

Dense Correspondence and Statistical Shape Reconstruction of Fractured, Incomplete Skulls



Master Thesis

by

Zhang Kun

under the guidance of

A/Prof. Leow Wee Kheng

SCHOOL OF COMPUTING

NATIONAL UNIVERSITY OF SINGAPORE

July 2014

Declaration

I hereby declare that this thesis is my original work and it has been written by me in its entirety. I have duly acknowledged all the sources of information which have been used in the thesis.

This thesis has also not been submitted for any degree in any university previously.

Signature: _____

Date: _____

Acknowledgements

First, I would like to thank my supervisor, Professor Leow Wee Kheng, for his guidance and continuous support during the whole period of my study. Prof. Leow guided me to embark on the research on Nov. 2012. He taught me the way of initiating a good and useful research topic, formulating a research problem, solving the problem, conducting experiments, and writing scientific reports, etc. He was always responsible and provided encouragements and insightful ideas to me whenever I faced problems. Without Prof. Leow's great help, it is impossible for me to complete this thesis.

Also, I am grateful to my friends and colleagues: Cheng Yuan, Zhang Li and Xie Shudong. These brilliant people provided me with insightful discussions on research and rendered my years in Singapore enjoyable.

Abstract

It is common to encounter fractured or incomplete skulls in surgery, forensics and archeology. Reconstructing normal, complete skulls from fractured, incomplete skulls is a very important task. In surgery planning, the normal appearance of the skull can help to determine the correct positions and orientations of the fractured bones required to restore the deformed skull. In forensic investigation and archeology, the normal, complete skull provides clues for the facial appearance of the person involved.

Unfortunately, reconstruction of skulls is very difficult and challenging. The original normal, complete skulls of the subjects with fractured, incomplete skulls are normally unavailable. The reconstruction method has to predict the normal shapes from the normal parts of the fractured, incomplete skulls. Although human skulls have the same global structure, they differ greatly in details among people with different races, genders and ages. Therefore, reconstructing the normal, complete skulls is a very difficult task.

Among existing skull reconstruction algorithms, symmetric-based reconstruction produces a reconstructed skull model based on left-right symmetry of human skulls. It requires the presence of healthy bones to reconstruct the fractured parts on the opposite side of the skull. When both sides of the skull are fractured, this approach cannot be applied. Geometric reconstruction reconstructs a deformed skull by deforming and registering a reference model to the deformed skull. It depends highly on the similarity between the reference model and the deformed skull to obtain accurate results. Statistical reconstruction overcomes the weaknesses of these approaches by using statistical models of human skulls.

Active shape model (ASM) is by far the most popular statistical reconstruction method. To build an ASM of the skull from a set of training samples, it is necessary

to first establish the correspondence between the training samples. Among existing correspondence building algorithms, mesh registration is the predominant approach and it uses landmarks to guide the deformation and matching process. Manual marking of landmarks is too tedious to perform on the entire skull. On the other hand, methods that automatically detect correspondence based on local geometric features are sensitive to noise and outliers.

This thesis aims to overcome the difficulties of existing methods for skull reconstruction. It proposes an accurate and efficient dense correspondence algorithm for 3D skull models. The proposed algorithm uses thin-plate spline (TPS) for non-rigid registration. It uses anatomical landmarks as hard constraints and control points sampled on the skull surfaces as soft constraints. Then, the proposed dense correspondence algorithm is used to improve the accuracy of ASM reconstruction for fractured, incomplete skulls. In addition, the source of error of ASM skull reconstruction is investigated, and a novel method is proposed to estimate the reconstruction error of real fractured, incomplete skulls. Such estimation of the reconstruction error of real samples has not been reported in existing literature.

Test results show that the proposed algorithm can achieve more accurate and robust dense correspondence than other TPS-based algorithms that use only hard constraints or soft constraints but not both. With the accurate dense correspondence, the accuracy of skull reconstruction also increases. Consequently, the main error source of skull reconstruction is restricted to the ASM fitting process. A linear relationship between ASM fitting error and reconstruction error is observed, which leads to a method for estimating the reconstruction errors of real fractured, incomplete skulls using ASM fitting error.

Contents

List of Publications	IV
List of Figures	V
List of Tables	VIII
1 Introduction	1
1.1 Motivation	1
1.2 Thesis Objectives	4
2 Related Work	7
2.1 Building Dense Correspondence	7
2.1.1 Mesh Registration	7
2.1.2 Surface Parameterization	9
2.1.3 Group Representation	10
2.2 Skull Reconstruction	11
2.2.1 Symmetry-based Reconstruction	11
2.2.2 Geometric Reconstruction	12
2.2.3 Statistical Reconstruction	13

2.3	Summary	15
3	Dense Correspondence with Thin-Plate Spline	18
3.1	Automatic Craniometric Landmark Detection	19
3.2	Non-Rigid Registration with Thin-Plate Spline	23
3.3	Correspondence and Resampling	25
3.4	Error Measures	28
3.5	Experiments and Discussions	30
3.6	Summary	34
4	Active Shape Skull Model	36
4.1	Active Shape Skull Model	36
4.2	Experiments and Discussions	38
4.2.1	Data Preparation and Procedures	38
4.2.2	Error Measures	40
4.2.3	Results and Discussions	40
4.3	Summary	43
5	Reconstruction of Skulls	45
5.1	Active Shape Skull Reconstruction	45
5.2	Reconstruction of Fractured, Incomplete Skulls	47
5.3	Experiments and Discussions	48
5.3.1	Data Preparation and Test Procedure	48

5.3.2	Resampling Error	53
5.3.3	ASM Fitting Error	62
5.3.4	Skull Reconstruction Error	66
5.3.5	Estimation of Reconstruction Error of Real Samples	74
5.4	Summary	77
6	Conclusion	79

List of Publication

K. Zhang, Y. Cheng and W.K. Leow. Dense correspondence of skull models by automatic detection of anatomical landmarks. In *Int. Conference on Computer Analysis of Images and Patterns (CAIP)*, 2013.

List of Figures

- 1.1 Skull deformities 2
- 1.2 3D skull models have inner and outer surfaces 5

- 2.1 Geometric reconstruction 13
- 2.2 Statistical shape model of lower jaw 14

- 3.1 Reference skull model 20
- 3.2 Landmarks used by correspondence algorithm 22
- 3.3 Control points sampling and correspondence searching 27
- 3.4 Resampling of a fractured, incomplete skull 28
- 3.5 Manually marked landmarks 31
- 3.6 Sample resampling results 33
- 3.7 Effect of control points on correspondence accuracy 35

- 4.1 Examples of skull models used for training 39
- 4.2 Residual plot of ASM skull modeling 41
- 4.3 Modes of variation 42
- 4.4 Influence of number of eigenvectors on reconstruction error 43

5.1	Shape vector generation	47
5.2	Examples of normal, complete skull models used for testing	49
5.3	Common skull fracture patterns	50
5.4	Synthetic fractured samples	50
5.5	Synthetic incomplete skulls	51
5.6	Synthetic testing samples with multiple defects	52
5.7	Real fractured, incomplete testing samples	52
5.8	Sample resampling results of fractured, incomplete skulls	55
5.9	Resampling results on training samples and normal, complete testing samples	55
5.10	Hole filling in the resampling process	56
5.11	Resampling results on synthetic fractured, incomplete testing samples . . .	57
5.12	Resampling results on real fractured, incomplete testing samples	58
5.13	Reconstructed results of training samples and normal, complete skull test- ing samples	67
5.14	Reconstruction results of synthetic fractured, incomplete testing samples .	68
5.15	Reconstruction results of ABM and AKM	69
5.16	Reconstruction results of NAV and OSL	70
5.17	Reconstruction results of SAMK and NGMSRE	71
5.18	A plot of resampling error vs. reconstruction error	75
5.19	Plots of ASM fitting error vs. reconstruction error	76

List of Tables

2.1	Comparison of skull reconstruction algorithms	16
3.1	Quantitative evaluation	32
5.1	Resampling errors of training samples, normal testing samples and synthetic incomplete testing samples	60
5.2	Resampling errors of synthetic fractured testing samples and testing samples with multiple defects	61
5.3	Resampling errors of real fractured, incomplete testing samples	62
5.4	ASM fitting errors of training samples, normal, complete testing samples and synthetic incomplete testing samples	64
5.5	ASM fitting errors of synthetic fractured testing samples and testing samples with multiple defects	65
5.6	ASM fitting errors of real fractured, incomplete testing samples	65
5.7	Reconstruction error of training samples, normal testing samples and synthetic incomplete samples	73
5.8	Reconstruction error of synthetic fractured testing samples and testing samples with multiple defects	73
5.9	Evaluation of fitted lines	77

Chapter 1

Introduction

1.1 Motivation

Many patients suffer from some forms of skull deformity which can be due to impact fracture or congenital defects (Figure 1.1). In North American, out of 1000 people, 2.9 suffer from head and facial injuries due to traffic accidents, work accidents, home accidents, sports injuries, and violence every year [GTH⁺03]. Worldwide, 2 in 100 children are born with facial asymmetry [Lit]. In Singapore, National University Hospital (NUH) alone receives about 250–350 patients with skull fractures every year. Fractured skulls may be incomplete because some bones may be pulverized on impact. In surgery planning of congenitally deformed skulls, the surgeons will cut away the deformed parts, resulting in incomplete skulls. In CT scanning, it is common to scan only the affected parts of the skulls resulting in incomplete skulls in the scanned images. In addition, in forensic investigation and archeology, the subjects' skulls are often fractured or incomplete due to criminal acts or natural processes.

Reconstructing normal, complete skulls from fractured, incomplete skulls is a very important task. In surgery planning, the normal appearance of the skull can help to determine the correct positions and orientations of the fractured bones required to restore

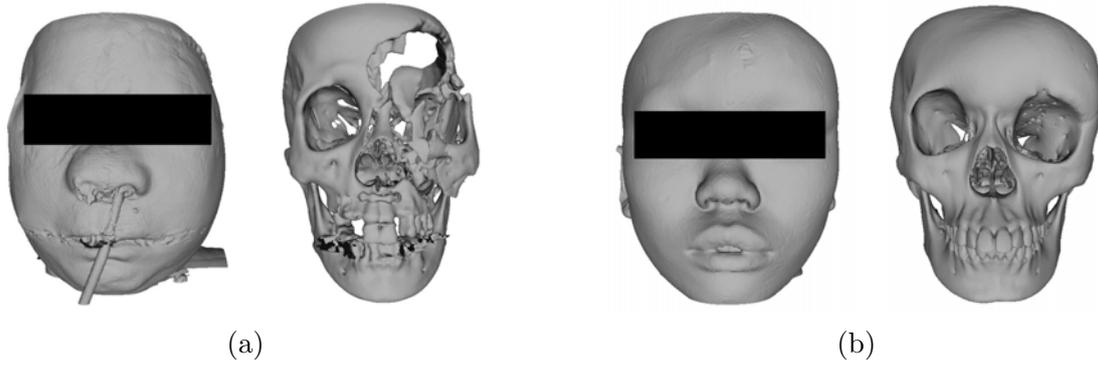


Figure 1.1: Skull deformities. (a) Skull fractures due to accident. (b) Congenital skull deformity due to brain tumor.

the deformed skull. In forensic investigation and archeology, the normal, complete skull provides clues for the facial appearance of the person involved. Unfortunately, reconstruction of skulls is very difficult and challenging. The original normal, complete skulls of the subjects with fractured, incomplete skulls are normally unavailable. The reconstruction method has to predict the normal shapes from the normal parts of the fractured, incomplete skulls. Although human skulls have the same global structure, they differ greatly in details among people with different races, genders and ages. Therefore, reconstructing the normal, complete skulls is a very difficult task.

Currently, several approaches have been used for skull reconstruction. A commonly used approach is symmetric-based reconstruction [GWL99, LCL⁺03, ECP⁺06], which produces a reconstructed skull model based on left-right symmetry of human skulls. Symmetric-based approach requires the presence of healthy bones to reconstruct the fractured parts on the opposite side of the skull. When both sides of the skull are fractured, which is common in impact injuries, this approach cannot be applied. Geometric reconstruction [WYLL11, GMBW04, BSMG09] has also been used. Geometric reconstruction deforms a reference skull model to estimate the normal parts of the deformed skull. This approach depends highly on the similarity between the reference skull and the deformed skull to obtain accurate results. Statistical reconstruction [Gun05, ZLES05, LAV09] overcomes the weaknesses of the other approaches. Instead of using a single reference skull

model, this approach constructs a statistical model of possible variations of human skulls and fits the statistical model to the normal parts of a fractured, incomplete skull and use the fitted model to reconstruct the normal appearance of the fractured, incomplete skull.

Active shape model (ASM) is by far the most popular statistical reconstruction method. To build an ASM of the skull from a set of training samples, it is necessary to first establish the correspondence between the training samples. Establishing correspondence refers to the process of ascertaining which part of one model corresponds to which part of another model. In some applications, such as face recognition, camera self-calibration, a set of sparse correspondences is sufficient [MC03, ZCJX12, LN07]. Skulls, on the other hand, have very complex shape and usually have more than 10000 mesh vertices. So dense correspondence is needed. Establishing accurate dense correspondence is a challenging and crucial part of ASM construction. As the ASM tries to capture the shape variations of the training samples, wrong correspondences will introduce unwanted shape variations.

There are three approaches for computing dense correspondence between two shape models: mesh registration, surface parameterization, and group representation. Among them, mesh registration is the predominant approach, which can be differentiated according to the registration method used. The simplest approach applies rigid registration [BM92, RHB97]. It is efficient but unable to match complex shapes closely. Non-rigid registration [STA98, STA99, ST03, ACP03, SU05, BDGY04, ZWY10, FL98, CR00] can achieve close matching but needs landmarks to guide the deformation and matching process. Methods that use manually marked landmarks are accurate [CR00, DZS⁺11, HBH03], but manual marking is too tedious to perform on the entire skull. On the other hand, methods that automatically detect correspondence based on local geometric features [TBK⁺05] are easy to apply but they are sensitive to noise and outliers, which can adversely affect their accuracy.

Surface parameterization [Ben91, FH05] and group representation [KT98, Dav02] are two other approaches for computing dense correspondence. Although they have good

theoretical foundations, they are applicable only to simple shapes. It is technically very difficult to apply them to complex shapes such as skulls.

In summary, ASM is a very promising method for skull reconstruction, but it requires accurate dense correspondence. Moreover, the accuracy of skull reconstruction methods have not been thoroughly investigated in existing literature.

1.2 Thesis Objectives

The overall goal of this research is to develop methods for reconstructing normal, complete skulls from fractured, incomplete skulls for various real-world applications. To achieve this goal and overcome some of the difficulties discussed in Section 1.1, this thesis focuses on three main contributions:

- Development of an accurate and efficient dense correspondence algorithm for 3D skull models. The proposed algorithm uses thin-plate spline (TPS) for non-rigid registration. It uses anatomical landmarks as hard constraints, and control points sampled on the skull surfaces as soft constraints. The landmarks ensure anatomically consistent correspondence, and the control points provide additional local shape constraints to ensure close matching of the reference and target surfaces.
- Application of the proposed dense correspondence algorithm to improve the accuracy of ASM reconstruction for fractured, incomplete skulls. As the proposed dense correspondence algorithm provides more accurate dense correspondence, the ASM is able to obtain more accurate reconstruction results.
- Investigation of the source of error of ASM skull reconstruction and development of a method to estimate the reconstruction error of real fracture, incomplete skulls without ground-truth. The source of error has never been thoroughly investigated in existing literature such as [Gun05, ZLES05, LAV09]. This thesis investigates the possible error sources due to correspondence building and ASM fitting. By

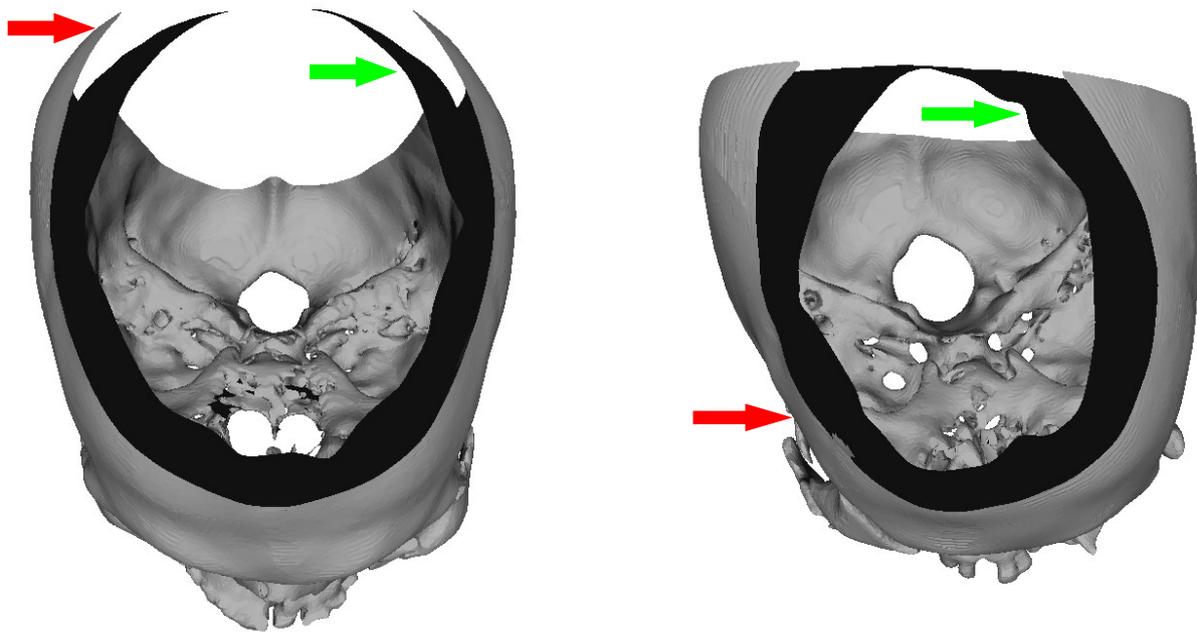


Figure 1.2: 3D skull models have inner and outer surfaces. The red arrows point to the outer surfaces, and the green arrows point to the inner surfaces.

studying the errors of a wild range of testing samples, a method for estimating the reconstruction error of real fractured, incomplete skull is developed. Such an error estimation method for real data has not been reported in the literature.

Note that bones have measurable thickness. Therefore 3D skull models have inner and outer surfaces (Figure 1.2). In applications such as surgery planning, forensic investigation and archeology, only the outer surfaces are important because they determine the shape of the face. So, this thesis focuses on the outer surfaces only. The outer surface is extracted based on the surface normal. For each vertex, a ray is shot outward from the center of the skull model to this vertex. If the surface normal at this vertex has the same direction with the ray, and the ray does not intersect any other mesh surfaces beyond this vertex, then the vertex is on the outer surface.

In Chapter 2, a detailed review of existing correspondence building and skull reconstruction algorithms is presented. The strength and weakness of these methods are

discussed. To overcome the weaknesses of existing correspondence algorithms, Chapter 3 presents a novel dense correspondence algorithm. Chapter 4 presents the active shape model (ASM) algorithm for normal, complete skull, and Chapter 5 presents the ASM skull reconstruction algorithm for fractured, incomplete skulls. Chapter 6 concludes with a discussion of possible future work.

Chapter 2

Related Work

This chapter reviews existing approaches for building dense correspondence and skull reconstruction. First, three types of dense correspondence approaches are discussed in Section 2.1, namely mesh registration, surface parameterization and group representation. Next, existing skull reconstruction methods, including symmetry-based reconstruction, geometric reconstruction and statistical reconstruction, are analyzed in Section 2.2.

2.1 Building Dense Correspondence

There are three existing approaches for building dense correspondences between mesh models, namely, mesh registration, surface parameterization and group representation. Each of these approaches is discussed in a following section.

2.1.1 Mesh Registration

Mesh registration approach registers a reference model to a target model, and then transfers the vertex connectivity of the reference to the target. The registration process usually includes rigid registration which gives a coarse registration, followed by non-rigid registration which gives an accurate registration.

Rigid registration is the task of registering two objects without changing their shapes. Shape is a property that is invariant to translation, rotating and scaling. Iterative Closest Points (ICP) by Besl and McKay [BM92] and the Softassign Procrustes by Rangarajan et al. [RHB97] are two popular algorithms for rigid registration. Both of them accept two meshes with potentially different number of vertices as input and compute the optimal similarity transformation from one to the other. These methods are simple to apply, but if the target has missing parts or outliers, the registration accuracy will be affected. A more robust method is the Fractional ICP (FICP) proposed by Philips et al. [PLT07]. Instead of using all reference vertices, FICP computes the similarity transformation using only a subset of target vertices whose distances to the reference model are the smallest. Thus the outliers are excluded in computing the similarity transformation. Several simple correspondence algorithms directly apply rigid registration methods such as ICP [JEK05, VDBA⁺04]. They are not expected to work well for complex shapes such as skulls.

Various non-rigid registration methods have been used to compute correspondence, for example, Thirion’s Demons registration [LAV09, Thi98], mass-spring model [ST03], local affine transformation [ACP03], free-form elastic transformation [SU05], trilinear transformation [BDGY04], graph and manifold matching [ZWY10], octree-splines [FL98], and thin-plate spline (TPS) [CR00, DZS⁺11, HDZ⁺12a, HBH03, LP00, LK00, RB02, TBK⁺05]. Most of these methods are applied to models with simple surfaces such as faces [HBH03, PHWZ10, ZWY10], human bodies [ACP03, ST03], knee ligaments [FL98], and lower jaws [BDGY04]. Among the various non-rigid registration methods, TPS is the most frequently used. TPS ensures smooth deformation and has few parameters that need manual tuning. It has close-form solutions so it is efficient. In addition, TPS has a sound physical explanation for its energy function. TPS method aims to determine a mapping function between an input mesh and a deformed mesh by minimizing an energy function that models the deformation as the bending of a thin sheet of metal [Wik]. It is particularly effective for mesh models with highly complex surfaces such as brain sulci [CR00], lumbar vertebrae [LK00], and skulls [DZS⁺11, HDZ⁺12b, LP00, RB02, TBK⁺05]. Skull models are particularly complex because they have holes, teeth, bones, a generally

convex outer surface and a generally concave inner surface (Figure 1.2).

All non-rigid registration algorithms require a set of landmarks on the reference and the target to guide the shape deformation process, which can be manually marked or automatically detected. The first approach manually marks landmarks on the reference and the target [DZS⁺11, LP00, RB02], and uses the landmarks as hard constraints in non-rigid registration. This approach is accurate, but manually marking a large set of landmarks is too tedious to apply on the entire skull and it needs expert knowledge. The second approach automatically detects surface points on the reference mesh, which are mapped to the target surface. These surface points can be randomly sampled points [HDZ⁺12b] or distinctive feature points such as local curvature maximals [TBK⁺05], and they serve as soft constraints in non-rigid registration. The detected surface points are prone to noise on the surfaces. Using them alone cannot ensure automatically consistent correspondence among all the training samples.

2.1.2 Surface Parameterization

Another approach to building dense correspondence is surface parameterization [Ben91, FH05]. Surface parameterization aims to compute a one-to-one mapping between an input mesh and a simple, uniform shape called the base domain. For example, the 2D base domain for closed contours can be a circle, and the 3D base domain for orientable closed 2D-manifolds of genus zero, i.e., objects without holes and self-intersections, can be a sphere. Harmonic map is used to compute the one-to-one mapping [FH05]. Harmonic map refers to the mapping with the smallest amount of elastic potential energy, and it is guaranteed to be one-to-one for convex regions.

In 2D, for closed contours, determining the correspondences using surface parameterization is equivalent to arc-length resampling [SD05]. Given a starting point and number of vertices, the vertices are uniformly resampled on the contours so that the arc-lengths between neighbouring vertices are the same. In 3D, parameterization of surfaces is far more

complex and usually dependent on the topology of the shape. Most of the applications of surface parameterization are limited to orientable closed 2D-manifolds of genus zero. Kelemen et al. [KSG99] proposes to generate a spherical harmonics mapping between input meshes and the base domain. After aligning all meshes by their first-order ellipsoids, the surface points that map to the same position on the sphere are considered to be corresponding points. Brett et al. [BT00] uses a similar method for 3D topological disks. In their approach, all shapes are mapped to 2D disks and aligned by optimizing the disk rotation to minimize distances between corresponding points determined by a preceding ICP step. While these methods provide a set of automatically determined corresponding points, they can only work on simple mesh surfaces such as hippocampus [KSG99] and ventricle [SD05]. They are not applicable for objects with complex 3D shapes such as skull models, whose surfaces are not genus zero.

2.1.3 Group Representation

Group Representation approach focuses on the properties of the resulting statistical shape model rather than the goodness of correspondence itself. An explicit objective function is employed to evaluate the goodness of the resulting statistical shape model.

Kotcheff et al. [KT98] propose a method that takes the determinant of the training set’s covariance matrix as the objective function. Their method tries to produce a statistical shape model that is as compact and specific as possible. Since the determinant of the covariance matrix concentrates the variance into a few modes with large variances and effectively measures the volume that the training set occupies in shape space, minimising the determinant should lead to a more compact model. In order to minimize the objective function, this method uses a genetic algorithm to modify the positions of the mesh vertices. While this method in some cases create good results, it suffers from a couple of problems. The genetic optimization scheme does not converge in all cases, and cannot cope with complex objects or large data sets. In addition, there is no sound theoretical foundation for the objective function [HM09].

Davies [Dav02] proposes another objective function based on the minimum description length (MDL). MDL is a principle from information theory which states that the best model should describe the training data as efficiently as possible. A MDL objective function is derived that favours models with good specificity, generalisation ability and compactness. Two methods are used to optimize the objective function. The first uses a piecewise linear parameterization scheme to place landmarks on the objects and a genetic algorithm to optimize correspondence. The second method uses a continuous re-parameterization with Cauchy kernels and an optimization using the Nelder-Mead algorithm. The objective function has a sound theoretical foundation. Tests on small-sized training sets of brain ventricles and rat kidneys produce good results. However, this method is computationally very expensive. It takes hours or even days to converge for a small set of simple surfaces. Therefore, it is not suitable for complex shapes such as skulls.

2.2 Skull Reconstruction

Skull reconstruction algorithms seek to estimate the normal shape of a fractured, incomplete skull model. There are currently three reconstruction approaches, namely, symmetry-based reconstruction, geometric reconstruction and statistical reconstruction. In addition, [Che14] has proposed a skull restoration method that recovers the normal shape by repositioning the fractured bone fragments. This is different from skull reconstruction approaches, which does not perform bone repositioning. And it is computationally expensive. So in the following subsections, only existing skull reconstruction approaches will be discussed.

2.2.1 Symmetry-based Reconstruction

Symmetry-based approach produces a reconstructed model based on left-right symmetry of the human skull [GWL99, LCL⁺03, ECP⁺06]. This approach requires the user to indicate the healthy regions of a deformed skull. Then it reflects the healthy part with

respect to the skull symmetry plane to serve as an estimation of the normal shape of the deformed parts. Symmetry-based approach has been applied in Brainlab [GWL99], a leading CMF surgery planning system. Brainlab does not reconstruct the whole skull model. Instead, it just reflects the healthy parts identified by the user, which are then regarded as the reference for actual surgery.

Symmetry-based reconstruction uses the natural approximate left-right symmetry of the human skull. It is simple and efficient to apply. However, it requires the presence of healthy bones to reconstruct the fractured parts on the opposite side of the skull. When both sides of the skull are fractured, which is common in impact injuries, this approach cannot be applied.

2.2.2 Geometric Reconstruction

Geometric reconstruction overcomes the weakness of symmetry-based reconstruction. Geometric reconstruction uses a generic shape model to estimate the normal shape of the deformed part [WYLL11, GMBW04, BSMG09]. It deforms the generic model to match the healthy parts of the deformed model. Then it applies a generic shape function to interpolate the fractured or missing parts to generate the reconstructed surface. For example, [GMBW04] applies thin-plate spline to deform a reference model and register it to the incomplete target model. After registration, the missing parts of the incomplete target model are filled in by the deformed reference model (Figure. 2.1).

Geometric reconstruction is relatively simple to apply. However, detailed shapes of human skulls vary significantly across people with different genders, ages and races. Selection of a reference model in the same gender, age and race is essential for reconstruction accuracy [GMBW04]. Lack of similar reference model will compromise the reconstruction accuracy. In addition, geometric reconstruction uses the correlation between the healthy parts and the deformed parts to perform skull reconstruction. This correlation is weak for severely deformed models. For example, in some severely deformed models, the whole

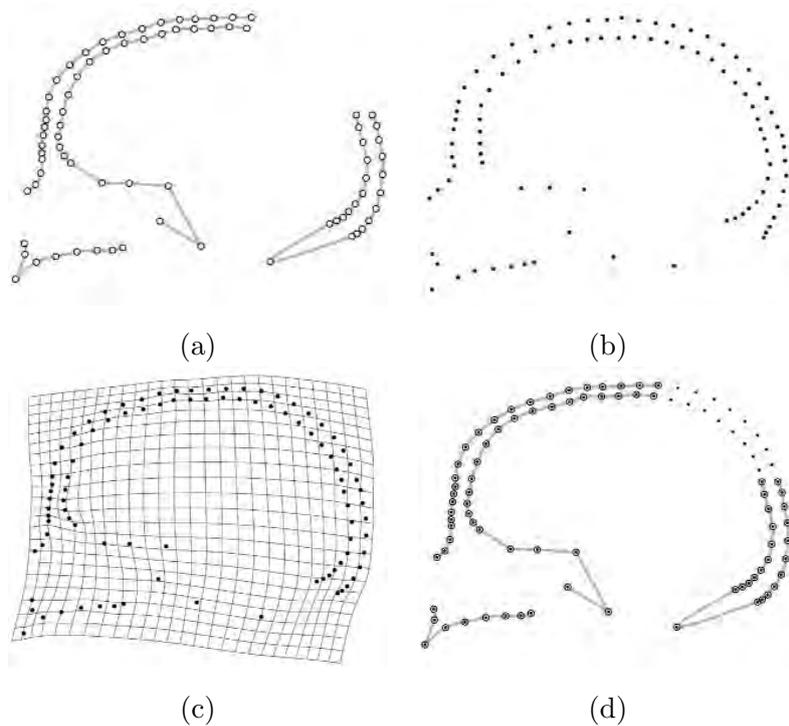


Figure 2.1: Geometric reconstruction [GMBW04]. (a) A model with missing parts. (b) The reference model. (c) The reference model is deformed and registered to (a). (d) Reconstructed model with the missing parts filled in by the registered reference model.

frontal face is deformed, leaving only the back of the skull healthy. The correlation between the back portion and the frontal portion of human skulls is weak. Therefore, in such case, the geometric reconstruction generates a model whose frontal face is close to the reference model instead of the patient's normal model.

2.2.3 Statistical Reconstruction

Statistical reconstruction overcomes the weakness of geometric reconstruction. Instead of using a single reference model, it constructs a statistical model from training samples to capture possible normal variations. Then it matches the statistical model to the normal parts of a deformed model and uses the matched statistical model to infer the fractured and incomplete parts of the deformed model ([Gun05, ZLES05, LAV09]). Given a set of training samples of healthy lower jaw, the method in [ZLES05] applies Principal Compo-

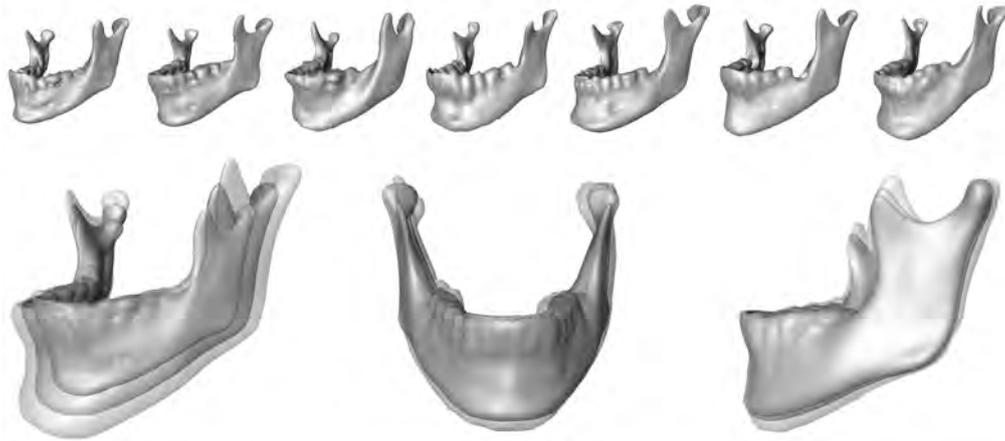


Figure 2.2: Statistical shape model of lower jaw [ZLES05] (Top) Training samples. (Bottom) Three main variation modes of the statistical model.

Principal Component Analysis (PCA) to compute the mean shape and the model matrix of the normal, complete training samples (Figure 2.2). Next, it computes the shape parameters that best match the active shape model (ASM) to the whole deformed model. Finally, it uses the computed shape parameters to generate a reconstructed model from the statistical model.

[Gun05] assumes the correspondence of the skull models are readily prepared and the outliers are already known. [LAV09] use the Thirion’s Demons registration algorithm to build dense correspondence. The Thirion’s Demons registration algorithm runs fast, but artifacts may be introduced as its deformation is unsmooth [HDR⁺13]. Next [LAV09] identifies the outlier vertices on the input skull surface, and the mesh vertices that are identified as outliers are marked as missing. Both of [Gun05] and [LAV09] reconstruct the skulls in an iterative process. An ASM is used to reconstruct the input skull model, and the reconstructed skull model is in turn used to re-train the ASM. Despite the complexity of this method, the re-training process may not improve the ASM. For severely deformed models with large missing parts, the reconstructed parts will always be close to the mean shape, which is not necessarily the same as the ideal reconstructed shape. For deformed models have small missing parts, the reconstruction in the first iteration may be good

enough, and re-training the ASM does not make much difference.

Statistical reconstruction overcomes the limitation of geometric reconstruction by capturing normal variations of human skulls in the statistical model. It can potentially produce a reconstructed model that is close to the normal model of a patient, provided that a good match is obtained between the patient’s deformed model and the statistical model. One potential difficulty is that the construction of the statistical reference model requires a large amount of training samples of normal, complete skull models. The reconstruction accuracy depends on how well the statistical model captures normal shape variation of the patients’ normal models. In the case that such normal shape variation is not adequately captured, then ASM reconstruction is likely to produce the mean shape or an inaccurate model superposition of the training samples. Then, statistical reconstruction degenerates to geometric reconstruction.

2.3 Summary

Based on the review presented in preceding sections, it can be concluded that thin-plate spline (TPS) is the most promising method for building dense correspondence. It ensures smooth deformation and has few parameters that need manual tuning, and it is efficient. In addition, TPS has a sound physical explanation of its energy function. Compared to surface parameterization and group representation, TPS is more effective for complex 3D shapes like skull surfaces.

TPS requires a set of landmarks to guide the shape deformation and registration process, which can be manually marked or automatically detected. The first approach uses manually marked landmarks as hard constraints [DZS⁺11, LP00, RB02]. This approach is accurate, but manually marking a large set of landmarks is too tedious to apply on the entire skull and it needs expert knowledge. The second approach uses automatically detected surface points as soft constraints [HDZ⁺12b, TBK⁺05]. The detected surface points are prone to noise on the surfaces. Using them alone cannot ensure automati-

Table 2.1: Comparison of skull reconstruction algorithms.

Method	Anatomy	Correspondence	Reconstruction	ASM Fitting
[ZLES05]	Lower Jaw	Surface Parameterization Applicable only to simple shapes	ASM	Fit to whole model Deformed parts introduce error
[Gun05]	Skull	Assumed known	ASM with retraining	Fit to normal parts
[LAV09]	Skull	Thirion’s Demons registration Unsmooth deformation introduces artifacts	ASM with retraining	Fit to normal parts
Proposed	Skull	TPS with hard and soft constraints	ASM	Fit to normal parts

cally consistent correspondence among all the training samples. This thesis overcomes these weaknesses by applying TPS in a new way that combines both hard and soft constraints, which ensures both anatomically consistent correspondence and close matching of reference and target surfaces.

Among the existing skull reconstruction approaches, statistical reconstruction is the most promising approach that overcomes the weaknesses of symmetry-based and geometric reconstruction. Among the various statistical reconstruction methods, active shape model (ASM) is the most widely used [ZLES05, Gun05, LAV09]. In particular, the method of [ZLES05] (Table 2.1) works on lower jaws. It uses surface parameterization to build dense correspondence. With the established correspondence, an ASM is constructed and fitted to the whole deformed lower jaw model. This method is restricted to simple shapes, because surface parameterization is applicable only to simple shapes. Moreover, fitting the ASM to the whole model can introduce error because the deformed part of the model contains noise and incorrect shapes. The method of [Gun05] (Table 2.1) assumes that correspondence is already known. In practice, an appropriate automatic algorithm needs to be used. [Gun05] fits an ASM to the normal part of the deformed skull model, which

is in turn used to re-train the ASM. Despite its complexity, the re-training process may not improve ASM because the re-training process does not use any new knowledge of the objects' shapes. The method of [LAV09] uses the Thirion's Demons registration algorithm to build dense correspondence. It also uses ASM with re-training to reconstruct skull models. The deformation of Thirion's Demons algorithms is unsmooth [HDR⁺13]. Therefore, artifacts may be introduced during registration. In contrast, the method proposed in this thesis applies TPS with both hard and soft constraints as the non-rigid registration method that ensures anatomical consistency and close matching of the reference and target models. It is more smooth compared to that of [LAV09] and is applicable to complex skull shape unlike that of [ZLES05]. It fits an ASM to the normal part of a deformed skull to reconstruct the skull. In summary, it is expected to perform more robustly than existing methods, especially for severely fractured and incomplete skulls.

Chapter 3

Dense Correspondence with Thin-Plate Spline

Correct correspondence among input samples is crucial for building Active Shape Model of 3D skulls. This chapter presents a new dense correspondence algorithm for both normal, complete skulls and fractured, incomplete skulls. The part for normal, complete skulls has been published in [ZCL13]. This chapter discusses an extended version of [ZCL13].

The proposed algorithm performs non-rigid registration of the reference model to a target model using Thin-Plate Spline algorithm (Section 3.2). It uses landmarks as hard constraints to ensure **anatomically consistent correspondence**, and samples control points on skull surfaces to serve as soft constraints, which provide additional local shape constraints for **close matching** of reference and target surfaces (Section 3.3). For normal, complete skulls, it can automatically detect landmarks (Section 3.1). On the other hand, for fractured, incomplete skulls, it requires manually marked landmarks, because the automatic algorithm is not accurately enough for severely fractured, incomplete skulls.

Quantitative evaluation of point correspondence is also a challenging task. The quantitative errors measured in [BDGY04, HBH03, WPS00] are non-rigid registration error instead of point correspondence error, because it is very difficult to measure actual point

correspondence error. Section 3.4 presents a method of estimating point correspondence error, and uses both correspondence error and registration error to evaluate the performance of various methods.

3.1 Automatic Craniometric Landmark Detection

In anatomy [SH] and forensics [Tay01], craniometric landmarks are feature points on a skull that are used to define and measure skull shapes. Automatic detection of craniometric landmarks is very difficult and challenging due to a form of cyclic definition. Many craniometric landmarks are defined according to three *anatomical orientations* of the skull (Fig. 3.1(a)): lateral (left-right), anterior-posterior (front-back), and superior-inferior (up-down). These orientations are defined by the *Frankfurt plane* (FP) and the *mid-sagittal plane* (MSP), which are in turn defined as the planes that pass through specific landmarks.

To overcome this cyclic definition, this section presents an automatic landmark detection algorithm based on [CLL12] that registers a reference skull model with known landmarks to a target skull model to automatically locate the landmarks on the target model. It then fits two planes to the landmarks to obtain good initial estimates of FP and MSP. Then, it iteratively refines the locations of the landmarks and the locations and orientations of FP and MSP. This algorithm is published in [CLL12] and it can be summarized as follows:

Craniometric Landmark Detection Algorithm

1. Register a reference model with known landmarks to the target model.
2. Locate the landmarks on the target based on the registered reference and fit FP and MSP to their landmarks on the target.
3. Repeat until convergence:

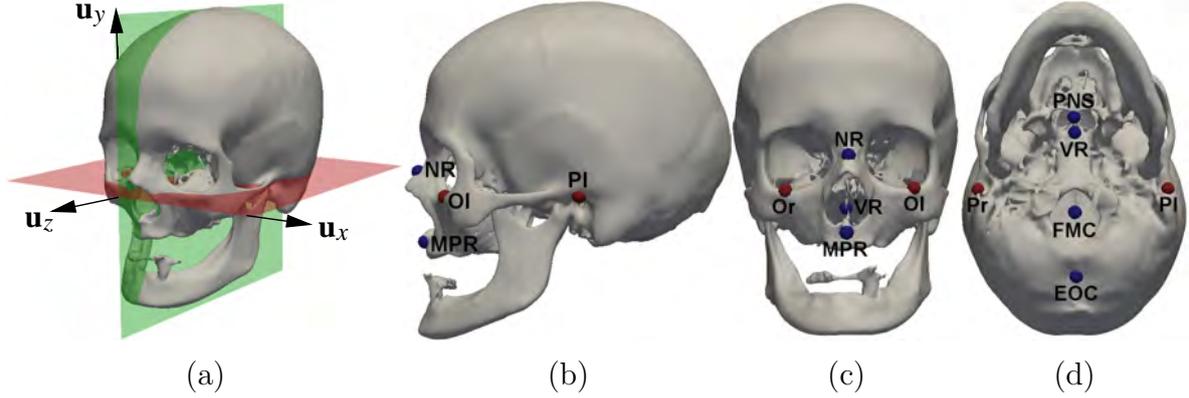


Figure 3.1: Reference skull model. (a) Frankfurt plane (FP) is the horizontal (red) plane and mid-sagittal plane (MSP) is the vertical (green) plane. (b, c, d) Red landmarks define FP and blue landmarks define MSP.

- (a) Refine the locations of the FP landmarks on the target, and fit FP to the refined FP landmarks.
- (b) Refine the locations of the MSP landmarks on the target, and fit MSP to the refined MSP landmarks, keeping it orthogonal to FP.

Step 1 registers the reference to the target using Fractional Iterative Closest Point (FICP) [PLT07], a variant of ICP robust to noise, outliers, and missing bones. Like ICP, FICP iteratively computes the best similarity transformation (scaling, rotation, and translation) that registers the reference to the target. The difference is that in each iteration, FICP computes the transformation using only a subset of reference points whose distances to the target model are the smallest. FICP determines the subset by minimizing the following error:

$$\frac{1}{f^\lambda} \sqrt{\frac{1}{|R_f|} \sum_{p \in R_f} \|T(p) - \mu(p)\|^2} \quad (3.1)$$

where R is the reference model, M is the target model, T is the transformation, $f \in (0, 1)$ is the fraction of the subset of reference points, $\mu : R \rightarrow M$ matches each point of R to the closest point of M . FICP compute the subset D_f and the similarity transformation T that aligns D_f to M .

After registration, Step 2 maps the known landmarks on the reference to the target. These known landmarks include the FP and MSP landmarks as well as other landmarks that help define the skull's shape (Figure 3.2(1)). First, closest points on the target surface to the reference landmarks are identified. These closest points are the initial estimates of the landmarks on the target, which may not be accurate due to shape variations of the target skull. Next, FP and MSP are fitted to the initial estimates using PCA. The mean of the landmark points gives the position of the plane, and the smallest eigenvector obtained by PCA gives the unit normal vector of the plane. So, the FP and MSP are each represented by its position and unit normal vector.

In Step 3, an elliptical landmark region R is identified around each initial estimate. The orientation and size of R are empirically predefined. R varies for different landmarks according to the shapes of the skull around the landmarks. These regions should be large enough to include the landmarks on the target model. Accurate landmark locations are searched within the regions according to their anatomical definitions. For example, the left and right porions (Pl, Pr in Fig. 3.1) are the most *lateral* points of the roofs of the ear canals [SH, Tay01]. After refining FP landmarks in Step 3(a), FP is fitted to the refined FP landmarks. Next, MSP landmarks are refined in Step 3(b) in a similar manner, and MSP is fitted to the refined MSP landmarks. Since MSP is orthogonal to FP, it projects to a line l on FP. The perpendicular distance of a point \mathbf{p} to MSP is equal to the distance of the projection of \mathbf{p} on FP to the line l . So, the line l is obtained by fitting it to the projections of MSP landmark points on FP. Then, the midpoint of l gives the location of MSP, and MSP's normal vector \mathbf{u}_x is obtained as the cross product of FP's normal vector \mathbf{u}_y and l 's unit direction vector \mathbf{u}_z : $\mathbf{u}_x = \mathbf{u}_y \times \mathbf{u}_z$.

As Step 3 is iterated, the locations and orientations of FP and MSP are refined by fitting to their respective landmarks, and the landmarks' locations are refined according to the refined FP and MSP. After the algorithm converges, accurate craniometric landmarks are detected on the target model (Figure 3.2(2)).

In addition to the landmarks on FP and MSP, other landmarks are also detected.

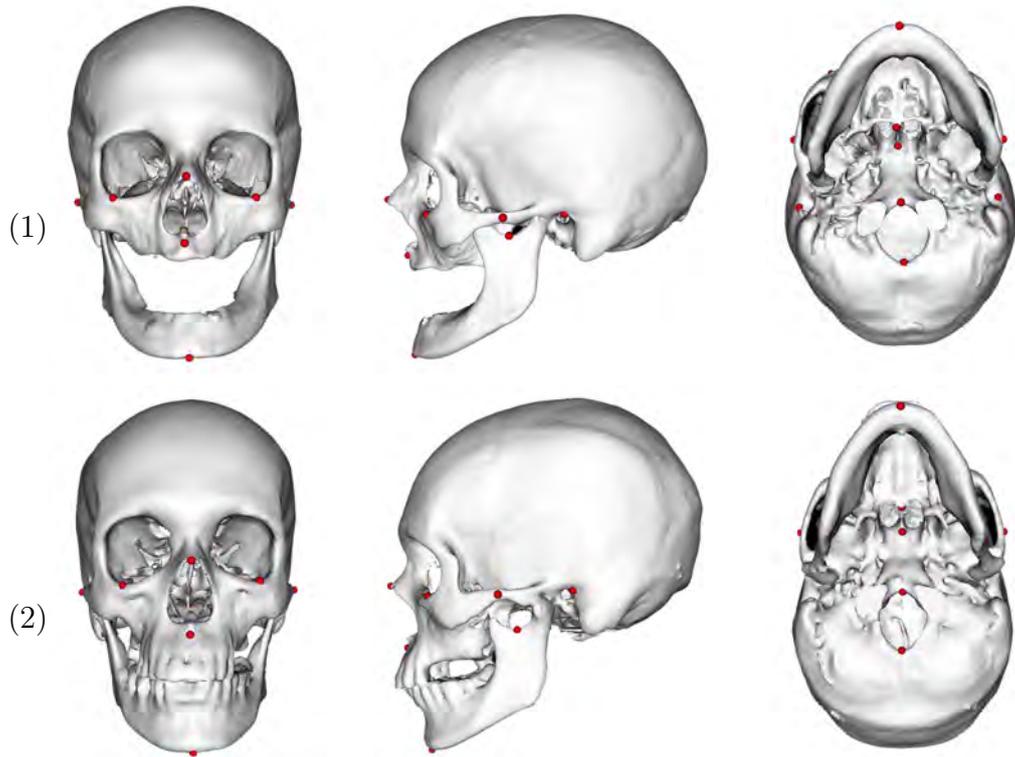


Figure 3.2: Landmarks used by correspondence algorithm. (1) Manually marked landmarks on the reference model. (2) Automatically detected landmarks on the target model. Landmarks in (2) are automatically located by registering the reference skull model with known landmarks (in (1)) to the target skull model.

These include points of extremum along the anatomical orientations defined by FP and MSP. These landmarks are detected in a similar manner as the FP and MSP landmarks, first by mapping known landmark regions on the reference to the target, and then searching within the regions for the landmarks according to their anatomical definitions.

3.2 Non-Rigid Registration with Thin-Plate Spline

Non-rigid registration is the predominant approach for computing the correspondence between two 3D meshes because it has the potential of closely registering one mesh to the other. It is typically preceded by rigid registration to first align the sizes, positions, and orientations of the meshes. Among the various non-rigid registration methods discussed in Section 2.1, Thin-Plate Spline (TPS) method is particularly effective for registering mesh models with highly complex surfaces such as brain sulci [CR00], lumbar vertebrae [LK00], and skulls [DZS⁺11, HDZ⁺12b, LP00, RB02, TBK⁺05].

TPS method aims to determine a mapping function f between an input surface and a deformed surface by minimizing an energy function E_B that models the deformation as the bending of a thin sheet of metal [Wik]. Let \mathbf{p}_i and \mathbf{p}'_i , $i = 1, \dots, n$, denote, respectively, the initial coordinates and deformed coordinates of points on the surface. Then, TPS deformation with hard constraints looks for the mapping function $f(\mathbf{p}_i) = \mathbf{p}'_i$ that minimizes the 3-dimensional second-order bending energy E_B [Wah90],

$$E_B = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (f_{xx}^2 + f_{yy}^2 + f_{zz}^2 + 2[f_{xy}^2 + f_{xz}^2 + f_{yz}^2]) dx dy dz, \quad (3.2)$$

where $f(\mathbf{p}_i) = \mathbf{p}'_i$, for $i = 1, \dots, n$, serve as the hard constraints, and $f_{xx}, f_{yy}, f_{zz}, f_{xy}, f_{xz}, f_{yz}$ are second derivative of f with respect to x, y, z .

The hard constraints are relaxed to obtain TPS with soft constraints which minimizes

the energy functional $E(f)$ that combines positional differences and the bending energy:

$$E(f) = \sum_{i=1}^n \left\| \mathbf{p}'_i - f(\mathbf{p}_i) \right\|^2 + \lambda E_B. \quad (3.3)$$

The positive regularization parameter λ controls the trade-off between minimization of positional differences and the bending energy.

TPS with hard constraints is not robust to the noise and outliers in the constraints, as it requires each initial coordinate to be mapped exactly to its deformed coordinate. On the other hand, TPS with soft constraints weakens the constraints. TPS with hard combine soft constraints overcomes their weaknesses. The constraints which are accurate are used as hard constraints so that they are fully used, and the constraints which may not be so accurate are used as soft constraints, so the possible noise of constraints cannot deteriorate the results too much. Let \mathbf{p}_i and \mathbf{p}'_i , $i = 1, \dots, m$, denote corresponding points that serve as hard constraints and \mathbf{p}_i and \mathbf{p}'_i , $i = m + 1, \dots, n$, denote those that serve as soft constraints. Then, TPS deformation with both hard and soft constraints aims to minimize the energy functional $E(f)$:

$$E(f) = \sum_{i=m+1}^n \left\| \mathbf{p}'_i - f(\mathbf{p}_i) \right\|^2 + \lambda E_B, \quad (3.4)$$

subject to the hard constraints $f(\mathbf{p}_i) = \mathbf{p}'_i$, for $i = 1, \dots, m$.

A solution f that minimizes Equation 3.4 is given by the equation:

$$f(\mathbf{p}_i) = a_0 + a_1 x_i + a_2 y_i + a_3 z_i + \sum_{j=1}^n w_j U(\|\mathbf{p}_j - \mathbf{p}_i\|), \quad (3.5)$$

where $U(r) = r^2 \log r$. Writing the coefficients $\mathbf{w} = (w_1, \dots, w_n)^\top$, $\mathbf{a} = (a_0, a_1, a_2, a_3)^\top$, Equation 3.5 is rearranged into a linear system [RSS⁺01]:

$$\begin{bmatrix} \mathbf{K} + \lambda \mathbf{\Sigma} & \mathbf{P} \\ \mathbf{P}^\top & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{w} \\ \mathbf{a} \end{bmatrix} = \begin{bmatrix} \mathbf{P}' \\ \mathbf{0} \end{bmatrix}. \quad (3.6)$$

\mathbf{P} and \mathbf{P}' are $n \times 4$ matrices such that their i -th rows are, respectively, $(1, \mathbf{p}_i^\top)$ and $(1, \mathbf{p}'_i^\top)$, $\mathbf{K}_{ij} = U(\|\mathbf{p}_i - \mathbf{p}'_j\|)$, $\mathbf{\Sigma} = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_n)$, and

$$\sigma_i = \begin{cases} 0 & \text{if } i \leq m, \\ 1 & \text{if } i \geq m + 1. \end{cases} \quad (3.7)$$

The first row of Equation 3.6 is

$$\mathbf{P}\mathbf{a} + (\mathbf{K} + \lambda\mathbf{\Sigma})\mathbf{w} = \mathbf{P}' \quad (3.8)$$

For $\lambda = 0$, the transformation is an interpolating spline, which means all constraints behave as hard constraints. For $\lambda > 0$, the transformation is an approximating spline. If λ is small, the solution can adapt to local shape structure, which is equivalent to having both hard constraints (for $\sigma_i = 0$) and soft constraints (for $\sigma_i > 0$). If λ is large, the solution is a smooth transformation and the constraints are soft. For $\lambda \rightarrow \infty$, the solution is an affine transformation. The second row $\mathbf{P}^\top \mathbf{w} = \mathbf{0}$ ensures that the non-rigid transformation is localized, i.e., is $\mathbf{0}$ at infinity.

3.3 Correspondence and Resampling

The dense correspondence algorithm adopts multi-stage coarse-to-fine approach. It uses TPS with combines hard and soft constraints to register a reference mesh to a target mesh. It consists of the following stages:

1. Apply FICP to register the reference model to the target model.
2. Identify landmarks on the target by either manual marking or applying the automatic landmark detection algorithm.
3. Apply TPS to register the reference to the target with craniometric landmarks as hard constraints.

4. Sample control points on the surfaces of the reference mesh and map them to the surfaces of the target mesh.
5. Apply TPS with craniometric landmarks as hard constraints and control points as soft constraints.
6. Resample target model by mapping reference mesh vertices and vertex connectivity to the target.

Stage 1 registers the reference model to the target model using FICP. FICP does a rigid registration to align scale, position and orientation of the two models. Like ICP, FICP iteratively computes the best similarity transformation (scaling, rotation, and translation) that registers the reference model to the target model. The difference is that in each iteration, FICP computes the transformation using only a subset of reference points whose distances to the target model are the smallest.

Stage 2 identifies landmarks on the target model. The landmarks are automatically detected using Craniometric Landmark Detection Algorithm (Section 3.1) or manually marked. For normal, complete skull models, automatically detected landmarks are used. As the automatic method is not robust for fractured, incomplete skull models, manually marked landmarks are used for fractured, incomplete skull models.

Stage 3 applies TPS to perform coarse registration with the landmarks as hard constraints, which ensures **anatomically consistent correspondence**. Based on the coarse registration result, the control points are sampled more accurately in the next step.

Stage 4 randomly selects m reference mesh vertices with the largest registration errors as the control points. For each control point, a nearest point on the target surface within a fixed distance and with a sufficiently similar surface normal is selected as the corresponding point (Figure 3.3(a)). In the current implementation, the fixed distance is set to 10 mm and the surface normals are similar enough if the cosine of the angle between them is larger than 0.86. If a corresponding point that satisfies these criteria cannot be found, then the control point is discarded. In this way, the algorithm is robust to missing

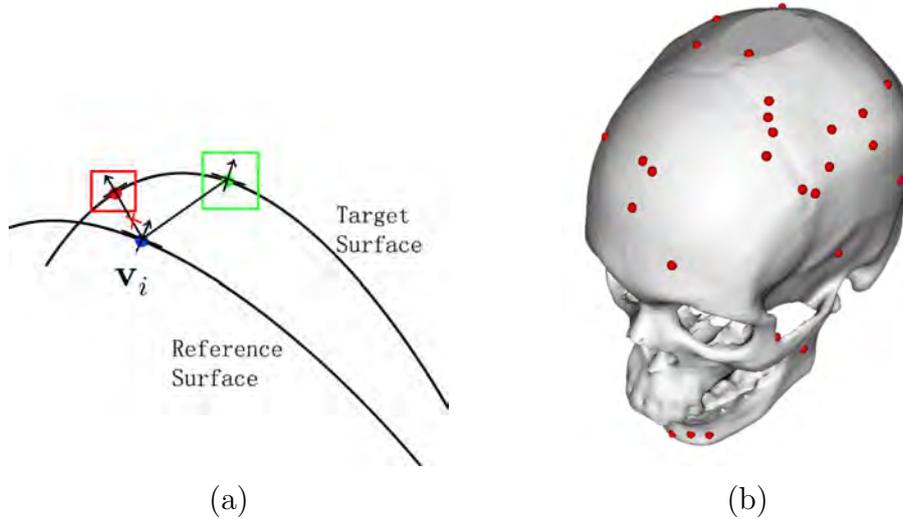


Figure 3.3: Control points sampling and correspondence searching. (a) Correspondence point searching. Although the red point is the closest point to the reference vertex v_i , it is not selected as its surface normal is quite different from that of the reference vertex v_i . On the other hand, the green point is selected as it is close to the reference vertex v_i and has similar surface normal. (b) Control points sampling. There are very few craniometric landmarks on the top of the skull, where large registration error occurs. Most control points (red dots) are sampled in this area.

parts in the target skulls. Figure 3.3(b) shows some sampled control points.

Stage 5 performs another TPS registration with craniometric landmarks as hard constraints and control points as soft constraints. These constraints ensure **close matching** of reference and target surfaces while maintaining anatomically consistent correspondence (Section 3.2).

After TPS registration, stage 6 maps the reference mesh vertices to the target surface in the same manner as mapping of control points in Stage 3. For each vertex on the reference mesh, its nearest point on the target mesh within a fixed distance and with a sufficiently similar surface normal is selected as the corresponding point (Figure 3.3(a)). In the current implementation, the fixed distance is set to 10 mm and the surface normals are similar enough if the cosine of the angle between them is larger than 0.86. For a normal, complete target mesh, if a corresponding point that satisfies these criteria cannot be found,

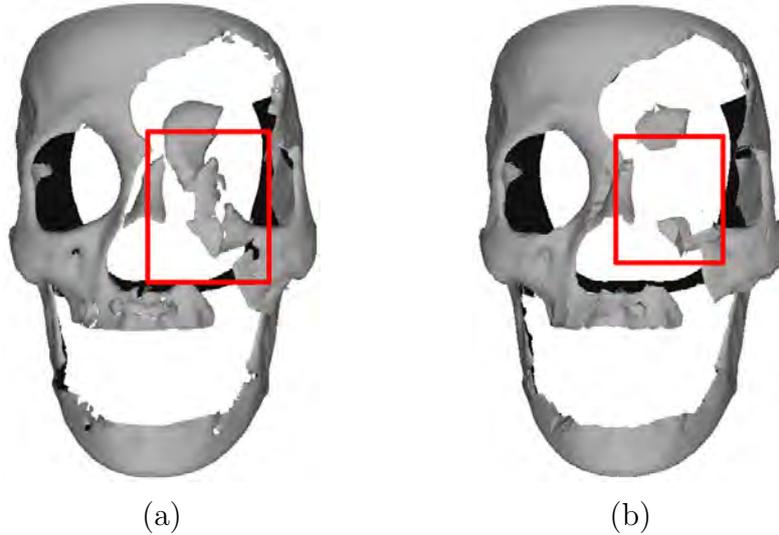


Figure 3.4: Resampling of a fractured, incomplete skull. (a) The fractured bones highlighted by the red boxes have no correspondence. (b) So, these fractured bones are omitted in the resampled model.

then its nearest point is used as corresponding point. This is because there should be no missing correspondence in normal, complete skulls. On the other hand, for a fractured, incomplete target mesh, if a corresponding point that satisfies these criteria cannot be found, then its corresponding point is regarded as missing (Figure 3.4). Fractured bones are typically displaced and rotated due to tension of the muscles attached to them. In this displaced and rotated state, their shapes with respect to the whole skull are incorrect. Therefore, they should be omitted in the resampled model. This method provides some robustness to the correspondence algorithm.

3.4 Error Measures

There are two ways to measure error of a correspondence algorithm, namely, registration error and correspondence error. Registration error E_R measures the difference between the registered reference surface and the target surface. It is computed as the mean distance

between the reference mesh vertices \mathbf{v}_i and the nearest surface points \mathbf{p}_i on the target:

$$E_R = \left[\frac{1}{n} \sum_{i=1}^n \|\mathbf{v}_i - \mathbf{p}_i\|^2 \right]^{1/2} \quad (3.9)$$

where n is the number of vertices. This is essentially the error measured in [BDGY04, HBH03, WPS00], although the actual formulations that they used differ slightly.

Correspondence error, on the other hand, should measure the error in computing point correspondence. One possible formulation of correspondence error is to measure the mean distance between the desired and actual corresponding target points of reference mesh vertices. The desired corresponding point $D(\mathbf{v}_i)$ is the ground-truth marked by a human expert, whereas the actual corresponding point $C(\mathbf{v}_i)$ is the one computed by dense correspondence algorithm. With this formulation, the correspondence error E_C can be computed as

$$E_C = \left[\frac{1}{n} \sum_{i=1}^n \|D(\mathbf{v}_i) - C(\mathbf{v}_i)\|^2 \right]^{1/2} \quad (3.10)$$

In practice, it is impossible to manually mark the desired corresponding points of reference mesh vertices accurately on the target mesh surface. An alternative formulation is to measure the mean distance between the desired and actual corresponding target landmarks of reference landmarks \mathbf{M}_i :

$$E_C = \left[\frac{1}{l} \sum_{i=1}^l \|D(\mathbf{M}_i) - C(\mathbf{M}_i)\|^2 \right]^{1/2} \quad (3.11)$$

where l is the number of landmarks used for evaluation. The desired target landmarks are manually marked whereas the actual target landmarks are computed by the dense correspondence algorithm. Given enough landmarks adequately distributed over the entire reference surface, Equation 3.11 is a good approximation of Equation 3.10.

3.5 Experiments and Discussions

11 normal, complete skull models reconstructed from CT images were used in the tests. One of them served as the reference model and the others were target models. For performance comparison, the following methods were tested for computing dense correspondence:

1. ICP: ICP rigid registration with mesh vertices as corresponding points.
2. FICP: FICP rigid registration with mesh vertices as corresponding points.
3. CP-S: TPS registration with automatically detected control points as soft constraints. This approach was adopted by [HDZ⁺12b].
4. LM-H: TPS registration with automatically detected craniometric landmarks as hard constraints.
5. LM-S/CP-S: TPS registration with automatically detected craniometric landmarks and control points as soft constraints. This approach is similar to the method of [TBK⁺05], except [TBK⁺05] adopted a more elaborate multi-stage, coarse-to-fine, and forward-backward registration scheme.
6. **LM-H/CP-S**: TPS registration with automatically detected craniometric landmarks as hard constraints and control points as soft constraints. This is the proposed algorithm described in Section 3.3.
7. MLM-H: TPS registration with manually marked craniometric landmarks as hard constraints. This approach was adopted by [DZS⁺11, LP00, RB02].

These test cases were equivalent to the proposed algorithm (Case 6) with different stages and constraints omitted. In these test cases, the TPS algorithm was implemented by modifying the VTK source code. All the TPS registrations were preceded by FICP rigid registration. TPS regularization parameter λ of Equation 3.4 was set to 0.8 where the

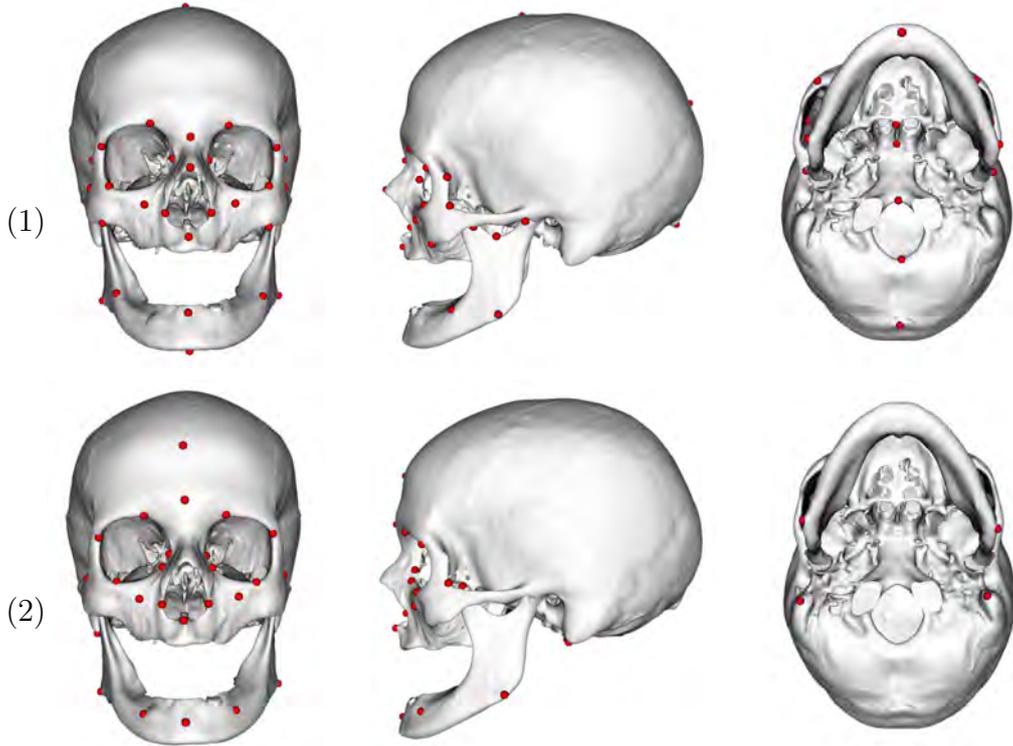


Figure 3.5: Manually marked landmarks. (1) Manually marked landmarks on target models for registration, (2) Manually marked landmarks on the target models for evaluation of the algorithm’s accuracy.

algorithms generally performed well. 15 landmarks (Figure 3.2) and 150 control points were used for registration for Cases 3–6, and 30 landmarks (Figure 3.5(1)) for Case 7. More landmarks could be used for Case 7 because they included landmarks that could be accurately marked manually but not detected automatically. 28 additional landmarks were used for evaluation (Figure 3.5(2)). Both registration error and correspondence error were measured.

Test results (Table 3.1) show that FICP performs more accurately than does ICP because it is more robust than ICP in rigid registration. The registration error of CP-S is smaller than those of LM-S/CP-S and LM-H/CP-S, but its correspondence error is larger. This shows that low registration error does not necessarily imply low correspondence error.

Table 3.1: Quantitative evaluation. E_R : registration error. E_{CR} , E_{CE} : correspondence errors for registration landmarks and evaluation landmarks, respectively. Errors are measured in mm.

Algorithm	E_R	E_{CR}	E_{CE}
ICP	2.22	7.09	7.42
FICP	1.97	5.55	6.35
CP-S	1.64	4.15	5.81
LM-H	2.69	3.51	5.94
LM-S/CP-S	1.76	3.68	5.73
LM-H/CP-S	1.76	3.58	5.56
MLM-H	2.42	0.00	4.66

CP-S and LM-S/CP-S use only soft constraints, which are inadequate for ensuring anatomically consistent correspondence. So, their correspondence errors are larger than those of LM-H/CP-S, which also uses registration landmarks as hard constraints. On the other hand, LM-H uses only landmarks, which are insufficient for ensuring close matching of reference and target surfaces, though consistent correspondence is somewhat achieved. So, its correspondence error for registration landmarks E_{CR} is very small, but its correspondence error for evaluation landmarks E_{CE} is large. LM-S/CP-S uses landmarks as soft constraints, which weakens the anatomical consistency of correspondence, though close matching of reference and target surfaces is achieved. Using landmarks as hard constraints, the proposed algorithm LM-H/CP-S ensures strong anatomically consistent correspondence. Together with control points as soft constraints, it achieves very low registration error and the lowest correspondence error for evaluation landmarks E_{CE} among the automatic methods (Cases 1–6).

MLM-H uses manually marked landmarks as hard constraints. So, it is not surprising that it has the smallest correspondence errors. Interestingly, its registration error is quite large compared to those of the other methods. This is because some parts of the skulls lack distinctive surface features for locating both registration and evaluation landmarks (Fig. 3.5), where most of the registration errors occur.

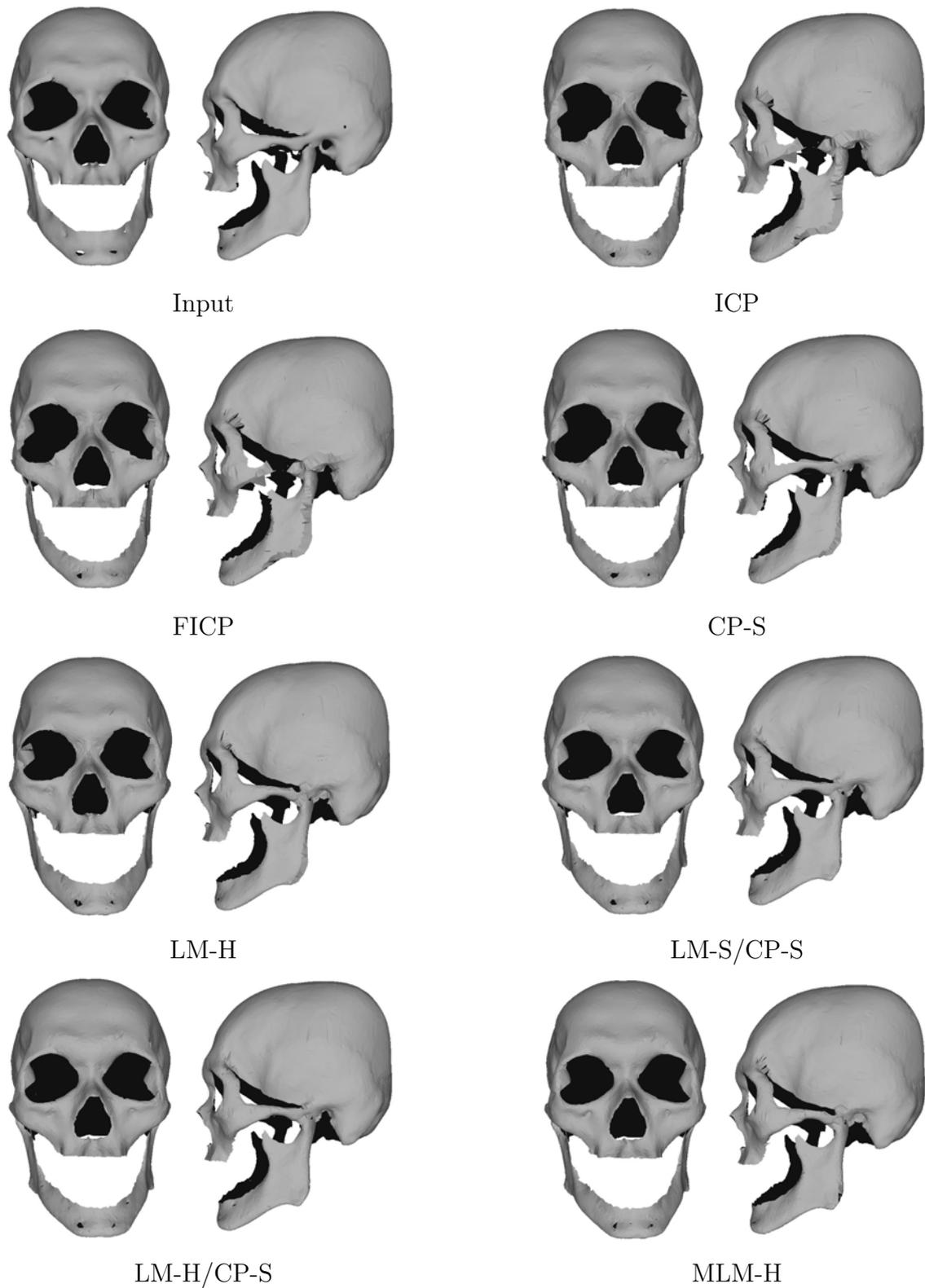


Figure 3.6: Sample resampling results. LM-H/CP-S and LM-S/CP-S have the best qualitative results. The resampling results of MLM-H is better than that of CP-S.

For normal, complete skulls, with small correspondence and registration error, LM-H/CP-S has the best resampling quality (Figure 3.6). The result of LM-S/CP-S is very close to that of LM-H/CP-S. Interestingly, the resampling result of MLM-H is better than that of CP-S (Figure 3.6), even though CP-S has smaller registration error. This shows that correspondence error is more important than registration error in measuring the accuracy of correspondence algorithms for normal, complete samples. More comprehensive evaluation of resampling accuracy will be presented in Section 5.3.

To investigate the stability of LM-H/CP-S algorithm, it is tested with varying numbers of control points and two different sampling schemes that are used by existing methods: low curvature [TBK⁺05] and large registration error [DZS⁺11]. Figure 3.7 shows that control points with large registration errors are more effective than those with low curvatures in reducing correspondence error. Compared to the accuracy of LM-H, which uses landmarks only, a small number of control points can already improve correspondence accuracy significantly. After sampling enough control points that cover various parts of the skulls, adding more control points do not reduce correspondence error significantly. This is due to the diminished quality of the additional control points.

3.6 Summary

This chapter presents a multi-stage, coarse-to-fine dense correspondence algorithm for mesh models of skulls that combines two key features. First, anatomical landmarks serve as hard constraints for TPS registration. They ensure **anatomically consistent correspondence**. Second, control points are sampled on the skull surfaces to serve as soft constraints for TPS registration. They provide additional local shape constraints to ensure **close matching** of reference and target surfaces. Test results show that, by combining both hard and soft constraints, the proposed algorithm can achieve more accurate and robust dense correspondence than other algorithms, for both normal, complete skull models and fractured, incomplete skull models. Nevertheless, with enough control points, adding

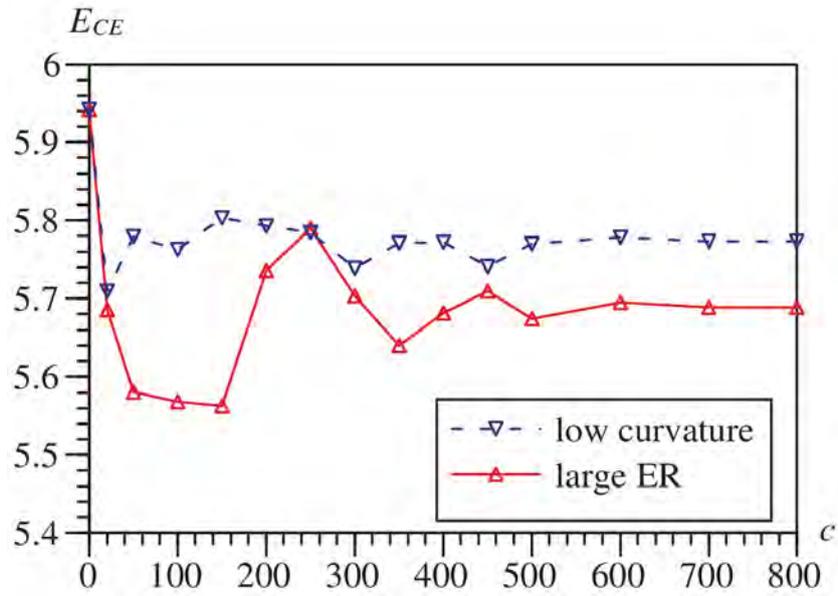


Figure 3.7: Effect of control points on correspondence accuracy. Increasing the number c of control points significantly reduces correspondence error E_{CE} . With enough control points, adding more control points do not reduce correspondence error significantly.

more control points do not reduce correspondence error significantly. Test results also show that low registration error does not always imply low correspondence error. So, both error measures should be used together to evaluate the accuracy of dense correspondence algorithms.

Chapter 4

Active Shape Skull Model

Active shape model (ASM) [CTC⁺95] is by far the most popular statistical shape model, which has been widely used to analyse medical images. Model building and fitting with ASM are quite efficient and robust [HM09]. This chapter first describes ASM in general, followed by application of ASM to model the shape of normal, complete skulls (Section 4.1). Next, Section 4.2 presents an experiment that constructs an ASM of normal, complete skulls. Application of ASM to the reconstruction of fractured, incomplete skulls will be discussed in Chapter 5.

4.1 Active Shape Skull Model

Active shape model works with n spatially-aligned shape vectors \mathbf{s}_i of the training samples, such that

$$\mathbf{s}_i = \{x_{i1}, y_{i1}, z_{i1}, x_{i2}, y_{i2}, z_{i2}, \dots, x_{im}, y_{im}, z_{im}\}^T \quad (4.1)$$

where m is the number of feature points, and (x_{ij}, y_{ij}, z_{ij}) is the spatial coordinate of the j -th feature point. The training samples should be spatially aligned so that variations due to position, scale and rotation are removed and only geometric shape variations are captured in the model. Generalized Procrustes Analysis [Gow75, Goo91] can be used to

align the training samples.

After alignment, the next step is to perform Principal Component Analysis (PCA) [Jol02] to identify the directions of major variations among the training samples. PCA is performed as follows: The mean shape of the training samples is computed as

$$\bar{\mathbf{s}} = \frac{1}{n} \sum_{i=1}^n \mathbf{s}_i, \quad (4.2)$$

and the covariance matrix of the training samples is computed as

$$\mathbf{\Sigma} = \frac{1}{n-1} \sum_{i=1}^n (\mathbf{s}_i - \bar{\mathbf{s}})(\mathbf{s}_i - \bar{\mathbf{s}})^\top. \quad (4.3)$$

Next, eigendecomposition is performed on the covariance matrix $\mathbf{\Sigma}$, which computes the eigenvectors ϕ_i and their corresponding eigenvalues λ_i in decreasing order. The eigenvectors ϕ_i represent the directions of major variations and each eigenvalue λ_i captures the amount of variation along the direction of its corresponding eigenvector ϕ_i .

Let $\mathbf{\Phi}$ denote the model matrix whose columns are the eigenvectors ϕ_i . Then, any shape vector \mathbf{s} can be generated by $\mathbf{\Phi}$ as follows:

$$\mathbf{s} = \bar{\mathbf{s}} + \mathbf{\Phi}\mathbf{b}, \quad (4.4)$$

where \mathbf{b} is the shape parameter vector. When \mathbf{b} is $\mathbf{0}$, \mathbf{s} is just the mean shape $\bar{\mathbf{s}}$. When the first component of \mathbf{b} is varied from $-3\sqrt{\lambda_1}$ to $+3\sqrt{\lambda_1}$, the first mode of variation of the model shape is produced (Figure 4.3). Variation of mode j can be produced similarly.

When the number n of training samples is smaller or equal to the number of dimensions $3m$ of the training samples, the model matrix $\mathbf{\Phi}$ has $n-1$ eigenvectors because n vectors span a space of at most $n-1$ dimensions. When $n > 3m$, the number of eigenvectors is at most $3m$. In the case of skull modeling, n is typically much smaller than $3m$ (Section 4.2). In practice, when n is sufficiently large, a smaller number of the first k eigenvectors is sufficient to capture most of the variations of the training samples. The

number k can be selected by determining the number of eigenvectors needed for a desired residual or ratio of unaccounted variance R :

$$R = \frac{\sum_{i=k+1}^{n-1} \lambda_i}{\sum_{i=1}^{n-1} \lambda_i}. \quad (4.5)$$

3D mesh models such as skulls are very complex and each of the training samples can have a different number of vertices and different vertex connectivity. Before apply ASM, the training samples must be resampled to obtain meshes of the same number of vertices and connectivity. To achieve this goal, the dense correspondence algorithm described in Section 3.3 is applied. In this case, manually marked landmarks are used because the automatic algorithm (Section 3.1) detects too few landmarks for accurate resampling. After resampling, Procrustes analysis is applied to spatially align the resampled training samples. Next, shape vectors of the form described in Equation 4.1 are generated from the aligned training samples. Finally, ASM model building algorithm is applied to the aligned samples to generate the model matrix Φ .

4.2 Experiments and Discussions

4.2.1 Data Preparation and Procedures

In active shape model construction, 34 skull models were segmented and reconstructed from CT volume images. The outer surfaces of the skull models were extracted. One of them was used as the reference skull model (Figure 4.1) as well as a training sample. As the ASM will be used for fractured, incomplete skulls, even though the training samples were normal and complete and there were no missing correspondences, to make it consistent with the testing samples, manually marked landmarks were used. 30 landmarks were marked on each of the training samples (Figure 3.5(1)). Then the dense correspondence

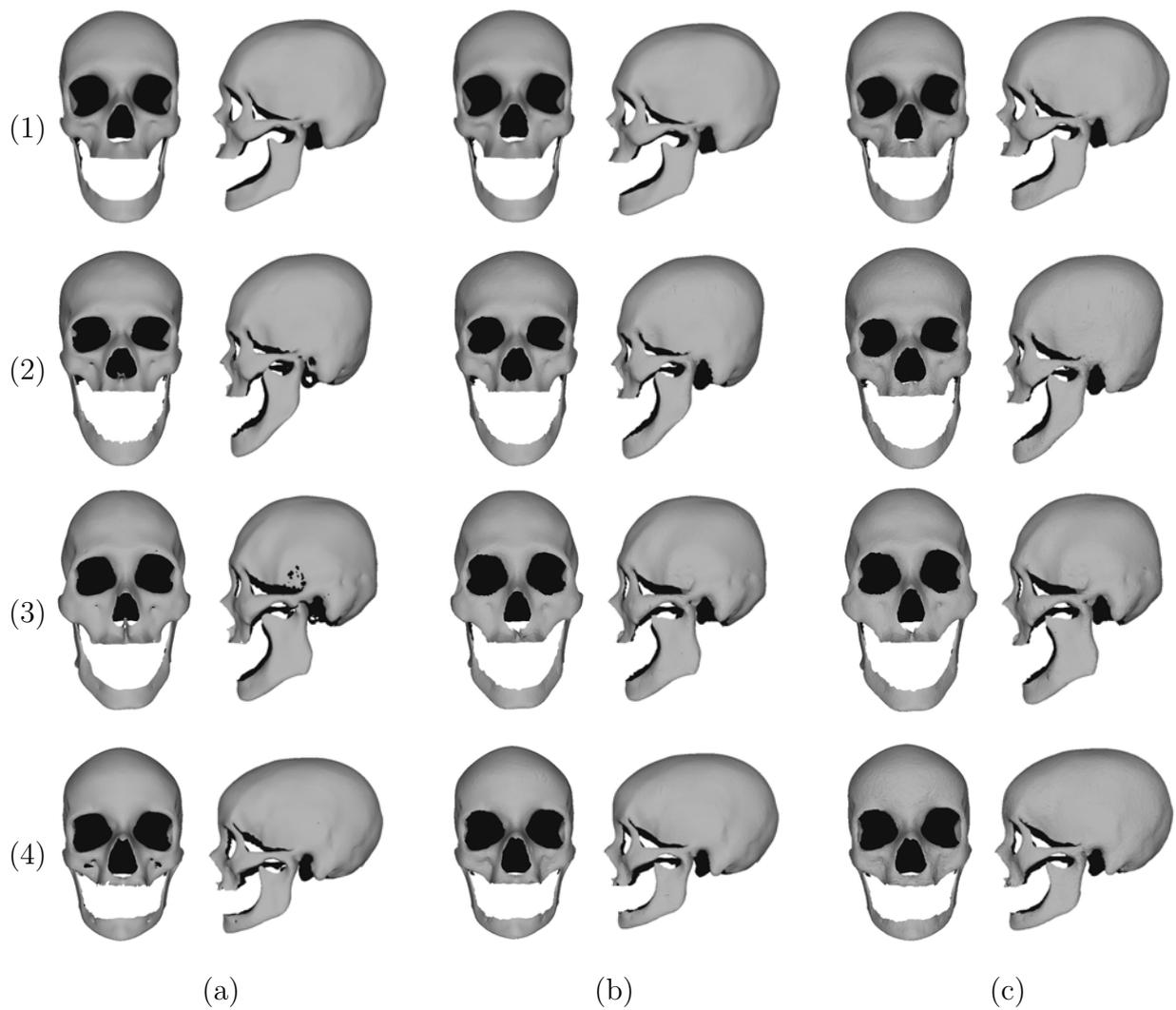


Figure 4.1: Examples of skull models used for training. (1) The reference model. (2–4) Other training samples. The resampled models (b) and reconstructed models (c) are very similar to the input models (a).

algorithm (Section 3.3) was applied to resample the training samples. Finally the ASM skull model was constructed using the active shape skull model construction algorithms described in Section 4.1.

4.2.2 Error Measures

Resampling error and reconstruction error of the training samples were computed for quantitative evaluation. Resampling error E_S was used to measure the difference between the training samples and their resampled meshes, which was computed as the mean surface distance:

$$E_S = \frac{1}{m} \sum_{j=1}^m \|\mathbf{u}_j - \mathbf{p}_j\|, \quad (4.6)$$

where \mathbf{u}_j was a vertex on a training sample, and \mathbf{p}_j was its nearest surface point on the resampled mesh, and m was the number of vertices on a training sample. The reconstruction error E_R , on the other hand, measured the difference between the training samples and their reconstructed meshes, which was also computed as the mean surface distance:

$$E_R = \frac{1}{m} \sum_{j=1}^m \|\mathbf{u}_j - \mathbf{q}_j\|, \quad (4.7)$$

where \mathbf{u}_j was a vertex on a training sample, and \mathbf{q}_j was its nearest surface point on the reconstructed mesh.

4.2.3 Results and Discussions

The training samples each has $m = 21786$ vertices. The number n of training samples is 34, which is much smaller than the number of dimensions $3m$. Therefore, the ASM constructed has $n - 1 = 33$ eigenvectors. A smaller number of eigenvectors is sufficient to capture most of the variations of the training samples. The number of eigenvectors k can be selected according to a desired residual. The residual plot in Figure 4.2 shows that the ratio of unaccounted variance decreases rapidly with increasing number of eigenvectors.

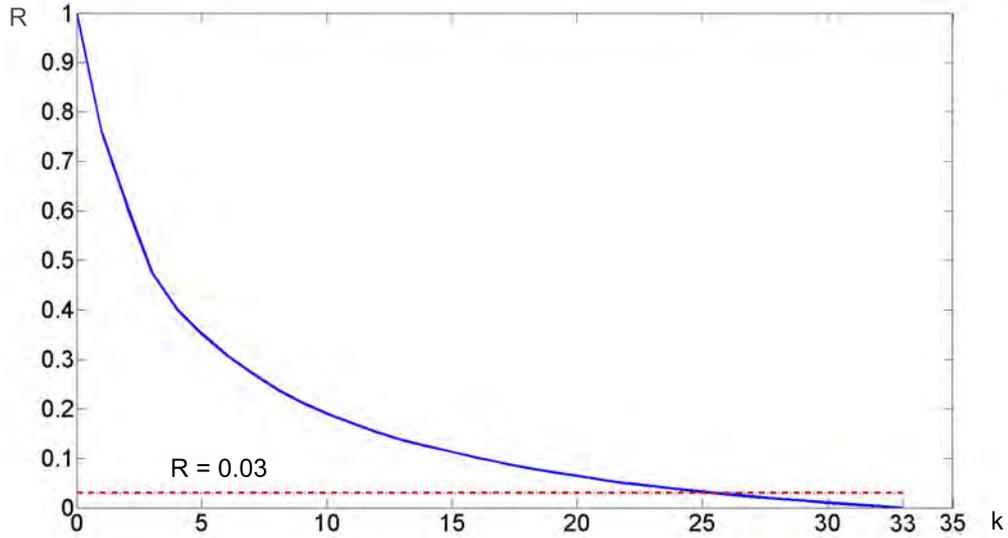


Figure 4.2: Residual plot of ASM skull modeling. The residual R decreases rapidly with increasing number k of eigenvectors used for ASM reconstruction. For a residual of 0.03, 26 eigenvectors are enough.

For a residual of 0.03, only 26 eigenvectors are required.

Figure 4.3 shows the first four modes of variation, with shape parameter values ranging from $-3\sqrt{\lambda_i}$ to $3\sqrt{\lambda_i}$. The first mode of variation shows the mouth opening or closing (Figure 4.3(1)). The second mode of variation shows shape variation at the back of the skull, the lower jaw, the eye bridge and the eye socket (Figure 4.3(2)). The third mode of variation also shows shape variation at the lower jaw and the eye bridge. In addition, it also shows shape change of the nose bone, the bottom part of the skull, the upper jaw and the change of width of the nose bridge (Figure 4.3(3)). The fourth mode of variation shows shape change of the top part of the skull and the cheek bone, together with the eye socket, the eye bridge, the nose bone, the lower jaw and the back part of the skull (Figure 4.3(4)).

The resampling error of the training samples is independent of the number k of eigenvectors used for ASM reconstruction (Figure 4.4). On the other hand, the reconstruction error is affected by both the resampling error and the number of eigenvectors used. With

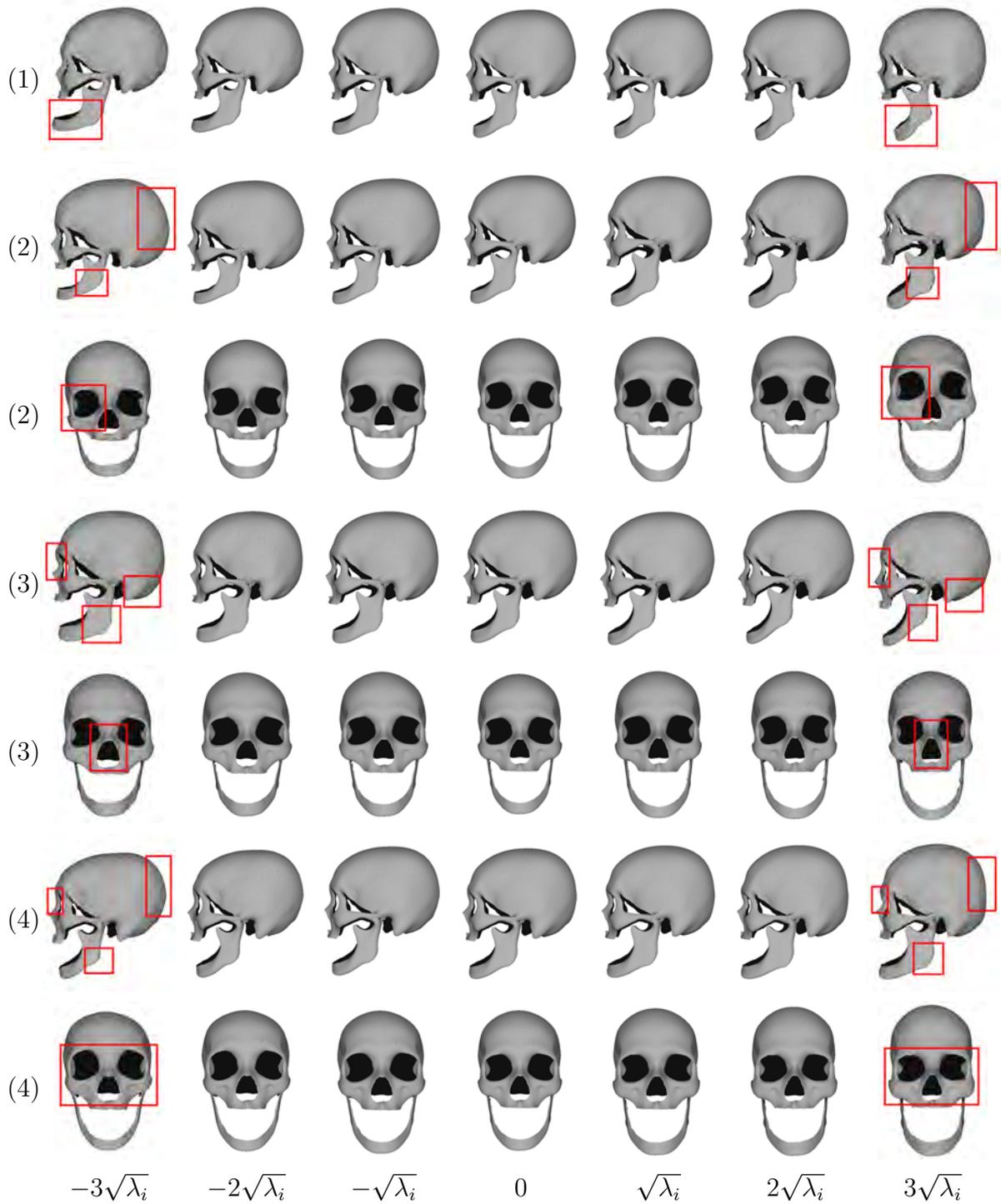


Figure 4.3: Modes of variation. (1) to (4) refer to the first four modes of variation of the skull model, ranging from $-3\sqrt{\lambda_i}$ to $3\sqrt{\lambda_i}$. The column with 0 variation illustrates the mean shape. Red boxes highlight shape variations.

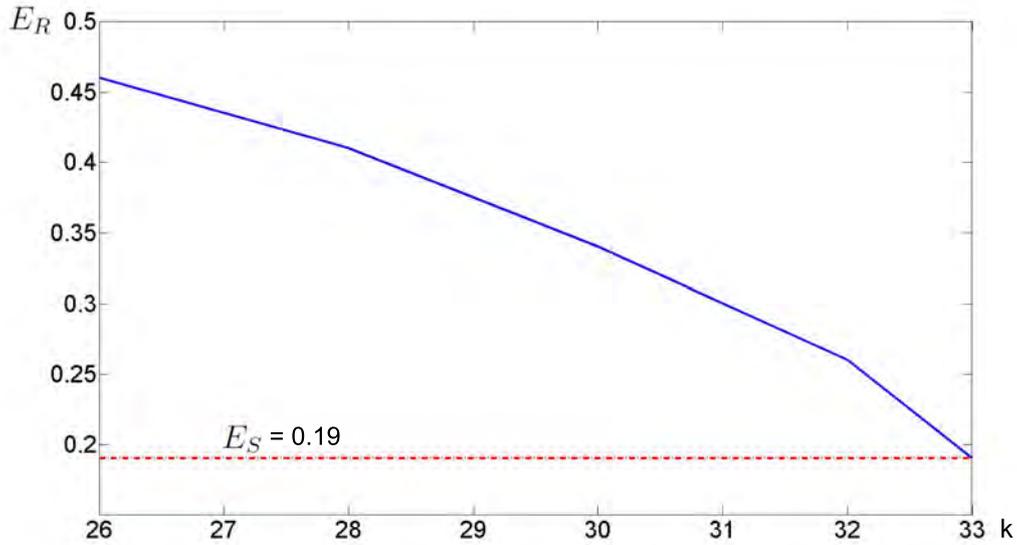


Figure 4.4: Influence of number of eigenvectors on reconstruction error. The resampling error E_S (horizontal line) is independent of the number k of eigenvectors used in ASM reconstruction. On the other hand, the reconstruction error E_R decreases with increasing k . Unit is mm.

increasing number of eigenvectors, the reconstruction error decreases. Using all 33 eigenvectors, the reconstruction error is 0.19 mm, which is the same as the resampling error (Figure 4.4), indicating practically no error contributed by ASM reconstruction. In contrast, using 26 eigenvectors, the reconstruction error increases to 0.46 mm, which is more than two times the resampling error, indicating significant error contributed by ASM reconstruction. Nevertheless, a reconstruction error of 0.46 mm is still very small compared to the average width of human skull.

4.3 Summary

This chapter first presents the active shape model (ASM) algorithm in general followed by application of ASM to model the shape of normal, complete skulls. The experiments show that only 26 of 33 eigenvectors are needed to capture 97% of the shape variations of the training samples. Moreover, the reconstruction error decreases with increasing number of

eigenvectors used in ASM reconstruction, whereas the resampling error is independent of the number of used eigenvectors. In the case of fractured, incomplete skulls (Chapter 5), both the resampling error and the reconstruction error are expected to increase.

Chapter 5

Reconstruction of Skulls

The reconstruction of a skull using active shape model is the process of searching for the shape parameters that give the best approximation of the skull. For ease of explanation, the special case of reconstructing normal, complete skulls is first introduced in Section 5.1. Next, the reconstruction algorithm is adapted to handle fractured, incomplete skulls in Section 5.2. Finally, these algorithms are evaluated on synthetic and real testing samples in Section 5.3.

5.1 Active Shape Skull Reconstruction

The reconstruction of input normal skull \mathbf{r} is basically projecting its shape vector \mathbf{s} to the eigen-space denoted by the model matrix Φ and recovering shape parameters \mathbf{b} such that

$$\mathbf{s} = \bar{\mathbf{s}} + \Phi\mathbf{b}, \quad (5.1)$$

where $\bar{\mathbf{s}}$ is the mean shape. In practice, the input skull \mathbf{r} may not be spatially aligned to the model represented by the mean shape $\bar{\mathbf{s}}$. So it is necessary to recover the similarity

transformation T that best aligns \mathbf{r} to $\bar{\mathbf{s}}$, giving

$$T(\mathbf{r}) = \bar{\mathbf{s}} + \Phi \mathbf{b}. \quad (5.2)$$

Therefore, the reconstruction of \mathbf{r} is formulated as the problem of determining the similarity transformation T and shape parameters \mathbf{b} that minimize the error E :

$$E = \|\bar{\mathbf{s}} + \Phi \mathbf{b} - T(\mathbf{r})\|^2. \quad (5.3)$$

To minimize E , Cootes [BG00] used a simple iterative algorithm as follows:

ASM Fitting Algorithm

1. Initialise the shape parameters \mathbf{b} to be zero, and set the reconstructed shape \mathbf{s} to be the mean shape $\bar{\mathbf{s}}$.
2. Repeat until convergence:
 - (a) Compute the similarity transformation T that best aligns \mathbf{r} to \mathbf{s} by minimizing $\|T(\mathbf{r}) - \mathbf{s}\|^2$.
 - (b) Compute shape parameters \mathbf{b} using the equation $\mathbf{b} = \Phi^\top (T(\mathbf{r}) - \bar{\mathbf{s}})$ derived from Equation 5.2.
 - (c) Compute the reconstructed shape \mathbf{s} as in Equation 5.1.

Initially, the reconstructed shape \mathbf{s} is initialized as the mean shape $\bar{\mathbf{s}}$. Then the best approximating T and \mathbf{b} are recovered from \mathbf{s} . In step 2(b), $\Phi^\top = \Phi^{-1}$, because the column vectors of Φ are orthogonal and have unit length. As the algorithm iterates, $\|T(\mathbf{r}) - \mathbf{s}\|^2$ decreases and \mathbf{s} converges to the best estimate of $T(\mathbf{r})$.

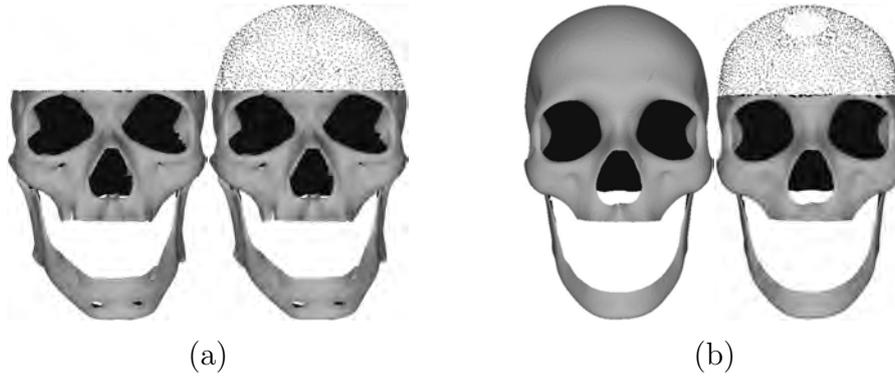


Figure 5.1: Shape vector generation. (a) The incomplete parts represented by the dots of a resampled incomplete skull are filled with $(0, 0, 0)$. (b) The corresponding vertices in the mean shape are also set to $(0, 0, 0)$.

5.2 Reconstruction of Fractured, Incomplete Skulls

In reconstructing fractured, incomplete skulls, the ASM fitting algorithm in Section 5.1 is adopted to fit to only the normal parts of a fractured, incomplete skull, and use the fitted shape parameters to generate the complete results of the skull. This approach is similar to those of [Gun05, LAV09], but is different from that of [ZLES05] which fits to the whole skull including the fractured parts. Since the fractured parts are typically displaced and rotated, their shapes are incorrect. So it is better not to use the fractured parts in ASM fitting.

For fractured, incomplete skulls, the dense correspondence algorithm is employed to resample them and identify the fractured or incomplete parts (Section 3.3). Reference vertices with no correspondence are identified as fractured or incomplete parts, and their coordinates in the shape vector are set to $(0, 0, 0)$, assuming that no normal mesh vertex lies at $(0, 0, 0)$. After resampling, the shape vector of the input skull is prepared. Let \mathbf{s}' denote the input shape vector whose coordinates of the fractured or incomplete parts are set to $(0, 0, 0)$. Correspondingly, let $\bar{\mathbf{s}}'$ and Φ' denote the mean shape and model matrix whose corresponding rows are set to $\mathbf{0}$ to remove unnecessary constraints on the fractured or incomplete parts (Figure 5.1). Then the reconstruction problem is to recover the shape

parameters \mathbf{b} that best fit \mathbf{s}' :

$$\mathbf{s}' = \bar{\mathbf{s}}' + \Phi' \mathbf{b}. \quad (5.4)$$

Since Φ' may not have an inverse, the shape parameter \mathbf{b} is recovered using the pseudo-inverse of Φ' :

$$\mathbf{b} = (\Phi'^{\top} \Phi')^{-1} \Phi'^{\top} (\mathbf{s}' - \bar{\mathbf{s}}'). \quad (5.5)$$

Then, the reconstructed complete skull \mathbf{s} is recovered using the complete mean shape $\bar{\mathbf{s}}$ and model matrix Φ :

$$\mathbf{s} = \bar{\mathbf{s}} + \Phi \mathbf{b}. \quad (5.6)$$

To handle the input skulls that are not spatially aligned to the mean shape, a best approximating similarity transformation should be recovered as described in Section 5.1. Therefore, the reconstruction algorithm for fractured, incomplete skulls is the same as the ASM Fitting Algorithm described in Section 5.1, except $\bar{\mathbf{s}}$ is replaced by $\bar{\mathbf{s}}'$ and Φ is replaced by Φ' in Step 2(b).

5.3 Experiments and Discussions

5.3.1 Data Preparation and Test Procedure

A comprehensive set of experiments was conducted to evaluate the skull reconstruction algorithm. ASM of skull was constructed using 34 normal, complete skulls (Figure 4.1). 8 other normal, complete skulls were used as testing samples (Figure 5.2). They were flipped about their lateral symmetric planes to create additional test samples. As human skulls are not exactly symmetric, the number of complete, normal test samples were double, in this way to 16.

The complete, normal testing samples were used to generate synthetic fractured, incomplete testing samples. To study how the level of severity affects the reconstruction results, synthetic skulls with four levels of severity were created: mild, moderate, severe



Figure 5.2: Examples of normal, complete skull models used for testing.

and very severe. The first three levels were either fractured or incomplete. For each of the first three levels, the skulls were manually fractured at three positions: cranial bones, facial bones and jaws (Figure 5.4), in a manner similar to real fractures (Figure 5.3). Incomplete skull samples were created by removing the fractured bone fragments (Figure 5.5). These incomplete cases may occur in forensic investigation due to criminal acts and surgery due to removal of deformed parts. Additional incomplete skulls were created by removing the top of the cranial bone or the bottom of the lower jaw. These situations could happen in real applications due to CT scanning limits while scanning the patients. The very severe testing samples had multiple defects, i.e., both fractures and missing parts (Figure 5.6). They fall into two types similar to real samples (Figure 5.7). In total, each of the first three levels of severity had 128 ($= 16 \times (3+5)$) synthetic samples, and the fourth level of severity had 32 ($= 16 \times 2$) synthetic samples. In addition, 6 other fractured, incomplete skulls constructed from patients' CT volumes were used as real fractured, incomplete testing samples. They had different degrees of fracture and incompleteness (Figure 5.7). In summary, there were 34 normal, complete training samples, 16 normal, complete testing samples, 416 ($= 128 \times 3 + 32$) synthetic fractured, incomplete testing samples and 6 real fractured, incomplete testing samples.

For the experimental procedure, in the first step, an ASM of skull was constructed using the training samples, as reported in Section 4.2. Next, the dense correspondence

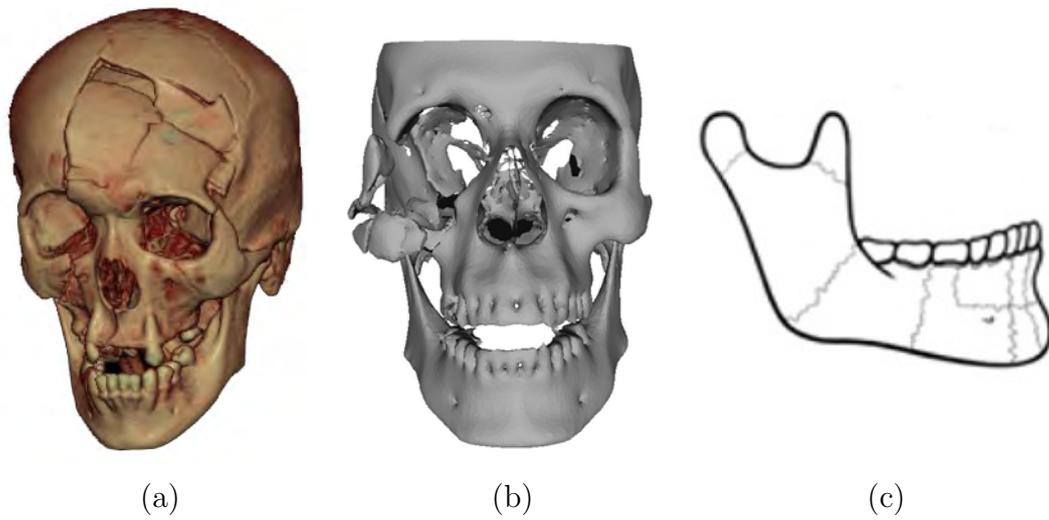


Figure 5.3: Common skull fracture patterns. (a) Cranial fractures [Neu]. (b) Facial fractures [Che14]. (c) Jaw fractures [Ini].

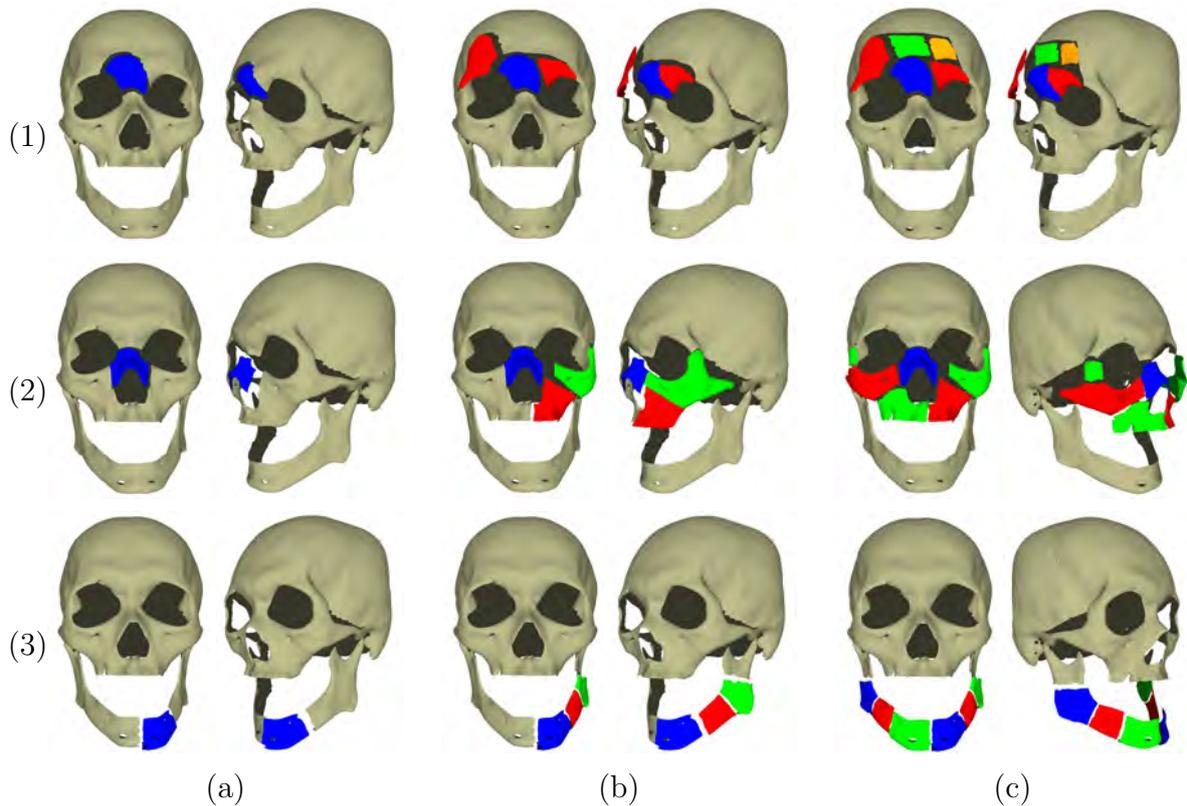


Figure 5.4: Synthetic fractured samples. (1), (2), (3) illustrate fractures at the cranial, facial and jaw bones respectively. (a), (b), (c) illustrate three levels of fracture severity: mild, moderate and severe.

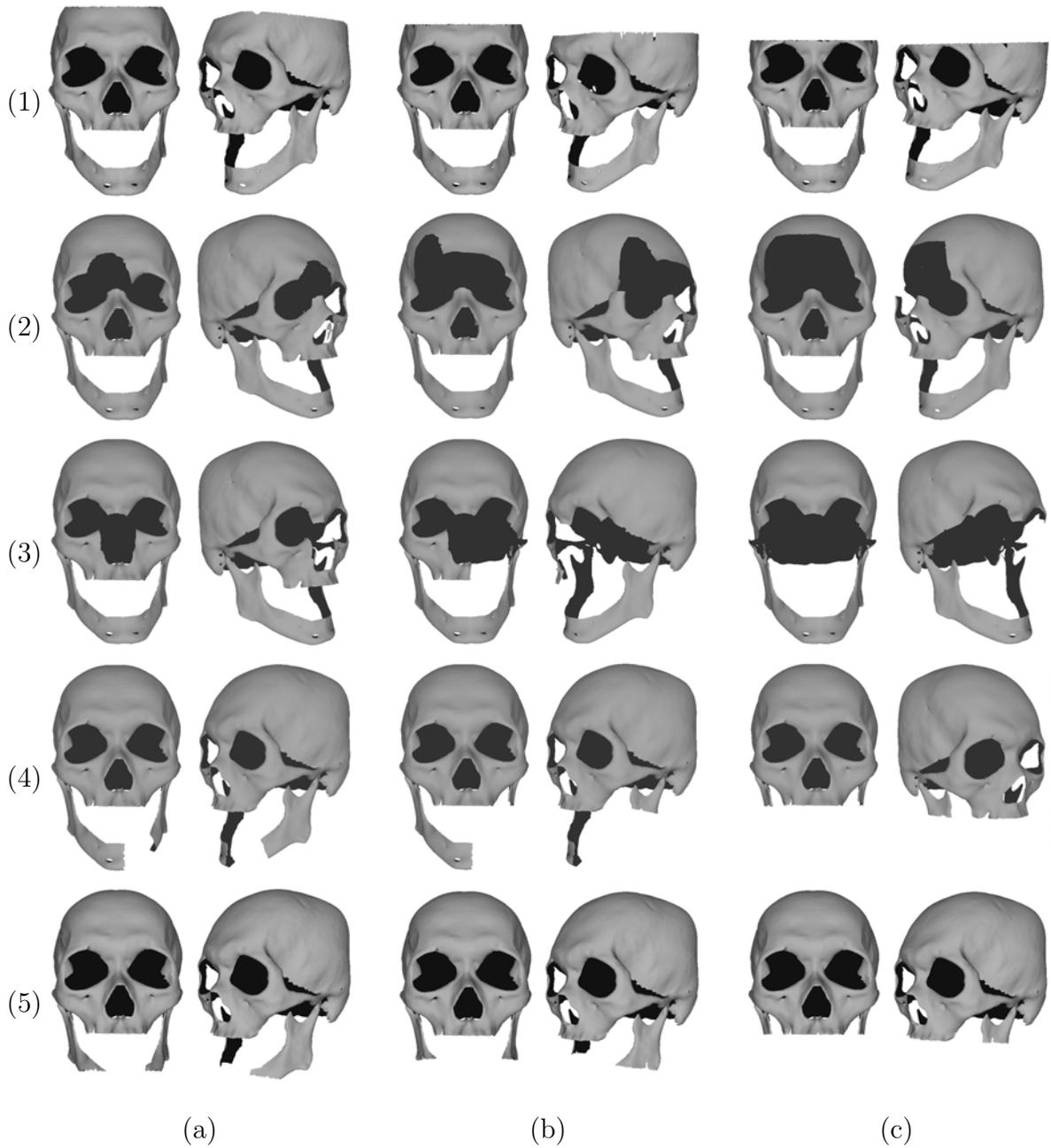


Figure 5.5: Synthetic incomplete skulls. (1) and (5) illustrate missing cranial and jaw bones due to scan limits. (2), (3), (4) illustrate incomplete cranial, facial and jaw bones due to fracture. (a), (b), (c) illustrate three levels of severity: mild, moderate and severe.

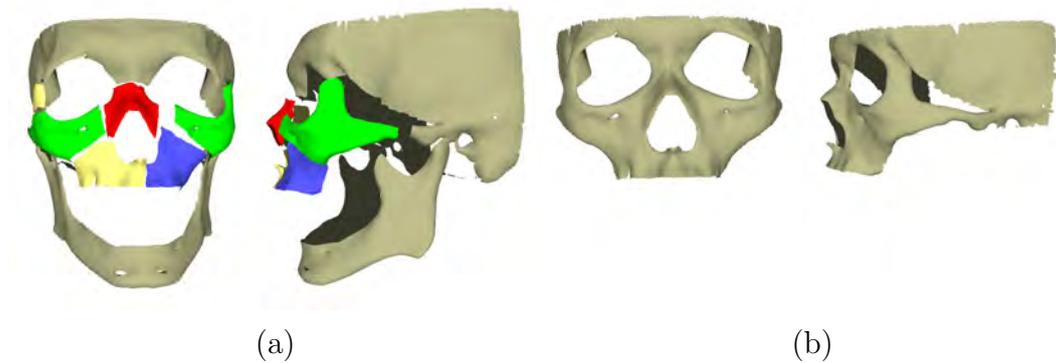


Figure 5.6: Synthetic testing samples with multiple defects. (a) Incomplete cranial bone due to scan limit and fractured facial bone. (b) Severely incomplete skull due to scan limits and minor facial fracture

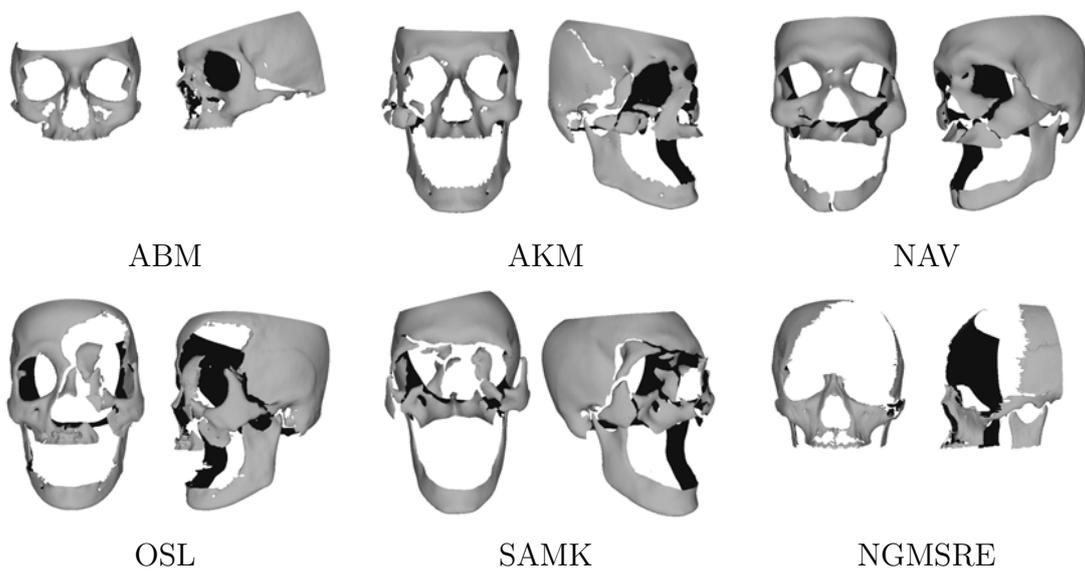


Figure 5.7: Real fractured, incomplete testing samples. All of them have different degrees of fracture and incompleteness. The first five have various fractures and NGMSRE has its congenitally deformed part removed.

algorithm was employed to resample the input skull models and identify the normal parts, if the input skulls were fractured or incomplete (Section 3.3). Then, the skulls were reconstructed by fitting the ASM to the resampled skulls using ASM fitting algorithm (Section 5.2, Section 5.3, Section 5.3).

To evaluate how the resampling process affected ASM fitting and reconstruction accuracy, three different resampling algorithms (Section 3.5) were tested, namely, LM-H, CP-S and LM-H/CP-S. LM-H used a very small set of sparsely distributed landmarks as hard constraints. CP-S used a large set of control points as soft constraints but without hard constraints. On the other hand, LM-H/CP-S used both. Three kinds of errors were measured, namely, resampling error, ASM fitting error and reconstruction error. For ease of explanation, test results for each error type is presented in a separate section.

5.3.2 Resampling Error

The qualitative results of resampling by LM-H/CP-S is compared with those of LM-H and CP-S. LM-H/CP-S uses landmarks to ensure anatomical consistency and control points to guarantee close matching. Thus, it generates more accurate resampled mesh than do LM-H and CP-S (Figure 5.8). LM-H does not use control points to guarantee close matching. So, some normal parts could be wrongly identified as fractured parts. For example, the input skull in Figure 5.8 has no landmark on the top of the skull. So, the top part of the skull is wrongly omitted in the resampled mesh. On the other hand, without landmarks, CP-S cannot ensure anatomically consistent correspondence. For example, the correspondence at the lower jaw is wrong in Figure 5.8. So, in the resampled results, the lower jaw is incomplete.

Figure 5.9 shows sample resampling results of training samples and normal, complete testing samples. The resampled meshes of LM-H/CP-S are very close to the inputs, except that the small holes of the inputs are filled up in the resampled results (Figure 5.10). This is because the reference mesh has no small holes. Normal, complete input skulls

are also supposed to have no small holes. But in practice these small holes occur due to segmentation error. For each vertex in the reference, LM-H/CP-S always finds a corresponding point on the normal, complete target. And the target model is resampled according to the mesh connectivity of the reference model. So, the small holes on the target are filled up.

For synthetic fractured, incomplete testing samples (Figure 5.11), LM-H/CP-S identifies correspondences based on two criteria: distance of the two points and similarity of the surface normals at the two points. If a corresponding point that satisfies these criteria cannot be found, then its corresponding point is regarded as missing. Therefore LM-H/CP-S is robust to fractured, incomplete skull models. The resampled incomplete testing samples are very close to the input models. For synthetic fractured testing samples, the bone fragments with large displacement are identified as fractured parts (Figure 5.11(1)(b)). On the other hand, bone fragments with small displacements are identified as normal parts (Figure 5.11(1)(a)).

Next, let us investigate the resampling results of real fractured, incomplete testing samples (Figure 5.12). Like the synthetic fractured testing samples, the bone fragments with large displacement are identified as fractured parts, and they do not appear in the resampled results. In addition, some bone fragments are fused together. This is because these cracks and the displacements of these bone fragments are small. And for the vertices in the corresponding region of the reference model, LM-H/CP-S still finds correspondence on the target skulls, and maps the mesh connectivity to the resampled model, making these bone fragments fused in the resampled model.

For quantitative evaluation of the accuracy of the correspondence and resampling process, the resampling error E_S is used to measure the difference between the input skull model and the resampled mesh using mean surface distance:

$$E_S = \frac{1}{m} \sum_{j=1}^m \|\mathbf{u}_j - \mathbf{p}_j\|, \quad (5.7)$$

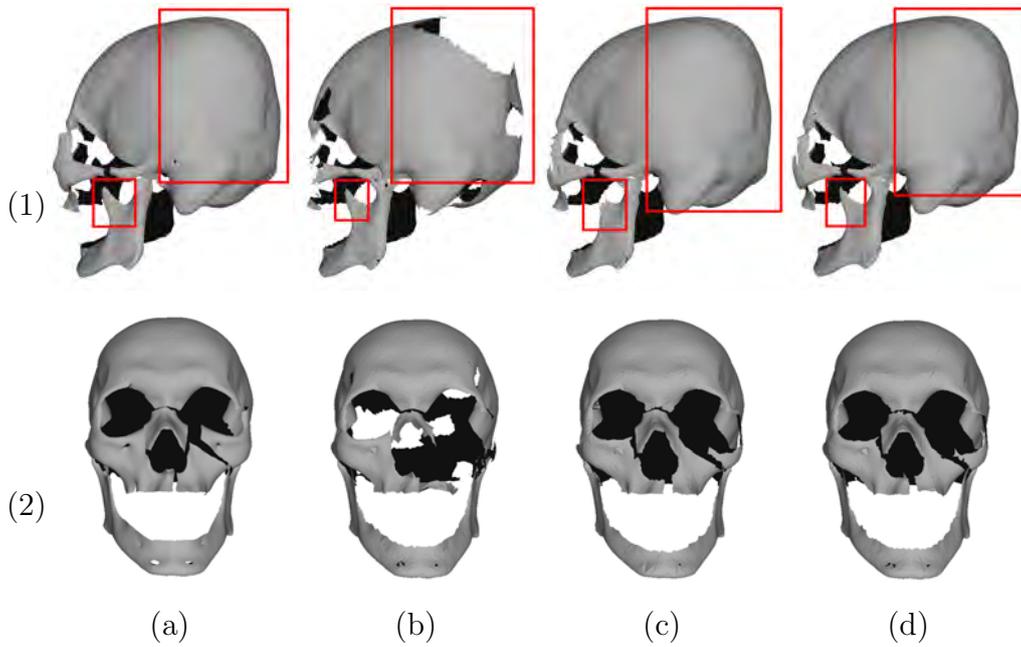


Figure 5.8: Sample resampling results of fractured, incomplete skulls. (a) Input model. (b) LM-H and (c) CP-S produce more errors than does (d) LM-H/CP-S.

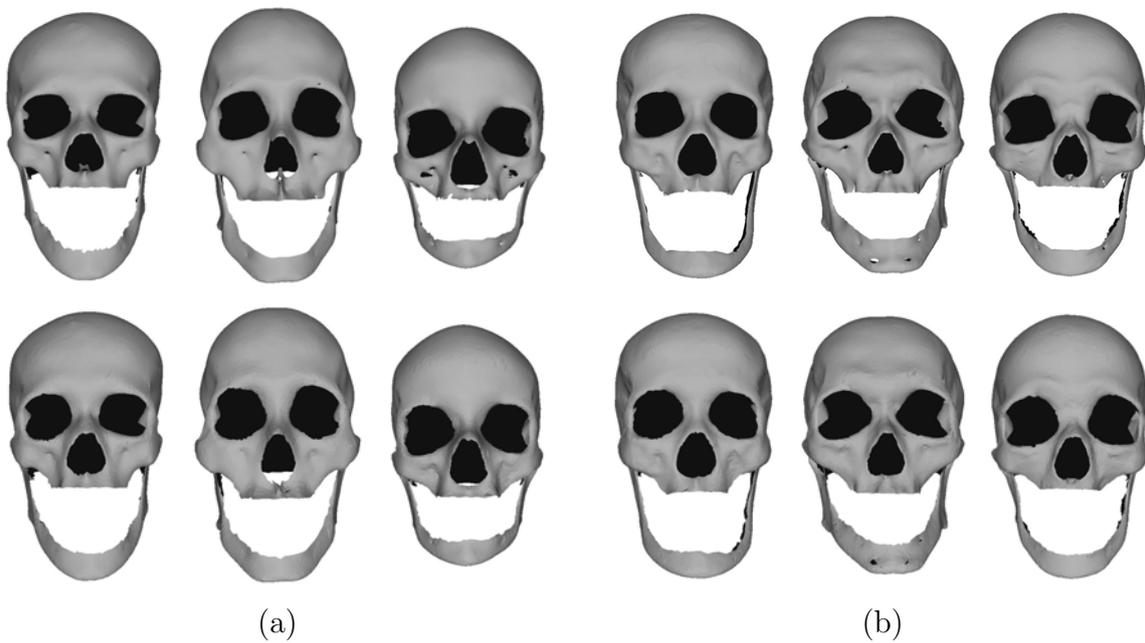


Figure 5.9: Examples of resampling results of LM-H/CP-S on training samples and normal, complete testing samples. (a) Training samples. (b) Normal, complete testing samples. The first row shows the input models and the second row shows the resampled results.

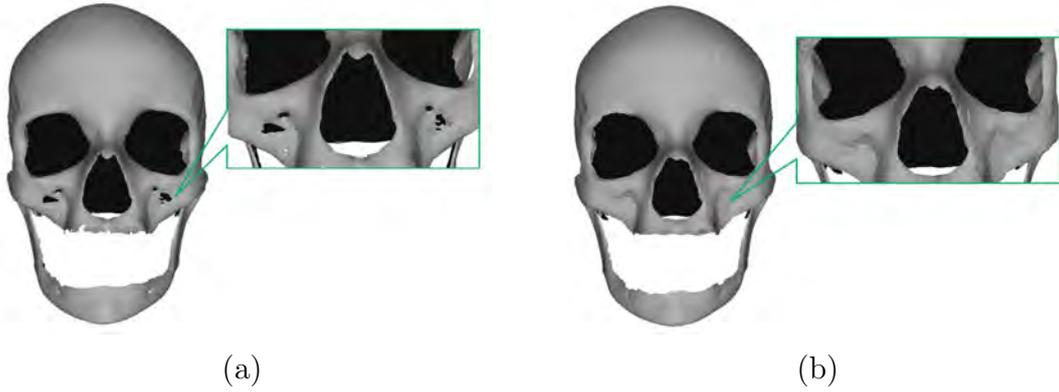


Figure 5.10: Hole filling in the resampling process. The small holes on (a) normal, complete input skulls are filled up in (b) the resampled results.

where \mathbf{u}_j is a mesh vertex on the input skull model and \mathbf{p}_j is its nearest surface point on the resampled mesh. The resampling error E_S measures the combined error of correspondence building and mesh resampling. As landmarks are used as hard constraints in the correspondence algorithm, landmarks in an input mesh remain at the same positions in the resampled mesh. On the other hand, the control points are used as soft constraints. Therefore, control points in an input mesh may change positions in the resampled mesh. Resampling errors of the fractured parts and normal parts of a fractured skull are measured separately because the fractured parts are expected to have larger resampling error than the normal parts.

First let us investigate resampling error across different sample types. For training samples and normal, complete testing samples, the resampling errors are very small (Table 5.1). The resampling errors of synthetic incomplete testing samples are very close to those of training samples and normal, complete testing samples for the results produced by LM-H/CP-S and CP-S. And they remain quite stable and are not affected significantly by the severity and position of incompleteness (Table 5.1). For the synthetic fractured testing samples (Table 5.2), their resampling errors of the normal parts are quite close to those of the synthetic incomplete testing samples. On the other hand, the resampling errors of the fractured parts vary a lot. For example, the resampling errors range from

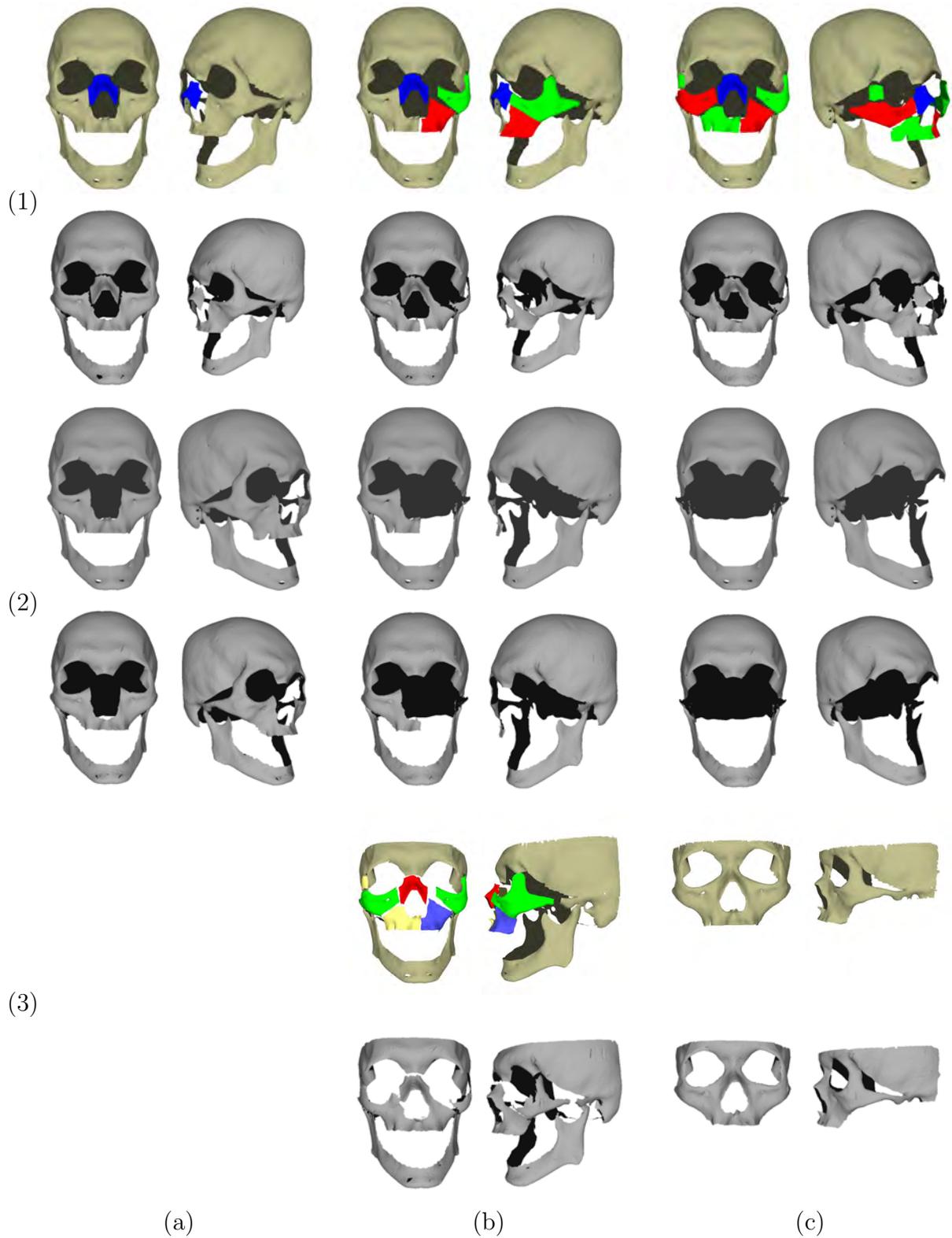


Figure 5.11: Examples of resampling results of LM-H/CP-S on synthetic fractured, incomplete testing samples. (1), (2), (3) show examples of fractured testing samples, incomplete testing samples and testing samples with multiple defects respectively. The first row of each category shows the input models and the second row shows the resampled results. (a), (b), (c) show examples with different levels of severity: mild, moderate and severe.

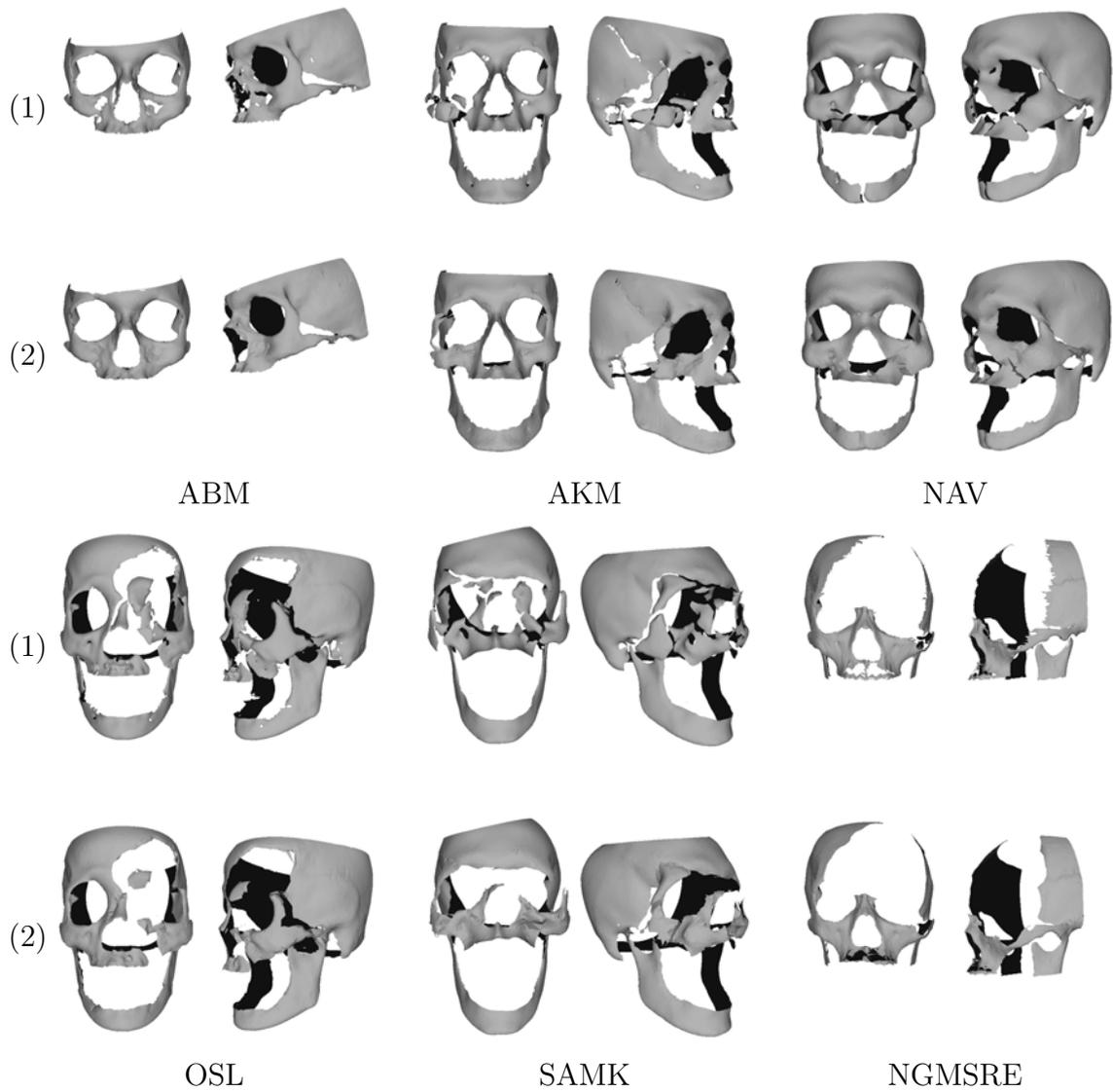


Figure 5.12: Resampling results of LM-H/CP-S on real fractured, incomplete testing samples. (1) Input models. (2) Resampled results.

0.2 to 1.68 for LM-H/CP-S, from 0.80 to 5.03 for LM-H and from 0.35 to 1.75 for CP-S. This variation is affected by the displacements of the fractured fragments. The testing samples which are fractured at the cranial bones usually have smaller resampling errors, as the fractured cranial bone fragments rotate less than the fractured facial bones and the jaw bones. For real fractured, incomplete testing samples (Table 5.3), the resampling errors are close to those of synthetic fractured testing samples.

Next, let us turn our attention to resampling error of different resampling algorithms. In comparison, the resampling errors of LM-H/CP-S are the smallest, which are almost half of the resampling errors of CP-S (Table 5.1, Table 5.2 and Table 5.3). LM-H/CP-S uses landmarks to ensure anatomical consistency and control points to guarantee close matching. Thus its resampled results are the most accurate. For normal, complete skulls, LM-H has slightly larger resampling errors than LM-H/CP-S, but still smaller than those of CP-S (Table 5.1). This shows that anatomical consistency is more important than close matching in the resampling of normal, complete skulls, which is consistent with the conclusion in Section 3.5, which states correspondence error is more important than registration error in measuring the accuracy of resampling algorithms for normal, complete samples. For fractured, incomplete skulls, the resampling errors of LM-H are almost six times those of CP-S (Table 5.1, Table 5.2 and Table 5.3). This is because with no control points to ensure close matching, the normal parts which have large registration errors are identified as fractured or incomplete parts by LM-H and are omitted, which results in large resampling error.

Table 5.1: Average resampling errors E_S of training samples, normal testing samples and synthetic incomplete testing samples. Errors are measured in mm.

Data	Severity	LM-H	CP-S	LM-H/CP-S
Training	–	0.21	0.33	0.19
Normal Testing	–	0.19	0.34	0.17
Incomplete cranial (scan limits)	mild	1.46	0.31	0.18
	moderate	1.12	0.37	0.20
	severe	1.25	0.40	0.22
Incomplete cranial (fractures)	mild	1.81	0.26	0.16
	moderate	1.81	0.25	0.16
	severe	1.84	0.26	0.16
Incomplete facial (fractures)	mild	1.82	0.26	0.16
	moderate	1.94	0.28	0.17
	severe	2.15	0.35	0.25
Incomplete jaws (fractures)	mild	1.84	0.26	0.16
	moderate	1.90	0.26	0.15
	severe	2.04	0.28	0.15
Incomplete jaws (scan limits)	mild	1.99	0.26	0.16
	moderate	2.07	0.27	0.16
	severe	2.05	0.29	0.15

Table 5.2: Average resampling errors E_S of synthetic fractured testing samples and testing samples with multiple defects. Errors are measured in mm.

Data	Severity	Part	LM-H	CP-S	LM-H/CP-S
Fractured cranial	mild	normal	1.81	0.26	0.16
		fractured	1.09	1.64	0.58
	moderate	normal	1.81	0.26	0.16
		fractured	0.93	0.71	0.28
	severe	normal	1.84	0.27	0.16
		fractured	0.80	0.53	0.22
Fractured facial	mild	normal	1.83	0.26	0.17
		fractured	1.05	0.35	0.20
	moderate	normal	1.95	0.27	0.18
		fractured	5.03	1.88	0.71
	severe	normal	2.08	0.26	0.18
		fractured	3.84	1.30	0.59
Fractured jaws	mild	normal	1.84	0.26	0.16
		fractured	2.14	1.75	0.48
	moderate	normal	1.90	0.26	0.15
		fractured	2.47	0.74	0.33
	severe	normal	2.04	0.25	0.14
		fractured	4.81	0.43	1.68
Multiple defects	moderate	normal	0.66	0.32	0.22
		fractured	4.40	1.43	0.72
	very severe	normal	0.62	0.27	0.19

Table 5.3: Average resampling errors E_S of real fractured, incomplete testing samples. Errors are measured in mm.

#	Data Name	Part	LM-H	CP-S	LM-H/CP-S
1	ABM	normal	0.26	0.28	0.24
2	AKM	normal	0.54	0.42	0.37
		fractured	7.28	0.75	3.33
3	NAV	normal	1.58	0.56	0.34
		fractured	2.62	0.74	0.30
4	OSL	normal	0.52	0.33	0.20
		fractured	7.40	1.14	2.20
5	SAMK	normal	2.09	1.71	0.30
		fractured	7.52	3.01	2.27
6	NGMSRE	normal	0.90	0.32	0.24

5.3.3 ASM Fitting Error

This section focuses on ASM fitting error. ASM fitting error measures the difference between the resampled mesh and the mesh reconstructed by ASM fitting. Qualitative results of the reconstructed meshes will be discussed in Section 5.3.4. There are two ways to measure ASM fitting error. The first way is to measure the mean displacement of mesh vertices:

$$E_F = \frac{1}{n} \sum_{i=1}^n \|\mathbf{v}_i - \mathbf{v}'_i\|, \quad (5.8)$$

where \mathbf{v}_i is a mesh vertex on the resampled mesh and \mathbf{v}'_i is its corresponding vertex on the reconstructed mesh. This formulation of ASM fitting error measures the average mesh vertex displacement as a result of ASM fitting. So, it can evaluate the amount of shape change incurred in the ASM fitting process.

Another formulation of ASM fitting error measures the mean surface distance:

$$E_G = \frac{1}{n} \sum_{i=1}^n \|\mathbf{v}_i - \mathbf{p}'_i\|, \quad (5.9)$$

where \mathbf{v}_i is a mesh vertex on the resampled mesh and \mathbf{p}'_i is its nearest surface point on the reconstructed mesh. This formulation of ASM fitting error is appropriate since it is consistent and comparable to the formulation of resampling error and reconstruction error.

First let us investigate the ASM fitting error across different sample types. Training samples are used to train the ASM. So, their ASM fitting errors are the smallest among all the samples (Table 5.4). Normal, complete test samples are not used in training, so their ASM fitting errors are larger than those of the training samples, but smaller than those of most fractured, incomplete samples (Table 5.4). For synthetic fractured, incomplete testing samples, ASM fitting error E_F increases with severity. Synthetic fractured testing samples have fractured bone fragments, which introduce noise and affect the reconstruction accuracy, making their ASM fitting errors larger than those of synthetic incomplete testing samples (Table 5.4 and Table 5.5). Real fractured, incomplete testing samples are both fractured and severely incomplete. So, they have larger ASM fitting errors than most synthetic fractured, incomplete testing samples (Table 5.6).

Next, let us turn our attention to ASM fitting errors of different resampling algorithms. In all test cases, reconstruction results using LM-H/CP-S for correspondence building have the smallest ASM fitting errors. This is because using the landmarks and control points, LM-H/CP-S ensures anatomically consistent correspondence so it can better capture the shape features of input skull models, and it also guarantees close matching so it is more robust to fractured, incomplete skull models. LM-H has smaller ASM fitting error than does CP-S. This is because using the information provided by the landmarks, LM-H can help to capture the shape variation more accurately than does CP-S, especially for skull models which are very different from the reference skull model.

For all samples, E_F is much larger than E_G . E_F measures the average displacement of the mesh vertices, whereas E_G measures the average surface distance between the resampled mesh and the reconstructed mesh. This shows that measurement of ASM fitting error by average surface distance underestimates the actual difference between the

Table 5.4: Average ASM fitting errors E_F and E_G of training samples, normal, complete testing samples and synthetic incomplete testing samples. Errors are measured in mm.

Data	Severity	E_F			E_G		
		LM-H	CP-S	LM-H/CP-S	LM-H	CP-S	LM-H/CP-S
Training	–	0.72	0.75	0.70	0.30	0.31	0.30
Testing	–	2.99	3.31	2.82	1.25	1.53	1.25
Incomplete cranial (scan limits)	mild	3.36	4.05	3.32	1.38	1.57	1.33
	moderate	3.84	4.39	3.84	1.42	1.62	1.41
	severe	3.88	4.44	3.84	1.44	1.63	1.41
Incomplete cranial (fractures)	mild	3.11	3.37	2.93	1.34	1.53	1.25
	moderate	3.14	3.32	3.04	1.35	1.52	1.25
	severe	3.15	3.38	3.07	1.35	1.51	1.27
Incomplete facial (fractures)	mild	3.11	3.21	2.92	1.33	1.50	1.24
	moderate	3.20	3.29	3.03	1.35	1.48	1.24
	severe	3.25	3.29	3.05	1.33	1.45	1.22
Incomplete jaws (fractures)	mild	3.12	3.29	2.93	1.34	1.52	1.25
	moderate	3.15	3.55	2.98	1.32	1.50	1.23
	severe	3.21	3.65	3.04	1.28	1.42	1.20
Incomplete jaws (scan limits)	mild	3.20	3.29	2.95	1.36	1.52	1.28
	moderate	3.18	3.60	2.99	1.34	1.47	1.26
	severe	3.21	3.61	3.04	1.29	1.44	1.21

Table 5.5: Average ASM fitting errors E_F and E_G of synthetic fractured testing samples and testing samples with multiple defects. Errors are measured in mm.

Data	Severity	E_F			E_G		
		LM-H	CP-S	LM-H/CP-S	LM-H	CP-S	LM-H/CP-S
Fractured cranial	mild	3.13	3.39	2.92	1.35	1.55	1.28
	moderate	3.20	3.22	3.03	1.44	1.58	1.37
	severe	3.18	3.25	3.05	1.46	1.60	1.40
Fractured facial	mild	3.10	3.23	2.99	1.36	1.56	1.30
	moderate	3.19	3.65	3.18	1.40	1.65	1.43
	severe	3.25	3.91	3.17	1.41	1.65	1.40
Fractured jaws	mild	3.11	3.25	2.92	1.35	1.55	1.28
	moderate	3.17	3.64	2.99	1.35	1.55	1.28
	severe	3.31	3.73	3.15	1.37	1.55	1.28
Multiple defects	moderate	4.99	5.27	4.91	2.04	2.10	1.91
	severe	6.40	6.65	6.00	2.18	2.37	1.97

Table 5.6: Average ASM fitting errors E_F and E_G of real fractured, incomplete testing samples. Errors are measured in mm.

#	Data	E_F			E_G		
		LM-H	CP-S	LM-H/CP-S	LM-H	CP-S	LM-H/CP-S
1	ABM	6.87	6.97	6.16	2.26	2.51	2.03
2	AKM	4.18	4.14	3.81	1.28	1.32	1.24
3	NAV	4.00	4.16	3.70	1.47	1.55	1.38
4	OSL	3.29	3.79	3.25	1.44	1.54	1.42
5	SAMK	4.60	4.94	4.10	1.70	1.72	1.61
6	NGMSRE	4.87	5.10	4.82	1.77	1.84	1.59

resampled mesh and the reconstructed mesh. Nevertheless, mean surface distance is still widely used [ZLES05, ZCL13] for error measurement because it can be computed easily and automatically.

5.3.4 Skull Reconstruction Error

Some qualitative results of skull reconstruction based on LM-H/CP-S are shown in Figure 5.13 to 5.17. The training samples are used in the construction of the ASM of skull. So, their reconstructed skull models are very close to the training samples (Figure 5.13(a)). For the testing samples, the reconstructed skull models are similar to the testing samples, except there are some mismatches and differences in details (Figure 5.13(b)). The reconstructed skull models are generally more flat than the testing samples. For example, the brow bones of the reconstructed skull models are usually not as obvious compared to those of the testing samples. This is because ASM works on the global structure and does not take into account detailed local surface fitting.

Synthetic fractured testing samples have fractured bone fragments (Figure 5.14(1)). Therefore, they have more noise and their reconstruction results are less accurate than those of the synthetic incomplete testing samples (Figure 5.14(2)). For example, the shapes of the reconstructed eye sockets of the fractured testing samples are less similar to those in the inputs. The testing sample with the severe defect does not have lower jaw bone (Figure 5.14(3c)). So, its reconstructed shape is essentially the mean shape.

For real fractured, incomplete testing samples (Figure 5.15 to 5.17), the reconstruction method can reconstruct the overall shape of the input skull and predict the shape of the fractured or incomplete parts. For example, the jaws of the reconstructed AKM (Figure 5.15), NAV, OSL (Figure 5.16) and SAMK (Figure 5.17) are open, just as for the inputs. There are small differences between the inputs and the reconstructed skulls, such as the turning point of the jaw bones. This is because ASM does not take into account detailed local surface fitting. In addition, for the bone fragments which are rotated

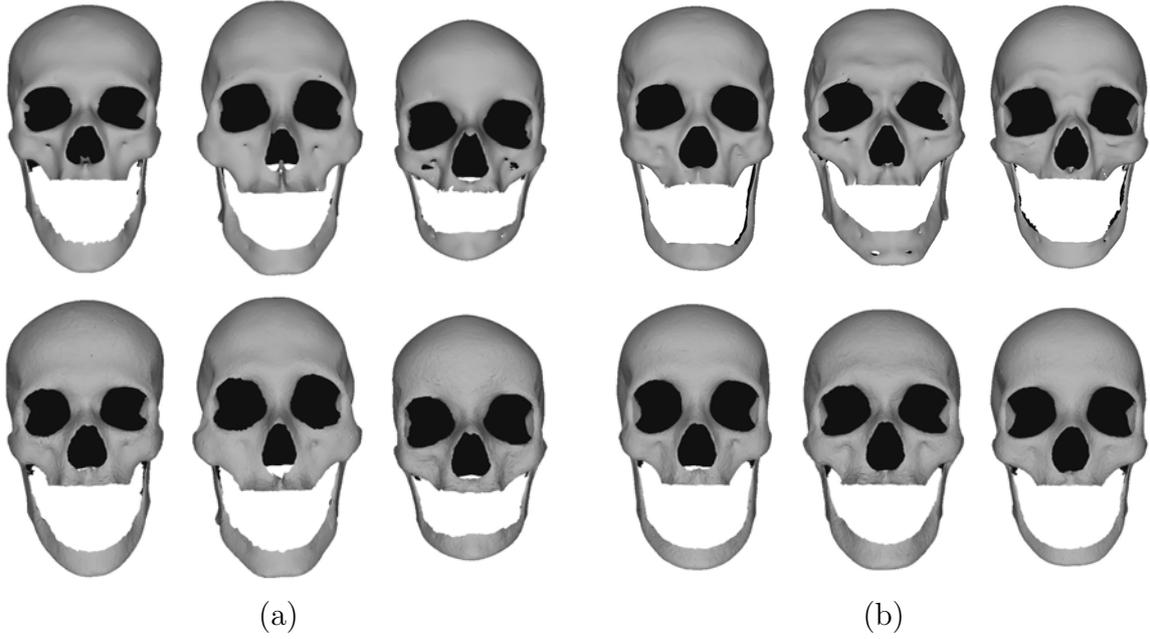


Figure 5.13: Examples of reconstructed results of training samples and normal, complete skull testing samples. (a) Training samples. (b) Normal, complete testing samples. The first row shows the input models and the second row shows the reconstructed results.

and displaced, the reconstruction algorithm can recover their correct look. For example, the left facial bones of OSL are fractured and rotated (Figure 5.16). The reconstructed model provides a good estimate of the normal appearance of the left facial bones. On the other hand, for bone fragments with small displacements and insignificant rotation, the reconstructed results just aligned to the surfaces of the bone fragments. For example, the displacement of NAV’s jaw bone is small (Figure 5.16). So, the reconstructed result just aligns to the bone fragments and fills up the crack between them.

To quantitatively measure the reconstruction results, reconstruction error is used, and it is computed as the mean surface distance between the ground truth and the reconstructed mesh:

$$E_R = \frac{1}{n} \sum_{i=1}^n \|\mathbf{v}_i^* - \mathbf{q}'_i\| \quad (5.10)$$

where \mathbf{v}_i^* is a mesh vertex on the ground truth and \mathbf{q}'_i is its nearest surface point on the reconstructed mesh.

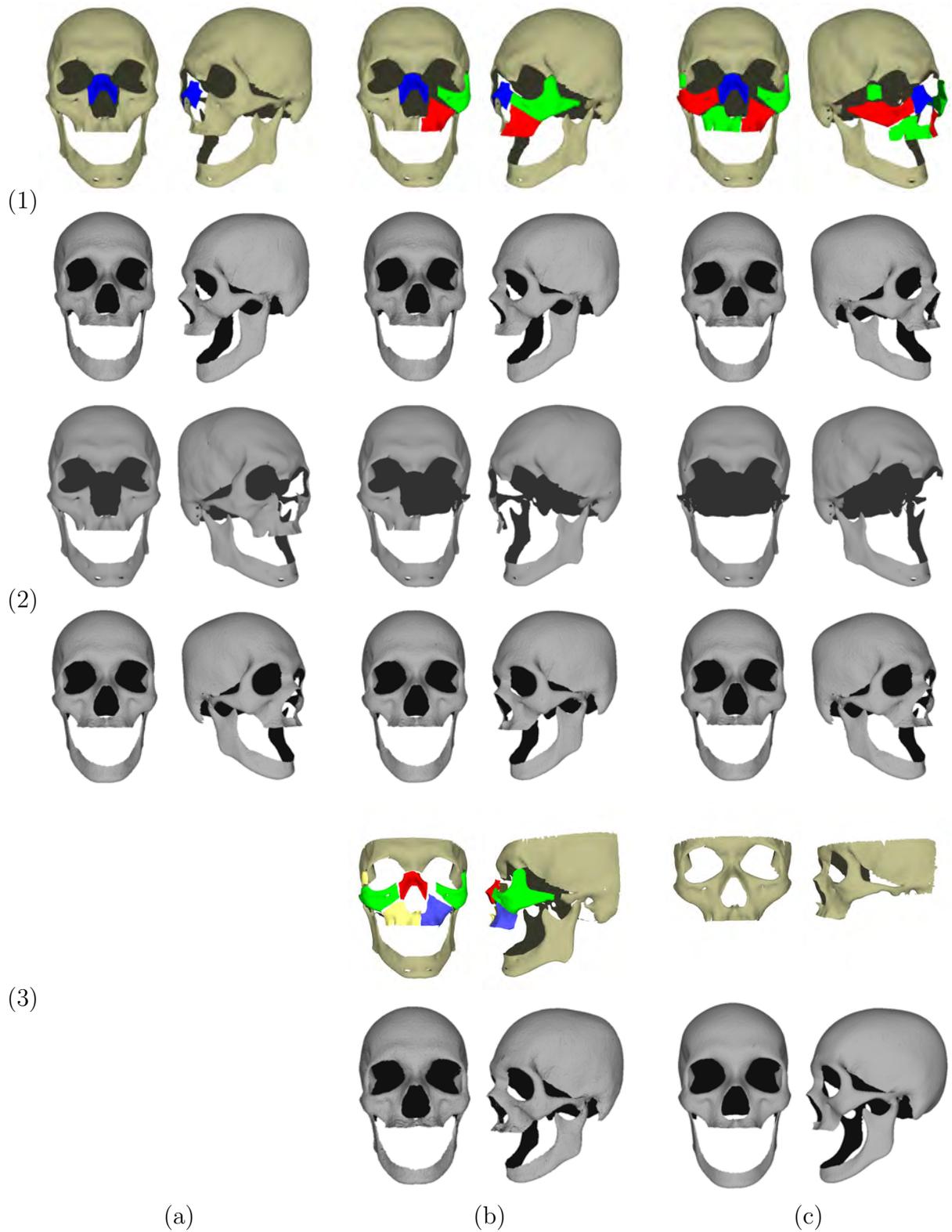


Figure 5.14: Examples of reconstruction results of synthetic fractured, incomplete testing samples. (1), (2), (3) show examples of fractured testing samples, incomplete testing samples and testing samples with multiple defects respectively. The first row of each category shows the input models and the second row shows the reconstructed results. (a), (b), (c) show examples with different levels of severity: mild, moderate and severe.

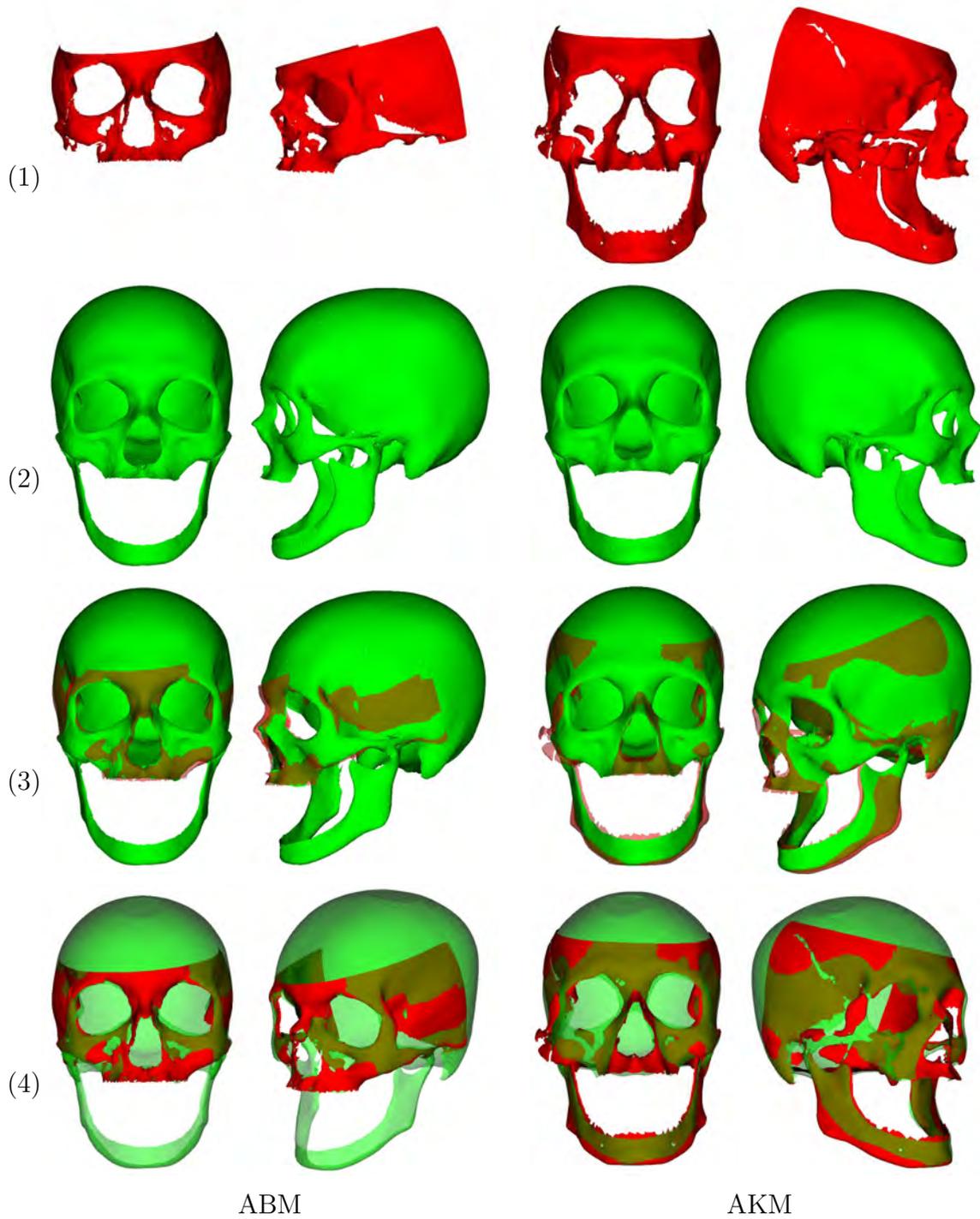


Figure 5.15: Reconstruction results of ABM and AKM. (1) Input models. (2) Reconstructed models. (3), (4) Comparison between inputs and reconstructed models.

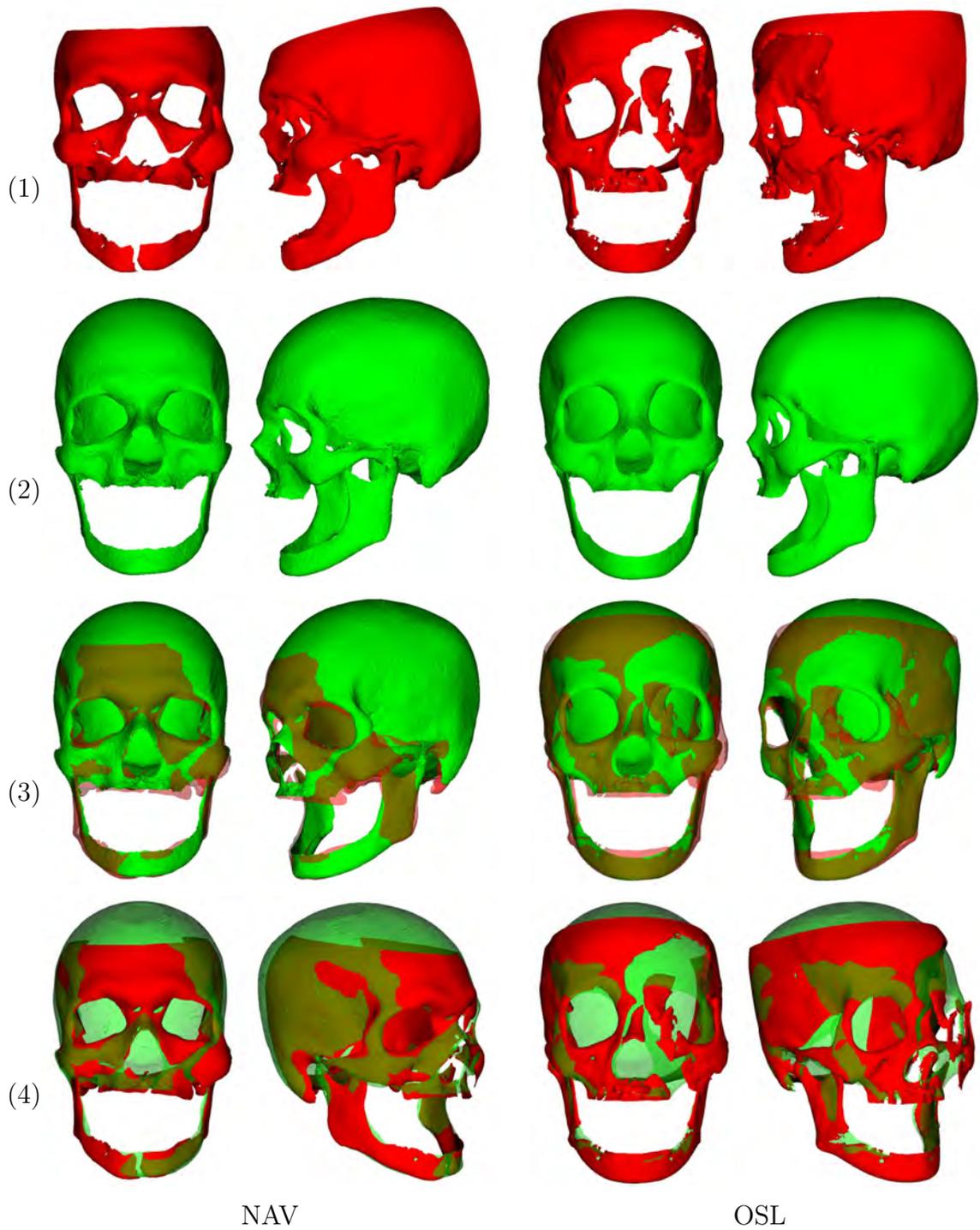


Figure 5.16: Reconstruction results of NAV and OSL. (1) Input models. (2) Reconstructed models. (3), (4) Comparison between inputs and reconstructed models.

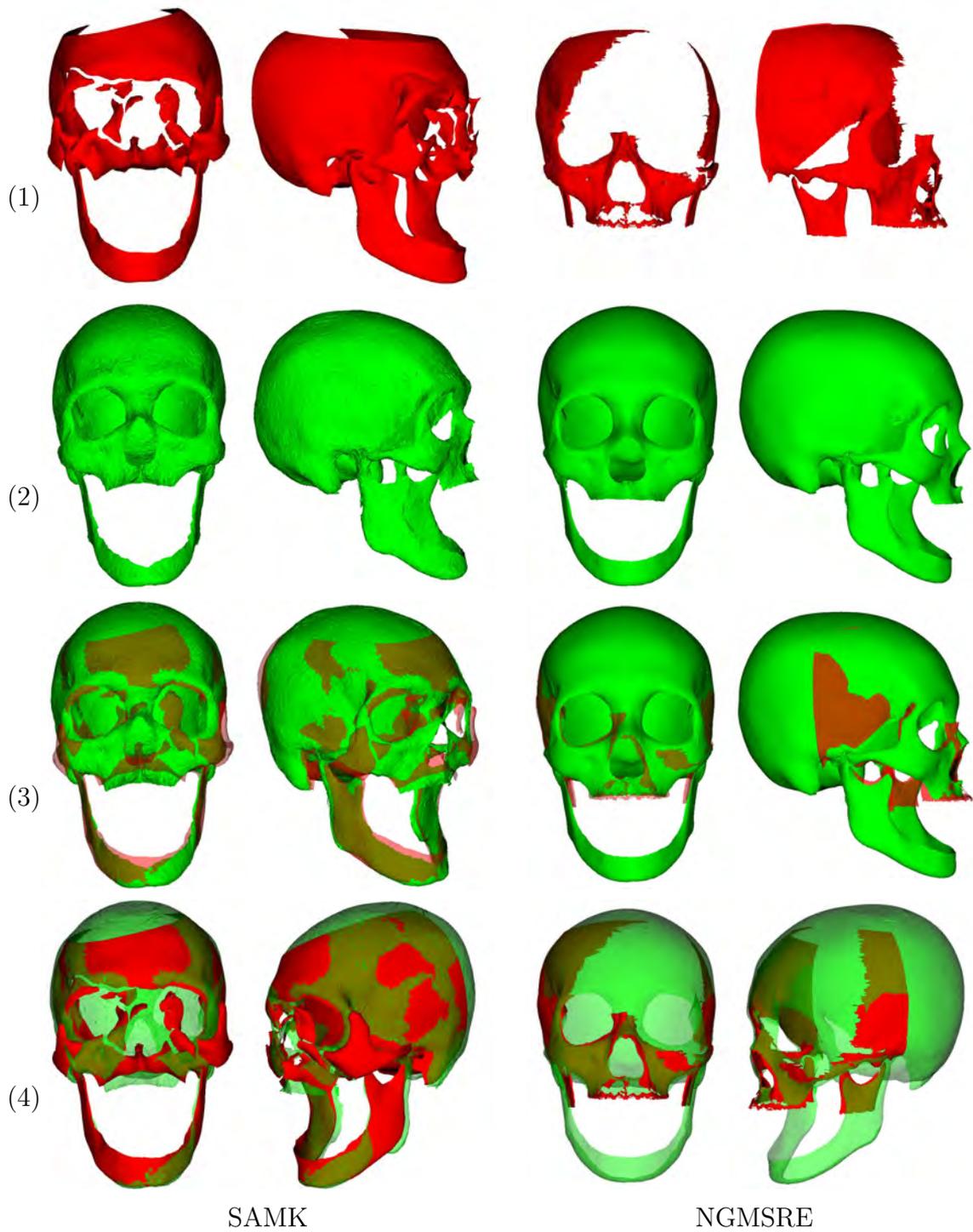


Figure 5.17: Reconstruction results of SAMK and NGMSRE. (1) Input models. (2) Reconstructed models. (3), (4) Comparison between inputs and reconstructed models.

First let us investigate reconstruction error across different sample types. As the training samples are used in the construction of skull ASM, their reconstruction errors are the smallest among all samples (Table 5.7). The normal, complete testing samples have larger reconstruction errors than do the training samples, but the smallest reconstruction errors among all testing samples. For synthetic incomplete testing samples, their reconstruction errors are affected by the severity of incompleteness (Table 5.7). With increasing severity from mild to severe, the reconstruction error consistently increases. In addition, incompleteness due to scan limits has larger reconstruction error than do incompleteness due to fracture (Table 5.7). This is because scan limits usually remove a larger portion of the skulls than do fractures (Figure 5.5). Interestingly, the synthetic fractured testing samples (Table 5.8) have larger reconstruction error than do their corresponding incomplete testing samples, which shows that bone fragments in fractured skull models introduce reconstruction errors.

Next, let us turn our attention to the reconstruction errors in relation to different correspondence and resampling algorithms. In all cases, LM-H/CP-S has the smallest reconstruction errors. With anatomical consistency and close matching, LM-H/CP-S resamples the input model more accurately and consistently across all samples. So the ASM based on LM-H/CP-S can capture more shape features than do LM-H and CP-S. LM-H has smaller reconstruction error than CP-S, which shows that anatomical consistency enforced by LM-H can help ASM to reconstruct input skulls more accurately than does close matching enforced by CP-S.

Table 5.7: Average reconstruction error E_R of training samples, normal testing samples and synthetic incomplete samples. Errors are measured in mm.

Data	Severity	LM-H	CP-S	LM-H/CP-S
Training	–	0.48	0.60	0.46
Testing	–	1.46	1.70	1.34
Missing cranial (scan limits)	mild	1.63	1.87	1.48
	moderate	1.77	1.97	1.65
	severe	1.78	1.96	1.70
Missing cranial (fractures)	mild	1.47	1.75	1.34
	moderate	1.49	1.74	1.34
	severe	1.52	1.75	1.37
Missing facial (fractures)	mild	1.47	1.72	1.34
	moderate	1.53	1.77	1.39
	severe	1.58	1.78	1.43
Missing jaws (due to fractures)	mild	1.48	1.75	1.35
	moderate	1.55	1.78	1.40
	severe	1.72	1.87	1.61
Missing jaws (scan limits)	mild	1.61	1.82	1.49
	moderate	1.66	1.83	1.54
	severe	1.72	1.87	1.60

Table 5.8: Average reconstruction error E_R of synthetic fractured testing samples and testing samples with multiple defects. Errors are measured in mm.

Data	Severity	LM-H	CP-S	LM-H/CP-S
Fractured cranial	mild	1.47	1.73	1.34
	moderate	1.47	1.74	1.35
	severe	1.49	1.74	1.37
Fractured facial	mild	1.47	1.77	1.36
	moderate	1.52	1.84	1.43
	severe	1.61	1.98	1.56
Fractured jaws	mild	1.47	1.78	1.36
	moderate	1.57	1.83	1.45
	severe	1.86	1.94	1.75
Multiple defects	moderate	2.80	2.98	2.58
	severe	2.87	2.94	2.75

5.3.5 Estimation of Reconstruction Error of Real Samples

Correspondence building and ASM fitting both contribute errors to the skull reconstruction process. It is important to investigate how much error each of them contributes to skull reconstruction so that the reconstruction process can be further improved in the future.

Resampling error E_S measures the error of both correspondence building and mesh resampling. It is small compared to reconstruction error E_R , typically about one tenth of the reconstruction error. It remains stable across all sample types. Most of them range from 0.1 to 0.35. As resampling error E_S is very small and remains stable, it contributes less to the reconstruction error E_R . A plot of resampling error E_S vs. reconstruction error E_R for LM-H/CP-S on all samples (Figure 5.18) shows that there is no significant correlation between E_S and E_R .

ASM fitting error is much larger compared to resampling error. So ASM fitting error must contribute to most of the reconstruction error. The plots of ASM fitting error E_F vs. reconstruction error E_R and E_G vs. E_R (Figure 5.19) show that most samples fall roughly on a straight line, with some outliers falling further away from the straight line. So, there is a strong linear relationship between ASM fitting error and reconstruction error.

In practice, the reconstruction errors of real fractured, incomplete samples are unknown because of the lack of ground-truth. To the best of our knowledge, nobody has attempted to estimate the reconstruction error of real fractured, incomplete samples. The linear relationship observed between the ASM fitting error and the reconstruction error makes it possible to estimate reconstruction error based on ASM fitting error. Using a robust regression method [HW77], the ASM fitting errors and the reconstruction errors of all samples can be fitted to a line. This robust regression method iteratively reweights each sample and minimizes the weighted sum of least squares. The weight of each sample ranges from 0 to 1. With a smaller weight, the sample is more likely to be an outlier.

As there are two ASM fitting errors E_F and E_G , three lines can be fitted: E_F vs. E_R ,

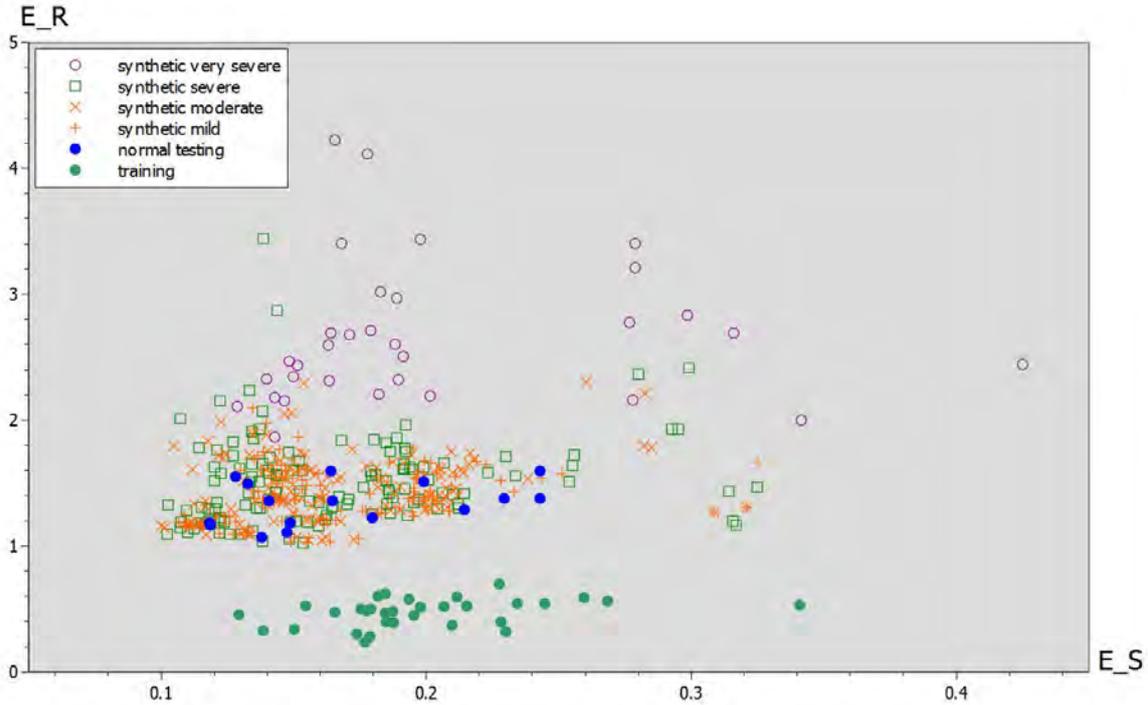


Figure 5.18: A plot of resampling error E_S vs. reconstruction error E_R .

E_S vs. E_R and E_F , E_G vs. E_R . The first two are single-fits, as each line uses one ASM fitting error only. The third line is a dual-fit. As we have no real fractured data with ground truth to tell which of these three methods is the best, all of them are presented. These fitted lines are evaluated on two criteria: fitting error and possibility of outlier. Fitting error computes the averaged weighted sum of absolute residuals of all samples. The possibility of outlier is computed as 1 minus the averaged weight of the samples. Note that this possibility measurement is not strictly a probability. The evaluation of fitted lines are shown in Table 5.9. Dual-fit has smaller fitting error and possibility of outlier than single-fit with E_G . Single-fit with E_F has the smallest percentage of outliers but the largest fitting error. Further study should be performed to investigate which one of these three fitting methods is more appropriate.

Using the fitted lines, the estimated reconstruction errors of real fractured, incomplete testing samples are computed (Table 5.10) and plotted in Figure 5.19. The estimated re-

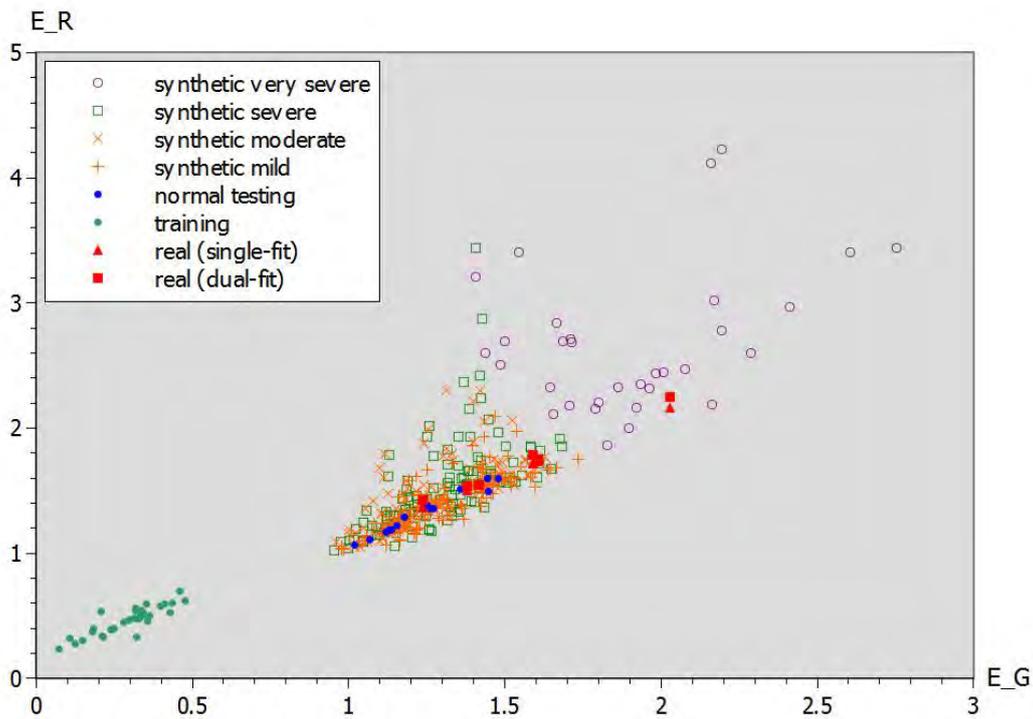
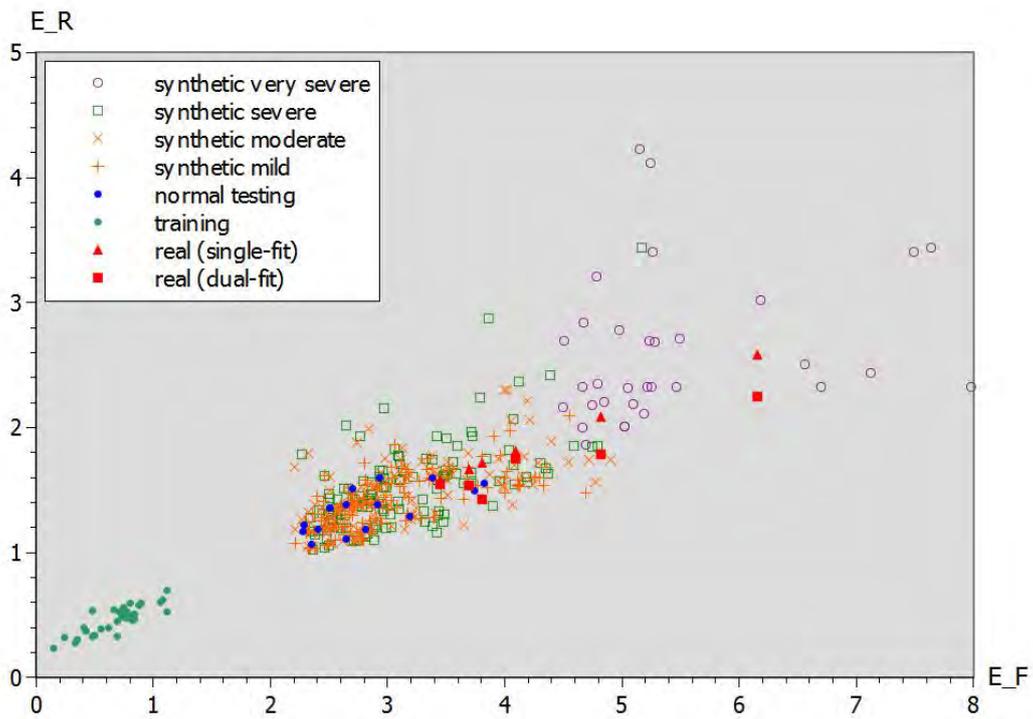


Figure 5.19: Plots of ASM fitting error E_F vs. reconstruction error vs. E_R and E_G vs. E_R .

Table 5.9: Evaluation of fitted lines.

Methods	Fitting error (mm)	Possibility of outliers
Single-Fit with E_F	0.14	0.11
Single-Fit with E_G	0.08	0.16
Dual-Fit	0.07	0.16

Table 5.10: Estimated reconstruction errors of real fractured, incomplete testing samples. Errors are measured in mm.

	Single-Fit given E_F	Single-Fit given E_G	Dual-Fit
ABM	2.58	2.16	2.24
AKM	1.71	1.36	1.42
NAV	1.66	1.5	1.53
OSL	1.57	1.54	1.54
SAMK	1.81	1.73	1.75
NGMSRE	2.08	1.71	1.78

construction errors vary with severity of fracture and incompleteness. For example, ABM and NGMSRE are the most severely incomplete and they have the largest reconstruction errors. The estimated reconstruction errors of these real fractured, incomplete samples are within the range of those of synthetic testing samples (Figure 5.19), indicating that the estimated errors are credible.

5.4 Summary

This chapter presents a skull reconstruction algorithm using Active Shape Model (ASM). First the dense correspondence algorithm is employed to resample an input skull model and identify the normal parts. Then the ASM is fitted to the resampled model to generate the reconstructed model. A comprehensive set of testing samples are used to test the skull reconstruction algorithm in relation to three resampling methods: LM-H, CP-S and LM-

H/CP-S.

Experimental results show that with anatomical consistency and close matching, LM-H/CP-S resamples input skulls the most accurately. LM-H used a small set of landmarks as hard constraints. Although it ensures anatomical consistency in the region around the landmarks, it cannot guarantee close matching of skull shapes. So the reconstruction algorithm that uses LM-H has larger reconstruction errors than that using LM-H/CP-S. However, it has smaller reconstruction errors than that using CP-S, which shows that anatomical consistency is more important than close matching in skull reconstruction.

ASM fitting error increases with increasing severity of fracture and incompleteness. Measurement of ASM fitting error measured according to average surface distance is smaller than ASM fitting error measured according to mesh vertex displacement. Therefore, it tends to underestimate the actual difference between the resampled mesh and the reconstructed mesh.

For the reconstruction process, reconstruction error also increases with increasing severity. The proposed skull reconstruction method can reconstruct the overall shape of the input skull and predict the shape of the fractured or incomplete parts. But there are some local mismatches between the normal parts of the input skull and the reconstructed skull, because ASM reconstruction does not account for detailed local surface fitting.

In comparison, resampling error is almost one tenth the ASM fitting error and reconstruction error. This shows that the resampling process is sufficiently accurate for skull reconstruction. Most of the error of reconstruction process is due to ASM fitting. Experiments show that there is a significant linear relationship between ASM fitting error and reconstruction error. This linear relationship leads to a method for estimating the reconstruction error of real fractured, incomplete skulls given ASM fitting error. The estimated reconstruction error also increases with increasing severity, which is consistent with the previous conclusion.

Chapter 6

Conclusion

This thesis has made three major contributions. First, an accurate and efficient dense correspondence algorithm for 3D skull models is developed. The proposed algorithm uses thin-plate spline (TPS) for non-rigid registration. It uses anatomical landmarks as hard constraints to ensure anatomically consistent correspondence, and control points sampled on the skull surfaces as soft constraints to provide additional local shape constraints to ensure close matching of the reference and target surfaces. Test results show that the proposed algorithm can achieve more accurate and robust dense correspondence than other TPS-based algorithms that use only hard constraints or soft constraints but not both. Test results also show that low registration error does not always imply low correspondence error. So, both error measures should be used together to evaluate the accuracy of dense correspondence algorithms.

Second, the proposed dense correspondence algorithm is applied to improve the accuracy of active shape model (ASM) reconstruction of fractured, incomplete skulls. An ASM of skull is constructed using normal, complete training samples. Next, the proposed dense correspondence algorithm is used to resample the input skull models and identify the normal parts. Then, the skulls are reconstructed by fitting the ASM to the resampled skulls. Test results show that accurate dense correspondence increases the accuracy of ASM reconstruction.

Third, the error source of ASM skull reconstruction is investigated. The two primary sources of ASM skull reconstruction error are resampling error, which includes correspondence error, and ASM fitting error. Test results show that resampling error is about 10 times smaller than ASM fitting error and reconstruction error. So, ASM fitting error is the main source of reconstruction error. Test results also show that there is a linear relationship between ASM fitting error and reconstruction error, which allows reconstruction error of real fractured, incomplete skulls to be estimated by ASM fitting error. In addition, error measured in terms of surface distance such as E_G and E_R tends to underestimate the error, because the actual error in terms of mesh vertex displacement such as E_F is larger than error based on surface distance E_G .

The proposed methods have been successfully applied to the reconstruction of fractured, incomplete skulls, and have overcome some weaknesses of existing methods. Nevertheless, some difficulties remain and should be further investigated. First, the proposed skull reconstruction method can reconstruct the overall shape of the input skull and predict the shape of the fractured or incomplete parts. But there are some local mismatches between the normal parts of the input skull and the reconstructed skull. This is because ASM works on the global structure and does not take into account local surface alignment. A possible solution to this problem is to apply a registration algorithm to complement ASM. A candidate registration algorithm is the surface continuity-constrained registration developed by Cheng [Che14]. This registration algorithm ensures that the reconstructed surface is flushed with the surface of the normal parts of the input skull, resulting in a continuous surface.

Second, human skulls differ greatly in details among people of different races, genders and ages. In some applications such as surgery planning, the race, gender and age of the subject is known. The current ASM cannot use this information. A possible way to use such information is to construct an ASM for each race, gender and age group. Then the ASMs can better capture the specific shape details of each group. Moreover, a clustering algorithm may be used to automatically divide the training samples into various coherent

groups. In this way, the reconstruction accuracy should be improved.

The most difficult problem is the estimation of the shapes corresponding to the missing part. When a small part of the skull is missing, ASM can reliably estimate the shape corresponding to the missing part. However, when a large part of the skull is missing, then the correlation between the shape of the part that is present and the missing part is very weak. ASM essentially just reproduces the mean shape, which is not expected to be the most accurate result. To estimate the shape corresponding to the large missing part accurately, additional information and constraints are required. For example, in surgery planning, a patient's photograph may be used to provide additional information. However, in forensic investigation and archeology where the person's identity is likely unknown, it would be a great challenge to obtain useful information for skull reconstruction. Further research in this direction would be very interesting and beneficial.

Bibliography

- [ACP03] B. Allen, B. Curless, and Z. Popović. The space of human body shapes: reconstruction and parameterization from range scans. In *Proc. of ACM Special Interest Group on Graphics and Interactive Techniques (ACM SIGGRAPH)*, 2003.
- [BDGY04] M. Berar, M. Desvignes, Bailly G., and Payan Y. 3D meshes registration: Application to statistical skull model. In *Proc. of Image Analysis and Recognition*, 2004.
- [Ben91] C. Bennis. Piecewise surface flattening for non-distorted texture mapping. *Computer Graphics*, 25(4):237–246, 1991.
- [BG00] E.R. Baldock and J. Graham. *Image Processing and Analysis*. Oxford University Press, 2000.
- [BM92] P.J. Besl and N.D. McKay. A method for registration of 3-D shapes. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 14(2):239–256, 1992.
- [BSMG09] S. Benazzi, E. Stansfield, C. Milani, and G. Gruppioni. Geometric morphometric methods for three-dimensional virtual reconstruction of a fragmented cranium: The case of Angelo Poliziano. *Int. Journal on Legal Medicine*, 123(4):333–344, 2009.

- [BT00] A.D. Brett and C.J. Taylor. Automated construction of 3d shape models using harmonic maps. In *Proc. of Medical Image Understanding and Analysis*, 2000.
- [Che14] Y. Cheng. *Computer-aided craniomaxillofacial surgery planning for fractured skulls*. PhD thesis, National University of Singapore, School of Computing, 2014.
- [CLL12] Y. Cheng, W. K. Leow, and T. C. Lim. Automatic identification of frankfurt plane and mid-sagittal plane of skull. In *Proc. on IEEE Workshop on Apps of Computer Vision (WACV)*, 2012.
- [CR00] H. Chui and A. Rangarajan. A new algorithm for non-rigid point matching. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2000.
- [CTC⁺95] T. Cootes, C. Taylor, D. Cooper, J. Graham, et al. Active shape models - their training and application. *Computer vision and image understanding*, 61(1):38–59, 1995.
- [Dav02] R.H. Davies. Learning shape: optimal models for analysing natural variability. Technical report, University of Manchester, 2002.
- [DZS⁺11] Q. Deng, M. Zhou, W. Shui, Z. Wu, Y. Ji, and R. Bai. A novel skull registration based on global and local deformations for craniofacial reconstruction. *Forensic Science Int.*, 208:95–102, 2011.
- [ECP⁺06] D.M. Elena, J. Chapuis, I. Pappas, G. Ferrigno, W. Hallermann, A. Schramm, and M. Caversaccio. Automatic extraction of the mid-facial plane for cranio-maxillofacial surgery planning. *Int. Journal of Oral And Maxillofacial Surgery*, 35(7):636–642, 2006.

- [FH05] M.S. Floater and K. Hormann. Surface parameterization: a tutorial and survey. *Advances in Multiresolution for Geometric Modelling*, pages 157–186, 2005.
- [FL98] M. Fleute and S. Lavallée. Building a complete surface model from sparse data using statistical shape models: Application to computer assisted knee surgery. In *Proc. of Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, 1998.
- [GMBW04] P. Gunz, P. Mitteroecker, F. Bookstein, and G. Weber. Computer aided reconstruction of human crania. In *Proc. of Computer Applications and Quantitative Methods in Archaeology*, 2004.
- [Goo91] C. Goodall. Procrustes methods in the statistical analysis of shape. *Journal of the Royal Statistical Society*, 53:285–339, 1991.
- [Gow75] J.C. Gower. Generalized procrustes analysis. *Psychometrika*, 40:33–51, 1975.
- [GTH⁺03] R. Gassner, T. Tuli, O. Hachl, A. Rudisch, and H. Ulmer. Craniomaxillofacial trauma: a 10 year review of 9543 cases with 21067 injuries. *Journal of Cranio-maxillofacial Surgery*, 31(1):51–61, 2003.
- [Gun05] P. Gunz. *Statistical and Geometric Reconstruction of Hominid Crania: Reconstructing Australopithecine Ontogeny*. Computational Imaging and Vision. Universität Wien, 2005.
- [GWL99] H. K. Gumprecht, D. C. Widenka, and C. B. Lumenta. BrainLab VectorVision Neuronavigation System: technology and clinical experiences in 131 cases. *Neurosurgery*, 44(1), January 1999.
- [HBH03] T.J. Hutton, B.F. Buxton, and P. Hammond. Automated registration of 3D faces using dense surface models. In *Proc. of British Machine Vision Conference (BMVC)*, 2003.

- [HDR⁺13] E. Hwuang, S. Danish, M. Rusu, R. Sparks, R. Toth, and A Madabhushi. Anisotropic smoothing regularization (ansr) in thirion’s demons registration evaluates brain mri tissue changes post-laser ablation. In *Proc. of IEEE Conference on Engineering in Medicine and Biology Society (EMBC)*, 2013.
- [HDZ⁺12a] Y. Hu, F. Duan, M. Zhou, Y. Sun, and B. Yin. Craniofacial reconstruction based on a hierarchical dense deformable model. *EURASIP Journal on Advances in Signal Processing*, 2012:217, 2012.
- [HDZ⁺12b] Y. Hu, F. Duan, M. Zhou, Y. Sun, and B. Yin. Craniofacial reconstruction based on a hierarchical dense deformable model. *EURASIP Journal on Advances in Signal Processing*, 217:1–14, 2012.
- [HM09] T. Heimann and H.P. Meinzer. Statistical shape models for 3D medical image segmentation: A review. *Medical Image Analysis*, 13(4):543–563, 2009.
- [HW77] P.W. Holland and R.E. Welsch. Robust regression using iteratively reweighted least-squares. *Communications in Statistics – Theory and Methods*, 6(9):813–827, 1977.
- [Ini] Inion. Patient Information Guide to Inion CMF (Craniomaxillofacial) Surgery. <http://www.inion.com/patienteducation/CMFSurgery/>.
- [JEK05] K. Josephson, A. Ericsson, and J. Karlsson. Segmentation of medical images using three-dimensional active shape models. In *Proc. of Scandinavian Conference on Image Analysis*, 2005.
- [Jol02] I. T. Jolliffe. *Principal Component Analysis*. Springer, second edition, 2002.
- [KSG99] A. Kelemen, G. Szekely, and G. Gerig. Elastic model-based segmentation of 3-D neuroradiological data sets. *IEEE Trans. Medical Imaging*, 1999.
- [KT98] A.C.W. Kotcheff and C.J. Taylor. Automatic construction of eigenshapemod-els by direct optimization. *Medical Image Analysis*, 2(4):303 – 314, 1998.

- [LAV09] M. Lthi, T. Albrecht, and T. Vetter. Building shape models from lousy data. In *Proc. of Medical Image Computing and Computer-Assisted Intervention (MICCAI)*. 2009.
- [LCL⁺03] M.Y. Lee, C.C. Chang, C.C. Lin, L.J. Lo, and Y.R. Chen. T medical rapid prototyping in custom implant design for craniofacial reconstruction. *IEEE Trans. on Systems, Man and Cybernetics*, 3:2903–2908, 2003.
- [Lit] LittleBabyFaceFoundation. Facial deformities. <http://www.littlebabyface.org>.
- [LK00] C. Lorenz and N. Krahnstöver. Generation of point-based 3D shape models for anatomical objects. *Computer Vision and Image Understanding*, 77:175–191, 2000.
- [LN07] M. Louw and F. Nicolls. Improved sparse correspondence resolution using loopy belief propagation with mrf clique based structure preservation. In *Proc. of Int. Conf. on Visualization, Imaging and Image Processing (VIIP)*, 2007.
- [LP00] R. J. A. Lapeer and R. W. Prager. 3D shape recovery of a newborn skull using thin-plate splines. *Computerized Medical Imaging & Graphics*, 24(3):193–204, 2000.
- [MC03] J. Maciel and J.P. Costeira. A global solution to sparse correspondence problems. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 25(2):187–199, Feb 2003.
- [Neu] Neurosurgery. Skull fracture. <http://www.neurosurgery.sg/index.php?id=69>.
- [PHWZ10] G. Pan, S. Han, Z. Wu, and Y. Zhang. Removal of 3D facial expressions: A learning-based approach. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2010.

- [PLT07] J.M. Phillips, R. Liu, and C. Tomasi. Outlier robust ICP for minimizing fractional RMSD. In *Proc. of Int. Conf. on 3-D Digital Imaging and Modeling*, 2007.
- [RB02] A. Rosas and M. Bastir. Thin plate spline analysis of allometry and sexual dimorphism in the human craniofacial complex. *American J. Physical Anthropology*, 117:236–245, 2002.
- [RHB97] A. Rangarajan, Chui H., and F.L. Bookstein. The softassign procrustes matching algorithm. In *Proc. of Information Processing in Medical Imaging*, 1997.
- [RSS⁺01] K. Rohr, H. S. Stiehl, R. Sprenkel, T. M. Buzug, J. Weese, and M. H. Kuhn. Landmark-based elastic registration using approximating thin-plate splines. *IEEE Trans. on Med Imaging*, 20(6):526–534, June 2001.
- [SD05] M.B. Stegmann and Pedersen D. Bi-temporal 3D active appearance models with applications to unsupervised ejection fraction estimation. In *Proc. of SPIE Medical Imaging*, 2005.
- [SH] D. F. Siwek and R. J. Hoyt. Anatomy of the human skull. <http://skullanatomy.info>.
- [ST03] H. Seo and N.M. Thalmann. An automatic modeling of human bodies from sizing parameters. In *Proc. of ACM Special Interest Group on Graphics and Interactive Techniques (ACM SIGGRAPH)*, 2003.
- [STA98] G. Subsol, J.P. Thirion, and N. Ayache. A scheme for automatically building 3d morphometric anatomical atlases: Application to a skull atlas. *Medical Image Analysis*, 2(1):37–60, 1998.
- [STA99] G. Subsol, J.P. Thirion, and N. Ayache. Incorporating a statistically based shape model into a system for computer-assisted anterior cruciate ligament surgery. *Med. Image Anal*, 3(3):209–222, 1999.

- [SU05] A. Souza and J.K. Udupa. Automatic landmark selection for active shape models. In *Proc. of SPIE Medical Imaging*, 2005.
- [Tay01] K. T. Taylor. *Forensic Art and Illustration*. CRC, 2001.
- [TBK⁺05] W. D. Turner, R. E. Brown, T. P. Kelliher, P. H. Tu, M. A. Taister, and K. W. Miller. A novel method of automated skull registration for forensic facial approximation. *Forensic Science Int.*, 154:149–158, 2005.
- [Thi98] Image matching as a diffusion process: an analogy with maxwell’s demons. *Medical Image Analysis*, 2(3):243 – 260, 1998.
- [VDBA⁺04] F.M. Vos, P. W. De Bruin, J. G M Aubel, G.J. Streekstra, M. Maas, L.J. van Vliet, and A.M. Vossepoel. A statistical shape model without using landmarks. In *Proc. of Int. Conf. on Pattern Recognition (ICPR)*, 2004.
- [Wah90] G. Wahba. *Spline models for observational data*. CBMS-NSF Regional Conference Series in Applied Mathematics. 1990.
- [Wik] Wikipedia. Thin plate spline. http://en.wikipedia.org/wiki/Thin_plate_spline.
- [WPS00] Y. Wang, B. S. Peterson, and L. H. Staib. Shape-based 3D surface correspondence using geodesics and local geometry. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2000.
- [WYLL11] L. Wei, W. Yu, M. Li, and X. Li. Skull assembly and completion using template-based surface matching. In *Proc. of Int. Conference on 3D Imaging, Modeling, Processing, Visualization and Transmission (3DIMPVT)*, 2011.
- [ZCJX12] Z. Zeng, T.H. Chan, K. Jia, and D. Xu. Finding correspondence from multiple images via sparse and low-rank decomposition. In *Proc. of European Conference on Computer Vision (ECCV)*. 2012.

- [ZCL13] K. Zhang, Y. Cheng, and W.K. Leow. Dense correspondence of skull models by automatic detection of anatomical landmarks. In *Int. Conference on Computer Analysis of Images and Patterns (CAIP)*, 2013.
- [ZLES05] S. Zachow, H. Lamecker, B. Elsholtz, and M. Stiller. Reconstruction of mandibular dysplasia using a statistical 3D shape model. In *Computer Assisted Radiology and Surgery*, pages 1238–1243, 2005.
- [ZWY10] Y. Zeng, C. Wang, and Wang Y. Dense non-rigid surface registration using high-order graph matching. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2010.