Observer Design for Structure from Motion using Concurrent Learning

Ghananeel Rotithor, Ryan Saltus, Rushikesh Kamalapurkar, Ashwin Dani

***Abstract*—In this paper, a concurrent learning based observer for a perspective dynamical system (PDS) is developed. The PDS is a widely used model for estimating the depth of the feature point from a sequence of camera images. Building on the current progress of concurrent learning (CL) for parameter estimation in adaptive control, a state observer is developed for a PDS model where the inverse depth appears as a time-varying parameter in the dynamics. Using the data-recorded over a sliding time window in the near past, information about the recent depth values is used in a CL term and an observer is developed. A Lyapunov-based stability analysis is carried out to prove the uniformly ultimately bounded (UUB) stability of the observer. Comparisons in simulations are presented with the existing observers in terms of convergence, and error statistics. Comparisons reveal that CL improves the convergence and accuracy of the presented observer.**

I. Introduction

Estimating the 3D coordinates of feature points using observations from a sequence of camera images is referred to as the Structure from Motion (SfM) problem in computer vision literature. The estimated 3D coordinates or structure information can be used in a variety of automatic control, autonomy, and intelligent control applications. Existing solutions to this problem include offline [1] and online [2]–[5] methods. The focus of this paper is on online methods, where the problem is formulated as a state estimation problem of a perspective dynamical system (PDS). The PDS is a class of nonlinear system that uses inverse depth parametrization, which is widely used in observer-based methods, and simultaneous localization and mapping (SLAM).

Building on our prior work in [6]–[8], a concurrent learning (CL)-based state observer is designed for a PDS. CL is used in adaptive control for parameter estimation, where the knowledge of past trajectory data is leveraged to estimate the constant parameter and achieve state tracking. In PDS, the inverse depth appears as a parameter in the dynamics of pixel coordinates. The inverse depth is time-varying with known dynamics associated with it. Inspired by recent advances in CL in adaptive control [9]–[11], in this paper, an observer is designed which uses CL terms in the observer design for PDS. To the authors' knowledge this is the first attempt to use CL terms in the state observer structure.

Online methods often rely on the use of an Extended Kalman Filter (EKF) [12]–[15]. In comparison to Kalman filter-based approaches, nonlinear observers are developed for SfM with analytical proofs of stability. Under the assumption that the camera motion is known, continuous and discontinuous observers are developed. A high-gain observer called the identifier-based observer (IBO) is presented for range estimation in [16]. A semi-globally asymptotically stable reduced-order observer is presented in [17] to estimate the range of a stationary object. A continuous observer which guarantees asymptotic range estimation is presented in [18] under the assumption that camera motion is known. In [19], an asymptotically converging nonlinear observer is developed based on Lyapunov's indirect method. In [20], a discontinuous sliding-mode observer is developed which guarantees uniformly ultimately bounded (UUB) result of estimation error to a small ball around the origin of the system. In [21], a nonlinear observer is developed that achieves exponentially fast convergence of estimation error provided a persistency of excitation (PE) condition is satisfied. A local exponential stability of estimation error dynamics is obtained for this observer, which means the initial condition of the observer needs to be close to the 'true' depth. This observer is used in conjunction with visual servoing for simultaneous depth estimation and VS control. In another recent work [22], an immersion and invariance (I&I) based approach is used to design a reduced-order observer to achieve global exponential convergence of the estimation error. The observer requires camera velocity and acceleration measurements along with feature point measurements from the image and the Extended Output Jacobian (EOJ) observability rank condition must be satisfied, which is more strict than the PE condition. In our prior work, we have developed a globally exponentially stable observer for feature point depth estimation of static objects using a moving camera [7], [6].

The key idea behind concurrent learning is to use recorded input and output data, also known as the history stack of the system, to make updates to the parameter estimation problem. The concurrent learning method is based on the premise that excitation will only be for a finite time and convergence can be guaranteed in finite time. This relaxes the PE condition to a finite excitation condition, wherein the minimum singular value of the regressor matrix needs to be positive [23]. Such a condition can obviously be monitored online in comparison to the PE condition. The advantage of concurrent learning for parameter estimation in model reference adaptive control (MRAC) is presented in [11]. A concurrent learning based method is developed in [10] for parameter estimation using dynamic state derivatives. Simultaneous state and parameter

Ghananeel Rotithor, Ryan Saltus, and Ashwin P. Dani are with the Department of Electrical and Computer Engineering at University of Connecticut, Storrs, CT 06269. Email: {ghananeel.rotithor; ryan.saltus; ashwin.dani}@uconn.edu. Rushikesh Kamalapurkar is with the Mechanical and Aerospace Engineering Department at the Oklahoma State Univeristy, OK.

estimation are demonstrated in [24]. Concurrent learning has been used in a variety of applications where parameter estimation is critical [25]–[27].

The existing observers designed for PDS in the literature require a strict PE condition to guarantee the stability of the observer. By adding the concurrent learning terms that keeps record of past history of depth values the PE condition requirement on the converge to the observer is somewhat relaxed. Even though unlike standard adaptive control, the depth parameter χ is time-varying, the recent past history of the system contains information about the current χ. From the simulation results of the new observer design, it can be seen that the inclusion of CL results in a faster convergence in finite time and has an obvious advantage when the PE condition cannot be satisfied.

II. Perspective Camera Motion Model

The movement of a camera capturing a scene results in a change of location of a static object in the image plane. The reference frame of the camera changes with the camera motion and can be transformed using translation and rotation. For such reference frames, let $\bar{m}(t) \in \mathbb{R}^3$ and $m(t) \in \mathbb{R}^3$ be the Euclidean and normalized Euclidean coordinates, expressed as

$$\bar{m}(t) = [X, Y, Z]^T \tag{1}$$

$$m(t) = [\frac{X}{Z}, \frac{Y}{Z}, 1]^T \tag{2}$$

where Z is the depth of a point. To estimate the depth, let's define an auxiliary vector $y(t) \in \mathbb{R}^3$, given by

$$y(t) = [\frac{X}{Z}, \frac{Y}{Z}, \frac{1}{Z}]^T \tag{3}$$

The auxiliary vector $y(t)$ is related to the feature points in the image frame as $s_i = Ay$, where $A \in \mathbb{R}^{3\times3}$ is a camera calibration matrix. Let $s \in \mathbb{R}^{2m}$ be a collection of m feature points where the first two components of s_i are considered. Taking the time derivative of (3), the dynamics of s as a function of linear and angular velocities of the camera can be expressed as

$$\dot{s} = L_s(s, \chi)u \tag{4}$$

where $L_s \in \mathbb{R}^{2m\times6}$ is the interaction matrix representing the dynamics associated with the feature point, $\chi \in \mathbb{R}^m$ is a vector containing the inverse depth values associated with all the feature points, $u \in \mathbb{R}^6$ is a vector of 3-dimensional linear and angular velocities of the camera, $u = [v_x, v_y, v_z, \omega_x, \omega_y, \omega_z]^T$, where $v = [v_x, v_y, v_z]^T \in \mathbb{R}^3$ is the linear velocity and $\omega = [\omega_x, \omega_y, \omega_z]^T \in \mathbb{R}^3$ is the angular velocity of the camera. The dynamics in (4) can also be expressed in the following form

$$\dot{s} = f_m(s, \omega) + \Omega^\intercal(s, v)\chi \tag{5}$$

where $f_m(s, \omega) \in \mathbb{R}^{2m}$ and $\Omega(s, v) \in \mathbb{R}^{m\times2m}$ are functions of measurable quantities or known quantities. The dynamics associated with the inverse depth χ can be modeled as

$$\dot{\chi} = f_u(s, \chi, u) \tag{6}$$

For a single feature point, by formulating the state as $s_j = [y_1(t), y_2(t)]^\intercal$ and inverse depth is $\chi = y_3(t)$, using (4) and (6) the dynamics for the motion of a single feature point can be written as

$$\begin{aligned}
\dot{y_1} &= y_3(y_1v_z - v_x) + y_1y_2\omega_x - \omega_y(1 + y_1^2) + y_2\omega_z \\
\dot{y_2} &= y_3(y_2v_z - v_y) + \omega_x(1 + y_2^2) - \omega_y y_2 y_1 - \omega_z y_1 \\
\dot{y_3} &= y_3^2 v_z + y_2y_3\omega_x - y_1y_3\omega_y
\end{aligned} \tag{7}$$

From (7), $f_m(s, \omega) = [y_1y_2\omega_x - \omega_y(1+y_1^2) + y_2\omega_z,\ \omega_x(1+y_2^2) - \omega_y y_2 y_1 - \omega_z y_1]^T$ and $\Omega(s, v) = [y_1v_z - v_x,\ y_2v_z - v_y]$.

Problem Definition: Given the feature point dynamics in (7), the measurements of feature points in images $z = [y_1 + \epsilon_x, y_2 + \epsilon_y]^T$ where ϵ_x and ϵ_y are additive Gaussian white noise in x and y direction, and linear and angular velocity of camera u, the depth of each feature point can be estimated using a sequence of image observations of the same feature point. To this end, a state observer is designed in Section III using concurrent learning (CL).

III. Concurrent Learning-based Observer Design

The depth estimation schemes in the existing literature require a strong observability condition called Persistence of Excitation (PE). For such estimators the estimation error converges to zero only if the PE condition is satisfied. In cases where PE cannot be satisfied the observer may be unstable. The PE condition is satisfied if there exists $T, \rho \in \mathbb{R}^+$ for the following integral condition

$$\int_t^{t+T} \Omega(\tau)\Omega^\intercal(\tau)d\tau \geq \rho I > 0, \forall t > t_0 \tag{8}$$

PE being a strict condition can be impractical to implement and continuously monitor. Concurrent learning based parameter estimation schemes use the recorded data generated by the system to make updates to the parameter estimation dynamics. Concurrent learning is based on the premise that PE will not be true and excitation will be available only for a finite amount of time. Thus, the machinery developed in [23] guarantees convergence in finite time. As a result, the integral PE condition is reduced to a rank based finite excitation condition. Thus, the convergence of the estimated value to the true value can be guaranteed in finite time.

$$\mathrm{rank}(\Sigma_{j=1}^{M-1}\Omega(x_j)\Omega^\intercal(x_j)) = m \tag{9}$$

where the history stack is $\mathcal{H} = \{(\dot{x}_j, x_j, u_j, \hat{\chi}_j)\}_{j=1}^{M-1}$ containing the past data points. If the history stack satisfies the rank condition given in (9), stability of the observer can be guaranteed. The estimates of s, χ are denoted by $\hat{s}, \hat{\chi}$, respectively.

Let's define the state and depth estimation errors as $z = \chi - \hat{\chi}$, and $\xi = s - \hat{s}$, respectively. Using the dynamics in (4) and (6), the observer for estimating the state and the depth is designed as follows.

$$\dot{\hat{s}} = f_m(s, \omega) + \Omega^\intercal(s, v)\hat{\chi} + H\xi \tag{10}$$

$$\dot{\hat{\chi}} = f_u(s, \hat{\chi}, u) + \alpha\Omega(s, v)\xi + K_{CL}\alpha\Sigma_{j=1}^{M}\Omega(s_j, v_j)(\dot{s}_j - f_m(s_j, \omega_j) - \Omega^\intercal(s_j, v_j)\hat{\chi}) \tag{11}$$

where $H \in \mathbb{R}^+$, $\alpha \in \mathbb{R}^+$ and $K_{CL} \in \mathbb{R}^+$ are suitable observer gains. The index M in (11) is for the signals at the current time instance. Using the system dynamics in (4)-(6) and the observer in (10)-(11), the estimation error dynamics can be written as follows.

$$\dot{\xi} = -H\xi + \Omega^\intercal(s, v)z \tag{12}$$

$$\dot{z} = -\alpha\Omega(s, v)\xi - K_{CL}\alpha\Sigma_{j=1}^{M}\Omega(s_j, v_j)(\dot{s}_j - f_m(s_j, \omega_j) - \Omega^\intercal(s_j, v_j)\hat{\chi}) + g(s, \chi, \hat{\chi}, u) \tag{13}$$

where $g(s, \chi, \hat{\chi}, u) = f_u(s, \chi, u) - f_u(s, \hat{\chi}, u)$. The state derivative term $\dot{s}(t)$, i.e., the optical flow, can be substituted as $\dot{s}_j = f_m(s_j, \omega_j) + \Omega^\intercal(s_j, v_j)\chi_j$. The error dynamics in (13) can be rewritten as

$$\dot{z} = -\alpha\Omega(s, v)\xi - K_{CL}\alpha\Sigma_{j=1}^{M}\Omega(s_j, v_j)(f_m(s_j, \omega_j) + \Omega^\intercal(s_j, v_j)\chi_j - f_m(s_j, \omega_j) - \Omega^\intercal(s_j, v_j)\hat{\chi}) + g(s, \chi, \hat{\chi}, u) \tag{14}$$

which can be written as

$$\dot{z} = -\alpha\Omega(s, v)\xi - K_{CL}\alpha\Sigma_{j=1}^{M}\Omega(s_j, v_j)\Omega^\intercal(s_j, v_j)(\chi_j - \hat{\chi}) + g(s, \chi, \hat{\chi}, u) \tag{15}$$

Adding and subtracting $\Omega(s_j, v_j)\Omega^\intercal(s_j, v_j)\chi$ in the summation term of (15) yields

$$\dot{z} = -\alpha\Omega(s, v)\xi - K_{CL}\alpha(\Sigma_{j=1}^{M}\Omega(s_j, v_j)\Omega^\intercal(s_j, v_j)(\chi_j - \hat{\chi}) + \Omega(s_j, v_j)\Omega^\intercal(s_j, v_j)\chi - \Omega(s_j, v_j)\Omega^\intercal(s_j, v_j)\chi) + g(s, \chi, \hat{\chi}, u) \tag{16}$$

Grouping χ and $\hat{\chi}$, the error dynamics can be written as

$$\dot{z} = -\alpha\Omega(s, v)\xi - K_{CL}\alpha\big(\Sigma_{j=1}^{M}\Omega(s_j, v_j)\Omega^\intercal(s_j, v_j)z + \Sigma_{j=1}^{M}\Omega(s_j, v_j)\Omega^\intercal(s_j, v_j)(\chi_j - \chi)\big) + g(s, \chi, \hat{\chi}, u) \tag{17}$$

Assumption 1: The motion of the camera is smooth such that the depth change is also smooth, i.e., $\|(\chi_j - \chi)\| \leq \bar{\chi}$, where χ_j are the true depth values during the period $t - t_1$ to t.

IV. Stability Analysis

The stability analysis is carried out for the initial phase where the data is being collected in the history stack of the concurrent learning term and the phase where the history stack is full. In Theorem 1, leveraging our prior work in [7], it is shown that the estimation error dynamics in (12)-(13) is stable under a PE condition. In Theorem 2, it is shown that the estimation error dynamics in (12)-(13) yields uniformly ultimately bounded error. The advantage of adding the CL term is that the error is bounded even if the PE is not satisfied.

Theorem 1. *When the history stack is incomplete, the error system in (12)-(13) is uniformly ultimately bounded if Assumption 1 and the PE condition in (8) are satisfied.*

Proof: The error system in (17) can be written

$$\dot{z} = -\alpha\Omega(s, v)\xi - K_{CL}\alpha\Omega(s_M, v_M)\Omega^\intercal(s_M, v_M)z + g(s, \chi, \hat{\chi}, u) - K_{CL}\alpha(\Sigma_{j=1}^{M-1}\Omega(s_j, v_j)\Omega^\intercal(s_j, v_j)z + \Sigma_{j=1}^{M}\Omega(s_j, v_j)\Omega^\intercal(s_j, v_j)(\chi_j - \chi)\big) \tag{18}$$

In the subsequent development, result of Theorem 1 of [7] is used which proves that the error system in (18) without last two terms is globally exponentially stable.

Consider a domain $\mathcal{D} \subset \mathbb{R}^{m+1}$ containing $e(0) = [\xi(0), z(0)]^T$ and a continuously differentiable, radially unbounded candidate Lyapunov function, $V(e) : \mathcal{D} \to \mathbb{R}^+$, defined as

$$V = \frac{1}{2}\xi^\intercal\xi + \frac{1}{2}z^\intercal\alpha^{-1}z \tag{19}$$

where $\alpha \in \mathbb{R}$ is a constant. The Lyapunov function can be upper and lower bounded by $c_1\|e\|^2 \leq \|V\| \leq c_2\|e\|^2$, where $c_1 \in \mathbb{R}$ and $c_2 \in \mathbb{R}$ are positive constants. Taking the time derivative of (19) and substituting the error dynamics in (12) and (18) yields

$$\dot{V} = -\xi^\intercal H\xi - K_{CL}z^\intercal(\Omega(s_M, v_M)\Omega^\intercal(s_M, v_M))z + \alpha^{-1}z^\intercal g(s, \chi, \hat{\chi}, u) - K_{CL}z^T(\Sigma_{j=1}^{M-1}\Omega(s_j, v_j)\Omega^\intercal(s_j, v_j)z + \Sigma_{j=1}^{M}\Omega(s_j, v_j)\Omega^\intercal(s_j, v_j)(\chi_j - \chi)) \tag{20}$$

When the history stack is incomplete, $\sum_{j=1}^{M-1}\Omega(s_j, v_j)\Omega^\intercal(s_j, v_j) \geq 0$. Using Assumption 1, (20) can be written as

$$\dot{V} \leq -\xi^\intercal H\xi - K_{CL}z^\intercal\Omega(s_M, v_M)\Omega^\intercal(s_M, v_M)z + \alpha^{-1}z^\intercal g(s, \chi, \hat{\chi}, u) + K_{CL}\sigma_1^2M\bar{\chi}||z|| \tag{21}$$

Since $\sigma_1^2M\bar{\chi} \geq 0$, using the result of Theorem 1 of [7], $\dot{V} \leq -c_1||e||^2 \quad \forall||e|| \geq \frac{K_{CL}\sigma_1^2M\bar{\chi}}{\theta}$, which yields an uniformly ultimately bound on estimation error $\|e(t)\|$ according to Theorem 4.18 of [28]. ■

Theorem 2. *When the history stack is complete and full rank, the error system in (12)-(13) is uniformly ultimately bounded if Assumption 1 is satisfied.*

Proof: Consider the same candidate Lyapunov function, $V(e):\mathcal{D}\to\mathbb{R}^{+}$, in (19). Taking the time derivative of (19) and substituting the error dynamics in (12) and (17) yields

$$\begin{aligned}\dot{V}=&-\xi^{\intercal}H\xi-K_{CL}z^{\intercal}(\Sigma_{j=1}^{M}\Omega(s_j,v_j)\Omega^{\intercal}(s_j,v_j)z\\&+\Sigma_{j=1}^{M}\Omega(s_j,v_j)\Omega^{\intercal}(s_j,v_j)(\chi_j-\chi))\\&+\alpha^{-1}z^{\intercal}g(s,\chi,\hat{\chi},u)\end{aligned}\tag{22}$$

The Lipschitz continuous term $g(s,\chi,\hat{\chi},u)$ can be upper bounded by $||g(s,\chi,\hat{\chi},u)||<L_g||z||$, where L_g is the Lipschitz constant. Using the Cauchy-Schwarz inequality and Lipschitz continuity assumption of $g(s,\chi,\hat{\chi},u)$, the upper bounds on the term $z^{\intercal}g(s,\chi,\hat{\chi},u)$ can be derived as follows.

$$||z^{\intercal}g(s,\chi,\hat{\chi},u)||\le L_g||z||^2\tag{23}$$

Since the history stack is complete, $\sum_{j=1}^{M-1}\Omega(s_j,v_j)\Omega^{\intercal}(s_j,v_j)>0$, the summation term in (22) containing $(\chi_j-\chi)$ can be upper bounded using the the Cauchy-Schwarz inequality as follows

$$\sigma_1^2(||\sum_{j=1}^{M}(\chi_j-\chi)||)\tag{24}$$

where σ_1^2 is the smallest singular value of $\sum_{j=1}^{M}\Omega\Omega^{T}$.

Using (23), and (24), $\dot{V}$ in (22) can be modified to

$$\begin{aligned}\dot{V}\le&-k_1||\xi||^2-K_{CL}\sigma_1^2||z||^2+K_{CL}\sigma_1^2M\bar{\chi}||z||\\&+\alpha^{-1}L_g||z||^2\end{aligned}\tag{25}$$

$$\begin{aligned}\le&-k_1||\xi||^2-(K_{CL}\sigma_1^2-\alpha^{-1}L_g)||z||^2\\&+K_{CL}\sigma_1^2M\bar{\chi}||z||\end{aligned}\tag{26}$$

If the following condition is satisfied the effect of

$$K_{CL}>\frac{\alpha L_g^{-1}}{\sigma_1^2}\tag{27}$$

then (27) can be written as

$$\dot{V}\le-k_1||\xi||^2-k_2||z||^2+k_3||z||\tag{28}$$

where $k_2=K_{CL}\sigma_1^2-\alpha^{-1}L_g$, and $k_3=K_{CL}\sigma_1^2M\bar{\chi}$. Adding and subtracting $\theta||z||^2$ in (28), yields

$$\begin{aligned}\dot{V}\le&-k_1||\xi||^2-\bar{k}_2||z||^2-\theta||z||^2+k_3||z||\\\le&-\min(\bar{k}_1,\bar{k}_2)||e||^2\qquad\forall||e||\ge\frac{k_3}{\theta}\end{aligned}\tag{29}$$

where $\bar{k}_1=k_1-\theta$, and $\bar{k}_2=k_2-\theta$, and $0<\theta<1$. Now, using the upper and lower bounds on $V(e)$, (29) and invoking Theorem 4.18 in [28], the error $\|e(t)\|$ is uniformly ultimately bounded with an ultimate bound according to Theorem 4.18. ■

Remark 3. The gain K_{CL} can be chosen to minimize the disturbance caused due to $\alpha^{-1}z^{\intercal}g(s,\chi,\hat{\chi},u)$. However, a choice of proper trajectories leading to a higher value of the minimum singular value would be appropriate in such a case. With the proper choice of trajectories and high minimum singular value the estimated depth will rapidly converge to the true depth. Although the presented analysis contains only two cases, i.e., before and after the history stack is full, new data can be continuously added to the stack after time T, as long as σ_1^2 stays positive. Using the Algorithm I, the σ_1^2 stays positive after time T, hence, the upper bound on the derivative of the Lyapunov function holds for all time after T. Thus, (19) is a common Lyapunov function (cf. [29]).

Algorithm 1: Depth Estimation using Concurrent Learning

Data: State vector s and velocity vector $u=[v,\omega]^{\intercal}$
Result: Estimates $\hat{s}$ and $\hat{\chi}$
Initialize $\hat{s},\hat{\chi},M$;
Define the estimator gains for K_{CL},α,H;
Initialize the History Stack and the Auxiliary Stack $\mathcal{H},\mathcal{G}$ with zeros of size $M-1$ and N respectively;
while *data for the current time step is present* **do**
    Measure the linear and angular velocity of the camera v,ω;
    Compute feature points and create state vector s;
    Compute the Optical flow and obtain $\dot{s}$;
    Compute $f_m(s,\omega),f_u(s,\hat{\chi},u),\Omega(s,\omega)$;
    Estimate the values for $\dot{\hat{s}},\dot{\hat{\chi}}$;
    Integrate $\dot{\hat{s}},\dot{\hat{\chi}}$ to obtain $\hat{s},\hat{\chi}$;
    if *Number of iterations < M+1* **then**
        Add data point to History Stack $\mathcal{H}$;
    end
    Add data point to G in a cyclic way;
    Search for M data points with maximum σ_1^2 in the G stack;
    Replace data in $\mathcal{H}$ with the selected M points from $\mathcal{G}$;
end

V. Simulation Results

A simulation was performed using a single feature point to verify the performance of the observer designed in Section 3. For simulation purposes, the value of the focal length λ is set to 1. An initial global point is selected as $x(t_0)=[5,2.5,3]^{\intercal}$ where $x=[x_1,x_2,x_3]^{\intercal}=[X,Y,Z]^{\intercal}$. An auxiliary state vector $y=[y_1,y_2,y_3]^{\intercal}=[\frac{x_1}{x_3},\frac{x_2}{x_3},\frac{1}{x_3}]^{\intercal}$ is constructed for the measurement of the inverse depth $\chi=\frac{1}{x_3}$. The point values are simulated by integrating the equation given in (7), where $v=[v_x,v_y,v_z]^{\intercal}$ are the linear velocities and $\omega=[\omega_x,\omega_y,\omega_z]^{\intercal}$ are the angular velocities. From the mathematics presented in Section 2, we can build our state vector as $s=[y_1,y_2]^{\intercal}$ and $\chi=y_3$.

A fourth order Runge-Kutta integrator with a fixed timestep of 0.001s is used to integrate the equations and generate trajectories for all values of y. The velocities used for the simulation are

$$v=[0.3,0.2cos(\frac{\pi t}{4}),-0.3]^{\intercal}$$

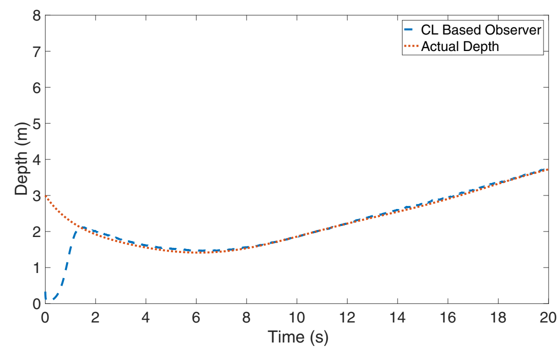

Figure 1. Comparison between the actual depth and estimated depth over a period of 20 seconds

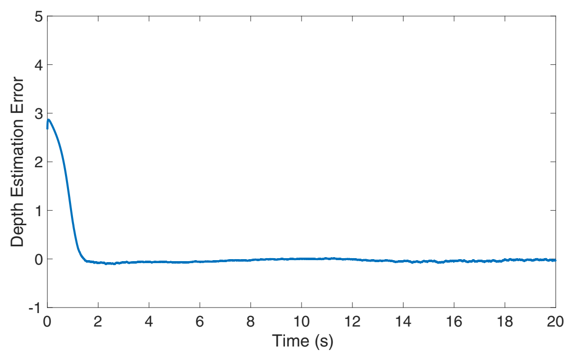

Figure 2. Error between the true and the estimated depth

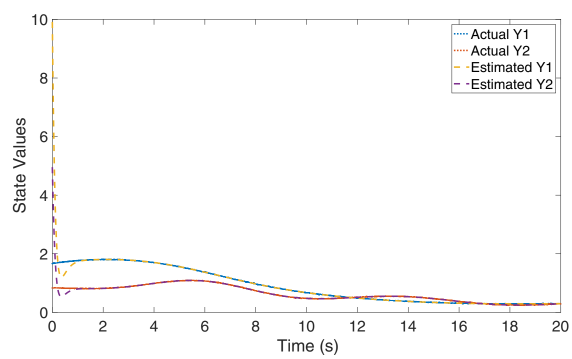

Figure 3. Comparison between the actual and estimated state values

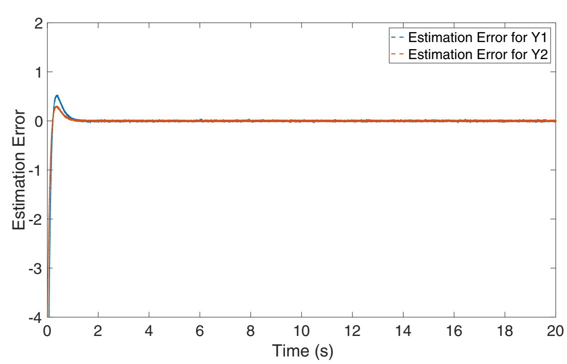

Figure 4. Error between the true and estimated state

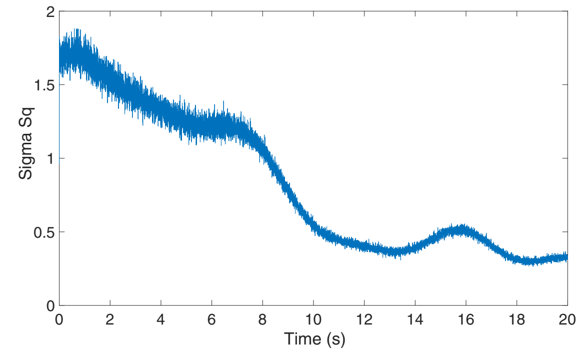

Figure 5. Change in minimum singular value of the history stack

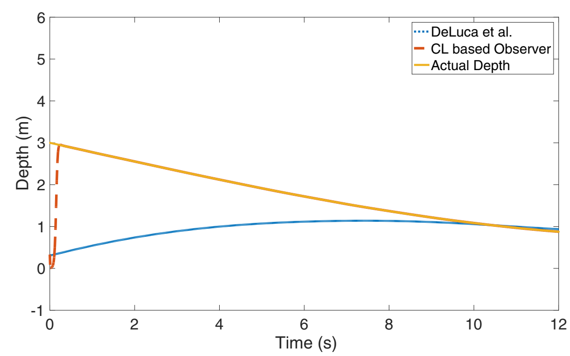

Figure 6. Comparison of the convergence of the CL based observer and the observer presented in [2]

$$\omega = [0, -\frac{\pi}{30}, 0]^\intercal$$

Additive white Gaussian noise with a signal to noise ratio (SNR) of 40 dB is added to the image pixel measurements, and noise with a variance of .01 and zero mean is added to the velocity signal. The gain values used for the simulation are $K_{CL} = 3$, $\alpha = 0.5$, $H = 10$. The initial values for the state estimate and inverse depth estimate are selected as $\hat{s}(t_0) = [10, 5]^\intercal$ and $\hat{\chi}(t_0) = 3$. The history stack and the auxiliary stack are initialized with two points for the purpose of simulation. The results of the simulation can be summarized by the plots shown in Figs. 1-5.

Fig. 1 shows the value of the actual depth compared with the estimated depth. After a high initial error the depth estimation of the CL Based Observer quickly converges to the true value in about 1.5 seconds, and then closely approximates the true depth as time progresses. In Fig. 2, the depth estimation error can be seen for the simulation. Fig. 3. shows the actual state trajectories compared to the estimated state trajectories. The state trajectories generated by the estimator perfectly follow the actual state trajectories

Table I
COMPARISON OF THE CL BASED OBSERVER WITH THE OBSERVER PRESENTED IN [2].

	CL based Observer	Observer in [2]
Total RMSE	0.1173	0.4003
Convergence time(seconds)	0.228	9.83
MAPE(%)	0.4352	6.0204

after initial error. The state estimation error in Fig. 4 goes to zero in approximately 0.8s. The value of the minimum singular value of the history stack $\mathcal{H}$ varies with time and in shown in Fig. 5. Due to the time varying nature of the depth signal and constant purging of the history stack, the minimum singular value varies. However, at all times it is evident that the minimum singular value stays positive, which is an essential condition for the convergence of the estimated depth to the true depth.

Fig 6. shows the comparison of the convergence speed of the CL based observer and the observer developed in [2] using the same velocity signal and initial conditions present in the first signal. The CL based observer converges rapidly in 0.228s compared 9.83s required by the other observer. The CL observer also achieves better error statistics which are demonstrated in Table I. The root mean squared error and the mean absolute percentage error values computed for the CL observer at 0.1172 and 0.4352 are much lower compared to the observer in [2]. The faster convergence of the CL observer can be associated to the recorded history stack which is absent in the compared observer.

VI. Conclusion

A concurrent learning based estimator is presented for the estimation of depth of a stationary feature point in an image using a moving camera. The estimator error is shown to be UUB. The simulation result shows the performance of the designed estimator and its comparison with the existing estimator in the presence of noise. The estimator designed in this paper shows faster convergence and achieves better error statistics in the presence of measurement noise.

References

[1] C. Godard, O. Mac Aodha, and G. J. Brostow, "Unsupervised monocular depth estimation with left-right consistency," in *CVPR*, vol. 2, no. 6, 2017, p. 7.

[2] R. Spica and P. R. Giordano, "A framework for active estimation: Application to structure from motion," in *IEEE Conference on Decision and Control, CDC'13*, 2013.

[3] R. Spica, P. R. Giordano, and F. Chaumette, "Active structure from motion: Application to point, sphere, and cylinder," *IEEE Transactions on Robotics*, vol. 30, no. 6, pp. 1499–1513, 2014.

[4] ——, "Coupling visual servoing with active structure from motion," in *IEEE Int. Conf. on Robotics and Automation, ICRA'14*, 2014.

[5] R. Spica, P. Robuffo Giordano, and F. Chaumette, "Coupling active depth estimation and visual servoing via a large projection operator," *The International Journal of Robotics Research*, vol. 36, no. 11, pp. 1177–1194, 2017.

[6] A. Dani, N. Fisher, and W. E. Dixon, "Single camera structure and motion," *IEEE Transactions on Automatic Control*, vol. 57, no. 1, pp. 241–246, 2012.

[7] A. Dani, N. Fischer, Z. Kan, and W. Dixon, "Globally exponentially stable observer for vision-based range estimation," *Mechatronics*, vol. 22, no. 4, pp. 381–389, 2012.

[8] A. P. Dani, S.-J. Chung, and S. Hutchinson, "Observer design for stochastic nonlinear systems via contraction-based incremental stability," *IEEE Transactions on Automatic Control*, vol. 60, no. 3, pp. 700–714, 2015.

[9] G. Chowdhary, T. Yucelen, M. Muhlegg, and E. N. Johnson, "Concurrent learning adaptive control of linear systems with exponentially convergent bounds," *International Journal of Adaptive Control and Signal Processing*, vol. 27, no. 4, pp. 280–301, 2013.

[10] R. Kamalapurkar, B. Reish, G. Chowdhary, and W. E. Dixon, "Concurrent learning for parameter estimation using dynamic state-derivative estimators," *IEEE Transactions on Automatic Control*, vol. 62, no. 7, pp. 3594–3601, 2017.

[11] M. Muhlegg, G. Chowdhary, and E. Johnson, "Concurrent learning adaptive control of linear systems with noisy measurements," in *AIAA Guidance, Navigation, and Control Conference*, 2012, p. 4669.

[12] S. Soatto and P. Perona, "Reducing" structure from motion": a general framework for dynamic vision. 1. modeling," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 9, pp. 933–942, 1998.

[13] L. Matthies, T. Kanade, and R. Szeliski, "Kalman filter-based algorithms for estimating depth from image sequences," *International Journal of Computer Vision*, vol. 3, no. 3, pp. 209–238, 1989.

[14] H. Kano, B. Ghosh, and H. Kanai, "Single camera based motion and shape estimation using extended kalman filtering," *Mathematical and Computer Modelling*, vol. 34, no. 5-6, pp. 511–525, 2001.

[15] A. Chiuso, P. Favaro, H. Jin, and S. Soatto, "Structure from motion causally integrated over time," *IEEE transactions on pattern analysis and machine intelligence*, vol. 24, no. 4, pp. 523–535, 2002.

[16] M. Jankovic and B. K. Ghosh, "Visually guided ranging from observations of points, lines and curves via an identifier based nonlinear observer," *Systems & Control Letters*, vol. 25, no. 1, pp. 63–73, 1995.

[17] D. Karagiannis and A. Astolfi, "A new solution to the problem of range identification in perspective vision systems," *IEEE Transactions on Automatic Control*, vol. 50, no. 12, pp. 2074–2077, 2005.

[18] W. E. Dixon, Y. Fang, D. M. Dawson, and T. J. Flynn, "Range identification for perspective vision systems," in *Proc. Am. Control Conf.*, Denver, Colorado, June 2003, pp. 3448–3453.

[19] O. Dahl, F. Nyberg, and A. Heyden, "Nonlinear and adaptive observers for perspective dynamic systems," in *American Control Conference, 2007. ACC'07.* IEEE, 2007, pp. 966–971.

[20] X. Chen and H. Kano, "State observer for a class of nonlinear systems and its application to machine vision," *IEEE Transactions on Automatic Control*, vol. 49, no. 11, pp. 2085–2091, 2004.

[21] A. De Luca, G. Oriolo, and P. B. Giordano, "Feature depth observation for image-based visual servoing: Theory and experiments," *Int J Robot Res*, vol. 27, no. 10, pp. 1093–1116, 2008.

[22] F. Morbidi and D. Prattichizzo, "Range estimation from a moving camera: an immersion and invariance approach," in *Robotics and Automation, 2009. ICRA'09. IEEE International Conference on*. IEEE, 2009, pp. 2810–2815.

[23] G. Chowdhary and E. Johnson, "Concurrent learning for convergence in adaptive control without persistency of excitation," in *Decision and Control (CDC), 2010 49th IEEE Conference on*. IEEE, 2010, pp. 3674–3679.

[24] R. Kamalapurkar, "Simultaneous state and parameter estimation for second-order nonlinear systems," in *Decision and Control (CDC), 2017 IEEE 56th Annual Conference on*. IEEE, 2017, pp. 2164–2169.

[25] A. Parikh, R. Kamalapurkar, H.-Y. Chen, and W. E. Dixon, "Homography based visual servo control with scene reconstruction," in *Decision and Control (CDC), 2015 IEEE 54th Annual Conference on*. IEEE, 2015, pp. 6972–6977.

[26] Z. Bell, A. Parikh, J. Nezvadovitz, and W. E. Dixon, "Adaptive control of a surface marine craft with parameter identification using integral concurrent learning," in *Decision and Control (CDC), 2016 IEEE 55th Conference on*. IEEE, 2016, pp. 389–394.

[27] G. Chowdhary and E. Johnson, "Theory and flight test validation of long term learning adaptive flight controller," in *AIAA Guidance, Navigation and Control Conference and Exhibit*, 2008, p. 6781.

[28] H. K. Khalil, *Nonlinear Systems*, 3rd ed. Prentice Hall, 2002.

[29] D. Liberzon, *Switching in systems and control*. Springer Science & Business Media, 2003.