Analysis of Rotational Deformations from Directional Data

Jörn Schulz[1] *, Sungkyu Jung[2], Stephan Huckemann[3], Michael Pierrynowski[4], J. S. Marron[5] and Stephen M. Pizer[5]

[1]University of Tromsø, [2]University of Pittsburgh, [3]University of Göttingen, [4]McMaster University, Hamilton [5]University of North Carolina at Chapel Hill

Abstract

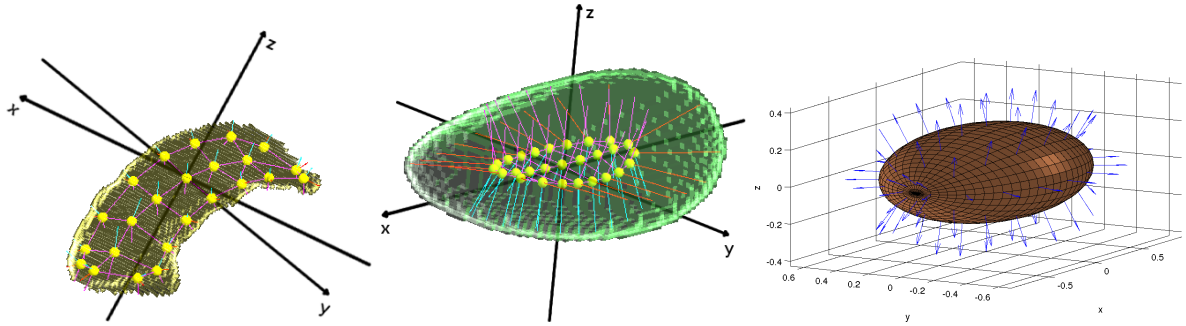
This paper discusses a novel framework to analyse rotational deformations of real 3D objects. The rotational deformations such as twisting or bending have been observed as the major variation in some medical applications, where the features of the deformed 3D objects are directional data. We propose modelling and estimation of the global deformations in terms of generalized rotations of directions. The proposed method can be cast as a generalized small circle fitting on the unit sphere. We also discuss the estimation of descriptors for more complex deformations composed of two simple deformations. The proposed method can be used for a number of different 3D object models. Two analyses of 3D object data are presented in detail: one using skeletal representations in medical image analysis as well as one from biomechanical gait analysis of the knee joint. Supplementary Materials are available online.

Keywords: 3D object, axis of rotation, directional statistics, skeletal model, small circle.

1 Introduction

Modeling deformations of a real object is a central issue in computer vision, biomechanics and medical imaging. In a number of applications, generalized rotations appear to be the

*This research was funded by the Norwegian Research Council through grant 176872/V30 in the eVita program and performed as part of Tromsø Telemedicine Laboratory.



(a) (b) (c)

Figure 1: 3D objects and their models. (a) S-rep of a hippocampus (b) S-rep of a rotationally deformed ellipsoid. (c) Attached boundary normals on the meshed surface of an ellipsoid.

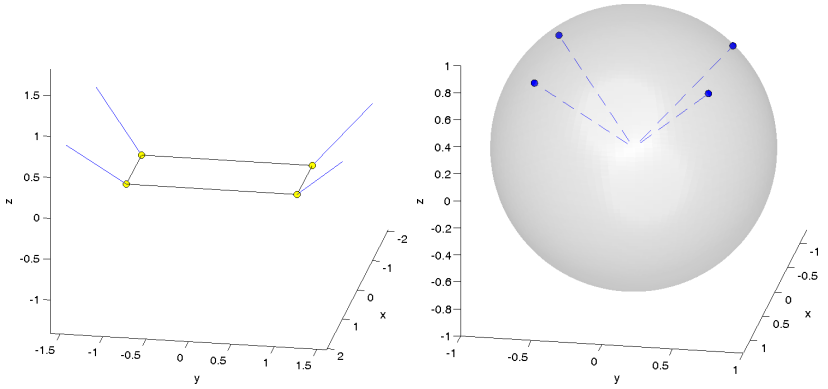
major forms of deformation. For instance, the major variation of shapes of hippocampi in the human brain has been shown to be bending of the object (Joshi et al., 2002; Pizer et al., 2013); Human joint movements, such as the motion of the knee or the elbow, consist of bending and twisting about the joint (Rivest, 2001; Rivest et al., 2008; Oualkacha and Rivest, 2012). A direct modeling of such rotational deformations will promote a precise description of object variation and will be important for surgery or treatment planning.

In this paper, we propose an estimation procedure for descriptors of underlying rotational deformations from a random sample of objects. Specifically, the descriptors are parameters of the model we introduce in Section 2; they include rotational axes of a rotational model. Our model embraces a number of different types of deformations including rigid rotation, bending, twisting and a mixture of the last two. Although we aim to analyze variations in sophisticated human organs such as the hippocampus (Fig. 1a), we work with a simpler object resembling ellipsoids (Fig. 1b) to show the validity of the proposed method.

A major challenge in modeling rotational deformation is that such variations are typically mixed with translational and scaling effects. We address this issue by only considering direction vectors, which are invariant to translation and size changes. It will be shown that the rotational deformation can be sufficiently modeled using directional data. Another advantage of our approach is that well-studied directional data techniques can be applied (Fisher et al., 1993; Mardia and Jupp, 2000; Chang and Rivest, 2001; Jung et al., 2011).

Before we introduce our method, we point out several modeling approaches of 3D objects that are relevant to our framework of directional data, as follows:

Point distribution model A solid object is modeled by the positions of the sampled surface points on which directions normal to the surface can be attached (Cootes et al., 1992; Dryden and Mardia, 1998; Kurtek et al., 2011). See Fig. 1c.



(a) A simple 3D object (b) Direction vectors

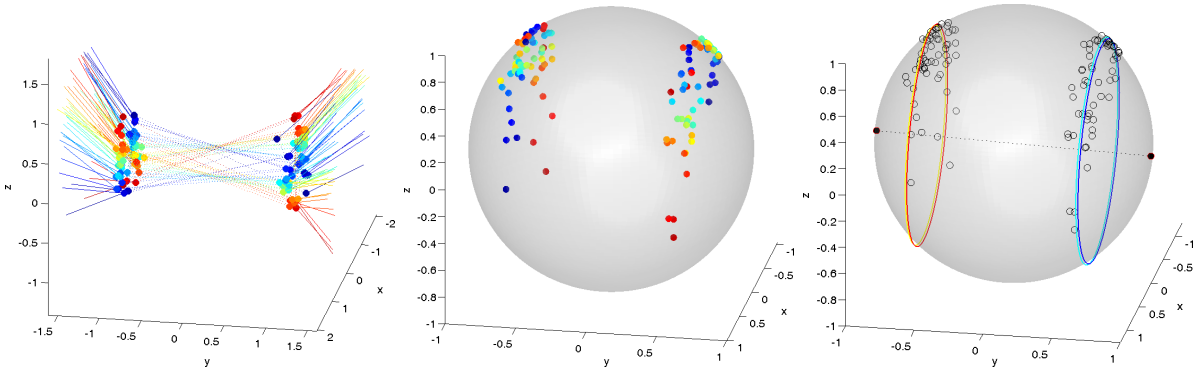
Figure 2: Toy example. (a) A toy object, to be deformed. (b) Each direction vector is a point on the surface of the unit sphere.

Large deformations The shape changes of an object in images are modeled by the deformations of a template image (Pennec, 2008; Rohde et al., 2008). The deformation can be understood as a vector field, where each vector contains the direction.

S-rep In skeletal representations (s-rep), a 3D object is modeled by skeletal positions lying inside of the object and spoke vectors pointing to the boundary of the object (Siddiqi and Pizer, 2008; Pizer et al., 2013). See Fig. 1a and Fig. 1b. We describe s-rep data analysis in more detail in Section 5.

The framework of our analysis can be understood by considering a simple example of a 3D object (Fig. 2). The object is modeled by four surface points (or skeletal positions) with attached direction vectors μ_j for $1 \leq j \leq 4$ (Fig. 2a). Consider random twists of the object, where the left and right sides are rotated about a common axis by a common angle but in opposite directions (Fig. 3a). After collecting the directional data (Fig. 3b), our method finds an estimate of the axis (overlaid in Fig. 3c) as well as the mean directions μ_j and the rotation angles. The proposed method provides an simple interpretation of the underlying rotational deformation and accurate estimation of the parameters. The estimated axis in the toy example is close to the true axis with relatively small sample size $n = 30$. See Section 2.3 for detailed discussion of this example.

To the best of our knowledge, this paper is the first attempt to model rotational deformations and to estimate the axis of rotations from directional data. There are, however, several methods available for the estimation of the axis of rotation based on 3D landmarks, especially in the area of biomechanics (Halvorsen et al., 1999; Rancourt et al., 2000; Rivest, 2001; Gamage and Lasenby, 2002; Teu and Kim, 2006; Rivest et al., 2008; Ball and Greiner,



(a) Random twists (b) Random directions (c) Estimated axis

Figure 3: Toy example. (a) $n = 30$ realizations of random twists. Different colors represent different observations. (b) Directional features. (c) The estimated axis of the twists (dotted line) is close to truth. The four concentric circles, discussed in more detail in Section 2.3, are also overlaid.

2012). In the statistical literature, estimation of rotation matrices has been studied in terms of spherical regression and its generalizations (Chang, 1986, 1988, 1989; Rivest, 1989, 2006; Chang and Rivest, 2001). In spherical regression both the axis c and the angle θ are estimated. In contrast, our model treats θ as a random variable. Hanna and Chang (2000) used the quaternion representation to model a smooth path of rotations. However, in our problem of estimation of axes of rotational deformations, it is straightforward to use the axis–angle representation. As will be pointed out later and may be guessed from Fig. 3c, the estimation of the axis of rotation based on directional vectors has a close relationship with the estimation of small circles on the unit sphere, which was studied in various contexts (Mardia and Gadsden, 1977; Rivest, 1999; Jung et al., 2011).

This paper is a part of a bigger project in modeling and estimation of deformations. We leave relevant asymptotic theory as future work.

The rest of the paper is organized as follows. We begin with introducing some necessary facts on rotations and their connections to circles on the unit sphere. In Section 2, a simple rotation model is introduced, and our estimation procedure is discussed. Non-rigid deformations such as twisting and bending of the object are also discussed in that section. A hierarchical rotation model is introduced in Section 3, where the object is assumed to be deformed sequentially by different rotations. In Section 4, simulation results are reported to show the effectiveness of the estimator. In Section 5 and 6, the merits of the proposed methods are further illustrated using applications from models that represent organs and knee motion during gait.

1.1 Rotations, circles and spheres

In the axis–angle representation of rotations, an axis c is the unit vector that is left fixed by the rotation and an angle θ gives the amount of rotation. A unit vector lies on the unit sphere $S^2 = \{x \in \mathbb{R}^3 : \|x\| = 1\}$. The axis–angle pair $(c, \theta) \in S^2 \times [0, 2\pi)$ represents a rotation in 3-space, where a vector $x \in \mathbb{R}^3$ is rotated by (c, θ) by applying $x \mapsto R(c, \theta)x$ with

$$R(c, \theta) = I_3 + \sin\theta[c]_\times + (1 - \cos\theta)(cc' - I_3), \tag{1}$$

where c' denotes the transpose of c, and $[c]_\times$ is the cross product matrix satisfying $[c]_\times v = c \times v$ for any $v \in \mathbb{R}^3$.

A useful observation in our analysis is that the direction vectors follow circles when they are rotated. In particular, when $x \in S^2$ is rotated about an axis $c \in S^2$, the trajectory of such rotation is precisely a circle, centered at c with radius $r = \arccos(x'c)$, and is denoted by $\delta(c, r) = \{x \in S^2 : x'c = \cos(r)\} \subset S^2$. Since $\delta(c, r) = \delta(-c, \pi - r)$ we may assume that $r \leq \pi/2$. We call $\delta(c, r)$ a great circle if $r = \pi/2$ and a small circle if $r < \pi/2$.

If a K-tuple of $K \geq 2$ direction vectors $\mathbf{x} = (x_1, \ldots, x_K) \in (S^2)^K$ are rotated together about a common axis c, then each of the rotated direction vectors is on a circle with common center c but with different radii $r_j = \arccos(c'x_j)$, $j = 1, \ldots, K$. Denote the collection of concentric circles with a common center c and radii tuple $\mathbf{r} = (r_1, \ldots, r_K) \in [0, \pi/2] \times [0, \pi]^{K-1}$ by

$$\delta(c, \mathbf{r}) = \{(x_1, \ldots, x_K) \in (S^2)^K : x_j'c = \cos(r_j),\ j = 1, \ldots, K\}.$$

To work with observations on S^2, the geodesic distance function $d_g : S^2 \times S^2 \to [0, \pi]$ is defined by the arc length of the shortest great circle segment joining $x, y \in S^2$, and is $d_g(x, y) = \arccos(x'y)$. We further define $d_g(x, A) = \inf_{y \in A} d_g(x, y)$ for $x \in S^2, A \subset S^2$. For a random element X whose domain is S^2, a sensible notion of mean $\mu(X)$ is defined by a minimizer of mean squared distance,

$$\mu(X) = \operatorname*{argmin}_{x \in S^2} \mathbb{E}\{d_g^2(x, X)\},$$

often called the geodesic or Fréchet mean (Fréchet, 1948; Karcher, 1977). A useful measure of dispersion is geodesic variance which is defined as $\mathrm{Var}(X) = \mathbb{E}\{d_g^2(\mu(X), X)\} = \min_{x \in S^2} \mathbb{E}\{d_g^2(x, X)\}$ provided that $\mu(X)$ exists.

2 Single rotational deformations

In this section, an estimation procedure for rotational deformation models is proposed. We begin with a discussion on the simpler rigid rotation model.

2.1 Rigid rotation model

Suppose we have a K-tuple of random direction vectors $\mathbf{X} = (X_1, \ldots, X_K)$. For some unknown constants $c, \mu_j \in S^2$ and a latent random variable $\theta \in [-\pi/2, \pi/2)$, we model $X_j \in S^2$ $(j = 1, \ldots, K)$ as noisy observations of rotations of μ_j by $R(c, \theta)$, that is,

$$X_j = R(c, \theta)\mu_j \oplus \epsilon_j \quad (j = 1, \ldots, K). \tag{2}$$

Here, the ϵ_j are independently distributed random error terms, and the $\oplus$ sign defines a specific action of the error distribution as defined in the following.

There are several ways to define random spherical points $X \sim \mu \oplus \epsilon \in S^2$. A natural way is to introduce an S^2-valued distribution, e.g., the von Mises–Fisher distribution (Mardia and Jupp, 2000, p. 36) with the density $f_{\text{vMF}}(x; \mu, \kappa) \propto \exp(\kappa \mu' x)$ with respect to the uniform measure on S^2 for $\mu \in S^2$, $\kappa > 0$. Alternatively, one can utilize the tangent space at $\mu \in S^2$, allowing a distribution on the tangent space to be mapped to S^2. Another approach is to use the embedding of S^2 into $\mathbb{R}^3$, by scaling a three-dimensional random vector to unit length. This approach is often called a perturbation model (Goodall, 1991). It is well-known that a perturbation model introduces a bias in the estimation of the geodesic mean unless the distribution is isotropic (Kent and Mardia, 1997; Le, 1998; Huckemann, 2011a). In this paper in Section 4, we use the von Mises–Fisher distribution and in Section 5 the perturbation model. The latter is justified because we consider only isotropic distributions. In the following discussion, we do not specify a particular distribution for ϵ, but require that the geodesic mean of $X \sim \mu \oplus \epsilon$ is uniquely found at μ, i.e., $\mu = \operatorname{argmin}_x \mathbb{E}\{d_g^2(x, X)\}$. The geodesic variance $\operatorname{Var}(\epsilon) := \operatorname{Var}(\mu \oplus \epsilon)$ is then well defined.

In model (2), several different combinations of θ and μ_j lead to the same model. Specifically, replacing θ and μ_j by $\theta^*(a) = \theta - a$ and $\mu_j^*(a) = R(c, a)\mu_j$ for any $a \in \mathbb{R}$ gives the same model as (2). Therefore, we assume

$$\mathbb{E}\theta = 0. \tag{3}$$

The trajectory of rotated direction vectors forms a small circle (*cf.* Section 1.1), which

is approximately true in the presence of the noise. In other words, the collection of X_j in (2) are distributed along concentric circles with common center at c, as the following lemma states.

Lemma 1. *Let the η-neighborhood of concentric circles $\delta(c, \mathbf{r})$ be*

$$\delta_\eta(c, \mathbf{r}) = \{(X_1, \ldots, X_K) \in (S^2)^K : d_g(\delta(c, r_j), x_j) < \eta \text{ for all } j = 1, \ldots, K\},$$

for $\eta > 0$. If $\mu \oplus \epsilon_j$ are independent and identically distributed and spherically symmetric about μ, then

$$P\{\mathbf{X} \in \delta_\eta(c, \mathbf{r})\} \geq \left\{1 - \frac{Var(\epsilon)}{\eta^2}\right\}^K.$$

The auxiliary parameters $r_j = \arccos(c'\mu_j)$ represent the radii of the concentric circles, and are obtained from c and μ_j, the parameters of (2). A proof of Lemma 1 is given in the Appendix.

Lemma 1 suggests that $\mathbf{X}$ and $\delta(c, \mathbf{r})$ are close with high probability, which motivates to define the population concentric circles $\delta(c_0, \mathbf{r}_0)$ as a minimizer of squared loss. In the view of this estimation strategy, the capability of identifying parameters as minimizers leads to a natural estimation strategy, namely the M–estimation or the sample Fréchet mean (Karcher, 1977; Huckemann, 2011b). The rest of this section is devoted to the identification of the population parameters c, r_j and μ_j as population Fréchet means.

First, the distance function ρ between $\delta(c, \mathbf{r})$ and $\mathbf{x}$ is defined as the Cartesian product metric based on d_g by

$$\rho^2(\delta(c, \mathbf{r}), \mathbf{x}) = \sum_{j=1}^K d_g^2\left(\delta(c, r_j), x_j\right) = \sum_{j=1}^K \left(\arccos(x_j'c) - r_j\right)^2.$$

The collection of population concentric circles $\delta(c_0, \mathbf{r}_0)$ is defined as the Fréchet ρ-mean set

$$\operatorname*{argmin}_{c \in S^2, \mathbf{r} \in [0, \pi/2] \times [0, \pi]^{K-1}} \mathbb{E}\rho^2(\delta(c, \mathbf{r}), \mathbf{X}), \tag{4}$$

where the expectation $\mathbb{E}$ is with respect to the random directions $\mathbf{X}$. We assume in the following that there is a unique minimizer $\delta(c_0, \mathbf{r}_0)$. The center c_0 of the circles also represents the axis of rotation.

It should be noted that there is no guarantee for the true axis of rotation $\check{c}$ of (2) to be the same as c_0 from (4). Simulation studies, reported in the Supplementary Material, have suggested that the case $c_0 = \check{c}$ occurs when $\mathrm{Var}(r_j\theta_j)$ is large enough compared to the error

variance $\mathrm{Var}(\epsilon_j)$ for all j. In our simulation studies in Section 4, the effect of this bias is shown to be small.

While the axis of rotation c is the center of the concentric circles $\delta(c, \mathbf{r})$, each base point μ_j is also a point on $\delta(c, r_j)$, $j = 1, \ldots, K$. The assumption of isotropy of ϵ_j implies that

$$\mu(Y_j^{\theta_0}) = R(c, \theta_0)\mu_j$$

for $Y_j^{\theta_0} = R(c, \theta_0)\mu_j + \epsilon_j$ with deterministic angle $\theta_0 \in [-\pi/2, \pi/2)$. In particular, $\mu(Y_j^0) = \mu_j$. For random θ define

$$\mu(X_j|\theta) := R(c, \theta)\mu_j \,.$$

With the distance function $\rho_{\delta(c,r)}(x, y)$ which measures the shortest arc-length between $x, y \in \delta(c, r)$ along the (small) circle via $\rho_{\delta(c,r)}(x, y) = \sin(r) \arccos[(x'y - \cos^2(r))/\sin^2(r)]$ (Jung et al., 2012) we have by definition

$$\psi_0 = \underset{\psi \in [-\pi/2, \pi/2)}{\operatorname{argmin}} \mathbb{E}\rho^2_{\delta(c,r_j)}(\mu(X_j|\theta), R(c, \psi)\mu_j) = \underset{\psi \in [-\pi/2, \pi/2)}{\operatorname{argmin}} \mathbb{E}(\theta - \psi)^2,$$

which leads to the minimizer $\psi_0 = 0$ due to the assumption (3). Thus,

$$\mu_j = \underset{\mu \in \delta(c,r_j)}{\operatorname{argmin}} \mathbb{E}\rho^2_{\delta(c,r_j)}(\mu(X_j|\theta), \mu). \tag{5}$$

Finally, we view $\mu(X_j|\theta)$ as the expectation of X_j conditioned on the unobserved random variable θ which represents the amount of rotation. Then, by solving the equation $\mu(X_j|\theta) = R(c, \theta)\mu_j$ for θ, using the Rodrigues' rotation formula (Gray, 1980; Altmann, 2005) $R(c, \theta)\mu_j = \mu_j \cos\theta + (c \times \mu_j) \sin\theta + \langle c, \mu_j \rangle c(1 - \cos\theta)$, we get

$$\theta = \mathrm{atan2}[\langle \mu(X_j|\theta), c \times \mu_j \rangle, \langle \mu(X_j|\theta), \mu_j - c\cos(r_j) \rangle], \quad (j = 1, \ldots, K), \tag{6}$$

where the two argument function $\mathrm{atan2}(x_2, x_1) \in (-\pi, \pi]$ is the signed angle between two vectors $e_1 = (1, 0)$ and $(x_1, x_2) \in \mathbb{R}^2$.

2.2 Estimation

Suppose we have n independent observations $\mathbf{X}_1, \ldots, \mathbf{X}_n$ from model (2). Each $\mathbf{X}_i$ is a collection of K directions $\mathbf{X}_i = (X_{ij})_{j=1,\ldots,K}$. The estimates of parameters c, r_j, μ_j are obtained

as the sample Fréchet means as follows:

$$(\hat{c}, \hat{\mathbf{r}}) = \underset{c, r_1, \ldots, r_K}{\operatorname{argmin}} \sum_{i=1}^{n} \sum_{j=1}^{K} d_g^2\{\delta(c, r_j), X_{ij}\}, \tag{7}$$

$$\hat{\mu}_j = \underset{\mu}{\operatorname{argmin}} \sum_{i=1}^{n} \rho^2_{\delta(\hat{c}, \hat{r}_j)}(P_{(\hat{c}, \hat{r}_j)} X_{ij}, \mu) \quad (j = 1, \ldots, K). \tag{8}$$

Note that in (8), we have used $P_{(\hat{c},\hat{r}_j)}X_{ij}$, the projection of X_{ij} onto $\delta(\hat{c}, \hat{r}_j)$, instead of $\mathbb{E}(X_{ij}|\theta_i)$ used in (5). The projection $P_{\delta(c,r)}x$ is a point on $\delta(c, r)$ with the minimal geodesic distance to x, given by (Mardia and Gadsden, 1977, Eq. (3.3))

$$P_{\delta(c,r)}x = \underset{v \in \delta(c,r)}{\operatorname{argmin}}\, d_g(v, x) = \frac{x \sin(r) + c \sin\{d_g(x, c) - r\}}{\sin\{d_g(x, c)\}}.$$

The predicted values of the latent variable θ_i are obtained using (6) by substituting the estimates for the parameters. The predictor for θ_i is $\hat{\theta}_i = K^{-1} \sum_{j=1}^{K} \theta_{ij}$ for each $i = 1, \ldots, n$ with

$$\theta_{ij} = \operatorname{atan2}\{\langle P_{(\hat{c},\hat{r}_j)}X_{ij}, \hat{c} \times \hat{\mu}_j \rangle, \langle P_{(\hat{c},\hat{r}_j)}X_{ij}, \hat{\mu}_j - \hat{c}\cos(\hat{r}_j) \rangle\}. \tag{9}$$

The least squares problems (7-8) do not have closed form solutions. The problem (8) is simpler and the same as finding the geodesic mean of angles, since both $P_{(\hat{c},\hat{r}_j)}X_{ij}$ and μ are on the one-dimensional circle $\delta(\hat{c}, \hat{r}_j)$. Solutions to this type of problem are combinatorial (Moakher, 2002) but also found efficiently by numerical methods (Le, 2001; Fletcher et al., 2003). The problem (7) is precisely the fitting of concentric (small) circles. Therefore, numerical algorithms for (7) are generalized algorithms of the well-studied fitting of small circles (Mardia and Gadsden, 1977; Rivest, 1999; Jung et al., 2011, 2012) and are discussed in the Appendix.

2.3 Rotational deformations

The single rotation model (2) describes rigid rotations of objects. We extend the model to more general cases so that the generalized rotational model can explain, for example, non-rigid twisting or bending.

Suppose two direction vectors x_1 and x_2 are rotated about the same axis c but by different angles θ_1 and θ_2. This allows the underlying object to deform. In general, the assumption of a single rotation angle θ in (2) is relaxed to possibly different angles $\theta_1, \ldots, \theta_K$, which may be either independent or dependent of each other. To incorporate such general situations,

the single rotation model is generalized to

$$X_j = R(c, \theta_j)\mu_j \oplus \epsilon_j \quad (j = 1, \ldots, K). \tag{10}$$

The relationships among the θ_j can be specified using prior knowledge about the specific rotational deformation. As a special case, when a rigid rotation is assumed, it is reasonable to set $\theta_1 = \cdots = \theta_K$, which goes back to the model (2). The general model (10) includes other important physical deformations. The twisting or bending of the object can be modeled by different rotations with a common axis of rotation. As an example, when an object (and its attached direction vectors) is twisted, one group of direction vectors is rotated clockwise, while the other group is rotated counter–clockwise. Let I_1 and I_2 be a partition of the indices $\{1, \ldots, K\}$ representing groups of the direction vectors that rotate together. A simple twisting or bending motion can be obtained by assuming $\theta_i = -\theta_j$ for all $i \in I_1, j \in I_2$. Another example is the scenario of independent rotations where all directions in the same group rotate together ($\theta_i = \theta_j, i, j \in I_l$) but two angles in different groups are independent (θ_i and θ_j are independent for $i \in I_l, j \in I_k, 1 \leq l \neq k \leq 2$).

In all cases above, we assume that some functions f_j are known in advance, so that the relationships between θ_j are modeled through known functions, i.e., $\theta_j = f_j(\theta)$.

In the estimation of the parameters in (10), we use the fact that the estimation procedure (7) does not depend on specific assumptions of the latent variable θ_j. Therefore, the same least squares estimators $\{\hat{c}, \hat{r}_j, \hat{\mu}_j\}$ can be used to estimate the parameters of (10). When $f_j(\theta)$ is known and invertible, the prediction of the ith sample of θ, $\theta_{(i)}$, can be obtained. Since each θ_{ij} of (9) is a perturbed version of $f_j(\theta_i)$, the prediction of $\theta_{(i)}$ is then

$$\hat{\theta}_i = \frac{1}{K} \sum_{j=1}^{K} f_j^{-1}(\theta_{ij}).$$

Remark 1. A misspecification of the function f_j does not affect the estimation procedure (7), i.e., the estimation of the rotation axis. Nevertheless, the specification of f_j models the relationships between the rotation angles θ_j and is therefore crucial for their prediction as elaborated in Section 4 of the Supplementary Material. The partition I_1 and I_2 models f_j and is not a parameter of (7).

Example 1. The toy example presented in Fig. 2 is now discussed in more detail. The dataset consists of $n = 30$ observations of random twisting. The axis of twist is $c = (0, 1, 0)'$. The random angle θ follows $N(0, \sigma^2)$ with $\sigma \approx 22.5^\circ$ with $\theta_1 = \theta_2 = \theta$ and $\theta_3 = \theta_4 = -\theta$. The noise is independently added by a perturbation of $N_3(0, 0.1^2)$ on both the head and tail of

the direction vectors and then projected onto S^2.

The estimate $(\hat{c}, \hat{\mathbf{r}})$ was obtained by (7). The corresponding four concentric circles and the axis estimates $\hat{c}$ are overlaid in Fig. 3c. The estimate $\hat{c} = (0.007, 0.999, -0.031)'$ is only 1.8 degrees away from the truth. The base point estimates $\hat{\mu}_j$, predictions of θ_i, and the estimate of σ are also obtained, which are close to the truth. For example, $\hat{\sigma} = 21.5°$.

Despite a relatively small sample size ($n = 30$), the proposed estimator successfully estimated the axis of rotation, and leads to a clear visualization of the underlying rotational deformation, as depicted in Fig. 3.

3 Hierarchical rotations

We now discuss an estimation procedure for rotational deformations that consist of two independent generalized rotations. Such deformations include twisting and bending of the objects about different axes.

Suppose a set of base points μ_j is rotated by $R(c_1, \theta_j)$ and then by $R(c_2, \psi_j)$. The rotated random direction vector X_j is represented by

$$X_j = R(c_2, \psi_j)R(c_1, \theta_j)\mu_j \oplus \epsilon_j \quad (j = 1, \ldots, K), \tag{11}$$

with some error ϵ_j as seen in (2). The axes c_1, c_2 and the base points μ_j are unknown parameters and θ_j, ψ_j are independent latent variables representing rotation angles. The random direction vectors X_j have the same distribution as in

$$R'(c_2, \psi_j)X_j = R(c_1, \theta_j)\mu_j \oplus \epsilon_j \quad (j = 1, \ldots, K), \tag{12}$$

provided that the distribution of $\mu_j \oplus \epsilon_j$ is spherically symmetric about μ_j. The order of these rotations is not interchangeable because $R(c_2, \psi_j)R(c_1, \theta_j) \neq R(c_1, \theta_j)R(c_2, \psi_j)$ in general. Therefore, call the first rotation operation $R(c_1, \theta_j)$ the primary rotation, and $R(c_2, \psi_j)$ the secondary rotation.

With n observations, we have $X_{ij} = R(c_2, \psi_{ij})R(c_1, \theta_{ij})\mu_j \oplus \epsilon_{ij}$ $(i = 1, \ldots, n, j = 1, \ldots, K)$ and we wish to estimate the axes of rotations c_1, c_2 and predict the unobserved random variables ψ_{ij} and θ_{ij}. It is required to constrain the relationship among the ψ_j as a function of ψ (also for θ_j). Otherwise, θ_{ij} and ψ_{ij} will catch all sample fluctuations, leading to an overfit of c_1 and c_2. Let $\theta_j = f_{1j}(\theta)$ and $\psi_j = f_{2j}(\psi)$ for some known functions f_{1j} and f_{2j} $(j = 1, \ldots, K)$. For example, when the deformation is composed of rigid rotation and

twisting, the two functions can be modeled by $f_{1j}(\theta) = \theta$ and $f_{2j}(\psi) = 1_{j\in I_1}\psi - 1_{j\in I_2}\psi$, where 1 denotes the indicator function for a partition I_1, I_2 of $\{1, \ldots, K\}$. See also Section 4.

The estimates of the axes of rotation c_1, c_2 might be obtained as a solution of the least-squares problem

$$\min_{c_1,c_2} \sum_{i=1}^{n} \sum_{j=1}^{K} \rho_g^2\{X_{ij}, R(c_2, \psi_{ij})R(c_1, \theta_{ij})\mu_j\}.$$

Since the above problem is challenging to solve directly, we divide it into two subproblems, which can be solved iteratively. In a search for two subproblems, suppose first that (c_2, ψ_{ij}) of the secondary rotation are known. Then we would de-rotate X_{ij} by the action of $R(c_2, -\psi_{ij}) = R(c_2, \psi_{ij})'$ so that the unrotated direction vectors $X_{ij}^* = R(c_2, \psi_{ij})'X_{ij}$ are solely expressed by the primary rotation of μ_j. In other words,

$$X_{ij}^* = R(c_1, \theta_{ij})\mu_j \oplus \epsilon_{ij}. \tag{13}$$

Noticing the structure of the model is the same as (2), the least squares estimators of c_1, μ_j and the auxiliary parameters $r_{1j} = \arccos(c_1'\mu_j)$, for $j = 1, \ldots, K$, are given by

$$\begin{aligned} (\hat{c}_1, \hat{r}_{11}, \ldots, \hat{r}_{1K}) &= \operatorname*{argmin}_{c_1,(r_{11},\ldots,r_{1K})} \sum_{i=1}^{n} \sum_{j=1}^{K} d_g^2\{\delta(c_1, r_{1j}), X_{ij}^*\}, \\ \hat{\mu}_j &= \operatorname*{argmin}_{\mu} \sum_{i=1}^{n} d^2_{\delta(\hat{c}_1,\hat{r}_{1j})}(P_{(\hat{c}_1,\hat{r}_{1j})}X_{ij}^*, \mu). \end{aligned} \tag{14}$$

On the other hand, suppose we know in advance (c_1, θ_{ij}) of the primary rotation as well as the base points μ_j. Then we would rotate μ_j so that the secondary rotation is only needed to reach for the observations X_{ij} from the rotated base points $\mu_{ij}^* = R(c_1, \theta_{ij})\mu_j$. That is,

$$X_{ij} = R(c_2, \psi_{ij})\mu_{ij}^* \oplus \epsilon_{ij}. \tag{15}$$

The model (15) is different from models (2) and (13) as the base points μ_{ij}^* are different for different observations and assumed to be known. To estimate c_2, we modify (14) to incorporate the fact that μ_{ij}^* are varying but known, which leads to

$$\hat{c}_2 = \operatorname*{argmin}_{c_2} \sum_{i=1}^{n} \sum_{j=1}^{K} d_g^2\{\delta(c_2, r_{ij}^*(c_2)), X_{ij}\}, \tag{16}$$

where $r_{ij}^*(c_2) = \arccos(c_2'\mu_{ij}^*)$ is a function of c_2. For the estimation of c_1 and c_2 in (11),

we iteratively update $\hat{c}_1$ given $\hat{c}_2$ and then update $\hat{c}_2$ given $\hat{c}_1$. With the kth estimates $\hat{c}_1^{(k)}, \hat{c}_2^{(k)}, \tilde{\theta}_{ij}^{(k)}, \tilde{\psi}_{ij}^{(k)}$ of $c_1, c_2, \theta_{ij}, \psi_{ij}$ and the pre-specified functions $f_{1j}(\theta)$, $f_{2j}(\psi)$, the $(k+1)$th estimate is obtained as follows.

Step 1 Using $\hat{c}_2^{(k)}$ and $\hat{\psi}_{ij}^{(k)}$, obtain the de-rotated observations $X_{ij}^{(k)} = R(\hat{c}_2, -\hat{\psi}_{ij}^{(k)})X_{ij}$.

Step 2 Obtain $\hat{c}_1^{(k+1)}, \hat{r}_{1j}^{(k+1)}, \hat{\mu}_j^{(k+1)}$ using (14) with $X_{ij}^{(k)}$ in place of X_{ij}^*, and also the predictions $\tilde{\theta}_{ij}^{(k+1)}$ of θ_{ij} using (9) with $X_{ij}^{(k)}, \hat{c}_1^{(k+1)}, \hat{r}_{1j}^{(k+1)}, \hat{\mu}_j^{(k+1)}$ in place of $X_{ij}, \hat{c}, \hat{r}_j, \hat{\mu}_j$. Afterwards, use the function f_{1j} to predict θ_i as $\hat{\theta}_i^{(k+1)} = K^{-1}\sum_{j=1}^K f_{1j}^{-1}(\tilde{\theta}_{ij}^{(k+1)})$, so the predictions for θ_{ij} are updated as $\hat{\theta}_{ij}^{(k+1)} = f_{1j}(\hat{\theta}_i^{(k+1)})$.

Step 3 Rotate the base points to $\tilde{\mu}_{ij}^{(k+1)} = R(\hat{c}_1^{(k+1)}, \hat{\theta}_{ij}^{(k+1)})\hat{\mu}_j^{(k+1)}$ using the partial estimates $\hat{c}_1^{(k+1)}$ and predictions $\hat{\theta}_{ij}^{(k+1)}$.

Step 4 Obtain $\hat{c}_2^{(k+1)}$ as the minimizer of (16) with $\tilde{\mu}_{ij}^{(k+1)}$ in place of μ_{ij}^*, and also the predictions of rotation angles $\tilde{\psi}_{ij}^{(k+1)}$ using

$$\tilde{\psi}_{ij}^{(k+1)} = \text{atan2}(\langle X_{ij}, \hat{c}_2^{(k+1)} \times \tilde{\mu}_{ij}^{(k+1)} \rangle, \langle X_{ij}, \tilde{\mu}_{ij}^{(k+1)} - \langle \hat{c}_2^{(k+1)}, \tilde{\mu}_{ij}^{(k+1)} \rangle \hat{c}_2^{(k+1)} \rangle), \qquad (17)$$

which is similar to (9). Given $\tilde{\psi}_{ij}^{(k+1)}$, use the function f_{2j} to predict ψ_i by $\hat{\psi}_i^{(k+1)} = K^{-1}\sum_{j=1}^K f_{2j}^{-1}(\tilde{\psi}_{ij}^{(k+1)})$, leading to updated predictions for ψ_{ij} as $\hat{\psi}_{ij}^{(k+1)} = f_{2j}(\hat{\psi}_i^{(k+1)})$.

Step 5 If both $d_g(\hat{c}_1^{(k)}, \hat{c}_1^{(k+1)})$ and $d_g(\hat{c}_2^{(k)}, \hat{c}_2^{(k+1)})$ are negligible, stop. Otherwise, return to Step 1.

Numerical algorithms to solve the least-squares optimizations (14, 16) are similar to the problem (7), and are presented in the Appendix. In updating $\hat{c}_1^{(k+1)}$ and $\hat{c}_2^{(k+1)}$, one can use $\hat{c}_1^{(k)}$ and $\hat{c}_2^{(k)}$ as initial values for the optimization.

Remark 2. The initial values $\hat{c}_1^{(0)}$ and $\hat{c}_2^{(0)}$ may be suggested by a practitioner or obtained from an exploratory analysis. A careful choice is important to avoid convergence into local minima and is a topic of further studies. If the two rotational deformations are uncorrelated, a principal component analysis (PCA) may be used to find two major components as initial values. We propose to use the principal arc analysis (PAA) of Jung et al. (2011), which is a generalized PCA for data lying on $(S^2)^K$. Jung et al. (2011) argued that non-linear variation along small circles is better captured by PAA than by other extensions of PCA including Fletcher et al. (2004) and Huckemann et al. (2010). PAA is well suited to our problem, since the X_{ij} are distributed along small circles.

We now discuss how to use PAA to obtain initial values. For $\mathbf{X}_i = (X_{i1}, \ldots, X_{iK})' \in (S^2)^K$ $(i = 1, \ldots, n)$, PAA gives the mean $\mu^{PAA} = (\mu_1^{PAA}, \ldots, \mu_K^{PAA})' \in (S^2)^K$ and the

projections $\widetilde{\mathbf{X}}_{i(m)} = (\widetilde{X}_{i1(m)}, \ldots, \widetilde{X}_{iK(m)})' \in (S^2)^K$ onto the mth component, $m \in \{1, \ldots, M\}$ and M is the minimum of $2K$ and $n-1$. The first two components will be used to provide initial values. Which component corresponds to which rotational motion depends on the variance of θ and ψ. If $\mathrm{Var}(\theta)$ of the primary rotation is assumed to be larger than $\mathrm{Var}(\psi)$ of the secondary rotation, then the first component will provide an initial value for the primary rotation. In such a case, the solutions of (7) and (9) with $\widetilde{X}_{ij(1)}$ in place of X_{ij} are used as the initial values of $\hat{c}_1^{(0)}, \tilde{\theta}_{ij}^{(0)}$. Likewise, $\widetilde{X}_{ij(2)}$ are used to evaluate $\hat{c}_2^{(0)}, \tilde{\psi}_{ij}^{(0)}$. On the other hand, if $\mathrm{Var}(\theta) < \mathrm{Var}(\psi)$, then $\widetilde{X}_{ij(1)}$ is used for $\hat{c}_2^{(0)}$, and $\widetilde{X}_{ij(2)}$ for $\hat{c}_1^{(0)}$.

Remark 3. In contrast to single rotational deformations the function f_j effects the estimation of the rotation axes by the iterative back-and-forward rotation between two deformations which depend on the angle predictions. The order of the hierarchical deformation is specified by the primary and secondary information as well as by the functions f_{1j}, f_{2j}. Simulation studies, reported in Section 4 of the Supplementary Material, discuss the misspecification of f_j and a misspecified deformation order.

4 Numerical studies

In this section, we turn to the numerical performance of the proposed estimators. As our modeling approach is novel, there is no competing method to compare with. We study performance over several different rotational deformation situations.

Two different objects are studied. The first object (body 1), illustrated in Fig. 2, consists of $K = 4$ directions, while the second object (body 2) contains $K = 8$ direction vectors. The von Mises–Fisher distribution (Mardia and Jupp, 2000, p. 36) with concentration parameter κ, denoted as $\mathrm{vMF}(\kappa)$, is used for the distribution of errors. Three models (indexed by equation number) are considered for each object:

- Model (2)–*Rigid rotation*: $c = (1, 0, 0)'$, $\theta \sim N(0, \sigma_\theta^2)$ and $\sigma_\theta = \pi/12 \approx 15°$.
- Model (10)–*Twisting*: $c = (0, 1, 0)'$, $\theta_j = f_j(\theta) = 1_{j \in I_1}\theta - 1_{j \in I_2}\theta$, where $\theta \sim N(0, \sigma_\theta)^2$, $\sigma_\theta = \pi/8 \approx 22.5°$. Here, I_1 and I_2 are the partitions of $\{1, \ldots, K\}$ according to the right and left sides.
- Model (11)–*Hierarchical deformations*: $c_1 = (1, 0, 0)'$, $c_2 = (1/\sqrt{2}, -1/\sqrt{2}, 0)'$, $\theta_j = \theta$ and $\psi_j = f_j(\psi) = 1_{j \in I_1}\psi - 1_{j \in I_2}\psi$, where $\theta \sim N(0, \sigma_\theta)^2$, $\sigma_\theta \approx 22.5°$ and $\psi \sim N(0, \sigma_\psi)^2$, $\sigma_\psi \approx 15°$. The I_1 and I_2 are the same partition used in the twisting model above.

Table 1: Numerical performance over 1000 replications, for single deformation models.

Rotation (unit: degrees)		Body 1 $d_g(\hat{c}, c)$	Body 1 $\hat{\sigma}_\theta$ ($\sigma_\theta = 15$)	Body 2 $d_g(\hat{c}, c)$	Body 2 $\hat{\sigma}_\theta$ ($\sigma_\theta = 15$)
$\kappa = 100$	$n = 30$	4.133 (2.269)	15.248 (1.909)	2.905 (1.602)	15.018 (1.891)
	$n = 100$	2.235 (1.182)	15.365 (1.138)	1.560 (0.836)	15.114 (1.138)
$\kappa = 1000$	$n = 30$	1.166 (0.641)	14.896 (1.974)	0.841 (0.466)	14.881 (1.966)
	$n = 100$	0.655 (0.344)	15.012 (1.040)	0.448 (0.227)	14.982 (1.041)
Twisting (unit: degrees)		Body 1 $d_g(\hat{c}, c)$	Body 1 $\hat{\sigma}_\theta$ ($\sigma_\theta = 22.5$)	Body 2 $d_g(\hat{c}, c)$	Body 2 $\hat{\sigma}_\theta$ ($\sigma_\theta = 22.5$)
$\kappa = 100$	$n = 30$	2.761 (1.510)	22.647 (2.820)	4.055 (3.624)	22.785 (2.702)
	$n = 100$	1.482 (0.781)	22.753 (1.687)	1.883 (1.336)	22.705 (1.672)
$\kappa = 1000$	$n = 30$	0.803 (0.439)	22.344 (2.939)	1.017 (0.743)	22.336 (2.935)
	$n = 100$	0.446 (0.234)	22.484 (1.564)	0.536 (0.362)	22.474 (1.561)

In the last model, c_1 and c_2 are not orthogonal. Recall that we do not require any orthogonality of the two axes. The model thus represents hierarchical deformations by a rigid rotation and oblique twist.

For each model, we generate $n = 30, 100$ rotationally deformed objects with different error concentration parameters $\kappa = 100, 1000$. The proposed method is then applied to obtain the estimates $\hat{c}, \hat{\mu}_j$ and the predictions $\hat{\theta}_i$. These are replicated 1000 times, and the estimation quality is measured by $d_g(\hat{c}, c)$ and $\hat{\sigma}_\theta = \sum_{i=1}^n \hat{\theta}_i^2/n$ because the empirical mean is zero by equation (3).

Table 1 reports the mean and standard deviations of the measures, for the single deformation models. The quality of estimation is improved upon larger sample size and smaller error variance (larger κ). In the rigid rotation model, the performance is better for body 2. A main difference between the two bodies is the number of directions: $K = 4$ for body 1 and $K = 8$ for body 2. The higher accuracy observed for body 2, can be explained as the additional directions yielding more information. On the other hand, the performance in fitting the twisting model is inferior for body 2. An explanation is the close proximity of the axis c and the base points μ_j (to be rotated) in body 2, which is further illustrated in Fig. 4.

Figure 4 compares the estimates $\hat{c}$ from the single deformation models (2) and (10). Since $\hat{c} \in S^2$, these points are approximated by their images under the inverse exponential map (see Appendix) on the tangent plane $T_c S^2$ centered at c. For comparison, the scatter of vMF(100) is plotted at the top left panel. The rest of the top row shows the scatter of $\hat{c}$ from the rigid rotation model. In the bottom row, where the scatter of $\hat{c}$ from the twist model is plotted, the estimates of body 2 show a different pattern of scatter compared to body 1. This is due to a smaller radius $r_j = d_g(c, \mu_j)$ for some j, where the dispersion of the error is large compared to small r_j. When fitting the concentric circles, the smaller radius

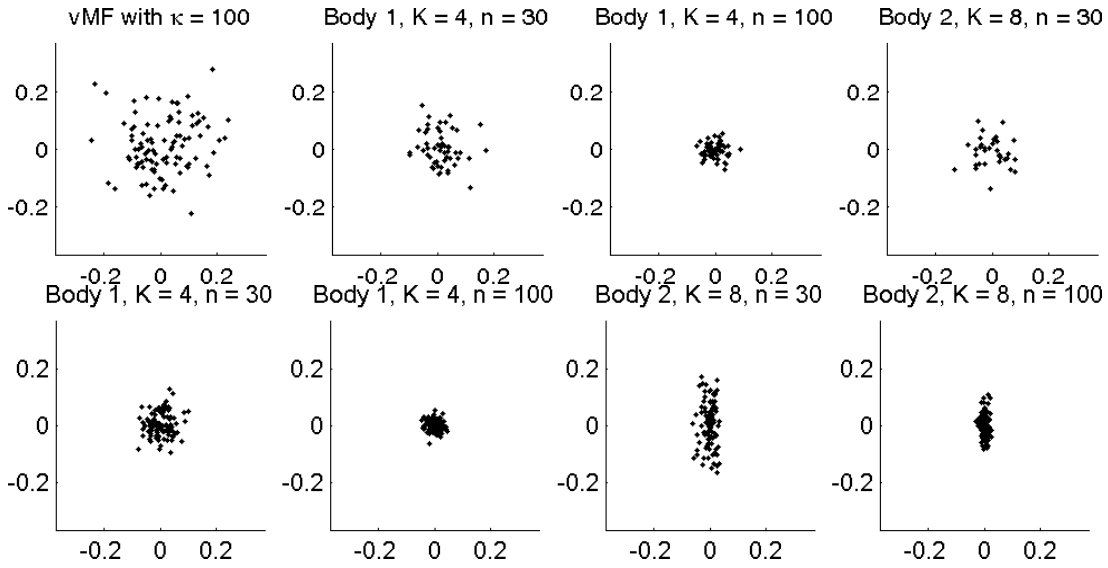


Figure 4: Scatters under their inverse exponential map of vMF(100) (top left panel) and of estimates $\hat{c}$ for the rigid rotation model (top row) and the twisting model (bottom row). The shape of the empirical covariance of $\hat{c}$ is different for the body 2–twist pair due to the proximity of the axis and base points.

$d_g(c, \mu_j)$ introduces large variance of the estimate $\hat{c}$.

The numerical performance in estimation of the hierarchical deformation model (11) is comparable to the simpler models. The results can be found in the Supplementary Material.

The asymptotic time complexity of the proposed algorithm is $O(nK)$ if a finite number of iterations is assumed as elaborated in Section 5 of the Supplementary Material. Furthermore, it is exemplified that the computation time increases, approximately linear in the number of samples n and the number of direction vectors K.

5 Analysis of s-rep deformations

In this section, an application of the proposed method to s-rep data is discussed.

5.1 S-reps of deformed ellipsoids

The skeletal representation (s-rep) gives a rich and efficient description of 3D objects (Siddiqi and Pizer, 2008; Pizer et al., 2013). The s-rep of human organs has been used to study structural and statistical properties and to promote precise segmentation of the organ from images. Accurate understanding of the deformations of prostates and hippocampi is crucial

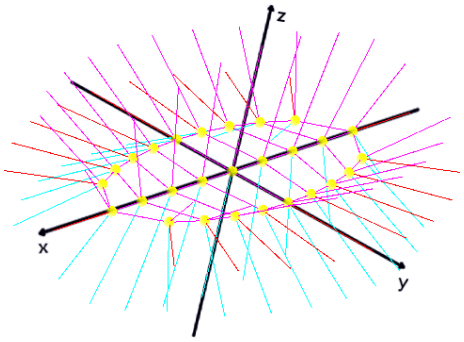


Figure 5: Fitted s-rep of a plain ellipsoid.

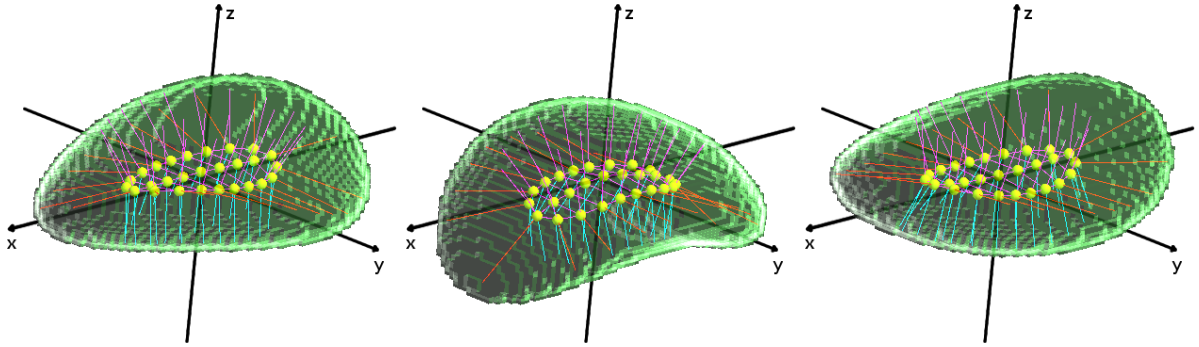
in medical operations. It has been observed that the major variation in the shape of these organs appears to be rotational deformations (Joshi et al., 2002; Jeong et al., 2008; Pizer et al., 2013), which motivated our analysis of s-rep data.

While our final goal is to analyze s-reps of real human organs, we work here with s-reps of deformed ellipsoids to validate the proposed method. A number of human organs, including hippocampi, prostates and bladders, are similar in shape to bent and twisted ellipsoids. Therefore, in our analysis, the rotationally deformed ellipsoids were used as a surrogate of real human organs. This enables us to the compare the estimate with the underlying truth.

An s-rep of a 3D object consists of a two-sided sheet of skeletal positions with spokes connecting the skeletal sheets to the boundary of the object. The skeletal sheet is nearly medial and the spokes are roughly normal to the boundaries, as defined in Pizer et al. (2013). We work with discrete s-reps, where the skeletal positions and the corresponding spokes are evaluated over a finite grid (9 by 3 in our analysis), as shown in Fig. 5.

The raw data are binary images of rotationally deformed ellipsoids, to which s-reps are fitted using the s-rep fitting procedure described in Pizer et al. (2013). The binary images are pre-processed by the anti-aliasing method of Niethammer et al. (2013). Figure 5 shows the fitted s-rep of a plain ellipsoid. The plain ellipsoid is centered at the origin with axes of lengths 3/4, 1/2 and 1/4, in x, y, z coordinates, and is a template for further deformation. The s-rep fitting for deformed ellipsoids is an iterative procedure using the s-rep of the plain ellipsoid as the initial value. The Supplementary Material contains a detailed description of the plain and deformed ellipsoids.

Three different rotational deformations of ellipsoids are examined: twisting, bending and a mix of those. In each experiment, $n = 30$ randomly deformed ellipsoids were obtained, followed by s-rep fitting. See Fig. 6 for the result of s-rep fitting for randomly deformed ellipsoids. The $K = 74$ spoke directions of each s-rep were recorded.



(a) Twisted ellipsoid (b) Bent ellipsoid. (c) Bent and twisted ellipsoid.

Figure 6: Examples of fitted s-reps for rotationally deformed ellipsoids. The surface of the raw ellipsoid, to which the s-rep is fitted, is overlaid.

5.2 Results

The first data set consists of fitted s-rep directions of twisted ellipsoids. The twisting was implemented as the rotation of the plain ellipsoid about the x-axis with angles proportional to $f_x(\theta) = \theta x$, where $x \in [-\frac{3}{4}, \frac{3}{4}]$ is the x-coordinate of the skeletal position of the spokes. The random angle θ is a zero-mean normal random variable with standard deviation $\sigma_\theta = 0.3$. From the $n = 30$ samples, we obtained an estimate of the rotation axis $\hat{c} = (0.99, 0.05, 0.12)'$, with $d_g(\hat{c}, c) = 7°$. The standard deviation estimate, $\hat{\sigma}_\theta = 0.29$, is close to the truth.

The bending deformation in the second experiment was realized as the rotation about the y-axis with angles proportional to $f_x(\alpha) = \alpha x$. Here, α follows $N(0, \sigma_\alpha^2)$ with $\sigma_\alpha = 0.4$. Similar to the estimation of twisting, an accurate estimate of the axis of rotation $\hat{c} = (0.01, 1.00, 0.01)'$ with distance $0.7°$ to the true axis was obtained from the sample of size $n = 30$. However, the estimate of σ_α was not consistent with the truth. This is so because the rotation angle of each direction is not consistent with $f_x(\alpha)$ due to the additional swing of the direction. The additional swing is a consequence of the change in surface curvature. Dealing with such an issue is beyond the scope of the current paper; it is discussed further in the Supplementary Material.

Finally, we report the results for bent and twisted ellipsoids. The raw ellipsoids were sequentially deformed by bending about the y-axis, then twisting about the x-axis. The initial values chosen by the data-driven method (see Remark 2 in Section 3) are $\hat{c}_1^0 = (-0.13, -0.99, -0.00)'$ and $\hat{c}_2^0 = (-0.07, -0.99, 0.02)'$, which are almost the same. A uniformly randomly chosen initial value for c_2 was used instead. In particular, a uniform random direction $\tilde{c}_2^0$ was used, provided that $\tilde{c}_2^0$ is at least 11 degrees away from $\hat{c}_1^0$. With this alternative initial value, the iterative estimation leads to estimates $\hat{c}_1 = (0.01, -1, -0.00)'$

and $\hat{c}_2 = (-0.99 - 0.05, 0.00)'$, both of which are close to their corresponding population counterparts. A simulation study, the report of which is omitted, confirms that the estimates are stable with respect to different choices of initial value $\tilde{c}_2^0$.

As we have pointed out in the introduction, the ellipsoid considered here can be understood as a template for many real human organs. The accurate estimation of the parameters of rotational deformations of ellipsoids indicates the potential of this type of analysis of deformed objects in real 3D images obtained from, e.g., magnetic resonance imaging. Further experiments cover surface point distribution models and a more general deformation; they are discussed in the Supplementary Material.

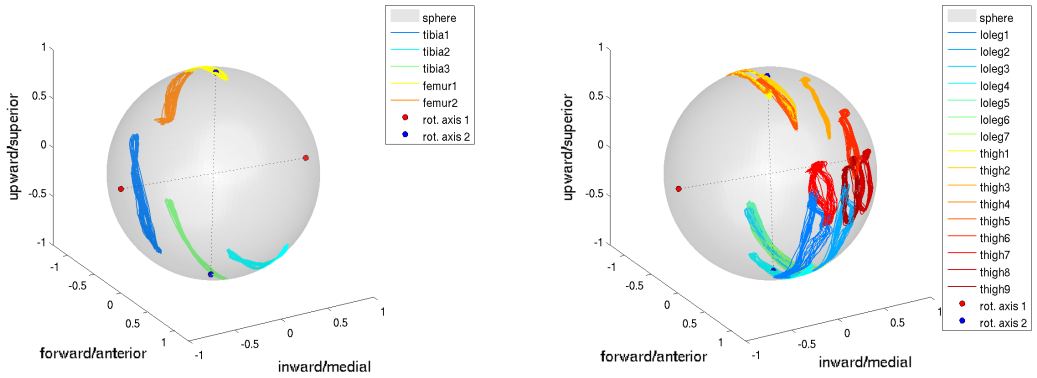
6 Application to knee motion during gait

In order to further support the validity of the proposed estimation procedure, this section presents findings from experimentally collected biomechanical data as a part of a larger project.

The estimation of two rotation axes of the knee joint is a well-studied problem in biomechanics (e.g., Ball and Greiner (2012)). The two estimated rotation axes model the primary and secondary rotation axes of the upper and lower leg relative to each other. The dominant rotation axis defines the flexion-extension motion at the knee. This axis is approximately directed right-to-left (lateral-to-medial). The secondary rotation axis defines the internal-external motion of the lower leg relative to the upper leg. This axis is approximately directed down-to-up (distal-to-proximal) along the long axis of the tibia (ankle-to-knee joint centers).

The motion of 25 markers placed on the right lower extremity of one healthy male volunteer was collected following informed consent. The volunteer consented to have two 6 mm pins surgically inserted into his femur and tibia. Insertion sites were selected to minimize neuro-muscular effects that could influence natural knee motion. Three and four markers were then attached to these rigid pins which allowed us to measure the true motion of the hidden femur and tibia bones. Additional markers were also placed on the surface of the thigh (10 skin markers) and lower leg (8 skin markers). In each of the four segments (femur, tibia, thigh, lower leg) one marker was chosen as a basis point and directions were derived between the basis point and the remaining markers of that segment. The coordinate system for this experiment was defined when the volunteer stood at attention and faced forwards. The XYZ axes were in the directions Forward, Inward, Upward (FIU).

The volunteer walked at 2.5 mph on a motor driven treadmill. After a familiarization



(a) bone (b) skin

Figure 7: Estimation of first (flexion-extension) and second (internal-external) rotation axes of the knee from bone and skin marker directions on the upper (femur; thigh) and lower (tibia; lower leg) extremity. In addition, the path of each marker direction is depicted. (a) Estimated rotation axes of directions derived from bone markers. (b) Estimated rotation axes of directions derived from skin markers.

time period, the motion of the markers were collected for approximately 20 seconds at 50 Hz. Within this data collection period, 16 complete gait cycles were identified. A gait cycle is defined from right foot contact with the floor to the next right foot contact. In total 976 time points were used within the following analyses.

Figure 7 shows the path of $K = 5$ bone directions and $K = 16$ skin directions of all 16 walking cycles. In addition to the least square estimations of (c_1, c_2), the rotation axes c_{1j} and c_{2j} were estimated for each marker $j = 1 \ldots, K$. Therewith, we can quantify the variation in the estimations by the dispersion measure $\sigma_{d_g(\hat{c}_1, \hat{c}_{1j})}$ of the geodesic distance $d_g(\hat{c}_1, \hat{c}_{1j})$ and $\sigma_{d_g(\hat{c}_2, \hat{c}_{2j})}$ respectively. The estimates are obtained by the least squares estimator (7) and the procedure is described in detail in Section 6 of the Supplementary Material.

The estimated rotation axes from the bone directions are $\hat{c}_1 = (-0.00, -1.00, 0.06)$ with dispersion $\sigma_{d_g(\hat{c}_1, \hat{c}_{1j})} = 9.41$ degrees and $\hat{c}_2 = (0.06, -0.06, -1.00)$ with $\sigma_{d_g(\hat{c}_2, \hat{c}_{2j})} = 12.30$ degrees. The unsigned directions of these two axes correspond to the anatomically expected flexion-extension and internal-external knee rotation axes (Grood and Suntay, 1983, Figure 6, page 142).

The two rotation axes estimated from the skin marker directions are very similar to the bone marker direction estimates. Quantitatively these were, $\hat{c}_1 = (0.01, -1.00, 0.07)$ with $\sigma_{d_g(\hat{c}_1, \hat{c}_{1j})} = 12.80$ degrees and $\hat{c}_2 = (0.01, 0.03, -1.00)$ with $\sigma_{d_g(\hat{c}_2, \hat{c}_{2j})} = 25.82$ degrees. In both cases, the higher dispersion of the second rotation axis can be explained by the smaller range of rotation angles. As expected, we observe more dispersion of the rotation axes using

the skin data because of the complex deformation of the skin surface. Estimation results of the rotation angles can be found in the Supplementary Material. Future work should improve the estimations and reduce the dispersion by a more careful time modelling of knee motions such as that proposed by Rivest (2001) who examined elbow motion.

7 Discussions

The paper proposes a novel method to estimate rotational deformations from directional data by concentric small circles. The estimation procedure is independent of the latent variable θ_j for single rotational deformations. In addition, the paper proposed an estimation procedure for hierarchical deformations depending on functions f_{1j}, f_{2j} and the ordering of the initial estimation. An interesting topic for future research is improvement of the prediction of θ_j, which includes the automatic classification of directions into a partition I_1 and I_2 as well as the development of methods to predict f_j from the data. The geometry of composing deformations has to be studied in further detail in order to extend the estimation method to more than two rotational deformations. A first step in decreasing the relevance of the order of deformations would be the implementation of an expectation maximization (EM) based optimization procedure.

Appendix

Proof of Lemma 1

Let $A_j = \{v_j \in \delta_\eta(c, r_j)\}$, where $\delta_\eta(c, r) = \{x \in S^2 : d_g(\delta(c, r), x) < \eta\}$. For $R = R(c, \theta)$,

$$\begin{aligned} P(A_j) &= P[d_g\{R\mu_j \oplus \epsilon_j, \delta(c, r_j)\} < \eta] = P[d_g\{R^T(R\mu_j \oplus \epsilon_j), R^T\delta(c, r_j)\} < \eta] \\ &= P[d_g\{\mu_j \oplus \epsilon, \delta(c, r_j)\} < \eta] \geq P[d_g(\mu_j \oplus \epsilon, \mu_j) < \eta] \geq 1 - \mathrm{Var}(\epsilon)/\eta^2, \end{aligned}$$

where $\mu_j \oplus \epsilon$ has the same distribution as $R^T(R\mu_j \oplus \epsilon_j)$ because of the spherical symmetry. A Markov inequality is used. Since the A_js are independent, $P(\bigcap_{j=1}^K A_j) = \prod_{j=1}^K P(A_j) \geq \{1 - \mathrm{Var}(\epsilon)/\eta^2\}^K$.

Numerical Algorithms for (7), (14), and (16)

The optimization problems (7) and (14) are identical and can be understood as fitting concentric circles on the unit sphere. The problem (16) is a more general nonlinear least squares problem, which however can be solved in a similar manner to the former two problems. We propose a variant of the doubly iterative algorithm used in fitting small circles in S^m (Jung et al., 2011, 2012).

We first introduce some notation. For $m \geq 2$, the tangent space at $c \in S^m$ is denoted by $T_c S^m$, which can be parametrized by $\mathbb{R}^m$. Let $c = e_{m+1}$ without loss of generality. The exponential map $\mathrm{Exp}_c : \mathbb{R}^m \to S^m$ is defined for $v_1 \in \mathbb{R}^m$ by

$$\mathrm{Exp}_c(v_1) = \left(\frac{v_1}{\|v_1\|} \sin \|v_1\|, \cos \|v_1\| \right),$$

with a convention of $\mathrm{Exp}_c(0) = c$. The exponential map has an inverse, called the log map, and is denoted by $\mathrm{Log}_c : S^m \to T_c S^m$.

For problems (7) and (14), the following iterative algorithm can be used. The algorithm finds a suitable point of tangency $\hat{c}_0$, which is also the center of the fitted concentric circles. Given the candidate $\hat{c}_0$, the data x_{ij} are mapped to the tangent space $T_{\hat{c}_0} S^2$ by the Log map. Let $x_{ij}^\dagger = \mathrm{Log}_{\hat{c}_0}(x_{ij})$. Since the Log map preserves distance, we have $\arccos(\hat{c}_0' x_{ij}) = \|x_{ij}^\dagger\|$. Then we solve a non-linear least-squares problem

$$\min_{c^\dagger, r_j} \sum_{i=1}^{n} \sum_{j=1}^{K} (\|x_{ij}^\dagger - c^\dagger\| - r_j)^2. \tag{18}$$

Since the optimization problem (18) does not have any constraint, it can be numerically solved by, e.g., the Levenberg–Marquardt algorithm (Scales, 1985). The solution $c^\dagger$ is then mapped to S^2 by the exponential map at c and becomes $\hat{c}_1$. This procedure is repeated until $\hat{c}$ converges.

The optimization problem (16) can be solved in a similar way. We use the fact that $\rho_g^2(\delta(c, r^*(c)), x) = (\arccos(c'x) - \arccos(c'\mu^*))^2 = (\|\mathrm{Log}_c x\| - \|\mathrm{Log}_c \mu^*\|)^2$. Thus for fixed c, $\rho_g^2(\delta(c, r^*(c)), x) \geq \min_y (\|\mathrm{Log}_c x - y\| - \|\mathrm{Log}_c \mu^* - y\|)^2$. The minimizer y leads to a better candidate for c through the exponential map. The algorithm to solve (16) follows the same lines as the algorithm to solve (7), except instead of (18) we minimize

$$\min_{c^\dagger} \sum_{i=1}^{n} \sum_{j=1}^{K} (\|\mathrm{Log}_{\hat{c}} x - c^\dagger\| - \|\mathrm{Log}_{\hat{c}} \mu - c^\dagger\|)^2.$$

Supplementary Materials

Additional discussions and data analyses: Article containing *i.)* additional data analysis results, *ii.)* simulation results for the hierarchical deformation model described in Section 4, *iii.)* further discussion of the model bias, brought up in Section 2.1, *iv.)* study of the estimator behaviour using misspecified parameters, *v.)* a computational complexity study of the algorithm and *vi.)* the estimation procedure for knee motion analysis during gait as discussed in Section 6. (SupplementaryMaterialSJ.pdf)

Matlab code: A set of Matlab code for application of the proposed method. The code also contains all datasets used as examples in the article. (estRotDeformation.zip)

References

Altmann, S. L. (2005), *Rotations, Quaternions, and Double Groups*, Dover books on mathematics, Dover Publications.

Ball, K. A. and Greiner, T. M. (2012), "A Procedure to Refine Joint Kinematic Assessments: Functional Alignment," *Computer Methods in Biomechanics and Biomedical Engineering*, 15, 487–500.

Chang, T. (1986), "Spherical Regression," *The Annals of Statistics*, 14, 907–924.

— (1988), "Estimating the Relative Rotation of Two Tectonic Plates from Boundary Crossings," *Journal of the American Statistical Association*, 83, 1178–1183.

— (1989), "Spherical Regression with Errors in Variables," *The Annals of Statistics*, 17, 293–306.

Chang, T. and Rivest, L.-P. (2001), "M-Estimation for Location and Regression Parameters in Group Models: A Case Study using Stiefel Manifolds," *The Annals of Statistics*, 29, 784–814.

Cootes, T. F., Taylor, C., Cooper, D., and Graham, J. (1992), "Training Models of Shape from Sets of Examples," in *Proc. British Machine Vision Conference*, eds. Hogg, D. and Boyle, R., Berlin. Springer-Verlag, pp. 9–18.

Dryden, I. L. and Mardia, K. V. (1998), *Statistical Shape Analysis*, Chichester: Wiley.

Fisher, N. I., Lewis, T., and Embleton, B. J. J. (1993), *Statistical Analysis of Spherical Data*, Cambridge: Cambridge University Press.

Fletcher, P., Lu, C., and Joshi, S. (2003), "Statistics of Shape via Principal Geodesic Analysis on Lie Groups," in *Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on*, vol. 1, pp. I–95–I–101 vol.1.

Fletcher, P. T., Lu, C., Pizer, S. M., and Joshi, S. (2004), "Principal Geodesic Analysis for the Study of Nonlinear Statistics of Shape," *IEEE Transactions on Medical Imaging*, 23, 995–1005.

Fréchet, M. (1948), "Les Éléments Aléatoires de Nature Quelconque dans un Espace Distancié," *Annales de l'Institut Henri Poincaré*, 10, 215–310.

Gamage, S. S. and Lasenby, J. (2002), "New Least Squares Solutions for Estimating the Average Centre of Rotation and the Axis of Rotation," *Journal of Biomechanics*, 35, 87–93.

Goodall, C. R. (1991), "Procrustes Methods in the Statistical Analysis of Shape (with discussion)," *Journal of the Royal Statistical Society: Series B*, 53, 285–339.

Gray, J. J. (1980), "Olinde Rodrigues' Paper of 1840 on Transformation Groups," *Archive for History of Exact Sciences*, 21, 375–385.

Grood, E. and Suntay, W. (1983), "A joint coordinate system for the clinical description of three-dimensional motions: Application to the knee," *Journal of Biomechanical Engineering*, 105, 136–144.

Halvorsen, K., Lesser, M., and Lundberg, A. (1999), "A new Method for Estimating the Axis of Rotation and the Center of Rotation," *Journal of Biomechanics*, 32, 1221–1227.

Hanna, M. S. and Chang, T. (2000), "Fitting Smooth Histories to Rotation Data," *Journal of Multivariate Analysis*, 75, 47–61.

Huckemann, S. (2011a), "Inference on 3D Procrusted Means: Tree Bole Growth, Rank Deficient Diffusion Tensors and Perturbation Models," *Scandinavian Journal of Statistics*, 38, 424–446.

— (2011b), "Intrinsic Inference on the Mean Geodesic of Planar Shapes and Tree Discrimination by Leaf Growth," *Annals of Statistics*, 39, 1098–1124.

Huckemann, S., Hotz, T., and Munk, A. (2010), "Intrinsic Shape Analysis: Geodesic PCA for Riemannian Manifolds modulo Isometric Lie Group Actions," *Statistica Sinica*, 20, 1–58.

Jeong, J.-Y., Stough, J. V., Marron, J. S., and Pizer, S. M. (2008), "Conditional-Mean Initialization Using Neighboring Objects in Deformable Model Segmentation," in *SPIE Medical Imaging*.

Joshi, S., Pizer, S. M., Fletcher, P., Yushkevich, P., Thall, A., and Marron, J. S. (2002), "Multiscale Deformable Model Segmentation and Statistical Shape Analysis Using Medial Descriptions," *IEEE Transactions on Medical Imaging*, 21, 538–550.

Jung, S., Dryden, I. L., and Marron, J. S. (2012), "Analysis of Principal Nested Spheres," *Biometrika*, 99, 551–568.

Jung, S., Foskey, M., and Marron, J. S. (2011), "Principal Arc Analysis on Direct Product Manifolds," *Annals of Applied Statistics*, 5, 578–603.

Karcher, H. (1977), "Riemannian Center of Mass and Mollifier Smoothing," *Communications on Pure and Applied Mathematics*, 30, 509–541.

Kent, J. T. and Mardia, K. V. (1997), "Consistency of Procrustes Estimators," *Journal of the Royal Statistical Society: Series B*, 59, 281–290.

Kurtek, S., Ding, Z., Klassen, E., and Srivastava, A. (2011), "Parameterization-Invariant Shape Statistics and Probabilistic Classification of Anatomical Surfaces," in *Information Processing in Medical Imaging*, vol. 22, pp. 147–158.

Le, H. (1998), "On the Consistency of Procrustean Mean Shapes," *Advances in Applied Probability*, 30, 53–63.

— (2001), "Locating Fréchet Means with Application to Shape Spaces," *Advances in Applied Probability*, 33, 324–338.

Mardia, K. V. and Gadsden, R. J. (1977), "A Circle of Best Fit for Spherical Data and Areas of Vulcanism," *Journal of the Royal Statistical Society: Series C*, 26, 238–245.

Mardia, K. V. and Jupp, P. E. (2000), *Directional Statistics*, Chichester: Wiley.

Moakher, M. (2002), "Means and Averaging in the Group of Rotations," *SIAM Journal on Matrix Analysis and Applications*, 24, 1–16 (electronic).

Niethammer, M., Juttukonda, M. R., Pizer, S. M., and Saboo, R. R. (2013), "Anti-Aliasing Slice-Segmented Medical Images via Laplacian of Curvature Flow," *In preparation*.

Oualkacha, K. and Rivest, L.-P. (2012), "On the Estimation of an Average Rigid Body Motion," *Biometrika*, 99, 585–598.

Pennec, X. (2008), "Statistical Computing on Manifolds: from Riemannian Geometry to Computational Anatomy," *Emerging Trends in Visual Computing*, 5416, 347–386.

Pizer, S. M., Jung, S., Goswami, D., Zhao, X., Chaudhuri, R., Damon, J. N., Huckemann, S., and Marron, J. S. (2013), "Nested Sphere Statistics of Skeletal Models," in *Innovations for Shape Analysis: Models and Algorithms*, eds. Breuß, M., Bruckstein, A., and Maragos, P., Springer Lecture Notes in Computer Science, pp. 93–115.

Rancourt, D., Rivest, L.-P., and Asselin, J. (2000), "Using Orientation Statistics to Investigate Variations in Human Kinematics," *Journal of the Royal Statistical Society: Series C*, 49, 81–94.

Rivest, L.-P. (1989), "Spherical Regression for Concentrated Fisher-Von Mises Distributions," *The Annals of Statistics*, 17, 307–317.

— (1999), "Some Linear Model Techniques for Analyzing Small-Circle Spherical Data," *Canadian Journal of Statistics*, 27, 623–638.

— (2001), "A Directional Model for the Statistical Analysis of Movement in Three Dimensions," *Biometrika*, 88, 779–791.

— (2006), "Regression and Correlation for 3×3 Rotation Matrices," *Canadian Journal of Statistics*, 34, 1–17.

Rivest, L.-P., Baillargeon, S., and Pierrynowski, M. (2008), "A Directional Model for the Estimation of the Rotation Axes of the Ankle Joint," *Journal of the American Statistical Association*, 103, 1060–1069.

Rohde, G. K., Ribeiro, A. J. S., Dahl, K. N., and Murphy, R. F. (2008), "Deformation-Based Nuclear Morphometry: Capturing Nuclear Shape Variation in HeLa Cells," *Cytometry A*, 73, 341–350.

Scales, L. E. (1985), *Introduction to Nonlinear Optimization*, New York: Springer-Verlag.

Siddiqi, K. and Pizer, S. (2008), *Medial Representations: Mathematics, Algorithms and Applications*, Computational Imaging and Vision, Vol. 37, Dordrecht, Netherlands: Springer, 1st ed.

Teu, K. K. and Kim, W. (2006), "Estimation of the Axis of a Screw Motion from Noisy Data–A New Method Based on Plücker Lines," *Journal of Biomechanics*, 39, 2857–2862.

Supplementary Material: Analysis of Rotational Deformations from Directional Data

Jörn Schulz[1] *, Sungkyu Jung[2], Stephan Huckemann[3], Michael Pierrynowski[4], J. S. Marron[5] and Stephen M. Pizer[5]

[1]University of Tromsø, [2]University of Pittsburgh, [3]University of Göttingen, [4]McMaster University, Hamilton [5]University of North Carolina at Chapel Hill

1 Additional data analysis

In this section, some additional data analyses are discussed, in particular using datasets from point distribution models.

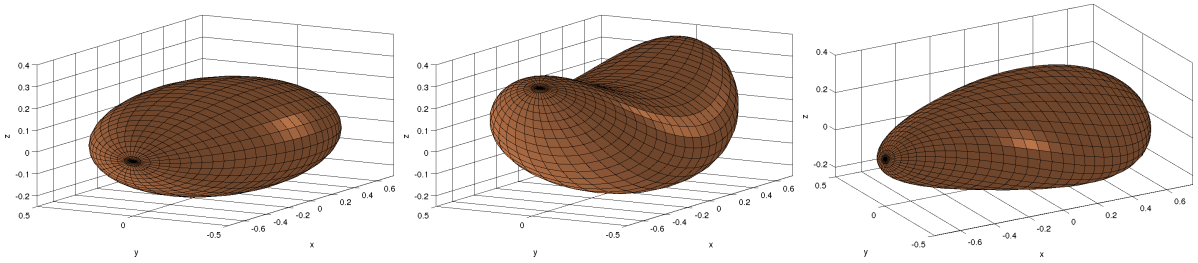
1.1 Rotationally deformed ellipsoids

We first discuss the procedures of obtaining the raw ellipsoid data. To avoid notational confusion, we denote a random vector by X_{ij} and their observed values by χ_{ij}, $i = 1, \ldots, n, j = 1, \ldots, K$. A point in $\mathbb{R}^3$ is described by $(x, y, z) \in \mathbb{R}^3$ in a fixed Cartesian coordinate system.

The surface of an ellipsoid can be parameterized by

$$\mathbf{s}(u,v) = \begin{pmatrix} x(u,v) \\ y(u,v) \\ z(u,v) \end{pmatrix} = \begin{pmatrix} r_a \sin(v) \\ r_b \sin(u)\cos(v) \\ r_c \cos(u)\cos(v) \end{pmatrix}, \qquad u \in [-\pi, \pi), v \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \tag{1.1}$$

*This research was funded by the Norwegian Research Council through grant 176872/V30 in the eVita program and performed as part of Tromsø Telemedicine Laboratory.



(a) Original (b) Bent (c) Twisted

Figure 1.1: Tube views of meshed surfaces of (a) an original ellipsoid with $r_a = 0.75$, $r_b = 0.5$ and $r_c = 0.25$, (b) bent ellipsoid with $\alpha = 0.6$, (c) twisted ellipsoid with $\theta = 0.6$.

where $r_a \geq r_b \geq r_c > 0$ are the length of the axes. We assume a default ellipsoid of $r_a = 0.75$, $r_b = 0.5$ and $r_c = 0.25$. For a parameter space $\Omega = [-\pi, \pi) \times [-\frac{\pi}{2}, \frac{\pi}{2}] \subset \mathbb{R}^2$, $\mathbf{s} : \Omega \to \mathbb{R}^3$ is a surface map in $\mathbb{R}^3$ with $(u, v) \xrightarrow{\mathbf{s}} (x(u, v), y(u, v), z(u, v))$. Without loss of generality, we assume that the first to third principal axes are x, y, z-axis in $\mathbb{R}^3$, respectively. The centroid of the ellipsoid is at origin $(0, 0, 0)'$.

Rotational bending around the y-axis ($c_b = (0, 1, 0)'$) is given by

$$B(\mathbf{s}) = R(c_b, f_b(x))\mathbf{s}(u, v), \tag{1.2}$$

and twisting around the x-axis ($c_t = (1, 0, 0)'$) is given by

$$T(\mathbf{s}) = R(c_t, f_t(x))\mathbf{s}(u, v), \tag{1.3}$$

where

$$\begin{aligned} R(c_b, f_b(x)) &= \begin{pmatrix} \cos(f_b(x)) & 0 & -\sin(f_b(x)) \\ 0 & 1 & 0 \\ \sin(f_b(x)) & 0 & \cos(f_b(x)) \end{pmatrix}, \\ R(c_t, f_t(x)) &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(f_t(x)) & -\sin(f_t(x)) \\ 0 & \sin(f_t(x)) & \cos(f_t(x)) \end{pmatrix}. \end{aligned}$$

Here, $f_b(x) = \alpha x$ and $f_t(x) = \theta x$ for some $\alpha, \theta \in [-\pi/2, \pi/2]$ representing the overall size of the deformation. The amount of bending and twisting depends on the location on the x-axis. In addition to the rotational bending, we also consider quadratic bending around the

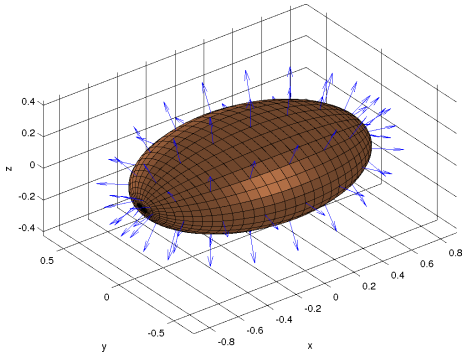


Figure 1.2: Visualization of a standard ellipsoid with attached boundary normals.

y-axis which is defined by

$$B_q(\mathbf{s}) = (x, y, z + \alpha x^2)', x, y, z, \in \mathbb{R} \tag{1.4}$$

using the ellipsoid parametrization (1.1). In the following, the term *bending* is used for rotational bending, and *quadratic bending* will be specified explicitly.

Figure 1.1 shows an example of an original ellipsoid, bent ellipsoid with $\alpha = 0.6$ and twisted ellipsoid with $\theta = 0.6$.

1.2 Point distribution and boundary normals

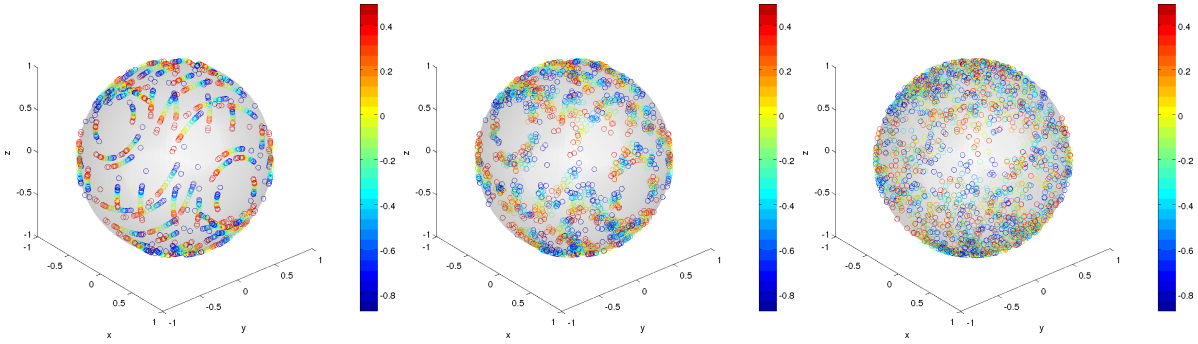
We now discuss a parameterization of ellipsoids by a point distribution model. In particular, a mesh grid and attached boundary normals of the surface $\mathbf{s}(u, v)$ will be used. See Fig. 1.2.

The size of surface mesh we used is 37×33. We chose $K = 9 \times 8 = 72$ vertices to attach normal direction vectors $\chi_{k_1 k_2} \in S^2, k_1 = 1, \ldots, 9, k_2 = 1, \ldots, 8$. For each k_1, the vertices where $\chi_{k_1 k_2}$ are attached have common x-coordinates

$$\{-0.738, -0.649, -0.482, -0.256, 0, 0.256, 0.482, 0.649, 0.738\},$$

the values of which influence the amount of deformation. Note that there are some directions that will not vary when the object is deformed. For example, the normals χ_{5k_2} $(1 \leq k_2 \leq 8)$, which are attached to vertices with zero x-coordinate, will stay still when the object is twisted or bent.

In the following we report four sets of experiments on the boundary normal ellipsoid data. As opposed to the s-rep data analysis, we are working with the raw data directly. The noise in the data is from the von Mises–Fisher distribution (Mardia and Jupp, 2000)



(a) without noise (b) $\kappa = 1000$ (c) $\kappa = 100$

Figure 1.3: (Twisting) Boundary normals using different concentration parameter κ. The rotation around the x-axis is clearly visible in (a). Different colors represent different amounts of twisting parameter θ.

with concentration parameter $\kappa > 0$. For each experiment, two levels of noise are used: $\kappa = 100, 1000$. The four sets of models are

- *Twisting* by (1.3), with $c_t = (1, 0, 0)'$, $\theta \sim N(0, \sigma_\theta^2)$, $\sigma_\theta = 0.3 \approx 17.2°$. See Fig. 1.3.
- *Bending* by (1.2) with $c_b = (0, 1, 0)'$, $\alpha \sim N(0, \sigma_\alpha^2)$, $\sigma_\alpha = 0.4 \approx 22.9°$. See Fig. 1.4.
- *Quadratic bending* by (1.4) with above c_b and σ_α. See Fig. 1.5.
- *Hierarchical deformation* by bending (primary) and twisting (secondary) with the same set of parameters above. See Fig. 1.7.

From each model a random sample of size $n = 30$ or 100 is obtained, from which the estimate $\hat{c}$ of the axis and $\hat{\sigma}$ are obtained. This is repeated for 1000 times.

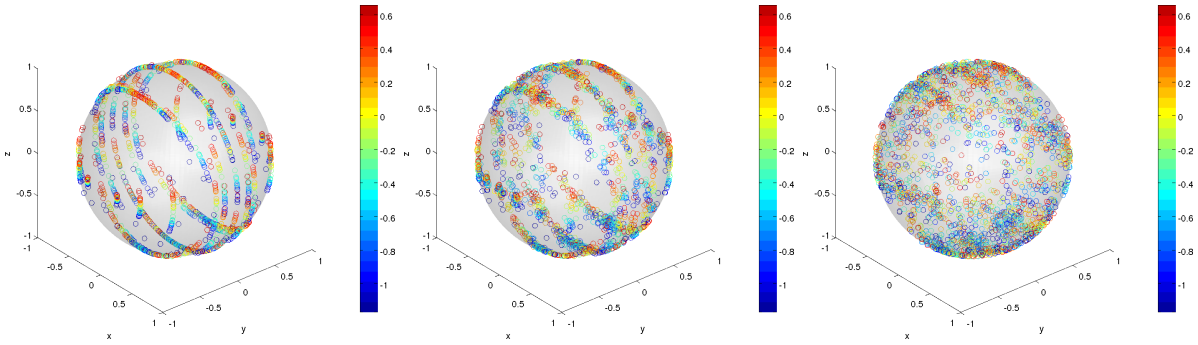
Twisting

Figure 1.3 shows 30 samples of 72 different normal directions from sets of twisted ellipsoids with different noise levels. Different colors represent different amounts of twisting parameter θ. Therefore, the number of colors correspond to the sample size.

Table 1.1 shows the performance of our estimator based on 1000 simulations. The performance is measured by the mean and standard deviation of the absolute error $d_g(\hat{c}, c)$ and those of the estimated twisting parameter $\hat{\sigma}_\theta$. The accuracy is increased for larger n or κ. In general, we observe quite accurate estimates even for a larger noise level ($\kappa = 100$).

Table 1.1: Twisting: Mean absolute error for $\hat{c}$, and the estimates $\hat{\sigma}_\theta$

Twisting (unit: degrees)			
κ	n	$d_g(\hat{c}, c)$	$\hat{\sigma}_\theta(\sigma_\theta = 17.189)$
100	30	3.174 (2.294)	17.209 (2.152)
100	100	1.563 (0.890)	17.324 (1.250)
1000	30	0.561 (0.317)	17.045 (2.133)
1000	100	0.289 (0.164)	17.173 (1.235)



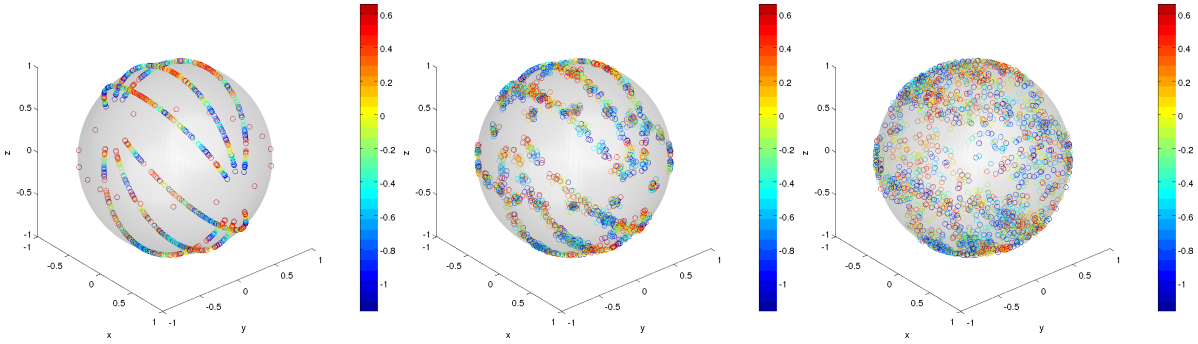
(a) without noise (b) $\kappa = 1000$ (c) $\kappa = 100$

Figure 1.4: (Rotational bending) Boundary normals using different concentration parameter κ. Different colors represent different amounts of bending parameter α.

Bending

Figure 1.4 shows 30 samples of 72 different normal directions after rotational bending using different noise levels. The case of quadratic bending is illustrated in Fig. 1.5. Different colors represent different amounts of bending parameter α. The different effects of rotational and quadratic bending to the boundary normals can be compared by examination of Fig. 1.4a and Fig. 1.5a. Rotationally bent directions are at concentric small circles (Fig. 1.4a) while quadratically bent directions are at circles with different centers (Fig. 1.5a). Despite the major violation of our assumption of concentric circles, the proposed method surprisingly works well for quadratic bending models, as Table 1.2 summarizes.

Table 1.2 shows the performance of our estimator for the rotational and quadratic bending models. The absolute errors between the true axis c and the estimated rotation axis $\hat{c}$ are virtually small for both models. The performance of the estimator is enhanced for larger κ and n. Moreover, as expected, the empirical errors are smaller for rotational bending than quadratic bending. Note that the estimates $\hat{\sigma}_\alpha$ of rotational bending model are biased upwards, which we discuss in section 1.2.1. The parameter σ_α affects the quality of estimators.



(a) without noise (b) $\kappa = 1000$ (c) $\kappa = 100$

Figure 1.5: (Quadratic bending) Boundary normals using different concentration parameter κ. Different colors represent different amounts of bending parameter α.

Table 1.2: Rotational bending: Mean absolute error for $\hat{c}$, and the estimates $\hat{\sigma}_\alpha$

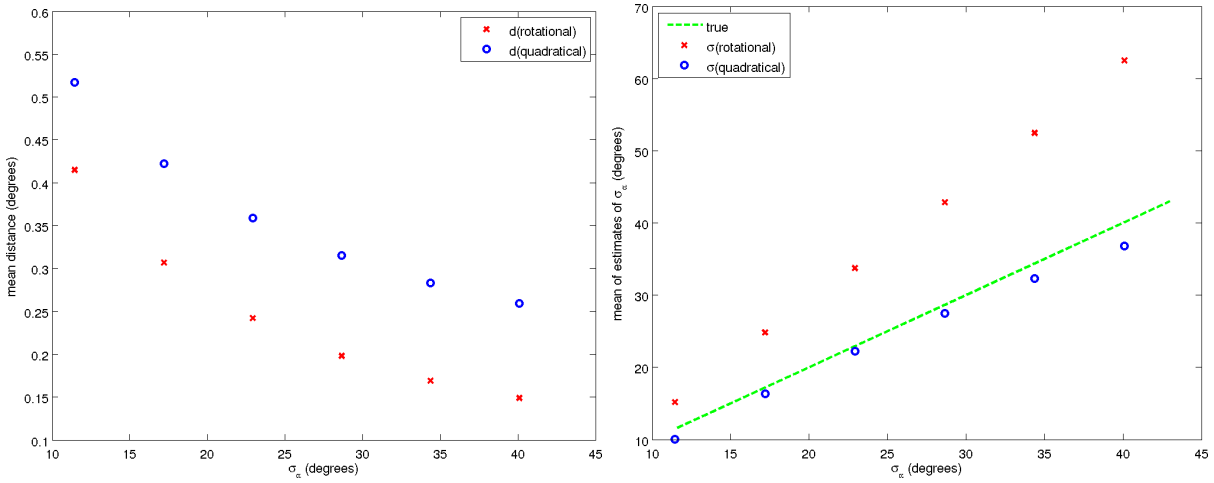
(unit: degrees)		Rotational bending		Quadratic bending	
κ	n	$d_g(\hat{c}, c)$	$\hat{\sigma}_\alpha(\sigma_\alpha = 22.918)$	$d_g(\hat{c}, c)$	$\hat{\sigma}_\alpha(\sigma_\alpha = 22.918)$
100	30	0.898 (0.492)	34.133 (4.429)	1.494 (0.871)	23.277 (3.389)
100	100	0.467 (0.261)	34.179 (2.592)	0.789 (0.470)	22.880 (2.454)
1000	30	0.242 (0.127)	33.739 (4.448)	0.359 (0.213)	22.203 (3.184)
1000	100	0.127 (0.069)	33.973 (2.581)	0.193 (0.112)	22.276 (1.891)

In particular, larger σ_α leads to a greater bias of $\hat{\sigma}_\alpha$, meanwhile it yields a better estimate of $\hat{c}$ (Fig. 1.6).

Hierarchical motion

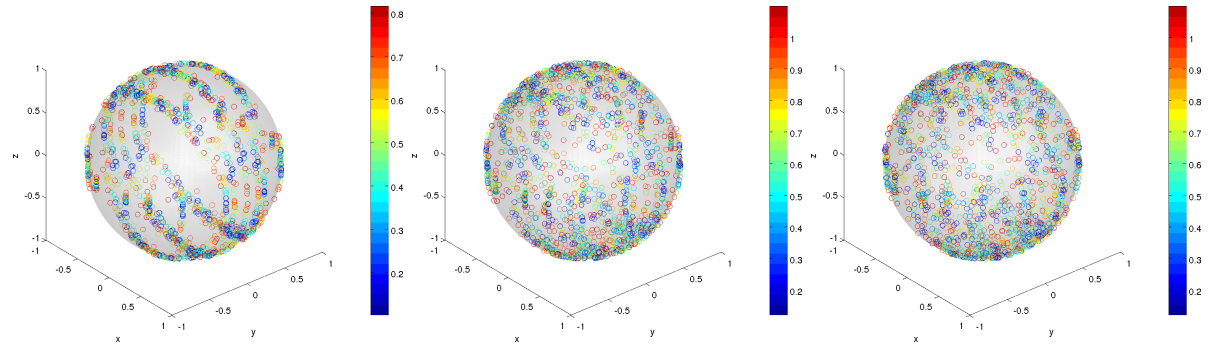
Figure 1.7 shows 30 samples of 64 different normal directions using different concentration parameters κ, twistings σ_θ and a fixed bending parameter $\sigma_\alpha = 0.4$. We have excluded 8 normal directions attached at $x = 0$. Different colors represent different amounts of absolute rotation, which are $|\alpha_j| + |\theta_j|$. In Fig. 1.7a we see rotations of normals along small circles around the y-axis. The clear motion pattern disappeared after an increase of σ_θ and κ in Figure 1.7b and Figure 1.7c.

Table 1.3 shows the performance of our estimator based on 1000 simulations under hierarchical rotational bending and twisting. As expected, the rotation axis estimates are less accurate than for single motions. Moreover, the estimate of the secondary rotation axis is less accurate than the estimate of the primary axis, particularly for $\kappa = 100$. The bias in the estimates of the rotation angle will be further discussed shortly.



(a) (b)

Figure 1.6: (a) Mean absolute error $d_g(\hat{c}, c)$ after quadratic and rotational bending for different σ_α values. (b) Mean of estimated bending parameter $\hat{\sigma}_\alpha$ after quadratic and rotational bending for different σ_α values. The green dashed line marks the first diagonal.



(a) $\epsilon_j = 0, \sigma_\theta = 0.1$ (b) $\epsilon_j = 0, \sigma_\theta = 0.3$ (c) $\kappa = 1000, \sigma_\theta = 0.3$

Figure 1.7: (Hierarchical deformation) Boundary normals using $\sigma_\alpha = 0.4$, different concentration parameter κ and twisting σ_θ.

Two initial value choices (from Principal Arc Analysis and random directions) are applied. In contrast to the s-rep analysis in the main article, the results in Table 1.3 are very similar for both choices.

In the four sets of analyses, we have shown accurate estimation results of the proposed method for the boundary normal data. In particular, the estimation from the quadratic bending model is surprisingly accurate despite the misspecified model. We believe that the proposed method will lead to robust estimates in other deformation models, which are similar to a rotational deformation.

Table 1.3: (Rotational bending and twisting) Mean absolute error for $\hat{c}_1$, $\hat{c}_2$, and the estimates $\hat{\sigma}_\alpha$, $\hat{\sigma}_\theta$.

(unit: degrees)		1st rotation axis		2nd rotation axis	
κ	n	$d_g(\hat{c}_1, c_1)$	$\hat{\sigma}_\alpha(\sigma_\alpha = 22.918)$	$d_g(\hat{c}_2, c_2)$	$\hat{\sigma}_\theta(\sigma_\theta = 17.189)$
initialization by 1st and 2nd principal component					
100	30	3.526 (2.775)	18.125 (2.370)	20.047 (16.717)	9.232 (1.598)
100	100	1.902 (1.444)	18.268 (1.337)	11.081 (13.477)	9.239 (1.143)
1000	30	2.683 (2.272)	17.785 (2.377)	8.570 (9.031)	9.119 (1.323)
1000	100	1.637 (1.126)	18.101 (1.349)	3.901 (2.459)	9.367 (0.691)
initialization by 1st principal component and a random direction					
100	30	3.496 (2.762)	18.125 (2.367)	19.133 (15.445)	9.295 (1.498)
100	100	1.866 (1.390)	18.260 (1.342)	8.944 (6.753)	9.390 (0.798)
1000	30	2.678 (2.272)	17.785 (2.377)	8.479 (8.702)	9.125 (1.299)
1000	100	1.635 (1.127)	18.102 (1.349)	3.892 (2.451)	9.367 (0.691)

1.2.1 Estimation of σ_α

A precise estimation of σ_α under the bending model is an interesting open problem. We have observed that the amount of swing is nonlinear, and conjecture that the change of the surface curvature in the object is responsible for the additional swing of the directions. Figure 1.8 exemplifies such a non-linear relationship.

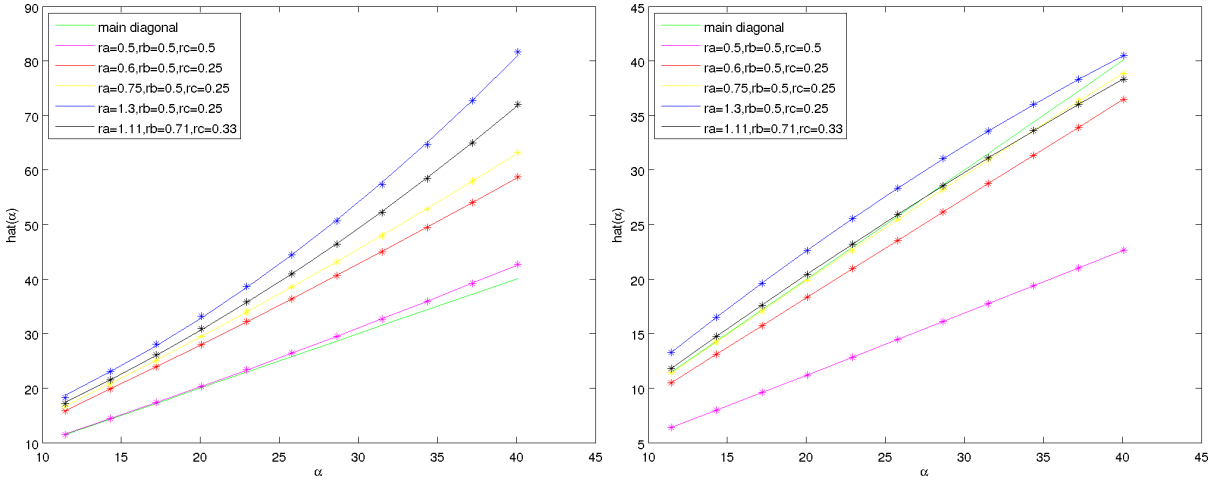
In the case of rotational bending, we may assume that our estimate $\hat{\sigma}_\alpha$ and the parameter σ_α are related through a quadratic function as Fig.1.8a suggests. If such a quadratic function $f(\sigma_\alpha) = p_0 + p_1\sigma_\alpha + p_2\sigma_\alpha^2 = \hat{\sigma}_\alpha$ is known or can be estimated efficiently by a least square quadratic polynomial, one can correct the estimate of σ_α for fixed r_a, r_b and r_c by

$$\hat{\sigma}_\alpha^{\text{new}} = -\frac{1}{2p_2} \pm \sqrt{\frac{1}{4p_2^2}(p_1^2 - 4p_0p_2 + 4p_2\hat{\sigma}_\alpha)},$$

with $f''(\hat{\sigma}_\alpha^{\text{new}}) \geq 0$. A similar modification can be made for quadratic bending (Fig.1.8b).

In general, we believe that modeling based on the curvatures will improve our current estimator. Such a modeling is beyond the scope of this paper, and we list a few important notions of curvature that have potential in modeling.

Most common curvature measures are the principal curvatures (κ_1, κ_2) with $\kappa_1 \geq \kappa_2$, Gaussian curvature, and mean curvature. These measures are calculated from the first and



(a) Rotational bending (b) Quadratic bending

Figure 1.8: Relationship between $\hat{\alpha}$ and α for various (r_a, r_b, r_c).

second fundamental form (see Gray (1998) and Kühnel (2006)). Koenderink (1990) indicated that Gaussian curvature and mean curvature are not representatives of local shapes because Gaussian curvature is identical for all local approximations for which the ratio of the principal curvatures (κ_1, κ_2) is equal. Therefore, he introduced the two alternative quantities: *shape index* S and *curvedness* C,

$$S = \frac{2}{\pi} \tan^{-1}\left(\frac{\kappa_2 + \kappa_1}{\kappa_2 - \kappa_1}\right), \qquad (\kappa_1 \geq \kappa_2) \tag{1.5}$$

$$C = \frac{2}{\pi} \ln\left(\sqrt{\frac{1}{2}(\kappa_1^2 + \kappa_2^2)}\right). \tag{1.6}$$

1.3 Quadratic bending of s-reps

The proposed method is also applied to the fitted s-reps of quadratically bent ellipsoids. After fitting s-reps to the raw images discussed in Section 1.2, we obtained the estimated axis of rotation of $\hat{c}_b = (-0.026, 0.999, 0.002)$ with a distance of 1.517° to the true y-axis $c_b = (0, 1, 0)$. Similar to mesh ellipsoid surfaces, the distance of the rotation axis estimate to the true axis is lower in case of rotational bending compared to quadratic bending.

Table 2.1: Numerical performance over 1000 replications, for hierarchical deformations.

Hierarchical rotations (unit: degrees)			$d_g(\hat{c}_1, c_1)$	$d_g(\hat{c}_2, c_2)$	$\hat{\sigma}_\theta(\sigma_\theta = 22.5)$	$\hat{\sigma}_\psi(\sigma_\psi = 15)$
Body 1	$\kappa = 100$	$n = 30$	4.527 (3.591)	4.929 (3.007)	22.693 (3.031)	15.890 (2.096)
		$n = 100$	2.201 (1.206)	2.944 (1.550)	22.698 (1.589)	15.844 (1.126)
	$\kappa = 1000$	$n = 30$	2.084 (1.364)	1.275 (0.701)	22.385 (2.984)	14.940 (1.941)
		$n = 100$	1.123 (0.741)	0.652 (0.338)	22.492 (1.542)	15.030 (1.110)
Body 2	$\kappa = 100$	$n = 30$	2.617 (1.762)	3.066 (3.099)	22.440 (2.959)	15.094 (2.011)
		$n = 100$	1.366 (0.746)	1.682 (0.870)	22.542 (1.562)	15.219 (1.073)
	$\kappa = 1000$	$n = 30$	1.099 (1.171)	0.921 (2.349)	22.339 (2.983)	14.872 (1.945)
		$n = 100$	0.568 (0.354)	0.438 (0.236)	22.470 (1.543)	14.981 (1.099)

2 Simulation results for the hierarchical deformation model

Table 2.1 summarizes the numerical performance of estimation of the hierarchical deformation model (11) as discussed in Section 4 of the main article. We have used the data-driven method to choose the initial values as described in Section 3 of the main article. The results are comparable to those from the simpler models in Section 4 and are fairly successful.

3 Discussion of model bias

A possibly important issue that is left unanswered is whether the parameters c_0 and $r = d_g(c_0, \mu)$ of the model

$$X = R(c_0, \theta)\mu \oplus \epsilon \tag{3.1}$$

are the minimizer $(\tilde{c}, \tilde{r})$ of the least squares problem

$$\min_{c,r} \mathbb{E}\rho^2\{\delta(c, r), X\} = \min_{c,r} \mathbb{E}\{d_g(c, X) - r\}^2. \tag{3.2}$$

Rivest (1999) has shown that when c_0 is known, the minimizer $\tilde{r} = \text{argmin}_r\, \mathbb{E}\{d_g(c_0, X) - r\}^2$ is biased towards $\pi/2$, i.e. $\tilde{r} > r = d_g(c_0, \mu)$ if $r < \pi/2$. In fact, given any c for the axis of rotation, $\tilde{r}_c = \mathbb{E}\{d_g(c, X)\}$ minimizes $\mathbb{E}\{d_g(c, X) - r\}^2$. Now to focus on c, $\min_{c,r} \mathbb{E}\{d_g(c, X) - r\}^2 = \min_c \mathbb{E}\{d_g(c, X) - \tilde{r}_c\}^2 = \min_c \text{Var}\{d_g(c, X)\}$. Therefore c_0 of (3.1) is the minimizer of (3.2) if for all $c \in S^2$,

$$\text{Var}\{d_g(c_0, X)\} \leq \text{Var}\{d_g(c, X)\}. \tag{3.3}$$

Table 3.1: Distance to true axis (measured in degrees)

	σ_θ				
r	0.01	0.1	0.2	0.5	1
$\pi/16 = 11.25$	11.25	11.19	10.99	9.10	2.29
$\pi/8 = 22.50$	22.50	22.36	21.89	2.55	0.20
$\pi/4 = 45.00$	44.98	44.81	42.73	0.22	0.00
$\pi/3 = 60.00$	59.97	59.79	3.72	0.16	0.00
$\pi/2 = 90.00$	90.00	90.00	0.49	0.02	0.02

This inequality may be satisfied when $r\sigma_\theta$ is large compared to the standard deviation of the error ϵ. If σ_θ or r is 0, there is no variation due to the rotation of $R(c_0, \theta)$, which makes the model unidentifiable. Heuristically, small σ_θ makes the estimation difficult. Likewise, the variation due to rotation is small if the rotation radius $r = d_g(c, \mu)$ is small. The standard deviation of the length $r\theta$ of the arc on $\delta(c, r)$ is $r\sigma_\theta$. Hence, the hypothesis:

(H) If (3.3) is not satisfied, the least-squares estimator may be biased.

While we have not succeeded in finding analytic forms of conditions that satisfy (3.3), a simulation study has been carried out to support our hypothesis (H). For $c_0 = e_3 = (0, 0, 1)'$, $\mu_r = \mu(r) = \cos(r)c_0 + \sin(r)e_1$, and $\theta \overset{\text{iid}}{\sim} N(0, \sigma_\theta^2)$, X is distributed as the von Mises–Fisher distribution with center $R(c, \theta)\mu_r$ and the concentration parameter $\kappa = 100$. We then evaluated the minimizer $\tilde{c}$ of $\text{Var}\{d_g(c, X)\}$ based on 5×10^5 Monte-Carlo random observations of X, for different combinations of $r \in \{\pi/16, \pi/8, \pi/4, \pi/3, \pi/2\}$ and $\sigma_\theta \in \{0.1, 0.2, 0.5, 1\}$ in radian. The result of the experiment is summarized as the distance between $\tilde{c}$ and c_0 in Table 3.1.

It can be checked from Table 3.1 that the distance between $\tilde{c}$ and c_0 is smaller for larger values of σ_θ and r. Moreover, for sufficiently large σ_θ and r, $d_g(\tilde{c}, c_0) = 0$, which leads to c_0 from the model (3.1) satisfying (3.3). On the other hand, when σ_θ and r are small with respect to the variance of the error, the rotation (3.1) becomes unidentifiable. This is further illustrated at Figure 3.1, which shows the scatter of 100 random observations from model (3.1), with $(r, \sigma_\theta) = (\pi/16, 0.2)$ on the left and $(\pi/4, 0.5)$ on the right panel. The left panel suggests that when (r, σ_θ) are small, the rotation about the axis c_0 is not distinguished and the optimal $\tilde{c}$ is near μ and $d_g(\tilde{c}, c_0)$ is large (cf. Table 3.1). The right panel illustrates that the rotation is visually identified for large values of (r, σ_θ), and for such a case, the parameter c_0 is the solution of the least squared problem (3.2).

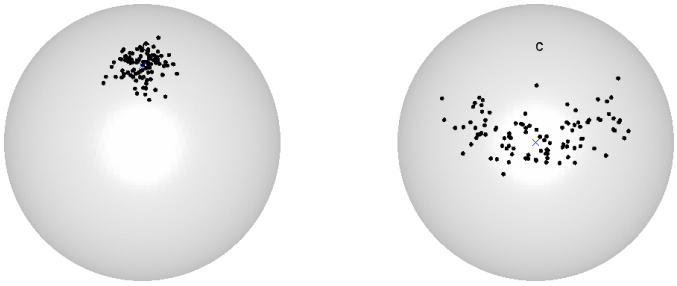


Figure 3.1: The scatter of 100 random observations from the model (3.1), with $(r, \sigma_\theta) = (\pi/16, 0.2)$ on the left and $(\pi/4, 0.5)$ on the right panel. The north pole, depicted as c in the figure, is the true axis of rotation c_0. The blue `x` visualizes μ.

4 Numerical studies with misspecified parameters

In this section, we study the impact of parameter misspecification of the estimator, particularly of the function f_j that model the relationships between the rotation angles θ_j. We study parameter misspecification over different rotational deformation situations as described in Section 4 in the main article.

Recall, Section 4 in the main article reports studies for two different objects. The first object (Body 1) consists of $K = 4$ directions, while the second object (Body 2) contains $K = 8$ directions. The von Mises–Fisher distribution is used for the distribution of errors. Three rotation models (indexed by equation number in the main article) are considered for each object:

- Model (2)–*Rigid rotation*: $c = (1, 0, 0)'$, $\theta_j = f_j(\theta) = \theta$, where $\theta \sim N(0, \sigma_\theta^2)$ and $\sigma_\theta = \pi/12 \approx 15°$.
- Model (10)–*Twisting*: $c = (0, 1, 0)'$, $\theta_j = f_j(\theta) = 1_{j\in I_1}\theta - 1_{j\in I_2}\theta$, where $\theta \sim N(0, \sigma_\theta)^2$, $\sigma_\theta = \pi/8 \approx 22.5°$. Here, I_1 and I_2 are the partitions of $\{1, \ldots, K\}$ according to the right and left sides whereas $I_1 = \{1, 2\}$ and $I_2 = \{3, 4\}$ for Body 1 and $I_1 = \{1, \ldots, 4\}$ and $I_2 = \{5, \ldots, 8\}$ for Body 2.
- Model (11)–*Hierarchical deformations*: $c_1 = (1, 0, 0)'$, $c_2 = (1/\sqrt{2}, -1/\sqrt{2}, 0)'$, $\theta_j = \theta$ and $\psi_j = f_j(\psi) = 1_{j\in I_1}\psi - 1_{j\in I_2}\psi$, where $\theta \sim N(0, \sigma_\theta)^2$, $\sigma_\theta \approx 22.5°$ and $\psi \sim N(0, \sigma_\psi)^2$, $\sigma_\psi \approx 15°$. The I_1 and I_2 are the same partition used in the twisting model above.

The hierarchical model represents deformations by a rigid rotation and oblique twist. For each model, we generate $n = 30, 100$ rotationally deformed objects with different error concentration parameters $\kappa = 100, 1000$. These are replicated 1000 times, and the estimation quality is measured by $d_g(\hat{c}, c)$ and $\hat{\sigma}_\theta = \sum_{i=1}^n \hat{\theta}_i^2/n$.

Table 4.1: Numerical performance over 1000 replications, for single deformation models.

Rotation (unit: degrees)		Body 1 $d_g(\hat{c}, c)$	Body 1 $\hat{\sigma}_\theta$ ($\sigma_\theta = 15$)	Body 2 $d_g(\hat{c}, c)$	Body 2 $\hat{\sigma}_\theta$ ($\sigma_\theta = 15$)
$\kappa = 100$	$n = 30$	4.133 (2.26)	3.314 (0.44)	2.905 (1.60)	6.771 (0.85)
	$n = 100$	2.235 (1.18)	3.308 (0.24)	1.560 (0.83)	6.816 (0.51)
$\kappa = 1000$	$n = 30$	1.166 (0.64)	1.037 (0.13)	0.841 (0.46)	6.698 (0.88)
	$n = 100$	0.656 (0.34)	1.039 (0.07)	0.448 (0.22)	6.744 (0.46)
Twisting (unit: degrees)		Body 1 $d_g(\hat{c}, c)$	Body 1 $\hat{\sigma}_\theta$ ($\sigma_\theta = 22.5$)	Body 2 $d_g(\hat{c}, c)$	Body 2 $\hat{\sigma}_\theta$ ($\sigma_\theta = 22.5$)
$\kappa = 100$	$n = 30$	2.761 (1.51)	3.669 (0.48)	4.062 (3.67)	17.150 (2.12)
	$n = 100$	1.482 (0.78)	3.658 (0.25)	1.889 (1.35)	17.055 (1.27)
$\kappa = 1000$	$n = 30$	0.803 (0.43)	1.139 (0.15)	1.017 (0.74)	16.760 (2.19)
	$n = 100$	0.446 (0.23)	1.147 (0.08)	0.536 (0.36)	16.857 (1.17)

The estimation results using correct parameters are reported for the single deformation models in Table 1 in the main article and for the hierarchical deformation model in Table 2.1 above.

4.1 Parameter f_j

The modelling of the function f_j can be challenging as discussed in Section 1.2.1 or for the s-rep model in case of bent, and bent and twisted ellipsoids in chapter 5 in the main article. Therefore, it is crucial to study the impact of misspecification of f_j.

Table 4.1 reports mean and standard deviations of the measures for the single deformation models. The true underlying rigid rotation deformation is indicated by $f_j(\theta) = \theta, j = 1, \ldots, 4$ for Body 1 and is misspecified by $f_j(\theta) = \theta, j = 1, 2$ and $f_j(\theta) = -\theta, j = 3, 4$. Body 2 is misspecified by $f_j(\theta) = (j/10)\theta$ instead of $f_j(\theta) = \theta$ for $j = 1, \ldots, 8$. In both cases the misspecification does not effect the estimated rotation axis $\hat{c}$ but leads to a poor prediction of $\hat{\sigma}_\theta$. The true underlying twisting motion of model (10) is given by $f_j(\theta) = \theta, j = 1, 2,\ f_j(\theta) = -\theta, j = 3, 4$ for Body 1 and is misspecified by $f_j(\theta) = \theta, j = 1, 4,\ f_j(\theta) = -\theta, j = 2, 3$. Body 2 is misspecified by $f_j(\theta) = \theta,\ f_l(\theta) = -0.5\theta$ whereas the true parameter is given by $f_j(\theta) = \theta,\ f_l(\theta) = -\theta$ for $j = 1, \ldots, 4$ and $k = 5, \ldots, 8$. The comparison of Table 4.1 to Table 1 in the main article shows that a misspecification of the parameter f_j does not effect the rotation axis estimation of a single deformation by fitting concentric small circles whereas the predicted rotation angle is biased by misspecification of f_j. The specification of f_j models the relationships between the rotation angles θ_j and is therefore crucial for their prediction.

On the other hand, the partition I_1 and I_2 is not explicitly used in the estimation proce-

Table 4.2: True and misspecified parameter f_j for Body 1 in the hierarchical deformation model (11).

	$f_{11}(\theta)$	$f_{12}(\theta)$	$f_{13}(\theta)$	$f_{14}(\theta)$	$f_{21}(\psi)$	$f_{22}(\psi)$	$f_{23}(\psi)$	$f_{24}(\psi)$
True	θ	θ	θ	θ	ψ	ψ	$-\psi$	$-\psi$
Table 4.4	θ	θ	$-\theta$	$-\theta$	ψ	ψ	ψ	ψ
Table 4.5	θ	$-\theta$	$-\theta$	θ	$-\psi$	ψ	ψ	$-\psi$

Table 4.3: True and misspecified parameter f_j for Body 2 in the hierarchical deformation model (11).

Hierarchical rotations		j							
		1	2	3	4	5	6	7	8
True	$f_{1j}(\theta)$	θ	θ	θ	θ	θ	θ	θ	θ
	$f_{2j}(\psi)$	ψ	ψ	ψ	ψ	$-\psi$	$-\psi$	$-\psi$	$-\psi$
Table 4.4	$f_{1j}(\theta)$	0.5θ	0.5θ	0.5θ	0.5θ	0.5θ	0.5θ	0.5θ	0.5θ
	$f_{2j}(\psi)$	ψ	ψ	ψ	ψ	$-\psi$	$-\psi$	$-\psi$	$-\psi$
Table 4.5	$f_{1j}\theta)$	0.1θ	0.2θ	0.3θ	0.4θ	θ	θ	θ	θ
	$f_{2j}(\psi)$	ψ	ψ	0.7ψ	0.7ψ	-0.2ψ	-0.2ψ	-0.3ψ	-0.3ψ

dure of the rotation axis. The partition I_1 and I_2 is implied by the function f_j. A partition I_1 and I_2 of indices $\{1, \ldots, K\}$ represents sets of direction vectors that rotate together, i.e., the partition models the deformation type. In the simulated examples, the partitions are chosen to model the bending and twisting deformation accordingly. The specification of f_j gives the possibility to incorporate additional prior knowledge about the statistical model of the rotation angle θ_j, e.g., linear or quadratic relation by the distance of the direction to the rotation axis. Nevertheless, the modelling of the function f_j can be challenging as discussed before. There are real data applications where the definition of a partition is naturally motivated, e.g., by the physical structure of the body. An example is to estimate the rotational deformation between the upper and lower leg as studied in Section 6 in the main article. This example can be extended to all joints inside the human body and to all objects which are connected by a joint. On the other hand, there is a group of data sets where such a partition is not obvious, e.g., in the s-rep model of the hippocampus. A first approach could be to define the partition on the basis of an observation of a medical expert. An automatic detection of partitions and clusters is an interesting future research direction.

In contrast to the single deformation model, a misspecification of f_j has an impact in the hierarchical deformation model by the iterative back-and-forward deformations of the random direction vectors. Therefore, a misspecification of the parameter might guide the iterative estimation procedure to fall in a local minimum, and leads to a poor estimation of the rotation axes and angles. Table 4.2 and Table 4.3 report the true and misspecified f_j for the simulation studies whereas the estimation results are summarized in Table 4.4

Table 4.4: Numerical performance over 1000 replications, for hierarchical deformations using moderate misspecified parameters.

Hierarchical rotations (unit: degrees)			$d_g(\hat{c}_1, c_1)$	$d_g(\hat{c}_2, c_2)$	$\hat{\sigma}_\theta(\sigma_\theta = 22.5)$	$\hat{\sigma}_\psi(\sigma_\psi = 15)$
Body 1	$\kappa = 100$	$n = 30$	45.223 (3.61)	44.792 (3.69)	15.013 (2.00)	22.725 (3.01)
		$n = 100$	45.188 (1.73)	44.423 (1.79)	14.992 (1.06)	22.731 (1.58)
	$\kappa = 1000$	$n = 30$	43.648 (1.29)	44.257 (2.67)	14.454 (1.89)	22.442 (2.99)
		$n = 100$	43.705 (0.72)	43.981 (1.42)	14.482 (1.06)	22.510 (1.54)
Body 2	$\kappa = 100$	$n = 30$	2.617 (1.76)	3.066 (3.09)	22.440 (2.96)	15.094 (2.01)
		$n = 100$	1.100 (1.17)	0.921 (2.35)	22.339 (2.98)	14.872 (1.94)
	$\kappa = 1000$	$n = 30$	1.366 (0.74)	1.683 (0.87)	22.542 (1.56)	15.219 (1.07)
		$n = 100$	0.569 (0.35)	0.438 (0.23)	22.470 (1.54)	14.981 (1.09)

Table 4.5: Numerical performance over 1000 replications, for hierarchical deformations using drastically misspecified parameters..

Hierarchical rotations (unit: degrees)			$d_g(\hat{c}_1, c_1)$	$d_g(\hat{c}_2, c_2)$	$\hat{\sigma}_\theta(\sigma_\theta = 22.5)$	$\hat{\sigma}_\psi(\sigma_\psi = 15)$
Body 1	$\kappa = 100$	$n = 30$	11.576 (4.66)	33.737 (4.94)	5.423 (2.58)	2.619 (3.45)
		$n = 100$	11.272 (2.38)	33.372 (2.48)	4.749 (1.02)	0.966 (1.44)
	$\kappa = 1000$	$n = 30$	11.228 (3.91)	33.586 (4.06)	3.597 (1.98)	1.186 (2.34)
		$n = 100$	11.183 (2.19)	33.445 (2.25)	3.295 (0.32)	0.413 (0.26)
Body 2	$\kappa = 100$	$n = 30$	12.204 (5.01)	33.456 (5.11)	15.917 (1.88)	214.219 (20.16)
		$n = 100$	11.337 (2.46)	33.959 (2.47)	16.180 (1.15)	214.553 (10.75)
	$\kappa = 1000$	$n = 30$	11.838 (4.40)	33.787 (4.46)	15.727 (2.01)	219.414 (20.99)
		$n = 100$	11.289 (2.30)	33.978 (2.34)	16.316 (1.20)	219.587 (10.45)

and Table 4.5. At first we have changed the order of the original parameters f_1, f_2 for Body 1 in Table 4.4. The first estimated rotation axis $\hat{c}_1$ is around $(1/\sqrt{2}, -1/\sqrt{2}, 0)'$ and the second estimated rotation axis $\hat{c}_2$ is around $(1, 0, 0)'$, i.e., the estimator has interchanged the true underlying deformations which results in a distance of approximately 45 degree of $\hat{c}_1$ to c_1 and $\hat{c}_2$ to c_2. Nevertheless, the order of generalized rotational deformations are not interchangeable in general, and a misspecification might bias the results. In a second set, we have misspecified f_{1j} globally by a factor of 0.5 and kept the original f_{2j} for Body 2. The means and standard deviations in Table 4.4 show only small changes compared to Table 2.1 and are very accurate. A global scale change does not change the relationships between the rotation angles. Both cases demonstrate the performance of the hierarchical estimation procedure in case of a moderate misspecification by reasonable estimates.

In addition, two cases with drastically misspecified parameters are reported. In a third scenario, f_1 and f_2 are misspecified so as to generate different deformations for Body 1 in Table 4.5. Both the means and the standard deviations show poor estimation results. In a last set we modified f_1 and f_2 by keeping the underlying deformation direction of each

Table 4.6: Numerical performance over 1000 replications, for hierarchical deformations with misspecified order of primary and secondary rotation axis.

Hierarchical rotations (unit: degrees)			$d_g(\hat{c}_1, c_1)$	$d_g(\hat{c}_2, c_2)$	$\hat{\sigma}_\theta(\sigma_\theta = 22.5)$	$\hat{\sigma}_\psi(\sigma_\psi = 15)$
Body 1	$\kappa = 100$	$n = 30$	4.624 (2.66)	4.876 (2.85)	22.720 (3.00)	15.896 (2.10)
		$n = 100$	2.258 (1.30)	2.949 (1.54)	22.701 (1.58)	15.847 (1.12)
	$\kappa = 1000$	$n = 30$	2.044 (1.27)	1.279 (0.70)	22.382 (2.98)	14.947 (1.94)
		$n = 100$	1.124 (0.72)	0.656 (0.33)	22.492 (1.54)	15.033 (1.11)
Body 2	$\kappa = 100$	$n = 30$	2.590 (1.47)	2.992 (1.75)	22.439 (2.96)	15.094 (2.01)
		$n = 100$	1.323 (0.72)	1.688 (0.87)	22.541 (1.56)	15.220 (1.07)
	$\kappa = 1000$	$n = 30$	1.063 (0.67)	0.849 (0.47)	22.336 (2.98)	14.873 (1.94)
		$n = 100$	0.567 (0.35)	0.438 (0.23)	22.470 (1.54)	14.982 (1.09)

direction vector for Body 2 but changing the amount of deformation locally. Also in this scenario the hierarchical estimator shows poor estimation results in Table 4.5.

4.2 Primary and secondary rotation

The hierarchical model is a first attempt to model and estimate more then one rotational deformation. The order of two rotations is not interchangeable and is considered as a part of the statistical model which attempts to describe the nature as well as possible. Therefore, we assume a primary rotation $R(c_1, \theta_j)$ and a secondary rotation $R(c_2, \theta_j)$ in our hierarchical model. The order of rotation can be misspecified in two different ways in the proposed approach. At first, we might interchange f_1 and f_2 as discussed in Section 4.1 above for Body 1 in Table 4.4. In addition, the primary and secondary rotation has to be specified for the initialization.

Table 4.6 summarizes the results in estimation of the hierarchical deformation model (11) with interchanged primary and secondary rotation for the initialization. The results are similar to Table 2.1. The estimator converges to the same results in this example. Nevertheless, a misspecification might increase the risk that the estimation procedure converges in a different local minimum.

4.3 Discussion

The introduction of f_j in our model has advantages and disadvantages. We study generalized rotational deformations on the basis of directional data, and the rotation of a direction from one to another place on the sphere is not uniquely defined in the hierarchical case. Therefore, different functions f_j can describe different rotational deformations. Prior knowledge is nec-

essary in order to restrict the rotational directions to avoid the convergence of the optimizer into a different local minimum and to avoid an overfit of c_1 and c_2 as mentioned Section 4 in the main article.

The development of a method to predict the function f_j from a given data set is an interesting open research question. Furthermore, an automatic classification of spoke directions into a partition I_1 and I_2 and a final classification of the deformation type into rigid rotation, bending or twisting are of future interest.

5 Computational complexity of the algorithm

The computational complexity of the proposed estimation procedure is now reported in two forms: the asymptotic time complexity and finite sample time measurements.

The asymptotic time complexity of the proposed algorithm depends on the number of samples n and the number of direction vectors K. In particular the optimization problem,

$$(\hat{c}, \hat{\mathbf{r}}) = \underset{c, r_1, \ldots, r_K}{\operatorname{argmin}} \sum_{i=1}^{n} \sum_{j=1}^{K} d_g^2\{\delta(c, r_j), X_{ij}\},$$

is the dominant part of the algorithm. Our algorithm to solve this nonlinear least squares problem is doubly iterative. The outer loop consists of applications of the exponential and inverse exponential maps whose time complexity is $O(nK)$. The inner loop iteratively updates $c^\dagger \in \mathbb{R}^3$ and $r_j^\dagger \in (0, \pi/2)$ by the Levenberg–Marquardt algorithm, where each iteration requires the asymptotic time complexity of $O(nK)$ mainly due to the computation of Jacobian matrix. Notice, that the computation time for inverting a 3×3 matrix is dominated by other terms and is ignored. Overall, if a finite number of iterations is assumed, then the asymptotic time complexity of the proposed algorithm is $O(nK)$.

We now turn to our attention to the complexity of the algorithm in real time, which we believe is more useful for practitioners. Computation times were measured by a personal computer with a Intel(R) Xeon(TM) 3.73GHz processor.

We have tested the estimations of three different rotational deformations

- *Model (2) - Rigid rotation*,
- *Model (10) - Twisting* and
- *Model (11) - Hierarchical deformations*

Table 5.1: Median time measurements in seconds over 100 replications.

n	Model	K					
		4	8	16	32	64	128
30	(2)	0.04	0.07	0.13	0.26	0.61	1.81
	(10)	0.11	0.16	0.26	0.40	0.79	2.00
	(11)	0.54	0.84	1.61	3.17	6.91	19.23
60	(2)	0.04	0.07	0.13	0.28	0.68	2.03
	(10)	0.10	0.18	0.29	0.42	0.85	2.18
	(11)	0.75	1.39	2.72	6.55	13.84	30.80
120	(2)	0.04	0.07	0.13	0.29	0.76	2.34
	(10)	0.12	0.18	0.27	0.44	0.92	2.50
	(11)	1.29	2.50	4.90	10.03	20.37	53.91

Table 5.2: Minimal time measurements in seconds over 100 replications..

n	Model	K					
		4	8	16	32	64	128
30	(2)	0.02	0.05	0.08	0.19	0.48	1.59
	(10)	0.03	0.07	0.12	0.25	0.61	1.75
	(11)	0.23	0.40	0.74	1.50	3.21	7.71
60	(2)	0.02	0.04	0.08	0.18	0.54	1.78
	(10)	0.03	0.05	0.14	0.27	0.68	1.89
	(11)	0.37	0.68	1.32	2.61	5.52	12.37
120	(2)	0.02	0.04	0.09	0.22	0.63	2.12
	(10)	0.03	0.09	0.15	0.33	0.79	2.38
	(11)	0.64	1.21	2.39	4.85	9.99	21.61

as described in Section 4 (the numbers correspond to the equation number in the main article), with two different types of objects to be deformed.

We first examined the computation times for a set of well-controlled objects. Using the Body 1 (as plotted in Fig. 2 of the main article) consisting of the original $K = 4$ directions, we duplicated the same direction vectors to increase the number of directions ($K = 8, 16, 32, 64, 128$) so that the effect on computation time of the different locations of direction vectors is minimized. For each sample size $n = 30, 60, 120$, we have repeated the estimation $R = 100$ times to measure the computation time required to obtain the estimates $\hat{c}$. Note that we have used von Mises-Fisher distribution with $\kappa = 1000$ for the i.i.d. errors.

Tables 5.1 and 5.2 report the median computation time and the minimal computation time among the R repetitions, respectively. With large numbers of sample and directions, the computation requires less than a minute on average, while it takes less than a second for moderate sizes of sample and directions.

From an inspection of Table 5.1, there is a trend for the computation time increasing

Table 5.3: Quality of the repeated rotation axis estimations for the time measurements by $d_g(\hat{c}, c)$ in degree.

n	Model	rotation axis	K					
			4	8	16	32	64	128
30	(2)	1	1.18	0.92	0.63	0.40	0.33	0.22
	(10)	1	0.88	0.61	0.40	0.29	0.20	0.14
	(11)	1	2.18	2.10	1.99	1.89	1.86	1.81
	(11)	2	1.43	1.09	0.77	0.60	0.48	0.43
60	(2)	1	0.74	0.63	0.42	0.30	0.21	0.16
	(10)	1	0.59	0.42	0.29	0.18	0.14	0.10
	(11)	1	1.42	1.28	1.24	1.20	1.16	1.17
	(11)	2	0.88	0.64	0.49	0.35	0.28	0.21
120	(2)	1	0.60	0.39	0.30	0.20	0.15	0.10
	(10)	1	0.37	0.30	0.20	0.14	0.10	0.07
	(11)	1	1.09	0.97	0.99	0.90	0.89	0.89
	(11)	2	0.61	0.44	0.29	0.22	0.18	0.12

approximately linear with K and also with n for all models.

By comparing the minimal time (Table 5.2) and the median time (Table 5.1), we have noticed that the computation time varies by a large amount. See for example model (11) with $n = 60$, $K = 128$; The median time is over 30 seconds while the shortest time is only 12 seconds. This is due to the slow convergence of the iterative algorithm for a particular choice of observations.

Finally, Table 5.3 reports the quality of the repeated rotation axis estimations by the mean distance of $d_g(\hat{c}, c)$, which shows higher accuracy for larger K and n as discussed in Section 4 of the article.

We also have examined the computation times with another set of objects whose base direction vectors are determined randomly. As shown in the following, the additional randomness leads to more variation in the computation times. Nevertheless, the computation time exhibits again the approximate linear increase for K and n.

The second set of objects to be deformed has $K = 8, 16, 32, 64, 128$ directions vectors, each of which is obtained from uniform distribution on a hemisphere. With $n = 30, 60, 120$ samples, we also report the time measurements from $R = 100$ replications.

Table 5.4 and 5.5 report the median and the minimum computation time in seconds, respectively. As expected, the time increases with larger K and larger n. Due to the uncontrolled model complexity (originated from the random base directions) the time difference between the median and the minimum time is larger than the previous controlled case. We conjecture that the computation times are heavily dependent on the convergence of the

Table 5.4: Median time measurements in seconds over 100 replications.

n	Model	K				
		8	16	32	64	128
30	(2)	0.08	0.14	0.26	0.63	1.83
	(10)	0.18	0.29	0.49	0.94	2.10
	(11)	1.99	6.39	17.18	51.44	130.85
60	(2)	0.08	0.14	0.28	0.70	2.04
	(10)	0.22	0.32	0.51	0.96	2.27
	(11)	2.74	6.75	24.25	63.20	265.95
120	(2)	0.08	0.14	0.30	0.77	2.40
	(10)	0.22	0.32	0.48	0.98	2.54
	(11)	9.79	18.86	42.81	129.75	305.54

Table 5.5: Minimal time measurements in seconds over 100 replications..

n	Model	K				
		8	16	32	64	128
30	(2)	0.05	0.08	0.20	0.58	1.78
	(10)	0.07	0.12	0.25	0.59	1.84
	(11)	0.42	0.84	2.30	3.45	8.03
60	(2)	0.05	0.10	0.23	0.59	1.80
	(10)	0.07	0.16	0.31	0.72	1.99
	(11)	0.72	1.36	2.75	5.99	18.78
120	(2)	0.04	0.10	0.24	0.63	2.23
	(10)	0.06	0.17	0.32	0.71	2.31
	(11)	1.25	2.43	4.91	15.24	33.16

Table 5.6: Quality of the repeated rotation axis estimations for the time measurements by $d_g(\hat{c}, c)$ in degree.

n	Model	rotation axis	K				
			8	16	32	64	128
30	(2)	1	0.92	0.64	0.45	0.33	0.21
	(10)	1	0.67	0.44	0.30	0.22	0.15
	(11)	1	1.20	0.59	0.40	0.23	0.16
	(11)	2	2.57	0.73	0.55	0.37	0.25
60	(2)	1	0.64	0.53	0.31	0.23	0.15
	(10)	1	0.43	0.26	0.22	0.14	0.10
	(11)	1	1.05	0.37	0.25	0.18	0.11
	(11)	2	2.67	0.60	0.35	0.23	0.17
60	(2)	1	0.42	0.33	0.23	0.14	0.11
	(10)	1	0.31	0.21	0.15	0.10	0.07
	(11)	1	0.86	0.27	0.17	0.11	0.07
	(11)	2	2.39	0.37	0.25	0.17	0.12

Levenberg–Marquardt algorithm.

In addition to the computation time, Table 5.6 reports the quality of the repeated rotation axis estimations by the mean distance of $d_g(\hat{c}, c)$, which again shows higher accuracy for larger K and n as discussed in Section 4 of the article.

6 Estimation procedure for knee motion during gait

In section 6 in the main article we have studied a real data example: the deformation of the upper and lower leg by two potentially non-orthogonal rotations at the knee joint during gait. These two rotations are flexion-extension about a right-to-left (medial-lateral) axis and the internal-external rotation of the lower-leg relative to the upper leg about an axis directed along the long axis of the lower leg. The data set consists of T time dependent observations $\mathbf{M}_1, \ldots, \mathbf{M}_T$ whereas each $\mathbf{M}_i$ is a collection of markers $\mathbf{M}_i = (M_{i1}, \ldots, M_{i\widetilde{K}})$ on the upper and lower leg with $M_{ij} \in \mathbb{R}^3, j = 1, \ldots, \widetilde{K}$. Let $\widetilde{I}_1$ and $\widetilde{I}_2$ be a partition of indices $\{1, \ldots, \widetilde{K}\}$ representing groups of $\widetilde{K}_1$ markers $\widetilde{I}_1$ on the upper leg and $\widetilde{K}_2$ markers $\widetilde{I}_2$ on the lower leg. Two markers $\mathbf{M}_{\iota_1}, \iota_1 \in \widetilde{I}_1$ and $\mathbf{M}_{\iota_2}, \iota_2 \in \widetilde{I}_2$ are chosen as basis points at the upper and lower leg. Set $I_1 = \widetilde{I}_1 \setminus \{\iota_1\}$, $I_2 = \widetilde{I}_2 \setminus \{\iota_2\}$, $K = \widetilde{K} - 2$, $K_1 = \widetilde{K}_1 - 1$ and $K_2 = \widetilde{K}_2 - 1$ then directions $\mathbf{X}_i = (X_{i1}, \ldots, X_{iK})$ are derived by

$$X_{ij} = \frac{M_{ij} - M_{i\iota_1}}{\|M_{ij} - M_{i\iota_1}\|}, j \in I_1, \quad X_{ij} = \frac{M_{ij} - M_{i\iota_2}}{\|M_{ij} - M_{i\iota_2}\|}, j \in I_2$$

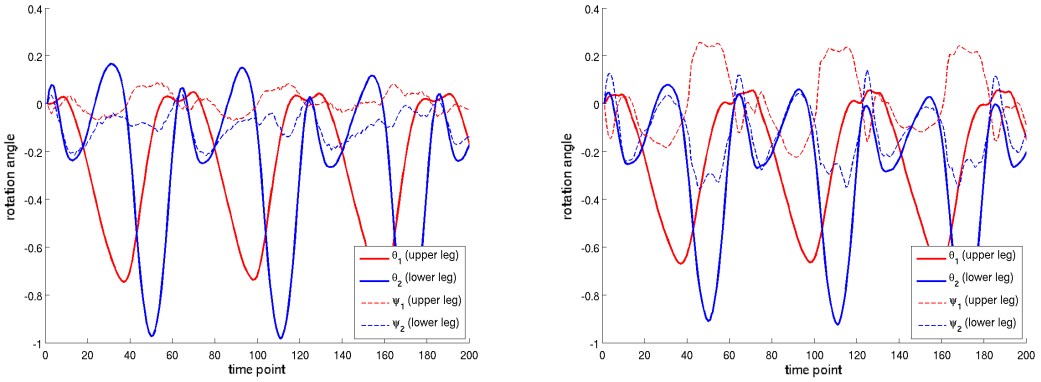
for $i = 1, \ldots, T$, which are invariant to translation and size changes.

For the sake of convenience, we assume the observations $\mathbf{X}_1, \ldots, \mathbf{X}_T$ are independent and modify the single rotation model (10) in the main article to

$$X_{ij} = R(c, \theta_{ij}) X_{1j} \oplus \epsilon_{ij} \quad (j = 1, \ldots, K). \tag{6.1}$$

A more careful modelling of the time dependencies is left for future work, e.g., by an autoregressive model as suggested by (Rivest, 2001, Section 4.1).

The first (flexion-extension) rotation axis c_1 is estimated by the estimation procedure (7) in the main article and describes a bending deformation of the upper and lower leg around the knee. The lower leg rotates relative to the upper leg whereas the upper leg rotates relative to the pelvis. At first, we estimate the rotation angles θ_i^u of the upper leg in order to estimate the rotation angles θ_i^l of the lower leg relative to the upper leg. The least squares



(a) bone (b) skin

Figure 6.1: Predicted rotation angles $(\theta_1, \psi_1) = (\hat{\theta}^u, \hat{\psi}^u)$ of directions on the upper leg and angles $(\theta_2, \psi_2) = (\hat{\theta}^l, \hat{\psi}^l)$ of directions on the lower leg for the first 200 time points. (a) Estimated rotation angles of directions derived from bone markers. (b) Estimated rotation angles of directions derived from skin markers. (angle units in radian)

estimators $(\hat{c}_1, \hat{\mathbf{r}}_1)$ can be used to estimate the parameters of (6.1) with

$$\theta_{ij} = \text{atan2}\{\langle P_{(\hat{c}_1,\hat{r}_{1j})} X_{ij}, \hat{c}_1 \times X_{1j} \rangle, \langle P_{(\hat{c}_1,\hat{r}_{1j})} X_{ij}, X_{1j} - \hat{c}_1 \cos(\hat{r}_j) \rangle\}. \tag{6.2}$$

The predictor of $\theta_i^u, i = 1, \ldots, T$ is

$$\hat{\theta}_i^u = \frac{1}{K_1} \sum_{j=1}^{K_1} \theta_{ij}, \quad j \in I_1.$$

Next, the directions are de-rotated by $\widetilde{X}_{ij} = R(\hat{c}_1, -\hat{\theta}_i^u) X_{ij}$ for $j = 1 \ldots, K$ and $i = 1, \ldots, T$. The directions $\widetilde{X}_{ij}, j \in I_1$ are directions on the upper leg and do not rotate about $\hat{c}_1$ after the inverse deformation. The remaining rotation of the lower leg relative to the upper leg is then

$$\hat{\theta}_i^l = \frac{1}{K_2} \sum_{j=1}^{K_2} \tilde{\theta}_{ij}, \quad j \in I_2 \text{ with}$$

$$\tilde{\theta}_{ij} = \text{atan2}\{\langle P_{(\hat{c}_1,\hat{r}_{1j})} \widetilde{X}_{ij}, \hat{c}_1 \times \widetilde{X}_{1j} \rangle, \langle P_{(\hat{c}_1,\hat{r}_{1j})} \widetilde{X}_{ij}, \widetilde{X}_{1j} - \hat{c}_1 \cos(\hat{r}_j) \rangle\}.$$

Finally, we obtain a set of de-rotated directions $\mathbf{Z}$ by $Z_{ij} = \widetilde{X}_{ij}, \ j \in I_1$ and $Z_{ij} = R(\hat{c}_1, -\hat{\theta}_i^l) \widetilde{X}_{ij}, \ j \in I_2$ for $i = 1, \ldots, T$.

The estimation of the second (internal-external) rotation axis $(\hat{c}_2, \hat{\mathbf{r}}_2)$ and $\hat{\psi}^u, \hat{\psi}^l$ follows the same steps of the above paragraph using using $\mathbf{Z}_i$ instead of $\mathbf{X}_i$ for $i = 1, \ldots, T$.

In addition to the estimates $\hat{c}_1$ and $\hat{c}_2$, we estimate the rotation axes $\hat{c}_{1j}$ and $\hat{c}_{2j}$ for each

marker $j = 1 \ldots, K$. Therewith, we can quantify the estimations by the dispersion $\sigma_{d_g(\hat{c}_1, \hat{c}_{1j})}$ of the geodesic distance $d_g(\hat{c}_1, \hat{c}_{1j})$ and $\sigma_{d_g(\hat{c}_2, \hat{c}_{2j})}$ respectively.

Figure 6.1 reports the predicted rotation angles $(\hat{\theta}_i^u, \hat{\psi}_i^u)$ of the upper leg relative to the laboratory reference frame and $(\hat{\theta}_i^l, \hat{\psi}_i^l)$ of the lower leg relative to the upper leg for the first 200 time points. The larger observed rotation angles around the second rotation axis for the skin data is due to the well-known deformation of the skin surface.

References

Gray, A. (1998), *Modern Differential Geometry of Curves and Surfaces with Mathematica*, CRC Press, 2nd ed.

Koenderink, J. J. (1990), *Solid Shape*, The MIT Press.

Kühnel, W. (2006), *Differential Geometry*, vol. 16, Student Mathematical Library, 2nd ed.

Mardia, K. V. and Jupp, P. E. (2000), *Directional Statistics*, Chichester: Wiley.

Rivest, L.-P. (1999), "Some Linear Model Techniques for Analyzing Small-Circle Spherical Data," *Canadian Journal of Statistics*, 27, 623–638.

— (2001), "A Directional Model for the Statistical Analysis of Movement in Three Dimensions," *Biometrika*, 88, 779–791.