

## 测向交叉定位系统中的 K-means 聚类融合算法

孙鹏, 熊伟

(海军航空工程学院信息融合研究所, 山东烟台 264001)

**摘要:** 利用多个无源观测站对多个目标进行测向交叉定位时,会产生大量的虚假交叉点。针对这个问题,提出一种新的基于改进 K-means 的聚类融合定位算法。算法对每条测向线上的交叉定位点进行聚类,获得每条测向上的目标位置估计,从而可以剔除大部分虚假交叉点,然后对各个观测站的聚类结果进行分步融合,充分利用各个观测站的聚类结果,进一步消除残余虚假交叉点的影响,最终获得目标的估计位置。仿真结果证明了该算法在进行多站交叉定位时具有更好的定位效果和鲁棒性。

**关键词:** 测向交叉定位; K-means 聚类; 虚假点; 数据融合

**中图分类号:** TN973      **文献标志码:** A      **文章编号:** 1671-637X(2016)10-0036-05

## K-means Cluster and Fusion Algorithm for Passive Bearing-Crossing Localization System

SUN Peng, XIONG Wei

(Research Institute of Information Fusion, Navy Aeronautical and Astronautical University, Yantai 264001, China)

**Abstract:** There are a large number of false intersection points when using multiple passive observation stations for making bearing-crossing localization to multiple targets. To solve the problem, a new algorithm is proposed, which is based on improved K-means cluster and data fusion. The intersection points on the same direction line are clustered, so that the targets in the direction line can be estimated. In this way, the most of false intersection points are eliminated. And then the results of clustering obtained from different stations are fused step by step. In the process of fusion, the algorithm makes full use of the clustering result of each observation station to reduce the influence of the residual false intersection points. Therefore, more accurate results of target location are obtained. The result of simulation show that the algorithm has better locating performance and robustness in multi-station cross positioning.

**Key words:** bearing-crossing localization; K-means cluster; false intersection point; data fusion

### 0 引言

测向交叉定位是伴随着电子对抗活动而产生的,早在 20 世纪 40 年代,人们就已经开始运用测向设备对目标进行多次测向,并通过交叉定位获得目标位置信息<sup>[1]</sup>。测向交叉定位凭借其作用距离远、不易受干扰等优势成为无源定位中运用最广泛、机制最成熟、研究时间最长的一种定位方法。随着应用环境的不断复杂,定位要求的不断提高,国内外学者对测向交叉定位的研究一直没有停止,如今多站多目标进行交叉定位

仍是研究的一大热点<sup>[2-4]</sup>。

在多站多目标测向交叉定位过程中,随着目标数量的增加,虚假交叉点数量会急速增加,如何对虚假交叉点进行快速、高效的剔除是测向交叉定位所面临的主要难点问题。针对这个问题,许多学者进行了研究<sup>[5-9]</sup>,现有方法大多是利用不同的聚类方法对目标位置进行估计,文献[6]提出一种基于最小距离的二次聚类法,通过对交叉点进行两次聚类,排除虚假点,从而获得目标的估计位置;文献[7]借鉴了多维分配的思想,结合整体相关和最近邻方法对多个观测站的测向数据进行降维处理,并通过 $\chi^2$ 分布对观测数据进行关联,从而达到剔除虚假交叉点的目的,实现对目标的定位跟踪;文献[8]首先利用基于聚类 MEM 模型剔除一部分静态虚假交叉点并确定目标数量,然后建立

收稿日期:2015-08-28

修回日期:2015-10-16

基金项目:国家自然科学基金(61471383)

作者简介:孙鹏(1990—),男,山东烟台人,硕士,研究方向为信息融合技术。

航迹树,通过动态跟踪的方法剔除剩余的虚假交叉点。上述方法在一定程度上减少了虚假交叉点的影响,但当测角误差较大而目标之间距离较近时,由于真实交叉点与虚假交叉点之间距离较近,这些方法都会遇到一个共同的问题:若门限设定较小,会遗漏目标;若门限设定较大,则虚假交叉点的消除效果太差。为解决上述问题,本文以三站交叉定位为例,针对多站交叉定位中虚假交叉点较多的问题,提出一种新的定位方法,并进行了仿真验证。

## 1 问题描述

当利用多个观测站对多个目标进行测向交叉定位时,由于存在观测误差并且无法获得径向距离信息,因此无法直接区分虚假交叉点和真实交叉点<sup>[10]</sup>。针对这个问题,现有的方法主要是通过一定的方法对相对集中的交叉点进行聚合,从而估计出目标的数量与位置<sup>[6-9]</sup>,如图 1 所示。图 1 中以 3 个固定观测站和 3 个目标作为示例,从图中可以看出,在目标距离观测站较近或者目标之间距离较远时,通过这种方法可以快速获得目标的位置估计。

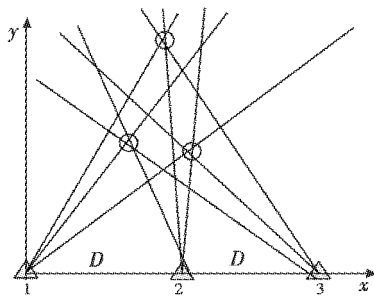


图 1 交叉定位模型

Fig. 1 The model of cross bearing

随着目标之间距离的缩小,观测站获得的测向线分布更为密集,由测向线交叉产生的交叉点分布也更加集中,使得真实交叉点与虚假交叉点之间的距离更近,模糊度增大,此时,若仍通过之前的方法对交叉点进行聚合,则得到的目标位置与数量均会有较大偏差。如图 2 所示。

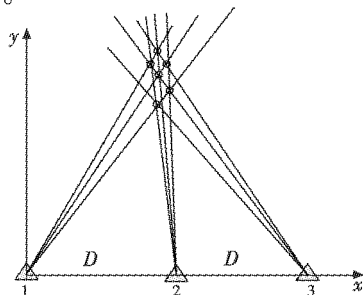


图 2 密集多目标交叉定位模型

Fig. 2 The locating model of dense targets

图中是由 3 个目标产生的测向线,而利用之前方法进行聚类时,获得的定位点远多于 3 个(图中仅为示意),若目标距离进一步靠近,聚类时的模糊度也更大,获得的目标数目估计与位置估计也会存在较大的偏差,因此需要有新的方法,针对性解决这类问题。

## 2 基于 K-means 的聚类融合算法

针对上文所提到的密集多目标在交叉定位中遇到的问题,本文通过对密集多目标的定位交叉点进行特征分析,针对传统交叉定位方法在对密集多目标的定位中所遇到的问题,提出一种新的定位算法。

新算法改变了传统对所有定位交叉点进行聚类的思想,在聚类时以每条测向线为基准对该测向线上的交叉点单独聚类,这不仅减少了其他交叉点对聚类的影响,而且在不考虑两个目标测向线有重合和漏检的情况下,由于每条测向线都对应且仅对应一个目标,算法还可解决传统定位方法中目标数目模糊的问题。为充分利用各个观测站的方位信息,新算法在聚类后对结果进行了分步融合,减少了偶然误差对结果的影响。算法过程主要分为两步。

### 2.1 基于测向线的分簇聚类方法

基于测向线的聚类有效避免了其他测向线交叉点对聚类的影响,将每次需要聚类的点数从  $(M-1) \cdot MN^2/2$  个缩减至  $(M-1)N$  个(其中,  $N$  为目标数量,  $M$  为观测站数量),很大程度上减少了不同测向线上虚假点之间的相互影响。

下面以观测站 1 为例,具体介绍聚类过程。

1) 设观测站 1 的第  $i$  条测向线与观测站 2 的各条测向线的交叉点  $X_{ij} = [\hat{x}_{ij}, \hat{y}_{ij}]'$ ,  $j = 1, 2, \dots, N$ , 其与观测站 3 的各条测向线交叉点为  $X_{il} = [\hat{x}_{il}, \hat{y}_{il}]'$ ,  $l = 1, 2, \dots, N$ ,  $N$  为目标数。分别求得  $X_{ij}$  中每个位置元素与  $X_{il}$  中每个位置元素的欧式距离为

$$R = \sqrt{(\hat{x}_{ij} - \hat{x}_{il})^2 + (\hat{y}_{ij} - \hat{y}_{il})^2} \quad R \leq \gamma \quad (1)$$

其中,  $\gamma$  为一个门限值,将两个交叉点归为一类,并求得两点的中心位置为

$$X_u = \frac{X_{ij} + X_{il}}{2} = \left[ \frac{\hat{x}_{ij} + \hat{x}_{il}}{2}, \frac{\hat{y}_{ij} + \hat{y}_{il}}{2} \right] \quad (2)$$

式中,  $t$  为聚类编号。将  $X_{it}$  定义为初始聚类中心。

2) 计算每个未聚类交叉点与聚类中心的距离,当该交叉点与聚类中心的距离有小于门限的情况时,将其归于与聚类中心距离最小的一类。

3) 根据新的分类情况,重新求得聚类中心

$$X_u = \left[ \frac{\hat{x}_{ij} + \hat{x}_{il} + \hat{x}_{ik} + \dots}{n}, \frac{\hat{y}_{ij} + \hat{y}_{il} + \hat{y}_{ik} + \dots}{n} \right] \quad (3)$$

式中:  $(\hat{x}_{ik}, \hat{y}_{ik})$  表示新归入该类的交叉点坐标;  $n$  表示

在这一类中交叉点的总数。

4) 将新获得的聚类中心与上一步的聚类中心进行比较,若两者一致,聚类结束;否则,返回步骤2)。

5) 当最终获得的聚类数大于1时,由于在静态聚类的过程中,无法进一步区分真实目标与虚假点,因此将其中交叉点数量最多一类的聚类中心当作目标位置的估计值。

通过上述聚类过程可获得目标的初步估计位置。

## 2.2 聚类点的分步融合

通过 K-means 方法对各观测线上的交叉点进行聚类后,由于最终得到的聚类结果中往往包含一定数量的虚假交叉点,受虚假交叉点的影响,所获得的聚类中心位置可能与真实位置有一定的偏差,而每个观测站在对同一个目标进行聚类时,所使用的聚类点是不一样的,因此所得聚类点与真实位置的偏差也不一致。为获得更为精确的定位结果,可对各个观测站的聚类结果进行融合<sup>[11]</sup>。

设观测站  $i$  在  $k$  时刻对目标  $t$  的定位结果为  $X_i^t(k|k)$ , 设此时目标的协方差矩阵为  $P_i^t(k|k)$ , 当将观测站 1 的定位结果与观测站 2 的定位结果进行融合时,可以求得其协方差为

$$P_u = \begin{bmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \end{bmatrix} \begin{bmatrix} \sigma_{\theta_i}^2 & 0 \\ 0 & \sigma_{\theta_j}^2 \end{bmatrix} \begin{bmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \end{bmatrix}^T \quad (4)$$

式中,  $\sigma_{\theta_i}^2$  和  $\sigma_{\theta_j}^2$  分别为观测站 1 和观测站 2 的测角误差;且

$$d_{11} = \frac{D_1 \sec^2 \theta_i}{(\tan \theta_i - \tan \theta_j)^2} \quad (5)$$

$$d_{12} = \frac{D_2 \sec^2 \theta_j}{(\tan \theta_i - \tan \theta_j)^2} \quad (6)$$

$$d_{21} = \frac{D_1 \sec^2 \theta_i \tan \theta_j}{(\tan \theta_i - \tan \theta_j)^2} \quad (7)$$

$$d_{22} = \frac{D_2 \sec^2 \theta_j \tan \theta_i}{(\tan \theta_i - \tan \theta_j)^2} \quad (8)$$

$\theta_i$  为通过观测站 1 测得的当前定位结果的方位角;  $\theta_j$  为通过观测站 2 测得的当前定位结果的方位角;另外

$$D_1 = (x_{s2} - x_{s1}) \tan \theta_j + (y_{s1} - y_{s2}) \quad (9)$$

$$D_2 = (x_{s1} - x_{s2}) \tan \theta_i + (y_{s2} - y_{s1}) \quad (10)$$

式中:  $(x_{s1}, y_{s1})$  为观测站 1 的位置坐标;  $(x_{s2}, y_{s2})$  为观测站 2 的位置坐标。

以观测站 1 为主站, 计算  $k$  时刻观测站 1 的第  $t_1$  ( $t_1 \in \{1, 2, \dots, N\}$ ) 个定位结果  $X_{t_1}^1(k|k)$  与观测站 2 的第  $t_2$  ( $t_2 \in \{1, 2, \dots, N\}$ ) 个定位结果  $X_{t_2}^2(k|k)$  之间的统计距离

$$\alpha_{t_1 t_2}(k) = [X_{t_1}^1(k|k) - X_{t_2}^2(k|k)]^T \cdot$$

$$[P_{t_1}^1(k|k) + P_{t_2}^2(k|k)]^{-1} [X_{t_1}^1(k|k) - X_{t_2}^2(k|k)] \quad (11)$$

如果

$$\alpha_{t_1 t_2}(k) \leq \delta \quad (12)$$

则判断观测站 1 的第  $t_1$  个定位结果与观测站 2 的第  $t_2$  个定位结果来自同一个目标, 其中  $\delta$  为一个门限值, 根据式(13)、式(14)对两个观测站的定位结果进行融合

$$X_{t_1 t_2}(k|k) = P_f^i(k|k) [P_{t_1}^1(k|k)^{-1} X_{t_1}^1(k|k) + P_{t_2}^2(k|k)^{-1} \hat{X}_{t_2}^2(k|k)] \quad (13)$$

$$P_f^{12}(k|k) = [P_{t_1}^1(k|k)^{-1} + P_{t_2}^2(k|k)^{-1}]^{-1} \quad (14)$$

若观测站 1 的一个定位结果与观测站 2 的多个定位结果均关联成功, 则将  $\alpha_{t_1 t_2}$  最小的一组定位结果进行融合。

同理, 根据式(11)计算观测站 1 的各个定位结果与观测站 3 各个定位结果的统计距离  $\alpha_{t_1 t_3}$ , 若满足式(12), 则利用式(13)与式(14)将定位结果进行融合得到融合后的位置  $X_{t_1 t_3}(k|k)$  与协方差矩阵  $P_f^{13}(k|k)$ 。然后将两次融合后的结果再次通过式(11)与式(12)进行判决, 关联成功的通过式(13)与式(14)获得最终融合结果。

算法的实现过程如图 3 所示。

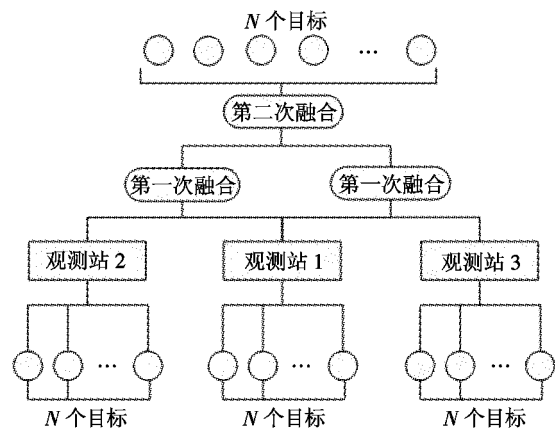


图3 分步融合模型

Fig. 3 The model of stepping fusion

## 3 仿真实验

仿真时, 在矩形区域内随机生成多个运动目标, 矩形区域坐标分别为  $(40, 200)$ ,  $(80, 200)$ ,  $(40, 300)$ ,  $(80, 300)$ , 单位为 km, 目标之间的距离在 5 ~ 10 km 之间。使用 3 个固定观测站对上述目标进行定位跟踪, 3 个观测站的位置坐标分别为  $(50, 0)$ ,  $(60, 0)$ ,  $(70, 0)$  处, 单位为 km。观测站的测角误差服从均值为零的高斯分布, 且 3 个观测站的测角误差之间相互独立, 采样间隔为 1 s, 观测时间为 200 s。

由于在观测过程中始终存在测量误差, 且误差服从均值为零、方差为  $\sigma^2$  的高斯分布, 因此最终获得的

目标位置应分布在真实位置附近,其与真实位置之间的距离应服从均值为零、方差为 $(R\sigma)^2$ 的高斯分布,其中, $R$ 为目标与观测站之间的距离。为直观表现算法的定位性能,本文定义正确定位门限 $\gamma = R\sigma$ ,当获得的目标位置与真实位置之间的距离小于门限 $\gamma$ 时,可以认为已经获得正确定位结果。正确定位次数的多寡,可以反映出算法定位性能的优劣。

### 3.1 仿真一

在测量误差 $\sigma = 0.5^\circ$ 的情况下,使用本文算法与基于聚类 MEM 算法对 4 个目标进行定位,通过 200 次 Monte Carlo 仿真实验,得到两种算法的正确定位次数,如表 1 所示。

表 1 两种算法对各个目标定位性能比较

Table 1 The locating performance of two algorithms

	目标 1	目标 2	目标 3	目标 4	平均
本文算法	186	175	182	179	180.5
基于聚类 MEM 算法	161	151	146	158	154

通过表 1 可以看出,本文算法对目标的平均正确定位次数达到 180.5 次,基于聚类 MEM 算法的平均正确定位次数仅为 154 次,相比于基于聚类 MEM 算法,本文算法对目标的定位更为准确。

为更直观地展现两种算法的定位效果,任意取出一个周期中两种算法的定位结果,图 4 给出第 30 个周期两种算法的定位结果。

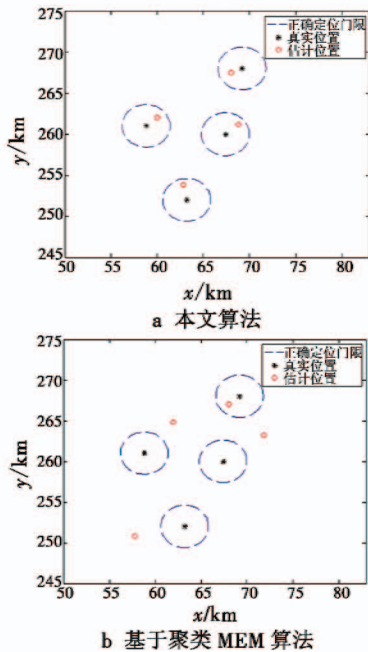


图 4 两种算法的定位结果

Fig. 4 The locating result of two algorithms

图 4 中,圆形区域代表认定的正确定位区域,估计位置位于圆形区域内则认为已经进行了正确的定位。在第 30 个周期的定位中,本文算法对目标均可以进行

正确的定位,而基于聚类 MEM 算法产生较大的定位偏差,由此可见,本文算法具有更好的收敛速度。这主要是由于本文算法的聚类是针对每条测向线上的交叉点进行的,相比于基于聚类 MEM 算法,大大减少了需要聚类的交叉点数量,从而减少了不必要的干扰,并在融合的过程中充分利用各个观测站的信息,进一步消除了由于单个观测站产生的偶然误差。

### 3.2 仿真二

为检验算法对密集多目标的定位性能,改变 4 个目标之间的距离,通过 200 次 Monte Carlo 仿真实验,比较两种算法对不同间距目标的定位性能,其中,测量误差 $\sigma = 0.5^\circ$ ,算法的平均正确定位次数如表 2 所示。

表 2 两种算法对不同间距目标定位性能比较

Table 2 Performance of two algorithms in locating targets with different spacing

	目标间隔/km			
	3	5	10	15
本文算法	158	175	189	195
基于聚类 MEM 算法	98	146	171	192

通过表 2 可以看出,当目标之间距离较大时,两种算法都可对目标进行准确的定位,且速度较快,这是因为目标之间距离较大,交叉点分布较为分散,在定位过程中,可以较为容易地剔除虚假点。随着目标间距的缩小,交叉点分布也更为集中,基于聚类 MEM 算法是对所有交叉点进行聚类,虚假点对算法性能的影响更大,虽然其后期通过动态跟踪的方法进一步去除“重影”,但由于目标之间距离较近,建立各自独立的航迹树也需要较长时间。而本文算法在此情况下,是对每条测向线上的交叉点进行聚类,大大减少了其他虚假交叉点的干扰,且在融合过程中,充分利用 3 个观测站的信息,因此可以更为准确地对目标进行定位。由此可见,在对密集多目标进行定位时,本文算法具有更为明显的优势。

### 3.3 仿真三

为检验算法的鲁棒性,设定不同的测量误差,通过 200 次 Monte-Carlo 仿真实验,比较两种算法对不同间距目标的定位性能,目标间距设置为 5 km,比较结果如表 3 所示。

表 3 两种算法在不同测量误差时的定位性能比较

Table 3 Performance of two algorithms under different measuring errors

	测量误差 $\sigma$			
	0.1°	0.3°	0.5°	1°
本文算法	193	190	177	161
基于聚类 MEM 算法	190	178	145	120

由表 3 可以看出,在测量误差 $\sigma = 0.1^\circ$ 的情况下,两种算法的定位性能接近,随着测量误差的增大,本文算法的优势逐渐体现,当测量误差 $\sigma \geq 0.5^\circ$ 时,本文算

法的定位性能明显优于基于聚类 MEM 算法,这主要是由于随着测量误差的增大,目标的真实交叉点分布更为分散,与虚假交叉点之间的模糊度更高,导致基于聚类 MEM 算法在第一步聚类时会产生较大的偏差,且对目标数目估计不准确,这使得在第二步动态跟踪时需要较长时间去除虚假航迹,而本文算法通过融合 3 个观测站的聚类信息,可很大程度上减少观测误差带来的影响。由此可见,在测量误差较大的情况下,本文具有更高的定位精度。

#### 4 结 语

本文针对多站交叉定位中存在的虚假点多的情况,提出一种新的定位跟踪方位,通过 K-means 聚类方法以测向线为基准对每条测向线上的交叉点进行聚类,减少了聚类时的交叉点数量且可准确估计目标数量;通过对 3 个观测站聚类结果的分步融合,综合 3 个观测站信息,减少了单一观测站偶然误差的影响,有效解决了多个观测站对多目标定位难的问题。仿真结果表明:相比于现有算法,本文算法在对密集多目标进行定位时,具有更好的定位效果和更快的收敛速度,且具有更强的环境适应能力。

#### 参 考 文 献

- [1] 胡来招. 无源定位[M]. 北京:国防工业出版社,2004. (HU L Z. Passive location[M]. Beijing:National Defense Industry Press, 2004. )
- [2] 何友,修建娟,关欣. 雷达数据处理及应用[M]. 3 版. 北京:电子工业出版社,2013:284-287. (HE Y, XIU J J, GUAN X. Radar data processing and application[M]. 3rd ed. Beijing:Publishing House of Electronics Industry, 2013:284-287. )
- [3] 修建娟,何友,王国宏,等. 测向交叉定位系统中的交叉角研究[J]. 宇航学报,2005, 26(3):282-286. ( XIU J J, HE Y, WANG G H, et al. Study on cut angle of direction finding location system[J]. Journal of Astronautics, 2005, 26(3):282-286. )
- [4] 沈晓峰,徐保根,邹继锋,等. 基于正性约束的测向交叉定位方法[J]. 电子科技大学学报,2014, 43(6):834-837. (SHEN X F, XU B G, ZOU J F, et al. A positive restriction method used in crossing location[J]. Journal of University of Electronic Science and Technology of China, 2014, 43(6):834-837. )
- [5] 张旻,罗争. 一种基于 RSS 估计的虚假点消除方法[J]. 电子学报,2012, 40(10):2117-2121. (ZHANG M, LUO Z. A new method for eliminating ghost based on RSS estimation[J]. Acta Electronica Sinica, 2012, 40(10):2117-2121. )
- [6] 蒋维特,杨露菁,杨亚桥. 测向交叉定位中基于最小距离的二次聚类算法[J]. 火力与指挥控制,2009, 34(10):25-28. (JIANG W T, YANG L J, YANG Y Q. Quadratic clustering algorithm based on least distance in DOA location[J]. Fire Control and Command Control, 2009, 34(10):25-28. )
- [7] 修建娟,何友,车志宇,等. 三维测向无源定位中的数据关联算法研究[J]. 系统工程与电子技术,2009, 31(2):329-332. ( XIU J J, HE Y, CHE Z Y, et al. Study on bearing measurements association algorithm in 3D passive location systems[J]. Systems Engineering and Electronics, 2009, 31(2):329-332. )
- [8] 刘梅,权太范,姚天宾,等. 多传感器多目标无源定位跟踪算法研究[J]. 电子学报,2006, 34(6):991-995. (LIU M, QUAN T F, YAO T B, et al. Multi-sensor multi-target passive locating and tracking[J]. Acta Electronica Sinica, 2006, 34(6):991-995. )
- [9] 丁文锐,李红光,李新军. 基于聚类最近数据关联的多目标跟踪算法[J]. 北京航空航天大学学报,2009, 35(12):1487-1490. (DING W R, LI H G, LI X J. Multi-target tracking based on clustermin-distance data association[J]. Journal of Beijing University of Aeronautics and Astronautics, 2009, 35(12):1487-1490. )
- [10] 辛云宏,杨万海. 被动多站多目标的测量数据关联算法研究[J]. 宇航学报,2005, 26(6):748-752. (XIN Y H, YANG W H. A method of the passive multi-sensor multi-target measurement data association[J]. Journal of Astronautics, 2005, 26(6):748-752. )
- [11] 陈世明,陈小玲,徐青刚,等. 基于自适应加权融合的分分布式滤波算法[J]. 控制与决策,2015, 30(4):605-610. (CHEN S M, CHEN X L, XU Q G, et al. Adaptive weighted fusion based linear distributed filtering algorithm[J]. Control and Decision, 2015, 30(4):605-610. )