

固定场景下的人体姿态识别

赵一祫

(辽宁科技大学, 辽宁 鞍山 114000)

摘要:近年来,随着信息技术的发展和智能科技的普及,全球科技变革正在进一步推进,云计算、物联网、大数据和人工智能等技术也在飞速发展,其中,人体姿态识别技术已开始在计算机视觉相关领域中广泛应用。就固定场景下的人体姿态识别做出研究分析。

关键词:人体姿态识别;云计算;人工智能

DOI:10.16184/j.cnki.comprg.2018.11.057

当前人工智能技术的热点,固定情景下对于人体姿态识别具有十分重要的研究意义,对于我国实现现代化建设具有一定的推动作用,因此加强技术分析研判就显得十分重要。早在上世纪70年代,我国已经开始了对人体行为分析方面的研究,这些研究对于我国人工智能的发展有了较强的推动作用,在特定情景下或者说在比较标准的场景中分析较为简单的姿态和动作已经成为了可能,但这些工作的开展大多停留在理论的层次,并没有付诸实践,要想将这些分析技术真正应用到实际场景中仍然需要大量的实验进行探索。

1 人体姿态识别

人体姿态识别主要在于研究描述人体姿态以及预测人体行为,其识别过程是指,在指定图像或视屏中,根据人体中关节点位置的变化,识别人体动作的过程。人体姿态识别的算法主要分为两类,一是基于深度图的算法,另一类直接基于RGB图像的算法。深度图是指由相机拍摄的图片,其每个像素值代表的是物体到相机XY平面的距离。这种算法的应用容易因采集设备的要求而受限,但基于RGB图像的算法直接通过对红、绿、蓝3个颜色通道的变化以及它们相互之间的叠加来得到的颜色进行识别,不会受到其他因素的干扰限制,因此更具有发展前景,同时也取得了不少成果。目前,即使是在较为复杂的、某种固定的场景中,基于RGB图像的人体姿态估计算法相较于基于深度图的人体姿态估计算法也能达到很好的识别效果。无论是深度图技术还是RGB图像技术,都是通过计算机强大的运算能力进行人体姿态的动作预算,通过这样的方式能够一定程度地实现人工图像的监测,并且能够为人工智能的普及奠定良好的基础。随着我国社会水平的不断提升,人们对于社会生活的质量要求也在不断增加,因此在实际的生活

过程中视频监控已经成为人们不可或缺的一种安全措施,基于视频分析的技术要求也越来越高。例如在智能家居、医疗领域及运动分析等行业中都得到了较为广泛的应用,固定场景下的人体姿态识别在各领域起到的作用显而易见。特别是近年来,我国安保工作的加强,对于大城市人口密集流动以及犯罪分子的甄别等都有较强需求。

2 人体姿态识别的实现

人体姿态是被主要分为基于计算机视角的识别和基于运动捕获技术的识别。基于计算机视觉的识别通过各种特征信息来对人体姿态动作进行识别,比如视频图像序列、人体轮廓、多视角等。基于计算机视觉的识别可以比较容易获取人体运动的轨迹、轮廓等信息,但没有办法具体实现表达人体的运动细节,以及容易存在因遮挡而识别错误等问题。基于运动捕获技术的人体姿态识别,则是通过定位人体的关节点、储存关节点运动数据信息来识别人体运动轨道。相较于计算机视角的人体姿态识别,基于运动捕获技术的人体姿态识别可以更好地反应人体姿态信息,也可以更好地处理和记录运动细节,不会因为物体颜色或被遮挡而影响运动轨道的识别。技术的革新对人体姿态的分析捕捉有较强的辅助作用,并且能够更好地展现动作的细节,对于专业人士进行动作分析的痕迹管理有较高的参考价值。通过良好的运动前景预算能够,在各种计算方法中做出合理的预测,并且在各种环境中的适应能力也能够得到一定程度的加强。由于未来的监控实现的方向是在全领域的视频监控,因此对于用户的特定化要求也应该及时进行技术革新,用户对于技术的需求就是技术革新的发展方向,

收稿日期:2018-08-15

主要在通过算法完善的过程中,还应该加强对于运动轨迹的采取方法的提升。

固定场景下的人体姿态识别主要通过前景提取、特征向量提取、姿态分类识别3个步骤来逐步实现。首先通过前景图像提取特征,通过计算前景图像中物体的位置、方向、周长、宽高比、面积与所围面积比、离心率等几何特征,再通过提取的几何特征参数形成的特征向量反映人体姿态。最后再根据具体算法要求,采集不同的人体姿态图像作为样本,提取其中的特征作为数据集,以此作为数据源参考,从而达到人体姿态识别的目的。整个数据从数据采集到数据整理再到数据计算,依靠的是计算机强大的运算能力和数学建模的合理性,因此想要加强人体姿态识别系统的准确性,一定程度上要依靠于计算机的运算能力和数学建模的科学性。想要实现这些应该加强专业人员的系统培训,并且做到有理有据的数学建模。对于运动前景的数学建模也要包括对于场景的背景建模,将整个建模过程合为一体,因此作为数学建模工程中的模块背景建模对于人体姿态的预测准确性有较强的影响。对于环境建模的科学性是保证人体姿态预测准确性的良好保证,由于环境是随着时间变化而不断变化的,其背景会受光线、风速,以及其他各种自然界因素的影响,因此,在其背景中,进行人体姿态的模拟,就需要将这些自然因素,尽可能地考虑到数学建模中去,并且及时更新模型,从而将这些自然现象进行更为客观的反映,进而保证预测的准确性。

姿态特征参数的获取是保证固定场景下对于人体姿态识别的又一大重任,同时也是影响预测结果准确性的重要环节。特征参数的选择具有一定的经验,选择正确的特征参数描述内容,能够最大限度地实现预测的准确性,并且对于分类过程有着较强的指导意义,在设计识别系统时可以将系统尽可能地细化,实现预处理、模型选取、参数选取以及分类判别等几大模块。对系统进行较为细化的预处理,能够实现特征参数选取的准确性,并且对于预测具有承前启后的作用,是保证整个系统准确性的关键一步。因此,姿态特征参数的选取对于整个系统的模型构成具有十分重要的意义。

3 人体姿态识别的应用

人体姿态识别的应用范围十分广泛,可用于人机交互、影视制作、运动分析、游戏娱乐等各种领域。人们可以利用人体姿态识别定位人体关节运动轨迹并记录

其运动数据,实现3D动画模拟人体运动来制作电影电视;也可以通过记录的轨道和数据对运动进行分析;还可以实现人机交互、游戏娱乐,比如体感游戏就是通过对人体运动姿态进行识别来实现游戏互动的。

人体姿态识别目前最为广泛的应用是在于智能监控中。智能监控与一般普通监控的区别主要在于将人体姿态识别技术嵌入视频服务器中,运用算法,识别、判断监控画面场景中的动态物体——行人、车辆的行为,提取其中关键信息,当出现异常行为时,及时向用户发出警报。同样,固定场景下的人体姿态识别技术可以应用于家庭监控,如为了预防独居老人摔倒情况的发生,可以通过在家中安装识别摔倒姿态的智能监控设备,对独居老年人摔倒情况的识别,当出现紧急情况时及时做出响应。人类社会的不断发展与生活质量的不断提升,视频监控已经被非常广泛地应用到各个领域,人们生活空间的不断扩大和延展,公共及私人场所的领域也在随之发展,遇到各种突发情况的概率在不断增加,尤其是在公共场所,由于其监控难度较大,人口密集。通过简单的监控,已经无法满足当今社会发展的要求,简单地依靠值班人员的坚守,想要真正做到人体姿态预测还有较大难度,对于社会资源也是一种潜在的浪费。因此,选用智能监控系统已经成为当前社会解决这一根本问题的必由之路,在社交的过程中,人类除了语言以外其肢体动作也能传递一定的信息,通过较为科学合理的计算机预测能够实现动作的含义解读,并且更好帮助人们的实现社交。

4 人体姿态识别面临的挑战与发展方向

当前固定场景下的人体姿态识别的实际运用中,仍面临着一些挑战。比如多种相似动作难以识别,人的肢体众多灵活且复杂,当出现交换频繁或是相似的肢体动作是难以进行人体姿态识别。除此之外,衣着的变化、视角的变化等视觉原因也为人体姿态识别造成了严重影响和很大的困难。对于提升固定场景下人体姿态识别技术的准确性预测,技术革新之路还很漫长。当然对于数学建模而言,也是其发展的必然选择,由于当今社会选择整容的人越来越多,这也一定程度上给人体姿态识别带来了更多的挑战。

随着我国科学技术的不断完善和计算机图形算法能运算能力的不断提升,人工智能领域的研究将会不断深入。高新型人工智能人才的培养已经成为了社会教育发



展必然趋势,因此将现有的简单姿势进行较为科学合理的扩展,进行更深层次的研究已经成为了未来固定场景下人体姿态识别发展的重要方向。在使用信息技术的前提下能够更高效地做到视觉处理,进而来描述更为复杂的动作,这将会是未来该行业发展的重要方向。

5 结语

目前固定场景下的人体姿态识别仍然处于在发展状态中,还存在着许多大大小小的缺陷,面临着来自不同方面的挑战和压力,但其所能取得的成果指日可待。基于姿态识别的智能监测系统的研究成果已经被人们所认可,这样不单单可以帮助客户更加快捷地寻找到监控的重点,并且能够做到资源的整合,有针对性地提取有效

(上接第 143 页)

量判断,被认为是评估视觉质量的最准确的方法。常用的主观评估算法主要包括单刺激方法 (Single-Stimulus)、双刺激方法 (Double-Stimulus)、强迫选择成对比较 (Forced-Choice Pairwise Comparison) 和相似性判断 (Similarity Judgement)。单刺激和双刺激方法代表分类等级,观察者以固定的 5 点尺度判断单个图像和一对图像的质量。强制选择成对比较是一种排序方法,其中观察者决定两个显示图像中的哪一个具有更高的质量。在相似性判断方法中,观察者不仅选择质量较高的图像,而且还在连续的尺度上估计质量的差异。

根据原始图像的可用性,并与失真图像进行比较,可以对客观图像质量评估度量进行分类。大多数现有的方法被称为全参考 (Full Reference, FR) 评估方法,意味着一个完整的参考图像被假定为已知的。然而,在许多实际应用中,参考图像不可用,因此希望采用无参考 (No Reference, NR) 或“盲”质量评估方法。全参考和无参考评估方法也可以分别称为有监督和无监督评估。有的参考图像仅部分可用,其形式是一组提取的特征,这些特征可用作辅助信息,以帮助评估失真图像的质量。这被称为半参考 (Reduced Reference, RR) 质量评估^[4,5]。其中全参考评估方法是目前可靠性最高的客观评估方法^[6,7]。

图像分割中被广泛使用的客观图像质量评估标准是基于像素的度量: MSE (Mean Squared Error) 和 PSNR (Peak Signal to Noise Ratio)。令 $X = \{x_i | i=1, 2, \dots, N\}$ 和 $Y = \{y_i | i=1, 2, \dots, N\}$ 分别表示大小为 N 的参考图像和测

试图像,同时还能节省人力物力,给人们的生产生活带来了极大的便利。

参考文献

- [1] 郑潇, 彭晓东, 王嘉璇. 基于姿态时空特征的人体行为识别方法 [J]. 计算机辅助设计与图形学学报, 2018, (09): 1615-1624.
- [2] 张承玺. 固定场景下的人体姿态识别 [D]. 哈尔滨工业大学, 2014.
- [3] 张苗辉. 基于视觉系统的行人检测与跟踪方法研究 [D]. 上海交通大学, 2013.
- [4] 潘华伟. 基于多视点的三维姿态运动重建与跟踪 [D]. 湖南大学, 2013.

试图像,则 MSE 和 PSNR 的计算公式如下表示:

$$MSE = \frac{1}{N} \sum_{i=1}^N (x_i - y_i)^2$$

$$PSNR = 10 \log_{10} \left(\frac{N \cdot \max(X)^2}{\sum_{i=1}^N (x_i - y_i)^2} \right)$$

其中 $\max(X) = 2^l - 1$ 表示参考图像像素强度的最大值, l 表示图像像素编码所需的位数。MSE 的值越小表示测试图像与参考图像的误差越小。相反, PSNR 的值越小则表示测试图像与参考图像的误差越大。

5 结语

阐述了研究基于视觉特性的计算机图像分割算法的必要性,具体介绍了图像分割的理论定义。介绍了视觉特性在图像分割算法中的应用,包括视觉注意机制、视觉显著性检测,以及视觉显著性检测与常用图像分割方法相结合的算法应用。在图像分割结束后,还需要对其分割结果进行质量评估。阐述了利用人类视觉的图像质量主观评估方法和基于定量分析的图像质量客观评估方法。对基于视觉特性的计算机图像分割算法应用研究为后续的进一步图像分割算法提供必要的研究基础。

参考文献

- [1] 欧中亚. 试论基于视觉特性的图形图像分割算法. 信息技术与信息化, 2018, 7: 101-104.
- [2] 常盼盼, 黄继风. 基于视觉感知特性的彩色图像分割. 计算机与数字工程, 2012, 2: 92-95.