ISSN 2616-5775 Vol. 8, Issue 4: 33-39, DOI: 10.25236/AJCIS.2025.080404

A study on assisted swing sports training based on computer vision estimation algorithm of human three-dimensional posture with enhanced image details

Chaobing Yan[1,a], Ting Zheng[2,b,*]

[1]*Department of Teacher Education, Lishui University, Lishui, 323000, Zhejiang, China*
[2]*School of Ecology, Lishui University, Lishui, 323000, Zhejiang, China*
[a]*yanchaobing@lsu.edu.cn*, [b]*zhengting06@126.com*
[*]*Corresponding Author*

Abstract: *The human swing sport plays an important role in tennis, badminton, table tennis and other sports, and its training effect depends on accurate sports posture evaluation. Traditional sports training methods are difficult to acquire athletes' 3D posture data in real time and efficiently, while the development of computer vision technology provides a new solution for sports posture analysis. In this paper, a computer vision estimation algorithm for human 3D posture based on enhanced image details is proposed and applied to swing sports training. This paper investigates the key techniques of computer vision in human posture estimation and discusses the effect of enhanced image details on the accuracy of 3D posture estimation. In this paper, a human 3D pose estimation algorithm integrating image enhancement and deep learning is designed, including key aspects such as data preprocessing, model training and optimization. In this paper, a computer vision-assisted swing sports training system is constructed to realize real-time acquisition of sports data, posture analysis and feedback, and to improve the science and accuracy of sports training. The experimental results show that the method is better than the traditional method in terms of posture estimation accuracy and sports training assistance effect, which provides effective support for intelligent sports training.*

Keywords: *enhanced image details, human 3D posture estimation, computer vision, swing sports training, motion analysis*

1. Introduction

Swinging movement occupies an important position in tennis, badminton, table tennis and other competitive sports, and its training effect directly affects the athletes' performance [1]. The traditional swing training method mainly relies on the coach's experience and the athlete's self-perception, and lacks scientific data analysis, making it difficult to realize accurate technical optimization. The rapid development of computer vision technology provides new possibilities for sports training, especially the development of human three-dimensional posture estimation technology, which allows athletes to analyze their movements more accurately [2]. Existing human pose estimation algorithms still face many challenges in complex sports scenarios, such as the loss of image details, the effect of lighting changes, and partial limb occlusion, which lead to large errors in pose prediction, thus affecting the science and accuracy of sports training.

In the field of computer vision, human pose estimation is mainly categorized into 2D image-based pose estimation and 3D pose estimation [3]. Two-dimensional pose estimation algorithms, such as OpenPose and HRNet, can extract the key point information of the human body, but lack depth information, making it difficult to accurately analyze the spatial trajectory in the swing motion. On the other hand, three-dimensional pose estimation algorithms, VideoPose3D and VIBE, can reconstruct the three-dimensional skeleton structure of the human body, which is widely used in motion analysis and ergonomics research. Limited by the quality of input images, existing 3D pose estimation methods still suffer from large prediction errors in complex motion scenes [4]. In order to improve the accuracy of human pose estimation, in recent years, some researchers have tried to optimize the input image quality by combining super-resolution reconstruction, image enhancement and other techniques to improve the robustness and accuracy of pose estimation.

ISSN 2616-5775 Vol. 8, Issue 4: 33-39, DOI: 10.25236/AJCIS.2025.080404

For the demand of human 3D pose estimation in swing sports training, this paper proposes a computer vision estimation algorithm based on enhanced image details, and constructs an intelligent swing sports training assistance system. The method optimizes the input data through image processing techniques such as super-resolution reconstruction and edge enhancement to improve the accuracy of 3D pose estimation, and combines with deep learning methods to efficiently extract motion features to achieve accurate human pose prediction [5]. Based on this method, this paper develops a computer vision-assisted swing sports training system, which is capable of real-time acquisition of sports data, providing feedback analysis, helping athletes optimize their technical movements, and improving the scientific and personalized level of training [6]. The results of this study not only promote the application of computer vision in the field of sports training, but also provide theoretical support and technical support for the development of intelligent motion analysis and personalized training programs.

2. Three-dimensional human posture estimation based on computer vision

Computer vision technology has been widely used in the field of human posture estimation, especially in sports training, motion capture and intelligent monitoring [7]. Human 3D posture estimation aims to recover the joint point information of human body in 3D space from 2D images or videos in order to accurately analyze the motion trajectories and movement patterns. Traditional pose estimation methods mainly rely on sensors or marker point tracking, but these methods are more costly and restrictive in use, making it difficult to meet the needs of daily sports training [8]. With the rapid development of deep learning, markerless 3D posture estimation methods based on computer vision have gradually become a research hotspot, capable of realizing high-precision posture reconstruction with the support of common camera equipment. Mean Per Joint Position Error (MPJPE):

$$\mathrm{MPJPE} = \frac{1}{N}\sum_{i=1}^{N}|\widehat{J_i} - J_i|_2 \tag{1}$$

Current human posture estimation algorithms are mainly categorized into monocular vision and multicameral vision based technical solutions. Monocular vision methods utilize deep learning models to predict the 3D coordinates of key points of the human body from a single image or video frame, and typical methods include VideoPose3D and VIBE [9]. Due to the lack of real depth information, monocular pose estimation is susceptible to viewpoint changes and occlusion in complex environments [10]. Multi-ocular vision methods, on the other hand, acquire image data of the human body from different angles by multiple cameras and use triangulation and other methods to reconstruct the 3D posture, which has higher accuracy, but its hardware deployment and data synchronization requirements are high, and it is not applicable to general sports training scenarios.

Image quality has an important impact on the accuracy of human 3D pose estimation, and factors such as low resolution, motion blur, and illumination variations may lead to increased prediction errors at key points. In recent years researchers have begun to focus on the application of image enhancement techniques in pose estimation. Methods such as super-resolution reconstruction, edge detection, and contrast enhancement have been used to optimize the quality of the input image so that the model can more accurately extract human keypoint information. Some deep learning methods also incorporate self-attention mechanisms to enhance the ability to capture motion details and improve the generalization performance of the model in different environments.

In the field of sports training, computer vision technology not only realizes high-precision 3D posture estimation, but also provides real-time feedback and motion correction. The use of posture estimation technology can automatically analyze the swing trajectory, postural stability and power distribution of the athlete, thus providing trainers with visualized movement data to help them adjust their movements and improve training efficiency. With the further development of computer vision and artificial intelligence technology, human 3D posture estimation will play a more important role in the fields of sports science, rehabilitation medicine and intelligent interaction.

3. Human posture estimation algorithm based on enhanced image details

In order to improve the accuracy and robustness of human 3D pose estimation, this paper proposes a pose estimation algorithm based on enhanced image details. The algorithm improves the image quality by performing various enhancement processes on the input image, such as super-resolution reconstruction, edge enhancement, and contrast enhancement, which in turn improves the recognition accuracy of the pose estimation model for the key points. The key aspects of this algorithm include data

ISSN 2616-5775 Vol. 8, Issue 4: 33-39, DOI: 10.25236/AJCIS.2025.080404

preprocessing and image enhancement, design of the human 3D pose estimation model, and model optimization and performance evaluation. Through these steps, the estimation accuracy can be effectively improved and the error caused by image quality difference or motion blur can be reduced, so as to realize more accurate pose estimation in complex motion scenes.

3.1. Data Preprocessing and Enhancement

Data preprocessing and enhancement are key steps to improve the accuracy of human 3D pose estimation, especially in motion scenes, where the improvement of image quality has a direct impact on the subsequent estimation results. In order to ensure that the data have high quality when they are input into the estimation algorithm, this paper first performs a series of pre-processing on the original image, including denoising, normalization and data enhancement. These processes can effectively improve the noise interference in the image, enhance the model's ability to recognize the joints, and provide a reliable data base for the subsequent 3D pose estimation.

Denoising is an important part of data preprocessing. Due to camera jitter, lighting changes or sensor noise during motion, the original image often contains a lot of noise, resulting in inaccurate joint point localization. In order to remove the noise, this paper adopts classical denoising techniques such as Gaussian filtering, mean filtering, etc., or combines the denoising self-encoder model with deep learning in order to extract clearer features from the image. The denoised image provides cleaner visual data for subsequent processing, which helps the algorithm to more accurately detect the key points of the human body. Loss Function for 3D Pose Estimation :

$$\begin{aligned}\mathcal{L} = \lambda_1 &\sum_{i=1}^{N} |\widehat{J_i^{2D}} - J_i^{2D}|_2 \\ +\lambda_2 &\sum_{i=1}^{N} |\widehat{J_i^{3D}} - J_i^{3D}|_2 \\ +\lambda_3 &\textstyle\sum_{i=1}^{M} R_m \end{aligned} \tag{2}$$

Normalization is an important step in the image enhancement process. The brightness, contrast and color distribution of an image may differ significantly under different shooting angles and lighting conditions. In order to make the input data have consistent features, this paper adopts the image normalization technique, which scales the pixel values of the image to a uniform range and adjusts the contrast through histogram equalization to enhance the detail performance of the image. This process helps to eliminate the interference of environmental factors, allowing the pose estimation model to focus more on the human motion features rather than being affected by background variations.

Data enhancement techniques are widely used to improve the generalization ability of models. In order to enhance the model's adaptability to different motion scenes and viewpoints, this paper introduces a variety of data enhancement techniques, such as random rotation, translation, scaling, and mirror flip. These operations can generate diverse training samples by simulating different shooting angles and environment changes, thus enhancing the robustness of the model to image deformation, occlusion and other problems. This paper also combines the image super-resolution reconstruction technique to reconstruct the low-resolution image with high quality, which further improves the detail performance of the image and provides clearer visual information for the pose estimation Through this series of data preprocessing and enhancement techniques, the image quality is significantly improved, which provides more accurate and robust data support for the subsequent human 3D pose estimation. These processes not only improve the performance of the estimation algorithm, but also enhance the model's ability to adapt to different training scenarios, thus realizing more accurate pose prediction in complex environments, showed in Figure 1 :

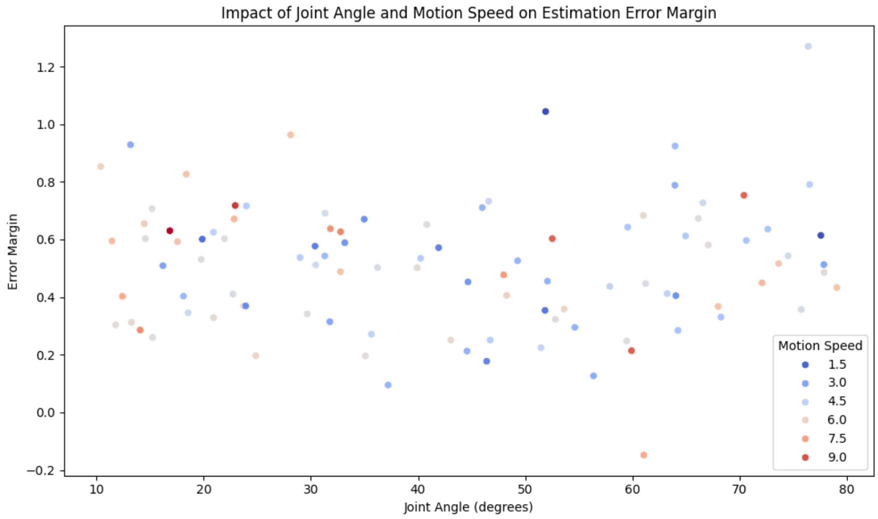

Figure 1 Impact of Joint Angle and Motion Speed on Estimation Error Margin

3.2. Human 3D posture estimation model

The human 3D pose estimation model is an important research direction in the field of computer vision, and its core objective is to accurately predict the 3D coordinate information of key points of the human body in 2D images or video sequences. Existing pose estimation methods mainly rely on deep learning models to extract human motion features through convolutional neural networks (CNN), graph neural networks (GNN), or self-attention mechanisms (Transformer), and use regression or optimization methods for 3D pose reconstruction. In order to improve the accuracy of pose estimation in swing motion training, this paper designs a deep learning model incorporating augmented image details, which is capable of achieving highly robust human pose recognition in complex motion scenarios.

The model employs a two-stream network architecture in order to extract local and global motion features simultaneously. One stream employs a CNN-based feature extraction module for capturing local detail information of human joints, while the other stream is based on a Transformer or Graph Convolutional Network (GCN) for modeling spatial relationships and temporal dependencies between key points of the human body. Through the fusion of the dual-stream networks, the model can effectively learn the hierarchical features of human posture, thus improving the adaptability to complex movements. In addition, for the high-speed motion characteristics in swinging motion, the model specifically optimizes the time series modeling capability to reduce the short-term instability problem in posture prediction.

To further improve the accuracy of 3D pose estimation, this paper introduces a multi-scale feature fusion mechanism into the network. Through the feature extraction modules at different scales, the model is able to pay attention to the overall structure of the human body and the detailed changes of the local joint points at the same time, which ensures that the pose can be accurately predicted under various motion states. Combined with the image super-resolution enhancement technique, the model can recognize the key point features more clearly, which ensures high prediction accuracy even under low-resolution or complex lighting conditions.

In order to optimize the training process of the model, a multi-task loss function is designed in this paper, which integrates the 2D keypoint prediction loss, 3D pose regression loss and physical constraint loss. The 2D key point loss is used to guide the model to accurately locate the joints in the image, the 3D pose regression loss is used to optimize the 3D reconstruction accuracy, and the physical constraint loss is used to ensure the reasonableness of the human body pose and avoid producing prediction results that do not conform to the physiological structure. The experimental results show that the model achieves good estimation results in several public datasets and swing motion training scenarios, providing a reliable technical support for intelligent motion training.

3.3. Model Optimization and Performance Evaluation

In order to improve the accuracy and robustness of the human 3D pose estimation model, this paper

ISSN 2616-5775 Vol. 8, Issue 4: 33-39, DOI: 10.25236/AJCIS.2025.080404

carries out various optimizations during the model training process. In order to improve the stability and convergence speed of the training process, this paper adopts an advanced optimization algorithm, Adam optimizer, and combines it with the learning rate decay mechanism to ensure that the model can converge more accurately at the late stage of training. In addition, by introducing techniques such as Batch Normalization and Residual Connections, the gradient vanishing problem during training is effectively reduced, allowing the network to be trained deeper and faster, thus improving the overall model performance,showed in Figure 2 :

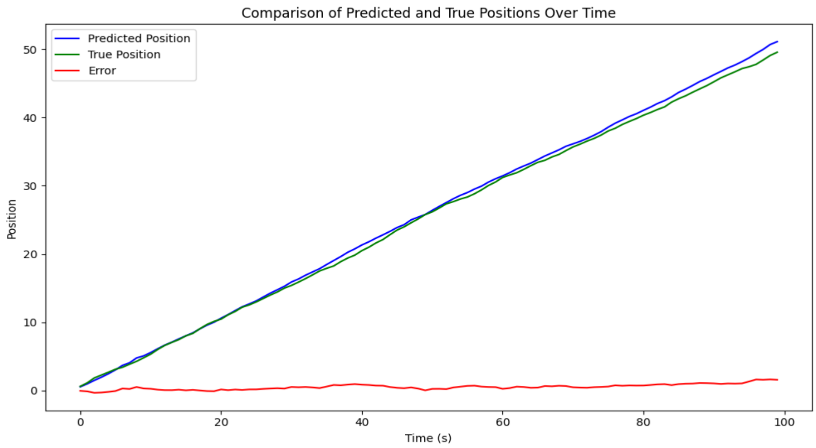

Figure 2 Comparison of Predicted and True Positions Over Time

In this paper, the data enhancement techniques are optimized and various means are used to expand the training data. These enhancement techniques include common methods such as random rotation, panning, cropping, scaling, etc., which are designed to simulate different shooting angles, background changes, and motion speeds. By diversifying the training dataset, the model is able to learn to handle more complex motion scenarios, reduce the risk of overfitting, and improve the generalization ability of the model. For the common occlusion problem in swinging motion, this paper specially designs an occlusion compensation mechanism, which generates a reasonable supplementary image of the occluded region by introducing a generative adversarial network (GAN), so that the model can still accurately predict the pose in the face of partial occlusion.

In terms of performance evaluation, this paper adopts several common evaluation metrics, including mean joint positioning error (MPJPE), relative position error (PCK), and 3D reconstruction error, to comprehensively assess the effectiveness of the model. These metrics quantify the performance of the model in different dimensions such as key point localization accuracy, joint point relationship preservation and 3D reconstruction accuracy, respectively. The experimental results show that the optimized model exceeds the existing mainstream algorithms in these metrics, especially in the complex motion state, and is able to effectively reduce the pose estimation error and improve the overall pose prediction capability.

To ensure the robustness of the model, this paper also validates it in several real-world scenarios. The adaptability of the model under different conditions was verified by testing the model in a number of different motion types (swinging, running, jumping, etc.) as well as in different environments (indoor light changes, background complexity, etc.). The model is still able to maintain high prediction accuracy in the face of multiple complex environments, providing stable technical support for swing sport training. In addition, this paper further validates the superiority of the optimized model through comparison experiments with other existing methods, demonstrating the enhancement of 3D pose estimation accuracy by enhanced image details.

4. Computer vision-assisted swing motion training system

The computer vision-assisted swing sports training system aims to utilize human 3D posture estimation technology to capture and analyze athletes' swing movements in real time and provide accurate feedback and improvement suggestions. The system combines a high-precision pose estimation model with sports training requirements, and is capable of comprehensively monitoring the athletes' trajectories and movement details. Supported by image enhancement and deep learning models, the system is able to operate stably in complex environments and provide athletes with real-time,

ISSN 2616-5775 Vol. 8, Issue 4: 33-39, DOI: 10.25236/AJCIS.2025.080404

personalized training guidance, thus significantly improving training efficiency and science.

The system captures the athletes' swing movements in real time by configuring multiple HD cameras. These cameras are reasonably arranged to ensure that they can record the athlete's body posture in all directions and provide stable image data input. The image data is processed by a computer vision algorithm to extract the coordinates of the athlete's 3D key points, which include the shoulders, elbows, wrists and other important parts. Based on the 3D pose estimation model with enhanced image details, the system is able to ensure high quality pose estimation results by reducing the effects of image noise and blurring while ensuring motion accuracy.

The system utilizes the motion data analysis module to evaluate the athlete's swing posture in real time. By comparing the athlete's current posture with the optimal movement pattern, the system provides an accurate posture score and makes targeted adjustment suggestions. The system is able to detect deviations in the position of key points during the athlete's swing, such as excessive elbow bending angle and insufficient shoulder rotation, and provide timely feedback on these deviations and suggest the athlete to adjust his/her movement. Through the detailed analysis report provided by the system, coaches and athletes can quickly identify technical shortcomings and make targeted adjustments and improvements.

The system also integrates a multi-dimensional training feedback mechanism. In the swing training process, in addition to posture analysis, the system also combines a variety of parameters, such as movement speed and swing trajectory, to conduct a comprehensive assessment. The system can calculate the speed, angle, power distribution and other information of the athlete's swing to provide trainers with comprehensive sports data support. In addition, the system can generate the historical data trend graph of the athlete, which helps the coach to track the training progress of the athlete and make personalized training plan. These features make training more intelligent and personalized, and can provide tailored guidance programs according to the specific needs of athletes.

The system provides remote monitoring and analysis functions through the interconnection of the cloud platform and mobile devices. Athletes and coaches can view training data and feedback through cell phones or tablet devices at any time, and adjust training strategies at any time. The system also supports the switching of training modes for a variety of sports types, including the analysis of swing movements for different sports such as tennis and badminton. Through the cloud platform, coaches can obtain data from multiple athletes in real time and conduct comprehensive analysis to provide scientific training decision support for the team. This convenient system deployment and real-time feedback function greatly improves the training efficiency and intelligence level, and promotes the development of traditional sports training in the direction of intelligent training.

5. Conclusion

In this paper, a computer vision estimation algorithm for human 3D posture based on enhanced image details is studied and applied to assisted swing sports training. The quality of the input image is improved by image enhancement techniques, which enables the 3D pose estimation model to extract human key point information more accurately. A posture estimation model combining deep learning and physical constraints is designed to achieve high-precision human posture recognition in complex motion environments. By optimizing the data preprocessing, network architecture and loss function, the robustness and generalization ability of the model are effectively improved.

The algorithm proposed in this paper shows high accuracy of pose estimation in a variety of sports scenarios, especially in swing sports training, which can accurately capture the details of athletes' movements and provide real-time technical analysis and feedback. Compared with traditional methods, the present algorithm can better adapt to complex environments, improve the stability of key point detection, and effectively reduce the errors due to motion blur or occlusion. By comprehensively evaluating the performance of different models, the role of enhanced image details in improving the accuracy of 3D pose estimation is verified.

The computer vision-assisted swing sports training system constructed based on this algorithm is able to analyze athletes' swing postures in real time and provide scientific evaluation and personalized guidance for sports training. The system can not only accurately record key sports parameters, but also provide data visualization support through cloud analysis, which provides a strong basis for coaches and athletes to formulate training plans. The system can further integrate multimodal sensing data, such as inertial sensors and EMG signals, to enhance the comprehensiveness and accuracy of motion analysis.

ISSN 2616-5775 Vol. 8, Issue 4: 33-39, DOI: 10.25236/AJCIS.2025.080404

The 3D pose estimation algorithm based on enhanced image details and its application in sports training proposed in this paper provides a new technical solution for intelligent motion analysis. Future research will continue to optimize the adaptability of the model in real-world scenarios and explore methods such as multi-view fusion and cross-modal data integration to further enhance the value of human pose estimation in sports training, rehabilitation medicine and other fields.

Acknowledgements

Supported by the Open Project Program of the State Key Laboratory of CAD&CG (Grant No. A2331), Zhejiang University

References

[1] Tran C , Trivedi M M .3-D Posture and Gesture Recognition for Interactivity in Smart Spaces[J]. IEEE Transactions on Industrial Informatics, 2012, 8(1):178-187.DOI:10.1109/TII.2011.2172450.
[2] Bilal S , Akmeliawati R , Shafie A A ,et al.Hidden Markov model for human to computer interaction: a study on human hand gesture recognition[J].Artificial Intelligence Review, 2013, 40(4):495-516. DOI:10.1007/s10462-011-9292-0.
[3] Tsai M J , Lung H Y .Two-phase optimized inverse kinematics for motion replication of real human models[J]. Journal of the Chinese Institute of Engineers, 2014, 37(7):899-914. DOI:10.1080/02533839.2014.904474.
[4] Murugan A S , Noh G , Jung H ,et al.Optimising computer vision-based ergonomic assessments: sensitivity to camera position and monocular 3D pose model[J].Ergonomics, 2025, 68(1):120-137. DOI:10.1080/00140139.2024.2304578.
[5] Bandouch J , Jenkins O C , Beetz M .A Self-Training Approach for Visual Tracking and Recognition of Complex Human Activity Patterns[J].International Journal of Computer Vision, 2012, 99(2):166-189. DOI:10.1007/s11263-012-0522-y.
[6] Aminian Modarres A F , Soryani M .Body posture graph: a new graph-based posture descriptor for human behaviour recognition[J].Iet Computer Vision, 2013, 7(6):488-499.DOI:10.1049/iet-cvi.2012.0121.
[7] Chwa D .Robust Distance-Based Tracking Control of Wheeled Mobile Robots Using Vision Sensors in the Presence of Kinematic Disturbances[J].IEEE Transactions on Industrial Electronics, 2016, 63(10):1. DOI:10.1109/TIE.2016.2590378.
[8] Trigueiros P , Ribeiro F , Reis L P .Generic System for Human-Computer Gesture Interaction: Applications on Sign Language Recognition and Robotic Soccer Refereeing[J].Journal of Intelligent & Robotic Systems, 2015, 80(3):573-594.DOI:10.1007/s10846-015-0192-4.
[9] Cho S Y, Wong J J. Human face recognition by adaptive processing of tree structures representation[J]. Neural Computing & Applications, 2008, 17(3):201-215.DOI:10.1007/s00521-007-0108-8.
[10] Hao Y, Guo S, Zhou X Y H .Underwater swimming fish mass estimation based on binocular vision[J]. Aquaculture International, 2024, 32(6):7973-7995.DOI:10.1007/s10499-024-01550-z.