DOI: 10.18686/ahe.v6i15.5196

AI Body Detection and Teaching System based on Mediapipe Machine Learning Platform and OpenCV Computer Vision Library

Ling Li[1], Huijuan Huang[1], Shaogeng Zeng[234*], Huiqi Cao[1], Rongrui Zheng[1], Shuimei Lin[5]

1. Lingnan Normal University School of Electronic and Electrical Engineering, Guangdong Zhanjiang 524048

2. Guandong Provincial Key Laboratory of Development and Education for Special Needs Children, Guangdong Zhanjiang 524048

3. Lingnan Normal University School of Computer and Intelligent Education, Guangdong Zhanjiang 524048

4. Guangdong University Digital Learning Engineering Technology Development Center, Guangdong Zhanjiang 524048

5. Lingnan Normal University School of Business, Guangdong Zhanjiang 524048

Abstract: To solve the problems of low intera ctivity, high cost, large amount of data, "difficult to quantify, difficult to record, difficult to supervise, difficult to analyze" of human motion detection correction devices on the market today, we designed an intelligent device based on Mediapipe machine learning platform and OpenCV computer based on Raspberry Pi, camera and display. We designed an intelligent device for AI body detection and teaching based on Mediapipe machine learning platform and OpenCV computer vision library. By combining chip, sensor, computing platform and technology level of computer vision, speech recognition and machine learning, the device can capture human movement in real time by using camera equipment, judge the accuracy and completeness of user's movement according to the comparison of standard movement, and give feedback to the user in real time by voice broadcast and image prompt. The test results show that the device has the advantages of low cost, simple structure, intelligence, unmanned, data and accuracy, which provides a feasible solution to further enhance the convenience and accuracy of unmanned movement teaching and rehabilitation training.

Keywords: Mediapipe; OpenCV; Rehabilitation training; Unmanned movement teaching

Project fund: Guangdong College Students' innovation and entrepreneurship training program in 2021, Project No.: Provincial Project No.25 of Lingnan Normal University, Project name: Flying shadow - JiangYiXinCheng studio.

Introduction

With the rise of the Internet of Things and artificial intelligence technology, the market for intelligent products is developing rapidly, and situational interactive virtual environment technology and human posture estimation technology at home and abroad are booming stages of development, and there are some physical detection devices with greater feasibility, such as the IREX situational interactive rehabilitation training system of the United States, ReoG Shangzhi rehabilitation robot, Huawei wisdom screen, billion key magic mirror, KEEP, etc. However, the above-mentioned domestic and foreign situational interactive action training systems are not perfect in terms of real-time interaction experience and humanization, and most of them only achieve action detection, with few real-time action guidance and correction functions, which are "difficult to quantify, record, monitor and analyze". It is difficult to be applied in the field of daily life. Therefore, there is an urgent need for an intelligent, unmanned, data-driven and accurate AI body detection and teaching device to solve the above problems.[1]

1. System Overview

AI body detection and teaching equipment can capture human movements in real time through the camera equipment, and can judge the accuracy and integrity of the user's actions according to the comparison of standard actions, and give real-time feedback to the user through voice broadcasting and image prompts. The system can be used in dance, yoga, martial arts and other movement teaching fields and rehabilitation training fields, can realize real-time detection and the human movement, further improve the standardization and guidance of movement teaching, to achieve visual, unmanned standard movement teaching.

2. System Design

2.1 Overall design

The system is a "deep learning" based body posture detection correction device, and through the use of Mediapipe based machine learning platform and OpenCV computer vision library for research and development, by the Raspberry Pi development board, atomic mirror, display, camera and speaker five main components.The atomic mirror and the display facilitate users to view their own movements and standard movements to compare the difference, Raspberry Pi as the core master of the project is used to collect and process a large amount of data, using the camera to collect the user's action posture pictures. The process is that the camera collects the user's action posture picture in real time, uses the image processing technology of Raspberry Pi for pre-processing (rotation operation, noise reduction operation, brightness operation, fuzzy processing), generates real-time human key node coordinate data according to the trained human body analysis network model, and outputs the human key node map on the display after calculation, and the voice broadcast prompts the user to correct the action to achieve better.

Based on the standard action demonstration video, the frame maps of key actions are processed and marked, and the algorithm is trained to obtain the key nodes of the human body for each key frame map. The video information of user's body posture is recorded by camera in real time, and GRB image and grayscale image are obtained by OpenCV pre-processing. The processed images are passed through Mediapipe machine learning platform to achieve real-time recognition of multi-human key points in 2D or 3D, predict the confidence S and affinity vector L of each key point, and compare the generated coordinate data of key nodes with the existing standard coordinate

data by frame, and use the formula dx, y,

$$z = (x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2$$

The calculated coordinate offset automatically judges whether the user's action is standardized in place, displays the degree of matching between the user's real-time action and the standard action, and gives real-time action correction suggestions. This project has the features of function innovation, easy operation and data stability, and also has high practicality and accuracy, which can be used as the testing and correction equipment for various sports movements such as dance and fitness.

2.2 Attitude Landmark Model Training and Data Acquisition

The system is based on Mediapipe machine learning platform as a model management and migration learning tool, which operates on three aspects of facial recognition, extraction of key point coordinates of body pose and gesture tracking, and then analyzes and judges the real-time pose and action coordinate data of a specific user. Mediapipe provides ML Kit pose detection API, which can accomplish the fast migration learning of deep learning models in natural language processing and computer vision scenarios, and complete the training of models in a shorter time.

2.3 Realize real-time pose image frame pre-processing

The framework based on Mediapipe machine learning platform can realize connecting the key points and limbs belonging to the same person to form a skeleton map of each person, using a certain number of videos of standard actions as a sample set, and getting the body posture data through the human body parsing network model that has completed pre-training. According to the standard action key nodes demonstrated by the body posture detection correction equipment compared with the user's human skeleton key nodes detected by the Mediapipe machine learning platform algorithm, automatically determine whether the user's action is standardized in place, and show the action correction suggestions on the display in real time.[2]

2.4 Normative posture discrimination methods and processing measures

This system realizes the action comparison by calculating the coordinates of key nodes. The head and torso joints, the center of the hip and the right wrist are selected as the main key nodes. After detecting the coordinate data of each key node according to the model, the distance between the key nodes and the angle of the line connecting any joint point with the other two joints are calculated according to the following equations (1) and (2), respectively.

The distance dx, y, z between two joints is calculated by the formula:

$$dx,y,z=\sqrt{(x_i - x_j)^2 + (y_i - y_j)_2 + (z_i - z_j)^2} \quad (1)$$

The formula for calculating the angle θ between the line connecting any one joint and the other two joints:

$$\theta=\arccos\frac{d_1^2 + d_2^2 - d_3^2}{2d_1 * d_2} \quad (2)$$

Where (xi, xj), (yi, yj), (zi, zj) are the coordinates of key nodes, θ is the angle between any joint in a group and the line connecting two numbered adjacent joints.

Specify that when the percentage deviation of coordinate data of key nodes is greater than 20%, the part with larger deviation is given with text and voice prompts for the purpose of action correction. Calculate the percentage of angular deviation between the adjacent key node linkage of a real-time detected action and the adjacent key node linkage of a standard action P:

$$P=1-\frac{|\theta_{current} - \theta_{standard}|}{\theta_{standard}} * 100\% \quad (3)$$

in the equation $\theta_{current}$ the action angle detected in real time, the $\theta_{current}$ standard action angle.

3. Summary

Based on Mediapipe machine learning platform and OpenCV computer vision library, and relying on the powerful motion capture ability and information processing ability of AI posture real-time teaching system, the device can accurately identify human posture and realize real-time detection and correction of human actions. The existing technology can be used to analyze and study the movement teaching field, which can help traditional manual teaching. The test results show that the performance of the system is stable, which further improves the standardization and guidance of movement teaching, and realizes the visual and unmanned standard movement teaching.

References:

[1] Wang Rubin, DOU Quanli, Zhang Qi, ZHOU Cheng. Gesture Recognition based on MediaPipe for Remote Operation Control of Excavator [J]. Information technology in civil engineering and construction: 1-8 [2022-04-23]. HTTP: / / http://kns.cnki.net/kcms/detail/11.5823.TU.20211125.1918.024.html

[2]Mei Zaixia, Yin Chun, ZHANG Lei. 6G Vision, Application Scenarios and key technologies analysis [C]//.Push the evolution of Promote the application of innovation - 5 g network conference (2021), vol., 2021:387-390. The DOI: 10.26914 / Arthur c. nkihy. 2021.039175.