Optimal Statistical Operators for 3-Dimensional Rotational Data: Geometric Interpretations and Application to Prosthesis Kinematics¶

OYVIND STAVDAHL*†‡§, ANNE KARIN BONDHUS*,
KRISTIN Y. PETTERSEN* and KJELL E. MALVIG*

Keywords: *Orientation statistics, rotational data, rotation matrix, quaternion, Euler parameter, Euler angle, orientation vector, attitude vector, kinematics, orthopaedics, prosthetics, biomechanics*

Rotational data in the form of measured three-dimensional rotations or orientations arise naturally in many fields of science, including biomechanics, orthopaedics and robotics. The cyclic topology of rotation spaces calls for special care and considerations when performing statistical analysis of rotational data. Relevant theory has been developed during the last three decades, and has become a standard tool in some areas. In relation to the study of human kinematics and motion however, these concepts have hardly been put to use. This paper gives an introduction to the intricacies of three-dimensional rotations, and provides a thorough geometric interpretation of several approaches to averaging rotational data. A set of novel, simple operators is presented. Simulations and a prosthetics-related real-world example involving wrist kinematics illuminate important aspects of the results. Finally generalizations and related subjects for further research are suggested.

1. Introduction

The overall function of human joints is of a mainly rotational nature. Consequently, the study of rotational movement is essential to the study of human motion and relevant to the design of prosthetic devices, planning of orthopaedic surgical procedures, ergonomics, sport science and biomechanics in general, as well as to technical fields like robotics and vessel control. The inherent cyclicity of rotations differs substantially from the properties of linear translations. Therefore many statistical methods cannot be translated into the rotational domain in a straightforward manner.

One objective of this paper is to briefly review the theoretical basis for conducting unbiased statistical analysis on general three-dimensional rotational data. Much of the relevant literature is very theoretical and abstract. A second purpose of this work is therefore to derive and analyse optimal operators related to orientational average and variance by using geometric considerations, in an attempt to give the reader a concrete understanding of the operators' practical significance, applications and implications. The resulting operators, which we refer to collectively as the Cosine Statistics, operate on data in rotation matrix form, and are closely related to theoretical results previously

derived by others. The performance of the Cosine Statistics is compared to that of more naive approaches by application to simulated data. In a real-world example we employ the Cosine Average to estimate the optimal static alignment of a prosthetic terminal device.

1.1. *On representations*

The choice of representation for body or joint orientations and rotations has occupied the biomechanics society for more than two decades, and several publications have been dedicated to the subject (Wu & Cavanagh, 1995; Chao, 1980; Grood & Suntay, 1983; Andrews, 1984; Woltring, 1994; Sheehan & Mitiguy, 1999). Any rotation in space can be represented by a single rotation through a suitable angle θ about a suitable fixed axis represented by a unit 3-vector $\mathbf{k}$. Similarly, any orientation in space can be obtained, and thus represented, by such a single rotation from a reference orientation. Each axis/angle pair $(\mathbf{k}, \theta)$ corresponds to a unique 3×3 rotation matrix which can be parameterized in many ways. One popular 'joint coordinate system' representation, often referred to as the Chao/Grood and Suntay convention (Chao, 1980; Grood & Suntay, 1983), in effect resembles an Euler or Cardan angles convention. This model has been opposed on the basis of singularities and poor metric properties, the latter which can cause the individual angles to be much larger than the total rotation in question (Woltring, 1994). Alternatives that have been suggested to alleviate these problems include the four-parameter 'unit quaternion' (Euler parameter) representation (Andrews, 1984) $\mathbf{q} = [q_0, q_1, q_2, q_3]^T \triangleq [\cos \frac{\theta}{2} \sin \frac{\theta}{2} \mathbf{k}^T]^T$, and the closely related three-parameter 'attitude vector' (Woltring, 1994) $\mathbf{\Theta} = [\theta_x, \theta_y, \theta_z]^T \triangleq \theta \mathbf{k}$, $\theta \in [0, \pi]$, which we will refer to as the 'orientation vector' to comply with pertinent robotics literature. Though superior to Euler angles in many respects, both these representations as well as the full rotation matrix representation may yield meaningless results if averaged arithmetically. This calls for a different approach to statistics that accounts for the special topology of three-dimensional rotations.

1.2. *Previous research*

Most prosthetic devices available exhibit very simple kinematic structures. A recent study explored compensatory movements in response to different static wrist alignments by comparing extremal values of scalar joint angles (Landry & Biden, 2002). The majority of literature relevant to upper-limp prosthesis kinematics, however, is of a mainly qualitative or non-statistical nature.

In the more general field of biomechanics, statistics has been applied to rotational data for example in conjunction with movement estimation from noisy marker position data (Woltring *et al.*, 1985; Veldpaus *et al.*, 1988). The reported results apply to differential (i.e. small) rotations, justifying certain simplifications which were actively exploited by the respective investigators. Several studies related to the human wrist have focused on joint range of motion in the well-known 'clinical angles' flexion/extension, adduction/abduction and internal/external rotation. The reported results include averaged extreme values for the clinical angles across a test population (Brumfield *et al.*, 1966; Sarrafian *et al.*, 1977; Ruy *et al.*, 1991); average arcs of motion for each clinical angle (Palmer *et al.*, (1985); and the 'centroid' of motion in terms of the average flexion/extension angle and adduction/abduction angle (Youm *et al.*, 1978). Such practice is acceptable when the clinical angles are viewed separately as such. However, pairs or triplets of average angles do not give a representative value for the joint's overall

average orientation; clinical angles resemble the Chao/Grood and Suntay convention (Chao, 1980; Grood & Suntay, 1983), and thus suffer the metric defects previously mentioned. Use of Euler parameters has been suggested for interpolating orientation data in animation applications (Ariel *et al.*, 2000). Rotation statistics has been applied for the analysis of variations of static postures (Rancourt *et al.*, 2000), but most of the specific methods used rely on the assumption of closely spaced data. However, there seems to have been little or no tradition in the study of human kinematics and motion, including orthopaedics, to consider the statistics of arbitrary sets of three-dimensional orientations.

The theoretical literature relevant to orientation statistics has been concisely reviewed by Krieger Lassen, Juul Jensen and Conradsen (1994). Important results are related to the Bingham and the matrix Fisher probability distributions, which are equivalent but related to quaternion vectors and rotation matrices, respectively. This theory has been applied for estimating average crystal orientation from data obtained in an electron microscope (Krieger Lassen *et al.*, 1994). Others who reported the same application calculated the average orientation as the normalized vectorial sum of individual quaternions (Humbert *et al.*, 1996; Humphreys *et al.*, 2001). This latter approach is computationally trivial; however, in some situations it may lead to absurdities which makes the approach less suitable for general application.

2. Mathematical preliminaries

The set of all three-dimensional rotation matrices i.e. all orthogonal 3×3 matrices R with $R = 1$, is referred to as the special orthogonal group of order 3 and is denoted SO(3). Inversion and transposition are equivalent operations on SO(3):

$$R^{-1} = R^T \Leftrightarrow R^T R = R R^T = I \tag{1}$$

The orientation of a coordinate frame S_1 relative to a frame S_2 can be represented by a rotation matrix R_2^1. Let $\mathbf{p}_1$ be a coordinate vector decomposed in frame S_1, and let $\mathbf{p}_2$ be the same vector decomposed in frame S_2 with the same origin as S_1. The transformation between $\mathbf{p}_1$ and $\mathbf{p}_2$ is then given by:

$$\mathbf{p}_2 = R_2^1 \mathbf{p}_1 \Leftrightarrow \mathbf{p}_1 = R_2^{1T} \mathbf{p}_2 \triangleq R_1^2 \mathbf{p}_2 \tag{2}$$

Let the basis vectors of the two frames be the orthonormal unit vector triples $(\mathbf{x}_1, \mathbf{y}_1, \mathbf{z}_1)$ and $(\mathbf{x}_2, \mathbf{y}_2, \mathbf{z}_2)$, respectively. Then R_2^1 can be written:

$$R_2^1 = \begin{bmatrix} \mathbf{x}_2 \cdot \mathbf{x}_1 & \mathbf{x}_2 \cdot \mathbf{y}_1 & \mathbf{x}_2 \cdot \mathbf{z}_1 \\ \mathbf{y}_2 \cdot \mathbf{x}_1 & \mathbf{y}_2 \cdot \mathbf{y}_1 & \mathbf{y}_2 \cdot \mathbf{z}_1 \\ \mathbf{z}_2 \cdot \mathbf{x}_1 & \mathbf{z}_2 \cdot \mathbf{y}_1 & \mathbf{z}_2 \cdot \mathbf{z}_1 \end{bmatrix} \tag{3}$$

Thus, the elements of R_2^1 are the direction cosines of the basis vectors of S_2 with respect to those of S_1.

As already stated, any rotation can be defined by an angle θ and a unit vector $\mathbf{k}$. The matrix representing this rotation will be denoted $R_{\mathbf{k}, \theta}$, which is equivalent to $R_{-\mathbf{k}, -\theta}$ and which is cyclic (mod 2π) with respect to θ. The formula for $R_{\mathbf{k}, \theta}$ is given in many kinematics textbooks (Spong & Vidyasagar, 1989).

2.1. *Metric on SO(3): angular distance*

Let S_1 and S_2 denote two coordinate frames whose orientations are given by the rotation matrices R_1 and R_2, respectively. The geodesic angular distance between R_1 and R_2 then is defined to be the minimum angle β by which any of the two frames must be rotated, about a suitable fixed axis, in order to make the two frames' axes parallel. The relationship between the two matrices can be expressed as $R_1 = R_{k,\beta}R_2$ so that $R_{k,\beta} = R_1 R_2^T$. The geodesic distance between R_1 and R_2 is given by the formula

$$\angle(R_1, R_2) = \beta = \arccos\left(\frac{\mathrm{tr}(R_1 R_2^T) - 1}{2}\right) \tag{4}$$

where tr is the *trace* operator, i.e. the sum of the argument's diagonal elements. This metric, or distance measure, is the core of many optimization problems related to rotations, including some of those discussed in this paper.

3. Derivation of the Statistics

Let $\{R_i\}_{i=1}^n$ be a set of rotation matrices. The aim of this section is to define an operator which calculates the single rotation matrix which best represents the whole set, i.e. the 'average rotation' of the set, as well as operators related to the dispersion of the data about this average. For an average quantity to be justifiable, it must satisfy certain criteria, e.g.: for trivial data sets it must yield the expected result; the average must be independent of the order of the data elements; and for a given data set it must provide a unique average value or an explanation of why a unique average cannot be found.

3.1. *Shortcomings of vectorial approaches*

We will start by showing how two techniques used in the literature, namely the arithmetic average of orientation vectors and normalized average of Euler parameter vectors, respectively, may come short in trivial cases. Without loss of generality we will assume rotations about the x-axis (i.e. $\mathbf{k} = [1, 0, 0]^T$), so that the two representations can be represented completely in the (θ_x, θ_y) and the (q_0, q_1) plane, respectively. These planes are illustrated in the left and right part of Figure 1, respectively. As the angle of rotation θ sweeps through the interval $[0, 2\pi]$ the orientation vector Θ goes through the path defined by the bold arrows labelled 1, 2 and 3, in that order. The dashed arrow 2 represents the sign shift occurring at $\theta = \pi$, i.e. the representation singularity of the orientation vector representation. In contrast, the plain Euler parameter representation has no intrinsic singularity. When averaging Euler parameter vectors, however, the double representation of this space creates a dilemma: each orientation can be equivalently represented by two oppositely directed vectors; which sign should be used for the averaging? A common convention is to choose the sign so that $q_0 \geq 0$ (Humbert *et al.*, 1998) which is the case in the right part of Figure 1 where the labelled arrows correspond to those in the left figure. As indicated by arrow 2, the sign convention has introduced a representation singularity very similar to that of the orientation vector.

We now quickly review a trivial case in which the data set contains only two elements, corresponding to $\theta_1 = \pi - \delta\theta$ and $\theta_2 = \pi + \delta\theta$, where $|\delta\theta| \ll \pi$ and $\mathbf{k} = [1, 0, 0]^T$ as before. It is then reasonable to expect the average of the rotations defined by $(\mathbf{k}, \theta_1)$ and $(\mathbf{k}, \theta_2)$ to be that corresponding to $(\mathbf{k}, \pi)$, i.e. a rotation through the angle π about the x-axis. The arithmetic average of the corresponding orientation vectors, however, is the null vector, which implies $\theta = 0$, i.e. no rotation at all.
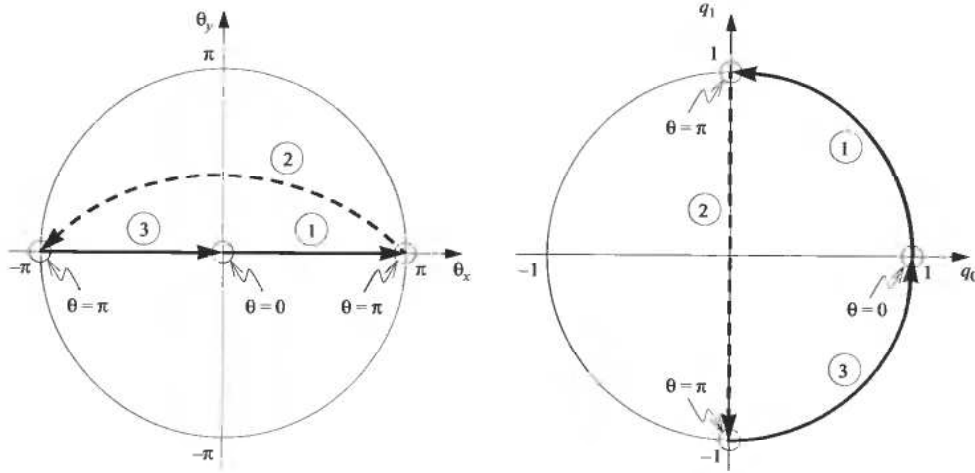
Figure 1.   Trajectory of the first two elements of the orientation vector (left) and the Euler parameter vector (right) when the angle of rotation is gradually changed from 0 to 2π. Rotations performed about the *x*-axis.

Normalizing the average of the corresponding Euler parameter vectors yields $q_0 = 1$ and $q_1 = q_2 = q_3 = 0$, also implying a zero angle of rotation. Data sets similar to that considered here may appear as subsets of larger data sets if the data is dispersed and/or if the reference frame is chosen such that the data appear with angles of rotation in the vicinity of π. The consequence is potential bias of the estimated average orientation. This is a very serious drawback that calls for extreme caution if applying these techniques.

3.2. *A least squares attempt*

A principal solution has been suggested for calculating the average of a finite group (Diaconis, 1988). We propose to generalize this definition to a class of average operators for finite samples (i.e. discrete subsets) on finite or infinite groups as follows:

Definition 1.   *Let* $\Gamma = \{\gamma_i\}_{i=1}^{n}$ *be a finite sample on (i.e. a discrete subset of) a group* G, *and let* ρ *be a metric on* G. *Furthermore, let* h *be a lower bounded scalar function which is monotonously increasing throughout the range of* ρ. *Define*

$$J(s; \Gamma) = \frac{1}{n} \sum_{i=1}^{n} h(\rho(s, \gamma_i)) \tag{5}$$

where $s \in G$. *The group element* $\eta \in G$ *is then said to be a candidate sample average of* Γ *if* η *minimizes* $J(s; \Gamma)$. *The number* $J(\eta; \Gamma)$ *is called the* ρh–*spread of* Γ.

Note that Definition 1 guarantees the existence, but not the uniqueness, of a minimum of J; hence the use of the term *candidate sample average*. By setting $\Gamma = \{R_i\}_{i=1}^{n}$, substituting the angular metric of Equation (4) for ρ and choose $h(\cdot) = (\cdot)^2$, we get what we may refer to as the Least Squares Average $\bar{R}_{LS}$, defined by

$$\bar{R}_{LS}\{R_i\} \triangleq \arg \left(\min_{\bar{R} \in SO(3)} J_{LS}(\bar{R}; \{R_i\}) \right)$$

$$J_{LS}(\bar{R}; \{R_i\}) \triangleq \frac{1}{n} \sum_{i=1}^{n} \angle (R_i, \bar{R})^2 \tag{6}$$
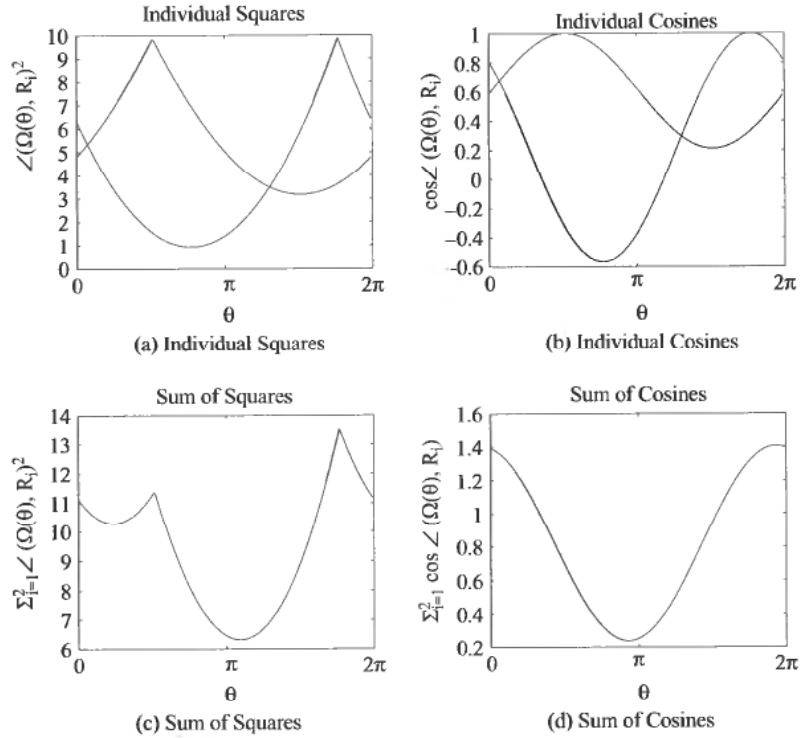
Figure 2.   Individual and summed squares and cosines, respectively, of $\angle(\bar{R}(\theta), R_i)$ for $i = 1, 2$.

Here, the operator yields the argument $\bar{R}$ corresponding to the extremum of J_{LS} rather than the extremal value itself. A major defect of this solution is the likely existence of local minima of J_{LS}. This is illustrated in Figure 2 for a random data set with $n = 2$. Figure 2(a) shows two graphs, each corresponding to one of the summed elements of (6). For each graph we have $\bar{R} = \Omega(\theta) = \bar{R}_{\mathbf{k}, \theta}\bar{R}_0$ with the unit vector $\mathbf{k}$ and the rotation matrix $\bar{R}_0$ random but fixed, and with $\theta \in [0, 2\pi)$ (i.e. $\bar{R}$ spans a one-dimensional subspace of SO(3)). Because of the cyclicity of SO(3) each parabola in the figure will repeat itself once every turn of θ. Figure 2(c), which shows the sum of the two functions above (i.e. it essentially shows J_{LS} as $\bar{R}$ is rotated by an angle of 2π about $\mathbf{k}$), clearly exhibits two minima, one for each element of $\{R_i\}$. There are a corresponding number of maxima at which J_{LS} is not differentiable, severely complicating the process of finding the possibly non-unique global minimum. Considering that Figure 2 is a one-dimensional section of a complex, three-dimensional problem, we conclude that a least squares approach is effectively useless for the present problem in the form of Equation (6).

3.3. *The cosine average*

Figure 2(b) and 2(d) show the functions we get by repeating the process just described after choosing $h(\cdot) = 1 - \cos(\cdot)$. In this case each of the individual functions (upper right pane) as well as their sum (lower right) constitutes a (generally scaled,

biased and angularly offset) cosine function, guaranteeing a unique minimum unless the scaling factor happens to be zero. This gives rise to the following definition:

Definition 2.   *Given a general sample* $\{R_i\}_{i=1}^n$, $R_i \in SO(3)$, *the* Cosine Average *(CA) of the sample is denoted* $\bar{R}_{\cos}$ *and is defined as*

$$\bar{R}_{\cos}\{R_i\} \triangleq \arg \left(\min_{\bar{R} \in SO(3)} J_{\cos}(\bar{R}; \{R_i\}) \right)$$

$$J_{\cos}(\bar{R}; \{R_i\}) \triangleq \frac{1}{n} \sum_{i=1}^n (1 - \cos(\angle(R_i, \bar{R}))) \tag{7}$$

whenever the solution is unique; if not, the CA is said to be indefinite.

Simplification of Equation (7) by substituting (4) and setting $\Sigma_{i=1}^n R_i = R_\Sigma$ yields

$$\bar{R}_{\cos}\{R_i\} = \arg \left(\max_{\bar{R} \in SO(3)} J_{tr}(\bar{R}; \{R_i\}) \right)$$

$$J_{tr}(\bar{R}; \{R_i\}) \triangleq tr(R_\Sigma \bar{R}^T) \tag{8}$$

(Note the change from *min* in (7) to *max* in (8)). This is exactly the maximum likelihood estimate of the mean orientation of a matrix Fisher probability distribution (Downs, 1972). It is also the projection, in a least squares sense, of R_Σ on SO(3) since if $\mathbf{r}_{ij}^T$ and $\bar{\mathbf{r}}_j^T$ are the j'th row of R_i and $\bar{R}$, respectively, we have (Krieger *et al.*, 1994; Stephens, 1979)

$$\sum_{i=1}^n \sum_{j=1}^3 |\mathbf{r}_{ij} - \bar{\mathbf{r}}_j|^2 = 6n - 2tr(R_\Sigma \bar{R}^T). \tag{9}$$

3.4. *Existence and uniqueness*

The function $J_{\cos}$ in Equation (7) is a continuous function on a closed, bounded set, and thus it must exhibit at least one minimum.

For the uniqueness properties of the CA, the picture is more complicated. A theoretical treatment of this problem has been given in several previous publications (Downs, 1972; Prentice, 1986; Khatri & Mardia, 1977; Stephens, 1979). The following paragraphs are built on these publications, and contain what to our best knowledge is a novel, thorough geometric interpretation of their results. For the discussion to follow we let $R_\Sigma = \Sigma_{i=1}^n R_i = [\mathbf{r}_{\Sigma 1} \ \mathbf{r}_{\Sigma 2} \ \mathbf{r}_{\Sigma 3}]^T$ and $\bar{R} = [\bar{\mathbf{r}}_1 \ \bar{\mathbf{r}}_2 \ \bar{\mathbf{r}}_3]^T$. Furthermore we recall that we can always write

$$R_\Sigma = USV^T \tag{10}$$

where (10) is the singular value decomposition of R_Σ, so that U and V are orthogonal (but not necessarily rotation) matrices and $S = \text{diag}(\sigma_1, \sigma_2, \sigma_3)$ where $\sigma_1 \geq \sigma_2 \geq \sigma_3 \geq 0$ are the singular values of R_Σ. We can then write

$$tr(R_\Sigma \bar{R}^T) = tr(S\tilde{R}^T) \tag{11}$$

with $\tilde{R}^T = V^T \bar{R}^T U$. The rows and columns of S are orthogonal since S is diagonal, and so are the rows and columns of $\tilde{R}$ because this matrix is orthogonal. Consequently, without loss of generality the problem can be discussed in terms of fitting one orthogonal vector triplet to another. We will do this in the following by assuming that R_Σ is diagonal
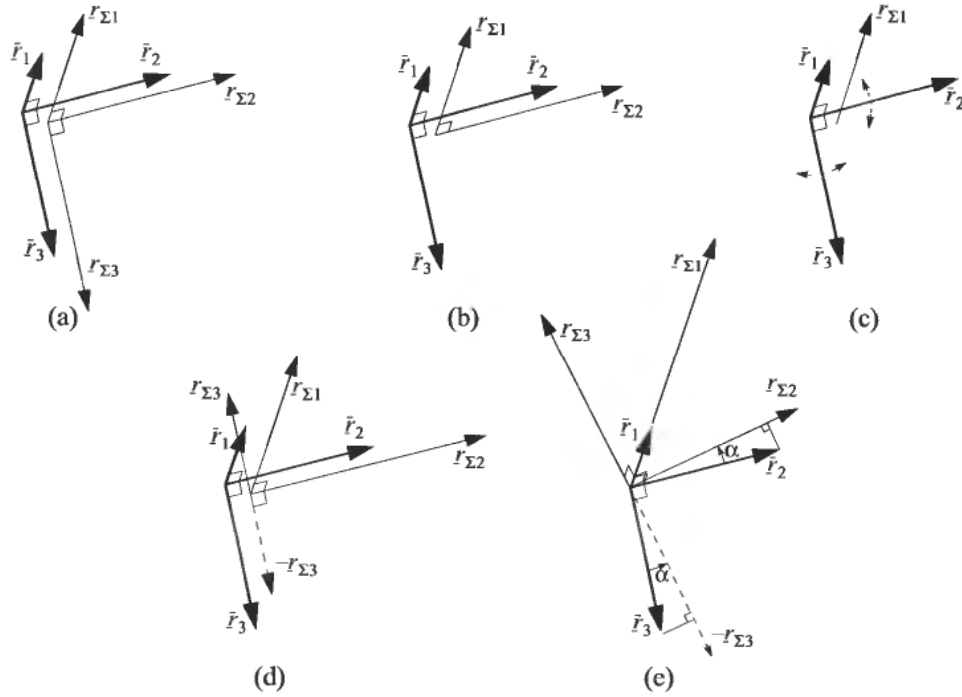
Figure 3. The five non-trivial uniqueness modes of the Cosine Average: geometric interpretations.

and fitting the orthogonal matrix $\bar{R}$ to it rather than explicitly involving S and $\tilde{R}$ in the analysis. (If R_Σ is negative, the transformation (11) will yield det $S>0$ and $\tilde{R}<0$, i.e. $\tilde{R}$ will no longer be a proper rotation matrix. This can easily be resolved and does not invalidate the discussion.) We discuss the different uniqueness modes one at a time:

(i) det $R_\Sigma >0$

R_Σ resembles a right-handed (though generally not orthonormal) coordinate frame. In this case the CA is uniquely defined; its row vectors are fitted to those of R_Σ in the least squares sense (Figure 3(a)).

(ii) rank $R_\Sigma =2$

The information contained in the sample $\{R_i\}$ is related to two coordinate axes only; the information pertaining to the third axis is cancelled out (Figure 3(b)). In this case $\bar{R}$ is uniquely defined by the two nonzero vectors of R_Σ; the third vector of $\bar{R}$ is determined by the right-hand rule.

(iii) rank $R_\Sigma =1$

In this case only the first vector $\bar{r}_1$ of $\bar{R}$ is determined by the data. The remaining vectors are restricted to forming an orthonormal right-handed triplet with $\bar{r}_1$, but their orientation in the plane normal to $r_{\Sigma 1}$ is undetermined (Figure 3(c)).

(iv) rank $R_\Sigma =0$

This degenerate case arises if the data are perfectly homogenously distributed on SO(3) so that all elements of R_Σ are equal to zero. In this case the notion of average is meaningless and is said to not exist.

(v) det $R_\Sigma <0$

R_Σ resembles a left-handed (though generally not orthonormal) coordinate

frame. In this case the uniqueness can be derived from the two smallest singular values of R_Σ as follows:

(a) $\sigma_2 \neq \sigma_3$

The vector of R_Σ corresponding to σ_3 is distinctly shorter than the other two. To obtain what we may call a right-handed data set, we change the sign of that shorter vector before calculating the solution (Stephens, 1979). Figure 3(d) illustrates the geometry of this case. Once the sign is changed, the problem becomes equivalent to that of case (1) above.

(b) $\sigma_2 = \sigma_3$

This situation is illustrated by Figure 3(e). The dominant axis of the data set is associated with the larger singular value σ_1, thus $\bar{\mathbf{r}}_1$ is aligned with it. The remaining two axes of R_Σ are of equal length; however, we will always have $\mathbf{r}_{\Sigma1}^T\bar{\mathbf{r}}_1 + \mathbf{r}_{\Sigma2}^T\bar{\mathbf{r}}_2 = 0$ so the orientation of the last two vectors of $\bar{R}$ in the plane normal to the dominant axis is undetermined. Consequently there are infinitely many solutions.

3.5. *Calculating the cosine average*

From the preceding discussion we can derive a simple formula for calculating the CA. Since S is diagonal with nonnegative elements, it is obvious that if $R_\Sigma > 0$ (and thus $\tilde{R} > 0$) the trace in Equation (11) is maximized when $\bar{R}^T = I$. This immediately yields $\bar{R}^T_{\cos} = VU^T$ where V and U are as in (10). If $R_\Sigma < 0$ and $\sigma_2 \neq \sigma_3$, the aforementioned sign shift must be carried out. It is the third and smallest diagonal element of S that needs to be changed, which can be done by setting $R_{\Sigma-} = USJV^T$ with $J = \mathrm{diag}\,(1, 1, -1)$ and S, U and V as in (10), and then calculating the CA from $R_{\Sigma-}$ instead of R_Σ. A general formula which covers both cases is therefore given by

$$\bar{R}^T_{\cos} = VJU^T \tag{12}$$

with $J = diag(1, 1, s)$, where $s = \mathrm{sgn}(\det R_\Sigma)$ and sgn is the *signum* function.

3.6. *Definity*

Whenever the CA is not uniquely defined (i.e. it is indefinite), this is so because the data set exhibits some kind of rotational symmetry. Obviously the data can be infinitely close to being rotationally symmetric without actually being so. When using the CA for analysing e.g. biomechanical data, we need a measure of *how far the CA is from being indefinite* in order to correctly interpret the results. We propose a simple measure which we will denote the *definity* and which can be calculated as

$$Def\,\{R_i\} \triangleq \begin{cases} \frac{1}{2}\,\dfrac{\sigma_2 + \sigma_3\,\mathrm{sgn}\,(\det R_\Sigma)}{\sigma_1} & \text{if rank } R_\Sigma > 0 \\ 0 & \text{otherwise} \end{cases} \tag{13}$$

where σ_i is the i'th singular value of R_Σ. As can be readily verified, the definity takes on the value 1 if all elements R_i of the data set are identical (i.e. the most definite case possible), and becomes zero if the CA is indefinite; in all intermediate cases the definity takes on intermediate values.

3.7. Cosine variance and standard deviation

We recall that in basic statistics, the average is chosen such that the variance, which is defined by means of the former, is minimized. We propose to define a variance and standard deviation associated with the CA in a similar way:

Definition 3. *Given a sample* $\{R_i\}_{i=1}^n$, $R_i \in SO(3)$ *with a CA* $\bar{R}_{\cos}\{R_i\}$. *The* Cosine Variance *of the sample is then denoted* $V_{\cos}$ *and defined as*

$$V_{\cos}\{R_i\} \triangleq \frac{1}{2n} \sum_{i=1}^n (1 - \cos \angle(R_i, \bar{R}_{\cos}\{R_i\})). \tag{14}$$

Likewise, the Cosine Standard Deviation *of the sample is denoted* $SD_{\cos}$ *and defined as*

$$SD_{\cos}\{R_i\} \triangleq \arccos (1 - 2V_{\cos}\{R_i\}). \tag{15}$$

Calculation of the concentration matrix for a sample assumed to be matrix Fisher distributed is known to be extremely complicated (Krieger *et al.*, 1994). The much simpler Cosine Variance, however, can be found by simplifying Equation (14), which yields

$$V_{\cos}\{R_i\} = \tfrac{1}{4}(3 - \operatorname{tr}(\bar{R}_\Sigma \bar{R}^T)) \tag{16}$$

where $\bar{R}_\Sigma = \frac{1}{n} \Sigma_{i=1}^n R_i$ and $\bar{R} = \bar{R}_{\cos}\{R_i\}$. Note that while the concentration parameters of a matrix Fisher distribution carry information about the three-dimensional dispersion of the data, the Cosine Variance is a scalar quantity and carries information about angular spread only. Therefore these two quantities are related but by no means equivalent.

3.8. Further geometrical aspects and their implications

In this section we illuminate the geometric significance of some of the results more or less explicitly stated by Downs (1972).

Let $\bar{R}$ be the CA of a set $\{R_i\}_{i=1}^n$ of rotation matrices, and let $\{\Delta R_i\}_{i=1}^n$ be the set of rotational residuals given by

$$\Delta R_i = R_i \bar{R}^T \tag{17}$$

i.e. the distance (in matrix form) from the average $\bar{R}$ to the data element R_i. Furthermore let S_i and $\bar{S}$ be the coordinate frames which orientations are given by R_i and $\bar{R}$, respectively. It can easily be shown that

$$\sum_{i=1}^n \Delta R_i = \sum_{i=1}^n \Delta R_i^T \tag{18}$$

i.e. the sum of the residuals is symmetric. We recall that ΔR_i represents the orientation of frame S_i with respect to the average frame $\bar{S}$, while ΔR_i^T represents the orientation of $\bar{S}$ relative to S_i. Equation (18) therefore implies that $\{R_i\}$ are spaced around $\bar{R}_{\cos}$ in a *rotationally neutral* manner: if the set $\{\Delta R_i\}$ has a rotational bias, the set $\{\Delta R_i^T\}$ will have a rotational bias in the opposite direction, and the sums of the elements in each of the two sets cannot be equal.

Let $Q \in SO(3)$. Then the following linear property holds for the Cosine Average:

$$R_{\cos}\{QR_i\} = QR_{\cos}\{R_i\}, \tag{19}$$

i.e. SO(3) is distributive over the Cosine Average operation.

Premultiplication by a rotation matrix Q implies a change of coordinate frame, hence the value of the CA is independent of the reference coordinate frame used during computation. This is a very important property in fields where the choice of coordinate frames may vary among investigators and applications, in that it decouples the choice of frames for intermediate data processing and eventual reporting.

4. Experimental Section

4.1. *Application to Simulated Data*

In this section we compare the CA with the values obtained by averaging the parameters of several different representations, namely Euler parameter (quaternion) vectors, Euler angle triples and orientation vectors. We also include separate averaging of angle and axis of rotation, the latter in the form of unit vectors $\mathbf{k}_i$. For the Euler parameters we choose the sign so that the first element q_0 of each vector is positive or zero. Also, the arithmetic mean of the Euler parameter vectors and that of the axis of rotation vectors are normalized to obtain results in the valid ranges.

We somewhat arbitrarily use the Euler angles ψ, θ and ϕ, taken about the reference x-, y- and z-axis, resectively. The angles are calculated such that $\phi \in [-\pi, \pi]$, $\theta \in \langle -\pi/2, \pi/2 \rangle$ and $\psi \in [-\pi, \pi]$.

We will calculate the average of the set $\{R_i\}_{i=1}^{n}$ with

$$R_i = R_{\mathbf{k}, \alpha_i} R_0 \triangleq R_{\mathbf{k}_i, \beta_i} \tag{20}$$

where the axis of rotation $\mathbf{k}$ and the 'rotational offset' R_0 are fixed, while α takes on values in the interval $[0, 2\pi\rangle$. This model spans a one-dimensional subspace of SO(3). We let α_i have a Normal distribution with given fixed expectation μ and standard deviation σ. Intuitively we then expect that

$$\lim_{n \to \infty} \bar{R} = R_{\mathbf{k}, \mu} R_0 \triangleq R_{\mathbf{k}_{exp}, \beta_{exp}} \tag{21}$$

Several data sets were generated and the average orientation estimated using the different techniques. The parameter values chosen were: $\mathbf{k} = [1, 2, -5]/\sqrt{3}$, $R_0 = R_{\mathbf{k}_0, \theta_0}$ with $\mathbf{k}_0 = [1, 1, 1]/\sqrt{3}$, $\theta_0 = 60°$ and $\mu = 70°$. The angular difference between the estimated average orientations and the expected value $R_{\mathbf{k}_{exp}, \beta_{exp}}$, i.e. the estimation error, were recorded for different values of the standard deviation σ and different sample sizes n. Figure 4 shows the average estimation error resulting from 50 different data sets with $\sigma = 25°$ and $\sigma = 60°$.

To obtain a conservative result with respect to the qualitites of the CA, the singularity of the Euler angle representation was resolved by adding or subtracting 360° from ψ_i for certain values of θ to make $\psi(\theta)$ a smooth function. Similar techniques were applied to the axis/angle and the orientation vector data prior to averaging.

In the figure legends 'CosAv' denotes the Cosine Average, 'Quat' denotes quaternion (Euler parameter) averaging, 'Orient' denotes orientation vector averaging, 'k-Angle' denotes separate axis and angle averaging and 'Euler' denotes Euler angle averaging.

In the present example the CA and the Euler parameter based estimates were virtually indistinguishable in quality for $\sigma = 25°$ (less than 1.5% difference), while the CA was superior when the angular spread of the data was increased to $\sigma = 60°$ (Figure 4). Of the other estimators, only the average orientation vector was comparable to the former two, and only when the spreading of the data was kept low. Furthermore, the Cosine Average converges towards the expected value as the data sets become larger, even as the data spreading is increased; this is not the case for any of the other estimators
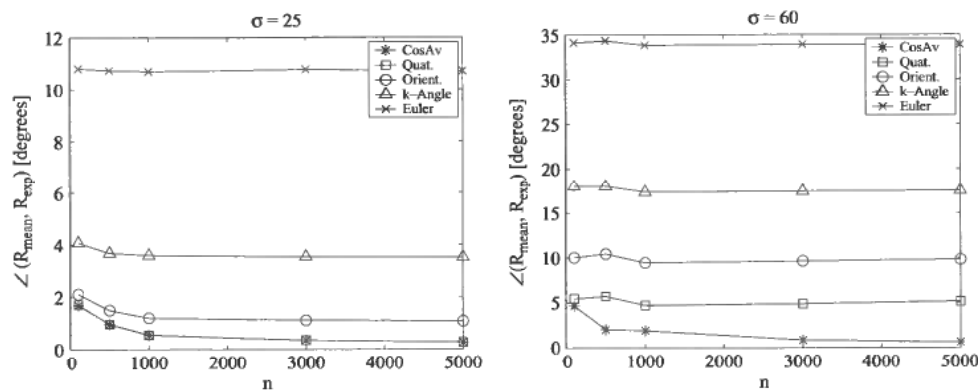
Figure 4.   Estimation error for $\sigma = 25°$ (left) and $\sigma = 60°$ (right) and different sample sizes n.

tested. The estimates based on average Euler angles display inferior performance in all scenarios investigated and is useless for most practical purposes.

4.2. *A real-world case*

To illustrate the relevance and applicability of the Cosine Statistics they were applied to authentic data from an experiment involving eight right-handed healthy subjects. Full details about this experiment are given in Stavdahl (2002); here, only a short description is given. The experiment aimed at investigating the average orientation of the wrist joint with respect to the forearm during activities of daily living (ADL). This quantity is of potential relevance to prosthetics, as it can be said to indicate the functionally optimal alignment of an anthropomorphic terminal device.

4.2.1. *Experimental set-up.*   A custom-made plaster of Paris gauntlet was firmly applied to the left hand and a socket made of the same material was applied to the left proximal forearm of each subject, and the four fingers were taped together to form a single unit. The purpose of these structures was to restrict the forearm, wrist and hand movements to those obtainable with a typical prosthetic terminal device. A MotionStar* electromagnetic motion capture system was employed to record the orientation of each socket during the performance of 15 predefined ADL. Figure 5 depicts an instrumented limb in the standard posture used to calibrate the sensors' orientation with respect to relevant body segments.

Four body-fixed coordinate frames were defined as indicated in the figure: a proximal frame S_u fixed to the ulna, and three hand-fixed frames S_h, $S_{h'}$ and S_w, the latter which was parallel to the S_u frame when the limb was in the reference posture. A global laboratory-fixed frame S_0 served as a reference for both sensors. Each sensor's orientation was sampled at 20 Hz during fifteen predefined activities of daily living (ADL). Each activity was performed twice, each run of 20s duration. Subsequently the data were transformed to yield one of the hand fixed frames' orientation with respect to S_u.

The average wrist orientation was estimated by using the CA operator, by averaging individual Euler angles, by averaging orientation vectors and finally by means of the normalized average Euler parameter vector. The calculations were carried out twice, first
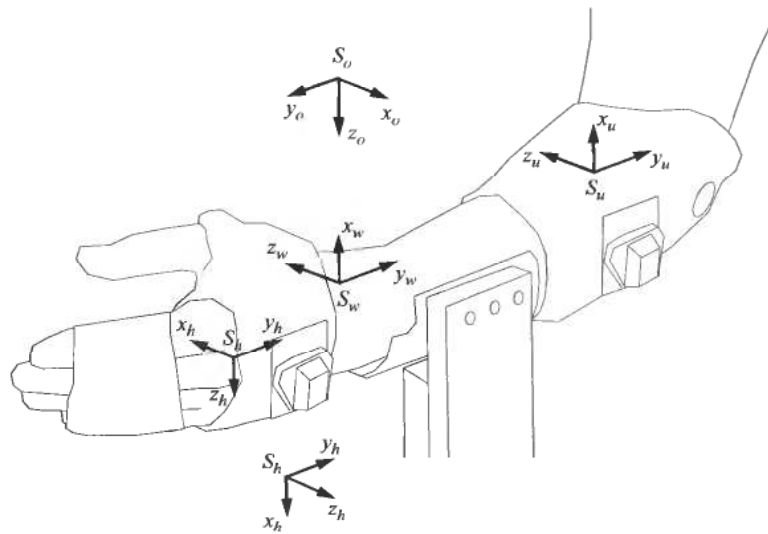

Figure 5.   Forearm and hand instrumented for tracking of wrist orientation; coordinate frame definitions.

by letting S_h represent the hand orientation and then by choosing $S_{h'}$ as the hand-fixed frame, to investigate the respective techniques' robustness with respect to this kind of choices. The average orientations calculated were all changed into rotation matrix form and transformed to express the orientation of the S_w frame with respect to S_u, a representation where all the clinical angles are zero when the overall angle of rotation is zero (i.e. in the calibration posture of Figure 5).

4.2.2. *Results.*   Table 1 displays two sets of data separated by slash symbols: quantities based on choosing S_h as the hand-fixed frame to the left, those obtained when choosing $S_{h'}$ as the hand-fixed frame to the right of the separators.

The leftmost column of Table 1 shows the angular deviation of the other average estimates from that obtained with the CA. The results based on S_h as the hand-fixed frame are in general agreement with the simulation results in that the Euler parameter and orientation vector based estimates only deviates moderately from the CA while the Euler angle based quantities show little agreement with the others. The orientation of $S_{h'}$ with respect to S_u was deliberately chosen such that their geodesic distance approximately equals π in the neutral posture of Figure 5 so that the Euler parameter vector representation of the data set exhibits the aforementioned sign shifts. The results clearly

Table 1.   Average orientation estimates: Deviation from the CA and implied clinical angles. All quantities in °.

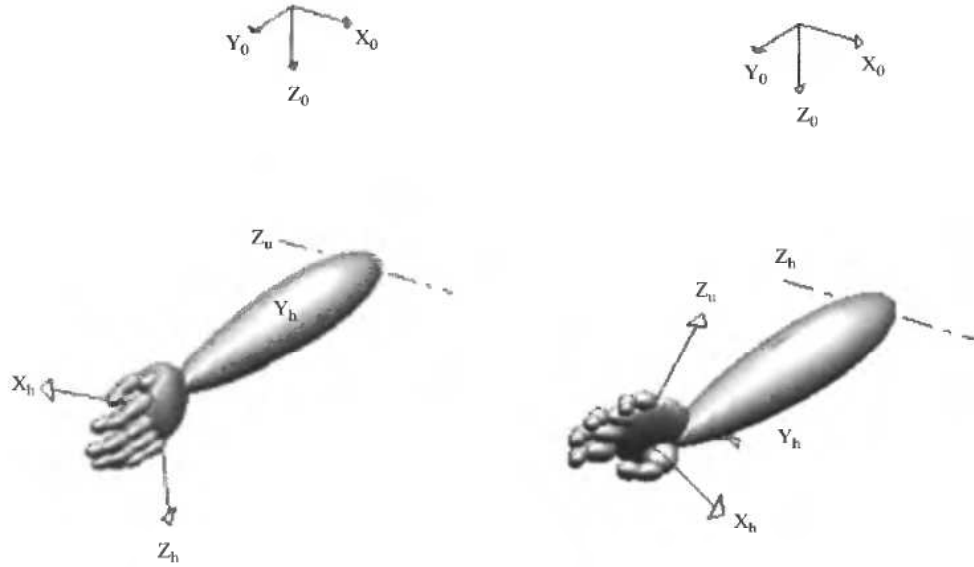|	Deviation from CA	Pronation	Extension	Ulnar deviation
Cosine Average	—	3.3	7.3	4.9
Euler Parameters	0.3/159	3.1/165	7.2/17.6	4.9/17.6
Orient. vectors	2.1/172	1.3/176	7.4/5.0	4.2/6.1
Euler Angles	104/147	− 24.5/132	− 23.9/ − 27	− 88.6/64

Figure 6.   Estimated average wrist orientation during ADL. Left: Cosine Average. Right: Estimate based on Euler parameters and hand orientation represented by $S_{h'}$. The line labelled Z_u represents the humeroulnar joint's axis of rotation.

illustrate the serious consequences of this singularity; the estimated average orientation in this case deviates from the CA by as much as 159°. The corresponding pronation angle is 165°, which is in fact outside the range of a healthy wrist. The non-physical nature of this result is further illustrated by Figure 6, where the left figure depicts the average orientation given by the CA and the right pane shows the wrist posture corresponding to the Euler parameter-based estimate. The former configuration deviates from the calibration posture by a net angle of 9.2° only. As implied by Equation (19), the CA yielded exactly the same result before and after the change of the hand-fixed frame.

5.   Generalizations and Future Research Issues

We would like to briefly mention an interesting generalization of the Cosine Average operator, namely time-domain filtering of rotational data. A slight generalization of Equation (8) yields

$$Y_k = \arg \left(\max_{Y \in SO(3)} \mathrm{tr}(U_\Sigma Y^T) \right)$$

$$U_\Sigma = \sum_{i=0}^{n} a_i U_{k-i} \tag{22}$$

where $a_0, a_2, \ldots, a_n \in \mathbb{R}$ are a set of constant coefficients. This construction implements an n'th order Finite Impulse Response (FIR) filter on SO(3). Generalization to an Infinite Impulse Response (IIR) filter is straightforward. We propose to employ this kind of filter e.g. in conjunction with interpolation or decimation of data sequences in rotation matrix form to avoid the complications associated with nonlinearities, double representations and singularities of other parametrizations.

The detailed properties of such filters remains to investigate, however, and constitutes a possible subject for future research. Possible further generalizations include what might be denoted rotationally non-symmetric filters, i.e. by substituting a set of 3×3 matrices for the filter parameters.

6. Discussion and Conclusions

This paper has pointed out the existence of an optimal average operator, referred to here as the Cosine Average, for three-dimensional rotational data. The operator is robust with respect to the dispersion of the data set and is completely insensitive to the choice of reference coordinate frame, making it ideal for biomechanics-related kinematic analysis where the appropriate choice of reference frame is often dictated by the application or by clinical tradition. It has clear geometric interpretations, yields a unique result except in degenerate data sets, and can be readily calculated using standard computer software tools. The definity statistic offers a measure of the degree of rotational symmetry in the data set, which is closely related to the degree of uniqueness of the Cosine Average. The associated Cosine Variance yields a simple measure of the degree of dispersion in the data.

Our prosthesis-related example illustrates the use of rotational statistics in a biomechanics and orthopaedics context, while also demonstrating the severe limitations of simpler approaches to orientation statistics.

We believe that rotation statistics represents a potential which is largely unexploited for the study of human kinematics and motion in general. The Cosine Statistics provide a simple toolbox that allows for explicit empirical analysis of complex rotational biomechanical data without any a priori assumptions with respect to the data set. Possible applications include optimization of kinematics of orthopaedic devices like orthoses and exoskeletal prostheses as well as prosthetic joint implants.

References

ANDREWS, J. G. (1984). On the specification of joint configurations and motions, *J. Biomech.*, 17(2), pp. 155–158.

ARIEL, G. B., BUIJS, R. J. C., PENNY, A. & CHUNG, S. G. (2000). Visualizing orientation using quaternions, *Proc. of the Sixth International Symposium on the 3D Analysis of Human Movement*, pages 21–24 (Cape Town, South Africa, May 2000, International Society of Biomechanics) pp. 21–24.

BRUMFIELD, R., NICKEL, V. & NICKEL, E. (1966). Joint motion in wrist flexion and extension, *South Med. J.*, 59, pp. 909–910.

CHAO, E. Y. S. (1980). Justification of triaxial goniometer for the measurement of joint rotation, *J. Biomech.*, 13, pp. 989–1006.

DIACONIS, P. (1988). *Group Representations in Probability and Statistics*, **volume 11** of *Lecture Notes—Monograph Series* (Institute of Mathematical Statistics, Hayward, CA, 1988).

DOWNS, T. D. (1972). Orientation statistics, *Biometrika*, 59, pp. 665–676.

GROOD, E. S. & SUNTAY, W. J. (1983). A joint coordinate system for the clinical description of three-dimensional motions: Application to the knee, *J. Biomech. Eng.*, 105(2), pp. 136–144.

HUMBERT, M., GEY, N., MULLER, J. & ESLING, C. (1996). Determination of a mean orientation from a cloud of orientations. Application to electron back-scattering pattern measurements, *J. Appl. Crystallogr.*, 29, pp. 662–666.

HUMBERT, M., GEY, N., MULLER, J. & ESLING, C. (1998). Response to Morawiec's (1998) Comment on determination of a mean orientation from a cloud of orientations. Application to electron back-scattering pattern measurements, *J. Appl. Crystallogr.*, 31, p. 485.

HUMPHREYS, F. J., BATE, P. S. & HURLEY, P. J. (2001). Orientation averaging of electron backscattered diffraction data, *J. Microsc.*, 201(Pt 1), pp. 50–58.

KHATRI, C. G. & MARDIA, K. V. (1977). The von Mises-Fisher matrix distribution in orientation statistics, *J. Roy. Stat. Soc.*, Ser B, 39(1), pp. 95–106.

LANDRY, J. S. & BIDEN, E. N. (2002). Optimal fixed wrist alignment for below-elbow, powered, prosthetic hands, *MEC'2002 Conference Proceedings* (Inst. of Biomedical Engineering, Univ. of New Brunswick, Fredericton, NB, Canada, August 2002) pp. 12–14.

KRIEGER LASSEN, N. C., JUUL Jensen, D. & CONRADSEN, K. (1994) On the statistical analysis of orientation data, *Acta Crystallogr.* A50, pp. 741–748.

PALMER, A. K., WERNER, F. W., MURPHY, D. & GLISSON, R. (1985). Functional wrist motion: A biomechanical study, *J. Hand. Surg. [Am.]* 10A(1), pp. 39–46.

PRENTICE, M. J. (1986). Orientation statistics without parametric assumptions, *J. Roy. Stat. Soc.* Ser B. 48(2), pp. 214–222.

RANCOURT, D., RIVEST, L.-P. & ASSELIN, J. (2000). Using orientation statistics to investigate variations in human kinematics, *Appl. Statist.*, 49, pp. 81–94.

RUY, J., COONEY, W. P., ASKEW, L. J., AN, K.-N. & CHAO, E. Y. S. (1991). Functional ranges of motion of the wrist joint, *J. Hand. Surg. [Am.]* 16A(3), pp. 409–419.

SARRAFIAN, S. K., MELAMED, J. L. & GOSHGARIAN, G. M. (1977). Study of wrist motion in flexion and extension, *Clin. Orthop*, 126, pp. 153–159.

SHEEHAN, F. T. & MITIGUY, P. (1999). In regards to the 'ISB recommendations for standardization in the reporting of kinematic data', *J. Biomech.*, 32(10), pp. 1135–1136 (Letter to the Editor).

SPONG, M. W. & VIDYASAGAR, M. (1989). *Robot Dynamics and Control* (John Wiley & Sons, 1989).

STAVDAHL, O. (2002) Optimal wrist prosthesis kinematics: threedimensional rotation statistics and parameter estimation, *Ph.D. thesis* (Norwegian University of Science and Technology, Trondheim, Norway, 2002. ISBN 82-471-5526-5).

STEPHENS, M. A. (1979). Vector correlation, *Biometrika*, 66(1), pp. 41–48.

VELDPAUS, F. E., WOLTRING, H. J. & DORTMANS, L. J. M. G. (1988). A least-squares algorithm for the equiform transformation from spinal marker co-ordinates, *J. Biomech.*, 21(1), pp. 45–54.

WOLTRING, H. J., HUISKES, R. & DE LANGE, A. (1985). Finite centroid and helical axis estimation from noisy landmarkmeasurements in the study of human joint kinematics, *J. Biomech.*, 18(5), pp. 379 389.

WOLTRING, H. J. (1994). 3-D attitude representation of human joints: a standardization proposal, *J. Biomech.*, 27(12), pp. 1399–1414.

WU, G. & CAVANAGH, P. R. (1995). ISB recommendations for standardization in the reporting of kinematic data, *J. Biomech.*, 28, pp. 1258–1261.

YOUM, Y., MCMURTRY, R. Y., FLATT, A. E. & GILLESPIE, T. E. (1978). Kinematics of the wrist. (I). An experimental study of radioulnar deviation and flexion-extension, *J. Bone Joint Surg.*, 60-A(4), pp. 423–431.