Proceedings of the 36th Chinese Control Conference
July 26-28, 2017, Dalian, China

Information Fusion Analysis of Multi-UAV System Based on Information Geometry

ZHAO Yunyun[1], WANG Xiangke[1], KONG Weiwei[2], SHEN Lincheng[1]

1. College of Mechatronic Engineering and Automation, National University of Defense Technology, Changsha 410073, P. R. China
E-mail: xkwang@nudt.edu.cn

2. Naval Academy of Armament, Beijing 100071, P. R. China
E-mail: kww_kww@163.com

Abstract: We study the problem of optimally coordinating multiple fixed-wing UAVs to perform target tracking, which entails that the UAVs are tasked with gathering the best joint radar-based measurements of ground targets. In order to obtain the maximum amount of information, we first studied the information expression in multi-sensor network system. Based on the Minkowski determinant theorem, the change of information volume caused by data fusion is analyzed with the relevant content of information geometry, and take it as the foundation of communication and action decision-making. Simulation results show that the data fusion can increase the amount of information, and make the result more accurate. Further, for specific scene, the simulation results illustrate that multi-UAV cooperative observation can improve the accuracy of target tracking. This sets the foundation for the further study of the information exchange between multi-UAV.

Key Words: Multi-UAVs, Decision-making, Information Geometry, Fisher Information

1 Introduction

Collaborate multi-UAV systems have received increased attention in the past and have applications in traffic supervision, autonomous robot navigation, and surveillance of large facilities.One of particular interest is that of using small, fixed-wing UAVs to perform target tracking, which entails that one or more sensor-equipped UAVs is responsible for autonomously tracking moving ground targets. In the presence of uncertainty in a target motion and uncertainty in the sensory readings, the true target state is not available for decision-making purposes. The purpose of this work is analyze the change of information volume caused by data fusion with the relevant content of information geometry, and take it as the foundation of communication and action decision-making. Finally, we study an effective solution to the problem of optimally coordinating two UAVs to track some moving ground targets under fairly realistic conditions.

In principle, UAVs could synchronize their beliefs and coordinate their actions by communicating because how UAV updates their beliefs in response to sending and receiving a communication affects future decision making. Considering the limited communication, it must decide whether to communicate this new information to its teammates, and if so, to whom, and at what time. Anticipatory behaviors have been identified as a key characteristic in effective human teams. A number of previous works have aimed to model communication as a form of anticipatory behavior [1]. [3] introduced a formalized model for reasoning about the cost and benefit of supportive actions, including communications, during the collaboration. Agents in [3] computed the expected value of communicating versus not communicating using a domain specific probability recipe tree (PRT) and acted to minimize expected cost. [2] proposed a hybrid Belief-Desire-Intention (BDI) and decision-theoretic approach that integrated PRT with an MDP formulation to decide, not only whether or not to communicate but also the best future time to communicate assuming full observation of the teammates actions. Both approaches, however, were evaluated using an omniscient agent with full observability of the world and its teammate actions. In our work, we evaluate our model with fully decentralized teams operating under partial observability of the world and teammate actions. UAVs need to not only reason over uncertainties in their teammates plans, but also over other teammates states. To communicate effectively, multi-UAV system must trade off the benefit that can be achieved through communication with the cost of communicating. Therefore, it is necessary to use a suitable expression to measure the benefits of communication in order to compare with the cost of communication.

That is to say, we seek optimally coordinated behavior between multiple UAVs aimed at improving the estimate of the target state. The decentralized partially observable Markov decision process (Dec-POMDP) is a general framework for representing multi-agent coordination problems. Dec-POMDPs have been widely studied in artificial intelligence as a way to address the fundamental differences in decision-making in decentralized settings [4–6]. Like the POMDP [7] model that it extends, Dec-POMDPs consider general dynamics, cost and sensor models. Any problem where multiple robots share a single overall reward or cost function can be formalized as a Dec-POMDP. Unfortunately, this generality comes at a cost: Dec-POMDPs are typically infeasible to solve except for small problems [5, 8].

2 Analysis of Fusion Filtering Based on Information

Consider the following tracking problem. A dynamic target is tracked by multiple disparate sensors mounted on different UAV, with different measurement dynamics and noise characteristics. How do we evaluate the approaches to combine the multi-sensor measurements to obtain a state estimates? There are various multisensor data fusion approaches, of which Kalman filtering is one of the most significant. Methods for Kalman filter based data fusion, including state-vector fusion and measurement fusion, have

This work is supported by National Natural Science Foundation (NNSF) of China under Grant 61403406.

been widely studied. Measurement fusion methods directly fuse the sensor measurements to obtain a weighted or combined measurement and then use a single Kalman filter to obtain the final state estimate based upon the fused observation. It generally provides better overall estimation performance. There are some commonly used methods for measurement fusion. One flexible and efficient method simply merges the multisensor data, which increases the dimension of the observation vector of the Kalman filter. We will analyze the change of the amount of information in the fusion filtering method from the point of view of information geometry information changes.

2.1 The Information Fusion Model

Assume a dynamical target which is tracked by N sensors. The observation processes are modeled by the following discrete-time state-space model:

$$\boldsymbol{z}_k(t) = \boldsymbol{h}_k\big(\mathbf{x}(t)\big) + \boldsymbol{w}_k(t), \quad k = 1, 2, \ldots, N \tag{1}$$

where $\boldsymbol{h}_k\big(\mathbf{x}(t)\big)$ denotes the error-free true target position in the kth sensor polar coordinates, and $\boldsymbol{w}_k(t)$ is the respective random measurement errors. It is normally assumed that these measurement errors in the sensor CS are zero-mean with Gaussian distributed, and the covariance matrix is $\mathbf{C}_k\big(\mathbf{x}(t)\big)$, i.e.,

$$\boldsymbol{z}_k(t)\,|\mathbf{x}(t) \sim \mathcal{N}\Big(\boldsymbol{h}_k\big(\mathbf{x}(t)\big), \mathbf{C}_k\big(\mathbf{x}(t)\big)\Big). \tag{2}$$

The corresponding probability density function is formula (3).

It is assumed that the measurement noise is independent. By measurement fusion, the N sensor models can be integrated into the following single model:

$$\boldsymbol{z}(t) = \boldsymbol{h}\big(\mathbf{x}(t)\big) + \boldsymbol{w}(t) \tag{4}$$

where $\boldsymbol{w}(t) \sim \mathcal{N}\Big(0, \mathbf{C}\big(\mathbf{x}(t)\big)\Big)$.

The measurement fusion integrates the sensor measurement information by augmenting the observation vector as follows:

$$\boldsymbol{z}(t) = \begin{bmatrix} \boldsymbol{z}_1(t) & \boldsymbol{z}_2(t) & \cdots & \boldsymbol{z}_N(t) \end{bmatrix}^{\mathrm{T}} \tag{5}$$

$$\boldsymbol{h}(\mathbf{x}) = \begin{bmatrix} \boldsymbol{h}_1(\mathbf{x}) & \boldsymbol{h}_2(\mathbf{x}) & \cdots & \boldsymbol{h}_N(\mathbf{x}) \end{bmatrix}^{\mathrm{T}} \tag{6}$$

$$\mathbf{C}(\mathbf{x}) = diag\begin{bmatrix} \mathbf{C}_1(\mathbf{x}) & \mathbf{C}_2(\mathbf{x}) & \cdots & \mathbf{C}_N(\mathbf{x}) \end{bmatrix} \tag{7}$$

2.2 Changes in the Amount of Information

It is well known that the value of the determinant of the Fisher information matrix represents the volume of the amount of information [14] that can be acquired by the underlying sensor system.

The Fisher information matrix $[\mathbf{G}(\mathbf{x})] = \big[g_{ij}(\mathbf{x})\big]$ is given by the following formula (8)

$$g_{ij}(\mathbf{x}) = E\left[\frac{\partial \ln p(\boldsymbol{z}|\mathbf{x})}{\partial \mathrm{x}_i} \cdot \frac{\partial \ln p(\boldsymbol{z}|\mathbf{x})}{\partial \mathrm{x}_i}\right]. \tag{8}$$

As the probability density function is nonnegative, the Fisher information matrix is a positive semidefinite symmetric matrix.

By noting that the observation model is a Gaussian distribution, the Fisher information matrix can be written as typical element:

$$\begin{aligned}\big[\mathbf{G}(\mathbf{x})\big]_{ij} =& \left[\frac{\partial \boldsymbol{h}(\mathbf{x})}{\partial \mathrm{x}_i}\right]^{\mathrm{T}} \mathbf{C}^{-1}(\mathbf{x}) \left[\frac{\partial \boldsymbol{h}(\mathbf{x})}{\partial \mathrm{x}_j}\right] \\ &+ \frac{1}{2} tr\left(\mathbf{C}^{-1}(\mathbf{x}) \frac{\partial \mathbf{C}(\mathbf{x})}{\partial \mathrm{x}_i} \mathbf{C}^{-1}(\mathbf{x}) \frac{\partial \mathbf{C}(\mathbf{x})}{\partial \mathrm{x}_j}\right)\end{aligned} \tag{9}$$

The determinant of the matrix represents the amount of the information, which can be used as the basis for decision-making to obtain a more accurate target state estimation. In order to compare the change of amount of information caused by multi-sensor measurement, we introduce the Minkowski determinant theorem as following:

Minkowski determinant theorem. *If A and B are all n-dimensional nonnegative hermitian matrices, then*

$$\left(\det(\mathbf{A} + \mathbf{B})\right)^{1/n} \geq \left(\det(\mathbf{A})\right)^{1/n} + \left(\det(\mathbf{B})\right)^{1/n}, \tag{10}$$

where $\det(\cdot)$ *indicates the determinant of a matrix.*

Come back to see the Fisher information with multi-sensor measurement, we can obtain that:

$$\det\big(\mathbf{G}(\mathbf{x})\big) \geq \sum_{k=1}^{N} \det\big(\mathbf{G}_k(\mathbf{x})\big). \tag{11}$$

Naturally,

$$\det\big(\mathbf{G}(\mathbf{x})\big) \geq \det\big(\mathbf{G}_k(\mathbf{x})\big), k = 1, 2, \cdots, N \tag{12}$$

Proof. Using the measurement fusion methods, we get the Fisher information with multi-sensor measurement, shown in (13).

Due to the sum of positive semidefinite hermitian matrix is also positive semidefinite hermitian matrix, the **Minkowski determinant theorem** can be extended to multiple matrix.

As $\mathbf{G}_k(\mathbf{x})\,(k = 1, 2, \cdots, N)$ are Fisher information matrix, they are positive semidefinite symmetric matrices. It names that:

$$\det\big(\mathbf{G}_k(\mathbf{x})\big) \geq 0,$$

and then, they conform to the hypothesis of the **Minkowski determinant theorem**. We have:

$$\det\big(\mathbf{G}(\mathbf{x})\big) \geq \left(\sum_{k=1}^{N} \left(\det\left(\mathbf{G}_k(\mathbf{x})\right)^{1/n}\right)^n. \tag{14}$$

Ulteriorly,

$$\left(\sum_{k=1}^{N} \left(\det\left(\mathbf{G}_k(\mathbf{x})\right)^{1/n}\right)^n \geq \sum_{k=1}^{N} \det\left(\mathbf{G}_k(\mathbf{x})\right) \tag{15}$$

□

From the point of view of information geometry, the mutual communication between sensors, namely data fusion, is able to increase the amount of information, which is better

$$p\left(\boldsymbol{z}_k|\mathbf{x}\right)=\frac{1}{\sqrt{|2\pi\mathbf{C}_k(\mathbf{x})|}}\exp\left\{-\frac{1}{2}\big(\boldsymbol{z}_k-\boldsymbol{h}_k(\mathbf{x})\big)^{\mathrm{T}}\mathbf{C}^{-1}(\mathbf{x})\big(\boldsymbol{z}_k-\boldsymbol{h}_k(\mathbf{x})\big)\right\} \tag{3}$$

$$\begin{aligned}\big[\mathbf{G}(\mathbf{x})\big]_{ij}=&\left[\begin{array}{ccc}\left(\frac{\partial\boldsymbol{h}_1(\mathbf{x})}{\partial\mathrm{x}_i}\right)^{\mathrm{T}} & \cdots & \left(\frac{\partial\boldsymbol{h}_N(\mathbf{x})}{\partial\mathrm{x}_i}\right)^{\mathrm{T}}\end{array}\right]diag\left[\begin{array}{ccc}\mathbf{C}_1^{-1}(\mathbf{x}) & \cdots & \mathbf{C}_N^{-1}(\mathbf{x})\end{array}\right]\left[\begin{array}{ccc}\frac{\partial\boldsymbol{h}_1(\mathbf{x})}{\partial\mathrm{x}_j} & \cdots & \frac{\partial\boldsymbol{h}_N(\mathbf{x})}{\partial\mathrm{x}_j}\end{array}\right]^T\\&+\frac{1}{2}tr\left\{diag\left[\begin{array}{ccc}\mathbf{C}_1^{-1}(\mathbf{x})\frac{\partial\mathbf{C}_1(\mathbf{x})}{\partial\mathrm{x}_i}\mathbf{C}_1^{-1}(\mathbf{x})\frac{\partial\mathbf{C}_1(\mathbf{x})}{\partial\mathrm{x}_j} & \cdots & \mathbf{C}_N^{-1}(\mathbf{x})\frac{\partial\mathbf{C}_N(\mathbf{x})}{\partial\mathrm{x}_i}\mathbf{C}_N^{-1}(\mathbf{x})\frac{\partial\mathbf{C}_N(\mathbf{x})}{\partial\mathrm{x}_j}\end{array}\right]\right\}\\=&\left[\sum_{k=1}^{N}\mathbf{G}_k(\mathbf{x})\right]_{ij}\end{aligned} \tag{13}$$

for target state estimation. However, information communication is limited by bandwidth and time, therefore it is necessary to consider the communication efficiency and cost. Next, we apply the above conclusion to the specific scenario, and compare the effect of multi-UAV tracking target with or without information exchange.

3 Application in Target Tracking

Consider the following scenario. Two dynamic targets are tracked by a couple of UAVs. The UAVs measure the state of targets using Tracking-While-Scan (TWS) radars, each with different measurement dynamics and noise characteristics. Since the sensing capabilities of UAVs are constrained by the UAV dynamics and the onboard sensor performances, we should research how to plan ahead for the sensor platform movement to maximize target localization accuracy at a future time. Assumes that the system state transition has the markov property.The original problem is presented as a decentralized operation of a group of decision-makers lacking full observability of the global state of the system which can be modeled as a Decentralized Partial Observable Markov Decision Process (Dec-POMDP). Then the criterion function is the amount of information (i.e. the determinant of Fisher information matrix).

Hypothesis two UAVs are homogeneous. The first k UAV evolves based on $\boldsymbol{x}_k^o(t+1)=f\big(\boldsymbol{x}_k^o(t),\boldsymbol{u}_k(t)\big)$, where the mapping function f can be specified as a collection of simple kinematic equations that govern the UAV motion. The states of the UAV is given by $\boldsymbol{x}_k^o=[x_k^o,y_k^o,V_k^o,\theta_k^o]^{\mathrm{T}}$, where (x_k^o,y_k^o) represents the position coordinates, V_k^o represents the speed, and θ_k^o represents the heading angle. The kinematic equations of the UAV motion (Geiger et al. 2006) are as follows:

$$\begin{cases}x_k^o(t+1)=x_k^o(t)+V_k^o(t)T\cos\big(\theta_k^o(t)\big)\\ y_k^o(t+1)=y_k^o(t)+V_k^o(t)T\sin\big(\theta_k^o(t)\big)\\ V_k^o(t+1)=\big[V_k^o(t)+a_k(t)T\big]_{V_{\min}}^{V_{\max}}\\ \theta_k^o(t+1)=\theta_k^o(t)+a_k(t)T\tan\big(\phi_k(t)\big)/V_k^o(t)\end{cases} \tag{16}$$

where $[v]_{V_{\min}}^{V_{\max}}=\max\{V_{\min},\min(V_{\max},v)\}$, $V_{\min}$ and $V_{\max}$ are the minimum and the maximum limits on the speed of the UAV, g is the acceleration due to gravity, and T is the length of the time step.

We consider that the first l target state evolves as a linearized target motion model with zero-mean noise. as given below:

$$\boldsymbol{x}_l^t(t+1)=\mathbf{A}\boldsymbol{x}_l^t(t)+\boldsymbol{v}_l,\boldsymbol{v}_l\sim\mathcal{N}(\mathbf{0},\mathbf{Q}_l), \tag{17}$$

where the state of the lth target is given by $\boldsymbol{x}_l^t=[x_l^t,y_l^t,\dot{x}_l^t,\dot{y}_l^t]^{\mathrm{T}}$, and

$$\mathbf{A}=\left[\begin{array}{cccc}1&0&T&0\\0&1&0&T\\0&0&1&0\\0&0&0&1\end{array}\right]. \tag{18}$$

The position of the target relative to the UAV $\mathbf{x}=(x,y)^{\mathrm{T}}$ be

$$\mathbf{x}=(x^t-x^o,y^t-y^o)^{\mathrm{T}}$$

The UAV states are assumed to be fully observable. Sensors (mounted on the UAV) used for target tracking provide measurements of a target in a natural sensor coordinate system (CS). In our case, for track-while-scan radar, this CS is polar in 2-D with range r, bearing φ. The true relative target position is

$$\boldsymbol{h}(\mathbf{x})=\left[\begin{array}{c}\sqrt{x^2+y^2}\\ \arctan\left(y/x\right)\end{array}\right]. \tag{19}$$

The measure covariance matrix $\mathbf{C}(\mathbf{x}_k)$ depends on the the distance between the UAV and the target, and the specific form is

$$\mathbf{C}(\mathbf{x})=\left[\begin{array}{cc}r^2\sigma_r^2&0\\0&r^2\sigma_\varphi^2\end{array}\right] \tag{20}$$

where $r\sigma_r$ and $r\sigma_\varphi$ are the standard deviations of the range and bearing, respectively. r is the estimated range between the UAV and the target.

Since the underlying sensor collects target information only at discrete sampling points, the accumulative information should only take into account those points where the sensor will take measurements. Therefore, the sum of determinants of Fisher information matrix at H time horizon sampling locations will be used to approximate the accumulative information along the state trajectory $(\mathbf{x}_1,\mathbf{x}_2,...,\mathbf{x}_H)$, i.e.,

$$\mathcal{D}(\mathbf{x}_1,\mathbf{x}_2,...,\mathbf{x}_H)\approx\sum_{i=1}^{H}\det\big(\mathbf{G}(\mathbf{x}_i)\big) \tag{21}$$

The value of the determinant of a FIM represents the volume of the amount of information that can be acquired by the sensor. Our decision-making process is to find the optimal sequence of actions which maximized the amount of cumulative information.

4 Simulation Experiment and Analysis

We simulate a scenario with two UAVs and two targets as shown in Fig. 3. We set both the parameters of the airborne sensors for distance measurement variance $\sigma_r = 10^{-1}$, and Angle measurement variance $\sigma_\varphi = 10^{-3}\pi$. Assume that the aircraft is flying at a fixed altitudewhich is 100. This means that the distance between the sensor and the target becomes $r = \sqrt{x^2 + y^2 + 100}$. First of all, the information distribution of a target measured by two sensors represented graphically. Then, under the framework of Dec-POMDP, the information of fusion as the decision-making basis is applied to the scenario of double UAV tracking two target, and compare with the method of the information non fusion.

4.1 Sensor Network Information Distribution

In the process of target tracking, the amount of information contained in the measured data of the sensor determines the accuracy of the target state estimation or filtering. In this subsection, we study the distribution of information in static dual sensor network.

Information distribution are given respectively in Fig. 1 and Fig. 2, and each figure according to the amount of information for each Fisher information matrix determinant, and adopts the logarithmic coordinates. The distribution amount of information is representation of two-dimensional position estimation accuracy, the greater the amount of information, the higher precision of state estimation.

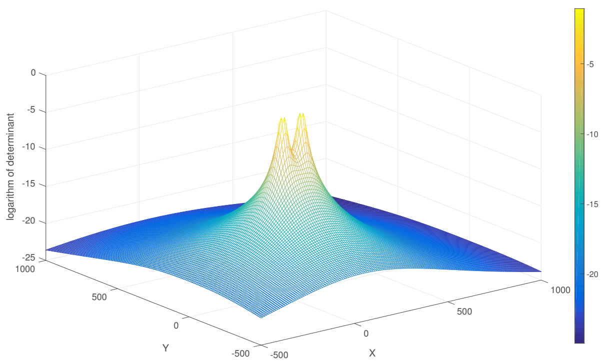

Fig. 1: The information distribution of data fusion

The Fig. 1 shows the distribution of the amount of information after the measured data fusion of the two sensors. The abscissa and ordinate represent the position of the target, and the vertical coordinate represents the amount information of data fusion, which is $\det\left(\mathbf{G}(\mathbf{x})\right) = \det\left(\sum_{k=1}^{2} \mathbf{G}_k(\mathbf{x})\right)$. The figure shows that for TWS radar, the amount of information has two maximum region, located in the vicinity of the two radars, respectively. When the target is located in the middle of the two radars, the amount of information is reduced, but the rate of reduction is significantly lower than that of the target when it is far away from the two radars.

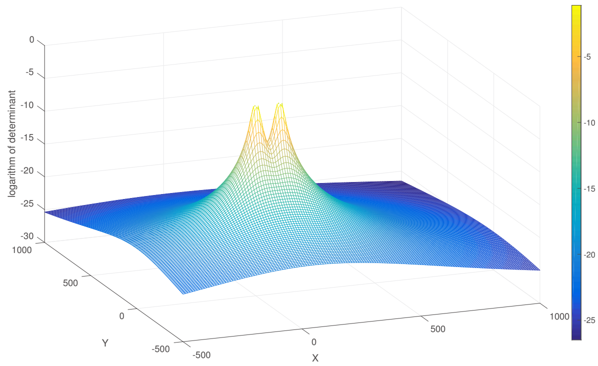

Fig. 2: The distribution of simple information accumulation

The Fig. 1 shows the distribution of the simple accumulation of information from the two sensors. It is similar to Fig. 2 in shape and trend, but it is obvious that the amount of information represented by the vertical axis is less than Fig. 2. As shown in theoretical derivation results, information fusion can improve the utilization of information contained in the data. On the other hand, due to the two figure approximation, it provides the idea of approximate tracking decision-making before data fusion, which is helpful to improve the efficiency of obtaining suboptimal strategy.

4.2 Dec-POMDP Approach vs. Greedy Approach

In the process of tracking and decision-making by Dec-POMDP, an unmanned aerial vehicle (UAV) must estimate the strategy of other aircraft, and then select the optimal strategy of the whole system according to the amount of information after data fusion.

In this subsection, we compare the performance of our Dec-POMDP approach with a greedy approach (defined as follows). In the greedy approach, the UAVs do not communicate with each other, and each UAV optimizes only its own kinematic controls over the time horizon ($H = 6$). Clearly, Dec-POMDP approach induces cooperation among the UAVs by letting each UAV optimize the joint kinematic controls (along with local communication decisions), and the greedy approach is non-cooperative in the sense that each UAV behaves in a non-cooperative manner by optimizing only its own kinematic controls. We apply these two approaches to a scenario with two UAVs and two targets.

Then we implement our approach in MATLAB, where we use the command $fmincon$ (an optimization tool in MATLAB) to minimize the objective function in (21). We perform 500-step simulations and every step is $0.1\ s$. In our simulations, the time horizon H is set to be $H = 6$ similar to [15]. It means that at the current time-step, we plan for 6 time-steps into the future, then implement the first control action for the current time-step, and discard the next 5 planned actions. At the very next time-step, we replan for 6 time-steps, and so forth. In addition, we set the speed of UAV in all simulations was between 20 m/s to 30 m/s. The target is moving in constant velocity in both simulation scenarios. We compare them using Dec-POMDP and greedy

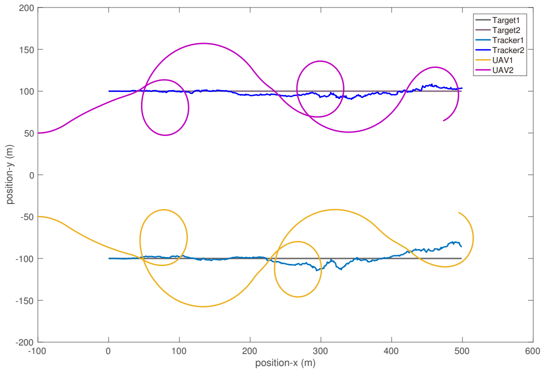

Fig. 3: The information distribution map of data fusion

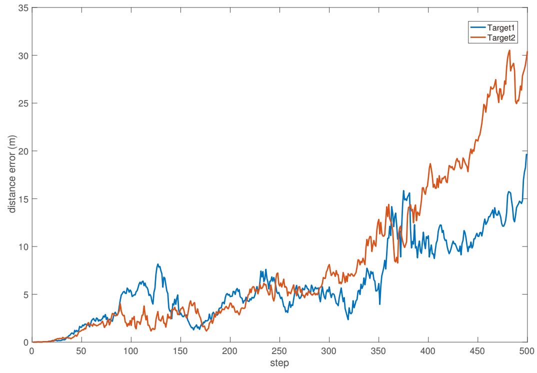

Fig. 6: The information distribution map of data fusion

approach. The Fig. 3 $\sim$ Fig. 6 show the simulation results. The meaning of every line in the figures are shown in the legend. Fig. 3 and Fig. 5 depict the behavior of the UAVs in the Dec-POMDP and greedy approaches respectively. We also compare the performances (with respect to average target-location error) of these two approaches, as depicted in Fig. 4 and Fig. 6.

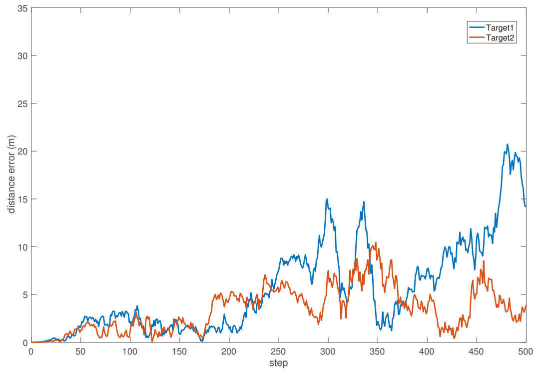

Fig. 4: The information distribution map of data fusion

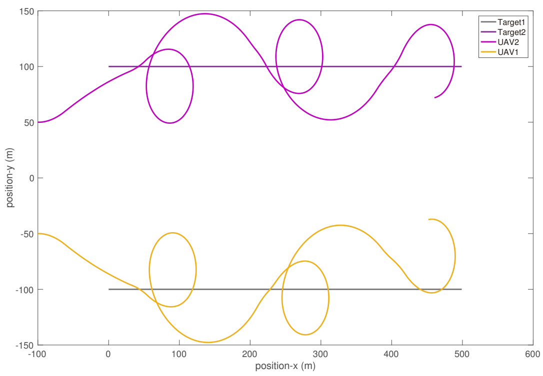

Fig. 5: The information distribution map of data fusion

Fig. 3 and Fig. 5 reveal that the two algorithms have little difference in flight trajectory. Because the two aircraft are close to a target respectively, and take into account that whether the information fusion has little effect on the distribution of information as shown in Fig. 1 and Fig. 2. It is unsurprising that the flight strategy is approximated, based on such an initial condition. Fig. 4 demonstrates that Dec-POMDP approach significantly outperforms the greedy approach in Fig. 6. Specifically, the Dec-POMDP approach results in average target location-error values that are approximately half small compared to that of the greedy approach. Further illustrate the validity of the information fusion for improving target positioning accuracy. That is to say, in order to improve the effect of tracking, it is necessary to exchange data through communication. But in real environments, communication is limited, and communication is costly, so combined with the results of the first simulation experiment, you can consider calculating the amount of information before communication and then weighing whether to communicate.

5 Conclusion and Future Work

This paper studies the change of the information quantity in the system before and after the information fusion, and indicates that information fusion can increase the amount of information. We also apply the analysis results to the multi-target tracking decision and compared the performance of our Dec-POMDP approach with a greedy approach. In the Dec-POMDP approach, each UAV optimizes joint actions and implements its local component. In contrast, in the greedy approach, each UAV only optimizes its own kinematic controls. We showed quantitatively how much our Dec-POMDP approach outperforms (with respect to average target-location error) the greedy approach. In the future, we will further study the relationship between the amount of information and the cost of communication in order to facilitate better multi-UAV collaborative decision.

References

[1] Fan X, Wang R, Sun S, et al. Context-centric needs anticipation using information needs graphs. *Applied Intelligence*, 2006, 24(1): 75-89.

[2] Amir O, Grosz B J, Stern R. To share or not to share? the single agent in a team decision problem. *Models and Paradigms for Planning under Uncertainty: a Broad Perspective*, 2014: 19.

[3] Kamar E, Gal Y, Grosz B J. Incorporating helpful behavior into collaborative planning. In *Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems*, 2009: 875-882.

[4] Amato C, Chowdhary G, Geramifard A, et al. Decentralized control of partially observable Markov decision processes, in

Proceedings of 52nd IEEE Conference on Decision and Control, 2013: 2398-2405.

[5] Bernstein D S, Givan R, Immerman N, et al. The complexity of decentralized control of Markov decision processes, *Mathematics of operations research*, 2002, 27(4): 819-840.

[6] Oliehoek F A. Decentralized POMDPs, in *Reinforcement Learning*. Springer Berlin Heidelberg, 2012: 471-503.

[7] Kaelbling L P, Littman M L, Cassandra A R. Planning and acting in partially observable stochastic domains, *Artificial intelligence*, 1998, 101(1): 99-134.

[8] Amato C, Konidaris G D, Kaelbling L P. Planning with macro-actions in decentralized POMDPs, in *Proceedings of the 2014 international conference on Autonomous agents and multi-agent systems*. International Foundation for Autonomous Agents and Multiagent Systems, 2014: 1273-1280.

[9] Cheng D, Controllability of switched bilinear systems, *IEEE Trans. on Automatic Control*, 50(4): 511–515, 2005.

[10] H. Poor, *An Introduction to Signal Detection and Estimation*. New York: Springer-Verlag, 1985, chapter 4.

[11] B. Smith, An approach to graphs of linear forms, accepted.

[12] D. Cheng, On logic-based intelligent systems, in *Proceedings of 5th International Conference on Control and Automation*, 2005: 71–75.

[13] D. Cheng, R. Ortega, and E. Panteley, On port controlled hamiltonian systems, in *Advanced Robust and Adaptive Control — Theory and Applications*, D. Cheng, Y. Sun, T. Shen, and H. Ohmori, Eds. Beijing: Tsinghua University Press, 2005: 3–16.

[14] Carter, Kevin M., et al. ”Fine: Fisher information nonparametric embedding.” *IEEE transactions on pattern analysis and machine intelligence*, 31(11) : 2093-2098, 2009.

[15] S. Ragi, and E. K. P. Chong, UAV path planning in a dynamic environment via partially observable Markov decision process, *IEEE Transactions on Aerospace and Electronic Systems*, 49(4): 2397–2412, 2013.