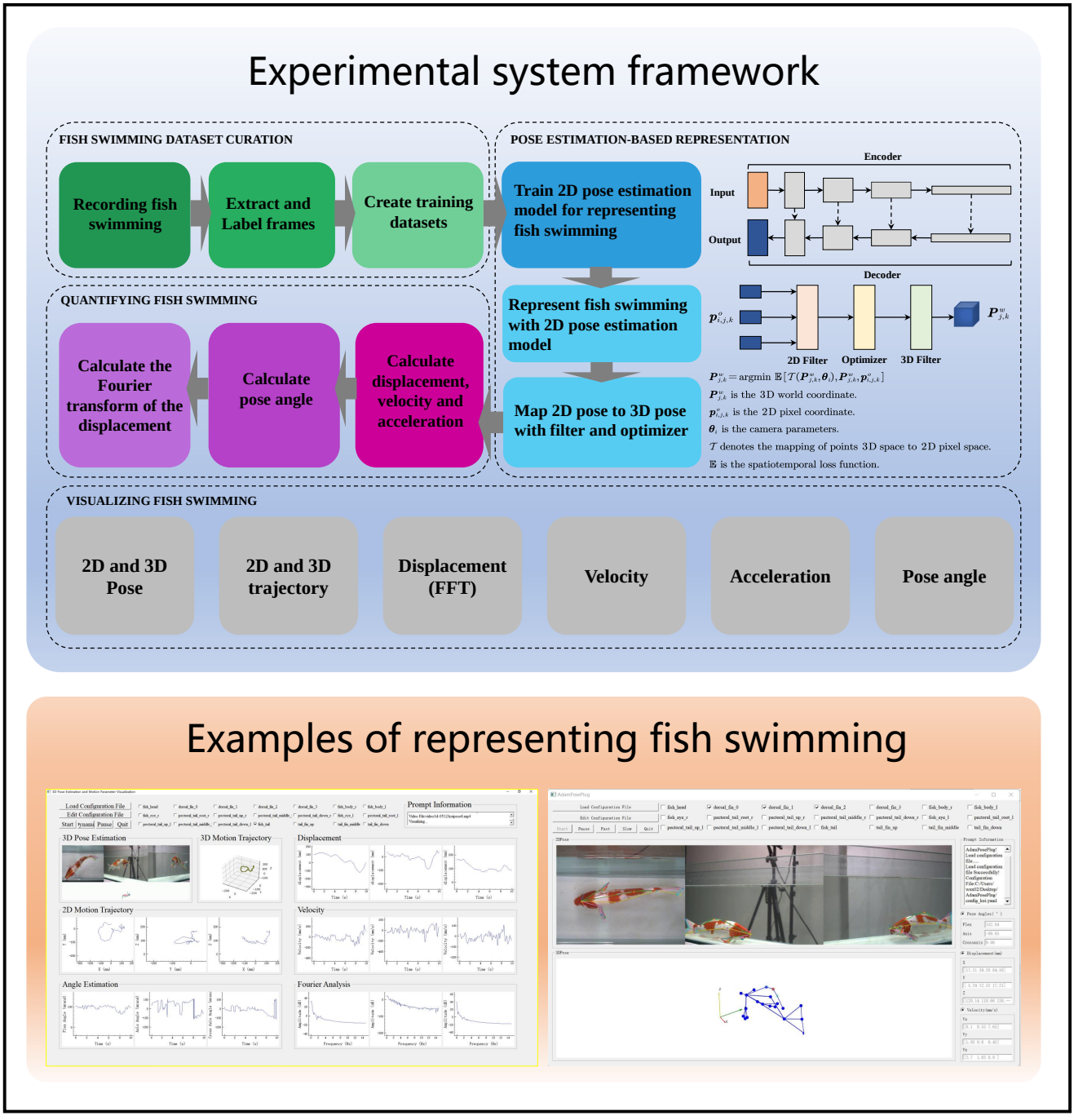
Graphical Abstract

**Pose estimation-based experimental system for analyzing fish swimming**

Xin Wu,Jipeng Huang,Yonghui Wang,Lianming Wang



# Highlights

## Pose estimation-based experimental system for analyzing fish swimming

Xin Wu,Jipeng Huang,Yonghui Wang,Lianming Wang

- A novel pose estimation-based experimental system for representing fish swimming is designed.
- A software that integrates experiments on fish swimming studying is designed.
- An insight into the synergistic of the fins of koi in four swimming states is proposed.

# Pose estimation-based experimental system for analyzing fish swimming⋆

Xin Wu[a], Jipeng Huang[a], Yonghui Wang[c] and Lianming Wang[b,*]

[a]School of Physics, Northeast Normal University, Chang Chun, 130024, China
[b]School of Marine Information Engineering, Hainan Tropical Ocean University, San Ya, 572022, China
[c]Department of Computer Science, Prairie View A&M University, Prairie View, 77446, TX, USA

ARTICLE INFO

*Keywords*:
Experimental System
Fish Swimming
Body-fin Cooperative Propulsion
Animal Pose Estimation

ABSTRACT

Advances in science and technology have led to the development of advanced underwater robots, significantly improving humanity's ability to understand, explore, exploit, and protect the oceans. The secret of efficient swimming in fish provides an essential theoretical basis for developing new underwater robots with low noise, high efficiency, high stability, and high maneuverability. The tools or methods to study the movement of fish are the key to unlocking the secret of efficient fish swimming. Therefore, this study builds an pose and motion estimation platform based on advanced computer vision technology for studying fish motion. The platform can quantify the motion of fish, enable 3D pose estimation of the fish body, and dynamically visualize the motion data of marked vital points. We used this experimental system to study the movement of koi (forward, backward, floating, and turning, etc.) to demonstrate the advantages of the experimental system we designed. The research results provide innovative ideas for fish motion research methods, which can be valuable in designing, developing, and optimizing modern underwater robots. The code and dataset of this study are available at https://github.com/wux024/AdamPosePlug.

## 1. Introduction

With the rapid development of science and technology, resource scarcity has increasingly become a pressing concern, and these problems have prompted humanity to turn its attention to the oceans. However, the complex marine environment greatly challenges human underwater operations. In order to solve these problems, researchers have developed many underwater robots to help human beings obtain information about the ocean. As a result, underwater robots play a vital role in various industries such as marine oil and gas exploration, national defense, maritime search and rescue, oceanographic research, underwater archaeology, and environmental monitoring (Bogue, 2015).

Fish have evolved over millions of years to develop highly efficient locomotor skills that far exceed traditional engineered marine locomotion systems. Their excellent maneuverability and coordination inspire engineers to design and control novel underwater robots within a bionic framework. (Bandyopadhyay, 2005; Vincent, 2006; Yu et al., 2018) The secret of efficient swimming in fish contributes to human understanding of fish behavior, physiology, and ecology and, at the same time, provides essential data and theoretical support for human design and control of novel underwater robots. Therefore, constructing a key that can unlock the secret of efficient fish swimming has become an important research direction. Recognition, detection, and measurement techniques based on advanced computer vision technologies have been developed recently. Researchers have constructed many fish motion analysis systems based on computer vision technology.

**2D vision system for fish movement** Jing et al. (2004) recorded the motion of fish using video and studied the turning kinematics of a frightened carp. Li et al. (2005) built on the previous research on the dorsal fin of carp and recorded richer 2D motion parameters. Yan et al. (2008)designed observation systems that can track the two-dimensional motion of fish more accurately. However, these systems mainly acquired or tracked the fish movement condition in pixel space and needed to realize three-dimensional movement observation to extract more practical features.

**3D vision system for fish movement** Wu (2010) constructed a video tracking system for 3D kinematics of fish using two cameras as well as a reflector, which was able to be used to extract ventral and lateral views of the fish, and with which the pitching and tailing motions of the fish were investigated. Zhang et al. (2011) used X-ray photography to determine the number and length of vertebral joints, which supported the design of joint control for a carp bionic robot. Wang et al. (2013)used a high-speed camera system to analyze the three-dimensional motion of a koi carp, focusing on its pectoral fins. Voesenek et al. (2016)proposed a method for tracking the three-dimensional swimming of fish using multiple high-speed cameras designed with a video system based on image morphological transformations. Mao et al. (2016)discussed the basic techniques of unobstructed tracking, investigated the

⋆This research was partially sponsored by the *Major Science and Technology Projects in Sanya* (ZDKJ-SY-2020-001), the *Scientific Research Foundation of Hainan Tropical Ocean University* (No. RHDRC202001), and the *Major Science and Technology Projects of Hainan Province* (ZDKJ202017-02-01).

*Corresponding author: Lianming Wang
wux024@nenu.edu.cn (X. Wu); huangjp848@nenu.edu.cn (J. Huang); yowang@pvamu.edu (Y. Wang); lmwang@hntou.end.cn (L. Wang)
ORCID(s): 0000-0002-6250-4218 (X. Wu); 0000-0001-5519-9561 (J. Huang); 0000-0003-4334-0921 (Y. Wang); 0000-0002-3695-3909 (L. Wang)

method of unobstructed fish tracking, and realized the three-dimensional unobstructed fish tracking. Saberioon and Cisar (2016) developed three-dimensional occlusion-free tracking of fish using the structured light sensor. Qian and Chen (2017)Built a simple and cost-effective 3D motion detection device for zebrafish, enabling researchers to track the 3D motion of multiple zebrafish at low cost and precise spatial locations. Cheng et al. (2018) proposed a method to acquire quantitative 3D motion trajectories of individual fish using a synchronized camera. These methods acquire more information than 2D vision systems. However, they are limited by traditional computer vision methods, resulting in the extraction and understanding of localized, complex motion features that still need to be improved. Liu et al. (2019) earlier proposed a method for three-dimensional tracking of fish using skeleton analysis and was able to acquire the three-dimensional trajectories of fish, but it was still limited to traditional vision methods and did not acquire more features.

These fish movement observation systems mainly focus on 2D/3D movement tracking and simple movement pattern analysis. However, due to the fish's disorganized movements, these methods cannot analyze localized complex motions more comprehensively, such as detailed characterization of the caudal or pectoral fins. Therefore, extracting local complex motion features of fish using traditional computer vision methods is challenging, and a more effective tool is needed to extract more information for use in order to analyze local complex motion.

Recently, the rapid development of deep learning-based computer vision technology has made it an efficient way of solving complex feature extraction problems. (Afsar et al., 2015; Sarkar et al., 2022; Zhang et al., 2024) Deep learning models can quickly obtain the position and pose of the target object in the target coordinate system. Numerous advanced visual inference tools for animal behavior analysis (Mathis et al., 2018; Lauer et al., 2022; Pereira et al., 2022; Chen et al., 2023) have been developed, and we also report on an algorithm for characterizing the three-dimensional motion of a target (Wu et al., 2022). These advanced computer vision algorithms provide an essential theoretical basis for developing a visual system for observing fish movements.

In this study, we constructed an experimental system specifically for the study of fish movement based on pose estimation, an advanced computer vision technique, which provides a new method for the study of fish movement and is expected to be further applied to the analysis of complex movements of fish. The main research contents and contributions include the following three aspects:

(1) We design an experimental system for fish movement research centered on a multi-view visual perception system for collecting fish movements, equipped with a high-speed camera, a high-performance computer, complementary light, and flow generation;

(2) design the fish motion visualization software based on our algorithms (Wu et al., 2022);

(3) and use the experimental system to study a variety of motions (forward, backward, floating, and turning, etc.) of koi in the tank.

The remainder of this paper is presented as follows: Section 2 describes the design of the experimental system for the study of fish locomotion. Section 3 presents the results of applying the experimental system to study the koi movement and discusses the strengths of the experimental system and details to be improved. Finally, Section 4 provides the concluding remarks of this research.

## 2. Materials and methods

In this section, we describe the technical details of our proposed framework.

### 2.1. Overall framework

In order to be able to record videos of fish swimming comprehensively and to represent fish swimming in a more detailed way, we constructed an experimental system for studying fish movement based on pose estimation. The experimental system integrates a video capture device designed to quickly record fish swimming, an advanced algorithm to represent fish swimming, and an easy-to-use visualization software.

The overall experimental framework is depicted in Figure 1. This consists of four main stages: (1) Fish swimming dataset curation (Section 2.2); (2) Pose estimation-based representation (Section 2.3); (3) Quantifying fish swimming (Section 2.4); and (4) Visualizing fish swimming (Section 2.4).

### 2.2. Fish swimming dataset curation

The platform consists of the following equipment: a water tank (120 cm × 20 cm × 30 cm in length, width, and depth, respectively), a high-performance computer, three high-frame-rate (the maximum of frame rate is 210 FPS), and high-resolution (1280 × 1024 pixels) CCD cameras, a wave maker, and two fill lights (the maximum of power is 100 W). The schematic diagram of the data acquisition equipment is shown in Figure 2.A.

We control the capture device to record the movement of the fish in the pool from three different viewpoints simultaneously (more viewpoints can be extended). In addition, the wavemaker could simulate some typical water currents, and the fill light could provide even light for the whole recording process. Figure 2.B shows the fish swimming images recorded simultaneously from different views by the capture device.

In addition, we need to produce fish swimming datasets that can be used to train 2D pose estimation models.We marked the koi carp with customized 21 key points, including fish head (1), fish eyes (1 × 2), dorsal fin (4), pectoral fin (4×2), fish body (1×2), and caudal fin (4). Figure 2.C shows an example of labeled koi. We also shared all of our data on Google Drive for researchers.

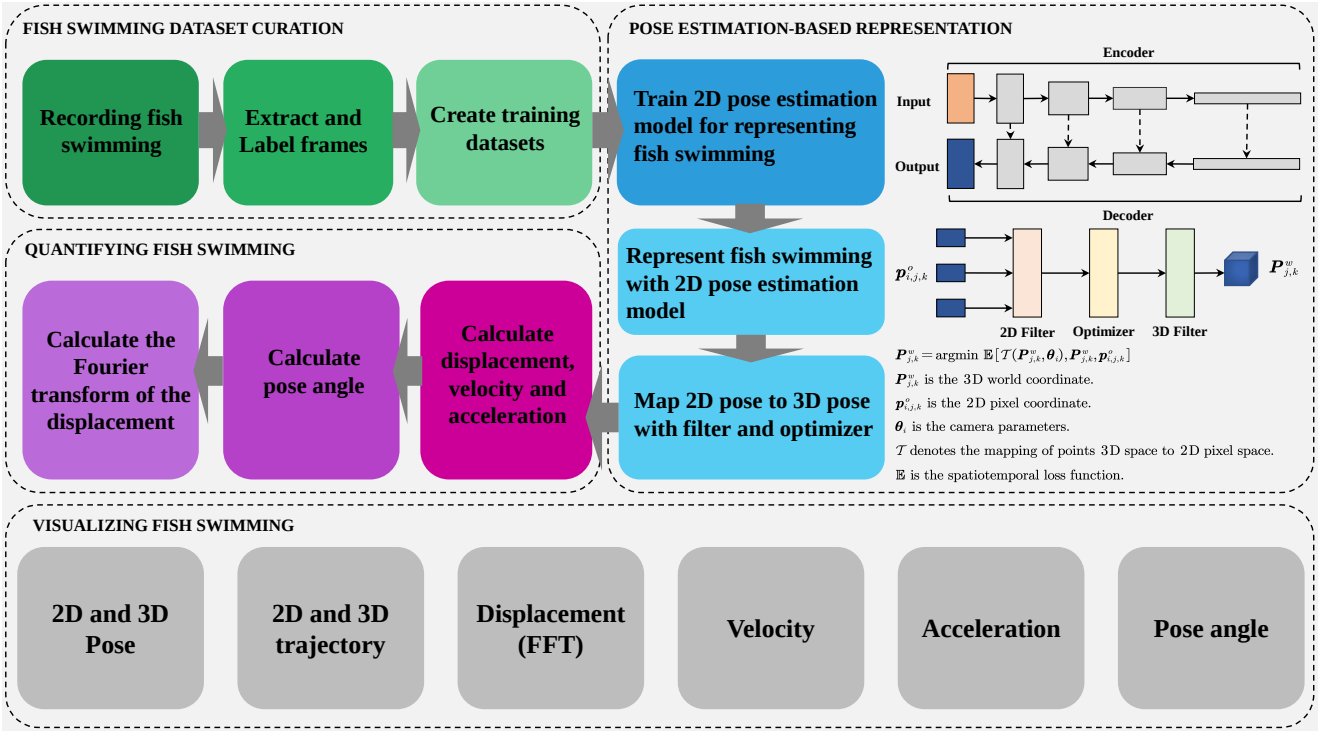



**Figure 1:** The proposed experimental framework consists of four phases. The dashed box includes four different phases, each of which, in turn, contains different steps. The first phase is fish swimming dataset curation, which mainly includes data acquisition, validation, proper extraction and labeling, and creating a dataset suitable for use in the second phase. The second phase is pose estimation-based representation, which mainly includes 2D pose estimation model training, using the 2D pose estimation model, and mapping 2D poses to 3D poses. The third phase quantifies fish swimming, mainly involving calculating various motion parameters. The fourth stage is visualizing fish swimming, which includes six visualization modules.

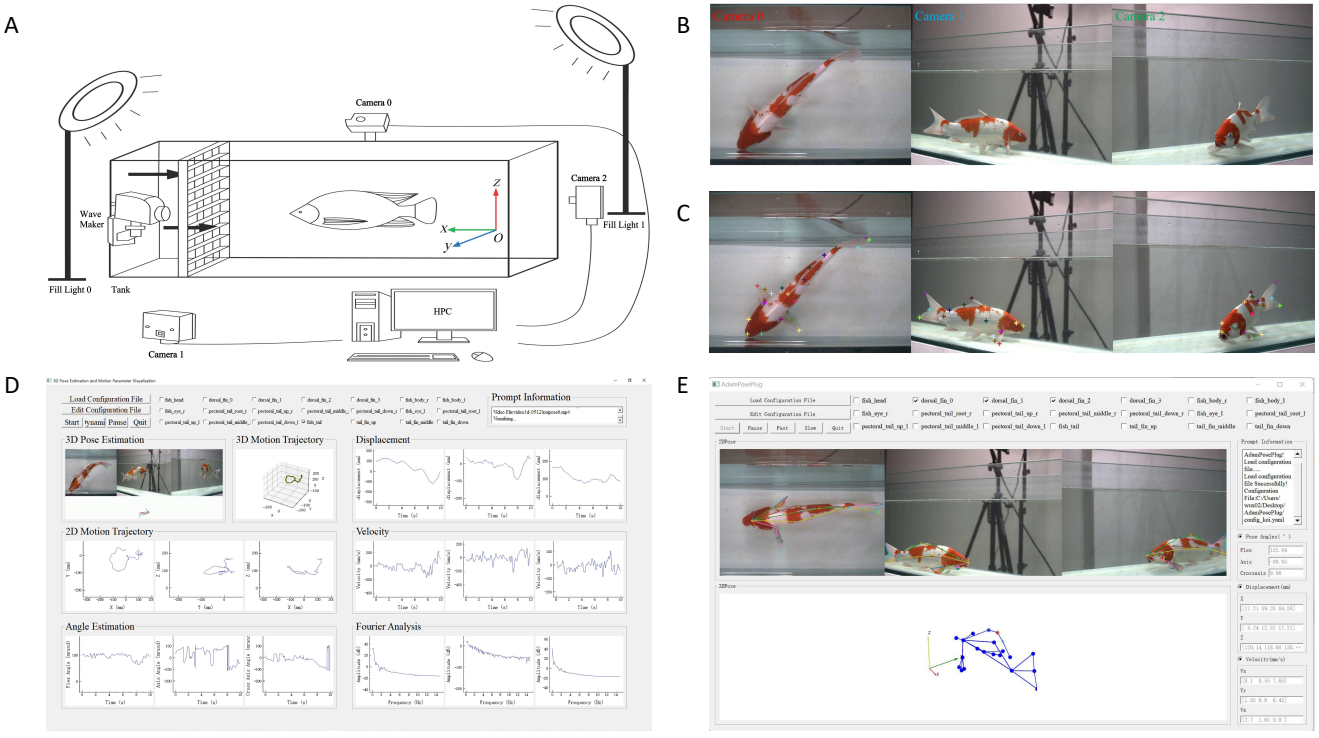



**Figure 2:** (A) This is a schematic of the fish movement research system; (B) an example of image acquisition using the experimental system; (C) an example diagram of koi key point labeling; (D)(E) and the software for quantifying and visualizing fish swimming. We shared an example of the visualization software run on Google Drive.

### 2.3. Pose estimation-based representation

The remarkable performance when using deep learning for human 2D and 3D pose estimation plus dense representations (Ionescu et al., 2014; Cao et al., 2021; Zhang et al., 2022) made this large body of work ripe for exploring its utility in animal behavior. Many machine learning-based tools (Mathis et al., 2018; Lauer et al., 2022; Pereira et al., 2022; Chen et al., 2023) for laboratory experiments have arrived in the past five years. However, these methods are not perfect for 3D pose representation, so we designed algorithms (Wu et al., 2022) to map 2D pose to 3D pose.

As quality tools for 2D animal pose estimation, DeepLabCut (Mathis et al., 2018) and SLEAP (Pereira et al., 2022) show different strengths, respectively; DeepLabCut supports more baseline models and is more accurate, while SLEAP is lighter and can adapt to high frame rates while maintaining accuracy. In addition, they still have in common that both use the encoder-decoder model structure. At the same time, the encoder and decoder all use skip connection (He et al., 2016) to avoid gradient vanishing or explosion during training. The encoder plays the role of feature extraction, while the decoder plays the role of feature re-representation. The encoder mainly consists of a series of convolution (downsampling) kernels, while the decoder consists of deconvolution (upsampling) kernels of the corresponding dimensions. We used ResNet-152 as a baseline model to train models by DeepLabCut that can represent the 2D pose of a fish.

After we obtain the 2D pose representation of the fish, we can map the 2D pose to the 3D pose by solving the optimization problem (1).

$$\boldsymbol{P}^{w}_{j,k} = \arg\min \mathbb{E}\left[\mathcal{T}\left(\boldsymbol{P}^{w}_{j,k}, \boldsymbol{\theta}_i\right), \boldsymbol{P}^{w}_{j,k}, \boldsymbol{p}^{o}_{i,j,k}\right] \tag{1}$$

In the formula (1), $\boldsymbol{P}^{w}_{j,k}$ is the 3D world coordinate, $\boldsymbol{p}^{o}_{i,j,k}$ is the 2D piexl coordinate, $\boldsymbol{\theta}_i$ is the camera parameters, $\mathcal{T}$ denotes the mapping of points 3D space to 2D piexl space, and $\mathbb{E}$ is the spatiotemporal loss function. The algorithm we designed (Wu et al., 2022) knows all the specific details of solving the optimization problem.

In summary, we have implemented a 3D pose representation of fish swimming by combining DeepLabCut and the designed optimization algorithm. In the next stage, we can quantify the fish swimming by calculating the 3D pose representation.

### 2.4. Quantifying and visualizing fish swimming

Computing on the time series composed of 3D key points, the displacement, velocity, and acceleration of each key point can be obtained, which are the most basic features to characterize fish swimming. Tracking any one or more of the 2D key points and 3D key points can obtain the trajectory of the target in 2D pixel space and 3D real space, which is the most essential feature representing the behavior of the fish. Fourier analysis of the time series composed of displacements can analyze the possible periodic motion of the fish swimming. In addition, the angles constituted by any three key points can be calculated based on the geometric relationship, and these angles can represent the local (fin) bending of the fish, which provides a new perspective for the analysis of the local complex motion of fish.

To facilitate easy and quick access to these motion features and to better present these results to the experimenter, we designed software (AdamPosePlug) to integrate these steps, and the software code has been open-sourced on GitHub.

The software runs as shown in Figure 2.D and Figure 2.E. The interface contains the six modules in the fourth stage. The software is divided into two primary interfaces, which mainly qualitatively display the 2D and 3D postures of the fish swimming and quantitatively display the various motion characteristics of the fish swimming, thus facilitating the experimentalists to analyze the fish swimming. The main interface displays the motion parameters that show changes over time, as well as the 2D pose and 3D pose of the fish. The plug-in interface shows more prominently the 2D pose and 3D pose of the fish, as well as displaying motion parameters for selected key points.

## 3. Results and discussions

This section presents the results of our experiments using the experimental system to characterize and visualize koi movement in the experimental tanks and provides some insights into koi movement. The koi (300 mm) used in the experiment were kept in a tank. The experimental platform recorded 40 minutes of koi movement from three different angles, divided into 10-second segments, for a total of 240 videos. The videos were processed using software, and all 3D pose data obtained were stored on Google Drive. We extracted several segments of these data for analysis, and the experimental analysis was divided into three main parts:

(1) We extracted 5 key points labeled with the dorsal fin at the top view angle for analyzing the fish's forward, backward, and turning and showed 2D and 3D pose representations of its floating up.

(2) We extracted 13 key points from the body, caudal, and pectoral fins to analyze koi's body-fin coordination in four different swimming states (forward, backward, turn, and upward).

(3) We extracted pectoral fin key points and then computed pose angles for analyzing pectoral fin flexion in koi under four different swimming states (forward, backward, turning, and upward) to understand the localized complex motion of the pectoral fins.

### 3.1. Overlooking the koi swimming

This section focuses on extracting information about the key points of the dorsal fins of koi in the three swimming states of forward, backward, and turn in the top view angle. A total of 5 points, marked on the back of the fish at the locations of $x = 0.11, 0.15, 0.17, 0.21, 0.30$ m, are selected for the analysis. The $y$-direction displacements of these points are shown in Figure 3. A, plotted along the time axis.

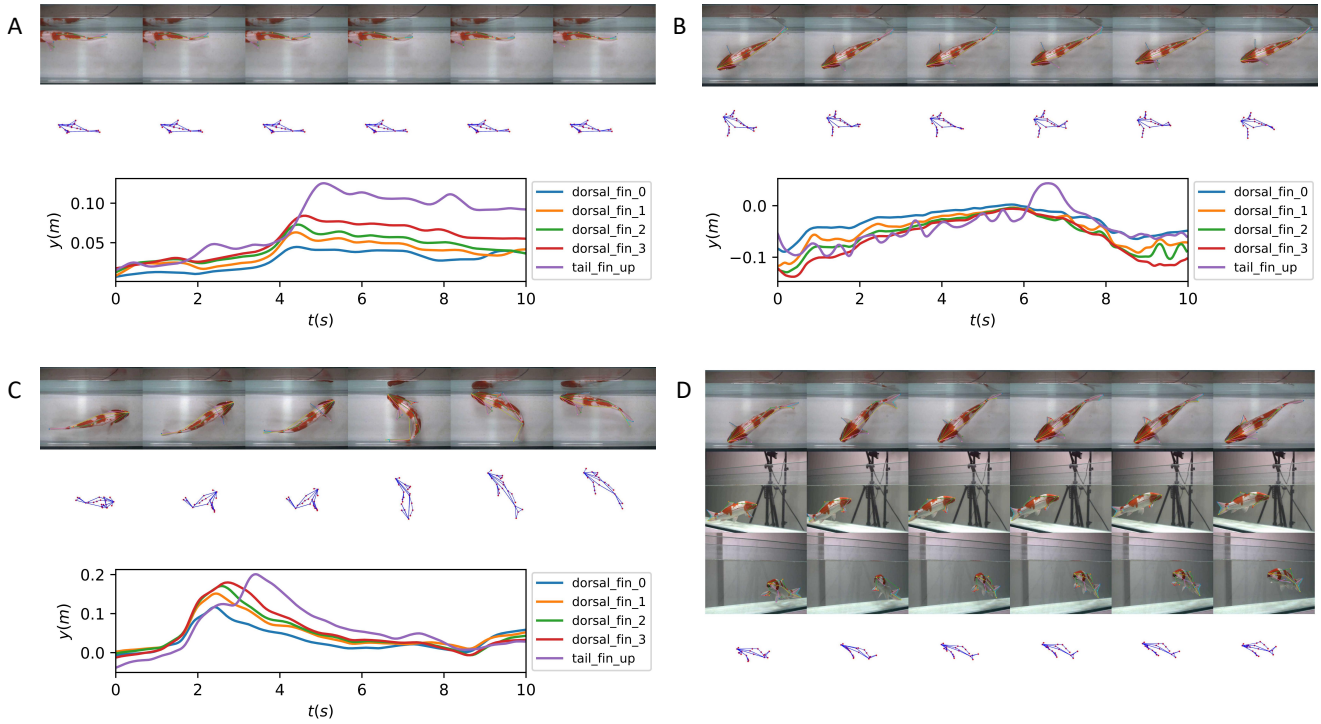



**Figure 3:** (A) Overlooking the koi advancing; (B) Overlooking the koi retreating; (C) Overlooking the koi advancing turning; (D) An example of 3D pose estimation from different views and a virtual perspective when the koi floats up.

**Overlooking the koi advancing** The advancing state of the koi over a while is shown in Figure 3. A. The first row in the figure shows consecutive frames of the actual viewing angle, and the second row shows the 3D pose estimations from a virtual perspective. Both rows show slight bending of the fish's body.

**Overlooking the koi retreating** The actual motion of the koi's retreating state during a period and the 3D pose estimation of the virtual perspective are shown in Figure 3. B. The same key points of observation are used. The displacement curves of these points in the $y$ direction along the time axis are shown in Figure 3. B. For retreating, the swing of the fish's body is not significant; instead, the body has a small amplitude quasi-periodic swing. The main driving parts for retreating are the left and right pectoral fins. Subsequent experiments will carry out a detailed analysis of this state.

**Overlooking the koi turning** The top view of the actual motion status of the koi turning state in a period and its 3D pose estimation are shown in Figure 3. C. The key points observed are the same as those in the previous two states. The displacements of the points in the $y$ direction along the time axis are shown in Figure 3. C. For turning, the fish body first has a slow adjustment process, then swings wildly, and finally, after the turning is completed, the fish body swings and adjusts slightly. The primary body parts responsible for turning are the fish body and the caudal fin, while the other fins have little effect and do not provide significant driving forces. Subsequent experiments will give more detailed body-fin coordination.

**Overlooking the koi floating up** The 2D and 3D poses of the fish surfacing are shown in Figure 3.D. We could observe the 2D poses from different viewpoints and the synthesized 3D pose representations when the koi surfaced.

## 3.2. Body-fin synergistic for fish swimming

In this section, 13 key points of the body, pectoral, and caudal fins were extracted using the experimental system for analyzing each fin's synergy in the koi's four swimming states, namely, forward, backward, turning, and upward floating. The process of body-fin coordination during sinking is similar to that of surfacing, and therefore, sinking is not measured.

**Body-fin synergistic for koi advancing** The plots of the movement parameters of the fish body's key points and each fin's key points (labeled as in Figure 2. C) when the koi moves forward. Figure 4.A shows the displacement curves of the four key points of the fish body relative to the fish head, the displacement curves of the three outer key points on the caudal fin relative to the base of the caudal fin, and the displacement curves of the outer key points on the left and right pectoral fins relative to the corresponding bases of the left and right pectoral fins. Each row shows the $s-t$ curves of different key points relative to the same body part in the three directions of $x$, $y$, and $z$. Fourier analysis of these curves approximates the flapping frequency.

From Figure 4. A, it can be seen that when the koi swims forward, the fish body slowly swings in the $y$ direction, and the maximum swing amplitude exceeds 7.5 cm. At the same time, the caudal fin either sways extensively in the $y$ direction

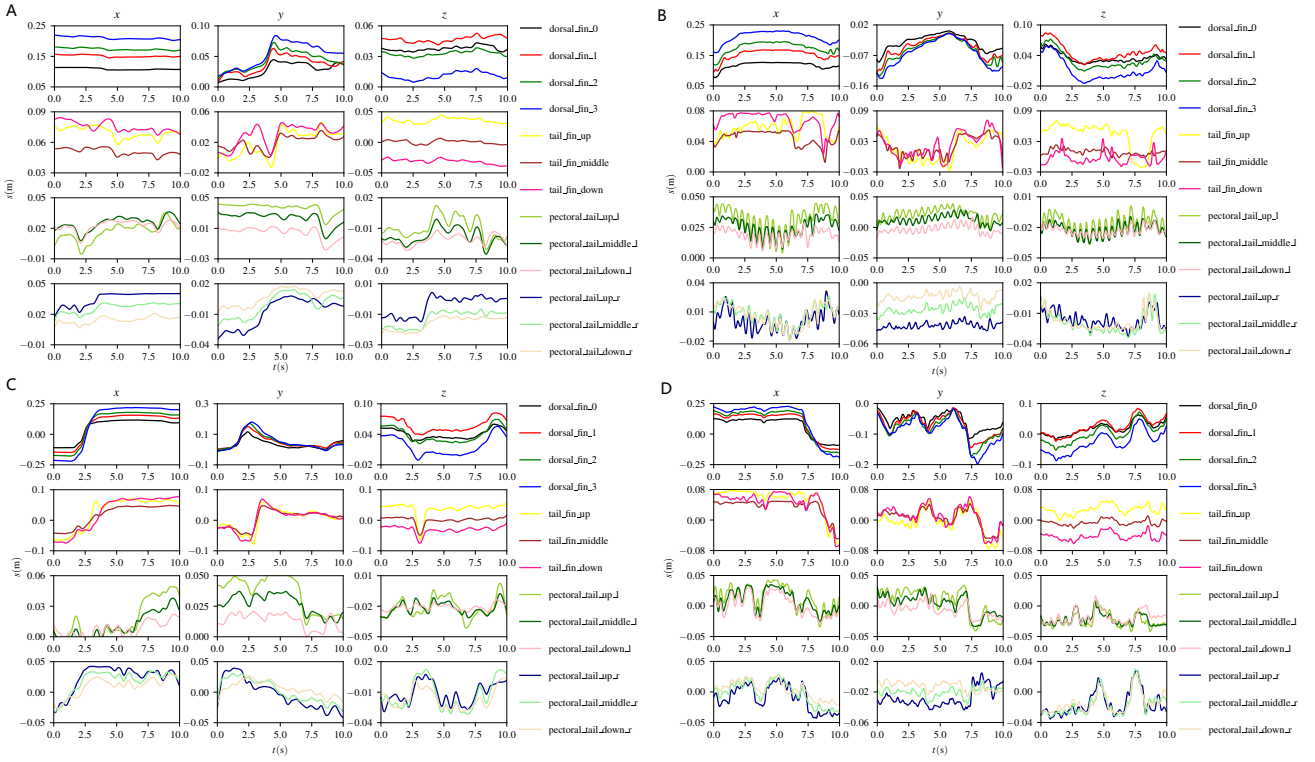



**Figure 4:** Relative displacement measurements of 13 key points as the koi moves (A) forward; (B) backward; (C) turn; and (D) floating up. The first row represents the change of the fish body; the second row represents the change of the caudal fin; the third row represents the change of the left pectoral fin; and the fourth row represents the change of the right pectoral fin.

with the fish body or sways slightly not synchronized with the fish body. The maximum sway of the caudal fin exceeds 3 cm, and the frequency of its big swings is about 0.25 Hz, while the frequency of the small swings is about 1.67 Hz. The left and right pectoral fins flap slightly back and forth and up and down. The maximum amplitude of the back-and-forth flaps exceeds 2 cm, and the flapping frequency is about 0.167 Hz. The maximum amplitude of the up-and-down flaps exceeds 2 cm, and the flapping frequency is about 0.89 Hz.

**Body-fin synergistic for koi retreating** It can be seen from Figure 4. B that when the koi swims backward, the caudal fin does not have apparent periodic movement but has a high deformation frequency. The left pectoral fin flaps back and forth with relatively high frequency. The maximum swing amplitude is about 2 cm, and the swing frequency is about 2 Hz. In addition, its up-and-down stroke amplitude exceeds 2 cm, and the stroke frequency is about 2.5 Hz. The right pectoral fin is not highly synchronized with the left pectoral fin. Its back-and-forth stroke frequency is high, the maximum swing amplitude is about 3 cm, and the swing frequency is about 1.32 Hz, while the up-and-down flap amplitude is about 2 cm and the flapping frequency is about 1.5 Hz. The asynchrony of the left and right pectoral fins is one of the main reasons for its tortuous moving path as the fish retreats.

**Body-fin synergistic for koi turning** The same fish body parts used to analyze fish advancing are used. As can be seen from Figure 4. C, together with the fish body, the tail fin bends considerably with a high frequency. The entire turning process is completed in about 4 seconds, and the swing range of the tail fin is about 12 cm. During this left-turning process, the right pectoral fin is retracted to its belly, and the left pectoral fin is kept away from the belly to maintain the balance of the fish body through deformations. The turning process depends on the body's sudden movement, while the fins mainly balance and control the turning direction.

**Body-fin synergistic for koi floating up** Figure 4.D shows consecutive frames from a video of the koi floating up. The first three rows are consecutive frames of the actual observation perspectives, and the last row is the 3D pose estimation under the virtual perspective.

It can be seen from Figure 4.D that when the koi floats up, the fish body swings left and right with large amplitude, the maximum swing amplitude is about 8 cm, and the swing frequency is about 0.4 Hz. The caudal fin moves with the fish body at a small amplitude; the maximum amplitude of the movement is about 5 cm, and the movement frequency is about 0.4 Hz. The swing amplitudes of the pectoral fins are negligible. Their maximum swing amplitude is about 2 cm, the swing frequency of the left pectoral fin is about 1.5 Hz, while the swing frequency of the right pectoral fin is 1.2 Hz.

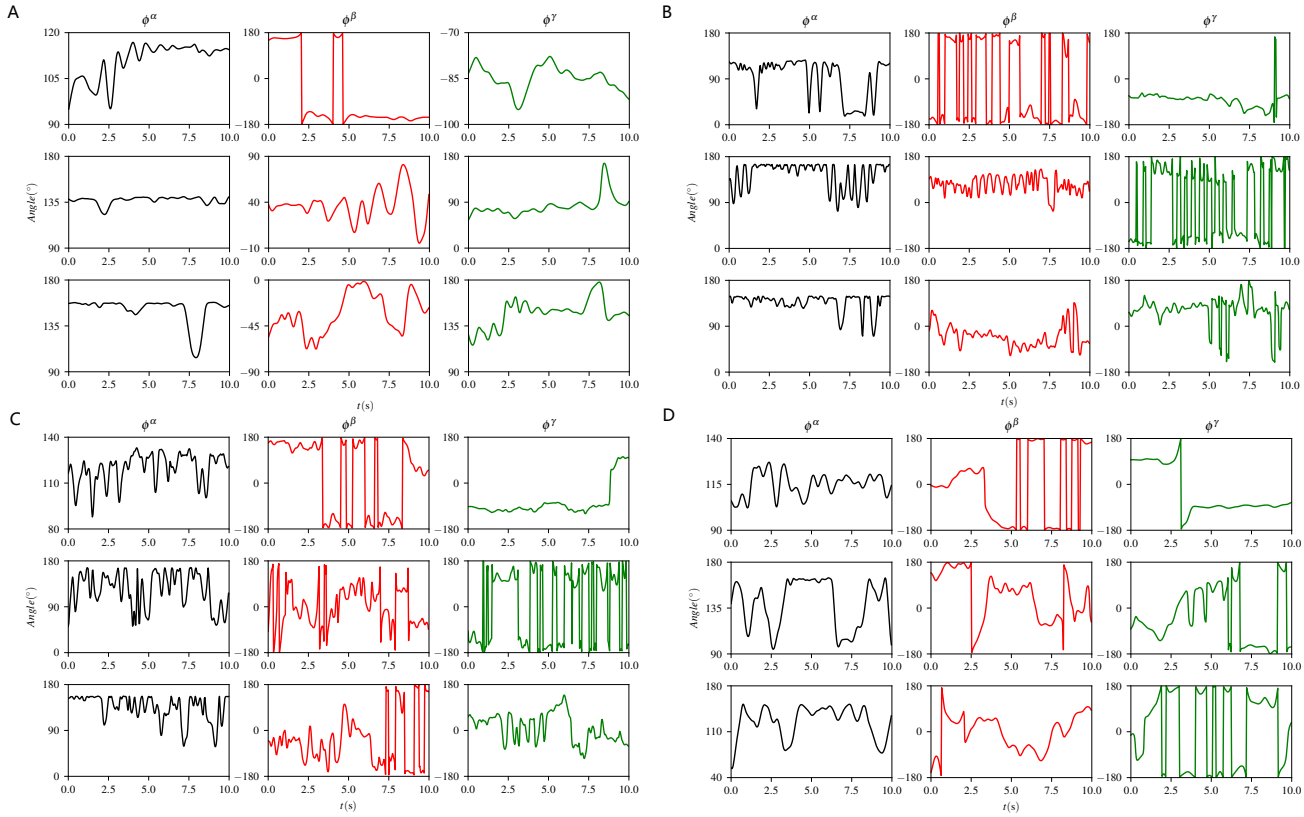



**Figure 5:** The deformation measurements of each fin when the koi moves (A) forward; (B) backward; (C) turns; and (D) floats up. The first row is the deformation of the caudal fin; and the second and third rows are the deformation of the left and right pectoral fins, respectively.

### 3.3. Representing pectoral fin morphology for fish swimming

In addition, when fish swims, each fin will deform, and the pose angle can be used to measure the bending degree of each fin. The visual result of the deformation measurement of each fin is shown in Figure 5.

**Representing pectoral fin morphology for koi advancing** From Figure 5. A, it can be seen that the koi fins deform when the fish swims forward. According to the definition of pose angle, the first column of the figure represents the flexion and extension angle, and the variation range of the caudal fin flexion and extension angle is 95.0° ∼ 116.8°. The variation range of the left pectoral fin flexion and extension angle is 123.0° ∼ 140.4°, and the range of the right pectoral fin flexion and extension angle is 103.9° ∼ 158.1°. The second and third columns of the figure represent complex rotations, which are directional and more complex deformations. These complex deformations are mainly to maintain the fish's balance and control its movement's direction.

**Representing pectoral fin morphology for koi retreating** It can be seen from Figure 5. B that, in retreating, the koi fins deform with relatively high deformation frequency. The first column of the figure shows that the range of flexion and extension angle of the caudal fin is 16.8° ∼ 126.9°; the range of flexion and extension angle of the left pectoral fin is 73.2° ∼ 165.3°; and the range of flexion and extension angle of the right pectoral fin is 70.5° ∼ 149.3°. The second and third columns of the figure represent complex rotations, which are directional and more complex deformations. These complex deformations are mainly to maintain the balance of the fish and control the direction of its retreat.

**Representing pectoral fin morphology for koi turning** Figure 5. C reflects the deformations of the koi fins in the turning process. The figure's first column shows that the variation range of caudal fin flexion and extension angle is 102.2° ∼ 127.1°, the variation range of left pectoral fin flexion and extension angle is 94.6° ∼ 164.6°, and the variation range of right pectoral fin flexion and extension angle is 53.7° ∼ 152.0°. The second and third columns of the figure represent complex rotations, which are directional and more complex deformations. These complex deformations mainly maintain the fish's balance and control its turning direction.

**Representing pectoral fin morphology for floating up** Figure 5.D reflects the koi fin deformations in the floating-up process. As shown in the first column, the variation range of caudal fin flexion and extension angle is 87.9° ∼ 133.1°, the variation range of left pectoral fin flexion and extension angle is 51.6° ∼ 167.0°, and the variation range of right pectoral fin flexion and extension angle is 58.1° ∼ 157.4°. The flexion and extension angle variations of the fins for floating up are far more severe than those for the previous

three motion states. The second and third columns of the figure represent complex rotations, which are directional and more complex deformations. These complex deformations are mainly to maintain the balance of the fish and control its direction of floating up.

### 3.4. Discussion

In this section, we extracted the characteristics of koi using an experimental system and combined these characteristics to analyze their four swimming states in still water. We analyzed koi swimming from different perspectives, with the highlight being the study of body-fin synergy in koi swimming. For koi swimming in static water with a body length of $L$, their body-fin synergies in the four swimming states (forward, backward, turn, and upward) are summarized in the following four points.

(1) When the koi moves forward, the fish body swings left and right with amplitude $A_1 < 0.34L$, and the swing frequency is 0.28 Hz< $f_1$ <0.67 Hz. At the same time, the tail fin swings together with the fish body at an amplitude of $A_2 \approx A_1$, and a swing frequency at $f_2 \approx f_1$; or it does not swing together with the fish body with an amplitude of $A_3 < 0.17L$, and a swing frequency at $f_3 = 3.0 \sim 7.0f_1$. The left and right pectoral fins move back and forth or flap up and down, and the flap frequencies are $f_{=}0.5 \sim 1.0f_1$ and $f_5 = 0.5 \sim 1.0f_3$.

(2) To retreat, the koi mainly depends on the back-and-forth and up-and-down strokes of the left and right pectoral fins, while the fish body and tail fin provide a low contribution. The left pectoral fin's back-and-forth stroke frequency is 2 Hz< $f_6$ <3 Hz, yet the right pectoral fin is not highly synchronized with the left pectoral fin, and its back-and-forth stroke frequency is 2 Hz< $f_7$ <3 Hz, where $f_6 \neq f_7$. The asynchrony of the left and right pectoral fins is one of the main reasons for the koi's tortuous retreating path.

(3) When the koi turns left (right), it is mainly done by the rapid fish body and tail fin movements. The fish body significantly bends, with a bending amplitude upper limit of $0.6L$. Together with the body, the caudal fin significantly bends with the bending amplitude upper limit of $0.4L$. At the same time, the right (left) pectoral fin is retracted to the fish belly, while the left (right) pectoral fin is kept away from the fish belly.

(4) When the koi floats up, the fish body swings left and right with the amplitude $A_4 < 0.27L$ and the swing frequency $f_8$ <0.5 Hz. At the same time, the caudal fin swings left and right with the amplitude $A_5 < 0.17L$ and the swing frequency $f_9$ <0.5 Hz, where $f_8 \approx f_9$. The left and right pectoral fins flap up and down, while the flapping frequency of the left pectoral fin is $f_{10}$ <2 Hz, and the flapping frequency of the right pectoral fin is $f_{11}$ <2 Hz, where $f_{10} \neq f_{11}$.

During all these processes, each fin undergoes complex deformation, and the deformation includes complex fin surface fluctuations. These complex deformations are mainly to maintain the balance of the fish body and control the moving directions.

In summary, the experimental system we designed integrates three modules of acquisition, processing, and visualization and has the following three main advantages:

(1) We utilized multiple high-frame-rate cameras to record fish swimming, which compensates for the incomplete recording of raw motion perspectives in other fish-observing systems, and the high-frame-rate cameras can capture faster forms of motion.

(2) We used an advanced animal pose estimation method to represent fish swimming, compensating for the shortcomings of other fish observation systems that remember a single feature to characterize fish swimming and thus are insufficient to describe the complex motions of fish swimming.

(3) We designed high-performance software to integrate the experimental steps, facilitating researchers to obtain quantitative and qualitative fish swimming representations quickly.

The experiments proved the effectiveness of the experimental system we designed and provided new ideas for studying fish movement. Despite the advantages of the experimental system we designed, there are still some areas for expansion:

(1) Designing more advanced algorithms for animal pose estimation. More advanced algorithms mainly refer to single-view animal 3D pose estimation and dense animal representation (Xu et al., 2023). The single-view animal 3D pose estimation can save the cost of acquisition equipment, and the animal-dense representation can provide richer information for the study of fish surface bending.

(2) Combining particle image velocimetry systems gives a novel representation of fish swimming with flow field and pose.

## 4. Conclusion

In this paper, a pose estimation-based experimental platform is constructed to study the movement of fish. The experimental platform can quickly capture videos of fish movements and extract the key points of the fish by pose estimation algorithm and compute the motion parameters of interest. We designed the visualization software (AdamPose-Plug) to enable all these steps. The experimental platform constructed in this study offers great potential for fish movement studies. In the future, this experimental platform may be combined with particle image velocimetry equipment to represent fish movement in more detail by jointly analyzing fish pose and flow field data.

## CRediT authorship contribution statement

**Xin Wu:** Conceptualization, Methodology, Software, Validation, Formal Analysis, Investigation, Data Curation, Writing – Original Draft Preparation, Visualization. **Jipeng Huang:** Investigation, Resources, Writing – Review & Editing, Supervision, Project Administration. **Yonghui Wang:** Writing – Review & Editing, Supervision. **Lianming Wang:** Investigation, Resources, Writing – Review & Editing, Supervision, Project Administration, Funding Acquisition.

## Ethics statement

The experimental procedures were carried out following animal ethics approval granted by Northeast Normal University. All experimental procedures in this study were approved by the National Animal Research Authority of Northeast Normal University, China (approval number NENU-20080416) and the Forestry Bureau of Jilin Province of China (approval number [2006]178).

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgment

This research was partially sponsored by the *Major Science and Technology Projects in Sanya* (ZDKJ-SY-2020-001), the *Scientific Research Foundation of Hainan Tropical Ocean University* (No. RHDRC202001), and the *Major Science and Technology Projects of Hainan Province* (ZDKJ202017-02-01).

## References

Afsar, P., Cortez, P., Santos, H., 2015. Automatic visual detection of human behavior: A review from 2000 to 2014. Expert Systems with Applications 42, 6935–6956. URL: https://www.sciencedirect.com/science/article/pii/S0957417415003516, doi:https://doi.org/10.1016/j.eswa.2015.05.023.

Bandyopadhyay, P.R., 2005. Trends in biorobotic autonomous undersea vehicles. IEEE Journal of Oceanic Engineering 30, 109–139. doi:10.1109/JOE.2005.843748.

Bogue, R., 2015. Underwater robots: a review of technologies and applications. Industrial Robot: An International Journal 42, 186–191. URL: https://doi.org/10.1108/IR-01-2015-0010, doi:10.1108/IR-01-2015-0010.

Cao, Z., Hidalgo, G., Simon, T., Wei, S.E., Sheikh, Y., 2021. Openpose: Realtime multi-person 2d pose estimation using part affinity fields. IEEE Transactions on Pattern Analysis and Machine Intelligence 43, 172–186. doi:10.1109/TPAMI.2019.2929257.

Chen, Z., Zhang, R., Fang, H.S., Zhang, Y.E., Bal, A., Zhou, H., Rock, R.R., Padilla-Coreano, N., Keyes, L.R., Zhu, H., Li, Y.L., Komiyama, T., Tye, K.M., Lu, C., 2023. Alphatracker: a multi-animal tracking and behavioral analysis tool. Frontiers in Behavioral Neuroscience 17. URL: https://www.frontiersin.org/articles/10.3389/fnbeh.2023.1111908, doi:10.3389/fnbeh.2023.1111908.

Cheng, X.E., Du, S.S., Li, H.Y., Hu, J.F., Chen, M.L., 2018. Obtaining three-dimensional trajectory of multiple fish in water tank via video tracking. Multimedia Tools and Applications 77, 24499–24519. URL: https://doi.org/10.1007/s11042-018-5755-5, doi:10.1007/s11042-018-5755-5.

He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition, in: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 770–778. doi:10.1109/CVPR.2016.90.

Ionescu, C., Papava, D., Olaru, V., Sminchisescu, C., 2014. Human3.6m: Large scale datasets and predictive methods for 3d human sensing in natural environments. IEEE Transactions on Pattern Analysis and Machine Intelligence 36, 1325–1339. doi:10.1109/TPAMI.2013.248.

Jing, J., Yin, X., Lu, X., 2004. Hydrodynamic analysis of c-start in crucian carp. Journal of Bionic Engineering 1, 102–107. URL: https://doi.org/10.1007/BF03399460, doi:10.1007/BF03399460.

Lauer, J., Zhou, M., Ye, S., Menegas, W., Schneider, S., Nath, T., Rahman, M.M., Santo, V.D., Soberanes, D., Feng, G., Murthy, V.N., Lauder, G., Dulac, C., Mathis, M.W., Mathis, A., 2022. Multi-animal pose estimation, identification and tracking with deeplabcut. Nature Methods 19, 496–504. URL: https://doi.org/10.1038/s41592-022-01443-0, doi:10.1038/s41592-022-01443-0.

Li, F., jiang Hu, T., ming Wang, G., cheng Shen, L., 2005. Locomotion of gymnarchus niloticus : Experiment and kinematics. Journal of Bionic Engineering 2, 115–121. URL: https://doi.org/10.1007/BF03399488, doi:10.1007/BF03399488.

Liu, X., Yue, Y., Shi, M., Qian, Z.M., 2019. 3-d video tracking of multiple fish in a water tank. IEEE Access 7, 145049–145059. doi:10.1109/ACCESS.2019.2945606.

Mao, J., Xiao, G., Sheng, W., Qu, Z., Liu, Y., 2016. Research on realizing the 3d occlusion tracking location method of fish's school target. Neurocomputing 214, 61–79. URL: https://www.sciencedirect.com/science/article/pii/S0925231216305215, doi:https://doi.org/10.1016/j.neucom.2016.05.067.

Mathis, A., Mamidanna, P., Cury, K.M., Abe, T., Murthy, V.N., Mathis, M.W., Bethge, M., 2018. Deeplabcut: markerless pose estimation of user-defined body parts with deep learning. Nature Neuroscience 21, 1281–1289. URL: https://doi.org/10.1038/s41593-018-0209-y, doi:10.1038/s41593-018-0209-y.

Pereira, T.D., Tabris, N., Matsliah, A., Turner, D.M., Li, J., Ravindranath, S., Papadoyannis, E.S., Normand, E., Deutsch, D.S., Wang, Z.Y., McKenzie-Smith, G.C., Mitelut, C.C., Castro, M.D., D'Uva, J., Kislin, M., Sanes, D.H., Kocher, S.D., Wang, S.S.H., Falkner, A.L., Shaevitz, J.W., Murthy, M., 2022. Sleap: A deep learning system for multi-animal pose tracking. Nature Methods 19, 486–495. URL: https://doi.org/10.1038/s41592-022-01426-1, doi:10.1038/s41592-022-01426-1.

Qian, Z.M., Chen, Y.Q., 2017. Feature point based 3d tracking of multiple fish from multi-view images. PLOS ONE 12, e0180254–. URL: https://doi.org/10.1371/journal.pone.0180254.

Saberioon, M.M., Cisar, P., 2016. Automated multiple fish tracking in three-dimension using a structured light sensor. Computers and Electronics in Agriculture 121, 215–221. URL: https://www.sciencedirect.com/science/article/pii/S0168169915003968, doi:https://doi.org/10.1016/j.compag.2015.12.014.

Sarkar, A., Banerjee, A., Singh, P.K., Sarkar, R., 2022. 3d human action recognition: Through the eyes of researchers. Expert Systems with Applications 193, 116424. URL: https://www.sciencedirect.com/science/article/pii/S0957417421017115, doi:https://doi.org/10.1016/j.eswa.2021.116424.

Vincent, J.F.V., 2006. Applications — influence of biology on engineering. Journal of Bionic Engineering 3, 161–177. URL: https://doi.org/10.1016/S1672-6529(06)60020-6, doi:10.1016/S1672-6529(06)60020-6.

Voesenek, C.J., Pieters, R.P.M., van Leeuwen, J.L., 2016. Automated reconstruction of three-dimensional fish motion, forces, and torques. PLOS ONE 11, e0146682–. URL: https://doi.org/10.1371/journal.pone.0146682.

Wang, L., Xu, M., Liu, B., Low, K.H., Yang, J., Zhang, S., 2013. A three-dimensional kinematics analysis of a koi carp pectoral fin by digital image processing. Journal of Bionic Engineering 10, 210–221. URL: https://www.sciencedirect.com/science/article/pii/S1672652913602176, doi:https://doi.org/10.1016/S1672-6529(13)

60217-6.

Wu, G., 2010. Measuring the three-dimensional kinematics of a free-swimming koi carp by video tracking method. Journal of Bionic Engineering 7, 49–55. URL: https://www.sciencedirect.com/science/article/pii/S1672652909601906, doi:https://doi.org/10.1016/S1672-6529(09)60190-6.

Wu, X., Wang, Y., Chen, L., Zhang, L., Wang, L., 2022. Motion parameters measurement of user-defined key points using 3d pose estimation. Engineering Applications of Artificial Intelligence 110, 104667. URL: https://linkinghub.elsevier.com/retrieve/pii/S0952197622000045, doi:10.1016/j.engappai.2022.104667.

Xu, J., Zhang, Y., Peng, J., Ma, W., Jesslen, A., Ji, P., Hu, Q., Zhang, J., Liu, Q., Wang, J., Ji, W., Wang, C., Yuan, X., Kaushik, P., Zhang, G., Liu, J., Xie, Y., Cui, Y., Yuille, A., Kortylewski, A., 2023. Animal3d: A comprehensive dataset of 3d animal pose and shape, in: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), pp. 9099–9109.

Yan, H., min Su, Y., Yang, L., 2008. Experimentation of fish swimming based on tracking locomotion locus. Journal of Bionic Engineering 5, 258–263. URL: https://doi.org/10.1016/S1672-6529(08)60033-5, doi:10.1016/S1672-6529(08)60033-5.

Yu, J., Wang, M., Dong, H., Zhang, Y., Wu, Z., 2018. Motion control and motion coordination of bionic robotic fish: A review. Journal of Bionic Engineering 15, 579–598. URL: https://doi.org/10.1007/s42235-018-0048-2, doi:10.1007/s42235-018-0048-2.

Zhang, H., Cao, J., Lu, G., Ouyang, W., Sun, Z., 2022. Learning 3d human shape and pose from dense body parts. IEEE Transactions on Pattern Analysis and Machine Intelligence 44, 2610–2627. doi:10.1109/TPAMI.2020.3042341.

Zhang, S., Yang, Y., Chen, C., Zhang, X., Leng, Q., Zhao, X., 2024. Deep learning-based multimodal emotion recognition from audio, visual, and text modalities: A systematic review of recent advancements and future prospects. Expert Systems with Applications 237, 121692. URL: https://www.sciencedirect.com/science/article/pii/S0957417423021942, doi:https://doi.org/10.1016/j.eswa.2023.121692.

Zhang, Y., He, J., Zhang, G., 2011. Measurement on morphology and kinematics of crucian vertebral joints. Journal of Bionic Engineering 8, 10–17. URL: https://www.sciencedirect.com/science/article/pii/S1672652911600061, doi:https://doi.org/10.1016/S1672-6529(11)60006-1.