USC-SIPI REPORT #118

# 3-D Motion Estimation Using a Sequence of Noisy Stereo Images Part I: Models and Motion Estimation

by

G. S. Young and R. Chellappa

Signal and Image Processing Institute
UNIVERSITY OF SOUTHERN CALIFORNIA
Department of Electrical Engineering-Systems
Powell Hall of Engineering
University Park/MC-0272
Los Angeles, CA 90089 U.S.A.

Supported in part by NSF Grant No. MIP-84-51010 and matching funds from TRW.

# 3-D Motion Estimation Using a Sequence of Noisy Stereo Images

## Part I : Models and Motion Estimation


### Abstract

We discuss a kinematic model based approach for the estimation of 3-D motion parameters from a sequence of noisy stereo images. The approach is based on representing the constant acceleration translational motion, and constant angular velocity or constant precession rotational motion in the form of a bilinear state space model using standard rectilinear states for translation and quaternions for rotation. Closed form solutions of the state transition equations are obtained to propagate the quaternions in both constant angular velocity and constant precession models. The measurements are noisy perturbations of 3-D feature points represented in an inertial coordinate system. It is assumed that the 3-D feature points are extracted from the stereo images and matched over the frames. Owing to the nonlinearity in the state model, nonlinear filters are designed for the estimation of motion parameters. A performance bound for the motion parameters is calculated. Simulation results are included. In a companion report [19], uniqueness of motion parameter estimates is established.

# 1 Introduction

A significant amount of work has been done in computer vision literature [1] for the estimation of 3-D motion and/or structure parameters from a monocular sequence of images. Much of the earlier work involved using two or three frames of noisy images. The nonrobustness and lack of numerical stability of these two or three view monocular image based algorithms have been noted by several researchers [1,2,3,4]. Some of these problems were partly alleviated using a long sequence of noisy image frames [5,6,7,8].

In this report we develop robust kinematic model based algorithms for 3-D motion parameter estimation from an arbitrarily long sequence of noisy stereo images. Only recently [3,8] this problem has received some attention in the literature. The use of a large number of stereo images improves on the robustness and numerical accuracy of the motion parameters. Our approach uses models for the motion of the rigid object, the time evolution of this motion and the observation of the object, as illustrated in Figure 1. The modeled motion has constant acceleration in the translational components, and constant angular velocity or constant precession [8] in the rotational components. As shown in Figure 2, the constant precession motion is modeled as the instantaneous angular velocity of the object rotating w.r.t. a spatially fixed axis with a constant angular velocity. Constant precession motion is an example of higher order rotation than the motion of constant angular velocity.

When standard rectilinear states are adopted for translational motion, our motion model results in a quadratic time evolution. The constant angular velocity motion or the constant precession motion is propagated using quaternion representation. In the constant angular velocity model, the combined translational and rotational motion is represented in the form of a state space model with nineteen states corresponding to nine translational motion states, seven rotational states, and three states which describe the position of the rotation center. Three more states for precession are needed in the constant precession case. Due to the propagation of rotation, the state space model tends to be bilinear. For the purposes of this report we assume that the structure is known.

The measurements are modeled as noisy perturbations of 3-D coordinates of a physical point on the rigid object represented in an inertial coordinate frame. The 3-D coordinates are assumed

to be obtained from the stereo pair. Although the ultimate goal is to characterize the noise distribution using the error analysis of a typical stereo matching algorithm [9,10], for the purposes of this report we assume the noise to be Gaussian. It is also assumed that matching of 3-D feature points and establishing correspondence between the matching points have been accomplished [2].

The 3-D motion estimation problem is thus reduced to one of identifying the states of a dynamical system with bilinear state model and trilinear measurement model. Due to the nonlinearities present, an Extended Kalman Filter (EKF) [11,12] is used for motion parameter identification. Our simulation results show that for heavy noise situations, i.e. when the noise level is more than 14% of the object size, modifications of EKF such as the iterated extended Kalman filters [11,12] may be required. Once the parameters or states in the motion model are estimated, prediction and interpolation of missing measurements can be accomplished. An objective evaluation of the robustness of our models w.r.t. additive Gaussian noise is obtained by generalizing the Cramér-Rao inequality for the time-varying motion parameters. The closeness between this objective performance bound and the sample variances from simulations shows that the parameter estimates are reasonably efficient.

One of the major questions to be answered in any motion estimation problem is the uniqueness or lack thereof of motion parameters. A proof for the uniqueness of motion parameters in our model is given in a companion report [19].

# 2 Models

The motion states/parameters are chosen to incorporate the general 3-D motion of constant acceleration, and constant angular velocity or constant precession. As will be seen, the states chosen will be different between the models for the unknown structure case and the models which incorporate structure information.

## 2.1 Observation Noise Model

The observation noise model, which describes the way how noise comes into the measurements, will be introduced.

It is well known that the 3-D coordinates of feature points can be decided from binocular

images, e.g. by stereo triangulation. Therefore the calculated 3-D coordinates of feature points from the binocular images will be taken as the measurements in our models. Define

- inertial coordinate system $I \equiv$ arbitrary 3-D world coordinate system.

Thus either $I_1$ or $I_2$ defined in Figure 1 can be chosen as the inertial coordinate system $I$. Consider a feature point $P$. Let

- $(x, y, z)^T \equiv$ the noisy 3-D coordinates of $P$ in $I$ calculated from binocular images, which are noisy, and
- $(x_{tru}, y_{tru}, z_{tru})^T \equiv$ the true(noise-free) 3-D coordinates of $P$ in $I$.

The measurements are taken at discrete time $t_i$. With the measurement time $t_i$ expressed explicitly, the *observation noise model* is

$$\begin{pmatrix} x(t_i) \\ y(t_i) \\ z(t_i) \end{pmatrix} = \begin{pmatrix} x_{tru}(t_i) \\ y_{tru}(t_i) \\ z_{tru}(t_i) \end{pmatrix} + \begin{pmatrix} n_x(t_i) \\ n_y(t_i) \\ n_z(t_i) \end{pmatrix} \tag{1}$$

where $n_x, n_y$,and $n_z$ are the noise components, assumed to be Gaussian processes with mean zero.

*The observation noise model allows the possibility to exploit overestimation in an optimal way.* If optimal treatments of noise and overestimation are desired, a model for the way how noise comes into the system must be established. Under this model, optimal motion estimator can be designed. Also the performance of the estimator can be analyzed in terms of the noise model.

## 2.2 Object and Motion Model

The object and motion model, which describes the structure of the rigid body and the motion of constant acceleration and constant angular velocity or precession, will be introduced. The motion states/parameters will be chosen from the physical quantities in this model.

As shown in Figure 1, under the known structure assumption, there exists a structure coordinate system $S$ which is fixed on the rigid body, and in which the 3-D coordinates of every feature point $P$ are known. Let

- $(x_s, y_s, z_s)^T \equiv$ the 3-D coordinates of $P$ in $S$.

*S facilitates the incorporation of structure information.* $S$ can also be used in the model of unknown structure.

We assume that the underlying motion can be decomposed into a translation of a point $O_b$ which is fixed on the object with a constant acceleration and a rotation w.r.t. $O_b$ with a constant angular velocity or precession. Let

- $\underline{\omega}(t) = (\omega_x(t), \omega_y(t), \omega_z(t))^T \equiv$ the angular velocity of the rigid body w.r.t. $O_b$ with components represented in $I$
- $L \equiv$ the rotation axis which is fixed on the rigid body and on which $\underline{\omega}(t)$ lies.
- object coordinate system $B \equiv$ the 3-D coordinate system fixed on the rigid body with the origin $O_b$, and with its coordinate axes oriented in the same direction as the coordinate axes of $S$.
- $(x_b, y_b, z_b)^T \equiv$ the 3-D coordinates of $P$ in $B$.
- $\underline{d} = (d_x, d_y, d_z)^T \equiv$ the 3-D coordinates of $O_b$ in $S$.
- $(x_R, y_R, z_R)^T \equiv$ the 3-D coordinates of $O_b$ in $I$.

It is clear that the transformation between $B$ and $S$ is

$$\begin{pmatrix} x_b(t_i) \\ y_b(t_i) \\ z_b(t_i) \end{pmatrix} = \begin{pmatrix} x_s(t_i) \\ y_s(t_i) \\ z_s(t_i) \end{pmatrix} - \begin{pmatrix} d_x \\ d_y \\ d_z \end{pmatrix} \tag{2}$$

This is due to the fact that $B$ and $S$ have mutually parallel axes. Let $R$ be the $3 \times 3$ alignment matrix that aligns the object coordinate system $B$ with the inertial coordinate system $I$. Thus $R$ is a rotation matrix.

Then the transformation between $I$ and $B$ is

$$\begin{pmatrix} x_{tru}(t_i) \\ y_{tru}(t_i) \\ z_{tru}(t_i) \end{pmatrix} = \begin{pmatrix} x_R(t_i) \\ y_R(t_i) \\ z_R(t_i) \end{pmatrix} + R(t_i) \begin{pmatrix} x_b(t_i) \\ y_b(t_i) \\ z_b(t_i) \end{pmatrix} \tag{3}$$

Substitution of (2) into (3) yields the *object motion model*

$$\begin{pmatrix} x_{tru}(t_i) \\ y_{tru}(t_i) \\ z_{tru}(t_i) \end{pmatrix} = \begin{pmatrix} x_R(t_i) \\ y_R(t_i) \\ z_R(t_i) \end{pmatrix} + R(t_i) \left[ \begin{pmatrix} x_s(t_i) \\ y_s(t_i) \\ z_s(t_i) \end{pmatrix} - \begin{pmatrix} d_x \\ d_y \\ d_z \end{pmatrix} \right] \tag{4}$$

where $d_x, d_y$, *and* $d_z$ *are chosen to be the three motion parameters which describe the center of rotation or precession.*

### 2.2.1 Translational Motion States

The motion parameters which describe the translational and rotational parts of the underlying motion will be defined. Closed form expressions for the time dependence of these parameters will also be given. These closed form expressions will be used in the Kalman filter to reduce the computational load.

Let

- $\underline{r} = (x_R, y_R, z_R)^T \equiv$ the position vector of $O_b$ w.r.t. $I$ and represented in $I$.
- $\underline{v} = (v_x, v_y, v_z)^T \equiv$ the velocity vector of $O_b$ w.r.t. $I$ and represented in $I$
- $\underline{a} = (a_x, a_y, a_z)^T \equiv$ the acceleration vector of $O_b$ w.r.t. $I$ and represented in $I$

Under the assumption of constant acceleration, we have

$$\begin{cases} \dot{\underline{r}}(t) &= \underline{v}(t) \\ \dot{\underline{v}}(t) &= \underline{a}(t) \\ \dot{\underline{a}}(t) &= 0 \end{cases} \tag{5}$$

By direct integration, the closed form solution of this equation is obtained as

$$\begin{cases} \underline{r}(t_2) &= \underline{r}(t_1) + (t_2 - t_1)\underline{v}(t_1) + (t_2 - t_1)^2\underline{a}(t_1)/2 \\ \underline{v}(t_2) &= \underline{v}(t_1) + (t_2 - t_1)\underline{a}(t_1) \\ \underline{a}(t_2) &= \underline{a}(t_1) \end{cases} \tag{6}$$

where $x_R(t), y_R(t), z_R(t), v_x(t), v_y(t), v_z(t), a_x, a_y$, *and* $a_z$ *are chosen to be the 9 motion states to describe the translational part of the underlying motion.*

### 2.2.2 Rotational Motion States

There are about eight commonly used alternatives to represent rotation [14]. The quaternion [14,15,16,17] is one of these. The quaternion

$$\underline{q}(t) = (q_1(t), q_2(t), q_3(t), q_4(t))^T$$

is defined as the solution of the differential equation

$$\dot{\underline{q}}(t) = \Omega\left[\underline{\omega}(t)\right] \underline{q}(t) \tag{7}$$

where

$$\Omega\left[\underline{\omega}(t)\right] = \frac{1}{2}\begin{pmatrix} 0 & -\omega_z & \omega_y & -\omega_x \\ \omega_z & 0 & -\omega_x & -\omega_y \\ -\omega_y & \omega_x & 0 & -\omega_z \\ \omega_x & \omega_y & \omega_z & 0 \end{pmatrix}_t \tag{8}$$

The relation between $\underline{q}$ and the "standard" rotation parameters $(n_1, n_2, n_3, \theta)^T$ is

$$\underline{q} = \left(n_1 \sin\frac{-\theta}{2}, n_2 \sin\frac{-\theta}{2}, n_3 \sin\frac{-\theta}{2}, \cos\frac{-\theta}{2}\right)^T$$

where $(n_1, n_2, n_3)^T$ is the unit directional vector of the rotation axis of the alignment matrix $R$ and $\theta$ is the rotation angle of $R$.

The standard rotation parameters are not used as it is not easy to obtain a closed form expression for time dependence.

After suppressing the time dependence, we get the alignment matrix $R$ expressed in $\underline{q}$ as

$$R = \begin{pmatrix} q_1^2 - q_2^2 - q_3^2 + q_4^2 & 2(q_1q_2 + q_3q_4) & 2(q_1q_3 - q_2q_4) \\ 2(q_1q_2 - q_3q_4) & -q_1^2 + q_2^2 - q_3^2 + q_4^2 & 2(q_2q_3 + q_1q_4) \\ 2(q_1q_3 + q_2q_4) & 2(q_2q_3 - q_1q_4) & -q_1^2 - q_2^2 + q_3^2 + q_4^2 \end{pmatrix} \tag{9}$$

**Case I. Constant Angular Velocity ($\underline{\omega}(t)$ = constant)**

In this case, the direction of the rotation axis $L$ does not change with time. If $\underline{\omega}$ is constant, (7) is a system of first order linear time invariant differential equation with a closed form solution [13],

$$\underline{q}(t_2) = \exp\left[\Omega \cdot (t_2 - t_1)\right] \underline{q}(t_1)$$

The above can be simplified to [6],[7]

$$\underline{q}(t_2) = \left[I_4 \cos\frac{|\underline{\omega}|(t_2 - t_1)}{2} + \frac{2}{|\underline{\omega}|}\Omega \sin\frac{|\underline{\omega}|(t_2 - t_1)}{2}\right] \underline{q}(t_1) \tag{10}$$

where $I_4$ is the $4 \times 4$ identity matrix.

$q_1(t), q_2(t), q_3(t), q_4(t), \omega_x, \omega_y$, *and* $\omega_z$ *are chosen to be the seven rotational states for the constant angular velocity model.*

**Case II. Constant Precession**

In this case, the direction of $L$ is no longer fixed. As shown in Figure 2, $L$ and the angular velocity vector $\underline{\omega}(t)$ are assumed to be rotating with a constant angular velocity[8]. Define

- precession vector $\underline{p}= (p_x, p_y, p_z)^T \equiv$ the angular velocity of $\underline{\omega}(t)$ (or $L$) w.r.t. $O_b$ and with components represented in $I$. Under the assumption of constant precession, $\underline{p}$ is a constant vector.
- precession axis $L_p \equiv$ the axis on which $O_b$ and the precession vector $\underline{p}$ lie. Under the assumption of constant precession, the direction of $L_p$ does not change with time.

Since the angular velocity vector $\underline{\omega}(t)$ rotates with the constant velocity $\underline{p}$, we have

$$\dot{\underline{\omega}}(t) = P(\underline{p})\underline{\omega}(t) \tag{11}$$

where

$$P(\underline{p}) = \begin{pmatrix} 0 & -p_z & p_y \\ p_z & 0 & -p_x \\ -p_y & p_x & 0 \end{pmatrix} \tag{12}$$

Using the standard approach in Linear System Theory [13], the closed form solution of (11) is obtained as

$$\begin{aligned} & \underline{\omega}(t_2) \\ = \; & \exp\left[P \cdot (t_2 - t_1)\right]\underline{\omega}(t_1) \\ = \; & \left[I_3 + \frac{\sin\left[|\underline{p}|(t_2 - t_1)\right]}{|\underline{p}|}P + \frac{1 - \cos\left[|\underline{p}|(t_2 - t_1)\right]}{|\underline{p}|^2}P^2\right]\underline{\omega}(t_1) \\ = \; & \begin{pmatrix} n_1^2 + (1 - n_1^2)\cos\theta & n_1 n_2(1 - \cos\theta) - n_3\sin\theta & n_1 n_3(1 - \cos\theta) + n_2\sin\theta \\ n_1 n_2(1 - \cos\theta) + n_3\sin\theta & n_2^2 + (1 - n_2^2)\cos\theta & n_2 n_3(1 - \cos\theta) - n_1\sin\theta \\ n_1 n_3(1 - \cos\theta) - n_2\sin\theta & n_2 n_3(1 - \cos\theta) + n_1\sin\theta & n_3^2 + (1 - n_3^2)\cos\theta \end{pmatrix}\underline{\omega}(t_1) \\ \equiv \; & \Phi_2(\underline{p}; t_2 - t_1)\underline{\omega}(t_1) \end{aligned} \tag{13}$$

where

$$(n_1, n_2, n_3)^T = \left(\frac{p_x}{|\underline{p}|}, \frac{p_y}{|\underline{p}|}, \frac{p_z}{|\underline{p}|}\right)^T$$

and

$$\theta = |\underline{p}|(t_2 - t_1)$$

It is shown, in Appendix I, that the closed form solution of $\underline{q}$ is

$$\underline{q}(t_2) = \Phi_1\left(\underline{\omega}(t_1), \underline{p}; t_2 - t_1\right)\underline{q}(t_1) \tag{14}$$

where

$$\Phi_1\left(\underline{\omega}(t_1), \underline{p}; t_2 - t_1\right)$$

$$= \begin{pmatrix} \phi_4 & \phi_3 & -\phi_2 & \phi_1 \\ -\phi_3 & \phi_4 & \phi_1 & \phi_2 \\ \phi_2 & -\phi_1 & \phi_4 & \phi_3 \\ -\phi_1 & -\phi_2 & -\phi_3 & \phi_4 \end{pmatrix} \left[ I_4 \cos\frac{|\underline{\lambda}|(t_2 - t_1)}{2} + \frac{2}{|\underline{\lambda}|}\Omega(\underline{\lambda}) \sin\frac{|\underline{\lambda}|(t_2 - t_1)}{2} \right]$$

in which

$$\begin{pmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \\ \phi_4 \end{pmatrix} = \begin{pmatrix} \frac{p_x}{|\underline{p}|} \sin\frac{-|\underline{p}|(t_2 - t_1)}{2} \\ \frac{p_y}{|\underline{p}|} \sin\frac{-|\underline{p}|(t_2 - t_1)}{2} \\ \frac{p_z}{|\underline{p}|} \sin\frac{-|\underline{p}|(t_2 - t_1)}{2} \\ \cos\frac{-|\underline{p}|(t_2 - t_1)}{2} \end{pmatrix}$$

and

$$\Omega(\underline{\lambda}) = \frac{1}{2}\begin{pmatrix} 0 & -\lambda_z & \lambda_y & -\lambda_x \\ \lambda_z & 0 & -\lambda_x & -\lambda_y \\ -\lambda_y & \lambda_x & 0 & -\lambda_z \\ \lambda_x & \lambda_y & \lambda_z & 0 \end{pmatrix} \tag{15}$$

$$\underline{\lambda} = (\lambda_x, \lambda_y, \lambda_z)^T = \underline{\omega}(t_1) - \underline{p}$$

$q_1(t), q_2(t), q_3(t), q_4(t), \omega_x(t), \omega_y(t), \omega_z(t), \underline{p}_x, \underline{p}_y$, *and* $\underline{p}_z$ *are chosen to be the ten rotational states for the constant precession model.*

*Remarks.* In the model of constant angular velocity, there is one extra degree of freedom involved in $d_x, d_y, d_z, x_R, y_R$, and $z_R$, because $O_b$ is not uniquely defined. In fact, if the rotation axis $L$ does not change direction with time, any point on $L$ can be chosen as the rotation center $O_b$ to decompose the underlying motion into a translational constant acceleration component and

a rotational constant angular velocity component as long as $O_b$ is fixed on the rigid body. This extra degree of freedom will be retained in order to avoid singularity. However, in the model for constant precession, there is no such extra degree of freedom involved if $L$ and $L_p$ are not coincident.

## 3 Recursive Filter Formulation

Recursive filters [5,6,11,12] will be used to estimate the motion parameters from noisy measurements. Both EKF and iterated extended Kalman filter (IEKF) are recursive estimators of the parameters in a nonlinear system. The initial guess of the parameters is recursively improved by EKF/IEKF as additional measurements are available. This feature is desired in tracking of moving objects, since motion parameters can be updated using newly obtained measurements without waiting for the collection of all data. The parameters to be estimated are allowed to be time-varying and denoted by a $n \times 1$ column vector $\underline{x}(t)$, called state vector. The time-varying behavior of the states is required to satisfy the *plant model*

$$\dot{\underline{x}}(t) = \underline{f}\left[\underline{x}(t), \underline{u}(t), t\right] + G(t)\underline{w}(t) \tag{16}$$

where $\underline{u}(t)$ is the deterministic input functions and $\underline{w}(t)$ a zero-mean white Gaussian noise process with covariance kernel

$$E\left\{\underline{w}(t)\underline{w}^T(t+\tau)\right\} = Q(t)\delta(\tau)$$

The measurements $\underline{z}$ at time $t_i$ are required to depend on the states according to the *measurement model*

$$\underline{z}(t_i) = \underline{h}\left[\underline{x}(t_i), t_i\right] + \underline{v}(t_i) \tag{17}$$

where $\underline{v}$ is a white Gaussian noise sequence of mean zero and covariance kernel

$$E\left\{\underline{v}(t_i)\underline{v}^T(t_j)\right\} = \begin{cases} R_v(t_i) & , t_i = t_j \\ 0 & , t_i \neq t_j \end{cases} \tag{18}$$

Once the plant and measurement models are specified, EKF updates the estimates $\hat{\underline{x}}$ for the true state vector $\underline{x}$ by a series of equations introduced in the following [12]. Since the nonlinearity of

$\underline{f}$ and $\underline{h}$ is linearized by the first order Taylor series expansions, the following matrix definitions are needed,

$$F \equiv \frac{\partial \underline{f}}{\partial \underline{x}} \tag{19}$$

$$H \equiv \frac{\partial \underline{h}}{\partial \underline{x}} \tag{20}$$

The estimate $\underline{\hat{x}}(t_i^+)$ for $\underline{x}(t_i^+)$ immediately after the measurements at $t_i$ is obtained by

$$\underline{\hat{x}}(t_i^+) = \underline{\hat{x}}(t_i^-) + K(t_i)\left\{\underline{z}(t_i) - \underline{h}\left[\underline{\hat{x}}(t_i^-), t_i\right]\right\} \tag{21}$$

where the gain matrix is

$$K(t_i) = P(t_i^-)H^T[t_i; \underline{\hat{x}}(t_i^-)]\left\{H[t_i; \underline{\hat{x}}(t_i^-)]P(t_i^-)H^T[t_i; \underline{\hat{x}}(t_i^-)] + R_v(t_i)\right\}^{-1} \tag{22}$$

The approximate covariance matrix is updated by

$$P(t_i^+) = P(t_i^-) - K(t_i)H^T[t_i; \underline{\hat{x}}(t_i^-)]P(t_i^-) \tag{23}$$

Then the estimate and the approximate covariance matrix are propagated from $t_i^+$, $\underline{\hat{x}}(t_i^+)$ and $P(t_i^+)$, forward to the next sample time $t_{i+1}^-$, $\underline{\hat{x}}(t_{i+1}^-)$ and $P(t_{i+1}^-)$, according to the following differential equations

$$\dot{\underline{\hat{x}}}(t) = \underline{f}[\underline{\hat{x}}(t), \underline{u}(t), t] \tag{24}$$

$$\dot{P}(t) = F[t; \underline{\hat{x}}(t)]\,P(t) + P(t)F^T[t; \underline{\hat{x}}(t)] + G(t)Q(t)G^T(t) \tag{25}$$

The initial conditions used to solve these equations are $\underline{\hat{x}}(t_i^+)$ and $P(t_i^+)$. If (24) has a closed form solution, called state transition equation, the computation is simplified since no numerical integration is needed to propagate $\underline{\hat{x}}(t_i^+)$ to $\underline{\hat{x}}(t_{i+1}^-)$.

If the noise level is very high, the EKF does not give satisfactory performance, and an IEKF is needed. In the IEKF [12], (21) and (22) are replaced by setting $\underline{\hat{x}}_0$ equal to $\underline{\hat{x}}(t_i^-)$ and doing iterations on

$$K(t_i) = P(t_i^-)H^T[t_i; \underline{\hat{x}}_k]\left\{H[t_i; \underline{\hat{x}}_k]P(t_i^-)H^T[t_i; \underline{\hat{x}}_k] + R_v(t_i)\right\}^{-1} \tag{26}$$

$$\underline{\hat{x}}_{k+1} = \underline{\hat{x}}(t_i^-) + K(t_i)\left\{\underline{z}(t_i) - \underline{h}[\underline{\hat{x}}_k, t_i] - H(t_i, \underline{\hat{x}}_k)\left[\underline{\hat{x}}(t_i^-) - \underline{\hat{x}}_k\right]\right\} \tag{27}$$

for $k = 0, 1, \ldots, N-1$ and then setting $\hat{\underline{x}}(t_i^+) = \hat{\underline{x}}_N$. The iterations are stopped when the improvement in $\hat{\underline{x}}_k$ is less than a preselected small threshold.

The standard plant equation (16), measurement equation (17), and state transition equation for our model are formulated in the following. These formulations are readily applied to both EKF and IEKF.

## 3.1 Plant Equation

### Case I. Constant Angular Velocity

Let $\underline{x}(t) \equiv (\underline{r}^T(t), \underline{v}^T(t), \underline{a}^T, \underline{d}^T, \underline{q}^T(t), \underline{\omega}^T)^T$ be the state vector whose components are formed by the 19 motion states described previously. The plant equation (16) for the recursive filter is obtained from (5) and (7) as

$$\underbrace{\frac{d}{dt}\begin{pmatrix} \underline{r}(t) \\ \underline{v}(t) \\ \underline{a} \\ \underline{d} \\ \underline{q}(t) \\ \underline{\omega} \end{pmatrix}}_{\dot{\underline{x}}(t)} = \underbrace{\begin{pmatrix} O_3 & I_3 & O_3 & O_3 & O & O_3 \\ O_3 & O_3 & I_3 & O_3 & O & O_3 \\ O_3 & O_3 & O_3 & O_3 & O & O_3 \\ O_3 & O_3 & O_3 & O_3 & O & O_3 \\ O^T & O^T & O^T & O^T & \Omega(\underline{\omega}) & O^T \\ O_3 & O_3 & O_3 & O_3 & O & O_3 \end{pmatrix}\begin{pmatrix} \underline{r}(t) \\ \underline{v}(t) \\ \underline{a} \\ \underline{d} \\ \underline{q}(t) \\ \underline{\omega} \end{pmatrix}}_{\underline{f}[\underline{x}(t)]} \tag{28}$$

where $\Omega$ is defined in (8) , $O_3$ is the $3 \times 3$ zero matrix, and $O$ is the $3 \times 4$ zero matrix. Note that the plant equation (28) has been put in the standard form (16) with $\underline{u}(t)$ being a zero vector, $G(t)$ being a zero matrix, and $\underline{f}$ being a time-invariant function.

### Case II. Constant Precession

Let $\underline{x}(t) \equiv (\underline{r}^T(t), \underline{v}^T(t), \underline{a}^T, \underline{d}^T, \underline{q}^T(t), \underline{\omega}^T(t), \underline{p}^T)^T$ be the state vector whose components are formed by the 22 motion states/parameters described previously. The plant equation for the

recursive filter is obtained from (5), (7), and (11) as

$$\underbrace{\frac{d}{dt}\begin{pmatrix} \underline{r}(t) \\ \underline{v}(t) \\ \underline{a} \\ \underline{d} \\ \underline{q}(t) \\ \underline{\omega}(t) \\ \underline{p} \end{pmatrix}}_{\dot{\underline{x}}(t)} = \underbrace{\begin{pmatrix} O_3 & I_3 & O_3 & O_3 & O & O_3 & O_3 \\ O_3 & O_3 & I_3 & O_3 & O & O_3 & O_3 \\ O_3 & O_3 & O_3 & O_3 & O & O_3 & O_3 \\ O_3 & O_3 & O_3 & O_3 & O & O_3 & O_3 \\ O^T & O^T & O^T & O^T & \Omega[\underline{\omega}(t)] & O^T & O^T \\ O_3 & O_3 & O_3 & O_3 & O & P(\underline{p}) & O_3 \\ O_3 & O_3 & O_3 & O_3 & O & O_3 & O_3 \end{pmatrix}\begin{pmatrix} \underline{r}(t) \\ \underline{v}(t) \\ \underline{a} \\ \underline{d} \\ \underline{q}(t) \\ \underline{\omega}(t) \\ \underline{p} \end{pmatrix}}_{\underline{f}[\underline{x}(t)]} \tag{29}$$

where $P$ is defined in (12).

*Remarks.* In motion estimation in which structure information is not used, the states do not contain $\underline{d}$ and $\underline{q}$, because the structure coordinate system $S$ should be defined appropriately and a simple way to do this is to align $S$ and $I$ when $t = 0$. In the model which incorporates structure information, the states $\underline{d}$ and $\underline{q}$ are used to account for the initial position and orientation of $S$.

### 3.2 Measurement Equation

Let

- $(x_j(t_i), y_j(t_i), z_j(t_i))^T \equiv$ the noisy 3-D coordinates of the $j^{th}$ feature point in $I$ calculated from binocular images at the measurement time $t_i$.
- $(x_{sj}(t_i), y_{sj}(t_i), z_{sj}(t_i))^T \equiv$ the 3-D coordinates of the $j^{th}$ feature point in $S$ at $t_i$.

From (1) and (4), the measurement equation (17) for EKF/IEKF is obtained as

$$\underbrace{\begin{pmatrix} x_1(t_i) \\ y_1(t_i) \\ z_1(t_i) \\ \vdots \\ x_m(t_i) \\ y_m(t_i) \\ z_m(t_i) \end{pmatrix}}_{\underline{z}(t_i)} = \underbrace{\left\{ \begin{pmatrix} x_R(t_i) \\ y_R(t_i) \\ z_R(t_i) \\ \vdots \\ x_R(t_i) \\ y_R(t_i) \\ z_R(t_i) \end{pmatrix} + \begin{pmatrix} & & \\ & R[\underline{q}(t_i)] & \\ & & \\ & \vdots & \\ & & \\ & R[\underline{q}(t_i)] & \\ & & \end{pmatrix} \begin{pmatrix} x_{s1}(t_i) - d_x \\ y_{s1}(t_i) - d_y \\ z_{s1}(t_i) - d_z \\ \vdots \\ x_{sm}(t_i) - d_x \\ y_{sm}(t_i) - d_y \\ z_{sm}(t_i) - d_z \end{pmatrix} \right\}}_{\underline{h}[\underline{x}(t_i), t_i]} + \underbrace{\begin{pmatrix} n_{x1}(t_i) \\ n_{y1}(t_i) \\ n_{z1}(t_i) \\ \vdots \\ n_{xm}(t_i) \\ n_{ym}(t_i) \\ n_{zm}(t_i) \end{pmatrix}}_{\underline{v}(t_i)} \tag{30}$$

where $n_{xj}(t_i), n_{yj}(t_i), n_{zj}(t_i)$ are Gaussian noise variables with zero mean, and $m$ is the total number of feature points at time $t_i$.

*Remarks.*

1. $m$ can be a function of $t_i$, i.e. the number of feature points is not necessarily the same for each measurement time $t_i$.

2. The feature points may be chosen differently at different time $t_i$.

3. Occlusion can be treated in a straight forward way because of 1. and 2.

4. The time intervals between two consecutive measurements need not be constant.

5. The variances of each noise components can be different. That is, the accuracy of each measurement is allowed to be different.

6. The noise components can be correlated. The interdependence between the noise components can be considered explicitly in EKF/IEKF through the measurement noise covariance matrix $R_v(t_i)$ in (18).

7. The noise distribution has been discussed by other authors. Blostein and Huang [10] analyzed the noise distributions in the 3-D coordinates obtained by stereo triangulation. They concluded that the noise levels in different coordinate directions ( horizontal, vertical and range) are different. The measurement model (30) has the flexibility to account for different noise levels if the variance of the noise in each coordinate direction is assigned a different value. Under Blostein and Huang's assumptions, the noise distributions obtained are not Gaussian. However, Matthies and Shafer [18] report that using 3-D Gaussian distributions to model triangulation

error can lead to good performance in motion analysis. For the purpose of this report we assume that the noise variables are distributed as Gaussian.

### 3.3 State Transition Equation

The closed forms of state transition equations will be given to propagate the states in the recursive filters. The computation to propagate the states is simplified since no numerical integration of (24) is needed.

**Case I. Constant Angular Velocity**

From (6) and (10), the state transition equation, which propagate $\hat{\underline{x}}(t_i^+)$ forward to $\hat{\underline{x}}(t_{i+1}^-)$, is

$$\begin{pmatrix} \underline{r}(t_{i+1}^-) \\ \underline{v}(t_{i+1}^-) \\ \underline{a}(t_{i+1}^-) \\ \underline{d}(t_{i+1}^-) \\ \underline{q}(t_{i+1}^-) \\ \underline{\omega}(t_{i+1}^-) \end{pmatrix} = \begin{pmatrix} I_3 & I_3(t_{i+1}^- - t_i^+) & I_3(t_{i+1}^- - t_i^+)^2/2 & O_3 & O & O_3 \\ O_3 & I_3 & I_3(t_{i+1}^- - t_i^+) & O_3 & O & O_3 \\ O_3 & O_3 & I_3 & O_3 & O & O_3 \\ O_3 & O_3 & O_3 & I_3 & O & O_3 \\ O^T & O^T & O^T & O^T & \Phi(\underline{\omega}; t_{i+1}^- - t_i^+) & O^T \\ O_3 & O_3 & O_3 & O_3 & O & I_3 \end{pmatrix} \begin{pmatrix} \underline{r}(t_i^+) \\ \underline{v}(t_i^+) \\ \underline{a}(t_i^+) \\ \underline{d}(t_i^+) \\ \underline{q}(t_i^+) \\ \underline{\omega}(t_i^+) \end{pmatrix} \tag{31}$$

where

$$\Phi(\underline{\omega}; t_{i+1}^- - t_i^+) = \begin{cases} I_4 \cos\frac{|\underline{\omega}|(t_{i+1}^- - t_i^+)}{2} + \frac{2}{|\underline{\omega}|}\Omega \sin\frac{|\underline{\omega}|(t_{i+1}^- - t_i^+)}{2} & \text{if } \underline{\omega} \neq \underline{0} \\ I_4 & \text{otherwise} \end{cases}$$

**Case II. Constant Precession**

From (6), (14), and (13) the state transition equation is

$$\begin{pmatrix} \underline{r}(t_{i+1}^-) \\ \underline{v}(t_{i+1}^-) \\ \underline{a}(t_{i+1}^-) \\ \underline{d}(t_{i+1}^-) \\ \underline{q}(t_{i+1}^-) \\ \underline{\omega}(t_{i+1}^-) \\ \underline{p}(t_{i+1}^-) \end{pmatrix} = \begin{pmatrix} I_3 & I_3(t_{i+1}^- - t_i^+) & I_3(t_{i+1}^- - t_i^+)^2/2 & O_3 & O & O_3 & O_3 \\ O_3 & I_3 & I_3(t_{i+1}^- - t_i^+) & O_3 & O & O_3 & O_3 \\ O_3 & O_3 & I_3 & O_3 & O & O_3 & O_3 \\ O_3 & O_3 & O_3 & I_3 & O & O_3 & O_3 \\ O^T & O^T & O^T & O^T & \Phi_1 & O^T & O^T \\ O_3 & O_3 & O_3 & O_3 & O & \Phi_2 & O_3 \\ O_3 & O_3 & O_3 & O_3 & O & O_3 & I_3 \end{pmatrix} \begin{pmatrix} \underline{r}(t_i^+) \\ \underline{v}(t_i^+) \\ \underline{a}(t_i^+) \\ \underline{d}(t_i^+) \\ \underline{q}(t_i^+) \\ \underline{\omega}(t_i^+) \\ \underline{p}(t_i^+) \end{pmatrix} \tag{32}$$

where $\Phi_1$ and $\Phi_2$ are defined in (14) and (13), respectively.

# 4 Filtering

The standard filtering algorithm is summarized as follows.

1. Since the plant model (16) is given as (28) or (29), and the measurement model (17) is given as (30), the linearized matrices $F$ and $H$ are obtained from (19) and (20).

2. The initial guess of the state vector is assigned to $\hat{\underline{x}}(t_1^-)$. The initial guess of the approximate covariance matrix is assigned to $P(t_1^-)$.

3. Set $i = 1$.

4. At time $t_i$, the measurement $\underline{z}(t_i)$ is taken.

5. At time $t_i^+$, immediately after the measurement time $t_i$, the estimate $\hat{\underline{x}}(t_i^+)$ for the state vector is obtained by (21) and (22). The approximate covariance matrix $P(t_i^-)$ is updated to $P(t_i^+)$ using (23).

   If an IEKF is used, $\hat{\underline{x}}(t_i^+)$ is obtained by the iteration defined in (26) and (27), instead of (21) and (22).

6. At time $t_{i+1}^-$, immediately before the measurement time $t_{i+1}$, the estimate $\hat{\underline{x}}(t_{i+1}^-)$ is propagates from $\hat{\underline{x}}(t_i^+)$ by the state transition equation (31) or (32). $P(t_{i+1}^-)$ is obtained by solving (25) using numerical integration [12].

7. Increase $i$ by 1, then go to step 4 to incorporate new measurements.

# 5 Simulation Results

Computer simulations are used to investigate the models. The object studied is a rigid transparent cube with corners on $(\pm 1, \pm 1, \pm 1)$ in the structure coordinate system $S$. The feature points are chosen from the corners of the rigid body. The object size is defined as the maximum length between two feature points. The simulated measurements are generated by the following scheme. First, the noise-free 3-D coordinates feature points in the inertial coordinate system $I$ are generated by a computer program. Then independent Gaussian noise variables with standard deviations $\sigma_x, \sigma_y$, and $\sigma_z$ are added to obtain the noisy measurements in the x, y, and z directions of the inertial coordinate system $I$, respectively. The time period between measurements is 0.5. The noise level in x-direction is defined as $\sigma_x$/(object size).

Two experimental results are shown in Figure 3 and Figure 4 corresponding to constant angular velocity and precession rates, respectively. In each experiment, only 3 feature points are used. The object size is $\sqrt{2}$. The ordinates equal the differences between the estimated motion parameter values and the true values.

*Case I. Constant Angular Velocity*

Figure 3 shows a typical sample path when $\sigma_x = \sigma_y = 0.14$ and $\sigma_z = 0.35$. It corresponds to the noise level of 10% of the object size in x- and y- directions and 25% of the object size in z- direction. An IEKF algorithm was used. This experiment can be interpreted as an example for the case when triangulation is used to obtain the 3-D measurements. As it has been noted [10] that in calculating the 3-D coordinates of features by applying triangulation to stereo image pairs, the noise level in the z-direction is much higher than the ones in the x- and y- directions. In this and the next experiment the initial conditions chosen to begin the recursive IEKF computations were intentionally chosen to be quite far away from the true values. The true and initial values used in this and the next experiment are given in Table 1.

*Case II. Constant Precession*

Figure 4. shows a typical sample path when $\sigma_x = \sigma_y = \sigma_z = 0.14$. This corresponds to the noise level of 10% of the object size.

As can be seen, the quaternions converge within 3 frames. This fast convergence rate is because only one stereo image pair is enough to determine the orientation of the object at the measurement time when the image pair is taken if 3 noncolinear feature points are used, as shown in [19]. It is also observed that, once the filter converges, the estimates reach different accuracy : acceleration highest, velocity second, and position lowest in the translational components; precession highest, angular velocity second, and quaternion lowest in the rotational components. A possible explanation is as follows: Since the measurements are taken only at discrete times, the velocity is meaningless if only one time instant is considered. Thus the velocity can be considered as a quantity involved in at least two image pairs at different times. By the same reason, the acceleration involves at least three image pairs, angular velocity two, and precession three. If more image pairs are based, higher ability to smooth out the noise is expected. Thus the above order of accuracy is formed.

*Remarks :*

(i) In the case of very highly deviated initial guesses, divergence is also observed, which is a well known feature for EKF and IEKF. A batch algorithm can be used to start the filter with better initial guess.

(ii) In the case of constant angular velocity, if started from other initial guesses, the filter has the possibility to converge to other states which are indistinguishable from the "true" states. This is an issue of the uniqueness of the motion states and is analyzed in [19].

(iii) In the case of constant angular velocity, since there is one extra degree of freedom involved in $\underline{r}(t)$ and $\underline{d}$, the estimates $\hat{\underline{r}}(t)$ and $\hat{\underline{d}}$ are different from the "true" states, but consistent with the "true"states. This is also analyzed in [19].

# 6 Performance Bounds

The estimation process formulated can be conceptually separated into two stages: establishing models to describe the underlying motion and measurements and applying an estimator to estimate the model parameters. After the first stage, it is desirable to know what the maximum obtainable accuracy of the estimates of the model parameters is. If the maximum attainable accuracy is not satisfactory, the models need to be changed before the design of an estimator. Once the estimator is designed, its performance can be compared with the maximum obtainable accuracy to give an insight into the efficiency of the estimator. Since variances quantify the accuracy of estimates if bias is small, the maximum obtainable accuracy can be quantified by the greatest lower bound of the variances of the estimates. In this section, this performance bound will be discussed using the Cramér-Rao inequality [20]. The bound is expected to be time-varying since the motion states are functions of time. The performance bound is intrinsic to the models, independent of the estimators applied, and gives an indication of the noise immunity of the motion model.

## 6.1 Fisher's Information Matrix

As needed in the Cramér-Rao inequality, the Fisher's information matrix, which is time varying, will be calculated first.

Recall that $\underline{x}(t)$ is the state vector at time $t$. Let

- measurement vector $\underline{Z}_t \equiv$ the collection of all measurements taken no later than $t$.

Then we have

$$\underline{Z}_t = \underline{h}_t[\underline{x}(t)] + \underline{n}_t \tag{33}$$

where $\underline{n}_t$ is the noise vector. This can be seen as follows. From the measurement model (30), the measurements at any time $t'$ are functions of $\underline{x}(t')$ plus a noise term. But $\underline{x}(t')$ is a function of $\underline{x}(t)$ according to the state transition equations (31) or (32). Therefore the measurement vector at $t'$ is a function of $\underline{x}(t)$ plus a noise term, as shown in (33).

The Fisher's information matrix is calculated directly from the definition [20],

$$J(t) = E\left\{ \left[\frac{\partial}{\partial \underline{x}(t)} \ln f\left(\underline{Z}_t \mid \underline{x}(t)\right)\right]^T \left[\frac{\partial}{\partial \underline{x}(t)} \ln f\left(\underline{Z}_t \mid \underline{x}(t)\right)\right] \middle| \underline{x}(t) \right\} \tag{34}$$

where $E$ is the expectation operator and $f$ is the conditional probability density function of $\underline{Z}_t$ given $\underline{x}(t)$.

A detailed calculation of $J(t)$ in the case of independent noise assumption can be found in Appendix II.

## 6.2 Cramér-Rao inequality

The Cramér-Rao inequality for the variances of parameter estimates will be reviewed.

Assume that $J(t)$ is nonsingular and $\hat{\underline{x}}(t)$ is an unbiased estimator of $\underline{x}(t)$. Then the Cramér-Rao inequality says

$$E\left[\left(\underline{x}(t) - \hat{\underline{x}}(t)\right)\left(\underline{x}(t) - \hat{\underline{x}}(t)\right)^T \middle| \underline{x}(t)\right] \geq \left[J(t)\right]^{-1} \tag{35}$$

Because the inequality in (35) holds in the sense that the difference of the left and the right hand sides is non-negative definite, the $i^{th}$ diagonal element of $[J(t)]^{-1}$ gives lower bounds of the variance of the $i^{th}$ component of $\hat{x}(t)$ given $\underline{x}(t)$. These lower bounds are called the Cramér-Rao lower bounds (CRLB). Expressions available [12] for CRLB of biased estimators need to know the bias term explicitly as a function of parameters. In our problem, we do not have explicit expressions for the bias term. Hence we are forced to use CRLB for unbiased estimators.

If $J(t)$ is singular, as the case in our model, then the above inequality needs to be modified. Let

- $\underline{x}'(t) \equiv$ the vector obtained from deleting the $j^{th}$ component of $\underline{x}(t)$.
- $\hat{\underline{x}}'(t) \equiv$ any unbiased estimator of $\underline{x}'(t)$.
- $J'(t) \equiv$ the square matrix obtained from deleting the $j^{th}$ row and column of $J(t)$.

Assuming that $J'(t)$ is nonsingular, then we have

$$E\left[\left(\underline{x}'(t) - \hat{\underline{x}}'(t)\right)\left(\underline{x}'(t) - \hat{\underline{x}}'(t)\right)^T \middle| \underline{x}(t)\right] \geq \left[J'(t)\right]^{-1} \tag{36}$$

This can be obtained by following the same derivations of Cramér-Rao inequality in [20] exactly except starting from the equality $E\left[(\underline{x}' - \hat{\underline{x}}')|\underline{x}\right] = 0$.

If more than one component of $\underline{x}(t)$ is deleted, the same result can be obtained.

*Remarks.*

(i) In the model of constant precession, $J(t)$ is nonsingular. But in the model of constant angular velocity, $J(t)$ is singular because of the extra degree of freedom in the states $\underline{r}(t)$ and $\underline{d}$. Detailed discussion about the relation between the uniqueness of motion states and the singularity of $J(t)$ can be found in [19].

(ii) If neither $J(t)$ nor $J'(t)$ is singular, the CRLB obtained by (36) is looser (smaller) than the one obtained by (35). A brief explanation for this can be found in Appendix III.

### 6.3 Performance Bounds and Estimator Efficiency

The CRLB obtained above will be used as a performance bound for our model. The following will be checked : (i) the bias of the output of EKF/IEKF. (This needs to be checked since the CRLB obtained from (35) is for unbiased estimators.) (ii) the tightness of the bound.

(i) the bias of the output of EKF/IEKF used in our model.

Since the CRLB mentioned above is for unbiased estimators, it approximately offer a lower bound only when the bias is small. Thus CRLB is expected to be a lower bound of the output of EKF/IEKF because of the observed small bias shown in Figure 5, which is the sample bias from Monte Carlo trials. ( In Figure 5, the ordinate equals the difference between the sample mean from 50 Monte Carlo trials and the true states in the constant angular velocity motion. 3 and 5 feature points are used. the object size is $\sqrt{2}$. The noise variance is $\sigma_x = \sigma_y = \sigma_z = 0.2$, which corresponds to the noise level of 14% of the object size. The true and initial values of states used to start the filters are given in Table 1. )

As mentioned in [11], the local iteration in EKF/IEKF produces biased estimates, because the mode of the posterior density is used for the conditional mean. But if the mean is close

to the mode, the bias is small. As shown in Figure 5, after the filter converges, the biases of the estimates for velocity, acceleration, and angular velocity tend to zero quickly. But the biases of quaternions are relatively large, and behave as sinusoidal functions of time with the period $4\pi/|\underline{w}|$, which is twice the period of rotation of the object. (Recall that $+\underline{q}$ and $-\underline{q}$ describe the same rotation.)

(ii) the tightness of the bound

The CRLB and the sample variance from 50 Monte Carlo trials are quite close, as can be seen in Figure 6. Therefore, it is believed that the CRLB gives very good lower bound for the variance of the estimators of state. ( The parameters used in Figure 6 are the same as those in Figure 5. )

Since there is one extra degree of freedom involved in states, according to the analysis in [19], $J(t)$ is singular. The first row and column of $J(t)$ are deleted to form $J'(t)$. Then the nonsingular matrix $J'(t)$ is used to compute the CRLB. The calculated CRLB and the sample variances of the states obtained from 50 Monte Carlo trials starting from the same initial guess are shown in Figure 6. As can be seen, the CRLB and sample variances are quite close.

## Acknowledgment

The authors are grateful to Dr. T.J. Broida for many helpful discussions.

## References


[1] J. K. Aggarwal. Motion and time-varying imagery-An overview. In *Proc. of IEEE Workshop on Motion: Representation and Analysis*, May 1986.

[2] J. Q. Fang and T. S. Huang. Some experiments on estimating the 3-D motion parameters of a rigid body from two consecutive image frames. *IEEE Trans. on Patt. Anal. Mach. Intell.*, PAMI-6:545–554, September 1984.

[3] T. S. Huang et. al. Motion detection and estimation from stereo image sequences: Some preliminary experimental results. In *Proc. of IEEE Workshop on Motion: Representation and Analysis*, May 1986.

[4] J. K. Aggarwal and A. Mitiche. Structure and motion from images: fact and fiction. In *Proc. Third Workshop on Computer Vision: Representation and Control*, Michigan, pp. 127–128, Oct. 1985.

[5] T. J. Broida and R. Chellappa. Estimation of object motion parameters from noisy images. *IEEE Transactions on Patt. Anal. and Mach. Intell.*, PAMI-8(1), pp. 90–99, January 1986.

[6] T. J. Broida and R. Chellappa. Kinematics and structure of a rigid object from a sequence of noisy images. In *Proc. of IEEE Workshop on Motion: Representation and Analysis*, May 1986.

[7] T. J. Broida and R. Chellappa. Kinematics and structure of a rigid object from a sequence of noisy images : A batch approach. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, Miami Beach, FL, June 1986, pp. 176-182.

[8] J. Weng et. al. 3-D motion estimation, understanding, and prediction from noisy image sequences. *IEEE Transactions on Patt. Anal. and Mach. Intell.*, PAMI-9(3), pp. 370–389, May 1987.

[9] T. S. Huang and S. D. Blostein. Robust algorithm for motion estimation based on two sequential stereo image pairs. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, San Francisco, CA, 1985, pp.518-523.

[10] S. D. Blostein and T. S. Huang. Error analysis in stereo determination of 3-D point positions. *IEEE Transactions on Patt. Anal. and Mach. Intell.*, PAMI-9(6), pp. 752-765, Nov. 1987.

[11] A. H. Jazwinski. *Stochastic Processes and Filtering Theory.* Academic Press, 1970.

[12] P. S. Maybeck. *Stochastic models, estimation, and control*, volume 2. Academic Press, 1982.

[13] C. T. Chen. *Linear System Theory and Design.* Holt, Rinehart and Winston, 1984.

[14] B. K. P. Horn. *Robot Vision.* The MIT Press, 1986.

[15] G. D. Niva. The use of quaternions with an all-attitude IMU. In *Proc. Ann. Rocky Mountain Guidance and Control Conf.*, pages 269–283, January 1982.

[16] B. Friedland. Analysis of strapdown navigation using quaternions. *IEEE Trans. on Aerospace and Elect. Systems*, AES-14(5), September 1978.

[17] J. R. Wertz, editor. *Spacecraft Attitude Determination and Control.* D. Reidel Publishing Co., 1978.

[18] L. Matthies and S. A. Shafer. Error modeling in stereo navigation. *IEEE Journal of Robotics and Automation,* June 1987

[19] G. S. Young and R. Chellappa. 3-D motion estimation using a sequence of noisy stereo images, Part II. uniqueness results. *SIPI report*, University of Southern California, Los Angeles, CA 90089. Nov. 1987.

[20] H. W. Sorenson. *Parameter estimation: principles and problems.* Marcel Deckker, NY.

[21] G. Strang. *Linear algebra and its application.* Academic press, 1980.

# Appendix I

In this appendix, the closed form solution (14) of the quaternion $\underline{q}(t)$ under the assumption of constant precession is derived.

For clarity, the quaternions will be denoted by $\hat{}$, e.g. $\hat{q}$, in this appendix.

The goal is to solve the following first order linear time-varying differential equation

$$\dot{\hat{q}}(t) = \Omega\left[\underline{\omega}(t)\right]\hat{q}(t) \tag{37}$$

where $\underline{\omega}(t)$ is obtained from (13).

We first (1) define the quaternion multiplication [14], (2) then do a change of variable to reduce (37) to a *time-invariant* equation, (3) and solve the time-invariant equation to get the solution. Intuitive interpretations of the derivations are given in the remarks at the end of this appendix.

(1) definition of the quaternion multiplication [14]

If $\hat{r}$ and $\hat{\phi}$ are two quaternions, the multiplication of $\hat{r}$ and $\hat{\phi}$ are defined as

$$\hat{r}\hat{\phi} \equiv \begin{pmatrix} r_4 & -r_3 & r_2 & r_1 \\ r_3 & r_4 & -r_1 & r_2 \\ -r_2 & r_1 & r_4 & r_3 \\ -r_1 & -r_2 & -r_3 & r_4 \end{pmatrix} \begin{pmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \\ \phi_4 \end{pmatrix} = \begin{pmatrix} \phi_4 & \phi_3 & -\phi_2 & \phi_1 \\ -\phi_3 & \phi_4 & \phi_1 & \phi_2 \\ \phi_2 & -\phi_1 & \phi_4 & \phi_3 \\ -\phi_1 & -\phi_2 & -\phi_3 & \phi_4 \end{pmatrix} \begin{pmatrix} r_1 \\ r_2 \\ r_3 \\ r_4 \end{pmatrix} \tag{38}$$

The conjugate of $\hat{q}$ is defined as

$$\hat{q}^* \equiv (-q_1, -q_2, -q_3, q_4)^T \tag{39}$$

The unit quaternion $\hat{I}$ is defined as

$$\hat{I} \equiv (0, 0, 0, 1)^T$$

Equations (38) and (39) yield

$$\hat{q}^*\hat{q} = \hat{q}\hat{q}^* = |\hat{q}|^2\hat{I} \tag{40}$$

and

$$\hat{q}\hat{I} = \hat{I}\hat{q} = \hat{q}$$

(2) reducing (37) to a time-invariant equation (46)

Use the vector $\underline{\omega}(t)$ to define two quaternions

$$\hat{\omega}_0 \equiv \begin{pmatrix} \underline{\omega}(0) \\ 0 \end{pmatrix}$$

and

$$\hat{\omega}_t \equiv \begin{pmatrix} \underline{\omega}(t) \\ 0 \end{pmatrix}$$

Also define

$$\hat{\phi}(t) \equiv \left( \frac{p_x}{|\underline{p}|} \sin \frac{-|\underline{p}|t}{2}, \frac{p_y}{|\underline{p}|} \sin \frac{-|\underline{p}|t}{2}, \frac{p_z}{|\underline{p}|} \sin \frac{-|\underline{p}|t}{2}, \cos \frac{-|\underline{p}|t}{2} \right)^T \tag{41}$$

A direct calculation shows that an equivalent form of (13) is

$$\hat{\omega_t}^* = \hat{\phi}^* \hat{\omega}_0^* \hat{\phi} \tag{42}$$

It is obvious that (37) can be put into an equivalent form

$$\dot{\hat{q}} = \frac{1}{2} \hat{q} \hat{\omega}_t^* \tag{43}$$

As a change of variable, define

$$\hat{r} \equiv \hat{q}\hat{\phi}^*$$

Post-multiplying both sides by $\hat{\phi}$, we have

$$\hat{r}\hat{\phi} = \hat{q} \tag{44}$$

Substituting of (44) and (42) into (43) yields

$$\frac{d}{dt}(\hat{r}\hat{\phi}) = \frac{1}{2} \hat{r} \hat{\omega}_0^* \hat{\phi} \tag{45}$$

After the substitution of

$$\frac{d}{dt}\left(\hat{r}\hat{\phi}\right) = \dot{\hat{r}}\hat{\phi} + \hat{r}\dot{\hat{\phi}}$$

into (45), we get

$$\dot{\hat{r}}\hat{\phi} = \frac{1}{2}\hat{r}\left(\hat{\omega}_0^* - 2\dot{\hat{\phi}}\hat{\phi}^*\right)\hat{\phi}$$

After post-multiplying both sides by $\hat{\phi}^*$ and the substitution of

$$2\dot{\hat{\phi}}\hat{\phi}^* = \left(-p_x, -p_y, -p_z, 0\right)^T$$

which is a result of (41), we get the time-invariant equation

$$\dot{\hat{r}} = \frac{1}{2}\hat{r}\begin{pmatrix} \underline{\omega}(0) - \underline{p} \\ 0 \end{pmatrix}^* \tag{46}$$

(3) final solution (14)

Note that (46) has the same form as (43), which is equivalent to (37). Therefore the solution of (46) is

$$\hat{r}(t) = \left[I_4 \cos\frac{|\underline{\lambda}|t}{2} + \frac{2}{|\underline{\lambda}|}\Omega(\underline{\lambda})\sin\frac{|\underline{\lambda}|t}{2}\right]\hat{r}(0) \tag{47}$$

where

$$\underline{\lambda} \equiv \underline{\omega}(0) - \underline{p}$$

and $\Omega$ is defined in (8). Then the closed form solution for $\hat{q}(t)$ is obtained from (44) as

$$\hat{q}(t) = \hat{r}(t)\hat{\phi}(t) \tag{48}$$

Setting $t = 0$, we get the value for $\hat{r}(0)$ as

$$\hat{r}(0) = \hat{q}(0) \tag{49}$$

Summarizing (47), (48), and (49) into forms of matrix multiplication by the quaternion multiplication defined in (38) gives the final solution (14).

*Remarks.*

(i) There is an intuitive interpretation for the derivation of the closed form solution. We first find a new coordinate system, denoted as $N$, in which the angular velocity vector of the rigid body is constant. The quaternion $\hat{r}$ used above actually describes the alignment

from the object coordinate system $B$ to $N$, and $\hat{\phi}$ describes the alignment from $N$ to the inertial coordinate system $I$. Thus the quaternion $\hat{q}$, which describes the alignment from $B$ to $I$, equals the multiplication of $\hat{r}$ and $\hat{\phi}$.

(ii) An intuitive understanding of (41) is as follows. When $t = 0$, $N$ is aligned with $I$. Furthermore, $N$ rotates w.r.t. $I$ with the constant angular velocity $\underline{p}$. Therefore we can write down the quaternion $\hat{\phi}$, which describes the alignment from $N$ to $I$, directly either from the result of Case I in Section 2.2.2 or by intuition.

(iii) The angular velocity vector of the rigid body relative to $N$ is $\underline{\lambda} \equiv \underline{\omega}(0) - \underline{p}$. This can be seen as follows. At $t = 0$, the relative angular velocity of the rigid body to $I$ is $\underline{\omega}(0)$. The relative angular velocity of $N$ to $I$ is $\underline{p}$. Thus we get the relative angular velocity of $B$ to $N$ as $\underline{\omega}(0) - \underline{p}$. Notice that, relative to $I$, both $N$ and $\underline{\omega}(t)$ rotate with the same angular velocity $\underline{p}$. Thus in $N$, the angular velocity vector of the rigid body is constant.

(iv) Equation (46) is a restatement of (iii). This can be seen by a comparison of (46) and (43), both of which describe the time propagation of the quaternion representing the alignment between two coordinates with constant relative angular velocity.

## Appendix II

In this appendix, we give more detailed calculations of the Fisher's information matrix $J(t)$ for the case of independent noises.

Assume that $0 \leq t_1 < t_2 < \ldots < t_i < \ldots < t_l \leq t$, and the measurements are taken at $t_1, \ldots, t_l$ before $t$. For simplicity, define

- $\underline{x} \equiv \underline{x}(t)$, the state at $t$
- $\underline{x}^i \equiv \underline{x}(t_i)$, the state at measurement time $t_i$
- $\underline{z}^i \equiv$ the collection of all measurements at $t_i$
- $\underline{v}^i \equiv$ the noise vector at $t_i$
- $\underline{Z} = (\underline{z}^{1T}, \ldots, \underline{z}^{lT})^T \equiv$ the measurement vector which collects all of the measurements before $t$.

Putting the state transition equation (31) or ( 32) in a concise form, we get

$$\underline{x}^i = \Phi(t_i - t; \underline{x})\underline{x} \tag{50}$$

where $\Phi(t_i - t; \underline{x})$ is the state transition matrix. Putting the measurement model (30) in a concise form, we have

$$\underline{z}^i = \underline{h}^i(\underline{x}^i) + \underline{v}^i \tag{51}$$

From (50) and (51), the more explicit form for equation (33) is

$$\underline{Z} = \underbrace{\begin{pmatrix} \underline{z}^1 \\ \vdots \\ \underline{z}^l \end{pmatrix}}_{\underline{Z}_t} = \underbrace{\begin{pmatrix} \underline{h}^1\left[\Phi(t_1 - t; \underline{x})\underline{x}\right] \\ \vdots \\ \underline{h}^l\left[\Phi(t_l - t; \underline{x})\underline{x}\right] \end{pmatrix}}_{\underline{h}_t\left[\underline{x}(t)\right]} + \underbrace{\begin{pmatrix} \underline{v}^1 \\ \vdots \\ \underline{v}^l \end{pmatrix}}_{\underline{n}_t} \tag{52}$$

Let

- $h_j^i \equiv$ the $j^{th}$ component of $\underline{h}^i$
- $m_i \equiv$ the number of components in $\underline{h}^i$ or the number of measurements taken at $t_i$
- $z_j^i \equiv$ the $j^{th}$ component of $\underline{z}^i$
- $v_j^i \equiv$ the $j^{th}$ component of $\underline{v}^i$.
- $f(\underline{Z}|\underline{x}) \equiv$ the conditional density function of $\underline{Z}$ given $\underline{x}$.

Assuming that $v_j^i$ are independent Gaussian random variables with mean 0 and variance ${\sigma_j^i}^2$, from (52), we have

$$\begin{aligned}
\frac{\partial}{\partial \underline{x}} \ln f(\underline{Z}|\underline{x}) &= \frac{\partial}{\partial \underline{x}} \ln \prod_{i=1}^{l} \prod_{j=1}^{m_i} f(z_j^i|\underline{x}) \\
&= \frac{\partial}{\partial \underline{x}} \ln \prod_{i=1}^{l} \prod_{j=1}^{m_i} \frac{1}{\sqrt{2\pi}\sigma_j^i} \exp\left[-\frac{1}{2{\sigma_j^i}^2}\left(z_j^i - h_j^i[\Phi(t_i - t; \underline{x})\underline{x}]\right)^2\right] \\
&= \sum_{i=1}^{l}\sum_{j=1}^{m_i} \frac{1}{{\sigma_j^i}^2}\left(z_j^i - h_j^i[\Phi(t_i - t; \underline{x})\underline{x}]\right) \frac{\partial}{\partial \underline{x}}\left(h_j^i[\Phi(t_i - t; \underline{x})\underline{x}]\right)
\end{aligned}$$

Noticing that

$$\begin{aligned}
E\left[\left(z_j^i - h_j^i[\Phi\left(t_i - t; \underline{x}\right)\underline{x}]\right)^2 \middle| \underline{x}\right] &= E\left[\left(v_j^i\right)^2 \middle| \underline{x}\right] \\
&= {\sigma_j^i}^2
\end{aligned}$$

from the definition of the Fisher's information matrix (34), we have

$$\begin{aligned} J(t) &= E\left\{ \left[\frac{\partial}{\partial \underline{x}} \ln f(\underline{Z} \mid \underline{x})\right]^T \left[\frac{\partial}{\partial \underline{x}} \ln f(\underline{Z} \mid \underline{x})\right] \middle| \underline{x} \right\} \\ &= \sum_{i=1}^{l} \sum_{j=1}^{m_i} \frac{1}{\sigma_j^{i\,2}} \left[\frac{\partial h_j^i}{\partial \underline{x}^i} \frac{\partial}{\partial \underline{x}} (\Phi(t_i - t; \underline{x})\, \underline{x})\right]^T \left[\frac{\partial h_j^i}{\partial \underline{x}^i} \frac{\partial}{\partial \underline{x}} (\Phi(t_i - t; \underline{x})\, \underline{x})\right] \end{aligned}$$

## Appendix III

In this appendix, we are going to show if neither $J(t)$ nor $J'(t)$ is singular, the CRLB obtained by (36) is looser (smaller) than the one obtained by (35).

The time dependence will be suppressed in the discussion. Assume that $\underline{x}'$ is obtained from deleting $k$ components of $\underline{x}$. Because we can always rearrange the order of the elements in $\underline{x}$, without loss of generality, it is assumed that the last $k$ elements of $\underline{x}$ are deleted. That is, $J$ has the form

$$J = \begin{pmatrix} J' & B \\ B^T & C \end{pmatrix},$$

where $C$ is a $k \times k$ matrix. Also note that if

$$J^{-1} = \begin{pmatrix} J' & B \\ B^T & C \end{pmatrix}^{-1} \equiv \begin{pmatrix} D & E \\ E^T & F \end{pmatrix},$$

where $F$ is a $k \times k$ matrix, then [20]

$$D = J'^{-1} + J'^{-1} B F B^T J'^{-1} \tag{53}$$

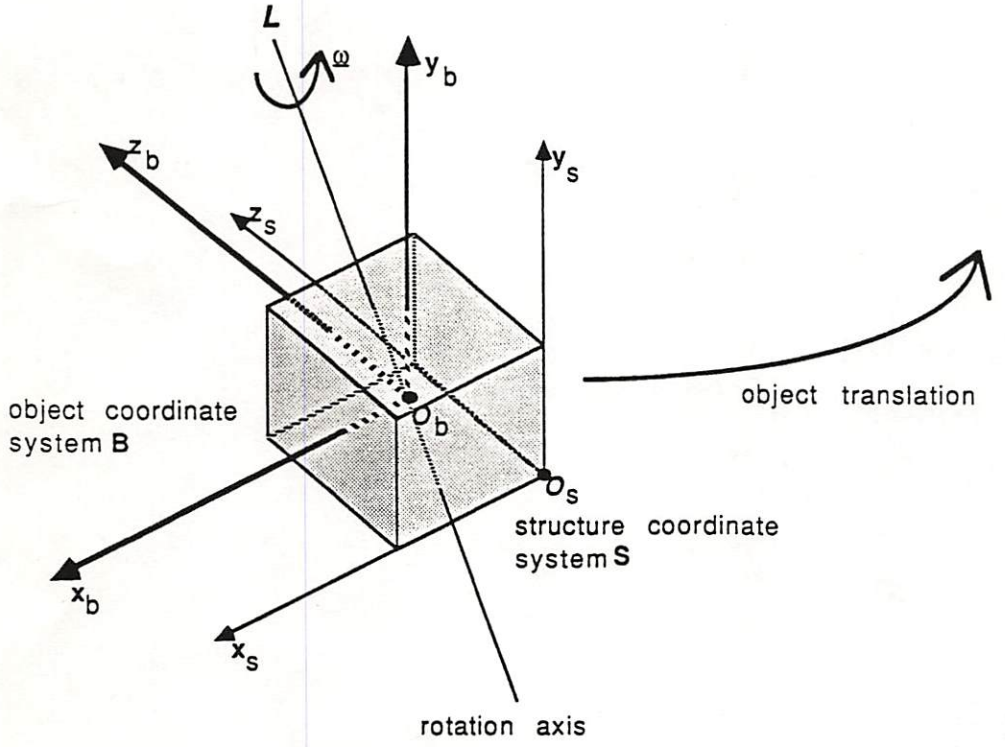
$J$ is positive definite, hence $J^{-1}$ is also positive definite. $F$ is a square submatrix of $J^{-1}$ along the diagonal, thus $F$ is positive definite. There exists a $k \times k$ matrix $G$ such that $F = GG^T$ [21]. Therefore,

$$\begin{aligned} D - J'^{-1} &= J'^{-1} B F B^T J'^{-1} \\ &= \left(J'^{-1} B G\right)\left(J'^{-1} B G\right)^T, \end{aligned}$$

which implies that the matrix $(D - J'^{-1})$ is positive definite. Thus the diagonal elements of $D$ are greater than the corresponding ones of $J'^{-1}$. As a result, $J^{-1}$ gives tighter bounds than $J'^{-1}$.

| | constant angular velocity (Figure 3) | | constant precession (Figure 4) | | Monte Carlo trials (Figure 5 & 6) | |
|---|---|---|---|---|---|---|
| state | true | initial guess | true | initial guess | true | initial guess |
| $x_R(0)$ | 4 | 2 | 4 | 3 | 4 | 2 |
| $y_R(0)$ | -3 | -1 | -3 | -2 | -3 | -1 |
| $z_R(0)$ | 1 | 0 | 1 | 0 | 1 | 0 |
| $v_x(0)$ | 1 | 0 | 2 | 0 | 1 | 0 |
| $v_y(0)$ | 0.8 | 0 | 1.8 | 0 | 0.8 | 0 |
| $v_z(0)$ | -0.7 | 0 | -1.7 | 0 | -0.7 | 0 |
| $a_x$ | -0.3 | 0 | -0.9 | 0 | -0.3 | 0 |
| $a_y$ | 0.2 | 0 | 1.2 | 0 | 0.2 | 0 |
| $a_z$ | 0.25 | 0 | 0.75 | 0 | 0.25 | 0 |
| $d_x$ | 2 | 0 | 0.5 | 0 | 2 | 0 |
| $d_y$ | 1 | 0 | 0.4 | 0 | 1 | 0 |
| $d_z$ | 1 | 0 | 0.2 | 0 | 1 | 0 |
| $q_1(0)$ | 0.2 | 0 | 0.2 | 0 | 0.2 | 0 |
| $q_2(0)$ | 0.1 | 0 | 0.1 | 0 | 0.1 | 0 |
| $q_3(0)$ | 0.5 | 0 | 0.5 | 0 | 0.5 | 0 |
| $q_4(0)$ | 0.837 | 1 | 0.837 | 0 | 0.837 | 1 |
| $w_x(0)$ | 0.2 | 0 | 0.4 | 0 | -0.4 | -0.2 |
| $w_y(0)$ | 0.1 | 0 | 0.3 | 0 | -0.3 | -0.2 |
| $w_z(0)$ | 0.2 | 0 | 0.4 | 0 | -0.4 | -0.2 |
| $p_x$ | | | 0.2 | 0 | | |
| $p_y$ | | | -0.3 | 0 | | |
| $p_z$ | | | 0.3 | 0 | | |

Table 1: True motion parameters and the initial conditions to start the IEKF in the simulations of constant angular velocity motion and constant precession motion



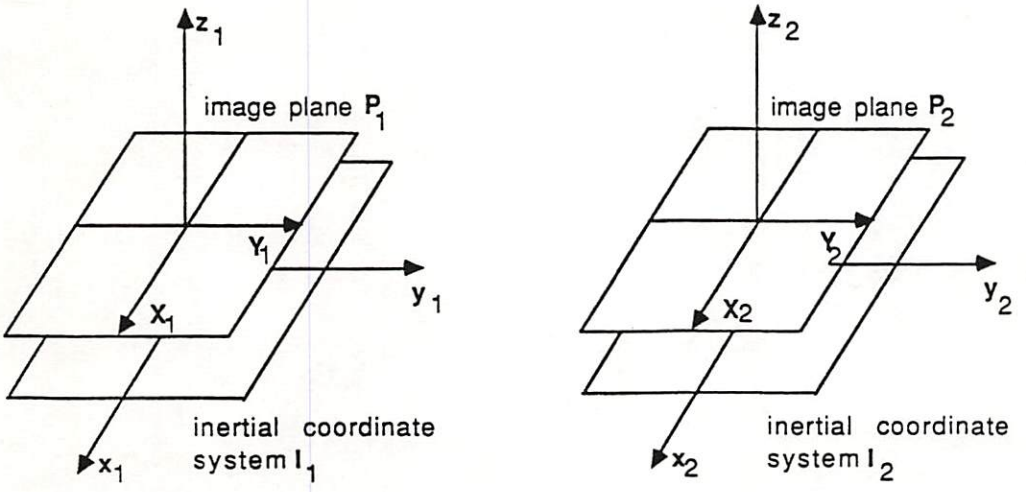




**Figure 1 : Geometry of motion and measurement models**

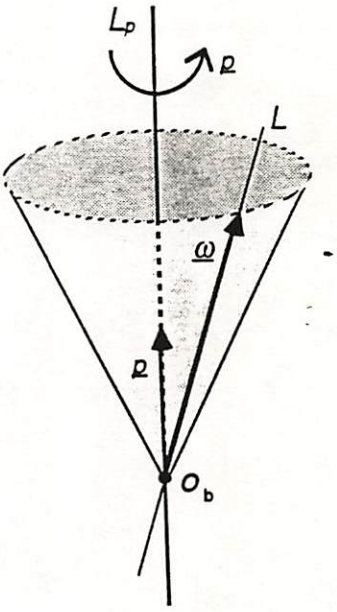



**Figure 2. Geometry of the constant precession model**

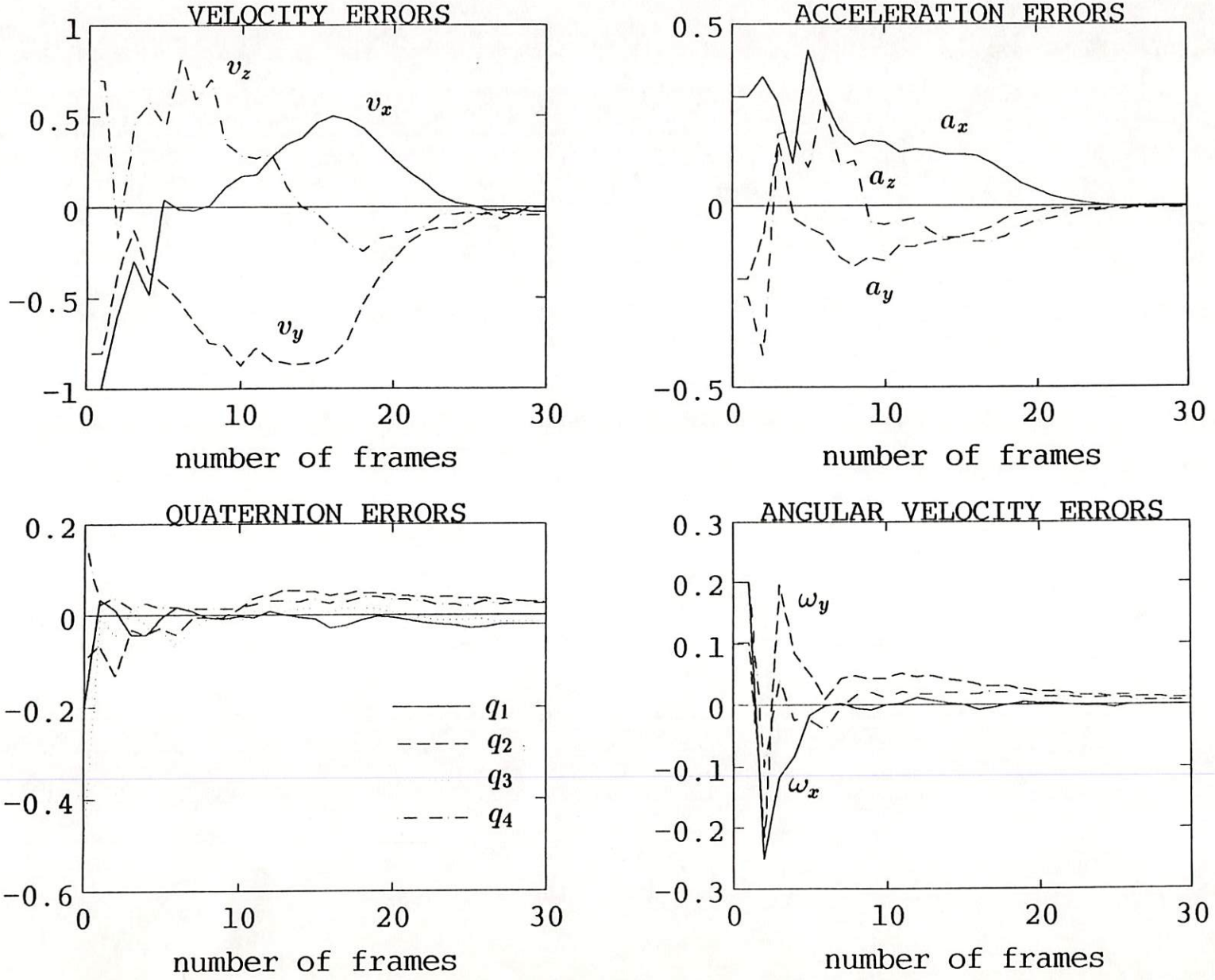
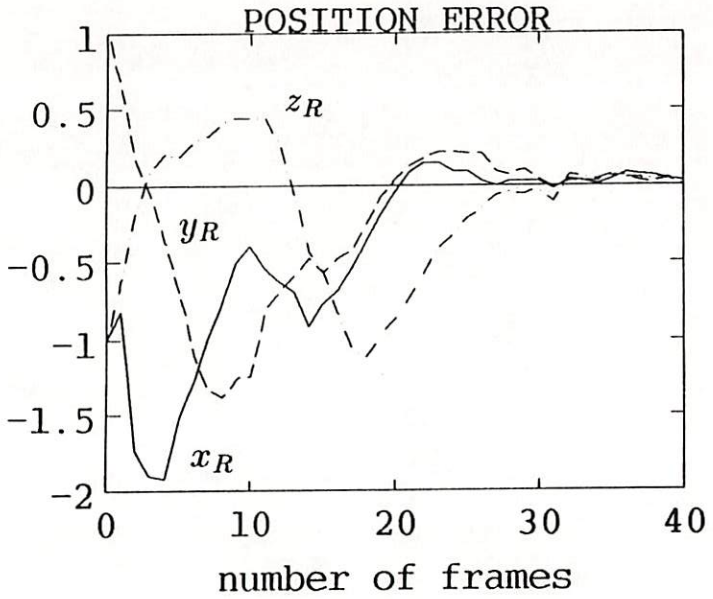
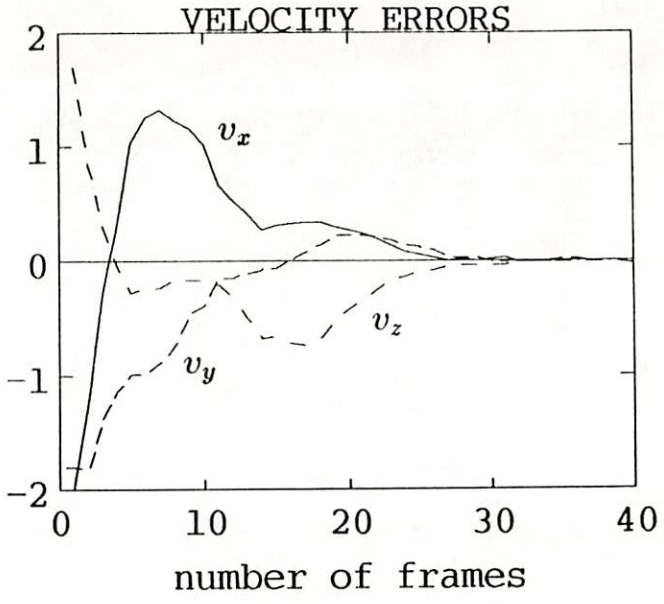
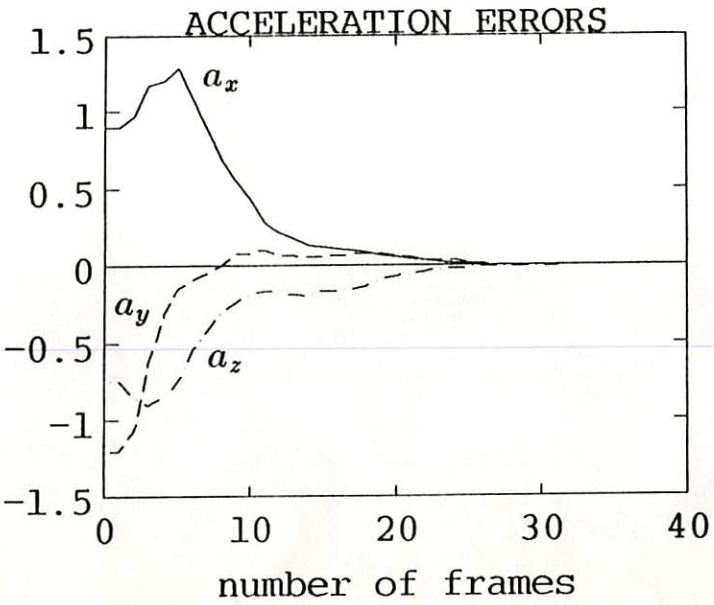



Figure 3: State estimate error in the constant angular velocity motion



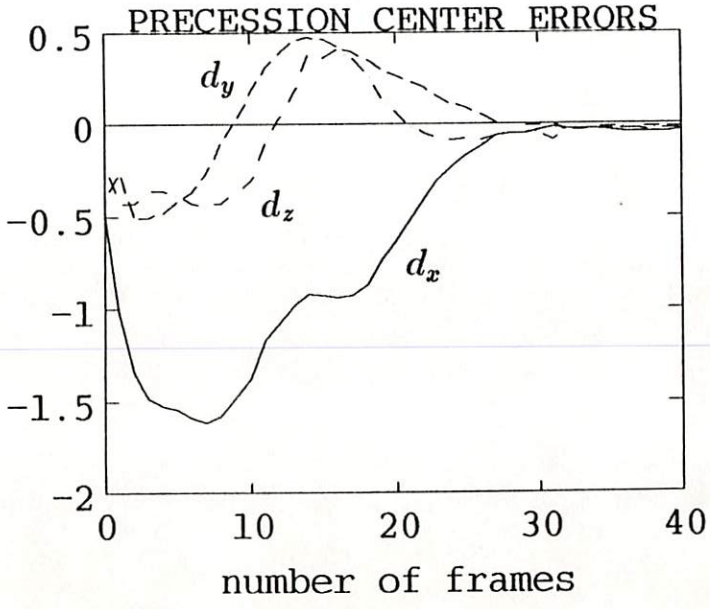
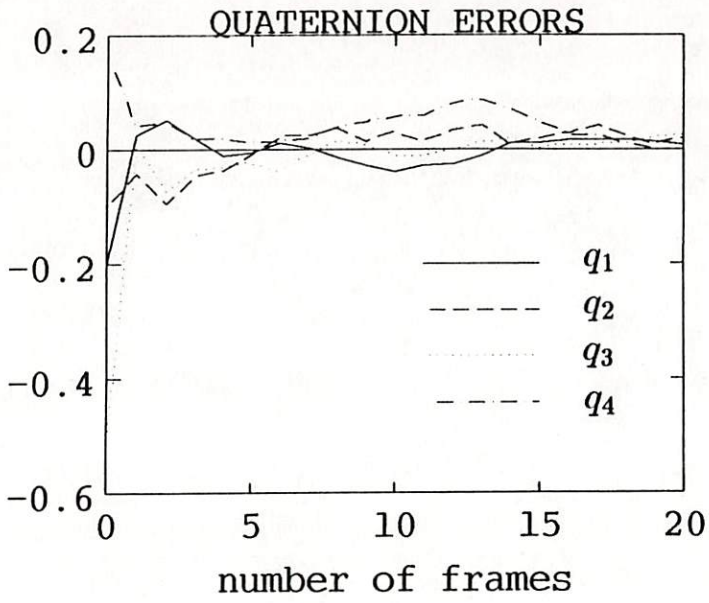
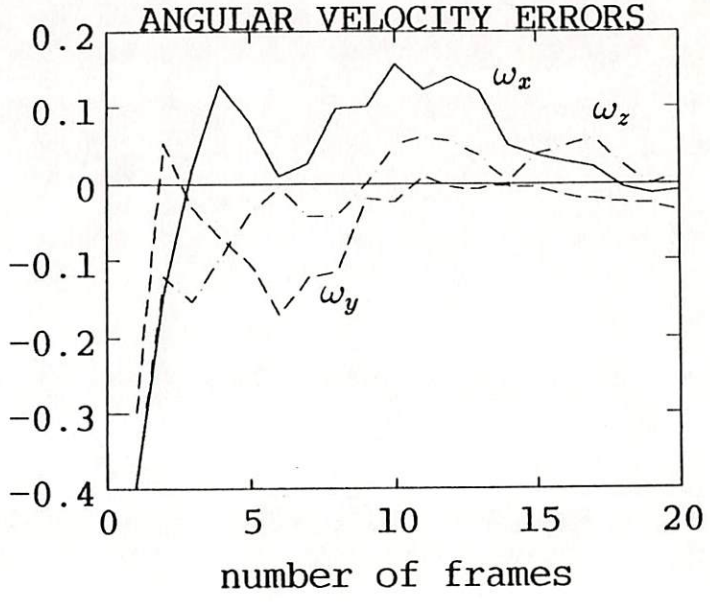




Figure 4: State estimate error in the constant precession motion



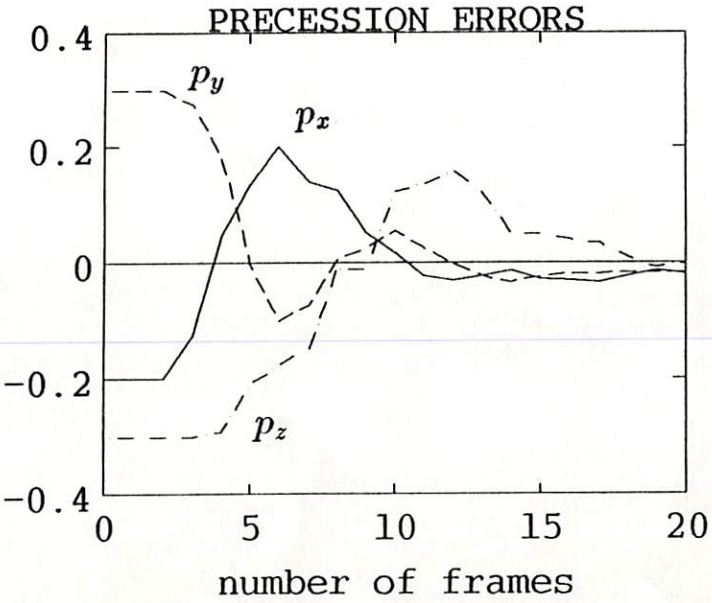




Figure 4: (Continued)

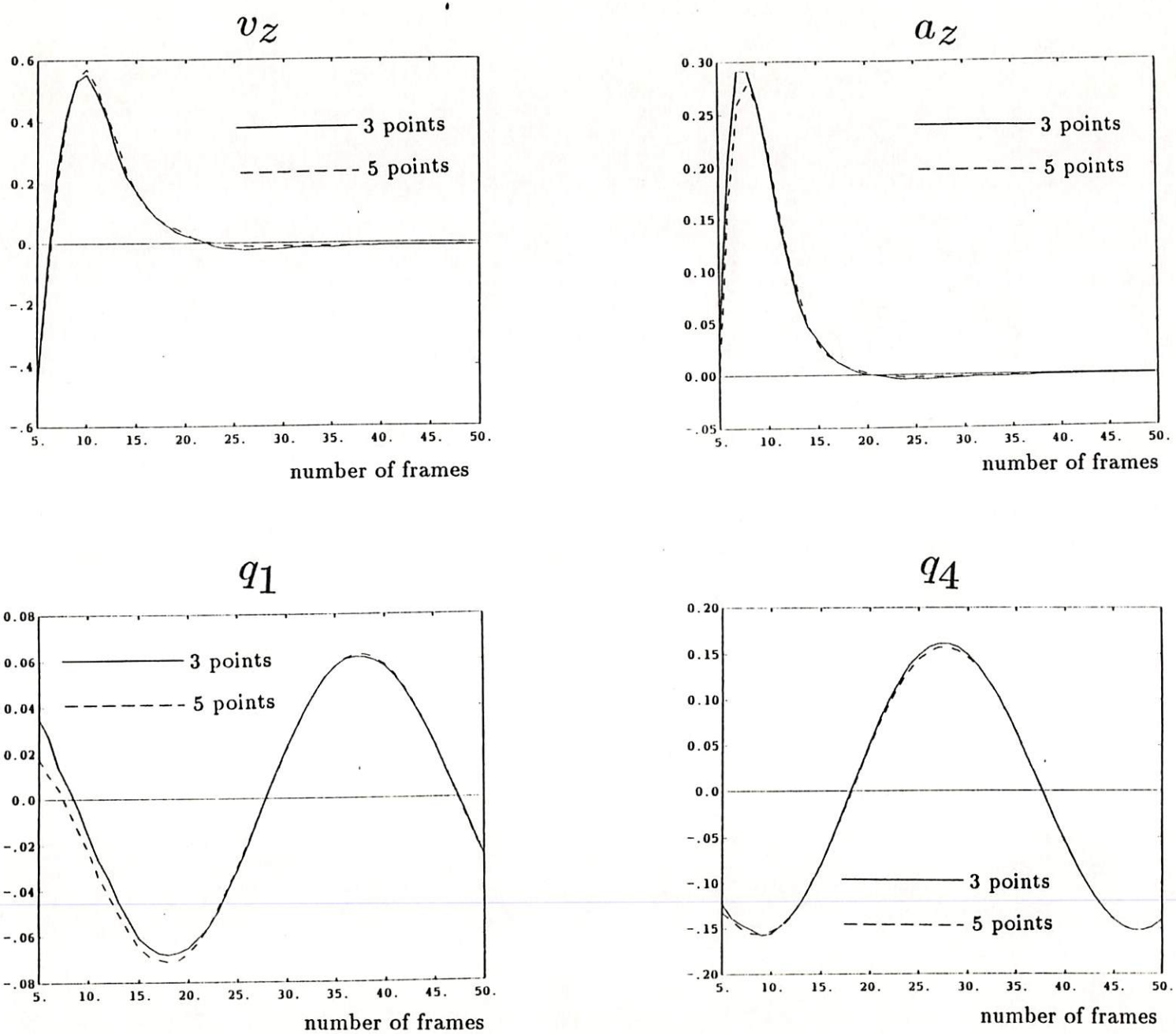



Figure 5: Sample bias from 50 Monte Carlo trials

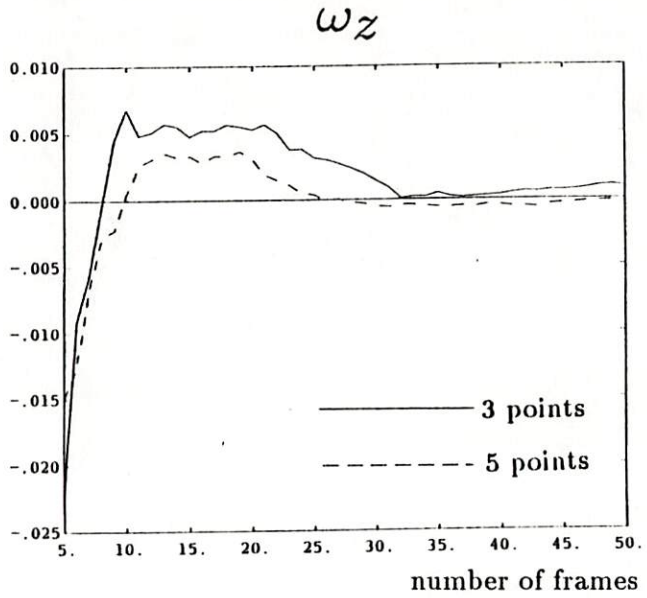
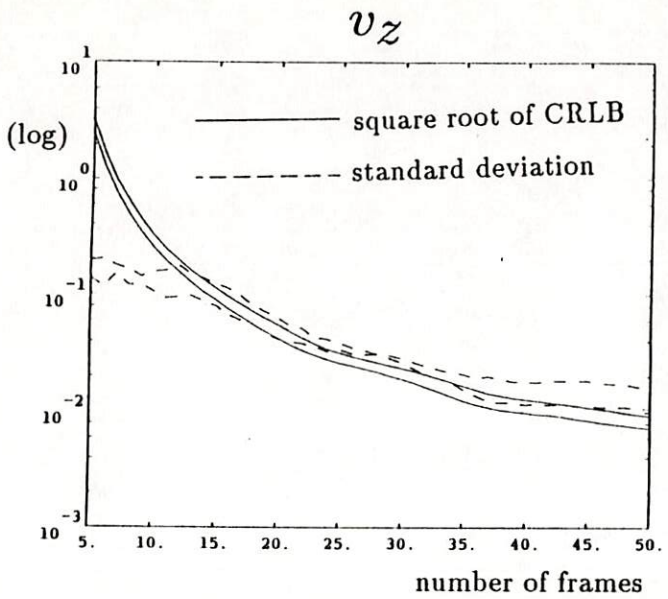
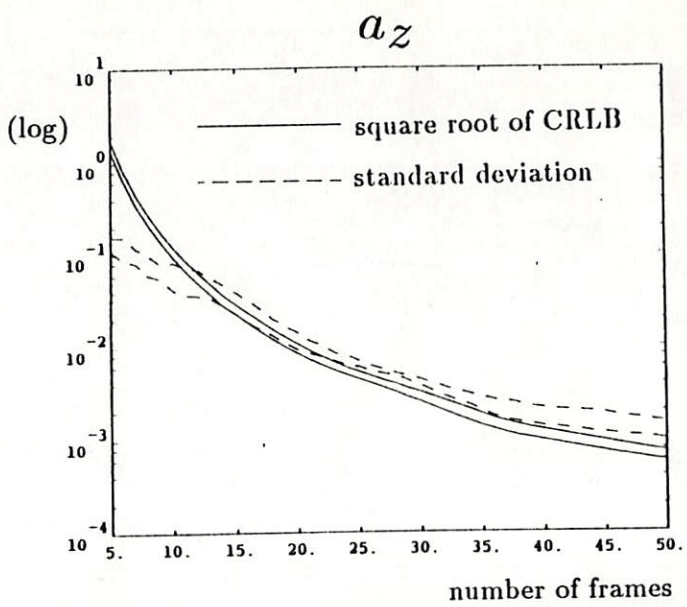
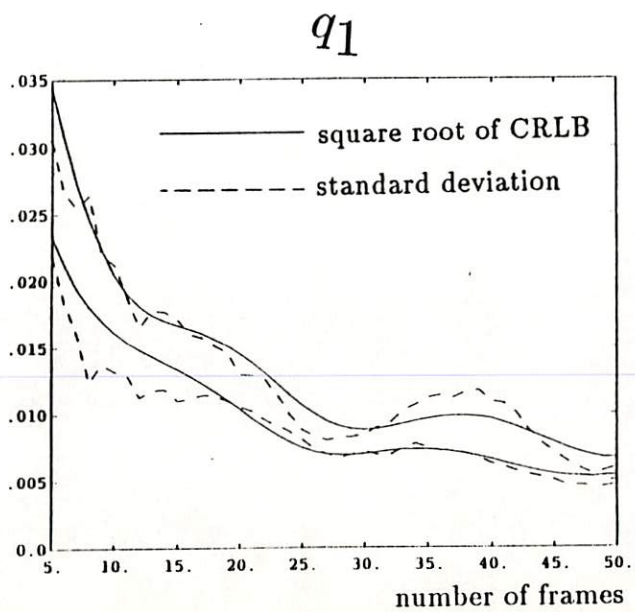



Figure 5: (Continued)



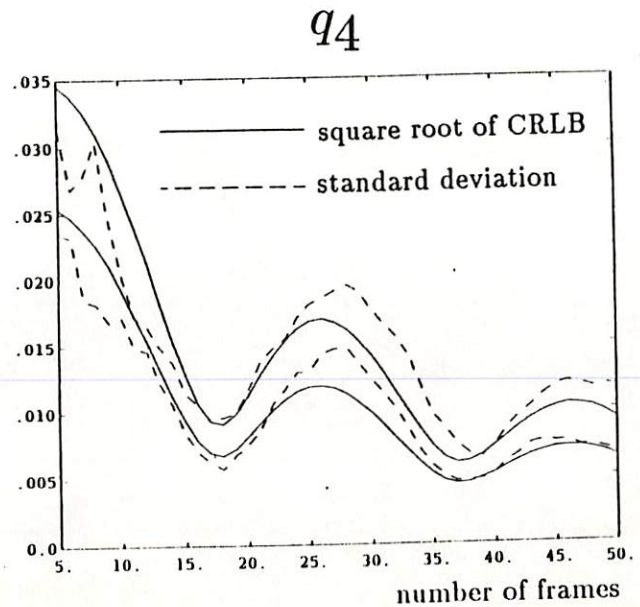




Figure 6: Square root of Cramér-Rao lower bound and sample standard deviation from 50 Monte Carlo trials. Upper curves correspond to 3 feature points. Lower curves correspond to 5 feature points.

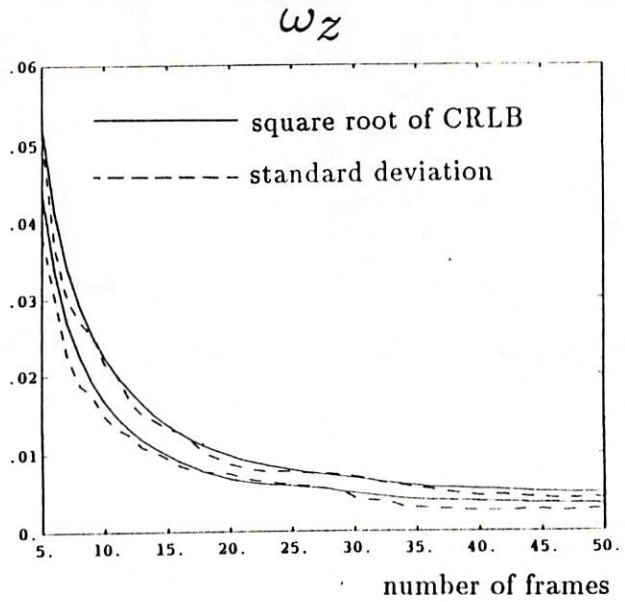



Figure 6: (Continued)