Discrete Nonlinear Observers for Inertial Navigation

Yong Zhao and Jean-Jacques E. Slotine

Nonlinear Systems Laboratory
Massachusetts Institute of Technology
Cambridge, Massachusetts, 02139, USA
yongzhao@mit.edu, jjs@mit.edu

Abstract

We derive an exact deterministic nonlinear observer to compute the continuous state of an inertial navigation system based on partial discrete measurements, the so-called strap-down problem. Nonlinear contraction is used as the main analysis tool, and the hierarchical structure of the system physics is systematically exploited. The paper also discusses the use of nonlinear measurements, such as distances to time-varying reference points.

1 Introduction

This paper derives an exact deterministic nonlinear observer to compute the continuous state of an inertial navigation system based on partial discrete measurements. The main analysis tool is nonlinear contraction theory [9, 10, 12, 11, 14]. Recent work on nonlinear observer design for mechanical systems based on nonlinear contraction theory can be found in [1, 3, 8, 7].

Specifically, we consider the classical strap-down problem in inertial navigation [4, 17], where angular position (Euler angles) $\mathbf{x} = (\psi, \theta, \phi)^{\mathbf{T}}$ and inertial position $\mathbf{r}$ are computed from the body turn rate ω and inertial acceleration γ, measured continuously in intrinsic (body-fixed) coordinates,

$$\begin{cases} \dot{\mathbf{x}} = \mathbf{H^{-1}}\,\omega \\ \dot{\mathbf{v}} = \mathbf{A}\,\gamma \\ \dot{\mathbf{r}} = \mathbf{v} \end{cases} \tag{1}$$

with

$$\mathbf{H} = \begin{bmatrix} 1 & 0 & -\sin\theta \\ 0 & \cos\psi & \cos\theta\sin\psi \\ 0 & -\sin\psi & \cos\theta\cos\psi \end{bmatrix}$$

and

$$\mathbf{A} = \begin{bmatrix} \cos\theta\cos\phi & \sin\psi\sin\theta\cos\phi - \cos\psi\sin\phi & \cos\psi\sin\theta\cos\phi + \sin\psi\sin\phi \\ \cos\theta\sin\phi & \sin\psi\sin\theta\sin\phi + \cos\psi\cos\phi & \cos\psi\sin\theta\sin\phi - \sin\psi\cos\phi \\ -\sin\theta & \cos\theta\sin\psi & \cos\theta\cos\psi \end{bmatrix}$$

As made precise in [12] such a system lies at the boundary between convergence and divergence, much like a triple integrator.

In this paper, the continuous measurements of ω and γ are augmented by *discrete* measurements of $\mathbf{x}$ and $\mathbf{r}$, leading to a globally exponentially convergent nonlinear observer design. Such combinations of measurements are typical in inertial navigation, whether for vehicles or robots (see e.g. [16] for a recent discussion). The human vestibular system also features a similar structure, with otolithic organs measuring linear acceleration and semi-circular canals estimating angular velocity through heavily damped angular acceleration signals, an information then combined with visual data at much slower update rate.

After a brief review of contraction theory, Section 2 introduces the basic observer design. We build simple observers to compute $(\mathbf{x}, \mathbf{v}, \mathbf{r})$ based on partial discrete measurements $\mathbf{x}_i$ and $\mathbf{r}_i$. In Section 3 we discuss extensions, such as the use of nonlinear measurements, and the effects of system disturbance and measurement disturbance [14]. We also study the case where the inertial navigation system is expressed in quaternion form [4, 5, 6]. Section 4 presents simulation results on a 3-dimensional system. Brief concluding remarks are offered in Section 5.

2 Basic Algorithm

In this section, we first briefly review basic results in contraction theory. We then construct a discrete observer for system (1), which consists of a hierarchy of three sub-systems, mirroring the hierarchical nature of systems physics (1).

2.1 Contraction Theory

The basic theorem of contraction analysis [9] can be stated as

Theorem 1 *Consider the deterministic system* $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, t)$ *, where* $\mathbf{f}$ *is a smooth nonlinear function. If there exist a uniformly positive definite metric*

$$\mathbf{M}(\mathbf{x}, t) \ = \ \mathbf{\Theta}(\mathbf{x}, t)^T \ \mathbf{\Theta}(\mathbf{x}, t)$$

such that the associated generalized Jacobian

$$\mathbf{F} \ = \ \left(\dot{\mathbf{\Theta}} + \mathbf{\Theta}\frac{\partial \mathbf{f}}{\partial \mathbf{x}}\right) \mathbf{\Theta}^{-1}$$

is uniformly negative definite, then all system trajectories then converge exponentially to a single trajectory, with convergence rate $|\lambda_{max}|$*, where* λ_{max} *is the largest eigenvalue of the symmetric part of* $\mathbf{F}$*. The system is said to be contracting.*

It can be shown conversely that the existence of a uniformly positive definite metric with respect to which the system is contracting is also a necessary condition for global exponential convergence of trajectories. In the linear time-invariant case, a system is globally contracting if and only if it is strictly stable, with $\mathbf{F}$ simply being a normal Jordan form of the system and $\mathbf{\Theta}$ the coordinate transformation to that form. Furthermore, since

$$\mathbf{\Theta}^{-1} \ \mathbf{F}_s \ \mathbf{\Theta} \ = \ \frac{1}{2} \ \mathbf{M}^{-1} \ (\dot{\mathbf{M}} + \ \mathbf{M} \ \frac{\partial \mathbf{f}}{\partial \mathbf{x}} \ + \ \frac{\partial \mathbf{f}}{\partial \mathbf{x}}^T \mathbf{M})$$

where is $\mathbf{F}_s$ the symmetric part of $\mathbf{F}$, all transformations $\mathbf{\Theta}$ corresponding to the same $\mathbf{M}$ lead to the same eigenvalues for $\mathbf{F}_s$, and therefore to the same contraction rate $|\lambda_{max}|$.

Consider now a hybrid case [11], consisting of a continuous system

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, t)$$

which is switched to a discrete system

$$\mathbf{x}_{i+1} = \mathbf{f}_i(\mathbf{x}_i, i)$$

every Δt_i for one discrete step. Letting, in the *same coordinate system* $\mathbf{\Theta}$, $\bar{\lambda}$ be the largest eigenvalue of the symmetric matrix $\mathbf{F}^T + \mathbf{F}$, and $\bar{\lambda}_i$ be the largest eigenvalue of $\mathbf{F}_i^T \mathbf{F}_i$ (the corresponding discrete-time quantity, where $\mathbf{F}_i = \mathbf{\Theta}_{i+1} \frac{\partial \mathbf{f}_i}{\partial \mathbf{x}_i} \mathbf{\Theta}_i^{-1}$, see [11]), a sufficient condition for the overall system to be contracting is

$$\exists \ \alpha < 1, \forall i, \quad 0 \leq \bar{\lambda}_i e^{\bar{\lambda} \Delta t_i} \leq \alpha \tag{2}$$

Contraction theory proofs and this paper make extensive use of *virtual displacements*, which are differential displacements at fixed time borrowed from mathematical physics and optimization theory. Formally, if we view the position of the system at time t as a smooth function of the initial condition $\mathbf{x}_o$ and of time, $\mathbf{x} = \mathbf{x}(\mathbf{x}_o, t)$, then $\delta\mathbf{x} = \frac{\partial \mathbf{x}}{\partial \mathbf{x}_o}\, d\mathbf{x}_o$.

2.2 A Basic Algorithm

The observer is based on the partial measurements of the state **x** and **r** at a series of instants $\{t_i\}$.

First, based on the discrete measurement $\mathbf{x}_i$, compute **x** with the observer

$$\begin{cases} \dot{\hat{\mathbf{x}}} = \mathbf{H}^{-1}(\hat{\mathbf{x}})\,\omega \\ \\ \hat{\mathbf{x}}_i^+ = k_{1i}\,\hat{\mathbf{x}}_i^- + (1 - k_{1i})\,\mathbf{x}_i \end{cases} \tag{3}$$

where the first equation describes a continuous update between measurements, and the second equation a discrete measurement incorporation.

Computing virtual displacements in (3) leads to

$$\begin{cases} \delta\dot{\hat{\mathbf{x}}} = \frac{\partial(\mathbf{H}^{-1}\omega)}{\partial\hat{\mathbf{x}}}\,\delta\hat{\mathbf{x}} \\ \\ \delta\hat{\mathbf{x}}_i^+ = k_{1i}\,\delta\hat{\mathbf{x}}_i^- \end{cases} \tag{4}$$

Based on [12], define $\delta\hat{\mathbf{z}} = \mathbf{\Theta}\,\delta\hat{\mathbf{x}}$ with $\mathbf{\Theta}(\hat{\mathbf{x}}, t) = \mathbf{A}\mathbf{H}$. This implies that $\begin{cases} \delta\hat{\mathbf{z}}_i^+ = \mathbf{\Theta}_i^+\,\delta\hat{\mathbf{x}}_i^+ \\ \delta\hat{\mathbf{z}}_i^- = \mathbf{\Theta}_i^-\,\delta\hat{\mathbf{x}}_i^- \end{cases}$
From (4), we have

$$\begin{cases} \delta\dot{\hat{\mathbf{z}}} = (\dot{\mathbf{\Theta}} + \mathbf{\Theta}\,\frac{\partial(\mathbf{H}^{-1}\omega)}{\partial\hat{\mathbf{x}}})\mathbf{\Theta}^{-1}\,\delta\hat{\mathbf{z}} = \mathbf{0} \\ \\ \delta\hat{\mathbf{z}}_i^+ = k_{1i}\,(\mathbf{\Theta}_i^+)(\mathbf{\Theta}_i^-)^{-1}\,\delta\hat{\mathbf{z}}_i^- \end{cases}$$

where $\delta\dot{\hat{\mathbf{z}}} = \mathbf{0}$ stems from the indifference property of the system [12].

Note that this indifference property can be understood intuitively from the physical meaning of the transformation $\mathbf{\Theta}$ used to define $\delta\hat{\mathbf{z}}$. Indeed, $\mathbf{\Theta} = \mathbf{A}\mathbf{H}$, where **H** is the transformation matrix from Euler angles to body-fixed coordinates and **A** is the transformation matrix from body-fixed coordinates to inertial coordinates. Thus, $\delta\hat{\mathbf{z}}$ simply describes the virtual system in inertial coordinates. Also note that the indifference property is actually immediate in a quaternion representation, as we shall discuss in Section 3.4.

From hybrid contraction condition (2) in Section 2.1, if

$$\bar{\lambda}_{1i}\, e^{0 \cdot \Delta t_i} = \bar{\lambda}_{1i} < 1 \qquad \text{uniformly} \tag{5}$$

where $\bar{\lambda}_{1i} = \lambda_{max}(\mathbf{F}_{1i}^T \mathbf{F}_{1i})$ and $\mathbf{F}_{1i} = k_{1i}\,(\mathbf{\Theta}_i^+)(\mathbf{\Theta}_i^-)^{-1}$, then both $\delta\hat{\mathbf{z}}$ and $\delta\hat{\mathbf{x}}$ tend to zero exponentially. So $\hat{\mathbf{x}}$ tends to $\mathbf{x}$ exponentially.

Second, based on the discrete measurement of $\mathbf{r}$, compute $\mathbf{v}$ with the observer

$$\begin{cases} \dot{\hat{\mathbf{v}}} = \mathbf{A}(\hat{\mathbf{x}})\,\gamma \\ \\ \hat{\mathbf{v}}_{i+1}^+ = \hat{\mathbf{v}}_{i+1}^- - \frac{1}{\Delta t_i}\int_{t_i}^{t_{i+1}} \hat{\mathbf{v}} dt + \frac{1}{\Delta t_i}\,(\mathbf{r}_{i+1} - \mathbf{r}_i) \end{cases} \tag{6}$$

From (6) and the first step, we get

$$\begin{cases} \frac{d}{dt}(\delta\hat{\mathbf{v}}) = \frac{\partial(\mathbf{A}\gamma)}{\partial\hat{\mathbf{x}}}\,\delta\hat{\mathbf{x}} \rightarrow \mathbf{0} \\ \\ \delta\hat{\mathbf{v}}_{i+1}^+ = \delta\hat{\mathbf{v}}_{i+1}^- - \frac{1}{\Delta t_i}\int_{t_i}^{t_{i+1}} \delta\hat{\mathbf{v}}\,dt \end{cases} \tag{7}$$

Since $\delta\hat{\mathbf{v}}$ tends exponentially to a constant, we have

$$\frac{1}{\Delta t_i}\int_{t_i}^{t_{i+1}} \delta\hat{\mathbf{v}}\,dt \rightarrow \frac{1}{\Delta t_i}\,(\delta\hat{\mathbf{v}}_{i+1}^-\,\Delta t_i) = \delta\hat{\mathbf{v}}_{i+1}^-$$

Using (7), this implies that $\delta\hat{\mathbf{v}}_{i+1}^+ \rightarrow \mathbf{0}$, which by continuity implies that the constant which $\delta\hat{\mathbf{v}}$ tends to must be zero. We thus have, exponentially,

$$\begin{cases} \delta\hat{\mathbf{v}} \rightarrow \mathbf{0} \\ \delta\hat{\mathbf{v}}_{i+1}^+ \rightarrow \mathbf{0} \end{cases}$$

Since by design $\hat{\mathbf{v}} = \mathbf{v}$ is a particular solution of (6), this implies that $\hat{\mathbf{v}}$ tends to $\mathbf{v}$ exponentially.

Third, based on the discrete measurement $\mathbf{r}_i$, use the observer

$$\begin{cases} \dot{\hat{\mathbf{r}}} = \hat{\mathbf{v}} \\ \\ \hat{\mathbf{r}}_i^+ = \mathbf{F}_{3i}\,\hat{\mathbf{r}}_i^- + (\mathbf{I} - \mathbf{F}_{3i})\,\mathbf{r}_i \end{cases} \tag{8}$$

Since we know $\delta\hat{\mathbf{v}}$ tends to zero exponentially, we have

$$\begin{cases} \frac{d}{dt}(\delta\hat{\mathbf{r}}) = \delta\hat{\mathbf{v}} \rightarrow \mathbf{0} \\ \\ \delta\hat{\mathbf{r}}_i^+ = \mathbf{F}_{3i}\,\delta\hat{\mathbf{r}}_i^- \end{cases}$$

If $\bar{\lambda}_{3i} \; < \; 1$, i.e.

$$\bar{\lambda}_{3i} \, e^{0 \cdot \Delta t_i} < 1 \qquad \text{uniformly} \tag{9}$$

where $\bar{\lambda}_{3i}$ is the largest eigenvalue of $\mathbf{F}_{3i}^T \mathbf{F}_{3i}$. So $\hat{\mathbf{r}}$ tends to $\mathbf{r}$ exponentially.

Extension 1: When we compute **v** and **r**, we only use the discrete-time measurement $\mathbf{r}_i$ without $\mathbf{x}_i$. This allows $\mathbf{x}_i$ and $\mathbf{r}_i$ to be measured at different instants, with the same computation.

Extension 2: The metric can also be written $\mathbf{\Theta}^T \mathbf{\Theta} \; = \; (\mathbf{AH})^T (\mathbf{AH}) \; = \; \mathbf{H}^T \mathbf{H}$ since $\mathbf{A}$ is orthogonal. So we can simply use $\mathbf{\Theta} = \mathbf{H}$.

Extension 3: Assume that in (3) we replace the discrete update law by the more general

$$\hat{\mathbf{x}}_i^+ = \mathbf{F}_{1i} \, \hat{\mathbf{x}}_i^- + (\mathbf{I} - \mathbf{F}_{1i}) \, \mathbf{x}_i$$

where $\mathbf{\Theta}_i$ and $\mathbf{F}_{1i}$ commute. Then

$$\begin{cases} \delta \dot{\hat{\mathbf{z}}} = (\dot{\mathbf{\Theta}} + \mathbf{\Theta} \, \frac{\partial (\mathbf{H^{-1}} \omega)}{\partial \hat{\mathbf{x}}}) \mathbf{\Theta^{-1}} \, \delta \hat{\mathbf{z}} = \mathbf{0} \\ \\ \delta \hat{\mathbf{z}}_i^+ = (\mathbf{\Theta}_i^+) \mathbf{F}_{1i} (\mathbf{\Theta}_i^-)^{-1} \, \delta \hat{\mathbf{z}}_i^- \end{cases}$$

The hybrid contraction condition (5) becomes

$$\bar{\lambda}_{1i} \, e^{0 \cdot \Delta t_i} = \bar{\lambda}_{1i} < 1 \qquad \text{uniformly}$$

where $\bar{\lambda}_{1i}$ is the largest eigenvalue of $[(\mathbf{\Theta}_i^+) \mathbf{F}_{1i} (\mathbf{\Theta}_i^-)^{-1}]^T \; [(\mathbf{\Theta}_i^+) \mathbf{F}_{1i} (\mathbf{\Theta}_i^-)^{-1}]$.

Note that because the generalized Jacobians are zero at each step of the hierarchy, the hybrid contraction conditions simply define the *metrics* in which the discrete measurement incorporation steps should be contracting. As we shall see later, the flexibility offered within this constraint will allow us to trade-off model error vs measurement error, similarly in spirit to a standard Kalman filter.

3 Extensions of the Basic Algorithm

Discussions about full discrete measurements, disturbance effects, nonlinear measurements, and quaternion representation are offered in this section. An observer based on full measurement is described in Section 3.1. Effects of system disturbance and measurement disturbance are discussed in Section 3.2. Section 3.3 we develop a more general discrete observer applicable to nonlinear measurements. Use of quaternions is studied in Section 3.4.

3.1 Computation with Full Discrete Measurement

Assume that *all* states $\mathbf{x}$, $\mathbf{v}$, and $\mathbf{r}$ are actually measured, at a series of discrete instants $\{t_i\}$. Then steps 1 and 3 are unchanged, but we can replace step 2 (the estimation of $\mathbf{v}$) by the observer

$$\begin{cases} \dot{\hat{\mathbf{v}}} = \mathbf{A}(\hat{\mathbf{x}})\,\gamma \\ \\ \hat{\mathbf{v}}_i^+ = \mathbf{F}_{2i}\,\hat{\mathbf{v}}_i^- + (\mathbf{I} - \mathbf{F}_{2i})\,\mathbf{v}_i \end{cases}$$

Since we know $\delta\hat{\mathbf{x}}$ tends to zero exponentially, we have

$$\begin{cases} \frac{d}{dt}(\delta\hat{\mathbf{v}}) = \frac{\partial(\mathbf{A}\gamma)}{\partial\hat{\mathbf{x}}}\,\delta\hat{\mathbf{x}} \to \mathbf{0} \\ \\ \delta\hat{\mathbf{v}}_i^+ = \mathbf{F}_{2i}\,\delta\hat{\mathbf{v}}_i^- \end{cases}$$

With $\bar{\lambda}_{2i} \ < \ 1$, we have

$$\bar{\lambda}_{2i}\,e^{0\cdot\Delta t_i} < 1 \qquad \text{uniformly}$$

where $\bar{\lambda}_{2i}$ is the largest eigenvalue of $\mathbf{F}_{2i}^T\mathbf{F}_{2i}$. So $\hat{\mathbf{v}}$ tends to $\mathbf{v}$ exponentially.

Note that in some cases one only needs to estimate orientation $\mathbf{x}$ and velocity $\mathbf{v}$, and that the discrete measurement of $\mathbf{v}$ may be obtained from optical flow, which can be computationally ”expensive” and thus infrequent.

3.2 Disturbance Effects

Effects of bounded inputs and measurement disturbances can be quantified and observer gains chosen accordingly.

Consider input disturbance $\mathbf{d}$ and measurement disturbance $\mathbf{n}$, with $\|\mathbf{d}\| \leq D$ and $\|\mathbf{n}\| \leq N$, leading to the modified system

$$\begin{cases} \dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \mathbf{d} \\ \mathbf{x}_i^{measure} = \mathbf{x}_i + \mathbf{n} \end{cases}$$

Using the basic robustness result in [9, 14], we can quantify the corresponding quadratic bounds R on the estimation error

$$R^{new} = \mid k_j \mid \ e^{\bar{\lambda}\ \Delta t_i}\,R^{old} + \mid k_j \mid \frac{D}{\bar{\lambda}}(e^{\bar{\lambda}\ \Delta t_i} - 1) + \mid k_j - 1 \mid N$$

where $\bar{\lambda}$ is the largest eigenvalue of the symmetric part of $\int_0^1 \frac{\partial \mathbf{f}}{\partial \mathbf{x}} (\hat{\mathbf{x}} + c(\mathbf{x} - \hat{\mathbf{x}}))\, dc$.

Define the objective function ($0 \leq k_j < 1$)

$$\begin{aligned} F(k_j) &= \mid k_j \mid \ e^{\bar{\lambda}\ \Delta t_i}\ R^{old} + \mid k_j \mid \frac{D}{\bar{\lambda}}(e^{\bar{\lambda}\ \Delta t_i} - 1) + \mid k_j - 1 \mid N \\ &= k_j\, e^{\bar{\lambda}\ \Delta t_i}\ R^{old} + k_j \frac{D}{\bar{\lambda}}(e^{\bar{\lambda}\ \Delta t_i} - 1) + (1 - k_j)N \end{aligned}$$

Then, $F(k_j) = (A + B - N)k_j + N$, where $A = e^{\bar{\lambda}\ \Delta t_i}\ R^{old}$ and $B = (e^{\bar{\lambda}\ \Delta t_i} - 1)D/\bar{\lambda}$.

We know k_j should also satisfy

$$k_j e^{\bar{\lambda} \Delta t_i} < 1 \qquad \text{uniformly}$$

Define k_{max} as an upper bound of k_j. Therefore,

$$0 \leq k_j \leq k_m$$

where $k_m = min(k_{max}, 1)$. Finally, we obtain the minimum of $F(k_j)$

$$F_{min} = \begin{cases} N, \ \text{ when } \ k_j = 0 & \text{if } \ A + B - N > 0 \\ (A + B - N)k_m + N, \ \text{ when } \ k_j = k_m & \text{if } \ A + B - N < 0 \\ N, \ \text{ when } \ 0 \leq k_j \leq k_m & \text{if } \ A + B - N = 0 \end{cases}$$

where $A = e^{\bar{\lambda}\ \Delta t_i}\ R^{old}$ and $B = (e^{\bar{\lambda}\ \Delta t_i} - 1)D/\bar{\lambda}$.

When different measurements are available, the above formulas can also be used to select *a priori* the most informative measurement. This can be the case for instance for selecting the direction of gaze of the eyes in hopping robot [15]. This can also be the case when the measurements are ”expensive”, for instance computationally.

Extension: The discussions above will still work when the bounds of input disturbance and measurement disturbance are time-varying. If $\|\mathbf{d}\| \leq D_i$ and $\|\mathbf{n}\| \leq N_i$ when $t \in [t_i, t_{i+1})$. Similar to the above, we have

$$F_{min} = \begin{cases} N_i, \ \text{ when } \ k_j = 0 & \text{if } \ A + B_i - N_i > 0 \\ (A + B_i - N_i)k_m + N_i, \ \text{ when } \ k_j = k_m & \text{if } \ A + B_i - N_i < 0 \\ N_i, \ \text{ when } \ 0 \leq k_j \leq k_m & \text{if } \ A + B_i - N_i = 0 \end{cases}$$

where $A = e^{\bar{\lambda}\ \Delta t_i}\ R^{old}$ and $B_i = (e^{\bar{\lambda}\ \Delta t_i} - 1)D_i/\bar{\lambda}$.

3.3 Nonlinear measurements

For the system $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$, consider the observer

$$\begin{cases} \dot{\hat{\mathbf{x}}} = \mathbf{f}(\hat{\mathbf{x}}) \\ \hat{\mathbf{x}}_i^+ = \hat{\mathbf{x}}_i^- + \mathbf{g}_i(\hat{\mathbf{y}}_i^-) - \mathbf{g}_i(\mathbf{y}_i) \end{cases} \tag{10}$$

where

$$\begin{cases} \mathbf{y}_i = \mathbf{y}_i(\mathbf{x}_i) \\ \hat{\mathbf{y}}_i^- = \mathbf{y}_i(\hat{\mathbf{x}}_i^-) \end{cases}$$

We have

$$\begin{cases} \delta\dot{\hat{\mathbf{x}}} = \frac{\partial \mathbf{f}}{\partial \hat{\mathbf{x}}}\,\delta\hat{\mathbf{x}} \\ \\ \delta\hat{\mathbf{x}}_i^+ = (\mathbf{I} + \frac{\partial \mathbf{g}_i}{\partial \hat{\mathbf{y}}_i}\frac{\partial \hat{\mathbf{y}}_i}{\partial \hat{\mathbf{x}}_i})\delta\hat{\mathbf{x}}_i^- \end{cases} \tag{11}$$

Defining $\delta\hat{\mathbf{z}} = \mathbf{\Theta}\,\delta\hat{\mathbf{x}}$, we have $\begin{cases} \delta\hat{\mathbf{z}}_i^+ = \mathbf{\Theta}_i^+\,\delta\hat{\mathbf{x}}_i^+ \\ \delta\hat{\mathbf{z}}_i^- = \mathbf{\Theta}_i^-\,\delta\hat{\mathbf{x}}_i^- \end{cases}$. Using Equation (11) yields

$$\begin{cases} \delta\dot{\hat{\mathbf{z}}} = \mathbf{F}\,\delta\hat{\mathbf{z}} \\ \delta\hat{\mathbf{z}}_i^+ = \mathbf{F}_i\,\delta\hat{\mathbf{z}}_i^- \end{cases}$$

where $\mathbf{F} = (\dot{\mathbf{\Theta}} + \mathbf{\Theta}\,\frac{\partial \mathbf{f}}{\partial \hat{\mathbf{x}}})\mathbf{\Theta}^{-1}$ and $\mathbf{F}_i = (\mathbf{\Theta}_i^+)\,(\mathbf{I} + \frac{\partial \mathbf{g}_i}{\partial \hat{\mathbf{y}}_i}\frac{\partial \hat{\mathbf{y}}_i}{\partial \hat{\mathbf{x}}_i})\,(\mathbf{\Theta}_i^-)^{-1}$. The sufficient contraction condition on hybrid systems can be written

$$\bar{\lambda}_i\, e^{\bar{\lambda}\,\Delta t_i} < 1 \tag{12}$$

where $\bar{\lambda}_i = \lambda_{max}(\mathbf{F}_i^T\mathbf{F}_i)$ and $\bar{\lambda}$ is the largest eigenvalue of the symmetric matrix $\mathbf{F}^T + \mathbf{F}$. If condition (12) is satisfied by an appropriate choice of $\mathbf{g}_i$, then $\hat{\mathbf{x}}$ will tend to $\mathbf{x}$ exponentially.

A a simple illustration, consider using distance measurements instead of direct cartesian position measurements. In the 3-dimensional space, measure the distances from one point $X = (x_1, x_2, x_3)^T$ to four time-varying reference points $A = [a_1(t), a_2(t), a_3(t)]^T$, $B = [b_1(t), b_2(t), b_3(t)]^T$, $C = [c_1(t), c_2(t), c_3(t)]^T$, and $D = [d_1(t), d_2(t), d_3(t)]^T$, $D = (d_1, d_2, d_3)^T$.

$$\begin{aligned} y_1 &= \mid XA \mid = \sqrt{(x_1 - a_1)^2 + (x_2 - a_2)^2 + (x_3 - a_3)^2} \\ y_2 &= \mid XB \mid = \sqrt{(x_1 - b_1)^2 + (x_2 - b_2)^2 + (x_3 - b_3)^2} \\ y_3 &= \mid XC \mid = \sqrt{(x_1 - c_1)^2 + (x_2 - c_2)^2 + (x_3 - c_3)^2} \\ y_4 &= \mid XD \mid = \sqrt{(x_1 - d_1)^2 + (x_2 - d_2)^2 + (x_3 - d_3)^2} \end{aligned} \tag{13}$$

The discrete-update part of observer (10) can be built up as below,

$$\begin{bmatrix} \hat{x}^+_{1,i} \\ \hat{x}^+_{2,i} \\ \hat{x}^+_{3,i} \end{bmatrix} = \begin{bmatrix} \hat{x}^-_{1,i} \\ \hat{x}^-_{2,i} \\ \hat{x}^-_{3,i} \end{bmatrix} - \frac{1}{2}\mathbf{K}_i \begin{bmatrix} (\hat{y}^-_{1,i})^2 - (\hat{y}^-_{2,i})^2 - (y^2_{1,i} - y^2_{2,i}) \\ (\hat{y}^-_{2,i})^2 - (\hat{y}^-_{3,i})^2 - (y^2_{2,i} - y^2_{3,i}) \\ (\hat{y}^-_{3,i})^2 - (\hat{y}^-_{4,i})^2 - (y^2_{3,i} - y^2_{4,i}) \end{bmatrix} \tag{14}$$

where $\mathbf{K}_i$ is a 3 by 3 time-varying gain matrix. Using equation (13) yields

$$\delta\hat{\mathbf{x}}^+_i = (\mathbf{I} - \mathbf{K}_i\, \mathbf{J}_i)\delta\hat{\mathbf{x}}^-_i \tag{15}$$

$$\text{where } \mathbf{J}_i = \begin{bmatrix} (b_{1i} - a_{1i}) & (b_{2i} - a_{2i}) & (b_{3i} - a_{3i}) \\ (c_{1i} - b_{1i}) & (c_{2i} - b_{2i}) & (c_{3i} - b_{3i}) \\ (d_{1i} - c_{1i}) & (d_{2i} - c_{2i}) & (d_{3i} - c_{3i}) \end{bmatrix}$$

where subscript i refers to the value at time t_i.

Assume $\mathbf{J}_i$ is non-singular. Then we can choose

$$\mathbf{K}_i = k_i\, \mathbf{J}_i^{-1} \tag{16}$$

With Equation (15), we have

$$\delta\hat{\mathbf{x}}^+_i = (1 - k_i)\delta\hat{\mathbf{x}}^-_i$$

By choosing k_i, we can make $\bar{\lambda}_i$ satisfy the following contraction condition that makes $\delta\hat{\mathbf{z}}$ tends to zero exponentially.

$$\bar{\lambda}_i\, e^{\bar{\lambda}\ \Delta t_i} < 1 \tag{17}$$

where $\bar{\lambda}_i = (1 - k_i)^2$ and $\bar{\lambda}$ is the largest eigenvalue of the symmetric matrix $\mathbf{F}^T + \mathbf{F}$. Therefore, $\delta\hat{\mathbf{x}}$ will tend to zero, and $\hat{\mathbf{x}}$ will tend to $\mathbf{x}$ exponentially.

Remark When $\mathbf{J}_i$ is singular, one has

$$\begin{vmatrix} (b_{1i} - a_{1i}) & (b_{2i} - a_{2i}) & (b_{3i} - a_{3i}) \\ (c_{1i} - b_{1i}) & (c_{2i} - b_{2i}) & (c_{3i} - b_{3i}) \\ (d_{1i} - c_{1i}) & (d_{2i} - c_{2i}) & (d_{3i} - c_{3i}) \end{vmatrix} = 0 \tag{18}$$

Equation (18) is equivalent to

$$[(b_{1i} - a_{1i})\mathbf{i} + (b_{2i} - a_{2i})\mathbf{j} + (b_{3i} - a_{3i})\mathbf{k}] \cdot \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ (c_{1i} - b_{1i}) & (c_{2i} - b_{2i}) & (c_{3i} - b_{3i}) \\ (d_{1i} - c_{1i}) & (d_{2i} - c_{2i}) & (d_{3i} - c_{3i}) \end{vmatrix} = 0 \tag{19}$$

which we can write

$$\overrightarrow{AB} \cdot (\overrightarrow{BC} \times \overrightarrow{CD}) = 0$$

This means that points A, B, C, and D are in the same plane, and therefore that the geometry does not contain enough information to infer position.

To compute velocity, one can rewrite observer (6) as

$$\begin{cases} \dot{\hat{\mathbf{v}}} = \mathbf{A}(\hat{\mathbf{x}})\,\gamma \\ \hat{\mathbf{v}}^+_{i+1} = \hat{\mathbf{v}}^-_{i+1} - \frac{1}{2}\frac{1}{\Delta t_i}\{\mathbf{K}_{i+1}\begin{bmatrix} (\hat{y}^-_{1,i+1})^2 - (\hat{y}^-_{2,i+1})^2 - (y^2_{1,i+1} - y^2_{2,i+1}) \\ (\hat{y}^-_{2,i+1})^2 - (\hat{y}^-_{3,i+1})^2 - (y^2_{2,i+1} - y^2_{3,i+1}) \\ (\hat{y}^-_{3,i+1})^2 - (\hat{y}^-_{4,i+1})^2 - (y^2_{3,i+1} - y^2_{4,i+1}) \end{bmatrix} - \mathbf{K}_i\begin{bmatrix} (\hat{y}^+_{1,i})^2 - (\hat{y}^+_{2,i})^2 - (y^2_{1,i} - y^2_{2,i}) \\ (\hat{y}^+_{2,i})^2 - (\hat{y}^+_{3,i})^2 - (y^2_{2,i} - y^2_{3,i}) \\ (\hat{y}^+_{3,i})^2 - (\hat{y}^+_{4,i})^2 - (y^2_{3,i} - y^2_{4,i}) \end{bmatrix}\} \end{cases} \tag{20}$$

where

$$\begin{aligned} y_1 &= \mid \mathbf{r}A \mid = \sqrt{(r_1-a_1)^2 + (r_2-a_2)^2 + (r_3-a_3)^2} \\ y_2 &= \mid \mathbf{r}B \mid = \sqrt{(r_1-b_1)^2 + (r_2-b_2)^2 + (r_3-b_3)^2} \\ y_3 &= \mid \mathbf{r}C \mid = \sqrt{(r_1-c_1)^2 + (r_2-c_2)^2 + (r_3-c_3)^2} \\ y_4 &= \mid \mathbf{r}D \mid = \sqrt{(r_1-d_1)^2 + (r_2-d_2)^2 + (r_3-d_3))^2} \end{aligned}$$

and

$$\mathbf{K}_i = \begin{bmatrix} (b_{1i}-a_{1i}) & (b_{2i}-a_{2i}) & (b_{3i}-a_{3i}) \\ (c_{1i}-b_{1i}) & (c_{2i}-b_{2i}) & (c_{3i}-b_{3i}) \\ (d_{1i}-c_{1i}) & (d_{2i}-c_{2i}) & (d_{3i}-c_{3i}) \end{bmatrix}^{-1} \text{ and } \mathbf{K}_{i+1} = \begin{bmatrix} (b_{1i+1}-a_{1i+1}) & (b_{2i+1}-a_{2i+1}) & (b_{3i+1}-a_{3i+1}) \\ (c_{1i+1}-b_{1i+1}) & (c_{2i+1}-b_{2i+1}) & (c_{3i+1}-b_{3i+1}) \\ (d_{1i+1}-c_{1i+1}) & (d_{2i+1}-c_{2i+1}) & (d_{3i+1}-c_{3i+1}) \end{bmatrix}^{-1} \tag{21}$$

We then have

$$\begin{cases} \frac{d}{dt}(\delta\hat{\mathbf{v}}) = \frac{\partial(\mathbf{A}\gamma)}{\partial\hat{\mathbf{x}}}\,\delta\hat{\mathbf{x}} \to \mathbf{0} \\ \delta\hat{\mathbf{v}}^+_{i+1} = \delta\hat{\mathbf{v}}^-_{i+1} - \frac{1}{\Delta t_i}(\delta\hat{\mathbf{r}}^-_{i+1} - \delta\hat{\mathbf{r}}^+_i) = \delta\hat{\mathbf{v}}^-_{i+1} - \frac{1}{\Delta t_i}\int_{t_i}^{t_{i+1}} \delta\hat{\mathbf{v}}\,dt \end{cases}$$

which is the same as equation (7). Similarly to the second step of Section 2, this shows that $\hat{\mathbf{v}}$ tends to $\mathbf{v}$ exponentially.

Note that the geometry problem of going from distances to positions is solved by a dynamic system, the observer, rather than explicitly at each instant. In general, one may also use linear measurements at some instants and nonlinear ones at others.

Note that if a measurement is delayed, the algorithms work similarly but the actual information is available after the delay (i.e. the measurement is incorporated at some past time and the forward simulation runs instantly to the current time).

Consider now, extending section 3.2, the effect of model and measurement errors. For the modified system,

$$\left\{ \begin{array}{l} \dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \mathbf{d} \\ \mathbf{x}_i^{measure} = \mathbf{x}_i + \mathbf{n} \end{array} \right.$$

with the following nonlinear observer,

$$\left\{ \begin{array}{l} \dot{\hat{\mathbf{x}}} = \mathbf{f}(\hat{\mathbf{x}}) \\ \hat{\mathbf{x}}_i^{+} = \hat{\mathbf{x}}_i^{-} + \mathbf{g}_j(\hat{\mathbf{y}}_i^{-}) - \mathbf{g}_j(\mathbf{y}_i) \end{array} \right.$$

where $\left\{ \begin{array}{l} \mathbf{y}_i = \mathbf{y}_i(\mathbf{x}_i) \\ \hat{\mathbf{y}}_i^{-} = \mathbf{y}_i(\hat{\mathbf{x}}_i^{-}) \end{array} \right.$ and $\left\{ \begin{array}{ll} \mathbf{d} - \text{model error} & \|\mathbf{d}\| < D \\ \mathbf{n} - \text{measurement error} & \|\mathbf{n}\| < N \end{array} \right.$

We know the quadratic bounds R on the estimation error

$$R^{new} = \sqrt{\bar{\lambda}_i}\, e^{\bar{\lambda}\Delta t_i}\, R^{old} + \sqrt{\bar{\lambda}_i}\, \frac{D}{\bar{\lambda}}(e^{\bar{\lambda}\Delta t_i} - 1) + \sqrt{\bar{\lambda}_{ei}}\, N$$

where $\bar{\lambda}_i = \lambda_{max}((\mathbf{I} + \frac{\partial \mathbf{g}_j}{\partial \hat{\mathbf{y}}_i}\frac{\partial \hat{\mathbf{y}}_i}{\partial \hat{\mathbf{x}}_i})^T(\mathbf{I} + \frac{\partial \mathbf{g}_j}{\partial \hat{\mathbf{y}}_i}\frac{\partial \hat{\mathbf{y}}_i}{\partial \hat{\mathbf{x}}_i}))$, $\bar{\lambda}_{ei} = \lambda_{max}((\frac{\partial \mathbf{g}_i}{\partial \mathbf{y}_i})^T(\frac{\partial \mathbf{g}_i}{\partial \mathbf{y}_i}))$, and $\bar{\lambda}$ is the largest eigenvalue of the symmetric part of $\int_0^1 \frac{\partial \mathbf{f}}{\partial \mathbf{x}}\,(\hat{\mathbf{x}} + c(\mathbf{x} - \hat{\mathbf{x}}))\, dc$.

We can choose the most relevant discrete update function $\mathbf{g}_j$ which will best contribute to improving the estimate $\hat{\mathbf{x}}$ (i.e., to minimize R^{new}).

3.4 Quaternion Representation

Angular position can be expressed in quaternion form, avoiding representation singularities [4, 5]. Quaternions express a rotation of angle θ about the unit vector $\mathbf{n}$ as $\mathbf{q} = (\cos(\theta/2), \mathbf{n}\sin(\theta/2))^T$. With $\mathbf{q} = (q_0, q_1, q_2, q_3)^T$ the quaternion vector, this leads to

$$\left\{ \begin{array}{rl} \dot{\mathbf{q}} = & \frac{1}{2}\,\mathbf{\Omega}\,\mathbf{q} \\ \dot{\mathbf{v}} = & \mathbf{A}\,\gamma \\ \dot{\mathbf{r}} = & \mathbf{v} \end{array} \right.$$

where

$$\mathbf{\Omega} = \begin{bmatrix} 0 & -\omega_1 & -\omega_2 & -\omega_3 \\ \omega_1 & 0 & -\omega_3 & \omega_2 \\ \omega_2 & \omega_3 & 0 & -\omega_1 \\ \omega_3 & -\omega_2 & \omega_1 & 0 \end{bmatrix}$$

and

$$\mathbf{A}(\mathbf{q}) = \begin{bmatrix} q_0^2+q_1^2-q_2^2-q_3^2 & 2(q_1q_2-q_0q_3) & 2(q_1q_3+q_0q_2) \\ 2(q_1q_2+q_0q_3) & q_0^2-q_1^2+q_2^2-q_3^2 & 2(q_2q_3-q_0q_1) \\ 2(q_1q_3-q_0q_2) & 2(q_2q_3+q_0q_1) & q_0^2-q_1^2-q_2^2+q_3^2 \end{bmatrix}$$

In this representation, the fact that the dynamics of $\mathbf{q}$ is indifferent is obvious, since $\mathbf{\Omega}$ is skew-symmetric.

The observers can be derived as earlier, simply by replacing (3) by

$$\begin{cases} \dot{\hat{\mathbf{q}}} = \frac{\mathbf{1}}{\mathbf{2}}\,\mathbf{\Omega}\,\mathbf{q} \\ \\ \hat{\mathbf{q}}_i^+ = \mathbf{F}_{1i}\,\hat{\mathbf{q}}_i^- + (\mathbf{I}-\mathbf{F}_{1i})\,\mathbf{q}_i \end{cases}$$

based on the discrete measurements $\mathbf{q}_i$. Computing virtual displacements

$$\begin{cases} \delta\dot{\hat{\mathbf{q}}} = \frac{\mathbf{1}}{\mathbf{2}}\,\mathbf{\Omega}\,\delta\hat{\mathbf{q}} \\ \\ \delta\hat{\mathbf{q}}_i^+ = \mathbf{F}_{1i}\,\delta\hat{\mathbf{q}}_i^- \end{cases}$$

and because the dynamics of $\mathbf{q}$ is indifferent, we only need

$$\bar{\lambda}_{1i}\,e^{0\cdot\Delta t_i} = \bar{\lambda}_{1i} < 1 \qquad \text{uniformly} \tag{22}$$

where $\bar{\lambda}_{1i}$ is the largest eigenvalue of $\mathbf{F}_{1i}^T\mathbf{F}_{1i}$. Under Condition (22), $\delta\hat{\mathbf{q}}$ tends to zero exponentially, and $\hat{\mathbf{q}}$ tends to $\mathbf{q}$ exponentially.

The other two steps are unchanged, with $\mathbf{A}(\hat{\mathbf{x}})$ being replaced by $\mathbf{A}(\hat{\mathbf{q}})$.

All the above variations and extensions can of course be combined.

4 Simulation

In this section, we will do a 3-dimensional simulation about system (1) based on the discrete measurement $\mathbf{x}_i$ and the nonlinear distance measurements $y_{1,i}$, $y_{2,i}$, $y_{3,i}$, and $y_{4,i}$, as in Section 3.3.

Consider System (1) in the 3-dimensional case. Where

$$\omega = \begin{bmatrix} \frac{2+\sin t}{3} \\ \frac{3+\cos t}{5} \\ \frac{2+\sin 2t}{3} \end{bmatrix} \text{ and } \gamma = \begin{bmatrix} \cos(2t) \\ \sin t \\ \frac{1+2\sin t}{3} \end{bmatrix}$$

Four time-varying reference points are chosen as below (all move on circular trajectories),

$$A(a_1,a_2,a_3)\ \begin{cases} a_1^2+a_2^2=1 \\ a_3=0 \end{cases} \qquad B(b_1,b_2,b_3)\ \begin{cases} (b_1-60)^2+b_2^2=1 \\ b_3=0 \end{cases}$$

$$C(c_1,c_2,c_3)\ \begin{cases} (c_2-60)^2+c_3^2=1 \\ c_1=60 \end{cases} \qquad D(d_1,d_2,d_3)\ \begin{cases} (d_1-60)^2+(d_3-60)^2=1 \\ d_2=60 \end{cases}$$

Observer (3) with $\bar{\lambda}_{1i}=1/9$ is used to compute $\mathbf{x}$. Observer (20) with gain (21) is used to compute $\mathbf{v}$. Using observer (10,14) and gain (16), we choose $k_i=\frac{2}{3}$ to satisfy Condition (12), thus we can compute $\mathbf{r}$. Figure 1 shows $(\hat{\mathbf{x}},\hat{\mathbf{v}},\hat{\mathbf{r}})^T$ tends to $(\mathbf{x},\mathbf{v},\mathbf{r})^T$ exponentially.

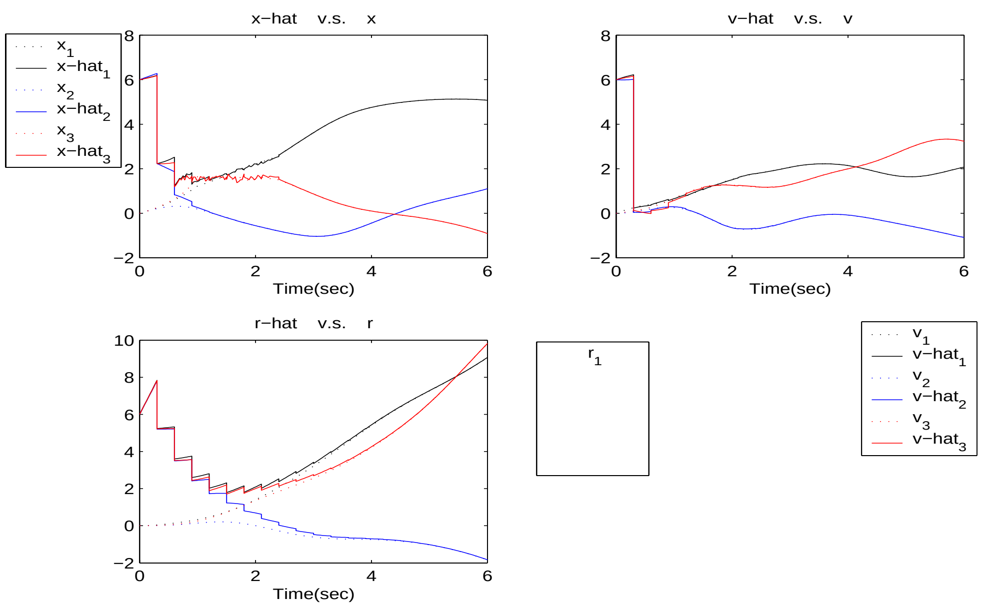

Figure 1: Simulation result of computing $\mathbf{x}$, $\mathbf{v}$, and $\mathbf{r}$ with the discrete measurements $\mathbf{x}_i$ and $\mathbf{r}_i$

5 Concluding Remarks

Observers similar to those developed in this paper can in principle be applied to other continuous nonlinear systems besides inertial navigation systems, although much simplification was afforded by exploiting the hierarchical structure of the system physics. An animation of the basic observer as applied to head stabilization [2] in a simulated robot hopper [15] can also be found in *http://web.mit.edu/nsl/www/hopping_robot.htm*.

Acknowledgement This paper benefited from stimulating discussions with Dr. Agostino Martinelli.

References

[1] Aghannan, N., Rouchon, P., An Intrinsic Observer for a Class of Lagrangian Systems, *I.E.E.E. Trans. Aut. Control*, **48(6)** (2003).

[2] Berthoz, A., The Brain's Sense of Movement, *Harvard University Press, Cambridge, Massachusetts* (2000).

[3] Egeland, O., Kristiansen, E., and Nguyen, T.D., Observers for Euler-Bernoulli Beam with Hydraulic Drive, *I.E.E.E. Conf. Dec. Control* (2001).

[4] Goldstein, H., Classical Mechanics, *Addison-Wesley* (1980).

[5] Grassia, F.S., Practical Parameterization of Rotations Using the Exponential Map, *The Journal of Graphics Tools*, Vol. 3.3 (1998).

[6] Hestenes, D., New Foundations for Classical Mechanics, *Kulwer Academic Publishers, Dordrecht, The Netherlands* (1986).

[7] Jouffroy, J., and J. Opderbecke, Underwater Vehicle Trajectory Estimation Using Contracting PDE-based Observers, *American Control Conference, Boston, Massachusetts* (2004).

[8] Kristiansen, D., and Egeland, O., Time and Spatial Discretization Methods that Preserve Passivity Properties for Systems Described by Partial Differantial Equations, *In Proceedings of the 2000 American Control Conference, Chiago, Illinois* (2000)

[9] Lohmiller, W., and Slotine, J.J.E., On Contraction Analysis for Nonlinear Systems, *Automatica*, **34(6)** (1998).

[10] Lohmiller, W., Contraction Analysis of Nonlinear Systems, *Ph.D. Thesis, Department of Mechanical Engineering, MIT* (1999).

[11] Lohmiller, W., and Slotine J.J.E., Nonlinear Process Control Using Contraction Theory, *A.I.Ch.E. Journal* (2000).

[12] Lohmiller, W., and Slotine J.J.E., Control System Design for Mechanical Systems Using Contraction Theory, *I.E.E.E. Trans. Aut. Control*, **45(5)** (2000).

[13] Slotine, J.J.E., and Li, W., Applied Nonlinear Control, *Prentice-Hall* (1991).

[14] Slotine, J.J.E., and Lohmiller, W., Modularity, Evolution, and the Binding Problem:A View from Stability Theory, *Neural Networks*, **14(2)** (2001).

[15] Raibert, M.H., Legged Robots That Balance, *The M.I.T. Press, Cambridge, Massachusetts* (1986).

[16] Ryu, J., Rossetter E.J., and Gerdes, J.C., Vehicle Sideslip and Roll Parameter Estimation Using GPS. *the Proceedings of AVEC 2002, Hiroshima, Japan* (2002).

[17] Varshalovich, D.A., Moskalev, A.N., and Khersonskii, V.K., Description of Rotation in Terms of the Euler Angles, *Singapore: World Scientific* (1988).