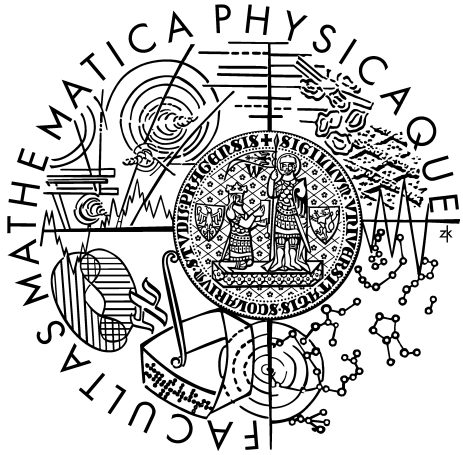
Charles University in Prague
Faculty of Mathematics and Physics

DOCTORAL THESIS



Václav Krajíček

Correspondence Problem in Geometrics Morphometric Tasks

Department of Software and Computer Science Education

Supervisor of the doctoral thesis: RNDr. Josef Pelikán
Study programme: Computer science
Specialization: 4I2 — Software systems

Prague 2015

At this point, I would like to express my sincere gratitude to my supervisor RNDr. Josef Pelikán for the continuous support in study and research, for his patience, motivation and enthusiasm, especially in the last years.

Beside my supervisor, I would like to thank RNDr. Jana Velemínská, PhD., from The Department of Anthropology and Human Genetics by Faculty of Life Sciences, Charles University in Prague, who allowed me to join her research team and taught me a great deal about practical science.

Also, my sincere thanks goes to all the people from Computer Graphics Group at The Department of Software and Computer Science Education by Faculty of Mathematics and Physics, Charles University in Prague, for many years of a great and motivating work atmosphere. Notably, I thank Mgr. Ján Dupej, my colleague and friend who joined me in our common effort to extend our knowledge in the field of shape analysis as well as in life in general.

I would also like to extend my thanks to all colleagues from The Department of Anthropology and Human Genetics who welcomed me as one of their own. Special thank goes to Mgr. Šárka Bejdová, Mgr. Hana Brzobohatá, PhD. and Mgr. Jana Koudelová for sharing their grant support with me and for helping me fulfill our common research goals.

This thesis and the related research were partially supported by research grants GAUK 309611 and GAUK 1388213.

Last but not the least, I would like to thank my parents who had never lost trust in me and supported me in all possible ways throughout my life, no matter what path I had chosen.

I declare that I carried out this doctoral thesis independently, and only with the cited sources, literature and other professional sources.

I understand that my work relates to the rights and obligations under the Act No. 121/2000 Coll., the Copyright Act, as amended, in particular the fact that the Charles University in Prague has the right to conclude a license agreement on the use of this work as a school work pursuant to Section 60 paragraph 1 of the Copyright Act.

In ..... date ..... signature of the author

Název práce: *Problém korespondence v úlohách geometrické morfometrie*
Autor: *Václav Krajíček*
Katedra / Ústav: *Katedra software a výuky informatiky*
Vedoucí doktorské práce: *RNDr. Josef Pelikán*
e-mail vedoucího: *pepca@cgg.mff.cuni.cz*
Abstrakt:

Analýza tvaru ve fyzické antropologii, biomedicíně a přidružených oborech je často prováděna s použitím landmarků nebo pomocí měření vzdáleností. Nové technické možnosti dovolují digitalizovat věrný vzhled objektu ve formě trojúhelníkových sítí nebo objemových dat. Tyto digitální obrazy jsou obzvláště užitečné v případech, kdy nemohou být vhodným způsobem použity landmarky k popisu tvaru.

Aby bylo možné statisticky analyzovat tvar na vzorku pozorování, které jsou reprezentovány zmíněnými zobrazovacími technikami, musí být identifikovány vzájemně si odpovídající body.

Registrace je klíčovým nástrojem k mapování reprezentací tvaru do společné souřadné soustavy, kde se hledají vzájemně si odpovídající body, v případě trojúhelníkových sítí na principu nejbližšího souseda a v případě objemových dat podle překrývajících se bodů. Elastická registrace založená na B-spline interpolaci byla vybrána kvůli své mnohostrannosti, relativní rychlosti a schopnosti registrovat trojúhelníkové sítě i objemová data. Zároveň byly prováděny experimenty i s alternativními registračními metodami — založenými na Thin-plate spline funkcích a Coherent point drift algoritmu. B-spline registrace byla modifikována, aby zvládala datové množiny různých morfometrických studií a zrychlena s využitím chytrého vzorkování během optimalizace registrační kriteriální funkce, což umožnilo její urychlení až o 2–3 řády.

Navržené algoritmy byly demonstrovány dvěma způsoby: (1) jako nástroj pro obecné morfometrické úlohy, jako je zkoumání variability tvaru, asymetrie nebo dopočítání chybějících dat; (2) v mnoha úlohách reálných morfometrických výzkumných projektů, kde byly studovány různé fenomény od struktury středověkého obyvatelstva po hodnocení lékařských procedur v dentální chirurgii.

Bylo prokázáno, že přímá analýza digitálních dat, bez redukce informace způsobené výběrem landmarků, odhalí mnohem víc o studovaném fenoménu, než když je použita pouze landmarková metodika. Například, úspěšnost při určení pohlaví podle obličeje se zlepšila o 22,7% za použití trojúhelníkových sítí v porovnání s landmarky. Vždy ale záleží na konkrétním projektu, zda nelandmarková metoda podá lepší výsledky, než řešení za použití landmarků.

Klíčová slova: *registrace, korespondence, analýza tvaru, geometrická morfometrie, lékařské zobrazovací metody, trojúhelníkové sítě, objemová data*

Title: *Correspondence Problem in Geometric Morphometrics Tasks*
Author: *Václav Krajíček*
Department / Institute: *Department of Software and Computer Science Education*
Supervisor of the doctoral thesis: *RNDr. Josef Pelikán*
Supervisor's e-mail address: *pepca@cgg.mff.cuni.cz*
Abstract:

Shape analysis in physical anthropology, biomedicine, and related disciplines is mostly done using landmarks or by measuring distances. New technological advancements allow the digitization of object's appearance in the form of triangular meshes or volume images. These digital images are especially beneficial in the cases when landmarks cannot be used to effectively describe the shape.

In order to statistically analyze shape in a sample of observations, which are represented by these modalities, correspondence has to be found.

Registration is a crucial tool in mapping the shape representations into a common space where correspondence is found by nearest neighbor principle in the case of triangular meshes or by overlaps in the case of volume images. B-spline based non-rigid registration is chosen because of its versatility, relative speed and ability to handle both meshes and volume images. Experiments were also performed with other alternatives — Thin-plate splines and Coherent point drift. The algorithm was modified to handle the data in various morphometric studies. It was also improved in speed by employing smart sampling for the optimization of the registration objective function, allowing a speed up of 2–3 orders of magnitude.

The proposed algorithms were demonstrated in two ways: (1) as a tool for generic morphometric tasks such as shape variability, asymmetry analysis, missing data imputation; (2) in many tasks of actual morphometric research investigating different phenomena ranging from the structure of the medieval population to an evaluation of treatment procedures in dental surgery.

It was confirmed that a direct analysis of digital images, without information reduction by landmark placement, is able to uncover more of the studied phenomena a method based solely on landmarks. For example, discrimination success rate of face with respect to sex has improved by 22.7% using meshes in comparison to landmarks. Of course, it always depends on the particular project whether a non-landmark method outperforms a landmark-based solution.

Keywords: *registration, correspondence, shape analysis, geometric morphometrics, medical imaging, triangular meshes, volume images*

Contents

Chapter 1

Introduction

Our visions begin with our desires.

Audre Lorde

Current research into life sciences uses highly elaborate modern methods, such as genetic analysis in molecular biology, but there has been an evolution in the traditional methods used for centuries, operating on a scale perceivable to the naked eye. Documenting the physical appearance of real world objects is one of the most important methodological approaches and lies at the foundation of many fields in life sciences. This approach has been extended by elements of quantitative research, i.e. integrating documentation of many appearances into one model to explain studied phenomena. With the advent of mathematical statistics, these approaches have become more and more sophisticated.

The presented thesis is concerned with an extension of the methodological framework for the analysis of shapes, called *geometric morphometry* (GMM) and its application in the fields of physical anthropology and biomedicine. The pivotal subject of this thesis is the problem of geometric correspondence, since the tasks laid out below involve various forms that compare different geometric representations of real world objects and knowing the correspondence is the first step for comparison. This problem becomes more challenging when it comes to new modern techniques of capturing shapes that produce new data modalities.

1.1 Structure of the thesis

The thesis is divided into five chapters. Throughout the chapters, several case studies are presented which either demonstrate the application of methods explained in currently published research that the author has participated on, or describe complex examples of still-to-be published original algorithms.

After the introductory chapter, which presents a layout of the work, the second chapter introduces the topic of the thesis along with a background to

the research. Several examples will show the method of statistical analysis for geometrical shapes used in scenarios where correspondences are explicitly known. This is important in determining the type of tasks which can be solved and motivates research into correspondence in the following chapters. At the end of the chapter, a study is presented which constructs correspondences in two-dimensional images, partially by explicitly defined points and partially by an algorithm.

The third chapter approaches the topic of correspondences on surfaces represented by triangular meshes. Mesh and point-cloud registration algorithms are described. Information is provided to show that B-spline-based registration, among other methods, is particularly suitable for precise and fast mesh-fitting, along with a description of how other options are used. Various tasks are discussed, including basic variability analysis, dimorphism analysis, asymmetry analysis, and even partial geometry fitting and missing data.

In the fourth chapter, voxel-based morphometrics is introduced, which uses volume data for shape representation. This is a natural extension to surface representation of meshes and allows for the most complicated shapes to be captured, which might otherwise be impossible to analyze while using surfaces.

The fifth chapter concludes the work and offers potential directions for future research.

1.2 Publications

In this section, a list of the author's original published work is presented, which is related to the topic of the thesis. The publications are divided into three groups, reflecting the order of their appearance in the following chapters.

Landmark data

The author has contributed to a number of research studies based on *landmark methodology*, which is considered the current mainstream methodology for shape description and analysis. However, there are still unresolved problems, e.g. missing data computation. These works are also important since they inspire new approaches to shape analysis.

- Brzobohatá, H., Krajíček, V., Velemínský, P., Poláček, L., Velemínská, J., The Shape Variability of Human Tibial Epiphyses in an Early Medieval Great Moravian Population (9th -10th Century AD): A Geometric Morphometric Assessment. *Anthropologischer Anzeiger*, 2014, **71**(3), pp. 219–236

- Brzobohatá, H., Krajíček, V., Horák, Z., Velemínská, J., Sex Classification Using the Three-Dimensional Tibia Form or Shape Including Population Specificity Approach. *Journal of Forensic Science*, 2015b, **60**(1), pp. 29–40

- Brzobohatá, H., Krajíček, V., Horák, Z., Sedlak, P., Velemínská, J., Diachronic changes in size and shape of human proximal tibia in the area of Central Europe during the latest 1200 years. *HOMO - Journal of Comparative Human Biology*, 2015a. (submitted)

- Chvojková, M., Krajíček, V., Velemínská, J., Kraniometrická Variabilita Historických Populací z Oblasti Údolí Nilu. *Slov. Antropol.*, 2010, **13**(2), pp. 19–23–5

- Bigoni, L., Krajíček, V., Sládek, V., Velemínský, P., Velemínská, J., Skull shape asymmetry and the socioeconomic structure of an early medieval Central European society. *American Journal of Physical Anthropology*, 2013b, **150**(3), pp. 349–364

- Bigoni, L., Krajíček, V., Sládek, V., Velemínský, P., Poláček, L., Velemínská, J., Different Subsistence Patterns and the Socioeconomic Structure of Medieval Society of Great Moravia. 1838emes Journées de la Société d'Anthropologie de Paris, Paris, 2013a

- Bejdová, S., Krajíček, V., Velemínská, J., Horák, M., Velemínský, P., Microevolution of mandible in the area of central Europe during the latest 1200 years using methods of 3D geometric morphometrics. Anthropologischer Anzeiger, 68, 4, 2011

- Bejdová, S., Krajíček, V., Velemínská, J., Horák, M., Velemínský, P., Changes in the sexual dimorphism of the human mandible during the last 1200 years in Central Europe. *HOMO - Journal of Comparative Human Biology*, 2013, **64**(6), pp. 437–53

- Bejdová, S., Krajíček, V., Velemínská, J., Horák, M., Velemínský, P., A Microevolution of upper face in the area of Central Europe during the latest 1200 years. 18th Congress of the European Anthropological Association, Ankara, Turkey, 2012b

- Velemínská, J., Krajíček, V., Dupej, J., Goméz-Valdés, J. A., Velemínský, P., Šefčáková, A., Pelikán, J., Sánchez-Mejorada, G., Brůžek, J., Geometric morphometrics and sexual dimorphism of the greater sciatic notch in adults from two skeletal collections: The accuracy and reliability of sex classification. *American Journal of Physical Anthropology*, 2013, **152**(4), pp. 558–565

Mesh data

A new trend in digitizing the shapes of objects is to capture them using a surface scanner. *Triangular meshes* might not only be used to preserve shape information and to extract landmark data, but they can also be used directly in

analysis since they contain complete geometrical information on given object. However, currently there is a lack of methods and software tools for tackling various tasks of shape analysis. This current state of play is the major motivation for the author's contribution to this thesis.

- Dupej, J., Krajíček, V., Pelikán, J., Low-Rank Matrix Approximations for Coherent Point Drift. *Pattern Recognition Letters*, 2014a, **52**, pp. 53–58

- Dupej, J., Krajíček, V., Velemínská, J., Pelikán, J., Statistical Mesh Shape Analysis with Nonlandmark Nonrigid Registration. 12th Symposium on Geometry Processing, Cardiff, UK, 2014b

- Velemínská, J., Bigoni, L., Krajíček, V., Borský, J., Šmahelová, D., Cagáňová, V., Peterka, M., Surface facial modelling and allometry in relation to sexual dimorphism. *HOMO - Journal of Comparative Human Biology*, 2012, **63**(2), pp. 81–93

- Bejdová, S., Krajíček, V., Trefný, P., Peterka, M., Velemínská, J., Variability in palatal shape and size in patients with bilateral complete cleft lip and palate assessed using dense surface model construction and 3D geometric morphometrics. *Journal of Cranio-Maxillofacial Surgery*, 2012a, **40**(3), pp. 201–208

- Rusková, H., Bejdová, S., Peterka, M., Krajíček, V., Velemínská, J., 3-D shape analysis of palatal surface in patients with unilateral complete cleft lip and palate. *Journal of Cranio-Maxillofacial Surgery*, 2014, **42**(5), pp. 140–147

- Krajíček, V., Dupej, J., Velemínská, J., Pelikán, J., Morphometric Analysis of Mesh Asymmetry. *Journal of WSCG*, 2012, **20**(1), pp. 65–72

- Dupej, J., Krajíček, V., Velemínská, J., Pelikán, J., Analysis of Asymmetry in Triangular Meshes. In *Proceedings of the 33 rd Conference on Geometry and Graphics*, VŠB-Technical University of Ostrava, 2013, pp. 65–78

- Krajíček, V., Dupej, J., Koudelová, J., Velemínská, J., Statistical Mesh Analysis of Longitudinal Shape Changes. In *Proceedings of the 33 rd Conference on Geometry and Graphics*, VŠB-Technical University of Ostrava, 2013, pp. 155–168

- Špačková, J., Cagáňová, V., Krajíček, V., Velemínská, J., Specification of child and juvenile identification: 3D modelling of facial ontogenetic development during the pubertal spurt. European Academy of Forensic Science Conference, Hague, Netherland, 2012

- Cagáňová, V., Borský, J., Krajíček, V., Hoffmannová, E., Černý, M., Velemínská, J., Three-dimensional facial morphology following neonatal cheiloplasty in six-old-years patients with unilateral cleft of the lip and palate. *Journal of Cranio-Maxillofacial Surgery*, 2014. (submitted)

- Koudelová, J., Brůžek, J., Cagáňová, V., Krajíček, V., Velemínská, J., Development of facial sexual dimorphism in children aged between 12 and 15 years: a three-dimensional longitudinal study. *Orthodontics & Craniofacial Research*, 2015. (in press)

- Trefný, P., Krajíček, V., Velemínská, J., Three-dimensional analysis of palatal shape in patients treated with SARME using a dense surface model. *Orthodontics & Craniofacial Research*, 2015. (submitted)

Volume data

The last group relates to the processing of *volume data*, especially their segmentation and registration. In the context of this thesis, this group of methods represents an attempt by the author to employ medical image-processing methods in the field of shape analysis. In the case of physical anthropology, published work focusing directly on volume data is scarce, since it is a relatively new modality and not widely available in the field.

- Krajíček, V., Pelikán, J., Horák, M., Measuring and Segmentation in CT Data Using Deformable Models. In Skala, V. (ed.), *WSCG' 2007 Short Communications Proceedings*, vol. 2, Union Agency, 2007, pp. 149–152

- Krajíček, V., *Volume measurement in 3D data*. Master's thesis, Faculty of Matematics and Physics, Charles University in Prague, April 2007

- Kolomazník, J., Horáček, J., Krajíček, V., Pelikán, J., Segmentation on CUDA Using Graph-Cuts and Watershed Transformation. In *WSCG Poster Proceedings*, Union Agency, 2012, pp. 35–38

- Krajíček, V., Design of Segmentation Algorithm for Volume Measuring CAD system. In *Proceedings of MIS 2008*, Matfyzpress, 2008b, pp. 47–57

- Krajíček, V., Analyzing Contrast Enhanced MRI Sequences for Mammography. In *Proceedings of Contributed Papers: Part I - Mathematics and Computer Sciences*, Matfyzpress, 2008a, pp. 195–201

- Krajíček, V., Dupej, J., Bejdová, S., Velemínská, J., Pelikán, J., Teeth and Jaw Segmentation Using Fast Level-set Algorithm and Local Region Anisotropic Priors. *Imaging Science Journal*, 2014. (submitted)

- Krajíček, V., Bejdová, S., Velemínská, J., Pelikán, J., Improving B-spline Deformation Based Fitting for Volume Registration. In *Proceedings of the 31st Conference on Geometry and Graphics*, VŠB-Technical University of Ostrava, 2011, pp. 139–154

1.3 Goals

In many fields of life sciences there is a noticable cross-over with image processing related disciplines as novel imaging techniques emerge and find application in various traditional areas. In particular, physical anthropology can significantly benefit from medical imaging techniques, such as computed tomography or surface scanning, as a source of shape information.

The characteristic requirement of life sciences is the quantitative approach to data analysis, which is, in the case of shape information, connected with the need for constructing corresponding primitives among all specimens in the sample. The goals of this thesis are to address this need, and specifically to:

- Research current approaches in landmark-based shape analysis including typical tasks.
- Develop effective registration procedures for dense correspondence construction of new types of data (triangular meshes, volumes).
- Apply these registration procedures to similar tasks in order to show their superiority, i.e. the advantages of including full shape information and eliminating decisions made by researchers which can affect the outcomes of experiments, as well as improve accuracy and repeatability.

1.4 Other contributions

Apart from the investigation into new methodological possibilities in the application fields, the author has also participated in the creation of a software tool — Morphome3cs (2015), which has gradually made these new methods practically accessible, so that they can be used by researchers and students in conducting their GMM tasks. The tool is under constant development and the ambition of the author is to make it available for the broader community.

For several years, Morphome3cs has been used for research and educational purposes in the Laboratory of 3D Visualization and Analytical Methods in the Department of Anthropology and Human Genetics, Faculty of Life Sciences at Charles University in Prague. A certain amount of the published work listed above has been carried out with the help of Morphome3cs.

There are many GMM software tools, among which the most popular belong to Morphologica, MorphoJ, PAST, TPS Suite and many others listed at SUNY (2015). The motivation for starting to develop Morphome3cs arose

from the fact that most current tools are either outdated, closed-source, single-purpose or offer no possibility of extending to new methods.

Morphome3cs combines the power of modern GUI with visualization methods for 2D images, meshes and volumetric data, embedded Python scripting for easy extending, capabilities of R, as well as a statistical computation environment (R Development Core Team, 2008).

Chapter 2

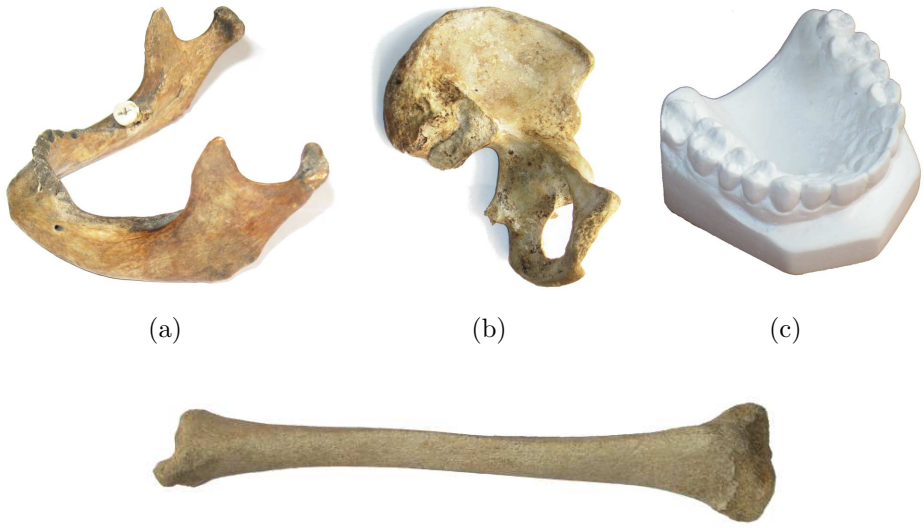
Geometrical methods for shape analysis

> Logic is the beginning of wisdom, not the end.
>
> Leonard Nimoy

The shapes and sizes of living, growing things can reveal much about their essential make-up, history and future, about their relation to other things and about their ability to accommodate to their ever-changing environments.

The ability of researchers to study the shapes and sizes of objects is connected to technical development. Not so long ago, researchers were limited to such rudimentary tools as rulers, calipers, protractors, weights or even restricted to the simpler approaches of filling cavities with mustard seeds or sand to obtain their volumes. Later, photography allowed planar projections of, often, three-dimensional shapes to be produced. However, these projections were connected with a certain information loss and had the potential to bias scientific results based on these data. Therefore, special care was required when using photography as an input for morphometric analysis.

With the advent of digitization techniques, two-dimensional images started to be processed on computers. An ideological shift came about when coordinates of points (landmarks) started to be processed instead of just distances, lengths and length ratios (indices). This resulted in the emergence of a whole new field of landmark data methods, *geometric morphometry* — GMM. Recently, new devices known as contact 3D scanners have appeared which can capture physical coordinates of a point in 3D, which has in turn enabled shape analysis of landmarks to be performed, unimpeded by 3D-to-2D bias. The most recent, relatively complicated devices, comparitively well-known in other fields, have started to be used for shape analysis in life sciences. Devices, such as surface lasers scanners, optical scanners and medical grade CT scanners, have revitalized a field that for hundreds of years had to make do with the caliper.



(a) (b) (c)

(d)

Figure 2.1: Examples of typical objects of interest in modern anthropology and biomedicine shape analysis: (a) jawbone (or mandible); (b) pelvic bone; (c) dental cast; (d) tibia.

Simultaneously with this change, the need for new software tools, algorithms and mathematics has become more urgent. Of course, the tools that now accompany these devices may be used, but they often lack the required functionality. In this chapter, basic methods of analyzing landmark data will be introduced as well as real-world examples of their application in actual research.

2.1 Statistical shape analysis in anthropology

Geometric morphometry (Zelditch *et al.*, 2004) has become an important tool in sub-fields of anthropology (evolutionary, forensic, physical) that are concerned with complex parts of the human body, their characteristic features and differences. See Figure 2.1 for examples.

In general, the methodology may be divided into two stages. The first stage is to directly process geometrical representation, which takes place on the output of an acquisition device or technique. The second, more traditional and common with the other fields of science is statistical processing and datamining (Hastie *et al.*, 2008), which must change to allow statistical results from actual geometry to be interpreted. Let us start with laying backgrounds.

2.2 Landmark-based methods

Landmarks are the basic GMM modality and incorporate information about shape as well as size. One of the central ideas in GMM is the separation of shape and size. Traditional methods have acknowledged that objects with similar shapes can be reflected differently in measured lengths by introducing length ratios. In GMM, the first step of analysis is to normalize measured coordinates by geometrically aligning them to some common frame, while the original size is measured and analyzed separately. The size of object is very important because it contains a lot of information about the specimen, for example, its age, sex, social status or health condition.

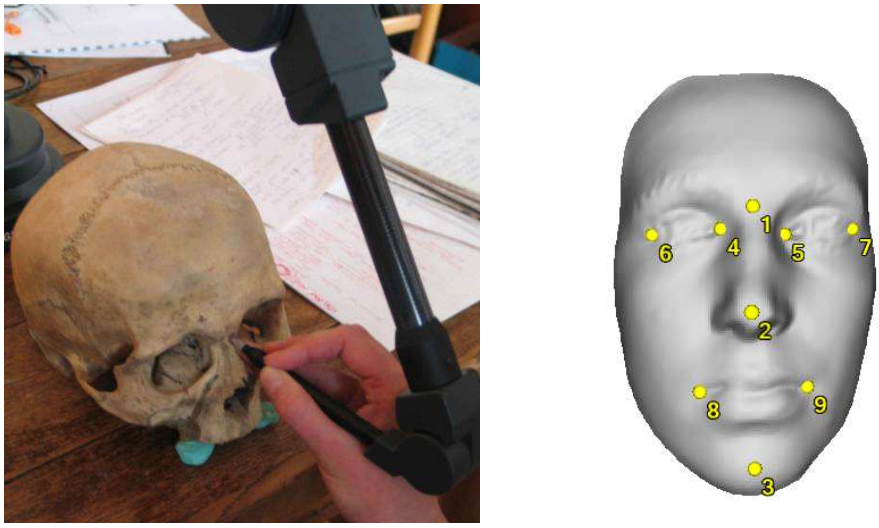
Size in GMM can be defined in various ways, but the most common is by measuring centroid size (CS), which is a linear and non-negative function of landmark configuration $L = \{l_{1,1}, \ldots, l_{1,d}, \ldots, l_{n,d}\}$ for d-dimensional landmarks

$$CS(L) \quad = \quad \sqrt{\sum_{i=1}^{n} \|l_i - \bar{l}\|^2}, \bar{l} = \frac{1}{n} \sum_{i=1}^{n} l_i$$

The shape of the object in relation to a sample of objects of the same kind $\mathbf{L} = \{L_j\}_{j=0}^{m}$ is defined in terms of all geometric information (landmark coordinates, L) remaining after differences in size, orientation and position are removed. Object geometry normalization employs various methods, but the most frequently used is Generalized Procrustes Analysis — GPA (Bookstein, 1997). This method rotates, translates and scales all of the L_j landmark configurations in the sample in order to minimize distances of all landmarks to a corresponding mean landmark configuration $\bar{L}$, which produces configurations L'_j with unit centroid size

$$\arg\min_{\mathbf{L}'} \sum_{j}^{m} \sum_{i}^{n} \|L'_{j,i} - \bar{L}_i\|^2, \bar{L} = \frac{1}{m} \sum_{j}^{m} L'_j$$

Configurations L'_j do not differ from each other in size, orientation or position; hence, they are numerical representations of shapes. The average $\bar{L}$ of L'_j is called the *mean shape*. Furthermore, it is difficult to make conclusions about the shape variations represented by the sample just from L'_j, since they are still coupled point clusters. In order to extract major trends in shape variations, Principal Component Analysis (PCA) is applied (Bishop, 2006). PCA is a statistical method that has many applications in many fields. In shape analysis, represented by a sample (a set of exemplars), it is used to create a Point Distribution Model (PDM), i.e. it finds a basis $\{w_k\}_{k=1}^{n \times d}$ that can be used to represent deviations of each individual in the sample from the mean shape. Moreover, the basis vectors are ordered so that the first represents directions of landmark points in which the variation through the sample is the largest.



(a) (b)

Figure 2.2: Landmark digitization: (a) Localizing landmarks on a physical object using the MicroScribe G2X contact scanner device; (b) Placing landmarks on a facial mesh using an interactive visualization software tool.

The second basis is orthogonal to the first and has the second largest variation and so on. Basis vectors $\{w_k\}_1^{n\times d}$ are called *modes of variation*.

$$L'_j = \bar{L} + \sum_{k=1}^{n\times d} \alpha_{j,k} w_k \tag{2.1}$$

The score vector $\alpha_j = \{\alpha_{j,k}\}_{k=1}^{n\times d}$ defines coordinates of specimen j in vector space $\mathbf{R}^{n\times d}$, which is simply a rotation of the original space containing coordinates of aligned landmarks. The condition on unit size reduces the occurrence of shapes to a curved subspace of $\mathbf{R}^{n\times d}$ called *shape space*. Sets of points in shape space, their mutual relations and statistical properties are evidence for various conclusions.

The number of specimens m is recommended to be larger then $n \times d$, otherwise, linear model (Equation 2.1) is overfitting. Since intrinsic dependency between the landmarks, not all modes of variation are statistically significant, i.e. not all of w_k are required to make reconstructed shape unambiguously identifiable from the others. A number of statistically significant modes of variation can by found by various criteria mentioned by Peres-Neto *et al.* (2005). In this work, broken-stick criterion is implicitly used for this purpose.

In the following paragraphs, examples of these types of research studies will be described.

Case study 1: The human face

The human face is a focal point of interest for researchers as well as artists and ordinary people, since reading the human face can reveal a great deal about the owner. Although the face is a relatively complex object for description due to the many features and characteristics it can be identified by, shape is the only characteristic of surface geometry that becomes relevant, to the exclusion of skin color, facial hair, eyebrows, eye color and hairline. Face shape information is traditionally reduced to a set of landmarks that are placed according to Type 1-3 landmark taxonomy (Zelditch *et al.*, 2004, chapter 2).

In a presented case study, a set of 101 faces of a young Czech population (50 male and 51 female) was captured by a surface scanner and pre-processed, after which landmarks were manually localized by an expert using a software tool under controlled conditions. Figure 2.2(b) shows landmark configurations, while Figures 2.3(a) and 2.3(b) show how the whole sample spread out through the space as well as the alignment and normalization by GPA. Answers to following frequent questions for GMM are sought:

- Mean shape — the mean shape of the sample can be constructed by averaging the aligned landmarks. Mean shapes of subgroups can also be constructed and compared. Thin-plate splines (TPS), introduced to GMM by Bookstein (1997, chapter 2.2), are often used to deform a regular grid to demonstrate spatial deformation required for the transition from one shape to another. See Figure 2.3(c).

- Form analysis — shapes are scaled back to their original size and analyzed when potentially different behavior is shown and provided the size of the specimens is an important feature. See Figure 2.3(d).

- Variability — the most important and mutually independent trends are isolated by PCA. The relation of shapes and groups of shapes is expressed by the amounts of these trends presented in a particular shape instance. See Figure 2.3(e).

- Separability — a property of a particular shape whose instances can be clearly distinguished from each other according to different classes they belong to. The first step is to prove the statistical significance of the differences using statistical measures, i.e. multivariate statistical tests. The second step is to visualize the subgroup, i.e. localize the differences. The third step is to investigate the discrimination power of the shape itself, i.e. compute a success rate while classifying a shape instance into a correct class based on shape features with the help of a suitable classifier.

- Sexual dimorphism — an example of separability according to sex affiliation. In the case of sexual dimorphism of facial landmarks, leave-one-out

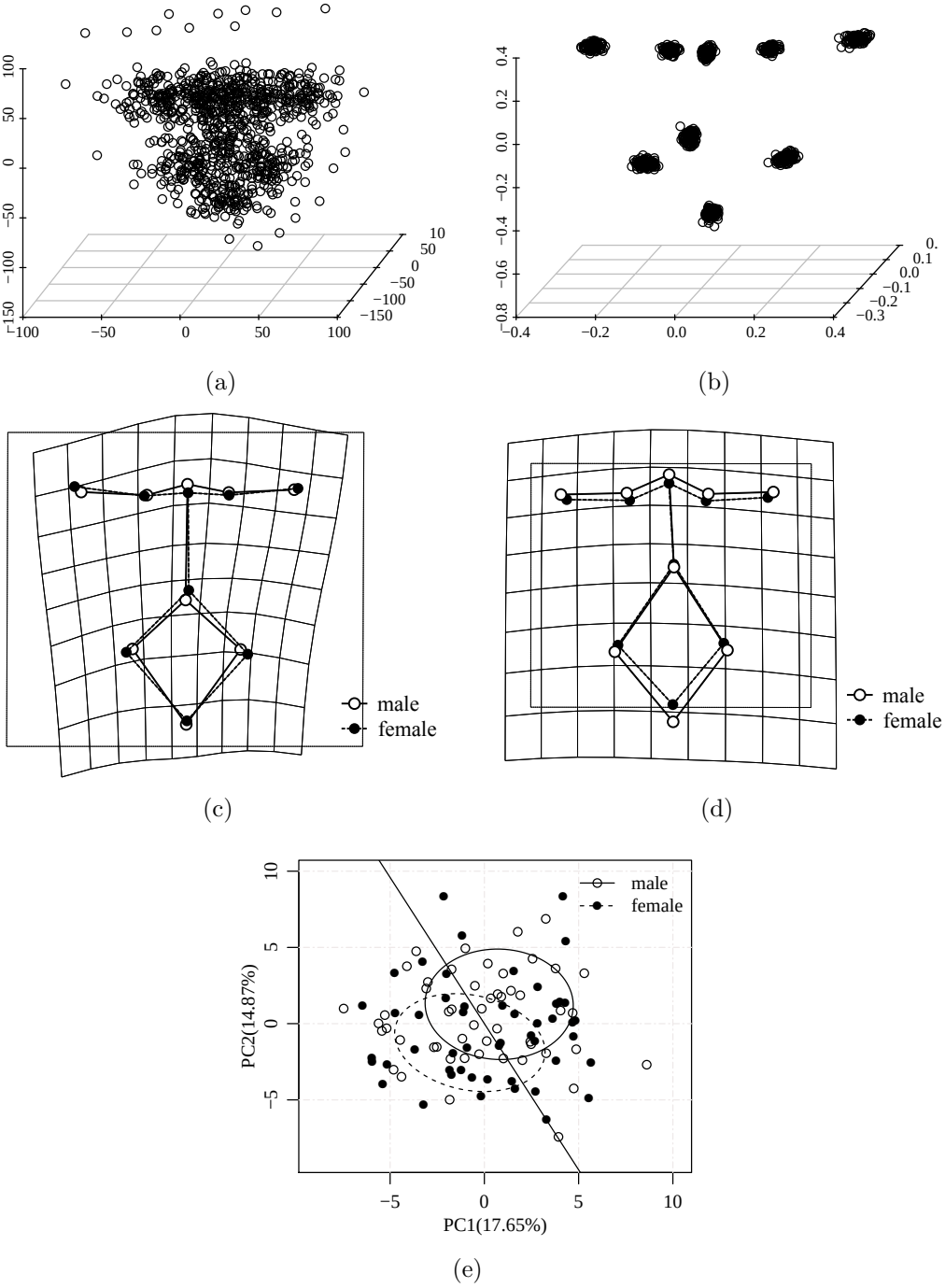


Figure 2.3: Facial landmark analysis: (a) a sample of 101 landmark configurations; (b) an aligned sample of landmark configurations; (c) visualization of space deformation caused by a transition from the female to the male average shape, exaggerated 3 times; (d) TPS visualization of form change from the average male to the average female, including size; (e) scatterplot of PC1, PC2 scores of facial landmarks. The figure also shows the LDA decision boundary, optimally separating male and female groups.

cross validation scored a success rate of 72.28% and a posterior success rate of 76.24% using the Linear Discriminant Analysis (LDA) classifier.

Case study 2: Analyzing digitized models of tibial epiphyses

An actual GMM study that demonstrates all the methods described in the previous paragraph was performed on shapes of tibial epiphyses (lower and upper ends of tibia - in Figure 2.1(d)), the skeletal remains of an early medieval Great Moravian population are shown (Brzobohatá *et al.*, 2014). In the study, bones were assigned to different classes according to three criteria: age of death (young - adultus, old - maturus), sex (male, female), social status (castle or sub-castle derived from the location of the grave).

The shapes of tibial epiphyses were described by landmarks (proximal - upper end by 13, distal - lower end by 9). After alignment by GPA and application of a PCA, differences according to individual criteria were tested. It was found that sex and age significantly influence proximal shape while distal shape is only affected by sex. However, social stratification could not be proved in shapes of tibial epiphyses described by landmarks.

Brzobohatá *et al.* (2015b,a) extended the study by using samples from the early 20-th century and 21-st century.

2.3 Missing landmarks

It is clear from the above description of the methodology that it is important to have a relatively high number of specimens in the sample in order to obtain better coverage of the population. More importantly, every specimen in the sample must define all landmarks. This condition is sometimes difficult to fulfill since the nature of the material might be hundreds of years old, damaged or incomplete. In general, there are many statistical methods (Little and Rubin, 2002) that deal with missing data in the experiment, e.g. when estimating PCA (Ilin and Raiko, 2010). However, it is more interesting to recover incomplete shape data than just to deal with their statistical contribution. Moreover, the geometrical nature and high redundancy of information in landmark coordinates provide an opportunity to deliver more suitable and specialized procedures (Mitteroecker and Gunz, 2009).

The geometry's symmetry is a property that is often exploited. In particular, if a part of the specimen is missing, it can be filled in with landmarks available on the corresponding symmetric part of the shape by transforming them with respect to the symmetry.

Another approach uses thin-plate spline (TPS) interpolation to fill in missing data (Mitteroecker and Gunz, 2009). As mentioned above, TPS is a method often used in GMM for visualization purposes (Bookstein, 1997). The method takes two landmark configurations s and t with n corresponding landmarks as an input and provides a transformation function $f_{TPS}^{s \to t}$ of the entire R^n. TPS

transforms points from s to points of t as well as the rest of the space in order to minimize second derivatives and enforce smoothness, i.e. in 2D

$$\iint \left(\frac{\partial^2 f_{TPS}^{s\to t}}{\partial x^2}\right)^2 + \left(\frac{\partial^2 f_{TPS}^{s\to t}}{\partial x \partial y}\right)^2 + \left(\frac{\partial^2 f_{TPS}^{s\to t}}{\partial y^2}\right)^2 dxdy$$

This formulation leads to a solution for $\vec{x} = (x, y)$ in the form of

$$f_{TPS}^{s\to t}(\vec{x}) = a_0 + a_x x + a_y y + \sum_{i=1}^{n} \phi(|\vec{x} - s_i|^2) w_i$$

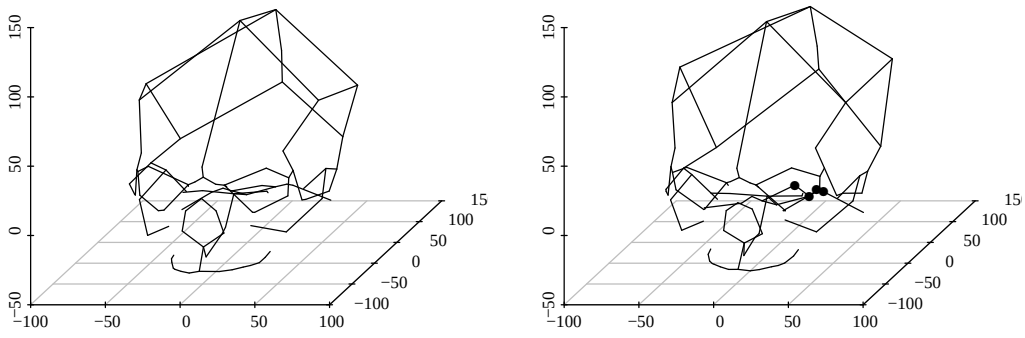
With $\phi(r) = r^2 \ln(r)$, $f_{TPS}^{s\to t}(s_i) = t_i$ and additional constraints $\sum_i^n w_i s_i = 0$ and $\sum_i^n w_i = 0$ it can be solved for a and w. There are more variants of TPS since it is a widely used method in various other fields. TPS can be used to estimate missing values by computing TPS between available landmarks of full configuration s and corresponding landmarks of partial configuration t. Afterwards, missing landmarks of t are filled in by transforming these landmarks from s using TPS interpolation. The advantage of this method is that only a single full landmark configuration (reference) is needed; however, it should be selected with care. The results are highly dependent on particular landmarks that are later computed. For example, they depend on their relative position with respect to the other existing landmarks, or on whether they are located inside the convex hull of the existing landmarks or not.

One recent study (Brown *et al.*, 2012) compared various methods for missing morphometric data imputation, while proposing a method based on the regression of available values against principal components using a Bayesian estimate of PCA (Arbour and Brown, 2014).

Case study 3: Nile Valley skull analysis

Missing landmark computation was used in a morphometric research study of skull remains (crania without jawbones) from two ancient populations of the Nile Valley (Chvojková *et al.*, 2010). The first population originated from the Mirgissa area of Lower Nubia between 1890 and 1580 B.C. and is represented by a sample of 205 skulls. The second population originated from Wadi Qitna in Upper Egypt during the period between the 3-rd and 5-th centuries A.D. and is represented by 38 skulls. Figure 2.4(a) shows an example of full landmark configuration. Both samples are described using 89 craniometric landmarks. For more details of landmark location, see Chvojková *et al.* (2010). The first sample has only 18 complete landmark configurations, while 6.70% of all landmarks are missing due to the poor condition of the remains. In the second sample, 13.60% of landmarks are missing while only one specimen shows a full landmark configuration.

Following the previously mentioned approach (Mitteroecker and Gunz, 2009), missing landmarks were computed by TPS interpolation. Subsequent



(a) (b)

Figure 2.4: Nile Valley skull analysis: (a) Full craniometric landmark configuration (89 landmarks); (b) Example of a specimen with incomplete landmark configuration. Full black points represent landmarks that were were missing and had to be computed (using the TPS method in this case).

analysis proved a significant difference in the shapes of the Mirgissa and Wadi Qitna populations. Figure 2.4(b) shows a typical example of missing landmarks that were computed using the method.

However, TPS does not reflect the variance of the sample, i.e. it might result in an undesirable shape configuration. If there are enough full landmark configurations in the sample, they can be used to create a statistical model of the shape and use it to predict the full configuration for partial input, in terms of least square differences.

Input sample $\mathbf{L}$ is split into $\mathbf{L}_f$ with full landmark configurations and $\mathbf{L}_m$ with incomplete specimens. $\mathbf{L}_f$ is aligned with GPA, and variation of $\mathbf{L}'_f$ is decomposed into principal components $\{w_k\}_{k=1}^{n}$. Incomplete specimens $L_{m,j}$ are aligned to a mean shape $\bar{L}'_f$ of $\mathbf{L}'_f$. A vector of coefficients $\tilde{\alpha}$ is optimized so the difference between corresponding existing landmarks is minimized, i.e.

$$\arg\min_{\tilde{\alpha}} \sum_{i \in O_j} |\bar{L}'_{f,i} + \sum_{k=1}^{m} \tilde{\alpha}_k w_{k,i} - L'_{m,j,i}|^2 \tag{2.2}$$

where O_j is a set of landmark indices that are not missing in a particular specimen j and m is the number of statistically significant principal components, e.g. determined by the broken-stick criterion (Peres-Neto *et al.*, 2005). Missing landmarks of incomplete specimens are replaced with the landmarks of estimated complete landmark configurations, computed using model and optimized coefficient vector $\tilde{\alpha}$, i.e.

$$\tilde{L}_j = \bar{L}'_f + \sum_{k=1}^{m} \tilde{\alpha}_k w_k$$

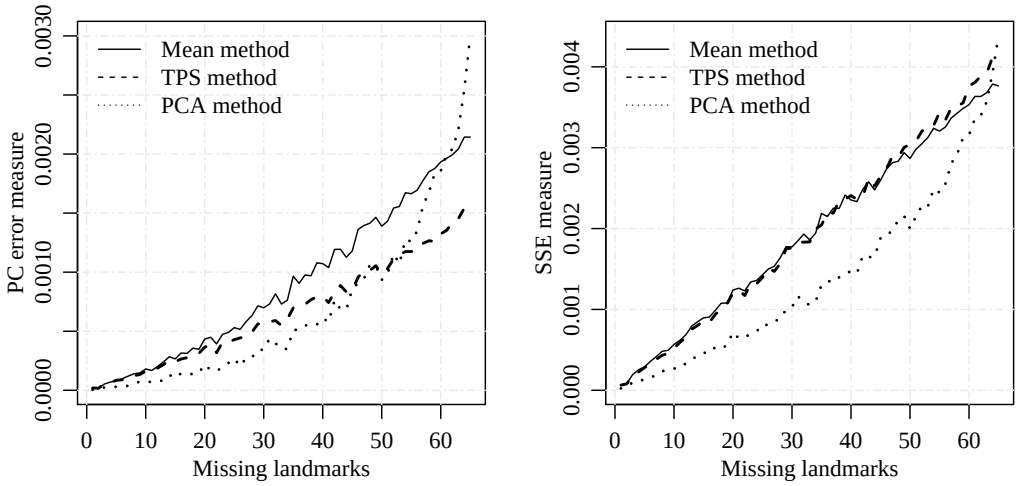


Figure 2.5: Results of missing landmark computation methods in a comparison experiment on Mirgissa skull samples: (a) PC error measure of statistically significant components (23); (b) SSE measure of all landmarks (89).

Three different approaches (mean shape substitution, TPS interpolation and the PCA model) are compared by randomly removing various numbers of landmarks and recomputing them. Two error measures are evaluated: the sum of squared errors (SSE) and the squared distance in a subspace of statistically significant principal components (PCE).

$$SSE = \sum_{i=1}^{n} |L_i - \tilde{L}_i|^2$$

$$PCE = \sum_{i=1}^{m} |\alpha_i - \tilde{\alpha}_i|^2$$

Figure 2.5 shows a comparison of the methods. Missing landmark computation is repeated 100 times for each number of missing landmarks and an average of these measurements is presented. It appears that the PCA method outperforms the TPS and mean shape methods for a reasonable number of missing landmarks, but overfits the available data for a higher number of missing landmarks. Adding a regularization term to minimization scheme 2.2 would solve this problem. The difference between the TPS and mean methods used for PCE and SSE measures is caused by the fact that PC error measure filters out noise in high-order components while keeping significant information relevant to group statistics.

2.4 Asymmetry

Object size and shape variability represent two upper layers of shape analysis and description. Another individual feature that is manifested in shapes of nature is symmetry, of which there is a great variety of types. On the one hand, there are very complex types — spherical symmetry, observed often in molecules, cells and unicelular organisms, which allows an object to be cut in an arbitrary direction through the center in order to obtain two symmetrical parts. On the other hand, there are simpler types such as *bilateral symmetry*, which is found in higher species defined by a single plane of symmetry, i.e. the medial plane (Conway *et al.*, 2008).

Generally, it holds that the more complex the organism is, the more likely it is to break the symmetry due to the specialization for the environment it lives in. For example, the human body is bilaterally symmetrical, but even this symmetry is often broken. Handedness is an example of broken bilateral symmetry, which is exhibited by larger bones and muscles in the left or right arm and connected with faster movement and improved physical strength. In many other cases, it can reflect various facts about the life of a specimen or population.

Traditionally, bilateral asymmetry has been studied by comparing measurements of corresponding symmetrical parts of an object (Palmer and Strobeck, 1986). Recently, bilateral asymmetry has begun to be studied with regard to its application to landmark data (Klingenberg *et al.*, 2002). In order to be able to compare symmetric features, landmarks in a configuration have to be divided into two classes:

- Paired landmark — located on a shape feature that has its symmetric counterpart in the shape according to its bilateral symmetry; a landmark located on the symmetric feature is also included in the configuration.
- Non-paired landmark — located on a feature that lies or should lie on the medial plane.

The asymmetry of an individual specimen is extracted as follows. In the first step of the analysis, mirror landmark configuration must be created by negating one coordinate of the landmark. The landmarks then have to be reordered so that each paired landmark swaps with its symmetric counterpart. Non-paired landmarks keep their position in the configuration. Subsequently, original and mirror configurations are aligned by GPA. Differences between landmarks in the original configuration and its paired counterpart in the mirror configuration arise due to shape asymmetry. If the shape is absolutely symmetric, the differences vanish; however, the greater the asymmetry, the greater the misalignment. Because GPA is used for alignment, properties of GPA are projected to asymmetry evaluation, i.e. large asymmetry in one landmark is distributed into small asymmetries of all landmarks and one larger asymmetry of the respective landmark, since GPA distributes misalignment error evenly.

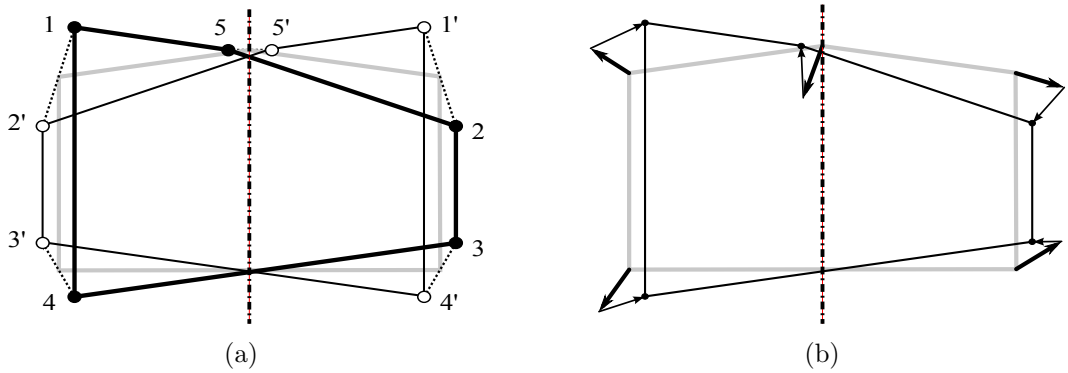


Figure 2.6: Explanation of bilateral asymmetry extraction from landmark data: (a) Original (solid thick line and black landmarks) and mirror shapes (solid thin and white landmarks) aligned by mirror-corresponding paired landmarks (connected by dotted lines), where the symmetrized shape is the average of the two (gray solid thick line); (b) Individual asymmetry (solid line) is composed of group directional asymmetry (thick arrows) and specimen fluctuation asymmetry (thin arrows).

The average of the aligned mirror and original landmark configuration is symmetric — a symmetric component of the shape; whereas the difference from the original configuration is called the asymmetry or asymmetric component of the shape. Figure 2.6(a) illustrates the procedure described above.

In *group asymmetry* analysis, all landmark configurations are aligned by GPA beforehand. The average of shape asymmetry components, also called *directional asymmetry* (DA), represents the group asymmetry. The difference in directional asymmetry from a particular specimen asymmetry is called *fluctuating asymmetry* (FA). An illustrative explanation is shown in Figure 2.6(b). The difference between individual shape asymmetry and FA is that shape asymmetry is distributed normally with a mean equal to DA, while FA resembles uniformly distributed noise with a zero mean. For DA interpretation, it is important to know its actual direction; whereas, in the case of FA, only its magnitude plays an important role.

The contrast of asymmetric components in group pairs can be confirmed by multivariate tests for mean equality, such as Hotelling's T^2 or non-parametric variants. Multivariate analysis of variances (MANOVA) can be used to compare multiple groups at once or to analyze the effect of independent continuous variables on an asymmetric component.

Asymmetry and, especially, FA are very sensitive to accurate measurement. Often picking landmarks is not a completely error-free process. Because asymmetry magnitudes can come very close to error as a result of improperly measured landmarks, it is necessary to make repeated measurements in order to prove that the variation caused by repeated measurements is much lower that the one reflected in FA. In order to isolate sources of variation caused

by asymmetry, measurement error and other sources from mis-alignement error, repeated GPA is required before applying multivariate statistics using Procrustes (M)ANOVA (Klingenberg *et al.*, 2002).

Case study 4: Asymmetry as a proof of socioecomomic structure

Asymmetry analysis in anthropology is frequently used to prove differences in life conditions of studied populations. For example, Bigoni *et al.* (2013a,b) used asymmetry as a major tool to analyze the socioeconomic structure of a particular area (the Mikulčice settlement) in Early Medieval society (Great Moravian period). The studied sample consisted of 200 well-preserved skulls (crania without jawbones).

The sample was further divided into castle (129 specimens, 61 female and 68 males) and sub-castle groups (71 specimens, 38 females and 33 males) according to the location of burial ground and under the assumption of a connection with social class affiliation. The Mikulčice sample was compared with a representative sample from the Pachner Collection from the 1930s, which examined lower socioeconomic classes of Prague. Three landmark configurations were designed to describe the facial area, base of the skull and cranium so that these areas could be analyzed separately.

It was shown that:

- Differences in asymmetry of the Mikulčice sample were more obvious among females (statistically significant between the castle and sub-castle groups).
- DA was more pronounced in the Mikulčice than in the Pachner sample due to the more solid diet examined, which required greater mastication of food.
- On the other hand, FA was stronger in the Pachner Collection which indicates that this population faced higher environmental stress then the Mikulčice population who lived in relatively favorable conditions.
- Inside the Mikulčice sample, DA demonstrated that levels of biomechanical stress connected to the quality of the diet were higher in the sub-castle population who probably had to eat grittier and more solid foods such as cereals, while in the castle group the population lived on a more meat-based diet.
- Taking sex into consideration, while looking at the Mikulčice sample, the difference between males in the castle and sub-castle groups was not so strong. However, females of the sub-castle group were much smaller and much more directionally asymmetric. This supports the hypothesis that sub-castle females ate lower amounts of protein and had a grittier diet than females of the castle group and that both male groups from the castle and sub-castle groups were from more mixed sub-classes.

- FA is connected to increased environmental stress and the simultaneously increased variability of a population (genetic stress). Males in the Mikulčice sample exhibited no significant differences, whereas, for females, the differences were noticeably higher in the castle population to the point that they were comparable with FA in the Pachner Collection sample. The higher population variability derived from this finding could be connected to the phenomenon of patrilocality (females move to their husbands' homes instead of staying near to their parents' or family residence), common in Early Medieval societies. Alternatively, the high FA could indicate that, in the castle region, many highly environmentally stressed, lower socioeconomic females (servants) were buried.

Case study 5: Jawbone microevolution over a 1200-year period

One important case study, which demonstrates the methodological contribution of this thesis, analyzed the microevolution of the jawbone (see Figure 2.1(a)) over the last 1200 years (Bejdová *et al.*, 2013). The study is based on four samples from four different eras:

- Great Moravian Empire sample (GM) — 900–1000 A.D.; the sample consists of 81 skulls with jawbones.
- High Middle Ages sample (HM) — 1100–1400 A.D.; 53 skulls with jawbones. It must be noted that the samples were not preserved very well due to burials in coffins whose top lids were often destroyed after years in the ground, resulting in damage to the interior remains.
- Early Modern Ages sample (EM) — 1600–1800 A.D.; 64 skulls with jawbones. Some specimens were also damaged for the same reason given for the HM sample.
- Recent Population sample (RE) — 2000 A.D.; 92 skulls with jawbones of living people whose images were obtained by CT imaging in hospital (Department of Radiology in Na Homolce Hospital, Prague).

Special care was taken so that no patients with abnormal jaw development were included in the sample. Other factors were also considered, such as age, sex and health condition, so that their comparison would be unbiased by these factors. Three pre-processing procedures are performed before data enter the analysis.

The first is the computation of missing landmarks, without which the study would not be even possible. This is because the HM and EM samples were so damaged that they could not provide enough complete individuals for GMM analysis. Table 2.7 summarizes the structure of the samples and the extent of their damage. For landmark computation, the method based on the PCA model of full landmark configurations was used.

Population	**Sample size (#)**	**Missing landmarks (%)**	**CV Raw (%)**	**CV Sym (%)**	**P-value**
GM	81	0	69.14	79.01	0.0010
HM	53	7.3	56.60	90.57	0.0130
EM	64	16.3	64.06	60.94	0.0280
RE	92	0	82.61	80.43	0.0001

Table 2.1: Jawbone sexual dimorphism discrimination results. The **CV Raw** column represents the LDA cross-validation success rate of raw landmark data, while the **CV Sym** column represents the success rate of data with asymmetry and allometry removed.

The second step is the elimination of asymmetry from the shape data. In this case, asymmetry is not expected to carry any information for distinguishing between sexes. On the contrary, it may act as a systematic noise or error which disguises the sex differences. Asymmetry is eliminated so that the symmetric component of each individual is used instead of raw landmark data.

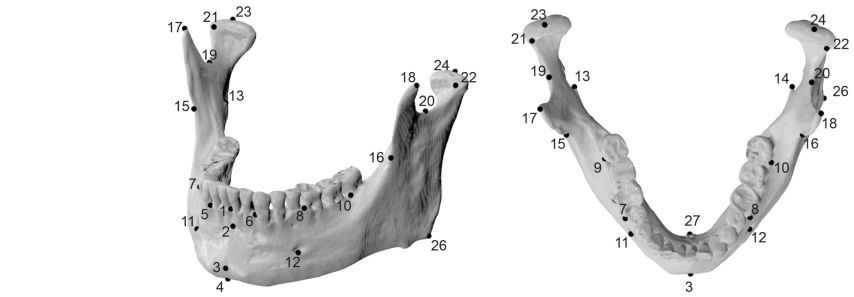
The third step is the removal of the effect of allometry. Allometry is the relation between the size of the specimen and its shape, i.e. a deviation from isometric scaling. Studying allometry is based on the principle that larger specimens have specifically different shapes compared to smaller ones, due to different rates of body part development while accommodating to the environment. In GMM, allometry is expressed by a linear model of size (e.g. centroid size, CS) as an independent variable and shape α as a multivariate-dependent variable.

$$\alpha_{CS} = \mathbf{a} + CS\mathbf{b} \tag{2.3}$$

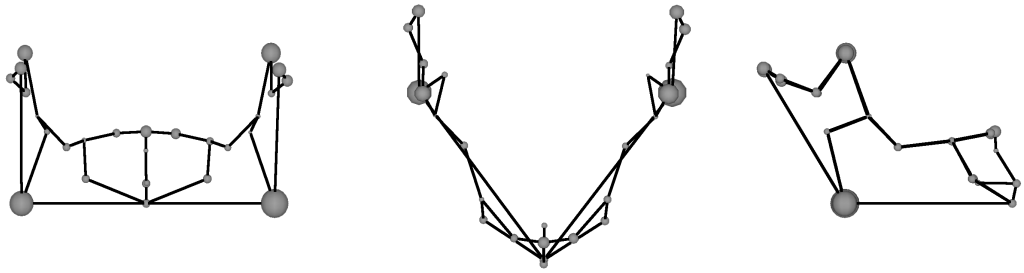
where α_{CS} is a vector of PC scores, which, if substituted to Equation 2.1, will give a landmark configuration of an average specimen with a centroid size equal to CS.

The effect of allometry on shape descriptors is removed when analyzing residuals $r_L = L - L(\alpha_{CS})$ instead of landmark configurations L. If there is no significant relation between size and shape, i.e. the average shapes of various sizes are similar to the average shape of the sample, then the variability of residuals $r_L = L - L(\alpha_{CS})$ will be identical to the variability of L.

After these steps, landmark residual r_L data were tested for separability with respect to sex using the LDA classifier. The results of leave-one-out cross-validation are shown in Table 2.1. In the same table, it is evident how the elimination of asymmetry and allometry elimination changed the separability of the data on the cross-validation rate of data not treated for asymmetry and allometry. It seems that these two additional factors display a decreasing tendency relative to sexual dimorphism. However, this theory was not



(a)



(b)

Figure 2.7: Jawbone sexual dimorphisms: (a) Full landmark configuration on all historical jawbone samples (27 landmarks; see Bejdová *et al.* (2013) for exact definitions); (b) An example of sexual dimorphism in a Great Moravian sample (900–1000 A.D.), where the size of the landmark discs shows the measure of the impact on sexual dimorphism.

investigated further.

The localization of sex differences is obtained from a discriminant function which defines a decision plane

$$c = \vec{v}\vec{x}$$

where $\vec{v}$, if used in a weighted sum of statistically significant principal components as a coefficient vector, i.e $\sum_{k=1}^{m} v_k w_k$, gives landmarks influence in discrimination (see Figure 2.7(b)). Apart form shape comparison, size differences between sexes were also compared in all the samples, which revealed an increasing trend over time. For the full analysis of all samples, the reader should refer to Bejdová *et al.* (2013). Although, differences in shape between sexes were statistically significant in all populations, they were always exceeded by the differences between populations. Therefore, a population-specific sex classifier is needed, in this case.

Sexual dimorphism is much stronger in some bones in the human body, to the extent that it can overcome population differences. This is demonstrated in the last case study of this chapter, which also shows a transition from methods

requiring user-localized, explicitly corresponding landmarks to algorithmically-computed correspondence.

2.5 Non-landmark methods

Until now, shape analysis has been based on the most simple geometrical primitives — landmarks — specifically defined by the researcher. The previous sections specified how to computationally fill in missing data, even when represented by landmarks. Naturally, more complex primitives can now be used to represent the shapes of real world objects. However, these shapes might be difficult to digitize, not to mention the complications involved in defining correspondence between different instances or group-wise correspondences. Overcoming these obstacles will allow shapes and shape features, which have previously been difficult or impossible to describe by landmarks, to be digitized.

In particular, the analysis of outlines in 2D images I_i will be described. It is assumed that the images are taken under controlled conditions in the context of preserving shape features while projecting into 2D space. Perspective projection caused by a pinhole-type camera sensor results in loss of information on the depth of the structure. In order to reduce the loss, analyzed features should lie in the plane parallel to the projection plane. If this is not possible, at least all of the shapes should be oriented identically, so that the distortion is similar. Features such as edges and landmarks, which define the analyzed outline, should be clearly visible in the image.

The first step of the analysis is to extract the shape from the image. In order to apply the method, which is described further on, the outline is defined as a curve starting and ending at points defined by a pair of landmarks L_s and L_e. The curve itself is automatically segmented as a sequence of neighboring pixels (p_i is a neighbor of p_{i+1}) with minimal overall cost

$$C_P = \sum_{i=0}^{k} c(p_i), \text{for} P = \{p0, \ldots, p_k\}, p_0 = L_s, p_k = L_e$$

The cost of a pixel p is defined as

$$c(p) = |\nabla I(p)|^2 - max_{q \in I}(|\nabla I(q)|^2)$$

The actual sequence P is found using dynamic programming according to Martelli (1976). The sequence itself cannot be used to compare between different specimens, since it is not invariant due to resolution and conditions of image acquisition. Therefore, the use of approximative curve representations is a reasonable advancement as a next step. Also, reduced information from curve representations might serve as a guideline for simplified in-field shape classification.

Six different approximative curve representations were used to analyze the outline shape. The first three representations use the term *semilandmark*. A semilandmark is a point-primitive, in that it is not defined by a user but generated algorithmically, e.g. by equidistant subdivision along the curve, and so on.

- Semilandmarks-by-angle — This method is used to represent simple convex arc-like curves with a center $\mathbf{c}$ in the middle of the line connecting L_b and L_e. For a given number of semilandmarks, all are distributed along the curve so that subsequent points hold equal angles with the center, i.e. $\angle p_i \mathbf{c} p_{i+1} = \angle p_j \mathbf{c} p_{j+1}$.

- Semilandmarks-by-arc — For a given number of semilandmarks, semilandmarks are distributed along the curve so that the curves between the subsequent landmarks have an equal length.

- Distances-by-angle — This method is similar to semilandmarks-by-angle. However, instead of coordinates of semilandmarks, which are redundant, only the distances from the center $\mathbf{c}$ are stored.

- Normal polyline subdivision — The curve is subdivided hierarchically by casting a normal from the center between the endpoints of one segment to the segment. The intersection divides the segment into two subsegments. For a given number of coefficients, distances $n_{i,j}$ of centers and intersections are recorded for i-th segment of j-th level of subdivision.

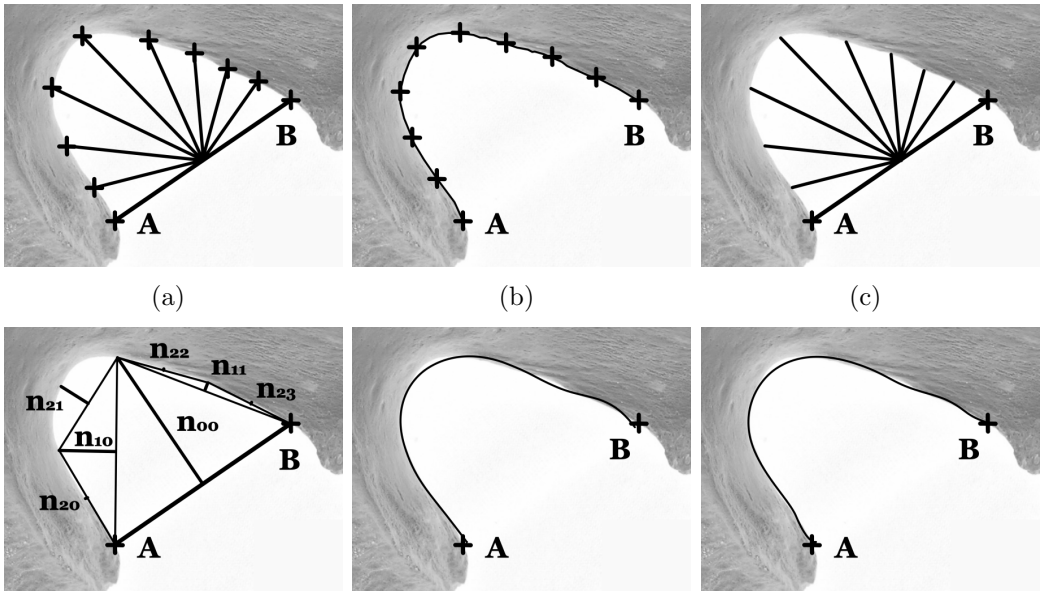
- Fourier coefficients of circular harmonics — This method makes use of the fact that the curve is a simple arc which can be expressed in polar coordinates, $r(\phi)$. Since only one angle parameter is involved in a polar representation of the curve, circular harmonics can be used to encode its shape very effectively, i.e. by a small number of coefficients. The circular harmonics are a set of basis functions that can efficiently represent a function defined over a circle, given by the Fourier series

$$B = \left\{ \frac{1}{\sqrt{2\pi}}, \frac{cos(\theta)}{\sqrt{\pi}}, \frac{sin(\theta)}{\sqrt{\pi}}, \frac{cos(2\theta)}{\sqrt{\pi}}, \frac{sin(2\theta)}{\sqrt{\pi}}, \dots \right\}$$

  It can be seen that the basis functions are orthonormal and orthogonal, which means they can be used to decompose function $r(\phi)$.

$$f_i = \int_0^{2\pi} r(\phi) B_i(\phi) d\phi$$

  Figure 2.8(e) shows how the curve is reconstructed using only five coefficients. The advantage of polar space descriptors lies in their invariance to rotation. Kazhdan *et al.* (2003) shown that their superclass — spherical harmonics outperform orientation-dependent descriptors on shapes roughly aligned by principal axes.



(a) (b) (c)

(d) (e) (f)

Figure 2.8: Various curve representations of segmented curves in 2D using an example of the *incisura ischiadica major*: (a) Semilandmarks-by-angle; (b) Semilandmarks-by-arc; (c) Distances-by-angle; (d) Normal polyline subdivision; (e) Curve approximation reconstructed using five Fourier coefficients of circular harmonics; (f) Curve approximation reconstructed using eight Legendre coefficients of circular harmonics.

- Legendre coefficients of circular harmonics — This method is identical to the previous one with the only difference being that, as a set of basis functions, Lagendre polynomials $P = \{P_0, \ldots, P_n, \ldots\}$ are used

  $$P_0(x) = 1, P_1(x) = x, P_{n+1}(x) = \frac{2n+1}{n+1} x P_n(x) - \frac{n}{n+1} P_{n-1}(x)$$

  Rao *et al.* (2010); Sharma and Dhole (2013) have demonstrated the high classification efficiency of Legendre moments in combination with SVM classifier.

All described representations are shown in Figure 2.8.

Case study 6: Sexual dimorphism of the *incisura ischiadica major*

The *incisura ischiadica major* (greater sciatic notch) is an anatomical feature that is found on the pelvic bone (see Figure 2.1(b)), which is well known for high sex discrimination ability. Velemínská *et al.* (2013) studied this ability in photographs of pelvic bones by using the representations described above.

In the experiment, two samples were used: The first one consisted of 114 photographs of adult human pelvic bones (57 males, 57 females) from a collection of Euroamericans at the Maxwell Museum (MM), University of New

method	**coefficients (#)**	**success rate (%)**		
		pooled	**MM**	**UNAM**
SL-by-angle	4	90.83	92.11	88.70
SL-by-arc	9	90.39	90.35	89.57
Distances	7	90.83	90.35	88.70
Normal	7	91.27	89.47	90.43
Fourier	5	92.14	92.11	92.11
Legendre	22	92.14	89.47	89.57

Table 2.2: The *incisura ischiadica major* classification results of SVM with a radial basis. The success rate of leave-one-out cross-validation for each population sample as well as for the pooled sample.

Mexico, Albuquerque; the second sample consisted of Hispanics from the Osteological Collection of the Universidad Nacional Autónoma de México (UNAM), made up of 115 photographs of adult human pelvic bones (55 females and 60 males). All photographs were take in the same position under controlled and identical conditions.

In the first step, landmarks indicating the beginning and end of the curve must be located. In this case, landmarks L_b and L_e were located at at the tip of the piriform tubercle and at the base of the ischial spine in each image (see Figure 2.8). Outlines were segmented out as mentioned above and the number of coefficients was optimized for each representation to establish the best discrimination results. To evaluate discrimination ability, a Support Vector Machine (radial basis) classifier was used (Cortes and Vapnik, 1995). Discrimination ability is expressed by the leave-one-out cross-validation success rate. Table 2.2 summarizes the results and the optimal number of coefficients for all the methods in individual as well as in pooled populations.

In order to visualize shape variation, a semilandmarks-by-arc representation with a high number of semilandmarks (32) was generated for all curves. Semilandmarks were considered to correspond with each other, since they appeared along the curve across the entire sample. Principal component analysis was performed in the same way as for landmarks in all examples of this chapter. It was found that the first two components represent 97.85% of the total variability. In this way, the curve can be parameterized with a loss of 2.15% of information by only two numbers — the first two principal component scores. In Figure 2.9, shapes matching part of the shape-space projection given by PC1 and PC2 are shown.

Incisura ischiadica major analysis revealed that male shapes tend to be higher and asymmetric, whereas female ones tend to be lower and symmetric. Female variability is much lower than male variability, which is connected to the specialized role of the pelvic bone in females. High variability in male

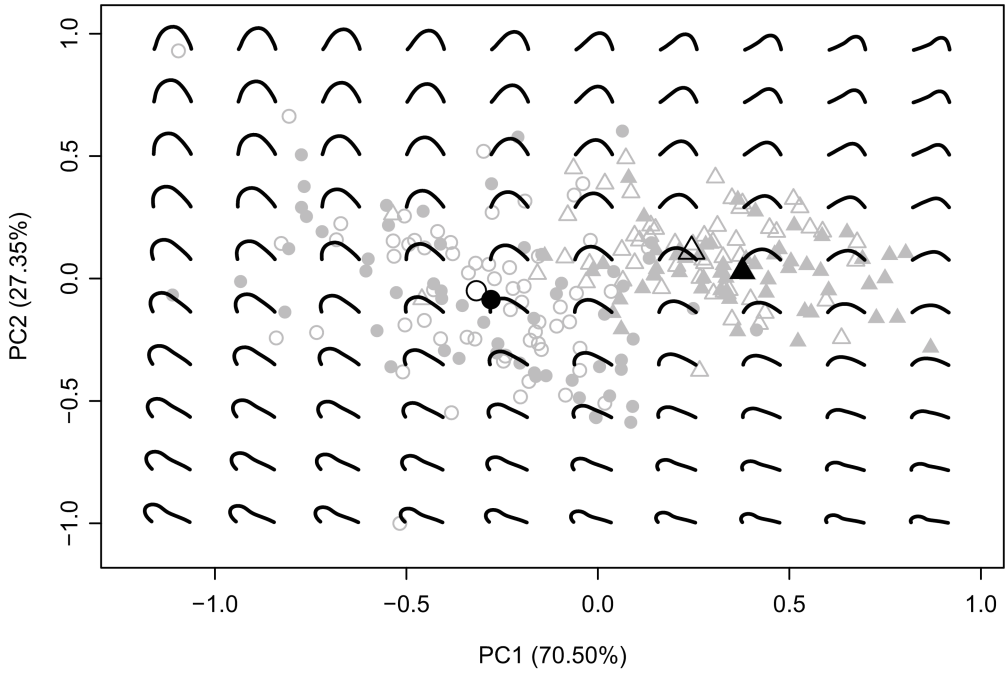


Figure 2.9: Shape variability of the *incisura ischiadica major*. Circles correspond to males; triangles correspond to females. Full symbols correspond to the MM Collection sample, hollow symbols correspond to the UNAM sample and symbols represent the subgroup mean.

shapes causes higher mis-classification. The overall discrimination rate exceeds 92% in the best case and is generally better for pooled samples in all representation, due to the better generalization of the classifier when a more variable training set is provided.

The analysis also proved that the methods are very robust to L_s and L_e localization precision. The endpoints were shifted randomly along the curve with a magnitude of 10% of the overall curve length, during which the success rate was measured. The Legendere polynomial method was the most susceptible to noise, with a drop in the success rate of 3.2%, when the noise was applied to both training and testing specimens. The most resilient representation was found to be the semilandmarks-by-arc method, with a drop on average of only 0.1%. All methods produced marginally better results when noise was applied to test specimens only.

2.6 Conclusions

In this chapter, the general topic of the thesis was introduced and demonstrated in many practical examples.

It was shown that shape analysis is performed on landmark data, requir-

ing explicitly user-defined landmarks and that the only difficult task is to normalize landmark configurations to obtain comparable vectors which can represent shapes. General statistical methods, such as multivariate statistical tests for mean equality, classification methods and linear modeling, are used to answer questions and prove theories. The author contributed by applying the methodology to an analysis of tibial epiphyses (Brzobohatá *et al.*, 2014, 2015b,a).

If some landmarks cannot be localized due to damaged specimens, there are statistical methods of dealing with this situation or by the imputation of missing data with extrapolated values. Missing landmark computation was performed by Chvojková *et al.* (2010); Bejdová *et al.* (2013) with the contribution of the author.

Asymmetry is a very sensitive shape property and it is used to prove phenomena connected to environmental impact, genetics and lifestyle. It is also an independent feature of shapes that mingle with shape variability and can distort sample separability. The author has analyzed the asymmetry of human skulls in connection with the socioeconomic structure of a Middle Ages society (Bigoni *et al.*, 2013a,b).

Landmarks are not the only geometrical primitive used to describe shapes. Two general approaches have been introduced to handle non-landmark data with the author's contribution (Velemínská *et al.*, 2013). The first uses shape transformation from a geometric representation to a parameter space and continues using these parameters with statistical methods. Alternatively, densely sampled geometry can be constructed based on the original shape representation and normalized, after which PCA can be used to extract trends in high dimensional data. The second approach will be investigated in the rest of the text using data that provide more information on shape, even though they do not explicitly define correspondence.

Other fields that benefit from statistical shape modeling are worth mentioning. For example, a statistical model could be used as a constraint in model-based image segmentation. Cootes and Taylor (1992) introduced *active shape models* to segment simple outlines in two-dimensional images. The advantage of model-based segmentation lies in its high robustness against noise or partial occlusion. Therefore, it could be used to segment noisy data such as ultrasound images (Cootes *et al.*, 1993) or for video tracking (Koschan *et al.*, 2003). It has subsequently been extended to model not only shape, but also intensity, in what are known as active appearance models (Cootes *et al.*, 1998).

On the other side of the spectrum, there is a popular and much older application of linear modeling of raster images of human faces, called eigenfaces (Sirovich and Kirby, 1987). The approach completely ignores geometrical interpretation of content and simply relies on a picture taken from a normalized position, where correspondences are made by an overlayed pixel.

Chapter 3

Triangular mesh analysis

> You are in pretty good shape for the shape you are in
>
> Dr. Seuss

Boundary is an integral part of shape abstraction and represents an interface between shape interiors and the background. Boundary is a set of points in space where various physical properties such as atomic density are subject to rapid change. In the case of many real-world objects, often it is the only part of their shape that is visible. Hence, capturing the boundary of an object provides complete unambiguous information about that object's shape geometry. In tree-dimensions, a boundary is actually a surface, 2D manifold in $\mathbb{R}^2$. One common way to represent a boundary is to densely sample the surface and connect the sample points to a *triangular mesh*.

The triangular mesh is a subclass of a general polygonal mesh. It is a data type used in many fields of computer graphics, geometry modeling and engineering in general due to its reliable approximation and adaptability properties. In practice, it is used as a design tool to create shapes from scratch since there are many algorithms which allow the creator to work with the representation of shapes naturally. It is also used to efficiently describe existing objects with a controllable amount of information loss. Either way, triangular meshes are easy to store and, due to advances in current graphics hardware, very easy to visualize on a large scale.

State-of-the-art shape analysis often works with surface- (range-) scanning technologies (see Figure 3.1) by exploiting optical properties of the object's surface in order to sample its boundary and produce a dense cloud of points connected into a triangular mesh. These devices are often able to capture a photographic image of the surface, and providing a surface texture that reproduces an even more realistic look of the digitized object. Since the focus of interest here is shape which is comprised of all geometric information, surface texture, color and any other non-geometric information associated with the mesh is unimportant. However, color variations can give practitioners infor-

(a) (b) (c)

Figure 3.1: Various surface-geometry capturing devices used for shape analysis: (a) box laser scanner (Roland LX-1200); (b) optical surface scanner (Canfield Vectra 3D); (c) hand-held optical scanner (Artec Eva).

mation about an object's geometry that goes beyond the resolution of scanning devices, which means that texture can be used for better shape-feature localization.

The benefits of shape digitization are obvious. It allows researchers to keep and preserve accurate information about the shape of a unique object's geometry, which could otherwise become unstable over time due to growth or decomposition. There are also drawbacks to the approach, which are connected to the limitation of scanning technologies and the imperfect way they are handled. Acquired triangular meshes sometimes exhibit artifacts, cracks, holes, non-manifoldness, loss of detail, etc. Therefore, they must be manually treated, which can lead to inaccuracies in shape approximation, especially when global operations such as smoothing or remeshing are used.

Current applications of geometric morphometry in anthropological research are aimed only at extracting the simplest primitives from these digitized shapes — landmarks. These landmarks are subsequently processed by the means described in the previous chapter. The historical motivation for this approach is understandable since its methodological framework existed before the appearance of surface scanning. On the other hand, the reduction in information by the selection of only a few landmark points is unnecessarily radical. Moreover, the selection of landmarks by the researcher in the planning phase of the study significantly influences the outcomes of the research at the end.

Before progressing further, the basic formalism for triangular meshes should be outlined. Triangular mesh T is defined as a tuple $T = (V, F)$, where V is a set of n vertices $v_i \in \mathbb{R}^3$

$$V = \{v_1, v_2, \ldots, v_n\}, v_i = (x_i, y_i, z_i)$$

F is a set of m triangular faces

$$F = \{f_1, f_2, \ldots, f_m\}, f_i \in V \times V \times V$$

For each f_i three edges can be defined

$$\{e_{i,1}, e_{i,2}, e_{i,3}\} = \{(f_{i,1}, f_{i,2}), (f_{i,2}, f_{i,3}), (f_{i,3}, f_{i,1})\}$$

The number of n vertices and the number of m faces are not usually equal for two triangular meshes representing the shapes of two different objects of the same kind. Therefore, the vertices and faces of these two meshes are not homologous, i.e they do not represent identical loci. In order to be able to compare two meshes or a set of meshes representing studied shapes, correspondences between these meshes must be found.

The first part of this chapter focused on various methods of correspondence construction, targeting practical usability in processing large data samples. The second part aims to utilize found correspondences through several tasks: shape variability and its visualization, shape asymmetry evaluation, longitudinal data processing and missing data computation.

3.1 Correspondence of mesh shapes

In the previous chapter, correspondences were explicitly constructed when the user selected a particular landmark manually. More importantly, the user defined the set of landmarks in consideration of what would reflect the features of the shape. These kinds of correspondences, *sparse correspondences*, are often used because they considerably reduce the amount of information and, furthermore, carry significant parts of that information. For example, in face recognition in 2D images, detection of facial landmarks is a key step.

However, in the case of featureless shapes such as the last example given in the previous chapter — the *incisura ischiadica major*, a different approach to correspondences must be taken. Apart from capturing shape descriptors, correspondences between multiple parts of a curve are created by interpolation of *semilandmarks*. Such correspondences, *dense correspondences*, apparently greatly increase the dimensionality of shape representation.

Matching the vertices of two meshes is a natural and often-used approach to create correspondences (van Kaick *et al.*, 2011). In the case of two meshes, T_Q and T_P, correspondences Σ between their vertices, $P = \{p_1, \ldots, p_{N_P}\}$ and $Q = \{q_1, \ldots, q_{N_Q}\}$, are defined as a subset of $P \times Q$, where each member $\{p_i, q_j\}$ represents a paired correspondence of i-th point from P and j-th point of Q.

The most common way to determine Σ is to find its nearest neighbor from the other point cloud with respect to the first point clouds (Figure 3.2(a)). This approach is asymmetric, since Σ with respect to the other point cloud

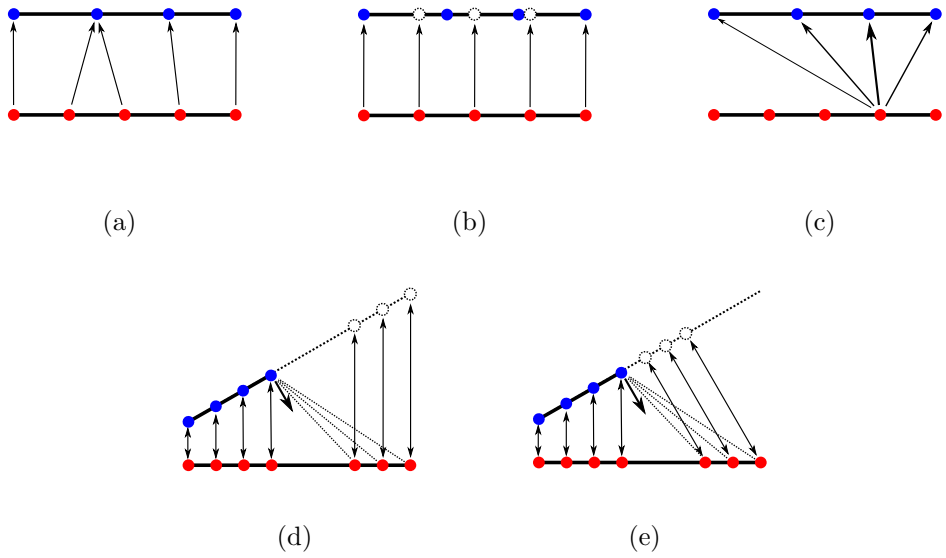


Figure 3.2: Approaches to mesh correspondences: (a) vertex-to-vertex correspondence; (b) vertex-to-point correspondence; (c) fuzzy correspondence; (d) plane intersection correspondence; (e) normal intersection correspondence.

can be completely different. Nevertheless, it is well-suited to the following algorithms.

In scenarios where there are sparse vertices and demands for higher precision or dissimilar point clouds, correspondences other than vertex-to-vertex type correspondences must be used. For any particular point in the first point cloud, a corresponding point is computed — a phantom point — based on the second cloud or mesh. Fuzzy correspondences, i.e. vertex-to-multiple-vertex, increase the robustness of the algorithms which are based on them, since their criteria are continuous and piecewise differentiable. See Figures 3.2(b), 3.2(c), 3.2(d) and 3.2(e) for examples.

Provided that the found corresponding point is too far from the vertex, due to the high disimilarity of the meshes or the missing partial geometry, it should be discarded. The decision to discard the correspondence can either be based on refined criteria or just on a threshold on the distance.

3.2 Mesh registration

Mesh and point cloud registration is a major tool in statistical mesh shape analysis. Comprehensive surveys on mesh and point cloud registration have been conducted by Audette *et al.* (2000); van Kaick *et al.* (2011); Tam *et al.* (2013). In this section, existing mesh and point cloud registration algorithms will be briefly introduced. Surface mesh registration maps multiple surface

meshes to the same coordinate system so that overlapping parts of the shape are aligned. First, the aligning of a pair of meshes starts with simplifying the surface representation by replacing triangular meshes with surface point samples, i.e. point clouds made by triangular mesh vertices (Figure 3.2(a)). Second, a point cloud P, representing the vertices of the first mesh, is called the *source* and a point cloud Q representing the vertices of the second mesh is called the *target*. Transformation f is sought in order to satisfy an alignment criterion, i.e. to minimize alignment error. Alignment error is expressed through correspondences as

$$Err = \frac{1}{|\Sigma|} \sum_{i=1}^{|\Sigma|} \|q_j - f(p_i)\|^2, \{p_i, q_j\} \in \Sigma \tag{3.1}$$

The algorithms for mesh and point cloud registration can generally be divided into groups with respect to various criteria. The whole registration task consists of three components: a transformation model, matching criteria and an optimization strategy. Classification of the algorithms is based on the realization of these three components.

Primary classification of the registration algorithms allows for a transformation to map one mesh or point cloud to the other. The transformation model is partially given by the ability of the model to perform *rigid* or *non-rigid* registration. The former represents a transformation that only consists of translation and rotation. Occasionally, scaling is also considered. Frequently, multiple calibrated scans of the same object are registered together so that the scaling factor can be omitted. The latter group of transformations allows deformations, i.e. elastic changes in shape or coherence in the mutual positions of potential rigid parts of one object.

Rigid transformation models are usually characterized by 6 degrees of freedom (DoF), for translation and rotation in a three-dimensional space. For non-rigid registration, there are several options for various DoF. If the number of DoF does not depend on the number of points in the point cloud, the transformation model is called *parametric*. Such a transformation depends on the geometric model of space deformation described by auxiliary primitives or parameters, e.g. control points, skeleton nodes, etc. On the contrary, if the number of DoF depends on the number of cloud points, it usually describes individual translation of each point in the cloud. Hence, it is called *non-parametric*, even though there could still be some parameters involved.

Matching criteria and constraints offer means of restricting the search space of all possible transformations to only those that make sense for a particular task. These criteria and constraints cover, for example:

- Landmark fitting — optimal transformation should obtain certain corresponding points (landmarks) as close as possible to each other.
- Closest point criterion - for all points in a source cloud, the transformation minimizes overall distance to the closest point (Besl and McKay,

1992).

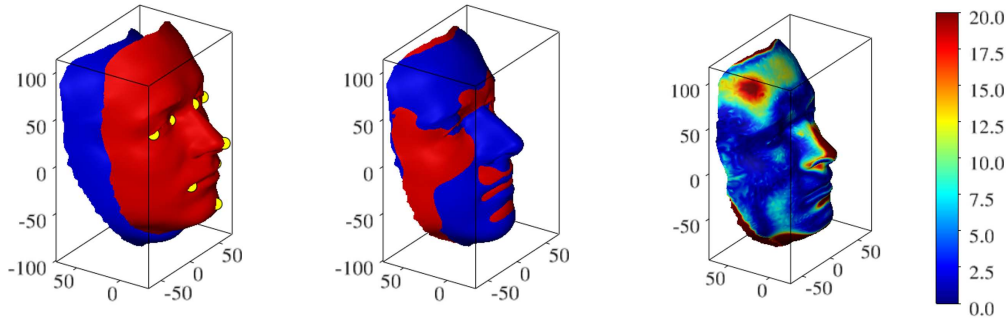
- Principal axes — the transformation aligns according to orthogonal axes, which are identified in the clouds and along which the clouds span with the highest variance.
- Saliency — the transformation fits points with matching properties based on local differential geometry (e.g. curvature)
- Features — fitting of locations with matching descriptors based on various criteria.
- Template fit — fitting to a predefined template shape or shape parts.
- Regularization — enforcing the smoothness of the transformation to avoid implausible space deformations.
- Diffeomorphism — transformation is sought in both directions, of which each is an inversion of the other. The concept of the source and target images becomes irrelevant.

Practical applications that employ registration methods do not always stand on fully automatic processing, especially in the case of non-rigid registration. Although non-rigid registration is a rapidly developing field, it still fails to deliver satisfactory results in all scenarios. By providing more information, however, the user can improve registration results, guide the algorithm and fix input data appropriately, all of which might help to successfully realize the final goal. Traditional shape analysis using landmarks depends on the very precise localization of anatomically important points. This precision is key to positive outcomes of the analysis. For example, it may be the reason why statistical tests of the difference between sample groups are not positive, since the magnitude of the noise caused by impressions surpasses the magnitude of the signal.

Despite being able to switch from landmark-based shape analysis methods to non-landmark methods applied on triangular meshes, landmarks should still be incorporated into the correspondence search as constraints. Similarly, other types of geometric primitives that define correspondences should be incorporated as constraints.

In order to compare and contrast methods with different levels of sophistication, as well as to cover methods that are used in actual research projects, several algorithms have been chosen for further presentation:

- Rigid registration based on surface landmarks selected by experts.
- Standard iterative closest point (ICP) rigid registration (Besl and McKay, 1992).
- Non-rigid registration based on TPS.



(a) (b) (c)

Figure 3.3: Dense correspondence using landmark-based registration: (a) input meshes with landmarks; (b) fitted meshes; (c) color-coded distances (in mm) of corresponding points with $Err = 9.92239$.

- Non-rigid parametric registration using free-form space warping.
- Non-rigid non-parametric registration using the coherent point drift algorithm (CPD) by Myronenko *et al.* (2006).

3.2.1 Landmark-based registration

Assume two landmark sets L_Q and L_P are placed on the surface of triangular meshes T_Q and T_P by an expert. These two landmark sets are aligned by GPA (see Section 2.2 [page 17]) and the same transformation is used to align the meshes' vertices of T_Q and T_P. Figure 3.3 shows the result of landmark-based rigid alignment using an example of facial meshes (see also Figure 2.2(b) for landmark configuration) as well as deviations of corresponding points found after alignment by the nearest neighbor principle, as described above.

3.2.2 Iterative closest point registration

Iterative closest point (ICP) is one of the most popular algorithms for mesh and point cloud registration. The original algorithm by Besl and McKay (1992) has been continuously improved and generalized for more and more applications, e.g. by Rusinkiewicz and Levoy (2001). For the sake of completeness, the original ICP will be explained here.

The input of ICP is a pair of point clouds, P and Q. First, centers of mass of P and Q are defined

$$\mu_p = \frac{1}{N_p}\sum_{i=1}^{N_p} p_i,\ \mu_q = \frac{1}{N_q}\sum_{i=1}^{N_q} q_i$$

The original ICP iteratively searches for the optimal transformation $t = (t_T, t_R)$, i.e. translation and rotation, which will minimize the objective function E_k in k-th iteration

$$E_k(t) = \frac{1}{N_p} \sum_{i=1}^{N_p} \|\hat{q}_i - R(t_R)p_i - t_T\|^2 \tag{3.2}$$

where t_T is a translation parameter vector and t_R is the quaternion describing the rotation. Vector $\hat{q}_i$ is a point of Q that corresponds to p_i based on the nearest neighbor principle for the current iteration k, i.e. after point p_i is transformed by the transformation found in the previous step $k-1$, starting with the identity. Rotation matrix $R(t_R)$, characterized by quaternion $t_R = (t_0, t_1, t_2, t_3)$, is defined

$$R(t) = \begin{bmatrix} t_0^2 + t_1^2 - t_2^2 - t_3^2 & 2(t_1t_2 - t_0t_3) & 2(t_1t_3 + t_0t_2) \\ 2(t_1t_2 + t_0t_3) & t_0^2 + t_2^2 - t_1^2 - t_3^2 & 2(t_2t_3 + t_0t_1) \\ 2(t_1t_3 - t_0t_2) & 2(t_1t_2 - t_0t_3) & t_0^2 + t_3^2 - t_1^2 - t_2^2 \end{bmatrix}$$

The optimal rotation is computed using a method by Horn (1987) from a 3×3 covariance matrix of P and Q

$$\Sigma_{pq} = \frac{1}{N_p} \sum_{i=1}^{N_p} (p_i - \mu_p)(\hat{q}_i - \mu_q)^T \tag{3.3}$$

and from the cyclic component of the anti-symmetric matrix $A = \Sigma_{pq} - \Sigma_{pq}^T$, $\Delta = [A_{2,3}, A_{3,1}, A_{1,2}]$. Σ_{pq} and Δ are used to form a $4x4$ matrix M

$$M(\Sigma_{pq}) = \begin{bmatrix} tr(\Sigma_{pq}) & \Delta^T \\ \Delta & \Sigma_{pq} + \Sigma_{pq}^T - tr(\Sigma_{pq})I_3 \end{bmatrix} \tag{3.4}$$

The unit eigenvector of $M(\Sigma_{pq})$ corresponding to the largest eigenvalue is a unit quaternion representing the optimal rotation, minimizing Equation 3.2 for the current set of corresponding points. The optimal translation vector is given by

$$t_T = \mu_q - R(t_R)\mu_p \tag{3.5}$$

A detailed derivation of this closed-form solution for optimal rotation between given pairs of points is described by Horn (1987). The algorithm described above is equivalent to Ordinary Procrustes Analysis (see Section 2.2 [page 17]) and can be used instead.

After the translation and rotation vector t is computed, it is used to transform P. Next, a new $\{\hat{q}_i | i = 1 \dots N_p\}$ is found and the process is repeated until changes in the mean-square matching error drop below a given threshold, i.e $Err_k - Err_{k+1} < \tau$.

The ICP procedure is summarized in Algorithm 1.

Algorithm 1: Iterative closest point

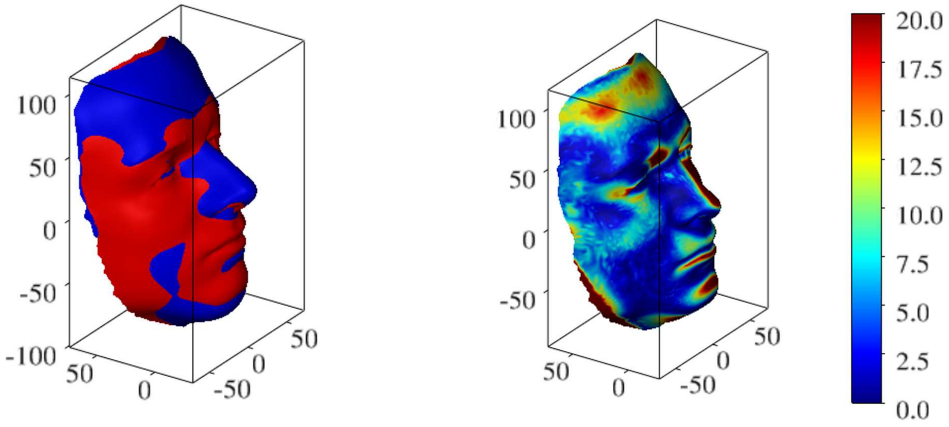
Input : target point cloud Q, source point cloud P
Output: aligned source P
begin
  $P_0 = P$
  for $k = 0, \ldots,$ *max_iterations* **do**
    For each $p_i \in P^k$ find closest point $\hat{q}_i \in Q$.
    Compute Σ_{pq} according to Equation 3.3.
    Compute rotation matrix $M(\Sigma_{pq})$ (Equation 3.4) and translation vector t_T (Equation 3.5)
    Compute $P_{k+1} = \{p_i^{k+1} | p_i^{k+1} = M(\Sigma_{pq}) p_i^k + t_T\}$
    if $Err(P_{k+1}, Q) - Err(P_k, Q) < \tau$ **then**
      break
    end
  end
end

The most time-consuming part (more then 95%) of the algorithm is computation of the closest points to the arbitrary point of Q. A typical approach to this problem is to employ space partitioning acceleration structures that are capable of rapidly executing the nearest neighbor operation. For example, a k-d tree has logarithmic nearest neighbor time complexity and is very effective for point cloud and vertex-to-vertex correspondence (Samet, 2005, Chapter 1.5). In the case of vertex-to-point correspondence in a triangular mesh, where the k-d tree is built on a triangle soup, it loses its efficiency due to the more complicated traversing of the tree. In these instances, an R-tree can be more efficient (Samet, 2005, Chapter 2.1.5).

The result of ICP applied to the reference face meshes is shown in Figure 3.4. Alignment is similar to landmark registration, but in this basic variant of ICP it is influenced by point distribution in the mesh or by missing parts of one of the meshes.

3.2.3 Registration by landmark-fitted TPS

Thin-plate spline interpolation was mentioned in Section 2.2 [page 19] as a non-rigid transformation model for analysis and visualization of the change in geometry between two shapes. TPS is fairly popular for non-rigid registration in 2D as well as in 3D. In this section, TPS will be utilized as a landmark-based non-rigid registration tool for mapping one triangular mesh onto another. Assume two landmark sets L_Q and L_P are placed on the surface of triangular meshes T_Q and T_P by an expert. These landmark sets will be used to construct the TPS transformation function $f_{TPS} : \mathbb{R}^3 \mapsto \mathbb{R}^3$ that maps L_P to L_Q and interpolates space between them, as described in Section 2.2 [page 19].



(a) (b)

Figure 3.4: Dense correspondence using ICP registration: (a) fitted data; (b) color-coded distances (in mm) of corresponding points with $Err = 8.55524$.

Function f_{TPS} can be used to transform vertices P of mesh T_P.

Figure 3.5 shows deviations of nearest neighbor corresponding points after TPS mesh alignment. The method causes large deviations at the margin of the meshes since they lie outside the convex hull of the landmarks, although the center of the face is relatively well aligned.

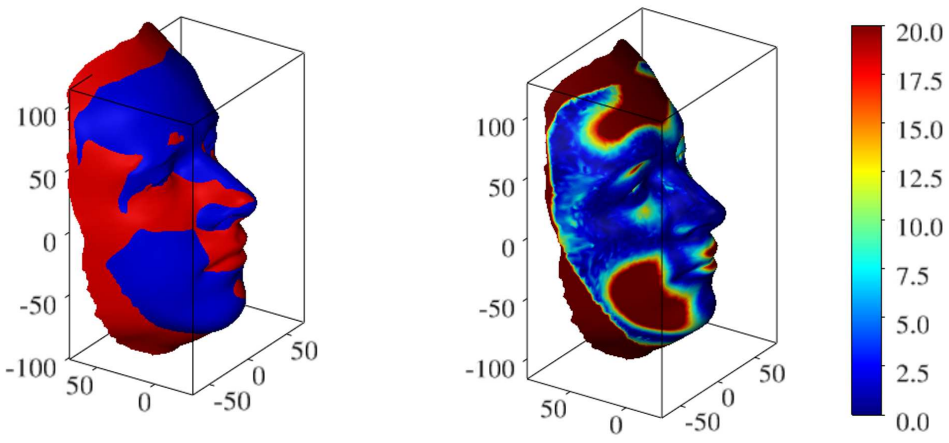
One disadvantage of using TPS arises in instances of very different shapes, where corresponding points can even cross. These cases cause singularities in TPS transformation, i.e. two close points near the singularity are transformed far from each other. Such behavior can be controlled to a degree by relaxing interpolation conditions and enforcing smoothness in reformulated *smoothing TPS* satisfying in 2D

$$E(f_{TPS}) = \sum_{i=1}^{K} \|y_i - f_{TPS}(x_i)\|^2 + \lambda \iint \left(\frac{\partial^2 f_{TPS}}{\partial x^2}\right)^2 + \left(\frac{\partial^2 f_{TPS}}{\partial x \partial y}\right)^2 + \left(\frac{\partial^2 f_{TPS}}{\partial y^2}\right)^2 dxdy$$

and where λ is the smoothness control parameter. This problem has a unique solution (Wahba, 1990).

3.2.4 Free-form space warping

Free-form space warping is a well-described parametric transformation model for medical image registration, introduced by Rueckert *et al.* (1998). It was



(a) (b)

Figure 3.5: Dense correspondence using TPS registration: (a) fitted data; (b) color-coded distances (in mm) of corresponding points with $Err = 49.3197$.

originally developed for registration of MRI images of deformable objects and inter-patient comparisons in 2D and in 3D images. However, in this subsection it will be used to register of point cloud data (Wenckebach *et al.*, 2005). It is based on the free-from deformation (FFD) function of d-th order in 3D

$$f_{FFD}(\vec{x}, c) = \sum_{i=l}^{l+d} \sum_{j=m}^{m+d} \sum_{k=n}^{n+d} c_{i,j,k} B^d_{i,j,k}(\vec{x})$$

Parameter c is a tensor containing the configuration of a rectangular grid of control points that span the space around the object, $c = \{c_{i,j,k} \in \mathbb{R}^3, \text{for } i \in \{0, \ldots, n_i\}, j \in \{0, \ldots, n_j\}, k \in \{0, \ldots, n_k\}\}$. The density of c defines the details of the deformation that the FFD can model. The other way around, the sparsity of c works as an intrinsic regularization. Moreover, FFD has natural subdivision properties that allow it to be adaptively refined. Function $B^d_{i,j,k}$ is non-zero only for control points in the subgrid of c; hence, there are only $(d+1)^3$ summands in the above equation. The subgrid depends on $\vec{x}$, $l = \lfloor x/(n_i - d) \rfloor$, $m = \lfloor y/(n_j - d) \rfloor$ and $n = \lfloor z/(n_k - d) \rfloor$. Weight function $B^d_{i,j,k}$ is composed of B-spline bases of d-th order over knots $\{t_i\}$ which are defined recursively

$$\begin{aligned} B^d_i(x) &= \frac{x - t_i}{t_{i+k-1}} B^{k-1}_i(x) + \frac{t_{i+k} - x}{t_{i+k} - t - i + 1} B^{k-1}_{i+1}(x) \\ B^1_i(x) &= \begin{cases} 1 & \text{if} \quad t_i \leq x < t_{i+1} \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

However, in practical applications, the particular base B for the cubic spline is defined by four third-order polynomials on four disjunctive intervals of the B-spline base, mapped to interval $[0, 1]$

$$\begin{aligned} B_0^3(x) &= (1-x)^3/6 \\ B_1^3(x) &= (3x^3 - 6x^2 + 4)/6 \\ B_2^3(x) &= (-3x^3 + 3x^2 + 3x + 1)/6 \\ B_3^3(x) &= x^3/6 \end{aligned} \tag{3.6}$$

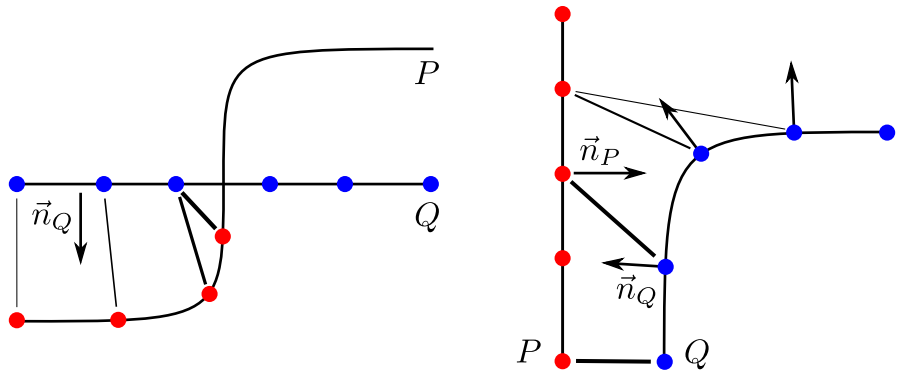
For $x' = x/(n_i - 3) - l$, $y' = y/(n_j - 3) - m$ and $z = z/(n_k - 3) - n$, the weight function is then defined as

$$B_{i,j,k}^3(\vec{x}) = B_{i-l}^3(x') * B_{j-m}^3(y') * B_{k-n}^3(z')$$

Non-rigid registration is defined as the minimization of a certain functional with respect to a transformation model. In the case of FFD, it is effectively reduced to a minimization of a cost function E with respect to control point grid c of f_{FFD} deformation. The cost function E consists of various terms that reflect the data correspondence or penalize the transformation's intrinsic properties, such as volume preservation (Rohlfing *et al.*, 2003) or minimal second derivatives. In the presented case, the functional E is defined as follows

$$\begin{aligned} E(c) &= E_d(c) + w_r E_r(c) \\ &= \frac{1}{N}\sum_{i=1}^{N_p} \|q_i - f_{FFD}(p_i, c)\|^2 + \\ &\quad w_r \frac{1}{M}\sum_{i=1}^{M} \left(\frac{\partial^2 f_{FFD}}{\partial x^2}\right)^2 + \left(\frac{\partial^2 f_{FFD}}{\partial y^2}\right)^2 + \left(\frac{\partial^2 f_{FFD}}{\partial z^2}\right)^2 + \\ &\quad 2\left[\left(\frac{\partial^2 f_{FFD}}{\partial x \partial y}\right)^2 + \left(\frac{\partial^2 f_{FFD}}{\partial y \partial z}\right)^2 + \left(\frac{\partial^2 f_{FFD}}{\partial z \partial x}\right)^2\right] \end{aligned} \tag{3.7}$$

The first sum reflects the total distance of corresponding vertices. Minimizing this term brings the corresponding vertices closer to each other. The second sum affects the second derivatives of the transformation. The transformation is stiffer during the optimization process, and thus more resistant to undesirable warping of the space and the possibility of falling into narrow local minima. Weight $w_r \geq 0$ allows the user to balance between these two effects. The main difference between the two sums is that the first sum counts samples of the registered point cloud, while the second sum counts samples everywhere in the range of the FFD transformation function. A simple approach for the second sum is to regularly sample the space transformed by the FFD function and then evaluate second derivatives in these samples. The number of space samples relative to the number of samples from the cloud affects the balance of the data fitting and the smoothing effect of the registration.



(a) (b)

Figure 3.6: Problems with point clouds registration with changing nearest neighbor correspondence: (a) normal direction weighting by target normal direction and target-to-source point direction consine; (b) normal direction weighting by cosine of source and target normal directions.

The disadvantage of nearest neighbor correspondence is that the corresponding pairs can change during the registration process due to the complexity of the point cloud shapes, non-uniform point distribution in the clouds and the general dissimilarity of the registered shapes. These factors can cause, for example, the surfaces (represented by the point clouds) to slide along each other introducing local contractions. Increasing the value of the regularization term weight w_r will help to avoid local contractions. However, such restrictions might not be advantageous everywhere. Therefore, it would be beneficial to exploit surface geometry features, such as normals, to weight the impact of particular correspondences on control point movement.

Integrating surface geometry requires additional geometrical information, i.e. a vertex normal or a vertex triangulation from which the vertex normal can be approximated. One possibility is to weight correspondence according to the angle between points and normals in target shape T_Q (see Figure 3.6(a)). The data term — i-th summand — is weighted by $w_{n,i}$

$$w_{n,i} = 1 - \frac{q_i - f_{FFD}(p_i, c)}{\|q_i - f_{FFD}(p_i, c)\|} \cdot n(q_i)$$

Another option is to use the normal of both the target and deformed source shape, as suggested by Wenckebach *et al.* (2005).

Normal direction weighting is integrated to the data term

$$E_d(c) = \sum_{i=1}^{N_p} (w_{n,i} w_n + (1 - w_n)) \|q_i - f_{FFD}(p_i, c)\|^2 \tag{3.8}$$

Weight w_n, ranging from 0 to 1, controls the overall impact of normal weighting

on the data term. If w_n is equal to 0, normal direction weighting is completely disabled.

Actual application usually employs various kinds of constraints or explicitly defined correspondences, which can be determined manually or by an automatic feature-matching algorithm. Either way, these correspondences can be naturally incorporated into the cost function above

$$E(c) \quad = \quad E_d(c) + w_r E_r(c) + w_c \frac{1}{C} \sum_{j=1}^{C} \|c_i^p - f_{FFD}(c_i^q)\|^2 \tag{3.9}$$

Weight w_c controls the effect of constraint correspondences from completely disabled ($w_c = 0$) to perfectly fitting $w_c \gg 0$. The optimal value for this particular study depends on the data.

Optimization strategies

The problem of cost function in Equation 3.9 minimization is well suited for various methods of local numerical optimization (Antoniou and Wu-Sheng, 2007, chapter 5). Due to higher order polynomials, partial derivatives and nearest neighbor distances, the optimization problem is highly non-linear. In general, local optimization starts at some approximation of the solution (initial solution) and continuously improves by iteratively exploring the adjacent problem space. The following list includes numerical methods for non-linear optimization which are often used in parametric image registration, as ordered by the amount of gradient information it depends on:

- Powell's method (PM) does not depend on partial derivatives at all. It aims to optimize along predefined directions d_i and their combinations as if one-dimensional problem $E(c + \gamma d_i)$ is being solved with respect to variable γ.

- The downhill simplex method (DS or Nelder-Mead method) does not depend on partial derivatives either. Cost function E is evaluated in all directions around the current position and the optimal direction towards lower values is approximated.

- The steepest descent method (SD) evaluates gradient ∇E and moves in its direction. There are various approaches for making decision on step size: either it is fixed, continuously decreasing, adaptively changing, gradient magnitude dependent or determined by a one-dimensional line search. This method has been used by Kybic and Unser (2003) for volume registration.

- The conjugate gradient method (CG) evaluates gradients of E in every step. Gradients are continuously accumulated and the direction in the current step is based on the gradient of E in the current step as well as on gradients in previous steps (conjugate directions).

- The Levenberg-Marquardt algorithm (LM) blends the SD and Newton's method using an approximation of the Hessian matrix inverse. The approximation works well only if E is quadratic. This condition is satisfied for general E when close to the minimum; it behaves like SD when far from the minimum. The balance between the two strategies is tuned by a damping parameter. Kabus *et al.* (2004) have shown that LM is superior to SD for volume data registration tasks.

- The Broyden-Fletcher-Goldfarb-Shanno algorithm (BFGS) evaluates gradient of E and continuously approximates its Hessian matrix (second derivatives) as well as its inverse, so that no repeated inverse matrix computation for high dimensional problems is needed, unlike for LM. Nevertheless, storage of full Hessian matrix approximation is required, which might be difficult in high dimensions, especially for applications with memory-limited environments (e.g. GPU implementations, mobile devices, etc.).

- The limited-memory BFGS variant (L-BFGS) stores only a few vectors for low-rank approximation of the Hessian matrix. Song *et al.* (2014) have successfully applied the L-BFGS approach to B-spline non-rigid multi-modal registration of MRI and CT images.

- Newton's method (NM) evaluates first and second derivatives of E. It advances in the direction of the gradient multiplied by the inverse of the Hessian matrix.

SD is the easiest to implement and works well. However, the convergence is rather slow if certain accuracy is required and the step parameter must be tuned accordingly. Memory requirements are linear with regard to the number of variables. However, L-BFGS is the state-of-the-art method for the task and also many existing implementations are available, e.g. by Okazaki (2010) and by Johnson (2014).

In order to speed up the convergence and improve robustness, a multi-scale approach is employed. The B-spline can be subdivided by increasing the number of control points while keeping the transformation identical. This property is used to let the algorithm converge on various levels of detail (starting with the most coarse), while fitting the point clouds with low-frequency, non-rigid space warping. After the convergence, the number of control points in one dimension is almost doubled, i.e. if n_k is the number of control points in one dimension on k-th level, $n_{k+1} = 2n_k - d$ for d equal to the B-spline order. In other words, the total number of control points is approximately 8 times larger in 3D and the size of the detail that space warping can fit is halved. For ordinary shapes, 3 to 5 levels of subdivision are usually used. For example, 5-th level of the control point grid subdivision, starting with 4^3 at the first level, accounts for the optimization problem with 0.13 milion variables. It might cause trouble for algorithms that must hold data with quadratic space

Algorithm 2: B-spline mesh registration algorithm

```
Input  : mesh P, Q
Output: mesh Pk
begin
    initialization of c
    for l := 1 to max_subdivision do
        optimize E with respect to c
        if l != max_subdivision-1 then
            subdivide c
        end
    end
end
```

complexity in a number of variables. The general structure of the multi-scale registration algorithm is summarized in Algorithm 2.

B-spline control grid subdivision essentially adds new knot points between each pair of knot points from the previous scale, i.e. if 4 control points in 2D defines the cubic B-spline with 2 knot points, by subdivision it becomes a B-spline with 5 control points and 3 knot points. See Figure 3.7(a).

A subdivision algorithm (Lane and Riesenfeld, 1980) for the B-spline grid of the B-spline of degree d starts by doubling the control points, so that $c'(2i, 2j, 2k) = c'(2i+1, 2j+1, 2k+1) = c(i, j, k)$. Actual control points representing space deformation identical to c are computed by averaging the adjacent control points d-th times. The algorithm is summarized in Algorithm 3. It can be seen that the computational complexity of the algorithm is $\mathcal{O}(8 \cdot d \cdot n_x n_y n_z)$

Nearest neighbor search

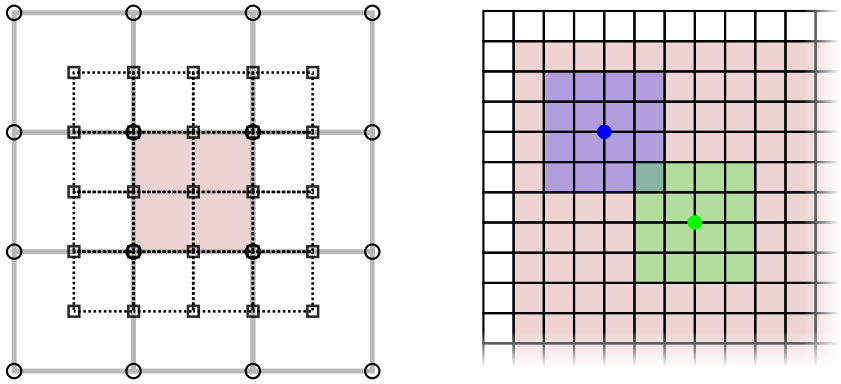
As mentioned above, most of the local optimization strategies evaluate partial derivatives of E with respect to variables represented by coordinates of points in control grid c

$$\begin{aligned} \frac{\partial E}{\partial c_{i,j,k,l}} &= \frac{\partial E_d}{\partial c_{i,j,k,l}} + w_r \frac{\partial E_r}{\partial c_{i,j,k,l}} + w_c \frac{\partial E_c}{\partial c_{i,j,k,l}} \\ (i,j,k,l) &\in \{1 \dots n_x\} \times \{1 \dots n_y\} \times \{1 \dots n_z\} \times \{x, y, z\} \end{aligned}$$

Since all terms of E integrate function f_{FFD}, the chain rule is applied while computing the derivatives.

$$\frac{\partial E_d}{\partial c_{i,j,k,l}} = \frac{\partial E_d}{\partial f_{FFD}} \frac{\partial f_{FFD}}{\partial c_{i,j,k,l}}$$

Due to the locality of f_{FFD}, its value is only influenced by a limited number of surrounding control points. Similarly, the particular control point participates in computating of the transformation of a limited area. The partial



(a) (b)

Figure 3.7: B-spline control points in 2D: (a) two levels of control points subdivision — first level (circles) defines the same deformation as the second level (squares) on the image space (red area); (b) the effect area (blue and green squares) of particular control points (blue and green discs) is limited in the image space (red area).

derivative of f_{FFD}, with respect to a coordinate of a particular control point ($\partial f_{FFD}/\partial c_{i,j,k,l}$), is non-zero on a limited area, on which transformation this control point participates (see Figure 3.7(b)). This behavior has a very advantageous speed-up factor.

The most time-consuming part of the computation of E_d is the evaluation of $(q_i - f_{FFD}(p_i, c))$ (see Equation 3.7), since it is not explicitly defined which q_i and $f_{FFD}(p_i, c)$ form a corresponding pair. Moreover, it differs in every iteration because grid c changes, which in turn affects $f_{FFD}(p_i, c)$. Therefore, the mutual distances to q_i change as well as the inferred correspondences. This is the same reason for the major computational burden faced by ICP registration. In order to deal with that nearest neighbor search, an acceleration structure must be implemented. Often, the data structure used for nearest neighbor queries in point clouds is the k-d tree, but octrees or R-trees can also be used (Elseberg *et al.*, 2012). However, a nearest neighbor query to k-d tree has computational complexity $\mathcal{O}(\log n)$, which is still too much for the B-spline surface registration of point clouds containing 10000+ points. The most successful ICP acceleration approaches are those that employ approximative nearest neighbor search structures. The first example of such a structure is locality-sensitive hashing, described by Datar *et al.* (2004), which improves query time complexity to $\mathcal{O}(n^{\epsilon})$ with ϵ given by the choice of hashing functions. A locality-sensitive hashing framework was previously used by Mayer and Greenspan (2007) for white matter trajectories registration in MRI. The second example is grid caching with a query complexity of $\mathcal{O}(1)$, proposed by Marden and Guivant (2012), which trades off some accuracy and higher memory demands. The last approach worth mentioning is a two-level data

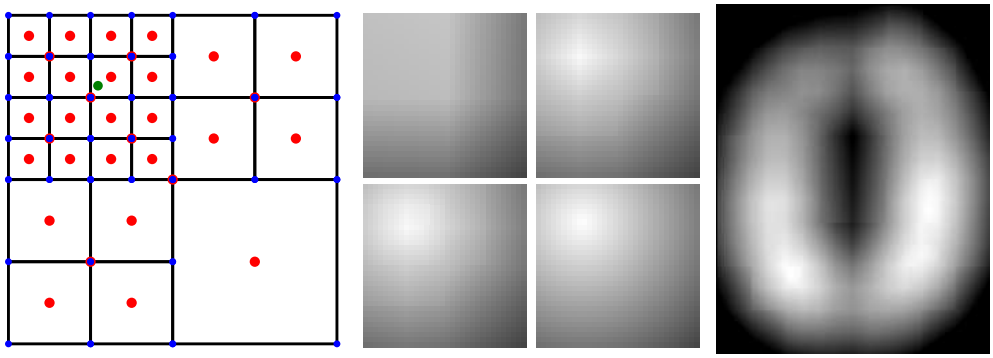
Algorithm 3: Degree d B-spline control points 3D grid subdivision

Input : grid c with $n_x \times n_y \times n_z$ control points
Output: grid c' with $(2n_x - d) \times (2n_y - d) \times (2n_z - d)$ control points
begin
  $C \longleftarrow$ grid with $2n_x \times 2n_y \times 2n_z$ control points
  for $(i, j, k) \in 2n_x \times 2n_y \times 2n_z$ **do**
    $C(i, j, k) := c(\lfloor i/2 \rfloor, \lfloor j/2 \rfloor, \lfloor k/2 \rfloor)$
  end
  $C' := C$
  for $l = 1$ **to** d **do**
    for $(i, j, k) \in (2n_x - l) \times (2n_y - l) \times (2n_z - l)$ **do**
      $C'(i, j, k) := \frac{1}{8} \sum_{i'=j'=k'=0}^{1,1,1} C(i + i', j + j', k + k')$
    end
    $C \longleftrightarrow C'$
  end
  $c' \longleftarrow$ grid with $(2n_x - d) \times (2n_y - d) \times (2n_z - d)$ control points
  for $(i, j, k) \in (2n_x - d) \times (2n_y - d) \times (2n_z - d)$ **do**
    $c'(i, j, k) := 2C(i, j, k)$
  end
end

structure which employs an octree on top of a Voronoi diagram giving the exact nearest neighbor in $\mathcal{O}(\log(\log(n))$ (Drost and Ilic, 2013).

The method of choice to speed up nearest neighbor searches in partial derivative computation is an approximative approach of caching the nearest neighbor in a fine grid structure. The actual result of the nearest neighbor query is linearly interpolated using values fetched from the *cache grid* in $\mathcal{O}(1)$ time. Refinement of the grid depends on the amount of memory available. The grid values can be either computed on the fly or pre-computed. Pre-computation is also rather computationally expensive, i.e. for a grid with 512^3 cells, 134 million nearest neighbor queries must be evaluated.

Pre-computation is still a possible approach since it can be performed approximately and faster for points that are further from the point cloud, without harming the accuracy of the registration. In other words, evaluation of nearest neighbor queries for points distant from the point cloud will not generate as many exact nearest neighbor searches as points queried closer to the point cloud. An octree data structure can be built upon points in the point cloud. Each cell of the octree evaluates the exact nearest neighbor search using the k-d tree eight times, once for every corner of the cell cube. The interior of the cell is approximated by a linear interpolation. The cells of the octree that are far from the cloud are larger, while the whole space they occupy generates only eight nearest neighbor searches. The sizes of the leaf cells in the octree are given by the *cell subdivision criterion*, which reflects the size of the cell in



(a) (b) (c)

Figure 3.8: B-spline registration acceleration structures in 2D: (a) Tree refinement on a point cloud containing a single point (green). Subdivision is performed if the distance of the cell center (red) from the nearest point (green) is lower then the tree cell diagonal length and if the tree cell diagonal length is higher than the cache grid cell diagonal length; (b) Corresponding distances to point cloud approximations interpolated from tree cell corners (blue) in 1 to 3 levels of subdivision compared to the true distance map; (c) An example of a slice through the cache grid for acceleration of face mesh registration.

the cache grid and the distance to the nearest neighbor point in the cloud, i.e. the leaf very close to the cloud is the size of the cache grid cell while the cell further apart is as large as its approximate distance to the cloud. For a clearer description, see Figure 3.8(a).

Sampling

Although the complexity of the nearest neighbor query is reduced to $\mathcal{O}(1)$ at the acceptable cost of sacrificed accuracy, the time needed to compute the partial derivatives is also determined by the number of points in the point cloud, i.e. the number of summands in evaluation of E_d. By employing various sampling strategies, the number of samples can be lowered without significant loss of registration accuracy and with a huge improvement in convergence speed. Various approaches to sampling are possible:

- Uniform random sampling — a certain fraction (e.g. 30%) of the original samples are chosen randomly and used in the registration. In the case of very dense meshes with few details, this strategy helps significantly.
- Poisson disk sampling — distributes points randomly while ensuring that no two samples are closer to each other than a certain minimum. This kind of handling is computationally expensive. A popular algorithm for approximation of Poisson disk sampling is Mitchell's best candidate algorithm (Mitchell, 1991).

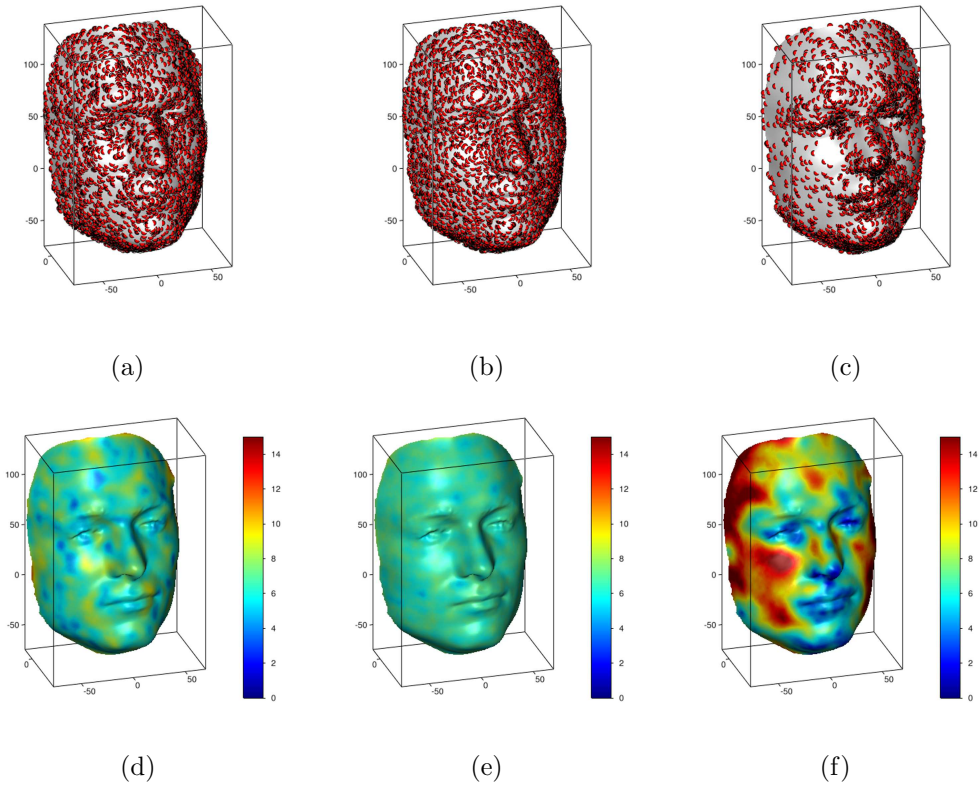
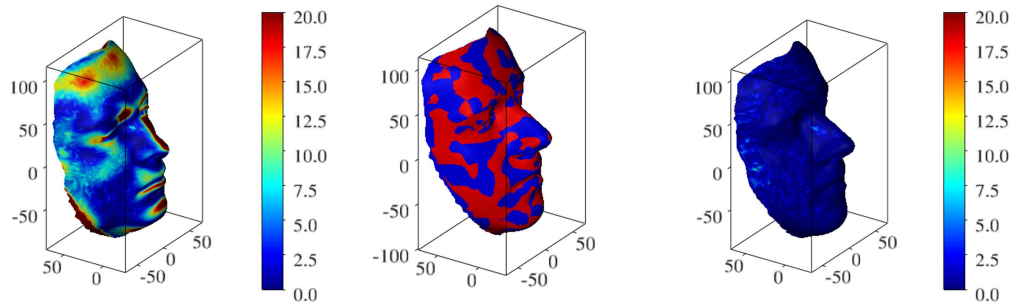


Figure 3.9: Mesh sampling approaches showing point distribution and color-coded point density of random uniform sampling (a) and (d); Poisson sampling (b) and (e); and importance sampling (c) and (f).

- Importance sampling — random sampling is weighted by data properties (local point density, geometrical features, etc.). This strategy might be very helpful in enabling decimated point clouds to describe objects that contain important features such as edges, ridges and borders that are difficult to use as constraints with explicit correspondence.
- Normal sampling — Rusinkiewicz and Levoy (2001) proposed uniform sampling with respect to the normal directions at sample points. This method is advantageous for large flat meshes with sparse features with significantly different surface orientation, e.g. edges, embossed structures, dents, etc. These surfaces have too many points in non-feature areas, but fewer points with different orientations on feature structures which, on the other hand, might help fitting considerably.

The first three mentioned sampling approaches are relevant to facial data registration and are used further in the experiments. See Figure 3.9 for a demonstration.

Various aspects of B-spline mesh registration of meshes were tested. Figure 3.10 shows results with reference face data that are comparable to the results



(a) (b) (c)

Figure 3.10: Dense correspondence using B-spline warping registration: (a) color-coded distances of corresponding points (in mm) of rigidly aligned meshes entering non-rigid registration; (b) a source mesh fitted to a target mesh; (c) color-coded distances (in mm) of corresponding points after non-rigid registration with $Err = 0.90659$.

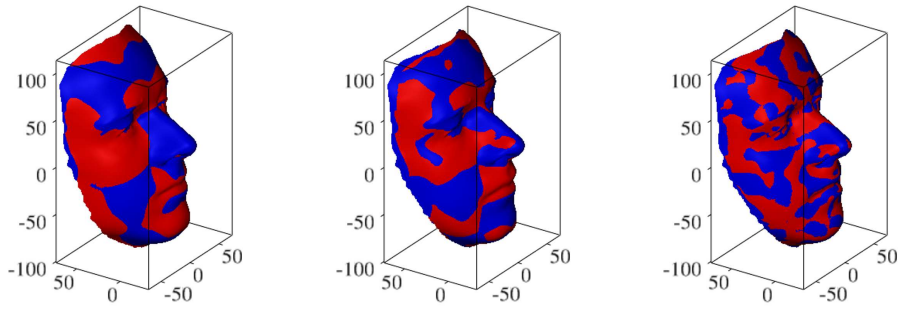
of the other algorithms.

Figure 3.11 shows the results of B-spline registration with various depths of B-spline control point grid subdivision. It shows how the density of the control points' grids influence local details of the deformations as well as resulting deviation criteria (Err). Another parameter that influences the precision of the fitting is resolution of the nearest neighbor approximation data structure — the cache grid, as shown in Table 3.1.

Sampling density has a significant impact on the speed of convergence as well as on the precision of the result. Fewer samples cause higher resulting deviations compared to results with all vertices included (see Table 3.2).

Table 3.3 shows how optimization strategy influences the speed of convergence and the deviation of the result. However, specific configuration of the optimization routine, especially its convergence criterion, has an important effect on convergence time, since the objective function tends to drop quickly at the beginning of the optimization and improve very slowly in later iterations. The convergence criterion of SD and L-BFGS, which stops optimization when improvement gain drops below a certain threshold is set identically to $\epsilon = 0.0001$. Note that in order to obtain comparable results, all Err evaluation is computed from all samples and not the sampled subset.

A comparison of various sampling strategies is shown in Table 3.4. The efficiency of the sampling approach is connected to the total number of samples the strategy utilizes. As shown, uniform sampling is outperformed by Poisson sampling, which is outperformed by importance sampling in equal sample sizes measured by the alignment error. But the convergence time is worsened in the exact same order. Either more samples could be used for an inferior sampling



(a) (b) (c)

Figure 3.11: B-spline registration using various levels of subdivision: (a) 1 level of subdivision resulted in surface deviation $Err = 1.734$; (b) 3 levels of subdivision, $Err = 1.218$; (c) 5 levels of subdivision, $Err = 0.907$.

# levels	**cache resolution**	**# control points**	**time (s)**	*Err*
1	256	64	54.85	1.9517
2	256	125	124.80	1.7767
3	256	343	249.81	1.3041
4	256	1331	362.35	0.9799
5	256	6859	619.65	0.8792
1	384	64	65.40	1.9044
2	384	125	105.82	1.7939
3	384	343	283.36	1.2174
4	384	1331	439.24	0.8859
5	384	6859	698.62	0.7733

Table 3.1: Evaluation of B-spline registration of faces with respect to subdivision levels of control point grids.

cache resolution	**samples (%)**	**time (s)**	*Err*
256	50	269.44	0.8933
256	25	147.82	0.8947
256	10	60.07	0.9210
256	5	34.01	0.9420
256	1	16.30	1.1263
384	50	303.94	0.7914
384	25	216.34	0.8011
384	10	128.65	0.8283
384	5	72.44	0.8632
384	1	53.30	1.0438

Table 3.2: Evaluation of uniform random sampling in B-spline registration of faces with 5 levels of control point grid subdivision.

samples (%)	**optimization**	**time (s)**	*Err*
100	SD	646.75	0.7730
10	SD	80.48	0.8073
100	L-BFGS	260.97	0.7571
10	L-BFGS	63.55	0.8038

Table 3.3: Evaluation of optimization methods in B-spline registration, with 5 levels of control point grid subdivision and 384^3 grid resolution.

samples (%)	**optimization**	**sampling method**	**time (s)**	*Err*
50	L-BFGS	uniform	199.13	0.8692
50	L-BFGS	Poisson	246.11	0.8650
50	L-BFGS	importance	240.91	0.8558
50	SD	uniform	385.85	0.8996
50	SD	Poisson	369.19	0.8910
50	SD	importance	328.57	0.8806
10	L-BFGS	uniform	51.45	0.9328
10	L-BFGS	Poisson	63.88	0.9138
10	L-BFGS	importance	79.92	0.9030
10	SD	uniform	88.01	0.9268
10	SD	Poisson	100.45	0.9286
10	SD	importance	80.20	0.9121
5	L-BFGS	uniform	32.29	0.9490
5	L-BFGS	Poisson	38.63	0.9729
5	L-BFGS	importance	52.76	0.9478
5	SD	uniform	43.04	0.9743
5	SD	Poisson	54.86	0.9496
5	SD	importance	64.11	0.9359

Table 3.4: Evaluation of sampling methods in B-spline registration, with 5 levels of control point grid subdivision and 256^3 grid resolution.

method, which would either match with the better method in the same time, or the better sampling method would converge earlier to get the same result faster.

3.2.5 Coherent point drift

Coherent point drift (CPD) is a novel rigid or non-rigid point cloud registration algorithm introduced by Myronenko *et al.* (2006) and extended by Myronenko and Song (2010). The previous algorithms depend on explicit point-to-point correspondence between source and target point clouds. CPD, on the other hand, treats these correspondences in a probabilistic way. The algorithm makes use of Gaussian mixture (GM) models to relate the source point cloud as a data points and target point cloud as a GM model centroids. The GM model is a probability density model, consisting of a certain number of components represented by Gaussian kernels, which allow fitting to an arbitrary number of data points. In other words, it enables two point clouds $X = \{x_n\}_{n=1}^{N}$ and $Y = \{y_m\}_{m=1}^{M}$ with a different number of points N and M to be registered, while expressing the correspondence between pairs of points in the clouds as a probability. GM probability density function (PDF) in CPD

is expressed as

$$p(x) = \sum_{m=1}^{M+1} P(m)p(x|m) \tag{3.10}$$

where $P(m) = 1/M$ for $m \in \{1, \ldots, M\}$ and correspondence probability of x given vertex m

$$p(x|m) = \frac{1}{(2\pi\rho^2)^{D/2}} \exp^{-\frac{|x-y_m|^2}{2\sigma^2}}$$

The last summand ($M+1$-th in Equation 3.10) is a uniform distribution added to the mixture model to account for noise and outliers, where $p(x, M+1) = 1/N$. Noise and outlier influence can be controlled by the weight w of the uniform distribution.

$$p(x) = w\frac{1}{N} + (1-w)\sum_{m=1}^{M} \frac{1}{M} p(x|m)$$

Increasing the value of w reduces the influence of fitting the GM model to the data by blending in a uniform distribution. This mitigates the effect of outliers and noise in the PDF.

The optimal GM model is found by looking for GM model centroid locations θ that maximize the value of the mixture model, which is equivalent to minimizing the negative log-likehood function E.

$$E(\theta, \sigma^2) = -\sum_{n=1}^{N} \log \sum_{m=1}^{M+1} P(m)p(x_n|m)$$

The expectation-maximization (EM) approach is then used to iteratively look for optimal transformation τ (rigid, affine or non-rigid) of GM model centroid locations θ, while progressively recomputing posterior probabilities $P^{prev}(m|x_n)$. In each iteration, it improves the objective function Q value which is partially evaluated with parameters found in the previous step.

$$\begin{aligned} Q(\theta, \sigma^2) &= -\sum_{n=1}^{N}\sum_{m=1}^{M+1} P^{prev}(m|x_n) \log(P^{next}(m)p^{next}(x_n|m)) \\ &= \frac{1}{2\sigma^2}\sum_{n=1}^{N}\sum_{m=1}^{M} P^{prev}(m|x_n)\|x_n - \tau(y_m, \theta)\|^2 + \frac{N_P D}{2}\log(\sigma^2) \end{aligned} \tag{3.11}$$

The probabilistic model is rewritten so that it uses Euclidean distance as an estimator for a negative logarithm of the correspondence probability. Posterior probability $P^{prev}(m|x_n)$ is calculated using Bayes' theorem, i.e. $P(m|x_n) = \frac{P(m)P(x_n|m)}{P(x_n)} = \frac{P(m)P(x_n|m)}{\sum P(k)P(x_n|k)}$. By substituting normal distribution, it results in

$$P^{prev}(m|x_n) = \frac{\exp\left(-\frac{1}{2}\|\frac{x_n - \tau(y_m, \theta^{prev})}{\sigma^{prev}}\|^2\right)}{\sum_{k=1}^{M}\exp\left(-\frac{1}{2}\|\frac{x_n - \tau(y_k, \theta^{prev})}{\sigma^{prev}}\|^2\right) + (2\pi\sigma^2)^{D/2}\frac{w}{1-w}\frac{M}{N}}$$

In non-rigid registration, the variant of CPD represents transformation τ by an arbitrary motion field v

$$\tau(Y, v) = Y + v(Y)$$

Further, non-rigid CPD is formulated as a variational problem by plugging τ into 3.11 and adding term $\|Pv\|^2$, which ensures the smoothness of the solution.

$$\begin{aligned} Q(\theta, \sigma^2) &= \frac{1}{2\sigma^2} \sum_{n=1}^{N} \sum_{m=1}^{M} P^{prev}(m|x_n) \|x_n - (y_m + v(y_m))\|^2 \\ &\quad + \frac{N_P D}{2} \log(\sigma^2) + \frac{\lambda}{2} \|Pv\|^2 \end{aligned} \tag{3.12}$$

The functional above is minimized by function v which satisfies the Euler-Lagrange equation with a form of radial basis function (see Myronenko and Song (2010) for details)

$$\begin{aligned} v(z) &= \frac{1}{2\sigma^2} \sum_{n,m=1}^{N,M} P^{prev}(m|x_n)(x_n - (y_m + v(y_m)))G(z, y_m) \\ &= \sum_{m=1}^{M} w_m G(z - y_m) \end{aligned} \tag{3.13}$$

where G is the Gaussian kernel and $w_m = \frac{1}{2\sigma^2} \sum_{n,m=1}^{N} P^{prev}(m|x_n)(x_n - (y_m + v(y_m)))$. The choice of Gaussian kernel G is connected to the regularization of the solution. It also makes the algorithm valid according to the Motion Coherence Theory by Yuille and Grzywacz (1989); hence, the name of the algorithm. Plugging equation 3.13 into 3.12 and taking derivatives with respect to $\mathbf{W}$, a linear system of equations is formed

$$\begin{aligned} \frac{\partial Q}{\partial \mathbf{W}} &= \frac{1}{\sigma^2} \mathbf{G}(d(\mathbf{P1}))((\mathbf{Y} + \mathbf{GW}) - \mathbf{PX}) + \lambda \mathbf{GW} = 0 \\ 0 &= d(\mathbf{P1})\mathbf{Y} + d(\mathbf{P1})\mathbf{GW} - \mathbf{PX} + \lambda\sigma^2 \mathbf{W} \\ \mathbf{PX} - d(\mathbf{P1})\mathbf{Y} &= (d(\mathbf{P1})\mathbf{G} + \lambda\sigma^2 \mathbf{I})\mathbf{W} \end{aligned} \tag{3.14}$$

where $d(\cdot)$ is a diagonal matrix made out of the argument, $\mathbf{1}$ is the column vector of 1, $\mathbf{W}$ is the $M \times D$ matrix of the coefficients and $\mathbf{G}$ is the $M \times M$ kernel matrix where $g_{ij} = \exp \frac{1}{2} \| \frac{y_i - y_j}{\beta} \|^2$. The new value of σ^2 is recomputed at the end of the maximization step of the EM algorithm by equating the derivative of Q with the new transformation τ of y_m to zero.

$$\begin{aligned} \sigma^2 &= \frac{1}{N_P D} \sum_{m,n=1}^{M,N} \|x_n - \tau(y, W)\|^2 \\ &= \frac{1}{N_P D} (tr(\mathbf{X}^T d(\mathbf{P}^T \mathbf{1})\mathbf{X}) - tr((\mathbf{PX})^T \mathbf{T}) + tr(\mathbf{T}^T d(\mathbf{P1})\mathbf{T})) \end{aligned} \tag{3.15}$$

Algorithm 4: Non-rigid coherent point drift registration

Input : target point cloud X, source point cloud Y
Output: aligned source T
begin
  initialize $W = 0$
  initialize $\sigma^2 = \frac{1}{DNM} \sum_{m,n=1}^{M,N} \|x_n - y_m\|^2$
  initialize $w \in (0, 1)$, $\beta > 0$, $\lambda > 0$
  construct G, $g_{ij} = \exp\left(-\frac{1}{2\beta^2}\|y_i - y_j\|^2\right)$
  while *not converged* **do**
    E-step: compute P^{prev}

$$p_{mn} = \frac{\exp\left(-\frac{1}{2}\left\|\frac{x_n - \tau(y_m, \theta^{prev})}{\sigma^{prev}}\right\|^2\right)}{\sum_{k=1}^{M} \exp\left(-\frac{1}{2}\left\|\frac{x_n - \tau(y_k, \theta^{prev})}{\sigma^{prev}}\right\|^2\right) + (2\pi\sigma^2)^{D/2} \frac{w}{1-w} \frac{M}{N}}$$

    M-step: solve linear system 3.14
    recompute σ^2 according to 3.15
  end
  Optimal transfomation τ is given by $\tau(Y, W) = T = Y + GW$
end

Complete non-rigid CPD is summarized in Algorithm 4. As shown, the computational complexity is derived from the complexity of the matrix operations, such as multiplication which is $\mathcal{O}(MN)$, and the solving of a linear system which is $\mathcal{O}(M^3)$.

The authors of the original algorithm (Myronenko and Song, 2010) lowered the time complexity in two ways. First, fast Gaussian transform (FGT) by Greengard and Strain (1991) is employed to compute matrix-vector multiplication of $\mathbf{P1}$, $\mathbf{P}^T\mathbf{1}$ and $\mathbf{PX}$ in $\mathcal{O}(M + N)$ time. This way, matrix $\mathbf{P}$ does not have to be evaluated and stored. Second, *low-rank matrix approximation* is used to speed up the solution of linear system 3.14. Equation 3.14 is rewritten in a form where the left-hand matrix is symmetric and positive definite

$$(\mathbf{G} + \lambda\sigma^2 d(\mathbf{P1})^{-1})\mathbf{W} \quad = \quad d(\mathbf{P1})^{-1}\mathbf{PX} - \mathbf{Y} \tag{3.16}$$

Matrix $\mathbf{G}$ can be replaced by matrix $\hat{\mathbf{G}}$ with rank K ($K < M$) and minimal difference from $\mathbf{G}$ expressed by the Frobenius norm, i.e. $\|\hat{\mathbf{G}} - \mathbf{G}\|_F^2 = min_{\bar{\mathbf{G}}}\{\|\bar{\mathbf{G}} - \mathbf{G}\|_F, rank(\bar{\mathbf{G}}) = K\}$. In other words, matrix $\hat{\mathbf{G}}$ is the low-rank approximation of matrix $\mathbf{G}$. The most popular way of finding low-rank approximation of a real symmetric matrix is to use eigenvector decomposition (Golub and Van Loan, 1996) (or Singular Value Decomposition with computational complexity $\mathcal{O}(n^3)$), which gives an optimal solution for k-rank approximation.

Matrix $\hat{\mathbf{G}}$ can be written as

$$\hat{\mathbf{G}} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^T$$

where $\mathbf{\Lambda}$ is the $K \times K$ diagonal matrix with K largest eigenvalues of $\mathbf{G}$ on its diagonal and $\mathbf{Q}$ is the $M \times K$ matrix of corresponding eigenvectors of $\mathbf{G}$. Using Woodbury's identity, the inversion of the left-hand term approximation of Equation 3.16 can be rewritten as

$$(\mathbf{Q\Lambda Q}^T + \lambda\sigma^2 d(\mathbf{P1})^{-1})^{-1} =$$
$$\frac{1}{\lambda\sigma^2} d(\mathbf{P1}) - \frac{1}{(\lambda\sigma^2)^2} d(\mathbf{P1})\mathbf{Q}(\mathbf{\Lambda}^{-1} + \frac{1}{\lambda\sigma^2}\mathbf{Q}^T d(\mathbf{P1})\mathbf{Q})^{-1}\mathbf{Q}^T d(\mathbf{P1})$$

Computing the matrix inversion this way leads to $\mathcal{O}(K^3)$. For $K \ll M$, time complexity is significantly improved. Authors have recommended using $K = M^{1/3}$, the largest eigenvalues of which make actual time complexity linear ($\mathcal{O}(M)$).

The only remaining difficult problem is the computation of eigenvectors of matrix $\mathbf{G}$. These vectors are computed once at the beginning and used repeatedly in EM algorithm iterations. The pre-computation can also be performed very effectively. Authors have suggested using the deflation technique, which involves matrix-vector products that can be accelerated via FGT. In general, there are many numerical algorithms capable of approximating eigenvectors of large matrices differing in numerical stability, computational complexity and precision. Currently, the most successful algorithms for eigenvector decomposition are based on a projection onto Krylov subspaces. One of these algorithms is Arnoldi's method (Saad, 1992), an iterative numerical algorithm for general matrices, which gives the first few eigenvalues after a relatively small number of iterations. Practical implementations (e.g. Matlab's eigs(), ARPACK) use implicitly restarted Arnoldi method (IRAM) which effectively handle high numbers of large eigenvectors. The computational complexity of Arnoldi's method is ($\mathcal{O}(M^2N)$).

One class of alternative method for low-rank matrix approximation, also known as randomized low-rank decomposition, attempts to build the approximation using columns sampled from the matrix itself and use them as eigenvectors instead of looking for actual eigenvectors (Deshpande and Vempala, 2006). The error of the approximation is only $(1 + \epsilon)$ times the error found by an algorithm that looks for true eigenvectors (i.e. IRAM). The algorithm iteratively adds columns to sample S in pace derived from the approximation error. In each iteration, the columns are chosen randomly based on a probability that is proportional to their size, relative to the norm of continuously reduced matrix $E_j = A - \pi_{span(S)}(A)$

$$P_i^{(j-i)} >= c\frac{\|E_{j-1}^{(i)}\|^2}{\|E_{j-1}\|_F^2}$$

Computational complexity of the described randomized k-rank decomposition is $\mathcal{O}(mk/\epsilon)$, where m is the number of non-zero elements of the matrix.

Another approach to approximate eigenvectors when constructing low-rank matrix approximation, is derived from the Nyström method. The Nyström

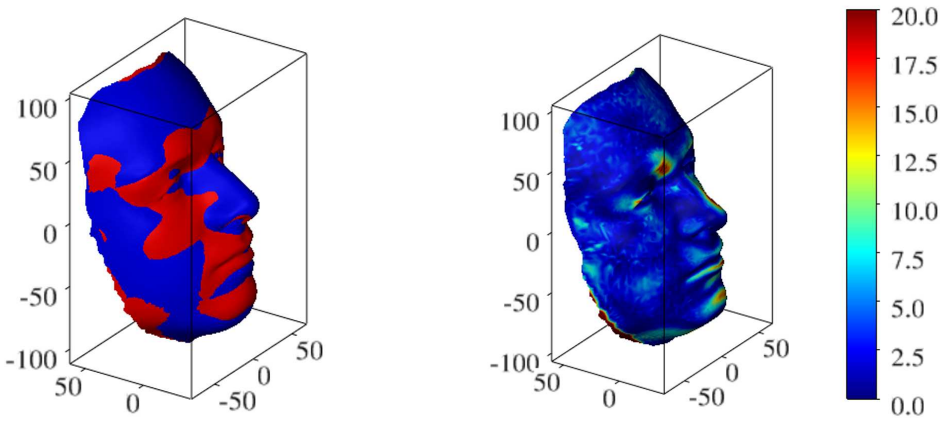
method	**dataset**	**init (s)**	**total (s)**	*Err*	**speedup**
IRAM	face	131	245	0.0144	1.00x
Nyström	face	41	196	0.0529	1.25x
k-means	face	14	119	0.0841	2.05x
IRAM	dragon	2204	2504	0.0014	1.00x
Nyström	dragon	1395	1695	0.0014	1.48x
k-means	dragon	82	394	0.0051	6.36x

Table 3.5: Evaluation of CPD speed-up methods: eigen decomposition using Arnoldi, Clustered Nyström and point cloud clustering. The face dataset consists of 13400 vertices; Stanford Dragon has 50016 vertices in the experiment. Data were taken from Dupej *et al.* (2014a).

method is a numerical method for solving integral equations by replacing them with a finite sum, which makes an integral equation into a system of linear equations. Various quadrature rules (interval approximations) for the Nyström method are used, including stochastic approximations and their error estimates (Kumar *et al.*, 2012). Williams and Seeger (2001) have used the Nyström method to approximate eigen decomposition simply by randomly choosing a $k \ll n$ number of columns of the full matrix K and using them as an input for the eigenproblem of size k. The solution of the k-sized problem is used as an approximation of the solution of the full-scale eigenproblem on matrix K.

Quantization of the full matrix column can be used to produce a more systematic choice of columns for eigenvector and eigenvalue approximation using the Nyström method, as suggested by Zhang *et al.* (2008). Matrix columns can be seen as points in a high-dimensional space. For better eigenvector approximation, it makes sense to choose points that are far apart from each other than points that are members of the same cluster, since they do not significantly contribute to the variation of the approximation. Error bounds of Nyström sampling can be estimated as quantization errors by choosing a particular representative column for its neighbors in a high-dimensional space. Authors have proposed the k-means clustering algorithm as an optimal sampling method with minimal quantization error for the desired k. This method has been further referenced as Clustered Nyström CPD.

The columns of the large matrix in the CPD algorithm correspond to points of the source point cloud in the three-dimensional space. If the points in the cloud are relatively far from each other, the norm of the difference between the corresponding columns in the matrix will also be large. Therefore, k-means clustering of points in the source point cloud provides meaningful sampling of the corresponding columns of the matrix using the Nyström method. Because k-means clustering is performed on data of much lower dimensionality than the dimensionality of the columns, it significantly speeds up the CPD algorithm



(a) (b)

Figure 3.12: Dense correspondence using coherent point drift registration: (a) fitted data; (b) color-coded distances (in mm) of corresponding points $Err = 3.9312$ ($\beta = 5.0, \lambda = 0.01, w = 0.1$).

while keeping the accuracy in acceptable bounds, compared to a method that uses exact low-rank approximation (Dupej *et al.*, 2014a).

As shown in Table 3.5, the authors of this idea also evaluated improved CPD registration on facial scan data. The measurement is divided into two phases of the algorithm. The initialization phase consists of pre-processing at the point at which eigen decomposition of the G matrix is computed, which is the actual bottleneck. This is followed by the registration phase where the algorithm iteratively drifts the points to the optimal result. The results show that speed-up scales quickly with the input data size; on the other hand, the difference in the error measure scales much faster to the algorithm's disadvantage. However, the time complexity of the improved algorithm — low-dimensional clustered Nyström CPD — is $\mathcal{O}(nk^2)$. Furthermore, the parts of the algorithm are implemented in CUDA and run on GPU.

Figure 3.12 shows how CPD handles registration of reference face data compared with other methods. It definitely manages to decrease differences in areas such as the cheeks and forehead, but the region around the lips is not fitted very well.

The disadvantage of CPD is that in the current state it does not allow explicit correspondences to be incorporated, i.e. landmarks set by a user to guide the registration. Nevertheless, the CPD algorithm, as a representative of non-rigid registration, has introduced the non-parametric transformation model as well as fuzzy-like correspondence into the convergence criterion.

3.3 Statistical models of mesh shapes

A statistical shape model is built from a sample that is large enough and appropriately distributed to cover the variability of a population. A procedure analogous to the creation of the point distribution model, previously mentioned in Section 2.2 [page 17], can be applied to mesh geometry if two conditions are fulfilled:

- there is an equal number of vertices
- there is group-wise point-to-point correspondence

A very basic algorithm for mesh modeling is called the dense correspondence model (DCM), published by Hutton *et al.* (2001). It is based on the user placing landmarks onto the surface of every model in the sample, which can result in very tedious and error-prone work. The algorithm itself makes use of GPA and TPS (Section 2.2 [page 17]) for correspondence construction. This results in the geometrical representation of each specimen by a mesh with the same number of vertices and those of identical topology, i.e. vertex-to-vertex correspondence. The procedure is summarized in Algorithm 5.

It is also important to note that TPS transformation does not have an implicitly defined inversion. Therefore, f_{TPS}^{-1} is computed numerically using iterative gradient-based optimization or the Gauss-Newton method by minimizing a squared distance $F(\vec{x}) = \|f_{TPS}(\vec{x}) - \vec{y}\|^2$

$$f_{TPS}^{-1}(\vec{y}) = \arg\min_{\vec{x}} \|f_{TPS}(\vec{x}) - \vec{y}\|^2$$

First derivatives of the squared distance expression should be equal to zero

$$\frac{\partial F}{\partial \vec{x}} = 2(f_{TPS}(\vec{x}) - \vec{y})^T J(f_{TPS})(\vec{x}) = 0 \tag{3.17}$$

where $J(f_{TPS})$ is the Jacobian matrix with partial derivatives of f_{TPS}, i.e. $J(f_{TPS})_{i,j} = \partial f_{TPS}(\vec{x})_i / \partial x_j$, which is specifically defined as

$$\begin{aligned} \frac{\partial f_{TPS,x}}{\partial x}(\vec{x}) &= \sum_{i=1}^{m} w_i^x \frac{\partial \varphi}{\partial r} \frac{\partial r}{\partial x}(c_i, \vec{x}) \\ r(c, \vec{x}) &= \|c - \vec{x}\|, \varphi(r) = r^3 \end{aligned}$$

The numerical solution of 3.17 is obtained by iteratively updating $\vec{x_i}$, initialized close to the expected solution (e.g. $\vec{x_0} = \vec{y}$). In other words, $x_{i+1} = x_i - \delta(\partial \|f_{TPS}(\vec{x_i}) - \vec{y}\|^2 / \partial \vec{x})$ until it converges for some small $\delta > 0$. Initialization is usually critical for local numerical optimization. Nevertheless, initializing the iteration in $\vec{y}$ is a good choice in this case because the landmarks defining TPS are already aligned by Procrustes analysis and the affine part of the TPS is close to the identity. Therefore, only the deformation part

Algorithm 5: Dense correspondence model construction procedure

Input : Input meshes $M = \{M_i\}$ and input landmarks $L = \{l_i\}$ for $i \in \{1, \ldots, n\}$

Output: Average mesh $\bar{M}$, $C = \{C_i\}$ meshes with identical topology and shape of a selected base mesh

begin

- Align L using GPA for each l_i producing matrix A_i and $\hat{l}_i$
- **for** $i \leftarrow 1$ **to** n **do**
  - $\hat{M}_i \leftarrow$ mesh with vertices of M_i transformed by matrix A_i
- **end**
- $B \leftarrow$ selected base mesh, e.g. $B = M_0$
- $l \leftarrow$ landmarks corresponding to B
- **for** $i \leftarrow 1$ **to** n **do**
  - $\tilde{M} \leftarrow$ mesh with vertices of $\hat{M}_i$ transformed by using $f_{TPS}^{\hat{l}_i \to l}$
  - $C_i \leftarrow$ copy of mesh B
  - **for** $j \leftarrow 1$ **to** $|C_i|$ **do**
    - $\vec{v} \leftarrow$ nearest point to j-th vertex of C_i on the surface of $\tilde{M}$
    - j-th vertex of $C_i \leftarrow (f_{TPS}^{\hat{l}_i \to l})^{-1}(\vec{v})$
  - **end**
  - **if** *we study shape* **then**
    - scale C_i by factor inverse to centroid size of l_i
  - **end**
- **end**
- $\bar{M} \leftarrow$ copy mesh B
- **for** $j \leftarrow 1$ **to** $|\bar{M}|$ **do**
  - j-th vertex of $\bar{M} \leftarrow \frac{1}{n} \sum_{i=1}^{n} j$-th vertex of C_i
- **end**

end

of the TPS is inverted, which means that the input coordinates are not as far from the output because the deformation is usually not so prominent.

For faster convergence (i.e. in a fewer number of iterations), a second-order method, e.g. the multidimensional Newton's method, can be used. However, it requires a computation of second derivatives and an inversion of the Hessian matrix, which might be slightly more time-consuming and less stable.

$$\vec{x}_{i+1} = \vec{x}_i - \delta H_F^{-1}(\vec{x}_i) \nabla_F(\vec{x}_i)$$

where H_F is the Hessian matrix of F, i.e. $H_F^{i,j}(\vec{x}) = \partial^2 F(\vec{x}_i) / \partial x_i \partial x_j$

$$\frac{\partial^2 F}{\partial x \partial y} = 2 \left(\frac{\partial f_{TPS}}{\partial x} \frac{\partial f_{TPS}}{\partial y} + (f_{TPS}(\vec{x}) - \vec{y})^T \frac{\partial^2 f_{TPS}}{\partial x \partial y} \right)$$

The Hessian of the TPS, $\partial^2 f_{TPSTPS}/\partial x_i \partial x_j$ is actually a tensor of order 3.

$$\frac{\partial^2 f_{TPS}}{\partial x \partial y} = \sum_{i=1}^{m} w_i^x \frac{\partial \varphi}{\partial r} \frac{\partial r}{\partial x}(c_i, \vec{x})$$

Unfortunately, due to the computational burden of Newton's method mentioned above, practical experiments have demonstrated that Newton's method takes twice as much computational time, compared to the first-order steepest descent method.

Alternative to a numerical TPS inversion approximation is to express the position of $\vec{y}$ on a triangular surface of $\hat{M}_i$ by barycentric coordinates of the point in a particular triangle. Since inverse transformation of the triangle vertices is known, the barycentric coordinates can be used with these known vertices. This approach works well in cases where meshes consisting of small triangles, which are not significantly non-linearly deformed.

The correspondence construction is explained visually in Figure 3.13 using a two-dimensional illustrative example.

Once the correspondences are found, PCA analysis can be executed as if the corresponding vertices were considered corresponding landmarks (Section 2.2 [page 17]). This is the basic principle of shape analysis applied on a sample of n meshes. The difference is that in the case of landmark-based analysis the number of landmarks is significantly lower than the number of vertices. Moreover, the number of variables p defined by $p/3$ vertices' coordinates is usually substantially higher than the number of observations in the sample. This imposes a fundamental limitation for PCA. A direct algorithm for PCA computation based on eigen decomposition assumes $p < n$; otherwise, it is not possible to compute the full rank covariance matrix, which is the reason many eigenvalues are zero. In addition, the computational cost is scaled with the dimension of the data, i.e. computational complexity is $\mathcal{O}(p^3)$.

PCA on mesh data is performed similarly to PCA on landmarks, i.e. the meshes' vertices are packed into one-dimensional vectors in the same order with respect to the found correspondences. The average of these vectors is an average mesh or a mean shape. Subtracting the mean shape from each specimen vector will center the dataset to origin.

In order to compute PCA on data with a high number of variables per observation ($p > n$), an algorithm for high-dimensional PCA (HDPCA) must be employed (Bishop, 2006, page 569). Input data are centered and represented by matrix X with $n \times p$ dimensions. First, $n \times n$ matrix $\tilde{S} = \frac{1}{n} X X^T$ is evaluated instead of computing the covariance matrix of input data X which would be $p \times p$. For eigenvectors u_i and coresponding eigenvalues λ_i of $\tilde{S}$, the following equation holds

$$\frac{1}{n} X X^T u_i = \lambda_i u_i$$

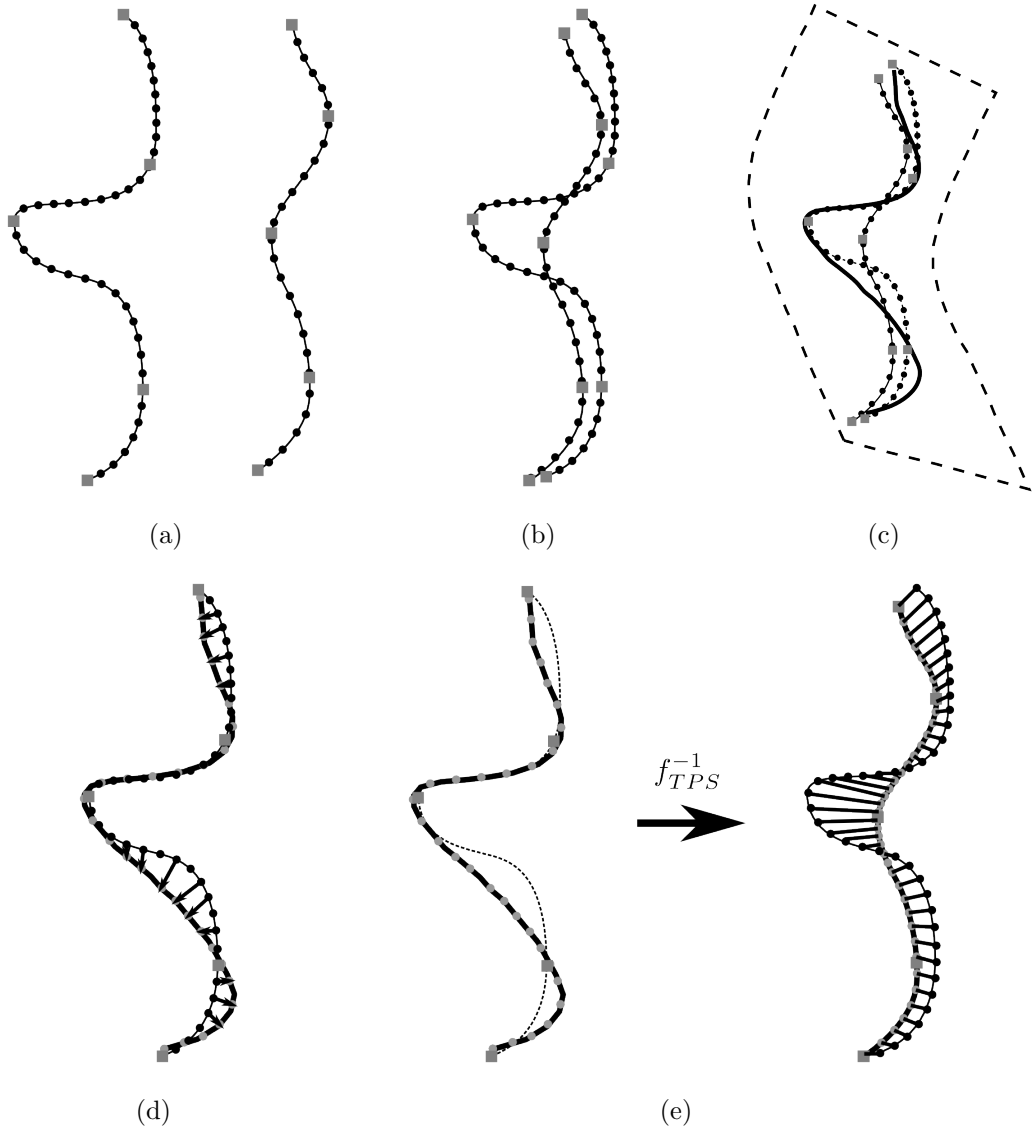


Figure 3.13: An explanation of a dense correspondence algorithm in 2D: (a) a pair of input polylines (the equivalent of a triangular mesh, where black dots represent vertices) with user-placed landmarks (gray squares); (b) rigid alignment using GPA; (c) TPS deformation; (d) nearest neighbor search; (e) inverse transformation.

By pre-multiplying by X^T

$$\frac{1}{n}X^T X(X^T u_i) = \lambda_i(X^T u_i)$$

The expression $X^T u_i = v_i$ is, by definition, the eigenvector with a corresponding eigenvalue λ_i of $S = \frac{1}{n}X^T X$, which is a covariance matrix of the original dataset. In this way, n eigenvectors v_i and eigenvalues λ_i of S can be obtained, while the rest of the eigenvalues are zero. The computational complexity is $\mathcal{O}(n^3)$. Eigenvectors, in descending order of the corresponding eigenvalues, represent a basis of the shape space where first dimensions account for the most variable components of the shape. PCA scores $\mathbf{s}_i$, i.e. coordinates of the inputted sample observations in this space can be computed.

$$\mathbf{s}_i = V^T x_i$$

where V is $p \times n$ matrix with unit eigenvectors v_i as columns, x_i is the p long vector and $x_{i,3j}, \ldots, x_{i,3j+2}$ = the coordinate deviation of j-th vertex of C_i from $\bar{M}$. The particular specimen i can then be computed

$$C_i = \bar{M} + V\mathbf{s}_i \tag{3.18}$$

Model properties

A certain degree of quality in any statistical model (Equation 3.18) is reflected in its ability to generate a given specimen in the population with the least error possible. This ability is connected with the complexity of the shape, the size of the training sample used for the creation of the model as well as the distribution of the sample in the population. All these factors are reflected in the properties of generalization and the specificity of the model (Davies *et al.*, 2008, chapter 4.2.2). Specificity indicates how well the distribution and density of the training sample correspond to the population, such that it is measured as the average shortest distance between randomly generated specimens to the nearest member of the training sample.

A generalization property, on the other hand, measures how well the model is able to represent unseen specimens. This ability is measured by the leave-one-out approach so that one specimen is removed from the training set, after which the model is created and used to reconstruct the removed specimen. The error between the reconstruction and the actual specimen is then measured. The whole process is repeated for all specimens in the training set.

$$G(M) = \frac{1}{|M|}\sum_{i=0}^{|M|} \|M_i - \tilde{C}_i\|^2 \tag{3.19}$$

The measure depends on the complexity of the shape and whether the size of the training sample covers all of the shape variations, even when one of

the specimens is removed. The quality of the model is also influenced by the quality of found correspondences. Better correspondences are able to reflect a more complex shape space, which means that more specimens are required to represent it properly.

3.4 Mesh shape analysis

The major goal of many life science disciplines is to describe the world around us. In the case of physical shapes, the ability to fulfill the task is very limited when using simple tools because the shapes of natural objects are very complex. Moreover, presenting these descriptions is also very difficult when only using numbers and tables. As already shown, geometric morphometry introduced ways of realizing these tasks on shapes represented by landmarks or similarly simple primitives (see Chapter 2).

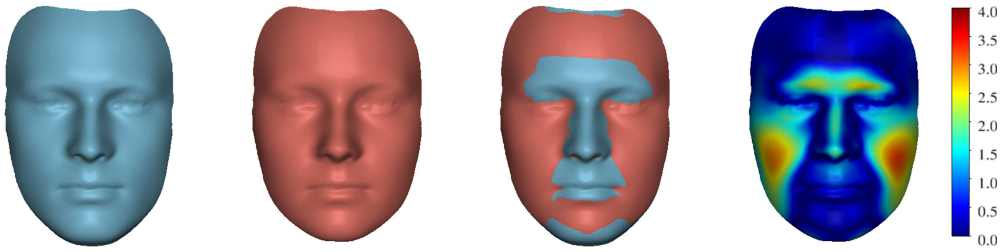
Using a statistical model of correspondences, created using non-rigid registration introduced above, the variability of samples of various three-dimensional triangular surfaces can now be studied in a geometrics morphometry fashion.

3.4.1 Mean shape and shape variability

The first goal is to study mean shape, referenced in the previous section where it was needed to perform high-dimensional PCA. As well as the mean shape of the whole sample, one can also switch focus to the mean shape of a subsample. For example, in many research tasks, the point of general interest is the difference between the average shape of the male and female variants of the object.

Comparison of the shapes of the male and female human face has been examined by Velemínská *et al.* (2012). In the study, a model was created with a dense correspondence algorithm using TPS for non-rigid registration (see Section 3.3), which is now considered inferior to B-spline-based non-rigid registration. The reason for this is addressed further. There are many traits that are typical for each sex and which are clearly noticeable to the experienced observer. Others, however, may not be so obvious. The advantage of a computational method is to help the observer to notice these differences and describe them impartially.

Figures 3.14(a) and 3.14(b) show average male and female faces where the extent of how diffcult it can be to spot, isolate and describe these differences is evident. Average shapes, on the other hand, have a disadvantage in that they can be very close to each other in the shape space. It is generally known that males have a protruding superorbital ridge, a deeper position of the eyes in relation to the facial plane, a larger and wider straight nose and nostrils, and larger lower face. These features can be observed if mean shapes are aligned and overlayed (see Figure 3.14(c)).



(a) (b) (c) (d)

Figure 3.14: Comparison of mean shapes of male and female faces. Subtle differences can be seen while meshes are overlayed and distances between corresponding points are measured: (a) mean shape of the male face; (b) mean shape of the female face; (c) overlayed mean shapes; (d) color-coded distance map between mean face meshes of males and females.

In addition to computing the group averages, the model can be used to extrapolate the extreme cases of hyper-masculine or hyper-feminine faces by constructing a shape corresponding to the parameters in the shape space lying on the line that connects male and female averages $\mathbf{s}_1$ and $\mathbf{s}_2$ in the space far away from their common mean $\bar{\mathbf{s}}$.

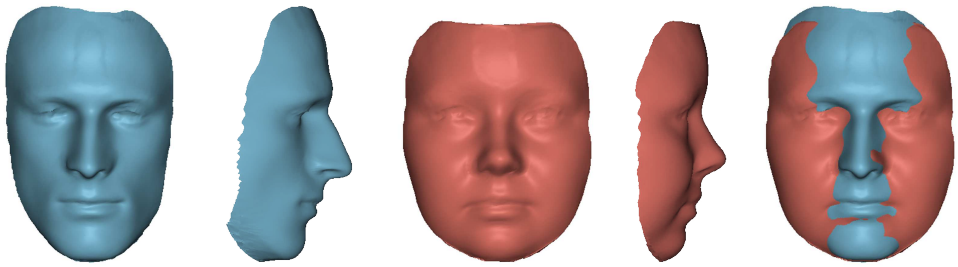
$$\mathbf{s}_e = \bar{\mathbf{s}} + \alpha(\mathbf{s}_1 - \mathbf{s}_2)$$

This solution will work if the groups have very similar variances. However, in the case of different variances and different orientations of the subgroups in the shape space, a simple line representing extrapolated shapes might miss a large part of one group or the other, i.e. the extrapolated shape is not very close to extreme cases in the subsample. In that case, the direction should be weighted by inversion of the covariance matrix, Σ^{-1}. The covariance matrix of the sample in the shape space obtained using PCA is a diagonal matrix.

$$\mathbf{s}_e = \bar{\mathbf{s}} + \alpha\Sigma^{-1}(\mathbf{s}_1 - \mathbf{s}_2)$$

Such a weighted direction generally does not progress via group means. More likely, it follows the shape of the sample in the shape space. It is worth mentioning that the weighted direction is equal to the normal of the groups separating plane, which comes out of linear discriminant analysis (Section 2.2 [page 19]). These artificial shapes are often used to visualize actual differences in cases both where actual differences are subtle and where comparing exact averages is not illustrative.

Moreover, the information-mining property of PCA can be exploited to extrapolate the shape, which represents group-specific traits beyond the mean, so that the only part of the shape's variability that is significant for shape description is used. In order to do this, only statistically significant components



(a) (b) (c)

Figure 3.15: Extrapolated shapes representing trends between mean shapes using only eight statistically significant principal components according to the broken-stick criterion: (a) extrapolated male; (b) extrapolated female frontal and lateral view; (c) overlayed frontal view with ever more prominent sexually dimorphic traits.

are used for extrapolation. The number of statistically significant components is determined by using the broken-stick criterion (Peres-Neto *et al.*, 2005). Figure 3.15, in comparison with Figure 3.14, shows the difference in a situation where the shapes are extrapolated using all information and where shapes are extrapolated using eight statistically significant components only.

In the case of more complex categorization or continuous categorical parameters, the least square-fitting of the various types of curves can be applied to the data points, which pass through mean configurations with respect to these parameters.

3.4.2 Form and allometry

Up to this point, only shape differences with sizes removed by registration have been considered. Instead of modeling differences between shapes of normalized size, triangular meshes can preserve their original size; hence, subsequent analysis becomes an analysis of form. Analysis of form and shape can differ significantly (see example in Figure 2.3(c) and 2.3(d)). If this is the case, the first component in PCA of form will reflect the size of the object, i.e. individual specimens will be sorted along the PC1 axis with respect to their actual size. Nevertheless, the shapes and sizes of natural objects are connected by allometry (for more details, see Case study 5 [page 29]). Therefore, certain shape characteristics connected with size are preserved even though the size is removed.

The relation of size and shape can be studied by constructing a linear model of dependent shape variables $\mathbf{s}_i$ and independent size variable s. Size can be computed using arbitrary methods depending on the information available in order to better reflect the size of the particular class of object, i.e. linear

measurement, surface area or landmark-based centroid size (Equation 2.3) can be used. In the case of facial surface, centroid size is used. The linear model of shape-by-size is defined as follows

$$\mathbf{s} = \mathbf{a}s + \mathbf{b}$$

Model parameters $\mathbf{a}$ and $\mathbf{b}$ are computed using multiple simple linear regressions as follows

$$\begin{aligned} \mathbf{b}_i &= \frac{\sum_{j=1}^{n}(\mathbf{s}_{i,j} - \bar{\mathbf{s}}_j) \cdot (s_j - \bar{s})}{\sum_{j=1}^{n}(s_j - \bar{s})^2} \\ \mathbf{a} &= \bar{\mathbf{s}} - \mathbf{b}\bar{s} \end{aligned}$$

where $\bar{\mathbf{s}}$ is the average shape and $\bar{s}$ is the average size. An example of relating the shape to the size property is shown in Figure 3.16.

Case study 7: Human face meshes

Using the tools described above, a study of a young adult Central European (Czech) population was carried out (Velemínská *et al.*, 2012). The aim of the study was to prove and describe the sex differences in facial shape as well as the relationship between size and shape in a population represented by a sample of 101 specimens (50 male, 51 female). The general observations are summarized as follows:

- Static allometry (individual variation within a population and age class) is a substantial part of sexual dimorphism. The cross-validation score of sex classification using linear discriminant analysis (see Section 2.2 [page 19]) exhibited a strong correlation with centroid size ($R = 0.578$, $R^2 = 0.334$, $p < 0.0001$). Apart from that, discrimination rates of shape variables (size-normalized) yielded 87.13% of correctly classified specimens using leave-one-out cross-validation. Large sizes are related to facial elongation, while small faces are more rounded. For a comparison, see Figures 3.16(a) and 3.16(c).

- The sexual dimorphism of the shape is expressed by the width proportionality of the face. Masculine faces display a low and narrow forehead in combination with a larger lower face (Figure 3.15(a)). For feminine faces, a high, vaulted and wide forehead is associated with a smaller and less prominent lower face (Figures 3.15(b)). Female variation is almost completely overlayed with part of the male variation (see Figure 3.17). Therefore, for female classification, shape in sexual dimorphic traits appears to be more important while, in males, there is a stronger association with size.

- The forehead in males is more sloping while the eyes are located deeper in relation to the periorbital area. The deeper position of the eyes in males,

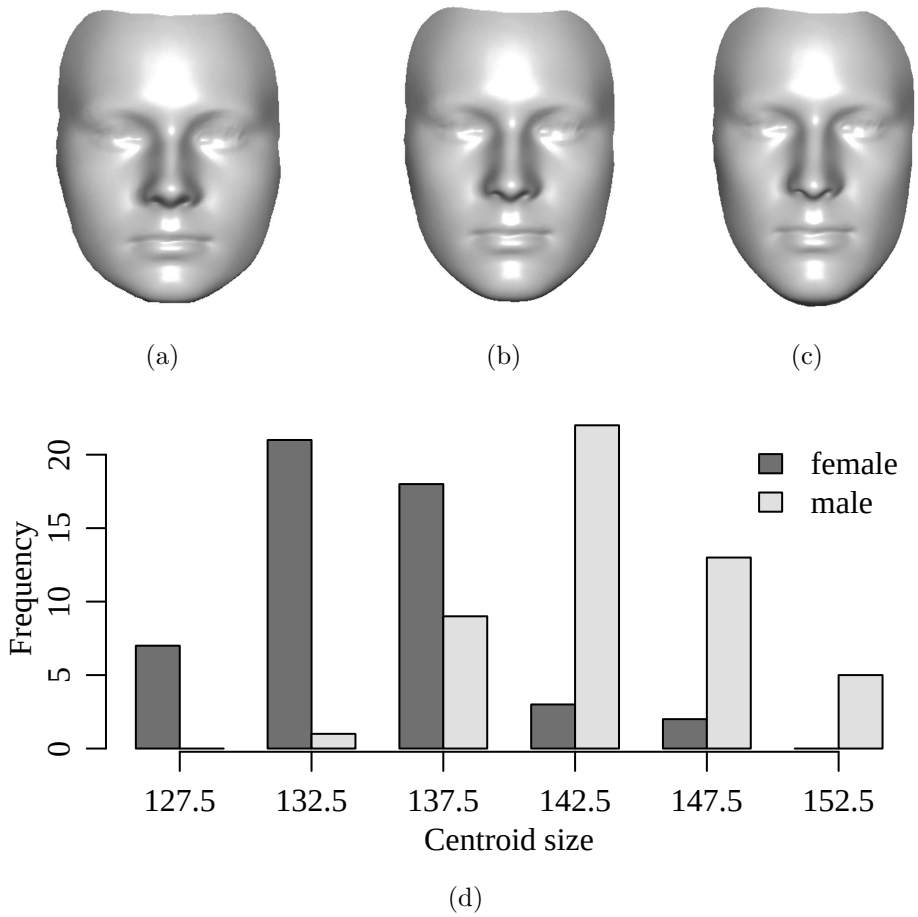


Figure 3.16: An example of allometry in facial meshes: (a) average small face; (b) Average medium-sized face; (c) average large face; (d) histogram of face sizes in the sample.

in combination with a narrower eye fissure, is probably the consequence of a larger and more protuberant frontal sinus (Figure 3.15(a)).

- In the mid-facial area, the male external nose is enlarged anteriorly (a larger convex nose) and the nostrils are strikingly larger and more flaring. The face is relatively narrower with less prominent cheeks and a smaller amount of soft (probably fatty) tissue in the cheeks. For a comparison, see Figures 3.15(a) and 3.15(b).

- The lower face in males is larger with massive, well-shaped musculature enveloping the mouth aperture and prominent chin (Figure 3.15(a)).

As mentioned above, the research was realized using a dense correspondence construction algorithm based on landmarks and TPS for non-rigid registration (Section 3.2.3). It is also interesting to see how B-spline based regis-

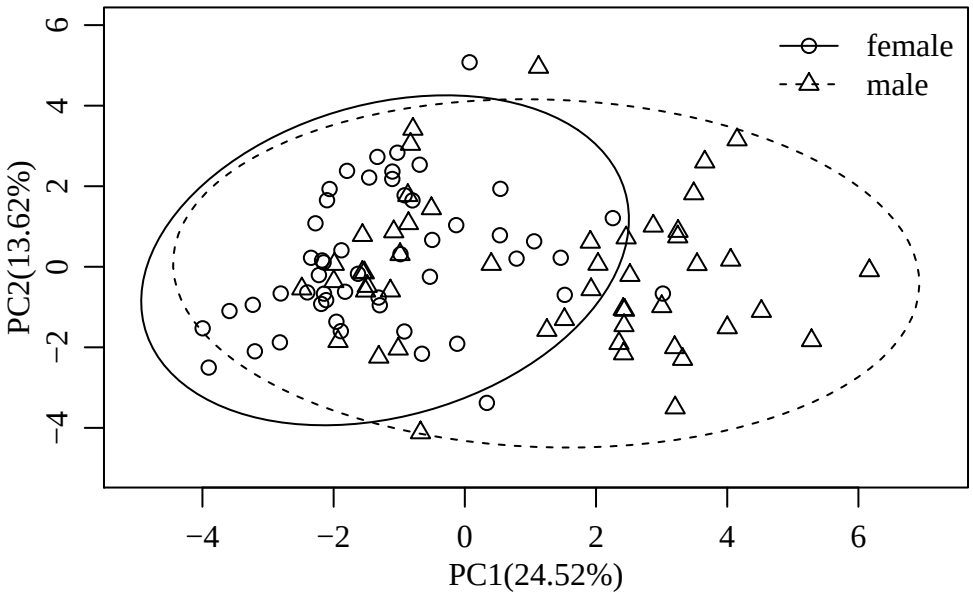
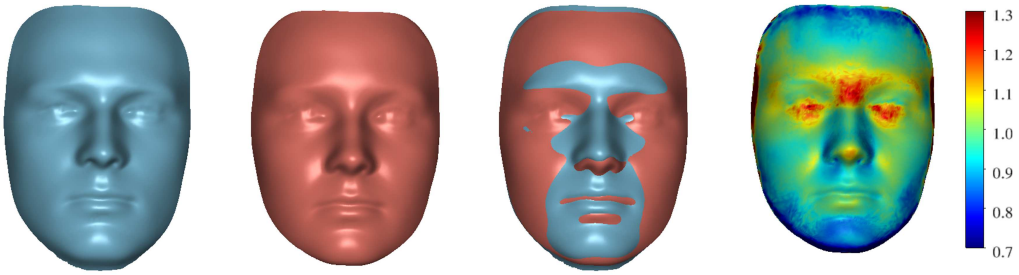


Figure 3.17: A scatterplot of specimen scores in the first two principal components (38.14% of the overall variability) showing projection of the shape space where the relation of sexes is the most distinguishable. Confidence ellipses in the plot approximate 95% of the represented population.

tration (Section 3.2.4) performs in the same tasks on the same datasets. The average shapes are shown in Figure 3.18. Yet, no ground truth exists against which the results can be compared with, so the assessment is based purely on visual observations.

In comparison to mean shapes acquired via TPS-based registration (Figure 3.14), the differences between males and females are also very subtle. However, they better reflect commonly known sex differences, i.e. high superorbital ridges; a larger nose positioned higher in the face; massive soft tissue surrounding the mouth along with a pronounced chin in the male face in contrast to the round face of the female with prominent cheeks and vaulted forehead. On the other hand, results of sex discrimination by LDA can provide strong evidence that additional computational and algorithmic labor are worth the effort. Actual results of LDA on the same data give a 94.1% discrimination success rate with leave-one-out cross-validation, which is approximately 7% better, i.e. 7 specimens were classified correctly in the sample in contrast to the inferior TPS-based registration method. This difference is particularly significant and suggests that LDA discrimination ability reaches its limit on this dataset and cannot be improved considerably, even when it is applied to much more precise feature descriptors. Experimenting with more complex classifiers, such as various modifications of support vector machines (Cortes and Vapnik, 1995), naive Bayesian classifier or artificial neural networks, yielded even better clas-



(a) (b) (c) (d)

Figure 3.18: Comparison of mean shapes of male and female faces created using B-spline non-rigid registration (compare with Figure 3.14): (a) male mean face; (b) female mean face; (c) overlayed mean faces; (d) color-coded difference in local size measure.

classifier	**B-spline** (%)	**TPS** (%)	**difference** (%)	**landmarks** (%)
LDA	94.1	87.1	+7.0	72.3
SVM(radial kernel)	97.0	83.2	+13.8	74.3
Naive Bayes	92.1	87.1	+5.0	68.3
ANN(hidden=4)	89.9	79.0	+10.9	60.8

Table 3.6: The success rate of sexual dimorphism classification of facial shapes using various classifiers of correspondences constructed using B-spline registration and TPS-based registration. In addition to surface analysis, sexual dimorphism evaluation on landmarks is included for comparison purposes.

sification success rates in some cases (see Table 3.6).

The variability of shape captured by the two models is better compared by visualizing the effect of principal components on the shape itself (see Figure 3.19). The difference in coverage of the principal components between models, acquired using two different registration methods, suggests that they reflect similar features of the shape, i.e. PC1 reflects border parts of the shape (not clearly visible in Figure 3.19(a)), PC2 reflects the forehead, PC3 reflects the forehead, chin and lateral parts of the face and PC4 reflects the forehead and cheeks. However, the components of the B-spline-based model go into more detail (central part of the face, asymmetry), probably because of the stability of found correspondences. On the other hand, component coverage of the TPS-based model is more general, mainly covering the forehead and marginal parts of the face. This is confirmed by generalization measure G (Equation 3.19), which is equal to $3.25 \cdot 10^4$ for the TPS-based model and $4.17 \cdot 10^4$ for the B-spline-based model.

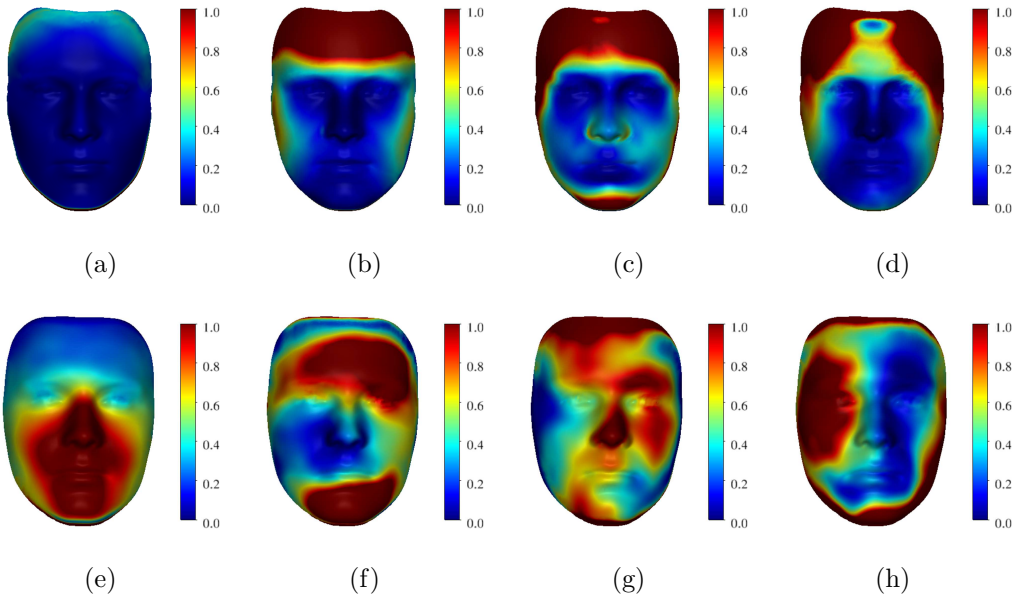


Figure 3.19: Comparison of face variability acquired using two registration methods: (a)-(d) PC1 - PC4 of the model based on TPS registration representing 24.5%, 13.6%, 9.4% and 7.5% of variability, respectively; (e)-(h) PC1 - PC4 of the model based on B-spline registration representing 23.3%, 15.5%, 11.9% and 5.7% of variability, respectively. Color-coded maps show the distance in mm by which the components affect a position of a particular point scaled by 10^4.

Case study 8: Cleft lip and palate

The second case study that demonstrates mesh shape variability analysis is an application used in the field of biomedicine, namely, the analysis of the shape of the human palate with respect to the evaluation of cleft lips and palate conditions. This analysis is performed on a sample of patients and compared to healthy population. A cleft lip and palate is a clefting deformity caused by abnormal facial development during the gestation period. This condition is relatively common in newborns, with a prevalence of approximately 0.15%, which is equal to 1 per 700 births worldwide and affected by various factors (ethnicity, genetic factors). In the Czech Republic, the incidence is approximately 1.8 per 1000 births (Peterka *et al.*, 2000). Apart from the esthetic aspects of the condition, there are many others that affect the quality of the patient's life, such as mastication and food intake difficulties, speech problems, hearing problems and psychosocial development problems. A severe case of cleft lip and palate is called *complete cleft palate*, since it continues above the upper lip into one or both nostrils. There are generally two types of complete cleft lip and palate: unilateral (UCLP) and bilateral (BCLP) cleft lip and palate. Due to the severity of this condition, a great deal of research has developed, with

(a) (b) (c)

Figure 3.20: An example of shape analysis based on a TPS registration-dense correspondence model of a human palate: (a) digitized model of a human palate with landmark configuration; (b) mean shapes of a healthy (gray) palate and a BCLP patient (green); (c) visualization of local size differences between means, i.e FESA (taken from Bejdová *et al.* (2012a)).

many procedures now capable of fixing the condition. Surgical procedures are even performed on infants in their first months of life. The advantage of the treatment at this very young age is that the natural growth will allow the palate to develop to a healthy looking shape. The measure of success of a particular protocol for early-age surgery can be evaluated over a decade or decades of natural growth.

A digital image of the human palate is captured by scanning a palate cast with a surface scanner, usually based on more precise laser technology. An example of a cast and its digital image is shown in Figure 2.1(c) and Figure 3.20(a). These digital scans are further processed using software tools (Rapidform, INUS Technology), mostly manually, to remove teeth arches so that the analysis will only consider the shape of the palate. Until recently, the condition of the palatal shape was analyzed only by linear measurements on palate casts or by constructing cross-sections of a digital model (Šmahel *et al.*, 2009). The first palate shape analysis of a complete surface model was performed on BCLP and UCLP by Bejdová *et al.* (2012a) and Rusková *et al.* (2014) using TPS-based registration.

Figure 3.20(b) shows the shape of an average BCLP palate of a sample, which forms a group of 29 Czech boys between 12.1 and 16.5 years, who underwent the same treatment procedure (modified Veau technique at an average age of 8 months followed by pharyngeal flap surgery at the age of 5 years). The control group of a healthy population consisted of 29 Czech boys with an average age of 14.7 years. One difficulty in the application of TPS registration is finding enough well-defined and stable landmarks. The shape of a palate basically recalls a half of a paraboloid without as many significant feature points. A configuration of four landmarks are suggested: the first is placed on the tip

of the palate, two more are positioned on the most medial cervical margins of the second premolars and the fourth is placed in the middle of a line connecting the first molars at the most medial cervical margins (see Figure 3.20(a)). Four points in a three-dimensional space do not allow for very complex deformation using TPS. Bejdová *et al.* (2012a) used the visualization of shape differences between treated cleft patients and controls as the key results. The comparison might answer the question whether the treatment is actually able to restore a healthy shape and, if not, focus the attention on what to change in future surgical and treatment protocols.

For a detailed depiction of the local differences, finite element scaling analysis was used (FESA), developed by Cheverud and Richtsmeier (1986). The method compares the geometry of individual corresponding triangles between two triangular meshes and assigns two values expressing the difference in local size and shape. In the case of the study, the pair of meshes correspond to a mean healthy palate and a mean clefted palate. Figure 3.20(c) shows the local differences in size. The figure appears to be rather noisy, which is due to element-wise FESA factors evaluation.

The summary of observations uncovered in the study are the following (for more details, see Bejdová *et al.* (2012a)):

- The variability of palatal size and shape in BCLP patients is much greater in comparison with that observed in the non-clefted population.
- The palatal size and shape of BCLP patients (ranging from 12.1 to 16.5 years) are not correlated with age. This means that there is no connection between how the palate looks and how much time it has taken to develop. Hence, there is no bias connected with the age structure of the sample.
- A comparison of the mean palatal shapes of the clefted and non-clefted groups shows that the BCLP is flatter and narrower. The most notable size difference is seen in the area between the maxilla and premaxilla. This phenomenon could be associated with the persisting separation of the premaxilla from the rest of the palate.
- The shape of the mean palatal configuration of the premaxilla and adjacent area in lateral projection is concave in the non-clefted population and convex in BCLP patients.

As mentioned above, a palatal shape is difficult to describe using anthropometrical landmarks. However, TPS registration heavily depends on a number of landmarks that cover the whole body of the shape. Regions away from landmarks and outside of their convex hull are not aligned very well due to the global behavior of the TPS interpolation function.

Free-form deformation, on the other hand, behaves locally with various levels of refinement. An advantage of B-spline-based registration is that it is

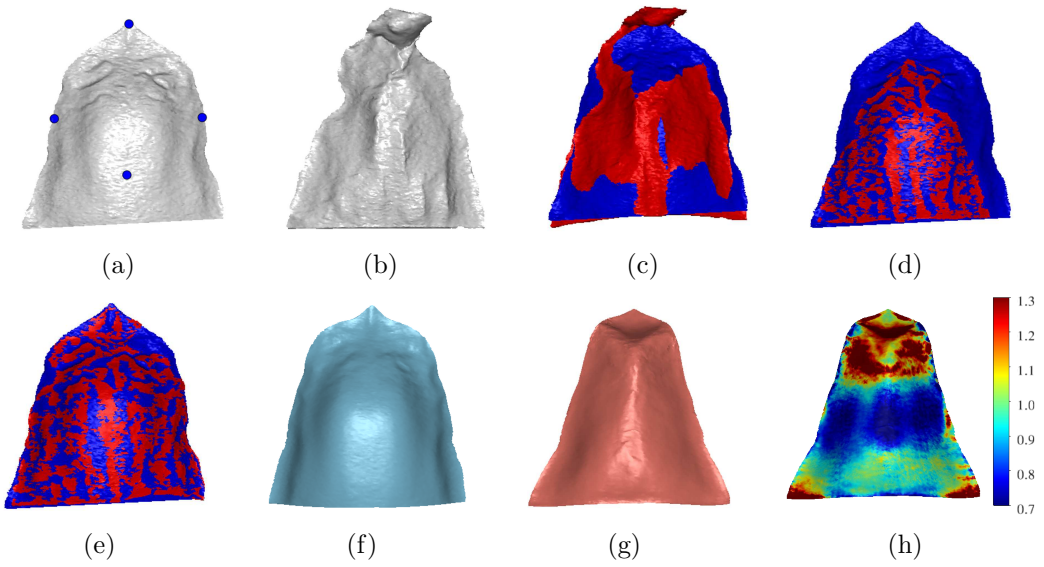


Figure 3.21: B-spline non-rigid registration of palates: (a) input of a healthy palate with landmarks; (b) an example of a clefted palate; (c) rigidly aligned palates; (d) non-rigid registration of a clefted palate to a healthy palate; (e) non-rigid registration with landmark constraints; (f) mean shapes of a healthy sample; (g) mean shapes of a BCLP sample; (h) local size differences of a mean healthy palate and a BCLP mapped onto a mean BCLP shape.

able to fit mesh borders, especially beneficial in cases where palate meshes are trimmed along the teeth line. The teeth line actually serves as a helpful feature for correspondence matching. An example of a pre-processed healthy palate and BCLP is shown in Figure 3.21(a) and 3.21(b). Applying B-spline-based registration and minimizing the sum of squared distances to the nearest neighbors will not result in the expected fitting, because the simple feature-less shape of the palate allows the source surface (clefted) to slide along the target surface (healthy) and shrink, while minimizing the distances before the movement is stopped by regularization. In order to deal with this problem, the B-spline registration framework makes use of landmark constraints (see Section 3.2.4). The landmarks keep the homologous places overlaid, while the sum of squared distances brings the respective surfaces closer to each other, where regularization prevents the undesired local deformation. Constrained B-spline registration results in almost a perfect fit (see Figure 3.21(e)). Consequently, meaningful correspondences are found which allow plausible mean shapes of healthy (see Figure 3.21(f)) and clefted palates (see Figure 3.21(g)) to be created, compared with means acquired using TPS-based registration (Figure 3.20(b)). Although the model basically reveals similar results, the shape of the mean clefted palate significantly differs from the TPS version. The lateral teeth line is concave along the entire side, making the middle section of the shape fairly narrow. The detail of the valley between the maxilla

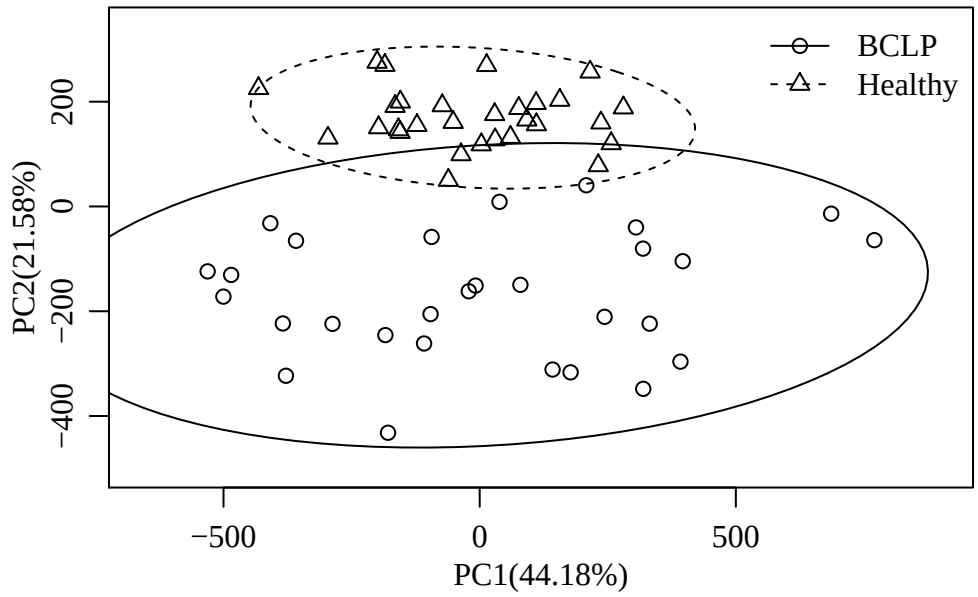


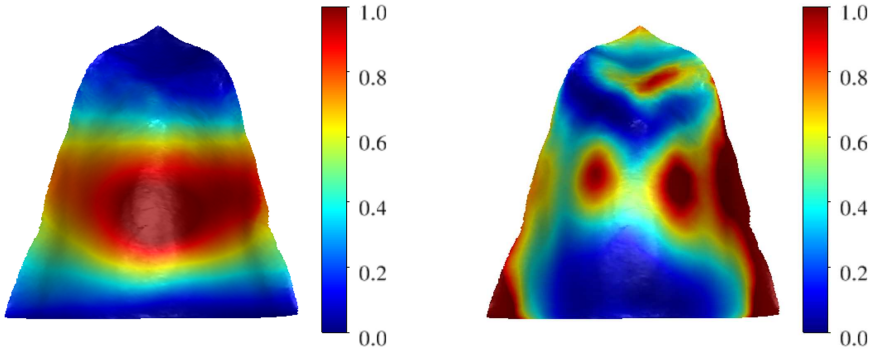
Figure 3.22: A scatterplot of specimen scores in the first two principal components containing subsamples of the healthy population and BCLP patients approximated with 95% confidence ellipses.

and the premaxilla is preserved in the mean shape. Instead of FESA, local size differences are computed using a k-nearest neighbor mean distance ratio, which proves the local size expansion due to the valley and the convex "bump" on the side of the premaxilla (see Figure 3.21(h)).

In agreement with published results, Figure 3.22 shows that the variability of the BCLP group is much larger than the variability of the healthy population and that these two subgroups are well separated by the score value related to the second principal component. The effect of the second component in the geometry of the palate is seen in the area of the lateral teeth line in the middle section and in the valley between the maxilla and the premaxilla, as shown in Figure 3.23.

Case study 9: Evaluation of a SARME procedure

The last case study of mesh shape variability analysis also focused on the shape of the human palate. However, in this study the primary goal is to visualize the effect of a surgically-assisted rapid maxillary expansion (SARME) procedure. SARME is a solution for a medical condition called "transverse maxillary deficiency", which occurs when the upper jaw is too short and narrow for the jawbone, resulting in misalignment of the upper teeth arch with the lower teeth arch. The condition causes esthetic and functional problems. SARME is applied to young adult and adult patients who have already reached skeletal



(a) (b)

Figure 3.23: Palatal shape analysis with a B-spline registration-based dense correspondence model of a human palate: (a) magnitude of PC1 affecting a palatal shape; (b) magnitude of PC2 affecting a palatal shape.

maturity. In practice, the surgery is based on fixing an appliance (a Hyrax expander) to the middle of the upper teeth arch in the patient's mouth. This is bound to the first and the third molar and screws are then used to apply lateral force while cutting the maxilla bone (osteotomy) horizontally above the teeth roots and vertically behind the teeth arch. The arch is realigned to a desired position in which the separated tissues grow together. After several weeks the appliance is removed.

In the case study (Trefný *et al.*, 2015), dental casts of 15 patients were made before the surgery and after the appliance was removed. The size of the sample was limited since the condition is relatively rare. The sample was compared with a control group of 50 healthy palates (23 males and 27 females). A TPS-based dense correspondence algorithm was used to evaluate palate variability of the control group and pre- (SARME T1) and post-treatment patients (SARME T2), while the statistical significance of differences between pre-treatment and control groups and post-treatment and control groups were tested (using a multivariate permutation test for mean equality).

Although the samples were relatively small, the results are rather conclusive. Figure 3.24 shows that SARME patients have a much wider variability than the control group. However, statistically significant differences were proved only between the control group and SARME T1, not between the control group and SARME T2. One possible interpretation is that the SARME procedure changed the shape of the palate to the extent that it could no longer be distinguished from the palate of a healthy person. The arrows in Figure 3.24 unambiguously demonstrate shape changes favoring the control group in all cases as a result of the SARME procedure. Figure 3.25 shows mean shapes of the aforementioned sub-groups.

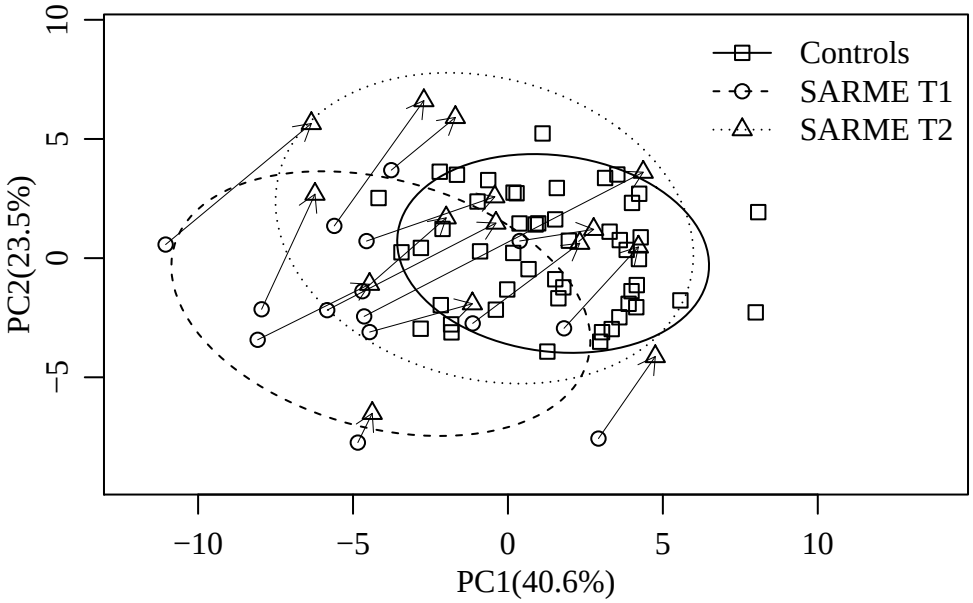
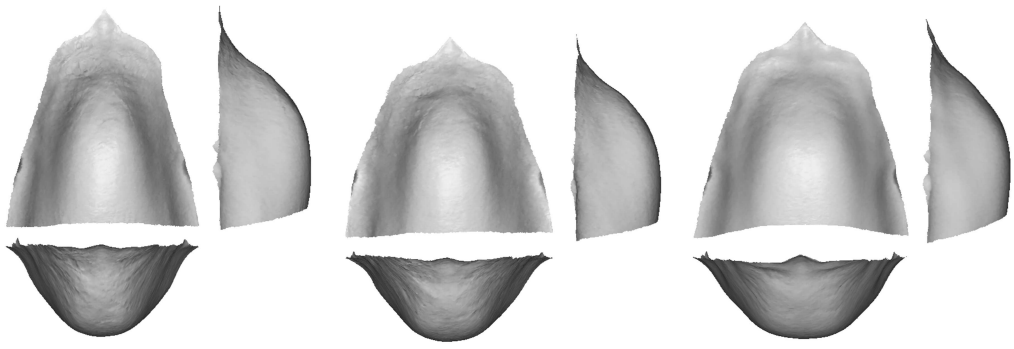


Figure 3.24: A scatterplot of specimen scores in the first two principal components of shape analysis from mixed SARME patients in both pre-, post-treatment (T1 and T2) and control groups. Groups are approximated with 80% confidence ellipses.



(a) (b) (c)

Figure 3.25: SARME pre- and post-treatment mean shapes compared to a control sample mean shape computed using a TPS-based correspondence construction algorithm: (a) SARME T1 mean shape (before surgery); (b) SARME T2 mean shape (after removal of the corrective appliance); (c) control sample mean shape.

3.5 Mesh asymmetry analysis

In this section, methods used to assess bilateral asymmetry — the deviation from ideal bilateral symmetry on mesh data — are presented. The importance of symmetry in life sciences was already highlighted in the introductory chapter (see Section 2.4).

In connection to non-landmark data, a couple of studies have been performed recently which evaluate bilateral asymmetry in medical images. For example, symmetry as an important feature of the human brain is a topic that has been intensively studied in the past. In the modern era there have been many attempts (e.g. by Fournier *et al.* (2011)) to analyze MRI images and interpret brain asymmetry in terms of their connection to illnesses, functional abilities and genetics.

However, there have not been many attempts at automatic analysis of asymmetries in mesh data. One particular approach developed by Liu and Palmer (2003), maps objects of interest onto a surrounding cylindrical surface. In the reference cylindrical coordinate system, the corresponding symmetric points are found with the help of manually placed landmarks. Asymmetry is then deduced from these pairs. This approach obviously works only for simple shapes that can be unambiguously projected onto a cylinder.

A different method by Ólafsdóttir *et al.* (2007) assumes the existence of an ideally symmetric template and then maps each subject in the study onto this template using B-spline-based non-rigid registration. Construction of the ideal or, indeed, any symmetric template for a given group is no trivial task. Constructing such a template requires a mid-sagittal plane, around which the template needs to be bilaterally symmetrical.

Another approach suggested by Combes *et al.* (2008) constructs a mid-sagittal plane using a modified EM algorithm (Combes and Prima, 2008) and uses it to mirror the studied shapes. Asymmetry is then represented as the distance between the corresponding points on the original and mirror shapes. Correspondence is determined using non-rigid registration, bending the mirror shape to the original.

In geometric morphometry, asymmetry has been analyzed on landmark datasets (e.g. in Bookstein (1997); Schaefer *et al.* (2006)) by mirroring landmark configuration and reordering the landmarks so that the mirror and the original can be realigned. The asymmetry is then defined as the difference of the ideally superimposed mirror and the original. More importantly, the approach decomposes the asymmetry, as has been traditionally studied in biological sciences (Van Valen, 1962; Palmer, 1994). The following presented approach utilizes some of these ideas.

The method captures group mesh asymmetry and uses the concept of decomposing the asymmetry into *directional* and *fluctuating* components (Krajíček *et al.*, 2012). Concurrently, dense triangular meshes allow the asymmetry to be expressed on a very localized level. Moreover, captured correspon-

dences help to explore the variability of the asymmetry using statistical tools such as PCA.

Before the local asymmetries can be compared in all meshes, they need to be transformed into a common coordinate space. This is done by applying a group-wise rigid landmark-based registration, specifically, the use of generalized Procrustes superimposition. The meshes' vertices are transformed the same way as the landmarks in the Procrustes superimposition.

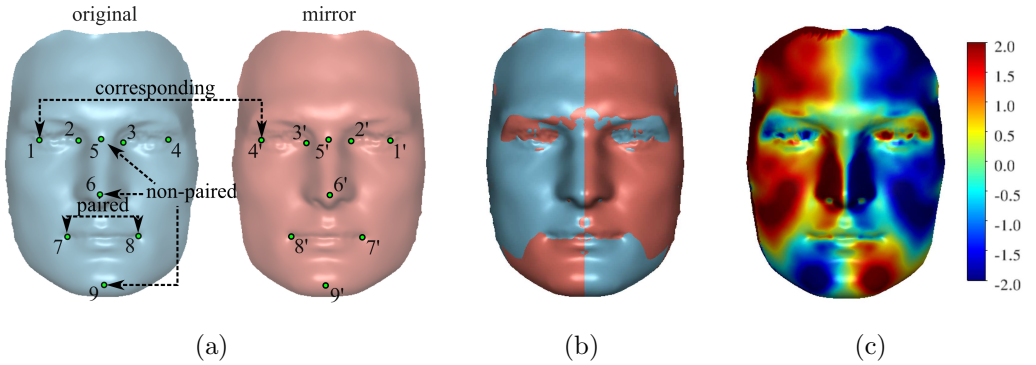
The next step is to recompute the meshes to the same number of vertices and the same topology. The dense correspondence construction algorithm based on TPS registration described in Section 3.3 is used.

At this point, the individual mesh asymmetry is computed as the difference between the mesh and its aligned mirror reflection. The result is a list of directions for every vertex. If all vertices are moved by their respective displacement, the mesh becomes ideally symmetric. The mirror mesh must be constructed by negating one of the vertex coordinates; the same is done with landmarks placed on the mesh.

The landmarks are then used to align the mirror mesh back to the original mesh. In order to do that, the landmarks must be reordered, since they change their homologous meaning after mirroring, e.g. some landmarks on the left side of a bilaterally symmetric mesh become the landmarks on the right side of the mirror mesh. These are known as paired landmarks; they swap their positions with their mirror counterparts. The others non-paired landmarks are not affected by mirroring (see Figure 3.26(a)).

After mirroring and landmark reordering, the mirror meshes are realigned by ordinary Procrustes superimposition (Figure 3.26(b)). Any non-rigid registration procedure can be used to deform the mirror meshes in order to get them closer to the original ones and to enable subsequent correspondence searching. Such a procedure may be different to the registration method used for group correspondences. The impact of the chosen procedure is discussed further below. The closest points on the mirror meshes to each vertex of the original correspondence meshes are found using search acceleration structures. Again, a k-d tree is used for this purpose.

The vectors defined by the difference between the original mesh vertices and their closest mirror mesh points are the local measures of the asymmetry. Completely symmetrical shapes have identical mirrors and, when aligned, the distances between the original mesh vertices and the closest points are zero. If asymmetries occur on the mesh, the difference between the left and the right part of the mesh appears and the distance between the part of the mesh and its mirrored counterpart becomes non-zero. Furthermore, the associated vector holds the information about the direction of the asymmetry, i.e. how the part of the mesh is moved to form the asymmetry. This information is exhibited on either side of a bilaterally symmetrical mesh in opposite directions. Therefore, it is not possible to determine which part of the mesh originates from a symmetric shape and which is altered, assuming that this is the way the analyzed



(a) (b) (c)

Figure 3.26: Principles of individual mesh asymmetry analysis: (a) the original mesh and mirror mesh with corresponding paired landmarks; (b) the aligned original mesh and mirror mesh; (c) individual asymmetry (scale in millimeters).

asymmetric shape was created. From this point of view, bilateral asymmetry is a symmetric feature. The vector field that represents the displacement of a point on a mesh, from where it would lie if the mesh were ideally symmetric, is called *individual asymmetry*.

The aforementioned vector field on the original mesh is visualized with color-coded signed distances (see Figure 3.26(c)). The sign is the same as that of the dot product of the mesh normal and the vector of the individual asymmetry in that point. The color images can be simply interpreted in the following way: red areas lie in front of the corresponding mirrored counterpart which means that they are larger than the corresponding paired counterpart, while blue areas are smaller and lie behind the aligned mirrored counterpart. The areas that are close to green are not significantly larger or smaller. This interpretation does not include any information about the direction of the asymmetry. This sort of visualization is also known as *clearance vector mapping* and is useful in quantifying the facial surface asymmetries in the areas where anthropometric landmarks are scarce. The volume of detected asymmetries is potentially significant in patients who have their unilateral facial deficiencies corrected using injections or implants (O'Grady and Antonyshyn, 1999).

All individual asymmetries are already aligned group-wise; therefore, directional asymmetry is computed as the average of all corresponding individual asymmetry vectors. The lengths of the respective individual asymmetry vectors are the same for all the meshes and they correspond to one another for every element since the meshes are reconstructed to have the same topology. The directional asymmetry is visualized in the same way as the individual asymmetry (see Figure 3.28(a)).

Fluctuating asymmetry FA_i is computed as the difference between individual asymmetry IA_i and directional asymmetry DA. Its visualization is also

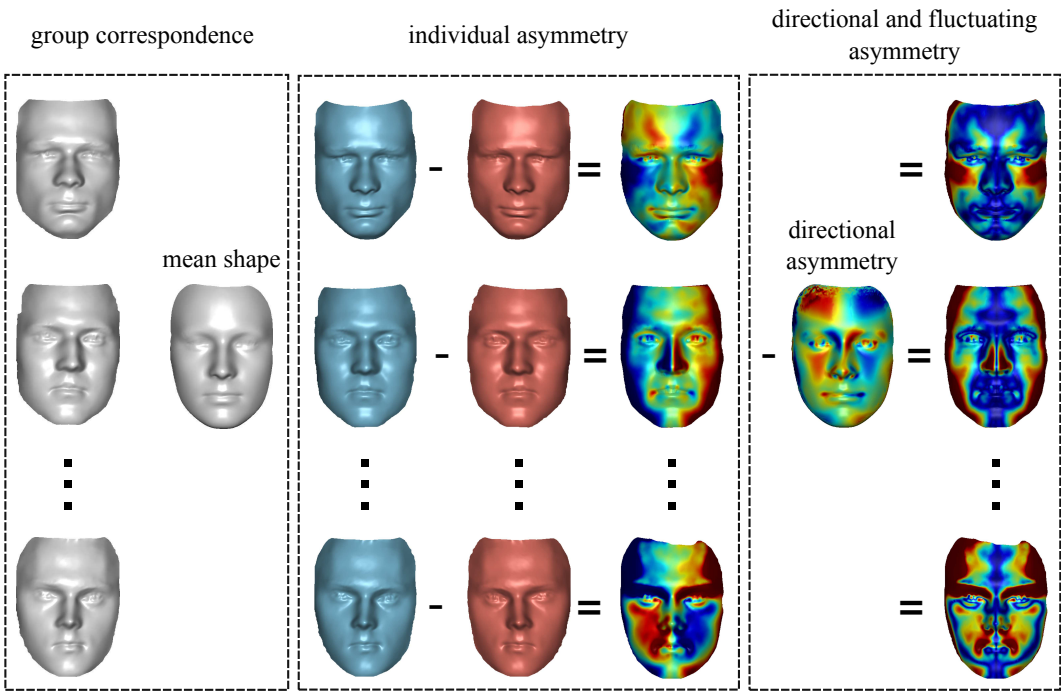


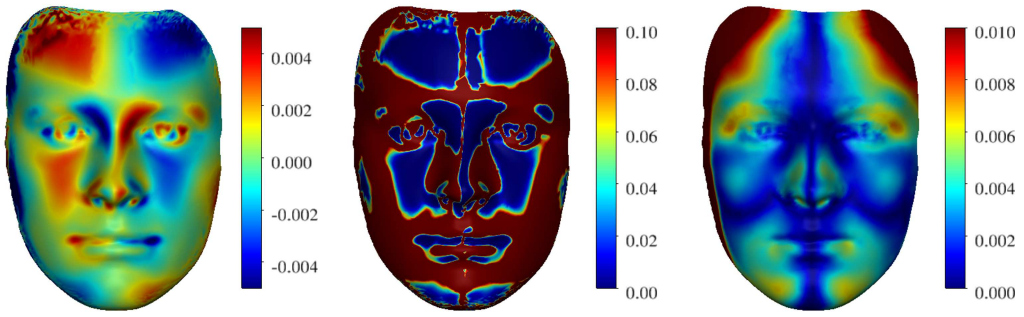
Figure 3.27: Schematics of statistical mesh asymmetry analysis.

based on color-coded distances without considering the sign of direction, as performed for individual and directional asymmetry above. As stated in Section 2.4, the point of interest is the overall magnitude of the fluctuating asymmetry, which is computed as the sum of the squared distances of the fluctuating asymmetry vectors. It can be compared across the group or, if normalized by the number of vectors, between groups just as effectively. Figure 3.27 shows explanatory schematics of mesh asymmetry analysis.

In order to prove that the directional asymmetry reflects the global trend of the group and is not the result of randomness in the group, it must be tested statistically. A standard t-test is performed on the signed lengths of corresponding individual asymmetry vectors. The significance map can then also be visualized (see Figure 3.28(b)). This way of interpretation is especially important for specific research in sub-fields of life sciences.

Apart from the directional asymmetry that reflects the overall asymmetry trend in the sample, the asymmetry can be decomposed into orthogonal components and analyzed in a similar way to the variability of the mesh shape, as discussed above. This is especially useful both when analyzing samples of two populations together and when searching for differences in their asymmetry trends. The approach is based on PCA of individual asymmetry vectors and construction of linear asymmetry models, i.e.

$$IA_i = DA + \sum_{j=1}^{n} \alpha_{i,j} A_j$$



(a) (b) (c)

Figure 3.28: An example of shape asymmetry analysis based on a dense correspondence model of a human face: (a) directional asymmetry with a general trend in the sample; (b) significance map (p-values); (c) the first component of asymmetry, i.e. an independent part of the asymmetry with the highest variance.

where $\alpha_\mathbf{i}$ is the asymmetry score vector of the particular specimen. Individual asymmetry components (modes of variation) demonstrate the variability of the asymmetry in the sample, similar to shape variability components (see an example of a component in Figure 3.28(c)). For visualization purposes, the shape of the specimen can be reconstructed

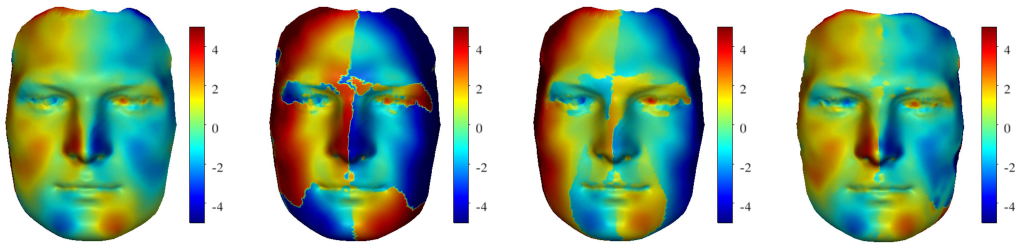
$$I_i = S_i + IA_i$$

where S_i is the individual symmetrical shape, which can be simply replaced for visual asymmetry inspection by symmetrical mean shape $\bar{S}$.

The direction of the individual asymmetry vector is another important property that should be taken into account. The *local orientation difference asymmetry measure* is defined as a cosine of the angle between the corresponding individual asymmetry vectors. The lengths of the individual asymmetry vectors and the local orientation differences of the asymmetry measures can be summarized in terms of *total asymmetry* (*TA*) and *total orientation asymmetry* (*TOA*).

$$\begin{aligned} TA &= \frac{1}{|C|}\sum_{i=1}^{|C|} \|\vec{a_i}\|^2 \\ TOA &= \frac{1}{|C|}\sum_{i=1}^{|C|} \vec{n_i}\cdot\vec{m}_i \end{aligned}$$

The approach described above gives a generic idea of how to extract information from asymmetries in a group of samples. The results can be improved using different registration procedures for group normalization as well as for



(a) (b) (c) (d)

Figure 3.29: Comparison of correspondence mapping deviations using various registration methods: (a) landmark-based rigid registration; (b) TPS-based original mesh-to-mirror fitting; (c) TPS-based original mesh and mirror mesh-to-symmetrized landmark configuration fitting; (d) B-spline-based non-rigid registration.

individual asymmetry computation. Coherent point drift has been used with satisfactory results by Dupej *et al.* (2013). The main advantage in this case is that no user-defined landmark points are needed for the registration, which is one step closer to automatic analysis.

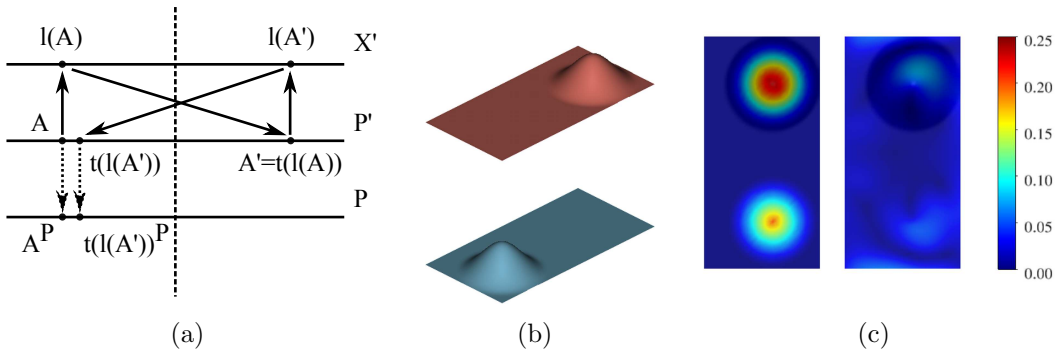
Further, the suitability of the registration procedure and correspondence matching between the original mesh and the mirror mesh is evaluated by comparing four approaches: rigid landmark-based alignment, TPS-based fitting of the mirror to the original mesh, TPS-based fitting of both the mirror and the original mesh to a symmetrized landmark configuration, and B-spline-based non-rigid registration (see Section 3.2.4). Symmetry is an important trait when observed in a color map of corresponding point distances. Although the method is not perfect due to the asymmetric nature of the specific mesh topology, the main features of the symmetry tend to be generally preserved because small features on one side appear as larger features on the other side and vice versa.

An individual asymmetry assessed by the four above-mentioned methods is visualized in Figure 3.29. The asymmetry of the specimen captured by the methods is generally the same. However, the difference is in the details and the magnitude as well as in the direction, of which the latter, although not visible, is the most significantly influenced by the correspondence mapping associated with the registration method.

In order to determine the quality of found correspondences with respect to the bilateral symmetry, the *correspondence symmetry error metric* must be defined.

$$ErrSym \quad = \quad \sum_{A \in P} \|A^P - t(l(A'))^P\|^2 \tag{3.20}$$

where P is a mesh and X is its mirrored copy rigidly aligned with P. P' and X' are the same meshes non-rigidly deformed to fit each other. A is a point on



(a) (b) (c)

Figure 3.30: An explanation of the *correspondence symmetry error metric*: (a) a diagram showing the principle of symmetry error metric computation; (b) an example of an asymmetrical sheet and its mirror; (c) an asymmetry error computed using correspondence mapping between rigidly aligned shapes (left) and between non-rigidly tightly fitted shapes with B-spline registration (right).

P'. A^P is a corresponding point on the original mesh P. Function $l(.)$ finds the closest point on X' to the argument. Function $t(.)$ finds a topologically identical point on P' to the argument via an equal triangle index and barycentric coordinates of the argument in X'. $A' = t(l(A))$ is a point on the other side of the mesh P'. $t(l(A'))^P$ is a point corresponding to A' on the other side, which is found by the procedure applied twice to A. If a symmetric point is found through nearest mapping, the distance between A^P and $t(l(A'))^P$ is zero. For a better explanation, see Figure 3.30(a). The rationale of the measure is that the points keep their homological meaning, i.e. the point on the mirror mesh represents the same feature and its correspondence on the other side of the mesh should be identical to the correspondence of that point on the original mesh.

Figures 3.30(b) and 3.30(c) also show a difference in the symmetry error metric of correspondences found after rigid alignment and after non-rigid fitting on artificial shapes. It seems that reliable asymmetry evaluation is related to reliable non-rigid registration, although it might not be true in the case of an actual application to face meshes. According to Figure 3.29, the asymmetry captured after rigid alignment and non-rigid fitting using B-spline-based registration is almost identical. However, as will be shown further, the value of the *correspondence symmetry error metric* is not very convincing when using B-spline-based non-rigid registration as is shown further. It appears that the mirror mesh is fitted tightly to the original mesh. However, symmetric points found by projecting the fitted mirror mesh onto the original mesh are shifted locally along the surface, which means that the total correspondence symmetry error metric is higher, even higher than in the case of TPS non-rigid registration. These results depend heavily on the various parameters that B-spline

specimen #	GPA	TPS	symmetric TPS	B-spline	symmetric B-spline
1	0.37	0.69	0.51	0.53	0.16
2	0.35	0.36	0.28	0.46	0.16
3	0.38	0.47	0.37	0.44	0.15
4	0.38	1.04	0.76	0.46	0.20
5	0.31	0.72	0.62	0.41	0.19
6	0.38	0.49	0.41	0.53	0.15
7	0.42	0.46	0.46	0.55	0.14
8	0.36	0.42	0.33	0.51	0.16
9	0.30	0.41	0.39	0.48	0.13
10	0.50	0.73	0.58	0.65	0.16
average	0.37	0.58	0.47	0.50	0.16
std. dev.	0.06	0.21	0.15	0.07	0.02

Table 3.7: Mean correspondence symmetry error metric ($ErrSym$) acquired using various registration methods, reflecting the reliability of found correspondences with respect to symmetry.

registration is controlled by:

- Regularization — higher regularization in a mirror-to-original registration process causes sub-optimal fitting, but the symmetry error metric is not significantly worse. However, it is not better than rigid registration alone.
- Normal direction weighting — high normal direction weighting influence (Equation 3.8) causes the moving mirror mesh to fit to the original mesh in a perpendicular direction. It limits sliding along the mesh surface, which causes higher correspondence asymmetry.
- Peripheral area weighting — the total symmetry error metric is biased by the fact that the mesh does not cover the same amount of area on both sides of the median plane. This factor can be reduced by weighting the samples in non-rigid registration by their initial distance after rigid pre-alignment.
- Minimization of the correspondence symmetry error metric — the B-spline framework allows the symmetry error metric to be included in the minimized cost function. Registration simultaneously minimizes symmetry error (Equation 3.20) as well as alignment error (Equation 3.1).

Ten face meshes were randomly chosen and all four registration methods were used to assess their individual asymmetries. The correspondence symmetry error metric and the respective alignment error, i.e. the overall distance

specimen #	GPA	TPS	symmetric TPS	B-spline	symmetric B-spline
1	0.89	1.23	1.23	0.59	0.27
2	1.21	1.04	1.04	0.65	0.21
3	1.22	1.22	1.22	0.68	0.21
4	1.17	1.62	1.63	0.59	0.26
5	1.19	1.99	1.98	0.57	0.25
6	1.12	1.13	1.13	0.63	0.20
7	1.14	1.64	1.64	0.62	0.18
8	0.92	0.87	0.87	0.53	0.23
9	0.86	1.53	1.52	0.59	0.21
10	1.18	1.44	1.43	0.68	0.19
average	1.09	1.37	1.37	0.61	0.22
std. dev.	0.14	0.34	0.33	0.05	0.03

Table 3.8: Mean alignment error measure (Err) using various registration methods, reflecting the reliability of found correspondences with respect to the distance from where they are found after registration.

on which correspondences are found by the nearest neighbor principle, were evaluated. The results are shown in Table 3.7 and Table 3.8. For B-spline registration, normal weight w_n is set equal to 0.8, regularization weight w_r is set equal to 0 and all samples of the mirror mesh further than 2mm from the original mesh, since initial rigid registration does not contribute to the registration criterion. The fifth experiment involves B-spline registration with a minimization of the correspondence symmetry error, which outperforms all previous methods in measured characteristics.

An example of symmetry error metric distribution over a face mesh, captured by various registration methods, is shown in Figure 3.31

In conclusion, correspondence between the mesh and its aligned mirror version is an effective tool for assessing asymmetry. However, in particular cases where the deviation from the symmetry is not extreme and the object of interest contains enough paired feature points (manually or automatically found), correspondences found by simple rigid alignment and nearest neighbor projection can be as good as, or even better than correspondences found using a sophisticated non-rigid registration procedure. At any rate, either the reliability of paired correspondences should be evaluated by the symmetry error metric before interpreting asymmetry results, or a registration method which directly optimizes the measure should be used. A surprising side effect of B-spline-based registration, the one which minimizes the symmetry error metric ($ErrSym$), is the better minimization of the alignement error (Err), i.e. fitting of the symmetric features improves the registration itself.

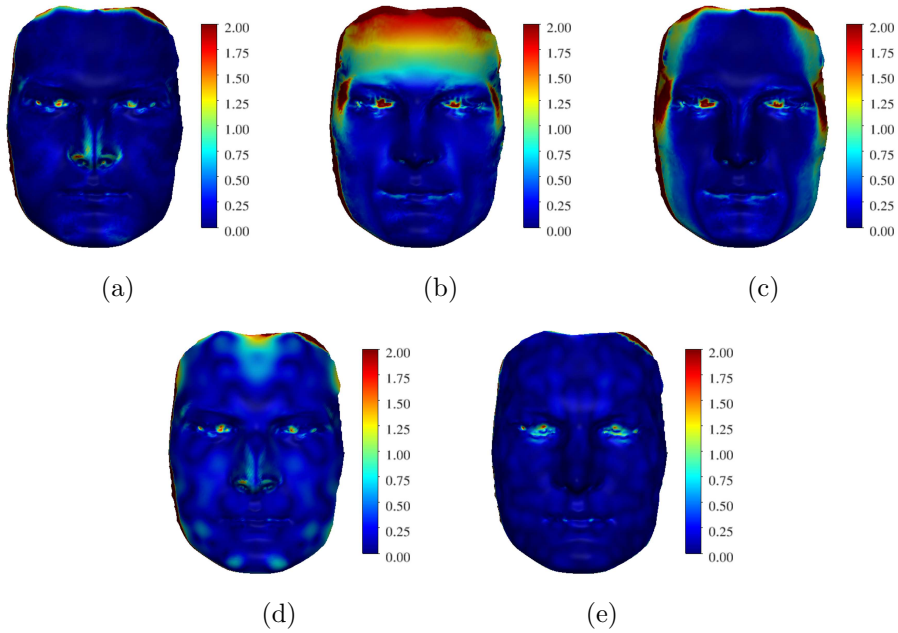


Figure 3.31: Comparison of correspondence symmetry error metric distribution using various matching methods: (a) landmark-based rigid registration; (b) TPS-based original mesh-to-mirror fitting; (c) TPS-based original mesh and mirror mesh-to-symmetrized landmark configuration fitting; (d) B-spline-based non-rigid registration; (e) B-spline-based non-rigid registration with minimization of the symmetry error metric.

Case study 10: Face asymmetry of UCLP patients

One particular example of mesh asymmetry assesment was presented in a study that evaluated the difference between unilateral cleft lip and palate patients (several years after treatment therapy) and a normal population sample (Cagáňová *et al.*, 2014). In Case study 8, the significance of the cleft lip and palate condition was explained in terms of how it affects many aspects of human life, both physiologically (how it influences the shape of the face and scarring up to the upper lip) and psychologically.

The study was based on facial scans of 12 children with complete UCLP at 5.5 years. All children had undergone *neonatal cheiloplasty* in the first week of their postnatal life. This surgical procedure is a promising technique, delivering very impressive immediate esthetic results along with many other benefits. The scans of the patient group were compared with a control group of 12 healthy children with an average mean age of 6.5 years.

The mesh variability analysis (see Section 3.4.1) of these two groups does not reveal significant differences, which may be attributed to the small size of the sample. However, the difference in the asymmetry score from these two samples appears to be statistically significant, after a permutation multivariate

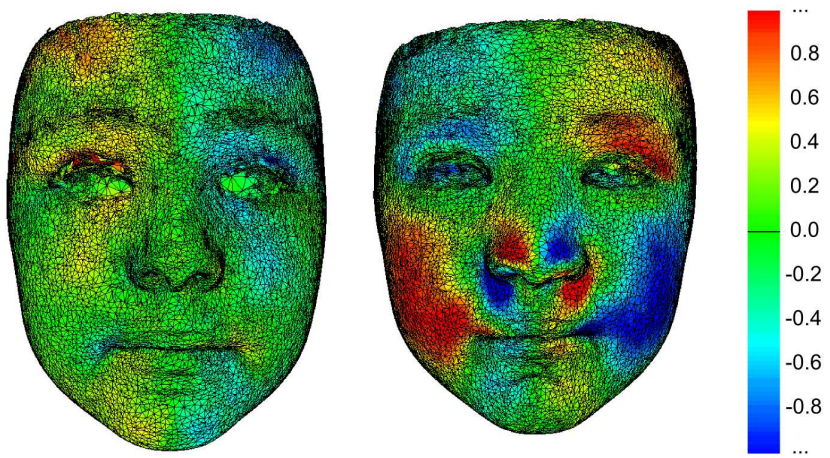


Figure 3.32: Face asymmetry in patients with UCLP: directional asymmetry of a healthy sample (left); directional asymmetry of a patient with UCLP (right).

test for means equality is performed ($p = 0.035$). In Figure 3.32, it is clear that the directional asymmetry is much stronger in UCLP patients than in the healthy sample.

3.6 Paired and longitudinal analysis

Extending the idea of searching for correspondences between related meshes allows various phenomena to be analyzed. When analyzed quantitatively, this offers a great opportunity to uncover behavior that would otherwise be unobservable in a single individual. One such example, analysis of the SARME procedure, was introduced in Case study 9. However, in that case, quantitative analysis was performed without searching for correspondences between the patient's shape before and after the procedure.

Paired analysis is defined as the quantitative analysis of the transition between the shape configuration of one shape at time t_1 and time t_2. A question of interest might be whether there are statistically significant differences in transitions between various groups. For example, is there a difference in the growth of male and female specimens over a period of one year, given their shapes are captured at the beginning and end of the defined period?

There are generally two types of scenarios for a correspondence search. The first is *inter-individual* correspondence matching, applied to the tasks in Section 3.4.1. This does not assume any relation or similarity between the two objects that are to be registered, not even in size or shape. Therefore, less precise correspondence matching can be expected because some parts of the shape might be missing in one shape of a pair or the other. This depends on the likelihood of similarity and regularization weights and landmark constraints (if

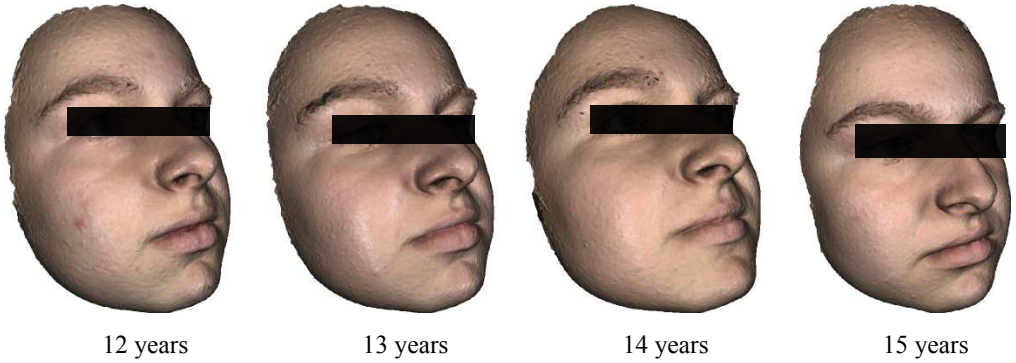


Figure 3.33: An example of longitudinal data — annual facial scans.

present) being set.

The second scenario is *intra-individual* correspondence matching, which assumes certain relations or similarities, at least in the parts of the shapes that overlap. It is also more likely for far greater detail to be captured since the analyzed group differences consist of many intra-individual specimen differences, as opposed to differences between two mean groups. Here, the registration algorithm should be chosen and attuned differently than for inter-individual correspondences. Intra-individual correspondence matching is the center of interest of this section.

To these methods of shape variability analysis (dense correspondence, PCA), *longitudinal data* can also be included (Krajíček *et al.*, 2013). Longitudinal data are multiple measurements of a single specimen collected in defined intervals over a period of time. An example of such data is shown in Figure 3.33. The goal is to analyze transitions from one stage of j-th specimen X_j^1 to another stage of specimen X_j^2. The first step of the analysis is to find correspondences across the first stages of the two for all the specimens in the sample. The base mesh is chosen arbitrarily. At this point, the mean mesh of the first stages $\overline{X^1}$ can also be computed. In the second step, correspondences between the two stages of each specimen are found individually, choosing the correspondence mesh (a mesh with common topology) found in the first step as the base mesh. Correspondences can be found using an arbitrary registration method, but this choice significantly influences the analysis. Krajíček *et al.* (2013) have used TPS registration-based correspondences.

Now the matrix M_c with n rows and $k \times 3$ columns can be defined, where n is the number of specimens in the sample and k is the number of vertices that are found during the correspondence search in the first step. The j-th row and $(i \times 3)$-th, $(i \times 3 + 1)$-th, $(i \times 3 + 2)$-th column of the matrix is a shift vector (dx, dy, dz) representing a shift of i-th vertex of the j-th specimen toward its second stage found during the second step (see Figure 3.34).

Matrix M_c captures complete information about the stage changes in the sample. The column average of M_c can be computed to get the average stage

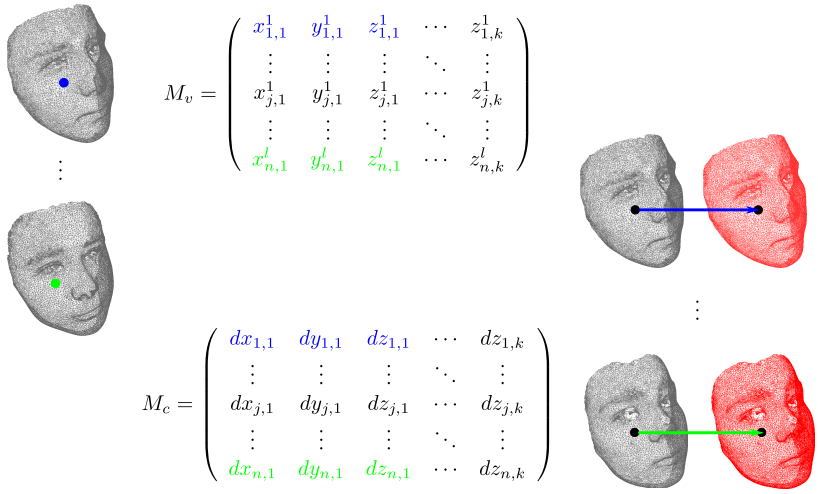


Figure 3.34: The difference between analyses of shape variability and analyses of stage changes and corresponding matrices.

change $\overline{X_c}$ or to study individual vertex shift variances, although this would be too much information to interpret. For this reason, PCA is applied to the matrix M_c to obtain principal components v_i^c of the changes manifested in the sample. These *principal stage changes* can be used to model a particular specimen's second stage

$$X_j^2 = X_j^1 + \overline{X_c} + \sum_{i=1}^{n} \alpha_{j,i}^c v_i^c$$

An approximation of the second stage can be created using a pre-computed mean shape of the first stage

$$\widetilde{X_j^2} = \overline{X^1} + \overline{X_c} + \sum_{i=1}^{n} \alpha_{j,i}^c v_i^c$$

The coefficients $\alpha_{j,i}^c$ have a similar meaning to $s_{i,j}$. In Equation 3.18, they represent coordinates of the specimen within *a space of stage changes*. The space is oriented so that the first coordinates allow the largest orthogonal differences to be captured.

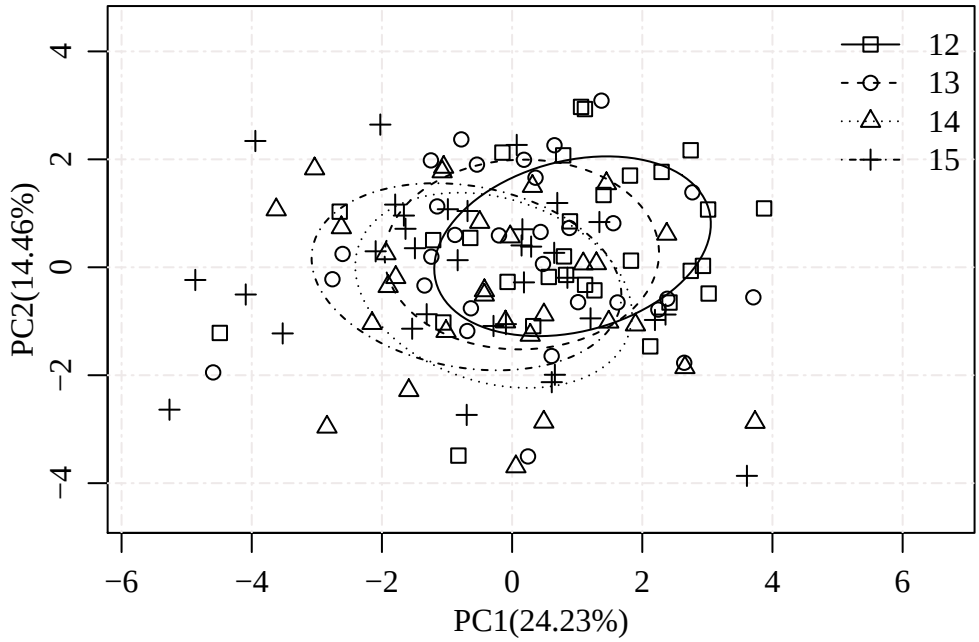
Case study 11: Development of sexual dimorphism in adolescent face morphology

The method described above was used by Krajíček *et al.* (2013) to analyze a longitudinal collection of human faces. In parallel, the same data were analyzed by Koudelová *et al.* (2015) with an emphasis on searching for sexual dimorphism in forms (shape representation with sizes preserved). The faces were collected between 2009 and 2012 by annually scanning approximately 60

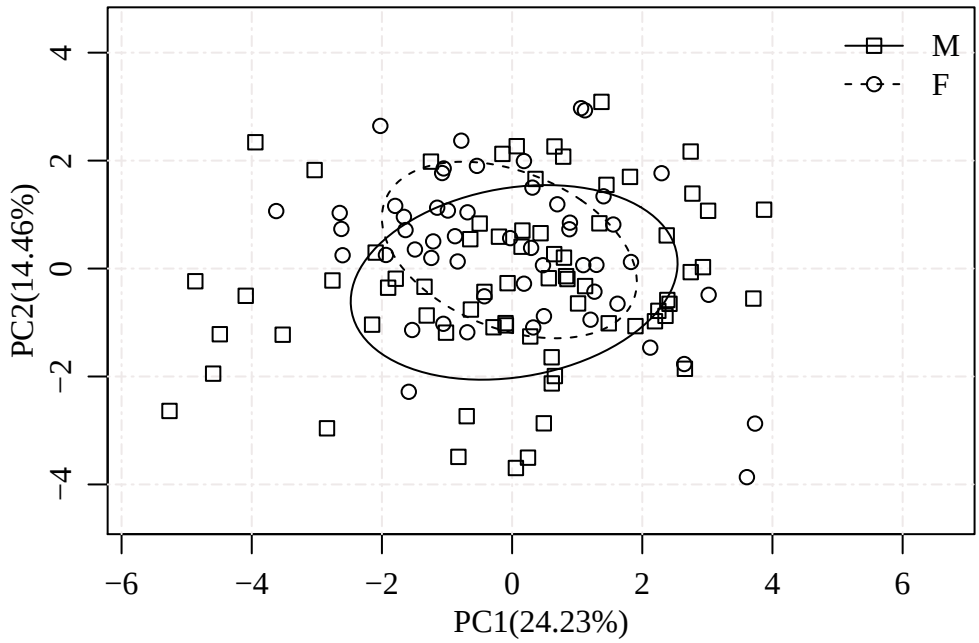
children in two schools in Prague and Kladno. The scanning started when the children were roughly about 12 years of age and continued up until the ages of 13, 14 and 15 years (see Figure 3.33). The scanning was performed in order to capture the continuous development and growth of the sample group. Data acquisition was performed with a Vectra 3D scanner (see Figure 3.1(b)) and manually pre-processed using 3D editing software to remove any scanning artifacts and to reduce the mesh density to a reasonable number of vertices per mesh for the analysis. The final meshes contained about 20k vertices. Sets of nine landmarks were placed on each mesh by an expert. The locations of the landmarks are also shown in Figure 2.2(b). The goal of the study was to find out at what point and in what way the shapes of the children's faces would start to exhibit sexually dimorphic traits.

In the first experiment, all age groups are put together and a shape variability analysis is performed by searching for correspondences across the combined sample and by subsequently applying PCA to the correspondences. Plotting a projection of the shape space (see Figure 3.35), which represents the scores of the first two principal components, records changes within age groups and determines whether there are any common features that they follow while aging. As shown in Figure 3.35(a), the centroid of the age groups slides towards negative parts of x-axes and moderately towards negative parts of y-axes with increasing age. What is not so visible is the slight increase in the size of the age group as the subjects grow older. This observation is consistent with the idea that as the specimen has more time to grow it has more time to reach a wider variety of shape configurations. Similar behavior is shown in the plot colored according to sex (Figure 3.35(b)). The male group exhibits larger variability because the male specimens are significantly larger in general and also have a better chance of reaching the larger part of the shape space than the female group. The groups move away from each other at the ends of the ellipses, which represent the shapes of older faces. In other words, growing differences are observable between sexes in this longitudinal population sample.

Equation 3.18 can be used to generate shapes that correspond to arbitrary points in the shape space. The shapes at extrema of x- and y-axes of the plot are generated and shown in Figure 3.35 in order to give a better understanding of the features that are represented by principal components. The features are complex and manifest themselves in various places in the mesh and are best seen on face profiles. Figure 3.36(a) shows that the first principal component is connected with the size of the nose and the size of the chin. The negative direction of PC1 is related to the larger nose and larger chin, a feature related to adult face shape, while the positive direction represents the more child-looking features. The second principal component (Figure 3.36(b)) is also connected with the size of the nose and additionally with the shape of the forehead. The negative PC2 direction represents a receding forehead which is a strong masculine feature, while the positive direction is more advanced which represents a more feminine feature.

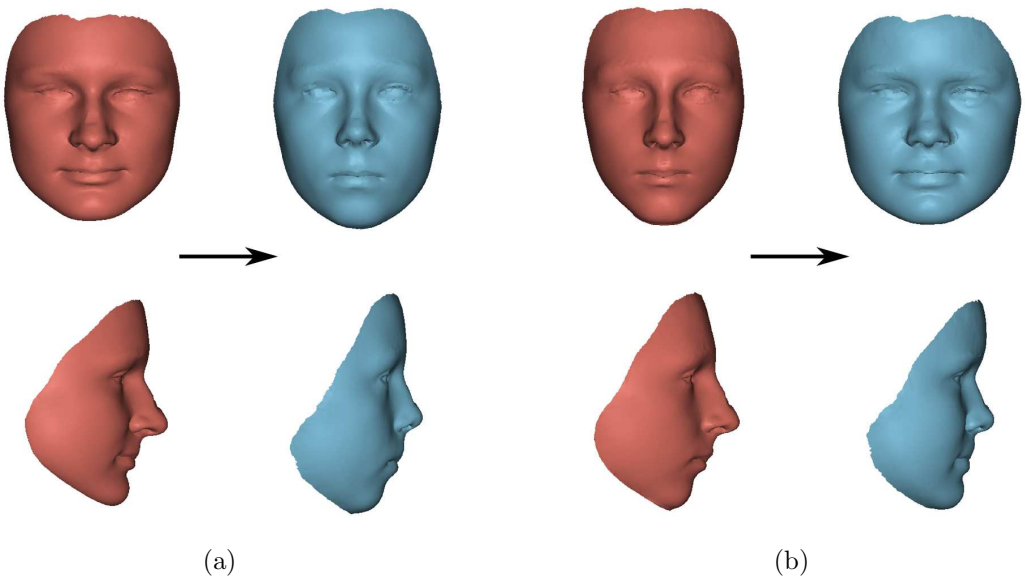


(a)



(b)

Figure 3.35: Shape variability of a combined dataset represented by the first two components: (a) grouped by age; (b) grouped by sex. The 50% ellipses approximate the shape of the population.



(a) (b)

Figure 3.36: Interpretation of shape features represented by the first two principal components of shape variability: (a) change with increasing PC1; (b) change with increasing PC2.

Another visible feature in Figures 3.36(a) and 3.36(b) is that the PC1 and PC2 control the overall width of the face: PC1 from a masculine width to a narrow width and PC2 from a thin oval, almost triangular to a wide round width. To conclude, interpreting the features is no simple task since individual PCs cannot be easily identified by a single anatomical feature. To that end, it could be useful to rate particular anatomical features in order to try to find a specific basis in the shape space.

A certain trend is found in the face shape changes in aging children, either identified as male or female. In order to prove in which age group the differences become statistically significant, an individual analysis is performed for each age group and the resulting PCA scores are tested with a multivariate permutation test for group equality between male and female subgroups. According to these results (shown in Table 3.9), the differences become statistically significant at the age of 15.

The previous experiment ignored the fact that the age groups were actually longitudinal data. Neither did it exploit the relation between the specimen stages contained in the groups. Therefore, in the second experiment, the analysis of stage changes is performed as described above. Figure 3.37(a) shows projections of the *stage change space* for 12 year-old to 13 year-old children. Although the plot is difficult to interpret, a trend is observable since the groups do not overlie it in the most ideal way. The sample size is also rather small so the difference in stage changes between the sexes is not statistically significant (see Table 3.10).

age group	# **specimens**	# **PCs**	**p-value**
12	32	6	0.3245
13	32	6	0.6215
14	32	5	0.1425
15	34	5	0.0380*

Table 3.9: Statistical significance of shape differences between males and females in various age groups from combined dataset analysis.

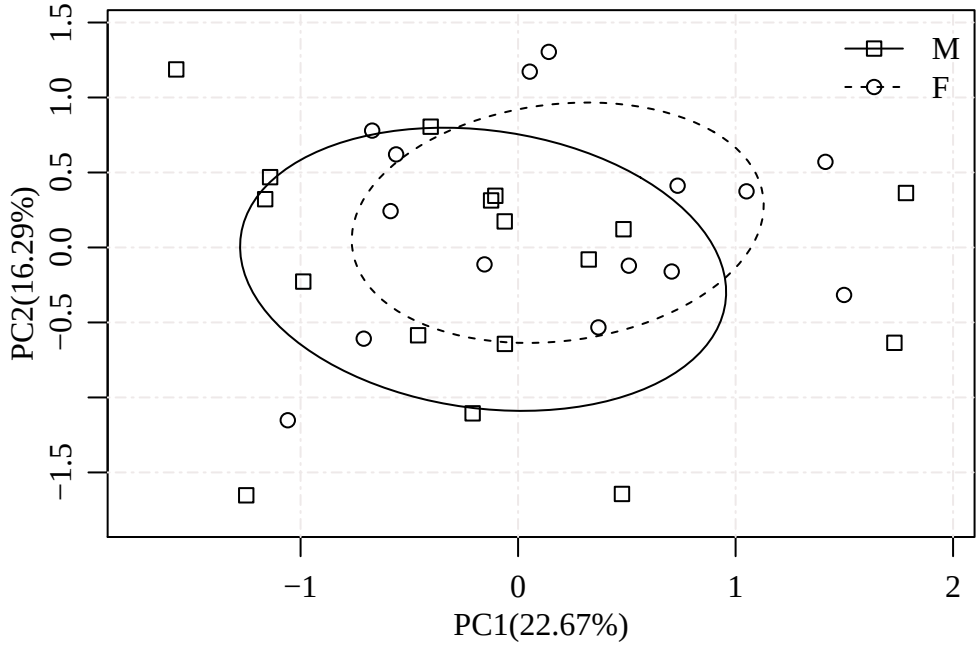
stage change	# **specimens**	# **PCs**	**p-value**
12-13	32	5	0.0961
12-14	32	5	0.0328*
12-15	32	5	0.0005****
13-14	32	5	0.0536
13-15	29	6	0.0145*
14-15	29	6	0.4496

Table 3.10: Statistical significance of differences between males and females in various stage change groups.

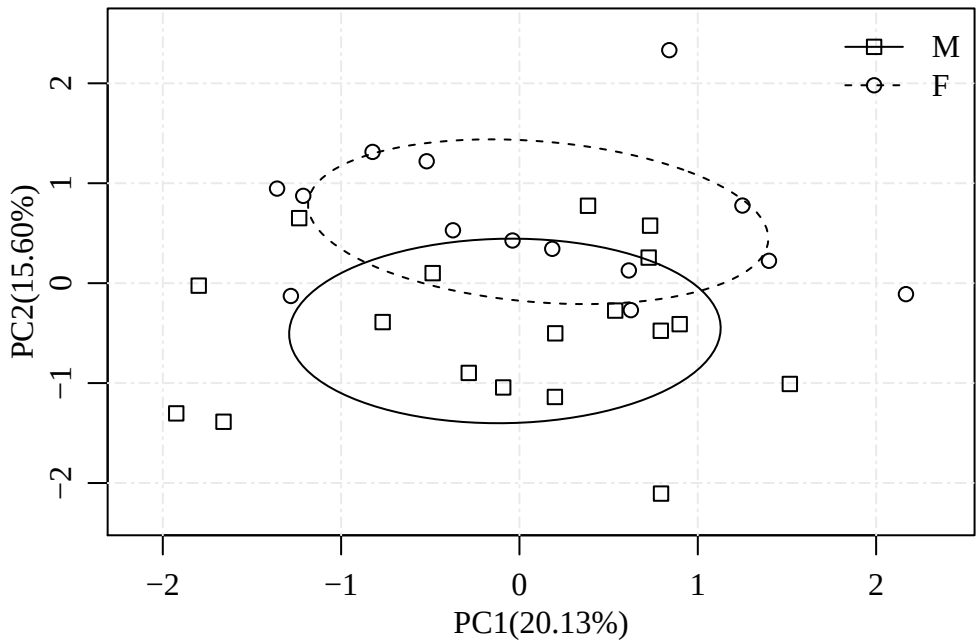
On the other hand, Figure 3.37(b), shows visibly separated groups, particularly along the x-axis. The first principal component can be interpreted in a similar way to shape variability analysis, i.e. it is connected with overall face width (see Figure 3.38).

Unfortunately, the projected data do not seem to be normally distributed and there are outliers, especially along the y-axis. Such results manifested in scatterplots often indicate errors in the data (e.g. switched landmarks, misplaced landmarks, etc.). Figure 3.38(b) shows that the y-axis is connected with shape changes along the jawbone towards the ears. In this case, the problem stemmed from improperly manually pre-processed data, i.e. the border of the meshes did not go through the same anatomical area in all the meshes. This may have been caused by excessive facial hair which the scanner failed to capture.

However, this particular error is isolated in the second principal component. The third principal component is the most prominent in the forehead area (see Figure 3.39(b)). Statistically significant differences between males and females were already found in the changes between the 12- and 14 year-old groups (see Table 3.10). This observation indicates that growth in this period is different for boys and girls.

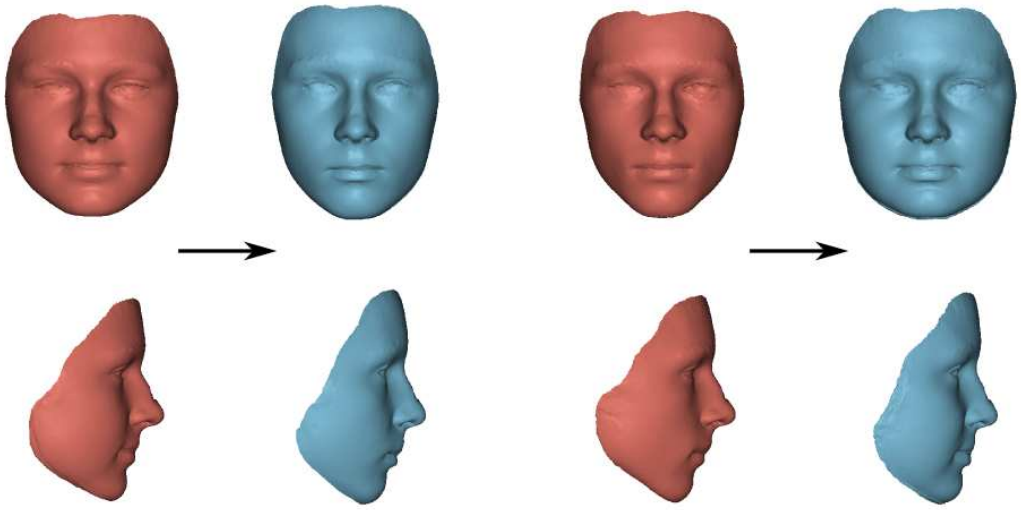


(a)



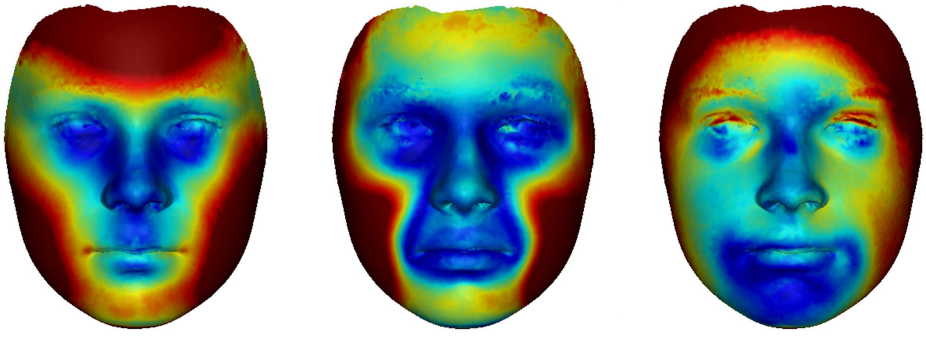
(b)

Figure 3.37: Variability of changes between stages: (a) between 12 and 13 years of age; (b) between 12 and 15 years of age. The 50% confidence ellipses approximate the shape of the population.



(a) (b)

Figure 3.38: Interpretation of paired analysis for stage changes between 12- and 15-year olds. (a) Increasing PC1 is connected with overall widening/narrowing of the face. (b) Increasing PC2 is connected with sexual differences, but also with a posterior jaw area change caused by improperly trimmed borders of the meshes.



(a) (b) (c)

Figure 3.39: Color-coded principal components influence magnitudes of stage changes between 12- and 15 year-olds: (a) PC1; (b) PC2; (c) PC3 — representing 20.13%, 15.60% and 10.30% in the variability of stage changes, respectively.

3.7 Missing data computation

A statistical shape model collects general information about the shape of a particular class of objects from the training sample. This information can be diverted back to recover the shape of an actual specimen based on much sparser or incomplete geometry. In other words, parameters of the model can be sought to optimize the relation to available data, similar to the computation of missing landmarks introduced in Section 2.3. An interesting and beneficial application of this approach is the approximation of missing geometry. The accuracy of this type of approximation is guaranteed by the amount and quality of the shapes the model is based on, as well as on the amount of the missing area.

Approximation of missing geometry could be used in various situations. For instance, imperfect scanning procedures might cause part of the mesh to be damaged. One limitation of scanning technology is that it does not allow parts of the object with particular material and structural properties to be captured, e.g. facial hair causes difficulty for laser as well as photogrammetric surface scanners. Facial scans of people with moustaches, beards or bushy eyebrows contain holes or very rough geometry. In other cases, parts of the geometry can be absent simply because the object of interest is incomplete at the time of scanning. However, it would be interesting to see how the object might look based on an approximation using a large sample of objects of the same kind. There is a large number of potential applications in the fields of forensic science and biomedicine connected with shape prediction and reconstruction.

Various methods deal with statistical modeling of shape to satisfy external conditions. For example, Blanz and Vetter (1999) attempted to reconstruct a 3D shape from a single two-dimensional image by fitting a statistical model of a textured triangular mesh to an imaginary one. This was achieved while optimizing the texture, pose and shape parameters as well as material and lighting properties of an illumination model (i.e. ambient, diffuse, specular reflection constants, light direction and intensities). To handle the computational complexity, the authors chose stochastic gradient descent to optimize the cost function, relating the rendered model to the input image. However, the study did not evaluate the precision of the reconstructed geometry. The radiometric properties of the actual object, along with the environment and the illumination model of choice (a Phong illumination model in this case) made it more difficult to precisely approximate the exact shape. Incorporation of multiple two-dimensional views into the computation can also improve approximation of the shape.

Blanz *et al.* (2004) solved fitting of the statistical model to the 3D mesh in order to reconstruct the shape, even with the missing parts. The approach assumes that there is a correspondence between the model $\mathbf{y}$ vertices and the

present vertices of the target mesh $\mathbf{x}$ that are to be fitted.

$$\mathbf{y} = \bar{\mathbf{v}} + \sum_{i=0}^{n} s_i \sigma_i \mathbf{v}_i$$

Parameters $\mathbf{s} = \{s_i\}_{i=0}^{n}$ of the model are searched for in order to minimize cost function E. Vectors $\mathbf{v}_i$ are normalized eigenvectors and σ_i is a variance of the individual coefficient for the members of the training set. These variances are equal to corresponding eigenvalues. As mentioned above, eigenvectors are ordered with respect to the amount of information they represent, which is equal to the variance and magnitude of the respective coefficients, i.e. the coefficients of the first components are higher in value and variation then the coefficients with higher indices. To normalize the space of the coefficient unit, eigenvectors must be scaled by respective variances.

$$\begin{aligned} E &= \|SV\mathbf{s} + \bar{\mathbf{v}} - \mathbf{x}\|^2 \\ &= \|Q\mathbf{s} + \bar{\mathbf{v}} - \mathbf{x}\|^2 \end{aligned} \tag{3.21}$$

where V is a matrix with columns representing components of a linear model (eigenvectors) and S is a diagonal matrix with variance, i.e. $S_{i,i} = \sigma_i$ and $Q = SV$. Minimization of E is achieved by solving the following equation

$$\begin{aligned} 0 &= \nabla E = 2Q^T Q\mathbf{s} - 2Q^T(\bar{\mathbf{v}} - \mathbf{x}) \\ 2Q^T Q\mathbf{s} &= 2Q^T(\bar{\mathbf{v}} - \mathbf{x}) \end{aligned} \tag{3.22}$$

Since V is a matrix with orthonormal columns and $V^T V = I$, the situation simplifies to

$$\mathbf{s} = (VS)^T(\bar{\mathbf{v}} - \mathbf{x})$$

Moreover, vector $\mathbf{s}$ computed this way has a useful property. Although $\mathbf{s}$ is one of the possible solutions to equation 3.22, it is the solution that has minimal norm $\|s\|$, which means that it is as close to average as possible, i.e. it has maximum prior probability.

Such a solution overcompensates in modeling the data at the expense of the missing parts, which might not benefit from the generalizing property and thus the approximation might not be as successful as expected. Blanz *et al.* (2004) improved the model by adding a regularization term to cost function 3.21.

$$E = \|Q\mathbf{s} + \bar{\mathbf{v}} - \mathbf{x}\|^2 + \eta\|\mathbf{s}\|^2$$

where $\|\mathbf{s}\|^2$ is the Mahalanobis distance of the configuration from the average. The amount of regularization is controlled by parameter η, even though the solution of 3.22 is still rather simple

$$\begin{aligned} \mathbf{s} &= (S + \eta S^{-1})^{-1} V^T(\bar{\mathbf{v}} - \mathbf{x}) = S^* V^T(\bar{\mathbf{v}} - \mathbf{x}) \\ S^*_{i,j} &= \begin{cases} \frac{\sigma_i}{\sigma_i^2 + \eta} & i = j \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

Another problem of fitting the model to an arbitrary mesh is that the model only embodies geometrical variation in shape, but the particular shape representation also includes pose properties, i.e. translation, orientation and scale in space. Only the shape variation needs to be modeled, not the shape pose using column vectors of V. The second improvement by Blanz *et al.* (2004) was to include additional terms for y that describe pose information of the particular shape. These terms represent translation, rotation and scale in a natural way without breaking the linearity of E.

$$\begin{aligned}\mathbf{y} &= \bar{\mathbf{v}} + \sum_{i=1}^{n}(\lambda s_i)\mathbf{v}_i + (1-\lambda)\bar{\mathbf{v}} + \\ &\quad s_{tx}\mathbf{v}_{tx} + s_{ty}\mathbf{v}_{ty} + s_{tz}\mathbf{v}_{tz} + s_\gamma\mathbf{v}_\gamma + s_\theta\mathbf{v}_\theta + s_\phi\mathbf{v}_\phi\end{aligned}$$

Translation is modeled by $\mathbf{v}_{tx}$, $\mathbf{v}_{ty}$, $\mathbf{v}_{tz}$ allowing arbitrary shifts of $\mathbf{y}$ along the axes.

$$\begin{aligned}\mathbf{v}_{tx} &= (1,0,0,1,0,0,\ldots,1,0,0)^T \\ \mathbf{v}_{ty} &= (0,1,0,0,1,0,\ldots,0,1,0)^T \\ \mathbf{v}_{tz} &= (0,0,1,0,0,1,\ldots,0,0,1)^T\end{aligned}$$

Scaling is allowed by multiplying the model with scaling factor $\lambda \approx 1$ which is equivalent to adding $1-\lambda$ times the mean shape $\bar{\mathbf{v}}$ and obtaining scaled coefficients λs_i. Rotation compensation for small angles is estimated by adding vectors that approximate multiplication by general rotation matrix R

$$R = R_\gamma R_\theta R_\phi = \begin{bmatrix} c_\gamma c_\theta - s_\gamma s_\phi s_\theta & -s_\gamma c_\phi & c_\gamma s_\theta + s_\gamma s_\phi c_\theta \\ s_\gamma c_\theta - c_\gamma s_\phi s_\theta & c_\gamma c_\phi & s_\gamma s_\theta - c_\gamma s_\phi c_\theta \\ c_\phi s_\theta & s_\phi & c_\phi c_\theta \end{bmatrix} \approx \begin{bmatrix} 1 & -s_\gamma & s_\theta \\ s_\gamma & 1 & -s_\phi \\ s_\theta & s_\phi & 1 \end{bmatrix}$$

where $s_\gamma = \sin(\gamma)$, $c_\gamma = \cos(\gamma)$, etc. Using a simplified rotation matrix, the vector rotation can be expressed as a linear combination of four vectors

$$R\begin{bmatrix} x \\ y \\ z \end{bmatrix} \approx \sin\gamma \begin{bmatrix} -y \\ x \\ 0 \end{bmatrix} + \sin\theta \begin{bmatrix} 0 \\ -z \\ y \end{bmatrix} + \sin\phi \begin{bmatrix} z \\ 0 \\ -x \end{bmatrix} + \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

The operation can be applied to the whole shape vector $\bar{\mathbf{v}}$

$$\begin{aligned}R_v\bar{\mathbf{v}} &= s_\gamma\mathbf{v}_\gamma + s_\theta\mathbf{v}_\theta + s_\phi\mathbf{v}_\phi + \bar{\mathbf{v}} \\ \mathbf{v}_\gamma &= (-\bar{y}_1, \bar{x}_1, 0, -\bar{y}_2, \bar{x}_2, 0, \ldots)^T \\ \mathbf{v}_\theta &= (0, -\bar{z}_1, \bar{y}_1, 0, -\bar{z}_2, \bar{y}_2, \ldots)^T \\ \mathbf{v}_\phi &= (\bar{z}_1, 0, -\bar{x}_1, \bar{z}_2, 0, -\bar{x}_2, \ldots)^T\end{aligned}$$

The method described above assumes that there is an ideal point-to-point correspondence as an input. Blanz *et al.* (2004) applied the method to compute

missing parts of a tooth crown while using an ad-hoc registration procedure. The algorithm works with manually placed landmarks on a two-dimensional height map in two subsequent refinement passes. The study evaluated RMS of the reconstructed mesh (43% of the missing area) and original surface, which was approximately 100 μm on shapes that were 1 cm in diameter, i.e. 1% relative error.

Mueller *et al.* (2011) presented a very interesting application of the above model-from-photograph approach. The issue presented in the paper is a computation of the shape and texture of a part of the human face for use in the creation of a natural-looking prosthetic appliance for a patient after a complete rhinectomy and resection of a large section of the facial area. As an input, the authors used a database of 200 3D facial scans, a photograph of the patient before the illness affected his face, and a 3D scan of his current face.

Another application of surface statistical modeling is face recognition framework presented by Bustard and Nixon (2010). Apart from face geometry, the authors included ear geometry as well because the ear is a very complex feature and differs from specimen to specimen. Therefore, it is able to increase the classification success rate significantly. The proposed framework also incorporated a number of manually placed landmarks, as well as initial rigid registration and optimization-based fitting.

Model fitting

A new algorithm presented here further combines the useful points from all the mentioned approaches. The goal is to use a statistical model constructed from a set of faces using non-rigid B-spline registration and to fit it to the mesh with missing parts. The main difference to the approach by Blanz and Vetter (1999) and Blanz *et al.* (2004) is that the algorithm does not construct correspondences prior to the fitting of the model. Instead, the model is optimized to fit the data while correspondences are computed every time they are needed for the evaluation of the minimized cost function.

$$
\begin{aligned}
E(\mathbf{s}) &= \sum_i w(\|\mathbf{y}_i^{\mathbf{s}} - \mathbf{x}_i\|)((\mathbf{y}_i^{\mathbf{s}} - \mathbf{x}_i) \cdot n(\mathbf{x}_i))^2 + \eta\|s\|^2 \qquad (3.23)\\
\mathbf{y}^{\mathbf{s}} &= \bar{\mathbf{v}} + \sum(\lambda s_i)\mathbf{v}_i + (1-\lambda)\bar{\mathbf{v}} + \\
&\quad s_{tx}\mathbf{v}_{tx} + s_{ty}\mathbf{v}_{ty} + s_{tz}\mathbf{v}_{tz} + s_\gamma\mathbf{v}_\gamma + s_\theta\mathbf{v}_\theta + s_\phi\mathbf{v}_\phi
\end{aligned}
$$

Points $\mathbf{x}_i$ corresponding to points $\mathbf{y}_i^{\mathbf{s}}$ of the model are searched for based on the nearest neighbor principle with the use of acceleration and approximation data structures described in Section 3.2.4 [page 52], namely the k-d tree, octree, grid cache, tri-linear interpolation and extrapolation outside of the grid cache.

The pose optimization of E with respect to $\mathbf{t} = \{s_{tx}, s_{ty}, s_{tz}, \lambda, s_\gamma, s_\theta, s_\phi\}$ is optional. It depends on the possibility and accuracy of pre-registration, by using landmarks or other variants of rigid registration of partially missing data to a model mean shape.

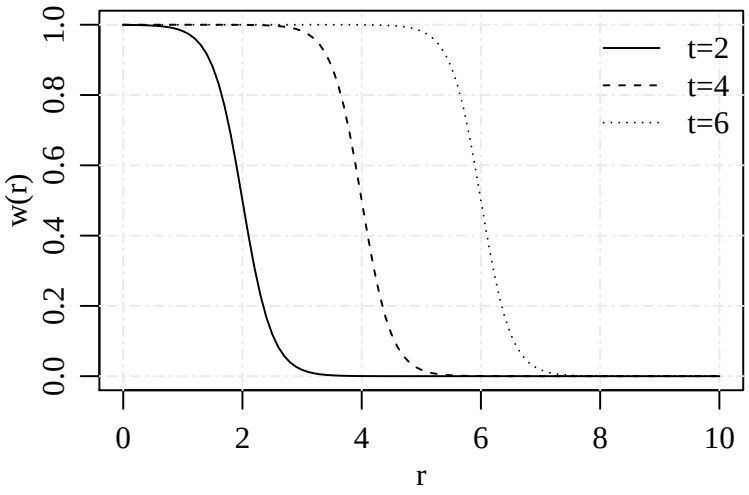


Figure 3.40: The model-fitting weighting function distinguishes correspondence from the missing geometry.

The weight of a particular point depends on three criteria. The first is the squared distance between model point $\mathbf{y}_i^{\mathbf{s}}$ and nearest neighbor $\mathbf{x}_i$.

The second criterion is the distance weighting function w which should cancel out points that are too far to be considered as corresponding, or which are more likely to be missing in the modeled surface mesh. The specific function w used in the computation is as follows

$$w(r) = \frac{1}{1 + \exp 4(r - t)}$$

where parameter t approximately reflects a distance, beyond which the correspondence is discarded and discontinued in order to attract model vertices. Figure 3.7 shows the behavior of the suggested weighting function for various values of parameter t.

The third criterion is normal weighting, i.e the squared cosine of an angle between the line segment connecting $\mathbf{y}_i^{\mathbf{s}}$ and $\mathbf{x}_i$ and the normal of the surface at point $\mathbf{x}_i$. The rationale behind the dot product is that the points are more attracted to the corresponding points on the surface of the modeled mesh in a perpendicular direction to the surface than in a direction parallel to the surface. Figure 3.41 shows how distance weighting and a normal-weighting influence model fitting in a particular example.

The last improvement is multi-stage fitting. The idea of fitting a PCA model to the data is based on preferring the first few components while damping the higher order components, since the first components represent more general shape features and the last components model noise. This is partially done by scaling eigenvectors $\mathbf{v}i$ by respective eigenvalues. However, the optimization procedure can still use noise components to improve the value of

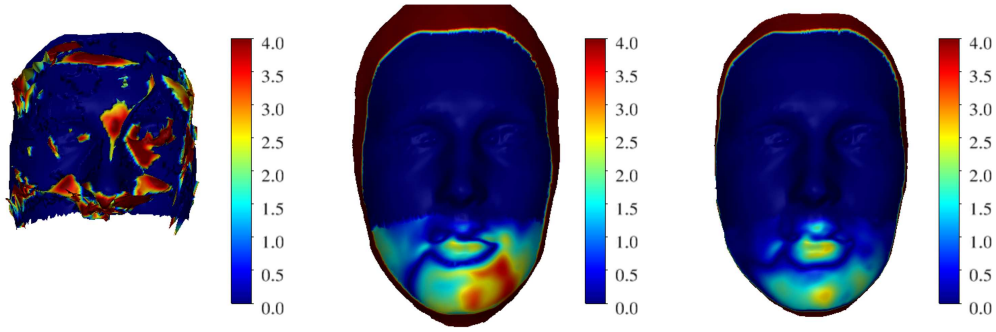


Figure 3.41: Effects of cost function components on model-fitting results displayed as a deviation of the modeled mesh and the original: (a) no normal and distance weighting; (b) distance weighting only (t=8); (c) distance and normal weighting in a single optimization stage.

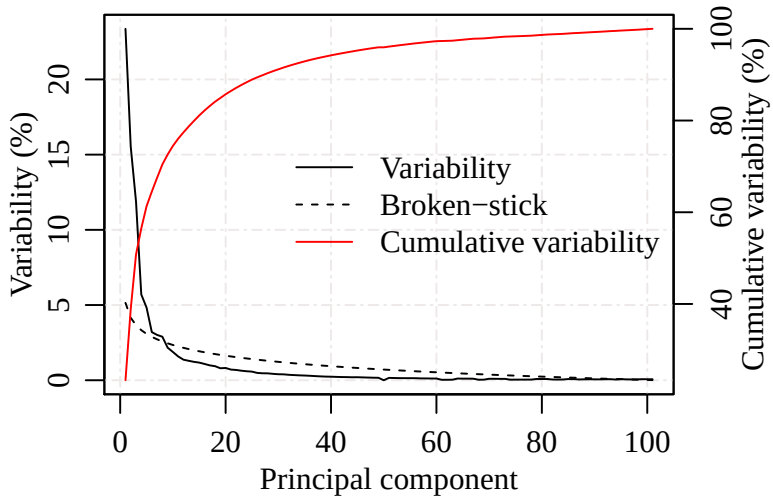


Figure 3.42: The model components' amount of variability in the face dataset.

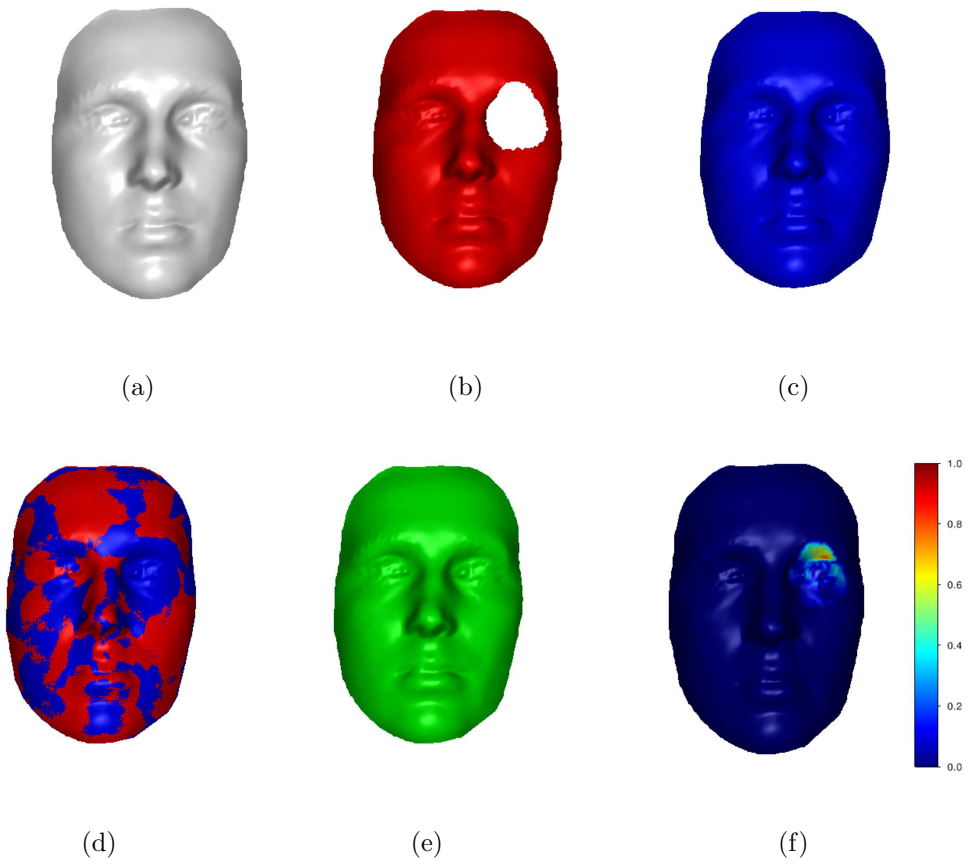
Algorithm 6: Optimization based missing geometry model fitting

Input : model components V
mean shape $\bar{\mathbf{v}}$, mean shape landmarks $\mathbf{k}$
missing data mesh $\mathbf{x}$, missing data mesh landmarks $\mathbf{l}$
number of components per level $\mathbf{c}$
Output: fitted model $\mathbf{y}$
begin
  pre-align $\mathbf{x}$ with $\bar{\mathbf{v}}$ (e.g. with use of landmarks)
  initialize $\mathbf{s}$ and $\mathbf{t}$ with zeros
  for $i \leftarrow 1$ **to** *number_of_levels* **do**
    $\mathbf{s}, \mathbf{t} \leftarrow$ Optimize Equation 3.23 initialized by $\mathbf{s}, \mathbf{t}$ with respect to $s_1, \ldots, s_{\mathbf{c}[i]}$ and $\mathbf{t}$
  end
  $\mathbf{y} = \bar{\mathbf{v}} + \lambda V \mathbf{s}^T + V_t \mathbf{t}^T$
end

the cost function. The optimization can be divided into multiple phases (2 or 3), in which the optimization procedure converges with a limited number of model components which are increased in a subsequent phase. A number of phases and a number of components used in each phase can be determined experimentally. A certain estimate can be made as to the variability distribution on model components. Figure 3.7 shows such a distribution for the face dataset. The multi-stage optimization of the PCA model is analogous to the multiresolution approach used in B-spline registration. Similarly, it could speed up the convergence as well as lower the chance of getting stuck in local minima caused by the overfitting of noise components.

For model optimization, local numerical optimization by L-BFGS is used, which requires computation of partial derivatives of E with respect to $\mathbf{s}$ and $\mathbf{t}$. Due to nearest neighbor correspondences constructed differently for different values of $\mathbf{s}$, E is not strictly continuous. However, for properly pre-aligned missing data meshes and multi-phase convergence, only a small area of the parameters is searched with low probability of large discontinuities. Partial derivatives are computed numerically by central differences, but an analytical solution is still possible. An overview of fitting the model to the missing mesh data is shown in Algorithm 6.

After the model is fitted to the data, as much of the original geometry as possible should be retained; only parts of the model should be used in place of missing geometry, where detected. Model vertices are considered missing from the original mesh if the distance to the nearest neighbor in the original mesh exceeds a certain threshold (e.g. 1mm). The missing geometry in the original model is replaced with the geometry of the model by adding missing vertices and triangles. Additional triangles are added by triangulating a seam between newly added vertices and the vertices of the original geometry using



(a) (b) (c)

(d) (e) (f)

Figure 3.43: Example of computing a missing shape based on a statistical model: (a) original mesh; (b) altered mesh with approximately 6.5% of the surface missing; (c) model fitted to the altered mesh; (d) overlap of (b) and (c); (d) altered shape with missing parts is fixed with parts of the fitted model shape; (e) shape deviation of (a) and (e). The scale is in millimeters.

Algorithm 7.

The presented algorithm was tested on a face dataset and the model was analyzed in Section 3.4.1. Three phases of optimization were used with 10, 20 and 100 components, respectively. A particular example of missing data computation is shown in Figure 3.43.

Additional five test cases were created by removing various parts and various amounts of the face geometry. The specific results of the computation are visualized in Figure 3.44. Results of the evaluation are also presented in Table 3.11.

According to the results, the quality of missing data recovery depends on the area which is missing more than the amount of the missing data. For example, if the paired feature is removed (case 1), the model fits the remaining data so that the recovered part does not match its mirror counterpart opti-

Algorithm 7: Recover missing geometry from model mesh

```
Input  : missing data mesh x, model mesh y
Output: stitched mesh y'
begin
    x' = x
    b = {}
    // add all distant triangles and collect border vertices
    foreach triangle t in y do
        if ∀v ∈ t, d(v, x) < th then
            add triangle t to x'
        else
            foreach vertex v in t do
                if d(v, x) < th then
                    add v to b
                end
            end
        end
    end
    // triangulate the seam between the original and border
       vertices
    foreach vertex v in b do
        o = nearest(x, v)
        n(v) = {w|w ∈ b ∨ (v, w) ∈ y}
        foreach vertex n in n(v) do
            no = nearest(x, o)
            if no == o then
                add triangle (v, n, o) to x'
            else if (no, o) ∈ x then
                add triangle (v, n, o) to x'
                add triangle (n, no, o) to x'
            else
                p = path from no to o
                for i ← 1 to len(p) − 1 do
                    add triangle (p[i], p[i + 1], v) to x'
                end
                add triangle (v, n, o) to x'
            end
        end
    end
end
```

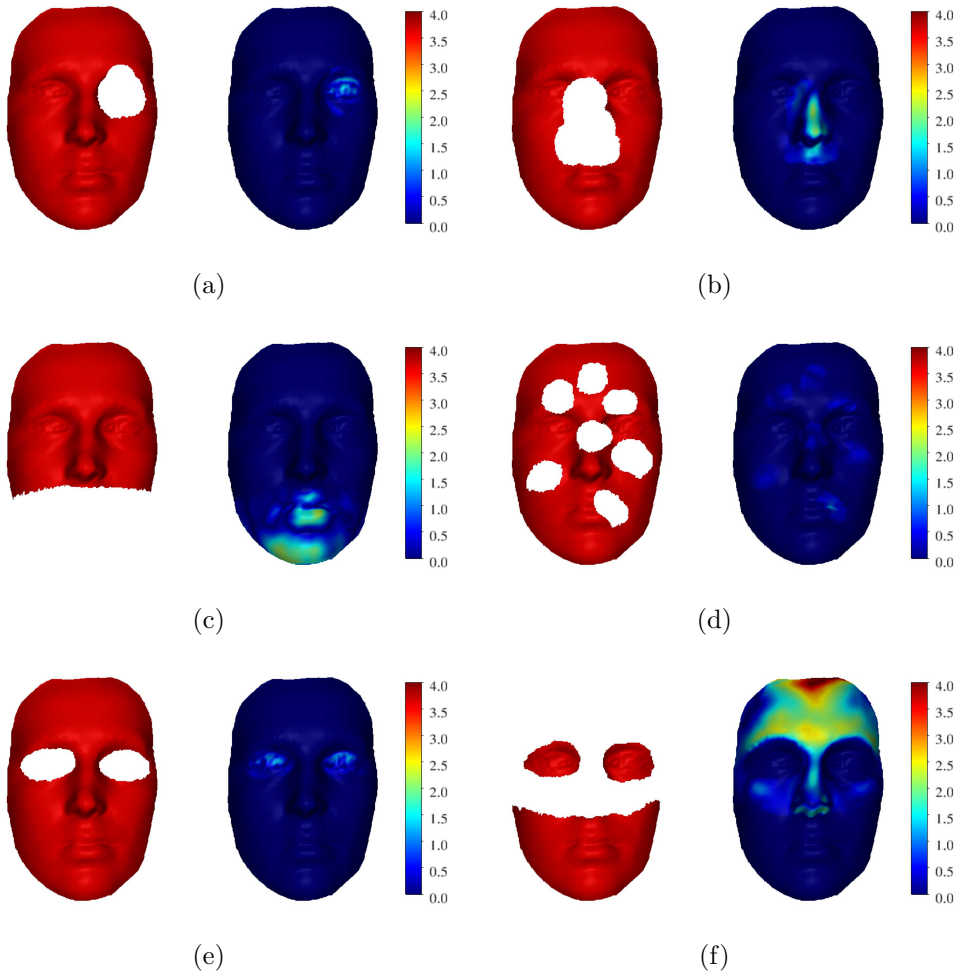


Figure 3.44: Comparison of the results of missing part computation with different missing parts and different amounts of missing areas: (a) one eye missing — paired feature; (b) nose missing — non-paired feature; (c) jaw missing — large area; (d) random holes — large area, but sparsely distributed; (e) eyes missing — both paired features; (f) large part of upper face missing — "unmasking Batman".

mally, causing higher mean deviation. But if both paired features are missing (case 5), the recovered geometry is symmetrical and mean deviation is lower, although the total missing area is larger. If the missing geometry is distributed relatively uniformly (case 4), its recovery is rather successful and mean deviation is very low, provided the missing parts are concentrated in one place in the object (case 2, 3). In all cases, no explicit regularization was used ($\eta = 0$). Employing regularization might become more important when a larger part of the surface is missing.

case no.	surface area (mm^2)	preserved area (%)	mean dev. (mm)	max. dev. (mm)
1	28279	93.5	1.297	9.618
2	25163	83.2	1.178	9.638
3	21632	71.5	1.174	8.338
4	24423	80.8	0.692	8.448
5	27757	91.8	0.888	7.607
6	13822	45.7	1.528	11.85

Table 3.11: Evaluation of missing geometry computation on a mesh with various parts and different amounts of missing areas. Mean deviation is computed only in the area of the recovered geometry.

3.8 Conclusions

In this chapter, basic tools for statistical shape analysis, represented by defining a boundary with a triangular mesh, were introduced. First, the measuring devices were discussed. Next, the properties of the data they provide and the procedures that should follow before entering the statistical shape analysis pipeline were also examined.

The most important step in shape analysis and comparison is mapping correspondences from one shape to another. For this purpose, various rigid alignment and non-rigid registration procedures were described. The main focus was on the following methods:

- Procrustes superimposition — landmark-based registration which minimizes the sum of squared distances. This method is essential in applications where precise landmarks can be defined by the user and is more important than automatic processing.
- Iterative closest point — an automatic procedure for rigid alignment in its basic variant. It can be used in cases where no landmarks can be specified and as a pre-processing step for more elaborate non-rigid matching.
- TPS-based registration — global non-rigid deformation defined by hand-placed landmarks.
- B-spline non-rigid registration — this method has great potential to tightly align very diverse meshes while incorporating various criteria characterized by permitted deformations and correspondences. Many novel improvements have already been described with the aim of improving speed and quality, e.g. evaluation of sampling approaches, use of specialized search and caching data structures, evaluation of numerical methods, etc. The author contributed by modifying the method

originally used for volume data registration to fit surfaces as well as by speeding up convergence with sampling approaches.

- Coherent point drift — this method differs from others by incorporating one-to-many correspondence while computing the regularized drift of one point set to the other. The author contributed by participating in improving the convergence rate of CPD by employing spatial clustering as a sampler using the Nyström method for low-rank matrix approximation (Dupej *et al.*, 2014a).

The second part of the chapter illustrated the utilization of registration methods in statistical shape analysis and its application to various research tasks. Most of the tasks were vital parts or actual research projects from areas of physical anthropology and biomedicine.

- Mean shape construction — mutual correspondences allow a mean mesh to be constructed. However, the process bears the disadvantage of averaging, i.e. details are smoothed out and the result does not look natural.

- Shape variability and sexual dimorphism of the human face — the superiority of mesh shape analysis over landmark-based methodology was demonstrated on higher discrimination rates in sexual dimorphism by nearly 15% in a TPS registration-based model. However, mesh registration is more effective when using a B-spline-based algorithm and outperforms TPS by 7% (see Table 3.6).

- Allometry — drawing the dependency of a shape represented by a triangular mesh in size. The author contributed to the application of the described method for dense correspondence model construction and the analysis of a human face with respect to allometry and sexual dimorphism (Velemínská *et al.*, 2012).

- Variability of the healthy human palate and the palates of BCLP and UCLP patients were analyzed with author's contribution (Bejdová *et al.*, 2012a; Rusková *et al.*, 2014). B-spline registration offers more accurate mean shapes than TPS.

- Post- and pre-treatment palatal shape variability of patients undergoing the SARME procedure in comparison to a healthy population sample. The author contributed by performing statistical shape analysis in a biomedical study (Trefný *et al.*, 2015).

- Mesh asymmetry analysis — the author contributed by developing a methodology for statistical analysis of shape asymmetry and its demonstration on human face asymmetry in a young adult sample from the Czech population (Krajíček *et al.*, 2012).

- Face asymmetry of UCLP patients — the author applied the methodology in biomedical research (Cagáňová *et al.*, 2014).

- Shape morphology of an annually scanned sample of children from 12 to 15 years of age with emphasis on the development of sexually dimorphic traits. The author contributed by proposing a method for paired analysis while applying it to longitudinal data (Krajíček *et al.*, 2013; Koudelová *et al.*, 2015).

- Missing data computation — a new method with high practical potential for filling holes in incorrectly scanned meshes and approximate missing data due to partially destroyed objects.

One possible avenue of further research in this area is the application of the above-described methods to analyze many other interesting phenomena in physical anthropology and biology in general. However, in order to obtain statistically reasonable conclusions, a sample of a certain size is required, which is not always possible for many interesting case studies. Furthermore, collecting multiple scans is a tedious, expensive and time-consuming process, involving the work of many other researchers.

From a methodical point of view, future work in the area of statistical triangular mesh analysis might benefit from the inclusion of the following topics:

- Incorporating texture into statistical analysis to extend the statistical analysis beyond examining just shape properties, while also including texture features. However, adding more information about the individual limits the ability to draw statistically reasonable conclusions from the sample size presented in this work, which requires a larger sample.

- Searching for precise correspondences in multiple-related scans, which might be useful for analyzing differences between groups in multiple phases of development.

- Modeling the relation between various parts of the meshes might help to uncover relations that are not apparent when analyzing whole shapes. Alternatively, it could be used when the sample is not large enough to describe the variability of detailed features. This idea could be extended to reflect multi-scale shape analysis.

Chapter 4

Statistical analysis of volume images

> Every man is a volume if you know how to read him.
>
> William Ellery Channing

For decades, volume imaging has served as an important tool for many disciplines including radiology, material science, and biochemistry. It allows capturing the volume of space and storing the information in grids, where each cell represents a single volume element of the space (a voxel). The type or resolution of the grid, as well as the type of information stored in the voxel, depends on the particular technology used. As such, volume imaging is a valuable source of information about shape, especially shapes of very complex objects that cannot be captured by surface scanning.

Specific scanning devices are specialized for specific types of objects, sizes, resolutions, and compositions. These constraints are even more limiting in the case of volume imaging. In general, for most types of objects there exists a solution for volume imaging that meets requirements. Nevertheless, the availability of the devices had traditionally been a problematic issue. Due to the high penetration of new devices in various institutions, they become accessible even for purposes they were not originally meant to.

Tomography is a common approach used for capturing regular 3D grids slice-by-slice. Each slice is a raster image containing an isometric pixel, where the distance between slices is different from the size of the slice pixels. The term *computed tomography* (CT) usually indicates a tomographic approach, which is combined with ionizing radiation as a physical phenomenon, enabling non-destructive imaging (Suetens, 2009).

A volume image is basically a regularly sampled space, which can be inspected via various volume visualization techniques, e.g. direct volume rendering (DVR) or iso-surface rendering. It is also possible to use volume data as a source for landmark-based morphometry. Similarly, triangular surfaces,

representing boundaries of an object in the sampled space, can be extracted using algorithms such as marching cubes. However, both approaches lead to a great reduction in information.

Computed tomography, as mentioned above, is suitable for capturing bone tissue. However, if applied to a living specimen where bone tissue is tightly surrounded by other types of tissues, suitable pre-processing steps need to be performed in order to obtain readable inputs for subsequent analysis. These pre-processing steps include mainly image enhancement and segmentation, which are discussed in the first part of this chapter. The main focus is then aimed at correspondence construction using non-rigid registration. In the last part, a statistical model is created based on correspondences found in the data sample and its particular application is described.

4.1 Volume segmentation

For the complex task of morphometric analysis of shapes captured in volume data, segmentation is an essential pre-processing step for subsequent measurements. Volume data measurements can cover various modalities, for example, landmark placing, contour extraction, surface extraction, surface area computation, volume computation and corresponding point matching. These measurements can also be influenced by the presence of the background.

Image segmentation, after image registration, is the second most important task in medical image processing. However, providing a full overview of current state-of-the-art segmentation methods is beyond the scope of this thesis. The reader should refer to existing surveys by Pham *et al.* (2000); Hu *et al.* (2009); Khan (2013). The following sub-sections describe two methods for segmentation of volume data acquired by computed tomography and their application to particular tasks. Both algorithms are based on deformable models and intensity priors: the first uses a B-spline curve for shape representation; the second evolves from the first by employing a more sophisticated representation and segmentation criterion.

4.1.1 Contour-based 2.5D segmentation

Computed tomography in radiology produces images of human body interiors which can be as poorly readable as X-Rays to the untrained eye. On the other hand, CT images offer much more information in the form of calibrated 3D data (position and density), organized into slices in an axial plane of the human body (see Figure 4.1). Thanks to these features, they can be used not only to gain a rough idea of the human interior, but also to obtain precise measurements of volume.

Physical measurements of the human body can provide a lot of information about human health. For example, the temperature of a healthy body is about 36.5°C. Having a higher temperature means that the body is fighting

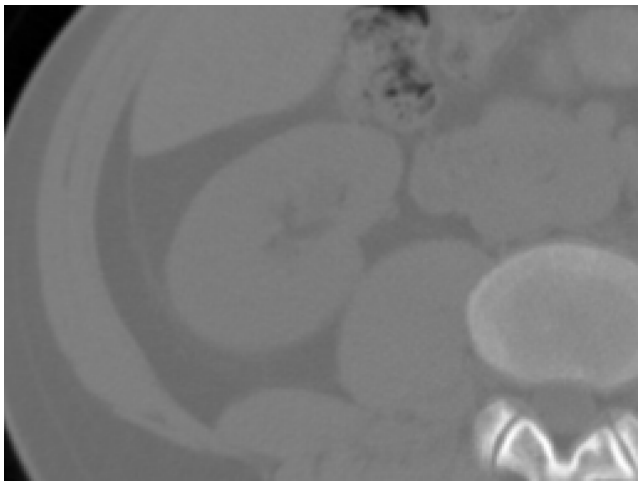


Figure 4.1: Slice of CT data from the abdominal area.

an illness. The same goes for volumes of organs. Noticeable changes in kidney volume in either an upward or downward direction indicate proper function or dysfunction. These changes can only be observed after a long period between CT scannings and also depend on correct and objective evaluation of CT images.

Krajíček *et al.* (2007) have made an effort to develop methods for fast, accurate and automatic kidney segmentation, which are a precondition for measuring volume. The nature of CT data (noise, partial volume effect, beam hardening, low contrast) as well as the variety of the human body make this task rather difficult.

Related work

Deformable models have been used and studied in computer vision for almost two decades. Their best-known application is called 'active snakes', introduced in the late 80's by Kass *et al.* (1988). This is based on evolving curves towards a lowest total energy value, defined mostly by the image gradient and some regularization properties.

Unfortunately, medical images are very noisy and have unreliable gradient information. Chan and Vese (2001) have introduced an image energy term which depends on region statistics only

$$\begin{aligned} E_{region1}(s) &= \alpha \int_{int(s)} (f(x,y) - \mu_{ext}(s))^2 dxdy \\ &\quad + \beta \int_{ext(s)} (f(x,y) - \mu_{ext}(s))^2 dxdy \end{aligned} \tag{4.1}$$

where s is a closed curve, f is an image and μ is a mean intensity of exterior and interior areas that are to be separated. In this case, μ can be known a

priori or computed from an initial position and updated during evolution.

Another improvement replaces the classical 'snake' curve representation with a set of points called 'snaxels' using a closed B-spline curve. The B-spline curve as a perfect tool for contour detection was introduced by Brigger *et al.* (2000). This approach reduces the number of optimized parameters to coordinates of control points only. Using B-spline implicitly brings the benefit of an intrinsic regularization property. B-spline snakes along with a very efficient computational scheme are presented in the work of Jacob *et al.* (2004).

Energy model

In the algorithm presented here, the flexibility of the snake framework is utilized, into which new terms are incorporated. The energy model is primarily based on regions while the contours defining them are represented by the B-spline curves mentioned above.

In the first attempt, the energy functional is extended by an additional term which describes a relation to the adjacent slices in the volume dataset. This relation is based on a shape similarity measure. The following measure is used to reduce the unwanted behavior of the evolving curve

$$E(s) = E_{region1}(s) + \gamma \cdot similarity(s, s_{neigbour}) \tag{4.2}$$

The idea is to find how to express local high-frequency changes of a shape in the similarity measure. This criterion can indicate overflow of the contour to an adjacent area through 'bridges' of similar intensity. These bridges should be narrow enough to be rendered as high frequency. High frequency can raise the similarity term in the energy functional (Equation 4.2). The minimization scheme can isolate and eliminate them because of their significant differences. A shape-describing vector is constructed as the Fourier power spectrum of the boundary orientation angle changes. The similarity of two shapes is equal to l_2-distance of their shape vectors. It is used as *similarity* term in Equation 4.2. Lee *et al.* (2003) have shown that this similarity measure is sensitive to significant differences in the corresponding parts of curves.

Chan and Vese's region energy scheme (Equation 4.1) incorporates the square distance from the mean, which is a very rough approximation of inverted probability. Instead of this, the probability of element classification to the exterior or interior region is directly used according to Jacob *et al.* (2004). Furthermore, weight coefficients α and β are simplified to α and $(1 - \alpha)$, expressing the balance between the exterior and interior. However, it is only of little importance in the new scheme

$$\begin{aligned} E_{region2}(s) &= -\alpha \int_{int(s)} log(P_{int}(f))dxdy \\ &\quad -(1-\alpha) \int_{ext(s)} log(P_{ext}(f))dxdy \end{aligned}$$

The probability in Equation 4.3 can be approximated by normal distribution based on analysis of the initial shape position. The algorithm completely relies on the user's initial input and uses normalized and optionally smoothed histograms of areas (interior, exterior) as the probability distributions. Other interesting region-based energy models can also be used, introduced by Yezzi *et al.* (1999).

$$\begin{aligned} E_{region3}(s) &= -\frac{1}{2}(\rho_{int}^2 - \rho_{ext}^2)^2 \\ E_{region4}(s) &= -\frac{1}{2}(\mu_{int}^2 - \mu_{ext}^2)^2 \end{aligned}$$

Minimization itself can be achieved using the gradient descent method. Partial derivatives of functional E by parameters $\mathbf{x}$ of the contour are the only terms to compute.

$$\mathbf{x_{n+1}} = \mathbf{x_n} + \delta \cdot \nabla_{\mathbf{x}} \mathbf{E}(\mathbf{s}(\mathbf{x}))$$

These parameters are, in this case, identical to the control points of the B-spline curve. The most important feature of region-based snake segmentation is the ability to reduce computational complexity, from computing an area integral to computing a curve integral, using Green's theorem.

The approach of Chan and Vese (2001) is not completely followed and no more than two types of energy functional mentioned above are used. It encounters the problem of energy equivalence which cannot always be solved by parameters tuning to full satisfaction. The question of how much region energy is equivalent to a unit of gradient-based edge energy, length or area energy (in terms of linear relation) cannot be simply answered.

B-spline curves

B-spline curves are well-known splines that use polygonal base functions (see Section 3.2.4) with limited support, which gives them great local control properties. There are a few problems that should be mentioned.

The first is that B-splines must be used in all energy functionals in place of contour s. If the points on the contour and their displacements are used during the minimization step, only the locations of control points need to be optimized further on. This is a significant advantage because they are fewer in number than the contour points, even when sampling of the contour is very sparse. The order of B-spline is another question. Cubic B-splines are well-suited for the purpose because of their shape variability and relatively low computational complexity.

The number of control points or number of B-spline sections depends on the overall length of the closed contour. Minimal and maximal lengths of a segment are defined along with additional parameters. During the iterations the lengths are checked: if the limits are exceeded, the particular segment is split or merged with adjacent segments as necessary.

The last problem relates to intersections. Sometimes an evolving curve happens to be self-intersecting. For example, if a shape with holes is segmented, the curve could wrap around the hole and cycle forever. One solution consists of detecting intersections and breaking off parts of the curve. Free ends of the larger part of the curve will be bound together and the rest can be thrown away. Alternatively, the rest of the curve can also be used as an inner structure segmentation. This can be seen when segmenting higher or lower parts of a kidney that has the pelvis in the middle (the pelvis has very different tissue density). It can be ignored for the first time and subsequently segmented out by simple thresholding. Intersection detection can use the convex wrapper property of B-spline segments for computation speed-up, i.e. to check convex hulls defined by control points for mutual collisions before checking B-spline arcs.

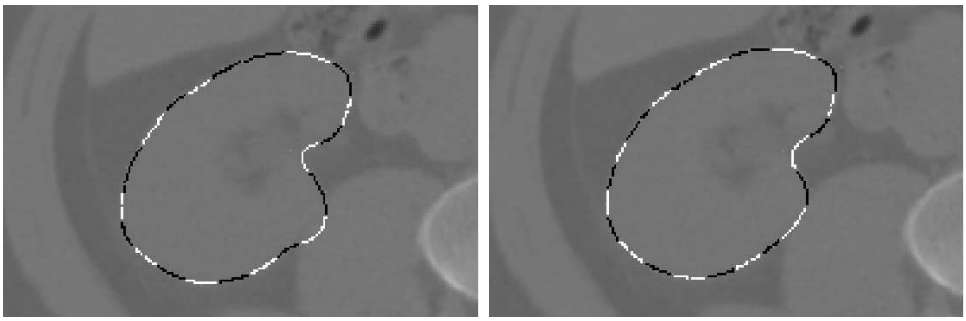
Jacob *et al.* (2004) used integration of angle changes which give various multiples of 2π depending on the number and direction of the loops. This method detects loops that implicate self-intersections, but is ineffective in the cases described above.

Segmentation procedure

In order to segment 3D volume, the described 2D segmentation method is used in each slice. A shape similarity measure is used as an inter-slice relation which binds adjacent slices together. The reasons for preferring this scheme (compared to a pure 3D deformable model technique like Active Surfaces) are that it consists of anisotropic data and exhibits computational efficiency, which is far better than in the case of full 3D techniques. Due to the properties of CT data, datasets almost always contain a voxel with one dimension larger than the other two (anisotropy). The 2.5D approach offers an almost interactive response and better user control (better then editing 3D surface control points).

The segmentation procedure starts with the user specifying top and bottom end points of a segmented organ. Optionally, the user can set other points in slices between the top and bottom ends, which can help to determine starting positions. Starting positions on other slices are placed automatically between those set manually by the user. An initial shape is defined as the small circle-like shape around these points. The probability distribution is estimated from the area inside these circles. It is important to put them in the right place inside the segmented organ. The user can optionally supply probability distribution from a manually segmented shape in one slice.

The iteration starts after shape initialization and placement on each slice. In each iteration of the procedure, one step of snake energy minimization is performed. Thanks to this scheme, the whole object changes its shape during the segmentation process and thus can be observed and possibly corrected by the user. After one of the slices converges, it becomes a steady point for its neighbors and available for use as a similarity term in Equation 4.2. Neighbors of previously converged slices are restricted by the similarity in their further



(a) (b)

Figure 4.2: Converged B-spline snakes: (a) results of converged B-spline snakes excluding similarity forces; (b) the same example of converged snakes including similarity forces using the parameter $\gamma = 0.3$.

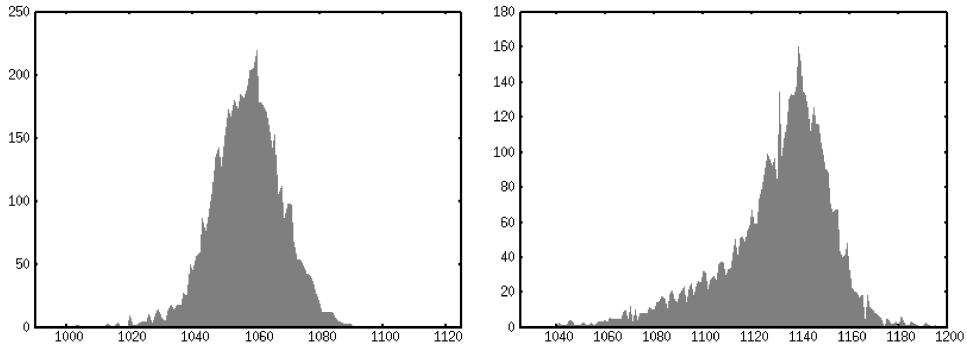
evolution. This scheme works well for rounded organs with homogeneous intensity whose axial slices do not change topology and shape variation too rapidly.

Results

Convergence of the basic model (Equation 4.2) is strongly dependent on parameters α and γ. If supplied with proper probability distribution, the model (Equation 4.3), delivers very promising results. A proposed improvement with a similarity measure, which can help to hold the curve in the correct shape, is shown in Figure 4.2.

The algorithm was tested on native images of a kidney check-up (without a contrast agent) and also with images of the same check-up with a contrast agent in several phases of saturation. In each case, different strategies for parameter settings and probability distribution are required to obtain optimal results. Native images work well with approximation by normal distribution. But in the case of thin patients lacking contrast in inter-organ fat, it is almost impossible to find a distribution which prevents overflowing to neighboring organs without losing too much tissue on the surface of the segmented organ. Using this scheme on a slightly saturated kidney, it is possible to segment the whole organ and satisfactorily determine its volume. Figure 4.4 shows final results of the segmentation. The approximation by normal distribution fails on a highly saturated kidney because it does not correspond to the reality, as shown in Figure 4.3. On the other hand, fine-tuned probability distribution works well for the majority of patients.

The speed of the segmentation procedure is satisfying. Processing of a five-millimeter thick slice dataset takes about 80-90 seconds (an intersection check is performed after every 40 iterations; convergence is achieved after about 400 iterations). A two-millimeter slice dataset takes 160 seconds, but



(a) (b)

Figure 4.3: Histograms of segmented shapes: (a) native CT corresponds to normal distribution; (b) the image with medium satiation of the contrast agent demonstrates slightly skewed distribution.

the measurement is more accurate. Time measurements are performed on a single core computer (AthlonXP 2600+, 2 GB RAM).

Subsequently, the algorithm is modified in two ways (Krajíček, 2007). First, the simultaneous convergence of individual slices is changed to continuous propagation from both ends of the organ. The converged shape in the current slice initializes the shape in the next adjacent slice. Meanwhile, the following slice segmentation is much faster which speeds up the whole algorithm for almost one order of magnitude. The overall segmentation time of the organ then takes approximately 15-17 seconds. Second, propagation from both ends can run independently which allows a multi-core system to be utilized. The modified algorithm was tested on a multi-core system (AMD Athlon 64 X2 Dual Core 3800+) and achieved 1.55x speed-up over a single core implementation (for more details, see Krajíček (2007)).

The results of the segmentation algorithm were compared to manual segmentation performed under controlled conditions. The average error of volume measurement was -0.68% with standard deviation of 2.15% in a collection of 12 datasets.

4.1.2 Locally adaptive level-set segmentation

The segmentation algorithm introduced in Section 4.1.1 has two major disadvantages. First, due to representation by surface contours, it is unable to segment objects with complex topologies, such as branching veins, objects with holes, and cavities. Second, it depends on global intensity characteristics. In the following text, an algorithm that confronts both of these problems is described. It is primarily designed to segment teeth and jawbone structures in volume images intended for anthropological research.

Automatic teeth segmentation and 3D reconstruction from CT images of

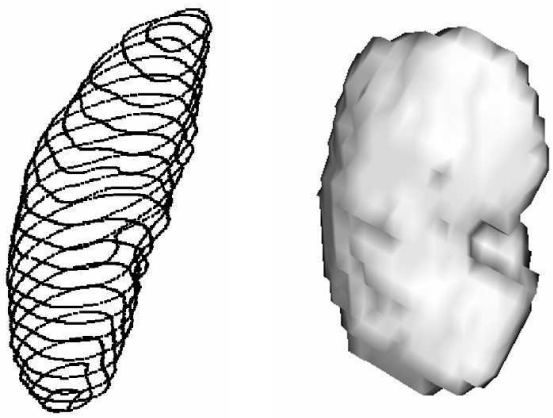


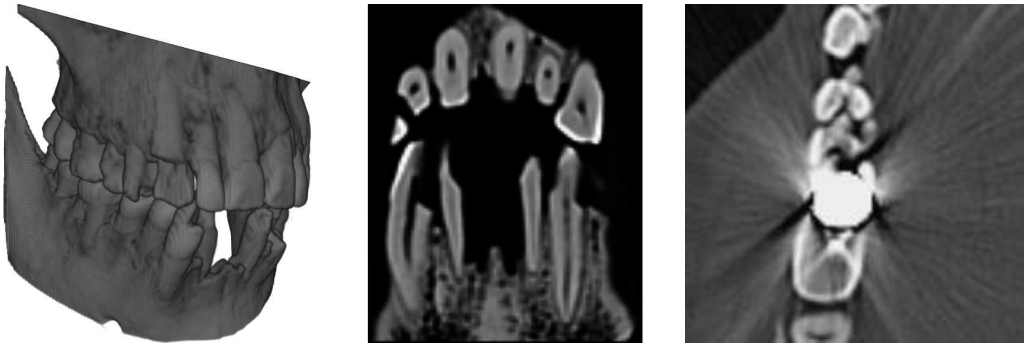
Figure 4.4: A schematic model (left) and a 3D model (right) of a segmented kidney.

patients have recently been addressed in a number of studies. Earlier approaches exploited the nature of data slices being orthogonal to the direction of tooth growth and also used the similarity of tooth profiles in adjacent slices. It is a simple process to initialize 2D segmentation of a tooth in the slice by segmenting the tooth from the previous slice, similar to the approach described in Section 4.1.1. Either a two-dimensional B-spline snake (Heo and Chae, 2004) or a variational level-set (Hosntalab *et al.*, 2008) is used for 2D segmentation. The 2D method has been improved by introducing a region competition principle that keeps adjacent teeth separated while simultaneously segmenting them (Gao and Chae, 2010). This assumes that the segmentation of neighboring teeth is initialized at approximately the same distance from their boundary.

Semi-automatic region-growing segmentation combined with thresholding in 3D has also been successfully used by Akhoondali *et al.* (2009). Concurrent with tooth volume data segmentation, there have also been attempts to segment range images captured with a surface scanner by Kondo *et al.* (2004).

One of the major problems of teeth segmentation is that tooth tissue varies widely in intensity as it consists of the hardest bone tissue in the body (enamel) as well as softer tissues (dentine, cementum). It also contains cavities filled with pulp which are usually completely missing from raw bone data. Moreover, the jaw tissue surrounding teeth has a density very close to the tissue of teeth themselves. This makes segmentation difficult for algorithms based purely on global area intensity priors. The relatively low resolution, partial volume effect, connected adjacent teeth, and jawbone have also proven difficult for algorithms based solely on edge information (see Figure 4.5).

Recent research in the area favors level-set based segmentation methods, introduced by Malladi *et al.* (1995), due to their versatile ability to segment arbitrary shapes from images of various modalities and to incorporate various



(a) (b) (c)

Figure 4.5: Input data examples: (a) raw jawbone volume; (b) a slice through a tooth with raw bone data showing a wide range of pixel intensity inside the tooth tissue; (c) an example of a tooth image demonstrating artifacts, caused by the medical condition of a recent population and particular configuration of the imaging device.

criteria depending on the given problem. The only significant disadvantage of level-sets might be their long processing times, caused by the iterative nature of the segmentation algorithm.

Level-set segmentation is based on solving a variational problem of energy minimization of a certain functional $E(\varphi)$ by looking for an optimal shape border. The functional $E(\varphi)$ incorporates particular constraints on the result of the segmentation. The shape border Γ is represented implicitly by using a function φ such that

$$\Gamma = \{x | \varphi(x) = 0\}$$

The function φ is negative for x lying inside the segmented area, i.e. belonging to the object, and positive for x lying outside, in the background. The variational problem $\arg\min_{\varphi} E(\varphi)$ is solved by introducing an artificial variable t and describing its behavior with respect to t by a partial differential equation (PDE)

$$\frac{\partial \varphi(x)}{\partial t} = -v \cdot |\nabla \varphi|$$

Therefore, the function φ is initialized as a signed distance function. Term v represents the speed and direction in which the border moves in a time step. Some research presents experiments with various methods in an attempt to solve this type of PDF effectively; while other methods deal with specific definitions of v in actual applications. Both problems are handled in the following text.

The Chan-Vese model

The Chan-Vese model by Chan and Vese (2001) is a very successful level-set segmentation scheme that relies solely on average area intensity priors.

$$\begin{aligned} E(c1, c2, \Gamma) &= \lambda_1 \int_{\Omega_f} (f(x) - c_1)^2 dx + \lambda_2 \int_{\Omega_b} (f(x) - c_2)^2 dx \\ &\quad + \mu \int_{\Gamma} dx + \nu \int_{\Omega} dx \end{aligned}$$

Parameters λ_1, λ_2, μ and ν are the weights that are tuned for specific tasks, i.e. prioritizing one region over another, the minimal length of the border, or the area of the foreground. Parameters c_1 and c_2 are the average intensities of the background and foreground. Their values are given a priori; either estimated from the initial segmentation or continuously recomputed during the evolution. For the latter case, c_i must be continuously dependent on φ to have a stable solution. In order to do that, regularization is added to the computation of c_i.

$$c_1(\varphi) = \frac{\int_\Omega f(x) H(\varphi(x)) dx}{\int_\Omega H(\varphi(x))}, c_2(\varphi) = \frac{\int_\Omega f(x)(1 - H(\varphi(x))) dx}{\int_\Omega (1 - H(\varphi(x)))} \tag{4.3}$$

Term H is the Heaviside function. $H(x) = 0$ for $x < -\epsilon$, $H(x) = 1$ for $x > \epsilon$ and for $x \in (-\epsilon, \epsilon)$ it smoothly transitions from 0 to 1.

Evolving φ to minimize $E(\varphi)$ leads to the flattening of φ which is no longer a signed distance function. This is solved by reinitialization of φ after a fixed number of iterations; however, the reinitialization interval may have a significant impact on the segmentation results. Another option, suggested by Li *et al.* (2005), is to include an additional self-regularization term in the functional $E(\varphi)$ which would penalize the deviation of $\nabla\varphi$ from 1. In fact, this only leads to the replacement of one parameter (a reinitialization interval) with another, which is the weight constant for the new regularization term.

Fast level-set segmentation

The key element of the presented approach is the fast level-set algorithm by Shi and Karl (2005), which replaced the PDE level-set segmentation based on the Chan-Vese model. It consists of a simple segmentation scheme that resembles region-growing-like propagation combined with border smoothing. The resulting algorithm is very efficient, but there is still room for improvement and for adaptation to specific applications.

The algorithm evolves an implicit function $\phi(x)$ by tracking the inner and outer borders of the segmented area, L_{in} and L_{out}. It alternates between an expansion phase, where it moves the border elements in n_a steps according to the external force F, and a regularization phase, where it smoothens the border elements in n_g steps by applying a smoothing filter on $\phi(x)$. Gaussian smoothing was used in the regularization phase of the original implementation. In Algorithm 8, the basic structure of the algorithm is summarized.

Algorithm 8: Fast level-set algorithm

Input : data D, seed S
Output: ϕ
begin
initialisation of L_{out}, L_{in}:
$L_{out} = \{x | x \notin S \vee \exists y \in N(x), y \in S\}$
$L_{in} = \{x | x \in S \vee \exists y \notin N(x), y \in S\}$
initialisation of $\phi(x)$, $\forall x \in I$:

$$\phi(x) = \begin{cases} 3 & \text{if } x \notin S \vee x \notin L_{out} \\ -3 & \text{if } x \in S \vee x \notin L_{in} \\ 1 & \text{if } x \notin S \vee x \in L_{out} \\ -1 & \text{if } x \in S \vee x \in L_{in} \end{cases}$$

while *not converged* **do**
for i **to** n_a **do**
for each $x \in L_{out}$: if($F(x) > 0 \vee cond(x)$): $checkin(x)$
for each $x \in L_{in}$: if($\forall y \in N(x), \phi(y) < 0$): L_{in}.remove(x), $\phi(x) = -3$
for each $x \in L_{in}$: if($F(x) < 0 \vee cond(x)$): $checkout(x)$
for each $x \in L_{out}$: if($\forall y \in N(x) \vee \phi(y) < 0$): L_{out}.remove(x), $\phi(x) = 3$
end
for i **to** n_g **do**
for each $x \in L_{out}$: if($filter(\phi) < 0$): $checkin(x)$
for each $x \in L_{in}$: if($\forall y \in N(x), \phi(y) < 0$): L_{in}.remove(x), $\phi(x) = 3$
for each $x \in L_{in}$: if($filter(\phi) > 0$): $checkout(x)$
for each $x \in L_{out}$: if($\forall y \in N(x), \phi(y) > 0$): L_{out}.remove(x), $\phi(x) = -3$
end
end
end

- The condition of optimality is defined as follows:
  $$cond(x) = \begin{cases} 1 & \exists y \in N(x), \phi(x)\phi(y) < 0 \vee F(x)F(y) > 0 \\ 0 & \text{otherwise.} \end{cases}$$

- $checkin(x)$ - x moves from the outer border to the inner border
  L_{out}.remove(x), L_{in}.insert(x), set $\phi(x) = -1$
  for each $y \in N(x), \phi(y) = 3$: L_{out}.insert(y), set $\phi(y) = 1$

- $checkout(x)$ - x moves from the inner border to the outer border
  L_{in}.remove(x), L_{out}.insert(x), set $\phi(x) = 1$
  for each $y \in N(x), \phi(y) = -3$: L_{out}.insert(y), $\phi(y) = -1$

Kratky and Kybic (2008) improved the scheme to simulate the behavior of the Chan-Vese level-set segmentation model by defining the external force:

$$F(x) = \begin{cases} -1 & \text{if } -\lambda_1(f(x)-c_1)^2 + \lambda_2(f(x)-c_2)^2 \geq 0 \\ 1 & \text{if } -\lambda_1(f(x)-c_1)^2 + \lambda_2(f(x)-c_2)^2 < 0 \end{cases} \tag{4.4}$$

where c_1 and c_2 are the expected average values of the foreground and background intensity. The parameters λ_1 and λ_2 are non-negative weights for the foreground and background elements, which can be interpreted as the foreground and background intensity variances. These values can be set experimentally or approximated in a pre-segmentation step. In conclusion, Kratky and Kybic (2008) also propose employing anisotropic filtering as a replacement for the isotropic Gaussian regularization phase.

Local level-set segmentation

A disadvantage of Chan-Vese segmentation is the use of c_1 and c_2 priors that hold for the background and foreground of the whole image. This is problematic in the case of images with intensity inhomogeneities, e.g. if a foreground object consists of tissues with a wide variety of densities. A tooth with surrounding jawbone is a typical example of such an image that contains varying foreground and background characteristics. This issue has been addressed in the following most recent papers by modifying the Chan-Vese model.

Wang *et al.* (2010) used the difference between an image and a Gaussian smoothed image to express a term for a local region-based criterion, which is added to the Chan-Vese energy functional. The approach was improved by Gao *et al.* (2012) using region initialization involving a pre-segmentation step based on k-means clustering. Another approach by Li *et al.* (2011) clustered the image into disjoint regions, in which intensity statistics for the region-based segmentation were measured. Local statistics can be computed for each pixel individually using some kind of averaging kernel, e.g. a Gaussian kernel (Wu and Yang, 2012). Zhang *et al.* (2013) reflected intensity inhomogeneity by modeling local average intensity with a bias field.

Local structure properties

Local structure properties in images are often used in various filtering techniques such as non-linear anisotropic filtering by Weickert (1998) and shock filtering by Osher and Rudin (1990); Vacavant *et al.* (2012). These methods describe image manipulation using PDF in a similar way to how level-set evolution works with a level-set function. Local information is usually constructed by evaluating the gradient, gradient magnitude, structure tensor T_ρ (Weickert, 1999, 2003), or Hessian matrix H of the image u (Descoteaux *et al.*, 2005).

$$T = T_0 = \nabla u^T \nabla u, T_\rho = T_0 * G_\rho \tag{4.5}$$

The eigenvalues and eigenvectors of tensors and Hessian matrices describe the intensity changes in orthogonal directions ordered by decreasing magnitudes. The eigenvalues λ_i ($\lambda_i > \lambda_j, i < j$) of tensors T, T_ρ or H can be combined to represent specific local behavior in the image such as sheet-like, tube-like, or blob-like appearances. The ratios of various eigenvalues in various image scales allow us to define the sheetness measure. Sheetness measure was used in image segmentation by Descoteaux *et al.* (2005) as a criterion for edge-based level-set segmentation of bones. Krcah *et al.* (2011) used a similar sheetness measure to estimate edge weights in a segmentation algorithm based on graph-cuts.

Local statistics processing

Tooth images are typical examples of inhomogeneous yet compact objects. They exhibit significant variations in intensity. Additionally, their global intensity statistics are burdened by a rather high variance and a mean intensity that is close to the mean of the surrounding tissue. This motivates the use of locally adaptive statistics for level-set segmentation. The local methods for level-set segmentation mentioned above often use one averaging kernel for the whole image, which has an isotropic spherical shape. Wider kernels provide greater stability against noise but are less adaptive to local intensity changes.

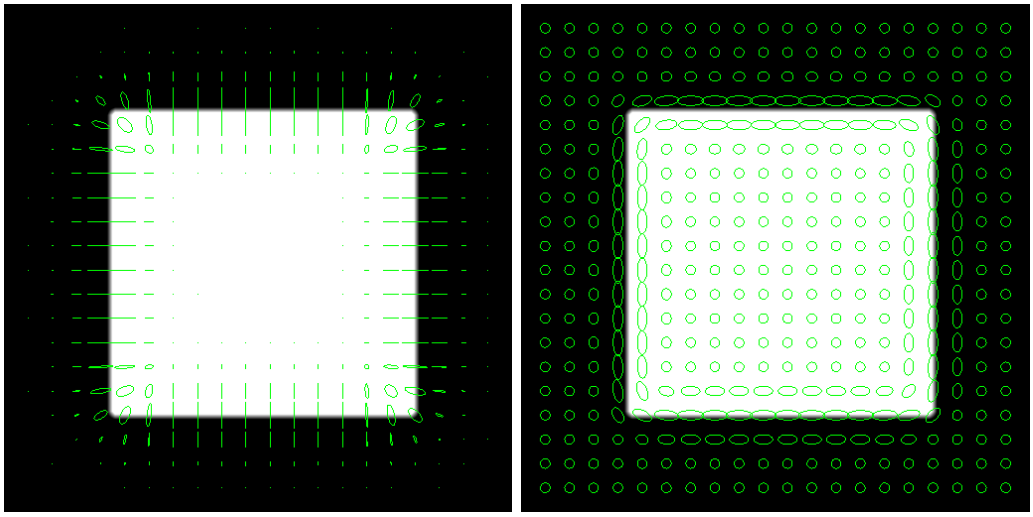
The shapes of the teeth are elongated and the individual teeth are separated from each other and the surrounding tissue by narrow gaps. A wide kernel can be very ineffective in describing the average intensity in the gap, as it will include the tissue behind the gap. Local structure information can be used to compute the shape of the kernel that adapts to the surrounding space.

In the first step, $T_\rho(x)$ of u_σ is computed, where u_σ is a smoothed version of the original image with a kernel of the width σ. Smoothing of the original image is performed to improve the stability of partial derivatives and, consequently, robustness to noise. The value of σ is usually set to around 1. Tensor T_ρ is a structure sensor, or *second-moment matrix*, and is constructed to describe the average shape structure in the neighbourhood of size ρ, which is usually chosen to be larger than σ. T_ρ is a symmetric real 3x3 matrix for 3D data or a 2x2 matrix for 2D data.

Subsequently, eigenvectors $w_1, ..., w_n$ of T_ρ ($n = 3$ for 2D or $n = 2$ for 3D data) are gathered and ordered with respect to decreasing corresponding eigenvalues $\nu_1, ..., \nu_n$. Eigenvectors w_i are used to construct a new diffusion tensor D with suitable properties. First, local coherence K is measured

$$K(x) = \sum_{i,j,i<j}^{n} (\nu_i - \nu_j)^2 \tag{4.6}$$

Second, the desired eigenvalues $\lambda_0, ..., \lambda_n$ of D are calculated and then used



(a) (b)

Figure 4.6: Structural information examples: (a) visualization of the smoothed structure tensor T_ρ, $\rho = 5$; (b) visualization of the diffusion tensor $D_{\alpha,c}(x)(\alpha = 0.3, c = 0.001)$. Major and minor axes of green ellipses represent the eigenvectors of the tensor at a particular point.

in combination with w_i to compute D

$$\begin{aligned} \lambda_1 &= \alpha \\ \lambda_i &= \alpha + (1-\alpha)\exp\left(-\frac{c}{K}\right), i = 2 \dots n \\ D_{\alpha,c}(x) &= \lambda_1 w_1 \cdot w_1^T + \lambda_2 w_2 \cdot w_2^T + \lambda_3 w_3 \cdot w_3^T \end{aligned}$$

The parameter α is responsible for a small regularization across the edge, serving as a threshold to guide smoothing along the edge depending on the local coherence of D, which is proportional to the variance of the eigenvalues of T_ρ (Weickert, 1999). If $c \ll K$, smoothing along the edges is higher; conversely, if $c \gg K$, smoothing along the edges is reduced. See Figure 4.6.

Modified fast level-set segmentation

In this study, a modified fast level-set algorithm (Kratky and Kybic, 2008) based on tissue density priors is used. The aim is to segment the teeth from the background, other bone, and soft tissue. While experimenting with the original algorithm, it was found that it was rather sensitive when faced with a choice between c_1, c_2, λ_1 and λ_2. Sometimes, it was not even possible to find the correct values despite a multitude of attempts.

Different c_1 and c_2 are used for different parts of the image and these constants are continuously recomputed during the iterations, i.e. local values

for c_1 and c_2 can be reinitialized once for every r iteration. The value of r is the reinitialization period. The major improvement in the solution is the way local statistics, such as average intensity, are computed

$$c_i(x) = \frac{\int_\Omega W(D(y),x,y)H_{i,\epsilon}(\psi(y))f(y)dy}{\int_\Omega W(D(y),x,y)H_{i,\epsilon}(\psi(y))} \approx \frac{\sum_j W(D(y_j),x,y_j)M_i(\psi(y_j))f(y_j)}{\sum_j W(D(y_j),x,y_j)M_i(\psi(y_j))}, i = 1,2$$

Image intensity variability can be computed similarly

$$\sigma_i(x)^2 = \frac{\int_\Omega W(D(y),x,y)H_{i,\epsilon}(\psi(y))(c_1(x)-f(y))^2dy}{\int_\Omega W(D(y),x,y)H_{i,\epsilon}(\psi(y))} \approx \frac{\sum_i W(D(y_i),x,y_i)M_i(\psi(y_i))(c_1(x)-f(y_i))^2}{\sum_i W(D(y_i),x,y_i)M_i(\psi(y_i))}, i = 1,2$$

In the case of fast level-set segmentation, the continuous Heaviside function $H_{i,\epsilon}$ is not needed as ψ is changed in discrete steps. The weight factor W is defined as the density of normal distribution with respect to use of the local structure, represented by the diffusion tensor $D = D_{c,a}$

$$W(D,x,y) = \frac{1}{\sqrt{(2\pi)^k|D|}} exp(-\frac{1}{2}(x-y)^T D^{-1}(x-y))$$

The masks for separating the inner and outer regions are defined as follows

$$M_1(x) = \frac{(sgn(x)+1)}{2}$$
$$M_2(x) = \frac{(sgn(-x)+1)}{2}$$

Diffusion tensor D can be pre-computed with pre-defined parameters a once, and used repeatedly for local intensity statistics $c_i(x)$ according to Equation 1.

Post-processing

A regrettable property of tooth data is that the individual teeth lie very close to each other and sometimes touch in large contact areas. Combined with the low resolution and partial volume effect, it is almost impossible to distinguish an edge between two parts of adjacent teeth. Hosntalab *et al.* (2008) separated teeth using panoramic projection before performing actual tooth segmentation. Gao and Chae (2010) extended the level-set segmentation with region competition to keep even and odd teeth separated along the tooth arch.

In order to make the volume segmentation less complicated and for the purpose of separating everything from slightly touching to an almost indistinguishable pair of adjacent teeth, it was decided to perform the separation in a

post-processing step. Graph-cut based segmentation proved to be very effective for the segmentation and separation of bones, provided the edges of the corresponding graph were chosen carefully (Krcah *et al.*, 2011). Morphological operations were performed (erosion, dilatation) to acquire sub-volumes of the individual teeth, which were then used as seeds for a simple 2-way graph-cut algorithm to iteratively separate one tooth from the other (Liu *et al.*, 2008; Kolomazník *et al.*, 2012).

Artificial data experiments

In the first experiment, the properties of local statistics are demonstrated via modified fast level-set segmentation on artificial data. Artificial data are represented by a volume containing a cube with gradient intensity on a gradient intensity background. The gradients lie in opposite directions. Moreover, the foreground mean intensity is equal to the mean intensity of the background (see Figure 4.7(a)). Formally, it is impossible to distinguish between the background and the foreground in this case. Figure 4.7(d) shows a plot of intensities in a cut through the artificial dataset. In some related papers, similar tests with inhomogeneous backgrounds were performed; however, the mean intensities of foreground and background were different, as illustrated by the dotted and dashed profiles in Figure 4.7(d).

The results of segmentation on such complicated data obtained by proposed method are shown in Figure 4.7(c). Because of the high ρ, the segmentation method cannot capture very sharp features, such as the corners of the box. However, in comparison to the locally modified Chan-Vese method by Gao *et al.* (2012), segmentation of the foreground was successful. Figures 4.7(b) and 4.7(e) show different approaches to incorporating local information. The first image demonstrates how the local neighborhood is weighted while computing the local statistics in different parts of the image. On the other hand, the second image shows intensity gradients, compensated by subtracting the original u image from its blurred version $k_\sigma * u$, which is an example of the difference-of-Gaussians method for edge enhancement. This image is a part of the local term that is linearly combined with the global Chan-Vese model.

The second experiment on the artificial data demonstrates how local anisotropy in a region statistics computation affects the segmentation results. The test data consist of a gradient spiral on a gradient background (see Figure 4.8(a)). Figure 4.8(d) shows that segmentation using the global criterion fails completely. Figures 4.8(b) and 4.8(e) show the difference between isotropic and anisotropic local statistics weighting using the properties of the diffusion tensor, as described above. The main advantage of anisotropically computed local statistics for region segmentation is that they help to overcome areas where the foreground object's intensity is similar to the intensity of the neighboring background (see Figure 4.8(f)).

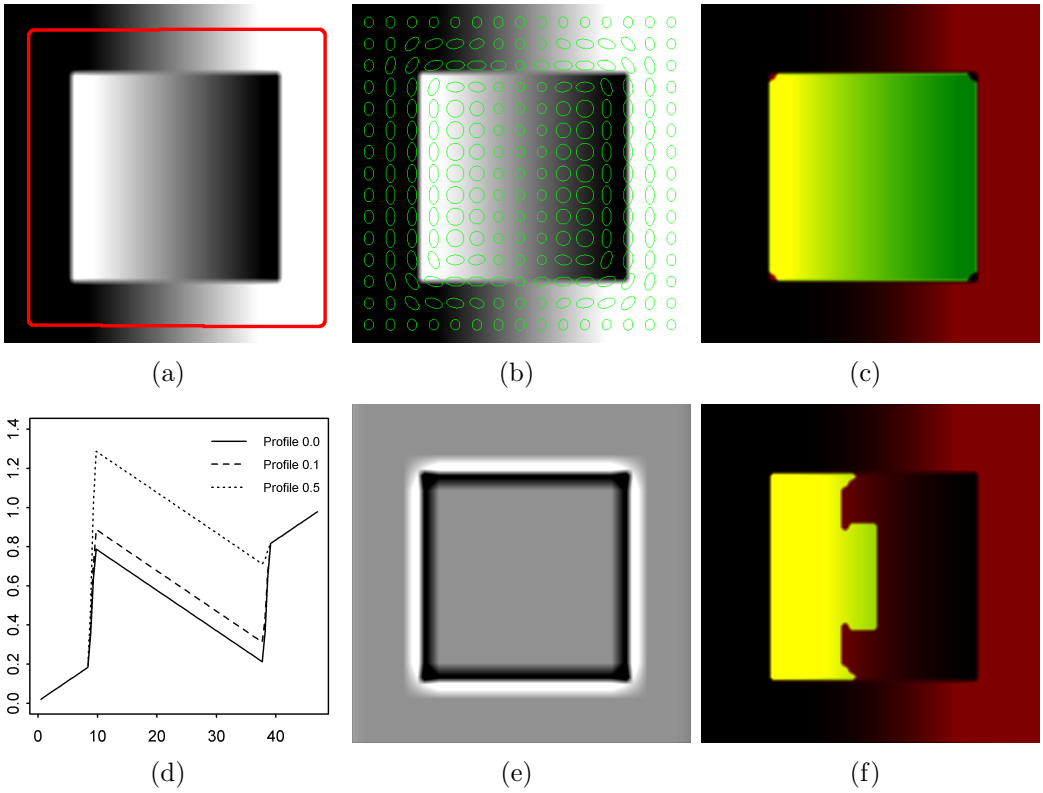


Figure 4.7: Results of the algorithm on artificial data: (a) original data with an initial segmentation (red line); (b) visualization of the D tensor field ($a = 0.5$, $c = 0.00001$, $\rho = 4.0$); (c) the result of segmentation using our algorithm; (d) the plot of data profiles with various additive constants for the foreground (zero is used in our tests); (e) visualization of $k_\sigma * u - u$, which is used to homogenize the image in local term provided by Gao *et al.* (2012); (f) the result of the algorithm by Gao *et al.* (2012).

Real data experiments

Real data experiments start with a small sub-volume of raw skull CT data containing a single tooth (see Figure 4.9(a)). The tooth tissue is clearly very inhomogeneous. Segmentation without the anisotropically computed statistics does not allow the level-set to spread along tissue layers of the same density (Figure 4.9(b)). On the other hand, anisotropy (Figure 4.9(d)) helps the algorithm to spread over the whole object successfully without leaking into the jawbone tissue (see Figure 4.9(c)). Of course, as with the other active contour algorithms, the initialization of the level-set plays a role in the resulting segmentation. The initial seed area should cover a wider range of intensities, but positioning the seed volume in an area of extreme intensity (e.g. the center of the tooth cavity, a dense area of enamel) would lead to undesirable results.

In the second real data experiment, the whole tooth arch is segmented

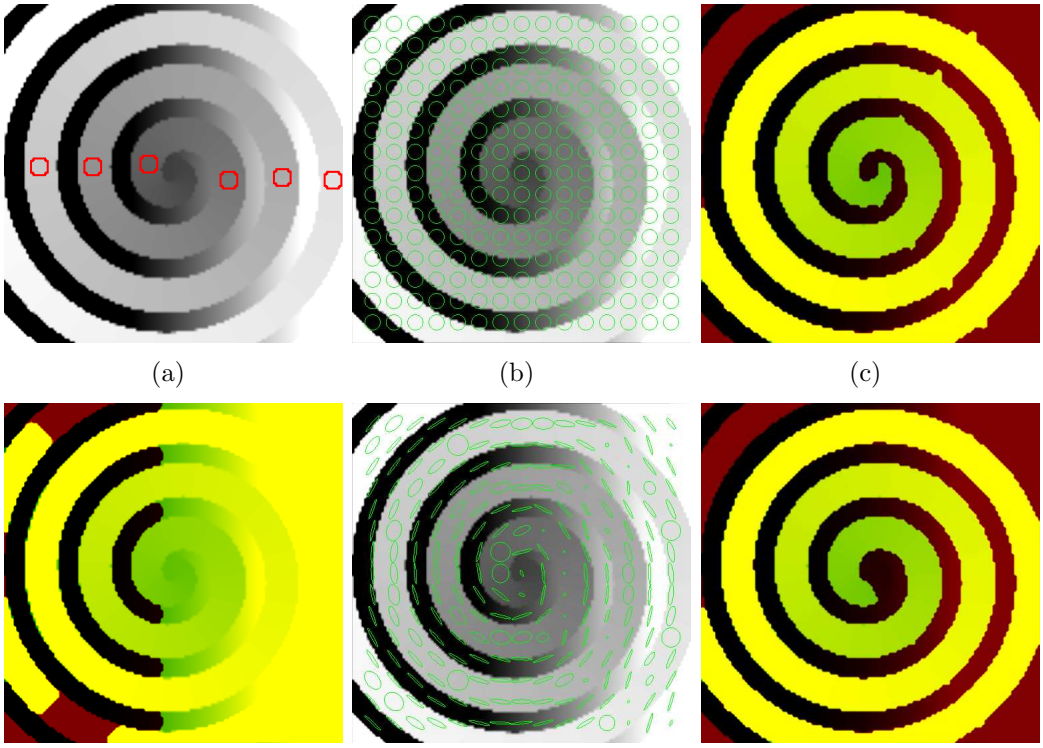
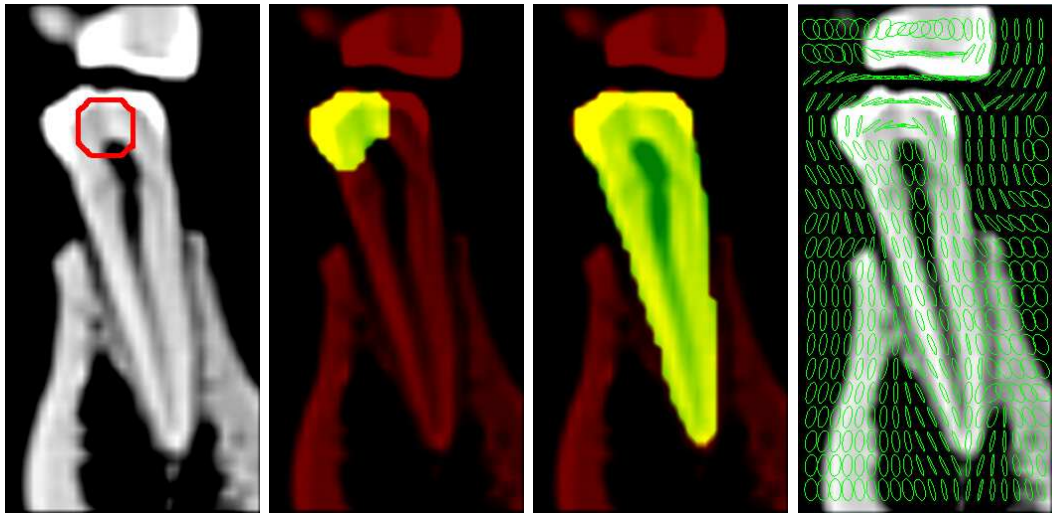


Figure 4.8: Artificial data segmentation demonstrating the anisotropic computation of statistics: (a) spiral data with segmentation initialization; (b) isotropic tensor visualization ($\alpha = 0.3$, $c = 0.1$, $\rho = 5.0$); (c) results of isotropic local segmentation; (d) the result of segmentation using global statistics; (e) Anisotropic tensor visualization ($\alpha = 0.3$, $c = 1e-9$, $\rho = 5.0$); (f) Result of anisotropic local segmentation.

from raw skull CT data. First, the seeds need to be placed manually or semi-automatically. This can be accomplished easily by looking for a slice passing horizontally through the upper parts of the teeth. Subsequently, simple thresholding is performed and connected components representing individual teeth are selected to obtain the seeds (see Figures 4.10(b) and 4.10(c)). Next, segmentation with anisotropic weighting for local statistics computation is performed. Successful separation of the tooth arch from the rest of the bone structure is then achieved (Figures 4.10(d) and 4.10(e)). Up until this point, the aim is not to segment individual teeth. Instead, the author concentrates on the separation of the whole arch from the jawbone, while splitting of the arch into individual teeth is left for a post-processing step. The results of the post-processing are shown in Figure 4.10(f).

In the last real data experiment, the algorithm was used to segment the tooth arch and the remaining jawbone from the rest of the skull. The motivation in this case is that CT datasets of a unique historical skull collection were



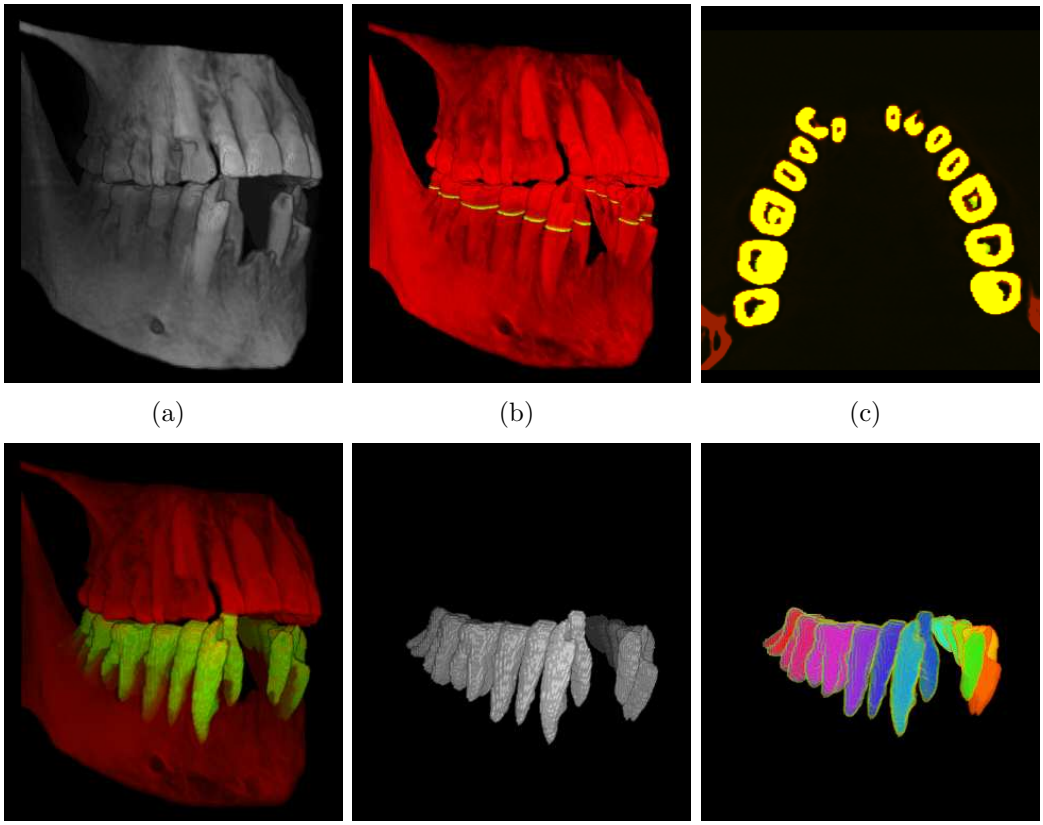
(a) (b) (c) (d)

Figure 4.9: An example of segmentation of real tooth data: (a) lateral input tooth image ($32 \times 64 \times 48$ voxels) with initialization; (b) results of segmentation using a local criterion with low anisotropy ($a = 0.1$, $c = 0.1$); (c) results of segmentation with high anisotropy ($a = 0.1$, $c = 1e - 8$); (d) visualization of a high anisotropy tensor.

improperly scanned: the cranium was placed on the jawbone in an incorrect anatomical position, which prevented the shapes from being analyzed individually. The major problem here is that the cranium and jawbone touch in several places, i.e. the area of the jaw joint condyles and teeth. Furthermore, the teeth are not important if the goal is to study the shape of the jawbone alone. Therefore, they must be removed, as in the previous experiment. The procedure of extracting the jawbone is based on the separation of bones with identical density and a highly porous structure. The best approach is to employ graph-cut segmentation seeded with the pre-segmented regions from the level-set algorithm. The pre-segmentation itself is not very precise due to the structure of the bone. However, it reaches all the parts of the jawbone and produces a good seed for graph-cut segmentation. The individual steps of the procedure are demonstrated in Figure 4.11.

Implementation issues

The fast level-set segmentation algorithm (Algorithm 8) evaluates $F(x)$, which depends on the current local foreground and background statistics of $c_1(x)$, $\sigma_1(x)$, $c_2(x)$ and $\sigma_2(x)$. These values are computed based on the given segmentation for the foreground and background. The values of the statistics can either be computed every time they are needed or, alternatively, pre-computed and fixed for iterations, after which they are recomputed again. The former



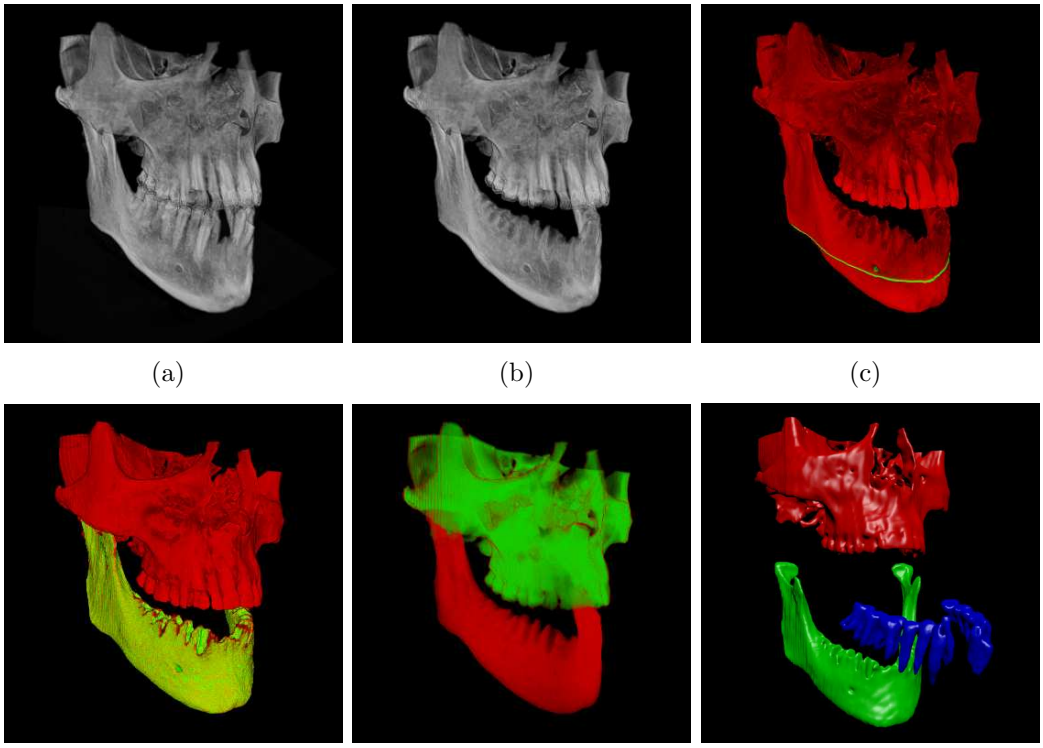
(a) (b) (c)

(d) (e) (f)

Figure 4.10: An example of segmentation of a whole tooth arch: (a) visualization of input tooth arch volume ($180 \times 128 \times 256$ voxels); (b) visualization of segmentation seeds in 3D intersecting a well-defined part of the arch; (c) a 2D slice with seeds; (d) segmentation results ($a = 0.05$, $c = 1e - 9$, $\rho = 3.0$); (e) a tooth arch without jawbone; (f) individual tooth separation in a post-processing step.

approach speeds up the convergence but makes segmentation less progressive over time. The computation of the statistics is very time-consuming. Its time complexity represents all the voxels in a cubic volume, where the side of length n is $\mathcal{O}(n^6)$. This can be reduced if only a limited neighborhood of each voxel is considered and where the weight values are far from zero. The size of the neighborhood depends on the value of ρ, which is the width of the Gaussian kernel used for weighting.

In order to reduce the time of computation in a real application, a parallel computing scheme is implemented using parallel graphics hardware. The computation of individual local statistics c for the individual voxels is performed in a CUDA kernel. Due to this improvement, the computation time is reduced by several orders of magnitude. Even with parallel computation, gathering local statistics is the slowest part of the whole algorithm.



(a) (b) (c)

(d) (e) (f)

Figure 4.11: Jaw segmentation: (a) input skull data; (b) skull data without teeth; (c) seed for level-set pre-segmentation; (d) results of the graph-cut segmentation; (e) separated teeth, jawbone, and cranium.

Summary

A segmentation algorithm for easy tooth arch extraction was described, which is based on a simplification of level-set segmentation by Kratky and Kybic (2008) using embedded anisotropic local intensity priors. The algorithm was designed to objectively segment teeth from bone relics, scanned with standard medical CT, as well as images of patients. However, it was not tested on extreme cases, such as patients with large tooth implants or deformations. It was assumed that because extreme cases can be so different, automatic algorithms might only help with their segmentation to a degree and that the rest would need to be manually segmented. It was decided to employ anisotropic local intensity priors since the teeth and jawbone area contains a wide variety of bone tissue densities with unreliable edge information. From this point of view, relying on a single type of feature e.g. edge or intensity priors, was not deemed to be sufficiently robust. In a broader sense, locally adaptive region information is another way of incorporating edge information into an active shape model criterion. A comparison with the recent locally modified Chan-Vese method (Gao *et al.*, 2012) was provided.

Additionally, the particular task of tooth segmentation was approached as a two-phase process. Initially, the tooth tissue is segmented and then the teeth are separated from each other in a post-processing procedure. This separation enables specialized methods for jaw-teeth and tooth-tooth segmentation to be used.

The segmentation procedure described here relies on a certain amount of user interaction and initialization. However, it is usually solved by a smart user interface or various support algorithms such as thresholding, region growing, small component removal, or clustering in a particular application. In this case, a slice parallel to the tooth arch was always available, containing clearly separated teeth to be used for the initial segmentation.

In the future, the algorithm will be tested on CT images of living people. It is expected that the major problem will lie not in the presence of soft tissue but in the medical condition of the teeth. In most cases, adults have replacements or fillings, which are made from a far less X-ray-penetrable material then the hardest bone tissue. This causes star-shape artifacts and beam hardening (see Figure 4.5(c)).

4.2 Non-rigid volume registration

Similar to the previous chapter, the primary tool for correspondence searches in volume data is non-rigid registration. However, the principle of correspondence is a little different from point clouds or meshes. With volume data, corresponding points are considered points that are mapped on each other by transformations found via the registration procedure.

Non-rigid or elastic registration has been intensively researched in the last two decades, mostly in the area of medical image processing (Mani and Arivazhagan, 2013). It is primarily characterized by transformation, one of the three components of registration (see Section 3.2), which allows non-linear deformation mapping between the source and target image.

A non-linear deformation function can be divided into two classes. The first is a *parametric approach*, which defines the transformation function using a smaller set of parameters in the form of feature points (landmarks) or control points, and an interpolation operator that computes transformation of intermediate points. The second is a *non-parametric approach*, which describes and optimizes the transformation of each individual point in the image while keeping them locally consistent. The latter is computationally more expensive; however, it allows advanced mathematical concepts to be incorporated, such as Navier equations with material constants, optical flow, PDE formulation of image relation, etc. The former group, on the other hand, is often based on the mathematics of splines or the piece-wise linear approximation of smooth non-linear mapping.

A typical example of the parametric approach is TPS-based registration, described in Section 2.2 [page 19]. The method defines the transformation

globally, i.e. each landmark point influences the transformation everywhere. By contrast, B-spline (introduced in Section 3.2.4) is only defined in the close vicinity of control points, which means that it is able to model localized deformations, while maintaining relatively low computational costs and general applicability. Hence, B-spline registration has gained in popularity in various registration tasks in medical image processing (Krajíček, 2008a).

In the following text, a registration method suitable for statistical analysis of volume data is described (Krajíček *et al.*, 2011). The core idea and mathematical background to the algorithm are the same as the one used for triangular mesh registration (see Chapter 3). The difference and novelty of the method described here are attributed to introduction of speed-up approaches, which could make it useful for tasks that would otherwise require vast computational power and a lot of time to perform.

4.2.1 B-spline-based registration

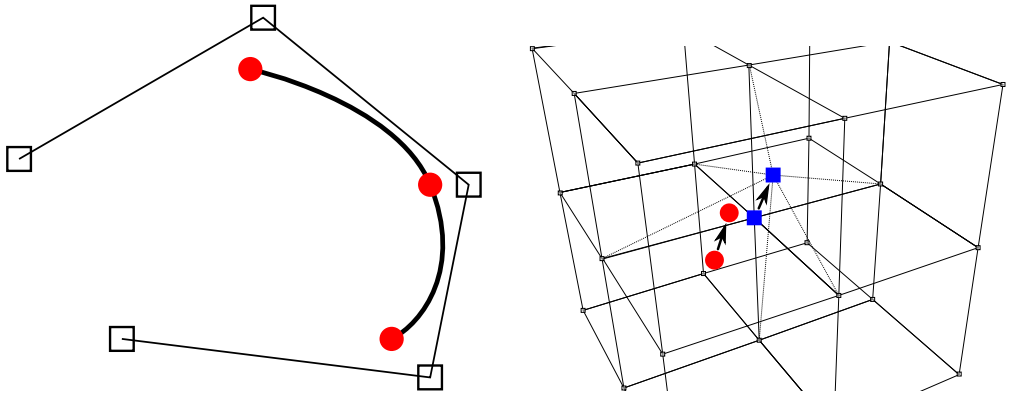
Registration is an optimization of the relation between source and target images by looking for the optimal transformation of the source image to the target image. In the case of B-spline-based registration, the transformation is modeled by the B-spline free-form space warping function f_{FFD}, which is derived from a B-spline curve. This curve is defined by control points and minimal support polynomial bases with respect to the particular degree and smoothness (see Figure 4.12). It was also referenced in Section 3.2.4 in relation to application in point cloud and mesh registration. For the sake of completeness, the f_{FFD} definition is repeated here

$$f_{FFD}(\vec{x}, c) = \sum_{i=l}^{l+d} \sum_{j=m}^{m+d} \sum_{k=n}^{n+d} c_{ijk} B_{i-l}^{d}(x') B_{j-m}^{d}(y') B_{k-n}^{d}(z')$$

where c is a grid of control points, d is the order of polynomial bases, and i, j, k are all indices of control points that influence the rectangular subset of the f_{FFD} domain given by $l = \lfloor x/(n_i - d) \rfloor$, $m = \lfloor y/(n_j - d) \rfloor$ and $n = \lfloor z/(n_k - d) \rfloor$. As a basis B_i^d, a cubic polynomial ($d = 3$) is usually used (see Equation 3.6).

With a continuous, differentiable function as a solution, registration can be formulated as a variational problem of minimizing functional $E(f_{FFD})$. However, registration, in general, is an ill-posed problem, which means that there are many solutions to such a problem and it cannot be decided which one is the best. Moreover, B-spline free-form space warping is an artificially designed function which can model spatially and physically undesirable space transformations (e.g. self-intersection, space dilation or expansion). In order to solve both problems, the regularization term f_r is required in variational formulation next to the data term f_d, which models the relation in the data

$$E(f_{FFD}) = \int_{\Omega} (1 - \alpha) f_d(f_{FFD}, x) + \alpha f_r(f_{FFD}, x) dx$$



(a) (b)

Figure 4.12: Demonstration of B-spline space deformation: (a) an example of a parametric cubic B-spline curve (thick line) with control points (squares) and control polygon (narrow lines); (b) an example of free-form space deformation with a control points (squares). The grid demonstrates space deformation at an arbitrary point (red) by shifting one of the control points (blue square).

In B-spline non-rigid registration, the main difference between the registration of raster data and point cloud data is in the basic formulation of the problem. It is an integral of smooth function over continuous image domain Ω instead of a discrete sum over a point set, which is discontinuous in principle.

One option for regularization term f_r is a sum of second derivatives (similar to Equation 3.7).

$$f_r(f_{FFD}, x) = \sum_{p=1}^{3}\sum_{q=1}^{3}\left(\partial^2 f_{FFD}(x)/\partial x_p \partial x_q\right)^2 \tag{4.7}$$

Another option is to carry out volume preservation while matching two images of the same specimen (Rohlfing *et al.*, 2003). This is done by adding incompressibility constraint base on the determinant of the Jacobi matrix (a matrix with partial derivatives), which is equal to one for volume preservation at point x.

$$f_r(f_{FFD}, x) = |log(J_{f_{FFD}}(x))|$$

This particular task also requires proper data term definition. The most commonly used is the sum of the squared difference between the source image sample at x and the target image sample at $T(x)$.

$$f_d(f_{FFD}, x) = (S(x) - T \circ f_{FFD}(x))^2$$

In order to fit multimodal images (CT and PET) or images that are not normalized to a particular intensity scale (MRI), more complicated metrics are used, such as mutual information (MI) or normalized mutual information (NMI) (Rueckert *et al.*, 1999).

The core of the registration process searches for f_{FFD}, which minimizes E. Due to the parametric nature of f_{FFD}, local numerical optimization can be used to find correct locations for control points c. Many numerical optimization algorithms can be used for this purpose. Particular examples are listed in Section 3.2.4 [page 50]. One interesting characteristic which could be used to distinguish between them is the amount of gradient information they depend on. In the following text, modified steepest descent is employed. Either way, almost all of the methods require computation of partial derivatives of E with respect to c_{ijk} and $\nabla E = \{\partial E/\partial c_{ijk}\}_{ijk}$, which is decomposed into partial derivatives of individual terms. These derivatives can be approximated by central differences, but for more precise values an analytical solution is preferred.

$$\nabla E = \left\{ \int_\Omega (1-\alpha)\frac{\partial f_d}{\partial c_{ijk}} + \alpha \frac{\partial f_r}{\partial c_{ijk}} dx \right\}_{ijk}$$

Partial derivatives are computed by applying the chain rule

$$\begin{aligned}
\frac{\partial f_d}{\partial c_{ijk}} &= \left[\frac{\partial f_d}{\partial c_{ijk,x}}, \frac{\partial f_d}{\partial c_{ijk,y}}, \frac{\partial f_d}{\partial c_{ijk,z}} \right] \\
\frac{\partial f_d}{\partial c_{ijk,x}} &= \frac{\partial f_d}{\partial f_{FFD,x}} \frac{\partial f_{FFD,x}}{\partial c_{ijk,x}} \\
&= -2(S - T \circ f_{FFD}) \frac{\partial T}{\partial x} B^d_{i-l} B^d_{j-m} B^d_{k-n}
\end{aligned}$$

Partial derivatives of f_r (Equation 4.7) are little bit more complicated

$$\begin{aligned}
\frac{\partial f_r}{\partial c_{ijk,x}} &= \frac{\partial}{\partial c_{ijk,x}} \sum_{p=1}^{3} \sum_{q=1}^{3} \left(\partial^2 f_{FFD,x}(x) / \partial x_p \partial x_q \right)^2 \\
&= \sum_{p=1}^{3} \sum_{q=1}^{3} 2 \frac{\partial^2 f_{FFD,x}}{\partial x_p \partial x_q} \frac{\partial^3 f_{FFD,x}}{\partial x_p \partial x_q \partial c_{ijk,x}} \\
\frac{\partial^2 f_{FFD,x}}{\partial x_p \partial x_q} &= \frac{1}{n_p - d} \frac{1}{n_q - d} \sum_{i=l}^{l+d} \sum_{j=m}^{m+d} \sum_{k=n}^{n+d} c_{ijk,x} \times \\
&\quad (B^d_{i-l})^{(r(p,q,1))} (B^d_{j-m})^{(r(p,q,2))} (B^d_{k-n})^{(r(p,q,3))} \\
\frac{\partial^3 f_{FFD,x}}{\partial x_p \partial x_q \partial c_{ijk,x}} &= \frac{1}{n_p - d} \frac{1}{n_q - d} \times \\
&\quad (B^d_{i-l})^{(r(p,q,1))} (B^d_{j-m})^{(r(p,q,2))} (B^d_{k-n})^{(r(p,q,3))}
\end{aligned}$$

where $(B^d_i)^{(r(\cdot))}$ means $r(\cdot)$-order partial derivative of B^d_i with $r(\cdot)$ defined as

$$r(p,q,s) = \begin{cases} 2 & p = q = s \\ 1 & p = q \veebar q = s \\ 0 & p \neq q \neq s \end{cases}$$

Algorithm 9: General numerical optimization based on gradients

Input : S - source, T - target, $initial\_guess$, ϵ - tolerance
Output: $solution$
begin
  $solution \longleftarrow initial\_guess$
  $converged \longleftarrow False$
1 **while** *not converged* **do**
2 **for** $(i, j, k) \in n_i \times n_j \times n_k$ **do**
     $\frac{\partial E}{\partial c_{ijk}} = 0$
3 **for** $x \in \Omega_{(i,j,k)}$ **do**
       $\frac{\partial E}{\partial c_{ijk}} += (1 - \alpha)\frac{\partial f_d(x)}{\partial c_{ijk}} + \alpha\frac{\partial f_r(x)}{\partial c_{ijk}}$
     end
    end
    $new\_solution \leftarrow solution - \delta * \nabla E$
    $converged \leftarrow \|new\_solution - solution\|^2 < \epsilon$
    $solution \leftarrow new\_solution$
  end
end

The integral in the previous equation can be approximated by the sum over all the voxels in the volume image. Fortunately, the derivatives of the terms are non-zero only in a subset of the f_{FFD} domain. For the general scheme of the optimization algorithm, see Algorithm 9.

The total time needed for computation depends on the speed of convergence and the number of iterations it performs (Line 1). On the other hand, the most computationally demanding part of the algorithm is for the loop on Line 3, where partial derivatives have to be computed. This results in at least two places which can be focused on in order to speed up the algorithm.

4.2.2 Speed-up approaches

Image registration, in general, is a computationally demanding process. Many methods of improving performance have been studied. Some of them are enhancements on the implementation level, while others work on the algorithm level. Since one of the goals of this work is to propose a solution for practical usability, the following paragraphs describe approaches that improve computation time.

Parallelization

The straightforward approach is to employ trivial parallelization to compute of partial derivatives, i.e. a loop on Line 2 in Algorithm 9 is parallelized, since these values are independent of each other and produce a single scalar.

Theoretical speed-up is, in this case, linear with the number of CPU cores.

Shackleford *et al.* (2010) employed GPU hardware to run CUDA implementation of B-spline registration, which was shown to be 15× faster than the highly optimized CPU version. Rohlfing and Maurer (2003) used a multiprocessor with shared memory to speed up the computation of the MI criterion, which requires continuous recomputation of joint histograms.

Multiresolution

Multiresolution is often used as a speed-up technique for image registration or even for other image processing algorithms. Input data are scaled down so that the algorithm converges on them faster. The solution of this scaled instance is then scaled up and used as an initialization of the algorithm on the original data. The idea is that the algorithm will make fewer steps before converging so that the total time of these two computations is shorter than the time of the original one. This can also work on more than two levels. There is also one additional advantage of numerical optimization: on scaled down data the algorithm easily overcomes some local minima in which the algorithm on the original data converges. In fact, this approach is not only faster but produces better results. The concept behind this approach is similar to multigrid methods used in numerical analysis.

In the case of B-spline-based registration, the resolution of the control grid can also be progressively increased. The algorithm with the sparser control grid displays more global deformation behavior, but works more locally with a denser grid. The only remaining problem is how to scale up the solution, i.e. to increase the number of control points. In the case of B-spline problems, there are many ways to subdivide the control grid and to insert control points. One particular solution proposed by Lane and Riesenfeld (1980) almost doubles the number of control points in a few simple steps (also see Algorithm 3). It is recommended that this operation be performed together with resolution scaling of the input image in Gaussian pyramid fashion.

Sampling

Smart sampling is a natural way to speed up algorithms which involve computations of multidimensional integrals over large sets, such as those evaluated when computing partial derivatives in Algorithm 9 (Line 3).

$$\frac{\partial E}{\partial c_{ijk}} = \int_{\Omega} (1-\alpha)\frac{\partial f_d(x)}{\partial c_{ijk}} + \alpha\frac{\partial f_r(x)}{\partial c_{ijk}} dx$$

By choosing the appropriate sampling strategy, a much lower number of samples is needed when evaluating the approximation of the integral and preserving accuracy, compared to traditional dense regular sampling. Moreover, the registration does not depend on one individual gradient estimation because it is recomputed several times and results improve gradually.

Stochastic sampling has been used for image registration by a few studies in recent years. Klein *et al.* (2006) were probably the first to study simple stochastic sampling as an acceleration technique for image registration. They compared it with smarter optimization procedures (Quasi-Newton method, conjugate gradient) and observed improved performance in one to two orders of magnitude without losing accuracy. Incorporating stochastic sampling in gradient descent results in *stochastic gradient descent* (SGD) or *mini-batch stochastic gradient descent*, which are very popular methods in the field of machine learning (Spall, 2003, chapter 5).

Bhagalia *et al.* (2006) described the use of importance sampling for intensity-based image registration with a mutual information data criterion of MRI images.

The Quasi-Monte Carlo sampling approach was used by Thevenaz *et al.* (2008) to speed up joint histogram computation for mutual information-based registration.

Apart from these scientific results, stochastic sampling has not yet been fully accepted as an efficient registration speed-up technique in the same manner as multiresolution. Also, no study has ever compared smarter sampling methods, such as importance sampling and Quasi-Monte Carlo sampling, when focusing on implementation, performance, and accuracy.

Convergence criterion

Numerical optimization procedures often converge in many iterations and do not improve the solution much in the last part of the process. Stopping the convergence at the right moment can significantly lower computation time. Stopping criteria for image registration algorithms is not an extensively researched area. Only the basic rules are mentioned in published texts.

For example, an algorithm should stop if there is almost no change between the last two solutions, if the gradient of the objective function with respect to the parameters is almost zero or, if there have been a fixed number of iterations. Such conditions are represented by thresholds which are not easy to set and which must often be found empirically. Stochastic sampling inside the optimized objective function changes the behavior of the optimization procedure as well as the criterion in order to stop the convergence. To stop stochastic gradient descent, a slowly decaying gain sequence can be used (Klein *et al.*, 2006).

4.2.3 Sampling methods

The basic principle behind Monte Carlo methods (Owen, 2003) is in repeated random generation of a certain result, assuming that the mean value of such a process converges to the exact value. Monte Carlo integration is an approx-

imation of an integral by a finite sum of n random function samples

$$I = \int_\Omega f(x)dx \approx I_n = |\Omega|\frac{1}{n}\sum_{i=1}^{n} f(x_i)$$

More importantly, $E(I_n) = I$, which is in fact used in many applications, e.g. image rendering and stochastic optimization. The variance of I_n is inversely proportional to the number of samples; hence, the root mean square error is $\mathcal{O}(1/\sqrt{n})$. However, it can be improved by choosing a proper sampling approach, i.e. a method used to select individual samples.

A similar idea is tested using stochastic gradient descent optimization where various sampling strategies are selected for experimentation. All stochastic methods are compared to regular sampling (see Figure 4.13(b)), which is usually used with raster-based registration algorithms.

The simplest of these, and the one with the largest variance, is pure random sampling (see Figure 4.13(c)). Samples are taken from the whole integration domain by generating pseudo-random numbers between boundaries of the domain. With a smaller number of samples, some empty areas or areas with too many samples can appear.

The previous result can be improved using stratified sampling (see Figure 4.13(d)), where the domain is subdivided into subdomains, in which a certain portion of random samples are generated. The variance of such sampling is reduced compared to pure random sampling.

If the approximate shape of the integrated function is known in advance, *proposal distribution* D can be created. Probability p can be assigned to each image element, which expresses how much it contributes to the whole integral and how important it is to include the element in the resulting integral approximation. Hence, this method is known as importance sampling (Figure 4.13(e)).

$$\int_\Omega f(x)dx = \int_\Omega \frac{f(x)}{p(x)}p(x)dx \approx I_n = \frac{1}{n}\sum_{i=1}^{n}\frac{f(x_i)}{p(x_i)}$$

It has the lowest variance but the disadvantage is that the basic importance sampling algorithm (inverse transform sampling of discreet CDF) can only generate a certain fixed number of pre-defined samples. Also, the structure that is used to hold them is as large as the image itself (Devroye, 1986, chapter 2.2).

The last sampling method presented here belongs to a group that are made up of what are known as Quasi-Monte Carlo (QMC) methods. These methods do not generate samples randomly; rather, they use a deterministic algorithm which produces points that cover the whole integral domain. Various Quasi-Monte Carlo sampling algorithms are compared with respect to quality of their coverage by a value of (star) *discrepancy* measure $D_n^\star$.

$$D_n^\star = \sup_{a\in[0,1)^d}\left|\frac{1}{n}\sum_{i=1}^{n} 1_{0\le x_i<a} - |[0,a)|\right|$$

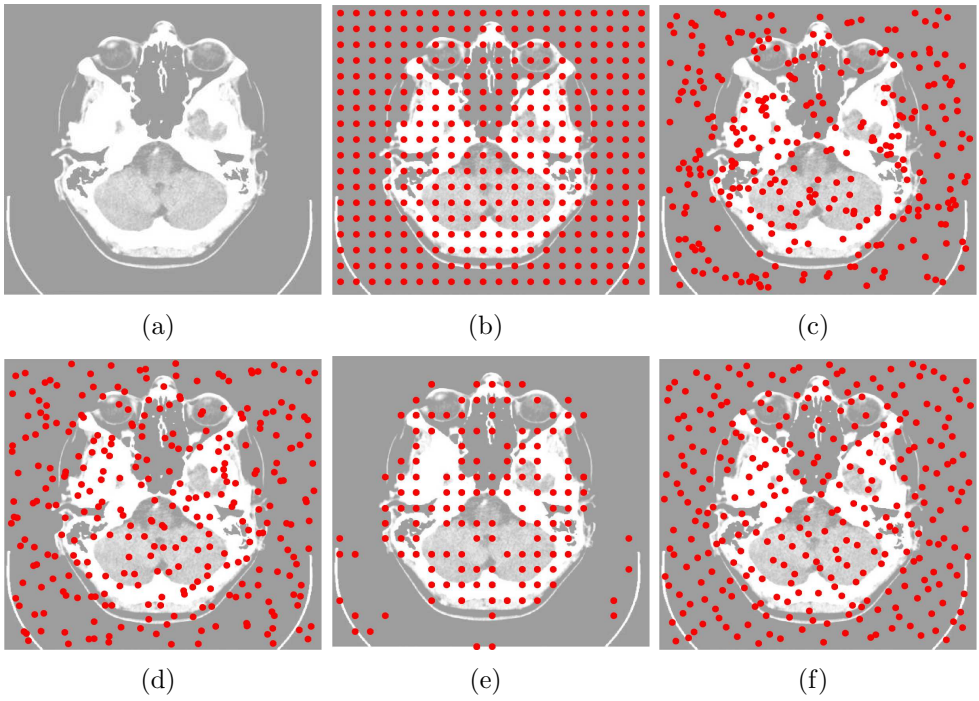


Figure 4.13: Various sampling approaches: (a) original image; (b) regular sampling (360 samples); (c) random sampling; (d) stratified sampling; (e) importance sampling; (f) Halton sequences. All stochastic schemes are presented with 200 samples.

One example of a QMC sampling algorithm with low discrepancy is based on Halton sequences (Figure 4.13(f)). The idea of generating a pseudo-random value is to iterate integer numbers from 0 to infinity and express them in a selected prime base, b.

$$n = \sum_{k=1}^{\infty} n_k b^{k-1}, n_k \in \{0, 1, ..., b-1\}$$

The individual digits are then combined together and weighted by inversions of a powered base, b^{-k}. Using different prime bases for individual coordinates, multidimensional vector sampling of the space, with discrepancy of Halton sequences, can be generated.

$$\phi_b(n) = \sum_{k=1}^{\infty} n_k b^{-k} \Longrightarrow \vec{x} = (\phi_5(0), \phi_7(0)), (\phi_5(1), \phi_7(1)), \ldots$$

4.2.4 Implementation

Theoretical foundations for implementing the B-spline non-rigid registration algorithm are now introduced. Before the algorithm is executed to fit the input

images, they are pre-aligned rigidly. Manually placed landmarks can be used for this purpose. Rigid alignment of landmarks can be performed by ordinary Procrustes superimposition described in Section 2.2 [page 17]. As a further step, non-rigid registration is employed to solve the deformation part of the transformation between the data sets.

Data criterion

The sum of squared differences (SSD) serves as a standard mean of comparison for all the experiments. The use of mutual information is rejected because it is computationally expensive and pointless since no multi-modal data are used. In addition to that, MI has a property to build a relation between element values in both images in every registration step, depending on the number of occurrences of element pairs in a joint histogram. In other words, if the initial images are badly aligned, they could relate black pixels in the first image with white pixels in the second image. This is a very real possibility in the case of bone-air images.

The very strong boundary information of the data is exploited as a replacement for MI. It is a good idea to relate image elements to the object boundary by the shortest distance. In order to store information on the shortest distance of each element to the nearest boundary element, the Euclidean distance map (EDM) is computed. Subsequently, these maps, S_{edm} and T_{edm}, are aligned in the context of the sum of squared differences.

$$f_d(x) = (S_{edm}(x) - T_{edm} \circ f_{FFD}(x))^2 \tag{4.8}$$

EDMs have already been used for non-rigid shape registration by Paragios *et al.* (2003) in an algorithm combining global and local transformations. Local non-rigid transformation, guided by lowering of the differences in EDT, are evaluated only in close vicinity of the shape boundary. Otherwise, differences in the distance from the shape are larger; hence, the optimization prioritizes them over the areas along the boundary, which is undesirable.

The advantage of EDMs is also that they can be computed very quickly using dynamic programming (see Meijster *et al.* (2002)). To find a boundary in the data or to divide the data into two groups of elements (the object and the background), any segmentation algorithm can be used.

Importance sampling

The importance sampling algorithm is based on random sampling of image domain Ω with respect to *proposed distribution* D, e.g. derived from smoothed image intensity levels or image gradient magnitudes. It is realized using inverse transform sampling of the discrete cumulative distribution function (CDF) of D (Devroye, 1986, chapter 2.2), which requires construction of CDF for the integration domain and storage of it in the memory. Moreover, during the partial derivative computation, only a sub-interval of Ω is sampled because

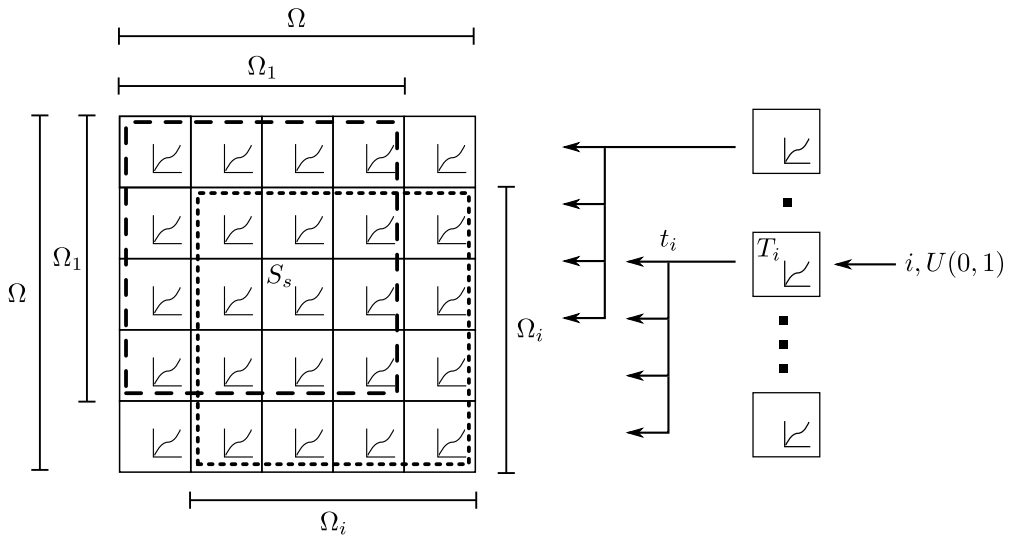


Figure 4.14: Schematics of importance sampling of image domain sub-intervals.

the partial derivative of f_{FFD} with respect to the control point i coordinate is non-zero only in their sub-intervals — Ω_i. The number of these intervals is equal to the number of control points (e.g. 7^3, 11^3, 19^3, etc.). Inverse transform sampling of CDF generally disallows effective sampling of arbitrary sub-intervals.

Fortunately, a number of these intervals' intersections is also finite, corresponding to disjunct intervals between knot points. Hence, CDF of intervals between knot points can be created and re-used in multiple image domain sub-interval CDFs that correspond to particular control points. The auxiliary data structure is needed to keep track of which knot intervals are part of which domain sub-intervals Ω_i (Figure 4.14). The data structure also holds a small CDF for each sub-interval Ω_i built over its knot interval CDFs, making it a two-level inverse transform sampling of discrete CDF sub-intervals.

Algorithm 10 describes non-uniform sampling with the use of the described data structure. This approach allows sub-intervals to be sampled without the need to store a CDF for each sub-interval, which would be memory-inefficient. The memory complexity of the proposed solution is $\mathcal{O}(n + m)$, where n is the number of samples in Ω and m is the number of control points (i.e. sub-intervals).

Stopping criterion

A stopping criterion of the convergence process should reflect the fact that the objective function is optimized in a stochastic manner. For this reason, the behavior of the stochastic objective function should be relativized of fluctuations caused by randomness in the integral computation in order to be able to

Algorithm 10: Two-level inverse transform sampling of discreete CDF in domain sub-intervals.

Input : i - index of interval Ω_i in Ω,
r - $U(0,1)$
Data : S - subdomain CDFs,
T_i - CDF of i-th interval,
t_i - interval i subdomain indices
Output: c - coordinates of a sample point,
p_i - probability of the point in interval i
begin
$j \leftarrow \text{binary\_search}(r, T_i)$
$s \leftarrow t_i(j)$, subdomain index
$r_j \leftarrow (r - T_i(j-1))/(T_i(j) - T_i(j-1))$, $U(0,1)$ in subdomain s
$k \leftarrow \text{binary\_search}(r_j, S_s)$
$c \leftarrow k$-th samples coordinates in subdomain s
$p \leftarrow P(S_j)(S_j(k) - S_j(k-1))$, probability in Ω
$p_i \leftarrow p/P(T_i)$, probability of k in interval i
end

set a certain threshold to a characteristic derived from it. Such a characteristic cannot be based on the changes between subsequent steps because they also have a propensity for unpredictable fluctuations. The obvious way to overcome this behavior is to average over several steps. Specifically, a line can fit a certain number of objective function values in a sequence and monitor its slope (see Figure 4.15). If it falls below a certain threshold, the convergence can be stopped. All is relative to the objective function value at the beginning of the iterative process.

4.2.5 Material

For the experiments, improvements to the registration algorithm were tested on CT images of human skulls, which were taken in hospitals as part of anthropological research. Together with CT scans of the contemporary population, three other samples from periods ranging from the 9-th to the 21-st century were collected (see Case study 5 for details). Figure 4.16 shows surface visualization of the skull of a specimen from the collection of the historical population and volume visualization of a living person from the contemporary population sample.

4.2.6 Benchmarking stochastic registration

It is difficult to assess the efficiency of the algorithm on real data. In order to deal with this issue, an artificial deformation is generated and applied to a sample of data. The algorithm is then created to recover the original shape by

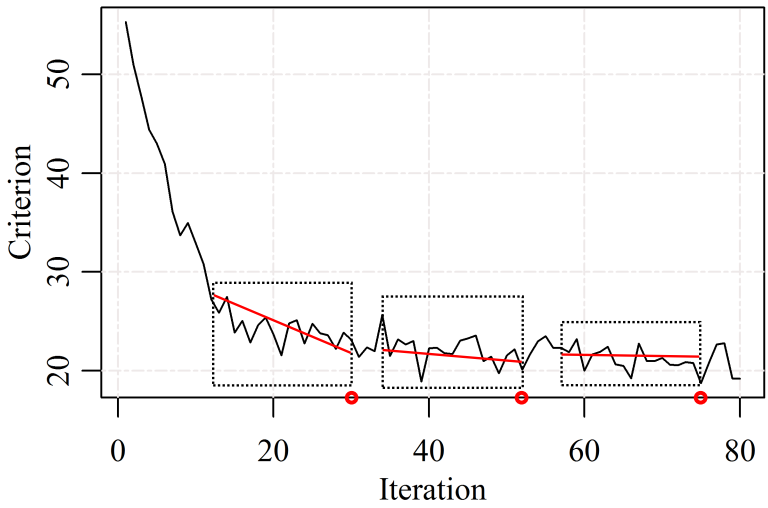
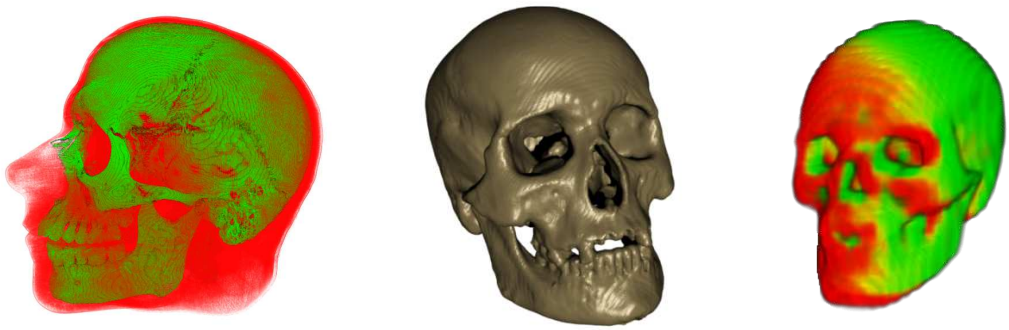


Figure 4.15: Criterion value behavior during the stochastic convergence process. The line (red) is fitted to the last 20 steps and its slope is compared to the given threshold.



(a) (b) (c)

Figure 4.16: A recent and historical dataset example: (a) visualization of CT volume data of a living patient; (b) the skull surface of a person who died hundreds of years ago; (c) a misaligned pair of down-sampled skulls (red — source; green – target).

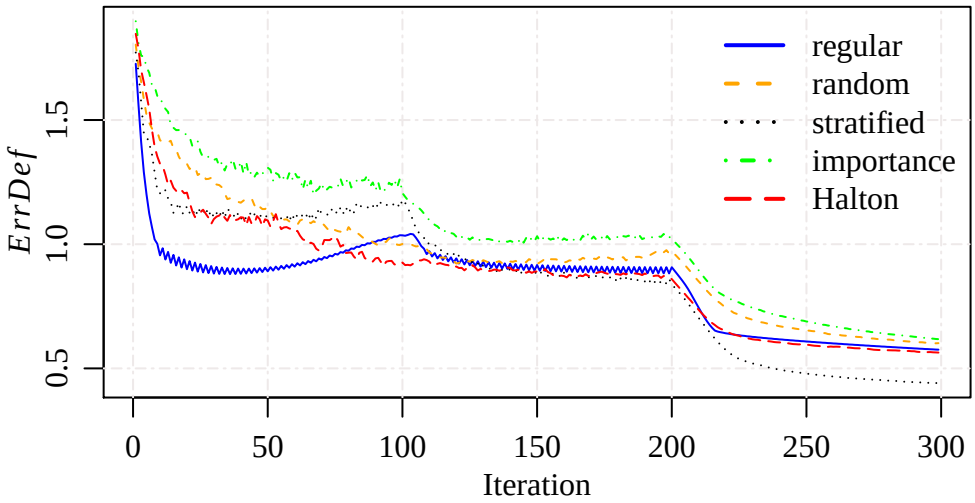


Figure 4.17: Error development throughout the convergence process, using the SSD data criterion and 1% of samples in stochastic methods. A noticeable transition between levels can be seen in the plot.

searching for an inverse deformation. By comparing the ground truth deformation T_{GT} with the deformation T found by the algorithm, an error measure is computed. This scenario is not equivalent to an intended real live application of the algorithm, but it is assumed that the results are highly correlated.

$$ErrDef(T_{GT}, T) = \frac{1}{N} \sum_{i=1}^{N} \|T_{GT}(x_i) - T(x_i)\|$$

The energy functional, including the regularization term with a weight $\alpha = 0.01$, is used in the algorithm, which is intended to prevent deformation from diverging. For the purpose of testing the sampling approaches, the original data sets are scaled from approximately 512^3 to $50 \times 50 \times 43$ voxels (Figure.4.16(c)). The algorithm uses three levels of multiresolution with $5 \times 5 \times 5$, $7 \times 7 \times 7$, and $11 \times 11 \times 11$ control point grids.

The algorithm with the SSD criterion is executed using all five sampling approaches discussed, with only 1% of the sample point. The results are shown in Figure 4.17. In order to observe the behavior of the algorithm with respect to various gradient approximation methods, a fixed number of steps are used with a constant step size. Stratified and Halton sampling outperform regular sampling, but generally all methods converge to the similar solution. The result of importance sampling is necessarily influenced by the data. For relatively sparse images with a lot of background elements, importance sampling should be more advantageous. It must be emphasized that although the algorithms perform the same number of iterations, stochastic sampling is 100× faster in real time than regular sampling. The behavior of several curves suggests that

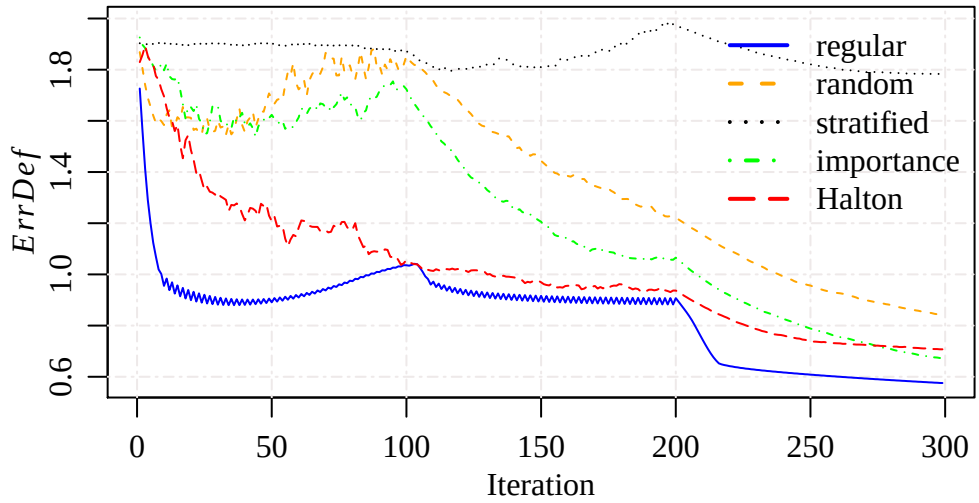


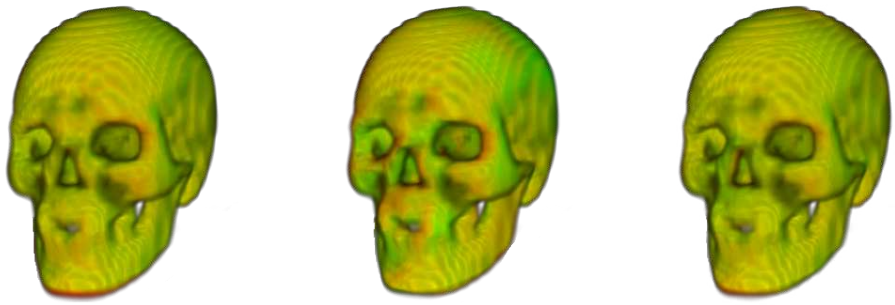
Figure 4.18: Error development throughout the convergence process, containing an SSD data criterion and 0.1% of samples using stochastic methods.

the convergence should be stopped earlier since the error gets worse if the algorithm overfits the SSD criterion. It is difficult to tune the stopping criterion with a noisy criterion function, but low variance gradient approximation can help significantly.

An extreme experiment with only 0.1% of samples was conducted with the intention of determining the difference in sampling strategies (see Figure 4.18). It is not surprising that sampling methods fail in the first phase when there are only 1-2 samples per volume. However, Halton sampling makes much better use of these samples than the other methods, and also performs consistently with the previous experiment. On the other hand, stratified sampling fails, which is caused by a fixed number of strata (sampling parameter) that are lower than the number of drawn samples. Therefore, a part of the space is not covered by the samples. Figure 4.19 shows results of the registration experiment, but the differences are very subtle.

By way of conclusion, properties of sampling methods used in stochastic optimization of non-rigid registration problem have been introduced. More importantly, it has been shown that the registration performs as well with a fraction of the sample points as with the traditionally regularly sampled image domain. It shortens, in real terms, the algorithm runtime from 5-6 hours to 3-4 minutes (1% samples) or 20 seconds (0.1% samples) with sub-optimal implementation, depending on the machine. In some cases, it performs even better because of the similar effect of simulated annealing, allowing it to jump out of local minima. The precision of stochastic optimization can be improved by slowly increasing the number of samples during the iterations.

There are other speed-up techniques that could also be employed. For



(a) (b) (c)

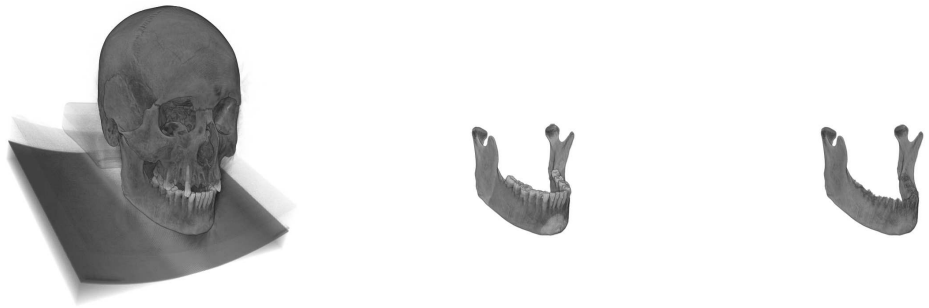
Figure 4.19: Sample skull data aligned by non-rigid registration using various sampling strategies after 3×100 iterations: (a) regular sampling ($DefErr$=0.58); (b) uniform random sampling with 0.1% of samples ($DefErr$=0.84); (c) importance sampling with 0.1% of samples ($DefErr$=0.67). In all figures, the source skull (red) overlaps the target skull (green), while the yellow color represents the overlap of tissues with similar density.

example, the often-used approach of control point inhibition might improve efficiency. The principle is to stop movement with some control points and also to stop computing partial derivatives with respect to these points' coordinates, provided the changes to their positions do not lower the value of the objective function. The decision when to stop moving the points must also be adapted to stochastic optimization.

The testing scenario described above was designed to compare the sampling approaches under artificial conditions. Especially, scaling data down 10 times does not produce practically useful results. Other aspects of the proposed stochastic registration algorithm are explained in the following practically-oriented case study.

Case study 12: Registration of the human jaw

In this case study, a registration method is demonstrated on the shape of a human jawbone acquired by standard medical CT in original resolution. A particular pair of images used in the case study is part of the Great Moravian sample (see Case study 5). Hence, only bone tissue is worked with in the images; no soft tissue, muscle, fat or cartilages are present. However, in order to perform accurate human jawbone registration, the jawbone must be segmented and separated from the rest of the bone tissue, with which it was acquired by the scanner, i.e. with the cranial bone and teeth (see Figure 4.16(b)). In particular, the teeth need to be removed due to their very different condition at the time of death; the extent of deterioration and damage can also be diverse.



(a) (b) (c)

Figure 4.20: Jawbone segmentation and pre-processing: (a) original dataset as obtained from a CT scanner (resolution $512 \times 512 \times 307$ with voxel dimensions $0.488 \times 0.488 \times 0.7$ mm); (b) cranial bone removed; (c) teeth removed. After segmentation, the whole volume can be trimmed to $260 \times 174 \times 171$ voxels.

These factors significantly affect results of registration. Figure 4.20 shows how a bare jawbone is extracted from original CT data. It can be performed using the segmentation algorithm suggested in Section 4.1.2 or manually if needed.

Jawbone tissue occupies only a fraction of the image space (approximately 5%), while a large part of the image samples actually does not contribute to the data criterion. To avoid unnecessary computation and in order to speed up the registration, the importance sampling scheme is chosen, which allows much fewer samples to be used when computing the criterion function and its derivatives than the number of discrete voxels in the volumes. The gradient magnitude of the highly blurred input image is chosen as a proposal probability density function, and is shown in Figure 4.21(b). It enables dense sampling of areas near the interface between bone and air. In contrast, the homogeneous areas inside the bone and primarily outside and far from the interface are sampled sparsely or not at all.

Since the registered image consists of a single type of tissue — bone — which has a sharp interface with background air and is easily obtainable by simple thresholding, it is possible to use a Euclidean distance criterion for the data term in the registration energy functional, according to Equation 4.8. The criterion will then precisely fit these interfaces so that the source and target images overlap perfectly in the ideal case. But it only works if combined with importance sampling of the area near the shape boundary. Moreover, it can be extended by adding the sum of squared differences between the overlapping image data to better overlap the tissues of the same density inside the bones. However, this option is not used further on. The threshold for distinguishing bone and air is set empirically so that artifacts of the partial volume effect and beam hardening (e.g. Figure 4.5(c)) are not visible. The effect of variation in the threshold is mostly visible inside the bone because of its porous nature.

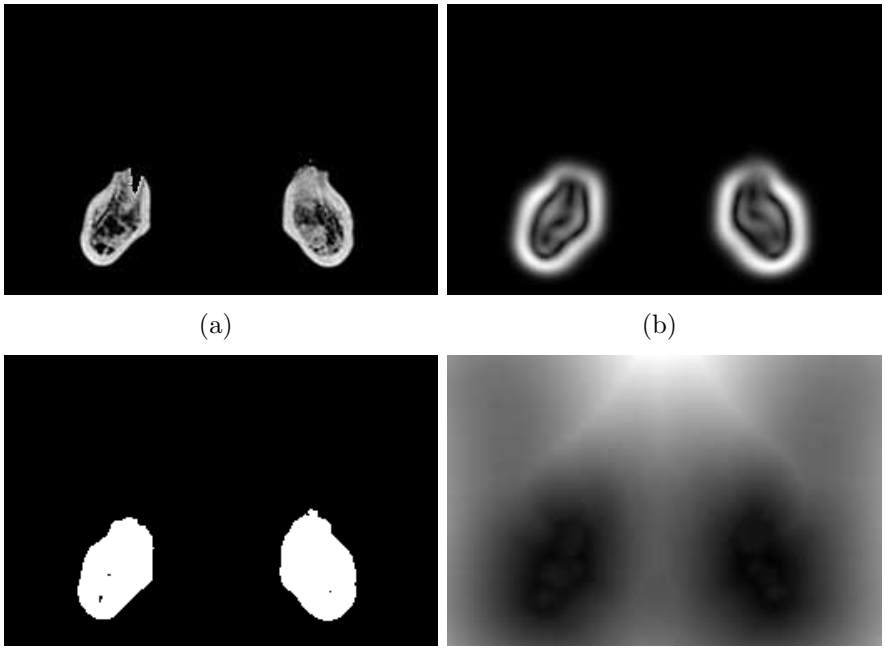
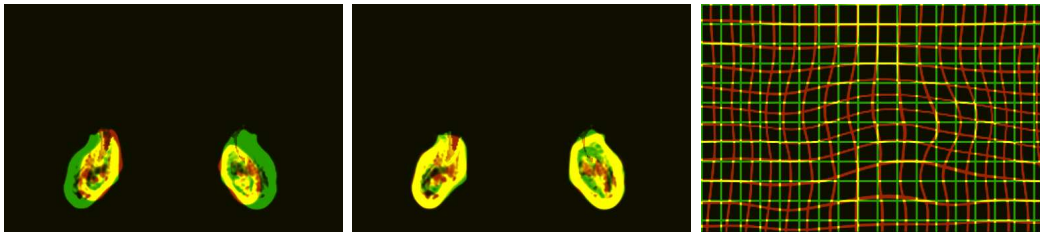


Figure 4.21: Jawbone registration criteria: (a) input slice; (b) thresholded image; (c) distance map, where black represents the most distant voxels inside the bone to the interface, and white represents the most distant voxel outside the bone; (d) sampling density.

Nevertheless, the impact of the precise threshold selection has a negligible impact on resulting registrations. Figure 4.21(c) shows thresholding of the original input image. Euclidean distances are pre-computed by the Euclidean distance transform (see Figure 4.21(d)).

The stochastic nature of Monte Carlo integration in the objective function as well as its derivative computation, which was introduced above, are not suitable for use with the numerical optimization procedures listed in Section 3.2.4 [page 50] other than steepest (gradient) descent. Because of the constantly changing value of E every time it is evaluated, it does not allow the local numerical optimization methods to work properly. In the following experiment, single random sampling of the space is performed and used for the entire optimization process in combination with L-BFGS (Okazaki, 2010). Therefore, the effect of error reduction by repeated stochastic approximation is lowered. This approach actually non-uniformly sub-samples the image data on top of the uniform sub-sampling by multiresolution.

The multiresolution approach is also employed in order to speed up the convergence as well as to lower the chance of the optimization procedure converging to the local minimum. The optimal number and resolution of levels are determined in further benchmarking. Figure 4.22 shows a single slice from



(a) (b) (c)

Figure 4.22: Results of jawbone registration: (a) a slice of volume data before non-rigid registration; (b) after non-rigid registration procedure convergence; (c) visualization of space deformation, enabling the source bone (red) to fit the target bone (green). Yellow pixels represent overlaying high-density structures in both the source and target bones.

a successful registration of a pair of human jaws. Figure 4.23 shows the same result in 3D, visualized by DVR.

For the purpose of comparing various parameters of registration with respect to registration results, and in order to determine the optimal parameters of a single registration procedure, the error measure must be defined

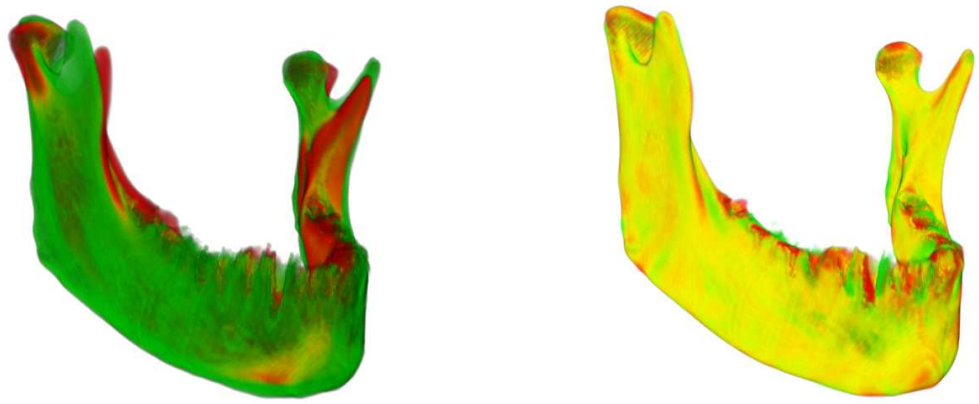
$$Err = \frac{1}{|I|} \sum_{x \in I} |S(x) - T \circ f_{FFD}(x))|^2$$

where S and T are the source and target image, respectively, and I is a set of the individual voxel's integer coordinates in full resolution. The value of T is linearly interpolated in case of transformation to non-integer values by B-spline function f_{FFD}.

Following the experience with mesh registration in Section 3.2.4, registration of the human jaw can be improved by incorporating additional user-defined constraints, which should be added to the data term. With respect to the demand for maximum precision, it is recommended to use as much input as possible, especially since expert-placed landmarks for the dataset are available (see Figure 2.1(a) for landmark configuration L). The landmark explicitly states which points correspond to which between the specific image data. The degree with which landmark constraints are considered is given by the value of multiplicative weight factor w_l, i.e. E is extended as follows

$$\begin{aligned} E(f_{FFD}) &= \int_\Omega (1-\alpha) f_d \circ f_{FFD}(x) + \alpha f_r \circ f_{FFD}(x) dx + \\ & \quad w_l \frac{1}{|L|} \sum_{i=1}^{|L|} \|L(S,i) - f_{FFD}(L(T,i))\|^2 \end{aligned}$$

Apart from that, a value of the landmark term alone can be used to measure the quality of the fit as an error measure, Err_C.



(a) (b)

Figure 4.23: Volume visualization of registration results: (a) a rigidly aligned pair of jaws; (b) a pair of jaws after non-rigid fitting of the source (red) to the target (green) volume. The yellow color represents ideally overlaying tissues of the same density.

data term	**time (s)**	Err	Err_C
SSD	741	$1.82 \cdot 10^{-3}$	20.3
EDT	329	$2.88 \cdot 10^{-3}$	18.1

Table 4.1: Comparison of stochastic gradient descent and sub-sampled L-BFGS for non-rigid B-spline registration of a jawbone. The landmark weight is set to zero ($w_l = 0$), regularization is also set to zero ($\alpha = 0$) and sampling is set to 0.5% of the total samples ($s = 0.005$).

In the first test, two registration algorithms are compared: the first uses an SSD data term and is optimized by stochastic gradient descent; the second uses an EDM data term and L-BFGS non-stochastic optimization. Both algorithms use importance sampling for generating samples, as explained above. Table 4.1 shows a comparison of these two methods with identical settings of other parameters (regularization, landmark weight, sample count).

In further tests, the EDT-based algorithm is evaluated with respect to various parameters. Table 4.2 shows how a value of w_l impacts on error measures and convergence time. It seems that when the weight is too high, it has a significant influence on overall convergence time and error measure. In contrast, omitting the constraints completely does not make convergence much worse than in the optimal case, which is approximately 0.1.

Regularization is often the important part of non-rigid registration, espe-

landmarks weight	**time (s)**	Err	Err_C
100	288	$8.76 \cdot 10^{-3}$	$3.9 \cdot 10^{-5}$
10	741	$5.31 \cdot 10^{-3}$	$5.5 \cdot 10^{-4}$
1	914	$3.72 \cdot 10^{-3}$	$9.1 \cdot 10^{-3}$
0.5	930	$3.81 \cdot 10^{-3}$	$1.6 \cdot 10^{-2}$
0.1	1395	$3.35 \cdot 10^{-3}$	$4.9 \cdot 10^{-3}$
0.01	1191	$3.18 \cdot 10^{-3}$	$2.1 \cdot 10^{-1}$
0	1285	$3.00 \cdot 10^{-3}$	$2.2 \cdot 10^{1}$

Table 4.2: Landmark weight in a jawbone registration evaluation. Regularization is set to zero ($\alpha = 0$) and sampling is set to 2% of the total samples ($s = 0.02$).

regularization weight	**time (s)**	Err	Err_C
0.1	749	$2.93 \cdot 10^{-3}$	22.2
0.01	1007	$2.94 \cdot 10^{-3}$	17.3
0.001	1355	$2.84 \cdot 10^{-3}$	16.2
0	320	$2.88 \cdot 10^{-3}$	18.1

Table 4.3: Regularization weight in jawbone registration evaluation. The landmark weight is set to zero ($w_l = 0$) and the sampling ratio is set to 0.5% of the total samples ($s = 0.005$).

cially in cases of transformation function, which allows unnatural space warping. This can occur in cases where there are very large differences between source and target images. This does not seem to be the case with human jawbones, which are all alike; neither does it occur in B-spline transformation, given its intrinsic regularization properties. However, the possibility of sliding surfaces in the source image along surfaces in the target image should be avoided, along with inconsistent local deformation caused by the fine internal structure of the bone containing cavities. Table 4.3 shows how the regularization weight, which enforces low second derivatives and the smooth behavior of the transformation function, influences convergence time and error measure. It should be noted that in the implementation, the computation of the regularization term is turned off when it is canceled out by setting the weight to zero. The impact on convergence time is significant because the computation of the regularization term is relatively expensive and makes the overall convergence time approximately $2.5 - 4\times$ longer. On the other hand, a high regularization weight stops convergence sooner with a slightly negative impact on error measure. Even so, regularization does not seem to be very important

sample size (%)	**time (s)**	Err	Err_C
2.0	1291	$3.00 \cdot 10^{-3}$	21.9
1.0	535	$3.02 \cdot 10^{-3}$	18.5
0.5	320	$2.88 \cdot 10^{-3}$	18.1
0.4	282	$2.93 \cdot 10^{-3}$	19.3
0.3	194	$3.23 \cdot 10^{-3}$	21.2
0.2	233	$2.85 \cdot 10^{-3}$	20.0
0.1	133	$3.27 \cdot 10^{-3}$	20.8

Table 4.4: Sample size in jawbone registration evaluation. The landmark weight is set to zero ($w_l = 0$) and regularization is also set to zero ($\alpha = 0$).

# levels	**init. grid resolution**	**time (s)**	Err	Err_C
2	11	243	$3.03 \cdot 10^{-3}$	16.1
3	7	338	$2.88 \cdot 10^{-3}$	18.1
4	5	469	$2.87 \cdot 10^{-3}$	21.1

Table 4.5: Multiresolution in jawbone registration evaluation. The landmark weight is set to zero ($w_l = 0$), regularization is set to zero ($\alpha = 0$), and the sampling ratio is set to 0.5% of the total samples ($s = 0.005$).

in this particular registration task.

By contrast, a more influential parameter with respect to speed and precision is the sampling ratio. In order to make registration usable in clinical applications, the optimal ratio must be sought for the particular registration task to keep the error measure and computation time low.

Another important attribute of the process is the number of levels in the multiresolution scheme and the resolution of the control grid. According to Table 4.5, it seems that additional levels do not greatly contribute to error measure. Nevertheless, the actual optimization criterion concerns minimizing the difference in Euclidean distances to shape interfaces. This is better achieved using three levels in the multiresolution scheme with $7 \times 7 \times 7$, $11 \times 11 \times 11$ and $19 \times 19 \times 19$ control points, respectively, represented by 1029, 3993, and 20577 individual variables. Four or more levels are unnecessary for resolution of the input data. It probably causes misalignment in low resolution on the first level, which is not corrected well on a higher level. Two levels of multiresolution do not overfit the thresholded data, which results in a lower landmark error measure (Err_C), but also in poor alignment of image intensities (Err).

4.3 Template-based group registration

Statistical analysis of volume images requires construction of correspondence across the whole sample. In this section, the registration procedure described above is applied to register multiple images to a common frame in order to obtain mutual correspondences. Registration of multiple volume images into a common frame of reference has gained a considerable amount of attention in recent years, especially in the field of neurology (Gholipour *et al.*, 2007). Volume images of the human brain can be taken using various imaging techniques (CT, MRI, PET), while different phenomena can be studied on a sample of patients. Based on these samples, various types of atlases can be created with the use of specialized registration methods.

There are two approaches to group registration:

- Templated fitting — one specimen is chosen as a template and the rest of the group is registered to this template (Rueckert *et al.*, 2001) or the template is registered to each specimen. An obvious disadvantage is that the group registration is biased by the choice of template.

- Groupwise registration — the sample is continuously fitted to the mean template, which is updated until it stops changing (Noblet *et al.*, 2012). There are various modifications of this approach (Bhatia *et al.*, 2004; Balci *et al.*, 2007).

Although the second group of methods is considered superior, template fitting is used for the purpose of jawbone group registration. In the first step, all images are aligned rigidly using any rigid registration method. If the landmark data are present, GPA is preferred. In fact, GPA is a landmark-based variant of groupwise registration; thus, at this point the images are rigidly aligned to the common mean.

Moreover, GPA can align the shapes to remove differences in orientation and position with respect to landmarks, or even to remove differences in size. This depends on whether the subsequent analysis interpretation expects the size component to be included or only shapes with size differences removed. Alternatively, size, position, and orientation can be removed by aligning all corresponding points found by non-rigid registration.

Rigidly aligned images are a starting point for non-rigid B-spline-based registration, as described above. The group registration pipeline is summarized in Algorithm 11.

Case study 13: Human jaw variability of a Great Moravian sample

Group registration of shape samples was demonstrated on a collection of jawbones from the Great Moravian period in Case study 5. These bones are unique because of their historical value and well-preserved condition. All bones in the

Algorithm 11: Template-based group registration.

Input : Images I_i, Landmarks L_i, Index of template image $t \in \{0..n\}$
Output: Aligned images A_i, B-spline based transformations t_i of I_t to I_i
begin
  Run GPA on L_i to get aligned landmarks M_i
  for $l := 1$ **to** n **do**
    Transform I_i according to rigid transformation defined by transition from L_i to M_i
  end
  for $l := 1$ **to** n **do**
    Run B-spline based non-rigid registration looking for optimal transfomation t_i that fits A_t to A_i;
  end
end

collection were treated as demonstrated in Figure 4.20. Figure 4.24 shows an example of rigid alignment followed by non-rigid fitting in a pair of jawbones.

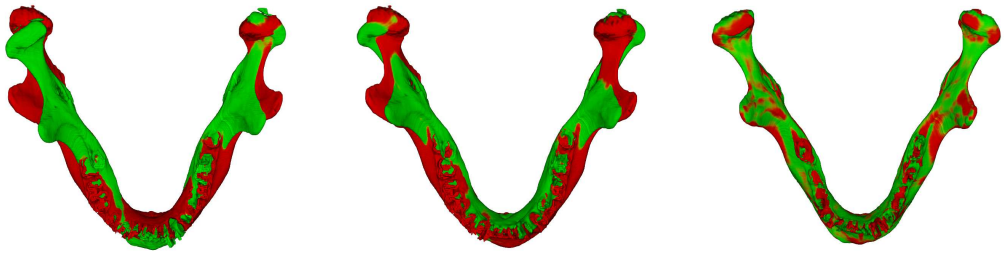
Group registration of shapes represented by volume data was performed by preserving the size of the bones, which were normalized to a mean size. The results can be used to study the following features:

- Mean shape and mean form — the mean shape and form are obtained by averaging deformation fields defined by all transformations t_i and by applying them to the template volume. Shape and form are distinguished according to whether the transformations have the uniform scaling component removed, as mentioned above. Moreover, if the sample is divided into groups according to a certain property (e.g. sex, age, etc.), the averaging can be performed in these sub-groups separately (see Figure 4.25(b)).

- Local variation — local variation in positions of corresponding points is computed by evaluating the variation of deformation fields per voxel of the template image. The magnitude of the variance vectors is shown in Figure 4.25(c).

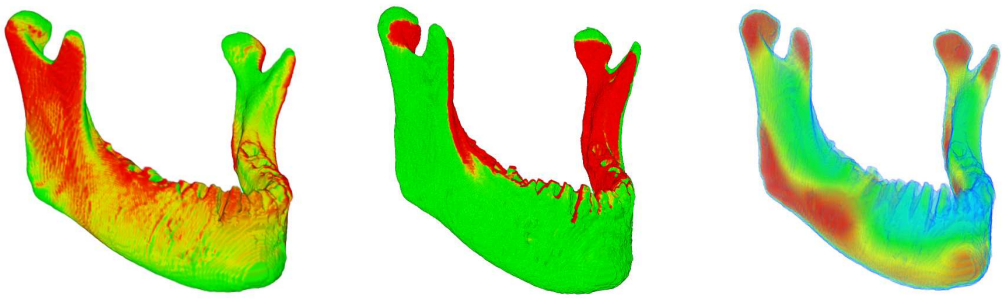
4.4 Statistical modeling of volume data

Once correspondence in the volume data sample is constructed by non-rigid registration, statistical processing of the data can then take place. One field that focuses on element-wise statistics in aligned volume images, mostly in neuroimaging, is called *voxel-based morphometry* (Ashburner and Friston, 2000). However, this particular methodology has received criticism due to the potential for bias results in relation to imprecise registration and Gaussian smoothing (Bookstein, 2001).



(a) (b) (c)

Figure 4.24: Jawbone group registration: (a) an individual jawbone before alignment (green) and a template (red) in reference position; (b) each jawbone (green) is rigidly aligned to an optimal position on the template (red), where, in this case, alignment is computed based on the GPA-aligned landmark; (c) transformation which fits the template (red) to each jawbone (red) is found so that the template can be deformed to the shape of the particular jaw.



(a) (b) (c)

Figure 4.25: Jawbone mean shapes and forms from the Great Moravian sample: (a) the mean shape of the sample sub-group — male (green), female (red); (b) the mean form of the sample sub-groups, where the male mean form (green) is clearly larger than the female form (red); (c) local corresponding points variance magnitude.

On the other hand, statistical analysis of shape geometry is based on highly correlated positions of densely corresponding points. Therefore, statistical shape modeling of volume data can be performed in the same way as it is performed for landmark and mesh data. Apart from vector fields obtained by registration, there are other representations that can be used for statistical shape analysis of volume images, e.g. level-set methods as suggested by Tsai *et al.* (2003).

Statistical modeling of volume data, in connection with skeletal remains and in its application to physical anthropology, is unexplored territory. Zheng and Lösch (2013) recently attempted to create a statistical model of a skull from a sample of 25 specimens. They used a diffeomorphic demons registration algorithm (Vercauteren *et al.*, 2007) to find correspondences in an attempt to use the model for incomplete shape reconstruction. However, only preliminary results were presented.

4.4.1 Shape variability

Variability of corresponding points is analyzed by decomposition into independent trends by Principal Component Analysis. PCA is applied directly to coordinates of corresponding voxels. The disadvantage here is that volume data have the highest dimensionality of all representations mentioned in this work. This means that a particular specimen is represented by a vector three times longer than the number of voxels of the volume image, which could reach tens of millions. In order to reduce dimensionality, parts of the volumes can be masked out in the template depending on the region of interest. For example, voxels clearly belonging to the background can be identified by simple thresholding and masked out. Another example is via analysis of voxels located along the interface between air/background and high-density tissues. Such voxels can be identified by a sequence of thresholding, flood-filling, and morphological operations. This approach reduces volume analysis to an analysis of closed surfaces. Figure 4.26 shows examples of various template masks.

Positions of voxels that are included in the analysis are inputs of PCA. Similar to mesh data, HDPCA (Section 3.3 [page 69]) is used in cases where dimensionality is higher than the number of specimens. The output of the analysis is a set of vectors $\{w_j\}_{j=1}^{n}$, or modes of variation. Linear combinations of these vectors can reconstruct deformation field t_i, which, if applied to the template, results in the shape of a particular specimen i.

$$t_i = \bar{t} + \sum_{j=1}^{n} \alpha_{i,j} w_j$$

Coefficient vectors α_i of these deformations represent the position of shape i in the shape space. Dimensions of the shape space display decreasing importance in the reconstruction when associated with the specified amount of information. The specific amount of information in relation to shape variability is

(a) (b) (c)

Figure 4.26: An example of various template masks: (a) masking only an interior of the tissue of interest, representing 406k elements (approx. 5% of the volume); (b) masking only an outline of the object, representing 61k elements (approx. 0.8% of the volume); (c) a thick object outline, representing 371k elements (approx 4.8% of the volume).

shown in Figure 4.27. Using the broken-stick criterion (see Section 3.4.1 [page 73]), it was found that the first nine principal components were statistically significant for shape reconstruction. It is likely that the rest of the components model noise.

Quickly decreasing variability and a relatively low number of statistically significant components are a good sign that the model represents a low number of important trends, which are hopefully associated with a simple geometrical interpretation. For clarity, statistically significant components and associated variability of shape and form in landmark-based and voxel-based models are summarized in Table 4.6. In comparison to shape variability distribution of the landmark-based model's components, the voxel-based model presented here displays very similar behavior. According to the broken-stick criterion and in the case of shape, a number of statistically significant components is identical for both the landmark method and the voxel method. The voxel-based model of form has nine significant components while the landmark-based model has only eight.

Projection of the space represented by two of the most important principal components of form variability is shown in Figure 4.28 as well as the assignment of individual specimens to male and female sub-groups. The size of the jaw is substantially important when distinguishing between sexes, which can be seen in the ordering of the sub-groups along PC1, which is associated with size (represented by centroid size measure) by regression ($p < 0.0001$, $R^2 = 0.2556$, $F = 28.46$, $df = 1, 79$).

Equivalently, scores in the first two shape components are shown in Figure 4.29. Both sexes almost overlap, whereas only a subtle separation can be observed along the PC1 axis.

Examining the most important (first) dimensions can provide an idea of the most pronounced, independent trends in the sample. The effect of a particular component on a form's appearance is shown in Figure 4.30, which displays the color-coded magnitude of the deformation modeled by the principal com-

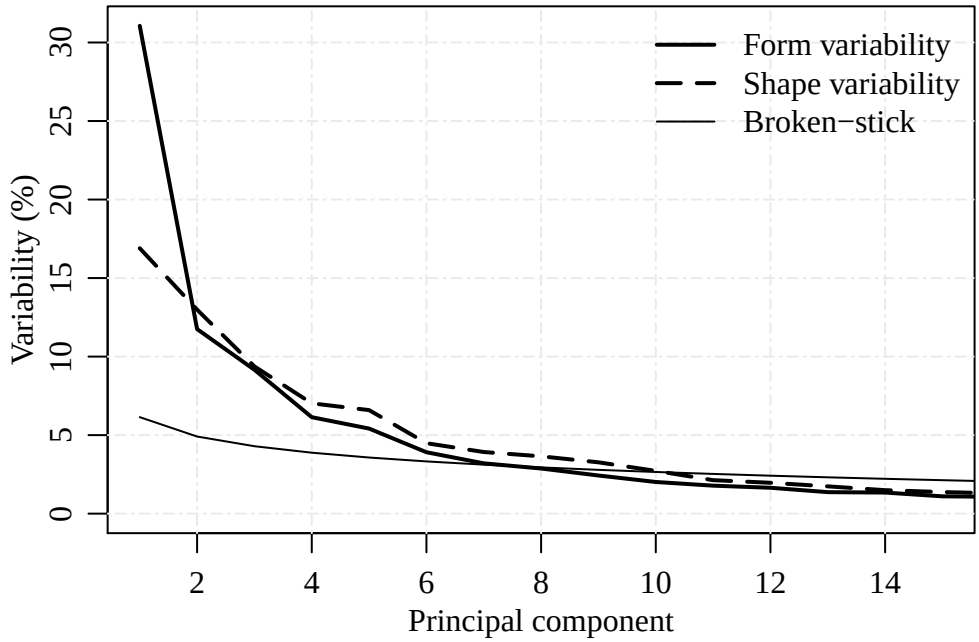


Figure 4.27: Amounts of variability (information) in individual principal components (black). The broken-stick criterion (red) reflects a minimum amount of statistically significant information for the component.

PC	form variability		shape variability	
	voxel (%)	**landmark** (%)	**voxel** (%)	**landmark** (%)
1	31.06*	33.50*	16.90*	18.74*
2	11.75*	11.83*	12.99*	13.70*
3	9.16*	8.23*	9.37*	8.16*
4	6.14*	5.81*	7.03*	6.80*
5	5.42*	4.90*	6.59*	5.95*
6	3.91*	4.16*	4.49*	5.47*
7	3.20*	3.35*	3.92*	4.30*
8	2.87	2.99*	3.65*	3.55*
9	2.43	2.48	3.27*	2.95*
sum	75.95	77.25	69.35	69.62

Table 4.6: Variability of statistically significant components of jawbone shape and form.

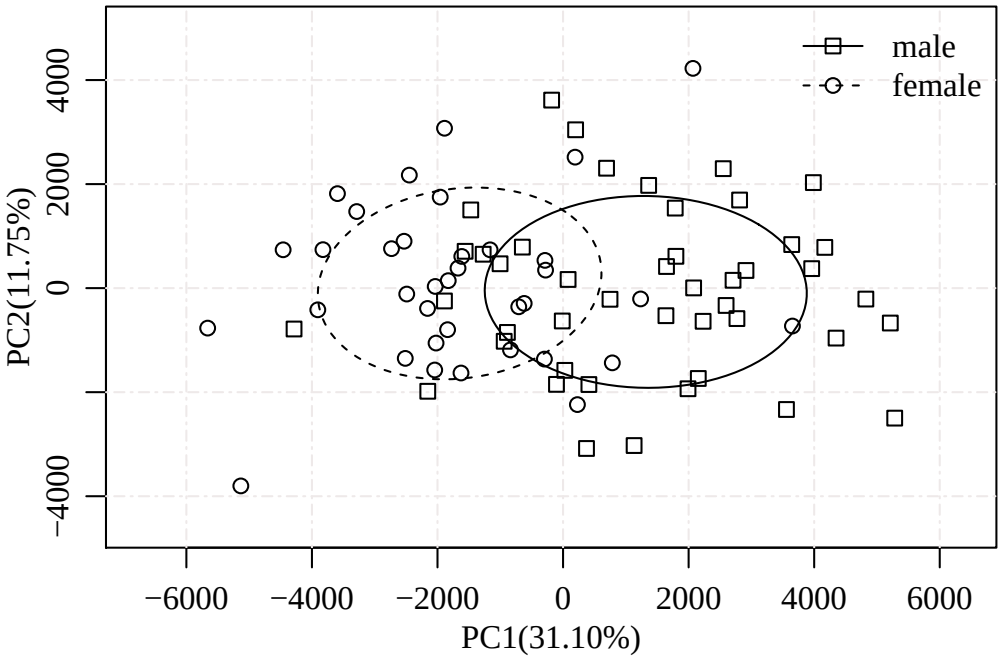


Figure 4.28: Jawbone form variability in the first two principal components. The ellipses represent 50% confidence intervals around the group centroids.

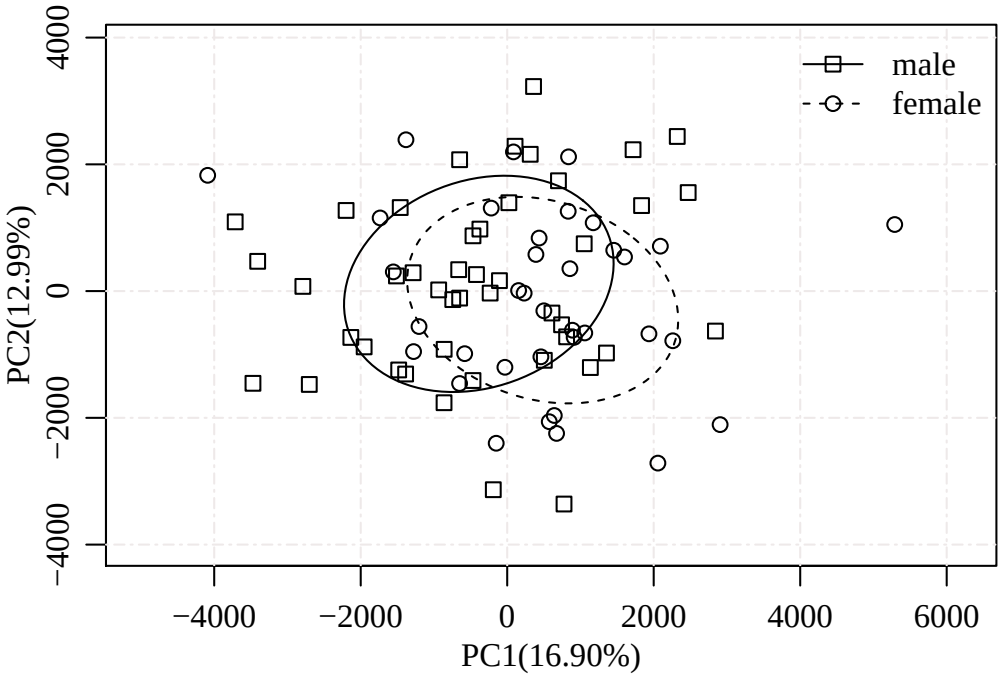
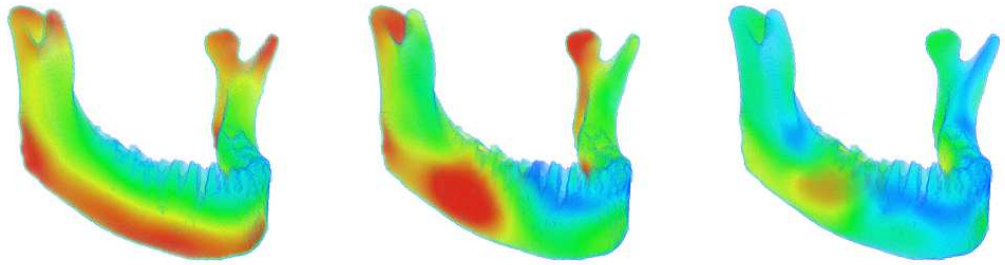


Figure 4.29: Jawbone shape variability in the first two principal components. The ellipses represent 50% confidence intervals around the group centroids.



(a) (b) (c)

Figure 4.30: Color-coded independent trends in jaw form variability: (a) the magnitude of PC1, which is the component associated with size; (b) the magnitude of PC2; (c) the magnitude of PC3.

ponent. For example, the first component models the lower part of the jaw, as well as the partially posterior condyle and coronoid process. The second principal component affects the condyles, posterior ramus, and middle part of the jaw body. Both first components are approximately symmetrical. By comparison, the third component affects the middle part of the jaw body much more on the left side than on the right side, which does not happen when only a symmetric component of the shape is analyzed.

4.4.2 Sexual dimorphism

It is difficult to interpret individual principal components with respect to some other factors (e.g. sex, age group, population) because such factors are exhibited in all components at once. For example, it should be avoided, to summarize, the difference between two sub-groups in one component as a prolongation while the second component represents widening because the combined shape change will not probably look like prolongation nor widening.

Some studies quote a high accuracy for sex estimation based on jawbone, for example, up to 97.1% (Bejdová *et al.*, 2013). However, sexual dimorphism is relatively difficult to prove in only the shapes of complex bones like the pelvis or jaw. Usually, size is the major factor, which is correlated with a particular sex; however, if size is not considered, discrimination rates are lowered considerably.

It was shown by Bejdová *et al.* (2013) and in Case study 5 that sexual dimorphism does exist when based on landmark methodology. Moreover, the discrimination rate of sexes in the symmetric component of jawbone shapes was relatively high, reaching around 80% in some populations.

Case study 14: Sexual dimorphism in a Great Moravian jaw sample

Voxel-based methodology was used to evaluate sexual dimorphism of a Great Moravian jaw sample (Case study 5). The high dimensionality of voxel shape

representation was already reduced to nine statistically significant components, i.e. seven numbers that are able to reconstruct a particular shape with approximately 69.35% of the overall information. In the first step, multivariate data are positively tested for statistical significance of differences between males and females by the multivariate Hotelling's two-sample T^2 test ($p = 0.0125$, $T^2 = 25.81$, $F = 2.5771$, $df = 9, 71$).

Subsequently, the ability to discriminate sexes in the sample of jawbones is demonstrated by the success rate in classification of shape and form variables. Also, in this case, shape variables are represented by the scores of statistically significant components. For evaluation of shape and form discrimination, the following classifiers are chosen:

- Linear discriminat analysis (LDA) — an assumption of a normally distributed sample is made, after which optimal weight $\vec{w}$ for linear discrimination function $\vec{w} \cdot \vec{x}$ is found (Bishop, 2006, page 186).

- Naive Bayesian classifier — an assumption of the classifier for individual features independence is fulfilled since they are PCA scores; hence, the classifier can optimally benefit from the setup (Bishop, 2006, page 380).

- Support vector machines (SVM) — radial basis $\exp^{\gamma|u-v|^2}$ is used with $\gamma = 1/p$, where p is the number of components involved (Cortes and Vapnik, 1995). In this configuration, the classifier is highly over-fitted and non-generalized.

- Artificial neural networks — network with one hidden layer and a softmax activation function in the output layer is used. The number of neurons in the hidden layer impacts significantly on the behavior of the classifier. The recommended number of neurons for the dimension of the input is about 4. Also, 10 neurons are tested. A higher number of neurons leads to over-fitting, as evidenced by its high posterior rate (Bishop, 2006, page 225).

Table 4.7 shows the posterior discrimination rate as well as the more objective leave-one-out cross-validation rate of sexual dimorphism in jawbone form (with size preserved). Table 4.8 shows results of discrimination by sex for jawbone shape.

It seems that the simplest linear classifier (LDA) has the better generalization property, whereas non-linear classifier scores (SVM, ANN) often over-fit when not configured properly. Although the naive Bayesian classifier offers a non-trivial decision boundary (given by the mixture of Gaussian probabilities) and maximizes posterior probability, it does not outperform the linear classifier. This would indicate that sexual dimorphism of the human jawbone shape appears to be linearly separable.

Using the LDA classifier's weights $\vec{w}$ in a linear combination of corresponding principal components (modes of variation), extreme cases of both classes'

classifier	posterior rate (%)	cross-validation rate (%)	female	male
LDA	82.72	79.01	28/8	36/9
Naive Bayes	82.72	79.01	27/9	37/8
SVM(radial kernel)	86.42	69.14	21/15	35/10
ANN(hidden=10)	87.65	77.78	26/10	37/8
ANN(hidden=4)	80.25	79.01	31/5	33/12

Table 4.7: Jawbone form sex classification results using various classifiers. The last two columns specify the number of females/males classified in the same class or separately, respectively.

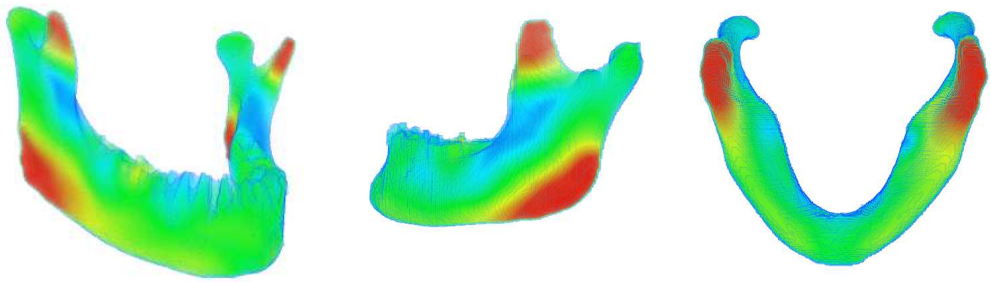
classifier	posterior rate (%)	cross-validation rate (%)	female	male
LDA	70.37	67.90	22/12	33/14
Naive Bayes	74.07	51.85	15/18	27/21
SVM(radial kernel)	82.72	50.62	10/14	31/26
ANN(hidden=10)	96.30	56.79	21/20	25/15
ANN(hidden=4)	72.84	65.43	24/16	29/12

Table 4.8: Jaw-bone shape sex classification results using various classifiers. The last two columns specify the number of females/males classified in the same class or separately, respectively.

representatives can be constructed. Alternatively, visualizing color-coded magnitudes of a vector field, which this linear combination represents, shows where the most important differences between both classes are. Figure 4.31 shows the differences between males and females reflected in the form of a jaw sample from the Great Moravian period.

In consensus with the results reported by Bejdová *et al.* (2013), where landmark methodology was employed using voxel-based methodology, sexual dimorphism was also proved in the Great Moravian jaw sample. By comparing Figures 2.7(b) and 4.31, it can be seen that the dimorphism manifestation reported by both methodologies is similar.

In order to interpret the results of discrimination, it must be noted that the published results achieved using landmark-based methodology were evaluated from symmetric components of data that were also reduced of allometric effects. In other words, part of the variability that is not related to sex was filtered out. Hence, voxel-based discrimination described in this section should be compared to equivalent results mentioned in Case study 5, which recorded a 69.14% cross-validation rate for discrimination by sex with respect to shape and a 79.01% rate with respect to form. This is only a 1.24% lower rate for



(a) (b) (c)

Figure 4.31: Color-coded effects of sexual dimorphism on a jaw-bone shape derived from the Great Moravian sample. The red color represents voxels whose positions are highly influenced by sex, the green color indicates mild influence, and the blue color indicates low influence: (a) free angle view; (b) sagittal view; (c) axial view.

shape, which is equivalent to one misclassified specimen. Discrimination success rates with respect to form are equal for landmark-based and voxel-based methodologies.

It is reasonable to claim that the new methodology approximately achieved the same discrimination results. The same results can be attributed to the fact that the jawbone was well-covered by landmarks, which completely reflects its variability in the size of the available sample. It is also reasonable to assume that the voxel-based discrimination results could have been further improved upon if the symmetric component of the voxel data had been analyzed. Separation of symmetrical and asymmetrical components under voxel-based analysis is one of the possible directions for future research.

4.5 Conclusions

In this chapter, several methods for use in morphometrical analysis of voxel data have been described. The motivation for employing voxel data is to incorporate more information into the analysis and to avoid making subjective decisions when choosing landmark points, lines, surfaces, or any other kind of low-dimensional primitives.

Even though input for analysis can take the form of raw data since it is issued from a scanner (e.g. CT, MRI), it is also recommended to perform a preprocessing step to isolate the shape that is to be morphometrically analyzed, i.e. segment it from the background in the image.

In the first part, a method for segmenting organs was presented. It shares a common mathematical basis with other parts of this thesis, particularly the B-spline function for shape representation. It also presents a variational

formulation of the problem as well as a numerical method for its solution. The author has applied this algorithm to measure the volume of organs in the human body (Krajíček *et al.*, 2007; Krajíček, 2007, 2008b).

The following segmentation method is targeted at processing specific data — jawbones images, which are pivotal for the rest of the chapter. The author has contributed by employing locally-adaptive intensity priors using region-based level-set segmentation applied to teeth and jawbone segmentation (Krajíček *et al.*, 2014).

Correspondence problems in voxel data are solved using a non-rigid registration framework, which has been improved in several ways by the author:

- Sampling as a speed-up technique to decrease the execution time of the algorithm from hours to minutes using contemporary machine (Krajíček *et al.*, 2011). As a result of this process, a stochastic registration algorithm variant was created.

- A suitable data criterion for bone registration, based on pre-computed Euclidean distance maps and non-uniform sub-sampling.

- Limited-memory BFGS — a cutting-edge numerical procedure for non-linear optimization was chosen to speed up numerical optimization. In combination with the previous points, a non-uniform sub-sampling registration algorithm variant was produced.

- Parameters of the whole pipeline were optimized by a sequence of tests and experiments.

The subsequent application of the registration method to template-based group registration of the Great Moravian jaw sample, along with statistical analysis of the deformation field, proves the applicability of the pipeline to similar tasks in physical anthropology and biomedicine.

Future work on the framework could be aimed in the following directions:

- Extend the registration procedure to register objects with the use of multiple interfaces (bone, skin, fat layers).

- Analysis of asymmetry, which will allow new phenomena to be studied and which will improve discrimination, as mentioned above.

- Reconstruction of missing data based on information collected in the statistical model in accordance with the ideas presented in Chapters 2 and 3.

Chapter 5

Conclusions and future work

> In the End, we will remember not the words of our enemies, but the silence of our friends.
>
> Martin Luther King, Jr.

Studying the shapes of living things and how they change have always been important problems in life sciences. New methods of shape description emerge with the new technologies for their acquisition. This thesis describes shape analysis methods that use two novel modalities — triangular meshes and volume imaging — and which are relevant to the field of physical anthropology and their specific tasks.

Up until now, the common approach of quantitatively analyzing various phenomena associated with the shape and form of things was connected to landmark analysis and geometric morphometric methodology (Section 2.2). It has been shown that certain shapes are difficult to describe using landmarks, which accounts for the different descriptions of shapes, e.g. curves, surfaces, or volume space sampling, where necessary (Section 2.5). The basic principles of data mining and the application of principal component analysis can be both transferred to triangular surfaces as well as volume images (Sections 3.4, 4.4). However, it is not easy to obtain correspondences using these representations.

Registration methods have been investigated with an emphasis on the speed of processing since they should be used for registering tens or hundreds of specimens. Thin-plate spline-based registration (Section 3.2.3) used in basic dense correspondence algorithms (Section 3.3) is fast, but it does not deliver consistent correspondence matching. The state-of-the-art coherent point drift algorithm exploits fuzzy correspondence in exchange for computationally expensive processing, which was possible to improve (see Section 3.2.5). B-spline-based non-rigid registration seems to be the most versatile since it can be used for mesh as well as for volume images registration, and also allows incorporating various criteria such as landmarks or nearest neighbor distances. Landmarks can still be used as a support for data-driven registration, while giving the user

confidence in the registration process and allowing exact fits of stable points to be controlled. Sampling appears to be an effective approach as a way of reducing registration time by orders of magnitude (Sections 3.2.4 and 4.2.2) while also maintaining accuracy.

Different kinds of relations can be uncovered by studying the asymmetry of shapes. The landmark-based method was described and used in a research study in Section 2.4. However, the methodology for mesh asymmetry needs to be developed (Section 3.5), but it has already found an application in actual research.

The method in which pairs of related data are registered beforehand, enabling their relations to be analyzed, has also found a use in applications (Section 3.6).

The path for future work on the non-rigid registration framework is two-fold. First, groupwise registration is thought to be superior to template-based registration since it removes bias caused by the selection of a particular template. It would be reasonable to prove this assumption in the case of an actual morphometric task. Template-based group registration has the advantage of using incremental model-building, whereas groupwise registration requires a complete set of specimens at the beginning.

Second, it would be beneficial to further improve registration algorithms in terms of speed. For example, by employing hardware acceleration using sampling approaches introduced in this work, another speed-up of at least one order of magnitude would be reasonable to expect. This approach has already been demonstrated in case of CPD (Section 3.2.5).

Apart from improving the registration framework, it would be especially useful to continue with the application of the shape variability analysis framework in actual research. Not only might this motivate further improvement but it might also have a multiplicative effect on the utilization of research presented in this text.

Extending asymmetry analysis to volume images might improve discrimination results in the Great Moravian jawbone case study. The results of shape analysis without the removal of asymmetry effects are not better than results acquired using landmark-based methodology.

Similarly, missing data computation (Section 2.3, 3.7) could also be extended to volume images. It would be beneficial for forensic sciences and archaeology to be able to restore damaged artifacts based on a statistical model created from volume data collections. There have already been attempts at performing such a procedure, but it is heavily dependent on collected data sets. These sets are currently available in relatively generous amounts thanks to the Department of Anthropology and Human Genetics at the Faculty of Life Sciences, Charles University in Prague.

Bibliography

Akhoondali, H., Zoroofi, R., Shirani, G., Rapid Automatic Segmentation and Visualization of Teeth in CT-Scan Data. *Journal of Applied Sciences*, 2009, **9**, pp. 2031–2044.

Antoniou, A., Wu-Sheng, L., *Practical optimization: algorithms and engineering applications*. Springer, 2007, ISBN 978-0-387-71106-5.

Arbour, J. H., Brown, C. M., Incomplete specimens in geometric morphometric analyses. *Methods in Ecology and Evolution*, 2014, **5**(1), pp. 16–26.

Ashburner, J., Friston, K. J., Voxel-Based Morphometry - The Methods. *NeuroImage*, 2000, **11**(6), pp. 805–821.

Audette, M. A., Ferrie, F. P., Peters, T. M., An algorithmic overview of surface registration techniques for medical imaging. *Medical Image Analysis*, 2000, **4**(3), pp. 201–217.

Balci, S. K., Golland, P., Shenton, M., Wells, W. M., Free-Form B-spline Deformation Model for Groupwise Registration. *Medical Image Computing and Computer-Assisted Intervention*, 2007, **10**(WS), pp. 23–30.

Bejdová, S., Krajíček, V., Trefný, P., Peterka, M., Velemínská, J., Variability in palatal shape and size in patients with bilateral complete cleft lip and palate assessed using dense surface model construction and 3D geometric morphometrics. *Journal of Cranio-Maxillofacial Surgery*, 2012a, **40**(3), pp. 201–208.

Bejdová, S., Krajíček, V., Velemínská, J., Horák, M., Velemínský, P., Microevolution of mandible in the area of central Europe during the latest 1200 years using methods of 3D geometric morphometrics. Anthropologischer Anzeiger, 68, 4, 2011.

Bejdová, S., Krajíček, V., Velemínská, J., Horák, M., Velemínský, P., A Microevolution of upper face in the area of Central Europe during the latest 1200 years. 18th Congress of the European Anthropological Association, Ankara, Turkey, 2012b.

Bejdová, S., Krajíček, V., Velemínská, J., Horák, M., Velemínský, P., Changes in the sexual dimorphism of the human mandible during the last 1200 years in Central Europe. *HOMO - Journal of Comparative Human Biology*, 2013, **64**(6), pp. 437–53.

Besl, P. J., McKay, N. D., A Method for Registration of 3-D Shapes. *IEEE Trans. Pattern Anal. Mach. Intell.*, 1992, **14**(2), pp. 239–256.

Bhagalia, R. R., Fessler, J. A., Kim, B., Gradient based image registration using importance sampling. In *3rd IEEE International Symposium on Biomedical Imaging: Nano to Macro*, IEEE, 2006, pp. 446–449.

Bhatia, K., Hajnal, J., Puri, B., Edwards, A., Rueckert, D., Consistent groupwise non-rigid registration for atlas construction. In *IEEE International Symposium on Biomedical Imaging: Nano to Macro*, vol. 1, IEEE, 2004, pp. 908–911.

Bigoni, L., Krajíček, V., Sládek, V., Velemínský, P., Poláček, L., Velemínská, J., Different Subsistence Patterns and the Socioeconomic Structure of Medieval Society of Great Moravia. 1838emes Journées de la Société d'Anthropologie de Paris, Paris, 2013a.

Bigoni, L., Krajíček, V., Sládek, V., Velemínský, P., Velemínská, J., Skull shape asymmetry and the socioeconomic structure of an early medieval Central European society. *American Journal of Physical Anthropology*, 2013b, **150**(3), pp. 349–364.

Bishop, C. M., *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Secaucus, NJ, USA: Springer-Verlag, 2006, ISBN 0387310738.

Blanz, V., Mehl, A., Vetter, T., Seidel, H.-P., A statistical method for robust 3D surface reconstruction from sparse data. In *Proceedings of the 2nd International Symposium on 3D Data Processing, Visualization, and Transmission*, IEEE, 2004, pp. 293–300.

Blanz, V., Vetter, T., A Morphable Model for the Synthesis of 3D Faces. In *Proceedings of the 26th Annual Conference on Computer Graphics and Interactive Techniques*, ACM, 1999, pp. 187–194.

Bookstein, F. L., *Morphometric tools for landmark data : geometry and biology*. Cambridge University Press, 1997, ISBN 0521585988.

Bookstein, F. L., 'Voxel-based morphometry' should not be used with imperfectly registered images. *NeuroImage*, 2001, **14(6)**, pp. 1454–1462.

Brigger, P., Hoeg, J., Unser, M., B-Spline Snakes: A Flexible Tool for Parametric Contour Detection. *IEEE Transactions on Image Processing*, 2000, **9**(9), pp. 1484–1496.

Brown, C. M., Arbour, J. H., Jackson, D. A., Testing of the Effect of Missing Data Estimation and Distribution in Morphometric Multivariate Data Analyses. *Systematic Biology*, 2012, **61**(6), pp. 941–954.

Brzobohatá, H., Krajíček, V., Horák, Z., Sedlak, P., Velemínská, J., Diachronic changes in size and shape of human proximal tibia in the area of Central Europe during the latest 1200 years. *HOMO - Journal of Comparative Human Biology*, 2015a. (submitted).

Brzobohatá, H., Krajíček, V., Horák, Z., Velemínská, J., Sex Classification Using the Three-Dimensional Tibia Form or Shape Including Population Specificity Approach. *Journal of Forensic Science*, 2015b, **60**(1), pp. 29–40.

Brzobohatá, H., Krajíček, V., Velemínský, P., Poláček, L., Velemínská, J., The Shape Variability of Human Tibial Epiphyses in an Early Medieval Great Moravian Population (9th -10th Century AD): A Geometric Morphometric Assessment. *Anthropologischer Anzeiger*, 2014, **71**(3), pp. 219–236.

Bustard, J., Nixon, M., 3D morphable model construction for robust ear and face recognition. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2010, pp. 2582–2589.

Cagáňová, V., Borský, J., Krajíček, V., Hoffmannová, E., Černý, M., Velemínská, J., Three-dimensional facial morphology following neonatal cheiloplasty in six-old-years patients with unilateral cleft of the lip and palate. *Journal of Cranio-Maxillofacial Surgery*, 2014. (submitted).

Chan, T., Vese, L., Active contours without edges. *IEEE Transactions on Image Processing*, 2001, **10**(2), pp. 266–277.

Cheverud, J. M., Richtsmeier, J. T., Finite-element scaling applied to sexual dimorphism in rhesus macaque (Macaca mulatta) facial growth. *Systematic Zoology*, 1986, **35**(3), pp. 381–399.

Chvojková, M., Krajíček, V., Velemínská, J., Kraniometrická Variabilita Historických Populací z Oblasti Údolí Nilu. *Slov. Antropol.*, 2010, **13**(2), pp. 19–23–5.

Combes, B., Hennessy, R., Waddington, J., Roberts, N., Prima, S., Automatic symmetry plane estimation of bilateral objects in point clouds. In *IEEE Conference on Computer Vision and Pattern Recognition*, IEEE, 2008, pp. 1–8.

Combes, B., Prima, S., New Algorithms to Map Asymmetries of 3D Surfaces. In *Proceedings of the 11th international conference on Medical Image*

Computing and Computer-Assisted Intervention - Part I, Springer, 2008, pp. 17–25.

Conway, J., Burgiel, H., Goodman-Strauss, C., *The Symmetries of Things*. Ak Peters Series, Taylor & Francis, 2008, ISBN 9781568812205.

Cootes, T. F., Edwards, G. J., Taylor, C. J., Active Appearance Models. In *Computer Vision — ECCV'98*, vol. 1407, Springer, 1998, pp. 484–498.

Cootes, T. F., Hill, A., Taylor, C. J., Haslam, J., The Use of Active Shape Models for Locating Structures in Medical Images. In *Information Processing in Medical Imaging*, Springer, 1993, pp. 33–47.

Cootes, T. F., Taylor, C. J., Active Shape Models - Smart Snakes. In *Proceedings of the British Machine Vision Conference*, Springer, 1992, pp. 266–275.

Cortes, C., Vapnik, V., Support-Vector Networks. *Machine Learning*, 1995, **20**(3), pp. 273–297.

Datar, M., Immorlica, N., Indyk, P., Mirrokni, V. S., Locality-sensitive Hashing Scheme Based on P-stable Distributions. In *Proceedings of the Twentieth Annual Symposium on Computational Geometry*, ACM, 2004, pp. 253–262.

Davies, R. H., Twining, C. J., Taylor, C. J., *Statistical models of shape - optimisation and evaluation*. Springer, 2008, ISBN 978-1-8480-0137-4, I-XII, 1-302 pp.

Descoteaux, M., Audette, M., Chinzei, K., Siddiqi, K., Bone Enhancement Filtering: Application to Sinus Bone Segmentation and Simulation of Pituitary Surgery. In *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2005*, vol. 3749, Springer, 2005, ISBN 978-3-540-29327-9, pp. 9–16.

Deshpande, A., Vempala, S., Adaptive Sampling and Fast Low-Rank Matrix Approximation. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*, vol. 4110, Springer, 2006, ISBN 978-3-540-38044-3, pp. 292–303.

Devroye, L., *Non-Uniform Random Variate Generation*. Springer-Verlag, 1986, ISBN 0387963057.

Drost, B., Ilic, S., A Hierarchical Voxel Hash for Fast 3D Nearest Neighbor Lookup. In *Pattern Recognition*, vol. 8142, Springer, 2013, pp. 302–312.

Dupej, J., Krajíček, V., Pelikán, J., Low-Rank Matrix Approximations for Coherent Point Drift. *Pattern Recognition Letters*, 2014a, **52**, pp. 53–58.

Dupej, J., Krajíček, V., Velemínská, J., Pelikán, J., Analysis of Asymmetry in Triangular Meshes. In *Proceedings of the 33 rd Conference on Geometry and Graphics*, VŠB-Technical University of Ostrava, 2013, pp. 65–78.

Dupej, J., Krajíček, V., Velemínská, J., Pelikán, J., Statistical Mesh Shape Analysis with Nonlandmark Nonrigid Registration. 12th Symposium on Geometry Processing, Cardiff, UK, 2014b.

Elseberg, J., Rol, S. M., Nüchter, S. A., Comparison of nearest-neighbor-search strategies and implementations for efficient shape registration. *Journal of Software Engineering for Robotics*, 2012, **3**(1), pp. 2–12.

Fournier, M., Combes, B., Roberts, N., Keller, S., Crow, T., Hopkins, W., Prima, S., Surface-based method to evaluate global brain shape asymmetries in human and chimpanzee brains. In *IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, IEEE, 2011, pp. 310–316.

Gao, H., Chae, O., Individual tooth segmentation from CT images using level set method with shape and intensity prior. *Pattern Recognition Letters*, 2010, **43**(7), pp. 2406–2417.

Gao, S., Yang, J., Yan, Y., A local modified Chan-Vese model for segmenting inhomogeneous multiphase images. *International Journal of Imaging Systems and Technology*, 2012, **22**(2), pp. 103–113.

Gholipour, A., Kehtarnavaz, N., Briggs, R., Devous, M., Gopinath, K., Brain Functional Localization: A Survey of Image Registration Techniques. *IEEE Transactions on Medical Imaging*, 2007, **26**(4), pp. 427–451.

Golub, G. H., Van Loan, C. F., *Matrix Computations (3rd Ed.)*. Baltimore, MD, USA: Johns Hopkins University Press, 1996, ISBN 0-8018-5414-8.

Greengard, L., Strain, J., The Fast Gauss Transform. *SIAM Journal on Scientific and Statistical Computing*, 1991, **12**(1), pp. 79–94.

Hastie, T., Tibshirani, R., Friedman, J., *The Elements of Statistical Learning: Data Mining, Inference and Prediction*. 2nd edn., Springer, 2008, ISBN 978-0-387-84858-7.

Heo, H., Chae, O.-S., Segmentation of tooth in CT images for the 3D reconstruction of teeth. In *Proceedings of the SPIE*, vol. 5298, SPIE, 2004, pp. 455–466.

Horn, B. K. P., Closed-form solution of absolute orientation using unit quaternions. *Journal of the Optical Society of America A*, 1987, **4**(4), pp. 629–642.

Hosntalab, M., Aghaeizadeh Zoroofi, R., Abbaspour Tehrani-Fard, A., Shirani, G., Segmentation of teeth in CT volumetric dataset by panoramic projection and variational level set. *International Journal of Computer Assisted Radiology and Surgery*, 2008, **3**(3-4), pp. 257–265.

Hu, Y. C., Grossberg, M. D., Mageras, G. S., *Survey of Recent Volumetric Medical Image Segmentation Techniques*, chap. 17. InTech, 2009, pp. 321–346.

Hutton, T. J., Buxton, B. F., Hammond, P., Dense Surface Point Distribution Models of the Human Face. In *IEEE Workshop on Mathematical Methods in Biomedical Image Analysis*, 2001, pp. 153–160.

Ilin, A., Raiko, T., Practical Approaches to Principal Component Analysis in the Presence of Missing Values. *Journal of Machine Learning Research*, 2010, **11**, pp. 1957–2000.

Jacob, M., Blu, T., Unser, M., Efficient energies and algorithms for parametric snakes. *IEEE Transactions on Image Processing*, 2004, **13**(9), pp. 1231–1244.

Johnson, S. G., The NLopt nonlinear-optimization package. http://ab-initio.mit.edu/wiki/index.php/NLopt, 2014.

Kabus, S., Netsch, T., Fischer, B., Modersitzki, J., B-Spline Registration of 3D Images with Levenberg-Marquardt Optimization. In *Proceedings of the SPIE*, vol. 5370, SPIE, 2004, pp. 304–313.

Kass, M., Witkin, A., Terzopoulos, D., Snakes: Active Contour Models. *International Journal of Computer Vision*, 1988, **1**(4), pp. 321–331.

Kazhdan, M., Funkhouser, T., Rusinkiewicz, S., Rotation Invariant Spherical Harmonic Representation of 3D Shape Descriptors. In *Eurographics Symposium on Geometry Processing*, Eurographics Association, 2003, pp. 156–164.

Khan, W., Image Segmentation Techniques: A Survey. *Journal of Image and Graphics*, 2013, **1**(4), pp. 166–170.

Klein, S., Staring, M., Pluim, J., A Comparison of Acceleration Techniques for Nonrigid Medical Image Registration. In *Biomedical Image Registration*, vol. 4057, Springer, 2006, pp. 151–159.

Klingenberg, C. P., Barluenga, M., Meyer, A., Shape analysis of symmetric structures: Quantifying variation among individuals and asymmetry. *Evolution*, 2002, **56**(10), pp. 1909–1920.

Kolomazník, J., Horáček, J., Krajíček, V., Pelikán, J., Segmentation on CUDA Using Graph-Cuts and Watershed Transformation. In *WSCG Poster Proceedings*, Union Agency, 2012, pp. 35–38.

Kondo, T., Ong, S. H., Foong, K. W. C., Tooth segmentation of dental study models using range images. *IEEE Transactions on Medical Imaging*, 2004, **23**(3), pp. 350–362.

Koschan, A., Kang, S., Paik, J., Abidi, B., Abidi, M., Color active shape models for tracking non-rigid objects. *Pattern Recognition Letters*, 2003, **24**(11), pp. 1751–1765.

Koudelová, J., Růžek, J., Cagáňová, V., Krajíček, V., Velemínská, J., Development of facial sexual dimorphism in children aged between 12 and 15 years: a three-dimensional longitudinal study. *Orthodontics & Craniofacial Research*, 2015. (in press).

Krajíček, V., *Volume measurement in 3D data*. Master's thesis, Faculty of Matematics and Physics, Charles University in Prague, April 2007.

Krajíček, V., Analyzing Contrast Enhanced MRI Sequences for Mammography. In *Proceedings of Contributed Papers: Part I - Mathematics and Computer Sciences*, Matfyzpress, 2008a, pp. 195–201.

Krajíček, V., Design of Segmentation Algorithm for Volume Measuring CAD system. In *Proceedings of MIS 2008*, Matfyzpress, 2008b, pp. 47–57.

Krajíček, V., Bejdová, S., Velemínská, J., Pelikán, J., Improving B-spline Deformation Based Fitting for Volume Registration. In *Proceedings of the 31st Conference on Geometry and Graphics*, VŠB-Technical University of Ostrava, 2011, pp. 139–154.

Krajíček, V., Dupej, J., Bejdová, S., Velemínská, J., Pelikán, J., Teeth and Jaw Segmentation Using Fast Level-set Algorithm and Local Region Anisotropic Priors. *Imaging Science Journal*, 2014. (submitted).

Krajíček, V., Dupej, J., Koudelová, J., Velemínská, J., Statistical Mesh Analysis of Longitudinal Shape Changes. In *Proceedings of the 33 rd Conference on Geometry and Graphics*, VŠB-Technical University of Ostrava, 2013, pp. 155–168.

Krajíček, V., Dupej, J., Velemínská, J., Pelikán, J., Morphometric Analysis of Mesh Asymmetry. *Journal of WSCG*, 2012, **20**(1), pp. 65–72.

Krajíček, V., Pelikán, J., Horák, M., Measuring and Segmentation in CT Data Using Deformable Models. In Skala, V. (ed.), *WSCG' 2007 Short Communications Proceedings*, vol. 2, Union Agency, 2007, pp. 149–152.

Kratky, J., Kybic, J., Three-dimensional segmentation of bones from CT and MRI using fast level sets. In *Medical Imaging 2008: Image Processing*, vol. 6914, SPIE, 2008, pp. 691 447–10.

Krcah, M., Szekely, G., Blanc, R., Fully automatic and fast segmentation of the femur bone from 3D-CT images with no shape prior. In *IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, IEEE, 2011, pp. 2087–2090.

Kumar, S., Mohri, M., Talwalkar, A., Sampling Methods for the Nyström Method. *J. Mach. Learn. Res.*, Apr. 2012, **13**(1), pp. 981–1006.

Kybic, J., Unser, M., Fast parametric elastic image registration. *IEEE Transactions on Image Processing*, 2003, **12**(11), pp. 1427–1442.

Lane, J. M., Riesenfeld, R. F., A Theoretical Development for the Computer Generation and Display of Piecewise Polynomial Surfaces. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 1980, **2**(1), pp. 35–46.

Lee, D., Antani, S., Long, L. R., Similarity Measurement Using Polygon Curve Representation and Fourier Descriptors for Shape-based Vertebral Image Retrieval. In *Medical Imaging: Image Processing*, vol. 5032, SPIE, 2003, pp. 1283–1291.

Li, C., Huang, R., Ding, Z., Gatenby, J., Metaxas, D., Gore, J., A Level Set Method for Image Segmentation in the Presence of Intensity Inhomogeneities With Application to MRI. *IEEE Transactions on Image Processing*, 2011, **20**(7), pp. 2007–2016.

Li, C., Xu, C., Gui, C., Fox, M. D., Level set evolution without re-initialization: A new variational formulation. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 1, IEEE, 2005, pp. 430–436.

Little, R. J. A., Rubin, D. B., *Statistical analysis with missing data (second edition)*. Wiley-Interscience, 2002, ISBN 978-0-471-18386-0.

Liu, L., Raber, D., Nopachai, D., Commean, P., Sinacore, D., Prior, F., Pless, R., Ju, T., Interactive Separation of Segmented Bones in CT Volumes Using Graph Cut. In *Proceedings of the 11th international conference on Medical Image Computing and Computer-Assisted Intervention - Part I*, Springer, 2008, pp. 296–304.

Liu, Y., Palmer, J., A quantified study of facial asymmetry in 3D faces. In *IEEE International Workshop on Analysis and Modeling of Faces and Gestures*, IEEE, 2003, pp. 222–229.

Malladi, R., Sethian, J., Vemuri, B., Shape modeling with front propagation: a level set approach. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 1995, **17**(2), pp. 158–175.

Mani, V., Arivazhagan, D., Survey of Medical Image Registration. *Journal of Biomedical Engineering and Technology*, 2013, **1**(2), pp. 8–25.

Marden, S., Guivant, J., Improving the Performance of ICP for Real-Time Applications using an Approximate Nearest Neighbour Search. In *Proceedings of Australasian Conference on Robotics and Automation*, 2012, pp. 1–6.

Martelli, A., An Application of Heuristic Search Methods to Edge and Contour Detection. *Communications of the ACM*, 1976, **19**(2), pp. 73–83.

Mayer, A., Greenspan, H., Direct Registration of White Matter Tractographies with Application to Atlas Construction. In *MICCAI 2007 Workshop Statistical Registration PairWise and GroupWise Alignment and Atlas Formation*, 2007, pp. 1–8.

Meijster, A., Roerdink, J. B. T. M., Hesselink, W. H., A General Algorithm for Computing Distance Transforms in Linear Time. In *Mathematical Morphology and its Applications to Image and Signal Processing*, vol. 18, Springer, 2002, ISBN 978-0-306-47025-7, pp. 331–340.

Mitchell, D. P., Spectrally Optimal Sampling for Distribution Ray Tracing. In *Proceedings of the 18th Annual Conference on Computer Graphics and Interactive Techniques*, ACM, 1991, pp. 157–164.

Mitteroecker, P., Gunz, P., Advances in Geometric Morphometrics. *Evolutionary biology*, 2009, **36**(2), pp. 235–247.

Morphome3cs, Morphome3cs Project at Computer Graphics Group At Charles University in Prague. http://cgg.mff.cuni.cz/Morpho, 2015. Accessed: May 2015.

Mueller, A., Paysan, P., Schumacher, R., Zeilhofer, H.-F., Berg-Boerner, B.-I., Maurer, J., Vetter, T., Schkommodau, E., Juergens, P., Schwenzer-Zimmerer, K., Missing facial parts computed by a morphable model and transferred directly to a polyamide laser-sintered prosthesis: an innovation study. *British Journal of Oral and Maxillofacial Surgery*, 2011, **49**(8), pp. e67 – e71.

Myronenko, A., Song, X., Point Set Registration: Coherent Point Drift. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 2010, **32**(12), pp. 2262–2275.

Myronenko, A., Song, X., Carreira-Perpinán, Á., Non-rigid point set registration: Coherent Point Drift (CPD). In *Advances In Neural Information Processing Systems*, vol. 19, MIT Press, 2006, pp. 1009–1016.

Noblet, V., Heinrich, C., Heitz, F., Armspach, J.-P., An efficient incremental strategy for constrained groupwise registration based on symmetric pairwise registration. *Pattern Recognition Letters*, 2012, **33**(3), pp. 283–290.

O'Grady, K. F., Antonyshyn, O. M., Facial asymmetry: three-dimensional analysis using laser surface scanning. *Plastic and Reconstructive Surgery*, 1999, **104**(4), pp. 928–37.

Okazaki, N., libLBFGS: a library of Limited-memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS). http://www.chokkan.org/software/liblbfgs/, 2010.

Ólafsdóttir, H., Lanche, S., Darvann, T. A., Hermann, N. V., Larsen, R., Ersbøll, B. K., Oubel, E., Frangi, A. F., Larsen, P., Perlyn, C. A., Morriss-Kay, G. M., Kreiborg, S., A point-wise quantification of asymmetry using deformation fields: application to the study of the Crouzon mouse model. In *Proceedings of the 10th international conference on Medical image computing and computer-assisted intervention*, Springer, 2007, pp. 452–459.

Osher, S., Rudin, L. I., Feature-oriented image enhancement using shock filters. *SIAM Journal on Numerical Analysis*, 1990, **27**(4), pp. 919–940.

Owen, A. B., Quasi-Monte Carlo Sampling. In *Monte Carlo Ray Tracing: Siggraph 2003 Course 44*, SIGGRAPH, 2003, pp. 69–88.

Palmer, A. R., Fluctuating Asymmetry Analyses: A Primer. In *Developmental Instability: Its Origins and Evolutionary Implications*, Springer, 1994, pp. 335–364.

Palmer, A. R., Strobeck, C., Fluctuating asymmetry: Measurement, analysis, patterns. *Annual Review of Ecology and Systematics*, 1986, **17**, pp. 391–421.

Paragios, N., Rousson, M., Ramesh, V., Non-rigid registration using distance functions. *Computer Vision and Image Understanding.*, 2003, **89**(2-3), pp. 142–165.

Peres-Neto, P. R., Jackson, D. A., Somers, K. M., How many principal components? stopping rules for determining the number of non-trivial axes revisited. *Computational Statistics & Data Analysis*, 2005, **49**(4), pp. 974–997.

Peterka, M., Peterková, R., Tvrdek, M., Kuderová, J., Likovský, Z., Significant differences in the incidence of orofacial clefts in fifty-two Czech districts between 1983 and 1997. *Acta Chirurgiae Plasticae*, 2000, **4**, pp. 124–129.

Pham, D. L., Xu, C., Prince, J. L., Current Methods In Medical Image Segmentation. *Annual Review of Biomedical Engineering*, 2000, **2**, pp. 315–337.

R Development Core Team, *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2008. ISBN 3-900051-07-0.

Rao, C. S., Kumar, S. S., Mohan, B. C., Content Based Image Retrieval Using Exact Legendre Moments and Support Vector Machine. *International journal of Multimedia & Its Applications*, 2010, **2**(2), pp. 69–79.

Rohlfing, T., Maurer, C., Bluemke, D., Jacobs, M., Volume-preserving non-rigid registration of MR breast images using free-form deformation with an incompressibility constraint. *IEEE Transactions on Medical Imaging*, 2003, **22**(6), pp. 730–741.

Rohlfing, T., Maurer, C. R., Nonrigid Image Registration in Shared-Memory Multiprocessor Environments With Application to Brains, Breasts, and Bees. *IEEE transactions on information technology in biomedicine*, 2003, **7**(1), pp. 16–25.

Rueckert, D., Frangi, A., Schnabel, J., Automatic Construction of 3D Statistical Deformation Models Using Non-rigid Registration. In *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2001*, vol. 2208, Springer, 2001, ISBN 978-3-540-42697-4, pp. 77–84.

Rueckert, D., Hayes, C., Studholme, C., Summers, P., Leach, M., Hawkes, D. J., Non-rigid Registration of Breast MR Images Using Mutual Information. In *Medical Image Computing and Computer-Assisted Intervention — MICCAI'98*, vol. 1496, Springer, 1998, pp. 1144–1152.

Rueckert, D., Sonoda, L. I., Hayes, C., Hill, D. L. G., Leach, M., Hawkes, D. J., Non-rigid registration using free-form deformations: Application to breast MR images. *IEEE Transactions on Medical Imaging*, 1999, **18**(8), pp. 712–721.

Rusinkiewicz, S., Levoy, M., Efficient Variants of the ICP Algorithm. In *Third International Conference on 3D Digital Imaging and Modeling (3DIM)*, IEEE Computer Society, 2001, pp. 145–152.

Rusková, H., Bejdová, S., Peterka, M., Krajíček, V., Velemínská, J., 3-D shape analysis of palatal surface in patients with unilateral complete cleft lip and palate. *Journal of Cranio-Maxillofacial Surgery*, 2014, **42**(5), pp. 140–147.

Saad, Y., *Numerical Methods for Large Eigenvalue Problems*. Manchester, UK: Manchester University Press, 1992, ISBN 0719033861.

Samet, H., *Foundations of Multidimensional and Metric Data Structures*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2005, ISBN 0123694469.

Schaefer, K., Lauc, T., Mitteroecker, P., Gunz, P., Bookstein, F. L., Dental arch asymmetry in an isolated Adriatic community. *American Journal of Physical Anthropology*, 2006, **129**(1), pp. 132–42.

Shackleford, J., Kandasamy, N., Sharp, G., *GPU Computing Gems 4*, chap. Deformable Volumetric Registration using B-splines. Elsevier, 2010, pp. 751–770.

Sharma, S., Dhole, A., Content Based Image Retrieval Based on Shape Feature using Accurate Legendre Moments and Support Vector Macines. *International Journal of Computer Science Engineering & Technolo*, 2013, **3**(5), pp. 194–199.

Shi, Y., Karl, W., A Fast Level Set Method Without Solving PDEs. In *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '05)*, vol. 2, IEEE, 2005, pp. 97–100.

Sirovich, L., Kirby, M., Low-dimensional procedure for the characterization of human faces. *Journal of the Optical Society of America A*, 1987, **4**(3), pp. 519–524.

Song, H., Li, J.-j., Wang, S.-l., Ma, J.-t., Multi-modality liver image registration based on multilevel B-splines free-form deformation and L-BFGS optimal algorithm. *Journal of Central South University*, 2014, **21**(1), pp. 287–292.

Spall, J. C., *Introduction to Stochastic Search and Optimization*. 1st edn., New York, NY, USA: John Wiley & Sons, Inc., 2003, ISBN 0471330523.

Suetens, P., *Fundamentals of Medical Imaging*. Cambridge University Press, 2009, ISBN 9780521519151.

SUNY, Morphometrics at SUNY Stony Brook. `http://life.bio.sunysb.edu/morph/`, 2015. Accessed: May 2015.

Tam, G., Cheng, Z.-Q., Lai, Y.-K., Langbein, F., Liu, Y., Marshall, D., Martin, R., Sun, X.-F., Rosin, P., Registration of 3D Point Clouds and Meshes: A Survey from Rigid to Nonrigid. *IEEE Transactions on Visualization and Computer Graphics*, 2013, **19**(7), pp. 1199–1217.

Thevenaz, P., Bierlaire, M., Unser, M., Halton Sampling for Image Registration Based on Mutual Information. *Sampling Theory in Signal and Image Processing*, 2008, **7**(2), pp. 141–171.

Trefný, P., Krajíček, V., Velemínská, J., Three-dimensional analysis of palatal shape in patients treated with SARME using a dense surface model. *Orthodontics & Craniofacial Research*, 2015. (submitted).

Tsai, A., Yezzi, A., Wells, W., Tempany, C., Tucker, D., Fan, A., Grimson, W., Willsky, A., A shape-based approach to the segmentation of medical imagery using level sets. *IEEE Transactions on Medical Imaging*, 2003, **22**(2), pp. 137–154.

Vacavant, A., Albouy-Kissi, A., Menguy, P., Solomon, J., Fast smoothed shock filtering. In *21st International Conference on Pattern Recognition (ICPR)*, IEEE, 2012, pp. 182–185.

van Kaick, O., Zhang, H., Hamarneh, G., Cohen-Or, D., A Survey on Shape Correspondence. *Computer Graphics Forum*, 2011, **30**(6), pp. 1681–1707.

Van Valen, L., A study of fluctuating asymmetry. *Evolution*, 1962, **16**(2), pp. 125–142.

Velemínská, J., Bigoni, L., Krajíček, V., Borský, J., Šmahelová, D., Cagáňová, V., Peterka, M., Surface facial modelling and allometry in relation to sexual dimorphism. *HOMO - Journal of Comparative Human Biology*, 2012, **63**(2), pp. 81–93.

Velemínská, J., Krajíček, V., Dupej, J., Goméz-Valdés, J. A., Velemínský, P., Šefčáková, A., Pelikán, J., Sánchez-Mejorada, G., Brůžek, J., Geometric morphometrics and sexual dimorphism of the greater sciatic notch in adults from two skeletal collections: The accuracy and reliability of sex classification. *American Journal of Physical Anthropology*, 2013, **152**(4), pp. 558–565.

Vercauteren, T., Pennec, X., Perchant, A., Ayache, N., Non-parametric Diffeomorphic Image Registration with the Demons Algorithm. In *Proceedings of the 10th International Conference on Medical Image Computing and Computer-assisted Intervention*, Springer, 2007, pp. 319–326.

Wahba, G., *Spline models for observational data*, vol. 59. Philadelphia, PA: Society for Industrial and Applied Mathematics (SIAM), 1990, ISBN 978-0-89871-244-5.

Wang, X.-F., Huang, D.-S., Xu, H., An efficient local Chan-Vese model for image segmentation. *Pattern Recognition*, 2010, **43**(3), pp. 603–618.

Weickert, J., *Anisotropic Diffusion In Image Processing*. Teubner-Verlag, Stuttgart, 1998, ISBN 3-519-02606-6.

Weickert, J., Coherence-Enhancing Diffusion Filtering. *International Journal of Computer Vision*, 1999, **31**(2-3), pp. 111–127.

Weickert, J., Coherence-enhancing shock filters. In *Pattern Recognition*, vol. 2781, Springer, 2003, pp. 1–8.

Wenckebach, T., Lamecker, H., Hege, H.-C., Capturing Anatomical Shape Variability Using B-Spline Registration. In *Information Processing in Medical Imaging*, vol. 3565, Springer, 2005, ISBN 978-3-540-26545-0, pp. 578–590.

Williams, C., Seeger, M., Using the Nyström Method to Speed Up Kernel Machines. In *Advances in Neural Information Processing Systems 13*, MIT Press, 2001, pp. 682–688.

Wu, B., Yang, Y., Local- and Global-Statistics-Based Active Contour Model for Image Segmentation. *Mathematical Problems in Engineering*, 2012, **2012**, p. 16.

Yezzi, J., A., Tsai, A., Willsky, A., A statistical approach to snakes for bimodal and trimodal imagery. In *The Proceedings of the Seventh IEEE International Conference on Computer Vision*, vol. 2, IEEE, 1999, pp. 898–903.

Yuille, A., Grzywacz, N., A mathematical analysis of the motion coherence theory. *International Journal of Computer Vision*, 1989, **3**(2), pp. 155–175.

Zelditch, M. L., Swiderski, D. L., Sheets, D. H., Fink, W. L., *Geometric Morphometrics for Biologists: A primer*. Elsevier Academic Press, 2004, ISBN 978-0-12-778460-1.

Zhang, K., Tsang, I. W., Kwok, J. T., Improved Nyström Low-rank Approximation and Error Analysis. In *Proceedings of the 25th International Conference on Machine Learning*, ACM, 2008, ISBN 978-1-60558-205-4, pp. 1232–1239.

Zhang, K., Zhang, L., Lam, K.-M., Zhang, D., A Local Active Contour Model for Image Segmentation with Intensity Inhomogeneity. *Computer Vision and Pattern Recognition*, 2013, **1**(1), pp. 1–?

Zheng, G., Lösch, S., Volumetric Morphometrics (VMM) for Physical Anthropology: Preliminary Results. *Bulletin der Schweizerischen Gesellschaft für Anthropologie*, 2013, **19**(2), pp. 25–31.

Šmahel, Z., Velemínská, J., Trefný, P., Müllerová, Ž., Three-dimensional morphology of the palate in patients with bilateral complete cleft lip and palate at the stage of permanent dentition. *Cleft Palate-Craniofacial Journal*, 2009, **46**, pp. 399–408.

Špačková, J., Cagáňová, V., Krajíček, V., Velemínská, J., Specification of child and juvenile identification: 3D modelling of facial ontogenetic development during the pubertal spurt. European Academy of Forensic Science Conference, Hague, Netherland, 2012.

List of Tables

List of Case Studies

List of Abbreviations

ANN Artificial neural networks.

CDF Cumulative distribution function.

CPD Coherent point drift.

CT Computed tomography.

DCA Dense correspondence analysis.

DCM Dense correspondence model.

DVR Direct volume rendering.

EDM Euclidean distance map.

EM Expectation-maximization.

FESA Finite element scaling analysis.

FFD Free-form deformation.

FGT Fast Gaussian transform.

GM Gaussian mixture models.

GMM Geometric morphometrics.

GPA Generalized Procrustes analysis.

HDPCA High-dimensional principal component analysis.

ICP Iterative closest point.

IRAM Implicitly restarted Arnoldi's method.

LDA Linear discriminant analysis.

MI Mutual information.

MRI Magnetic resonance imaging.

OPA Ordinary Procrustes analysis.

PCA Principal component analysis.

PDF Probability density function.

PDM Point distribution model.

RBF Radial basis function.

SARME Surgically-assisted rapid maxillary expansion.

SGD Stochastic gradient descent.

SVD Singular vector decomposition.

SVM Support vector machines.

TPS Thin-plate spline.