Received February 25, 2018, accepted April 4, 2018, date of publication April 11, 2018, date of current version May 16, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2825887

Clustering Statistic Hough Transform Based Estimation Method for Motion Elements of Multiple Underwater Targets

ZHEPING YAN, GENGSHI ZHANG, JIAN XU, TAO CHEN, XUE DU, AND JUAN LI

College of Automation, Harbin Engineering University, Harbin 150001, China

Corresponding author: Xue Du (duxue2012cc@163.com)

This work was supported in part by the National Nature Science Foundation of China under Grant 51679057, Grant 51709062, Grant 51409055, and Grant 51609046, in part by the Distinguished Youth Scholars Foundation of Heilongjiang Province under Grant J2016JQ0052, and in part by the National Nature Science Foundation of Heilongjiang province under Grant E2015050.

ABSTRACT Estimation for motion elements is one of the core functional components of the multiple targets tracking system. Aiming at estimation for motion elements of multiple underwater targets, a Clustering Statistic Hough Transform (CSHT) method is proposed in order to overcome the false alarm and missing detection effects as well as positioning errors of the sonar data and improve the accuracy and reliability of feature extraction. First, the distance-direction data from the multi-beam forward-looking sonar mounted on the unmanned underwater vehicle are transformed to position curves of multiple targets in the earth-fixed frame, and the position curves appear to be sampling points that form the data space. Second, parameter space is constructed by applying Hough transform to the sampling points in the data space, and then the votes of each cell in the rasterized parameter space are accumulated. Finally, fuzzy iterative self organizing data analysis techniques algorithm clustering method is exploited for extraction of multiple peaks in the parameter space to realize estimation for motion elements. The application of CSHT method in the underwater multiple targets tracking system is further explained in this paper. Simulation results demonstrate that CSHT method is insensitive to environment noise, false alarm and missing detection effects of the sonar and offers favorable estimation accuracy and tracking performance, indicating engineering reliability.

INDEX TERMS Clustering statistic, estimation for motion elements, Hough transform, multiple targets tracking, multiple underwater targets, unmanned underwater vehicle.

I. INTRODUCTION

Unmanned underwater vehicle (UUV) is an important strategic equipment to explore and develop the ocean, and it increasingly arouses research interests among marine powers. As UUV steps towards deeper and farther in the ocean, its ability of safe autonomous navigation faces higher demands. Besides comprehension of the nearby static environment, correct perception of the surrounding maneuvering targets is an essential factor that provides effective guidance for movements in view of safety. Consequently, it is of practical engineering significance to study the estimation method for motion elements of multiple underwater targets [1]–[3].

The sonar is an essential instrument for UUV to obtain information about underwater environment [4], [5]. Estimation for motion elements of multiple underwater targets means to analyze and process the distance-direction data from the forward-looking sonar for resolving the states of the maneuvering targets, including estimation of the target number as well as the position and velocity of each target. Estimation for motion elements would be further used for multiple targets tracking jointly with data association technology in order to ensure grasp of motion states of the surrounding targets for UUV throughout the voyage.

Various filtering methods have been applied to multiple targets tracking problem for improvement of accuracy and reliability, including Kalman filter [6], [7], particle filter [8], [9] and probability hypothesis density (PHD) filter [10], [11]. In the marine environment, the continuity and reliability of sonar signals are greatly reduced resulting from reverberation, internal wave and multi-path effect [12]. Consequently, underwater multiple targets tracking is more difficult in comparison with aerial and terrestrial environments. Although the

above filtering methods have been intensively investigated and some research results have been applied to engineering practice, the application in the complex underwater environment still faces various limitations.

The detected data accumulate with time and form an image, so some researchers apply image processing methods to multiple targets tracking. Hough transform is a typical image processing method proposed by Hough [13] in 1962. This method exploits the mapping between curves in the image space and points in the parameter space, recognizing basic shapes by extracting the peaks in the parameter space. Line detection is most widely applied [14].

Hough transform has aroused widespread concern among the scholars since it was proposed, and several variants have been put forward against the problems of high computational cost and low accuracy in cluttered condition. Typical variants are Hierarchical Hough Transform (HHT) [15], Fast Hough Transform (FHT) [16], Random Hough Transform (RHT) [17], Probabilistic Hough transform (PHT) [18], Dynamic Combinatorial Hough Transform (DCHT) [19], and Connective Hough Transform (CHT) [20].

Recently the standard Hough transform(SHT) and its variants have been combined with other methods for improvement in computing efficiency and accuracy, and part of the research results have been applied to multiple targets tracking.

Some researchers firstly use other time-domain or frequency-domain methods to pre-process the data for denoising, and then exploit Hough transform for estimation. Aiming at detecting moving targets from Infrared (IR) imagery sequences, Huber-Shalem *et al.* [21] introduces a parametric temporal compression incorporating Gaussian fit and polynomial fit with possibly simplest representation for the original sample data. The compression method is evaluated by using the variance estimation ratio score (VERS), which is a signal-to-noise ratio (SNR) based measure for point target detection that scores each pixel and yields an SNR scores image. The target location is extracted from the SNR scores image by Hough transform with high detection probability and low false alarm probability. Islam and Chong [22] applies Hough transform to radar signals that has been processed by an improved wavelet threshold function for denoising, which effectively detects moving targets in strong noise condition. Lei and Huang [23] applies the entropy weighted coherent integration (EWCI) algorithm to the data collected from a radar system, which suppresses the time-varying clutter due to multi-path effects in the foliage-penetration environment. Based on the obtained image with high visual quality, the radial velocity of a moving human target is accurately estimated by Hough transform, and then the target trajectory is detected.

Another group of researchers firstly use Hough transform to pre-process data or estimate some parameters, and then exploit other methods for state estimation for targets. In order to track multiple moving targets by dual-frequency continuous-wave through-wall radar (DF-CW TWR), Ding *et al.* [24] uses Hough transform to decompose the echo with a progressive geometrical model and applies a modified high-order ambiguous function (MHAF) method to the estimation of target instantaneous frequency. The target trajectories are synthesized based on the estimation results. In order to focus moving targets by single-antenna synthetic aperture radar, Yang *et al.* [25] uses Hough transform to estimate the slope of the range walk trajectory, and then the cross-track velocities are obtained and the Doppler ambiguity problem is solved. Polynomial Fourier transform is further used for estimation of motion parameters of the moving targets.

Track initialization is crucial to multiple targets tracking problem, and Hough transform is selected as a satisfying method to accomplish this task for its insensitivity to noise. Xu *et al.* [26] proposes an ant colony optimization (ACO) based algorithm for the initiation of bearing only multiple targets tracking process, where the cost function is derived from the thought of Hough transform. A satisfying performance of track initiation is achieved against a heavy clutter environment with computing time effectively reduced. Hadjira *et al.* [27] proposes a Real Time Hough Transform (RTHT) towards real time automatic initiation of tracks in clutter by radar signals, where only the measurements that satisfy the velocity and acceleration tests are transformed. As a result, the processing time is significantly reduced.

Extension of Hough transform into multiple dimensions proves efficient in the state estimation process. Moyer *et al.* [28] proposes a multi-dimensional Hough transform (MHT) technique for the track-before-detect (TBD) processing, where the data dimensions can be the target position, its range, range rate and the first-threshold crossing times. The detection of the moving targets are efficiently enhanced by combining multiple first-threshold crossings. Moqiseh and Nayebi [29] extends the SHT to a 3-D data space, where information of planar positions and time is involved. The data space is constructed from surveillance radar signal using the distance and direction information of several successive scans. The proposed 3-D Hough detector is then used to detect the existent targets in the 3-D data space, which effectively improves the detection performance.

The application of Hough transform theory in the multiple targets tracking system has been improving, whereas the majority of the methods are radar signal oriented, infrared signal oriented, or video image oriented. Hough transform usually only works in the track initiation stage. Few of the Hough transform based multiple targets tracking solutions process sonar signal that severely suffers from ambient noise.

When Hough transform is used in estimation for motion elements of multiple underwater targets, the target number is unknown, and the sonar data are affected by measurement noise as well as false alarm and missing detection problems, so the votes in the parameter space do not appear as isolated peaks. A cluster of peaks are distributed near the cell corresponding to a certain pair of parameters. It is difficult to accomplish parameter extraction by simply setting

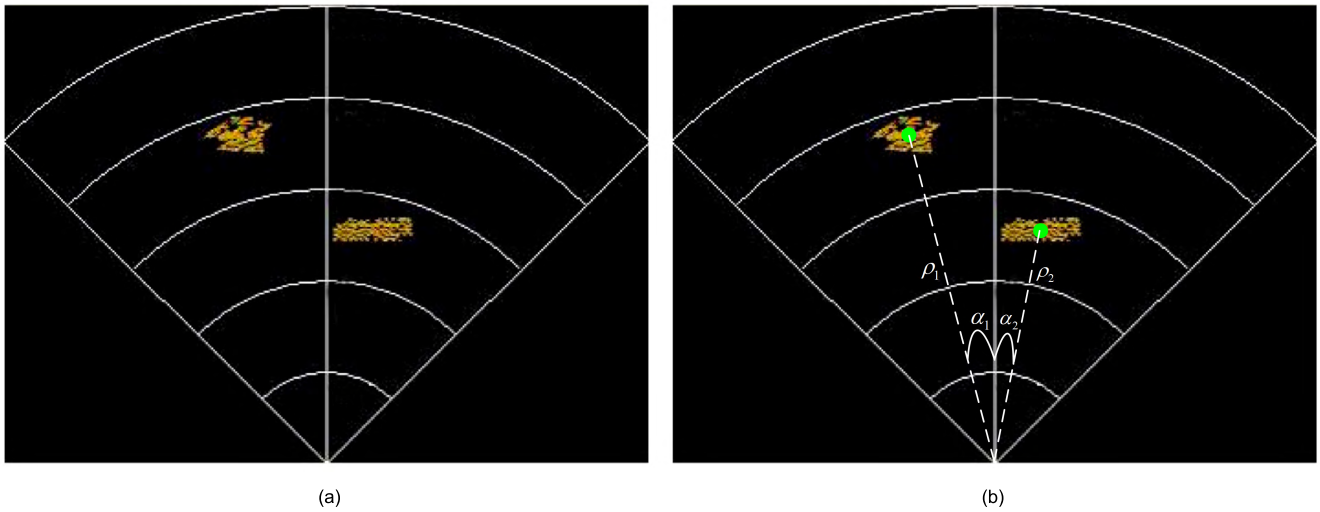


FIGURE 1. Diagrammatic sketch of sonar detection. (a) Original sonar image. (b) Detected targets by sonar.

a threshold, because an improper threshold may lead to several false targets or missing targets.

Clustering analysis is an unsupervised learning process that divides a set of abstract objects into classes comprise of similar objects. The purpose of clustering is to emphasize the similarity within the same class as well as the diversity between different classes. Without training data in the implementation procedure, classification is realized by exploiting the inherent characteristics of the data and the similarity relation. In explorative data analysis and data mining, the clustering is often exploited to discover part or all of the patterns hidden in the data [30]. The clustering algorithm applies to recognition of peaks in the parameter space for Hough transform. Clustering methods have been widely applied in the engineering field with the rapid development of machine learning technology. Fuzzy ISODATA is a typical and effective clustering method, which starts from an initial classification and updates the fuzzy membership matrix by iterative computation until the terminating condition is satisfied [31].

Aiming at underwater target detection by multi-beam forward looking sonar, a Clustering Statistic Hough transform (CSHT) method is proposed in order to improve the accuracy and reliability of extraction for multiple peaks in the parameter space. The distance-direction data from the forward-looking sonar are mapped to sampling points in the earth-fixed frame, which are further mapped to parameter space by Hough transform. The votes of each cell in the parameter space are accumulated, and the clustering sample data are obtained according to the vote matrix. Fuzzy ISODATA algorithm is used for clustering of parameters to estimate motion elements because of its class adjustment function, which makes it appropriate for the condition of unknown target number. CSHT method is further exploited in the underwater multiple targets tracking system, and it continuously works for restraint of false alarm and missing detection effects in the whole tracking process rather than merely in the track initiation stage.

The remainder of this paper is organized as follows. In Section 2, The mathematical model of sonar vision field is established, and the coordinate conversion mechanism of sonar data is introduced; In Section 3, the principle of the CSHT method is explained, including Hough transform, clustering statistic and spatial correlation; In Section 4, the application of CSHT method in the underwater multiple targets tracking system is explained; Simulation results on different methods are compared and discussed in Section 5; Finally, conclusions are drawn in Section 6.

II. MATHEMATICAL MODEL OF SONAR VISION FIELD

So far in underwater environment, sound wave is the only carrier that can remotely transmit information, and thus sonar is an ideal instrument for underwater detection and perception. The studied UUV in this paper uses an active multi-beam forward-looking sonar for target detection.

Sound beams are sent and received by phased array and the echo intensity signals form a sonar image. The function of image processing is integrated in the studied sonar. The distance and direction information of the targets can be obtained by applying a series of image processing techniques like segmentation, interpolation and enhancement, as is shown in Fig. 1. The distance and direction information are used to calculate northern and eastern coordinates of the targets that act as the input data for motion elements estimation.

The key parameters of the forward-looking sonar are listed in Table 1.

The forward looking sonar is horizontally mounted in the front of the UUV. The vertical beam width of the sonar is only 6°, and it is reasonable to assume that the detected objects exist in the same plane with the UUV.

TABLE 1. Parameters of the multi-beam forward-looking sonar.

Parameters	Values
Number of beams M	90
Horizontal beam width α_{H}	90°
Vertical beam width α_{V}	6°
Direction resolution α_{R}	1°
Distance detection scope ρ_{D}	150 m

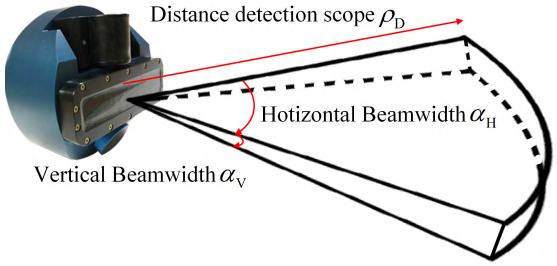


FIGURE 2. The multi-beam forward looking sonar.

The sonar comprises of 90 ceramic receivers that form 90 beams corresponding to 90 directions. If a target is detected at one beam, the sonar provides the distance of the target, which would be combined with the corresponding direction in order to locate the target. The sonar sends the detection information to the UUV at 1 Hz.

The detection space of the multi-beam forward-looking sonar is a polar coordinate system that provides distance and direction information of multiple targets. The sonar is mounted in the front of the UUV. A polar frame is established to describe the sonar detection information, where the origin is the mount point of the sonar, and the polar axis direction accords with the longitudinal direction of the body fixed frame of UUV. The sonar detection information is expressed as a distance-direction mode, namely $[\rho, \alpha]^{\mathrm{T}}$, where ρ and α respectively represent the distance and direction from the sonar to the target in the sonar frame. The direction pointing to the starboard of the UUV is positive.

The coordinate systems describing the movement of the targets are shown in Fig. 3. The axes of the earth-fixed frame represent north and east directions.

The target coordinates in the body-fixed frame are:

$$\begin{bmatrix} x \\ y \end{bmatrix} = \rho \begin{bmatrix} \cos\alpha \\ \sin\alpha \end{bmatrix} + \begin{bmatrix} x_S \\ y_S \end{bmatrix} \tag{1}$$

where $[x_{\mathrm{s}}, y_{\mathrm{s}}]^{\mathrm{T}}$ are the coordinates of the sonar mount point P_{s} in the body-fixed frame. The target coordinates in the earth-fixed frame are:

$$\boldsymbol{\eta} = \begin{bmatrix} n \\ e \end{bmatrix} = \boldsymbol{S} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} n_{\mathrm{V}} \\ e_{\mathrm{V}} \end{bmatrix} \tag{2}$$

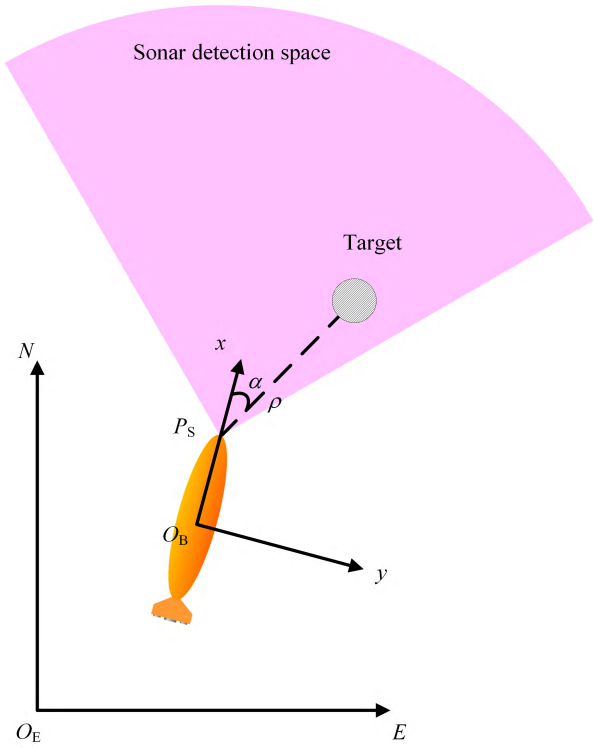


FIGURE 3. Coordinate systems and sonar detection space.

where $[n_{\mathrm{v}}, e_{\mathrm{v}}]^{\mathrm{T}}$ are the coordinates of the UUV in the earth-fixed frame, and $S \in R^{2\times 2}$ is the transformation matrix:

$$\boldsymbol{S} = \begin{bmatrix} \cos\psi & -\sin\psi \\ \sin\psi & \cos\psi \end{bmatrix} \tag{3}$$

where Ψ is the heading of the UUV. A set of target position data with time-stamp are obtained over time:

$$\left(t_i, \{\boldsymbol{\eta}_i\}\right), \quad i \in \mathbf{N} \tag{4}$$

where $\{\boldsymbol{\eta}_i\}$ is the set of coordinates in the earth-fixed frame of the targets that are calculated from the sonar data at the moment of t_i.

The state space model of a single target is:

$$\begin{aligned} \boldsymbol{X}(i) &= \boldsymbol{\phi}\boldsymbol{X}(i-1) + \boldsymbol{\Gamma}\boldsymbol{W}(i) \\ \boldsymbol{Y}(i) &= \boldsymbol{\Psi}\boldsymbol{X}(i) + \boldsymbol{V}(i) \end{aligned} \tag{5}$$

where $\boldsymbol{X}(i) = \left[\boldsymbol{\eta}_i^{\mathrm{T}} \ \dot{\boldsymbol{\eta}}_i^{\mathrm{T}}\right]^{\mathrm{T}}$, $\boldsymbol{Y}(i)$ is the observation value of $\boldsymbol{X}(i)$.

$$\boldsymbol{\phi} = \begin{bmatrix} 1 & 0 & T_s & 0 \\ 0 & 1 & 0 & T_s \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad \boldsymbol{\Gamma} = \begin{bmatrix} \frac{1}{2}T_s^2 & 0 \\ 0 & \frac{1}{2}T_s^2 \\ T_s & 0 \\ 0 & T_s \end{bmatrix},$$

$$\boldsymbol{\Psi} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

$\boldsymbol{W}(i) \in \mathbf{R}^2$ is the process noise and $\boldsymbol{V}(i) \in \mathbf{R}^2$ is the observation noise. $\boldsymbol{Q}_W$ and $\boldsymbol{Q}_V$ are the covariance matrices

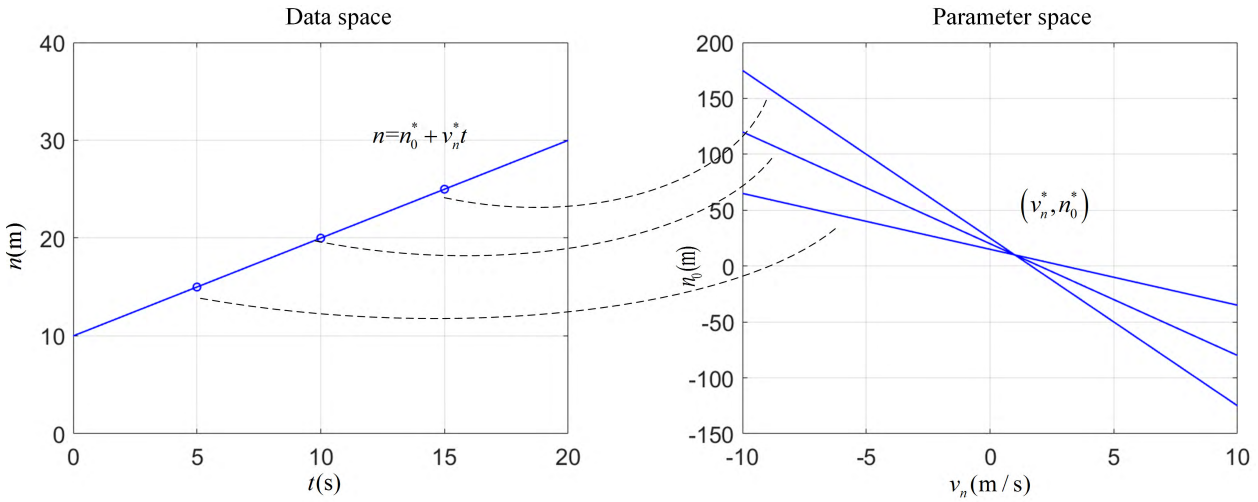


FIGURE 4. Fundamental principle of Hough Transform for straight line.

of $\boldsymbol{W}(i)$ and $\boldsymbol{V}(i)$ respectively.

$$\boldsymbol{Q}_W = \begin{bmatrix} \sigma_W^2 & 0 \\ 0 & \sigma_W^2 \end{bmatrix}, \quad \boldsymbol{Q}_V = \begin{bmatrix} \sigma_V^2 & 0 \\ 0 & \sigma_V^2 \end{bmatrix}$$

III. ESTIMATION FOR MOTION ELEMENTS OF MULTIPLE UNDERWATER TARGETS

Northern and eastern position-time sampling points of the targets are obtained by coordinate conversion of the sonar data, which form binary images. Typical underwater targets like suspended materials and other cruising UUVs move at relatively low velocities, and the motion curve of a given target usually appears as a straight line. Consequently, Hough transform for lines detection applies to the position-time curves of the multiple targets for estimation of velocity components at each degree of freedom (DOF). Although the function of image processing is integrated in the studied forward-looking sonar, the performances of the sonar are sensitive to hydrologic conditions like temperature, salinity and depth. Consequently, the false alarm and missing detection effects as well as positioning errors still exist in the sonar data. In order to overcome these defects, clustering statistic mechanism is introduced for SHT to realize parameter extraction.

In the implementation process, Hough transform is applied to the sampling data of the latest W instants for estimation of motion elements of multiple targets, including initial positions, velocities and current positions. W is the width of the time window. Motion elements at one DOF of the multiple targets are estimated at first, and that of other DOFs could be obtained by exploiting the spatial correlation of sonar data and single line extraction for each DOF. Without loss of generality, when the principle of CSHT method is explained, northern motion elements are estimated at first.

A. HOUGH TRANSFORM FOR BINARY SONAR IMAGE

The slope intercept form for lines is adopted in this paper, where the slope indicates the velocity, and the intercept indicates the initial position. The target velocities in discuss are finite values, so the problem of infinite slope is avoided. Take the northern movement as an example, the target motion equation is:

$$n = n_0 + v_n t \tag{6}$$

where t is the time variable, n indicates the northern position, n_0 indicates the northern initial position, v_n indicates northern velocity. The straight line in (6) is defined by the parameters n_0 and v_n. Accordingly, the detection data (t_i, n_i) define a straight line in the parameter space:

$$n_0 = -t_1 v_n + n_1 \tag{7}$$

The parameter space is discretized into cells, and a cell corresponds to a set of parameters representing image features. One curve in the parameter space contributes votes to the cells it passes by. After transformed, the collinear detection points in the data space form a cluster of straight lines that intersect at one point in the parameter space. Each straight line contributes one vote to the cell corresponding to the crossover point, and the crossover point corresponds to the parameters of the straight line, as is shown in Fig. 4. Consequently, parameter extraction for lines in the data space could be realized by votes statistics of the cells in the parameter space.

The parameter space is discretized at a certain resolution:

$$v_n \in \left\{v_n^i\right\}, \quad v_n^i = v_n^{\min} + h_{v_n}(i-1), \ i = 1, 2, \cdots, N_{v_n} \tag{8}$$

$$n_0 \in \left\{n_0^i\right\}, \quad n_0^i = n_0^{\min} + h_{n_0}(i-1), \ i = 1, 2, \cdots, N_{n_0} \tag{9}$$

The above symbols are defined as:

h_{vn}	: step length of parameter v_n;
N_{vn}	: number of discretized values of parameter v_n;
v_n^i	: the ith value of parameter v_n;
$v_n^{\min} = v_n^1$	: the minimum value of parameter v_n;

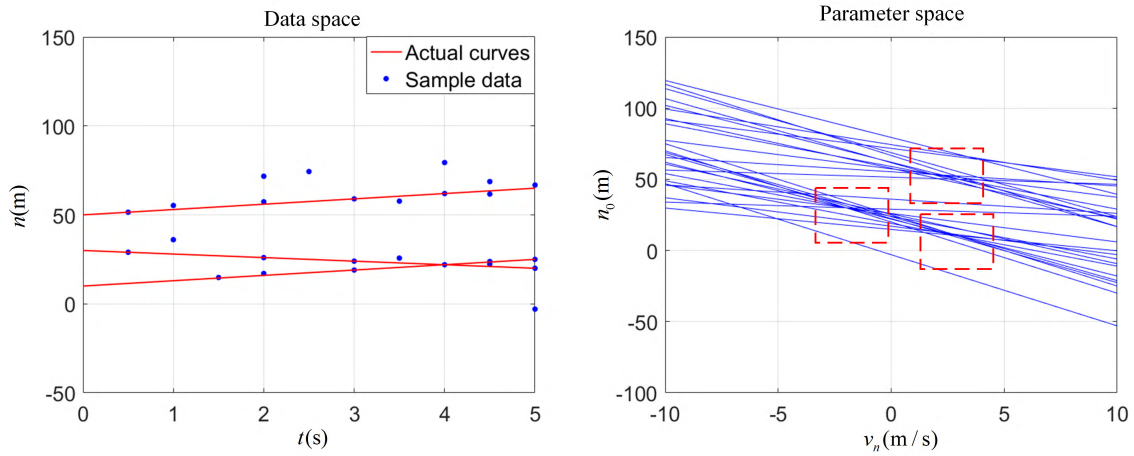


FIGURE 5. Cluttered sampling data in the data space and the corresponding Hough transformed lines in the parameter space.

$v_n^{\max} = v_n^{N_{vn}}$: the maximum value of parameter v_n;
h_{n_0} : step length of parameter n_0;
N_{n_0} : number of discretized values of parameter n_0;
n_0^i : the ith value of parameter n_0;
$n_0^{\min} = n_0^1$: the minimum value of parameter n_0;
$n_0^{\max} = n_0^{N_{n_0}}$: the maximum value of parameter n_0.

The velocity v_n is discretized into N_{vn} values at the resolution of h_{vn}, ranging from $v_n^{\min} = v_n^1$ to $v_n^{\max} = v_n^{N_{vn}}$; The initial position n_0 is discretized into N_{n_0} values at the resolution of h_{n_0}, ranging from $n_0^{\min} = n_0^1$ to $n_0^{\max} = n_0^{N_{n_0}}$. The parameter space is divided into $N_{vn} \times N_{n_0}$ cells, and it corresponds to a $N_{vn} \times N_{n_0}$ matrix $\boldsymbol{V}$, whose v_{ij} element corresponds to the parameter pair (v_n^i, n_0^j).

I pairs of position sampling data are obtained within W instants:

$$\left(t_i,\ \boldsymbol{\eta}_i\right), \quad i = 1, 2, \cdots, I \tag{10}$$

where $\boldsymbol{\eta}_i = [n_i, e_i]^{\mathrm{T}}$ is the position coordinates of a target at the moment of t_i. Northern sampling data are:

$$(t_i, n_i), \quad i = 1, 2, \cdots, I \tag{11}$$

and eastern sampling data are:

$$(t_i, e_i), \quad i = 1, 2, \cdots, I \tag{12}$$

Data matrix $\boldsymbol{D}$ is defined from northern sampling data for estimation of northern motion elements:

$$\boldsymbol{D} = \begin{bmatrix} t_1 & t_2 & \cdots & t_I \\ n_1 & n_2 & \cdots & n_I \end{bmatrix} \tag{13}$$

Transform matrix $\boldsymbol{H}$ is defined from discretized velocity values:

$$\boldsymbol{H} = \begin{bmatrix} -v_n^1 & 1 \\ -v_n^2 & 1 \\ \vdots & \vdots \\ -v_n^{N_{vn}} & 1 \end{bmatrix} \tag{14}$$

The product of $\boldsymbol{H}$ and $\boldsymbol{D}$ form a $N_{vn} \times I$ matrix $\boldsymbol{R}$:

$$\boldsymbol{R} = \boldsymbol{HD} = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1I} \\ r_{21} & r_{22} & \cdots & r_{2I} \\ \vdots & \vdots & \ddots & \vdots \\ r_{N_{vn},1} & r_{N_{vn},2} & \cdots & r_{N_{vn},I} \end{bmatrix} \tag{15}$$

$$r_{ij} \in \left[n_0^{j'}, n_0^{j'+1}\right) \tag{16}$$

The elements in $\boldsymbol{R}$ are values of the parameter n_0 in the parameter space. The element r_{ij} corresponds to the parameter pair (v_n^i, n_0^j), namely r_{ij} contributes one vote to the v_{ij}, element of the vote matrix $\boldsymbol{V}$. The vote matrix $\boldsymbol{V}$ is computed according to the matrix $\boldsymbol{R}$, and is further used for extraction of image features from which the motion elements are obtained.

B. CLUSTERING STATISTIC FOR EXTRACTION OF PARAMETERS

Parameters of straight lines could be extracted through peak detection by setting a threshold according to the vote matrix in the condition of accurate and adequate sensor information as well as weak noise. However, in the problem of estimation for motion elements of multiple underwater targets, because of the relatively low accuracy, the position-time curves of the targets provided by sonar appear as broken lines instead of standard straight lines. The straight lines in the parameter space transformed from the points in the broken lines in the data space form a cluster of points distributed in a certain area, rather than intersect at one point, as is shown in Fig. 5. In addition, the number of the targets is unknown, and the false alarm and missing detection effects are prominent, so the data space suffers from severe noise pollution and information loss. Several false targets might appear by simply setting a threshold for straight line parameter extraction. The votes belonging to a true target may be assigned to several neighboring cells and the true target may be submerged by noise. Failure may occur in the parameter extraction even if Hough Transform is adopted despite its robustness and fault tolerance.

The peaks corresponding to a target are closely distributed in the parameter space, so clustering method is considered for recognition of peaks in the parameter space corresponding to the multiple targets. The combination, division and delete operations of ISODATA clustering method could effectively overcome the adverse effects resulting from noise pollution and loss of information, and the unsupervised clustering mode solves the problem of unknown target number. In contrast with the widely used K-means method, ISODATA method could dynamically adjust the clustering centers and automatically correct the class number, which would restrain the noise to some extent. ISODATA method would improve its performance in objectivity, flexibility and simplicity if combined with fuzzy logic.

One curve in the parameter space contributes one vote to each of the cells it passes by. Many cells irrelevant to motion parameters of the targets have some votes that would not work for parameter extraction. Consequently, it is necessary to set a relatively low threshold for rejection of irrelevant votes before applying the Fuzzy ISODATA clustering method, which would also improve the efficiency of parameter extraction. A new vote matrix $\boldsymbol{V}'$ is obtained by truncating the initial vote matrix $\boldsymbol{V}$ with the primary threshold λ:

$$v'_{ij} = \begin{cases} v_{ij}, & v_{ij} > \lambda \\ 0, & v_{ij} \le \lambda, \end{cases} \quad i = 1, 2, \cdots, N_{v_n},\ j = 1, 2, \cdots, N_{n_0} \tag{17}$$

The clustering sample data are obtained from the vote matrix $\boldsymbol{V}'$. The number of clustering sample data is:

$$N = \sum_{i=1}^{N_{vn}} \sum_{j=1}^{N_{n_0}} v'_{ij} \tag{18}$$

The dimension of the sample data is $n = 2$. The element v'_{ij} of matrix $\boldsymbol{V}'$ provides v'_{ij} sample data:

$$\boldsymbol{x}_{ij}^{k} = \begin{bmatrix} v_n^i \\ n_0^j \end{bmatrix}, \quad k = 1, 2, \cdots, v'_{ij} \tag{19}$$

For convenience, the clustering sample data are rewritten as:

$${}^{\mathrm{s}}\boldsymbol{x}_i = \begin{bmatrix} {}^{\mathrm{s}}v_n^i \\ {}^{\mathrm{s}}n_0^i \end{bmatrix}, \quad i = 1, 2, \cdots, N \tag{20}$$

where the left superscript s denotes sample data. Symbols related to the Fuzzy ISODATA clustering algorithm are defined as follows:

N	: Number of sample data;
$K \in \mathbf{N}^*$	: Number of patterns, $K < N$;
P_i	: The ith pattern, $i = 1, 2, \cdots, K$;
$Z_i \in \mathbf{R}^{n\times 1}$	: The center of the ith pattern, $i = 1, 2, \cdots, K$;
$\boldsymbol{U} \in \mathbf{R}^{K\times N}$	: Fuzzy membership matrix, and the element μ_{ij} indicates the membership degree of the sample x_j belonging to the pattern P_i, $i = 1, 2, \cdots, K$, $j = 1, 2, \cdots, N$.

In the Fuzzy ISODATA method, a fuzzy membership matrix is established to indicate the membership between the sample data and the patterns. The fuzzy membership matrix is iteratively updated, and the combination, division and delete operations are performed for class adjustment until the convergence condition is satisfied.

Clustering criterion function measures the similarity and difference between the patterns, and it is a function of sampling data and patterns, which is the basis to realize the clustering process:

$$J = \sum_{i=1}^{K} \sum_{j=1}^{N} \mu_{ij}^{m} \left\| \boldsymbol{x}_j - \boldsymbol{Z}_i \right\| \tag{21}$$

For optimization of clustering criterion function, the elements of the fuzzy membership matrix are calculated as:

$$\mu_{ij}(n_{\mathrm{iter}}) = \frac{1}{\sum_{k=1}^{K(n_{\mathrm{iter}})} \left(\frac{d_{ij}}{d_{kj}}\right)^{2/(m-1)}}, \quad i = 1, 2, \cdots, K(n_{\mathrm{iter}});\ j = 1, 2, \cdots, N(n_{\mathrm{iter}}) \tag{22}$$

where d_{ij} denotes the distance from the sample x_j to the clustering center Z_i. The center of each pattern is updated as:

$$\boldsymbol{Z}_i(n_{\mathrm{iter}}) = \frac{\sum_{j=1}^{N(n_{\mathrm{iter}})} \mu_{ij}(n_{\mathrm{iter}})^m \boldsymbol{x}_j}{\sum_{j=1}^{N(n_{\mathrm{iter}})} \mu_{ij}(n_{\mathrm{iter}})^m}, \quad i = 1, 2, \cdots, K(n_{\mathrm{iter}}) \tag{23}$$

Fuzzy ISODATA method combines the advantages of fuzzy logic and ISODATA clustering algorithm, and the core of this method is the function of class adjustment, namely combination, division and delete operations are applied to the patterns. The number of patterns is variable in the clustering process. The function of class adjustment improves the clustering performance, especially for severely noisy sample data.

1) COMBINATION OPERATION

If the distance between the two clustering centers $\boldsymbol{Z}_i$ and $\boldsymbol{Z}_j$ is less than the combination threshold M_{ind}, $\boldsymbol{Z}_i$ and $\boldsymbol{Z}_j$ will be combined and a new clustering center $\boldsymbol{Z}^*$ is obtained:

$$\boldsymbol{Z}^* = \frac{\sum_{k=1}^{N(n_{\mathrm{iter}})} \mu_{ik}(n_{\mathrm{iter}}) \boldsymbol{Z}_i + \sum_{k=1}^{N(n_{\mathrm{iter}})} \mu_{jk}(n_{\mathrm{iter}}) \boldsymbol{Z}_j}{\sum_{k=1}^{N(n_{\mathrm{iter}})} \mu_{ik}(n_{\mathrm{iter}}) + \sum_{k=1}^{N(n_{\mathrm{iter}})} \mu_{jk}(n_{\mathrm{iter}})} \tag{24}$$

The combination threshold is:

$$M_{ind} = D[1 - F(K)] \tag{25}$$

where D is the average distance between the clustering centers. $F(K) \in [0, 1]$ is an artificially constructed decreasing function of K, and is usually set as:

$$F(K) = \frac{1}{K^{\alpha_c}} \tag{26}$$

where α_c is a parameter that could be designed.

2) DIVISION OPERATION

The Fuzzy variance of the jth feature of P_i is:

$$S_{ij}^2 = \frac{1}{N-1}\sum_{k=1}^{N}\mu_{ik}(x_{kj}-z_{ij})^2, \quad i=1,2,\cdots,K, \quad j=1,2,\cdots,n \tag{27}$$

where x_{kj} and z_{ij} respectively denote the jth feature of the sample data $\boldsymbol{x}_k$ and the clustering center $\mathbf{Z}_i$. The fuzzy variance threshold is:

$$F_{\text{std}} = S\,[1+G(K)] \tag{28}$$

where S is the mean of the fuzzy variances and $G(K)$ is an artificially constructed increasing function of K that is usually set as:

$$G(K) = \frac{1}{K^{\gamma}} \tag{29}$$

where γ is a parameter that could be designed to adjust the decomposition sensitivity. For check of aggregation degree of each pattern, C_i is calculated for P_i:

$$C_i = Q_i/T_i \tag{30}$$

$$Q_i = \sum_{k=1}^{N} t_{ik}\mu_{ik} \tag{31}$$

$$t_{ik} = \begin{cases} 0, & \mu_{ik} \le \theta \\ 1, & \mu_{ik} > \theta \end{cases} \tag{32}$$

$$T_i = \sum_{k=1}^{N} t_{ik} \tag{33}$$

where θ and A are parameters. If $C_i > A$, division operation is unnecessary because it implies the aggregation degree of P_i is high. Otherwise division operation is needed.

If $S_{ij} > F_{\text{std}}$, it needs to respectively add and minus S_{ij} on the jth feature of the clustering center $\mathbf{Z}_i$ of the pattern P_i, and two new clustering centers $\mathbf{Z}_i^+$ and $\mathbf{Z}_i^-$ are obtained:

$$Z_{ij}^+ = Z_{ij} + \lambda_{\text{d}} S_{ij} \tag{34}$$

$$Z_{ij}^- = Z_{ij} - \lambda_{\text{d}} S_{ij} \tag{35}$$

where λ_{d} is a division coefficient.

3) DELETE OPERATION

If the ith pattern satisfies either of the following conditions, it would be deleted:

Delete condition1:

$$T_i \le \delta \cdot \frac{N}{K} \tag{36}$$

This condition indicates that there are little sample data with high membership belonging to the pattern P_i. The parameter δ could be designed to adjust the deletion sensitivity.

Delete Condition2:

$$C_i \le A, \quad \textit{and } S_{ij} \le F_{\text{std}}, \ \forall j = 1,2,\cdots,n \tag{37}$$

This condition indicates that some sample data exist near the clustering center $\mathbf{Z}_i$, while the aggregation degree is relatively low, namely $\mathbf{Z}_i$ is not an ideal clustering center.

By finishing the clustering algorithm, ${}^*K$ clustering centers are obtained that correspond to the velocities and initial positions of the ${}^*K$ targets:

$${}^*\mathbf{Z}_i = \begin{bmatrix} {}^*v_n^i \\ {}^*n_0^i \end{bmatrix}, \quad i = 1,2,\cdots,{}^*K \tag{38}$$

The northern motion equations of the targets obtained from CSHT method are:

$$n^i = {}^*n_0^i + {}^*v_n^i t, \quad i = 1,2,\cdots,{}^*K \tag{39}$$

The current northern positions of the targets are obtained by substituting the current time ${}^*t$ into (39):

$${}^*n^i = {}^*n_0^i + {}^*v_n^i {}^*t, \quad i = 1,2,\cdots,{}^*K \tag{40}$$

Spatial correlation is performed to assign eastern sampling data to the targets according to the shortest distance principle by calculating the errors between northern sample data and estimated northern positions:

$$(t_j,e_j) \in D_i, \quad s.t. \ \left\| n^j - {}^*n^i \right\| = \min\left\{ \left\| n^j - {}^*n^k \right\|, k=1,2,\cdots,{}^*K \right\}, \quad j=1,2,\cdots,I \tag{41}$$

D_i is the eastern sampling data set of the ith target that forms a binary sonar image. CSHT is exploited for extraction of a single line to estimate the eastern motion elements of the ith target, including the velocity ${}^*v_e^i$, initial position ${}^*e_0^i$ and current position ${}^*e^i$.

The CSHT based estimation method for motion elements of multiple underwater targets is described in Algorithm 1.

IV. CSHT BASED UNDERWATER MULTIPLE TARGETS TACKING SYSTEM

Estimation for motion elements of multiple underwater targets could be realized by CSHT method at every moment. The voyage of the UUV is a continuous process, and CSHT based estimation for motion elements would be further used for multiple targets tracking jointly with data association technology in order to ensure grasp of motion states of the surrounding targets in real time in view of safe route planning [32], [33]. The estimation results of CSHT method involve not only the current positions of the targets, but also the velocities and the initial positions, which would be considered in the data association process for improvement of robustness and accuracy of the association results. However, certain errors still exist in the estimation results of CSHT method, and the errors at different moments are independent of each other. The tracking errors would increase if the position estimation results at different moments are simply connected. Consequently, a smoothing procedure is needed.

Algorithm 1 CSHT Based Estimation Method for Motion Elements of Multiple Underwater Targets

Given: The target detection data in the earth-fixed frame as is expressed in (11) and (12).

1. Hough transform.

- Define the data matrix $\boldsymbol{D}$ from the sampling data of one DOF, as is expressed in (13);
- Define the transform matrix $\boldsymbol{H}$ from the discretized velocity values, as is expressed in (14), and apply the Hough transform by multiplying the data matrix $\boldsymbol{D}$ with the transform matrix $\boldsymbol{H}$ to obtain the initial position matrix $\boldsymbol{R}$ as is expressed by (15);
- Compute the discretized initial position parameter corresponding to each element in the matrix $\boldsymbol{R}$;
- Compute the vote matrix $\boldsymbol{V}$ according to the matrix $\boldsymbol{R}$.

2. Clustering statistic.

- Obtain the clustering sample data according to the vote matrix $\boldsymbol{V}$, as is expressed in (20);
- Apply the fuzzy ISODATA algorithm to the clustering sample data to achieve the target number and velocities, initial positions and current positions of one DOF.

3. Spatial correlation.

- Assign the sampling data of other DOFs to the targets according to the shortest distance principle;
- Estimate the motion elements of other DOFs for each target by single line extraction through CSHT.

Output: The target number ${}^*K$, northern velocities ${}^*v_n^i$, eastern velocities ${}^*v_e^i$, northern initial positions ${}^*n_0^i$, eastern initial positions ${}^*e_0^i$, northern current positions ${}^*n^i$ and eastern current positions ${}^*e^i$ of the targets, $i = 1, 2, \cdots, {}^*K$.

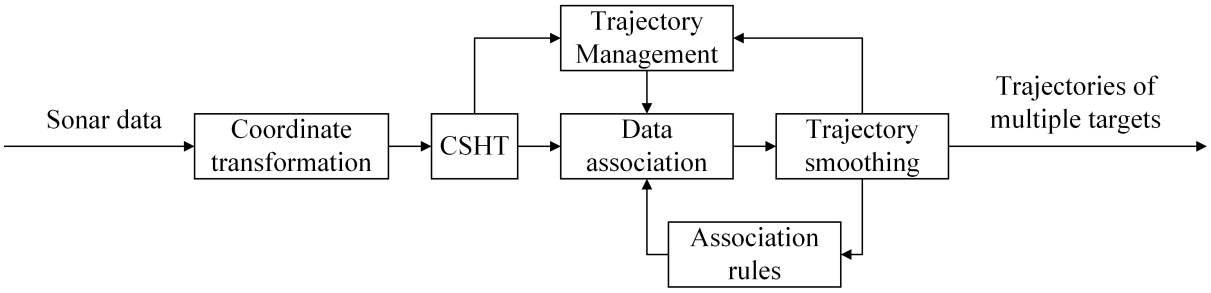


FIGURE 6. Functional block diagram of the CSHT based underwater multiple targets tacking system.

The functional block diagram of CSHT based underwater multiple targets tacking system is shown in Fig. 6. Firstly, sampling data in the earth-fixed frame are obtained from the sonar data by coordinate transform. Secondly, CSHT method is applied to the sampling data for estimation of motion elements of the multiple targets. Thirdly, data association is performed to incorporate the states of the targets into the set of existing targets, new targets or false targets. Fourthly, the associated trajectories of the targets are smoothened and the association rules are corrected according to the updated states of the targets. Fifthly, trajectory management is performed according to the motion elements and the smoothened states of the targets, including initiation, termination and quality of the trajectories, which would guide the subsequent data association. Finally, the motion states of the targets are output.

V. SIMULATION RESULTS AND DISCUSSION

Simulation of estimation for motion elements of multiple underwater targets is performed to verify the effectiveness of the proposed CSHT method. In order to verify the practicability of the CSHT method, simulation of underwater multiple targets tracking is performed as a further application of this method.

In the process of estimation for motion elements and targets tracking, for comprehensive evaluation of the effectiveness of target number estimation and target states estimation, Optimal Sub-Pattern Assignment (OSPA) distance is adopted as the evaluation criterion [34]. Assuming that $\boldsymbol{X} = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_m\}$ and $\boldsymbol{Y} = \{\boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_n\}$ are two finite sets, and $m, n \in \mathbf{N}$ are the element numbers of $\boldsymbol{X}$ and $\boldsymbol{Y}$. If $m \leq n$, OSPA distance is defined as:

$$\bar{d}_p^{(c)}(\boldsymbol{X}, \boldsymbol{Y}) = \left(\frac{1}{n} \cdot \left(\min_{\pi \in \Pi_n} \sum_{i=1}^{m} d^{(c)}\left(\boldsymbol{x}_i, \boldsymbol{y}_{\pi_i}\right)^p + c^p(n-m) \right) \right)^{1/p} \tag{42}$$

If $m > n$:

$$\bar{d}_p^{(c)}(\boldsymbol{X}, \boldsymbol{Y}) = \bar{d}_p^{(c)}(\boldsymbol{Y}, \boldsymbol{X}) \tag{43}$$

where $d^{(c)}$ is defined as:

$$d^{(c)}(\boldsymbol{x}, \boldsymbol{y}) = \min\left(c, d\left(\boldsymbol{x}, \boldsymbol{y}\right)\right) \tag{44}$$

In the context of motion elements estimation and tracking of multiple targets, assuming that $\boldsymbol{X}, \boldsymbol{Y} \subset \mathbf{R}^K$, $K \in \mathbf{N}^*$ and $d(\cdot)$ is a metric defined in $\mathbf{R}^K$, typically the Euclidean metric is adopted. Π_k is the set of permutations on $\{1, 2, \cdots, k\}$ for any $k \in \mathbf{N}^*$. c and p are cardinality and localization error sensitive parameters. $\bar{d}_p^{(c)}$ is called the OSPA metric of order p with cut-off c, where p stands for the order of the metric and c determines the weighting of how the metric penalizes cardinality errors as opposed to localization errors. The parameters are set as $c = 20$, $p = 2$ in the simulations. OSPA distance comprises of cardinality error $\bar{e}_{p,\text{card}}^{(c)}$ and localization error $\bar{e}_{p,loc}^{(c)}$ that respectively indicate the estimation effectiveness of target number and target states:

$$\bar{e}_{p,\text{card}}^{(c)}(\boldsymbol{X}, \boldsymbol{Y}) = \left(\frac{1}{n} \cdot c^p(n-m)\right)^{1/p} \tag{45}$$

$$\bar{e}_{p,\text{loc}}^{(c)}(\boldsymbol{X}, \boldsymbol{Y}) = \left(\frac{1}{n} \cdot \min_{\pi \in \Pi_n} \sum_{i=1}^{m} d^{(c)}\left(\boldsymbol{x}_i, \boldsymbol{y}_{\pi_i}\right)^p\right)^{1/p} \tag{46}$$

In the simulation process, several moving targets are set. UUV sails along the scheduled route. The sonar provides the distance and direction information of the targets that appear in the detection scope in theory. Northern and eastern coordinates of the targets are calculated according to (1) and (2), which further act as the input data for motion elements estimation.

In order to accord with the actual environmental conditions, false alarm and missing detection effects as well as positioning errors are considered in the generation of sonar data. The detection result of one beam at a certain instant is one of the following situations:

1) When some target exists, the sonar feeds back one target, which is correct. This situation is expressed as **Detection**, and is denoted as event E_D.
2) When some target exists, the sonar feeds back no targets, which is incorrect. This situation is expressed as **Missing Detection**, and is denoted as event E_M.
3) When no target exists, the sonar feeds back no targets, which is correct. This situation is expressed as **None Detection**, and is denoted as event E_N.
4) When no target exists, the sonar feeds back one target, which is incorrect. This situation is expressed as **False Alarm**, and is denoted as event E_F.

$E_\text{D}, E_\text{M}, E_\text{N}, E_\text{F}$ are mutually exclusive random events, and the occurrence probabilities are $P(E_\text{D}), P(E_\text{M}), P(E_\text{N}), P(E_\text{F})$:

$$P(E_D) + P(E_M) + P(E_N) + P(E_F) = 1 \tag{47}$$

Note that $P(E_\text{M})$ is missing detection rate, and $P(E_\text{F})$ is false alarm rate, namely $P(E_\text{M}) = P_\text{M}$, $P(E_\text{F}) = P_\text{F}$.

Aiming at one beam at a certain instant: when some target exists in the detection scope, one and only one of E_D and E_M will happen. If E_D happens, the sonar provides the target distance including positioning errors. The positioning error is white Gaussian noise with the mean of zero and the variance of σ^2. If E_M happens, no valid target distance will be feed back. When no target exists in the detection scope, one and only one of E_N and E_F will happen. If E_N happens, no valid target distance will be feed back. If E_F happens, the feedback target distance ρ obeys uniform distribution in the interval $(0, \rho_\text{D})$, namely $\rho \sim \text{U}(0, \rho_\text{D})$.

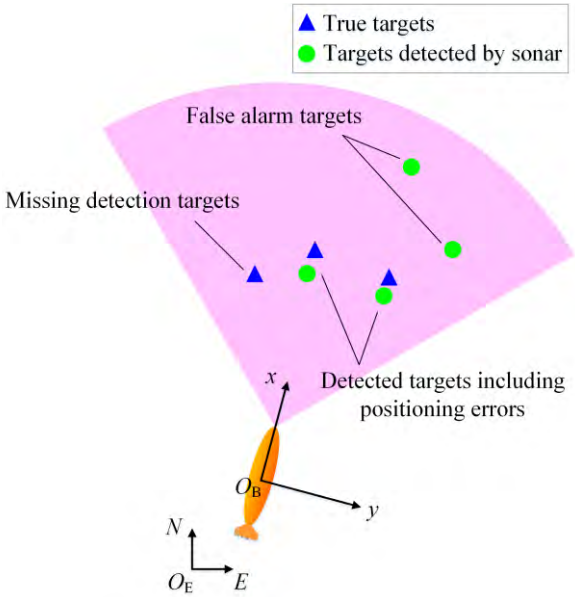


FIGURE 7. Simulation of sonar detection.

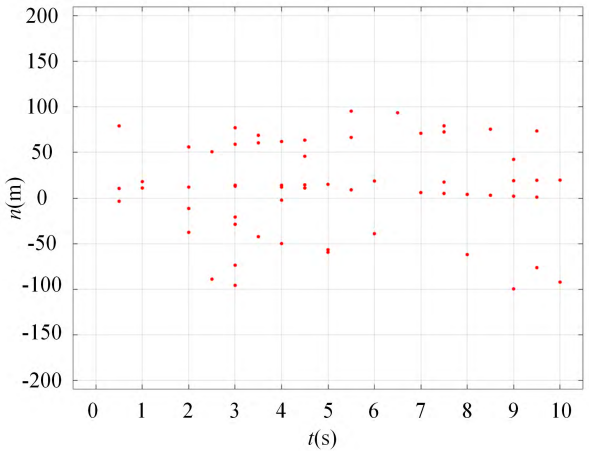


FIGURE 8. Northern position sampling data.

Fig. 7 shows the simulation of sonar detection.

The simulations are performed under the platform of Matlab 2016a on a computer with 4GB running memory and a 2.3 GHz processor.

A. SIMULATION OF FEATURE EXTRACTION

At each moment in the voyage of the UUV, motion elements of multiple targets could be estimated by CSHT method according to the sonar data of the latest W moments. The number of surrounding maneuvering targets and their velocities and initial positions as well as their current positions could be obtained.

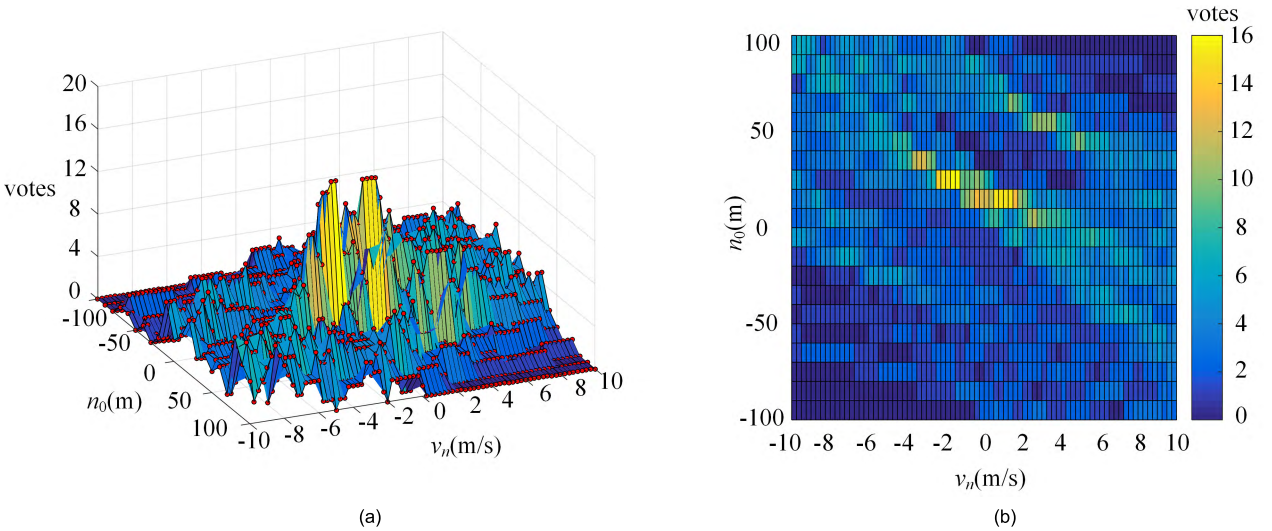


FIGURE 9. Vote distribution and its plane view in the parameter space. (a) Vote distribution. (b) Plane view of vote distribution.

In the simulation of estimation for motion elements, 3 targets moving at constant velocities along straight lines are set. The motion parameters of the UUV and the targets are shown in Table 2.

TABLE 2. Motion parameters of the UUV and the targets.

Moving Objects	Velocities(m/s)	Initial Positions(m)
UUV	(0.5,0.5)	(-10,-10)
Target 1	(1,4)	(10,80)
Target 2	(-2,1)	(20,30)
Target 3	(3,-3)	(50,60)

The false alarm rate and missing detection rate of the sonar are respectively set as $P_{\mathrm{F}} = 0.3$ and $P_{\mathrm{M}} = 0.3$, the standard deviation of the positioning error is as $\sigma = 3$ m, the sampling time is set as $T_{\mathrm{s}} = 0.5$ s, and the width of the time window is set as $W = 20$. Without loss of generality, the sonar data of the first W instants are selected for estimation. During the first W instants, the targets exist within the detection scope of the forward looking sonar. CSHT method is exploited for estimation of northern velocities and initial positions of the targets as well as their positions at the Wth instant. Furthermore, spatial correlation is performed and then the eastern velocity and initial position as well as current position of each target are obtained by single straight line extraction through CSHT.

The northern position sampling data of the targets are shown in Fig. 8. The vote distribution in the parameter space and the corresponding plane view are shown in Fig. 9. The estimation results of SHT method are shown in Fig. 10. The clustering results in the parameter space are shown in Fig. 11. The estimation results of CSHT method are shown in Fig. 12. The eastern position sampling data of the targets are shown in Fig. 13. The eastern estimation results of SHT and CSHT methods are respectively shown in Fig. 14 and Fig. 15.

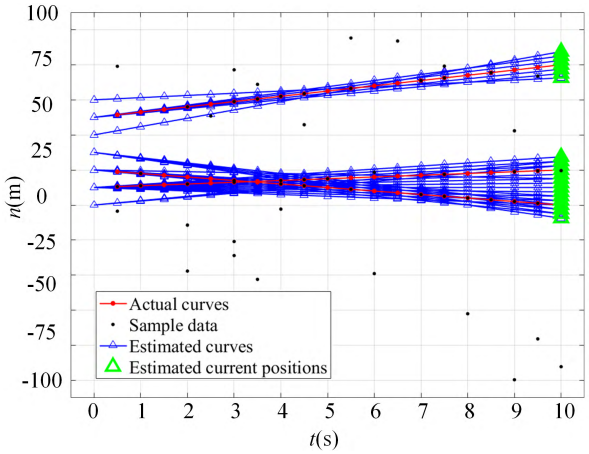


FIGURE 10. Estimation results by SHT.

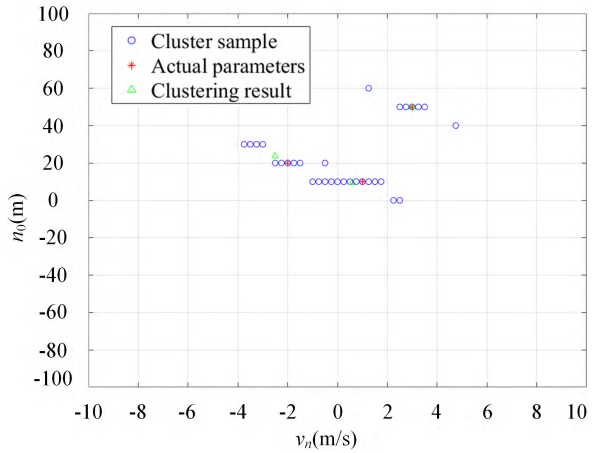


FIGURE 11. Clustering results in the parameter space.

It can be drawn from the simulation results that recognition for motion curves of the targets can be hardly accomplished according to the sonar data affected by false alarm and

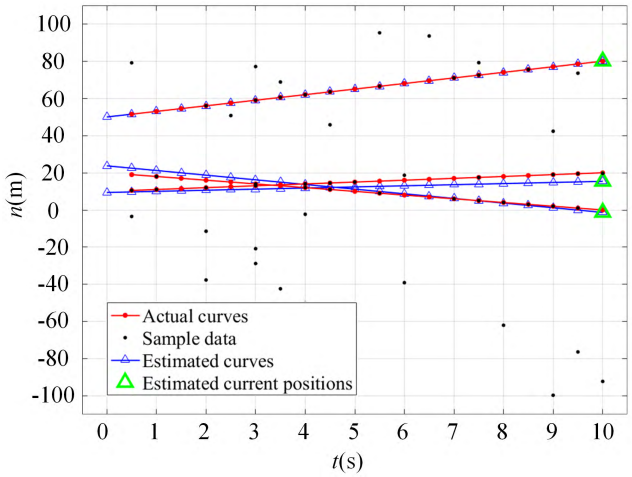


FIGURE 12. Estimation results by CSHT.

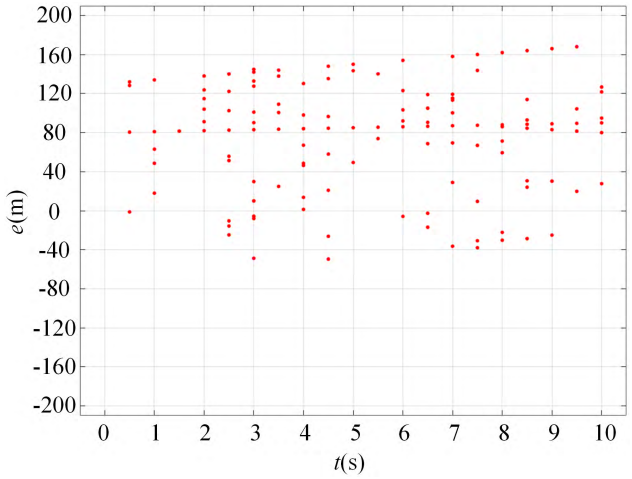


FIGURE 13. Eastern position sampling data.

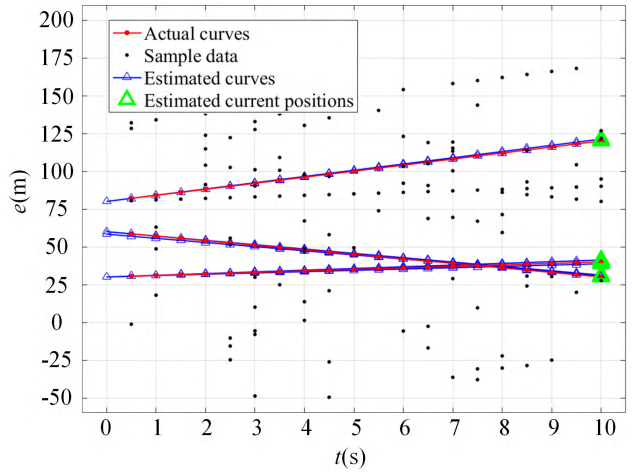


FIGURE 14. Eastern estimation results by SHT.

missing detection problems. Because of the massive clutter and measurement errors, even if Hough transform is applied, it is difficult to extract motion elements of the targets by simply setting a threshold for the vote matrix. Although the northern estimation results of SHT method contain the true targets, a large number of false targets also exist. The errors of eastern estimation results of SHT method are relatively small, because Hough transform is applied for single line detection for each target, which could be achieved by peak extraction. But redundancy is serious due to the false alarm problem of the northern estimation results. In contrast, the motion elements of the targets could be accurately extracted from the parameter space by CSHT method, including both the target number and the target states. The estimation results coincide with the actual situation, indicating that estimation for motion elements of multiple underwater targets could be accomplished by CSHT method in false alarm and missing detection condition.

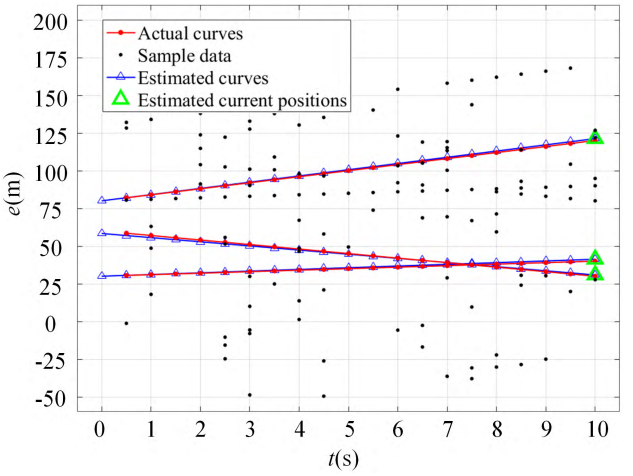


FIGURE 15. Eastern estimation results by CSHT.

For adaptability analysis against different false alarm and missing detection conditions as well as positioning errors, comparative simulations of CSHT and SHT methods are performed. The motion parameters of the targets are the same as Table 2. Firstly, the standard deviation of the positioning error is set at a medium level, $\sigma = 3$ m. The false alarm rate P_{F} and missing detection rate P_{M} respectively vary from 0.1 to 0.7, and 100 Monte Carlo simulations are performed for each condition. The velocity OSPA distances and position OSPA distances of the two methods in different conditions are shown in Table 3 and Table 4.

Accordingly, the variation tendency of velocity OSPA distances and position OSPA distances of the two methods are shown in Fig. 16 and Fig. 17.Secondly, the false alarm rate P_{F} and missing detection rate P_{M} are both set at medium levels, $P_{\mathrm{F}} = 0.3$ and $P_{\mathrm{M}} = 0.3$. The standard deviation of the positioning error σ varies from 1 m to 5 m. 100 Monte Carlo simulations are performed for each condition. The velocity OSPA distances and position OSPA distances of the two methods in different conditions are shown in Table 5 and Table 6. Accordingly, the variation tendency of velocity OSPA distances and position OSPA distances of the two methods are shown in Fig. 18 and Fig. 19.

TABLE 3. Velocity OSPA distances of motion elements estimation for multiple underwater targets in different false alarm and missing detection conditions.

$\bar{d}^{(c)}_{p,\mathrm{vel}}$(m/s)	P_F=0.1		P_F=0.3		P_F=0.5		P_F=0.7	
	CSHT	SHT	CSHT	SHT	CSHT	SHT	CSHT	SHT
P_M=0.1	0.13	2.89	0.15	3.28	0.18	3.76	0.26	4.72
P_M=0.3	0.16	3.78	0.18	4.79	0.21	5.42	0.33	5.59
P_M=0.5	0.19	5.22	0.21	5.21	0.26	6.37	0.35	6.72
P_M=0.7	0.22	5.19	0.26	5.36	0.32	6.83	0.42	7.85

TABLE 4. Position OSPA distances of motion elements estimation for multiple underwater targets in different false alarm and missing detection conditions.

$\bar{d}^{(c)}_{p,\mathrm{pos}}$(m)	P_F=0.1		P_F=0.3		P_F=0.5		P_F=0.7	
	CSHT	SHT	CSHT	SHT	CSHT	SHT	CSHT	SHT
P_M=0.1	1.33	12.42	1.98	15.32	2.22	16.63	2.51	18.62
P_M=0.3	1.82	17.37	2.14	22.43	2.79	21.39	2.96	24.33
P_M=0.5	1.95	18.22	2.95	23.37	3.11	26.53	3.45	25.44
P_M=0.7	2.36	15.79	3.52	26.45	3.78	30.53	4.33	32.51

TABLE 5. Velocity OSPA distances of motion elements estimation for multiple underwater targets in different positioning error conditions.

$\bar{d}^{(c)}_{p,\mathrm{vel}}$(m/s)	σ=1 m	σ=2 m	σ=3 m	σ=4 m	σ=5 m
CSHT	0.13	0.17	0.18	0.22	0.25
SHT	3.26	4.03	4.79	5.12	5.58

TABLE 6. Position OSPA distances of motion elements estimation for multiple underwater targets in different positioning error conditions.

$\bar{d}^{(c)}_{p,\mathrm{pos}}$(m)	σ=1 m	σ=2 m	σ=3 m	σ=4 m	σ=5 m
CSHT	1.29	1.75	2.14	2.34	2.85
SHT	18.93	20.19	22.43	25.18	27.26

The simulation results show that the estimation performance of SHT are satisfactory by properly selecting a threshold in low-noise condition, because the peaks corresponding to the true targets in the parameter space are concentratively distributed. However, noise points in the data space increase as the environmental condition becomes worse, and the votes corresponding to the true targets in the parameter space appear to be relatively scattered. Setting a threshold cannot guarantee to filter out all false alarm targets, and a number of redundant targets appear near a true target, which results in an inaccurately estimated target number and degrades the estimation performance. In contrast, CSHT is more adaptable to positioning errors, false alarm and missing detection problems. The estimated target number is correct, and the accuracy of the estimation results is barely degraded with the growing positioning errors, false alarm rate and missing detection rate. The velocity OSPA distances are lower than 0.5 m/s and the position OSPA distances are lower than 5 m.

B. SIMULATION OF UNDERWATER MULTIPLE TARGETS TRACKING

The proposed CSHT method and the widely used Kalman filter are applied to underwater multiple targets tracking problem along with data association technology for comparative simulations. Kalman filter for each target works according to the state space model expressed by (5). σ_W is set as $2\sigma/T_s^2$, and σ_V is set as σ. The initial state for each target is set as $\boldsymbol{X}_0 = \begin{bmatrix} 0 & 0 & 0 & 0 \end{bmatrix}^{\mathrm{T}}$, the initial covariance is set as:

$$\boldsymbol{P}_{0|0} = \begin{bmatrix} 5^2 & 0 & 0 & 0 \\ 0 & 5^2 & 0 & 0 \\ 0 & 0 & 2^2 & 0 \\ 0 & 0 & 0 & 2^2 \end{bmatrix}$$

5 targets moving at constant velocities are set, and the simulation process lasts for 100 s. The targets are not always visible throughout the simulation process because of the limitation of detection distance of the forward looking sonar. Each target appears and disappears at two certain moments.

TABLE 7. Motion parameters of the UUV and the targets

Moving Objects	Velocities(m/s)	Initial Positions(m)	Appearing moment(s)	Disappearing moment(s)
UUV	(0.87,0.5)	(0,0)	1	100
Target 1	(-0.5,-1)	(120,80)	1	69
Target 2	(2,0.25)	(20,0)	1	100
Target 3	(-0.25,1.5)	(70,-10)	1	59
Target 4	(-1.5,0.5)	(200,20)	22	88
Target 5	(0.25,-0.75)	(120,150)	32	100

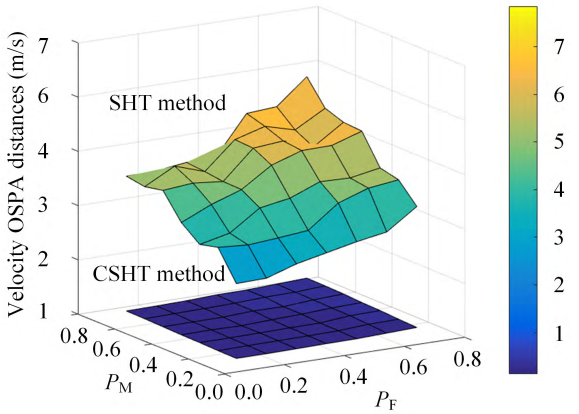


FIGURE 16. Velocity OSPA distances in different false alarm and missing detection conditions.

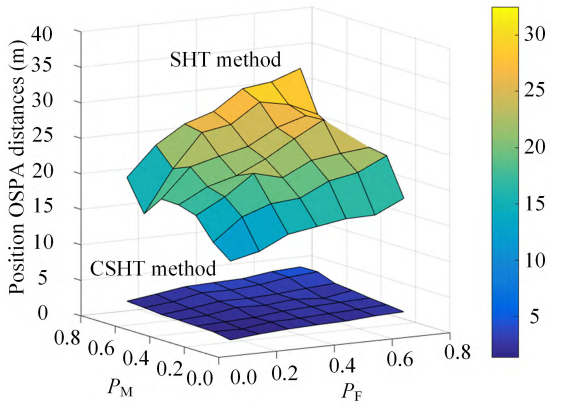


FIGURE 17. Position OSPA distances in different false alarm and missing detection conditions.

The motion parameters of the UUV and the targets are shown in Table 7.

The false alarm rate and missing detection rate of the sonar are respectively set as $P_F = 0.5$ and $P_M = 0.5$, the standard deviation of the positioning error is as $\sigma = 3$ m, the sampling time is set as $T_s = 0.5$ s, and the width of the time window is set as $W = 20$. The target position curves of CSHT method and Kalman method are shown in Fig. 20. The tracking results of CSHT method and Kalman method are shown in Fig. 21. The cardinality and localization errors as well as the OSPA distances of the two methods are compared in Fig. 22.

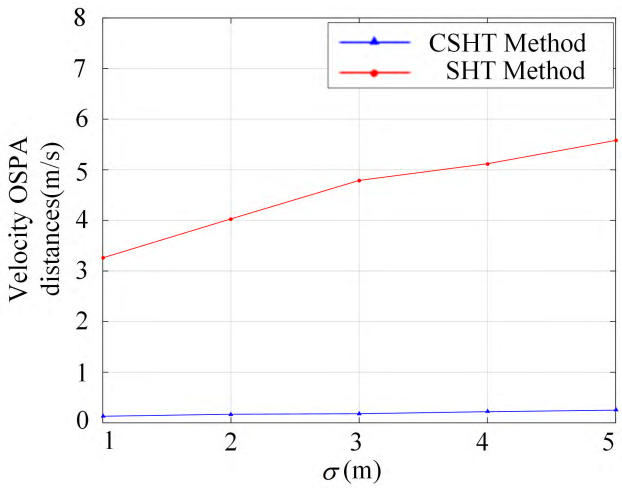


FIGURE 18. Velocity OSPA distances in different positioning error conditions.

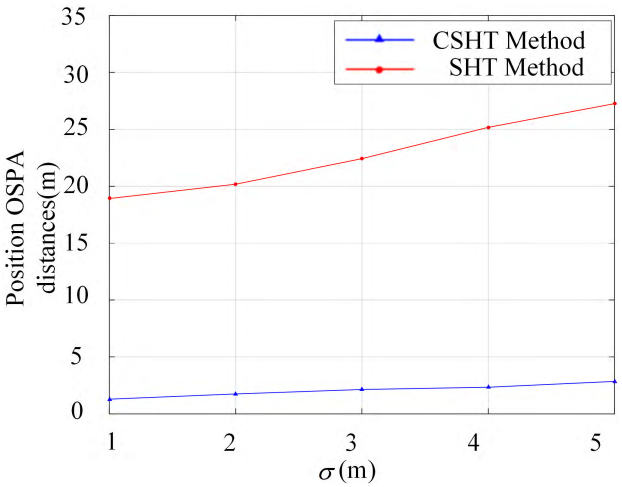


FIGURE 19. Position OSPA distances in different positioning error conditions.

Obviously, when CSHT method is applied to the multiple targets tracking problem, the OSPA distances are relatively large within the first 10 seconds or so since a new target appears. Once the targets are locked, the cardinality errors remain zero and the targets are consistently and stably tracked

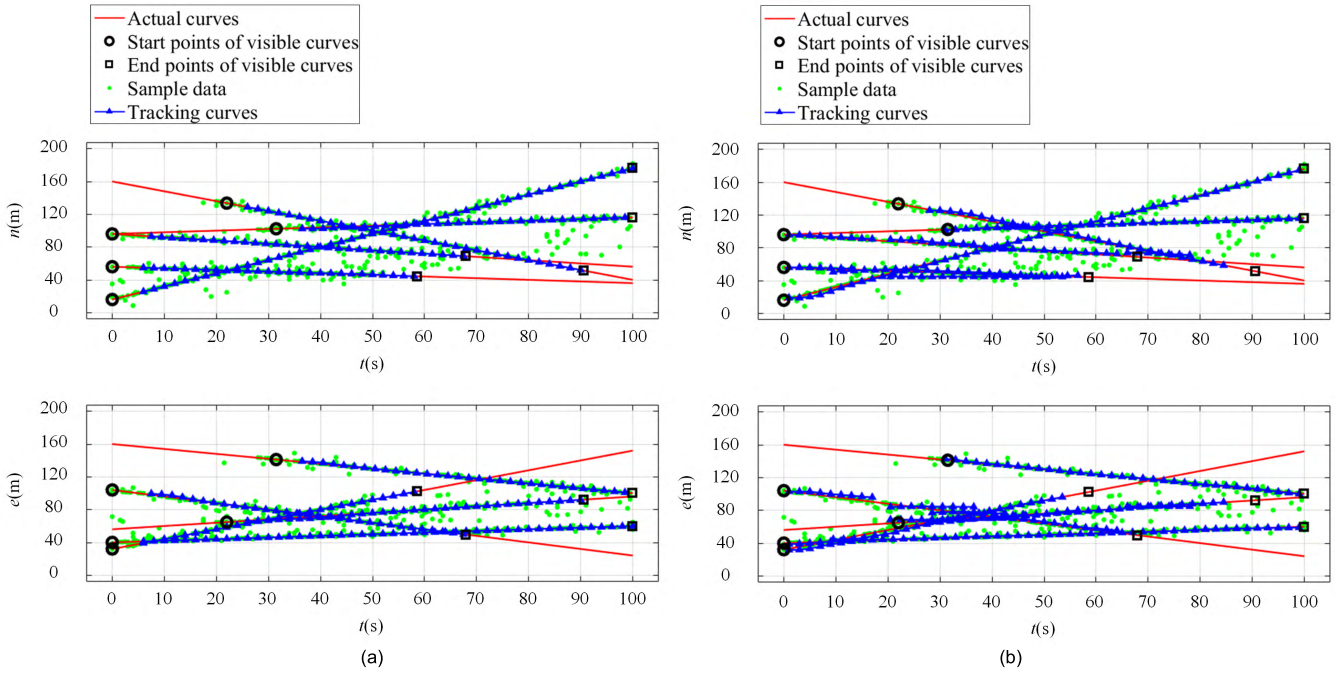


FIGURE 20. The target position curves. (a) CSHT method. (b) Kalman method.

TABLE 8. Cardinality errors of multiple targets tracking in different false alarm and missing detection conditions.

$\bar{e}^{(c)}_{p,\mathrm{card}}$(m)	P_F=0.1		P_F=0.3		P_F=0.5		P_F=0.7	
	CSHT	Kalman	CSHT	Kalman	CSHT	Kalman	CSHT	Kalman
P_M=0.1	1.39	3.14	1.36	3.50	1.46	2.68	1.47	5.15
P_M=0.3	1.50	4.36	1.53	3.45	1.65	5.73	2.34	6.32
P_M=0.5	2.19	4.05	2.56	7.15	2.57	6.37	2.78	5.75
P_M=0.7	2.75	5.42	3.03	5.33	3.77	4.85	3.33	9.18

TABLE 9. Localization errors of multiple targets tracking in different false alarm and missing detection conditions.

$\bar{e}^{(c)}_{p,\mathrm{loc}}$ / m	P_F=0.1		P_F=0.3		P_F=0.5		P_F=0.7	
	CSHT	Kalman	CSHT	Kalman	CSHT	Kalman	CSHT	Kalman
P_M=0.1	1.14	1.01	1.05	1.16	1.18	1.08	1.22	1.17
P_M=0.3	1.22	1.34	1.31	1.48	1.28	1.74	1.46	1.97
P_M=0.5	1.32	1.80	1.33	2.67	1.42	2.77	1.55	2.82
P_M=0.7	1.37	1.91	1.49	2.31	1.91	2.75	1.93	3.02

with low localization errors. The average OSPA distance is lower than 3.5 m, which is lower than 50% of that of Kalman method. When Kalman method is applied, the localization errors are relatively low owning to the intrinsic practicability and effectiveness of Kalman method. But the cardinality errors are large due to the severe false tracking and missing tracking that are caused by the clutter. The tracking trajectories appear to be incoherent and the OSPA distances are relatively large. As to time consumption, the average computing time for each instant of CSHT method and Kalman method are 0.1537 s and 0.0110 s. So CSHT method improves the estimation performance at the cost of great time consumption. However, the computational efficiency of CSHT method is sufficient to meet the application requirement of 0.5 s sampling time. Moreover, a more effective programming language would be selected to realize the algorithm in practical engineering applications, which would further reduce the computing time. Consequently, computational efficiency would not limit the application of CSHT method.

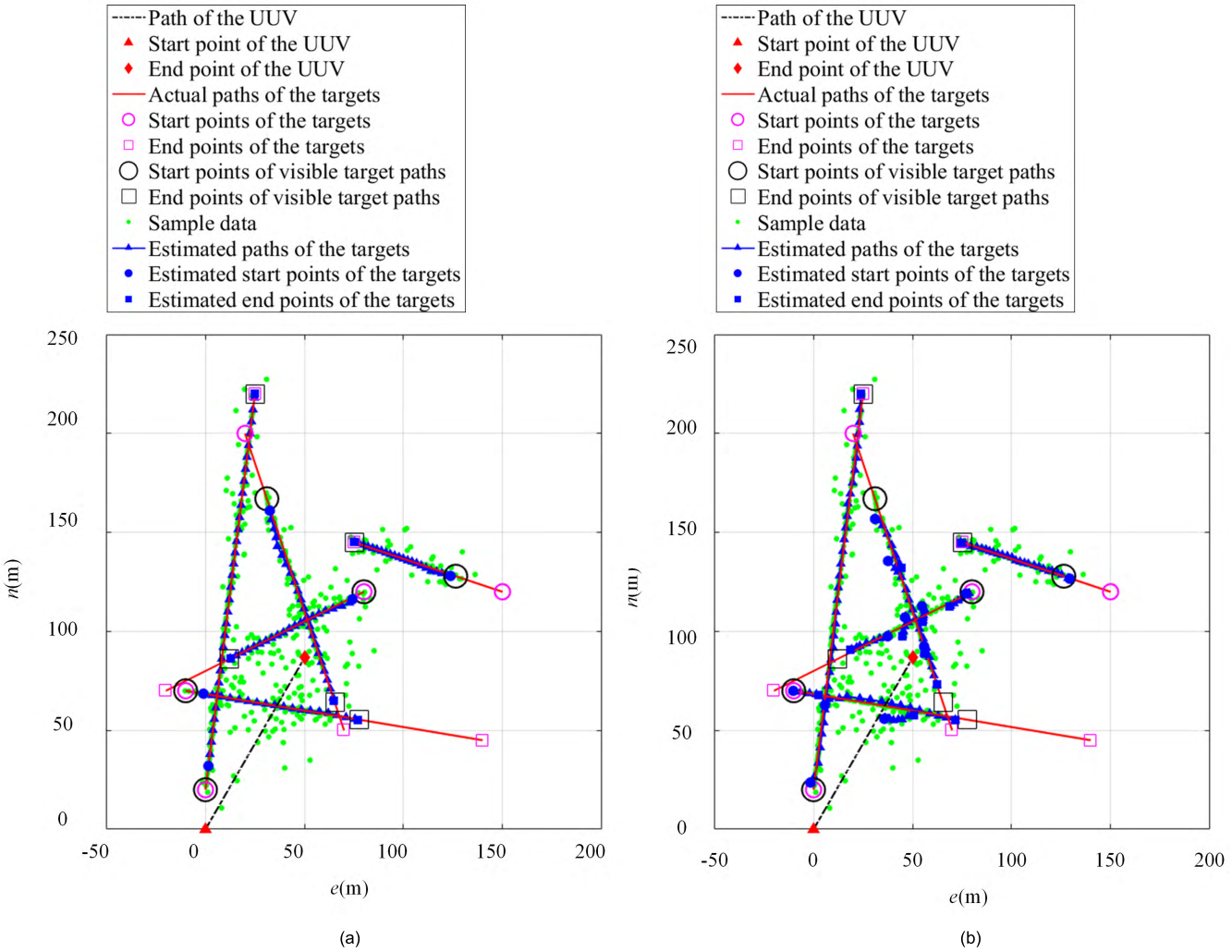


FIGURE 21. The tracking results. (a) CSHT method. (b) Kalman method.

TABLE 10. OSPA distances of multiple targets tracking in different false alarm and missing detection conditions.

$\bar{d}_p^{(c)}$ / m	P_F=0.1		P_F=0.3		P_F=0.5		P_F=0.7	
	CSHT	Kalman	CSHT	Kalman	CSHT	Kalman	CSHT	Kalman
P_M=0.1	1.80	3.30	1.72	3.69	1.88	2.89	1.91	5.28
P_M=0.3	1.93	4.56	2.01	3.75	2.09	5.99	2.76	6.62
P_M=0.5	2.56	4.43	2.88	7.63	2.94	6.95	3.18	6.40
P_M=0.7	3.07	5.75	3.38	5.81	4.23	5.58	3.85	9.66

TABLE 11. Cardinality errors of multiple targets tracking in different positioning error conditions.

$\bar{e}_{p,\mathrm{card}}^{(c)}$(m)	σ=1 m	σ=2 m	σ=3 m	σ=4 m	σ=5 m
CSHT	0.98	1.12	1.53	1.79	2.02
Kalman	1.68	2.17	3.45	3.92	5.23

To capture the adaptability against different positioning errors, false alarm and missing detection conditions, CSHT method and Kalman method are applied to the multiple targets tracking problem in different conditions. The motion parameters of the targets are the same as Table 7. Firstly, the standard deviation of the positioning error is set at

TABLE 12. Localization errors of multiple targets tracking in different positioning error conditions

$\bar{e}_{p,\mathrm{loc}}^{(c)}$ / m	σ=1 m	σ=2 m	σ=3 m	σ=4 m	σ=5 m
CSHT	0.82	1.07	1.31	1.33	1.86
Kalman	0.93	1.22	1.48	1.93	2.29

TABLE 13. OSPA distances of multiple targets tracking in different positioning error conditions.

$\bar{d}_{p}^{(c)}$ / m	σ=1 m	σ=2 m	σ=3 m	σ=4 m	σ=5 m
CSHT	1.27	1.55	2.01	2.23	2.75
Kalman	1.92	2.49	3.75	4.37	5.71

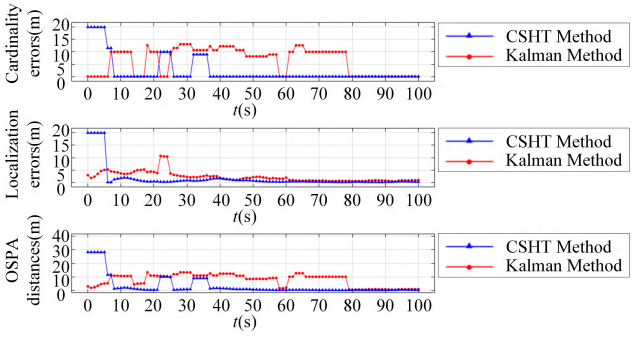


FIGURE 22. The cardinality errors, localization errors and OSPA distances.

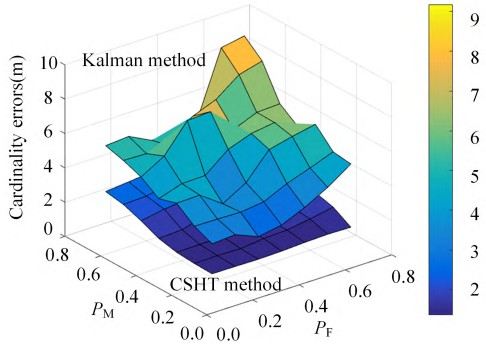


FIGURE 23. Cardinality errors in different false alarm and missing detection conditions.

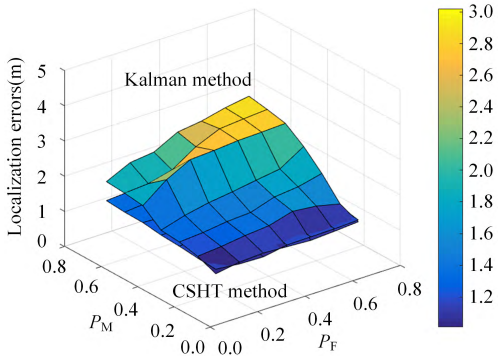


FIGURE 24. Localization errors in different false alarm and missing detection conditions.

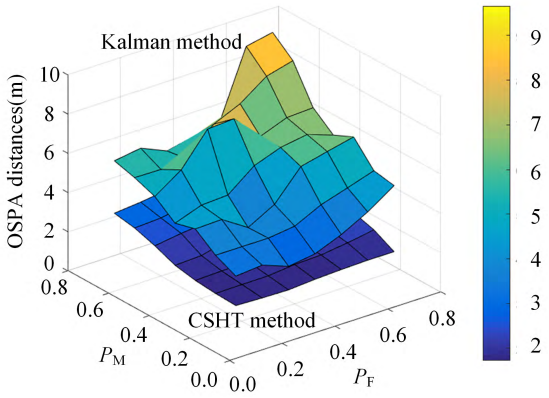


FIGURE 25. OSPA distances in different false alarm and missing detection conditions.

a medium level, $\sigma = 3$ m. The false alarm rate P_{F} and missing detection rate P_{M} respectively vary from 0.1 to 0.7, and 100 Monte Carlo simulations are performed for each method in each condition. The cardinality errors, localization errors and OSPA distances in different scenarios are listed in Table 8, Table 9 and Table 10. Accordingly, the variation tendency of cardinality errors, localization errors and OSPA distances of the two methods are shown in Fig. 23, Fig. 24 and Fig. 25. Secondly, the false alarm rate P_{F} and missing detection rate P_{M} are both set at medium levels, $P_{\mathrm{F}} = 0.3$ and $P_{\mathrm{M}} = 0.3$. The standard deviation of the positioning error σ varies from 1 m to 5 m. 100 Monte Carlo simulations are performed for each method in each condition. The cardinality errors, localization errors and OSPA distances in different scenarios are listed in Table 11, Table 12 and Table 13. Accordingly, the variation tendency of cardinality errors, localization errors and OSPA distances are shown in Fig. 26, Fig. 27 and Fig. 28.

The simulation results show that CSHT method provides satisfactory tracking performances in different positioning errors, false alarm and missing detection conditions. The cardinality errors and localization errors are lower than that

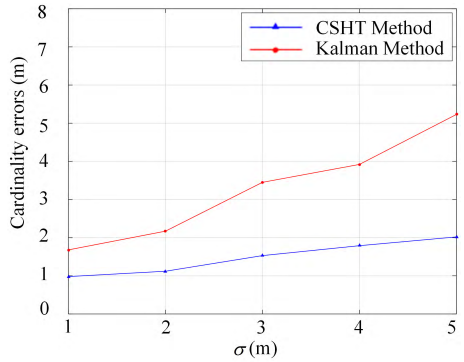


FIGURE 26. Cardinality errors in different positioning error conditions.

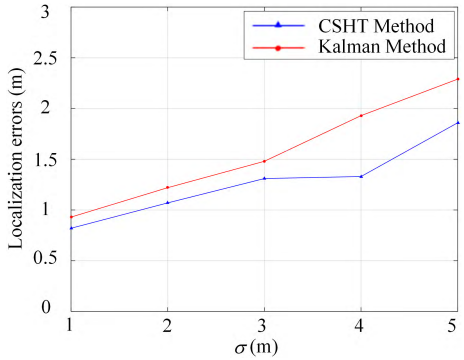


FIGURE 27. Localization errors in different positioning error conditions.

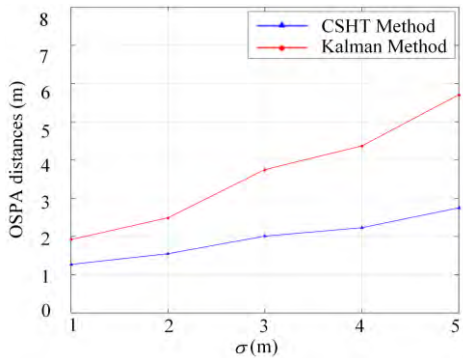


FIGURE 28. OSPA distances in different positioning error conditions.

of Kalman method, and the advantage of cardinality errors is particularly obvious. The OSPA distances remain below 5 m, indicating excellent adaptability to environment. Kalman method performs well in multiple targets tracking in favorable environment conditions. However, the false tracking and missing tracking problem become more and more prominent as the positioning errors, false alarm rate and missing detection rate increase. The OSPA distances increase and the tracking performances are degraded obviously, indicating poor adaptability to environment.

VI. CONCLUSIONS

In this paper we have presented the Clustering Statistic Hough Transform (CSHT) method for estimation of motion elements of multiple underwater targets in order to overcome the positioning errors, false alarm and missing detection effects of the sonar data and improve the accuracy and reliability of the estimation results. The data from the forward-looking sonar mounted on the UUV are used. Firstly the mathematical model of sonar vision field as well as the processing method of sonar data is introduced; secondly the principle of CSHT is elaborated, which applies fuzzy ISODATA clustering method to feature extraction in the Hough space; and then the application of CSHT method in the underwater multiple targets tracking system is explained. Simulation results show that CSHT method could accurately estimate motion elements of multiple underwater targets in noisy and incomplete information conditions, indicating insensitiveness to positioning errors, false alarm and missing detection effects of sonar. Furthermore, when applying CSHT method to underwater multiple targets tracking, the false and missing tracking phenomena are significantly reduced in contrast with the widely used Kalman filter method. The OSPA distances in the tracking process by CSHT method are lower than 5 m even in a severe condition, indicating high positioning and tracking accuracy. The computational efficiency of CSHT method meets the application requirement, so it is applicable in engineering practice. The method presented in this paper would contribute to enhancement of situation assessment capability of the UUV towards the surrounding environment and provide basis for further decision-making behaviors like collision avoidance and path planning.

It is assumed in this paper that underwater targets move at constant velocities, which is reasonable and accords with the actual underwater environment condition to some extent. However, maneuvering underwater targets with accelerations indeed exist. In order to guarantee situation assessment capability of the UUV, future studies will focus on motion elements estimation of maneuvering underwater targets with accelerations.

REFERENCES

[1] M. N. V. S. Kumar, N. Modalavalasa, L. Ganesh, K. S. Prasad, and G. S. Rao, "A new approach for tracking moving objects in underwater environment," *Current Sci.*, vol. 110, no. 7, pp. 1315–1323, Apr. 2016.

[2] X. Li, Y. Li, J. Yu, X. Chen, and M. Dai, "PMHT approach for multi-target multi-sensor sonar tracking in clutter," *Sensors*, vol. 15, no. 11, pp. 28177–28192, Nov. 2015.

[3] L. Zhang, L. Zhang, S. Liu, J. Zhou, and C. Papavassiliou, "Three-dimensional underwater path planning based on modified wolf pack algorithm," *IEEE Access*, vol. 5, pp. 22783–22795, 2017.

[4] Y. Zhang and F. Xian, "Multiplication-based pulse integration for detecting underwater target in impulsive noise environment," *IEEE Access*, vol. 4, pp. 6894–6900, 2016.

[5] L. Bjorno, "Developments in sonar technologies and their applications," in *Proc. IEEE Int. Underwater Technol. Symp.*, Tokyo, Japan, Mar. 2013, pp. 1–8.

[6] T. Wagner, R. Feger, and A. Stelzer, "Radar signal processing for jointly estimating tracks and micro-doppler signatures," *IEEE Access*, vol. 5, pp. 1220–1238, 2017.

[7] K. Naus and A. Nowak, "The positioning accuracy of BAUV using fusion of data from USBL system and movement parameters measurements," *Sensors*, vol. 16, no. 8, p. 1279, Aug. 2016.

[8] L. Úbeda-Medina, A. F. García-Fernández, and J. Grajal, "Adaptive auxiliary particle filter for track-before-detect with multiple targets," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 53, no. 5, pp. 2317–2330, Oct. 2017.

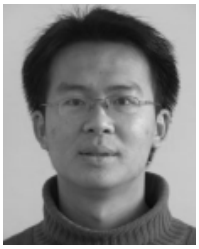
[9] S. Kim, J. Cho, and D. Park, "Moving-target position estimation using GPU-based particle filter for IoT sensing applications," *Appl. Sci.*, vol. 7, no. 11, p. 1152, Sep. 2017.
[10] R. P. Mahler, "A theoretical foundation for the stein-winter 'probability hypothesis density (PHD)," in *Proc. MSS Nat. Symp. Sensor Data Fusion Multitarget Tracking Approach*, Jun. 2000.
[11] Y. Zheng, Z. Shi, R. Lu, S. Hong, and X. Shen, "An efficient data-driven particle PHD filter for multitarget tracking," *IEEE Trans. Ind. Informat.*, vol. 9, no. 4, pp. 2318–2326, Nov. 2013.
[12] Y. Luo, L. Pu, H. Mo, Y. Zhu, Z. Peng, and J. Cui, "Receiver-initiated spectrum management for underwater cognitive acoustic network," *IEEE Trans. Mobile Comput.*, vol. 16, no. 1, pp. 198–212, Jan. 2016.
[13] P. V. C. Hough, "Method and means for recognizing complex patterns," U.S. Patent Dec. 3069654, Dec. 18, 1962.
[14] P. Mukhopadhyay and B. B. Chaudhuri, "A survey of Hough transform," *Pattern Recognit.*, vol. 48, no. 3, pp. 993–1010, Mar. 2014.
[15] L. S. Davis, "Hierarchical generalized Hough transforms and line-segment based generalized Hough transforms," *Pattern Recognit.*, vol. 15, no. 4, pp. 277–285, Dec. 1982.
[16] C. Kannan and H. Chuang, "Fast Hough transform on a mesh connected processor array," *Inf. Process. Lett.*, vol. 33, no. 5, pp. 243–248, Jan. 1990.
[17] L. Xu, E. Oja, and P. Kultanen, "A new curve detection method: Randomized Hough transform (RHT)," *Pattern Recognit. Lett.*, vol. 11, no. 5, pp. 331–338, 1990.
[18] N. Kiryati, Y. Eldar, and A. M. Bruckstein, "A probabilistic Hough transform," *Pattern Recognit.*, vol. 24, no. 4, pp. 303–316, 1991.
[19] V. F. Leavers, D. Ben-Tzvi, and M. B. Sandier, "A dynamic combinatorial Hough transform for straight lines and circles," in *Proc. Alvey Vis. Conf.*, Jan. 1989, pp. 163–168.
[20] S. Y. K. Yuen, T. S. L. Lam, and N. K. D. Leung, "Connective Hough transform," in *Proc. Brit. Mach. Vis. Conf.*, Glasgow, U.K., Jan. 1991, pp. 127–135.
[21] R. Huber-Shalem, O. Hadar, S. R. Rotman, and M. Huber-Lerner, "Parametric temporal compression of infrared imagery sequences containing a slow-moving point target," *Appl. Opt.*, vol. 55, no. 5, pp. 1151–1163, Feb. 2016.
[22] M. S. Islam and U. Chong, "Improvement in moving target detection based on Hough transform and wavelet," *IETE Tech. Rev.*, vol. 32, no. 1, pp. 46–51, Jan. 2015.
[23] P. Lei and X. Huang, "Robust detection of moving human target in foliage-penetration environment based on Hough transform," *Radioengineering*, vol. 23, no. 1, pp. 3–10, Apr. 2014.
[24] D. Yipeng, S. Kehui, and X. Xuemei, "Hough-MHAF localization algorithm for dual-frequency continuous-wave through-wall radar," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 52, no. 1, pp. 111–121, Feb. 2016.
[25] J. Yang, C. Liu, and Y. Wang, "Imaging and parameter estimation of fast-moving targets with single-antenna SAR," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 2, pp. 529–533, Feb. 2014.
[26] B. Xu, Q. Chen, and Z. Wang, "Track initiation with ant colony optimization," *Commun. Nonlinear Sci. Numer. Simul.*, vol. 14, nos. 9–10, pp. 3629–3644, Sep. 2009.
[27] B. Hadjira, M. Abdelkrim, M. Keche, A. Ouamri, and M. S. Woolfson, "Real time hough transform based track initiators in clutter," *Inf. Sci.*, vols. 337–338, pp. 82–92, Apr. 2016.
[28] L. R. Moyer, J. Spak, and P. Lamanna, "A multi-dimensional hough transform-based track-before-detect technique for detecting weak targets in strong clutter backgrounds," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 47, no. 4, pp. 3062–3068, Oct. 2011.
[29] A. Moqiseh and M. M. Nayebi, "3-D Hough detector for surveillance radars," *IEICE Trans. Commun.*, vols. E93-B, no. 3, pp. 685–695, Mar. 2010.
[30] A. K. Jain, M. N. Murty, and P. J. Flynn, "Data clustering: A review," *ACM Comput. Surv.*, vol. 31, no. 3, pp. 264–323, Sep. 1999.
[31] Q. Liu, Z. Zhao, Y.-X. Li, and Y. Li, "Feature selection based on sensitivity analysis of fuzzy ISODATA," *Neurocomputing*, vol. 85, pp. 29–37, May 2012.
[32] P. Dong, Z. Jing, D. Gong, and B. Tang, "Maneuvering multi-target tracking based on variable structure multiple model GMCPHD filter," *Signal Process.*, vol. 141, no. 12, pp. 158–167, Dec. 2017.
[33] W. Si, L. Wang, and Z. Qu, "Multi-target tracking using an improved Gaussian mixture CPHD filter," *Sensors*, vol. 16, no. 11, p. 1964, Nov. 2016.
[34] D. Schuhmacher, B.-T. Vo, and B.-N. Vo, "A consistent metric for performance evaluation of multi-object filters," *IEEE Trans. Signal Process.*, vol. 56, no. 8, pp. 3447–3457, Aug. 2008.

ZHEPING YAN was born in Longyou, Zhejiang, China, in 1972. He received the B.S. degree in nuclear power equipment, the M.S. degree in special auxiliary device and system in naval architecture and ocean engineering, and the Ph.D. degree in control theory and control engineering from Harbin Engineering University, China, in 1994, 1997, and 2001, respectively. He was the Post-Doctoral Researcher in mechatronic engineering with the Harbin Institute of Technology, China, in 2004.

In 2015, he was a Visiting Scholar with the National University of Singapore, Singapore. In 2016, he was a Visiting Scholar with The Australian National University, Canberra, Australia. In 2017, he was a Visiting Scholar with the University of Guelph, Guelph, ON, Canada. In 1997, he joined Harbin Engineering University and was promoted to Professor in 2005. His current research interests include system design of unmanned underwater vehicle, identification of non-linear system, and multi-sensors data fusion and intelligent control.

Dr. Yan is the Deputy Secretary-General of the National Standardization Committee of Submersibles, and the Associate Supervisor of the State Key Laboratory of Underwater Robot Technology.

GENGSHI ZHANG was born in Harbin, China, in 1986. He received the B.S. degree in automation and the M.S. degree in control theory and control engineering from Harbin Engineering University, China, in 2010 and 2014 respectively, where he is currently pursuing the Ph.D. degree in control science and engineering. His research interests include intelligent planning and control of unmanned underwater vehicle.



JIAN XU was born in Wuhan, China, in 1980. He received the B.S. degree in heat engine, and the M.S. and Ph.D. degrees in control theory and control engineering from Harbin Engineering University, China, in 2002, 2004, and 2008, respectively. He was a Post-Doctoral Researcher in mechanical engineering with the Harbin Institute of Technology, China, in 2008.

In 2008, he joined Harbin Engineering University and was promoted to Professor in 2017. His current research interests include autonomous control of unmanned systems and motion and operation control of ships.

Dr. Xu is the member of the Chinese Association of Automation, and the Evaluation Expert of the National Nature Science Foundation of China.

TAO CHEN was born in Hebi, Henan, China, in 1983. He received the B.S. degree in nuclear engineering and technology, the M.S. and Ph.D. degrees in control theory and control engineering from Harbin Engineering University, China, in 2005, 2008, and 2011, respectively. He was a Post-Doctoral Researcher in naval architecture and ocean engineering with the Harbin Engineering University, China, in 2013.

In 2011, he joined Harbin Engineering University and was promoted to Associate Professor in 2017. His current research interests include planning and decision and control of unmanned systems.

XUE DU was born in Harbin, Heilongjiang, China, in 1987. She received the B.S. degree in measure control technology and instrument, the M.S. degree in control engineering, and the Ph.D. degree in navigation guidance and control from Harbin Engineering University, China, in 2010, 2012, and 2016, respectively.

In 2016, she joined Harbin Engineering University. Her current research interests include underwater image processing and autonomous control of unmanned underwater vehicle.

JUAN LI was born in 1976. She received the B.S. degree in nuclear engineering and technology, and the M.S. and Ph.D. degrees in control theory and control engineering from Harbin Engineering University, China, in 2000, 2003, and 2008, respectively.

In 2003, she joined Harbin Engineering University and was promoted to Associate Professor in 2008. Her current research interests include autonomous control of unmanned systems and motion and operation control of ships.