

Vision-Based Localization and Mapping for an Autonomous Mower

Junho Yang¹, Soon-Jo Chung¹, Seth Hutchinson¹, David Johnson², and Michio Kise²

Abstract—This paper presents a vision-based localization and mapping algorithm for an autonomous mower. We divide the task for robotic mowing into two separate phases, a teaching phase and a mowing phase. During the teaching phase, the mower estimates the 3D positions of landmarks and defines a boundary in the lawn with an estimate of its own trajectory. During the mowing phase, the location of the mower is estimated using the landmark and boundary map acquired from the teaching phase.

Of particular interest for our work is ensuring that the estimator for landmark mapping will not fail due to the nonlinearity of the system during the teaching phase. A nonlinear observer is designed with pseudo-measurements of each landmark's depth to prevent the map estimator from diverging. Simultaneously, the boundary is estimated with an EKF. Measurements taken from an omnidirectional camera, an IMU, and a ground speed sensor are used for the estimation. Numerical simulations and offline teaching phase experiments with our autonomous mower demonstrate the potential of our algorithm.

I. INTRODUCTION

In this paper, we present a vision-based localization and mapping algorithm for an autonomous mower. The objective for the mowing robot is to maintain turf health and aesthetics of the lawn autonomously by performing complete and uniform coverage of a desired area. To safely achieve this goal, the system must determine and maintain containment inside of the permitted area. Vision sensors are an attractive option for mowing robots for their potential to enable no-infrastructure installations and perform sensing functions beyond positioning such as determining and diagnosing turf problems.

There have been several research prototypes as well as manufactured products developed for robotic lawn mowing [1]–[4]. Boundary wires are widely used to ensure containment in available products [1]. However, they require users to add infrastructure to the environment which increases set-up time and decreases portability. GPS has been widely used for navigation purposes [5] but performs best in a wide open area. It can be difficult to get accurate position estimation results with the GPS in a residential area occluded by walls and tree canopies. RF and infrared signal based methods have also been demonstrated for localization [4], [6] but they can require costly infrastructure.

¹ Junho Yang is with the Department of Mechanical Science and Engineering, Soon-Jo Chung and Seth Hutchinson are with the Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, Urbana, IL, 61801, USA {yang125, sjchung, seth}@illinois.edu

² David Johnson is with the Turf & Utility Advanced Development, John Deere, Cary, NC, 27513, USA and Michio Kise is with the Intelligent Solution Group, John Deere, Urbandale, IA, 50322, USA {johnsondavid2, kisemichio}@johndeere.com



Fig. 1. Our autonomous mower modified with an omnidirectional camera and an IMU for experiments

In this work, we used a robotic mower from John Deere which is shown in Figure 1. Our autonomous mower is equipped with a ground speed sensor and modified with an omnidirectional vision sensor and an IMU. The task for robotic mowing can be divided into two phases, a teaching phase and a mowing phase as shown in Figure 2. During the teaching phase, the mower can follow a boundary wire temporarily set-up in the lawn or can be tele-operated by a user. Our algorithm defines a boundary by estimating the trajectory of the mower with an EKF while generating a point feature-based map of its surrounding landmarks with an observer. We designed a nonlinear observer to estimate the 3D positions of landmarks with respect to the robot's body frame. Hybrid contraction analysis [7]–[9] is used to guarantee the convergence of the estimates with a continuous dynamic model and a discrete measurement model. We use high frequency IMU readings for the dynamic model and slow vision measurements for the measurement model. Instead of using an inverse depth parametrization [10] which is known to cause a negative depth problem [11], [12], we derived pseudo-measurements of each landmark's depth by exploiting different views. During the mowing phase, the landmark and boundary maps can be provided to the mower to estimate the location of the mower. Further, the localization results can be used to determine whether the mower is contained in an area permitted for mowing.

A. Related Work

Simultaneous localization and mapping (SLAM) problems have been studied extensively in the past decade [13]–[15] for autonomous navigation of robots. Robot-centric

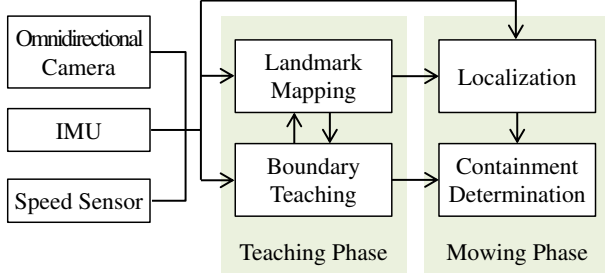


Fig. 2. Block diagram of our localization and mapping strategy developed for an autonomous mowing

estimation opposed to world-centric estimation has also been exploited [16]–[20]. Location estimation with a monocular camera is a challenging problem in part because the distance to a landmark from the camera cannot be estimated with a single measurement. Earlier research [21] solved this problem by sequentially taking measurements from different locations. The position of a landmark was initialized into the map after having enough camera motion to use the Gaussian assumption.

EKF has been widely used for SLAM but it does not cope well with severe nonlinearities. The problem typically occurs when the estimation of landmarks is done in Cartesian coordinates with a monocular camera [22]. Alternatively, the inverse depth parametrization [10] and the anchored homogeneous point method [22] have been proposed. Instead of estimating the location of the landmarks in Cartesian coordinates, the methods use spherical coordinates and homogeneous coordinates respectively. They parameterize each landmark with the location of the camera where it first observes a landmark, and the inverse of the landmark's depth in a direction from the camera to the landmark. The inverse depth is used to alleviate the nonlinearity of the measurement model and let the EKF to perform better with a Gaussian assumption. However, it was found that the estimation results using an inverse depth parametrization can diverge due to linearization error [11]. An iterative Kalman filter was used in [11] to manage the nonlinearity more effectively. In [12], a logarithm-based parametrization was proposed to prevent such a problem.

There have also been various approaches to designing nonlinear observers to solve the structure from motion (SFM) problem with a monocular camera. In [23], a nonlinear observer was designed with Lyapunov theory to estimate the range of a point feature considering the use of a monocular camera and an inertial sensor. The system was composed of the normalized coordinates of a point feature and the inverse of its depth along the optical axis of the camera. The work was extended in [24] to a structure and motion problem where a reduced-order observer was designed to estimate the motion of the camera as well as the depth of a point feature. They assumed that only one of the linear velocities and a corresponding acceleration of the camera are

known. A new parametrization that involves projection onto a sphere was introduced in [25]. An adaptive observer was designed for the perspective system to estimate the inverse depth of a point feature and identify the linear velocity of the camera. In [26], a locally asymptotically stable observer was designed for a similar perspective system to estimate a constant initial depth of a point feature and the angular velocity of the perspective system. Their method integrates the motion of the camera to recover the time-varying depth in an open loop fashion. In [27], a nonlinear observer was designed to estimate the 6-DOF pose of a vehicle in the special Euclidean group $SE(3)$ with bearing measurements of features and inertial measurements. The locations of the features were assumed to be known.

The work presented in this paper establishes a nonlinear estimation framework to the localization and mapping problem for autonomous mowing. The mapping portion of our work is in the scope of observer based structure from motion research. But in contrast with the aforementioned prior work, we derive pseudo-measurements of feature depth instead of using an unmeasurable state of the inverse depth. We then use hybrid contraction analysis [7]–[9] to design a nonlinear observer and guarantee its global exponential stability. In parallel, we estimate the states of the mower with a robot-centric EKF estimator. The estimated location of the mower is used to define a boundary in the lawn during the teaching phase. This information can be used by the mower for positioning and determining containment in the mowing phase.

B. Organization

The rest of this paper is organized as follows. In Section II, we describe the dynamics and measurements for landmark mapping, and design a nonlinear observer with hybrid contraction analysis. In Section III, an EKF estimator is presented to estimate the trajectory of the mower and define a boundary in the lawn. In Section IV, the landmarks are used to estimate the location of the autonomous mower and solve the containment problem with the boundary information. Numerical simulations are shown in Section V. Offline experiments of the teaching phase are presented in Section VI. We conclude with plans for future work in Section VII.

II. OBSERVER DESIGN FOR ROBOT-CENTRIC LANDMARK MAPPING

In this section, we describe the dynamics of the vision system for robot-centric mapping and derive pseudo-measurements of the depth of each landmark. A nonlinear observer is designed for mapping, and the convergence of the estimates is proved with hybrid contraction analysis.

A. Dynamic Model for Landmark Mapping

The dynamic model of each landmark in the robot's body frame is given by [28]

$$\dot{\mathbf{x}}_i = -[\boldsymbol{\omega}]_{\times} \mathbf{x}_i - \mathbf{v} \quad (1)$$

where $\mathbf{x}_i \in \mathbb{R}^3$ is the location of the i -th landmark with respect to the robot's body frame. The x -axis of the robot's

body frame is pointing towards the front of the mower, and the z-axis is pointing up from the mower. For simplicity, all the variables without a superscript are in the robot's body frame.

The linear and angular velocities of the robot measured in the robot's body frame are denoted by $\mathbf{v} \in \mathbb{R}^3$ and $\boldsymbol{\omega} \in \mathbb{R}^3$. The skew-symmetric matrix $[\boldsymbol{\omega}]_{\times} \in so(3)$ is formed from the angular velocity vector $\boldsymbol{\omega}$.

We estimate the location of each landmark in the robot's body frame and let the measurements be linear with respect to the states. Similar to [26], a landmark can be described with a unit vector $\mathbf{y} = \mathbf{x}_i / \|\mathbf{x}_i\|_2 \in \mathbb{R}^3$ from the robot and its distance $d = \|\mathbf{x}_i\|_2 \in \mathbb{R}$. The state vector of each landmark is $\mathbf{z}^T = (d, \mathbf{y}^T)^T$, and the dynamics of the system is given by

$$\frac{d}{dt} \begin{pmatrix} d \\ \mathbf{y} \end{pmatrix} = \begin{pmatrix} -\mathbf{v}^T \mathbf{y} \\ -(I - \mathbf{y} \mathbf{y}^T) \mathbf{v} d^{-1} - [\boldsymbol{\omega}]_{\times} \mathbf{y} \end{pmatrix} + \boldsymbol{\eta} \quad (2)$$

where $\boldsymbol{\eta} \in \mathbb{R}^4$ denotes the disturbance.

Assumption 1: The Euclidean distance d between the camera and a landmark is lower bounded by a known positive constant. Therefore, we assume that $\underline{d} \leq d$, where $\underline{d} \in \mathbb{R}^+$ is a known constant parameter.

Remark 1: Assumption 1 is satisfied due to physical constraints of the system.

Measurements of system (2) is linear to the states. The unit vector $\mathbf{y}$ is directly measurable, whereas the depth d of a landmark is not directly measurable from a single image. However, pseudo-measurements of d can be formulated as described in Section II-B.

B. Pseudo-Measurements of a Landmark's Depth

We formulate pseudo-measurements $d_k \in \mathbb{R}$ at current timestep k to acquire the depth of each landmark as

$$d_k = \|\hat{\mathbf{x}}_b^{b_j}\| \left(1 - \left(\frac{\hat{\mathbf{x}}_b^{b_j}}{\|\hat{\mathbf{x}}_b^{b_j}\|} \cdot \mathbf{y}_j \right)^2 \right)^{\frac{1}{2}} \left(1 - (R^T(\hat{\mathbf{q}}_{b_j}) \mathbf{y}_k \cdot \mathbf{y}_j)^2 \right)^{-\frac{1}{2}} + \xi_{d,k} \quad (3)$$

where $\mathbf{y}_k$ and $\mathbf{y}_j$ are the direction vector measurements of a landmark at timestep k and at a previous timestep j . Noise in the pseudo-measurements d_k is denoted by $\xi_{d,k} \in \mathbb{R}$, and $R(\cdot)$ is a rotation matrix. In Section III, we estimate the location $\hat{\mathbf{x}}_{b_j}$ and orientation quaternions $\hat{\mathbf{q}}_{b_j}$ of the robot's body frame at timestep j with respect to its current body frame. The current location of the mower with respect to its body frame at timestep j , which is used in (3), is $\hat{\mathbf{x}}_b^{b_j} = -R^T(\mathbf{q}_{b_j}) \hat{\mathbf{x}}_{b_j}$.

C. Observer Design with Hybrid Contraction Analysis

The observer presented in this section updates the estimates of the states using vision measurements at discrete-time instances and propagates the motion between the measurements in continuous-time. We use dwell-time $\Delta t_k = t_k - t_{k-1}$ for vision measurements $\mathbf{y}_k$ since image processing can be much slower than the inertial measurements which is used in the dynamic model. One can also consider using

the direction estimates $\hat{\mathbf{y}}_k$ to improve the vision tracking algorithm. We can allow sufficiently long dwell-time to track landmarks which are instantaneously occluded in images.

Estimation of landmarks is decoupled by using separate observers. This gives us a potential to increase the number of landmarks for the map. It was shown in [29] that increasing the number of landmarks is more profitable than increasing the measurement rate in order to enhance the accuracy of the estimates. We prove that our observer is guaranteed to be globally exponentially stable by using contraction theory.

1) Observer Design and Stability Analysis: The observer for each landmark is given by

$$\frac{d}{dt} \begin{pmatrix} \hat{d} \\ \hat{\mathbf{y}} \end{pmatrix} = \begin{pmatrix} -\mathbf{v}^T \hat{\mathbf{y}} \\ -(I - \hat{\mathbf{y}} \hat{\mathbf{y}}^T) \mathbf{v} \hat{d}^{-1} - [\boldsymbol{\omega}]_{\times} \hat{\mathbf{y}} \end{pmatrix} \quad (4)$$

$$\begin{pmatrix} \hat{d}_k^+ \\ \hat{\mathbf{y}}_k^+ \end{pmatrix} = \begin{pmatrix} \hat{d}_k^- + l_{k,1}(d_k - \hat{d}_k^-) \\ \hat{\mathbf{y}}_k^- + l_{k,2}(\mathbf{y}_k - \hat{\mathbf{y}}_k^-) \end{pmatrix} \quad (5)$$

where $\hat{d} \in \mathbb{R}$ is an estimate of a landmark's depth between vision measurements, $\hat{d}_k^- \in \mathbb{R}$ and $\hat{d}_k^+ \in \mathbb{R}$ are estimates of the depth before and after the measurement update at timestep k , $\hat{\mathbf{y}} \in \mathbb{R}^3$ is an estimate of a unit vector from the robot's body frame to the landmark between its measurements, $\hat{\mathbf{y}}_k^- \in \mathbb{R}^3$ and $\hat{\mathbf{y}}_k^+ \in \mathbb{R}^3$ are estimates of the unit vector before and after the measurement update at timestep k , and $I \in \mathbb{R}^{3 \times 3}$ is an identity matrix. The user can select observer gains $l_{k,1} \in \mathbb{R}$ and $l_{k,2} \in \mathbb{R}$ for $\hat{d}_k$ and $\hat{\mathbf{y}}_k$ respectively.

The continuous system (4) for prediction of the states $\hat{\mathbf{z}}^T = (\hat{d}, \hat{\mathbf{y}}^T)^T \in \mathbb{R}^4$ is switched to the discrete system (5) at every Δt_k to update the states with measurements.

Estimation error for the hybrid system is defined as $\mathbf{e}_k \triangleq \mathbf{z}_k - \hat{\mathbf{z}}_k \in \mathbb{R}^4$ where $\mathbf{z}_k \in \mathbb{R}^4$ is the ground truth of the states.

Theorem 1: The observer in (4) and (5) is globally exponentially stable such that

$$\|\mathbf{e}_{k+1}\| \leq \bar{\sigma}_k \|\mathbf{e}_k\| \exp(\bar{\lambda} \Delta t_k) \quad (6)$$

if Assumption 1 is satisfied and the states are constrained by $\|\hat{\mathbf{y}}\| = 1$ and $\hat{d} > \underline{d}$, and if the observer gains are given by

$$l_{k,m} = 1 - \exp\left(\frac{1}{2}(\gamma_m - \bar{\lambda}) \Delta t_k\right), \quad m \in \{1, 2\} \quad (7)$$

where $\gamma_m \in \mathbb{R}^-$ is defined by the user for $\hat{d}_k$ and $\hat{\mathbf{y}}_k$. The convergence rate of the system in the prediction stage (4) is given by $\bar{\lambda} = \lambda_{\max}(F^T + F)$, where $\lambda_{\max}$ denotes the maximum eigenvalue and $F \in \mathbb{R}^{4 \times 4}$ is the Jacobian matrix in (8).

Proof: We can write the first variation of system (4) and (5) as

$$\frac{d}{dt} \begin{pmatrix} \delta \hat{d} \\ \delta \hat{\mathbf{y}} \end{pmatrix} = \begin{pmatrix} 0 \\ (I - \hat{\mathbf{y}} \hat{\mathbf{y}}^T) \mathbf{v} \hat{d}^{-2} - [\boldsymbol{\omega}]_{\times} + (\hat{\mathbf{y}} \mathbf{v}^T + \mathbf{v}^T \hat{\mathbf{y}} I) \hat{d}^{-1} \end{pmatrix} \begin{pmatrix} \delta \hat{d} \\ \delta \hat{\mathbf{y}} \end{pmatrix} \quad (8)$$

$$\begin{pmatrix} \delta \hat{d}_k^+ \\ \delta \hat{\mathbf{y}}_k^+ \end{pmatrix} = \begin{pmatrix} 1 - l_{k,1} & 0 \\ 0 & I(1 - l_{k,2}) \end{pmatrix} \begin{pmatrix} \delta \hat{d}_k^- \\ \delta \hat{\mathbf{y}}_k^- \end{pmatrix} \quad (9)$$

where the virtual displacement $\delta \hat{d} \in \mathbb{R}$ and $\delta \hat{\mathbf{y}} \in \mathbb{R}^3$ are infinitesimal displacements [7] at a fixed time instance and

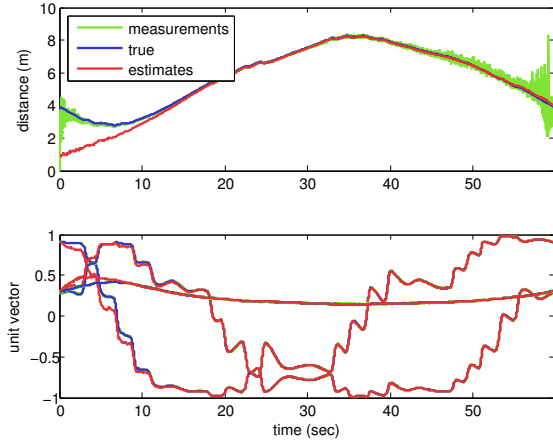


Fig. 3. Simulation results of landmark depth and direction estimation

$I \in \mathbb{R}^{3 \times 3}$. Let $F_k \in \mathbb{R}^{4 \times 4}$ denote the Jacobian matrix in (9), and let $\gamma = \max\{\gamma_1, \gamma_2\}$. The convergence rate of the measurement update stage is given by

$$\bar{\sigma}_k = \lambda_{\max}(F_k^T F_k) = (1 - l_k)^2 \quad (10)$$

where $l_k = l_{k,m}$ with $\gamma_m = \gamma$.

Consider the observer given by (4) and (5). The condition for the hybrid system to be contracting [7]–[9] is satisfied since

$$\bar{\lambda} + \frac{\ln \bar{\sigma}_k}{\Delta t_k} = \bar{\lambda} + \frac{\ln \exp(\gamma \Delta t_k - \bar{\lambda} \Delta t_k)}{\Delta t_k} = \gamma \quad (11)$$

where γ is selected to be negative. We then have

$$\|\delta \hat{\mathbf{z}}_{k+1}\| \leq \bar{\sigma}_k \|\delta \hat{\mathbf{z}}_k\| \exp(\bar{\lambda} \Delta t_k) \quad (12)$$

and since $\delta \hat{\mathbf{z}}_k$ tends to zero, the estimated states converge to their true values globally exponentially fast. ■

2) *Uncertainty Bound on the Estimation Error:* Uncertainty bound on estimation error can be analyzed by considering the disturbance η and measurement noise $\xi_k = (\xi_{d,k}, \xi_{y,k}^T)^T \in \mathbb{R}^4$, where $\xi_{y,k} \in \mathbb{R}^3$ is the noise in unit sphere vision measurements. Let $\bar{\mathbf{e}}_k = \int_{\bar{\mathbf{z}}}^{\mathbf{z}} \|\delta \mathbf{x}_k\| \in \mathbb{R}$ be the quadratic bound of the observer error [7] which considers the uncertainties. Then

$$\bar{\mathbf{e}}_{k+1} \leq \bar{\sigma}_k \bar{\mathbf{e}}_k \exp(\bar{\lambda} \Delta t_k) + \|\eta \Delta t_k + l_k \xi_k\|_{\infty} \quad (13)$$

where l_k is the observer gain. The observer gain can be designed to take into account the magnitude of the noise in the measurements. When the observer gain l_k is increased, the estimation error converges to zero faster and the estimates are less sensitive to disturbance η . When the observer gain l_k is decreased, the estimates will be less affected by the measurement noise ξ_k .

III. ROBOT-CENTRIC LOCALIZATION FOR BOUNDARY ESTIMATION

In this section, we estimate the trajectory of the mower with a robot-centric system which allows us to use a linear

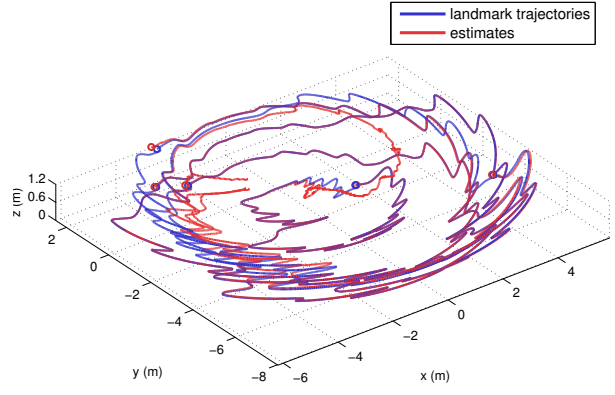


Fig. 4. Simulation results of estimation of landmarks in the robot's body frame

motion model. The mower traverses a boundary in the lawn during the teaching phase. It can either follow a boundary wire temporarily set-up in the lawn or be tele-operated by a user.

Consider a model with a state vector $\mathbf{x} = ((\mathbf{x}_{b_j})^T, (\mathbf{q}_{b_j})^T)^T$, where $\mathbf{x}_{b_j} \in \mathbb{R}^3$ and $\mathbf{q}_{b_j} \in \mathbb{R}^4$ are the location and orientation quaternions of the mower at an instance for timestep j with respect to its current body frame.

The motion model is given by

$$\frac{d}{dt} \begin{pmatrix} \mathbf{x}_{b_j} \\ \mathbf{q}_{b_j} \end{pmatrix} = \begin{pmatrix} -[\omega]_{\times} & 0 \\ 0 & \frac{1}{2} \Omega(\omega) \end{pmatrix} \begin{pmatrix} \mathbf{x}_{b_j} \\ \mathbf{q}_{b_j} \end{pmatrix} - \begin{pmatrix} \mathbf{v} \\ 0 \end{pmatrix} \quad (14)$$

where $\Omega(\omega) \in \mathbb{R}^{4 \times 4}$ is a skew symmetric matrix

$$\Omega(\omega) = \begin{pmatrix} 0 & \omega^T \\ -\omega & -[\omega]_{\times} \end{pmatrix} \quad (15)$$

The measurement model $h(\mathbf{x})$ is a stacked vector of measurements of each landmark given by

$$\begin{aligned} h_i(\mathbf{x}) &= \mathbf{y}_i^{b_j} \\ &= \mathbf{x}_i^{b_j} / \|\mathbf{x}_i^{b_j}\|_2, \quad \forall i \in \{1, 2, \dots, n\} \end{aligned} \quad (16)$$

Here, n is the number of landmarks. The direction vector measurement $\mathbf{y}_i^{b_j}$ of the i -th landmark is taken when the landmark was first observed at timestep j .

Note that $\mathbf{x}_i^{b_j} = R^T(\mathbf{q}_{b_j})(\mathbf{x}_i - \mathbf{x}_{b_j})$, where, $\mathbf{x}_i = \mathbf{y}_i d \in \mathbb{R}^3$ is the position of the i -th landmark we estimate in the robot's body frame with the nonlinear observer given by (4) and (5).

An EKF estimator for boundary estimation is written as

$$\begin{aligned} \dot{\hat{\mathbf{x}}} &= A \hat{\mathbf{x}} + \mathbf{u} + K(\mathbf{h} - h(\hat{\mathbf{x}})) \\ \dot{P} &= AP + PA^T - PH^T V^{-1} H P + W \end{aligned} \quad (17)$$

where P is the covariance of the states, $A \in \mathbb{R}^{7 \times 7}$ is the state transition matrix in (14), $\mathbf{u} = (-\mathbf{v}^T, 0)^T \in \mathbb{R}^7$ is the velocity input, $\mathbf{h}$ is the vision measurement, H is the Jacobian of the measurement function $h(\hat{\mathbf{x}})$, and V and W are the covariance matrices that approximate the measurement noise and the process noise.

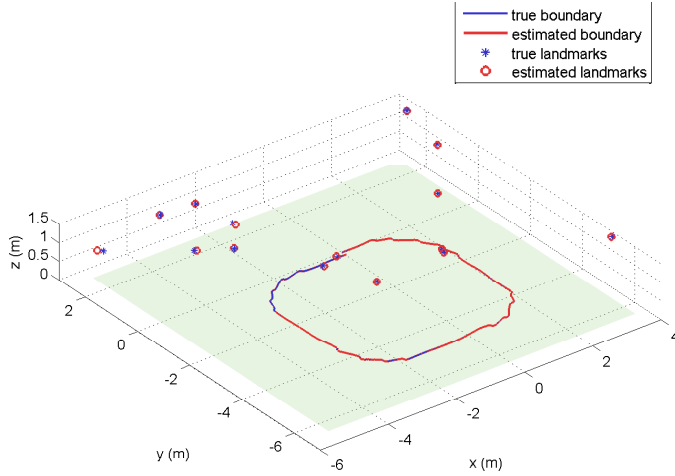


Fig. 5. Simulation results of boundary estimation and landmark mapping

The estimator gain K is given by

$$K = PH^T V^{-1} \quad (18)$$

The location and orientation of the world reference frame with respect to the current body frame are updated by

$$\begin{aligned} \hat{\mathbf{x}}_w &= \hat{\mathbf{x}}_{b_j} + R(\hat{\mathbf{q}}_{b_j}) \hat{\mathbf{x}}_w^{b_j} \\ \hat{\mathbf{q}}_w &= \hat{\mathbf{q}}_{b_j} \otimes \hat{\mathbf{q}}_w^{b_j} \end{aligned} \quad (19)$$

where $\otimes$ is a quaternion multiplication.

Finally, the results of mower localization $\hat{\mathbf{x}}_b^w$ and landmark mapping $\hat{\mathbf{x}}_i^w$ are represented in the world reference frame by

$$\begin{pmatrix} \hat{\mathbf{x}}_b^w \\ \hat{\mathbf{x}}_i^w \end{pmatrix} = \begin{pmatrix} -R^T(\hat{\mathbf{q}}_w) \hat{\mathbf{x}}_w \\ R^T(\hat{\mathbf{q}}_w) (\hat{\mathbf{x}}_i - \hat{\mathbf{x}}_w) \end{pmatrix} \quad (20)$$

The boundary can be defined in the lawn based on the history of the estimated trajectory $\hat{\mathbf{x}}_b^w$ of the mower.

IV. LOCALIZATION DURING AUTONOMOUS MOWING

During the mowing phase, localization of an autonomous mower can be performed with the information of landmarks acquired through the teaching phase. The estimated location of the autonomous mower can be used to determine whether the mower is contained inside the estimated boundary.

The state vector of the autonomous mower is given by $((\mathbf{x}_b^w)^T, (\mathbf{q}_b^w)^T)^T$, where $\mathbf{x}_b^w$ is the current location of the mower, and $\mathbf{q}_b^w$ is its orientation in quaternions. The pose $\mathbf{x}_b^w$ and $\mathbf{q}_b^w$ are described in the world reference frame.

The motion model of the system is given by

$$\frac{d}{dt} \begin{pmatrix} \mathbf{x}_b^w \\ \mathbf{q}_b^w \end{pmatrix} = \begin{pmatrix} R(\mathbf{q}_b^w) \mathbf{v} \\ \frac{1}{2} \Omega(\omega) \mathbf{q}_b^w \end{pmatrix} \quad (21)$$

The measurement model $g(\mathbf{x}_b^w, \mathbf{q}_b^w)$ is a stacked vector of measurements of each landmark given by

$$\begin{aligned} g_i(\mathbf{x}_b^w, \mathbf{q}_b^w) &= \mathbf{y}_i \\ &= \mathbf{x}_i / \|\mathbf{x}_i\|_2, \quad \forall i \in \{1, 2, \dots, n\} \end{aligned} \quad (22)$$

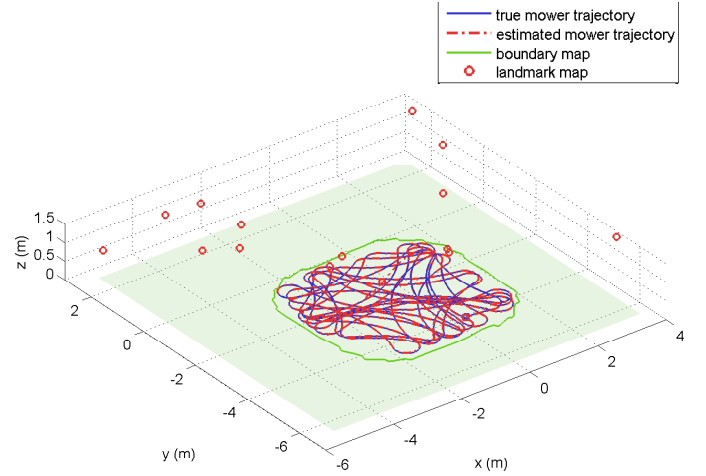


Fig. 6. Simulation results of robot containment based on localization with information provided from the teaching phase

where the i -th landmark in the robot's body frame $\mathbf{x}_i$ is

$$\mathbf{x}_i = R(\mathbf{q}_w) (\mathbf{x}_i^w - \mathbf{x}_b^w) \quad (23)$$

the location of the i -th landmark $\mathbf{x}_i^w$ in the world reference frame is provided by the teaching phase results (20).

The states of the autonomous mower can be estimated with an EKF estimator using the motion model (21) and the measurement model (22). A world-centric representation is used since the location of the landmarks $\mathbf{x}_i^w$ are provided by the teaching phase algorithm and the problem becomes a standard localization problem.

Containment of the robot can be determined by applying the estimated boundary and the estimate of the mower's current location to a point-in-polygon algorithm [30]. Consider spreading a set of rays from the mower's estimated location. The number of times the rays encounter the predefined boundary can be denoted as a winding number when the boundary is a single loop. If the winding number is odd, the mower is determined to be contained inside the boundary and it is permitted to continue mowing. If the winding number is even, the mower is outside of the boundary and mowing should be halted.

V. SIMULATION RESULTS

Numerical simulations are presented in this section to analyze the performance of our algorithm. We distributed 15 simulated landmark points randomly in a 3D space. Gaussian white noise with standard deviation of 3 was added to the camera pixel measurements.

Figure 3 shows the estimated depth and direction vector of one of the landmarks converging to their true values. Figure 4 shows the trajectories of the landmarks in the robot's body frame converging to their true position. The motion of the robot and the scene can be understood when the estimates are converted to the world reference frame through (20). Figure 5 shows simulation results of boundary estimation and landmark mapping represented in the world reference frame. The estimated boundary follows the true trajectory

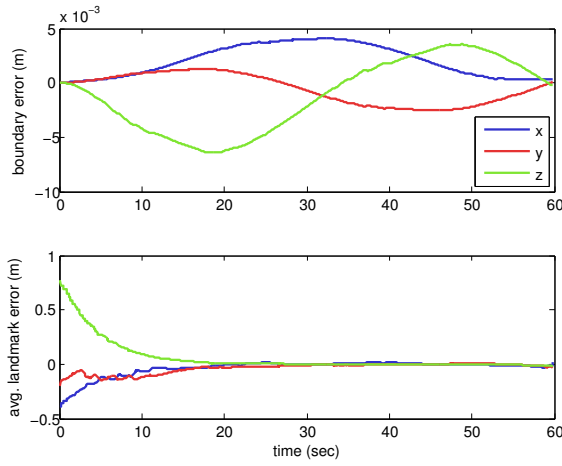


Fig. 7. Estimation error in the location of the boundary and average of the landmarks' positions during the teaching phase

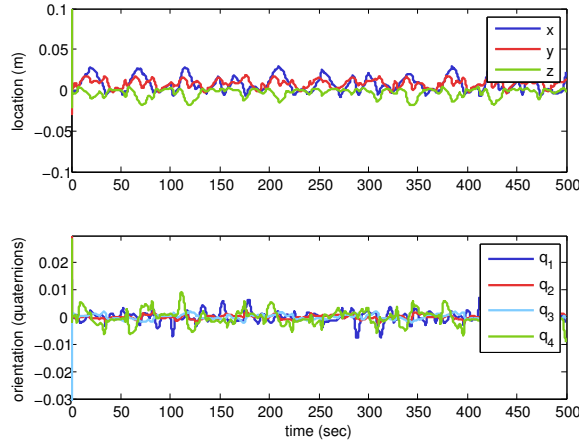


Fig. 8. Estimation error in the location and orientation quaternions of the mower during the mowing phase

of the mower, and the estimated positions of the landmarks converge to their true locations. The estimated trajectory of the mower is red, and its true trajectory is blue. The red circles are the estimated positions of the landmarks, and the blue stars are their true positions.

Figure 6 shows simulation results of containment during autonomous mowing. The estimated trajectory of the robot is red, and its true trajectory is blue. The boundary estimated during the teaching phase is green. The red circles are the positions of the landmarks estimated in the teaching phase. We applied bang-bang control and randomly changed the heading direction of the mower when it approached the boundary. Threshold value of 0.4m was used in the simulation. Once the robot declared that it is inside the boundary, the mower randomly covered the given area while estimating the location of itself in the map. It is shown in Figure 6 that the estimated trajectory of the robot converges to its true trajectory.

Figure 7 shows the estimation error of the mowers location and orientation that are used to define the boundary and



Fig. 9. Our autonomous mower following a boundary set-up in the backyard of our research building for map estimation and boundary teaching



Fig. 10. Tracking landmarks in a sequence of omnidirectional camera images

generate the map during the teaching phase. Figure 8 shows the error in the estimated location and quaternion orientation of the mower during the mowing phase. The errors converged towards zero rapidly but oscillated continually because the mower abruptly changed its heading direction whenever it approached the boundary.

VI. TEACHING PHASE EXPERIMENTS

In this section, we demonstrate the potential of our teaching algorithm using a data set collected with our autonomous mower. Our autonomous mower is equipped with a 0-360 Panoramic Optic omnidirectional camera which can persistently capture landmarks in the scene with 360 degree field of view. A VM-100 Rugged IMU is mounted on the bottom of the camera. Ground speed measurements are provided from the mower. During the teaching phase, our mower followed a boundary wire set-up in the lawn as shown in Figure 9. The mower made a full loop while traversing a

Parameter	Value
Focal Length (f_u, f_v)	(476.60667, 476.74991)
Principal Point (u_0, v_0)	(775.11715, 778.91684)
Mirror Transformation ξ	0.92036
Skew α	0
Distortion (k_1, k_2, k_3, k_4)	(-0.17357, 0.02025, -0.00209, 0.00091, 0)
Pixel Error (e_x, e_y)	(0.50050, 0.51821)

TABLE I
CAMERA CALIBRATION RESULTS

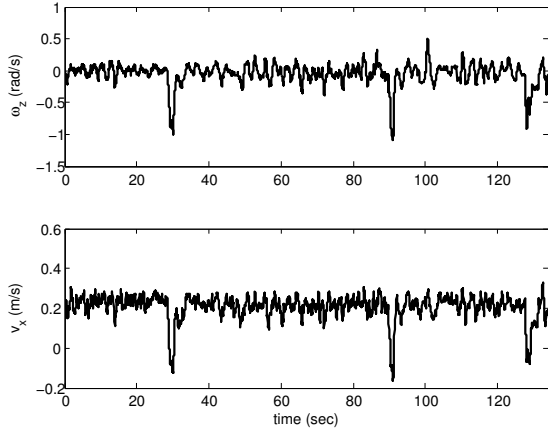


Fig. 11. Angular and linear velocities of the mower collected during the experiments

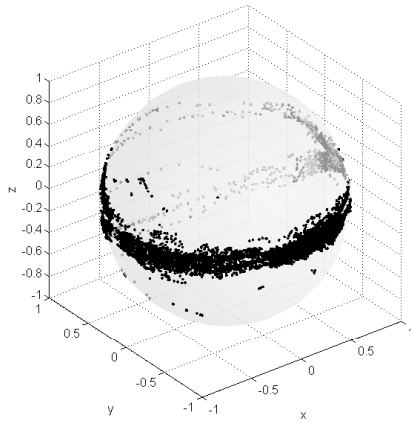


Fig. 12. Unit sphere projection of landmark measurements at each timestep

boundary and captured 1540×1540 pixels omnidirectional camera images at 4Hz. We collected inertial measurements at 100Hz and filtered the measurements from the ground speed sensor at the same rate. Figure 11 shows the angular and linear velocities measured from the IMU and the ground speed sensor.

To demonstrate our algorithm, we manually selected 10 corners from the windows near the lawn as landmarks and tracked the points with the pyramid Lucas Kanade optical flow method [31]. Figure 10 shows a set of images collected from our autonomous mower. The landmarks are marked in Figure 10 with red dots. To extract the direction vector measurements $\mathbf{y}$, the pixel coordinates $\mathbf{p} = (p_u, p_v, 1)^T$ of each feature were transformed to normalized image coordinates with $\mathbf{p}_n = C^{-1}\mathbf{p} = (p_x, p_y, 1)^T$, where C is the camera projection matrix. The calibration parameters for our camera are shown in Table I. The normalized coordinates $\mathbf{p}_n$ were projected on a unit sphere through the method described in

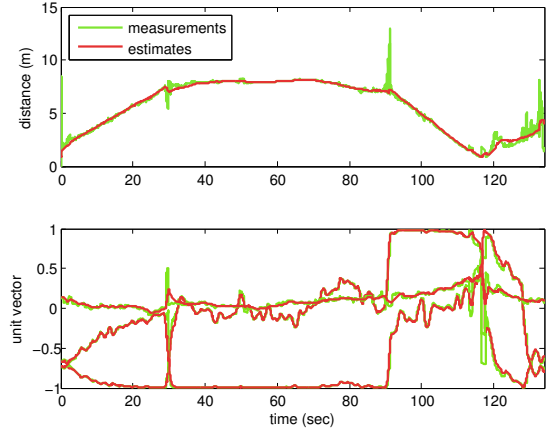


Fig. 13. Experimental results of landmark depth and direction estimation

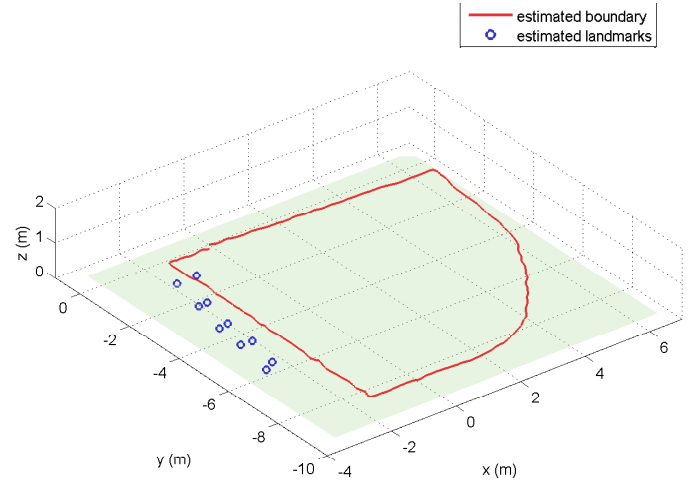


Fig. 14. Experimental results of boundary teaching and landmark mapping with the data set collected with our autonomous mower

[32] which is given by

$$\mathbf{y} = \begin{pmatrix} \frac{\zeta + \sqrt{1 + (1 - \zeta^2)(p_x^2 + p_y^2)}}{p_x^2 + p_y^2 + 1} p_x \\ \frac{\zeta + \sqrt{1 + (1 - \zeta^2)(p_x^2 + p_y^2)}}{p_x^2 + p_y^2 + 1} p_y \\ \frac{\zeta + \sqrt{1 + (1 - \zeta^2)(p_x^2 + p_y^2)}}{p_x^2 + p_y^2 + 1} - \zeta \end{pmatrix} \quad (24)$$

where ζ is a mirror transformation parameter. Figure 12 shows the unit sphere projection of the landmarks measured from the omnidirectional camera at each timestep.

Figure 13 shows the depth of a landmark estimated with pseudo-measurements and the direction vector of the landmark filtering the unit sphere vision measurements. Figure 14 demonstrates the landmark mapping and boundary teaching with the data set collected using our mower. The estimated boundary, which the mower traversed, is marked with red. The blue circles are the estimated positions of the landmarks. Shape of the estimated boundary and the estimated positions of the landmarks resemble the actual configuration. We plan to acquire ground truth information to analyze the

experimental results quantitatively in the near future.

VII. CONCLUSION

A vision based localization and mapping algorithm for an autonomous mower was presented in this paper. A nonlinear observer was designed for robot-centric landmark mapping, and a boundary estimation strategy using localization results was described. We proposed to use the estimated boundary and landmark map to estimate the location of the mower for autonomous mowing. Numerical simulations illustrated the convergence of the estimates and the capability of using the estimates for containment of the mower. Preliminary experimental results showed boundary estimation and landmark mapping with a set of data collected with our autonomous mower.

In the future, we plan to use a larger set of vision data with a more robust tracking algorithm. Loop closing techniques will also be applied to further improve the estimation results. We aim to provide the experimental data processed through the teaching phase to the mower and demonstrate fully autonomous mowing and containment with no infrastructure.

ACKNOWLEDGMENT

This material is based on work supported by John Deere. The authors acknowledge Dr. Ashwin Dani for useful discussions on designing the observer. The authors also thank Colin Das for his help on the experiments.

REFERENCES

- [1] ConsumerSearch, "Best robotic mowers," [accessed 5-June-2013]. [Online]. Available: <http://www.consumersearch.com/robotic-lawn-mowers/best-robotic-mowers>
- [2] R. W. Hicks II and E. L. Hall, "Survey of robot lawn mowers," in *Proc. SPIE Intelligent Robots and Computer Vision XIX: Algorithms, Techniques, and Active Vision*, Oct. 2000, pp. 262–269. [Online]. Available: <http://dx.doi.org/10.1117/12.403770>
- [3] H. Sahin and L. Guvenç, "Household robotics: autonomous devices for vacuuming and lawn mowing [applications of control]," *IEEE Control Systems*, vol. 27, no. 2, pp. 20–96, 2007.
- [4] A. Smith, H. Chang, and E. Blanchard, "An outdoor high-accuracy local positioning system for an autonomous robotic golf greens mower," in *Proc. IEEE International Conference on Robotics and Automation*, May 2012, pp. 2633–2639.
- [5] S. Sukkarieh, E. Nebot, and H. Durrant-Whyte, "A high integrity IMU/GPS navigation loop for autonomous land vehicle applications," *IEEE Transactions on Robotics and Automation*, vol. 15, no. 3, pp. 572–578, June 1999.
- [6] B. Ferris, D. Haehnel, and D. Fox, "Gaussian processes for signal strength-based location estimation," in *Proc. Robotics: Science and Systems*, Philadelphia, USA, Aug. 2006.
- [7] W. Lohmiller and J.-J. E. Slotine, "Nonlinear process control using contraction theory," *AICHE Journal*, vol. 46, pp. 588–596, 2000.
- [8] K. Rifai and J.-J. Slotine, "Compositional contraction analysis of resetting hybrid systems," *IEEE Transactions on Automatic Control*, vol. 51, no. 9, pp. 1536–1541, 2006.
- [9] Y. Zhao and J.-J. E. Slotine, "Discrete nonlinear observers for inertial navigation," *Systems & Control Letters*, vol. 54, no. 9, pp. 887–898, 2005.
- [10] J. Civera, A. Davison, and J. Montiel, "Inverse depth parametrization for monocular SLAM," *IEEE Transactions on Robotics*, vol. 24, no. 5, pp. 932–945, 2008.
- [11] S. Tully, H. Moon, G. Kantor, and H. Choset, "Iterated filters for bearing-only SLAM," in *Proc. IEEE International Conference on Robotics and Automation*, May 2008, pp. 1442–1448.
- [12] M. Parsley and S. Julier, "Avoiding negative depth in inverse depth bearing-only SLAM," in *Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems*, Sept. 2008, pp. 2066–2071.
- [13] H. Durrant-Whyte and T. Bailey, "Simultaneous localization and mapping: part I," *Robotics Automation Magazine, IEEE*, vol. 13, no. 2, pp. 99–110, 2006.
- [14] T. Bailey and H. Durrant-Whyte, "Simultaneous localization and mapping (SLAM): part II," *IEEE Robotics Automation Magazine*, vol. 13, no. 3, pp. 108–117, 2006.
- [15] H. Choset, K. M. Lynch, S. Hutchinson, G. A. Kantor, W. Burgard, L. E. Kavraki, and S. Thrun, *Principles of Robot Motion: Theory, Algorithms, and Implementations*. Cambridge, MA: MIT Press, 2005.
- [16] J. Yang, A. Dani, S.-J. Chung, and S. Hutchinson, "Inertial-aided vision-based localization and mapping in a riverine environment with reflection measurements," in *Proc. AIAA Guidance, Navigation, and Control Conference*, August 2013, to appear.
- [17] J. A. Castellanos, R. Martinez-cantin, J. D. Tardos, and J. Neira, "Robocentric map joining: Improving the consistency of EKF-SLAM," *Robotics and Autonomous Systems*, vol. 55, pp. 21–29, 2007.
- [18] A. Boberg, A. Bishop, and P. Jensfelt, "Robocentric mapping and localization in modified spherical coordinates with bearing measurements," in *Proc. International Conference on Intelligent Sensors, Sensor Networks and Information Processing*, Dec. 2009, pp. 139–144.
- [19] J. Civera, O. G. Grasa, A. J. Davison, and J. M. M. Montiel, "1-point RANSAC for extended Kalman filtering: Application to real-time structure from motion and visual odometry," *Journal of Field Robotics*, vol. 27, no. 5, pp. 609–631, 2010.
- [20] B. Williams and I. Reid, "On combining visual SLAM and visual odometry," in *Proc. IEEE International Conference on Robotics and Automation*, May 2010, pp. 3494–3500.
- [21] A. Davison, I. Reid, N. Molton, and O. Stasse, "MonoSLAM: Real-time single camera SLAM," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 6, pp. 1052–1067, 2007.
- [22] J. Sola, T. Vidal-Calleja, J. Civera, and J. M. M. Montiel, "Impact of landmark parametrization on monocular EKF-SLAM with points and lines," *The International Journal of Computer Vision*, vol. 97, no. 3, pp. 339–368, 2012.
- [23] W. Dixon, Y. Fang, D. Dawson, and T. Flynn, "Range identification for perspective vision systems," *IEEE Transactions on Automatic Control*, vol. 48, no. 12, pp. 2232–2238, 2003.
- [24] A. Dani, N. Fischer, and W. Dixon, "Single camera structure and motion," *IEEE Transactions on Automatic Control*, vol. 57, no. 1, pp. 238–243, 2012.
- [25] O. Dahl, F. Nyberg, and A. Heyden, "Nonlinear and adaptive observers for perspective dynamic systems," in *Proc. American Control Conference*, July 2007, pp. 966–971.
- [26] A. Heyden and O. Dahl, "Provably convergent structure and motion estimation for perspective systems," in *Proc. IEEE Conference on Decision and Control*, Dec. 2009, pp. 4216–4221.
- [27] G. Baldwin, R. Mahony, and J. Trumpf, "A nonlinear observer for 6 DOF pose estimation from inertial and bearing measurements," in *Proc. IEEE International Conference on Robotics and Automation*, May 2009, pp. 2237–2242.
- [28] S. Hutchinson, G. Hager, and P. Corke, "A tutorial on visual servo control," *IEEE Transactions on Robotics and Automation*, vol. 12, no. 5, pp. 651–670, Oct. 1996.
- [29] H. Strasdat, J. Montiel, and A. Davison, "Real-time monocular SLAM: Why filter?" in *Proc. IEEE International Conference on Robotics and Automation*, May 2010, pp. 2657–2664.
- [30] W. Li, E. Ong, S. Xu, and T. Hung, "A point inclusion test algorithm for simple polygons," *Lecture Notes in Computer Science*, vol. 3480, pp. 769–775, 2005.
- [31] J. Bouguet, "Pyramidal implementation of the Lucas Kanade feature tracker: Description of the algorithm. intel corporation microprocessor research labs," *OpenCV documents*, 2002.
- [32] C. Mei and P. Rives, "Single view point omnidirectional camera calibration from planar grids," in *Proc. IEEE International Conference on Robotics and Automation*, April 2007, pp. 3945–3950.