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

March 26, 2014

Abstract

This paper discusses a novel framework to analyze 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 modeling 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 major forms of deformation. For instance, the major variation of shapes of hippocampi
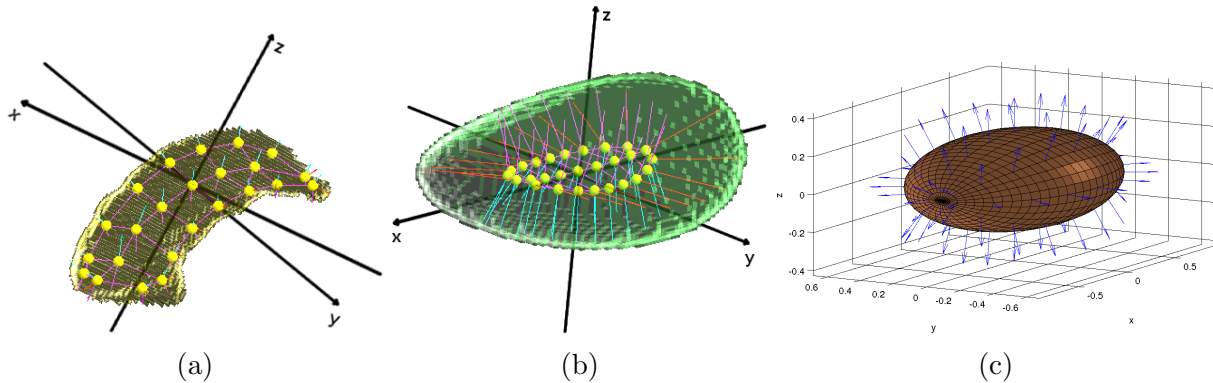
Figure 1: 3D objects and their models. (a) S-rep of a hippocampus (b) S-rep of a rotationally deformed ellipsoid. (c) Attached boundary normals on the meshed surface of an ellipsoid.

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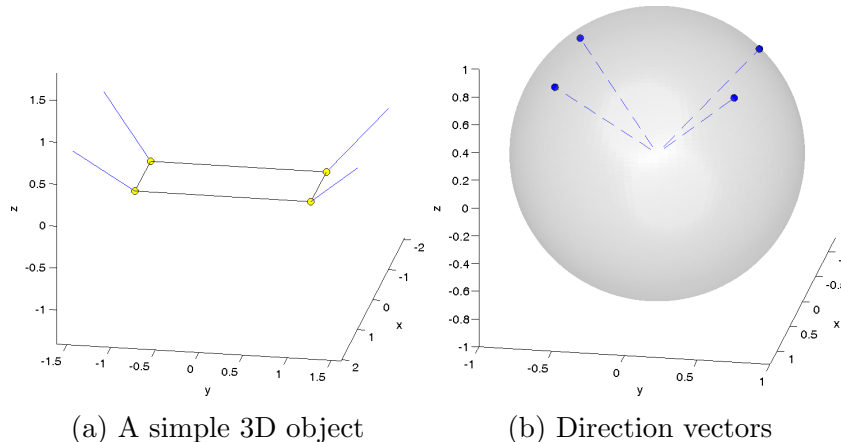
(a) A simple 3D object   (b) Direction vectors

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.

In all three cases, the direction vectors are predetermined by the shape models.

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,

3

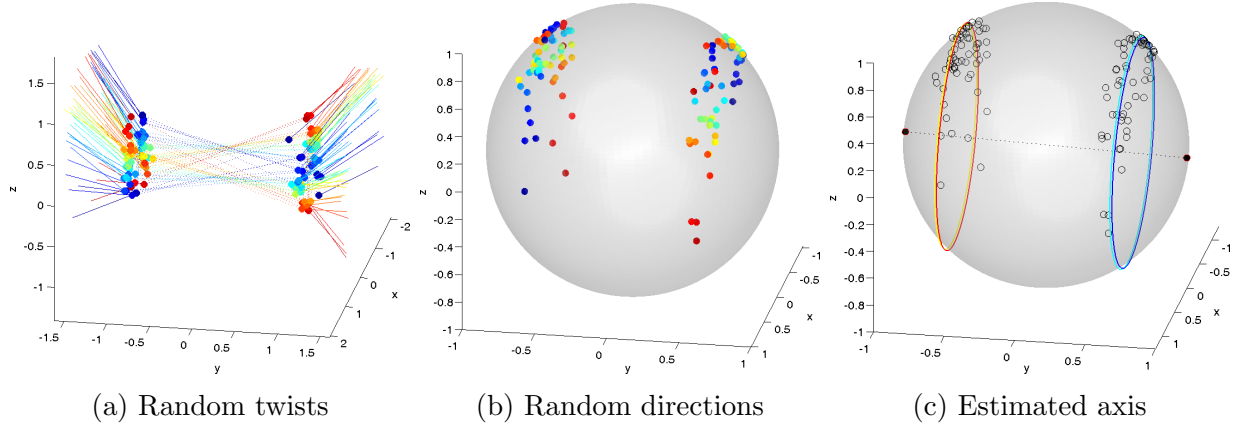| (a) Random twists | (b) Random directions | (c) Estimated axis |

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.

2001; Gamage and Lasenby, 2002; Teu and Kim, 2006; Rivest et al., 2008; Ball and Greiner, 2012). In the statistical literature, estimation of rotation matrices has been studied in terms of spherical regression and its generalizations (Chang, 1986, 1988, 1989; Chapman et al., 1995; Rivest, 1989, 1998, 2006; Chang and Rivest, 2001). In spherical regression coupled observations (x_j, y_j) are assumed where both the axis c and the angle θ are estimated. In contrast, our model is comparable with a 1-factor model and 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.

4

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 $\delta(c, r) = \{x \in S^2 : x'c = \cos(r)\} \subset S^2$, which is a set of equidistant points from c. We call c the center and $r = \arccos(x'c)$ the radius of the circle. 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,

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

often called the geodesic or Fréchet mean (Fréchet, 1948; Karcher, 1977; Huckemann, 2012). A useful measure of dispersion is the geodesic variance which is defined as $\operatorname{Var}(X) = \mathbb{E}\{d_g^2(\mu(X), X)\} = \min_{x \in S^2} \mathbb{E}\{d_g^2(x, X)\}$ 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. 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

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

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

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

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 $\text{Var}(r_j\theta_j)$ is large enough compared to the error variance $\text{Var}(\epsilon_j)$ 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

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

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

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

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

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

where the two argument function $\operatorname{atan2}(x_2, x_1) \in (-\pi, \pi]$ 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:

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

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

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

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

Note that (8) and (9) lead to $\sum_{i=1}^{n} \theta_{ij} = 0$ $(j = 1, \ldots, K)$, which is an empirical statement of (3).

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 $\theta_{j_1} = -\theta_{j_2}$ for all $j_1 \in I_1, j_2 \in I_2$. Another example is the scenario of independent rotations where all directions in the same group rotate together $(\theta_{j_1} = \theta_{j_2}, j_1, j_2 \in I_l)$ but two angles in different groups are independent $(\theta_{j_1}$ and θ_{j_2} are independent for $j_1 \in I_l, j_2 \in I_k, 1 \le l \ne k \le 2)$.

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., $\theta_j = f_j(\theta)$. Due to the identifiability issue, which leads to the assumption (3), we assume $\mathbb{E}(\theta_j) = 0$.

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 $\sigma \approx 22.5°$ 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, $\hat{\sigma} = \sum_{i=1}^{n} \hat{\theta}_i^2 / n = 21.5°$.

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

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

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

$$(\hat{c}_1, \hat{r}_{11}, \ldots, \hat{r}_{1K}) \quad = \quad \underset{c_1, (r_{11}, \ldots, r_{1K})}{\operatorname{argmin}} \sum_{i=1}^{n} \sum_{j=1}^{K} d_g^2 \{\delta(c_1, r_{1j}), X_{ij}^*\}, \tag{14}$$

$$\hat{\mu}_j \quad = \quad \underset{\mu}{\operatorname{argmin}} \sum_{i=1}^{n} d_{\delta(\hat{c}_1, \hat{r}_{1j})}^2 (P_{(\hat{c}_1, \hat{r}_{1j})}X_{ij}^*, \mu).$$

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

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

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

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

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 $\text{Var}(\theta)$ of the primary rotation is assumed to be larger than $\text{Var}(\psi)$ 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 $\text{Var}(\theta) < \text{Var}(\psi)$, 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 affects 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 vMF(κ), 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 $\theta \sim N(0, \sigma_\theta^2)$,

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

Rotation (unit: degrees)		Body 1		Body 2	
		$d_g(\hat{c}, c)$	$\hat{\sigma}_\theta$ ($\sigma_\theta = 15$)	$d_g(\hat{c}, c)$	$\hat{\sigma}_\theta$ ($\sigma_\theta = 15$)
$\kappa = 100$	$n = 30$	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		Body 2	
		$d_g(\hat{c}, c)$	$\hat{\sigma}_\theta$ ($\sigma_\theta = 22.5$)	$d_g(\hat{c}, c)$	$\hat{\sigma}_\theta$ ($\sigma_\theta = 22.5$)
$\kappa = 100$	$n = 30$	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)

$\sigma_\theta = \pi/8 \approx 22.5°$. 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.

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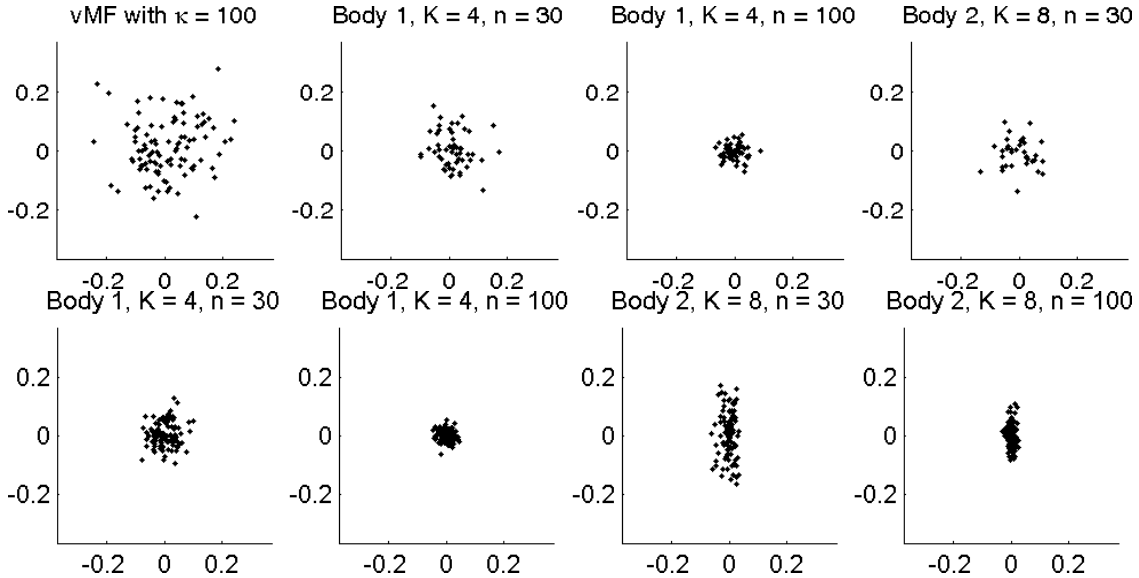
Figure 4: Scatters under their inverse exponential map of vMF(100) (top left panel) and of estimates $\hat{c}$ for the rigid rotation model (top row) and the twisting model (bottom row). The shape of the empirical covariance of $\hat{c}$ is different for the body 2–twist pair due to the proximity of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

(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 $d_g(c, \mu_j)$ 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

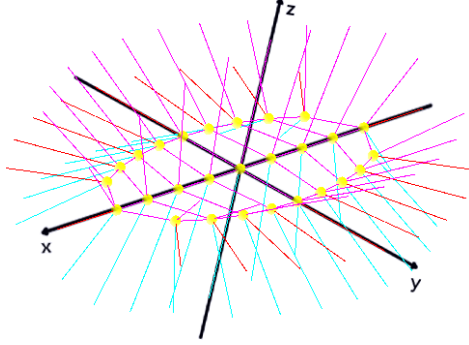In this section, an application of the proposed method to s-rep data is discussed.

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

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 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

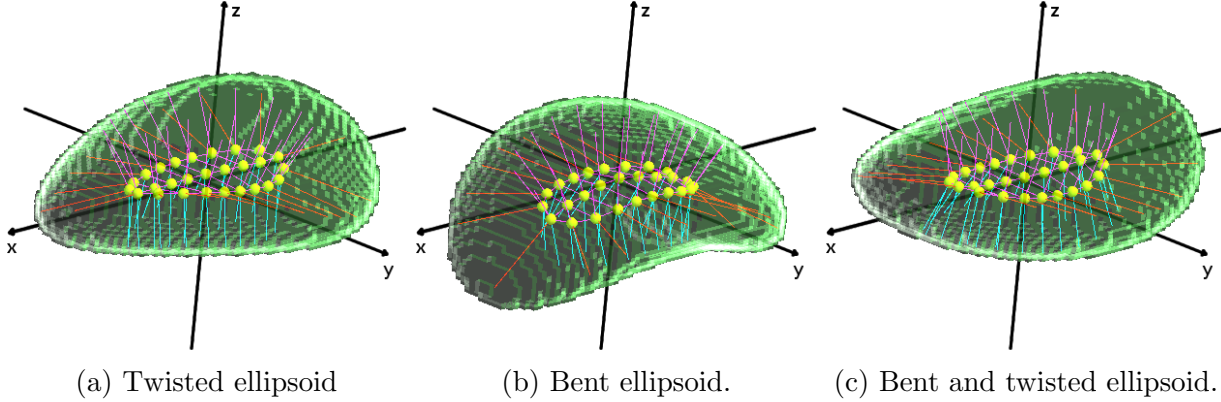| (a) Twisted ellipsoid | (b) Bent ellipsoid. | (c) Bent and twisted ellipsoid. |

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.

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.

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 standard error estimates, $\hat{se}_B(\hat{c}) = 3.44°$ and $\hat{se}_B(\hat{\sigma}) = 0.036$, are small and obtained by using the bootstrap algorithm as explained in Section 1.4 in the Supplementary Material.

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 and standard error of $\hat{se}_B(\hat{c}) = 0.49°$ 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

18

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. The estimated standard error $\hat{\text{se}}_B(\hat{c}_1) = 3.34°$ of the first rotation axis is small in contrast to the second rotation axis with $\hat{\text{se}}_B(\hat{c}_2) = 27.50°$. This indicates that a precise estimation of the rotation angle is critical for hierarchical deformations. In contrast to single deformations the rotation angles affect the estimation of the rotation axes (cf. Remark 3 in Section 3). Thus, the large standard error is a result of the inaccurate estimate of the rotation angle α in case of bending as discussed in the previous paragraph.

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 (Pierrynowski et al., 2010).

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. To provide measures of the motion of the bones within the lower leg and thigh the volunteer consented to have two 6 mm pins surgically inserted into his femur and tibia using local anesthesia. The methodological concerns cited by Ramsey et al. (2003) were followed to guide pin placement. The insertion sites were selected to minimize neuro-muscular effects that could influence natural knee motion. Self-report by the volunteer stated that the pins were not painful and did not influence his walking pattern. 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, described in detail in Section 6 in the Supplementary Material. 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). During movement, the leg XYZ axes relative to the thigh XYZ axes were used for the analysis.

The volunteer walked at 2.5 mph on a motor driven treadmill. After a familiarization 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 standard errors were estimated obtained by bootstrapping (cf. Section 1.4 of the Supplementary Material). 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 $\hat{c}_1 = (0, -1, 0.06)$ with standard error $\hat{se}_B(\hat{c}_1) = 0.23°$ and $\hat{c}_2 = (0.06, -0.06, -1)$ with standard error $\hat{se}_B(\hat{c}_2) = 20.34°$. 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 $\hat{c}_1 = (0.01, -1, 0.07)$ with standard error $\hat{se}_B(\hat{c}_1) = 0.35°$ and $\hat{c}_2 = (0.01, 0.03, -1)$ with standard error $\hat{se}_B(\hat{c}_2) = 16.58°$. Quite consistent with the simulations in Section 5, the standard error estimates are larger
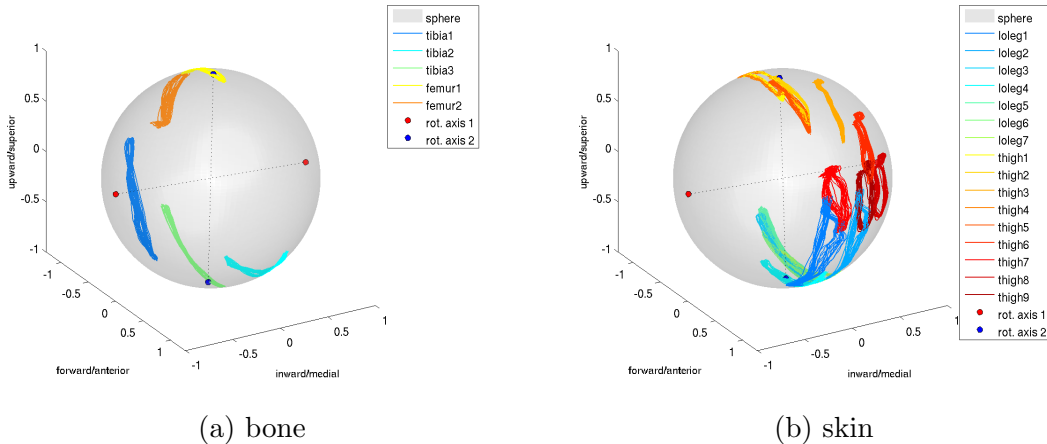
(a) bone   (b) skin

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.

for the second axis. In both cases (bone and skin data), the higher standard error of the second rotation axis is due to the small range of rotation angles about this axis. The larger standard error of the second axis for the bone data is deemed to be a result of the small number of observed directions. Estimation results of the rotation angles can be found in the Supplementary Material. Future work lies in the improved estimation based on a more careful time modeling of knee motions such as that proposed by Rivest (2001) who examined elbow motion.

7   Discussion

The paper proposes a novel nonparametric method to estimate rotational deformations from directional data. For the simple, single rotational deformation, the estimation procedure does not depend on the latent variable θ_j. In addition, the paper proposed an estimation procedure for hierarchical deformations, which depends on the order and the specification of functions f_{1j}, f_{2j}. An important future research is the improved prediction of the rotation angles θ_j and ψ_j, in particular for hierarchical rotations as introduced in Section 3. This includes the classification of directions into carefully chosen partitions I_1 and I_2 as well as the development of methods to predict f_j from the data as noted in Remarks 1 and 3. In order to extend the estimation method to more than two rotational deformations, characterization of geometry related to composition of hierarchical deformations is essential. 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 $d_g^2(\delta(c, r^*(c)), x) = (\arccos(c'x) - \arccos(c'\mu^*))^2 = (\|\text{Log}_c x\| - \|\text{Log}_c \mu^*\|)^2$. Thus for fixed c, $d_g^2(\delta(c, r^*(c)), x) \geq \min_y (\|\text{Log}_c x - y\| - \|\text{Log}_c \mu^* - y\|)^2$. 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

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

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)

Acknowledgments

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.

Chapman, G. R., Chen, G., and Kim, P. T. (1995), "Assessing Geometric Integrity Through Spherical Regression Techniques," *Statistica Sinica*, 5, 173–220.

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.

— (2012), "On the Meaning of Mean Shape: Manifold Stability, Locus and the Two Sample Test," *Annals of the Institute of Statistical Mathematics*, 64, 1227–1259.

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.

Pierrynowski, M., Costigan, P., Maly, M., and Kim, P. (2010), "Patients with Osteoarthritic Knees Have Shorter Orientation and Tangent Indicatrices During Gait," *Clinical Biomechanics*, 25, 237–241.

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.

Ramsey, D. K., Wretenberg, P. F., Benoit, D. L., Lamontagne, M., and Nemeth, G. (2003), "Methodological concerns using intra-cortical pins to measure tibiofemoral kinematics," *Knee surgery, sports traumatology, arthroscopy : official journal of the ESSKA*, 11, 344–349.

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.

— (1998), "Some Linear Models for Estimating the Motion of Rigid Bodies with Applications to Geometric Quality Assurance," *J. Amer. Statist. Assoc.*, 93, 632–642.

— (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.