

Body Turning Detection Method Based on AI Image Recognition for VR Space Motion

Zhuo Liu

Guangzhou Zhuoyuan Virtual Reality Technology Co., Ltd., Guangzhou, Guangdong, 511450, China

Abstract

In VR+sports application, to provide a good experience, it is necessary to accurately map real human movement actions to virtual characters in the VR space. The current action acquisition and recognition technology is not user-friendly enough. AI action recognition technology based on video images provides new ideas for consumer-level action acquisition. Aiming at the demand for users' body turning perception in the scenarios of miniaturization and limited space of the VR dynamic platform, this paper designs a specific image acquisition scheme, uses MediaPipe to extract the position coordinates of key points of the human body to calculate the rotation angle, adopts low-pass filtering to optimize the continuous body turning angle data, and conducts control optimization. Experimental results show that this method can run stably in the existing product hardware experimental environment and meet the current VR experience requirements.

Keywords

VR Experience; Body Turning; MediaPipe; Pose Recognition

基于 AI 图像识别的 VR 空间身体转向检测方法

刘卓

广州卓远虚拟现实科技股份有限公司, 中国·广东 广州 511450

摘要

为了给VR+体育运动的用户提供良好体验,需将用户的真实人体运动动作精准映射到VR空间虚拟角色。当前动作获取识别技术易用性不足,基于视频图像的AI姿态识别技术为消费级动作获取提供新思路。针对VR动感平台小型化和空间有限场景中使用者身体转向感知需求,本文设计特定图像采集方案,利用MediaPipe提取人体关键点位置坐标计算旋转角度,采用低通滤波优化连续体转角度数据,并进行控制优化。实验结果表明,该方法在现有产品硬件实验环境下能稳定运行,满足现有VR体验需求。

关键词

VR体验; 身体转向; MediaPipe; 姿态识别

1 背景与现状

在VR+体育运动的发展中,如何准确获取和识别现实空间中的人体运动动作并精准映射到VR空间中的虚拟角色是提供好的体验的基础。目前的动作获取识别技术基本上都需要在用户穿戴传感器或者标志物来辅助识别,这对于面向普通用户消费级产品来说易用性不强。随着人工智能技术的发展,基于视频图像的AI动作识别技术进入了工程应用阶段,为非接触式、无干扰的消费级动作获取和识别提供了新的思路 and 手段。Google开发的跨平台多媒体处理框架

MediaPipe 是其中的代表性开源 AI 工具平台。MediaPipe 人体姿态检测模型设计用于实时处理,可以在高帧率下进行快速姿态检测,支持跨多种平台,包括 PC、移动设备和嵌入式设备,使得人体姿态检测可以在不同设备上应用。使用 MediaPipe 的人体姿态检测模型,无需从头开始训练自己的模型,可以直接使用预训练的模型和 API 进行快速开发。针对上述行业共性难题,本研究依托广州市番禺区创新领军团队重点支持项目《智能 VR 交互设备关键技术的研发与产业化》,即采用 MediaPipe 的人体姿态估计模型,针对 VR 动感平台上使用者身体转向感知的场景需求,提出一种基于机器学习图像识别的 VR 空间身体转向检测方法,为使用者带来更好的 VR 空间运动体验。

【基金项目】番禺区创新领军团队项目《智能 VR 交互设备关键技术的研发与产业化》(项目编号: 2021-R01-5)。

【作者简介】刘卓(1980-),男,中国河北行唐人,博士,从事建模仿真与人工智能研究。

2 方法思路

针对设备自身小型化和放置空间有限的具体应用场景需求,在采集运动人体图像时,无法像普通 MediaPipe 应用

的摄像头设置方式一样在较大范围内正面采集图像，本文专门设计了适用于应用场景的图像采集解决方案，在获取图像后根据实际场景特点设计了专门计算方法，确保能够准确检测识别身体转向的量化数值。

2.1 检测转向角度

在平台后端部分安装视频摄像头，摄像头自下向上实时采集真实运动人体图像，在前端部分安装有一定偏移角度的反光镜，在垂直地面方向保证摄像头视野在反光镜内能够实时地、清晰准确地采集完整的上半身图像，在平行地面方向确保当人体自然竖直站立是，获取的上半身图像肩部处于水平状态，镜像摄像头与地面水平夹角为。摄像头刷新性能不低于 30 帧/s，获取的图像像素分辨率不低于 640×480。

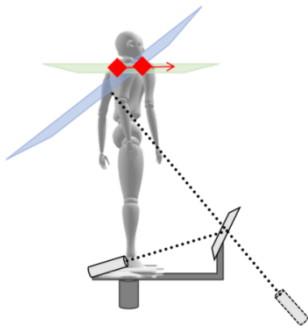


图 1 检测场景与方法示意

对采集到的视频图像进行逐帧处理，利用开源的 MediaPipe 通过 MediaPipe Pose Marker，提取图像中人体的关键点位置坐标，包括：头部中心点 $head_center(x,y)$ ，左右肩部点 $left_shoulder(x,y)$ 、 $right_shoulder(x,y)$ ，左右髋部点 $left_hip(x,y)$ 、 $right_hip(x,y)$ 等关键点。依据关键点位置，利用以下方法求出旋转的角度 $twist_angle$ ：

(1) 当 $x_{left_shoulder} = x_{right_shoulder}$ 时，

$$twist_angle = \begin{cases} & 90, & \text{当 } y_{right_shoulder} < x_{right_shoulder} \\ & -90, & \text{当 } y_{right_shoulder} > x_{right_shoulder} \end{cases}$$

(2) 当 $x_{left_shoulder} \neq x_{right_shoulder}$ 时，

$$twist_angle = \arctan\left(\frac{y_{right_shoulder} - y_{left_shoulder}}{x_{right_shoulder} - x_{left_shoulder}} \cdot \frac{1}{\sin\alpha}\right)$$

计算得到转体角度后再通过控制优化，最终控制平台进行相应的旋转运动。

2.2 控制数据优化

在基于 MediaPipe 的人体姿态识别系统中，图像采集和姿态解算过程不可避免地受到环境噪声、帧率波动、关键点定位误差的影响，导致最终输出的转体角度数据存在高频抖动现象，直接影响 VR 动感平台对使用者动作响应的平滑性，如果进行有效优化，将造成设备控制信号不稳定，进而影响沉浸式体验质量。当前平台采用 50ms 的控制更新周期（即 20Hz 刷新频率），而人体运动的数据变化速率远高于该控

制采样频率，尤其在快速转身或连续变向的情况下，原始数据会出现明显的瞬时偏差与跳变情况。为解决上述问题，本文引入低滤波算法，动态处理旋转角度序列，通过对历史角度数据加权平均，抑制短时间内的异常波动，提升控制系统对于身体转向的响应稳定性。系统采用一阶 Butterworth 低通滤波器，其差分形式如下：

$$\theta_{filtered}[n] = \alpha \cdot \theta_{raw}[n] + (1 + \alpha) \cdot \theta_{filtered}[n-1] \quad (1)$$

其中， $\theta_{filtered}[n]$ 是当前时刻的滤波后角度值， $\theta_{raw}[n]$ 是当前原始角度输入， α 为滤波系数 ($0 < \alpha < 1$)，通过实验选取最优值，科学平衡滤波效果。

针对图像采集过程中遮挡引发的关键点丢失问题，设计了基于线性插值的角度补偿策略。当连续 k 帧内某关键点缺失时，采用前后有效帧的角度值进行插值估计：

$$\theta_{interpolated}[t] = \frac{(t_2 - t)\theta[t_1] + (t - t_1)\theta[t_2]}{t_2 - t_1} \quad (2)$$

其中， t_1 和 t_2 分别表示缺失前后的有效帧时刻， $\theta[t_1]$ 和 $\theta[t_2]$ 为其对应角度值， t 为当前缺失帧时刻。

系统采用一阶 Butterworth 低通滤波器，其截止频率依据人体动作的典型频域特性设定，在滤除高频干扰的基础上，不会引入过大的相位延迟（见图 2）。滤波参数通过多轮实验验证，在不同运动速度下选取最优衰减系数，并结合滑动窗口机制对数据流进行局部均值平滑处理。针对图像采集过程中遮挡引发的关键点丢失问题，设计基于线性插值的角度补偿策略，使平台在短暂信息缺失情况下仍能保持合理姿态预测。

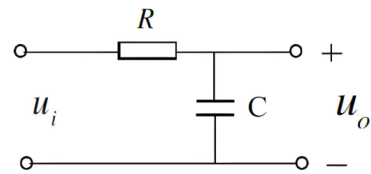


图 1 一阶 Butterworth 低通滤波器

3 实验结果

本研究在 Intel Core i5-1135G7 @ 2.40GHz、集成显卡及 Windows 10 操作系统环境下构建测试平台，运行基于 MediaPipe 姿态估计模型的检测程序，实测系统资源占用约为 450MB 内存，CPU 使用率维持在 35% 以内，GPU 负载较低，符合轻量化边缘计算部署要求。从图像采集到姿态解算的整体流程可在满足 640×480 分辨率与 30fps 帧率条件下稳定运行，实际处理效率达到 20fps，足以支撑 VR 动感平台对身体转向动作的实时捕捉与反馈需求。测试数据显示，系统在常规光照条件下的姿态识别精度误差控制在 ±3° 以内，且在连续转身动作中具有良好的轨迹一致性，没有出现明显漂移现象。

图 1 展示了系统在不同转向状态下的检测效果，左

侧为摄像头采集的真实人体图像，右侧为 MediaPipe Pose Marker 提取的关键点分布情况，其中头部中心、双肩与髋部点位清晰可见，具备较高的空间定位准确率。图 3 呈现了左右转身动作的原始角度检测曲线，可见在未经优化前，角度数据存在较为显著的毛刺，尤其在快速转动阶段，最大瞬态误差可达 $\pm 6^\circ \sim 8^\circ$ ，影响控制系统的响应平稳性。图 4 则为经过低通滤波处理后的角度变化曲线，结果显示，滤波后角度数据更加平滑，去除了高频扰动，同时保留了主要动作趋势特征，使得平台在驱动过程中不会产生急启、急停

等非预期行为，大幅度提高用户的操控感受。此外，测试还验证了在长时间运行状态下系统的稳定性，连续工作 2 小时后，内存占用无明显增长，CPU 温度控制在安全范围内，未发生因资源过载导致的性能下降问题。研究表明，该套基于 AI 视觉识别的身体转向检测方案不仅具备良好的工程可行性，还能在现有硬件条件下实现高效、稳定的 VR 交互体验，为消费级 VR 动感平台提供了一种低成本、高适配性的技术路径。

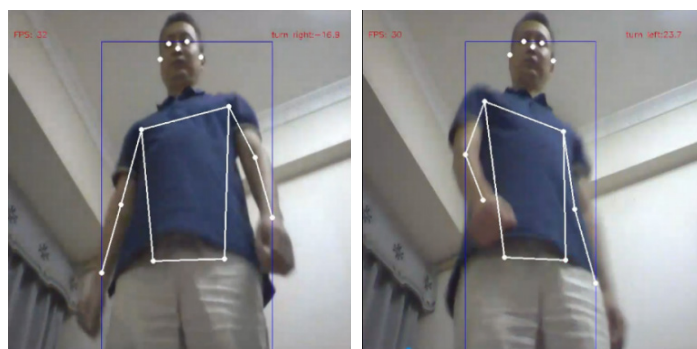


图 3 左右体转的检测结果

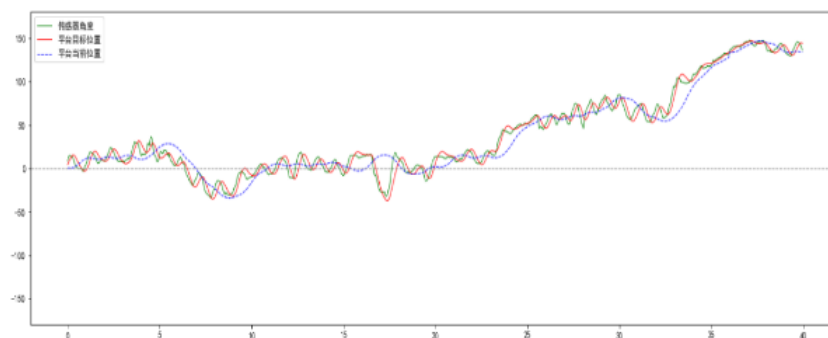


图 4 优化后控制结果

4 结论

本文针对 VR 动感平台设备自身小型化和放置空间有限场景中，无接触感知使用者身体转向的场景需求，设计实现适用该场景的基于机器学习图像识别身体转向检测方法，实时检测出身体的关键点，根据关键点位置计算出身体左右转体的角度，再进行控制优化后发送到 VR 空间中的虚拟角色实现角色的对应运动。应用结果表明，该方法能够满足现有 VR 体验应用场景的需求，提供较好的 VR 空间运动体验。后续将在本文工作基础上，继续深入研究探索相关机器学习图像识别的更深入更广泛的应用，例如提供更多动作的检测

识别，为使用者提供更高沉浸感、更强交互的 VR 体验。

参考文献

- [1] 柯健,张量.用于虚拟现实/增强现实人机交互的实时手势识别方法的研究[J].现代计算机, 2023, 29(1):59-62.
- [2] 田翀,杨孟妹,周扬诗宇等.基于MediaPipe人体姿态模型的步态分析方法:CN202211165984. X[P]. CN115644853A[2024-08-14].
- [3] 蔡历.面向虚拟现实的三维人体关键点识别与交互研究[D].浙江工业大学,2020.
- [4] 杨燕婷,张新松,马云莺.基于人体姿态估计的健身指导技术[J].信息与电脑:理论版, 2022(12):182-184.