sensors

MDPI

*Article*

# Estimation of a Human-Maneuvered Target Incorporating Human Intention

**Yongming Qin [1], Makoto Kumon [2] and Tomonari Furukawa [3,*]**

[1] Department of Electrical and Computer Engineering, University of Virginia, Charlottesville, VA 22903, USA; yq9vc@virginia.edu
[2] Faculty of Advanced Science and Technology, International Research Organization of Advanced Science and Technology, Kumamoto University, 2-39-1 Kurokami, Chuo-ku, Kumamoto 860-8555, Japan; kumon@gpo.kumamoto-u.ac.jp
[3] Department of Mechanical Engineering, University of Virginia, Charlottesville, VA 22903, USA
* Correspondence: tomonari@virginia.edu

**Abstract:** This paper presents a new approach for estimating the motion state of a target that is maneuvered by an unknown human from observations. To improve the estimation accuracy, the proposed approach associates the recurring motion behaviors with human intentions, and models the association as an intention-pattern model. The human intentions relate to labels of continuous states; the motion patterns characterize the change of continuous states. In the preprocessing, an Interacting Multiple Model (IMM) estimation technique is used to infer the intentions and extract motions, which eventually construct the intention-pattern model. Once the intention-pattern model has been constructed, the proposed approach incorporate the intention-pattern model to estimation using any state estimator including Kalman filter. The proposed approach not only estimates the mean using the human intention more accurately but also updates the covariance using the human intention more precisely. The performance of the proposed approach was investigated through the estimation of a human-maneuvered multirotor. The result of the application has first indicated the effectiveness of the proposed approach for constructing the intention-pattern model. The ability of the proposed approach in state estimation over the conventional technique without intention incorporation has then been demonstrated.

**Keywords:** estimation; tracking; human intention; motion pattern; prediction; multiple model; interacting multiple model; multirotor; drone

check for updates

**Citation:** Qin, Y.; Kumon, M.; Furukawa, T. Estimation of a Human-Maneuvered Target Incorporating Human Intention. *Sensors* **2021**, *21*, 5316. https://doi.org/10.3390/s21165316

Academic Editor: Ondrej Krejcar

Received: 9 June 2021
Accepted: 2 August 2021
Published: 6 August 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

## 1. Introduction

Most dynamic targets to track or engage are either human-maneuvered or humans themselves. Estimating the state of such a human-maneuvered target is essential and important, and has attracted tremendous interest in the last decades [1–4]. Despite the importance, difficulty in the estimation of the human-maneuvered target lies in the motion uncertainty. Even though the motion model of the target may be well or precisely known, the control of the human is often unknown [5]. The motion, as a result, becomes considerably different from the expectation. This gives rise to need for the ability to handle motion uncertainty [6].

For a human-maneuvered target, estimation techniques proposed in the past to handle motion uncertainty can be classified into two types. In the first, a single accurate motion model is developed and used to describe the motion behavior. Due to their robust estimation upon past observations, various Bayesian methods, including the parametric Kalman filters (KFs) and the nonparametric particle filters, have been applied by characterizing the estimation problem and identifying the best estimation technique for the problem [7–13]. Steckenrider and Furukawa [14] proposed to introduce higher-order terms to the motion model through Taylor series expansion and adaptively estimated the target state. Gindele et al. [15] improved the motion model by incorporating the situational

context and extending the state space. As human control is unknown most of the time, conservative motion behaviors such as constant velocity (CV) and constant acceleration (CA) have been incorporated as the most probable human controls [5]. Instead of the motion model, Mehra [16] estimated the covariances of motion noise and observation noise when the filter is detected not working optimally. Almagbile et al. [17] evaluated three adaptation methods of noise covariances and showed improvements over the conventional Kalman filter. It is effective to control uncertainty when the deterministic motion accuracy can no longer be improved. In addition to the model and its uncertainty, other work has dealt with unknown human control and its uncertainty from the motion noise [5]. The human control dominates the motion behavior when the target has a large unconstrained workspace. Bogler et al. [18] represented the time-varying human control deterministically by piecewise constants and estimated the control in addition to the state. Chakrabarty et al. [19] assumed the exogenous input and its derivative to be bounded for a class of nonlinear systems in state estimation. Conte and Furukawa [20] used head motion as an additional indicator when the target is a human and improved the estimation accuracy. While they are more detailed and more adaptively represented, these motion models cannot keep capturing the target motion and estimating its state well particularly if the motion is drastically changed by a human. This is due to the limited representation of a single model.

In the second, multiple models, which are either superposed or switched, have been used to estimate more varying motion behavior [21–25]. The multiple-model (MM) estimation methods extend existing techniques to handle multiple models and cover a wider range of motion behavior [26]. Henk et al. [21] proposed the interacting MM (IMM) method that uses a fixed set of motion models with Markovian switching coefficients. The transition probability and model likelihood were introduced to recursively adapt the model probabilities. Li et al. [27] proposed the variable-structure MM (VSMM) method to overcome the limitations of using a fixed set of models in describing the motion. The VSMM method introduces model set adaption besides the model adaptation and thus can describe and estimate even a broader range of motion behavior. Recently, Xu et al. [28] has engaged with estimating varying motion behaviors by adapting parameters where a fixed coarse grid and an adaptive fine grid of the parameters were combined to determine the models that best match the target motion behavior. Despite the wider covering, it is still insufficient to capture and estimate the target if the human control changes considerably. The MM methods are rather formulated to cover a larger state space given the most probable human control. As the drastic control change may magnify changes in state space, the resulting target state could be beyond the permissible space of the MM estimation. In addition, the use of the deterministic control makes the estimation underestimated as the human control is most uncertain.

This paper presents an approach for estimating the state of a human-maneuvered target by associating the recurring motion behaviors with human intentions. The proposed approach consists of a preprocess, which constructs the so-called intention-pattern model to encapsulate the human intention, and the main process, which allows state estimation using the intention-pattern model. In the preprocess, the intention-pattern model is constructed from the prior observations by running a revised IMM estimation, extracting motion behaviors of each human intention, aligning them, and probabilistically representing its behavior. The main process, then, uses standard state estimation such as KF extensively using the probabilistically represented intention-pattern model. The strength of the proposed approach lies in the incorporation of the intention-pattern model as the incorporation can make the estimation not only accurate in mean but also precise in covariance.

The paper is organized as follows. The next section describes the estimation problem and its solution using the IMM estimation method, which is not only a generalized formulation but also the technique used in the preprocess of the proposed approach. Section 3 presents the proposed estimation approach including the preprocess and the main process. Numerical validation investigating the effectiveness of both the intention-pattern model

and the state estimation is presented in Section 4. Conclusions are summarized in the final section.

## 2. Estimation of a Human-Maneuvered Target Using IMM Estimation

### *2.1. Estimation Problem Formulation*

Figure 1 shows a schematic diagram of the problem of estimating the state of a human-maneuvered target in case the target is a multirotor. When maneuvering a target, a human operator interacts with a controller using an interface device such as a vehicle panel or a joystick. The controller may be implemented in the interface device, in the target, or both. Some parameters of the controller, such as the maximum speed of the target, are usually configurable to realize different motion behavior. The information of human operation and configurable parameters are not known as no communication with the target is available. The estimator does not affect the human operator and target as well. Having the target observed in the field of view (FOV) of a fixed sensor such as a stereo camera, the goal of the problem is to design the estimator to estimate the target state from observations. The discrete motion model of the target and the observation model are generically given by

$$\boldsymbol{x}_k = \boldsymbol{f}(\boldsymbol{x}_{k-1}, \boldsymbol{u}_k, \boldsymbol{w}_k) \tag{1}$$

$$\boldsymbol{z}_k = \boldsymbol{h}(\boldsymbol{x}_k, \boldsymbol{v}_k) \tag{2}$$

where $\boldsymbol{f}$ and $\boldsymbol{h}$ are the motion and the observation models, respectively; $\boldsymbol{x}_k$ is the state of the target at step $k$ to estimate; $\boldsymbol{u}_k$ is the input; $\boldsymbol{z}_k$ is the observation; and $\boldsymbol{w}_k$ and $\boldsymbol{v}_k$ are the motion and observation noises, respectively. Because it is a target maneuvered by a human, $\boldsymbol{f}$ and $\boldsymbol{w}_k$ may not be well known, while $\boldsymbol{u_k}$ is fully unknown. On the other hand, $\boldsymbol{h}$ and $\boldsymbol{v}_k$ are fully known as they are with the sensor(s) of the estimator. With short time interval, it is valid to assume that $\boldsymbol{w}_k$ and $\boldsymbol{v}_k$ are Gaussian. The problem is resultantly defined as the estimation of $\boldsymbol{x}_k$ with no knowledge on $\boldsymbol{u}_k$ and some knowledge on $\boldsymbol{f}$ and $\boldsymbol{w}_k$, given $\boldsymbol{h}$, $\boldsymbol{z}_k$ and $\boldsymbol{v}_k$.

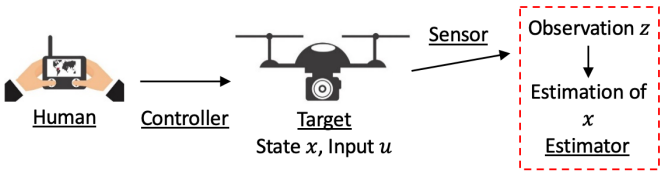



**Figure 1.** The problem of estimating a human-maneuvered target from observations.

### *2.2. IMM Estimation*

Lacking information of $\boldsymbol{f}$ and $\boldsymbol{u}_k$ results in high motion uncertainty. The MM estimation methods deal with motion uncertainty by describing the motion with several motion behaviors called modes, which are denoted by $\mathbb{S} = \{s^j\}, j \in \mathbb{N}$ where $\mathbb{N}$ represents all natural integers. With the definition, a mode $s^j$ is used to represent the motion at step $k$ when it approximates the motion behavior well, i.e., $s_k = s^j$. To describe the behavior with minimum complexity, a mode $s^j$ is most commonly described by a mathematical model $m^i$, which is collectively represented by $\mathbb{M} = \{m^i\}, i \in \mathbb{N}$. A model $m^i$ thus represents the motion behavior at a step, i.e., $s_k = s^j = m^i$. Li et al. [29] is referred for more description. Most MM estimation methods utilize models of known motion behaviors with different parameters such as variants of the CV and CA models. As an example, suppose that $\boldsymbol{f}$ of Equation (1) at step $k$ is approximated by a single linear Gaussian model. The motion model $m^i$ is given by

$$\boldsymbol{x}_k = \boldsymbol{A}_k^{(i)} \boldsymbol{x}_{k-1} + \boldsymbol{B}_k^{(i)} \boldsymbol{u}_k^{(i)} + \boldsymbol{w}_k^{(i)}. \tag{3}$$

where $\boldsymbol{A}_k^{(i)}$ is a system matrix, $\boldsymbol{B}_k^{(i)}$ is a control matrix, and $\boldsymbol{w}_k^{(i)}$ is Gaussian with mean $\mathbf{0}$ and covariance $\boldsymbol{Q}_k^{(i)}$. The symbol $^{(i)}$ indicates that the model $m^i$ is used. The observation model (2) is also supposed to be linear Gaussian:

$$\boldsymbol{z}_k = \boldsymbol{C}_k \boldsymbol{x}_k + \boldsymbol{v}_k \tag{4}$$

where $\boldsymbol{C}_k$ is the observation matrix, and $\boldsymbol{v}_k$ is Gaussian with mean $\mathbf{0}$ and covariance $\boldsymbol{R}_k$.

Figure 2 shows the framework of the IMM estimation method to estimate the target state. The motion behavior is described with a set of models $\{m^1, m^2, m^3, ...\}$. Having the observation $\boldsymbol{z}_k$ of the state $\boldsymbol{x}_k$, KF is applied for each model (3) with the linear Gaussian assumption. Each KF updates the target state of mean $\hat{\bar{\boldsymbol{x}}}_{k-1|k-1}^{(i)}$ and covariance $\hat{\boldsymbol{\Sigma}}_{k-1|k-1}^{(i)}$ to mean $\bar{\boldsymbol{x}}_{k|k}^{(i)}$ and covariance $\boldsymbol{\Sigma}_{k|k}^{(i)}$. Each $\hat{\bar{\boldsymbol{x}}}_{k-1|k-1}^{(i)}$ and $\hat{\boldsymbol{\Sigma}}_{k-1|k-1}^{(i)}$ are derived from the IMM filter re-initialization incorporating all of $\bar{\boldsymbol{x}}_{k-1|k-1}^{(i)}$ and $\boldsymbol{\Sigma}_{k-1|k-1}^{(i)}$. The output $\bar{\boldsymbol{x}}_{k|k}$ and covariance $\boldsymbol{\Sigma}_{k|k}$ are calculated by the estimate fusion of all of $\bar{\boldsymbol{x}}_{k|k}^{(i)}$ and $\boldsymbol{\Sigma}_{k|k}^{(i)}$.

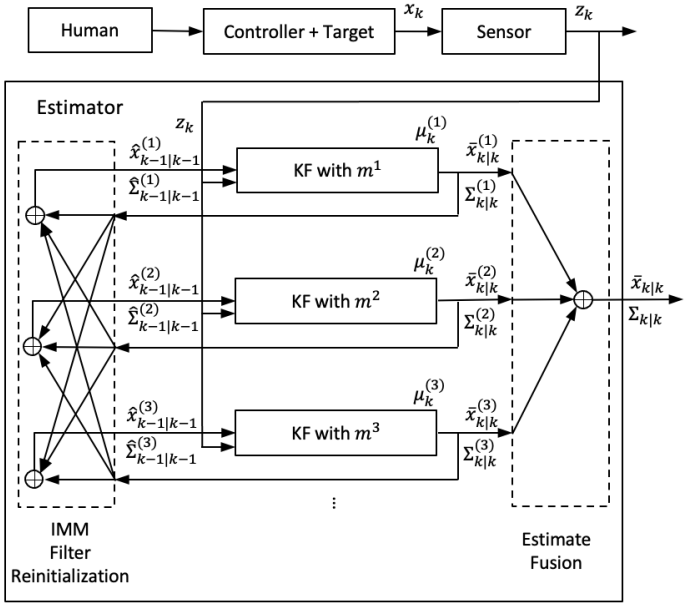



**Figure 2.** The IMM estimation method.

The mathematical derivation is as follows. The event that the model $m^i$ matches the mode at step $k$ is denoted as $m_k^{(i)} \triangleq \{s_k = m^i\}$. The probability of $m_k^{(i)}$ is denoted as $\mu_k^{(i)} \triangleq Pr\{m_k^{(i)}\}$ where $Pr\{\cdot\}$ indicates the probability of an event. The IMM estimator assumes the probability of transitioning from a model $m^i$ at step $k$ to a model $m^j$ at step $k+1$ is constant and known as $\pi_{ij}$:

$$Pr\{m_{k+1}^{(j)} | m_k^{(i)}\} = Pr\{s_{k+1} = m^j | s_k = m^i\} = \pi_{ij}, \tag{5}$$

where $i, j \in \mathbb{N}$. For one cycle, the predicted model probability, $\hat{\mu}_{k|k-1}^{(i)}$, is given by

$$\hat{\mu}_{k|k-1}^{(i)} \triangleq Pr\{m_k^{(i)} | \boldsymbol{z}_{1:k-1}\} = \sum_j \mu_{k-1}^{(j)} \pi_{ji} \tag{6}$$

where $\boldsymbol{z}_{1:k-1}$ are the observations from step 1 to step $k-1$. The weight that $m_{k-1}^{(j)}$ contributes to $m_k^{(i)}$ is derived as

$$\mu_{k-1}^{j|i} \triangleq Pr\{m_{k-1}^{(j)}|m_k^{(i)}, \boldsymbol{z}_{1:k-1}\} = \mu_{k-1}^{(j)}\pi_{ji}/\hat{\mu}_{k|k-1}^{(i)} \tag{7}$$

The KF of each model starts with the derivation of input:

$$\hat{\bar{\boldsymbol{x}}}_{k-1|k-1}^{(i)} \triangleq E[\boldsymbol{x}_{k-1}|m_k^{(i)}, \boldsymbol{z}_{1:k-1}] = \sum_j \bar{\boldsymbol{x}}_{k-1|k-1}^{(j)}\mu_{k-1}^{j|i}, \tag{8a}$$

$$\hat{\boldsymbol{\Sigma}}_{k-1|k-1}^{(i)} = \sum_j \Big[\boldsymbol{\Sigma}_{k-1|k-1}^{(j)} + (\bar{\boldsymbol{x}}_{k-1|k-1}^{(i)} - \hat{\boldsymbol{x}}_{k-1|k-1}^{(j)})(\bar{\boldsymbol{x}}_{k-1|k-1}^{(i)} - \hat{\boldsymbol{x}}_{k-1|k-1}^{(j)})^\top\Big]\mu_{k-1}^{(j|i)}. \tag{8b}$$

According to the KF formulation, the predicted mean and covariance are derived as

$$\bar{\boldsymbol{x}}_{k|k-1}^{(i)} = \boldsymbol{A}_k^{(i)}\hat{\boldsymbol{x}}_{k-1|k-1}^{(i)} + \boldsymbol{B}_k^{(i)}\boldsymbol{u}_k^{(i)}, \tag{9a}$$

$$\boldsymbol{\Sigma}_{k|k-1}^{(i)} = \boldsymbol{A}_k^{(i)}\hat{\boldsymbol{\Sigma}}_{k-1|k-1}^{(i)}(\boldsymbol{A}_k^{(i)})^\top + \boldsymbol{Q}_k^{(i)}. \tag{9b}$$

For correction, the KF gain is first computed through

$$\boldsymbol{K}_k^{(i)} = \boldsymbol{\Sigma}_{k|k-1}^{(i)}(\boldsymbol{C}_k)^\top(\boldsymbol{S}_k^{(i)})^{-1} \tag{10}$$

where the residual covariance is given by

$$\boldsymbol{S}_k^{(i)} = \boldsymbol{C}_k\boldsymbol{\Sigma}_{k|k-1}^{(i)}(\boldsymbol{C}_k)^\top + \boldsymbol{R}_k \tag{11}$$

The corrected mean and covariance are derived as

$$\bar{\boldsymbol{x}}_{k|k}^{(i)} = \bar{\boldsymbol{x}}_{k|k-1}^{(i)} + \boldsymbol{K}_k^{(i)}\tilde{\boldsymbol{z}}_k^{(i)}, \tag{12a}$$

$$\boldsymbol{\Sigma}_{k|k}^{(i)} = (\boldsymbol{I} - \boldsymbol{K}_k^{(i)}\boldsymbol{C}_k)\boldsymbol{\Sigma}_{k|k-1}^{(i)}, \tag{12b}$$

where the observation residual is given by

$$\tilde{\boldsymbol{z}}_k^{(i)} = z_k - \boldsymbol{C}_k\bar{\boldsymbol{x}}_{k|k-1}^{(i)} \tag{13}$$

For IMM estimation, the model likelihood $L_k^{(i)}$ is assumed as

$$\begin{aligned} L_k^{(i)} &\triangleq Pr\{\tilde{\boldsymbol{z}}_k^{(i)}|m_k^{(i)}, \boldsymbol{z}_{1:k-1}\} \\ &\overset{\text{assume}}{=} |2\pi\boldsymbol{S}_k^{(i)}|^{-\frac{1}{2}}\exp[-\frac{1}{2}(\tilde{\boldsymbol{z}}_k^{(i)})^\top(\boldsymbol{S}_k^{(i)})^{-1}\tilde{\boldsymbol{z}}_k^{(i)}] \end{aligned} \tag{14}$$

and the model probability $\mu_k^{(i)}$ is given by

$$\mu_k^{(i)} \triangleq Pr\{m_k^{(i)}|\boldsymbol{z}_{1:k}\} = \frac{\hat{\mu}_{k|k-1}^{(i)}L_k^{(i)}}{\sum_j \hat{\mu}_{k|k-1}^{(j)}L_k^{(j)}}. \tag{15}$$

The overall mean and covariance are derived as

$$\bar{\boldsymbol{x}}_{k|k} \triangleq E[\boldsymbol{x}_k|\boldsymbol{z}_{1:k}] = \sum_i \bar{\boldsymbol{x}}_{k|k}^{(i)}\mu_k^{(i)} \tag{16a}$$

$$\boldsymbol{\Sigma}_{k|k} = \sum_i [\boldsymbol{\Sigma}_{k|k}^{(i)} + (\boldsymbol{x}_{k|k}^{(i)} - \bar{\boldsymbol{x}}_{k|k})(\boldsymbol{x}_{k|k}^{(i)} - \bar{\boldsymbol{x}}_{k|k})^\top]\mu_k^{(i)}. \tag{16b}$$

Owing to the introducing of transition probability $Pr\{m_{k+1}^{(j)}|m_k^{(i)}\}$ of Equation (5) and the likelihood $L_k^{(i)}$ of Equation (14), the model probabilities $\mu_k^{(i)}$ adapt to match the current motion. Suppose the model $m^{(i)}$ matches the current mode better, the filter of $m^{(i)}$ contributes more on $\bar{\mathbf{x}}_{k|k}$ and $\mathbf{\Sigma}_{k|k}$ by having a higher model probability $\mu_k^{(i)}$.

For estimation with one motion model, the single motion model (1) cannot impair its inconsistency from the actual motion when the human has changed the target motion considerably. The IMM method estimates in a larger state space due to the usage of multiple models, but it still uses the most probable deterministic human control such as the CV and CA. If the control is different, the multiple models of the IMM method may not be able to cover the space of estimation and could lead to a wrong estimation. The uncertainty could also be underestimated since the unknown human control, which is most uncertain, is handled deterministically. This limitation of the conventional techniques affects the quality of estimation when the target is human-maneuvered.

## 3. Proposed Approach Using Intention-Pattern Model

As the contributions of this paper are the construction of a intention-pattern model and the state estimation using the intention-pattern model, this section describes each contribution in a subsection. Section 3.1 presents the overview of the construction of an intention-pattern model, followed by the details of the two major components, which are the intention inference and the intention-pattern modeling. The implementation of the constructed intention-pattern models into the state estimation is then detailed in Section 3.2.

### *3.1. Construction of Intention-Pattern Model*

#### 3.1.1. Overview

Figure 3 shows the construction of the intention-pattern model where an example illustration is given on the right side. The prior analysis of the target behavior leads to the extraction of a set of human intentions $H = \{\eta^{(i)}|\forall i\}$ and the corresponding approximate control terms, $\boldsymbol{U}^{(i)}$. Each human intention $\eta^{(i)}$ is an expression describing an aim or a plan, such as "moving forward" and "turning right".

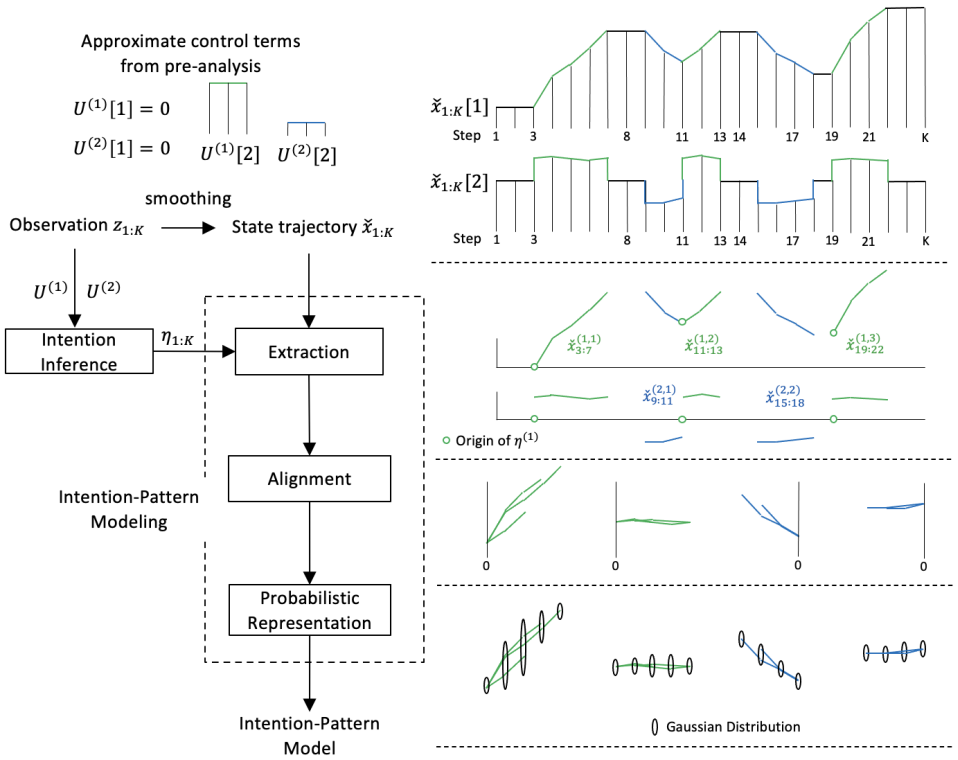



**Figure 3.** Construction of intention-pattern model. $(\cdot)[i]$ represents the $i$th dimension of $(\cdot)$.

Each human intention at step $k$ is defined as a function of the recent states of $N_h$ steps:

$$\eta_k = \alpha(\boldsymbol{x}_{k-N_h+1:k}). \tag{17}$$

Compared with the human actions which are of one step and are associated with the control input $\boldsymbol{u}_k$, the human intentions are labels of continuous states of multiple steps.

The corresponding control term could vary, but let it be constant for simplicity. Because the intention is defined for a period, the figure illustratively shows each control term with two steps. Given a sequence of observations $\boldsymbol{z}_{1:K}$, the human intention at step $k$, $\eta_k = \eta^{(i)} \in H$, is first inferred for all steps, i.e., $\eta_{1:K} = \{\eta_1, ..., \eta_K\}$. We assume that the observations are fully observable for simplicity. After smoothing the observations and deriving the state trajectory $\check{\boldsymbol{x}}_{1:K}$, the proposed construction technique identifies segments in state space exhibiting the extracted intention, $\check{\boldsymbol{x}}_{k_{s(i,j)}:k_{e(i,j)}}, j \in \mathbb{N}$, where $k_{s(i,j)}$ and $k_{e(i,j)}$ are the starting and the ending steps. Three segments of intention $\eta^{(1)}$ and two segments of intention $\eta^{(2)}$ are shown in the example illustration. The segments of the same intention are aligned to characterize the pattern of motion probabilistically. The intention-pattern model, describing the relationship between the input intention and the output motion pattern, is finally represented by a set of Gaussian distributions.

#### 3.1.2. Intention Inference

As states $\boldsymbol{x}_{k-N_h+1:k}$ are not directly available, this section proposes the approach to infer the human intention $\eta_k$ based on observations. Given observations $\boldsymbol{z}_{1:k}$ in addition to the control terms corresponding to the extracted intentions, $\boldsymbol{U}^{(i)}, \forall i$, the first process of the intention inference is to run the IMM estimation. There is only one motion model, but the motion is simulated for each intention $\boldsymbol{U}^{(i)}$:

$$\boldsymbol{x}_k = \boldsymbol{A}_k \boldsymbol{x}_{k-1} + \boldsymbol{U}^{(i)} + \boldsymbol{w}_k, \tag{18}$$

where $\boldsymbol{A}_k$ and $\boldsymbol{w}_k$ are determined from the analysis of target motion. Having Equation (4) as an observation model, the KF updates the mean $\boldsymbol{x}_{k|k}$ and the covariance $\boldsymbol{\Sigma}_{k|k}$ similarly to Equations (9)–(13) for each intention:

$$\bar{\boldsymbol{x}}^{(i)}_{k|k-1} = \boldsymbol{A}_k \bar{\boldsymbol{x}}^{(i)}_{k-1|k-1} + \boldsymbol{U}^{(i)} \tag{19a}$$

$$\boldsymbol{\Sigma}^{(i)}_{k|k-1} = \boldsymbol{A}_k \boldsymbol{\Sigma}^{(i)}_{k-1|k-1} (\boldsymbol{A}_k)^\top + \boldsymbol{Q}_k. \tag{19b}$$

$$\boldsymbol{K}^{(i)}_k = \boldsymbol{\Sigma}^{(i)}_{k|k-1} (\boldsymbol{C}_k)^\top (\boldsymbol{S}^{(i)}_k)^{-1} \tag{19c}$$

$$\boldsymbol{S}^{(i)}_k = \boldsymbol{C}_k \boldsymbol{\Sigma}^{(i)}_{k|k-1} (\boldsymbol{C}_k)^\top + \boldsymbol{R}_k \tag{19d}$$

$$\tilde{\boldsymbol{z}}^{(i)}_k = z_k - \boldsymbol{C}_k \bar{\boldsymbol{x}}^{(i)}_{k|k-1}, \tag{19e}$$

$$\bar{\boldsymbol{x}}^{(i)}_{k|k} = \bar{\boldsymbol{x}}^{(i)}_{k|k-1} + \boldsymbol{K}^{(i)}_k \tilde{\boldsymbol{z}}^{(i)}_k, \tag{19f}$$

$$\boldsymbol{\Sigma}^{(i)}_{k|k} = (\boldsymbol{I} - \boldsymbol{K}^{(i)}_k \boldsymbol{C}_k) \boldsymbol{\Sigma}^{(i)}_{k|k-1}. \tag{19g}$$

The likelihood of the motion model at step $k$ is given by Equation (14). As the intention is determined for a period, let the number of steps that defines an intention be $N_h$ steps. The intention likelihood is defined and derived as the joint likelihood of the model likelihoods $L^{(i)}_k$:

$$\mathcal{L}_k(\boldsymbol{U}^{(i)}) = \prod_{\kappa=k-N_h+1}^{k} L^{(i)}_k(\boldsymbol{U}^{(i)}) = \prod_{\kappa=k-N_h+1}^{k} |2\pi \boldsymbol{S}^{(i)}_\kappa|^{-\frac{1}{2}} \exp[-\frac{1}{2} (\tilde{\boldsymbol{z}}^{(i)}_\kappa)^\top (\boldsymbol{S}^{(i)}_\kappa)^{-1} \tilde{\boldsymbol{z}}^{(i)}_\kappa]. \tag{20}$$

The control term that maximizes the intention likelihood is then selected:

$$i_k = \arg\max_i \{\mathcal{L}_k(\boldsymbol{U}^{(i)}) | \forall i\}, \tag{21}$$

if the intention likelihood is above the threshold

$$\mathcal{L}_k(\boldsymbol{U}^{(i_k)}) \geq \mathcal{L}^*(\boldsymbol{U}^{(i_k)}).$$

The corresponding intention $\eta_k$ is given by

$$\eta_k = \begin{cases} \eta^{(i_k)} & \mathcal{L}_k(\boldsymbol{U}^{(i_k)}) \geq \mathcal{L}^*(\boldsymbol{U}^{(i_k)}) \\ \text{ø} & \text{Otherwise} \end{cases} \tag{22}$$

where ø indicating an empty element means that there is no matching intention. The recursive operation infers intention for all steps, $\eta_{1:K}$.

#### 3.1.3. Intention-Pattern Modeling

The first process of the intention-pattern modeling, the extraction of the intended motions, checks the intention $\eta_k$ and identifies its period. Let the $j$th segment of the $i$th intention extracted from the smoothed state trajectory $\check{\boldsymbol{x}}_{1:K}$ be $\check{\boldsymbol{x}}_{k_s(i,j):k_e(i,j)}$. The second process of alignment aligns the extracted segments by co-locating their origins $\check{\boldsymbol{x}}_{k_0(i,j)}$:

$$\begin{aligned} k_0(i,j) &= 0, \\ \check{\boldsymbol{x}}_{k_0(i,j)} &= \check{\boldsymbol{x}}_{k_0(i)} = \text{const.}, \end{aligned}$$

where the step of the origin $k_0(i,j) \in \{k_s(i,j), ..., k_e(i,j)\}$.

The final process of motion pattern characterization derives the intention-pattern model by probabilistically characterizing the aligned segments. Figure 4 shows the characterization after three segments—green, blue, and purple—are aligned. As the number of segments increases, it is valid to assume that the variation of the motion follows a Gaussian distribution:

$$\check{\boldsymbol{x}}_\kappa^{(i)} \sim \mathcal{N}\left(\bar{\check{\boldsymbol{x}}}_\kappa^{(i)}, \check{\boldsymbol{\Sigma}}_\kappa^{(i)}\right),$$

where $\kappa$ is a step of the intention-pattern model after alignment, and the mean and the covariance are

$$\begin{aligned} \bar{\check{\boldsymbol{x}}}_\kappa^{(i)} &= \frac{1}{n_\kappa^{(i)}} \sum_j \check{\boldsymbol{x}}_\kappa^{(i,j)}, \\ \check{\boldsymbol{\Sigma}}_\kappa^{(i)} &= \frac{1}{n_\kappa^{(i)}} \sum_j (\check{\boldsymbol{x}}_\kappa^{(i,j)} - \bar{\check{\boldsymbol{x}}}_\kappa^{(i)})(\check{\boldsymbol{x}}_\kappa^{(i,j)} - \bar{\check{\boldsymbol{x}}}_\kappa^{(i)})^\top. \end{aligned}$$

$n_\kappa^{(i)}$ is the number of segments at step $\kappa$ for the $i$th intention [30]. The intention-pattern model is finally derived as

$$\check{\boldsymbol{x}}_k^{(i)} \sim \mathcal{N}\left(\bar{\check{\boldsymbol{x}}}_\kappa^{(i)}, \check{\boldsymbol{\Sigma}}_\kappa^{(i)}\right)\delta(k-\kappa), \tag{23}$$

or

$$\begin{aligned} \bar{\check{\boldsymbol{x}}}_k^{(i)} &= \bar{\check{\boldsymbol{x}}}_\kappa^{(i)}\delta(k-\kappa), \\ \check{\boldsymbol{\Sigma}}_k^{(i)} &= \check{\boldsymbol{\Sigma}}_\kappa^{(i)}\delta(k-\kappa), \end{aligned}$$

where $\delta(\cdot)$ is a Dirac delta function. This means that the intention-pattern model is defined by a set of Gaussian distributions:

$$\mathcal{N}^{(i)} = \{\mathcal{N}(\bar{\check{\boldsymbol{x}}}_{\kappa}^{(i)}, \check{\boldsymbol{\Sigma}}_{\kappa}^{(i)}) | \forall \kappa\}. \tag{24}$$

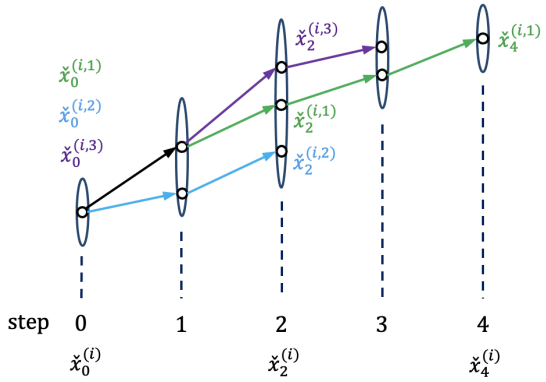



**Figure 4.** A set of Gaussian distributions representing the motion pattern which is the output of intention-pattern model.

### *3.2. Estimation Using Intention-Pattern Model*

Figure 5 shows the schematics of the proposed state estimator using the intention-pattern model. Given a new observation $\boldsymbol{z}_k$, Equations (19)–(22) output the intention of the current step, $\eta_k = \eta^{(i_k)}$. The estimator then checks the corresponding Gaussian distributions $\mathcal{N}^{(i_k)}$ to find the matching step $\kappa_k$ with respect to the recent estimate state $\boldsymbol{x}_{k-1|k-1}$:

$$\mathcal{L}(\boldsymbol{x}_{k-1|k-1}|\mathcal{N}^{(i_k)}) = \max_{\kappa} Pr\{\boldsymbol{x}_{k-1|k-1}|\mathcal{N}(\bar{\check{\boldsymbol{x}}}_{\kappa}^{(i_k)}, \check{\boldsymbol{\Sigma}}_{\kappa}^{(i_k)})\} \tag{25}$$

$$\kappa_k = \arg\max_{\kappa} Pr\{\boldsymbol{x}_{k-1|k-1}|\mathcal{N}(\bar{\check{\boldsymbol{x}}}_{\kappa}^{(i_k)}, \check{\boldsymbol{\Sigma}}_{\kappa}^{(i_k)})\}. \tag{26}$$

The step $\kappa_k$ of $\mathcal{N}^{(i_k)}$ is chosen if the intention-pattern model of the $i_k$th intention is satisfactory:

$$\mathcal{L}(\boldsymbol{x}_{k-1|k-1}|\mathcal{N}^{(i_k)}) \geq \mathcal{L}^*(\mathcal{N}^{(i_k)}), \tag{27}$$

and the step $\kappa_k$ matches the current step $k$.

Having the step $\kappa_k$ of the intention $i_k$ identified, the state is predicted by

$$\bar{\boldsymbol{x}}_{k|k-1} = \boldsymbol{A}_k \bar{\boldsymbol{x}}_{k-1|k-1} + \bar{\boldsymbol{U}}_k^{(i_k)}, \tag{28a}$$

$$\boldsymbol{\Sigma}_{k|k-1} = \boldsymbol{A}_k \boldsymbol{\Sigma}_{k-1|k-1} (\boldsymbol{A}_k)^\top + \boldsymbol{T}_k^{(i_k)} + \boldsymbol{Q}_k. \tag{28b}$$

where the control term $\boldsymbol{U}_k^{(i_k)}$ and the covariance of the control uncertainty, $\boldsymbol{T}_k^{(i_k)}$, are derived from the intention-pattern model as

$$\bar{\boldsymbol{U}}_k^{(i_k)} = \bar{\check{\boldsymbol{x}}}_{\kappa_k+1}^{(i_k)} - \boldsymbol{A}_k \bar{\check{\boldsymbol{x}}}_{\kappa_k}^{(i_k)}, \tag{29a}$$

$$\begin{aligned} \boldsymbol{T}_k^{(i_k)} &= \mathbf{E}[(\boldsymbol{U}_k^{(i_k)} - \bar{\boldsymbol{U}}_k^{(i_k)})(\boldsymbol{U}_k^{(i_k)} - \bar{\boldsymbol{U}}_k^{(i_k)})^\top] \\ &= \mathbf{E}\{[\check{\boldsymbol{x}}_{\kappa_k+1}^{(i_k)} - \bar{\check{\boldsymbol{x}}}_{\kappa_k+1}^{(i_k)} - \boldsymbol{A}_k(\check{\boldsymbol{x}}_{\kappa_k}^{(i_k)} - \bar{\check{\boldsymbol{x}}}_{\kappa_k}^{(i_k)})][\check{\boldsymbol{x}}_{\kappa_k+1}^{(i_k)} - \bar{\check{\boldsymbol{x}}}_{\kappa_k+1}^{(i_k)} - \boldsymbol{A}_k(\check{\boldsymbol{x}}_{\kappa_k}^{(i_k)} - \bar{\check{\boldsymbol{x}}}_{\kappa_k}^{(i_k)})]^\top\} \\ &= \check{\boldsymbol{\Sigma}}_{\kappa_k+1}^{(i_k)} + \boldsymbol{A}_k \check{\boldsymbol{\Sigma}}_{\kappa_k}^{(i_k)} \boldsymbol{A}_k^\top, \end{aligned} \tag{29b}$$

Equation (28) shows that the covariance propagation is more than that of the conventional KF-based estimation by the addition of $\boldsymbol{T}_k^{(i_k)}$. The correction is conducted by Equation (19c)–(19g) of KF with $i_k$ on behalf of $i$.

Because the proposed approach estimates the state incorporating the human intention, the mean of the estimated state is potentially more accurate than the conventional state estimation if observations are not available or reliable. The covariance of the estimated state is also more precise as it is updated adding the uncertainty of the human intention and prevents underestimation. Finally, note that the proposed state estimation allows future prediction with human intention in addition to the current estimation by recursively predicting the state with Equation (28).

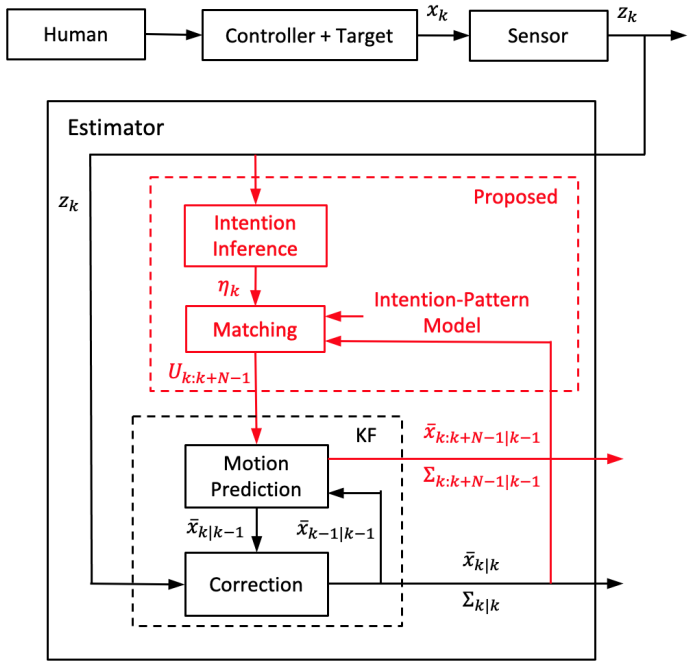



**Figure 5.** Estimation taking advantage of the proposed intention-pattern model. The red indicates the proposed parts compared with the conventional KF.

## 4. Numerical Validation

Having the strength of the intention-pattern model identified, it is essential to test the proposed approach numerically and identify the capability and limitations. The approach was evaluated by applying to the state estimation of a human-maneuvered multirotor, which is one of the applications of this class with high demand. To identify the capability and limitations in depth, a simulated environment was created and used.

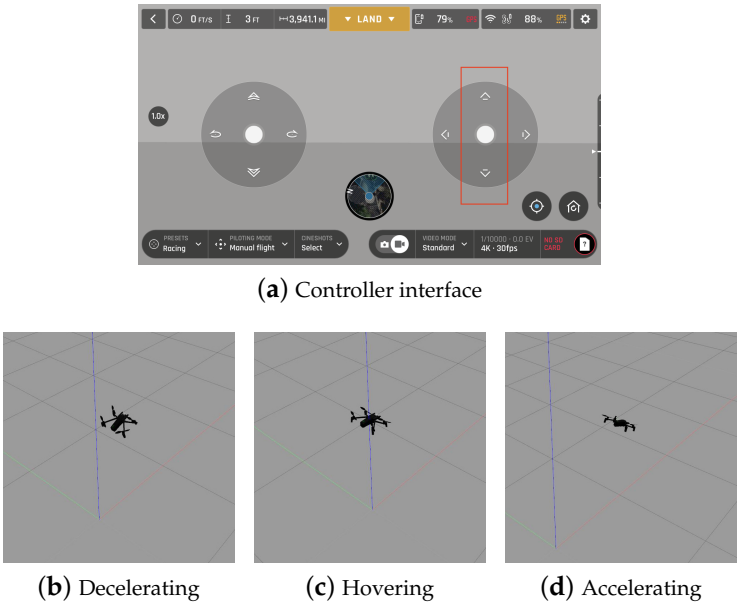
Figure 6 shows the controller interface used to create the multirotor motion and the resulting hovering, accelerating, and decelerating motions in the software-in-the-loop (SITL) simulation environment, whereas Table 1 lists the parameters used for simulation. With the right joystick of the controller interface, the human issues void command for hovering and forward or backward command for accelerating or decelerating.

The multirotor dynamics was calculated in Gazebo, which also created motion noise artificially. As the most fundamental and typical motion, the linear horizontal motion of the multirotor was considered. The multirotor's state, $\boldsymbol{x}$, is given by

$$\boldsymbol{x} = [p, \dot{p}, \theta, \dot{\theta}]^\top$$

where $p$ is the position in the moving direction, $\dot{p}$ is the linear velocity, $\theta$ is the attitude (pitch angle), and $\dot{\theta}$ is the angular velocity. The estimator was assumed to observe all the state variables of the multirotor, i.e., $\boldsymbol{z} = [z^p, z^{\dot{p}}, z^\theta, z^{\dot{\theta}}]^\top$. The observations were created by

adding noise to the true state where the noise variances are indicated as $[o_{s1}^2, o_{s1}^2, o_{s2}^2, o_{s2}^2]$ as the variances are varied in the parametric study. Figure 7 shows the time-varying human command, true state and observation. The observation was created with $[o_{s1}, o_{s2}] = [1, 0.05]$. The observation noise was set high as the proposed approach is effective when the observation is uncertain or unavailable. The first 100 s was used to construct the intention-pattern model, and the state estimation using the proposed approach was conducted with the observation of the remaining 60 s. The command varies dynamically, and the multirotor motion is seen to reflect the commands of forward, void and backward.



(**a**) Controller interface

(**b**) Decelerating (**c**) Hovering (**d**) Accelerating

**Figure 6.** The controller interface of the SITL simulation environment and the motion examples of the multirotor for three intentions.

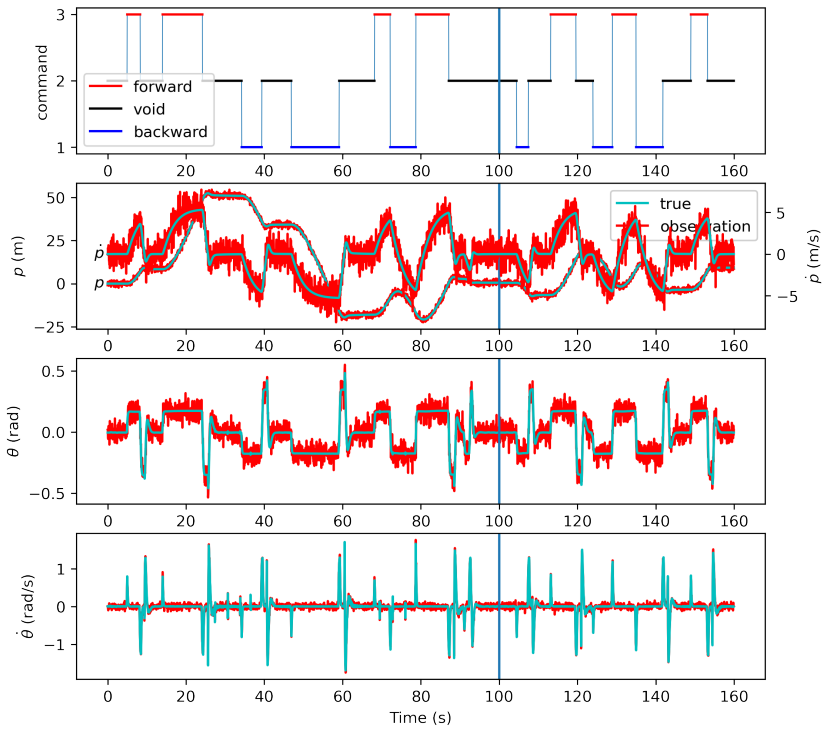



**Figure 7.** The human command and the true state and observation trajectories.

**Table 1.** Parameters for simulation.

| Parameter | Value |
|---|---|
| Human commands | Void, Forward, Backward |
| Simulation motion noise | Specified in Gazebo |
| Simulation observation noise variances | $[o_{s1}^2, o_{s1}^2, o_{s2}^2, o_{s2}^2]$ |
| Cruising inclination limit [rad] | 0.17 |

Through the analysis of the multirotor state estimation problem, the motion and observation models used by the proposed approach were linear. The motion matrix $\boldsymbol{A}_k$ is given by

$$\boldsymbol{A}_k = \begin{bmatrix} 1 & dt & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & dt \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{30}$$

whereas the observation matrix $\boldsymbol{C}_k$ is a four-dimensional identity matrix. Table 2 lists the parameters of the proposed approach for both the intention-pattern model construction and the state estimation. The number of prediction steps between observations is denoted as $n_p$ as it takes a different value for each process/study. While the variances of the observation noise is known, those of the motion noise were determined from the theoretical and experimental analyses. $\boldsymbol{U}^{(1)}$, $\boldsymbol{U}^{(2)}$, and $\boldsymbol{U}^{(3)}$ were chosen to infer the decelerating intention $\eta^{(1)}$, the hovering intention $\eta^{(2)}$, and the accelerating intention $\eta^{(3)}$, respectively. $\theta_U$ is a parameter to control the value of $\boldsymbol{U}^{(i)}$ for parametric study.

Section 4.1 investigates the validity of the construction process of intention-pattern model through the parametric study. Section 4.2 then validates the estimation performance using the intention-pattern model.

**Table 2.** Parameters of the proposed approach.

| Parameter | Value |
|---|---|
| Time step [s] | $\Delta t = 0.05$ |
| Number of predictions between two observations | $n_p$ |
| Variances of motion noise $\boldsymbol{Q}$ | $[0, (1 - o_{s1}/4)^2 \Delta t^2, 0, 2^2 \Delta t^2]$ |
| Variances of observation noise $\boldsymbol{R}$ | $[o_{s1}^2, o_{s1}^2, o_{s2}^2, o_{s2}^2]$ |
| Human intention | Decelerating, Hovering, Accelerating |
| $\boldsymbol{U}^{(1)}$ | $[0, 0, -\bar{x}_{k-1\|k-1}^{(i)}[3] - \theta_U, 0]^\top$ |
| $\boldsymbol{U}^{(2)}$ | $[0, 0, -\bar{x}_{k-1\|k-1}^{(i)}[3], 0]^\top$, |
| $\boldsymbol{U}^{(3)}$ | $[0, 0, -\bar{x}_{k-1\|k-1}^{(i)}[3] + \theta_U, 0]^\top$ |
| Smoothing technique | Triangular moving average |
| Duration of construction [s] | 100 |
| Duration of estimation [s] | 60 |

### *4.1. Construction of Intention-Pattern Model*

Figure 8 shows the inferred intentions and those in the corresponding smoothed trajectories when $\theta_U$ was 0.2. The smoothed trajectories are segmented based on the inferred intentions. The position is seen to appropriately increase and decrease when the human intention is with accelerating and decelerating respectively. As $\boldsymbol{U}^{(1)}$, $\boldsymbol{U}^{(2)}$, and $\boldsymbol{U}^{(3)}$ differ from each other in the pitch angle $\theta$, the pitch angle plot also shows intentions clearly: $\theta$ near 0 indicates hovering; positive $\theta$ with large magnitude indicates accelerating; negative $\theta$ with large magnitude indicates decelerating.

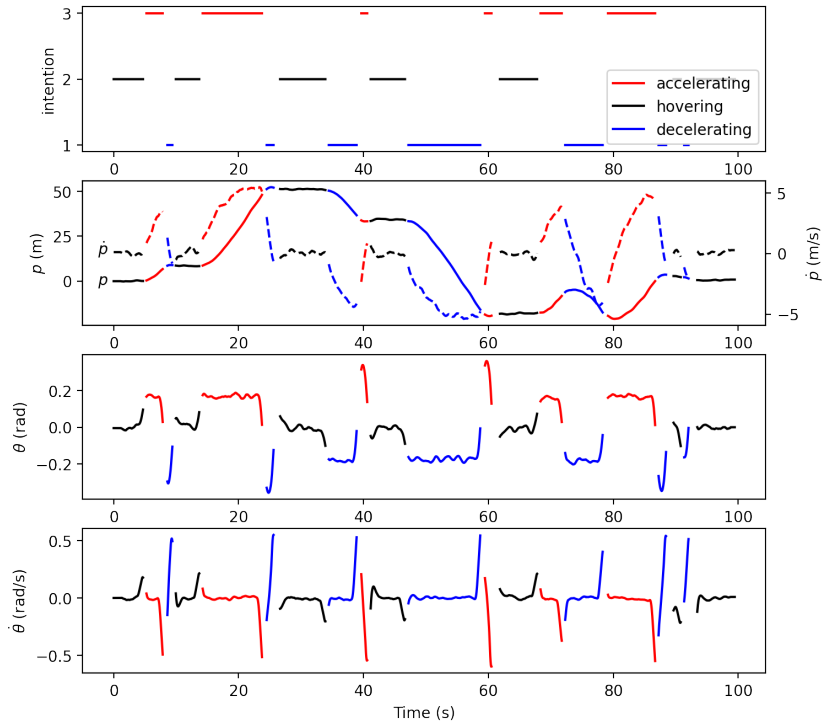


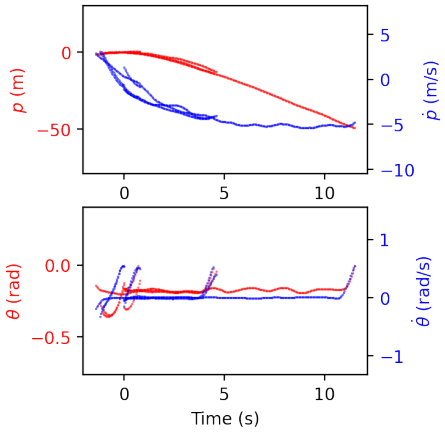

**Figure 8.** The inferred intention and the corresponding smoothed trajectory.

Figure 9 shows the aligned segments and the variances of each resulting intention-pattern model when $n_p = 1$. It is first seen that the aligned segments are consistent, which indicates that the proposed intention inference is valid. More consistency is shown in position than in pitch angle partly because the Gaussian assumption is not flexible enough to describe the pitch angle. The derived variances show that the intention-pattern models are modeled probabilistically from observations and could be used to perform state estimation more precisely.

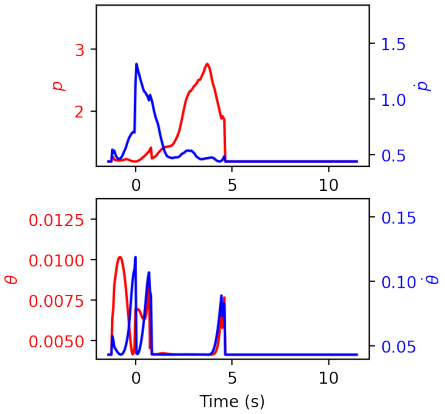
To analyze the dependency of the intention inference, the F1 score [31], evaluating the inference performance was derived with different levels of observation noises and control terms. The parameters varied were $o_{s2}$ for the observation noise and $\theta_U$ for the control term as the pitch angle $\theta$ characterizes the intention. The ground truth intention was defined based on the real $\theta$ value: hovering when $|\theta| \leq 0.05$; accelerating when $\theta > 0.05$; decelerating when $\theta < -0.05$. The F1 score is calculated as

$$\frac{2}{\frac{\mathrm{TP}+\mathrm{FP}}{\mathrm{TP}} + \frac{\mathrm{TP}+\mathrm{FN}}{\mathrm{TP}}}$$

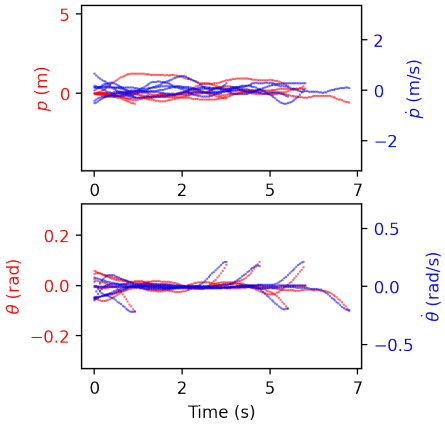
where TP, FP, and FN correspond to the number of steps of true positive, false positive, and false negative [31]. The F1 score which is closer to 1 indicates better inference. Figure 10 shows the distribution of the F1 score over $o_{s2}$ and $\theta_U$. As seen from the figure, the smaller the noise, the better the inference. For $\theta_U$, there is a best value in the middle; either too large or too small will result in poor inference.



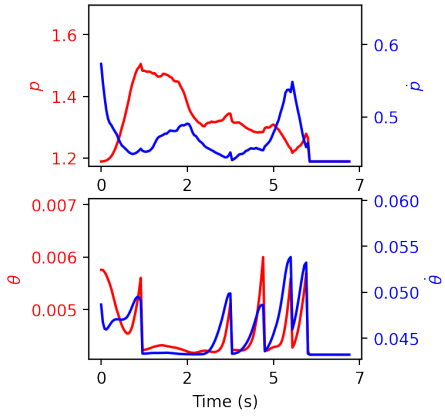

(**a**) Aligned segments (decelerating)



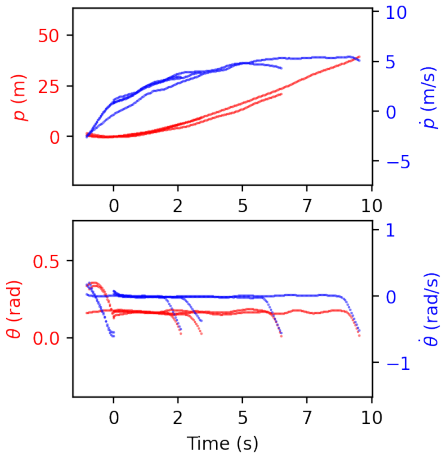

(**b**) Variances of Gaussians (decelerating)



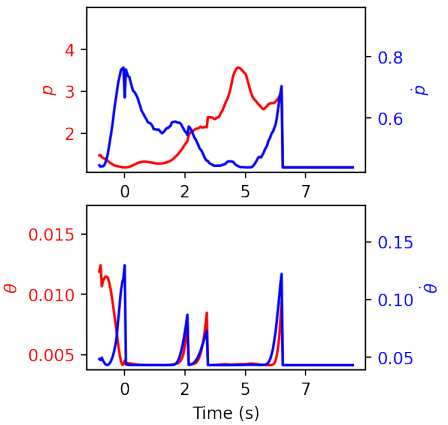

(**c**) Aligned segments (hovering)




(**d**) Variances of Gaussians (hovering)




(**e**) Aligned segments (accelerating)




(**f**) Variances of Gaussians (accelerating)

**Figure 9.** Construction of the intention-pattern model for three intentions.

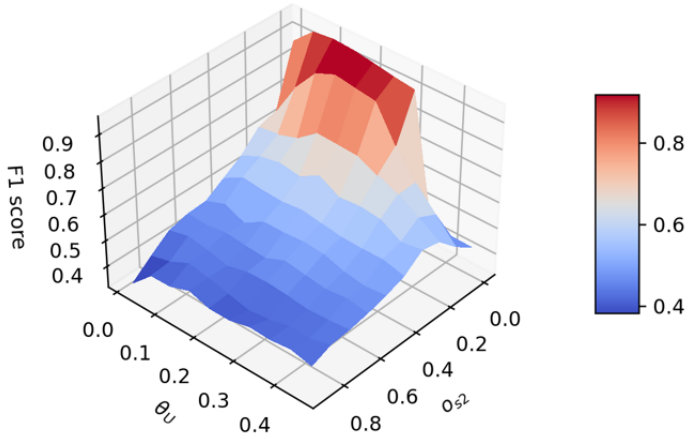
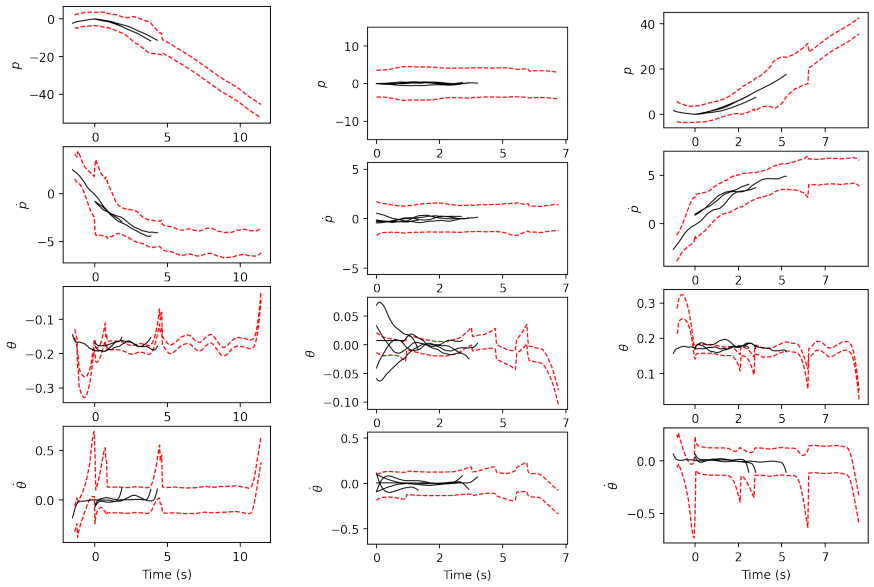



**Figure 10.** F1 score evaluating the intention inference accuracy with respect to simulation observation noise and the control term $\boldsymbol{U}^{(i)}$.

Figure 11 shows the resulting performance of each intention-pattern model. The two red broken lines show the range of motion pattern defined by the variance of the intention-pattern model constructed from the first 100 s whereas the solid black lines the motions of the same intention-pattern model identified in the next 60 s. It is seen that motions extracted in the 60 s are well along with the range of the intention-pattern model. This verifies the validity of the probabilistically represented intention-pattern model.




(**a**) Decelerating (**b**) Hovering (**c**) Accelerating

**Figure 11.** Validation of the constructed intention-pattern model.

### *4.2. Estimation Using Intention-Pattern Model*

Having the intention-pattern model constructed using the first 100 s, Figure 12 shows the result of state estimation incorporating the constructed intention-pattern model in the subsequent 60 s. Unlike the intention-pattern model, the state estimation uses $n_p = 5$ as the effect of the proposed approach can be seen with the motion prediction. The ground truth and the result of the conventional KF estimation without intention incorporation are also shown for comparison. The estimation result of the proposed approach is seen to be closer to the ground truth than that of the conventional approach. The estimation of $p$ and $\dot{p}$ particularly shows the responsive estimation of the proposed approach when the target motion is changed by the human while the conventional estimation exhibits notable delay.

The faster response is due to the use of the intention-pattern model. The conventional approach could improve estimation by frequent accurate observation, but observations are often uncertain or unavailable.

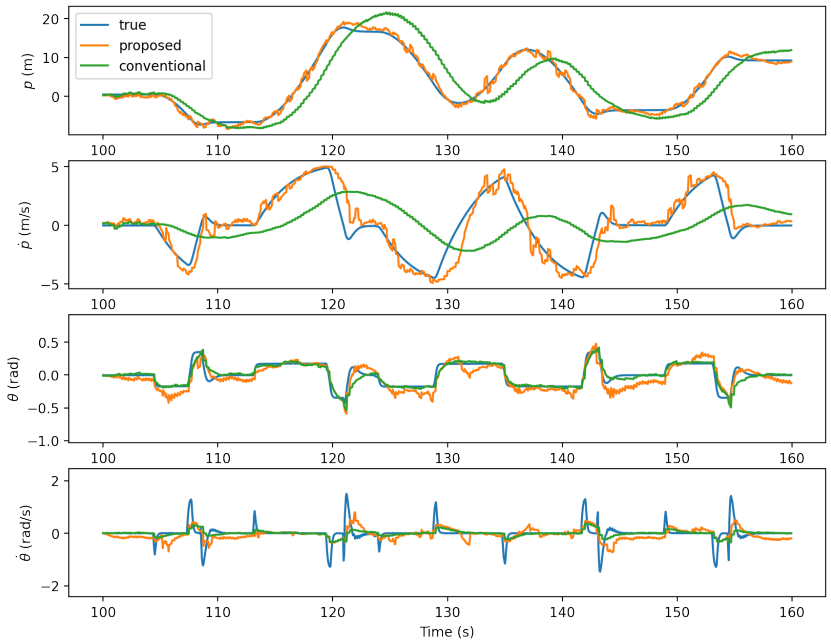



**Figure 12.** State estimation by the proposed and the conventional approaches.

Figure 13 shows the absolute error of estimated mean of each state variable with respect to time. While seeing less difference in $\theta$ and $\dot{\theta}$, the error of the proposed approach in $p$ and $\dot{p}$ consistently and significantly stays low compared to the conventional approach. The difference is particularly large when the human changes the target motion as the conventional approach does not take the human intention into account. The maximum error and the mean squared error (MSE), integrating the absolute errors to a single quantity, are improved by almost three times and 8.7 times, respectively, when the proposed approach was deployed.

Figure 14 shows the variance of each state variable estimated by the proposed and the conventional approaches. The result shows that the proposed approach exhibits larger variances than those of the conventional approach when the error is large. As the proposed approach infers human intentions and adds their uncertainties, its variance is estimated more precisely and adequately. The variance of the conventional approach, on the other hand, is significantly smaller though the mean estimation is wrong. Having the human control deterministically treated without inferring intentions, the uncertainty of the conventional approach is markedly underestimated.

The performance of the proposed approach in state estimation was further examined through the parametric study. Figure 15 shows the MSE of the proposed approach when $o_{s1}$ and $n_p$ were varied. $o_{s1}$ was varied to examine the effect of the observation noise as it contributed less at the construction of the intention-pattern model. The result of the conventional approach is also shown for comparison. It is first seen that the MSE of the proposed approach is significantly lower than that of the conventional approach when $o_{s1}$ is large. The large $o_{s1}$ increases the dependency of state estimation onto the prediction. As a result, the proposed approach, incorporating human intention and effective in prediction, can thus keep the MSE low. The result also shows that the MSE of the proposed approach remains low even when $n_p$ is large. $n_p$ also increases the dependency of state estimation onto the prediction, so the proposed approach becomes better than the conventional approach in accuracy. Meanwhile, the proposed and the conventional approaches exhibit

a similar MSE when $o_{s1}$ is low and $n_p$ is one. This is because the estimation becomes correction-driven as the frequency of the accurate correction becomes high.

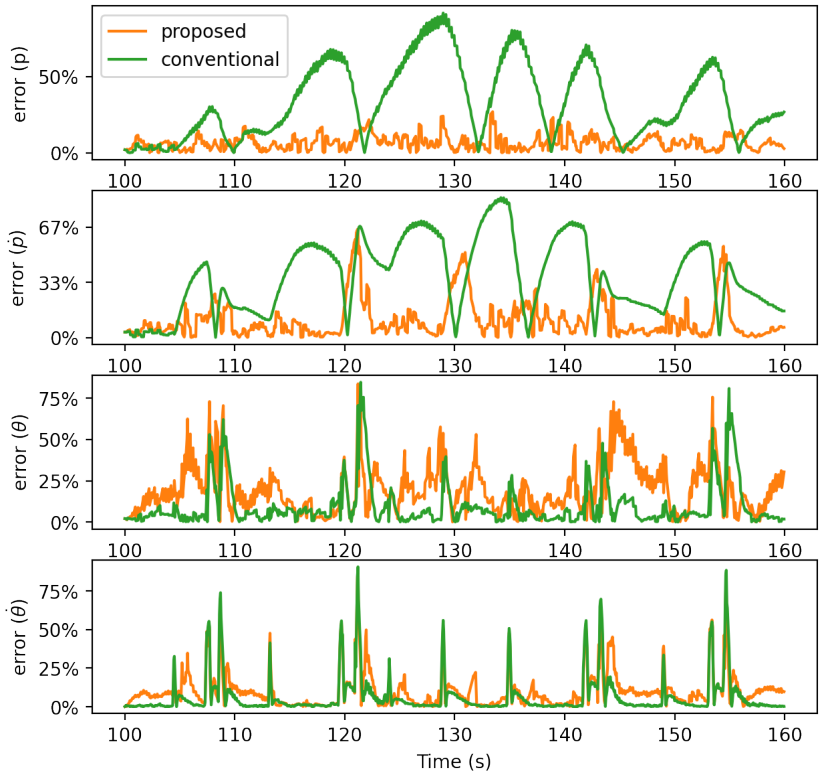



**Figure 13.** Absolute error of mean estimated by the proposed and the conventional approaches.

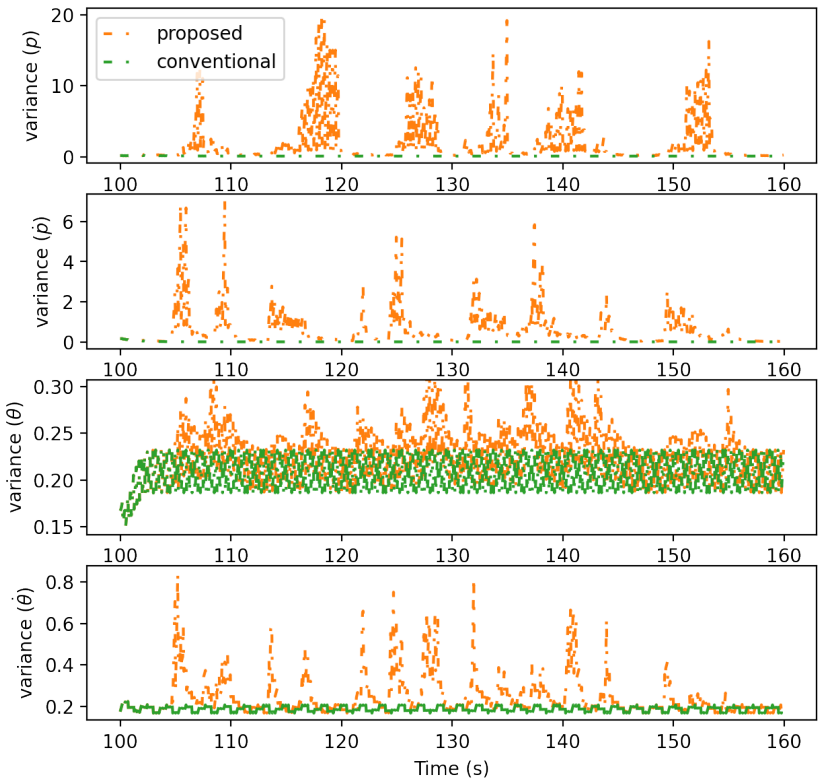
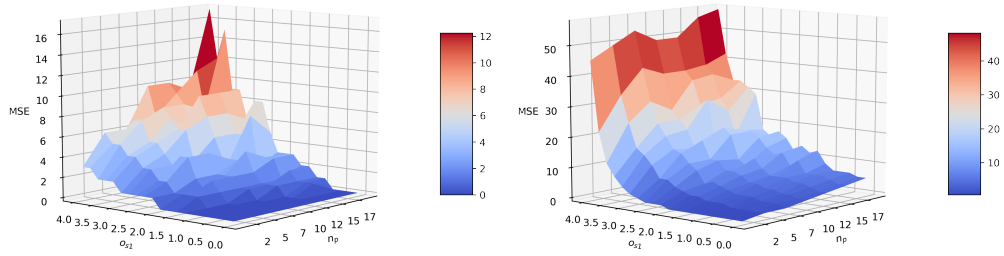



**Figure 14.** Variance estimated by the proposed and the conventional approaches.




(**a**) Proposed approach (**b**) Conventional approach

**Figure 15.** MSE with different observation noise and the number of prediction steps between observations.

## 5. Conclusions

This paper has presented an approach that estimates the state of a human-maneuvered target incorporating human intention, which consists of a preprocess constructing an intention-pattern model, and the main process allowing state estimation using the intention-pattern model. The preprocess constructs the intention-pattern model from the prior observations and probabilistically represents the model. The main process, then, uses standard state estimation such as KF extensively leveraging the probabilistically represented intention-pattern model. In the application of the proposed approach to the state estimation of a human-maneuvered multirotor, the numerical result has first shown that the constructed intention-pattern model represents the human intention appropriately. The result of state estimation of the human-maneuvered multirotor then shows that the proposed approach estimates the state more accurately than the conventional approach particularly when observations are uncertain or unavailable. The proposed approach has also demonstrated that it can estimate the covariance more precisely.

The paper has reported the first progress of the state estimation of a human-maneuvered target using human intention, and much future work is possible. Ongoing work includes the extension of the proposed approach for partially observable problems and model predictive control. Observations are necessary for the construction of the intention-pattern model, but the state may not be fully observable. The proposed approach is effective in prediction-driven estimation, so the model predictive control of an autonomous robot becomes one of the most effective extensions. The outcomes will be summarized and published in the form of papers as soon as they are ready.

**Author Contributions:** Conceptualization, Y.Q. and T.F.; methodology, Y.Q., M.K., and T.F.; software, Y.Q.; validation, Y.Q., M.K., and T.F.; formal analysis, Y.Q. and M.K.; investigation, Y.Q. and M.K.; resources, Y.Q., M.K. and T.F.; data curation, Y.Q.; writing—original draft preparation, Y.Q.; writing—review and editing, M.K. and T.F.; visualization, Y.Q.; supervision, M.K. and T.F.; project administration, T.F.; funding acquisition, T.F. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work is sponsored by US Office of Naval Research (N00014-20-1-2468).

**Acknowledgments:** The authors would like to express their sincere gratitude to Tom McKenna at Office of Naval Research for his invaluable support and advice.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Bar-Shalom, Y.; Li, X.R.; Kirubarajan, T. *Estimation with Applications to Tracking and Navigation: Theory Algorithms and Software*; John Wiley & Sons: Hoboken, NJ, USA, 2004.
2. Vrba, M.; Heřt, D.; Saska, M. Onboard marker-less detection and localization of non-cooperating drones for their safe interception by an autonomous aerial system. *IEEE Robot. Autom. Lett.* **2019**, *4*, 3402–3409. [CrossRef]

3. Pessanha Santos, N.; Lobo, V.; Bernardino, A. Two-stage 3D model-based UAV pose estimation: A comparison of methods for optimization. *J. Field Robot.* **2020**, *37*, 580–605. [CrossRef]
4. Jin, R.; Jiang, J.; Qi, Y.; Lin, D.; Song, T. Drone detection and pose estimation using relational graph networks. *Sensors* **2019**, *19*, 1479. [CrossRef]
5. Li, X.R.; Jilkov, V.P. Survey of maneuvering target tracking. Part I. Dynamic models. *IEEE Trans. Aerosp. Electron. Syst.* **2003**, *39*, 1333–1364.
6. Thrun, S.; Burgard, W.; Fox, D. *Probabilistic Robotics*; MIT Press: Cambridge, MA, USA, 2005.
7. Kawase, T.; Tsurunosono, H.; Ehara, N.; Sasasc, I. Two-stage kalman estimator using advanced circular prediction for maneuvering target tracking. In Proceedings of the 1998 IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP'98 (Cat. No. 98CH36181), Seattle, WA, USA, 12–15 May 1998; Volume 4, pp. 2453–2456.
8. Costa, P.J. Adaptive model architecture and extended Kalman-Bucy filters. *IEEE Trans. Aerosp. Electron. Syst.* **1994**, *30*, 525–533. [CrossRef]
9. Leven, W.F.; Lanterman, A.D. Unscented Kalman filters for multiple target tracking with symmetric measurement equations. *IEEE Trans. Autom. Control.* **2009**, *54*, 370–375. [CrossRef]
10. Liu, C.; Shui, P.; Li, S. Unscented extended Kalman filter for target tracking. *J. Syst. Eng. Electron.* **2011**, *22*, 188–192. [CrossRef]
11. Arulampalam, M.S.; Maskell, S.; Gordon, N.; Clapp, T. A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. *IEEE Trans. Signal Process.* **2002**, *50*, 174–188. [CrossRef]
12. Martínez-García, M.; Kalawsky, R.S.; Gordon, T.; Smith, T.; Meng, Q.; Flemisch, F. Communication and interaction with semiautonomous ground vehicles by force control steering. *IEEE Trans. Cybern.* **2020**, *51*, 3913–3924. [CrossRef] [PubMed]
13. Martínez-García, M.; Zhang, Y.; Gordon, T. Memory pattern identification for feedback tracking control in human–machine systems. *Hum. Factors* **2021**, *63*, 210–226. [CrossRef] [PubMed]
14. Steckenrider, J.J.; Furukawa, T. A Probabilistic Model-adaptive Approach for Tracking of Motion with Heightened Uncertainty. *Int. J. Control. Autom. Syst.* **2020**, *18*, 2687–2698. [CrossRef]
15. Gindele, T.; Brechtel, S.; Dillmann, R. A probabilistic model for estimating driver behaviors and vehicle trajectories in traffic environments. In Proceedings of the 13th International IEEE Conference on Intelligent Transportation Systems, Madeira Island, Portugal, 19–22 September 2010; pp. 1625–1631.
16. Mehra, R. On the identification of variances and adaptive Kalman filtering. *IEEE Trans. Autom. Control.* **1970**, *15*, 175–184. [CrossRef]
17. Almagbile, A.; Wang, J.; Ding, W. Evaluating the performances of adaptive Kalman filter methods in GPS/INS integration. *J. Glob. Position Syst.* **2010**, *9*, 33–40. [CrossRef]
18. Bogler, P.L. Tracking a maneuvering target using input estimation. *IEEE Trans. Aerosp. Electron. Syst.* **1987**, *3*, 298–310. [CrossRef]
19. Chakrabarty, A.; Corless, M.J.; Buzzard, G.T.; Żak, S.H.; Rundell, A.E. State and unknown input observers for nonlinear systems with bounded exogenous inputs. *IEEE Trans. Autom. Control.* **2017**, *62*, 5497–5510. [CrossRef]
20. Conte, D.; Furukawa, T. Autonomous Robotic Escort Incorporating Motion Prediction and Human Intention. In Proceedings of the 2021 IEEE International Conference on Robotic and Automation, Xi'an, China, 30 May–5 June 2021; 7p.
21. Blom, H.A.; Bar-Shalom, Y. The interacting multiple model algorithm for systems with Markovian switching coefficients. *IEEE Trans. Autom. Control.* **1988**, *33*, 780–783. [CrossRef]
22. Akca, A.; Efe, M.Ö. Multiple model Kalman and Particle filters and applications: A survey. *IFAC-PapersOnLine* **2019**, *52*, 73–78. [CrossRef]
23. Wang, X.; Xu, M.; Wang, H.; Wu, Y.; Shi, H. Combination of interacting multiple models with the particle filter for three-dimensional target tracking in underwater wireless sensor networks. *Math. Probl. Eng.* **2012**, *2012*, 829451. [CrossRef]
24. Ma, Y.; Zhao, S.; Huang, B. Multiple-model state estimation based on variational Bayesian inference. *IEEE Trans. Autom. Control.* **2018**, *64*, 1679–1685. [CrossRef]
25. Yu, M.; Oh, H.; Chen, W.H. An improved multiple model particle filtering approach for manoeuvring target tracking using airborne GMTI with geographic information. *Aerosp. Sci. Technol.* **2016**, *52*, 62–69. [CrossRef]
26. Li, X.R.; Jilkov, V.P. Survey of maneuvering target tracking. Part V. Multiple-model methods. *IEEE Trans. Aerosp. Electron. Syst.* **2005**, *41*, 1255–1321.
27. Li, X.R.; Bar-Shalom, Y. Multiple-model estimation with variable structure. *IEEE Trans. Autom. Control.* **1996**, *41*, 478–493.
28. Xu, L.; Li, X.R.; Duan, Z. Hybrid grid multiple-model estimation with application to maneuvering target tracking. *IEEE Trans. Aerosp. Electron. Syst.* **2016**, *52*, 122–136. [CrossRef]
29. Li, X.R.; Bar-Shalom, Y.; Blair, W.D. Engineer's guide to variable-structure multiple-model estimation for tracking. *Multitarg.-Multisens. Tracking Appl. Adv.* **2000**, *3*, 499–567.
30. Bishop, C.M. *Pattern Recognition and Machine Learning*; Springer: Berlin/Heidelberg, Germany, 2006.
31. Géron, A. *Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*; O'Reilly Media: Newton, MA, USA, 2019