RESEARCH ARTICLE

Comparison of different pseudo-linear estimators for vision-based target motion estimation

Zian Ning[1,2] · Yin Zhang[2] · Shiyu Zhao[2]

Abstract

Vision-based target motion estimation based Kalman filtering or least-squares estimators is an important problem in many tasks such as vision-based swarming or vision-based target pursuit. In this paper, we focus on a problem that is very specific yet we believe important. That is, from the vision measurements, we can formulate various measurements. Which and how the measurements should be used? These problems are very fundamental, but we notice that practitioners usually do not pay special attention to them and often make mistakes. Motivated by this, we formulate three pseudo-linear measurements based on the bearing and angle measurements, which are standard vision measurements that can be obtained. Different estimators based on Kalman filtering and least-squares estimation are established and compared based on numerical experiments. It is revealed that correctly analyzing the covariance noises is critical for the Kalman filtering-based estimators. When the variance of the original measurement noise is unknown, the pseudo-linear least-squares estimator that has the smallest magnitude of the transformed noise can be a good choice.

Keywords Pseudo-linear measurements · Kalman filter · Least-squares estimator · Vision-based target motion analysis · Fisher information

1 Introduction

The work in this paper is motivated by the aerial target pursuit task, where one micro aerial vehicle (MAV) uses its onboard camera to detect, localize, and then pursue another flying MAV [1, 2]. The study of this task can be potentially applied to the defence of malicious MAVs. In particular, while MAVs have been applied in many domains nowadays such as video shooting, surveillance, and inspection [3], misused ones have caused serious security problems. Aerial target pursuit may provide a promising technical solution that is complementary to the existing countermeasures [1, 4].

One key problem in the aerial target pursuit task is to use vision measurements to estimate the motion of the target. This problem is the prerequisite for subsequent tasks such as pursuit control. The existing approaches for vision-based motion estimation can be classified into two categories.

In the first category, the vision measurement is modelled as a bearing-only vector that points from the camera to the target. Using bearing-only vectors to estimate the target's motion is usually called *bearing-only* target motion estimation [2, 5, 6]. One fundamental problem about bearing-only estimation is observability. In particular, since the depth/distance of the target cannot be measured directly, it is necessary to ensure proper relative motion between the camera and the target to recover the target's motion. In fact, the camera's motion must have high-order components than the target's motion [5]. Otherwise, the target's motion is unobservable.

In the second category, in addition to the bearing vector, other vision measurements are also assumed to be available [1, 7]. For example, if the exact geometric information of the target is known, its relative position and attitude can be recovered by using a pose estimation algorithm [1]. If the

---

✉ Shiyu Zhao
zhaoshiyu@westlake.edu.cn

Zian Ning
ningzian@westlake.edu.cn

Yin Zhang
zhangyin@westlake.edu.cn

1 Department of Computer Science and Technology, Zhejiang University, Hangzhou 310058, Zhejiang, China

2 School of Engineering, Westlake University, Hangzhou 310024, Zhejiang, China

target can be approximated as a sphere and the diameter of the sphere is known, then the distance of the target can be recovered based on the pin-hole camera model [8].

The advantage of the bearing-only approach is that it does not require knowing the size of the target. As long as the target can be recognized by a vision algorithm, its bearing can be readily obtained from the pixel coordinate of the target in the image. However, the disadvantage is that it requires that the camera's motion must be of higher order than the target's motion. This requirement may be difficult to achieve in practice. In this case, it is favourable to directly estimate the target's distance from the size information of the target. The size information may be available in practice. For example, a vision algorithm may not only be able to detect an MAV but also can classify its model. When we know it is a specific type of commercial MAV, then the size information can be known. In addition, the size information is available in cooperative tasks such as vision-based formation control since all the robots are known in advance [9, 10].

Once we have obtained vision measurements, we can readily establish an estimator such as a Kalman filter to estimate the target's motion. It is notable that vision measurements such as bearing vectors are highly nonlinear functions of the target's motion. The conventional extended Kalman filter shows unstable performances when handling bearing measurements [11]. Therefore, pseudo-linear measurements and pseudo-linear Kalman filters have been widely used in bearing-related estimation problems [2, 12, 13].

This paper focuses on a very specific problem. That is, once we have obtained the vision measurements, there are different ways to convert the nonlinear vision measurements to pseudo-linear. Which pseudo-linear measurement should be used and how to set the covariance matrix for the pseudo-linear noise? These problems are important for achieving high-performance motion estimation. However, we notice that practitioners usually do not pay special attention to them and often make mistakes. Motivated by this, we compare different types of pseudo-linear measurements and provide advice for practitioners on how to properly utilize vision measurements.

First of all, three different pseudo-linear measurements are established based on the bearing and angle measurements. Then, three pseudo-linear Kalman filter (PL-KF) estimators and three pseudo-linear recursive least square (PL-RLS) estimators are built up based on proposed pseudo-linear measurement equations, respectively. To reveal the roles of different measurements, we compared the performances of these estimators based on numerical experiments for three different scenarios.

In the first set of numerical experiments, it is revealed that all the estimators show similar performances even though they are built based on different measurement equations. The experimental results are supported by the theoretical analysis

that the three pseudo-linear measurement equations have the same Fisher information about the states. Although the estimators show similar performances, some differences are still observed. In particular, the three PL-RLS estimators have different performances. That is because the magnitudes of transformed noises in the three different measurement equations are different. However, the PL-RLS estimators do not well incorporate this property. By contrast, the three PL-KF estimators show almost the same performance when the covariances of the measurement noises can be correctly setup.

Furthermore, the covariance setups in the PL-KF estimators are studied in the second set of numerical experiments. Because of the pseudo-linear transformation, the covariance of the pseudo-linear measurement noises are different in different measurement equations. It is revealed that incorrect values of the covariance parameters in PL-KF estimators would downgrade the estimation performance.

In summary, the contribution of this paper is to clarify some mistakes that practitioners may make when conducting vision-based motion estimation in practice. The following advice is given to practitioners. First, if the variances of the original measurement noises are known in advance, any PL-KF estimator is the best choice as long as the calculation of the covariance of the noises is correct. Second, if the variance of the original measurement noise is unknown, the PL-RLS estimator that has the smallest magnitude of the transformed noise can be a good choice.

2 Measurements obtained from visual sensing

This section presents the measurements obtained from vision and the corresponding nonlinear measurement equations.

2.1 Vision measurements

Consider a target MAV flying in 3-D space. Its position and velocity are denoted as $p_T, v_T \in \mathbb{R}^3$ respectively. Let $\ell \in \mathbb{R}$ be the physical size of the wheelbase of the target multirotor MAV. Suppose that we know ℓ. There is an observer carrying a monocular camera to observe the target. The position of the observer camera is denoted as $p_o \in \mathbb{R}^3$.

A bounding box surrounding the target in the image can be obtained if the target MAV can be detected by a vision algorithm. There are two measurements that can be extracted from the bounding box.

First, the position of the bounding box in the image $q_{\text{pix}} = [x_{\text{pix}}, y_{\text{pix}}, 1]^T \in \mathbb{R}^3$ can be used to calculate the bearing vector that points from the observer to the target (see Fig. 1). Denote $g \in \mathbb{R}^3$ as the unit bearing vector, and $\hat{g} \in \mathbb{R}^3$ as the noise corrupted bearing measurement. According to the
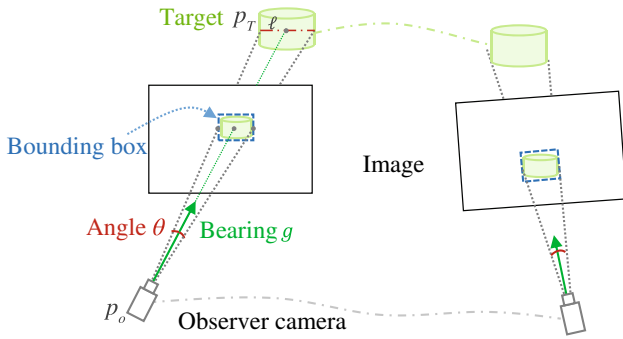
Fig. 1 The bearing vector g and angle θ can be obtained from the bounding box in the image

pin-hole camera model [8], it can be calculated as

$$\hat{g} = \frac{R_c^w P_{cam}^{-1} q_{pix}}{\| R_c^w P_{cam}^{-1} q_{pix} \|},$$

where $P_{cam} \in \mathbb{R}^{3 \times 3}$ is the intrinsic parameter matrix of the camera [8], $R_c^w \in \mathbb{R}^{3 \times 3}$ is the rotation matrix from the camera frame to the world frame.

Second, the size of the bounding box can be used to calculate the angle subtended by the target object in the camera's field of view. Here, the size can be either the width or the height of the bounding box. In this paper, we consider the width which corresponds to the wheelbase of the multi-copter (Fig. 1). Denote $\theta \in (0, \pi/2)$ as the subtended angle, and $\hat{\theta}$ as the noisy angle measurement. According to the pin-hole camera model [8], the angle can be calculated as

$$\hat{\theta} = \arccos \left(\frac{l_{left}^2 + l_{right}^2 - s_{pix}^2}{2 l_{left} l_{right}} \right),$$

where $l_{left} = \sqrt{(f/\alpha)^2 + (\delta x_{pix} - s_{pix}/2)^2 + \delta y_{pix}^2}$ and $l_{right} = \sqrt{(f/\alpha)^2 + (\delta x_{pix} + s_{pix}/2)^2 + \delta y_{pix}^2}$ are distances in pixels from the center of the camera to the middle points of the left and right sides of the bounding box respectively (Fig. 1). Here, f and α are the camera's focal length and single-pixel size, respectively. Here, $\delta x_{pix} = \| x_{pix} - i_{width}/2 \| \in \mathbb{R}$ and $\delta y_{pix} = \| y_{pix} - i_{height}/2 \| \in \mathbb{R}$ are the distances between the center of the bounding box and the center of the image, respectively.

The goal of the vision-based target motion analysis is to estimate the target's position and velocity, p_T and v_T, based on the noisy bearing $\hat{g}$ and angle $\hat{\theta}$ measurements, with known target's physical size ℓ and observer's position p_o.

2.2 Nonlinear measurement equations

We next establish the relationship between the vision measurements and the state of the target.

Denote

$$x = \begin{bmatrix} p_T \\ v_T \end{bmatrix} \in \mathbb{R}^6$$

as the state vector that we aim to estimate. Then, the bearing $\hat{g}$ and angle $\hat{\theta}$ measurements are non-linear functions of the state vector.

The work in [14] proved that the measurement noises from the camera are Gaussian distributed. In the vision-based target motion analysis problem, many works treat the vision-based measurement noises as Gaussian distribution and perform well [1, 2, 7]. In particular, the noise-corrupted non-linear measurement equations are

$$\hat{g} = \frac{p_T - p_o}{r} + \mu, \tag{1}$$

$$\hat{\theta} \approx \frac{\ell}{r} + w, \tag{2}$$

where

$$r = \| p_T - p_o \|$$

is the distance between the target and the observer. Here, $\mu \sim \mathcal{N}(0, \sigma_\mu^2 I_{3 \times 3})$ and $w \sim \mathcal{N}(0, \sigma_w^2)$ are the measurement noises. The approximation in (2) is accurate. Specifically, when $r > 3\ell$, the approximation error is less than 0.08%.

It should be noted that the original noise of bearing measurements is productive $\hat{g} = R(\eta, \epsilon)g$, where $R(\eta, \epsilon) \in \mathbb{R}^{3 \times 3}$ is a rotation matrix that perturbs g. This rotation matrix rotates the vector g by an angle ϵ around the axis η. This productive noise can be transformed into an additive one in (1). This kind of operation has been widely used and shown stable performance in bearing-only localization [2, 15, 16].

3 Pseudo-linear measurement equations

It is noted that the measurement equations in (1) and (2) are nonlinear. When the conventional extended Kalman filter is directly applied, the estimation is usually unstable [11]. Therefore, we convert them to pseudo-linear in this section. We will see that there are three different ways to obtain pseudo-linear measurements and equations. They will be

compared later in this paper.

Three pseudo-linear measurements are proposed first. Then, three pseudo-linear measurement equations are proposed based on transformed pseudo-linear measurements.

3.1 Three pseudo-linear measurements

In the following, we convert the non-linear measurement Eqs. of (1) and (2) into pseudo-linear.

First, we convert the *bearing measurement* in (1) to be pseudo-linear. Multiplying $r P_{\hat{g}}$ on both side of (1) yields

$$P_{\hat{g}} p_o = P_{\hat{g}} p_T + r P_{\hat{g}} \mu, \tag{3}$$

where $P_{\hat{g}} = I - \hat{g}\hat{g}^{\mathrm{T}} \in \mathbb{R}^{3\times3}$ is an orthogonal projection matrix, which plays an important role in the bearing-related estimation and control problems [17]. The matrix $P_{\hat{g}}$ has an important property: $P_{\hat{g}}\hat{g} = 0_{3\times1}$. Here, $P_{\hat{g}} p_o$ is the transformed pseudo-linear measurements, and $r P_{\hat{g}} \mu$ is the transformed noise.

Second, we convert the *angle measurement* in (2) to be pseudo-linear. There are two types of pseudo-linear measurements.

a) Multiplying $r\hat{g}$ on both side of Eq. (2) yields

$$r\hat{\theta}\hat{g} = \ell\hat{g} + r w\hat{g}. \tag{4}$$

Equation (1) implies $r\hat{g} = p_T - p_o + r\mu$. Substituting it into the left-hand of (4) gives

$$\hat{\theta} p_o + \ell\hat{g} = \hat{\theta} p_T + r(\hat{\theta}\mu - w\hat{g}). \tag{5}$$

Here, $\hat{\theta} p_o + \ell\hat{g}$ is the transformed pseudo-linear measurements, and $r(\hat{\theta}\mu - w\hat{g})$ is the corresponding transformed noise.

b) Bearing and angle measurements can be used to calculate the target's position directly. Then, we have another pseudo-linear transformation of the *angle measurement*. Dividing both sides of (5) by $\hat{\theta}$ yields

$$p_o + \frac{\ell}{\hat{\theta}}\hat{g} = p_T + r\left(\mu - \frac{w}{\hat{\theta}}\hat{g}\right). \tag{6}$$

Here, $p_o + \frac{\ell}{\hat{\theta}}\hat{g}$ is the transformed pseudo-linear measurements, and $r\left(\mu - \frac{w}{\hat{\theta}}\hat{g}\right)$ is the corresponding transformed noise.

To sum up, $P_{\hat{g}} p_o$, $\hat{\theta} p_o + \ell\hat{g}$, and $p_o + \frac{\ell}{\hat{\theta}}\hat{g}$ in the left-hand of Eqs. (3), (5), and (6) are three transformed pseud-linear measurements. It should be noted that the corresponding transformed noises are different.

3.2 Three pseudo-linear measurement equations

Choosing one or combining two of the proposed pseudo-linear measurements properly yields the pseudo-linear measurement equations that can be used in Kalman filtering or least square algorithms. To fully compare the difference of transformed noises under different transformations, we propose three pseudo-linear measurement equations that correspond to three PLKF or PLLS estimators.

First, the pseudo-linear measurements in (6) directly give the target's position. As a result, (6) itself can be seen as an independent measurement equation. Rewriting equation (6) in terms of the state vector x yields the first pseudo-linear measurement equation:

$$z_1 = H_1 x + v_1, \tag{7}$$

where

$$z_1 = p_o + \frac{\ell}{\hat{\theta}}\hat{g} \in \mathbb{R}^3,$$
$$H_1 = \begin{bmatrix} I_{3\times3} & 0_{3\times3} \end{bmatrix} \in \mathbb{R}^{3\times6},$$
$$v_1 = r\left(\mu - \frac{w}{\hat{\theta}}\hat{g}\right) \in \mathbb{R}^3,$$

are the corresponding pseudo-linear measurement, pseudo-linear measurement matrix, and transformed noises, respectively. Although the expression of v_1 is complex, it can be approximately treated as a zero mean Gaussian noise $v_1 \sim \mathcal{N}(0, \Sigma_{v_1})$. This kind of approximation has been widely used in many works [2, 12, 13, 18]. Here Σ_{v_1} can be calculated by

$$\Sigma_{v_1} = r^2\left(\sigma_\mu^2 I_{3\times3} + \frac{\sigma_w^2}{\hat{\theta}^2}\hat{g}\hat{g}^{\mathrm{T}}\right) \in \mathbb{R}^{3\times3}. \tag{8}$$

Here, the distance r is unknown. We use $\hat{r} = \|\hat{p}_T - p_o\|$ to replace it [2, 19]. Here $\hat{p}_T$ is the estimated target's position.

Second, the pseudo-linear measurements in (5) have a pseudo-linear relationship with the target's position. Rewriting (5) in terms of the state vector x yields the second pseudo-linear measurement equation:

$$z_2 = H_2 x + v_2, \tag{9}$$

where

$$z_2 = \hat{\theta} p_o + \ell\hat{g} \in \mathbb{R}^3,$$
$$H_2 = \begin{bmatrix} \hat{\theta} I_{3\times3} & 0_{3\times3} \end{bmatrix} \in \mathbb{R}^{3\times6},$$
$$v_2 = r(\hat{\theta}\mu - w\hat{g}) \in \mathbb{R}^3,$$

are the corresponding pseudo-linear measurement, pseudo-linear measurement matrix, and transformed noise, respectively.

Similar to ν_1, we approximate ν_2 as a Gaussian noise $\nu_2 \sim \mathcal{N}(0, \Sigma_{\nu_2})$, where the covariance matrix Σ_{ν_2} can be calculated as:

$$\Sigma_{\nu_2} = \hat{r}^2(\hat{\theta}^2 \sigma_\mu^2 I_{3\times3} + \sigma_w^2 \hat{g}\hat{g}^{\mathrm{T}}) \in \mathbb{R}^{3\times3}. \tag{10}$$

Third, the pseudo-linear measurements in (3) alone can not recover the target's position, because $P_{\hat{g}}$ does not have full column rank. However, combining pseudo-linear measurements in (3) and (5) together can. Combining (3) and (5) together and rewriting it in terms of the state vector x yields the third pseudo-linear measurement equation:

$$z_3 = H_3 x + \nu_3, \tag{11}$$

where

$$z_3 = \begin{bmatrix} P_{\hat{g}} p_o \\ \hat{\theta} p_o + \ell \hat{g} \end{bmatrix} \in \mathbb{R}^6,$$

$$H_3 = \begin{bmatrix} P_{\hat{g}} & 0_{3\times3} \\ \hat{\theta} I_{3\times3} & 0_{3\times3} \end{bmatrix} \in \mathbb{R}^{6\times6},$$

$$\nu_3 = r \begin{bmatrix} P_{\hat{g}} \mu \\ \hat{\theta} \mu - w\hat{g} \end{bmatrix} \in \mathbb{R}^6,$$

are the corresponding pseudo-linear measurement, pseudo-linear measurement matrix, and transformed noise, respectively.

Similarly, the transformed noise ν_3 is treated as a Gaussian noise $\nu_3 \sim \mathcal{N}(0, \Sigma_{\nu_3})$, where the covariance matrix Σ_{ν_3} can be calculated by

$$\Sigma_{\nu_3} = \hat{r}^2 \begin{bmatrix} \sigma_\mu^2 P_{\hat{g}} & \hat{\theta} \sigma_\mu^2 P_{\hat{g}} \\ \hat{\theta} \sigma_\mu^2 P_{\hat{g}} & \hat{\theta}^2 \sigma_\mu^2 I_{3\times3} + \sigma_w^2 \hat{g}\hat{g}^{\mathrm{T}} \end{bmatrix} \in \mathbb{R}^{6\times6}. \tag{12}$$

To sum up, the key specifications of three pseudo-linear measurement equations are listed in Table 1.

4 Pseudo-linear estimators

This section proposes three PL-KF estimators and three PL-RLS estimators based on the proposed three pseudo-linear measurement equations.

4.1 PL-KF estimators

To estimate the motion of the target, it is common to model its motion as a noise-driven double integrator:

$$x(t_{k+1}) = F x(t_k) + q(t_k), \tag{13}$$

where

$$F = \begin{bmatrix} I_{3\times3} & \delta t I_{3\times3} \\ 0_{3\times3} & I_{3\times3} \end{bmatrix} \in \mathbb{R}^{6\times6},$$

where δt is the sampling time, and I and 0 are identity and zero matrices, respectively. Here, $q \sim \mathcal{N}(0, \Sigma_q)$ is the process noise, whose covariance is

$$\Sigma_q = \mathrm{diag}(0, 0, 0, \sigma_v^2, \sigma_v^2, \sigma_v^2) \in \mathbb{R}^{6\times6},$$

where σ_v is the standard deviation of the target's velocity.

With the state transition Eq. (13) and three pseudo-linear measurement equations, three different PL-KF estimators can be realized. They all share the same filtering procedure as shown below. The differences between these three PL-KF estimators are reflected in z_i, H_i, and Σ_{ν_i}. The prediction steps are

$$\hat{x}^-(t_k) = F\hat{x}(t_{k-1}),$$
$$P_{\mathrm{kf}}^-(t_k) = F P_{\mathrm{kf}}(t_{k-1}) F^T + \Sigma_q,$$

where $\hat{x}^-(t_k) \in \mathbb{R}^6$ and $P^-(t_k) \in \mathbb{R}^{6\times6}$ are the prior estimated states and covariance matrix, respectively. The correction and update steps are

$$K_{\mathrm{kf}}(t_k) = P_{\mathrm{kf}}^-(t_k) H_i^{\mathrm{T}}(t_k) [H_i(t_k) P_{\mathrm{kf}}^-(t_k) H_i^{\mathrm{T}}(t_k) + \Sigma_{\nu_i}]^{-1},$$
$$\hat{x}(t_k) = \hat{x}^-(t_k) + K_{\mathrm{kf}}(t_k)[z_i(t_k) - H_i(t_k)\hat{x}^-(t_k)],$$
$$P_{\mathrm{kf}}(t_k) = [I - K_{\mathrm{kf}}(t_k) H_i(t_k)] P_{\mathrm{kf}}^-(t_k),$$

where $K(t_k)$ is the gain matrix of Kalman filter, $\hat{x}(t_k)$ and $P(t_k)$ are posterior estimated states and corresponding covariance matrix.

4.2 PL-RLS estimators

Similarly, three PL-RLS estimators can be proposed based on the proposed three pseudo-linear measurement equations.

The form of the least square algorithm is as follows:

$$x(t_k) = (A_i^{\mathrm{T}} A_i)^\dagger A_i^{\mathrm{T}} b_i,$$

where

$$A_i = \begin{bmatrix} H_i(t_1) F^{-(k-1)} \\ H_i(t_2) F^{-(k-2)} \\ \vdots \\ H_i(t_{k-1}) F^{-1} \\ H_i(t_k) \end{bmatrix}$$

Table 1 Comparison of key specifications of three pseudo-linear measurement equations

	z_i	H_i	v_i	Σ_{v_i}
PL-KF1/PL-RLS1	$p_o + \dfrac{\ell}{\hat{\theta}}\hat{g}$	$[I_{3\times3} \ 0_{3\times3}]$	$r\left(\mu - \dfrac{w}{\hat{\theta}}\hat{g}\right)$	$r^2\left(\sigma_\mu^2 I_{3\times3} + \dfrac{\sigma_w^2}{\hat{\theta}^2}\hat{g}\hat{g}^{\mathrm{T}}\right)$
PL-KF2/PL-RLS2	$\hat{\theta}p_o + \ell\hat{g}$	$[\hat{\theta}I_{3\times3} \ 0_{3\times3}]$	$r(\hat{\theta}\mu - w\hat{g})$	$r^2(\hat{\theta}^2\sigma_\mu^2 I_{3\times3} + \sigma_w^2\hat{g}\hat{g}^{\mathrm{T}})$
PL-KF3/PL-RLS3	$\begin{bmatrix} P_{\hat{g}}p_o \\ \hat{\theta}p_o + \ell\hat{g} \end{bmatrix}$	$\begin{bmatrix} P_{\hat{g}} & 0_{3\times3} \\ \hat{\theta}I_{3\times3} & 0_{3\times3} \end{bmatrix}$	$r\begin{bmatrix} P_{\hat{g}} \\ \hat{\theta}\mu - w\hat{g} \end{bmatrix}$	$r^2\begin{bmatrix} \sigma_\mu^2 P_{\hat{g}} & \hat{\theta}\sigma_\mu^2 P_{\hat{g}} \\ \hat{\theta}\sigma_\mu^2 P_{\hat{g}} & \hat{\theta}^2\sigma_\mu^2 I_{3\times3} + \sigma_w^2\hat{g}\hat{g}^{\mathrm{T}} \end{bmatrix}$

and

$$b_i = \begin{bmatrix} z_i(t_1) \\ z_i(t_2) \\ \vdots \\ z_i(t_{k-1}) \\ z_i(t_k) \end{bmatrix}.$$

Here, t_1 is the initial time, and t_k is the current time.

It should be noted that the original form of the least square algorithm uses all times of measurements, which shows unstable performance if the target has time-varying velocity. We use the recursive least square (RLS) algorithm together with a time-decay factor, which can solve this problem [20].

The process of the recursive least square algorithm is as follows:

$$\hat{x}^-(t_k) = F\hat{x}(t_{k-1}),$$
$$P_{\mathrm{ls}}^-(t_k) = FP_{\mathrm{ls}}(t_{k-1})F^{\mathrm{T}},$$
$$K_{\mathrm{ls}}(t_k) = P_{\mathrm{ls}}^-(t_k)H_i^{\mathrm{T}}(t_k)[H_i(t_k)P_{\mathrm{ls}}^-(t_k)H_i^{\mathrm{T}}(t_k) + \lambda I]^{-1},$$
$$\hat{x}(t_k) = \hat{x}^-(t_k) + K_{\mathrm{ls}}(t_k)[z_i(t_k) - H_i(t_k)\hat{x}^-(t_k)],$$
$$P_{\mathrm{ls}}(t_k) = \lambda^{-1}[I - K_{\mathrm{ls}}(t_k)H_i(t_k)]P_{\mathrm{ls}}^-(t_k),$$

where $\lambda \in \mathbb{R}$ is the decay factor.

As can be seen, the differences between the three PL-RLS estimators are reflected in z_i and H_i.

5 Numerical results

This section presents the simulation results to verify the performances of three PL-KF estimators and three PL-RLS estimators. Three different scenarios are considered in the simulation: 1) MAV tracking scenario; 2) MAV guidance scenario; 3) ground camera monitoring scenario. 1000 independent experiments are done for each scenario. Deep analyses of the results are also given. We analyze the performances of six estimators from the perspective of transformed noises, the Fisher information, and the architecture of algorithms.

5.1 Simulation setup

First, we give the simulation setup. All experiments use the same setup and parameters.

The variances of measurements' noises are set to $\sigma_\mu = 0.01$, and $\sigma_w = 0.01$. The variance of the process noise is set to $\sigma_v = 0.1$. The decay factor in the PL-RLS estimators is $\lambda = 0.8$. The initial estimated states are set to $\hat{p}_T = p_o + \hat{r}\hat{g}$ and $\hat{v}_T = [0, 0, 0]^{\mathrm{T}}$. The initial variance of estimated states is set to $\hat{P} = I$. The simulation time interval is set to dt = 0.1 s.

1) MAV tracking scenario. The initial target's position is at $p_T = [10, 0, 5]^{\mathrm{T}}$. The target's motion is a maneuver with two 90°'s turns, whose radius is 10 m. The magnitude of the target's velocity is 2 m/s. Then, the magnitude of the target's lateral acceleration can be calculated as around 0.4 m/s². The initial observer's position is set to $p_o = [0, 0, 10]^{\mathrm{T}}$. The observer's velocity is controlled by a tracking algorithm

$$v_o = \hat{v}_T + k\frac{r^2 - d_{\mathrm{track}}^2}{d_{\mathrm{track}}^2}\hat{g},$$

where $k = 1$ is the control parameter, and $d_{\mathrm{track}} = 5$ is the desired tracking distance. As a result, the initial horizontal distance between two MAVs is 10 m. With the observer tracking the target, their horizontal distance gradually narrowed to 5 m. The trajectory of two MAVs is shown on the left-hand of Fig. 2.

2) MAV guidance scenario. The target moves along a straight line from the initial position $p_T = [10, 0, 5]^{\mathrm{T}}$ with the speed of 2 m/s. The initial observer's position is set to $p_o = [0, 0, 0]^{\mathrm{T}}$. The observer's velocity is controlled by a PNG guidance law to pursue the target with a speed of 2.4 m/s. As a result, their distance is getting smaller and smaller. The trajectory of two MAVs is shown on the left-hand of Fig. 3.

3) Ground camera monitoring scenario. The setup of the target MAV is the same as the MAV tracking scenario. While, the observer remains stationary at the position of $p_o = [20, 10, 0]^{\mathrm{T}}$. This scenario simulates the ground monitoring camera. As a result, the target MAV flies overhead the observer, and their distance first decreases and then increases.

The trajectory of two MAVs is shown in the left-hand of Fig. 4.

5.2 Simulation 1: Comparison between different estimators

The estimation performances of three PL-KF estimators and three PL-RLS estimators for three scenarios are shown in Figs. 2, 3, and 4 respectively.

We use the norm error of estimated states $\|\hat{x} - x\|$ as metrics to evaluate six estimators. The error of estimated states is shown in the middle-upper of Figs. 2, 3, and 4. In the meantime, the average error of estimated states for 1000 times of independent experiments is shown in the middle-down of Figs. 2, 3, and 4. Both results show that all six estimators have similar performance.

In detail, three PL-KF estimators have the same performance, while the performances of three PL-RLS estimators are different. The PL-RLS estimator 3 has the highest performance, PL-RLS estimator 2 is the second, and PL-RLS estimator 1 is the worst. The performance differences of three PL-RLS estimators are affected by the distance between the target and the observer. Bigger distance results in greater differences.

We next analyze the simulation results from three perspectives: the transformed noises, the Fisher information, and the architecture of the algorithms.

First, the results of similar performance between PL-KF and PL-RLS estimators are analyzed from the perspective of Fisher information, which indicates how much information the measurements contain about the states. The Fisher information matrix is calculated by

$$FIM_i = \sum_{t=1}^{t_k} \left[H_i^{\mathrm{T}}(t) \Sigma_{v_i}^{-1}(t) H_i(t) \right].$$

Here, we only consider the target's position as states. As a result, only the first three columns of H_i are extracted.

The average determinant of the Fisher information matrix for 1000 times experiments are shown in the right-upper of Figs. 2, 3, and 4. As can be seen, three different pseudo-linear measurement equations have the same Fisher information, and the determinant of the Fisher information matrix increases with time. It makes sense because the pseudo-linear transformation does not bring new information about the states. The exactly same Fisher information means that six estimators use the same amount of information to estimate states. That is why different estimators have similar performance.

Second, the result of different performances between the three PL-RLS estimators is analyzed from the perspective of transformed noises. We plot the average norm of pseudo-linear measurements' noises for 1000 times of experiments in (3), (5), and (12) in the right-down of Figs. 2, 3, and 4. For a quick review, the average norms of pseudo-linear measurements' noises are calculated by $\|r P_{\hat{g}} \mu\|/3$, $\|r(\hat{\theta}\mu - w\hat{g})\|/3$, and $\|r(\mu - (w/\hat{\theta})\hat{g})\|/6$.

As can be seen, the noises are influenced by the distance, but in most time, the noise of pseudo-linear measurement 3 is the biggest, while the noises of pseudo-linear measurements 1 and 2 are smaller. Compared to the performances of PL-RLS estimators, we can conclude that smaller transformed noises have better performance. From another perspective, the transformed noises are actually the residual error of the least square algorithm.

Third, the result that three PL-KF estimators have exactly the same performance is analyzed from the perspective of algorithm architecture. Comparing the estimation process of PL-KF estimators and PL-RLS estimators, the calculation of the gain matrix in PL-KF estimators considers the covariance of measurements' noises, while PL-RLS estimators do not. Together with the same Fisher information, the performance of the PL-KF estimator is not affected by transformed noises.

5.3 Simulation 2: incorrect parameters in PL-KF estimators

Sometimes, the variances of the original measurement noises are difficult to estimate. Some users may overlook or make a mistake in the calculation of the covariance of the transformed noises when applying PL-KF estimators. This section investigates the influence of inappropriate setting of the parameter (covariance of the pseudo-linear measurement noise) in the PL-KF estimator.

Many new users directly use I as the covariance matrix of the measurement noises. The second simulation experiment setups $\Sigma_{v_i} = I$, and all other parameters and measurement data are the same as in simulation 1.

The simulation results are shown in Fig. 5. As can be seen, inappropriate measurement covariance can bring uncertain effects on the estimation performance in all three scenarios. As a result, our advice is to use the correct calculation of the covariance matrix of the pseudo-linear measurement noises when applying PL-KF estimators.

5.4 Simulation 3: wrong variance of original noises

Sometimes, it is difficult to estimate the variance of the original measurement noises. However, wrong value may bring bad effects. The third simulation experiment setups $\sigma_\mu = 0.2$ and $\sigma_w = 0.2$ when calculating the covariance matrix of Σ_{v_i} in the PL-KF estimators, while the real variances of measurement noises remain $\sigma_\mu = 0.01$ and $\sigma_w = 0.01$. Other parameters and measurement data are the same as in simulation 1.
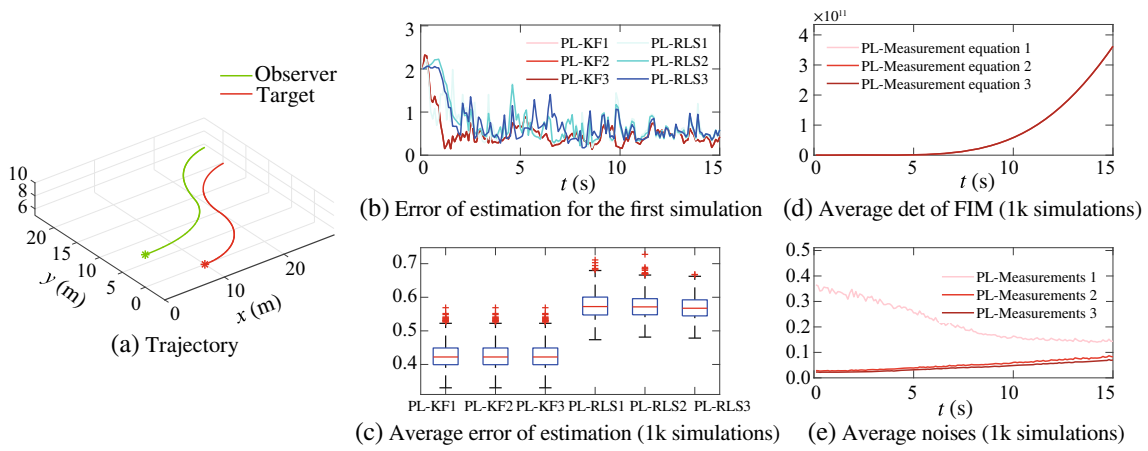
Fig. 2 Simulation results for MAV tracking scenario. The symbol * means the start position
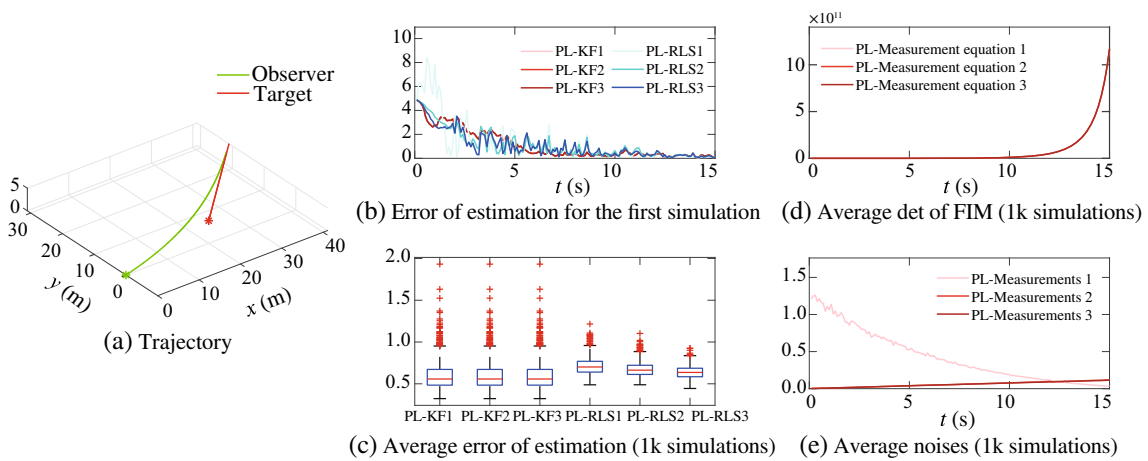

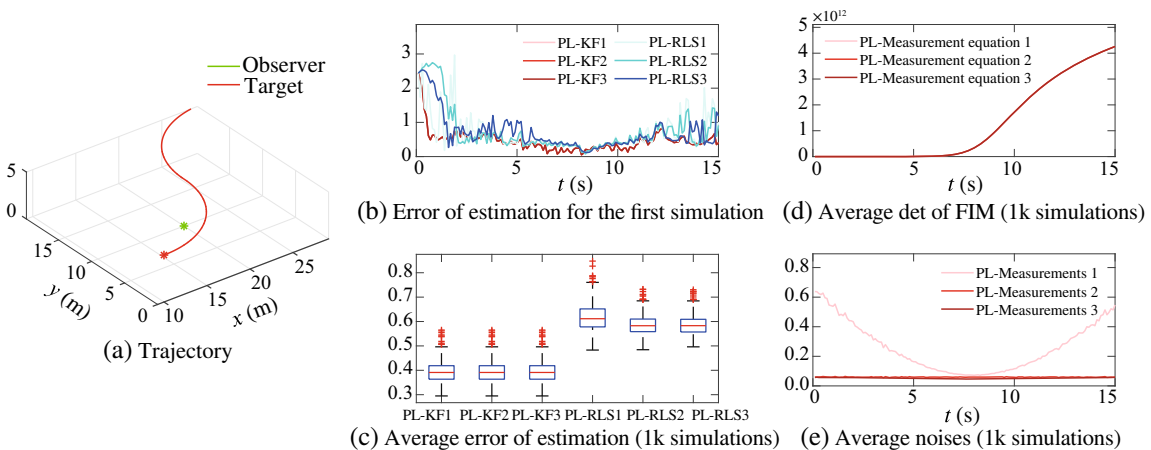**Fig. 3** Simulation results for MAV guidance scenario. The symbol * means the start position

Fig. 4 Simulation results for ground camera monitoring scenario. The symbol * means the start position
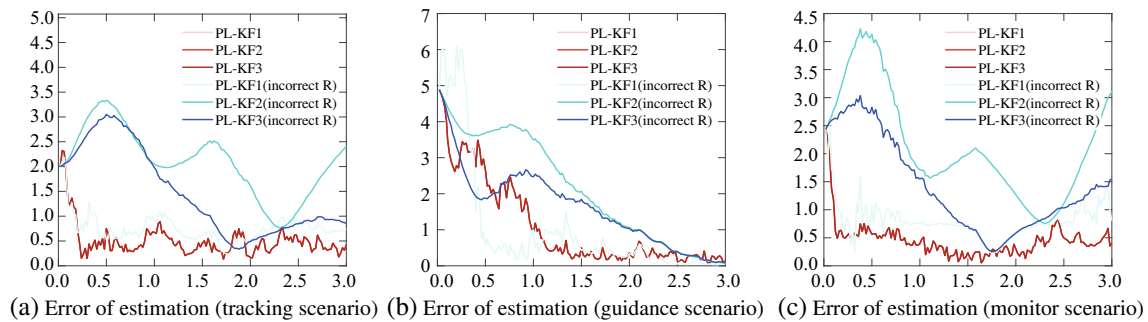
(a) Error of estimation (tracking scenario) (b) Error of estimation (guidance scenario) (c) Error of estimation (monitor scenario)

Fig. 5 Comparison of PL-KF estimators with the same incorrect measurement covariance in three scenarios

(a) Error of estimation (tracking scenario) (b) Error of estimation (guidance scenario) (c) Error of estimation (monitor scenario)
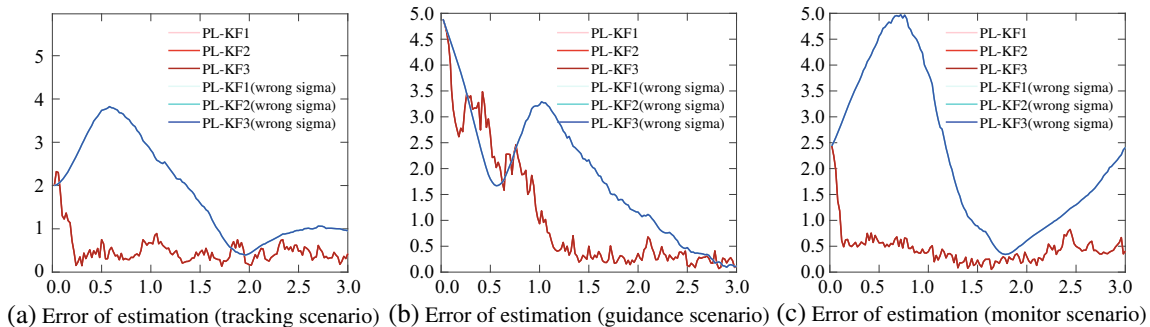
Fig. 6 Comparison of PL-KF estimators with the same wrong original measurement variance

The simulation results are shown in Fig. 6. As can be seen, the performances become worse but remain the same in all three scenarios. As a result, our advice is to use PL-RLS estimators if the variances of the original measurement noises are difficult to estimate.
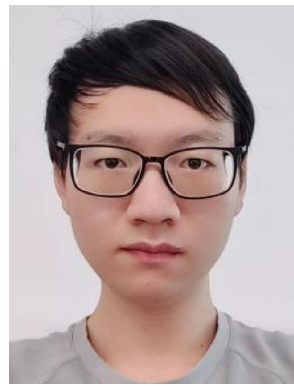
The final advice is to use the PL-KF estimator with correct parameters if the variance of the original measurement noise is known in advance. However, if the variance of the original measurement noise is unknown, the PL-RLS estimator with the smallest transformed noises is the best choice.

6 Conclusion

This paper investigates the effectiveness of using different pseudo-linear measurements on the estimation performance. We choose the vision-based target motion estimation problem as an example to introduce. The vision measurements are modelled as bearing and angle, and are then transformed into three different pseudo-linear measurements. Three different pseudo-linear measurement equations with corresponding transformed noises are then proposed. Based on the proposed pseudo-linear measurement equations, three PL-KF and three PL-RLS estimators are proposed.

Simulation results and corresponding analysis show that a) three PL-KF and three PL-RLS estimators have similar performance; b) the smaller transformed noises, the better performance of the PL-RLS estimator; c) three PL-KF estimators have the same performance.

For the inappropriate parameters in PL-KF estimators, the simulation results show that a) incorrect covariance brings unknown influence; b) wrong original variance of the measurement noises makes the performances worse but the same.

References

1. Vrba, M., & Saska, M. (2020). Marker-less micro aerial vehicle detection and localization using convolutional neural networks. *IEEE Robotics and Automation Letters, 5*(2), 2459–2466. https://doi.org/10.1109/lra.2020.2972819

2. Li, J., Ning, Z., He, S., Lee, C.-H., & Zhao, S. (2022). Three-dimensional bearing-only target following via observability-enhanced helical guidance. *IEEE Transactions on Robotics*, 39(2), 1509–1526. https://doi.org/10.1109/tro.2022.3218268

3. Kang, H., Joung, J., Kim, J., Kang, J., & Cho, Y. S. (2020). Protect your sky: A survey of counter unmanned aerial vehicle systems. *IEEE Access, 8*, 168671–168710. https://doi.org/10.1109/access.2020.3023473

4. Dressel, L., & Kochenderfer, M.J.: (2019). Hunting drones with other drones: Tracking a moving radio target. In: *2019 International Conference on Robotics and Automation (ICRA)*. Montreal, QC, Canada. https://doi.org/10.1109/icra.2019.8794243

5. Fogel, E., & Gavish, M. (1988). Nth-order dynamics target observability from angle measurements. *IEEE Transactions on Aerospace and Electronic Systems, 24*(3), 305–308. https://doi.org/10.1109/7.192098

6. He, S., Shin, H.-S., & Tsourdos, A. (2019). Trajectory optimization for target localization with bearing-only measurement. *IEEE Transactions on Robotics, 35*(3), 653–668. https://doi.org/10.1109/tro.2019.2896436

7. Griffin, B.A., & Corso, J.J. (2021). Depth from camera motion and object detection. In: *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Nashville, TN, USA. https://doi.org/10.1109/cvpr46437.2021.00145

8. Ma, Y., Soatto, S., & Kosecká, J.,& Sastry, S.S. (2012). *An Invitation to 3-D Vision, Interdisciplinary Applied Mathematics* (1st ed., Vol. 26). New York: Springer. https://doi.org/10.1007/978-0-387-21779-6

9. Walter, V., Staub, N., Franchi, A., & Saska, M. (2019). UVDAR system for visual relative localization with application to leader/follower formations of multirotor UAVs. *IEEE Robotics and Automation Letters, 4*(3), 2637–2644. https://doi.org/10.1109/lra.2019.2901683

10. Schilling, F., Schiano, F., & Floreano, D. (2021). Vision-based drone flocking in outdoor environments. *IEEE Robotics and Automation Letters, 6*(2), 2954–2961. https://doi.org/10.1109/lra.2021.3062298

11. Lin, X., Kirubarajan, T., Bar-Shalom, Y., Maskell, S. (2002). Comparison of EKF, pseudomeasurement, and particle filters for a bearing-only target tracking problem. In: *O.E. Drummond (Ed.), SPIE Proceedings*. Orlando, FL, USA. https://doi.org/10.1117/12.478508

12. Lingren, A., & Gong, K. (1978). Position and velocity estimation via bearing observations. *IEEE Transactions on Aerospace and Electronic Systems, AES–14*(4), 564–577. https://doi.org/10.1109/taes.1978.308681

13. Aidala, V., & Nardone, S. (1982). Biased estimation properties of the pseudolinear tracking filter. *IEEE Transactions on Aerospace and Electronic Systems, AES–18*(4), 432–441. https://doi.org/10.1109/taes.1982.309250

14. Boie, R. A., & Cox, I. J. (1992). An analysis of camera noise. *IEEE Transactions on Pattern Analysis & Machine Intelligence, 14*(6), 671–674. https://doi.org/10.1109/34.141557

15. Zhao, S., Chen, B. M., & Lee, T. H. (2013). Optimal sensor placement for target localisation and tracking in 2D and 3D. *International Journal of Control, 86*(10), 1687–1704. https://doi.org/10.1080/00207179.2013.792606

16. Bishop, A. N., Fidan, B., Anderson, B. D. O., Doğançay, K., & Pathirana, P. N. (2010). Optimality analysis of sensor-target localization geometries. *Automatica, 46*(3), 479–492. https://doi.org/10.1016/j.automatica.2009.12.003

17. Zhao, S., & Zelazo, D. (2019). Bearing rigidity theory and its applications for control and estimation of network systems: Life beyond distance rigidity. *IEEE Control Systems Magazine, 39*(2), 66–83. https://doi.org/10.1109/mcs.2018.2888681

18. Dogancay, K. (2015). 3D pseudolinear target motion analysis from angle measurements. *IEEE Transactions on Signal Processing, 63*(6), 1570–1580. https://doi.org/10.1109/tsp.2015.2399869

19. He, S., Wang, J., & Lin, D. (2018). Three-dimensional bias-compensation pseudomeasurement kalman filter for bearing-only measurement. *Journal of Guidance, Control, and Dynamics, 41*(12), 2678–2686. https://doi.org/10.2514/1.g003785

20. Islam, S. A. U., & Bernstein, D. S. (2019). Recursive least squares for real-time implementation [lecture notes]. *IEEE Control Systems, 39*(3), 82–85. https://doi.org/10.1109/mcs.2019.2900788

Zian Ning received the B.Sc. degree in electrical engineering from Northwestern Polytechnical University, Xi'an, China, in 2015, and the M.Sc. degree in aeronautical engineering from Shanghai Jiao Tong University, Shanghai, China, in 2018. He is currently working toward the Ph.D. degree with the Intelligent Unmanned Systems Laboratory, Westlake University, Hangzhou, China. His research interests include vision-based multi-robot cooperation, state estimation, and filtering.

Yin Zhang received the B.Sc. degree in measurement and control technology and instrumentation from the Tianjin University, Tianjin, China, in 2017, and the M.Sc. degree in instrument science and technology from the Beihang University, Beijing, China, in 2020. She is currently working toward the Ph.D. degree with the Intelligent Unmanned Systems Laboratory, Westlake University, Hangzhou, China. Her research interests include domain adaptation, drone detection, and depth estimation.

Shiyu Zhao (Member, IEEE) received the B.Eng. and M.Eng. degrees in electrical engineering from the Beijing University of Aeronautics and Astronautics, Beijing, China, in 2006 and 2009, respectively, and the Ph.D. degree in electrical engineering from the National University of Singapore, Singapore, in 2014. From 2014 to 2016, he was a Post-Doctoral Researcher with the Technion – Israel Institute of Technology, Haifa, Israel, and the University of California at Riverside, Riverside, CA, USA. From 2016 to 2018, he was a Lecturer with the Department of Automatic Control and Systems Engineering, The University of Sheffield, Sheffield, U.K. He is currently an Associate Professor with the School of Engineering, Westlake University, Hangzhou, China. His research interests include theories and applications of robotic systems.