Pose Estimation Using Line-Based Dynamic Vision and Inertial Sensors

Henrik Rehbinder, *Member, IEEE,* and Bijoy K. Ghosh, *Fellow, IEEE*

***Abstract*—In this paper, an observer problem from a computer vision application is studied. Rigid body pose estimation using inertial sensors and a monocular camera is considered and it is shown how rotation estimation can be decoupled from position estimation. Orientation estimation is formulated as an observer problem with implicit output where the states evolve on $SO(3)$. A careful observability study reveals interesting group theoretic structures tied to the underlying system structure. A locally convergent observer where the states evolve on $SO(3)$ is proposed and numerical estimates of the domain of attraction is given. Further, it is shown that, given convergent orientation estimates, position estimation can be formulated as a linear implicit output problem. From an applications perspective, it is outlined how delayed low bandwidth visual observations and high bandwidth rate gyro measurements can provide high bandwidth estimates. This is consistent with real-time constraints due to the complementary characteristics of the sensors which are fused in a multirate way.**

***Index Terms*—Dynamic vision, implicit output, inertial sensors, lie group, observers.**

I. Introduction

THE fundamental problem of rigid body state estimation is that there is no single sensor that measures position and orientation (pose) with respect to an inertial frame with high bandwidth and long-term stability. The most known way of deriving position and orientation is perhaps inertial navigation where rate gyros and accelerometers are integrated. The bandwidth of such a system is typically good, but long-term stability cannot be obtained due to integrated errors. This problem is further enhanced for cheap sensors where drifting zero-level offsets will result in an approximately linear error growth in orientation and a quadratic error growth in position. There are long-term stable gravity based sensors for attitude estimation (pitch and roll) such as liquid-filled inclinometers. The working principle of inclinometers is simply that of a water level. All gravity based attitude sensors are sensitive to translational accelerations as it is impossible to distinguish between gravity and inertial forces. Inclinometers are also subject to low bandwidth. There have been numerous attempts to combine the above mentioned sensors by using different heuristics. Aerospace applications are considered by Greene [13] and Lefferts *et al.* [18]. Foxlin *et al.* [10], [11] has studied a virtual reality helmet application and among the robotics applications we would like to point out Vaganay *et al.* [33], Madni *et al.* [20] and also Barshan and Durrant-Whyte [4]. Sakaguchi *et al.* [28] and Smith *et al.* [29] have also considered the problem. Theoretically justified linear approaches have been studied by Baerveldt and Klang [2] and Balaram [3] while nonlinear approaches have been investigated by Rehbinder and Hu [25]–[27]. The dominating algorithm among the heuristic studies is the extended Kalman filter (EKF). Regardless of which method is used, yaw can never be obtained and position is not considered.

This paper proposes the use of vision and inertial sensors, fused with a provably stable observer.

A. Vision

Using computer vision as a stand alone sensor for pose estimation is quite a standard task; see, for example, the surveys by Huang and Netravali [14] and by Olensis [23]. The main body of research has been devoted to point correspondence based algorithms. An alternative would be a line correspondence based algorithm such as proposed by Spetsakis [31], [32], Dornaika and Garcia [9] and by Christy and Horaud [7], [8]. Line correspondences have the advantage of being more robust than point correspondences. The disadvantages are that line tracking algorithms are computationally more intensive and low sampling frequency and long time delays can therefore be expected. Furthermore, they are mathematically more complicated. A specific field of computer vision, where feature tracking and correspondences are studied, is *dynamic vision* [15], [21], [30] which utilizes control of dynamical systems where an underlying dynamic model for the pose is used. Standard filtering or observer techniques are often used in dynamic vision. The control theoretic approach has also been successfully used inline based visual servoing by Andreff *et al.* [1]. In dynamic vision, the above mentioned time delays and low sampling frequency typically pose a serious problem when implementing the algorithms developed. This is especially true if a continuous-time approach has been used. Such a continuous-time approach is what we will propose, but we claim to have an algorithm that can be implemented in such a way that the estimates are high bandwidth, real-time estimates. The reason for this is that our method uses both vision and inertial sensors and the implementation proposed is a multirate implementation where the vision is run at a low frequency and the inertial sensors at a high frequency. Our motivation for why this will work is that the inertial sensors provide a very accurate high-frequency information about the motion. Visual information is needed only to estimate the low frequency components of the motion and can therefore be sampled

Manuscript received August 17, 2001; revised July 9, 2002. Recommended by Associate Editor K. Gu. This work was supported by the SSF through its Centre for Autonomous Systems, by the Royal Institute of Technology, and by the National Science Foundation under Grants ECS9720357 and ECS9976174.

H. Rehbinder is with Optimization and Systems Theory, Royal Institute of Technology, 10044 Stockholm, Sweden (e-mail: henrik.rehbinder@raysearchlabs.com).

B. K. Ghosh is with the Systems Science and Mathematics, Washington University, Saint Louis, MO 63130-4899, USA (e-mail: ghosh@zach.wustl.edu).

Digital Object Identifier 10.1109/TAC.2002.808464

at a low frequency. Approaches along this line of thinking has been proposed earlier by Kurazume and Hirose [17], Lobo and Dias [19] and by Rehbinder and Ghosh [24]. Apart from these real-time oriented advantages of using inertial sensors and vision the robustness of the system will also be enhanced as two different sensors are used, instead of just one.

B. Observers

Pose estimation is a delicate subject since the set of all orientations is not a vector space, but a manifold and a Lie group, $SO(3)$. In order to circumvent these mathematical difficulties, local representations such as Euler angles are often used. A consequence of this is complicated nonlinear equations and most authors resort to *EKF*-based solutions (see the aforementioned references). In this paper, we consider observability and propose an observer where such local representations are not used. The use of a global representation lends the paper a geometric flavor, following many ideas proposed by Koditschek [16], Bullo and Murray [6] and by Soatto *et al.* [30]. However we do not make explicit use of the differential geometric framework and stick to a more standard exposition. The state evolves on $SO(3)$ and the output is implicitly characterized. The observability analysis for this system reveals interesting group theoretic generalizations of standard linear observability and the observer is designed in such a way that the estimated state evolves on $SO(3)$. The rotation estimation is solved independently of the position and these estimates can subsequently be used to formulate the position estimation as a straightforward problem with linear implicit output function.

We are of the opinion that the major contributions of this paper is a new, theoretically sound algorithm that fuses data from vision and rate gyros and an implementation which enables fast tracking of ego-motion with a slow vision system. We also outline how these results on orientation estimation can be used to formulate the position estimation problem as a linear implicit output problem. From a control theoretic point of view, the observer proposed is interesting as it evolves not on a vector space, but on a manifold and the results of the observability analysis has quite an interesting group structure. A major drawback is the lack of experimental results and that we do not consider how to actually find the line correspondences. Further, we consider the observed lines to have known orientations. The outline of this paper is as follows. In Section II, we derive the mathematical statement of the problem and in Section III, we study observability. In Section IV, we describe the orientation estimation algorithm and prove its convergence and in Section V, we derive the formulation of the position estimation problem. Section VI is devoted to the multirate implementation and to handling the problems with low vision sampling frequency and delays. We provide simulations in Section VII. In Appendix A, we provide details of the calculations in Section III.

II. Problem Formulation

We will now derive the mathematical formulation of the problem and describe a moving rigid body equipped with a camera and a strap-down inertial measurement unit (IMU), i.e., a body fixed rate-gyro/accelerometer package.

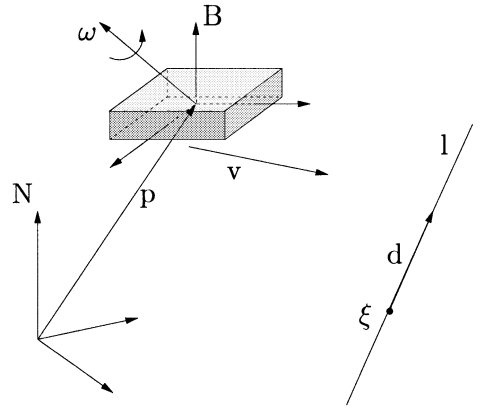

Fig. 1. Rigid body moving in space.

A. Rigid Body Kinematics

Consider a rigid body moving in an inertial space as shown in Fig. 1. We are interested in estimating its orientation and position relative to the surrounding world. Our primary concern here will be rotation estimation. We denote the inertial frame by N and the body fixed frame by B. Let x be an arbitrary point in space and denote by x^N its coordinates in the N-frame and by x^B its coordinates in the B-frame (we will use the same notation for other variables as well). The two coordinate vectors are, due to the rigid body motion related via

$$x^B = R\left(x^N - r^N\right) \tag{1}$$

where r is the vector from the N-origin to the B-origin and where R is a rotation matrix. If we now denote $p = r^N$, $v = dr^N/dt$ and $a = d^2r^N/dt^2$ then rigid body kinematics gives rise to the following description of the camera motion:

$$\begin{cases} \dot{R} = \Omega(t)R \\ \dot{p} = v \\ \dot{v} = a \end{cases} \tag{2}$$

where

$$\Omega = S(\omega) = \begin{pmatrix} 0 & \omega_3 & -\omega_2 \\ -\omega_3 & 0 & \omega_1 \\ \omega_2 & -\omega_1 & 0 \end{pmatrix} \tag{3}$$

is called the wedge matrix due to the fact that $\Omega x = -\omega \wedge x$. The symbol ω is the angular velocity expressed in the B-frame. Now let us consider lines l_i fixed in inertial space with the representation

$$l_i = \left\{x^N \in \mathbb{R}^3 : x^N = \xi_i^N + d_i^N s, s \in \mathbb{R}\right\} \tag{4}$$

where ξ_i is an arbitrary point on l_i and d_i its direction vector. The B-coordinates of the lines will be time varying due to camera motion. Using (1) and (4), we obtain

$$l_i = \left\{x^B \in \mathbb{R}^3 : x^B = R\left(\xi_i^N - p + d_i^N s\right), s \in \mathbb{R}\right\} \tag{5}$$

which describes the line in B-coordinates.

Remark 2.1: The representations of rigid body motion and lines can be cast in a differential geometric framework where pose is described as an element in the Lie group of Euclidean transformations

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

and the lines are represented as points in the Grassmanian manifold Grass(2,4).

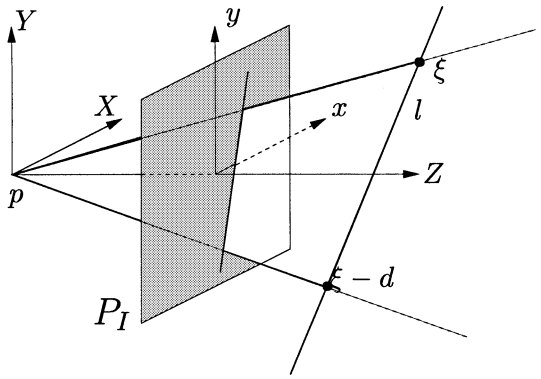

Fig. 2. Line projected on the image plane.

B. Sensor Measurements

We consider the body to be equipped with a strap-down IMU. The rate gyros provide measurements of $\omega(t)$ and we will therefore take the angular velocity as a known entity in the problem. A strap-down accelerometer measures inertial forces and gravity in the B-frame so the output can be written as

$$u = R\left(a - g^N\right) \tag{6}$$

where $g^N = (00 - g)$ and $g = 9.81$ m/s^2. It is clear that we can write a from (6) as

$$a = R^T u + g^N. \tag{7}$$

For a discussion on inertial sensor offsets, see Section VII. Finally, we consider the camera. We let the camera be a perfect normalized pinhole camera. For simplicity, we let the camera focal point coincide with the B-origin and let the optical axis coincide with the z-axis of the B-frame. To see how a line l projects onto the image plane P_I, consider the plane P_l defined by the camera focal point and the line l (Fig. 2). The line on the image plane is given by the intersection of P_l and P_I. Two vectors in P_l are $R(\xi - p)$ and Rd. Thus, a normal vector η to P_l is given by

$$\eta = [Rd] \wedge [R(\xi - p)] = R[d \wedge (\xi - p)] \tag{8}$$

where $\wedge$ is the wedge (cross) product. On the image plane $z = 1$, the line is described by the equation

$$\eta^T \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = 0 \tag{9}$$

where x and y are the image plane coordinates. From (9), it may be concluded that we can only derive η up to length from measurements of the projected line. To get a parameterization of this measurement it is necessary to make some assumption on the length or orientation of η. If we, for example, consider the case where we use room corners for indoor navigation it is quite reasonable to assume that the lines we observe never coincide with the B-frame z-axis. If they would, that corresponds to the camera actually being inside the walls. We therefore assume that η is nonzero. An alternative approach would be to make assumptions on the orientation of η. This is studied in [15]. To avoid technical difficulties we also assume that $p(t)$ is bounded. We make the following assumption.

Assumption 2.1: There is a constant $\kappa_1 > 0$ such that $||d_i \wedge (p(t) - \xi_i)|| \geq \kappa_1 \; \forall \, t$, i.e., d_i and $p(t) - \xi_i$ are not parallel. Furthermore, there is a $\kappa_2 < \infty$ such that $||p(t)|| \leq \kappa_2 \; \forall \, t$.

Under this assumption, it follows that $||\eta|| \neq 0$ and that we can take the following vectors as our observations:

$$y_i = \mu_i \frac{\eta_i}{||\eta_i||} = \mu_i \frac{R\rho_i(t)}{||\rho_i(t)||} \tag{10}$$

where $\mu_i \in \{-1, 1\}$ is an unknown sign ambiguity parameter and where

$$\rho_i(t) = d_i \wedge (p(t) - \xi_i). \tag{11}$$

It is worth pointing out that even though the line parameters are to a certain extent arbitrary, the problem actually is well-posed. The sign parameter μ_i accounts for unknown line directions and the normalization for magnitude. Regarding the point ξ_i on the line, take another point which can then be written as $\zeta_i = \xi_i + sd_i$ for some s. It follows that $d_i \wedge (p(t) - \zeta_i) = d_i \wedge (p(t) - \xi_i)$, so the output equation is well defined. Note especially that without any loss of generality we can take $||d_i|| = 1$ which we shall assume in the sequel.

Assumption 2.2: $||d_i|| = 1$

C. Formulation as an Implicit Output System

The input–output system we have from (2), (7), and (10) is

$$\begin{cases} \dot{R} = \Omega(t)R \\ \dot{p} = v \\ \dot{v} = R^T u(t) + g^N \\ y_i = \mu_i \frac{R\rho_i(t)}{||\rho_i(t)||} \end{cases} \tag{12}$$

and our interest is to estimate R and p based on the measurements y_i. The sensor readings ω and u will be used as known inputs to the system. If the goal is to construct a standard Luenberger-type observer, the parameters μ_i would pose a serious problem as they are unknown. This problem can be solved by noting that, from (10), we have

$$h_i(R, y_i) = {d_i}^T R^T y_i = 0 \tag{13}$$

due to orthogonality of the vectors d_i and ρ_i. The formulation of (12) that we can work with, instead, is therefore an implicit output problem

$$\begin{cases} \dot{R} = \Omega(t)R \\ \dot{p} = v \\ \dot{v} = R^T u(t) + g^N \\ 0 = h(R, y) \end{cases} \tag{14}$$

where $h(R, y) = [h_1(R, y_1), \ldots, h_m(R, y_m)]^T$. It has the advantage that there are no unknown terms in the output equation. It will also turn out that we will be able to estimate the rotation R without knowing or estimating p. We will, therefore, in Sections III and IV, consider the less complex problem.

Problem 2.1 (Rotation Estimation): Given the system

$$\begin{cases} \dot{R} = \Omega(t)R \\ 0 = h(R, y) \end{cases} \tag{15}$$

and measurements of $y(t)$ and $\Omega(t)$, the problem is to estimate $R(t)$.

Note that for an observer design based on the system model (15), $p(t)$ is not required.

III. Observability

In order to understand the structure and fundamental limitations of Problem 2.1, we propose to study observability of (15). It will turn out that a careful treatment of the observability properties of the problem will result in insights that are very reasonable and intuitive from the perspective of computer vision. Before showing the actual calculations of observability, we will comment on the structure of the problem. We consider $R(t)$ as the states and $\mu_i, p(t)$ are the unknown parameters in the output (10). We also note that $p(t)$ is time varying. The output equation is linear in R and the state dynamics appear to be linear. However, since the states live on $SO(3)$ and not on a vector space, such a statement lacks meaning. If R is not confined to $SO(3)$, then (15) would have linear state dynamics. This fact will be used later on, in the observability analysis. For a linear system evolving on a linear space, the *unobservable subspace* $\mathcal{O}^c$ is derived by considering all the initial state values that produce the same output. As the states here evolve on $SO(3)$, a C^∞-manifold and a Lie group, what would correspond to the unobservable subspace could be anticipated to be either an unobservable subgroup or a submanifold. We will see that an unobservable subgroup will be the answer.

Now, we will briefly review a standard observability derivation for a normal linear system. The reason for this is that the derivation, to a large extent, can be mimicked for system (15) that evolves on $SO(3)$. Let Σ be a linear system

$$\Sigma: \begin{cases} \dot{x}(t) = A(t)x(t) \\ y(t) = C(t)x(t) \end{cases} \tag{16}$$

evolving on a vector space X. Let us consider the time interval $[t_0, t_0 + T]$ and assume two initial conditions $x(t_0) = x_0$ and $x(t_0) = \xi_0$ such that

$$y_{x_0}(t) = y_{\xi_0}(t) \quad \forall\, t \in [t_0, t_0 + T] \tag{17}$$

where $y_{x_0}(t)$ is the output corresponding to the initial value x_0, and likewise for $y_{\xi_0}(t)$. It is easily derived that

$$C(t)\Phi(t, t_0)(x_0 - \xi_0) = 0 \qquad \forall\, t \in [t_0, t_0 + T] \tag{18}$$

where $\Phi(t, t_0)$ is the transition matrix for (16). Define the operator U by

$$\begin{aligned} U: X &\to L_2(t_0, t_0 + T) \\ x &\mapsto C(t)\Phi(t, t_0)x \end{aligned} \tag{19}$$

where $L_2(t_0, t_0 + T)$ is the set all functions

$$f: [t_0, t_0 + T] \to \mathbb{R}$$

such that

$$\int_{t_0}^{t_0+T} f^2(\tau)d\tau < \infty.$$

We see that (18) can be written as

$$-\xi_0 + x_0 \in \ker U \tag{20}$$

and we define the *unobservable subspace* as

$$\mathcal{O}^c = \ker U. \tag{21}$$

Furthermore, the finite-dimensional observability Gramian $M(t_0, T)$ for which it holds that $\ker M = \ker U$ is derived and we obtain

$$\begin{aligned} M(t_0, T) =& U^*U \\ =& \int_{t_0}^{t_0+T} \left[C(\tau)e^{A(\tau - t_0)}\right]' \\ & \times \left[C(\tau)e^{A(\tau - t_0)}\right] d\tau \end{aligned} \tag{22}$$

where U^* is the adjoint operator of U. If there is a $T < \infty$ and an $\epsilon > 0$ such that $M(t_0, T) \geq \epsilon I\, \forall\, t_0$ then the system is called strongly observable. If we now return to (20) and pay particular attention to the fact that the vector space X is also a group under vector addition and that, $-\xi_0$ is the inverse element of ξ_0, then (20) can be written as

$$(\xi_0)^{-1} \oplus x_0 \in \mathcal{O}^c \tag{23}$$

where $\mathcal{O}^c$ is now viewed as the unobservable *subgroup* and where $\oplus$ is the group operation (vector addition). This group theoretic thinking provides us with the abstraction needed to study (15).

Now, we go back to our system (15) and consider a time interval $[t_0, t_0 + T]$. Let us assume $R(t_0) = R_0$ and take some $p(t), \mu_i$ (which defines the output $y(t)$) and assume that we are observing m lines l_i. Let the lines be ordered in such a way that the first M lines have linearly independent direction vectors. Obviously $M \leq 3$. The outputs are given, as in (10), by $y_i(t) = R(t)(\rho_i(t)/\|\rho_i(t)\|)\, \mu_i : i = 1, \ldots, m$. Let $R(t_0) = S_0$ be a different initial value and consider parameterizing all choices of S_0 that satisfy the implicit output equation

$$0 = h_i(R_{S_0}(t), y_{i,R_0}(t)) \qquad \forall\, t \in [t_0, t_0 + T],\ \ i = 1, \ldots, m. \tag{24}$$

In (24), $R_{S_0}(t)$ is the solution of the equation

$$\dot{R} = \Omega(t)R \tag{25}$$

assuming $R(t_0) = S_0$. $y_{i,R_0}(t)$ is the output of (12) assuming $R(t_0) = R_0$. This is the natural implicit output generalization of (17). Let $\Phi(t, t_0)$ be the transition matrix of (25). We obtain

$$R_{S_0}(t) = \Phi(t, t_0)S_0 \tag{26}$$

and

$$y_{i,R_0}(t) = \Phi(t, t_0)R_0 \frac{\rho_i(t)}{\|\rho_i(t)\|}\mu_i. \tag{27}$$

The implicit output (24) is equivalent to

$$0 = d_i^T S_0^T R_0 \rho_i(t) \qquad \forall\, t \in [t_0, t_0 + T], \qquad i = 1 \ldots m \tag{28}$$

where the fact that $\Phi(t, t_0) \in SO(3)$ has been used. If we, in analogy with (19), redefine U as

$$\begin{aligned} U: \mathbb{R}^{3\times 3} &\to L_2(t_0, t_0 + T) \\ X &\mapsto \begin{pmatrix} d_1^T X \rho_1(t) \\ \vdots \\ d_m^T X \rho_m(t) \end{pmatrix} \end{aligned} \tag{29}$$

then (28) can be written equivalently as

$$\left(S_0^{-1}\right) \oplus R_0 \in \ker U \cap SO(3) \tag{30}$$

where the group structure of $SO(3)$ is now used ($S_0^T = S_0^{-1}$ and $\oplus$ is matrix multiplication). This formula should be compared to (23) and it seems clear that if $\ker U \cap SO(3)$ is a group, then (30) is the natural generalization of (23). It will be made clear that, if strong observability holds, then $\ker U \cap SO(3)$ actually is a subgroup and it, therefore, makes sense to define the following.

Definition 3.1: The *unobservable subgroup* of the implicit system (15) is given by

$$\mathcal{O}^c_{SO(3)} = \Big\{ \overline{R} \in SO(3) \colon d_i^T \overline{R} \rho_i(t) = 0 \quad \forall\, t \in [t_0, t_0+T],\ i = 1 \dots m \Big\}. \tag{31}$$

From (30), we understand that we can first compute $\ker U$ using linear methods and then intersect with $SO(3)$. The associated calculations are sketched in Appendix A.1 and here we only state the main results.

Definition 3.2: The *observability sub-Gramians* for (15) are

$$M_i(t_0, T) = \int_{t_0}^{t_0+T} \rho_i(\tau)\rho_i^T(\tau) d\tau \Big|_{\{d_i\}^\perp}, \qquad i \in \{1 \dots m\} \tag{32}$$

where $\{d_i\}^\perp$ is the orthogonal complement of $\operatorname{span}\{d_i\}$. The following definition corresponds to standard strong observability for time-varying systems.

Definition 3.3: The implicit output system (15) is called *strongly observable* if $\exists\, \bar{M} \le M$ such that $\forall\, i = 1, \dots, M$, $\exists\, \epsilon_i > 0, T < \infty$ such that $\forall\, t_0$

$$M_i(t_0, T) \ge \epsilon_i I. \tag{33}$$

The introduction of $\bar{M}$ in the definition is merely a technicality and for clarity of presentation we will assume that $\bar{M} = M$. For an interpretation of (33), see Remark 3.3 and the derivations in Appendix A.2. For a strongly observable system, analytic expressions for the unobservable subgroups can be found.

Theorem 3.1: For a strongly observable (15) the unobservable subgroup is given by one of the following three alternatives.

1) One linearly independent line ($M = 1$)

$$\mathcal{O}^c_{SO(3)} = \Big\{ I, D^{-T} \begin{pmatrix} 1 & 0 & 0 \\ 0 & r_1 & r_2 \\ 0 & -r_2 & r_1 \end{pmatrix} D^T \quad D^{-T} \begin{pmatrix} -1 & 0 & 0 \\ 0 & r_1 & r_2 \\ 0 & r_2 & -r_1 \end{pmatrix} D^T \quad \forall\, r_1^2 + r_2^2 = 1 \Big\}. \tag{34}$$

2) Two linearly independent lines ($M = 2$)

$$\mathcal{O}^c_{SO(3)} = \left\{ I, D^{-T} J_i D^T \quad i = \left\{ \begin{array}{c} 3 \text{ if } d_1 \not\perp d_2 \\ 1, 2, 3 \text{ if } d_1 \perp d_2 \end{array} \right\} \right. . \tag{35}$$

3) Three linearly independent lines ($M = 3$)

$$\mathcal{O}^c_{SO(3)} = \{ I, D^{-T} J_i D^T \ \forall\, i \colon d_i \perp d_j \& d_k \ \forall\, k, j \ne i \} \tag{36}$$

where

$$D = \begin{cases} [d_1 \dots d_M f_{M+1} \dots f_3] & \text{if } M \le 2 \\ [d_1 \dots d_3] & \text{if } M = 3 \end{cases} \tag{37}$$

and where $f_i \in \mathbb{R}^3$ are such that D is of full rank. The matrices J_i are rotation matrices describing rotations of π radians around the ith axis. Thus, they are diagonal matrices with 1 on the iith position and -1 on the remaining two positions.

Observation 3.1: For strongly observable systems, $\mathcal{O}^c_{SO(3)}$ is an Abelian subgroup of $SO(3)$.

Proof: First, consider case 1). The identity element $I \in \mathcal{O}^c_{SO(3)}$. For every R, $D^{-T} R^T D^T \in \mathcal{O}^c_{SO(3)}$ and as $RR^T = I$ and as $SS = I$, every element has an inverse. A straightforward calculation shows that $\mathcal{O}^c_{SO(3)}$ is closed under multiplication. For the cases 2) and 3), $J_i J_i = I$ so every element is its own inverse. It is also straightforward that it is closed under multiplication.

Observation 3.2: Let $R, S \in SO(3)$. Let $R \sim S$, i.e., R is similar to S, if $SR^T \in \mathcal{O}^c_{SO(3)}$. Then $\sim$ is an equivalence relation on $SO(3)$.

Proof: The fact that subgroups define equivalence relations on groups in the above way is a standard fact from group theory [12, p. 120]. □

Given the previous observations, we can now formulate a theorem about observability.

Theorem 3.2 (Observability): If the system (15) is strongly observable in the sense of (33) then it is observable up to the equivalence relation defined by $\mathcal{O}^c_{SO(3)}$.

It should be noted that, as for linear systems, the observable part of the configuration space can be viewed as a quotient space [5, p. 60] defined by the equivalence relation $\sim$.

Remark 3.1: The unobservable subgroups (34)–(36) have many intuitive interpretations. Note first of all that the D^T-matrix corresponds to a change of basis reflecting the line orientations. For the case $M = 1$, the unobservable states consist of arbitrary rotations from the true state around an axis parallel to the observed line. A typical example is that if only vertical lines are observed, then the heading (yaw) cannot be estimated. For the cases $M = 2, 3$ it makes a big difference whether or not the observed lines are orthogonal. We first consider the orthogonal case i.e when $d_i \perp d_j$, $i \ne j$. When $M = 2$, then $\mathcal{O}^c_{SO(3)}$ consists of all the four elements corresponding to I, J_1, J_2, J_3. The states that cannot be distinguished from the true states are those corresponding to π-rotations around the two axes d_1, d_2 and around the axis perpendicular to those. For $M = 3$ the same fact is true so that when all the lines are orthogonal, it does not matter if we observe two lines or three. Considering the other extreme, when all line pairs are nonorthogonal, we have observability for the $M = 2$ case up to a π-rotation around the axis perpendicular to the observed two. For $M = 3$ we have full observability as the only element in $\mathcal{O}^c_{SO(3)}$ is the identity element I.

From an applications perspective, it is rather unfortunate that the scenario with orthogonal lines provides the least information. This is because, in a typical man-made environment, such line configurations are most likely to be present. However, even for the case with orthogonal lines ($M = 2$ or $M = 3$), we can argue that we have a certain practical observability. Recall that the orientations that cannot be distinguished from each other correspond to rotations of π rad (180°); so with a rough knowledge of the orientation, we can single out the correct alternative.

Remark 3.2: If (33) does not hold then it can be argued that the system is still locally observable in the following sense: There are many elements of $\mathcal{O}^c_{SO(3)}$ that depend on $p(t)$. They are however bounded away from I so there is a region around the true state where observability holds. However since $p(t)$ as unknown, it is hard to quantify how large this region is. Furthermore, $\mathcal{O}^c_{SO(3)}$ is no longer a group and the appealing algebraic structure is lost. In Section IV, we present a local observer which does not require strong observability and we provide some numerical estimates of the domain of attraction where it is clear that this domain depends on the position of $p(t)$.

Remark 3.3: To understand the observability condition (33) consider the equation with $\epsilon = 0$. It can be shown (see Appendix A.1) that this would correspond to $(p(\tau) - \xi_i)$ being *confined to a plane containing* d_i. An alternative formulation is that

$$d_i^T((p(t) - \xi_i) \wedge v(t)) = 0. \tag{38}$$

To ensure that this is not the case, we require that $\exists\, T > 0$ and $\epsilon' > 0$ such that $\forall\, t$

$$\int_t^{t+T} \left(d_i^T((p(t) - \xi_i) \wedge v(t))\right)^2 dt \geq \epsilon' \quad \forall\, i \leq M \tag{39}$$

which can be viewed as a condition alternative to (33).

IV. Observer

In this section, we sketch a heuristic motivation for an observer, prove local convergence and numerically estimate its domain of attraction. It should be pointed out that the observer is local in the errors, not in the representation of rigid body rotation. The fact that the observer is local explains why the results on observability, discussed in Section III, is not of importance here. This has been commented on in Remark 3.3

A. Heuristic Motivation

An observer for our orientation estimation problem must be designed in such a way that the estimated states $\hat{R}(t)$ evolve on $SO(3)$. It is known that elements of $SO(3)$ obey

$$\dot{\hat{R}}(t) = S(w(t))\hat{R}(t) \tag{40}$$

for some $w(t)$. $S(w)$ is defined as in (3). As the true states obey (2) it is reasonable to choose

$$w(t) = \omega(t) + v(t) \tag{41}$$

where $v(t)$ acts as a correction term. The resulting observer can be written as

$$\dot{\hat{R}}(t) = \Omega(t)\hat{R}(t) + V(t)\hat{R}(t) \tag{42}$$

where $V(t) = S(v(t))$. This is very reasonable if compared to the standard Luenberger observer for linear systems

$$\dot{\hat{x}}(t) = A(t)\hat{x} + L(t)[y(t) - C(t)\hat{x}(t)]. \tag{43}$$

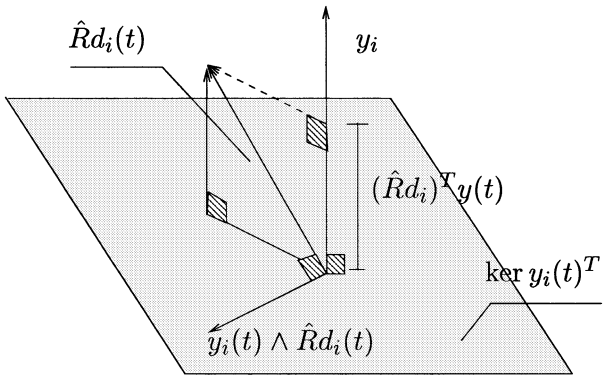

Fig. 3. Geometry underlying the observer design.

In order to describe the observer, what remains is to decide $v(t)$. We propose the following observer:

$$\boxed{\begin{cases} \dot{\hat{R}}(t) = \left[\Omega(t) - k_i \sum_{i=1}^{m} V_i(y_i(t), \hat{R}(t))\right] \hat{R}(t) \\ \hat{R}(0) = \hat{R}_0 \in SO(3) \\ V_i = -S(y_i \wedge \hat{R}d_i)(\hat{R}d_i)^T y_i \end{cases}} \tag{44}$$

where the rationale for the choice of V_i is the following, as illustrated in Fig. 3. From the implicit output equation, it follows that ideally $\hat{R}d_i \in \ker y_i^T$ so if it is not satisfied then a corrective angular velocity should be applied. The angular velocity needed for this can be directed along $y_i \wedge (\hat{R}d_i)$ which constitutes the first factor in the V_i-expression. The second factor is simply $h_i(\hat{R}, y_i)$ and is the magnitude of the correction.

B. Local Convergence

We now prove that the observer (44) is locally convergent provided that the following condition of trivial observability is satisfied together with Assumption 2.1.

Definition 4.1: For every index set $I \subseteq \{1, 2, \ldots, m\}$ let us define the information matrix

$$Q(I, t) = \sum_{i \in I} \frac{p_i(t) p_i^T(t)}{||p_i(t)||^2} \tag{45}$$

where $p_i(t)$ is the shortest vector from $p(t)$ to l_i. Furthermore, if there is a $q > 0$ such that for each t there is an index set $I_t \subseteq \{1, 2, \ldots, m\}$ such that

$$\det Q(I_t, t) \geq q \tag{46}$$

then the system is called *trivially observable*.

Note: p_i is the vector $p(t) - \xi_i(t)$ where $\xi_i(t)$ is chosen such that $d_i^T(p(t) - \xi) = 0$.

Theorem 4.1: Let us assume that Assumption 2.1 holds and let (15) be trivially observable. Then, the observer (44) is locally exponentially convergent.

Corollary 4.1: Let us assume that for each t there are three integers $i_j \in I$ such that the lines l_{i_j} intersect in a common point ξ. Assume furthermore that

$$p(t) - \xi = \sum_{j=1}^{3} \alpha_{i_j}(t) d_{i_j} \tag{47}$$

where $|\alpha_{i_j}| \geq \alpha > 0$ for some α. Then the observer is convergent.

Proof of Corollary 4.1: It is easy to see that the system is trivially observable for this simple case.

Proof of Theorem 4.2: Define the error rotation

$$\tilde{R} \doteq \hat{R}R^T \tag{48}$$

and the estimation error

$$X = I - \tilde{R}. \tag{49}$$

The error rotation dynamics are

$$\dot{\tilde{R}} = \dot{\hat{R}}R^T + \hat{R}\dot{R}^T = S(\omega)\tilde{R} - \tilde{R}S(\omega) + \sum_{i\in I} V_i\tilde{R}. \tag{50}$$

Consider the Lyapunov function candidate from [16], given as

$$\begin{aligned} V(X) =& \frac{\|X\|_F^2}{2} = \frac{\mathrm{tr}\{X^TX\}}{2} \\ =& \frac{\mathrm{tr}\{(I-\tilde{R}^T)(I-\tilde{R})\}}{2} = \mathrm{tr}\{X\}. \end{aligned} \tag{51}$$

The total derivative is

$$\begin{aligned} \dot{V} =& \frac{d}{dt}\frac{\|X\|_F^2}{2} = \mathrm{tr}\{\dot{X}\} = -\mathrm{tr}\{\dot{\tilde{R}}\} \\ =& -\mathrm{tr}\{S(\omega)\tilde{R} - \tilde{R}S(\omega) + \sum V_i\tilde{R}\} = \sum \mathrm{tr}\{-V_i\tilde{R}\} \\ =& \sum \mathrm{tr}\{v_i \wedge \tilde{R}\}. \end{aligned} \tag{52}$$

In order to show local exponential convergence, we must show that $-\dot{V} \geq c\|x\|^p$ for some constants $c > 0,\ p \geq 1$. A simple calculation shows that

$$\begin{aligned} \mathrm{tr}\{v_i \wedge \tilde{R}\} =& -k_i\mathrm{tr}\left\{\{(y_i \wedge \hat{R}d_i)y_i^T\hat{R}d_i\} \wedge \tilde{R}\right\} \\ =& -\frac{k_i}{\|\rho_i\|^2}\rho_i^TR^T\tilde{R}Rd_i \\ & \times \mathrm{tr}\{(R\rho_i \wedge \hat{R}d_i) \wedge \tilde{R}\} \\ =& -\frac{k_i}{\|\rho_i\|^2}\rho_i^TR^T\tilde{R}Rd_i \\ & \times \mathrm{tr}\left\{-(R\rho_id_i^T\hat{R}^T - \hat{R}d_i\rho_i^TR^T)\tilde{R}\right\} \\ =& -\frac{k_i}{\|\rho_i\|^2}\rho_i^TR^T\tilde{R}Rd_i\rho_i^TR^T\tilde{R}^2Rd_i \end{aligned} \tag{53}$$

so that we have

$$-\dot{V} = \sum \frac{k_i}{\|\rho_i\|^2}\rho_i^TR^T\tilde{R}Rd_i\rho_i^TR^T\tilde{R}^2Rd_i. \tag{54}$$

As a consequence of Assumption 2.1, there are $\underline{k} > 0$ and $\overline{k} < \infty$ such that

$$\underline{k} \leq \frac{k_i}{\|\rho_i\|^2} \leq \overline{k}. \tag{55}$$

Use now Rodriques' representation for rotation matrices (see [22, p. 28])

$$\tilde{R} = I - S(w)\sin\overline{v} + S^2(w)(1-\cos\overline{v}) \tag{56}$$

where $w \in \mathbb{R}^3$, $\|w\| = 1$ and where $\overline{v} \in \mathbb{R}$. By noting that

$$\|X\|^2 = 2\mathrm{tr}(I-\tilde{R}) = 4(1-\cos\overline{v}) \tag{57}$$

we have

$$\tilde{R} = I - S(w)\frac{\|X\|}{\sqrt{2}} + O(\|X\|^2) \tag{58}$$

and

$$\tilde{R}^2 = I - \sqrt{2}S(w)\|X\| + O(\|X\|^2). \tag{59}$$

O is the Ordo-symbol, that is, $O(x^k)$ is a possibly infinite polynomial where the term with the lowest degree is of degree k. Now, if we let $\overline{w} = Rw$ and use the fact that $R^TS(w)R = S(Rw)$ then we obtain

$$\dot{V}(X,t) = \left\{\sum \frac{k_i}{\|\rho_i\|^2}(\rho_i^T(t)S(\overline{w})d_i)^2\right\}\|X\|^2 + O(\|X\|^3). \tag{60}$$

Note that due to (55), the Ordo-term is bounded as a function of time. What now remains is to show that the quadratic term is bounded below in time by a strictly positive constant. We have

$$\rho_i^T(t)S(\overline{w})d_i = -\rho_i^T(t)(\overline{w} \wedge d_i) = (d_i \wedge \rho_i)^T\overline{w} \tag{61}$$

and also that

$$d_i \wedge \rho_i = d_i \wedge (d_i \wedge (p - \xi_i)). \tag{62}$$

Let us now assume that the point ξ_i is a time varying point such that $p_i(t) = p(t) - \xi_i(t)$ (see Definition 4.1). Then, it follows that

$$d_i \wedge (d_i \wedge (p - \xi_i)) = -p_i. \tag{63}$$

Thus, (60) can be written as

$$-\dot{V}(X,t) = \overline{w}^T\left\{\sum k_i\frac{p_ip_i^T}{\|p_i\|^2}\right\}\overline{w}\|X\|^2 + O(\|X\|^3) \tag{64}$$

where $\|\overline{w}\| = 1$. Finally, as a simple consequence of that $\sum(p_ip_i^T)/(\|p_i\|^2) \geq qI > 0$ we have that

$$-\dot{V}(X,t) \geq q\min\{k_i\}\|X\|^2 + O(\|X\|^3) \tag{65}$$

which concludes the proof. □

C. Numerical Estimation of the Domain Attraction

The observer (44) can, according to Theorem 4.1, be shown to be exponentially convergent. It is of interest to estimate how large initial errors can be tolerated and how this is connected to the position vector $p(t)$. The following theorems from [34] can be used to ascertain this. Let us consider the autonomous system

$$\dot{x} = f(x). \tag{66}$$

Definition 4.2: Suppose 0 is an equilibrium point of (66). The domain of attraction D is defined as

$$D = \{x_0 \in \mathbb{R}^n : x(t,x_0) \to 0 \text{ as } t \to 0\}. \tag{67}$$

Theorem 4.2: Suppose there exist a Lyapunov function V for (66). Let c be any positive constant such that the level set $L_V(c)$ defined to be the largest connected set containing the origin such that $V(x) \leq c$ is contained in the domain $S = \{x \neq 0 : V(x) > 0,\ \dot{V}(x) < 0\}$ and is bounded. Then $L_V(c)$ is a subset of $D(0)$.

From (54), we conclude that $\dot{V}$ is a function of p but that it can be taken to have unit length. We now use, once again, the Rodriques' representation (see [22, p. 28]) and write $X = S(w)\sin v - S^2(w)(1-\cos v)$ where $\|w\| = 1$ should be

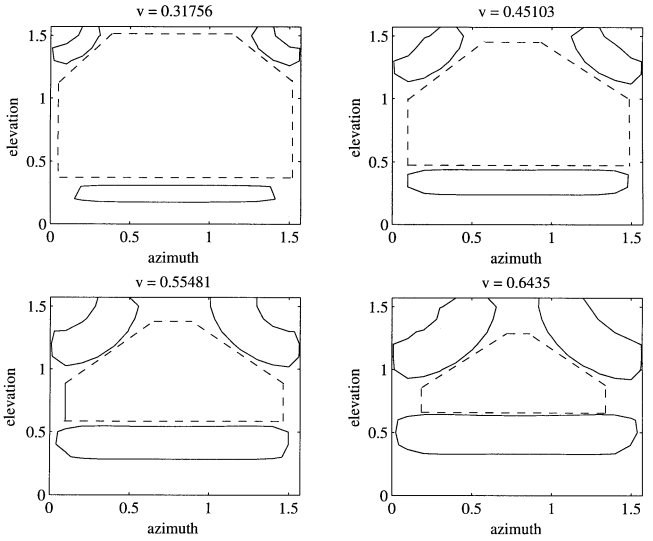

Fig. 4. Level curves $\Lambda(v)$ (solid) and estimates of $V_p(v)$ (dashed).

thought of as a direction of rotation and v as the angle of rotation. By virtue of (57), $L_V(c)$ can be written as

$$L_V(c) = \Big\{X\colon X = S(w)\sin v - S^2(w)(1-\cos v)\colon \|w\| = 1,\ v \le \arccos\Big(1-\frac{c}{4}\Big)\Big\}. \quad (68)$$

Now, we define $\mathcal{V}_p(v)$ as the set of p such that V is decreasing on the level sets $L_V(c(v))$ where $c(v) = 4(1-\cos v)$, so that

$$\mathcal{V}_p(v) \doteq \{\|p\| = 1\colon -\dot{V}(X,t) > 0 \ \forall\, X \in L_V(c(v^*)) \text{ where } v^* \le v\}. \quad (69)$$

An estimate of the domain of attraction is obtained by studying the sets $\mathcal{V}_p(v)$. To be able to obtain the estimate numerically, define

$$h(v,p) = \min_{\|w\|=1} -\dot{V}(X(v,w),p). \quad (70)$$

With a slight abuse of notation and using the spherical coordinate representation of p, we obtain the following:

$$h(v,\alpha_{az},\alpha_{el}) = \min_{\|w\|=1} -\dot{V}(v,w,p(\alpha_{az},\alpha_{el})) \quad (71)$$

which can be visualized graphically by considering the level curves $\Lambda(v)$ where $h = 0$

$$\Lambda(v) = \{\alpha_{az},\alpha_{el}\colon h(v,\alpha_{az},\alpha_{el}) = 0\}. \quad (72)$$

In Fig. 4, a sample of these level curves can be seen along with a conservative estimate of $\mathcal{V}_p(v)$. The figure illustrates how the domain of attraction gets smaller and smaller, the closer the camera is to the lines observed. This is of course natural as we then get closer and closer to violating the assumption 2.1 that $\|\eta\| \neq 0$. The minimization in (71) is carried out by exhaustive search.

V. Translation Estimation

In this section, we turn to the problem of estimating position and velocity assuming that the rotation matrix has already been estimated. We will show how the orientation estimates can be used to formulate the position estimation problem as a linear implicit output problem. To actually solve the problem, that is to find an observer, remains an open issue. We consider the case where three lines intersect in a common point ξ and we estimate the position relative to this point. The point ξ will now be used instead of the arbitrary points ξ_i on the lines. The translational parts of (12) are

$$\begin{cases} \dot{p} = v \\ \dot{v} = R^T(t)u(t) + g^N \\ y_i = \mu_i R(t)\frac{\rho_i}{\|\rho_i\|} \end{cases} \quad (73)$$

where $u(t) = R\left(a - g^N\right)$ is the accelerometer output. We also have the rotation estimate $\hat{R}$ and we propose to use this estimate as if it was a perfect estimate to transform (73) to a linear implicit output system. Let us define

$$\begin{aligned} r(t) &= \hat{R}^T u(t) + g^N \\ z_i &= \hat{R}^T y_i. \end{aligned} \quad (74)$$

Observe now that

$$\begin{aligned} r(t) &= a(t) + (I - \hat{R}^T R)\left(g^N - a\right) \\ z_i &= \mu_i \hat{R}^T R(t)\frac{\rho_i}{\|\rho_i\|}. \end{aligned} \quad (75)$$

Therefore

$$\begin{aligned} r(t) &= a + f_1(t) \\ (p-\xi_i)^T z_i &= 0 + f_2(t) \end{aligned} \quad (76)$$

where $f_1(t) \to 0$, $f_2(t) \to 0$ since the observer presented in Section IV is convergent. As these unknown error terms vanish we use the model

$$\begin{cases} \frac{d}{dt}(p(t) - \xi) = v(t) \\ \frac{d}{dt}v(t) = r(t) \\ (p(t) - \xi)^T z_i = 0 \end{cases} \quad (77)$$

as a model for (73). The quantities $r(t)$ and $z_i(t)$ are measurable quantities derived from the accelerometer and from the camera. Let us introduce the state vector $x = \left[(p-\xi)^T v^T\right]^T$ and rewrite $(p-\xi)^T z_i = \left[z_i^T 0^T\right] x$. We can write (77) as the linear implicit output problem

Problem 5.1 (Position Estimation Using Lines): We are given the system

$$\begin{cases} \dot{x} = Ax + g(t) \\ 0 = B(z)x \end{cases} \quad (78)$$

where

$$A = \begin{pmatrix} 0 & I \\ 0 & 0 \end{pmatrix} \quad g(t) = \begin{pmatrix} 0 \\ r(t) \end{pmatrix} \quad (79)$$

and

$$B(z) = \begin{pmatrix} z_1^T & 0 \\ z_2^T & 0 \\ z_3^T & 0 \end{pmatrix}. \quad (80)$$

Assume that the measurements of $g(t)$ and $z(t)$ are known in a certain interval of time, the problem is to estimate x.

The state dynamics for system (78) are quite simple (three parallel double integrators). A Luenberger type observer for critically stable linear implicit output systems has been shown to be convergent [21] by Matveev *et al.* That particular choice of observer is not applicable here since the system is unstable. How-

ever, the observability Gramian for linear implicit output systems is also given [21] and it is

$$M(t_0,T)=\int_{t_0}^{t_0+T} e^{A\tau}B(z(\tau))B(z(\tau))^T e^{A^T\tau}d\tau \qquad (81)$$

and strong observability amount to the existence of $q,T>0$ such that $M(t_0,T)\geq qI\ \forall\ t_0$. In order to understand this condition, we consider the equation $M(t_0,t_0+T)\eta=0$. Calculations similar to those in Appendix A.2 shows that if $p(t)-\xi$ is confined to a plane, then observability is lost. This is quite natural as in order to achieve depth information, the camera must move in such a way that a stereo effect is achieved.

VI. Multirate Implementation

In this paper, we have so far assumed that sensor data is continuously available. In an implementation, sensor data will of course be sampled and the observer has to be integrated in discrete time. Here, as in the rest of this paper we mainly consider rotation estimation but the implementation presented could just as well be used for position estimation. As the line detection algorithms will take a considerable amount of time, it will be necessary to make an implementation with this in mind in order to get a high-bandwidth system. We would like to point out that the observer itself is convergent to any trajectory $R(t)$ and can thus be said to be of infinite bandwidth. What will limit the achievable bandwidth here is the sampling frequency of the *inertial sensors*. It might be argued that what should limit bandwidth is the computational time associated with the *vision data* (which will induce low sampling frequency and time delayed data) but that is actually not the case. The key insight is that the IMU provides excellent high-frequency information and the vision data will only be used to compensate for the slowly varying errors induced by the integrated gyro signals. In the light of this, the low vision sampling frequency and the time delay is no longer a problem as it will implicitly only be used for slowly changing signals. Regarding the time delayed vision data, this problem can also be solved by a proper use of the IMU. Predicting the orientation ahead in time for, say a few seconds, is easy, using integrated gyros. This prediction will be associated with a small error but for the time span considered here, this error can for most practical cases be neglected. The observer can now be run with old data, estimating old orientations which are used, together with the predictions, to obtain the actual orientation.

To put all this on a firmer basis, let the IMU sampling frequency be $1/T_I$, the vision sampling frequency be $1/T_V$ and let the vision data be delayed by T_d. Let also, for simplicity, $T_V=pT_I$ and $T_d=qT_I$ where p and q are integers. The structure of the implementation is given in Fig. 5. The multirate observer is

$$\begin{aligned}\hat{R}(\tau_{k+1}) &= e^{T_I(\Omega_k+V_k)}\hat{R}(\tau_k),\ \text{for}\ \frac{k}{p}\in Z\\ \hat{R}(\tau_{k+1}) &= e^{T_I\Omega_k}\hat{R}(\tau_k),\ \text{else}\end{aligned} \qquad (82)$$

where $\tau_k=kT_I-T_d$, $\Omega_k=\Omega(\tau_k)$, $V_k=V(y_k(\tau_k),\hat{R}(\tau_k))$. Based on this estimate, $\hat{R}(t_k)$ where $t_k=kT_I$ is computed as

$$\hat{R}(t_k)=\Phi(t_k,t_k-T_d)\hat{R}(t_k-T_d) \qquad (83)$$

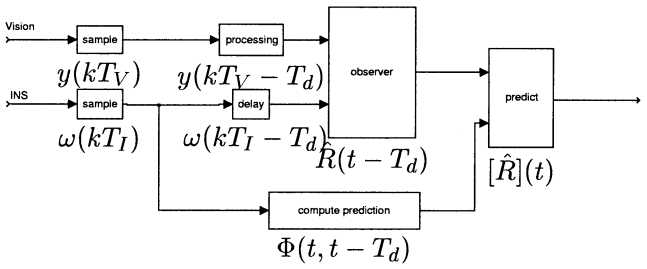

Fig. 5. Data flow in the implementation.

where $\Phi(t_k,t_k-T_d)$ is the prediction, obtained by solving

$$\begin{cases}\frac{d}{ds}\Phi(s,t-T_d)=\Omega(s)\Phi(s,t-T_d)\\ \Phi(t-T_d,t-T_d)=I\end{cases} \qquad (84)$$

for $\Phi(t_k,t_k-T_d)$. With this multirate prediction based architecture it will be possible to obtain high-bandwidth (due to the IMU), long-term stable (due to vision) orientation estimates with a slow, line-based vision algorithm. This is due to the complementary bandwidth characteristics of vision and inertial sensors.

VII. Simulations

To demonstrate the observer convergence and to show the effect of gyro offsets and of the multirate implementation we will consider some numerical simulations. We let the camera move in front of an orthogonal corner which for simplicity is taken as the origin of the inertial frame. Thus, $\xi=0$ and d_i are the three unit vectors along the coordinate axes. The unknown camera position is given by

$$p(t)=\begin{pmatrix}6\\7\\6\end{pmatrix}+\begin{pmatrix}3\sin\left(\frac{2\pi}{5}t\right)\\ 1\sin\left(\frac{2\pi}{5}t+\frac{\pi}{6}\right)\\ 3\sin\left(\frac{2\pi}{5}t+\frac{\pi}{3}\right)\end{pmatrix} \qquad (85)$$

and the orientation trajectory by the Euler angles yaw (ψ), pitch (θ), roll (ϕ)

$$\begin{aligned}\psi &= \frac{\pi}{4}\sin\left(\frac{\pi}{2}t\right)\\ \theta &= \frac{\pi}{4}\sin\left(\frac{\pi}{2}t+\frac{\pi}{4}\right)\\ \phi &= \frac{\pi}{4}\sin\left(\frac{\pi}{2}t+\frac{\pi}{2}\right)\end{aligned} \qquad (86)$$

which are translated into a rotation matrix $R(t)$. We would like to point out that *the only reason* for using yaw/pitch/roll here is pedagogic. It is hard to visualize rotations in terms of 3×3-matrices. The observer is initialized with an error corresponding to a rotation of $\pi/8$ rad around a randomly chosen direction. The gain parameters are taken as $k_i=2/\sqrt{3}$. The IMU sampling interval is $T_I=0.01$ s, the vision sampling interval is $T_v=0.1$ s as is the delay $T_d=0.1$ s. We will show two simulations with the above parameters. In Fig. 6, true and estimated ψ,θ,ϕ can be seen when there is no gyro offset and in Fig. 7 when there is a rate gyro offset of 0.1 rad/s. It can be seen from both figures that the estimated angles tend to the true. It appears however that in the case with gyro offsets, the convergence is slower. This is confirmed by plotting $\|X\|_F$ [see (51)] in Fig. 8. It is clear that the convergence is slower when the offset is present. Note however that the estimates still converge to the true values. It is also apparent that the error decreases every T_v when visual data arrives.

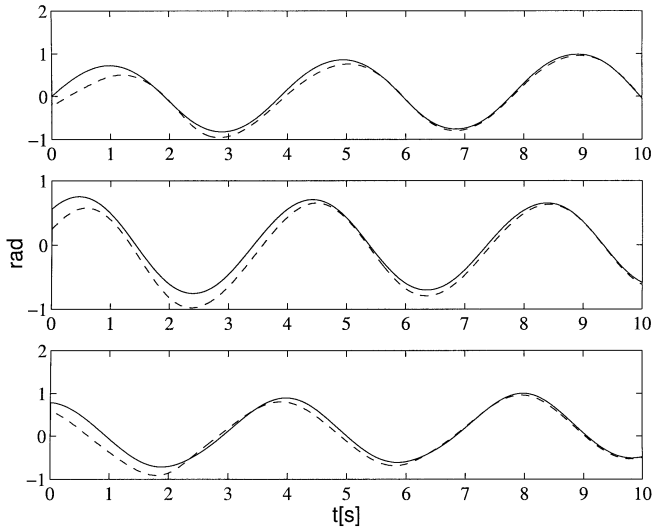

Fig. 6. Estimated and true yaw pitch and roll angle. There is no rate gyro offset used in this simulation.

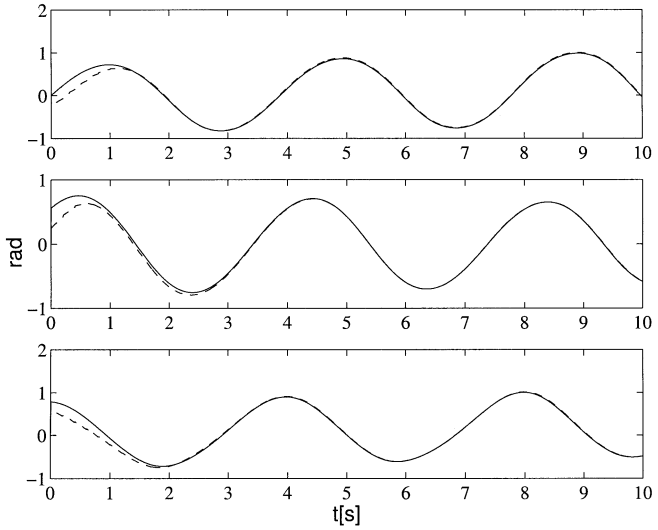

Fig. 7. Estimated and true yaw pitch and roll angle. There is a rate gyro offset of 0.1 rad/s used in this simulation.

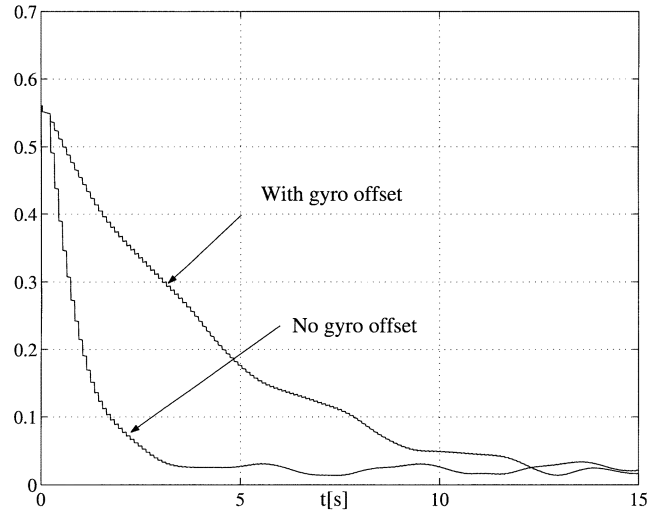

Fig. 8. $\|X(t)\|_F$ for the case with and without rate gyro offsets.

VIII. Summary and Discussion

In this paper, we have discussed a control theoretic approach to camera pose estimation. We have studied how to simultaneously use inertial sensors and computer vision and have paid particular attention to an implementation that is consistent with real-time demands. We have solved the orientation estimation problem with a locally convergent observer where the estimated states evolve on the Lie group of rotation matrices. When discussing observability we were able to show that the system itself actually is not only locally observable and it would therefore be interesting to design an observer with nonlocal convergence. This is one issue of further work. On a more general level, this problem could be studied in the context of implicit output systems evolving on manifolds. Position estimation is not at all solved in this paper and an observer for system (78) should be designed. A first attempt could be to try an implicit extended Kalman filter such as proposed by Soatto *et al.* in [30]. It might also be possible to generalize the ideas by Matveev *et al.* to unstable systems. The final test of the ideas presented is to implement the algorithms on a mobile robot. For the ideas presented here to be implemented there are a number of questions that must be answered. Consider the perhaps simplest realistic scenario, that of a mobile robot moving in a structured indoor environment. We are assuming that the direction of the lines are known. In an indoor environment, the number of orthogonal triplets corresponding to different corners can rather safely be assumed to dominate over those lines that do not correspond to these corners. With the robot starting in a standstill position, the accelerometer can be used to determine initial pitch and roll and as zero yaw anyway is arbitrary, good lines with their directions along the x, y, z-directions should be possible to find. Once this initialization phase is completed and the robot starts moving, these lines can be tracked using the IMU to help solving the line correspondence problem. The IMU should also prove helpful when handling the problem of disappearing lines. If the lines are obscured by other objects, the IMU can be used for inertial navigation while searching for new lines. Handling these problems is far from easy but the simultaneous use of vision and an IMU provides new possibilities for doing it.

Appendix A
Observability Calculations

A. The Unobservable Subgroup

The unobservable subgroup has been defined in (31) as

$$\mathcal{O}^c_{SO(3)} = \Big\{ \bar{R} \in SO(3) \colon d_i^T \bar{R} \rho_i(t) = 0 \quad \forall\, t \in [t_0, t_0 + T], \qquad i = 1, \ldots, m \Big\} \tag{87}$$

and it has been clarified that $\mathcal{O}^c_{SO(3)} = \ker U \cap SO(3)$ with U given by (29). Here, we will first compute $\ker U$ and later consider the intersection with $SO(3)$. Let us, with a slight abuse of notation, redefine U as an operator from $\mathbb{R}^9$ as opposed to $\mathbb{R}^{3\times 3}$ by

$$\begin{aligned} U \colon \mathbb{R}^9 &\to L_2(t_0, t_0 + T) \\ x &\mapsto \begin{pmatrix} d_1^T X \rho_1(t) \\ \vdots \\ d_m^T X \rho_m(t) \end{pmatrix} = C(t) x \end{aligned} \tag{88}$$

where $x = \vec{X}$ and where

$$C(t) = \begin{pmatrix} \rho_1^T \otimes d_1^T \\ \vdots \\ \rho_m^T \otimes d_m^T \end{pmatrix}. \tag{89}$$

$\otimes$ is the Kronecker product. A brief summary of some useful properties of Kronecker products and vectorized matrices is found in Appendix B. It easy to derive the adjoint operator U^* of U and the observability Gramian is given by

$$\begin{aligned} M(t_0,T) =& U^*U = \int_{t_0}^{t_0+T} C^T(\tau)C(\tau)d\tau \\ =& \int_{t_0}^{t_0+T} Q(\tau)d\tau. \end{aligned} \tag{90}$$

In the sequel, we will consider the case with $M = m$, i.e., when all the lines are linearly independent. In Remark A.3, we comment on the $m \geq M$ case.

If the observed lines are oriented along the coordinate axes, calculating $\ker M$ would be straightforward. The general case can be transformed to that special case by applying a change of basis. Recall that the lines are ordered in such a way that the first M d_i-vectors are linearly independent. Take $3-M$ vectors $f_j,\ j = M+1,\ldots,3$ such that the matrix

$$D = \begin{cases} [d_1,\ldots,d_M f_{M+1},\ldots,f_3] & \text{if } M \leq 2 \\ [d_1,\ldots,d_3] & \text{if } M = 3 \end{cases} \tag{91}$$

is full rank. Let V be a 3×3 block matrix of 3×3-blocks such that that the $l,\ n$-block is

$$V_{l,n} = D^{-T} e_n e_l^T. \tag{92}$$

V is full rank. The identity

$$D^{-1}d_i = e_i \tag{93}$$

will be used repeatedly. Now, let us define

$$P(t) = V^T C(t)^T C(t) V. \tag{94}$$

We have

$$C(t)^T C(t) = \sum_{i=1}^{M} (\rho_i(t) \otimes d_i)\left(\rho_i^T(t) \otimes d_i^T\right) \tag{95}$$

and using (132)–(135) from Appendix B, we have the (k,l)th block

$$[C(t)^T C(t)]_{k,l} = \sum_{i=1}^{M} (\rho_i)_k (\rho_i)_l d_i d_i^T. \tag{96}$$

Now, the (q,r)th block of the matrix P can be derived as

$$\begin{aligned} P_{qr} =& \sum_{i=1}^{M}\sum_{k,l=1}^{3} \left(V^T\right)_{qk} Q_{kl} V_{kr} \\ =& \sum_{i=1}^{M}\sum_{k,l=1}^{3} e_k e_q^T D^{-1} (\rho_i)_k (\rho_i)_l d_i d_i^T D^{-T} e_r e_l^T \\ =& \sum_{i=1}^{M}\sum_{k,l=1}^{3} (\rho_i)_k (\rho_i)_l e_k e_q^T D^{-1} d_i d_i^T D^{-T} e_r e_l^T \\ =& \sum_{k,l=1}^{3} e_k e_l^T \sum_{i=1}^{M} (\rho_i)_k (\rho_i)_l \delta_{i=q}\delta_{i=r} \\ =& \sum_{k,l=1}^{3} e_k e_l^T (\rho_q)_k (\rho_q)_l \delta_{r=q}\delta_{q\leq M} \\ =& \rho_q \rho_q^T \delta_{q=r}\delta_{q\leq M} \end{aligned} \tag{97}$$

and consequently we have the block diagonal matrix

$$P = \begin{pmatrix} \rho_1\rho_1^T & 0 & 0 \\ 0 & \rho_2\rho_2^T\delta_{2\leq M} & 0 \\ 0 & 0 & \rho_3\rho_3^T\delta_{3\leq M} \end{pmatrix}. \tag{98}$$

It is also clear that

$$\begin{aligned} \ker U = \Big\{ & x = Vy : y = \left[y_1^T y_2^T y_3^T\right]^T \\ & \text{where } \int_{t_0}^{t_0+T} |\rho_i^T(\tau) y_i|^2 d\tau = 0 \quad \forall\, i \leq M \Big\}. \end{aligned} \tag{99}$$

Regarding $\ker U$ it is clear that it is nonempty because $y_i = d_i$, as defined above, is a solution. As a matter of fact, it is at least three-dimensional. The best observability that can be achieved is, thus, given by the existence of $T,\ \epsilon_i > 0$ such that $\forall\, t$

$$M_i(t,T) = \int_t^{t+T} \rho_i(t)\rho_i^T(t)dt \Big|_{\{d_i\}^\perp} \geq \epsilon I, \qquad i \leq M \tag{100}$$

which is what has been defined in Definition 3.3 as strong observability. We have provided an interpretation of this condition in Remark 3.3 and in Appendix A.2 we have provided the calculations supporting this interpretation. From now on, we assume that (100) holds and proceed to the problem of computing $\ker U$. Given (100) in (99) we see that for $l \leq M$, $y_l = d_l\alpha_l$ are solutions to the integral constraint for some $\alpha_l \in \mathbb{R}$. For $l > M$, y_l is free and can be written in the basis D as $y_l = D\xi_l$ for arbitrary $\xi_l \in \mathbb{R}^3$. In a shorthand notation

$$y_l = d_l\alpha_l\delta_{l\leq M} + D\xi_l\delta_{l>M}, \qquad l = 1,2,3. \tag{101}$$

Inserting (101) in $x = Vy$ from (32), we obtain

$$\begin{aligned} (Vy)_k =& (V)_{kl} y_l \\ =& D^{-T} e_l e_k^T \left(d_l\alpha_l\delta_{l\leq M} + D\xi_l\delta_{l>M}\right) \\ =& D^{-T} e_l \left(\alpha_l\delta_{l\leq M} d_l^T D^{-T} + \delta_{l>M}\xi_l^T\right) e_k \\ =& D^{-T} A D^T e_k \end{aligned} \tag{102}$$

where

$$A = \begin{pmatrix} \alpha_1\delta_{1\leq M} e_1^T + \delta_{1>M}\xi_1^T \\ \alpha_2\delta_{2\leq M} e_2^T + \delta_{2>M}\xi_2^T \\ \alpha_3\delta_{3\leq M} e_3^T + \delta_{3>M}\xi_3^T \end{pmatrix}. \tag{103}$$

We can now formulate $\ker U$ and we do this for the original definition (29) of U, that it is as an operator from $\mathbb{R}^{3\times 3}$. As we consider $x = \vec{X}$, by reverting the vec operation for an $X \in \ker U$, we have the following:

$$X = D^{-T} A D^T [e_1 e_2 e_3] = D^{-T} A D^T \tag{104}$$

where A is given by (103). We now proceed to intersect with $SO(3)$. For $X \in SO(3)$, it must hold that

$$\begin{aligned} X^T X =& DA^T D^{-1} D^{-T} A D^T = DA^T D^{-1} D^{-T} A D^T \\ =& DA^T (D^T D)^{-1} A D^T = I \end{aligned} \tag{105}$$

which is equivalent to

$$A^T (D^T D)^{-1} A = (D^T D)^{-1} \tag{106}$$

and that

$$1 = \det X = \det A. \tag{107}$$

The (106) and (107) constrains A, and will be explored further separately for each of the M-cases.

1) Three Linearly Independent Lines: Let us assume that $M = 3$ in (103). Then, we obtain

$$A = \text{diag}([\alpha_1\alpha_2\alpha_3]) \tag{108}$$

for some $\alpha_i \in \mathbb{R}$ that will be determined. Denoting $W = (D^TD)^{-1}$ we get from (106)–(108) that

$$\alpha_i\alpha_j W_{ij} = W_{ij} \tag{109}$$

$$\alpha_1\alpha_2\alpha_3 = 1. \tag{110}$$

Since W is of full rank and $W_{ii} = 0$, we have

$$\alpha_i = \pm 1. \tag{111}$$

It now follows from (110) that, either two or none of the α_i is -1. Thus, we have at most four elements. The solution that all $\alpha_1 = 1$ corresponds to the identity solution

$$X = D^{-T}ID^T = I. \tag{112}$$

Consider now the solution candidate $\alpha_2 = \alpha_3 = -1$. We note that it exists if and only if $d_1 \perp d_2 \& d_3$. From the (109) for $i = 1, j = 2,3$ we infer that $W_{12} = W_{13} = 0$ which in light of symmetry of W and D^TD demands that $d_1^Td_2 = d_1^Td_3 = 0$. The if part of the statement is equally straightforward. As the choice of indices is arbitrary we have showed that for $M = 3$

$$\mathcal{O}^c_{SO(3)} = \left\{I, D^{-T}J_iD^T \,\forall\, i\text{: } d_i \perp d_j \& d_k \,\forall\, k, j \neq i\right\} \tag{113}$$

where J_i are diagonal matrices with 1 on position i and -1 on the two remaining positions.

2) Two Linearly Independent Lines: Let us now consider $M = 2$ in (103). It follows that

$$A = \begin{pmatrix} \alpha_1 e_1^T \\ \alpha_2 e_2^T \\ \xi_3^T \end{pmatrix} \tag{114}$$

for some $\alpha_1, \alpha_2 \in \mathbb{R}$, $\xi_3 \in \mathbb{R}^3$. Let f_3 in (91) be orthogonal to d_1 and d_2 and of unit length. Then, we have

$$D^TD = \begin{pmatrix} \Delta & 0 \\ 0 & 1 \end{pmatrix} \tag{115}$$

where

$$\Delta = [\, d_1 \quad d_2 \,]^T [\, d_1 \quad d_2 \,] \tag{116}$$

and (106) can be written as

$$\begin{pmatrix} \begin{pmatrix} \alpha_1 & 0 \\ 0 & \alpha_2 \end{pmatrix} \Delta^{-1} \begin{pmatrix} \alpha_1 & 0 \\ 0 & \alpha_2 \end{pmatrix} & \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\ (0 \quad 0) & 0 \end{pmatrix} + \xi_3\xi_3^T = \begin{pmatrix} \Delta^{-1} & 0 \\ 0 & 1 \end{pmatrix}. \tag{117}$$

From (117), it can be deduced that

$$\xi_3 = \begin{pmatrix} 0 \\ 0 \\ \pm 1 \end{pmatrix} \tag{118}$$

and by using (114) we see that $\alpha_1\alpha_2$. Using the same thinking for finding the sign combinations as in the $M = 3$-case we get that

$$\mathcal{O}^c_{SO(3)} = \left\{ I, D^{-T}J_iD^T \quad i = \left\{ \begin{array}{c} 3 \text{ if } d_1 \not\perp d_2 \\ 1,2,3 \text{ if } d_1 \perp d_2 \end{array} \right\} \right. . \tag{119}$$

3) One Line: Let $M = 1$ in (103). Then

$$A = \begin{pmatrix} \alpha_1 e_1^T \\ \xi_2^T \\ \xi_3^T \end{pmatrix}. \tag{120}$$

By letting f_2 and f_3 in (91) be such that D is orthogonal, (106) simply states that A be a rotation matrix. It is then clear from (120) that

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r_1 & r_2 \\ 0 & -r_2 & r_1 \end{pmatrix} \quad \begin{pmatrix} -1 & 0 & 0 \\ 0 & r_1 & r_2 \\ 0 & r_2 & -r_1 \end{pmatrix}$$

$$\text{where } r_1^2 + r_2^2 = 1 \tag{121}$$

are the solutions. We have

$$\mathcal{O}^c_{SO(3)} = \left\{ I, D^{-T} \begin{pmatrix} 1 & 0 & 0 \\ 0 & r_1 & r_2 \\ 0 & -r_2 & r_1 \end{pmatrix} D^T \right. D^{-T} \begin{pmatrix} -1 & 0 & 0 \\ 0 & r_1 & r_2 \\ 0 & r_2 & -r_1 \end{pmatrix} D^T \left. \forall\, r_1^2 + r_2^2 = 1 \right\}. \tag{122}$$

B. Interpretation of Strong Observability

In Definition 3, strong observability was defined and in Remark 3.3 we gave an interpretation. The calculations supporting this interpretation are given below. Consider (33) with $\epsilon = 0$. Take basis vectors f_i, g_i for $\{d_i\}^\perp$ such that $\{d_i, f_i, g_i\}$ is a right-oriented ON-base and study the kernel of

$$\begin{pmatrix} f_i^T \\ g_i^T \end{pmatrix} M_i(t_0, T) \begin{pmatrix} f_i & g_i \end{pmatrix} = \begin{pmatrix} \int_{t_0}^{t_0+T} |\rho_i^T(\tau)f_i|^2 d\tau & \int_{t_0}^{t_0+T} \rho_i^T(\tau)f_i\rho_i^T(\tau)g_i d\tau \\ \int_{t_0}^{t_0+T} \rho_i^T(\tau)f_i\rho_i^T(\tau)g_i d\tau & \int_{t_0}^{t_0+T} |\rho_i^T(\tau)g_i|^2 d\tau \end{pmatrix}. \tag{123}$$

There is a nonempty kernel if the determinant of the right hand side is zero, i.e.,

$$\int_{t_0}^{t_0+T} |\rho_i^T(\tau)f_i|^2 d\tau \int_{t_0}^{t_0+T} |\rho_i^T(\tau)g_i|^2 d\tau - \left(\int_{t_0}^{t_0+T} \rho_i^T(\tau)f_i\rho_i^T(\tau)g_i d\tau \right)^2 = 0. \tag{124}$$

According to the Cauchy–Schwartz inequality, (124) is satisfied if and only if

$$\begin{aligned} &\rho_i^T(t)g_i = 0 \quad \forall\, t \in [t_0, t_0+T] \text{ or} \\ &\rho_i^T(t)f_i = \lambda\rho^T(t)g \quad \forall\, t \in [t_0, t_0+T] \end{aligned} \tag{125}$$

for some $\lambda \in \mathbb{R}$. Using (11), the definition of ρ, and that $\{d_i, f_i, g_i\}$ is a right-oriented basis, this is rewritten as

$$\begin{aligned}(p(t)-\xi_i)^T f_i =&0 \quad \forall\, t \in [t_0, t_0+T] \text{ or}\\ (p(t)-\xi_i)^T g_i =&\lambda (p(t)-\xi_i)^T f_i\\ &\forall\, t \in [t_0, t_0+T].\end{aligned} \tag{126}$$

As f_i and g_i are fixed, this means that $(p(\tau)-\xi_i)$ is confined to a plane containing d_i. An alternative formulation is that

$$d_i^T((p(t)-\xi_i)\wedge v(t)) = 0 \quad \forall\, t \in [t_0, t_0+T]. \tag{127}$$

To prevent that this is not the case we demand that $\exists\, T > 0$ and $\epsilon' > 0$ such that $\forall\, t$

$$\int_t^{t+T} \left(d_i^T((p(t)-\xi_i)\wedge v(t))\right)^2 dt \geq \epsilon', \qquad i = 1, \ldots, m. \tag{128}$$

C. Linearly Dependent Lines

In Appendix A.1, it was assumed that $m = M$, i.e., that all lines have linearly independent direction vectors. Here, we comment on the $m > M$-case and motivate why the structural results still hold. Consider for example the case $M = 3, m = 4$ and let $d_4 = d_1$, that is two of the four lines are parallel. The matrix P from (98), for this case is given by

$$P = \begin{pmatrix} \rho_1\rho_1^T + \rho_4\rho_4^T & 0 & 0\\ 0 & \rho_2\rho_2^T & 0\\ 0 & 0 & \rho_3\rho_3^T \end{pmatrix}.$$

As $d_4 = d_1$, $\ker U$ from (99) is still the same. $\ker U$ define the unobservable subgroup so the fourth line does not have any influence on the observability structure of the problem. Regarding the conditions of strong observability the extra lines does have the influence that could be expected. The more lines that are observed the easier it is to achieve strong observability. The observability sub-Gramians (32) for $i = 1$ is for this example

$$M_1(t_0, T) = \int_{t_0}^{t_0+T} \rho_1(\tau)\rho_1^T(\tau) + \rho_4(\tau)\rho_4^T(\tau) d\tau \Big|_{\{d_1\}^\perp}. \tag{129}$$

As ρ_1 and ρ_4 typically are not parallel, it is easier to achieve full rank with the ρ_4-term than without it.

Appendix B
Vectorized Matrices and Kronecker Products

Let A be a $n \times m$ matrix and B a $k \times l$ matrix. The *Kronecker product* of A and B is the $nk \times ml$ matrix $A \otimes B$ given by

$$A \otimes B = [A_{ij}B]. \tag{130}$$

An example is that if $A = \begin{pmatrix} a_{11} & a_{12}\\ a_{21} & a_{22}\end{pmatrix}$, then

$$A \otimes B = \begin{pmatrix} a_{11}B & a_{12}B\\ a_{21}B & a_{22}B\end{pmatrix}.$$

The vec operation is defined by stacking matrix columns in a column vector according to

$$\mathrm{vec}A = \begin{pmatrix} A_{\cdot,1}\\ \vdots\\ A_{\cdot,m}\end{pmatrix} \tag{131}$$

where $A_{\cdot,j}$ is A's j:th column. For matrices of compatible dimensions, the following rules apply:

$$\mathrm{vec}ABC = (C^T \otimes A)\mathrm{vec}B \tag{132}$$

$$(A \otimes B)(C \otimes D) = (AC) \otimes (BD) \tag{133}$$

$$(A \otimes B)^{-1} = A^{-1} \otimes B^{-1} \tag{134}$$

$$(A \otimes B)^T = A^T \otimes B^T. \tag{135}$$

References

[1] N. Andreff, B. Espiau, and R. Horaud, "Visual servoing from lines," in *Proc. 2000 IEEE Int. Conf. Robotics Automation*, San Francisco, CA, Apr. 2000, pp. 2070–2075.

[2] A. J. Baerveldt and R. Klang, "A low-cost and low-weight attitude estimation system for an autonomous helicopter," in *Proc. IEEE Int. Conf. Intelligent Engineering Systems*, 1997, pp. 391–395.

[3] J. Balaram, "Kinematic observers for articulated rovers," in *Proc. 2000 IEEE Int. Conf. Robotics Automation*, San Francisco, Apr. 2000, pp. 2597–2604.

[4] B. Barshan and H. F. Durrant-Whyte, "Inertial navigation systems for mobile robots," *IEEE Trans. Robotics Automation*, vol. 11, pp. 328–342, June 1995.

[5] W. M. Boothby, *An Introduction to Differentiable Manifolds and Riemannian Geometry*. New York: Academic, 1975.

[6] F. Bullo and R. M. Murray, "Tracking for fully actuated mechanical systems: A geometric framework," *Automatica*, vol. 35, no. 1, pp. 17–34, 1999.

[7] S. Christy and R. Horaud, "Fast and reliable object pose estimation from line correspondences," in *Proc. 7th Int. Conf. Computer Analysis Images Patterns*, 1997, pp. 432–9.

[8] ——, "Iterative pose computation from line correspondences," *Comput. Vision Image Understanding*, vol. 73, no. 1, pp. 137–144, 1999.

[9] F. Dornaika and C. Garcia, "Pose estimation using point and line correspondences," *Real-Time Imag.*, vol. 5, pp. 215–230, 1999.

[10] E. Foxlin, "Inertial head-tracker sensor fusion by a complementary separete-bias Kalman filter," in *Proc. IEEE 1996 Virtual Reality Ann. Int. Symp.*, 1996, pp. 185–94.

[11] E. Foxlin, M. Harrington, and Y. Altshuler, "Miniature 6-DOF inertial system for tracking HMD," *Proc. SPIE*, vol. 3362, pp. 214–228, 1998.

[12] J. B. Fraleigh, *A First Course in Abstract Algebra*. Reading, MA: Addison-Wesley, 1999.

[13] M. Greene, "A solid state attitude heading reference system for general aviation," in *Proc. 1996 IEEE Conf. Emerging Technologies Factory Automation*, vol. 2, 1996, pp. 413–17.

[14] T. S. Huang and A. B. Netravali, "Motion and structure from feature correspondences: A review," *Proc. IEEE*, vol. 82, pp. 252–268, Feb. 1994.

[15] M. Jankovic and B. K. Ghosh, "Visually guided ranging from observations of points, lines and curves via an identifier based nonlinear observer," *Syst. Control Lett.*, vol. 25, pp. 63–73, 1995.

[16] D. E. Koditschek, "The application of total energy as a Lyapunov function for mechanical control systems," in *Dynamics and Control of Multibody Systems*, J. E. Marsden, P. S. Krishnaprasad, J. C. Simo, and editors, Eds. Providence, RI: AMS, 1989, vol. 97, pp. 131–157.

[17] R. Kurazume and S. Hirose, "Development of image stabilization system for remote operation of walking robots," in *Proc. 2000 IEEE Int. Conf. Robotics Automation*, San Francisco, CA, Apr. 2000, pp. 1856–1861.

[18] E. Lefferts, F. Markley, and M. Shuster, "Kalman filtering for spacecraft attitude estimation," *J. Guid., Control, Dyna.*, vol. 5, no. 5, pp. 417–429, Sept.–Oct. 1982.

[19] J. Lobo and J. Dias, "Integration of inertial information with vision," in *Proc. 24th Ann. Conf. IEEE Industrial Electronics Soc.*, vol. 3, 1998, pp. 1263–1267.

[20] A. M. Madni, D. Bapna, P. Levin, and E. Krotkov, "Solid-state six degree of freedom, motion sensor for field robotic applications," in *Proc. 1998 IEEE/RSJ Int. Conf. Intelligent Robots System*, vol. 3, 1998, pp. 1389–9.

[21] A. Matveev, X. Hu, R. Frezza, and H. Rehbinder, "Observers for systems with implicit output," *IEEE Trans. Automat. Contr.*, vol. 45, pp. 168–173, Jan. 2000.
[22] R. M. Murray, Z. Li, and S. S. Sastry, *A Mathematical Introduction to Robotic Manipulation*. Boca Raton, FL: CRC Press, 1994.
[23] J. Olensis, "A critique of structure-from-motion algorithms," *Comput. Vision Image Understand.*, vol. 80, pp. 172–214, 2000.
[24] H. Rehbinder and B. K. Ghosh, "Multi-rate fusion of visual and inertial data," presented at the Proc. IEEE Conf. Multi-Sensor Fusion Integration Intelligent Systems, Baden-Baden, Germany, 2001.
[25] H. Rehbinder and X. Hu, "Nonlinear pitch and roll estimation for walking robots," in *Proc. 2000 IEEE Int. Conf. Robotics Automation*, vol. 3, San Francisco, CA, 2000, pp. 2617–2622.
[26] ——, "Nonlinear state estimation for rigid body motion with low-pass sensors," *Syst. Control Lett.*, vol. 40, no. 3, pp. 183–190, 2000.
[27] ——, "Drift-free attitude estimation for accelerated rigid bodies," in *Proc. 2001 IEEE Int. Conf. Robotics Automation*, vol. 4, Seoul, Korea, 2001, pp. 4244–4249.
[28] T. Sakaguchi, K. Tsutomu, H. Katayose, K. Sato, and S. Inokuchi, "Human motion capture by integrating gyroscopes and accelerometers," in *Proc. 1996 IEEE/SICE/RSJ Int. Conf. Multisensor Fusion Integration Intelligent Systems*, 1996, pp. 470–475.
[29] R. Smith, A. Frost, and P. Probert, "Gyroscopic data fusion via a quartenion-based complementary filter," *Proc. SPIE*, vol. 3067, pp. 148–59, 1997.
[30] S. Soatto, R. Frezza, and P. Perona, "Motion estimation via dynamic vision," *IEEE Trans. Automat. Contr.*, vol. 41, pp. 393–413, Mar. 1996.
[31] M. Spetsakis, "Structure from motion using line correspondences," *Int. J. Comput. Vision*, vol. 4, pp. 171–183, 1990.
[32] ——, "A linear algorithm for point and line-based structure from motion," *Comput. Vision, Graph., Image Processing*, vol. 56, pp. 230–241, 1992.
[33] J. Vaganay, M. Aldon, and A. Fournier, "Mobile robot attitude estimation by fusion of inertial data," in *Proc. 1993 IEEE Int. Conf. Robotics Automation*, vol. 1, 1993, pp. 277–82.
[34] M. Vidyasagar, *Nonlinear Systems Analysis*. Upper Saddle River, NJ: Prentice-Hall, 1993.

Henrik Rehbinder (S'00–M'01) received the M.Sc. degree in engineering physics from the Royal Institute of Technology (KTH), Stockholm, Sweden, and the Ph.D. degree from the Division of Optimization and Systems Theory, KTH, where he was affiliated with the *Centre for Autonomous Systems*, in 1996 and 2001, respectively.

His research interests are of various applications of control theory to problems in sensor fusion for mobile robots and real-time systems. He is currently Director of Research at RaySearch Laboratories AB, Stockholm, Sweden, and his main responsibility is optimization and control design for radiation therapy for cancer.

Bijoy K. Ghosh (S'79–M'79–SM'90–F'00) received the B.Tech. and M.Tech. degrees in electrical and electronics engineering from BITS, Pilani, India, and the Indian Institute of Technology, Kanpur, India, in 1977 and 1979, respectively, and the Ph.D. degree in engineering from the Decision and Control Group of the Division of Applied Sciences, Harvard University, Cambridge, MA, in 1983.

Since 1983, he has been a Faculty Member in the Systems Science and Mathematics department at Washington University, St. Louis, MO, where he is currently a Professor and directs the Center for BioCybernetics and Intelligent Systems. His research interests are in multivariable control theory, machine vision, and biosystems and control.