Asian Journal of Control, Vol. 16, No. 3, pp. 657–668, May 2014
Published online 19 March 2014 in Wiley Online Library (wileyonlinelibrary.com) DOI: 10.1002/asjc.846

ACTIVE COOPERATIVE OBSERVATION OF A 3D MOVING TARGET USING TWO DYNAMICAL MONOCULAR VISION SENSORS

Feng Gu, Yuqing He, and Jianda Han

ABSTRACT

Vision-based 3D moving target observation (3D-VMTO) is a central problem in the field of mobile robotics. In this paper, a new 3D-VMTO method called active cooperative observation is proposed through cooperation among two dynamical monocular vision sensors (MVSs). Under this method, an algorithm based on an extended set-membership filter is designed to fuse observations from different MVSs. An optimal observation condition is then introduced as the constraint into the relative velocity coordinates based path planning scheme, so as to reduce the influence of relative positioning among MVSs and their target on cooperative observation results. Experiments on a multiple rotor-flying-robots testbed demonstrate the validity and feasibility of the proposed method.

Key Words: 3D moving target localization, active cooperative observation, extended set-membership filter.

I. INTRODUCTION

Visual data has been crucial to the field of mobile robotics because it provides not only external environmental information, such as the distribution of obstacles and hazards, but also inner state information, such as the posture of the mobile robot itself. As a result, visual information processing, as well as its applications in control and planning, has become a major field of research in its own right, and vision-based moving target observation (VMTO) in particular has emerged as a central topic for investigation [1–4].

VMTO has been used in many applications, including border patrol systems and multi-robot coordination systems, etc. Although there are some cases in which monocular vision is sufficient for estimation and tracking of moving targets (*i.e.* if movement is mostly restricted to two dimensions), in most other cases, high-precision, real time 3D VMTO is essential. Thus, innovations in 3D VMTO technology can have broad positive impact.

A number of methods for 3D VMTO have been proposed, including depth from focus (DFF) [5], depth from defocus (DFD) [6], and stereo vision. DFF and DFD are both monocular vision sensors based 3D reconstruction methods that model the relationship between the depth change and blur of images. However, systems using these methods achieve good accuracy only at short range because of the small physical size of optical components [7]. By contrast, stereo vision can work well at longer ranges and is more often used in real application of mobile robotics. For example, in [8] and [9], stereo vision is used to detect a dynamic target and to create a 3D terrain map of the ground environment, respectively. Unfortunately, stereo vision is still of insufficient precision because the length of the stereoscopic baseline cannot be increased as desired. This makes it unsuitable for some typical applications, such as 3D VMTO in large areas.

Given recent developments in multi-robot systems, an effective alternative to stereo vision has emerged: track target objects by constructing a vision system composed of multiple dynamical vision sensors, *i.e.*, vision-based cooperative observation. Over the last few decades, this topic has attracted a significant amount of research, most of which has focused on how to improve observational accuracy using data fusion algorithms. For example: in [10], re-parameterization of 2D Gaussian distributions was used to combine information from more than two cameras; in [11], the simultaneous localization and mapping (SLAM) problem was addressed using two cooperative monocular vision sensors; and in [12], a Bayes estimator-based algorithm was employed to fuse the bearing information from multiple cameras.

In contrast to information fusion, a strategy largely ignored by researchers is to regulate the coordinated behavior of vision-sensor equipped robots. The strategy is derived from two fundamental observations. First, it has been shown that the measurement accuracy of single vision systems, both binocular and monocular, is a function of the relative position between the sensor and the target, and that the relative position will strongly influence the fused results, especially when

Manuscript received December 30, 2012; revised April 22, 2013; accepted October 10, 2013.

The authors are with the State Key Laboratory of Robotics, Shenyang Institute of Automation Chinese Academy of Sciences, Nanta Street, 114 Shenyang, 110016, China.

This work is supported by National Nature Sciences Foundation of China (61203340, 61535005).

both the sensor and the target are in motion. Second, in complex dynamical environments, occlusion significantly deteriorates observational results, and can only be avoided by moving dynamical vision sensors to maintain line of sight.

This strategy has gradually received attention in recent research. For example, in [13], the so-called Gauss-Seidel Relaxation method was introduced to regulate the trajectory of a team of heterogeneous robots to minimize the observation uncertainty of a target object. However, this work was based on a stochastic model, and therefore assumed that process and measurement error are stochastic variables with known prior mean and covariance. These assumptions are difficult to satisfy in real applications [14].

In this paper, the 3D-VMTO problem with two dynamical monocular vision sensors (MVSs) is researched by solving the following two sub-problems: (i) the data fusion algorithm; and (ii) the coordination behavior optimization algorithm. The first sub-problem has been researched elsewhere in the literature, but as in [13], most existing solutions are based on a stochastic model. In our proposed method, data fusion is accomplished using a more robust method, called Set Theory-based Estimation (STE), which assumes that the measurement error and model error are both unknown but bounded (UBB) [15] and can produce an estimated uncertain set to which the actual system state belongs [16]. For the second sub-problem, we construct a path planning algorithm in relative velocity coordinates (RVCs) so as to guide the movement of MVSs. These two algorithms are connected by an optimal observation condition to form the so-called active cooperative observation.

II. PROBLEM DESCRIPTION

The sketch of 3D VMTO using two dynamical monocular vision sensors (MVSs) is shown in Fig. 1.

Based on the basic principle of stereo vision, the relationship between the coordinates of the target in pixel coordinate system (PCS) (u_i, v_i) and world coordinate system (WCS) (x_T, y_T, z_T) is denoted as the following (1) and (2):

$$z_{c1}\begin{pmatrix} u_1 \\ v_1 \\ 1 \end{pmatrix} = \mathbf{F}_1 \cdot \mathbf{ROT}_1 \begin{pmatrix} x_T \\ y_T \\ z_T \\ 1 \end{pmatrix} \triangleq \mathrm{M}_1(\mathbf{X}_1)\begin{pmatrix} x_T \\ y_T \\ z_T \\ 1 \end{pmatrix} \tag{1}$$

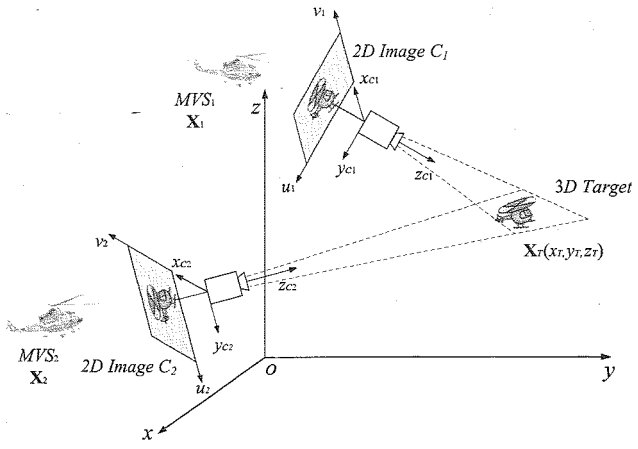


Fig. 1. Sketch of the 3D VMTO using two MVSs.
C_1 and C_2 are the image planes of two monocular vision systems; (x, y, z) is the world coordinate system (WCS); (x_{ci}, y_{ci}, z_{ci}) and (u_i, v_i) are the camera coordinate system (CCS) and pixel coordinate system (PCS) of each vision sensor respectively.

$$z_{c2}\begin{pmatrix} u_2 \\ v_2 \\ 1 \end{pmatrix} = \mathbf{F}_2 \cdot \mathbf{ROT}_2 \begin{pmatrix} x_T \\ y_T \\ z_T \\ 1 \end{pmatrix} \triangleq \mathrm{M}_2(\mathbf{X}_2)\begin{pmatrix} x_T \\ y_T \\ z_T \\ 1 \end{pmatrix} \tag{2}$$

where,

$$\mathbf{F}_i = \begin{bmatrix} f_{xi} & 0 & u_{0i} & 0 \\ 0 & f_{yi} & v_{0i} & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

$$\mathbf{ROT}_i = \begin{bmatrix} \mathbf{R}_i(\mathbf{X}_i) & \mathbf{T}_i(\mathbf{X}_i) \\ \mathbf{0}^T & 1 \end{bmatrix}$$

$$\mathbf{M}_i(\mathbf{X}_i) = \mathbf{F}_i\mathbf{ROT}_i \triangleq \begin{bmatrix} m_{11}^i(\mathbf{X}_i) & m_{12}^i(\mathbf{X}_i) & m_{13}^i(\mathbf{X}_i) & m_{14}^i(\mathbf{X}_i) \\ m_{21}^i(\mathbf{X}_i) & m_{22}^i(\mathbf{X}_i) & m_{23}^i(\mathbf{X}_i) & m_{24}^i(\mathbf{X}_i) \\ m_{31}^i(\mathbf{X}_i) & m_{32}^i(\mathbf{X}_i) & m_{33}^i(\mathbf{X}_i) & m_{34}^i(\mathbf{X}_i) \end{bmatrix}$$

z_{ci} denotes the z-coordinate of the target in the camera coordinate system (CCS) of the ith camera; x_T, y_T, and z_T compose the position vector of the target in the WCS; f_{xi}, f_{yi}, and u_{0i}, v_{0i} are all the inter parameters of the ith vision sensor; $\mathbf{R}_i(\mathbf{X}_i)$ and $\mathbf{T}_i(\mathbf{X}_i)$ are rotation matrix and translation vector between CCS and WCS. It should be noted that, since the two monocular vision sensors are both movable, rotation matrix

and translation vector $\mathbf{R}_i(\mathbf{X}_i)$ and $\mathbf{T}_i(\mathbf{X}_i)$ are also dynamical and can be denoted as the functions of each MVSs state $\mathbf{X}_i$.

By combing (1) and (2), the unknown variable of z_{c1} and z_{c2} can be counteracted, and we have,

$$\begin{cases} u_1 = \\ \dfrac{m_{14}^1(\mathbf{X}_1)+m_{11}^1(\mathbf{X}_1)x_T+m_{12}^1(\mathbf{X}_1)y_T+m_{13}^1(\mathbf{X}_1)z_T}{m_{34}^1(\mathbf{X}_1)+m_{31}^1(\mathbf{X}_1)x_T+m_{32}^1(\mathbf{X}_1)y_T+m_{33}^1(\mathbf{X}_1)z_T} \\ v_1 = \\ \dfrac{m_{24}^1(\mathbf{X}_1)+m_{21}^1(\mathbf{X}_1)x_T+m_{22}^1(\mathbf{X}_1)y_T+m_{23}^1(\mathbf{X}_1)z_T}{m_{34}^1(\mathbf{X}_1)+m_{31}^1x_T(\mathbf{X}_1)+m_{32}^1(\mathbf{X}_1)y_T+m_{33}^1(\mathbf{X}_1)z_T} \\ u_2 = \\ \dfrac{m_{14}^2(\mathbf{X}_2)+m_{11}^2(\mathbf{X}_2)x_T+m_{12}^2(\mathbf{X}_2)y_T+m_{13}^2(\mathbf{X}_2)z_T}{m_{34}^2(\mathbf{X}_2)+m_{31}^2(\mathbf{X}_2)x_T+m_{32}^2(\mathbf{X}_2)y_T+m_{33}^2(\mathbf{X}_2)z_T} \\ v_2 = \\ \dfrac{m_{24}^2(\mathbf{X}_2)+m_{21}^2(\mathbf{X}_2)x_T+m_{22}^2(\mathbf{X}_2)y_T+m_{23}^2(\mathbf{X}_2)z_T}{m_{34}^2(\mathbf{X}_2)+m_{31}^2(\mathbf{X}_2)x_T+m_{32}^2(\mathbf{X}_2)y_T+m_{33}^2(\mathbf{X}_2)z_T} \end{cases} \tag{3}$$

In this paper, the ACO problem is divided into the following two problems.

2.1 Estimation problem

Since the target is movable, the state estimation of it can be denoted as the following dynamical estimation problem:

$$\mathbf{X}_T(k+1)=\mathbf{f}(\mathbf{X}_T(k),\mathbf{w}(k)) \tag{4}$$

$$\mathbf{y}(k)=\mathbf{h}(\mathbf{X}_1(k),\mathbf{X}_2(k),\mathbf{X}_T(k),\mathbf{n}(k)) \tag{5}$$

where (4) is the motion equation of the moving target; (5) models the measurements of two MVSs; $\mathbf{X}_T(k)$ and $\mathbf{X}_i(k)(i=1, 2)$ are the state vector of target and two MVSs, respectively; $y(k)$ is the measurement vector that can be denoted as (6); $w(k)$ and $n(k)$ denote model and measurement errors; k and $k+1$ means that the corresponding variable is at time instant k and $k+1$.

$$\mathbf{h}(\mathbf{X}_1(k),\mathbf{X}_2(k),\mathbf{X}_T(k),\mathbf{n}(k)) \triangleq \begin{bmatrix} u_1 \\ v_1 \\ u_2 \\ v_2 \end{bmatrix} + \mathbf{n}(k) \tag{6}$$

where u_1, v_1, u_2, v_2 are defined as in Eq. (3).

Then the target states can be estimated based on the system equation (4) and the measurement equation (6) through some nonlinear estimation algorithms, which can be denoted as a map as follows:

$$\mathbf{X}_T(k+1)=Z(\mathbf{X}_T(k),\mathbf{y}(k),\mathbf{w}(k),\mathbf{n}(k)) \tag{7}$$

Remark 1. In most applications, the detailed motion equation of the target is difficult to obtain. Thus the noise-driven integral or double integral equation is usually used to model the target in-order to predict its motion, *i.e.*,

$$\begin{cases} \mathbf{X}_T(k+1)=\mathbf{X}_T(k)+\mathbf{v}_T(k)\Delta T+\mathbf{w}_1(k) \\ \mathbf{v}_T(k+1)=\mathbf{w}_2(k) \end{cases} \tag{8}$$

where $X_T(k)=(x_T(k),\ y_T(k),\ z_T(k))^{\mathrm{T}}$ and $v_T(k)=(v_x(k),\ v_y(k),\ v_z(k))^{\mathrm{T}}$ are the position vector and velocity vector of the target at time instant k respectively; ΔT is sampling time; $w_k=(w_1(k),\ w_2(k))$ denotes the modeling errors.

Remark 2. In the proposed dynamical stereo system, the relative posture of the two vision sensors is not fixed and has only limited precision. In this work, we consider these by rewriting the states of each MVS system as follow,

$$\begin{aligned} &\mathbf{y}(k)= \\ &\mathbf{h}(\mathbf{X}_1(k)+e_1(k),\mathbf{X}_2(k)+e_2(k),\mathbf{X}_T(k),\mathbf{n}(k)) \end{aligned} \tag{9}$$

In order to avoid dealing with the extra errors, $e_1(k)$ and $e_2(k)$ is included into the errors $n(k)$ by linearization. That means, $n(k)$ is an integrated error that considers not only the measurement error from vision sensors but also the modeling errors from the state of each MVS system.

2.2 Planning problem

Another problem is how to make full use of the dynamics performance of the MVS to improve the observation accuracy. Assume that the motion of MVS is dominated by the following equation,

$$\mathbf{X}_i(k+1)=\mathbf{f}(\mathbf{X}_i(k),\mathbf{u}_i(k)) \tag{10}$$

where $f(*,\ *)$ are some pre-defined functions (linear or nonlinear); $\mathbf{X}_i(k)$and $\mathbf{X}_i(k+1)$ are the states of the ith MVS at the time instant k and $k+1$, respectively. In this paper, the aim of the planning problem is to find some mapping which connects the current states of both MVSs and the target to the optimal behavior with respect to some cost function by considering the data fusion results (the so-called optimal observation condition in Section IV). This sub-problem can be modeled as constructing explicit or implicit mapping as follows:

$$\mathbf{u}(k)^{optimal}=\Xi(\mathbf{X}_1(k),\mathbf{X}_2(k),\mathbf{X}_T^e(k),\mathbf{y}(k)) \tag{11}$$

where $\mathbf{X}_T^e(k)$ is the predicted state of the moving target at time instant k.

Remark 3. In this paper, the following kinematic equation is used to denote the action of MVS system for the common applications:

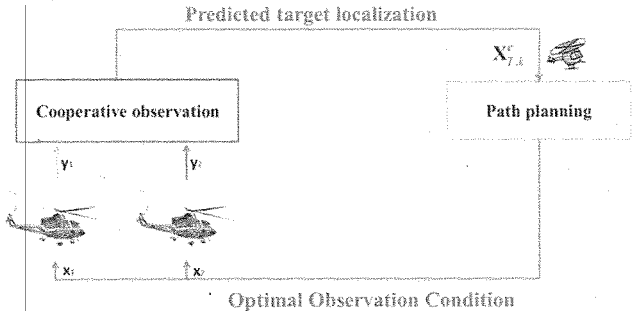


Fig. 2. Main idea of ACO.

$$\begin{cases}\mathbf{X}_i(k+1)=\mathbf{X}_i(k)+\mathbf{v}_i(k)\Delta T\\ \mathbf{v}_i(k+1)=\mathbf{v}_i(k)+\mathbf{a}_i(k)\Delta T\\ \mathbf{a}_i(k+1)=\mathbf{u}_i(k)\end{cases} \tag{12}$$

$$\Gamma(\mathbf{X}_i(k),\mathbf{v}_i(k),\mathbf{a}_i(k))\le 0 \tag{13}$$

where, $\Gamma(*, *, *)$ is the motion constraint inequality of the MVS.

With the above two sub-problems, the ACO problem in this paper can be described as follows.

ACO Problem. With system model (4) and the measurement equation (6), based on vision information from two monocular vision sensors, design the data fusion estimation algorithm $Z(*, *, *, *)$ and the coordinate path planning algorithm $\Xi(*, *, *)$ to obtain 3D observations of the moving target.

The whole process and main idea of ACO is sketched in Fig. 2.

III. ESMF BASED COOPERATIVE OBSERVATION

In this section, the estimation sub-problem is solved based on the set theory based filter, *i.e.*, extended set-member filter (ESMF). ESMF is a kind of nonlinear bound guarantee estimation and filter method which supposes the error is unknown but bounded (UBB) and often modeled as an ellipsoid. That means, model errors $w(k)$ and measurement errors $n(k)$ satisfied

$$\begin{aligned}&[\mathbf{w}(k)]^{\mathrm{T}}\mathbf{Q}(k)^{-1}[\mathbf{w}(k)]\le 1\\ &[\mathbf{n}(k)]^{\mathrm{T}}\mathbf{R}(k)^{-1}[\mathbf{n}(k)]\le 1\end{aligned} \tag{14}$$

where $\mathbf{Q}(k)$ and $\mathbf{R}(k)$ are both symmetric and positive definite matrices.

Similarly, the estimated states are also denoted as a super ellipsoid set given by the following equation [16],

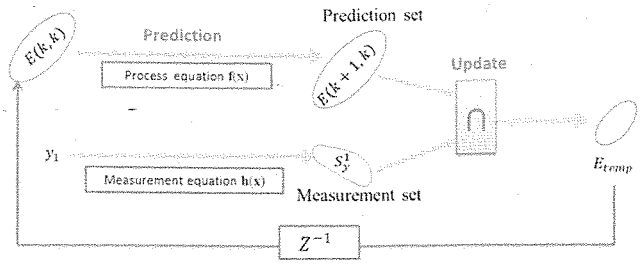


Fig. 3. Basic idea of ESMF algorithm.

$$\begin{aligned}&E(\hat{\mathbf{X}},P)=\\ &\left\{\mathbf{X}\in R^n\middle|(\mathbf{X}-\hat{\mathbf{X}})^{\mathrm{T}}\,\mathrm{P}^{-1}(\mathbf{X}-\hat{\mathbf{X}})\le 1\right\}\end{aligned} \tag{15}$$

where $\hat{\mathbf{X}}$ is the central point of the ellipsoid; P is an envelope matrix satisfying symmetric and positive definite condition.

Generally, the ESMF algorithm executes cyclically the following two steps (Fig. 3).

- **Prediction:** to predict the new system state by computing an ellipsoid of the new state $\mathbf{X}(k)$ based on the states uncertainty set $\mathbf{X}(k-1)$ at the last time instant and the model error set $w(k)$ through the nonlinear process equation.
- **Update:** to revise the prediction state set using the newly obtainable measurement uncertainty set and the measurement equation.

The nonlinearity in measurement equation (5) makes the update step difficult to be implemented because set computing of nonlinear functions is highly complicated. Thus linearization is firstly conducted with respect to the target state variable $\mathbf{X}_T(k)$.

$$\begin{aligned}&\mathbf{y}(k)=\mathbf{h}|_{\mathbf{X}_T(k)=\hat{\mathbf{X}}_T(k)}+\\ &\nabla_{\mathbf{X}_T(k)}\mathbf{h}|_{\mathbf{X}_T(k)=\hat{\mathbf{X}}_T(k)}\Delta\mathbf{X}_T(k)+\\ &O\left[\Delta\mathbf{X}_T(k)^{\mathrm{T}}\Delta\mathbf{X}_T(k)\right]+\mathbf{n}(k)\end{aligned} \tag{16}$$

where h is defined as (6), $\nabla_{\mathbf{X}_T(k)}\mathbf{h}$ indicates the gradient of $h(\cdot)$ with respect to $\mathbf{X}_T(k)$; $\hat{\mathbf{X}}_T(k)$ is the current estimated result and $\Delta\mathbf{X}_T(k)=\left(\mathbf{X}_T(k)-\hat{\mathbf{X}}_T(k)\right)$.

The higher order terms (HOT) are considered as part of the measurement error in ESMF algorithm, so as to ensure that the real states are in the estimated set. Thus the new measurement equation can be rewritten as,

$$\begin{aligned}&\mathbf{y}(k)=\mathbf{h}|_{\mathbf{x}_T(k)=\hat{\mathbf{x}}_T(k)}+\\ &\nabla_{\mathbf{X}_T(k)}\mathbf{h}|_{\mathbf{X}_T(k)=\hat{\mathbf{X}}_T(k)}\Delta\mathbf{X}_T(k)+\hat{\mathbf{n}}(k)\end{aligned} \tag{17}$$

where $\hat{\mathbf{n}}(k)$ is the new measurement error term.

The new error uncertainty set can be denoted as,

$$[\hat{\mathbf{n}}(k+1)]^{\mathrm{T}}\,\hat{\mathbf{R}}(k+1)^{-1}[\hat{\mathbf{n}}(k+1)]\leq 1$$

with

$$\hat{\mathbf{R}}(k+1)=\frac{\bar{\mathbf{R}}(k+1)}{(1-\beta_R)}+\frac{\mathbf{R}(k+1)}{\beta_R}$$

$$\beta_R=\frac{\sqrt{tr(\mathbf{R}(k+1))}}{\sqrt{tr(\bar{\mathbf{R}}(k+1))}+\sqrt{tr(\mathbf{R}(k+1))}}$$

where $\bar{\mathbf{R}}(k+1)$, which can be obtained using interval analysis [16], are the envelope matrix of the HOT: $O[\Delta\mathbf{X}_T(k)^{\mathrm{T}}\Delta\mathbf{X}_T(k)]$, $tr(\cdot)$ means the trace of the matrix.

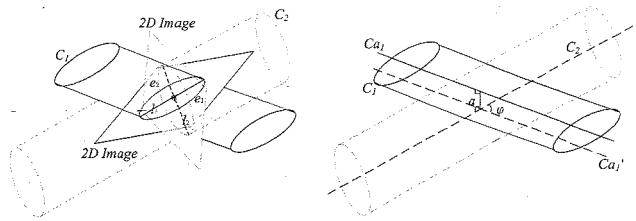
Remark 4. From the above analysis, it is clear that the robustness of the ESMF algorithm is mainly because: all kinds of errors and noises can be included in the algorithm, which make the real value to be always a member of the estimated uncertainty set. Of course, these procedures maybe result in conservativeness. This problem, fortunately, can be alleviated by the ACO algorithm discussed later.

With the linearized system, ESMF algorithm can be used to estimate the state of system (8) and (6). The detailed process can be found in [15].

IV. PATH PLANNING ALGORITHM FOR ACTIVE COOPERATIVE OBSERVATION

4.1 Optimal observation condition

The optimal observation condition of the ACO problem means a formation of the two MVSs and the target. With this formation, the best cooperative observation results can be obtained. In this section, we will explain and compute the optimal observation condition. Firstly, the procedure of monocular vision observation can be shown as in Fig. 4. In Fig. 4a, the real position of target is denoted as a '*'. Based on the UBB condition as in the last section, each monocular vision sensors measurement can be denoted as an ellipse, within which the real position of the target must lie. Furthermore, because the observation is two-dimensional, the real observation uncertain set in 3D space can be denoted as a cylinder whose axis is perpendicular to the corresponding image plane. Thus, the data fusion result is actually the intersection of the two cylinders. That means, the optimal observation condition is the one at which the centerline of the two cylinders are perpendicular, and the aim of the path planning here is to make the intersection of the two cylinders smallest through by regulating the relative posture of the two cylinders, *i.e.*, relative position among the target and the two MVSs.



(a) Sketch of the data fusion. (b) Relative posture of two cylinders.

Fig. 4. Relative relationship between two observation uncertainty set.
*C_1–*cylinder1*; C_2–*cylinder2*; Ca_1–*axis of cylinder1*; Ca_2–*axis of cylinder2*; Ca_1', *is intersecting line which is parallel to* Ca_1; a–*the smallest relative distance between two cylinders*; φ–*angle between two cylinders*; e_1, e_2–*uncertainty sets in the image plane*; l_1, l_2–*the length of long axis of* e_1 *and* e_2.

Two variables will influence the intersection of the two cylinders: one is the relative position, *i.e.*, the smallest distance of the two central axes; the other is the relative posture, *i.e.*, the angle between the two axes. The relative posture can be easily changed through regulating the relative position of two MVSs. However, the relative position cannot be defined beforehand because the measurement result of each monocular vision sensor are uncertain in an directed ellipsoid, thus our so-called optimal observation condition can only be denoted as the relative posture of the two cylinder and thus can be computed by solving the following min-max optimization problem:

$$\min_{\varphi}\max_{a\leq(l_1+l_2)/2} V(C_1\bigcap C_2) \tag{18}$$

On the optimal observation condition, the following theorem can be shown:

Theorem I. Assume the measurement of each MVS is isotropous, *i.e.*, the uncertainty set in the image plane shown as e_1 and e_2 in Fig. 4a is a circle, then the optimal observation condition, *i.e.*, the solution of (18), is $\varphi=\pi/2$.

Proof. The condition of the Theorem I means the two 2D measurement uncertainty sets are two circles. This is usually reasonable since the difference between two directions is very small.

Theorem I can be shown using the following two steps: (i) the volume of intersection between the two cylinders is maximal when $a=0$ regardless of the angle φ; (ii) the optimal angle φ is $\pi/2$ when $a=0$.

The first step can be directly obtained from [17]. Now we will show the second step. Without considering the absolute posture of the two cylinders, the following equation can be used to denote them:

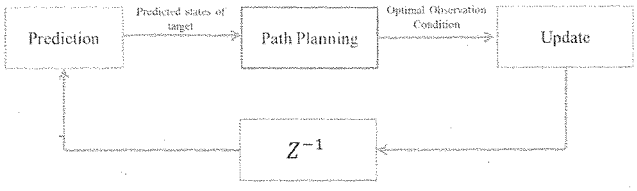


Fig. 5. Scheme of active cooperative observation.

$$x^2 + y^2 \le r_1^2 \tag{19}$$

$$x^2 + (y\cos\alpha + z\sin\alpha)^2 \le r_2^2 \tag{20}$$

Thus, the intersection between them can be computed using the following double integration in the area denoted as (19). Described in the cylinder coordination, the area of intersection can be denoted as,

$$\frac{2}{\sin\alpha}\int_0^{2\pi} d\theta \int_0^{r_1} \sqrt{r_2^2 - r_1^2\sin^2\theta}\, dr \tag{21}$$

where r and θ denote the coordinates in cylinder coordinate of 3D space. From (21), it is easy to conclude that the volume of the intersection is smallest if $\varphi = \pi/2$. This completes the proof of Theorem I.

4.2 Relative velocity coordinates-based path planning

The implementation of path planning for the ACO problem can be shown as Fig. 5, where the path planning algorithm is embedded into the cooperative observation algorithm between prediction step and update step so that the MVSs can plan their trajectories to achieve the optimal observation condition and to satisfy some other objectives.

In this paper, we use the RVC-based method in [18] to optimize the coordinate behavior of the two MVSs. As discussed in Section 4.1, the optimal observation condition, combined to target tracking and collision avoidance, composes the optimization objective for ACO algorithm.

4.2.1 Optimal observation condition

We have shown that the optimal observation condition is the one where the two cylinders are perpendicular. That means the cooperative observation angle (COA) φ shown in Fig. 6 should be, or approaches, as far as, possible 90 degree. From Fig. 6, $\cos\varphi$ can be denoted as:

$$\cos\varphi = \frac{(\mathbf{X}_1(k+1) - \mathbf{X}_T(k+1,k))\cdot(\mathbf{X}_2(k+1) - \mathbf{X}_T(k+1,k))}{|\mathbf{X}_1(k+1) - \mathbf{X}_T(k+1,k)||\mathbf{X}_2(k+1) - \mathbf{X}_T(k+1,k)|} \tag{22}$$

where |*| denotes the Euclidean norm. From system equation of the MVS as (12), we have,

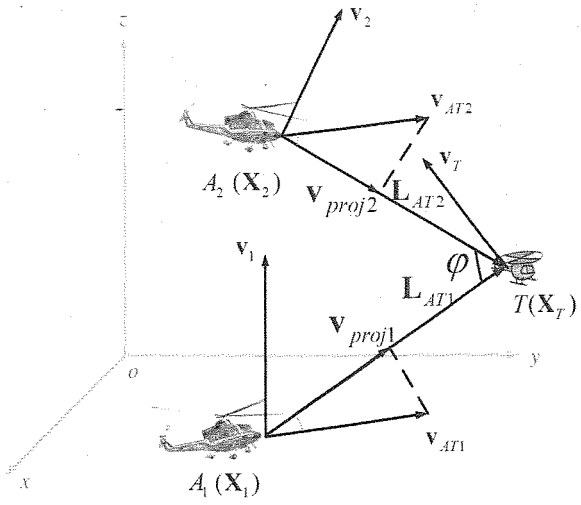


Fig. 6. Main idea of RVCs based path planning for ACO.
** A_i (i = 1 or 2) and T denote the two MVSs and the moving target, respectively; v_1, v_2 and v_T are the absolute velocity of the two MVSs and the moving target; v_{ATi} denotes the relative velocity between the MVS and the target; L_{ATi} means the directed lines from the ith MVS to the moving target; v_{proji} is the projection of v_{ATi} along the direction of L_{ATi}; φ is the angle between L_{AT1} and L_{AT2}, that is, cooperative observation angle.*

$$\mathbf{X}_i(k+1) = \mathbf{X}_i(k) + \mathbf{v}_i(k)\Delta T + \frac{\Delta T^2}{2}\mathbf{u}_i(k) \tag{23}$$

Substituting (23) into (22) and conducting linearization, we have

$$\cos\varphi = c_0 + (\mathbf{c}_1, \mathbf{c}_2)\cdot\begin{pmatrix}\mathbf{u}_1(k)\\ \mathbf{u}_2(k)\end{pmatrix} \tag{24}$$

where

$$c_0 = \frac{\mathbf{E}_1\cdot\mathbf{E}_2}{|\mathbf{E}_1||\mathbf{E}_2|}$$

$$\mathbf{c}_1 = \frac{\Delta T^2}{2|\mathbf{E}_1||\mathbf{E}_2|}\cdot\mathbf{E}_2^{\mathrm{T}} - \frac{(\mathbf{E}_2\cdot\mathbf{E}_1)\Delta T}{2|\mathbf{E}_1|^3|\mathbf{E}_2|}\cdot\mathbf{E}_1^{\mathrm{T}}$$

$$\mathbf{c}_2 = \frac{\Delta T^2}{2|\mathbf{E}_1||\mathbf{E}_2|}\cdot\mathbf{E}_1^{\mathrm{T}} - \frac{(\mathbf{E}_2\cdot\mathbf{E}_1)\Delta T}{2|\mathbf{E}_1||\mathbf{E}_2|^3}\cdot\mathbf{E}_2^{\mathrm{T}}$$

$$\mathbf{E}_1 = \mathbf{X}_1(k) + \Delta T\mathbf{v}_1(k) - \mathbf{X}_T(k+1,k)$$

$$\mathbf{E}_2 = \mathbf{X}_2(k) + \Delta T\mathbf{v}_2(k) - \mathbf{X}_T(k+1,k)$$

Thus, the cost function of the cooperative observation can be denoted as,

$$J_1 = \left| c_0 + (\mathbf{c}_1, \mathbf{c}_2) \cdot \begin{pmatrix} \mathbf{u}_1(k) \\ \mathbf{u}_2(k) \end{pmatrix} \right| \tag{25}$$

Then the optimization of COA is to minimize J_1 by adjusting the $u_i(k)$. In order to transform (25) into a linear function, we define a positive variable z_1 as follows [18]:

$$-z_1 \le c_0 + (\mathbf{c}_1, \mathbf{c}_2) \cdot \begin{pmatrix} \mathbf{u}_1(k) \\ \mathbf{u}_2(k) \end{pmatrix} \le z_1 \tag{26}$$

Then, minimizing z_1 subjected to the inequality of (26) is equivalent to minimizing the following new J_1,

$$J_1 = w_1 z_1 \tag{27}$$

where $w_1 \ge 0$ is the to be designed weight value of this objective.

4.2.2 Target pursuit

Target pursuit is necessary for the MVS to track the target and obtain its sustaining state. It is naturally that the optimal target pursuit can be implemented to maximum, *i.e.* maximize the following v_{proji} (Fig. 6):

$$\mathbf{v}_{proji}(k+1) = \mathbf{L}_{ATi}(k+1) \cdot \mathbf{v}_{ATi}(k+1) \tag{28}$$

where $\mathbf{L}_{ATi}(k+1)$ is the directed line from the *ith* MVS to the target at time instant $k+1$ and can be calculated as:

$$\mathbf{L}_{ATi}(k+1) = \mathbf{X}_T(k+1, k) - \mathbf{X}_i(k) \tag{29}$$

By assuming that the target velocity is constant during each sampling time ΔT, the following equation is always satisfied:

$$\begin{aligned} \Delta \mathbf{v}_{ATi}(k) &\triangleq \mathbf{v}_{ATi}(k+1) - \mathbf{v}_{ATi}(k) \\ &= \mathbf{v}_T(k+1) - \mathbf{v}_i(k+1) - \mathbf{v}_T(k) + \mathbf{v}_i(k) \\ &= -\mathbf{v}_i(k+1) + \mathbf{v}_i(k) \triangleq -\Delta \mathbf{v}_i(k) \\ &= -\mathbf{u}_i(k) \Delta T \end{aligned} \tag{30}$$

Then the relative velocity at time instant $k+1$ can be denoted as,

$$\begin{aligned} \mathbf{v}_{ATi}(k+1) &= \mathbf{v}_{ATi}(k) - \Delta \mathbf{v}_i(k) \\ &= \mathbf{v}_{ATi}(k) - \mathbf{u}_i(k) \Delta T \end{aligned} \tag{31}$$

The goal of target pursuit is to regulate the $u_i(k)$ to maximize $v_{pfoji}(k+1)$. Thus the cost function of target pursuit can be denoted as,

$$\begin{aligned} &J_2 = \\ &-(\mathbf{L}_{AT1}(k) \cdot \mathbf{v}_{AT1}(k) + \mathbf{L}_{AT2}(k) \cdot \mathbf{v}_{AT2}(k)) + \\ &\Delta T(w_2 \mathbf{L}_{AT1}(k) \cdot \mathbf{u}_1(k) + w_3 \mathbf{L}_{AT2}(k) \cdot \mathbf{u}_2(k)) \end{aligned} \tag{32}$$

where $w_2 \ge 0$ and $w_3 \ge 0$ are two to be designed weight values used to trade off among other performance.

4.2.3 Collision avoidance

Collision avoidance is one of the most important problems in multiple robots cooperation. Here the collision can be explained as the following two aspects: collision between any MVS and the target, and that between the two MVSs. However, in our path planning method, because of the cost function J_1, the ACO φ can be kept near 90 by the two MVSs as shown in Fig. 6. Then, combined with J_1, keeping the MVSs apart from the target can not only avoid the collision between MVS and the target but also that between the two MVSs. Thus, the desired distance between each MVS and target are taken as the collision avoidance cost function.

The distance between each MVS and the target can be denoted as

$$\begin{aligned} &D_1(\mathbf{u}_1(k)) \triangleq d_1(k+1) \\ &= |\mathbf{X}_1(k+1) - \mathbf{X}_T(k+1, k)| \end{aligned} \tag{33}$$

$$\begin{aligned} &D_2(\mathbf{u}_2(k)) \triangleq d_2(k+1) \\ &= |\mathbf{X}_2(k+1) - \mathbf{X}_T(k+1, k)| \end{aligned} \tag{34}$$

Substituting (23) into (33) and (34), and linearizing, we have

$$D_1(\mathbf{u}_1(k)) = |\mathbf{E}_1| + \left(\frac{\Delta T^2}{|\mathbf{E}_1|} \right) \cdot \mathbf{E}_1^T \cdot \mathbf{u}_1(k)$$

Then the cost function of collision avoidance can be denoted as,

$$J_3 = |D_1(\mathbf{u}_1(k)) - Dc_1| \tag{35}$$

$$J_4 = |D_2(\mathbf{u}_2(k)) - Dc_2| \tag{36}$$

where Dc_1, Dc_2 are the desired distance between each MVS and the target and they can be always set larger than the safety distance. Using the same method as in the cost function of optimal observation condition (25), minimizing J_3 and J_4 is equivalent to minimizing

$$J_3 = z_2 \tag{37}$$

$$J_4 = z_3 \tag{38}$$

subjected to the following inequality:

$$-z_2 \le |\mathbf{E}_1| + w_4 \frac{\Delta T^2}{|\mathbf{E}_1|} \mathbf{E}_1^{\mathrm{T}} \cdot \mathbf{u}_1(k) - Dc_1 \le z_2 \tag{39}$$

$$-z_3 \le |\mathbf{E}_2| + w_5 \frac{\Delta T^2}{|\mathbf{E}_2|} \mathbf{E}_2^{\mathrm{T}} \cdot \mathbf{u}_2(k) - Dc_2 \le z_3 \tag{40}$$

by regulating $u_1(k)$ and $u_2(k)$, where z_2 andz_3 are both nonnegative. In (39) and (40), $w_4 \ge 0$ and $w_5 \ge 0$ are weight values.

With the preceding three cost functions, the motion planning can be modeled as a linear programming (LP) problem:

Minimizing: $J = J_1 + J_2 + J_3 + J_4$

Subjecting to:

$$\begin{cases} -z_1 \le c_0 + (\mathbf{c}_1, \mathbf{c}_2) \cdot \begin{pmatrix} \mathbf{u}_{1,k} \\ \mathbf{u}_{2,k} \end{pmatrix} \le z_1 \\ -z_2 \le |\mathbf{E}_1| + w_4 \dfrac{\Delta T^2}{|\mathbf{E}_1|} \mathbf{E}_1^{\mathrm{T}} \cdot \mathbf{u}_1(k) - Dc_1 \le z_2 \\ -z_3 \le |\mathbf{E}_2| + w_5 \dfrac{\Delta T^2}{|\mathbf{E}_2|} \mathbf{X}_2^{\mathrm{T}} \cdot \mathbf{u}_2(k) - Dc_2 \le z_3 \\ \mathbf{u}_{1,lower} \le \mathbf{u}_1(k) \le \mathbf{u}_{1,upper} \\ \mathbf{u}_{2,lower} \le \mathbf{u}_2(k) \le \mathbf{u}_{2,upper} \\ \mathbf{v}_{1,lower} \le \mathbf{v}_1(k) \le \mathbf{v}_{1,upper} \\ \mathbf{v}_{2,lower} \le \mathbf{v}_2(k) \le \mathbf{v}_{2,upper} \\ z_1 \ge 0 \\ z_2 \ge 0 \\ z_3 \ge 0 \end{cases}$$

where J_1,J_2, J_3 and J_4 are defined by (27), (32), (37), and (38) $u_{i,upper}$ and $u_{i,lower}$ are the upper and lower limits of the optimizing variable $u_i(k)$, respectively, $v_{i,lower}$ and $v_{i,upper}$ are the maximum and minimum value of the MVS velocity.

V. SIMULATIONS AND EXPERIMENTS

In this section, simulations and experiments are conducted to verify the proposed ACO algorithm.

5.1 Simulations

Simulations are conducted using MATLAB on the PC with Pentium Core2 processor. During the simulation, the target is designed to move in 3D space freely and two MVSs tried to observe and track the target actively and cooperatively using the proposed method.

The simulation parameters are as follows:

1. Envelope matrix of modeling error:
   $\mathbf{Q}$ = diag(0.0001, 0.0001, 0.0001);
2. Envelope matrix of measurement error:
   $\mathbf{R}$ = diag(0.001, 0.001, 0.001);
3. Initial envelope matrix of target state:
   $\mathbf{P}_0$ = diag (0.1, 0.1, 0.1).

The parameters of ACO algorithm are as follows:

1. Envelope matrix of modeling error:
   $\mathbf{Q}$ = diag(0.0001, 0.0001, 0.0001);
2. Envelope matrix of measurement error:
   $\mathbf{R}$ = diag(0.001, 0.001, 0.001);
3. Weight values: $w_1 = 0.01$, $w_2 = 0.0005$, $w_3 = 2.7$, $w_4 = 1$, $w_5 = 1.1$.

The target moves along the following trajectory during simulation,

$$\begin{cases} x_T(k) = 5 + 0.6k \\ y_T(k) = 5 + 0.6k \\ z_T(k) = 5 + 0.6\sin\left(\dfrac{\pi}{15}k\right) \end{cases} \quad (k < 5s)$$

$$\begin{cases} x_T(k) = 5 + 0.6k - 3 \\ y_T(k) = 5 + 0.6k \\ z_T(k) = 5 + 0.6\sin\left(\dfrac{\pi}{15}k\right) \end{cases} \quad (k \ge 5s)$$

The simulation results are shown in Fig. 7 and Fig. 8. Fig. 7 presents the observation uncertainty set, as well as trajectory of both the two MVSs and the target. From Fig. 7, it can be seen clearly that: (i) the two MVSs can keep tracking the target to obtain continuous observations; and (ii) the observation uncertainty ellipsoid becomes smaller and smaller and finally converges to a steady value. This indicates

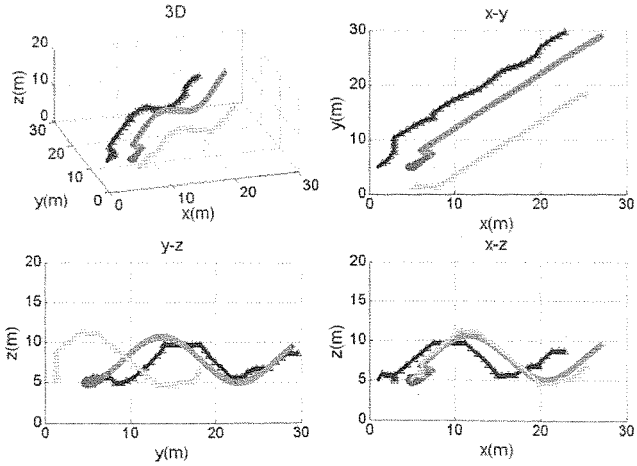


Fig. 7. Simulation result: trajectory and uncertainty set.
**the middle line is the trajectory of the target; the other two lines are the trajectory of the two MVSs, respectively; the blue ellipsoids are the observation result, i.e. the uncertainty sets of the target position at each time instant.*

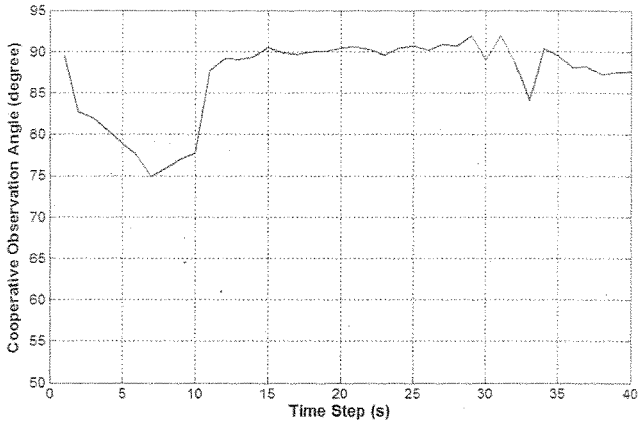


Fig. 8. Cooperative observation angle.

that the cooperative observation algorithm can be effectively used to estimate the position of the target.

Fig. 8 shows the COA during the simulation. From it we can see that by using the proposed algorithm, the two MVSs can modify their motion properly so as to optimize the COA even when there exists a sudden change of target trajectory at time $k = 5s$.

5.2 Experiments

In order to verify the proposed algorithm in real systems, a demonstration experiment is conducted on a multiple rotor-flying-robots (MRFRs) testbed, which is a newly-designed platform to verify the cooperation algorithms of multiple rotor-flying-robot and is introduced in detail in [19].

5.1.1 Testbed setup

The experimental platform is shown in Fig. 9, which is composed of three arms with one side fixed and the other side equipping a small rotor flying robot. In this experiment, the middle manually controlled robot (defined as R1) is used as a moving target whose motion state is unknown for the other robots, and the other two robots (defined as R2 and R3, respectively) are taken as two MVSs required to cooperatively observation R1. Furthermore, an LED and two optical filters (both with wavelength of 850 nm) are mounted on R1 and R2/R3, respectively, so that the detection and identification algorithm, which is not the main concern in this paper, can be simplified as far as possible.

It should be noted that the motion of the target (R1) as well as the two MVSs (R2, R3) can be directly measured using the encoders equipped in the rotary joints and the measurement result can be described as vertical angle α_1 and horizontal angle β_1 as shown in Fig. 9. Due to the motion limitation as shown in (41), each arm has only two DOFs. In this experiment, (41) is only used to obtain the real 3D position of the target, while in the estimation algorithm, the 3D motion equation discussed in Section II, (4) and (6) are still utilized to model the motion of the target.

$$\begin{cases} x_T = -l_1 \cos\alpha_1 \cos\beta_1 \\ y_T = -l_1 \cos\alpha_1 \sin\beta_1 \\ z_T = h_1 - l_1 \sin\alpha_1 \end{cases} \tag{41}$$

where the meaning of symbol l_1 and h_1 can be found in the right figure of Fig. 9.

5.2.2 Experiment result and analysis

In the experiment, the target is controlled to move along the following desired trajectory,

$$\begin{cases} x_T = -l_1 \cos\alpha_1 \cos\beta_1 \\ y_T = -l_1 \cos\alpha_1 \sin\beta_1 \\ z_T = h_1 - l_1 \sin\alpha_1 \end{cases} \tag{42}$$

where,

$$\alpha_1 = 0^o$$
$$\beta_1 = \begin{cases} \beta_I & k \le 20s \\ 5k & 20s < k \le 95s \\ 475 + 3k & 95s \le k \end{cases} \tag{43}$$

β_I is the initial horizontal angle of the target as shown in Fig. 9

Remark 5. The collision avoidance cost function, *i.e.* J_3 and J_4, are not considered in the experiment because the collision avoidance can be achieved by the arm that fixed the MVS in the tesebed.

Remark 6. The real position of the target can be obtained from (41) according to its encoder measurement value (α_1, β_1).

Remark 7. According to the pose parameters of the cameras, the COA equals to 90° when the horizon angle between the arm of R2 and that of R3 is 110°. That means, the optimal observation condition is

$$|\beta_2(k) - \beta_3(k)| = 110° \tag{44}$$

Remark 8. As discussed in Section III, the estimation result is the uncertainty set (expressed as an ellipsoid) that is described as (15). Then, define that,

$$F \triangleq [\mathbf{X} - \hat{\mathbf{x}}]^{\mathrm{T}} \mathbf{P}^{-1} [\mathbf{X} - \hat{\mathbf{x}}] \tag{45}$$

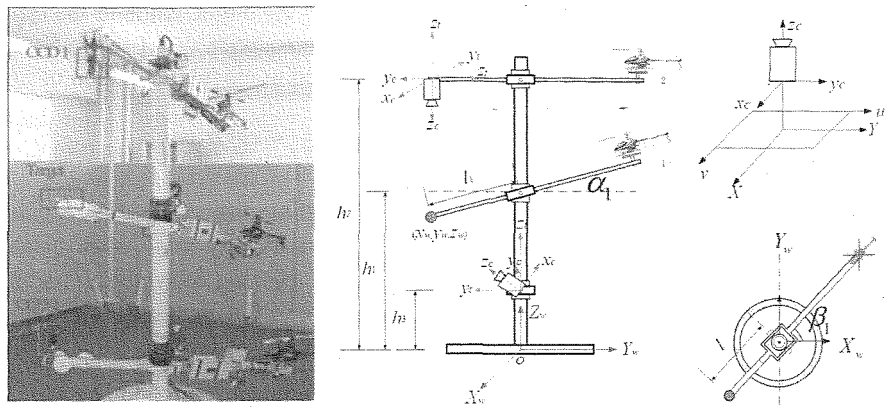


Fig. 9. MRFRs testbed and coordination system.

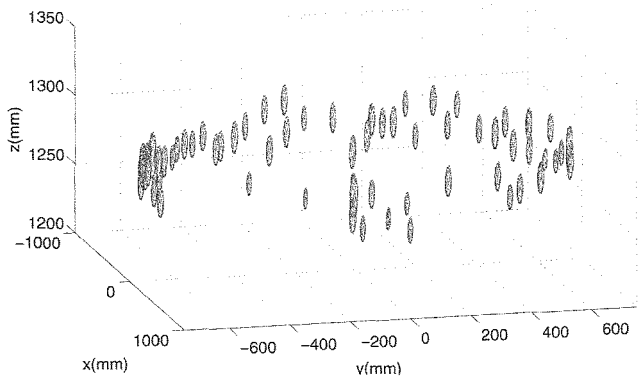


Fig. 10. Overview of the cooperative observation result. **The ellipsoids are the estimated uncertainty sets that contain the real state of the target.*

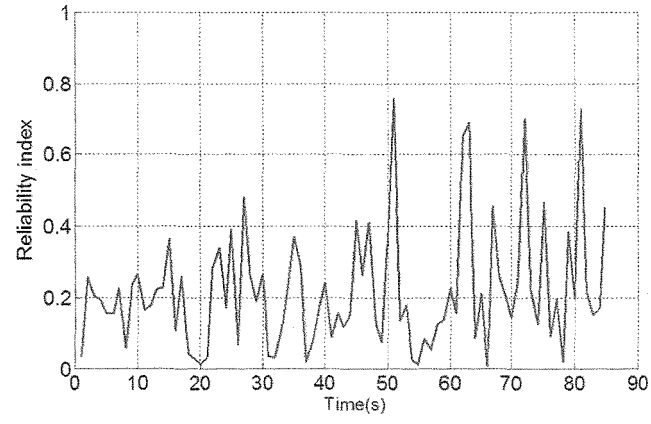


Fig. 11. Reliability index.

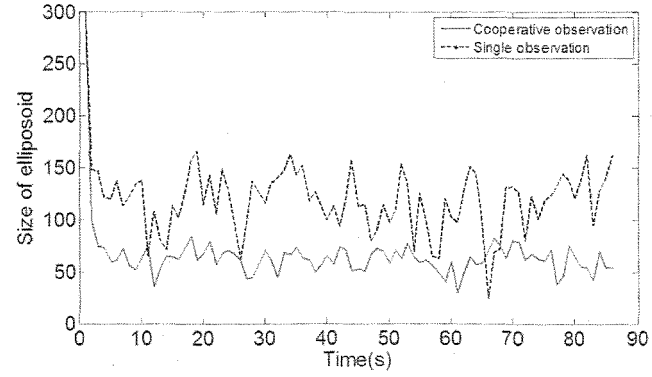


Fig. 12. Comparison between single and cooperative observation.

$0 \le F \le 1$ means X lie within the ellipsoid $E(\hat{\mathbf{x}}, \mathbf{P})$. In our experiment, we use this value to check the reliability of the proposed algorithm.

Experimental results are as shown in Figs 10–13. Fig. 10 is the overview of the cooperative observation procedure.

Fig. 11 gives out the value of F in (45) at each time instant. From this figure we can see that it is always in the interval (0, 1), that means all the true value point of the target $X_T(k)$ are lie within the uncertainty set.

From Fig. 10 and Fig. 11, the following conclusion can be obtained: the estimated states always contain the true state of the target. This verifies the reliability of the method. The observation errors are no more than 4 mm.

Since the observation result of the proposed algorithm is an uncertainty set, the size of it should also be an index to evaluate the observation performance. In the experiment, the size of uncertainty sets from single MVS and two cooperative MVSs are compared to show the validity of the proposed method. (Due to the motion limitation of the testbed, the target is actually a 2D motion. That means single vision sensor is theoretically enough to obtain 3D position of the target) Fig. 12 shows the results. From Fig. 12, the

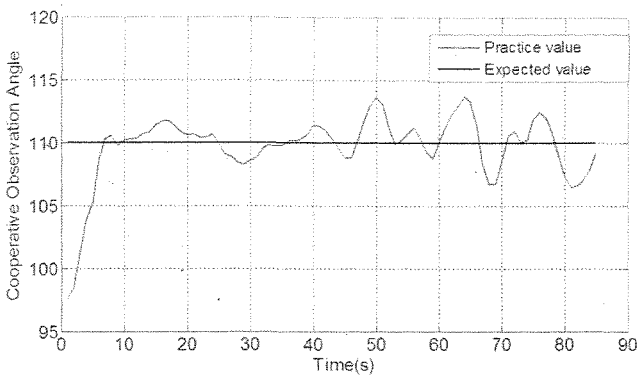


Fig. 13. Cooperative observation angle.

uncertainty size of cooperative observation algorithm is much smaller and smoother than that of the observation algorithm using only one sensor, which indicates the proposed method has better accuracy and robustness.

Fig. 13 shows the COA during the experiment. It can be seen that the COA can be kept much closer to the expected angle during the whole experiment.

VI. CONCLUSION AND FUTURE WORK

In this paper, we studied the problem of localizing and tracking moving targets in 3D space using only two dynamical monocular vision sensors, and proposed a new algorithm called ACO. The primary contributions of this paper are as follows:

1. Proposal of a framework for ACO in which data fusion and path planning are connected by the optimal observation condition, allowing improvement of fusion results through adjustment of relative orientations among movable sensors and their target.
2. Proposal of an ESMF-based data fusion algorithm and a relative velocity coordinates (RVCs) based path planning algorithm for implementation of the ACO algorithm.
3. Showing that the optimal observation condition of the two movable monocular vision sensors is the one at which the centerline of their uncertainty cylinders are perpendicular.
4. Experiments with multiple rotor-flying-robots (MRFRs) testbed demonstrated the validity and feasibility of the above methods in realistic 3D-VMTO conditions. It should be noted that the proposed methods are not limited to multiple monocular vision systems; they can also be applied in other multiple movable sensor systems, or even heterogeneous ones.

In future work, we plan to integrate camera dynamics with target dynamics during the estimation step to account for the influence of the uncertainty of monocular vision sensors. Furthermore, convergence of the proposed algorithm is also a problem to be researched.

REFERENCES

1. Papanikolopoulos, N., "Visual tracking of a moving target by a camera mounted on a robot: a combination of control and vision," *IEEE Trans. Robotics and Automation*, Vol. 9, No. 1, pp. 14–35 (1993).
2. Hyeun, M., "Feature-based covariance matching for a moving target in multi-robot following," *19th Mediterranean Conf. Control Autom., Corfu, Greece*, pp. 163–168 (2011).
3. Kamiyal, T., M. Doi, H. Mochiyama, and Y. Mori, "Visual tracking for references generated by a stochastic model," *Asian J. Control,* Vol. 5, No. 3, pp. 437–444 (2003).
4. Lin, C., F. Hsiao, and F. Hsiao, "Vision-based tracking and position estimation of moving targets for unmanned helicopter systems." *Asian J. Control*, Vol. 15, No. 5, pp. 1270–1283 (2013).
5. Illah, R., A. David, T. Carlo, and R. Michael, "Mobile robot obstacle avoidance via depth from focus," *Robot. Auton. Syst.*, Vol. 22, pp. 151–158 (1997).
6. Sabnis, A. and L. Vachhani, "Single image based depth estimation for robotic applications," *Recent Advances in Intelligent Computational Systems (RAICS), Trivandrum, India*, pp. 102–106 (2011).
7. Wang, H. and H. Chung, "Design and implementation of dynamic target tracking with stereo vision system," *International MultiConf. of Engineers and Computer Scientists*, Vol. I, pp. 102–106 (2011).
8. Kim, T., W. Bahn, C. Lee, T. Lee, T. Shaikh, and K. Kim, "Vision system for mobile robots for tracking moving targets, based on robot motion and stereo vision information," *IEEE/SICE International Symposium System Integratio, Kyoto, Japan*, pp. 634–639 (2011).
9. Stefanik, K., J. Gassaway, K. Kochersberger, and A. Abbott, "UAV-based stereo vision for rapid aerial terrain mapping," *GIScience and Remote Sensing*, Vol. 48, pp. 24–49 (2011).
10. Stroupe, W., C. Martin, and T. Balch, "Distributed sensor fusion for object position estimation by multi-robot systems," *IEEE Conf. on Robot and Automation, Seoul, South Korea*, pp. 1092–1098 (2001).
11. Sola, J., A. Monin, M. Devy, and T. Vidal-Calleja, "Fusing monocular information in multicamera SLAM," *IEEE Trans. Robot.*, Vol. 24, No. 5, pp. 958–968 (2008).
12. Montesano, L., J. Gaspar, J. Santos-Victor, and L. Montano, "Cooperative localization by fusing vision-based bearing measurements and motion," *IEEE/RSJ*

Conf. Intelligent Robots and Systems, Alberta, Canada, pp. 2333–2338 (2005).
13. Zhou, K. and S. Roumeliotis, "Multirobot active target tracking with combinations of relative observations," *IEEE Trans. Robot.*, Vol. 27, No. 4, pp. 678–698 (2012).
14. Zhou, B., J. Han, and G. Liu, "A UD factorization-based nonlinear adaptive set-membership filter for ellipsoidal estimation," *Int. J. Robust and Nonlinear Control*, Vol. 18, No. 16, pp. 1513–1531 (2008).
15. Gu, F., Y. He, J. Han, and Y. Wang, "ESMF based multiple UAVs active cooperative observation method in relative velocity coordinates," *International Conference on Decision and Control, Shanghai, China*, pp. 3008–3013 (2009).
16. Scholte, E. and M. Campbell, "A nonlinear set-membership filter for on-line applications," *Int. J. Robust and Nonlinear Control*, Vol. 13, No. 15, pp. 1337–1358 (2003).
17. Ros, L., A. Sabater, and F. Thomas, "An ellipsoidal calculus based on propagation and fusion," *IEEE Trans. on Systems. Man Cybernetics. Part B – Cybern.*, Vol. 32, No. 4, pp. 430–442 (2002).
18. Zu, D., J. Han, and D. Tan, "MILP-based optimal trajectory generation in relative velocity coordinates," *46th IEEE Conf. on Decision and Control, New Orleans, USA*, pp. 1975–1980 (2007).
19. Wang, Z., F. Gu, Y. He, J. Han, and Y. Wang, "Design and implementation of multiple-rotorcraft-flying-robot testbed," *IEEE Conf. on Robotics and Biomimetics*, Phuket, Thailand, pp. 1291–1296 (2011).

Feng Gu is Research Assistant at the State Key Laboratory of Robotics, Shenyang Institute of Automation, Chinese Academy of Sciences. He received B.S. and M.S. degree in system engineering from Nanjing University of Science and Technology in 2005 and 2007 respectively, and Ph.D. degree in pattern recognition and intelligent systems from Shenyang Institute of Automation, CAS in 2011. Subsequently, he worked at the same institute until now. His current researches focus on the set theory based estimation and control, cooperation of multiple robots and also include the researches of unmanned surface vehicle.

Yuqing He is Associate Professor at the State Key Laboratory of Robotics, Shenyang Institute of Automation, Chinese Academy of Sciences. He obtained the bachelor degree at North-eastern University at Qinhuangdao in 2002. Subsequently, he entered into the Shenyang Institute of Automation and obtained the Ph.D. degree in the year of 2008. After that, he worked at the same institute until now. In 2012, he was a visiting researcher at Institute for automatic control theory in Technique University of Dresden. Currently, his researches focus on the Nonlinear Estimation and Control of Mechatronic Systems, especially set- and statistics- based filter, acceleration feedback control, nonlinear predictive control and Control Lyapunov functions based control. Also, he is involved into the researches of mobile robotics.

Jianda Han received his Ph.D. from Harbin Institute of Technology in 1998. Currently he is Professor and Vice-director of the State Key Laboratory of Robotics, Shenyang Institute of Automation, Chinese Academy of Sciences. His research interests include nonlinear estimation and control, robotics, and mechatronic systems.