

动作识别系统和方法

信息和通信

计算机/人工智能/数据处理和信息技术

传感器

机遇

从视频中识别人体动作对于监控、游戏和人机交互等应用至关重要。然而，一个主要挑战是视角变化——随着相机角度移动或受试者旋转，其外观会发生剧烈变化。现有方法依赖于固定、人工定义的预处理（例如，将骨架居中并对齐），或需要庞大的多视角数据集。这些预处理策略对于无人机或监控录像等真实场景（相机和受试者相对运动）不够灵活。它们也没有为动作识别显式地学习最优视角，限制了其鲁棒性和准确性。因此，需要一种更具适应性、可学习的方法。

技术

本专利提出一种具有视角自适应神经网络的动作识别系统，可动态学习并修正视角变化。系统从RGB-D相机或姿态估计中接收3D骨架数据（关节位置）。对于每一帧，视角适配模块应用无监督学习算法来确定最优变换参数（旋转角度 α, β, γ 和平移向量 b ）。然后，它使用3D旋转矩阵变换整个骨架，有效地将受试者重定向到规范的“最佳视角”，无需手动预处理。

转换后的骨架序列随后被送入图神经网络。图神经网络将骨架转换为图（关节=节点，骨骼=边）。利用自适应图卷积，它学习关节之间的物理连接性和潜在关系。多个残差块处理时空特征，分类器输出识别的动作。整个网络（视角适配+图神经网络）进行端到端训练，使视角参数能够专门针对动作识别精度进行优化。

优势

- 视角不变识别：无需手动预处理即可动态适应任何相机角度，显著提高视角变化下的识别精度。
- 无监督视角学习：视角适配模块仅使用分类损失学习最优变换，无需真实视角标签。
- 端到端训练：视角适配和图神经网络联合优化，比单独的预处理流程学习到更好的特征。
- 超越现有水平：在NTU60数据集上的准确率（CV 94.18%, CS 86.21%）高于使用固定预处理的A-GCN-P（CV 92.70%, CS 84.30%）。
- 参数高效：与堆叠更多图卷积层相比，视角适配模块增加的参数极少，却带来更好的性能提升。

应用

- 安防监控：从固定或移动摄像机识别可疑动作或手势，无需关注拍摄角度。
- 人机交互：为智能电视、游戏主机或VR/AR头显实现手势控制，不要求用户正对设备。

备注

CIMDA: P00037

IP状态

已申请专利



技术成熟度等级 (TRL) ?

4

发明人

严洪教授

Dr. Ali Raza SHAHID

范歆琦博士

查询: kto@cityu.edu.hk

Follow-on

Develop
ConceptProof
Concept

Build Value

- 自动驾驶：从不同角度的车载摄像头识别行人手势或驾驶员动作。
- 医疗康复：从床边摄像头监控患者锻炼或日常活动，不限制摄像头位置。
- 体育分析：从多个变焦、变角度的转播画面分析运动员动作。

