|  |
| --- |
| **项目创新性分析报告**  项目名称 RCDancer  小组名称 RC.Dancer  **2018.10.20** |

**一、项目实施的背景以及意义**

随着计算机视觉和人工智能逐渐发达，基于视频图像对人体动作的分析变得十分便捷。其次，基于视频图像的人体识别无需其他外部硬件设备，在精度要求不非常高的情况下，这类人体动作识别便可在生活，工作，学习的方方面面使用。在识别到动作后，利用训练好的人工智能即可对动作进行分析，提取关键点、找出动作差异、判断动作完成度、提供纠错信息等。

社会上许多人群有着动作判断和指导的需求，比如跳舞动作，某些领域上的专业动作（交警的指挥动作，运动员的技术动作等）。尤其现代社会越来越多的青少年热衷于歌舞，也激发了更多人跳舞的兴趣。但通常情况下，他们不会因为某一个舞就请老师进行系统的舞蹈教学，更多的是通过视频进行参照学习。但当他们自学舞蹈时，却苦于没有动作上的提示和指导；即便有舞蹈基础的人，在自主练习也没有有效的方法纠正不标准的动作。其次，这些人群可能只有摄像头和已有的视频图像，而在这样轻度使用的情况下，他们并不会去专门购买动作捕捉设备来达到他们的需求。因此这类人群对基于视频图像的动作分析的需求是存在的。本项目从最贴近生活的跳舞入手，着力解决在舞蹈练习上存在的动作纠正的需求，并以此为出发点，将技术逐渐渗透到其他有这类需求的领域。

本项目通过基于图像的人体动作识别技术，结合人工智能分析，为自主练习舞蹈的人群提供舞蹈动作的提示，纠正，引导用户练习标准的舞蹈动作。项目采用Web应用的形式，使用B/S架构，用户只需要在有网络、摄像头的情况下即可使用。

**二、技术发展趋势及国内外发展现状**

**2.1涉及技术领域**  
 本项目涉及的主要领域有：深度学习、浏览器Web应用、视频图像处理、舞蹈教学、动作矫正；

**2.2技术介绍**

本项目主要涉及到的主要技术包含如下几个方面：

（1）姿态估计（pose estimation）：包括关键关节动作识别以及某些具体部位（如面部、手指）等的精细动作识别；OpenPose是一个利用OpenCV和Caffe并以C++写成的开源动作识别库，通过使用OpenPose能够得到运动中人体的躯干关键点，并以此构建骨架模型；具体到本项目中，我们期望从分析过的标准舞蹈视频中提取出关键帧（根据音乐节奏点以及物理图像辨析），建立标准参照骨架模型；在用户学习的过程中，我们能够实时获取用户姿态，并输入模型进行比对；

（2）多分类方法：目前机器学习领域解决分类问题的方法有许多，单一分类方法常见的有：决策树、朴素贝叶斯、人工神经网络、K-近邻等；多分类问题的实现可以通过构造合适的支持向量机（SVM）多类分类器来实现，而构造此类多类分类器的主要方法有两类：直接法与间接法，直接法计算复杂度高，适用范围窄，不适合应用于大型问题中；间接法的主要原理是通过组合多个二分类器来实现多分类器的构造，常见方法又有一对多（one against all）和多对一法（all against one）以及层次支持向量（LSVM）等等；针对本项目而言，我们需要对用户输入的动作信息进行处理，与“标准”动作模型进行比对，得到未经分类的原始偏差数据，随后根据上述技术对数据进行多分类处理，最终输出分类完毕的偏差信息；

（4）视频数据处理：关键帧提取是视频数据处理中一个主要部分，目的在于降低计算机的运算量，并提取出有效信息。关键帧代指一段镜头中最重要、最富有代表性的一副或多幅图像；对于动作特征明显的视频，主要提取方法有：基于运动的分析方法，通过比对摄像机镜头焦距变化，角度移动前后的帧差异程度来提取关键帧；基于聚类的分析方法，此方法原理在于将视频聚成n个类并进行比对，选出差异较大的类中的代表帧作为关键帧，如果一个类的帧数过少，则直接与相邻帧进行合并；我们项目需要对用户提供的标准舞蹈视频进行关键帧提取，既大致判断出舞蹈中的关键动作，并通过第一点的姿态预测建立标准模型；其余方法还包括：灰度变化、图像阈值化、图像分割（得到运动图像）、背景减除及降噪等；

（5）图片对比：（建立对比评估模型）本项目图片对比的主要目的是确定用户动作与标准动作模型的偏差值，主要方法是通过OpenCV配合pHash算法或SIFT算法等建立对比评估模型；

**2.3可能涉及到的其余技术**

（1）人体立体模型深度处理：项目可能需要对2D模型进行三维重建，最新技术有基于CVPR 2017 Lifting from the Deep 等论文提出的实现模型Lifting-from-the-Deep-release；

**三、项目创新点**

**3.1题材创新性分析**

（1）针对有动作分析需求的人群提出解决方案

动作识别的信息可以作为输入内容，对这些信息进行比较，得出用户动作和标准动作(标准来源多样，主要为用户认可的标准)的差距信息，这些信息特化于不同的场景，就可以满足多类人群对动作分析的需求。

（2）利用现有技术开发新技术

要满足上述需求，首先能够满足在视频图像上进行人体动作识别，而此技术已经较为成熟，有多种方案可供选择，其次是得到动作信息后的比对分析，这时候使用人工智能，进行深度学习训练，得出结论。这一过程的整合可视作一个新技术，用于解决特定的动作分析问题。

（3）对具体使用场景和人群进行需求分析，最大程度解决跳舞入门难，练习无指导的问题

项目确定了从参照视频自学舞蹈这一具体的使用场景出发，对这类人群进行了具体的需求分析，在满足基本的需求基础上，能够最大程度的为用户解决普遍存在的问题，并且在设计用户界面和交互方式时有更实际的需求可以遵循。

（4）从舞蹈动作入手，逐渐扩大使用场景和人群

项目选择从跳舞这一大众兴趣入手，着力于解决舞蹈动作分析纠正，并且当该技术成熟运用后，逐步扩大使用场景，即不仅仅在跳舞动作上，可以拓展到其他技术动作的分析纠正上，并可以将这一套技术应用于其他各类领域，各类人群中使用。

**3.2功能创新性分析**

（1）对原视频进行关键动作的识别、提取和分析

系统能够分析用户选择作为参照学习的视频，根据乐曲节奏，动作变化，提取关键动作，结合利用大量人体动作训练出的基本模型，建立一个针对于该视频的动作模型，以便后续指导和纠正用户动作。

（2）客户端进行实时人体动作识别

采用摄像头捕获用户动作，就Web应用来说，若动作的识别过程在服务器进行，由于网络延迟和其他影响因素的情况下，并不能达到实时捕捉的要求，因此，本系统采用在浏览器上，即客户端进行实时的动作识别，完全消除了网络延迟的时间误差，识别速度取决于用户电脑本身的性能，并能采用GPU加速，达到实时的目的。

（3）人工智能判断动作，给予实时的动作提示和纠正信息

识别到动作后，系统能够使用人工智能对该动作和参照视频进行比对，判断动作是否做对，是否卡到节拍，是否规范等，并能直接反馈给用户精确到错误信息，如具体到手臂伸错方向，动作慢了，腿未伸直。

（4）更人性化的Web界面和交互方式

根据具体使用场景的分析结果，项目提出了更加符合使用场景的界面和更加人性化的交互方式，在界面上，摒弃传统视频播放界面，重新设计更符合舞蹈练习的播放器，此外，在交互方面，基于动作识别可以设计利用体感控制应用，利用肢体动作和界面进行简单交互，大大增强了用户体验。

**四、项目预期目标**

项目从舞蹈动作纠正入手，逐步完善动作纠正技术，并能渗透在各个领域，起到帮助人们规范动作，辅助训练的作用。其次是打造一个较为成熟的舞蹈动作纠正应用，并建立足够的用户群，与其他相关网络内容的深度合作，扩大用户群体，亦可实现商业价值。

项目的初期目标是能够帮助喜爱舞蹈的用户能够随时随地根据自身情况进行自学；我们希望通过这个项目让更多的人能喜欢上舞蹈，为生活增添健康和乐趣。