DEVELOPMENT AND COMPARISON OF STRATEGIES FOR THE RECONSTRUCTION OF FULL AND PARTIAL SKULL GEOMETRIES

by

M.J.R. Schoeman

Submitted in partial fulfilment of the requirements for the degree

Master of Engineering (Mechanical Engineering)

in the

Department of Mechanical and Aeronautical Engineering
Faculty of Engineering, Built Environment and Information Technology
UNIVERSITY OF PRETORIA

2016

ABSTRACT

DEVELOPMENT AND COMPARISON OF STRATEGIES FOR THE RECONSTRUCTION OF FULL AND PARTIAL SKULL GEOMETRIES

by

M.J.R. Schoeman

Supervisor(s):	Prof. Schalk Kok
	Dr. Nico Wilke
Department:	Mechanical and Aeronautical Engineering
University:	University of Pretoria

This thesis presents the development and comparison of strategies for the reconstruction of full and partial surface mesh based skull geometries. The intended application is to aid the South African Police Service Victim Identification Centre (SAPS VIC) with forensics, specifically prediction of a mandible when only the cranium is available.

Various methods for the registration of surface meshes are outlined. A new non-rigid iterative closest point (NR-ICP) algorithm based on an adaptively refined least square Radial Basis Function (RBF) approximation of the forward and backward nearest neighbour correspondence is developed. The newly developed non-rigid registration strategy is demonstrated and characterised for various parameters using an artificial mandible dataset constructed through Monte-Carlo (MC) sampling of a quadratic displacement field. Various suitable parameters are shown to result in imperceptible visual registration differences, with the correspondence error mainly distributed in-surface.

Multivariate regression techniques suited to the application of geometry prediction are considered, specifically for cases where the data is expected to be multi-collinear and the number of variables are far greater than the number of observations. Two regression approaches based on spatial information are considered. The first is the classical use of Procrustes Analysis where the Cartesian coordinates are used directly for regression. The second is a new Euclidean distance based approach utilizing pair-wise distances to consistent reference points. The proposed regression methods' time-space scaling is investigated to limit system sizes that result in time tractable cross-validation and model comparison. Pre- and post-processing required for tractability considerations

are also developed for both approaches.

Proof of concept of the registration based prediction strategies are demonstrated. This is accomplished through the use of an artificial dataset with embedded covariance and the use of registration targets without point-wise correspondence. The registration based prediction strategy is shown to be capable of accurate predictions for data with strong underlying structure/covariance.

The proposed registration based prediction strategy is demonstrated on a real cranium and mandible dataset, where the mandible geometry is predicted from the cranium geometry. Marginal improvement over the geometric mean is obtained. Observation scaling suggests that model accuracy is improved for increased observations, which merits expanding the dataset.

The proposed registration strategy has the limitation that it is not capable of registration of significant partial/incomplete geometries. A new regression-registration hybrid strategy is developed for use with partial geometries, when a full dataset of the given geometry is available. The regression-registration hybrid strategy is demonstrated on a real mandible dataset and mandible fossil.

ACKNOWLEDGEMENTS

I would like to acknowledge the following persons and entities:

- S. Kok and D.N. Wilke for their supervision.
- The staff from the Anthropology Department at the University of Pretoria for the logistics management, digitizing and handling of the skulls.
- The staff at the South African National Centre for Radiography and Tomography for scanning of the Cranium-Mandible dataset.
- The Pretoria Bone Repository for use of the skulls in their collection.
- C. and T. Cronje for proof-reading.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

2D	Two-dimensional
3D	Three-dimensional
CT	Computed Tomography
CV	Cross-Validation
CPU	Central Processing Unit
GPA	Generalised Procrustes Analysis
ICP	Iterative Closest Point
LMO	Leave-m-out
LOO	Leave-one-out
MC	Monte Carlo
MMDS	Metric Multidimensional Scaling
NR-ICP	Non-rigid Iterative Closest Point
PC	Principal Component
PCA	Principal Component Analysis
PLS	Partial Least Squares
RBF	Radial Basis Function
R-ICP	Rigid Iterative Closest Point
SAPS	South African Police Service
SPC	Superset Principal Component
SVD	Singular Value Decomposition
VIC	Victim Identification Centre

NOTATION

The notation in this study is given as follows:

1. Bold capital letters denote a matrix i.e. $\boldsymbol{X}$,

2. bold lower-case letters denote a vector i.e. $\boldsymbol{x}$ and

3. all non-bold letters denote scalars i.e. x_{ij} denotes the scalar in the i^{th} row and j^{th} column of matrix $\boldsymbol{X}$.

CHAPTER 1

INTRODUCTION

1.1 BACKGROUND

The goal of this thesis is to present the development and comparison of strategies for the registration and prediction of full and partial geometries, represented by triangulated surface meshes. The South African Police Service Victim Identification Centre (SAPS VIC) presented a need for predicting missing geometry to aid in forensics, specifically the prediction of a mandible when only the cranium is available. The SAPS VIC already acquired a 3D surface scanner for forensic facial reconstruction but do not have access to medical imaging modalities, which prompted the specific development of strategies for registration of triangular surface meshes.

Triangular surface mesh representations of geometries generated through surface scanning are generated arbitrarily. The mesh topology varies significantly between scans. Different scans of the same geometry will thus not have correspondence between meshes. The method of generating correspondence, which is a requirement for regression/prediction, is defined as registration. Therefore the thesis starts by developing and implementing a registration method, followed by regression methods which facilitate prediction of missing geometry.

Other possible applications for this work include:

1. Forensics identification, where a face (the missing part) can be statistically reconstructed based on skull shape (the existing part) [1].

2. Human anthropology, especially where the relationship between cranial and mandibular covariation is studied [2].

3. Palaeoanthropology where researchers try to reconstruct an incomplete fossil [3, 4].

The work in this study makes part of an inter-disciplinary project. The scope of work presented in this study is limited to the numerical modelling component. Work outside this scope was conducted by the University of Pretoria Department of Anthropology and the SAPS VIC, who can be contacted for any queries regarding these components.

1.2 OUTLINE OF THESIS

Chapter 2: Registration Methods

Various registration methods used to generate heuristic correspondence between similar geometries are documented. Registration is accomplished by transforming a reference geometry to represent a target geometry, resulting in a mapping with approximated correspondence. Each registration method is documented and demonstrated by means of a simplified example.

Chapter 3: Registration Application

A non-rigid iterative closest point (NR-ICP) registration strategy based on the registration methods in Chapter 2 is proposed. The non-rigid registration strategy is demonstrated and characterised for various parameters using an artificial mandible dataset constructed through Monte-Carlo (MC) sampling of a quadratic displacement field. Various suitable parameters are shown to result in imperceptible visual registration differences, with the correspondence error mainly distributed in-surface.

Chapter 4: Regression Methods

Multivariate regression techniques suited to the application of geometry prediction are considered. This requires methods that can handle multi-collinear data and where the number of predictor and response variables are many more than the number of observations. Pre- and post-processing, to facilitate regression for the purpose of prediction, of the registered geometry point-sets are proposed and discussed. The proposed regression methods' time-space scaling is investigated to limit system sizes that result in time tractable cross-validation and model comparison.

Chapter 5: Regression Application

Proof of concept of the registration based prediction strategy is demonstrated. An artificial dataset with embedded covariance is generated, registration targets without correspondence are simulated and the prediction error due to the heuristically generated correspondence and regression space approaches/models are evaluated and discussed.

Chapter 6: VIC Cranium-Mandible Dataset Application

The proposed registration based prediction strategy is demonstrated on a real cranium and mandible dataset. The objective is to predict a mandible when only a cranium is made available. This is a specific need of the South African Police Service Victim Identification Centre (SAPS VIC), who has a growing archive of cases with craniums without matching mandibles, rendering them unable to perform forensic facial reconstructions. Feature-cropping, a method proposed for improved prediction results, is introduced and applied. Observation scaling is performed to estimate model improvement for expansion of the current dataset.

Chapter 7: Regression-Registration Hybrid

The limitation of the proposed registration strategy, to only be capable of registration of full/complete geometries, is discussed. A new regression based registration and prediction strategy, for when a fully registered dataset of the geometry is available, is also proposed and demonstrated. This strategy is shown to be capable of registration of partial geometries.

Chapter 8: Conclusion

Final remarks and possible future work are discussed.

CHAPTER 2

REGISTRATION METHODS

2.1 CORRESPONDENCE

The ultimate purpose of this dissertation is to develop and compare strategies for the reconstruction of full and partial skull geometries. These geometries are expressed as triangular surface meshes, each composed of a point-set of vertices and a list of triangle connectivity. The goal is to predict a response point-set $\mathcal{P}_Y$ from a prediction point-set $\mathcal{P}_X$. Computationally this requires the point-sets to have one-to-one correspondence between any 2 observations. Triangular surface meshes used to represent each of the scanned geometries do not satisfy this condition. This is because the meshes generated through 3D surface scanning are arbitrary, with varying number of vertices and connectivity. The mesh registration methods discussed in this chapter are heuristic techniques for generating one-to-one correspondence. This is because no true one-to-one correspondence exists and approaches are largely based on the assumption of an appropriate mappings. This process comprises of transforming a source mesh, with a fixed number of vertices and known connectivity, to fit a target geometry. This approximated representation with consistent one-to-one correspondence allows for linear algebraic manipulation enabling the use of regression for the purpose of prediction.

2.2 NOTATION

The notation for point-sets of triangular meshes is described in this section. A point-set is given by

$$\mathcal{P} \triangleq \{\boldsymbol{p}_i\}_{i=1}^{N}, \quad \boldsymbol{p}_i \in \mathbb{R}^m. \tag{2.1}$$

The meaning of subscripts are explained in Table 2.1.

Table 2.1: Point set subscript notation in order of appearance.

Designation	Description
T or S	The designated point-set - Target or Source
$A, B, C, ..$	Different mesh representations of the designated point-set
J	The J'th transformation or approximation of the designated point-set
n	Direction/coordinate component n
i or $c(i)$	The i'th or mapped i'th point of the designated point-set
$(\quad)$	The designated point-set at the iteration number given in brackets

2.3 RIGID POINT SET REGISTRATION

This section describes various methods for the rigid registration of point-sets, which align point-sets using pure translation and rotation transformations. This section considers source and target point-sets given by $\mathcal{P}_S \triangleq \{\boldsymbol{p}_{Si}\}_{i=1}^{N_S}$ and $\mathcal{P}_T \triangleq \{\boldsymbol{p}_{Ti}\}_{i=1}^{N_T}$ respectively.

2.3.1 Principal Component Analysis (PCA) Based Registration

PCA registration consists of translating and rotating a source point-set such that it's principal axes and centroid are aligned with that of a target point-set. This method does not require correspondence between point-sets and is thus often used as a pre-alignment of similar geometries when correspondence is not known. The procedure for PCA registration is set out as follows [5]:

1. Consider a source and target set of points $\mathcal{P}_S$ and $\mathcal{P}_T$ with centroid location $\boldsymbol{c}_S$ and $\boldsymbol{c}_T$ respectively.

2. Perform principal component analysis of both $\mathcal{P}_S$ and $\mathcal{P}_T$ to determine the rotation matrix $\boldsymbol{R}$ that aligns the PCA axes as follows:

   (a) Construct $\boldsymbol{P}_S$ whose i'th column is given by $\boldsymbol{p}_{Si} - \boldsymbol{c}_S$.

   (b) Construct $\boldsymbol{P}_T$ whose i'th column is given by $\boldsymbol{p}_{Ti} - \boldsymbol{c}_T$.

   (c) Evaluate the covariance matrices $\boldsymbol{M}_S$ and $\boldsymbol{M}_T$, given by

$$\boldsymbol{M}_S = \boldsymbol{P}_S \boldsymbol{P}_S^T \quad \text{and} \quad \boldsymbol{M}_T = \boldsymbol{P}_T \boldsymbol{P}_T^T. \tag{2.2}$$

(d) Eigenvalue decompositions of the covariance matrices $\boldsymbol{M}_S$ and $\boldsymbol{M}_T$ are performed. The eigenvectors represent principal directions of variance. The eigenvalues indicate the variance along each eigenvector. The loadings $\boldsymbol{L}_S$ and $\boldsymbol{L}_T$ are matrices whose columns are the eigenvectors of the respective source and target point-set covariance matrices.

(e) The rotation matrix $\boldsymbol{R}$ which aligns two sets of PCA axes is evaluated as follows:

$$\begin{aligned} \boldsymbol{R}\boldsymbol{L}_S &= \boldsymbol{L}_T \\ \boldsymbol{R} &= \boldsymbol{L}_T\boldsymbol{L}_S^{-1}. \end{aligned} \tag{2.3}$$

By noting that the eigenvector matrices $\boldsymbol{L}_S$ and $\boldsymbol{L}_T$ are orthonormal:

$$\boldsymbol{R} = \boldsymbol{L}_T\boldsymbol{L}_S^{T}. \tag{2.4}$$

3. The aligned source point-set is evaluated as follows:

$$\boldsymbol{p}'_{Si} = \boldsymbol{c}_T + \boldsymbol{R}(\boldsymbol{p}_{Si} - \boldsymbol{c}_S). \tag{2.5}$$

PCA registration is thus only suitable for geometries of which

1. the principal directions of variation are expected to be consistent,

2. the eigenvalues (amount of variation along each principal direction) are not similar (which may lead to inconsistent alignment) and

3. noise does not affect the centroid or principal axes significantly.

Figure 2.1 illustrates an example of 2D source and target point-sets and the corresponding PCA registration. These point-sets do not have point correspondence as source and target points sets are composed of $N_S = 50$ and $N_T = 25$ points respectively.

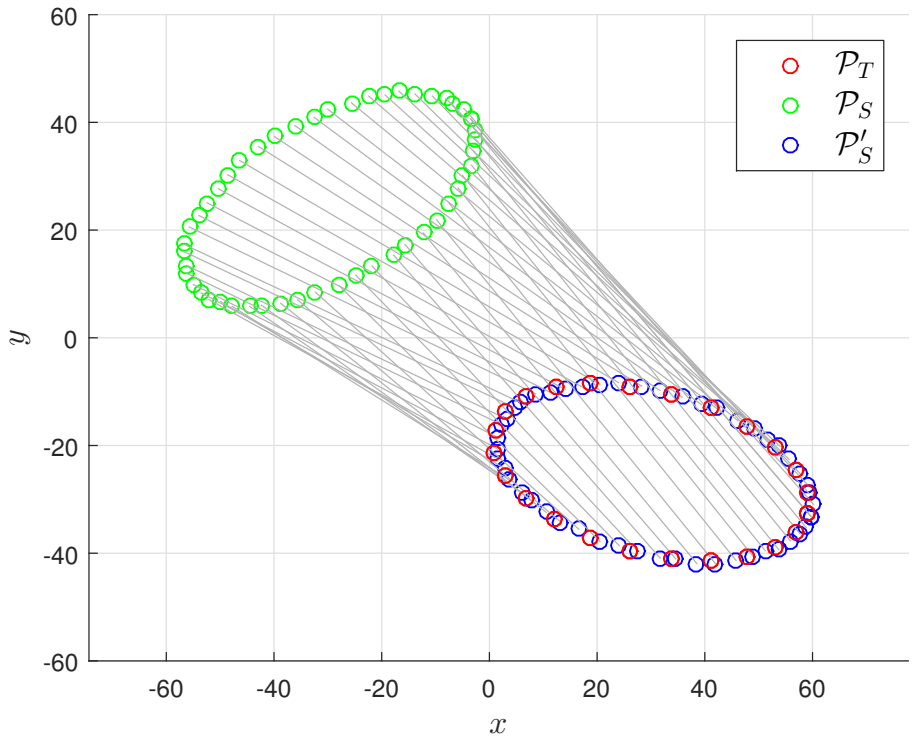


Figure 2.1: Source, target and PCA aligned point-sets.

2.3.2 Singular Value Decomposition (SVD) Based Registration

SVD based registration consists of finding the optimal alignment when correspondence between the source and target point-sets is known. After centroid alignment and rotation by the rotation matrix $\boldsymbol{R}$, a transformed point-set is given by

$$\boldsymbol{p}'_{Si} = \boldsymbol{c}_T + \boldsymbol{R}(\boldsymbol{p}_{Si} - \boldsymbol{c}_S). \tag{2.6}$$

SVD registration is based on finding the rotation matrix $\boldsymbol{R}$ that minimizes the sum of the pairwise distances:

$$\begin{aligned} \boldsymbol{R} &= \arg\min_{\boldsymbol{R}} \left(\sum_{i=1}^{N_S} \|\boldsymbol{p}'_{Si} - \boldsymbol{p}_{Ti}\|_2^2 \right) \\ &= \arg\min_{\boldsymbol{R}} \left(\sum_{i=1}^{N_S} \|\boldsymbol{c}_T + \boldsymbol{R}(\boldsymbol{p}_{Si} - \boldsymbol{c}_S) - \boldsymbol{p}_{Ti}\|_2^2 \right), \end{aligned} \tag{2.7}$$

with

$$\boldsymbol{N}_S = \boldsymbol{N}_T, \quad \boldsymbol{R}^T\boldsymbol{R} = \boldsymbol{I}_m, \quad \det(\boldsymbol{R}) = 1, \quad \boldsymbol{R} \in \mathbb{R}^{m \times m}.$$

An analytical solution for finding $\boldsymbol{R}$ is possible through an equivalent formulation using the cross-covariance called the orthogonal Procrustes problem [6]:

1. Let $\boldsymbol{P}_S$ be a matrix whose $i-$th column is the vector $\boldsymbol{p}_{Si} - \boldsymbol{c}_S$

2. Let $\boldsymbol{P}_T$ be a matrix whose $i-$th column is the vector $\boldsymbol{p}_{Ti} - \boldsymbol{c}_T$

3. The cross-covariance is given by

$$\boldsymbol{M} = \boldsymbol{P}_S \boldsymbol{P}_T^T. \tag{2.8}$$

4. The equivalence is finding the orthonormal matrix $\boldsymbol{R}$ that maximizes the trace

$$\text{tr}(\boldsymbol{R}\boldsymbol{M}). \tag{2.9}$$

5. The solution is given by

$$\boldsymbol{R} = \boldsymbol{V}\boldsymbol{U}^T, \tag{2.10}$$

where $\boldsymbol{U}$ and $\boldsymbol{V}$ are obtained through the SVD of the cross-covariance $\boldsymbol{M}$

$$\boldsymbol{M} = \boldsymbol{U}\boldsymbol{W}\boldsymbol{V}^T, \tag{2.11}$$

where

- $\boldsymbol{U}$ and $\boldsymbol{V}$ are orthonormal matrices (i.e. rotations) and
- $\boldsymbol{W}$ is diagonal matrix of singular values.

Figure 2.2 illustrates an example of 2D source and target point-sets with the corresponding SVD registration.

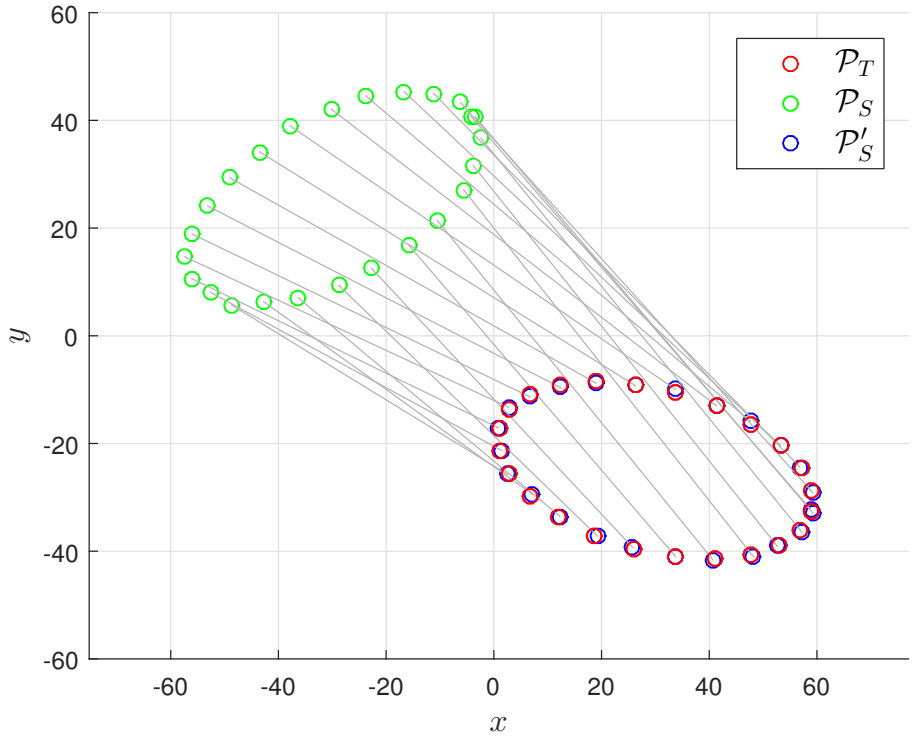


Figure 2.2: Source, target and SVD aligned point-sets.

2.3.3 Rigid Iterative Closest Point (R-ICP) Registration

The R-ICP registration algorithm, by Besl and Mckay [7], approaches registration as a least squares problem. The general ICP algorithm seeks the transformation $\boldsymbol{T}$ that best aligns source point-set $\mathcal{P}_S$ to a target point-set $\mathcal{P}_T$ in $\mathbb{R}^m$:

$$\min_{\boldsymbol{T},j\in\{1,2,\ldots,N_T\}} \left(\sum_{i=1}^{N_S} \|\boldsymbol{T}(\boldsymbol{p}_{Si}) - \boldsymbol{p}_{Tj}\|_2^2 \right). \tag{2.12}$$

The R-ICP algorithm seeks the transformation $\boldsymbol{T}$ that consists of a rotation matrix $\boldsymbol{R}$ and translation vector $\boldsymbol{t}$:

$$\min_{\boldsymbol{R},\boldsymbol{t},j\in\{1,2,\ldots,N_T\}} \left(\sum_{i=1}^{N_S} \|(\boldsymbol{R}\boldsymbol{p}_{Si}) + \boldsymbol{t} - \boldsymbol{p}_{Tj}\|_2^2 \right), \tag{2.13}$$

with

$$\boldsymbol{R}^T\boldsymbol{R} = \boldsymbol{I}_m, \quad \det(\boldsymbol{R}) = 1, \quad \boldsymbol{R} \in \mathbb{R}^{m\times m} \quad \text{and} \quad \boldsymbol{t} \in \mathbb{R}^m.$$

The R-ICP procedure performs two steps iteratively:

1. Correspondence is built up between the source and target point-set, with correspondence for point $\boldsymbol{p}_{Si(k-1)}$ given by:

$$c_k(i) = \underset{j\in\{1,2,\ldots,N_T\}}{\arg\min} \left(\left\| (\boldsymbol{p}_{Si(k-1)} - \boldsymbol{p}_{Tj}) \right\|_2^2 \right), \quad i = 1, 2, \ldots, N_S. \tag{2.14}$$

   This may be done using a $k-d$ tree nearest neighbour search algorithm.

2. The rotation matrix $\boldsymbol{R}$ and translation vector $\boldsymbol{t}$ are determined from

$$(\boldsymbol{R}_k, \boldsymbol{t}_k) = \underset{\boldsymbol{R},\boldsymbol{t}}{\arg\min} \left(\sum_{i=1}^{N_S} \left\| \boldsymbol{R}\boldsymbol{p}_{Si(k-1)} + \boldsymbol{t} - \boldsymbol{p}_{Tc_k(i)} \right\|_2^2 \right), \tag{2.15}$$

   with

$$\boldsymbol{R}^T\boldsymbol{R} = \boldsymbol{I}_m, \quad \det(\boldsymbol{R}) = 1, \quad \boldsymbol{R} \in \mathbb{R}^{m\times m} \quad \text{and} \quad \boldsymbol{t} \in \mathbb{R}^m,$$

   and used to update $\mathcal{P}_S$ before performing another iteration

$$\boldsymbol{p}_{Si(k)} = \boldsymbol{R}_k\boldsymbol{p}_{Si(k-1)} + \boldsymbol{t}_k. \tag{2.16}$$

Shaoyi et al. [8] showed that the translation vector in Equations (2.15) and (2.16) can be eliminated:

Theorem 2.1 *Given two point-sets* $\{\boldsymbol{p}_{Si}\}_{i=1}^N$ *and* $\{\boldsymbol{p}_{Ti}\}_{i=1}^N$ *the function* $F(\boldsymbol{t}) = \sum_{i=1}^N \|\boldsymbol{p}_{Si} + \boldsymbol{t} - \boldsymbol{p}_{Ti}\|_2^2$ *has the minimum when* $\boldsymbol{t} = \frac{1}{N}\sum_{i=1}^N \boldsymbol{p}_{Ti} - \frac{1}{N}\sum_{i=1}^N \boldsymbol{p}_{Si}$.

Proof 2.1 *The minimum of* $F(\boldsymbol{t})$ *must satisfy* $\frac{dF(\boldsymbol{t})}{d\boldsymbol{t}} = 2\sum_{i=1}^N (\boldsymbol{p}_{Si} + \boldsymbol{t} - \boldsymbol{p}_{Ti}) = 0$, *from which* $\boldsymbol{t} = \frac{1}{N}\sum_{i=1}^N \boldsymbol{p}_{Ti} - \frac{1}{N}\sum_{i=1}^N \boldsymbol{p}_{Si} = \boldsymbol{c}_T - \boldsymbol{c}_S$.

This simplifies Equation (2.15) to

$$\begin{aligned} \boldsymbol{R}_k &= \underset{\boldsymbol{R}}{\arg\min} \left(\sum_{i=1}^{N_S} \left\| \boldsymbol{R}\boldsymbol{p}_{Si(k-1)} + \boldsymbol{c}_T - \boldsymbol{c}_{S(k)} - \boldsymbol{p}_{Tc_k(i)} \right\|_2^2 \right) \\ &= \underset{\boldsymbol{R}}{\arg\min} \left(\sum_{i=1}^{N_S} \| \boldsymbol{c}_T + \boldsymbol{R}(\boldsymbol{p}_{Si(k-1)} - \boldsymbol{c}_{S(k-1)}) - \boldsymbol{p}_{Tc_k(i)} \|^2 \right), \end{aligned} \tag{2.17}$$

with

$$\boldsymbol{R}^T\boldsymbol{R} = \boldsymbol{I}_m, \quad \det(\boldsymbol{R}) = 1, \quad \boldsymbol{R} \in \mathbb{R}^{m\times m},$$

which is identical to the orthogonal Procrustes problem in Equation (2.7), giving an analytical solution to the k-th rotation matrix $\boldsymbol{R}_k$ as described in Section 2.3.2.

The use of this algorithm for 20 iterations is demonstrated in Figure 2.3 for a data set without correspondence. The corresponding sum of squared errors (SSE), as given by

$$\varepsilon = \sum_{i=1}^{N_S} \|\boldsymbol{p}_{Si(k)} - \boldsymbol{p}_{Tc_k(i)}\|^2, \tag{2.18}$$

is shown in Figure 2.4. These point-sets do not have point correspondence as source and target points sets are composed of $N_S = 50$ and $N_T = 25$ points respectively.

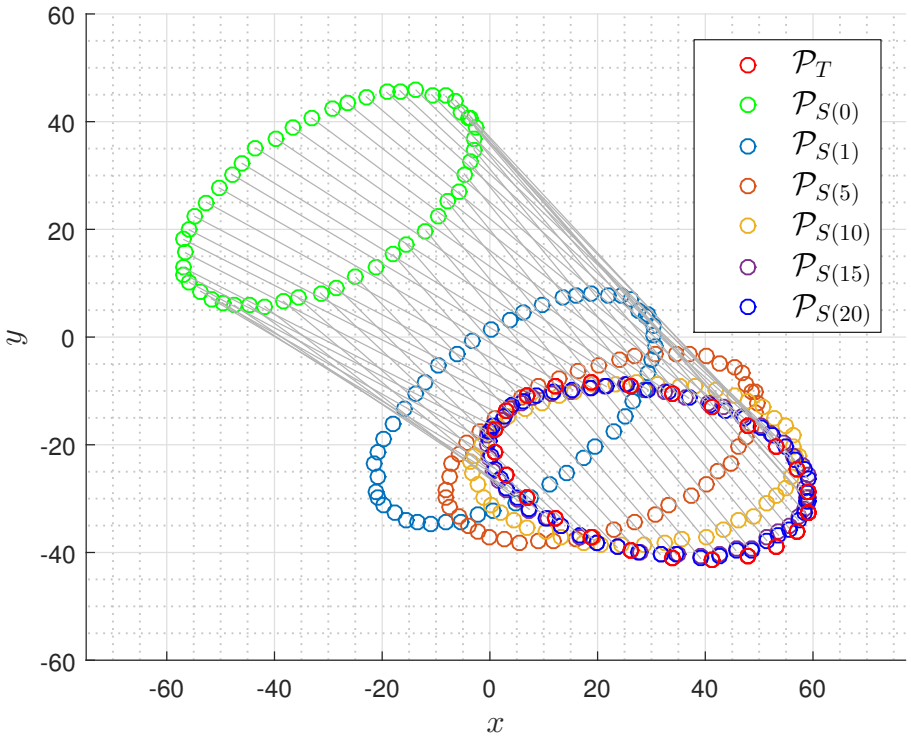


Figure 2.3: Source, target and R-ICP registered points set at iteration number k.

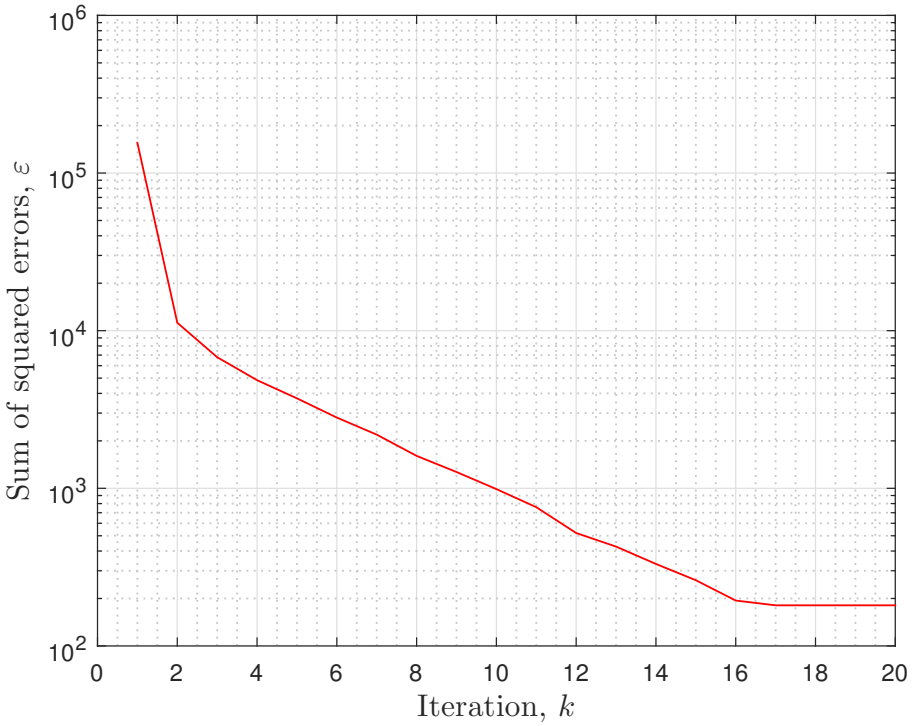


Figure 2.4: Sum of squared errors vs. iteration number.

2.4 NON-RIGID POINT SET REGISTRATION

This section describes methods for non-rigid registration of meshes. The first method entails a mesh movement scheme based on Radial Basis Function (RBF) interpolation between points with known correspondence. The second method consists of a Non-rigid iterative closest point registration (NR-ICP) algorithm based on an iterative least squares RBF estimation of the displacement field.

2.4.1 Radial Basis Function (RBF) Interpolation

RBF interpolation is a technique used for the interpolation of scalar fields of which the values are known at discrete locations [9]. In this work it is used to create an interpolated displacement vector field, from the known displacement at discrete points with known correspondence between the source and target meshes.

The interpolation function to describe a component of the displacement vector field is

approximated by the sum of basis functions:

$$S(\boldsymbol{p}) = \sum_{i=1}^{N_b} \alpha_i \phi\left(\|\boldsymbol{p} - \boldsymbol{p}_{bi}\|\right) + \mathbb{P}(\boldsymbol{p}), \tag{2.19}$$

where $\boldsymbol{p}_{b_i}$ are points with known displacement (centers), $\mathbb{P}(\boldsymbol{p})$ is a polynomial and N_b is the number of points with known field values. The function $\phi(d)$ is a basis function with respect to the Euclidean norm, d.

The minimal degree of polynomial $\mathbb{P}(\boldsymbol{p})$ depends on the choice of basis function. The use of conditionally positive definite basis functions of order $m \leq 2$ allows the use of a linear polynomial, which has the added advantage that rigid body translations are exactly recovered [9]. This is not an essential feature as rigid-alignment strategies (translation and rotation transformations) are often employed before non-rigid registration.

The RBF and linear polynomial coefficients, $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, are determined by solving the system

$$\begin{bmatrix} \mathbf{M}_{b,b} & \mathbf{P}_b \\ \mathbf{P}_b^T & 0 \end{bmatrix} \begin{Bmatrix} \boldsymbol{\alpha} \\ \boldsymbol{\beta} \end{Bmatrix} = \begin{Bmatrix} \mathbf{d}_b \\ 0 \end{Bmatrix}, \tag{2.20}$$

where $\mathbf{d_b}$ is the known displacements and $\mathbf{M}_{b,b}$ is a $N_b \times N_b$ matrix containing the evaluations of the basis function

$$\phi_{b_i b_j} = \phi\left(\|\boldsymbol{p}_{bi} - \boldsymbol{p}_{bj}\|\right). \tag{2.21}$$

$\mathbf{P}_b$ is an $N_b \times 4$ matrix with row i given by $\mathbf{P}_{b_i} = \{1 \quad x_{b_i} \quad y_{b_i} \quad z_{b_i}\}$. It is also seen in Equation (2.20) that the extra degrees of freedom are taken up by moment conditions on $\boldsymbol{\alpha}$, which is a regularization that guarantees unique coefficients $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ [10]. With the coefficients solved, displacements at unprescribed locations are evaluated using the interpolation function in Equation (2.19).

Various radial basis functions are available and can be divided into functions with global support and functions with local support. Local or compact support functions are scaled with a chosen support radius r such that $\xi = d/r$. The original basis function $\phi(d)$ is then replaced with

$$\phi_r = \begin{cases} \phi(\xi) & \text{if } 0 \leq \xi \leq 1, \\ 0 & \text{if } \xi > 1. \end{cases} \tag{2.22}$$

This has the effect of only displacing coordinates in a confined radius about each

center $\boldsymbol{p}_b$. Global support functions on the other hand cover the entire interpolation space, and often have a parameter ε, which controls the overall shape/bandwidth of the function.

Global support functions, and larger radii for local support functions, generally result in more accurate results [9]. This results in denser matrix systems which are solved less efficiently. (Local support functions are generally required for very large system sizes to result in sparse matrix systems which are efficient to solve.) An example of this is in the Finite Element Method where meshes with a large number of small elements, i.e. many points, are used for improved accuracy. The surface meshes used for this study are coarse in comparison, with fewer prescribed displacements. The computational overhead is thus small enough to allow the use of global support functions.

Various examples of local and global support functions are given in Table 2.2. The radial basis functions with global support in Table 2.2 are demonstrated on a dolphin mesh obtained from the Inria model repository [11] in Figure 2.5. Here the tips of the dolphin's nose and tail were displaced vertically by an arbitrary amount, and zero displacements prescribed at the tips of it's dorsal and pectoral fins. This clearly demonstrates how each basis function results in a unique displacement field.

Table 2.2: Various basis functions with local support $\phi(\xi)$ and global support $\phi(d)$.

Name	Function
CP C^0	$(1-\xi)^2$
CP C^2	$(1-\xi)^4(4\xi+1)$
CTPS C^0	$(1-\xi))^5$
CTPS C^1	$1+\frac{80}{3}\xi^2-40\xi^3+15\xi^4-\frac{8}{3}\xi^5+20\xi^2\log(\xi)$
Linear	d
Cubic	d^3
Thin plate spline (TPS)	$d^2\log(d)$
Multi-quadratic bi-harmonic (MQB)	$\sqrt{\varepsilon^2+d^2}$
Inverse multi-quadratic bi-harmonic (IMQB)	$(\varepsilon^2+d^2)^{-\frac{1}{2}}$
Gaussian (G)	$e^{-(\varepsilon d)^2}$

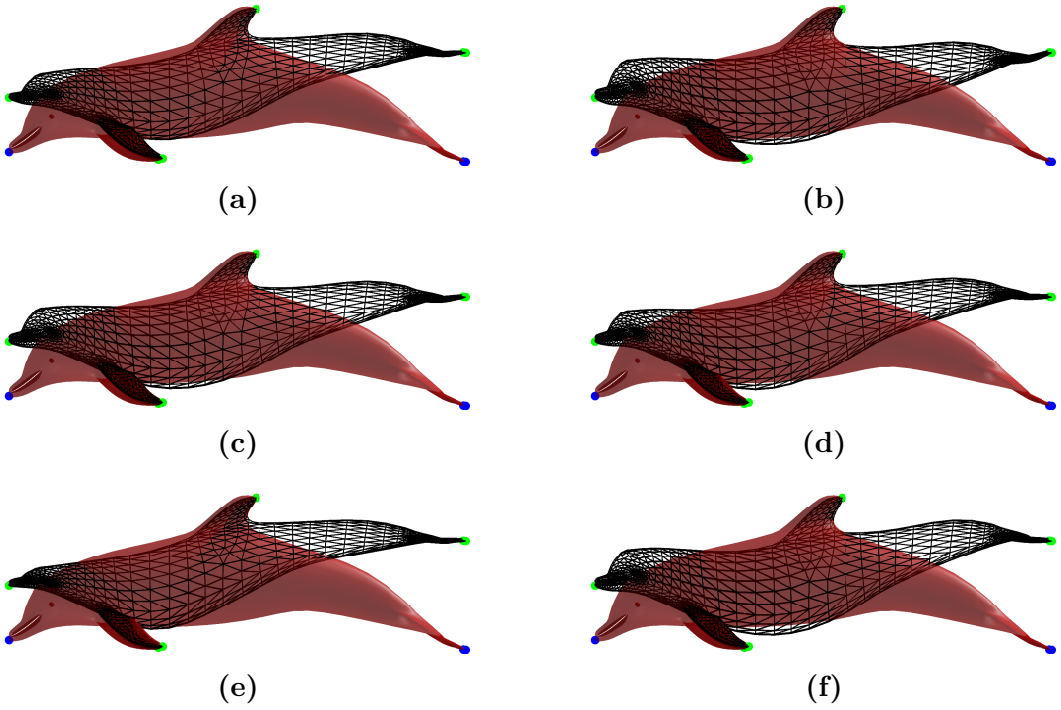


Figure 2.5: Plots of various RBF functions: (a) Linear, (b) Cubic, (c) TPS, (d) MQB with $\varepsilon = 100$, (e) IMQB with $\varepsilon = 100$ and (f) G with $\varepsilon = 0.0025$. The deformed mesh is displayed by the black wire-frame, with the original mesh given by the semi-opaque red surface. The scattered points reflect prescribed displacements at the nose and fin tips.

2.4.2 Non-Rigid Iterative Closest Point (NR-ICP) Registration

Various non-rigid registration algorithms exist [12, 13, 14, 15]. These all differ based on the assumption of transformations that result in appropriate mappings. A validation approach which results in a proof of concept is followed in this study. This allowed for the development of a validated new generic fit-for-purpose registration method similar to a well-known point matching algorithm [12]. This avoids having to compare many different methods/implementations, which would be intractable for the purpose this study. The new Non-Rigid Iterative Closest Point (NR-ICP) method developed in this study regresses the RBF displacement field $\mathcal{U}$ by solving a least squares problem and resolves correspondence between the source and target point-sets iteratively. It is important to note that the flexibility of the regressed RBF is adaptively enhanced. This approach and encompasses evaluating the following steps iteratively:

1. Correspondence from the source to target point-set is built up, with correspondence for point $\boldsymbol{p}_{Si(k-1)}$ given by:

$$c_{Sk}(i) = \underset{j\in\{1,2,\ldots,N_T\}}{\arg\min} \left(\left\| (\boldsymbol{p}'_{Si(k-1)} - \boldsymbol{p}_{Tj}) \right\|_2^2 \right), \quad i = 1, 2, \ldots, N_S. \tag{2.23}$$

2. Correspondence from the target to source point-set is built up, with correspondence for point $\boldsymbol{p}_{Ti(k-1)}$ given by:

$$c_{Tk}(i) = \underset{j\in\{1,2,\ldots,N_S\}}{\arg\min} \left(\left\| (\boldsymbol{p}_{Ti} - \boldsymbol{p}'_{Sj(k-1)}) \right\|_2^2 \right), \quad i = 1, 2, \ldots, N_T. \tag{2.24}$$

3. A number, N_b, of uniformly sampled centres $\mathcal{P}_{Sb}$ are uniformly selected out of the source point-set $\mathcal{P}_S$. This can be accomplished through Poisson Disk sampling. These control the number of fitting coefficients and thus the complexity of the RBF regression surface. Choosing a small number initially results in a displacement field that captures global shape changes (lower complexity fitting), with the number of centres increased with iterations to allow for better local fitting (higher complexity fitting).

4. The RBF coefficients are evaluated by solving the following overdetermined system:

$$\begin{bmatrix} \mathbf{M}_{S,Sb} \\ \mathbf{M}_{TS,Sb} \end{bmatrix} \left\{ \boldsymbol{\alpha}_n \right\} = \begin{Bmatrix} \mathbf{u}_{Sn} \\ -\mathbf{u}_{Tn} \end{Bmatrix}, \tag{2.25}$$

with:

- $\mathbf{M}_{S,Sb}$ a $N_S \times N_b$ matrix that contains the basis function evaluations of the source mesh points w.r.t. the RBF centres $\mathcal{P}_{Sb}$,

$$\phi_{S_iSb_j} = \phi\left(\|\boldsymbol{p}_{Si} - \boldsymbol{p}_{Sbj}\| \right), \tag{2.26}$$

with

$$i = 1, 2, \ldots, N_S; \quad j = 1, 2, \ldots, N_b.$$

- $\mathbf{M}_{TS,Sb}$ a $N_T \times N_b$ matrix that contains the basis function evaluations of the source mesh correspondence of the target mesh point-set w.r.t. the RBF centres $\mathcal{P}_{Sb}$

$$\phi_{S_{c_T i}Sb_j} = \phi\left(\|\boldsymbol{p}_{Sc_{Tk}(i)} - \boldsymbol{p}_{Sbj}\| \right), \tag{2.27}$$

with

$$i = 1, 2, \ldots, N_T; \quad j = 1, 2, \ldots, N_b.$$

- $\boldsymbol{\alpha}_n$ the fit coefficient vector for dimension/direction n
- $\mathbf{u}_{Sn}$ the source to target mesh correspondence displacement field vector for dimension/direction n evaluated as

$$\mathbf{u}_{Sn} = \boldsymbol{p}_{Tnc_{Sk}} - \boldsymbol{p}_{Sn}. \tag{2.28}$$

- $\mathbf{u}_{Tn}$ the target to source mesh correspondence displacement field vector for dimension/direction n evaluated as

$$\mathbf{u}_{Tn} = \boldsymbol{p}_{Snc_{Tk}} - \boldsymbol{p}_{Tn}. \tag{2.29}$$

5. The displacement field estimation is then used to update the registered mesh as follows

$$\mathcal{P}'_{S(k)} = \mathcal{P}_S + \begin{bmatrix} \mathbf{M}_{S,Sb}\boldsymbol{\alpha}_1 & \ldots & \mathbf{M}_{S,Sb}\boldsymbol{\alpha}_n \end{bmatrix}. \tag{2.30}$$

6. The iterative method is stopped when the number of iterations, k, reaches the specified maximum iterations $k_{\max}$ or when the sum of point wise distances of the registered mesh reaches a specified convergence tolerance

$$\sum_{i=1}^{N_S} \left\| \boldsymbol{p}'_{Si(k)} - \boldsymbol{p}'_{Si(k-1)} \right\|_2^2 < \varepsilon_{\text{tol}}. \tag{2.31}$$

7. Correspondence with obtuse normals may be removed from the approximation by performing a dot product of the normals for the forward (source to target) and backward (target to source) correspondences at each iteration. The row entries may then be removed from Equation (2.25) where the angular difference is more than a specified threshold θ_{thresh}.

This method is applied here in 2D to register two different parametrised asteroid curves. The parametric equations for the asteroid geometries are given by

$$\begin{aligned} X &= (\cos t)^3 \,, \\ Y &= (\sin t)^3 \,, \end{aligned} \tag{2.32}$$

with source and target geometries discritised using 1000 points for $0 \leq t < 2\pi$. The target geometry is obtained by applying a linear deformation to the source geometry. The source and target point-sets are given by the following:

$$\begin{aligned} \mathcal{P}_S &= \begin{bmatrix} X & Y \end{bmatrix} , \\ \mathcal{P}_T &= \begin{bmatrix} X & Y \end{bmatrix} \begin{bmatrix} 1.2 & 0.1 \\ 0.1 & 0.9 \end{bmatrix} . \end{aligned} \tag{2.33}$$

The resulting source and target geometries are presented in Figure 2.6.

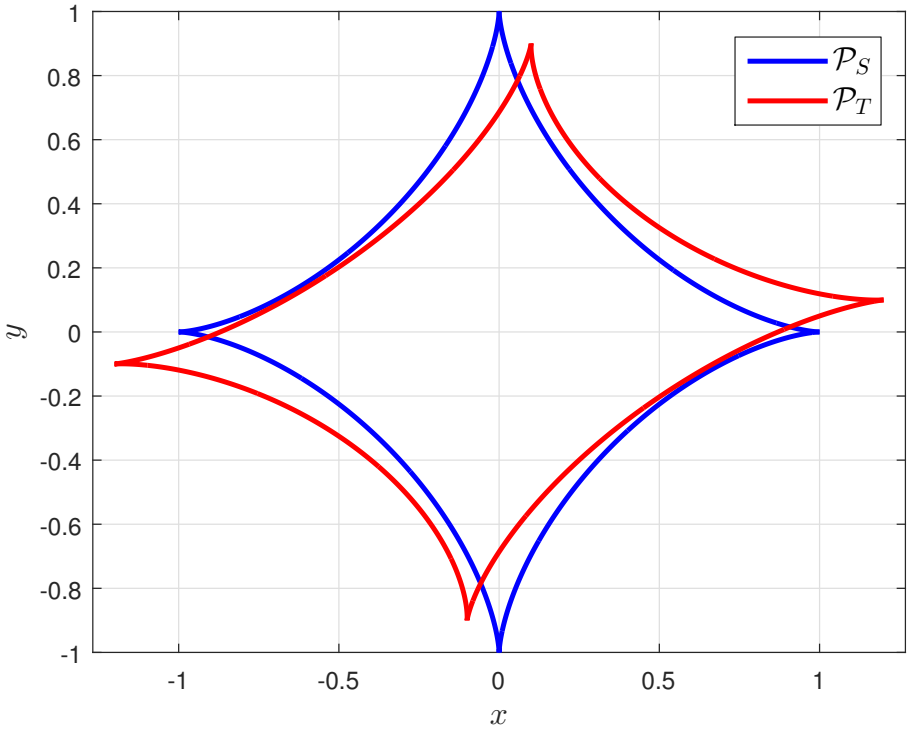


Figure 2.6: 2D Source and target geometries.

The registration results for application of this method to the generated 2D geometry are shown for various iterations in Figure 2.7 with the use of a linear basis function.

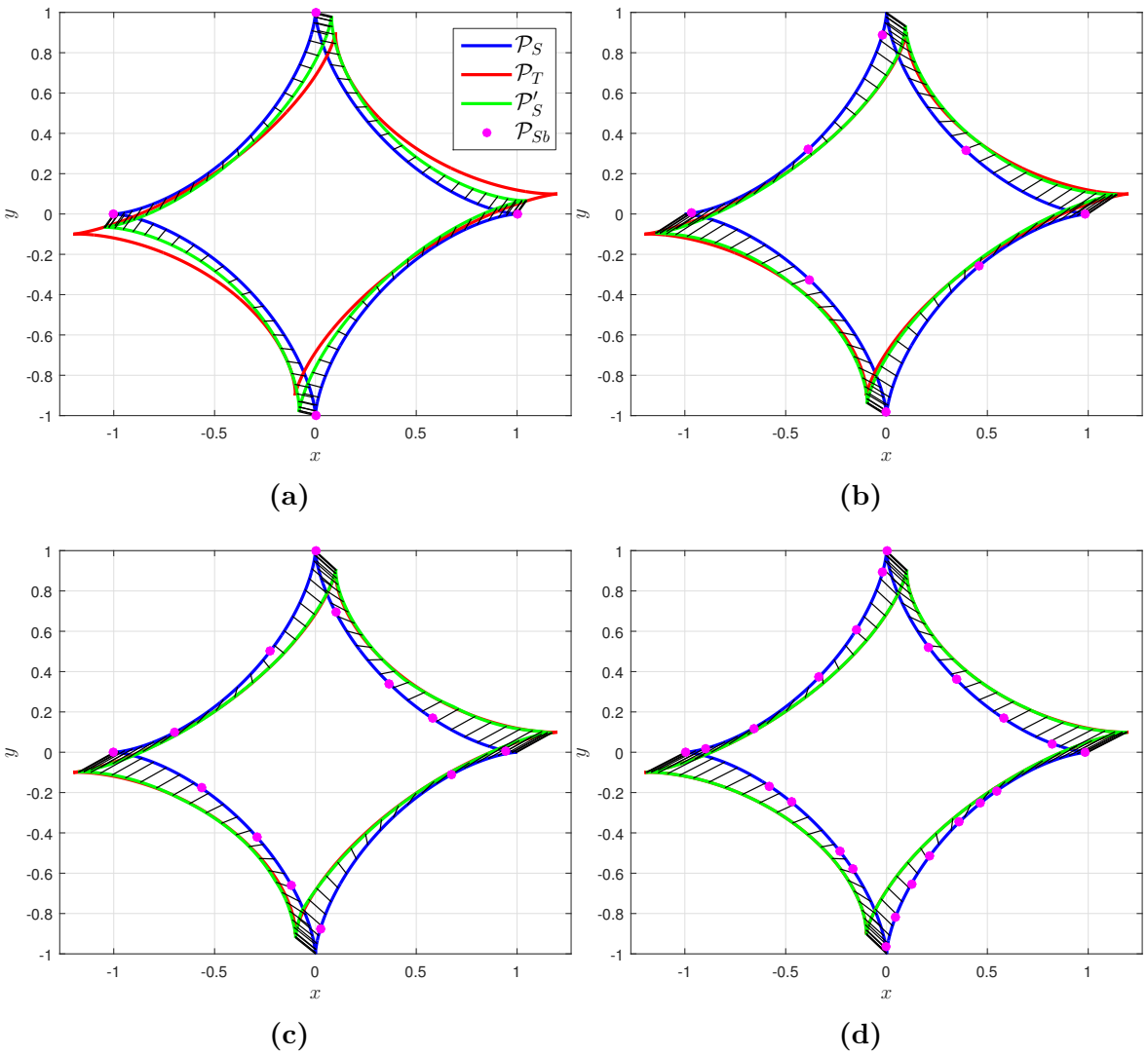


Figure 2.7: Registration result at various iterations, refer to Table 2.3. The black lines indicate the correspondence between registered and source geometries.

Table 2.3: Iteration and centre numbers in Figure 2.7.

Caption	Iteration, k	Centres, N_b
(a)	1	5
(b)	5	8
(c)	10	13
(d)	20	23

Three error functions are given at different iteration numbers in Figure 2.8. The first two errors relate to convergence properties of the method, while the third error relates directly to the least squares error which is aimed to be minimized:

1. The sum of the point-wise distances between each consecutive iteration

$$\varepsilon_{k-1} = \sum_{i=1}^{N_S} \left\| \mathbf{p}'_{Si(k)} - \mathbf{p}'_{Si(k-1)} \right\|_2^2 . \tag{2.34}$$

2. The sum of the point-wise distances between the current and final iteration

$$\varepsilon_{k_f} = \sum_{i=1}^{N_S} \left\| \mathbf{p}'_{Si(k_f)} - \mathbf{p}'_{Si(k)} \right\|_2^2 . \tag{2.35}$$

3. The sum of the point-wise distances between the current iteration and the true position on the target point-set

$$\varepsilon = \sum_{i=1}^{N_S} \left\| \mathbf{p}_{Ti} - \mathbf{p}'_{Si(k)} \right\|_2^2 . \tag{2.36}$$

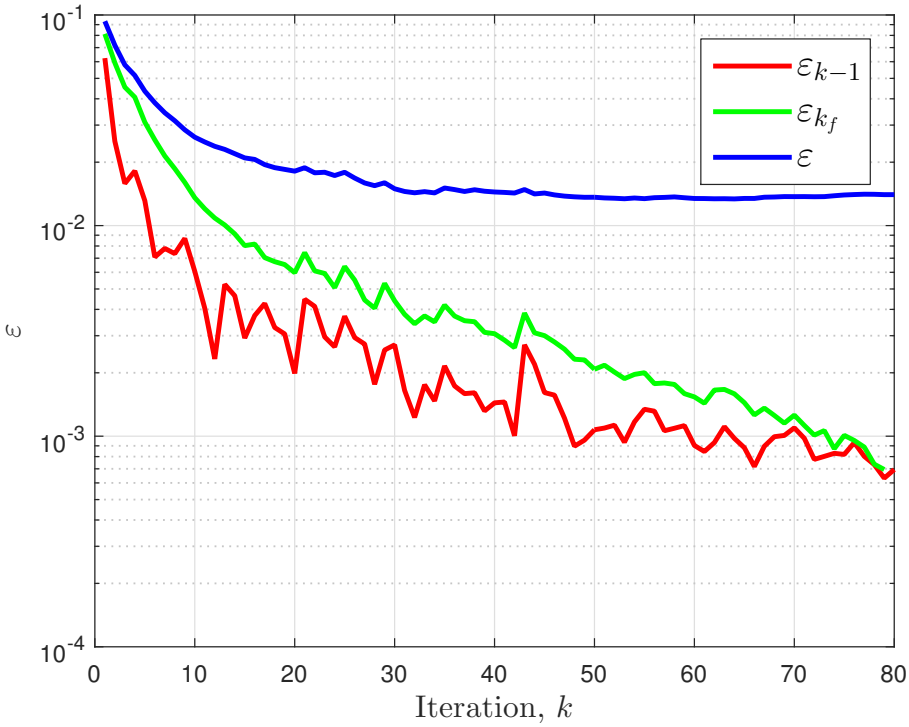


Figure 2.8: Various error functions vs. iteration number.

The following may be observed from Figure 2.8:

1. The sum of the point-wise distances between each consecutive iteration, Equation (2.34), is decreasing non-monotonically. It is therefore recommended that the termination criterion be altered to have the mean of a number of consecutive final iterations below the convergence tolerance.

2. The sum of the point-wise distances between the current and final iteration, Equation (2.35), is also decreasing non-monotonically. This shows that the last few iterations are closer to the final result than the preceding iterations, suggesting unique convergence.

3. The sum of the point-wise distances between the current iteration and the true position on the target point-set, Equation (2.36), is initially a decreasing function with a minimum at around 60 iterations where after it is slightly increasing. This shows that the approximation does not converge to the target point-set. This function is however never practically available and only shows that the heuristic correspondence approximation is non-exact.

CHAPTER 3

REGISTRATION APPLICATION

3.1 CHAPTER OVERVIEW

Firstly a NR-ICP registration strategy based on the methods in Chapter 2 is proposed. This strategy is characterized, where the parameter effects of the NR-ICP registration algorithm are studied. Termination criteria and algorithm parameters for the NR-ICP registration strategy, to be used for the remainder of the study, are proposed based on the characterisation findings.

3.2 REGISTRATION STRATEGY

The proposed NR-ICP registration strategy, which will be used in the remainder of this study, is set out as follows:

1. Manual pre-alignment, where the geometries' centres are moved to the origin and roughly rotated consistently along chosen directions. This is because PCA based registration of the used cranium and mandible geometries result in principal directions/components with similar eigenvalues, where the geometries are often aligned incoherently.

2. Rigid registration through R-ICP, Section 2.3.3, where a finer rigid alignment is obtained.

3. Non-rigid registration through NR-ICP, Section 2.4.2, where the final approximated geometry with correspondence is obtained.

3.3 REGISTRATION STRATEGY CHARACTERISATION

The intent of this section is to investigate the effect of the NR-ICP parameters on the resulting generated correspondence. In general NR-ICP registration displacement field approximations are not unique - several displacement fields/transformations exist which result in the same perceived geometry i.e. a target geometry can be represented by different deformations of a source geometry. This is clearly demonstrated in this section, as a number of configurations of the NR-ICP parameters, with different resulting displacement fields, are capable of accurate representations of the target geometries. The ideal is a registration result with a deformation mapping that is homologous between different observations i.e. landmarks between observations are represented by the same vertices/points. Practically this may not be attainable, therefore the intent is to determine whether some configurations of the NR-ICP registration parameters are better at approximating true correspondence. In this subsection this is investigated by application of known non-linear displacement fields and evaluating how close the generated correspondence is compared to known true deformation correspondence. The NR-ICP parameters studied in this section are:

1. the choice of basis function $\phi(d)$,
2. the number of iterations k and
3. the number of RBF centres N_b.

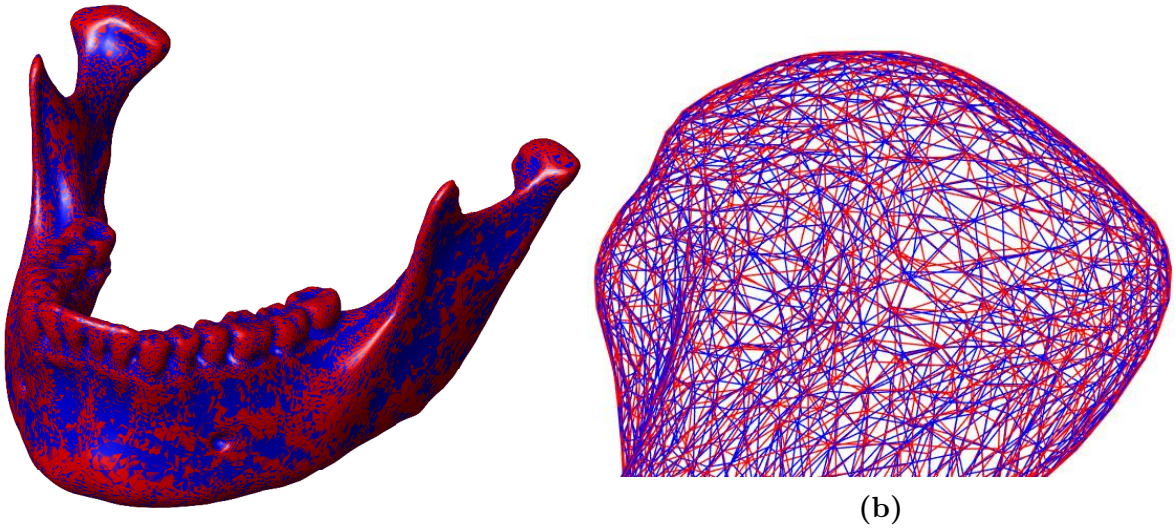
3.3.1 Procedure

1. Two different mesh representations of an arbitrary mandible geometry, described by point-sets $\mathcal{P}_{SA}$ and $\mathcal{P}_{SB}$, are generated. These meshes do not have the same connectivity, vertex numbers or vertex positions as illustrated in Figure 3.1. Consequently these meshes do not have correspondence, which is representative of two different surface scans of the same geometry.

2. An artificial dataset is generated where both $\mathcal{P}_{SA}$ and $\mathcal{P}_{SB}$ are transformed non-linearly $N_M = 15$ times to generate a Monte-Carlo (MC) sample:

$$\mathcal{P}_{SA} \overset{f_J}{\longmapsto} \mathcal{P}_{TAJ}, \quad \mathcal{P}_{SB} \overset{f_J}{\longmapsto} \mathcal{P}_{TBJ}. \tag{3.1}$$

   The non-linear deformation f_J for each sample is given by the following quadratic



(a) **(b)**

Figure 3.1: (a) Two superimposed mesh representations of the same mandible geometry, (b) Close-up of superimposed meshes.

deformation mapping:

$$\begin{bmatrix} x & y & z \end{bmatrix} = \begin{bmatrix} X \circ X & 2X \circ Y & 2X \circ Z & Y \circ Y & 2Y \circ Z & Z \circ Z & X & Y & Z \end{bmatrix} \mathbf{C} \tag{3.2}$$

with

$$\begin{aligned} C_{ij} &\sim \mathbf{U}\{-0.001, 0.001\}, \quad i = 1 \ldots 6, \quad j = 1 \ldots 3 \\ C_{72}, C_{73}, C_{81}, c_{83}, C_{91}, C_{92} &\sim \mathbf{U}\{-0.05, 0.05\} \\ C_{71}, C_{82}, C_{93} &\sim \mathbf{U}\{0.95, 1.05\} \end{aligned} \tag{3.3}$$

where

x, y, z	Column vectors with coordinates of the deformed geometry p_T
X, Y, Z	Column vectors with coordinates of the undeformed geometry p_S
$\boldsymbol{C}$	Matrix containing polynomial deformation constants C_{ij}
$\circ$	Denotes the Hadamard/point-wise product
$\mathbf{U}\{a, b\}$	Sampled from the uniform distribution with bounds a and b

The upper and lower bounds of the uniform distributions, from which the coefficients c_{ij} are sampled in Equation (3.3), are chosen so that the resulting deformation generates reasonable variations of the chosen mandible geometry. The superimposed Monte-Carlo samples are given in Figure 3.2.

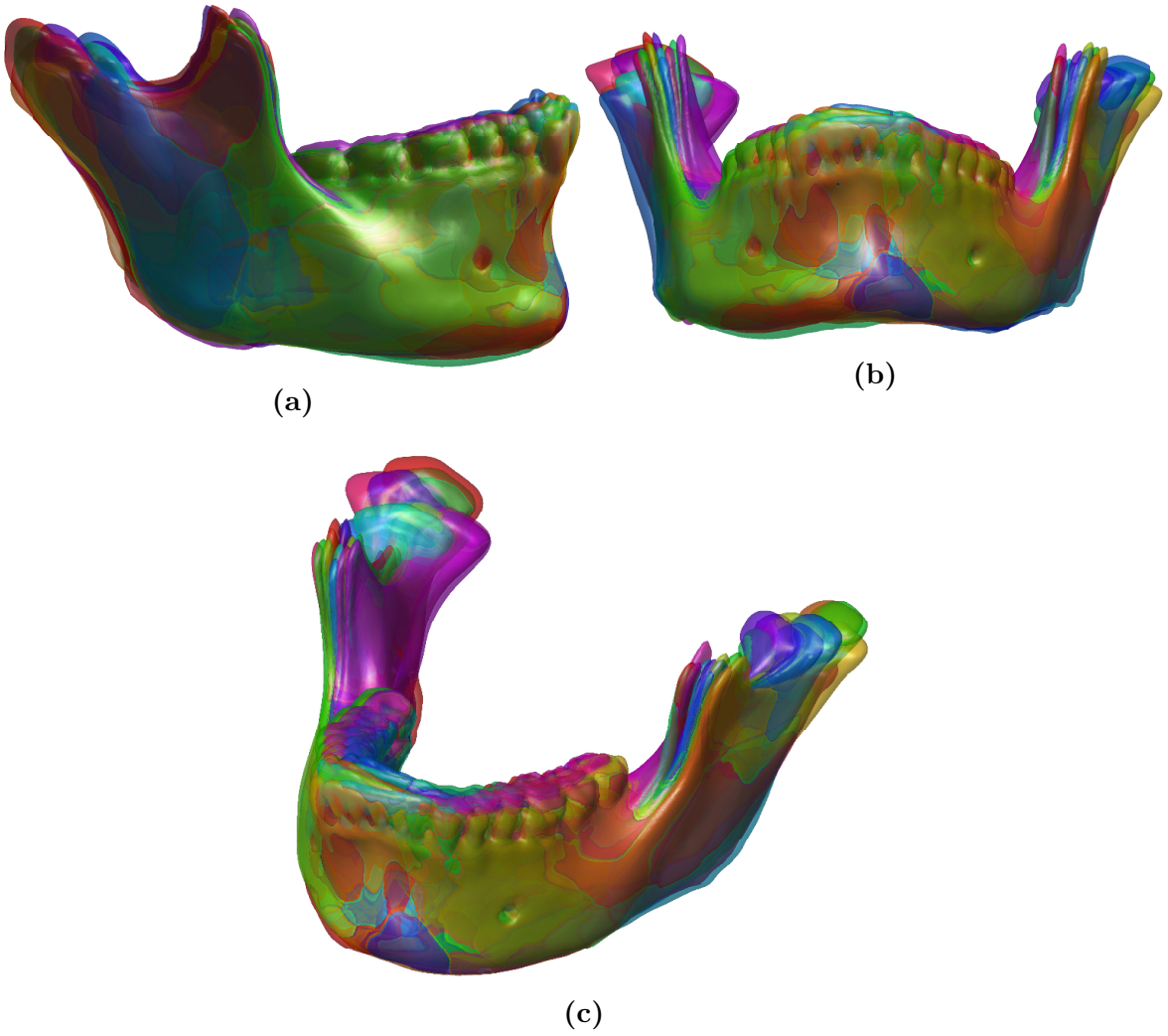


Figure 3.2: Monte-Carlo non-linear deformation sample views. (a) Right side view, (b) Frontal/Anterior view and (c) Isometric View.

3. The source mesh point-set p_{SA} is registered to each target mesh point-set p_{TBJ} to give p'_{SAJ}.

4. The mean point-wise distance error between the actual transformation/deformation mapping and registration approximation results (averaged over the number of registrations/Monte-Carlo samples N_M) is used to generate error distributions for various basis functions with respect to the number of RBF centres N_b and number of iterations k. The actual correspondence error is given by

$$\varepsilon_{(iJ)} = \|(\boldsymbol{p}_{TAJi} - \boldsymbol{p}'_{SAJi})\|_2^2 . \tag{3.4}$$

The nearest neighbour correspondence error, nearest point in $\mathbf{p}_{TBJi}$ to each point in $\mathbf{p}'_{SAJi}$, is given by

$$\varepsilon_{nn(iJ)} = \left\|(\boldsymbol{p}_{TBJc_S(i)} - \boldsymbol{p}'_{SAJi})\right\|_2^2 , \tag{3.5}$$

with

$$i = 1, 2, .., N_S; \quad J = 1, 2, .., N_M .$$

The nearest neighbour error is included, as it is the only available function during the registration process. The intent is to determine whether it can be used as a proxy to the actual correspondence error. The resultant error surfaces are given in Section 3.3.2 for $k \in [0; 25; 50 \ldots 150]$, $N_b \in [25; 50; 75, \ldots 200]$ and the following commonly used basis functions:

(a) Linear
(b) Cubic
(c) Thin plate spline
(d) Gaussian

Due to the high computational cost of evaluating the error distributions, the Gaussian basis function is only evaluated for a single value of ε. The Gaussian half-width h is chosen as the maximum distance from any point towards its closest centre so as to remain affected by the approximated displacement field, with a basis function evaluation of at least $\phi(d) = 0.5$. This implies that the Gaussian function parameter ε is given by

$$\varepsilon = \sqrt{\frac{\ln 2}{h^2}} . \tag{3.6}$$

3.3.2 Results

The mean and maximums of the actual correspondence error distributions, Equation (3.4), are rendered as surfaces in Figures 3.5, 3.8, 3.11 and 3.14 for each respective basis function. The maximum of the distribution is given by the translucent surface and the mean by the opaque surface. For a qualitative error comparison consider the geometry relative size bounding box in Figure 3.4. The following is observed from these surfaces:

1. The true correspondence error distributions vary with respect to the choice of basis function $\phi(d)$, number of iterations k and number of centres N_b.

2. The correspondence error is generally reduced with increased centres distributed over an increased number of iterations, regardless of basis function. The maximum and mean of the error distributions are given for each respective basis function in the approximate minimum region at $N_b = 200$ and $k = 200$ in Table 3.1. This shows that the mean of the true correspondence error distributions are comparable, but the maximum error differs significantly between basis functions.

Table 3.1: Correspondence error distribution properties for each basis function.

Basis function	mean(ε) [mm]	max(ε) [mm]
Linear	0.0658	1.1215
Cubic	0.0517	0.2567
TPS	0.0512	0.3172
Gaussian	0.0768	2.1046

3. The linear and Gaussian basis functions require many more centres and iterations to achieve comparable maximum error. This behaviour of the Gaussian basis function may however be an artefact of the choice of the Gaussian half-width.

4. The cubic and thin plate spline basis functions have distinctive transitions from higher to lower maximum error, illustrating the highly path dependant behaviour of the NR-ICP algorithm. These transitions are concomitant with the separation of regions with higher maximum error, due to a larger increase in the number of fitting centres/degrees of freedom per iteration, and lower maximum error, due to a lower increase in the number of fitting centres/degrees of freedom per iteration.

This illustrates that the NR-ICP algorithm is highly dependant on the rate at which fitting complexity is increased between iterations.

5. The fitting complexity increase between iterations also explains the general decrease in true correspondence error for an increase in the number iterations when a fixed number of centres are considered, and vice versa. This is because higher fitting complexity between iterations allows too much fitting freedom and shape change, altering correspondence built up in consequent iterations.

6. The general true correspondence error decrease along unit centre increase per iteration, $N_b = k$, may be due to the fact that this is the slowest possible whole numbered rate at which fitting complexity/degrees of freedom can be increased per iteration i.e. without repeating the number of fitting centres between iterations.

The mean and maximums of the nearest neighbour error distributions, Equation (3.5), are rendered as surfaces in Figures 3.6, 3.9, 3.12 and 3.15 for each respective basis function. The maximum of the distribution is given by the translucent surface and the mean by the opaque surface. The following is observed from these surfaces:

1. The nearest neighbour error distributions are comparably invariable with respect to the choice of basis function $\phi(d)$, number of iterations k and number of centres N_b. The mean and maximums of these distributions present as near flat surfaces except at $k = 0$ (only rigid alignment). This is because the registration result at the lowest evaluated iteration number $k = 25$ already results in a visually indiscernible difference between the registrations and target geometries. This suggests that the majority of the nearest neighbour error is in-surface error. (The mean maximum in-surface error is expected to be approximately half the average triangle side-width, which is 1.06mm in this instance). The maximum and mean of the nearest neighbour error distributions are given for each respective basis function in the approximate minimum region at $N_b = 200$ and $k = 200$ in Table 3.2, and shows an insignificant difference:

2. The nearest neighbour error illustrates that it is not a good proxy of the true correspondence error.

The true correspondence error distributions vs. iterations k, with $N_b = 200$, are given for each respective basis function in Figures 3.7, 3.10, 3.13 and 3.16. The following is

Table 3.2: Nearest neighbour error distribution properties for each basis function.

Basis function	mean(ε) [mm]	max(ε) [mm]
Linear	0.2884	1.7755
Cubic	0.2844	1.7949
TPS	0.2847	1.7887
Gaussian	0.2869	1.7553

observed from these surfaces:

1. The resulting error distributions are of the same overall shape for iterations beyond $k = 100$, regardless of basis function. This justifies using the mean and maximums of these distributions as error descriptors in the region of minimum true correspondence error as done earlier.

2. Beyond $k = 100$ the maximum tail-end is shortened considerably for each basis function (consider the logarithmic scaling of the error axis).

3. Investigation of these tails shows that the long maximum tail-ends are generally concomitant with a region of higher true correspondence error for a single registration result within the MC sample. This is illustrated in Figure 3.3 which shows the registration error result for NR-ICP parameters $N_b = 100$, $k = 100$ and $\phi(d) = d^3$. Considering again the mean triangle side-width of 1.06mm, the nearest neighbour error, in Figure 3.3 (a), with a mean of 0.3mm suggests registration with largely in-surface true correspondence error, which is corroborated by not being able to visually discern a difference between the target and registration result. This is again corroborated in Figure 3.3 (b), where the nearest neighbour error is weighted by the absolute of the dot-product between the unit correspondence vector (source to target point) and the target point unit normal. This weighting approximately removes the in-surface error and retains the perceptible normal error. Comparison with the true correspondence error, (c) in Figure 3.3 , with a mean of 0.07mm, also corroborates and explains this. The true correspondence error clearly illustrates that the high in-surface-dominant error associated with the maximum tail-end is confined to a region of the registration result. This shows that high in-surface true correspondence errors cannot be identified from available error measures during registration. Manual visual inspection may be an option, but the practicality and success of exhaustive visual cross validation however remains questionable for the following reasons:

(a) Manual inspection of multiple registrations will be tedious, especially having to repeat this for the purpose of building a dataset.

(b) There is uncertainty as to whether it would be humanly possible to even discern the subtle differences of the largely in-surface correspondence errors (a maximum of 0.69mm in Figure 3.3 (c)).

(c) It is uncertain if a human generated guess would practically be any better in areas which are not geometrically distinct. It may on the contrary, for the purpose of regression, be better to describe these through registration, rather than relying on human perception to enforce biological homology where it does not strictly exist.

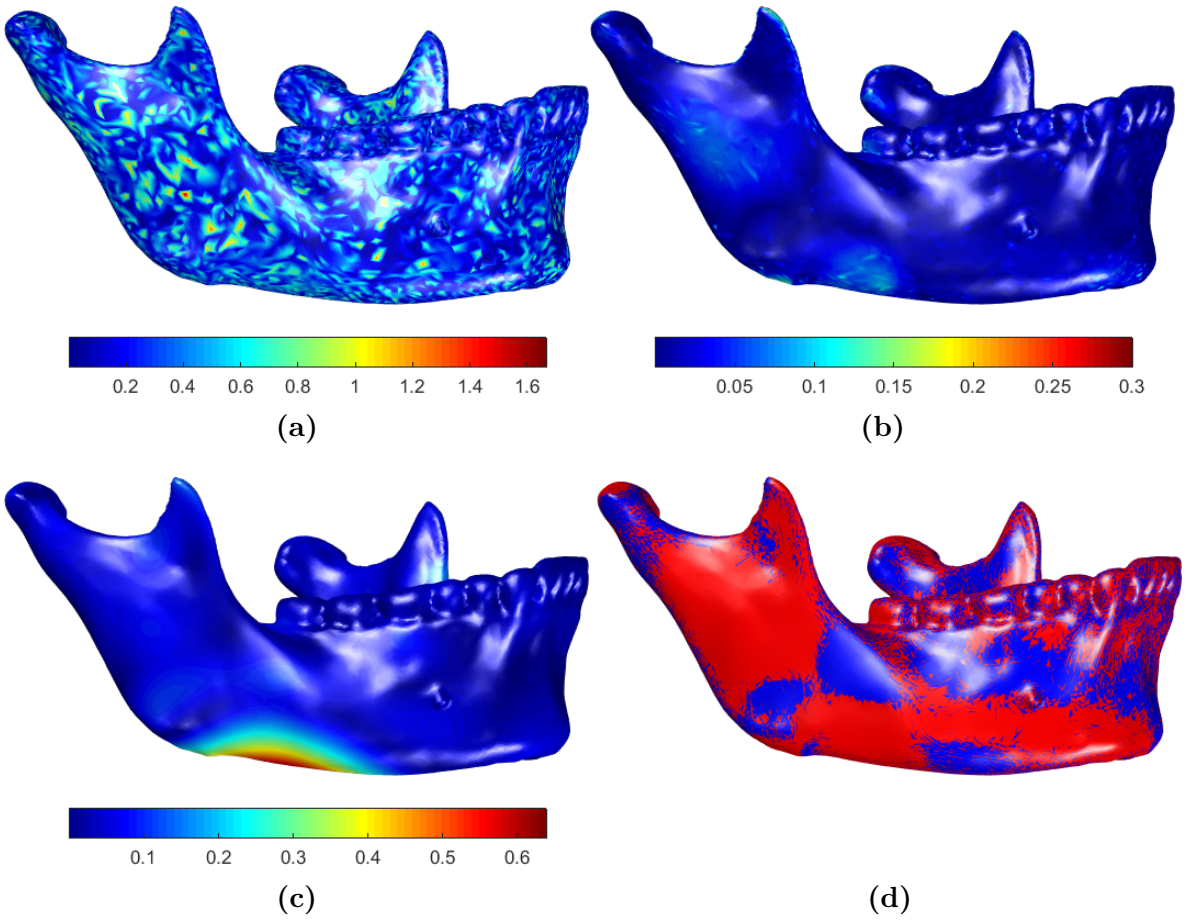


Figure 3.3: Registration error of the MC observation responsible for high true correspondence error, all given in mm. (a) Nearest neighbour correspondence error ε_{nn}, (b) Normal weighted nearest neighbour correspondence error $\varepsilon_{nn} w_n$, (c) True correspondence error ε and (d) Registration result (blue) and target (red).

The errors have thus far only been used comparatively with the magnitudes remaining unaddressed. A discussion on the relative size of the error w.r.t. the size of the geometry and shape variation may be insightful at this point. For reference consider the dimensions of the MC sample given in Figure 3.4. Comparison of the rough maximum dimensions with the errors in Tables 3.1 and 3.2 serves as indication that the errors are insignificant even if they were out-of-surface errors i.e. they would remain visually imperceptible. Comparison with the mean point wise difference, $\varepsilon = 2.77$mm at $k = 0$ in Figure 3.5, shows that the mean true correspondence errors in Table 3.1 are small, $\varepsilon \approx 0.06$mm, with correspondingly small added variance i.e. the variance due to registration error (largely in-surface error) is roughly 2% compared to the variance due to shape differences/variation.

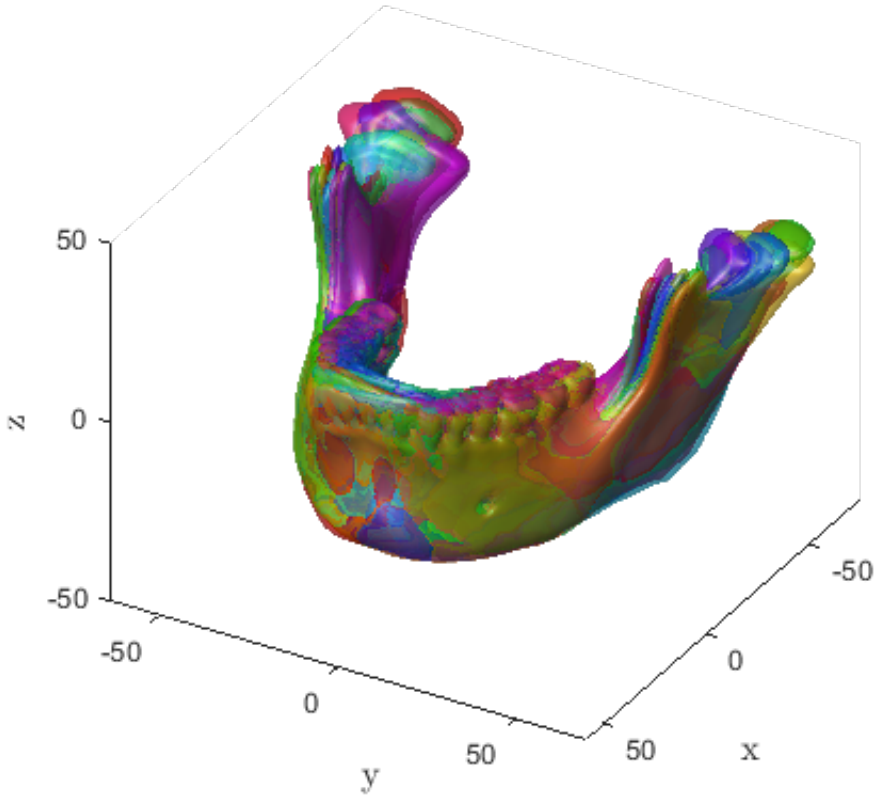


Figure 3.4: Superimposed MC sample.

3.3.2.1 Linear basis function

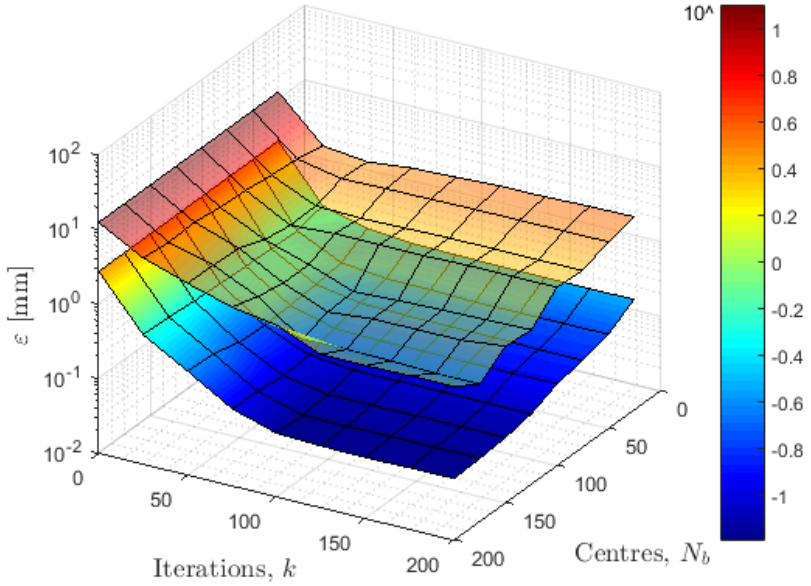


Figure 3.5: Actual correspondence error surface [mm per point] of Linear basis function for MC sample. The maximum of the distribution is given by the translucent surface and the mean by the opaque surface.

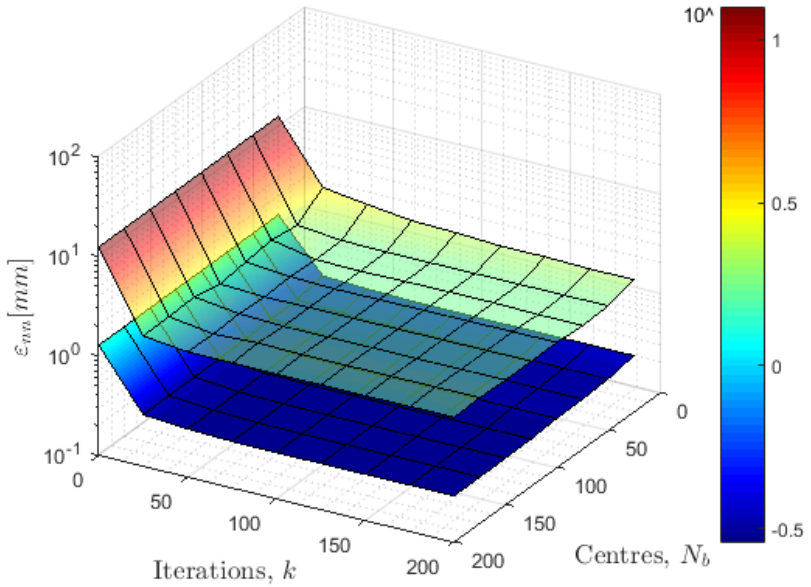


Figure 3.6: Nearest neighbour error surface [mm per point] of Linear basis function for MC sample. The maximum of the distribution is given by the translucent surface and the mean by the opaque surface.

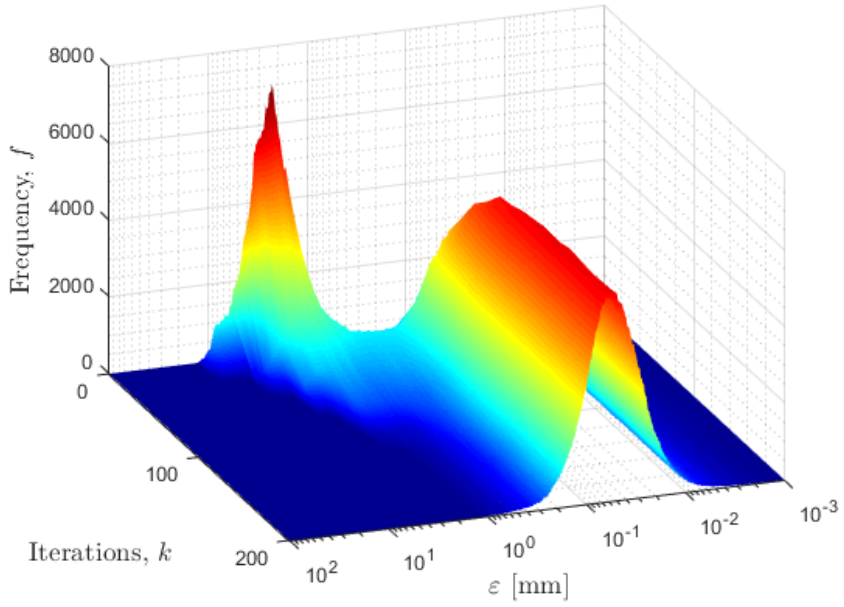


Figure 3.7: Error distribution for various iteration numbers k and number of centres $N_b = 200$ for use of the linear basis function on the MC sample.

3.3.2.2 Cubic basis function

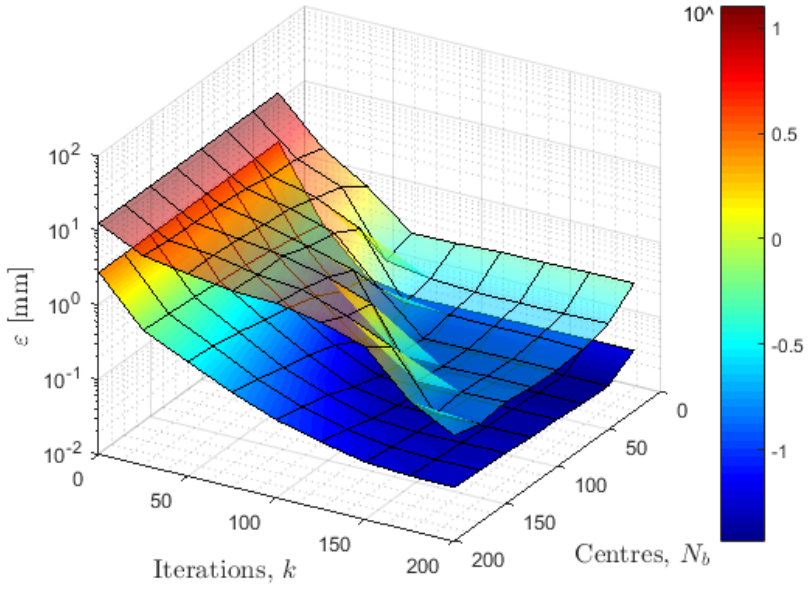


Figure 3.8: Actual correspondence error surface [mm per point] of cubic basis function for MC sample. The maximum of the distribution is given by the translucent surface and the mean by the opaque surface.

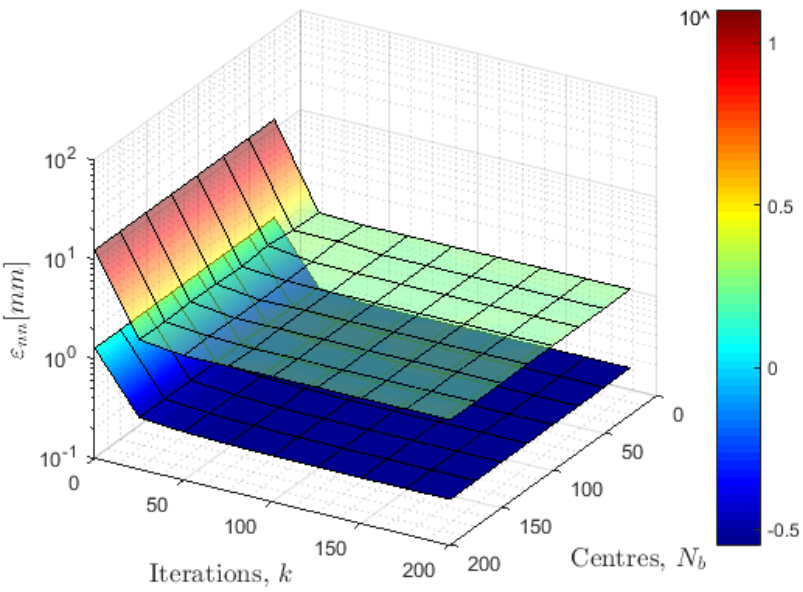


Figure 3.9: Nearest neighbour error surface [mm per point] of cubic basis function for MC sample. The maximum of the distribution is given by the translucent surface and the mean by the opaque surface.

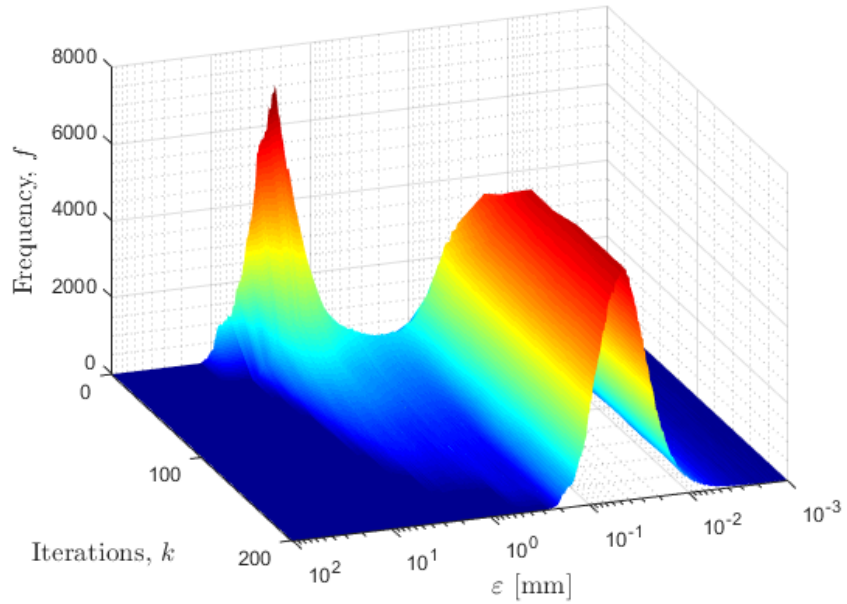


Figure 3.10: Error distribution for various iteration numbers k and number of centres $N_b = 200$ for use of the cubic basis function on the MC sample.

3.3.2.3 TPS basis function

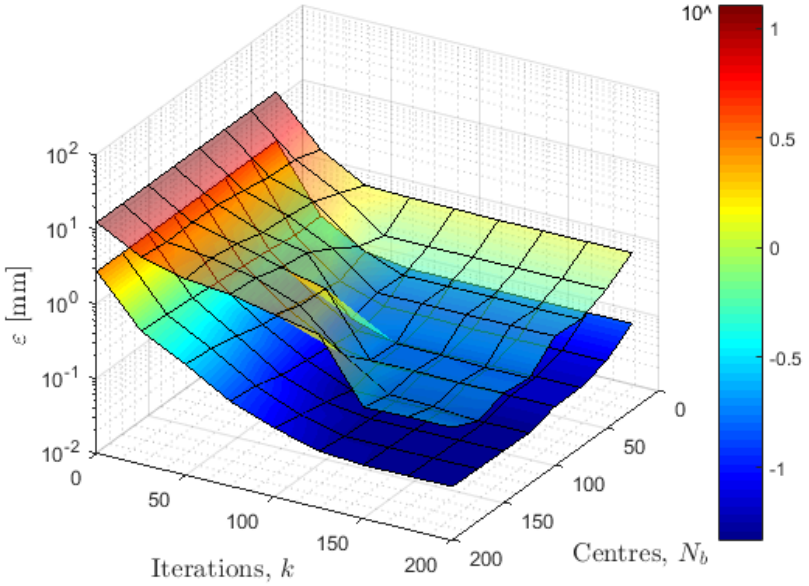


Figure 3.11: Actual correspondence error surface [mm per point] of TPS basis function for MC sample. The maximum of the distribution is given by the translucent surface and the mean by the opaque surface.

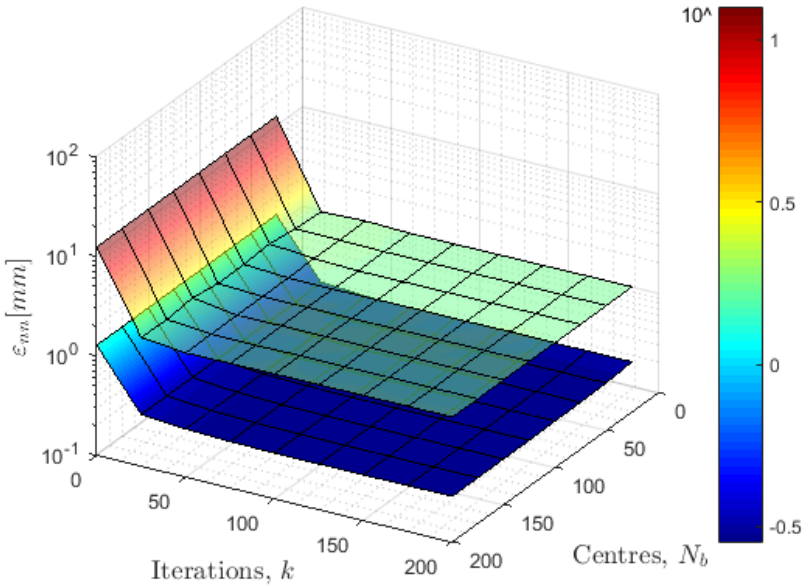


Figure 3.12: Nearest neighbour error surface [mm per point] of TPS basis function for MC sample. The maximum of the distribution is given by the translucent surface and the mean by the opaque surface.

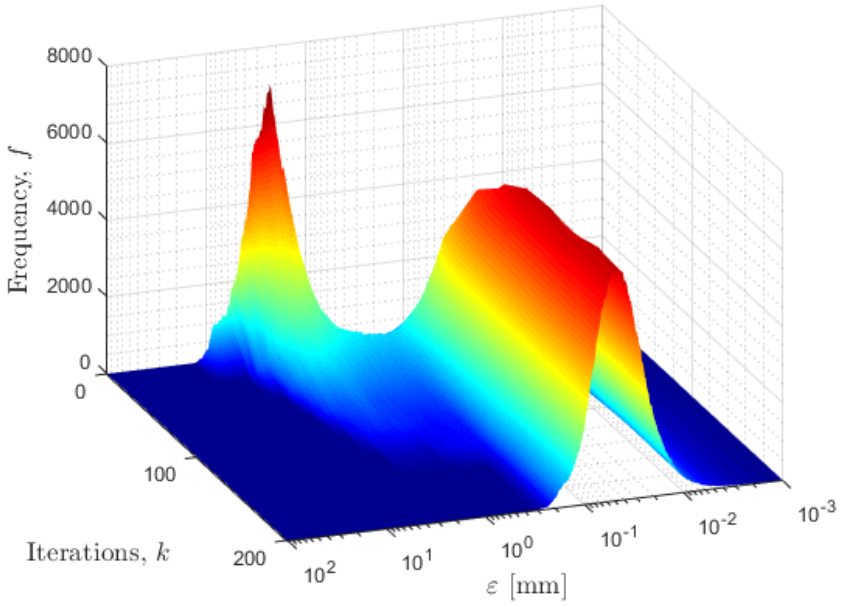


Figure 3.13: Error distribution for various iteration numbers k and number of centres $N_b = 200$ for use of the TPS basis function on the MC sample.

3.3.2.4 Gaussian basis function

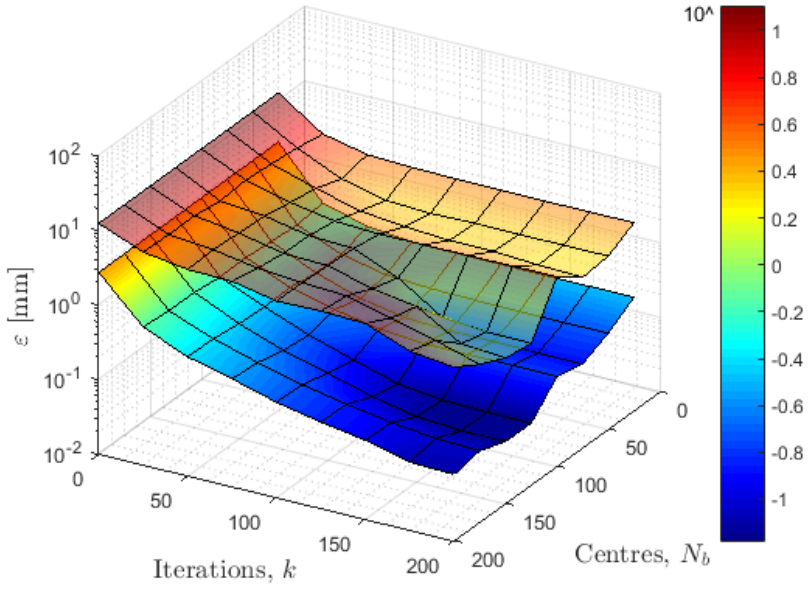


Figure 3.14: Actual correspondence error surface [mm per point] of Gaussian basis function for MC sample. The maximum of the distribution is given by the translucent surface and the mean by the opaque surface.

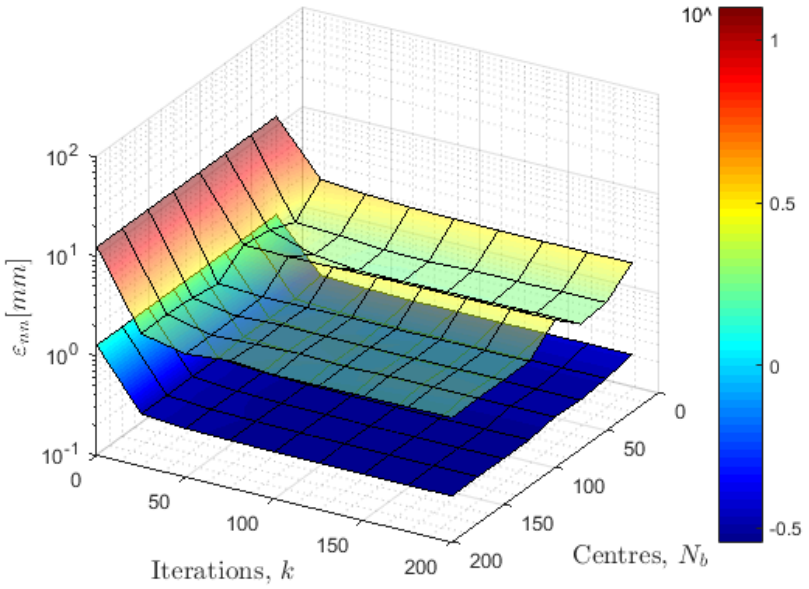


Figure 3.15: Nearest neighbour error surface [mm per point] of Gaussian basis function for MC sample. The maximum of the distribution is given by the translucent surface and the mean by the opaque surface.

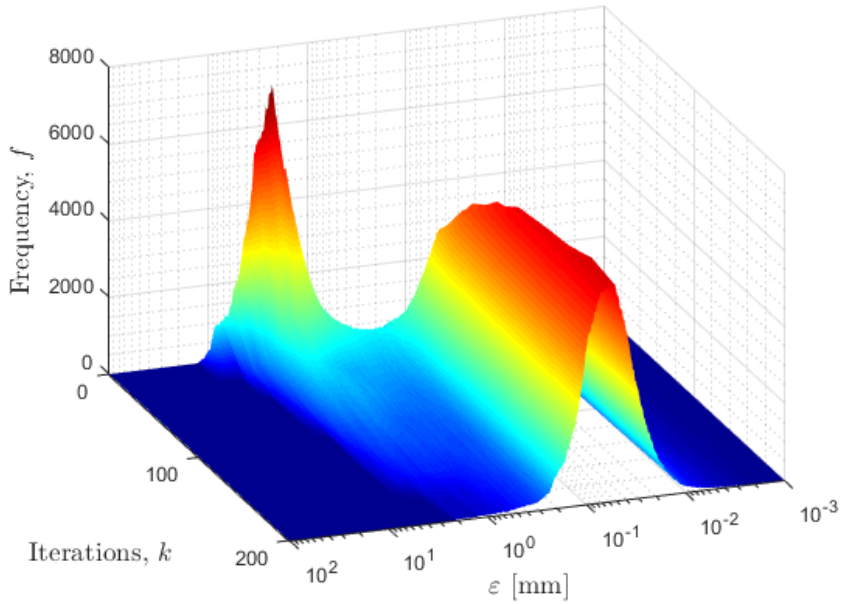


Figure 3.16: Error distribution for various iteration numbers k and number of centres $N_b = 200$ for use of the Gaussian basis function on the MC sample.

3.3.3 Conclusion

The following is concluded from the study in this section:

1. The true correspondence error distributions vary with respect to the choice of basis function $\phi(d)$, number of iterations k and number of centres N_b.

2. The registration procedure requires few centres and iterations to produce visually indiscernible registration results with regard to the target geometry i.e. the lowest sample of $N_b = k = 25$ results in a "good" visual registration regardless of the basis function used. The numerical error between configurations is largely due to in-surface error, which is not visually perceptible.

3. The registration algorithm requires many more centres and iterations (basis function dependant) than what is required for a "good" visual registration to result in sub-millimetre true correspondence error.

4. The true correspondence error generally decreases along unit centre increase per iteration i.e. $N_b = k$ for increasing N_b and k, regardless of the basis function used.

5. The available nearest neighbour error measures during registration are not indicative of the true correspondence error and can therefore not be used as a proxy.

6. The nearest neighbour error is highly invariable in the region of concern $N_b, k \geq 25$, where the true correspondence error is lowered significantly. This and the previous point demotivates using this error measure as termination criteria for the registration algorithm.

Although these results are based on artificial quadratic deformation, which may not be representative of true shape variation of mandibles or crania, the results may still serve as a guideline. Ideally cross-validation should be used i.e. different datasets should be generated using the registration procedure, regressed and the parameters of the datasets which results in the least error should be used. This process may however only be computationally tractable for a few configurations of NR-ICP parameters, as each regression would also first require cross-validation. For the lack of any other alternative these may be restricted to:

1. Use of only the cubic and thin plate spline basis functions, which resulted in significantly lower true correspondence error for reduced N_b and k.

2. The number of centres and iterations may be restricted to $N_b = k$ for numbers which are $1\times$ and $2\times$ the minimum required for a "good" visual registration. These may then be repeated with doubling until the visual results after regression differ insignificantly.

It is also important to note that should the magnitudes of the errors in this study be indicative of true correspondence errors for real data, cross-validation would be meaningless with imperceptible visual differences. This is because the added variance related with in-surface error would be insignificant compared to the variance of shape difference/variation.

CHAPTER 4

REGRESSION METHODS

Regression requires the use of features/variables with one-to-one correspondence, which fundamentally requires that a sampled variable consistently represents the same quantitative feature. Chapter 2 discussed why raw surface-scan geometry does not fulfil this requirement and demonstrated how heuristic correspondence may be generated using proposed registration techniques. This chapter describes how registered point-sets may be pre- and post-processed to facilitate regression for the purpose of prediction. The intended application of this work will often be limited to datasets with few observations due to a culmination of ethical considerations, availability of physical data-sets (bone-collections) and cost/time/labour of digitisation, manipulation and processing. Only regression techniques that are suitable for collinear data with few observations are thus considered and documented.

4.1 REGRESSION

The regression algorithms used in this study are limited to multivariate regression techniques that are suited to the case where the number of predictor and response variables are many more than the number of observations $n_X, n_Y \gg n$, with high multicollinearity expected amongst both predictor and response variables. Partial Least Squares (PLS) regression is a regression technique suited to this application [16]. A better scaling regression technique based on Principal Components (PC) regression of a combined superset is also proposed.

4.1.1 Partial Least Squares (PLS)

The PLS algorithm is a statistical method, based on the use of derived input directions, that models the relationship between predictor and response matrices, $\boldsymbol{X}$ and $\boldsymbol{Y}$ respectively. This is done by decomposition of both $\boldsymbol{X}$ and $\boldsymbol{Y}$ into linear combinations of latent variables and a residual term [17]:

$$\boldsymbol{X} = \boldsymbol{T}\boldsymbol{P}^T + \boldsymbol{E} \tag{4.1}$$

$$\boldsymbol{Y} = \boldsymbol{U}\boldsymbol{Q}^T + \boldsymbol{F} \tag{4.2}$$

with

$\boldsymbol{T}, \boldsymbol{U}$	Latent variables
$\boldsymbol{P}^T, \boldsymbol{Q}^T$	Loading vectors associated with the respective latent variables
$\boldsymbol{E}, \boldsymbol{F}$	Residual terms that remain unexplained by decomposition

The decomposition is done in a manner so as to find directions of maximum covariance between predictor and response spaces [18]. This work makes use of a PLS implementation based on the SIMPLS algorithm [19]:

1. The covariance $\boldsymbol{S}_1$ is evaluated using

$$\boldsymbol{S}_1 = \boldsymbol{X}_0^T \boldsymbol{Y}_0, \tag{4.3}$$

   with zero subscripts indicating centering.

2. The a-th weight vectors $\boldsymbol{r}_a$ and $\boldsymbol{q}_a$ are determined by

$$\underset{\boldsymbol{r}_a, \boldsymbol{q}_a}{\arg\max} \left(\boldsymbol{r}_a^T \boldsymbol{S}_a \boldsymbol{q}_a \right), \tag{4.4}$$

   subject to the weights being orthonormal. These are evaluated as the left and right singular vectors of $\boldsymbol{S}_a$.

3. Each maximisation results in an a set of weighting vectors, latent vectors and loading vectors. In addition, latent vectors are required to be orthogonal to result in more than one solution.

4. The loading vectors are transformed into an orthogonal basis and used to deflate the covariance before extraction of the next component (repeating from step 2):

$$\boldsymbol{S}_{a+1} = \boldsymbol{S}_a - \boldsymbol{P}(\boldsymbol{P}^T\boldsymbol{P})^{-1}\boldsymbol{P}^T\boldsymbol{S}_a \tag{4.5}$$

5. This process may be repeated until the rank of the deflated covariance reduces to zero i.e. when the number of components, a, equals the lesser of the number of predictor variables, n_X, or the number of observations, n, minus one:

$$a_{\max} = \min(n_X, n-1). \tag{4.6}$$

4.1.2 Superset Principal Component (SPC) regression

This technique is developed for improved computational scaling over PLS, with the intuitive assumption of two partial sets $\boldsymbol{X}$ and $\boldsymbol{Y}$ forming part of a larger set $\boldsymbol{Z}$

$$\boldsymbol{Z} = \begin{bmatrix}\boldsymbol{X} & \boldsymbol{Y}\end{bmatrix}. \tag{4.7}$$

The maximal directions of variance or loadings $\boldsymbol{Z}_L$ are evaluated as the left singular vectors of the combined set $\boldsymbol{Z}_0$, which can be decomposed into vector components $\boldsymbol{X}_L$ and $\boldsymbol{Y}_L$

$$\boldsymbol{Z}_L = \begin{bmatrix}\boldsymbol{X}_L \\ \boldsymbol{Y}_L\end{bmatrix}. \tag{4.8}$$

The scores $\boldsymbol{Z}_S$ are evaluated from solving:

$$\boldsymbol{Z}_0 = \boldsymbol{Z}_S\boldsymbol{Z}_L^T, \tag{4.9}$$

$$\begin{bmatrix}\boldsymbol{X}_0 & \boldsymbol{Y}_0\end{bmatrix} = \boldsymbol{Z}_S\begin{bmatrix}\boldsymbol{X}_L \\ \boldsymbol{Y}_L\end{bmatrix}^T, \tag{4.10}$$

which shows that $\boldsymbol{Z}_S$ can be solved using only a single subset, i.e. $\boldsymbol{X}_0(\boldsymbol{X}_L^T)^+$. Assuming $\boldsymbol{X}_n$ to be a new observation n, corresponding estimations/predictions $\widehat{\boldsymbol{Z}}_{n0}$ and $\widehat{\boldsymbol{Y}}_{n0}$ are available through:

$$\widehat{\boldsymbol{Z}}_{n0} = \boldsymbol{X}_{n0}(\boldsymbol{X}_L^T)^+\boldsymbol{Z}_L^T, \tag{4.11}$$

and

$$\widehat{\boldsymbol{Y}}_{n0} = \boldsymbol{X}_{n0}(\boldsymbol{X}_L^T)^+\boldsymbol{Y}_L^T, \tag{4.12}$$

where $^+$ indicates the pseudo-inverse and zero subscripts refer to centering with means of corresponding training sets $\boldsymbol{X}$, $\boldsymbol{Y}$, and $\boldsymbol{Z}$.

4.2 REGRESSION TIME COMPLEXITY

The proposed regression methods' time complexity is now considered. The aim is to set system size limits that result in timeous/tractable solutions for cross-validation (especially during prototyping). This is because triangular mesh representation accuracy/complexity is limited by the number of vertices, which is inherently geometry dependent. This study is based on a maximum of $n = 50$ observations, which is expected to be the maximum due to the time, cost and labour involved in acquisition, digitisation and manipulation of the biological samples obtained from the Pretoria bone collection for application in Chapter 6, [20]. The maximum system sizes for the respective regression algorithms are limited so as to result in less than 90% memory usage of 64GB, to prevent any virtual memory usage and associated transfer overhead/paging time.

4.2.1 PLS Complexity

Time complexity of PLS based on the SIMPLS algorithm is studied in this subsection. The implementation is based on a sparse SVD solver which only obtains the first singular value and corresponding singular vectors (Equation (4.4)). This results in significantly improved time complexity for $n_X, n_Y > 5,000$ at the cost of additional memory.

4.2.1.1 Retained component time complexity

The cumulative time vs. component complexity for the SIMPLS algorithm is given for $n_X, n_Y = 5,000$ and $n = 50$ in Figure 4.1, with a linear fit for the first 25 components. It is observed that the behaviour of the algorithm with a sparse SVD solver is initially linear, whereafter decreased time per component is observed as the number of retained components approaches the number of observations $a \to n - 1$. It is assumed that the sparse solver works more efficiently as the rank of the cross-covariance is significantly reduced through iterative deflation. Although not entirely realistic, full component retention is expected to be the worst case scenario. An upper-bound estimate for this scenario is expected to be $t_{n-1} = t_1(n - 1)$, $t_1 = 10$s.

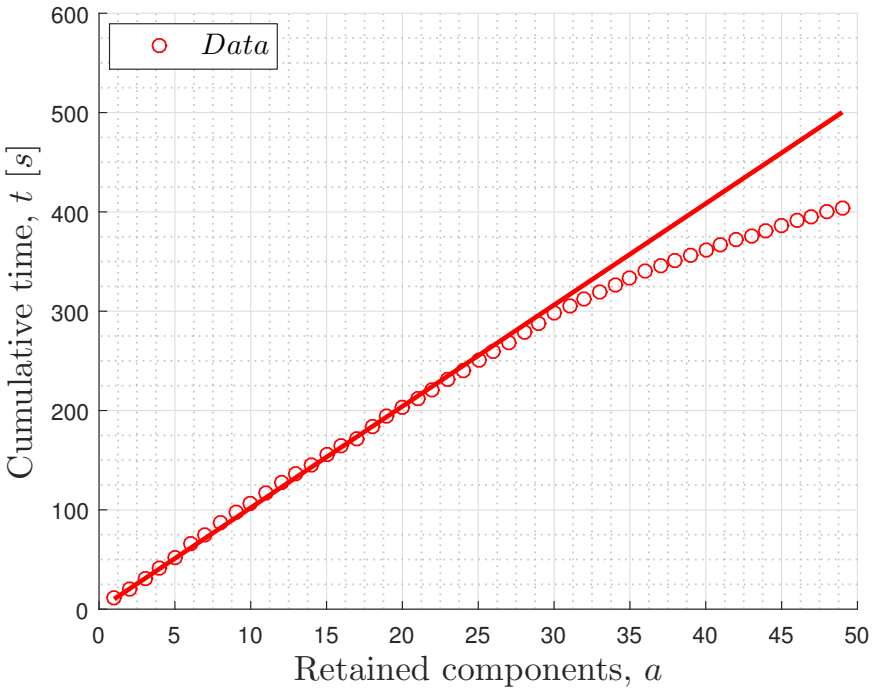


Figure 4.1: Time-component complexity for PLS with fit $t = ca$ for first 25 components.

4.2.1.2 System size time complexity

The time vs. size complexity for a single evaluated component is given for various combinations of n_X and n_Y with $n_X + n_Y \leq 50,000$ (memory limitation to prevent paging) in Figure 4.2. Polynomial scaling of the form $t \propto a n_X^b n_Y^c$, with $b, c \approx 1$ (quadratic polynomial) seems to give a good qualitative approximation of the asymptotic size complexity. The use of the sparse SVD solver is motivated by the fact that a direct exact SVD solver exhibits cubic asymptotic time-size complexity of the form $\mathcal{O}(min(n_X n_Y^2, n_X^2 n_Y))$ [21]. Recall that the upper bound for full component retention is expected to be $t_{n-1} = t_1(n-1)$, with total times of $t_{\max} \leq 49 t_1$ as given in Figure 4.2. This clearly demonstrates the limitation of PLS in terms of the CPU time required for the large systems associated with geometric data. Consider that evaluation of a single PLS component regression of 10,000 points in $\mathbb{R}^3$ $(n_X, n_Y = 30,000)$ for both predictor and response geometries takes approximately 10 minutes, the worst case scenario of $a = n - 1$ with 50 observations results in an upper-bound approximate total time of 8 hours. This with the use of non-exhaustive k-fold cross-validation using 10 folds increases run-time to over 3 days, whereas exhaustive leave-one-out cross-validation increases run-time to 3 weeks! This clearly illustrates the need for curvature driven mesh simplification or sub-sampling to result in timeous/tractable PLS regression models for

geometry data.

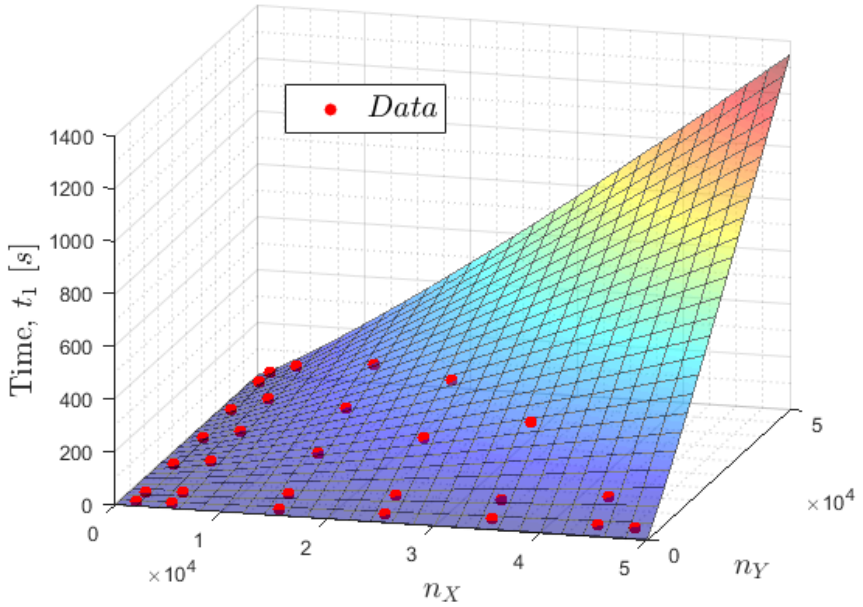


Figure 4.2: Time-size complexity for PLS with fit $t = an_X^b n_Y^c$.

4.2.2 SPC regression complexity

The asymptotic time-size complexity of SPC is given by the asymptotic complexity of the SVD of Equation (4.7), a $n \times (n_X + n_Y)$ matrix. The theoretical complexity is given by $\mathcal{O}(min((n_X + n_Y)n^2, (n_X + n_Y)^2 n))$ [21]. Given $n << (n_X + n_Y)$ for this study, the asymptotic complexity is given by $\mathcal{O}((n_X + n_Y)n^2)$. This is verified by the time-observation complexity in Figure 4.3, where the time complexity fit of the form $t \propto an^b$ results in exponent $b \approx 2$ for $n_X, n_Y = 100,000$. The time-size complexity is given in Figure 4.4, where the time complexity fit of the form $t \propto a(n_X + n_Y)$ also gives a good qualitative approximation of the asymptotic time-size complexity for $n = 50$. Note that all components are evaluated through a single decomposition and reduced rank regression is facilitated through truncation of the loading vectors, whereas for PLS only the truncated loading vectors are evaluated through iterative deflation.

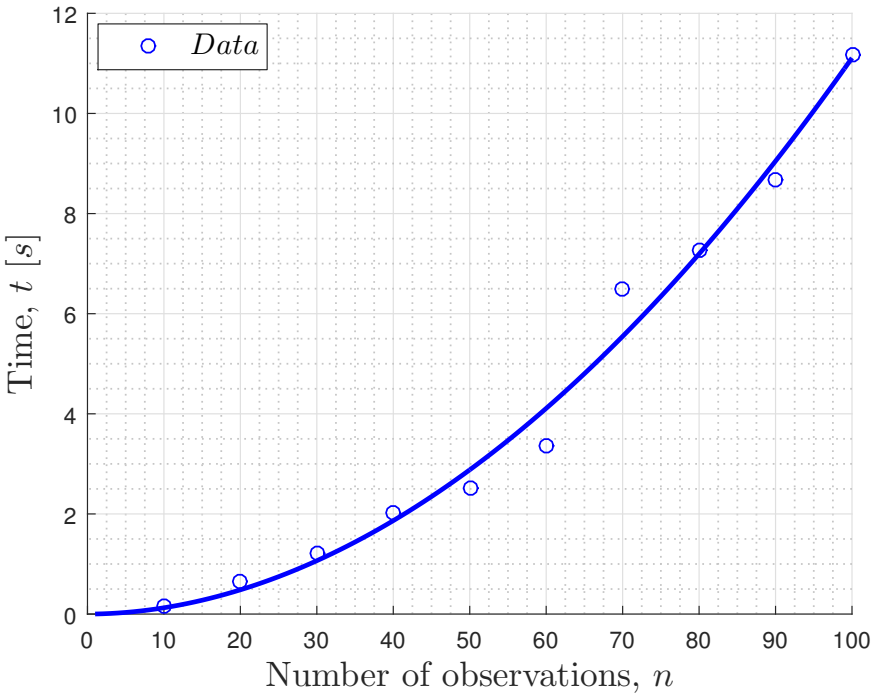


Figure 4.3: Time-size complexity for SPCR with fit $t = an^b$.

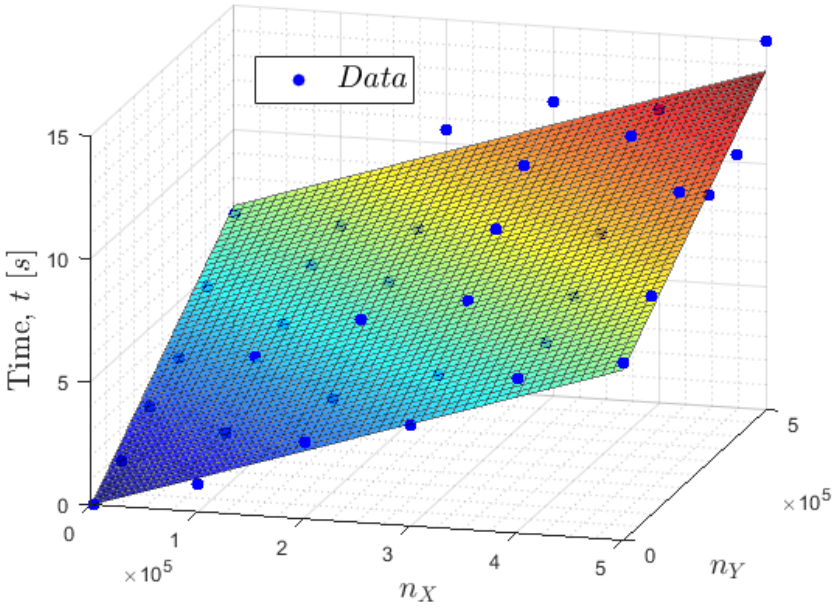


Figure 4.4: Time-size complexity for SPCR with fit $t = a(n_x + n_Y)$.

4.2.3 Conclusion

SPC regression results in far more desirable time-size complexity compared to PLS. The drawback however is that the resulting loading vectors are not evaluated to result in dually maximised cross-covariance, which depending on the data, may result in inferior predictions compared to PLS. The superior time-size complexity of SPC regression allows much larger systems to be modelled directly e.g. system sizes of $n_X, n_Y = 500,000$ and $n = 50$ can be regressed in 15 seconds as opposed to 20 times smaller systems: $n_X, n_Y = 25,000$ taking approximately 5 minutes for single component retention using PLS.

Due to time constraints, the remainder of this study is limited to the use of system sizes which result in 10-fold cross-validated models executable within 24 hours.

For PLS this equates to system sizes of $n = 50$ observations with a first component evaluation time smaller than 3 minutes:

$$\begin{aligned} 180 &\geq (4.512 \times 10^{-7}) n_X n_Y \\ n_X n_Y &\lesssim 400 \times 10^6 \\ n_X, n_Y &\approx 20,000. \end{aligned} \tag{4.13}$$

This implies roughly 6,500 predictor and response geometry points in $\mathbb{R}^3$. PLS may in some instances, depending on the geometry, require the use of simplification/sub-sampling to meet this constraint.

SPC regression may equivalently, however improbable, be limited to full component extraction times of approximately 2.5 hours:

$$\begin{aligned} 8640 &\geq (1.378 \times 10^{-5})(n_X + n_Y) \\ n_X + n_Y &\lesssim 600 \times 10^6 \\ n_X, nY &\approx 300 \times 10^6. \end{aligned} \tag{4.14}$$

This implies roughly 100 million predictor and response geometry points in $\mathbb{R}^3$.

4.3 DATA PROCESSING

This section describes the pre- and post-processing of the point-sets to enable the use of regression. Two approaches are used for this study:

1. Regression of the raw Cartesian coordinates of the registered point-sets, from here on referred to as "Cartesian space based regression".

2. Regression of the Euclidean distances of the registered point-sets, from here on referred to as "Euclidean space based regression". The motivation for this method is rotation and translation invariance i.e. this method does not require alignment techniques.

Each of these are outlined in the subsequent sections.

4.3.1 Cartesian space regression

This regression approach is based on using the Cartesian coordinates of the registered point-sets. This requires the following steps:

1. Pre-processing:

   (a) alignment of the registered point-sets (including the test/prediction sets),

   (b) sub-sampling of the aligned point-sets and

   (c) stacking of the coordinates to form a single row vector for predictor and response respectively.

2. Regression modelling and prediction.

3. Post-processing of prediction:

   (a) alignment of predicted and true sub-sampled point-sets for cross-validation and

   (b) reconstruction of the full geometry from predicted sub-sampled point-set via interpolation.

Detailed discussion of these points are given in the relevant subsequent sections.

4.3.1.1 Alignment

Due to the heuristic one-to-one correspondence of the registered geometries, alignment is facilitated easily through the use of SVD based-alignment (Section 2.3.2). Alignment of the dataset, as opposed to only a target and source, is known as Generalised Procrustes Analysis (GPA). The procedure for GPA is outlined as follows [22]:

1. A reference shape is chosen arbitrarily.

2. All observations are aligned with the reference shape (Section 2.3.2).

3. The mean shape of the superimposed observations are evaluated.

4. The Procrustes distance between the reference shape and mean, computed using Equation (2.18), is evaluated. If this is not below a threshold, the mean shape is chosen as the new reference shape and the process is repeated from step 2.

Note that the GPA formulation allows for the removal of isotropic scaling, by modification of the singular values in Section 2.3.2. Implementation of GPA in this study does not include explicit isotropic scale removal, as the regression methods employed implicitly regress/predict isotropic scaling i.e. if scaling contributes significantly to the total variance, a superposition of dominant loadings will include scaling.

4.3.1.2 Sub-sampling

Sub-sampling in this study is facilitated through using the nearest neighbours of a simplified mesh representation of the same geometry, generated through quadratic edge collapse decimation. Low curvature areas can be sub-sampled even more sparsely by further simplification through user selection. Figure 4.5 illustrates the sub-sampled points (blue) of a cranium and mandible pair using this method: the number of points are reduced from $25,000 \rightarrow 5,000$ and $12,500 \rightarrow 500$ for the cranium and mandible geometries respectively, resulting in $n_X = 15,000$ and $n_Y = 1,500$ respectively. Note the sparser sampling around the top and back of the cranium through further manual simplification. The effect of the level of simplification on accuracy is not included in this work and may require further study.

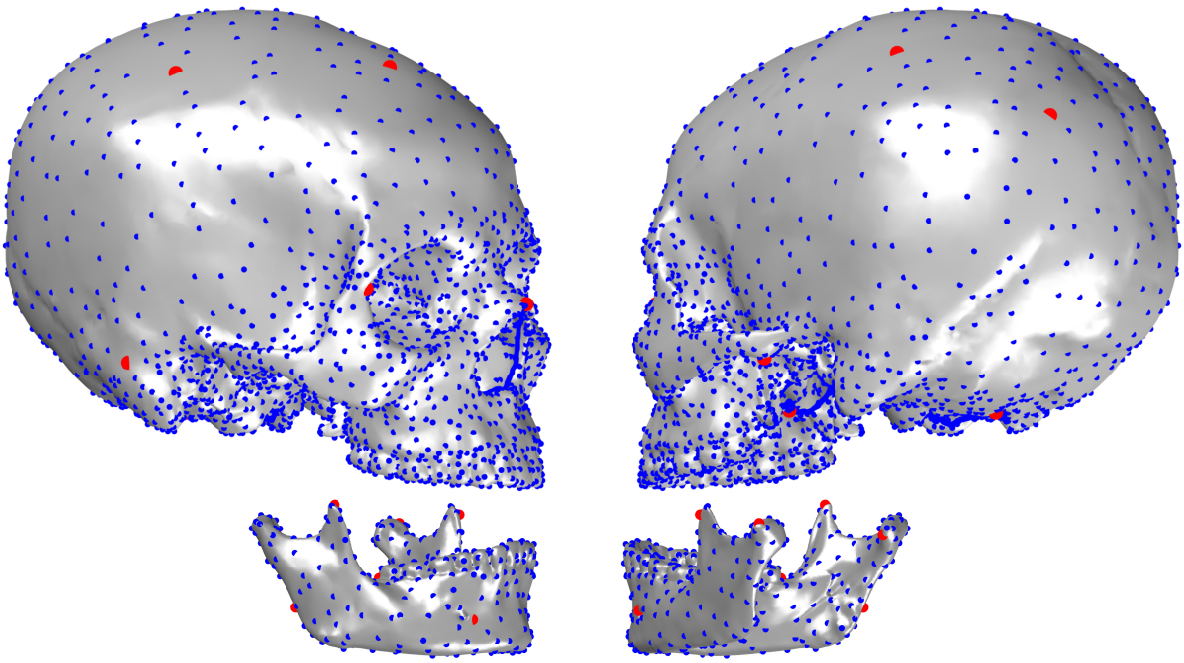


Figure 4.5: Sub-sampled cranium and mandible point-sets (blue) with reference coordinates for Euclidean distance regression approach (red).

4.3.1.3 Stacking

Stacking is used to transform the matrix of coordinates into row vector form to enable regression. Transformation of the n-th prediction point-set observation into the n-th row of a dataset matrix is obtained by the transpose of the vec operator [23]

$$\boldsymbol{X}_n = \text{vec}(\mathcal{P}_n)^T, \tag{4.15}$$

with the n-th observation point-set

$$\mathcal{P} \triangleq \{\boldsymbol{p}_i\}_{i=1}^{N}, \quad \boldsymbol{p}_i \in \mathbb{R}^m. \tag{4.16}$$

4.3.1.4 Full Geometry Reconstruction

The full geometry is reconstructed by RBF interpolation, given in Section 2.4.1, of the reference geometry full point-set, with prescribed displacement at the predicted sub-sampled point-set locations. The sub-sampled predicted point-sets are assumed to be dense enough for a linear RBF interpolation to be appropriate/accurate.

4.3.2 Euclidean space regression

This regression approach is based on using Euclidean pairwise distances w.r.t. reference points of the registered point-sets. This requires the following steps:

1. Pre-processing:
   (a) sub-sampling of the point-sets and selection of reference points,
   (b) calculation of the pairwise distance matrix w.r.t. reference points and
   (c) stacking of the distance matrix to form a single row vector for predictor and response respectively.
2. Regression modelling and prediction.
3. Post-processing of prediction:
   (a) reconstruction of the predicted sub-sampled geometry from the predicted pairwise distance matrix,
   (b) alignment of predicted and true sub-sampled point-sets for cross-validation and
   (c) reconstruction of the full geometry from predicted sub-sampled point-set via interpolation.

Detailed discussion of these points are given in the relevant subsequent sections.

4.3.2.1 Sub-sampling

Same as described in Section 4.3.1.2.

4.3.2.2 Reference points

Due to the high dimensionality of inter-point distances versus Cartesian coordinates for a given point-set in $\mathbb{R}^3$ ($(N^2 - N)/2$ vs. $3N$ respectively), regression time complexity becomes problematic for even small point-sets. Considering Equation (4.13) limits the number of points for the use of the full inter-point distance matrix to $N = 200$ for

equal sized predictor and response point-sets $n_X = n_Y = (N^2 - N)/2 \approx 20,000$. This may arguably only be adequate for the representation of simplistic 3D geometries. It is for this reason that the use of pairwise distances of the sub-sampled point-set w.r.t. an inclusive reference subset, N_r is rather considered. The number of pairwise distance entries then reduces to $N \times N_r - (N_r^2 - N_r)/2$, where the subtractive term represents the repeated symmetric distances. For $N_r \ll N$, the complexity reduces to approximately linear for the pairwise distances vs. quadratic for the inter-point distances, $N \times N_r$ vs. N^2. With $N_r = 10$ the maximum number of sub-sampled points increases tenfold to $N \approx 2,000$ for equal sized predictor and response point-sets. This proposal is solely implemented due to the necessity to have the capability of handling more complex geometries/larger sub-sampled point-sets. The number and choice of reference points and the effect on prediction accuracy is therefore not part of this work and may constitute work for further study. Selection of the reference points in this study is done through random MC sampling of N_r reference points of the sub-sampled point-set and selecting the sample with the largest minimum inter-point distance of the sampled reference points. This heuristic aims to select the most uniform spread of reference points in the sub-sampled point-set. Figure 4.5 illustrates $N_r = 10$ reference points (red) of the sub-sampled point-set (blue) obtained through this method.

4.3.2.3 Stacking

Stacking is used to transform the pairwise-distance matrix into row vector form to enable regression. Transformation of the n-th pairwise distance matrix observation into the n-th row of a dataset matrix is also obtained by the transpose of the vech operator [23], which only keeps the non-repeated entries:

$$\boldsymbol{X}_n = \text{vech}(\boldsymbol{D}_n)^T, \tag{4.17}$$

with the n-th pairwise distance matrix $\boldsymbol{D}_n$ given by

$$D_{ir} = \|\boldsymbol{p}_i - \boldsymbol{p}_r\|, \tag{4.18}$$

and i the i-th point and r the r-th reference point in the sub-sampled observation point-set $\mathcal{P}_n$ respectively.

4.3.2.4 Sub-sampled point-set reconstruction from pairwise distances

Once the pairwise distances have been predicted through the regression model, the sub-sampled point-set of Cartesian coordinates which relates to the same pairwise distances needs to be generated. This is accomplished through metric multidimensional scaling (MMDS), which seeks the point-set, $\mathcal{P}$ in $\mathbb{R}^m$, through minimisation of a "stress" function:

$$\sigma(\mathcal{P}) = \sum_{i<j} w_{ij} \left(\hat{D}_{ij} - ||\boldsymbol{p}_i - \boldsymbol{p}_j|| \right)^2, \tag{4.19}$$

with $\hat{D}_{ij}$ the predicted inter-point distance matrix. Note that $\hat{D}_{ij}, w_{ij} = 0$ for $j \neq r$ as only predictions for the pairwise entries D_{ir} of the inter-point distance matrix entries D_{ij} are available. The "stress" function is minimised through an iterative steepest descent approach [24].

4.3.2.5 Full geometry reconstruction

Same as described in Section 4.3.1.4.

CHAPTER 5

REGRESSION APPLICATION

This chapter demonstrates proof of concept for a registration based prediction strategy. The sequence of steps followed are:

1. Generate a true correspondence dataset with known structure/covariance,

2. simulate the absence of true correspondence by using arbitrary point-sets which represent the same geometric data as the registration targets and

3. quantitatively and qualitatively evaluate and discuss the prediction error due to

   (a) the proposed regression space limitations, which are evident when using true correspondence, and

   (b) the use of heuristically generated registration correspondence.

5.1 PROCEDURE

The procedure for validation of the proposed registration-regression strategy consists of the following steps:

1. The proposed strategy is tested by means of creating an artificial geometry dataset, where the underlying structure is embedded/imposed by means of using the same quadratic deformation mapping of an arbitrarily chosen prediction and corresponding response geometry i.e. cranium and mandible.

2. The geometries are generated using Monte Carlo sampling of the same quadratic deformation mapping given by Equation (3.2) in Chapter 3. The same sampled

deformation matrix $\mathbf{C}$ is used for the deformation mapping of both the prediction and response target geometries.

3. Registration of the prediction and response target geometries are performed using a different mesh configuration of the corresponding geometry. This is to simulate the absence of true correspondence.

4. Registration is performed for two basis functions i.e. cubic and TPS, for four choices for the number of centres and iterations: $N_b, k = 25, 50, 100, 200$. Each of these registrations are checked to ensure that the resulting geometries are visually accurate approximations of the target geometry. This is done to investigate the influence of the in-surface true correspondence errors on the prediction error. The registration results at the sub-sampled locations are given in Section 5.2.1.

5. The registration results are sub-sampled, as depicted in Figure 5.1, to keep the cross-validation computationally tractable. This is done using a lower density mesh of the reference geometry and selecting point indices using a nearest neighbour search of the reference dense mesh. The number of sub-samples are 5000 and 500 for the prediction and response point-sets respectively.

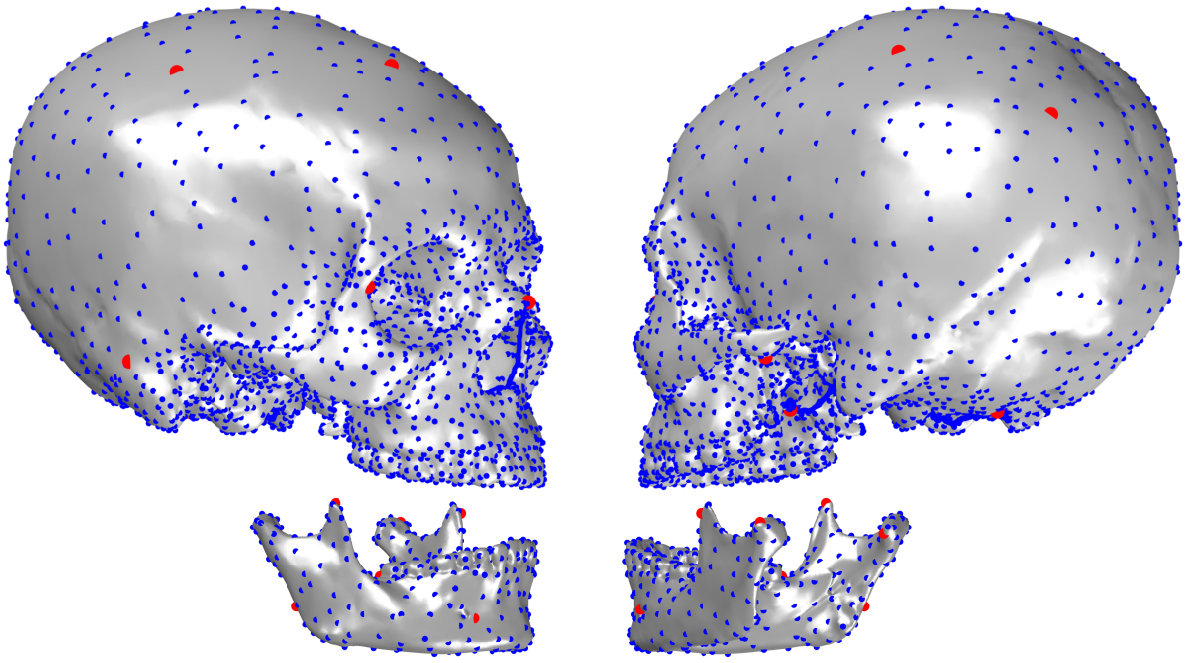


Figure 5.1: Sub-sampled reference cranium/predictor and mandible/response point-sets (blue) with reference coordinates for Euclidean distance regression approach (red).

6. The different sub-sampled registration point-sets are aligned using Procrustes alignment. Two different datasets are generated using the sub-sampled true correspondence point-set i.e. one with Procrustes alignment and the other unaltered.

7. The Cartesian and Euclidean distance based regression approach datasets are then generated accordingly for the 8 different registrations' sub-sampled point-sets. This is also done for the true correspondence configuration that will serve as a benchmark.

8. The number of training observations (MC samples) is chosen to be significantly larger than the expected number of retained components for an exact prediction of the Cartesian true correspondence dataset (which is expected to be $N_{PC} = 27$ retained components). This is because the Cartesian true correspondence artificial geometries are represented by superposition of 27 linearly independent quadratic polynomial transformations of the geometries' point-sets in $\mathbb{R}^3$. Due to the non-linear transformation of the Cartesian coordinates, the Euclidean distance based approach is not expected to be able to produce an exact prediction i.e. anything other than the full component retention, $N_{PC} = N_{\text{train}} - 1$, will render an inexact prediction. This is verified by SVD of the cross-covariance matrix with centered $\boldsymbol{X}$ and $\boldsymbol{Y}$, $\boldsymbol{X}_0$ and $\boldsymbol{Y}_0$ accordingly, given by:

$$\boldsymbol{P} = \boldsymbol{X}_0^T \boldsymbol{Y}_0. \tag{5.1}$$

The squared singular values represent the variance explained by each component. The cumulative explained and unexplained variance as a fraction of the totals vs. the retained number of components are presented in Figure 5.2 for training sets of $N_{train} = 40$ observations. Figure 5.2 clearly shows that, as expected, the Cartesian based approach only has $N_{PC} = \text{rank}(\boldsymbol{P}_{\text{Cart}}) = 27$ unique components whereas the Euclidean distance approach has $N_{PC} = \text{rank}(\boldsymbol{P}_{\text{Eucl}}) = N_{train} - 1$. Although it seems tempting to directly compare the explained variance between the two approaches, direct comparison should only be done in a consistent space, as done for the prediction results in Section 5.2. Training sets with $N_{train} = 40$ observations with Cross-Validation (CV) datasets of $N_{test} = 10$ observations are consequently chosen for the remainder of this study, as $N_{train} > N_{PC}$, for the Cartesian true correspondence dataset.

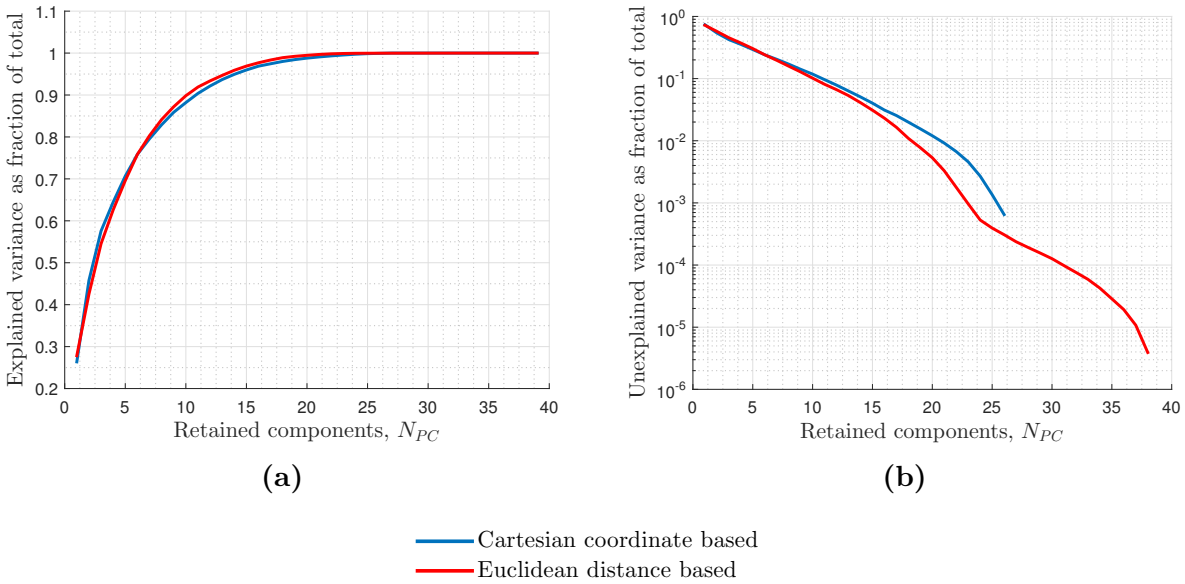


Figure 5.2: (a) Explained and (b) unexplained variance as fraction of total variance vs. the retained number of principal components.

9. The datasets are submitted to PLS and SPC regression and corresponding predictions are generated for $N_{test} = 10$ CV test observations. The Euclidean distance prediction results are transformed back to Cartesian coordinates by means of MMDS, Section 4.3.2.4, to enable comparison of the two prediction results for both regression spaces. The Cartesian space, mean point-wise, CV prediction error estimates for both approaches are presented in Section 5.2.2.

5.2 RESULTS

5.2.1 Registration Results

The point-wise error distributions for the different registrations of the MC sampled geometries are given in Figures 5.3 and 5.4 for the sub-sampled point-sets of the cranium and mandible geometries. The corresponding normal weighted error distributions are also given. The corresponding mean and maximums of the error distributions for the different registrations are given in Table 5.1. From these the following observations are made:

1. The mean of the point-wise and normal weighted error distributions of the cranium and mandible geometries are comparable for the different registration parameters. These decrease with increased registration complexity, as expected.

2. The mean of the normal weighted error distributions suggest the out-of-surface errors account for approximately $20-25\%$ of the overall/un-weighted error. This results in near imperceptible visual registration difference (< 0.2mm).

3. The maximums of the point-wise error distributions for the different registrations are not comparable between cranium and mandible geometries i.e. maximums of 25.0019 and 0.3934mm respectively for the best registrations. There is also no significant decrease in the maximums of the distributions with increasing registration complexity. This is attributed to using the same quadratic mapping coefficients, which results in much larger complex displacements of the cranial geometries due to their larger relative size (See Figure 5.1). The registration algorithm results in distorted correspondence in some instances due to the fact that correspondence is built up using a nearest neighbour search in areas of large complex displacement. The comparatively large maximums are thus not unexpected.

4. Although large, the normal weighted error distribution maximums for the cranium registrations are a poor indication of visually perceptible out-of-surface error. This is because this error weighting is only appropriate for either justifiably "small" error or regions of near zero curvature, where out-of-plane is approximately out-of-surface.

For qualitative evaluation consider the comparatively worst considered registration (cubic basis function with $N_b, k = 25$) for an arbitrary observation cranium and mandible pair, within the MC sample, in Figures 5.5 and 5.6 respectively. These clearly show the registrations to be good visual representations of the target geometries relative to the difference between the source and target geometries.

Table 5.1: Error distribution mean and maximums for various registrations.

Geometry	Registration	N_b, k	ε [mm]		$\varepsilon \cdot w_n$ [mm]	
			mean	max	mean	max
Crania	None	0	7.5187	33.2523	–	–
	Cubic	25	0.7343	25.7047	0.1599	8.3015
		50	0.4286	25.5117	0.0705	7.3887
		100	0.2717	25.3749	0.0433	7.6515
		200	0.1730	25.9126	0.0276	7.2498
	TPS	25	0.9433	25.6560	0.2598	7.9734
		50	0.5350	24.6643	0.1191	7.0652
		100	0.2885	24.2036	0.0572	6.9827
		200	0.1543	25.0019	0.0330	5.8906
Mandibles	None	0	3.1997	19.2893	–	–
	Cubic	25	0.2773	4.9131	0.0513	1.3939
		50	0.1003	4.3637	0.0200	0.4283
		100	0.0518	3.8805	0.0146	0.3108
		200	0.0555	2.8264	0.0142	0.1193
	TPS	25	0.3613	4.4665	0.0926	1.4561
		50	0.1339	4.0625	0.0409	0.6304
		100	0.0689	2.7167	0.0267	0.4118
		200	0.0530	0.3934	0.0151	0.2511

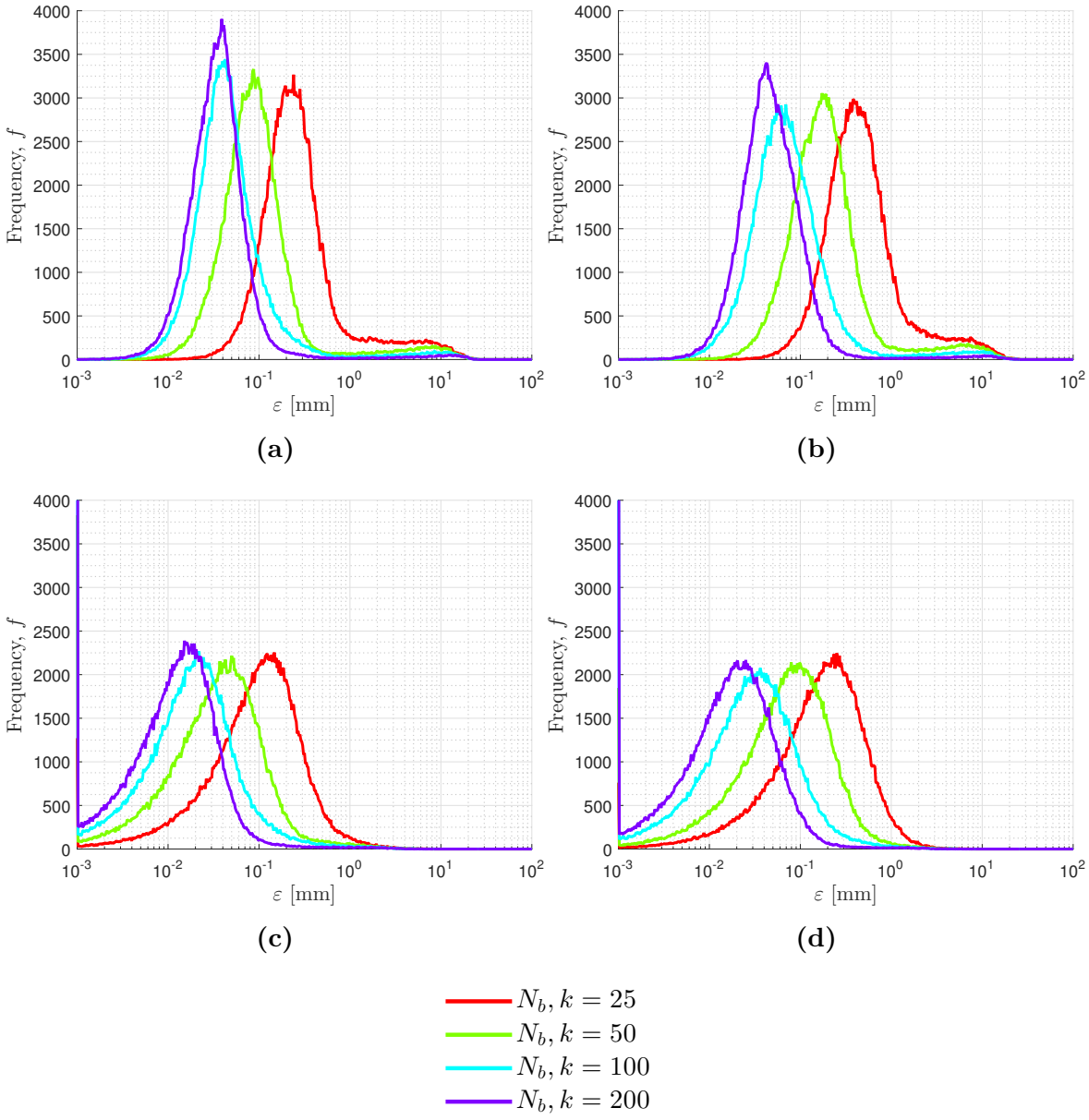


Figure 5.3: Registration error distributions for the Cranium MC sample using (a) Cubic and (b) TPS basis functions. Normal weighted registration error distributions using (c) Cubic and (d) TPS basis functions.

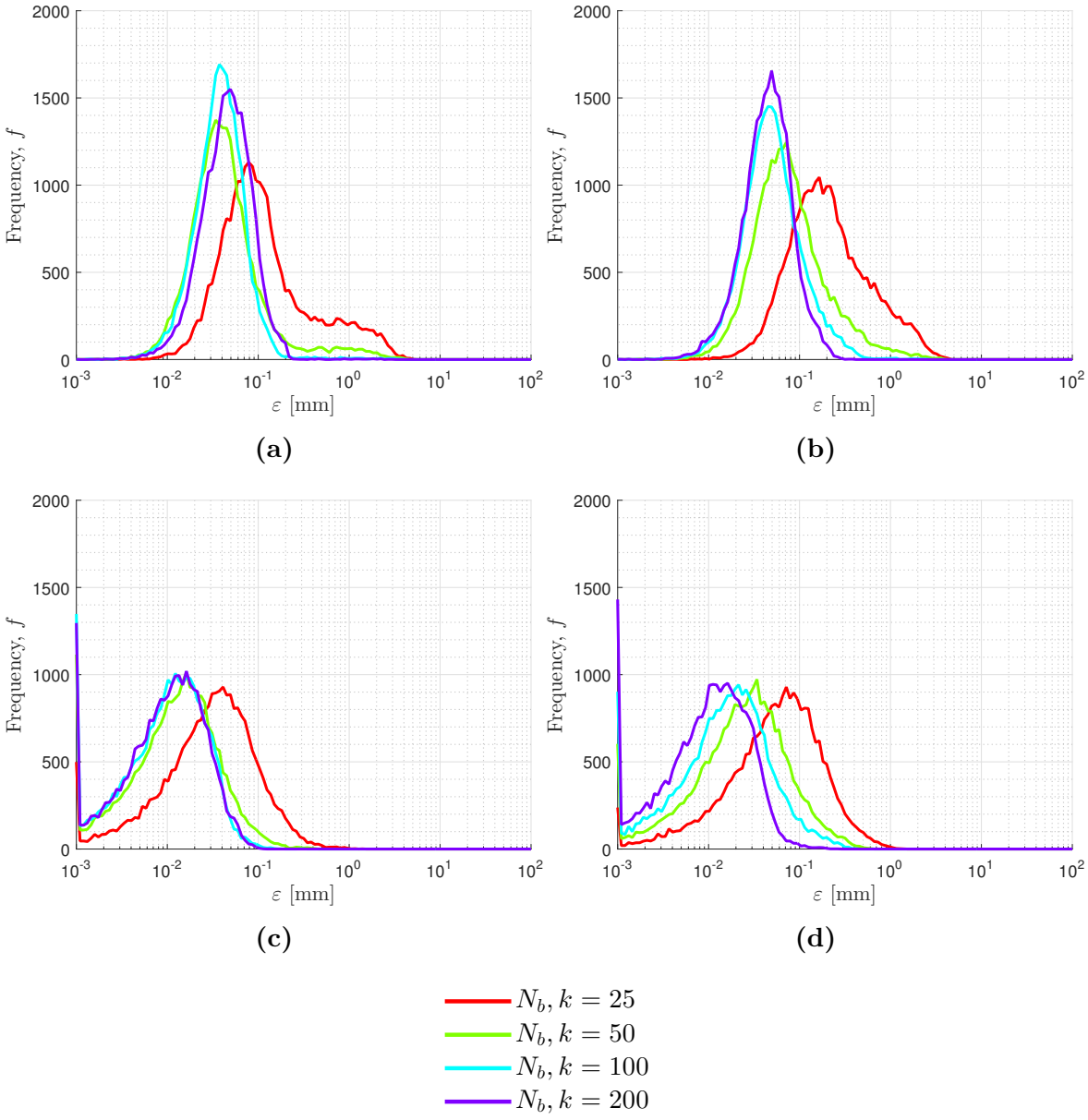


Figure 5.4: Registration error distributions for the Mandible MC sample using (a) Cubic and (b) TPS basis functions. Normal weighted registration error distributions using (c) Cubic and (d) TPS basis functions.

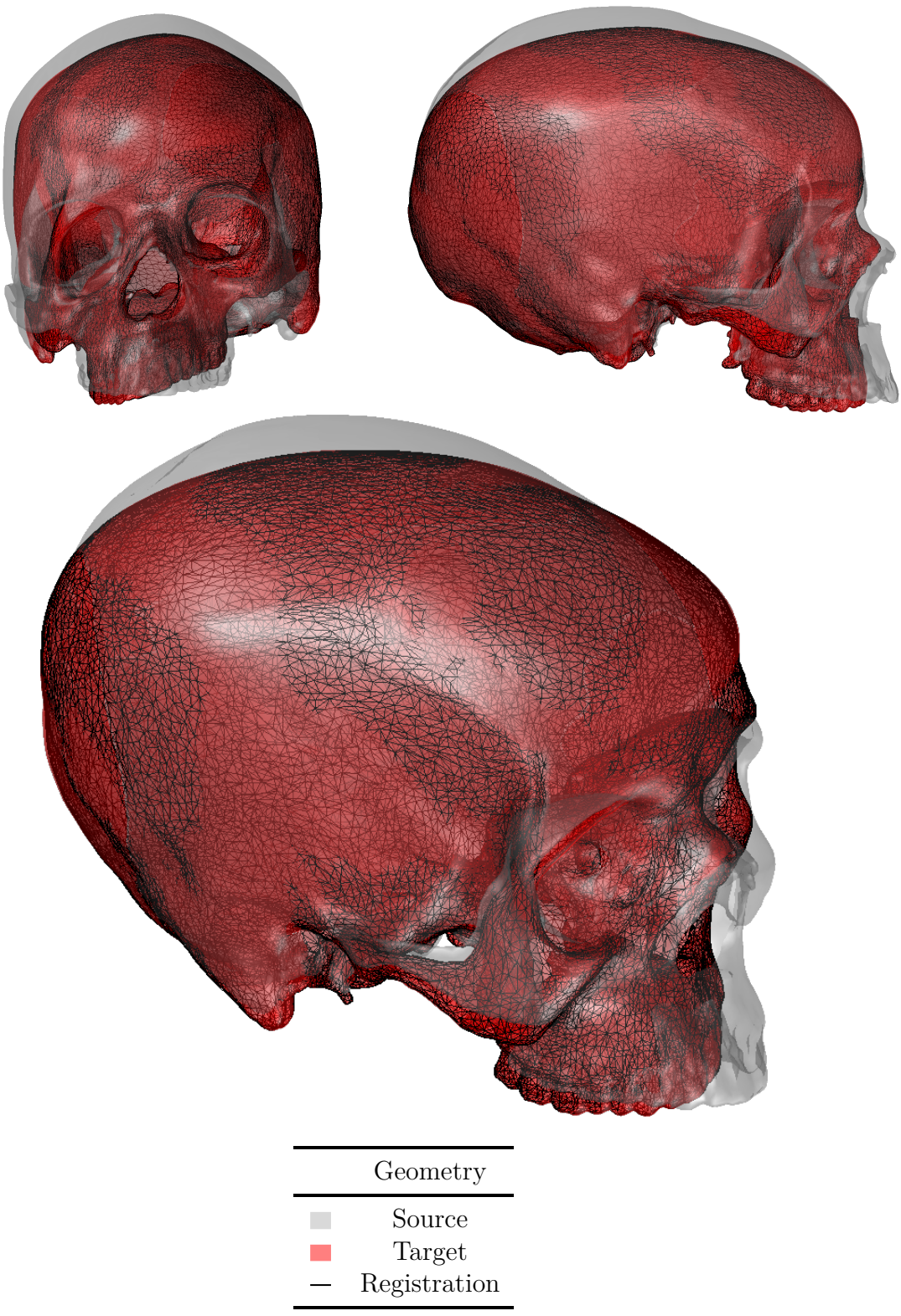


Figure 5.5: Cranium registration result of an arbitrary observation in the MC sample for the use of a cubic basis function and $N_b, k = 25$.

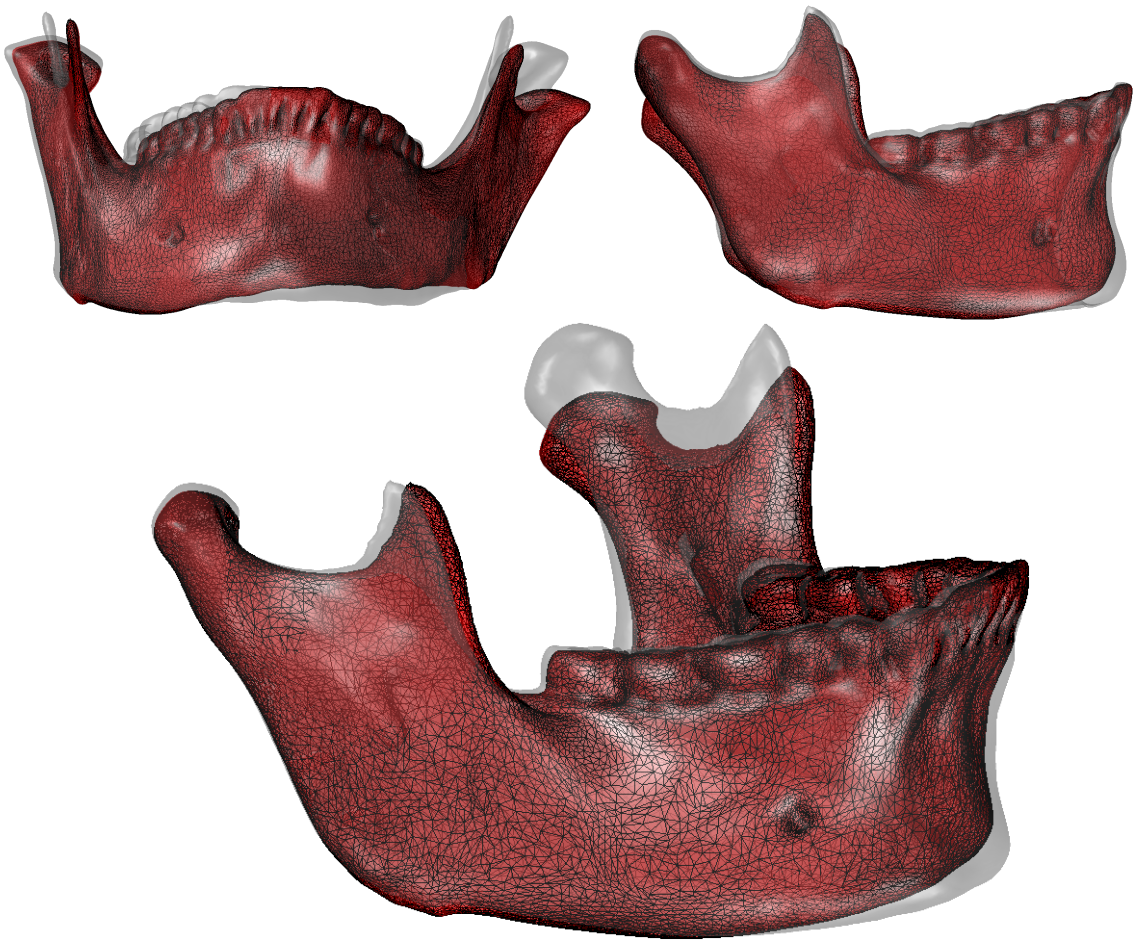


Figure 5.6: Mandible registration result of an arbitrary observation in the MC sample for the use of a cubic basis function and $N_b, k = 25$.

5.2.2 Prediction Results

The CV results of the Cartesian and Euclidean PLS and SPC regressions are given in Figures 5.7 and 5.8 respectively, after reconstruction in $\mathbb{R}^3$. This is given as the mean of the point-wise error over the $N_{test} = 10$ test samples vs. the number of retained components. The results of the Cartesian PLS and SPC regressions in Figure 5.7 are summarised as follows:

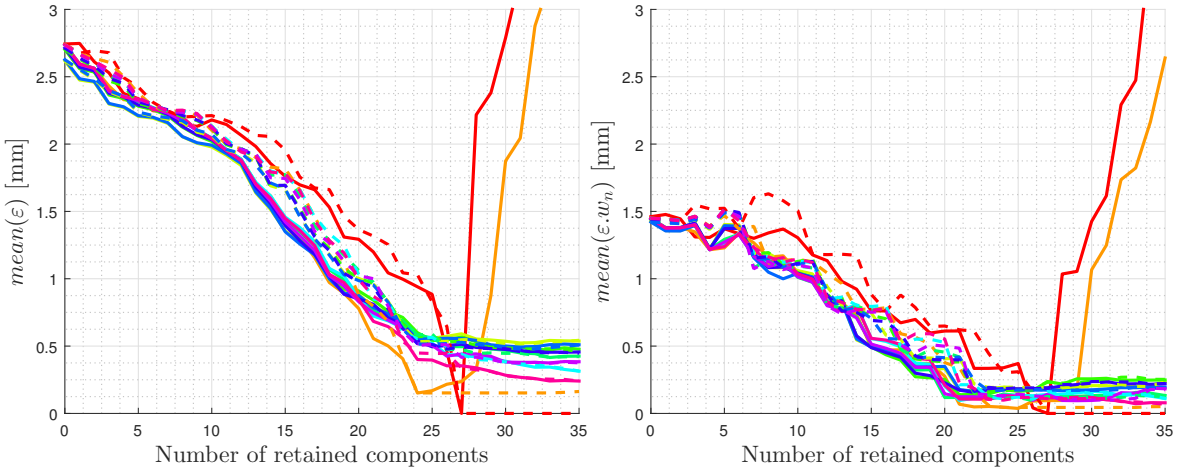
1. PLS and SPC regression of the true correspondence dataset are exact when using $N = 27$ retained PLS components, as expected. The result of this dataset is only included for interest's sake and will never be available for real data.

2. PLS and SPC regression of the true correspondence Procrustes aligned dataset have associated error and are unable to exactly predict the solution. This is attributed to unexplained covariance due to inconsistent alignment between prediction and response geometries i.e. a cranium and mandible pair will not undergo the same rotation when being aligned with their respective geometry datasets using Procrustes alignment. This is presumably only due to shear-like deformations, which are rotated to result in minimum point-wise variance during Procrustes alignment. This is expected to be the theoretical minimum mean point-wise error of mean$(\varepsilon) = 0.15$mm.

3. All the registered datasets result in minimum error for full component retention. The prediction accuracy is increased for more complex registrations, as expected, with mean point-wise error ranging from mean$(\varepsilon) \approx 0.5$ to 0.25mm for the least to most complex registrations respectively ($N_b, k = 25$ and 200).

4. The normal weighted mean point-wise error shows that out-of-surface error, or visually perceptible error, equals approximately half the unweighed error towards the respective minimums. Note that this weighting is not completely appropriate for the relatively larger errors, due to the implicit assumption that in-plane errors are in-surface errors.

5. The final prediction errors obtained through the registration-regression strategy are small compared to the imposed deformations, typically $10 - 20\%$ compared to the mean geometry mean$(\varepsilon) \approx 2.75$mm (zero retained components).



(a) **(b)**

	Registration description	N_b, k
—	True Correspondence	–
—	True Correspondence, aligned	–
—	Cubic	25
—		50
—		100
—		200
—	TPS	25
—		50
—		100
—		200

Figure 5.7: Cartesian regression - (a) Mean point wise error and (b) mean normal weighted point wise error vs. retained number of components. PLS and SPC regression results are indicated by solid and dashed lines respectively.

The results of the Euclidean PLS and SPC regressions in Figure 5.8 are summarised as follows:

1. PLS and SPC regression of the true correspondence dataset are not capable of exactly prediction the solution and plateaus after $N = 24$ retained components. This is due to the non-linear transformation of the Cartesian space to Euclidean distances. The linear deformations are expected to occupy a larger space after the transformation (higher rank cross-covariance matrices) and thus requires more training observations/components to better span the space for improved predictions.

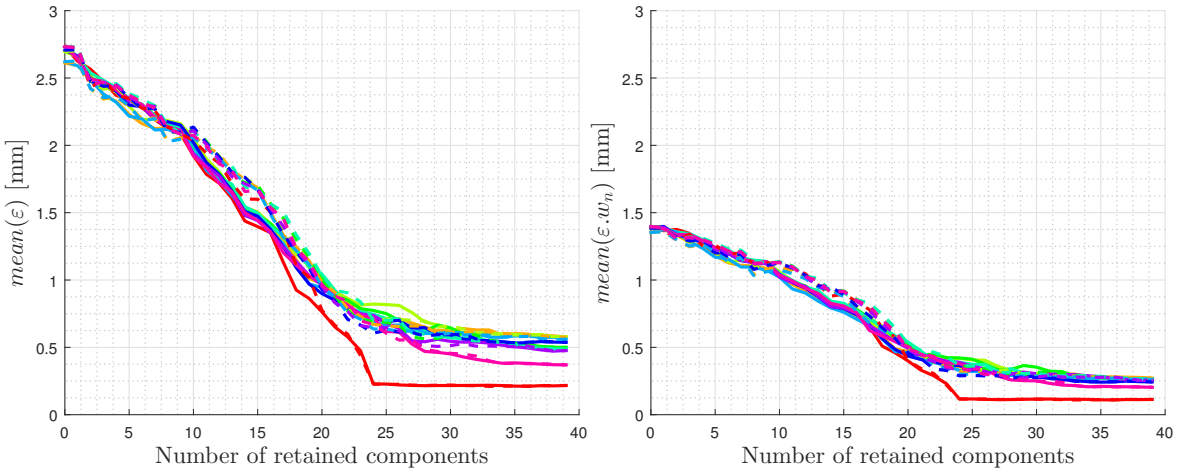
2. All the registered datasets result in minimum error at full component retention. The prediction accuracy is increased for more complex registrations, as expected, with mean point-wise error ranging from mean$(\varepsilon) \approx 0.6$ to 0.37mm for the least to most complex registrations respectively ($N_b, k = 25$ and 200).

3. Again the normal weighted mean point-wise error shows that out-of-surface error, or visually perceptible error, equals approximately half the unweighed error towards the respective minimums.

4. The final prediction errors obtained through the registration-regression strategy are also small compared to the imposed deformations, all between $15 - 22\%$ compared to the reconstructed geometry of Euclidean distance mean, mean$(\varepsilon) \approx$ 2.75mm (zero retained components).

5. The reconstructed mean point-wise errors of the Euclidean regression predictions are $10 - 25\%$ higher than the corresponding Cartesian regression based predictions.

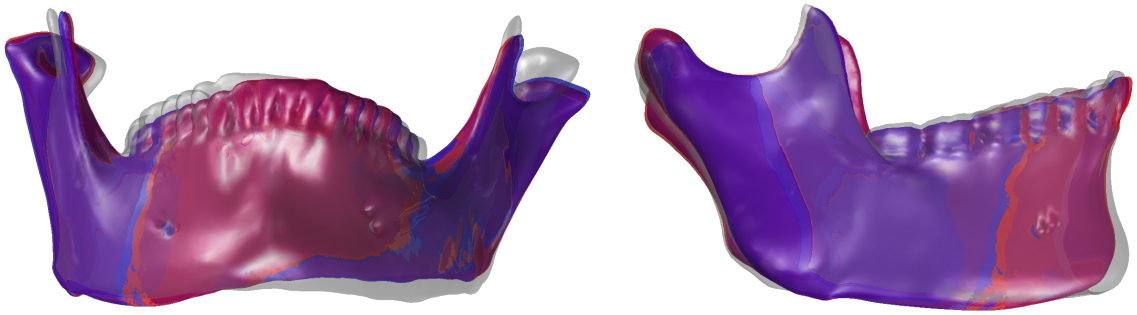
Qualitative comparison of the regression prediction results for the quantitatively best and worst registration datasets are given in Figure 5.9. These are the predictions for the corresponding arbitrarily chosen cranium and mandible pair in Figures 5.5 and 5.6. The decreased prediction error with increased registration complexity intuitively results in decreased visual error. This is however barely discernible due to the small relative magnitude. These are only included for reference, as visual cross-validation is not realistically tractable.



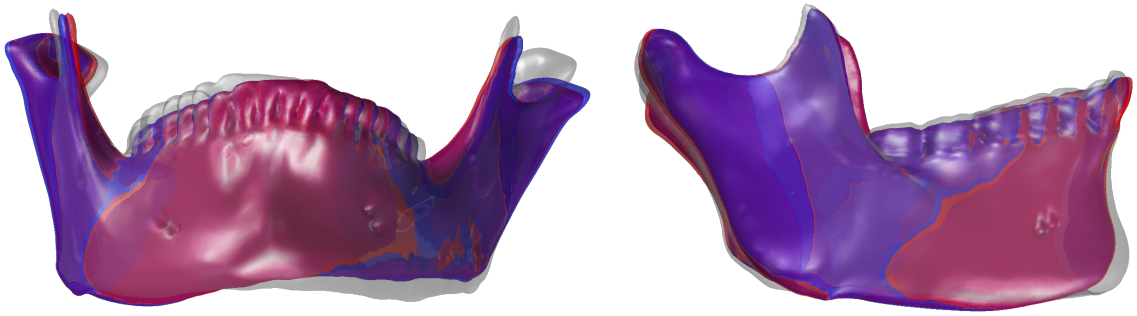
(a) **(b)**

	Registration description	N_b, k
—	True Correspondence	–
—	Cubic	25
—		50
—		100
—		200
—	TPS	25
—		50
—		100
—		200

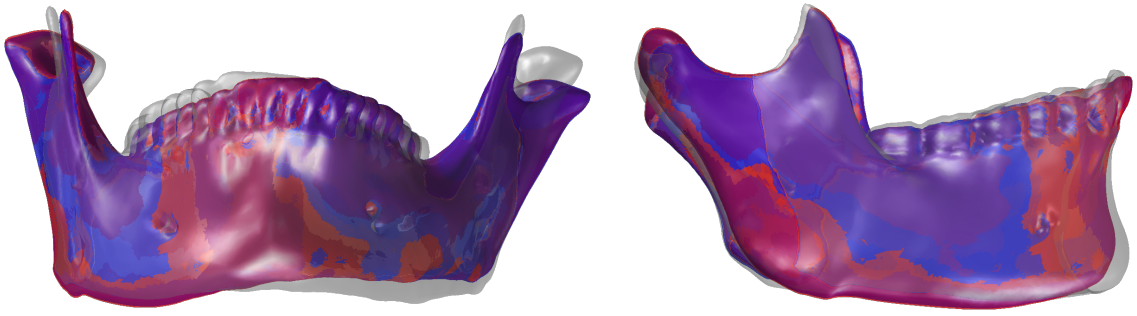
Figure 5.8: Euclidean regression - (a) Mean point wise error and (b) mean normal weighted point wise error vs. retained number of components. PLS and SPC regression results are indicated by solid and dashed lines respectively.



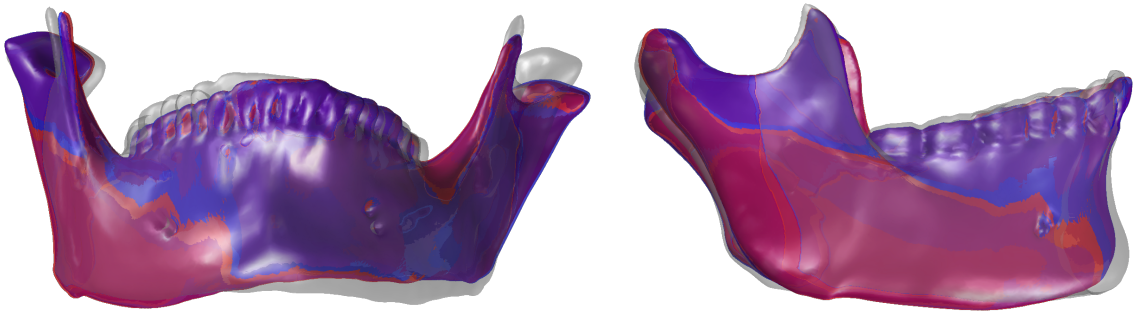
Cartesian, Cubic, $N_b, k = 25$



Euclidean, Cubic, $N_b, k = 25$



Cartesian, TPS, $N_b, k = 200$



Euclidean, TPS, $N_b, k = 200$

Figure 5.9: Qualitative results for various registrations with targets (red), prediction (blue) and mean (grey) geometries.

5.3 CONCLUSION

In this chapter the use of a proposed registration-regression strategy on an artificial dataset constructed from Monte Carlo sampling of quadratic deformation is demonstrated. The study found the registration-regression strategy to be capable of good qualitative and quantitative predictions for at least quadratic deformations. It was also found that prediction results do not differ significantly for varying degrees of registration complexity and basis functions. This suggests that the registration complexity required for a good visual registration to be adequate. This is also beneficial for application on real data, as this acts to interpolate in missing and inconsistent regions of geometries e.g. damage and tooth loss. Although the Cartesian regression space prediction results were found to be slightly more accurate, this result is specific to this study, as the embedded covariance was arbitrarily based on deformation mapping of the Cartesian coordinates. The study also found both PLS and SPC regression to have similar predictive accuracy.

Comparison of the two regression space approaches is summarised as follows:
The Cartesian regression space based approach is comparatively computationally economical, requiring very little pre- and post processing. It is however limited to geometric datasets with small magnitude shear-like deformation, as this results in unexplained covariance due to inconsistent rotation alignment. The Euclidean regression space based approach is rotation invariant, but relies heavily on sub-sampling to remain computationally tractable, depending on the regression algorithm. In addition, the computational cost due to having to reconstruct/reverse-map predictions is significant. This approach may also, depending on the data and regression algorithm, result in additional computational cost due to possible higher rank cross-covariance, which requires extraction of more components to obtain similar prediction accuracies. More observations may also be necessary to obtain similar accuracies due to the larger space spanned following this approach.

CHAPTER 6

VIC CRANIUM-MANDIBLE DATASET APPLICATION

This chapter demonstrates the application of the automated mesh registration based prediction strategy on an actual cranium and mandible dataset. This dataset was obtained by the University of Pretoria Department of Anthropology with the specific purpose of mandible reconstruction to aid in forensics. The South African Police Service (SAPS) Victim Identification Centre (VIC) has a growing archive of cases with craniums without matching mandibles [25]. This results in most of these cases remaining unclosed since facial reconstruction cannot to be performed. To facilitate identification through digital facial reconstruction, the VIC presented a requirement for a technique to predict mandibles when only the craniums are available.

6.1 DATASET

The dataset was obtained from the Pretoria Bone Repository and consists of 50 skulls (cranium and mandible pairs) of the South African black male population group [20]. This population group was identified by the SAPS VIC as the most representative group of their archive demographics [25]. The bone dataset was scanned using Micro-focus Computed Tomography (CT) at the South African National Centre for Radiography and Tomography, Necsa [26]. The resulting volume files are converted to a triangulated mesh format and the outer visible surface is extracted using an ambient occlusion filter in Meshlab [27]. This is done to simulate a surface scan, which is the only scanning technology SAPS VIC has procured for digital facial reconstruction [25].

6.2 REGISTRATION

The dataset of 50 complete skulls was filtered to exclude cranium and mandible pairs with significant damage or a majority of missing teeth. Although the least square RBF registration is less sensitive to locally damaged areas, as compared to full RBF interpolation, these areas still have to represent an insignificant portion of the squared global displacement correspondence. A simple way of reduced fitting in these regions is to reduce the number of RBF centres, which results in reduced approximation complexity. Therefore each registration has it's own number of RBF centers which results in the best visual registration. Figure 6.1 illustrates a typical example of a cranium and mandible pair in the dataset where a lower complexity approximation is used to result in reduced fitting of missing teeth. A cubic basis function was used for registration of the dataset with the number of centres varied between $250 \leq N_b \leq 400$ to obtain the best visual registration for each observation.

During registration of the crania, significant correspondence distortion was generated in areas surrounding the nasal cavity in some of the observations. This may be attributed to a combination of physical differences and CT density thresholding that results in dissimilar structures within the nasal cavity. Either way, the dissimilarities within the nasal cavity result in distorted correspondence due to the nearest neighbour correspondence mechanism of the registration technique. Resolving this whilst keeping the internal nasal structures would require modification of the registration algorithm. This is not considered in this study due to time considerations and may constitute work for further study. Removal of the internal nasal structures by post-processing was initially attempted, but proved to be too time-consuming considering time limitations. Due to only being in the developmental stage of this project for the SAPS VIC, proof of concept is considered as the primary goal. The number of observations were thus limited to 33, omitting registrations with correspondence distortion due to internal nasal cavity structures. A procedural change in the database construction for the final application may thus include covering the nasal cavity to prevent the surface scanner from scanning the internal nasal cavity structures. This would prevent the time consuming process of having to remove these structures through post-processing to recover undistorted correspondence.

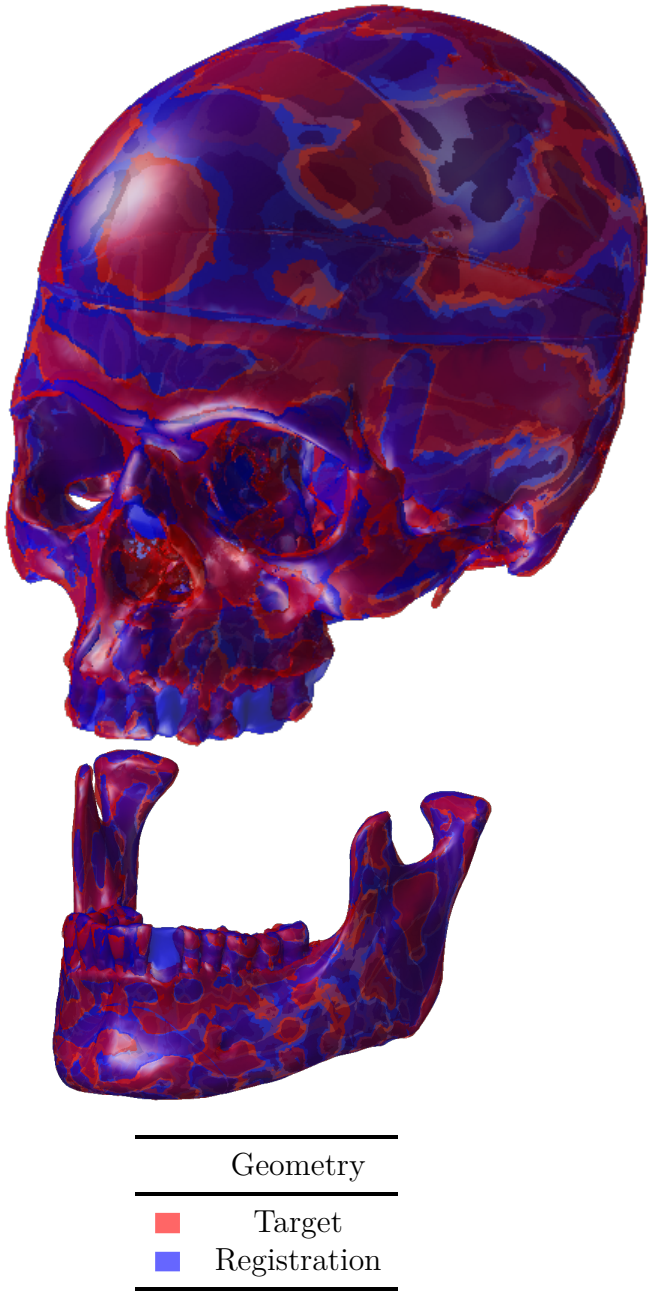


Figure 6.1: Typical cranium-mandible pair registrations with the number of RBF centres chosen to result in fill-in of missing teeth. The reduced number of centres still provide qualitatively good visual registrations.

6.3 REGRESSION

The registered geometries are sub-sampled as illustrated in Figure 5.1 to meet the cross-validation time tractability requirements, as set out in Chapter 4. The coordinates are stacked and submitted to both PLS and SPC regression and cross-validated by k-fold cross validation with $k = 11$. An additional model for both PLS and SPC regression is included where each variable/feature is normalised to result in unit variance. Variance normalisation in this instance equates to obtaining loading vectors through eigen decomposition of the correlation matrices, as opposed to the covariance matrices. This aims to prevent relative greater variable/feature variances from dominating the primary/principal loading vectors.

6.4 RESULTS

The mean point-wise error for the 11-fold cross validation regression results of both the Cartesian and Euclidean regression models are presented in Figures 6.2 and 6.3 respectively. The following is observed from these figures:

1. All regression models, irrespective of space, result in $\approx 5-9\%$ improvement over only mean prediction i.e. using no components.

2. For the optimally retained components, the Cartesian based regression models do consistently better than the optimal Euclidean regression models. These improvements are however only marginal and the two methods remain comparable. This is of great significance as the Euclidean regression models, as described in Chapter 4, are computationally expensive due to space scaling, as well as further post-processing in the form point-set reconstruction.

3. Models with normalised unit variance result in marginally improved model performance, except for the PLS regression of the Euclidean based approach.

4. In this application, SPC regression models with normalised unit variance result in monotonically decreasing mean point-wise errors up to the CV minimum.

5. PLS and SPC regression of the Cartesian space, results in comparable minimum mean point wise errors, although the optimal number of components may differ significantly. This is of great significance as SPC regression has insignificant computational cost compared to PLS for the system sizes of interest.

The comparatively best and worst predictions as compared to using the geometric means, are visualised in Figures 6.4 and 6.5 respectively. These are for Cartesian SPC regression with normalised unit variance and the 5 retained component 11-fold CV minimum. The comparatively best is a vast visual improvement over the geometric mean, while the geometric mean is not a vast improvement over the comparatively worst. For visual comparison, the mean point wise error is visualised in Figure 6.6 for different scalings of the primary loading vector for Cartesian SPC regression of the complete dataset. This serves as an overly simplistic gauge for visualising mean point-wise errors in the expected error range and may not be representative of complex actual errors.

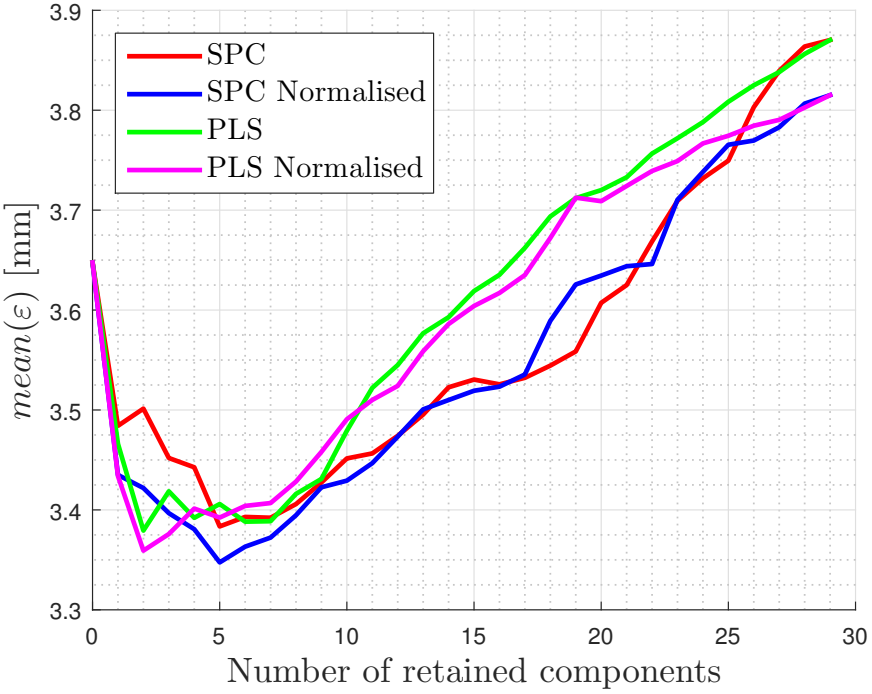


Figure 6.2: Mean point wise error vs. retained components for Cartesian regression.

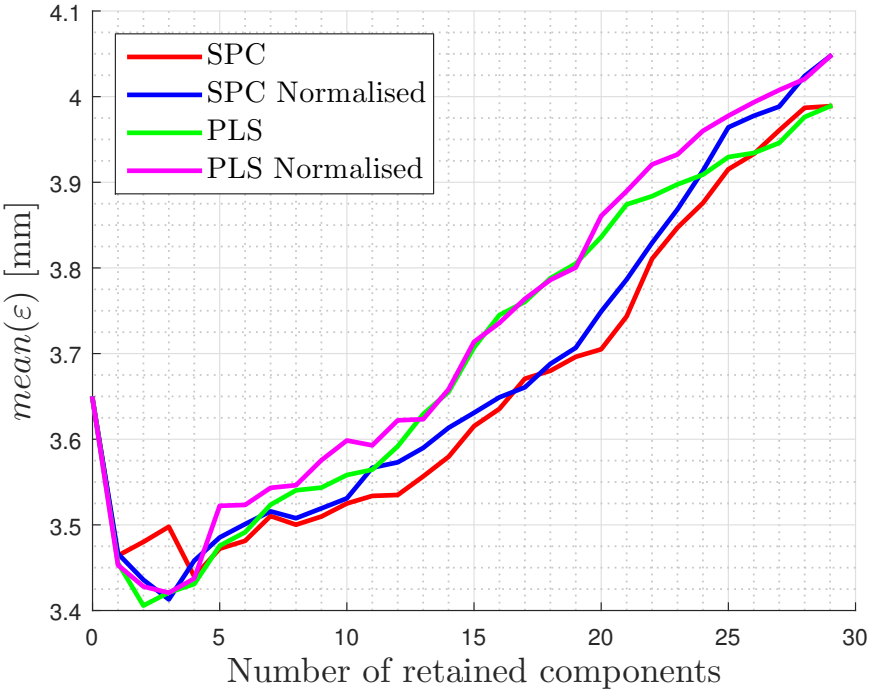
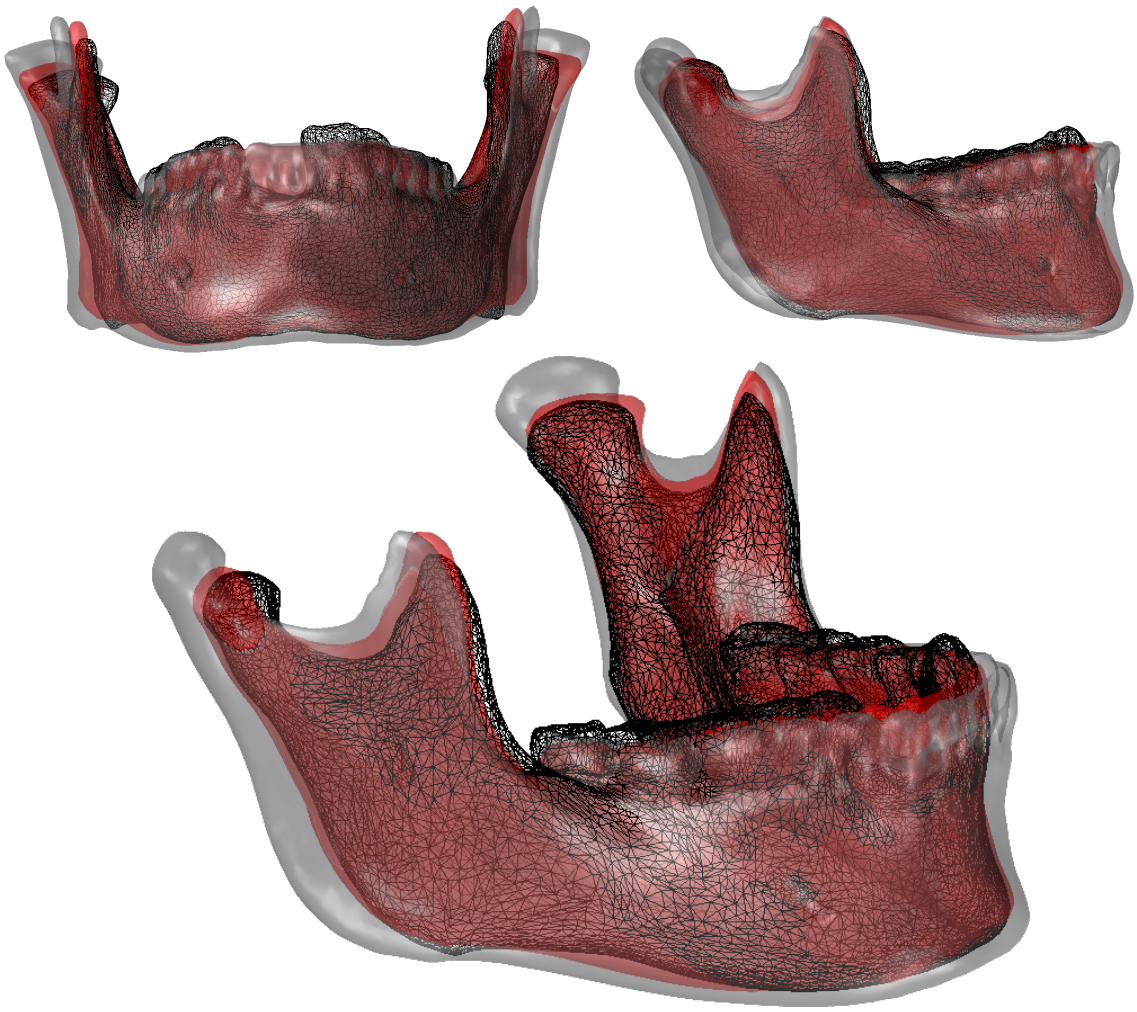
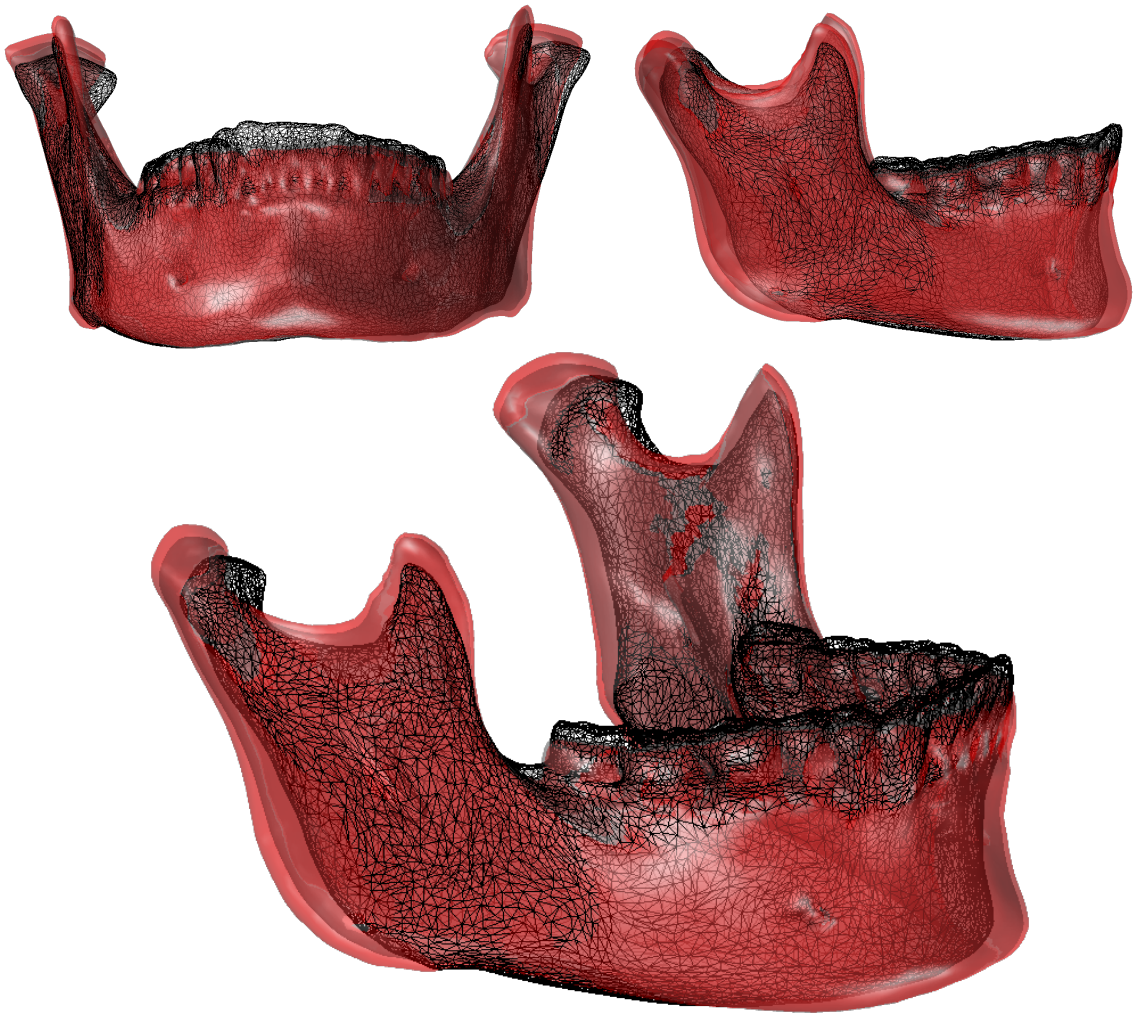


Figure 6.3: Mean point wise error vs. retained components for Euclidean regression.



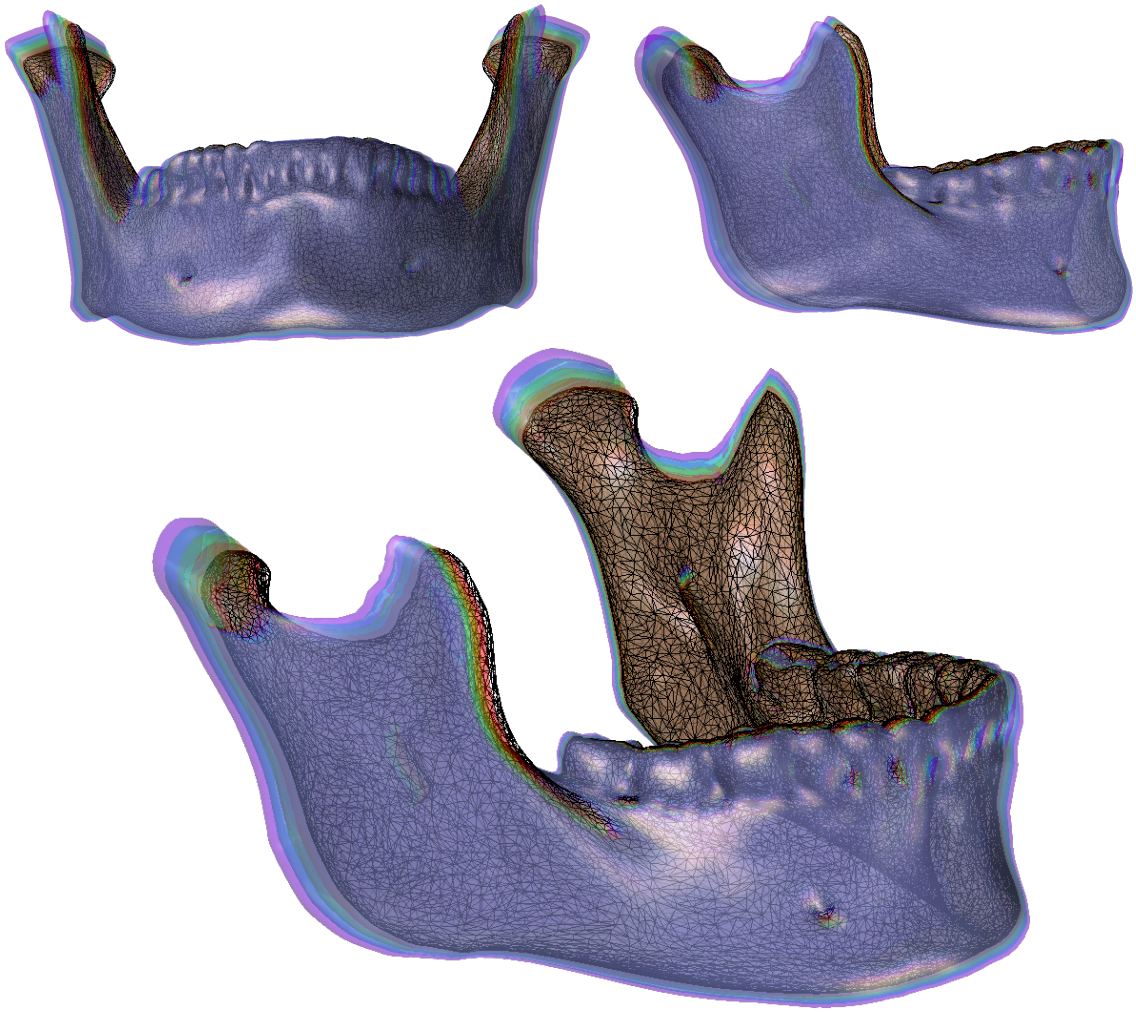
Geometry	mean(ε) [mm]
Geometric mean	5.023
SPC prediction	3.005
Target	

Figure 6.4: Best 5 retained components Cartesian SPC prediction of the 11-fold CV w.r.t. the geometric mean.



Geometry	mean(ε) [mm]
Geometric mean	3.363
SPC prediction	4.055
Target	

Figure 6.5: Worst 5 retained components Cartesian SPC prediction of the 11-fold CV w.r.t. the geometric mean.



	mean(ε) [mm]
—	0
	1
	2
	3
	4

Figure 6.6: Mean point-wise error visualisation by scaling the first load vector, $\mathbf{y}_{L1}^{T}$, of SPC regression of the full dataset.

6.5 ERROR VS. OBSERVATION SCALING

This section serves to demonstrate how the model error behaves as the number of observations in the dataset is increased. The aim is to ascertain to what extent model performance is increased for further expansion of the dataset. This is important due to the time and financial considerations associated with dataset expansion. This study is conducted by performing 1,000 repeated randomly sampled leave-m-out cross-validations (LMO CV) of the full dataset, with m chosen to result in the desired number of training observations n. This investigation is only performed using Cartesian space SPC regression with normalised unit variance, due to its time tractability when compared to PLS, whilst resulting in comparable model performance, as shown in Figure 6.2.

The LMO CV results are given in Figure 6.7 and shows that the geometric mean and predictive model CV minimums result in decreasing mean point-wise error for increasing observations. In addition, the optimal number of components increased from 3 to 5 as the number of observations increased from 10 to 30 in the data set. The mean point-wise error vs. number of observation are given for the geometric mean and predictive model CV minimums in Figure 6.8. The mean point-wise error of the geometric mean is expected to be a convergent quantity as the sample mean approaches the population mean. The improvement of the predictive CV minimum over the geometric mean is given in Table 6.1, which suggests a sub-linear relationship.

Table 6.1: Improvements of the predictive CV minimums over the geometric mean for increasing observations.

Observation, n	Improvement [mm]
10	0.121
15	0.167
20	0.206
25	0.243
30	0.274

This result suggests that expanding the dataset from the current size will lower mean point-wise error for both the mean model and predictive model minimums. Predictive model improvements over mean models are also expected to increase sub-linearly with increased observations i.e. doubling model performance w.r.t. the mean model requires doubling the number of observations at the very least, realistically requiring more. This indicates that there is merit in increasing the number of observations in the data set.

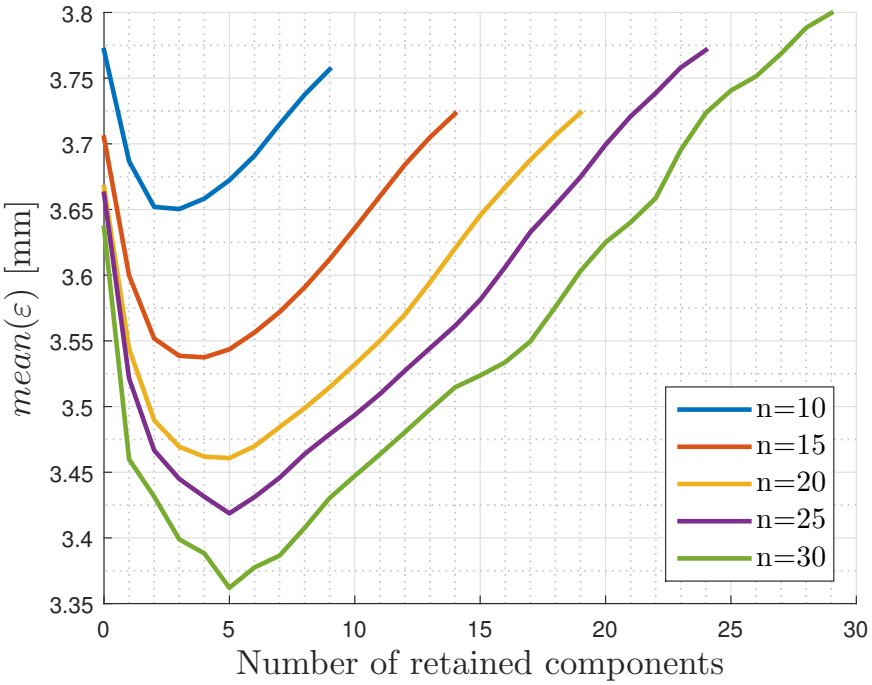


Figure 6.7: Mean point-wise error vs. number of retained components for five data sets with increasing number of observations, using LMO CV.

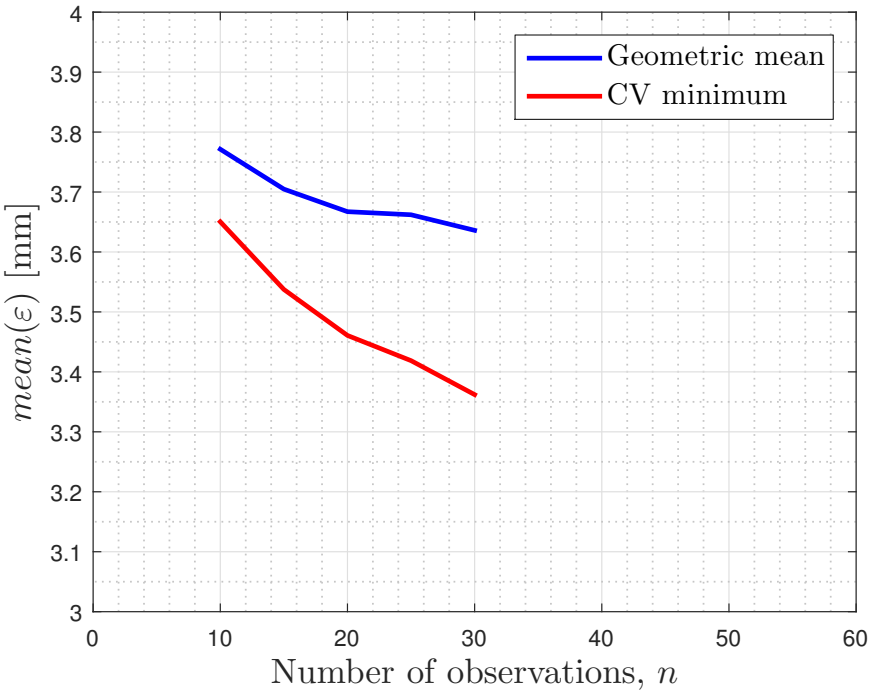
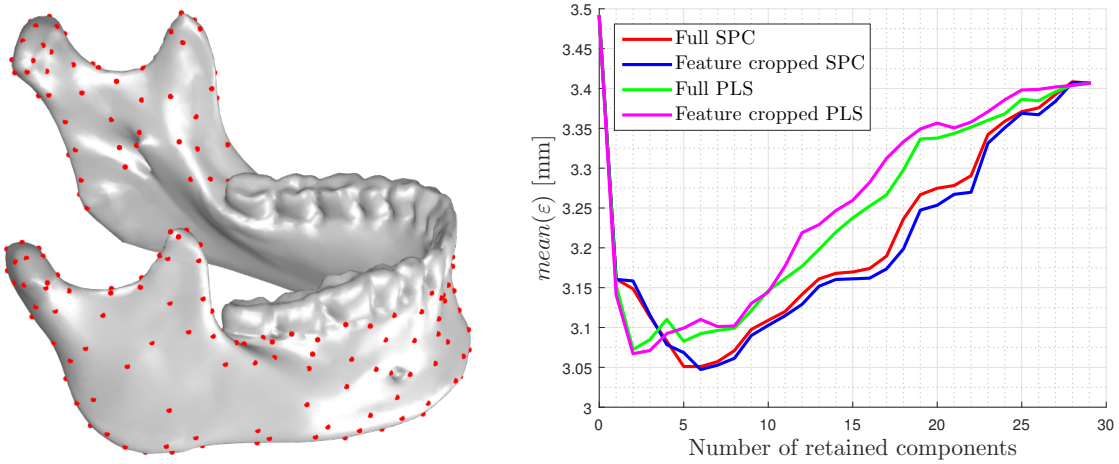


Figure 6.8: Mean point-wise error vs. the number of observations for geometric mean and predictive CV minimum models.

6.6 FEATURE CROPPING

Exclusion of less important features may possibly result in increased explained variance in primary loadings of the remaining features, resulting in improved prediction performance. For application of cranio-facial reconstruction, only the outer visible surfaces of the mandible are important i.e. everything visible from the frontal/anterior and side/lateral views. For this demonstration, feature cropping is utilised to result in an output space consisting of the features relating to these visible areas, illustrated in Figure 6.9. Due to some observations in the dataset having multiple missing teeth that could not be "interpolated" during registration, features relating to teeth are also cropped. The results of an 11-fold CV using normalised Cartesion SPC and PLS regression of the full and feature cropped output spaces are given in Figure 6.9. The evaluated error is only given at the preserved features/points and shows that negligible improvement is obtained over the CV minimum. This is because the cropped covariance does not result in significantly differing output space primary loadings. The feature cropped error is observed to be lower, when considering the post optimal domain of between 10 and 25 retained components. This is where the loadings are assumed to differ more significantly. Although no significant improvements are obtained in this instance, feature cropping may result in improved predictions on datasets where important features constitute weaker covariance, compared to less important features.



(a) **(b)**

Figure 6.9: (a) Preserved features and (b) mean point-wise error vs. retained components of full and cropped output spaces.

6.7 CONCLUSION

This chapter demonstrated the use of the developed registration-regression based strategy on a real cranium and mandible dataset containing 33 observations. The results are summarised as follows:

1. 11-fold cross-validation shows marginal and comparable prediction improvements of the proposed regression strategies over using mean models (null component retention of regression models based on derived latent vectors).

2. The marginal prediction improvements for application on this dataset is not entirely unexpected, since facial features, which are inherently tied to underlying bone structure, are often employed as biometric indicators [28].

3. The various proposed regression models require greatly varying computational time for comparable prediction performance. This suggests the most computationally efficient prediction model, Cartesian SPC regression, to be best suited for further development. This conclusion may be re-evaluated for expanded datasets.

4. Observation scaling suggests that further dataset expansion, from the current $n = 33$, will result in increased model performance. Prediction performance over mean models improve sub-linearly i.e. doubling model performance w.r.t. the mean model requires doubling the number of observations at the very least, realistically requiring many more.

5. Feature cropping does not result in significantly reduced prediction error in this application.

CHAPTER 7

REGRESSION-REGISTRATION HYBRID

Registration of geometries could up to this point only be performed on full/complete geometries. This limitation of the proposed registration strategy is discussed in this chapter. A new regression based registration and prediction strategy, for when a fully registered dataset of the geometry is available, is also proposed and demonstrated. This strategy is shown to also work for partial geometries.

7.1 ON THE PROPOSED REGISTRATION ALGORITHM

The registration algorithm proposed in Chapter 3 can only register full/complete geometries. If a geometry is too dissimilar or has significant missing representation due to damage (partial geometries), the resulting registration may not be representative. Referring back to Section 2.4.2, recall that a least square RBF estimation of the spatially overdetermined forward and backward displacement correspondence is utilised. This mechanism dictates nearest neighbour correspondence in regions that strictly don't have correspondence, resulting in deformation of regions where interpolation is more suitable. Modification of the registration algorithm may be required to result in interpolation of these regions. Possible modifications, that may constitute work for further study due to limited scope of this study, may include filters to remove correspondence from the least square RBF displacement field estimation based on:

1. Normal based thresholding, to remove correspondence with obtuse normals.

2. Similarity based thresholding, to remove correspondence of regions that are not geometrically similar.

7.2 REGRESSION BASED REGISTRATION AND PREDICTION

The proposed modifications to the registration algorithm in the previous section may in general only act to improve registration of damaged geometries where interpolated regions equate to small spatial extrapolations. As the spatial extrapolation becomes greater, the resulting deformation in these regions become an artefact of the chosen basis function and may in general not be representative of actual biological geometric variation. This section proposes a technique where the resulting correspondence is built up through observed deformation. This dictates that geometric differences in missing regions, that require "spatial extrapolations" of the registration algorithm in Chapter 3, are described by observed geometric differences rather than an arbitrarily chosen basis function.

The proposed registration strategy is based on regression of a registered dataset of the given geometry i.e. regression based registration. This strategy results in a single step registration/prediction for a geometry that does not have correspondence, i.e. if a dataset for the given geometry is available, registration is explicitly facilitated through fitting the deformational modes (loadings) of the dataset. A similar approach is used by Buchaillard *et. al.* for the reconstruction of teeth [29]. The strategy is set out as follows:

1. A registered dataset of the given geometry is required.

2. A PCA of the dataset is performed to obtain the associated deformational loadings, as discussed in Chapter 4.

3. A numerical optimisation routine is utilised to obtain the scores which best fit the reconstructed deformational loadings to the target geometry. This is implemented as follows:

   (a) Gradient descent optimization is utilised.

   (b) The target geometry is registered to the perturbed geometry through R-ICP, outlined in Section 2.3.3.

   (c) The sum of the squared point-wise distances (Procrustes distance) of the nearest neighbour correspondence is used as the minimising function.

The aim of this section is solely to demonstrate the concept of fitting deformational loadings when correspondence is not available. Problem properties, modifications and alterations to the proposed technique may constitute work for further study.

The proposed technique is demonstrated on two examples that both make use of the mandible dataset used in Chapter 6. The first example removes an observation in the dataset and simulates a partial mandible, where a significant portion has been removed. A leave-one-out cross-validation (LOO CV) is performed to select the number of retained components (deformational modes). The superset, $\boldsymbol{Z}$, mean point-wise error is used for this i.e. a combination of both fitting the input space, $\boldsymbol{X}$, and predicting the output space, $\boldsymbol{Y}$. Note that the predictor and response sets constitute approximately 70% and 30% of the superset respectively i.e. $n_X \approx 0.7n_Y$ and $n_Y \approx 0.3n_Z$. This is done using the registered dataset and results in a full component retention minimum, as shown in Figure 7.1. The regression-registration result for full component retention for the simulated partial mandible is given in Figure 7.2. The second example is the reconstruction of a partial mandible fossil, Homo Naledi [30] obtained from MorphoSource [31] with the resulting reconstruction given in Figure 7.3. This is only included as an interesting application to demonstrate the concept. The resulting reconstruction may not be representative due to the potentially unsuitable assumption of the same geometric mean and deformational modes between different species. A more representative result may be obtained by using a registered dataset of different species within the Homo genus.

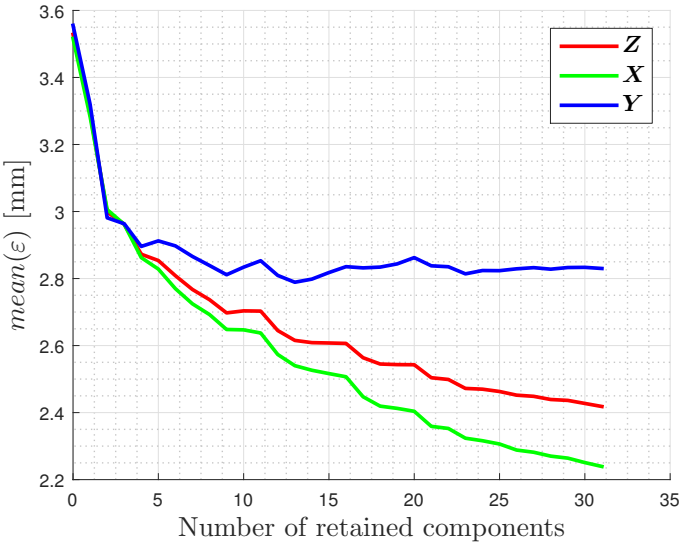
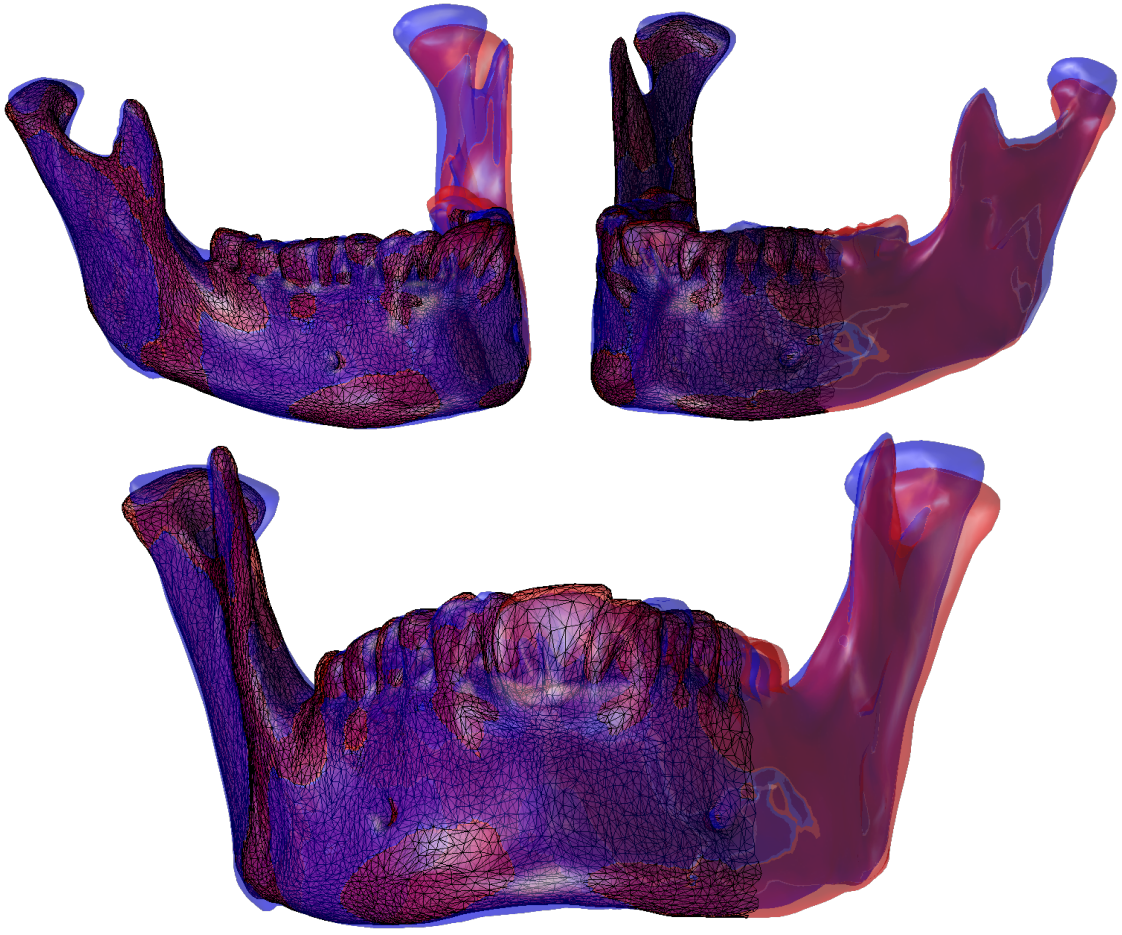


Figure 7.1: LOO CV results using the superset mean point-wise error of the simulated partial mandibles.



Geometry	
—	Partial
	Target
	Registration/Prediction

Figure 7.2: Reconstruction of a simulated partial mandible using regression based registration, with full component retention.

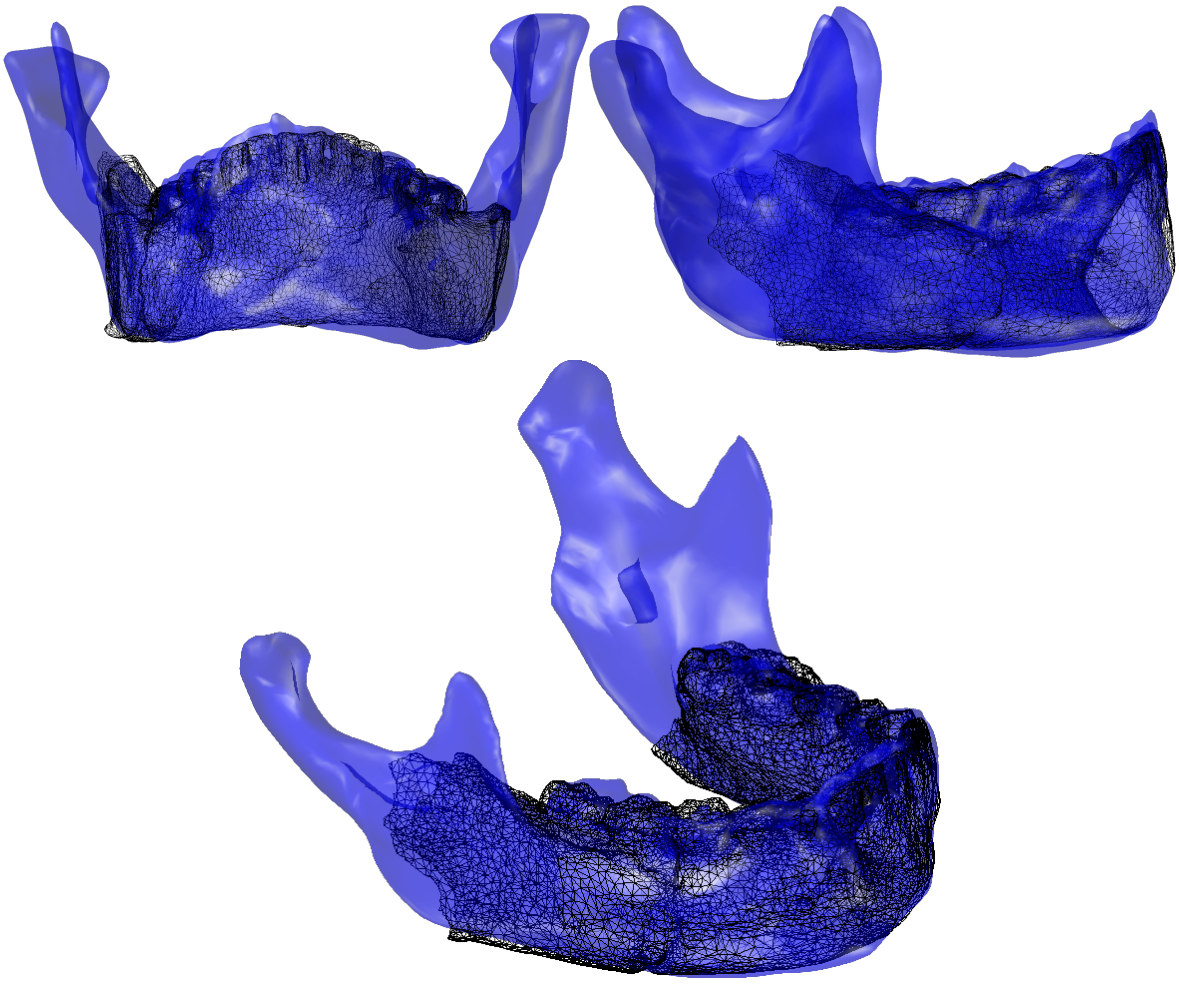


Figure 7.3: Mandible reconstruction of a Homo Naledi partial mandible fossil, using regression based registration, with full component retention.

CHAPTER 8

CONCLUSION

8.1 REMARKS AND FUTURE WORK

An NR-ICP registration method based on a least square RBF approximation of the displacement field is developed for use with surface meshes. The objective of the technique is to construct a mapping from one manifold/surface to a similar other. This is accomplished by deforming the point-set of one to the other by minimising the sum of the squared point-wise nearest neighbour displacements. The mechanism showed to be capable of reasonable mappings for most of the considered geometries, but is highly geometry/manifold and similarity dependant, as convergence of the minimising function is not guaranteed to result in a representative mapping. This may however be regularised/corrected by using normals and/or a similarity/probabilistic approach to only allow reasonable correspondences in the least square approximation, as described in Section 7.1.

A recent non-peer-reviewed study can be used to sound board the methods and approach followed throughout this study. This internal technical report provides a comparison of suitable non-rigid registration frameworks: *Development and Comparison of Non-Rigid Surface Registration Algorithms and Extensions* by Snyders *et. al.* [15]. They provide several suitable refined frameworks as compared to the generic non-rigid registration developed in this study. They for example utilise an outlier detection method, where detected outlier correspondences are interpolated [15]. This method can be implemented as is in the existing registration where outliers are removed from the least squares displacement correspondence approximation. Further an iterative correspondence update is possible by registering the dataset geometric mean shape towards each observation after each full registration of the dataset [15, 32]. A single registration step was utilised in this study. The use of a similar update may be invest-

igated and is expected to result in improved correspondence as the total registration deformation is expected to be lowered i.e. the possibility of registering a lower outlier to an upper outlier is avoided - these large deformations are expected to result in more distorted correspondence. Further comparison of this study with that of Snyders *et. al.* [15] shows that the registration techniques developed in this study are generic versions of existing intuitive concepts:

1. Displacement transformation regularisation, in this study, is accomplished through regression of a chosen RBF displacement.. Snyders *et. al.* also considered this transformation regularisation (quasi-interpolation).

2. The issue of large deformation registration which requires protrusion of surfaces, is dealt with in this study by using a single-step least square approximation of forward and backward displacement correspondences i.e. source to target and target to source. Snyders *et. al.* use a multi-step symmetric registration procedure.

3. The study in Chapter 3, that characterised the non-rigid registration algorithm, showed that the slowest possible registration complexity increase resolves the optimal correspondence. This is a heuristic manual approach that aims to accomplish what the formal framework of deterministic annealing [33] accomplishes i.e. the convergence of registration is regularised by keeping the registration as rigid as possible (minimum fitting complexity increase at each iteration).

4. Snyders *et. al.* also utilise Principle Component Analysis to develop a statistical based registration technique variant as done in Chapter 7.

The artificial dataset application serves as proof of concept and demonstrated that the proposed registration-regression approach is capable of reasonable prediction when the dataset contains underlying linear structure. The real world application demonstrated that an improvement over using the geometric mean is possible, however marginal. The use of kernalised regression methods were not investigated due to the relative limited dataset size, but these methods may be capable of regressing non-linear relationships when the dataset is expanded vastly.

Cartesian coordinate based alignment is facilitated through standard Generalised Procrustes Analysis (GPA) in this study. Other Cartesian coordinate based alignment techniques, such as resistant-fit analysis [34], may be investigated. Resistant-fit differs from GPA by reducing the distribution of high variance regions over the entire

geometry. This is an inherent trait of the least squares approach of GPA.

The characteristics of the proposed regression based registration in Chapter 7 may need further study. Although it was only demonstrated as a registration/reconstruction/prediction strategy suited for partials, this method may also be suited to registration of full/complete geometries or at least as an initial correspondence guess for NR-ICP registration. The registration noise of regression based registration may be potentially lower for use of the dominant/primary retained modes and associated fixed linear deformation, when compared to the non-linear deformations possible through NR-ICP registration, i.e. regression based registration deformation is better regularised than that of NR-ICP registration. This may result in improved prediction performance for dataset expansion using this method.

BIBLIOGRAPHY

[1] J. G. Clement and M. K. Marks, *Computer-graphic facial reconstruction*. Academic Press, 2005.

[2] M. Bastir, A. Rosas, and H. D. Sheets, "The morphological integration of the hominoid skull: a partial least squares and pc analysis with implications for european middle pleistocene mandibular variation," in *Modern morphometrics in physical anthropology*. Springer, 2005, pp. 265–284.

[3] S. Senck, M. Coquerelle, G. W. Weber, and S. Benazzi, "Virtual reconstruction of very large skull defects featuring partly and completely missing midsagittal planes," *The Anatomical Record*, vol. 296, no. 5, pp. 745–758, 2013.

[4] R. Neeser, R. R. Ackermann, and J. Gain, "Comparing the accuracy and precision of three techniques used for estimating missing landmarks when reconstructing fossil hominin crania," *American journal of physical anthropology*, vol. 140, no. 1, pp. 1–18, 2009.

[5] T. Ju. (2014) Registration. [Online]. Available: http://www.cse.wustl.edu/~taoju/cse554/lectures/lect08_Alignment.pdf

[6] K. Arun, T. Huang, and S. Blostein, "Least-squares fitting of two 3-d point sets," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 5, pp. 698–700, 1987.

[7] P. Besl and H. McKay, "A method for registration of 3d shapes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, pp. 236–256, 1992.

[8] D. Shaoyi, Z. Nanning, Y. Shihui, and L. Jianyi, "Affine iterative closest point algorithm for point set registration," *Pattern recognition letters*, vol. 31, pp. 791–799, 2010.

[9] A. de Boer, M. van der Schoot, and H. Bijl, "Mesh deformation based on radial basis function interpolation," *Computers and Structures*, vol. 85, pp. 784–795, 2007.

[10] M. Powell, *The theory of radial basis function approximation.* Oxford: Oxford University Press, 1990, pp. 105–210.

[11] Inria. Gamma model repository. [Online]. Available: https://www.rocq.inria.fr/gamma/gamma/download/download.php

[12] H. Chui and A. Rangarajan, "A new point matching algorithm for non-rigid registration," *Computer Vision and Image Understanding*, vol. 89, no. 2, pp. 114–141, 2003.

[13] B. J. Brown and S. Rusinkiewicz, "Non-rigid range-scan alignment using thin-plate splines," in *3D Data Processing, Visualization and Transmission, 2004. 3DPVT 2004. Proceedings. 2nd International Symposium on.* IEEE, 2004, pp. 759–765.

[14] J. Feldmar and N. Ayache, "Rigid, affine and locally affine registration of free-form surfaces," *International journal of computer vision*, vol. 18, no. 2, pp. 99–119, 1996.

[15] J. Snyders, P. Claes, V. D., and S. P., "Development and comparison of non-rigid surface registration algorithms and extensions," KU Leuven, Technical Report, 2014.

[16] K. Varmuza and P. Filzmozer, *Introduction to Multivariate Statistical Analysis in Chemometrics.* Boca Raton: CRC Press, 2009, p. 105.

[17] R. Rosipal and N. KrÃđmer, *Overview and recent advances in partial least squares.* Springer, 2006, ch. Subspace, latent structure and feature selection, pp. 34–51.

[18] T. Hastie, J. Tibshirani, and J. Friedman, *The elements of statistical learning: data mining, inference, and prediction,.* Springer, 2009.

[19] S. De Jong, "Simpls: an alternative approach to partial least squares regression," *Chemometrics and intelligent laboratory systems*, vol. 18, no. 3, pp. 251–263, 1993.

[20] E. L'Abbe, M. Loots, and J. Meiring, "The pretoria bone collection: A modern south african skeletal sample," *Journal of Comparative Human Biology*, vol. 56,

pp. 197–205, 2005.

[21] M. P. Holmes, J. Isbell, C. Lee, and A. G. Gray, "Quic-svd: Fast svd using cosine trees," in *Advances in Neural Information Processing Systems 21*, D. Koller, D. Schuurmans, Y. Bengio, and L. Bottou, Eds. Curran Associates, Inc., 2009, pp. 673–680.

[22] J. Gower, "Generalised procrustes analysis," *Psychometrika*, vol. 40, pp. 33–51, 1975.

[23] J. Magnus and H. Neudecker, *Matrix Differential Calculus with Applications in Statistics and Econometrics*. New York: John Wiley and Sons, 1999, p. 34.

[24] J. Kruskal, "Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis," *Psychometrika*, vol. 29, pp. 1–27, 1964.

[25] E. van der Westhuizen, Personal communication, 2015, lieutenant, Facial Reconstruction, Victim Identification Center, Forensic Science Laboratory, South African Police Service.

[26] J. Hoffman and F. De Beer, "Characteristics of the micro-focus x-ray tomography facility (mixrad) at necsa in south africa," *World Conference on Nondestructive Testing*, 2012.

[27] P. Cignoni, M. Callieri, M. Corsini, F. Dellepiane, F. Ganovelli, and G. Ranzuglia, "Meshlab: an open-source mesh processing tool," *Eurographics Italian Chapter Conference*, 2008.

[28] B. Klare and A. Jain, "On a taxonomy of facial features," *Biometrics: Theory Applications and Systems, IEEE International Conference*, 2010.

[29] S. I. Buchaillard, S. H. Ong, Y. Payan, and K. Foong, "3d statistical models for tooth surface reconstruction," *Computers in Biology and Medicine*, vol. 37, no. 10, pp. 1461–1471, 2007.

[30] L. Berger *et al.*, "Homo naledi, a new species of the genus homo from the dinaledi chamber, south africa," *eLife*, 2015.

[31] L. Berger *et al.* Rising star - surface scan of u.w. 101-1261 mandible (dh1).

[Online]. Available: http://morphosource.org/index.php/Detail/ProjectDetail/Show/specimens_order_by/taxon/project_id/124

[32] A. Dierckz, "Automated implant design for facial reconstructive surgery," Master's thesis, KU Leuven, 2014.

[33] S. Gold, A. Rangarajan, C. Lu, S. Pappu, and E. Mjolsness, "New algorithms for 2d and 3d point matching: Pose estimation and correspondence," *Pattern recognition*, vol. 31, no. 8, pp. 1019–1031, 1998.

[34] F. Rohlf and D. Slice, "Extensions of the procrustus method for the optimal superposition of landmarks," *Systematic Zoology*, 1990.