On Robust Observer Design for System Motion on SE(3) Using Onboard Visual Sensors

Tong Zhang *, Ying Tan **, Xiang Chen *, Zike Lei ***

* Department of Electrical and Computer Engineering, University of Windsor, Windsor, Canada N9B3P4 (email: zhang14h@uwindsor.ca, email: xchen@uwindsor.ca)
** Department of Mechanical Engineering, University of Melbourne, Melbourne, Australia 3010 (email: yingt@unimelb.edu.au)
*** Department of Automation, School of Information Science and Engineering, Wuhan University of Science and Technology, China 430081 (email: leizike@wust.edu.cn)

Abstract: Onboard visual sensing has been widely used in the unmanned ground vehicle (UGV) and/or unmanned aerial vehicle (UAV), which can be modeled as dynamic systems on SE(3). The onboard sensing outputs of the dynamic system can usually be applied to derive the relative position between the feature marks and the system, but bearing with explicit geometrical constraint. Such a visual geometrical constraint makes the design of the visual observer on SE(3) very challenging, as it will cause a time-varying or switching visible set due to the varying number of feature marks in this set along different trajectories. Moreover, the possibility of having mis-identified feature marks and modeling uncertainties might result in a divergent estimation error. This paper proposes a new robust observer design method that can accommodate these uncertainties from onboard visual sensing. The key design idea for this observer is to estimate the visible set and identify the mis-identified features from the measurements. Based on the identified uncertainties, a switching strategy is proposed to ensure bounded estimation error for any given trajectory over a fixed time interval. Simulation results are provided to demonstrate the effectiveness of the proposed robust observer.

1. INTRODUCTION

Localization problem is an important issue for autonomous vehicular systems such as UAVs and UGVs, as accurate attitude information is critical for maneuvering the vehicles in different applications. The problem of indoor localization is very challenging due to the lack of GPS signals and the limited available sensing devices. One approach is to construct a GPS-like environment to track the attitude of the vehicles. In Ocana et al. (2005); Zhang et al. (2020), the method of localizing robots is proposed based on the received signal strengths from several WiFi transmitters. The localization methods based on data fusion of multiple sensors are developed in Bischoff et al. (2012); Amri et al. (2015). In addition, positioning the robotic system Bischoff et al. (2012); Delibasis et al. (2013) with stationary cameras has been observed in many applications. However, all of these methods do not use onboard sensors, which causes a significant restriction on the effective operating region of vehicles. In addition, the control of vehicular systems with this kind of sensing approach cannot be conducted locally either, hence, further limiting the varieties of autonomous applications.

The other approach to positioning the vehicular systems is to apply the observer/filter to estimate the pose, based on the kinetic or dynamic models of the system and data captured with onboard sensors. One advantage is, obviously, the controlling algorithm can be implemented onboard using onboard sensing. In this regard, Park et al. (1995); Hervé (1999) demonstrate observer designs in Lie Group/Algebra formulation of rigid robotic motion to render position information on Riemannian Manifold. In Salcudean (1991), a global convergence observer is designed on Special Orthogonal Group SO(3) to estimate the rotation of the system. Nonlinear observers Lageman et al. (2009); Hua et al. (2011); Baldwin et al. (2007); Vasconcelos et al. (2010) and Extended Kalman Filters Cheng (2019); Heo and Park (2018); Bourmaud et al. (2013) are presented to estimate of the system states on Special Euclidean Group SE(3).

Different from traditional sensors which usually measure just a single value, e.g., temperature meters, humidity meters, pressure sensors, etc, visual sensors observe and provide information about important features in certain areas with geometrical constraint. These constraints in space pose limitations on the sensing capability and, to the authors’ best knowledge, are not considered in the literature mentioned above in observer designs. In some other applications, such as coverage optimization and formation control, the model of the geometrical constraint is considered. For example, in Mavrinac and Chen (2013); Zhang et al. (2018); Gusrialdi et al. (2008), the geometrical constraint of cameras is modeled and considered in the parameterized cost function for coverage optimization, while a formation control method is proposed considering the geometrical property of sensors in Li et al. (2018).

This paper focuses on designing an observer with the consideration of the explicit geometrical constraint of the visual sensor. The main results demonstrate how the proposed observer can improve performance when modeling uncertainty, measurement noise, and detection error arise, which is not the case in existing work.

2. NOTATION AND PROBLEM FORMULATION

2.1 Notations

Let $\mathbb{R}$ and $\mathbb{R}^n$ denote the set of real numbers and an $n$-dimensional Euclidean space respectively. For an $n$-
The adjoint map $\text{Ad}$ is defined as $\text{Ad}(A) = \sum_{i,j=1}^{n} a_{ij}$. For any two matrices, $A, B \in \mathbb{R}^{n \times m}$, the Euclidean inner product and Frobenius norm are defined as

$$\langle A, B \rangle = \text{tr}(A^T B), \quad ||A||_F = \sqrt{\langle A, A \rangle},$$

respectively.

A continuous function $\Psi: \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is said to be class $\mathcal{K}$ if its initial condition $\Psi(0) = 0$ and strictly increasing. The function $\Psi$ is a class $\mathcal{K}_\infty$ if it is class $\mathcal{K}$ and $\lim_{a \to \infty} \Psi(a) \to \infty$. A continuous function $\beta: [0,a] \times [0,\infty) \to [0,\infty)$ is said to belong to class $\mathcal{K}_\infty$ if, for each fixed $s$, the mapping $\beta_r(s)$ belongs to class $\mathcal{K}$ with respect to $r$ and for each fixed $r$, the mapping $\beta(r,s)$ is decreasing with respect to $s$ and $\beta(r,s) \to 0$ as $s \to \infty$ (Khalil, 2002, Chapter 4).

Let $I$ and $B$ be the inertial frame fixed to the world and the frame fixed to the rigid body. The notion of $R$ is the rotation matrix defined on a special orthogonal group $SO(3)$, which satisfies

$$SO(3) := \{ R \in \mathbb{R}^{3 \times 3} | R^T R = I_{3 \times 3}, \det(R) = 1 \}.$$  

A special Euclidean group $SE(3)$ is defined as

$$SE(3) := \left\{ X = \begin{bmatrix} R & P \\ 0_3 & 1 \end{bmatrix} \in \mathbb{R}^{4 \times 4} | R \subset SO(3), P \in \mathbb{R}^3 \right\},$$

where $R$ represents the attitude, and $P$ represents the position of the rigid body both in the Inertial frame. The Lie algebra $se(3)$ is defined as

$$se(3) := \left\{ A = \begin{bmatrix} \omega_x & v \\ 0_3 & 0 \end{bmatrix} \in \mathbb{R}^{4 \times 4} | \omega_x \in \mathbb{R}^{3 \times 3}, v \in \mathbb{R}^3 \right\}.$$  

The notion $v$ denotes the linear velocity, and $\omega_x$ is a skew-symmetric matrix obtained from the angular velocity $\omega$ of the rigid body. Its dot product satisfies

$$\omega_x k = \omega \times k, \forall k \in \mathbb{R}^3.$$  

The anti-symmetric projection and the symmetric projection of matrix $L \in \mathbb{R}^{n \times n}$ are defined as $P_a(L) = 0.5(L - L^T)$ and $P_s(L) = 0.5(L + L^T)$ respectively. Let $P: \mathbb{R}^{4 \times 4} \to se(3)$ be the projection of $\mathbb{R}^{4 \times 4}$ onto $se(3)$. For all $A \in se(3)$ and $L \in \mathbb{R}^{4 \times 4}$, it satisfies

$$\langle A, L \rangle = \langle A, P(L) \rangle = \langle P(L), A \rangle,$$

where $\langle \cdot, \cdot \rangle$ is Euclidean inner product. For all $L_1 \in \mathbb{R}^{3 \times 3}$, $\omega_1, \omega_2 \in \mathbb{R}^3$ and $I_1, I_2 \in \mathbb{R}$, the following condition holds

$$P \begin{bmatrix} L_1 & I_2 \\ \omega_1^T & 0 \end{bmatrix} = \begin{bmatrix} P_a(L_1) & I_2 \\ 0_{3 \times 3} & 0 \end{bmatrix}.$$  

The adjoint map $Ad: SE(3) \times se(3) \to se(3)$ is formed as $Ad_A := XAX^{-1}$, for all $X \in SE(3), A \in se(3)$.

### Geometrical constraint of the visual sensor

Assume that the visual sensor, such as the camera, is in the fixed position of the body frame. For any given system (or state $X$) in a 3D space, next will define the geometrical constraints coming from the fixed visual sensor in the local frame $B$ as shown in Figure 1.

**Fig. 1. Geometrical Constraint in Body Frame $B$**

Let $f_{min}$, $f_{max}$ be two positive real numbers representing the minimum and maximum detectable distance for the visual sensor. The notions of $\theta_H, \theta_V \in (0, \pi)$ represent the visual sensor’s horizontal and vertical field of view (FoV) for the directional sensor shown in the above figure. The vision distance and FoV define a geometrical constraint as shown in Figure 1. We use the notation
\[ \theta = [f_{\min}, f_{\max}, \theta_H, \theta_V]^T \] to present the parameters of the geometrical constraint.

If the robot system is located at \( X \in SE(3) \), which is the state of the system (7), the center of the geometrical constraint is denoted as \( O_c(X) \) in \( \mathbb{I} \), which is

\[ O_c(X) = XO_c^B, \quad O_c^B := \left[ 0, \frac{f_{\max} + f_{\min}}{2}, 0, 1 \right]^T. \tag{10} \]

The notion of \( O_c^B \) is the center of the geometrical constraint local frame \( \mathbb{B} \). Next will define a visible region.

**Definition 1.** For a rigid body robotic system with the state \( X \in SE(3) \), its visible region \( \Omega_V(X) \) contains all the points whose relative positions measured in frame \( \mathbb{B} \) satisfies the following inequalities:

\[ \begin{align*}
  f_{\min} & \leq py_v \leq f_{\max} \\
  |pz_v| & \leq p_{y,\tan}(\theta_H^v) \\
  |px_v| & \leq p_{y,\tan}(\theta_H^v),
\end{align*} \tag{11} \]

where \( px_v, py_v, pz_v \) is the relative position between the point \( V \) and \( X \in SE(3) \) in a 3D space.

The relative position between a point \( V \) and the system \( X \) in coordinate \( \mathbb{B} \) is denoted as

\[ P'_V = [px_v \ py_v \ pz_v \ 1]^T. \tag{12} \]

Consequently, its position in coordinate \( \mathbb{I} \) denoted as \( P_I(X) = XP'_V \).

**Property 1.** For any geometry parameter \( \theta \), there exists a positive constant \( m_\theta \), which is a function of \( \theta \), such that for any point \( V \in \Omega_V \), the following inequality holds:

\[ \|P'_V - O_c^B\|_F \leq m_\theta, \tag{13} \]

where \( P'_V \) is defined in (12).

**Proof 1.** See Appendix 6.1.

Let \( \mathbb{E} \) contain \( N \) feature marks in the environment of interest. The set \( \mathcal{O}(\bar{Y}) \) contains the feature marks identified by the visual sensor, where \( \bar{Y} \) is some function related to the measurements of the visual sensor and the errors coming from the feature identification algorithm. This set \( \mathcal{O}(\bar{Y}) \) is available from the visual sensor and the location of the fixed feature marks. The set \( \mathcal{O}^c(\bar{Y}) \) is the complementary set \( \mathcal{O}(\bar{Y}) \) with respect to \( \mathbb{E} \). Two sets satisfy the following relations:

\[ \mathcal{O}(\bar{Y}) \cap \mathcal{O}^c(\bar{Y}) = \emptyset, \quad \mathcal{O}(\bar{Y}) \cup \mathcal{O}^c(\bar{Y}) = \mathbb{E}, \tag{14} \]

where \( \emptyset \) is the empty set. For any feature mark in \( \mathcal{O}(X) \), with the consideration of mis-identified marks, it is defined that

\[ \mathcal{G}(X, \bar{Y}) := \mathcal{O}(\bar{Y}) \cap \mathcal{O}_V(X), \tag{15} \]

while the set \( \mathcal{G}^c(X, \bar{Y}) \) contains all mis-identified marks in \( \mathcal{O}(\bar{Y}) \). For any feature mark in \( \mathcal{G}(X, \bar{Y}) \), Property 1 holds.

With the consideration of geometrical constraint from the visual sensor, for the rigid robotic system with the state \( X \) coming from (7), the output for the \( i \)th feature mark becomes

\[ Y_i = \left\{ \begin{array}{ll}
  C^i + \lambda_i, & i \in \mathcal{G}(X, \bar{Y}) \\
  0, & i \in \mathcal{O}(\bar{Y})
\end{array} \right. \tag{16} \]

where \( \lambda_i = [\lambda_p \ 0]^T, \lambda_p \in \mathbb{R}^3 \) indicates the measurement noises. For a fixed feature landmark arrangement, the set \( \mathcal{O}(\bar{Y}) \) is always trajectory dependent. Due to the geometrical constraint from the visual sensor, for a given trajectory, some landmarks will be within this set when \( t = t_1 \) and will leave this set at \( t = t_2 \). Such a switching behavior from the practical setting of visual sensors will make the observer design much harder. It is highlighted that the problem settings used in Hua et al. (2011); Zhang et al. (2021) assume that \( \mathcal{G}(X, \bar{Y}) = \mathcal{O}(\bar{Y}) = \mathbb{E} \), which completely ignores the geometrical constraint from the visual sensor.

The modeling uncertainty \( d \) is in \( se(3) \) and the measurement noises \( \lambda_i, i = 1, \ldots, N \) is in \( \mathbb{R}^3 \). It is assumed that they satisfy the following assumption.

**Assumption 1.** \( D(t), \lambda_i(t) \) are uniformly bounded. That is, there exist two positive constants \( d_D \) and \( d_\lambda \) such that the following inequalities hold

\[ \text{esssup}_{t \geq 0} \|D(t)\|_F \leq d_D, \tag{17} \]

\[ \text{esssup}_{t \geq 0} \max_{i = 1, \ldots, N} \{\|\lambda_i(t)\|_F\} \leq d_\lambda. \tag{18} \]

**Objective**

This work is to design an appropriate observer to track the state of the system (7) using the noisy measurements \( Y_i \) with the consideration of geometrical constraints coming from (16) when the modeling uncertainties \( d \) and measurement noises \( \lambda_i \) satisfying Assumption 1.

**Remark 2.** It is noted that the role of the observer is to estimate the state \( X \) of the system (7) from the output \( Y_i, i = 1, \ldots, N \) in (16) coming from \( N \) feature marks. However, as \( X \) is unknown in (16), the set \( \mathcal{G}(X, \bar{Y}) \) is unknown, though \( Y_i, i = 1, \ldots, N \) and \( \mathcal{O}(\bar{Y}) \) are measured. In this work, the set \( \mathcal{G}(X, \bar{Y}) \) is also estimated from the measurements so that the state \( X \) can be estimated correctly.

### 3. MAIN RESULT

This section discusses how to estimate the state \( X \) from the dynamics (7) when the output \( Y_i \) in (9) has the form defined in (16). As \( X \) is unknown, its visible region \( \Omega_V(X) \) is unknown. Assume that, the state of the observer can estimate the state \( X \) by \( \bar{X} \). Let \( \Omega_V(\bar{X}) \) denote the set of all points in the visible region for estimation \( X \in SE(3) \) and \( \Omega^c_V(\bar{X}) \) denote the complementary set with respect to \( \Omega(\bar{Y}) \). This leads to the following two sets that are available.
\[ \mathcal{G}(\hat{X}, \hat{Y}) := \mathcal{O}(\hat{Y}) \cap \Omega \mathcal{V}^c(\hat{X}) \]  
(19)

\[ \mathcal{G}^c(\hat{X}, \hat{Y}) := \mathcal{O}(\hat{Y}) \cap \Omega \mathcal{V}^c(\hat{X}). \]  
(20)

Next will show how to use the information of these two sets as well as the knowledge of \( \Omega \mathcal{V} \) to design an appropriate observer.

The \( E_o \) denotes the estimation error and has the following form:

\[ E_o := \hat{X}X^{-1} - I. \]  
(21)

With the consideration of the dynamics (7) and the output (9) or (16), it satisfies the following relations:

\[ E_o C_i = \hat{X}(Y_i - \hat{Y}_i) = \Xi_i \]  
(22)

\[ (E_o + I)C_i = \hat{X}Y_i = \Gamma_i. \]  
(23)

Although \( E_o \) is unknown, both \( \Xi_i \) and \( \Gamma_i \) are measurable for any \( i = 1, \ldots, n \).

### 3.1 Observer Design

The observer has the following form:

\[ \hat{X} = \hat{X}(U - \Psi - Q(\hat{Y}, Y)) \]  
(24)

\[ \hat{Y}_i = \hat{X}^{-1}C_i, \]  
(25)

where

\[ \Psi = Ad_{\hat{X}^{-1}}^T \left( \ell \sum_{i \in \mathcal{G}(\hat{X}, Y)} \Xi_i C_i^T \right) \]  
(26)

\[ Q(\hat{Y}, Y) = hAd_{\hat{X}^{-1}}^T \left( \ell \cdot K_o \sum_{i \in \mathcal{G}^c(\hat{X}, \hat{Y})} \Xi_i \Gamma_i^T \right). \]  
(27)

Here \( \ell \in \mathbb{R}^+ \) is a positive constant, \( K_o = K_o^T \geq I_{4 \times 4} \) is a high gain damping term (see more details about this term in Zhang et al. (2021)), \( \Xi_i \) and \( \Gamma_i \) are defined in (22) and (23) respectively, and the sets \( \mathcal{G}(\hat{X}, \hat{Y}) \) and \( \mathcal{G}^c(\hat{X}, \hat{Y}) \) are defined in (19) and (20) respectively. It is highlighted that the positive constant \( h \) plays an important role in this observer design. It has the following form:

\[ h = h \left( \| \hat{Y}_i - O^B \|_F \right) = e^{-\frac{\| \hat{Y}_i - O^B \|_F}{\eta}}. \]  
(28)

where \( \eta \) is a positive design parameter and \( O^B \) is defined in (10). In this proposed observer, we need to design two positive constants: \( (\ell, \eta) \) and one positive definite matrix \( K_o \).

The following lemma is used to design the parameter \( \eta \) needed in the observer (24).

**Lemma 1.** For the dynamic system (7), which satisfies Assumption 1 and the proposed observer (24) with an estimation error \( E_o \) defined in (21) that satisfying \( \| E_o \|_F \leq \Delta \). If \( \hat{Y}_i \) in (25) of satisfying \( i \in \mathcal{G}(X, Y) \), there exists a positive constant, which depends on \( \Delta, m_{\theta} \) and \( \| O^B \|_F \). Such that the following inequality holds:

\[ \| \hat{Y}_i - O^B \|_F \leq \eta. \]  
(29)

**Proof 2.** See Appendix 6.2

From the definition of \( \Omega \mathcal{V}( \hat{X} ) \), it is known that from any estimated output \( \hat{Y}_i \), if it is within the set \( \mathcal{G}(\hat{X}, \hat{Y}) \), it will satisfy Property 1. Due to the existence of uncertainties, there might be some feature marks, which are actually with the visible region \( \Omega \mathcal{V}(X) \) and \( \Omega \mathcal{V}(X) \). The role of \( h \) (28) is to try to compensate for this mismatch. The design parameter \( \eta \) is based on the knowledge of the system model, the geometrical constraint model, and the worst case of uncertainties. For any \( i \in \mathcal{G}(\hat{X}, \hat{Y}) \), we have \( \| \hat{Y}_i - O^B \|_F \geq m_{\theta} \) from Property 1, moreover, it has

\[ \| \hat{Y}_i - O^B \|_F \leq \eta \Rightarrow h(\cdot) \geq 1 \]  
(30)

\[ \| \hat{Y}_i - O^B \|_F > \eta \Rightarrow 0 < h(\cdot) < 1, \]  
(31)

indicating if \( \| \hat{Y}_i - O^B \|_F \) is too large, it can be confirmed that it is a point in \( \mathcal{G}^c(\hat{X}, \hat{Y}) \), the influence of the damping term \( Q \) will decay quickly.

**Fig. 3. The Geometrical Constraint and The Point Sets**

Due to the existence of estimation error, the components in \( \mathcal{G}(X, Y) \) and \( \mathcal{G}^c(\hat{X}, \hat{Y}) \) are not the same as Figure 3 shows. However, by comparing whether \( \| \hat{Y}_i - O^B \|_F > \eta \) of all points in \( \mathcal{O}(X, Y) \), an estimation of which components in \( \mathcal{G}^c(\hat{X}, \hat{Y}) \) are due to the estimation error (\( P_1 \) in the Figure), and which components are the detection error \( \mathcal{G}^c(\hat{X}, \hat{Y}) \) (\( P_3 \) in the figure) can be obtained. Based on the estimation, \( h(\cdot) \geq 1 \) will be assigned to the components due to estimation error, and \( 0 < h(\cdot) < 1 \) will be assigned to the components confirmed to be in \( \mathcal{G}^c(\hat{X}, \hat{Y}) \).

**Remark 3.** Since the geometrical constraint is fixed for each sensor, larger \( \eta \) corresponds to a larger upper bound of the estimation error \( \| E_o \|_F \) for the system.

**Remark 4.** In literature, it is always assumed that the on-board sensors can capture all landmarks correctly as in Hua et al. (2011); Zhang et al. (2021), i.e, \( \mathcal{G}(X, Y) = \mathcal{O}(\hat{Y}) = E \). For example, the observer design in Hua et al. (2011) has the similar form as (24) with the operators \( \Psi \) without the damping term \( Q(\cdot, \cdot) \).

\[ \Psi = Ad_{\hat{X}^{-1}}^T \left( \ell \sum_{i \in E} \Xi_i C_i^T \right), \]  
(32)

The difference between two sets: \( E \) and \( \mathcal{G}(\hat{X}, \hat{Y}) \) comes from the geometrical constraint from the visual sensor. In Zhang et al. (2021), other than using the similar \( \Psi \) as in
(32), the damping term $Q(\hat{Y}, Y)$ similar to (27) was added, which has the following form:

$$Q(\hat{Y}, Y) = A_d X, Y P \left( \ell \cdot K_a \sum_{i \in E} \frac{\Xi_i \Gamma_i^T}{n} \right).$$

It is noted that in the proposed observer, the damping term is deactivated when it detects that one landmark is mis-identified from the estimated state, i.e., $i \in \mathbb{G}^c(X, Y)$.

The closed-loop error dynamics $\hat{E}_o$, which is computed from the system in (7), the output in (16), the observer in (24), and the estimated output in (25), thus becomes

$$\dot{\hat{E}}_o = -A_d (\Psi + Q(\hat{Y}, Y) + D)(E_o + I).$$

Similarly, the dynamics of $\hat{E}_o$ will be

$$\dot{\hat{E}}_o = -A_d (\Psi + Q(\hat{Y}, Y) + D)(E_o + I)C_i,$$

where $i = 1, 2, \ldots, N$, and $\Gamma_i = \Xi_i + C_i = (E_o + I)C_i$.

### 3.2 Asymptotic Analysis of Estimation Error

This section shows how to choose the parameters and the matrix of the proposed observer $(\ell, h, K_a)$ so that the estimation error $E_o$ defined in (21) is uniformly bounded for any $t \geq 0$ under some sufficient conditions.

It is noted that the set $\mathbb{G}(X, Y)$ is time-varying, indicating that the number of identified feature marks might change over time. It is also possible that the number of feature marks is the same at $t = t_1$ and $t = t_2$, but the feature marks within $\mathbb{G}(X(t_1), Y)$ and $\mathbb{G}(X(t_2), Y)$ are different. For the convenience of notation, we denote a sequence of time instants $t_k$. For any $t \in [t_k, t_{k+1})$, $\mathbb{G}(X(t), Y)$ is not changing. At $t = t_{k+1}$, the set $\mathbb{G}(X(t_{k+1}), Y)$ changes, i.e., $\mathbb{G}(X(t_{k+1}), Y) \neq \mathbb{G}(X(t_k), Y)$. Such changes lead to a sequence of switches of the set $\mathbb{G}(X(t), Y)$. In the stability analysis, we will analyze the trajectories of $E_o(t)$ within a given time interval $[0, T]$ by carefully checking each such interval $[t_k, t_{k+1})$. To simplify our analysis, it is assumed that for any given time interval $[0, T]$, the number of switches is finite.

**Assumption 2.** For a given trajectory defined $[0, T]$, and a given set up of feature marks, there exists a positive integer $N_T$ such that

$$\bigcup_{k=1}^{N_T} [t_k, t_{k+1}) = T.$$  

(35)

The main result is presented in Theorem 1.

**Theorem 1.** Let the dynamic system (7) satisfy Assumption 1 with a given trajectory defined on $[0, T]$ satisfying Assumption 2, and the observer come from (24) with parameter $\eta$ coming from Lemma 1. Let the positive pair $(\Delta, \Delta_0)$ satisfying $\Delta_b > \Delta$. For any given positive parameters $\Delta, D_o, D_X$, there exists a positive pair $k^*_\eta$ and $h^*$ such that for any $K_o \geq k^*_\eta I$, the solutions of (33) for any $t \in [0, T]$ satisfy

$$\|E_o(t)\|_F \leq \Delta_b$$

for all $\|E_o(t_0)\| \leq \Delta$.

**Proof 3.** The proof consists of three steps. Step 1 shows that the estimation error decreases within the interval $[t_k, t_{k+1})$. Step 2 shows that when a switching happens at the jump point $\tau^+_{k+1}$, the estimation error increases but is bounded. Step 3 concludes the result with the help of Assumption 2.

**Step 1**

The first cost function is used as a Lyapunov candidate:

$$L(E_o) = \frac{1}{2} \sum_{i \in \mathbb{G}(X, Y)} \|\Xi_i\|^2_\ell + \frac{1}{2} \sum_{i \in \mathbb{G}(X, Y)} h\|\Phi_i\|^2_{\ell_p}$$

$$= \sum_{i \in \mathbb{G}(X, Y)} tr(\Xi_i \Xi_i^T) + l \sum_{i \in \mathbb{G}(X, Y)} htr(K_o \Xi_i \Xi_i^T)$$

where $\Xi_i$ is in (22) and $\Phi_i = K_o$ is the Cholesky decomposition of $K_o$. The derivative of the Lyapunov candidate along the trajectory of $E_o$ is

$$\dot{L}(E_o) = l \sum_{i \in \mathbb{G}(X, Y)} tr(\Xi_i \Xi_i^T) + l \sum_{i \in \mathbb{G}(X, Y)} htr(K_o \Xi_i \Xi_i^T)$$

Substituting $\Xi_i$ by (34) yields:

$$\dot{L}(E_o) = -l \sum_{i \in \mathbb{G}(X, Y)} tr(\Xi_i \Xi_i^T) + l \sum_{i \in \mathbb{G}(X, Y)} htr(K_o \Xi_i \Xi_i^T)$$

(37)

where $\mathcal{R} = A_d (\Psi + Q(\hat{Y}, Y) + D)$. Substituting $\Psi(26)$, $Q(\hat{Y}, Y)$ (27) into the above equation yields

$$\mathcal{R} = \mathcal{R}_n + A_d D$$

$$= \mathbb{P} \left( \ell \sum_{i \in \mathbb{G}(X, Y)} \Xi_i C_i^T \right) + h \mathbb{P} \left( \ell \cdot K_o \sum_{i \in \mathbb{G}(X, Y)} \Xi_i \Gamma_i^T \right)$$

$$+ A_d D$$

(38)

Since $\mathcal{R} \in se(3)$, using the property in (5), the derivative of the Lyapunov function becomes

$$\dot{L}(E_o) = -tr \left( \mathbb{P}(\ell) \sum_{i \in \mathbb{G}(X, Y)} \Xi_i \Gamma_i^T \mathcal{R} \right)$$

$$-tr \left( h \mathbb{P}(\ell \cdot K_o \sum_{i \in \mathbb{G}(X, Y)} \Xi_i \Gamma_i^T) \mathcal{R} \right).$$

Note that $\mathbb{P}(\Xi_i \Gamma_i^T) = \mathbb{P}(\Xi_i \Xi_i^T + C_i^T)$ and $\mathbb{P}(\Xi_i \Xi_i^T) = 0$, leading to $\mathbb{P}(\Xi_i \Gamma_i^T) = \mathbb{P}(\Xi_i \Xi_i^T)$. Consequently,

$$\dot{L}(E_o) = -tr \left( \mathbb{R}_n^T \mathcal{R} \right) = tr \left( \mathbb{R}_n^T (\mathcal{R}_n + A_d D) \right)$$

$$= -\|\mathbb{R}_n\|_F^2 - tr(\mathbb{R}_n^T A_d D)$$

$$\leq -\|\mathbb{R}_n\|_F^2 + \|\mathbb{R}_n\|_F \| A_d D \|_F$$

$$= -\|\mathbb{R}_n\|_F \left( \| A_d D \|_F - \| A_d D \|_F \right).$$

For a given $d_D$, there always exist a positive definite matrix $K_o$ and a positive parameter $h$ such that $\dot{L}(E_o) < 0$ when $\|E_o\|_F \geq \Delta$, indicating the boundedness of the estimation error.

**Step 2**
Next will prove the estimation error is diverging at time $t = t_{k+1}$ but still bounded, as

$$
\|E_0(\tau_{k+1})\|_F \leq (\|E_0(\tau_k)\|_F + 1)e^{d_D(\tau_{k+1} - \tau_{k+1})} \tag{39}
$$

By applying the Gronwall-Bellman inequality, the solution to equation (33) with the observer in (24) can guarantee the following inequality:

$$
\|E_0(\tau_{k+1})\|_F \leq \left| \begin{array}{c}
E_0(\tau_k) - e^{-\int_{\tau_{k+1}}^{\tau_k} R(\tau)d\tau} \\
+ e^{-\int_{\tau_{k+1}}^{T_k} D(\tau)d\tau}
\end{array} \right|_F \tag{40}
$$

According to Lemma 1, if $E_0 > A$, all components will be in the set $\mathcal{G}^*(X, Y)$, and $\|Y - O_B\| > \eta$. In this case, the $0 < h(\cdot) < 1$ for all $i \in \mathcal{O}(Y)$. It yields $h\mathcal{P} (\ell, K, \Sigma_{i \in \mathcal{O}(X, Y)} \Xi_{i \in \mathcal{T}_1}) \to 0$ and $e^{-\int_{\tau_{k+1}}^{\tau_k} R(\tau)d\tau} \to I_{4 \times 4}$ Hence we can get

$$
e^{-\int_{\tau_{k+1}}^{\tau_k} R(\tau)d\tau} \approx e^{-\int_{\tau_{k+1}}^{\tau_k} A d\xi D(\tau)d\tau}
$$

Equation (40) becomes

$$
\|E_0(\tau_{k+1})\|_F \leq \left| \begin{array}{c}
E_0(\tau_k) - e^{-\int_{\tau_{k+1}}^{\tau_k} D(\tau)d\tau} \\
+ e^{-\int_{\tau_{k+1}}^{T_k} D(\tau)d\tau}
\end{array} \right|_F
$$

Since $\text{esssup}_{[0, T]} \|D(t)\|_F \leq d_D$ according to Assumption 1, it yields

$$
\left| \begin{array}{c}
e^{-\int_{\tau_{k+1}}^{\tau_k} A d\xi D(\tau)d\tau}
\end{array} \right|_F \leq e^{d_D(\tau_{k+1} - \tau_{k+1})}, \text{which leads to}
$$

$$
\|E_0(\tau_k + T_k)\|_F \leq (\|E_0(\tau_k)\|_F + 1)e^{d_D(\tau_{k+1} - \tau_{k+1})} \tag{41}
$$

Step 3

As Assumption 2 holds, there is only $N_T$ switch times during the time interval $t \in [0, T]$. We can show the boundedness of the estimation error within $[0, T]$, completing the proof.

4. SIMULATION

This section presents simulation examples to evaluate the performance of the proposed observer in the presence of modeling uncertainties, measurement noises, visual geometrical constraints, and mis-identification.

The system in (7) with the output in (16) is used. The space of interest is a $40 \times 40 \times 40$ cubic space, and the number of feature marks is 150. They are randomly distributed in this space.

Modeling Uncertainty and Measurement Noise

The modeling uncertainty $D(t)$ in (7) is given

\[
D = \begin{bmatrix}
0 & \frac{w\pi}{2500} & 0 & \frac{w\cos}{2} \\
\frac{w\pi}{2500} & 0 & \frac{w\pi}{1500} & 0 \\
0 & -\frac{w\pi}{2} & \frac{w\sin}{2} & \frac{t}{300} - \frac{\pi}{6} \\
0 & 0 & \frac{w\sin}{2} & \frac{t}{700} + \frac{\pi}{7}
\end{bmatrix}
\]

where $w \in \mathbb{R}$ can be selected. The measurement noise signal in (16) is selected as:

$$
A_i = sr \begin{bmatrix}
0.3 \cos \left(\frac{t \cdot i}{2}\right) & 0.2 \sin \left(\frac{t \cdot i}{5}\right) & 0.1 \sin \left(\frac{t \cdot i}{4}\right)
\end{bmatrix}^T
$$

where $sr \in \mathbb{R}$ is the amplitude of this noise. Here $t \in [0, 3000]$ and $i = 1, \ldots, 150$. By calculation, the upper bound of $\|D\|_F$ and $\|A_i\|_F$ are $d_D = |w| \sqrt{2(\frac{\pi^2}{2500^2} + \frac{\pi^2}{1500^2})}$ and $D_A = |sr| \sqrt{0.15}$.

Mis-identification

To examine the performance of the proposed observer with respect to the mis-identification, several false feature marks are forced to be recognized and evaluated as $\mathcal{G}^*(X, Y)$, during the time duration $t \in [400, 700]$. They can be represented as

$$
Y_{(1, 150, 20, 77, 11, 67)} = \begin{bmatrix}
4 & 4.5 & -1 & -1 & 1 \\
2.8 & -2 & -1.8 & 1 \\
2 & -2 & 1.5 & 1 \\
3 & -1.5 & -1 & 1 \\
2.5 & -2 & -1 & 1 \\
3.5 & -2 & 1 & 1
\end{bmatrix}
$$

in practice, mis-identification happens quickly at some time instants. In the simulations, a longer mis-identification period to use to test the robustness of the proposed observer.

Visual Geometrical Constraints

The parameters of the geometrical constraint denoted as $\theta = [f_{\min}, f_{\max}, \theta_H, \theta_V]^T = \left[1, 8, \frac{3}{4}, \frac{2}{3}\pi\right]^T$ in the simulation setup.

Initial Condition

The initial condition of the system and the observer are given as $\tilde{X} = I_{4 \times 4}, X = \begin{bmatrix}
0.9211 & -0.3817 & 0.0774 & 4 \\
0.3894 & 0.9027 & -0.1830 & 3 \\
0.1987 & 0.9801 & 2 \\
0 & 0 & 0 & 1
\end{bmatrix}$.

Control Input

The control input in the simulation is given as

\[
U = \begin{bmatrix}
0 & 0 & \pi & -4\pi\cos(2\pi t/3000) \\
\pi & 4000 & 0 & 3000 \\
4000 & 0 & 0 & 2\pi\sin(2\pi t/3000) \\
0 & -5000 & 0 & 3000
\end{bmatrix}
\]

The number of switches in $[0, 3000]$ for the given trajectory is $N_T = 76$ during the given time interval.

Figure 4 shows the simulation setup. It includes the distribution of the feature marks, the true trajectories of $X(t)$ over $t \in [0, T]$, and the switch of the detected object due to the visual geometrical constraint.
Two cases are used to validate the effectiveness of the proposed observer. Case 1 shows that the existing observers in Hua et al. (2011); Zhang et al. (2021) cannot work well in the presence of visual geometrical constraints and mis-identification while the proposed observer can. Case 2 further shows the robustness of the proposed observer in the presence of all uncertainties. It also discusses how the choice of parameters ($\eta, K_o$) of the proposed observer will affect its performance.

**Cases 1.** The performances of three observers: OB1 (Hua et al. (2011)), OB2(Zhang et al. (2021)), and proposed observer (OB3) are presented in Figure 5 in the presence of visual geometrical constraints without the mis-identification and noises/uncertainties while Figure 6 shows their performance in the presence of both visual geometrical constraints and the mis-identification.

The parameters of the observers are given as $l = 0.004$, $K_o = 1.5 \cdot I_4$ and $\eta = 6$ are selected based on the estimation of $\Delta$ and the value $m_\theta$ from the given geometrical constraint. It is observed from Figure 5 that both OB1 and OB2 exhibit large estimation errors while the proposed observer still works well with much smaller estimation errors even though the parameters are not well-selected. In addition, the mis-identification has less impact on the proposed observer.

**Cases 2.** Figure 7 shows that the proposed observer can deal with disturbances from the modeling uncertainty, measurement noise, visual geometrical constraints, and mis-identification while neither OB1 nor OB2 has good estimation performance. In this case, the modeling uncertainty and measurement noise are selected to be $w = 0.04$ and $sr = 4$. By carefully selecting $\eta = 6$ from Lemma 1 and $K_o = 1.5 \cdot I$ from Theorem 1, the estimation error is bounded within $[0, 3000]$ for the proposed observer.

Next, different choices of $K_o$ are used in the proposed observer as shown in Figure 8 when $\eta$ is fixed at 6 while the performance of the proposed observer with different choice of $\eta$ when $K_o$ is fixed at $1.5 \cdot I$ is shown in Figure 9. It is observed that larger $K_o$ and $\eta$ will have better robustness.

Fig. 5. The performance comparison in terms of estimation error for three observers in the presence of visual geometrical constraints

Fig. 6. The performance comparison in terms of estimation error for three observers in the presence of visual geometrical constraints and mis-identification

Fig. 7. The performance of estimation error in the presence of all uncertainties within $[0, 3000]$ within $[0, 3000]$

Fig. 8. The estimation errors of the proposed observer with different choice of $K_o$ when $\eta = 6$

Fig. 9. The estimation errors of the proposed observer with different choice of $K_o$ when $K_o = 1.5 \cdot I$

It is pointed out that experiments using UGVs with camera sensors have been conducted to validate the result, but not presented here due to page limit.

5. **CONCLUSION**

This paper proposed a new robust observer dynamic system modeled on Special Euclidean Group (SE3) when...
visual observer can handle modeling uncertainties, measurement noises, visual geometrical constraints, and misidentification to ensure the boundness of estimation error over a finite time interval. Our future work will focus on validating the proposed observer in an experimental setup along with the appropriate observer-based controller design.

REFERENCES

Amri, M.H., Becis, Y., Aubry, D., and Ramdani, N. (2015). Indoor human/robot localization using robust multi-modal data fusion. In 2015 IEEE International Conference on Robotics and Automation (ICRA), 3456–3463. IEEE.

Baldwin, G., Mahony, R., Trumpf, J., Hamel, T., and Cheviron, T. (2007). Complementary filter design on the special euclidean group se (3). In 2007 European Control Conference (ECC), 3763–3770. IEEE.

Bischoff, B., Nguyen-Tuong, D., Streichert, F., Ewert, M., and Knoll, Å. (2012). Fusing vision and odometry for accurate indoor robot localization. In 2012 12th International Conference on Control Automation Robotics & Vision (ICARCV), 347–352. IEEE.

Bourmaud, G., Mégret, R., Giremus, A., and Berthomieu, Y. (2013). Discrete extended kalman filter on lie groups. In 21st European Signal Processing Conference (EUSIPCO 2013), 1–5. IEEE.

Cheng, X. (2019). Stability, Robustness and Switching of Vibration Control Systems. Ph.D. thesis, The University of Melbourne.

Delibasis, K.K., Plagianakos, V.P., and Maglogiannis, I. (2013). Real time indoor robot localization using a stationary fisheye camera. In IFIP International Conference on Artificial Intelligence Applications and Innovations, 245–254. Springer.

Gusrialdi, A., Hatanka, T., and Fujita, M. (2008). Coverage control for mobile networks with limited-range anisotropic sensors. In 2008 47th IEEE Conference on Decision and Control, 4263–4268. IEEE.

Heo, S. and Park, C.G. (2018). Consistent ekf-based visual-inertial odometry on matrix lie group. IEEE Sensors Journal, 18(9), 3780–3788.

Hervé, J.M. (1999). The lie group of rigid body displacements, a fundamental tool for mechanism design. Mechanism and Machine theory, 34(5), 719–730.

Hua, M.D., Zamani, M., Trumpf, J., Mahony, R., and Hamel, T. (2011). Observer design on the special euclidean group se (3). In 2011 50th IEEE Conference on Decision and Control and European Control Conference, 8169–8175. IEEE.

Khalil, H.K. (2002). Nonlinear systems third edition. Patrick Hall, 115.

Lagman, C., Trumpf, J., and Mahony, R. (2009). Gradient-like observers for invariant dynamics on a lie group. IEEE Transactions on Automatic Control, 55(2), 367–377.

Li, X., Tan, Y., Mareels, I., and Chen, X. (2018). Compati bile formation set for uavs with visual sensing constraint. In 2018 Annual American Control Conference (ACC), 2497–2502. IEEE.

Mavrinac, A. and Chen, X. (2013). Modeling coverage in camera networks: A survey. International journal of computer vision, 101(1), 205–226.

Ocana, M., Bergasa, L., Sotelo, M., Nuevo, J., and Flores, R. (2005). Indoor robot localization system using wi signal measure and minimizing calibration effort. In Proceedings of the IEEE International Symposium on Industrial Electronics, volume 4, 1545–1550.

Park, F.C., Bobrow, J.E., and Ploen, S.R. (1995). A lie group formulation of robot dynamics. The International journal of robotics research, 14(6), 609–618.

Salcudean, S. (1991). A globally convergent angular velocity observer for rigid body motion. IEEE transactions on Automatic Control, 36(12), 1493–1497.

Vasconcelos, J.F., Cunha, R., Silvestre, C., and Oliveira, P. (2010). A nonlinear position and attitude observer on se (3) using landmark measurements. Systems & Control Letters, 59(3-4), 155–166.

Zhang, L., Chen, Z., Cui, W., Li, B., Chen, C., Cao, Z., and Gao, K. (2020). Wifi-based indoor robot positioning using deep fuzzy forests. IEEE Internet of Things Journal, 7(11), 10773–10781.

Zhang, T., Tan, Y., and Chen, X. (2021). A damping-based robust nonlinear observer on se(3). In 2021 American Control Conference (ACC), 78–83. IEEE.

Zhang, X., Chen, X., Farzadpour, F., and Fang, Y. (2018). A visual distance approach for multicamera deployment with coverage optimization. IEEE/ASME Transactions on Mechatronics, 23(3), 1007–1018.

6. APPENDICES

6.1 Proof of Property 1

Since $i \in G(X, Y)$ and the output $Y_i = X^{-1}C_i = C_i'$, it has

$$C_i' - O_x^B = \begin{bmatrix} x_v & y_v & \frac{f_{\min} + f_{\max}}{2} & z_v & 0 \end{bmatrix}^T.$$ 

Note that according to (11)

$$\left\| C_i' - O_x^B \right\|_F = \sqrt{\text{tr}} \left( (C_i' - O_x^B)^T (C_i' - O_x^B) \right)$$

$$= x_v^2 + (y_v - \frac{f_{\min} + f_{\max}}{2})^2 + z_v^2.$$ 

If $i \in G$, the system output $Y_i$ has to satisfy the constraint in (11), which will provide the bound of each term in (42). This completes the proof.

6.2 Proof of Lemma 1

For any $i \in G(X)$, the following inequality holds:

$$\left\| \dot{Y}_i - O_c^B \right\|_F = \left\| Y_i - Y_i + Y_i - O_c^B \right\|_F \leq \left\| Y_i - Y_i \right\|_F + M_\theta,$$

from Property 1. It is noted that

$$\left\| \dot{Y}_i - Y_i \right\|_F = \left\| (\dot{X}^{-1}X - I)X^{-1}C_i \right\|_F \leq \left\| (\dot{X}^{-1}X - I) \right\|_F \left\| Y_i \right\|_F$$

$$\Rightarrow \left\| Y_i \right\|_F \leq \left\| Y_i - O_c^B + O_c^B \right\|_F \leq M_\theta + \left\| O_c^B \right\|_F.$$

As the Frobenius norm of $\left\| \dot{X}^{-1}X - I \right\|_F$ can be approximated by

$$\left\| \dot{X}^{-1}X - I \right\|_F \approx \left\| \dot{X}X^{-1} - I \right\|_F \leq \Lambda$$

as $\dot{X}X^{-1} \to I$. Consequently, it leads to

$$\left\| \dot{Y}_i - O_c^B \right\|_F \leq \left\| (\dot{X}^{-1}X - I) \right\|_F \left\| Y_i \right\|_F + M_\theta \leq \Lambda (M_\theta + \left\| O_c^B \right\|_F) + M_\theta = \eta,$$

which completes the proof.