Research Article

Algorithm for Swimmers' Starting Posture Correction Based on Kinect

Zheng Chang[1] and Yu Zhao [2]

[1]*Physical Education Teaching and Research Department, Northeast Agricultural University, Harbin 150030, China*
[2]*School of Electrical Engineering and Information, Northeast Agricultural University, Harbin 150030, China*

Correspondence should be addressed to Yu Zhao; zhaoyu@neau.edu.cn

Effective identification and correction of swimmers' improper postures can significantly improve athletes' weekday swimming training quality. The human body's affine deformation is prone to occur during swimming movements when performing posture recognition and correction, resulting in the creation of low-brightness action feature locations. The inability of coaches to identify and correct athletes' improper posture in real time is a result of a lack of detection and correction. Additionally, the human skeleton motion data from the depth camera Kinect contains a high amount of noise and fewer skeleton nodes, and the data level of detail is low. To overcome this issue, this research proposes a network for enhancing Kinect skeleton motion data. The network is composed of six bidirectional cyclic autoencoder stacks. The stacking structure improves the smoothness and naturalness of the data, and the training phase includes hidden variable limitations to ensure that the bone motion data preserve a genuine bone shape when the degree of detail is raised. The trials demonstrate that the optimized data from the network have a better degree of smoothness and can keep a more realistic bone structure, enabling the goal of obtaining high-precision motion capture data with low-precision Kinect equipment to be met.

1. Introduction

Swimming starts mark the beginning of a swimming competition, and they are a crucial aspect in gaining an advantage and ultimately winning a swimming tournament. When it comes to short-distance swimming competitions, the quality of the starting technique is a critical aspect in determining the final result. The beginning time at a distance of 25 meters determines 25% of the final results of the race. The beginning time at a distance of 50 m determines the outcome of the race by around 20% [1, 2]. Approximately 10% of the race outcomes are determined by the starting time from a distance of 100 meters. When it comes to short-distance swimming competitions, the squat start is a crucial beginning technique that is commonly utilized in the freestyle and butterfly events. Because of the ongoing increase in the world's competitive swimming level, as well as the increasingly fierce rivalry, the difference between high-level players is frequently measured in millimeters. The

result will be determined by any little modification or improvement in any technological link, which will result in a change in the competition ranking. Although swimmers spend less effort on starting than they do on strokes and turns, proper starting is still the most important factor in achieving success in short-distance swimming competitions [3, 4].

While competing in the World Swimming Championships, coaches and athletes from every country and every continent place a high value on each individual competition's performance, and the progression of each competition's performance is inextricably linked to the advancement of each athlete's skills. Having proper swimming posture can assist athletes in improving their swimming capabilities; however, incorrect swimming posture hinders the advancement of athletes' abilities. It will not be able to obtain optimal competition results. Because of this, proper assessment and adjustment of swimmers' posture is critical for athletes who want to enhance their skills and attain positive

competition outcomes. Under the influence of special medium water, the coach's spoken instruction will not be effective in improving all athletes' abilities, and athletes will have difficulty reproducing the coach's movements accurately. Furthermore, athletes find it difficult to correct their exact technical motions through hearing because of the limitations of the use of human body demonstrations in swimming technique instruction. They are unable to get self-feedback information and hence are unable to attain the goal of enhancing their own abilities [5–7].

As a result, it is critical to use multimedia technology to correct athletes' incorrect posture as they compete. The recognition and correction of swimmers' posture have emerged as a significant research topic among relevant experts and scholars and have significant research value. Methods such as video recording, front and back comparison of action pictures, on-site shooting, and replay of video technology, among others, have made it a popular topic of discussion. The research into athletes' incorrect posture recognition methods has yielded some promising results, according to the findings of the associated research. Some researchers use swimming movements to identify and fix the posture of swimmers, while others use a different method. Other researchers employ a dynamic module matching algorithm to detect and identify incorrect athlete posture, as well as to complete the recognition and repair of swimmers' incorrect postures. When the swimming posture is affected by scale, the image is susceptible to distortion. When the swimming posture is affected by factors such as noise and distance, the image is susceptible to distortion [8, 9].

Game creation, film and television production, sports training, and medical rehabilitation have all benefited from the availability of human skeletal motion data. Individuals can get human skeleton motion data by the use of high-precision motion capture technologies, such as the Vicon optical motion capture system and the Xsens sensor motion capture system, among others. However, because it is both pricey and cumbersome to wear, it is unlikely to become widely popular. The Kinect depth camera is capable of acquiring real-time data on the mobility of the human skeleton. It is inexpensive and simple to use; however, the motion data acquired contain just 25 skeletal nodes and contain a significant amount of noise [10].

As a result, before the Kinect data acquired from the human skeleton can be used meaningfully, the data must be optimized. Motion data optimization is classified into two categories: surface optimization, which refers to the process of filling in missing values in motion data, and noise and outlier removal, which requires that the optimized data retain the spatiotemporal pattern and human kinematics information contained in the motion data. This paper aims to increase the quality of Kinect data by utilizing deep neural networks. Not only can deep neural networks fill in gaps in the Kinect's skeleton node information, but they can also eliminate noise and outliers, raise the level of detail and accuracy of Kinect skeletal motion data, and permit acquisition with low-precision equipment. The purpose of high-precision motion capture data is to accurately correct

the swimmer's starting position, which is then communicated to the coach.

2. Background

2.1. Related Work. Some researchers developed a real-time probabilistic framework based on the Gaussian process model to increase the accuracy of motion data recorded by the Kinect in order to optimize the incomplete motion data captured by the device. In the situation of self-occlusion, the approach is also capable of generating high-quality motion. However, because of the high computing complexity of the Gaussian process, this strategy is only appropriate for limited datasets. Other researchers have used Gaussian process-based local mixture approaches to enhance the solution speed, and their methods allow the local model to be progressively updated in real time, which allows them to handle a larger number of data samples than previous researchers. Both of these strategies generate smooth motion by controlling the speed change between successive frames; however, it is impossible to recover the motion of turning with either of these two methods. Other researchers employ the dimensionality reduction method based on principal component analysis to search for suitable motion data from the motion database in order to fill in the missing values and then employ the proportional-derivative controller to generate physically credible motion in order to generate physically credible motion [4, 8, 11].

The accuracy of this motion recovery method is highly influenced by the content and amount of the database used in the analysis. Other researchers have presented a method that combines local pose estimation with global retrieval technology, employing a voting approach based on the Hausdorff distance to combine the two hypotheses in order to achieve the final motion pose, as described in detail below [12–14]. It takes into account the continuity of motion and can efficiently tackle the problem of data missing caused by occlusion and noise; nevertheless, it has a poor recovery effect for fast motion or motion that rotates more than 45° around the vertical axis. The non-deep learning Kinect data denoising methods discussed above can be used to process some motion segments; however, the effect of processing large continuous motion sequences that may contain arbitrary motion is not immediately apparent [15–18].

Methods for optimizing motion data based on deep learning have also been developed steadily over the last few years. While some studies improve motion data using EBD and EBF neural networks, others do not. By employing joint information and temporal correlation, EBD fills in gaps in missing data. EBF is essential because the motion data denoised by EBD contain faults such as jitter and poor noise reduction, implying that EBD alone is insufficient. Additional smoothing is possible, but this method cannot be used to increase the amount of information included in movement data [19–24]. Other studies map and invert Kinect motion data using CNN-based autoencoders, but the resulting motion data retain some jitter relative to the original data. Deep RNNs are used to process the three-dimensional coordinates and velocity of joint points in

Kinect data, and the outputs are stored as pDRNN and vDRNN, respectively. To assure the naturalness of motion, some researchers employ the Kalman filter and K-nearest neighbor algorithm to integrate and enhance the findings of pDRNN and vDRNN. This method requires additional postprocessing procedures to achieve smooth and natural motion [25–29]. A number of academics have asserted that the BRA network can optimize the Kinect bone motion data segment, but it has been demonstrated that when the amount of detail in the data needs to be increased, it is difficult for the BRA network to ensure optimization. The authenticity of the data in the post needs to be perceived [30–32].

2.2. Motion Characteristics of the Start of the Grab Table. In swimming competitions, the goal of any starting posture and technique is to cause the athlete's center of gravity to achieve the maximum horizontal speed as quickly as possible, and the swimmer's reaction and technical movements will play a critical role in the speed with which the initiation occurs. The horizontal force of the body's supporting reaction force is the most important source of horizontal speed, and the characteristics of the prepared posture and action define the size of the force value. As the body is being prepared to step down from a platform, it should be fully stretched and its center of gravity should be pushed forward as far as it can possibly go. However, if the technical fundamentals are not learned, the athlete's body will assume a sitting-back position when the starting signal is given, which will increase the time required to exit the stage and have a negative impact on the technology's effect and function. The initial power increases proportionally to how long it takes to depart the level. According to data from athletes' daily training sessions, the average athlete's time to leave the stage for the start of a competition is approximately 0.75 seconds. One of the most crucial factors for evaluating an athlete's starting technique is the athlete's response to the starting signal and the action that follows as a result. The force will be exerted in proportion to the amount of time spent away from the stage; conversely, the greater the amount of time spent away from the stage is, the lower the force will be.

When swimming begins, the first action is a movement of the human body, and several uncontrollable factors must be taken into account. For example, some swimmers will have a short flight time but a long distance during daily training, while others will have a long flight time but a short distance. Additionally, the time period is lengthy. When it comes to movement parameters following takeoff, the swimmer's speed upon departure from the platform is crucial; thus, the grab table's starting method has a relatively fair departure angle, as a good departure angle results in an improved aerial. Athlete's posture and time in the air can be drawn as a graphs. One of the most critical parts of the table start technique is the departure angle. Similar to how the law of motion for a parabola has been developed, the whole movement path following the athlete's exit from the stage has been established. When the athlete's center of

gravity cannot be changed while in the air, the athlete's posture can be shifted into the water stance. When entering the water, it is necessary to completely reverse one's posture, from up to down, as well as to adjust the angle of entry to the water to one's position. The vertical axis remains unchanged from the previous configuration.

The table grab start technique is the most often utilized beginning technique in the training and competition of young swimmers. It is also the most effective starting technique. Swingarm starting technology and grasping table starting technology are two of the most often used beginning techniques today. This technology for grabbing table beginning offers several advantages, as can be demonstrated in this example. There are several essential elements of the technology used to initiate the grab table's operation. As a starting table is being prepared, two hands grab hold of the preparation position, allowing the arm to be employed as a support point, resulting in the athlete's body being more solid and making it harder for him or her to seem offset. The use of fouls is minimized to some extent. Second, the center of gravity of the body can be shifted forward to the greatest extent possible during preparation, allowing the takeoff speed to be significantly increased with minimal effort. Third, the entry angle of young swimmers can be decreased to a significant degree throughout the process of taking off and leaving the water. Fourth, the initial speed can be increased to the greatest extent possible because the athlete's time in the air is relatively short when using the table start technique, and the speed with which the athlete leaves the table and enters the water is also relatively quick.

Because of their inexperience with grab-and-go starting procedures, which have been developed through relevant research, young athletes frequently make a series of mistakes in their everyday training sessions. Examples include competitors who fail to grasp the starting platform, which is wholly unacceptable behavior in the sport. Because the athlete is required to grasp the starting platform with both hands, when the starting signal is issued, the arm can be quickly pulled up and the elbow can be bent, causing the athlete's body to be tilted downward and the distance between the thigh and upper body to be maximized, and then the center of gravity of the body can be moved forward and to the front of the starting platform, resulting in an acceleration of time to enter the water. Furthermore, grasping the starting platform with both hands has another vital role, which is to maintain the athlete's body when the athlete's center of gravity departs from the support surface during the race. It can also maintain a relatively static state for an extended length of time, allowing it to efficiently manage the forward thrust force stored by the body during that time. Athletes simply need to flatten their bodies at this point because the body will rush forward with the help of the already stored forward force when they do. Meanwhile, there is another critical element that athletes must learn in order to be successful in their endeavors. When the horizontal plane is at a 45-degree angle between the athlete's center of gravity and the support points of both feet, the best takeoff angle is achieved.

In addition, when the athlete begins to take off, his body's center of gravity has already shifted away from the starting platform and his body has already exceeded the table of the support platform. Athletes that utilize the grab table starting technique just need to drop their heads, stretch their legs and pull their arms, allow their bodies to lean forward as much as they can, and then conduct actions such as bending their knees and lying down to begin. When the angle between the knee joint and the ankle joint achieves the most appropriate angle, the athlete executes activities such as extending the knee and ankle, swinging the arm, and extending the hip all at the same time, resulting in the formation of a resultant force by the entire body. However, if the athlete's hands do not grab the starting platform, the body will not be able to optimize forward leaning, resulting in a flat shot into the water and a relatively slow speed of entry into the water, both of which are harmful to the athlete's overall performance. As a result, while young swimmers are training, the coach should emphasize the importance of effective training in the grab table starting technique. First and foremost, ensure that the athletes have a good understanding of the key tactics and their characteristics, so that the young swimmer may more thoroughly begin the grab table. Master the advantages of technology, employ the fundamentals of movement in a flexible manner, and then effectively boost your sporting performance [31].

3. Method

3.1. Motion Dataset. To accomplish the study's objective, it is necessary to first construct a motion dataset that is simultaneously collected by two motion data collecting devices, namely, the Kinect and motion capture devices, and then analyze that dataset. This project establishes a synchronized database through the use of the Kinect and Noitom Perception Neuron motion capture sensors. This enables performers to practice swimming grabbing and starting actions continuously while in the acquisition environment, resulting in a succession of longer motion sequences. The data collected by Kinect are stored in joint point coordinate format, and each frame of data contains the three-dimensional coordinates of 25 joint points, resulting in a motion data frame with a dimension of 75 ($25 \cdot 3 = 75$). Noitom converts the motion data it collects into a format called joint coordinates. Each frame of data contains the three-dimensional coordinates of 59 joint points, yielding a total dimension of 177 points per frame of motion data. As illustrated in Figures 1 and 2, a skeleton structure has been generated from motion data collected by two motion collection devices [32].

Following the acquisition of the motion sequence with two different types of equipment, the data must be preprocessed in a synchronized manner. Prior to any motion analysis, spatial processing must be completed, and the two kinds of data must be translated into relative coordinate data based on node 0, which is the root node, before the two kinds of motion analysis can begin. In this case, the data are time-synchronized because the sample frequency of Noitom motion capture is 60 frames per second and the sampling frequency of Kinect is unstable, around 30 frames per
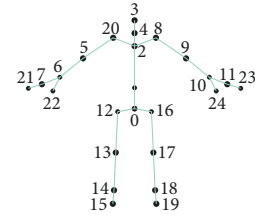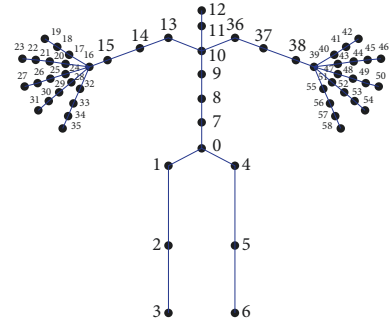

Figure 1: Kinect.

Figure 2: Motion capture.

second, and thus the sampling rate and time domain of Noitom are used as the benchmarks. The data from the Kinect were upsampled to 60 frames per second using natural spline interpolation. It is acquired after preprocessing the data to obtain a synchronized long-sequence motion dataset, from which 80% of it is utilized as a training set and 20% of it is used as a test set.

Take the Kinect motion data segment as the noise data X_N and denote the Nuoyiteng motion capture data as X_C.

The network structure of this paper is shown in Figure 3.

The input data of the network refer to the Kinect motion data segment, and the label data refer to the motion capture motion data segment. For this reason, this paper defines the mean square error L_p as

$$L_p(Y, X_C) = \frac{1}{f} \|Y - X_C\|_2, \tag{1}$$

where $1/f = 1/120$.

Let the bone length of the stacked autoencoder output Y in the ith frame of the bone t be

$$I_{t_i}' = \|p_t^{i,1} - p_t^{i,2}\|_2, \tag{2}$$

where $p_t^{i,1}$ and $p_t^{i,2}$ are the position coordinates corresponding to the two end nodes, respectively, and the loss function about the bone length is

$$L_b(Y) = \frac{1}{f(J-1)} \sum_{i=1}^{f} \sum_{t=1}^{J-1} |L_{t_i}' - L_t|. \tag{3}$$

To ensure that the optimized motion sequence transitions smoothly between adjacent frames, this paper imposes a smoothness constraint on the output data of the stacked autoencoder:
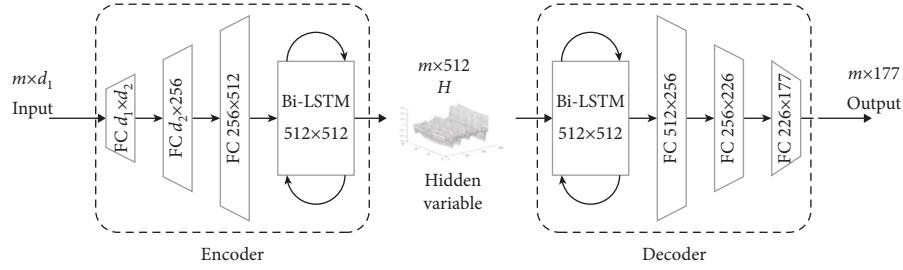
FIGURE 3: Structure of our method.

$$O = \begin{bmatrix} 1 & -1 & & & \\ -2 & 1 & & & \\ 1 & & & & \\ \ddots & \ddots & \ddots & 1 & \\ 1 & -2 & & & \\ 1 & -1 & & & \end{bmatrix} (f+2) \times (f+2).$$ (4)

The loss is

$$L_s(y) = \frac{1}{177(f+2)} \|OY'\|_2.$$ (5)

The perceptual loss is

$$L_C(X_{S_6}) = \frac{1}{120f} \|E_C(X_{S_6}) - E_C(X_C)\|_2.$$ (6)

The first stage only uses all X_C in the training dataset to train the perceptual autoencoder and imposes the mean square error constraint, bone length constraint, and smoothness constraint on the output Y_C of the perceptual autoencoder. The loss is

$$L_1 = \alpha_{C1}L_p(Y_C) + \alpha_{C2}L_b(Y_C) + \alpha_{C3}L_s(Y_C).$$ (7)

On the basis of hidden variable constraints, the second stage trains a stacked autoencoder, which then applies a mean squared error constraint to the output of each autoencoder within the stacked bidirectional recurrent autoencoder. There are bone length limitations and smoothness constraints, and the function also imposes hidden variable constraints, which results in loss.

$$L_2 = \sum_{i=1}^{6} \left(\alpha_p L_p(X_{S_i}) + \alpha_b L_b(X_{S_i}) + \alpha_s L_s(X_{S_i}) \right).$$ (8)

In the process of athlete pose recognition and correction, $Q = (Q_R, Q_G, Q_B)$ represents the foreground pixels of the athlete pose image, and $W = (W_R, W_G, W_B)$ represents background pixels of the athlete pose image. The positioning of each joint point when the athlete swims is

$$G = (x, y, z) = \frac{Q \cdot W}{(\Delta_T, \Delta_D, \Delta_0, \Delta_{L0}, \Delta_{L1}, \Delta_{R0}, \Delta_{R1})}.$$ (9)

Analyze the located joint points. Use equation (10) to extract the feature points of each joint of the athlete's limbs during swimming:

$$x^2 + y^2 + z^2 = \sqrt{\frac{\text{shead}}{\pi}} \times \frac{(x, y)}{R} \times (x - r \cdot \Phi, y + r),$$ (10)

where $\Phi = 1 + \sqrt{5}/2$.

As part of the preprocessing of the image, median filtering is employed because it is capable of successfully reducing the non-linear signal of noise in essence. The median of each value is represented. Individual noise points are excluded from the analysis using this strategy, which involves making the relationship between the adjacent real value and the pixel value identical. The advantages of this technology are that it removes noise quickly, it is quick, and it is simple to operate. It is capable of effectively preprocessing digital photos under particular situations [28–30].

4. Results

Parallel to the continuous upgrading and improvement of domestic information technology, the development of DV video screen-related hardware is reaching completion, with the recognition and monitoring of moving target image features being widely used in a number of industries. Swimming activities also have a broader range of applications. The first stage is to acquire images using Kinect technology; the second step is to implement timely detection and recognition of swimming target poses using a method given in the third part. The athlete's goal action is estimated using deep image interception, which reduces the challenging prediction problem to a relatively straightforward classification problem based on the image deep interception and tracking method. The swimming posture of the human body is employed as the tracking and detection objective, and the Kinect technology is used to collect deep intercepted images in order to more effectively monitor moving targets and obtain better results. The position information for the moving target can be calculated using the approach outlined above, and this position information can serve as a critical foundation for later retrieval and recognition of swimming improper postures.

The validity and stability of the algorithm are checked by measuring the precision and rate with which swimming posture photos of moving targets are captured. The long sequence data of different athletes are recorded in test experiment 1, as shown in Figures 4–6, when they first begin competing.
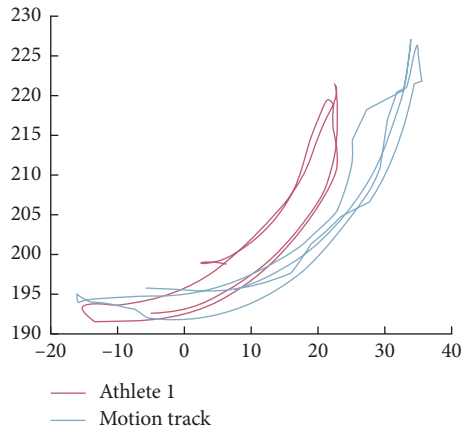
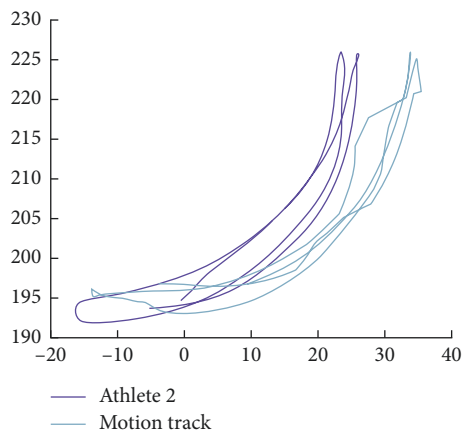FIGURE 4: Motion capture track of athlete 1.

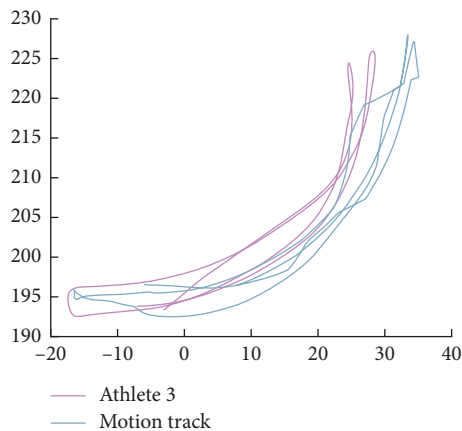FIGURE 5: Motion capture track of athlete 2.

FIGURE 6: Motion capture track of athlete 3.

It can be seen that for three athletes, our method can better capture the pose of the athletes.

To further study the distinctions between the method described in this research and KPC, a standard posture capture algorithm, we compare the two methods' success rate and matching time (again, as shown in Figures 7 and 8).
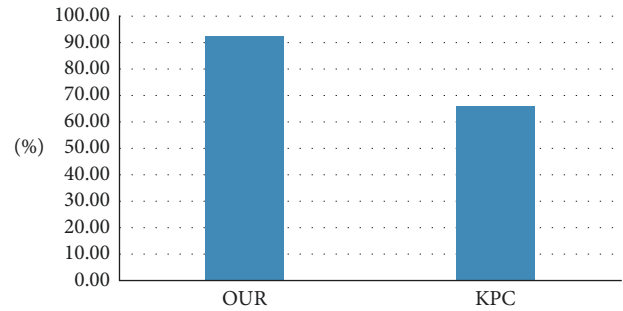

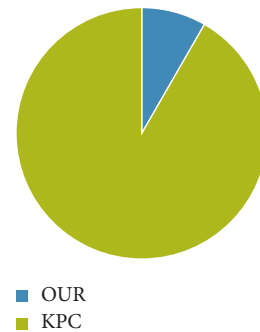FIGURE 7: Comparison of success rates of different methods.

FIGURE 8: Comparison of time consumption of different methods.

As can be observed, our solution outperforms the KPC method in terms of success rate (more than 92%) and matching time (much less than the KPC method).

5. Conclusion

Effectively identifying and correcting swimmers' incorrect postures can considerably improve the quality of athletes' daily swimming training. Coaches' inability to notice and correct athletes' poor posture in real time is due to a deficiency in detection and correction. Additionally, the human skeleton motion data from the depth camera Kinect are noisy, with fewer skeleton nodes and a poor level of resolution. To address this issue, this study presents an enhanced network using Kinect skeleton motion data. Six bidirectional cyclic autoencoder stacks comprise the network. The stacking structure enhances the smoothness and naturalness of the data, and the training phase incorporates hidden variable constraints to ensure that the bone motion data retain its real shape when the degree of detail is increased. The studies demonstrate that the optimized data from the network have a higher degree of smoothness and can maintain a more realistic bone structure, enabling the goal of high-precision motion capture using low-precision Kinect equipment to be accomplished.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] H. Haggag, M. Hossny, and S. Haggag, "Safety applications using Kinect technology," in *Proceedings of the 2014 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pp. 2164–2169, San Diego, CA, USA, October 2014.

[2] H. Alaoui, M. T. Moutacalli, and M. Adda, "Ai-enabled high-level layer for posture recognition using the azure Kinect in unity3d," in *Proceedings of the 2020 IEEE 4th International Conference on Image Processing, Applications and Systems (IPAS)*, pp. 155–161, Genova, Italy, December 2020.

[3] S. Örücü and M. Selek, "Design and validation of rule-based expert system by using Kinect V2 for real-time athlete support," *Applied Sciences*, vol. 10, no. 2, p. 611, 2020.

[4] P. Soltani, P. Figueiredo, R. J. Fernandes, and J. P. Vilas-Boas, "Do player performance, real sport experience, and gender affect movement patterns during equivalent exergame?" *Computers in Human Behavior*, vol. 63, pp. 1–8, 2016.

[5] B. Milosevic, A. Leardini, and E. Farella, "Kinect and wearable inertial sensors for motor rehabilitation programs at home: state of the art and an experimental comparison," *BioMedical Engineering Online*, vol. 19, no. 1, pp. 1–26, 2020.

[6] J. E. Muñoz, J. F. Villada, and S. Casanova, "Dynamic systems theory in human movement exploring coordination patterns by angle-angle diagrams using Kinect," in *Proceedings of the 2018 10th International Conference on Virtual Worlds and Games for Serious Applications (VS-Games)*, pp. 1–4, IEEE, Würzburg, Germany, September, 2018.

[7] J. Yin, "The method of table tennis players' posture recognition based on a genetic algorithm," *International Journal of Biometrics*, vol. 13, no. 2-3, pp. 243–257, 2021.

[8] T. Seno, F. Funatsu, and S. Palmisano, "Virtual swimming - breaststroke body movements facilitate vection," *Multisensory Research*, vol. 26, no. 3, pp. 267–275, 2013.

[9] Z. Huang, A. Nagata, M. Kanai-Pak et al., "Self-help training system for nursing students to learn patient transfer skills," *IEEE Transactions on Learning Technologies*, vol. 7, no. 4, pp. 319–332, 2014.

[10] S. Xu and S. Lee, "An inertial sensing-based approach to swimming pose recognition and data analysis," *Journal of Sensors*, p. 2022, 2022.

[11] L. Yang, B. Yang, and H. Dong, "3-D markerless tracking of human gait by geometric trilateration of multiple Kinects," *IEEE Systems Journal*, vol. 12, no. 2, pp. 1393–1403, 2016.

[12] Q. Liu, X. Wu, Y. Jiang et al., "A fully connected deep learning approach to upper limb gesture recognition in a secure FES rehabilitation environment," *International Journal of Intelligent Systems*, vol. 36, no. 5, pp. 2387–2411, 2021.

[13] E. Matsas and G.-C. Vosniakos, "Design of a virtual reality training system for human-robot collaboration in manufacturing tasks," *International Journal on Interactive Design and Manufacturing*, vol. 11, no. 2, pp. 139–153, 2017.

[14] J. Beacon, *Assessing 2D and 3D Motion Tracking Technologies for Measuring the Immediate Impact of Feldenkrais Training on the Playing Postures of pianists*, Université d'Ottawa/ University of Ottawa, Ottawa, Canada, 2015.

[15] T. L. Le and M. Q. Nguyen, "Human posture recognition using human skeleton provided by Kinect," in *Proceedings of the 2013 International Conference on Computing, Management and Telecommunications*, pp. 340–345, Ho Chi Minh City, Vietnam, January 2013.

[16] B. F. Mentiplay, R. A. Clark, and A. Mullins, "Reliability and validity of the Microsoft Kinect for evaluating static foot posture[J]," *Journal of Foot and Ankle Research*, vol. 6, no. 1, pp. 1–10, 2013.

[17] Z. Zhang, Y. Liu, and A. Li, "A novel method for user-defined human posture recognition using Kinect," in *Proceedings of the 2014 7th International Congress on Image and Signal Processing*, pp. 736–740, Dalian, China, October 2014.

[18] J. A. Diego-Mas and J. Alcaide-Marzal, "Using Kinect sensor in observational methods for assessing postures at work," *Applied Ergonomics*, vol. 45, no. 4, pp. 976–985, 2014.

[19] W.-J. Wang, J.-W. Chang, S.-F. Haung, and R.-J. Wang, "Human posture recognition based on images captured by the Kinect sensor," *International Journal of Advanced Robotic Systems*, vol. 13, no. 2, p. 54, 2016.

[20] M. U. Islam, H. Mahmud, and F. B. Ashraf, "Yoga posture recognition by detecting human joint points in real time using microsoft Kinect," in *Proceedings of the 2017 IEEE Region 10 Humanitarian Technology Conference*, pp. 668–673, Dhaka, Bangladesh, December 2017.

[21] L. Zhou, Z. Liu, and H. Leung, "Posture reconstruction using Kinect with a probabilistic model," in *Proceedings of the 20th ACM Symposium on Virtual Reality Software and Technology*, pp. 117–125, Edinburgh Scotland, November 2014.

[22] Y. Wang and R. Yang, "Real-time hand posture recognition based on hand dominant line using Kinect," in *Proceedings of the 2013 IEEE International Conference on Multimedia and Expo Workshop*, pp. 1–4, San Jose, CA, USA, July 2013.

[23] Z. Xiao, F. Mengyin, and Y. Yi, "3D human postures recognition using Kinect," in *Proceedings of the 2012 4th International Conference on Intelligent Human-Machine Systems and Cybernetics*, pp. 344–347, Washington, DC, USA, August 2012.

[24] M. C. Shih, R. Y. Wang, and S. J. Cheng, "Effects of a balance-based exergaming intervention using the Kinect sensor on posture stability in individuals with Parkinson's disease: a single-blinded randomized controlled trial," *Journal of NeuroEngineering and Rehabilitation*, vol. 13, no. 1, pp. 1–9, 2016.

[25] B. Li, C. Han, and B. Bai, "Hybrid approach for human posture recognition using anthropometry and BP neural network based on Kinect V2," *EURASIP Journal on Image and Video Processing*, vol. 2019, no. 1, pp. 1–15, 2019.

[26] S. Monir, S. Rubya, and H. S. Ferdous, "Rotation and scale invariant posture recognition using Microsoft Kinect skeletal tracking feature," in *Proceedings of the 2012 12th International Conference on Intelligent Systems Design and Applications*, pp. 404–409, Kochi, India, November 2012.

[27] W. J. A. Grooten, L. Sandberg, and J. Ressman, "Reliability and validity of a novel Kinect-based software program for measuring posture, balance and side-bending," *BMC Musculoskeletal Disorders*, vol. 19, no. 1, pp. 1–13, 2018.

[28] S. Mealin, I. X. Domínguez, and D. L. Roberts, "Semi-supervised classification of static canine postures using the Microsoft Kinect," in *Proceedings of the Third International Conference on Animal-Computer Interaction*, pp. 1–4, Milton Keynes, UK, November 2016.

[29] C. K. Lee and V. Y. Lee, "Fall detection system based on Kinect sensor using novel detection and posture recognition algorithm," in *Proceedings of the International Conference on Smart Homes and Health Telematics*, pp. 238–244, Hammamet, Tunisia, June 2013.

[30] A. Dutta, S. Chugh, A. Banerjee, and A. Dutta, "Point-of-care-testing of standing posture with Wii balance board and microsoft Kinect during transcranial direct current stimulation: a feasibility study," *NeuroRehabilitation*, vol. 34, no. 4, pp. 789–798, 2014.

[31] E. Rocha-Ibarra, M.-I. Oros-Flores, D.-L. Almanza-Ojeda et al., "Kinect validation of ergonomics in human pick and place activities through lateral automatic posture detection," *IEEE Access*, vol. 9, Article ID 109067, 2021.

[32] B.-L. Oh, J. Kim, J. Kim, J.-M. Hwang, and J. Lee, "Validity and reliability of head posture measurement using Microsoft Kinect," *British Journal of Ophthalmology*, vol. 98, no. 11, pp. 1560–1564, 2014.