Essential Matrix Estimation Using Gauss-Newton Iterations on a Manifold

UWE HELMKE
Department of Mathematics, University of Würzburg, D-97074 Würzburg, Germany
helmke@mathematik.uni-wuerzburg.de

KNUT HÜPER, PEI YEAN LEE AND JOHN MOORE
*National ICT Australia Ltd. Locked Bag 8001, Canberra ACT 2601, Australia; and Department of Information
Engineering, RSISE, The Australian National University, Canberra, ACT 0200, Australia*
knut.hueper@nicta.com.au
peiyean.lee@nicta.com.au
john.moore@anu.edu.au

Abstract. A novel approach for essential matrix estimation is presented, this being a key task in stereo vision processing. We estimate the essential matrix from point correspondences between a stereo image pair, assuming that the internal camera parameters are known. The set of essential matrices forms a smooth manifold, and a suitable cost function can be defined on this manifold such that its minimum is the desired essential matrix. We seek a computationally efficient optimization scheme towards meeting the demands of on-line processing of video images. Our work extends and improves the earlier research by Ma et al., who proposed an intrinsic Riemannian Newton method for essential matrix computations. In contrast to Ma et al., we propose three Gauss-Newton type algorithms that have improved convergence properties and reduced computational cost. The first one is based on a novel intrinsic Newton method, using the normal Riemannian metric on the manifold consisting of all essential matrices. The other two methods are Newton-like methods, that are more efficient from a numerical point of view. Local quadratic convergence of the algorithms is shown, based on a careful analysis of the underlying geometry of the problem.

1. Introduction

In the seminal work by Longuet-Higgins (1981) it is shown that the pose information between a pair of calibrated camera images is algebraically captured by an essential matrix. Several techniques have been developed for estimation of this matrix and are reviewed e.g. in Huang and Faugeras (1989) and Huang and Netravali (1994). We mention only two such methods. The first is the celebrated normalized eight-point algorithm (Hartley and Zisserman, 2003), that works well for noise free data but presents only an approximate solution in the important noisy data case. The second one is due to Ma et al. (2001), who proposed an intrinsic Newton method on the manifold of essential matrices, in order to improve

the estimate obtained by the eight-point algorithm. The algorithm in Ma et al. (2001), although locally quadratically convergent, has shortcomings that will be resolved in the present paper. First, it converges only in a possibly small neighborhood of the true solution, while secondly, it is confined to a particular Riemannian metric structure.

In this paper, which builds on our earlier conference paper (Helmke et al., 2004), a new optimization approach for essential matrix computations is proposed that is based on the homogeneous space structure of the manifold $\mathcal{E}$ of essential matrices. Using different local parametrizations for $\mathcal{E}$, a new class of Newton-like methods is introduced that is shown to converge locally quadratically fast to the desired solution. By combining

this method with suitable gradient descent steps, new variants of the Gauss-Newton method on a manifold are obtained. This leads to an improvement of the Riemannian Newton methods by enlarging the domain of attraction for the algorithm. Our approach may also be of interest in other optimization applications, where a greater flexibility in algorithm design may prove beneficial. Moreover, appropriate choices of the local parametrization lead to more efficient numerical implementations, as shown by computer simulations with complexity per iteration considerably less than in the work of Ma et al. (2001).

We begin with the problem formulation in Section 2. This is followed in Section 3 by a geometric analysis of the manifold of essential matrices, including tangent space descriptions, local parametrisations and subsequent discussion of the cost function. The algorithm, together with a proof of local quadratic convergence is discussed in Section 4. Simulation results are presented in Section 5. The conclusion can be found in Section 6 and a detailed comparison between the theoretical approach by Ma et al. and ours is detailed in the Appendix.

2. Problem Formulation

Two images of the same scene are related by epipolar geometry (Longuet-Higgins, 1981), as we now explain (see Fig. 1). Although the subsequent material in this section is certainly well known in the computer vision community, we include the derivation of the epipolar constraint, to enhance readability of the paper for a wider audience.

For simplicity, we assume that images are taken by two identical pin-hole cameras with focal length equal to one. The more general case of arbitrary cameras can be treated similarly, provided the camera calibration matrices are known.

Thus the two cameras are specified by the camera centers $C_1, C_2 \in \mathbb{R}^3$ and attached orthogonal camera frames $\{e_1, e_2, e_3\}$ and $\{e'_1, e'_2, e'_3\}$, respectively. Let $t := C_2 - C_1$

be the translation vector of the first camera into the second and $R \in SO_3$ the uniquely determined real orthogonal transformation that transforms the basis vectors e_1, e_2, e_3 into e'_1, e'_2, e'_3, expressed with respect to the basis e_1, e_2, e_3. Let $[t_1 t_2 t_3]^\top$ denote the coordinates of t with respect to the first camera basis. Given an object point $M \in \mathbb{R}^3$ let $m_i := [X_i Y_i Z_i]^\top$, $i = 1, 2$, denote the coordinates of M with respect to the two camera bases. Thus

$$m_1 = Rm_2 + t. \qquad (1)$$

Let

$$\Omega := \begin{bmatrix} 0 & -t_3 & t_2 \\ t_3 & 0 & -t_1 \\ -t_2 & t_1 & 0 \end{bmatrix} \qquad (2)$$

be the skew-symmetric matrix satisfying $\Omega t = 0$. Thus, by premultiplying (1) from the left by Ω and noting that $m_1^\top \Omega m_1 = 0$, we obtain

$$m_1^\top \Omega R m_2 = 0. \qquad (3)$$

Let $\widehat{m}_1 = [u_1 \ v_1 \ 1]^\top$, $\widehat{m}_2 = [u_2 \ v_2 \ 1]^\top$ be the camera image points, expressed with respect to the camera basis in pixel image coordinates, respectively. Since the vectors $\widehat{m}_i$ and m_i are proportional we obtain the so-called epipolar constraint

$$\widehat{m}_1^\top \Omega R \widehat{m}_2 = 0 \qquad (4)$$

for the camera image coordinates. The matrix

$$E = \Omega R \qquad (5)$$

is called the essential matrix. More generally, any (3×3)-matrix of the form $E := \Omega R$, where Ω, R denote a (3×3)- skew-symmetric matrix and orthogonal rotation matrix, respectively, is called an essential matrix. From its knowledge the pose information defined by the rotation R and the translation vector t are easily determined
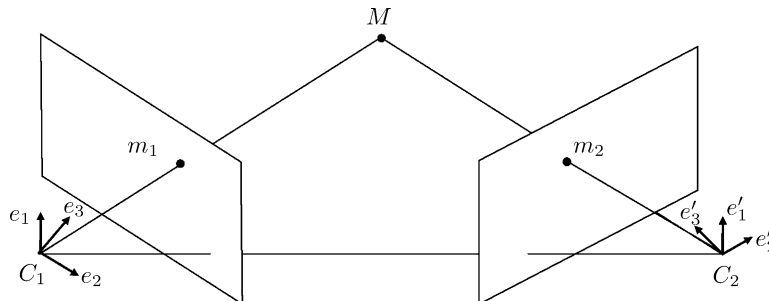

Figure 1. Epipolar geometry.

(see Section 3.1). The essential matrix plays a crucial role in motion analysis. Thus, given a set of point correspondences between a pair of calibrated intensity images, the task is to estimate the essential matrix which encapsulates the relative pose between the images.

It is known that with eight or more generic point correspondences in the noise-free case, the essential matrix is uniquely determined. For a discussion of the number of solutions for five generic point correspondences we refer to Horn (1990).

3. Optimization on the Essential Manifold

3.1. Geometry of the Essential Manifold

Since the novel aspects of our algorithms depend on a detailed understanding of the differential geometry of essential matrices, we begin by reviewing such geometric details. Note that the authors of Soatto et al. (1996) already suggest to exploit the geometry of the essential manifold.

Orientation preserving rotational transformations in $\mathbb{R}^3$ are represented by elements of the special orthogonal group SO_3, which consists of (3×3) orthogonal matrices of determinant equal to 1:

$$SO_3 := \{X \in \mathbb{R}^{3 \times 3} \mid \det X = 1, \, X^\top X = I_3\}. \quad (6)$$

The set SO_3 is a 3-dimensional Lie group and its associated Lie algebra $\mathfrak{so}_3$ is the set of (3×3) skew-symmetric matrices:

$$\mathfrak{so}_3 := \{X \in \mathbb{R}^{3 \times 3} \mid X = -X^\top\}. \quad (7)$$

There is a well known isomorphism from the Lie algebra $(\mathbb{R}^3, \times)$ to the Lie algebra $(\mathfrak{so}_3, [., .])$, where $\times$ denotes the vector cross product and $[A, B] := AB - BA$ denotes the matrix commutator. This allows one to identify $\mathfrak{so}_3$ with $\mathbb{R}^3$ using the linear map

$$\Omega : \mathbb{R}^3 \to \mathfrak{so}_3,$$

$$\omega = \begin{bmatrix} \omega_x \\ \omega_y \\ \omega_z \end{bmatrix} \mapsto \begin{bmatrix} 0 & -\omega_z & \omega_y \\ \omega_z & 0 & -\omega_x \\ -\omega_y & \omega_x & 0 \end{bmatrix}. \quad (8)$$

Notice that Ω can be written as,

$$\Omega(\omega) = \omega_x Q_x + \omega_y Q_y + \omega_z Q_z, \quad (9)$$

where

$$Q_x := \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}, \quad Q_y := \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix},$$

$$Q_z := \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \quad (10)$$

According to the definition above, an essential matrix is defined as the product $E = \Omega R$ of a skew-symmetric matrix Ω and a rotation matrix R. It is well-known from the computer vision literature (see Hartley and Zisserman, 2003), that the essential matrices are characterized by the property that they have exactly one positive singular value of multiplicity two, consequently E must be rank 2. In particular, normalized essential matrices of Frobenius norm equal to $\sqrt{2}$ are therefore characterized by having the set of singular values $\{1, 1, 0\}$. Consequently, the set of normalized essential matrices is given as

$$\mathcal{E} := \{\Omega R \mid \Omega \in \mathfrak{so}_3, \, R \in SO_3, \, \|\Omega\|^2 = 2\}. \quad (11)$$

This is the basic nonlinear constraint set on which the proposed optimization algorithms are defined, where $\|\Omega\|$ denotes the Frobenius norm $\sqrt{\mathrm{tr}(\Omega \Omega^\top)}$.

3.1.1. SVD Characterization of Normalized Essential Matrices. There is an alternative characterization of essential matrices as the homogeneous space of the compact Lie group $SO_3 \times SO_3$. This description is fundamental to our subsequent algorithm development and is therefore explained in more detail. A key property of essential matrices for this characterization is as follows.

Lemma 3.1. *A non-zero (3×3)-matrix E is essential if and only if there exists a singular value decomposition (SVD),*

$$E = U \Sigma V^\top \quad \text{with} \quad \Sigma = \begin{bmatrix} s & 0 & 0 \\ 0 & s & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad s > 0$$

$$\text{and} \quad U, V \in SO_3. \quad (12)$$

Moreover, E is a normalized essential matrix, if and only if $s = 1$. In particular, for

$$E_0 := \begin{bmatrix} I_2 & 0 \\ 0 & 0 \end{bmatrix}, \quad (13)$$

the normalized essential manifold is equal to the homogeneous space

$$\mathcal{E} = \{U E_0 V^\top \mid U, V \in SO_3\}. \qquad (14)$$

Proof: Assume that $E = \Omega R$ is essential with $\Omega \in \mathfrak{so}_3$ and $R \in SO_3$. Then the equality $E = \Omega R$ implies

$$EE^\top = \Omega R R^\top \Omega^\top$$
$$= -\Omega^2,$$

with a corresponding set of eigenvalues

$$\lambda(EE^\top) = \{s^2, s^2, 0\}, \qquad \text{where} \qquad s := \tfrac{1}{\sqrt{2}}\|\Omega\|.$$

The set of singular values of E is then $\sigma(E) = \{s, s, 0\}$. Thus there exists an SVD, see (12), with $U, V \in O_3$. But it is easily seen that there exists also a factorisation (12) with $U, V \in SO_3$.

For the converse, consider

$$\Psi := \begin{bmatrix} 0 & -s & 0 \\ s & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \in \mathfrak{so}_3, \quad \Gamma := \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \in SO_3.$$

One has

$$\Psi^\top = \Gamma \begin{bmatrix} s & 0 & 0 \\ 0 & s & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Hence by the singular value decomposition of E, with $U, V \in SO_3$,

$$\begin{aligned} E &= U \begin{bmatrix} s & 0 & 0 \\ 0 & s & 0 \\ 0 & 0 & 0 \end{bmatrix} V^\top \\ &= U\Gamma^\top \Gamma \begin{bmatrix} s & 0 & 0 \\ 0 & s & 0 \\ 0 & 0 & 0 \end{bmatrix} \Gamma U^\top U \Gamma^\top V^\top \\ &= (U\Gamma^\top \Psi^\top \Gamma U^\top)(U\Gamma^\top V^\top) \\ &= \Omega_1 R_1, \end{aligned} \qquad (15)$$

with

$$\Omega_1 = U\Gamma^\top \Psi^\top \Gamma U^\top \quad \text{and} \quad R_1 = U\Gamma^\top V^\top \quad (16)$$

as required. A second decomposition of E is $E = \Omega_2 R_2$ with $\Omega_2 = -\Omega_1$ and $R_2 = U\Gamma V^\top$.   □

The orthogonal matrices appearing in the above SVD of a given essential matrix are not uniquely determined. Thus, given factorising matrices $U, V \in SO_3$, then

$$U \begin{bmatrix} \Theta & 0 \\ 0 & \det\Theta \end{bmatrix}, \quad V \begin{bmatrix} \Theta & 0 \\ 0 & \det\Theta \end{bmatrix} \qquad (17)$$

are also factors, where $\Theta \in O_2$. Moreover, this describes all possible factors of an essential matrix.

From the above characterization of essential matrices we see that there are two different descriptions of $\mathcal{E}$: (i) The characterization (14) via the SVD and (ii) the characterization as the image of the map

$$\begin{aligned} \pi : \mathbb{S}^2 \times SO_3 &\to \mathcal{E}, \\ (\Omega, R) &\mapsto \Omega R. \end{aligned} \qquad (18)$$

The second one has been extensively used in Ma et al. (2001), whereas the first one will be exploited here. By the next theorem we clarify how these two factorisations can be computed from each other.

Theorem 3.1.   *Let*

$$\mathbb{S}^2 := \left\{ \Omega \in \mathfrak{so}_3 \,\middle|\, \|\Omega\| = \sqrt{2} \right\} \qquad (19)$$

denote the two-sphere in $\mathfrak{so}_3$ with radius equal to $\sqrt{2}$. The map

$$\begin{aligned} \pi : \mathbb{S}^2 \times SO_3 &\to \mathcal{E}, \\ (\Omega, R) &\mapsto \Omega R, \end{aligned} \qquad (20)$$

is a two-to-one covering map with preimages

$$\begin{aligned} &\pi^{-1}\left(U \begin{bmatrix} I_2 & 0 \\ 0 & 0 \end{bmatrix} V^\top \right) \\ &= \left(U \begin{bmatrix} 0 & \varepsilon & 0 \\ -\varepsilon & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} U^\top, U \begin{bmatrix} 0 & -\varepsilon & 0 \\ \varepsilon & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} V^\top \right), \end{aligned} \qquad (21)$$

with $U, V \in SO_3$ and $\varepsilon \in \{+1, -1\}$. Thus if $E = U E_0 V^\top$, then $E = \Omega R$ with $\varepsilon = \pm 1$,

$$\Omega = U \begin{bmatrix} 0 & \varepsilon & 0 \\ -\varepsilon & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} U^\top \quad \text{and}$$

$$R = U \begin{bmatrix} 0 & -\varepsilon & 0 \\ \varepsilon & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} V^\top. \qquad (22)$$

Conversely, if $E = \Omega R$ is given, then the eigenvalue decomposition of

$$\Omega = U \begin{bmatrix} 0 & \varepsilon & 0 \\ -\varepsilon & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} U^\top \qquad (23)$$

yields U, and V is given as

$$V := R^\top U \begin{bmatrix} 0 & -\varepsilon & 0 \\ \varepsilon & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \qquad (24)$$

Proof: The proof is obvious from the above discussion. □

For a geometric interpretation of the two possible factorisations $E = \Omega_1 R_1 = \Omega_2 R_2$ (see Hartley and Zisserman, 2003), Figs. 9.12 (a), (d) in Section 9.6.3. Note that Figs. (b), (c) in that section correspond to the factorisation of $-E$.

The next result details the manifold structure of $\mathcal{E}$.

Theorem 3.2. *The set $\mathcal{E}$ is a smooth five-dimensional compact manifold diffeomorphic to $\mathbb{RP}^2 \times SO_3$.*

Proof: Recall, that the real projective plane $\mathbb{RP}^2$ can be identified with the isospectral manifold

$$\left\{ U \begin{bmatrix} I_2 & 0 \\ 0 & 0 \end{bmatrix} U^\top \,\middle|\, U \in SO_3 \right\}.$$

It can be shown that the map

$$\phi : \mathcal{E} \to \mathbb{RP}^2 \times SO_3,$$

$$U \begin{bmatrix} I_2 & 0 \\ 0 & 0 \end{bmatrix} V^\top \mapsto \left(U \begin{bmatrix} I_2 & 0 \\ 0 & 0 \end{bmatrix} U^\top, UV^\top \right) \qquad (25)$$

is smooth and bijective. Moreover, the inverse

$$\phi^{-1}(X, Y) = XY,$$

is smooth as well. Therefore, ϕ is a diffeomorphism. Since $\mathbb{RP}^2$ is of dimension 2, and SO_3 is of dimension 3, then $\mathbb{RP}^2 \times SO_3$ is of dimension 5. □

3.1.2. Tangent Space of the Essential Manifold.
One advantage of the characterization of essential matrices by the singular value decomposition is that it allows for an elegant description of associated tangent spaces.

Theorem 3.3. *The tangent space at the normalized essential matrix $E = U E_0 V^\top$ is*

$$T_E \mathcal{E} = \{U(\Omega E_0 - E_0 \Psi)V^\top | \Omega, \Psi \in \mathfrak{so}_3\}$$

$$= \left\{ U \begin{bmatrix} 0 & \omega_{12} - \psi_{12} & -\psi_{13} \\ \psi_{12} - \omega_{12} & 0 & -\psi_{23} \\ -\omega_{13} & -\omega_{23} & 0 \end{bmatrix} V^\top \,\middle|\, \omega_{ij}, \right.$$

$$\left. \psi_{ij} \in \mathbb{R}, i, j \in \{1, 2, 3\} \right\} \qquad (26)$$

with $\Omega = (\omega_{ij})$ and $\Psi = (\psi_{ij})$.

Proof: For any $E = U E_0 V^\top \in \mathcal{E}$, let $\alpha_E : SO_3 \times SO_3 \to \mathcal{E}$ be the smooth map defined by $\alpha_E(\widehat{U}, \widehat{V}) = \widehat{U} E \widehat{V}^\top$. The tangent space $T_E \mathcal{E}$ is the image of the linear map

$$D\alpha_E(I_3, I_3) : \mathfrak{so}_3 \times \mathfrak{so}_3 \to \mathbb{R}^{3 \times 3},$$
$$(\widehat{\Omega}, \widehat{\Psi}) \mapsto \widehat{\Omega} E - E \widehat{\Psi}, \qquad (27)$$

i.e., the image of the derivative of α_E evaluated at the identity $(I_3, I_3) \in SO_3 \times SO_3$. By setting $\Omega := U^\top \widehat{\Omega} U$ and $\Psi := V^\top \widehat{\Psi} V$ the first equation is shown (see Helmke and Moore, 1994) (Helmke and Moore, 1994), pp. 89 for details. The result follows from a straightforward computation. □

Corollary 3.1. *The kernel of the mapping $D\alpha_{E_0}(I_3, I_3) : \mathfrak{so}_3 \times \mathfrak{so}_3 \to \mathbb{R}^{3 \times 3}$ is the set of matrix pairs $(\Omega, \Psi) \in \mathfrak{so}_3 \times \mathfrak{so}_3$ with*

$$\Omega = \Psi = \begin{bmatrix} 0 & x & 0 \\ -x & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad x \in \mathbb{R}. \qquad (28)$$

Proof: Let $\Omega = (\omega_{ij})$, $\Psi = (\psi_{ij})$ with $\Omega, \Psi \in \mathfrak{so}_3$ and E_0 as defined in (13), then

$$\Omega E_0 - E_0 \Psi = 0$$

$$\iff \begin{bmatrix} 0 & \omega_{12} - \psi_{12} & -\psi_{13} \\ -(\omega_{12} - \psi_{12}) & 0 & -\psi_{23} \\ -\omega_{13} & -\omega_{23} & 0 \end{bmatrix} = 0_3$$

$$\iff \omega_{12} = \psi_{12} \quad \text{and} \quad \omega_{13} = \psi_{13} = \omega_{23} = \psi_{23} = 0.$$
$$(29)$$

Setting $\omega_{12} = \psi_{12} = x$ gives the result. □

Since the affine tangent space $T_E^{\text{aff}} \mathcal{E}$ at $E = U E_0 V^\top \in \mathcal{E}$ can be identified with the tangent space $T_E \mathcal{E}$ via
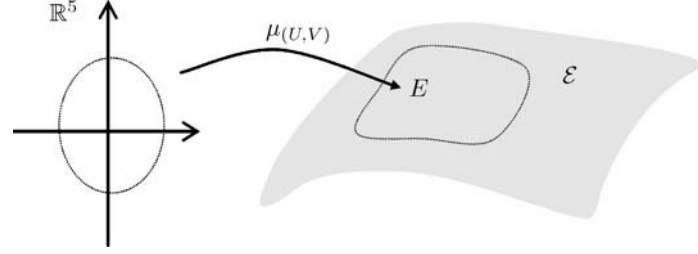
Figure 2.   Local parameterization of essential manifold.

translating $T_E \mathcal{E}$ by E, we obtain

$$
T_E^{\text{aff}} \mathcal{E} = \left\{ U \begin{bmatrix} 1 & -x_3 & -x_5 \\ x_3 & 1 & x_4 \\ -x_2 & x_1 & 0 \end{bmatrix} V^\top \,\middle|\, x_1, \dots, x_5 \in \mathbb{R} \right\}.
\tag{30}
$$

3.1.3. Parameterization of the Essential Manifold

Computations on a manifold are often conveniently carried out in terms of a local parameterization. For our later convergence analysis we therefore need a local parameterization of the essential manifold (see Fig. 2).

We introduce a class of parameterizations for $\mathcal{E}$, that is basic for the subsequent formulation of our algorithms. Given an essential matrix $E \in \mathcal{E}$ let $U, V \in SO_3$ with $U E_0 V^\top = E$. Recall, that any other factorisation $E = \widetilde{U} E_0 \widetilde{V}^\top$ with $\widetilde{U}, \widetilde{V} \in SO_3$ is given as $(\widetilde{U}, \widetilde{V}) = (U\Gamma, V\Gamma)$, where

$$
\Gamma \in \mathcal{S} := \left\{ \begin{bmatrix} \Theta & 0 \\ 0 & \det \Theta \end{bmatrix} \,\middle|\, \Theta \in O_2 \right\}.
\tag{31}
$$

Here, $\mathcal{S}$ is a subgroup of SO_3, called the stabiliser subgroup of E_0.

Definition 3.1.   Let $U, V \in SO_3$ with $U E_0 V^\top = E$. A local parameterization for $\mathcal{E}$ at E is a smooth map

$$
\mu_{(U,V)} : \mathbb{R}^5 \to \mathcal{E}
\tag{32}
$$

such that the following conditions hold:

(i) $\mu_{(U,V)}(0) = E$.
(ii) $\mu_{(U,V)}$ is a local diffeomorphism around 0.
(iii)   There exists a map $L : \mathcal{S} \to GL_5$ such that

$$
\mu_{(U\Gamma, V\Gamma)}(x) = \mu_{(U,V)}(L(\Gamma)x)
\tag{33}
$$

for all $\Gamma \in \mathcal{S}$ and $x \in \mathbb{R}^5$.

Here

$$
GL_5 := \{ X \in \mathbb{R}^{5 \times 5} \mid \det X \neq 0 \}.
\tag{34}
$$

In the sequel we consider three specific examples.

To describe them, let $U, V \in SO_3$ be arbitrary, let $x = [x_1, \dots, x_5]^\top \in \mathbb{R}^5$, and let E_0 be defined as in (13). Consider the mappings

$$
\Omega_1 : \mathbb{R}^5 \to \mathfrak{so}_3,
$$
$$
[x_1, \dots, x_5]^\top \mapsto \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & -\frac{x_3}{\sqrt{2}} & x_2 \\ \frac{x_3}{\sqrt{2}} & 0 & -x_1 \\ -x_2 & x_1 & 0 \end{bmatrix}
\tag{35}
$$

and

$$
\Omega_2 : \mathbb{R}^5 \to \mathfrak{so}_3,
$$
$$
[x_1, \dots, x_5]^\top \mapsto \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & \frac{x_3}{\sqrt{2}} & x_5 \\ -\frac{x_3}{\sqrt{2}} & 0 & -x_4 \\ -x_5 & x_4 & 0 \end{bmatrix}.
\tag{36}
$$

Note, that the $\sqrt{2}$-factors are important for subsequent Riemannian geometry interpretations.

Example 3.1.   The *exponential parameterization* is defined as

$$
\mu_{(U,V)} : \mathbb{R}^5 \to \mathcal{E},
$$
$$
\mu_{(U,V)}(x) := U \, e^{\Omega_1(x)} E_0 \, e^{-\Omega_2(x)} V^\top.
\tag{37}
$$

This is a local parameterization of $\mathcal{E}$ at E. Here $\Omega_1(x)$ and $\Omega_2(x)$ are defined as in (35) and (36), respectively. For the verification that conditions (i) and (ii) in Definition 3.1.3 are satisfied we proceed as follows. Smoothness of $\mu_{(U,V)}$ and $\mu_{(U,V)}(0) = E$ are obvious. To see that $\mu_{(U,V)}$ is a local diffeomorphism around $0 \in \mathbb{R}^5$, we have to show that the derivative

$$
D\mu_{(U,V)}(0) : \mathbb{R}^5 \to T_E \mathcal{E},
\tag{38}
$$

is invertible. For arbitrary $h = [h_1, \ldots, h_5]^\top \in \mathbb{R}^5$, then

$$
\begin{aligned}
&\mathrm{D}\mu_{(U,V)}(0)h \\
&= U \left(\frac{1}{\sqrt{2}} \begin{bmatrix} 0 & -\frac{h_3}{\sqrt{2}} & h_2 \\ \frac{h_3}{\sqrt{2}} & 0 & -h_1 \\ -h_2 & h_1 & 0 \end{bmatrix} E_0 \right. \\
&\quad \left. -E_0 \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & \frac{h_3}{\sqrt{2}} & h_5 \\ -\frac{h_3}{\sqrt{2}} & 0 & -h_4 \\ -h_5 & h_4 & 0 \end{bmatrix} \right) V^\top \\
&= \frac{1}{\sqrt{2}} U \begin{bmatrix} 0 & -\sqrt{2}h_3 & -h_5 \\ \sqrt{2}h_3 & 0 & h_4 \\ -h_2 & h_1 & 0 \end{bmatrix} V^\top, \quad (39)
\end{aligned}
$$

which implies bijectivity in an obvious manner. This shows (i) and (ii).

To verify the third property we consider an arbitrary element

$$
\Gamma_\pm := \begin{bmatrix} R_\pm & 0 \\ 0 & \pm 1 \end{bmatrix} \in SO_3 \qquad (40)
$$

where

$$
R_\pm = \begin{bmatrix} c & -s \\ \pm s & \pm c \end{bmatrix} \quad \text{with} \quad c^2 + s^2 = 1, \quad (41)
$$

i.e. $R_+ \in SO_2$ is a rotation and $R_- \in O_2$ is a reflection, i.e. $\det R_- = -1$ holds. Therefore, $L : \mathcal{S} \to GL_5$, with $L(\Gamma_\pm) = L_\pm$

$$
\begin{aligned}
L_\pm : \mathbb{R}^5 &\to \mathbb{R}^5, \\
L_\pm = \pm &\begin{bmatrix} R_\pm & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & R_\pm \end{bmatrix}, \qquad (42)
\end{aligned}
$$

satisfies

$$
\begin{aligned}
\mu_{(U\Gamma_\pm, V\Gamma_\pm)}(x) &= U\Gamma_\pm \, e^{\Omega_1(x)} E_0 \, e^{-\Omega_2(x)} \Gamma_\pm^\top V^\top \\
&= U \, e^{\Omega_1(L_\pm x)} E_0 \, e^{-\Omega_2(L_\pm x)} V^\top \\
&= \mu_{(U,V)}(L_\pm x) \qquad (43)
\end{aligned}
$$

where $x \in \mathbb{R}^5$. Thus $\mu_{(U,V)}$ satisfies (i)–(iii).

Subsequently, we consider the essential manifold as an orbit of the group $SO_3 \times SO_3$ acting on E_0 by equivalence

$$
\begin{aligned}
SO_3 \times SO_3 \times \mathbb{R}^{3\times3} &\to \mathbb{R}^{3\times3}, \\
(U, V, E) &\mapsto UEV^\top. \qquad (44)
\end{aligned}
$$

By the derivative of this group action the usual canonical Riemannian metric on $SO_3 \times SO_3$ induces a Riemannian metric on the essential manifold which is called the normal Riemannian metric on $\mathcal{E}$ (see e.g. Helmke and Moore, 1994) for details about this construction. Moreover, by exploiting Corollary 3.1 one can show that via this group action geodesics on $SO_3 \times SO_3$ are mapped to geodesics on $\mathcal{E}$. We refer to Mahony (1994), Theorem 5.9.2, for a proof of this fact in a more general context. Moreover, in the Appendix we will discuss the main differences to the Riemannian metric introduced by Ma et al. (2001).

The mapping $\mu_{(U,V)}$ maps straight lines in $\mathbb{R}^5$ through the origin onto geodesics of $\mathcal{E}$

$$
t \mapsto U \, e^{\Omega_1(tx)} E_0 \, e^{-\Omega_2(tx)} V^\top \qquad (45)
$$

with respect to the normal Riemannian metric on $\mathcal{E}$. Thus $\mu_{(U,V)}^{-1}$ defines the Riemannian normal coordinates. Such a chart has the feature that the Riemannian metric expressed in this chart evaluated at zero is represented by the identity.

Example 3.2. In the same spirit as in Example 3.1 we can define the *Cayley parameterization*. Recall, that the Cayley transformation $\mathrm{Cay}(\Omega)$

$$
\begin{aligned}
\mathrm{Cay} : \mathfrak{so}_3 &\to SO_3, \\
\Omega &\mapsto \left(I + \tfrac{1}{2}\Omega\right)\left(I - \tfrac{1}{2}\Omega\right)^{-1}
\end{aligned}
$$

serves as a second order approximation to the matrix exponential e^Ω of skew-symmetric matrices. This Cayley mapping on $\mathfrak{so}_3$ is well known to be a local diffeomorphism around $0 \in \mathfrak{so}_3$. The Cayley parameterization is then

$$
\begin{aligned}
\mu_{(U,V)}^{\mathrm{Cay}} : \mathbb{R}^5 &\to \mathcal{E}, \\
\mu_{(U,V)}^{\mathrm{Cay}}(x) &:= U \, \mathrm{Cay}(\Omega_1(x)) E_0 \, \mathrm{Cay}(-\Omega_2(x)) V^\top. \qquad (46)
\end{aligned}
$$

Analogously to Example 3.1, one can show that $\mu_{(U,V)}^{\mathrm{Cay}}$ satisfies condition (i)–(iii) of Definition 3.1.

Example 3.3. The third example we consider is the *SVD parameterization*. Let $\Omega_1(x)$ and $\Omega_2(x)$ be defined as in (35) and (36), respectively. The orthogonal projection of $\mathbb{R}^{3\times3}$ onto $\mathcal{E}$ with respect to the Frobenius norm is achieved via the singular value decomposition

$$
\begin{aligned}
\pi^{\mathrm{SVD}} : \mathbb{R}^{3\times3} &\to \mathcal{E}, \\
X = U\Sigma V^\top &\mapsto UE_0V^\top, \qquad (47)
\end{aligned}
$$

where $U, V \in SO_3$ are computed such that the singular values $\Sigma = \mathrm{diag}(\sigma_1, \sigma_2, \sigma_3)$ are ordered as $\sigma_1 \geq \sigma_2 \geq$

$\sigma_3 \geq 0$. Thus define a map

$$\mu^{\mathrm{SVD}}_{(U,V)} : \mathbb{R}^5 \to \mathcal{E},$$
$$x \mapsto U\pi^{\mathrm{SVD}}(E_0 + \Omega_1(x)E_0 - E_0\Omega_2(x))V^\top. \tag{48}$$

Note that the smallest singular value of $E_0 + \Omega_1(x)E_0 - E_0\Omega_2(x)$ is simple for x in a sufficiently small neighborhood of 0 in $\mathbb{R}^5$. Thus $\mu^{\mathrm{SVD}}_{(U,V)}$ is smooth on a neighborhood of 0, as is seen by the following lemma, see the Appendix for a proof. The Properties (i)–(iii) of Definition 3.1 are then easily checked.

Lemma 3.2.   Let

$$\mathcal{U} := \{X \in \mathbb{R}^{3\times 3} \mid \text{smallest singular value is simple}\}. \tag{49}$$

$\mathcal{U} \subset \mathbb{R}^{3\times 3}$ is an open subset and the projection

$$\pi^{\mathrm{SVD}} : \mathcal{U} \to \mathcal{E},$$
$$X \mapsto \widehat{X}$$

is smooth.

3.2.   Cost Function

3.2.1. Cost Function on the Essential Manifold.   Let $M^{(i)} := \widehat{m}_2^{(i)}\widehat{m}_1^{(i)\top}$, where $\widehat{m}_1^{(i)}, \widehat{m}_2^{(i)} \in \mathbb{R}^3$ correspond to the normalized i-th point image pair in the left and in the right camera, respectively, for which the correspondence is assumed to be known. Our task is to develop a constrained least squares approach for solving the linear equations

$$\widehat{m}_1^{(i)\top}E\widehat{m}_2^{(i)} = 0, \quad i = 1, \dots, n. \tag{50}$$

It is desirable that the possibly large number of data points be preprocessed to achieve a useful compressed form. Thus let

$$\overline{M} := \begin{bmatrix} \mathrm{vec}^\top(M^{(1)\top}) \\ \vdots \\ \mathrm{vec}^\top(M^{(n)\top}) \end{bmatrix}, \quad \mathbb{R}^{9\times 9} \ni \mathcal{M} := \tfrac{1}{n}(\overline{M}^\top\overline{M}) \geq 0. \tag{51}$$

Consider the smooth function

$$f : \mathcal{E} \to \mathbb{R},$$
$$f(E) = \tfrac{1}{2n}\sum_{i=1}^n \left(\widehat{m}_1^{(i)\top}E\widehat{m}_2^{(i)}\right)^2$$
$$= \tfrac{1}{2n}\sum_{i=1}^n \mathrm{tr}^2\left(M^{(i)}E\right)$$
$$=: \tfrac{1}{2}\|\mathrm{vec}(E)\|_{\mathcal{M}}^2. \tag{52}$$

The global minimization of this cost function presents the main computational task addressed in this paper. Note that, according to the symmetry of f, with an essential matrix E also $-E$ is a critical point of f. Thus critical points of the cost function come in pairs $E, -E$ and therefore the number of critical points is even. Note, by the Mountain Pass Lemma and the nontrivial topology of $\mathcal{E}$, that f must have saddle points for generic data. These saddle points are local attractors for any Newton method, unless gradient descent modifications such as e.g. Gauss-Newton methods are used.

3.2.2. Noise-free Solution.   The value of this cost function attains zero if and only if there is an essential matrix which fulfills the epipolar constraint $\widehat{m}_1^{(i)\top}E\widehat{m}_2^{(i)} = 0$ for each image point pair. That is, in the noise free case the global minimum value is zero and yields the desired solutions. It should be clear also that in the noise free case $\mathrm{vec}(E)$ is an eigenvector corresponding to the zero eigenvalue of $\mathcal{M}$, and so there is a means to calculate this, as in the so-called normalized 8-point algorithm (Hartley and Zisserman, 2003), within a scale factor.

3.2.3. Closed Form Initial Estimate in the Noisy Case.
In the noisy case, which is of main interest in applications, there is in general no zero eigenvalue of $\mathcal{M}$, and a zero value cost function will not be attained. It nevertheless makes sense to search for a minimum of this cost function even in the presence of noise. Heuristically, the global minima can be expected to be least squares approximations to the true essential matrix solution, thus motivating the use of the 8-point algorithm for initialization. Thus one can start the optimization algorithm with an estimate $\widehat{E}$ of E, $\|\widehat{E}\| = \sqrt{2}$, such that $\mathrm{vec}(\widehat{E})$ is an eigenvector that corresponds to the smallest eigenvalue $\lambda_{\min}(\mathcal{M})$ of $\mathcal{M}$. In general, $\widehat{E}$ is not on the manifold $\mathcal{E}$, so we project $\widehat{E}$ to the manifold of essential matrices via a singular value decomposition. Thus we achieve a first estimate E_1 of E via

$$\mathcal{M}\mathrm{vec}(\widehat{E}) = \lambda_{\min}(\mathcal{M})\mathrm{vec}(\widehat{E}),$$
$$\widehat{E} = \widehat{U}\Sigma\widehat{V}^\top, \quad E_1 = \widehat{U}E_0\widehat{V}^\top, \tag{53}$$

that we use to initialize our algorithm. This is the normalized eight point algorithm solution (see also Hartley and Zisserman, 2003), Section 11.2. Note that the 8-point algorithm gives numerically meaningful results only in the case where the smallest eigenvalue is well separated from the others. This can be assumed, unless the noise level is too high.

3.2.4. Quadratic Model of the Cost Function.

The cost function $f : \mathcal{E} \to \mathbb{R}$ in a neighborhood of $E \in \mathcal{E}$ can be expressed in local parameter space coordinates (37) as follows.

$$f \circ \mu_{(U,V)} : \mathbb{R}^5 \to \mathbb{R},$$
$$f \circ \mu_{(U,V)}(x) = \tfrac{1}{2} \left\| \mathrm{vec}(U\, e^{\Omega_1(x)}\, E_0\, e^{-\Omega_2(x)}\, V^\top) \right\|_{\mathcal{M}}^2$$
$$= \tfrac{1}{2} \left\| (V \otimes U)\mathrm{vec}(e^{\Omega_1(x)}\, E_0\, e^{-\Omega_2(x)}) \right\|_{\mathcal{M}}^2.$$
(54)

Here "$\otimes$" and "vec" denote the matrix Kronecker product and vec-operation, that is stacking columns, respectively.

The second order Taylor polynomial of $f \circ \mu_{(U,V)}$ around $0 \in \mathbb{R}^5$ is

$$f \circ \mu_{(U,V)}(0) + \mathrm{D}\left(f \circ \mu_{(U,V)}\right)(0)h$$
$$+ \tfrac{1}{2}\, \mathrm{D}^2\left(f \circ \mu_{(U,V)}\right)(0)(h, h). \quad (55)$$

(i) The constant term is

$$\left(f \circ \mu_{(U,V)}\right)(0) = \tfrac{1}{2}\|\mathrm{vec}(E)\|_{\mathcal{M}}^2. \quad (56)$$

(ii) The linear term is

$$\tfrac{\mathrm{d}}{\mathrm{d}t} f \circ \mu_{(U,V)}(th)\Big|_{t=0}$$
$$= \mathrm{D}\left(f \circ \mu_{(U,V)}\right)(0)h$$
$$= \left(\nabla(f \circ \mu_{(U,V)})(0)\right)^\top \cdot h$$
$$= \langle \mathrm{grad}\, f(U E_0 V^\top), U(\Omega_1(h)E_0 - E_0\Omega_2(h))V^\top \rangle$$
$$= \mathrm{vec}^\top(E)\mathcal{M}\mathrm{vec}(U(\Omega_1(h)E_0 - E_0\Omega_2(h))V^\top).$$
(57)

Here, $\nabla(f \circ \mu_{(U,V)})(0)$ denotes the standard Euclidean gradient of $f \circ \mu_{(U,V)} : \mathbb{R}^5 \to \mathbb{R}$ evaluated at zero. Moreover,

$$\mathrm{grad}\, f(U E_0 V^\top)$$
$$= \mathrm{D}\,\mu_{(U,V)})(0) \cdot \nabla(f \circ \mu_{(U,V)})(0) \quad (58)$$

is the Riemannian gradient of $f : \mathcal{E} \to \mathbb{R}$ with respect to the normal Riemannian metric $\langle\,,\,\rangle$ on $\mathcal{E}$.

To calculate the Euclidean gradient and Hessian, respectively, let

$$\mathrm{vec}(\Omega_1(x)) := Q_1 x,$$
$$\mathrm{vec}(\Omega_2(x)) := Q_2 x,$$
(59)

where

$$Q_1 := \left[\mathrm{vec}\left(\tfrac{1}{\sqrt{2}}Q_x\right)\ \mathrm{vec}\left(\tfrac{1}{\sqrt{2}}Q_y\right)\ \mathrm{vec}\left(\tfrac{1}{2}Q_z\right)\ 0\ 0\right] \in \mathbb{R}^{9\times 5},$$
$$Q_2 := \left[0\ 0\ \mathrm{vec}\left(-\tfrac{1}{2}Q_z\right)\ \mathrm{vec}\left(\tfrac{1}{\sqrt{2}}Q_x\right)\ \mathrm{vec}\left(\tfrac{1}{\sqrt{2}}Q_y\right)\right] \in \mathbb{R}^{9\times 5},$$
(60)

and Q_x, Q_y, Q_z from (10). Then, the explicit formula for the Euclidean gradient of the local cost function at zero is

$$\nabla\left(f \circ \mu_{(U,V)}\right)(0) = J^\top \mathcal{M}\mathrm{vec}(E), \quad (61)$$

where

$$J := (V \otimes U)((E_0 \otimes I)Q_1 - (I \otimes E_0)Q_2). \quad (62)$$

(iii) The quadratic term in h consists of a sum of two terms. The first one,

$$\|\mathrm{vec}(U(\Omega_1(h)E_0 - E_0\Omega_2(h))V^\top)\|_{\mathcal{M}}^2$$
$$=: h^\top \widehat{\mathrm{H}}_{f\circ\mu_{(U,V)}}(0)h,$$

is a quadratic form on $\mathbb{R}^5$ with the corresponding matrix $\widehat{\mathrm{H}}_{f\circ\mu_{(U,V)}}(0)$ being positive (semi) definite for all $U, V \in SO_3$. Note, that by inspection $\widehat{\mathrm{H}}_{f\circ\mu_{(U,V)}}(0)$ is seen as positive definite if and only if the data matrix $\mathcal{M}$ is positive definite.

The second term is given as

$$h^\top \widetilde{\mathrm{H}}_{f\circ\mu_{(U,V)}}(0)h$$
$$:= \mathrm{vec}^\top(E)\mathcal{M}\mathrm{vec}\left(U\left(\Omega_1^2(h)E_0 + E_0\Omega_2^2(h)\right.\right.$$
$$\left.\left. - 2\Omega_1(h)E_0\Omega_2(h)\right)V^\top\right). \quad (63)$$

Hence, the Hessian matrix of the local cost function evaluated at zero is,

$$\mathrm{H}_{f\circ\mu_{(U,V)}}(0) = \widehat{\mathrm{H}}_{f\circ\mu_{(U,V)}}(0) + \widetilde{\mathrm{H}}_{f\circ\mu_{(U,V)}}(0), \quad (64)$$

where

$$\widehat{\mathrm{H}}_{f\circ\mu_{(U,V)}}(0) = J^\top \mathcal{M}J \geq 0, \quad (65)$$

and by denoting

$$\mathrm{vec}(D) := (V^\top \otimes U^\top)\mathcal{M}\mathrm{vec}(E) \quad (66)$$

we have

$$\widetilde{H}_{f \circ \mu_{(U,V)}}(0)$$
$$= \begin{bmatrix} Q_1^\top & Q_2^\top \end{bmatrix}$$
$$\begin{bmatrix} -(DE_0 \otimes I) & (D \otimes E_0) \\ (D^\top \otimes E_0) & -(E_0 D \otimes I) \end{bmatrix} \begin{bmatrix} Q_1 \\ Q_2 \end{bmatrix}. \quad (67)$$

The Newton direction for optimizing $f \circ \mu_{(U,V)}$ is

$$-H_{f \circ \mu_{(U,V)}}(0)^{-1} \nabla \big(f \circ \mu_{(U,V)}\big)(0) \quad (68)$$

while the Gauss-Newton direction coincides with the gradient descent

$$-\widehat{H}_{f \circ \mu_{(U,V)}}(0)^{-1} \nabla \big(f \circ \mu_{(U,V)}\big)(0). \quad (69)$$

The quadratic term in (55)

$$\frac{d^2}{dt^2} \big(f \circ \mu_{(U,V)}\big)(th)\big|_{t=0} = \mathcal{H}_{f(E)}(\dot\gamma(0), \dot\gamma(0)), \quad (70)$$

has the geometric interpretation that $\mathcal{H}_{f(E)}$ is the Riemannian Hessian form of $f : \mathcal{E} \to \mathbb{R}$ with respect to the normal Riemannian metric. Here $\gamma(t) = \mu_{(U,V)}(th)$ with $\gamma(0) = E$ denotes the geodesic emanating from E in direction $\dot\gamma(0) = D\mu_{(U,V)}(0)h$ (see e.g. Lang, 1999), pp. 342.

4. The Algorithm

The Newton-type algorithm we propose is defined by iterating a map

$$s : \mathcal{E} \to \mathcal{E}. \quad (71)$$

For any $U, V \in SO_3$ let $x^{\text{opt}}(U, V)$ denote the solution of

$$H_{f \circ \mu_{(U,V)}}(0) x = -\nabla \big(f \circ \mu_{(U,V)}\big)(0). \quad (72)$$

Alternatively, let $x_{\text{GN}}^{\text{opt}}(U, V)$ denote the solution of

$$\widehat{H}_{f \circ \mu_{(U,V)}}(0) \cdot x = -\nabla \big(f \circ \mu_{(U,V)}\big)(0), \quad (73)$$

i.e. the *Gauss-Newton* direction.

The algorithmic map s in the first case is given as

$$s(U E_0 V^\top) = \nu_{(U,V)}(x^{\text{opt}}(U, V)) \quad (74)$$

where

$$\nu_{(U,V)} \in \big\{ \mu_{(U,V)}, \mu_{(U,V)}^{\text{Cay}}, \mu_{(U,V)}^{\text{SVD}} \big\} \quad (75)$$

denotes one of the three previously introduced local parameterizations. We now describe the implementation details of the proposed algorithm. First pre-process the data to form a 9×9 matrix $\mathcal{M}$ as in (51). Let $\lambda_{\min}(A)$ denote the smallest eigenvalue of the symmetric matrix A.

Step 0.
    Determine an initial estimate of the essential matrix $E_1 = U_1 E_0 V_1^\top$.
Step 1.
    Set $k = 1$ and set $\varepsilon > 0$ to prescribed accuracy.
Step 2. (**Newton step**) Given (U_k, V_k) with $E = U_k E_0 V_k^\top$.
    Compute the gradient $\nabla(f \circ \mu_{(U_k, V_k)})(0) \in \mathbb{R}^5$.
    Compute the Hessian $H_{f \circ \mu_{(U_k, V_k)}}(0) \in \mathbb{R}^{5 \times 5}$.
    If $\lambda_{\min}\big(H_{f \circ \mu_{(U_k, V_k)}}(0)\big) > \varepsilon$

    solve for x_k^{opt}:

$$H_{f \circ \mu_{(U_k, V_k)}}(0) x_k^{\text{opt}} = -\nabla\big(f \circ \mu_{(U_k, V_k)}\big)(0).$$

    Else

    solve for x_k^{opt}:

$$\widehat{H}_{f \circ \mu_{(U_k, V_k)}}(0) x_k^{\text{opt}} = -\nabla\big(f \circ \mu_{(U_k, V_k)}\big)(0).$$

Step 3. (**Projection back onto $\mathcal{E}$**):
    Set $U_{k+1} = U_k e^{\Omega_1(x_k^{\text{opt}})}, \quad V_{k+1} = V_k e^{\Omega_2(x_k^{\text{opt}})}$.
    Set $E_{k+1} = U_{k+1} E_0 V_{k+1}^\top$.
Step 4.
    If $\|\nabla\big(f \circ \mu_{(U_{k+1}, V_{k+1})}\big)(0)\| < \varepsilon$

    terminate.

    Else

    Set $U_k = U_{k+1}, \quad V_k = V_{k+1}$.
    Set $E_k = E_{k+1}$.

Step 5.
    Set $k = k + 1$ and goto Step 2.

Note that the above algorithm implements the Riemannian Newton algorithm for the normal metric, at least on a sufficiently small neighborhood of the set of local minima of f. Outside of such a neighborhood, where the Riemannian Hessian is either nearly degenerate of indefinite, we switch to a Riemannian gradient algorithm, using the Riemannian metric defined by the approximate Hessian $\widehat{H}$. Thus s implements effectively a Gauss-Newton algorithm on the manifold $\mathcal{E}$. This modification of the Newton algorithm is of crucial importance for enlarging the domain of attraction of the local minima.

Alternative descriptions of the algorithm replace the exponential matrix in the update step 3 by Cayley transforms or SVD substitutions; see Section 4.1.

4.1. Geometric Interpretation of the Algorithm

We show that the above algorithm has the geometric interpretation of iterating the map

$$s = \pi_2 \circ \pi_1 : \mathcal{E} \to \mathcal{E}, \tag{76}$$

consisting of an optimization step π_1 followed by a projection π_2. The optimization step that appears in the algorithm defines a map

$$\pi_1 : \mathcal{E} \to T_E^{\mathrm{aff}}\mathcal{E} \subset \mathbb{R}^{3\times 3},$$
$$E = U E_0 V^\top \mapsto U(E_0 + \Omega_1(x^{\mathrm{opt}}(U, V))E_0 - E_0\Omega_2(x^{\mathrm{opt}}(U, V)))V^\top. \tag{77}$$

Here, $x^{\mathrm{opt}} \in \mathbb{R}^5$ is a function of (U, V) and is given by the solution of a Newton step. The mapping π_2 involves one of three projections defined below, with relative merits as discussed in the Simulation section.

4.1.1. Orthogonal Projection.   The orthogonal projection with respect to the Frobenius norm

$$\pi_2^{\mathrm{SVD}} : T_E^{\mathrm{aff}}\mathcal{E} \subset \mathbb{R}^{3\times 3} \to \mathcal{E},$$
$$X = U\Sigma V^\top \mapsto U E_0 V^\top, \tag{78}$$

is given via the singular value decomposition, where $U, V \in SO_3$ are defined such that the singular values $\Sigma = \mathrm{diag}(\sigma_1, \sigma_2, \sigma_3)$ of X are ordered as $\sigma_1 \geq \sigma_2 \geq \sigma_3 \geq 0$.

4.1.2. Exponential Map Projection.   Let Ω_1 and Ω_2 be defined as in (35) and (36), respectively. Note that for $x \in \mathbb{R}^5$ the element $U(E_0 + \Omega_1(x)E_0 - E_0\Omega_2(x))V^\top$ paramterizes arbitrary affine tangent vectors and therefore defines a smooth mapping

$$\pi_2^{\mathrm{exp}} : T_E^{\mathrm{aff}}\mathcal{E} \to \mathcal{E},$$
$$U(E_0 + \Omega_1(x)E_0 - E_0\Omega_2(x))V^\top \mapsto U\, e^{\Omega_1(x)} E_0\, e^{-\Omega_2(x)} V^\top$$
$$= \mu_{(U,V)}(x). \tag{79}$$

4.1.3. Cayley Projection.   Let Ω_1 and Ω_2 be defined as in (35) and (36), respectively. Consider the smooth mapping

$$\pi_2^{\mathrm{Cay}} : T_E^{\mathrm{aff}}\mathcal{E} \to \mathcal{E},$$
$$U(E_0 + \Omega_1(x)E_0 - E_0\Omega_2(x))V^\top \mapsto U\, \mathrm{Cay}(\Omega_1(x))E_0$$
$$\mathrm{Cay}(-\Omega_2(x))V^\top$$
$$= \mu_{(U,V)}^{\mathrm{Cay}}(x). \tag{80}$$

In summary, one algorithmic step of s consists of two partial steps, namely π_1 sending a point E on the essential manifold $\mathcal{E}$ to an element of the affine tangent space $T_E^{\mathrm{aff}}\mathcal{E}$, followed by π_2 that projects that element back to $\mathcal{E}$. Our key theoretical contribution is a local quadratic convergence result for the Newton method, assumed to be started in the domain of attraction of a local minimum.

4.2. Local Quadratic Convergence

Before stating and proving the main convergence result of this paper we begin with some preliminaries. Consider the local parameterization

$$\mu_{(U,V)} : \mathbb{R}^5 \to \mathcal{E},$$
$$\mu_{(U,V)}(x) := U\, e^{\Omega_1(x)} E_0\, e^{-\Omega_2(x)} V^\top \tag{81}$$

which is smooth with respect to $(U, V) \in SO_3 \times SO_3$ and $x \in \mathbb{R}^5$. Moreover, it is a local diffeomorphism at $0 \in \mathbb{R}^5$. Let $\mathcal{U} \subset \mathbb{R}^5$ be an open neighborhood of 0. The Newton step is defined as the map

$$\mathcal{N}\big(f \circ \mu_{(U,V)}\big) : \mathcal{U} \to \mathbb{R}^5,$$
$$x \mapsto \big(\mathrm{H}_{f\circ\mu_{(U,V)}}(x)\big)^{-1} \nabla\big(f \circ \mu_{(U,V)}\big)(x). \tag{82}$$

Here, $\nabla g(x)$ and $\mathrm{H}_g(x)$ denote the standard gradient vector and Hessian matrix, respectively, of the smooth function $g : \mathbb{R}^5 \to \mathbb{R}$. In the sequel, we will always make the assumption that the Hessian $\mathrm{H}_{f\circ\mu_{(U,V)}}(0)$ at the point of interest (U, V) is invertible. Certainly, this is locally true around a nondegenerate critical point of f.

Consider the map

$$\bar{s} : SO_3 \times SO_3 \to \mathcal{E},$$
$$\bar{s}(U, V) := \mu_{(U,V)}\big(\mathcal{N}\big(f \circ \mu_{(U,V)}\big)(0)\big) \tag{83}$$

and the corresponding algorithmic map

$$s : \mathcal{E} \to \mathcal{E},$$
$$s(U E_0 V^\top) := \bar{s}(U, V). \tag{84}$$

We define

$$\Gamma_\pm := \begin{bmatrix} R_\pm & 0 \\ 0 & \pm 1 \end{bmatrix} \in SO_3 \tag{85}$$
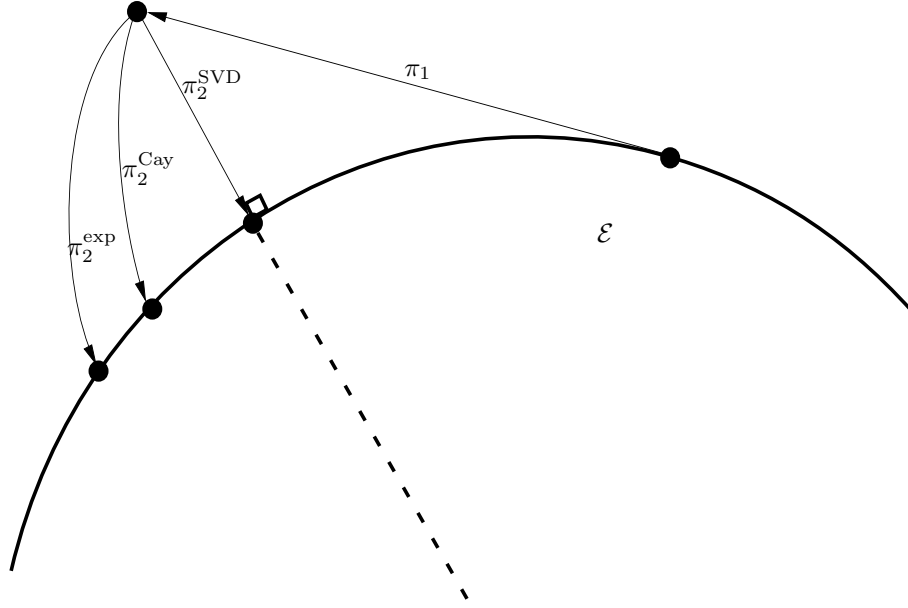
Figure 3.   Three different projection steps.

where

$$R_{\pm} = \begin{bmatrix} c & -s \\ \pm s & \pm c \end{bmatrix} \quad \text{with} \quad c^2 + s^2 = 1, \quad (86)$$

i.e. $R_+ \in SO_2$ is a rotation and $R_- \in O_2$ is a reflection of determinant $\det R_- = -1$. Using $\Omega_1, \Omega_2 : \mathbb{R}^5 \to \mathfrak{so}_3$, defined by (35) and (36), respectively, we have

$$\Gamma_{\pm}\Omega_1(x)\Gamma_{\pm}^{\top} = \frac{1}{\sqrt{2}} \left[\begin{array}{cc|c} 0 & \mp\frac{x_3}{\sqrt{2}} & \pm R_{\pm} \cdot \begin{bmatrix} x_2 \\ -x_1 \end{bmatrix} \\ \pm\frac{x_3}{\sqrt{2}} & 0 & \\ \hline \mp[x_2 \ -x_1] \cdot R_{\pm}^{\top} & & 0 \end{array} \right]$$

(87)

$$\Gamma_{\pm}\Omega_2(x)\Gamma_{\pm}^{\top} = \frac{1}{\sqrt{2}} \left[\begin{array}{cc|c} 0 & \pm\frac{x_3}{\sqrt{2}} & \pm R_{\pm} \cdot \begin{bmatrix} x_5 \\ -x_4 \end{bmatrix} \\ \mp\frac{x_3}{\sqrt{2}} & 0 & \\ \hline \mp[x_5 \ -x_4] \cdot R_{\pm}^{\top} & & 0 \end{array} \right].$$

Therefore,

$$\Gamma_{\pm}\Omega_i(x)\Gamma_{\pm}^{\top} = \Omega_i(L_{\pm}x), \quad \text{for } i = 1, 2, \quad (88)$$

where

$$L_{\pm} : \mathbb{R}^5 \to \mathbb{R}^5,$$

$$L_{\pm} = \pm \begin{bmatrix} R_{\pm} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & R_{\pm} \end{bmatrix}. \quad (89)$$

Thus we obtain for any $\Gamma_{\pm}$ and $L_{\pm}$ as in (85) and (89), respectively, the invariance property

$$\begin{aligned} \mu_{(U\Gamma_{\pm}, V\Gamma_{\pm})}(x) &= U\Gamma_{\pm} \, e^{\Omega_1(x)} \, E_0 \, e^{-\Omega_2(x)} \, \Gamma_{\pm}^{\top} V^{\top} \\ &= U \, e^{\Omega_1(L_{\pm}x)} \, E_0 \, e^{-\Omega_2(L_{\pm}x)} \, V^{\top} \\ &= \mu_{(U,V)}(L_{\pm}x). \end{aligned} \quad (90)$$

We can now state the following lemma.

Lemma 4.1. *The algorithmic map s is well defined and smooth on an open neighborhood of the set of nondegenerate critical points of f.*

Proof: Using (90) we have

$$U E_0 V^{\top} = \bar{U} E_0 \bar{V}^{\top} \iff (\bar{U}, \bar{V}) = (U\Gamma_{\pm}, V\Gamma_{\pm}). \quad (91)$$

Note that for any $\Gamma_{\pm}$ and $L_{\pm}$ as in (85) and (89), respectively, the gradient and Hessians are

$$\begin{aligned} \nabla\big(f \circ \mu_{(U\Gamma_{\pm}, V\Gamma_{\pm})}\big)(x) &= L_{\pm}^{\top} \cdot \nabla\big(f \circ \mu_{(U,V)}\big)(L_{\pm}x), \\ \mathrm{H}_{f\circ\mu_{(U\Gamma_{\pm}, V\Gamma_{\pm})}}(x) &= L_{\pm}^{\top} \cdot \mathrm{H}_{f\circ\mu_{(U,V)}}(L_{\pm}x) \cdot L_{\pm}. \end{aligned} \quad (92)$$

Therefore

$$\mathcal{N}\big(f \circ \mu_{(U\Gamma_{\pm}, V\Gamma_{\pm})}\big)(0) = L_{\pm}^{\top} \cdot \mathcal{N}\big(f \circ \mu_{(U,V)}\big)(0) \quad (93)$$

and thus

$$
\begin{aligned}
\bar{s}(U\Gamma_\pm, V\Gamma_\pm) &= U\Gamma_\pm\, e^{\Omega_1(L_\pm^\top \mathcal{N}(f\circ\mu_{(U,V)})(0))} \\
&\quad \times E_0\, e^{-\Omega_2(L_\pm^\top \mathcal{N}(f\circ\mu_{(U,V)})(0))}\, \Gamma_\pm^\top V^\top \\
&= U\Gamma_\pm \Gamma_\pm^\top\, e^{\Omega_1(\mathcal{N}(f\circ\mu_{(U,V)})(0))} \\
&\quad \times E_0\, e^{-\Omega_2(\mathcal{N}(f\circ\mu_{(U,V)})(0))}\, \Gamma_\pm \Gamma_\pm^\top V^\top \\
&= \bar{s}(U, V).
\end{aligned}
\tag{94}
$$

This shows that s is well defined. Smoothness of s follows from the smoothness of $\bar{s}$ and the fact that

$$
\begin{aligned}
\pi : SO_3 \times SO_3 &\to \mathcal{E}, \\
(U, V) &\mapsto U E_0 V^\top
\end{aligned}
\tag{95}
$$

is a principle bundle with commuting diagram

From now on we assume that

$$
E_* := U_* E_0 V_*^\top
\tag{96}
$$

is a nondegenerate critical point of

$$
f : \mathcal{E} \to \mathbb{R}.
\tag{97}
$$

Recall, that

$$
\begin{aligned}
D\,\mu_{(U_*, V_*)}(0) : \mathbb{R}^5 &\to T_{E_*}\mathcal{E}, \\
x &\mapsto U_*(\Omega_1(x)E_0 - E_0\Omega_2(x))V_*^\top
\end{aligned}
\tag{98}
$$

is a linear isomorphism. Moreover,

$$
\begin{aligned}
\bar{\mu}_{(U_*, V_*)} : \mathbb{R}^5 &\to SO_3 \times SO_3, \\
x &\mapsto \left(U_* e^{\Omega_1(x)}, V_* e^{\Omega_2(x)}\right)
\end{aligned}
\tag{99}
$$

is smooth and the derivative

$$
D\,\bar{\mu}_{(U_*, V_*)}(0) : \mathbb{R}^5 \to T_{(U_*, V_*)}(SO_3 \times SO_3)
\tag{100}
$$

maps $\mathbb{R}^5$ linearly and isomorphically onto the 5-dimensional linear subspace

$$
\mathcal{T} := \{(U_*\Omega_1(x), V_*\Omega_2(x)) | x \in \mathbb{R}^5\}.
\tag{101}
$$

In the subsequent proof of local quadratic convergence of the algorithm we need to compute the Fréchet derivative of the map

$$
(U, V) \mapsto \nabla\big(f \circ \mu_{(U,V)}\big)(0) \in \mathbb{R}^5 \quad \text{at} \quad (U_*, V_*).
\tag{102}
$$

This Fréchet derivative is denoted as

$$
D_{(U,V)} \nabla\big(f \circ \mu_{(U,V)}\big)(0)\big|_{(U_*,V_*)} : T_{(U_*,V_*)}(SO_3 \times SO_3) \to \mathbb{R}^5.
\tag{103}
$$

Lemma 4.2. *For any $x \in \mathbb{R}^5$ one has*

$$
\begin{aligned}
&D_{(U,V)}\nabla\big(f \circ \mu_{(U,V)}\big)(0)\big|_{(U_*,V_*)}(U_*\Omega_1(x), V_*\Omega_2(x)) \\
&= H_{f\circ\mu_{(U_*,V_*)}}(0)x
\end{aligned}
\tag{104}
$$

Proof: For any $y, z \in \mathbb{R}^5$ and $(U, V) \in SO_3 \times SO_3$

$$
\begin{aligned}
Df\big(\mu_{(U,V)}(y)\big) \cdot D\,\mu_{(U,V)}(y) \cdot z &= D\big(f \circ \mu_{(U,V)}\big)(y) \cdot z \\
&= \big\langle \nabla\big(f \circ \mu_{(U,V)}\big)(y), z\big\rangle.
\end{aligned}
\tag{105}
$$

Differentiating (105) with respect to y at $y = 0$ we get for any $x \in \mathbb{R}^5$

$$
\begin{aligned}
&D^2 f(E_*)\big(D\,\mu_{(U_*,V_*)}(0) \cdot z, D\,\mu_{(U_*,V_*)}(0) \cdot x\big) \\
&= \big\langle H_{f\circ\mu_{(U_*,V_*)}}(0)x, z\big\rangle,
\end{aligned}
\tag{106}
$$

while differentiating (105) for $y = 0$ with respect to (U, V) at $(U, V) = (U_*, V_*)$ in direction $(\dot{U}, \dot{V}) := (U_*\Omega_1(x), V_*\Omega_2(x))$ yields

$$
\begin{aligned}
&\big\langle D_{(U,V)} \nabla(f \circ \mu_{(U,V)})(0)\big|_{(U_*,V_*)}(\dot{U}, \dot{V}), z\big\rangle \\
&= D^2 f(E_*)\big(D_{(U,V)}\,\mu_{(U,V)}(0)\big|_{(U_*,V_*)}(\dot{U}, \dot{V}), \\
&\qquad D\,\mu_{(U_*,V_*)}(0) \cdot z\big) \\
&= D^2 f(E_*)\big(D\,\mu_{(U_*,V_*)}(0) \cdot x, D\,\mu_{(U_*,V_*)}(0) \cdot z\big),
\end{aligned}
\tag{107}
$$

since

$$
\mu_{(U,V)}(0) = U E_0 V^\top
\tag{108}
$$

and therefore

$$
\begin{aligned}
&D_{(U,V)}\,\mu_{(U,V)})(0)\big|_{(U_*,V_*)}(U_*\Omega_1(x), V_*\Omega_2(x)) \\
&= U_*(\Omega_1(x)E_0 - E_0\Omega_2(x))V_*^\top \\
&= D\,\mu_{(U_*,V_*)}(0) \cdot x.
\end{aligned}
\tag{109}
$$

Since $x, z \in \mathbb{R}^5$ are arbitrary the result follows. $\square$

We can now state and prove the main local quadratic convergence result.

Theorem 4.1. *Let $E_* = U_* E_0 V_*^\top \in \mathcal{E}$ be a nondegenerate critical point of $f : \mathcal{E} \to \mathbb{R}$. Then the Fréchet derivative of the algorithmic map $s : \mathcal{E} \to \mathcal{E}$ satisfies*

$$\mathrm{D}\, s(E_*) = 0 \tag{110}$$

and therefore our algorithm converges locally quadratically fast to E_*.

Proof: The tangent map

$$\mathrm{D}\,\pi(U_*, V_*) : T_{(U_*, V_*)}(SO_3 \times SO_3) \to T_{E_*}\mathcal{E} \tag{111}$$

maps $\mathcal{T}$ defined by (101) linearly isomorphically onto $T_{E_*}\mathcal{E}$. Thus $\mathrm{D}\, s(E_*) = 0$ if and only if

$$\mathrm{D}\,\bar{s}(U_*, V_*)(U_*\Omega_1(x), V_*\Omega_2(x)) = 0 \tag{112}$$

for all $x \in \mathbb{R}^5$. By differentiating

$$\bar{s}(U, V) = \mu_{(U,V)}\big(\mathcal{N}(f \circ \mu_{(U,V)})(0)\big) \tag{113}$$

with respect to (U, V) at (U_*, V_*) we get for all $x \in \mathbb{R}^5$

$$
\begin{aligned}
&\mathrm{D}\,\bar{s}(U_*, V_*)(U_*\Omega_1(x), V_*\Omega_2(x))\\
&= \left. \mathrm{D}_{(U,V)}\, \mu_{(U,V)}(0)\right|_{(U_* V_*)} (U_*\Omega_1(x), V_*\Omega_2(x))\\
&\quad + \left. \mathrm{D}\,\mu_{(U_*, V_*)}(0)\, \mathrm{D}_{(U,V)}\, \mathcal{N}(f \circ \mu_{(U,V)})(0)\right|_{(U_* V_*)}\\
&\qquad (U_*\Omega_1(x), V_*\Omega_2(x))\\
&= U_*(\Omega_1(x) E_0 - E_0 \Omega_2(x)) V_*^\top - \mathrm{D}\,\mu_{(U_*, V_*)}(0)\\
&\qquad \Big((\mathrm{H}_{f \circ \mu_{(U_*, V_*)}}(0))^{-1} \left. \mathrm{D}_{(U,V)}\, \nabla(f \circ \mu_{(U,V)})(0)\right|_{(U_* V_*)}\\
&\qquad (U_*\Omega_1(x), V_*\Omega_2(x)) \Big)\\
&= U_*(\Omega_1(x) E_0 - E_0 \Omega_2(x)) V_*^\top\\
&\quad - \mathrm{D}\,\mu_{(U_*, V_*)}(0) \left((\mathrm{H}_{f \circ \mu_{(U_*, V_*)}}(0))^{-1} \mathrm{H}_{f \circ \mu_{(U_*, V_*)}}(0) \cdot x \right)\\
&= U_*(\Omega_1(x) E_0 - E_0 \Omega_2(x)) V_*^\top\\
&\quad - \mathrm{D}\,\mu_{(U_*, V_*)}(0) \cdot x\\
&= 0.
\end{aligned}
$$

The result follows by a standard Taylor series argument.
 The same argument holds for

$$\bar{s}(U, V) = \nu_{(U,V)}\big(\mathcal{N}(f \circ \mu_{(U,V)})(0)\big) \tag{114}$$

where $\nu_{(U,V)}$ denotes one of the three maps defined in (75). This can easily be verified using the fact that

$$\mathrm{D}\,\mu_{(U_*, V_*)}(0) = \mathrm{D}\,\nu_{(U_*, V_*)}(0). \tag{115}$$

$\square$

It should be clear from our development that the particular parameterization scheme of Ma et al. (2001) can also be implemented using our more general homogeneous space approach. We have not done this here as it would not lead to any substantial improvement. However, there is substantial benefit from using Gauss-Newton iterations on the manifold when the Hessian is not positive semidefinite, and numerical advantage from using the SVD or Cayley projections, as well as the computational advantages of essentially preprocessing the data to one 9×9 matrix, irrespective of the number of data points.

In subsequent implementations of the Ma et al. algorithm we found it necessary to use the Gaussian step when the Hessian is not positive definite to achieve sensible results, since simulation experiments showed otherwise that the domain of attraction of the Newton method can be surprisingly small.

5.   Simulations

All simulations are carried out in Matlab. Each trial:

- Generates 3D points randomly in a field of view (FOV) $60°$ and depth varying from 100 to 400 units of focal length,
- Computes the first projection matrix $P = K[I_3 \ \mathbf{0}]$, where K is the camera calibration matrix,
- Randomly generates a rotation R and a translation $\mathbf{t}$,
- Computes the second projection matrix $P' = K[R \ \mathbf{t}]$,
- Projects the 3D points onto the two 512×512 image planes using P and P'. Only visible points will be used in the algorithm,
- Adds Gaussian noise of mean zero and standard deviation σ to the image points,
- Normalizes the image points with K and solves for essential matrix E,
- Stops when $\|\nabla(f \circ \mu)(0)\| < 10^{-12}$,
- Computes the symmetric epipolar distance, i.e., the average absolute distance between a point's epipolar line and its corresponding point in other image, defined as

$$
\frac{1}{n} \sum_{i=1}^{n} \left| \frac{m_2^{(i)^\top} E^\top m_1^{(i)}}{\sqrt{\left(e_1^\top E^\top m_1^{(i)}\right)^2 + \left(e_2^\top E^\top m_1^{(i)}\right)^2}} \right|
$$
$$
+ \left| \frac{m_1^{(i)^\top} E m_2^{(i)}}{\sqrt{\left(e_1^\top E m_2^{(i)}\right)^2 + \left(e_2^\top E m_2^{(i)}\right)^2}} \right|,
$$

where

$$
e_1 := \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad e_2 := \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}.
$$

In the following, the algebraic cost function is denoted f as defined in (52). We also work with the weighted measure,

$$f_w(E) = \frac{1}{2n} \sum_{i=1}^{n} w^{(i)}(E)^2 \operatorname{tr}^2 \left(M^{(i)} E \right)$$

$$= \frac{1}{2} \|\operatorname{vec}(E)\|_{\mathcal{M}_w}^2, \qquad (116)$$

where

$$W := \operatorname{diag}\{w_1, w_2, \ldots, w_n\},$$
$$\mathcal{M}_w := \frac{1}{n} (\bar{M}^\top W \bar{M}) \geq 0 \in \mathbb{R}^{9 \times 9}. \qquad (117)$$

Here the weighting is ideally given in terms of the essential matrix E, but this complicates the calculation of the gradient and Hessian to such an extent that we here merely express the weightings "adaptively" at each iteration as the current estimate $\widehat{E}$. The term f_{dist} denotes the cost function that penalizes the socalled square symmetric epipolar distance adaptively, as

$$f_{\text{dist}}(E, \widehat{E}) = \frac{1}{2n} \sum_{i=1}^{n} w_{\text{dist}}^{(i)}(\widehat{E})^2 \operatorname{tr}^2 \left(M^{(i)} E \right), \quad (118)$$

where, denoting the previous estimate of E by $\widehat{E}$,

$$w_{\text{dist}}^{(i)}(\widehat{E})^2 = \left((e_1^\top \widehat{E} m_2^{(i)})^2 + (e_2^\top \widehat{E} m_2^{(i)})^2 \right)^{-1}$$
$$+ \left((e_1^\top \widehat{E}^\top m_1^{(i)})^2 + (e_2^\top \widehat{E}^\top m_1^{(i)})^2 \right)^{-1}.$$

Finally, the term f_{grad} denotes a weighted version of the gradient based nonlinear cost function,

$$f_{\text{grad}}(E, \widehat{E}) = \frac{1}{2n} \sum_{i=1}^{n} w_{\text{grad}}^{(i)}(\widehat{E})^2 \operatorname{tr}^2 \left(M^{(i)} E \right) \quad (119)$$

where

$$w_{\text{grad}}^{(i)}(\widehat{E})^2 = \left((e_1^\top \widehat{E} m_2^{(i)})^2 + (e_2^\top \widehat{E} m_2^{(i)})^2 \right.$$
$$\left. + (e_1^\top \widehat{E}^\top m_1^{(i)})^2 + (e_2^\top \widehat{E}^\top m_1^{(i)})^2 \right)^{-1} \quad (120)$$

Of course, we caution that this bootstrapping approach is not necessarily guaranteed to converge since the cost function changes from iteration to iteration.

We denote by "8p" the standard 8-point algorithm referred to as above. For comparison purposes, we implement an enhanced version of the Riemannian Newton algorithm proposed by Ma et al. (2001), which exploits a Gauss-Newton step when the Hessian is not positive definite. This modified version of their algorithm is termed "MaGN" here. Furthermore, "GNsvd", "GNexp", "GNcay" stand for our proposed Gauss-Newton on manifold algorithms with orthogonal, exponential and Cayley projection, respectively.

Sensitivity Analysis. The number of point correspondences is fixed at 20 and Gaussian noise of varying standard deviation $\sigma \in [0.5, 5]$ is introduced. The simulation is repeated 100 times and the averaged results are shown in Fig. 4. It is observed that the standard 8-point method (dotted line) is relatively sensitive to noise compared to
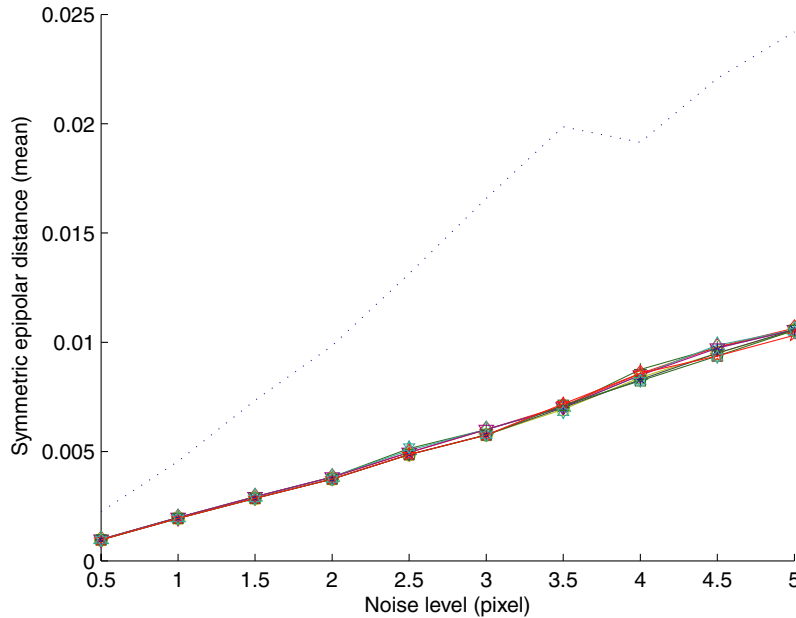

Figure 4. Symmetric epipolar distance (mean) vs. noise level for f–8p algorithm (dotted line) and other algorithms (f–MaGN, f–GNsvd, f–GNexp, f–GNcay, f_{grad}–GNsvd, f_{grad}–GNexp, f_{grad}–GNcay, f_{dist}–GNsvd, f_{dist}–GNexp, f_{dist}–GNcay).
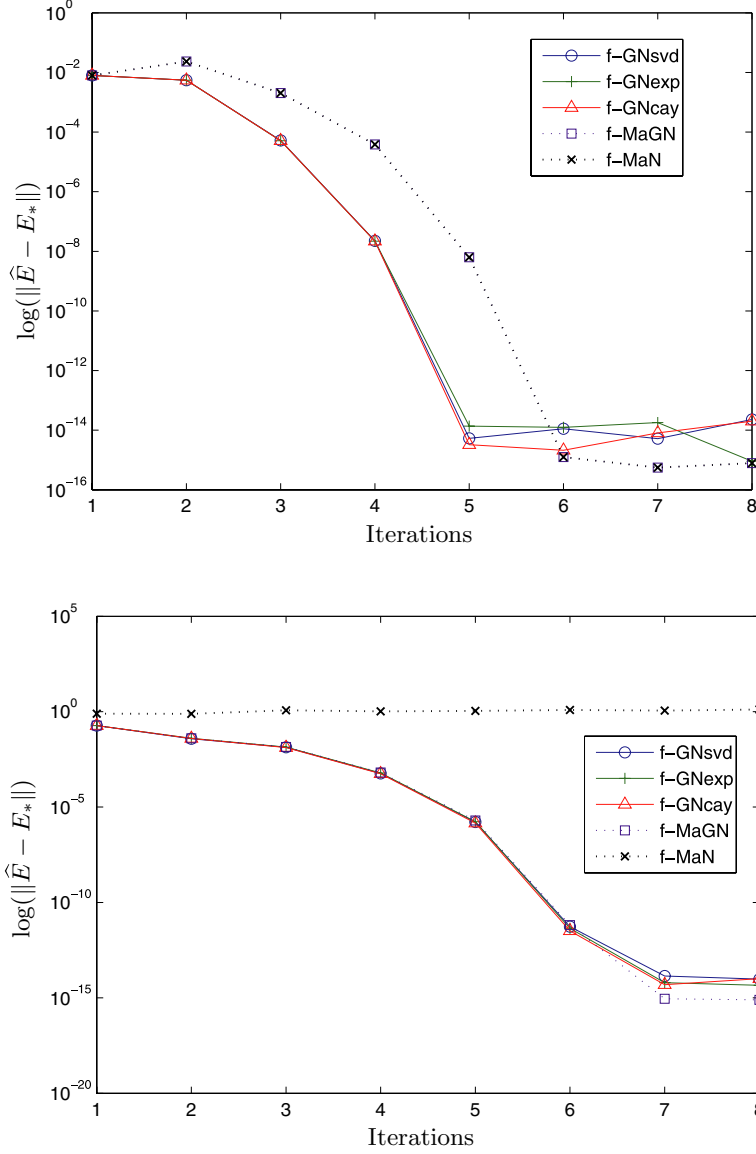
Figure 5. $\log(\|\widehat{E} - E_*\|)$ vs. iterations.

other methods in minimizing the cost function (52). The plot also shows that all three proposed algorithms and "MaGN" have equal sensitivity properties.

Rate of Local Convergence.   The number of point correspondences is fixed at 20 and noise level $\sigma = 2$. Figure 5 illustrates the speed at which the estimate $\widehat{E}$ of various algorithm converges to its corresponding fixed point E_*. It is clear that all three proposed techniques, together with the "MaGN" method converge at a local quadratic rate. The algorithm "MaN" in Fig. 5(b) is the Riemannian Newton presented by Ma et al. It illustrates that for a certain initial guess, "MaN" does not converge to the global minimum.

Computing Time.   The number of matched points is fixed at 20, noise level $\sigma = 2$ and the simulation is repeated 100 times. Figure 6 shows the average computing time (implemented using Matlab's *tic* and *toc* commands) using different algorithms and minimizing different cost functions. When the algebraic cost function f is used, the standard 8-point method is the fastest technique. This is expected since it does not involve iterations but of course achieves an approximate solution only, as shown in Fig. 4.

Although all three proposed methods are slightly slower than the standard 8-point method, they are at least 25 times faster than 'MaGN'. Moreover, our proposed methods minimizing adaptive cost function f_{dist} and f_{grad} are at least 3 times faster than "MaGN" using the function f. Note, that although the "MaGN" method is locally
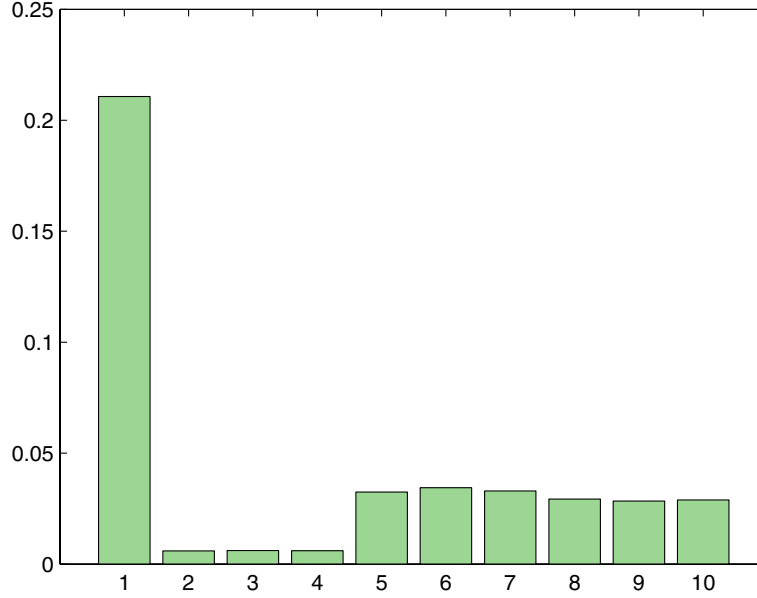
Figure 6. Computing time for various methods minimizing different cost functions: (1) f–MaGN, (2) f–GNsvd, (3) f–GNexp, (4) f–GNcay, (5) f_{grad}–GNsvd, (6) f_{grad}–GNexp, (7) f_{grad}–GNcay, (8) f_{dist}–GNsvd, (9) f_{dist}–GNexp, (10) f_{dist}–GNcay.

quadratically convergent as is ours (Fig. 5), its computational effort per iteration is data dependent, and so can be prohibitive for large data sets.

6. Conclusions

This paper puts forward a new geometric optimization framework, new to the area of computer vision, to compute the essential matrix between two calibrated intensity images, which subsequently allows the pose information to be determined. The approach is based on successive local parameterization of the constraint manifold. Research contributions of this paper are:

- The algorithm uses a prior data compression technique, so its computational effort per iteration is essentially independent of the number of point correspondences used.
- The iterations use Gauss or Gauss/gradient steps as well as the Newton step, which is a well known approach for optimizing in Euclidean space, but appears relatively new for optimizing on manifolds. We found it necessary to apply to the Ma et al. algorithm (Ma et al., 2001), in order to achieve convergence in the presence of reasonable noise.
- Our algorithm achieves the same local quadratic convergence rate as Ma et al. (2001), but our computational effort per iteration is significantly lower.
- We have proposed algorithms with three different projections: the exponential map, the Cayley transformation and the orthogonal projection.

- We prove local quadratic convergence rate of the proposed algorithm. This goes beyond earlier proofs for intrinsic Newton methods on manifolds, as we use parameterizations that are not necessarily induced by any Riemannian normal coordinates.

Appendix

A.1. Comparison with the Approach by Ma et al.

If π_2^{exp} is used for the second algorithmic step π_2 then one can easily show that the overall algorithm is nothing else than a Riemannian manifold version of Newton's method, the Riemannian metric being the so-called normal one. Despite the well-known fact that under mild assumptions, the Riemannian manifold version of Newton's method is locally quadratically convergent (see Smith, 1994), Theorem 3.4, p. 57, we claim that our results go further than being just an application of this nice result, since they cover more general situations, such as when π_2^{SVD} or π_2^{Cay} is used rather than π_2^{exp}. We would like to mention that our algorithm is also different from the approach taken in Ma et al. (2001). The Riemannian metric those authors use is different, as are their geodesics. In Ma et al. (2001) the local structure of the essential manifold being a product of Stiefel manifolds is exploited. Here, we think of this manifold as an orbit of $SO_3 \times SO_3$ acting on $\mathbb{R}^{3 \times 3}$ by equivalence. Some features about these different approaches are summarized as follows.

A.1.1. Manifold Structure

- Ma et al.: The essential manifold $\mathcal{E}$ is *locally* diffeomorphic to the product of two Stiefel manifolds

$$\mathcal{E} \cong_{\text{local}} \mathbb{S}^2 \times SO_3 \qquad (121)$$

- Our approach: We exploit the *global* diffeomorphism of $\mathcal{E}$ to the set of matrices having singular values $\{1, 1, 0\}$

$$\mathcal{E} \cong SO_3 \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \cdot SO_3 \qquad (122)$$

A.1.2. Geodesics Emanating from $E = \Omega R = U E_0 V^\top$

- Ma et al.:

$$t \mapsto (e^{\Delta t} \Omega e^{-\Delta t}, R e^{\Gamma t}) \qquad (123)$$

where $\Delta, \Gamma \in \mathfrak{so}_3$ and Δ is restricted by $[\Delta, [\Delta, \Omega]] = -\frac{1}{2}\|\Delta\|^2 \Omega$.

- Our approach:

$$t \mapsto U e^{\Omega_1 t} E_0 e^{-\Omega_2 t} V^\top \qquad (124)$$

where

$$\Omega_1(x) = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & -\frac{x_3}{\sqrt{2}} & x_2 \\ \frac{x_3}{\sqrt{2}} & 0 & -x_1 \\ -x_2 & x_1 & 0 \end{bmatrix} \qquad (125)$$

and

$$\Omega_2(x) = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & \frac{x_3}{\sqrt{2}} & x_5 \\ -\frac{x_3}{\sqrt{2}} & 0 & -x_4 \\ -x_5 & x_4 & 0 \end{bmatrix}. \qquad (126)$$

and $x_1, \ldots, x_5 \in \mathbb{R}$.

A.1.3. Riemannian Metric $g : T_E\mathcal{E} \times T_E\mathcal{E} \to \mathbb{R}$

- Ma et al.: The Euclidean Riemannian metric induced by the canonical submanifold structure of each factor

$$\mathbb{S}^2 \subset \mathfrak{so}_3 \quad \text{and} \quad SO_3 \subset \mathbb{R}^{3\times3}, \qquad (127)$$

or equivalently, the normal metric induced by the similarity group action on the first factor

$$SO_3 \times \mathbb{S}^2 \to \mathbb{S}^2, \\ (U, \Omega) \mapsto U\Omega U^\top \qquad (128)$$

and right translation on the second factor

$$SO_3 \times SO_3 \to SO_3, \\ (V, R) \mapsto RV^\top. \qquad (129)$$

Explicitly, for two elements of the tangent space $\xi_1, \xi_2 \in T_{(\Omega,R)}\mathcal{E}, \xi_i = ([\Delta_i, \Omega], R\Gamma_i)$, the Riemannian metric is

$$g(([\Delta_1, \Omega], R\Gamma_1), ([\Delta_2, \Omega], R\Gamma_2)) \\ = \text{tr}\left(\Delta_1^\top \Delta_2\right) + \text{tr}\left(\Gamma_1^\top \Gamma_2\right) \qquad (130)$$

with $\Delta_i, \Gamma_i \in \mathfrak{so}_3$ and $[\Delta_i, [\Delta_i, \Omega]] = -\frac{1}{2}\|\Delta_i\|^2\Omega$ for $i = 1, 2$.

- Our approach: The normal Riemannian metric induced by the equivalence group action

$$SO_3 \times SO_3 \times \mathbb{R}^{3\times3} \to \mathbb{R}^{3\times3}, \\ ((U, V), E) \mapsto U E V^\top. \qquad (131)$$

Explicitly, for two elements of the tangent space $\xi_1, \xi_2 \in T_{UE_0V^\top}\mathcal{E}$ with $\xi_i = U(\Delta_i E_0 - E_0\Gamma_i)V^\top$, the Riemannian metric is

$$g(U(\Delta_1 E_0 - E_0\Gamma_1)V^\top, U(\Delta_2 E_0 - E_0\Gamma_2)V^\top) \\ = \text{tr}\left(\Delta_1^\top \Delta_2\right) + \text{tr}\left(\Gamma_1^\top \Gamma_2\right) \qquad (132)$$

where for $i = 1, 2$

$$\Delta_i = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & -\frac{x_3^{(i)}}{\sqrt{2}} & x_2^{(i)} \\ \frac{x_3^{(i)}}{\sqrt{2}} & 0 & -x_1^{(i)} \\ -x_2^{(i)} & x_1^{(i)} & 0 \end{bmatrix},$$

$$(133)$$

$$\Gamma_i = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & \frac{x_3^{(i)}}{\sqrt{2}} & x_5^{(i)} \\ -\frac{x_3^{(i)}}{\sqrt{2}} & 0 & -x_4^{(i)} \\ -x_5^{(i)} & x_4^{(i)} & 0 \end{bmatrix}$$

and $x_1^{(i)}, \ldots, x_5^{(i)} \in \mathbb{R}$. In fact, the tangent map of $\mu_{(U,V)}$ defined by (39) maps the standard basis vectors $\{e_1, \ldots, e_5\}$ in $\mathbb{R}^5$ into basis vectors of $T_E\mathcal{E}$, orthonormal with respect to the normal Riemannian metric:

$$e_i \mapsto D\mu(0) \cdot e_i = U(\Omega_1(e_i)E_0 - E_0\Omega_2(e_i))V^\top \\ = U\xi_i V^\top, \qquad (134)$$

with

$$\langle U\xi_i V^\top, U\xi_j V^\top\rangle_{\text{n.RM}} = \operatorname{tr}\Omega_1^\top(e_i)\Omega_1(e_j)$$
$$+ \operatorname{tr}\Omega_2^\top(e_i)\Omega_2(e_j)$$
$$= e_i^\top e_j$$
$$= \delta_{ij}. \qquad (135)$$

One can easily check by an example that the two Riemannian metrics are indeed not the same.

A.2. Proof of Lemma 3.2

Consider

$$\mathcal{M} := \left\{ (U, \Sigma, V) \in O_3 \times \mathbb{R}^{3\times 3} \times O_3 \,\middle|\, \Sigma = \begin{bmatrix} \alpha & \beta & 0 \\ \beta & \gamma & 0 \\ 0 & 0 & \delta \end{bmatrix}, \right.$$
$$\left. 2|\delta| < \alpha + \gamma - \sqrt{(\alpha - \gamma)^2 + 4\beta^2} \right\}.$$

Note that

$$\sigma_{\min}\left(\begin{bmatrix} \alpha & \beta \\ \beta & \gamma \end{bmatrix} \right) = \tfrac{1}{2}\left(\alpha + \gamma - \sqrt{(\alpha - \gamma)^2 + 4\beta^2}\right). \quad (136)$$

Thus the condition on δ implies that δ is the eigenvalue of Σ with smallest absolute value. Let

$$\mathcal{M}_0 := \left\{ (U, E_0, V) \in O_3 \times \mathbb{R}^{3\times 3} \times O_3 \,\middle|\, E_0 = \begin{bmatrix} I_2 & 0 \\ 0 & 0 \end{bmatrix} \right\}$$

and

$$\Gamma := \left\{ S = \begin{bmatrix} \Delta & 0 \\ 0 & 1 \end{bmatrix} \in \mathbb{R}^{3\times 3} \,\middle|\, \Delta \in O_2 \right\}.$$

Then

$$\sigma : \Gamma \times \mathcal{M} \to \mathcal{M},$$
$$(S, (U, \Sigma, V)) \mapsto (US, S^\top \Sigma S, VS)$$

defines a smooth, proper Lie group action with smooth orbit space $\mathcal{M}/\Gamma$ and the quotient map

$$P : \mathcal{M} \to \mathcal{U},$$
$$(U, \Sigma, V) \mapsto U\Sigma V^\top,$$

is a principal fibre bundle with structure group Γ. Obviously, σ leaves $\mathcal{M}_0$ invariant and therefore restricts to a smooth quotient map

$$P : \mathcal{M}_0 \to \mathcal{E},$$
$$(U, E_0, V) \mapsto U E_0 V^\top.$$

Moreover, the projection map

$$F : \mathcal{M} \to \mathcal{M}_0,$$
$$F(U, \Sigma, V) = (U, E_0, V)$$

is smooth and the diagram

$$
\begin{array}{ccc}
\mathcal{M} & \xrightarrow{\ F\ } & \mathcal{M}_0 \\
P \downarrow & & \downarrow P \\
\mathcal{U} & \xrightarrow[\pi^{\text{SVD}}]{} & \mathcal{E}
\end{array}
\qquad (137)
$$

is commutative. By standard arguments this implies that π^{SVD} is smooth. This completes the proof. $\qquad\square$

References

Hartley, R. and Zisserman, A. 2003. *Multiple View Geometry*, 2nd edition. Cambridge: Cambridge University Press.

Helmke, U., Hüper, K., Lee, P.Y., and Moore, J.B. 2004. Essential matrix estimation via Newton-type methods. In *Proceedings of the MTNS*, Leuven.

Helmke, U. and Moore, J.B. 1994. *Optimization and Dynamical Systems*. London: CCES. Springer.

Horn, B. 1990. Relative orientation. *International Journal of Computer Vision*, 4(1):59–78.

Huang, T.S. and Faugeras, O.D. 1989. Some properties of the E matrix in two-view motion estimation. *IEEE Transactions on Pattern Analysis and Machine Inteligence*, 11(12):1310–1312.

Huang, T.S. and Netravali, A.N. 1994. Motion and structure from feature correspondence: A review. *Proceedings of the IEEE*, 82(2):252–268.

Lang, S. 1999. *Fundamentals of Differential Geometry*. New York: Springer.

Longuet-Higgins, H. 1981. A computer algorithm for reconstructing a scene from projections. *Nature*, 293:133–135.

Ma, Y., Košecká, J., and Sastry, S. 2001. Optimization criteria and geometric algorithms for motion and structure estimation. *International Journal of Computer Vision*, 44(3):219–249.

Mahony, R. 1994. *Optimization Algorithms on Homogeneous Spaces*. PhD thesis, Australian National University, Canberra.

Smith, S.T. 1994. Optimization techniques on Riemannian manifolds. In *Hamiltonian and Gradient Flows, Algorithms and Control*, A. Bloch (ed.), Fields institute communications, pp. 113–136. American Math. Soc., Providence.

Soatto, S., Frezza, R., and Perona, P. 1996. Motion estimation via dynamic vision. *IEEE Transactions on Automatic Control*, 41(3):393–413.