2010 First International Conference on Pervasive Computing, Signal Processing and Applications

The Application of Unsupervised Clustering in Radar Signal Preselection based on DOA Parameters

Xiang-Peng Zhu, Ming Jin, Wei-Qiang Qian, Shuai Liu, Yu-Mei Wei
Harbin Institute of Technology
Harbin, China
E-mail：johnny06@126.com

***Abstract*- With the deterioration of electronic environment，using signals DOA (direction of arrival) parameters has great significance to preselect multiple radar pulses. Cluster analysis as an important means of data classification, is gradually applied to radar signal sorting. In this paper, a novel method of signal sorting flowsheet is proposed based on Fuzzy Clustering to sort emitters DOA as data objects, with dynamic clustering for reference and Gaussian distance function instead of Euclidean distance. This method avoids establishing enormous similar matrix and adapt to the change of the number of emitters. The result of simulation demonstrates that this method is effective.**

Keywords-radar signal preselection; direction of arrival; fuzzy clustering

I. INTORDUCTION

As the upgrading of electronic warfare, reconnaissance receiver receives the forms of radar signal more changeful and dense; therefore, it is very difficult to sort them. As for Anti-radiation seeker, the target of attack is the relatively fixity of the location of radars on the ground, or there is a regularity. The DOA would be the favorable foundations for radar signal sorting and orientation. However, the intercept and capture of pluses usually lack the number of pluses categories. So radar emitter angles sorting belong to the unsupervised clustering typically. There are important research meanings to solve the unknown categories for radar signal sorting with unsupervised clustering algorithm.

Traditional signal sorting generally uses rectangle evenly division or rectangle non-uniform division [1], however, how to appropriate choose interval length is difficult. Cluster analysis is a typical way of unsupervised classification. With the development of this technique, it has been used in the field of radar signal sorting gradually. In recent years some methods have been used, such as clustering based on neural networks, ,STING algorithm [3]，CURE algorithm [4]and CLIQUE algorithm [5]. Though these methods can sort signals, they all have a fatal disadvantage: the extremely huge calculation task, impossible to achieve real-time requirements, thus failing to be applied to the practical separation system. Besides, the two basic parameters are very difficult to control. So it is necessary to propose a practical algorithm to solve the current problems.

Fuzzy clustering is used in statistics commonly. The cluster process is based on the similarity of the statistics samples. There is no need to set cluster centers and cluster radius. Based on fuzzy clustering, a novel method of signal sorting flowsheet is proposed to sort emitters DOA of elevation angle and azimuth angle mean, with dynamic clustering for reference and Gaussian distance Function instead of Euclidean distance. This method avoids establishing enormous similar matrix, saves the sorting time and adapts to the change of the number.

II. FUZZY CLUSTERING ALGORITHM

Clustering can divide data objects into several classes or clusters based on data objects comparability. Due to the uncertainty of the classification in science and technology, it is more natural and effective to cluster objects with fuzzy mathematic method. It mainly includes several steps as follows:

A. Data standardization

In practice, different data have different dimensions. In order to divide the different dimension data, there will need to compress the data to interval $[0,1]$ so that they have the same dimension classified. Supposing that there are n objects to be classified. Characteristic of each dimension x_k has n original data for each feature

978-0-7695-4180-8/10 $26.00 © 2010 IEEE
DOI 10.1109/PCSPA.2010.236

Characteristics of each dimension. $x_{1k}, x_{2k}, \cdots, x_{nk}$ are referred as the elements of this feature.

In order to standardize these data, firstly, compute:

$$\bar{x}_k = \frac{1}{n}\sum_{i=1}^{n} x_{ik} \tag{1}$$

$$S_k^2 = \frac{1}{n}\sum_{i=1}^{n}(x_{ik} - \bar{x}_k)^2 \tag{2}$$

Do standardization of the standard deviation:

$$x'_{ik} = \frac{x_{ik} - \bar{x}_k}{S_k} \tag{3}$$

Such data x'_{ik} is not always in the interval $[0,1]$.So x'_{ik} is compressed by the extremum standardized formula to interval $[0,1]$.

The extremum standardized formula:

$$x'_{ik} = \frac{x'_{ik} - x'_{k\min}}{x'_{k\max} - x'_{k\min}} \tag{4}$$

Here $x'_{k\max}$ and $x'_{k\min}$ are respectively the maximum and the minimum values among $x'_{1k}, x'_{2k}, \cdots, x'_{nk}$.

B. The establishment of similar matrix

Similarity coefficient is used to describe the similarity between the objectives. r_{ij} indicates the degree of similarity between vector x_i and vector x_j The bigger r_{ij} is, the more vectors similar are. Fuzzy similar matrix $R = (r_{ij})_{n\times n}$ can be got after computing all r_{ij}. According to actual situation, the method of calculating similar matrix can be chosen from a lot of formula. The algorithm in this paper uses the Gaussian function to determine the similarity matrix, and the definition of formula is as follow:

$$r_{ij} = \begin{cases} 1 & i = j \\ \exp(-\| x_i - x_j \|^2 / m\times\sigma^2) & i \neq j \end{cases} \tag{5}$$

The m is the dimension of the sample clustering, σ is the cluster radius.

C. Clustering

Fuzzy clustering analysis should be at a certain confidence, with intercepting fuzzy similar matrix through a threshold. The principle of interception is as follow:

$$r_{ij} = \begin{cases} 1 & r_{ij} > \partial \\ 0 & r_{ij} < \partial \end{cases} \quad \partial \in [0,1] \tag{6}$$

Because of the matrix symmetry, operate lower triangular part of the R_α .Check out the location where the value is 1 line by line in the principle, and then select all types.

In preference [6], it can be seen that the method of tracking can make it effectively transferring the matrix. The process of extracting pulse numbers is very simple, it is suitable for computer programming, especially for analysis of large-scale data, and is practically helpful for the promotion of application of the fuzzy cluster analysis.

III NEW SORTING METHOD BASED ON CLUSTERING

The signal DOA has two characteristics of Elevation angle and Azimuth angle. The number of Radar pulse samples is assumed n $X = (x_1, x_2, \cdots, x_n)$, each sample x_i has two characteristics. So x_i can be described as $x_i = (x_{i1}, x_{i2})$.

When the emitter angle is sorted by the Fuzzy Clustering, the characteristics of DOA parameters have the same dimension so the Data Standardization can be omitted. There will be only establishing fuzzy similar matrix, using tracking algorithm searching after calculate-eng ∂ -cut matrix. The mean of each category shall be the estimated location of each radar emitter. This algorithm has good performance and the estimated locations have small errors with the actual location of emitters. But it has to accumulate a certain amount of pulse, establish large fuzzy similar matrix, waste amount of calculation when extracting the same category pulse.

Dynamic clustering [6] selects the object of classification randomly, which represents initial mean of each cluster. For the remaining each object, it will be assigned to the nearest cluster according to the distance between the object and cluster center. The mean of cluster, adding a new object, needs to be calculated. This process is repeated till the criterion function converges.

A new algorithm based on above method can be proposed in order to overcome the limitation of establishing large similarity matrix. It uses fuzzy clustering estimating each cluster center, with dynamic clustering for reference and Gaussian distance Function instead of Euclidean distance. Define Gaussian distance Function d as follows:

$$d(x, y) = 1 - \exp(-\| x - y \|^2 / 2\sigma^2) \tag{7}$$

x is observation data, y is some emitter estimated center. σ represents the width of Gaussian function, which can be changed.

After fuzzy clustering, the number of category reaching a certain number can be selected. The cluster center can be estimated by these selected categories' mean. The pulses only need to compute the distance d between the DOA and each cluster center. If satisfying $d \le d_\lambda$ to some category, the pulse is divided into the category. If the distance between the DOA and each cluster center meet $d > d_\lambda$, the pulse is preserved in surplus section. When the number of emitters in different direction increases, the pulses of new emitters will be preserved in surplus section. The pulses in surplus section need to be sorted by fuzzy clustering reaching a certain number. After fuzzy clustering, the mean of category, whose pulse reach a certain number, can be added to the cluster center. Specific flow chart is as follows:

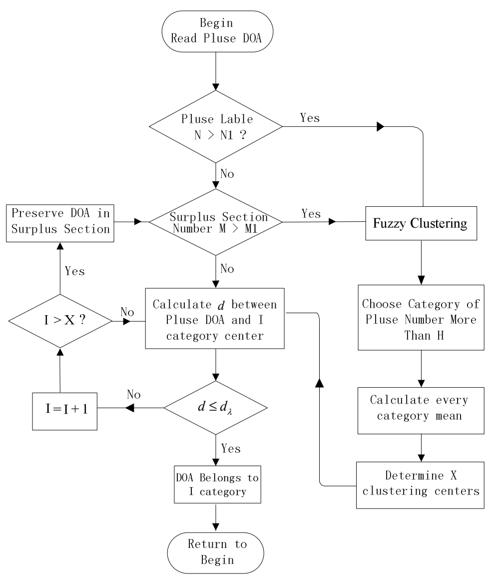

Figure 1.Program flow chart

Each emitter receiving pulses' intervals are marked. If the category doesn't receive signal pulse longer than the given time, the emitter can be viewed as stopping work. So this cluster center can be deleted the number of emitters decreasing can be tested.

IV. SIMULATED EXPERIMENT AND ANALYSIS

In order to verify the validity of the above methods, four radars signals in different direction are simulated on Matlab 2008, whose PRI is 1:1:1:1. The pulse steam is in accordance with 1:1:1:1 aliasing. Single-pulse angle measurement of dual-baseline is used. The Sinc function antenna is used as receiving antenna.

Simulation parameters are set as follows: Signal frequency 5GHz, Short baseline antenna spacing 0.049m, long baseline antenna spacing 0.179m, Antenna half-power beam width $(-37.5^\circ, 37.5^\circ)$, the Receiver SNR 10dB, each pulse sampled 15 points, pulse width 15us, pulse duration ratio 1:100.

The maximum unambiguous angle of short baseline angle is 37.75°, and the maximum unambiguous angle of long baseline angle is 9.65°. Angle measurement was carried out by long short baseline. The position of emitters and pulse number are set as Tab.1

TABEL I. The Position of Emitters and Number Setting

Radar Label	**Elevation Angle(°)**	**Azimuth Angle(°)**	**Pulse Number**
1	26	23	142
2	30	30	242
3	35	30	242
4	30	35	178

The pulse stream overlap is set as Tab.2. The pulse from 1 to 192 determine cluster center with Fuzzy Clustering. The pulses from 193 to 504 test the performance of algorithm when emitters increase. The pulses from 505 to 804 test the performance of algorithm when emitters decrease.

TABLE II. Pulse Stream Overlap Setting

Pulse Label	**Including Radar Label**	**PRI Proportion**
1 to 192	1，2，3	1:1:1
193 to 504	1，2，3，4	1:1:1:1
505 to 804	2，3，4	1:1:1

In the simulation, 20 T_i are used as the time measurement. When some category doesn't receive pulse more than 20 T_i, it will be deleted form cluster center.

The threshold ∂ of fuzzy clustering is set as 0.86 when determining cluster center. Particle Size d_λ is set as 0.96. The simulation result as table III. The simulation result of sorting the pulse stream from 1 to 804 is statistical as table IV.

According to Tab.3, it is known that fuzzy clustering algorithm can determine the location of emitters, which have a small error with the actual location of emitters. After determining the cluster centers, pulse DOA only compares with cluster centers' distance. This method does not require the establishment of a large similarity matrix and avoids bigger error of the Initial cluster centers with dynamic clustering. From figure 4 and figure 5, the number of emitters can be tested in time. From tab.4, the sorting accuracy is high.

TABLE III. Result Of Sorting Pulse Label From 1 To 504

Radar label	1	2	3
Total number	64	64	64
Right number	63	51	55
Elevation angle mean(°)	22.8	30.4	29.7
Azimuth angle mean(°)	26.1	30	35.1
Right rate	93.2%		

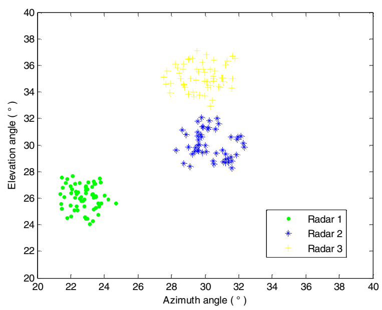

Figure 2 Pulse label from 1 to 192 clustering figure

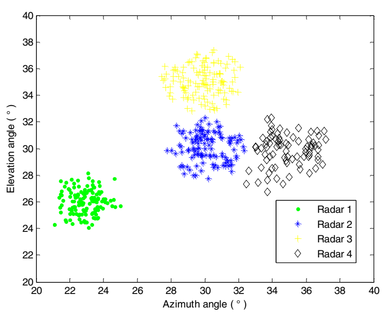

Figure 3 Pulse label from 1 to 504 clustering figure

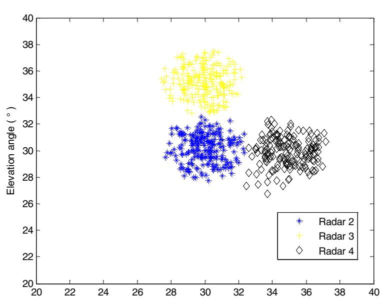

Figure 4 Pulse label from 1 to 804 clustering figure

TABLE IV. Result Of Sorting Pulse Label From 1 To 804

Radar label	1	2	3	4
Total number	142	242	242	178
Sorting number	138	224	219	168
Right number	138	222	216	156
Error number	0	2	3	12
Right rate	91.04%			

V. CONCLUTION

This paper introduces a novel method of signal sorting flowsheet based on Fuzzy Clustering with dynamic clustering for reference. This method does not require the establishment of a large similarity matrix, saves the sorting time, tests the number change of emitters. The sorting accuracy rate is high. The experiment results indicate that the algorithm is effective for radar signal sorting in time.

REFERENCE

[1] Wang Fan, Kohonen T. , The Self-organizing Map , IEEE ,1990 ,78 (9):pp.1464-1480.

[2] Wang Wei, Yang Jiong, Muntz R. STING: A Statistical Information Grid Approach to Spatial Data Mining , Proc. of the 23rd International Conference on very Large Data Bases, 1997.

[3] Agrawal R, Gehrke J ,Gunopulcs D. Automatic subspace clustering of high dimensional data for data mining application, Proc. of ACM SIGMOD Intconfon Management on Data , Seattle , WA ,1998:pp.94-205.

[4] Guha S, Rastogi R, Shim K. Cure: An Efficient Clustering Algorithm for Large Database, Proc. of ACM2SIGMOND Int . Conf, Management on Data, Seattle Washington ,1998: pp.73-84.

[5] Zhang Xing-hua, A New Algorithm for Fuzzy Clustering, Mathematics In Practice and Theory, 2005-03: pp.138~141.

[6] Wang Yong-gang, Radar Signal Sorting Method based on Fuzzy Clustering, Electronic Warfare,2007-02.