The Associated Algorithm of Target Batch Number Based on Gaussian Mixture Clustering

Ning Liu(✉), Shufen Liu(✉), and Xinjia Zhang(✉)

Software College, Jilin University, Changchun, China
frank93@sina.cn, {liusf,zhxj}@jlu.edu.cn

Abstract. This article provides some questions about the large quantities of target track information, unclear sensor batch number, etc. referring to target fusion, and provide the clustering algorithm of Gaussian Mixture Distribution to realize the division of all the batch numbers of target track and associated operation. According to the processing to the Gaussian distribution of information by maximum likelihood estimation and through continuous iterative and refinement of the clusters divided, the tracking information detected by all radar sensors of each target batch number can be obtained precisely in the end. This algorithm provides rapid and precise working conditions for target fusion.

Keywords: Gaussian mixture distribution · Associated operation Maximum likelihood estimation · Clustering algorithm

1 Introduction

When the formation of ship combat synergistically, naval and radars in the fleet will report all the detected objectives about the main naval, which will create influences on the warship equipment attacked precisely according to the information.

Not only this, the objectives detected by radars can only feedback its physical information and can not precisely judge the target batch number defined by human. In the current literature, radars' objective fusion rarely have research, which aims at the two sides. In order to simplify the target quantity, avoid unnecessary redundant target information. The associated algorithm of target batch number in this article will connect the prototype cluster in the ensemble learning to solve the related problems between target information and target batch number.

2 The Data Analysis

The command system of formation of ship receives periodically signals from different radars in formation. Here it is set as M pieces of sensors, with each sensor as m. It can periodically make up instructions and send N pieces of tracking information processed by itself and under the unified coordinate system. Here the ensemble of communication of multiple batches of tracking information sent by No.m sensor is set as D_m and among

Q. Zu and B. Hu (Eds.): HCC 2017, LNCS 10745, pp. 23–27, 2018.
https://doi.org/10.1007/978-3-319-74521-3_3

it, some batch of tracking information is d_{mn}(PS: m = 1, 2,…… as the tracking information detected by NO.m sensor. n = 1, 2…… as the NO.n target detected).

3 Target Batch Correlation

Target batch correlation is the basis of target fusion. Exactly differentiating DS is the precondition of fusion algorithm accuracy. In the war of formation of ship, because it is very difficult to recognize the target without exactly differentiating a large amount of tracking information detected, the accuracy of the algorithm is very important.

Firstly define some basic knowledge. As it is mentioned earlier, target tracking information includes a great amount of data, such as a sensor' serial number, target nationality, target type, longitude and latitude, height and a series of data. Here the number of attributes is set as k and set $d_{mn} = \{x_{mn1}, x_{mn2}, \ldots, x_{mnk}\}$, the NO.n target tracking information detected by NO.m radar as a k-dimensional vector. x_{mnk} is the value of NO.k of a some batch tracking information (such as, the above target type's value on x_{ijk} is "the air"), while k is the dimension of the tracking information vector.

The result of batch correlation is a list, with a batch number on each line. Batch numbers correspond to the tracking information collection under the target detected by the related sensor. The collections can not be controlled strictly or intersected and the parallel operation of these collections is D, which is all tracking information detected by all radars. Here we call every subset as a "cluster". Through this divide, each cluster stands for tracking information collection of the same target detected by the radar sensor. After the target batch correlates with tracking information, the list formed by all clusters report to fusion center to perform blend operation.

It is formally said that D can be divided into h pieces of disjoint clusters $\{C_i | i = 1, 2\ldots, h\}$, with $C_i \cap C_{i'} \neq \varnothing$. Accordingly, each cluster's marker is set as the target batch number. We use $\lambda = \{\lambda_1, \lambda_2, \ldots, \lambda_h\}$ to represent tracking information clusters' marker, including the value of λ_h as the target batch number.

Above we introduce is some basic knowledge about association algorithm. The association algorithm the article studies will be researched by connecting clusters under unsupervised learning. We just need to divide the D collection formed by all tracking information into multiple joint clusters. But we need to solve the problem of distance calculation.

3.1 Gaussian Mixture Clustering Algorithm

We know that Gaussian distribution is fixed by two parameters among it. The two parameters are mean vector μ and covariance matrix G. We define Gaussian mixture distribution first:

$$PM(x) = \sum_{i=1}^{h} \alpha_i p\left(x | \mu_i, G\right) \tag{3.1.1}$$

Among it, the latter part $p(x|\mu_i, G)$ is the probability density of Gaussian distribution. Suppose that all tracking information is produced by Gaussian mixture distribution, the specific process is to fix different Gaussian mixture distribution probability density according to $\alpha_1, \alpha_2, \ldots, \alpha_h$. After fix the probability density function, put different tracking information in different Gaussian distributions. In here we can see that Gaussian mixture distribution is the mixed composition of multiple Gaussian distribution. Each Gaussian distribution's probability is controlled by a parameter α_h. The probability is how much the the size is about some tracking information has been specifically related to one Gaussian distribution.

It can be seen that three parameters α_i, μ_i, G among it should be fixed if defining Gaussian mixture distribution. Suppose that all tracking information has been completed the correlation, how to fix the target batch number of some tracking information, which means how to fix cluster marker of the tracking information cluster some tracking information is in. Suppose that the tracking information cluster of some tracking information d_{ij} is in is z. In Gaussian mixture distribution, z_{mn} is the Gaussian distribution some sample is in. The value range is 1~h. But we have no way to know its specific value. According to Bayes' theorem, posteriori distribution of z_{ij} is

$$\begin{aligned} PM\left(z_{mn} = h|d_{mn}\right) &= \frac{p\left(z_{mn} = h\right) \cdot PM\left(d_{mn}|z_{mn} = h\right)}{PM\left(d_{mn}\right)} \\ &= \frac{\alpha_i \cdot p(d_{mn}|\mu_i, G)}{\sum_{l=1}^{h} \alpha_l \cdot p(d_{mn}|\mu_i, G)} \end{aligned} \tag{3.1.2}$$

Which means we have got the posteriori probability of each tracking information cluster d_{mn} is in, marked as γ_{mnz}. When we want to know the batch number of some target tracking information, it is only to find the max one the d_{mn} corresponds to.

3.2 Maximum Likelihood Estimation

To all tracking information collection D, perform maximum likelihood estimation.

$$LL(D) = \ln\left(\prod_{j=1}^{m} PM(x)\right) = \sum_{j=1}^{m} \ln\left(\sum_{i=1}^{h} \alpha_i \cdot p(d_{mn}|\mu_i, G)\right) \tag{3.2.1}$$

Maximum likelihood estimation is to maximize the above formula. The derivative is 0 with the maximum value. The partial derivatives of three parameters are 0 through the formula 3.2.3.

Firstly, get out the partial derivative to μ_i through the formula 3.2.3, with the values of derivatives as 0, and according to the formula 3.2.2, get out

$$\mu_i = \frac{\sum \gamma_{mnz} d_{mn}}{\sum \gamma_{mnz}} \tag{3.2.2}$$

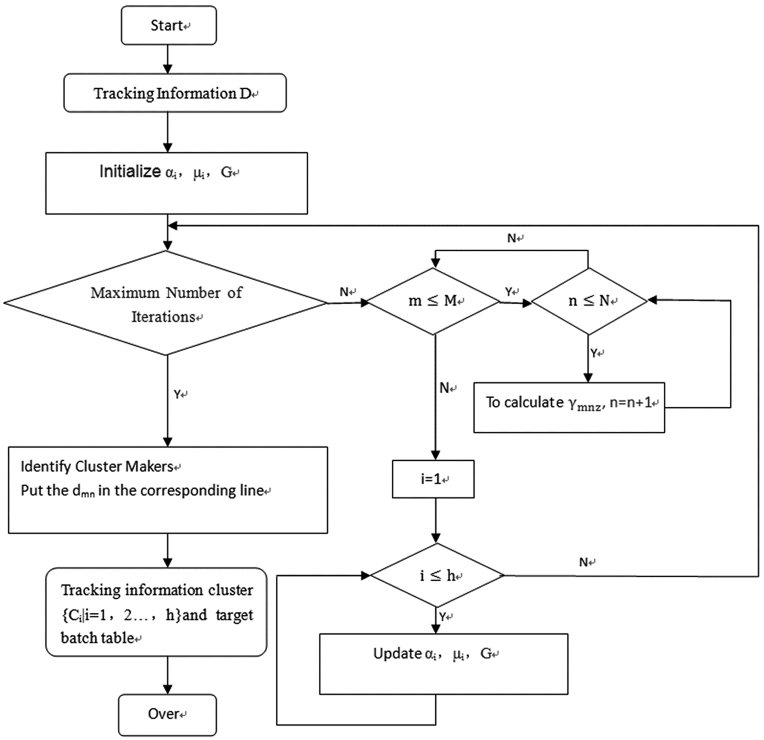
Similarly, through the derivation of G, get out of

$$G = \frac{\sum \gamma_{mn}\left(d_{mn} - \mu_i\right)\left(d_{mn} - \mu_i\right)^T}{\sum \gamma_{mnz}} \tag{3.2.3}$$

Because α_i is the related probability of target batch selected by tracking information, the sum of $\alpha_1, \alpha_2, \ldots, \alpha_h$ should be 1, which all are greater than and equal to 0. According to the lagrange formalism of maximum likelihood estimation, perform α_i partial derivatives and get the values with the partial derivative as 0, get out

$$\alpha_i = \frac{1}{mn} \sum \gamma_{mnz} \tag{3.2.4}$$

After get every parameter's calculation method and use iterative process of Gaussian mixture clustering, which means getting posterior probability of each Gaussian distribution through current parameters and current tracking information. Then through the formula, update three parameters till reach the maximum value of iterative or LL(D) with a little improvement or no improvement, stop updating the parameters and finally make sure the tracking information included among every tracking information. And the corresponding marker of every tracking information is target batch. In this way, the overview that flow the target tracking information relates is shown in following figure.



4 Justification and Summary

We define a measure that evaluates the clarity of the tracking information to prove the advantages of this algorithm. Davies-Bouldin index:

$$DB = \frac{1}{h}\sum_{i=1}^{h}\max_{j \neq i}\left(\frac{avg(C_i) + avg(C_j)}{d_{cen}(\mu_i, \mu_j)}\right)$$

Among them, $d_{cen}(\ldots)$ is to calculate the distance between the center of two clusters, avg(C) is the average distance of each trace information in the cluster C. So, we can see, the denominator in the DB index is the distance between the two tracking information clusters, the larger the denominator, the lower the similarity of the two clusters, and the molecules represent the tracking information aggregation in two clusters. So, the smaller the value of the DB index, the better the result of the algorithm. In a tracking information data set of 11 ships, the results are obtained 0.32~0.33, it's better than the general information.

The algorithm promoted in this article controls all the tracking information within the k-dimension space, proceeds loop iteration and fix the collection of tracking information detected by multi sensors each target corresponds to, which realizes the connective work of precise target batch number and complete the work within the quadratic time.

References

1. Wan, S.: Clustering of multisensor data fusion method. J. Syst. Eng. Theory Pract. (05) (2008)
2. Xinrui, Z., he, J.: Multi-sensor information fusion technology. Comput. Meas. Control (09) (2007)
3. Hongfeng ailin, W., Single Showers: Research progress of foreign military information fusion theory and application. Lightning Control (04) (2007)
4. Ma, J., Jixiang, S.: Information fusion model of neural network to describe. J. Avionics Technol. (01) (1998)
5. Xin, W., Qidan, Z., Shuli, S.: Unfettered global optimal weighted measurement fusion estimation. Comput. Eng. Appl. (24) (2010)
6. Hu, W., Shenyuan, Y.: The multi-sensor track correlation. Appl. Sci. Technol. (12) (2005)