RESEARCH ARTICLE

Portable and Feasible Device to Evaluate the Technique of Swimming Practitioners

ZHENGLONG FENG[1], AIFENG REN[1], (Member, IEEE), ZHIYUAN CHEN[1], ADNAN ZAHID[2], LEI YE[1], MUHAMMAD ALI IMRAN[3,4], (Senior Member, IEEE), AND QAMMER H. ABBASI[3], (Senior Member, IEEE)

[1]School of Electronic Engineering, Xidian University, Xi'an, Shaanxi 710071, China
[2]School of Engineering & Physical Sciences, Heriot Watt University, Edinburgh EH14 4AS, U.K.
[3]James Watt School of Engineering, University of Glasgow, Glasgow G12 8QQ, U.K.
[4]Artificial Intelligence Research Center (AIRC), Ajman University, Ajman, United Arab Emirates

Corresponding author: Aifeng Ren (afren@mail.xidian.edu.cn)

ABSTRACT Based on people's enthusiasm for swimming and their pursuit of swimming skill at a higher level, this paper proposes a non-contact motion guidance system based on optical motion capture technology, which aims to improve their swimming technique conveniently, scientifically and quickly. The swimming actions can be captured using one smartphone underwater, and the app on the phone will process the video and give suggestions based on the reference postures derived from statistics of professional swimmers for improvement of the swimming performance. In this system, the Canny operator was applied to detect the edge of the image captured from the swimming video, and the Hough line was implemented to recognize and locate the key position information which refers to the angle between knee and hip joints, and the angle between the trunk and horizontal plane. Furthermore, by comparing the key position information with the data of standard swimming, improvement measures can be given to improve the swimming skills of swimmers. As the result of the experiments, the system can accurately capture the swimmer's movement state, locate the incorrect posture of the swimmer, and provide the corrective information in real-time. Although the paper only took breaststroke as an example, the proposed method can be extended to the other swimming styles.

INDEX TERMS Motion capture, non-contact, canny operator, swimming, image processing.

I. INTRODUCTION

With the improvement of people's attention to health, swimming is becoming more and more popular because it is mild and not easy to be injured [1]. At the same time, swimming can exercise every muscle in the body, which is conducive to deeper movement. The pursuit of being able to swim nearly professional is further rising, but most swimmers would like to improve their body posture through self-study. Learning the swimming action on land other than in the actual swimming pool makes it easy to master the body position, and it's

The associate editor coordinating the review of this manuscript and approving it for publication was Senthil Kumar.

also tough to find the real problems in the pose. Regarding addressing the human motion capture, five significant principles have been suggested, including 2D systems employed with image processing technologies and 3D capture systems equipped with optoelectronic, electromagnetic, ultrasonic localisation, or inertial sensory technologies [2]. However, the swimming motion capture will be more challenging due to the non-stationary noisy background in the aquatic environment [3], [4]. Professionally speaking, optical and inertial types are two effective ways to realize motion capture for body position analysis and recognition [5], [6], [7]. Optical motion capture (OMC) includes marker-based optical mode and markerless optical mode, and inertial motion

capture (IMC) belongs to non-optical mode. The optical marker means that multiple cameras can photograph the object with tracking points. Subsequently, according to the corresponding relationship between 3D coordinates and 2D coordinates in the camera, the object's motion trajectory is obtained by an algorithm [8]. For example, In the production of science fiction films and TV works such as Avatar, optical markers often play an important role in capturing the actor's movement [9]. Accordingly, markers would be placed on the actors, and the desired shape would be calculated through algorithms. Meanwhile, because of the requirements of markers to mark the joints of athletes, it will increase the complexity of the system application and reduce the performance of the user experience. Nevertheless, the non-optical marker mode uses multiple high-speed cameras to monitor and track the target feature points from different angles to capture the action [10], [11]. Devices with artificial intelligence would be an example of the non-optical marker mode. The built-in algorithm would calculate the motion based on the identified feature points and then output the motion tracking. The Inertial motion capture method is to wear the integrated accelerometer, gyroscope, magnetometer, and other inertial sensor equipment at the crucial nodes of the body and realize the capture action through the algorithm [12]. This method has been used in vehicle stability control, VR training and simulation, medical and other scenes, and is more widely used in different sports. In swimming, the technical details of athletes could be accurately recorded by placing sensors on all parts of their bodies [13]. In addition, interial measurement units (IMU) have been validated for analyzing and assessing the joint kinematics and coordination dynamics of the human aquatic behavior during swimming [14], [15].

Considering the sensory characteristics of the inertial motion capture system in swimming activity, three-dimensional human motion trajectory can be built through the collected data, and more accurate motion suggestions can be provided to the swimmers. Meanwhile, this method is less affected by the environment or light. Nevertheless, a reference sensor node, and at least two sensor units, has to be placed on the wearable biofeedback suit in order to track joint angles [16], the inertial motion capture system can affect the swimmer's movements due to the wearing of the sensors and fixed devices on the body, which will make the measurement inaccurate. Furthermore, the sensors worn on the body are susceptible to the water flow during the swimming activity, leading to displacement and affecting the data collection. In terms of wearing position of wearable devices, the position of the sensor has a direct impact on motion measurement, and the optimal position of the sensor depends on the user's acceptance and the accuracy of motion classification [17]. For a more accurate measurement, the device needs to be close enough to the measurement location. And because of the need to wear sensors, the system is more suitable for professional use. For ordinary recreational swimmers, these applications are complicated and not user-friendly. So it is hard to be

extended to a broader range of swimming enthusiasts to improve their swimming skills.

In the optical marker motion capture system, the special markers attached to the subject are implemented for data acquisition, including the passive markers and active markers. The advantage of the optical marker motion capture system is high capture frame rates and measurement resolution. At present, dozens of frames of video can be captured by recording equipment, and the information provided is enough to calculate the user's actions. Whereas, if the optical marker motion capture system is used underwater in the swimming scene, being the light is too dark, it is not conducive to the distinction between the body and water. In addition, the marked points are easily occluded and mismatched, leading to a significant deviation of the calculated information. One markerless optical motion capture system for swimming motion has been proposed in [18], which was focused on the butterfly stroke and the variable thresholding was applied to obtain the participant's silhouettes from the segmented swimming images.

In this paper, one smartphone-based markerless optical motion capture system is proposed to guide swimmers to improve their swimming skill, with low cost and high flexibility, which can be more efficient to capture the swimming movement accurately. A portable mobile phone is used for underwater shooting, and then processed by an algorithm to feed back the swimming posture information to the user. This method is not only convenient to use, but also has a low threshold for users to use, and does not require certain technical knowledge. The rest of the paper is organized as follows: Section II presents the method and theory applied in the motion capture system. Section III describes the results and discussion on the findings, while Section IV concludes the paper.

II. MATERIALS AND METHODS
A. MOTION CAPTURE SYSTEM
In this markerless and non-contact system, a water-proof smartphone is placed about 15cm below the water surface because the position of the human body during swimming is generally about 15cm below the water surface, where the optical motion capture is not easily affected by the fluctuation of the water surface. The placement position is 15cm below the water surface, and the observer is about 2m away from the smartphone so that body parts 2m below the water surface can be comprehensively observed as shown in Figure 1. The APP will detect the images in the video every 0.1 seconds after a complete swimming cycle is performed by the observer with the smartphone. The most important feature of this method is that it is portable and convenient. The popularization of smartphones is more conducive to the promotion and application of this system. Furthermore, the proposed method lowers the user's barrier to entry without wearing redundant devices, which needs to employ smartphones to complete the swimming video recording. Meanwhile, the algorithms integrated

into smartphones can distinguish all parts of the body in a short time, and the information for swimmers extracted from the swimming video is sufficient and valuable. While the critical position of the human body is identified by the algorithm and the required objective information is obtained, it can be compared with the reference angle range for analysis and assessment. Subsequently, the scores are given to positions with different weights based on the deviation between the user's motion information and reference motion information. Finally, the improvement suggestions will be proposed for movements with great differences accordingly. In this regard, we closely monitored and analyzed the movements of 8 male professional swimmers, including seven national first-class athletes and one second-class athlete (age 21.8±2.6 years, height 1.886±0.6 meters, weight 76.02±2.32 kg, best individual swimming time over 200m breaststroke long with 145.60±2.04s; data format with mean±standard deviation(SD)) to statistically extract the reference angles of each important part of the breaststroke as standard angle knowledge [19]. All participants were informed about the research procedure and signed a consent in advance.

Subsequently, All measurements were performed in the university swimming pool with water temperature of approximately 28.5°C. One smartphone equipped with EMUI based on the Android 10 source code and three rear cameras, including one 40 MP ultrawide lens main camera, one 8MP telephoto lens with 3x optical zoom, and one 16 MP ultrawide lens camera, was applied to capture the swimming videos in a waterproof bag. One entire breaststroke cycle can be divided into three different stages according to the time process, which includes stroke, kick, and glide. In the experiment, we captured at least five complete swimming motion cycles of 8 athletes and denoised the data collected for different posture features through mean and median filtering algorithm processing for reference. In practice, this system should guarantee to collect at least three complete action cycles of swimmers.

B. SELECTION OF KEY INFORMATION

Breaststroke is the most technically demanding and slowest of the four swimming styles, but its stroke is more complicated than other swimming movements. The ordination and symmetry of the movements are focus and difficulty in the learning and training of breaststroke, which involves the harmonization of the arms, legs, and body movements during the underwater stroke. The key factors affecting the swimming speed are the kinematic information as joint angles and displacement of the participant. The propulsive force of breaststroke mainly comes from kicking, so three indexes are observed primarily in this paper: the angle between the trunk and horizontal plane, the angle between thigh and trunk, and the angle between thigh and calf. The purpose of observing the angle between the trunk and horizontal plane is to confirm whether the arm and leg are moving in a coordinated manner. And the measurement of the thigh and the trunk is to judge whether the thigh is too close to the belly with legs together.

Meanwhile, the angle measured between thigh and calf is applied to determine whether the subject's legs are in place. These three indicators largely determine whether the user's breaststroke is standard or not.

C. VIDEO PROCESSING

After the user's swimming action is shot, images in the video are preprocessed and enhanced every 0.1 second. The human body is then detected, after which the key angle is calculated. A comprehensive score and targeted optimization suggestions will be given to the user after the key angle is compared with the standard angle range. Processing flow is shown in Figure 2.

Similar to LOG (Laplacian of Gaussian) operator, canny operator is a method by smoothing and derivative [20]. The principle can be divided into the following four parts: Gaussian Filtering, calculate gradient amplitude and direction, double threshold method for edge closure, and detection of straight line by Hough transform.

1) GAUSSIAN FILTERING

The filtering is to weighted average the gray values of each pixel and its neighboring points according to certain parametric rules in order to effectively filter the high-frequency noise superimposed in the image [21].

Firstly, the 2D Gaussian filter template is used for convolution to eliminate the noises. Among all filtering methods, how to balance the conflict between de-noising and the accuracy of edge detection is the most important issue to be considered. The practical engineering experience indicates that the kernel determined by Gaussian function can provide a better compromise scheme [22].

The discrete window sliding convolution we selected is mainly realized by Gaussian kernel, which is an odd size Gaussian template [23]. There are two commonly used Gaussian kernel templates, which are 3 * 3 and 5 * 5, as shown in Figure 3.

The Gauss partial derivative is as:

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \tag{1}$$

The parameters are calculated by Gaussian function, which represents the distance between the pixel and the center pixel, and σ in the formula represents the standard deviation. If σ in the formula is too large, it will deepen the filtering degree, thus blurring the edge of the image. Furthermore, it is not conducive to edge detection in the next step. If the Sigma in the formula is too small, the filtering effect is not as good as expected [24].

In conclusion, σ should be set appropriately for targeted scenarios. Normalization is needed for calculating the parameters of Gaussian template because the central pixel of template calculated by convolution is limited to the gray range of 0 to 255 after normalization. If the gray value of all pixels within a neighborhood is 255, the gray value of the central pixel of the template is still 255 after the template is used
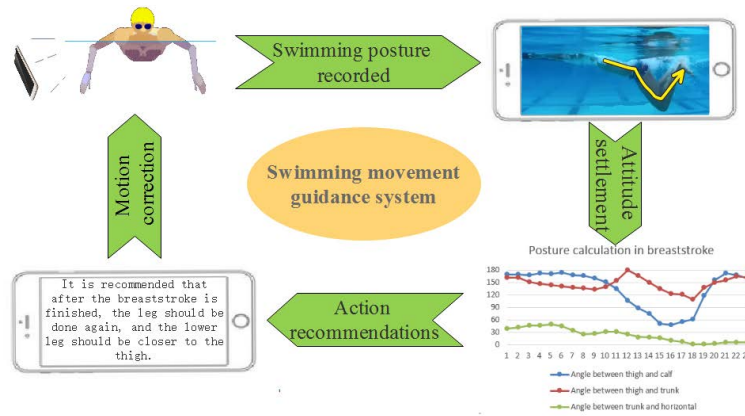
Swimming posture recorded

Motion correction

Swimming movement guidance system

Attitude settlement

It is recommended that after the breaststroke is finished, the leg should be done again, and the lower leg should be closer to the thigh.

Action recommendations

Posture calculation in breaststroke

Angle between thigh and calf
Angle between thigh and trunk
Angle between trunk and horizontal

FIGURE 1. Experimental system setup.

Swimming video

Video preprocessing

Edge detection and extraction

Trunk detection

Edge detection and extraction

Characteristics of Hough transform

Extraction of body information

Motion information of subject

Professional athlete or not

Y

N

Widen standard set

Comparison with standard angle range

Suggestions and comprehensive score

FIGURE 2. Video processing flow.

$$\frac{1}{16} \times \begin{array}{|c|c|c|} \hline 1 & 2 & 1 \\ \hline 2 & 4 & 2 \\ \hline 1 & 2 & 1 \\ \hline \end{array} \qquad \frac{1}{273} \times \begin{array}{|c|c|c|c|c|} \hline 1 & 4 & 7 & 4 & 1 \\ \hline 4 & 16 & 26 & 16 & 4 \\ \hline 7 & 26 & 41 & 26 & 7 \\ \hline 4 & 16 & 26 & 16 & 4 \\ \hline 1 & 4 & 7 & 4 & 1 \\ \hline \end{array}$$
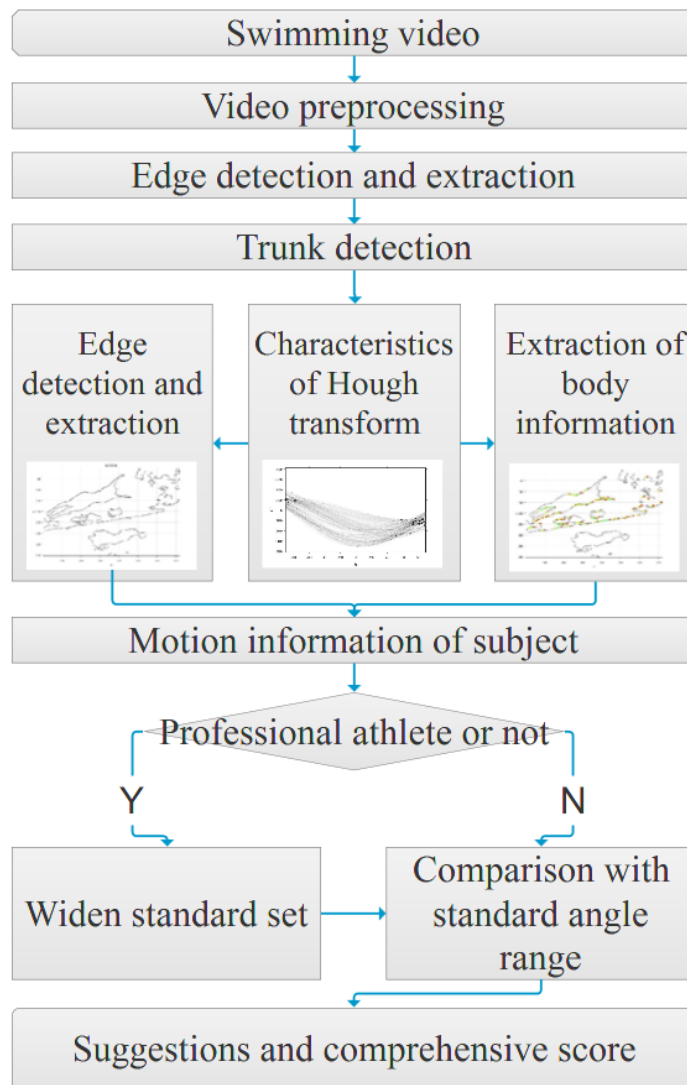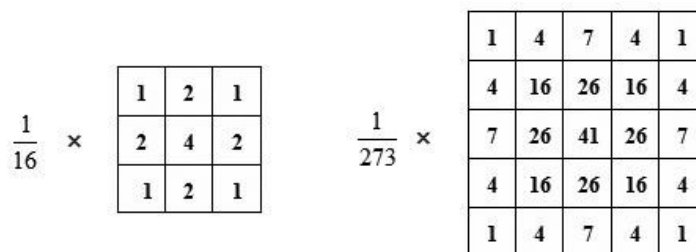
FIGURE 3. Gaussian kernel templates: 3 * 3 and 5 * 5.

for convolution; if the sum of the parameters of the Gaussian template calculated is less than 1, the gray value of the central pixel of the template will be less than 255 after the template for convolution is used, thus deviating from the actual gray value and further resulting in errors [25].

2) CALCULATE GRADIENT AMPLITUDE AND DIRECTION

The derivative operator is used to find the derivative of the gray areas of the image. Hence, the gradient amplitude is calculated by

$$|G| = \sqrt{G_x^2 + G_y^2} \tag{2}$$

Then, the calculation results of above formula are further used to calculate the direction of the gradient as shown in

$$\theta = Arc \tan\left(\frac{G_y}{G_x}\right) \tag{3}$$

The convolution operators used by Canny operator are as follows in equation (4). The gradient amplitude and gradient direction are calculated as follows, where P and Q respectively represent the first-order partial derivative matrixes in x direction and in y direction, M represents the gradient amplitude and θ represents the gradient direction.

$$P_{ij} = \frac{f_{i,j+1} - f_{ij} + f_{i+1,j+1} - f_{i+1,j}}{2}$$
$$Q_{ij} = \frac{f_{ij} - f_{i+1,j} + f_{i,j+1} - f_{i+1,j+1}}{2}$$
$$M[i,j] = \sqrt{P[i,j]^2 + Q[i,j]^2}$$
$$\theta[i,j] = \arctan(Q[i,j]/p[i,j]) \tag{4}$$

Once the direction of the edge is known, the gradient direction of the edge can be roughly divided into four kinds: horizontal, vertical, 45-degree direction, and 135-degree direction [26].

The adjacent pixels can be detected in the gradient direction of the pixel. If the gray value of a pixel is not larger than those of two pixels in the gradient direction, the pixel value is set to 0 and it is not an edge.

3) DOUBLE THRESHOLD METHOD FOR EDGE CLOSURE

In the matrix of gradient amplitude, the element with a large value represents great gradient, but it may not be an edge pixel [27]. Therefore, it is necessary to suppress non-maximum value, which is to find the local maximum value of pixel point, and set the gray value of non-maximum value point to 0 so that most non-edge points can be eliminated.

Some of the edges obtained are false with fractures. If an edge image is obtained according to the high threshold value, such an image contains few false edges [28]. However, the image edge may not be closed because the threshold value is high. Under such circumstance, a low threshold value is needed. When reaching the end point of the contour, the point satisfying the low threshold value can be found in the 8 neighborhood points of the breakpoint, and then the edge of the whole image is closed according to the newly collected edge.

As a basic link of image processing, skin detection provides more easily obtained information for subsequent processing. In order to apply the results of skin detection to the real-time image processing system, the skin area in the image has to be quickly and efficiently detected. And it is also important to choose the color space featuring cognitive attribute and strong ability of separately expressing brightness and chroma. Furthermore, a skin model with efficient classification should be selected.

At present, most methods for detecting skin color are mainly based on the statistical model of skin color [29]. The key steps are to select color space and establish a skin color mode. The color spaces proposed based on different application requirements exhibit different application characteristics. Selection of a proper color space from more than ten common color spaces can have a profound effect on the skin detection performance [30]. RGB tricolor space is hardware-oriented, whose major defects include brightness and chrominance mixing and device correlation. Normalized RGB (nrgb) space irreversibly loses information about brightness of RGB space. The perceived color spaces such as HSV and HIS is color-oriented and represented by three attributes of color. But currently, there is no definition about the color on its brightness axis. The color spaces such as YIQ and YES are linearly correlated with RGB, and these color spaces are mainly applied in the television industry.

The commonly used methods of skin color modeling include skin color range definition, non- parametric model and parametric model. The range of skin color can be simply,

quickly and easily defined. Meanwhile, the roughly estimated distribution of color skin is more accurate because the threshold value is based on analysis results and features statistical characteristics. But the difficulty of this method is to choose the appropriate color space and threshold, which is largely dependent on the statistical samples. The nonparametric skin color model is characterized by independent color space and convenient calculation, but it requires a large number of training samples, so fitting is likely to occur. Moreover, a unified standard is still not available for selecting the samples with a proper size. However, the parametric skin color model does not require a large number of training samples, which is only a fitting based on the distribution of skin color in color space rather than in its true sense [31], so the detection results using such methods are not often that accurate.

Regarding the color space, which is widely used in human detection, this paper chooses four typical color spaces: HSV, YIQ, YCbCr, and TSL as research objects. In order to select a better color space as soon as possible, this paper applies the chroma space model, which is a model that clearly defines color range [32]. The distribution of skin color can be greatly affected by ambient light because images are often shot in different scenarios. The skin color in an image can be even different for the same person because of different light sources, so a robust detection algorithm of skin color is required to detect the pixels of skin color in different environments. However, it is very difficult to precisely calculate the illumination parameters because of the difficulty in obtaining the parameters of the camera parameters and the spectra of light sources. Therefore, a simple and effective adaptive skin color model is applied to detect the pixels of skin color under different illumination conditions. The distribution of skin color in an image is determined by two factors, which are the intensity of ambient light and the natural color of human skin. The specific sources of color light will not be further discussed because a vast majority of natural images are not captured from specific sources of color light. In order to adapt to different skin colors and lighting conditions, this paper uses the Gauss skin color model as in equation (5), which is based on color space.

$$P(skin|c) = \frac{1}{2\pi |\Sigma s|^{1/2}} e^{-\frac{1}{2}(c-\mu_s)^T \Sigma_s^{-1}(c-\mu_s)} \quad (5)$$

where c is the color vector corresponding to the pixel. μ_s and Σs are two distribution parameters, in which μ_s is the mean vector, and Σs is the covariance matrix.

After the image is smoothed, each pixel in the skin region is transformed through color space to get (Cb, Cr) value, and then the skin color region is detected by the Gaussian skin color model. As the skin area is brighter than other parts of the image, it can be segmented from the image by thresholding.

4) DETECTION OF STRAIGHT LINE BY HOUGH TRANSFORM

Considering the key information extraction from the discrete frames of the camera sensors, the edge detection is simplified to locate the joints of the body by transforming the irregular edge into interconnected straight lines with nodes. Hough transform is a global method of detecting lines, which can be applied to implement the transformation between the image space of the frames and the parameterized Hough space [33].

Once the coordinates of all the points on the line segment are obtained, the angle of the corresponding body part can be calculated by inverse Hough transformation [34].

D. ALGORITHM OF THE SCORING SYSTEM

In order to enable users to have a clearer understanding of their own action standards and a more quantitative perception of their own improvement in action standards, the scoring system has been added. The system establishes a quantitative evaluation algorithm to give users a clearer understanding of their swimming skills and a more quantitative perception of their improvement in the action standards. In this case, through the analysis of captured swimming video and the main angle information that the system focuses on, the following quantization algorithm can be used as a reference for evaluating the performance of swimming movements [35], [36].

There are two important indexes in calculating the comprehensive score of the user's swimming action. To be specific, they are the proportion of the part exceeding the standard included angle in the whole cycle and in the standard data respectively. As nonstandard actions are often associated, such as, the extremely small angle between the thigh and the trunk or between the trunk and the horizontal plane, the former index is more important than the latter. Here, we define the weight of the first index as 60%, and that of the second index as 40%. Therefore, the formula of the comprehensive score (CS) is shown as in equation (6).

$$CS = FI \times 60\% + SI \times 40\% \quad (6)$$

where CS represents the comprehensive score. FI refers to the first indicator, and SI indicates the second indicator.

The proportion of the part exceeding the standard included angle in the whole period refers to that of the part exceeding or below the standard included angle in the whole period the formula is as (7).

$$FI = \sum_{n=1}^{3} \frac{S_n}{S} \times Q_n \quad (7)$$

in which S_n represents the time beyond the standard angle range; S means a complete stroke cycle, and Q_n refers to the weight without position information.

The weight of the specific key information varies according to the different swimming movements. In the breaststroke, the angle between the thigh and the trunk is the key to the correct rhythm of the whole breaststroke movement. The correct action will make it easier to keep the data of the thigh, the calf and the trunk at the water level correct. Therefore, the weight of specific key information is defined as the angle between the thigh and the trunk, which is 50%; the weight between the thigh and the calf is 25%, and that between the trunk
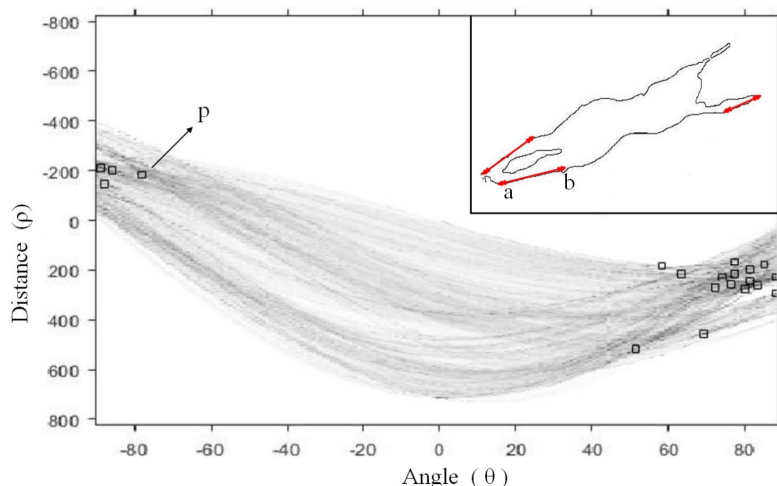
FIGURE 4. Represents the feature points.

and the horizontal plane is 25%. Therefore, in breaststroke, the formula of the first index is as in (8).

$$FI = \frac{S_1}{S} \times 50\% + \frac{S_2}{S} \times 25\% + \frac{S_3}{S} \times 25\% \qquad (8)$$

where S_1 shows the angle between the thigh and the trunk; S_2 represents the angle between the trunk and the horizontal plane, while S_3 refers to the angle between the thigh and the calf.

The ratio of the excess of the part exceeding the standard included angle range to the standard data refers to that of the excess concerning the extreme value exceeding the standard range to the standard value. Here, the standard value is the middle value of the included angle range which is as (9).

$$SI = \sum_{n=1}^{3} \frac{A_n - A}{A} \times Q_n \qquad (9)$$

where A_n is the extreme value exceeding the standard range, and A is the middle value of the standard included angle range.

III. RESULTS AND DISCUSSION

From the swimming video recording of the smartphone, the integrated algorithms can process the frames for the appropriate posture recognition and motion detection of human activity under the water. After the Hough Transform, theoretically, any point in the image such as point a and point b shown as in the embedded subfigure in Figure 4, can correspond to the lines in an infinite number of directions in a parameter space of transformation. All lines passing a point in the original image space will become a sine curve in the transformed parameter space. The sine curves transformed from the frames of the swimming video into Hough space are shown in Figure 4.

Subsequently, the intersection points of multiple lines such as point p shown as in Figure 4, can be identi-

fied in parameter space, which correspond to the different line segments in image space such as a-b in embedded figure in Figure 4. In the actual application, in order to reduce the computation, the limited number of directions of each pixel in one video frame will be considered to search its adjacent points, such as -45 degrees, 0 degrees, 45 degrees, 90 degrees and their opposite directions. Therefore, when all the pixels in the image space have been mapped to the parameter space, the accumulated counters of all intersection points, for example the point p shown in Figure 4, can calculate the number of intersecting curves at each point. According to the predetermined maximum intersection point threshold, some intersection points, which provide the same parameters of distance ρ and angle θ in the Hough space, are selected to determine the line segments in the image space, which can be applied to extract the key positions of swimming movements for the further video processing and shown as in Figure 5 [37].

In a full swimming cycle, the upper and lower limit of the standard angle is taken as the extreme angle value for the main sections of 8 athletes, and the standard angle curves are obtained shown in Figure 6. The advantage of this method is that the more athletes are observed, the more stable the angle of the statistical part of the action becomes. Thus, the mean range of the angles obtained can be regarded as the standard angle range of this posture for reference value. Concurrently, in this way, the inaccessible behaviour of athletes can be prevented which leads to the inaccuracy of the standard angle range. Our observations for this standard angle are only the preliminary version. In the future, the movement of more professional athletes will be recorded and analyzed their actions to make the standard angle range more authoritative.

Upon close analysis of Figure 6, during the stroke stage in the breaststroke event, the arm drives the upper body to rise or fall from the water while the legs remain still, which makes
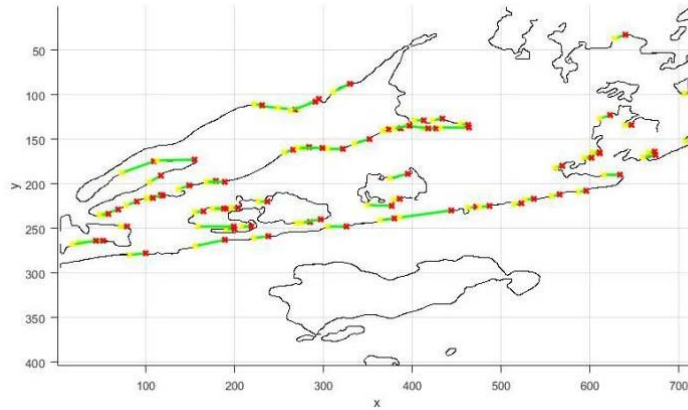
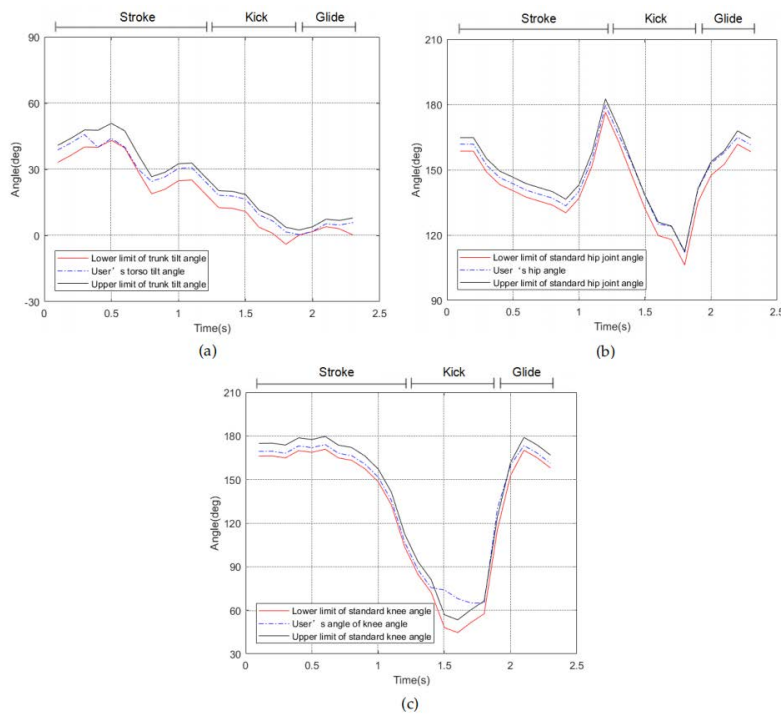**FIGURE 5.** The result of the detection of the line.

FIGURE 6. Angles detected from professional athletes and recreational swimmers:
(a) Angle between trunk and horizontal plane, (b) Angle between the thigh and the torso
and (c) Angle between the thigh and the calf.

the angle between the trunk and the horizontal plane larger first and then smaller. Due to the immobility of the legs, the angle between the thigh and the trunk, and the angle between the thigh and the calf is almost unchanged. In the kick stage, the legs will be brought up and pushed the water back while the arm keeps still. Meanwhile, it is mainly the heel close to the buttock and the calf close to the thigh, which results in the change of the angle between the leg and the thigh first reduced and then increased. At this stage, the angle between the thigh and the trunk will be also changed due to the action of the calf, but the change speed is lower than

that of the angle between the thigh and the calf. And there is nearly no change of the angle between the trunk and the horizontal plane. During the gliding stage, the arm and leg of the swimmer keep still, so there is only a little change between angles.

Figure 6(a) shows the upper and lower limit of the standard angle between the torso and the horizontal plane in one complete breaststroke cycle as solid lines. Likewise, in Figure 6(b), the solid lines denote the upper and lower limit of the standard angle between the thigh and the torso. Figure 6(c) illustrates the upper and lower limit of the
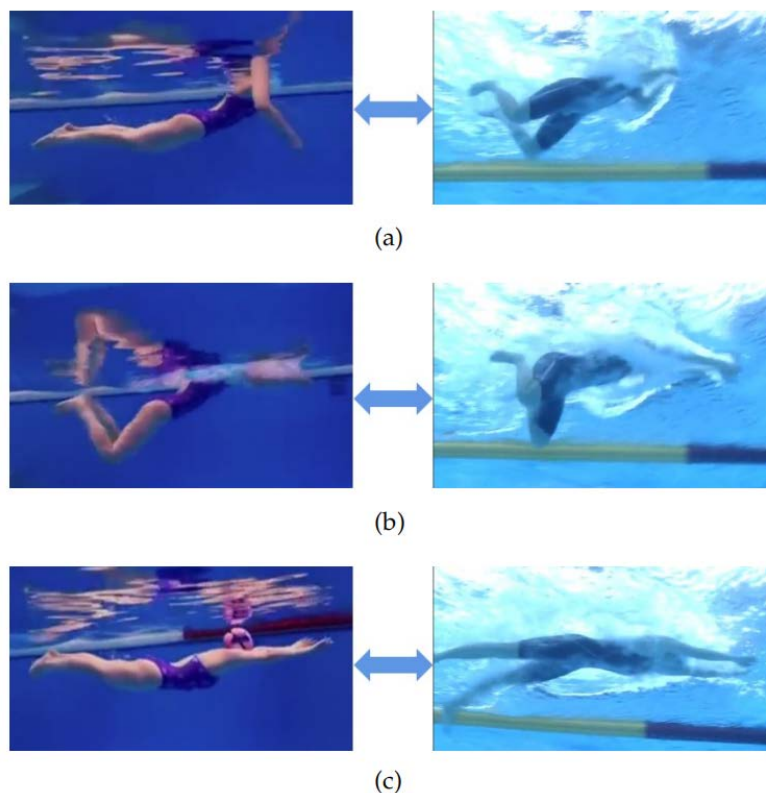
FIGURE 7. The movement display of breaststroke from different viewing angles in three stages (The three views in the left column are shot from the 15cm underwater, and the three views in the right column are shot from the elevation angle of 45 degrees underwater): (a) Arm movement in stroke stage, (b) Leg movement in kick stage and (c) Floating movement in glide stage.

standard angle between the thigh and the calf as solid lines, respectively. By contrast, the dashed lines shown as in Figure 6(a) to (c) were the results of the same positions on the amateur swimmer.

From Figure 6, it is analyzed that breaststroke cycle remains up to 2 seconds. As the cycle drops from 0 to 1.2 seconds, the arm action stage is identified. In addition, the leg action is observed as it approaches within the range of 1.2 to 1.8 seconds, the leg action is recognized. Eventually, it indicates the floating stage if it falls in the range of 1.2 to 1.8 seconds.

Furthermore, by comparing the amateur data with the standard data, the comprehensive score is 72 points calculated from equation (6). The comprehensive score indicates the quantitative result between the joint angles of each body and the reference standard values, making it easier for ordinary swimmers to understand quantitatively how far their swimming actions are from the desired action. The abnormality of the data allows us to realize that the user's thigh movement amplitude is too large when his calf moves close to his thigh, which causes the user's cross-section to be too large in the process of moving forward, which brings more resistance [38]. Consequently, the more strength the user uses, but the slower his swimming speed will be. Therefore, we can draw a conclusion that the user should pay attention

to keeping the thigh as motionless as possible when the calf is close to the thigh. Eventually, the users are advised to follow the suggestions to improve their swimming movement. First of all, in the stroking stage, the trunk tilt angle starts to drop too early, which will lead to lower stroking efficiency. Then, it is recommended to lengthen the stroke [36]. Secondly, the hip joint angle is too large during the kick. It is recommended that when the leg is in the recovery phase, the thigh should be folded up properly. At the same time, the knee joint angle is too large as well. It is recommended to keep your heels as close to your hips as possible [38].

Figure 6(a) to (c) demonstrate the comparison of the results extracted from three different stages of the complete breaststroke cycle from the amateur and professional swimming athletes. The data of the curves were calculated from the swimming video captured by smartphone in Figure 1, which includes stroke, kick, and glide three stages. In contrast to the actual action, In Figure 7(a) to (c), the movement of breaststroke from different viewing angles in three stages are displayed only for showing the standard movements of the hand and leg of the athlete corresponding to three stages of curves shown in Figure 6. By close investigations of the comparison with standard curves and coordination with standard movements of the opening and closing degree of the hand and leg of the athlete, the swimmer can get more

accurate guidance on different stages through this proposed system.

The whole system uses an external device, such as a smartphone, which does not need to be attached to the subject, to collect swimming movement information, and uses video image information to analyze the swimmer's posture, accordingly, to inform the swimmer's posture and the gap with the professionals, and practice to correct the posture. For swimmers, the more standard the trainer's posture, the greater the acceleration, the faster the speed, the better the performance the more ideal.

During the whole experiment, it was found that the location of the mobile phone is an important factor affecting the accuracy of motion capture. Make sure your smartphone's camera is as parallel as possible to the swimmer when capturing video to reduce the impact of systematic errors. As this research progresses, we will consider setting motion trajectories to keep the cameras of professional athletes relatively still, further improving the accuracy of standard data. At the same time, we will introduce video stabilization technology to correct different variables to improve the accuracy of extracting information from different strokes.

IV. CONCLUSION

In this study, an optical non-marking motion capture system is proposed to assist swimmers to enhance their swimming technology, with low cost and high flexibility technique, which can be more reliable to adequately capture the swimming movement of the swimmer. In this regard, canny operator technique was applied to detect the edge of the image obtained from the swimming video, which was captured using the smartphone under the water. Intriguingly, it was found that the detection image using technique proven to be more effective compared to LOG operator. Furthermore, the Hough line was employed to distinguish and identify the key location details relating to the angle between knee joint, hip joint and the angle between the trunk and horizontal plane In addition, by comparing the key position information with the data of standard swimming results, an optimized measures were established that can enable the swimmers to improve the swimming skills by identifying and locating the swimmer's inaccurate postures and offering the corrective information in real-time. In conclusion, the motion capture system proposed in this paper, which is based on the actual needs of swimmers, can be used conveniently to accurately improve the body position in swimming for further achieving the effect of fitness. Because the system focuses on the swimmer's posture and coordination underwater, it mainly provides improvement suggestions for users according to the joint angles of the human body that reflect this information. In addition, it can also provide information such as overall swimming speed and acceleration after part of the action. The proposed system's limitations include the sensitivity to the brightness of light from the underwater environment and waves generated by swimming movements, which may reduce the recognition accuracy of the joint angles. Through further

experimentation, this can be mitigated by adjusting the distance between the smartphone camera and the swimmer to keep the camera steady relative to the swimmer. To conclude, the system based on the already ubiquitous smartphone as motion capture devices, is very suitable for a wide range of swimmers to improve their swimming skills, including professional athletes and coaches. Swimming enthusiasts can learn the postures and coordination of their movements through the device and compare them with the reference data of professional athletes. More importantly, the system provides corresponding improvement suggestions for each action according to the differences related to the reference. Athletes and coaches can establish their own sports database through this system, further refine the postures according to their characteristics, and guide trainees to improve swimming skills.

This paper demonstrates the potential and offers a high flexibility to swimmers by not wearing any hefty equipment, and only requires a smartphone to capture his imperative actions. The proposed algorithm applied in the study can clearly distinguish the parts of body in a very short time and can provide valuable and useful knowledge to swimmers.

REFERENCES

[1] C. Perez-Diaz, J. C. De la Cruz-Marquez, N. Rico-Castro, and B. Cueto-Martín, "Challenges and needs of national swimming federations in the promotion of health," *J. Sports Sci.*, vol. 4, no. 6, pp. 353–360, Dec. 2016.

[2] E. van der Kruk and M. M. Reijne, "Accuracy of human motion capture systems for sport applications; state-of-the-art review," *Eur. J. Sport Sci.*, vol. 18, no. 6, pp. 806–819, Jul. 2018.

[3] K.-L. Chan, "Detection of swimmer based on joint utilization of motion and intensity information," in *Proc. 12th IAPR Conf. Mach. Vis. Appl. (MVA)*, 2011, pp. 450–453.

[4] W. Zh, X. Shi, J. Wang, F. Gao, J. Li, M. Guo, H. Zhao, and S. Qiu, "Swimming motion analysis and posture recognition based on wearable inertial sensors," in *Proc. IEEE Int. Conf. Syst., Man Cybern. (SMC)*, Oct. 2019, pp. 3371–3376.

[5] P.-G. Jung, G. Lim, and K. Kong, "A mobile motion capture system based on inertial sensors and smart shoes," in *Proc. IEEE Int. Conf. Robot. Autom.*, May 2013, pp. 692–697.

[6] H. Wang and S. Nguang, "Intelligent and comprehensive monitoring system for swimming pool," *Int. J. Sensors Wireless Commun. Control*, vol. 3, no. 2, pp. 85–94, May 2014.

[7] T. Schubert, K. Eggensperger, A. Gkogkidis, F. Hutter, T. Ball, and W. Burgard, "Automatic bone parameter estimation for skeleton tracking in optical motion capture," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2016, pp. 5548–5554.

[8] Y. Lee and H. Yoo, "Low-cost 3D motion capture system using passive optical markers and monocular vision," *Optik*, vol. 130, pp. 1397–1407, Feb. 2017.

[9] A. Tahir, J. Ahmad, S. A. Shah, G. Morison, D. A. Skelton, H. Larijani, Q. H. Abbasi, M. A. Imran, and R. M. Gibson, "WiFreeze: Multiresolution scalograms for freezing of gait detection in Parkinson's leveraging 5G spectrum with deep learning," *Electronics*, vol. 8, no. 12, p. 1433, Dec. 2019.

[10] F. Fioranelli, J. Le Kernec, and S. A. Shah, "Radar for health care: Recognizing human activities and monitoring vital signs," *IEEE Potentials*, vol. 38, no. 4, pp. 16–23, Jul. 2019.

[11] H. Zhang, L. Wang, S. Chu, S. Chen, H. Meng, and G. Liu, "Application of optical motion capture technology in power safety entitative simulation training system," *Opt. Photon. J.*, vol. 6, no. 8, pp. 155–163, 2016.

[12] D. Haider, A. Ren, D. Fan, N. Zhao, X. Yang, S. A. Shah, F. Hu, and Q. H. Abbasi, "An efficient monitoring of eclamptic seizures in wireless sensors networks," *Comput. Electr. Eng.*, vol. 75, pp. 16–30, May 2019.

[13] J. Lecoutere and R. Puers, "Wireless communication with miniaturized sensor devices in swimming," *Proc. Eng.*, vol. 72, pp. 398–403, Jan. 2014.

[14] S. Fantozzi, A. Giovanardi, F. A. Magalhães, R. Di Michele, M. Cortesi, and G. Gatta, "Assessment of three-dimensional joint kinematics of the upper limb during simulated swimming using wearable inertial-magnetic measurement units," *J. Sports Sci.*, vol. 34, no. 11, pp. 1073–1080, Jun. 2016.

[15] B. Guignard, A. Rouard, D. Chollet, and L. Seifert, "Behavioral dynamics in swimming: The appropriate use of inertial measurement units," *Frontiers Psychol.*, vol. 8, p. 383, Mar. 2017.

[16] S. Qiu, H. Zhao, N. Jiang, Z. Wang, L. Liu, Y. An, H. Zhao, X. Miao, R. Liu, and G. Fortino, "Multi-sensor information fusion based on machine learning for real applications in human activity recognition: State-of-the-art and research challenges," *Inf. Fusion*, vol. 80, pp. 241–265, Apr. 2022.

[17] G. Marta, F. Simona, C. Andrea, B. Dario, S. Stefano, V. Federico, B. Marco, B. Francesco, M. Stefano, and P. Alessandra, "Wearable biofeedback suit to promote and monitor aquatic exercises: A feasibility study," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 4, pp. 1219–1231, Apr. 2020.

[18] F. Ferryanto and M. Nakashima, "Development of a markerless optical motion capture system for daily use of training in swimming," *Sports Eng.*, vol. 20, no. 1, pp. 63–72, Mar. 2017.

[19] A. Conceição, A. J. Silva, T. Barbosa, J. Campaniço, A. Costa, and H. Louro, "Neuromuscular and motor patterns in breaststroke technique," *Revista Brasileira Cineantropometria Desempenho Humano*, vol. 21, May 2019, Art. no. e56408.

[20] H. Cho, M. Sung, and B. Jun, "Canny text detector: Fast and robust scene text localization algorithm," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 3566–3573.

[21] D. Raihan and S. Chakravorty, "Particle Gaussian mixture filters-II," *Automatica*, vol. 98, pp. 341–349, Dec. 2018.

[22] R. Matei, "Gaussian 2D IIR filters with multiple orientation," in *Proc. Signal Process., Algorithms, Archit., Arrangements, Appl. (SPA)*, Sep. 2016, pp. 265–269.

[23] F. Cabello, J. Leon, Y. Iano, and R. Arthur, "Implementation of a fixed-point 2D Gaussian filter for image processing based on FPGA," in *Proc. Signal Process., Algorithms, Archit., Arrangements, Appl. (SPA)*, Sep. 2015, pp. 28–33.

[24] X. Chen, Y. Zhu, J. Hu, and D. Ma, "Kernel recursive least-squares algorithm with sliding-window approximately linear dependence sparsification," *Comput. Eng.*, vol. 8, p. 13, Jan. 2016.

[25] H. Huang and L.-Y. Hou, "Speed limit sign detection based on Gaussian color model and template matching," in *Proc. Int. Conf. Vis., Image Signal Process. (ICVISP)*, Sep. 2017, pp. 118–122.

[26] S. Guiming and S. Jidong, "Remote sensing image edge-detection based on improved Canny operator," in *Proc. 8th IEEE Int. Conf. Commun. Softw. Netw. (ICCSN)*, Jun. 2016, pp. 652–656.

[27] G. Cheng, F. Zhu, S. Xiang, and C. Pan, "Road centerline extraction via semisupervised segmentation and multidirection nonmaximum suppression," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 4, pp. 545–549, Apr. 2016.

[28] J. Wang, W. Yi, R. Hoseinnezhad, and L. Kong, "An agile multi-frame detection method for targets with time-varying existence," *Signal Process.*, vol. 165, pp. 133–143, Dec. 2019.

[29] R. F. Rahmat, T. Chairunnisa, D. Gunawan, and O. S. Sitompul, "Skin color segmentation using multi-color space threshold," in *Proc. 3rd Int. Conf. Comput. Inf. Sci. (ICCOINS)*, Aug. 2016, pp. 391–396.

[30] M. Khaliluzzaman, M. F. Alam, and T. Ahsan, "Human facial feature detection based on skin color and edge labeling," in *Proc. Int. Conf. Innov. Sci., Eng. Technol. (ICISET)*, Oct. 2016, pp. 1–4.

[31] S. A. Shah, D. Fan, A. Ren, N. Zhao, X. Yang, and S. A. K. Tanoli, "Seizure episodes detection via smart medical sensing system," *J. Ambient Intell. Humanized Comput.*, vol. 11, no. 11, pp. 4363–4375, Nov. 2020.

[32] M. W. Roche, "Method for detecting body parameters," U.S. Patent 9 451 919, 2016.

[33] A. O. Djekoune, K. Messaoudi, and K. Amara, "Incremental circle Hough transform: An improved method for circle detection," *Optik*, vol. 133, pp. 17–31, Mar. 2017.

[34] W. Xu-Chen, L. Xin-Chen, Z. Heng-Sheng, X. Ya-Min, and X. Mei, "A lane detection method based on parallel coordinate system," *J. Univ. Electron. Sci. Technol. China*, vol. 47, no. 3, pp. 362–367, 2018.

[35] L. Seifert, H. M. Toussaint, M. Alberty, C. Schnitzler, and D. Chollet, "Arm coordination, power, and swim efficiency in national and regional front crawl swimmers," *Hum. Movement Sci.*, vol. 29, no. 3, pp. 426–439, Jun. 2010.

[36] A. Callaway, "Quantification of performance analysis factors in front crawl using micro electronics: A data rich system for swimming," Ph.D. thesis, Dept. Rehabil. Sport Sci., Bournemouth Univ., Poole, U.K., 2014.

[37] A. Y. Shkanaev, D. V. Polevoy, A. V. Panchenko, Y. Y. Borisov, and M. I. Kolomeychenko, "Straw row position and orientation reconstruction through image segmentation," *Int. J. Appl. Eng. Res.*, vol. 11, no. 24, pp. 11803–11810, 2016.

[38] B. H. Olstad, J. R. Vaz, C. Zinner, J. M. H. Cabri, and P.-L. Kjendlie, "Muscle coordination, activation and kinematics of world-class and elite breaststroke swimmers during submaximal and maximal efforts," *J. Sports Sci.*, vol. 35, no. 11, pp. 1107–1117, Jun. 2017.

ZHENGLONG FENG is currently pursuing the master's degree with Xidian University, Xi'an, China. His research interests include FPGA parallel computing, electronic design, and data processing.

AIFENG REN (Member, IEEE) received the Ph.D. degree in information and telecommunication engineering from Xi'an Jiaotong University, Xi'an, China, in 2007. He is currently a Professor with the School of Electronic and Engineering, Xidian University. He has authored/coauthored over 40 journals and conference publications. His research interests include embedded wireless sensor networks, terahertz sensing technology, complex brain networks, and signal and image processing.

ZHIYUAN CHEN received the master's degree from Xidian University, China, in June 2021. His research interests include machine learning in wireless sensing, image processing, radar technology, and software defined radios. His current research is machine learning enabled terahertz sensing for precision agriculture technology at cellular level.

ADNAN ZAHID received the B.Sc. degree (Hons.) in electronics and communications engineering from Glasgow Caledonian University, U.K., in 2012, the M.Sc. degree in electronics and electrical engineering from the University of Strathclyde, in 2016, and the Ph.D. degree in electronics and electrical engineering from the University of Glasgow, U.K., in 2021. He is currently an Assistant Professor with the School of Engineering and Physical Sciences, Heriot-Watt University. His research interests include machine learning enabled terahertz sensing for precision agriculture technology at cellular level, RF, and multisensing for activity recognition supported by artificial intelligence, wearable devices, and the Internet of Things.

LEI YE received the master's degree from Xidian University, China, in June 2020. His research interests include machine learning in image processing and software defined radios. His current research is aimed at making swimming training more scientific and accessible.
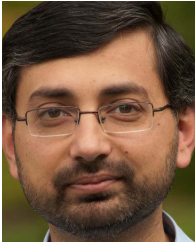
MUHAMMAD ALI IMRAN (Senior Member, IEEE) received the M.Sc. (Hons.) and Ph.D. degrees from the Imperial College London, U.K., in 2002 and 2007, respectively. He is currently the Vice Dean of the Glasgow College, UESTC, and a Professor in communication systems with the School of Engineering, University of Glasgow. He is also an Affiliate Professor at the University of Ajman, United Arab Emirates, and a Visiting Professor with the Artificial Intelligence Research Center (AIRC). He is also an Affiliate Professor at The University of Oklahoma, USA, and a Visiting Professor with the 5G Innovation Centre, university academic and industry experience, working primarily in the research areas of cellular communication systems. He has been awarded 15 patents, has (co)authored over 400 journals and conference publications.

QAMMER H. ABBASI (Senior Member, IEEE) received the B.Sc. and M.Sc. degrees in electronics and telecommunication engineering from the University of Engineering and Technology (UET) Lahore, Lahore, Pakistan, and the Ph.D. degree in electronic and electrical engineering from the Queen Mary University of London (QMUL), U.K., in 2012. In 2012, he was a Postdoctoral Research Assistant with the Antenna and Electromagnetics Group, QMUL. He is currently a Lecturer (an Assistant Professor) with the School of Engineering, University of Glasgow, U.K. He has contributed to a over 250 leading international technical journals and peer reviewed conference papers, and eight books. He received several recognitions for his research, which includes appearance in BBC, STV, dawnnews, local and international newspaper, cover of MDPI journal, most downloaded articles, U.K. exceptional talent endorsement by the Royal Academy of Engineering, National Talent Pool Award by Pakistan, International Young Scientist Award by NSFC China, URSI Young Scientist Award, National interest waiver by USA, four best paper awards, and best representative image of an outcome by QNRF. He is also an Associate Editor of IEEE JOURNAL OF ELECTROMAGNETICS, RF, AND MICROWAVES IN MEDICINE AND BIOLOGY, IEEE SENSORS, IEEE OPEN JOURNAL OF ANTENNA AND PROPAGATION, and IEEE ACCESS journal, and acted as a guest editor for numerous special issues in top notch journals.

• • •