JOURNAL OF GUIDANCE, CONTROL, AND DYNAMICS
Vol. 28, No. 5, September–October 2005

Nonlinear Moving Horizon State Estimation with Continuation/Generalized Minimum Residual Method

Yusuke Soneda*
Keyence Corporation, Osaka 533-8555, Japan
and
Toshiyuki Ohtsuka†
Osaka University, Osaka 565-0871, Japan

We propose a fast algorithm for nonlinear moving horizon state estimation (MHSE). An optimization problem to be solved at each time step is formulated in a deterministic setting to estimate the unknown state and the unknown disturbance over a finite past, which leads to a nonlinear two-point boundary-value problem (TPBVP) to be solved at each time. Instead of solving the nonlinear TPBVP with an iterative method, the estimate is updated by integrating a differential equation to trace the time-varying solution of the TPBVP, which is a kind of continuation method and needs to solve a linear algebraic equation only once at each sampling time. Moreover, the linear algebraic equation involved in the differential equation is solved efficiently by the generalized minimum-residual method, one of the Krylov subspace methods. The proposed algorithm is evaluated by a numerical simulation and experiment with a hovercraft model the dynamics of which are nonlinear. MHSE by the proposed algorithm generates reasonable estimates even when the extended Kalman filter fails, and the proposed algorithm is sufficiently fast for real-time implementation with a sampling period in the order of milliseconds.

Introduction

NONLINEAR state estimation has been an important and challenging problem for many years. Until now, various methods such as the extended Kalman filter (EKF), the statistical approximation approach, and the extended Luenberger observer have been proposed for nonlinear state estimation problems.[1–3] There are also sophisticated methods based on higher-order approximation than the EKF in the stochastic setting.[4,5] However, those methods often result in poor performance because they are formulated with linearization or approximation of nonlinearities. Another approach to state estimation is nonlinear observer design in the deterministic setting via, for example, exact linearization[6–10] or passivation[11] of error dynamics. However, only a restricted class of nonlinear systems satisfies the conditions required for those observer design methods.

In contrast to the preceding methods, nonlinear moving horizon state estimation (MHSE) is formulated for nonlinear systems without linearization and, moreover, can handle constraints. In MHSE, an optimization problem over a finite past is solved at each time step to determine the optimal estimate of the state. As long as the optimization problem can be solved numerically, MHSE is applicable to a wide class of nonlinear systems. Because a nonlinear optimization problem is computationally demanding, nonlinear MHSE has been mainly aimed at application to chemical processes,[12–15] where the sampling interval is sufficiently large for solving the optimization problem at every sampling time. It is still a challenging problem to apply nonlinear MHSE to mechanical systems, where the sampling intervals are much shorter than those in chemical processes, even if the newest computers are used.

Received 12 July 2004; revision received 28 February 2005; accepted for publication 8 March 2005.

*Engineer, FA-IN Products Division, Higashiyodogawa; soneday@keyence.co.jp.

†Associate Professor, Department of Mechanical Engineering, Graduate School of Engineering, Suita; ohtsuka@mech.eng.osaka-u.ac.jp. Senior Member AIAA.

Because the optimization problem in MHSE and its solution depend on time, it is a natural idea to trace the infinitesimal variation of the optimal solution with respect to time for avoiding computationally demanding iterative optimization methods, which is a kind of continuation method.[16,17] In fact, the continuation method has been used to derive a differential equation for the state estimate without any discretization,[18,19] which can be used as a real-time algorithm. However, that differential equation involves the complicated Riccati differential equation for solving a linear two-point boundary-value problem (TPBVP) associated with the time-dependent variation in the optimal trajectory. Therefore, there is still room for improvement in the amount of computation even if iterative optimization methods are avoided.

In this paper, we propose a fast algorithm, called the continuation/generalized minimum residual C/(GMRES) method, for nonlinear MHSE, in which the continuation method is combined with a fast algorithm for linear algebraic equation instead of the complicated Riccati differential equation. The nonlinear MHSE problem is formulated in a deterministic setting, and the optimization problem is discretized over the horizon. Then, an associated TPBVP to be solved at each time step is derived, and a differential equation for updating the optimized variable is obtained through the use of the continuation method. Because the differential equation involves a large linear algebraic equation, the GMRES method[20] is employed to solve the linear equation efficiently.

C/GMRES was originally proposed for the nonlinear receding horizon control (model predictive control) problem,[21,22] in which the control performance over a finite future is optimized, and it has been shown that C/GMRES is not only faster but also more numerically robust than the conventional algorithm involving the Riccati differential equation.[21] Therefore, it is meaningful to adopt C/GMRES into the nonlinear MHSE problem, which is the dual of the nonlinear receding horizon control problem in the sense that the horizon is time reversed. As shown in this paper, C/GMRES is also applicable to the MHSE problem with suitable formulation and modification. Some features inherent in the MHSE problem are also discussed.

To examine the practical applicability and computational time, the proposed algorithm is applied to a numerical simulation and an experiment with a hovercraft model. It is shown that MHSE by the proposed algorithm generates reasonable estimates even when the EKF fails. Results of simulation and experiment also show that nonlinear MHSE is possible in real time with the proposed algorithm.

Problem Formulation

Nonlinear Moving Horizon State Estimation Problem

This section briefly summarizes the nonlinear MHSE problem. All functions are assumed to be differentiable as many times as necessary hereafter. We consider a generic nonlinear system expressed by the following state equation and measurement equation:

$$\begin{cases} \dot{x}(t) = f[x(t), u(t), w(t)] \\ y(t) = g[x(t), u(t), v(t)] \end{cases} \tag{1}$$

where $x(t) \in \boldsymbol{R}^n$ denotes the state vector, $u(t) \in \boldsymbol{R}^{m_u}$ the input vector, $w(t) \in \boldsymbol{R}^{m_w}$ the vector of unknown disturbances, $y(t) \in \boldsymbol{R}^{m_y}$ the measurement vector, and $v(t) \in \boldsymbol{R}^{m_v}$ the vector of unknown measurement noise, respectively. Equality constraints are also imposed in general as

$$C[x(t), w(t), p(t)] = 0 \tag{2}$$

where C is an m_c-dimensional vector-valued function. In the case of an inequality constraint, it is necessary to introduce a penalty function method or some heuristic modification of the problem including a dummy input.[21] For the sake of simple notation, all known quantities are put together into a vector $p(t) \in \boldsymbol{R}^{m_p}$, for example, $p(t) = [y^T(t)\ u^T(t)]^T$. If there are any other known time-variant parameters in the problem setting, they can also be included in $p(t)$.

In general, state estimation can be regarded as a problem to find, for a certain time horizon, a state trajectory that is consistent with the model and measurement for some possible disturbance and noise. More specifically, it is a problem to find functions $x(t)$, $w(t)$, and $v(t)$ that satisfy Eqs. (1) and (2) for the given functions $y(t)$ and $u(t)$. To determine the estimate uniquely among the consistent estimates, we minimize the disturbance and noise under a certain criterion in this paper, regarding $w(t)$ and $v(t)$ in Eqs. (1) and (2) as the deviation of the estimate from the ideal noise-free model and measurement. To formulate the problem as a standard optimization problem, we assume that $v(t)$ can be represented with the state vector $x(t)$ and the known quantity $p(t)$ from the measurement equation in Eq. (1) as

$$v(t) = h[x(t), p(t)]$$

This assumption holds trivially in the case of additive measurement noise, where the noise represents the estimation residual, and is not a significant restriction on the problem. Then, the quantities to be minimized are expressed as functions of p, x, and w.

In the MHSE problem, the estimates of the state and disturbance are determined at each time t so as to minimize a performance index over the horizon $[t-T, t]$:

$$J = \eta[\hat{x}(t), p(t)] + \phi\{x^w[t-T; t, \hat{x}(t)], p(t-T)\} + \int_{t-T}^{t} L\{x^w[t'; t, \hat{x}(t)], w(t'), p(t')\}\,\mathrm{d}t' \tag{3}$$

where $\hat{x}(t)$ denotes the estimate of $x(t)$ and $x^w(t'; t, \hat{x})$ $(t - T \le t' \le t)$ the state trajectory for the input functions u and w such that $x^w(t; t, \hat{x}) = \hat{x}$ (Fig. 1). Note that $x^w[t-T; t, \hat{x}(t)]$ is not necessarily identical to $\hat{x}(t-T)$ obtained in the past estimation. The functions η, ϕ, and L are chosen to be appropriate penalties on the estimation residual, the estimated disturbance and noise, or unnecessary variation in the estimates. The problem to be solved at each time is essentially a generalization of a deterministic smoothing problem for a linear system.[23] The optimized quantities at time t are the terminal state $\hat{x}(t)$ and the time history of the disturbance $w(t')$ $(t-T \le t' \le t)$, and the optimal value of $\hat{x}(t)$ is regarded as the resultant estimate of the MHSE problem.

By introducing a fictitious time τ, the MHSE problem can be rewritten as a family of finite horizon optimization problems along the τ axis as follows.

Minimize:

$$J = \eta[x^*(0, t), p(t)] + \phi[x^*(-T, t), p(t-T)] + \int_{-T}^{0} L[x^*(\tau, t), w^*(\tau, t), p(t+\tau)]\,\mathrm{d}\tau$$

Subject to:

$$\begin{cases} x^*_\tau(\tau, t) = f[x^*(\tau, t), u(t+\tau), w^*(\tau, t)] \\ C[x^*(\tau, t), w^*(\tau, t), p(t+\tau)] = 0 \end{cases}$$

where x^*_τ denotes the partial derivative of x^* with respect to τ. The new state vector $x^*(\tau, t)$ denotes the trajectory along the τ axis such that $x^*(0, t) = \hat{x}(t)$ (Fig. 1). Then, the problem reduces to a kind of optimal control problem in which the terminal state is unknown, and the resultant estimates are expressed as $\hat{x}(t) = x^*(0, t)$ and $\hat{w}(t) = w^*(0, t)$. The horizon length T is a function of time, $T = T(t)$, in general, as is explained later.

Discretized Problem

At each time t, we find the optimal trajectory over $[t-T, t]$ that minimizes the performance index and employ its terminal state at time t as the present estimate $\hat{x}(t)$. To this end, we divide the horizon into N equal steps and discretize the optimal estimation problem with the backward difference as follows:

$$x^*_i(t) = x^*_{i+1}(t) - f\left[x^*_{i+1}(t), u^*_{i+1}(t), w^*_{i+1}(t)\right]\Delta\tau \tag{4}$$

$$C\left[x^*_i(t), w^*_i(t), p^*_i(t)\right] = 0 \tag{5}$$

$$J = \eta\left[x^*_N(t), p^*_N(t)\right] + \phi\left[x^*_0(t), p^*_0(t)\right] + \sum_{i=1}^{N} L\left[x^*_i(t), w^*_i(t), p^*_i(t)\right]\Delta\tau \tag{6}$$

where the discretization step is given by $\Delta\tau := T/N$, $x^*_i(t)$ corresponds to the state at time $t - T + i\Delta\tau$ on the optimal trajectory (Fig. 2), and $p^*_i(t)$ and $u^*_i(t)$ are given by $p(t - T + i\Delta\tau)$ and $u(t - T + i\Delta\tau)$, respectively. The optimized quantities at each time t are the terminal state $\hat{x}(t) = x^*_N(t)$ and the discretized sequence of the disturbance $\{w^*_i(t)\}_{i=1}^N$.

Let H denote the Hamiltonian defined by

$$H(x, \lambda, w, \mu, p) := L(x, w, p) + \lambda^T f(x, w, u) + \mu^T C(x, w, p)$$

where $\lambda \in \boldsymbol{R}^n$ denotes the costate and $\mu \in \boldsymbol{R}^{m_c}$ denotes the Lagrange multiplier associated with the equality constraint. The first-order necessary conditions for the optimal estimates $\hat{x}(t) = x^*_N(t)$ and $\{w^*_i(t)\}_{i=1}^N$ are readily obtained as a TPBVP by the calculus of

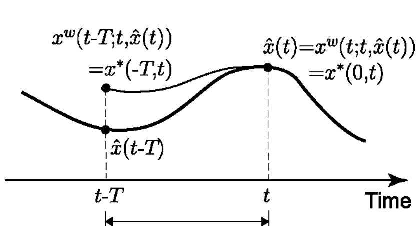

Fig. 1 Estimated state trajectory over the moving horizon.

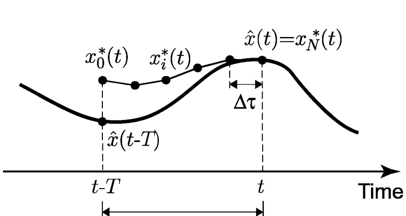

Fig. 2 Estimated state trajectory over the discretized horizon.

variations as

$$H_w^T\left[x_i^*(t), \lambda_{i-1}^*(t), w_i^*(t), \mu_i^*(t), p_i^*(t)\right] = 0 \tag{7}$$

$$\lambda_i^*(t) = \lambda_{i-1}^*(t) - H_x^T\left[x_i^*(t), \lambda_{i-1}^*(t), w_i^*(t), \mu_i^*(t), p_i^*(t)\right]\Delta\tau \tag{8}$$

$$\lambda_N^*(t) = \eta_x^T\left[x_N^*(t), p_N^*(t)\right] \tag{9}$$

$$\lambda_0^*(t) = -\phi_x^T\left[x_0^*(t), p_0^*(t)\right] \tag{10}$$

where H_w denotes the partial derivative of H with respect to w. The sequences of the optimal estimates $\{w_i^*(t)\}_{i=1}^N$, $\hat{x}(t) = x_N^*(t)$ and the multiplier $\{\mu_i^*(t)\}_{i=1}^N$ must satisfy Eqs. (4) and (5) and (7–10). The TPBVP for the discretized problem is identical to a finite difference approximation of the TPBVP for the original continuous-time problem. Therefore, the solution of the discretized problem converges to the solution of the continuous-time problem as $N \to \infty$ under mild conditions.[24]

We define a vector of the estimates and multipliers and a vector of known quantities as

$$W(t) := \left[w_1^{*T}(t)\, \mu_1^{*T}(t) \quad w_2^{*T}(t)\, \mu_2^{*T}(t)\right.$$
$$\left. \ldots w_N^{*T}(t)\, \mu_N^{*T}(t)\, x_N^{*T}(t)\right]^T \in \boldsymbol{R}^m$$

$$P(t) := \left[p_1^{*T}(t) \quad p_2^{*T}(t) \ldots p_N^{*T}(t)\right]^T \in \boldsymbol{R}^{m_p N}$$

where $m := (m_w + m_c)N + n$. We also define a projection $\Pi_N : \boldsymbol{R}^m \to \boldsymbol{R}^n$ as

$$\Pi_N(W) := x_N^*$$

For given $W(t)$ and $P(t)$, $\{x_i^*(t)\}_{i=1}^N$ is calculated backward recursively using Eq. (4), and then $\{\lambda_i^*(t)\}_{i=1}^N$ is also calculated recursively from $i = 0$ to N by using Eqs. (8) and (10). Therefore, Eqs. (5), (7), and (9) can be regarded as one equation defined as

$$F[W(t), P(t), t]$$

$$:= \begin{bmatrix} H_w^T\left[x_1^*(t), \lambda_0^*(t), w_1^*(t), \mu_1^*(t), p_1^*(t)\right] \\ C\left[x_1^*(t), w_1^*(t), p_1^*(t)\right] \\ \vdots \\ H_w^T\left[x_N^*(t), \lambda_{N-1}^*(t), w_N^*(t), \mu_N^*(t), p_N^*(t)\right] \\ C\left[x_N^*(t), w_N^*(t), p_N^*(t)\right] \\ \lambda_N^*(t) - \eta_x^T\left[x_N^*(t), p_N^*(t)\right] \end{bmatrix} = 0 \tag{11}$$

If the equation is solved with respect to $W(t)$ for the known quantities $P(t)$, then the optimal estimate $\hat{x}(t) = \Pi_N[W(t)]$ is obtained.

Other high-order discretization schemes can also be employed to obtain the equation corresponding to Eq. (11), at the expense of simplicity and computational burden. In particular, generalization of Eqs. (4) and (8) to any explicit scheme is straightforward as long as the state and costate can be calculated recursively.

Continuation/GMRES Method

Instead of solving the nonlinear equation $F(W, P, t) = 0$ itself at each time with such an iterative method as Newton's method, we find the derivative of W with respect to time such that $F[W(t), P(t), t] = 0$ is satisfied identically. Namely, we choose $W(0)$ so that $F[W(0), P(0), 0] = 0$ and determine $\dot{W}(t)$ so that

$$\dot{F}(W, P, t) = A_s F(W, P, t) \tag{12}$$

where A_s is a stable matrix introduced to stabilize $F = 0$. By total differentiation, we have

$$F_W \dot{W} = A_s F - F_P \dot{P} - F_t$$

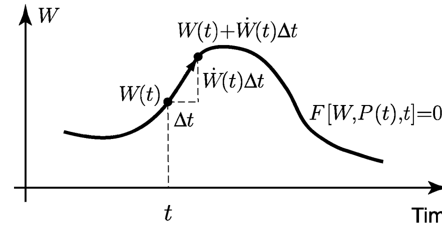

Fig. 3 Continuation method for tracing the solution curve.

which can be regarded as a linear algebraic equation, with a coefficient matrix F_W, to determine $\dot{W}$ for given W, P, $\dot{P}$, and t. Then, if the Jacobian F_W is nonsingular, we obtain a differential equation for $W(t)$ as

$$\dot{W} = F_W^{-1}(A_s F - F_P \dot{P} - F_t) \tag{13}$$

We can update the solution $W(t)$ of $F[W(t), P(t), t] = 0$ without iterative optimization methods by integrating Eq. (13) in real time as, for example, $W(t + \Delta t) = W(t) + \dot{W}(t)\Delta t$, where Δt denotes the sampling period. The present approach is a kind of continuation method[16,17] in the sense that the solution curve $W(t)$ is traced by integrating a differential equation (Fig. 3).

In the present deterministic formulation, we assume the estimate $W(t)$ and the measured quantity $P(t)$ are differentiable with respect to time, while actual disturbances, noise, and inputs are not necessarily differentiable. However, this assumption is not a practical limitation because the actual computation process is discretized with respect to time t. That is, differentiation is approximated by finite difference, which can be performed formally even for noisy or discontinuous data.

From the computational point of view, the differential equation (13) still involves expensive operations, that is, Jacobians F_W, F_P, and F_t and the linear algebraic equation associated with F_W^{-1}. In the present definition of F in Eq. (11), the Jacobian F_W is dense because $x_i^*(t)$ and $\lambda_i^*(t)$ are functions of $W(t)$, and therefore the sparse finite difference is not applicable. To reduce the computational cost in the Jacobians and the linear equation, we employ two techniques, that is, the forward difference approximation for products of Jacobians and vectors, and the GMRES method[20,25] for the linear algebraic equation.

First, we approximate the products of the Jacobians and some $U_1 \in \boldsymbol{R}^m$, $U_2 \in \boldsymbol{R}^{m_p N}$, and $U_3 \in \boldsymbol{R}$ with the forward difference as follows:

$$F_W(W, P, t)U_1 + F_P(W, P, t)U_2 + F_t(W, P, t)U_3$$
$$\simeq D_h F(W, P, t : U_1, U_2, U_3)$$
$$:= [F(W + hU_1, P + hU_2, t + hU_3) - F(W, P, t)]/h$$

where h is a positive real number. Then Eq. (12) is approximated by

$$D_h F(W, P, t : \dot{W}, \dot{P}, 1) = A_s F(W, P, t)$$

which is equivalent to

$$D_h F(W, P + h\dot{P}, t + h : \dot{W}, 0, 0) = b(W, P, \dot{P}, t) \tag{14}$$

where

$$b(W, P, \dot{P}, t) := A_s F(W, P, t) - D_h F(W, P, t : 0, \dot{P}, 1)$$

The forward difference approximation is different from the finite difference approximation of the Jacobians themselves. The forward difference approximation of the products of the Jacobians and vectors can be calculated with only an additional evaluation of the function, which requires notably less computational burden than an approximation of the Jacobians themselves.

Because Eq. (14) approximates a linear equation with respect to $\dot{W}$, we apply GMRES to Eq. (14), which is called FDGMRES,[20] with a certain initial guess $\hat{\dot{W}}$.

Algorithm 1

$\dot{W} := \text{FDGMRES}(W, P, \dot{P}, t, \hat{\dot{W}}, h, k_{\max})$.

1) $r_0 := b(W, P, \dot{P}, t) - D_h F(W, P + h\dot{P}, t + h : \hat{\dot{W}}, 0, 0)$, $v_1 := r_0/\|r_0\|$, $\rho := \|r_0\|$, $\beta := \rho$, $k := 0$.

2) While $k < k_{\max}$, do

a) $k := k + 1$.

b) $v_{k+1} := D_h F(W, P + h\dot{P}, t + h : v_k, 0, 0)$, for $j = 1, \ldots, k$.

i) $h_{jk} := v_{k+1}^T v_j$,

ii) $v_{k+1} := v_{k+1} - h_{jk} v_j$.

c) $h_{k+1,k} := \|v_{k+1}\|$.

d) $v_{k+1} := v_{k+1}/\|v_{k+1}\|$.

e) For $e_1 = [1\ 0 \cdots 0]^{\mathrm{T}} \in \boldsymbol{R}^{k+1}$ and $H_k = (h_{ij}) \in \boldsymbol{R}^{(k+1)\times k}$ ($h_{ij} = 0$ for $i > j + 1$), minimize $\|\beta e_1 - H_k z^k\|$ to determine $z^k \in \boldsymbol{R}^k$.

f) $\rho := \|\beta e_1 - H_k z^k\|$.

3) $\dot{W} := \hat{\dot{W}} + V_k z^k$, where $V_k = [v_1 \cdots v_k] \in \boldsymbol{R}^{m\times k}$.

GMRES is a kind of Krylov subspace method for such a linear equation as $Ax = b$ with A not necessarily symmetric or positive definite. GMRES at the kth iteration minimizes the residual $\rho := \|b - Ax\|$ with $x \in x_0 + \mathcal{K}_k$, where x_0 is the initial guess and $\mathcal{K}_k$ denotes the Krylov subspace defined by $\mathcal{K}_k := \mathrm{span}\{r_0, Ar_0, \ldots, A^{k-1}r_0\}$ with $r_0 := b - Ax_0$. GMRES also successively generates an orthonormal basis $\{v_j\}_{j=1}^k$ for $\mathcal{K}_k$. Minimization in step 2e is executed efficiently through the use of Givens rotations. In principle, GMRES reduces the residual monotonically and converges to the solution within the same number of iterations as the dimension of the equation. However, an important advantage of GMRES for a large linear equation is that a specified error tolerance, for example, $\rho \leq \xi \|r_0\|$ $(0 < \xi < 1)$, can be achieved with much fewer iterations.

In step 2b, the product of a matrix and a vector in the usual GMRES, $F_W(W, P, t)v_j$, is replaced with its forward difference approximation, $D_h F(W, P + h\dot{P}, t + h : v_k, 0, 0)$. It is clear that the orthonormal basis $\{v_i\}_{i=1}^k$, k vectors in $\boldsymbol{R}^m$, must be stored during the execution of FDGMRES, which might require a huge amount of data storage for a large problem. Moreover, many iterations can be impossible from the viewpoint of execution time in real-time implementation. Therefore, the iteration number $k_{\max}$ should be chosen as small as possible. Fortunately, in the present real-time application the solution $\dot{W}$ at the previous sampling time is often a good initial guess for FDGMRES, and a small $k_{\max}$ often suffices to obtain an accurate solution. If a large number of iterations are necessary with a limited amount of data storage, it is a common technique to restart FDGMRES from the current iterate by resetting the orthonormal basis.

With $\dot{W}$ obtained approximately through the use of FDGMRES, $W(t)$ is updated by integrating $\dot{W}$ in real time. The continuation/GMRES method for nonlinear MHSE is summarized as follows.

Algorithm 2 (C/GMRES)

1) Let the horizon $T(t)$ be a smooth function such that $T(0) = 0$ and $T(t) \to T_f$ $(t \to \infty)$. Let Δt be the sampling period, and let $\hat{x}(0)$ be an appropriate initial guess. Let $t := 0$, obtain the initial measurement $p(0)$, and let $x_i^*(0) := \hat{x}(0)$, $\lambda_i^*(0) := -\phi_x^T[\hat{x}(0), p(0)]$ $(i = 0, \ldots, N)$. Find $\hat{w}(0)$ and $\mu(0)$ analytically or numerically such that

$$\begin{bmatrix} H_w^T\left\{\hat{x}(0), -\phi_x^T[\hat{x}(0), p(0)], \hat{w}(0), \mu(0), p(0)\right\} \\ C[\hat{x}(0), \hat{w}(0), p(0)] \end{bmatrix} = 0$$

Let $w_i^*(0) = \hat{w}(0)$ and $\mu_i^*(0) = \mu(0)$ $(i = 1, \ldots, N)$, which gives the initial condition $W(0)$ such that $\|F[W(0), P(0), 0]\| = \|\phi_x^T[\hat{x}(0), p(0)] + \eta_x^T[\hat{x}(0), p(0)]\|$.

2) At time $t + \Delta t$, obtain the measurement $p(t + \Delta t)$, and let $\Delta P := P(t + \Delta t) - P(t)$. Compute $\dot{W}(t)$ by $\dot{W} :=$ FDGMRES $(W, P, \Delta P/\Delta t, t, \hat{\dot{W}}, h, k_{\max})$, where the initial guess $\hat{\dot{W}}$ is appropriately chosen, for example, $\hat{\dot{W}} := 0$ or $\hat{\dot{W}} := \dot{W}(t - \Delta t)$ with $\dot{W}(-\Delta t) := 0$. Let $W(t + \Delta t) = W(t) + \dot{W}(t)\Delta t$. The estimate is obtained as $\hat{x}(t + \Delta t) = \Pi_N[W(t + \Delta t)]$.

3) Let $t := t + \Delta t$, and go back to step 2.

The iterative method is used only to solve the linear algebraic equation (13) with respect to $\dot{W}$, and through the use of its solution the solution of the nonlinear equation $F(W, P, t) = 0$ is traced without any line search or Newton iteration. In particular, C/GMRES solves the linear equation (13) only once at each sampling time and, therefore, requires much less computational burden than such iterative methods as Newton's method, which solves a linear equation several times to determine search directions.

In step 1, because $T(0) = 0$, $W(0)$ can be initialized by finding only $\hat{w}(0)$ and $\mu(0)$, which requires much less computational burden than the original Eq. (11). One can also introduce a search of $\hat{x}(0)$ and an appropriate continuation in the penalties ϕ and η so that $\phi_x^T[\hat{x}(0), p(0)] + \eta_x^T[\hat{x}(0), p(0)] = 0$. When increasing $T(t)$ to T_f, the increasing rate $\mathrm{d}T(t)/\mathrm{d}t$ must be limited so that the measured data are available for the horizon $[t - T(t), t]$. If $t_d := \inf_{t\geq 0}[t - T(t)]$ is negative, data acquisition must start at $t = t_d$ before estimation starts at $t = 0$. Moreover, the vector $P(t)$ must be constructed by interpolating the measured data, unless the step on the horizon $\Delta\tau = T(t)/N$, is identical to the sampling period Δt.

In comparison to the case of nonlinear receding horizon control, the present algorithm has such inherent features that the terminal state is included in the unknown vector W, the measured data must be differentiated, and the horizon length is limited by the acquired data length. However, error analysis of C/GMRES in a general form,[21] taking discretization errors into account, is also valid for the present case of nonlinear MHSE. That is, under some assumptions such that $\|\dot{P}\|$ (or its finite difference approximation), $\|F_W\|$, $\|F_W^{-1}\|$, $\|F_P\|$, and $\|F_t\|$ are bounded and F_W, F_P, and F_t are Lipschitz continuous, the error $\|F\|$ is bounded if $A_s = -\zeta I$ with $0 < \zeta\Delta t \leq \bar{\zeta}\Delta t$ for some $\bar{\zeta}$ such that $1 \leq \bar{\zeta}\Delta t < 2$.

Numerical Example

Model and Performance Index

To evaluate the effectiveness of C/GMRES for nonlinear MHSE, we apply the algorithm to the state estimation of a hovercraft model. Figure 4 shows the coordinate system. The state equation of the hovercraft model is given by

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \\ \dot{x}_1(t) \\ \dot{x}_2(t) \\ \dot{x}_3(t) \end{bmatrix} = \begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \\ \dot{x}_3(t) \\ \dfrac{u_1(t) + u_2(t)}{M}\cos x_3(t) + w_1(t) \\ \dfrac{u_1(t) + u_2(t)}{M}\sin x_3(t) + w_2(t) \\ \dfrac{u_1(t) - u_2(t)}{I_c} r + w_3(t) \end{bmatrix}$$

$$\begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix} = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} v_1(t) \\ v_2(t) \end{bmatrix} = Dx(t) + v$$

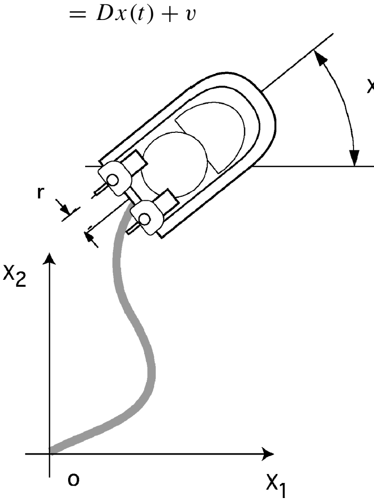

Fig. 4 Hovercraft model.

where $(x_1(t), x_2(t))$ denotes the position of the center of mass of the hovercraft and $x_3(t)$ the attitude angle. The parameters I_c, M, and r denote the moment of inertia, the mass, and the distance between the thrusters and the center of mass, respectively, $u_1(t)$ and $u_2(t)$ are outputs of the thrusters; $w_1(t)$, $w_2(t)$, and $w_3(t)$ are unknown disturbances; and $v_1(t)$ and $v_2(t)$ are unknown measurement noise. Hence, the only available measurement is the position of the hovercraft, and it is necessary to estimate the attitude angle and velocities.

The performance index with a moving horizon is chosen in the form of Eq. (3) with

$$p = [y^T \; u^T \; \hat{x}^T]^T$$

$$\eta = 0$$

$$\phi = \tfrac{1}{2}\{\hat{x}(t-T) - x^w[t-T; t, \hat{x}(t)]\}^{\mathrm{T}} S_f \{\hat{x}(t-T) - x^w[t-T; t, \hat{x}(t)]\}$$

$$L = \tfrac{1}{2}(\{y(t') - Dx^w[t'; t, \hat{x}(t)]\}^{\mathrm{T}} Q \{y(t') - Dx^w[t'; t, \hat{x}(t)]\} + w^T(t') W w(t'))$$

where S_f, Q, and R are weighting matrices and $\hat{x}(t-T)$ denotes the estimate in the past, which has already been obtained through optimization over $[t-2T, t-T]$ and is not necessarily identical to $x^w[t-T; t, \hat{x}(t)]$ in the current optimization over $[t-T, t]$. The penalty ϕ is added because it is preferable that the optimal trajectory be consistent with the past estimate. Such a penalty is also often effective for attaining numerical stability of the computation.

Accuracy of Estimation

To confirm the effectiveness for nonlinearity, we compare the proposed algorithm with an EKF. Inputs are given by

$$u_1(t) = 0.3\,\sin(0.2t) + 0.1$$

$$u_2(t) = 0.3\,\sin(0.2t + 0.1) + 0.1$$

The disturbances $w_1(t)$, $w_2(t)$, and $w_3(t)$ are given by discretized random sequences of the normal distributions of average 0 and standard deviation 0.001 N, and the measurement noise $v_1(t)$ and $v_2(t)$ is given by discretized random sequences of the normal distributions of average 0 and standard deviation 0.005 m. They are treated as constant over the integration step Δt of simulation.

The physical parameters of the simulation are the moment of inertia $I_c = 0.0125\ \mathrm{kg\cdot m^2}$, the mass $M = 0.86$ kg, and the distance $r = 0.0485$ m. For MHSE, the following weighting matrices are chosen:

$$S_f = \mathrm{diag}\,[10, 5, 0.1, 1, 1, 0.5]$$

$$Q = \mathrm{diag}\,[5, 5]$$

$$W = \mathrm{diag}\,[15, 15, 20]$$

These weighting matrices are chosen by trial and error. If the weight W on the disturbance is large in comparison with the weight Q on the noise, the unknown disturbance is assumed to be small, and the discrepancy between the measured output y and the estimated output Dx^w is attributed mainly to the measurement noise v. Conversely if the weight on the noise is large in comparison with the weight on the disturbance, the noise is assumed to be small, and the disturbance w is determined to make the state trajectory consistent with the measured output. If the weight S_f on the difference between the past estimate $\hat{x}(t-T)$ and the corresponding current estimate $x^w[t-T; t, \hat{x}(t)]$ is large, the estimate $\hat{x}(t)$ is determined relying mainly on the past estimate rather than the currently available measurement.

The computation is performed on a personal computer (CPU: Power PC, 600 MHz) with the integration step $\Delta t = 1/120$ s, the number of grids $N = 20$, $k_{\max} = 20$, and $\zeta = 1/\Delta t$. The length of the horizon is chosen so that $T(0) = 0$ and $T(t) \to T_f$ as $t \to \infty$, namely, $T(t) := T_f(1 - e^{-\alpha t})$ with $T_f = 2$ s and $\alpha = 0.5$. As discussed in the preceding section, if data before $t = 0$ are not available, $t - T(t)$ must always be nonnegative. Therefore, the parameter α must satisfy the condition $\dot{T}(0) = \alpha T_f \le 1$. The initial estimate is given as $\hat{x}(0) = [0, 0, 0.2, 0, 0, 0]^T$, and the actual initial state is set as $x(0) = [0, 0, 0.05, 0, 0, -0.01]^T$.

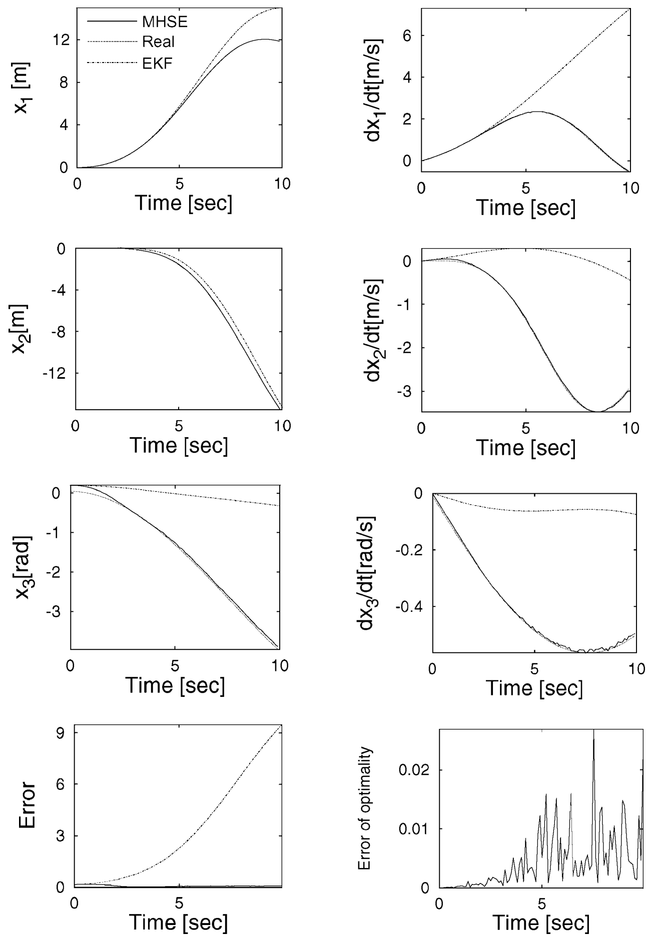

Fig. 5 Results of estimation.

Figure 5 shows the results of estimation by the proposed MHSE algorithm and the EKF. The figure shows the estimation error $\|\hat{x} - x\|$ for both algorithms as well as the error of optimality $\|F\|$ in MHSE. The estimates by MHSE (solid lines) agree well with the real states (broken lines). In particular, the estimates of x_1, $\dot{x}_1$, x_2, and $\dot{x}_2$ are indistinguishable from the corresponding real states. Although there are initial errors in x_3 and $\dot{x}_3$, the errors attenuate within 2.5 s in MHSE. In contrast, the EKF fails to generate reasonable estimates, and the estimates by EKF (dash–dotted lines) deviate from the real states. The estimation error for EKF goes beyond nine during the simulation of 10 s, while the estimation error for MHSE is always less than 0.2.

Computational Cost

To confirm the reduction of computational cost, we compare the proposed algorithm C/GMRES with a general-purpose algorithm. We employ MINPACK‡ to solve the nonlinear equation (11). The simulation settings are the same as described in the preceding subsection. Another algorithm based on the continuation method,[18,19] which involves the complicated Riccati differential equation, is not considered here because it is not only slower than C/GMRES but also sensitive to numerical error and often fails, as shown in the case of receding horizon control.[21] Table 1 shows average computational times per update and the average error of optimality per grid point

‡Data available online at http://www.netlib.org/minpack/ [cited 28 October 2000].

Table 1 Average computational time per update (ms) and the average error of optimality per grid point $\|F\|/N$

N	C/GMRES time, ms	Error	MINPACK time, ms	Error
10	0.4	4.3×10^{-3}	6.26	5.3×10^{-3}
20	0.82	1.9×10^{-3}	24.0	2.8×10^{-3}
30	1.1	1.1×10^{-3}	65.2	1.9×10^{-3}

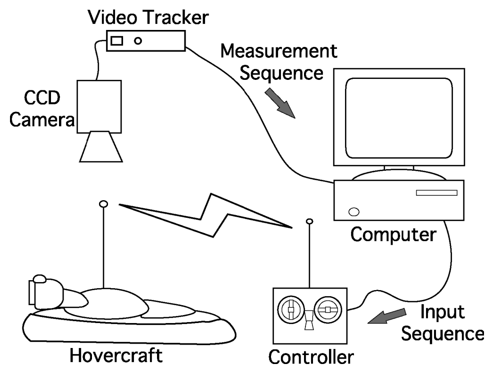

Fig. 6 Experimental apparatus.

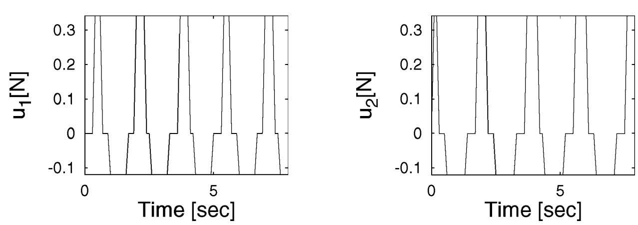

Fig. 7 Control inputs in the experiment.

$\|F\|/N$. Tuning parameters in MINPACK are chosen so that errors are of an order comparable to those in C/GMRES. Table 1 indicates that the computational time of the proposed algorithm is 1.1 ms even when $N = 30$, and it increases linearly with respect to the number of grid points, whereas MINPACK takes 6.26 ms on average even for the smallest number of grids $N = 10$ and takes 65.2 ms for $N = 30$. The computational time for update in C/GMRES is constant because it involves no iterative search, while the computational time for update in an iterative optimization method varies in general. Therefore, the proposed algorithm can be implemented with a sampling period not exceeding 1.1 ms in the present computational conditions.

Experiment

To evaluate the actual applicability of the proposed algorithm, we apply C/GMRES to MHSE in an experimental apparatus (Fig. 6). A hovercraft model (Taiyo Toy Ltd., Typhoon T-3; 356 mm × 212 mm × 142 mm) is controlled by a computer through a radio controller. The motion of the hovercraft model is measured by a charge-coupled-device camera, and the measurements are sent to the computer via a processing unit (OKK Inc., Video Tracker G280). In the experiment, the measurement equation is given as

$$\begin{bmatrix} y_1(t) \\ y_2(t) \\ y_3(t) \end{bmatrix} = \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix} + \begin{bmatrix} v_1(t) \\ v_2(t) \\ v_3(t) \end{bmatrix}$$
$$= Dx(t) + v$$

State estimation is carried out for a set of experimental data over 8 s. The physical parameters of the experiment are the same as those in the numerical example. Figure 7 shows the outputs of the two thrusters. Only three values, -0.121, 0, and 0.342, are available as the input in the experiment. The weighting matrices for MHSE are chosen as

$$S_f = \mathrm{diag}\,[0.2, 0.2, 0.2, 0.001, 0.001, 0.001]$$
$$Q = \mathrm{diag}\,[50, 50, 50]$$
$$W = \mathrm{diag}\,[3, 3, 0.1]$$

The computation is performed on a personal computer (CPU: PowerPC, 240 MHz) with the integration step $\Delta t = 1/120$ s, the number of grid points $N = 15$, $k_{\max} = 15$, and $\zeta = 1/\Delta t$. The length of the horizon is chosen so that $T(0) = 0$ and $T(t) \to T_f$ as $t \to \infty$, namely, $T(t) := T_f(1 - e^{-\alpha t})$ with $T_f = 0.2$ s and $\alpha = 1$. The initial estimates for the position and attitude are given by the measured data, and the initial estimates for the velocities are set zero.

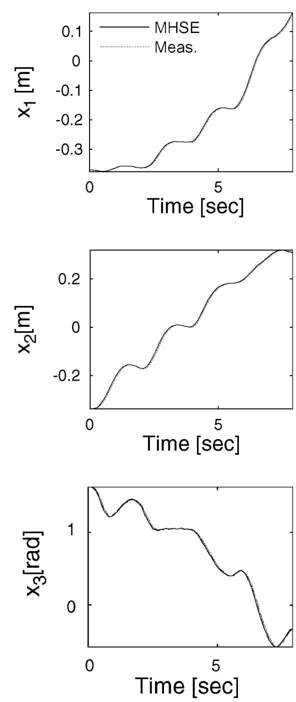

Fig. 8 Estimation of the position and attitude.

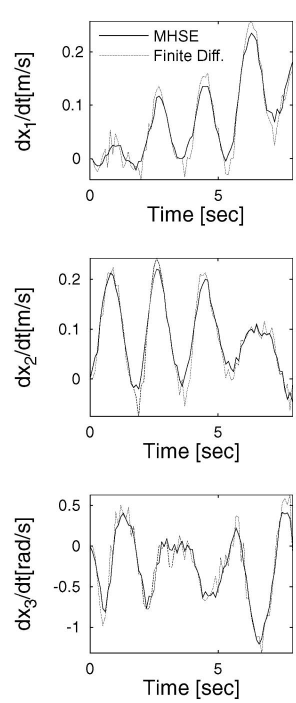

Fig. 9 Estimation of the velocities.

Figures 8 and 9 show the results of computation by C/GMRES and measurement data. Figure 8 shows the estimates of the position and attitude as well as the measured values. The estimates of position and attitude almost overlap with the measurements.

Figure 9 shows the estimates of the velocities, for which true values are not available in the experiment. The solid lines represent the estimates calculated by MHSE, and the broken lines represent the estimates calculated by processing the finite difference of the measured positions with a low-pass filter. The estimates by MHSE are smoother than the filtered data in Fig. 9, which indicates a smoothing property of MHSE based on the model. The total computational time for the estimation over 10 s is 5.55 s, which implies that the computational time for updating the estimate is $5.55 \div (10 \times 120) = 4.6 \times 10^{-3}$ s and the algorithm can be implemented in real time with a smaller sampling period than 1/120 s.

Conclusions

We proposed a real-time algorithm for nonlinear moving horizon state estimation by combining the continuation method with generalized minimum residual (GMRES). A state estimation problem has been formulated in a deterministic setting as a minimization of a performance index with a moving horizon over a finite past. To solve the optimization problem efficiently, the continuation method has been introduced to trace the solution without any line search or Newton iteration. As a result, the estimate can be updated at each sampling time by solving a linear algebraic equation only once. Moreover, the linear algebraic equation can be solved efficiently through the use of GMRES, one of the Krylov subspace methods.

The proposed algorithm has been demonstrated in a simulation and an experiment with a hovercraft model whose dynamics is nonlinear. The simulation results reveal that the proposed algorithm generates reasonable estimates even when the extended Kalman filter fails. Moreover, the results of the simulation and experiment confirm that the proposed algorithm is executable in real time.

Acknowledgments

This work is partially supported by the Research Fund of the Kurata Memorial Science and Technology Promotion Foundation, Grants for Receivers of Doctorates by the Foundation for C & C Promotion, and a Grant-in-Aid for Scientific Research (No. 15760319) from the Ministry of Education, Culture, Sports, Science and Technology, Japan.

References

[1]Gelb, A. (ed.), *Applied Optimal Estimation*, MIT Press, Cambridge, MA, 1974, Chap. 6.

[2]Misawa, E. A., and Hedrick, J. K., "Nonlinear Observers—A State-of-the-Art Survey," *Journal of Dynamic Systems, Measurement and Control*, Vol. 111, Sept. 1989, pp. 344–352.

[3]Muske, K. R., and Edgar, T. F., "Nonlinear State Estimation," *Nonlinear Process Control*, edited by M. A., Henson and D. E. Seborg, Prentice–Hall, Upper Saddle River, NJ, 1997, Chap. 6, pp. 311–370.

[4]Julier, S., Uhlmann, J., and Durrant-Whyte, H. F., "A New Method for the Nonlinear Transformation of Means and Covariances in Filters and Estimators," *IEEE Transactions on Automatic Control*, Vol. 45, No. 3, 2000, pp. 477–482.

[5]Nørgaard, M., Poulsen, N. K., and Ravn, O., "New Developments in State Estimation for Nonlinear Systems," *Automatica*, Vol. 36, No. 11, 2000, pp. 1627–1638.

[6]Krener, A. J., and Isidori, A., "Linearization by Output Injection and Nonlinear Observers," *Systems and Control Letters*, Vol. 3, No. 1, 1983, pp. 47–52.

[7]Krener, A. J., and Respondek, W., "Nonlinear Observers with Linearizable Error Dynamics," *SIAM Journal on Control and Optimization*, Vol. 23, No. 2, 1985, pp. 197–216.

[8]Levine, J., and Marino, R., "Nonlinear System Immersion, Observers and Finite-Dimensional Filters," *Systems and Control Letters*, Vol. 7, No. 2, 1986, pp. 133–142.

[9]Jouan, P., "Immersion of Nonlinear Systems into Linear Systems Modulo Output Injection," *SIAM Journal on Control and Optimization*, Vol. 41, No. 6, 2003, pp. 1756–1778.

[10]Back, J., and Seo, J. H., "Immersion of Non-linear Systems into Linear Systems up to Output Injection: Characteristic Equation Approach," *International Journal of Control*, Vol. 77, No. 8, 2004, pp. 723–734.

[11]Shim, H., Seo, J. H., and Teel, A. R., "Nonlinear Observer Design via Passivation of Error Dynamics," *Automatica*, Vol. 39, No. 5, 2003, pp. 885–892.

[12]Jang, S.-S., Joseph, B., and Mukai, H., "Comparison of Two Approaches to On-Line Parameters and State Estimation of Nonlinear Systems," *Industrial and Engineering Chemistry Process and Development*, Vol. 25, No. 3, 1986, pp. 809–814.

[13]Tjoa, I.-B., and Biegler, L. T., "Simultaneous Solution and Optimization Strategies for Parameter Estimation of Differential-Algebraic Equation Systems," *Industrial and Engineering Chemistry Research*, Vol. 30, No. 2, 1991, pp. 376–385.

[14]Robertson, D., and Lee, J. H., "Integrated State Estimation, Fault Detection and Diagnosis for Nonlinear Systems," *Proceedings of the American Control Conference*, Inst. of Electrical and Electronics Engineers, Piscataway, NJ, 1993, pp. 389–392.

[15]Meadows, E. S., and Rawlings, J. B., "Model Predictive Control," *Nonlinear Process Control*, edited by M. A. Henson and D. E. Seborg, Prentice–Hall, Upper Saddle River, NJ, 1997, Chap. 5, pp. 233–310.

[16]Richter, S. L., and DeCarlo, R. A., "Continuation Methods: Theory and Applications," *IEEE Transactions on Automatic Control*, Vol. AC-28, No. 6, 1983, pp. 660–665.

[17]Allgower, E. L., and Georg, K., *Numerical Continuation Methods*, Springer-Verlag, Heiderberg, Germany, 1990.

[18]Ohtsuka, T., and Fujii, H. A., "Nonlinear Receding-Horizon State Estimation by Real-Time Optimization Technique," *Journal of Guidance, Control, and Dynamics*, Vol. 19, No. 4, 1996, pp. 863–870.

[19]Ohtsuka, T., "Nonlinear Receding-Horizon State Estimation with Unknown Disturbances," *Transactions of the Society of Instrument and Control Engineers*, Vol. 35, No. 10, 1999, pp. 1253–1260.

[20]Kelley, C. T., *Iterative Methods for Linear and Nonlinear Equations*, *Frontiers in Applied Mathematics*, Vol. 16, Society for Industrial and Applied Mathematics, Philadelphia, 1995, Chap. 3.

[21]Ohtsuka, T., "A Continuation/GMRES Method for Fast Computation of Nonlinear Receding Horizon Control," *Automatica*, Vol. 40, No. 4, 2004, pp. 563–574.

[22]Seguchi, H., and Ohtsuka, T., "Nonlinear Receding Horizon Control of an Underactuated Hovercraft," *International Journal of Robust and Nonlinear Control*, Vol. 13, No. 3–4, 2003, pp. 381–398.

[23]Bryson, A. E., Jr., and Ho, Y.-C., *Applied Optimal Control*, Hemisphere, New York, 1975, Chap. 13.

[24]Ascher, U. M., Mattheji, R. M. M., and Russel, R. D., *Numerical Solution of Boundary Value Problems for Ordinary Differential Equations*, Society for Industrial and Applied Mathematics, Philadelphia, 1995, Chap. 5.

[25]Barrett, R., Berry, M., Chan, T. F., Demmel, J., Donato, J., Dongarra, J., Eijkhout, V., Pozo, R., Romine, C., and der Vorst, H. V., *Templates for the Solution of Linear Systems: Building Blocks for Iterative Methods*, 2nd ed., Society for Industrial and Applied Mathematics, Philadelphia, PA, 1994, Chap. 2.