Proceeding of the IEEE
International Conference on Robotics and Biomimetics
Dali, China, December 2019

Clutter Filtering Algorithm in Dense Clutter Environment

Sheng Mou, Jianhui Guo*

School of Computer Science and Engineering
Nanjing University of Science and Technology
Nanjing, Jiangsu Province, China
{mousheng & guojianhui}@njust.edu.cn

***Abstract* - Multi-Target Tracking (MTT) in dense clutter environment has always been a research difficulty in the field of radar target tracking, the key is to effectively combine state filtering with data association. In the dense clutter environment, in addition to the echo of the target point, there are also a large number of clutter interference from unknown scatters, so it is difficult to process the data. In this paper, we propose a clutter filtering algorithm in dense clutter environment based on Track-Oriented Multiple Hypothesis Tracking (TOMHT) and Support Vector Machine (SVM), which is used to filter clutters, and to provide prior environmental information for subsequent target tracking. It reduces the density of clutter and improves the efficiency of data association under the premise of satisfying the tracking accuracy. The results show that the algorithm can effectively suppress clutter and improve tracking performance.**

Index Terms – Multi-Target Tracking, Data Association, Track-Oriented Multiple Hypothesis Tracking, Support Vector Machine.

I. INTRODUCTION

Radar target tracking is usually carried out in an environment where the detection source is unknown. In addition to receiving the echo of the target, there will inevitably be clutter interference from the scattering body, such as thermal noise of the target, cloud layer and terrain reflection [1] etc. Clutter seriously affects radar tracking results and performances. As the clutter density increases, the tracking filtering and data association performance of radar decreases rapidly. Researchers have proposed a variety of different methods to solve this problem. It is commonly used to take corresponding clutter suppression measures in the signal processing stage of radar, such as moving target detection, moving target indication, constant false alarm rate detection [2] etc. However, these methods still fail to filter out all clutters. Inevitably, there will be clutters left in the data processing stage. Therefore, research on radar data processing in dense clutter environment is of great significance.

There are two main methods of clutter suppression in radar data processing stage [3]. One is to use data association algorithm to directly identify and eliminate clutters [4]. However, due to the complex characteristics of clutters, it is difficult to apply this method in practice. Another method is to convert the clutter suppression problem into a data association problem, and improve the accuracy of the association by improving the data association algorithm. The most classic data association algorithm is the Nearest Neighbor (NN) method, which is computationally feasible and easy to implement, but this method is only suitable for the case where the clutter density is low and the target is sparse. The joint probabilistic data association (JPDA) algorithm proposed by Bar-Shalom [5] has been widely used to calculate the joint posterior association probability of multiple targets in clutter, but the shortcoming of the algorithm is the tendency to merge adjacent tracks. Researchers have been working to improve the algorithm. Fitzgerald [6] proposed a cheap joint probability data association (CJPDA) algorithm, also known as the nearest neighbour joint probabilistic data association (NNJPDA) algorithm, to prevent track merging by pruning. Roecker [7] proposed a fast suboptimal algorithm for JPDA to optimize it by introducing weights for possible hits. Reid [8] proposed the Multiple Hypothesis Tracking (MHT) algorithm in 1979, which is based on a hypothetical structure from scan to scan. The echo not only considers the possibility that it comes from the target and false alarm, but also considers the possibility that it comes from the new target. The newly received data is continuously expanded and pruned [9]. Cox [10] modified the algorithm, reducing the number of low probability assumptions. Streit [11] proposed a Bayesian data association and tracking algorithm in 1995, called probabilistic multi-hypothesis tracking (PMHT). Kurien's track-oriented multiple hypothesis tracking (TOMHT) [12] algorithm is the most popular algorithm for target tracking in high computing resources and clutter environments. Instead of using explicit assumptions like MHT, it uses a tree structure to represent hypotheses. Each node of the tree represents an observation, and the path between the root and the leaf represents a possible track [13-15]. TOMHT algorithm is considered to be the best solution to multi-target tracking problem because of its high compatibility and consistency.

In addition to the above traditional clutter suppression methods, we note that radar targets and clutters can be regarded as a binary classification problem, and SVM algorithm is a classical algorithm to deal with classification problems. This paper proposes a clutter filtering algorithm in dense clutter environment combining TOMHT and SVM. The track points and clutter points received by the radar are trained as positive and negative samples, and the resulting classification model can eliminate the received clutter and predict the target track. The experimental results show that the algorithm has a good performance in filtering clutter and improving data correlation. The rest of this paper is organized

This work was supported by a grant from the National Natural Science Foundation of China (No.61603190).

*Corresponding author: Jianhui Guo (E-mail: guojianhui@njust.edu.cn)

978-1-7281-6321-5/19//$31.00 ©2019 IEEE

as follows: Section II introduces the TOMHT algorithm and optimization, Section III introduces the SVM-based clutter classification method. Section IV introduces experiments based on real radar data. Section V gives a summary of the work of this paper.

II. TRACK-ORIENTED MHT AND ADAPTIVE OPTIMIZATION

The main idea of the MHT method is that in each scan cycle, the acquired target information has various possible hypotheses (these hypotheses correspond to possible tracks). The new measurement results of this scan are associated with the set of hypothesis scanned last time. The hypothesis that the probability are greater than the given threshold are preserved, impossible hypothesis are deleted, and the hypothesis of the same effect are combined at the same time [16][17].

The workflow for multiple hypothesis tracking is as follows:

1) Receiving echo: Receive the echo, associate the echo with the target.

2) Associate echo with track: Calculate the number and density of echoes in the track gate. If it is greater than the given threshold, the echo is weighted and averaged to obtain the equivalent echo, and the equivalent echo is used to generate the track hypothesis branch. Conversely, the original echo forms a branch.

3) Pruning: Calculate track scores, perform N-scan pruning, form new hypotheses and remove unlikely hypotheses.

4) Track update: Echoes that are not associated form an equivalent echo if their number and density are greater than the threshold. The echo is related to the unconfirmed track area to determine whether to start a new track. Update track information if new track starts.

Considering the complexity of MHT workflow, it is often accompanied by a large number of iterative calculations. In this paper, the Logarithmic Likelihood Ratio (LLR) of track is used as the assumption of track score to replace the matching probability used in traditional MHT.

The following iterative formula can be used to calculate the likelihood ratio (LR) of the track.

$$LR_k = \Delta LR_k \cdot LR_{k-1} \tag{1}$$

where LR_{k-1} and LR_k are the likelihood ratios of the track $k-1$, k respectively, and ΔLR_k is the increment term at time k.

$$\Delta LR_k = \begin{cases} \dfrac{P_D N\left(v_k;0,S_k\right)}{P_{FA} P_{FN}(z_k)}, & \text{if a track is associated with an echo} \\ \dfrac{1-P_D}{1-P_{FA}}, & \text{if no track is associated with an echo} \end{cases} \tag{2}$$

where P_D stands for detection probability and P_{FA} stands for false alarm probability, P_{FN} is the probability density function of clutter (generally assumed to be a uniform distribution or a Poisson distribution), z_k is the observation vector, v_k is the new interest, S_k is the new interest covariance and N is a Gaussian distribution.

When the track starts, the Likelihood Ratio can be initialized to

$$LR_0 = \lambda_{NT} / \lambda_{FA} \tag{3}$$

where $\lambda_{NT}, \lambda_{FA}$ indicates the average spatial density of new targets and clutter, respectively.

Then, the Logarithmic Likelihood Ratio (LLR) at time k can be obtained by the following recursive calculation

$$LLR_k = LLR_{k-1} + \ln\left(\Delta LR_k\right) \tag{4}$$

The above method simplifies the calculation process of track score. In order to limit the number of track hypotheses, a track branch with a very low track score may not be generated.

The MHT algorithm has a good data correlation effect in a complex echo environment. However, due to the use of delay decision to improve the correlation accuracy, the MHT algorithm will inevitably face the two problems of hypothetical branch number and computational complexity index increase when the number of tracking targets and the number of clutter increases [18].

In our algorithm, the filtered clutter map can provide a prior environmental information. Based on these prior information, we can perform adaptive processing as follows:

1) Track initiation limiting: When the clutter density exceeds a certain threshold, the beginning of the track is limited.

2) Update track score: Calculate the scores for each track hypothesis, perform N-scan pruning, retain the best track hypothesis branch, and update the information.

3) Report limiting: A region with an average clutter density exceeding an acceptable threshold is considered a clutter region, and reports in that region will not be processed.

4) Adaptive track hypothesis pruning: The hypothesis of low scores is directly removed by adjusting the deletion thresholds (e.g. those with higher clutter densities) during track pruning.

Through the above adaptive processing method, the real-time performance and robustness of the algorithm are improved without affecting the tracking performance.

III. CLASSIFICATION METHOD OF CLUTTER BASED ON SVM

A. Introduction to SVM Algorithm

The Support Vector Machine (SVM) theory proposed by Vapnik [19] has been successfully applied in many fields. Compared with the traditional classification method, SVM has better generalization ability and higher classification accuracy in practical application, especially for nonlinear data sets [20]. SVM adopts a method that the input vector is nonlinearly mapped to a high-dimensional feature space, and constructs a linear decision surface in the feature space by this method. The special features of the decision surface ensure that the SVM has a high generalization ability. The standard SVM learning algorithm problem can be attributed to solving a constrained

Quadratic Programming (QP) problem. By finding an optimal classification hyper plane that satisfies the classification requirements, the hyper plane can best achieve the "tolerance" of the local disturbance of the training samples while ensuring the classification accuracy. The classification result produced by this hyper plane is the most robust, and the generalization ability of unknown data is the strongest.

Target tracking of radar in dense clutter environment can be regarded as a binary classification problem, given the training sample set $(x_i, y_i),\ i=1,2,...,l$, $x \in R^n$,$y \in \{\pm 1\}$. x_i represents the echo received by the radar. If the echo comes from the target, the value of y_i is $+1$, and if it comes from clutter, y_i is -1. The division hyper plane is described by a linear equation $w^{\mathrm{T}} x + b = 0$. The task of the SVM is to classify the samples by finding the optimal hyper plane through normal vector w and displacement term b . In order for the hyper plane to correctly classify all samples and have a maximum margin, the following constraints need to be met.

$$y_i\left(w^T x_i + b\right) \geq 1, i = 1,2,\ldots,m \tag{5}$$

In order to maximize the interval, only need to maximize $\|w\|^{-1}$, which is equivalent to minimize $\|w\|^2$, so the problem of constructing the optimal hyper plane is transformed into the following formula under the condition that the constraint (5) is satisfied.

$$\begin{aligned} &\min_{w,b} \frac{1}{2}\| w\|^2 \\ &s.t.\ \ y_i(w^T x_i + b) \geq 1, i = 1,2,...,m \end{aligned} \tag{6}$$

To solve this constraint optimization problem, introduce the Lagrange function

$$L(w,b,\alpha) = \frac{1}{2}\| w\|^2 + \sum_{i=1}^{m} \alpha_i \left(1 - y_i\left(w^{\mathrm{T}} x_i + b\right)\right) \tag{7}$$

where $\alpha = \left(\alpha_1; \alpha_2; \ldots; \alpha_m\right)$, let the partial derivative of w and b in $L(w,b,\alpha)$ be zero then

$$w = \sum_{i=1}^{m} \alpha_i y_i x_i \tag{8}$$

$$0 = \sum_{i=1}^{m} \alpha_i y_i \tag{9}$$

By replacing formula (8) with (7), we can eliminate w and b in $L(w,b,\alpha)$, then by adding the constraints in (9) we can get the dual problem of equation (6) as follows

$$\begin{aligned} &\max_{\alpha} \sum_{i=1}^{m} \alpha_i - \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m} \alpha_i \alpha_j y_i y_j x_i^{\mathrm{T}} x_j \\ &s.t.\ \ \sum_{i=1}^{m} \alpha_i y_i = 0, \alpha_i \geq 0, i = 1,2,\ldots,m \end{aligned} \tag{10}$$

B. SVM Data Processing

When the SVM performs data processing, it needs to divide the data into a training set and a test set. Each instance in the training set contains a category label and several attributes. The goal of the SVM is to generate a model based on the training data that will predict the category labels of the test data that only give the attributes. We apply this idea to the clutter classification by marking the track points as positive and the clutter points as negative classes for training. The new data received by the radar continuously expands the training set and adjusts the parameters adaptively. The classification performance of the model for track points and clutter points is getting better and better.

Since our training samples are linearly inseparable, we project them from the original space to a higher dimensional feature space, making the samples linearly separable within this space. We use $\phi(x)$ to represent the eigenvectors after mapping x , so the model corresponding to the hyper plane in the feature space can be expressed as

$$f(x) = w^T \phi(x) + b \tag{11}$$

where w and b are model parameters, similar to equation (6).

$$\begin{aligned} &\min_{w,b} \frac{1}{2}\| w\|^2 \\ &\text{s.t. } y_i(w^T \phi(x_i) + b) \geq 1, i = 1,2,...,m \end{aligned} \tag{12}$$

Its dual problem is

$$\begin{aligned} &\max_{\alpha} \sum_{i=1}^{m} \alpha_i - \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m} \alpha_i \alpha_j y_i y_j \phi(x_i)^T \phi(x_j) \\ &s.t.\ \ \sum_{i=1}^{m} \alpha_i y_i = 0, \alpha_i \geq 0, i = 1,2,\ldots,m \end{aligned} \tag{13}$$

$\phi(x_i)^T \phi(x_j)$ in equation (13) is the inner product of samples x_i and x_j mapped to the feature space. It is difficult to calculate directly. To avoid this obstacle, we envision such a function

$$k(x_i, x_j) = \left\langle \phi(x_i), \phi(x_j) \right\rangle = \phi(x_i)^T \phi(x_j) \tag{14}$$

That is, the inner product of x_i and x_j in the feature space is equal to the result that they are calculated by the function $k(\cdot,\cdot)$ in the original sample space, where $k(\cdot,\cdot)$ is the kernel function. In this paper, we adopt the Radial Basis Function (RBF) kernel

$$k(x_i, x_j) = \exp(-\left\|x_i - x_j\right\|^2 \big/ 2\sigma^2) \quad \sigma > 0 \tag{15}$$

σ is the width of RBF kernel. Using kernel functions, we avoid calculating inner products in feature spaces of high or even infinite dimensions. The optimal solution of the model can be expressed as follows

$$\begin{aligned} f(x) &= w^T \phi(x) + b = \sum_{i=1}^{m} a_i y_i \phi(x_i)^T \phi(x) + b \\ &= \sum_{i=1}^{m} a_i y_i k(x, x_i) + b \end{aligned} \tag{16}$$

IV. EXPERIMENT

A. *Evaluating Metrics*

We evaluate our method by Accuracy, Precision, Recall, F1-score and the area under the ROC curve (AUC-ROC). We take the valid track point as positive sample and the clutter point as negative sample. For each point to be detected, we define TP and TN as the number of positive samples correctly classified and the number of negative samples correctly classified. Meanwhile, FP is the number of the negative samples which are recognized as the positive samples and FN is the number of the positive samples that are classified into negative samples.

Accuracy is the most common metric, which is the percentage of the sample size of the correct samples predicted by the classifier, which can be calculated by

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{17}$$

In most cases, Accuracy can be used as an indicator to evaluate the performance of the classifier. But for an unbalanced data set, it does not describe the performance of the classifier very well. For the dense clutter environment of this experiment, the samples in the data set are seriously unbalanced, and the number of clutter points is much larger than the number of track points. Therefore, Precision and Recall are proposed to evaluate the performance of the classifier

$$Precision = \frac{TP}{TP+FP} \tag{18}$$

$$Recall = \frac{TP}{TP+FN} \tag{19}$$

F1-score take into account both precision and recall to achieve a balance between them and maximize the two at the same time, as shown in the following formula

$$F1 = \frac{2*Precision*Recall}{Precision+Recall} \tag{20}$$

Before introducing the ROC curve, we first introduce the following two indicators, True Positive Rate (TPR) and False Positive Rate (FPR).They can be calculated by the following formula

$$TPR = \frac{TP}{TP+FN} \tag{21}$$

$$FPR = \frac{FP}{FP+TN} \tag{22}$$

TPR is actually the same expression as Recall, indicating the ratio that is correctly judged to be positive in all samples that are actually positive, and FPR is the ratio that is erroneously judged to be positive in all samples that are actually negative. Both of these metrics are for actual samples and can be used to avoid sample imbalances. Thus, by using TPR as the vertical axis and FPR as the horizontal axis, given a two-classification model and its threshold, a coordinate

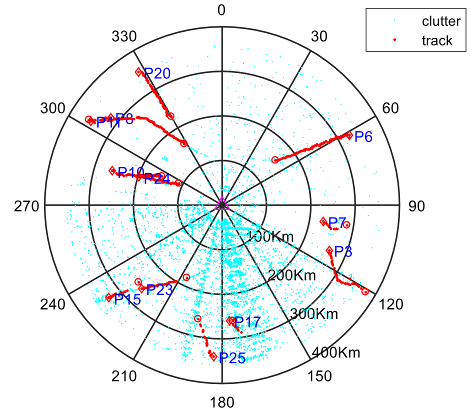

Fig. 1 Multi-target tracking results using TOMHT algorithm

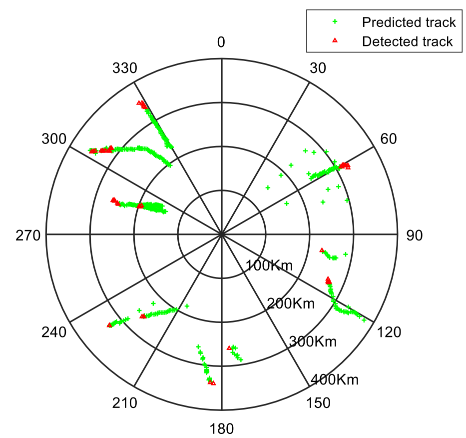

Fig. 2 Clutter filtering performance using our TOMHT-SVM algorithm

$(X = FPR, Y = TPR)$ can be calculated from all sample's real values and predicted values. The coordinates of each threshold of the model are drawn in the ROC space to form the ROC curve of the model, and the area under it can be used as a criterion for judging the quality of the classifier. The closer the value is to 1, the better the performance of the classifier is.

B. *Experiment Results*

The echo data were obtained from a down looking ball-borne radar, so the meteorological and ground clutter density is very high. There are 576 track points (positive samples) and 11980 clutter points (negative samples) in the echo. For each sample, we select three attributes, which are the azimuth, x and y coordinates of the echo in the radar Cartesian coordinate system.

Fig. 1 shows the results of radar multi-target tracking using the TOMHT algorithm. The red dots represent the track points and the bright blue dots represent the clutter points.

We filtered the clutter with the proposed TOMHT-SVM algorithm. After the track started, we marked the 10% of the data received by the radar as the training set. The track points were marked as positive samples and the clutter points were marked as negative samples. These training set generated our model. The remaining 90% of the radar data received later was used as a test set to predict the track through our model, as shown in Fig. 2.

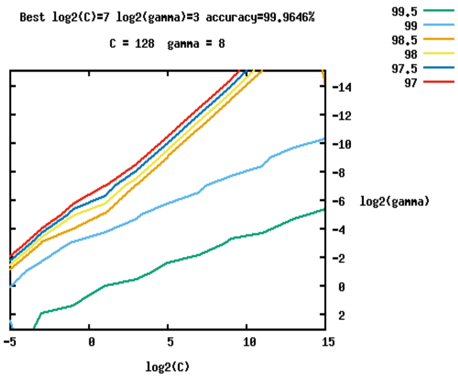

Fig. 3 Find the best parameters c, g by cross-validation

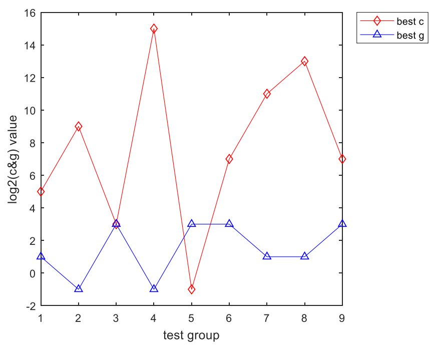

Fig. 4 Adaptive optimization of best c and best g

The red dot in Fig. 2 represents the starting track, which is the positive sample in our training set, and the green dot is the predicted track, which is the sample that we have classified as positive in our test set. Fig. 2 shows that our classification model already has a good classification performance at the beginning of the track, which can eliminate a lot of clutter and make a rough prediction of the subsequent track. However, due to the small number of training sets, the model misjudges a few clutter points as track points.

We conducted nine groups of comparative experiments. The training set data was increased by 10% each group compared with the previous group experiment, and the remaining echo data was used as a test set, which was reduced by 10% each group compared with the previous group. The experiment we designed above corresponds to the process of radar receiving data continuously.

Each group of experiments adaptively adjusts the optimal parameters to produce the best model. Fig. 3 shows that through 5-fold cross-validation using libsvm [21], we found the best penalty parameter c and the RBF kernel parameter g. Fig. 4 shows the best parameters for nine groups of experiments through adaptive adjustment.

TABLE I lists the experimental results of our TOMHT-SVM algorithm for different groups of radar operation.

As shown in Fig. 5, with the continuous increase of the training data received by the radar and the adaptive optimization of the model parameters, the model obtains better classification performance and the prediction of the track is more accurate.

TABLE I
PERFORMANCE AT DIFFERENT GROUPS

Group	Accuracy	Precision	Recall	F1-score	AUC
1	99.03%	96.79%	100%	98.37%	0.9875
2	98.50%	97.26%	100%	98.61%	0.9979
3	99.43%	98.62%	100%	99.31%	0.9591
4	98.41%	97.47%	100%	98.72%	0.9982
5	99.28%	98.38%	100%	99.18%	0.9647
6	99.38%	99.50%	100%	99.75%	0.9764
7	97.96%	96.15%	100%	98.04%	0.9980
8	99.16%	98.95%	100%	99.47%	0.9999
9	**99.36%**	**100%**	**100%**	**100%**	**0.9996**

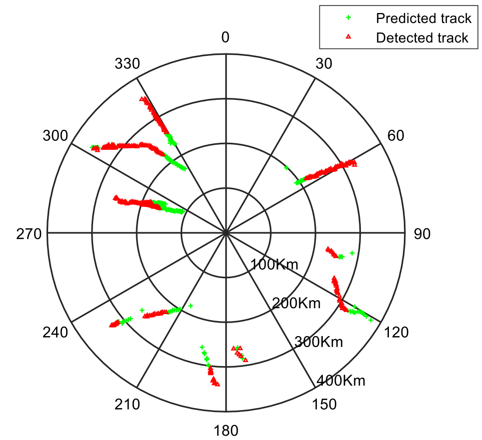

Fig. 5 Better performance after adaptive optimization

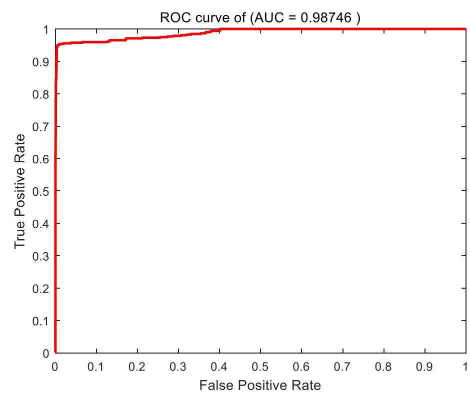

Fig. 6 ROC curve at the beginning of the track

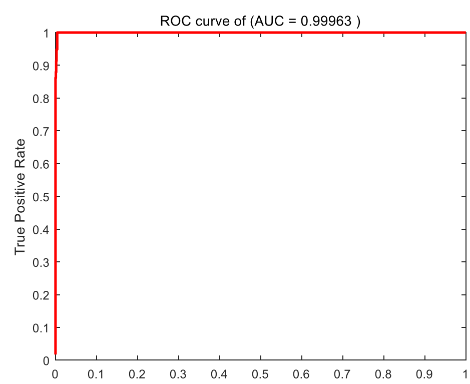

Fig. 7 ROC curve at the end of the track

Fig. 6 and Fig. 7 show the ROC curves at the beginning and end of the track. It can be seen that as time goes by, the model is continuously improved through adaptive training, and the classification performance is gradually perfect.

V. CONCLUSION

To further improve the tracking performance of radar in dense clutter environment, we propose a TOMHT-SVM based clutter filtering algorithm. The real radar application verifies that the algorithm can effectively remove clutter and accurately predict the target's track. It should be noted that our model is trained with labeled samples, which requires a certain amount of human and material costs to mark the training samples. Our future work will be modeled based on semi-supervised learning or unsupervised learning.

ACKNOWLEDGMENT

This work was supported by a grant from the National Natural Science Foundation of China (No.61603190).

REFERENCES

[1] Onunka, Chiemela, and G. Bright. "Modelling Radar Signal Error Performance under Atmospheric Refraction and Clutter Attenuation." *13th International Conference on Informatics in Control, Automation and Robotics* 2016.

[2] RADIOENGINEERING, VOL, and APRIL. "An Adaptive Clutter Suppression Technique for Moving Target Detector in Pulse Doppler Radar." *Radioengineering* 23(1):84-95, 2014.

[3] Xing-Wang, Luo, et al. "Researches on the method of clutter suppression in radar data processing." *Systems Engineering & Electronics* (2016).

[4] Z. H. Wu et al., "An Emulation Method of Sensor Clutter Plots Elimination Using Statistical Characteristics Analysis", Applied Mechanics and Materials, Vols. 170-173, pp. 3345-3348, 2012

[5] T. Fortmann, Y. Bar-Shalom and M. Scheffe, "Sonar tracking of multiple targets using joint probabilistic data association," in *IEEE Journal of Oceanic Engineering*, vol. 8, no. 3, pp. 173-184, July 1983.

[6] R. J. Fitzgerald, "Development of Practical PDA Logic for Multitarget Tracking by Microprocessor," *1986 American Control Conference*, Seattle, WA, USA, 1986, pp. 889-898.

[7] J. A. Roecker and G. L. Phillis, "Suboptimal joint probabilistic data association," in *IEEE Transactions on Aerospace and Electronic Systems*, vol. 29, no. 2, pp. 510-517, April 1993.

[8] Reid D.B., "An Algorithm for Tracking Multiple Targets," IEEE Trans. Autom. Control, Vol. AC-24, Dec. 1979, pp.843-854

[9] Kim, Chanho, et al. "Multiple Hypothesis Tracking Revisited." 2015 IEEE International Conference on Computer Vision (ICCV) IEEE, 2015.

[10] I. J. Cox and S. L. Hingorani, "An Efficient Implementation of Reid's Multiple Hypothesis Tracking Algorithm and It's Evaluation for the Purpose of Visual Tracking." IEEE Transaction on Pattern Analysis and Machine Intelligence, 18(2):138-150, 1996.

[11] Streit R.E, Lugjinbuhl T E. Probabilistic multi-hypothesis tracking[R]. NLI WC-NPT Technical Report, 1995.

[12] Kurien T, "Issues in the Design of Practical Multitarget Tracking Algorithm," Multitarget-Multisensor Tracking: Advanced Applications, Y.Bar-Shalom (Ed.), Norwood, MA: Artech House, 1990.

[13] Demos G.C., Ribas R.A., et al," Applications of MHT to Dim Moving Targets," Signal and Data Processing of Small Targets, Proc. SPIE ,Vol. 1305, Apr. 1990, pp:297-309

[14] Chan D.S.K., et al., "Performance Results of the Bilevel MHT Tracking Algorithm for Two Crossing Targets in a High Clutter Enviroment," Signal and Data Processing of Small Targets, Proc. SPIE ,Vol. 1954, 1994, pp:406-416

[15] A. Frank, P. Smyth and A. Ihler, "A graphical model representation of the track-oriented multiple hypothesis tracker," *2012 IEEE Statistical Signal Processing Workshop (SSP)*, Ann Arbor, MI, 2012, pp. 768-771.

[16] Guo, Jianhui, and R. Zhang. "Efficient radar data processing algorithm for dense cluttered environment." IEEE Cie International Conference on Radar IEEE, 2012.

[17] Moran, B.A., J. J. Leonard, and C. Chryssostomidis. "Curved shape reconstruction using multiple hypothesis tracking." *IEEE Journal of Oceanic Engineering* 22.4(1997):625-638.

[18] Bradley K Norman, Brian A Cronin, Samuel S Blackman and Robert J Dempster, " Adaptive processing to ensure practical application of a multiple hypothesis tracking system," Sensors, and Command, Control, Communications, and Intelligence (C3I) Technologies for Homeland Security and Homeland Defense V , edited by Edward M. Carapezza, Proc. of SPIE Vol. 6201, 62010N, 2006

[19] Cortes, C., & Vapnik, V. (1995). Support-vector networks. Machine learning: Vol. 20(pp. 273–297).

[20] Gecheng Chen, Zhiqiang Ge,SVM-tree and SVM-forest algorithms for imbalanced fault classification in industrial processes,IFAC Journal of Systems and Control,Volume 8,2019,100052.

[21] Chang, C.C. and Lin, C.J., 2011. LIBSVM: a library for support vector machines. ACM transactions on intelligent systems and technology (TIST), 2(3), p.27.