Application of Computer Vision and Deep Learning in Swimming Action Recognition and Evaluation

Wenzhi Hou[1 *], Pingyang Wang[1], Xiumin Lv[1], Jiting Yang[2], Wei Man[3] and Xiaoyi Zhao[4]

{*Corresponding author: houwenzhiswimming@163.com}

[1]Basic Teaching Department, Weihai Campus of Shandong Communications University, Jinan City, Shandong Province, 250000, China

[2]Hunan Women's University, Changsha, Hunan Province, 410013, China

[3]Department of Ophthalmology, Jiaozuo Hospital, Tongji University, Qingdao, Shandong Province, 266000, China

[4]Public Sports Department of Hebei North University, Zhangjiakou, Hebei Province, 075000, China

Abstract. China's position in the world swimming industry is becoming increasingly high, which is also an important achievement in the swimming industry. Improving the strength of athletes is urgent. This article used a combination of sliding window technology and mixed Gaussian model to model the underwater video background of a swimming pool. A video monitoring model based on computer vision and deep learning was used to recognize the actions of swimmers. This article extracted breaststroke motion data from various stages, as well as motion data from freestyle, butterfly stroke, and different stages of breaststroke, and conducted classification and recognition work. This article evaluated the recognition accuracy of different stages of breaststroke and the specific stage actions that were easily misclassified as freestyle and butterfly stroke, in order to achieve overall recognition and stage recognition of swimming posture and expand the application range of recognition models in swimming motion recognition and evaluation. Swimming motion recognition found that in butterfly stroke, the maximum pitch angle between the head and hips was 110°, and the maximum pitch angle between the head and hips was 42°. This article provided a sports guidance system suitable for swimming, which helped empower skills for training and enables most athletes to grow rapidly.

Keywords: Computer Vision, Deep Learning, Swimming Action Recognition, Action Evaluation, Background Modeling

1. Introduction

With the improvement of people's living standards and the rapid development of science and technology, people are paying more and more attention to the body.

ICSETPSD 2023, November 17-18, Coimbatore, India

DOI 10.4108/eai.17-11-2023.2342701

Swimming, which can enhance the functions of the human respiratory and cardiovascular systems, has also developed rapidly and is one of the most popular sports and fitness projects in the 21st century. Swimming is a commonly used medical rehabilitation method, and its changes in movement coordination and symmetry can better reflect the body's sports injuries, spinal diseases, etc. The swimming posture monitoring and recognition based on computer vision and deep learning carried out in this article could directly observe and measure the movement status of the human body in water, providing strong support for people to better understand the movement status of the human body in water. At the same time, detailed monitoring data could also help athletes track and analyze the movement status in real-time during the exercise process.

In terms of swimming skills, the competitive level of swimming competitions at all levels is gradually improving, and the gap between top athletes is gradually narrowing. The number of athletes in the same level is also increasing, which brings great pressure to athletes and coaches. In this context, competition among athletes at various levels revolves more around the control of technical movements and how to cultivate excellent swimming rhythms in daily training. Swimming techniques are classified according to specific events, and the details of each event are deeply explored to gain a clearer understanding of the sports characteristics of each action. In order to solve the core problem of swimming behavior recognition in the digital swimming monitoring system, Xiao Ju planed to use a combination of a single attitude sensor and a decision tree classifier to achieve autonomous recognition of swimming behavior. He first used a single wireless pose sensor to collect real-time three-dimensional acceleration and angular velocity during swimming, and combined them with real swimming categories to establish a swimming behavior database. On this basis, he proposed a decision tree classification algorithm based on genetic algorithm and improved it. On the basis of the classification and recognition results, he could accurately estimate the moment of the swimming action switching point, with an average estimated time error of 0.186s for training data and 0.209s for testing data [1]. Based on existing research, Xiang Hongbiao utilized improved algorithms to extract obstacle information from complex scenes for trajectory planning, and combined arc smooth labeling to achieve real-time updates of the center and corner of the swimming robot in complex scenes. He achieved dual closed-loop follow-up control of the swimming posture and posture of the magnetic controlled micro robot in complex environments, and verified the effectiveness and smoothness of the proposed method through experimental verification, as well as the effectiveness and accuracy of the identification and tracking methods [2]. Wang Song studied swimming data collection and monitoring systems based on attitude sensors, and he believed that detailed recording and analysis of swimming movements was of great significance. At present, wearable products on the market mainly focus on sports such as running, walking, and cycling, while there are very few that have swimming monitoring functions [3]. Although they have developed a more advanced understanding of competitive swimming, there is still room for improvement in the recognition of swimmer movements.

In water, moving objects would change their position according to changes in time. In addition, due to the different postures, postures, and clothing of different athletes, it is difficult to provide a unified feature description for athletes, which increases the difficulty of detecting athlete movement targets. Compared to general detection environments, there are certain differences between underwater environments. This article analyzed the principles of different background modeling algorithms, and then conducted corresponding experimental simulations. Based on the experimental results, the modeling accuracy and real-time performance were analyzed. This article explored the background simulation method based on sliding model and makes corresponding improvements to it. The video monitoring model used in this article included video data collection end, management monitoring end, and mobile monitoring end.

2. Exploring Methods for Identifying and Evaluating Swimming Movements

2.1 Modeling of Underwater Video Background in Swimming Pools Based on Moving Mean Model

This article presents a method based on the combination of sliding window technology and mixed Gaussian model, which achieves adaptive ability to real-time changes in underwater scenes and irregular interference distribution.

On the T-period, a moving average model is established using a mixed Gaussian distribution for continuous swimming action video sequences in time [4-5]. This method utilizes the latest observation data to continuously update and replace parameters in the pattern, in order to reduce errors caused by long-term observations.

The probability density function of each pixel in a swimming action video is as follows [6]:

$$P_M = \sum w * \beta(X, \mu_m, \sum \theta) \tag{1}$$

μ_m represents the mean of the m-th Gaussian function of the pixel.

In a swimming action image, an adjacent image in the image is taken and compared with the K Gaussian distributions of the image in the background model [7-8]. In the background model, if the deviation between the current pixel value and the average coefficient does not exceed 2.5 times, it indicates that the current pixel value satisfies the Gaussian distribution feature. If the current pixel value matches the background model, the parameters of the model must be updated:

$$w_{i,k} = (1-\theta) * w * \beta(X, Y) + \beta P_L \tag{2}$$

P_L is the efficiency of background update.

After setting the parameters, a background screen would be generated. The accuracy of initializing the background model is closely related to the number of frames involved. When the system is initialized, due to external disturbances or objects entering, it would have a significant impact on the accuracy of background modeling. In order to obtain an accurate initial background model, it is necessary to manually set an appropriate initial time.

2.2 Video Surveillance Based on Computer Vision

Intelligent video surveillance is achieved through computer vision technology. It has transformed from a single function of recording videos to two functions of recording and extracting information, enabling real-time analysis and interpretation. With the continuous development of video surveillance technology, its application scope is also becoming wider and wider.

The digitization and networking of computer vision monitoring mode enables the conversion of collected images from analog to digital transmission. On this basis, a matching DHDV (Digital Highway Data Vehicle) is also equipped, and wireless transmission is adopted, greatly improving the performance of the system. In addition, digital matching of intelligent video surveillance models based on computer vision has a driving effect on the overall intelligent video surveillance industry.

The so-called "networking" refers to connecting various modules through wireless networks, thereby improving the reliability of the entire system and creating more flexible product forms. The intelligent video surveillance system has left great room for improvement in subsequent device access and system upgrades.

This article presents a design method for a multi-level access model based on computer vision. According to different task requirements, different modules are combined together. In terms of intelligent video surveillance technology, it has been able to mimic the human brain for corresponding analysis and recognition. Video surveillance technology based on computer vision would play a leading role in the future video surveillance field. The front-end of the system integrates cameras and image algorithms, and transmits the digital signals collected by the monitoring device over the network to the terminal in a long-distance manner. During this process, the transmission quality of images is improved, while also reducing the overall operating cost of the system.

With the continuous development and popularization of drowning rescue technology, underwater intelligent monitoring technology based on digital image processing technology has become a current research hotspot. In practical applications, the complexity of underwater environments and the differences between different types of objects have brought great difficulties to image processing [9]. Therefore, the research on detection methods for moving objects is particularly important.

This article presents a method for underwater acoustic signal processing based on underwater acoustic signal processing, and studies the methods of underwater acoustic signal processing [10-11].

At the video data collection end, it consists of 8 impermeable cameras, a video server, and a virtual private network line.

The management monitor is a part of terminal control.

The mobile monitoring terminal mainly consists of a handheld tablet computer.

Workflow: The underwater swimming posture video captured by the high-definition camera is returned to the video server, and then fed back to the main control server through a virtual private network dedicated line [12]. On the server side, it is combined with specific deep learning algorithms to perform object detection and alarm judgment. Finally, the alarm signal is wirelessly transmitted to the handheld tablet on site through the network.

(1) Video data acquisition end

Data collection is mainly completed by cameras installed in the water. The instruments used in this article are all provided by third-party outsourcing companies and are equipped with relevant interfaces, which can be used for preliminary analysis and storage of data. At the same time, the video data collection end also provides interface services: the video secondary development memory package (WIN7/WIN1064 bit, and WIN means windows) includes: video selection, video playback, video parameter settings, etc.; it provides a web page (browser) for obtaining and playing raw videos.

(2) Management monitoring end

This end is centered around a single machine, mainly achieving real-time monitoring of images and image recognition and alarm. At the same time, these management and monitoring terminals can also be easily ported to public clouds and other service platforms, providing support for future cloud computing architectures. The interface relationship between the management supervision end and other parts: it calls the interface service of the video data collection end for video playback and acquisition; it provides real-time video monitoring and alarm data services for mobile supervision ends.

(3) Mobile monitoring terminal

Its main body is a mobile monitoring terminal based on the Win10 tablet computer. It is an extension of the management monitoring terminal and can provide users with services such as querying alarm information and basic control. The interface relationship between the mobile monitoring end and other parts: This mainly obtains real-time monitoring and alarm data from the management monitoring end, and can be set to stop alarms.

In order to adapt to the interaction between multiple endpoints and the migration of future cloud services, a network structure of an underwater swimming posture intelligent detection model was adopted [13-14]. This model constructs a unified support service layer that can achieve unified support services for desktop, mobile, and future web pages. Firstly, in order to facilitate future multi end expansion, basic services and computing would be uniformly placed in the support service layer. Secondly, computing and display are separated, with a focus on optimizing and improving the interface and display functions on each terminal (mobile phone, desktop). Ultimately, all complex calculations are handled by the backend, which allows for the full utilization of computing resources on the server and makes it easy for the server to expand and migrate.

This model adopts a unified data hierarchy to manage various types of resources uniformly. Firstly, in order to achieve unified management and service of data, a data source platform has been constructed. Secondly, the existing “swimming posture system” serves as a monitoring and management terminal (tool), and attempts to reuse the existing interface to reduce changes to the model [15]. On this basis, this article builds a mobile monitoring terminal based on supporting business. Based on a unified service interface, lightweight applications have been implemented, greatly reducing the workload and investment of construction. At one end of the application, it can be easily expanded. In the design, this article uses parallel processing to ensure that the work of each channel does not interfere with each other. When the final control computer detects that the swimming posture collection conditions are met, it can transmit the swimming posture signal to the on-site terminal in a timely manner [16-17].

3. Swimming Movement Recognition and Evaluation Collection Experiment

To classify and recognize swimming postures, it is necessary to collect a large amount of motion data [18]. This article would conduct land simulation experiments on swimming postures, thus using data collection models to collect and analyze leg, waist, and joint angle data of four common competitive swimming postures, thereby providing data support for the construction of subsequent classification and recognition models.

3.1 Collection Experiment

Based on the following considerations, the scheme of simulating leg movements in land swimming posture is adopted.

(1) On land, people can focus all their attention on their actions without being disturbed by external factors such as arm movements and water flow.

(2) The shore experiment can comprehensively observe the movement of the subjects, thereby ensuring the accuracy of the data.

(3) Some of the experimental subjects had not learned to swim. Therefore, during underwater testing, due to external conditions, the test objects cannot be conducted underwater, so the test objects are all conducted on land [19-20].

According to the needs of swimming posture classification and identification, a total of 10 participants (6 males and 4 females) were invited in this article. All invitees have signed informed consent forms, and the basic information of the invitees is shown in Table 1. The age range of the subjects is between 23 and 25 years old.

Table 1. Basic information of the inviter

Serial number	Age	Body mass index(kg/m^2)
1	24	23
2	23	21
3	25	21
4	25	19
5	23	21
6	24	21
7	24	20
8	23	23
9	23	23
10	23	20

The experimental equipment used in this article includes: sensitive pants, laptop (Window operating system), square stool, yoga mat, and portable collection device. The experimental site is located in the laboratory, and within 24 hours of the start of the experiment, all participants did not engage in any intense exercise. The participants were trained in four different swimming postures in strict accordance with the given training plan according to the order of numbering. The sampling frequency of the collection model is set to 20 Hz, and the collection time for each stroke is 20s for each candidate. After the data collection for each stroke is completed, there would be a 4-minute rest, so that after the candidate recovers their physical strength, they can continue to collect data for another stroke.

Freestyle swimming: Candidates lie on their backs on a square stool, with their legs extended downwards and suspended in the air. They bend their legs downwards and straighten their legs upwards, alternating between the two legs. The range of motion is approximately 30-40 cm.

Backstroke: The candidate sits on a square stool and leans back slightly. The student's two legs are straight together and suspended in the air. The instep is taut and the legs alternate up and down. The height of the movement fluctuates above and below the horizontal plane, with an amplitude of approximately 30-40 cm.

Breaststroke: The candidate lies on a square stool, with both hands extended forward. The candidate's legs should be folded towards their abdomen, and their heels should be as close to their thighs as possible; when flipping outward, the candidate also bends their knees slightly inward and their legs are in a "W" shape from behind; when stepping back, the candidate puts their feet together; the candidate lifted their waist and made a sliding move.

Butterfly stroke: Candidates lie prone on a yoga mat, with their elbows supporting their upper body (tools can be used to assist in supporting their upper body). The candidate's hip joint leans forward and drives the thigh to press down, allowing the knee to come into contact with the yoga mat. At the same time, candidates can also lift their feet. When the foot reaches its highest point, the candidate can extend their knees, lower legs, and feet to water. At this point, the candidate's thighs are lifted up until the knee joint is straightened. At the same time, the candidate's waist should be lifted up to complete a cycle of movements.

In each swimming posture, the cross-sectional angles of each athlete's legs and waist were collected using a portable collection device and transmitted to LabVIEW (Laboratory Virtual Instrument Engineering Workbench) through flexible circuit board wires, achieving synchronous calculation of hip and knee angles. During the experiment, the changes in the action parameters of the tested object can be observed in real-time on the display interface. LabVIEW writes each received information into a CSV (Comma-Separated Values) file and stores the collected information as the initial input for offline identification models.

3.2 Experimental Development

On the internet, instructional videos of land leg movements for four swimming styles have been searched for. Based on the information of the candidates, 5 candidates learned the leg movements of swimming posture by watching teaching videos and conducted practical training until they fully mastered the essentials of these movements. Five participants only watched 2-3 teaching videos. Ten participants wore sensing pants in sequence and connected them to a portable collection system to perform four different swimming movements according to the experimental plan.

4. Exploration Results of Swimming Posture

In the experiment, the legs of the experimental subject were suspended in mid air, and the movement space during simulated water kicking was limited. The movement amplitude of the legs was too small, and the node angle data of the experimental subject's leg sensors had little change. The sampling point when the left leg hits the lowest point is the same as the sampling point when the right leg hits the water close to the horizontal. That is to say, when the right leg hits the water upwards, the left leg hits the water downwards, which is consistent with the cross hit action of the left and right legs in freestyle. The analysis results of joint angle data in freestyle swimming are shown in Figure 1. Under small movements, there was no significant change in the angles of the left and right hip joints. The range of changes in the left knee joint is relatively large.

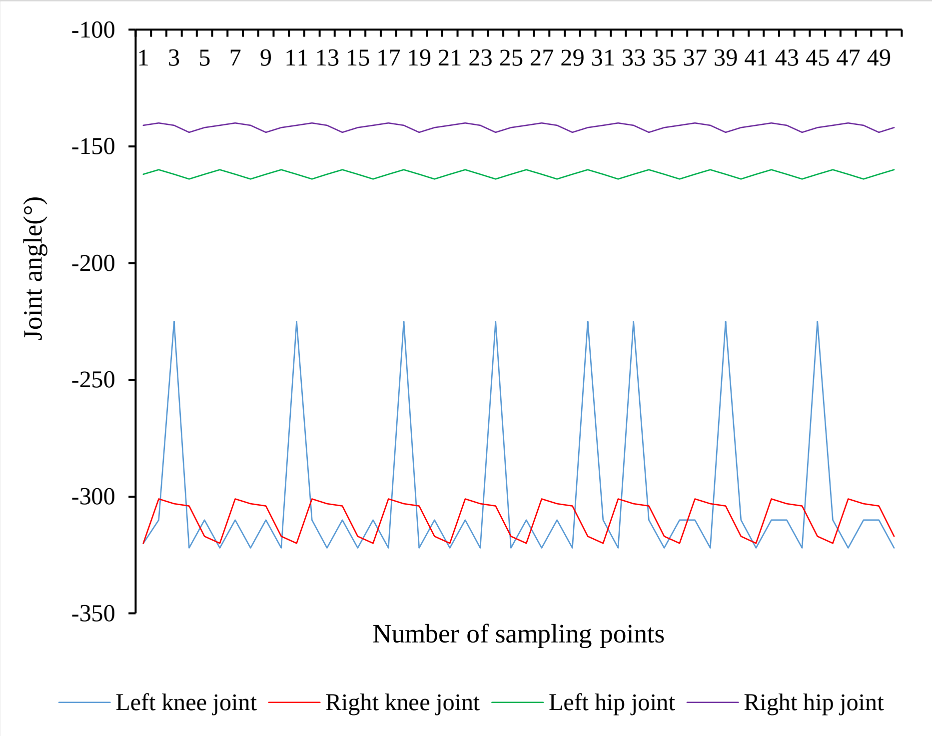

Fig.1 Analysis results of joint angle data in freestyle strokes

The sensor node angle in the backstroke stroke is shown in Figure 2. Due to the fact that the subject is in a semi bent state during the swimming process, the waist sensor collects a large amount of node angle data during the actual swimming process. During the experiment, the subjects' legs were partially suspended and their range of motion was very small. In a motion cycle, during the process of lifting water up, the thigh drives the calf up. Therefore, the sensor node angle data of the thigh and calf would change downward. After the thigh stops, the calf would continue to move up due to inertia. Therefore, the sensor node angle data of the thigh reaches the trough earlier than the calf.

During the process of diving, after the thigh stops pressing down, the calf is still pressing down, and the thigh reaches its peak earlier than the calf; The movement of the left leg of this subject is slightly greater than that of the right leg. When the peak valley values of the same leg change at the same time, and the peak valley values of both legs change at different times, the left and right feet cross kick at the same time.

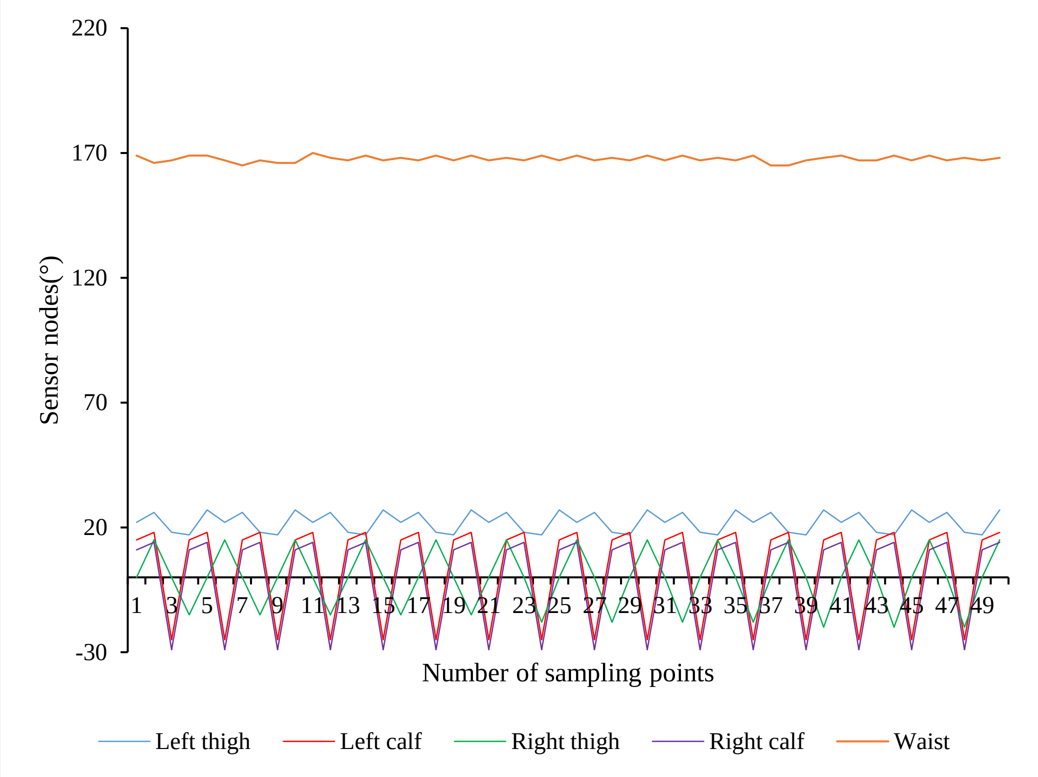

Fig.2 Sensor node angle in backstroke stroke

The pitch angle between the head and hips in butterfly stroke is shown in Figure 3, with a maximum value of 110° for the head and 42° for the hips. The athlete's hips are in a periodic state of motion, with each peak being the highest and each valley being the lowest. From the characteristics of hip data, it can be observed that the high and low peaks in the hip of butterfly stroke correspond to different pitch angle peaks caused by the two legs. The peak elevation angle of the head occurs during the period of the lowest elevation angle of the buttocks.

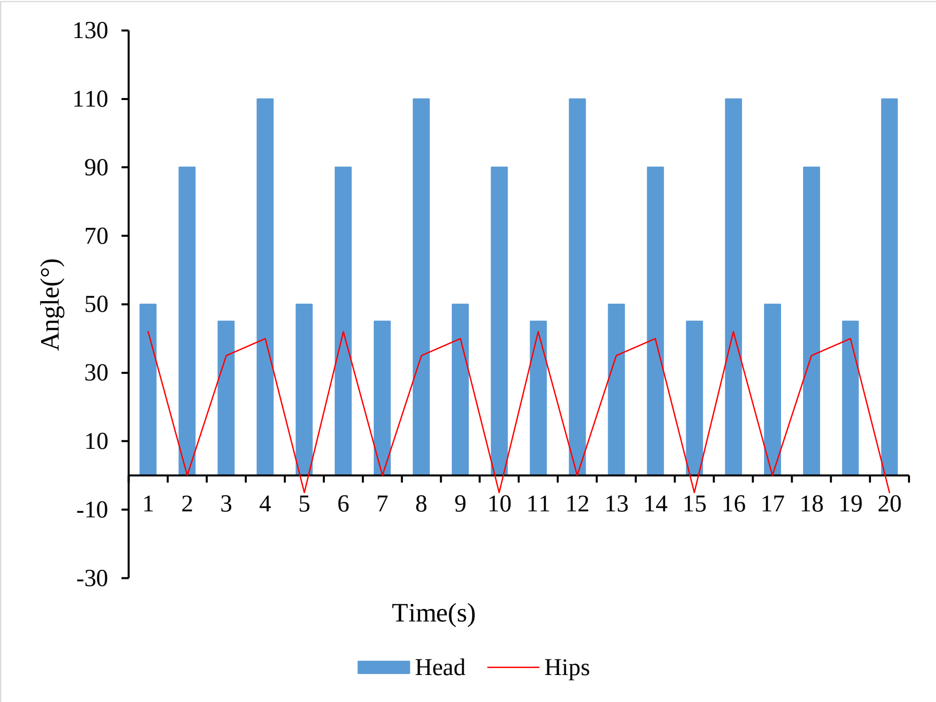

Fig.3 Head and hip pitch angle in butterfly stroke

This article defines the pitch period of butterfly stroke as the period from the lowest point of the first maximum peak to the lowest point of the second maximum peak. This article divides the head angle period of butterfly stroke into T1-T4. Through statistical analysis, it is found that the time proportions of the four stages are: 15%, 30%, 15%, and 40%. The head angle cycle of butterfly stroke is shown in Table 2. The proportion of time required to divide the hip joint pitch angle into K1-K4 and 4 hip joint pitch angles is 29%, 12%, 39%, and 20%. The pitch angle of the hip joint is shown in Table 3.

Table 2. Ratio of head angle cycles in butterfly stroke

Butterfly head angle cycle	Time proportion(%)
T1	15
T2	30
T3	15
T4	40

Table 3. Proportion of hip joint pitch angle

Hip pitch angle	Time proportion(%)
K1	29
K2	12
K3	39
K4	20

The paired sample T-test results of the peak and valley values of the head pitch angle in butterfly stroke are shown in Table 4. The peak value of head pitch angle is 100±5°, the valley value of head pitch angle is 15±6°, and the total amplitude of head pitch angle is 90±8°. The peak value of hip pitch angle is 50±7°, the valley value of hip pitch angle is 5±4°, and the total amplitude of hip pitch angle is 50±6°.

Table 4. Paired sample T-test results of peak and valley values of butterfly stroke head pitch angle

Corresponding angle	Head	Hip
Pitch angle peak(°)	100±5	50±7
Pitch valley value(°)	15±6	5±4
Total pitch angle amplitude(°)	90±8	50±6

5. Conclusions

The swimming monitoring and recognition used in this article mainly referred to the monitoring, data collection, and recognition of swimming posture through sensors. This provided convenience for swimming coaches to track and analyze athlete status, and helped swimmers monitor their movements. Compared with existing research results, the swimming posture monitoring and recognition model used in this article could collect various types of data. Based on this, a fabric like sensor network was designed and fabricated. This article effectively protected the sensors and the wires they were connected to, enabling monitoring and recognition of swimming posture. However, there was still room for improvement. The constructed collection board was located at the waist of the human body, and its hardware volume was large, which could cause discomfort to the human body. In the future, flexible circuit boards can be considered to alleviate discomfort to the human body.

References

[1] Xiao Ju. Swimming motion recognition method based on single pose sensing components [J]. Electronic Devices, 2022, 45(5):1264-1271.

[2] Xiang Hongbiao, Yang Dahu, Yang Lu, et al. Path planning and recognition tracking of magnetoelastic micro swimming robots in complex environments [J]. Journal of Mechanical Engineering, 2023, 59(5):89-99.

[3] Wang Song, Zhang Yufei, Huo Meimei. Swimming Data Analysis System Based on IMU Attitude Sensor [J]. Modern Computer (Professional Edition), 2021, 027(024):154-158.

[4] Yang Lihong. Analysis of the importance of leg movements in swimming teaching and training [J]. Education Modernization, 2018, 5(29):178-179.

[5] Mei Jiwei. Design and Strategy of Integrating Basic Swimming Actions into Theme Games for Preschool Children [J]. Journal of Shaanxi Institute of Education, 2022, 38(9):53-62.

[6] Crespi A, Karakasiliotis K, Guignard A, et al. Salamandra robotica II: an amphibious robot to study salamander-like swimming and walking gaits[J].

IEEE Transactions on Robotics, 2013, 29(2): 308-320.
[7] Cust E E, Sweeting A J, Ball K, et al. Machine and deep learning for sport-specific movement recognition: A systematic review of model development and performance[J]. Journal of sports sciences, 2019, 37(5): 568-600.
[8] Escobar,David,Simbana,et al.Functional Role of Movement and Performance Variability: Adaptation of Front Crawl Swimmers to Competitive Swimming Constraints[J].Journal of applied biomechanics, 2018, 34(1):53-64.
[9] Fan J, Wang S, Yu Q, et al. Swimming performance of the frog-inspired soft robot[J]. Soft robotics, 2020, 7(5): 615-626.
[10] Jeng C C .A Low-cost Mobile Real-time Monitoring System for Analyzing Head Position and Breathing Patterns in Front Crawl Swimming[J].Journal of Computers (Taiwan), 2021, 32(2):8-22.
[11] Li L, Xin X, Wei M, et al. Research on the action design of team artistic swimming[J]. International Journal of Sports Science and Physical Education, 2020, 5(1): 5-9.
[12] Nguyen D Q, Ho V A. Anguilliform swimming performance of an eel-inspired soft robot[J]. Soft Robotics, 2022, 9(3): 425-439.
[13] Qiu X, Veiga S, Calvo A L, et al. A kinematics comparison of different swimming relay start techniques[J]. Journal of Sports Sciences, 2021, 39(10): 1105-1113.
[14] Turdaliyevich A F, Pulatovna A B. Organization of Swimming Lessons in Preschool Institutions[J]. The american journal of social science and education innovations, 2020, 2(07): 322-330.
[15] Wang Z, Wang J, Zhao H, et al. Using wearable sensors to capture posture of the human lumbar spine in competitive swimming[J]. IEEE Transactions on Human-Machine Systems, 2019, 49(2): 194-205.
[16] Wang J, Wang Z, Gao F, et al. Swimming stroke phase segmentation based on wearable motion capture technique[J]. IEEE Transactions on Instrumentation and Measurement, 2020, 69(10): 8526-8538.
[17] Lucas K N, Lauder G V, Tytell E D. Airfoil-like mechanics generate thrust on the anterior body of swimming fishes[J]. Proceedings of the National Academy of Sciences, 2020, 117(19): 10585-10592.
[18] Kos A, Umek A. Wearable sensor devices for prevention and rehabilitation in healthcare: Swimming exercise with real-time therapist feedback[J]. IEEE internet of things journal, 2018, 6(2): 1331-1341.
[19] Rusdiana A .3D BIOMECHANICAL ANALYSIS OF SWIMMING START MOVEMENTS USING A PORTABLE SMART PLATFORM WITH ANDROID PIE[J].Journal of Engineering Science and Technology, 2021, 16(1):571-585.
[20] Mcgibbon K E , Pyne D B , Shephard M E ,et al.Contemporary practices of high-performance swimming coaches on pacing skill development and competition preparation:[J].International Journal of Sports Science & Coaching, 2020, 15(4):495-505.