Motion Recovery from Image Sequences: Discrete Viewpoint vs. Differential Viewpoint*

Yi Ma, Jana Košecká, and Shankar Sastry

Electronics Research Laborotory
University of California at Berkeley,
Berkeley CA 94720-1774, USA,
{mayi,janka,sastry}@robotics.eecs.berkeley.edu

Abstract. The aim of this paper is to explore intrinsic geometric methods of recovering the three dimensional motion of a moving camera from a sequence of images. Generic similarities between the discrete approach and the differential approach are revealed through a parallel development of their analogous motion estimation theories.

We begin with a brief review of the (discrete) essential matrix approach, showing how to recover the 3D displacement from image correspondences. The space of normalized essential matrices is characterized geometrically: the unit tangent bundle of the rotation group is a double covering of the space of normalized essential matrices. This characterization naturally explains the geometry of the possible number of 3D displacements which can be obtained from the essential matrix.

Second, a differential version of the essential matrix constraint previously explored by [19, 20] is presented. We then present the precise characterization of the space of differential essential matrices, which gives rise to a novel eigenvector-decomposition-based 3D velocity estimation algorithm from the optical flow measurements. This algorithm gives a unique solution to the motion estimation problem and serves as a differential counterpart of the SVD-based 3D displacement estimation algorithm from the discrete case.

Finally, simulation results are presented evaluating the performance of our algorithm in terms of bias and sensitivity of the estimates with respect to the noise in optical flow measurements.

Keywords: optical flow, epipolar constraint, motion estimation.

* Research is supported by ARO under the MURI grant DAAH04-96-1-0341, "An Integrated Approach to Intelligent Systems".

1 Introduction

The problem of estimating structure and motion from image sequences has been studied extensively by the computer vision community in the past decade. Various approaches differ in the types of assumptions they make about the projection model, the model of the environment, or the type of algorithms they use for estimating the motion and/or structure. Most of the techniques try to decouple the two problems by estimating the motion first, followed by the structure estimation. In spite of the fact that the robustness of existing motion estimation algorithms has been studied quite extensively, it has been suggested that the fact that the structure and motion estimation are decoupled typically hinders their performance [12]. Some algorithms address the problem of motion and structure (shape) recovery simultaneously either in batch [16] or recursive fashion [12].

The approaches to the motion estimation only, can be partitioned into the discrete and differential methods depending on whether they use as an input set of point correspondences or image velocities. Among the efforts to solve this problem, one of the more appealing approaches is the *essential matrix approach*, proposed by Longuet-Higgins, Huang and Faugeras *et al* in 1980s [7]. It shows that the relative 3D displacement of a camera can be recovered from an *intrinsic* geometric constraint between two images of the same scene, the so-called *Longuet-Higgins constraint* (also called the *epipolar or essential constraint*). Estimating 3D motion can therefore be decoupled from estimation of the structure of the 3D scene. This endows the resulting motion estimation algorithms with some advantageous features: they do not need to assume any *a priori* knowledge of the scene; and are computationally simpler (comparing to most non-intrinsic motion estimation algorithms), using mostly linear algebraic techniques. Tsai and Huang [18] then proved that, given an essential matrix associated with the Longuet-Higgins constraint, there are only two possible 3D displacements. The study of the essential matrix then led to a three-step SVD-based algorithm for recovering the 3D displacement from noisy image correspondences, proposed in 1986 by Toscani and Faugeras [17] and later summarized in Maybank [11].

Being motivated by recent interests in dynamical motion estimation schemes (Soatto, Frezza and Perona [14]) which usually require smoothness and regularity of the parameter space, the geometric property of the essential matrix space is further explored: the unit tangent bundle of the rotation group, *i.e.* $T_1(SO(3))$, is a double covering of the space of normalized essential matrices (full proofs are given in [9]).

However, the essential matrix approach based on the Longuet-Higgins constraint only recovers *discrete* 3D displacement. The velocity information can only be approximately obtained from the inverse of the exponential map, as Soatto *et al* did in [14]. In principle, the displacement estimation algorithms obtained by using epipolar constraints work well when the displacement (especially the translation) between the two images is relatively large. However, in real-time applications, even if the velocity of the moving camera is not small, the relative displacement between two consecutive images might become small due to a high sampling rate. In turn, the algorithms become singular due to the small translation and the estimation results become less reliable.

A differential (or continuous) version of the 3D motion estimation problem is to recover the 3D velocity of the camera from optical flow. This problem has also been explored by many researchers: an algorithm was proposed in 1984 by Zhuang *et al* [20] with a simplified version given in 1986 [21]; and a first order algorithm was given by Waxman *et al* [8] in 1987. Most of the algorithms start from the basic bilinear constraint relating optical flow to the linear and angular velocities and solve for rotation and translation separately using either numerical optimization techniques (Bruss and Horn [2]) or linear subspace methods (Heeger and Jepson [3,4]). Kanatani [5] proposed a linear algorithm reformulating Zhuang's approach in terms of essential parameters and twisted flow. However, in these algorithms, the similarities between the discrete case and the differential case are not fully revealed and exploited.

In this paper, we develop in parallel to the discrete essential matrix approach developed in the literature, as a review see Ma *et al* [9] or Maybank [11], a *differential essential matrix approach* for recovering 3D velocity from optical flow. Based on the differential version of the Longuet-Higgins constraint, so called *differential essential matrices* are defined. We then give a complete characterization of the space of these matrices and prove that there exists exactly one 3D velocity corresponding to a given differential essential matrix. As a differential counterpart of the three-step SVD-based 3D displacement estimation algorithm, a four-step eigenvector-decomposition-based 3D velocity estimation algorithm is proposed.

2 Discrete Essential Matrix Approach Review

We first introduce some notation which will be frequently used in this paper. Given a vector $p = (p_1, p_2, p_3)^T \in \mathbb{R}^3$, we define $\hat{p} \in so(3)$ (the space of skew symmetric matrices in $\mathbb{R}^{3\times 3}$) by:

$$\hat{p} = \begin{pmatrix} 0 & -p_3 & p_2 \\ p_3 & 0 & -p_1 \\ -p_2 & p_1 & 0 \end{pmatrix}. \tag{1}$$

It then follows from the definition of cross product of vectors that, for any two vectors $p, q \in \mathbb{R}^3$: $p \times q = \hat{p}q$. The matrices of rotation by θ radians about y-axis and z-axis are respectively denoted by:

$$R_Y(\theta) = \begin{pmatrix} \cos(\theta) & 0 & \sin(\theta) \\ 0 & 1 & 0 \\ -\sin(\theta) & 0 & \cos(\theta) \end{pmatrix}, \quad R_Z(\theta) = \begin{pmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{2}$$

The camera motion can be modeled as a rigid body motion in $\mathbb{R}^3$. The displacement of the camera belongs to the special Euclidean group $SE(3)$:

$$SE(3) = \{(p, R) : p \in \mathbb{R}^3, R \in SO(3)\} \tag{3}$$

where $SO(3)$ is the space of 3×3 rotation matrices (unitary matrices with determinant $+1$) on $\mathbb{R}$. An element $g = (p, R)$ in this group is used to represent the 3D translation and orientation (the displacement) of a coordinate frame F_c attached to the camera relative to an inertial frame which is chosen here as the initial position of the camera frame F_o. By an abuse of notation, the element $g = (p, R)$ serves both as a specification of the configuration of the camera and as a transformation taking the coordinates of a point from F_c to F_o. More precisely, let $q_o, q_c \in \mathbb{R}^3$ be the coordinates of a point q relative to frames F_o and F_c, respectively. Then the coordinate transformation between q_o and q_c is given by:

$$q_o = Rq_c + p. \tag{4}$$

In this paper, we use bold letters to denote quantities associated with the image. The image of a point $q \in \mathbb{R}^3$ in the scene is then denoted by $\mathbf{q} \in \mathbb{R}^3$. As the model of image formation, we consider both *spherical projection* and *perspective projection*. For the spherical projection, we simply choose the imaging surface to be the unit sphere: $S^2 = \left\{q \in \mathbb{R}^3 : \|q\| = 1\right\}$, where the norm $\|\cdot\|$ always means 2-norm unless otherwise stated. Then the spherical projection is defined by the map π_s from $\mathbb{R}^3$ to S^2:

$$\pi_s : q \mapsto \mathbf{q} = \frac{q}{\|q\|}.$$

For the perspective projection, the imaging surface is chosen to be the plane of unit distance away from the optical center. The perspective projection onto this plane is then defined by the map π_p from $\mathbb{R}^3$ to the projective plane $\mathbb{RP}^2 \subset \mathbb{R}^3$:

$$\pi_p : q = (q_1, q_2, q_3)^T \mapsto \mathbf{q} = (\frac{q_1}{q_3}, \frac{q_2}{q_3}, 1)^T.$$

The approach taken in this paper only exploits the intrinsic geometric relations which are preserved by both projection models. Thus, theorems and algorithms to be developed are true for both cases. We simply denote both π_s and π_p by the same letter π. The image of the point q taken by the camera at the initial position then is $\mathbf{q}_o = \pi(q_o)$, and the image of the same point taken at the current position is $\mathbf{q}_c = \pi(q_c)$. The two corresponding image points $\mathbf{q}_o$ and $\mathbf{q}_c$ have to satisfy an intrinsic geometric constraint, the so-called *Longuet-Higgins or epipolar constraint* [7]:

$$\mathbf{q}_c^T R^T \hat{p} \mathbf{q}_o \equiv 0. \tag{5}$$

The matrices which have the form $E = R^T \hat{p}$ with $R \in SO(3)$ and $\hat{p} \in so(3)$ play an important role in recovering the displacement (p, R). Such matrices are called *essential matrices*; and the set of all essential matrices is called the *essential space*, defined to be

$$\mathcal{E} = \{RS \mid R \in SO(3), S \in so(3)\}. \tag{6}$$

The following theorem is a stronger version of Huang and Faugeras' theorem and gives a characterization of the essential space:

Theorem 1. (Characterization of the Essential Matrix)
A non-zero matrix E is an essential matrix if and only if the singular value decomposition (SVD) of E: $E = U\Sigma V^T$ satisfies: $\Sigma = diag\{\lambda, \lambda, 0\}$ for some $\lambda > 0$ and $U, V \in SO(3)$.

The condition $U, V \in SO(3)$ was not in the original theorem given by Huang or Faugeras, but it is convenient for the following theorem which shows how to explicitly recover the displacement from an essential matrix. One may refer to the full paper [9] for the proof of this extra condition.

Theorem 2. (Uniqueness of the Displacement Recovery from the Essential Matrix)
There exist exactly two 3D displacements $g = (p, R) \in SE(3)$ corresponding to a non-zero essential matrix $E \in \mathcal{E}$. Further, given the SVD of the matrix $E = U\Sigma V^T$, the two displacements (p, R) that solve $E = R^T\hat{p}$ are given by:

$$
\begin{aligned}
(R_1^T, \hat{p}_1) &= (UR_Z^T(+\frac{\pi}{2})V^T, VR_Z(+\frac{\pi}{2})\Sigma V^T) \\
(R_2^T, \hat{p}_2) &= (UR_Z^T(-\frac{\pi}{2})V^T, VR_Z(-\frac{\pi}{2})\Sigma V^T).
\end{aligned} \tag{7}
$$

This theorem is a summary of results presented in [18, 14]. A rigorous proof of this theorem is given in [9]. A natural consequence of Theorem 1 and 2 is the three-step SVD-based displacement estimation algorithm proposed by Toscani and Faugeras [17], which is summarized in [11] or [9].

Motivated by recent interests in dynamic (or recursive) motion estimation schemes [14], differential geometric properties of the essential space $\mathcal{E}$ have been explored. Since the Longuet-Higgins condition is an homogeneous constraint, the essential matrix E can only be recovered up to a scale factor. It is then customary to set the norm of the translation vector p to be 1. Thus the normalized essential space, defined to be

$$\mathcal{E}_1 = \{RS \mid R \in SO(3), S = \hat{p}, ||p|| = 1\}, \tag{8}$$

is of particular interest in motion estimation algorithms.

Theorem 3. (Characterization of the Normalized Essential Space)
The unit tangent bundle of the rotation group $SO(3)$, i.e. $T_1(SO(3))$, is a double covering of the normalized essential space $\mathcal{E}_1$, or equivalently speaking, $\mathcal{E}_1 = T_1(SO(3))/\mathbb{Z}_2$.

The proof of this theorem, as well as a more detailed differential geometric characterization of the normalized essential space is given in [9]. As a consequence of this theorem, the normalized essential space $\mathcal{E}_1$ is a 5-dimensional connected compact manifold embedded in $\mathbb{R}^{3\times3}$. This property validates estimation algorithms which require certain smoothness and regularity on the parameter space, as dynamic algorithms usually do.

3 Differential Essential Matrix Approach

The differential case is the infinitesimal version of the discrete case. To reveal the similarities between these two cases, we now develop the *differential essential matrix approach* for estimating 3D velocity from optical flow in a parallel way as developed in the literature for the discrete essential matrix approach for estimating 3D displacement from image correspondences [9, 11]. After deriving a differential version of the Longuet-Higgins constraint, the concept of differential essential matrix is defined; we then give a thorough characterization for such matrices and show that there exists exactly one 3D velocity corresponding to a non-zero differential essential matrix; as a differential version of the three-step SVD-based 3D displacement estimation algorithm [11], a four-step eigenvector-decomposition-based 3D velocity estimation algorithm is proposed.

3.1 Differential Longuet-Higgins Constraint

Suppose the motion of the camera is described by a smooth curve $g(t) = (p(t), R(t)) \in SE(3)$. According to (4), for a point q attached to the inertial frame F_o, its coordinates in the inertial frame and the moving camera frame satisfy: $q_o = R(t)q_c(t) + p(t)$. Differentiating this equation yields: $\dot{q}_c = -R^T\dot{R}q_c - R^T\dot{p}$.

Since $-R^T\dot{R} \in so(3)$ and $-R^T\dot{p} \in \mathbb{R}^3$ (see Murray, Li and Sastry [13]), we may define $\omega = (\omega_1, \omega_2, \omega_3)^T \in \mathbb{R}^3$ and $v = (v_1, v_2, v_3)^T \in \mathbb{R}^3$ to be:

$$\hat{\omega} = -R^T\dot{R}, \quad v = -R^T\dot{p}. \tag{9}$$

The interpretation of these velocities is: $-\omega$ is the angular velocity of the camera frame F_c relative to the inertial frame F_i and $-v$ is the velocity of the origin of the camera frame F_c relative to the inertial frame F_i. Using the new notation, we get:

$$\dot{q}_c = \hat{\omega}q_c + v. \tag{10}$$

From now on, for convenience we will drop the subscript c from q_c. The notation q then serves both as a point fixed in the spatial frame and its coordinates with respect to the current camera frame F_c. The image of the point q taken by the camera is given by projection: $\mathbf{q} = \pi(q)$, and it's optical flow $\mathbf{u}$, $\mathbf{u} = \dot{\mathbf{q}} \in \mathbb{R}^3$. The following is the differential version of the Longuet-Higgins constraint, which has been independently referenced and used by many people in computer vision.

Theorem 4. (Differential Longuet-Higgins Constraint)
Consider a camera moving with linear velocity v and angular velocity ω with respect to the inertial frame. Then the optical flow $\mathbf{u}$ at an image point $\mathbf{q}$ satisfies:

$$(\mathbf{u}^T, \mathbf{q}^T)\begin{pmatrix}\hat{v} \\ s\end{pmatrix}\mathbf{q} = 0 \tag{11}$$

where s is a symmetric matrix defined by $s := \frac{1}{2}(\hat{\omega}\hat{v} + \hat{v}\hat{\omega}) \in \mathbb{R}^{3\times 3}$.

Proof. From the definition of the map π's, there exists a real scalar function $\lambda(t)$ ($\|q(t)\|$ or $q_3(t)$, depending on whether the projection is spherical or perspective) such that: $q = \lambda \mathbf{q}$. Take the inner product of the vectors in (10) with $(v \times \mathbf{q})$:

$$\dot{q}^T(v \times \mathbf{q}) = (\hat{\omega} q + v)^T (v \times \mathbf{q}) = q^T \hat{\omega}^T \hat{v} \mathbf{q}. \tag{12}$$

Since $\dot{q} = \dot{\lambda}\mathbf{q} + \lambda\dot{\mathbf{q}}$ and $\mathbf{q}^T(v \times \mathbf{q}) = 0$, from (12) we then have: $\lambda\dot{\mathbf{q}}^T \hat{v}\mathbf{q} - \lambda\mathbf{q}^T\hat{\omega}^T\hat{v}\mathbf{q} = 0$. When $\lambda \neq 0$, we have: $\mathbf{u}^T\hat{v}\mathbf{q} + \mathbf{q}^T\hat{\omega}\hat{v}\mathbf{q} \equiv 0$. For any skew symmetric matrix $A \in \mathbb{R}^{3\times 3}$, $\mathbf{q}^T A \mathbf{q} = 0$. Since $\frac{1}{2}(\hat{\omega}\hat{v} - \hat{v}\hat{\omega})$ is a skew symmetric matrix, $\mathbf{q}^T \frac{1}{2}(\hat{\omega}\hat{v} - \hat{v}\hat{\omega})\mathbf{q} = \mathbf{q}^T s\mathbf{q} - \mathbf{q}^T\hat{\omega}\hat{v}\mathbf{q} = 0$. Thus, $\mathbf{q}^T s\mathbf{q} = \mathbf{q}^T\hat{\omega}\hat{v}\mathbf{q}$. We then have: $\mathbf{u}^T\hat{v}\mathbf{q} + \mathbf{q}^T s\mathbf{q} \equiv 0$.

3.2 Characterization of the Differential Essential Matrix

We define the space of 6×3 matrices given by:

$$\mathcal{E}' = \left\{ \begin{pmatrix} \hat{v} \\ \frac{1}{2}(\hat{\omega}\hat{v} + \hat{v}\hat{\omega}) \end{pmatrix} \middle| \omega, v \in \mathbb{R}^3 \right\} \subset \mathbb{R}^{6\times 3}. \tag{13}$$

to be the *differential essential space.* A matrix in this space is called a *differential essential matrix.* Note that the differential Longuet-Higgins constraint (11) is homogeneous in the linear velocity v. Thus v may be recovered only up to a constant scale. Consequently, in motion recovery, we will concern ourselves with matrices belonging to *normalized differential essential space*:

$$\mathcal{E}'_1 = \left\{ \begin{pmatrix} \hat{v} \\ \frac{1}{2}(\hat{\omega}\hat{v} + \hat{v}\hat{\omega}) \end{pmatrix} \middle| \omega \in \mathbb{R}^3, v \in S^2 \right\} \subset \mathbb{R}^{6\times 3}. \tag{14}$$

The skew-symmetric part $\hat{v}$ of a differential essential matrix simply corresponds to the velocity v. The characterization of the (normalized) essential matrix only focuses on the characterization of the symmetric part of the matrix: $s = \frac{1}{2}(\hat{\omega}\hat{v} + \hat{v}\hat{\omega})$. We call the space of all the matrices of such form the *special symmetric space*:

$$\mathcal{S} = \left\{ \frac{1}{2}(\hat{\omega}\hat{v} + \hat{v}\hat{\omega}) \middle| \omega \in \mathbb{R}^3, v \in S^2 \right\} \subset \mathbb{R}^{3\times 3}. \tag{15}$$

A matrix in this space is called a *special symmetric matrix.* The motion estimation problem is now reduced to the one of *recovering the velocity* (ω, v) *with* $\omega \in \mathbb{R}^3$ *and* $v \in S^2$ *from a given special symmetric matrix* s.

The characterization of special symmetric matrices depends on a characterization of matrices in the form: $\hat{\omega}\hat{v} \in \mathbb{R}^{3\times 3}$, which is given in the following lemma. This lemma will also be used in the next section to prove the uniqueness of the velocity recovery from special symmetric matrices. Like the (discrete) essential matrices, matrices with the form $\hat{\omega}\hat{v}$ are characterized by their singular value decomposition (SVD): $\hat{\omega}\hat{v} = U\Sigma V^T$, and moreover, the unitary matrices U and V are related.

Lemma 1. *A matrix $Q \in \mathbb{R}^{3\times3}$ has the form $Q = \hat{\omega}\hat{v}$ with $\omega \in \mathbb{R}^3$, $v \in S^2$ if and only if the SVD of Q has the form:*

$$Q = -VR_Y(\theta)diag\{\lambda, \lambda\cos(\theta), 0\}V^T \tag{16}$$

for some rotation matrix $V \in SO(3)$. Further, $\lambda = \|\omega\|$ and $\cos(\theta) = \omega^T v/\lambda$.

Proof. We first prove the necessity. The proof follows from the geometric meaning of $\hat{\omega}\hat{v}$: for any vector $q \in \mathbb{R}^3$, $\hat{\omega}\hat{v}q = \omega \times (v \times q)$. Let $b \in S^2$ be the unit vector perpendicular to both ω and v: $b = \frac{v\times\omega}{\|v\times\omega\|}$ (if $v \times \omega = 0$, b is not uniquely defined. In this case, pick any b orthogonal to v and ω, then the rest of the proof still holds). Then $\omega = \lambda e^{\hat{b}\theta}v$ for some $\lambda \in \mathbb{R}_+$ and $\theta \in \mathbb{R}$ (according this definition, λ is the length of ω; θ is the angle between ω and v, and $0 \leq \theta \leq \pi$). It is direct to check that if the matrix V is defined to be: $V = (e^{\hat{b}\frac{\pi}{2}}v, b, v)$. Q has the form given by (16).

We now prove the sufficiency. Given a matrix Q which can be decomposed in the form (16), define the unitary matrix $U = -VR_Y(\theta) \in O(3)$. For matrix $\Sigma_\sigma = diag\{\sigma, \sigma, 0\}$ with $\sigma \in \mathbb{R}$, it is direct to check that matrices $R_Z(+\frac{\pi}{2})\Sigma_\sigma$ and $R_Z(-\frac{\pi}{2})\Sigma_\sigma$ are skew matrices. So $WR_Z(\pm\frac{\pi}{2})\Sigma_\sigma W^T$ are also skew for any $W \in O(3)$. Let $\hat{\omega}$ and $\hat{v}$ given by the formulae:

$$\hat{\omega} = UR_Z(\pm\frac{\pi}{2})\Sigma_\lambda U^T, \quad \hat{v} = VR_Z(\pm\frac{\pi}{2})\Sigma_1 V^T \tag{17}$$

where $\Sigma_\lambda = diag\{\lambda, \lambda, 0\}$ and $\Sigma_1 = diag\{1, 1, 0\}$. Then:

$$\begin{aligned}\hat{\omega}\hat{v} &= UR_Z(\pm\frac{\pi}{2})\Sigma_\lambda U^T VR_Z(\pm\frac{\pi}{2})\Sigma_1 V^T = UR_Z(\pm\frac{\pi}{2})\Sigma_\lambda(-R_Y^T(\theta))R_Z(\pm\frac{\pi}{2})\Sigma_1 V^T \\ &= Udiag\{\lambda, \lambda\cos(\theta), 0\}V^T = Q. \end{aligned} \tag{18}$$

Since ω and v have to be, respectively, the left and the right zero eigenvectors of Q, the reconstruction given in (17) is unique.

The following theorem gives a characterization of the special symmetric matrix.

Theorem 5. (Characterization of the Special Symmetric Matrix) *A matrix $s \in \mathbb{R}^{3\times3}$ is a special symmetric matrix if and only if s can be diagonalized as $s = V\Sigma V^T$ with $V \in SO(3)$ and: $\Sigma = diag\{\sigma_1, \sigma_2, \sigma_3\}$, with $\sigma_1 \geq 0, \sigma_3 \leq 0$ and $\sigma_2 = \sigma_1 + \sigma_3$.*

Proof. We first prove the necessity. Suppose s is a special symmetric matrix, there exist $\omega \in \mathbb{R}^3, v \in S^2$ such that $s = \frac{1}{2}(\hat{\omega}\hat{v} + \hat{v}\hat{\omega})$. Since s is a symmetric matrix, it is diagonalizable, all its eigenvalues are real and all the eigenvectors are orthogonal to each other. It then suffices to check its eigenvalues satisfy the given conditions.

Let the unit vector b and the rotation matrix V be the same as in the proof of Lemma 1, so are θ and γ. Then according to Lemma 1:

$$\hat{\omega}\hat{v} = -VR_Y(\theta)diag\{\lambda, \lambda\cos(\theta), 0\}V^T.$$

Since $(\hat{\omega}\hat{v})^T = \hat{v}\hat{\omega}$, it yields

$$s = \frac{1}{2}V\left(-R_Y(\theta)diag\{\lambda, \lambda\cos(\theta), 0\} - diag\{\lambda, \lambda\cos(\theta), 0\}R_Y^T(\theta)\right)V^T.$$

Define the matrix $D(\lambda,\theta) \in \mathbb{R}^{3\times 3}$ to be

$$\begin{aligned} D(\lambda,\theta) &= -R_Y(\theta)diag\{\lambda, \lambda\cos(\theta), 0\} - diag\{\lambda, \lambda\cos(\theta), 0\}R_Y^T(\theta) \\ &= \lambda\begin{pmatrix} -2\cos(\theta) & 0 & \sin(\theta) \\ 0 & -2\cos(\theta) & 0 \\ \sin(\theta) & 0 & 0 \end{pmatrix}. \end{aligned} \tag{19}$$

Directly calculating its eigenvalues and eigenvectors, we obtain that

$$\begin{aligned} D(\lambda,\theta) = R_Y&\left(\frac{\theta}{2}-\frac{\pi}{2}\right) \\ &\times diag\left\{\lambda(1-\cos(\theta)), -2\lambda\cos(\theta), \lambda(-1-\cos(\theta))\right\}R_Y^T\left(\frac{\theta}{2}-\frac{\pi}{2}\right) \end{aligned} \tag{20}$$

Thus $s = \frac{1}{2}VD(\lambda,\theta)V^T$ has eigenvalues:

$$\left\{\frac{1}{2}\lambda(1-\cos(\theta)), \quad -\lambda\cos(\theta), \quad \frac{1}{2}\lambda(-1-\cos(\theta))\right\}, \tag{21}$$

which satisfy the given conditions.

We now prove the sufficiency. Given $s = V_1 diag\{\sigma_1, \sigma_2, \sigma_3\}V_1^T$ with $\sigma_1 \geq 0, \sigma_3 \leq 0$ and $\sigma_2 = \sigma_1 + \sigma_3$ and $V_1^T \in SO(3)$, these three eigenvalues uniquely determine $\lambda, \theta \in \mathbb{R}$ such that the σ_i's have the form given in (21):

$$\begin{cases} \lambda = \quad \sigma_1 - \sigma_3, & \lambda \geq 0 \\ \theta = \arccos(-\sigma_2/\lambda), & \theta \in [0,\pi] \end{cases}$$

Define a matrix $V \in SO(3)$ to be $V = V_1 R_Y^T\left(\frac{\theta}{2}-\frac{\pi}{2}\right)$. Then $s = \frac{1}{2}VD(\lambda,\theta)V^T$. According to Lemma 1, there exist vectors $v \in S^2$ and $\omega \in \mathbb{R}^3$ such that $\hat{\omega}\hat{v} = -VR_Y(\theta)diag\{\lambda, \lambda\cos(\theta), 0\}V^T$. Therefore, $\frac{1}{2}(\hat{\omega}\hat{v} + \hat{v}\hat{\omega}) = \frac{1}{2}VD(\lambda,\theta)V^T = s$.

3.3 Uniqueness of 3D Velocity Recovery from the Special Symmetric Matrix

Theorem 5 is given in Kanatani [6] as exercise 7.12. However, we are going to use this property and its constructive proof to propose a new motion recovery algorithm. This algorithm is based upon the following theorem whose proof explicitly gives all the possible ω's and v's which can be recovered from a special symmetric matrix.

Theorem 6. (Uniqueness of the Velocity Recovery from the Special Symmetric Matrix)
There exist exactly four 3D velocities (ω, v) with $\omega \in \mathbb{R}^3$ and $v \in S^2$ corresponding to a non-zero special symmetric matrix $s \in \mathcal{S}$.

Proof. Suppose (ω_1, v_1) and (ω_2, v_2) are both solutions for $s = \frac{1}{2}(\hat{\omega}\hat{v} + \hat{v}\hat{\omega})$, we have: $\hat{v}_1\hat{\omega}_1 + \hat{\omega}_1\hat{v}_1 = \hat{v}_2\hat{\omega}_2 + \hat{\omega}_2\hat{v}_2$. From Lemma 1, we may write:

$$\begin{aligned} \hat{\omega}_1\hat{v}_1 &= -V_1 R_Y(\theta_1) diag\{\lambda_1, \lambda_1\cos(\theta_1), 0\} V_1^T \\ \hat{\omega}_2\hat{v}_2 &= -V_2 R_Y(\theta_2) diag\{\lambda_2, \lambda_2\cos(\theta_2), 0\} V_2^T. \end{aligned} \tag{22}$$

Let $W = V_1^T V_2 \in SO(3)$, then: $D(\lambda_1, \theta_1) = W D(\lambda_2, \theta_2) W^T$. Since both sides have the same eigenvalues, according to (20), we have: $\lambda_1 = \lambda_2, \theta_2 = \theta_1$. We then can denote both θ_1 and θ_2 by θ. It is direct to check that the only possible rotation matrix W which satisfies the preceding equation is given by $I_{3\times 3}$ or:

$$\begin{pmatrix} -\cos(\theta) & 0 & \sin(\theta) \\ 0 & -1 & 0 \\ \sin(\theta) & 0 & \cos(\theta) \end{pmatrix} \quad or \quad \begin{pmatrix} \cos(\theta) & 0 & -\sin(\theta) \\ 0 & -1 & 0 \\ -\sin(\theta) & 0 & -\cos(\theta) \end{pmatrix}. \tag{23}$$

From the geometric meaning of V_1 and V_2, all the cases give either $\hat{\omega}_1\hat{v}_1 = \hat{\omega}_2\hat{v}_2$ or $\hat{\omega}_1\hat{v}_1 = \hat{v}_2\hat{\omega}_2$. Thus, according to the proof of Lemma 1, if (ω, v) is one solution and $\hat{\omega}\hat{v} = U diag\{\lambda, \lambda\cos(\theta), 0\} V^T$, then all the solutions are given by:

$$\begin{aligned} \hat{\omega} &= U R_Z(\pm\frac{\pi}{2}) \Sigma_\lambda U^T, \quad \hat{v} = V R_Z(\pm\frac{\pi}{2}) \Sigma_1 V^T; \\ \hat{\omega} &= V R_Z(\pm\frac{\pi}{2}) \Sigma_\lambda V^T, \quad \hat{v} = U R_Z(\pm\frac{\pi}{2}) \Sigma_1 U^T \end{aligned} \tag{24}$$

where $\Sigma_\lambda = diag\{\lambda, \lambda, 0\}$ and $\Sigma_1 = diag\{1, 1, 0\}$.

Given a non-zero differential essential matrix $E \in \mathcal{E}'$, its special symmetric part gives four possible solutions for the 3D velocity (ω, v). However, only one of them has the same linear velocity v as the skew-symmetric part of E does. We thus have:

Theorem 7. (Uniqueness of Velocity Recovery from the Differential Essential Matrix)
There exists only one 3D velocity (ω, v) with $\omega \in \mathbb{R}^3$ and $v \in \mathbb{R}^3$ corresponding to a non-zero differential essential matrix $E \in \mathcal{E}'$.

In the discrete case, there are two 3D displacements corresponding to an essential matrix. However, the velocity corresponding to a differential essential matrix is unique. This is because, in the differential case, the twist-pair ambiguity (see Maybank [11]), which is caused by a 180° rotation of the camera around the translation direction, is avoided.

It is clear that the normalized differential essential space $\mathcal{E}'_1$ is a 5-dimensional differentiable submanifold embedded in $\mathbb{R}^{6\times 3}$. Further considering the symmetric and anti-symmetric structures in the differential essential matrix, the embedding space can be naturally reduced from $\mathbb{R}^{6\times 3}$ to $\mathbb{R}^9$. This property is useful when using estimation schemes which require some regularity on the parameter space, for example, the dynamic estimation scheme proposed by Soatto *et al* [14].

3.4 Algorithm

Based on the previous study of the differential essential matrix, in this section, we propose an algorithm which recovers the 3D velocity of the camera from a set of (possibly noisy) optical flow vectors.

Let $E = \begin{pmatrix} \hat{v} \\ s \end{pmatrix} \in \mathcal{E}_1'$ with $s = \frac{1}{2}(\hat{\omega}\hat{v} + \hat{v}\hat{\omega})$ be the essential matrix associated with the differential Longuet-Higgins constraint (11). Since the submatrix $\hat{v}$ is skew symmetric and s is symmetric, they have the following forms:

$$v = \begin{pmatrix} 0 & -v_3 & v_2 \\ v_3 & 0 & -v_1 \\ -v_2 & v_1 & 0 \end{pmatrix}, \quad s = \begin{pmatrix} s_1 & s_2 & s_3 \\ s_2 & s_4 & s_5 \\ s_3 & s_5 & s_6 \end{pmatrix}. \tag{25}$$

Define the (differential) *essential vector* $\mathbf{e} \in \mathbb{R}^9$ to be:

$$\mathbf{e} = (v_1, v_2, v_3, s_1, s_2, s_3, s_4, s_5, s_6)^T. \tag{26}$$

Define a vector $\mathbf{a} \in \mathbb{R}^9$ associated to optical flow $(\mathbf{q}, \mathbf{u})$ with $\mathbf{q} = (x, y, z)^T \in \mathbb{R}^3, \mathbf{u} = (u_1, u_2, u_3)^T \in \mathbb{R}^3$ to be[1]:

$$\mathbf{a} = (u_3 y - u_2 z, u_1 z - u_3 x, u_2 x - u_1 y, x^2, 2xy, 2xz, y^2, 2yz, z^2)^T. \tag{27}$$

The differential Longuet-Higgins constraint (11) can be then rewritten as: $\mathbf{a}^T\mathbf{e} = 0$. Given a set of (possibly noisy) optical flow vectors: $(\mathbf{q}^i, \mathbf{u}^i)$, $i = 1, \dots, m$ generated by the same motion, define a matrix $A \in \mathbb{R}^{m\times 9}$ associated with these measurements to be: $A = (\mathbf{a}^1, \mathbf{a}^2, \dots, \mathbf{a}^m)^T$, where $\mathbf{a}^i$ are defined for each pair $(\mathbf{q}^i, \mathbf{u}^i)$ using (27). In the absence of noise, the essential vector $\mathbf{e}$ has to satisfy: $A\mathbf{e} = 0$. In order for this equation to have a unique solution for $\mathbf{e}$, the rank of the matrix A has to be eight. Thus, *for this algorithm, in general, the optical flow vectors of at least eight points are needed to recover the 3D velocity, i.e.* $m \geq 8$, although the minimum number of optical flows needed is 5 (see Maybank [11]).

When the measurements are noisy, there might be no solution of $\mathbf{e}$ for $A\mathbf{e} = 0$. As in the discrete case, we choose the solution which minimizes the error function $||A\mathbf{e}||^2$. This can be mechanized using the following lemma. It is straight forward to see that (Theorem 6.1 of Maybank [11]):

Lemma 2. *If a matrix* $A \in \mathbb{R}^{n\times n}$ *has the singular value decomposition* $A = U\Sigma V^T$ *and* $c_n(V)$ *is the* n^{th} *column vector of* V *(the singular vector associated to the smallest singular value* σ_n*), then* $\mathbf{e} = c_n(V)$ *minimizes* $||A\mathbf{e}||^2$ *subject to the condition* $||\mathbf{e}|| = 1$.

Since the differential essential vector $\mathbf{e}$ is recovered from noisy measurements, the symmetric part s of E directly recovered from $\mathbf{e}$ is not necessarily a special symmetric matrix. Thus one can not directly use the previously derived results for special symmetric matrices to recover the 3D velocity. In the algorithms

[1] For perspective projection, $z = 1$ and $u_3 = 0$ thus the expression for $\mathbf{a}$ can be simplified.

proposed in Zhuang [20, 21], such s, with the linear velocity v obtained from the skew-symmetric part, is directly used to calculate the angular velocity ω. This is a over-determined problem since three variables are to be determined from six independent equations; on the other hand, erroneous v introduces further error in the estimation of the angular velocity ω.

We thus propose a different approach: first extract the special symmetric component from the first-hand symmetric matrix s; then recover the four possible solutions for the 3D velocity using the results obtained in Theorem 6; finally choose the one which has the closest linear velocity to the one given by the skew-symmetric part of E. In order to extract the special symmetric component out of a symmetric matrix, we need a projection from the space of all symmetric matrices to the special symmetric space $\mathcal{S}$.

Theorem 8. (Projection to the Special Symmetric Space)
If a symmetric matrix $F \in \mathbb{R}^{3\times3}$ is diagonalized as $F = V diag\{\lambda_1, \lambda_2, \lambda_3\} V^T$ with $V \in SO(3)$, $\lambda_1 \geq 0, \lambda_3 \leq 0$ and $\lambda_1 \geq \lambda_2 \geq \lambda_3$, then the special symmetric matrix $E \in \mathcal{S}$ which minimizes the error $\|E - F\|_f^2$ is given by $E = V diag\{\sigma_1, \sigma_2, \sigma_2\} V^T$ with:

$$\sigma_1 = \frac{2\lambda_1 + \lambda_2 - \lambda_3}{3}, \quad \sigma_2 = \frac{\lambda_1 + 2\lambda_2 + \lambda_3}{3}, \quad \sigma_3 = \frac{2\lambda_3 + \lambda_2 - \lambda_1}{3}. \tag{28}$$

Proof. Define $\mathcal{S}_\Sigma$ to be the subspace of $\mathcal{S}$ whose elements have the same eigenvalues: $\Sigma = diag\{\sigma_1, \sigma_2, \sigma_3\}$ with $\sigma_1 \geq \sigma_2 \geq \sigma_3$. Thus every matrix $E \in \mathcal{S}_\Sigma$ has the form $E = V_1 \Sigma V_1^T$ for some $V_1 \in SO(3)$. To simplify the notation, define $\Sigma_\lambda = diag\{\lambda_1, \lambda_2, \lambda_3\}$. We now prove this theorem by two steps.

Step One: We prove that the special symmetric matrix $E \in \mathcal{S}_\Sigma$ which minimizes the error $\|E - F\|_f^2$ is given by $E = V\Sigma V^T$. Since $E \in \mathcal{S}_\Sigma$ has the form $E = V_1 \Sigma V_1^T$, we get:

$$\|E - F\|_f^2 = \|V_1 \Sigma V_1^T - V \Sigma_\lambda V^T\|_f^2 = \|\Sigma_\lambda - V^T V_1 \Sigma V_1^T V\|_f^2. \tag{29}$$

Define $W = V^T V_1 \in SO(3)$. Then:

$$\|E - F\|_f^2 = \|\Sigma_\lambda - W\Sigma W^T\|_f^2 = tr(\Sigma_\lambda^2) - 2tr(W\Sigma W^T \Sigma_\lambda) + tr(\Sigma^2). \tag{30}$$

Using the fact that $\sigma_2 = \sigma_1 + \sigma_3$ and W is a rotation matrix, we get:

$$\begin{aligned} tr(W\Sigma W^T \Sigma_\lambda) &= \sigma_1(\lambda_1(1 - w_{13}^2) + \lambda_2(1 - w_{23}^2) + \lambda_3(1 - w_{33}^2)) \\ &\quad + \sigma_3(\lambda_1(1 - w_{11}^2) + \lambda_2(1 - w_{21}^2) + \lambda_3(1 - w_{31}^2)). \end{aligned} \tag{31}$$

Minimizing $\|E - F\|_f^2$ is equivalent to maximizing $tr(W\Sigma W^T \Sigma_\lambda)$. From (31), $tr(W\Sigma W^T \Sigma_\lambda)$ is maximized if and only if $w_{13} = w_{23} = 0$, $w_{33}^2 = 1$, $w_{21} = w_{31} = 0$ and $w_{11}^2 = 1$. Since W is a rotation matrix, we also have $w_{12} = w_{32} = 0$ and $w_{22}^2 = 1$. All possible W give a unique matrix in $\mathcal{S}_\Sigma$ which minimizes $\|E - F\|_f^2$: $E = V\Sigma V^T$.

Step Two: From step one, we only need to minimize the error function over the matrices which have the form $V\Sigma V^T \in \mathcal{S}$. The optimization problem is then converted to one of minimizing the error function:

$$||E - F||_f^2 = (\lambda_1 - \sigma_1)^2 + (\lambda_2 - \sigma_2)^2 + (\lambda_3 - \sigma_3)^2 \quad (32)$$

subject to the constraint: $\sigma_2 = \sigma_1 + \sigma_3$. The formula (28) for $\sigma_1, \sigma_2, \sigma_3$ are directly obtained from solving this minimization problem.

An important property of this projection is that it is statistically unbiased [9]. That is, if components of the essential vector $\mathbf{e}$ are corrupted by identically independent (symmetric) zero-mean noise, this projection gives an unbiased estimate of the true special symmetric matrix.

Remark 1. For symmetric matrices which do not satisfy conditions $\lambda_1 \geq 0$ or $\lambda_3 \leq 0$, one may simply choose $\lambda_1' = max(\lambda_1, 0)$ or $\lambda_3' = min(\lambda_3, 0)$.

We then have an eigenvector-decomposition based algorithm for estimating 3D velocity from optical flow:

Four-Step 3D Velocity Estimation Algorithm:

1. **Estimate Essential Vector:** For a given set of optical flows: $(\mathbf{q}^i, \mathbf{u}^i)$, $i = 1, \ldots, m$, find the vector $\mathbf{e}$ which minimizes the error function $V(\mathbf{e}) = ||A\mathbf{e}||^2$ subject to the condition $||\mathbf{e}|| = 1$;
2. **Recover the Special Symmetric Matrix:** Recover the vector $v_0 \in S^2$ from the first three entries of $\mathbf{e}$ and the symmetric matrix $s \in \mathbb{R}^{3\times3}$ from the remaining six entries.[2] Find the eigenvalue decomposition of the symmetric matrix $s = V_1 diag\{\lambda_1, \lambda_2, \lambda_3\} V_1^T$ with $\lambda_1 \geq \lambda_2 \geq \lambda_3$. Project the symmetric matrix s onto the special symmetric space $\mathcal{S}$. We then have the new $s = V_1 diag\{\sigma_1, \sigma_2, \sigma_3\} V_1^T$ with: $\sigma_1 = (2\lambda_1 + \lambda_2 - \lambda_3)/3$, $\sigma_2 = (\lambda_1 + 2\lambda_2 + \lambda_3)/3$, and $\sigma_3 = (2\lambda_3 + \lambda_2 - \lambda_1)/3$.
3. **Recover Velocity from the Special Symmetric Matrix:** Define $\lambda = \sigma_1 - \sigma_3 \geq 0$ and $\theta = \arccos(-\sigma_2/\lambda) \in [0, \pi]$. Let $V = V_1 R_Y^T\left(\frac{\theta}{2} - \frac{\pi}{2}\right) \in SO(3)$ and $U = -VR_Y(\theta) \in O(3)$. Then the four possible 3D velocities corresponding to the special symmetric matrix s are given by:

$$\begin{aligned} \hat{\omega} &= UR_Z(\pm\frac{\pi}{2})\Sigma_\lambda U^T, \quad \hat{v} = VR_Z(\pm\frac{\pi}{2})\Sigma_1 V^T \\ \hat{\omega} &= VR_Z(\pm\frac{\pi}{2})\Sigma_\lambda V^T, \quad \hat{v} = UR_Z(\pm\frac{\pi}{2})\Sigma_1 U^T \end{aligned} \quad (33)$$

   where $\Sigma_\lambda = diag\{\lambda, \lambda, 0\}$ and $\Sigma_1 = diag\{1, 1, 0\}$;
4. **Recover Velocity from the Differential Essential Matrix:** From the four velocities recovered from the special symmetric matrix s in step 3, choose the pair (ω^*, v^*) which satisfies: ${v^*}^T v_0 = \max_i v_i^T v_0$. Then the estimated 3D velocity (ω, v) with $\omega \in \mathbb{R}^3$ and $v \in S^2$ is given by: $\omega = \omega^*, v = v_0$.

[2] In order to guarantee v_0 to be of unit length, one needs to "re-normalize" $\mathbf{e}$, *i.e.* multiply $\mathbf{e}$ by a scalar such that the vector determined by the first three entries is of unit length.

Both v_0 and v^* contain recovered information about the linear velocity. However, experimental results show that, statistically, within the tested noise levels (next section), v_0 always yields a better estimate than v^* . We thus simply choose v_0 as the estimate. Nonetheless, one can find statistical correlations between v_0 and v^* (experimentally or analytically) and obtain better estimate, using both v_0 and v^*. Another potential way to improve this algorithm is to study the systematic bias introduced by the least square method in step 1. A similar problem has been studied by Kanatani [5] and an algorithm was proposed to remove such bias from Zhuang's algorithm [20].

Remark 2. Since both $E, -E \in \mathcal{E}_1'$ satisfy the same set of differential Longuet-Higgins constraints, both $(\omega, \pm v)$ are possible solutions for the given set of optical flows. However, one can discard the ambiguous solution by adding the "positive depth constraint".

Remark 3. By the way of comparison to the Heeger and Jepson's algorithm [3], note that the equation $A\mathrm{e} = 0$ may be rewritten to highlight the dependence on optical flow as: $[A_1(\mathbf{u}) \mid A_2]\,\mathrm{e} = 0$, where $A_1(\mathbf{u}) \in \mathbb{R}^{m\times 3}$ is a linear function of the measured optical flow and $A_2 \in \mathbb{R}^{m\times 6}$ is a function of the image points alone. Heeger and Jepson compute a left null space to the matrix A_2 ($C \in \mathbb{R}^{(m-6)\times m}$) and solve the equation: $CA_1(\mathbf{u})v = 0$ for v alone. Then they use v to obtain ω. Our method simultaneously estimates $v \in \mathbb{R}^3, s \in \mathbb{R}^6$. We make a simulation comparison of these two algorithms in section 4.

Note this algorithm is not optimal in the sense that the recovered velocity does not necessarily minimize the originally picked error function $||A\mathrm{e}(\omega, v)||^2$ on $\mathcal{E}_1'$ (same for the three-step SVD based algorithm in the discrete case [9]). However, this algorithm only uses linear algebra techniques and is thus simpler and does not try to optimize on the submanifold $\mathcal{E}_1'$.

4 Experimental Results

We carried out initial simulations in order to study the performance of our algorithm. We chose to evaluate it in terms of bias and sensitivity of the estimate with respect to the noise in the optical flow measurements. Preliminary simulations were carried out with perfect data which was corrupted by zero-mean Gaussian noise where the standard deviation was specified in terms of pixel size and was independent of velocity. The image size was considered to be 512x512 pixels.

Our algorithm has been implemented in Matlab and the simulations have been performed using example sets proposed by [15] in their paper on comparison of the egomotion estimation from optical flow[3]. The motion estimation was performed by observing the motion of a random cloud of points placed in front

[3] We would like to thank the authors in [15] for making the code for simulations of various algorithms and evaluation of their results available on the web.

of the camera. Depth range of the points varied from 2 to 8 units of the focal length, which was considered to be unity. The results presented below are for fixed field of view (FOV) of 60 degrees. Each simulation consisted of 500 trials with a fixed noise level, FOV and ratio between the image velocity due to translation and rotation for the point in the middle of the random cloud. Figures 1 and 2 compare our algorithm with Heeger and Jepson's linear subspace algorithm. The presented results demonstrate the performance of the algorithm while translating along X-axis and rotating around Z-axis with rate of 23^o per frame. The analysis of the obtained results of the motion estimation algorithm was performed using benchmarks proposed by [15]. The bias is expressed as an angle between the average estimate out of all trails (for a given setting of parameters) and the true direction of translation and/or rotation. The sensitivity was computed as a standard deviation of the distribution of angles between each estimated vector and the average vector in case of translation and as a standard deviation of angular differences in case of rotation. We further evaluated the

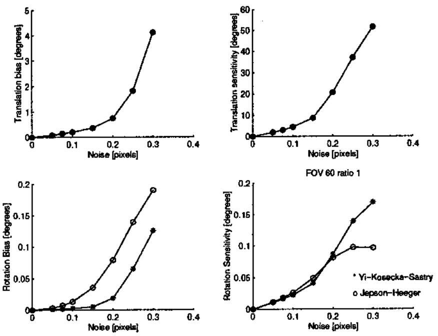

Fig. 1. The ratio between the magnitude of linear and angular velocity is 1.

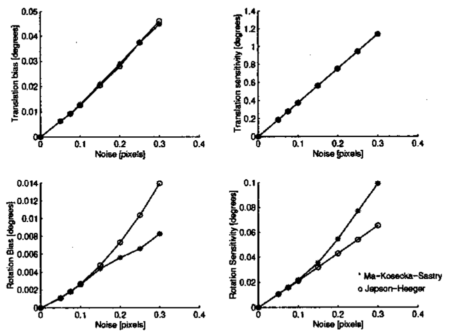

Fig. 2. The ratio between the magnitude of linear and angular velocity is 10.

algorithm by varying the direction of translation and rotation and their relative speed. The choice of the rotation axis did not influence the translation estimates. In the case of the rotation estimate our algorithm is slightly better compared to Heeger and Jepson's algorithm. This is due to the fact that in our case the rotation is estimated simultaneously with the translation so its bias is only due to the bias of the initially estimated differential essential matrix obtained by linear least squares techniques. This is in contrary to the rotation estimate used by Jepson and Heeger's algorithm which uses another least-squares estimation by substituting already biased translational estimate to compute the rotation. The translational estimates are essentially the same since the translation was estimated out from v_0, skew symmetric part of the differential essential matrix. Increasing the ratio between magnitudes of translational and angular velocities improves the bias and sensitivity of both algorithms.

The evaluation of the results and more extensive simulations are currently underway. We believe that through thorough understanding of the source of translational bias we can obtain even better performance by utilizing additional information about linear velocity, which is embedded in the symmetric part of the differential essential matrix. In the current simulations translation was estimated only from v_0 skew symmetric part of $\mathbf{e}$.

5 Conclusions and Future Work

This paper presents a unified view of the problem of egomotion estimation using discrete and differential Longuet-Higgins constraint. In both (discrete and differential) settings, the geometric characterization of the space of (differential) essential matrices gives a natural geometric interpretation for the number of possible solutions to the motion estimation problem. In addition, in the differential case, understanding of the space of differential essential matrices leads to a new egomotion estimation algorithm, which is a natural counterpart of the three-step SVD based algorithm developed for the discrete case by [17].

In order to exploit temporal coherence of motion and improve algorithm's robustness, a dynamic (recursive) motion estimation scheme, which uses implicit extended Kalman filter for estimating the essential parameters, has been proposed by Soatto *et al* [14] for the discrete case. The same ideas certainly apply to our algorithm.

In applications to robotics, a big advantage of the differential approach over the discrete one is that it can make use of nonholonomic constraints (*i.e.* constraints that confine the infinitesimal motion of the mobile base but not the global motion) and simplify the motion estimation algorithms [9]. An example study of vision guided nonholonomic system can be found in [10]. In this paper, we have assumed that the camera is ideal. This approach can be extended to uncalibrated camera case, where the motion estimation and camera self-calibration problem can be solved simultaneously, using the differential essential constraint [19, 1]. In this case, the essential matrix is replaced by the fundamental matrix which captures both motion information and camera intrinsic parameters. It is shown in [1], that the space of such fundamental matrices is a 7-dimensional algebraic variety in $\mathbb{R}^{3\times3}$. Thus, besides five motion parameters, only two extra intrinsic parameters can be recovered.

References

1. Michael J. Brooks, Wojciech Chojnacki, and Luis Baumela. Determining the egomotion of an uncalibrated camera from instantaneous optical flow. *in press*, 1997.
2. A. R. Bruss and B. K. Horn. Passive navigation. *Computer Graphics and Image Processing*, 21:3–20, 1983.
3. D. J. Heeger and A. D. Jepson. Subspace methods for recovering rigid motion I: Algorithm and implementation. *International Journal of Computer Vision*, 7(2):95–117, 1992.
4. A. D. Jepson and D. J. Heeger. Linear subspace methods for recovering translation direction. *Spatial Vision in Humans and Robots, Cambridge Univ. Press*, pages 39–62, 1993.

5. K. Kanatani. 3D interpretation of optical flow by renormalization. *International Journal of Computer Vision*, 11(3):267–282, 1993.
6. Kenichi Kanatani. *Geometric Computation for Machine Vision.* Oxford Science Publications, 1993.
7. H. C. Longuet-Higgins. A computer algorithm for reconstructing a scene from two projections. *Nature*, 293:133–135, 1981.
8. Waxman A. M., Kamgar-Parsi B., and Subbarao M. Closed form solutions to image flow equations for 3d structure and motion. *International Journal of Computer Vision 1*, pages 239–258, 1987.
9. Yi Ma, Jana Košecká, and Shankar Sastry. Motion recovery from image sequences: Discrete viewpoint vs. differential viewpoint. *Electronic Research Laboratory Memorandum, UC Berkeley*, UCB/ERL, June 1997.
10. Yi Ma, Jana Košecká, and Shankar Sastry. Vision guided navigation for a nonholonomic mobile robot. *Electronic Research Laboratory Memorandum, UC Berkeley*, UCB/ERL(M97/42), June 1997.
11. Stephen Maybank. *Theory of Reconstruction from Image Motion.* Springer Series in Information Sciences. Springer-Verlag, 1993.
12. Philip F. McLauchlan and David W. Murray. A unifying framework for structure and motion recovery from image sequences. In *Proceeding of Fifth International Conference on Computer Vision*, pages 314–320, Cambridge, MA, USA, 1995. IEEE Comput. Soc. Press.
13. Richard M. Murray, Zexiang Li, and Shankar S. Sastry. *A Mathematical Introduction to Robotic Manipulation.* CRC press Inc., 1994.
14. S. Soatto, R. Frezza, and P. Perona. Motion estimation via dynamic vision. *IEEE Transactions on Automatic Control*, 41(3):393–413, March 1996.
15. T. Y. Tian, C. Tomasi, and D. J. Heeger. Comparison of approaches to egomotion computation. In *Proceedings of 1996 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 315–20, Los Alamitos, CA, USA, 1996. IEEE Comput. Soc. Press.
16. Carlo Tomasi and Takeo Kanade. Shape and motion from image streams under orthography. *Intl. Journal of Computer Vision*, 9(2):137–154, 1992.
17. G. Toscani and O. D. Faugeras. Structure and motion from two noisy perspective images. *Proceedings of IEEE Conference on Robotics and Automation*, pages 221–227, 1986.
18. Roger Y. Tsai and Thomas S. Huang. Uniqueness and estimation of three-dimensional motion parameters of rigid objects with curved surfaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-6(1):13–27, January 1984.
19. T. Vieville and O. D. Faugeras. Motion analysis with a camera with unknown, and possibly varying intrinsic parameters. *Proceedings of Fifth International Conference on Computer Vision*, pages 750–756, June 1995.
20. Xinhua Zhuang and R. M. Haralick. Rigid body motion and optical flow image. *Proceedings of the First International Conference on Artificial Intelligence Applications*, pages 366–375, 1984.
21. Xinhua Zhuang, Thomas S. Huang, and Narendra Ahuja. A simplified linear optic flow-motion algorithm. *Computer Vision, Graphics and Image Processing*, 42:334–344, 1988.