

Manifold Statistics for Essential Matrices

Gijs Dubbelman¹, Leo Dorst², and Henk Pijls²

¹Carnegie Mellon University, The Robotics Institute, Pittsburgh, USA
gijsdubb@cs.cmu.edu

²University of Amsterdam, Faculty of Science, Amsterdam, The Netherlands
{L.Dorst, H.G.J.Pijls}@uva.nl

Abstract. Riemannian geometry allows for the generalization of statistics designed for Euclidean vector spaces to Riemannian manifolds. It has recently gained popularity within computer vision as many relevant parameter spaces have such a Riemannian manifold structure. Approaches which exploit this have been shown to exhibit improved efficiency and accuracy. The Riemannian logarithmic and exponential mappings are at the core of these approaches.

In this contribution we review recently proposed Riemannian mappings for essential matrices and prove that they lead to sub-optimal manifold statistics. We introduce correct Riemannian mappings by utilizing a multiple-geodesic approach and show experimentally that they provide optimal statistics.

1 Introduction

Statistical inference on a set of samples is at the basis of many computer vision algorithms. These algorithms typically require a metric to be expressed in the parameter space in which the samples reside. The most often used metrics are the Euclidean distance or one of its generalizations such as the Mahalanobis distance. All are based on the usual dot product between vectors and assume that the parameter space is a Euclidean vector space. Recent work in computer vision has pointed out that many parameter spaces are *differentiable manifolds*, i.e. they are differentiable subsets of Euclidean spaces. Neglecting the manifold structure of parameter spaces within statistics typically leads to reduced accuracy and efficiency, as one is optimizing over, and measuring distances through, irrelevant degrees of freedom.

Improved accuracy and efficiency can be obtained by restricting computations to the meaningful degrees of freedom, i.e. those of the manifold. This can be achieved by exploiting the paradigm of *manifold statistics* based on *metric charts*. A metric chart is a local linearization (or flattening) of the manifold for one particular point on the manifold, called the *charting point*, such that distances over the manifold with respect to this point can be computed using the Euclidean distance formula in this chart. Algorithms using metric charts to compute statistical properties of samples residing on manifolds are also referred to as *intrinsic statistical* algorithms and usually are iterative in nature. Crucial to this approach are *charting functions* which define the metric charts for each point on a particular manifold. They can be used to transfer samples residing in the manifold to the metric charts and back. Riemannian geometry provides the theory and tools to develop such charting functions (where they are called logarithmic and exponential mappings) and is discussed in Sec. 2.

The use of Riemannian geometry to disclose statistical information on manifolds can be traced back to [1]. This methodology has recently received considerable attention from the computer vision research community. For a theoretical overview see [2, 3] and the references therein. An example of a statistical algorithm that was generalized to certain manifolds is mean-shift [4, 5]. In [6] it was used for robust pose estimation, in [7] for robust essential matrix estimation and in [8] for simultaneous multiple motion estimation. In [9] a Riemannian approach was taken to estimate the dimension and entropy of shape spaces, particularly the shape space of handwritten digits. Riemannian geometry has also been used to estimate statistical properties of diffusion tensor data [10, 11]. In [12] Riemannian geometry was used to estimate properties from image point configurations. A Riemannian clustering approach was proposed in [13] and applied to 2D motion segmentation and to diffusion tensor segmentation. In [14] a Riemannian clustering approach was proposed for pedestrian detection from image data.

This work focuses on *epipolar configurations*, i.e. *scale-free* Euclidean motions, which are encoded by *essential matrices*. In Sec. 3 we develop our charting functions for essential matrices by using methods from Riemannian geometry. They are inspired by mappings proposed earlier in [15–17]. As is explained in Sec. 2 and in Sec. 3 the mappings in [15–17] are designed for manifold optimization (estimating essential matrices from image point correspondences) and cannot be used *as is* within manifold statistics. This was already pointed out in [5, 7] which introduced alternative mappings. In Sec. 3.2 it is proven that the mappings in [5, 7] are, strictly speaking, mathematically incorrect and in Sec. 4 it is shown that they result in sub-optimal statistics. This stands in contrast to our mappings for which it is shown that they lead to optimal statistics. Although our focus is mostly theoretical, the broader relevance of our findings is demonstrated by showing that *state-of-the-art* robust estimators based on the RANSAC [18–20] paradigm can potentially benefit from incorporating our mappings. Our conclusion are summarized in Sec. 5.

2 Riemannian geometry and manifold statistics

Manifold statistics conceptually requires a distance metric for every point pair on the manifold. In Riemannian geometry the distance $d(\mathbf{p}, \mathbf{u})$ between two points \mathbf{p} and \mathbf{u} residing on a differentiable manifold \mathcal{M} is defined as the length of the shortest curve over the manifold connecting the two points, i.e.

$$d(\mathbf{p}, \mathbf{u}) = \min_{\gamma \in \Gamma} l_{\mathbf{p}}^{\mathbf{u}}(\gamma), \quad (1)$$

where Γ is the set of all possible curves residing on the manifold joining \mathbf{p} and \mathbf{u} and $l_{\mathbf{p}}^{\mathbf{u}}(\gamma)$ is the length of one such curve between these two points. Under this definition $d(\cdot, \cdot)$ is a metric [21] and obeys to all its axioms, i.e. *positive definiteness*, *identity of indiscernibles*, *symmetry* and the *triangle inequality*. Curves which provided paths of least distance over manifolds are often called *geodesics*.

Just as in Euclidean spaces, the length of a curve γ in \mathcal{M} is the integral of all the local contributions of all points on the curve. Let the curve γ pass through the point \mathbf{p} at $\gamma(0)$ and pass through the point \mathbf{u} at $\gamma(1)$, then the length of the curve segment between

\mathbf{p} and \mathbf{u} is defined as

$$l_{\mathbf{p}}^{\mathbf{u}}(\gamma) = \int_0^1 \sqrt{\left\langle \frac{d\gamma}{dt}, \frac{d\gamma}{dt} \right\rangle_{\gamma(t)}} dt. \quad (2)$$

Each local contribution $\left\langle \frac{d\gamma}{dt}, \frac{d\gamma}{dt} \right\rangle_{\gamma(t)}$ is computed by the *Riemannian metric* $\langle \cdot, \cdot \rangle$ which is an inner product defined on *tangent vectors* of \mathcal{M} . A differentiable manifold equipped with a Riemannian metric is called a *Riemannian manifold*. As the Riemannian metric can vary over the manifold, the subscript $\gamma(t)$ denotes the point on the curve at which the Riemannian metric is computed and $\frac{d\gamma}{dt}$ is the tangent vector of the curve at this point. Note that the locally defined Riemannian metric should not be confused with the globally defined Riemannian distance $d(\mathbf{p}, \mathbf{u})$.

Fortunately, we need not work with these equations as for certain manifolds, including that of essential matrices, the globally defined metric $d(\mathbf{p}, \mathbf{u})$ can be computed efficiently through a construct known as the *Riemannian logarithmic map*. The Riemannian logarithmic map $\text{Log}_{\mathbf{p}}(\mathbf{u}) = \mathbf{u}$ takes the point \mathbf{u} on the manifold to the *tangent space* at \mathbf{p} , i.e. $T_{\mathbf{p}}\mathcal{M}$, producing the tangent vector \mathbf{u} . $T_{\mathbf{p}}\mathcal{M}$ is the collection of all tangent vectors of the manifold at \mathbf{p} , we denote its elements in gothic font. The Riemannian logarithmic map does so such that

$$d(\mathbf{p}, \mathbf{u}) = \sqrt{\text{Log}_{\mathbf{p}}(\mathbf{u})^{\top} \text{Log}_{\mathbf{p}}(\mathbf{u})} = \|\mathbf{u}\|, \quad (3)$$

i.e. the distance over the manifold can be computed as the Euclidean length of the tangent vector \mathbf{u} produced by $\text{Log}_{\mathbf{p}}(\mathbf{u})$. It is clear that the tangent spaces constructed by the logarithmic map are the metric charts of Sec. 1. Note that only for specific manifolds Eq. 3 holds globally for all \mathbf{u} , given \mathbf{p} . In general, it only holds for \mathbf{u} sufficiently close to \mathbf{p} on the manifold.

The Riemannian logarithmic map is differentiable and invertible. Its inverse, the *Riemannian exponential map*, takes the tangent vector \mathbf{u} defined in the tangent space (or metric chart) at \mathbf{p} back to the manifold such that $\text{Log}_{\text{Exp}_{\mathbf{p}}(\mathbf{u})}(\mathbf{p}) = -\mathbf{u}$.

By using the direct product between Riemannian manifolds one can combine existing manifolds into new manifolds. Under the direct product points on the manifold $\mathcal{M} = \mathcal{M}_1 \times \mathcal{M}_2$ are ordered pairs of points on \mathcal{M}_1 and points on \mathcal{M}_2 , i.e. $\mathbf{p} = (\mathbf{p}_1, \mathbf{p}_2)$ with $\mathbf{p}_1 \in \mathcal{M}_1$ and $\mathbf{p}_2 \in \mathcal{M}_2$. The exponential and logarithmic mappings of the direct product manifold are simply the combination of the individual mappings of each sub-manifold, i.e.

$$\begin{aligned} \text{Log}_{\mathbf{p}}(\mathbf{u}) &= (\text{Log}_{\mathbf{p}_1}(\mathbf{u}_1), \text{Log}_{\mathbf{p}_2}(\mathbf{u}_2)) \quad \text{and} \\ \text{Exp}_{\mathbf{p}}(\mathbf{u}) &= (\text{Exp}_{\mathbf{p}_1}(\mathbf{u}_1), \text{Exp}_{\mathbf{p}_2}(\mathbf{u}_2)). \end{aligned} \quad (4)$$

Every property of the sub-manifolds \mathcal{M}_1 and \mathcal{M}_2 are automatically transferred to the direct product manifold \mathcal{M} . The intrinsic distance between \mathbf{p} and \mathbf{u} over the direct product manifold \mathcal{M} is defined as

$$\begin{aligned} d(\mathbf{p}, \mathbf{u}) &= \sqrt{d(\mathbf{p}_1, \mathbf{u}_1)^2 + d(\mathbf{p}_2, \mathbf{u}_2)^2} \\ &= \sqrt{\|\text{Log}_{\mathbf{p}_1}(\mathbf{u}_1)\|^2 + \|\text{Log}_{\mathbf{p}_2}(\mathbf{u}_2)\|^2}, \end{aligned} \quad (5)$$

it is the square root of the sum of the squared lengths of the individual minimal length curves. Such a definition is called a *multiple geodesics* or a *product metric* approach [21, 22] and automatically obeys the axioms for metrics.

When the mappings of Eq. 4 are available for all combinations of \mathbf{p} and \mathbf{u} of interest to the statistical algorithm, the Riemannian logarithmic map can be used to generalize statistical algorithms based on the Euclidean distance formula. This generalization is performed by applying the substitution

$$(\mathbf{u} - \mathbf{p})^\top \Sigma^{-1} (\mathbf{u} - \mathbf{p}) \rightarrow \text{Log}_{\mathbf{p}}(\mathbf{u})^\top \Sigma^{-1} \text{Log}_{\mathbf{p}}(\mathbf{u}), \quad (6)$$

to a statistical objective function. It assures that all statistical calculations are performed in the metric charts of the manifold and therefore are constrained to the actual degrees of freedom of this manifold. The manifold statistical objective functions in [2–7] are all expressed similarly in terms of a logarithmic map, whereas the manifold optimization methods in [15–17] only require the exponential map. The required logarithmic mappings for essential matrices are also not provided in [15–17].

3 Riemann mappings for Essential Matrices

The essential matrix \mathbf{E} [23] is a 3×3 matrix which algebraically describes the *epipolar geometry* between two calibrated cameras. Let \mathbf{x} and \mathbf{x}' be the homogeneous normalized projections of the same world point on the imaging planes of the first and the second camera respectively. The essential matrix is then defined algebraically as the matrix that satisfies

$$\mathbf{x}'^\top \mathbf{E} \mathbf{x} = 0 \quad (7)$$

for all such correspondences \mathbf{x} and \mathbf{x}' .

The epipolar geometry of a normalized camera setup is defined by the translation \mathbf{t} and rotation \mathbf{R} between the two cameras. Given this translation and rotation a geometric composition of the essential matrix is $\mathbf{E} = [\mathbf{t}]_{\times} \mathbf{R}$, realizing that it can only be determined up to a global scale ambiguity from corresponding image points. The nine values in \mathbf{E} have five degrees of freedom, 3 for rotation and 2 for translation modulo global scale. Due to this scale ambiguity the length of the translation vector can be normalized to have unit length $\|\mathbf{t}\| = 1$ without loss of information. The normalized essential matrix is therefore defined as

$$\mathbf{E} = [\mathbf{t}]_{\times} \mathbf{R}, \quad \text{with } \|\mathbf{t}\| = 1. \quad (8)$$

This definition allows us to specify the parameter space of normalized essential matrices as the direct product space $S^2 \times SO(3)$ [16, 17]. Both S^2 and $SO(3)$ are well known Riemannian manifolds and so is their direct product space $S^2 \times SO(3)$. We call points on $S^2 \times SO(3)$ *epipolar configurations*. The challenge of this parametrization is that $S^2 \times SO(3)$ is a four-fold covering of the space of essential matrices, every essential matrix maps to four epipolar configurations. This is a well known ambiguity which is of less concern to manifold optimization than it is to manifold statistics.

Manifold optimization methods like those in [15–17] are only concerned with estimating a single essential matrix at a time. For any particular essential matrix one can

randomly pick any of its four epipolar configurations and start optimizing from there. The challenge is to parametrize $S^2 \times SO(3)$ (at each iteration of the optimizer) such that it cannot jump between epipolar configurations which all are related to the same essential matrix and hence all have the same *summed squared error* (sse).

In manifold statistics we typically have an empirical distribution consisting of several hundred essential matrices. When each of them is randomly mapped to one of its four epipolar configurations, almost all statistical information of the distribution is lost. This is demonstrated in Sec. 4.1. One creates a situation where the distance between two essential matrices depends purely on a random choice and is no longer related to the epipolar configurations that they encode. The challenge for manifold statistics is therefore to make a consistent choice when mapping each essential matrix to a point on $S^2 \times SO(3)$. As pointed out in [5, 7] the mappings in [15–17] do themselves not provide mechanism to perform this mapping consistently for multiple essential matrices as is required.

3.1 The normalized essential manifold $S^2 \times SO(3)$

We start with the normalized essential manifold $S^2 \times SO(3)$ of [16, 17]. In this direct product space the sphere S^2 is related to normalized translations, which we will call *directions*, and $SO(3)$ is related to rotations in 3D. When considering the *unit quaternion* representation of rotations, then it is clear that rotational sub-space $SO(3)$ has the shape of a unit hyper-sphere S^3 .

Each essential matrix must be taken to a unique point on this manifold $S^2 \times S^3$ in a differentiable and invertible fashion. This requires resolving the *four-fold ambiguity* and can be performed efficiently by enforcing the *positive depth constraint* [24]. In contrast to the remarks in [5, 7] we found that this can be done efficiently and robustly, we come back to this in Sec. 4. The image data and the *positive depth constraint* together with Eq. 8 provide an invertible and differentiable mapping between essential matrices and points in $S^2 \times S^3$. Conceptually speaking, the knowledge that the observed image data is a projection of a Euclidean space allows for ‘lifting’ the essential matrix from a projective entity to an epipolar configuration. The ability to do so is the main difference between a computer vision approach, where image data is always available, and a pure mathematical approach like [15].

Each essential matrix \mathbf{E} can be represented uniquely by a seven dimensional vector $\mathbf{E} \equiv (\mathbf{d}, \mathbf{q}) = (d_x, d_y, d_z, q, q_i, q_j, q_k)^\top$ with $\mathbf{d} \in S^2$ (modeling directions) and $\mathbf{q} \in S^3$ (modeling rotations by unit quaternions). Due to the fact that $\|\mathbf{d}\| = 1$ and $\|\mathbf{q}\| = 1$ there are only five degrees of freedom as desired. Using a multiple geodesic approach on $S^2 \times S^3$ results in the Riemannian logarithmic and exponential mappings

$$\begin{aligned} \mathbf{e}_2 &= \text{Log}_{\mathbf{E}_1}(\mathbf{E}_2) = (\text{Log}_{\mathbf{d}_1}(\mathbf{d}_2), \text{Log}_{\mathbf{q}_1}(\mathbf{q}_2)) \quad \text{and} \\ \mathbf{E}_2 &= \text{Exp}_{\mathbf{E}_1}(\mathbf{e}_2) = (\text{Exp}_{\mathbf{d}_1}(\mathbf{d}_2), \text{Exp}_{\mathbf{q}_1}(\mathbf{q}_2)). \end{aligned} \tag{9}$$

The Riemannian distance between two essential matrices can then be computed with $\|\text{Log}_{\mathbf{E}_1}(\mathbf{E}_2)\|$. These mappings are all that is required to generalize statistical algorithms according to the substitution of Eq. 6. Our contribution is providing explicit and efficient functions for these mappings.

For S^3 we choose the quaternion identity $(1, 0, 0, 0)^\top$ as the origin. For this choice the Riemannian exponential map wrt. the origin is

$$\mathbf{q} = \text{Exp}(\mathbf{q}) = \begin{cases} (\cos(\|\mathbf{q}\|), \sin(\|\mathbf{q}\|) \frac{\mathbf{q}}{\|\mathbf{q}\|}), & \|\mathbf{q}\| \neq 0 \\ (1, 0, 0, 0) & , \|\mathbf{q}\| = 0 \end{cases} . \quad (10)$$

For S^2 we take $(1, 0, 0)^\top$ as the origin which gives

$$\mathbf{d} = \text{Exp}(\mathfrak{d}) \equiv \begin{cases} (\cos(\|\mathfrak{d}\|), \sin(\|\mathfrak{d}\|) \frac{\mathfrak{d}}{\|\mathfrak{d}\|}), & \|\mathfrak{d}\| \neq 0 \\ (1, 0, 0) & , \|\mathfrak{d}\| = 0 \end{cases} . \quad (11)$$

These mappings are well known [25] and very similar to that of [16, 17] except that we choose to use S^3 (unit quaternions) instead of $SO(3)$ (rotation matrices).

In contrast to [15–17] we also require the logarithmic mappings of S^3 and S^2 . At the origin of S^3 with $\mathbf{q} = (q, q_i, q_j, q_k)^\top$ it is provided by

$$\mathbf{q} = \text{Log}(\mathbf{q}) = \begin{cases} \frac{\arccos(q)}{\|(q_i, q_j, q_k)\|} (q_i, q_j, q_k), & q \neq 1 \\ (0, 0, 0) & , q = 1 \end{cases} \quad \text{with } \underline{\arccos}(q) = \begin{cases} \arccos(q) - \pi, & q < 0 \\ \arccos(q) & , 0 \leq q \end{cases} \quad (12)$$

and at the origin of S^2 with $\mathbf{d} = (d_x, d_y, d_z)^\top$ it is

$$\mathfrak{d} = \text{Log}(\mathbf{d}) \equiv \begin{cases} \arccos(d_x) \frac{(d_y, d_z)}{\|(d_y, d_z)\|}, & d_x \neq 1 \\ (0, 0) & , d_x = 1 \end{cases} . \quad (13)$$

Note that arccos assures that antipodal quaternions \mathbf{q} and $-\mathbf{q}$ are mapped to the same point in the tangent space such that $d(\mathbf{q}, -\mathbf{q}) = 0$. This is required as S^3 is a double cover of $SO(3)$, i.e. \mathbf{q} and $-\mathbf{q}$ both represent the same rotation. Apart from arccos, these mappings are again well known [25].

The next step is to express these mappings wrt. general points on the manifolds S^3 and S^2 and not only wrt. to their origins. For S^3 this can be performed by exploiting the Lie group structure of unit quaternions [3]. The general logarithmic and exponential mappings for S^3 are defined by:

$$\begin{aligned} \mathbf{q}_2 &= \text{Log}_{\mathbf{q}_1}(\mathbf{q}_2) = \text{Log}(\mathbf{q}_1^{-1} \star \mathbf{q}_2) \quad \text{and} \\ \mathbf{q}_2 &= \text{Exp}_{\mathbf{q}_1}(\mathbf{q}_2) = \mathbf{q}_1 \star \text{Exp}(\mathbf{q}_2). \end{aligned} \quad (14)$$

The special case S^2 : It is well known that S^2 does not allow for a Lie group structure [26] which makes obtaining the general mappings significantly more challenging. Correct manifold statistics requires that all tangent space bases of all points on S^2 , except that of the antipodal of $(1, 0, 0)^\top$ which is $(-1, 0, 0)^\top$, are smoothly aligned with the tangent space basis at $(1, 0, 0)^\top$ such that minimal length curves originating from $(1, 0, 0)^\top$ are geodesics (i.e. it requires a torsion-free metric alignment related to the *Levi-Civita* connection [21] of S^2). This is because manifold statistics must simultaneously handle multiple empirical distributions on S^2 consisting out of multiple points and therefore requires a globally consistent approach to relate all tangent space bases and all minimal length curves of all these points to each other.

In manifold optimization this is not required as it is concerned with the exponential map of one point at a time. As such, [16, 17] can use a Gram-Schmidt process to define the bases of tangent spaces at each iteration. Their approach does not provide the required alignment of tangent space bases, does not provide a general logarithmic map such that $\text{Log}_{\text{Exp}_{\mathbf{d}_1}(\mathfrak{d}_2)}(-\mathfrak{d}_2) = \mathbf{d}_1$ with $d(\mathbf{d}_1, \mathbf{d}_2) = \|\text{Log}_{\mathbf{d}_1}(\mathbf{d}_2)\|$ and therefore cannot be used *as is* for manifold statistics. The same holds for the approach in [15].

Our main theoretical contribution is that we provide a product structure on S^2 such that its general Riemannian mappings can be expressed as

$$\begin{aligned} \mathfrak{d}_2 &= \text{Log}_{\mathbf{d}_1}(\mathbf{d}_2) = \text{Log}(\mathbf{d}_1^{-1} \star \mathbf{d}_2) \quad \text{and} \\ \mathbf{d}_2 &= \text{Exp}_{\mathbf{d}_1}(\mathfrak{d}_2) = \mathbf{d}_1 \star \text{Exp}(\mathfrak{d}_2). \end{aligned} \quad (15)$$

Our product structure coincides with the torsion-free metric alignment of S^2 at \mathbf{e} , see Fig. 1. Similarly to the quaternion product, its purpose is to provide an invertible map between geodesics originating from general points on S^2 to geodesics originating at \mathbf{e} by *parallel transport* [21]. The existence of such a product structure is often contributed to a Lie group structure which S^2 does not allow [26]. Still we show that S^2 has many useful *Lie-group-like* properties similar to that of the Lie group of unit quaternions. Our product structure for S^2 is provided in Theorem 1.

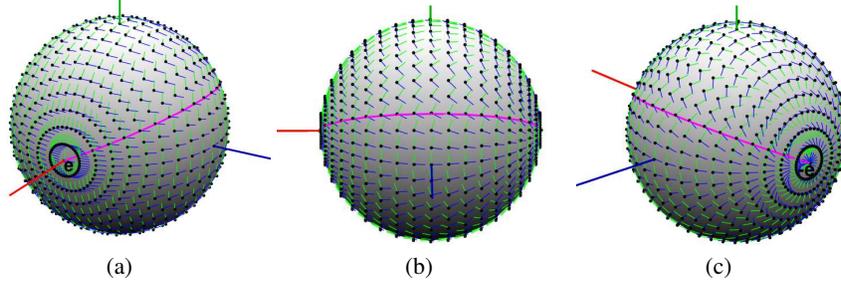


Fig. 1. The alignment of tangent space bases on S^2 enforced by our product with $\mathbf{e} = (1, 0, 0)^\top$. Only at $-\mathbf{e} = (-1, 0, 0)$ it has a discontinuity for all other points it is smooth. A minimal length curve through \mathbf{e} is plotted in magenta. Note that it does not change direction with respect to the tangent space bases of all points except that of $-\mathbf{e}$, i.e. the antipodal and *cut locus* of \mathbf{e} . Therefore, it is a geodesic everywhere except at $-\mathbf{e}$. The same holds for any other geodesic starting at \mathbf{e}

Theorem 1: The product \star between general points $\mathbf{d}_1 \in S^2$ and $\mathbf{d}_2 \in S^2$ is

$$\star : S^2 \times S^2 \rightarrow S^2, \quad \mathbf{d}_1 \star \mathbf{d}_2 = R_{\mathbf{d}_1} R_{\mathbf{d}_2} (1, 0, 0)^\top \quad (16)$$

where $R_{\mathbf{d}} = \mathbf{I}$ when $\mathbf{d} = (1, 0, 0)^\top$ and $R_{\mathbf{d}} = -\mathbf{I}$ when $\mathbf{d} = (-1, 0, 0)^\top$ and for all other $\mathbf{d} \in S^2$, $\mathbf{d} \notin \{(1, 0, 0)^\top, (-1, 0, 0)^\top\}$,

$$R_{\mathbf{d}} = \mathbf{I} + \sin(\theta_{\mathbf{d}}) [\mathbf{r}_{\mathbf{d}}]_{\times} + (1 - \cos(\theta_{\mathbf{d}})) [\mathbf{r}_{\mathbf{d}}]_{\times}^2 \quad (17)$$

with

$$\mathbf{r}_d = \frac{(0, -d_z, d_y)^\top}{\sqrt{d_y^2 + d_z^2}} \quad \text{and} \quad \theta_d = \arccos(d_x). \quad (18)$$

The identity element \mathbf{e} belonging to this product is $(1, 0, 0)^\top$ and inversion of an element \mathbf{d} is provided by $\mathbf{d}^{-1} = \mathbf{R}_d^{-1}\mathbf{e}$. We define the product to be *left associative*, i.e. $\mathbf{d}_1 \star \mathbf{d}_2 \star \mathbf{d}_3 = (\mathbf{d}_1 \star \mathbf{d}_2) \star \mathbf{d}_3$.

Proof: First note that we defined \mathbf{R}_d such that $\mathbf{d} = \mathbf{R}_d\mathbf{e}$ for all $\mathbf{d} \in S^2$. The axiom of *closure* is satisfied as both rotation matrices and $\pm\mathbf{I}$ map points on S^2 onto S^2 . The axiom of *identity* follows from $\mathbf{d} \star \mathbf{e} = \mathbf{R}_d\mathbf{Ie} = \mathbf{R}_d\mathbf{e} = \mathbf{d}$ and $\mathbf{e} \star \mathbf{d} = \mathbf{I}\mathbf{R}_d\mathbf{e} = \mathbf{d}$. The axiom of *invertibility* is guaranteed by $\mathbf{d} \star \mathbf{d}^{-1} = \mathbf{R}_d\mathbf{R}_d^{-1}\mathbf{e} = \mathbf{Ie} = \mathbf{e}$ and $\mathbf{d}^{-1} \star \mathbf{d} = \mathbf{R}_d^{-1}\mathbf{R}_d\mathbf{e} = \mathbf{Ie} = \mathbf{e}$. \square

The profound relation between minimal length curves starting at \mathbf{e} , the exponential map at \mathbf{e} and our product structure is best illustrated by Theorem 2.

Theorem 2: The product structure of Theorem 1 is related to the exponential map of S^2 at \mathbf{e} , which traces out minimal length curves with $\text{Exp}(t\mathfrak{d})$, by the Taylor series of the exponential function, i.e.

$$e^{\mathfrak{d}} \equiv \sum_{n=0}^{\infty} \frac{\mathfrak{d}^n}{n!} = (\cos(\|\mathfrak{d}\|), \sin(\|\mathfrak{d}\|) \frac{\mathfrak{d}}{\|\mathfrak{d}\|}) \quad (19)$$

where $\mathfrak{d} \in T_e\mathcal{M}$ and $0 < \|\mathfrak{d}\|$.

Proof: In order to evaluate \mathfrak{d}^n , with $n \in \{0, \mathbb{Z}^+\}$, the product structure must be expressed on tangent vectors in $\mathfrak{d} \in T_e\mathcal{M}$ for which $0 < \|\mathfrak{d}\|$. We can do so by defining the map $\mathcal{S} : T_e\mathcal{M} \rightarrow S^2$, $\mathcal{S}(\mathfrak{d}) = (0, \frac{\mathfrak{d}}{\|\mathfrak{d}\|})$. For $n = 0$ the power is then defined by $\mathfrak{d}^0 = \mathbf{e}$ and for $0 < n$ it is defined by $\mathfrak{d}^n = \|\mathfrak{d}\|(\mathfrak{d}^{n-1} \star \mathcal{S}(\mathfrak{d}))$. The powers of order 1 and higher evaluate to

$$\begin{aligned} \mathfrak{d}^1 &= \|\mathfrak{d}\|(0, \frac{\mathfrak{d}}{\|\mathfrak{d}\|}), & \mathfrak{d}^2 &= -\|\mathfrak{d}\|^2\mathbf{e}, & \mathfrak{d}^3 &= -\|\mathfrak{d}\|^3(0, \frac{\mathfrak{d}}{\|\mathfrak{d}\|}), & \mathfrak{d}^4 &= \|\mathfrak{d}\|^4\mathbf{e}, \\ \mathfrak{d}^5 &= \|\mathfrak{d}\|^5(0, \frac{\mathfrak{d}}{\|\mathfrak{d}\|}), & \mathfrak{d}^6 &= -\|\mathfrak{d}\|^6\mathbf{e}, & \mathfrak{d}^7 &= -\|\mathfrak{d}\|^7(0, \frac{\mathfrak{d}}{\|\mathfrak{d}\|}), & & \dots \end{aligned} \quad (20)$$

In geometric terms, every power \mathfrak{d}^n rotates $(0, \frac{\mathfrak{d}}{\|\mathfrak{d}\|})$ over its geodesic, which starts at \mathbf{e} and goes through $(0, \frac{\mathfrak{d}}{\|\mathfrak{d}\|})$, with angle $\frac{(n-1)}{2}\pi$. From the equalities in Eq. 20 it follows that

$$\begin{aligned} e^{\mathfrak{d}} &\equiv \sum_{n=0}^{\infty} \frac{\mathfrak{d}^n}{n!} \\ &= \mathbf{e} + \mathfrak{d} - \frac{\|\mathfrak{d}\|^2}{2!}\mathbf{e} - \frac{\|\mathfrak{d}\|^3}{3!}(0, \frac{\mathfrak{d}}{\|\mathfrak{d}\|}) + \frac{\|\mathfrak{d}\|^4}{4!}\mathbf{e} + \frac{\|\mathfrak{d}\|^5}{5!}(0, \frac{\mathfrak{d}}{\|\mathfrak{d}\|}) - \frac{\|\mathfrak{d}\|^6}{6!}\mathbf{e} - \dots \\ &= (1 - \frac{\|\mathfrak{d}\|^2}{2!} + \frac{\|\mathfrak{d}\|^4}{4!} - \frac{\|\mathfrak{d}\|^6}{6!} + \dots)\mathbf{e} + (\|\mathfrak{d}\| - \frac{\|\mathfrak{d}\|^3}{3!} + \frac{\|\mathfrak{d}\|^5}{5!} - \dots)(0, \frac{\mathfrak{d}}{\|\mathfrak{d}\|}) \quad (21) \\ &= \cos(\|\mathfrak{d}\|)\mathbf{e} + \sin\|\mathfrak{d}\|(0, \frac{\mathfrak{d}}{\|\mathfrak{d}\|}) \\ &= (\cos(\|\mathfrak{d}\|), \sin(\|\mathfrak{d}\|) \frac{\mathfrak{d}}{\|\mathfrak{d}\|}). \quad \square \end{aligned}$$

Similar proofs are well known for the unit circle and for unit quaternions and many other useful Lie groups [27]. We have extended this to include the *non-Lie-group* S^2 . The deeper mathematical foundations (i.e. the Levi-Civita connection [21] of S^2) related to our product on S^2 are beyond the scope of this text. Our purpose here is to prove that such a product exists and that it has useful applications to manifold statistics.

3.2 Alternative approach

Alternative Riemannian mappings for essential matrices were proposed earlier in [5, 7] within the context of manifold statistics. In this section we prove that their Riemannian mappings are incorrect as they are ambiguous, i.e. they do not map essential matrices which represent the same epipolar configuration to a unique point on the manifold.

Performing a singular value decomposition on an essential matrix results in $E = U\Sigma V^T$. The first two singular values of Σ are identical elements of \mathbb{R}^+ and the last one is zero [24]. Since an essential matrix is a projective entity, it is not affected by scaling, hence Σ can be set to $\Sigma_1 = \text{diag}\{1, 1, 0\}$. When $\det(U) = \det(V) = 1$ the four possible epipolar configurations on $S^2 \times \text{SO}(3)$ belonging to the essential matrix $E = U\Sigma V^T$ are $(+\mathbf{u}_3, UWV^T)$, $(+\mathbf{u}_3, UW^T V^T)$, $(-\mathbf{u}_3, UWV^T)$ and $(-\mathbf{u}_3, UW^T V^T)$ where $W = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ and \mathbf{u}_3 is the last column of U [24]. As pointed out in [24] performing a rotation around the z-axis by an angle γ , denoted as $R_z(\gamma)$, does not change the essential matrix. Furthermore, due to its projective nature the essential matrices E and $-E$ both relate to the same epipolar configurations. This equivalence relation is denoted as $E =_{\text{H}} -E$. For the SVD of E we thus have the equality

$$U R_z(\gamma) \Sigma_1 (V R_z(\gamma + k\pi))^T = (-1)^k U \Sigma_1 V^T =_{\text{H}} \pm E, \quad (22)$$

with $k \in \mathbb{Z}$. This shows that there is a more general equivalence class of essential matrices, denoted as \mathcal{R}_z , which all encode the same epipolar configuration. The claim in [5, 7] is that they can map all elements of \mathcal{R}_z to a single point on a Riemannian manifold without using image data or the positive-depth constraint. We now prove that their claim is incorrect pointing out an unavoidable ambiguity in their mappings.

Consider that when computing the SVD there is an arbitrariness in the choice for the basis vectors being used. For a 3×3 matrix there are 8 such choices. The choice being made often depends on considerations driven by numerical stability but it could just as well be random. This arbitrary choice results in the following equality

$$E = UB\Sigma(V(\pm B))^T = \pm U\Sigma V^T =_{\text{H}} \pm E \quad (23)$$

where $B = \text{diag}\{\pm 1, \pm 1, \pm 1\}$. In [24] it is implicitly assumed that $\det(U) = \det(V) = 1$ i.e. U and V are both elements of $\text{SO}(3)$. Due to B both U and V are merely orthogonal transformations, i.e. they are combinations of rotations and reflections. Since the last singular value of Σ is zero, the last column of both U and V can be multiplied by -1 such that their determinants become positive without changing the essential matrix. This ensures that U and V become elements of $\text{SO}(3)$ and ensures that retrieving the correct epipolar configuration by using the positive-depth constraint is unaffected by B . This is not true for the approach in [5, 7].

The task in [5, 7] is to take the essential matrix E_2 (with SVD $U_2 \text{diag}\{1, 1, 0\} V_2^\top$), to a unique point in the tangent space of E_1 (with SVD $U_1 \text{diag}\{1, 1, 0\} V_1^\top$). The general form of their logarithmic map is

$$\text{Log}_{E_1}(E_2) = (\text{Log}(U_1^\top U_2), \text{Log}(V_1^\top V_2)) = (\mathbf{u}, \mathbf{v}) \quad (24)$$

where Log on the right side of the equality denotes the Riemannian logarithmic map on $\text{SO}(3)$ and $(\mathbf{u}, \mathbf{v}) \in \mathfrak{so}(3) \times \mathfrak{so}(3)$. The exponential map of the SM method is defined as

$$\text{Exp}_{E_1}((\mathbf{u}, \mathbf{v})) = (U_1 \text{Exp}(\mathbf{u}), V_1 \text{Exp}(\mathbf{v})) \quad (25)$$

where Exp on the right side of the equality denotes the Riemannian exponential map on $\mathfrak{so}(3)$. This is conceptually similar to a multiple geodesic approach of two copies of $\text{SO}(3)$. The challenge is that both E_1 and E_2 are general elements of their own equivalence classes but the logarithmic map must return the same tangent vector for all combinations consistently.

In [5, 7] the equivalence class \mathcal{R}_z is resolved by enforcing $\text{log}_{E_1}(E_2)$ to have the form $[u_x, u_y, z, v_x, v_y, -z]^\top$. It does so by an initialization procedure followed by iteratively updating the development point according to $E_1 = \text{exp}_{E_1}([0, 0, \delta, 0, 0, \delta])$ for some δ . This basically applies z -axis rotations which assure that all essential matrices in \mathcal{R}_z are mapped to a single point. The resulting tangent vector $[u_x, u_y, z, v_x, v_y, -z]^\top$ has the same number of degrees of freedom as an essential matrix, i.e. five, and therefore provides a minimal representation. This representation is not unique however.

The focus is on B related to choice of basis when computing the SVD. If the claim in [5, 7] is correct, then every choice for B must result in the same point on their manifold. By ensuring a positive unit determinant of U and V the four choices for B related to reflections, i.e. $\text{diag}\{-1, -1, -1\}$, $\text{diag}\{-1, 1, 1\}$, $\text{diag}\{1, -1, 1\}$ and $\text{diag}\{1, 1, -1\}$, are mapped to $\text{diag}\{-1, -1, 1\}$, $\text{diag}\{-1, 1, -1\}$, $\text{diag}\{1, -1, -1\}$ and $\text{diag}\{1, 1, 1\}$ respectively. Note that the identity $I = \text{diag}\{1, 1, 1\}$ and $\text{diag}\{-1, -1, 1\}$ are both elements of \mathcal{R}_z and can therefore be resolved to I . The remaining two elements of \mathcal{B} , i.e. $\text{diag}\{1, -1, -1\} = R_x(\pm\pi)$ and $\text{diag}\{-1, 1, -1\} = R_y(\pm\pi)$, relate to rotations by an angle $\pm\pi$ around respectively the x -axis and y -axis. They are not elements of \mathcal{R}_z . Indeed, $R_x(\pm\pi) R_z(\pm\pi) = R_y(\pm\pi)$ and therefore the initialization procedure in [5, 7] should ideally be able to map both to a single element, e.g. $R_y(\pm\pi)$ (but [5, 7] provide no control over which one). When nevertheless assuming that $R_x(\pm\pi)$ is mapped to $R_y(\pm\pi)$ consistently, then there are still two choices for B . They are: I and $R_y(\pm\pi)$. They clearly do not differ by a z -axis rotation. Therefore, both will be mapped to tangents of the form $[u_x, u_y, z, v_x, v_y, -z]^\top$ but these tangent vectors will not be equal. Hence, there remains at least a twofold ambiguity in the approach of [5, 7].

Instead of trying to resolve the four-fold ambiguity algebraically as in [5, 7], we exploit the positive-depth constraint before applying our mappings.

4 Experimental evaluation

Our novel Riemannian mappings and those in [5, 7] are evaluated by incorporating them into two intrinsic statistical algorithms. One algorithm computes the intrinsic mean [3]

the other computes the dominant mode by using intrinsic mean-shift [4–8]. When reporting results of our mappings we use the prefix *Our* and for the mappings in [5, 7] we use the prefix *SM* (Subbarao and Meer). The mappings in [15–17] are not evaluated as they cannot be used *as is* for manifold statistics.

4.1 Maximum likelihood lower bound

Artificial image data is generated by projecting 3D world points onto two camera planes. The poses between the cameras are randomly generated with translation directions uniformly distributed over the unit sphere, translation distance varying uniformly between 1 m and 10 m and rotation yaw, pitch and roll parameters distributed uniformly in the interval $\pm 45^\circ$. Gaussian image noise with $\sigma = 0.2$ pixel is added to the projections. Essential matrices are estimated on minimal subsets, i.e. of size 5 point correspondences.

Fig.2(a,d) depicts the results (on a logarithmic scale) of computing the intrinsic mean on an increasing number of essential matrices. For this experiment the essential matrices are obtained using an ML estimator which minimizes reprojection errors.

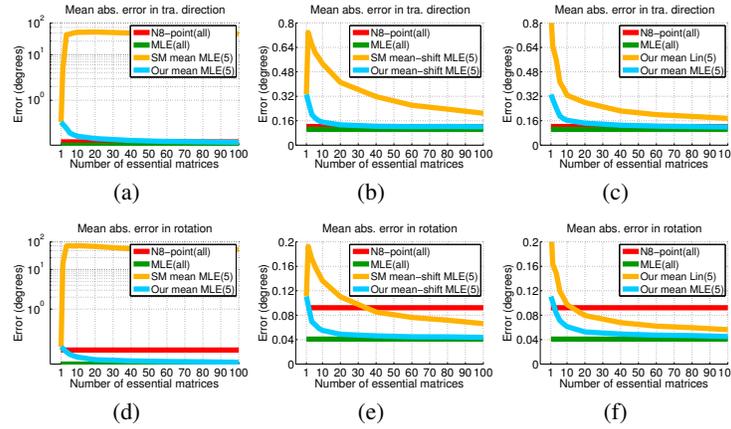
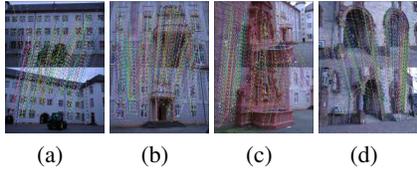


Fig. 2. Result averaged over 500 poses. Errors in translation direction are visualized in the top row and rotational errors in the bottom row. See Sec. 4.1 for details.

It can clearly be seen that when using our mappings, the accuracy asymptotically reaches the ML lower bound, i.e. the accuracy of an ML estimator on all image points. For reference, the accuracy of the normalized 8-point algorithm on all image points is also shown. When using the mappings of [5, 7], the mean does not produce a useable result. This is because all essential matrices, which only differ by a small random perturbation from the ground truth, are mapped to two clusters far apart from each other instead of to one. This is due to the two-fold ambiguity in their mappings. Computing the mean on these two clusters will result in a point in between the two clusters



	Max. subset	Tra. dir.	Rot.
Cov-RANSAC	14	0.39°	0.13°
LO-RANSAC	all inliers	0.27°	0.08°
Our mean-shift	5	1.21°	0.43°
SM mean-shift	5	1.47°	0.77°
Our top10-mean	5	0.26°	0.07°
SM top10-mean	5	24.9°	28.4°

Fig. 3. Example image pairs from the benchmark images provided by [28]. **Table 1.** Absolute error averaged over 10 runs on 44 benchmark images of [28].

which does not correspond to the ground truth. This experiment clearly illustrates our theoretical findings of Sec. 3.2.

The experimental evaluation in [5, 7] did not reveal this because the statistical algorithm used in [5, 7] is mean-shift. In Fig.2(b,e) we therefore show the result obtained when using mean-shift. When using the mappings in [5, 7] accuracy converges but results are still factors less accurate than when using our mappings. This is because mean-shift with the mappings in [5, 7] converges to the mode of the most dominant of two clusters and ignores the other cluster. It will therefore not use all available statistical information, though it will produce a result. Also note that its error on two essential matrices is larger than that of a single essential matrix. This is because for two essential matrices, there can be no dominant cluster. Our mappings do not exhibit this behavior.

In Fig.2(c,f) we show the result of our method when using a linear estimator (5-point algorithm) to generate the essential matrices. Again we see that the mean converges to satisfactory accuracy. Its rotational estimate even becomes more accurate than that of the normalized 8-point algorithm on all image points. Resolving the four-fold ambiguity is taken care of within our ML estimator, but for the 5-point algorithm we have to do it explicitly before applying our Riemannian mappings. We do this efficiently by immediately rejecting a configuration as soon as one point from the minimal subset is behind a camera. We found that there is practically no difference in the computation time between this approach and the iterative approach of [5, 7].

4.2 Robust pose clustering and RANSAC

We now show the practical relevance of our theoretical work by comparing against dominant state-of-the-art robust estimators, i.e. Cov-RANSAC [19] and LO-RANSAC [20]. We evaluate on 44 benchmark images of [28], see Fig. 3, which are accompanied by ground-truth pose information. All approaches use ML estimators for improved accuracy and extensive parameters sweeps are performed to assure that their accuracy is not limited by sub-optimal parameter settings.

The satisfactory accuracy of Cov-RANSAC and LO-RANSAC averaged over ten runs (i.e. 440 image pairs) is depicted in the top two rows of Table 1. Note the favorable accuracy of LO-RANSAC as it computes inner ML estimates on all inliers. This however counterbalances its reduction in the number of iterations which causes this LO-RANSAC approach to be around 1.5 times slower than regular RANSAC while

our Cov-RANSAC implementation is around 2.5 times faster. The outlier percentage of the data sets estimated by these methods is 70%. The next two rows in Table 1 are of mean-shift applied to models obtained from regular RANSAC as in [5, 7]. These errors are factors higher than those of the state-of-the-art RANSAC methods. However, it can again be observed that when using our Riemannian mappings, improved accuracy is obtained. While mean-shift clearly has great practical relevance, it is no match for Cov-RANSAC and LO-RANSAC in this experimental context.

The final two rows in Table 1 show the accuracy when computing the intrinsic mean on the top-10 ranking RANSAC models, which is a *RANSAC-manifold hybrid* method. It can be observed that when using our Riemannian mappings this obtains optimal accuracy for both directional and rotational estimates. This is because there are typically models estimated on points which are on the boundary of being inliers but are not real outliers either. These models can be identified based on their relative high number of inliers and computing their intrinsic mean can be as accurate as an ML estimate on the union of their subsets. This was demonstrated in Fig.2(a,d,c,f). The hybrid approach has practically the same computation time as regular RANSAC. When using the mappings in [5, 7] to compute the mean, no satisfactory accuracy is obtained.

The purpose of this experiment is to illustrate the advantage of using our mappings in a realistic setting. Their purpose is explicitly not to provide a comparison against all *state-of-the-art* RANSAC methods.

5 Conclusions

We have provided correct Riemannian logarithmic and exponential mappings for essential matrices. They allow the generalization of statistical algorithms designed for Euclidean vector spaces to the manifold of essential matrices. Mapping functions proposed earlier were proven to be incorrect. Using them within statistical calculation leads to sub-optimal or even erroneous results. Our theoretical findings are shown to be relevant to whoever is interested in developing robust statistical algorithms for essential matrices.

References

1. Karcher, H.: Riemannian Center of Mass and Mollifier Smoothing. *Communications on Pure and Applied Mathematics* **30** (1977) 509–541
2. Pennec, X., Ayache, N.: Uniform Distribution, Distance and Expectation Problems for Geometric Features Processing. *Journal of Mathematical Imaging and Vision* **9** (1998) 49–67
3. Pennec, X.: Intrinsic Statistics on Riemannian Manifolds: Basic Tools for Geometric Measurements. *Journal of Mathematical Imaging and Vision* **25** (2006) 127–154
4. Subbarao, R., Meer, P.: Nonlinear Mean Shift for Clustering over Analytic Manifolds. In: *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Volume 1. (2006) 1168–1175
5. Subbarao, R., Meer, P.: Nonlinear Mean Shift over Riemannian Manifolds. *International Journal of Computer Vision* **84** (2009) 1–20
6. Subbarao, R., Genc, Y., Meer, P.: Nonlinear Mean Shift for Robust Pose Estimation. In: *IEEE Workshop on Applications of Computer Vision*. (2007) 6

7. Subbarao, R., Genc, Y., Meer, P.: Robust Unambiguous Parametrization of the Essential Manifold. In: IEEE Conference on Computer Vision and Pattern Recognition. (2008) 1–8
8. Tuzel, O., Subbarao, R., Meer, P.: Simultaneous Multiple 3D Motion Estimation via Mode Finding on Lie Groups. In: Proceedings of the Tenth IEEE International Conference on Computer Vision. Volume 1. (2005) 18–25
9. Costa, J., Hero, A.O.: Learning I. In: Statistics and Analysis of Shapes. MA: Birkhauser, Cambridge (2006) 231–252
10. Fletcher, P.T., Joshi, S.: Riemannian Geometry for the Statistical Analysis of Diffusion Tensor Data. *Signal Processing* **87** (2007) 250–262
11. Pennec, X., Guttman, C.R.G., Thirion, J.: Feature-based Registration of Medical Images: Estimation and Validation of the Pose Accuracy. *Lecture Notes in Computer Science* **1496/1998** (1998) 1107–1114
12. Begelfor, E., Werman, M.: Affine Invariance Revisited. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition. (2006) 2087–2094
13. Goh, A., Vidal, R.: Clustering and Dimensionality Reduction on Riemannian Manifolds. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition. (2008) 1–7
14. Tuzel, O., Porikli, F., Meer, P.: Pedestrian Detection via Classification on Riemannian Manifolds. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **30** (2008) 1713–1727
15. Helmke, U., Hüper, K., Lee, P.Y., Moore, J.: Essential Matrix Estimation Using Gauss-Newton Iterations on a Manifold. *International Journal of Computer Vision* **74** (2007) 117–136
16. Košecká, J., Ma, Y., Sastry, S.: Optimization Criteria, Sensitivity and Robustness of Motion and Structure Estimation. *Lecture Notes in Computer Science* **1883** (2000) 166–182
17. Ma, Y., Košecká, J., Sastry, S.: Optimization Criteria and Geometric Algorithms for Motion and Structure Estimation. *International Journal of Computer Vision* **44** (2001) 219–249
18. Fischler, M.A., Bolles, R.C.: Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM* **24** (1981) 381–395
19. Raguram, R., Frahm, J., Pollefeys, M.: Exploiting Uncertainty in Random Sample Consensus. In: International Conference on Computer Vision. (2009)
20. Chum, O., Matas, J., Kittler, J.: Locally Optimized RANSAC. *Lecture Notes in Computer Science* **2781** (2003) 236–243
21. do Carmo, M.P.: Riemannian Geometry. 1st edn. Mathematics: Theory and Applications. Birkhäuser (1992)
22. Cullen, H.F.: Introduction to General Topology. Heath, Boston (1967)
23. Longuet-Higgins, H.C.: A Computer Algorithm for Reconstructing a Scene from Two Projections. *Nature* **293** (1981) 133–135
24. Hartley, R.I., Zisserman, A.: Multiple View Geometry in Computer Vision. Second edn. Cambridge University Press, ISBN: 0521540518 (2004)
25. Park, F.C.: Distance Metrics on the Rigid-Body Motions with Applications to Mechanism Design. *Transactions of the ASME* **117** (1995) 48–54
26. Eisenberg, M., Guy, R.: A Proof of the Hairy Ball Theorem. *The American Mathematical Monthly* **86** (1979) 571–574
27. Selig, J.M.: Geometrical Methods in Robotics. Springer (1996)
28. Strecha, C., von Hansen, W., van Gool, L., Fua, P., Thoennessen, U.: On Benchmarking Camera Calibration and Multi-View Stereo for High Resolution Imagery. In: IEEE Conference on Computer Vision and Pattern Recognition. (2008) 1–8