

5.1 相关理论及技术介绍

5.1.1 姿态模型介绍

(1) MediaPose

MediaPose 是一个用于人体姿态估计 (Human Pose Estimation, HPE) 的开源深度学习框架和模型, 主要用于从单张图片或视频帧中预测人体的关节点 (如头部、肩膀、手肘、膝盖等) 的位置和姿态。

(2) OpenPose

OpenPose 是一个开源的实时多人姿态估计系统, 最初由卡内基梅隆大学的 Perceptual Computing Lab 开发。它能够从图像或视频中检测和跟踪人体的关键点 (关节和关键区域), 并对人体的姿势进行精确估计。

5.1.2 相似度度量方法

根据不同特征向量可以采取不同的相似度度量方法, 来计算两个向量的距离, 判断两个向量之间的相似性。常见的相似度度量方法如下所示。

(1) 欧几里得距离

欧几里得距离是计算两个点之间的直线距离, 广泛用于向量空间中。对于两个 n 维向量 $X = (x_1, x_2, \dots, x_n)$ 和 $Y = (y_1, y_2, \dots, y_n)$, 欧几里得距离为:

$$d(X, Y) = \sqrt{\sum_{i=1}^n (X_i - Y_i)^2} \quad (1)$$

(2) 余弦距离

余弦相似度用于衡量两个向量在方向上的相似度。其值介于 -1 到 1 之间, 1 表示完全相同, 0 表示无关, -1 表示完全相反。对于两个 n 维向量 $a = (x_1, x_2, \dots, x_n)$ 和 $b = (y_1, y_2, \dots, y_n)$, 余弦相似度计算公式为:

$$\cos \theta = \frac{\sum_{i=1}^n x_i y_i}{\sqrt{\sum_{i=1}^n x_i^2} \sqrt{\sum_{i=1}^n y_i^2}} \quad (2)$$

(3) 曼哈顿距离

对于两个 n 维向量 $X = (x_1, x_2, \dots, x_n)$ 和 $Y = (y_1, y_2, \dots, y_n)$, 曼哈顿距离为:

$$d(X, Y) = \sum_{i=1}^n |x_i - y_i| \quad (3)$$

5.1.3 动作匹配方法

(1) 动态时间规整算法

动态时间规整 (Dynamic Time Warping, 简称 DTW) 算法是一种用于衡量和对齐两个时间序列相似度的算法。它通过寻找最佳匹配路径, 使得两个时间序列可以在不同的时间步长下对齐, 从而解决不同速度或不同时间尺度下的数据匹配问题。

5.2 开发技术栈

5.2.1 前端: UniApp 框架

UniApp 是一个基于 Vue.js 的跨平台应用开发框架, 主要用于构建 Web、iOS、Android、小程序 (如微信小程序、支付宝小程序、字节跳动小程序等) 和桌面应用等

多平台的应用。UniApp 提供了一套统一的开发方式，可以让开发者通过一次编码，发布到多个平台，极大地提高了开发效率，降低了多平台开发的成本。

5.2.2 后端：Django 框架

Django 是一个高级的 Python Web 框架，旨在帮助开发者快速而高效地构建 Web 应用。它遵循 MVC（模型—视图—控制器）设计模式，并且以 DRY（Don't Repeat Yourself）原则为核心，强调代码复用和简洁性。

5.2.3 数据库：MySQL

MySQL 是一个开源的关系型数据库管理系统（RDBMS），它使用结构化查询语言（SQL）来管理数据。它可以将数据存在不同的表中，利用关系关联各个数据，而不是将所有数据都存放在一个表中，这样能够使数据更清晰，也能提高数据操作的灵活性。

5.3 核心功能实现

本项目中轻度认知障碍老人运动跟练功能，内容为手机摄像头捕获轻度认知障碍老年人所做的运动动作，实时识别出轻度认知障碍老年人的动作，随后与模板样本库进行动作匹配，最后形成运动报告，评估轻度认知障碍老年人本次所做的动作，让轻度认知障碍老年人发现自身所做运动的不足之处。

5.3.1 收集标准康复训练动作模板视频以形成标准样本库

拍摄康复训练动作的标准模板视频，以及轻度认知障碍老人根据 Vivifrail 多组分运动康复训练方案所进行的训练视频，为后续轻度认知障碍老人动作识别和动作评估充当数据集。

5.3.2 姿态估计实现

本项目使用 MediaPose 模型进行人体姿态估计，检测人体的关键点。MediaPose 在移动手机上表现出了出色的性能，比 20 核桌面 CPU 上的 OpenPose 还要快 25-75 倍，全身 33 个关键点拓扑可以更准确地捕捉人体的全身姿势。

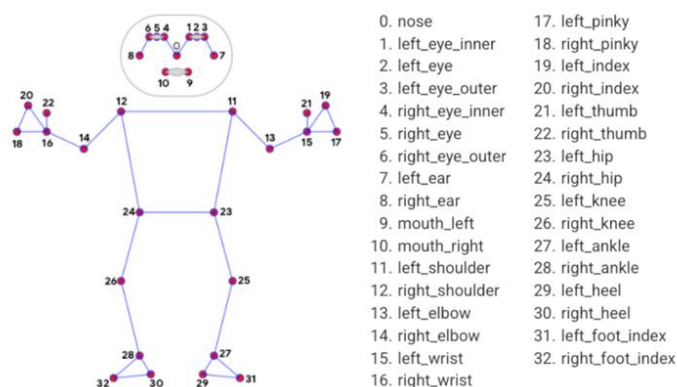


图 9 mediaPose 模型人体 33 个关键点示意图

5.3.3 动作识别实现

本部分内容的主要作用是识别轻度认知障碍老年人做的 Vivifrail 多组分运动动作，以让轻度认知障碍老人了解自身当前所做的动作以及确定后续动作评估时序的起始帧。根据多组分运动中描述的动作，可以把动作总结为 8 类：举水瓶、起身、跨越、手臂伸展、腿部伸展、坐下、走路、握力抗阻。

基于人体骨架数据的动作识别流程，由视频图像采集、骨架数据获取、特征提取与模型训练以及动作识别四个步骤组成。

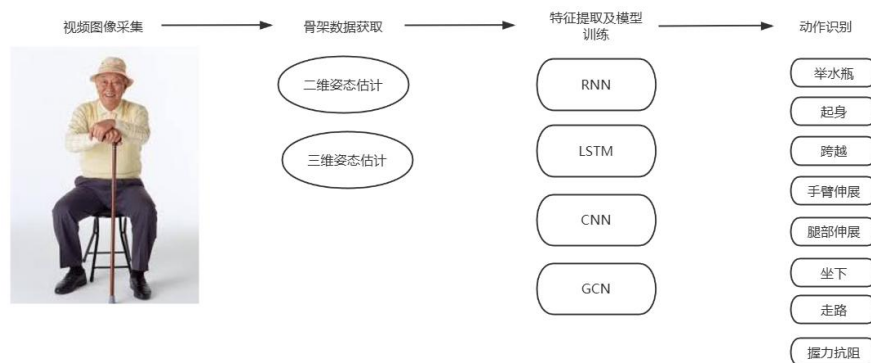


图 10 动作识别流程图

在基于人体骨架数据的动作识别过程中，首先，需要获取人体的骨架信息，即人体各关节点坐标。通常使用摄像头或视频录制设备捕获视频，尽可能地保证较高的视频质量，有利于后续步骤中人体骨架数据的准确提取；然后，使用基于深度学习的人体姿态估计算法从视频图像中提取人体的骨架数据，为后续的动作识别提供基础；接着，从获取到的骨架数据中提取丰富的运动特征；最后，根据提取到的特征识别出具体的动作或行为。

5.3.4 动作评估实现

利用训练模板视频中的标准动作作为模板，轻度认知障碍老年人所做的动作会实时跟标准动作视频进行匹配，并实时指出轻度认知障碍老年人动作中存在的问题，当完成指定目标数量和时间的运动后，会形成此动作的运动报告。运动报告为轻度认知障碍老年人提供此动作的相关指导和建议。

首先通过姿态估计算法提取轻度认知障碍老年人训练视频中的关键骨骼点信息，计算人体主要关节的活动角度，用于描述动作的核心状态。然后使用动态时间规整算法（DTW），将轻度认知障碍老年人动作序列与标准模板动作进行匹配，根据人体主要关节角度与模板视频主要关节角度偏差，对轻度认知障碍老年人进行动作指导，最后根据相似度评级量表及小程序提供的动作指导来生成运动报告。

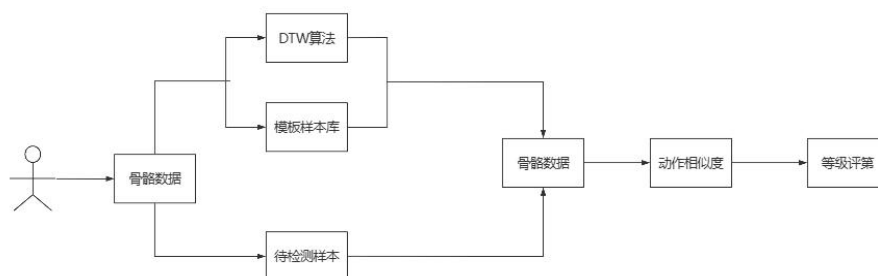


图 11 动作识别流程图

具体实现如下：

（1）构建人体肢体向量

图像中关键骨骼点坐标会受到人体形态（如高矮胖瘦），人体与摄像之间的距离（成像大小）等因素的影响，因此使用关键骨骼点坐标进行动作匹配前，需要进行标准化处理，增加了计算难度的同时，也降低了数据的真实性。因此，可以将检测到的人体关键点表示为人的肢体向量，以身体部分向量为例，见图 12。

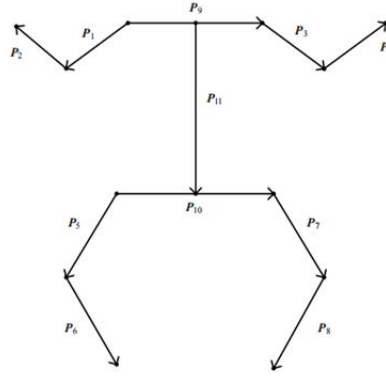


图 12 肢体向量

肢体向量定义，见表 5。

表 5 肢体向量定义

肢体	向量	向量定义
右上臂	P ₁	右肩->右肘
右小臂	P ₂	右肘->右手腕
左上臂	P ₃	左肩->左肘
左小臂	P ₄	左肘->左手腕
右大腿	P ₅	右髋->右膝
右小腿	P ₆	右膝->右脚踝
左大腿	P ₇	左髋->左膝
左小腿	P ₈	左膝->左脚踝
双肩	P ₉	右肩->左肩
腰部	P ₁₀	右髋->左髋
躯干	P ₁₁	双肩中点->腰部中点

(2) 关节活动角度特征向量

通过计算公式可计算右肩关节活动角度：

$$a_1 = \frac{180}{\pi} \arccos\left(\frac{P_1 P_{11}}{|P_1 P_{11}|}\right) \quad (4)$$

其中， a_1 表示右肩关节角度， P_1 表示右上臂， P_{11} 表示躯干。

通过计算公式可计算左肩关节活动角度：

$$a_2 = \frac{180}{\pi} \arccos\left(\frac{P_3 P_{11}}{|P_3 P_{11}|}\right) \quad (5)$$

其中， a_2 表示左肩关节角度， P_3 表示左上臂， P_{11} 表示躯干。

通过计算公式可计算右肘关节活动角度：

$$a_3 = \frac{180}{\pi} \arccos\left(\frac{P_1 P_2}{|P_1 P_2|}\right) \quad (6)$$

其中， a_3 表示右肘关节角度， P_1 表示右上臂， P_2 表示右小臂。

通过计算公式可计算左肘关节活动角度：

$$a_4 = \frac{180}{\pi} \arccos\left(\frac{P_3 P_4}{|P_3 P_4|}\right) \quad (7)$$

其中， a_4 表示左肘关节角度， P_3 表示左上臂， P_4 表示左小臂。
通过计算公式可计算右髋关节活动角度：

$$a_5 = \frac{180}{\pi} \arccos\left(\frac{P_5 P_{11}}{|P_5 P_{11}|}\right) \quad (8)$$

其中， a_5 表示右髋关节角度， P_5 表示右大腿， P_{11} 表示躯干。
通过计算公式可计算左髋关节活动角度：

$$a_6 = \frac{180}{\pi} \arccos\left(\frac{P_7 P_{11}}{|P_7 P_{11}|}\right) \quad (9)$$

其中， a_6 表示左髋关节角度， P_7 表示左大腿， P_{11} 表示躯干。
通过计算公式可计算右膝关节活动角度：

$$a_7 = \frac{180}{\pi} \arccos\left(\frac{P_5 P_6}{|P_5 P_6|}\right) \quad (10)$$

其中， a_7 表示右膝关节角度， P_5 表示右大腿， P_6 表示右小腿。
通过计算公式可计算左膝关节活动角度：

$$a_8 = \frac{180}{\pi} \arccos\left(\frac{P_7 P_8}{|P_7 P_8|}\right) \quad (11)$$

其中， a_8 表示左膝关节角度， P_7 表示左大腿， P_8 表示左小腿。

以上 8 个关节活动角度为 MCI 老人做肢体训练时的常用关节，可以得到 MCI 老人训练动作视频在第 i 帧的关节活动角度特征为：

$$A_i = (a_{1i}, a_{2i}, a_{3i}, a_{4i}, a_{5i}, a_{6i}, a_{7i}, a_{8i})$$

一组 n 帧长度的动作序列关节活动角度特征为：

$$Q = (A_1, A_2, \dots, A_i, \dots, A_n)$$

(3) DTW 相似度计算

假设有两个不同帧数的动作序列：

$$Q_1 = (A_1, A_2, \dots, A_i, \dots, A_n)$$

$$Q_2 = (B_1, B_2, \dots, B_i, \dots, B_m)$$

Q_1 表示第一个动作序列，其中每个 A_i 是第 i 帧的关节活动角度特征， Q_2 表示第二个动作序列，其中每个 B_i 是第 i 帧的关节活动角度特征。对于序列 Q 和 P 中的每一帧 i 和 j ，首先计算它们之间的局部距离 $d(A_i, B_j)$ 。使用欧式距离来计算它们的差异：

$$d(A_i, B_j) = \sqrt{\sum_{k=1}^m (a_{ik} - b_{jk})^2} \quad (12)$$

其中 d 是每个关节特征的维度， a_{ik} 和 b_{jk} 是分别对应于第 i 帧的关节特征 A_i 和第

j 帧的关节特征 B_j 中第 k 个关节角度的数值。

然后构建一个累积距离矩阵 $D(i,j)$,即为截止到 Q_1 的第 i 帧和 Q_2 的第 j 帧两视频的最小累积差异:

$$D(i,j) = \begin{cases} 0 & i=0, j=0 \\ d(A_i, B_j) + \min(D(i-1, j), D(i, j-1), D(i-1, j-1)) & i>1, j>1 \end{cases} \quad (13)$$

其中 $D(i-1, j)$ 表示从序列 Q_1 的第 i-1 帧对齐到序列 Q_2 的第 j 帧。 $D(i, j-1)$ 表示从序列 Q_1 的第 i 帧对齐到序列 Q_2 的第 j-1 帧。 $D(i-1, j-1)$ 表示从序列 Q_1 的第 i-1 帧对齐到序列 Q_2 的第 j-1 帧。

最后为了使计算的相似度更具有可比性,将 DTW 距离进行归一化处理,即可得到两动作序列之间的相似度,相似度公式如下:

$$s = 1 - \frac{D(n, m)}{D_{\max}} \quad (14)$$

其中 s 表示两动作的相似度, $D_{\max}$ 是可能的最大 DTW 距离。

(4) 等级评第

为将轻度认知障碍老年人所做动作和标准动作的相似度直观地呈现,本小程序设置了一个相似度评级量表,详见表 6。

表 6 相似度评级量表

相似度	等级
0.9-1	完美
0.75-0.9	良好
0.6-0.75	及格
0-0.6	不及格