A Track Initiation Algorithm Using Residual Threshold for Shore-Based Radar in Heavy Clutter Environments

Yaru Zhang 1, Aichao Liu 2, Chao Liu 1, Bo Ai 1,3,* and Xiao Zhang 1

1 College of Geodesy and Geomatics, Shandong University of Science and Technology, Qingdao 266590, China; sugar_zyr@163.com (Y.Z.); liuchao951230@163.com (C.L.); zhangxiao9809@163.com (X.Z.)
2 North China Sea Marine Forecasting Center of State Oceanic Administration, Qingdao 266590, China; liuaichao@ncs.mnr.gov.cn
3 Key Laboratory of Urban Land Resources Monitoring and Simulation, MNR, Shenzhen 518060, China
* Correspondence: aibo@sdust.edu.cn; Tel.: +86-532-86057285

Received: 8 July 2020; Accepted: 14 August 2020; Published: 17 August 2020

Abstract: There is a large error in the actual radar trajectory tracking process. Track initiation is the primary problem in trajectory tracking and the first step in target tracking. The current track initiation algorithms are greatly affected by heavy clutter environments, so it is necessary to propose an algorithm to solve the problem of low track initiation efficiency. This paper presents a track initiation algorithm using a residual threshold in heavy clutter environments. The falling probability of measured value and decision threshold are used to determine the correlation window. The angle limiting condition is added to establish the track association, and the residual threshold is used to further eliminate the false tracks. The initial track experiment with the trajectory data in the sea near Rizhao Port shows that the algorithm is superior to the traditional logic method and Kalman filter method in track quality. The experiment uses the AIS buffer zone to calculate track initiation probability and uses the multi-region AIS trajectory data for verification. The experimental result shows that track initiation probability with the proposed algorithm in this paper can reach 92.31%.

Keywords: track initiation; logic algorithm; correlation window; residual threshold; Automatic Identification System (AIS)

1. Introduction

In recent years, a large amount of information contained in trajectories has attracted more and more attention with the development of trajectory information collection technology. By analyzing the ship’s historical trajectory, inferring the ship’s sailing time and fuel consumption, and determining the ship’s collision and distress, it will have a positive impact on ship operation and safety [1,2]. Track initiation is the premise and foundation of the follow-up processing of the radar tracks [3]. The so-called track initiation specifically refers to the process of track establishment before entering stable tracks (track keeping), which includes track head selection, the track initiation process, and track segment formation [4]. At the beginning of track initiation, the accuracy has always been a difficult problem because of the following problems: the long target distance, the low sensor detection resolution, the poor measurement accuracy, and no real statistical appearance of true and false targets. Especially, the above factors have a greater impact on the track initiation effect in a heavy clutter environment. The heavy clutter environment refers to the environment in which the radar is affected by ground clutter, sea clutter, and atmospheric factor. In addition, clutter in this article also refers to ship points that are not involved in track association, which will cause a great interference to the process of track association and affect the radar data processing and subsequent initial process. Therefore, it is
necessary to process the generated tracks to reduce the appearance of a large number of false tracks in a cluttered environment.

The existing track initiation algorithms can be divided into two categories, sequential processing technology (SPT) and batch processing technology (BPT). Generally, SPT is used to initialize the target trajectories in a relatively weak clutter background, while BPT has a good effect on initializing the target trajectories in a heavy clutter environment, but the use of BPT will increase computational burden [5]. The traditional initiation algorithms include an intuitive method, logic method, and Hough transform-based method. The intuitive method [6], also known as the heuristic rule method, is a simple initiation algorithm. The principle of the algorithm is that if there are M consecutive points of observations that met the initial conditions in N scans, the M consecutive points of observations can be considered as an initial track. The intuitive method is effective in dealing with the lack of prior target information. But it depends on the two measurement results in adjacent periods, only using speed and acceleration constraints so that the detection results are relatively rough with the simple determination rule. A logic method was proposed by F. R. Castella from the intuitive method [7]. Its principle is consistent with the intuitive method, but the difference between them is that the logic method predicts the possible trajectories using a correlation window and multiple hypotheses. The so-called correlation window and multiple hypotheses refer to setting the correlation window by using the predicted value of the target’s position at the next moment, and then judging whether the track meets the logic method. This method improves the situation where the intuitive algorithm is only suitable for ideal conditions. It is possible to find all correct tracks by traversing trajectories, but the disadvantage is that the ability to distinguish targets is poor, resulting in a high probability of false alarms. To improve the above situation, Sedehi et al. proposed an improved M/N logic method [8], which added a restriction to the measured value falling into the correlation window to eliminate the measured points that are V-shaped with the tracks in the state of track initiation. It optimizes the track initiation procedure and can be used to detect fast-moving observation targets, reducing the false alarm probability effectively. The Hough transform-based method was first used for image processing. It decides trajectory quality and starting time by selecting the appropriate transformation equation parameters [9]. This method is suitable for the targets in clutter environments, but it is difficult to initialize the tracks of the maneuvering targets due to the characteristics of the algorithm itself.

With the application of mathematical methods such as probability theory and stochastic processes, Kalman filter technology using linear system state equations began to appear, supporting a new research idea for target tracking. Welch proposed a classic Kalman filter algorithm using linear stochastic difference equations [10]. The algorithm used the state of the previous system to estimate the state of the current system. It combined these two values in a proportional combination to obtain a posterior state estimation and it iterated process to complete the optimal estimation of the true values finally. Sebesta and Boizo proposed a real-time adaptive high-gain EKF method [11], based on the minimum mean square error criterion, using statistical models of state equation and measurement equation, and using an innovative index to change high-gain parameters to achieve good filtering performance and sensitivity. Based on the above methods, Bahari et al. proposed a target tracking algorithm combining Kalman filter with fuzzy logic [12], which used a fuzzy system to transform a definite or fuzzy input into a fuzzy set output non-linearly. At the same time, it limited the output interval of the fuzzy controller to complete the mapping from input to output. This algorithm greatly improved the possibility of system uncertainty through the selection of estimation parameters and the division of fuzzy systems and achieved a more accurate target estimation and follow-up tracking.

In recent years, people have paid more and more attention to non-traditional tracking problems such as multi-target tracking and target tracking in dense clutter environments. Therefore, it is necessary to continuously improve traditional algorithms to further satisfy target tracking. Many innovative track initiation algorithms have been proposed as researchers continue to explore. Clark proposed a method for group target tracking using Probabilistic Hypothesis Density (PHD) filter and Gaussian mixture method [13]. The algorithm explicitly identified group targets and constituent members by creating
a graph of target state estimation, which allowed to limit the evolution of the group and used PHD filters to maintain the overall target tracking identity over time, greatly improving the target tracking effect of dense formation targets. Musicki et al. proposed an IPDA-MAP algorithm considering the uncertainty of the measured value and the existence of additional noise [14]. It proposed a recursive expression for the existence probability and data correlation. The approximate prior clutter density was obtained by the existence probability of the targets in each scan, which reflected and improved the target tracking quality. The algorithm was applied in a strong cluttered environment, which can better deal with the real track situation and identify false tracks. Yang et al. changed the selection of all possible measurement values in traditional track initiation methods and adopted a random sample consensus algorithm [15]. It used the spatial and temporal distribution of the measurement as a priori information for direct sampling and extracted the minimum measurement values randomly to establish a hypothesis, which can reduce the randomness in the sampling process and improve the track initiation efficiency. Compared with traditional algorithms, these algorithms are improved from the target tracking to ensure the quality and time of subsequent track initiation, and at the same time, they are suitable for multi-target tracking in heavy clutter environments.

In this paper, a track initiation algorithm using a residual threshold in heavy clutter environments is proposed with the integration of the existing track initiation algorithms. First, we determine the shape and size of the correlation window through the falling probability of the measured value and the decision threshold and predict the possible trajectories for the measured values falling into the correlation window in a multi-hypothetical logical manner. Second, an angle restriction condition is added to eliminate the V-shaped tracks. Finally, we further identify and remove the false tracks using the residual threshold. The innovation of the algorithm is that after the track association is established, the residual threshold is added to remove the false tracks that do not meet the conditions, ensuring the quality and accuracy of the track initiation.

2. Data Preparation

2.1. Research Area

The algorithm proposed in the paper needs to be verified by experiments. Experiments are divided into a comparison experiment and two auxiliary experiments. Three research areas are corresponding to the experiments, which are located in the sea near Rizhao Port, the sea near Jiaozhou Bay, and the sea near Weihai. The research area is shown in Figure 1.

Rizhao Port is located in the middle of China’s 18,000-km coastline, facing the Yellow Sea to the east, and in the recessed area where Subei Shoal and Shandong Peninsula meet; Jiaozhou Bay is a semi-enclosed bay with an area of nearly 500 square kilometers, located in the middle of China’s Yellow Sea, Qingdao City, Shandong Province. It is a large natural and fine harbor in China, so the throughput of ships is huge. Radar detection in this area can obtain a large number of ship data points, and the experimental tracking effect is good; the sea area near Weihai is located in the northern part of China’s Yellow Sea and the eastern part of Weihai City in Shandong Province. As the main channel for communication with the outside world, ships frequently come and go, and a large number of ship data points can also be obtained as experimental data.

2.2. Research Data

The data selected for the experiment are radar trajectory data and AIS trajectory data from three research areas, and the acquisition time is 22 August, 2019. Radar trajectory data comes from the image generated by shore-based radar scanning, signal-regulating, and data processing [16]. The data attribute information comes from all relevant parameters extracted after processing, and the data interval is set to be 5 min; AIS trajectory data comes from automatic information reports sent by the ship’s AIS equipment in the range of high-frequency channels [17], and the time interval is determined by the ship’s AIS equipment. The data are stored in the database in the form of a two-dimensional...
Radar trajectory data and AIS trajectory data from 8:00–11:00 on 22 August, 2019, are visually displayed on the map, and their distribution in the research areas is shown in Figure 2.

**Figure 1.** The research areas selected for the experiment, which are the sea area near Weihai, the sea area near Jiaozhou Bay and the sea area near Rizhao Port.

**Figure 2.** Data required for the experiment, including radar trajectory data and Automatic Identification System (AIS) trajectory data in three areas. Radar trajectory data is more than AIS trajectory data.
2.2.1. Radar Trajectory Data

Radar trajectory data comes from the scan results of shore-based radar. Shore-based radar is an all-weather shore-based surveillance radar, mainly used for monitoring small ships on the sea. The radar adopts a quasi-continuous wave system and a dual-beam, dual-channel design. The outstanding feature is the ability to detect targets under strong clutter background and high reliability. The technical indicators of shore-based radar are shown in Table 1.

<table>
<thead>
<tr>
<th>Name</th>
<th>Specific Indicator Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transmit frequency</td>
<td>C band</td>
</tr>
<tr>
<td>Antenna form</td>
<td>Dual-beam coverage</td>
</tr>
<tr>
<td>Range resolution</td>
<td>Better than 30 m</td>
</tr>
<tr>
<td>Azimuth resolution</td>
<td>Better than 1°</td>
</tr>
<tr>
<td>Positioning accuracy</td>
<td>Better than 10 m</td>
</tr>
<tr>
<td>Orientation accuracy</td>
<td>Better than 0.5°</td>
</tr>
<tr>
<td>Scan azimuth</td>
<td>0° ~ 360°</td>
</tr>
<tr>
<td>Target detection speed</td>
<td>V 108km/h (60 knots)</td>
</tr>
<tr>
<td>Operating temperature</td>
<td>−20 ~ +700 °C</td>
</tr>
</tbody>
</table>

Radar trajectory data used in this experiment in August 2019 comes from the North China Sea Data & Information Service of State Oceanic Administration (SOA), and the interval is 5 min. The total number of radar trajectory data in the three regions is 222,120, including 176,573 in the sea near Jiaozhou Bay, 44,648 in the sea near Weihai, and 899 in the sea near Rizhao Port. The attribute information of radar trajectory data includes radar type, target number, Maritime Mobile Service Identify (MMSI), longitude, latitude, target speed, azimuth, and scan time [18].

2.2.2. AIS Trajectory Data

AIS trajectory data used in this experiment in August 2019 comes from the North China Sea Data & Information Service of State Oceanic Administration (SOA), and the data interval depends on the information report returned by the ship’s AIS equipment. The total number of ship AIS data in the three regions is 63,655, including 52,405 in the sea near Jiaozhou Bay, 10,663 in the sea near Weihai, and 587 in the sea near Rizhao Port. The attribute information of AIS trajectory data includes ship number, ship length, longitude, latitude, speed, azimuth, and observation time [17].

2.3. Algorithm Basis

2.3.1. Radar Data Processing

Radar trajectory data required for experiments comes from the North China Sea Data & Information Service of State Oceanic Administration (SOA). SOA uses radar to receive the reflected original echo signal, and then performs signal-regulating [19,20] and data processing to enhance the original weak echo signal [21], through beam-forming [22], pulse compression, clutter filtering, and Doppler processing, etc. Finally, the target is imaged and the target attribute information is returned to facilitate the target tracking and target recognition process.

This paper uses the radar data products provided by SOA, including 222,120 radar trajectory data in three areas near Rizhao Port, the sea near Jiaozhou Bay and the sea near Weihai to verify the track initiation algorithm using a residual threshold. The experiment focuses on the distribution status and attribute information of the radar trajectory data, which starts with the radar trajectory data falling into the correlation window, infers the ship’s movement trend by the distribution status, and uses the algorithm to calculate the tracks with the longitude, latitude, target speed, target azimuth, and observation time in the attribute information. The focus of the experiment is to start and follow-up the radar trajectory data to realize the efficient track initiation in a heavy clutter environment.
2.3.2. Clutter Environment Description

Clutter refers to the radar scattered echo of other objects except for the interested target including noise and interference signals [23], which will affect the normal radar operation. Clutter is the inherent environment of radar signal detection and processing. It is one of the basic radar tasks to process signals in a cluttered environment. Usually, the clutter signal intensity far exceeds the target signal and clutter spectrum is often close to the target spectrum [23], which increases the difficulty of radar clutter processing.

“Clutter” in this article, on the one hand, refers to the fact that there are some stationary ships and immovable ships with damage. These ships cannot participate in the data association process to form tracks. These ships that produce interfering signals are called “clutter”, which will affect the ship data observation and the track formation. The track initiation algorithm will consume a lot of time if there are a large number of stationary ships; on the other hand, it refers to other noises in the acquisition process of radar signals, such as ground signals, ocean signals, and signals generated by atmospheric scattering, which have no relation with ship data. Therefore, these environmental noises are also called “clutter”. There are a lot of ship data in the selected areas of the experiment, so it is in a heavy clutter environment. The algorithm is proposed for the heavy clutter environment, and experiments are carried out to verify whether the algorithm can initialize the track efficiently and accurately in the heavy clutter environment and whether the performance is improved compared with the traditional algorithms.

3. Track Initiation Algorithm Using Residual Threshold

The conventional logic method, whose principle is to determine possible tracks through correlation window and position prediction. In this paper, the scanning volume falling into the correlation window is continuously improved to ensure that it can be associated with the tracks with the greatest probability. The improvement of the proposed algorithm in this paper is made by restricting the association rules and introducing the residual threshold. The idea of the algorithm is as follows, first, the appropriate correlation window is selected according to whether the measured value at the first moment is derived from the decision threshold of the target and the falling probability. Then the possible trajectories for the second time measurement set that falls within the correlation window are established, and each possible trajectory is extrapolated at the same time. The next step is to continue to establish the correlation window with the extrapolation point as the center, whose size is determined by the trajectory error covariance. An angle limit association rule is added, if the association rule is met, the track association will be given. Otherwise, the possible tracks will be canceled, or the correlation window will be expanded through the acceleration limit. Recursively, stable tracks are formed until the end of the fourth scan, and the free measurements (those that do not fall into the correlation window to participate in the track association) are used as new track heads, and the above steps are repeated. Finally, the proposed algorithm calculates the residual threshold and compares the residual threshold with the measured value residual. If the measured value residual is greater than the residual threshold, then all false measured values will be discarded to further suppress the possibility of false tracks.

A graphical description method is used to clearly describe the thought and process of the algorithm. The flow chart of the algorithm is shown in Figure 3.

3.1. Correlation Window Selection

Correlation window refers to an area that centered on the predicted position of the targets and it is used to determine the possible range of the observed values of the target. The echo that falls within the correlation window is called the candidate echo [24]. Common correlation windows include ring wave gates, elliptical (spherical) wave gates, rectangular wave gates, and fan-shaped wave gates in polar coordinates [25,26].
Figure 3. Flow chart of track initiation algorithm using the residual threshold. The data comes from the Service of State Oceanic Administration (SOA), which is processed in the process.

In practical applications, the selection of an appropriate correlation window is related to many factors. When selecting the correlation window for the trajectory data at the initial time, it is necessary to consider that the initial trajectory data belongs to the free points where the track has not been established for the first time. When starting the tracks, the target is generally far away, and the sensor has low detection resolution and poor measurement accuracy. The correlation window should generally establish a large wave gate for capturing the target and initializing the tracks, and it should be a circular non-directional wave gate [14,25,27].

The ring wave gate is generally used as the initial wave gate at the beginning of the tracks. It is a 360° ring wave gate with the track head as the center and it is determined by the maximum speed, minimum speed, and sampling interval of the detection target. The inner diameter and outer diameter of the ring wave gate satisfies:

\[ R_1 = V_{\text{min}} T \]  
\[ R_2 = V_{\text{max}} T \]  

Among the above formula, \( R_1 \) and \( R_2 \) are the inner diameter and outer diameter, respectively; \( V_{\text{min}} \) and \( V_{\text{max}} \) are the minimum speed and maximum speed of the detection target; \( T \) represents the sampling.
interval. In this algorithm, the speed judgment criterion is adopted, that is, the detection target speed should meet:

\[ v_{\text{min}} < v < v_{\text{max}} \]  

(3)

At this time, we make \( v_{\text{min}} = \frac{2}{3} \overline{v} \), \( v_{\text{max}} = \frac{3}{2} \overline{v} \). After taking the above parameters, the shape of the ring gate is shown in Figure 4.

![Figure 4. Ring correlation window.](image)

3.2. Track Association Establishment

The traditional logic method involves the processing of the received observations sequence during the continuous scan of the radar. The observations sequence represents the input of the time window containing N radar scans. When the number of the detection in the time window reaches the specified threshold, a successful trajectory is generated, otherwise, the time window is moved one scan time in the direction of increasing time, which is the principle of the typical sliding window method [7].

In the proposed algorithm of this paper, angle and acceleration limiting conditions are added to improve the track association based on the logic method. Although this algorithm is in the same order of magnitude as the traditional logic algorithm, it can effectively start the target tracks under the condition of high false alarm probability and it has great practical value. The search process of the algorithm is carried out as follows:

(1) Let the measurement set of ship points at the first moment be \( Z(1) = \{ Z_1(1), Z_2(1), \ldots, Z_{m1}(1) \} \), the ship points measurement set at the second moment is \( Z(2) = \{ Z_1(2), Z_2(2), \ldots, Z_{m2}(2) \} \). \( \forall Z_i(1) \in Z(1) \), \( i = 1, 2, \ldots, m_1 \), \( \forall Z_j(2) \in Z(2) \), \( j = 1, 2, \ldots, m_2 \), according to

\[
\begin{align*}
    d_{ij}^t &= \max \left[ 0, z_j^l(k+1) - z_i^l(k) - v_{\text{max}} t \right] + \\
    d_{ij}^b &= \max \left[ 0, z_j^l(k+1) - z_i^l(k) - v_{\text{max}} t \right] +
\end{align*}
\]

(4)

The distance vector between measurement points \( z_i^l(k) \) and \( z_j^l(k+1) \) is \( d_{ij}^t(t) \). In the formula, \( t \) is the time interval between two measurement points [28].

If the observation error is assumed to be independent, zero mean and Gaussian distributed, and the covariance is \( R_i(k) \), then the normalized distance squared is

\[
D_{ij}(k) \triangleq d_{ij}^t \left[ R_i(k) + R_j(k+1) \right]^{-1} d_{ij}
\]

(5)

The normalized distance squared between the measurement points \( z_i^l(k) \) and \( z_j^l(k+1) \) is found to be \( D_{ij}(k) \), where \( D_{ij}(k) \) is a random variable that follows the \( \chi^2 \) distribution with \( \rho \) degrees of freedom. A threshold probability is determined according to the actual situation, and then the \( \chi^2 \) distribution table with a degree of freedom of \( \rho \) is checked to get the threshold \( \gamma \). If \( D_{ij}(k) \leq \gamma \), you can determine \( z_i^l(k) \) and \( z_j^l(k+1) \) association, and then establish all possible tracks, which are \( O_{s1}, s1 = 1, \ldots, Q1 \) [29].
(2) For each possible track $O_{s1}$ extrapolation, you can continue to establish the correlation window $\Omega_i(2)$ using the extrapolation point as the center; the ship point falling into the measurement set should be $Z_j(3)$, while adding the angle limit criterion, that is: assuming that the line connecting $Z_j(3)$ and the second point of the track $O_{s1}$ and the angle between line and track is $\alpha$, another reference angle $\sigma$ is set. If $\alpha < \sigma$, then $Z_j(3)$ can be regarded as associated with this track. This reference angle is determined by the measurement error while adding the angle limit criterion. So, the measurement point traces that form a V-shape with the tracks can be removed to a certain extent [29,30].

(3) If there are no measurement sets for subsequent correlation windows, then the acceleration limitation method is used to expand the correlation windows, which is

$$|v_{i+1} - v_i| \leq a_{max}(t_{i+1} - t_i)$$

(6)

Continue to check whether the third scan falls into it, otherwise, this possible track will be canceled.

(4) It continues to extrapolate the track with the extrapolation point as the center and establishes the subsequent correlation window $\Omega_i(3)$, which is also determined by the covariance of the track extrapolation error. For the ship point measurement set $Z_j(4)$ falling at the fourth moment of the correlation windows, it continues to use the angle recursion judgment. If the angle $\beta$ of connection between $Z_j(4)$ and the first point of the track $O_{s1}$ is less than $\sigma$, then the measurement is considered to be associated with the track.

(5) In previous scans, the shipping point measurement set that did not fall into the correlation windows to participate in the data interconnection discrimination was used to start a new trajectory. The above steps were repeated until stable trajectories were formed. So far, the track initiation has finished, forming multiple ship trajectories.

3.3. False Tracks Elimination

The two basic requirements for track initiation are (1) the true target tracks are started as many as possible; (2) the false target tracks appear as little as possible. Therefore, the occurrence probability of false tracks greatly affects the pros and cons of a track initiation algorithm. A good track initiation algorithm needs to suppress the occurrence of false tracks as much as possible and reduce the false alarm probability of the tracks. False track suppression is a difficult problem in radar trajectory tracking. On the one hand, false tracks are inevitably formed for the unavoidable occurrence of false traces in radar scanning; on the other hand, it is necessary to spend a lot of time to confirm whether the trajectory is correct because false tracks will reduce the credibility of the trajectory data largely. There is inevitably a contradiction between the start time and the start accuracy rate [28,31]. In the research work of this paper, after completing the selection of the correlation window and starting the tracks, the probability of false alarm is further reduced to reduce the number of false tracks as much as possible [32].

Firstly, a false track evaluation index should be selected before carrying out the research work, namely the correct track rate [33]. The correct track rate refers to the ratio of the correct trajectories to the total trajectories, and it is an evaluation criterion that reflects the number of correct trajectories (or false trajectories). Suppose that at time $t$, in the track result graph formed by the radar scan data, the correct number of tracks is $Q_{dc}(t)$, the total number of tracks is $Q_d(t)$, and the correct track start rate $\rho_{dc}(t)$ for

$$\rho_{dc}(t) = \frac{Q_{dc}(t)}{Q_d(t)}$$

(7)

In the period from $t_a$ to $t_b$, the correct track rate $\rho$ is

$$\rho = \frac{1}{t_b - t_a} \int_{t_a}^{t_b} \frac{Q_{dc}(t)}{Q_d(t)} dt$$

(8)
3.3.1. Residual Threshold Calculation

In the above track search process, the concept of the threshold was introduced. The threshold is a term in physics. The original meaning is that when the input signal-to-noise ratio of the detector is reduced to a specific value, the output signal-to-noise ratio of the detector drops sharply. The input signal-to-noise ratio at this time is called the threshold. Applied to radar data processing, it can be understood as the minimum value of the number of spot detections in each period of radar scanning [34]. At the same time, there is a difference between the real value and the estimated value in each point sequence. This difference is called the residual [35].

Radar measurement values are obtained using polar coordinates. The distance and azimuth detection accuracy are set to \( \delta_r \) and \( \delta_\theta \). In the calculation, the radar distance and azimuth detection accuracy are converted into a rectangular coordinate system, and the calculation formula is as follows:

\[
\begin{align*}
\delta_x &= \sqrt{\delta_r^2 \cos^2(A) + \delta_\theta^2 R^2 \sin^2(A)} \\
\delta_y &= \sqrt{\delta_r^2 \sin^2(A) + \delta_\theta^2 R^2 \cos^2(A)}
\end{align*}
\]  

(9)

In the formula, \( \delta_x \) and \( \delta_y \) represent the conversion of radar distance and azimuth detection accuracy in a rectangular coordinate system; \( R \) is the distance measurement value; \( A \) is the azimuth measurement value.

In a rectangular coordinate system, the residual thresholds of each coordinate component are \( \kappa \sigma_x \) and \( \kappa \sigma_y \), respectively, \( \kappa \) is usually taken as 3 [33].

3.3.2. Residual Threshold Comparison

The initial correlation sequence is set to be \( \{Z_1, Z_2 \ldots \ldots, Z_m\} \), and \( \hat{Z}_i \) is used to be the fitted value of \( Z_i \). \( v_i = \hat{Z}_i - Z_i \) is the residual of the i-th point in the measured sequence, and the residual of the entire series of measurement values is represented by \( E \). The calculation formula is as follows:

\[
E = \sqrt{\frac{\sum v_i^2}{m-1}}
\]  

(10)

After the measurement sequence residual and threshold residual are calculated, the two can be compared, and the following two results may occur. If the measurement sequence residual is less than the threshold residual, the correct track will be started. If the measurement sequence residual is greater than the threshold residual, the set of measured value sequence is discarded to avoid the formation of a false track [33]. The specific performance is as follows:

- \( E_x \ll \kappa \sigma_x \) and \( E_y \ll \kappa \sigma_y \): start the correct track;
- \( E_x > \kappa \sigma_x \) or \( E_y > \kappa \sigma_y \): discard the measured value sequence to avoid the formation of false tracks;

By using the residual threshold comparison method, the measurement sequence that does not meet the requirements of the residual can be discarded at the beginning of the track to avoid the occurrence of false tracks with poor quality. Then the correct track initiation rate of the radar is calculated again to ensure that the initial accuracy rate has been accurately increased and false tracks have been effectively eliminated.

4. Experiment

This paper compares the traditional logic method and Kalman filter method with the track initiation algorithm using residual threshold through real data to show the improvement effect and track quality. The experiment uses radar trajectory data in the sea near Rizhao Port, including attributes such as ship number, position (latitude and longitude), speed, angle, and scanning time [36]. Time interval \( \Delta t \) is set to be 5 min. The distribution status of the radar trajectory data within the specified view field is displayed with the attribute information, which is shown in Figure 5. Radar trajectory
data are calculated by traditional logic, Kalman filter, and the track initiation algorithm using a residual threshold to obtain corresponding tracks, which are shown in Figure 6.

**Figure 5.** The distribution of the radar trajectory data is in the specified field of view.

**Figure 6.** Tracks calculated by three methods: (a) is the radar track by traditional logic; (b) is the radar track by traditional Kalman filter method; (c) is the radar track by track initiation algorithm using a residual threshold.
Figure 5 shows that there are many radar trajectory data within. Trajectory data distribution is scattered and irregular, so the echo signal of each ship will affect other ships, which results in a heavy clutter environment. By comparing the tracks obtained by three different methods in Figure 6, it can be found that the tracks are quite different even using the same radar data. The track quality depends on the relationship between points and tracks and whether there is a situation where multiple tracks intersect. It can be expressed in the reflected image whether the line of ships can be distinguished, whether the tracks have excessive crossover and fusion and whether one point is shared by multiple lines. In Figure 6a, the radar tracks based on traditional logic is the worst. The distribution of points and the tracks are chaotic and have not be paired correctly. In addition, one point is shared by multiple lines, and the tracks cannot even be initialized correctly; In Figure 6b, the quality of the radar tracks based on the traditional Kalman filter is slightly improved. The track quality is better at the area of the lower left and upper right, but the track quality is still not correct in the other areas. There are situations where the tracks intersect; In Figure 6c, the radar tracks by track initiation algorithm using residual threshold has the best quality. The track distribution is clear, and the correct pairing with corresponding ships is realized. There is no condition as cross-tracks and multi-line sharing. It can be seen that in a heavy clutter environment with a large number of ships interference and noise, the traditional logic method and Kalman filter can no longer initialize correct tracks, but the proposed algorithm in this paper can simplify track formation and have a better track quality.

The principle of traditional logic is to identify the possible trajectory through prediction and correlation window. If M points in the N observation sequence meet a provided angle criterion, it can be considered M points as a track, and the iterative process traverses all possible trajectories; the principle of traditional Kalman filter is to create a state vector by selecting ships at the first moment, and use the linear state equation to estimate ship state at next moment. It combines measured value and estimated value in proportion to obtain the posterior state estimation; the principle of the track initiation algorithm using residual threshold is to select the correlation window and provide angle-acceleration restriction rules to the measured value falling into it. Then it initializes the track that meets the restriction and compares with the residual threshold to further eliminate the false tracks.

Whether a track initiation algorithm has a good performance should be evaluated from three aspects, namely temporary track formation, trace association, and track false alarm probability [28]. In track formation, three methods can initialize all possible tracks comprehensively and form stable temporary tracks. In track association, the association effect of logic is poor because it connects all tracks without restriction, which results in computational burden. While Kalman filter and track initiation algorithm using residual threshold can set correct track association. In false alarm probability, traditional logic and Kalman filter are not suitable for heavy clutter environments because they haven’t considered false tracks possibility. The algorithm in this paper reduces false alarm probability by adding the residual threshold to discard the poor quality measurement sequence and eliminate false tracks. Therefore, the verification through experiments can show that the algorithm proposed in this paper has a good performance in initializing the tracks correctly and effectively.

5. Algorithm Accuracy Analysis

The following work is carried out for the accuracy analysis of the algorithm. It finds the AIS trajectory data that are consistent with the radar trajectory data used in the experiment and uses the AIS scattered data as attributes to obtain the tracks. Then a buffer zone is further established for the track formed by the AIS data, and the similarity of radar tracks and AIS tracks is acquired by calculating the radar tracks that fall in the buffer zone. If the overlapping probability of the radar tracks and the AIS tracks is high, it can indicate that the initial efficiency of the tracks is high, and it can further indicate that the algorithm has high accuracy.
5.1. AIS Tracks Formation

In the process of measuring the trajectory traces and predicting the tracks, radar scanning technology, and ship-borne AIS technology are the two main technical means for locating ship information [4,37]. Radar scanning technology can capture the real-time tracks of the ship, but it is susceptible to the surrounding air humidity, density, and actual installation conditions of the radar. Therefore, the azimuth accuracy, range accuracy, and target resolution of the radar are low compared with the ship-borne AIS equipment. The AIS equipment is an automatic communication device which is used between vessels and shore stations. It can automatically transmit and receive automatic identification information with competent authorities, shore stations, other ships and aircrafts, and there is no influence of occlusion factors. So, the positioning of ship-borne AIS is more accurate and it contains more complete data, but the real-time update is slower [38–40]. Both technologies can provide more accurate ship positioning information, which can complement and verify each other.

As the ship-borne AIS data is based on GPS positioning information, it is necessary to preprocess the AIS data first before using AIS data as an attribute data to verify algorithm [40]. The approach is to convert the latitude and longitude coordinates of the AIS data into plane rectangular coordinates through Mercator projection. The conversion formula for Mercator projection is

\[
\begin{align*}
    x &= r_0 \lambda \\
    y &= r_0 q
\end{align*}
\]

(11)

Among them, the parameters \( r_0 \) and \( q \) are expressed by the formula:

\[
    r_0 = N_0 \cos \varphi_0
\]

(12)

\[
    q = \ln \tan \left( \frac{\pi}{4} + \frac{\varphi}{2} \right) - \frac{e}{2} \ln \left( \frac{1 + e \sin \varphi}{1 - e \sin \varphi} \right)
\]

(13)

The parameter \( N_0 \) is expressed by the formula:

\[
    N_0 = \frac{a}{\sqrt{1 - e^2 \sin^2 \varphi_0}}
\]

(14)

In the formula, \( r_0 \) is the circle radius of the reference dimension, \( q \) is the equivalent dimension, \( N_0 \) is the ellipsoid curvature radius at the reference dimension, \( \varphi_0 \) is the reference dimension of the Mercator projection transformation, \( e \) is the first eccentricity of the ellipsoid, \( a \) is Earth’s long radius, \((x, y)\) is the Cartesian coordinate of the Cartesian plane, and \((\varphi, \lambda)\) is the latitude and longitude of the WGS-84 coordinate system.

After the preprocessing is completed, the AIS data is entered as attribute data. The time interval is set to be 300 s, which is consistent with the ship point data scanned by the radar. The search process refers to the above algorithm, and the navigation of the same batch of ship data points using the AIS data can be obtained. The distribution of the AIS data in the visual field is shown in Figure 7. The result of the ship’s trajectories obtained by the AIS data according to the proposed algorithm in this paper is shown in Figure 7, where the same color points represent the same number of ships. It is judged whether the same trajectory contains ships with different numbers through color, to judge whether the AIS trajectories are correct. However, the obtained trajectories are less than that of the radar scan due to the limited AIS data in the sea area selected during the tracking experiment.
It can be seen from the tracks in 0 that there are 16 tracks in total, 13 of which contain the ship data with the same color, indicating that the ship numbers in the tracks are the same and the initial tracks are correct. While there are another 3 tracks where there are differences in the color of ship data, which means that the ship numbers in the tracks are inconsistent and the formed tracks are incorrect. It can be calculated from the result that the accuracy rate of this algorithm when forming the AIS track is 81.25%. So it is regarded that a higher accuracy rate can be obtained. It can be concluded from this experiment that the proposed algorithm in this paper can correctly start the ship data from AIS, display the ship data trace distribution, and form good and stable tracks.

In theory, the radar measurement values and AIS measurement values based on the same batch of ship data should form approximately the same tracks [41]. To verify whether the two types of tracks are consistent, this paper establishes a buffer to quantitatively explain the overlapping probability of the two types of tracks and calculates the initial efficiency of the tracks to further analyze the accuracy of the algorithm.

5.2. Track Initial Efficiency Verification

The above inferences indicate that there is a certain overlap between the AIS tracks and the radar tracks. The higher the overlap probability of the two, the higher the efficiency of the track initiation, and the more accurate and effective the algorithm can be guaranteed. The AIS track and the radar track are derived to be distributed on the real map to quantitatively calculate the probability of overlapping tracks. The specific distribution is shown in Figure 8. A buffer zone with a buffer distance of 500 m is established for the AIS track. The radar tracks that completely fall within the buffer zone are regarded as completely overlapping with the AIS tracks. The AIS tracks where the buffer zone is located are regarded as overlapping tracks, according to the formula

\[ \rho = \frac{\gamma_0}{\gamma} \times 100\% \]
Calculate the overlapping probability of the track, where $\rho$ indicates the overlap probability of the track, $\gamma_0$ indicates the AIS track overlapped with the radar track, and $\gamma$ indicates all correct AIS tracks. After establishing the buffer, the buffer will be represented by classified colors according to whether the radar track falls into the buffer. The judgment result is shown in Figure 9, where red indicates that the radar track has not fallen into the buffer, and blue indicates the radar track has fallen into the buffer. The overlap probability is calculated to be $\rho = \frac{12}{13} \times 100\% \approx 92.31\%$.

Analysis of the above results shows that after the buffer zone is established for the correct 13 AIS tracks, the buffer zone formed by 12 tracks completely includes the radar tracks, so these 12 tracks can be regarded as overlapping tracks. The remaining one buffer zone formed by the track does not completely include the radar track, so this track does not consider the overlapping phenomenon. Therefore, the calculated track overlap probability is 92.31%, which can be considered as having a larger target track initiation probability. According to the research conclusion in the literature [42], it is described as follows: the track initiation probability is 89% by using traditional 3/4 logic algorithm,
which the number of clutter is set to be 300 per shot (the radar ship data used in this article is in a complex environment, there is a large probability of false alarm, and the number of clutter is far greater at 300), and the target detection probability is 0.85 that consistent with the experiment in this article. The specific parameter settings and algorithm results are shown in Table 2 below. Comparing the initial probability of the target track obtained in this paper, the initial probability is increased by 3.31%, indicating that the trajectories formed by AIS data and radar data are basically the same, and then it further quantitatively verifies that the algorithm of this paper has a higher accuracy at the beginning.

<table>
<thead>
<tr>
<th>Performance Index / Algorithm</th>
<th>Clutter Number (/Beat)</th>
<th>Target Detection Probability</th>
<th>Track Start Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Existing 3/4 logic algorithm</td>
<td>300</td>
<td>0.85</td>
<td>89%</td>
</tr>
<tr>
<td>Improved logic algorithm</td>
<td>Greater than 300</td>
<td>0.85</td>
<td>92.31%</td>
</tr>
</tbody>
</table>

5.3. Multi-Region Comparison Experiment Verification

Although the above experiment can quantitatively prove the accuracy of the algorithm in this paper, the amount of detected AIS data is insufficient for the limitation of the location of the nearby sea area and the slow real-time update of the ship’s AIS data within the current time. There may be a large accidental error in the training effect, and this error cannot be reduced in the subsequent series of results. This paper selects two additional areas with their radar scan data and AIS data to improve the problem of sparse data distribution, reduce the accidental error, and further verify the accuracy of the algorithm in this paper. The two types of data have relatively consistent properties and can be controlled, where the radar scan data and the AIS data belong to the same ship with the consistent acquisition time, scan interval, target detection probability, and the threshold probability. The two areas selected in this paper are the sea near Jiaozhou Bay and Weihai. The operation is as follows: First, the radar data and AIS data are stored in the database to read the corresponding latitude and longitude coordinates, speed, and azimuth. Next, the radar data and AIS data are displayed on the actual map using the latitude and longitude coordinates to form the distribution results of the radar and AIS points in the two areas. Then the trajectories for data points are started. There is the ship number in the AIS attribute, so you can connect the data points to form the AIS trajectories according to the same ship number. Substituting the radar scan data into the logic initiation algorithm based on the comparison of residual thresholds proposed in this paper, and you can use the attributes of latitude, longitude, angle, and azimuth to form the corresponding radar tracks after being processed by the algorithm. The results of AIS trajectories and radar trajectories near Jiaozhou Bay are shown in Figure 10, and the results of AIS trajectories and radar trajectories near Weihai are shown in Figure 11.

It can be seen from the Figure 10 that the number of ship points monitored by radar is much greater than that obtained by AIS in the sea area near Jiaozhou Bay. This is because AIS is ship-borne equipment. There are some situations that some ships do not have AIS system installed or AIS system is not enabled, which has resulted in far more radar tracks than AIS tracks in the area. However, because ships near the port require the AIS system to be turned on, there are corresponding AIS tracks and radar tracks near the port. By focusing on the analysis of the track distribution at the port, especially around the Huangdao district and Jimo district highlighted in the figure, it can be found that the radar tracks formed by the algorithm and the AIS tracks by the shipping number have a consistent trend and distribution (only the tracks shared by the two are compared). The distribution of the two is shown in Figure 10a,b where the AIS tracks and the radar tracks of Jiaozhou Bay are displayed in the same map level and the scale is set to 1:3 km. In the area near Weihai, the trajectories are in good condition, and the AIS trajectories and radar trajectories are clear and distinct. There are also cases where the radar trajectories are more than the AIS trajectories. It can be found that the distribution and location of the radar tracks and the AIS tracks are almost the same without considering the extra radar tracks by analyzing the northern area marked in the figure. The distribution of the two is shown in
Figure 11a,b where the AIS tracks and the radar tracks of Weihai are displayed in the same map level and the scale is set to 1:5 km.

Figure 10. AIS tracks and radar tracks near Jiaozhou Bay. (a) AIS tracks of Jiaozhou Bay, (b) Radar tracks of Jiaozhou Bay.

Figure 11. AIS tracks and radar tracks near Weihai. (a) AIS tracks of Weihai, (b) radar tracks of Weihai.
The comparison between the AIS tracks and the radar tracks in the above two areas can make up for the lack of data points in the buffer experiment, and further qualitatively shows that the algorithm of this paper has high accuracy at the beginning of the track. There is no effect of the accidental error so that the result is more accurate and reliable. It can be seen from the experiment that there is no AIS monitoring signal in some areas, which has a great impact on the monitoring activities of ships at sea. Therefore, this paper proposes a new inspiration for the dynamic monitoring of ships at sea. We can monitor the missing AIS data in real-time and determine the corresponding ship through data attributes according to radar tracks using the proposed algorithm in this paper, which can greatly avoid the occurrence of certain illegal situations.

6. Discussion

This paper proposed a track initiation algorithm using residual threshold and compared it with traditional logic and Kalman filter. The logic method increases false track possibility without track association restriction. Kalman filter uses previous state value to eliminate subsequent state, whose process will cause error accumulation, so it is also not suitable for environments with noise and interference. The proposed algorithm can identify possible tracks as comprehensively as possible to avoid the occurrence of missing tracks and eliminate false tracks through restriction, ensuring the track initiation accuracy and initializing tracks effectively in high false alarms. According to the comparison experiment, the proposed algorithm in this paper has improved track quality in a heavy clutter environment. At the same time, verification experiments with real AIS data are carried out to analyze track initiation accuracy and precision. The result shows that it has an initial probability of 92.31%, while the traditional logic method has an initial probability of 89%, which indicates the track initiation accuracy can be improved.

At the same time, the algorithm in this paper inevitably has shortcomings due to the limitations of data and equipment. The radar trajectory data used in the experiment all come from images generated by shore-based radar scan data, which originate from reflection echo signals of targets. The echo signal with a lot of noise and interference is affected by environmental factors, and it still has many errors even after signal processing. Meanwhile, how to select the number of samples is also a problem that needs to be considered. Few radar data with too many accidental errors will reduce the experimental accuracy and credibility, but too many samples will increase processing difficulty and computing time. So, it is necessary to select an appropriate sample number after comprehensive consideration. In addition, there are the following problems in algorithm processing.

1. There is no rule to follow when selecting the correlation window since the initial points belong to free points (that is, a free point that has not yet established a track). In the beginning, the detection resolution of the radar track is low, and the measurement accuracy is poor. Therefore, a relatively large correlation window is established. The significance of a large correlation window is to contain as many points as possible, but it also increases the burden of distinguishing effective tracks, making the probability of false alarms larger and the number of clutters increasing, which greatly affects the quality and time of track initiation.

2. In the study of track initiation performance, track reaction time is also an important factor in addition to the track quality. The track reaction time refers to the time from when the target enters the radar detection area to the establishment of tracks. The track reaction time should be reduced as much as possible and the tracks should be initialized efficiently and quickly based on ensuring the quality of the tracks. This paper focuses on the improvement of the track initial quality but spends less effort on the improvement of track reaction time.

The proposed algorithm in this paper still has a lot of room for improvement and development. For example, the shape of the correlation window can be customized, which can contain all the measurement points and traces, and try to avoid the occurrence of redundancy. Secondly, the sample of radar ship data should be expanded to ensure enough samples, but this also means that the processing
difficulty and the calculation time will increase. So the appropriate number of samples needs to be selected after comprehensive consideration. Finally, the track reaction time is also considered as a factor of performance. It is necessary to conduct experiments to verify the average number of scans when the track initiation and compare it with the research results in the existing logic algorithm to observe its improvement in track reaction time.

Through the analysis of the advantages and disadvantages of the algorithm, it can provide a reference for future research and improve the corresponding disadvantages, so that the algorithm can be further improved and developed.

7. Conclusions

In this paper, a track initiation algorithm using residual threshold is proposed to initialize the tracks quickly and efficiently in heavy clutter environments. The algorithm compares the residual thresholds from the logic method to eliminate tracks that do not meet the residual criterion on the premise of fully identifying all possible tracks. Experiments show that the algorithm in this paper has improved track quality compared with traditional logic and Kalman filter, and meanwhile, it has a great track initiation possibility of 92.31%, which much larger than the traditional method.

In summary, the track initiation algorithm using a residual threshold for shore-based radar can initialize the tracks effectively and quickly, meanwhile, it can improve the track quality and accuracy. All conclusions have been verified through experiments. Although the algorithm still has a lot of room for improvement, such as the determination of the shape and size for correlation window, the accuracy of the ship data and the consideration of the track response time, it can effectively reduce the false alarms and missed alarms probability to achieve an efficient and fast track initiation, which has a great practical value in subsequent radar trajectory tracking. With the in-depth exploration of the important information that implicit in ship trajectories in recent years, the ship trajectories have increasingly become an important means for people to monitor maritime traffic, analyze historical trajectories, predict the sailing time and solve distress problems. Through the analysis and verification of the proposed algorithm in this paper, it can help initialize tracks quickly and accurately and lay a solid foundation for the entire trajectory tracking process, which is more conducive to the analysis and research of ship trajectories.

Author Contributions: Conceptualization, B.A.; data curation, A.L.; methodology, Y.Z.; software, C.L.; validation, C.L. and X.Z.; writing—original draft, Y.Z.; writing—review and editing, B.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Key Research and Development Program of China, grant number 2017YFC1405004, the Major Science and Technology Innovation Projects of Shandong Province, grant number 2019JZZY020103, the Open Fund of Key Laboratory of Urban Natural Resources Monitoring and Simulation, Ministry of Natural Resources, grant number KF-2019-04-070, and the SDUST Research Fund, grant number 2019TDJH103.

Conflicts of Interest: We declare that we have no financial and personal relationships with other people or organizations that can inappropriately influence our work, there is no professional or other personal interest of any nature or kind in any product, service and/or company that could be construed as influencing the position presented in, or the review of, the manuscript entitled.

References

2. Szlapczynski, R. Evolutionary planning of safe ship tracks in restricted visibility. *J. Navig.* 2015, 68, 39–51. [CrossRef]


38. Zhang, L.; Meng, Q. Probabilistic ship domain with applications to ship collision risk assessment. *Ocean Eng.* 2019, 186, 106130. [CrossRef]


