

Intelligent Skiing Posture Detection and Recognition Through Internet of Bodies

Peihua Liu, Beihua University, China*

ABSTRACT

The training of skiing should start from the control of body posture ability to highlight the characteristics of the sports. Thus, the athletes can have the sports ability in the process of high-speed skiing. This paper establishes a system to automatically recognize the skiing posture which can help athletes grasp the skiing postures. First, the skiing images are collected by distributed camera. Second, the skeleton features are extracted to learn a classification model which is used to recognize and adjust skiing postures. Lastly, the analytical results of posture recognition are returned to athletes through internet of bodies. The framework can effectively recognize the skiing postures and provide athletes with training advice.

KEYWORDS

Classification Model, Internet of Bodies, Posture Recognition, Skeleton Features

1. INTRODUCTION

The technique of skiing, especially alpine skiing, is very difficult, which requires athletes to have a high level of special ability to ensure the exertion of the technique. Special ability refers to various abilities which are suitable for the characteristics of the project and can improve the performance in this project. For skiing, the cultivation of special ability should not only enhance the overall physical quality (Ahmed et al. 2021) based on strength, but also improves the skill to control body posture. The ability to control ability body posture (Horak et al. 2002) is a special ability of skiers. The improvement of this ability has an important impact on the competition.

From the development process of skiing projects, we need to constantly explore new skills and strategies, then combine and improve them. With the progress of sports biomechanics and sports training (Ying & Huang 2021, Pellegrini et al. 2018), the method of sports technology analysis has been greatly improved. The research on posture training and correction (Han et al. 2017, Kashuba et al. 2019) is a hot topic in the scientific research of skiing. The distortion of body posture may lead to skiing failure and induce injury (Rentenberger et al. 2021, Yoshioka et al. 2018). When passing through various uneven areas, the speed of athletes who lose their balance due to large distortion of body posture will be significantly slower than the athletes who maintain a stable skiing posture. It is necessary to pay attention on reducing leaping movements and various leaping body movements during skiing. The key is to keep the balance of body.

In the skiing competition, body posture has a direct impact on competition results. Through the repeated the experimental results in training practice, it is proved that the distortion of posture is induced by the compound participation of posture reflex (Yoshioka et al. 2018, Alhammad

DOI: 10.4018/IJMCMC.293746

*Corresponding Author

This article published as an Open Access Article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>) which permits unrestricted use, distribution, and production in any medium, provided the author of the original work and original publication source are properly credited.

2020). For example, an adult athlete who wants to improve his sports performance will raise his hand unconsciously, close his eyes, and cause the excitement of anterior tibial muscle related to joint movement. Therefore, it can be said that the distortion of body posture is directly related to the participation of conditioned reflex. This change in body posture itself is the instinctive physiological mechanism to maintain the best movement. However, in the skiing competition, even if it becomes a physiological mechanism to restore balance, unconscious actions will also cause distortion in posture and affect competition results. This contradiction is mainly solved in training through two methods. One is to train athletes to get used to the conditions of physical changes. The other is to improve the ability to predict physical changes which can be improve with the help of automatic posture recognition and correction system.

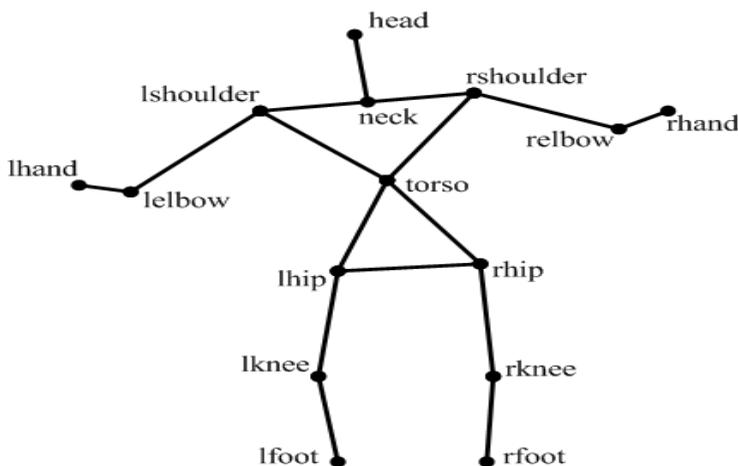
Kinect somatosensory device developed by Microsoft can easily collect human joint point data, scene depth information and color images (Li & Xiao 2018, Wang et al. 2019). It has attracted the attention of many researchers due to its excellent performance. The Kinect has been used in many applications, such as robot navigation (Li & Wang 2020), target tracking (Pang & Liang 2019), object recognition (Yang et al. 2021), and 3D modeling (Cai 2019).

Recently, using Kinect to study human behavior and posture recognition has become a hot spot in the field of behavior recognition (Ma & Guo 2019, Li 2020). In this paper, the Kinect is used to collect joint point data of skiing players. The joint point data is used to analyze the posture of skiing players. According to the characteristics of human structure, human structure vector is constructed by using joint point data as human model in three-dimensional space. The posture features are extracted by calculating the vector angle and vector modulus ratio between human structure vectors. Then, the features of skiing player posture is matched and corrected by the dynamic time warping (DTW) algorithm (Permanasari et al. 2019, Yadav & Alam 2018).

2. SKIING PLAYER JOINT POINTS FROM KINECT SENSOR AND FEATURE REPRESETATION

When an observer stands toward Kinect camera, the positive direction of X-axis points to the right of the observer, the positive direction of Y-axis points to the observer's head, the positive direction of Z-axis is consistent with the observation direction of Kinect camera, and the Z-axis represents depth information. The Kinect camera can collect 15 joint points to represent different parts of the human body. The distribution of each joint point is shown in Figure 1.

Figure 1. The architecture of document type recognition framework



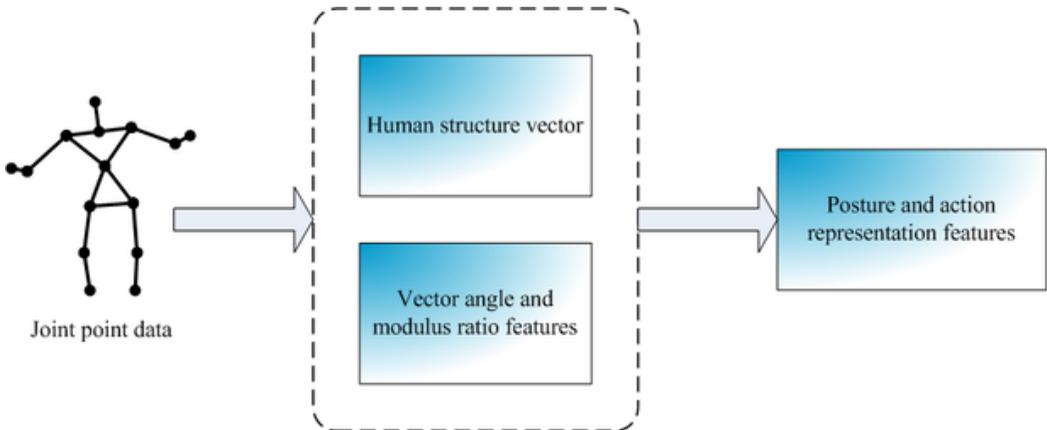
Each joint point is assigned as a name. The details are illustrated in Table 1.

Table 1. The performance comparison of document type recognition

| Position | Name | Position | Name | Position | Name |
|-------------|--------|----------------|-----------|-------------|-------|
| Head | Head | Neck | Neck | Left hand | Lhand |
| Left elbow | Leibow | Left shoulder | Lshoulder | Right hand | Rhand |
| Right elbow | Relbow | Right shoulder | Rshoulder | Center node | Torso |
| Left hip | Lhip | Left knee | Lknee | Left foot | Lfoot |
| Right hip | Rhip | Right knee | Rknee | Right foot | rfoot |

Due to the difference of human posture and the change of relative position relationship with Kinect, it may increase the difficulty of behavior recognition and classification to directly use joint point information as behavior representation feature. In order to use Kinect joint point data for human behavior analysis, the joint point data should be preprocessed first. Considering human cognitive characteristics and human structure characteristics, this paper first uses joint point data to construct human structure vector to express the posture of human limbs, trunk and head. The human behavior information is mostly contained in these posture information. Based on the structure vector, the features of behavior representation is obtained by calculating the angle and modulus ratio between vectors. The procedure is summarized in the following figure.

Figure 2. The architecture of document type recognition framework



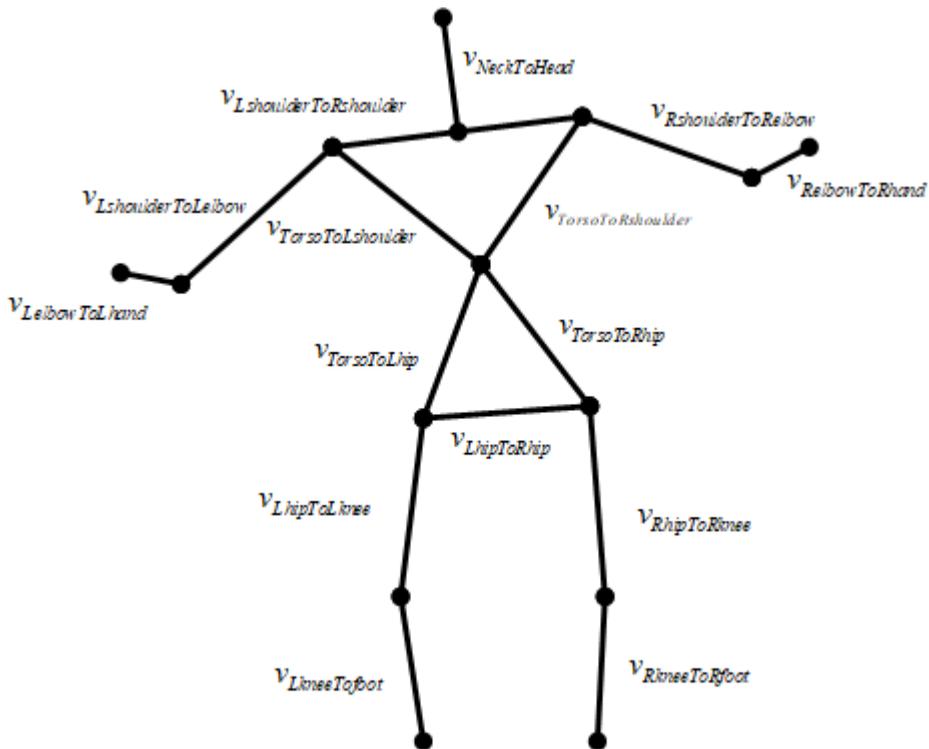
By analyzing human posture, the human structure vector is constructed by selecting appropriate joint point data according to the characteristics of human structure. The human structure vector is very simple. Let $A(x_1, y_1, z_1)$ and $B(x_2, y_2, z_2)$ be two joint points, the corresponding human structure vector is represented as follows:

$$v_{AB} = (x_2 - x_1, y_2 - y_1, z_2 - z_1) \tag{1}$$

According to Equation (1) $A(x_1, y_1, z_1)$ and $B(x_2, y_2, z_2)$ represent rshoulder and relbow, the vector in Equation (1) describes the posture of the right boom. By combing with the actual situation, the posture may contain more human action information.

In order to describe human actions during skiing, 15 groups of human structure vectors are constructed in this paper. These human structure vectors are divided into three types according to different body parts, including upper limb part, lower limb part, and trunk part. An illustration of human structure vectors is shown in Figure

Figure 3. The architecture of document type recognition framework



The upper limb part includes the human structure vectors from left shoulder to left elbow, from left elbow to left hand, right shoulder to right elbow, from right elbow to right hand. These human structure vectors are denoted as $v_{LshoulderToLelbow}$, $v_{LelbowToLhand}$, $v_{RshoulderToRelbow}$, and $v_{RelbowToRhand}$, respectively. The lower limb part includes the human structure vectors from left hip to left knee, from left knee to left foot, right hip to right knee, from right knee to right foot. These human structure vectors are denoted as $v_{LhipToLknee}$, $v_{LkneeToFoot}$, $v_{RhipToRknee}$, and $v_{RkneeToRfoot}$, respectively. The trunk part includes the human structure vectors from neck to head, from left shoulder to right shoulder, from torso to left shoulder, from torso to right shoulder, from torso to left hip, from torso to right hip, and from left hip to right hip. These human structure vectors are denoted as $v_{NeckToHead}$, $v_{LshoulderToRshoulder}$, $v_{TorsoToLshoulder}$, $v_{TorsoToRshoulder}$, $v_{TorsoToLhip}$, $v_{TorsoToRhip}$, $v_{LhipToRhip}$ respectively.

When the human body is moving, its posture and the distance to Kinect camera often change, which results in large changes in the defined human body structure vector. However, the actual situation

is that the change trend of limbs during human movement is roughly similar. By selecting the angle between structural vectors, it can effectively eliminate the differences between human structural vectors of different people or different positions. The angle between two vectors

$a = (a_1, a_2, a_3)$ and $b = (b_1, b_2, b_3)$ is written as follows:

$$\theta = \arccos \frac{a \cdot b}{\|a\| \|b\|}, \|a\| \neq 0 \ \& \ \|b\| \neq 0 \quad (2)$$

In the Equation (2), $a \cdot b$, $\|a\|$, and $\|b\|$ are defined as follows:

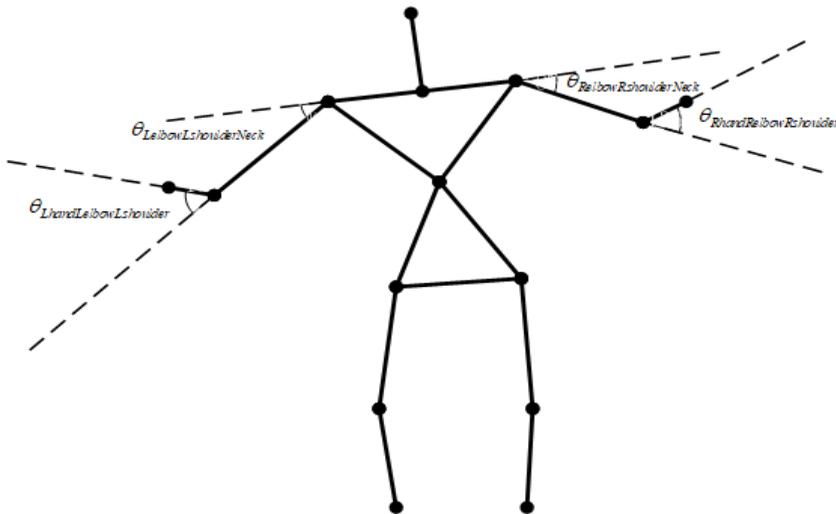
$$a \cdot b = a_1 b_1 + a_2 b_2 + a_3 b_3 \quad (3)$$

$$\|a\| = \sqrt{a_1^2 + a_2^2 + a_3^2} \quad (4)$$

$$\|b\| = \sqrt{b_1^2 + b_2^2 + b_3^2} \quad (5)$$

According to the human structure vectors, the angles are classified as three parts: upper limb angles, lower limb angles, and intermediate junction angles. The upper limb angles are used to describe the posture or movement of arms. The lower limb angles are used to describe the posture or movement of legs. The angle information at the middle junction is mainly used to represent the changes of human limbs relative to the trunk and the changes of human trunk itself during movement. Figure 4 shows the angle information of upper limbs

Figure 4. The architecture of document type recognition framework



In the Figure 4, the angles of upper limbs include the angle between $v_{LshoulderToLelbow}$ and $v_{LelbowToLhand}$, the angle between $v_{LshoulderToLelbow}$ and $v_{LshoulderToRshoulder}$, the angle between $v_{RshoulderToRelbow}$ and $v_{RelbowToRhand}$, and the angle between $v_{RshoulderToRelbow}$ and $v_{LshoulderToRshoulder}$. These angles are denoted as $\theta_{LhandLelbowLshoulder}$, $\theta_{LelbowLshoulderNeck}$, $\theta_{RhandRelbowRshoulder}$, and $\theta_{RelbowRshoulderNeck}$.

The lower limb angles include the angle between $v_{LhipToLknee}$ and $v_{LhipToRhip}$, the angle between $v_{RkneeToRfoot}$ and $v_{RhipToRknee}$, and the angle between $v_{RhipToRknee}$ and $v_{LhipToRhip}$. These angles are denoted as $\theta_{LhipLkneeLfoot}$, $\theta_{LkneeLhipRhip}$, $\theta_{LhipRhipRknee}$, and $\theta_{RhipRkneeRfoot}$.

The intermediate junction angles include the angle between $v_{HeadToNeck}$ and $v_{LshoulderToRshoulder}$, the angle between $v_{LshoulderToRshoulder}$ and $v_{TorsoToLshoulder}$, the angle between $v_{LshoulderToRshoulder}$ and $v_{TorsoToRshoulder}$, the angle between $v_{TorsoToLshoulder}$ and $v_{TorsoToRshoulder}$, the angle between $v_{TorsoToLshoulder}$ and $v_{TorsoToLhip}$, the angle between $v_{TorsoToRshoulder}$ and $v_{TorsoToRhip}$, the angle between $v_{TorsoToLhip}$ and $v_{LhipToRhip}$, and the angle between $v_{TorsoToRhip}$ and $v_{LhipToRhip}$. These angles are denoted as $\theta_{HeadNeckShoulder}$, $\theta_{RshoulderLshoulderTorso}$, $\theta_{LshoulderRshoulderTorso}$, $\theta_{LshoulderTorsoRshoulder}$, $\theta_{TorsoLhipRhip}$, $\theta_{TorsoRhipLhip}$, and $\theta_{LhipTorsoRhip}$.

In some cases, only using the angle information between structural vectors cannot describe the details of the posture and action. For instance, when determining the relative position between the upper limb and the trunk or head, the angle information cannot provide overall information of the posture. In order to solve this problem, this paper adopts the ratio between three groups of human structural vector modules as auxiliary information to improve the processing of joint point data. An illustration of ratio between human structure vector is shown in Figure 5.

In Figure 5, the a represents the vector from torso to head, the b represents the vector from right hand to head, the c represents the vector from left hand to head, the d represents the vector from torso to left hand, and e represents the vector from torso to right hand. Then, some ratios between human structure vectors are listed in the following equations.

$$\begin{aligned} rat_{ba} &= \frac{\|b\|}{\|a\|} \\ rat_{ca} &= \frac{\|c\|}{\|a\|} \\ rat_{da} &= \frac{\|d\|}{\|a\|} \\ rat_{ea} &= \frac{\|e\|}{\|a\|} \end{aligned} \tag{6}$$

The human posture p refers to the limb state at a specified time. In this paper, the human posture is represented as the combination of processed vector angle and modulus ratio, $p = (p_1, \dots, p_1, 5)$. The $p_i, i = 1, \dots, 11$ is an angle value, while $p = 12, \dots, 15$ is the modulus ratio. The vector p is the human description vector to depict human posture.

In the research of human behavior recognition, the behavior is usually represented by extracting the posture sequence in the key frame. A continuous posture sequence over a period of time is selected

to represent human action. For action A , $A = (A_1, \dots, A_n)$ in which A_1 is a human posture description vector.

The dynamic time warping (DTW) algorithm is a nonlinear dynamic warping technology that combines distance measurement with time warping method. By using DTW algorithm, it can easily implement the matching between templates with different lengths.

Let $M = \{M(1), \dots, M(m)\}$ represent reference template set, $T = \{T(1), \dots, T(n)\}$ represent text template set. The elements in M and T have the same length. The $d(M(i), T(j))$ represents the bias vector between vector $M(i)$ and vector $T(j)$ to calculate the distortion between $M(i)$ and $T(j)$. The $D(M(i), T(j))$ represents the cumulative distortion which calculates from $i = 1, j = 1$. Then, the $D(M(i), T(j))$ is written as following equation.

$$D(M(i), T(j)) = \sum_{p=1}^{p=j} d(M(p), T(j)) \quad (7)$$

From the Equation (7), it can be found that $D(M(i), T(j))$ is equivalent to calculating the cumulative distortion of a path connecting several eigenvector node pairs.

The purpose of DTW algorithm is to search an optimal path to ensure that $D(M(m), T(n))$ is minimized. The iterative process is implemented through the following equation.

$$D(M(i), T(j)) = d(M(i), T(j)) + D(M(i-1), T(j-1)) \quad (8)$$

In the Equation (8), $D(M(i-1), T(j-1))$ is obtained by the following equation.

$$D(M(i), T(j)) = \min \{D(M(i), T(j-1)), D(M(i-1), T(j-1)), D(M(i-2), T(j-1))\} \quad (9)$$

According to the characteristics of posture description vector, the following similarity calculation formula is adopted in this paper.

$$d(p, q) = \sum_{i=1}^n \frac{(p_i - q_i)^2}{p_i + q_i} \quad (10)$$

In the Equation (10), p and q are two eigenvectors without negative values. In this paper, $p, q \in \mathbb{R}^{15}$. The lower the $d(p, q)$ is, the higher the similarity is. When $d(p, q) = 0$, p complete matches q .

According to the above iterative process, the minimum cumulative distortion between the reference template and the test template can be obtained by repeating recursion from $M(1)$ and $T(1)$ to $M(n)$ and $T(m)$. If the test template is matched with all reference templates one by one, the test template belongs to the category with the smallest cumulative distortion

3. EXPERIMENTS AND SIMULATIONS

In this section, we first collect joint point data, vector angle and modulus ratio of skier in the process of skiing through Kinect camera to collect the regularization processed joint point data. In general,

even for the same behavior of the same person, the distribution of the change trajectory of human joint points in space does not have good aggregation. In addition, massive rotation and translation operations are carried out on the coordinate system, and the data distribution of related nodes does not have good aggregation. The variation trends of vector angle and modulus ratio are roughly the same, and their variation trajectories have good aggregation. The vector angle and modulus ratio obtained after regularization can better meet the invariance of translation and scaling. It provides a good basis for the design of classifier in the next step to recognize the postures of skiers.

The experimental environment of human posture and action recognition is set as follows: the human body faces Kinect camera to ensure that all joint points can be detected, and large-scale movement should be covered by the Kinect camera. We collect six kinds of skiing posture or action, including standing on the ground with skis, standing position on snowboard slope, straight skiing, plow brake, plough turn, and semi plow swing from 50 volunteers. Each volunteer repeats all actions 20 times. Thus, there are 6,000 actions. In each action, 100 sequences are used to establish matching pattern library, while other sequences are used as test set. The corresponding confusion matrix is reported in Table 2.

Table 2. The confusion matrix of the skiing posture and action

| | standing on the ground with skis | standing position on snowboard slope | straight skiing | plow brake | plough turn | semi plow swing | recognition rate |
|--------------------------------------|----------------------------------|--------------------------------------|-----------------|------------|-------------|-----------------|------------------|
| standing on the ground with skis | 869 | 26 | 1 | 0 | 0 | 0 | 96.99% |
| standing position on snowboard slope | 29 | 872 | 8 | 1 | 1 | 1 | 95.61% |
| straight skiing | 1 | 0 | 878 | 1 | 0 | 2 | 99.55% |
| plow brake | 1 | 0 | 10 | 881 | 17 | 9 | 95.97% |
| plough turn | 0 | 1 | 1 | 13 | 879 | 3 | 97.99% |
| semi plow swing | 0 | 1 | 2 | 4 | 3 | 885 | 98.88% |
| recognition rate | 96.56% | 96.89% | 97.56% | 97.89% | 97.67% | 98.33% | 97.48% |

When using DTW algorithm for template matching, the selection of reference template plays an important role in posture or action matching. From the result in Table 2, it can be found that recognition rate of the proposed method can achieve 96.56%, 96.89%, 97.56%, 97.89%, 97.67% and 98.33% for ground with skis, standing position on snowboard slope, straight skiing, plow brake, plough turn, and semi plow swing. The overall recognition rate achieve 97.48% for overall actions.

4. CONCLUSION

In this paper, we propose a framework to recognize the posture and action during skiing sports with the help of Kinect camera. First, the Kinect camera is used to collect joint points of skier's skeleton.

Second, the joint points of skier's skeleton data is represented as the human structure vector which consists of the human intermediate junction angles and modulus ratios. Compared with original joint point data, the intermediate junction angles and modulus ratios have better aggregation for the same behavior. The extracted features are used to construct a matching pattern library. The matching pattern library is used in dynamic time warping (DTW) algorithm to match the future gesture and action. The experimental results show that the proposed framework can identify more 97% skiing postures and actions. The framework can be further used to correct the skiing action and improve athlete's ability.

ACKNOWLEDGMENT

This research is supported by Social Science Project of Jilin Province Education Department in 13th Five-Year (No. JJKH20180395SK), Youth Teaching Reform in Beihua University in 2019 (No. XJQN2019032), and the Program for Promotion of Young Teachers in Beihua University.

REFERENCES

- Ahmed, M. D., & Leung, W. C. W. (2021). Using Wearable Devices to Enhance Quality of Physical Education for Students. *Strategies (La Jolla, Calif.)*, 34(2), 54–56. doi:10.1080/08924562.2021.1867458
- Alhammoud, M., Hansen, C., Meyer, F., Hautier, C., & Morel, B. (2020). On-field ski kinematic according to leg and discipline in elite alpine skiers. *Frontiers in Sports and Active Living*, 2, 56.
- Cai, D. (2019). Ecological Study of the Application of Flowers Plant Real-time Observation and 3D Reconstruction based on Kinect. *Ekoloji*, 28(107), 3143–3149.
- Han, S. H., Kim, H. G., & Choi, H. J. (2017, February). Rehabilitation posture correction using deep neural network. In *2017 IEEE international conference on big data and smart computing (BigComp)* (pp. 400–402). IEEE.
- Horak, F. B., Buchanan, J., Creath, R., & Jeka, J. (2002). Vestibulospinal control of posture. In *Sensorimotor control of movement and posture* (pp. 139–145). Springer. doi:10.1007/978-1-4615-0713-0_17
- Kashuba, V., Andrieieva, O., Goncharova, N., Kyrychenko, V., Karp, I., Lopatskyi, S., & Kolos, M. (2019). *Physical activity for prevention and correction of postural abnormalities in young women*. Academic Press.
- Li, N. F., & Xiao, Z. (2018). A Fire Drill Training System Based on VR and Kinect Somatosensory Technologies. *International Journal of Online Engineering*, 14(4).
- Li, Q., & Wang, S. (2020, August). Design of Tai-Chi Push-Hands Robot Control System and Construction of Visual Platform. In *2020 Chinese Control And Decision Conference (CCDC)* (pp. 2047–2052). IEEE.
- Li, Y., Chu, Z., & Xin, Y. (2020). Posture Recognition Technology Based on Kinect. *IEICE Transactions on Information and Systems*, 103(3), 621–630.
- Ma, T., & Guo, M. (2019, September). Research on Kinect-based Gesture Recognition. In *2019 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC)* (pp. 1–5). IEEE.
- Pang, X., & Liang, B. (2019, April). Motion recognition based on Kinect for human-computer intelligent interaction. *Journal of Physics: Conference Series*, 1187(3), 032028.
- Pellegrini, B., Stöggl, T. L., & Holmberg, H. C. (2018). Developments in the biomechanics and equipment of Olympic cross-country skiers. *Frontiers in Physiology*, 9, 976.
- Permanasari, Y., Harahap, E. H., & Ali, E. P. (2019, November). Speech recognition using dynamic time warping (DTW). *Journal of Physics: Conference Series*, 1366(1), 012091.
- Rentenberger, C., Salzmann, S. N., Shue, J., & Hughes, A. P. (2021). Heel Lift for Skiing to Compensate for Corrected Sagittal Vertical Axis After Spinal Surgery: A Case Report. *The International Journal of Spine Surgery*, 14(s4), S33–S36. doi:10.14444/7162 PMID:33900942
- Wang, L., Liu, J., & Lan, J. (2019). Feature Evaluation of Upper Limb Exercise Rehabilitation Interactive System Based on Kinect. *IEEE Access: Practical Innovations, Open Solutions*, 7, 165985–165996.
- Yadav, M., & Alam, M. A. (2018). Dynamic time warping (DTW) algorithm in speech: A review. *International Journal of Research in Electronics and Computer Engineering*, 6(1), 524–528.
- Yang, W., Peng, Y., & Xie, H. (2021). Action Recognition Based on Kinect Deep Learning. *Journal of Frontiers of Society. Science and Technology*, 1(2), 11–15.
- Ying, W., & Huang, W. (2021). Research on Methods of Improving Teaching Efficiency of Physical Training Based on Sports Biomechanics. *Curriculum and Teaching*, 4, 14–18.
- Yoshioka, S., Fujita, Z., Hay, D. C., & Ishige, Y. (2018). Pose tracking with rate gyroscopes in alpine skiing. *Sports Engineering*, 21(3), 177–188. doi:10.1007/s12283-017-0261-y