different semilandmarking approaches on the comparison of allometric trajectories between sexes or species, because no such comparison was carried it with these data.

2.4.2 Significance and implications of findings

In the present and previous studies (Boyer et al., 2015; Gonzalez et al., 2016; Rolfe

et al., 2021), differences are found in the locations of semilandmarks produced using different approaches and these have consequences of subsequent analyses. Thus, mean landmark co-ordinates, centroid sizes and distributions of specimens in Kendall's shape space are all impacted by the locations of semilandmarks. Just as different landmarking choices impact the results of subsequent analyses, so do variations in the number and locations of semilandmarks.

Previous studies have noted that with increasing semilandmark density, there is increasing consistency of scores on PC1 (Boyer et al., 2015; Vitek et al., 2017), while (Gonzalez et al., 2016) found that increasing density of semilandmarks did not necessarily result in greater group separation. These studies did not assess consistency of results with increasing density, among alternative semilandmarking methods. Here we find that the results generated by sliding TPS and TPS&NICP approaches are most consistent, especially where, as should always be the case with the sliding TPS algorithm, the surfaces are delimited by true landmarks. Greater differences are found between landmarks and semilandmarks from LS&ICP and the other two approaches especially for the more complex surfaces of the ape crania. For the head surface data, within semilandmarking approaches, the distributions of specimens as assessed by distance matrices and PCAs are most consistent across varying semilandmark densities for TPS&NICP and least for sliding TPS (Tables 2.8 and 2.11). Omitting the scalp, sliding TPS performs similarly to TPS&NICP when considering only the facial semilandmarks and the (facial) landmarks (Table 2.14). For the ape cranial data, sliding TPS and TPS&NICP produce distributions that are consistent within semilandmarking approaches across densities (Table 2.22). As such, TPS&NICP produces the most consistent results with both surface data sets, and between semilandmarking densities. Sliding TPS is almost as consistent, when applied to surfaces with landmarks over their entirety e.g. the ape crania and the facial region in the head surface data.

However, consistency does not relate to how well the homology map is represented by the resulting semilandmarks. Methods may be consistently wrong in identifying homology, and so, in describing differences. Here, for instance, both sliding TPS and TPS&NICP use a triplet of thin plate splines to achieve an initial fit between template and each specimen. In consequence these algorithms begin with initial placements of semilandmarks that are identical. This could well underlie why these two approaches achieve very similar results, rather than because they are both converge on 'the correct solution'. Each method estimates equivalent points in terms of its specific algorithm, but each estimate is different. In fact, all estimates of mean coordinate configurations, of the distance matrices and other statistical results are correct in each analysis, insofar as they are the correct results obtained from the landmarks and semilandmarks. Differences arise because of differences in the data; in the landmark and semilandmark locations. The issue

in studies that aim to describe and compare developmental or evolutionary transformations is which, if any, of the semilandmarking approaches correctly marks up homologies. How well the resulting semilandmarks represent homologies among specimens is limited by the extent to which knowledge of homology is applied in locating them, by the fact that such point homology is largely unknowable and, indeed, may not exist in reality because points at one stage may not actually turn into points at another. These considerations also apply to landmarks themselves, albeit arguably to lesser degree.

Semilandmarks have deficient coordinates, and so are located on the surface of interest, but with uncertainty regarding the equivalence (e.g. homology) of their position. Oxnard and O'Higgins (2009), Mitteroecker & Schaefer (2022) and Bastir et al. (2019) argue that the locations of the semilandmarks themselves should not be interpreted but rather the form of the surface mesh or curve that they describe should be the basis of comparison. This recognises their deficient coordinates in focusing on the form of the surface itself, however it also raises an important point and a question.

Thus, semilandmarks describe surfaces, but different semilandmarking approaches achieve this through different locations of semilandmarks. These differences in locations have very real effects on subsequent statistical analyses, here resulting in estimates of mean configurations, distributions and principal modes of shape variation (PCs) and covariation (e.g. allometry) that differ to some degree. This is an important point, because we rely on statistical results to test our hypotheses and yet, where these concern developmental or evolutionary transformations, the extent to which analyses of shape variation and covariation using any one method or density of semilandmarking respect and reflect homology is also unknowable.

Some insight into this, might be gleaned from a consideration of how well findings from analyses of landmarks and semilandmarks match findings based on presumed homologous points alone. Here several analyses have compared results obtained by different semilandmarking approaches and densities with those from landmarks alone. For ape crania, where landmarks are located, albeit sparsely, over the whole surface, Procrustes distances from both sliding TPS and TPS&NICP correlate strongly (>0.9) with distances from landmarks (Table 2.22). For the head surface data, correlations are weaker, but for the face alone, where landmarks are present, correlations (Table 2.15) are moderate to high. These findings suggest that analyses of landmarks and semilandmarks are consistent with those of landmarks alone, when landmarks are sufficient in number and located such that they delimit the surfaces that are to be semilandmarked. There are some consistent differences, which as noted above, may be due to better description (additional information) of surface form, or to shared error. However, given that high dimensional data, such as arises with semilandmarks presents serious analytical issues (Cardini et al., 2019), the potential benefits of semilandmarks in, particularly,

visualisation, should be set against the potential pitfalls of statistical analysis of such data (Cardini, 2020). The statistical gains are at best, unclear in the analyses presented here and there is an unresolvable doubt that the ‘gains’ may in fact not be gains at all, but rather due to erroneous identification of homologies. There may, however, be gains in applications to discrimination, identification and discrimination (Schlager & Rüdell, 2017; Shui et al., 2017), but these topics are not considered here, and further studies need to be conducted to assess this possibility.

With regard to visualisation, semilandmarks are often applied to enable detailed high quality representation of results as surface warping. Surface mesh form, rather than the form of a landmark and semilandmark configuration over the surface, is relevant in many practical circumstances. For example, surfaces are often visualised by warping a template to statistical estimates of e.g. mean landmark and semilandmark configuration form (Klingenberg, 2013). Beyond this, surfaces representing statistical results, such as the mean, might be used in the clinic to compare patient cranial form with that of the wider population, using clinic and condition specific (re)parameterisations of reference and patient surfaces (Duncan et al., 2022). Another increasingly common application of surfaces arising from geometric morphometric analyses is to use them to build finite element models (O’Higgins et al., 2019; O’Higgins et al., 2011). Thus, an important question arises, which is considered in a follow on study (Shui et al., 2023): how do different semilandmarking approaches perform in characterising the form, variation and covariations of the shape of the surface mesh itself, rather than the locations of semilandmarks on it?

What are the implications of this study for future work using semilandmarks? The results indicate that sliding TPS and TPS&NICP approaches produce very similar results in both data sets, and with the simpler surfaces of the head surface data, TPS&NICP produces the most consistent results between low and high density semilandmarks. As expected, sliding TPS performs less well where landmarks are inadequate to control their sliding. The LS&ICP approach appears least consistent with other approaches. However, as has been noted above, consistency does not necessarily indicate reliable identification of homologous points. This echoes (Cardini, 2020), who noted that consistency might be thought of as suggesting precision (repeatability of measures) but does not equate with accuracy (i.e. correctly marking up homologous points).

Our results cannot, support or refute the possibility that the consistency of results between sliding TPS and TPS&NICP reflects shared accuracy of identification of homologous semilandmarks. However, estimation of means depends on what is measured. They are a statistical, rather than biological entity and are ‘correct’ for the semilandmarks from each approach. The issue is not whether or not statistical results are correct, but rather if the semilandmarks are homologous in studies of transformations. The present

analyses have shown that differences in semilandmark locations (hypothesized homologies) among specimens will lead to differences in statistical results (Shui et al., 2023), and whether or not these differences are important depends on the context and the likelihood of reaching erroneous conclusions.

As such, it is not possible to state that any one method is superior to any other in identifying homologous semilandmarks, but it is clear from our findings with LS&ICP that some methods result in semilandmarks that clearly do not represent homologies (ape crania), while estimates from sliding TPS and TPS & NICP appear more reasonable in anatomical terms. Thus, in applying any method, extrinsic anatomical knowledge can guide assessment of accuracy of semilandmarking *sensu* (Cardini, 2020), but this is subjective. Some approaches will clearly fail this test while others will not. However, every approach will give rise to different statistical results. The extent to which differences due to choice of semilandmarking approach depends on the how large these are in relation to the aspects of variation among specimens that are of interest.

The degree to which results from semilandmarks are correlated with those from landmarks alone might be used as a basis for identifying ‘good’ methods (that yield results consistent with those based on homologous landmarks), but perfect association between methods would rather undermine the need for analyses of semilandmarks in the first place, because landmarks are obtained with much less effort. This is similar to the situation with true landmarks, in that landmarks can be located with error and choices of landmarks need to be made, both affect statistical results. However, landmarks, unlike semilandmarks, being defined based on prior anatomical knowledge, do not require an algorithm to be chosen as the basis of locating them. As noted by (Oxnard & O’Higgins, 2009), the number and locations of landmarks chosen in any particular study can and should be based on the question at hand. Many questions can be sensibly and fully addressed using a few well-chosen landmarks. However as noted in the introduction, landmarks may be sparse on (homologous) structures of interest, they also can have doubtful homology or be difficult to locate.

A cautious approach to working with landmark and semilandmark data would be, first to design a landmark configuration that relates to the hypothesis under test (Oxnard & O’Higgins, 2009), and then semilandmark the sample. Statistical testing might then be seen as distinct from visualisation, and proceed on the basis of the landmark configurations alone. Visualisations (warped surface meshes between e.g. means or representing a vector of transformation) can then be estimated based on parallel analyses using the landmarks and semilandmarks. This avoids the philosophical issues that arise concerning semilandmark homology and it avoids the statistical issues that arise when many variables are taken on small samples than arise with semilandmarks (p (number of variables) / n (sample size) ratio) (Cardini, 2020; Cardini et al., 2019). However, this

approach limits the analysis to identifiable landmarks and so, omits what might be useful ‘signal’ from the surface between landmarks. However, in semilandmarking surfaces there is a decision to be made regarding the balance between likelihood of erroneous results (‘noise’- inaccurate identification of homologous points and the ratio of the number variables to the number of specimens - p/n ratio) from semilandmarking and the potential gains in ‘signal’. To large degree this is a judgement call. However, the p/n ratio issue can be mitigated by minimizing the number of semilandmarks used, while the issue of homology of semilandmarks cannot.

With regard to the suggestion that visualisations based on landmarks and semilandmarks be estimated in a separate step from statistical testing using landmarks, for this to be reasonable, the choice of semilandmarking approach needs to have little effect on subsequent visualisations. As noted earlier, configurations of landmarks and semilandmarks lie on surfaces and, in some applications, the form of the semilandmark configuration is of secondary interest to that of the surface mesh itself. Different configurations might reasonably define the form of the same surface. While these different configurations may lead to different statistical findings (e.g. mean shape coordinates, patterns of variation and covariation), it is unclear if visualisations of surface mesh form warped to configurations derived from different semilandmarking algorithms also differ and if so, to what extent. This is an important consideration in certain applications and it is considered in (Shui et al., 2023).

2.4.3 Limitations and future work

The differences identified in the present study reflect the differences in semilandmark locations achieved by different algorithms for semilandmarking. As noted in the introduction, previous studies have compared the performance of different semilandmarking approaches based on different criteria: distance matrices (Boyer et al., 2011; Boyer et al., 2015; Pomidor et al., 2016), principal components (PCs) (Boyer et al., 2015; Gonzalez et al., 2016; Gunz & Mitteroecker, 2013; Rolfe et al., 2021) and differences between template and transformed meshes (Rolfe et al., 2021; Shui et al., 2020). Such sensitivity studies are useful in understanding sources of error and in guiding eventual parameterisation in a particular context, but it is not clear how generalisable their findings are. This caveat also applies to the present study and so, its findings cannot be considered as definitive, rather they offer insights into the consistency of statistical findings based on a limited range of alternative semilandmarking approaches. This study is limited in its scope, having examined only three possible semilandmarking approaches applied to head and cranial surface data. Future work should consider alternative approaches, the effects of varying numbers of landmarks on semilandmarking and a wider range of surface data. Additionally, studies need to be conducted using simulated data, created by perturbing a known surface in known ways to allow assessment of the accuracy

of estimation of means and other statistical parameters.

2.5 Conclusion

In summary, this study utilized three different semilandmarking approaches to yield semilandmarks at different densities. The effects of different semilandmarking approaches and densities of semilandmarks on semilandmark locations and on subsequent statistical results were then considered. It is not possible to assess the extent to which the different approaches yield semilandmarks that accurately reflect homology, but it was possible to assess consistency (= precision) (Cardini, 2020) between approaches and densities of semilandmarks. The TPS&NICP approach yields the most consistent results across varying semilandmark densities applied to both the head surface and ape cranial data and sliding TPS produces results that are most consistent with those from TPS&NICP, especially at higher semilandmarking densities and if only regions bounded by landmarks are considered. However, consistency is not the same as accuracy and so it is not possible to say which if any method produces semilandmarks that accurately represent homologies among specimens. This is a significant issue in applications to the study of developmental or evolutionary transformations, but less so in other applications, such as identification/discrimination. By focusing on landmarks with more secure homology for statistical analyses and employing semilandmarks for visualisation, these issues are minimised.

Further work is needed to assess alternative semilandmarking approaches in different contexts, but for now interpretations of statistical results based on semilandmarks should be made with due caution regarding the potential errors in semilandmarking and serious consideration should be given to why semilandmarking is being undertaken, given that simpler landmark data may well yield the same results, with less uncertainty about homology and so interpretation of studies of transformation of form.

2.6 Acknowledgments

I am grateful to Nick Pears and Christian Duncan for kindly providing access to the headspace dataset (<https://www-users.cs.york.ac.uk/~nep/research/Headspace/>, accessed on 16 January 2023). Thank Olivia Smith for providing landmarks for every head and to the Smithsonian Institution (www.vertebrates.si.edu) for kindly providing access to the ape crania.

2.7 References

- Adams, D, Rohlf, J, & Slice, D. (2004). Geometric morphometrics: ten years of progress following the 'revolution'. *Italian Journal of Zoology*, 71(1), 5-16.
- Amberg, B, Romdhani, S, & Vetter, T. (2007). Optimal step nonrigid ICP algorithms for surface registration. *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR'07)*, 1-8.
- Bardua, C, Felice, R, Watanabe, A, Fabre, A, & Goswami, A. (2019). A practical guide to sliding and surface semilandmarks in morphometric analyses. *Integrative Organismal Biology*, 1(1), obz016.
- Bastir, M, García-Martínez, D, Torres-Tamayo, N, Palancar, CA, Fernández-Pérez, FJ, Riesco-López, A, Osborne-Márquez, P, et al. (2019). Workflows in a Virtual Morphology Lab: 3D scanning, measuring, and printing. *Journal of Anthropological Sciences*, 97, 107-134.
- Besl, PJ, & McKay, ND. (1992). A method for registration of 3-D shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2), 239-256.
- Blanz, V, & Vetter, T. (1999). A morphable model for the synthesis of 3D faces. *The Proceedings of the 26th annual conference on Computer graphics and interactive techniques*, 187-194.
- Bookstein, FL. (1989). Principal warps: Thin-plate splines and the decomposition of deformations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(6), 567-585.
- Bookstein, FL. (1991). *Morphometric tools for landmark data: geometry and biology*: Cambridge University Press.
- Bookstein, FL. (1997). Landmark methods for forms without landmarks: morphometrics of group differences in outline shape. *Medical Image Analysis*, 1(3), 225-243.
- Booth, J, Roussos, A, Ponniah, A, Dunaway, D, & Zafeiriou, S. (2018). Large scale 3D morphable models. *International Journal of Computer Vision*, 126(2-4), 233-254.
- Boyer, DM, Lipman, Y, Clair, ES, Puente, J, Patel, BA, Funkhouser, T, Jernvall, J, et al. (2011). Algorithms to automatically quantify the geometric similarity of anatomical surfaces. *Proceedings of the National Academy of Sciences*, 108(45), 18221-18226.
- Boyer, DM, Puente, J, Gladman, JT, Glynn, C, Mukherjee, S, Yapuncich, GS, & Daubechies, I. (2015). A new fully automated approach for aligning and comparing shapes. *The Anatomical Record*, 298(1), 249-276.
- Canavan, S, Liu, P, Zhang, X, & Yin, L. (2015). Landmark localization on 3D/4D range data using a shape index-based statistical shape model with global and local constraints.

Computer Vision Image Understanding, 139, 136-148.

Cardini, A. (2020). Less tautology, more biology? A comment on “high-density” morphometrics. *Zoomorphology*, 139(4), 513-529.

Cardini, A, O’Higgins, P, & Rohlf, FJ. (2019). Seeing distinct groups where there are none: spurious patterns from between-group PCA. *Evolutionary Biology*, 46(4), 303-316.

Chakravarty, M, Aleong, R, Leonard, G, Perron, M, Pike, B, Richer, L, Veillette, S, et al. (2011). Automated analysis of craniofacial morphology using magnetic resonance images. *PLoS One*, 6(5), e20241.

Chen, X, Lian, C, Deng, HH, Kuang, T, Lin, H-Y, Xiao, D, Gateno, J, et al. (2021). Fast and Accurate Craniomaxillofacial Landmark Detection via 3D Faster R-CNN. *IEEE Transactions on Medical Imaging*, 40(12), 3867-3878.

Cootes, TF, Taylor, CJ, Cooper, DH, & Graham, J. (1995). Active shape models-their training and application. *Computer vision and image understanding*, 61(1), 38-59.

Dai, H, Pears, N, Smith, W, & Duncan, C. (2020). Statistical Modeling of Craniofacial Shape and Texture. *International Journal of Computer Vision*, 128(2), 547-571.

Duncan, C, Pears, NE, Dai, H, Smith, WAP, & O’Higgins, P. (2022). Applications of 3D Photography in Craniofacial Surgery. *Journal of Paediatric Neurosciences*, 17(S 1), 21-28.

Dutilleul, P, Stockwell, JD, Frigon, D, & Legendre, P. (2000). The Mantel test versus Pearson's correlation analysis: Assessment of the differences for biological and environmental studies. *Journal of Agricultural, Biological, Environmental statistics*, 131-150.

Gao, T, Yapuncich, GS, Daubechies, I, Mukherjee, S, & Boyer, DM. (2018). Development and assessment of fully automated and globally transitive geometric morphometric methods, with application to a biological comparative dataset with high interspecific variation. *The Anatomical Record*, 301(4), 636-658.

Gonzalez, PN, Barbeito-Andrés, J, D’Addona, LA, Bernal, V, & Perez, SI. (2016). Performance of semi and fully automated approaches for registration of 3D surface coordinates in geometric morphometric studies. *American Journal of Physical Anthropology*, 160(1), 169-178.

Gonzalez, PN, Perez, SI, & Bernal, V. (2010). Ontogeny of robusticity of craniofacial traits in modern humans: a study of South American populations. *American Journal of Physical Anthropology*, 142(3), 367-379.

Goswami, A, Watanabe, A, Felice, RN, Bardua, C, Fabre, A-C, & Polly, PD. (2019). High-density morphometric analysis of shape and integration: the good, the bad, and the not-really-a-problem. *Integrative Comparative Biology*, 59(3), 669-683.

- Gu, X, Wang, Y, Chan, TF, Thompson, PM, & Yau, S-T. (2004). Genus zero surface conformal mapping and its application to brain surface mapping. *IEEE transactions on Medical Imaging*, 23(8), 949-958.
- Gunz, P, & Mitteroecker, P. (2013). Semilandmarks: a method for quantifying curves and surfaces. *Hystrix, the Italian Journal of Mammalogy*, 24(1), 103-109.
- Gunz, P, Mitteroecker, P, & Bookstein, FL. (2005). Semilandmarks in three dimensions. In *Modern morphometrics in physical anthropology*, 73-98: Springer.
- Gunz, P, Mitteroecker, P, Neubauer, S, Weber, GW, & Bookstein, FL. (2009). Principles for the virtual reconstruction of hominin crania. *Journal of Human Evolution*, 57(1), 48-62.
- Harper, CM, Goldstein, DM, & Sylvester, AD. (2022). Comparing and combining sliding semilandmarks and weighted spherical harmonics for shape analysis. *Journal of Anatomy*, 240(4), 678-687.
- Klingenberg, CP. (2013). Visualizations in geometric morphometrics: how to read and how to make graphs showing shape changes. *Hystrix, the Italian Journal of Mammalogy*, 24(1), 15-24.
- Klingenberg, CP. (2016). Size, shape, and form: concepts of allometry in geometric morphometrics. *Development Genes Evolution*, 226(3), 113-137.
- Koehl, P, & Hass, J. (2015). Landmark-free geometric methods in biological shape analysis. *Journal of The Royal Society Interface*, 12(113), 20150795.
- Le, V-L, Beurton-Aimar, M, Zemmari, A, Marie, A, & Parisey, N. (2020). Automated landmarking for insects morphometric analysis using deep neural networks. *Ecological Informatics*, 60, 101175.
- Li, M, Cole, JB, Manyama, M, Larson, JR, Liberton, DK, Riccardi, SL, Ferrara, TM, et al. (2017). Rapid automated landmarking for morphometric analysis of three-dimensional facial scans. *Journal of Anatomy*, 230(4), 607-618.
- MacLeod, N. (1999). Generalizing and extending the eigenshape method of shape space visualization and analysis. *Paleobiology*, 25(1), 107-138.
- Marshall, AF, Bardua, C, Gower, DJ, Wilkinson, M, Sherratt, E, & Goswami, A. (2019). High-density three-dimensional morphometric analyses support conserved static (intraspecific) modularity in caecilian (Amphibia: Gymnophiona) crania. *Biological Journal of the Linnean Society*, 126(4), 721-742.
- Mitteroecker, P, & Gunz, P. (2009). Advances in geometric morphometrics. *Evolutionary Biology*, 36(2), 235-247.
- Mitteroecker, P, & Schaefer, K. (2022). Thirty years of geometric morphometrics:

Achievements, challenges, and the ongoing quest for biological meaningfulness. *American Journal of Biological Anthropology*, 178(S 74), 181-210.

Musilová, B, Dupej, J, Velemínská, J, Chaumoitre, K, & Bruzek, J. (2016). Exocranial surfaces for sex assessment of the human cranium. *Forensic Science International*, 269, 70-77.

Mydlová, M, Dupej, J, Koudelová, J, & Velemínská, J. (2015). Sexual dimorphism of facial appearance in ageing human adults: A cross-sectional study. *Forensic Science International*, 257, 519. e511-519. e519.

Negrillo-Cárdenas, J, Jiménez-Pérez, J-R, Cañada-Oya, H, Feito, FR, & Delgado-Martínez, AD. (2022). Hybrid curvature-geometrical detection of landmarks for the automatic analysis of the reduction of supracondylar fractures of the femur. *Computer Methods Programs in Biomedicine*, 226, 107177.

O'Higgins, P. (2000). The study of morphological variation in the hominid fossil record: biology, landmarks and geometry. *Journal of Anatomy*, 197(1), 103-120.

O'Higgins, P, Fitton, LC, & Godinho, RM. (2019). Geometric morphometrics and finite elements analysis: Assessing the functional implications of differences in craniofacial form in the hominin fossil record. *Journal of Archaeological Science*, 101, 159-168.

O'Higgins, P, Cobb, SN, Fitton, LC, Gröning, F, Phillips, R, Liu, J, & Fagan, MJ. (2011). Combining geometric morphometrics and functional simulation: an emerging toolkit for virtual functional analyses. *Journal of Anatomy*, 218(1), 3-15.

Oxnard, C, & O'Higgins, P. (2009). Biology clearly needs morphometrics. Does morphometrics need biology? *Biological Theory*, 4(1), 84-97.

Perez, SI, Bernal, V, & Gonzalez, PN. (2006). Differences between sliding semi-landmark methods in geometric morphometrics, with an application to human craniofacial and dental variation. *Journal of Anatomy*, 208(6), 769-784.

Pomidor, BJ, Makedonska, J, & Slice, DE. (2016). A landmark-free method for three-dimensional shape analysis. *PloS one*, 11(3), e0150368.

Porto, A, Rolfe, S, & Maga, AM. (2021). ALPACA: A fast and accurate computer vision approach for automated landmarking of three-dimensional biological structures. *Methods in Ecology and Evolution*, 12(11), 2129-2144.

Profico, A, Piras, P, Buzi, C, Di Vincenzo, F, Lattarini, F, Melchionna, M, Veneziano, A, et al. (2017). The evolution of cranial base and face in Cercopithecoidea and Hominoidea: Modularity and morphological integration. *American Journal of Primatology*, 79, e22721.

Rolfe, S, Davis, C, & Maga, AM. (2021). Comparing semi-landmarking approaches for analyzing three-dimensional cranial morphology. *American Journal of Physical Anthropology*, 175(1), 227-237.

- Rusinkiewicz, S, & Levoy, M. (2001). Efficient variants of the ICP algorithm. *The Proceedings of the Third IEEE international conference on 3-D digital imaging and modeling*, 145-152.
- Schlager, S. (2017). Morpho and Rvcg—Shape Analysis in R: R-Packages for geometric morphometrics, shape analysis and surface manipulations. In *Statistical shape and deformation analysis*, 217-256.
- Schlager, S, & Rüdell, A. (2017). Sexual Dimorphism and population affinity in the human zygomatic structure—comparing surface to outline data. *The Anatomical Record*, 300(1), 226-237.
- Shui, W, Profico, A, & O'Higgins, P. (2023). A comparison of semilandmarking approaches in the visualisation of shape differences. *Animals*, 13 (3), 385.
- Shui, W, Zhang, Y, Wu, X, & Zhou, M. (2021). A computerized facial approximation method for archaic humans based on dense facial soft tissue thickness depths. *Archaeological Anthropological Sciences*, 13(11), 186.
- Shui, W, Zhou, M, Maddock, S, He, T, Wang, X, & Deng, Q. (2017). A PCA-Based method for determining craniofacial relationship and sexual dimorphism of facial shapes. *Computers in Biology and Medicine*, 90, 33-49.
- Shui, W, Zhou, M, Maddock, S, Ji, Y, Deng, Q, Li, K, Fan, Y, et al. (2020). A computerized craniofacial reconstruction method for an unidentified skull based on statistical shape models. *Multimedia Tools and Applications*, 79, 25589-25611.
- Smith, OA, Nashed, YS, Duncan, C, Pears, N, Profico, A, & O'Higgins, P. (2021). 3D Modeling of craniofacial ontogeny and sexual dimorphism in children. *The Anatomical Record*, 304, 1918-1926.
- Van Kaick, O, Zhang, H, Hamarneh, G, & Cohen-Or, D. (2011). A survey on shape correspondence. *Computer Graphics Forum*, 30(6), 1681-1707.
- Velemínská, J, Bigoni, L, Krajíček, V, Borský, J, Šmahelová, D, Cagánová, V, & Peterka, M. (2012). Surface facial modelling and allometry in relation to sexual dimorphism. *HOMO-Journal of Comparative Human Biology*, 63(2), 81-93.
- Viscosi, V, & Cardini, A. (2011). Leaf morphology, taxonomy and geometric morphometrics: a simplified protocol for beginners. *PloS one*, 6(10), e25630.
- Vitek, NS, Manz, CL, Gao, T, Bloch, JI, Strait, SG, & Boyer, DM. (2017). Semi-supervised determination of pseudocryptic morphotypes using observer-free characterizations of anatomical alignment and shape. *Ecology Evolution*, 7(14), 5041-5055.
- Wang, S, Wang, Y, Jin, M, Gu, XD, & Samaras, D. (2007). Conformal geometry and its applications on 3D shape matching, recognition, and stitching. *IEEE Transactions on*

Pattern Analysis and Machine Intelligence, 29(7), 1209-1220.

White, JD, Ortega-Castrillón, A, Matthews, H, Zaidi, AA, Ekrami, O, Snyders, J, Fan, Y, et al. (2019). MeshMonk: Open-source large-scale intensive 3D phenotyping. *Scientific Reports*, 9(1), 1-11.

Supplementary material

Table 2.S1 Ape cranium specimen

Specimen number	Species	Sex	Repository
USNM 176211	Gorilla	Male	Smithsonian Institution
USNM 252578	Gorilla	Male	Smithsonian Institution
USNM 590949	Gorilla	Female	Smithsonian Institution
USNM 590952	Gorilla	Female	Smithsonian Institution
USNM 599170	Gorilla	Female	Smithsonian Institution
USNM 083262	Hylobates lar	Female	Smithsonian Institution
USNM 083263	Hylobates lar	Female	Smithsonian Institution
USNM 083264	Hylobates lar	Female	Smithsonian Institution
USNM 111970	Hylobates lar	Male	Smithsonian Institution
USNM 111988	Hylobates lar	Male	Smithsonian Institution
USNM 174701	Pan troglodytes	Female	Smithsonian Institution
USNM 174704	Pan troglodytes	Male	Smithsonian Institution
USNM 220062	Pan troglodytes	Female	Smithsonian Institution
USNM 220327	Pan troglodytes	Male	Smithsonian Institution
USNM 395820	Pan troglodytes	Male	Smithsonian Institution
USNM 143590	Pongo abelii	Male	Smithsonian Institution
USNM 267325	Pongo abelii	Male	Smithsonian Institution
USNM 270807	Pongo abelii	Female	Smithsonian Institution
USNM 283737	Pongo abelii	Female	Smithsonian Institution
USNM 293165	Pongo abelii	Male	Smithsonian Institution

Chapter 3 A comparison of semilandmarking approaches in the visualisation of shape differences

Abstract: In landmark based analyses of size and shape variation and covariation among biological structures, regions lacking clearly identifiable homologous landmarks are commonly described by semilandmarks. Different algorithms may be used to apply semilandmarks, but little is known about the consequences of analytical results. Here, we assess how different approaches and semilandmarking densities affect the estimates and visualisations of mean and allometrically scaled surfaces. The performance of three landmark driven semilandmarking approaches is assessed, using two different surface mesh data sets with different degrees of variation and complexity: adult human head and ape cranial surfaces. Surfaces fitted to estimates of the mean and allometrically scaled landmark and semilandmark configurations arising from geometric morphometric (GM) analyses of these datasets are compared between semilandmarking approaches and different densities as well with those from warping to landmarks alone. We find that estimates of surface mesh shape (i.e. after re-semilandmarking and then re-warping) made with varying numbers of semilandmarks are generally consistent, while warping of surfaces using landmarks alone yields surfaces that can be quite different to those based on semilandmarks, depending on landmark coverage and choice of template surface for warping. The extent to which these differences are important depends on the particular study context and aims.

Article details:

Wuyang Shui, Antonio Profico and Paul O'Higgins. A comparison of semilandmarking approaches in the visualisation of shape differences. *Animals*. 2023, 13(3): 385.

3.1 Introduction

Over the last three decades, landmark-based geometric morphometric (GM) methods have been increasingly applied to quantify and compare size and shape variation and covariation (Adams et al., 2004; Mitteroecker & Schaefer, 2022; O'Higgins & Jones, 1998; Rohlf, 1998; Viscosi & Cardini, 2011). Before performing GM analyses, the definition of a suitable configuration of landmarks in relation to the research aim is required (Cardini, 2020; Oxnard & O'Higgins, 2009). A simple landmark configuration might be perfectly adequate to quantify shape differences appropriate to the question in hand. In studies of biological transformations such as growth or evolution, the landmarks define equivalent points; that are 'the same' in terms of development or evolution ('this point turns into that point'; is homologous), but the locations of homologous landmarks and their density are limited by the extent to which they can be identified and usually on the presence of identifiable anatomical features as well as preservation of material and available time for digitization.

In many biological applications, landmarks cannot readily be identified, e.g. over smooth regions such as the human cranial vault or tooth crowns. In an attempt to provide detailed information on such regions, different approaches have been proposed to marking up semilandmarks (or dense point correspondences) among curves or surfaces between landmarks (Bardua et al., 2019; Van Kaick et al., 2011). The method of sliding semilandmarks that locates semilandmarks by minimising the bending energy of thin-plate splines (TPS) or Procrustes distance (Bookstein, 1997; Gunz & Mitteroecker, 2013; Gunz et al., 2005) is most commonly used in biology. Alternative semilandmarking methods include rigid registration approaches, e.g. the *auto3dgm* package (Boyer et al., 2015) based on the iterative closest points (ICP) algorithm (Besl & McKay, 1992), and non-rigid registration approaches, e.g. non-rigid ICP (NICP) (Amberg et al., 2007; Booth et al., 2018) and the optical flow algorithm (Blanz & Vetter, 1999), among others. The fundamental task of these semilandmarking approaches is to transfer the semilandmarks from a template surface (e.g. a mean surface) to the target specimen. It is worth noting that semilandmarks rely primarily on mathematical mappings and/or the similarity of topographic features, rather than developmental or evolutionary equivalences based on prior knowledge.

Recent studies have assessed the performance of different semilandmarking approaches based on principal components (PCs) (Boyer et al., 2015; Gonzalez et al., 2016; Gunz & Mitteroecker, 2013; Perez et al., 2006; Rolfe et al., 2021), distance matrices (Boyer et al., 2011; Boyer et al., 2015; Pomidor et al., 2016), the geometric deviation between template and transformed meshes (Rolfe et al., 2021; Shui et al., 2020). These have found that different approaches yield different semilandmark locations and so, result in analytical results that differ to some degree. This has been further investigated in a

prior study (Shui et al., 2023) that provided the starting point for the present one. The performance of three of the semilandmarking approaches described above was systematically examined. These included the sliding TPS approach outlined above. The second approach employed a hybrid rigid registration combining least squares (LS) (Arun et al., 1987) and ICP algorithms (LS&ICP). After using the LS algorithm to fit the template landmarks to those of each specimen, the ICP algorithm rigidly refitted the template to the target, minimising the sum of squared distances between landmarks and estimated semilandmarks, found by searching for the nearest points on the target from the registered template semilandmarks. The third approach (TPS&NICP) (Shui et al., 2021), used TPS to perform an initial non-rigid registration of the template landmarks and surface to specimens and then the NICP algorithm (Amberg et al., 2007) was applied to further warp the deformed template surface to each specimen as rigidly as possible, optimizing the cost function by assigning an affine transformation to each vertex, rather than an interpolation function as used in TPS, before transferring the semilandmarks from the template to the nearest point of the specimen surfaces. We compared among semilandmarking approaches, differences in the locations of semilandmarks, Procrustes distances between landmark and semilandmark configurations, estimates of mean landmark and semilandmark configurations, PCs of configuration shape, and estimates of allometry.

Because homology is unknown for regions that were semilandmarked, it is not possible to assess how well semilandmarks represent homology, rather the focus was on comparing the results of analyses based on semilandmarks between and within methods, with increasing semilandmark density. The analyses showed that each semilandmarking approach yields different locations of semilandmarks, which result in differences in each of the comparisons (Shui et al., 2023). The sliding TPS algorithm and TPS&NICP yielded results that are more similar to each other than with those based on LS&ICP. Further, we assessed consistency within methods among results obtained using different densities of semilandmarks, finding that sliding TPS and TPS&NICP approaches are most consistent, especially where true landmarks are dense. The extent to which these differences are important depends on the context, the question being addressed and the purpose of the study, but all semilandmarking approaches estimate homology with error, the extent of which is unknowable. Therefore, all subsequent statistical analyses that aim to describe developmental or evolutionary transformation are subject to that error, and should be treated with an appropriate degree of caution (Cardini, 2020; Shui et al., 2023).

Geometric morphometric analyses enable visualisation of statistical findings, generating landmark and semilandmark configurations that represent shapes or forms (sizes and shapes) of interest such as the mean, or allometrically scaled configurations. Surfaces or regular grids are often warped to these configurations to aid visualisation of

shape differences, and where applicable, changes. This is most commonly done using TPS (Bookstein, 1989; Mitteroecker & Gunz, 2009). However, (Oxnard & O’Higgins, 2009) noted that ‘With sliding semilandmarks, their relative positions on equivalent curves, surfaces, etc. are not singly interpretable, but rather should be read as a whole, respecting the fact that the underlying assumption in their construction is one of equivalence of the curve or surface patch as a whole’. This has recently been reiterated (Bastir et al., 2019; Mitteroecker & Schaefer, 2022), ‘the coordinates of semilandmarks along the surface are meaningless, and one cannot interpret the position of single semilandmarks, only the surface geometry that all semilandmarks describe together’. Thus, although semilandmarks are treated as landmarks in statistical analyses, and so ‘errors’ in their locations (or differences using different methods to locate them) influence statistical outcomes, as was demonstrated in the previous study (Shui et al., 2023), visualisation and interpretation of differences should ignore their locations and focus on the shape of the curve or surface they describe.

It is therefore of interest to know the extent to which the shapes of surfaces warped to fit semilandmark and landmark configurations varying in semilandmark density and locations (e.g. arising from different approaches to placing them) are consistent. If different densities and approaches yield identical or very similar visualisations, this may be reassuring in certain practical applications. For instance, a mean surface might be used in clinical work to compare measurements taken on a patient with an estimate of the population mean (Duncan et al., 2022) and facial approximation from the skull alone in the realm of forensic science (Shui et al., 2016). Additionally, surfaces from GM analyses are used to virtually repair and reconstruct fossil material (Gunz et al., 2009) and to build 3D models for functional analyses such as finite element analysis (FEA) (O’Higgins et al., 2011). The extent to which such estimated surfaces differ when derived using different semilandmark densities and semilandmarking approaches is unknown, yet it is important in that it may affect subsequent morphometric or functional analysis. This question is addressed in the present study.

The main purpose of this study is to empirically test two hypotheses using surface scans of human heads and ape crania: that there are no differences in surface mesh shape (the shape of the configuration of surface vertices, the nodes of the surfaces mesh, rather than the landmark and semilandmark configuration) between estimates, derived using different semilandmarking densities and approaches, applied to surfaces representing (a) the mean of a sample (b) allometrically scaled shapes.

To these ends, a template surface mesh is warped to fit estimated mean and allometrically scaled mean landmark and semilandmark configurations derived from different semilandmarking densities and approaches and the resulting surfaces are compared. Additionally, these surfaces are compared with surfaces warped using

landmarks alone. Of interest is the extent to which these surfaces differ and how they differ. The focus is on the comparison of shape of the surface rather than the geometry of the underlying mesh.

3.2 Materials and Methods

3.2.1 Materials

3.2.1.1 Datasets and templates

We used two datasets comprising surface meshes that exhibit varying degrees of complexity: 100 adult human male heads comprising 16 anatomical landmarks from the Liverpool-York Headspace dataset (Dai et al., 2020; Smith et al., 2021) and 20 ape crania consisting of 41 anatomical landmarks (5 *Gorilla*, 5 *Hylobates lar*, 5 *Pan troglodytes*, and 5 *Pongo abelii*) (Profico et al., 2017). More details of the ape crania can be seen in Table 3.S1. We extracted the external surfaces of heads and ape crania using the method developed in the previous study (Shui et al., 2020). It may avoid the internal surface interfering with projection of semilandmarks. The 3D meshes were post-processed by smoothing surfaces, removing the irrelative discrete vertices, and repairing self-intersecting triangle meshes. Compared to human heads, the sample of ape crania of different species shows greater size and shape variation and presents more complex surfaces.

Similar to the previous study (Shui et al., 2023), the mean surfaces of heads and ape crania were estimated and used as templates (after landmarking and semilandmarking them) for each dataset to yield semilandmarks among every specimen. For the human head, we selected an arbitrary head as the initial template and then used NICP (Amberg et al., 2007) to align all of the human heads based on landmarks and establish dense point correspondences (identify points on the target surface that match each vertex of the template surface). Next, the mean head was estimated by averaging correspondences among heads. For the ape crania, which vary far more in form, an alternative approach was required. The k-means clustering algorithm was employed to sample 800 points over a *Gorilla* cranium, then the sliding TPS approach (Gunz & Mitteroecker, 2013) was used to yield semilandmarks among specimens. Following this, the mean form of the landmark and semilandmark configurations was calculated, and the mean ape cranial surface was estimated by warping the surface mesh of an ape cranium specimen (USNM 176211) to fit this configuration. Note that the Procrustes distance between the selected specimen and the mean shape based on the mean landmarks and 800 semilandmarks was 0.0988, which is very similar to the average difference between individuals and the mean, estimated using sliding TPS semilandmarking method. This process of making an initial estimate of the mean follows that commonly used to compute semilandmarks, where an arbitrary specimen is used as an initial template to estimate semilandmark coordinates, and the

mean of the resulting landmarks and semilandmarks is used to estimate a new mean template by re-warping of the original template to them, before re-semilandmarking the sample.

Figure 3.1a shows the human template head with 16 landmarks and Figure 3.1b shows the ape template cranium with 41 landmarks. Notably, the scalp surface in the headspace data lacks identifiable landmarks, while the ape crania present landmarks over the whole surface. These differences are expected to affect how well semilandmarking is controlled, particularly for sliding TPS, because landmarks are required to control sliding, which is not the case for the other approaches. Sliding TPS was applied over the scalp for consistency of analyses and comparability of results.

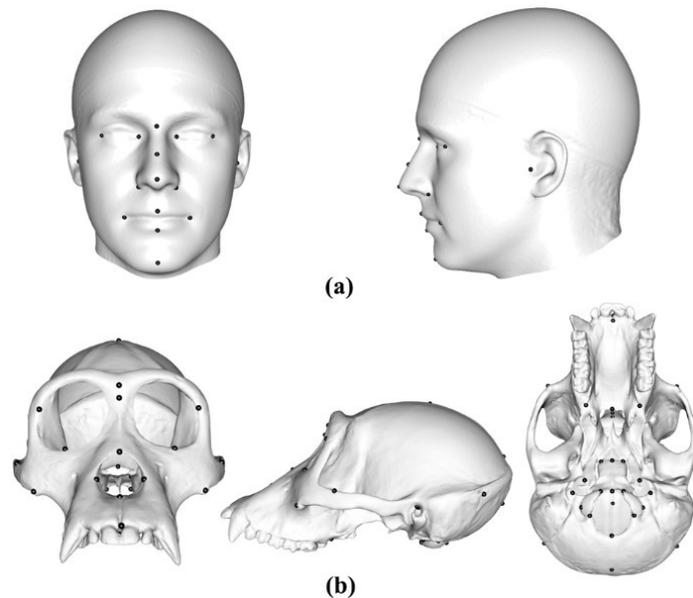


Figure 3.1 Two datasets. (a) The human template head with 16 landmarks. (b) The ape template cranium with 41 landmarks.

3.2.1.2 Semilandmarks

The *k*-means algorithm was used to sample, as evenly as possible, 10 different densities of semilandmarks per square centimetre from the template head; 20 (0.017/cm²), 40 (0.034/cm²), 60 (0.052/cm²), 80(0.069/cm²), 100 (0.086/cm²), 200 (0.172/cm²), 400 (0.343/cm²), 600 (0.515/cm²), 800 (0.688/cm²) and 1000 (0.858/cm²) semilandmarks. Additionally, the above procedure was repeated to generate five different semilandmark densities among ape crania; 50 (0.129/cm²), 100 (0.258/cm²), 200 (0.517/cm²), 400 (1.033/cm²) and 800 (2.067/cm²). Following creation of the templates, three different semilandmarking approaches were employed to project semilandmarks from the template to every specimen, to yield semilandmarks, as follows (Shui et al., 2023). The software prototype used to generate the semilandmarks using three different approaches can be downloaded from the Internet (https://github.com/sissun/Geometric_morphometrics.git).

a) Sliding TPS

The sliding TPS approach is the most commonly used approach in biological studies, to yield semilandmarks by sliding semilandmarks projected from the template along the tangent direction of a curve or the tangent plane of a surface, minimising the bending energy of TPS (Gunz et al., 2005; Mitteroecker et al., 2013). In this study, we used the patching (*placePatch*) and sliding (*slider3d*) procedures in the *Morpho* R package (version 2.10) to yield sliding semilandmarks at varying densities based on the template (Schlager, 2017). The sliding step minimises bending energy and so, depends on landmarks to control the sliding. For the headspace data, no landmarks are present over the scalp and so we expect sliding to be poorly controlled. This situation does not arise with the ape cranial data.

b) Rigid registration

We used the rigid LS&ICP method to register the template to every specimen based on the fixed landmarks and then projected semilandmarks from the template to each specimen. First, the initial rigid alignment calculated by LS, constrained by landmarks, was performed to fit the template to each specimen. Second, the ICP algorithm rigidly refitted the template to the target, minimising the sum of squared Euclidean distances between landmarks and semilandmarks on the template and specimen. The alignment generated by LS speeds up convergence of the ICP algorithm. Finally, we projected different densities of semilandmarks from the registered template to each specimen. This was carried out using purpose-built code in the C++ programming language using Microsoft Visual Studio 2015.

c) Non-rigid registration

We used the non-rigid TPS&NICP method (Shui et al., 2021) to yield semilandmarks on every specimen. This comprised two steps: first, a triplet of TPS was used to warp the template to every specimen based on the fixed landmarks. Second, the NICP algorithm (Amberg et al., 2007) was applied to warp the deformed template surface to each specimen and establish dense point correspondences based on locally affine regularizations and adjustable stiffness parameters. In this process, preliminary correspondences are established by searching for the nearest points between two surfaces and then the cost function is optimized. It comprises a landmark term, a local affine regularization term and a stiffness term and assigns an affine transformation to each vertex. New correspondences are obtained by searching the deformed template surfaces. Registration loops are carried out in which stiffness weights are iteratively decreased and the template is incrementally deformed. This non-rigid method, in contrast to the rigid registration used in LS&ICP, matches the warped template surface closely to each specimen. This was carried out using purpose-built code in the C++ and Matlab programming language.

Figure 3.2a shows 100 semilandmarks generated by sliding TPS (black points), LS&ICP (red points), and TPS&NICP (green points) on the mean form of head surface generated by sliding TPS. While semilandmark locations differ among methods of semilandmarking, the differences are small between sliding TPS and TPS&NICP approaches and a little larger between these and the LS&ICP approach (Shui et al., 2023). In contrast, differences are much greater among methods in the ape cranial dataset. Thus, Figure 3.2b shows 100 semilandmarks on the mean ape cranium generated by sliding. Semilandmarks generated by Sliding TPS appear to be in similar locations to those generated by TPS&NICP, but the locations of semilandmarks generated by LS&ICP are quite different.

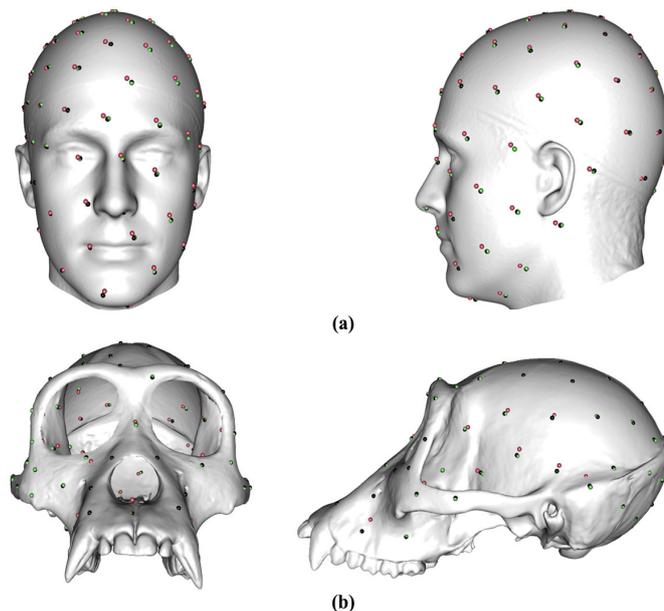


Figure 3.2 Different densities of semilandmarks generated by the sliding TPS (black points), LS&ICP (red points) and TPS&NICP (green points) approaches. (a) Mean forms of adult human heads. (b) Mean forms of ape crania.

3.2.2 Methods

3.2.2.1 Comparisons of mean surface meshes between different approaches

For each dataset, we applied GPA to the landmark and semilandmark configurations from each semilandmarking approach and density and then computed the Procrustes mean configurations (centroid size = 1.0). Subsequently, the surface of the template specimen was warped using TPS to fit each mean configuration, thereby generating a ‘mean surface’ consisting of the coordinates of the full set of vertices with identical topology but different relative vertex locations for each estimate of the mean (from each semilandmarking method and density). A vertex is a node of the mesh and the connections among nodes describe mesh topology. It should be noted that the template surface was already warped to an estimate of the mean during the semilandmarking process and as such, undergoes little further deformation in this step. Next, the differences among these estimates of the mean surface shape were quantified and visualized. A hybrid approach

was used to quantify global and regional differences in mean surface estimates generated by different semilandmarking approaches. The global comparison used Procrustes superimposition to register mean shape surface mesh vertices generated by different semilandmarking approaches, following which the Procrustes distance between mean surfaces was calculated and a principal components analyses (PCA) of mean surfaces was carried out. Additionally, regional differences between estimated mean surfaces were visualized based on (registration independent) colour maps (see example in Figure 3.3) of surface area differences between each equivalent triangle of the two surface meshes (Profico et al., 2021). While these are registration free depictions of differences in surface area, they incompletely describe the differences between surfaces and should be interpreted in conjunction with the surface renderings of reference and target shapes.

However, differences in semilandmark locations and densities between approaches resulted in different locations of mesh vertices, even if the shapes of surfaces being compared are identical. This affected visualisations and computations of distances and PCs based on the vertices.

This is related to the point made by Oxnard and O’Higgins (2009) , Mitteroecker and Schaefer (2022), and Bastir et al. (2019) that semilandmark locations on surfaces should not be interpreted singly. In warping the mesh to each semilandmark, the locations of semilandmarks directly control where mesh vertices are located, and so affect the local geometry of the mesh. Warping transfers differences in individual semilandmark locations to mesh vertices. This is evident from Figure 3.3 which presents colour maps of differences in mesh triangle areas among mean surface shapes generated using different semilandmarking approaches.

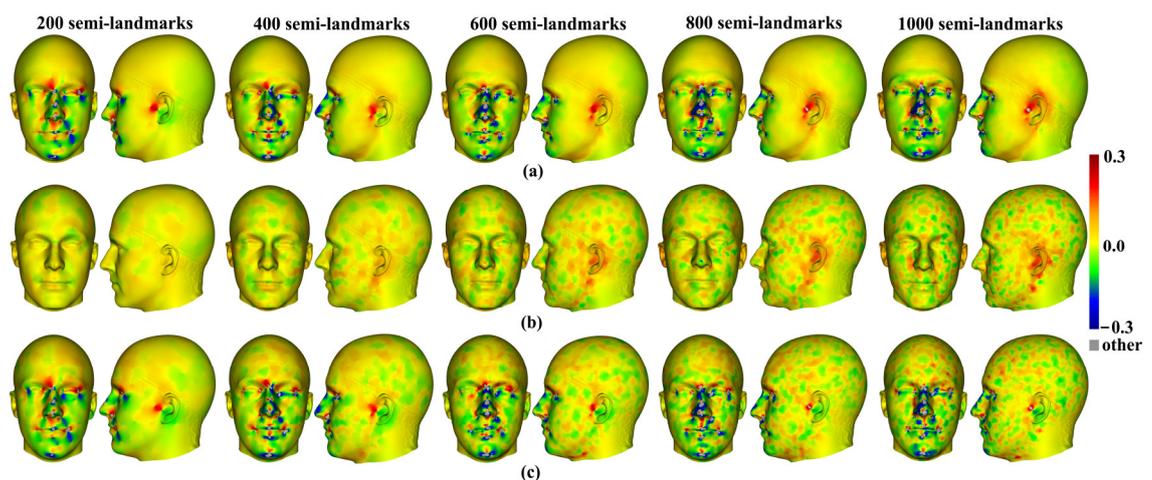


Figure 3.3 Visualisation of the differences in mesh triangle surface areas among mean surface shapes generated using different semilandmarking approaches. (a) Differences between sliding TPS (reference) and LS&ICP approaches (target). (b) Differences between sliding TPS (reference) and TPS&NICP (target) approaches. (c) Differences between TPS&NICP (reference) and LS&ICP (target) approaches. Scale bar indicates difference in local area between reference and target surfaces expressed as a proportion of the reference area. Other refers to the values outside the range of the scale bar.

In the comparisons of Figure 3.3, numerous punctate regions of localised differences in areas of triangle meshes are evident, particularly between sliding TPS and TPS&NICP approaches, where semilandmarks located over the vault are in slightly different places. These lead to the punctate appearance of the colour map. The resulting Procrustes distances between mesh vertices are illustrated in Figure 3.4. These distances increase between the lowest and highest semilandmarking densities, but this is not directly related to the number of semilandmarks used to warp the meshes. Rather the figure shows a generally increasing trend but with increases or decreases in Procrustes distance between successive increments of semilandmark density. While some part of these Procrustes distances relates to differences in surface shape, distances are inflated to unknown degree by the differences in semilandmark locations over the surface.

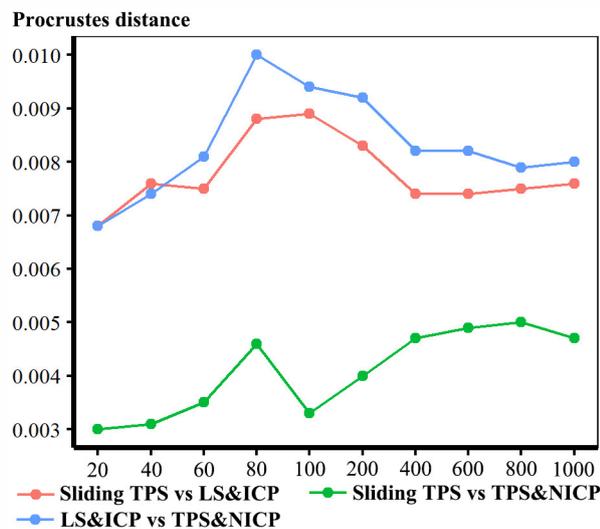


Figure 3.4 Procrustes distances computed between vertices of the mean surfaces of human heads generated by warping the template mesh to the semilandmarks obtained by different approaches.

To avoid this problem, a second semilandmarking step followed by a re-warping of the template surface to these new semilandmarks and the landmarks is required on all surfaces to be compared. First, the semilandmarks of the template surface were projected onto the estimated surfaces (e.g. mean or allometrically scaled shapes) generated by different approaches and densities to generate new semilandmarks based on the fixed landmarks (re-semilandmarking). Second, the template surface is warped to fit the original landmarks and new semilandmarks generated by different approaches and densities to produce the surface (re-warping). This eliminates the localised effects on mesh geometry (e.g. more or less deformed triangles within the meshes, while topology remains constant) of differences in semilandmark locations due to choice of semilandmarking approach. It focuses the comparison on the shapes of the re-warped surfaces, rather than mesh geometry. The sliding TPS and TPS&NICP semilandmarking approaches result in very similar semilandmark locations and consistent statistical results (Figure 3.2 and (Shui et al., 2023)). Either could be chosen as the basis of re-semilandmarking and re-warping of meshes, with little or no effect on the outcome of

comparisons. In this study, the sliding TPS approach was chosen because it is most commonly applied in such work. The result of re-semilandmarking and re-warping is shown in Figure 3.7 (results section) and the Procrustes distances between the vertices of these re-semilandmarked and re-warped surfaces is illustrated in Figure 3.6a (results section).

The resulting visualisations of differences and Procrustes distances between estimates of the mean surface mesh indicate smaller differences after re-sliding (or re-semilandmarking) and re-warping, as expected. Thus, in Figure 3.7, differences between the mean surface mesh derived by LS&ICP and the other two approaches are relatively large in the face, especially around irregular features such as the nose and mouth, while between sliding TPS and TPS&NICP the mesh differences are small, and diffuse. In Figure 3.6a, Procrustes distances, generally increase with increasing density as in Figure 3.4, but are smaller than those from the original fitting of the template mesh to the semilandmarks from each approach.

This re-semilandmarking and re-warping allows mesh surface shape to be compared between semilandmarking methods. It ignores the local differences in surface mesh triangle areas that will affect the colour maps of differences in mesh triangle surface areas, and refocuses the analysis on the shape of the surface (in the sense of its topography). It was applied to all subsequent comparisons of mean surfaces and allometrically scaled surfaces arising from different semilandmarking approaches in this study. It was also applied to the comparison of surfaces derived by each semilandmarking approach using different densities of semilandmarks.

3.2.2.2 Comparisons of allometrically scaled surface meshes

In the previous study (Shui et al., 2023), the predicted landmark and semilandmark configurations representing the extreme limits (smallest and largest) of the allometric vector, derived by multivariate regression of shapes (the scores of specimens on the full set of PCs) on the natural logarithm of centroid size, were computed using each semilandmarking method and density. This was done for both datasets. To investigate how differences in semilandmark locations between approaches affect predictions of allometrically scaled surfaces, the template surface was warped to these configurations. Next, as for the comparisons of mean surfaces, these surfaces were re-semilandmarked and re-warped to yield surface meshes before calculating Procrustes distances between mesh vertices, PCs and visualisations of differences in mesh triangle areas.

3.2.2.3 Comparisons of surface meshes resulting from different semilandmarking densities

The previous analyses focused on differences in surface mesh predictions arising from the use of different semilandmarking approaches. Further analyses were directed to

assessing the extent to which predicted surfaces differ when produced by each semilandmarking approach using different densities of semilandmarks. It was applied to both datasets. As for the comparisons of mean and allometrically scaled surfaces, the surfaces produced by each semilandmarking density using each semilandmarking approach were re-semilandmarked and re-warped and then GPA and PCA were carried based on the vertices of the surfaces generated by different densities of semilandmarks from each approach. Procrustes distances and PCAs were used to assess overall shape difference. Colour map visualisations of differences in mesh area were also produced, but these first required scaling of the meshes. Because the number of semilandmarks varies, the centroid sizes of the full set of vertices of the surfaces fitted to each mean semilandmark and landmark configuration are inversely related to the density of semilandmarks, i.e. surfaces generated using low-densities of semilandmarks are larger than those using high-densities. Therefore, to visualise differences in predicted surface mesh triangle areas, the surfaces (configuration of the full set of vertices) were scaled to the same centroid size.

3.2.2.4 Comparisons of mean and allometrically scaled surface meshes resulting from landmarks alone.

In order to assess what, if anything, is gained by using landmarks and semilandmarks to compute mean and allometrically scaled surfaces, the surfaces from the analyses described above were compared with warped surfaces derived using only the landmarks by computing Procrustes distances among vertices of the template surface mesh warped to fit the mean landmarks or allometrically scaled landmarks from each dataset. The differences between these surfaces and those derived using landmarks and semilandmarks were visualized using colour maps, as described above.

The template mesh for each dataset is an initial estimate of the average surface and so it is expected that fitting it to the mean landmarks will yield a surface not very dissimilar to the mean surfaces estimated using landmarks and semilandmarks. In practice, it is common to use the surface of an individual close to the mean for visualisation as the template, yet the effects of choice of template surface are unclear. Therefore, surfaces derived using landmarks and semilandmarks were compared with those derived using landmarks alone, this time using the head surface with the smallest Procrustes distance to the mean (based on landmarks and the maximum number of semilandmarks) and the ape cranial surface (USNM 176211) used to generate the template cranium, respectively. The resulting predictions of mean and allometrically scaled surfaces were compared with those based on the template surfaces.

3.3 Results

The effects of different semilandmarking approaches and densities on estimates of mean and allometrically scaled surfaces of human heads was assessed after the surfaces were re-semilandmarked and re-warped, and then key analyses were repeated using the ape cranial surfaces to compare the performance of approaches on surfaces that exhibit a greater degree of variation and complexity of surface size and shape. Additionally, these surfaces were compared with those warped to fit the landmark configurations (without semilandmarks).

3.3.1 Comparison of estimates of mean surfaces

The differences in shape of estimated mean surfaces generated by 1) different semilandmarking approaches and 2) densities are quantified. All of these comparisons, and those of allometrically scaled surfaces are based on surfaces derived by re-semilandmarking and re-warping as described in the methods section.

3.3.1.1 Different semilandmarking approaches

The mean head surfaces from each semilandmarking approach derived using varying numbers of semilandmarks, after re-semilandmarking and re-warping are shown in Figure 3.5. The surface mesh renderings before re-semilandmarking and re-warping are not noticeably different in shape and so, are not shown.

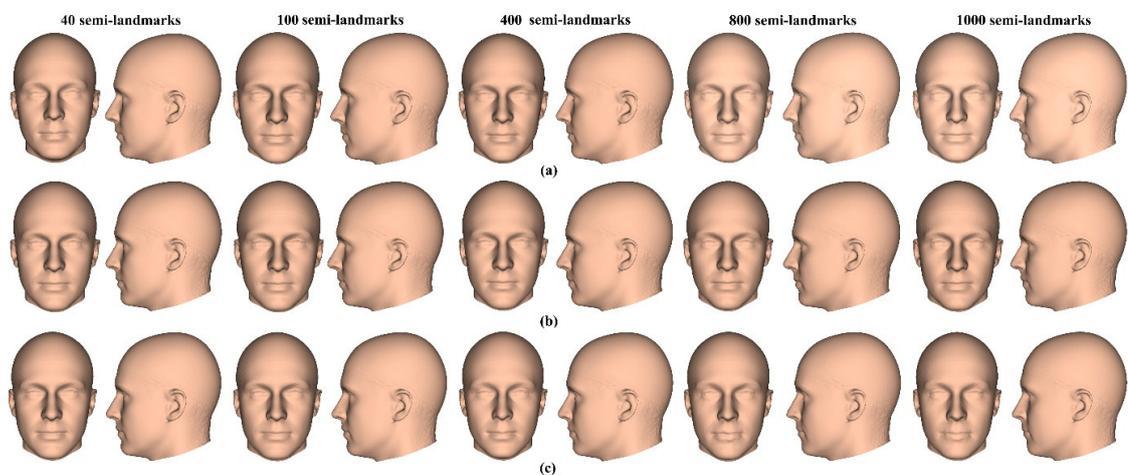


Figure 3.5 Visualisation of the head mesh surfaces generated using different semilandmarking approaches after re-semilandmarking and re-warping. (a) Sliding TPS. (b) TPS&NICP approaches. (c) LS&ICP approaches. Increasing semilandmark density from left to right.

In Figure 3.5, all head surfaces after re-semilandmarking and re-warping appear very similar. The main differences are in the detail of complex regions of the surfaces, where those from LS&ICP appear less sharp and more smooth, especially around the eyes and mouth of heads generated by high-density semilandmarks. In order to compare these in detail, Procrustes distances were computed between the coordinates of all vertices of the surface meshes of the mean human head surfaces estimated using different semilandmarking approaches and densities (Figure 3.6a). For comparison, for the same

surfaces, the Procrustes distances were also computed between the mean landmarks and semilandmarks (Figure 3.6b). These distances are very similar, indicating that the re-warping of meshes preserves differences between the landmark and semilandmark sets. In contrast, the re-warping has a marked effect on Procrustes distances between meshes compared to those warped to the original landmark and semilandmark configurations (see Methods; Figure 3.4 vs Figure 3.6a).

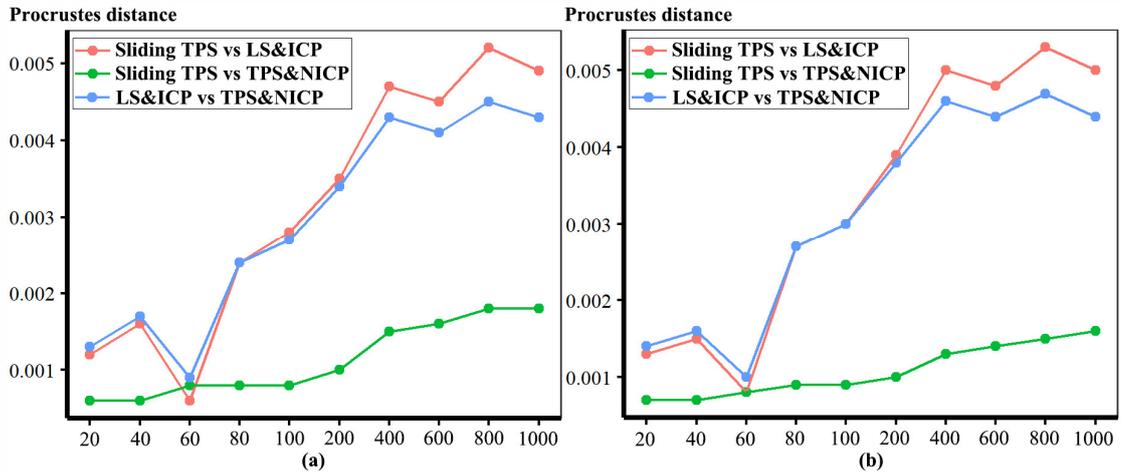


Figure 3.6 Procrustes distances computed between the mean surfaces of human heads obtained by different approaches after re-semilandmarking and re-warping the template mesh. (a) Procrustes distances computed between all the vertices (b) Procrustes distances between mean landmarks and semilandmarks.

Comparison of these distances among different semilandmarking approaches (Figure 3.6a), indicates that the full set of vertices of the mean surface generated from sliding TPS are, in general, most similar (smallest Procrustes distances) to those from TPS&NICP at all semilandmarking densities and these distances increase with increasing semilandmark density. The Procrustes distances between mean surfaces based on semilandmarks from LS&ICP and both sliding TPS and TPS&NICP are, in general, larger and also tend to increase with increasing semilandmark density.

Differences between mean surfaces of human heads derived from different semilandmarking approaches and densities of semilandmarks are illustrated in Figure 3.7. This visualizes differences in areas of equivalent triangles of the template surface mesh derived from each semilandmarking approach and density, after re-semilandmarking and re-warping. Figure 3.7a visualizes the differences in shape among mean surface meshes from sliding TPS (reference) and LS&ICP.

Differences in local surface areas between sliding TPS and TPS&NICP (Figure 3.7b) are very small at all semilandmark densities. The scalp region smoothly presents slightly smaller local surface areas (\sim ratio of difference in area $\sim 0.01 = 1\%$; light green) from TPS&NICP relative to sliding TPS. In comparisons between LS&ICP and the other semilandmarking approaches (Figures 3.7a and c) differences increase markedly with increasing semilandmark number and are mostly found in the face in regions of complex topography e.g. eyes, nose, mouth and chin, and in which semilandmarks are closer to

fixed landmarks. They are much less marked over the scalp. These visualisations reflect the Procrustes distances between surfaces presented in Figure 3.6a.

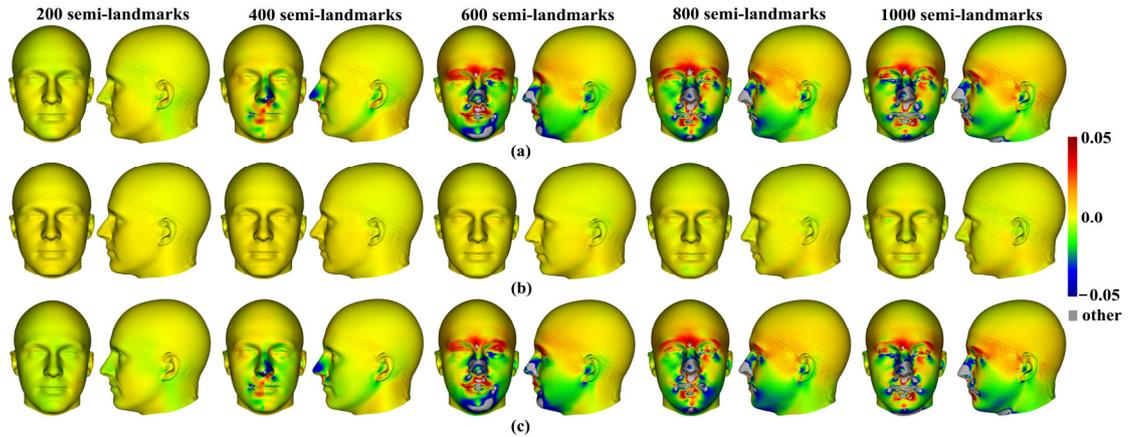


Figure 3.7 Visualisation of the differences in mesh triangle surface areas among mean surface shapes generated using different semilandmarking approaches after re-semilandmarking and re-warping. Differences between (a) sliding TPS (reference) and LS&ICP (target) approaches. (b) Sliding TPS (reference) and TPS&NICP (target) approaches. (c) TPS&NICP (reference) and LS&ICP (target) approaches. Scale bar indicates difference in local area between reference and target surfaces expressed as a proportion of the reference area. Other refers to the values outside the range of the scale bar.

The analyses described above were repeated with the ape cranial surfaces generated by mean landmarks and semilandmarks. Figure 3.8 presents the mean surfaces estimated by each semilandmarking approach at varying densities of semilandmarking. As with the headspace data, they appear very similar to the naked eye, with that from LS&ICP appearing slightly different (e.g. zygomatic region), from those derived by sliding TPS and TPS&NICP, especially at higher semilandmarking densities.

Because LS&ICP yields unreasonable semilandmarks among ape crania (red points in Figure 3.2b) and results in distinctive estimates of mean ape cranial shape, especially at higher densities of semilandmarking (Figure 3.8c), we focus on comparison of mean surfaces based on semilandmarks of varying density from sliding TPS and TPS&NICP. Procrustes distances between the coordinates of all vertices of the surface mesh of ape crania warped to the mean landmark and semilandmark configurations are presented in Table 3.1. These indicate that differences between the full sets of vertices of the mean surfaces generated from sliding TPS and TPS&NICP become greater with increasing density. As with the headspace data (Figure 3.6a, b), Procrustes distances based on the mean landmarks and semilandmarks of ape crania (Table 3.1) are similar to those based on the vertices of the surface meshes warped to fit them (Table 3.2), and the Pearson correlation between these vectors of distances is 0.9940.

Table 3.1 Procrustes distances computed between the mean surfaces of ape crania generated by Sliding TPS and TPS&NICP after re-semilandmarking and re-warping the template mesh.

	50	100	200	400	800
dist	0.0018	0.0025	0.0024	0.0027	0.0030

Table 3.2 Procrustes distances computed between the mean landmarks and semilandmarks of ape crania generated by Sliding TPS and TPS&NICP after re-semilandmarking the warped mesh.

	50	100	200	400	800
dist	0.0023	0.0029	0.0029	0.0032	0.0034

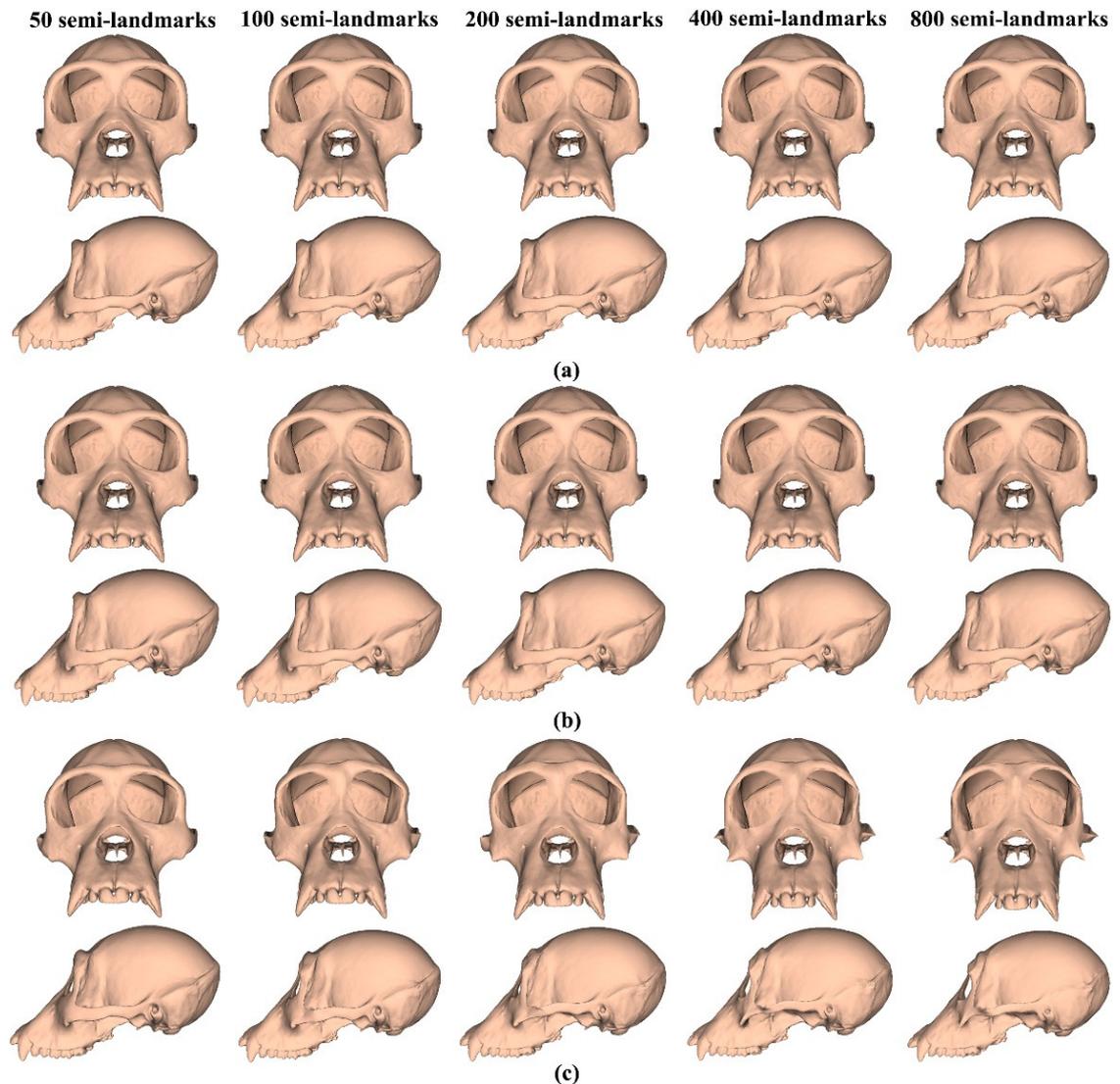


Figure 3.8 Surface meshes of the estimated mean ape cranium generated using different semilandmarking approaches after re-semilandmarking and re-warping. (a) Sliding TPS. (b) TPS&NICP. (c)LS&ICP.

The regional differences between mean surfaces of ape crania derived from sliding TPS (reference surface) and TPS&NICP approaches are illustrated in Figure 3.9. This figure reflects the Procrustes distances of Table 3.1 in indicating that differences in mean surfaces become greater with increasing semilandmarking density. The differences are concentrated in the vicinity of more complex surface regions, e.g. sagittal crests, supraorbital ridges, zygomatic arch, temporal fossa and nuchal crest.

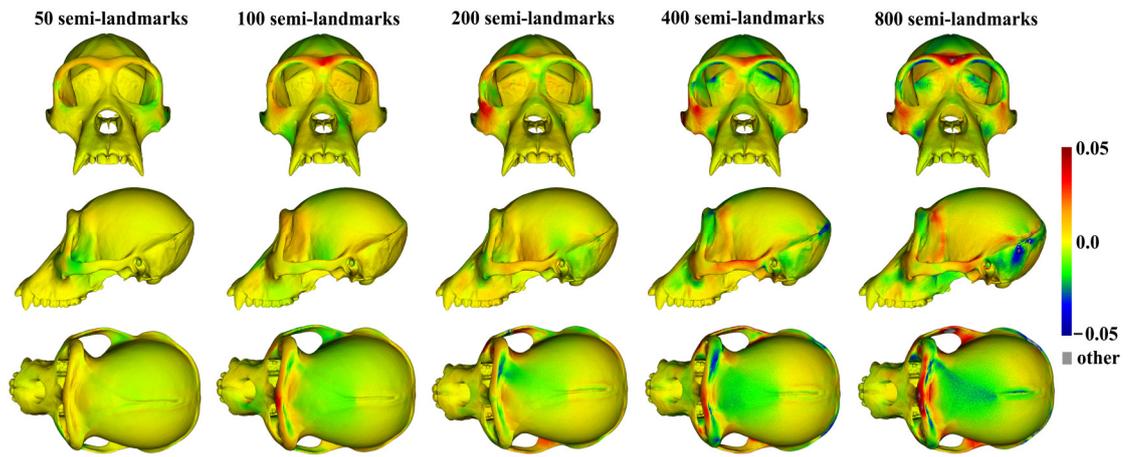


Figure 3.9 Visualisation of the regional differences in local surface areas of mean ape cranial shapes from sliding TPS (reference) and TPS&NICP (target) approaches. Scale bar indicates difference in local area between reference and target surfaces expressed as a proportion of the reference area. Other refers to the values outside the range of the scale bar.

3.3.1.2 Different densities of semilandmarks

The vertices of estimated mean surfaces from every semilandmarking approach and density, were submitted to separate GPA and PCA. Superimposed scatterplots of the first two PCs from each analysis are presented in Figure 3.10, and the proportion of the total variance explained by each axis is expressed as a percentage and tabulated in Table 3.3. Superimposition facilitates visual appraisal of differences in PC scores derived using each semilandmarking approach and density.

Table 3.3 Percentages of total variance explained by PC 1 and PC 2 of mean surface shape.

	Human heads		Ape crania	
	PC1	PC2	PC1	PC2
Sliding TPS	79.96%	7.37%	42.12%	27.51%
LS&ICP	68.32%	9.68%	-	-
TPS&NICP	64.15%	13.48%	42.33%	24.76%

Figure 3.10a shows the superimposed scatterplots of PC1 vs PC2 from separate PCAs of the estimates of the mean surface of the human heads obtained using each semilandmarking approach. The sliding TPS and TPS&NICP approaches result in very similar PC plots, while the PCA of estimated mean surfaces generated by LS&ICP results in a plot showing a similar pattern of variation among means, but with greater variance on both PCs (larger scatter). Similarly, for estimates of the mean surface among the ape crania derived using the sliding TPS and TPS&NICP approaches, the first two PCs from each separate PCA are superimposed in Figure 3.8b. These plots indicate that sliding TPS and TPS&NICP produce very similar scatters of estimated means. Both plots of Figure 3.10 present ‘U’-shaped curves, with the means estimated using the lowest and highest densities of semilandmarks having higher scores on PC2, although they are widely separated on PC1.

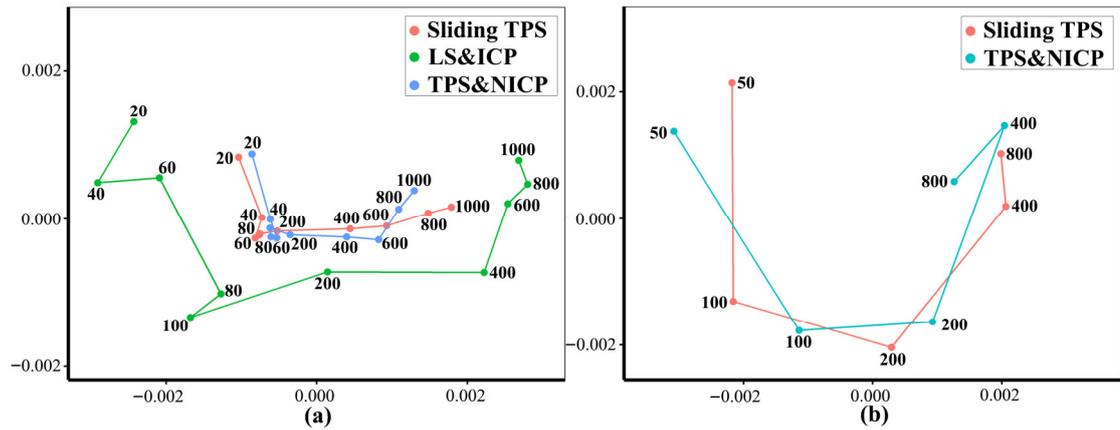


Figure 3.10 Superimposed scatterplots of PC 1 and PC 2 of mean shape surfaces using sliding TPS (red points), LS&ICP (green points) and TPS&NICP (cyan points) approaches. (a) Human heads. (b) Ape crania.

These results are supported by Procrustes distances, computed between mean surface mesh vertices derived from each lower density and the maximum density of semilandmarks, as shown in Tables 3.4 and 3.5, and Figure 3.11. For both datasets, all semilandmarking approaches show convergence between the surfaces based on increasing numbers of semilandmarks and that based on the maximum number. For the headspace data (Table 3.4, Figure 3.11a) sliding TPS and TPS&NICP perform similarly, in that they result in mean surfaces based on <1000 semilandmarks that are closer to that based on 1000 semilandmarks than their equivalents from LS&ICP. Likewise sliding TPS and TPS&NICP perform similarly and show convergence for the ape cranial dataset (Table 3.5, Figure 3.11b)

Table 3.4 Procrustes distances between the vertices of the estimated mean human head surfaces using 1000 semilandmarks and those using increasing numbers of semilandmarks from each approach after re-semilandmarking and re-warping the template mesh. Sliding TPS (first row), LS&ICP (middle row), and TPS&NICP (bottom row).

	20	40	60	80	100	200	400	600	800
Sliding TPS	0.0029	0.0026	0.0027	0.0026	0.0026	0.0024	0.0016	0.0011	0.0007
LS&ICP	0.0055	0.0058	0.0050	0.0046	0.0050	0.0037	0.0027	0.0020	0.0015
TPS&NICP	0.0023	0.0020	0.0020	0.0021	0.0021	0.0019	0.0014	0.0012	0.0008

Table 3.5 Procrustes distances between the vertices of mean ape cranial surfaces estimated by each approach, using 800 semilandmarks and those estimated using increasing numbers of semilandmarks after re-semilandmarking and re-warping the template mesh.

	50	100	200	400
Sliding TPS	0.0050	0.0051	0.0044	0.0034
TPS&NICP	0.0049	0.0044	0.0038	0.0036

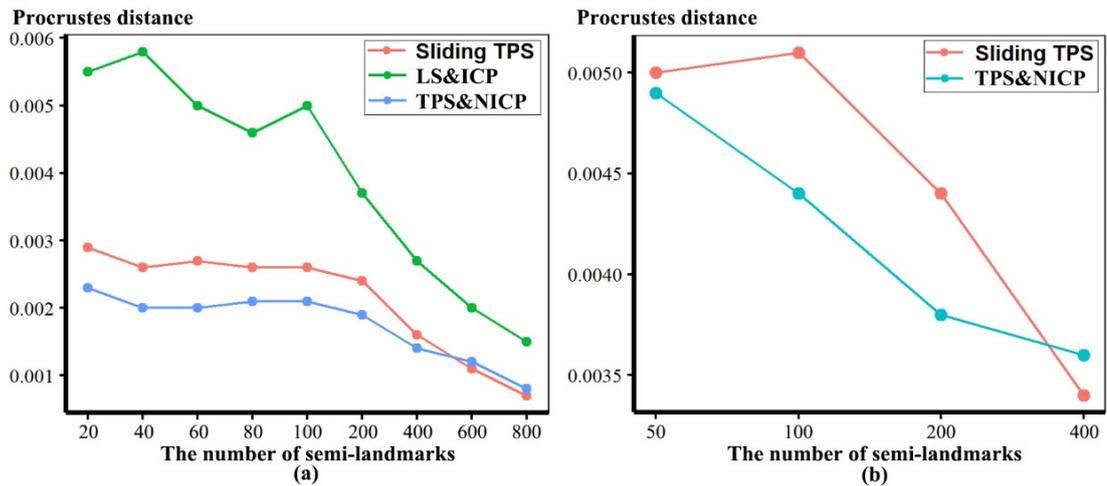


Figure 3.11 Procrustes distances of estimated mean surfaces between every density and the maximum density from each approach to semilandmarking. (a) Human heads. (b) Ape crania.

Within each approach to semilandmarking, the local variations in area between human head mean surfaces estimated by increasing densities of semilandmarks and the surface from 1000 semilandmarks were visualised as colour maps. These are presented in Figures 3.12a-c. Consistent with the Procrustes distances presented in Table 3.4, the closest fitting surfaces are between the surfaces derived using semilandmarks from sliding TPS and TPS&NICP. The colour maps comparing these surfaces with those from 1000 semilandmarks are relatively smooth (Figures 3.12a, b). Further, as semilandmarking density increases, the surfaces based on lower densities of semilandmarks converge with the surface from 1000 semilandmarks. Differences are more pronounced between surfaces derived using lower densities of semilandmarks, and 1000 semilandmarks generated by the LS&ICP approach. This reflects the generally greater Procrustes distances presented in Table 3.4, and visually, differences are most evident in the face (Figure 3.12c). The nasal, ocular and perioral regions show localized large differences but converge with increasing semilandmarking density on the surface derived using 1000 semilandmarks, particularly around the nose and eyes. However, with increasing densities of semilandmarks generated by LS&ICP the quality of the mean surfaces is poor (i.e. less sharp features around eyes and mouth in Figure 3.5c), because equivalent semilandmarks lie in different anatomical locations.

Similar comparisons were undertaken for the ape crania. Figures 3.13a, b show regional differences in area of mean surfaces computed between lower densities and the maximum density of 800 semilandmarks generated by the sliding TPS and TPS&NICP approaches. In both, the smallest differences are found in the cranial vault, where the colour map is smooth and indicative of small local area differences. Larger differences are observed around the frontal bone, supraorbital ridges, zygomatic arches, malar region, nasal bones and maxillae. Consistent with Table 3.5, with increasing semilandmark density, a degree of convergence occurs with the surface based on 800 semilandmarks.

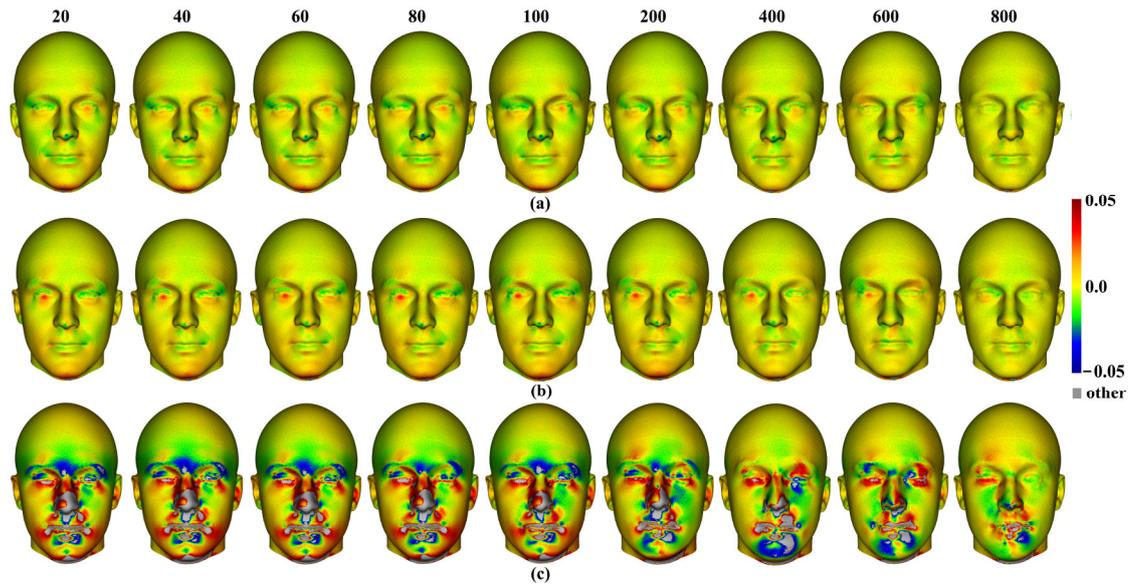


Figure 3.12 Colour maps of local surface mesh area differences among mean surfaces of human heads derived using lower densities (target) and that from 1000 semilandmarks (reference) using different semilandmarking approaches. (a) Sliding TPS. (b) TPS&NICP (c) LS&ICP. Scale bar indicates difference in local area between reference and target surfaces expressed as a proportion of the reference area. Other refers to the values outside the range of the scale bar.

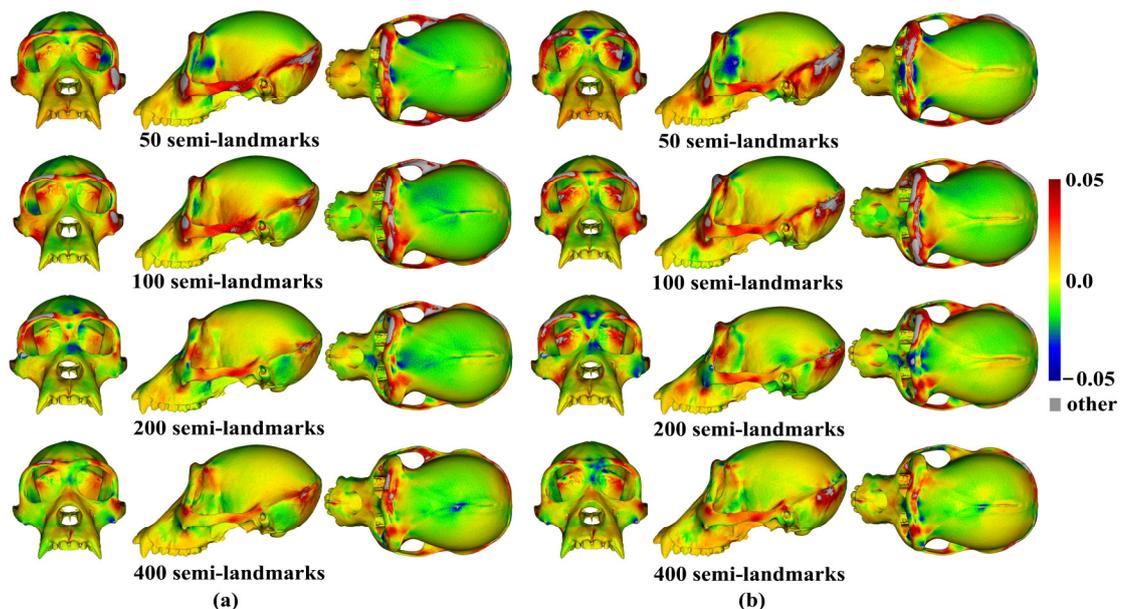


Figure 3.13 Colour maps of local surface mesh area differences among mean surfaces of ape crania derived using lower densities (target) and 800 semilandmarks (reference) based on different semilandmarking approaches. (a) Sliding TPS. (b) TPS&NICP. Scale bar indicates difference in local area between reference and target surfaces expressed as a proportion of the reference area. Other refers to the values outside the range of the scale bar.

3.3.2 Comparison of estimates of allometrically scaled surfaces

We generated surface meshes warped by TPS to the predicted landmark and semilandmark configurations representing the extreme limits (maximum and minimum centroid sizes) of the allometric vector then, after re-sliding and re-warping, we assessed the overall and regional differences among surfaces generated by different semilandmarking approaches and densities.

3.3.2.1 Different semilandmarking approaches

Procrustes distances between the vertices of the allometrically scaled surfaces of human heads representing the maximum centroid size generated by different approaches are illustrated in Figure 3.14a. Likewise, Procrustes distances between the fitted surfaces representing the minimum centroid size are illustrated in Figure 3.14b. In both cases, in comparisons between LS&ICP and the other two approaches, Procrustes distances between surface meshes increase with increasing numbers of semilandmarks while those between sliding TPS and TPS&NICP decrease. Sliding TPS and TPS&NICP approaches result in the most similar predictions as semilandmarking density increases. The distances among predicted shapes at minimum size are somewhat greater those at maximum size because of the skewed distribution of centroid sizes (see Figure 3.17).

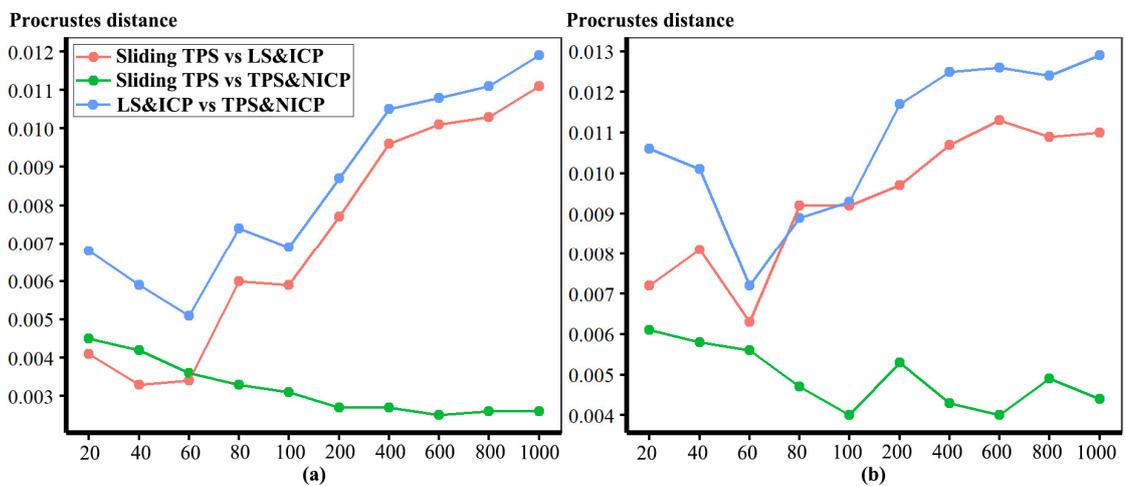


Figure 3.14 Procrustes distances computed between the vertices of human head surfaces, allometrically scaled to the maximum and minimum centroid sizes, based on different semilandmarking approaches after re-semilandmarking and re-warping the template mesh. (a) Maximum. (b) Minimum.

Additionally, the differences among allometric predictions of large and small surfaces were visualised between different semilandmarking approaches and different densities of semilandmarks in Figure 3.15. The visualizations show differences in surface area of equivalent triangles among re-warped to re-semilandmarked surface meshes predicted for the maximum centroid size, in Figure 3.15a, and those corresponding to the minimum centroid size are illustrated in Figure 3.15b. In both cases the differences between surface mesh predictions based on landmarks and semilandmarks from sliding TPS and TPS&NICP are small (middle rows in Figures 3.15a, b). They reflect the Procrustes distances in Figures 3.14a, b in becoming more similar with increasing semilandmark density and in being more similar for comparisons among predictions of the surface at the maximum centroid size than at the minimum. The differences among surface meshes predicted by LS&ICP and the other approaches (Top and bottom rows in Figures 3.15a, b) also reflect the Procrustes distances in Figures 14a, b in being large, becoming larger with increasing density, and in being larger for comparisons of the predicted surfaces at the minimum centroid size.

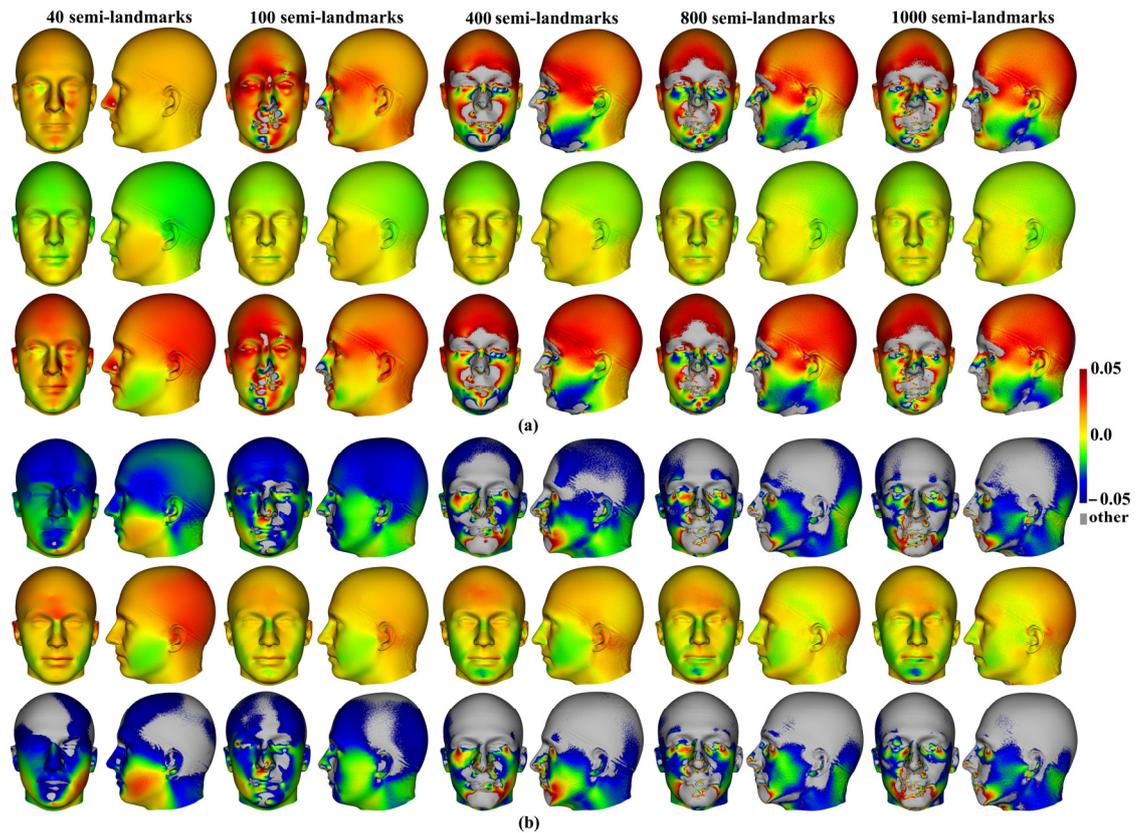


Figure 3.15 Visualisation of the differences in mesh triangle surface areas among predicted allometrically scaled surfaces of human heads representing the a) maximum and b) minimum centroid size generated by different semilandmarking approaches. In each figure, top row: sliding TPS (reference) vs LS&ICP (target), middle row: sliding TPS (reference) and TPS&NICP (target) and bottom row: TPS&NICP (reference) vs LS&ICP (target). Scale bar indicates difference in local area between reference and target surfaces expressed as a proportion of the reference area. Other refers to the values outside the range of the scale bar.

Similar visualisations compared allometrically scaled surfaces of the ape cranial dataset. The LS&ICP approach was not evaluated, because it failed to produce sensible semilandmarks when applied to these more complex and variable surfaces. Procrustes distances, between the mesh vertices of predicted cranial surface corresponding to the maximum and minimum centroid size estimated using the sliding TPS and TPS&NICP approaches are compared in Table 3.6. These distances indicate that differences among both allometric predictions of the surface increase with increasing semilandmark density, as with the comparison among means from the ape data estimated using sliding TPS and TPS&NICP approaches (Table 3.1). From Table 3.6, Procrustes distances at the maximum centroid size are less than those at the minimum, consistent with the skewing of the distribution of centroid sizes towards the maximum, which results in greater allometric warping of the mean shape towards the minimum than the maximum centroid size (see Figure 3.18). Further, the Procrustes distances are somewhat larger than those between estimated mean surfaces in Table 3.1, indicating greater differences in among allometrically scaled surfaces.

Table 3.6 Procrustes distances computed between vertices of ape cranial surfaces, allometrically scaled to the maximum (Max) and minimum (Min) centroid size, from sliding TPS and TPS&NICP semilandmarking approaches after re-semilandmarking and re-warping the template mesh.

	50	100	200	400	800
Max	0.0040	0.0039	0.0044	0.0052	0.0055
Min	0.0072	0.0056	0.0088	0.0077	0.0100

These differences are visualized in Figure 3.16 where consistent with the Procrustes distances of Table 3.6, differences in mesh triangle surface areas increase with semilandmarking density, are greater for the estimates of the mean ape cranium scaled to the minimum centroid size and are most pronounced around more complex surface regions e.g. periorbital region, crests, infratemporal region.

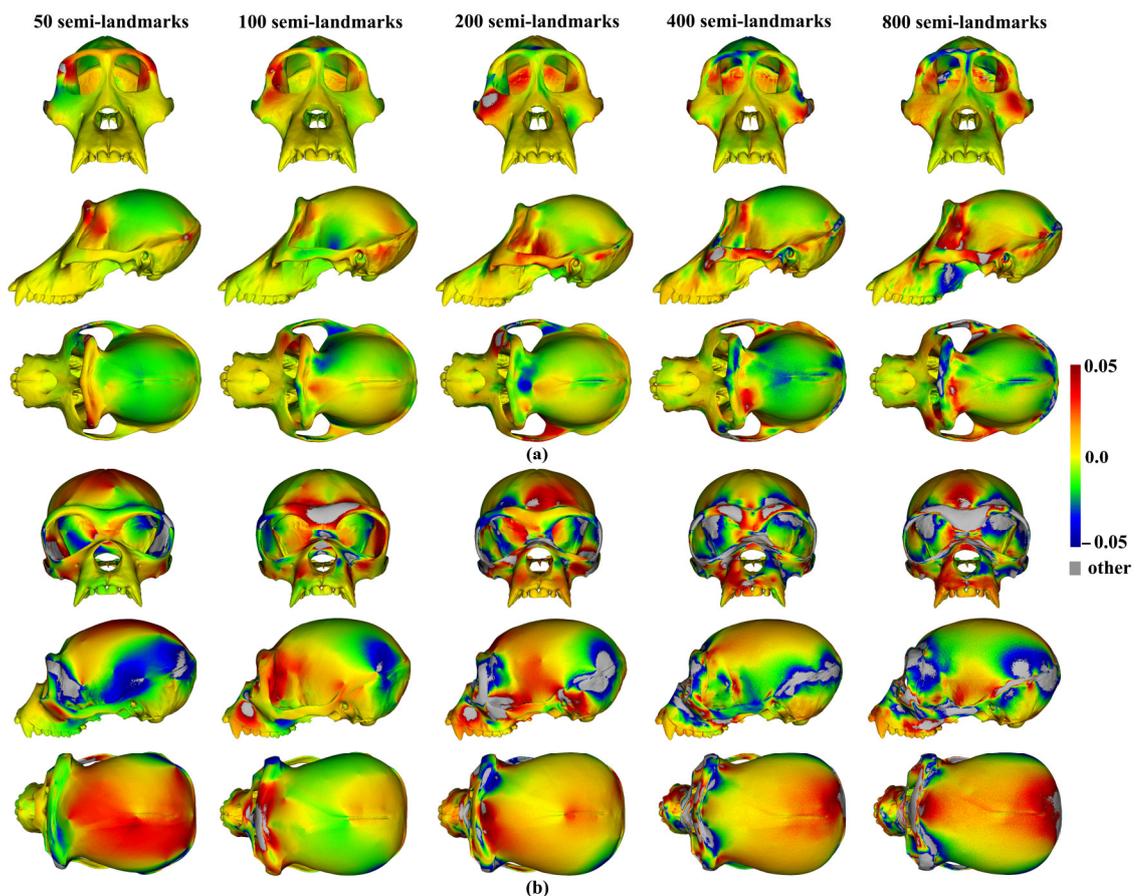


Figure 3.16 Visualisation of the differences in mesh triangle surface areas among fitted surface shapes of ape crania generated by the sliding TPS (reference) and TPS&NICP (target) approaches after re-semilandmarking and re-warping of the template. (a) Comparison of predictions corresponding to the maximum centroid size. (b) Comparison of predictions corresponding to the minimum centroid size. Scale bar indicates difference in local area between reference and target surfaces expressed as a proportion of the reference area. Other refers to the values outside the range of the scale bar.

Finally, these predictions are compared through GPA and PCA of allometrically scaled mesh vertices created using varying numbers of semilandmarks from each semilandmarking approach. The first two PCs from PCAs of the mean and allometrically scaled head surfaces are presented in Figure 3.17, and those of the ape surfaces in Figure 3.18. The first two PCs in both of these analyses account for nearly all of the variance

among surfaces (heads 97%, ape crania >99%) and so they well represent the differences among them well.

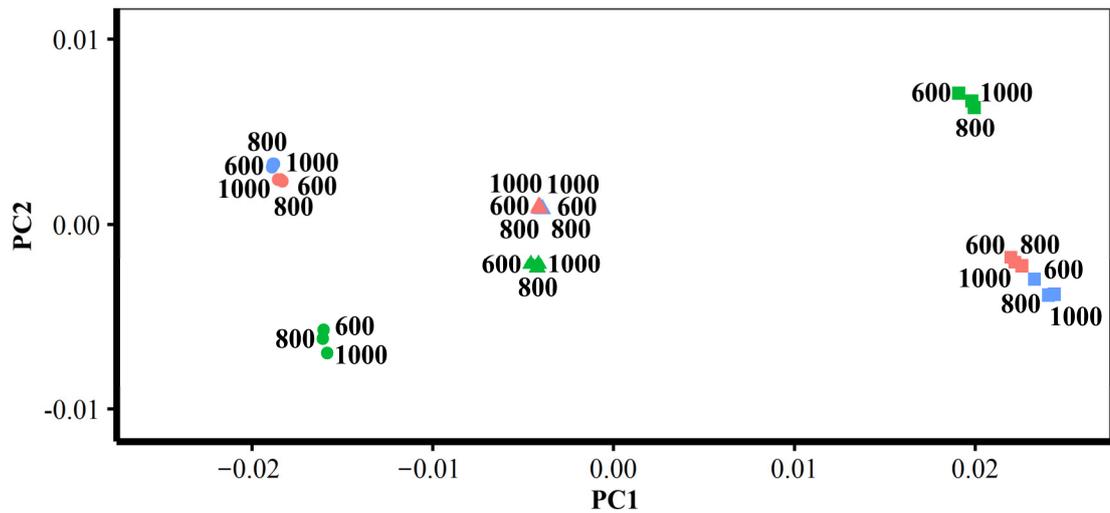


Figure 3.17 PC1 (92.4% of total variance) vs PC2 (4.7%) from PCA of the mean and allometrically scaled head surfaces derived using varying densities of semilandmarks and each semilandmarking approach. Triangles = means, squares = allometric predictions of surfaces at the sample minimum centroid size, circles = allometric predictions of surfaces at the sample maximum centroid size. Red = sliding TPS, blue = TPS&NICP, green = LS&ICP. Numbers indicate number of semilandmarks. The sliding TPS and TPS&NICP means are nearly superimposed.

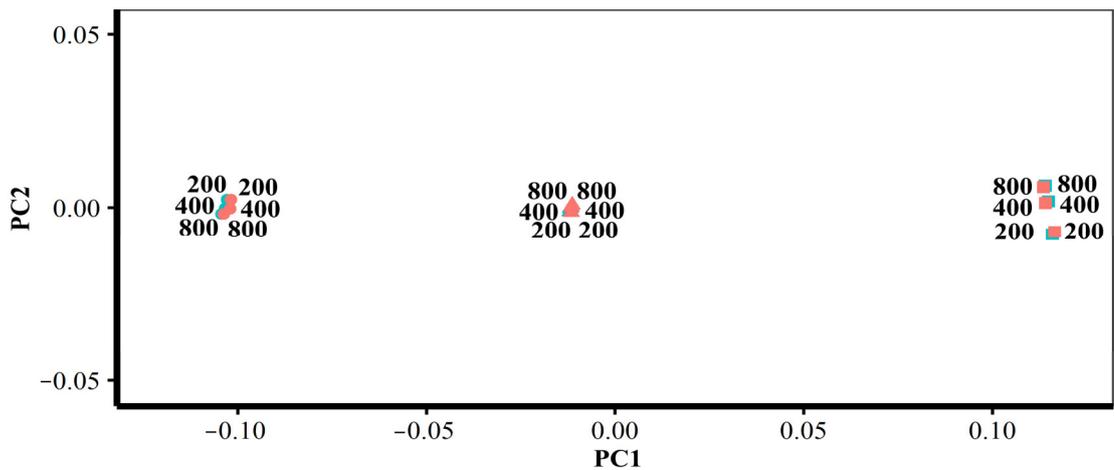


Figure 3.18 PC1 (99.5% of total variance) vs PC2 (0.15%) from PCA of the mean and allometrically scaled head surfaces derived using varying densities of semilandmarks and the sliding TPS and TPS&NICP semilandmarking approaches. Triangles = means, squares = allometric predictions of surfaces at the sample minimum centroid size, circles = allometric predictions of surfaces at the sample maximum centroid size. Red = sliding TPS, blue = TPS&NICP. Numbers indicate number of semilandmarks. Means are nearly superimposed.

Consistent with the visualisations of Figure 3.15 and the Procrustes distances in Table 3.4 and Figure 3.14, the PC plot of head data (Figure 3.17) shows that sliding TPS and TPS&NICP achieve very similar results (surfaces) with the means plotting on top of each other, the allometric predictions of the mean surface at the sample maximum centroid size (PC1 left, circles) grouping closely and those at the sample minimum centroid size (PC1 right, rectangles) being more variable. The mean and allometrically scaled surfaces from

LS&ICP (green) are somewhat dissimilar based on the PC plots, Procrustes distances and colour maps. Likewise, the PC plot of ape cranial surfaces (Figure 3.18) is consistent with the visualisations of Figure 3.16 and the Procrustes distances in Tables 3.5 and 3.6. It shows that the mean and allometrically scaled surfaces of ape crania derived using sliding TPS and TPS&NICP are very similar to each other, with smaller variance among the predictions of surface mesh shape at the sample maximum centroid size (PC1 left, circles) than those at the sample minimum centroid size (PC1 right, rectangles).

Further, the plots of Figures 3.17 and 18 serve to give perspective to differences seen in Table 3.6 and Figures 3.14-16. While the colour maps are highly sensitive to differences in surfaces and identify many regions of difference, when they are set against the differences among the estimates of the means and allometrically scaled means in the PC plots, they appear much more similar, especially for comparisons of results obtained using the sliding TPS and TPS&NICP approaches at all semilandmarking densities.

3.3.2.2 Different densities of semilandmarks

For each semilandmarking approach and dataset, the differences in shape between the allometrically scaled surfaces derived from lower densities of semilandmarks and those from the maximum density were assessed by computing the Procrustes distances among their vertices and visualizing differences in local surface areas. For the head surfaces, Table 3.7 presents, and Figure 3.19 plots these Procrustes distances. In both cases, the sliding TPS and TPS&NICP approaches consistently result in surfaces from lower semilandmarking densities being more similar (smaller Procrustes distances) to the surface with the maximum semilandmarking density than is the case for those derived using the LS&ICP approach. Further, at lower semilandmarking densities distances from the TPS&NICP approach are slightly smaller than those from sliding TPS. Procrustes distances are a little larger among predicted surfaces at the sample minimum centroid size, especially at lower semilandmarking densities, than among those at the sample maximum centroid size, because of the skewed distribution of centroid sizes (see Figure 3.17).

Table 3.7 Procrustes distances between vertices of the allometrically scaled surfaces of heads at the maximum and minimum centroid sizes using the landmarks and highest density of semilandmarks and surfaces estimated using the landmarks and lower densities of semilandmarks, after re-semilandmarking and re-warping. Sliding TPS (first two rows), LS&ICP (middle two rows), and TPS&NICP (other rows).

Size	20	40	60	80	100	200	400	600	800
Max	0.0075	0.0063	0.0066	0.0059	0.0056	0.0048	0.0031	0.0019	0.0014
Min	0.0137	0.0102	0.0100	0.0091	0.0084	0.0083	0.0048	0.0031	0.0022
Max	0.0140	0.0124	0.0120	0.0112	0.0103	0.0082	0.0050	0.0044	0.0030
Min	0.0182	0.0160	0.0128	0.0145	0.0137	0.0104	0.0073	0.0067	0.0037
Max	0.0067	0.0049	0.0050	0.0041	0.0038	0.0036	0.0025	0.0019	0.0012
Min	0.0119	0.0084	0.0079	0.0081	0.0083	0.0074	0.0056	0.0036	0.0017

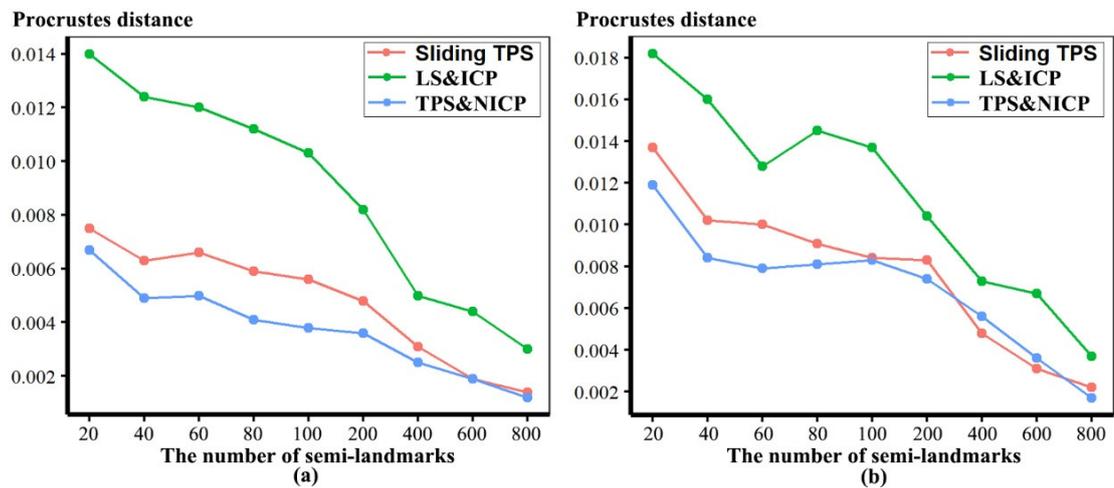


Figure 3.19 Procrustes distances, after re-semilandmarking and re-warping, between the allometrically scaled head surfaces derived from the maximum density and those from lower densities of semilandmarks. (a) Procrustes distances between predicted surfaces at the maximum centroid size. (b) Procrustes distances between predicted surfaces at the minimum centroid size.

Additionally, local differences in area between allometric predictions of head surfaces derived using the maximum density semilandmarks and lower densities, from each semilandmarking approach are visualized in the colour maps of Figure 3.20. Figures 3.20a, b present the visualisations corresponding to the sample maximum and minimum centroid size, respectively. Consistent with the Procrustes distances presented in Table 3.7 and Figure 3.19, the LS&ICP approach shows the greatest differences between surfaces derived from lower densities and the maximum while sliding TPS and TPS&NICP perform similarly. In all cases, shape differences between lower and the maximum semilandmarking densities become smaller with increasing density. The greatest differences between densities of semilandmarking are found around nose mouth, ears and chin, where topography is complex and the smallest around the forehead, scalp, where the surface is smooth and lacks identifiable landmarks.

These analyses were repeated using the allometric predictions of ape cranial surfaces between every density and the maximum density of semilandmarks generated by sliding TPS and TPS&NICP, respectively. The Procrustes distances between allometrically scaled predictions of the ape crania from varying semilandmarking densities and those from the maximum semilandmarking density are presented in Table 3.8 and plotted in Figure 3.21. These are very similar in magnitude for surfaces derived using both sliding TPS and TPS&NICP approaches at all densities and, with increasing density, show a similar trend of convergence on the surface derived using 800 semilandmarks. Procrustes distances between this surface and those derived using lower density semilandmarks are greater for estimates of the allometric predictions of surfaces at the minimum centroid size than at the maximum. This reflects the skewed distribution of centroid sizes, in particular, the greater difference between the overall mean and the predicted mean surface at the minimum than at the maximum centroid size (see Figure 3.18).

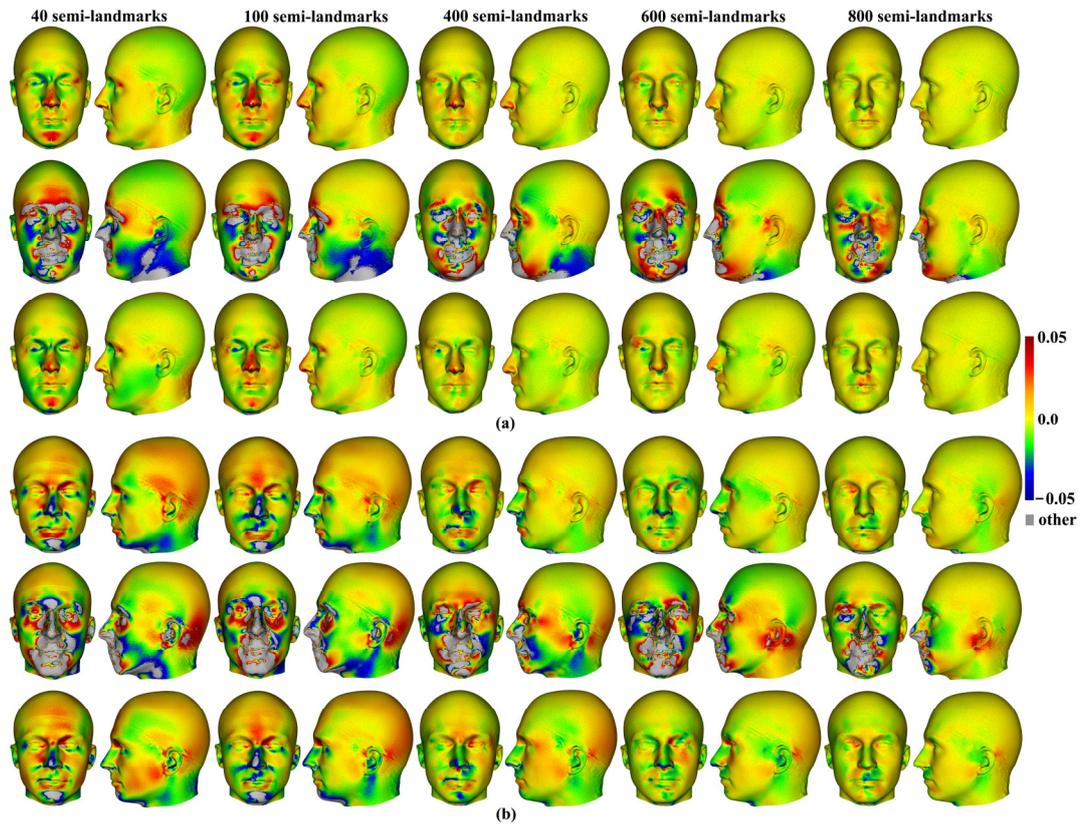


Figure 3.20 Colour map of local area differences computed between the re-semilandmarked and re-warped allometric predictions of surfaces of human heads at the (a) maximum and (b) minimum sample centroid sizes, computed between lower densities (reference) and the maximum density (target) of semilandmarking. In each figure, top row: sliding TPS, middle row: LS&ICP and bottom row: TPS&NICP. Scale bar indicates difference in local area between reference and target surfaces expressed as a proportion of the reference area. Other refers to the values outside the range of the scale bar.

Table 3.8 Procrustes distances between vertices of the estimated predictions of ape cranial surfaces at the maximum and minimum centroid sizes derived from the maximum density of semilandmarks and those from lower densities of semilandmarks, after re-semilandmarking and re-warping.

	Size	50	100	200	400
Sliding TPS	Max	0.0139	0.0094	0.0076	0.0058
	Min	0.0267	0.0191	0.0146	0.0104
TPS&NICP	Max	0.0141	0.0099	0.0075	0.0058
	Min	0.0272	0.0199	0.0157	0.0119

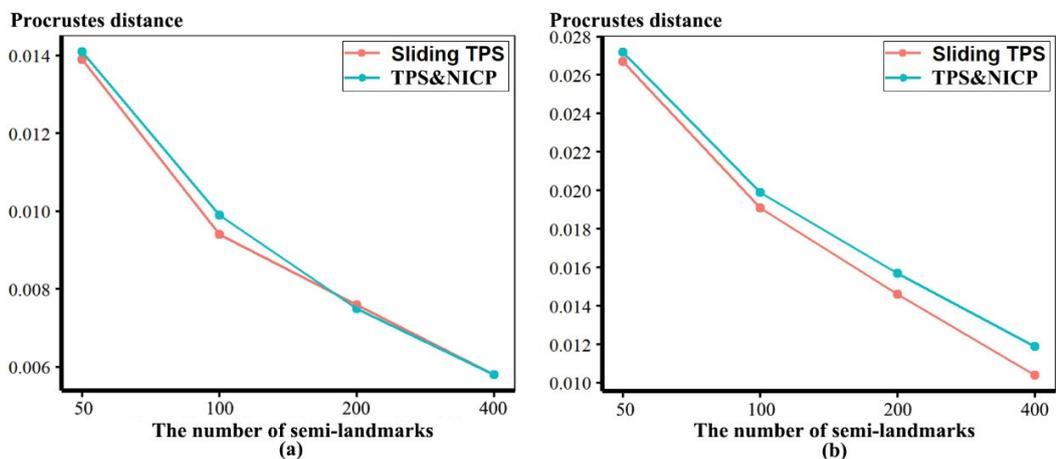


Figure 3.21 Procrustes distances, after re-semilandmarking and re-warping, between the allometrically

scaled ape cranial surfaces derived from the maximum density and those from lower densities of semilandmarks. (a) Procrustes distances between predicted surfaces at the maximum centroid size. (b) Procrustes distances between predicted surfaces at the minimum centroid size.

These localised variations in surface areas of the allometrically scaled surfaces are visualised in Figure 3.22. These visualisations reflect the Procrustes distances presented in Table 3.8 and Figure 3.21 in showing greater differences among densities of semilandmarking for the allometric predictions of the ape crania at the minimum centroid size than those at the maximum and convergence with increasing semilandmarking density. The largest shape differences are observed in the facial region, zygomatic arches, supraorbital, temporal and nuchal regions, where surface topography is most complex, and the least are observed over the cranial vault.

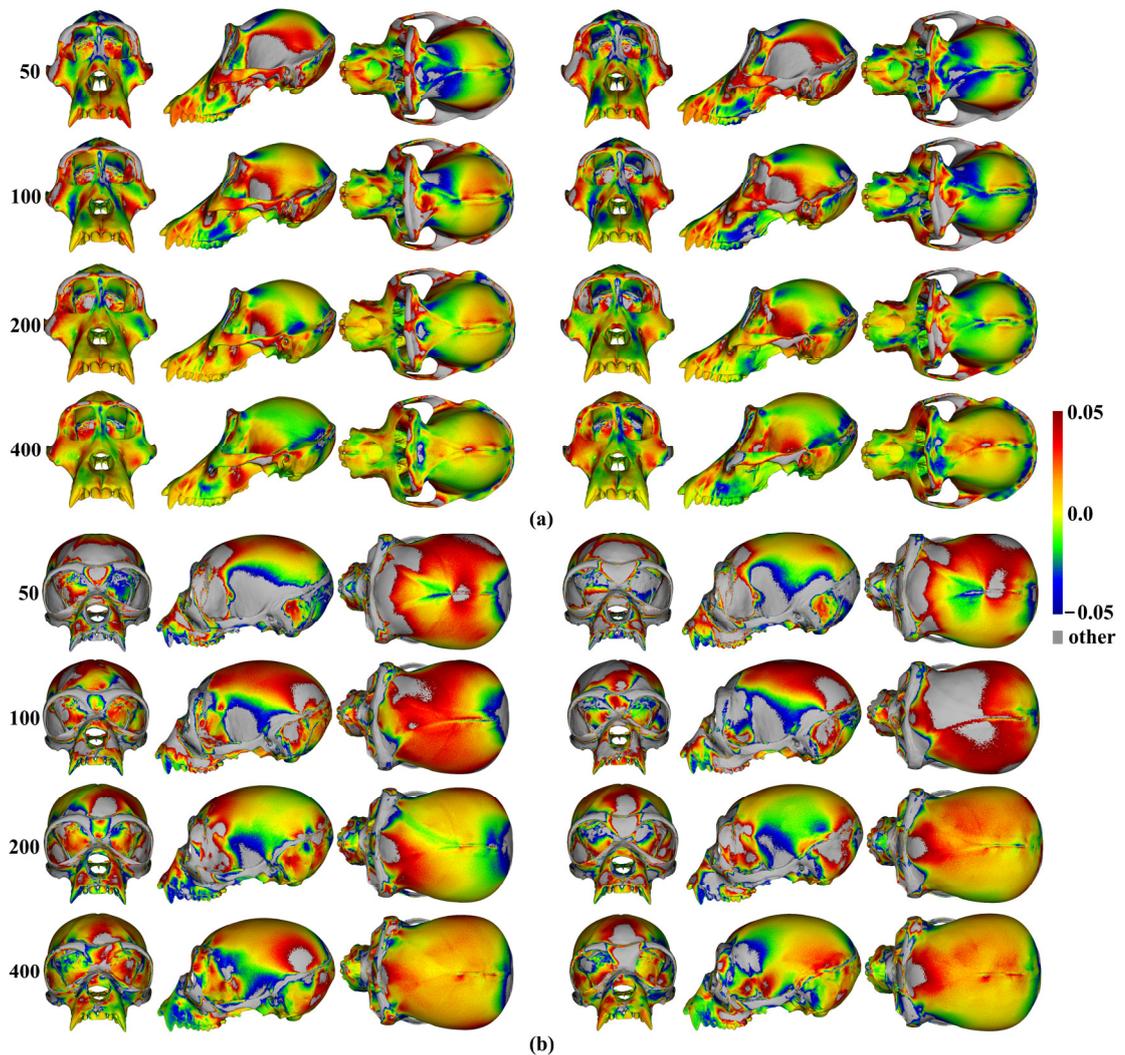


Figure 3.22 Colour map of local area differences computed between the re-semilandmarking and re-warped allometric prediction of surfaces of ape crania at the (a) maximum and (b) minimum sample centroid sizes, computed between lower densities (reference) and the maximum density (target) of semilandmarking. Left: sliding TPS and right: TPS&NICP. Scale bar indicates difference in local area between reference and target surfaces expressed as a proportion of the reference area. Other refers to the values outside the range of the scale bar.

3.3.3 Comparisons of mean and allometrically scaled surface resulting from landmarks alone

For each dataset, the mean surface from sliding TPS and 1000 semilandmarks was compared with warped surfaces derived using only the landmarks. The template surfaces, which are themselves an initial estimate of the average surface (see methods), were fitted to the mean landmarks (Figure 3.23 for heads and Figure 3.24 for apes). In practice, it is common to use the surface of an individual close to the mean for visualisation, yet the effects of choice of surface are unclear. Therefore, the estimation of mean surfaces was repeated, this time using the head surface with the smallest Procrustes distance to the mean and the ape surface (USNM 176211) used to generate the ape template, respectively. The resulting predictions of mean head surfaces for each dataset were compared using colour maps of local mesh surface area changes (Figure 3.23). For the ape surface comparison, two different colour maps were drawn, the first using the same colour scale range used in preceding analyses, to allow direct comparison with them, the second using an extended range to better visualise the full range of local area differences (Figure 3.24).

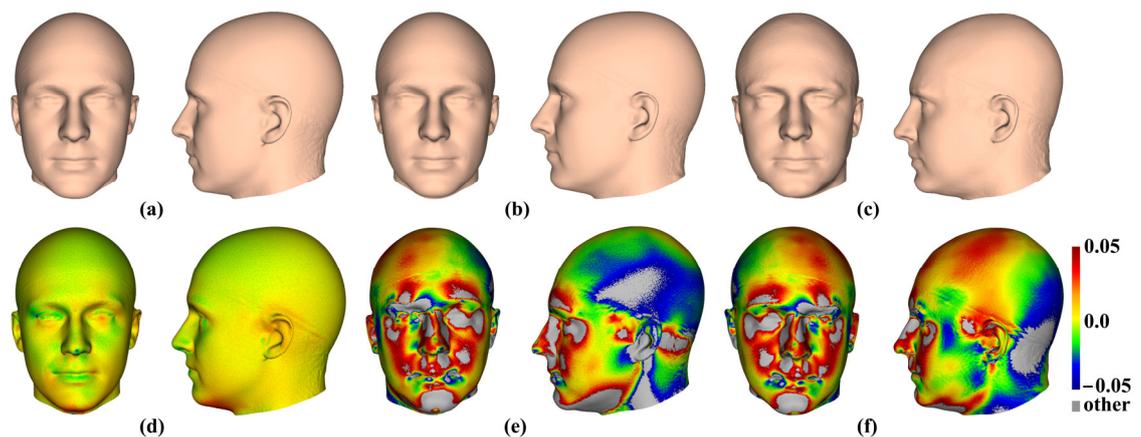


Figure 3.23 The effects of different template surfaces and semilandmarks on the visualisations of mean head. (a) Mean head surface estimated by warping the template surface to the mean configuration of landmarks and 1000 semilandmarks from sliding TPS. (b) Mean head surface estimated by warping the template surface to the mean landmark configuration (c) Mean head surface estimated by warping the surface of the head with minimum Procrustes distance from the mean to the mean landmark configuration. (d) Colour map between the surfaces a (reference) and b (target). (e) Colour map between the surfaces a (reference) and c (target). (f) Colour map between the surfaces b (reference) and c (target). Other refers to the values outside the range of the scale bar.

To the naked eye, the mean head surfaces (Figures 3.23a-c) differ but to a lesser degree than the mean ape cranial surfaces (Figures 3.24a-c). In both cases the greatest similarity (d, in each figure) is between (a), the surface estimated by warping the template surface to the mean configuration of landmarks and 1000 semilandmarks from sliding TPS and (b), the surface estimated by warping the template surface to the mean landmark configuration. More marked differences (e) are found comparing (a) with (c). In the case of the head dataset, both comparisons with (Figure 3.23c) show very similar differences in the face (Figure 3.23e compared to Figure 3.23f), especially around nose, mouth and

eyes where landmarks are present, but there are contrasting differences over the scalp, which lacks identifiable landmarks to control warping. In the case of the ape cranial dataset where landmarks are distributed over the entire surface, both comparisons with (Figure 3.24d) are similar (Figure 3.24e compared to Figure 3.24f) with the main differences concentrated around crests and ridges.

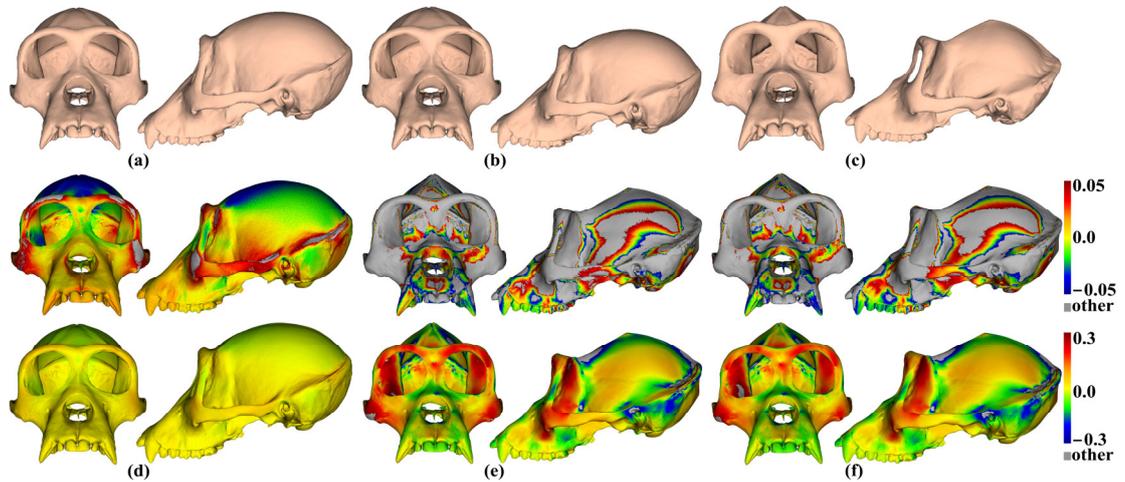


Figure 3.24 The effects of different template surfaces and semilandmarks on the visualisations of mean ape cranium. Top Row: (a) Mean ape cranial surface estimated by warping the template surface to the mean configuration of landmarks and 1000 semilandmarks from sliding TPS. (b) Mean ape cranial surface estimated by warping the template surface to the mean landmark configuration (c) Mean ape cranial surface estimated by warping the surface of the cranium used to generate the template to the mean landmark configuration. Colour maps between the surfaces using different colour ranges (see text) in the middle row and bottom rows: (d) Colour map between the surfaces a (reference) and b (target). (e) Colour map between the surfaces a (reference) and c (target). (f) Colour maps between the surfaces b (reference) and c (target). Other refers to the values outside the range of the scale bar.

Similar analyses are conducted to assess how landmarks alone perform in predicting allometrically scaled surfaces, as might be carried out where no initial estimate of the mean surface is possible (e.g. hand collected landmark data), but a surface mesh is available for warping. These focus on the comparison of the individual head surface closest to the mean and the ape surface (USNM 176211) used to generate the template, warped to the allometrically scaled landmark configurations, with those from allometric scaling of the template surfaces based on all landmarks and the maximum densities of semilandmarks. The results for the predictions of surfaces at the maximum sample centroid sizes are presented in Figure 3.25, and for the minimum centroid sizes in Figure 3.26. Note that the scale bar used to compare ape cranial means is wider than that used elsewhere, because the differences are greater. In both cases, the surfaces of individuals warped to fit the allometrically scaled landmark configurations (Figures 3.25b, e, and Figures 3.26b, e), are superficially similar to those derived by warping the template to the allometrically scaled landmark and highest density semilandmark configurations for each dataset. However, they differ in detail such that the human head surfaces estimated using landmarks alone and the surface of the individual nearest to the mean (Figures 3.25b and 3.26b) present more rounded faces with subtle differences around the eyes, mouth and

nose when compared with the template surfaces warped to the allometrically scaled landmark and semilandmark configurations (Figures 3.25a and 3.26a). The same comparisons for the ape crania (Figure 3.25d vs Figure 3.25e and Figure 3.26d vs Figure 3.26e) present more obvious differences, particularly around sagittal and nuchal crests, orbits and temporal fossae.

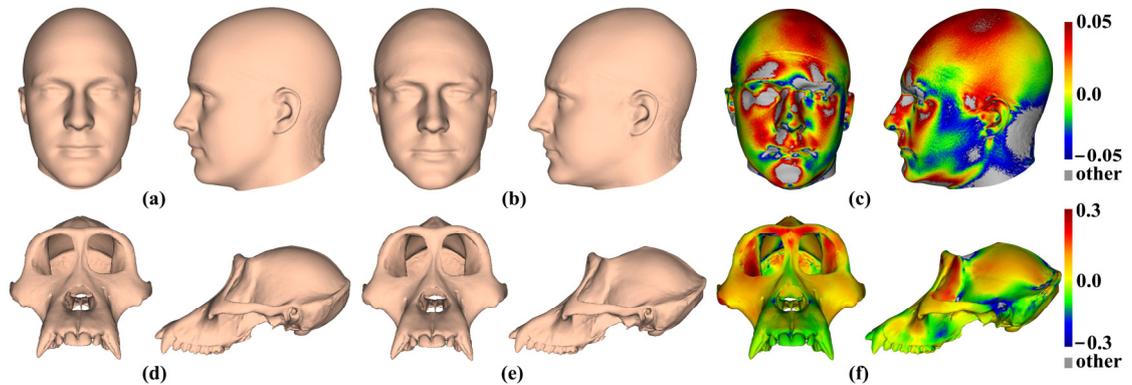


Figure 3.25 The effects of different template surfaces and semilandmarks on the visualisations of allometric prediction of surface at the maximum centroid size. (a) Allometric prediction of head surface at the sample maximum centroid size using the template surface, estimated using landmarks and 1000 semilandmarks from sliding TPS. (b) Allometric prediction of head surface at the sample maximum centroid size using the surface of the head with minimum Procrustes distance to the mean, warped using landmarks alone. (c) Colour map between the surfaces a (reference) and b (target). (d) Allometric prediction of ape surface at the sample maximum centroid size using the template surface, estimated using landmarks and 800 semilandmarks from sliding TPS. (e) Allometric prediction of ape surface at the sample maximum centroid size using the ape cranium used to generate the template, estimated using landmarks alone. (f) Colour maps between the surfaces d (reference) and e (target) using different ranges. Other refers to the values outside the range of the scale bar.

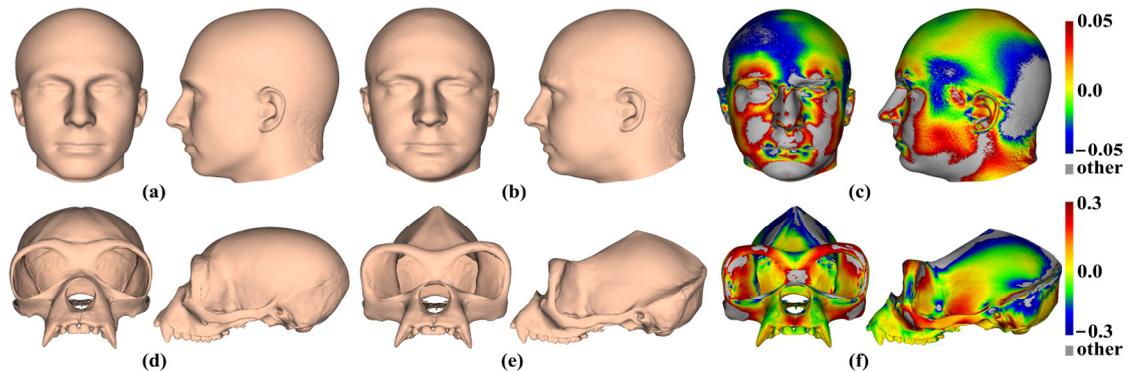


Figure 3.26 The effects of template surface and semilandmarks on the visualisation of allometric prediction of surface at the minimum centroid size. (a) Allometric prediction of mean head surface at the sample minimum centroid size using the template surface, estimated using landmarks and 1000 semilandmarks from sliding TPS. (b) Allometric prediction of mean head surface at the sample minimum centroid size using the surface of the head with minimum Procrustes distance to the mean, warped using landmarks alone. (c) Colour map between the surfaces a (reference) and b (target). (d) Allometric prediction of mean ape surface at the sample minimum centroid size using the template surface, estimated using landmarks and 800 semilandmarks from sliding TPS. (e) Allometric prediction of mean ape surface at the sample minimum centroid size using the ape cranium used to generate the template, estimated using landmarks alone. (f) Colour maps between the surfaces d (reference) and e (target) using different ranges. Other refers to the values outside the range of the scale bar.

These warped surfaces were added to the PCAs of the mean and allometrically scaled head surfaces derived using varying densities of semilandmarks and each semilandmarking approach in Figures 3.17 and 3.18. Figure 3.27 presents plots of PC1 vs PC2 and PC1 vs PC3 (accounting for 95% of the total variance) of the mean and allometrically scaled head surfaces. A further 3% of the total variance is explained by PC4. The plot of PC1 vs PC4 from this analysis (not shown) is very similar to that in Figure 3.17 in indicating a difference in allometric vector direction between LS&ICP and all other approaches. It is clear that the surfaces estimated by warping the surface of the individual head with minimum Procrustes distance from the mean to the mean and allometrically scaled landmark configurations are distinct from those estimated using the template surface and semilandmarks. Additionally, the vector connecting this mean and allometrically scaled means is not parallel to the vector connecting the semilandmark derived mean and scaled surfaces. Further, the template surfaces warped to fit the mean and scaled landmark configurations are arranged along a vector parallel to them, but with the mean near the mean of the surfaces warped using semilandmarks. Thus, while these surfaces are warped to exactly fit the overall mean and the allometrically scaled mean landmark configurations, the regions between the landmarks are deformed in the same way for both surfaces, but differently to the template surface warped to fit the landmark and semilandmark configurations. This is consistent with the visual comparisons of Figure 3.23a vs Figure 3.23c, Figure 3.25a vs Figure 3.25b and Figure 3.26a vs Figure 3.26b.

Similarly, Figure 3.28 presents a plot of PC1 vs PC2 of the mean and allometrically scaled ape cranial surfaces. This plot accounts for 99% of the total variance. The surfaces estimated by warping the ape cranial surface used to generate the template and allometrically scaled landmark configurations are again distinct from those using the template surface and semilandmarks. As with the head surfaces, the vector connecting the semilandmark derived mean and allometrically scaled surfaces is not parallel to the vector connecting these estimates of the mean and allometrically scaled surfaces and the surfaces obtained by warping the template to fit the mean and allometrically scaled landmark configurations lie along a parallel vector to the latter, with the mean near the means of surfaces derived using semilandmarks. Thus, as with the head surfaces, the surface between the landmarks is different to the template surface warped to fit the landmark and semilandmark configurations, and it is deformed differently. Again, this is consistent with the visual comparisons of Figure 3.24a vs Figure 3.24c, Figure 3.25d vs Figure 3.25e and Figure 3.26d vs Figure 3.26e.

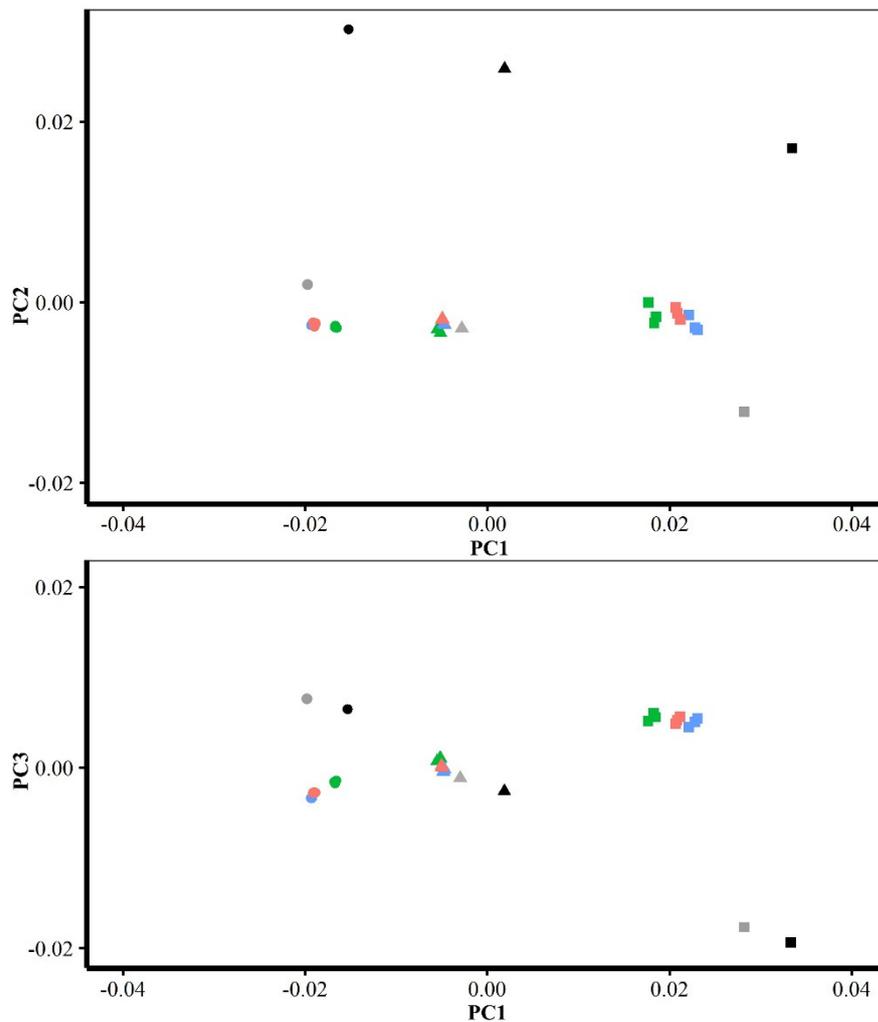


Figure 3.27 Scatter plots of the mean and allometrically scaled head surfaces. Top: PC1 (72.1% of total variance) vs PC2 (17.5%) and bottom: PC1 vs PC3 (5.36%) from PCA of the mean and allometrically scaled head surfaces derived using maximum density of semilandmarks from each semilandmarking approach (from Figure 3.14). Red = sliding TPS, blue = TPS&NICP, green = LS&ICP. Also included in this PCA are surfaces warped to the mean and scaled landmark configurations; the head surface with minimum Procrustes distance from the mean, black; and the template surface, grey. Triangles = means, squares = allometric predictions of surfaces at the sample minimum centroid size, circles = allometric predictions of surfaces at the sample maximum centroid size.

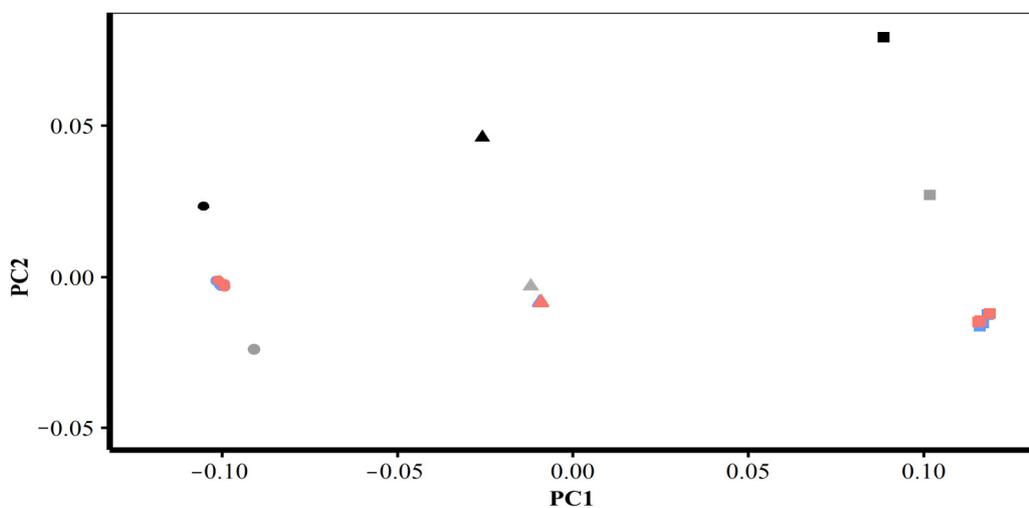


Figure 3.28 Scatter plots of the mean and allometrically scaled ape surfaces. PC1 (93.2% of total variance) vs PC2 (6.04%) from PCA of the mean and allometrically scaled ape cranial surfaces derived using

maximum density of semilandmarks and each semilandmarking approach (from Figure 3.15). Red = Sliding TPS, blue = TPS&NICP, also included in this PCA are surfaces warped to the mean and scaled landmark configurations; the ape cranial surface used to generate the template, black; and the template surface, grey. Triangles = means, squares = allometric predictions of surfaces at the sample minimum centroid size, circles = allometric predictions of surfaces at the sample maximum centroid size.

3.4 Discussion

The use of digital surface meshes of biological and anthropological specimens in 3D GM studies has become increasingly common, as has the use of landmarks and semilandmarks generated by different semilandmarking approaches to compare the details of morphology (Boyer et al., 2015; Dai et al., 2020; Gunz & Mitteroecker, 2013; Koehl & Hass, 2015; Pomidor et al., 2016; Rolfe et al., 2021). While dense coverage by semilandmarks allows more detailed description of form and, potentially, biological signal (Goswami et al., 2019), it introduces several difficulties in comparing forms. Further, given that semilandmarks are treated as equivalent between specimens in GM analyses and are given the same weight as landmarks, the basis of equivalence is an important consideration. In studies of biological transformations such as occur during development and evolution, the equivalences required to model and compare them are developmental or evolutionary. Landmarks and semilandmarks at each stage need mark-up points that are equivalent between specimens in terms of development or evolution at another stage (homologous points). For landmarks this matching is based on prior knowledge, but for semilandmarks it is algorithmic and relies on mathematical mappings and topographic features. As such the extent to which semilandmarks can be considered homologous has contributed to the debate about their validity and usefulness in relation to the study of developmental or evolutionary transformations (Cardini, 2020; Goswami et al., 2019; Oxnard & O'Higgins, 2009).

It has been noted by previous researchers that, because the locations of semilandmarks on surfaces and curves are uncertain, they should not be interpreted singly, but rather as a whole (Bastir et al., 2019; Mitteroecker & Schaefer, 2022; Oxnard & O'Higgins, 2009). While this avoids over interpreting differences in individual semilandmark locations, it does not avoid statistical issues. Thus, differences in semilandmark locations will lead to different distance matrices among specimens and so, to different analytical results. The extent of this issue has been explored in several previous studies (Boyer et al., 2011; Boyer et al., 2015; Gonzalez et al., 2016; Gunz et al., 2005; Perez et al., 2006; Pomidor et al., 2016; Rolfe et al., 2021). Additionally, the use of high-density semilandmarks raises statistical issues related to the ratio of variables to specimens (i.e. high p and low n) and in assessing covariances within landmark and semilandmark configurations (Cardini, 2020; Cardini et al., 2019).

Statistical considerations aside, high-density semilandmarks are routinely used to assess shape variations and covariations and to perform classification (Goswami et al.,

2019; Musilová et al., 2016; Schlager et al., 2018; Verhelst et al., 2021) with results presented as visualisations of a warped surface mesh. It is therefore of interest to know how different semilandmarking approaches, and densities of semilandmarks affect visualisations. This has been addressed by the analyses presented here.

In this study, we compare surface meshes warped to configurations of landmarks and semilandmarks arising from GM analyses that represent the overall mean and allometrically scaled surfaces. The aim is to compare the surface meshes used for visualisation rather than the statistical outcomes of analyses of the landmark and semilandmark configurations. These were compared in (Shui et al., 2023). Three different semilandmarking approaches were used with varying semilandmark densities. These are the method of sliding semilandmarks, minimising the bending energy of a set of thin-plate splines or Procrustes distances (Mitteroecker et al., 2013), the non-rigid combined approach of TPS&NICP (Shui et al., 2021), and the rigid LS&ICP approach. These lead to semilandmark configurations that differ in the locations of semilandmarks. These differences are least between sliding TPS and TPS&NICP approaches, and larger when comparing these with the LS&ICP approach (Figure 3.2). However, the locations of individual semilandmarks are not interpretable and, as noted above, they lie on the surface and so should be interpreted as a whole, in terms of differences between surfaces that fit them.

This study aimed to do this by empirically testing two hypotheses using surface scans of human heads and ape crania: that there are no differences in surface mesh shape between estimates, derived using different semilandmarking densities and approaches, applied to surfaces representing (a) the mean of a sample (b) allometric scaling of the mean. The surfaces were quantitatively compared using the coordinates of their vertices after re-semilandmarking and re-warping (see methods) to calculate Procrustes distances between them and, where relevant by extracting and comparing principal components. They were visually compared using colour maps of differences in local surface areas. Both hypotheses are falsified, differences clearly exist between estimated mean and allometrically scaled surfaces, but the degree of difference between semilandmarking approaches is small to moderate, with the non-rigid semilandmarking approaches (sliding TPS and TPS&NICP) showing a high degree of consistency.

Because landmarks have more secure homology, than semilandmarks and should be chosen with respect to the question at hand (Cardini, 2020; Oxnard & O'Higgins, 2009), they are likely few in number and less likely to result in statistical issues arising from large numbers of variables relative to the number of specimens. Additionally, surfaces can be warped to landmarks to visualise analytical results, albeit with less detail than warping based on dense correspondences. Thus, the present study also assessed

differences between warped surfaces based on landmarks and semilandmarks and those based on the landmark configuration alone, using different reference surfaces.

The three semilandmarking approaches were used to estimate sample mean surface meshes, by warping the template (an initial estimate of the average surface in each dataset) to the mean landmark and semilandmark coordinates arising from each method using varying semilandmark densities. For the head surfaces the means are visually quite similar (Figure 3.4) but differ in detail (Figures 3.6 and 3.7). The resulting mean surfaces from sliding TPS and TPS&NICP are most similar, and those from LS&ICP most different. Similar results are obtained in estimating the mean surface of the ape crania (Table 3.1 and Figures 3.8 and 3.9), but the LS&ICP approach performed poorly in locating semilandmarks in reasonably corresponding locations with the more complex ape cranial surfaces. In both datasets, estimated mean surfaces converge with increasing semilandmarking density on the surface from the highest density (Tables 3.4 and 3.5, and Figures 3.10-3.12). For the head surface data, warping the template surface to the mean landmark configuration (Figure 3.23b) resulted in a surface that was quite similar in general to that warped to landmarks and high-density semilandmarks, but which differed in detail from the semilandmark based mean (Figure 3.23d). This similarity is in large part due to the fact that the template surface is already an initial estimate of the mean. Repeating the analysis using the surface of the individual nearest to the mean landmarks and semilandmarks, resulted in an estimate of the mean surface (Figure 3.23c) that presented greater differences from the semilandmark based mean surface (Figure 3.23e). Visually this approach worked reasonably despite the lack of identifiable landmarks to guide warping of the scalp, however this is likely because the template scalp was not an initial estimate of, and very similar to the mean.

The mean ape surfaces estimated using sliding TPS and TPS&NICP with varying densities of semilandmarks are also visually quite similar (Figure 3.8), although the surface from LS&ICP shows some obvious differences. Focusing on sliding TPS and TPS&NICP, the mean surfaces resulting from these methods using varying numbers of semilandmarks are very similar, with differences increasing with semilandmarking density, especially where surface topography is complex (Figure 3.9 and Table 3.1). Surfaces estimated using increasing numbers of semilandmarks converge on the surface estimated using the maximum number of semilandmarks (Figures 3.10, 3.11, 3.13).

It should be noted that in the implementation of NICP used here, the initial registration of surfaces between template and target uses a triplet of TPS. This is also the case for the sliding TPS approach. This shared initial, non-rigid registration doubtless contributes to the similarities in results obtained using these approaches, when compared to the rigid, least squares registration employed in the LS&ICP approach. However, even the LS&ICP approach used the same landmark set for registration. It would be of interest

in future work to assess the impact of using different landmark configurations to estimate semilandmarks.

Using the mean landmark configuration alone to warp the template surface mesh results in a visually similar surface to the mean based on landmarks and high-density semilandmarks, but which differs in detail, especially around crests and ridges (Figures 3.24a, b and d). Visualisation of the mean by warping the ape surface used to generate the template results in a more different surface (Figures 3.24c, e and f), which in some ways resembles the mean based on landmarks and high-density semilandmarks (Figure 3.24a), but which differs particularly in regions with complex topography (Figures 3.24c and f). These landmark based warping differ in detail from the landmark and semilandmark based ones, but also bear a resemblance. Whether or not they are adequate depends on the purpose to which they are put. They may be sufficient to describe general aspects of shape variation, but would likely yield different results if used to build finite element models (FEM). Warping of a surface that is an initial estimate of the mean to the landmarks alone, inevitably yields a surface more like that based on landmarks and semilandmarks than warping a surface from an individual, even if close to the mean. This also applies to the comparisons of mean surfaces resulting from semilandmarking approaches and densities.

Predicted allometrically scaled mean surfaces were also compared among semilandmarking approaches and densities. With the head surface dataset, sliding TPS and TPS&NCP produced very similar surfaces particularly at the highest semilandmarking densities (Figures 3.14 and 3.15). The surfaces from LS&ICP were dissimilar. Likewise, for ape cranial surfaces, the allometrically scaled mean surfaces from sliding TPS and TPS&NCP are similar, but differ in detail, especially around ridges and crests (Table 3.6 and Figure 3.16). They become more dissimilar in the regions of crests and ridges as semilandmarking density increases, reflecting the more detailed controlled of warping by greater densities of semilandmarks. Both semilandmarking approaches show a similar pattern of convergence on the surface derived from the highest density, of surfaces with increasing densities of semilandmarking (Figures 3.21 and 3.22, and Table 3.8).

These differences among allometrically scaled means from both datasets and the different approaches and densities of semilandmarking are summarized by the PC plots of Figures 3.17 and 3.18. Figure 3.17 presents for the head surface data, the first two PCs from an analysis of mean and allometrically scaled mean surfaces derived from varying densities of semilandmarks and each approach. It shows that sliding TPS and TPS&NCP achieve very similar results, with many points overlapping, but that LS&ICP results in quite different estimates of the same surfaces that vary along a different vector from the other two approaches. The comparable analysis, for the ape crania compared only sliding TPS and TPS&NCP, and the resulting PC plot shows that these achieve very similar

results. These findings give perspective to the differences identified in the Procrustes distance matrices and visual comparisons in the analyses described above. Thus, the Procrustes distances between mean surfaces from varying semilandmarking approaches and densities are small compared to those between surfaces allometrically scaled to the maximum and minimum sample centroid sizes. The colour maps are very sensitive, identifying and emphasising what are in reality very small differences.

Allometrically scaled ape cranial surfaces from sliding TPS with 800 semilandmarks are compared with surfaces derived by warping to the allometrically scaled landmark configurations the template surface and the surface used to generate the template. The resulting predictions of surfaces at both the sample maximum and minimum centroid sizes share general similarities with, but differ in detail from the surfaces based on semilandmarks (Figures 3.25d, e and f and Figures 3.26d, e and f). As with the head surfaces, these differences reflect similar aspects of scaling, which may be adequate in describing general scaling trends but would likely lead to differences in FEA results among models based upon them.

The differences in scaling are emphasised by the PCAs of Figures 3.27 and 3.28, where for both datasets, the surfaces derived by warping the surface of the individual nearest to the mean to the allometrically scaled mean landmark configurations, are distant from the semilandmark based surfaces and are arranged along a vector that is not parallel to the vector between surfaces scaled using semilandmarks. Warping the template surface to the mean and allometrically scaled means in both datasets results in a vector parallel to that derived by warping the head surface of the individual nearest to the mean or the ape cranium used to generate the template, but with the mean close to the means from semilandmark based approaches. This indicates that these different surfaces scale in very similar ways. Thus, the choice of template surface determines where in the shape space the allometric vector is located while the landmarks and semilandmarks used to deform the surface determine how it is deformed. Semilandmarks result in the surface regions between landmarks being deformed differently to what is achieved through warping to the landmark configurations alone. This is not surprising and underlines how semilandmarks contribute to control of surface deformation.

The results of this study have shown that different semilandmarking approaches and densities achieve different visualisations of mean and allometrically scaled surfaces. The degree of difference depends on the approach, with non-rigid semilandmarking (sliding TPS and TPS&NICP) producing surfaces that are consistently more similar to each other than to those derived using the rigid LS&ICP approach. Additionally, the non-rigid approaches show consistency of surfaces produced using semilandmarks of varying densities. While Procrustes distances and colour maps emphasise differences among approaches, PCAs comparing scaled mean surfaces show that the differences between

surfaces from non-rigid semilandmarking approaches are very small when compared to the differences among allometrically scaled means. The differences between surfaces derived using LS&ICP are greater.

Semilandmarking involves a great deal of extra effort compared to landmarking alone, and as has been noted earlier, brings with it some severe statistical issues. This has led to questioning of their benefits and criticism that they may lead to erroneous conclusions (Cardini, 2020; Cardini et al., 2019). Thus, this study compared surfaces warped using landmark configurations alone with those from landmark and semilandmarking configurations. These comparisons have shown that if a surface that is an initial estimate of the mean surface is used, then mean surfaces are well estimated. This is to be expected, since the mean landmarks have little warping to do. This finding likely explains why LS&ICP results in more similar mean surfaces to those from sliding TPS and TPS&NICP at lower rather than higher semilandmarking densities (Figures 3.6 and 3.7). When an alternative surface is used, the surface visualisation is different, having inherited features of this new surface. Surfaces warped to scaled landmark configurations show differences and some similarities to those warped to landmarks and semilandmarks in combination. Such analyses and visualisations based on landmarks alone may be perfectly adequate for many questions, they involve less work to produce and avoid the statistical issues that can arise with many semilandmarks and few specimens. However, compared to surfaces from semilandmarks, they would likely lead to different results if used to build functional or biomechanical models.

Finally, we should emphasise that consistency is not the same as accuracy (Cardini, 2020). It is tempting to conclude that the remarkable consistency of surface shapes derived using sliding TPS and TPS&NICP reflects accuracy of estimation of means. Our results cannot, however, support or refute this possibility since no ‘true mean’ is known (or knowable). Estimates of means depend on what quantities are measured and compared because means are a statistical, rather than biological entity, particular to the data used to calculate the mean. The results are ‘correct’ for the variables (semilandmark locations) resulting from each method. However, with semilandmarks there is inevitable uncertainty about the extent to which they are equivalent between specimens in terms of homology. Our studies have shown that differences in semilandmark locations among specimens will lead to differences in statistical results (Shui et al., 2023) and visualisations (present study). In these studies, these differences are quite small relative to the differences among specimens, but it is not clear to what extent these empirical results apply to diverse datasets and semilandmarking approaches (e.g. minimisation of Procrustes distances by sliding (Gunz & Mitteroecker, 2013)); morphometric ‘fishnets’ (Polly, 2008). This can only be addressed by further extensive studies of real data and by simulation experiments, in which an initial ‘mean’ is perturbed and then estimated from the perturbed data.

For now, we have shown that the two non-rigid semilandmarking approaches yield consistent estimates of mean and scaled surfaces. Semilandmarking involves a great deal of additional work and runs statistical risks in analyses. With these things in mind, the investigator should carefully consider if semilandmarking is necessary to answer the question at hand and balance this need against the statistical and biological (e.g. regarding homology) downsides and the time involved in gathering and using semilandmarks to assess shape variance and covariances. It may be a more secure strategy to base statistical tests on homologous landmarks and visualisations on landmarks and semilandmarks from parallel analyses.

It should be borne in mind that homology is often also uncertain for landmarks, and that different sets of landmarks will lead to different results. However, three approaches that we compare in this study led to visually similar estimates of surface meshes that may be adequate for visualisation and functional simulation, in the sense that they are likely to be fair representations of average and scaled surfaces, but there is no single ‘true’ representation against which to assess this (see above). Their applicability depends on how much error in the estimation of surface shapes is judged acceptable, given the context of the particular study.

Finally, it should be noted that this study is limited in its scope, being based on only human heads and ape crania, different datasets need to be examined to assess the reliability of the findings. Studies also need to be conducted using simulated data in which true mean and allometrically scaled surfaces are known, in order to assess the accuracy of the estimates of these surfaces. Additionally, this study compared a limited range of possible approaches to semilandmarking and future work needs to extend these comparisons to include other methods and ‘landmark free’ approaches.

3.5 Conclusion

This study examined the effects of different semilandmarking approaches and densities of semilandmarking on estimates of mean and allometrically scaled mean surfaces. These are investigated by assessing overall and regional shape differences based on Procrustes distances and colour maps of local surface mesh area differences. The results show that the mean and fitted surfaces generated by the sliding TPS and TPS&NICP approaches are very similar, while the LS&ICP approach yields surfaces that differ most. Surfaces warped to landmark configurations differ from these, depending on the degree of similarity of the surface to the mean, and show a different vector of allometric scaling, reflecting the differences between TPS interpolation and semilandmark control of surfaces between landmarks. In conclusion, visualizations derived using, especially semilandmarks from non-rigid semilandmarking approaches are likely to fairly represent surfaces and differences between them, but are not identical. The

extent to which these differences are important depends on the particular study context and aims.

3.6 Acknowledgments

I am grateful to Nick Pears and Christian Duncan for kindly providing access to the headspace dataset (<https://www-users.cs.york.ac.uk/~nep/research/Headspace/>, accessed on 16 January 2023). Thank Olivia Smith for providing landmarks for every head and to the Smithsonian Institution (www.vertebrates.si.edu) for kindly providing access to the ape crania.

3.7 References

- Adams, D, Rohlf, J, & Slice, D. (2004). Geometric morphometrics: ten years of progress following the ‘revolution’. *Italian Journal of Zoology*, 71(1), 5-16.
- Amberg, B, Romdhani, S, & Vetter, T. (2007). Optimal step nonrigid ICP algorithms for surface registration. *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR'07)*, 1-8.
- Arun, KS, Huang, TS, & Blostein, SD. (1987). Least-squares fitting of two 3-D point sets. *IEEE Transactions on Pattern Analysis Machine Intelligence*(5), 698-700.
- Bardua, C, Felice, R, Watanabe, A, Fabre, A, & Goswami, A. (2019). A practical guide to sliding and surface semilandmarks in morphometric analyses. *Integrative Organismal Biology*, 1(1), obz016.
- Bastir, M, García-Martínez, D, Torres-Tamayo, N, Palancar, CA, Fernández-Pérez, FJ, Riesco-López, A, Osborne-Márquez, P, et al. (2019). Workflows in a Virtual Morphology Lab: 3D scanning, measuring, and printing. *Journal of Anthropological Sciences*, 97, 107-134.
- Besl, PJ, & McKay, ND. (1992). A method for registration of 3-D shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2), 239-256.
- Blanz, V, & Vetter, T. (1999). A morphable model for the synthesis of 3D faces. *The Proceedings of the 26th annual conference on Computer graphics and interactive techniques*, 187-194.
- Bookstein, FL. (1989). Principal warps: Thin-plate splines and the decomposition of deformations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(6), 567-585.
- Bookstein, FL. (1997). Landmark methods for forms without landmarks: morphometrics of group differences in outline shape. *Medical Image Analysis*, 1(3), 225-243.
- Booth, J, Roussos, A, Ponniah, A, Dunaway, D, & Zafeiriou, S. (2018). Large scale 3D morphable models. *International Journal of Computer Vision*, 126(2-4), 233-254.
- Boyer, DM, Lipman, Y, Clair, ES, Puente, J, Patel, BA, Funkhouser, T, Jernvall, J, et al. (2011). Algorithms to automatically quantify the geometric similarity of anatomical surfaces. *Proceedings of the National Academy of Sciences*, 108(45), 18221-18226.
- Boyer, DM, Puente, J, Gladman, JT, Glynn, C, Mukherjee, S, Yapuncich, GS, & Daubechies, I. (2015). A new fully automated approach for aligning and comparing shapes. *The Anatomical Record*, 298(1), 249-276.
- Cardini, A. (2020). Less tautology, more biology? A comment on “high-density” morphometrics. *Zoomorphology*, 139(4), 513-529.

- Cardini, A, O'Higgins, P, & Rohlf, FJ. (2019). Seeing distinct groups where there are none: spurious patterns from between-group PCA. *Evolutionary Biology*, 46(4), 303-316.
- Dai, H, Pears, N, Smith, W, & Duncan, C. (2020). Statistical Modeling of Craniofacial Shape and Texture. *International Journal of Computer Vision*, 128(2), 547-571.
- Duncan, C, Pears, NE, Dai, H, Smith, WAP, & O'Higgins, P. (2022). Applications of 3D Photography in Craniofacial Surgery. *Journal of Paediatric Neurosciences*, 17(S 1), 21-28.
- Gonzalez, PN, Barbeito-Andrés, J, D'Addona, LA, Bernal, V, & Perez, SI. (2016). Performance of semi and fully automated approaches for registration of 3D surface coordinates in geometric morphometric studies. *American Journal of Physical Anthropology*, 160(1), 169-178.
- Goswami, A, Watanabe, A, Felice, RN, Bardua, C, Fabre, AC, & Polly, PD. (2019). High-density morphometric analysis of shape and integration: the good, the bad, and the not-really-a-problem. *Integrative Comparative Biology*, 59(3), 669-683.
- Gunz, P, & Mitteroecker, P. (2013). Semilandmarks: a method for quantifying curves and surfaces. *Hystrix, the Italian Journal of Mammalogy*, 24(1), 103-109.
- Gunz, P, Mitteroecker, P, & Bookstein, FL. (2005). Semilandmarks in three dimensions. In *Modern morphometrics in physical anthropology*, 73-98: Springer.
- Gunz, P, Mitteroecker, P, Neubauer, S, Weber, GW, & Bookstein, FL. (2009). Principles for the virtual reconstruction of hominin crania. *Journal of Human Evolution*, 57(1), 48-62.
- Koehl, P, & Hass, J. (2015). Landmark-free geometric methods in biological shape analysis. *Journal of The Royal Society Interface*, 12(113), 20150795.
- Mitteroecker, P, & Gunz, P. (2009). Advances in geometric morphometrics. *Evolutionary Biology*, 36(2), 235-247.
- Mitteroecker, P, Gunz, P, Windhager, S, & Schaefer, K. (2013). A brief review of shape, form, and allometry in geometric morphometrics, with applications to human facial morphology. *Hystrix, the Italian Journal of Mammalogy*, 24(1), 59-66.
- Mitteroecker, P, & Schaefer, K. (2022). Thirty years of geometric morphometrics: Achievements, challenges, and the ongoing quest for biological meaningfulness. *American Journal of Biological Anthropology*, 178(S 74), 181-210.
- Musilová, B, Dupej, J, Velemínská, J, Chaumoitre, K, & Bruzek, J. (2016). Exocranial surfaces for sex assessment of the human cranium. *Forensic Science International*, 269, 70-77.
- O'Higgins, P, & Jones, N. (1998). Facial growth in *Cercocebus torquatus*: an application

of three-dimensional geometric morphometric techniques to the study of morphological variation. *Journal of Anatomy*, 193(2), 251-272.

O'Higgins, P, Cobb, SN, Fitton, LC, Gröning, F, Phillips, R, Liu, J, & Fagan, MJ. (2011). Combining geometric morphometrics and functional simulation: an emerging toolkit for virtual functional analyses. *Journal of Anatomy*, 218(1), 3-15.

Oxnard, C, & O'Higgins, P. (2009). Biology clearly needs morphometrics. Does morphometrics need biology? *Biological Theory*, 4(1), 84-97.

Perez, SI, Bernal, V, & Gonzalez, PN. (2006). Differences between sliding semi-landmark methods in geometric morphometrics, with an application to human craniofacial and dental variation. *Journal of Anatomy*, 208(6), 769-784.

Polly, PD. (2008). Adaptive zones and the pinniped ankle: a three-dimensional quantitative analysis of carnivoran tarsal evolution. In *Mammalian evolutionary morphology*, 167-196: Springer.

Pomidor, BJ, Makedonska, J, & Slice, DE. (2016). A landmark-free method for three-dimensional shape analysis. *PloS one*, 11(3), e0150368.

Profico, A, Buzi, C, Castiglione, S, Melchionna, M, Piras, P, Veneziano, A, & Raia, P. (2021). Arothron: An R package for geometric morphometric methods and virtual anthropology applications. *American Journal of Physical Anthropology*, 176(1), 144-151.

Profico, A, Piras, P, Buzi, C, Di Vincenzo, F, Lattarini, F, Melchionna, M, Veneziano, A, et al. (2017). The evolution of cranial base and face in Cercopithecoidea and Hominoidea: Modularity and morphological integration. *American Journal of Primatology*, 79, e22721.

Rohlf, FJ. (1998). On applications of geometric morphometrics to studies of ontogeny and phylogeny. *Systematic Biology*, 47(1), 147-158.

Rolfe, S, Davis, C, & Maga, AM. (2021). Comparing semi-landmarking approaches for analyzing three-dimensional cranial morphology. *American Journal of Physical Anthropology*, 175(1), 227-237.

Schlager, S. (2017). Morpho and Rvcg–Shape Analysis in R: R-Packages for geometric morphometrics, shape analysis and surface manipulations. In *Statistical shape and deformation analysis*, 217-256.

Schlager, S, Profico, A, Di Vincenzo, F, & Manzi, G. (2018). Retrodeformation of fossil specimens based on 3D bilateral semi-landmarks: Implementation in the R package “Morpho”. *PloS one*, 13(3), e0194073.

Shui, W, Profico, A, & O'Higgins, P. (2023). A comparison of semilandmarking approaches in the analysis of size and shape. *Animals*, 13(7), 1179.

Shui, W, Zhang, Y, Wu, X, & Zhou, M. (2021). A computerized facial approximation

method for archaic humans based on dense facial soft tissue thickness depths. *Archaeological Anthropological Sciences*, 13(11), 186.

Shui, W, Zhou, M, Deng, Q, Wu, Z, Ji, Y, Li, K, He, T, et al. (2016). Densely Calculated Facial Soft Tissue Thickness for Craniofacial Reconstruction in Chinese Adults. *Forensic Science International*, 266, 573.e571-e512.

Shui, W, Zhou, M, Maddock, S, Ji, Y, Deng, Q, Li, K, Fan, Y, et al. (2020). A computerized craniofacial reconstruction method for an unidentified skull based on statistical shape models. *Multimedia Tools and Applications*, 79, 25589-25611.

Smith, OA, Nashed, YS, Duncan, C, Pears, N, Profico, A, & O'Higgins, P. (2021). 3D Modeling of craniofacial ontogeny and sexual dimorphism in children. *The Anatomical Record*, 304, 1918-1926.

Van Kaick, O, Zhang, H, Hamarneh, G, & Cohen-Or, D. (2011). A survey on shape correspondence. *Computer Graphics Forum*, 30(6), 1681-1707.

Verhelst, P-J, Matthews, H, Verstraete, L, Van der Cruyssen, F, Mulier, D, Croonenborghs, T, Da Costa, O, et al. (2021). Automatic 3D dense phenotyping provides reliable and accurate shape quantification of the human mandible. *Scientific Reports*, 11(1), 1-10.

Viscosi, V, & Cardini, A. (2011). Leaf morphology, taxonomy and geometric morphometrics: a simplified protocol for beginners. *PloS one*, 6(10), e25630.

Supplementary material

Table 3.S1 Ape cranium specimen

Specimen number	Species	Sex	Repository
USNM 176211	Gorilla	Male	Smithsonian Institution
USNM 252578	Gorilla	Male	Smithsonian Institution
USNM 590949	Gorilla	Female	Smithsonian Institution
USNM 590952	Gorilla	Female	Smithsonian Institution
USNM 599170	Gorilla	Female	Smithsonian Institution
USNM 083262	Hylobates lar	Female	Smithsonian Institution
USNM 083263	Hylobates lar	Female	Smithsonian Institution
USNM 083264	Hylobates lar	Female	Smithsonian Institution
USNM 111970	Hylobates lar	Male	Smithsonian Institution
USNM 111988	Hylobates lar	Male	Smithsonian Institution
USNM 174701	Pan troglodytes	Female	Smithsonian Institution
USNM 174704	Pan troglodytes	Male	Smithsonian Institution
USNM 220062	Pan troglodytes	Female	Smithsonian Institution
USNM 220327	Pan troglodytes	Male	Smithsonian Institution
USNM 395820	Pan troglodytes	Male	Smithsonian Institution
USNM 143590	Pongo abelii	Male	Smithsonian Institution
USNM 267325	Pongo abelii	Male	Smithsonian Institution
USNM 270807	Pongo abelii	Female	Smithsonian Institution
USNM 283737	Pongo abelii	Female	Smithsonian Institution
USNM 293165	Pongo abelii	Male	Smithsonian Institution

Chapter 4 A Landmark-free Approach for Surface Asymmetry Detection and Profile Drawings from Bilaterally Symmetrical Geometry

Abstract: Bilaterally symmetrical objects represent a large and important proportion of archaeological artefacts and biological objects. The identification of the plane of symmetry plays a vital role in quantifying surface asymmetry and producing profile drawings in archaeology and anthropology. The correct recognition of symmetry provides evidence to allow experts to restore damaged artefacts, assess consistency in artefact manufacture, and examine morphological variability in human development. With the increasing availability of archaeological and anthropological 3D meshes, landmark-based and landmark-free morphometric methods for detecting planes of symmetry have both been proposed. However, the landmark-based approach requires manual identification of landmark locations, hence time consuming and prone to error. Additionally, the landmark independent morphometric method is influenced by missing data. This study presents an effective landmark-free approach to approximate the plane of symmetry from nearly bilaterally symmetrical objects by means of finding the best-fitted plane with the minimum geometric differences between the original and mirrored meshes. Subsequently, a global and regional method is carried out to quantify surface asymmetry, reducing the effect of the size and orientation of 3D meshes on gross asymmetry detection. Finally, profile drawings are produced by computing the intersections of the plane of symmetry and 3D meshes. Both synthetic and real objects are used to evaluate the effectiveness and robustness of the proposed method. Our results show the approximated plane of symmetry generated by the proposed method is consistent with that determined by anatomical landmarks, and no significant differences in asymmetry ratio (AR) representing the degree of gross asymmetry are found between the landmark-based and proposed methods. These results demonstrate that the proposed method provides a suitable plane of symmetry from a bilaterally symmetrical object with small geometric distortion or simple missing geometry, thereby speeding up asymmetry detection and profile drawings.

Article details:

Wuyang Shui, Pianpian Wei, Xia Zheng, Shengling Geng. A landmark-free approach for surface asymmetry detection and profile drawings from bilaterally symmetrical geometry. *ACM Journal on Computing and Cultural Heritage*. 2022. (accepted)

At the time of submission, publication of this chapter was pending. It was published by *ACM Journal on Computing and Cultural Heritage* on June 24, 2023. See the link: <https://dl.acm.org/doi/pdf/10.1145/3589247>.

4.1 Introduction

Bilateral symmetry is one of the most important characteristics of archaeological artefacts and human fossil specimens. In recent decades, asymmetry detection has wide applications in archaeology, anthropology, palaeontology, biology and craniofacial surgery (Klingenberg, 2015; Mitra et al., 2013). These analytical results provide considerable guidance and reliable evidence to allow experts to examine historical manufacturing capacities (McNabb et al., 2018; Saragusti et al., 1998), repair damaged artefacts (Shui & Gao, 2021), evaluate human developmental variability and instability (Hou & Fagan, 2021), and perceive the evolution of human cognition (Wynn, 2002). For example, an examination of asymmetry in pottery can enable researchers to assess whether potters employed a fast wheel, thereby indicating the level and scale of manufacturing in past societies. In addition, the recognition of fluctuating asymmetry of biological specimens contributes to examining the effects of genetic stress and environmental changes on human evolution.

The issue of quantitative analysis of asymmetry from archaeological and anthropological objects has created considerable interest. Much effort has gone into landmark-based approaches, because landmarks describe the shape and size of specimens in terms of anatomical, developmental, biomechanical or evolutionary knowledge (O'Higgins, 2000). A landmark-based geometric morphometrics (GM) method is the preferred way of quantifying asymmetry from a large collection of samples (McGrath et al., 2022; Neubauer et al., 2020; Profico et al., 2021). Generalized Procrustes Analysis (GPA) is applied to symmetric and asymmetric components comprising landmark and semilandmark configurations separately, removing the effect of location, orientation and scaling factors. Subsequently, Principal Component Analysis (PCA) is used to project every sample into shape space. Quantitative analysis, e.g. multivariate regression, and a colour map of geometric differences in fitted shapes along the extreme limits of the principal components of interest can be used to explore the patterns of asymmetry.

Current methods to quantify surface asymmetry of each individual involve the comparison of geometric differences between the original and registered mirrored landmark configurations or surfaces with regard to the plane of symmetry which passes through the mid-points of the landmark configuration, e.g. mid-sagittal plane (MSP) (Cassi et al., 2019; Claes et al., 2011; Damstra et al., 2012; Gibelli et al., 2018; Green et al., 2017; Verhoeven et al., 2016; Willing et al., 2013). Notably, the identification of the plane of symmetry plays a vital role in quantifying asymmetry. However, the manual placement of landmarks consumes the amount of time and the locations are primarily dependent upon the experts' anatomical knowledge and visual perception. Consequently, placement errors are inevitable, especially in featureless and smooth objects, thereby leading to inaccurate identifications of the plane of symmetry. To tackle the issue, the

landmark independent method needs to be further explored to improve the effectiveness of surface asymmetry detection.

In archaeological reports, a conventional manual 2D line drawing is commonly used as a standard method to record geometric properties and artefact cross-sections. However, this procedure is a time-consuming task and profile drawings are primarily dependent upon experts' skills (Karasik & Smilansky, 2008). Recent studies have proposed computer-assisted approaches to extract line drawings and to capture the dimensional features in archaeological potteries (Angelo et al., 2020; Wilczek et al., 2018). These procedures are very fast and the results are accurate and reproducible. Notably, the most challenging task in these applications is to find the axis of rotation from 3D meshes to accomplish the orientation process automatically. Likewise, the plane of symmetry is the fundament for orientating 3D meshes and sketching profile drawings of bilaterally symmetrical objects.

Archaeological and anthropological objects frequently exhibit asymmetry and contain missing geometry. In this study, we propose a landmark-free approach for effectively recognizing the best-fitted plane of symmetry from the symmetrical object. Furthermore, we quantify the extent of surface asymmetry and produce profile drawings regarding the approximated plane of symmetry. The main contributions are as follows:

- We identify the symmetrical regions from 3D meshes through the examination of the extent of asymmetry and plane clustering.
- We improve the morphometric method to approximate the plane of symmetry by finding the best-fitted plane with the minimum geometric differences between the original and mirrored 3D meshes.
- We present a generalized method to quantify gross and regional surface asymmetry with regard to the plane of symmetry.

4.2 Related work

Our literature review focuses on landmark-free approaches for approximating the plane of symmetry from 3D meshes. It comprises four categories: the landmark-free morphometric method, the extrinsic-based method, the intrinsic-based method, and deep learning method.

Extending the framework of the landmark-based morphometric approach which identifies the plane by fitting the mid-points of the original and registered mirrored landmark configurations (Damstra et al., 2012), the landmark-free morphometric method has been proposed based on a full set of vertices of 3D meshes (Di Angelo et al., 2019; Noori et al., 2020). It comprises three steps: first, the mirrored meshes were generated based on an arbitrary plane or one of the principal directions passing through the centroid

point of 3D meshes (Zhang et al., 2006). Second, the iterative closest point (ICP) algorithm (Besl & McKay, 1992) was conducted to register the original and mirrored meshes by minimizing the sum of the squared Euclidean distance between correspondences established by searching the nearest points. Finally, the plane of symmetry was fitted based on the mid-points of the original and registered mirrored meshes. It is worth noting that the ICP algorithm is likely to suffer from a local minima problem especially when the original and mirrored meshes have a great difference in initial location and orientation. To tackle the issue, an improved ICP algorithm has been proposed to speed up the convergence through the assignment of weight to paired correspondences, the rejection of outliers, and the optimization of the point-to-plane distance metric (Padia & Pears, 2011).

In computer vision, an effective approach to identifying the plane of symmetry is dependent upon extrinsic features that refer to invariance under rigid transformations. In Euclidean space, local extrinsic shape signatures, e.g. principal curvature (Mitra et al., 2006), orientation histogram (Sun & Sherrah, 1997) and spherical harmonics (Kakarala et al., 2013) have been used to find symmetrical correspondences of 3D meshes. Subsequently, the plane of symmetry was recognized based on all these reliable matches. Tevs et al. (2014) segmented the whole shapes into overlapping regions and then performed graph matching to establish the semantically meaningful symmetrical parts of man-made objects in terms of symmetric transformations. However, they might lead to failure when objects exhibit non-rigid transformation and distortion (Ovsjanikov et al., 2008).

Intrinsic spectral descriptors rely on the analysis of the eigenvalues and eigenfunctions of the Laplace-Beltrami (LB) operator that is a symmetrical negative semidefinite on a Riemannian manifold (Biasotti et al., 2016; Li & Hamza, 2014). Unlike extrinsic features, they preserve geodesic distances and remain invariant under an isometric mapping on the manifold (Xu et al., 2009). For example, Sun et al. (2009) presented a heat kernel signature (HKS) to represent the heat transfer from one point to another in time parameter t due to the diffusion process. Recent studies have shown that intrinsic features support measuring the similarity between non-rigid shapes (Zhang et al., 2021) and detecting the plane of symmetry. Sipran et al. (2014) employed a spectral descriptor to approximate the plane of symmetry from partial geometry. In their work, an intrinsic feature that balance the global and local characterization was used to find a set of distinctive points with the local maxima. Subsequently, the vote-based algorithm was conducted to obtain positive votes to improve the accuracy of identification. However, this method may be less effective when applied to featureless geometry.

With the progress towards the accessibility of a large collection of 3D meshes, recent studies have employed deep neural networks to detect the plane of symmetry. Ji and Liu

(2019) employed PointNet++ architecture to classify the point clouds to capture the reflection symmetry property and then used the random sample consensus (RANSAC) and least-squares algorithms to find the initial plane of symmetry. Based on this initial identification, the landmark-free morphometric method was employed to improve the accuracy of the symmetry. However, the annotation of training data takes the amount of time. To address this problem, Gao et al. (2020) proposed an unsupervised 3D convolutional neural network to extract global model features and then identify global planar reflective symmetry automatically.

4.3 Materials and methods

4.3.1 Materials

A variety of 3D acquisition technologies, e.g. medical imaging techniques (Zhang et al., 2012), 3D scanning (Kuzminsky & Gardiner, 2012) and photogrammetry (Gruen, 2012) have been used to digitize archaeological and anthropological objects. In this study, we used different technologies to capture 3D meshes. The 3D dataset comprises three parts and more details are provided in the supplementary material (Table 4.S1):

(a) In our previous study (Shui et al., 2016), a clinical multi-slice CT scanner system (Siemens Sensation 16) belonging to the Affiliated Hospital of Shaanxi University of Chinese Medicine was used to generate 3D meshes of 100 modern human male skulls, most of whom were residents of Shaanxi province, China. This work has been approved by the Institutional Review Board (IRB) of the Affiliated Hospital of the Shaanxi University of Chinese Medicine. The CT images of every skull were archived in standard DICOM 3.0 with a resolution of 512×512 . Subsequently, digital models were reconstructed from the segmented images using the threshold algorithm (Shui et al., 2017) to represent hard tissues and then the external surface of the skull was extracted by computing the intersection points between every cross-section and 3D meshes (Shui et al., 2020). This approach reduces the computational complexity and minimize the effect of internal structures and cervical spine on the shape analysis. Finally, the irrelative discrete meshes were manually removed and the reconstructed surfaces were processed smoothly. Each skull usually comprises of more than 160,000 vertices and 300,000 triangles. In this study, we selected thirty complete skulls aged 20-30 years old to test the effectiveness of the proposed method. Additionally, the CT scanner was used to acquire 3D meshes of a damaged modern human skull of a patient. Asymmetry analysis of this damaged skull can aid in formulating the surgery plan. Finally, a complete femur found at Jiangzhuang site in Jiangsu, north of the Yangtze, China was scanned using a clinical multi-slice CT scanner system to assess the accuracy of the proposed method to recognize the plan of symmetry from asymmetrical objects.

(b) We used two portable 3D scanners to obtain digital models of archaeological

artefacts. First, a Handyscan 700 scanner with a resolution of 0.03 mm was used to capture 3D meshes of a terracotta warrior head and a bronze mask. Second, an Artec spider scanner with a resolution of 0.05 mm was applied to capture 3D meshes of a pottery spoon, a ceramic plate, a clay pot, a tripod, a helmet and a Tibetan Buddha statue. The digital models were processed smoothly to improve the quality using the open-source software MeshLab (Cignoni et al., 2008).

(c) Two damaged crania have been digitized from casts by photogrammetry technique and then uploaded into the online digital repository on Sketchfab by the Research Labs of Archaeology, University of Chapel Hill, North Carolina. We used them to test the effectiveness in partial geometry. The first cranium (2501.1rp35) was known as Rhodesian Man, dating back to 125,000 years B.P (<https://sketchfab.com/3d-models/homo-sapiens-25011rp35-cranium-96ce6fa86a01479d907f7032e3e8f9b1>). The second model (2501.1rp20-1) was a cranium fragment of *Paranthropus boisei*, dating back to 1.8 million years ago (<https://sketchfab.com/3d-models/paranthropus-boisei-25011rp20-1-cranium-cd285c22e12c4cbaaf1684f80aed5d2e>).

4.3.2 Methods

This study presented a landmark-free computerized approach to extract symmetrical regions and approximate the best-fitted plane of symmetry, thereby quantifying surface asymmetry and producing profile drawings. Figure 4.1 summarizes the pipeline of the proposed method. It comprises three stages: First, the plane of symmetry was approximated through four steps: (a) the intrinsic feature of every vertex was calculated and visualized as a colour map; (b) 3D segmentation algorithm was used to separate 3D meshes into different regions through a comparison of the similarity of the neighbouring vertices; (c) the symmetrical regions (blue points) were extracted from the segmented results, each of which can be used to identify the plane of symmetry. (d) the best-fitted plane of symmetry was approximated (red colour) and then the accuracy was assessed by synthetic and real objects. Second, we registered the original (grey colour) and mirrored meshes (peach colour) and then quantified the extent of surface asymmetry by means of computing the geometric deviation between these two meshes. Finally, we orientated the 3D meshes and then sketched the profile (red colour) by calculating the intersections of the plane of symmetry and 3D meshes.

We used Matlab 2019 to calculate intrinsic features of 3D meshes (Li & Hamza, 2014) and employed C++ to recognize the plane of symmetry, quantify surface asymmetry and produce profile drawings. The software prototype can be downloaded from the Internet (https://github.com/sissun/Asymmetry_detection.git).

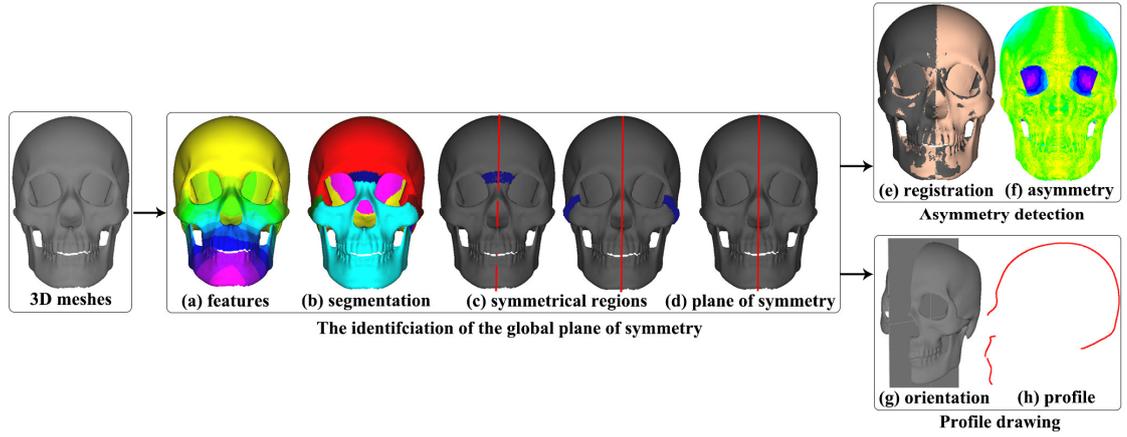


Figure 4.1 The workflow of asymmetry detection and profile drawings based on the identification of the plane of symmetry. (a) Visualisation of intrinsic feature of every vertex. (b) The segmented results. (c) The extraction of symmetrical region (blue points) accompanying the plane of symmetry (red colour). (d) The approximated best-fitted plane of symmetry (red colour) of 3D meshes. (e) The superimposition of original (grey colour) and registered mirrored meshes (peach colour). (f) Visualisation of asymmetry detection. (g) The orientated 3D meshes associated with the plane of symmetry. (h) Profile drawings.

4.3.2.1 An approximation of the plane of symmetry

An intrinsic wave kernel signature (WKS). WKS describes the average probability distribution of quantum mechanical particles with different energy densities (Aubry et al., 2011). It is invariant under isometric transformation (translations, reflections and rotations) and robust to small non-isometric deformations. Unlike HKS, WKS uses a band-pass filter to capture the geometric information from different frequencies. The evolution of a quantum particle can be governed at a location x at time t by wave function $\varphi(x, t)$ that is a solution of Schrödinger's equation

$$\frac{\partial \varphi}{\partial t}(x, t) = i\Delta \varphi(x, t) \quad (4.1)$$

where Δ denotes the Laplace-Beltrami operator, and i indicates the imaginary number.

The wave kernel function can be represented as follows:

$$\varphi(x, t) = \sum_{k=0}^{\infty} e^{i\lambda_k t} \phi_k(x) f_E(\lambda_k) \quad (4.2)$$

where $f_E(\lambda_k)$ represents the initial energy probability distribution and E denotes the energy at time $t = 0$. λ_k and $\phi_k(x)$ represent the k index eigenvalue and corresponding eigenvector.

WKS at a location x can be represented as follows:

$$\text{WKS}(E, x) = \sum_{k=0}^{\infty} \phi_k(x)^2 f_E(\lambda_k)^2 \quad (4.3)$$

To further support the calculation of WKS at a point of the manifold, WKS can be represented in the logarithmic energy scale $e_k = \log E_k$

$$\text{WKS}(e_k, x) = C_e \sum_k \phi_k(x)^2 e^{-\frac{(e_k - \log \lambda_k)^2}{2\delta^2}} \quad (4.4)$$

where $C_e = (\sum_k e^{-\frac{(e_k - \log \lambda_k)^2}{2\delta^2}})^{-1}$ and δ represents the variance, respectively.

The WKS values at the first k frequencies can be denoted by

$$\text{WKS} = [\text{WKS}_{i,1} \quad \text{WKS}_{i,2} \quad \dots \quad \text{WKS}_{i,k}] \quad (4.5)$$

where $\text{WKS}_{i,k}$ denotes the WKS value at the k index energy frequency of the i index vertex.

The WKS values at lower frequencies indicate the global topological information, and the values at the highest frequencies indicate the local topological information. In this study, we discriminated symmetrical regions of 3D meshes based on the intrinsic WKS features at the first k lowest frequencies (e.g. $k=100$). Figure 4.2 shows the WKS values at four different frequencies, respectively, including $k=1$, $k=10$, $k=50$, and $k=100$ (from left to right). Red colour represents the higher value and blue colour indicates the smaller value. Figure 4.2a displays the WKS values of a plane that exhibits perfect symmetry. Figure 4.2b illustrates the WKS values of a skull that exhibits nearly bilateral symmetry. The results show that symmetrical regions can be observed based on the WKS values.

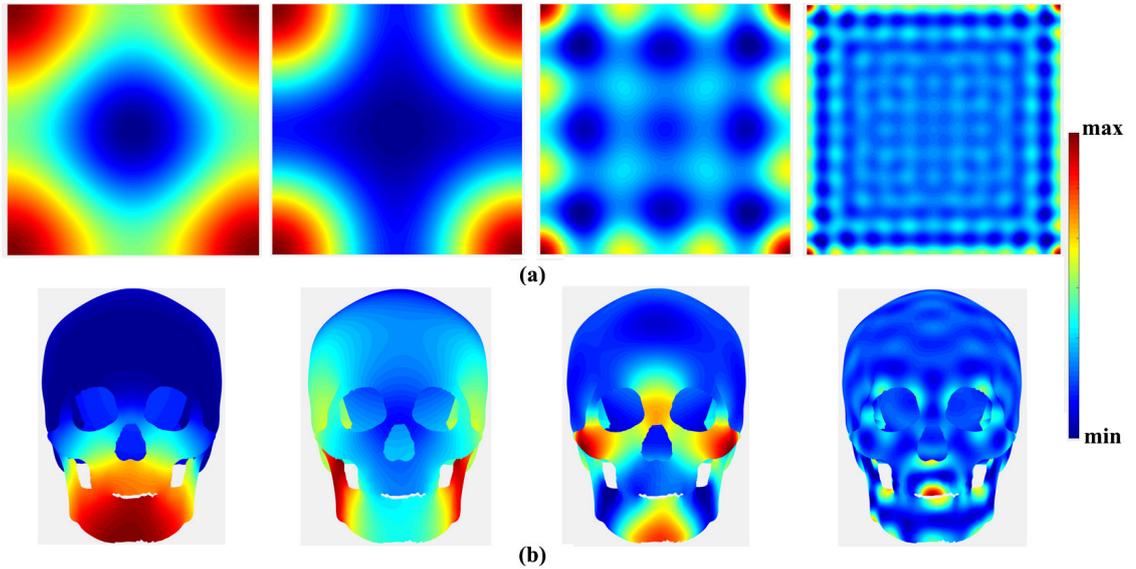


Figure 4.2 The WKS values of two models at different four energy frequencies, including $k=1$, $k=10$, $k=50$ and $k=100$ (from left to right). (a) A plane. (b) A skull.

Mesh segmentation. The region growing and merging algorithm (Wan & Higgins, 2003) is a promising means of extracting a region of interest. The choice of seed points and the growing and merging criterion affect the segmented results, e.g. inappropriate seeds might lead to under-segmentation or over-segmentation. To address this problem, this study used multiple sampling points as the seeds and then measured the cosine

similarity of the WKS values of the neighbouring vertices to segment 3D meshes into different regions. It comprised three steps:

Step 1: We used the Poisson-disk sample algorithm (Corsini et al., 2012) generate a set of points as seeds which covered the entire meshes evenly.

Step 2: In the growing process, the regions were iteratively grown through a comparison of the cosine similarity of WKS values between the current region and the unlabelled one-ring neighbouring vertices. The vertex that has a greater cosine similarity value than a fixed threshold ε_g would be assigned to the respective region. This process was performed until all the vertices were reached.

Step 3: The merging process was performed when it satisfied the condition that the cosine similarity of two different adjacent growing regions was larger than a fixed threshold ε_m , i.e. only m merging regions $S = \{S_1, S_2, \dots, S_m\}$ were extracted from n growing regions ($m \leq n$). It was an iterated process until all the growing regions were labelled.

The extraction of symmetrical regions. We assumed that the symmetrical regions comprised of two categories: first, the segmented region that exhibits symmetry itself, e.g. nasal bone and anterior mandible. Second, two different segmented regions exhibit symmetry with regard to the plane of symmetry, e.g. condylar process and zygomatic bones. We presented a method to extract the symmetrical regions from segmented regions by analysing the extent of symmetry and performing a clustering technique, as shown in Figure 4.3.

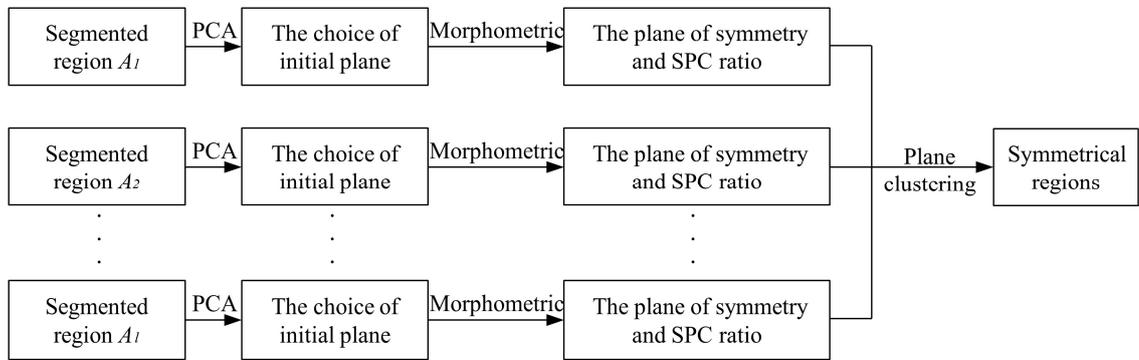


Figure 4.3 The workflow of the extraction of symmetrical regions from segmented regions.

Suppose every segmented region and pairs of two different segmented regions with similar mean WKS values were represented by $A = \{A_i, i=1, 2, \dots, l\}$, the landmark-free morphometric method was applied to identify the plane of symmetry. PCA was first applied to generate three initial planes corresponding to principal eigenvectors of each region. Then we recreated the mirrored meshes of A_i with regard to every initial plane as follows:

$$q_i = p_i - 2\vec{n}(\vec{n} \cdot p_i + d) \quad (4.6)$$

where $p_i=(x_i, y_i, z_i) \in \mathbb{R}^3$ and $q_i=(x_i, y_i, z_i) \in \mathbb{R}^3$ represent every vertex of the original and mirrored meshes, respectively. a, b, c , and d are the coefficients of the equation of a plane, and $\vec{n}=[a \ b \ c]$ denotes the normal vector.

Then the ICP algorithm was used to register the original and mirrored meshes and establish the correspondences between these two meshes by searching the nearest points. The correspondences with a smaller geometric deviation were regarded as symmetrical pairwise correspondences (SPC), and afterward the SPC ratio defined as the percentage of the number of SPC to the number of a full set of vertices was calculated to examine the extent of asymmetry. Notably, the choice of the initial plane affects the accuracy of the identification of the plane of symmetry. In this study, one of the planes derived from PCA with the condition of the maximum SPC ratio should be regarded as the more appropriate initial plane to identifying the plane of symmetry. A_i with a greater SPC ratio was regarded as a potentially symmetrical region (PSR). Otherwise, it was labelled as a falsely symmetrical region.

The next step was to refine the symmetrical regions through plane clustering. The cluster accumulated the consistent planes of symmetry through a comparison of the difference in orientation of normal vectors as follows:

$$\theta = \arccos(\vec{n}_i \cdot \vec{n}_j) \quad (4.7)$$

where \vec{n}_i and \vec{n}_j denote the normal vectors of two different planes, and \cdot denotes the dot product. If θ is smaller, these two planes are regarded as the consistent planes, thereby belonging to the same cluster.

In this study, if the plane of PSR belonged to the cluster that had the maximum consistent planes (hereafter called the consistent plane cluster), PSR would be regarded as the symmetrical regions, denoted by $C=\{C_j\}$, which satisfies $C \subseteq A$. Figure 4.4 shows an example of recognizing the symmetrical regions of the mandible. Figure 4.4a shows five segmented regions using the region growing and merging algorithm. Figure 4.4b shows a graph with five segmented regions as nodes and ten pairs of the segmented regions as edges. It represents possible combinations between the segmented regions to detect symmetrical regions. Figure 4.4c shows the SPC ratio and the plane of symmetry (red colour) of every node (top row) or every pair of nodes (middle and bottom rows). The results show that three separately segmented regions and four pairs of segmented regions (in green and black rectangles) with a greater SPC ratio (> 0.7) are regarded as PSR. Furthermore, only four of them (in green rectangles) with a consistent orientation of normal vectors are regarded as symmetrical regions by performing plane clustering (θ is set to 6° in this experiment).

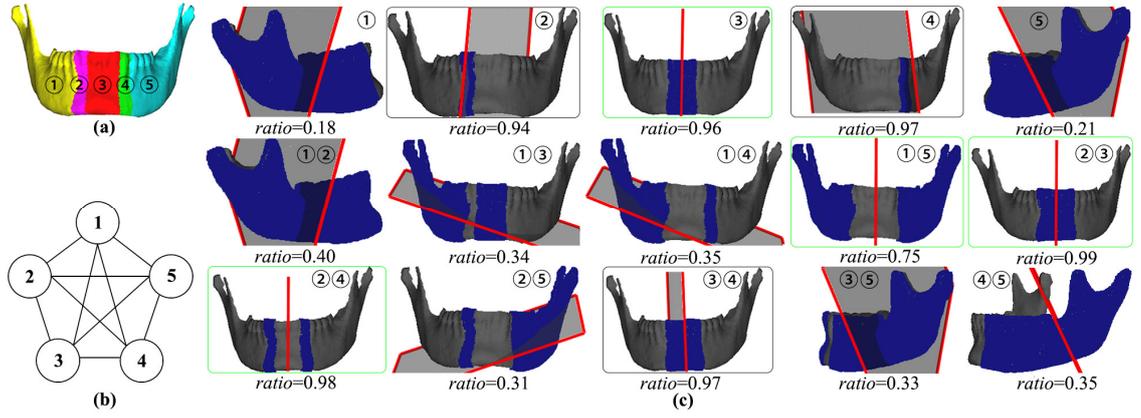


Figure 4.4 The extraction of symmetrical regions. (a) Five segmented regions of the mandible. (b) The graph representing with five segmented regions as nodes and ten pairs of the segmented regions as edges. (c) Every node (top row) and pairs of nodes (middle and bottom rows) accompanying the SPC ratio and the plane of symmetry (red colour). Seven of them (in green and black rectangles) with greater ratios, and only four of them (in green rectangles) within the consistent plane cluster.

The approximation of the plane of symmetry of 3D meshes. The plane with the minimum geometric differences between the original and mirrored meshes should be the best-fitted plane of symmetry of 3D meshes. We approximated the plane of symmetry taking into account two categories: first, the plane of symmetry in each symmetrical region coincided with the plane of symmetry of 3D meshes; Second, the fitted plane derived from at least three centroid points of $C=\{C_j\}$ was consistent with the plane of symmetry of 3D meshes.

Let $co_j=(x_j, y_j, z_j) \in \mathbb{R}^3$ denotes a centroid point of C_j calculated by averaging all the vertices within this symmetrical region, the coefficients of the fitted plane from centroid points regarding a subset of $C=\{C_j\}$ can be optimized as follows

$$\begin{cases} \arg \min_{\bar{\mathbf{n}}_l} \|\mathbf{M} \cdot \bar{\mathbf{n}}_l\| \\ s.t. \|\bar{\mathbf{n}}_l\|^2 = 1 \end{cases} \quad (4.8)$$

where $\bar{\mathbf{n}}_l=[a_l \ b_l \ c_l]^T$ denotes the normal vector of the fitted plane, $co=(x, y, z) \in \mathbb{R}^3$ represents the centroid point of the selection of C_j , and $\mathbf{M}=[co_j-co]$.

Evaluation. Since an archaeological and anthropological object often lacks an actual plane of symmetry, the plane determined by the manually placed landmarks was always regarded as the ground truth to evaluate the accuracy of the proposed method. The difference in orientation of normal vectors between the approximated and the landmark-based planes was calculated to validate the accuracy of the proposed method. It is an acknowledged that the smaller angle indicates the approximated plane is consistent with the manually placed planes. Additionally, we computed geometric differences between the original and mirrored meshes by calculating the average value (*dist*) of Euclidean distances between every pair of corresponding vertices. The student's *t*-test was further carried out to examine the consistency in geometric differences to validate the accuracy.

4.3.2.2 Asymmetry detection

We presented a global and regional method to detect surface asymmetry with regard to the plane of symmetry. An asymmetry ratio (AR) was used to quantify the gross asymmetry of bilaterally symmetrical objects. A greater AR indicates that the meshes exhibited obvious asymmetry. It comprised four steps: first, we used the ICP algorithm to register the original and mirrored meshes, removing the location and orientation. Second, we calculated an average value of geometric differences between the original and registered mirrored meshes. Next, 3D meshes were separated into two halves with regard to the plane of symmetry. Euclidean distance between every vertex and the plane of symmetry was calculated and then maximum distance of each side was computed. Finally, we computed a ratio of the average geometric differences divided by the summed maximum distances. Additionally, the colour map of Euclidean distances of correspondences between the original and registered mirrored meshes was used to quantify the regional asymmetry.

Figure 4.5 gives an example of the qualification of surface asymmetry of the mandible. Figure 4.5a shows the original (grey colour) and mirrored meshes (peach colour), and the colour map of geometric differences between these two meshes. Black colour represents a large difference and yellow colour represents a small difference. Since the differences in location and orientation have not been removed, geometric differences between these two meshes are rather large. We registered the original and mirrored meshes together. Figure 4.5b shows the average value (error) of Euclidean distance between the original and registered mirrored meshes with an increase of iterations. The results show that error declined dramatically and then it gradually converged. Figure 4.5c shows the superimposition of the original (grey colour) and registered mirrored (peach colour) meshes after 100 iterations and the colour map of geometric differences between these two meshes. The results show that asymmetry can be observed around the condylar process.

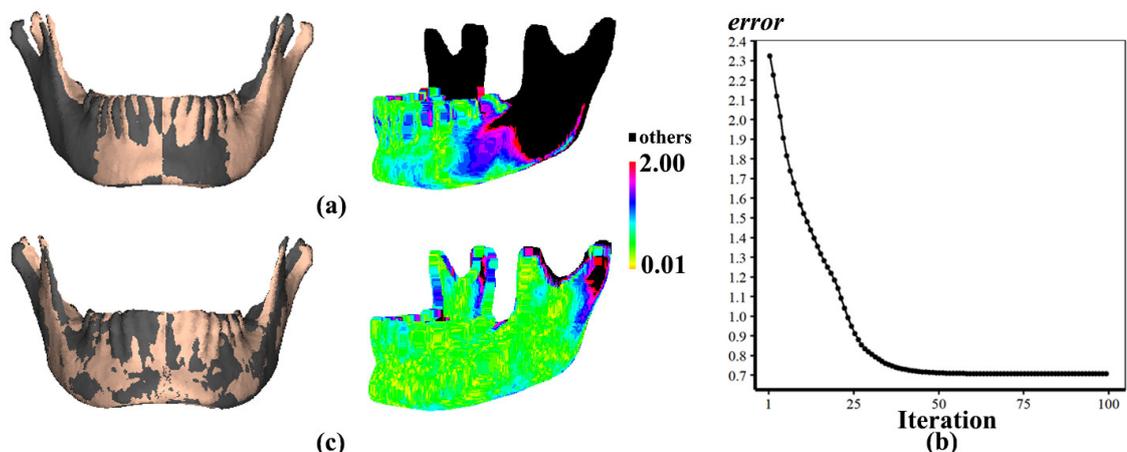


Figure 4.5 The qualification of surface asymmetry of the mandible. (a) The colour map of geometric differences between the original (grey colour) and mirrored meshes (peach colour). (b) Geometric difference between the original and registered mirrored meshes with an increase of iteration. (c) The colour

map of geometric differences between the original (grey colour) and registered mirrored meshes (peach colour) after 100 iterations.

4.3.2.3 Profile drawings

3D meshes were transformed to make the approximated plane of symmetry coincide with the YOZ plane of the coordinate system. Then, the profile was produced by calculating the intersections of the plane of symmetry and every triangle mesh. The intersection points were projected onto the plane of symmetry to improve the smoothness of profile drawings. Additionally, different profile drawings can be sketched when planes of interest are provided.

4.4 Results

4.4.1 The approximation of the plane of symmetry

The identification of the plane of symmetry is the fundamental step in surface asymmetry detection and profile drawings. Both synthetic objects with the actual planes of symmetry and real objects were used to evaluate the accuracy of the proposed method.

4.4.1.1 The plane of symmetry in synthetic data

We compared the approximated plane of symmetry generated by the proposed method with the actual one of synthetic data. Figure 4.6a shows the identification of MSP in a synthetic skull that exhibits nearly perfect bilateral symmetry. The leftmost two figures show twenty-eight growing regions ($\varepsilon_g = 0.99$) and twenty segmented regions by merging the adjacent growing regions ($\varepsilon_m = 0.985$). The third figure shows the consistent plane cluster (red colour) comprising nine planes of symmetry derived from each symmetrical region. The next two figures display two examples of symmetrical regions (blue colour) consisting of a segmented region alone (left) and a pair of segmented regions (right), respectively. The rightmost figure shows the approximated MSP (red colour) generated by the proposed method and the actual MSP (green colour). The result showed that the approximated MSP almost coincided with the actual one. Figure 4.6b shows the identification of MSP in a synthetic skull containing asymmetrical regions around parietal and zygomatic bones and mandible. From left to right, each figure displays the segmented regions, the consistent plane cluster (red colour), and the approximated (red colour) and actual (green colour) MSPs. Figure 4.6c shows the identification of MSP in a synthesized skull that contains a missing region around left parietal bone. In these experiments, there was a small difference in orientation of normal vectors ($\theta < 0.2^\circ$) between the approximated and actual MSPs. Notably, the extracted symmetrical regions, e.g. nasal bone, strongly contribute to recognizing the MSPs.

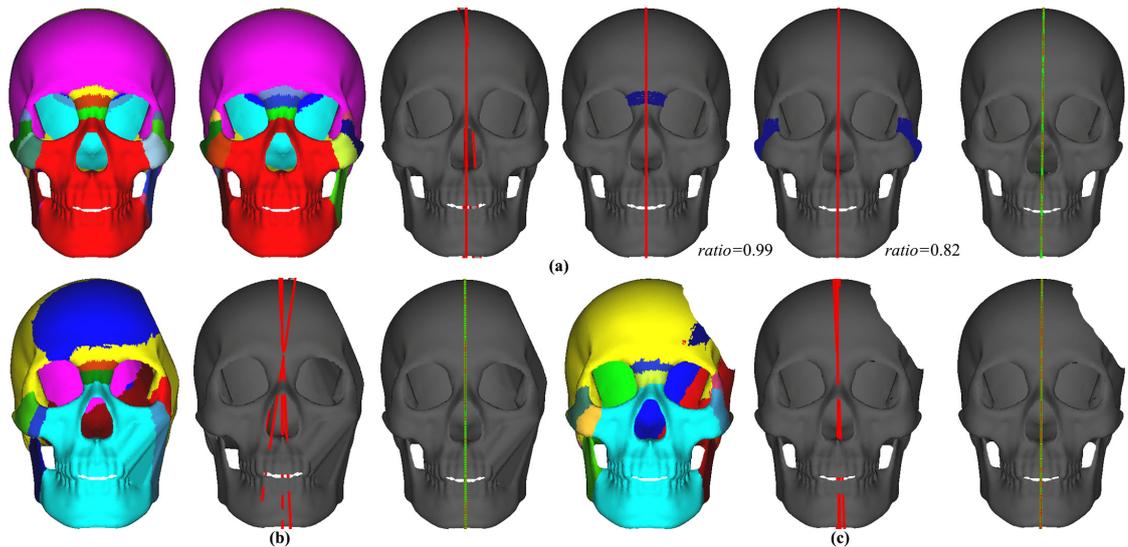


Figure 4.6 The comparison of the approximated and actual MSPs in synthetic skulls. (a) A nearly perfect skull; (b) An asymmetrical skull. (c) A damaged skull with missing data.

4.4.1.2 The plane of symmetry in real objects

The MSPs in the complete skulls. To assess the reliability of the proposed method, we identified the MSPs of thirty real skulls and then compared these approximated MSPs with those derived from the landmark-based morphometric approach (Damstra et al., 2012) which are determined by anatomical landmarks including nasion, nasospinale, prosthion (midline landmarks) and incisura supraorbitalis, frontozygomatic sutures, orbitale and porion (bilateral landmarks). Figure 4.7a shows the differences in orientation of normal vectors between the approximated and landmark-based MSPs. The average value of angles was about 1.24° and the standard deviation was about 0.65° . To examine the degree of difference in orientation of normal vectors, Figure 4.7b displays the approximated (red colour) and landmark-based (green colour) MSPs corresponding to a maximum angle (2.34°) and a minimum angle (0.13°), respectively. These results indicated there was a small orientation difference in normal vectors between two methods. Figure 4.7c shows the geometric differences between the original and mirrored meshes of every skull generated by the proposed (cyan colour) and the landmark-based morphometric (red colour) methods, respectively. The results showed there was a strong Pearson correlation ($r = 0.67$) and no significant difference ($p = 0.357$) between geometric differences generated by these two methods. Additionally, almost 56.67% ($n = 17$) of all the skulls had a smaller geometric deviation using the proposed method. These results indicated that the approximated MSPs generated by the proposed method were consistent with those derived from the landmark-based morphometric method.

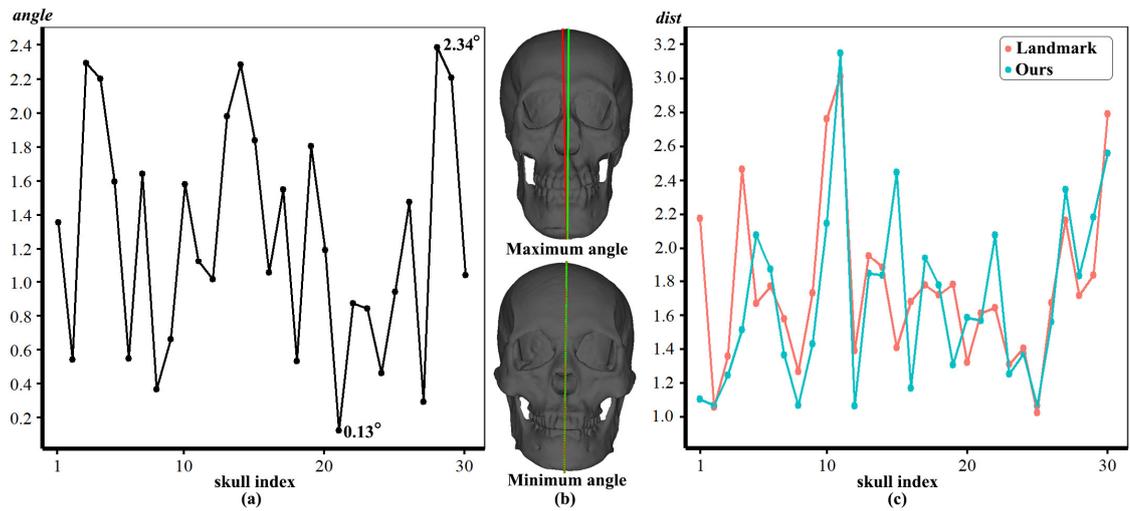


Figure 4.7 The comparison of the approximated and landmark-based MSPs. (a) The difference in orientation between the approximated and landmark-based MSPs. (b) The approximated (red colour) and landmark-based (green colour) MSPs corresponding to a maximum angle (2.34°) and a minimum angle (0.13°), respectively. (c) The comparison of geometric differences between the original and mirrored skulls using the proposed (cyan colour) and the landmark-based morphometric (Damstra et al., 2012) (red colour) methods.

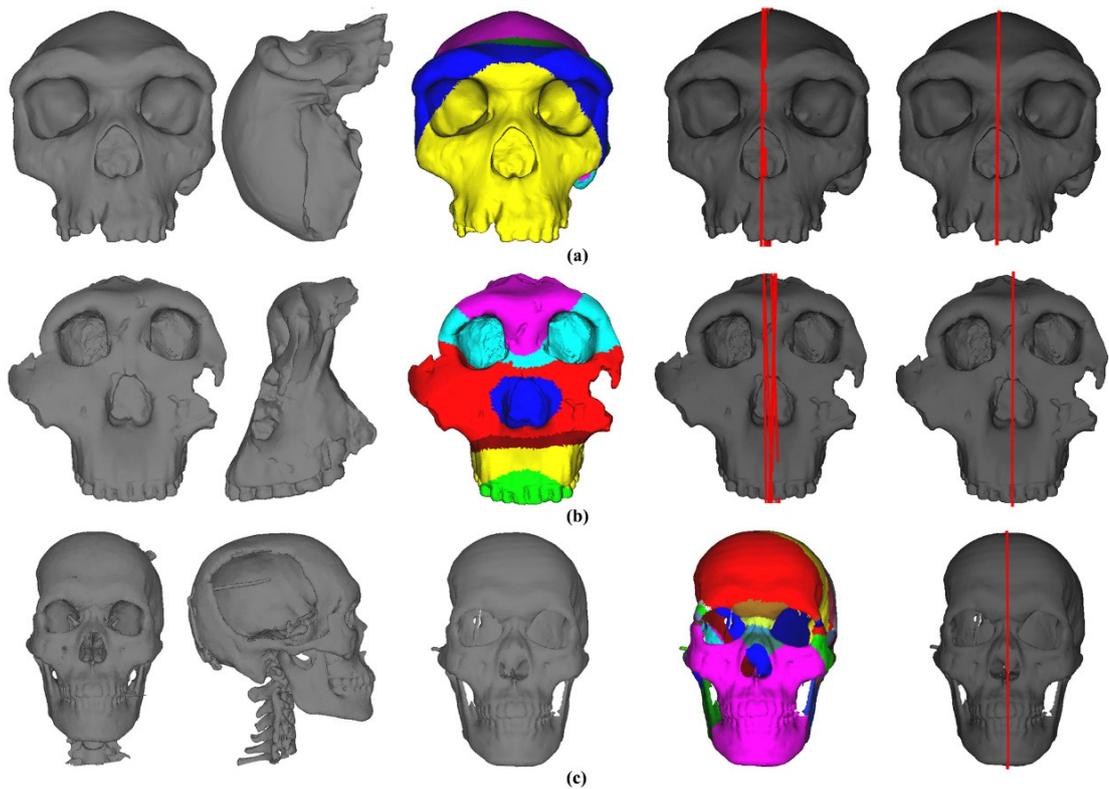


Figure 4.8 The identification of MSPs in three damaged skulls. (a) 2501.1rp35 cranium. (b) 2501.1rp20-1 cranium. (c) The damaged modern human skull.

The MSPs in partial geometry. We identified the MSPs of three damaged bilaterally symmetrical skulls. Figures 4.8a and b show the approximated MSPs from No.2501.1rp35 and No.2501.1rp20-1 crania. From left to right, each column shows the incomplete cranium, the segmented regions, the consistent plane cluster (red colour), and the approximated MSP (red colour), respectively. An example of the identification of the

MSP in a damaged modern human skull was illustrated in Figure 4.8c. To reduce the effect of the inner structure and spine on the identification and speed up computation, we used the method mentioned in literature (Shui et al., 2020) to extract the external surface (middle figure) and then portioned this surface into different regions. The MSP (red colour in rightmost figure) was identified based on the symmetrical regions. These results show the proposed method can identify the plane of symmetry in damaged bilaterally symmetrical objects.

The plane of symmetry in artefacts. We identified the planes of symmetry in the bilaterally symmetrical artefacts. Because there was no clear definition of landmarks over the artefact to identify the plane of symmetry, we used the landmark-free morphometric method to identify the planes and then compared them with those generated by the proposed method. Figure 4.9 shows the approximated planes of six artefacts generated by the proposed (red colour) and landmark-free morphometric (green colour) methods, respectively. Table 4.1 reports geometric differences (mm) between the original and mirrored meshes, and the difference in orientation of normal vectors between the approximated planes of symmetry using these two methods. The smaller geometric deviation and orientation difference indicated that the planes of symmetry of the terracotta warrior head, pottery spoon, ceramic plate and clay pot (Figures 4.9a-d) generated by two methods were almost identical. However, the planes of symmetry (green colour) of the tripod and helmet (Figures 4.9e-f) generated by the landmark-free morphometric method seemed less confidence, because these artefacts contained missing parts and exhibited asymmetry. Additionally, the identification of the plane of symmetry from an asymmetrical artefact was depicted in Figure 4.10. Figure 4.10a shows the segmented regions of a terracotta warriors hand based on the WKS values. The hand was portioned into four main parts, including fingers, thumb, the end part of hand and others. Figure 4.10b shows the approximated plane (red colour) of symmetry generated by the proposed method and that (green colour) derived from the landmark-free morphometric method. The results showed these two approximated planes were different, because the approximated plane using the landmark-free morphometric method was primarily determined by the entire meshes.

Table 4.1 The approximated planes of symmetry of different artefacts

Artefact	<i>dist</i> (ours)	<i>dist</i> (morphometric method)	Angle (°)
terracotta warriors head	2.16	2.95	1.81
pottery spoon	0.59	0.54	1.38
ceramic plate	0.34	0.37	0.88
clay pot	0.45	0.46	1.16
tripod	1.81	3.13	6.85
helmet	3.14	13.19	10.26

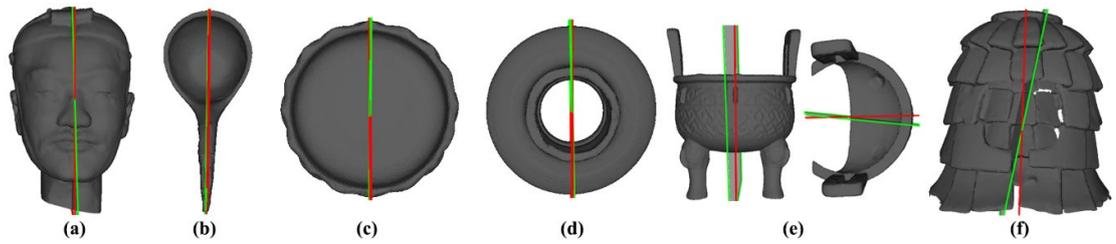


Figure 4.9 The comparison of the approximated planes of symmetry generated by the proposed (red colour) and the landmark-free morphometric (green colour) methods. (a) Terracotta warriors head. (b) Pottery spoon. (c) Ceramic plate. (d) Clay pot. (e) Tripod. (f) Helmet.

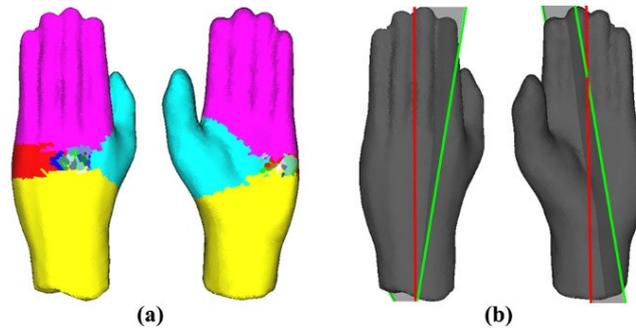


Figure 4.10 The identification of the plane of symmetry in a terracotta warriors hand. (a) Segmented regions. (b) The approximated planes generated by the proposed (red colour) and the landmark-free morphometric (green colour) methods.

4.4.2 Asymmetry detection

Based on the approximated plane of symmetry, we used the proposed method to quantify the gross and regional surface asymmetry of thirty skulls, as shown in Figure 4.11. Figure 4.11a shows the AR to capture the gross asymmetry of every skull. Figure 4.11b displays the most asymmetrical skull corresponding to the highest AR (0.0191) and the colour map of geometric differences (mm) between the original and registered mirrored skulls. Red colour represents the obviously asymmetrical regions and yellow colour indicates the nearly symmetrical regions. The result showed the greatest asymmetrical regions were found around parietal and occipital bones. Figure 4.11c displays the nearly symmetrical skull corresponding to the smallest AR (0.0060) and the colour map of geometric differences. To validate the reliability of the proposed method, we compared AR generated by the proposed method to those derived from the landmark-based morphometric method (Damstra et al., 2012). The results showed there was a strong Pearson correlation ($r = 0.94$) and no significant difference ($p = 0.763$) between AR generated by these two methods. Additionally, we used the proposed method to quantify the surface asymmetries of artefacts, as shown in Figure 4.12.

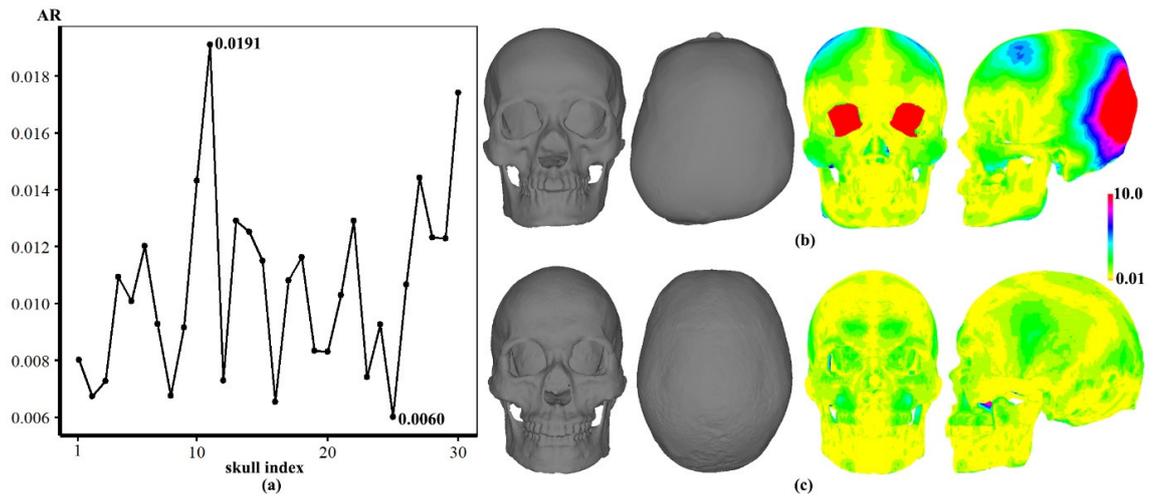


Figure 4.11 Surface asymmetry detection of thirty skulls. (a) AR of every skull. (b) The most asymmetrical skull corresponding to the highest AR and the colour map of geometric difference. (c) The nearly symmetrical skull corresponding to the smallest AR and the colour map of geometric difference.

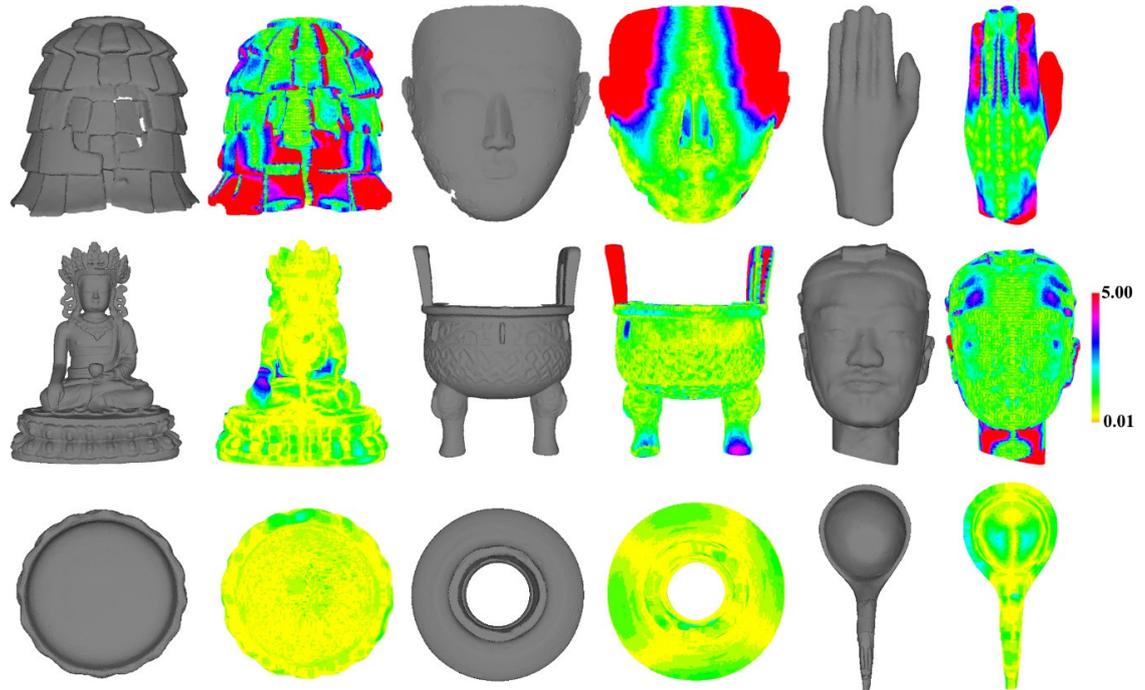


Figure 4.12 Surface asymmetry detection of archaeological artefacts.

4.4.3 Profile drawings

We produced profile drawings of nearly bilaterally symmetrical objects based on the plane of symmetry or an arbitrary plane. Figure 4.13a shows profile drawings of the terracotta warriors head, bronze mask, ceramic plate, pottery spoon and Tibetan Buddha statue with regard to the plane of symmetry. Figure 4.13b displays profile drawings of the skull, human endocast and femur. Figure 4.13c shows the profile drawing of the skull associated with the transverse plane and that of the tripod with regard to an arbitrary plane.

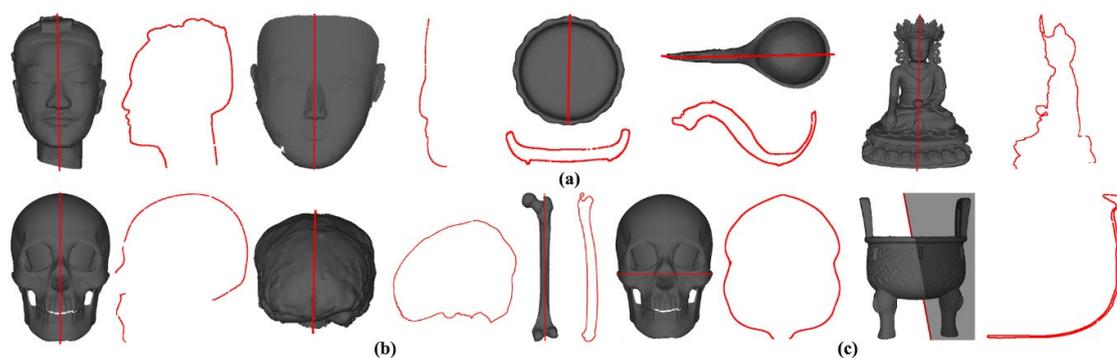


Figure 4.13 Profile drawings of symmetrical objects. (a) Profile drawings of artefacts with regard to the plane of symmetry. (b) Profile drawings of anthropological objects with regard to the plane of symmetry. (c) Profile drawings of objects with regard to different planes.

4.5 Discussion

The increased accessibility of digital models and the development of computerized technologies have led researchers to quantify surface asymmetry and produce profile drawings of archaeological artefacts and fossil specimens automatically, objectively and effectively. This study presented a landmark-free approach for approximating the plane of symmetry from nearly bilaterally symmetrical objects. Subsequently, we quantified surface asymmetry and produced profile drawings with regard to the plane of symmetry.

The identification of the plane of symmetry is a crucial step in quantifying asymmetry using the morphometric method. As natural objects always exhibit different extents of asymmetry, e.g. fluctuating asymmetry, the actual plane of symmetry always does not exist. The landmark-based morphometric method is the most commonly used method to identify the plane of symmetry (Damstra et al., 2012). Nevertheless, the debate over the choice of landmark configuration has attracted great interest. For example, central landmarks and bilateral landmarks have been suggested for approximating the MSP (Willing et al., 2013). On the contrary, central landmarks including nasion, basion and incisive foramen are recommended in other study (Green et al., 2017). Instead, this study attempts to find the best-fitted plane that yields the minimum geometric differences between the original and mirrored meshes. The results clearly demonstrate that the proposed and landmark-based methods can yield consistent MSPs in skulls. This finding suggests the use of high-density SPC does not significantly alter MSPs. Additionally, the proposed method avoids landmark placements, especially in featureless and smooth regions, e.g. clay pot. Hence, it contributes to identifying the plane of symmetry in archaeological artefacts.

Compared to nearly perfectly symmetrical objects, the identification of the plane of symmetry in partial geometry and asymmetrical objects faces a challenging problem using the landmark-free morphometric method (Di Angelo et al., 2019; Zhang et al., 2006). Because the centroid point and principal directions generated by PCA before and after being broken are different, it is likely to yield an inaccurate mid-point configuration

between original and registered mirrored meshes, thereby leading to an unreasonable plane of symmetry, e.g. the tripod and helmet. To address this problem, a set of symmetrical regions were manually segmented to approximate the plane of symmetry (De Momi et al., 2006; Xu et al., 2022). The idea behind is that the planes of symmetry in the extracted symmetrical regions are consistent with that of the entire meshes. This study employed a coarse-to-fine strategy to extract the symmetrical regions, avoiding human intervention. The extraction of symmetrical regions through the examination of the extent of asymmetry and plane clustering strongly contributes to identifying the plane of symmetry of objects with geometric distortions or simple missing geometry. However, our method might not be suitable for objects with substantial missing geometry due to lack of sufficiently symmetrical regions. In case that the entire meshes are regarded as a single region, the identification of the plane of symmetry derived from the proposed method would be consistent with that derived from the landmark-free morphometric method.

The extraction of symmetrical regions is partially dependent upon the segmentation results using the growing and merging algorithm. The smaller growing threshold is likely to lead to under-segmentation and ignore PSR. The insufficient symmetrical regions might result in an inaccurately approximated plane, especially in missing geometry and partial geometry. In contrast, a higher growing threshold can yield more segmented regions, thereby facilitating the extraction of PSR from partial and asymmetrical geometry. It is noteworthy that the over-segmentation increases the complexity of computation in extracting the symmetrical regions and identifying the plane of symmetry. In practice, the threshold values need to be investigated taking into account the extent of completeness of geometry and the degree of asymmetry. The same threshold values can always be used to extract symmetrical regions and identify the planes of symmetry from the same group of objects, e.g. skulls, reducing human intervention.

Several methods have been proposed to quantify the magnitude of surface asymmetry by measuring geometric differences between the original and registered mirrored meshes (Cassi et al., 2019; Claes et al., 2011; Gibelli et al., 2018; Melchionna et al., 2021; Verhoeven et al., 2016) or the normal distance between centroids of corresponding triangles (Willing et al., 2013). To examine the extent of asymmetry across a collection of specimens, an asymmetry index defined as the geometric differences between the original and the mirrored meshes divided by the diagonal length of the bounding box has been proposed to quantify the global asymmetry (Bockey et al., 2018). However, the orientation of the object affects the diagonal length, thereby leading to an inaccurate analytical result. In this study, AR regarding geometric differences divided by the length associated with the plane of symmetry was used to measure the gross asymmetry, minimizing the effect of the size and orientation of 3D meshes on asymmetry detection.

4.6 Conclusion

In archaeology and anthropology, the recognition of the plane of symmetry plays an important role in virtual restoration, asymmetry detection and profile drawings. Previous studies have proposed the landmark-free morphometric method to approximate the plane of symmetry. However, missing data and asymmetry might lead to inaccurate identifications. This study proposes an effective landmark-free approach to find the best-fitted plane of symmetry with the minimum geometric differences between the original and mirrored meshes. Subsequently, we present a global and regional method to quantify surface asymmetry. Finally, we produce profile drawings based on the plane of symmetry or an arbitrary plane. The experiments demonstrated that the proposed method is effective and robust for symmetrical objects with small geometric distortions or simple missing geometry. We expect the proposed method will have a range of applications.

4.7 Acknowledgements

I would like to thank Beijing Key Laboratory of Digital Preservation and Virtual Reality for Cultural Heritage, Beijing Normal University for granting access to their human skull database and artefacts. I would like to express my gratitude to the Research Labs of Archaeology at the University of Chapel Hill, North Carolina, and the online digital repository on Sketchfab for kindly providing their crania.

4.8 References

- Angelo, LD, Stefano, PD, Guardiani, E, & Pane, C. (2020). Automatic Shape Feature Recognition for Ceramic Finds. *Journal on Computing Cultural Heritage*, 13(3), 1-21.
- Aubry, M, Schlickewei, U, & Cremers, D. (2011). The wave kernel signature: A quantum mechanical approach to shape analysis. *The 2011 IEEE international conference on computer vision workshops*, 1626-1633.
- Besl, PJ, & McKay, ND. (1992). A method for registration of 3-D shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2), 239-256.
- Biasotti, S, Cerri, A, Bronstein, A, & Bronstein, M. (2016). Recent trends, applications, and perspectives in 3d shape similarity assessment. *Computer Graphics Forum*, 35(6), 87-119.
- Bockey, S, Berssenbrügge, P, Dirksen, D, Wermker, K, Klein, M, & Runte, C. (2018). Computer-aided design of facial prostheses by means of 3D-data acquisition and following symmetry analysis. *Journal of Cranio-Maxillofacial Surgery*, 46(8), 1320-1328.
- Cassi, D, Battistoni, G, Magnifico, M, Di Blasio, C, Pedrazzi, G, & Di Blasio, A. (2019). Three-dimensional evaluation of facial asymmetry in patients with hemifacial microsomia using stereophotogrammetry. *Journal of Cranio-Maxillofacial Surgery*, 47(1), 179-184.
- Cignoni, P, Callieri, M, Corsini, M, Dellepiane, M, Ganovelli, F, & Ranzuglia, G. (2008). MeshLab: an Open-Source Mesh Processing Tool. *The Eurographics Italian Chapter Conference*, 129-136.
- Claes, P, Walters, M, Vandermeulen, D, & Clement, G. (2011). Spatially-dense 3D facial asymmetry assessment in both typical and disordered growth. *Journal of anatomy*, 219(4), 444-455.
- Corsini, M, Cignoni, P, & Scopigno, R. (2012). Efficient and flexible sampling with blue noise properties of triangular meshes. *IEEE Transactions on Visualization and Computer Graphics*, 18(6), 914-924.
- Damstra, J, Fourie, Z, De Wit, M, & Ren, Y. (2012). A three-dimensional comparison of a morphometric and conventional cephalometric midsagittal planes for craniofacial asymmetry. *Clinical Oral Investigations*, 16(1), 285-294.
- De Momi, E, Chapuis, J, Pappas, I, Ferrigno, G, Hallermann, W, Schramm, A, & Caversaccio, M. (2006). Automatic extraction of the mid-facial plane for cranio-maxillofacial surgery planning. *International Journal of Oral Maxillofacial Surgery*, 35(7), 636-642.
- Di Angelo, L, Di Stefano, P, Governi, L, Marzola, A, & Volpe, Y. (2019). A robust and automatic method for the best symmetry plane detection of craniofacial skeletons.

Symmetry, 11(2), 245.

Gao, L, Zhang, L, Meng, H, Ren, Y, Lai, Y, & Kobbelt, L. (2020). PRS-Net: Planar reflective symmetry detection net for 3D models. *IEEE Transactions on Visualization Computer Graphics Forum*, 27(6), 3007-3018.

Gibelli, D, Cellina, M, Gibelli, S, Oliva, AG, Termine, G, Pucciarelli, V, Dolci, C, et al. (2018). Assessing symmetry of zygomatic bone through three-dimensional segmentation on computed tomography scan and “mirroring” procedure: a contribution for reconstructive maxillofacial surgery. *Journal of Cranio-Maxillofacial Surgery*, 46(4), 600-604.

Green, MN, Bloom, JM, & Kulbersh, R. (2017). A simple and accurate craniofacial midsagittal plane definition. *American Journal of Orthodontics Dentofacial Orthopedics*, 152(3), 355-363.

Gruen, A. (2012). Development and status of image matching in photogrammetry. *The Photogrammetric Record*, 27(137), 36-57.

Hou, M, & Fagan, M. (2021). Assessments of bilateral asymmetry with application in human skull analysis. *Plos one*, 16(10), e0258146.

Ji, P, & Liu, X. (2019). A fast and efficient 3D reflection symmetry detector based on neural networks. *Multimedia Tools Applications*, 78, 35471-35492.

Kakarala, R, Kaliamoorthi, P, & Premachandran, V. (2013). Three-dimensional bilateral symmetry plane estimation in the phase domain. *The 2013 IEEE Conference on Computer Vision and Pattern Recognition*, 249-256.

Karasik, A, & Smilansky, U. (2008). 3D scanning technology as a standard archaeological tool for pottery analysis: practice and theory. *Journal of Archaeological Science*, 35(5), 1148-1168.

Klingenberg, CP. (2015). Analyzing fluctuating asymmetry with geometric morphometrics: concepts, methods, and applications. *Symmetry*, 7(2), 843-934.

Kuzminsky, SC, & Gardiner, MS. (2012). Three-dimensional laser scanning: potential uses for museum conservation and scientific research. *Journal of Archaeological Science*, 39(8), 2744-2751.

Li, C, & Hamza, AB. (2014). Spatially aggregating spectral descriptors for nonrigid 3D shape retrieval: a comparative survey. *Multimedia Systems*, 20(3), 253-281.

McGrath, K, Eriksen, AB, García-Martínez, D, Galbany, J, Gómez-Robles, A, Massey, JS, Fatica, LM, et al. (2022). Facial asymmetry tracks genetic diversity among Gorilla subspecies. *Proceedings of the Royal Society B: Biological Sciences*, 289(1969), 20212564.

- McNabb, J, Cole, J, & Hoggard, CS. (2018). From side to side: Symmetry in handaxes in the British Lower and Middle Palaeolithic. *Journal of Archaeological Science: Reports*, 17, 293-310.
- Melchionna, M, Profico, A, Buzi, C, Castiglione, S, Mondanaro, A, Del Bove, A, Sansalone, G, et al. (2021). A New Integrated Tool to Calculate and Map Bilateral Asymmetry on Three-Dimensional Digital Models. *Symmetry*, 13(9), 1644.
- Mitra, NJ, Guibas, LJ, & Pauly, M. (2006). Partial and approximate symmetry detection for 3D geometry. *ACM Transactions on Graphics* 25(3), 560-568.
- Mitra, NJ, Pauly, M, Wand, M, & Ceylan, D. (2013). Symmetry in 3d geometry: Extraction and applications.
- Neubauer, S, Gunz, P, Scott, NA, Hublin, JJ, & Mitteroecker, P. (2020). Evolution of brain lateralization: A shared hominid pattern of endocranial asymmetry is much more variable in humans than in great apes. *Science Advances*, 6(7), eaax9935.
- Noori, SMR, Farnia, P, Bayat, M, Bahrami, N, Shakourirad, A, & Ahmadian, A. (2020). Automatic detection of symmetry plane for computer-aided surgical simulation in craniomaxillofacial surgery. *Physical Engineering Sciences in Medicine*, 43(3), 1087-1099.
- O'Higgins, P. (2000). The study of morphological variation in the hominid fossil record: biology, landmarks and geometry. *Journal of Anatomy*, 197(1), 103-120.
- Ovsjanikov, M, Sun, J, & Guibas, L. (2008). Global intrinsic symmetries of shapes. *Computer Graphics Forum*, 27(5), 1341-1348.
- Padia, C, & Pears, N. (2011). A review and characterization of ICP-based symmetry plane localisation in 3D face data. *Technical report, Department of Computer Science, University of York, UK*.
- Profico, A, Zeppilli, C, Micarelli, I, Mondanaro, A, Raia, P, Marchi, D, Manzi, G, et al. (2021). Morphometric maps of bilateral asymmetry in the human humerus: An implementation in the R package morphomap. *Symmetry*, 13(9), 1711.
- Saragusti, I, Sharon, I, Katzenelson, O, & Avnir, D. (1998). Quantitative analysis of the symmetry of artefacts: Lower Paleolithic handaxes. *Journal of Archaeological Science*, 25, 817-825.
- Shui, W, & Gao, F. (2021). A geometric completion and shape analysis method for damaged bilaterally symmetrical artifacts. *Journal of Cultural Heritage*, 52, 118-127.
- Shui, W, Zhou, M, Chen, S, Pan, Z, Deng, Q, Yao, Y, Pan, H, et al. (2017). The production of digital and printed resources from multiple modalities using visualization and three-dimensional printing techniques. *International Journal of Computer Assisted Radiology and Surgery*, 12(1), 1-11.

- Shui, W, Zhou, M, Deng, Q, Wu, Z, Ji, Y, Li, K, He, T, et al. (2016). Densely Calculated Facial Soft Tissue Thickness for Craniofacial Reconstruction in Chinese Adults. *Forensic Science International*, 266, 573.e571-e512.
- Shui, W, Zhou, M, Maddock, S, Ji, Y, Deng, Q, Li, K, Fan, Y, et al. (2020). A computerized craniofacial reconstruction method for an unidentified skull based on statistical shape models. *Multimedia Tools and Applications*, 79, 25589-25611.
- Sipiran, I, Gregor, R, & Schreck, T. (2014). Approximate symmetry detection in partial 3d meshes. *Computer Graphics Forum*, 33(7), 131-140.
- Sun, C, & Sherrah, J. (1997). 3D symmetry detection using the extended Gaussian image. *IEEE Transactions on Pattern Analysis Machine Intelligence*, 19(2), 164-168.
- Sun, J, Ovsjanikov, M, & Guibas, L. (2009). A concise and provably informative multi-scale signature based on heat diffusion. *Computer Graphics Forum*, 28(5), 1383-1392.
- Tevs, A, Huang, Q, Wand, M, Seidel, HP, & Guibas, L. (2014). Relating shapes via geometric symmetries and regularities. *ACM Transactions on Graphics*, 33(4), 1-12.
- Verhoeven, T, Xi, T, Schreurs, R, Bergé, S, & Maal, T. (2016). Quantification of facial asymmetry: a comparative study of landmark-based and surface-based registrations. *Journal of Cranio-Maxillofacial Surgery*, 44(9), 1131-1136.
- Wan, SY, & Higgins, WE. (2003). Symmetric region growing. *IEEE Transactions on Image Processing*, 12(9), 1007-1015.
- Wilczek, J, Monna, F, Jébrane, A, Chazal, CL, Navarro, N, Couette, S, & Smith, CC. (2018). Computer-assisted orientation and drawing of archaeological pottery. *Journal on Computing and Cultural Heritage*, 11(4), 1-17.
- Willing, RT, Roumeliotis, G, Jenkyn, TR, & Yazdani, A. (2013). Development and evaluation of a semi-automatic technique for determining the bilateral symmetry plane of the facial skeleton. *Medical Engineering Physics*, 35(12), 1843-1849.
- Wynn, T. (2002). Archaeology and cognitive evolution. *Behavioral Brain Sciences*, 25(3), 389-402.
- Xu, J, Cao, W, Liu, B, & Jiang, K. (2022). Object restoration based on extrinsic reflective symmetry plane detection. *The Visual Computer*, 38, 3627-3642.
- Xu, K, Zhang, H, Tagliasacchi, A, Liu, L, Li, G, Meng, M, & Xiong, Y. (2009). Partial intrinsic reflectional symmetry of 3d shapes. *ACM Transactions on Graphics*, 28(5), 1-10.
- Zhang, D, Wu, Z, Wang, X, Lv, C, & Liu, N. (2021). 3D skull and face similarity measurements based on a harmonic wave kernel signature. *The Visual Computer*, 37(4), 749-764.

Zhang, L, Razdan, A, Farin, G, Femiani, J, Bae, M, & Lockwood, C. (2006). 3D face authentication and recognition based on bilateral symmetry analysis. *The Visual Computer*, 22(1), 43-55.

Zhang, X, Blaas, J, Botha, C, Reischig, P, Bravin, A, & Dik, J. (2012). Process for the 3D virtual reconstruction of a microcultural heritage artifact obtained by synchrotron radiation CT technology using open source and free software. *Journal of Cultural Heritage*, 13(2), 221-225.

Supplementary material

Table 4.S1 Archaeological and anthropological objects

Artefact	Scanner	Number	Repository
Human skulls	CT	30	Beijing Normal University
Damaged skull	CT	1	Beijing Normal University
Femur	CT	1	Fudan University
Bronze mask	Handyscan 700	1	Beijing Normal University
Tripod	Artec Spider	1	Beijing Normal University
Helmet	Artec Spider	1	Beijing Normal University
Pottery spoon	Artec Spider	1	Zhejiang University
Ceramic plate	Artec Spider	1	Zhejiang University
Clay pot	Artec Spider	1	Zhejiang University
Tibetan Buddha statue	Artec Spider	1	Qinghai Normal University
<i>Rhodesian Man</i>	Photogrammetry	1	Sketchfab
<i>Paranthropus boisei</i>	Photogrammetry	1	Sketchfab

Chapter 5 A geometric completion and shape analysis method for damaged bilaterally symmetrical artefacts

Abstract: Archaeological artefacts are important forms of tangible cultural heritage, providing evidence of cultural signification, human values, and the development of manufacturing technologies. The reconstruction of damaged bilaterally symmetrical artefacts is frequently undertaken for research and public display. However, traditional restoration technology requires physical intervention, which can cause secondary damage, and raise questions of authenticity. In this study, we proposed a hybrid non-rigid deformation approach to repair damaged bilaterally symmetrical artefacts based on a study of a heavily corroded metal mask of the Liao dynasty from the Palace Museum, China. In addition, since the mask is always placed over the head of the deceased person, it is an interesting question to capture the characteristic features. We employed geometric morphometrics (GM) and anthropometric measurements (AM) to investigate the geometric variations between the mask and human faces. The experimental results demonstrated the effectiveness of the proposed approach in the application of virtual restoration and shape analysis. These methods have wider application to other forms of archaeological and anthropological data.

Article details:

Wuyang Shui and Fei Gao. A geometric completion and shape analysis method for damaged bilaterally symmetrical artefacts. 2021. *Journal of Cultural Heritage*. 2021, 52, 118-127.

5.1 Introduction

Archaeological artefacts provide reliable historic evidence for the past organization of human society, and allow us study the knowledge and skills of our ancestors. However, the majority of artefacts have been damaged due to weathering, erosion, physical stress, human activity, etc. In order to provide reliable evidence and clues for scientific research and preserve the value of artefacts, archaeologists and conservators need to undertake restoration. Existing approaches are generally concerned with two aspects: fragment reassembly and completion of missing geometry (Papaioannou et al., 2017). During this procedure, conservators often repair damaged artefacts by hand, but they face several challenges, including the cost of lengthy time-consuming processes, risks of physical intervention, and the introduction of subjective biases (Gilboa et al., 2013). These have raised the interest of conservators in improving the reliability and speed of restoration, and preventing secondary damage. Exploring computerized approaches offers a promising perspective for the completion of missing geometry, in particular, with regard to those artefacts that exhibit bilateral symmetry.

Over recent decades, cutting-edge digital technologies provide new insights for the analysis, interpretation and repair of artefacts. They present a potential way to address the limitations of the traditional approach to restoration and attract great interest from both conservators and computer scientists. Previous studies have employed 3D data acquisition technologies, e.g. laser scanning (Kuzminsky & Gardiner, 2012), photogrammetry (Sapirstein, 2018), computerized tomography (CT) (Zhang et al., 2012), etc., to record the entire geometric shape and texture of the artefacts. These high-quality 3D models of cultural resources can be accessed and investigated, rather than the actual artefacts. For example, conservators utilize interactive software to reassemble fragments and fill in the gaps. In recent years, 3D printing (also known as additive manufacturing) has been applied to produce replicas and repair damaged artefacts (Arbace et al., 2013; Jo et al., 2020). Without considering the characteristics of the materials, this technology provides a complementary component for the traditional casting technology.

A damaged metal mask (21.00 cm length, 19.00 cm width and 0.5 mm depth) of the Liao dynasty from the Palace Museum, China, shown in Figure 5.1a, had been broken into a large fragment and four small ones, each of which was covered in a green patina. To avoid secondary damage and further deterioration, and display the original appearance to the general public, it was decided to repair the mask by means of interdisciplinary technologies. Before the physical restoration, the composition and corrosion of the mask were examined using different analytical technologies, including X-Ray diffraction, Raman Spectroscopy, and Micro-Fourier Transform Infrared Spectroscopy (Qu et al., 2018). Subsequently, the conservator (the second author, Fei Gao) cleaned the surface of the mask manually, removed the corrosion, and consolidated the

mask using Cyclododecane. Finally, all the fragments were repositioned and reassembled by hand. However, the right cheek of the repaired mask still had two missing areas (Figure 5.1b). A computerized completion method was required to produce digital models that could be embedded within the damaged regions, such that the boundary curves of the embedded models are consistent with those of the damaged regions.



Figure 5.1 The metal mask. (a) The mask consisted of five fragments. (b) Fragment reassembly.

The Liao dynasty was founded by nomadic Qidan tribes in 907 A.D. During this period, the preservation of the remains of the deceased formed the central part of the funerary process that was an important cultural event and burial custom. The metal masks used to cover the faces of deceased are often recovered from the tombs (Mayberger, 2020; Peng, 2003). The masks were personalized and hence always have different characteristic features. For example, Shimada described a mask from the collection of the archaeological museums of Tokyo University with wide open eyes, comparatively large nose, indicated nostrils, and heavy upper and lower lips (Shimada, 1950). Both the general public and the conservators wanted to understand the extent to which the mask resembled a human face.

In this study, we presented a computerized approach to repair missing geometry and analyse the shape variations of archaeological artefacts that exhibit bilateral symmetry. It comprises the estimation of the plane of symmetry, the restoration of the missing geometry based on geometric deformation, and the landmarks-based shape analysis. It improves the reliability of completion and minimizes the effect of corrosion and distortion on the accuracy of deformation. The key contributions are as follows:

- We propose a framework with minimal human intervention to accomplish the completion of the damaged artefact and shape analysis.
- We present a hybrid non-rigid deformation combining thin-plate splines (TPS) (Bookstein, 1989) and Laplacian deformation (Sorkine et al., 2004) to repair the damaged artefact with the predicted boundary curve as a constraint.
- We employ landmark-based geometric morphometrics (GM) and anthropometric measurements (AM) to capture the characteristic features of the mask. The findings using GM are almost consistent with those using AM.

5.2 Related work

In recent years, there has been an increasing demand for repairing broken artefacts and capturing geometric features using advanced computer technologies. Geometric completion is one of the most important tasks in repairing damaged bilaterally symmetrical artefacts. In general, it is a straightforward way to warp the template model to the damaged region by means of non-rigid deformation methods.

During completion, the first step is to extract the plane of symmetry. Biologists often estimate the plane based on paired landmarks placed over the fossil manually (Gunz et al., 2009). Such landmarks can also guide the deformation and allow shape analysis to be performed using GM. To avoid intervention, many computerized approaches have attempted to extract the plane of symmetry automatically based on intrinsic and extrinsic features. Sipiran et al. (2014) recognized the symmetric plane from a set of candidate symmetries based on heat diffusion geometric features and a vote-based strategy. However, this algorithm is not suitable for models with less geometric features. Li et al. (2011) designed a descriptor that consisted of the principal curvatures and a shape diameter function to identify the plane of symmetry. But the distortion and noise data within the surface may influence the accuracy of estimation.

After the template model was recreated, a thin-plate splines (TPS) interpolation function was often used to repair the damaged model and correct the distortion (Gunz et al., 2009). During the interpolation, a landmark configuration is used as control points that need to be placed manually or automatically. It is of note that the number and positions of the landmarks will influence the resulting deformation. When the control points were limited, the resulting deformation was always inadequate. Schlager et al. (2018) placed 3D bilateral semilandmarks on the curves and surfaces to cover the entire fossil to enhance the quality of deformation. In addition, Deng et al. (2011) generated extra geometric correspondences and employed the compactly supported radial basis functions interpolation algorithm to closely align the two surfaces. But above interpolation methods are often sensitive to the degree of complex surface.

To improve the accuracy of deformation, a variety of non-rigid registration methods have been proposed by means of computing the transformation of every vertex of the template model. Amberg et al. (2007) proposed the non-rigid ICP (NICP) algorithm to enhance the accuracy of deformation by optimizing a cost function that consisted of different energy terms. In this process, a landmark term was used as a soft constraint to speed up the convergence. Sorkine et al. (2004) proposed the Laplacian deformation based on the discrete Laplacian operator. The geometric coordinates of every vertex can be calculated by optimizing an error function that comprised a fitting term that measured the distance of landmarks between two surfaces and a distortion term that quantified the extent to which the template model changed. Myronenko and Song (2010) presented a

coherent point drift (CPD) method for two point clouds alignment by solving a probability density estimation problem. Dai et al. (2020). employed the improved CPD method combining Laplace-Beltrami projection and Gaussian process posterior model to accomplish the deformation and establish the correspondences We need to investigate an approach to improve the accuracy of deformation for archaeological artefacts with heavy erosion.

It is an important question to characterize the shape features of artefacts. Researchers have generally employed an AM approach based on analysing the geometric measurements, e.g. distances, angles, ratios, surface area, etc. (Munn & Stephan, 2018). However, such variables are not suitable for the analysis of shape variations and covariation among different specimens. In biology, GM that combines quantitative statistical analysis and visualisation technology is a powerful tool to capture the main patterns of the specimen and interpret the development of fossil evolution (O'Higgins & Jones, 1998). Landmarks are central to the study, because different landmarks always result in different distance matrices and lead to ambiguous findings (Oxnard & O'Higgins, 2009). During this procedure, Generalized Procrustes Analysis (GPA) and Principal Component Analysis (PCA) are used to transform the Cartesian coordinates of the landmark configuration in shape space. Visualisation and quantitative analysis technologies are used to capture the shape features.

Few studies have employed GM to analyse archaeological artefacts (Okumura & Araujo, 2019). The main challenges are the limited number of suitable artefacts and the definition of landmarks among different groups. In archaeology, researchers have begun to employ the quantitative analysis method to capture the features of stone tools based on 2D landmarks during previous decades (Brande & Saragusti, 1996; Gero & Mazzullo, 1984). More recently, 3D GM has been used to investigate the morphological variations of stone tools based on landmarks only or landmarks and semilandmarks together. Shott and Brian (2010) employed GM to investigate shape variations of lithics and demonstrated the effectiveness of shape analysis. Okumura and Araujo (2014) adopted GM to interpret the geometric shape of lithic stemmed bifacial points over time. In their work, the relative warp ordination plot of each specimen was visualized in shape space, and partial least squares was used to examine the correlation among different groups. We need to investigate whether GM can be used when the group of artefacts is limited.

5.3 Materials and methods

5.3.1 Materials

A Creaform Handyscan 700 portable laser scanner was used to obtain the digital model of the mask. The open-source tool Meshlab (Cignoni et al., 2008) was then utilized to process the model, removing outlier points and filling in the small holes. The high-

quality model was archived as a .obj format file. Figure 5.2a shows the Gaussian curvatures of all the vertices of the mask. It can be seen that the surface is heavily corroded, in particular, the surface is not smooth around the right cheek. The distortion and erosion may lead to an inaccurate restoration during the deformation using a TPS interpolation function. Because landmarks are used to guide the completion and shape analysis, we employed eighteen landmarks according to the commonly known landmarks in the literature (Li et al., 2017). Six landmarks were on the midline and twelve landmarks were bilateral (Table 5.1). Figure 5.2b shows the landmark configuration on the mask.

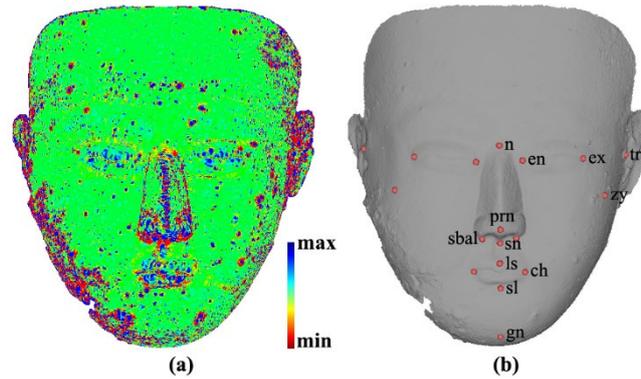


Figure 5.2 The digital mask as repaired through the reassembled fragments. (a) Visualisation of Gaussian curvatures of all the vertices. (b) The digital mask with eighteen landmarks labelled.

Table 5.1 Eighteen Landmarks

No.	Landmark definition	Position
1	Nasion (n)	midline
2	Pronasale (prn)	midline
3	Subnasale (sn)	midline
4	Labiale superius (ls)	midline
5	Sublabiale (sl)	midline
6	Gnathion (gn)	midline
7	Exocanthion (ex)	bilateral
8	Endocanthion (en)	bilateral
9	Subalare (sbal)	bilateral
10	Chelion (ch)	bilateral
11	Zygion (zy)	bilateral
12	Tragion (tr)	bilateral

5.3.2 Methods

Based on the landmarks, we proposed a computerized approach to repair the missing areas of the mask and employed AM and GM to capture the characteristic features (Figure 5.3). This process involves two stages. In the first stage, we approximated the missing

geometry of the damaged mask: (a) the plane of symmetry was estimated based on six landmarks on the midline and the reflection of the intact side was calculated; (b) based on the landmarks, the template model was produced using least squares and TPS methods, and the boundary curve of missing areas was predicted; (c) Laplacian deformation approach was used to deform the template model, minimizing the effect of heavy erosion and geometric distortion on the restoration; (d) a quantitative comparison of geometric deviation were used to validate the reliability of completion. In the second stage, we performed AM to analyse nine geometric measurements calculated by landmarks and employed landmark-based GM to visualize the geometry along the first two PCs in shape space. Matlab was used to code the algorithm to repair the missing geometry and perform geometric morphometric analysis. The software prototype can be downloaded from the Internet (https://github.com/sissun/Bilateral_Restoration.git).

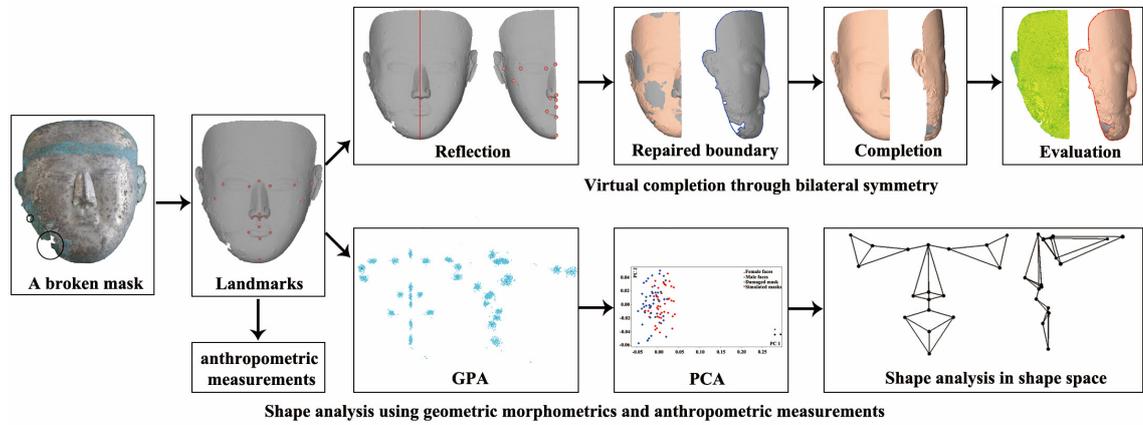


Figure 5.3 The pipeline of the proposed method.

5.3.2.1 The restoration of the damage mask

a) The plane of symmetry

Assuming that the equation of the symmetric plane was defined as $Ax + By + Cz + D = 0$ and the landmarks on the midline were $ml = \{ml_i\}, ml_i = (x_i, y_i, z_i) \in \mathbb{R}^3$, the least squares fitting method (Shui et al., 2015) was used to compute the coefficients of the symmetric plane by minimising the following equation

$$E = \min \sum_{i=0}^5 \|(Ax_i + By_i + Cz_i + D)\|^2 \quad (5.1)$$

where $A, B, C,$ and D denoted the coefficients of the plane.

Once the plane of symmetry was given, the left intact region $LM = \{lp_i\}, lp_i = (x_i, y_i, z_i) \in \mathbb{R}^3$ and the right broken side of the mask $RM = \{rp_i \in \mathbb{R}^3\}$ were separated. The reflection of the intact side $RLM = \{refp_i\}, refp_i = (refx_i, refy_i, refz_i) \in \mathbb{R}^3$ was generated by the following equation:

$$\begin{cases} refx_i = lx_i - 2At \\ refy_i = ly_i - 2Bt \\ refz_i = lz_i - 2Ct \end{cases} \quad (5.2)$$

$$\text{where } t = \frac{Ax_i + By_i + Cz_i + D}{A^2 + B^2 + C^2}.$$

The landmarks on the reflection of the intact region can also be calculated using Eq. 5.2. To validate the symmetric state of the mask, the Euclidean distance between each vertex of the mask and the plane of symmetry was calculated and visualized. Based on the symmetry of plane, the two new masks can be produced by reflecting the left and right sides, respectively.

b) The repaired boundary curve of the missing areas

The boundary curve of the mask consisted of two parts: the original curve and the break curve that was a newly generated curve when the artefact was broken. Hence, the first step of virtual restoration is to generate the template model and predict the boundary curve of the missing areas. It contained five steps:

Step 1: The boundary curves of the damaged mask and the reflection of the intact side were recognized based on the assumption that the one-ring neighbouring points of every vertex belonging to the boundary cannot form a closed loop. The boundary curve was equidistantly sampled and the corner points were identified based on the comparison of the angles of sampling points. In Figure 5.4a, the left figure shows sampling points (red points) and the angle θ of v_3 that was determined by $\overrightarrow{v_3v_2}$ and $\overrightarrow{v_3v_4}$. The right figure displays the identifications of the corner points (red points) based on the angles that compared to an appropriate threshold.

Step 2: The break curve was separated from the original curve in four steps. First, least squares algorithm based on singular value decomposition was used to rigidly register two meshes by minimizing the distance of labelled landmarks (Yu et al., 2019). Then, several extra points were automatically sampled on the damaged mask and the corresponding points were calculated on the reflection of the intact side by searching the nearest points. Subsequently, a TPS function was used to deform the reflection of the intact side to the damaged mask. Finally, the break curve was recognized through comparison of the Euclidean distance between the boundary curve of the damaged mask and the boundary curve of the resulting deformation. Once the distance was greater than the threshold, the vertex was suggested to belong to the break curve. In Figure 5.4b, the left figure displays the resulting deformation and the right figure illustrates the coarsely extracted break curve (yellow points).

Step 3: We extracted start and end points (fuchsia colour) from the corner points that covered the coarse break curve, as shown in Figure 5.4c. Then the improved break curve (blue points) and the original curve (red points) were separated.

Step 4: The sampling points located on the original curve only were chosen as the curve semilandmarks. We projected these curve semilandmarks on the boundary curve of the reflection of the intact side to identify the corresponding curve semilandmarks.

Step 5: Based on the landmarks and curve semilandmarks, a TPS function was used to deform the reflection of the intact side as the template model. Then the start and end points of the damaged mask could be projected on the boundary curve of the template model. The curve segments of the template model between start and end points were suggested to be the predicted missing boundary (Figure 5.4d).

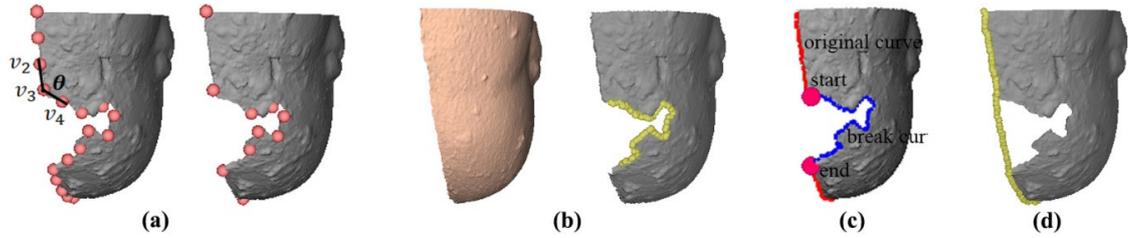


Figure 5.4 The repaired boundary curve of the missing area. (a) Corner points detection (red points) by means of computing the angle of every sampling vertex. (b) The extraction of the coarse break curve. The left figure displays the resulting deformation via TPS and the right figure illustrates the coarsely extracted break curve (yellow points). (c) The recognitions of start and end points (fuchsia colour), the break curve (blue points) and the original curve (red points). (d) The prediction of the boundary curve.

c) The deformation of the template model

Having formed the closed curve of the missing areas, Laplacian deformation was conducted to accomplish the completion (Sorkine et al., 2004). Assuming the template model can be represented as $TM=\{p_i\}, p_i \in \mathbb{R}^3$, the discrete Laplacian operator of p_i with the uniform weights can be as follows:

$$\delta(p_i)=p_i - \frac{1}{d_i} \sum_{q_j \in \xi_i} q_j \quad (5.3)$$

where ξ_i denoted the one-ring neighbouring vertices of p_i and d_i represented the number of vertices within ξ_i .

Geometric coordinates of every vertex on $TemMask$ can be calculated by minimizing a weighted error function E that consisted of a fitting term of boundary curve $E_{boundary}$, a fitting term of geometric correspondences $E_{closest}$, and a distortion term $E_{distortion}$.

$$E=E_{boundary} + \lambda E_{closest} + \gamma E_{distortion} \quad (5.4)$$

where λ and γ denoted weights that were used to balance the error function of every term.

$E_{boundary}$ was used as the hard constraint that allowed the boundary curve of the deformation to match the repaired boundary curve of the damaged artefact.

$$E_{boundary} = \sum_{i=1}^k \|u_i - v_i\|^2 \quad (5.5)$$

where v_i denoted the i -th vertex of the repaired and original boundary curves of the damaged mask and u_i represented the i -th corresponding point by searching the nearest point on the boundary curve of TM; k denoted the number of the corresponding vertices of boundary curves.

$E_{closest}$ expressed the distance between geometric correspondences between the template model and the damaged mask.

$$E_{closest} = \sum_{i=1}^m \|p_i - q_i\|^2 \quad (5.6)$$

where p_i denoted the remaining vertices of TM except for the boundary curve and q_i represented the corresponding vertices of the damaged mask by searching the nearest points. The distance between two points can be used to evaluate the reliability of correspondences. m denoted the number of geometric correspondences.

$E_{distortion}$ quantified the changes of $TemMask$ between before and after deformations.

$$E_{distortion} = \sum_{i=1}^t \|\delta(p_i') - \delta(p_i)\|^2 \quad (5.7)$$

where p_i' denoted every vertex after deformation, and $\delta(p_i')$ represented Laplacian coordinates using Eq. 5.3, and t denoted the number of vertices of $TemMask$.

During the optimization of the error function, we increased λ and decreased γ to allow the template model to match the damaged mask tightly. After the deformation was performed, we extracted the digital model of the missing geometry. Then, 3D printing technology was applied to convert the digital model to the physical model. Having considered the depth of the mask and the instrumental precision of 3D printer, an industrial resin 3D printer with a high precision, iSLA 880 produced by ZRapid Tech, China, was used. It is a type of Stereolithography that utilized an ultraviolet laser to draw a pre-programmed design and harden the liquid material to form a layer (Zhang et al., 2016). This process is repeated until every layer has been printed.

d) The evaluation of the reliability

To verify whether the resulting deformation can be used to fill in the gaps, we compared the geometric differences of surfaces and boundaries. The geometric deviation (*error*) between the damaged mask and the resulting deformation were quantitatively calculated as follows:

$$error = \frac{1}{2} error_b + \frac{1}{2} error_d \quad (5.8)$$

Where $error_b$ denoted the average fitting discrepancy of every vertex of the boundary curve and $error_d$ represented average distortion discrepancy of every corresponding vertex.

5.3.2.2 Shape analysis of the mask

It is an interesting question to capture the characteristic features of the mask and investigate the geometric variations between the mask and human faces. To capture the features of the mask, the dataset consisted of two groups: a) we selected 100 human faces (50 males and 50 females) from a dataset of 140 individuals, which was reconstructed using CT images acquired by a clinical multi-slice CT scanner system. More details of the dataset and the procedure of reconstruction 3D meshes were provided in our previous study (Shui et al., 2017). b) the actual damaged mask and two simulated masks via reflection. Then, we employed GM and AM to capture the shape features based on the Cartesian coordinates of eighteen landmarks.

a) Shape analysis using GM

To minimise the effects of translation, rotation, and scaling on the shape analysis, GPA was used to superimpose all the eighteen landmarks of human faces and the masks by minimising the overall sum of squares of the landmark configurations with regard to the Procrustes mean shape (Mitteroecker & Gunz, 2009). The procedure was as follows: first, an arbitrary specimen within the dataset was selected as the template model, and the other specimens were then aligned to the template model using Procrustes superimposition. The centroid size (CS) of each specimen, defined as the square root of the sum of squares Euclidean distances between every landmark and centroid, was calculated. Second, the Procrustes mean shape consisting of landmarks was calculated by averaging the Procrustes shape coordinates of all the specimens, and each specimen was then aligned to the Procrustes mean shape using Procrustes superimposition. This process was repeated through several iterations, i.e., the Procrustes mean shape was repeatedly calculated until the change in the Procrustes mean shape was less than a predetermined threshold or a maximum number of iterations. Finally, the Procrustes shape coordinates of every specimen were represented as discrete points in a hyper hemisphere, which approximated a Riemannian manifold, called Kendall's shape space (Slice, 2001).

We applied PCA to convert the landmark configuration of all the specimens to a tangent shape space. The landmark configuration of each specimen was represented by the average landmark configuration and a linear combination of principal component (PC) scores and the corresponding orthogonal PCs. Because all the PCs described the independent modes of shape variations in shape space, the visualisation of PCs of interest can further be conducted to interpret and explain the shape variability (O'Higgins, 2000; Shui et al., 2017). Student's *t*-test was performed on the PC scores to verify the significance level between the PC scores of human faces and the masks.

b) Shape analysis using AM

Craniofacial anthropometry is a conventional approach for morphological shape

analysis and identification in the realm of anthropology. Based on the given eighteen landmarks, we calculated nine anthropometric measurements of human faces and the masks, which included tr-tr, zy-zy, en-ex, n-prn, sbal-sbal, prn-sn, ch-ch, ls-sl and n-gn. Then, statistical analysis was performed and Student's t-test was carried out to assess the differences of the measured variables.

5.4 Results

5.4.1 The restoration of the damaged mask

To repair the broken mask, the first step was to estimate the plane of symmetry and recreate the reflection of the intact region. Based on six landmarks on the midline, we calculated the symmetric plane (red colour) of the mask (Figure 5.5a). Figure 5.5b shows the Euclidean distance (mm) from every vertex of the mask to the plane of symmetry. It indicated that the mask was asymmetric, particularly around the ears. Subsequently, the intact and damaged sides were separated and the reflection of the intact side was generated. Figure 5.5c shows the reflection of the intact side (grey colour) and the damaged side (peach colour) in the common coordinate system. Based on the plane of symmetry, two digital masks can be simulated by means of reflecting the digital models of the intact and damaged sides, respectively (Figures 5.5d and e).

In addition, we compared the plane of symmetry with other methods. Figure 5.5f shows the symmetric plane (red colour) through the identification of the principal axes using PCA, which is greatly dependent on the distribution of the vertices of the mask (Koutsoudis et al., 2010). Due to the existence of the missing areas and shape distortion, the plane of symmetry is not accurate. The symmetric plane was extracted by the calculation of the intrinsic wave kernel signature (WKS) (Aubry et al., 2011), as shown in Figure 5.5g. The left figure displayed the WKS descriptor of every vertex. The middle figure showed the symmetric region whose WKS values were relatively minimum. The right figure illustrated the extract symmetric plane by means of identification of the principal direction (red colour). The estimated plane was relatively closer to the plane that was recognized based on landmarks.

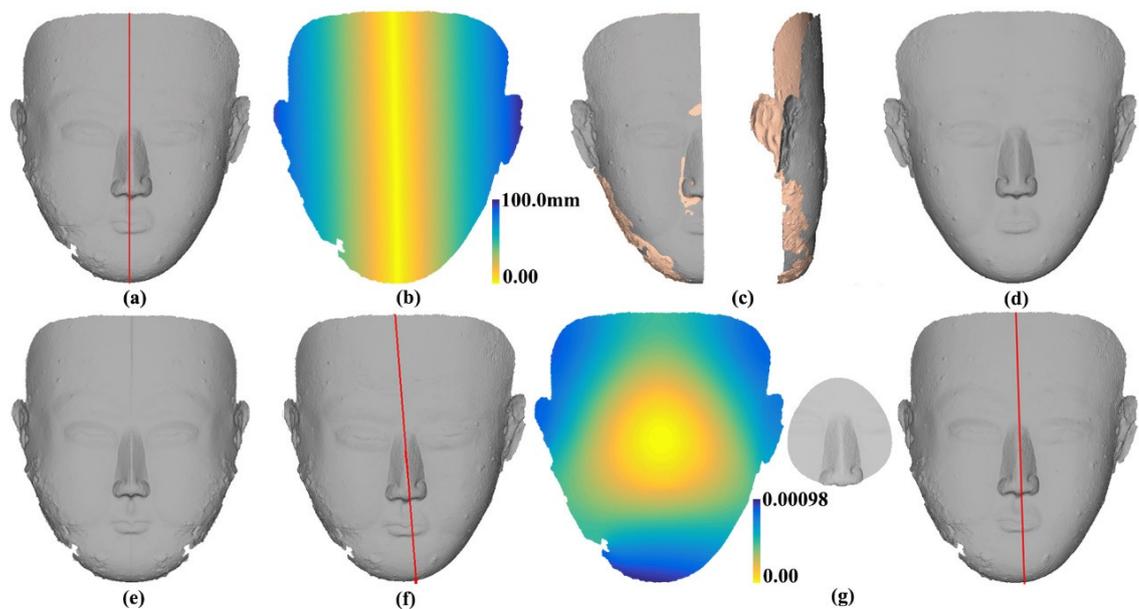


Figure 5.5 The reflection structure of the intact side of the mask. (a) The plane of symmetry (red colour) using six landmarks on the midline. (b) The Euclidean distance between every vertex and the plane of symmetry. (c) The damaged side (peach colour) and the reflection of the intact side (grey colour) in the common coordinate system. (d) The simulated digital mask by reflecting the surface of the intact side. (e) The simulated digital mask by reflecting the surface of the damaged side. (f) The estimated plane using the principal axis. (g) The estimated plane using the WKS descriptor and the principal axis. The left figure displayed the WKS descriptor of every vertex, and the middle figure showed the region that existed symmetry and the right figure illustrated the estimated symmetric plane (red colour).

We employed a hybrid non-rigid deformation approach to recreate the template model and accomplish the completion of the missing areas. Figure 5.6a shows the damaged side of the mask (peach colour) and the registered reflection of the intact side (grey colour) using the least squares method. The geometric deviation between these two meshes was visualized in Figure 5.6b, where the error was 3.2 mm and $error_b$ was 4.07 mm. Figure 5.6c illustrated 43 curve semilandmarks of the damaged mask (grey colour) and those of the reflection of the intact side (peach colour). Based on landmarks and semilandmarks, a TPS function was used to deform the reflection of the intact side to generate the template model. Figure 5.6d shows the template model and the geometric difference between the template model and the damaged side, where the $error$ was 0.66 mm and $error_b$ was 0.55 mm. The areas whose deviations were the greatest can be observed around the missing areas. Figure 5.6e shows the repaired boundary curve (blue points) that was used as the hard constraint in the further deformation. Figure 5.6f shows the resulting deformation via Laplacian deformation. The left two figures displayed the front and profile views of the resulting deformation and the right figure illustrated the geometric deviation between the resulting deformation and the damaged side. The $error$ was 0.21 mm and $error_b$ was 0.27 mm. It indicated the deformed template model fitted the damaged mask well and the boundary curve was almost consistent with that of the damaged mask. Thus, we can extract the digital model from the resulting deformation and it can be tightly embedded in the missing areas.

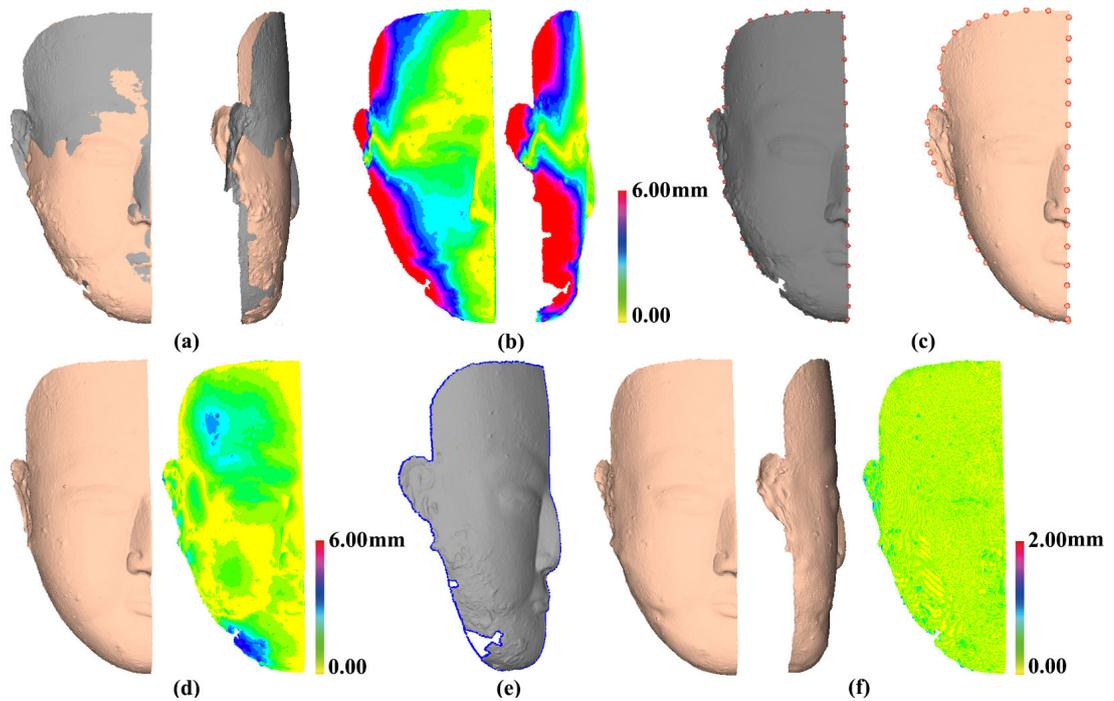


Figure 5.6 The deformation of the template model using our method. (a) the registered template model (grey colour) via least squares method and the damaged mask (peach colour). (b) The visualisation of geometric deviation between these two meshes. (c) Curve semilandmarks (red points) of the damaged mask (left figure) and the template model (right figure). (d) The resulting deformation via TPS based on landmarks and curve semilandmarks (left figure) and the visualisation of geometric deviation (right figure). (e) The repaired boundary curve and the original curve. (f) The deformed template model via Laplacian deformation (left and middle figures) and the visualisation of geometric deviation (right figure).

We extracted the missing geometry from the resulting deformation. Figure 5.7a shows the frontal and profile views of the damaged mask (peach colour) and the resulting deformation (grey colour). Figure 5.7b illustrates the repaired missing geometry is adhered to the damaged mask. We solidified the surfaces of the missing geometry and then utilized a 3D printer to produce the physical model. Figure 5.7c shows the replicas of the damaged mask and the missing geometry that can be reassembled together. In addition, the proposed method was compared with other non-rigid deformation approaches. Figure 5.7d shows the deformed template model using Laplacian deformation method (Sorkine et al., 2004), but the predicted boundary curve was not used. The deformed template model tightly matched the damaged mask; however, the boundary curve of the missing areas was inconsistent with that of the damaged mask. Figure 5.7e shows the deformed template model via the NICP method (Amberg et al., 2007). The majority of the resulting deformation closely matched the damaged mask; however, great variation was located around the ear and the boundary curve of the resulting deformation (grey colour) was inconsistent with that of the damaged mask (peach colour). Figure 5.7f shows the deformed template model via the CPD method (Myronenko & Song, 2010). It can be seen that the deformed template model (grey colour) and the damaged mask (peach colour) have not overlapped and the boundary curve of the deformed template model was inconsistent with that of the damaged mask.

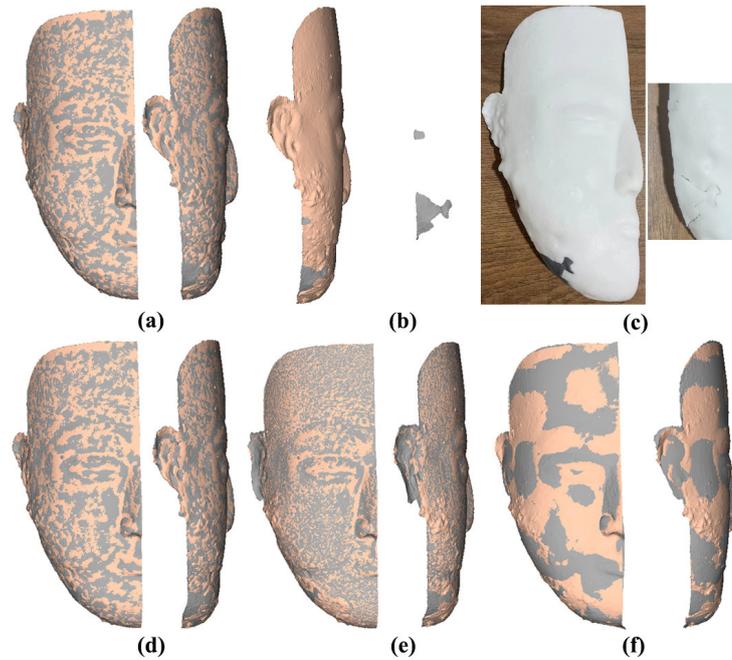


Figure 5.7 The evaluation of the reliability of the restoration. (a) The frontal and profile views of the damaged mask (peach colour) and the resulting deformation (grey colour) using the proposed method. (b) The digital surfaces of the predicted missing areas and the damaged mask. (c) The replicas of the repaired missing geometry and the damaged mask. (d) The resulting deformation via the Laplacian deformation without the predicted boundary curve as constraint. (e) The resulting deformation via the NICP method. (f) The resulting deformation via the CPD method.

We also examined the effects of the weighting coefficients λ and γ on the deformation by fixing one coefficient and adjusting the other. Geometric deviations between the deformed template models using different weights and the damaged mask were listed in Table 5.2. As shown in Figure 5.8, the left figure displayed the resulting deformation, and the middle figure showed the resulting deformation and the damaged mask were located together, and the right figure illustrated the geometric deviation between two meshes. Figure 5.8a shows the resulting deformation when λ was set 0.0 and γ was set 10.0 and Figure 5.8b shows the resulting deformation when λ was set 1.0 and γ was set 10.0. It is of note that when all the reliable correspondences between two meshes were used, the resulting deformation was closer to the damaged mask. Figure 5.8c displays the resulting deformation when λ was set 1.0 and λ was set 0.01 and Figure 5.8d displays the resulting deformation when λ was set 1.0 and γ was set 10000. The resulting deformation was much closer to the damaged mask when γ approached to zero, and the resulting deformation approximated the template model when γ became greater.

Table 5.2 Geometric deviation between the damaged mask and the resulting deformation using different coefficients

The weigh coefficients	$error_b$ (mm)	$error_d$ (mm)	$error$ (mm)
$\lambda=0.0$ and $\gamma=10.0$	0.31	0.78	0.55
$\lambda=1.0$ and $\gamma=10.0$	0.39	0.13	0.26
$\lambda=1.0$ and $\gamma=0.01$	0.21	0.05	0.13
$\lambda=1.0$ and $\gamma=10000$	0.50	0.85	0.68

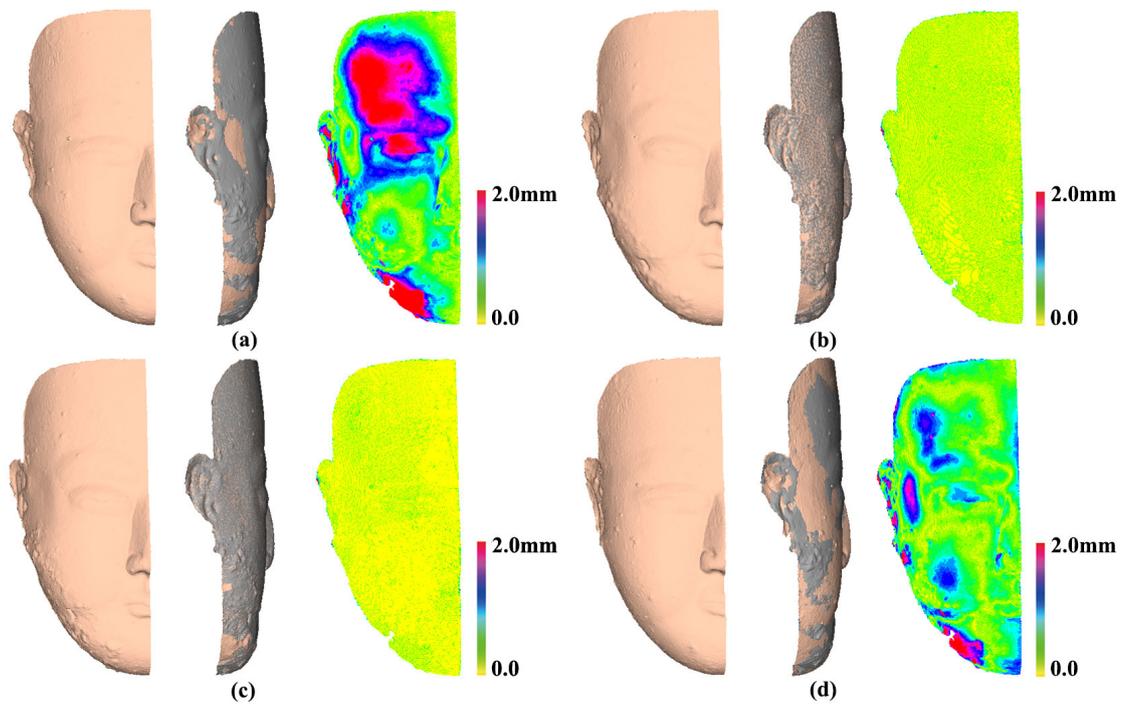


Figure 5.8 The effect of every weight on the deformation. The left figure displayed the resulting deformation, and the middle figure showed the resulting deformation and the damaged mask were located together, and the right figure illustrated the geometric deviation between two meshes. (a) λ was set 0.0 and γ was set 10.0. (b) λ was set 1.0 and γ was set 10.0. (c) λ was set 1.0 and γ was set 0.01. (d) λ was set 1.0 and γ was set 10000.

We used synthetic data that almost exhibited bilateral symmetry to test the reliability of the proposed method. To remove the partial geometry, the open-source software MeshLab (Cignoni et al., 2008) was applied. Figure 5.9a displays the restoration of a terracotta warrior head, where we removed the right cheek manually. This artefact was captured by an Artec spider scanner and provided by Beijing Key Laboratory of Digital Preservation and Virtual Reality for Cultural Heritage, Beijing Normal University. Figure 5.9b shows the restoration of a Qihe human cranium, where we removed the left eyebrow and forehead. This fossil was scanned using an industrial computerized tomography scanner at the Institute of Vertebrate Paleontology and Paleoanthropology, Chinese Academy of Sciences. From left to right, each figure illustrates the actual models, the damaged models, the intact and damaged sides of the models, the resulting deformation of the intact side, the predicted missing geometry (peach colour) and the damaged models (grey colour), the predicted geometry missing and the visualisation of geometric deviation between the predicted and the actual models. The average geometric value of the terracotta warrior head was 1.12 mm and the average geometric value of cranium was 0.71 mm. These results demonstrate that the repaired models adhere closely to the damaged artefacts.

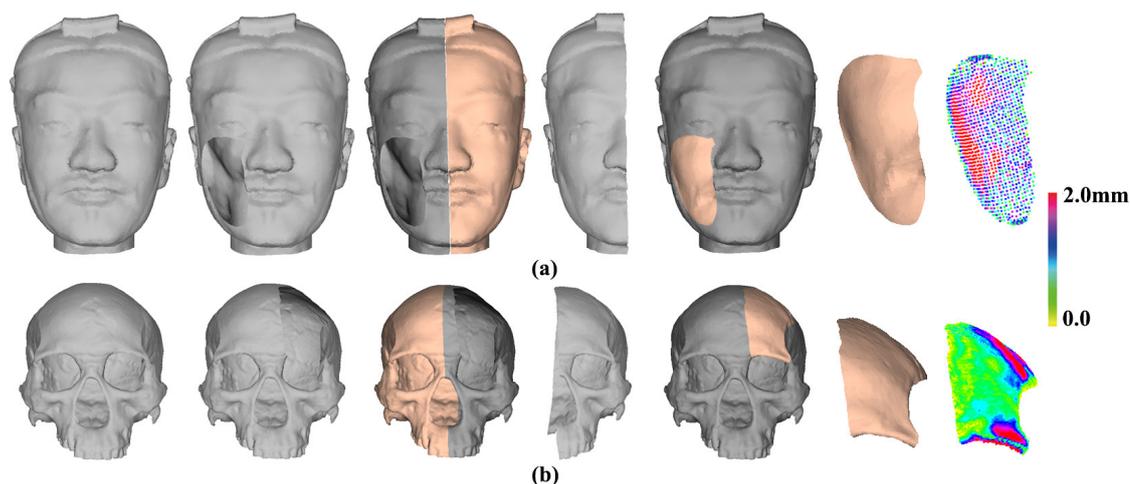


Figure 5.9 The restoration of synthetic data. From left to right, each figure illustrated the actual models, the damaged artefacts that removed partial regions manually, the intact and damaged sides of the models, the resulting deformation of the intact side, the predicted geometry missing (peach colour) and the damaged models (grey colour), the predicted geometry missing and the visualisation of geometric deviation between the predicted and the actual models. (a) A terracotta warrior head. (b) An ancient human cranium.

5.4.2 Shape analysis of the mask

A total of 30 PCs accounted for almost 98% of the total morphological variance. The first two PCs accounted for 28.05% and 11.78% of the total variance, respectively. Scatterplots of PC 1 and PC 2 of human faces and three masks were depicted in Figure 5.10a. The masks appeared to be located along the extreme positive PC 1, which provided a clear impression that the masks and human faces were greatly different in shape space. To identify the main patterns of shape variance along PC 1 and PC 2, four new shapes were produced along the extreme positive and negative PC 1 and PC 2, respectively (Figure 5.10b). The positive PC 1 (PC 1+) connected with an upper wide face that has a long, narrow and flat nose, broad eyes, and a narrow, short and small mouth. In contrast, the negative PC 1 (PC 1-) represented a narrow upper face with a short, wide and protruding nose, narrow eyes, and a wide and long mouth. Student's *t*-test was applied to the PC scores of two groups. There was a highly significant difference on PC 1 (p -value < 0.005). In contrast, the positive PC 2 (PC 2+) was related to a short nose and a short mouth. The negative PC 2 (PC 2-) represented a long nose and a long mouth. It is of note that PC 1 seems to be the inverse of the main patterns of morphological variation in the nose and mouth along PC 2. According to the analysis and the visualisation of the shape changes on PC 1 and PC 2, the mask has a broader upper face, broader eyes, a longer, narrower and flatter nose, and a shorter, narrower and smaller mouth than the human faces.

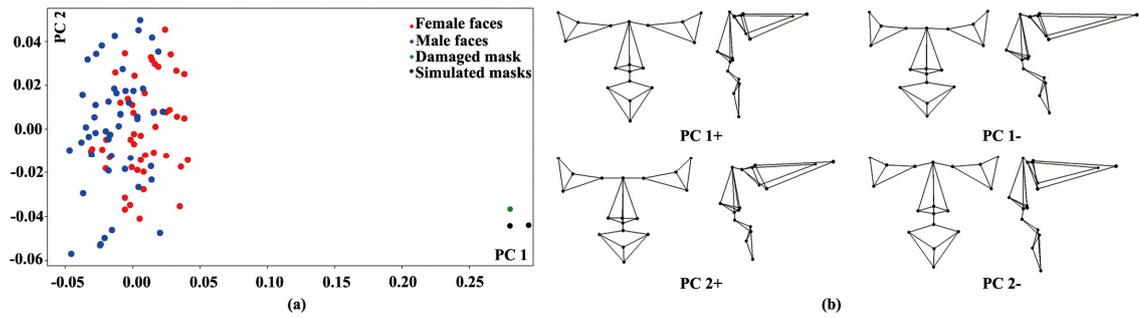


Figure 5.10 Shape variations between human faces and the masks in shape space. (a) Scatterplots of PC 1 and PC 2 of the human faces and three masks. (b) The visualisation of shape variabilities along the extreme positive and negative PC 1 and PC 2.

We employed AM to capture the main features of the mask. The average values of nine anthropometric measurements were calculated (Table 5.3). There are significant differences among the six anthropometric measurements, which represented the width of the head (tr-tr), the length (n-prn), width (sbal-sbal) and depth (prn-sn) of the nose, the length (ls-sl) and width (ch-ch) of the mouth. In addition, the width of the face (zy-zy) and the eyes (en-ex) of the masks were greater than the average values for human faces. These findings were consistent with those obtained using GM.

Table 5.3 The comparison of anthropometric measurements

Geometric Measurements	The average value of human faces (mm)	The actual Mask (mm)	The average value of three masks (mm)
tr-tr	150.40	166.97	168.42**
zy-zy	126.28	129.54	129.19
en-ex	31.65	37.77	35.63
n-prn	42.75	51.16	51.16**
sbal-sbal	30.36	20.86	20.55**
prn-sn	18.38	12.88	12.85**
ch-ch	52.31	31.13	30.55**
ls-sl	28.75	16.21	16.21**
n-gn	109.27	118.59	118.59*

* p -value < 0.05, ** p -value < 0.005

5.5 Discussion

The reconstruction of bilaterally symmetrical artefacts is an important task for conservators in order to maintain and restore damaged artefacts. It is a complex field that involves not only the technologies of restoration and shape analysis, but also historic concerns, culture, aesthetics, arts, education, etc. (Fowles et al., 2003; Qiao et al., 2020). Due to its vulnerable conservation status, it is necessary to repair the damaged mask featured in our case study in order to describe the aesthetic value that raises public interest in learning history, and disseminating and communicating cultural heritage. Therefore, this study presented a hybrid non-rigid deformation approach to repair the missing areas, minimizing human intervention.

When artefacts need to be repaired, conservators must respect the appearance, form, design, material, function, traditions and techniques of the past. Historically, several policies have proposed the fundamental principles to preserve authenticity and discernibility in the restoration of cultural heritage. In 1964, the Venice Charter pointed out that “It is our duty to hand them on in the full richness of their authenticity”. To avoid misunderstanding, the replacement of missing parts must be different from the original artefact. In the digital era, these principles still require compliance. Both the London Charter and Seville principles highlight the need for authenticity in scientific visualisation and virtual restorations (Lopez-Menchero & Grande, 2011; Vico, 2018). But different countries have different interpretations and principles, and such different understandings have led to different approaches to restoration. For example, “Restoring the old as it was” is the core principle in China (Zhu, 2017). As a consequence, conservators attempt to recover the geometry and texture of the original appearance so that the general public have a complete appreciation of past artefacts from the perspectives of aesthetics and beauty. Under these circumstances, the original artefact might be manipulated or corrected manually, and the replacements are not always distinguished. Virtual restoration and 3D printing technologies provide a promising way to integrate aesthetics, beauty and authenticity during the restoration and minimize the physical contact with the artefact. Geometric documentation that consists of digital models, texture, metadata and paradata records information about artefacts before and after restoration authentically and permanently. In addition, the physical model of the missing geometry through the use of 3D printing can be used to repair the missing geometry and can also be recognized easily. It is of note that the authenticity of 3D printing material is a challenge during the restoration. An improved physical model of the missing geometry would be reproduced when new scientific technologies and authorized 3D printing materials are used.

Both the production of the template model and the performance of the deformation greatly impact on the reliability of the virtual restoration of damaged artefacts. It is note of that the plane of symmetry influences the generation of the template model. To improve the reliability, the least squares approach based on the landmarks is used, instead the automatic extraction approaches (Li et al., 2011; Sipiran et al., 2014). When there is no other historical evidence, it is acknowledged that template model provides a considerable reference and guidance to repair the damaged artefact. The closer the resulting deformation of the model and the damaged artefact match, the more confidence there is in the reliability of the repaired model. In general, either the reflection of the retained geometry or the geometric-similar model is chosen as the template model based on whether more than half the original artefact survives (Jo et al., 2020; Shui et al., 2020). Because only small regions of the mask have not survived, the resulting deformation of the retained side via least squares and TPS methods is chosen as the template model that is relatively similar to the damaged side. It remains challenging to allow the resulting

deformation and the unsmooth and distorted surface to match tightly, particularly around the boundary curves. In our previous work, we had put great effort to employ 3D interactive graphical software to recreate and modify the digital model (Qu et al., 2018), but it is a tedious and time-consuming task and the repaired missing geometry might still not tightly match the damaged artefact. Using the repaired boundary curve as a hard constraint and the weighting coefficients to balance every term of the error function, we improved the resulting deformation more successfully than other approaches (Amberg et al., 2007; Deng et al., 2011; Myronenko & Song, 2010; Sorkine et al., 2004). But one of the limitations is that the details of the template model will transfer to the repaired artefact during the deformation. In addition, when the artefact is asymmetric or the missing area is greater than half, the regression based on statistical shape model (SSM) is a promising approach based on the establishment of the dense correspondences of artefacts within the same group (O'Higgins et al., 2019). Then the best-fitting parameters of the damaged artefact are used to estimate the missing geometry in shape space.

Since more software and packages, e.g. R packages (Adams & Otárola-Castillo, 2013), etc., can be freely used to perform shape analysis of 3D models, both AM and GM have been widely used to capture the characteristic features in biology. It is of note that landmarks play a central role that needs to be carefully defined in relation to a specific question. In this study, we used only landmarks to ensure the reliability of correspondences among different groups. Because no other metal masks of the Liao dynasty can be collected, we had to use both the actual mask and two simulated masks as a group to investigate morphological variations. This study offers a promising quantitative analysis and visualisation approach to describe the geometric shape of the artefact. We believe that a large dataset of the metal masks of the Liao dynasty will help archaeologists and general public to appreciate the main features of the masks and their cultural value.

5.6 Conclusion

In this study, we presented a computerized hybrid non-rigid deformation approach to accomplish the completion of a bilaterally symmetrical metal mask with heavy erosion. Our method integrated the estimation of the plane of symmetry, the production of the template model, and the deformation of the template model. The quantitative comparison of the resulting deformation and the damaged mask was used to validate the reliability of the restoration. The experimental results demonstrated that the proposed method improved the accuracy of the geometric completion. In addition, we used AM and GM to characterize the main patterns of the mask. The mask has a wider upper face, broader eyes, a longer, narrower and flatter nose, and a shorter, narrower and smaller mouth. Our method can further be applied to other archaeological and anthropological data that exhibit near bilateral symmetry. However, this study requires the placements of

landmarks to guide the virtual restoration and shape analysis. In future work we plan to develop an automatic method for identifying landmarks and geometric features to minimize human intervention.

5.7 Acknowledgements

I express my gratitude to Beijing Normal University for granting access to the terracotta warrior head. Additionally, I would like to thank Institute of Vertebrate Paleontology and Paleoanthropology for their provision of access to the Qihe cranium.

5.8 References

- Adams, DC, & Otárola-Castillo, E. (2013). geomorph: an R package for the collection and analysis of geometric morphometric shape data. *Methods in Ecology and Evolution*, 4(4), 393-399.
- Amberg, B, Romdhani, S, & Vetter, T. (2007). Optimal step nonrigid ICP algorithms for surface registration. The *IEEE Conference on Computer Vision and Pattern Recognition (CVPR'07)*, 1-8.
- Arbace, L, Sonnino, E, Callieri, M, Dellepiane, M, Fabbri, M, Idelson, AI, & Scopigno, R. (2013). Innovative uses of 3D digital technologies to assist the restoration of a fragmented terracotta statue. *Journal of Cultural Heritage*, 14(4), 332-345.
- Aubry, M, Schlickewei, U, & Cremers, D. (2011). The wave kernel signature: A quantum mechanical approach to shape analysis. The *2011 IEEE international conference on computer vision workshops*, 1626-1633.
- Bookstein, FL. (1989). Principal warps: Thin-plate splines and the decomposition of deformations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(6), 567-585.
- Brande, S, & Saragusti, I. (1996). A morphometric model and landmark analysis of acheulian hand axes from Northern Israel. In *Advances in morphometrics*, 423-435: Springer.
- Cignoni, P, Callieri, M, Corsini, M, Dellepiane, M, Ganovelli, F, & Ranzuglia, G. (2008). MeshLab: an Open-Source Mesh Processing Tool. The *Eurographics Italian Chapter Conference*, 129-136.
- Dai, H, Pears, N, Smith, W, & Duncan, C. (2020). Statistical Modeling of Craniofacial Shape and Texture. *International Journal of Computer Vision*, 128(2), 547-571.
- Deng, Q, Zhou, M, Shui, W, Wu, Z, Ji, Y, & Bai, R. (2011). A novel skull registration based on global and local deformations for craniofacial reconstruction. *Forensic Science International*, 208(1), 95-102.
- Fowles, PS, Larson, JH, Dean, C, & Solajic, M. (2003). The laser recording and virtual restoration of a wooden sculpture of Buddha. *Journal of cultural heritage*, 4, 367-371.
- Gero, J, & Mazzullo, J. (1984). Analysis of artifact shape using Fourier series in closed form. *Journal of Field Archaeology*, 11(3), 315-322.
- Gilboa, A, Tal, A, Shimshoni, I, & Kolomenkin, M. (2013). Computer-based, automatic recording and illustration of complex archaeological artifacts. *Journal of Archaeological Science*, 40(2), 1329-1339.
- Gunz, P, Mitteroecker, P, Neubauer, S, Weber, GW, & Bookstein, FL. (2009). Principles

for the virtual reconstruction of hominin crania. *Journal of Human Evolution*, 57(1), 48-62.

Jo, YH, Hong, S, Jo, SY, & Kwon, YM. (2020). Noncontact restoration of missing parts of stone Buddha statue based on three-dimensional virtual modeling and assembly simulation. *Heritage Science*, 8, 103.

Koutsoudis, A, Pavlidis, G, Liami, V, Tsiafakis, D, & Chamzas, C. (2010). 3D pottery content-based retrieval based on pose normalisation and segmentation. *Journal of Cultural Heritage*, 11(3), 329-338.

Kuzminsky, SC, & Gardiner, MS. (2012). Three-dimensional laser scanning: potential uses for museum conservation and scientific research. *Journal of Archaeological Science*, 39(8), 2744-2751.

Li, M, Cole, JB, Manyama, M, Larson, JR, Liberton, DK, Riccardi, SL, Ferrara, TM, et al. (2017). Rapid automated landmarking for morphometric analysis of three-dimensional facial scans. *Journal of Anatomy*, 230(4), 607-618.

Li, X, Yin, Z, Wei, L, Wan, S, Yu, W, & Li, M. (2011). Symmetry and template guided completion of damaged skulls. *Computers & Graphics*, 35(4), 885-893.

Lopez-Menchero, VM, & Grande, A. (2011). The principles of the Seville Charter. *The CIPA symposium proceedings*, 2-6.

Mayberger, E. (2020). Flaming Pearls and Flying Phoenixes: Materiality, Research, and Stewardship of Liao Dynasty Metalwork. *Journal of the American Institute for Conservation*, 59(1), 65-76.

Mitteroecker, P, & Gunz, P. (2009). Advances in geometric morphometrics. *Evolutionary Biology*, 36(2), 235-247.

Munn, L, & Stephan, CN. (2018). Changes in face topography from supine-to-upright position—and soft tissue correction values for craniofacial identification. *Forensic Science International*, 289, 40-50.

Myronenko, A, & Song, X. (2010). Point set registration: Coherent point drift. *IEEE transactions on Pattern Analysis and Machine Intelligence*, 32(12), 2262-2275.

O'Higgins, P. (2000). The study of morphological variation in the hominid fossil record: biology, landmarks and geometry. *Journal of Anatomy*, 197(1), 103-120.

O'Higgins, P, Fitton, LC, & Godinho, RM. (2019). Geometric morphometrics and finite elements analysis: Assessing the functional implications of differences in craniofacial form in the hominin fossil record. *Journal of Archaeological Science*, 101, 159-168.

O'Higgins, P, & Jones, N. (1998). Facial growth in *Cercocebus torquatus*: an application of three-dimensional geometric morphometric techniques to the study of morphological

variation. *Journal of Anatomy*, 193(2), 251-272.

Okumura, M, & Araujo, AG. (2014). Long-term cultural stability in hunter–gatherers: a case study using traditional and geometric morphometric analysis of lithic stemmed bifacial points from Southern Brazil. *Journal of Archaeological Science*, 45, 59-71.

Okumura, M, & Araujo, AG. (2019). Archaeology, biology, and borrowing: A critical examination of Geometric Morphometrics in Archaeology. *Journal of Archaeological Science*, 101, 149-158.

Oxnard, C, & O’Higgins, P. (2009). Biology clearly needs morphometrics. Does morphometrics need biology? *Biological Theory*, 4(1), 84-97.

Papaioannou, G, Schreck, T, Andreadis, A, Mavridis, P, Gregor, R, Sipiran, I, & Vardis, K. (2017). From reassembly to object completion: A complete systems pipeline. *Journal on Computing and Cultural Heritage*, 10(2), 1-22.

Peng, S. (2003). An archaeological view on the burial customs of the Khitan aristocrats in the Liao dynasty (in Chinese). *Research of China's Frontier Archaeology*, 2, 298-308.

Qiao, C, Zhang, W, Gong, D, & Gong, Y. (2020). In situ virtual restoration of artifacts by imaging technology. *Heritage Science*, 8, 110.

Qu, L, Gao, F, Liu, J, He, L, Lei, Y, & Shui, W. (2018). Museum as an Example Study of Metal Artefacts Reparation Using Modern Technology and Traditional Technology: Taking Liao Dynasty Metal Mask in the Palace (in Chinese). *Museum*, 8(2), 119-127.

Sapirstein, P. (2018). A high-precision photogrammetric recording system for small artifacts. *Journal of Cultural Heritage*, 31, 33-45.

Schlager, S, Profico, A, Di Vincenzo, F, & Manzi, G. (2018). Retrodeformation of fossil specimens based on 3D bilateral semi-landmarks: Implementation in the R package “Morpho”. *PloS one*, 13(3), e0194073.

Shimada, M. (1950). A Death-Mask of the Liao Period. *Artibus Asiae*, 13(4), 250-253.

Shott, MJ, & Trail, BW. (2010). Exploring new approaches to lithic analysis: laser scanning and geometric morphometrics. *Lithic Technology*, 35(2), 195-220.

Shui, W, Zhou, M, Gao, F, & Qu, L. (2015). A digital restoration method for broken bronze artifacts having a symmetrical structure (in Chinese). *Sciences of conservation and archaeology*, 27(4), 113-117.

Shui, W, Zhou, M, Maddock, S, He, T, Wang, X, & Deng, Q. (2017). A PCA-Based method for determining craniofacial relationship and sexual dimorphism of facial shapes. *Computers in Biology and Medicine*, 90, 33-49.

Shui, W, Zhou, M, Maddock, S, Ji, Y, Deng, Q, Li, K, Fan, Y, et al. (2020). A computerized

craniofacial reconstruction method for an unidentified skull based on statistical shape models. *Multimedia Tools and Applications*, 79, 25589-25611.

Sipiran, I, Gregor, R, & Schreck, T. (2014). Approximate symmetry detection in partial 3d meshes. *Computer Graphics Forum*, 33(7), 131-140.

Slice, DE. (2001). Landmark coordinates aligned by Procrustes analysis do not lie in Kendall's shape space. *Systematic Biology*, 50(1), 141-149.

Sorkine, O, Cohen-Or, D, Lipman, Y, Alexa, M, Rössl, C, & Seidel, H-P. (2004). Laplacian surface editing. The *The 2004 Eurographics/ACM SIGGRAPH symposium on Geometry processing*, 175-184.

Vico, L. (2018). Authenticity and realism: virtual vs physical restoration. In *In Authenticity and Cultural Heritage in the Age of 3D Digital Reproductions*, 25-33: McDonald Institute: Cambridge, UK.

Yu, J, Lin, Y, Wang, B, Ye, Q, & Cai, J. (2019). An Advanced Outlier Detected Total Least-Squares Algorithm for 3-D Point Clouds Registration. *IEEE Transactions on Geoscience and Remote Sensing*, 57(7), 4789-4798.

Zhang, L, Dong, H, & Saddik, AE. (2016). From 3D sensing to printing: a survey. *ACM Transactions on Multimedia Computing, Communications, and Applications*, 12(2), 27.

Zhang, X, Blaas, J, Botha, C, Reischig, P, Bravin, A, & Dik, J. (2012). Process for the 3D virtual reconstruction of a microcultural heritage artifact obtained by synchrotron radiation CT technology using open source and free software. *Journal of Cultural Heritage*, 13(2), 221-225.

Zhu, Y. (2017). Authenticity and heritage conservation in China: Translation, interpretation, practices. In *Authenticity in architectural heritage conservation*, 187-200: Springer.

Chapter 6 A computerized facial approximation method for archaic humans based on dense facial soft tissue thickness depths

Abstract: Facial approximation (FA) is a common tool used to recreate the possible facial appearance of a deceased person based on the relationship between soft tissue and the skull. Although this technique has been primarily applied to modern humans in the realm of forensic science and archaeology, only a few studies have attempted to produce FAs for archaic humans. This study presented a computerized FA approach for archaic humans based on the assumption that the facial soft tissue thickness depths (FSTDs) of modern living humans are similar to those of archaic humans. Additionally, we employed geometric morphometrics (GM) to examine the geometric morphological variations between the approximated faces and modern human faces. Our method has been applied to the Jinniushan (JNS) 1 archaic human, which is one of the most important fossils of the Middle Pleistocene, dating back to approximately 260,000 BP. The overall shape of the approximated face has a relatively lower forehead and robust eyebrows; a protruding, wider, and elongated middle and upper face; and a broad and short nose. Results also indicate skull morphology and the distribution of FSTDs influence the approximated face. These experiments demonstrate that the proposed method can approximate a plausible and reproducible face of an archaic human.

Article details:

Wuyang Shui, Yameng Zhang, Xiujie Wu and Minquan Zhou. A computerised facial approximation method for archaic humans based on dense facial soft tissue thickness depths. *Archaeological and Anthropological Science*. 2021, 13:186.

6.1 Introduction

Facial approximation (FA) or craniofacial reconstruction aims at recreating a potential facial appearance from a dry skull. This technique is often the last hope in the realm of forensic science when no other clues and evidence support the investigation and identification (Wilkinson, 2010). Based on the assumed relationship between soft tissue and the bony structure, FA has been applied in archaeology to reconstruct the portraits and facial appearances of people in the past (Benazzi et al., 2009; Hayes et al., 2017; Kustar, 2004; Marić et al., 2020; Shui & Wu, 2018). It has sometimes been applied to named individuals, but more usually unnamed people from the past. Nonetheless, these applications always focused on modern humans, and they are less commonly applied to archaic humans, where differences in skull and facial morphologies make the approximation more challenging. In recent years, the approximated appearance of our fossil ancestor has become an area of study for anthropologists and has also captured the imagination of the general public, influencing perceptions of how “like us” and how “human” Neanderthals were. The visualisation of the approximated face rather than an imaginary approximation provides an effective 3D presentation to help us perceive and understand the characteristic features of human fossils. In addition, FA offers a new insight to investigate the morphological shape variations between archaic humans and *Homo sapiens*.

3D manual facial approximation approaches have been widely used to recreate facial appearances (Hayes, 2016). Anthropologists collaborated with artists to recreate a possible likeness by means of modelling clay or plasticine over the replica of the skull and adding the facial features, e.g. eyes, nose, mouth, etc. During this procedure, muscle structures and facial soft tissue thickness depths (FSTDs) at anatomical landmarks can be used to represent the craniofacial relationship between soft tissue and skull. The manual FA approaches can be divided into three main categories: The Russian anatomical approach, the American anthropometrical approach, and the combination Manchester approach (Verzé, 2009). However, they are heavily dependent on the degree of anthropological interpretation, and the practitioners’ subjective experience. Under such circumstances, multiple approximated faces of the same skull can be produced. For instance, three portraits of Ferrante Gonzaga, an Italian nobleman of the Renaissance, have been recreated (Fatuzzo et al., 2016). Such various approximations with inconsistent facial features might probably lead to less public confidence when no convincing hypothesis and evidence can be provided.

With the rapid progress in computer science and medical image acquisition, computerized FA technology has been gradually developed to increase the level of accuracy and reliability of the approximated face. The basic idea is to mimic the manual FA approach using the computer (Wilkinson, 2005). Using the FSTDs at anatomical

landmarks and knowledge of facial muscle, both 2D and 3D interactive graphics technologies which mimic the Manchester approach have been developed. In 2D interactive FA, the frontal and profile portraits were recreated using Adobe Photoshop software (Hayes et al., 2012). Likewise, 3D interactive FA was used to recreate a 3D probable likeness through a haptic feedback device and 3D software, e.g. Autodesk 3ds Max, ZBrush, Blender, etc. (Lee et al., 2014; Miranda et al., 2018; Short et al., 2014; Wilkinson et al., 2006). In their work, the tissue depth pegs which represented the FSTDs at anatomical landmarks were attached to the correspondence vertices of the dry skull, and the facial muscles were revised and attached to the surface of the skull. Then the facial features were added and sculpted to improve the reliability of the approximated face. However, all these technologies require both anatomical knowledge and expertise in modelling skills. Anthropologists have to invest great effort in manual modelling when they wish to produce a range of multiple candidate faces that use the FSTDs of different samples.

Pioneering work on 3D graphical computerized FA was first proposed by Vanezis (1989). The average FSTDs at a limited number of anatomical landmarks were used to produce a coarse mask and then the generic face was deformed to recreate a facial appearance over the dry skull. It is acknowledged that the greater number of FSTDs is acquired, the greater reliability of the approximation is achieved. Another effective computerized FA employed the deformation-based approach based on the assumption the verified craniofacial relationship of the template model is similar to that of the dry skull, removing the skull morphology variations (De Buhan & Nardoni, 2018; Deng et al., 2011; Nelson & Michael, 1998; Quatrehomme et al., 1997; Turner et al., 2005). In this procedure, either a generic face or a specific face based on the properties of the dry skull, e.g. age, sex, ethnic group, etc. was chosen as the template model. Then the template face was deformed following the same transformation that was calculated by deforming the template skull to the dry skull. This approach is simple and easy-to-use, because it does not require the FSTDs table at anatomical landmarks. In recent years, with the increasing availability of skull and face datasets of modern living humans, a regression-based method has been applied to study the craniofacial relationship based on principal components (PC) scores of every skull and face in the shape space (Berar et al., 2011; Deng et al., 2016; Paysan et al., 2009). Then this predicted craniofacial relationship can be used to recreate the facial appearance.

With regard to the similarities in the craniofacial relationship between relatively recent modern human remains and modern living humans, the majority of existing studies focused mainly on approximating the appearance of archaeological human fossils. In contrast, only a few publications concerned the investigation of archaic humans (Hayes, 2016). Hayes et al. estimated the frontal and lateral appearances of

Liang Bua, the holotype of *Homo floresiensis* (Hayes et al., 2013). In their work, the 2D profile outlines of the approximated face were created based on the FSTDs at landmarks. Then muscle images were deformed and attached to the surface of the skull. Finally, the reliability of the reconstructed face was evaluated using geometric morphometrics (GM). Because 3D facial morphology might allow anthropologists to better elucidate the facial characteristics of archaic humans and investigate evolutionary changes in the face, a 3D computerized FA approach still needs to be further investigated.

The Jinniushan (JNS) 1 cranium, dating back to 260,000 years BP, was discovered in Yingkou County, Liaoning Province in northeast China in 1984 (Rosenberg et al., 2006; Wu, 1988). It is one of the most important fossils in East Asia and has been used to investigate morphological features and shape variations with other fossils (Athreya & Wu, 2017; Hublin, 2013). It appears that its supraorbital shape, superciliary arch thickness and shape, postorbital constriction, and paranasal inflation are somewhat closer to those of Dali and Maba individuals, who are considered to represent population immigration from outside of China, and to be the result of an admixture with archaic humans (Andrews, 1986; Rightmire, 1998). Although a manual approach has been used to produce the facial appearance of JNS 1, considerable interest has been shown in investigating the approximated face of JNS 1 based on reasonable assumption and supporting data, rather than experience and imagination. This paper aims to provide a computerized FA method to approximate the plausible and reproducible face of the archaic human.

6.2 Materials and methods

6.2.1 Materials

6.2.1.1 The archaic human fossil

The JNS 1 cranium retained most of the maxillary dentition although the bone has been broken into more than one hundred pieces (Wu, 1988). In an attempt to perform FA successfully, the cranium required careful examination and restoration. It has been manually repaired by researchers from the Institute of Vertebrate Paleontology and Paleoanthropology (IVPP) in Beijing, China. The restoration procedures were as follows: firstly, every fragment of JNS 1 fossil was cleaned and strengthened. Secondly, the fractured fragments were carefully matched together based on the similarity of the boundary of every fragment following the experience of the researchers. Thirdly, superglue was used to adhere fragments to each other. Finally, plaster was used to fill in the missing region of the cranium guided by geometric constraints. Anthropologists predicted the sex and age of JNS 1 through the analysis of morphological features, sutures, dental wear, etc. In recent years, JNS 1 was suggested to be female because of two important features, the subpubic concavity and the medial aspect of the ischiopubic

ramus (Rosenberg et al., 2006). Likewise, based on the comparison of dentition and the analysis of tooth wear, an early study suggested that JNS 1 was over 30 years old (Wu, 1988), but more recently it was suggested to have been approximately about 20-30 years old (Herrera & Garcia-Bertrand, 2018).

Because only the JNS 1 cranium remained and the mandible was not preserved, a well preserved late archaic human mandible was required to assemble the JNS 1 cranium. But it remains challenging to find a suitable mandible with similar age and features. We have to decide to use two archaic human mandibles whose ages covered the age of JNS 1, i.e. one mandible is more recent than JNS 1 and the other is older than JNS 1, to repair JNS 1.

The Tabun 2 mandible was found in stratigraphic layer C of the Tabun cave, one of the paleoanthropological sites in the Near East. It was reconstructed and virtually recovered in six fragments, but lacking the left condyle, part of basilar symphysis, etc. (Quam & Smith, 2002; Schwartz & Tattersall, 2000). Morphologically, Tabun 2 is relatively large and robust, and it indicates a strong development of anterior marginal tubercle, a triangular basal corpus profile at the symphysis, mandibular foramina, etc. It exhibited a mixture of morphological features of Neanderthals and early modern humans (Harvati & Lopez, 2017). Figure 6.1a shows the JNS 1 cranium (peach colour) and the Tabun 2 mandible (grey colour).

In addition, the well preserved and complete Mauer 1 mandible (Wagner et al., 2010), a holotype of *Homo heidelbergensis*, was found near Mauer, southeast of Heidelberg, Germany in 1907. It is the oldest hominin fossil reported to date from central and northern Europe. It is of note that Mauer 1 exhibits a mixture of both primitive and modern features (Mounier et al., 2009). In this study, Tabun 2 and Mauer 1 were selected to fit with the JNS 1 cranium. Then, these two reassembled skulls (called JNS 1 using Tabun 2 and JNS 1 using Mauer 1) were used to approximate the face of JNS 1. Figure 6.1b displays the JNS 1 cranium (peach colour) and the Mauer 1 mandible (grey colour).

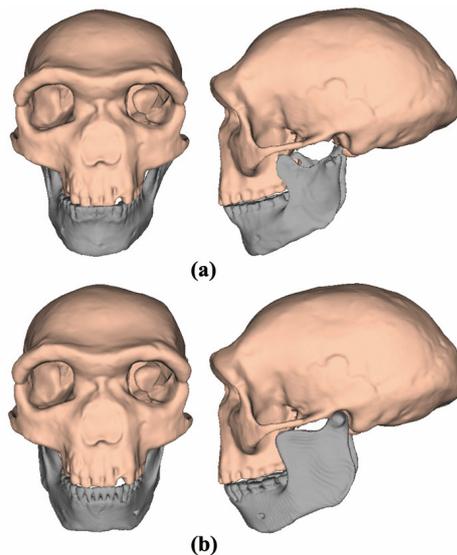


Figure 6.1 JNS 1 cranium (peach colour) and two different mandibles (grey colour). (a) Tabun 2. (b) Mauer 1.

6.2.1.2 Skull and face datasets of modern living humans

In our previous study (Shui et al., 2017), we constructed a skull and face dataset consisting of 140 individuals who lived in Shaanxi province in northern China. Each individual had normal morphology and had never undergone any orthodontic treatment. Medical images were acquired by means of a clinical multi-slice CT scanner system (Siemens Sensation 16). The CT images of each individual were archived in standard DICOM 3.0 with a resolution of 512×512. All participants were provided with full details of the study and written informed consent. This work has been approved by the Institutional Review Board (IRB) of the Affiliated Hospital of the Shaanxi University of Chinese Medicine. To obtain the craniofacial relationship between soft tissue and skull for recreating facial appearance of JNS 1, we selected 60 modern humans (30 females and 30 males aged 20-30 years old) from the constructed skull and face dataset (Shui et al., 2017) in the present study. This research was approved by the Ethics Review Committee of Department of Archaeology, University of York.

Our previous studies constructed dense corresponding vertices among skulls and faces, respectively. The procedure was as follows: firstly, image segmentation and the well-known Marching Cubes algorithm (Lorensen & Cline, 1987) were used to convert a series of CT images to the digital skull (or face). Secondly, the external surface of every skull (or face) within our dataset was computed. Thirdly, anatomical landmarks of the skulls and faces were defined and placed. Fourthly, the average skull (or face) estimated in our previous study (Shui et al., 2017) was chosen as the template. To register the template model and every skull (or face), the hybrid non-rigid registration method was applied by combining Iterative closest point (ICP), Thin-plate splines (TPS), and compact support radial basis function (CSRBF) algorithms. This step allows the warped template model to match the target skull (or face) closely. Next, every vertex of

the warped template model was projected onto the target skull (or face) by finding the closest points. This step can establish dense point correspondences, i.e. every skull (or face) had the same number of vertices and each vertex of every skull (or face) was located approximately in corresponding positions. To remove the effects of location, orientation and scaling, generalized Procrustes analysis (GPA) and principal component analysis (PCA) were carried out to construct the skull and face statistical shape model. Every skull (or face) can be represented by the coordinates of the average skull (or face), and the linear combinations of PC scores and corresponding orthogonal PCs (Shui et al., 2020).

6.2.1.3 Anatomical landmark definitions

In order to estimate the overall shape of the facial appearance, an anatomical landmark and semilandmark configuration was defined. A total of 91 anatomical landmarks were chosen and their 3D coordinates were acquired using Landmark Editor software (Wiley et al., 2005), where 17 anatomical landmarks were located on the midline and 74 anatomical landmarks were bilateral, respectively (Table 6.1). Most of these anatomical landmarks were defined according to Martin's definitions (Martin, 1928). Then, 404 semilandmarks were placed on JNS 1, which were identified in 16 patches based on the given landmarks. Here the semilandmarks of each patch were equally spaced within a 3×3 as a patch, and each patch with less than 9 anatomical landmarks was replenished with the middle points of two adjacent anatomical landmarks (Table 6.2). Finally, these semilandmarks were projected to the modern human skull, i.e. we established geometric correspondences between two skulls (Gunz & Mitteroecker, 2013). These landmarks and semilandmarks can facilitate the subsequent registration of the template skull and JNS 1, thereby improving the reliability of establishing dense point correspondences between two skulls.

Table 6.1 Anatomical landmarks

No.	Landmark definition	Position
1	Nasion	midline
2	Rhinion	midline
3	Nasospinale	midline
4	Subspinale	midline
5	Prosthion	midline
6	Infradentale anterius	midline
7	Gnathion	midline
8	Pogonion	midline
9	Glabella	midline
10	Ophryon	midline
11	Metopion	midline
12	Bregma	midline
13	Vertex	midline

14	Lambda	midline
15	Inion	midline
16	Intersection between inferior nuchal line and external occipital crest*	midline
17	Opisthion	midline
18	Maxillofrontale	bilateral
19	Orbitale	bilateral
20	Dacryon	bilateral
21	Superior orbital fissure (foramen), the closest point superior to the superior fissure*	bilateral
22	Most prominent point of superciliary above superior orbital fissure*	bilateral
23	Root of superciliary above superior orbital fissure*	bilateral
24	Ectoconchion	bilateral
25	Frontomalare orbitale	bilateral
26	Frontomalare temporale	bilateral
27	Zygion	bilateral
28	Zygomaxillare	bilateral
29	Jugale	bilateral
30	Most concave point on the inferior margin of maxilla*	bilateral
31	Vertical projection from jugale to lower margin of zygomatic*	bilateral
32	Infraorbital foramen	bilateral
33	Middle pyriform point, horizontal projection from infraorbital foramen to pyriform aperture*	bilateral
34	Lower pyriform point, lowest point of pyriform aperture*	bilateral
35	Stephanion	bilateral
36	Frontotemporale	bilateral
37	Auriculare	bilateral
38	Most prominent point on supramastoid crest*	bilateral
39	Asterion	bilateral
40	Mastoidale	bilateral
41	Ectomalare	bilateral
42	Coronion	bilateral
43	Lowest point of mandibular notch*	bilateral
44	Gonion	bilateral
45	Condylion laterale	bilateral
46	Ramus posterius, most concave point on the posterior margin of ramus*	bilateral
47	Ramus anterior, most concave point on the anterior margin of ramus*	bilateral
48	Vertical projection from lowest point of mandibular notch to lower margin of mandible along ramus	bilateral
49	Vertical projection from alveolare of lower m2 to lower margin of mandible*	bilateral
50	Temporale anterior, most anterior point of temporal squama*	bilateral
51	Temporale superius, most superior point of temporal squama*	bilateral
52	Alveolare of upper P3*	bilateral
53	Alveolare of lower m1*	bilateral

54	Mental foramen	bilateral
55	Mental laterale, turning point from mental to mandibular body on the inferior margin*	bilateral

Table 6.2 Eight patches on left side of JNS 1.

No.	Patch of anatomical region	Semilandmark density	Numbers of anatomical landmarks used in the patch
1	Zygomatic	5×5	19, 24, 25, 28, 29, 31
2	Maxilla	5×5	5, 28, 30, 32, 33, 34, 41, 51
3	Mandible	5×5	42, 43, 44, 45, 46, 47, 48, 49
4	Mental protuberance	5×4	6, 7, 8, 53, 54, 55
5	Nasal	5×3	1, 2, 18, 20, 33
6	Superciliary arches	5×5	9, 10, 18, 21, 22, 23, 25, 26, 36
7	Frontal	9×5	10, 11, 12, 35, 36
8	Parietal and occipital	11×6	12, 14, 15, 39, 50, 51

*Anatomical landmarks are defined by the authors and the rest of anatomical landmarks without special notice are from Martin (1928).

6.2.2 Methods

Figure 6.2 summarizes the framework of the proposed FA method. Firstly, the JNS 1 cranium and the selected mandible were virtually reassembled and the missing geometry was repaired. Secondly, a coarse-to-fine computerized FA approach was proposed to recreate the possible likeness of JNS 1 based on the assumption that the distribution of average FSTDs of the modern humans within the dataset is similar to that of JNS 1. This procedure comprised four steps: a) A hybrid non-rigid registration approach was carried out to establish the dense geometric correspondences between the template skull and JNS 1, where 495 landmarks and semilandmarks were used to guide the transformation mapping; b) The dense FSTDs of the template were calculated and visualized in a graphical format; c) The coarsely approximated face of JNS 1 was recreated by assigning dense FSTDs to the corresponding vertices of JNS 1. The TPS interpolation function was used to improve the approximation. Due to the absence of the actual mandible, multiple approximations can also be mathematically calculated by interpolating the surfaces of the approximated faces based upon Tabun 2 and Mauer 1. d) Quantitative evaluation was used to validate the reliability of the approximation through comparison of the distributions of FSTDs. Finally, we employed GM to examine the morphological shape variations between the approximated faces and modern human faces. We examined the effects of skull morphology and FSTDs on the approximated faces. Taking the approximation of JNS 1 using Tabun 2 as an example, we introduce the proposed method. All these methods were programmed using C++ and Matlab 2019 and software prototype can be downloaded from the Internet (<https://github.com/sissun/FAarchaic.git>).

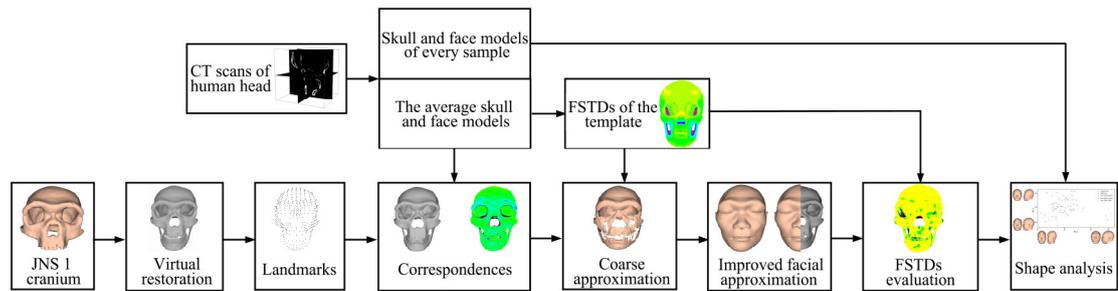


Figure 6.2 The pipeline of the computerized facial approximation method

6.2.2.1 The restoration of JNS 1

Because the temporomandibular joint (TMJ) that connected the JNS 1 cranium and Tabun 2 was insufficiently accurate, and the left condyle of Tabun 2 seemed incomplete, the first step was to predict the missing geometry of Tabun 2 and match the cranium and mandible closely. Before the virtual restoration, the external surface of JNS 1 was extracted. This comprised three steps: firstly, JNS 1 was transformed into the Frankfort coordinate system based on the left porion, right porion, left orbitale and the glabella. Secondly, an external point cloud was generated based on the cylindrical sampling algorithm. In this procedure, a couple of cross-section planes were generated between the bottom and top of JNS 1. For every cross-section plane, the external points were obtained by calculating the intersection points between JNS 1 and a set of given rays, starting at centroid of every cross-section along equally spaced angle vectors (Shui et al., 2020). Finally, the external point clouds were converted to a set of triangular meshes. Subsequently, we employed the mirror restoration method to repair the external surface of Tabun 2 (Gunz et al., 2009). Figure 6.3 shows the external surface of JNS 1 using Tabun 2 that comprised the anatomical landmark and semilandmark configuration.

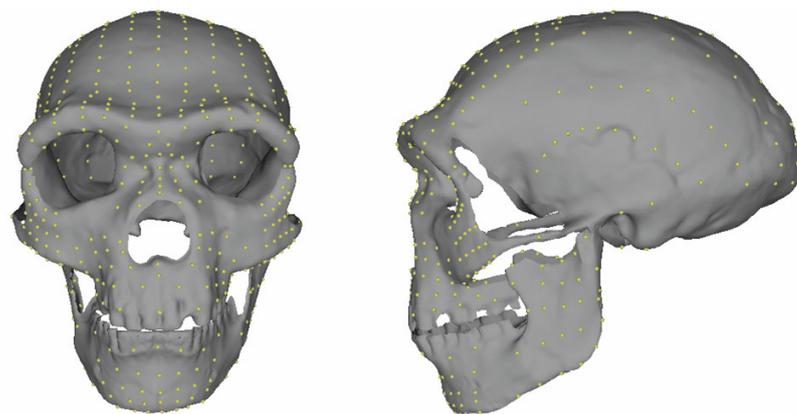


Figure 6.3 The restoration of JNS 1 using Tabun 2 that comprised the anatomical landmark and semilandmark configuration.

6.2.2.2 Computerized facial approximation

In anthropology, it is widely accepted that facial surface has a close relationship to the bony structure and that the overall shape of the face can be approximated based on

skull morphology and the craniofacial relationship between soft tissue and skull (Wilkinson, 2005). We proposed a coarse-to-fine FA approach to produce a reproducible and objective approximation.

a) Geometric correspondences between the template skull and JNS 1

We employed two steps to establish high-quality geometric correspondences between JNS 1 and the template skull. We assumed that all the landmarks and semilandmarks of the template skull and JNS 1 were represented by $\mathbf{P} = \{p_1, p_2, \dots, p_l\}$, $p_i = (x_i, y_i, z_i)$ and $\mathbf{Q} = \{q_1, q_2, \dots, q_l\}$, $q_i = (x_i, y_i, z_i)$, respectively, where l denoted the number of anatomical landmarks and semilandmarks. A popular non-rigid registration TPS function was first conducted to deform the template skull to JNS 1. During the deformation, it enabled the bending energy of the function $f(p_i) = q_i$ minimized (Bookstein, 1989). TPS can be represented by affine transformation parameters and non-affine warping parameters as the following linear equation

$$\begin{bmatrix} \mathbf{K} + \varphi \cdot \mathbf{I} & \mathbf{P} \\ \mathbf{P}^T & \mathbf{O} \end{bmatrix} \begin{bmatrix} \mathbf{w} \\ \boldsymbol{\alpha} \end{bmatrix} = \begin{bmatrix} \mathbf{Q} \\ \mathbf{A} \end{bmatrix} \quad (6.1)$$

where the radial basis kernel can be represented by $\mathbf{K} = [K_{i,j}]$, $K_{i,j} = \|p_i - p_j\|$ and $\|\cdot\|$ denoted the Euclidean distance. φ was the regularization parameter that used to balance the smoothness. \mathbf{I} denoted the $l \times l$ identity matrix; \mathbf{O} denoted the 4×4 zero matrix; \mathbf{A} denoted the 4×1 zero matrix; $\boldsymbol{\alpha} = [a_0 \ a_1 \ a_2 \ a_3]^T$ and $\mathbf{w} = [\varpi_i]^T$ represented the affine and non-affine parameters, respectively.

Following the same transformation that was computed by warping the template skull to JNS 1, the template face was deformed to produce a possible likeness as a candidate face. Next, we employed non-rigid registration to allow the deformed template skull and JNS 1 to match closely by assigning an affine transformation to every vertex of the deformed template skull. Assumed affine transformations $\mathbf{X} = [X_1 X_2 X_3 \dots X_n]^T$ of all the vertices, we defined the cost function, $\mathbf{E}(\mathbf{X})$, which consists of anatomical landmarks and semilandmarks term $E_l(\mathbf{X})$, a local affine regularization $E_d(\mathbf{X})$, and a stiffness term $E_s(\mathbf{X})$. To evaluate the accuracy of the skull match, the geometric deviation between the deformed template skull and JNS 1 was quantitatively calculated and depicted in a graphical format.

The cost function was as follows:

$$\mathbf{E}(\mathbf{X}) = \alpha E_l(\mathbf{X}) + \beta E_d(\mathbf{X}) + \lambda E_s(\mathbf{X}) \quad (6.2)$$

where α , β and λ denoted the weights that guided the optimization process.

$E_l(\mathbf{X})$ was used to initialize and guide the registration as follows:

$$\mathbf{E}_l(\mathbf{X}) = \sum \| \mathbf{X}_i v_i - m_i \|^2 \quad (6.3)$$

where m_i was the i -th landmark and semilandmark of JNS 1 and v_i was the i -th

corresponding landmark and semilandmark of resulting deformation of the template skull via TPS.

The local affine regularization term expressed the distance between a vertex of JNS 1 and the corresponding vertex of the resulting deformation of the template skull as follows:

$$\mathbf{E}_d(\mathbf{X}) = \sum \mu_i \text{dist}^2(rq_i, \mathbf{X}_i rp_i) \quad (6.4)$$

where $\text{dist}()$ denoted the distances between the corresponding points of JNS 1 and the resulting deformation of the template skull, and μ_i denoted the reliability of the correspondences between these two meshes, denoted by rq_i and rp_i . We assumed that the nearest points between two meshes were the correspondences. In this procedure, the angles between normal vectors of the corresponding points and the Euclidean distance of the corresponding points can be used to improve reliability and reject the outliers.

The stiffness term was applied to regularize the deformation as follows

$$\mathbf{E}_s(\mathbf{X}) = \sum \|(\mathbf{X}_i - \mathbf{X}_j) \cdot \mathbf{G}\|_F^2 \quad (6.5)$$

where $\|\cdot\|_F$ was the Frobenius norm. \mathbf{X}_i and \mathbf{X}_j were the transformations of neighbouring vertices, which were connected by an edge that belonged to the resulting deformation of the template skull. $\mathbf{G} = \text{diag}(1, 1, 1, \gamma)$ denoted a weighting matrix.

b) Dense FSTDs of the template

During the acquisition of FSTDs of the template, the normal vector that was almost perpendicular to the surface of the bony structure was considered to be the measurement direction. A ray that started at a vertex of the template skull along the normal vector often passed through the template face, thus the intersection point can be calculated. The FSTDs were defined as the Euclidean distances between pairs of corresponding vertices. It is of note that the normal vector of a given point that was determined by the geometric coordinates and topologies of the neighbouring vertices influenced the accuracy of the FSTDs measurement. When the surface contained noise and sharp features, e.g. boundary of the surface, normal estimation remained a challenge.

We extracted stable regions with robust normal estimation from the whole skull and then used FSTDs of the vertices within these stable regions to accomplish FA. It comprised two steps: firstly, we calculated the FSTDs of all the vertices along the closest distance vectors (Huempfer-Hierl et al., 2015). For every vertex of the template skull, the nearest point on the template face was searched and the FSTDs were defined as the Euclidean distances between every pair of corresponding vertices. Secondly, the discrepancies between FSTDs along the normal vectors and those along the closest distance vectors were calculated. Once the deviation was less than the threshold, the

vertex was suggested to be a stable vertex. Figure 6.4a shows stable regions (red points) and unstable regions (blue points). In addition, the boundary vertices of the skull often were not considered to belong to the stable region. They can be extracted from the triangle meshes based on the assumption that the one-ring adjacent points of every boundary vertex cannot form a closed loop (Shui et al., 2020). Figure 6.4b shows the boundary vertices of the template skull (green points). Neither the unstable regions nor the boundary vertices were used to generate the coarsely approximated face.

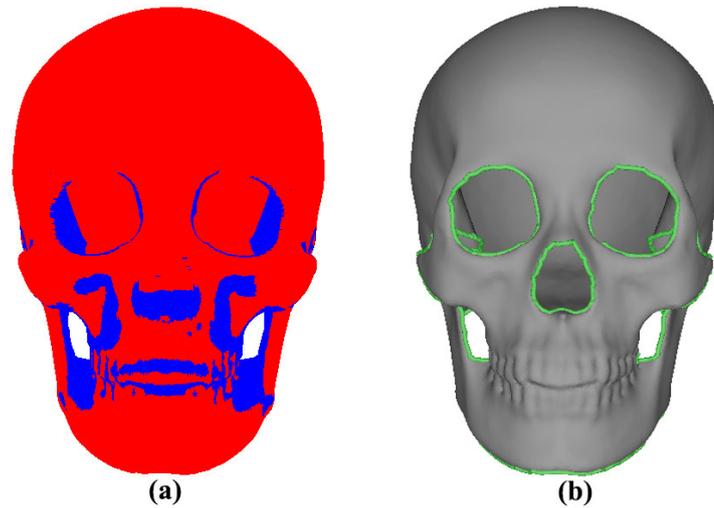


Figure 6.4 Segmentation of template skull. (a) Stable vertices (red points) and unstable vertices (blue points). (b) Boundary vertices (green points).

c) A coarse-to-fine facial approximation

The overall shape of the facial approximation can be coarsely produced based on the dense FSTDs of the template using the following equation:

$$f_i = s_i + \mathbf{V}_i \cdot d_i \quad (6.6)$$

where f_i and s_i denoted the geometric coordinates of the i -th vertex of the approximated face and JNS 1, respectively. \mathbf{V}_i represented the normal vector of the i -th vertex of JNS 1, and d_i denoted the soft tissue thickness of the corresponding vertex of the template.

The predictions of FSTDs and corresponding measurement directions were always inconsistent with the actual ones, thus the approximated face would be unsmooth. In addition, there always existed some voids, such as the eyes, nose, cheeks, etc. We employed a TPS function to warp the candidate face to the coarsely approximated face to improve the approximation. Because the position of the control point located on the two faces will greatly influence the deformation, we calculated the corresponding intersection points as control points based on the known anatomical landmarks and semilandmarks of JNS 1.

Additionally, we offered a tool to mathematically calculate multiple approximations

that simulated the mandible morphology changes using the following equation

$$\mathbf{JNSFace}(\eta) = \eta \cdot \mathbf{Face}_1 + (1-\eta) \cdot \mathbf{Face}_2 \quad (6.7)$$

where $\mathbf{JNSFace}$ represented the interpolated approximation. \mathbf{Face}_1 and \mathbf{Face}_2 denoted the approximated faces of JNS 1 using Tabun 2 and JNS 1 using Mauer 1, respectively. $\eta \in [0,1]$ represented the weight coefficient.

d) Evaluation of the reliability

We validated the reliability of the approximated face by means of examining whether or not the distribution of FSTDs of the template was consistent with that of the approximated face. The FSTDs defined along the closest distance vectors were used, because they were insensitive to measurement direction and data noise (Gietzen et al., 2019). The FSTDs deviation between the template and the approximation was calculated and visualized.

6.2.2.3 Morphological shape variations of facial approximation

GM was carried out to capture the main features of the approximated faces and examine the geometric morphological variations between the approximated faces and modern human faces. GPA was first used to register all the vertices of the approximated faces and modern human faces, removing translation, rotation, and scaling (O'Higgins & Jones, 1998). Thus all the faces can be represented in the non-linear Kendall's shape space. The centroid size (CS) which is defined as the square root of the summed squared distances between all corresponding vertices and their centroid was calculated. Then, PCA was conducted on the Procrustes aligned coordinates to construct a facial shape tangent space. In this shape space, every sample was represented by the average face and the linear combinations of PC scores and corresponding independent orthogonal PCs. The Matlab were used to apply geometric morphometric analysis. Next, Student's *t*-test was carried out to verify the significant level of PC of interest between the modern faces and the approximated faces. Finally, the visualisation technique was used to investigate the extent to which PC greatly explained the main patterns of morphological variation. In this process, two new faces along the positive and negative PC of interest were generated as follows

$$\mathbf{Face}(\gamma) = \overline{\mathbf{Mean}} + 3\gamma\delta_i\phi_i \quad (6.8)$$

where $\overline{\mathbf{Mean}}$ and γ denoted the average face and weighting coefficient (it was set to 1 or -1), and δ_i denoted the standard deviation of the *i*-th PC and ϕ_i represented the *i*-th PC.

6.2.2.4 The effects of FSTDs and skull morphology on the approximated face

It is noted that skull morphology and the distribution of FSTDs are the two fundamental components of FA. We examined how the choice of FSTDs affected the approximated face. The approximated faces of the same JNS 1 were produced based on the average FSTDs of the females and males within our dataset. Then the FSTDs deviation between two different approximations was calculated and depicted in a graphical format.

As the real mandible of JNS 1 was not survived, we investigated the effect of different mandibles on the approximated faces. Different approximated faces of JNS 1 using Tabun 2 and JNS 1 using Mauer 1 were produced based on the same distributions of FSTDs, respectively. Then the geometric deviation between the approximations was used to examine the shape difference.

6.3 Results

6.3.1 Facial approximation of JNS 1

Since JNS 1 is suggested to be a female, the average skull and face of the female group (Figure 6.5a) were chosen as the template to approximate the facial appearance. Based on 495 anatomical landmarks and semilandmarks, we first used the TPS deformation approach to approximate the facial appearance of JNS 1 using Tabun 2. Figures 6.5b and c show the deformed template skull and the candidate face. Figure 6.5d shows the template skull (grey colour) and JNS 1 (peach colour). It can be seen that the deformed template skull (black points) does not match JNS 1 (peach colour) closely, as shown in Figure 6.5e. The geometric difference between the deformed template skull and JNS 1 was calculated and depicted in Figure 6.5f, where the Euclidean distance of every corresponding vertex was calculated. The average distance value of the vertex was 0.98 mm. The smaller deviation regions were located around the top of the cranial vault, whereas the largest geometric deviation regions were found around superciliary arches, the bottom of frontal, zygomatic arch, lateral mastoid process, the bottom of the occipital bone, partial regions of maxilla and teeth, the greater wing of the sphenoid, etc. It is noted that the approximated face might be inaccurate around these regions and the conventional deformation-based FA approach needs to be improved.

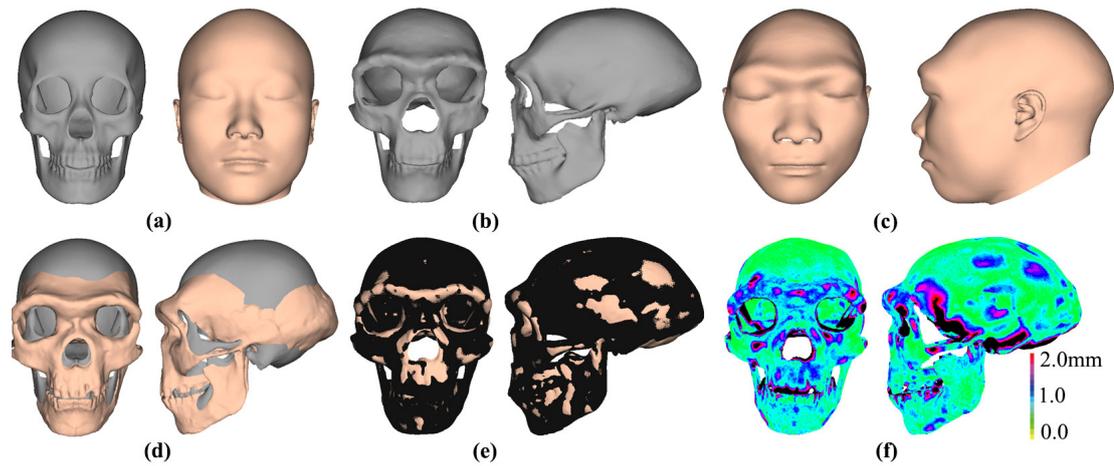


Figure 6.5 The deformation of the template skull and face using TPS. (a) The template skull and face. (b) The deformed template skull. (c) The deformed template face. (d) The template skull (grey colour) and JNS 1 using Tabun 2 (peach colour). (e) The superimposition of the deformed template skull (black points) and JNS 1 (peach colour). (f) Visualisation of geometric deviation between the deformed template skull and JNS 1 using Tabun 2.

We proposed a coarse-to-fine facial approximation approach by attaching the dense FSTDs of the template to the corresponding vertices of JNS 1. The non-rigid registration was used to warp the deformed template skull to JNS 1 to generate a better deformation result (Figure 6.6a). Figure 6.6b shows JNS 1 (peach colour) and the deformation result (black points). The geometric difference between the deformed template skull and JNS 1 was calculated and visualized in Figure 6.6c, where the Euclidean distance of every corresponding vertex was calculated. Almost 96.9% of overall vertices on JNS 1 showed the deviation within a deformation error of 1.0 mm, and the average deviation value was 0.54 mm. The area (black points) where the deviation was greater than 2.0 mm can be mainly observed around the boundaries of the bony structure. The region where the geometric deviation was greater can be observed around the superciliary arches and nasal bone.

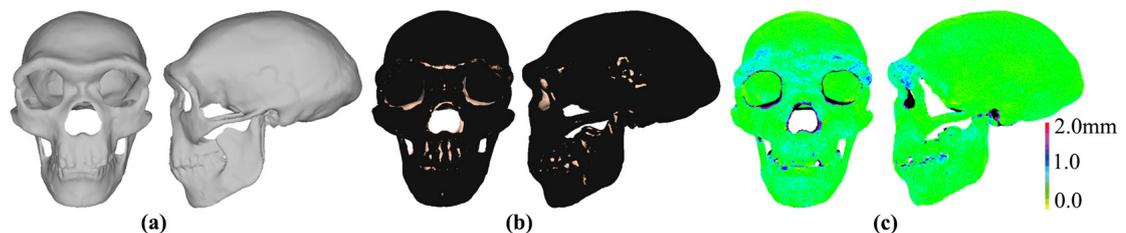


Figure 6.6 The deformation of the template skull using the proposed method. (a) The deformed template skull. (b) The superimposition of the deformed template skull (black points) and JNS 1 using Tabun 2 (peach colour). (c) The visualisation of geometric deviation between the deformed template skull and JNS 1 using Tabun 2.

Dense FSTDs of the template that were calculated along the normal vector were visualized in Figure 6.7a. It appeared that the FSTDs were almost distributed symmetrically along the midsagittal plane. The thinner FSTDs were observed at vertices located around the forehead, superciliary arches, nasal bone, and the cranial vault. The

thickest FSTDs were observed around the cheek region, the greater wing of the sphenoid, the bottom of the occipital bone, etc. Figures 6.7b and c display the point clouds and the triangle meshes of the coarsely approximated face. For every landmark and semilandmark of JNS 1, we calculated the corresponding points on the candidate face (left figure) and the coarsely approximated face (right figure), as seen in Figure 6.7d. Figures 6.7e and f show the improved resulting approximation that is created by warping the candidate face to the coarse approximation using TPS. The result shows that the approximated face has a lower forehead, and a protruding, wider and elongated middle and upper face. Also, it has robust eyebrows, a broad and short nose, and a wide mouth.

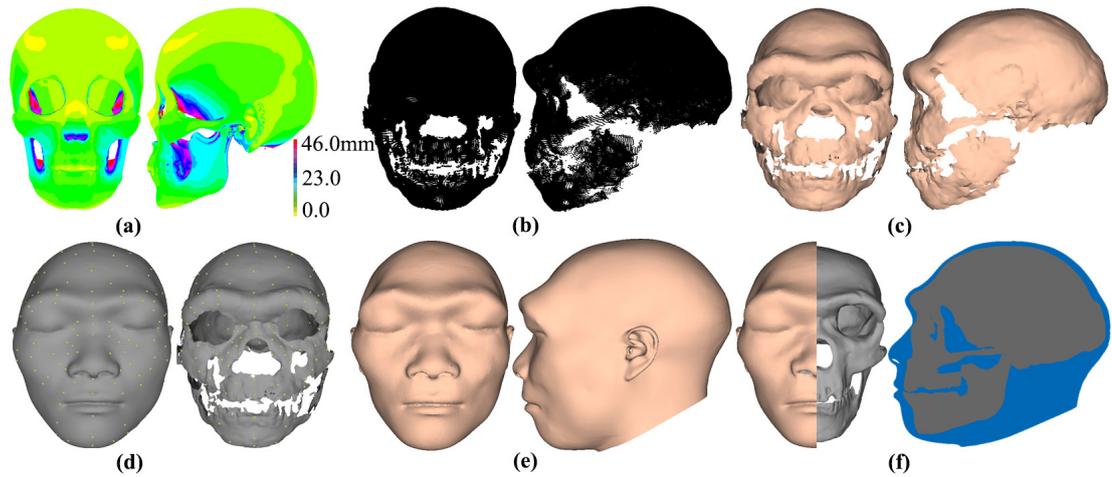


Figure 6.7 The facial approximation of JNS 1 using the proposed method. (a) The visualisation of dense FSTDs of the template. (b) The point clouds of the coarsely approximated face. (c) The surface of the coarsely approximated face. (d) The corresponding points on the coarsely approximated face and the deformed face using TPS. (e) The improved resulting approximation by warping the deformed template face. (f) The superimposition of JNS 1 and the facial approximation.

We compared the approximated faces using two different methods based on the comparison of the distribution of FSTDs. Figures. 6.8a-c illustrate the distributions of FSTDs of the candidate face, the coarsely approximated face and the improved resulting approximation, respectively. Figure 6.8d shows the FSTDs deviation between the candidate face and the template (Table 6.3). Almost 62.9% of overall vertices on the candidate face showed geometric deviation within a discrepancy of 1.0 mm, and the average deviation value was 1.17 mm. The area where the deviation was greater than 2.5 mm can be observed around the coronoid process, ramus, greater wing of the sphenoid, zygomatic, etc. The deviation of the vertex within the superciliary arches and nasal bone was greater as well. Figure 6.8e shows the FSTDs deviation between the coarsely approximated face and the template. Almost 87.7% of overall vertices on the coarse approximation showed the deviation within a discrepancy of 1.0 mm, and the average deviation value was 0.39 mm. It can be seen that distribution of average FSTDs of the template was very close to that of the coarse approximation, except for the side of the teeth, mandibular condyle, etc. Figure 6.8f illustrates FSTDs deviation between the resulting approximation and template. Almost 76.4% of overall vertices on the

improved approximation showed the deviation within a discrepancy of 1.0 mm, and average deviation value was 0.85 mm. The area where the deviation was greater than 2.5 mm can be observed around the coronoid process, both sides of the teeth, the greater wing of the sphenoid, etc.

Table 6.3 FSTDs deviation (mm) and the percentage distribution (%) between three approximations and the template

FSTDs deviation	The candidate face		Coarsely approximated face		Improved approximation	
	%	deviation	%	deviation	%	deviation
(0.0,0.5]	40.5%	0.23	80.6%	0.08	51.6%	0.22
(0.5,1.0]	22.4%	0.73	7.1%	0.71	24.8%	0.71
(1.0,2.5]	24.5%	1.59	8.5%	1.62	17.3%	1.51
>2.5	12.6%	4.16	3.8%	3.74	6.3%	4.74
Total	100%	1.17	100%	0.39	100%	0.85

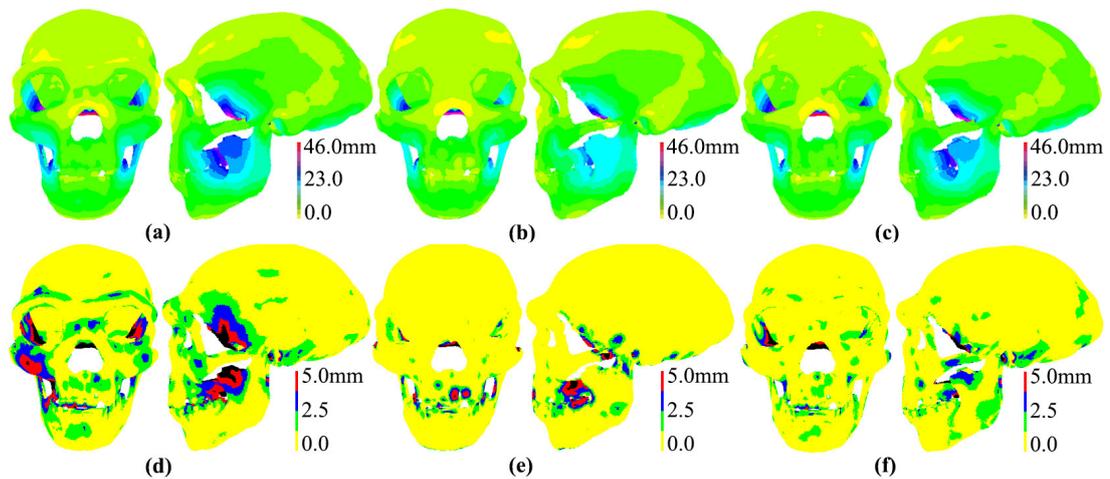


Figure 6.8 Comparisons of the facial approximations generated by different methods. (a) The distribution of FSTDs of the candidate face. (b) The distribution of FSTDs of the coarse approximation. (c) The distribution of FSTDs of the improved approximation. (d) The FSTDs deviation between the template and the candidate face. (e) The FSTDs deviation between the template and the coarse approximation. (f) The FSTDs deviation between the template and the improved approximation.

Because the FSTDs are the fundamental basis of the facial approximation in our method, we employed the FSTDs that were calculated using different methods to predict the overall shape of the coarsely approximated faces. Figure 6.9 shows the point clouds of the coarse approximation based on the normal vectors of the overall vertices, rather than the stable vertices. It appeared that much noise and outliers occurred around the coarsely approximated face. Figure 6.10a displays the FSTDs that are computed using a cylindrical sampling method (Shui et al., 2016). Figure 6.10b shows the coarsely approximated faces that are recreated by attaching these FSTDs to JNS 1 along the cylindrical sampling vector. Figure 6.10c shows the FSTDs deviation between the template and the approximation. Almost 60.5% of overall vertices on the approximation showed the deviation within a discrepancy of 1.0 mm, and the average deviation value was 1.3 mm. The area where the deviation was greater than 2.5 mm can be observed

around the side of teeth, greater wing of the sphenoid, the bottom of zygomatic region, mandibular condyle, superciliary arches, etc. These results indicated the proposed FA method can accurately assign the FSTDs to the corresponding vertices of JNS 1.

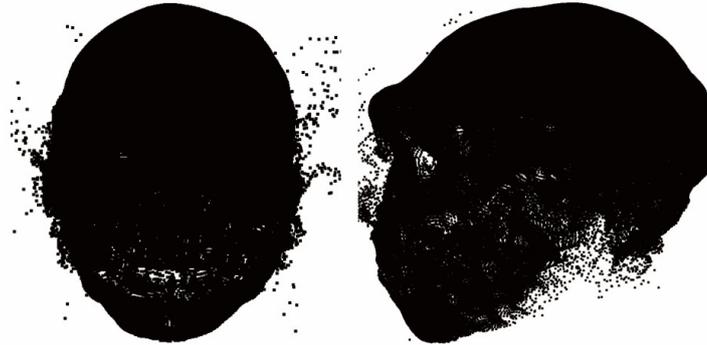


Figure 6.9 The coarsely approximated face using the normal vectors of overall the vertices.

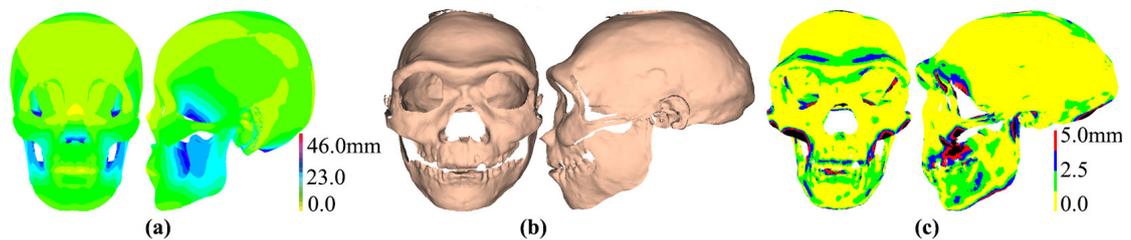


Figure 6.10 The coarsely approximated face using the FSTDs along the cylindrical sampling vectors. (a) The distribution of average FSTDs of the template using the cylindrical sampling method. (b) The coarsely approximated face. (c) The FSTDs deviation between the template and the coarse approximation.

6.3.2 Facial approximation of modern humans

In order to evaluate the reliability of the proposed FA approach, we approximated faces of two modern female skulls (01 and 03 skulls) and two modern male skulls (02 and 04 skulls). Then we compared the FSTDs deviation between the approximated and actual faces, and the geometric variations between the approximated and actual faces, as shown in Figure 6.11. From left to right, the leftmost column displayed the skull, and the next two columns illustrated the approximated face and the distribution of FSTDs. The middle two columns displayed the actual face and the distribution of FSTDs. The sixth column depicted the FSTDs deviation (*error*) for every vertex between the approximated and actual faces (Table 6.4). Additionally, the geometric deviation (*error*) between the approximated and actual faces was calculated (Table 6.5). The rightmost column displayed the geometric discrepancy of every vertex between the approximated and actual faces. Figure 6.11a shows the approximation of the 01 skull based on the average FSTDs of the females. Almost 59.3% of overall vertices showed the FSTDs deviation within 2.5 mm, and the average deviation value was 2.12 mm. etc. Figure 6.11b shows the approximation of the 02 skull based on the average FSTDs of the females. Almost 63.5% of overall vertices showed the FSTDs deviation within 2.5 mm, and the average deviation value was 2.19 mm. Figure 6.11c displays the approximation of the 03 skull based on the average FSTDs of the males. Almost 80.8% of overall

vertices showed the FSTDs deviation within a discrepancy of 2.5 mm, and the average deviation value was 1.43 mm. Figure 6.11d shows the approximation of the 04 skull based on the average FSTDs of the males. Almost 94.4% of overall vertices showed the FSTDs deviation within a discrepancy of 2.5 mm, and the average deviation value was 0.95 mm. A potential reason is that the distribution of the average FSTDs of the males is consistent with that of the 04 skull. However, the approximated chin of the 04 skull bore little visual resemblance to the actual chin, because the mental protuberance and tubercle seemed to be incomplete.

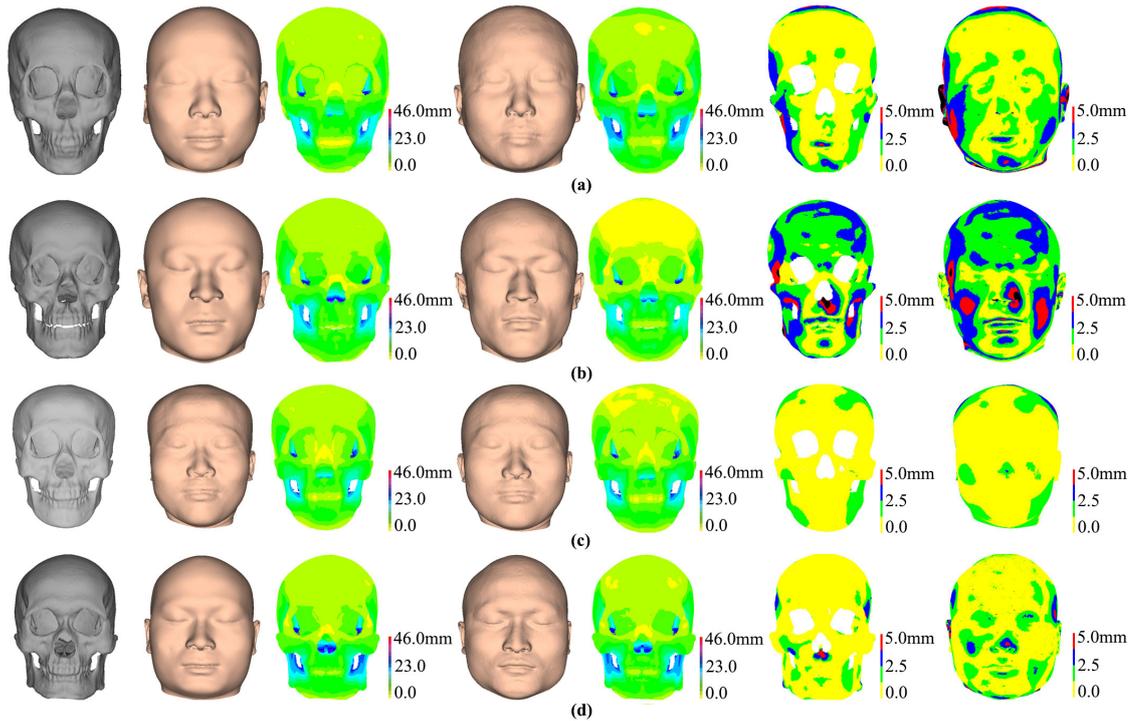


Figure 6.11 Four approximation examples of modern humans. From left to right, the left column displayed the skull and the next two columns illustrated the approximated face and the distribution of FSTDs. The middle two columns displayed the actual face and the distribution of FSTDs. The rightmost two columns depicted the FSTDs deviation and geometric discrepancy between the approximated and actual faces. (a) The 01 skull. (b) The 02 skull. (c) The 03 skull. (d) The 04 skull.

Table 6.4 FSTDs deviation (mm) and percentage distribution between the approximated and actual faces

FSTDs deviation	01 skull		02 skull		03 skull		04 skull	
	%	<i>error</i>	%	<i>error</i>	%	<i>error</i>	%	<i>error</i>
(0.0,1.0]	29.8%	0.44	17.0%	0.52	52.1%	0.42	63.8%	0.47
(1.0,2.5]	30.5%	1.73	46.5%	1.84	28.7%	1.64	30.6%	1.57
(2.5,5.0]	36.4%	3.40	33.7%	3.17	16.9%	3.68	5.6%	3.02
>5.0	3.3%	5.73	2.8%	6.40	2.3%	5.35	-	-
Total	100%	2.09	100%	2.19	100%	1.43	100%	0.95

Table 6.5 Geometric deviation (mm) and percentage distribution (%) between the approximated and actual faces

Geometric deviation	01 skull		02 skull		03 skull		04 skull	
	%	<i>error</i>	%	<i>error</i>	%	<i>error</i>	%	<i>error</i>
(0.0,1.0]	27.4%	0.57	16.9%	0.57	52.3%	0.47	53.0%	0.59
(1.0,2.5]	33.9%	1.71	44.9%	1.81	29.4%	1.61	35.8%	1.57
(2.5,5.0]	33.4%	3.41	32.5%	3.28	15.6%	3.75	9.1%	3.16
>5.0	5.3%	6.44	5.7%	6.78	2.7%	6.46	2.1%	6.58
Total	100%	2.21	100%	2.36	100%	1.49	100%	1.30

These results indicate that the overall shapes of the approximated faces bore resemblances to the actual faces. The area where the deviation was thinner can be observed around the forehead, nasal bone, maxillary bone, etc. But the greater deviation can be found around the cheeks, nose tip, zygomatic region, parietal and temporal regions, mental protuberance, etc. The distribution of FSTDs deviation was relatively consistent with that of the geometric discrepancy, but the average value of the FSTDs deviation was less than the geometric deviation. Thus the proposed FA approach can be applied to modern humans and the FSTDs deviation is a promising evaluation indicator to access the reliability of the approximated face.

6.3.3 The effects of FSTDs and skull morphology on facial approximation

We can recreate multiple approximated faces based on the distributions of FSTDs of different templates. Figure 6.12a shows another approximation of JNS 1 based on the average FSTDs of the males within our dataset. Figure 6.12b displays the FSTDs of the approximated face. The FSTDs deviation between two approximations based on the average FSTDs of the females and males was calculated, as shown in Figure 6.12c. Almost 69.9% of overall vertices showed the FSTDs deviation within a discrepancy of 1.0 mm, and the average FSTDs deviation value was 0.81 mm. The area (blue, red and black points) where the FSTDs deviation was greater than 1.0 mm can be observed around the mouth, nasal base, zygomatic, parietal and occipital regions, cheeks, etc.

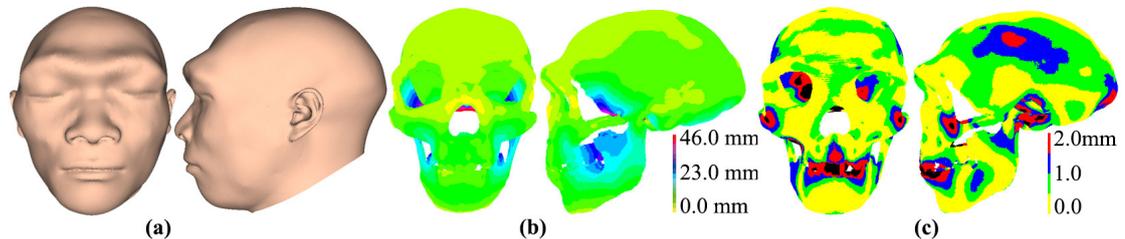


Figure 6.12 Facial approximation based on the average FSTDs of the males. (a) The approximated face. (b) The distribution of FSTDs of the approximation. (c) The FSTDs deviation between two approximations.

Using the above FA method, we recreated the approximated faces of JNS 1 using Mauer 1. Figure 6.13a shows the approximated face based on the average FSTDs of the females. Figure 6.13b shows the geometric variations between this approximation and

the approximation of JNS 1 using Tabun 2. We also used the distribution of average FSTDs of the males to recreate an approximation of JNS 1 using Mauer 1, as shown in Figure 6.13c. Figure 6.13d shows the geometric deviation between this approximation and the approximation of JNS 1 using Tabun 2. The approximation of JNS 1 using Mauer 1 has a wider face and a more robust chin than the approximations of JNS 1 using Tabun 2. These results showed that the geometric shape of the skull greatly influences the overall shape of the approximated faces. Based on the distribution of average FSTDs of the females, we employed the linear interpolation method to mathematically approximate three faces. Figure 6.14 displays three approximations using $\eta=0.25$, $\eta=0.50$ and $\eta=0.75$.

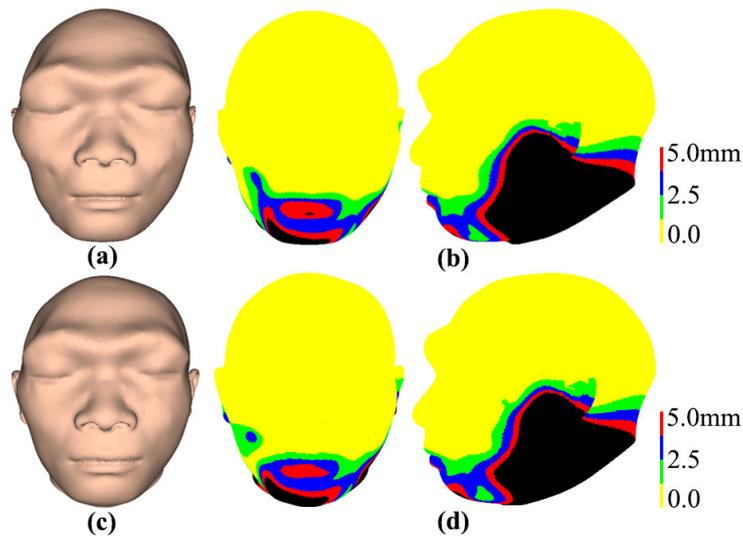


Figure 6.13 Facial approximation of JNS 1 using Mauer 1. (a) The approximated face using average FSTDs of the females. (b) Geometric deviation between two approximated faces. (c) The approximated face using the average FSTDs of the males. (d) Geometric deviation between two approximated faces.

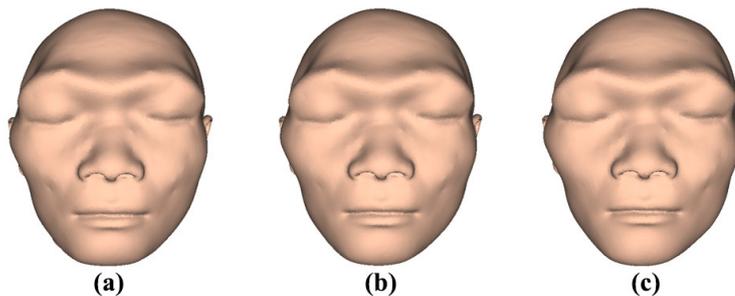


Figure 6.14 Three approximated faces using linear interpolation. (a) $\eta=0.25$. (b) $\eta=0.50$. (c) $\eta=0.75$.

6.3.4 Shape analysis of the approximated faces

GM analysis was conducted to capture the geometric features of the approximated faces. A total of 27 PCs accounted for over 95% of the morphological variance in the shape space. The first PC (PC 1) accounted for 31.6% of the morphological variance and the second PC (PC 2) accounted for 13.3% of the morphological variance. Figure 6.15 shows the plots of the first two PCs of four different approximations (Figure 6.S1), including the approximated face that used average FSTDs of the females and Tabun 2

(green point), the approximated face that used average FSTDs of the males and Tabun 2 (yellow point), the approximated face that used average FSTDs of the females and Mauer 1 (black point), the approximated face that used average FSTDs of the males and Mauer 1 (cyan point), 30 modern female faces (red points), and 30 modern male faces (blue points). It is of note that PC 1 ($p < 0.005$) and PC 2 ($p < 0.05$) have significant differences between modern human faces and the approximated faces. To identify the main patterns of shape variance, four new faces were recreated along the positive and negative directions of PC 1 and PC 2. The positive PC 1 connected with the approximated face with a relatively lower forehead, and robust and wide eyebrows; a protruding, wider and elongated middle and upper face; a broad and short nose, a wider mouth and robust chin. By contrast, the negative PC 1 represented a face with a prominent and protruding forehead, a narrower middle and upper face, relatively narrower mouth, nose and chin. All the approximated faces were located at the extreme positive end of PC 1. It indicated that the approximated face was greatly different from modern human faces and verified the characteristic features of the approximations. When the same FSTDs were performed, the approximated faces that used Mauer 1 were located on the right side along the positive PC 1. It indicates that these approximations have wider cheeks and robust chins.

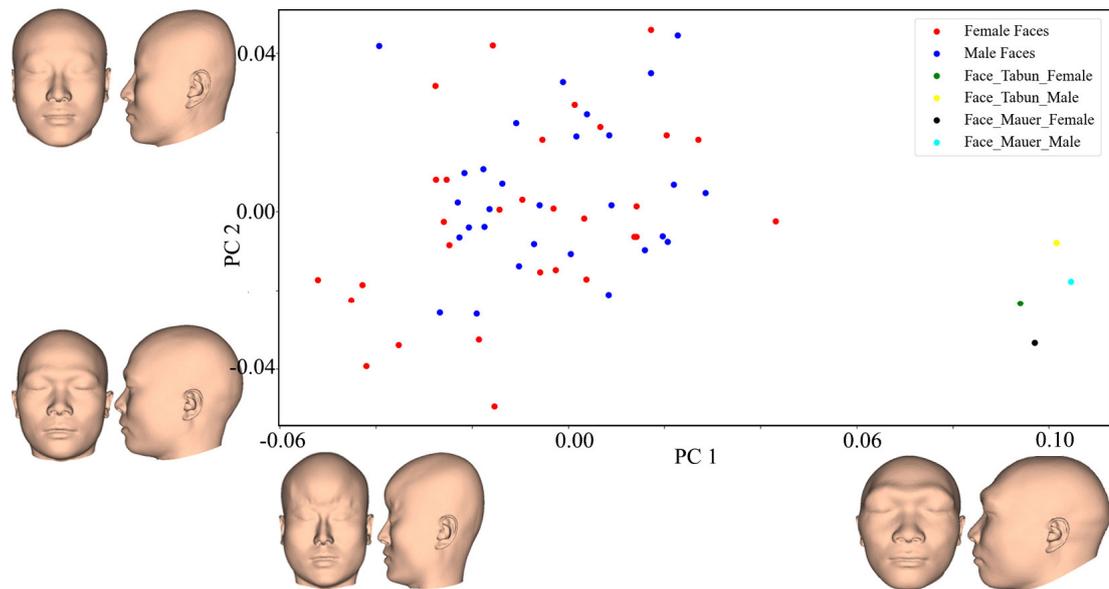


Figure 6.15 Comparison of the approximated faces and modern human faces in the shape space. Scatterplots of 30 female faces (red points), 30 male faces (blue points), the approximated face (green point) that used Tabun 2 and the FSTDs of the females, the approximated face (yellow point) that used Tabun 2 and the FSTDs of the males, the approximated face (black point) that used Mauer 1 and the FSTDs of the females, the approximated face (cyan point) that used Mauer 1 and FSTDs of the males. The frontal and profile views of new generated faces corresponding to the extreme limits of PC 1 and PC 2.

6.4 Discussion

Many previous studies have employed manual and computerized FA approaches to modern humans in archaeology and anthropology (Claes et al., 2010; Wilkinson, 2010).

However, when FA is applied to archaic humans, the main challenges are the poor preservation of archaic human fossils and the absence of their craniofacial relationship and anatomical knowledge. In this study, we propose a computerized coarse-to-fine FA approach to attach the distribution of average FSTDs of modern humans to archaic humans. The resulting approximation is promising, objective and repeatable, and the reliability can be evaluated through a quantitative comparison of the distributions of FSTDs. Furthermore, we investigate the effects of skull geometry and the distribution of FSTDs on the approximated face.

The first stage of FA is to examine and restore the dry skull. When the skull had missing parts or distortion, the TPS function was always used to deform the reflection structure of the intact side to fill in the gaps based on a landmark and semilandmark configuration. Other studies attempted to use computerized approaches to virtually reassemble the skull fractures together (Yu et al., 2012). In the worst case, the fossil specimen, as in the case of the mandible of JNS 1, had not survived. In these cases, there is a potential solution, using a mandible that has similar age and morphological features to match the cranium. But since the mandibles of archaic humans are rarely found, and the geometric shapes of the mandibles are always unique, it remains challenging to select an appropriate mandible to provide a good fit with the JNS 1 cranium. We attempted to use different mandibles to match the JNS 1 cranium and recreated different approximations based on the repaired skull. According to these approximated facial appearances, multiple approximations can further be mathematically recreated to provide some references to describe the overall shape. Although these approximations cannot be interpreted anatomically, they might provide a new perspective for researchers to illustrate the facial appearance.

During FA, the prediction of the craniofacial relationship between facial soft tissues and the dry skull, and the assignment of the predicted craniofacial relationship are two fundamental questions. The average FSTDs at landmarks and muscle structures of modern humans are always suggested to be the craniofacial relationship of people in the past (Hamre et al., 2017). In the study previously mentioned, the FSTDs of chimpanzees can be used to depict a thinner mid-face of archaic humans because FSTDs around the cheek were almost half that of modern humans (Hayes et al., 2013). However, due to the lack of evidence, there is a particular challenge to decide which FSTDs are confidently reasonable for JNS 1. This study considers the distribution of average FSTDs of modern humans as that of the archaic human and strongly suggests that FSTDs of the reliable regions along the normal vectors are appropriate for approximating the facial appearances of archaic humans.

It is worth mentioning that the FSTDs make a great contribution to improving the accuracy and reliability of the approximated face (Claes et al., 2006; Starbuck & Ward,

2007). It is acknowledged that many factors, sex, age, nutrition status, body mass index, ethnic groups will impact the FSTDs of landmarks. Previous studies always placed the landmarks on the skull manually and then acquired the FSTDs at these landmarks (Stephan, 2017). We investigated the effects of the distribution of FSTDs on FA. Approximated faces of the same JNS 1 using different FSTDs shared a resemblance. The larger deviation areas of two approximated faces between using FSTDs of the males and females were almost consistent with the regions of the FSTDs discrepancy between males and females (Shui et al., 2016). Once a reliable and verified craniofacial relationship can be obtained, a more reliable and accurate approximation would be recreated.

Because of the greater shape variations between the modern and archaic human skulls, it remains challenging to assign the FSTDs of the template to those of JNS 1. Previous studies often employed the deformation-based FA approach (Deng et al., 2011; Nelson & Michael, 1998; Turner et al., 2005). It is accepted that the closer the deformed template and the dry skulls matched, the more confidence there was in the reliability of the approximated face. However, the performance of the TPS deformation even incorporating a regularization is inadequate accuracy when two skulls have great geometric differences. To address this problem, this study employed a hybrid non-rigid registration algorithm to establish a high-quality set of geometric correspondences. Additionally, this approach can make use of FSTDs for every sample within the dataset to recreate a range of multiple approximated faces and then use PCA to construct a tailored approximation-space for JNS 1. In this context, the missing areas of the coarsely approximated face can be better repaired (Gietzen et al., 2019) and a range of possible approximation can be recreated by using appropriate coefficients of PCs of interests (Shui & Wu, 2018).

As previous studies mentioned (Oxnard & O'Higgins, 2009; Wärmländer et al., 2019), the anatomical landmarks and semilandmarks should be very carefully designed with regard to the research question. The purpose of this study is to approximate the overall shape of the facial appearance. Previous studies employed different numbers of landmarks (Deng et al., 2011; Vandermeulen et al., 2006) and topographic features, e.g. crest lines (Turner et al., 2005) to guide the deformation. But there is no standard criterion for performing FA. We recommend that landmarks and semilandmarks need to cover the entire skull, particularly around the region where the template and dry skulls are quite different. Such definitions will improve the certainties of the establishment of dense geometric correspondences, and so enhance the reliability of the approximated face. Additionally, we employed GM to capture the main features of the approximated faces. The dense corresponding vertices are used to provide better visualisation and interpretation, rather than the use of landmarks and semilandmarks. But due to the

complexity of biological structures, we need to carefully examine the effect of landmarks, semilandmarks and high-density correspondences on the shape analysis, and then decide which type of corresponding points can provide a reliable interpretation.

The geometric comparison of the actual and approximated faces has been used to validate the reliability of FA. In this process, the shell-to-shell deviation, surface-to-surface deviation and the shortest distance between two point clouds have been conducted (Miranda et al., 2018; Short et al., 2014; Wilkinson et al., 2006). But, the resulting registration impacts the comparisons of discrepancy between the approximated and actual faces. Additionally, since no verified actual faces can be provided, the comparison of the actual and approximated faces cannot be used to archaic humans. Thus, the reliability of the approximated face is mainly evaluated based upon the experience and knowledge of the experts. Even though the experts marked the areas of the approximated faces that need to be improved, researchers still do not know how to revise them exactly. To tackle this problem, we can integrate the approximation of facial appearance and recognition of the less confidence in approximated regions iteratively. This method provides a promising tool to allow researchers to examine the extent to which the resulting approximation is unreliable and where it needs to be further revised to improve the reliability.

However, uncertainty still remains since there are no actual archaic human faces and we cannot know the actual relationship between soft tissue and skull of archaic human. Also, it remains challenging to understand the relationship between facial features, e.g. eyes, nose, mouth, ears, etc., and bony structures. Thus, the approximated features need to be further improved, e.g. the size and orientation of nasal aperture. In addition, the ages and features of Tabun 2 and Mauer 1 are different from JNS 1, and an appropriate mandible needs to be used to enhance the reliability of the approximated face.

6.5 Conclusion

The approximated face of JNS 1 is a typical case of interdisciplinary study that provides anthropologists and the general public with an improved visual interpretation of the facial morphology of archaic human. This study proposed a coarse-to-fine FA approach based on dense FSTDs of modern humans and presented an evaluation approach to validate the reliability of FA. We also investigated the effects of skull morphology and the distribution of FSTDs on the approximated faces. Since the mandibles of archaic humans are rarely found, it remains challenging to select an appropriate mandible for the JNS 1 cranium. In the future, we will attempt to collect the different mandibles of archaic humans to improve the reliability of the approximation of JNS 1.

6.6 Acknowledgements

I would like to thank Beijing Key Laboratory of Digital Preservation and Virtual Reality for Cultural Heritage, Beijing Normal University for granting access to their human skull and face database. I acknowledge and thank Dr. Antonio Profico for providing access to the Mauer 1 mandible. Reproduced with permission from Springer Nature.

6.7 References

- Andrews, P. (1986). Fossil evidence on human origins and dispersal. The *Cold Spring Harbor symposia on quantitative biology*, 419-428.
- Athreya, S, & Wu, X. (2017). A multivariate assessment of the Dali hominin cranium from China: Morphological affinities and implications for Pleistocene evolution in East Asia. *American Journal of Physical Anthropology*, 164(4), 679-701.
- Benazzi, S, Fantini, M, Crescenzo, FD, Mallegni, G, Mallegni, F, Persiani, F, & Gruppioni, G. (2009). The face of the poet Dante Alighieri reconstructed by virtual modelling and forensic anthropology techniques. *Journal of Archaeological Science*, 36(2), 278-283.
- Berar, M, Tilotta, FM, Glaunès, JA, & Rozenholc, Y. (2011). Craniofacial reconstruction as a prediction problem using a Latent Root Regression model. *Forensic Science International*, 210, 228-236.
- Bookstein, FL. (1989). Principal warps: Thin-plate splines and the decomposition of deformations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(6), 567-585.
- Claes, P, Vandermeulen, D, De Greef, S, Willems, G, Clement, JG, & Suetens, P. (2010). Computerized craniofacial reconstruction: Conceptual framework and review. *Forensic Science International*, 201(1-3), 138-145.
- 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.
- De Buhan, M, & Nardoni, C. (2018). A facial reconstruction method based on new mesh deformation techniques. *Forensic Sciences Research*, 3(3), 256-273.
- Deng, Q, Zhou, M, Shui, W, Wu, Z, Ji, Y, & Bai, R. (2011). A novel skull registration based on global and local deformations for craniofacial reconstruction. *Forensic Science International*, 208(1), 95-102.
- Deng, Q, Zhou, M, Wu, Z, Shui, W, Ji, Y, Wang, X, Liu, CY, et al. (2016). A regional method for craniofacial reconstruction based on coordinate adjustments and a new fusion strategy. *Forensic Science International*, 259, 19-31.
- Fatuzzo, G, Sequenzia, G, & Oliveri, SM. (2016). Virtual anthropology and rapid prototyping: a study of Vincenzo Bellini's death masks in relation to autopsy documentation. *Digital Applications in Archaeology and Cultural Heritage*, 3(4), 117-125.
- Gietzen, T, Brylka, R, Achenbach, J, zum Hebel, K, Schömer, E, Botsch, M, Schwanecke, U, et al. (2019). A method for automatic forensic facial reconstruction

based on dense statistics of soft tissue thickness. *Plos one*, 14(1), e0210257.

Gunz, P, & Mitteroecker, P. (2013). Semilandmarks: a method for quantifying curves and surfaces. *Hystrix, the Italian Journal of Mammalogy*, 24(1), 103-109.

Gunz, P, Mitteroecker, P, Neubauer, S, Weber, GW, & Bookstein, FL. (2009). Principles for the virtual reconstruction of hominin crania. *Journal of Human Evolution*, 57(1), 48-62.

Hamre, SS, Erstrand, GA, Daux, V, Parson, W, & Wilkinson, C. (2017). Three individuals, three stories, three burials from medieval Trondheim, Norway. *Plos one*, 12(7), e0180277.

Harvati, K, & Lopez, EN. (2017). A 3-D look at the Tabun C2 jaw. In *Human paleontology and Prehistory*, 203-213: Springer.

Hayes, S. (2016). Faces in the museum: revising the methods of facial reconstructions. *Museum Management and Curatorship*, 31(3), 218-245.

Hayes, S, Bradley, R, Milne, N, & Dennison, J. (2012). Approximating the Face of 'Aunty': A Question of Likeness. *Journal of Archaeological Method & Theory*, 19(2), 306-321.

Hayes, S, Shoocongdej, R, Pureepatpong, N, Sangvichien, S, & Chintakanon, K. (2017). A Late Pleistocene woman from Tham Lod, Thailand: the influence of today on a face from the past. *Antiquity*, 91(356), 289-303.

Hayes, S, Sutikna, T, & Morwood, M. (2013). Faces of Homo floresiensis (LB1). *Journal of Archaeological Science*, 40(12), 4400-4410.

Herrera, RJ, & Garcia-Bertrand, R. (2018). *Ancestral DNA, Human Origins, and Migrations*: Academic Press.

Hublin, JJ. (2013). The Middle Pleistocene record: On the ancestry of Neandertals, modern humans and others. In *A Companion to Paleoanthropology*, 517-537: Blackwell Publishing Ltd.

Huempfer-Hierl, H, Bohne, A, Schaller, A, Wollny, G, & Hierl, T. (2015). Does facial soft tissue protect against zygomatic fractures? Results of a finite element analysis. *Head Face Medicine*, 11(1), 1-8.

Kustar, A. (2004). The facial restoration of Antal Simon, a Hungarian priest-teacher of the 19th c. *HOMO-Journal of Comparative Human Biology*, 55(1-2), 77-90.

Lee, WJ, Yoon, AY, Mi, KS, Wilkinson, CM, & Dong, HS. (2014). The archaeological contribution of forensic craniofacial reconstruction to a portrait drawing of a Korean historical figure. *Journal of Archaeological Science*, 49, 228-236.

- Lorensen, WE, & Cline, HE. (1987). Marching cubes: A high resolution 3D surface construction algorithm. *ACM Siggraph Computer Graphics*, 21(4), 163-169.
- Marić, J, Bašić, Ž, Jerković, I, Mihanović, F, Anđelinović, Š, & Kružić, I. (2020). Facial reconstruction of mummified remains of Christian Saint-Nicolosa Bursa. *Journal of Cultural Heritage*, 42, 249-254.
- Martin, R. (1928). *Lehrbuch der Anthropologie in systematischer Darstellung*. In: Stuttgart: Gustav Fischer Verlag.
- Miranda, GE, Wilkinson, C, Roughley, M, Beaini, TL, & Melani, R. (2018). Assessment of accuracy and recognition of three-dimensional computerized forensic craniofacial reconstruction. *Plos One*, 13(5), e0196770.
- Mounier, A, Marchal, F, & Condemi, S. (2009). Is Homo heidelbergensis a distinct species? New insight on the Mauer mandible. *Journal of Human Evolution*, 56(3), 219-246.
- Nelson, L, & Michael, S. (1998). The application of volume deformation to three-dimensional facial reconstruction: a comparison with previous techniques. *Forensic Science International*, 94(3), 167-181.
- O'Higgins, P, & Jones, N. (1998). Facial growth in *Cercocebus torquatus*: an application of three-dimensional geometric morphometric techniques to the study of morphological variation. *Journal of Anatomy*, 193(2), 251-272.
- Oxnard, C, & O'Higgins, P. (2009). Biology clearly needs morphometrics. Does morphometrics need biology? *Biological Theory*, 4(1), 84-97.
- Paysan, P, Lüthi, M, Albrecht, T, Lerch, A, Amberg, B, Santini, F, & Vetter, T. (2009). Face Reconstruction from Skull Shapes and Physical Attributes. The *Lecture Notes in Computer Science, Pattern Recognition*, 232-241.
- Quam, RM, & Smith, FH. (2002). A reassessment of the Tabun C2 mandible. In *Neandertals and modern humans in Western Asia*, 405-421: Springer.
- Quatrehomme, G, Cotin, S, Subsol, G, Delingette, H, Garidel, Y, Grévin, G, Fidrich, M, et al. (1997). A fully three-dimensional method for facial reconstruction based on deformable models. *Journal of Forensic Science*, 42(4), 649-652.
- Rightmire, GP. (1998). Human evolution in the Middle Pleistocene: the role of Homo heidelbergensis. *Evolutionary Anthropology*, 6(6), 218-227.
- Rosenberg, KR, Zuné, L, & Ruff, CB. (2006). Body size, body proportions, and encephalization in a Middle Pleistocene archaic human from northern China. *Proceedings of the National Academy of Sciences*, 103(10), 3552-3556.
- Schwartz, JH, & Tattersall, I. (2000). The human chin revisited: what is it and who has

it? *Journal of Human Evolution*, 38(3), 367-409.

Short, LJ, 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. e141-147. e148.

Shui, W, & Wu, X. (2018). Three-dimensional craniofacial reconstruction of an ancient Qihe human skull (in Chinese) *Chinese Science Bulletin*, 63(8), 745-754.

Shui, W, Zhou, M, Deng, Q, Wu, Z, Ji, Y, Li, K, He, T, et al. (2016). Densely Calculated Facial Soft Tissue Thickness for Craniofacial Reconstruction in Chinese Adults. *Forensic Science International*, 266, 573.e571-e512.

Shui, W, Zhou, M, Maddock, S, He, T, Wang, X, & Deng, Q. (2017). A PCA-Based method for determining craniofacial relationship and sexual dimorphism of facial shapes. *Computers in Biology and Medicine*, 90, 33-49.

Shui, W, Zhou, M, Maddock, S, Ji, Y, Deng, Q, Li, K, Fan, Y, et al. (2020). A computerized craniofacial reconstruction method for an unidentified skull based on statistical shape models. *Multimedia Tools and Applications*, 79, 25589-25611.

Starbuck, JM, & Ward, RE. (2007). The affect of tissue depth variation on craniofacial reconstructions. *Forensic science international*, 172(2), 130-136.

Stephan, CN. (2017). 2018 tallied facial soft tissue thicknesses for adults and sub-adults. *Forensic Science International*, 280, 113-123.

Turner, WD, Brown, RE, Kelliher, TP, Tu, PH, Taister, MA, & Miller, KW. (2005). A novel method of automated skull registration for forensic facial approximation. *Forensic Science International*, 154(2-3), 149-158.

Vandermeulen, D, Claes, P, Loeckx, D, Greef, SD, Willems, G, & Suetens, P. (2006). Computerized craniofacial reconstruction using CT-derived implicit surface representations. *Forensic Science International*, 159(1), S164-S174.

Vanezis, P, Blowes, R, Linney, A, Tan, A, Richards, R, & Neave, R. (1989). Application of 3-D computer graphics for facial reconstruction and comparison with sculpting techniques. *Forensic Science International*, 42(1-2), 69-84.

Verzé, L. (2009). History of facial reconstruction. *Acta Biomedica* 80, 5-12.

Wagner, GA, Krbetschek, M, Degering, D, Bahain, J-J, Shao, Q, Falguères, C, Voinchet, P, et al. (2010). Radiometric dating of the type-site for *Homo heidelbergensis* at Mauer, Germany. *Proceedings of the National Academy of Sciences*, 107(46), 19726-19730.

Wärmländer, SK, Garvin, H, Guyomarc'h, P, Petaros, A, & Sholts, SB. (2019). Landmark typology in applied morphometrics studies: What's the point? *The*

Anatomical Record, 302(7), 1144-1153.

Wilkinson, C. (2005). Computerized forensic facial reconstruction. *Forensic Science, Medicine, and Pathology*, 1(3), 173-177.

Wilkinson, C. (2010). Facial reconstruction--anatomical art or artistic anatomy? *Journal of Anatomy* 216(2), 235-250.

Wilkinson, C, Rynn, C, Peters, H, Taister, M, Kau, CH, & 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(2), 179-187.

Wu. (1988). The Reconstruction of the fossil human skull from Jinniushan, Yingkou, Liaoning Province, and its main features (in Chinese). *Acta Anthropologica Sinica*, 7, 97-102.

Yu, W, Li, M, & Li, X. (2012). Fragmented skull modeling using heat kernels. *Graphical Models*, 74(4), 140-151.

Supplementary material

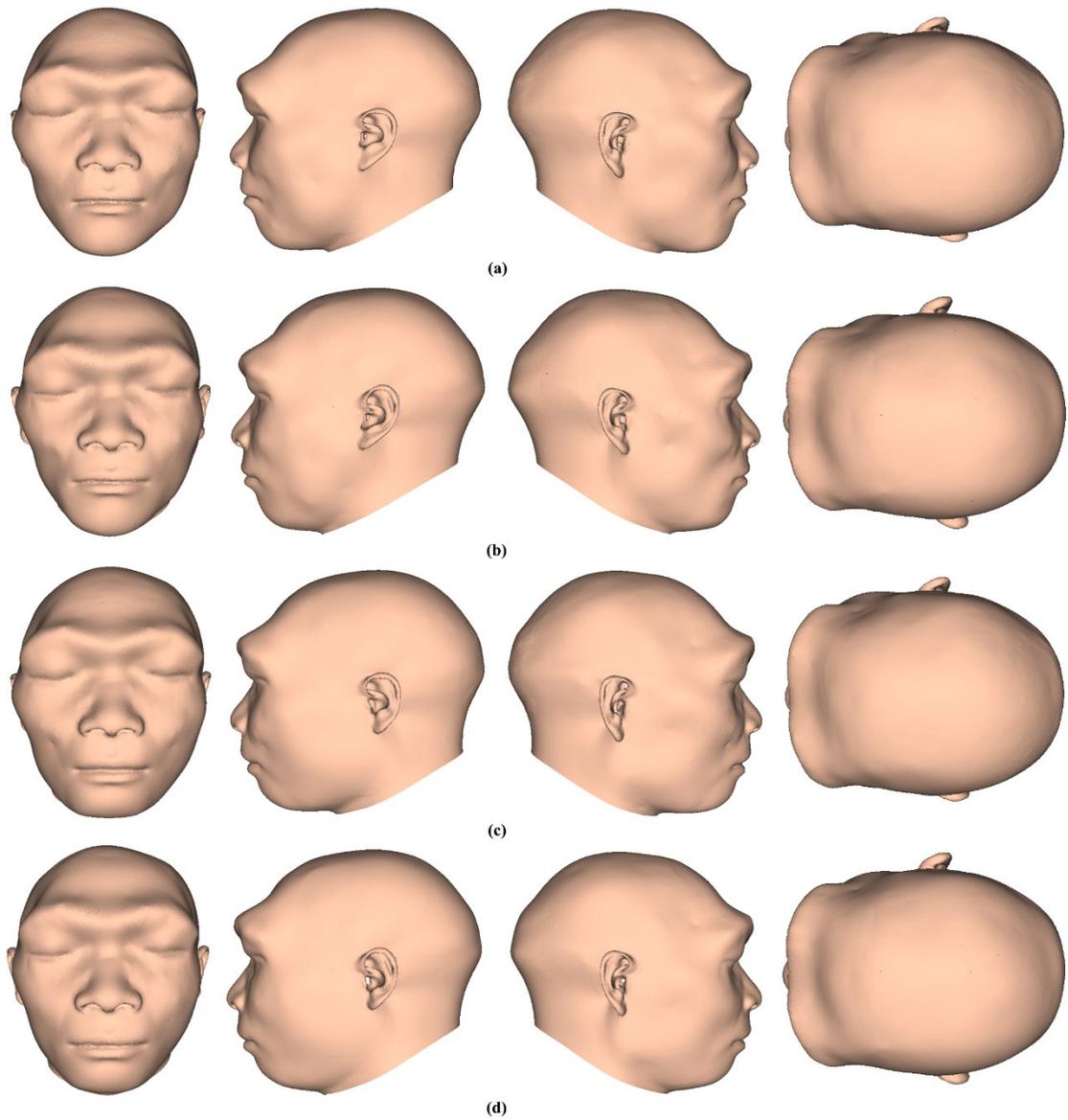


Figure 6.S1 Different approximated faces of JNS 1 based on different FSTDs and mandibles. (a) The approximated face that used the average FSTDs of the females and the Tabun 2 mandible. (b) The approximated face that used the average FSTDs of the males and the Tabun 2 mandible. (c) The approximated face that used the average FSTDs of the females and the Mauer 1 mandible. (d) The approximated face that used the average FSTDs of the males and the Mauer 1 mandible.

Chapter 7 A computerized facial approximation method for *Homo sapiens* based on facial soft tissue thickness depths and geometric morphometrics

Abstract: Facial approximation (FA) provides a promising means of generating the possible facial appearance of a deceased person. It facilitates exploration of the evolutionary forces driving anatomical changes in ancestral humans and can capture public attention. Despite the recent progress made toward improving the performance of FA methods, a limited understanding of detailed quantitative craniofacial relationships between facial bone and soft tissue morphology may hinder accuracy, and hence subjective experience and artistic interpretation are required. In this study, we proposed a computerized method for exploring craniofacial relationships among human populations and assigning the learnt craniofacial relationships to generate a probable facial appearance of *Homo sapiens*. The craniofacial relationships were dependent upon average facial soft tissue thickness depths (FSTDs) and covariations between nasal and oral hard and soft tissue morphological features using geometric morphometrics. A smaller resemblance comparison (average Procrustes distance was 0.0258 and geometric difference was 1.79 mm) between approximated and actual faces and greater recognition rate (91.67%) tested by a face pool indicated average dense FSTDs contributed to raising the accuracy of approximated faces. Results of partial least squares (PLS) analysis showed that nasal (or oral) hard tissues have an effect on the soft tissue shapes. However, relatively weak covariations (<0.4) and greater approximation errors suggested that we need to be cautious about the accuracy of approximated nose and mouth soft tissue shapes from bony structures alone. The proposed method was applied to approximate the facial appearance of the Upper Cave (UC) 101 skull, an important 30,000-year-old human skull excavated from the village of Zhoukoudian in northern China. Overall, the proposed method can facilitate investigations of craniofacial relationships and potentially improve the reliability of the approximated faces for use in numerous applications in forensic science, archaeology, and anthropology.

Article details:

Wuyang Shui, Xiujie Wu and Minquan Zhou. A computerized facial approximation method for *Homo sapiens* based on facial soft tissue thickness depths and geometric morphometrics. *Journal of Anatomy*. (under review)

At the time of submission, this chapter was pending. It was published by *Journal of Anatomy* on June 27, 2023.

See the link: <https://onlinelibrary.wiley.com/doi/epdf/10.1111/joa.13920>.

7.1 Introduction

Facial approximation (FA, or known as facial reconstruction) is a commonly used technique to recreate the probable facial appearance of a deceased person from skeletal remains. In the absence of other clues, it has served as a method of last resort to recover possible facial likenesses for use in triggering memories to aid forensic identification of a seriously decomposed cadaver (Baldasso et al., 2021; Nelson & Michael, 1998). It also seeks to recreate sculptural portraits of undocumented *Homo sapiens*, e.g. famous historical figures and our recent ancestors, in the realms of archaeology and anthropology (Benazzi et al., 2009; Marić et al., 2020). The approximated faces provide new insights in understanding the characteristic features of human fossils, exploring the evolutionary forces driving anatomical changes in ancestral humans and capturing public attention. Until recently, FA methods for creating three-dimensional (3D) facial appearances have involved manual, virtual sculpture and computer-based techniques.

The conventional 3D manual FA method is primarily based upon clay sculpturing techniques over skull casts, which has been used to approximate hundreds of facial sculptures of *Homo sapiens* (Hayes, 2016). This method can be divided into three main approaches (Verzé, 2009): the American method based on facial soft tissue thickness depths (FSTDs) at landmarks, the Russian method based on muscle knowledge, and the combined Manchester method that is the most commonly used in forensic and archaeology fields. However, it is argued the results are dependent upon anatomical interpretation and artistic skill of the individual practitioner (Campbell et al., 2021; Stephan, 2015). For example, subjective and artistic interpretation may lead to biased and inconsistent results, and hence the practitioner requires anthropological and artistic training and experience in practice. Additionally, the manual method is a time-consuming task, especially in the production of multiple possible faces of the same skull (Claes et al., 2010).

To address shortcomings of the manual method, a 3D interactive computerized method was developed by graphics software (e.g. ZBrush and Blender) and haptic feedback devices, mimicking the conventional Manchester method in the computer environment (Wilkinson et al., 2006). It has been successfully used in different case studies in archaeology and forensic science (Hamre et al., 2017; Lee et al., 2014). In practice, several virtual pegs representing FSTDs are attached to the surface of the digital skull at landmarks and then pre-modelled facial muscle models are interactively placed onto the skull. A critical point is to allow altered muscles to meet the specifications of the skull following anatomical guidelines. Subsequently, the appropriate facial features (e.g. eyes, nose, mouth and ears) are chosen from a dataset to place over the facial model. Finally, skin layers are added over the muscle structures and the detailed morphology and features were sculpted. A small geometric difference between the approximated and

actual faces has verified the accuracy and effectiveness (Lee et al., 2015; Miranda et al., 2018). Such a method offers a workflow standardization and the pre-modelled muscles and organs provide valuable anatomical guidance to reduce approximation bias and subjective interpretation. Nevertheless, such a system still requires subjective interpretation and modelling skill training.

Progress in the development of medical imaging technologies, such as computed tomography (CT), cone-beam computed tomography (CBCT) and magnetic resonance imaging (MRI), has led to the development of semi- and fully-automatic computerized FA methods (Claes et al., 2010; Wilkinson, 2005). Moreover, the accuracy of each method can be quantitatively assessed by computing the geometric differences between the approximated and actual faces. The earliest proposed 3D computerized FA efforts involved the deformation of the template face, e.g. an average human face, to fit the estimates of facial points which were generated by assigning average FSTDs at landmarks to the dry skull (Vanezis et al., 2000). Due to the fact that limited numbers of facial points cannot adequately capture detailed morphology, the approximated face might be inaccurate. To address this problem, a recent study has employed FSTDs at landmarks and high-density semilandmarks to build a facial envelope and then recreate a detailed facial appearance by deforming the template face (Shui et al., 2021). To avoid huge computation costs of dense FSTDs, an alternative deformation-based computerized FA method has been developed based on deformation of the template face followed by transformation to warp the template skull to the dry skull (De Buhan & Nardoni, 2018; Quatrehomme et al., 1997; Turner et al., 2005). A non-rigid registration algorithm, e.g. thin-plate splines (TPS) (Bookstein, 1989), was carried out to align these two skulls based on a set of landmarks or geometric features. Nonetheless, the ultimate approximation result might resemble the template face and a large skull registration error adversely affects the accuracy of the approximation (Deng et al., 2011; Wilkinson, 2005).

Since a large collection of skull and face models facilitates the investigation of detailed craniofacial relationships amongst human populations to achieve the anatomical modelling, several recent studies used machine learning algorithms, e.g. multiple regression, to learn the relationships of principal component (PC) scores between hard and soft tissues and then applied them to the dry skull to generate the ultimate approximation (Berar et al., 2011; Jia et al., 2021; Madsen et al., 2018; Shui et al., 2017). Furthermore, the craniofacial relationships associated with personal attribute information, such as age, sex and body mass index, were explored to improve the accuracy of FA. Nevertheless, facial morphology is complex and the craniofacial relationships of different regions are inconsistent, such that a global regression-based method cannot provide detailed relationships in facial features. To tackle the issue, covariations between bony and soft tissue structures of each component was regressed separately (Deng et al., 2016;

Guyomarc'h et al., 2014). Alternatively, a facial statistical shape model can facilitate the prediction of feature features and missing geometry over the facial envelope (Gietzen et al., 2019). Additionally, a recent deep learning method has been proposed to generate craniofacial images from 2D skull images, thereby achieving 3D FA (Hu et al., 2021; Li et al., 2022; Zhang et al., 2022).

The Upper Cave 101 (UC101) skull, which dates back to almost 30,000 years BP, was discovered in the upper cave of Zhoukoudian in northern China in 1930. As one of the most important East Asian human fossils, this specimen has been used to investigate morphological changes associated with human evolution and origins. Morphological analysis of UC 101 has revealed it to be derived from an adult male with shared ancestry with members of primitive Mongoloid, Easter Island or Upper Paleolithic European populations (Cunningham & Jantz, 2003). It exhibits a longer and lower cranial vault, a broader forehead, a more pronounced superciliary arch and a higher nasal bridge compared with the corresponding features of modern humans. The approximated face facilitates the investigation of human evolutionary processes, enhances scientific knowledge of the factors underlying the development of modern humans and captures public attention. More recently, we developed a computerized method to repair the damaged zygomatic arch and approximate the facial appearance based on dense FSTDs of modern humans (Shui et al., 2020a). In that work, average dense FSTDs were assigned to UC 101 to approximate the overall facial envelope. However, the resulting approximated facial features were predominantly derived from the template face, ignoring morphological traits of the skull.

Anatomical modelling to represent craniofacial relationships between hard and soft tissues, and determination of facial features based on bony morphology are the fundamental steps in enhancing the accuracy of FA. In this study, we present a computerized FA method for exploring detailed craniofacial relationships between bony structures and facial soft tissues and afterwards developing a coarse-to-fine method to generate a probably facial appearance. A quantitative method based on resemblance comparison between approximated and actual faces and recognition rate tested by a face pool was used to validate the accuracy of the proposed method.

7.2 Materials and Methods

7.2.1 Materials

7.2.1.1 A skull and face dataset of modern living humans

In our previous studies (Deng et al., 2016), we constructed a skull and face dataset of modern living humans exhibiting normal morphological features without prior orthodontic treatment. To ensure consistency and accuracy in the measurements, each participant kept their mouth closed throughout the data acquisition process. Medical

images of each individual were acquired using a clinical multi-slice CT scanner system (Siemens Sensation 16) then CT images were archived as standard DICOM 3.0 files of resolution 512×512 . Each head within the database comprised digital models of the skull and the corresponding face, and personal information (e.g. sex, age, ethnicity, etc.). Data processing procedures were as follows: first, we converted CT images of each individual to digital models of skull and face and then extracted external surfaces of skull and face, respectively. Each skull comprised more than 160,000 vertices and 300,000 triangle meshes, and each face consisted of more than 200,000 vertices and 500,000 triangle meshes. Subsequently, we manually placed skull landmarks including left porion (*Lp*), right porion (*Rp*), left orbitale (*Lo*) and glabella (*G*), and transformed them into the common Frankfort coordinate system. Thereafter, we estimated average skull and facial form (including shapes and sizes) surfaces. In the current study, 48 male adults aged 20-30 years were selected to explore craniofacial relationships and test performance of the proposed method.

Next, we established high-density point correspondences among skulls. The approach comprised four steps: first, we manually separated the average skull of the male group within the dataset into three components: oral and nasal hard tissues, and bony envelope. Second, the Poisson-disk algorithm (Corsini et al., 2012) was used to obtain 14,933 semilandmarks from the average bony envelope (excluding the temporal bone mastoid process and acoustic canal) in order to reduce noise and numbers of outliers around approximated ear soft tissues when assigning FSTDs to semilandmarks of the dry skull. Additionally, a total of 556 and 215 semilandmarks were sampled from oral and nasal hard tissues, respectively. Third, according to the definition used in our previous study (Shui et al., 2021), 91 landmarks were placed over the estimates of average skull and every skull within the dataset. Finally, we used the hybrid non-rigid approach combining TPS and non-rigid iterative closest point (NICP) (Amberg et al., 2007) to align each skull to the average skull, and then projected semilandmarks of each component of the deformed average skull onto each skull to yield dense point correspondences. All these methods were programmed using Matlab (Please see Chapters 2 and 3).

Additionally, we established high-density point correspondences among faces using the same four steps: first, we manually partitioned the average face into three components: mouth, nose soft tissues and other facial envelope. Second, the Poisson-disk sampling algorithm was used to obtain 498, 497 and 150,000 semilandmarks from each component, respectively. Third, a total of 52 anatomical landmarks in total were chosen on every face (Table 7.1), where 8 landmarks were located on the midline and 22 landmarks were bilaterally located. Finally, the average face was deformed to align with surfaces of every face and then semilandmarks of each component of the deformed average face were projected onto each face in order to establish high-density correspondences.

Based on the dense point correspondences among skulls and faces, we estimated the average skull and face forms (shape and size). Figure 7.1a shows the average skull and the landmark and semilandmark configuration. The leftmost two figures show the average skull including bony envelope (grey), nasal (red) and oral (blue) hard tissue forms. The other three figures display landmarks (black) and semilandmarks of skull envelope (yellow), nasal (red) and oral (blue) hard tissue forms. Likewise, Figure 7.1b shows each component of the average face and the landmark and semilandmark configuration.

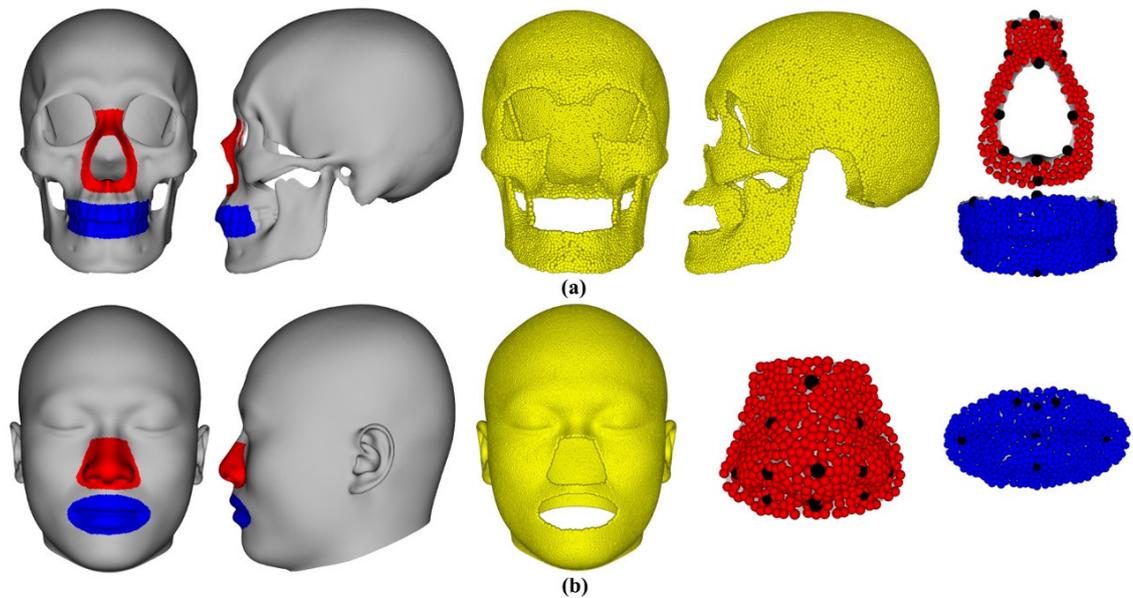


Figure 7.1 Average skull and face including the landmark and semilandmark configuration. (a) Skull envelope (grey), nasal (red) and oral (blue) hard tissue forms including landmarks (black) and semilandmarks of each component. (b) Facial envelope (grey), nasal (red) and oral (blue) soft tissue forms including landmarks (black) and semilandmarks of each component.

Table 7.1 Anatomical facial landmarks

No.	Definition	Nature	Position
1 and 2	Medial canthus	Bilateral	Eyes
3 and 4	Lateral canthus	Bilateral	Eyes
5	Nasal bridge	Median	Nose
6	Middle of nose	Median	Nose
7	Tip of nose	Median	Nose
8	Subnasale	Median	Nose
9 and 10	External alar curvature	Bilateral	Nose
11 and 12	Superior alar curvature	Bilateral	Nose
13 and 14	Alare	Bilateral	Nose
15 and 16	Alar curvature point	Bilateral	Nose
17 and 18	Corner of mouth	Bilateral	Mouth
19 and 20	crista philtra	Bilateral	Mouth
21	Middle of cupid's bow upper lip	Median	Mouth
22	Middle of oral fissure	Median	Mouth

23	Middle of bottom lip	Median	Mouth
24	Tip of chin	Median	Chin
25 and 26	Otobasion superius	Bilateral	Ears
27 and 28	Superior auricle	Bilateral	Ears
29 and 30	Posterior auricle	Bilateral	Ears
31 and 32	Inferior auricle	Bilateral	Ears
33 and 34	Anterior cymba concha	Bilateral	Ears
35 and 36	Superior cymba concha	Bilateral	Ears
37 and 38	Posterior concha	Bilateral	Ears
39 and 40	Intertragic incisure	Bilateral	Ears
41 and 42	Incisura intertragica	Bilateral	Ears
43 and 44	Tragion	Bilateral	Ears
45 and 46	Medial concha	Bilateral	Ears
47 and 48	Superior cavum concha	Bilateral	Ears
49 and 50	Otobasion posterius	Bilateral	Ears
51 and 52	Otobasion inferius	Bilateral	Ears

7.2.1.2 Digitisation of UC 101

A portable 3D scanner was employed to capture a digital model of a high-quality cast of UC 101 archived at the Institute of Vertebrate Paleontology and Paleoanthropology (IVPP) in Beijing, China. Subsequently, small discrete and self-intersecting meshes were removed and remaining meshes were smoothly processed. We repeated abovementioned semilandmarking method to align the average skull to the UC 101 meshes based on skull landmarks and then established high-density point correspondences. Because missing dental crowns existed, we labelled the correspondences locating at the intact surface, thereby establishing point correspondences among modern human skulls. Figure 7.2a shows the external surface. Figure 7.2b displays the semilandmarks of skull envelope and nasal hard tissues. Figure 7.2c shows semilandmarks of teeth in UC 101 (left) and average skull (right), respectively.

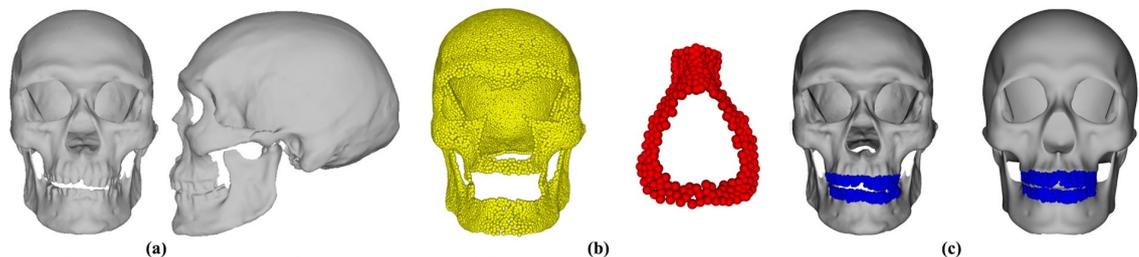


Figure 7.2 Semilandmarks of UC 101. (a) The external surface. (b) Semilandmarks of skull envelope (yellow) and nasal (red) hard tissues. (c) Semilandmarks of teeth (blue) in UC 101 (left) and average skull (right).

7.2.2 Methods

The proposed FA method includes two stages (Figure 7.3): in the first stage, we explored craniofacial relationships of modern humans following three steps: first, we

calculated average FSTDs at landmarks and high-density semilandmarks among human populations to represent overall craniofacial relationships. Second, the covariation of landmarks and semilandmarks between nasal hard and soft tissue shapes was explored using geometric morphometrics, thereby approximating nose soft tissue shapes. Third, we repeated the abovementioned step to explore the covariation between oral hard and soft tissue shapes and to generate mouth shapes. In the second stage, we presented a coarse-to-fine method to generate the facial appearance. First, the coarsely approximated face was generated by reassembling approximated facial envelope, nose and mouth shapes depended upon the learnt craniofacial relationships. Second, the improved approximated face was generated by fitting the facial statistical shape model (SSM) to the coarse approximation. Finally, we used resemblance comparison and recognition rate to evaluate the accuracy of the proposed method. All these methods were programmed using C++ and Matlab and software prototype can be downloaded from the Internet (<https://github.com/sissun/FAModern.git>).

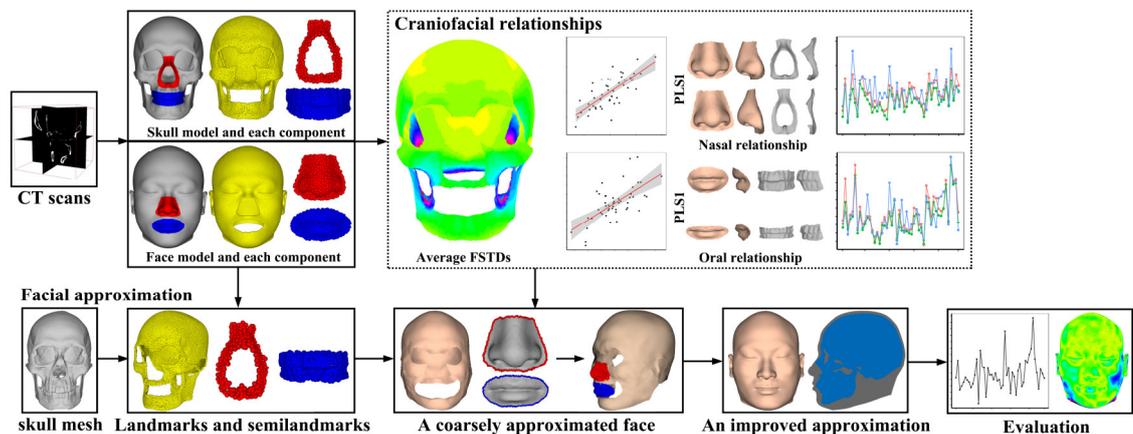


Figure 7.3 Workflow of quantification of craniofacial relationships and FA method.

7.2.2.1 Quantification of craniofacial relationships among human populations

7.2.2.1.1 Average dense FSTDs

FSTDs were defined as Euclidean distances between skull landmarks (or semilandmarks) and corresponding facial points. In this study, the average FSTDs among humans were used to represent overall craniofacial relationships. We used the normal vector (Toneva et al., 2021) of every landmark and semilandmark of the average skull as the measured orientation and then employed the same constant orientation to record FSTDs across samples. Next, descriptive statistics of FSTDs among samples (e.g. mean and standard deviation) were calculated and a colour map was used to show distributions of FSTDs.

Thereafter, average FSTDs at landmarks and semilandmarks were assigned to external surface of the dry skull to obtain a facial envelope, as follows:

$$h_i = s_i + \overline{n}_i \cdot d_i \quad (7.1)$$

where $s_i \in \mathbb{R}^3$ represents geometric coordinates of i -th landmarks and semilandmarks of the dry skull, $\{h_i\} \in \mathbb{R}^3$ denotes the facial points that can be used to generate the facial envelope. \overline{n}_i and d_i represent the measured vector and average FSTDs, respectively. These facial points were then converted to a set of triangle meshes by the ball-pivoting algorithm (Bernardini et al., 1999).

7.2.2.1.2 Quantification of relationships between nasal (or oral) hard and soft tissue shapes

To explore relationships between nasal (or oral) hard and soft tissue shapes, a generalized Procrustes analysis (GPA) algorithm was applied to the landmark and semilandmark configurations across samples, removing location, orientation and scaling. Next, a two-block partial least squares (2B-PLS) analysis (Rohlf & Corti, 2000) was used to explain the covariation between Procrustes shape coordinates of nasal (or oral) hard and soft tissue shapes using the R package “geomorph” (Adams & Otárola-Castillo, 2013), with Block-1 defined here as hard tissue shapes and Block-2 defined as soft tissue shapes. Unlike principal component analysis (PCA), 2B-PLS maximizes covariance patterns, as derived from the cross-covariance matrix, to produce pairs of component axes. Subsequently, correlation coefficients and significance were determined based on 1000 random permutations for each PLS dimension in order to quantify the level of covariation. Additionally, RV coefficients (Robert & Escoufier, 1976) were used to measure overall associations between the two matrices as generated using PLS scores derived from two different configurations by qualifying the multivariate extension of the correlation coefficient. To examine morphological variations of shapes representing the extreme limits of PLS of interest, fitted landmarks and semilandmarks along the positive (+) and negative (-) limits of PLS 1 were generated and then fitted shapes were generated by warping the mean shape using TPS. Regional variations of two different fitted shapes were found by observing main patterns of movements of landmarks and semilandmarks.

The next step was to approximate nasal (or oral) soft tissue shapes of the dry skull. To reduce data dimensionality, PCA was first carried out to project Procrustes shape coordinates of every sample onto the shape space, whereby PC scores were used to quantify shape variations of hard and soft tissue shapes. Subsequently, a linear multiple regression was used to quantify relationships between PC scores of hard (independent variable) and soft tissue shapes (dependent variable) as follows:

$$\mathbf{M}_{organ} = \arg \min \left\| \mathbf{n}\boldsymbol{\beta}_h \cdot \mathbf{M}_{organ} - \mathbf{n}\boldsymbol{\beta}_s \right\|^2 + \eta_{organ}^2 \left\| \mathbf{M}_{organ} \right\|^2 \quad (7.2)$$

where $\mathbf{n}\boldsymbol{\beta}_h$ and $\mathbf{n}\boldsymbol{\beta}_s$ represent PC scores of nasal (or oral) hard and soft tissues, respectively, and \mathbf{M}_{organ} denotes the craniofacial relationships. η_{organ} denotes the weighting

coefficient that is related to the number and standard deviation of PCs.

Based on the learnt relationships and the calculated PC scores derived from landmark and semilandmarks of hard tissue, coefficients of approximated nose (or mouth) soft tissue shapes were computed. Next, geometric coordinates of approximated landmarks and semilandmarks of soft tissues were obtained by adding the mean shape of nose (or mouth) soft tissues among samples to the linear combination of the computed coefficients and PCs.

Here we used leave-one-out cross validation (LOOCV), i.e. one sample within the dataset was selected for use as test data, while the other samples served as training data, to validate the accuracy of nose (or mouth) relationship and approximating results. This process was then repeated with each sample used once as the test data. The Procrustes distance between the approximated and actual shapes was calculated to quantify gross difference. The ratio of the Procrustes distance divided by the average Procrustes distance between soft tissues of every sample and mean shape was calculated to examine the degree of difference. Additionally, nose (or mouth) soft tissue shapes were generated by deforming the mean shape using TPS. Geometric differences in landmarks and semilandmarks between the approximated and actual nose (or mouth) shapes were transferred to the shapes to generate a colour map for use in comparing regional shape differences.

7.2.2.2 Facial approximation

7.2.2.2.1 A coarsely approximated face

Once landmarks and semilandmarks of the dry skull were generated, we were able to approximate the facial envelope based on average FSTDs and recreate the probable nose and mouth shapes based on the learnt nose and mouth relationships, respectively. However, they have different locations, orientations and size. A critical step involved the calculation of the transformation (translation, rotation and scaling) of approximated nose and mouth soft tissue shapes to fit them onto the approximated facial envelope, thereby generating the coarsely approximated faces.

The method used to compute such alignments comprised three steps: first, we used a deformation-based approach to warp the average facial envelope followed by transformation to align the average skull with the dry skull. Second, boundary curves of the deformed facial envelope with approximated nasal and mouth shapes were extracted, respectively. Finally, we used the least-squares algorithm to register approximated soft tissues of each component by minimizing Euclidean distances between boundary curves of the facial envelope and nose (or mouth) shapes as follows:

$$\arg \min_{\mathbf{R}, \mathbf{t}, \mathbf{S}} \sum_{i=1}^m \left\| q_i - (\mathbf{R} \cdot \mathbf{S} \cdot p_i + \mathbf{t}) \right\|^2 \quad (7.3)$$

where $p_i \in \mathbb{R}^3$ and $q_i \in \mathbb{R}^3$ represent corresponding boundary vertices of the facial envelope and nose (or mouth) shapes, m denotes the number of correspondences and $\mathbf{R} \in SO(3)$, $\vec{t} = (t_x, t_y, t_z)$ and S represent rotation, translation and uniform scaling matrices, respectively.

Figures 7.4a and b show the deformed average facial envelope with boundary curves and approximated nose (red) and mouth (blue) shapes with boundary curves, respectively. Figure 7.4c shows the superimposition of the deformed average facial envelope (grey) onto the transformed approximated nose (red) and mouth (blue) soft tissue forms. Figure 7.4d shows the coarsely approximated face comprising approximated facial envelope (grey), and the transformed nose (red) and mouth (blue) soft tissue forms.

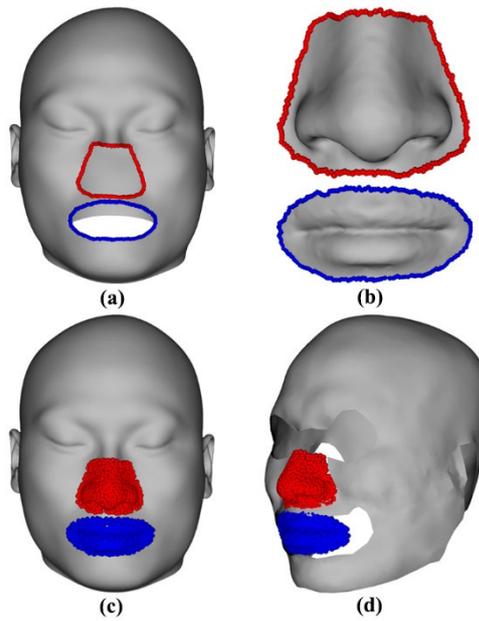


Figure 7.4 The coarsely approximated face. (a) Deformed average facial envelope with boundary curves. (b) Approximated nasal and mouth shapes with boundary curves. (c) Superimposition of deformed average facial envelope (grey) and transformed approximated nose (red) and mouth (blue) soft tissue forms. (d) The coarsely approximated face comprising approximated facial envelope (grey), and the transformed nose (red) and mouth (blue) soft tissue forms.

7.2.2.2.2 An improved approximated face

The improved approximated face was recreated by fitting the facial SSM to the alignment of approximated soft tissue shapes.

a) Facial SSM

GPA was applied to landmarks and semilandmarks of every face to generate Procrustes shape coordinates in Kendall's shape space, and then PCA was used to analyse the resulting Procrustes shape coordinates to construct the facial SSM, which represents the probability distribution of faces as a prior knowledge (Brunton et al., 2014). Thus, every face can be represented as follows:

$$F(\alpha) = \bar{F} + \sum_{i=1}^d \mu_i \alpha_i \quad (7.4)$$

where $F(\boldsymbol{\alpha})$ represents an arbitrary face within the database, $\bar{F} = [x_1, y_1, z_1, \dots, x_n, y_n, z_n]^T$ denotes the average face among samples, and n denotes the number of landmarks and semilandmarks. $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \dots, \alpha_d)$ and $\mathbf{U} = (\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \dots, \boldsymbol{\mu}_d) \in \mathbb{R}^{3 \times d}$ represent PC scores and corresponding orthogonal PCs derived from the covariance matrix, and d represents the number of PCs as determined from the cumulative proportion.

b) Facial fitting

Since the approximated facial envelope and SSM were located in different coordinate systems, we transformed the approximated facial envelope to fit the average face of SSM based on point correspondences over facial envelope. This was achieved using a method comprising three steps: first, we evenly sampled semilandmarks from the dry skull envelope. Second, we calculated intersection points from each semilandmark of the skull to facial envelope and the deformed average facial envelope (Shui et al., 2021). Finally, we recognized the index of every intersection point of the deformed average face then selected the vertices from the average face of SSM using the same indices, thereby considering the intersection points of the facial envelope and selected points of every face with the same indices serving as point correspondences. Afterwards, the coefficients were optimized to allow the coarsely approximated face to resemble the alignment of approximation as follows:

$$\arg \min_{\boldsymbol{\beta}} (\|\mathbf{U}_s \cdot \boldsymbol{\beta} - (tp_i - \bar{F}_i)\|^2 + \lambda^2 \|\boldsymbol{\beta}\|^2) \quad (7.5)$$

where $\boldsymbol{\beta} = (b_1, b_2, \dots, b_d)$ represents coefficients of the fitted approximated face in the facial SSM, tp_i and \bar{F}_i represent the i -th correspondence between the transformed facial envelope and the average face of SSM, $\mathbf{U}_s \subset \mathbf{U}$ represents the subset of PCs of SSM and λ denotes the weighting coefficient.

Then, geometric coordinates of every vertex of the fitting approximation were determined as follows:

$$Q(\boldsymbol{\beta}) = \bar{F} + \sum_{i=1}^d \boldsymbol{\mu}_i b_i \quad (7.6)$$

To minimize fitting errors and utilize the originally approximated nose and mouth soft tissues in place of those derived from the fitted SSM, the Laplacian deformation (Sorkine et al., 2004) was used to deform $Q(\boldsymbol{\beta})$ to the coarsely approximated soft tissues. The established point correspondences over facial envelope and landmarks and semilandmarks of nose and mouth soft tissues are regarded as fixed constant anchors to constrain the deformation. Hence an improved approximated face was generated as follows:

$$\mathbf{E} = \sum_{i=1}^n \|\delta_i(g_i) - \delta_i(v_i)\|^2 + \sum_{j=1}^m \|v_j - l_j\|^2 \quad (7.7)$$

where $g_i \in \mathbb{R}^3$ and $v_i \in \mathbb{R}^3$ represent the i -th vertex of improved approximation (after

deformation) and fitting approximation (before deformation), and δ_i represents Laplacian coordinates of every vertex derived from one-ring adjacent vertices, v_j and l_j represent the j -th correspondence of the fitting approximation and the coarsely approximated soft tissues.

7.2.2.3 Assessment of the proposed method

LOOCV was performed repeatedly with the skull of each sample used once as the test data and other samples as the training data to quantify the craniofacial relationships. Subsequently, the approximated face of every tested skull was generated by the proposed method. We used the resemblance comparison (i.e. geometric difference between approximation and actual faces) to test the accuracy. In this process, Procrustes superimposition (Gunz & Mitteroecker, 2013) was used to register the approximated face and actual face based on the landmark and semilandmarks configuration reported in the previous study (Smith et al., 2021). The Procrustes distances between landmarks and semilandmarks of the registered approximated and actual faces were calculated to quantify gross surface difference. Additionally, a geometric difference was calculated by computing the average value of Euclidean distances between dense point correspondences established by searching the nearest point between these two faces. Small differences would indicate that the approximated face bears a strong resemblance to the actual one, whereas large differences would reflect dissimilarity. A colour map of Euclidean distances between correspondences of the approximated and actual faces was used to recognize the regional surface difference.

Additionally, a recognition rate (i.e. a comparison of an approximated face to every face within the face pool comprising every actual face) was used to evaluate the accuracy. A correct match indicates that the approximated face is one of the k most resembled faces regarding the actual face through a comparison of the Procrustes distance, called top- k rank. The recognition rate (%) was calculated as the percentage of the number of correct matches divided by the number of samples.

7.3 Results

We analysed craniofacial relationships among modern humans and tested the accuracy of the proposed FA method. Afterwards, the proposed method was applied to UC 101.

7.3.1 Quantification of craniofacial relationships among modern humans

7.3.1.1 FSTDs at landmarks and semilandmarks of modern humans

FSTDs at landmarks and semilandmarks of modern humans were analysed and visualised using the colour map. Figure 7.5a shows average dense FSTDs for all samples. The average value of FSTDs was almost 9.29 mm with a standard deviation of almost

6.48 mm. Our results indicated that FSTDs were almost distributed symmetrically with regard to the mid-sagittal plane. The smallest FSTDs were observed around the frontal bone and cranial vault, while the largest FSTDs were found around the lateral maxillary bone and mandible (corresponding to cheek tissues), the greater wing of sphenoid bone and the base of the occipital bone. Figure 7.5b shows measured orientations (red arrows) of landmarks (yellow) located within the mid-sagittal plane and at intersection points (blue) on the average face. Line segments (red arrow) between skull landmarks and intersection points represent FSTDs.

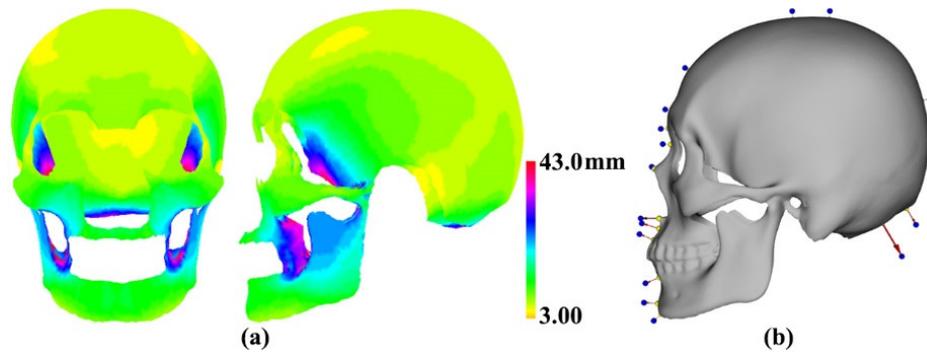


Figure 7.5 Average FSTDs at landmarks and semilandmarks. (a) A colour map of average FSTDs. (b) Measured orientations (red arrows) of landmarks (yellow) located within the mid-sagittal plane and at intersection points (blue) on the average face.

7.3.1.2 Quantification of the nasal relationship

To explore associations between nasal hard and soft tissue shapes, we performed 2B-PLS analysis of hard versus soft tissues. During this analysis, results obtained using hard tissue landmarks alone were compared to results derived from hard tissue landmarks and semilandmarks together in order to determine whether increased use of semilandmarks could enhance covariation. A summary of PLS correlation coefficients obtained using the two different configurations is presented in Table 7.2. Because only the first three PLS dimensions explained more than 10% the total shape variance, RV coefficients of these three PLS dimensions and all dimensions were calculated separately. The results showed that a weak covariation was found between the two blocks, and use of landmarks and semilandmarks together increased PLS and RV correlations (r), while an increase in number of PLS dimensions reduced the strength of the RV correlation.

Table 7.2 The 2B-PLS analysis between nasal hard and soft tissue shapes

Configurations of hard tissues	PLS axes				RV (first three PLS)		RV (all the PLS)	
	PLS	r	p value	%	r	p value	r	p value
landmarks	PLS1	0.748	<0.01	32.84%	0.389	<0.01	0.297	<0.01
	PLS2	0.599	<0.01	20.20%				
	PLS3	0.634	<0.01	15.37%				
landmarks and semilandmarks	PLS1	0.753	<0.01	34.69%	0.525	<0.01	0.383	<0.01
	PLS2	0.793	<0.01	25.41%				
	PLS3	0.645	<0.01	13.25%				

Figure 7.6a shows scatterplots of nasal hard tissue landmark and semilandmarks along the horizontal axis and soft tissue landmark and semilandmarks along the vertical axis in PLS1. The red line represents linearly fitted regression of soft tissues (dependent variables) onto hard tissues (independent variables) with confidence intervals. Figure 7.6b and e display frontal and profile views of two fitted hard and soft tissue shapes (PLS1+ and PLS1-), respectively. These fitted shapes were superimposed onto the common coordinate system, as shown in Figures 7.6c and f. Figures 7.6d and g show the main patterns of movements (red arrows) of landmarks and semilandmarks with PLS1- serving as the reference. These results indicate nose soft tissues tend to widen and shorten with increasing PLS1 score, whereby PLS1+ exhibits a relatively shorter, wider and more protruding rhinion and bridge, a wider and bigger nasal alar, a slightly protruding infratip lobule and columella, and a more protruding subnasale compared to PLS1-. Additionally, as the PLS score becomes increasingly positive, nasal hard tissues tended to widen and shorten and protrude, while the region around the nasal spine becomes smaller and contracted. These results indicate that morphological shape changes in hard tissues along two axes, including the width along the right-left axis and height along the superior-inferior axis, are almost identical to soft tissue shape changes.

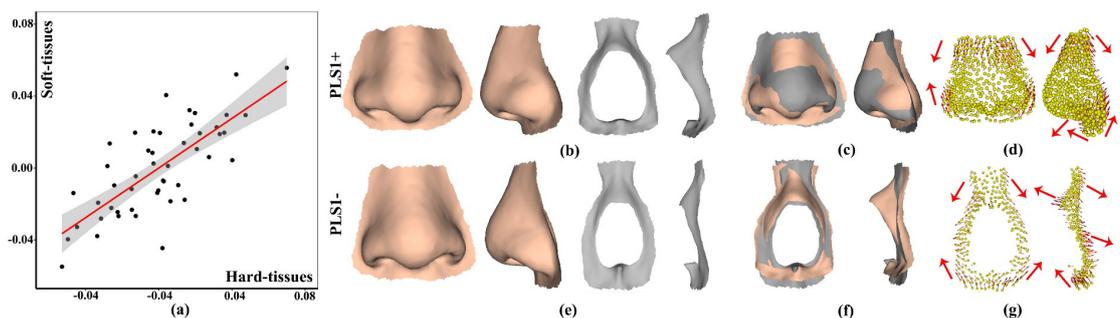


Figure 7.6 Covariation of PLS1 between nasal hard and soft tissue shapes. (a) Scatterplots of PLS1 axes of landmark and semilandmark configurations of nasal hard and soft tissues. The red line represents the regression of soft versus hard tissue shapes with confidence intervals. (b) Fitted shapes along the positive limit. (c) Superimposition of PLS1+ (peach) and PLS1- (gray) of nasal soft tissues. (d) The movements of landmarks and semilandmarks. (e) Fitted shapes along the negative limit. (f) Superimposition of PLS1+ (peach) and PLS1- (gray) of nasal hard tissues. (g) The movements of landmarks and semilandmarks.

We used the LOOCV strategy to test the accuracy of approximated soft tissue shapes then calculated statistical descriptive of Procrustes distance, e.g. mean, standard deviation, maximum Procrustes distance ratio (*maxRatio*) and minimum Procrustes distance ratio (*minRatio*), as listed in Table 7.3. Figure 7.7a shows the Procrustes distance between approximated and actual soft tissue shapes with green curves using landmark and semilandmark configurations as obtained by our method, while the red curve represents the result obtained using the partial least square regression (PLSR) method (Shrimpton et al., 2014). Because our method was associated with a smaller geometric deviation, as reflected by the lowest descriptive measures, the proposed method may provide an improved approximation. However, a great variation in the error ratio shows there is a greater error in the approximated nasal soft tissue shapes. Additionally, we examined the

effect of hard tissue landmarks alone on the approximation with cyan curves depicting results obtained using our method. The results suggest that the proposed method based on landmark and semilandmark configuration perform better. Figure 7.7b shows that the approximated nose (second left) bears the greatest resemblance to the actual one (left), as reflected by the minimum Procrustes distance. The rightmost two figures display a colour map of geometric differences between these two shapes, whereby yellow represents a smaller difference and red and blue represent greater deviation. Figure 7.7c shows the approximation corresponding to the maximum Procrustes distance. These results indicate that greatest approximated differences are observed around the nasal alar and tip, rhinion, bridge and subnasale.

Table 7.3 A comparison of the approximated and actual nasal soft tissues based on Procrustes distance

Methods	Different configurations of hard tissues	Shape			
		Mean	SD	<i>MaxdR</i>	<i>MindR</i>
Our method	landmarks and semilandmarks	0.0761	0.0131	1.4627	0.7079
PLSR	landmarks and semilandmarks	0.0842	0.0148	1.6788	0.7485
Our method	landmarks	0.0874	0.0179	2.0323	0.7390

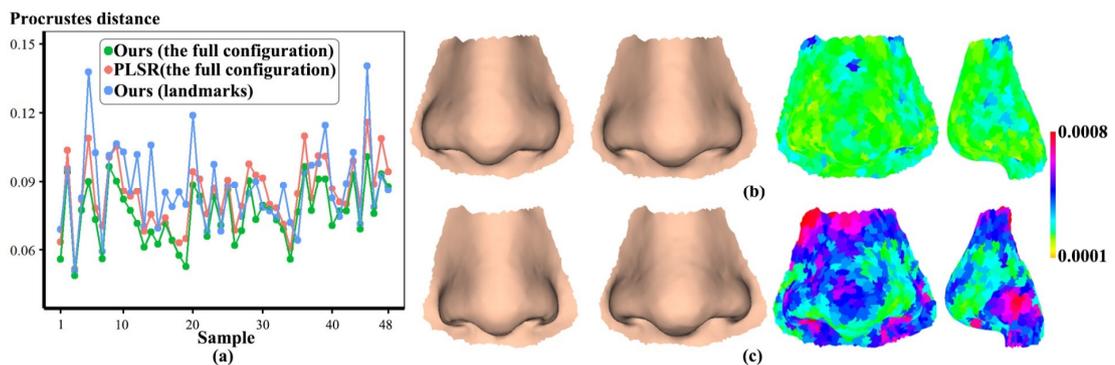


Figure 7.7 Geometric discrepancy of approximated and actual nasal soft tissue shapes. (a) Geometric discrepancies of every sample as derived using different approaches and configurations. (b) Comparisons of approximated (second left) and actual (left) shapes corresponding to the minimum Procrustes distance. (c) Comparison of approximated (second left) and actual (left) shapes corresponding to the maximum Procrustes distance.

7.3.1.3 Quantification of the mouth relationship

2B-PLS analysis was used to explore covariations of oral hard and soft tissue shapes. A summary of PLS correlation coefficients and RV coefficients obtained using two different hard tissue configurations is presented in Table 7.4. Due to the fact that only the first two PLS dimensions explained more than 10% of the total shape variance, RV coefficients of these PLS dimensions and all dimensions were calculated, with this analysis revealing that weak correlations exist between the two blocks. Each PLS and RV correlation (r) derived from hard tissue landmark and semilandmark configurations was greater than those derived from either landmark alone, respectively. However, an increase in the number of dimensions led to reduced RV correlations.

Figure 7.8a shows scatterplots of hard tissue shapes along the horizontal axis and soft

tissue shapes along the vertical axis of PLS1. The red line represents linearly fitted regression of soft tissues (dependent variables) onto hard tissues (independent variables) with confidence intervals. Figures 7.8b and e display frontal and profile views of fitted hard and soft tissue shapes (PLS1+ and PLS1-). Figures 7.8c and f show superimposed fitted. Figures 7.8d and g illustrate movement (red arrows) of every landmark and semilandmark (yellow points) of PLS1- (as the reference) to describe shape changes. These results show that mouth soft tissue shapes tend to narrow and enlarge with an increasing PLS score, with an increasing PLS1+ associated with an increasingly narrow, protruding upper and lower lip, an increasingly protruding and enlarged open mouth and an oral fissure located in a more backward position. Meanwhile, an increasingly positive PLS score is associated with increasingly narrow and enlarged mouth hard tissue shapes and increasingly protruding anterior teeth. Taken together, these results indicate that main patterns of hard tissue shape changes were almost always consistent with shape changes observed for soft tissue shape.

Table 7.4 2B-PLS analysis between mouth hard and soft tissue shapes

Configurations of hard tissues	PLS axes				RV (The first two PLS)		RV (all the PLS)	
	PLS	<i>r</i>	<i>p value</i>	%	<i>r</i>	<i>p value</i>	<i>r</i>	<i>p value</i>
Landmark	PLS1	0.633	<0.01	63.83%	0.374	<0.01	0.271	<0.01
	PLS2	0.704	<0.01	12.91%				
landmarks and semilandmarks	PLS1	0.661	<0.01	56.55%	0.412	<0.01	0.325	<0.01
	PLS2	0.768	<0.01	21.16%				

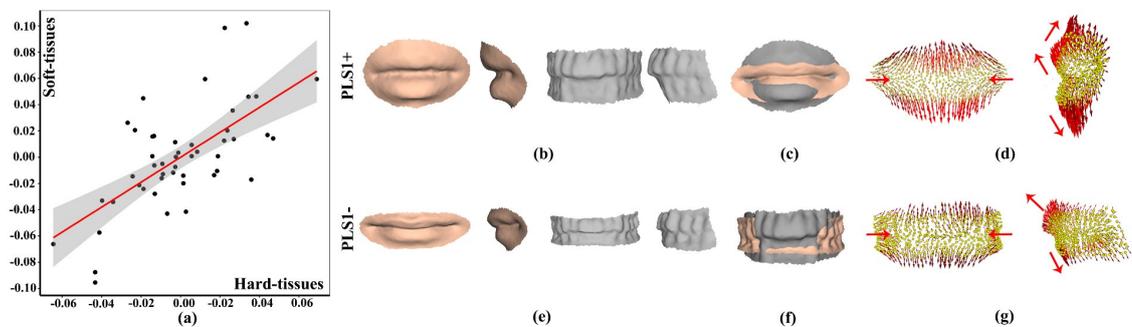


Figure 7.8 Covariation of PLS1 between mouth hard and soft tissue shapes. (a) Scatterplots of PLS1 axes of landmark and semilandmark configurations of hard and soft tissue shapes. The red line represents the regression of soft tissue shapes on hard tissues with confidence intervals. (b) Fitted shapes along the positive limit. (c) Superimposition of PLS1+ (peach) and PLS1- (gray) shapes. (d) Movement (red row) of every landmark and semilandmark of PLS1- (as the reference). (e) Fitted shapes along the negative limit. (f) Superimposition of fitted shapes along positive (peach) and negative (gray) limits. (g) Movement (red row) of every landmark and semilandmark of PLS1- (as the reference).

Testing of approximated mouth soft tissues using LOOCV revealed the Procrustes distance of approximated and actual soft tissue shapes (Figure 7.9a). Procrustes distances between the approximated and actual shapes are reported in Table 7.5. The smaller average values and standard deviations of geometric deviations indicate that our method may improve approximating results. However, a great variation in the error ratio shows there is a greater error in the approximated mouth shapes. 62.5% ($n = 30$) of samples

(obtained using landmarks and semilandmarks together) yielded smaller Procrustes distances as compared to results obtained using landmarks alone. The result suggested that the use of landmark and semilandmarks improve approximation accuracy. Figure 7.9b shows the approximated result (middle) that bears the greatest resemblance to the actual result (left), as reflected by its minimum Procrustes distance. The right column displays the colour map of geometric differences. Figure 7.9c shows regional approximation differences corresponding to the maximum Procrustes distance. The results indicated that the greatest changes were observed around boundary regions, as well as of philtrum, mouth corners, cupid’s bow and lower lip.

Table 7.5 Comparison of approximated and actual mouth soft tissue shapes based on Procrustes distance

Methods	Different configurations of hard tissues	Shape			
		Mean	SD	<i>MaxRatio</i>	<i>MinRatio</i>
Our method	landmarks and semilandmarks	0.0684	0.0221	2.0157	0.5731
PLSR	landmarks and semilandmarks	0.0734	0.0226	2.0169	0.5521
Our method	landmarks	0.0714	0.0229	2.1619	0.5715

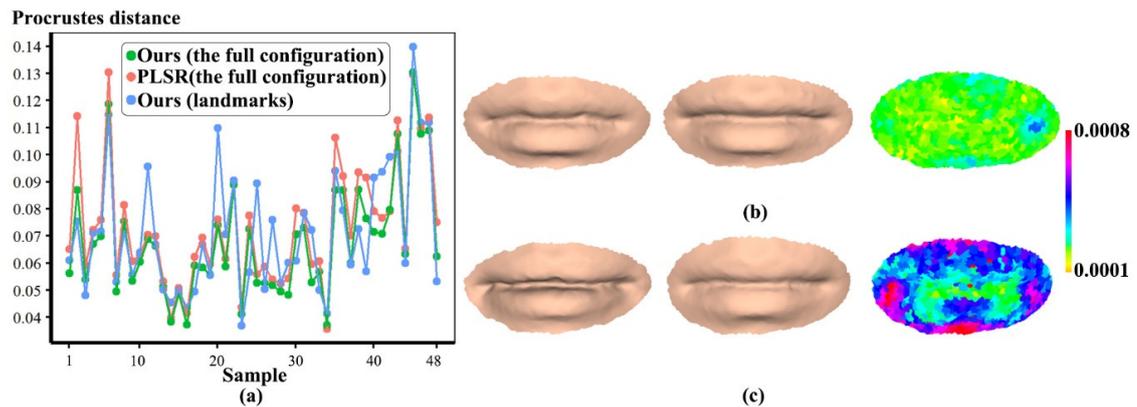


Figure 7.9 Geometric discrepancy of the approximated and mouth soft tissue shapes. (a) Geometric discrepancy of every sample as derived using different approaches and configurations. (b) Comparison of approximated (middle) and actual (left) shapes corresponding to the minimum Procrustes distance. (c) Comparison of approximated (middle) and actual (left) shapes corresponding to the maximum Procrustes distance.

7.3.2 Evaluation of the proposed method

We computed the Procrustes distance and geometric difference between the approximated and actual faces. Figure 7.10a shows the Procrustes distance of every sample, while the average Procrustes distance was 0.0258. The geometric difference between the approximated and actual faces was 1.79 mm, while 72.92% ($n = 35$) of samples showed a geometric difference less than 2.0 mm. Figure 7.10b shows the approximated face corresponding to the minimum Procrustes distance of 0.0164, in which the geometric difference was 1.12 mm. The approximated face closely resembled the actual face. The result showed that the approximated face closely resembled the actual face. Figure 7.10c shows the approximated face corresponding to the maximum Procrustes distance of 0.0433, while geometric difference was 3.28 mm. The greatest

shape differences were observed around cheeks, lateral temporal-cheek, nasolabial, medial inferior orbital, chin, lateral forehead, ears, and the back of head and neck.

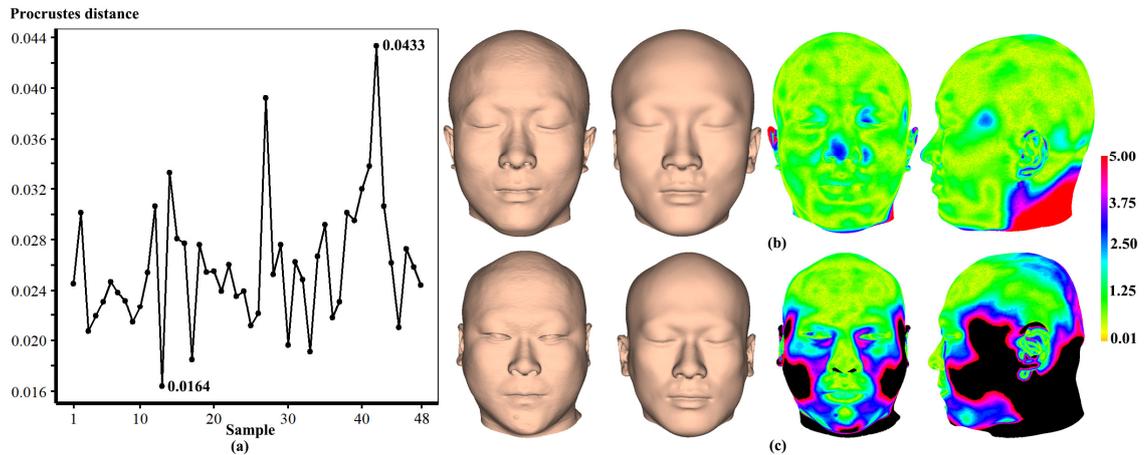


Figure 7.10 A comparison of the approximated and actual faces. (a) Procrustes distance between the approximated and actual face of every sample. (b) Comparison of approximated (second left) and actual (left) faces corresponding to the smallest deviation. (c) Comparison of approximated (second left) and actual (left) faces corresponding to the greatest deviation.

To present the extent of approximation bias in different regions, the mean and standard deviation of Euclidean distance of every vertex among all samples are illustrated in Figures 7.11a and b, respectively. The results showed that the greatest differences were found around the cheeks, lateral temporal-cheek, ears, and the back of head and neck. Figure 7.12 shows eight examples generated by the proposed method.

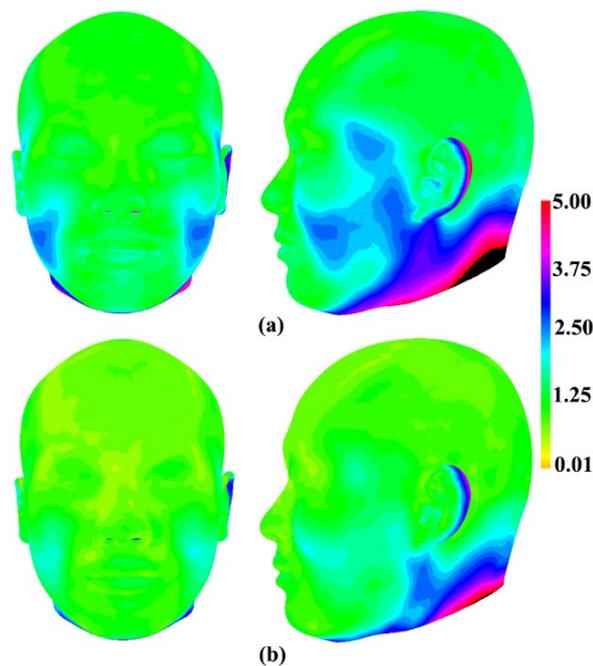


Figure 7.11 Colour maps of the approximation difference between the approximated and actual faces. (a) Mean. (b) Standard deviation.

Additionally, we calculated the Procrustes distance between the approximated face and every face within the face pool. The top-1 recognition rate of 91.67% ($n = 44$) and the top-3 recognition rate of 95.83% ($n = 46$) demonstrated the effectiveness of the

proposed method. Each example includes the tested skull and the actual faces (the leftmost two columns), the approximated faced generated by the proposed (middle column), and a colour map of geometric deviations (rightmost column) with the Procrustes distance.

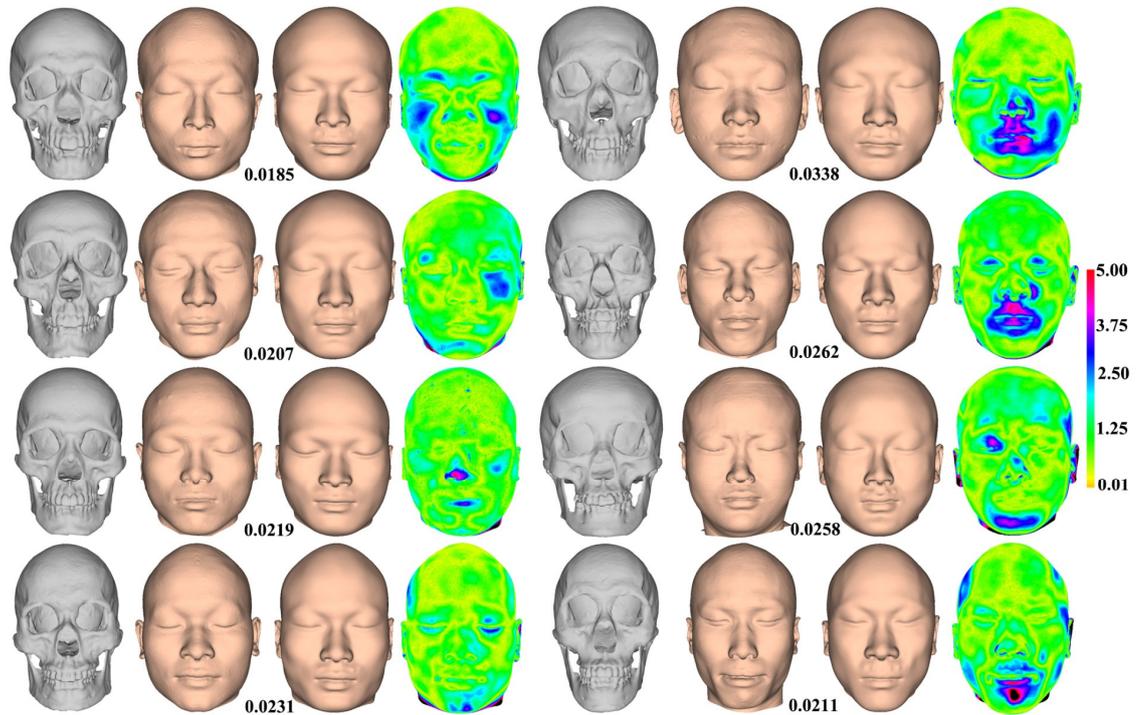


Figure 7.12 The approximated and actual faces of eight test samples. Each example includes the test skull (gray), actual (second column), and approximated (third column) faces, and a color map of geometric deviation (rightmost).

7.3.3 FA of the UC 101 skull

We approximated the facial appearance of UC 101 using the proposed method. Figure 7.13a shows the external surface of the approximated facial envelope based on average FSTDs. Figure 7.13b shows approximated nose and mouth soft tissue shapes (left) and the coarsely approximated face (right). However, it is still lacking eyes and ears, and including missing geometry. Figure 7.13c shows the improved approximated face and Figure 7.13d displays the profile views of half of the approximated face and UC 101, respectively. The result indicates that the main patterns and profile of the approximated face is almost consistent with those of UC 101. As compared to modern human faces, the approximated face exhibited an elongated shape, a sloped forehead, stronger and wider eyebrows and a wider nose bridge.

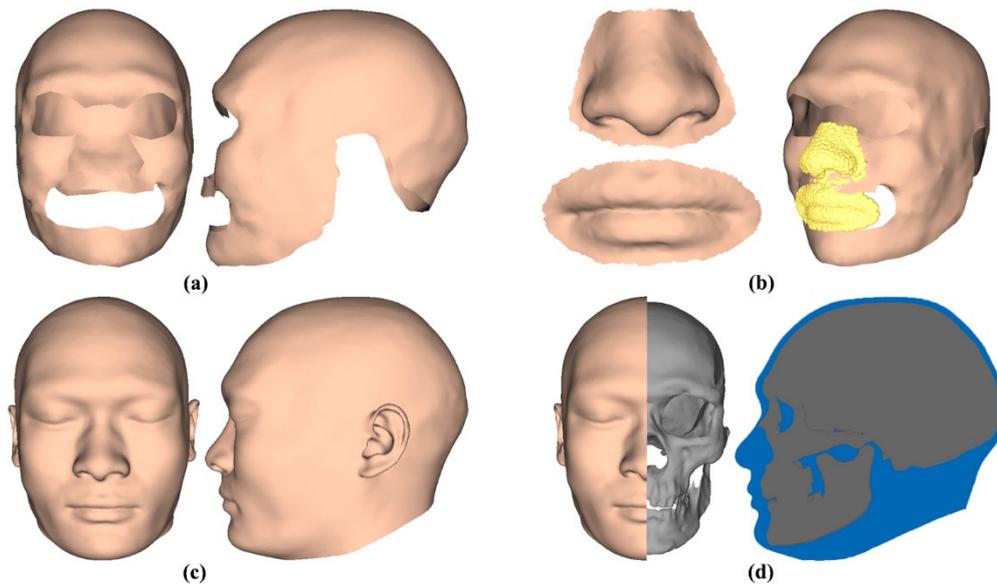


Figure 7.13 Facial approximation of UC 101. (a) External surface of the approximated facial envelope. (b) Alignment of approximated soft tissues of each component. (c) Improved approximated face. (d) Frontal and profile views of UC 101 and half of the approximated face.

7.4 Discussion

Despite recent progress in promoting the accuracy of the approximated face using computerized methods, existing methods suffer from several issues, including less confident understanding of detailed craniofacial relationships between hard and soft tissues and challenges related to the assignment of the learnt relationships to obtain a reliable and accurate facial appearance. Unlike the most computerized FA method for representing craniofacial relationship in a holistic way (Berar et al., 2011; De Buhan & Nardoni, 2018; Jia et al., 2021; Turner et al., 2005), this study has presented a computerized method for quantifying the craniofacial relationships based on average FSTDs and covariation of landmarks and semilandmarks between nasal and oral hard and soft tissues. Furthermore, we applied the learnt relationships to approximate the facial envelope, nose and mouth soft tissues and afterwards employed facial SSM to enhance the accuracy of approximating results. The smaller resemblance comparison and greater recognition rate demonstrate the effectiveness of the proposed method.

7.4.1 The quantification of craniofacial relationships

Anatomical modelling of craniofacial relationships plays a significant role in predicting facial soft tissues from the dry skull. Over the past decades, average FSTDs at landmarks reported in tabular forms provide a promising method to explore craniofacial relationships and to generate facial envelope of the dry skull (Hayes et al., 2013; Wilkinson, 2010). It facilitates prediction of the shape and size of muscle attachments. Thus many studies have collected and updated FSTDs datasets with regard to ethnicity, sex, age and nutritional status (Stephan, 2017). Because landmarks cannot capture the detailed information of skull morphology, FSTDs at landmarks cannot provide sufficient

information to represent detailed craniofacial relationships. Hence, the practitioner may determine the muscle structures based on experience and imagination. To reduce misinterpretation, FSTDs at landmarks and high-density semilandmarks provide a promising means for intuitively explaining the overall craniofacial relationships. Notably, the choice of FSTDs affects the ultimate FA results. A previous study highlighted average FSTDs at landmarks increased the accuracy of FA using virtual sculpture technique (Lee et al., 2015). In this study, a smaller resemblance comparison between approximated and actual faces verifies a positive effect of average dense FSTDs on the quantification of overall craniofacial relationships and the improvement of FA.

Facial features are capable of enhancing the realism of the approximated face. The conventional anthropomorphic measurements of skulls and faces have been applied to examine the relationships (e.g. the corner positions of mouth, nasal tip, inner and outer canthi) (Wilkinson, 2010). Nevertheless, the lack of detailed morphology means that the facial feature surfaces cannot be directly produced and ambiguous findings in interpreting the relationships may negatively affect the accuracy of FA. Recently, a landmark-based geometric morphometric method was developed to examine the relationships between bony facial features and soft tissue morphologies (Guyomarc'h et al., 2014; Kustár et al., 2013; Ridel et al., 2020), thereby generating surfaces by deforming the template facial features. A previous study indicated that there was a relatively higher correlation in nose and moderate correlation in mouth (RV coefficient <0.25) (Guyomarc'h et al., 2014). Nevertheless, approximation errors of nose (3.1 mm) and mouth (4.5 mm) were greater. A similar result was also found in the approximated nose (~2.7 mm) of South Africans (Ridel et al., 2020).

Based on the analysis of PLS 1 of nose and mouth, we find the shapes of nose (or mouth) hard and soft tissues seem to be changing with each other and bony structure appears to have an effect on the soft tissues in a nearly linear manner (Figures 7.6 and 7.8), e.g. a narrow nasal bone and enlarged and protruding anterior teeth are correlated with a mouth with narrow and protruding upper and lower lips. Hence, the multiple regression can be used to approximate the nasal and mouth soft tissues shapes. However, the approximation errors are inevitable (Figures 7.7 and 7.9) and hence the approximating results need to be used with caution. The potential reason is that soft tissues of facial features cannot be totally determined by bone morphology, e.g. nasal protrusion seems less likely to be associated with the nasal bone (Kustár et al., 2013). Additionally, since a limited number of landmarks may not capture the detailed biological structures, we used high-density geometric morphometrics to examine the nose and mouth relationships between hard and soft tissues, respectively. The slightly higher correlation (Tables 7.2 and 7.4) suggests high-density semilandmarks enhance the quantification of craniofacial relationships, thereby predicting better nose and mouth shapes with smaller

approximation errors (Figures 7.7 and 7.9).

7.4.2 The computerized FA method

Based on the learnt craniofacial relationships, we developed a coarse-to-fine strategy to approximate the entire facial appearance of the dry skull. Like our previous work (Shui et al., 2021), average dense FSTDs with regard to age and sex were assigned to the external surface of the skull to generate the facial envelope. The approximated facial envelope almost bears a resemblance to the actual one, especially in the regions exhibiting small FSTDs, e.g. forehead and scalp. This guarantees morphological consistency between the ultimate approximated face and dry skull and prevents overlap of the approximated within bony structure. However, geometric differences between the approximated and actual faces indicate that the greatest variations were observed around cheeks, chin and temporal regions, which are consistent with the results of a previous study that demonstrated the FSTDs variations were greatest around these regions (Dong et al., 2012). Hence, different choices of FSTDs can be used to generate multiple facial appearances of the same skull to assist facial recognitions.

As noted in previous studies (Guyomarc'h et al., 2014; Kustár et al., 2013; Ridel et al., 2020), the regression method was used to predict eyes, nose and mouth morphologies. Here we found that the proposed regression method seems to provide a slight smaller approximation error than PLSR. The choice of landmarks and semilandmarks also affects the approximating results. High-density semilandmarks are capable of improving the classification of specimens (Schlager & Rüdell, 2017). In this study, high-density semilandmarks seem to better approximate nose and mouth shapes than the use of landmarks alone, although the differences between them are relatively small. However, the degree of equivalences of semilandmarks among skulls and faces needs to be cautious. Additionally, the large errors in approximated nose and mouth morphologies are observed, and hence indicate that nose and mouth soft tissues needs to be very carefully produced. The approximation might require consideration of the extent of artistic interpretation to show detailed facial feature morphologies (Wilkinson, 2010).

A quantitative assessment enables us to validate the accuracy of the proposed FA method and to continue development and improvement, focusing on regions with the greatest discrepancies (Miranda et al., 2018). In recent years, the most frequently used method for quantitative assessment is to compare the geometric difference between the approximated and actual faces (Stephan & Henneberg, 2006). Several protocols, e.g. shell-to-shell deviation, surface-to-surface deviation, between the approximated and actual faces have been employed in previous studies (Decker et al., 2013; Lee et al., 2012; Miranda et al., 2018; Short et al., 2014). To superimpose these two faces together, a straightforward method is to perform registration based on the landmark configurations (at least three landmarks), thereby visually observing the profile variations and computing

the geometric difference. For example, the nasion, and the deepest lateral points of the orbits have been used (Decker et al., 2013; Miranda et al., 2018). Because the registration results greatly affect the evaluation of the level of approximation errors and resemblance, they need to be carefully examined before performing assessment. In this study, the landmarks and semilandmarks covering the whole head are used to register the approximated and actual faces, and then the gross and regional difference are quantified.

Additionally, the recognition rate is a preferred way for evaluating the accuracy of FA in forensic science (Stephan & Henneberg, 2006). Assessors are asked to select the most resembled face within the face pool and then the recognition rates with regard to different top ranks are shown (Parks & Monson, 2018). A higher recognition rate indicates the effectiveness of the method. However, a few studies have criticized this process as being too subjective and dependent upon assessors' intuition and experience (Li et al., 2022). A more objective evaluation method needs to be explored. For example, Parks and Monson used an automated facial recognition method to evaluate the accuracy of eye placement in facial approximations (Parks & Monson, 2016). In this study, a quantitative method followed by the resemblance comparison of digital models was developed for determining the correct match and afterwards calculating the recognition rate, avoiding human intervention. We suggest the same test cases, dataset and evaluation method (e.g. resemblance comparison and recognition rate) should be used for validating the level of effectiveness of different FA methods, avoiding the dependencies on the dataset.

A collection of skull and face models of the same person using medical imaging technique is the fundamental basis for exploring craniofacial relationships. It is worth mentioning that an increasing number of these models is capable of enhancing the accuracy of the approximated faces. However, a large acquisition of the whole head to learn craniofacial relationships remains challenging taking into account imaging time, cost, and health risks. In this study, the proposed method provides a promising means for generating the facial appearance from incomplete head CT scan, i.e. the approximated soft tissue shape of each component can be obtained from different datasets separately and then the entirely approximated face can be recreated by reassembling all these approximating results together. It reduces the number and quality of skull and face datasets to achieve FA.

7.4.3 Facial approximation of UC 101

As described in previous studies (Hayes et al., 2013; Wilkinson, 2010), the first step of FA involves an investigation of the UC101 skull to estimate personal information, such as age, sex and ethnicity, and examine morphological traits. Such information can facilitate the selection of appropriate human populations for obtaining craniofacial relationships. Additionally, distorted geometric surfaces and missing regions need to be carefully repaired based on the verified evidence, e.g. intact regions on the mirrored side

(Gunz et al., 2009; Guyomarc'h et al., 2018; Shui & Gao, 2021). In our previous study (Shui et al., 2020a), we restored the missing regions by deforming the reflection of intact regions to damaged regions based on manually placed landmarks.

Modern human faces have consistent patterns of craniofacial relationships and muscle structures (shapes, sizes and locations) associated with the bony structures. Previous studies used the computer-based FA method (Coutinho Nogueira et al., 2019) and virtual sculpture technique (Hamre et al., 2017) to approximate the facial appearance of named people in the past based on assigning the craniofacial relationships of modern humans to the dry skull. The idea behind these case studies is based on an intuitive method for estimating the past in light of present observations. It is acknowledged that approximation biases are inevitable (Benazzi et al., 2009; Wilkinson, 2010). However, the approximated faces provide new insights in understanding the characteristic features of human fossils and enable an exploration of the evolutionary forces driving anatomical changes in ancestral humans. In most cases, it is recommended to generate an average facial appearance of historical figure in absence of other reliable evidence (Marić et al., 2020). Hence, the verified craniofacial relationships among human populations is regarded as those of UC 101 to approximate facial appearance.

Since UC 101 exhibits a considerably longer and lower cranium compared with modern human skulls, there might be a great approximation error in prediction of facial envelope using regression method. In a worse situation, the approximated face is likely to penetrate the skull, e.g. nasal bridge and back of the head. Thus, the facial envelope is approximated based on average FSTDs instead of the multiple regression method (Berar et al., 2011; Jia et al., 2021; Madsen et al., 2018; Shui et al., 2020b). Additionally, the nasal and mouth shapes derived from the deformation-based method (Deng et al., 2011; Turner et al., 2005) are not regarded as those of UC 101, because individual nasal and mouth bony variation affects the facial morphological traits. Hence, the multiple regression was used to predict the nose and mouth soft tissues of UC 101, making use of skull morphology. However, the accuracy of approximated nasal and mouth shapes should be considered with caution. In the absence of the actual face of UC 101, it was not possible to quantify the extent of resemblance and the approximation error of modern humans is regard as that of UC 101. We suggest that scientifically tested FA methods should be used to generate the facial appearance of *Homo sapiens* and the approximation error, and advantages and limitations of the method should be elucidated.

7.5 Conclusion

In this study, we developed a computerized FA method for exploring the craniofacial relationships between hard and soft tissues and developing a coarse-to-fine strategy for generating facial appearances. Average FSTDs at landmarks and semilandmarks are used to quantify the overall craniofacial relationship and contribute to enhancing the accuracy

of the approximated face. Additionally, nasal and oral hard tissues have an effect on their soft tissue shapes separately, and hence the multiple regression can be used to approximate the nasal and mouth shapes. However, relatively weak covariations and greater approximation errors suggested we need to be cautious about the approximation of nose and mouth soft tissue shapes. This method should be useful for a broad range of applications in forensic science, anthropology and archaeology. Future studies are need to explore the novel regression method to improve the accuracy of the approximated nose and mouth soft tissue shapes.

7.6 Acknowledgements

I would like to thank Beijing Key Laboratory of Digital Preservation and Virtual Reality for Cultural Heritage, Beijing Normal University for granting access to their human skull and face database. Additionally, I would like to thank the Institute of Vertebrate Paleontology and Paleoanthropology for providing UC 101 and thank Dr. Yameng Zhang for placing 91 landmarks of UC 101.

7.7 References

- Adams, DC, & Otárola-Castillo, E. (2013). geomorph: an R package for the collection and analysis of geometric morphometric shape data. *Methods in Ecology and Evolution*, 4(4), 393-399.
- Amberg, B, Romdhani, S, & Vetter, T. (2007). Optimal step nonrigid ICP algorithms for surface registration. The *IEEE Conference on Computer Vision and Pattern Recognition (CVPR'07)*, 1-8.
- Baldasso, RP, Moraes, C, Gallardo, E, Stumvoll, MB, Crespo, KC, Strapasson, RAP, & de Oliveira, RN. (2021). 3D forensic facial approximation: Implementation protocol in a forensic activity. *Journal of Forensic Sciences*, 66(1), 383-388.
- Benazzi, S, Fantini, M, Crescenzo, FD, Mallegni, G, Mallegni, F, Persiani, F, & Gruppioni, G. (2009). The face of the poet Dante Alighieri reconstructed by virtual modelling and forensic anthropology techniques. *Journal of Archaeological Science*, 36(2), 278-283.
- Berar, M, Tilotta, FM, Glaunès, JA, & Rozenholc, Y. (2011). Craniofacial reconstruction as a prediction problem using a Latent Root Regression model. *Forensic Science International*, 210, 228-236.
- Bernardini, F, Mittleman, J, Rushmeier, H, & Taubin, G. (1999). The Ball-Pivoting Algorithm for Surface Reconstruction. *IEEE Transactions on Visualization & Computer Graphics*, 5(4), 349-359.
- Bookstein, FL. (1989). Principal warps: Thin-plate splines and the decomposition of deformations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(6), 567-585.
- Brunton, A, Salazar, A, Bolkart, T, & Wuhrer, S. (2014). Review of Statistical Shape Spaces for 3D Data with Comparative Analysis for Human Faces. *Computer Vision and Image Understanding*, 128(11), 1-17.
- Campbell, RM, Vinas, G, Henneberg, M, Diogo, R, & Evolution. (2021). Visual depictions of our evolutionary past: A broad case study concerning the need for quantitative methods of soft tissue reconstruction and art-science collaborations. *Frontiers in Ecology and Evolution*, 9, 639048.
- Claes, P, Vandermeulen, D, De Greef, S, Willems, G, Clement, JG, & Suetens, P. (2010). Computerized craniofacial reconstruction: Conceptual framework and review. *Forensic Science International*, 201(1-3), 138-145.
- Corsini, M, Cignoni, P, & Scopigno, R. (2012). Efficient and flexible sampling with blue noise properties of triangular meshes. *IEEE Transactions on Visualization and Computer Graphics*, 18(6), 914-924.

- Coutinho Nogueira, D, Dutailly, B, Comte, F, Vasil'iev, A, Khokhlov, A, Shvedchikova, T, Berezina, N, et al. (2019). "Gueule cassée"(facial injuries): a 3D paleotraumatology study and facial approximation of a Napoleonic soldier who died in 1812 at Königsberg during the Russian Campaign. *International Journal of Osteoarchaeology*, 29(2), 191-197.
- Cunningham, DL, & Jantz, RL. (2003). The morphometric relationship of Upper Cave 101 and 103 to modern Homo sapiens. *Journal of Human Evolution*, 45(1), 1-18.
- De Buhan, M, & Nardoni, C. (2018). A facial reconstruction method based on new mesh deformation techniques. *Forensic Sciences Research*, 3(3), 256-273.
- 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 Science International*, 229(1), 161. e161-e168.
- Deng, Q, Zhou, M, Shui, W, Wu, Z, Ji, Y, & Bai, R. (2011). A novel skull registration based on global and local deformations for craniofacial reconstruction. *Forensic Science International*, 208(1), 95-102.
- Deng, Q, Zhou, M, Wu, Z, Shui, W, Ji, Y, Wang, X, Liu, CY, et al. (2016). A regional method for craniofacial reconstruction based on coordinate adjustments and a new fusion strategy. *Forensic Science International*, 259, 19-31.
- 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(1-3), 396. e391-e397.
- Gietzen, T, Brylka, R, Achenbach, J, zum Hebel, K, Schömer, E, Botsch, M, Schwanecke, U, et al. (2019). A method for automatic forensic facial reconstruction based on dense statistics of soft tissue thickness. *Plos one*, 14(1), e0210257.
- Gunz, P, & Mitteroecker, P. (2013). Semilandmarks: a method for quantifying curves and surfaces. *Hystrix, the Italian Journal of Mammalogy*, 24(1), 103-109.
- Gunz, P, Mitteroecker, P, Neubauer, S, Weber, GW, & Bookstein, FL. (2009). Principles for the virtual reconstruction of hominin crania. *Journal of Human Evolution*, 57(1), 48-62.
- Guyomarc'h, P, Dutailly, B, Charton, J, Santos, F, Desbarats, P, & Coqueugniot, H. (2014). Anthropological Facial Approximation in Three Dimensions (AFA3D): Computer-Assisted Estimation of the Facial Morphology Using Geometric Morphometrics. *Journal of Forensic Sciences*, 59(6), 1502-1516.
- Guyomarc'h, P, Velemínský, P, Brůžek, J, Lynnerup, N, Horák, M, Kučera, J, Rasmussen, KL, et al. (2018). Facial approximation of Tycho Brahe's partial skull based on estimated data with TIVMI-AFA3D. *Forensic science international*, 292, 131-137.

- Hamre, SS, Erslund, GA, Daux, V, Parson, W, & Wilkinson, C. (2017). Three individuals, three stories, three burials from medieval Trondheim, Norway. *Plos one*, 12(7), e0180277.
- Hayes, S. (2016). Faces in the museum: revising the methods of facial reconstructions. *Museum Management and Curatorship*, 31(3), 218-245.
- Hayes, S, Sutikna, T, & Morwood, M. (2013). Faces of Homo floresiensis (LB1). *Journal of Archaeological Science*, 40(12), 4400-4410.
- Hu, Y, Wang, Z, Pan, Y, Xie, L, & Wang, Z. (2021). Craniofacial Reconstruction via Face Elevation Map Estimation Based on the Deep Convolution Neural Network. *Security Communication Networks*, 2021, 9987792.
- Jia, B, Zhao, J, Xin, S, Duan, F, Pan, Z, Wu, Z, Li, J, et al. (2021). Craniofacial reconstruction based on heat flow geodesic grid regression (HF-GGR) model. *Computers & Graphics*, 97, 258-267.
- Kustár, A, Forro, L, Kalina, I, Fazekas, F, Honti, S, Makra, S, & Friess, M. (2013). FACE-R—a 3D database of 400 living individuals' full head CT- and face scans and preliminary GMM analysis for craniofacial reconstruction. *Journal of Forensic Sciences*, 58(6), 1420-1428.
- Lee, WJ, Wilkinson, CM, & Hwang, HS. (2012). An accuracy assessment of forensic computerized facial reconstruction employing cone-beam computed tomography from live subjects. *Journal of Forensic Sciences*, 57(2), 318-327.
- Lee, WJ, Wilkinson, CM, Hwang, HS, & Lee, SM. (2015). Correlation Between Average Tissue Depth Data and Quantitative Accuracy of Forensic Craniofacial Reconstructions Measured by Geometric Surface Comparison Method. *Journal of Forensic Sciences*, 60(3), 572-580.
- Lee, WJ, Yoon, AY, Mi, KS, Wilkinson, CM, & Dong, HS. (2014). The archaeological contribution of forensic craniofacial reconstruction to a portrait drawing of a Korean historical figure. *Journal of Archaeological Science*, 49, 228-236.
- Li, Y, Wang, J, Liang, W, Xue, H, He, Z, Lv, J, & Zhang, L. (2022). CR-GAN: Automatic craniofacial reconstruction for personal identification. *Pattern Recognition*, 124, 108400.
- Madsen, D, Lüthi, M, Schneider, A, & Vetter, T. (2018). Probabilistic joint face-skull modelling for facial reconstruction. *The IEEE conference on computer vision and pattern recognition*, 5295-5303.
- Marić, J, Bašić, Ž, Jerković, I, Mihanović, F, Anđelinović, Š, & Kružić, I. (2020). Facial reconstruction of mummified remains of Christian Saint-Nicolosa Bursa. *Journal of Cultural Heritage*, 42, 249-254.
- Miranda, GE, Wilkinson, C, Roughley, M, Beaini, TL, & Melani, R. (2018). Assessment of accuracy and recognition of three-dimensional computerized forensic craniofacial

reconstruction. *Plos One*, 13(5), e0196770.

Nelson, L, & Michael, S. (1998). The application of volume deformation to three-dimensional facial reconstruction: a comparison with previous techniques. *Forensic Science International*, 94(3), 167-181.

Parks, CL, & Monson, KL. (2016). Automated facial recognition and candidate list rank change of computer generated facial approximations generated with multiple eye orb positions. *Forensic science international*, 266, 396-398.

Parks, CL, & Monson, KL. (2018). Recognizability of computer-generated facial approximations in an automated facial recognition context for potential use in unidentified persons data repositories: Optimally and operationally modeled conditions. *Forensic science international*, 291, 272-278.

Quatrehomme, G, Cotin, S, Subsol, G, Delingette, H, Garidel, Y, Grévin, G, Fidrich, M, et al. (1997). A fully three-dimensional method for facial reconstruction based on deformable models. *Journal of Forensic Science*, 42(4), 649-652.

Ridel, A, Demeter, F, L'abbé, E, & Oettlé, A. (2020). Nose approximation among South African groups from cone-beam computed tomography (CBCT) using a new computer-assisted method based on automatic landmarking. *Forensic Science International*, 313, 110357.

Robert, P, & Escoufier, Y. (1976). A unifying tool for linear multivariate statistical methods: the RV-coefficient. *Journal of the Royal Statistical Society: Series C*, 25(3), 257-265.

Rohlf, FJ, & Corti, M. (2000). Use of two-block partial least-squares to study covariation in shape. *Systematic Biology*, 49(4), 740-753.

Schlager, S, & Rüdell, A. (2017). Sexual Dimorphism and population affinity in the human zygomatic structure—comparing surface to outline data. *The Anatomical Record*, 300(1), 226-237.

Short, LJ, 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. e141-147. e148.

Shrimpton, S, Daniels, K, De Greef, S, Tilotta, F, Willems, G, Vandermeulen, D, Suetens, P, et al. (2014). A spatially-dense regression study of facial form and tissue depth: towards an interactive tool for craniofacial reconstruction. *Forensic Science International*, 234, 103-110.

Shui, W, & Gao, F. (2021). A geometric completion and shape analysis method for damaged bilaterally symmetrical artifacts. *Journal of Cultural Heritage*, 52, 118-127.

Shui, W, Zhang, Y, Wu, X, & Zhou, M. (2020a). The three-dimensional facial

reconstruction of a male Upper Cave 101 skull (in Chinese). *Acta Anthropologica Sinica*, 39(04), 659-670.

Shui, W, Zhang, Y, Wu, X, & Zhou, M. (2021). A computerized facial approximation method for archaic humans based on dense facial soft tissue thickness depths. *Archaeological Anthropological Sciences*, 13(11), 186.

Shui, W, Zhou, M, Maddock, S, He, T, Wang, X, & Deng, Q. (2017). A PCA-Based method for determining craniofacial relationship and sexual dimorphism of facial shapes. *Computers in Biology and Medicine*, 90, 33-49.

Shui, W, Zhou, M, Maddock, S, Ji, Y, Deng, Q, Li, K, Fan, Y, et al. (2020b). A computerized craniofacial reconstruction method for an unidentified skull based on statistical shape models. *Multimedia Tools and Applications*, 79, 25589-25611.

Smith, OA, Nashed, YS, Duncan, C, Pears, N, Profico, A, & O'Higgins, P. (2021). 3D Modeling of craniofacial ontogeny and sexual dimorphism in children. *The Anatomical Record*, 304, 1918-1926.

Sorkine, O, Cohen-Or, D, Lipman, Y, Alexa, M, Rössl, C, & Seidel, H-P. (2004). Laplacian surface editing. *The 2004 Eurographics/ACM SIGGRAPH symposium on Geometry processing*, 175-184.

Stephan, C, & Henneberg, M. (2006). Recognition by forensic facial approximation: case specific examples and empirical tests. *Forensic science international*, 156(2-3), 182-191.

Stephan, CN. (2015). Facial approximation—from facial reconstruction synonym to face prediction paradigm. *Journal of Forensic Sciences*, 60(3), 566-571.

Stephan, CN. (2017). 2018 tallied facial soft tissue thicknesses for adults and sub-adults. *Forensic Science International*, 280, 113-123.

Toneva, D, Nikolova, S, Harizanov, S, Zlatareva, D, & Hadjidekov, V. (2021). A dense approach for computation of facial soft tissue thickness data. *Forensic Imaging*, 200460.

Turner, WD, Brown, RE, Kelliher, TP, Tu, PH, Taister, MA, & Miller, KW. (2005). A novel method of automated skull registration for forensic facial approximation. *Forensic Science International*, 154(2-3), 149-158.

Vanezis, P, Vanezis, M, McCombe, G, & Niblett, T. (2000). Facial reconstruction using 3-D computer graphics. *Forensic Science International*, 108(2), 81-95.

Verzé, L. (2009). History of facial reconstruction. *Acta Biomedica* 80, 5-12.

Wilkinson, C. (2005). Computerized forensic facial reconstruction. *Forensic Science, Medicine, and Pathology*, 1(3), 173-177.

Wilkinson, C. (2010). Facial reconstruction--anatomical art or artistic anatomy? *Journal*

of Anatomy 216(2), 235-250.

Wilkinson, C, Rynn, C, Peters, H, Taister, M, Kau, CH, & 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(2), 179-187.

Zhang, N, Zhao, J, Duan, F, Pan, Z, Wu, Z, Zhou, M, & Gu, X. (2022). An End-to-End Conditional Generative Adversarial Network Based on Depth Map for 3D Craniofacial Reconstruction. *The 30th ACM International Conference on Multimedia*, 759-768.

Chapter 8 Conclusion and future research

8.1 General summary

With the increased accessibility of large collections of 3D meshes, a wide range of data and statistical analysis and visualisation methods is now available to researchers, facilitating studies of classification and identification. Nevertheless, there are some common problems that need to be considered when using these methods (including geometric morphometrics) to examine morphological variations. The general aim of the thesis is to understand and compare how different semilandmarking approaches impact geometric morphometric analysis. To this end, the thesis explores the application of novel statistical analysis and visualisation methods to anthropological and archaeological data.

The thesis is divided into eight chapters (Figure 8.1). The first chapter provides background, motivation, research questions, and context, while introducing three specific applications that need to be solved in archaeology and anthropology fields. Chapters 2 and 3 lay the groundwork by elucidating the fundamental principles of geometric morphometrics and then investigate an impact of different semilandmarking approaches and densities of semilandmarks on morphometric analysis and visualisation, respectively. These two chapters provide new insights into selecting the appropriate landmark and semilandmark configuration for statistical testing and shape visualisation. Moreover, they assess the efficacy of the proposed semilandmarking approach (Shui et al., 2021), which combines the Thin-plate splines (TPS) and non-rigid iterative closest points (NICP) algorithms. Notably, these two chapters form the theoretical basis of the thesis, providing the guidance and conceptual framework to quantify surface asymmetry with regard to the plane of symmetry from 3D meshes (Chapter 4), restore missing geometry (Chapter 5), examine shape difference of artefacts (Chapter 5), establish dense point correspondences between dry skulls and modern human skulls (Chapters 6 and 7), and explore craniofacial relationships to perform facial approximation (Chapters 6 and 7).

Bilaterally symmetrical objects are common within the realms of archaeology and anthropology. Drawing inspiration from the framework of conventional landmark-based morphemic method (Damstra et al., 2012), Chapter 4 presents an innovative landmark-free method to identify the plane of symmetry from bilaterally symmetrical objects. This method facilitates the quantification of asymmetry and production of profile drawings of the bronze mask (Chapter 5). Additionally, guided by the principles of the deformation-based method commonly employed in the field of anthropology, Chapter 5 presents a novel virtual restoration method to restore the missing geometry of damaged objects. This method enables the warped template model to match the damaged object closely, thereby enhancing the restoration performance. These two chapters evolving the asymmetry detection and virtual restoration approaches can significantly assist in facial

approximation of *Homo sapiens* (Chapter 6) and archaic humans (Chapter 7). In Chapter 6, average facial soft tissue thickness depths (FSTDs) of modern humans at landmarks and high-density semilandmarks are used to represent craniofacial relationships. This, in turn, facilitates the estimation of the general face of an archaic human. Given the notable similarities among modern humans and difference from archaic humans, Chapter 7 extends the computerised approach developed in Chapter 6, in which nose and mouth craniofacial relationships between hard and soft tissues are quantified and nose and mouth soft tissues of modern humans are estimated. Ultimately, the facial statistical shape model is constructed to enhance the accuracy of facial approximation. Chapter 8 summaries the key findings and describes the potential for future work.

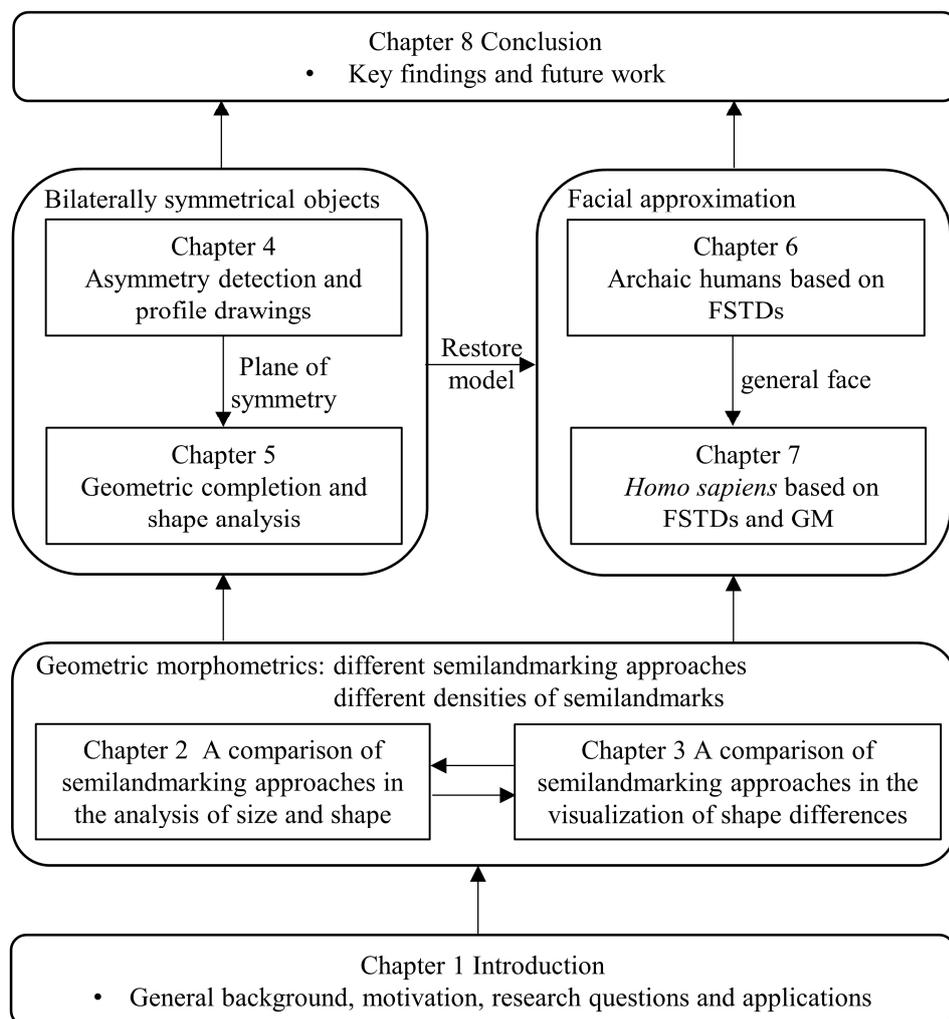


Figure 8.1 The structure of the thesis

In summary, the main objectives of the thesis have been threefold:

- Access how different semilandmarking approaches and densities of semilandmarks affect morphometric analysis (Chapter 2) and the visualisation of mean and allometrically scaled shapes (Chapter 3).
- Quantify the extent of gross and regional asymmetry in 3D anthropological and

archaeological data (Chapter 4) and then develop an approach for virtual restoration of missing geometry based on reflection (Chapter 5).

- Explore craniofacial relationships between bony structure and facial soft tissues among modern humans and then perform facial approximation by applying the learnt relationships to the dry skull of archaic humans (Chapter 6) and *Homo sapiens* (Chapter 7).

8.2 Key findings and discussion

8.2.1 Different semilandmarking approaches and densities

Landmarks characterized by their adherence to homology and reliable identification play a fundamental role in anthropological and archaeological studies. Semilandmarks have become increasingly common due to their potential for capturing detailed information about anthropological and archaeological materials. However, it is important to note that semilandmarks are mainly determined through mathematical algorithms based on topographic features, rather than developmental or evolutionary equivalences. Consequently, uncertainty about homology affects the semilandmarks and interpretations of morphometric analysis results derived from semilandmarks should be made with caution. It is crucial to highlight that the several factors influence the selection of a semilandmarking approach, such as coverage of landmarks, preservation of material, complexity of surface, time available for placement. Recently, landmark-based semilandmarking approaches have been presented in the field of anthropology and computer science. In these approaches, landmarks provide an initial map of equivalences, guiding subsequent mapping. However, a notable gap exists in the systematic investigation of the influence of various semilandmarking approaches and densities on the comparison of morphometric analysis and the visualization of mean and allometrically scaled surfaces. To the best of my knowledge, no previous studies have comprehensively addressed this specific aspect. In Chapters 2 and 3, I endeavour to answer this question by testing different hypothesis.

Chapter 2 evaluates the performance of three landmark-based semilandmarking approaches within the context of morphometric analysis. These approaches are tested using two different datasets with different degrees of variation and complexity. The findings of this analysis reveal that different approaches produce different semilandmark locations, subsequently leading to diverse statistical outcomes. More specifically, sliding TPS and TPS&NICP approaches produce very similar results, especially when landmarks are placed over the entire surfaces, such as ape crania and the facial region in head surface data. These similarities encompass semilandmark locations, mean landmark and semilandmark configurations, centroid sizes, distance matrices, principal components of shape variation, and allometrically scaled landmarks and semilandmark configurations.

The fundamental reason lies in the shared utilization of a triplet of Thin-plate splines (TPS) in both non-rigid registration methods, facilitating an initial deformation between template and each specimen. Conversely, semilandmarks generated by LS&ICP exhibit discrepancies, especially when analysing the ape crania dataset. The results of morphometric analyses conducted using these semilandmarks substantially differ from those derived through the application of the two non-rigid registration methods. This discrepancy can be attributed to the considerable variation in size between the template surface and each individual ape cranium. As a countermeasure, the isotropically scaled registration algorithm (Porto et al., 2021; Shui et al., 2016) can be performed as a prerequisite step, removing the size difference. In cases where no landmarks are available, sliding TPS and TPS&NICP encounter limitations. In such scenarios, ICP or NICP can be applied to register the template and each individual specimen while iteratively establishing dense point correspondences. However, we need to be cautious about the homology of correspondences given the absence of landmarks to guide semilandmarking.

In anthropology and biology, estimates of mean and allometrically scaled shapes often involves warping the template surface using TPS. When using the estimates of mean landmarks and high-density semilandmarks, the warping surface would be very similar to the true mean surface. However, this approach may introduce inaccuracies and incorrect distortions, particularly in regions distant from landmarks, where bending energy has less effect (Klingenberg, 2013; Schlager et al., 2018). Consequently, caution is warranted in interpreting the estimates of surfaces. To address this problem, landmarks should be placed to cover the entire surface of interest and then semilandmarks over curves and surfaces, especially in smooth regions, are required to ensure accurate surface warping. For example, in Chapter 6, the average skull is deformed to the dry skull based on 91 landmarks and 404 semilandmarks together and then NICP are applied to enable the warped template to match the dry skull closely, thereby enhancing the accuracy of correspondences.

Additionally, Chapter 3 extends this investigation to examine the effects of three different landmark-based semilandmarking approaches and densities of semilandmarks on estimates of mean and allometrically scaled surfaces. The results indicate consistent findings between surfaces generated by sliding TPS and TPS&NICP approaches, whereas surfaces derived from LS&ICP approach exhibit more pronounced discrepancies. This concurrence with the findings of the morphometric analyses in Chapter 2 emphasizes the importance of employing both statistical tests and visualization analyses for robust interpretation. Furthermore, the warping of surfaces using solely landmarks can yield surfaces that can be notably different from those created through landmarks and semilandmarks. This discrepancy becomes particularly pronounced when the selected template surface differs from the mean surface, as observed in the case of the ape cranium.

The primary underlying reason is that a limited number of landmarks may be insufficient to offer comprehensive information, particularly in cases where specimens display complex structures.

Based on the results of Chapters 2 and 3, a strategic guideline for selecting landmarks and semilandmarks is presented for the application of geometric morphometrics. When landmarks with regard to the research question are sufficient, such as when they can encompass the entire surface of interest, statistical analysis may proceed based solely on landmarks. This strategy can mitigate the uncertainties of homology of semilandmarks in shape analysis. In instances where semilandmarks are needed to provide detailed information, a suitable method for semilandmarking could involve landmark-based non-rigid registration, such as sliding TPS, TPS&NICP, and the mean surface generated by the estimates of landmarks and semilandmarks are regarded as the template surface to yield semilandmarks. In terms of visualisation, high-density semilandmarks are frequently employed to minimise the template bias in the generation of mean and allometrically scaled surfaces. It is important to note that morphometric analysis and visualization should complement and maintain consistency with each other.

8.2.2 Asymmetry detection and virtual restoration

The recognition of the plane of symmetry plays a crucial role in asymmetry detection and virtual restoration. While a variety of computational methods has been developed, the diverse applications and contexts make it difficult to use off-the-shelf tools and methods. For example, landmark-based approaches are unsuitable for analysing large samples. In this context, Chapter 4 presents an effective landmark-free approach to approximate the plane of symmetry from nearly bilaterally symmetrical objects by means of finding the plane with the minimum geometric differences between the original and mirrored meshes. Notably, this method can be applied to both complete objects and partial geometries. An essential step involves the automatic extraction of the symmetrical regions, such as the nasal bridge, from 3D meshes. This extraction mainly relies on intrinsic features that remain invariant under isometric mapping. Upon identifying the symmetric regions, the proposed method can improve the accuracy of plane of symmetry.

Nevertheless, it is important to acknowledge that the process of region growing and merging, and plane clustering is computationally demanding and consumes a significant amount of time. As a result, the experiments conducted in Chapter 4 primarily focused on analysing external surfaces, with interior surfaces omitted. Several factors, such as surface complexity, the selection of intrinsic features, and thresholds employed in the region growing, can influence the accuracy of extracted symmetrical regions. This, in turn, leads to an inaccurate recognition of plane of symmetry. To address this challenge, researchers can utilise an interactive tool for manually marking up symmetric regions and then identify the plane (Xu et al., 2022). Alternatively, novel methods for recognising

symmetrical pairwise correspondences can be developed, which take advantage of both extrinsic and intrinsic features (Fotouhi et al., 2021). These strategies provide avenues for enhancing the accuracy of symmetrical region identification and ultimately addressing the challenge of accurately identifying the symmetry plane.

By means of reflection of the intact side of damaged objects as a template surface, non-rigid deformation, such as TPS, are commonly employed to repair small regions of missing geometry. As noted in an earlier study (Schlager et al., 2018), it is worth noting that regions situated far from landmarks could potentially yield inaccurate results. Chapter 5 presents a hybrid non-rigid deformation to closely align the warped template surface with the damaged object, thereby enhancing the accuracy of restoration. In this process, an initial warping outcome is achieved using TPS and subsequently Laplacian deformation is applied to refine the deformation accuracy with the predicted boundary curve as a constraint. These deformation-based methods exhibit effectiveness when dealing with small missing regions. However, when handling large missing geometries, particularly in cases where a substantial portion of the object is damaged, our method seems ineffective. This is because the restored meshes would resemble the chosen template surface, and local constraints may fail to offer adequate information to guide accurate restoration.

To tackle this challenge, a statistical shape model offers a promising solution by providing a prior knowledge to potentially capture statistical variations and shape distributions within a training dataset. As noted in the earlier study (Cootes et al., 1995), this method was initially used in computer vision and recently it has been increasingly applied in the fields of biology, anthropology and medical surgery (Brunton et al., 2014). This is because advancements in data acquisition allow access to a large number of fossil specimens. For example, statistical shape models have been employed in pelvic defects (Meynen et al., 2020). In repairing missing geometry, the process involves the establishment of dense point correspondences between the damaged object and training dataset. Subsequently, the registered damaged object is projected onto the shape space to compute the coefficients based on the prior distribution. Regularization techniques are frequently used to optimise the coefficients, minimising the overfitting and prediction errors. Nevertheless, one of the main challenges is the necessity for an extensive and diverse dataset used to represent actual shape variability. Therefore, there is an urgent need to encourage researchers to openly share and reuse data to construct robust statistical shape models, thereby enhancing restoration accuracy.

8.2.3 Facial approximation

Computerised facial approximation is a promising means for generating the possible facial appearance of a dry skull, significantly reducing subjective interpretation and the demands for specialised training in physical modelling skills. Recent studies (Gietzen et

al., 2019; Li et al., 2022) have demonstrated these methods can produce repeatable and objective results. Notably, achieving this objective involves two critical aspects: the quantification of craniofacial relationships between bony structures and facial soft tissues of modern humans, and the assignment of the learnt relationships to the dry skull. Chapter 6 presents a novel computerised facial approximation method for archaic humans based on average FSTDs of modern humans. Building upon this analysis, Chapter 7 focuses on quantification of the relationships between nasal and oral hard and soft tissues using geometric morphometrics and then develops a novel computerised approach for *Homo sapiens*.

In the process of facial approximation, the main challenges include the poor preservation of human fossils and the lack of anatomical knowledge of their craniofacial relationships. Thus, the prerequisite step is the examination and restoration of the dry skull. Since the mandibles of archaic humans are often not found, a common task involves locating a well-preserved mandible from the same species that fits well with the cranium. In such situations, conventional deformation-based (Chapters 4 and 5) or statistical shape model approaches cannot be applied. As a result, the restoration of archaic human skulls presents unique challenges. Mandibles of similar age, such as Montmaurin-La Niche (Violet et al., 2018), were not available, therefore two available mandibles, Tabun 2 and Mauer 1, were instead used with the Jinniushan (JNS) 1 cranium in Chapter 6. Notably, the ages of these two mandibles differ significantly from JNS1, potentially leading to incorrect bony restoration and a less confident approximation result. To select the most suitable mandible, several factors, such as temporal proximity, regional relevance, and the available and quality of the fossil material, should be taken into account. In the future, I will endeavour to gather different mandible models and then reproduce and reassess the facial appearance of JNS1.

It is important to note that skull geometry plays a significant role in determining facial structure. Bony structures influence the positions and proportions of facial soft tissues and different regions of the skull, such as the size and shape of the orbits, nasal aperture, zygomatic arches, and mandible, can provide valuable clues about facial features. To examine how skull morphology affects soft tissues, Chapter 6 computes the geometric differences between two approximated faces that are recreated using Tabun 2 and Mauer 1, respectively. The greatest changes can be observed around the lower face. Specifically, the approximation of JNS 1 using Mauer 1 has a wider face and a more robust chin compared to the approximation using Tabun 2. This result is consistent with geometric difference between these two mandibles. When different mandibles are used to fit with the JNS 1 cranium, the approximated face varies. Furthermore, Chapter 7 examines the extent to which nasal and oral hard tissues influence their soft tissues. The results indicate that overall changes in nose and mouth soft tissues appear to correlate

with those of the hard tissues. For instance, narrow and enlarged oral hard tissue shapes with increasingly protruding anterior teeth are correlated with a mouth having narrow and protruding upper and lower lips. However, approximation errors mean that we need to be cautious about the accuracy of the approximated nose and mouth soft tissue shapes. Therefore, artistic interpretation and imagination may require consideration as approaches to enhance reconstructions of facial features.

The accuracy of facial approximation is influenced by the learnt craniofacial relationships, such as average FSTDs at landmarks and semilandmarks. In practical applications, average FSTDs with regard to age, sex, body mass index (BMI), ethnicity have been used for facial approximation of modern humans. The results of Chapters 6 and 7 demonstrate that average FSTDs at landmarks and semilandmarks contribute to raising the accuracy of the approximated face. However, the limited sample size of modern humans may lead to imprecise descriptive statistics of FSTDs and statistical issues in regression of soft on hard tissues. To obtain more accurate craniofacial relationships, a comprehensive skull and face model dataset taking into account age, sex, BMI, ethnicity is needed. Given that facial morphology can be influenced by evolutionary, genetic and environmental factors, the relationship between soft and hard tissues of archaic humans probably differs significantly from that of modern humans. Hence, the prediction of facial appearance of archaic humans poses a significant challenge compared to *Homo sapiens*. To enhance the accuracy of prediction in archaic humans, an intriguing solution is to explore whether chimpanzee FSTDs can perform well for JNS 1 which might share more similarities with chimpanzees than with modern humans.

A previous study (Hayes et al., 2013) suggests that chimpanzee FSTDs are approximately half of modern humans in the area of the cheeks, resulting in production of a thinner mid-face. If chimpanzee FSTDs at landmarks and semilandmarks are available, I will endeavour to reproduce and reassess the facial appearance of JNS1. In Chapter 6, average FSTDs of modern humans are regarded as those of JNS 1, and the warped facial features of the average human face are applied to JNS1. Approximation errors are inevitable and we need to be cautious about the accuracy of approximated facial features. Leveraging knowledge from comparative anatomy, biomechanics, and evolutionary biology might contribute to improving accuracy of the approximated face. Moreover, in the application of the anatomically modern human fossils, whose craniofacial relationships closely resemble those of modern humans, the approximated faces should be more confident than those of archaic humans.

The approximated faces provide new insights into the characteristic features of human fossils and the evolutionary forces driving anatomical changes in ancestral humans. Based on the average FSTDs of modern humans, Chapters 6 and 7 recreate the approximated faces of the Upper cave (UC) 101 and JNS 1 skulls, respectively. The

approximated face of UC 101 skull exhibits an elongated shape, a sloped forehead, and stronger and wider eyebrows. Conversely, the approximated face of JNS 1 presents a relatively lower forehead, robust eyebrows, a protruding, wider, and elongated middle and upper face. Compared to the general profile of the approximated face of JNS 1, the approximated face of UC 101 appears plausible and lifelike, and bears a stronger resemblance to modern human faces. This distinction arises from the impact of skull morphology on the resulting approximated face. For instance, UC 101 more closely resemble modern human skulls, whereas JNS 1 exhibits distinct shapes and features.

8.3 Future research

Landmarks serve as a crucial foundation of comparison of anthropological specimens. Often they are used to guide and control semilandmarking, i.e. by controlling the deformation of the template surface when marking up semilandmarks. However, manual placement of landmarks consumes time and requires anatomical knowledge. They are also prone to error, particularly in regions lacking distinctive features among a substantial collection of specimens. Thus, an automated landmarking approach is desirable. Although there are different strategies and methods for automatically estimating the locations of landmarks among specimens (e.g. human heads), these methods seem not to work well for all surfaces, especially in complex structures. For instance, the deformation-based methods mainly rely on the warping results and topographical features, and the Active shape model constructed by the Principal component analysis algorithm may not capture intricate shape variations and be sensitive to the initial estimate of the shape. Future work should consider the use of an end-to-end deep learning approach for landmark detection based on a small number of specimens with expertly placed landmarks as a training dataset.

Landmarks and high-density semilandmarks (or dense point correspondences) have been used to estimate the mean and allometrically scaled shapes, predict missing geometry and accomplish facial approximation. One concern is to assess the estimating results by means of visually interpreting the local difference between surfaces (e.g. the difference in estimates of mean shapes generated by different approaches, and resemblance similarity between the approximated and actual faces). Colour maps of geometric similarity between the aligned two surfaces are often used to quantify the regional difference in many different applications. However, the alignment results are extremely sensitive to the choice of landmarks and semilandmarks, thereby leading to inevitably different results. In my thesis, registration free colour maps of the distribution of surface area differences between equivalent triangles (Chapter 3) and the difference in FSTDs between the approximated and template faces (Chapter 6) are visualised. However, this method cannot work when the generated surfaces have different vertices and mesh topologies. Thus, novel visualisation methods still need to be developed and applied to

visualise regional differences.

Facial approximation has wide applications in forensic science, anthropology and archaeology. The quantification of craniofacial relationships is an important and inevitable step in facial approximation. In the present study, average FSTDs at landmarks and semilandmarks are used to quantify the overall relationships, and the results of partial least squares (PLS) suggest there seems to be a linear relationship between nasal (and oral) hard and soft tissues separately. However, multiple linear regression appears not to generate more accurate nose and mouth soft tissue shapes from bony structure alone. Deep learning with artificial neural networks is a powerful tool which has been increasingly applied in many tasks, achieving superior accuracy when trained a large dataset. Further work should increase the collection of skull and face data and design novel neural network models to enhance the approximation accuracy of facial features.

8.4 References

- Brunton, A, Salazar, A, Bolkart, T, & Wuhrer, S. (2014). Review of Statistical Shape Spaces for 3D Data with Comparative Analysis for Human Faces. *Computer Vision and Image Understanding*, 128(11), 1-17.
- Cootes, TF, Taylor, CJ, Cooper, DH, & Graham, J. (1995). Active shape models-their training and application. *Computer vision and image understanding*, 61(1), 38-59.
- Damstra, J, Fourie, Z, De Wit, M, & Ren, Y. (2012). A three-dimensional comparison of a morphometric and conventional cephalometric midsagittal planes for craniofacial asymmetry. *Clinical Oral Investigations*, 16(1), 285-294.
- Fotouhi, J, Taylor, G, Unberath, M, Johnson, A, Lee, SC, Osgood, G, Armand, M, et al. (2021). Exploring partial intrinsic and extrinsic symmetry in 3D medical imaging. *Medical image analysis*, 72, 102127.
- Gietzen, T, Brylka, R, Achenbach, J, zum Hebel, K, Schömer, E, Botsch, M, Schwanecke, U, et al. (2019). A method for automatic forensic facial reconstruction based on dense statistics of soft tissue thickness. *Plos one*, 14(1), e0210257.
- Hayes, S, Sutikna, T, & Morwood, M. (2013). Faces of Homo floresiensis (LB1). *Journal of Archaeological Science*, 40(12), 4400-4410.
- Klingenberg, CP. (2013). Visualizations in geometric morphometrics: how to read and how to make graphs showing shape changes. *Hystrix, the Italian Journal of Mammalogy*, 24(1), 15-24.
- Li, Y, Wang, J, Liang, W, Xue, H, He, Z, Lv, J, & Zhang, L. (2022). CR-GAN: Automatic craniofacial reconstruction for personal identification. *Pattern Recognition*, 124, 108400.
- Meynen, A, Matthews, H, Nauwelaers, N, Claes, P, Mulier, M, & Scheys, L. (2020). Accurate reconstructions of pelvic defects and discontinuities using statistical shape models. *Computer Methods in Biomechanics and Biomedical Engineering* 23(13), 1026-1033.
- Porto, A, Rolfe, S, & Maga, AM. (2021). ALPACA: A fast and accurate computer vision approach for automated landmarking of three-dimensional biological structures. *Methods in Ecology and Evolution*, 12(11), 2129-2144.
- Schlager, S, Profico, A, Di Vincenzo, F, & Manzi, G. (2018). Retrodeformation of fossil specimens based on 3D bilateral semi-landmarks: Implementation in the R package “Morpho”. *PloS one*, 13(3), e0194073.
- Shui, W, Zhang, Y, Wu, X, & Zhou, M. (2021). A computerized facial approximation method for archaic humans based on dense facial soft tissue thickness depths. *Archaeological Anthropological Sciences*, 13(11), 186.

Shui, W, Zhou, M, Deng, Q, Wu, Z, Ji, Y, Li, K, He, T, et al. (2016). Densely Calculated Facial Soft Tissue Thickness for Craniofacial Reconstruction in Chinese Adults. *Forensic Science International*, 266, 573.e571-e512.

Vialet, A, Modesto-Mata, M, Martín-Torres, M, Martínez de Pinillos, M, & Bermudez de Castro, J-M. (2018). A reassessment of the Montmaurin-La Niche mandible (Haute Garonne, France) in the context of European Pleistocene human evolution. *PLoS One*, 13(1), e0189714.

Xu, J, Cao, W, Liu, B, & Jiang, K. (2022). Object restoration based on extrinsic reflective symmetry plane detection. *The Visual Computer*, 38, 3627-3642.