

# Data Association Algorithm Overview and Performance Evaluation

Xu Wanli<sup>1, a</sup>, Bao Shuo<sup>1</sup>, Liu Zhun<sup>1</sup>

<sup>1</sup>Academy of Armored Force Engineering, Beijing, China

<sup>a</sup>xuwanli631101@sina.com

**Keywords:** Information Fusion, Data Association, Performance Evaluation

**Abstract.** With the continuous development of sensor technology, the sensor has been widely used in modern war. More and more information makes the information fusion has become a research focus in the future battlefield in computer field. In this paper, first of all, we studied the current variety of data association method, and experiment it. On the basis of comparing the experimental data, we analyzed the performance of the associated methods.

## Introduction

With the development of computer technology, a lot of researches are carried out for the data association problem, so far, there has been a lot of data association algorithm. For example, the Nearest Neighbor Data Association (NNDA), Probabilistic Data Association (PDA), the Joint Probabilistic Data Association (JPDA), Fuzzy Data Association (FDA) etc...

Associated methods are diverse, their effect and applicable environment is different, how to quickly find the most suitable association method is what we need most, here we will compare lots of methods which performance has been evaluated.

## The Nearest Neighbor Data Association (NNDA)

So far, many data association algorithm can achieve better association, the Nearest Neighbor Data Association (NNDA) algorithm is proposed first, also is one of the simplest method, but in certain circumstances is the most effective one. This method first set the associated gate to limit the potential number of decision making, and the associated gate's preliminary screen make the echoes become a candidate. Associated door is a subspace in tracking space, the center is located in the prediction of tracked target status, and the size of the design should ensure that can receive the correct probability of echo in a certain extent. The nearest neighbor method always chooses the point trace which falls into the association gate and nearest to the tracked target predicted position. Usually we judge it according to the statistical distance.

Through the analysis it is not difficult to find, the nearest neighbor data association is mainly suitable for tracking domain targets but just a small number of targets 'cases, or only for sparse environment's target tracking. The main advantages are: the computational complexity is small, and is easy to realize. The main disadvantage is that: the environmental limitations.

The definition of statistical distance: hypothesis before the first k times scanning, we have established the N path. New observations for the first k times is  $Z_j(k)$ ,  $j=1,2,\dots,N$ . In the association gate of track i, the difference vector of observed j and track i is defined as the difference between the measured value and predicted value, the filter residue,

$$e_{ij}(k) = Z_j(k) - H\hat{X}_i(k / (k - 1)) \quad (1)$$

Where H is the observation matrix, let S(k) be the covariance matrix. Then the statistical distance (square) is

$$d_{ij}^2 = e_{ij}(k)S_{ij}^{-1}(k)e_{ij}^T(k) \quad (2)$$

It is the judgment of nearest neighbor points metrics.

## The Probabilistic Data Association (PDA)

Probability Data Association believes that: as long as it is effective echo, they are likely to be derived from the target, only the probability of each echo source to target is different. In the first to the first  $k$  times obtained by scanning all the effective echo is known, the  $k$  time's scan, the echo  $i$  ( $i=1,2,3,\dots, m_k$ ) are correct echo, called the correct association probability, it represent by  $P_i(k)$ ,

$$P_i(k) \equiv P\{\theta_i(k) / Z_k\} \quad (3)$$

In the formula,  $\theta_i(k)$ ----the echo  $i$  for the correct echo scanning at time  $k$ ;

$Z_k$ ----all effective echoes from the first times to  $K$  times obtained by scanning;

$m_k$ ----the number of echoes measured by obtaining at time  $k$ .

According to the full probability formula, the target state estimation at  $k$  time, namely the optimal estimate for the mean square sense is

$$\hat{X}(k/k) \equiv \sum_{i=0}^{m_k} P_i(k) \hat{X}_i(k/k) \quad (4)$$

Among them  $\hat{X}_i(k/k)$ ,  $i=1,2,3,\dots, m_k$  That is to estimate the effective echo from the target under the condition of the target state value;  $\hat{X}_0(k/k)$  is the estimated echo from interference or clutter under the target state value.

## The Joint Probability Data Association (JPDA)

The basic idea of joint probabilistic data association: measure the situation that fall into the intersection area of the tracking gates, corresponding to some observation may be due to multiple targets, the purpose of JPDA is to calculate all possible target association probability of each observation, and think all the effective echo is probably originated from every specific goals, they're just different is the result of different target probability.

To establish a mixed system for describe the linear state equation and measurement equation:

$$\begin{cases} X^t(k+1) = F^t(k)X^t(k) + W^t(k) & k = 1, 2, 3 \dots \\ Z^t(k) = H^t(k)X^t(k) + V^t(k) & k = 1, 2, 3 \dots \end{cases} \quad (5)$$

The  $X(k)$  and  $Z(K)$  respectively the state and observation vector at time  $k$ ;  $F$ ,  $H$  respectively represents the state transfer matrix and observation matrix at time  $k$ ;  $V(k)$  and  $W(k)$  is the Gauss white noise with zero mean independent of each other.

$$X^t(k/k) = E[X^t(k) / Z^k] = E[X^t \cup_{j=0}^{m_k} \theta_{jt}(k) / Z^k] = \sum_{j=0}^{m_k} \hat{X}_j^t(k/k) \beta_{jt}(k) \quad (6)$$

The  $m_k$  indicates the number of measurements confirmed at time  $k$ ;  $\beta_{jt}(k)$  is the associated probability of the first  $j$  measurement and target  $t$ ,  $\sum_{j=0}^{m_k} \beta_{jt}(k)$ ;  $\hat{X}_j^t(k/k)$  is the resulting state estimation at the time  $k$  for the first  $j$  measurement to target  $t$  by filter.

## Actual data evaluation and conclusion

The performance evaluation using the two sets of real radar data, respectively by a coast radar and a radar for air control.

Test target state: the speed of 500 km, the turning acceleration is 1g.

The use of actual testing process, including NNDA, PDA, JPDA, NNPDPA etc. its evaluation purpose is to test a variety of data association method of performance, and compares them.

The first set of data is obtained from coastal radar, which is mainly used for long range air

defense and navigation. During the actual test, the working states of the radar are as follows:

The scanning speed: 5 r/min; Frequency: 1215 ~ 1400 MHZ; Distance: 410 km; Detection probability: 90%; Pulse width: 2 $\mu$ s; Beam width: 2 $^\circ$ .

In the first set of data, we define the following performance metrics:

$N_T$ : the number of the real track confirmation.

$N_F$ : the number of the false track confirmation.

$L_T$ : track to scan number representation of the duration, or track life.

$L_F$ : false track life to scan number representation.

$R_{MC}$ : error rate (error: association ratio association number and track life).

$N_A$ : all confirmed track number.

$L_A$ : all tracks life.

$N_{TCT}$ : the termination of track number.

$TE$ : the execution time.

The first group of the experimental data is listed in Table 1, see the table below:

Table 1 Using the Actual Radar Data to Evaluate Various Data Association Method (1)

Methods	$N_T$	$N_F$	$L_T$			$L_F$		$R_{MC}$	$TE$	
			Min	Max	Average	Max	Average		Max	Average
NNDA	36	14	4	91	60.33	59	14.08	0.0313	0.2167	0.0980
PDA	38	13	4	91	58.31	59	13.14	0.0306	0.1833	0.0987
JPDA	38	13	4	91	58.42	59	13.14	0.0307	0.2000	0.1013
NNPDA	36	14	4	91	59.84	59	13.14	0.0313	0.1730	0.0957

Through the experimental data we can see:

(1) Due to the calculated amount of JPDA and NNPDA is relatively large, so the completion of the computation time compared with other methods is much longer.

(2) From the tracking duration, NNDA and NNPDA tracking duration is long.

(3) Listed in the table is the result of all sorts of data association method, it has the similar performance, this is because the actual track is formed by the radar data of wave gate slightly overlapping. The method JPDA is closer to the tracking quality, but the processing time required for the longest, while NNDA required the shortest processing time.

(4) The PDA method has better ability to suppress clutter, generates less false trace points.

The second set of data is obtained from air traffic control radar. In the test and the data acquisition stage, the radar working states are as follows:

Scanning speed: 12 r/min; Frequency: 1300 MHZ; Distance: 150 km; Detection probability: 80%; Pulse width: 2 $\mu$ s; Beam width: 2 $^\circ$ .

The second group of the experimental data is listed in Table 2, see the table below:

Table 2 Using the Actual Radar Data to Evaluate Various Data Association Method (2)

Methods	$N_A$	$N_{TCT}$	$L_A$			$R_{MC}$	$TE$	
			Min	Max	Average		Max	Average
NNDA	36	29	2	0.360	93.08	0.2324	0.0500	0.0163
PDA	78	72	2	0.119	50.00	0.1990	0.0500	0.0182
JPDA	57	51	2	0.250	60.73	0.2029	7887	97.36
NNPDA	40	34	2	0.360	78.93	0.2122	0.0500	0.0145

Through the experimental data we can see:

(1) The processing time of JPDA is very long, and compares to other several kinds of data association method, its time cost is about 2000 times that of the other.

(2) No matter use what kind of method, almost all of the clutter is removed.

(3) NNDA and NNPDA has better tracking performance, NNDA has the longest track life and minimum breakpoint, but by contrast has produced larger correct rate; JPDA has the least error rate, but the processing time required for too long.

Considering the performance index of various kinds of data association method by using two groups of data, it should be said that JPDA is the best, NNDA also has a more outstanding

performance; in addition to the processing time of JPDA requires a longer processing time while the other has a smaller processing time.

## **Conclusion**

In this paper, we introduced several common methods of data association, and summarizes the main idea and the basic principle of data processing, and also summarized respectively. On this basis, in order to further understand their performance, using the several methods for processing in the experiment, and do data records, through the comparison to the experimental data, we investigated the performance of the various methods.

## **References:**

- [1] Ma Jie, Tian Lifeng. The Review of data association [J]. Journal of communication university of China natural science edition, 2006
- [2] Luan Zhuhui. Nearest neighbor joint probability data association algorithm in radar tracking
- [3] Groves P D. GNSS Inertial and Multisensor Integrated Navigation Systems[M]. Norwood, MA: Artech House Inc, 2008.
- [4] Zhang Quan, Zhang Shu White Features Beidou navigation system and the command and control system application [J] information with the computer, 2011 (3): 165-166.
- [5] Julier S, Uhlmann J. A General Method for Approximating Nonlinear Transformations of Probability Distributions[R]. Robotics Research Group, Department of Engineering Science, University of Oxford, 1996. 136-147.
- [6] Julier S, Uhlmann J. Unscented Filtering and Nonlinear Estimation[J]. PROCEEDINGS OF THE IEEE, 2004, 92(3): 401-422.
- [7] Julier S. The Scaled Unscented Transformation[C]. In Proceedings of the American Control Conference, 2002, 6(5): 4555-4559.