

Received October 23, 2021, accepted December 23, 2021, date of publication January 4, 2022, date of current version January 13, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3140331

Vision-Based 3D Aerial Target Detection and Tracking for Maneuver Decision in Close-Range Air Combat

LEISHENG ZHONG¹, LEIMING ZHAO¹, CHENCONG DING¹, XUESHI GE¹, JIALIN CHEN¹,
YU ZHANG², AND LI ZHANG³

¹Naval Research Institute, Shanghai 200436, China

²School of Astronautics, Beihang University, Beijing 102206, China

³Department of Electronic Engineering, Tsinghua University, Beijing 100084, China

Corresponding author: Leiming Zhao (cmzlm@sina.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 62132002 and Grant 61871248.

ABSTRACT Automatic maneuver decision in close-range air combat depends on the situation awareness of the 3D aerial space. Optimal decision could only be made when the 3D state (e.g. 3D position, orientation and velocity) of the target aircraft is accurately provided. Together with the state of the aircraft in our side, optimal maneuver decision could be made by maximizing the situation advantage or utilizing deep reinforcement learning. On the other hand, vision-based 3D sensing methods are ideal for acquiring the 3D state of the target aircraft in close-range air combat, since radar and other sensors work badly in such short range. In this paper, we propose a novel pipeline for vision-based maneuver decision in close-range air combat. The proposed pipeline contains three main modules: 3D target detection based on Augmented Autoencoder, 3D target tracking based on segmentation and optimization, and maneuver decision based on advantage maximization and Deep Q Networks (DQN). The proposed method effectively handles the difficulties in air combat environment, such as fast movement, occlusion from cloud, etc. Experiments demonstrate that our method could robustly detect and track the target aircraft in complex environment, which provides strong priors for maneuver decision and helps to significantly improve the winning rate of short-range air combat.

INDEX TERMS 3D target detection and tracking, reinforcement learning, maneuver decision, air combat.

I. INTRODUCTION

Automatic maneuver decision in close-range air combat has become a popular research topic in recent years [1]–[5]. Vision-based sensors play an important role in the state assessment of close-range air combat. Airborne radar and laser sensors could only measure the distance between the target aircraft and our aircraft, and are incapable of measuring the full 6-dof pose of the target aircraft, especially the orientation (heading, pitch and roll). Acquiring the full 6-dof pose of the enemy aircraft is essential for estimating its moving trend, after which optimal maneuver decision is possible to be made accordingly. As a result, 3D vision-based sensing from airborne cameras is often used for situation awareness in close-range air combat. In order to robustly and accurately obtain the 3D position, orientation and velocity of the target

aircraft, a vision-based 3D target detection module and a vision-based 3D target tracking module need to be developed. After that, a maneuver decision module could be learned based on the state of both sides. Most of the previous air combat maneuver decision methods assume that the state of the target aircraft (i.e. the position, orientation and velocity) is already known and do not consider the problem of state estimation (or situation awareness, such as target detection and tracking). However, we believe situation awareness is crucial in building an effective maneuver decision model. As a result, we focus on the front end, and try to solve the problem of 3D aerial target detection and tracking. The main pipeline for vision-based maneuver decision in close-range air combat is illustrated in Fig. 1. In this paper, we propose a new pipeline, which features three novel modules: 3D target detection, 3D target tracking, and maneuver decision. These three modules are specially tailored for the characteristics of short-range air combat, which is detailed as follows.

The associate editor coordinating the review of this manuscript and approving it for publication was Rosario Pecora^{1b}.

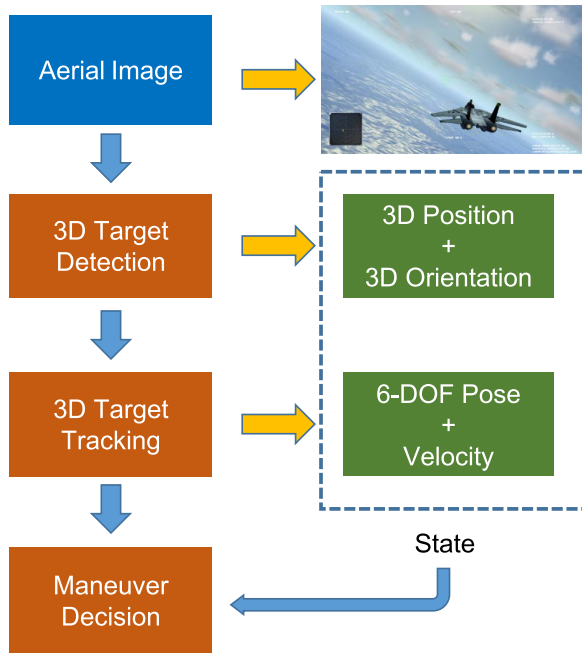


FIGURE 1. The proposed vision-based maneuver decision method.

Although 3D target detection and tracking have been widely studied and a lot of methods have been proposed in recent years [6]–[13], very few of them have considered the specific scenario of close-range air combat. Firstly, the environment in upper air is very complicated. For example, the visibility and dynamic range are very different from the experimental environment (such as indoor environment which the previous methods and datasets usually tested or recorded in). Moreover, the cloud often occludes the target, making it only partly visible or even hardly visible. Therefore, the previous methods would frequently fail in such complex environment. In this paper, we propose to detect the target in upper air using an Augmented Autoencoder, which is robust to different background and occlusion. Once detected, the target aircraft is then tracked in the subsequent frames using a robust segmentation-based 3D tracker, which is especially robust to cloud occlusions. Secondly, in close-range air combat, the target aircraft moves very fast, and the scale changes dramatically in short time period. This is very challenging for 3D target tracking algorithms. To handle this issue, we propose to iteratively optimize the tracker, which improves the convergence rate of the tracker and makes it more robust to fast movement and scale change. Thirdly, in air combat tasks, we usually need to fight with different kinds of enemy aircrafts. In previous methods, the deep detection and tracking models need to be trained individually for each aircraft, which makes it infeasible in real practice. In this paper, we make use of the so-called ‘separated latent rotation space’ and develop an Augmented Autoencoder capable of detecting different kinds of target aircraft in a single model. The tracking module is also universally trained and works for multiple aircraft types. Finally, based on the 3D target detection and tracking results, full information of the target is

obtained, and optimal maneuver decision could be made by maximizing the advantage score or utilizing deep reinforcement learning.

The main contributions of this paper are:

- 1) We propose a novel maneuver decision pipeline in close-range air combat based only on vision sensors. The proposed pipeline could effectively and efficiently make optimal maneuver decisions and helps to improve the winning rate.
- 2) We propose a robust vision-based 3D aerial target detection method and a 3D aerial target tracking method which are able to handle complex upper air environment and provide accurate state information of the target aircraft in close-range air combat.

II. RELATED WORK

In this section, we briefly introduce the related researches on vision-based 3D target detection, vision-based 3D target tracking, and maneuver decision in close-range air combat.

A. VISION-BASED 3D TARGET DETECTION

Vision-based 3D target detection (also known as 6-dof object pose estimation) aims to estimate the 6-dof pose of a known object from a single image. This task is very different from 2D target detection, in which the target is only detected in a 2D image, instead in the 3D space. Due to the high dimensionality, 3D target detection is much more difficult than 2D detection. In early days, monocular 3D target detection is usually achieved through template matching [14]–[16]. In these early methods, template images of the target object in various different poses are recorded or generated (rendered) to build a template image library. At test time, the input image is compared with all of the template images in the library using fast comparing algorithms (usually in gradient domain). Recently, deep learning-based methods have dominated the research for vision-based 3D target detection [6]–[11]. These methods could be classified into three main categories. Methods in the first category [6]–[8] train a deep neural network to directly regress the 6-dof pose parameters, which seems simple in spirit but relatively difficult to train. Methods in the second category [9]–[11] take another route by first predicting the keypoint locations in the image and then estimating the 6-dof pose using the PnP algorithm. Methods in the third category [17], [18] employ the Augmented Autoencoder to train a latent code for different poses of the target. The autoencoder-based architecture is relatively easy to train with only augmented synthetic training data. With the help of data augmentation and adversarial training, these methods are able to solve the domain gap and generalize to real test images. Currently, the deep learning-based methods are the state-of-art methods in the field of 3D target detection. Therefore, in this paper, we utilize an Augmented Autoencoder-based method to detect the target aircraft in close-range air combat, which is capable of robustly detecting different kinds of target aircraft in a single model.

B. VISION-BASED 3D TARGET TRACKING

Vision-based 3D target tracking aims to estimate the 6-dof pose of a known object in each frame of a video sequence by utilizing temporal information. For tracking tasks, the 6-dof pose of the first frame is given as the prior. This prior is usually provided by a 3D target detection algorithm introduced in Section II-A. When tracking is lost due to some difficult factors, the tracker needs to be reset by re-detecting the target in the following frames. In early years, people use feature-based and edge-based methods to accomplish the task of 3D target tracking [19]. However, feature-based methods only work for well-textured targets and could not work for textureless objects. Also, edge-based methods often struggle with background clutters and only work well for simple objects in simple background. On the other hand, region-based methods have been popular for monocular 3D target tracking in recent years [20]–[24]. These methods are built upon a statistical formulation aiming to maximize the discrimination of foreground and background regions. The earliest region-based method is PWP3D [20]. After that, a lot of region-based methods have been proposed based on PWP3D [21]–[25], and each of them seeks to tackle some of the problems in the original PWP3D algorithm. Specifically, a recent method [25] combines region-based method with learning-based video object segmentation, which achieves very good performance in cases with heavy occlusions. In this paper, we borrow the idea of [25] and propose to track the target aircraft in 3D space using segmentation and optimization. We have made several changes in order to make it work better for fast aircraft movement and upper air environment with cloud occlusions.

C. MANEUVER DECISION IN CLOSE-RANGE AIR COMBAT

Maneuver decision methods aim to automatically make the best decision of movement for the aircraft. Making the best maneuver decision is more crucial for the survival and winning of the aircraft in close-range air combat than in middle-range or long-range air combat, since it is much more time-sensitive in close-range air combat. Previously, maneuver decision is usually made by rules, which lacks generalization ability. In recent years, researchers have been trying to utilize reinforcement learning to tackle the problem of maneuver decision. Yang *et al.* propose a deep reinforcement learning-based maneuver decision model for UAV in short-range air combat [2]. Their method mainly includes the aircraft motion model, one-to-one short-range air combat evaluation model and the maneuver decision model based on Deep Q Network (DQN). Zhang *et al.* propose three efficient training techniques for a multi-agent combat problem in UAV combat scenario in [3]. The proposed method is able to train multiple agents simultaneously using multi-agent deep Q-learning and multi-agent deep deterministic policy gradient algorithms, but only 2D cases are discussed in their paper. Wang *et al.* propose to improve maneuver strategy in air combat by alternating freeze games with a deep reinforcement learning algorithm [4]. They focus on

the training stability of the self-play training problem in deep reinforcement learning. Kong *et al.* also propose an UAV autonomous aerial combat maneuver strategy generation method based on state-adversarial deep deterministic policy gradient and inverse reinforcement learning [5]. All of the above methods assume that the state of the target aircraft (i.e. the position, orientation and velocity) is already known and do not consider the problem of state estimation. However, we believe situation awareness is crucial in building an effective maneuver decision model, so we mainly focus on the problem of 3D aerial target detection and tracking. After that, we also build a simple but effective maneuver decision model to prove the advantage of our detection and tracking algorithms and the effectiveness of the proposed whole pipeline.

III. METHOD

In this section, we present the vision-based maneuver decision pipeline in detail. We first introduce the main workflow of our pipeline in Section III-A. After that, we present the proposed 3D aerial target detection module, the 3D aerial target tracking module, and the maneuver decision module in detail in Section III-B, III-C and III-D respectively.

A. OVERVIEW

The overview of the proposed pipeline is illustrated in Fig. 2. To accomplish vision-based automatic maneuver decision, we need to make use of the video stream from airborne cameras. When the target aircraft appears in the sight of view, it is detected by the proposed 3D target detection module. The detector estimates the 3D position and 3D orientation of the target aircraft, and provides initial 6-dof pose of the aircraft to the 3D target tracking module. With the initial pose estimate, the 3D target tracker then estimates the 6-dof pose of the aircraft in each of the subsequent frames. When tracking is lost, the tracker needs to be re-initialized by the detector. Apart from 6-dof poses, the 3D tracker could also provide 3D velocity of the target aircraft through temporal differential and filtering. After that, full state of the target aircraft, including the position, orientation and velocity, is provided to the maneuver decision model. We develop a simple one-step maneuver decision model and a second maneuver decision model based on Deep Q Networks (DQN). The one-step model is realized by maximizing the advantage score in the next single step. The DQN-based model is a simple neural network which takes the state vector of both sides and outputs the Q value of each possible maneuver (action). After training, optimal maneuver could be decided by choosing the action with the highest Q value.

B. VISION-BASED 3D AERIAL TARGET DETECTION

The main idea of the proposed 3D aerial target detection module is to first localize the target in the 2D image, and then estimate its 6-dof poses using an Augmented Autoencoder. The first step is accomplished using an on-the-shelf fast 2D detector, YOLO v3 [26]. In the following, we mainly

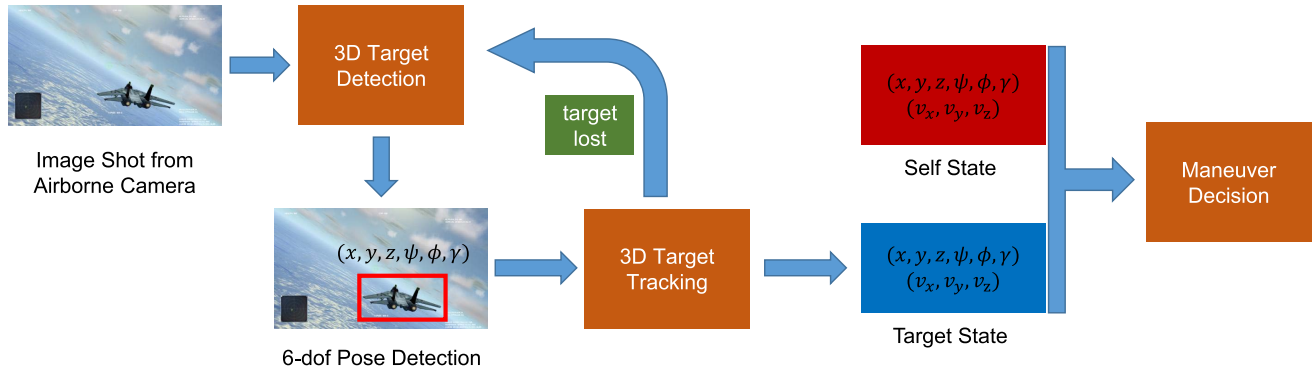


FIGURE 2. Overview of the proposed pipeline.

introduce the second step: 6-dof pose estimation using Augmented Autoencoder.

The network architecture is illustrated in Fig. 3. By using an autoencoder, we want to learn the latent feature space of the 3D rotation space. In order to make the model work for multiple aircraft types, we propose to learn a ‘separated latent rotation space’, which encodes the rotation of different aircrafts into different separated regions of the latent space. To achieve this, the training samples are prepared in a one-to-one manner. The prior distribution is selected as a mixture of Gaussians distribution in order to shape the distribution of the latent code into N separated regions. Also, the encoder plays a cooperative game with the classifier for better separation. As a result, the trained network could simultaneously estimate the class and orientation of the input image.

1) ONE-TO-ONE CONDITIONAL AUTOENCODER

Here a one-to-one mapping scheme is adopted for learning the separated rotation space. For an input image of object i ($i = 1, 2, \dots, N$) in rotation j , only itself is required to be reconstructed by the generator with the label $y = i$.

The reconstruction target can be written as:

$$\hat{x}_{ij} = G(z_{ij}, y) = G(E(x_{ij}), y = i) \quad (1)$$

Since there are no extra constraints on the latent code z , the learned rotation representation is not shared among objects. Instead, they are likely to be naturally separated and clustered by classes because of similar features within each object class. The reconstruction objective is:

$$\min_{E, G} \mathcal{L}_R(E, G) = \min_{E, G} L_2(x, G(E(x), y)) \quad (2)$$

2) DISTRIBUTION DISCRIMINATOR ON z

To better shape the distribution of the latent code of different object classes, we impose a mixture of Gaussians distribution on the latent code z . We incorporate the class label as the input of D_z . The one-hot class label could act as a switch that selects the corresponding decision boundary of the discriminative network given the class label y . Here the positive samples for D_z is randomly drawn from one mixture component of

the prior distribution $p(z)$ according to a given class label y : $z^* \sim p(z|y)$. The latent code generated by the encoder and the corresponding class label y are also provided to D_z as negative samples. The encoder E and discriminator D_z then play a min-max game with the following objective:

$$\min_E \max_{D_z} \mathcal{L}_D(E, D_z) \quad (3)$$

where

$$\mathcal{L}_D(E, D_z) = \mathbb{E}_{z^* \sim p(z|y)} [\log D_z(z^*, y)] + \mathbb{E}_{x \sim p_{data}(x)} [\log (1 - D_z(E(x), y))] \quad (4)$$

3) CLASSIFIER ON z

Since we want to completely separate different objects in the latent space, the encoder plays a cooperative game with the classifier C_z . E and C_z work together for better classification performance:

$$\begin{aligned} & \max_{E, C_z} \mathcal{L}_C(E, C_z) \\ & = \max_{E, C_z} \mathbb{E}_{x \sim p_{data}(x), y} \left[\sum_c \mathbb{I}(c = y) \log(C_z(c|E(x))) \right] \end{aligned} \quad (5)$$

4) THE FINAL OBJECTIVE FUNCTION

Combining the above objectives gives the final objective for learning separated latent rotation space:

$$\min_{E, G} \max_{D_z} [\mathcal{L}_R(E, G) + \lambda \mathcal{L}_D(E, D_z)] + \max_{E, C_z} \gamma \mathcal{L}_C(E, C_z) \quad (6)$$

The whole network is trained on each mini-batch with SGD sequentially in three phases:

(1) In the reconstruction phase, the encoder and the generator are updated by minimizing $\mathcal{L}_R(E, G)$.

(2) In the regularization phase, the distribution discriminator D_z is first updated by minimizing $-\mathcal{L}_D(E, D_z)$ to tell apart true samples (generated using prior distribution of z) from generated samples (the latent code computed by the encoder) with the given class label y . Then the encoder E is updated by minimizing $\mathcal{L}_D(E, D_z)$ to confuse the discriminator network.

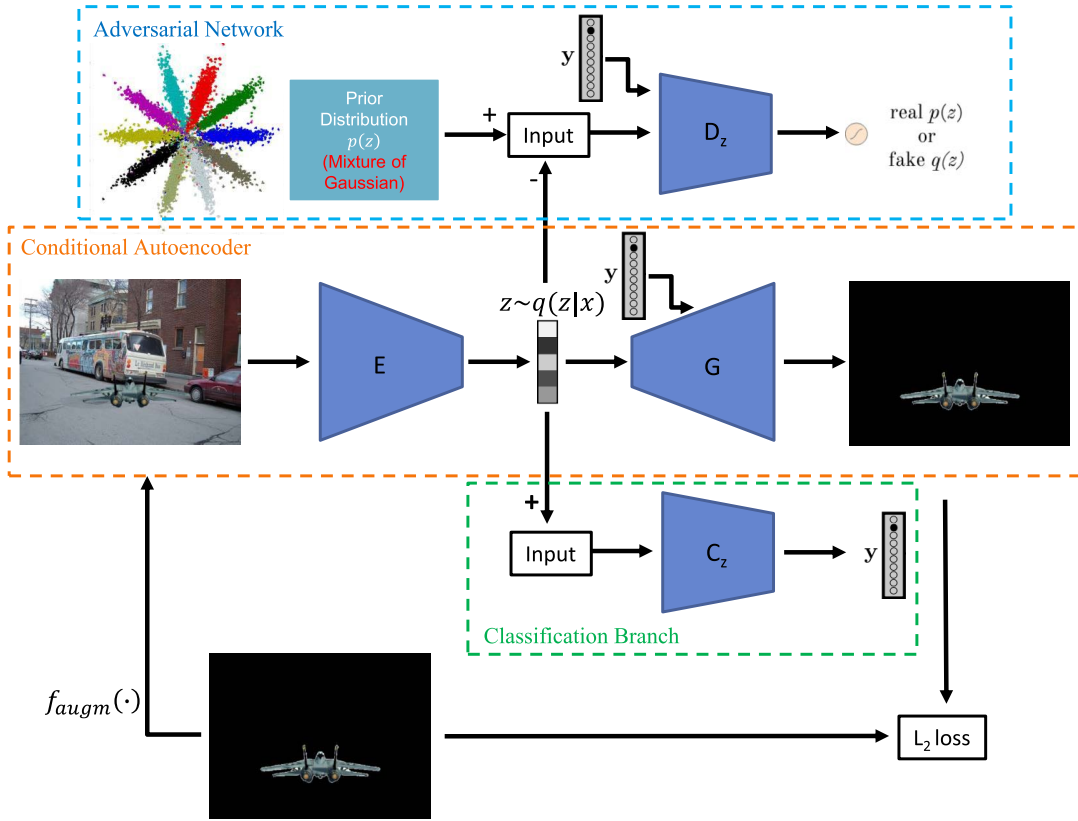


FIGURE 3. Vision-based 3D aerial target detection based on augmented autoencoder.

(3) In the classification phase, the encoder E and the classifier C_z are cooperatively updated by minimizing $-\mathcal{L}_C(E, C_z)$.

5) EXTENDING TO 6D TARGET DETECTION

After training the Augmented Autoencoder, the rotation of the target could be obtained through searching the latent code of the test image in the codebook [17]. If we want to detect N kinds of aircrafts, N codebooks need to be built by rendering all kinds of the aircrafts in different rotation and feeding them into the encoder to obtain the latent codes. For the i -th kind of aircraft, the latent code of the test image z_{test} is compared to all of the latent codes in codebook i :

$$\cos_j^i = \frac{z_j^i z_{test}}{\|z_j^i\| \|z_{test}\|} \quad (7)$$

The rotation of the test image is determined through KNN search:

$$R_{test} = R_{jmax} \quad (8)$$

where

$$jmax = \max_j \cos_j^i \quad (9)$$

After that, the translation part $t = (t_x, t_y, t_z)$ could be determined as follows and the detection is extended to 6D:

$$t_z = t_{syn-z} \times \frac{l_{syn}}{l_{test}} \times \frac{f_{test}}{f_{syn}} \quad (10)$$

$$t_x = \frac{t_z}{f_{test}} \times (x_{test} - x_c) \quad (11)$$

$$t_y = \frac{t_z}{f_{test}} \times (y_{test} - y_c) \quad (12)$$

where f_{syn} and f_{test} are the focal lengths of the synthetic camera and the test camera; l_{syn} and l_{test} are the diagonal distance of the 2D rectangle of the synthetic image and the detected 2D rectangle of the test image; t_{syn-z} is the rendering distance for the synthetic images; x_{test} , y_{test} , x_c , y_c are the 2D coordinates of the detection center and the coordinates of the optical center.

C. VISION-BASED 3D AERIAL TARGET TRACKING

Given the initial pose estimated by the 3D detector, we then track the target in the following frames. Once the tracker is lost, we reset the pose by re-detecting the target, and then re-start tracking. The proposed vision-based 3D aerial target tracking pipeline is shown in Fig. 4. We use a similar tracking pipeline as our previous work [25], but some modifications have been made according to the characteristics of the aerial tracking environment. The proposed tracking pipeline contains two steps: video object segmentation and region-based 6-dof pose estimation.

1) VIDEO OBJECT SEGMENTATION NETWORK

Firstly, before estimating the 6-dof pose of the target in the current frame, we try to segment the target in the image using a video-object segmentation network. Specifically, we use

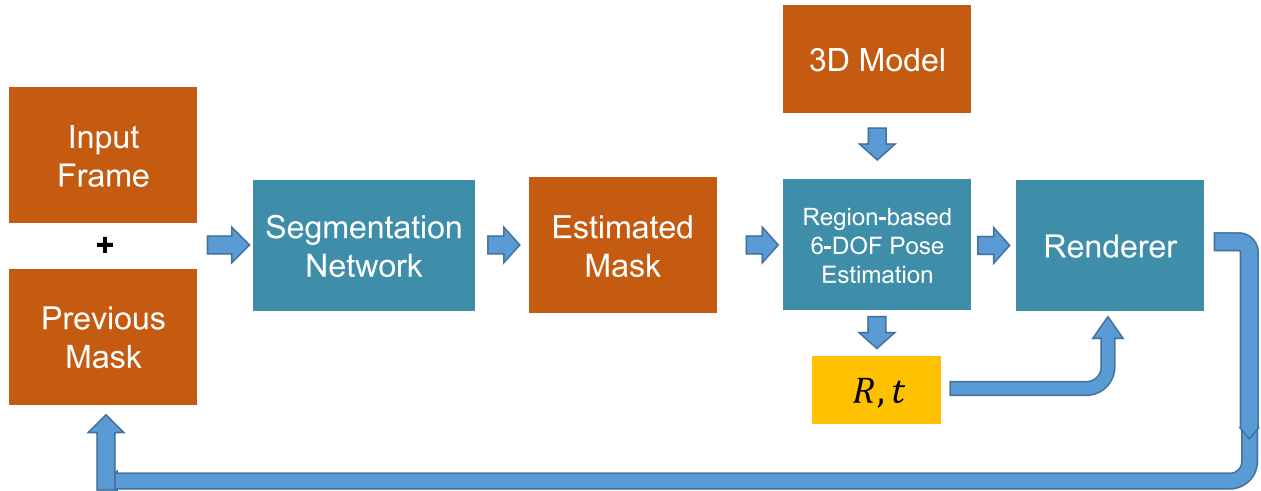


FIGURE 4. Vision-based 3D aerial target tracking pipeline.

the DeepLab v3+ [27] network, and modify the input from 3 channels to 4 channels (current frame RGB channels + the previous segmentation mask). By including the previous mask as input, the network could obtain some clues about the approximate current position of the target as temporal prior. On the other hand, the previous mask helps to handle occlusions, thus the segmentation network could predict the full mask of the target in the current frame, even in heavy occlusion cases. The training details for the video object segmentation network could be found in [27].

2) REGION-BASED 6-DOF POSE ESTIMATION

Next, we estimate the 6-dof pose of the target based on the segmentation result using non-linear optimization. We follow the state-of-art region-based 3D pose tracking formulation:

$$E(\mathbf{p}) = - \sum_{\mathbf{x} \in \Omega} \log[H_e(\Phi(\mathbf{x}(\mathbf{p})))P_f(\mathbf{x}) + (1 - H_e(\Phi(\mathbf{x}(\mathbf{p}))))P_b(\mathbf{x})] \quad (13)$$

where $\mathbf{p} = (\omega_1, \omega_2, \omega_3, t_1, t_2, t_3)^T \in \mathbb{R}^6$ is the 6-dof pose vector in Lie algebra representation. Ω is the image region, $\mathbf{x} = (x, y)^T \in \Omega$ are the pixel coordinates in the image. $\Phi(\mathbf{x})$ is the level-set embedding function (the signed distance function):

$$\Phi(\mathbf{x}) = \begin{cases} -d(\mathbf{x}, C), & \forall \mathbf{x} \in \Omega_f \\ d(\mathbf{x}, C), & \forall \mathbf{x} \in \Omega_b \end{cases} \quad (14)$$

where C is the contour, Ω_f and Ω_b are the foreground and background regions, $d(\mathbf{x}, C)$ calculates the nearest distance from \mathbf{x} to contour C . H_e is the smoothed Heaviside step function. $P_f(\mathbf{x})$ and $P_b(\mathbf{x})$ are the posterior distributions of foreground and background pixels, which are calculated from the soft segmentation mask provided by the segmentation network in our case.

The aim of region-based pose tracking is to minimize the energy function (13). Similar to previous works [22], [24],

we use a Gauss-Newton method by re-writing the energy function as a non-linear re-weighted least squares problem:

$$E(\mathbf{p}) = \frac{1}{2} \sum_{\mathbf{x} \in \Omega} \psi(\mathbf{x}) F^2(\mathbf{x}, \mathbf{p}) \quad (15)$$

where

$$F(\mathbf{x}, \mathbf{p}) = - \log[H_e(\Phi(\mathbf{x}(\mathbf{p})))P_f(\mathbf{x}) + (1 - H_e(\Phi(\mathbf{x}(\mathbf{p}))))P_b(\mathbf{x})] \quad (16)$$

and

$$\psi(\mathbf{x}) = \frac{1}{F(\mathbf{x}, \mathbf{p})} \quad (17)$$

Then the non-linear optimization problem could be solved iteratively by fixing and alternatingly updating the weights $\psi(\mathbf{x})$.

After pose tracking, the velocity is calculated using temporal differential with Gaussian filtering (smoothing).

D. MANEUVER DECISION IN CLOSE-RANGE AIR COMBAT

After 3D detection and tracking, the state of the target aircraft is obtained. Together with the state of our aircraft, the next best maneuver of our aircraft could be decided by maximizing the advantage of our aircraft compared to the target aircraft. In this section, we first introduce the aircraft motion model, then we present two maneuver decision methods based on the estimated state. Both two methods are simple yet effective. Our method could achieve similar results compared to the state-of-art methods with much simpler configurations (such as the dimension of state vector, action space, and training strategy), which proves the effectiveness of our detection and tracking modules.

1) AIRCRAFT MOTION MODEL

Similar to previous works [2], [3], we use a simplified aircraft motion model as shown in Fig. 5. We mainly consider

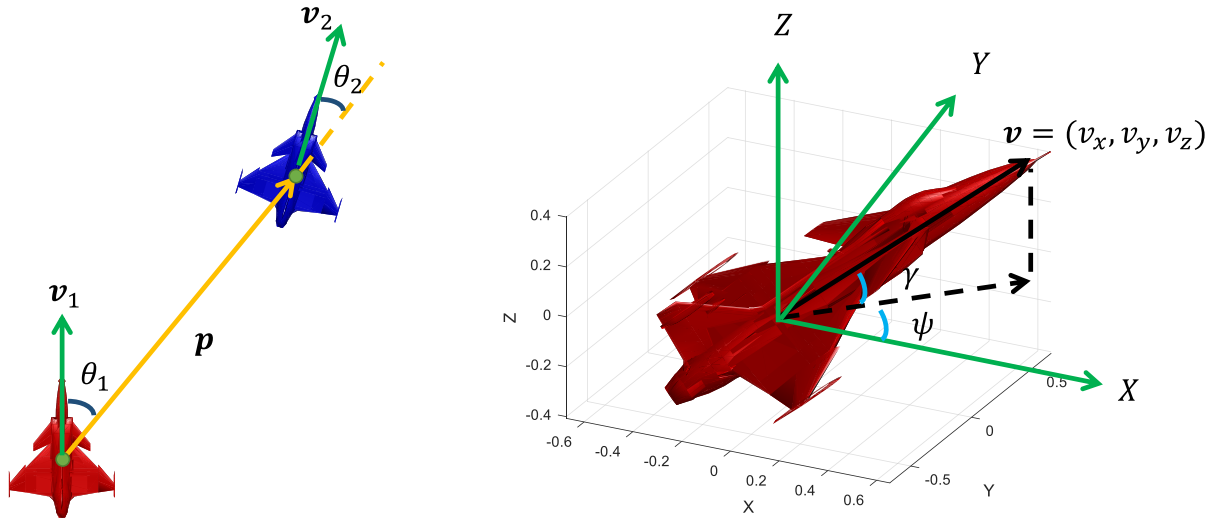


FIGURE 5. Aircraft motion model.

the position, orientation and velocity of the two aircrafts. We assume that the velocity direction coincides with the aircraft body axis. The velocity direction $\mathbf{v} = (v_x, v_y, v_z)$ is decided by the heading angle ψ and the pitch angle γ , and the roll angle ϕ is ignored ($\phi = 0$).

The relative position vector of two aircrafts is depicted as \mathbf{p} . The angle between the velocity of our aircraft \mathbf{v}_1 and \mathbf{p} is θ_1 , which is called the Antenna Train Angle (ATA). The angle between the velocity of target aircraft \mathbf{v}_2 and \mathbf{p} is θ_2 , which is called the Aspect Angle (AA).

The motion model can be described as:

$$\begin{cases} \dot{x} = v \cos \gamma \cos \psi \\ \dot{y} = v \cos \gamma \sin \psi \\ \dot{z} = v \sin \gamma \end{cases} \quad (18)$$

where $\dot{x}, \dot{y}, \dot{z}$ are the velocity components in three coordinate axis directions.

Here we ignore the dynamic model of the aircraft and only consider the motion model. To make the model more realistic, we have set limitation to the velocity change rate in each coordinate axis. In general, we want to find the best next velocity direction and magnitude in order to obtain the largest advantage towards the target aircraft, which is discussed in detail in the following.

2) ONE-STEP OPTIMAL MANEUVER DECISION

First, we present a simple one-step maneuver decision method to quickly test the effectiveness of the proposed 3D aerial target detection and tracking modules. The idea is to maximize the advantage score in the next step. Since only one step is considered, this maneuver decision algorithm could be seen as a greedy algorithm.

In order to win a close-range air combat, the key is to enter the tail zone of the target aircraft and avoid letting the target aircraft to enter our tail zone. To achieve this goal, both the

relative angle and relative distance between the two aircrafts need to be considered.

For the angle advantage, as shown in Fig. 5, the smaller θ_1 and θ_2 are, the larger advantage we have towards the target aircraft. The angle advantage score is defined as:

$$S_{angle} = \frac{1}{2} \left(\frac{\mathbf{v}_1 \mathbf{p}}{\|\mathbf{v}_1\| \|\mathbf{p}\|} + \frac{\mathbf{v}_2 \mathbf{p}}{\|\mathbf{v}_2\| \|\mathbf{p}\|} \right) \quad (19)$$

For the distance advantage, we want to enter the attack zone, which is a region defined by the maximum attack distance d_{max} and the minimum attack distance d_{min} . When the distance $\|\mathbf{p}\|$ is larger than d_{max} , the next movement should narrow the distance $\|\mathbf{p}\|$ as far as possible. On the contrary, when the distance $\|\mathbf{p}\|$ is smaller than d_{min} , the next movement should enlarge the distance $\|\mathbf{p}\|$. The distance advantage score is define as:

$$S_{dist} = \begin{cases} \frac{\|\mathbf{p}_{t-1}\| - \|\mathbf{p}_t\|}{2v_{max}}; & \|\mathbf{p}_t\| > d_{max}, \|\mathbf{p}_t\| < \|\mathbf{p}_{t-1}\| \\ 1 - \frac{\|\mathbf{p}_t\| - \|\mathbf{p}_{t-1}\|}{2v_{max}}; & \|\mathbf{p}_t\| > d_{max}, \|\mathbf{p}_t\| \geq \|\mathbf{p}_{t-1}\| \\ 1 - \frac{\|\mathbf{p}_{t-1}\| - \|\mathbf{p}_t\|}{2v_{max}}; & \|\mathbf{p}_t\| < d_{min}, \|\mathbf{p}_t\| < \|\mathbf{p}_{t-1}\| \\ \frac{\|\mathbf{p}_t\| - \|\mathbf{p}_{t-1}\|}{2v_{max}}; & \|\mathbf{p}_t\| < d_{min}, \|\mathbf{p}_t\| \geq \|\mathbf{p}_{t-1}\| \\ 1; & d_{min} \leq \|\mathbf{p}_t\| \leq d_{max} \end{cases} \quad (20)$$

where \mathbf{p}_t and \mathbf{p}_{t-1} are the distance between two aircrafts in the current time stamp and the previous time stamp, v_{max} is the maximum velocity.

Comining the angle score and the distance score, the advantage score is defined as follows:

$$S_{adv} = S_{angle} + \alpha S_{dist} \quad (21)$$

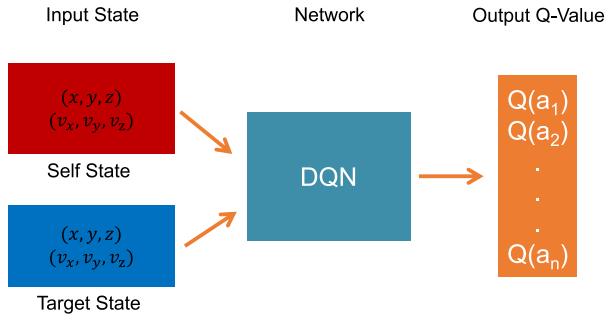


FIGURE 6. Deep Q network (DQN) for maneuver decision.

where α is a weight factor to balance the angle advantage score and distance advantage score.

Since we only consider the motion model, the optional action is to increase, decrease or maintain the velocity in each of the X, Y, Z axes. By changing the velocity in each axis individually, the heading angle and pitch angle are changed accordingly. For the velocity in each axis v_i , it could be changed to $v_i + \delta_v$, $v_i - \delta_v$, or maintain the same, where δ_v is the discretized value for velocity change. As a result, we have $3 \times 3 \times 3 = 27$ kinds of actions to choose. For each of the possible actions, we evaluate the advantage score as in (21) and choose the action with the largest advantage score as the optimal action.

The criterion of winning the air combat is as follows:

$$\begin{cases} \theta_1 < \frac{\pi}{6} \\ \theta_2 < \frac{\pi}{3} \\ d_{\min} \leq \|\mathbf{p}\| \leq d_{\max} \end{cases} \quad (22)$$

Although here we only consider one step to make maneuver decision, satisfactory results have been achieved in various different cases thanks to the robust and accurate state estimation provided by our 3D detection and tracking modules. The experimental results will be presented in detail in Section IV-D.

3) MANEUVER DECISION BASED ON DQN

Apart from the simple one-step decision model, we also test another model based on Deep Q Network (DQN). The DQN-based maneuver decision model is shown in Fig. 6. The input to the network is the state vector of both sides:

$$State = (x_1, y_1, z_1, v_{x1}, v_{y1}, v_{z1}; x_2, y_2, z_2, v_{x2}, v_{y2}, v_{z2}) \quad (23)$$

The network is a 3-layer fully connected network, with 128, 64 and 32 units respectively.

The output is the Q value of different actions. The actions are defined as follows:

$$Action = \{v_{x+}, v_{x-}, v_x =\} \times \{v_{y+}, v_{y-}, v_y =\} \times \{v_{z+}, v_{z-}, v_z =\} \quad (24)$$



FIGURE 7. 3D tracking results for an aircraft model. Left: input image; Right: 6-dof tracking results. The 3D CAD model of the aircraft is rendered according to the estimated pose parameters and superimposed on the original image. The estimated 3D position and orientation are also displayed in the image.

The possible action is to increase, decrease or maintain the velocity in each of the X, Y, Z axes, resulting in 27 possible actions in total.

The reward includes advantage reward and terminate reward:

$$Reward = R_{advantage} + \beta R_{terminate} \quad (25)$$

where the advantage reward is the same as the advantage score S_{adv} in Section III-D2), and the terminate reward is defined as:

$$R_{terminate} = \begin{cases} 1, & \text{if win} \\ -1, & \text{if lose} \\ 0, & \text{otherwise} \end{cases} \quad (26)$$

The loss function for DQN is:

$$L(\theta) = \|Q(s_i, a_i; \theta) - (r_i + \gamma \max_{a'} Q'(s_{i+1}, a'; \theta'))\| \quad (27)$$

where Q is the online Q-network, and Q' is the target Q-network. The parameters of the online Q-network is updated to the target Q-network periodically. r_i is the reward, and γ is the discount factor.

IV. EXPERIMENT

To simulate realistic air combat scenes, we use an open-source python-based 3D air combat simulator [28], which is developed upon the HARFANG 3D framework [29]. In all of the following experiments, we use a commodity desktop

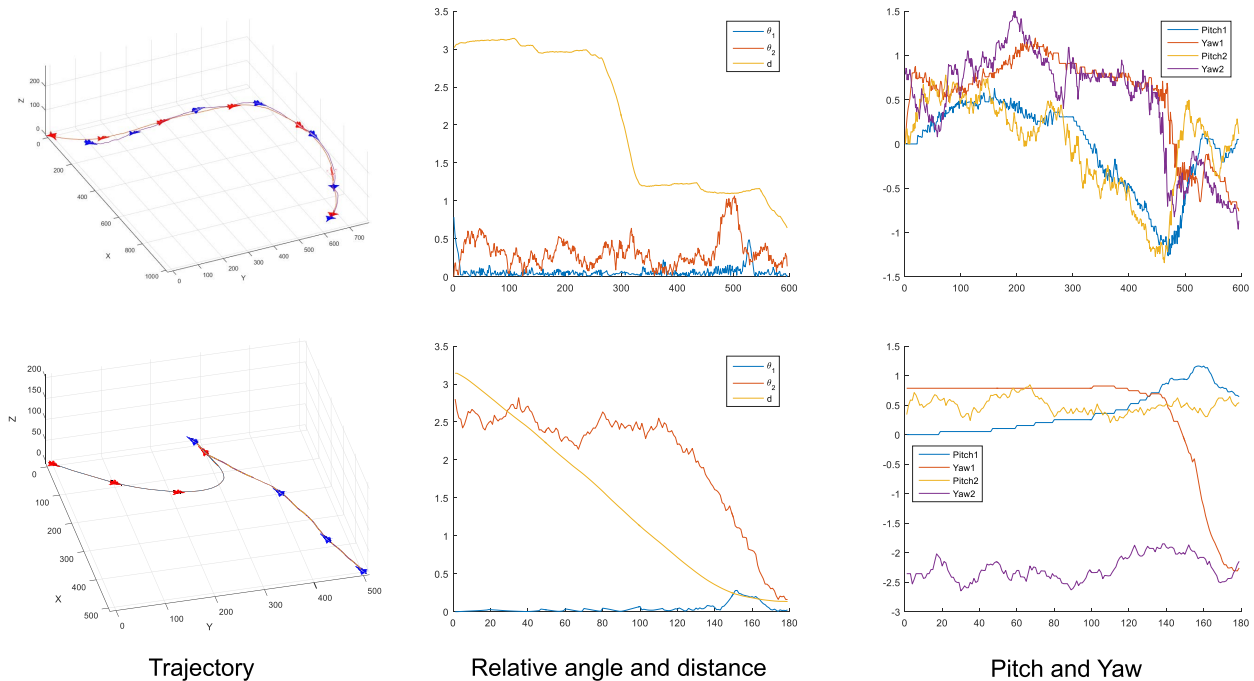


FIGURE 8. Curves for the relative angle and distance, and the pitch/yaw angles of both sides along the whole trajectory in two experiments.

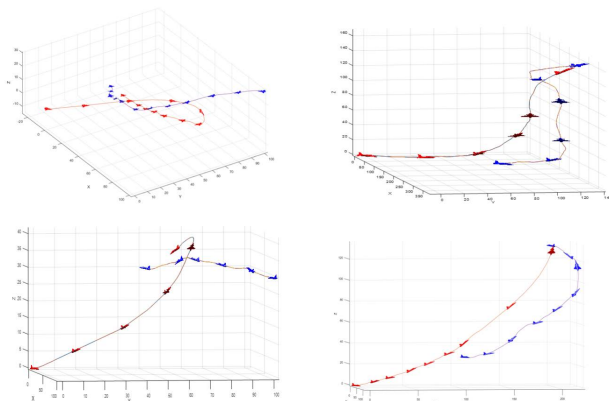


FIGURE 9. Maneuver trajectory of some typical cases in our experiment for one-step maneuver decision model. The blue aircraft is moving randomly, and the red aircraft tries to beat the blue one using the proposed one-step maneuver decision model.

computer with Intel i7 quad core CPU @4.0GHz and a single NVIDIA GeForce GTX1080Ti GPU.

A. RESULTS ON 3D TARGET DETECTION

In order to test the effectiveness of the proposed 3D target detection module, we evaluate our method on a subset of the famous public dataset, LineMod dataset [14]. We compare with two state-of-art methods: SSD-6D, a 6-dof object pose estimation method, which is an extension of the SSD detection framework [7]; and Ori-Learn [17], a 3D orientation estimation method based on Augmented Autoencoder. We use the ADD metric as in previous works. The evaluation results are summarized in Table 1.

TABLE 1. Detection accuracy in a subset of LineMod dataset using ADD metric, best scores are in bold.

Object	SSD-6D [7]	Ori-Learn [17]	Ours
Can	1.35	35.87	37.41
Duck	0.00	4.86	15.92
Iron	8.86	32.03	46.91

The results show that our method obtains higher ADD scores compared to the other two state-of-art methods. Specifically for the Duck object, our method performs much better, which proves the effectiveness of the proposed 3D target detection module.

We have also tested our 3D target detection method on a synthetic aircraft and a real aircraft model. The detection results are shown in Section IV-B together with the tracking process.

B. RESULTS ON 3D TARGET TRACKING

Next, we evaluate the performance of our 3D target tracking module. Firstly, in order to test the capability of handling occlusions, we evaluate our method on the occlusion sequences of the synthetic Rigid Pose Dataset [30]. We compare our method with 4 other methods [20], [23], [31], [32] and the results are summarized in Table 2. We measure the tracking success rate (SR), which is defined as the proportion of frames that are successfully tracked (in %). The results show that our method performs better than the other state-of-art methods in occlusion cases.

Secondly, we test our 3D target tracking module in real environment. We use a webcam to capture a real aircraft

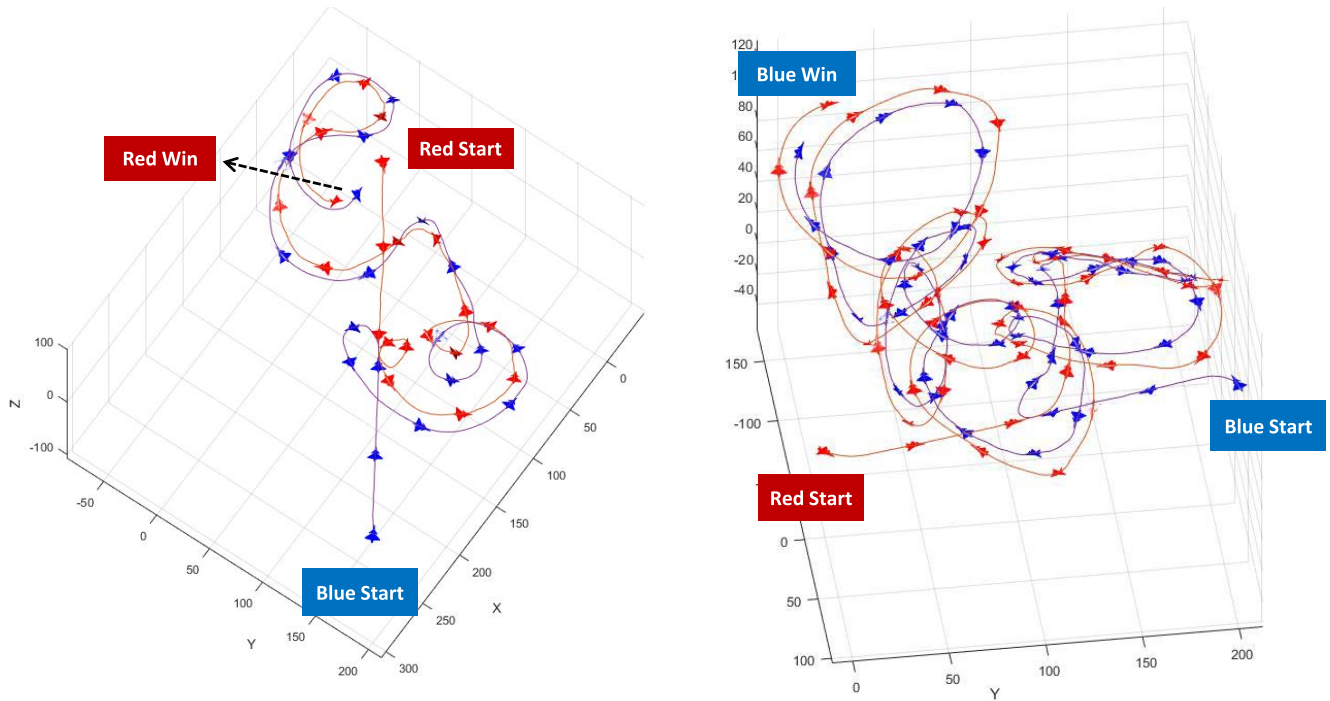


FIGURE 10. Two sample trajectories of the DQN-based model (red) fighting the one-step model (blue). Left: DQN-based model win; Right: One-step model win.

TABLE 2. Evaluation results on occluded sequences of rigid pose dataset. (Tracking success rate in %, best scores are in bold.)

Method	soda	soup	clown	candy	cube	edge	average
[20]	44.0	44.0	43.8	39.2	37.8	38.9	41.3
[31]	73.0	82.0	81.0	84.0	76.0	64.0	76.7
[32]	76.3	86.7	84.1	80.9	80.6	68.1	79.5
[23]	75.1	79.0	84.1	83.6	73.0	66.9	77.0
Ours	78.8	88.4	88.5	89.9	94.5	95.1	89.2

model of the F-16 fighter, and try to detect and track the aircraft model in a cluttered background. The tracking results are demonstrated in Fig. 7.

To demonstrate the tracking accuracy, we render the 3D CAD model of the aircraft according to the estimated 6-dof poses and superimpose the rendered model on the image. If the rendered model fits the aircraft on original image, the composed image would be visually satisfied and the tracking results are qualitatively accurate. As shown in Fig. 7, our 3D tracking method could robustly and accurately track the aircraft model in cluttered background, which proves the effectiveness of the proposed tracking module.

C. RESULTS ON MANEUVER DECISION IN CLOSE-RANGE AIR COMBAT

At last, we could finally test the maneuver decision model using the state information provided by the 3D detection and tracking modules.

Firstly, we test the proposed simple one-step maneuver decision model. In this experiment, we have red and blue aircrafts in two sides. The blue aircraft is moving at random,

and the red aircraft tries to beat the blue one using the proposed one-step maneuver decision model. In each trial, the two aircrafts are placed at two random positions in a limited region, and are set with random velocities. Some trajectories of the red and blue aircrafts in different trials are demonstrated in Fig. 9. In all these trials, the red aircraft armed with the proposed one-step maneuver decision model successfully chase the tail of the blue aircraft and win the air combat.

We have also drawn the curves of the relative angle θ_1 (ATA), θ_2 (AA), the relative distance d , together with the curves of pitch and yaw angles of both sides in Fig. 8. It is shown that, through the chasing process, the relative angle θ_1 , θ_2 and the relative d tend to decrease gradually, so that the advantage score would increase. Also, the pitch and yaw angles of the red aircraft generally try to follow the blue aircraft in order to chase its tail.

Secondly, we test the proposed DQN-based maneuver decision model. In this part of experiment, the DQN-based model fights with both the random moving model and the one-step maneuver decision model. When fighting with the random moving model, the DQN-based model acts similarly to the one-step model, but obtain higher success rate. When fighting with the one-step maneuver decision model, the DQN-based model also wins in most cases (but still loses in about 15% of the trails). Two sample trajectories of the DQN-based model (red) fighting the one-step model (blue) are shown in Fig. 10.

TABLE 3. Winning rate of different configurations.

	Red	Blue	Red Winning Rate
One-step	Random	Random	92.4%
DQN	Random	Random	99.1%
DQN	One-step	One-step	84.6%

In the end, we test the winning rate of 3 different cases: (1) One-step model vs. Random moving model; (2) DQN model vs. Random moving model; (3) DQN model vs. One-step model. For each case, we run 1000 trials and record the winning rate. The results are shown in Table 3. The DQN-based model performs best, which proves the advantage of considering more than one step by using deep Q learning.

V. CONCLUSION

In this paper, we have proposed a novel vision-based maneuver decision method in closed-range air combat. The proposed method contains three main modules: 3D target detection, 3D target tracking, and maneuver decision. The proposed method effectively handles the difficulties in air combat environment, such as fast movement, occlusion from cloud, etc. By robustly detecting and tracking the target aircraft in complex environment, our method could provide optimal maneuver decision which significantly improves the winning rate of short-range air combat. Experiments show that the proposed 3D detection and tracking methods perform better than previous methods, and the proposed maneuver decision models could effectively provide optimal decisions to win the close-range air combat.

REFERENCES

- [1] Y. Li, W. Han, and Y. Wang, "Deep reinforcement learning with application to air confrontation intelligent decision-making of manned/unmanned aerial vehicle cooperative system," *IEEE Access*, vol. 8, pp. 67887–67898, 2020.
- [2] Q. Yang, J. Zhang, G. Shi, J. Hu, and Y. Wu, "Maneuver decision of UAV in short-range air combat based on deep reinforcement learning," *IEEE Access*, vol. 8, pp. 363–378, 2019.
- [3] G. Zhang, Y. Li, X. Xu, and H. Dai, "Efficient training techniques for multi-agent reinforcement learning in combat tasks," *IEEE Access*, vol. 7, pp. 109301–109310, 2019.
- [4] Z. Wang, H. Li, H. Wu, and Z. Wu, "Improving maneuver strategy in air combat by alternate freeze games with a deep reinforcement learning algorithm," *Math. Problems Eng.*, vol. 2020, pp. 1–17, Jun. 2020.
- [5] W. Kong, D. Zhou, Z. Yang, Y. Zhao, and K. Zhang, "UAV autonomous aerial combat maneuver strategy generation with observation error based on state-adversarial deep deterministic policy gradient and inverse reinforcement learning," *Electronics*, vol. 9, no. 7, p. 1121, Jul. 2020.
- [6] Y. Xiang, T. Schmidt, V. Narayanan, and D. Fox, "PoseCNN: A convolutional neural network for 6D object pose estimation in cluttered scenes," in *Robotics, Science and Systems (RSS)*, 2018.
- [7] W. Kehl, F. Manhardt, F. Tombari, S. Ilic, and N. Navab, "SSD-6D: Making RGB-based 3D detection and 6D pose estimation great again," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 1521–1529.
- [8] T.-T. Do, M. Cai, T. Pham, and I. Reid, "Deep-6DPose: Recovering 6D object pose from a single RGB image," 2018, *arXiv:1802.10367*.
- [9] B. Tekin, S. N. Sinha, and P. Fua, "Real-time seamless single shot 6D object pose prediction," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 292–301.
- [10] M. Oberweger, M. Rad, and V. Lepetit, "Making deep heatmaps robust to partial occlusions for 3D object pose estimation," in *Eur. Conf. Comput. Vis. (ECCV)*, 2018, pp. 119–134.
- [11] S. Peng, Y. Liu, Q. Huang, X. Zhou, and H. Bao, "PVNet: Pixel-wise voting network for 6DoF pose estimation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 4561–4570.
- [12] K. Eickenhoff, Y. Yang, P. Geneva, and G. Huang, "Tightly-coupled visual-inertial localization and 3-D rigid-body target tracking," *IEEE Robot. Autom. Lett.*, vol. 4, no. 2, pp. 1541–1548, Apr. 2019.
- [13] K. M. Judd and J. D. Gammell, "The Oxford multimotion dataset: Multiple SE(3) motions with ground truth," *IEEE Robot. Autom. Lett.*, vol. 4, no. 2, pp. 800–807, Apr. 2019.
- [14] S. Hinterstoisser, S. Holzer, C. Cagniart, S. Ilic, K. Konolige, N. Navab, and V. Lepetit, "Multimodal templates for real-time detection of texture-less objects in heavily cluttered scenes," in *Proc. Int. Conf. Comput. Vis.*, Nov. 2011, pp. 858–865.
- [15] S. Hinterstoisser, C. Cagniart, S. Ilic, P. Sturm, N. Navab, P. Fua, and V. Lepetit, "Gradient response maps for real-time detection of texture-less objects," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 5, pp. 876–888, May 2012.
- [16] S. Hinterstoisser, V. Lepetit, S. Ilic, S. Holzer, G. Bradski, K. Konolige, and N. Navab, "Model based training, detection and pose estimation of texture-less 3D objects in heavily cluttered scenes," in *Proc. Asian Conf. Comput. Vis. (ACCV)*, Springer, 2012, pp. 548–562.
- [17] M. Sundermeyer, Z.-C. Marton, M. Durner, M. Brucker, and R. Triebel, "Implicit 3D orientation learning for 6D object detection from RGB images," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, Sep. 2018, pp. 699–715.
- [18] M. Sundermeyer, Z.-C. Marton, M. Durner, and R. Triebel, "Augmented autoencoders: Implicit 3D orientation learning for 6D object detection," *Int. J. Comput. Vis.*, vol. 128, no. 3, pp. 714–729, Mar. 2020.
- [19] V. Lepetit and P. Fua, *Monocular Model-Based 3D Tracking Rigid Objects*. Norwell, MA, USA: Now Publishers Inc, 2005.
- [20] V. Prisacariu and I. Reid, "PWP3D: Real-time segmentation and tracking of 3D objects," *Int. J. Comput. Vis.*, vol. 98, no. 3, pp. 335–354, 2012.
- [21] J. Hexner and R. R. Hagege, "2D–3D pose estimation of heterogeneous objects using a region based approach," *Int. J. Comput. Vis.*, vol. 118, no. 1, pp. 95–112, 2016.
- [22] H. Tjaden, U. Schwanecke, E. Schomer, and D. Cremers, "A region-based gauss-Newton approach to real-time monocular multiple object tracking," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 41, no. 8, pp. 1797–1812, Aug. 2019.
- [23] L. Zhong and L. Zhang, "A robust monocular 3D object tracking method combining statistical and photometric constraints," *Int. J. Comput. Vis.*, vol. 127, no. 8, pp. 973–992, Aug. 2019.
- [24] L. Zhong, X. Zhao, Y. Zhang, S. Zhang, and L. Zhang, "Occlusion-aware region-based 3D pose tracking of objects with temporally consistent polar-based local partitioning," *IEEE Trans. Image Process.*, vol. 29, pp. 5065–5078, 2020.
- [25] L. Zhong, Y. Zhang, H. Zhao, A. Chang, W. Xiang, S. Zhang, and L. Zhang, "Seeing through the occluders: Robust monocular 6-DOF object pose tracking via model-guided video object segmentation," *IEEE Robot. Autom. Lett.*, vol. 5, no. 4, pp. 5159–5166, Oct. 2020.
- [26] J. Redmon and A. Farhadi, "YOLOv3: An incremental improvement," 2018, *arXiv:1804.02767*.
- [27] L.-C. Chen, Y. Zhu, G. Papandreou, F. Schroff, and H. Adam, "Encoder-decoder with atrous separable convolution for semantic image segmentation," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, Sep. 2018, pp. 801–818.
- [28] HARFANG. (2021). *Game-Dogfight*. [Online]. Available: <https://github.com/harfang3d/game-dogfight>
- [29] HARFANG. (2021). *HARFANG 3D*. [Online]. Available: <https://www.harfang3d.com/>
- [30] K. Pauwels, L. Rubio, J. Diaz, and E. Ros, "Real-time model-based rigid object pose estimation and tracking combining dense and sparse visual cues," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2013, pp. 2347–2354.
- [31] S. Zhao, L. Wang, W. Sui, H.-Y. Wu, and C. Pan, "3D object tracking via boundary constrained region-based model," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Oct. 2014, pp. 486–490.
- [32] L. Zhong, M. Lu, and L. Zhang, "A direct 3D object tracking method based on dynamic textured model rendering and extended dense feature fields," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 28, no. 9, pp. 2302–2315, Sep. 2018.



LEISHENG ZHONG received the B.S. and Ph.D. degrees in signal and information processing from Tsinghua University, Beijing, China, in 2013 and 2020, respectively. He is currently an Engineer with NVRI. His research interests include 3D computer vision, vision for robotics, and machine learning.



JIALIN CHEN received the B.S., M.S., and Ph.D. degrees from Naval Aeronautical University, Shandong, China. He is currently an Engineer with NVRI.



LEIMING ZHAO received the B.S. and M.S. degrees in electronic engineering from Naval Aeronautical University, Shandong, China. He is currently a Senior Engineer (a Professor) with NVRI.



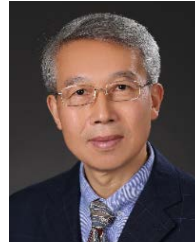
YU ZHANG received the B.S. degree in engineering from Dalian Maritime University, Dalian, China, in 2012, the M.S. degree in pattern recognition and intelligent systems from Beihang University, Beijing, China, in 2015, and the Ph.D. degree in signal and information processing from Tsinghua University, Beijing. He is currently a Postdoctoral Researcher with the Image Processing Center, School of Astronautics, Beihang University. His research interests include computer vision, medical image processing, and image fusion.



CHENCONG DING received the B.S. degree from Naval Aeronautical University and the M.S. degree from Beihang University, China. He is currently a Senior Engineer with NVRI.



XUESHI GE received the B.S. and M.S. degrees in software engineering from Shanghai Jiaotong University, Shanghai, China. He is currently an Engineer with NVRI.



LI ZHANG received the B.S., M.S., and Ph.D. degrees in signal and information processing from Tsinghua University, Beijing, China, in 1987, 1992, and 2008, respectively. In 1992, he joined the Faculty of the Department of Electronic Engineering, Tsinghua University, where he is currently a Professor. His research interests include image processing, computer vision, pattern recognition, and computer graphics.

...