Autonomous Maneuvering Decision Research of UAV Based on Experience Knowledge Representation

Xuemei He[1], Wei Zu[2], Hongxing Chang[3], Jie Zhang[4], Yang Gao[5]

1. Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China
E-mail: xuemei.he@ia.ac.cn

2. Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China
E-mail: wei.zu@ia.ac.cn

3. Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China
E-mail: hongxing.chang@ia.ac.cn

4. Chengdu Aircraft Design and Research Institute, Chengdu 610041, China
E-mail: zjzj611@gmail.com

5. Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China
E-mail: yang.gao@ia.ac.cn

Abstract: An autonomous maneuvering decision algorithm based on experiences of pilots and tactics theories in one-to-one air combat is proposed. Firstly a method of situation assessment based on näive Bayes is addressed, in which the historical situation is taken into consideration when predicting the next moment situation. Then preferences of experienced pilots and tactics theories are mathematically modeled. This model is more flexible and extensible than that models made up from a bunch of rules, because the mathematical model is easier to code and can adaptively adjust flying state of aircraft according to the output of situation assessment. Finally a maneuvering scoring function and a maneuvering matrix are described. The whole process of combating is regarded as a static complete information game where an optimal maneuvering solution by Nash equilibrium solution exists. Simulations are devised by one unmanned aircraft running the proposed algorithm against one manned aircraft adversary. Results show that the unmanned aircraft will reach an advantageous situation regardless of starting from offensive or defensive state.

Key Words: Experiential Knowledge, Bayes, Nash Equilibrium, Air Combat

1 INTRODUCTION

One-to-one air combat is a pursuer-evader game where roles of pursuer and evader change fast. The purpose of the game is to maneuver one's aircraft into advantageous position over adversary, and at the same time minimize own risks of attack from the adversary.

Research of autonomous maneuvering decision of unmanned aerial vehicle(UAV) has gone for nearly 30 years. Differential game [1–4] is the first proposed algorithm to solve the pursuer-evader game. During the process of of differential game, roles of pursuer and evader are fixed, which is quite distinct from real combat. Rule-based algorithm [5, 6] is effective in specific scenes. TacAir-Soar system developed by DARPA is the most typical rule-based system, some of its performances are similar to experienced pilots, yet the rule-based algorithm is time consuming, deficient in reusability and extendibility. Every time the same efforts are required when applied on any other type of aircrafts with different flying parameters. Fuzzy reasoning algorithms [7–10] output tactics in a high level with long decision cycle, yet this method is unpractical if there is none well-developed mechanism of re-planning. In recent years, approximate dynamic programming [11], particle swarm optimization [12], neural network [13, 14] and reinforcement learning [15] are applied in air combat field as well, yet most of these intelligent algorithms are sill in laboratory research, due to requirements for large training set.

At the beginning of 2015, DARPA proposed a new mode of collaborative operation of multi UAVs in denied environment(CODE), that UAV will develop from composite type into unitary type. Each aircraft in a group focuses on unique task to unload burden of situation assessment system and communication system that are ever integrated in one aircraft, because there are specific aircrafts taking in charge of identification of friend and foe, situation assessment and long range communication with ground station. On the other hand, assault aircrafts require more flexible ability of dogfight focuses on attacking the adversaries [16].

Algorithms mentioned above are not flexible enough to implement close combat, neither rule-based system nor expert system includes all cases in air combat, because it is improbable to enumerate all combat rules and tactics. An algorithm that is easier to code, more timesaving and reusable is proposed to describe experiences of pilots and tactics theories introduced in [17, 18] in a mathematical way. Firstly, situation assessment model based on näive Bayes is addressed [3], it outputs a distribution of the cur-

This work is supported by National Nature Science Foundation of China (Grant No.61304096).

rent situation instead of a determined value. Then several continuous functions are discussed to present preferences of pilots and tactics theories. Continuous functions are able to cover all cases in air combat, because the relative position of the unmanned aircraft and the manned aircraft falls into a certain definition domain of the continuous functions. Finally a decision method based on Nash equilibrium [5, 17] is described to output sequences of optimal maneuvering at discrete time moment.

2 ANALYSIS OF AUTONOMOUS MANEUVERING DECISION OF UAV

In the one-to-one air combat, outputs of maneuvering decision system are influenced by many factors, for example, the preferences of pilots, accuracy of situation assessment, time-consuming of algorithm, a good algorithm ought to take all of them into consideration as inputs for the autonomous decision system. We assume that the unmanned (the blue in simulation) and the manned adversary (the red in simulation) are both able to get precise position and orientation each other, and the continuous combat process is regarded as a discrete process, therefore the decision system has a fixed decision cycle, and it outputs the optimal maneuvering at t_k moment, where $t_k = k\Delta t$.

2.1 Analysis of Model in One-to-one Aircombat

Figure 1 is a diagram describing dependence between various factors in decision process. Nodes S_k^u, S_k^m are respectively pose of the unmanned and the manned, node $C_k = \begin{bmatrix} \omega_k & \theta_k & d_k \end{bmatrix}^T$ (bearing, back-off, distance) is the relative position of the two aircrafts. $N = \{n_1, n_2, \cdots, n_l\}$ is a maneuvering base including twenty-seven types of basic fighter maneuvering(BFM), node Θ_k is the output of situation assessment, node Q_k is evaluation of maneuvering utility, which describes preferences of pilots and tactics introduced in [17]. $\forall S_k^u,\ S_k^m \in \{S\}$, $\exists f$ that $f : \{C_k, \Theta_k\} \to N$, where $S = \begin{bmatrix} v & \phi & \gamma & x & y & z \end{bmatrix}$ (velocity, yaw, pitch, x coordinate, y coordinate, z coordinate). $\exists f_c, f_\theta, f_q$ that $C_k = f_c(S_k^u, S_k^m)$, $\Theta_k = f_\theta(C_k, \Theta_{k-1})$, $Q_k = f_q(C_k, \Theta_k)$. The purpose of the proposed algorithm is to find out proper f_c, f_θ, f_q and f to respectively calculate the relative position of two aircrafts C_k, situation assessment Θ_k, influence of each maneuvering to the future state Q_k and autonomous air combat strategy.

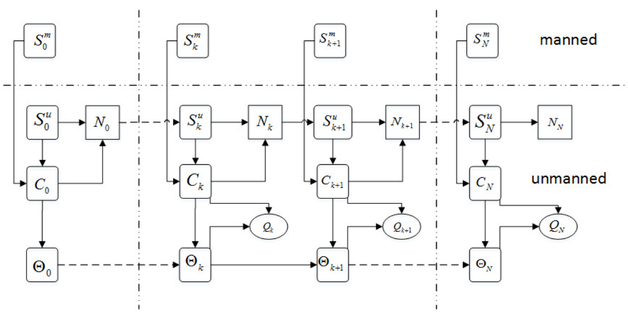


Fig. 1 Decision state diagram of UAV

Relative position C_k of the two aircrafts is one of the most imperative inputs of autonomous decision system, assume the body orientation is in accordance with speed diretion, there is

$$\begin{aligned} C_k &= \begin{bmatrix} \arccos(\frac{\vec{v^u}\cdot\vec{d}}{\|\vec{v^u}\|\|\vec{d}\|}) & \arccos(\frac{\vec{v^m}\cdot\vec{d}}{\|\vec{v^m}\|\|\vec{d}\|}) & \|\vec{d}\| \end{bmatrix} \\ \vec{v^i} &= \begin{bmatrix} \cos(\gamma^i)\cos(\phi^i) & \cos(\gamma^i)\sin(\phi^i) & \sin(\phi^i) \end{bmatrix} \\ \vec{d} &= \begin{bmatrix} x^m - x^u & y^m - y^u & z^m - z^u \end{bmatrix}^T \end{aligned} \quad (1)$$

2.2 Situation Assessment

Situation assessment outcome is an imperative input for autonomous decision system, incorrect assessment outcome may lead to task failure or even air crash. Herein the situation assessment model is based on näive Bayes, and the probability of situation given relative position C_k is denoted as $P(\Theta_k = j|C_k)$. Θ_k is classified into four types, neutral, advantage, disadvantage and mutual disadvantage [5]. The types are respectively quantified by 1, 2, 3 and 4 that satisfy $\sum_{j=1}^4 P(\Theta_k = j|C_k) = 1$, where $j = \{1, 2, 3, 4\}$. Situation is easily inferred from C_k, for example, when the manned is tail bited by the unmanned, it is obvious the unmanned is in advantage, and otherwise in disadvantage if the unmanned is tail bited by the manned. Hence there is

$$\begin{aligned} &P(\Theta_{k+1}|C_{k+1}) \\ &= \frac{P(\Theta_{k+1} = j)P(C_{k+1}|\Theta_{k+1} = j)}{\sum_1^4 P(\Theta_{k+1} = l)P(C_{k+1}|\Theta_{k+1} = l)} \\ &\approx \frac{P(\Theta_k = j)P(C_{k+1}|\Theta_{k+1} = j)}{\sum_1^4 P(\Theta_k = l)P(C_{k+1}|\Theta_{k+1} = l)} \end{aligned} \quad (2)$$

The prior probability of $\Theta_{k+1} = j$ is unknown, yet the flying state is improbable to change heavily in a short time, due to the reality that aircraft is a large inertia object, it is reasonable to use $P(\Theta_k = j)$ as prior probability $P(\Theta_{k+1} = j)$. At the same time, assume that probabilities of ω, θ and d given C_k all obey uniform distribution and independent with each other, there is

$$\begin{aligned} &P(C_k|\Theta_k = j) \\ &= p(\omega_k|\Theta_k = j)p(\theta_k|\Theta_k = j)p(d_k|\Theta_k = j) \end{aligned} \quad (3)$$

3 MODEL OF EXPERIENCES OF PILOTS AND TACTICS KNOWLEDGE

3.1 Representation Model of Tactics Theories

Typical tactics of one-to-one air combat, including composite maneuvering and basic maneuvering, are introduced in [17]. Life cycle of composite maneuvering are longer compared with basic maneuvering, because composite maneuvering primarily aim at changing defensive state into offensive state and maintaining offensive state to reach a proper weapon launch position. Most often, it is difficult to build a mathematical model for composite maneuvering at the tactical level, yet bearing ω, angle-off θ and distance d change constantly due to maneuvering implementation. ω, θ and d are the bottom states that constantly change during the process of air combat, thus the proposed algorithm focuses on a general model for tactics introduced in [17] and preferences of pilots in terms of ω, θ and d. In order to better model tactics, a weapon launch area is introduced firstly

$$T_i = \{C_k = \psi(C_k) \le 0\} \; (i = u, m)$$
$$\psi(C_k) = \left[\begin{array}{ccc} \omega_k - \omega_T & \theta_k - \theta_T & d_k - d_T \end{array} \right] \quad (4)$$

ω_T, θ_T, d_T are fixed value, the larger the value is, the stronger the strick capability of the airborne weapon is. The purpose of the decision system is to maneuver the unmanned to force the manned fly into T_i.

A position utility function $u^j(\cdot)$ is introduced to quantize influences of ω, θ and d on future situation. Larger $u^j(\cdot)$ indicates more benefits for future state after implementing a certain maneuvering in situation state j. In any situation state, both ω and θ should be reduced to achieve T_i, that is $u^j(\omega) \propto \frac{1}{\omega}$, $u^j(\theta) \propto \frac{1}{\theta}$ (see Eq. (5) and Eq. (6)). It reduces risks to be attacked by the adversary when in disadvantage state and increases probability of reaching weapon launch position when in advantage state. Case of d is complex, d should be reduced when the unmanned and the manned are far away, because influences of ω and θ are much less imperative for future state than that of d when the two aircrafts are wide apart. When the two aircrafts are much close, especially when one is pursuing the other, ω and θ should be adjusted instead of d for better weapon launching position. In a summarization, $u^j(\cdot)$ is

$$u^j(\omega) = 1 - \frac{\omega}{\pi} \quad (5)$$

$$u^j(\theta) = 1 - \frac{\theta}{\pi} \quad (6)$$

$$u^j(d) = \begin{cases} 1 - \frac{d}{R_D} & j = 1, 4 \\ \frac{1}{R_D} & j = 3 \\ u^2(d) & j = 2 \end{cases}$$
$$u^2(d) = \left(1 - \left(\frac{\beta_1}{\beta_1 + d_{k+1}}\right)^2\right) e^{-\left(\frac{d_{k+1} - R_i}{\beta_2}\right)^2} +$$
$$\left(1 - \frac{d_{k+1}}{R_D}\right) \varepsilon\left(d_{k+1} - R_p\right) \quad (7)$$

R_D is a parameter to control the rate of convergence of $u^2(d)$, R_i, R_p are respectively the minimum and optimum range of weapon launching. Figure 2 is curve of $u^2(d)$, in which $R_i = 3000m$, $R_p = 5000m$ and $R_D = 15000m$. Figure 2 indicates that in order to get larger $u^2(d)$, decision system will reduce range d when $d > 3000m$, and maintain in $d = 3000m$ when $d < 3000m$, this regulation prevents crashing when the unmanned is tail biting the manned. $\varepsilon(d_{k+1})$ is a step function that enables the unmanned to fast approach the manned otherwise due to slow attenuation of $u^2(d)$, speed of the unmanned is small when d is large.

3.2 Adjusted Weight of C_k

Pilots have various preferences on various situation for the same maneuvering, herein a weight vector α_j is applied to quantify such preferences. Effects on maneuvering are finally represented by changes of ω ,θ and d. When two aircrafts are flying in different directions, d and ω should

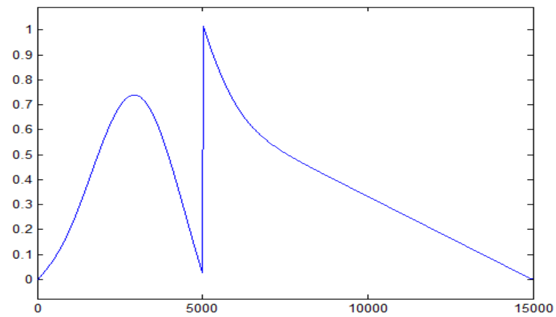


Fig. 2 Curve of $u^2(d)$

be reduced to avoid being followed by the other, because θ is less imperative to situation variation at this moment. When the unmanned is in advantage and at the same time the two aircrafts get closer, effects of ω and θ on situation state increase as well as that of d decreases. Table 1 is detail value of α_j.

$$a^j = \left[\begin{array}{ccc} \alpha^{j,\omega} & \alpha^{j,\theta} & \alpha^{j,d} \end{array} \right]$$
$$\alpha^{j,\omega} + \alpha^{j,\theta} + \alpha^{j,d} = 1 \quad (8)$$
$$\alpha^{2,\omega} = e^{-\frac{d}{R_P}}, \alpha^{2,d} = 1 - \alpha^{2,\omega}$$

Table 1 Weight of C_k in various situation state

Θ_k	j	$\alpha^{j,\omega}$	$\alpha^{j,\theta}$	$\alpha^{j,d}$
Neutral	1	0.2	0.1	0.7
Advantage	2	**	0.0	**
Mutual disadvantage	3	0.0	0.7	0.3
Disadvantage	4	0.2	0.1	0.7

3.3 Representation Model of Maneuvering Utility

With utility function $u^j(\cdot)$ and weight vector α^j of various maneuvering in situation $\Theta_k = j$, a weighted sum denoted in Q_k is introduced to describe the utility of a certain maneuvering in different situation state, that is

$$Q_k = f_q(\Theta_k, C_k) =$$
$$[U(1, C_k) \;\; U(2, C_k) \;\; U(3, C_k) \;\; U(4, C_k)]^T$$
$$U(j, C_k) = \alpha^{j,\omega} u^j(\omega) + \alpha^{j,\theta} u^j(\theta) + \alpha^{j,d} u^j(d) \quad (9)$$

The current situation state is not a determined value but obeys a probability distribution, here we take $P(\Theta_{k+1} = j | C_{k+1})$ as the weight of $U(j, C_k)$ to generate a weighted sum denoted in J_{k+1}, and J_{k+1} is the final utility of maneuvering at $k+1$ moment given k moment situation.

$$J_{k+1} = P(\Theta_{k+1} | C_{k+1})^T Q(\Theta_{k+1} | C_{k+1}) =$$
$$\sum_{j=1}^{4} P(\Theta_{k+1} = j | C_{k+1})^T Q(\Theta_{k+1} = j | C_{k+1}) \quad (10)$$

the larger J_{k+1} is, the more probably the unmanned will be superior to the adversary after certain maneuvering is implemented.

3.4 Autonomous Decision Based on Nash Equilibrium

The basic seven kinds of maneuvering in NASA standard [19] are extended to twenty-seven in this paper, assuming the unmanned is able to adjust its vertical overload N_x, tangential overload N_z and roll angle ϕ at the same time.

$$
\begin{aligned}
&N_k = \{\Delta n(r_1, r_2, r_3) \in \{-1, 0, 1\}\} \\
&\Delta n(r_1, r_2, r_3) = \begin{bmatrix} r_1 N_x & r_2 N_z & r_3 \phi \end{bmatrix} \\
&N_x = 3, N_z = 7, \phi = \frac{\pi}{3}
\end{aligned} \tag{11}
$$

Various maneuvering of the unmanned and the manned compose a game matrix $M_{27\times27}$. At moment k, $M_{i,j}$ consists of two elements, J_{k+1} of the unmanned implementing the ith maneuvering and J_{k+1} of the manned implementing the jth maneuvering. The optimal solution falls into three cases: The unique Nash equilibrium solution of $M_{27\times27}$, or the Nash equilibrium solution that maximizes J_{k+1} of the unmanned if $M_{27\times27}$ has more than one Nash equilibrium solutions, or solution based on Max-min method [3] if $M_{27\times27}$ has none Nash equilibrium solution.
Researchers tend to use Nash equilibrium in high order as the optimal solution, yet computational complexity of Nash equilibrium increases exponentially as the depth of game tree grows. Herein the optimal solution is in the first order, because the situation of air combat changes fast, and it is impractical to predict the optimal payoff of maneuvering several steps later at cost of real-time.

4 EXPERIMENTS

Aircrafts in all simulations were one unmanned aircraft (blue) running autonomous decision algorithm and one manned (red) aircraft handled by operator with rocker and accelerator equipments. Parameters of the two aircrafts are detailed in Table 2.

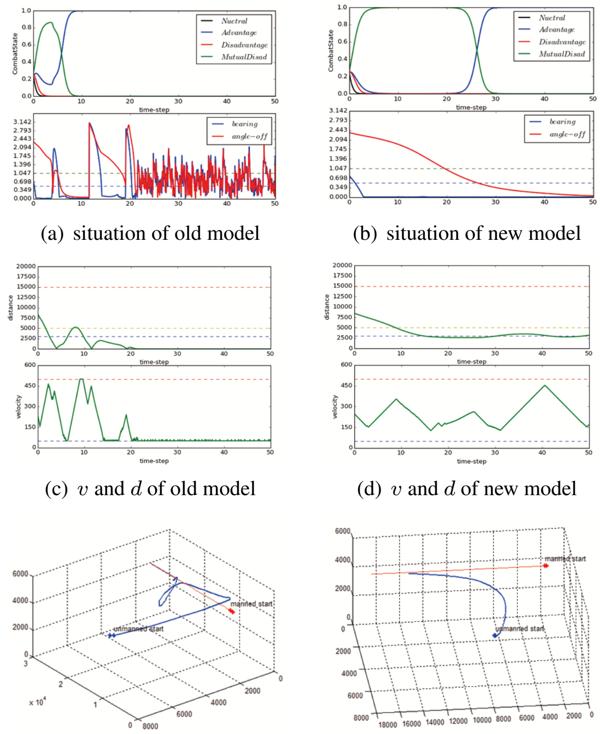
Table 2 Parameters of the manned and the unmanned

parameters	manned	unmanned
max speed	300m/s	300m/s
min speed	50m/s	50m/s
weapon(1km)	[15 4 2]	[15 4 2]
T	$[\pi/6\ \ \pi/6\ \ 3000]$	$[\pi/6\ \ \pi/3\ \ 3000]$

The three parameters in weapon vector are respectively the value of R_D , R_p and R_i (see Eq. (7)). The three parameters in T are respctively the value of ω_T, θ_T and d_T (see Eq. (4)).

4.1 Experiments of Experiences Knowledge Model

In this experiment, we compared the influence on flying state of two different $u^2(d)$, one was introduced in [17] where $u^2(d) = 1 - \frac{d}{R_D}$, one was proposed in this paper. Outputs of these two model are respectively the left three subfigures (see Fig. 3(a), 3(c), 3(e)) and the right subfigures (see Fig. 3(b), 3(d), 3(f)). Bearing and angle-off in Fig. 3(a) and 3(b) show that when the two aircrafts got closer, flying state of the unmanned aircraft running new model was more stable, and the unmanned aircraft was able to maintain a proper position of weapon for longer time. Because of instability, the unmanned running old model missed opportunity to attack the adversary.



(a) situation of old model (b) situation of new model

(c) v and d of old model (d) v and d of new model

(e) 3D output of old model (f) 3D output of new model

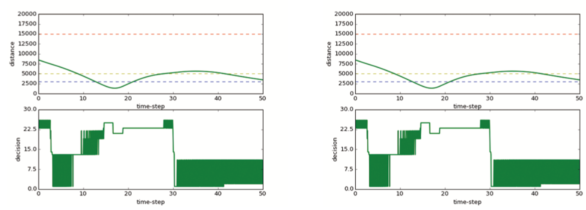
Fig. 3 Model comparison of experiences knowledge

4.2 Decision Effect of Various Order Solution

Table 3 Time consumption in various order(ms)

group	1	2	3	4	5	6
first order	15	16	16	15	15	16
second order	760	780	768	770	781	769

There were six groups of experiments with two decision systems respectively running the first-order algorithm and the second-order algorithm. Table 3 is comparison of time consumption of these two algorithms. Figure 4(a) and 4(b) show that solution of the first-order is only approximately 3 seconds delay to the second-order solution, yet time consumption of the first-order is one over forty-five of the second-order, therefore the first-order decision system is more competitive in air combat which requires real-time.



(a) maneuvering of second-order (b) maneuvering of first-order

Fig. 4 Maneuver output of various order

4.3 Autonomous Decision Based Nash Equilibrium

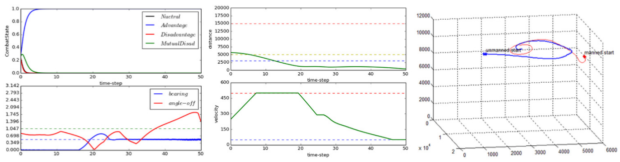
There were three groups simulations in which the unmanned is respectively in advantage state, mutual disadvantage state and disadvantage state.

4.3.1 Advantage Initial State

Table 4 Initial state of the unmanned and the manned

Role	Velocity	Yaw	Pitch	X/km	Y/km	Z/km
blue	250m/s	0.75π	0	2	2	8
red	300m/s	0	0	5	5	8

In this experiment, the unmanned initiated in advantage and at the same time the target was in weapon launch range. Figure 5(a) shows that the unmanned maintained advantageous state in the whole combat duration. When the manned tended to change from defensive to offensive by climbing up, the unmanned quickly climbed up to follow the manned. The two aircrafts were in stalemate state at the end of simulation time.



(a) situation assessment (b) d and v of blue (c) 3D output

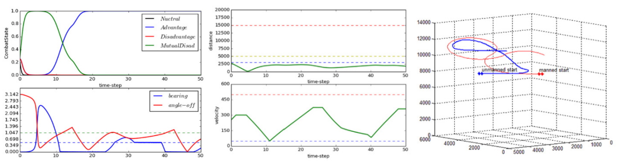
Fig. 5 Simulation of the unmanned in advantage state

4.3.2 Mutual disadvantage initial state

Table 5 Initial state of the unmanned and the manned

Role	Velocity	Yaw	Pitch	X/km	Y/km	Z/km
blue	250m/s	-0.75π	0	4	4	8
red	300m/s	0.25π	0	2	2	8

In this experiment, the two aircraft initiated in different directions, both of them took lateral climbing after tactical intervention. Figure 6(a) and 6(c) show that the unmanned quickly gained advantageous state from mutual disadvantage state, and the unmanned successfully pursued the manned with a lead-pursuit attack at the twelfth second.



(a) situation assessment (b) d and v of blue (c) 3D output

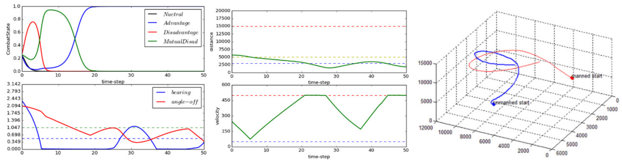
Fig. 6 Simulation of the unmanned in mutual disadvantage state

4.3.3 Disadvantage state

Table 6 Initial state of the unmanned and the manned

Role	Velocity	Yaw	Pitch	X/km	Y/km	Z/km
blue	250m/s	0	0	6	6	8
red	300m/s	0.25π	0	2	2	8

In this experiment the situation state gradually changed from disadvantage to mutual disadvantage, and finally reached advantage at the fifteenth second, in Fig. 7(a) is the detail. The unmanned firstly reduced bearing and angle-off of the manned until reached in weapon launch range. At the twentieth second, the unmanned got control of the air, and followed the manned with a rear-pursuit at the thirtieth second.



(a) situation assessment (b) d and v of blue (c) 3D output

Fig. 7 Simulation of the unmanned in disadvantage state

5 CONCLUSION

An autonomous maneuvering decision algorithm based on experiences of pilots and tactics theories in one-to-one air combat is proposed. Situation is more rigorously reflected by a probability distribution from the näive Bayes situation assessment method, it provides more reasonable outcomes for autonomous decision. The mathematical model of preferences of pilots and tactics theories is able to cover all probable cases in air combat, and at the same time this model is reusable and extendable. Aircraft running this algorithm is able to automatically maneuver into advantageous positions over the adversary, and reduce risks of being attacked by the adversary. When the proposed algorithm is applied on any other type of aircrafts, only maneuvering parameters and weapon launch parameters need to be changed. Statical data of all repeated experiments shows that the winning rate of the unmanned aircraft running the proposed algorithm is more than 90%, the proposed algorithm is proved to be capable for one-to-one air combat.

REFERENCES

[1] Isaacs R. Games of pursuit [J]. 1951.

[2] FU L, WANG X G. Research on Close Air Combat Modeling of Differential Games for Unmanned Combat Air Vehicles [J]. Acta Armamentarii, 2012, 10:1210-1216.

[3] Austin F, Carbone G, Falco M. Automated maneuvering decisions for air-to-air combat [J]. AIAA, 1987: 87-2393.

[4] Guo H, Zhou D Y, Zhang K. Study on UCAV Autonomous Air Combat Maneuvering Decision-Making [J]. Electronics Optics and Control, 2010, 08:28-32.

[5] Virtanen K, Karelahti J, Raivio T. Modeling air combat by a moving horizon influence diagram game [J]. Journal of Guidance, Control, and Dynamics, 2006, 29(5):1080-1091.

[6] Burgin G H, Sidor L B. Rule-based air combat simulation[R]. Titan systems inc la jolla, 1988.

[7] SUN T Z, SUN J B. Tactical Decision-Making for Formation Air Combat Based on Hierarchical Fuzzy Petri Net [J]. Electronics, Optics and Control, 2011, 18(11):22-26.

[8] ZHANG Y, YANG R N, WU M. Air combat tactics decision-making based on intuitionistic fuzzy Petri net. Computer Engineering and Applications, 2012, 48(30)224-228.

[9] Geng W, Kong F, Ma D. Study on tactical decision of UAV medium-range air combat [C]//Control and Decision Conference (2014 CCDC), The 26th Chinese. IEEE, 2014:135-139.

[10] Pawlak Z. Rough sets: Theoretical aspects of reasoning about data [M]. Springer Science and Business Media, 1991.

[11] McGrew J S, How J P, Williams B. Air-combat strategy using approximate dynamic programming [J]. Journal of Guidance, Control, and Dynamics, 2010, 33(5):1641-1654.

[12] GU J J, ZHAO J J, LIU W H. Air Combat Maneuvering Decision Framework Based on Game Theory and Memetic Algorithm [J]. Electronics, Optics and Control, 2015, 01:20-23.

[13] ZHU D F, JI F B, GUAN Y Y. Study On Air Combat Tactics Decision-making for the Fourth Generation Fighters [J]. Command Control and Simulation, 2012, 01:41-43.

[14] ZHONG L, TONG M A, ZHONG W Y. Cooperative Team Air Combat Decision based on Integration of Rough Sets and Neural Networks [J]. Fire Control and Command Control, 2006, 06:881-884.

[15] Li D, Jiang J, Xu H. Reinforcement learning methods for finding equilibria and tracking evolution paths in conflicts [C] //Systems, Man and Cybernetics, 2008. SMC 2008. IEEE International Conference on. IEEE, 2008:3292-3297.

[16] Wang X G, Zhang W G, Chen W. BVR air combat decision making and simulation for UAV formation [J]. Control and Decision, 2015, 30(02):328-334.

[17] Shaw R L. Fighter combat: Tactics and maneuvering [M]. Annapolis, Maryland: Naval Institute Press, 1985.

[18] Tan C L, Huang C Q, Ding D L, Du H W, Cai J. A Method of Intelligent Tactical Decision Making for UCAV Autonomous Air Combat [J]. Command Control and Simulation, 2015, 05:5-11.

[19] Lewis M S, Aiken E W. Piloted Simulation of One-on-One Helicopter Air Combat at NOE (Nap-of-the-Earth) Flight Levels [P]. America:ADA160538, 1985-4.