An Algorithm for Tracking Multiple Fish Based on Biological Water Quality Monitoring

XIAOQIANG ZHAO[1,2], SHENG YAN[1,2], AND QIANG GAO[3]
[1]School of Communication and Information, Xi'an University of Posts and Telecommunications, Xi'an 710121, China
[2]Shaanxi Key Laboratory of Information Communication Network and Security, Xi'an University of Posts and Telecommunications, Xi'an 710121, China
[3]College of Mechanical and Electronic Engineering, Northwest Agriculture and Forestry University, Yangling 712100, China

Corresponding author: Qiang Gao (hifinder@163.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 61473235, in part by the Xianyang Science and Technology Research Project under Grant 2017k01-25-7, in part by the Shaanxi Science and Technology Achievements Promotion Project under Grant 2018CG-007, and in part by the Industrialization Cultivation Project of the Shaanxi Province Education Department under Grant 18JC029.

ABSTRACT Abnormal water quality will increase the occlusion rate among fish schools, which causes difficulties in fish detection and tracking. In order to solve this problem, a multiple fish tracking algorithm for red snapper is proposed in this paper. In the detection stage, we use the Otsu adaptive segmentation algorithm to extract fish targets based on the background subtraction method, following which the fish tracking feature parameters can be obtained based on the fish geometric features. In the tracking stage, the Kalman filter is employed to first estimate the motion state, and then the cost function is constructed from the position of the fish body, target area, and the direction information. Finally, fish school tracking is realized by the interframe relationship matrix. We applied several tracking methods with various schemes to experimental videos of swimming fish schools in different environments. The experimental results show that the proposed tracking algorithm exhibits improved performance and robustness.

INDEX TERMS Data association, feature detection, multi-object tracking, water quality monitoring.

I. INTRODUCTION

The stress response of aquatic organisms to pollutants can be used to develop biological early warning systems, and behavioral changes of aquatic organisms reflect the aquatic environments and ecosystems [1]–[4]. Recently, with the considerable improvements in computer vision and image processing technology, the visual monitoring of fish school behavior changes has become an important topic in biological water quality monitoring [5]–[9]. In order to achieve this goal, it is necessary to first obtain the trajectory data for each fish target and then perform various statistical processes to realize biological water quality monitoring and early warning [10]. However, the shape of the fish target changes during swimming owing to its non-rigid attributes. In addition, environmental influences, detection errors, and the mutual occlusion of fish targets also result in many difficulties for fish school tracking in video images [11]. Therefore, fish school tracking is the key step in biowater quality monitoring.

Several fish tracking algorithms exist that are based on the detection of single fish [3], [6]; however, the tracking of fish schools is much more relevant. The existing fish school tracking algorithms are mainly based on feature points, which typically treat a single fish as a single point. At the same time, multiple fish tracking is achieved by feature point coordinates. For example, Miller and Gerlai [12] developed software to track multiple fish whereby the user first clicks on the body of each target manually, which then quantifies the characteristic parameters of the fish target. Unfortunately, this method is highly labor intensive when the fish group is large. On this basis, an automatic fish tracking system was proposed [13], which verified the feasibility of using computer vision technology to achieve the automatic tracking of multiple fish. However, the accuracy of this algorithm needs further improvement. Delcourt *et al.* [14] achieved the fish school detection and tracking using visual implant elastomer tags, but these tags may potentially affect the social behavior of the fish target, which in turn has a certain impact on the tracking results.

In recent years, more sophisticated multiple fish tracking algorithms have been developed to improve tracking performance [15]–[18]. In particular, the integration of motion direction has improved the fish school tracking performance. Qian *et al.* [19] proposed an effective fish school tracking algorithm based on fish head detection, which uses the fish body shape and gray features combined with extreme value detection and an ellipse fitting method to obtain the fish motion direction. Xia *et al.* [20] presented a multiple fish tracking algorithm based on fish swimming posture, which was used to identify the fish tail and head points via a grayscale histogram. The fish tracking was achieved by the feature points using a combination of swimming posture and the nearest neighbor algorithm. However, all the above mentioned algorithms have some shortcomings related to the fish detection of occlusion and motion direction.

Thus, a new multiple fish tracking algorithm is proposed in this paper. The features of the algorithm are as follows: first, a new method for detecting fish movement direction is presented based on the geometric characteristics of fish in the detection stage. Second, the data association between consecutive frames is realized by combining the centroid coordinates, target area, and motion direction of the fish target. Third, the application of a Kalman filtering algorithm effectively improves the performance of multiple fish tracking in the data association stage.

The rest of this paper is constructed as follows. The fish target information acquisition platform is introduced in section 2. Section 3 proposes the multiple fish tracking algorithm, including fish detection and fish tracking. Section 4 presents the experimental results and evaluates the performance of the fish detection and tracking method. The conclusions are then discussed in Section 5.

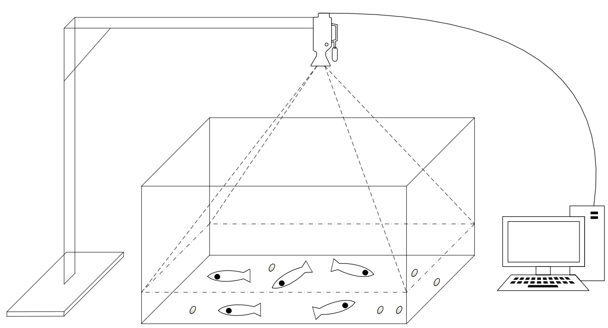

FIGURE 1. Experiment setup of the proposed fish school tracking method.

II. MATERIALS AND SETUP

Our experimental device platform, shown in Fig. 1, is composed of three parts: a charge-coupled device camera, experimental fish tank, and a computer. During the video acquisition process, the experimental device is located in a fixed illumination and environmentally stable isolation space in order to reduce the experimental errors caused by illumination changes and human interference. The video acquisition uses the Hikvision MV-CA030-10GC industrial camera, which is located directly above the experimental tank to ensure that the swimming area of the fish school is within the surveillance range of the camera. The resolution of the machine is 1360 pixels × 650 pixels in order to achieve clear detection. The size of the experimental tank is 40 cm × 20 cm × 20 cm, the top view is 40 cm × 20 cm, and the water depth is 3 cm. In order to prevent the fish body swimming in a vertical direction, the fish body can be approximated as swimming on a two-dimensional plane. Additionally, some colored pebbles were added to the experimental tank to simulate the interference problems that occur in the real environment. The proposed tracking algorithm was implemented on a personal computer (Intel(R) Core(TM)-i5-4210U CPU @ 1.70 GHz 2.39 GHz) using Matlab R2017b software. Red snapper was used as a water quality indicator organism and was about 1.5-3 cm in length. Due to its sensitivity to water quality, it has been widely used in biological water quality monitoring.

III. FISH SCHOOL TRACKING METHOD

The proposed algorithm is mainly composed of two parts: fish target detection and target tracking. The algorithm flowchart is shown in Fig. 2.

A. FISH DETECTION

1) MOTION REGION SEGMENTATION

In the laboratory environment, the illumination is fixed and the acquisition environment is stable. Therefore, the background subtraction method is used to obtain the moving target. The experimental video only contains the foreground target of the fish and the background that closes to static, and the moving target stays in a certain area for a short time. The background model estimation [21], [22] is performed using the Median filtering algorithm in the time domain. Assuming that the video input image is $I(i, j)$, the background image $B(i, j)$ can be obtained by statistical calculation through the m frame input images:

$$B(i, j) = Median(I_1(i, j), I_2(i, j), \ldots, I_m(i, j)). \tag{1}$$

The difference between the background image and the input image is the difference image $D(i, j)$, which is defined as

$$D(i, j) = |B(i, j) - D(i, j)|. \tag{2}$$

The image threshold segmentation method is used to convert the difference image into a binary image. The position $D(i, j)$ is determined as the foreground pixel point, whose value is greater than the threshold value. The binary image is defined as

$$R(i, j) = \begin{cases} 1, & if\ D(i, j) \geq T_g \\ 0, & if\ D(i, j) < T_g. \end{cases} \tag{3}$$

Then, the Otsu adaptive segmentation algorithm [23], [24] is used to determine the image segmentation threshold T_g in order to further extract the features of the fish target. First, the morphological operation and the hole filling process are performed on the binary image to smooth the target contour of the fish body. Second, the interference noise of the small block is filtered out by the threshold T_l to reduce the detection error.

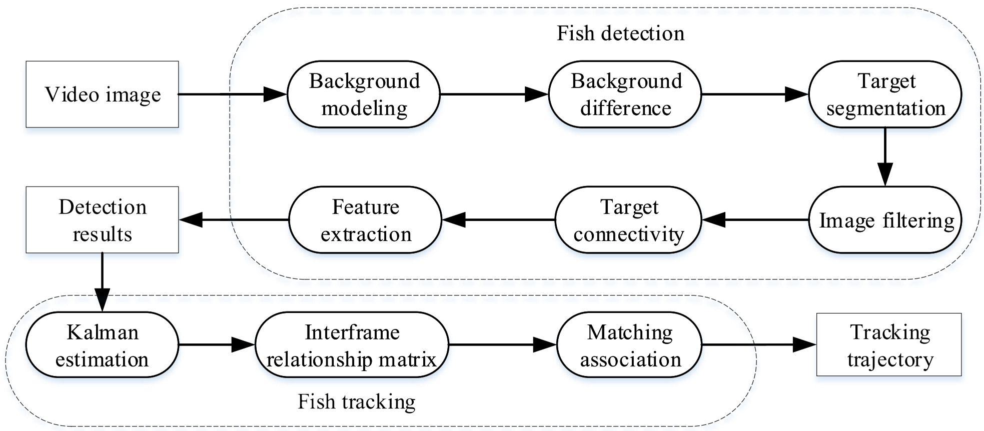

FIGURE 2. **A flowchart of the fish school tracking algorithm.**

2) CENTROID AND AREA DETECTION

After obtaining the filtered image, the connected region in the image represents the corresponding region of the fish body target, and the connected region marking method is then used to obtain the regional information of the fish moving target. Then, the external rectangular box and the area of the connected region are obtained according to the connected region information. The area of the connected region is equivalent to the target area of the fish body. The center position of the connected region is equivalent to the centroid position of the fish moving target [25]. The order moment in the target region of each fish body is as shown in Eq. (4).

$$\begin{cases} M_{00} = \sum_{i=0}^{M} \sum_{j=0}^{N} R(i,j) \\ M_{10} = \sum_{i=0}^{M} \sum_{j=0}^{N} i \times R(i,j) \\ M_{01} = \sum_{i=0}^{M} \sum_{j=0}^{N} j \times R(i,j), \end{cases} \tag{4}$$

where M_{00} represents the zeroth order moment of the connected region; M_{10}, M_{01} represent the first-order moment; and R is the filtered image. Then, the centroid position coordinates of the target area of the fish body are as follows:

$$\begin{cases} x_c = \dfrac{M_{10}}{M_{00}} \\ y_c = \dfrac{M_{01}}{M_{00}}. \end{cases} \tag{5}$$

3) MOTION DIRECTION DETECTION

As shown in Fig. 3(a), the direction of the movement of the red snapper is measured by the centroid and head of the individual target. The direction of motion is θ, whose range is $[0, 2\pi]$. According to the cosine theorem, the motion direction is

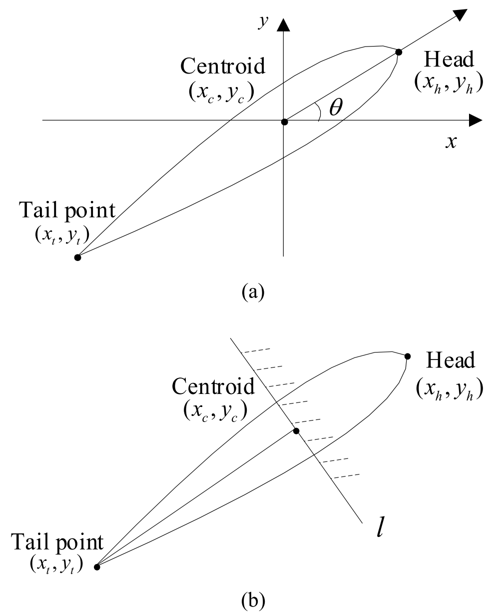

FIGURE 3. **An illustration of the fish body. (a) An illustration of the motion direction of the fish target, (b) an illustration of the calculation of the head point.**

$$\theta = \begin{cases} arctan\dfrac{y_h - y_c}{x_h - x_c}, & x_h - x_c > 0 \\ \pi + arctan\dfrac{y_h - y_c}{x_h - x_c}, & x_h - x_c < 0 \\ \dfrac{\pi}{2}, & x_h = x_c, y_h > y_c \\ \dfrac{3\pi}{2}, & x_h = x_c, y_h < y_c. \end{cases} \tag{6}$$

Therefore, the motion direction of the fish depends on the acquisition of the centroid and the head, in which the centroid position has been acquired. It can be seen from the experimental observation that the target of the fish body has a conical structure and the tail point is farthest from the centroid in all the contour points of the fish body. Thus, using the traversal method, the coordinates of the tail point from the contour points of the fish body can be obtained.

Suppose the set of fish body edge contour points is

$$M = \{(x_1, y_1), (x_2, y_2), \cdots, (x_n, y_n)\}. \tag{7}$$

Then the distance set between the contour point and the centroid is

$$N = \{dis_{c1}, dis_{c2}, \cdots, dis_{cn}\}, \tag{8}$$

where dis_{ci} is the Euclidean distance between the i[th] coordinate point in set M and the centroid, and the index value t corresponding to the maximum value in the set N is the index value of the tail point, then

$$t = \underset{i\in[1,n]}{arg} \; max(dis_{ci}). \tag{9}$$

Thus, the coordinates of the tail point are

$$(x_t, y_t) = M(x_t, y_t). \tag{10}$$

Based on the regulation of fish swimming characteristics, connect the centroid and the tail point and draw a line passing (x_c, y_c) and perpendicular to the line of the centroid and the tail point. As shown in Fig. 3(b), the side of the dotted line in l is the distribution range of the fish head, and the point farthest from the centroid in the outline of the fish body edge of this range is the head position.

The vertical line l is obtained from the centroid and tail point coordinates:

$$y_l = (-\frac{x_c - x_t}{y_c - y_t})(x_l - x_c) + y_c. \tag{11}$$

By judging the tail point and the vertical line, the contour point set on the side of the dotted line is

$$P = \{(x_1, y_1), (x_2, y_2), \cdots, (x_z, y_z)\}. \tag{12}$$

Thus, the head coordinates of the fish body satisfies:

$$h = \underset{h\in[1,z]}{arg} \; max(\sqrt{(x_h - x_c)^2 + (y_h - y_c)^2}). \tag{13}$$

B. FISH TRACKING

1) MOVING TARGET STATE ESTIMATION

As for the optimal autoregressive data processing algorithm, Kalman filtering is one of the most effective methods for estimating motion state [26], [27]. As the time interval between consecutive frames is small in the fish school swimming video, the target can be considered to be a uniform linear motion between adjacent frames, and thus the system can be approximated as a linear dynamic model. It is described by a state equation and an observation equation, which are defined as

$$x_k = A_k x_{k-1} + \omega_k, \tag{14}$$

$$z_k = C_k x_k + \nu_k, \tag{15}$$

where, x_k, z_k are the state vector and observation vector, respectively; A_k, C_k are the target state transition matrix and the observation matrix, respectively, and ω_k, ν_k are the system state noise and observed variable noise, respectively.

In the time update part, the Kalman filter theory can be used to obtain the prediction equations:

$$\hat{x}'_k = A_k \hat{x}_{k-1}, \tag{16}$$

$$P'_k = A_k P_{k-1} A_k^T + Q_{k-1}, \tag{17}$$

where $\hat{x}'_k$ is the prior state estimated value; P'_k is the updated error covariance matrix; Q_{k-1} is the covariance matrix of system state noise ω_k; $\hat{x}_{k-1}$ is the state estimate vector in $k-1$; and P_{k-1} is the error covariance matrix in $k-1$.

The current state estimation can be obtained by the state update equations based on data association, which is defined as:

$$H_k = P'_k C_k^T (C_k P'_k C_k^T + R_k)^{-1}, \tag{18}$$

$$\hat{x}_k = \hat{x}'_{k-1} + H_k(z_k - C_k A_k \hat{x}'_{k-1}), \tag{19}$$

$$P_k = (I - H_k C_k) P'_k, \tag{20}$$

where H_k is the Kalman gain value at time k; and R_k is the covariance matrix of the system observation noise ν_k. When the background is disturbed or the fish target is occluded, the target state variable estimation can be effectively realized by the prediction equations and the updated equations due to the adaptability of the Kalman algorithm.

2) COST FUNCTION CALCULATION

The data association between the state vector and the observation vector is a key step in the tracking of the fish school. In order to improve the association accuracy, the feature matching method is used for data association [28]. A certain relationship exists between the motion state characteristics of the fish target in adjacent frames; namely, that the centroid location, target area, and movement direction of the same fish vary little, and the parameters between the different targets vary greatly. Therefore, the cost function is introduced for feature matching. Define the previous frame tracking target list as $T = \{O_{k-1,1}, O_{k-1,2}, ..., O_{k-1,n}\}$, the current frame detection target list is $D = \{O_{k,1}, O_{k,2}, ..., O_{k,m}\}$, where $O_{k,m}$ represents the feature information of the m[th] moving target in the k[th] frame. Then the cost function between the target i of the previous frame and the target j of the current frame is defined as

$$cf_{ij} = \omega_1(\frac{pc_{ij}}{pc_{max}}) + \omega_2(\frac{dc_{ij}}{dc_{max}}) + (1 - \omega_1 - \omega_2)(\frac{ac_{ij}}{ac_{max}}), \tag{21}$$

where, pc_{max}, dc_{max}, ac_{max} represent the maximum values of positional change, direction change, and area change between consecutive frames, respectively. pc_{ij}, dc_{ij}, ac_{ij} represent the value of position change, direction change, and area change between the target i of the previous frame and the target j of the current frame, and ω_1, ω_2 represent the weights corresponding to the centroid and motion direction.

TABLE 1. The results of the parameter setting for fish detection and tracking.

Detection parameter			Tracking parameter			
m	T_g	T_l	ω_1	ω_2	d_0	N
100	/	800	0.60	0.20	300	5

TABLE 2. The evaluation results of fish detection.

Video	P(%)	R(%)	Number of occlusions	OR(%)	ODR(%)
D1(pH = 6)	99.41	99.26	395	3.95	92.41
D2(pH = 7)	99.75	99.44	205	2.05	92.68
D3(pH = 8)	98.23	98.85	325	3.25	91.38

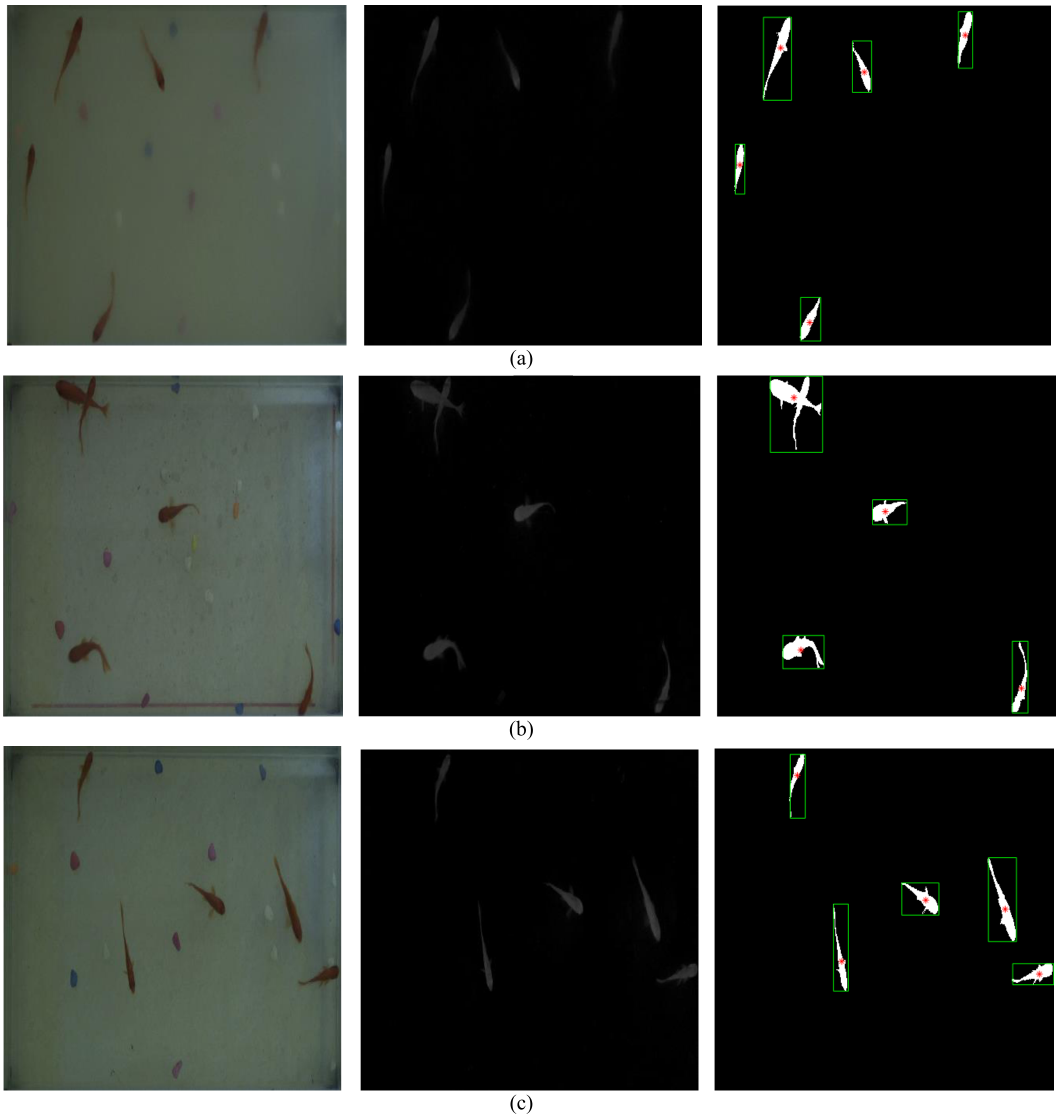

FIGURE 4. Fish detection effect on the 1000th frame in different pH environments. On the left side of the figure are the original images, in the middle of the figure are the difference images, and on the right side are the tag images. The experimental environment is (a) pH = 6, (b) pH = 7, and (c) pH = 8.

3) TARGET MATCHING ASSOCIATION

It is a very important way to achieve data association between multiple targets based on the nearest neighbor algorithm [29]. After obtaining the cost function values between the targets of adjacent frames, the fish body targets are matched by the interframe relationship matrix. The interframe relationship matrix between the n tracking targets in the previous frame and the m detection targets in the current frame is

TABLE 3. Comparison of the several tracking methods under different schemes.

Number	State estimation	Cost function calculation	Data association
Method1	None	Proposed	Proposed
Method2	Proposed	Only Centroid	Proposed
Method3	Proposed	Proposed	Proposed

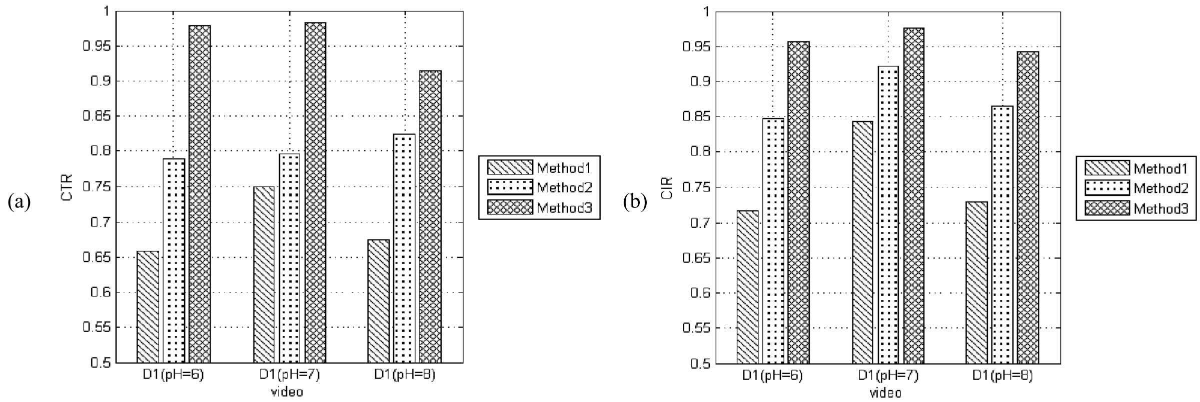

FIGURE 5. Tracking performance of the compared methods on two evaluation metrics, (a) CTR, (b) CIR.

defined as

$$P = \begin{bmatrix} cf_{11} & cf_{12} & \cdots & cf_{1m} \\ cf_{21} & cf_{22} & \cdots & cf_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ cf_{n1} & cf_{n2} & \cdots & cf_{nm} \end{bmatrix}. \tag{22}$$

To reduce the number of matching associations and improve the tracking performance, the values in the inter-frame relationship matrix can be simplified as follows:

$$cf_{ij} = \begin{cases} cf_{ij}, & if\ D(i,j) < d_0 \\ +\infty, & otherwise, \end{cases} \tag{23}$$

where, d_0 is the maximum overlap distance between fish targets, and $D(i,j)$ is the Euclidean distance between the target i in the previous frame and the target j in the current frame. Thus, the matching association can only be performed when the Euclidean distance between two targets in consecutive frames is less than a specified threshold.

Therefore, the following NP-hard optimization problem is established:

$$min \sum_{i=1}^{n}\sum_{j=1}^{m} cf_{ij}x_{ij}$$

$$s.t. \begin{cases} \sum_{i=1}^{n} x_{ij} = 1 & (j = 1, \cdots, m) \\ \sum_{j=1}^{m} x_{ij} = 1 & (i = 1, \cdots, n) \\ x_{ij} = 1\ or\ 0 & (i = 1, \cdots n, j = 1, \cdots m), \end{cases} \tag{24}$$

where, $x_{ij} = 1$ indicates that the target i is associated with the target j, and $x_{ij} = 0$ indicates that the target i is not associated with the target j. The optimal model is solved by the Hungarian algorithm [30], and thus the optimal matching scheme between consecutive frames can be obtained.

IV. EXPERIMENTAL RESULTS

To evaluate the performance of the proposed tracking algorithm, the experiments were performed under the experimental setup shown in Fig.1. The pH of the water environment was changed by oxalic acid and sodium hydroxide to simulate the water quality under different environments. Three groups of red snapper experimental videos in different environments were selected as test data: D1 (pH = 6), D2 (pH = 7), and D3 (pH = 8). The video for each group contained 2000 frames and five fish, and the video frame rate was 25 fps.

A. PARAMETER SETTINGS

There are some parameters that need to be set up for optimal detection and tracking performance. The parameters of m, T_g, T_l are set in the detection part. For the parameter m, which reflects the true degree of the background image, the larger the value is, the closer the background image is to the real background. The parameter T_g is used to segment the foreground object and the background image, which is obtained by the Otsu algorithm. For the threshold T_l, its value is determined by the size of the noise point in the image, and the larger the value is, the more likely the fish target will be filtered out. Otherwise, more noise points will be present in the binary image. The parameters of $\omega_1, \omega_2, d_0, N$ are set in the tracking part. The parameters ω_1, ω_2 are determined by the position, area, and direction change rate of the same fish target between consecutive frames. The parameter d_0 is determined by the maximum occlusion distance among the fish targets in the image. The longer the occlusion distance,

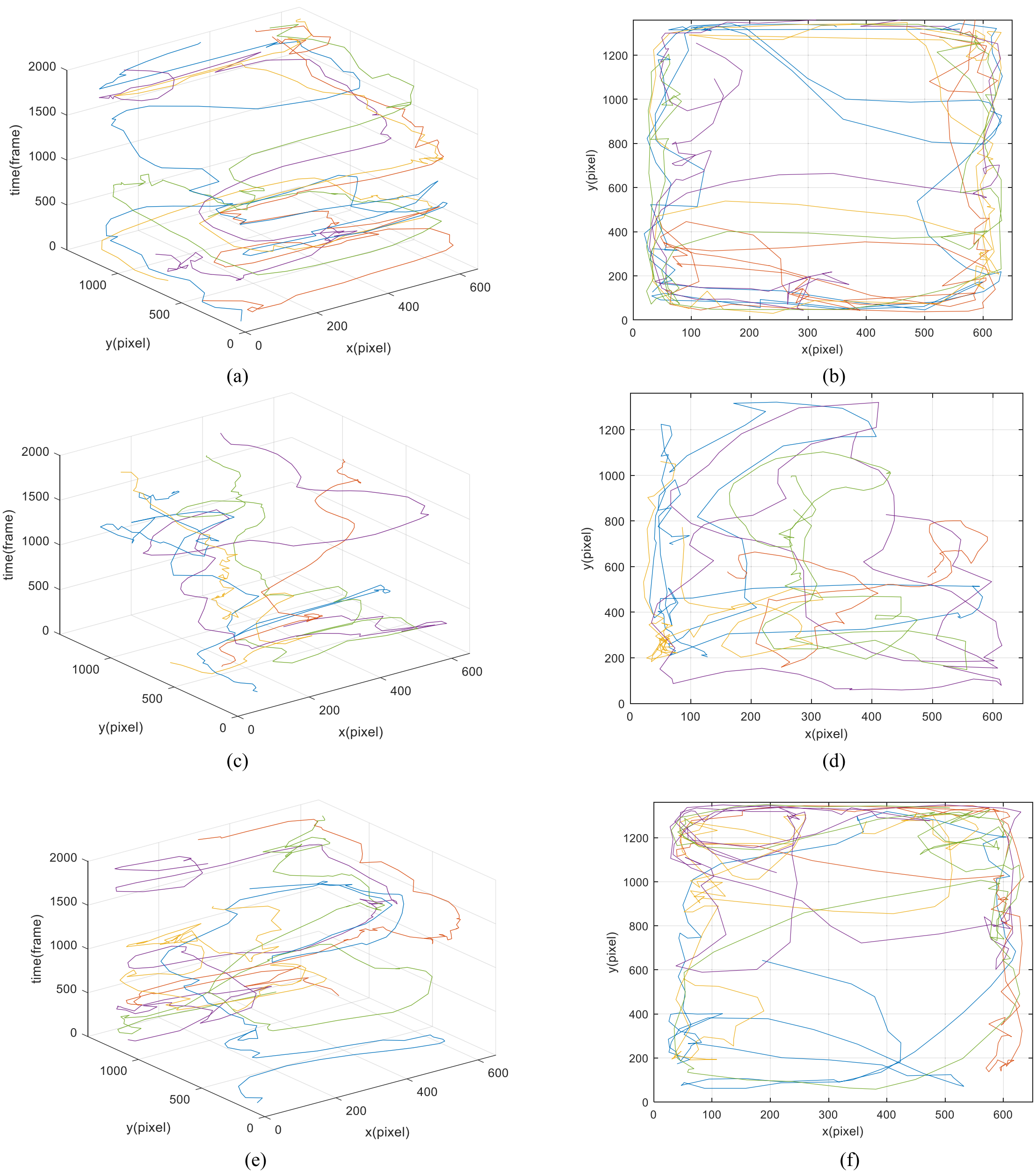

FIGURE 6. **Fish school tracking trajectory in different environments. (a), (c), and (e) are three-dimensional tracking trajectories of fish schools under pH = 6, 7, and 8, (b), (d), and (f) are two-dimensional tracking trajectories of fish schools under pH = 6, 7, and 8. The lines of different colors correspond to the swimming trajectories of different fish targets.**

the larger the value is. The parameter N represents the number of fish in the video.

In order to set the best parameters, we selected 500 frames from three test sets as training samples. Analyzing and comparing the detection and tracking results obtained under different parameters, the selected parameter results are shown in Table 1.

B. EVALUATION OF DETECTION PERFORMANCE

In the process of fish target detection, it is necessary to quantify the performance of the proposed method using the accuracy rate and recall rate to measure the fish school detection effect [31]. The definition is as follows:

$$P = \frac{TP}{TP + FP}, \tag{25}$$

$$R = \frac{TP}{TP + FN}, \tag{26}$$

where TP is the total number of fish targets that are correctly identified in all frames, FP is the total number of fish targets that are error-detected and FN is the total number of mis-detected fish targets.

In addition, in order to better evaluate the detection performance in the case of the occlusion of fish, two evaluation criteria of occlusion rate and occlusion detection rate [32] are proposed:

$$OR = \frac{\text{total number of occlusions}}{\text{total number of targets}}, \tag{27}$$

$$ODR = \frac{\text{successful number of occlusion detection}}{\text{total number of occlusions}}. \tag{28}$$

The detection results are shown in Table 2. It can be seen that compared with the normal water quality environment (pH = 7), the number of occlusions between fish targets in the abnormal water environment is significantly increased, resulting in a decrease in detection performance. The erratic and highly rapid movement of the fish in the acid-base environment is also an important factor influencing the decline in detection performance. Despite this, the accuracy of the three experimental videos was greater than 98.23%, and the recall rate was greater than 98.85%, which verified the effectiveness of the detection method. In addition, the occlusion detection rate indicates that the occlusion between the fish targets causes a decrease in detection performance, but the proposed method can effectively perform target detection in the case of occlusion. Fig. 4 shows the effect of the 1000$^{\text{th}}$ frame image detection in three groups of experimental videos. It can effectively detect fish targets in an acid-base environment and verify the applicability of the detection method for water quality monitoring experiments.

C. EVALUATION OF TRACKING PERFORMANCE

In the evaluation process of tracking performance based on the parameters used in [33], the correct tracking rate and correct recognition rate of the three sets of experimental videos were evaluated based on manual labeling, which is defined as

$$CTR = \frac{\sum(\text{number of correct frames of a single fish})}{\text{number of fish} \times \text{number of frames}}, \tag{29}$$

$$CIR = \frac{\text{times that all fish get correct identity after occlusion}}{\text{number of occlusion events}}, \tag{30}$$

where CTR represents the proportion of individual fish being tracked correctly, with a large value indicating better tracking performance. CIR measures the probability of the correct identification of all fish following an occlusion event, the smaller the value is, the lower the accuracy is.

The fish school tracking algorithm proposed in this paper contains three parts: moving target state estimation, cost function calculation and data association. Three methods were analyzed and compared using different schemes to better evaluate the proposed method, and the compared methods are shown in Table 3.

The comparison results of the three methods are shown in Fig. 5. As can be seen from the comparison of moving target state estimation, CTR and CIR using the state estimation method perform much better than the use of feature matching alone. By adding state estimation into feature matching, the matching accuracy was improved, and thus the tracking performance was improved significantly. As evident in the cost function calculation comparison, the tracking algorithm in this paper has the advantage of introducing the feature of the target area and motion direction, a better ability to match the target in the occlusion case, and thus a higher CTR and CIR value. When the water quality is abnormal and the fish body overlap increases, the introduction of new features is required to improve the tracking performance. The tracking trajectory results of the proposed method in different test videos are shown in Fig. 6.

V. CONCLUSION

A fish school tracking algorithm based on biological water quality monitoring is proposed in this paper, in which the data association is critical between the state variables and observation variable. Therefore, on the basis of the data association of the distance-based nearest neighbor algorithm, the data association is realized by introducing a fish target area and the moving direction feature, which effectively improve the tracking performance of the fish school under occlusion. The experimental results show that the proposed tracking algorithm exhibits improved tracking performance. In future work, the movement regulation of fish schools under different water quality environments will be further evaluated, and the relationship between fish movement behavior and water quality will be determined.

REFERENCES

[1] M.-J. Bae and Y.-S. Park, "Biological early warning system based on the responses of aquatic organisms to disturbances: A review," *Sci. Total Environ.*, vols. 466–467, no. 1, pp. 635–649, Jan. 2014.
[2] C. Xia, L. Fu, Z. Liu, H. Liu, L. Chen, and Y. Liu, "Aquatic toxic analysis by monitoring fish behavior using computer vision: A recent progress," *Int. J. Toxicol.*, vol. 2018, Apr. 2018, Art. no. 2591924.
[3] J. Mao, G. Xiao, W. Sheng, Z. Qu, and Y. Liu, "Research on realizing the 3D occlusion tracking location method of fish's school target," *Neurocomputing*, vol. 214, pp. 61–79, Nov. 2016.
[4] Z. Ren and Z. Wang, "Differences in the behavior characteristics between *Daphnia magna* and Japanese madaka in an on-line biomonitoring system," *J. Environ. Sci.*, vol. 22, no. 5, pp. 703–708, Aug. 2010.
[5] S. Nema, W. Hasan, A. Bhargava, and Y. Bhargava, "A novel method for automated tracking and quantification of adult zebrafish behaviour during anxiety," *J. Neurosci. Meth.*, vol. 271, pp. 65–75, Sep. 2016.
[6] H. Ma, T.-F. Tsai, and C.-C. Liu, "Real-time monitoring of water quality using temporal trajectory of live fish," *Expert Syst. Appl.*, vol. 37, no. 7, pp. 5158–5171, Jul. 2010.
[7] I. Kuklina, A. Kouba, and P. Kozák, "Real-time monitoring of water quality using fish and crayfish as bio-indicators: A review," *Environ. Monit. Assessment*, vol. 185, no. 6, pp. 5043–5053, Oct. 2013.
[8] V. M. Papadakis, I. E. Papadakis, F. Lamprianidou, A. Glaropoulos, and M. Kentouri, "A computer-vision system and methodology for the analysis of fish behavior," *Aquaculture Eng.*, vol. 46, no. 1, pp. 53–59, Jan. 2012.
[9] X. Zhao, W. Cheng, and Y. Yu, "Study on movement of the fish body related water pollution based on image processing," *Biotechnology*, vol. 7, no. 11, pp. 435–439, 2013.
[10] V. Lecheval, L. Jiang, P. Tichit, C. Sire, C. K. Hemelrijk, and G. Theraulaz, "Social conformity and propagation of information in collective U-turns of fish schools," *Proc. Roy. Soc. B, Biol. Sci.*, vol. 285, no. 1877, p. 20180251, Apr. 2018.
[11] R. Dolado, E. Gimeno, F. S. Beltran, V. Quera, and V. F. Pertusa, "A method for resolving occlusions when multitracking individuals in a shoal," *Behav. Res. Methods*, vol. 47, no. 4, pp. 1032–1043, Oct. 2014.
[12] N. Miller and R. Gerlai, "Quantification of shoaling behaviour in zebrafish (*Danio rerio*)," *Behav. Brain Res.*, vol. 184, no. 2, pp. 157–166, Dec. 2007.

[13] J. Green *et al.*, "Automated high-throughput neurophenotyping of zebrafish social behavior," *J. Neurosci, Methods*, vol. 210, no. 2, pp. 266–271, Sep. 2012.

[14] J. Delcourt *et al.*, "Video tracking in the extreme: A new possibility for tracking nocturnal underwater transparent animals with fluorescent elastomer tags," *Behav. Res. Methods*, vol. 43, no. 2, pp. 590–600, Jun. 2011.

[15] A. Pérez-Escudero, J. Vicente-Page, R. C. Hinz, S. Arganda, and G. G. de Polavieja, "idTracker: Tracking individuals in a group by automatic identification of unmarked animals," *Nature Methods*, vol. 11, no. 7, pp. 743–752, Jun. 2014.

[16] Z. Xu and X. E. Cheng, "Zebrafish tracking using convolutional neural networks," *Sci. Rep.*, vol. 7, no. 42815, Feb. 2017.

[17] A. Rodriguez, H. Zhang, J. Klaminder, T. Brodin, P. L. Andersson, and M. Andersson, "*ToxTrac*: A fast and robust software for tracking organisms," *Methods Ecol. Evol.*, vol. 9, no. 3, pp. 460–464, Aug. 2017.

[18] J. Delcourt, M. Denoël, M. Ylieff, and P. Poncin, "Video multitracking of fish behaviour: A synthesis and future perspectives," *Fish Fisheries*, vol. 14, no. 2, pp. 186–204, Jun. 2013.

[19] Z.-M. Qian, S. H. Wang, X. E. Cheng, and Y. Q. Chen, "An effective and robust method for tracking multiple fish in video image based on fish head detection," *BMC Bioinform.*, vol. 17, no. 251, pp. 1–11, Jun. 2016.

[20] C. Xia, T.-S. Chon, Y. Liu, J. Chi, and J. Lee, "Posture tracking of multiple individual fish for behavioral monitoring with visual sensors," *Ecol. Informat.*, vol. 36, no. 1, pp. 190–198, Nov. 2016.

[21] X. Ye, J. Yang, X. Sun, K. Li, C. Hou, and Y. Wang, "Foreground–background separation from video clips via motion-assisted matrix restoration," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 25, no. 11, pp. 1721–1734, Nov. 2015.

[22] T. Bouwmans, "Traditional and recent approaches in background modeling for foreground detection: An overview," *Comput. Sci. Rev.*, vols. 11–12, pp. 31–36, May 2014.

[23] Q. Chen, L. Zhao, J. Lu, G. Kuang, N. Wang, and Y. Jiang, "Modified two-dimensional Otsu image segmentation algorithm and fast realisation," *IET Image Process.*, vol. 6, no. 4, pp. 426–433, Jun. 2012.

[24] H. Wang and Y. Dong, "An improved image segmentation algorithm based on Otsu method," *Proc. SPIE*, vol. 6625, pp. 1–8, 2008.

[25] X. Zhao, H. Peng, and Q. Gao, "A video image detection method for fish motion characteristics," *J. Xian Univ. Posts Telecommun.*, vol. 22, no. 2, pp. 38–43, Feb. 2017.

[26] G. Evensen, "The ensemble Kalman filter: Theoretical formulation and practical implementation," *Ocean Dyn.*, vol. 53, no. 4, pp. 343–367, Nov. 2003.

[27] F. Auger, M. Hilairet, J. M. Guerrero, E. Monmasson, T. Orlowska-Kowalska, and S. Katsura, "Industrial applications of the Kalman filter: A review," *IEEE Trans. Ind. Electron.*, vol. 60, no. 12, pp. 5458–5471, Dec. 2013.

[28] Y. T. Tesfaye, E. Zemene, M. Pelillo, and A. Prati, "Multi-object tracking using dominant sets," *IET Comput. Vis.*, vol. 10, no. 4, pp. 289–297, Jul. 2016.

[29] J. Delcourt, C. Becco, N. Vandewalle, and P. Poncin, "A video multitracking system for quantification of individual behavior in a large fish shoal: Advantages and limits," *Behav. Res. Methods.*, vol. 41, no. 1, pp. 228–235, Feb. 2009.

[30] R. Ardekani *et al.*, "Three-dimensional tracking and behaviour monitoring of multiple fruit flies," *J. Roy. Soc. Interface*, vol. 10, no. 78, pp. 4409–4420, Jan. 2013.

[31] K. Muhammad, J. Ahmad, I. Mehmood, S. Rho, and S. W. Baik, "Convolutional neural networks based fire detection in surveillance videos," *IEEE Access*, vol. 6, pp. 18174–18183, 2018.

[32] Z. M. Qian, X. E. Cheng, and Y. Q. Chen, "Automatically detect and track multiple fish swimming in shallow water with frequent occlusion," *PLoS ONE*, vol. 9, no. 9, p. e106506, Sep. 2014.

[33] S. H. Wang, X. E. Cheng, Z.-M. Qian, Y. Liu, and Y. Q. Chen, "Automated planar tracking the waving bodies of multiple zebrafish swimming in shallow water," *PLoS ONE*, vol. 11, no. 4, p. e0154714, Apr. 2016.

XIAOQIANG ZHAO received the Ph.D. degree from the Xi'an University of Technology, Xi'an, China, in 2015. He is currently a Professor with the Xi'an University of Posts and Telecommunications, Xi'an. His current research interests include the Internet of Things technology and environmental engineering.

SHENG YAN received the B.S. degree in mathematics and application mathematics from Xianyang Normal University, Shaanxi, China, in 2015. He is currently pursuing the master's degree in electronic and communication engineering with the Xi'an University of Posts and Telecommunications, Xi'an, China. His research interests include computer vision and machine learning.

QIANG GAO received the M.S. degree from the Xi'an University of Posts and Telecommunications, Xi'an, China, in 2017. He is currently pursuing the Ph.D. degree with the College of Mechanical and Electronic Engineering, Northwest Agriculture and Forestry University. His research interests include digital signal processing, computer vision, and automatic control.