



(12)发明专利申请

(10)申请公布号 CN 110781765 A

(43)申请公布日 2020.02.11

(21)申请号 201910942510.3

(22)申请日 2019.09.30

(71)申请人 腾讯科技(深圳)有限公司

地址 518000 广东省深圳市南山区高新区
科技中一路腾讯大厦35层

(72)发明人 姚永强 葛彦昊 张伟 曹煊
倪辉 汪铖杰 李季霖 黄飞跃

(74)专利代理机构 北京派特恩知识产权代理有
限公司 11270

代理人 刘晖铭 张颖玲

(51)Int.Cl.

G06K 9/00(2006.01)

G06K 9/62(2006.01)

G06N 20/00(2019.01)

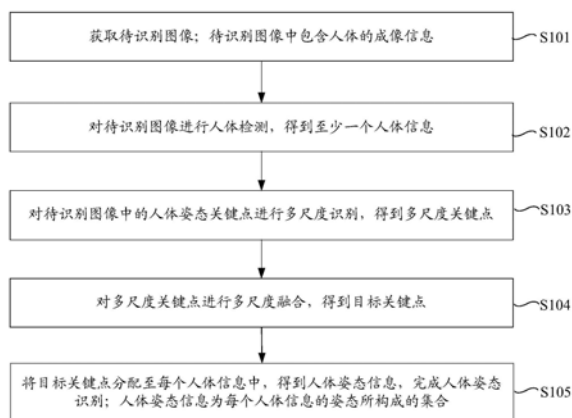
权利要求书3页 说明书18页 附图13页

(54)发明名称

一种人体姿态识别方法、装置、设备及存储
介质

(57)摘要

本发明实施例提供了一种人体姿态识别方法、装置、设备及存储介质；该方法包括：获取待识别图像；待识别图像中包含人体的成像信息；对待识别图像进行人体检测，得到至少一个人体信息；对待识别图像中的人体姿态关键点进行多尺度识别，得到多尺度关键点；对多尺度关键点进行多尺度融合，得到目标关键点；将目标关键点分配至每个人体信息中，得到人体姿态信息，完成人体姿态识别；人体姿态信息为每个人体信息的姿态所构成的集合。通过本发明实施例，能够提升人体姿态识别的准确度。



1. 一种人体姿态识别方法,其特征在于,包括:
获取待识别图像;所述待识别图像中包含人体的成像信息;
对所述待识别图像进行人体检测,得到至少一个人体信息;
对所述待识别图像中的人体姿态关键点进行多尺度识别,得到多尺度关键点;
对所述多尺度关键点进行多尺度融合,得到目标关键点;
将所述目标关键点分配至每个人体信息中,得到人体姿态信息,完成人体姿态识别;所述人体姿态信息为所述每个人体信息的姿态所构成的集合。
2. 根据权利要求1所述的方法,其特征在于,所述对所述待识别图像中的所述人体姿态关键点进行多尺度识别,得到多尺度关键点,包括:
对所述待识别图像进行多尺度特征提取,得到目标多尺度特征;
对所述目标多尺度特征进行所述人体姿态关键点识别,得到所述多尺度关键点。
3. 根据权利要求1或2所述的方法,其特征在于,所述对所述待识别图像中的人体姿态关键点进行多尺度识别,得到多尺度关键点,包括:
利用预设关键点识别模型,对所述待识别图像中的所述人体姿态关键点进行多尺度识别,得到所述多尺度关键点;
其中,所述预设关键点识别模型用于从至少两种尺度对所述待识别图像中的所述人体姿态关键点进行识别。
4. 根据权利要求1或2所述的方法,其特征在于,所述对所述多尺度关键点进行多尺度融合,得到目标关键点,包括:
获取所述多尺度关键点的局部最大值,得到预测多尺度关键点;
基于所述多尺度关键点,对所述预测多尺度关键点进行线性组合处理,得到初始多尺度关键点;
对所述初始多尺度关键点进行多尺度融合处理,得到所述目标关键点。
5. 根据权利要求4所述的方法,其特征在于,所述基于所述多尺度关键点,对所述预测多尺度关键点进行线性组合处理,得到初始多尺度关键点,包括:
从所述多尺度关键点上,获取预设数量的与所述预测多尺度关键点的距离最小的相邻关键点;
对所述预测多尺度关键点和所述相邻关键点进行线性组合,得到所述初始尺度关键点。
6. 根据权利要求4所述的方法,其特征在于,所述初始多尺度关键点包括至少一个多尺度目标关键点,每个多尺度目标关键点由至少两种尺度下每种尺度对应的关键点构成的集合组成;所述对所述初始多尺度关键点进行多尺度融合处理,得到所述目标关键点,包括:
获取所述初始多尺度关键点中所述每个多尺度目标关键点的目标中心点;
计算所述每个多尺度目标关键点中每种尺度对应的关键点,到所述目标中心点的距离,得到目标距离;
将所述目标距离中小于预设距离阈值的关键点的均值,作为与所述每个多尺度目标关键点对应的子目标关键点,从而得到与所述初始多尺度关键点对应的所述目标关键点。
7. 根据权利要求1或2所述的方法,其特征在于,所述将所述目标关键点分配至每个人体信息中,得到人体姿态信息,包括:

从所述目标关键点中,确定属于所述每个人体信息的个体关键点;

从所述个体关键点中,获取所述人体姿态关键点中每个关键点类型所对应的个体类型关键点;

从所述个体类型关键点中,获取目标个体类型关键点,得到与所述每个人体信息对应的个体姿态关键点;

根据所述个体姿态关键点,得到所述每个人体信息对应的个体姿态,从而得到所述至少一个人体信息对应的所述人体姿态信息。

8. 根据权利要求3所述的方法,其特征在于,所述对所述待识别图像进行人体检测,得到至少一个人体信息之前,所述方法还包括:

获取目标样本图像;所述目标样本图像用于训练得到所述预设关键点识别模型;

搭建原始关键点识别模型;所述原始关键点识别模型为多尺度层次结构的模型;

利用所述目标样本图像,对所述原始关键点识别模型进行训练,得到所述预设关键点识别模型。

9. 根据权利要求8所述的方法,其特征在于,所述获取目标样本图像,包括:

获取人体部位样本图像和待识别样本图像;

对所述人体部位样本图像的人体部位进行识别,得到人体部位识别结果;

基于所述人体部位识别结果,从所述人体部位样本图像中提取人体部位,得到人体部位数据库;

将所述人体部位数据库中的人体部位与所述待识别样本图像中对应的人体部位进行融合,得到所述目标样本图像。

10. 根据权利要求7所述的方法,其特征在于,所述获取目标样本图像之后,所述方法还包括:

搭建原始人体检测模型;所述原始人体检测模型为用于识别出所述待识别图像中的每个人体信息的初始模型;

利用所述目标样本图像,训练所述原始人体检测模型,得到人体检测模型;所述人体检测模型用于识别出所述待识别图像中的每个人体信息;

相应地,所述对所述待识别图像进行人体检测,得到至少一个人体信息,包括:

利用所述人体检测模型,对所述待识别图像进行人体检测,得到所述至少一个人体信息。

11. 根据权利要求3所述的方法,其特征在于,所述对所述待识别图像进行人体检测,得到至少一个人体信息之后,所述方法还包括:

利用所述预设关键点识别模型,对所述每个人体信息的所述人体姿态关键点进行多尺度识别,得到个体姿态关键点,从而得到与所述至少一个人体信息对应的至少一个个体姿态关键点;

根据所述至少一个个体姿态关键点,确定所述人体姿态信息,完成人体姿态识别。

12. 根据权利要求1或11所述的方法,其特征在于,所述得到所述人体姿态信息,完成人体姿态识别之后,所述方法还包括:

根据所述人体姿态信息,确定个体运动轨迹;所述个体运动轨迹用于确定决策信息。

13. 一种人体姿态识别装置,其特征在于,包括:

图像获取模块,用于获取待识别图像;所述待识别图像中包含人体的成像信息;
人体检测模块,用于对所述待识别图像进行人体检测,得到至少一个人体信息;
关键点识别模块,用于对所述待识别图像中的人体姿态关键点进行多尺度识别,得到多尺度关键点;

关键点融合模块,用于对所述多尺度关键点进行多尺度融合,得到目标关键点;
关键点分配模块,用于将所述目标关键点分配至每个人体信息中,得到人体姿态信息,完成人体姿态识别;所述人体姿态信息为所述每个人体信息的姿态所构成的集合。

14.一种人体姿态识别设备,其特征在于,包括:

存储器,用于存储可执行指令;

处理器,用于执行所述存储器中存储的可执行指令时,实现权利要求1至12任一项所述的方法。

15.一种存储介质,其特征在于,存储有可执行指令,用于引起处理器执行时,实现权利要求1至12任一项所述的方法。

一种人体姿态识别方法、装置、设备及存储介质

技术领域

[0001] 本发明涉及人工智能领域中的计算机视觉技术,尤其涉及一种人体姿态识别方法、装置、设备及存储介质。

背景技术

[0002] 人工智能(Artificial Intelligence, AI),是研究和开发用于模拟、延伸和扩展人的智能的理论、方法、技术及应用系统的一门新的科学技术。近年来,随着人工智能的发展,基于人工智能的计算机视觉技术也得到了快速发展,人体姿态识别作为其中的一个重要方法,在智慧商超、安防和监控等多个领域都有较大的应用前景。

[0003] 一般来说,进行人体姿态识别时,通常先检测出待识别图像中的关键点,然后将关键点分配给每个人,进而得到每个人的姿态信息。然而,上述人体姿态识别过程中,由于关键点的获取是从网络模型的最后一层输出的信息获得,而网络模型的第一层到最后一层的处理过程中,需要依次对待识别图像进行下采样操作,存在像素损失;因此,关键点的精度低,从而根据关键点实现人体姿态识别时,识别结果的准确度低。

发明内容

[0004] 本发明实施例提供一种人体姿态识别方法、装置及存储介质,能够提升人体姿态识别的准确度。

[0005] 本发明实施例的技术方案是这样实现的:

[0006] 本发明实施例提供一种人体姿态识别方法,包括:

[0007] 获取待识别图像;所述待识别图像中包含人体的成像信息;

[0008] 对所述待识别图像进行人体检测,得到至少一个人体信息;

[0009] 对所述待识别图像中的人体姿态关键点进行多尺度识别,得到多尺度关键点;

[0010] 对所述多尺度关键点进行多尺度融合,得到目标关键点;

[0011] 将所述目标关键点分配至每个人体信息中,得到人体姿态信息,完成人体姿态识别;所述人体姿态信息为所述每个人体信息的姿态所构成的集合。

[0012] 在上述方案中,所述利用所述目标样本图像,对所述原始关键点识别模型进行训练,得到所述预设关键点识别模型之后,所述方法还包括:

[0013] 获取新的目标样本图像;

[0014] 利用所述新的目标样本图像,优化所述预设关键点识别模型;

[0015] 相应地,所述利用预设关键点识别模型,对所述待识别图像中的所述人体姿态关键点进行多尺度识别,得到所述多尺度关键点,包括:

[0016] 利用优化后的预设关键点识别模型,对所述待识别图像中的所述人体姿态关键点进行多尺度识别,得到所述多尺度关键点。

[0017] 本发明实施例提供一种人体姿态识别装置,包括:

[0018] 图像获取模块,用于获取待识别图像;所述待识别图像中包含人体的成像信息;

- [0019] 人体检测模块,用于对所述待识别图像进行人体检测,得到至少一个人体信息;
- [0020] 关键点识别模块,用于对所述待识别图像中的人体姿态关键点进行多尺度识别,得到多尺度关键点;
- [0021] 关键点融合模块,用于对所述多尺度关键点进行多尺度融合,得到目标关键点;
- [0022] 关键点分配模块,用于将所述目标关键点分配至每个人体信息中,得到人体姿态信息,完成人体姿态识别;所述人体姿态信息为所述每个人体信息的姿态所构成的集合。
- [0023] 本发明实施例提供一种人体姿态识别设备,包括:
- [0024] 存储器,用于存储可执行指令;
- [0025] 处理器,用于执行所述存储器中存储的可执行指令时,实现本发明实施例提供的方法。
- [0026] 本发明实施例提供一种存储介质,存储有可执行指令,用于引起处理器执行时,实现本发明实施例提供的方法。
- [0027] 本发明实施例具有以下有益效果:由于用于确定待识别图像中每个人体信息的姿态的目标关键点,是由待识别图像的至少两种尺度下的关键点融合得到的;因此,目标关键点的准确度高;从而当根据目标关键点实现人体姿态识别时,识别结果的准确度高,也就提升了人体姿态识别的准确度。

附图说明

- [0028] 图1是本发明实施例提供的人体姿态识别系统100的一个可选的架构示意图;
- [0029] 图2是本发明实施例提供的服务器200的结构示意图;
- [0030] 图3是本发明实施例提供的人体姿态识别方法的一个可选的流程示意图;
- [0031] 图4是本发明实施例提供的人体检测的示例性示意图;
- [0032] 图5是本发明实施例提供的人体姿态识别方法的流程处理示意图;
- [0033] 图6是本发明实施例提供的提取多尺度特征的流程示意图;
- [0034] 图7是本发明实施例提供的多尺度融合的示例性示意图;
- [0035] 图8是本发明实施例提供的获取目标个体类型关键点的示例性示意图;
- [0036] 图9是本发明实施例提供的获取目标样本图像的一个可选流程示意图;
- [0037] 图10是本发明实施例提供的人体部位识别结果的示例性示意图;
- [0038] 图11是本发明实施例提供的人体部位数据库的示例性示意图;
- [0039] 图12是本发明实施例提供的示例性的融合示意图;
- [0040] 图13是本发明实施例提供的人体姿态识别方法的另一个可选流程示意图;
- [0041] 图14是本发明实施例提供的人体姿态识别方法的再一个可选的流程示意图;
- [0042] 图15是本发明实施例提供的人体姿态识别的示例性流程示意图;
- [0043] 图16是本发明实施例提供的人体姿态应用的示例性示意图一;
- [0044] 图17是本发明实施例提供的人体姿态应用的示例性示意图二。

具体实施方式

- [0045] 为了使本发明的目的、技术方案和优点更加清楚,下面将结合附图对本发明作进一步地详细描述,所描述的实施例不应视为对本发明的限制,本领域普通技术人员在没有

做出创造性劳动前提下所获得的所有其它实施例,都属于本发明保护的范围。

[0046] 在以下的描述中,涉及到“一些实施例”,其描述了所有可能实施例的子集,但是可以理解,“一些实施例”可以是所有可能实施例的相同子集或不同子集,并且可以在不冲突的情况下相互结合。

[0047] 除非另有定义,本发明实施例所使用的所有的技术和科学术语与属于本发明的技术领域的技术人员通常理解的含义相同。本发明实施例中所使用的术语只是为了描述本发明实施例的目的,不是旨在限制本发明。

[0048] 对本发明实施例进行进一步详细说明之前,对本发明实施例中涉及的名词和术语进行说明,本发明实施例中涉及的名词和术语适用于如下的解释。

[0049] 1) 人工神经网络,是一种模仿生物神经网络结构和功能的数学模型,包括输入层、中间层和输出层,每个层由大量处理单元相互联结构成,每个节点使用激励函数对输入的数据进行处理并输出给其他节点,激励函数的示例性的类型包括阈值型、线性型和S生长曲线(Sigmoid)型等。

[0050] 2) 机器学习(Machine Learning,ML),是一门多领域交叉学科,涉及概率论、统计学、逼近论、凸分析、算法复杂度理论等多门学科。专门研究计算机怎样模拟或实现人类的学习行为,以获取新的知识或技能;重新组织已有的知识结构使之不断改善自身的性能。机器学习是人工智能的核心,是使计算机具有智能的根本途径,其应用遍及人工智能的各个领域。机器学习和深度学习通常包括人工神经网络、置信网络、强化学习、迁移学习和归纳学习等技术。

[0051] 3) 深度学习(Deep Learning,DL),是机器学习领域中一个新的研究方向;深度学习是学习样本数据的内在规律和表示层次,最终目标是让机器能够像人一样具有分析学习能力,能够识别文字、图像和声音等数据;深度学习是一个复杂的机器学习算法。

[0052] 4) 卷积神经网络(Convolutional Neural Networks,CNN),是一类包含卷积计算且具有深度结构的前馈神经网络,是深度学习的代表算法之一。

[0053] 5) 部分亲和场(Partial Affinity Field,PAF),指一组可变数量的人体部位非结构化对应关系编码的流场表示;利用PAF对每对身体部位进行置信度测量,即属于同一个人体。

[0054] 6) 热度图,又称热力图,本发明实施例中特指将人体姿态关键点用热力点来展现的图像。

[0055] 随着人工智能技术的研究和进步,人工智能技术在多个领域展开了研究和应用;例如,常见的智能家居、智能穿戴设备、虚拟助理、智能音箱、智能营销、无人驾驶、自动驾驶、无人机、机器人、智能医疗和智能客服等;随着技术的发展,人工智能技术将在更多的领域得到应用,并发挥越来越重要的价值;比如,还可以将人工智能应用在计算机视觉领域。

[0056] 这里,需要说明的是,人工智能是计算机科学的一个综合技术,它企图了解智能的实质,并生产出一种新的能以人类智能相似的方式做出反应的智能机器。人工智能也就是研究各种智能机器的设计原理与实现方法,使机器具有感知、推理与决策的功能。

[0057] 另外,人工智能技术是一门综合学科,涉及领域广泛,既有硬件层面的技术也有软件层面的技术。人工智能基础技术一般包括如传感器、专用人工智能芯片、云计算、分布式存储、大数据处理技术、操作/交互系统、机电一体化等技术。人工智能软件技术主要包括计

计算机视觉技术、语音处理技术、自然语言处理技术以及机器学习/深度学习等几大方向。

[0058] 发明人在实施本发明的过程中发现,相关技术提供的方案中,人体姿态识别时,通常采用自上而下的识别方法和自下而上的识别方法两种。其中,自上而下的识别方法指先定位人体大概位置,再对每个人体信息的人体姿态进行具体识别;比如,先检测出待识别图像中每个人体信息的位置框,再在位置框的基础上针对每个人体信息进行人体姿态关键点检测,最终得到每个人体信息的人体姿态。而自下而上的识别方法指先检测出待识别图像中所有人体姿态的关键点,再将所有人体姿态的关键点聚类到不同的人体上;比如,先采用人体姿态热度图计算人体姿态的每个关键点,再采用部分亲和场将计算的关键点之间进行连接,最后采用图论的二分图求解方法得到每个人体信息的人体姿态。

[0059] 然而,对于自上而下的识别方法,当待识别图像中的人体靠近而产生空间干扰时检测效果差。所以,由于人体姿态识别大多为多个人体的场景,从而多选用自下而上的识别方法。而对于自下而上的识别方法,由于是先检测出图像中所有人体姿态的关键点,在将人体姿态的关键点分配给每个人体信息,因此,人体姿态的关键点的获取与人体姿态识别的结果至关重要,而人体姿态的关键点获取时,与网络模型的识别能力和网络模型的结构相关,通常网络模型对于易混淆肢体的区分能力差,并且,网络模型对待识别图像处理时,通常采用下采样的方式输出数据,以及将网络模型最后一层的输出结果作为模型的输出结果,存在像素损失;故,获得的人体姿态关键点的准确性差,从而根据关键点实现人体姿态识别时,识别结果的准确度低。

[0060] 基于此,本发明实施例提供一种人体姿态识别方法、装置、设备及存储介质,能够在人工智能的基础上实现人体姿态的识别,提升人体姿态识别的准确度。另外,本发明实施例提供的方案涉及人工智能的计算机视觉技术,比如,根据模型获得人体姿态信息;具体将在下文进行说明。

[0061] 下面说明本发明实施例提供的人体姿态识别设备的示例性应用,本发明实施例提供的人体姿态识别设备可以实施为智能手机、平板电脑、笔记本电脑等各种类型的用户终端,也可以实施为服务器。下面,将说明人体姿态识别设备实施为服务器时的示例性应用。

[0062] 参见图1,图1是本发明实施例提供的人体姿态识别系统100的一个可选的架构示意图,为实现支撑一个人体姿态识别应用,其中,服务器200为本发明实施例提供的人体姿态识别设备,终端400(示例性示出了终端400-1和终端400-2)通过网络300连接服务器200,网络300可以是广域网或者局域网,又或者是二者的组合。

[0063] 终端400用于搜集待识别图像并通过网络300将待识别图像发送至服务器200,并将通过网络300从服务器200获得的人体姿态信息通过在图形界面410(示例性示出了图形界面410-1和图形界面410-2)显示;服务器200用于通过网络300从终端400获得待识别图像进行人体姿态识别,并将识别出的人体姿态信息通过网络300发送至终端400进行显示。

[0064] 参见图2,图2是本发明实施例提供的服务器200的结构示意图,图2所示的服务器200包括:至少一个处理器210、存储器250、至少一个网络接口220和用户接口230。服务器200中的各个组件通过总线系统240耦合在一起。可理解,总线系统240用于实现这些组件之间的连接通信。总线系统240除包括数据总线之外,还包括电源总线、控制总线和状态信号总线。但是为了清楚说明起见,在图2中将各种总线都标为总线系统240。

[0065] 处理器210可以是一种集成电路芯片,具有信号的处理能力,例如通用处理器、数

字信号处理器 (DSP, Digital Signal Processor), 或者其他可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件等, 其中, 通用处理器可以是微处理器或者任何常规的处理器等。

[0066] 用户接口230包括使得能够呈现媒体内容的一个或多个输出装置231, 包括一个或多个扬声器和/或一个或多个视觉显示屏。用户接口230还包括一个或多个输入装置232, 包括有助于用户输入的用户接口部件, 比如键盘、鼠标、麦克风、触屏显示屏、摄像头、其他输入按钮和控件。

[0067] 存储器250包括易失性存储器或非易失性存储器, 也可包括易失性和非易失性存储器两者。其中, 非易失性存储器可以是只读存储器 (ROM, Read Only Memory), 易失性存储器可以是随机存取存储器 (RAM, Random Access Memory)。本发明实施例描述的存储器250旨在包括任意适合类型的存储器。存储器250可选地包括在物理位置上远离处理器210的一个或多个存储设备。

[0068] 在一些实施例中, 存储器250能够存储数据以支持各种操作, 这些数据的示例包括程序、模块和数据结构或者其子集或超集, 下面示例性说明。

[0069] 操作系统251, 包括用于处理各种基本系统服务和执行硬件相关任务的系统程序, 例如框架层、核心库层、驱动层等, 用于实现各种基础业务以及处理基于硬件的任务;

[0070] 网络通信模块252, 用于经由一个或多个 (有线或无线) 网络接口220到达其他计算设备, 示例性的网络接口220包括: 蓝牙、无线相容性认证 (Wi-Fi)、和通用串行总线 (USB, Universal Serial Bus) 等;

[0071] 显示模块253, 用于经由一个或多个与用户接口230相关联的输出装置231 (例如, 显示屏、扬声器等) 使得能够呈现信息 (例如, 用于操作外围设备和显示内容和信息的用户接口);

[0072] 输入处理模块254, 用于对一个或多个来自一个或多个输入装置232之一的一个或多个用户输入或互动进行检测以及翻译所检测的输入或互动。

[0073] 在一些实施例中, 本发明实施例提供的人体姿态识别装置可以采用软件方式实现, 图2示出了存储在存储器250中的人体姿态识别装置255, 其可以是程序和插件等形式的软件, 包括以下软件模块: 图像获取模块2551、人体检测模块2552、关键点识别模块2553、关键点融合模块2554、关键点分配模块2555、训练模块2556、识别模块2557和应用模块2558, 将在下文中说明各个模块的功能。

[0074] 在另一些实施例中, 本发明实施例提供的人体姿态识别装置可以采用硬件方式实现, 作为示例, 本发明实施例提供的人体姿态识别装置可以是采用硬件译码处理器形式的处理器, 其被编程以执行本发明实施例提供的人体姿态识别方法, 例如, 硬件译码处理器形式的处理器可以采用一个或多个应用专用集成电路 (ASIC, Application Specific Integrated Circuit)、DSP、可编程逻辑器件 (PLD, Programmable Logic Device)、复杂可编程逻辑器件 (CPLD, Complex Programmable Logic Device)、现场可编程门阵列 (FPGA, Field-Programmable Gate Array) 或其他电子元件。

[0075] 下面, 将结合本发明实施例提供的人体姿态识别设备实施为服务器的示例性应用和实施, 说明本发明实施例提供的人体姿态识别方法。

[0076] 参见图3, 图3是本发明实施例提供的人体姿态识别方法的一个可选的流程示意

图,将结合图3示出的步骤进行说明。

[0077] S101、获取待识别图像;待识别图像中包含人体的成像信息。

[0078] 在本发明实施例中,人体姿态识别设备在进行人体姿态识别时,所获得的对象即待识别图像;这里,待识别图像中包含人体的成像信息,比如,人物图像。

[0079] 需要说明的是,待识别图像可以是拍摄装置拍得的图像,也可以是拍摄装置拍得的视频对应的图像序列中的图像,等等,本发明实施例对此不作具体限定。另外,人体姿态识别指通过检测人体的关键点(如关节,五官等),并通过检测到的关键点描述人体骨骼信息的过程。另外,待识别图像为包含人体的成像信息。

[0080] S102、对待识别图像进行人体检测,得到至少一个人体信息。

[0081] 需要说明的是,待识别图像中包含至少一个人体信息的成像信息,因此,当人体姿态识别设备对待识别图像中的人体进行检测时,能够获得至少一个人体信息。

[0082] 在本发明实施例中,人体姿态识别设备对待识别图像进行人体检测时,返回的每个人体所处待识别图像中的位置信息(比如,区域框)即至少一个人体信息(比如,至少一个区域框)。

[0083] 示例性地,参见图4,图4是本发明实施例提供的人体检测的示例性示意图,如图4所示,待识别图像4-1上检测出了4-21和4-22两个人体信息。

[0084] S103、对待识别图像中的人体姿态关键点进行多尺度识别,得到多尺度关键点。

[0085] 在本发明实施例中,人体姿态识别设备获得了待识别图像之后,从多个尺度对待识别图像中的人体姿态关键点进行识别,从而也就获得了每种尺度下所识别出的关键点,该每种尺度下所识别出的关键点构成的集合即多尺度关键点。

[0086] 需要说明的是,人体姿态识别设备进行人体姿态识别时,预先设定了人体姿态关键点,即S101中描述的人体的关键点;相应地,预设关键点识别模型的训练过程中是针对人体姿态识别设备预先设定的人体姿态关键点训练的;因此,多尺度关键点为人体姿态关键点在至少两种尺度下的信息,并且,人体姿态关键点中包含至少一个关键点。比如,人体姿态关键点包括:头、颈、左肩、左肘、左腕、右肩、右肘、右腕、左髋、左膝、左踝、右髋、右膝和右踝14个关键点;如此,当从3种尺度对待识别图像进行人体姿态关键点识别时;例如,对于头关键点,则对应3种尺度下的关键点信息;对于右踝,则对应3种尺度下的关键点信息。

[0087] S104、对多尺度关键点进行多尺度融合,得到目标关键点。

[0088] 在本发明实施例中,人体姿态识别设备获得了多尺度关键点之后,由于多尺度关键点中包括了人体姿态关键点对应的至少两种尺度下的关键点信息;因此,人体姿态识别设备将至少两种尺度下的关键点信息进行融合,也就能够获得待识别图像中人体姿态关键点的识别结果,即目标关键点。也就是说,目标关键点为人体姿态识别设备通过对待识别图像从至少两种尺度下进行的人体姿态关键点进行识别所确定的人体的关键点。

[0089] 需要说明的是,S102与S103-S104在执行顺序上不分先后,S102与S103-S104可以同时执行,还可以交叉执行,又可以延迟执行,也可以先后执行,等等,本发明实施例对此不作具体限定。而图3示出的为S102与S103-S104先后执行的示意图。

[0090] S105、将目标关键点分配至每个人体信息中,得到人体姿态信息,完成人体姿态识别;人体姿态信息为每个人体信息的姿态所构成的集合。

[0091] 在本发明实施例中,人体姿态识别设备获得了待识别图像中的目标关键点和至少

一个人体信息之后,由于目标关键点是从待识别图像中识别出的每个人体的人体姿态关键点所构成的集合;因此,人体姿态识别设备确定了目标关键点中属于至少一个人体信息中的每个人体信息的关键点之后,也就完成了目标关键点向至少一个人体信息的每个人体信息的分配;此时,人体姿态识别设备也就获得了每个人体信息对应的关键点,也就获得了每个人体信息的姿态,从而也就获得了至少一个人体信息的姿态,即人体姿态信息,完成了人体姿态识别。也就是说,人体姿态信息为每个人体信息的姿态所构成的集合。

[0092] 这里,每个人体信息为至少一个人体信息中的人体信息,人体姿态信息为每个人体信息的姿态所构成的集合。

[0093] 参见图5,图5是本发明实施例提供的人体姿态识别方法的流程处理示意图,如图5所示,待识别图像5-1一方面经过多尺度识别5-2和多尺度融合5-3处理之后,得到目标关键点5-4;另一方面经过人体检测5-5处理之后,得到至少一个人体信息5-6;接下来,将目标关键点5-4与至少一个人体信息5-6进行绑定5-7处理,也就得到了人体姿态信息5-8,从而也就完成了对待识别图像5-1的人体姿态识别。

[0094] 可以理解的是,人体姿态识别设备通过融合从待识别图像中识别出的多尺度下的人体姿态关键点,来获得待识别图像中人体的所有关键点即目标关键点,实现了有效利用待识别图像的像素的方案,避免了获取目标关键点过程中的像素损失;故,所获得的目标关键点能够有效表征待识别图像中的人体姿态关键点,准确度高;从而当根据目标关键点实现人体姿态识别时,识别结果的准确度高,也就提升了人体姿态识别的准确度。

[0095] 进一步地,在本发明实施例中,S103可以通过S1031-S1032实现;也就是说,人体姿态识别设备对待识别图像中的人体姿态关键点进行多尺度识别,得到多尺度关键点,包括S1031-S1032,下面将结合各步骤进行说明。

[0096] S1031、对待识别图像进行多尺度特征提取,得到目标多尺度特征。

[0097] 在本发明实施例中,人体姿态识别设备获取待识别图像的多尺度下的图像,并对多尺度下的图像分别进行特征提取,每种尺度下的图像中所提取的特征共同组合成目标多尺度特征。

[0098] S1032、对目标多尺度特征进行人体姿态关键点识别,得到多尺度关键点。

[0099] 在本发明实施例中,人体姿态识别设备获得了目标多尺度特征之后,对目标多尺度特征中的人体姿态关键点进行识别,也就能够获得具有多尺度特征的多尺度关键点。这里,多尺度特征指每个关键点对应存在多个尺度下的关键点信息。

[0100] 进一步地,在本发明实施例中,S103中人体姿态识别设备对待识别图像中的人体姿态关键点进行多尺度识别,得到多尺度关键点还可以通过以下步骤实现:利用预设关键点识别模型,对待识别图像中的人体姿态关键点进行多尺度识别,得到多尺度关键点;其中,预设关键点识别模型用于从至少两种尺度对待识别图像中的人体姿态关键点进行识别。

[0101] 需要说明的是,人体姿态识别设备中预先通过训练获得了预设关键点识别模型,用于从至少两种尺度对待识别图像中的人体姿态关键点进行识别;因此,当人体姿态识别设备获得了待识别图像之后,就能够利用预设关键点识别模型,对待识别图像中的人体姿态关键点进行多尺度识别了,所得到的识别结果即多尺度关键点。

[0102] 另外,由于预设关键点识别模型用于从至少两种尺度对待识别图像中的人体姿态

关键点进行识别,因此,识别出的多尺度关键点中的每个关键点对应至少两种尺度下的关键点信息。

[0103] 这里,预设关键点识别模型包括多尺度特征提取模型和多尺度关键点输出模型,多尺度特征提取模型用于从至少两种尺度对待识别图像进行特征提取,多尺度关键点输出模型用于输出至少两种尺度下的关键点信息。

[0104] 相应地,S1031即人体姿态识别设备利用多尺度特征提取模型,对待识别图像进行多尺度特征提取,得到目标多尺度特征。

[0105] 在本发明实施例中,预设关键点识别模型是多任务的、多阶段的,这里的多任务或多阶段是通过多尺度特征提取模型实现的;这里,多尺度特征提取模型对待识别图像进行多尺度转换,得到多尺度待识别图像,此时,人体姿态识别设备通过多尺度特征提取模型对多尺度待识别图像中每个尺度下的待识别图像进行特征提取,也就实现了对待识别图像进行多尺度特征提取,而所提取到的待识别图像下的每个尺度下的特征所构成的集合即目标多尺度特征。

[0106] 需要说明的是,多尺度特征提取模型为金字塔分支结构的卷积神经网络,用于从多个尺度对待识别图像进行特征提取;因此,多尺度特征提取模型包括多个网络分支,其中,每个网络分支用于对待识别图像进行一种尺度下的特征提取;从而通过多个网络分支交互,也就获得了目标多尺度特征。

[0107] 示例性地,参见图6,图6是本发明实施例提供的提取多尺度特征的流程图,如图6所示,多尺度特征提取模型6-2对待识别图像6-1进行了三种尺度(6-21、6-22和6-23)下的特征提取,来获得目标多尺度特征6-3。

[0108] 可以理解的是,通过多尺度特征提取模型从多个尺度对待识别图像进行特征提取,能够在人体姿态识别的初始阶段提取到更丰富的图像特征,更充分利用待识别图像的有效像素,为后续人体姿态识别的准确度的提升提供了前提条件;同时,由于人体姿态识别的准确度高,也消除了肢体混淆导致识别结果不准确的问题。

[0109] 另外,S1032即人体姿态设备利用多尺度关键点输出模型,对目标多尺度特征进行人体姿态关键点识别,得到多尺度关键点。

[0110] 在本发明实施例中,预设关键点识别模型还包括多尺度关键点输出模型,用于根据目标多尺度特征,输出识别出的待识别图像的多尺度下的人体姿态关键点所对应的热度图,即多尺度关键点;因此,当人体姿态识别设备获得了目标多尺度特征之后,利用多尺度关键点输出模型对目标多尺度特征进行人体姿态关键点识别,也就获得了多尺度关键点。

[0111] 需要说明的是,多尺度关键点输出模型为至少两个沙漏结构模型组成的模型,这里,至少两种尺度所对应的尺度的数量,与多尺度关键点输出模型中沙漏结构模型的数量一致;其中,每个沙漏结构模型分成上下两个分支,上分支结果作为下一个沙漏结构模型的输入,下分支结果作为输出结果;如此,当目标多尺度特征经过多尺度关键点输出模型之后,也就能够获得至少两个下分支结果,也就获得了多尺度关键点。这里,下分支结果为识别出的人体姿态关键点对应的热度图,如此,也就获得了至少两种尺度下的人体姿态关键点对应的热度图。

[0112] 进一步地,在本发明实施例中,S104可以通过S1041-S1043实现;也就是说,人体姿态识别设备对多尺度关键点进行多尺度融合,得到目标关键点,包括S1041-S1043,下面将

结合各步骤进行说明。

[0113] S1041、获取多尺度关键点的局部最大值,得到预测多尺度关键点。

[0114] 在本发明实施例中,由于多尺度关键点为多尺度下的热度图;因此,人体姿态识别设备获得了多尺度关键点之后,基于每张热度图中各关键点的位置信息计算多尺度关键点的局部最大值,也就预测出了待识别图像的人体姿态关键点,即预测多尺度关键点。

[0115] S1042、基于多尺度关键点,对预测多尺度关键点进行线性组合处理,得到初始多尺度关键点。

[0116] 在本发明实施例中,人体姿态识别设备预测出了预测多尺度关键点之后,为了进一步准确地确定待识别图像的人体姿态关键点,利用多尺度关键点对预测出的预测多尺度关键点的准确性进行提升;这里,人体姿态识别设备通过基于多尺度关键点,对预测多尺度关键点进行线性组合处理,来提升预测多尺度关键点的准确性,所提升后的预测多尺度关键点即初始多尺度关键点。也就是说,初始多尺度关键点的准确性比预测多尺度关键点的准确性高。

[0117] S1043、对初始多尺度关键点进行多尺度融合处理,得到目标关键点。

[0118] 在本发明实施例中,人体姿态识别设备获得了初始多尺度关键点之后,以初始多尺度关键点中每个关键点为单位进行多尺度融合,也就获得了目标关键点。

[0119] 进一步地,在本发明实施例中,S1042可以通过S10421-S10422实现;也就是说,人体姿态识别设备基于多尺度关键点,对预测多尺度关键点进行线性组合处理,得到初始多尺度关键点,包括S10421-S10422,下面将结合各步骤进行说明。

[0120] S10421、从多尺度关键点上,获取预设数量的与预测多尺度关键点的距离最小的相邻关键点。

[0121] 在本发明实施例中,人体姿态识别设备提升预测多尺度关键点的准确性,是通过将预测多尺度关键点,与多尺度关键点中预测多尺度关键点附近的关键点进行比较实现;这里,人体姿态识别设备将多尺度关键点上预设数量(比如,4)的与预测多尺度关键点的距离最小的相邻关键点,作为多尺度关键点中预测多尺度关键点附近的点。

[0122] S10422、对预测多尺度关键点和相邻关键点进行线性组合,得到初始尺度关键点。

[0123] 需要说明的是,人体姿态识别设备获得了相邻关键点之后,对预测多尺度关键点和相邻关键点进行线性组合,所采用的线性组合方式取决于预测多尺度关键点和相邻关键点的比较结果;比如预测多尺度关键点和相邻关键点中的右关键点距离较近(小于等于阈值),而和相邻关键点中的其他关键点距离较远(大于阈值),则线性组合方式为取预测多尺度关键点和相邻关键点中的右关键点的均值,来获得初始尺度关键点。

[0124] 进一步地,在本发明实施例中,初始多尺度关键点包括至少一个多尺度目标关键点,至少一个多尺度目标关键点的每个多尺度目标关键点由至少两种尺度下每种尺度对应的关键点构成的集合组成;这里,S1043可以通过S10431-S10433实现;也就是说,人体姿态识别设备对初始多尺度关键点进行多尺度融合处理,得到目标关键点,包括S10431-S10433,下面将结合各步骤进行说明。

[0125] S10431、获取初始多尺度关键点中每个多尺度目标关键点的目标中心点。

[0126] 在本发明实施例中,人体姿态识别设备针对初始多尺度关键点中每个多尺度目标关键点,计算每个多尺度目标关键点对应的至少两种尺度下每种尺度对应的关键点的中心

位置,也就获得了每个多尺度目标关键点的目标中心点。

[0127] 需要说明的是,由于初始多尺度关键点包括至少一个多尺度目标关键点,当获得了每个多尺度目标关键点对应的目标中心点之后,也就获得了至少一个目标中心点。另外,至少一个多尺度目标关键点的数量大于等于预设的人体姿态关键点的数量;比如,当预设的人体姿态关键点的数量为14时,至少一个多尺度目标关键点的数量可以为14或15。

[0128] 示例性地,参见图7,图7是本发明实施例提供的多尺度融合的例子性示意图,如图7所示,在位置坐标中,11个星型点7-1为至少一个多尺度目标关键点中的一个多尺度目标关键点,实心原点7-2为这一个多尺度目标关键点对应的目标中心点。

[0129] S10432、计算每个多尺度目标关键点中每种尺度对应的关键点,到目标中心点的距离,得到目标距离。

[0130] 在本发明实施例中,人体姿态识别设备获得了目标中心点之后,为了提高所获得的目标关键点的准确度,计算每个多尺度目标关键点中每种尺度对应的关键点,到目标中心点的距离,也就得到了目标距离;以根据目标距离排除离群点。

[0131] S10433、将目标距离中小于预设距离阈值的均值的均值,作为与每个多尺度目标关键点对应的子目标关键点,从而得到与初始多尺度关键点对应的目标关键点。

[0132] 在本发明实施例中,人体姿态识别设备中设置了预设距离阈值,用来排除每个多尺度目标关键点中的离群点;这里,人体姿态识别设备将目标距离中大于等于预设距离阈值所对应的关键点,从每个多尺度目标关键点中排除掉,并将目标距离中小于预设距离阈值的均值的均值,作为与每个多尺度目标关键点对应的子目标关键点。

[0133] 需要说明的是,获得了每个多尺度目标关键点对应的子目标关键点之后,也就获得了至少一个多尺度目标关键点对应的至少一个子目标关键点,即目标关键点。

[0134] 示例性地,继续参见图7,图7中的7-3为预设距离阈值,圆形区域7-4为以目标中心点7-2为圆心,以预设距离阈值7-3为半径所获得的区域,易知,11个星型点7-1中落在圆形区域7-4中的星型点为所需星型点,对落在圆形区域7-4中的星型点求均值也就获得了一个子目标关键点7-5。

[0135] 可以理解的是,人体姿态识别设备通过将多尺度下识别的关键点进行融合,考虑了多尺度下每个关键点的置信度,有效排除了离群点,提升了目标关键点的准确度。

[0136] 进一步地,在本发明实施例中,S105可以通过S1051-S1054实现;也就是说,人体姿态识别设备将目标关键点分配至每个人体信息中,得到人体姿态信息,包括S1051-S1054,下面将结合各步骤进行说明。

[0137] S1051、从目标关键点中,确定属于每个人体信息的个体关键点。

[0138] 在本发明实施例中,人体姿态识别设备基于至少一个人体信息的每个人体信息的位置信息和目标关键点的位置信息,确定属于至少一个人体信息的每个人体信息的个体关键点。这里,人体姿态识别设备依据至少一个人体信息的每个人体信息,将目标关键点划分成了至少一个个体关键点;并且,由于每个人体之间存在交叠等情况,因此,至少一个个体关键点中每个个体关键点之间可能存在相同的点,所以,需要对每个个体关键点进行筛选处理。

[0139] 示例性地,当至少一个人体信息为至少一个人体区域框时,目标关键点中落在每个人体区域框的关键点即为该人体对应的个体关键点。

[0140] S1052、从个体关键点中,获取人体姿态关键点中每个关键点类型所对应的个体类型关键点。

[0141] 在本发明实施例中,人体姿态识别设备中的人体姿态关键点包括至少一种关键点类型,当获得了个体关键点之后,由于个体关键点中还包括关键点类型信息,用于表明该个体关键点中的每个关键点所属的关键点类型;从而,人体姿态识别设备能够从个体关键点中,获取人体姿态关键点中每个关键点类型所对应的关键点,即个体类型关键点;从而,针对个体关键点,也就获得了至少一个个体类型关键点。

[0142] S1053、从个体类型关键点中,获取目标个体类型关键点,得到与每个人体信息对应的个体姿态关键点。

[0143] 在本发明实施例中,由于个体关键点中每个关键点类型下仅有一个关键点,比如,一个人体仅有一个头、一个颈、一个左肩、一个左肘、一个左腕、一个右肩、一个右肘、一个右腕、一个左髋、一个左膝、一个左踝、一个右髋、一个右膝和一个右踝;从而,人体姿态识别设备需要对每个个体类型关键点中多余的关键点进行删除,此时,也就从个体类型关键点中获得了目标个体类型关键点,该目标个体类型关键点即每个关键点类型下仅有的一个关键点。而得到了每个人体信息的每个关键点类型对应的目标个体类型关键点,至少一个目标个体类型关键点也就构成了每个人体信息对应的个体姿态关键点;这里,至少一个目标个体类型关键点的数量与人体姿态关键点的类型的数量一致。

[0144] 这里,人体姿态识别设备可以通过训练筛选目标个体类型关键点的模型,训练时针对每个关键点类型输入多个关键点,并在输入的多个关键点上标注目标个体类型关键点,也就能够获得用于筛选目标个体类型关键点的模型。

[0145] 参见图8,图8是本发明实施例提供的获取目标个体类型关键点的示例性示意图,如图8所示,人体信息8-1的头部关键点有3个(黑色实心圆圈),通过筛选,也就获得了头部对应的目标个体类型关键点8-2。

[0146] S1054、根据个体姿态关键点,得到每个人体信息对应的个体姿态,从而得到所述至少一个人体信息对应的人体姿态信息。

[0147] 需要说明的是,个体姿态关键点包括至少一个目标个体类型关键点,人体姿态识别设备将至少一个目标个体类型关键点按照预设连接顺序进行连接,也就获得了每个人体信息对应的姿态,从而也就获得了至少一个人体信息对应的人体姿态信息。这里,预设连接顺序为人体姿态关键点之间的连接顺序,比如头与颈相连。

[0148] 进一步地,在本发明实施例中,人体姿态识别设备还包括训练获得预设关键点识别模型的过程,即在S102之前还包括S106-S108;也就是说,人体姿态识别设备对所述待识别图像进行人体检测,得到至少一个人体信息之前,该人体姿态识别方法还包括S106-S108,下面将结合各步骤进行说明。

[0149] S106、获取目标样本图像;目标样本图像用于训练得到预设关键点识别模型。

[0150] 在本发明实施例中,人体姿态识别设备通过训练获得预设关键点识别模型时,需先获取训练样本,即目标样本图像。这里,目标样本图像与待识别图像类型,同样也包含人体的成像信息。

[0151] S107、搭建原始关键点识别模型;原始关键点识别模型为多尺度层次结构的模型。

[0152] 在本发明实施例中,由于通过训练最终获得的预设关键点识别模型用于从至少两

种尺度对待识别图像中的人体姿态关键点进行识别;因此,人体姿态识别设备需要搭建具有多尺度层级结构的模型,即上述的金字塔分支结构的卷积神经网络,也就获得了原始关键点识别模型。

[0153] 需要说明的是,由于预设关键点识别模型包括多尺度特征提取模型和多尺度关键点输出模型;因此,原始关键点识别模型包括原始多尺度特征提取模型和原始多尺度关键点输出模型。

[0154] S108、利用目标样本图像,对原始关键点识别模型进行训练,得到预设关键点识别模型。

[0155] 在本发明实施例中,人体姿态识别设备获得了目标样本图像和原始关键点识别模型之后,就能够利用目标样本图像持续对原始关键点识别模型进行训练了,直至训练的原始关键点识别模型达到收敛平衡时结束训练;此时,也就完成了对原始关键点识别模型的训练,也就获得了预设关键点识别模型。

[0156] 另外,训练的过程与S103中描述的使用过程类型,本发明实施例在此不再赘述。

[0157] 需要说明的是,人体姿态识别设备对原始关键点识别模型进行训练的过程,即对原始多尺度特征提取模型和原始多尺度关键点输出模型训练的过程。

[0158] 进一步地,参见图9,图9是本发明实施例提供的获取目标样本图像的一个可选流程示意图,如图9所示,在本发明实施例中,S106可以通过S1061-S1064实现;也就是说,人体姿态识别设备获取目标样本图像,包括S1061-S1064,下面结合图9示出的步骤进行说明。

[0159] S1061、获取人体部位样本图像和待识别样本图像。

[0160] 在本发明实施例中,目标样本图像为经过数据增强处理后的图像,这里,数据增强处理所用到的图像包括人体部位样本图像和待识别样本图像,因此,人体姿态识别设备需要获得人体部位样本图像和待识别样本图像。

[0161] 需要说明的是,人体部位样本图像为用于对待识别样本图像进行数据增强处理得到目标样本图像的图像;而待识别样本图像为人体姿态识别设备初始获得的样本图像。

[0162] S1062、对人体部位样本图像的人体部位进行识别,得到人体部位识别结果。

[0163] 在本发明实施例中,人体姿态识别设备利用人体部位样本图像对待识别样本图像进行数据增强处理时,是利用人体部位样本图像中的各人体部位进行的;因此,人体姿态识别设备需先对人体部位样本图像的人体部位进行识别,所识别出的各人体部位即人体部位识别结果。

[0164] 这里,人体姿态识别设备可以基于人体部位的像素特征对人体部位样本图像的人体部位进行识别;比如,通过卷积神经网络基于人体部位的像素特征实现对人体部位样本图像的人体部位的识别。

[0165] 示例性地,参见图10,图10是本发明实施例提供的人体部位识别结果的示例性示意图,如图10所示,人体部位样本图像10-1的识别结果为人体部位识别结果10-2。

[0166] S1063、基于人体部位识别结果,从人体部位样本图像中提取人体部位,得到人体部位数据库。

[0167] 在本发明实施例中,人体姿态识别设备获得了人体部位识别结果之后,将识别出的人体部位识别结果作为掩膜,在人体部位样本图像中进行人体部位的分割,所分割出的各人体部位也就组成了人体部位数据库。

[0168] 示例性地,参见图11,图11是本发明实施例提供的人体部位数据库的示例性示意图,如图11所示,人体部位数据库11-1中包括头部11-11、上肢11-12、脚部11-13、腿部11-14和上身11-15。

[0169] S1064、将人体部位数据库中的人体部位与待识别样本图像中对应的人体部位进行融合,得到目标样本图像。

[0170] 在本发明实施例中,人体姿态识别设备获得了人体部位数据库之后,将人体部位数据库中的人体部位与待识别样本图像中对应的人体部位进行融合,也就得到了目标样本图像。

[0171] 需要说明的是,人体部位数据库中的各人体部位对应着类型,人体姿态识别设备基于人体部位的像素特征对待识别样本图像的人体部位进行识别,并根据待识别样本图像的识别结果与各人体部位对应的类型,实现人体部位数据库中的人体部位与待识别样本图像中对应的人体部位的融合。

[0172] 示例性地,参见图12,图12是本发明实施例提供的示例性的融合示意图,如图12所示,将上肢12-1贴合在待识别样本图像12-2中人体的上部附近(小于预设贴合距离的位置),将脚部12-3贴合在待识别样本图像12-2中人体的脚部附近,将腿部12-4贴合在待识别样本图像12-2中人体的腿部附近,从而也就获得了目标样本图像12-5。

[0173] 可以理解的是,通过采用数据增强的处理方式获得目标样本图像,从而在利用目标样本图像训练得到的预设关键点识别模型,对于易混淆人体关键点检测时,精度和准确度高。

[0174] 进一步地,在本发明实施例中,S108之后还包括S109-S110;也就是说,人体姿态识别设备利用目标样本图像,对原始关键点识别模型进行训练,得到预设关键点识别模型之后,该人体姿态识别方法还包括S109-S110,下面将结合各步骤进行说明。

[0175] S109、获取新的目标样本图像。

[0176] 在本发明实施例中,为了提升预设关键点识别模型的泛化能力,人体姿态识别设备还能够获取到新的目标样本图像;这里,新的目标样本图像,与目标样本图像类似,包含人体的成像信息;另外,新的目标样本图像用于优化预设关键点识别模型。

[0177] S110、利用新的目标样本图像,优化预设关键点识别模型。

[0178] 需要说明的是,人体姿态识别设备获得了新的目标样本图像之后,就可以利用新的目标样本图像,持续对预设关键点识别模型进行优化了,直到达到优化停止条件(比如,损失函数低于预设损失值)时结束优化,此时,也就获得了优化后的预设关键点识别模型。

[0179] 另外,优化的过程与S108中的训练过程一致,本发明实施例在此不再赘述。

[0180] 相应地,S103可通过以下步骤实现,即人体姿态识别设备利用预设关键点识别模型,对待识别图像中的人体姿态关键点进行多尺度识别,得到多尺度关键点,包括:利用优化后的预设关键点识别模型,对待识别图像中的人体姿态关键点进行多尺度识别,得到多尺度关键点。也就是说,人体姿态识别设备利用最新的模型进行多尺度识别。

[0181] 可以理解的是,人体姿态识别设备通过获取新的目标样本图像实现预设关键点识别模型的优化,提升了预设关键点识别模型的泛化能力;从而,当人体姿态识别设备利用优化后的预设关键点识别模型,对待识别图像中的人体姿态关键点进行多尺度识别,所获得的多尺度关键点的准确度得到了进一步提高。

[0182] 进一步地,在本发明实施例中,S106之后还包括S111-S112;也就是说,人体姿态识别设备获得了目标样本图像之后,该人体姿态识别方法还包括S111-S112,下面将结合各步骤进行说明。

[0183] S111、搭建原始人体检测模型;原始人体检测模型为用于识别出待识别图像中的每个人体信息的初始模型。

[0184] 在本发明实施例中,人体姿态识别设备搭建用于识别出待识别图像中的每个人体信息的初始模型,也就完成了原始人体检测模型的搭建。这里,原始人体检测模型比如为深度神经网络模型。

[0185] S112、利用目标样本图像,训练原始人体检测模型,得到人体检测模型;人体检测模型用于识别出待识别图像中的每个人体信息。

[0186] 在本发明实施例中,目标样本图像还用于训练得到用于识别出待识别图像中的每个人体信息的模型;因此,但人体姿态识别设备搭建好原始人体检测模型之后,利用目标样本图像,持续对原始人体检测模型进行训练,直到达到收敛时结束训练,此时也就获得了人体检测模型,也就能够利用人体检测模型进行待识别图像中的每个人体信息的识别了。

[0187] 相应地,S102中人体姿态识别设备对待识别图像进行人体检测,得到至少一个人体信息,可以通过以下步骤实现:人体姿态识别设备利用人体检测模型,对待识别图像进行人体检测,得到至少一个人体信息。

[0188] 进一步地,参见图13,图13是本发明实施例提供的人体姿态识别方法的另一个可选流程示意图,如图13所示,人体姿态识别设备还可以采用自上而下的识别方法进行人体姿态的识别。也就是说,S102之后还包括S113-S114,即人体姿态识别设备对待识别图像进行人体检测,得到至少一个人体信息之后,该人体姿态识别方法还包括S113-S114,下面将结合图13示出的步骤进行说明。

[0189] S113、利用预设关键点识别模型,对至每个人体的人体姿态关键点进行多尺度识别,得到个体姿态关键点,从而得到与至少一个人体信息对应的至少一个个体姿态关键点。

[0190] 在本发明实施例中,人体姿态识别设备获得了至少一个人体信息之后,利用预设关键点识别模型逐个对至少一个人体信息的每个人体信息的人体姿态关键点进行多尺度识别,也就获得了与每个人体信息对应的个体姿态关键点,从而也就获得了与至少一个人体信息对应的至少一个个体姿态关键点。

[0191] 需要说明的是,人体姿态识别设备针对每个个体姿态关键点采用本发明实施例提供的多尺度融合处理以及筛选(S1052-S1053)处理,来获得每个人体信息对应的人体姿态关键点,本发明实施例在此不再赘述。

[0192] S114、根据至少一个个体姿态关键点,确定人体姿态信息,完成人体姿态识别。

[0193] 在本发明实施例中,人体姿态识别设备获得了至少一个个体姿态关键点之后,依据人体姿态关键点之间的预设连接顺序进行连接,也就获得了每个人体信息对应的姿态,从而也就获得了至少一个人体信息对应的人体姿态信息,完成了人体姿态识别。

[0194] 进一步地,参见图14,图14是本发明实施例提供的人体姿态识别方法的再一个可选的流程示意图,如图14所示,基于图3,S105之后还包括S115;也就是说,人体姿态识别设备确定人体姿态信息,完成人体姿态识别之后,该人体姿态识别方法还包括S115:根据人体姿态信息,确定个体运动轨迹;个体运动轨迹用于确定决策信息。

[0195] 需要说明的是,当待识别图像为至少一张图像时,人体姿态设备获得了人体姿态信息之后,也就能够获得每个人体信息的运动轨迹即个体运动轨迹;这里,个体运动轨迹用于确定决策信息,比如,根据人体姿态信息获取目标用户在超市货物架边的脸的朝向,从而获知目标用户所关注的货物。另外,人体姿态识别设备获得了个体运动轨迹之后还可以通过显示设备进行显示。

[0196] 参见图15,图15是本发明实施例提供的人体姿态识别的示例性流程示意图,如图15所示,一方面,多尺度特征提取模型15-2对待识别图像15-1从三种尺度进行特征的提取,之后,多尺度关键点输出模型15-3输出多张每种尺度下的热度图,之后,多尺度融合模块15-4对多张每种尺度下的热度图上的关键点进行融合,得到目标关键点。另一方面,人体检测模块15-5对待识别图像15-1中的人体进行检测,得到两个人体信息。此时,人体姿态识别设备将目标关键点分配在这两个人体信息上,进而针对每个人体信息进行关键点的确定和筛选,使得每个人体信息中,对于人体姿态关键点的每个关键点类型仅保留一个关键点,也就得到了人体姿态信息15-6。

[0197] 需要说明的是,人体姿态识别是计算机视觉的研究,在计算机视觉的其他相关领域的研究中都起到了基础性的作用,如行为识别、人物跟踪、步态识别等相关领域。人体姿态识别的应用主要集中在智能视频监控、病人监护系统、人机交互、虚拟现实、人体动画、智能家居、智能安防和运动员辅助训练等。

[0198] 可以理解的是,本发明实施例通过人体姿态的识别,一方面,可以实现动作识别,即追踪一段时间内一个人体姿态的变化,也可以应用在动作、手势和步态识别上,比如用于检测一个人是否摔倒或疾病,或者用于健身、体育和舞蹈等的自动教学,或者用于理解全身的肢体语言(如机场跑道信号、交警信号等),或者用于增强安保和监控。另一方面,可以实现运动捕捉和增强现实,即将图形、风格、特效增强、设备和艺术造型等加载在检测到的人体姿态信息上,通过追踪人体姿态的变化(人体姿态运动轨迹),渲染的图形可以在人动的时候“自然”地与人“融合”。再一方面,还可以实现训练机器人。

[0199] 还需要说明的是,本发明实施例所提供的人体姿态识别方法,可以应用于其他的生物体上,以获得其他生物体的姿态信息。

[0200] 下面,将说明本发明实施例在一个实际的应用场景中的示例性应用。

[0201] 参见图16,图16是本发明实施例提供的人体姿态应用的示例性示意图一,如图16所示,目标用户(人体)16-1在街道16-2上的个体运动轨迹是16-31、16-32、16-33、16-34、16-35和16-36;其中,显示区域16-4显示的为目标用户16-1的基本信息,显示区域16-5显示的为目标用户16-1的个体运动轨迹的详细信息,显示区域16-6显示的为目标用户16-1在街道16-2上的个体运动轨迹,并且,显示区域16-6中的播放轨迹按钮被触发时,能够播放目标用户16-1的个体运动轨迹。如此,比如,对进入商场的顾客进行人体追踪及轨迹分析,能够分析还原出顾客真实的购物兴趣和购买行为,如此,能够促进后续定制化的端到端服务及广告推送,在进一步促进营收的同时帮助顾客提升购物体验。

[0202] 参见图17,图17是本发明实施例提供的人体姿态应用的示例性示意图二,如图17所示,个体运动轨迹还可以辅助人头人体绑定跟踪:目标用户(人体)17-1、目标用户17-2和目标用户17-3是要出电梯,而目标用户17-4和目标用户17-5是要进电梯。如此,通过人体姿态识别,能够准确丰富地获得人体运动轨迹。

[0203] 下面继续说明本发明实施例提供的人体姿态识别装置255的实施为软件模块的示例性结构,在一些实施例中,如图2所示,存储在存储器250的人体姿态识别装置255中的软件模块可以包括:

[0204] 图像获取模块2551,用于获取待识别图像;所述待识别图像中包含人体的成像信息;

[0205] 人体检测模块2552,用于对所述待识别图像进行人体检测,得到至少一个人体信息;

[0206] 关键点识别模块2553,用于对所述待识别图像中的人体姿态关键点进行多尺度识别,得到多尺度关键点;

[0207] 关键点融合模块2554,用于对所述多尺度关键点进行多尺度融合,得到目标关键点;

[0208] 关键点分配模块2555,用于将所述目标关键点分配至每个人体信息中,得到人体姿态信息,完成人体姿态识别;所述人体姿态信息为所述每个人体信息的姿态所构成的集合。

[0209] 进一步地,所述关键点识别模块2553,还用于对所述待识别图像进行多尺度特征提取,得到目标多尺度特征;对所述目标多尺度特征进行所述人体姿态关键点识别,得到所述多尺度关键点。

[0210] 进一步地,所述关键点识别模块2553,还用于利用预设关键点识别模型,对所述待识别图像中的所述人体姿态关键点进行多尺度识别,得到所述多尺度关键点;其中,所述预设关键点识别模型用于从至少两种尺度对所述待识别图像中的所述人体姿态关键点进行识别。

[0211] 进一步地,所述关键点融合模块2554,还用于获取所述多尺度关键点的局部最大值,得到预测多尺度关键点;基于所述多尺度关键点,对所述预测多尺度关键点进行线性组合处理,得到初始多尺度关键点;对所述初始多尺度关键点进行多尺度融合处理,得到所述目标关键点。

[0212] 进一步地,所述关键点融合模块2554,还用于从所述多尺度关键点上,获取预设数量的与所述预测多尺度关键点的距离最小的相邻关键点;

[0213] 对所述预测多尺度关键点和所述相邻关键点进行线性组合,得到所述初始尺度关键点。

[0214] 进一步地,所述初始多尺度关键点包括至少一个多尺度目标关键点,每个多尺度目标关键点由至少两种尺度下每种尺度对应的关键点构成的集合组成;所述关键点融合模块2554,还用于获取所述初始多尺度关键点中所述每个多尺度目标关键点的目标中心点;计算所述每个多尺度目标关键点中每种尺度对应的关键点,到所述目标中心点的距离,得到目标距离;将所述目标距离中小于预设距离阈值的关键点的均值,作为与所述每个多尺度目标关键点对应的子目标关键点,从而得到与所述初始多尺度关键点对应的所述目标关键点。

[0215] 进一步地,所述关键点分配模块2555,还用于从所述目标关键点中,确定属于所述每个人体信息的个体关键点;从所述个体关键点中,获取所述人体姿态关键点中每个关键点类型所对应的个体类型关键点;从所述个体类型关键点中,获取目标个体类型关键点,得

到与所述每个人体信息对应的个体姿态关键点;根据所述个体姿态关键点,得到所述每个人体信息对应的个体姿态,从而得到所述至少一个人体信息对应的所述人体姿态信息。

[0216] 进一步地,所述人体姿态识别装置255还包括训练模块2556,用于获取目标样本图像;所述目标样本图像用于训练得到所述预设关键点识别模型;搭建原始关键点识别模型;所述原始关键点识别模型为多尺度层次结构的模型;利用所述目标样本图像,对所述原始关键点识别模型进行训练,得到所述预设关键点识别模型。

[0217] 进一步地,所述训练模块2556,还用于获取人体部位样本图像和待识别样本图像;对所述人体部位样本图像的人体部位进行识别,得到人体部位识别结果;基于所述人体部位识别结果,从所述人体部位样本图像中提取人体部位,得到人体部位数据库;将所述人体部位数据库中的人体部位与所述待识别样本图像中对应的人体部位进行融合,得到所述目标样本图像。

[0218] 进一步地,所述训练模块2556,还用于搭建原始人体检测模型;所述原始人体检测模型为用于识别出所述待识别图像中的每个人体信息的初始模型;利用所述目标样本图像,训练所述原始人体检测模型,得到人体检测模型;所述人体检测模型用于识别出所述待识别图像中的每个人体信息。

[0219] 相应地,所述人体检测模块2552,还用于利用所述人体检测模型,对所述待识别图像进行人体检测,得到所述至少一个人体信息。

[0220] 进一步地,所述人体姿态识别装置255还包括识别模块2557,用于利用所述预设关键点识别模型,对所述至少一个人体信息的所述每个人体信息的所述人体姿态关键点进行多尺度识别,得到个体姿态关键点,从而得到与所述至少一个人体信息对应的至少一个个体姿态关键点;根据所述至少一个个体姿态关键点,确定所述人体姿态信息,完成人体姿态识别。

[0221] 进一步地,所述人体姿态识别装置255还包括应用模块2558,用于根据所述人体姿态信息,确定个体运动轨迹;所述个体运动轨迹用于确定决策信息。

[0222] 本发明实施例提供一种存储有可执行指令的存储介质,其中存储有可执行指令,当可执行指令被处理器执行时,将引起处理器执行本发明实施例提供的人体姿态识别方法,例如,如图3示出的人体姿态识别方法。

[0223] 在一些实施例中,存储介质可以是FRAM、ROM、PROM、EPROM、EEPROM、闪存、磁表面存储器、光盘、或CD-ROM等存储器;也可以是包括上述存储器之一或任意组合的各种设备。

[0224] 在一些实施例中,可执行指令可以采用程序、软件、软件模块、脚本或代码的形式,按任意形式的编程语言(包括编译或解释语言,或者声明性或过程性语言)来编写,并且其可按任意形式部署,包括被部署为独立的程序或者被部署为模块、组件、子例程或者适合在计算环境中使用的其它单元。

[0225] 作为示例,可执行指令可以但不一定对应于文件系统中的文件,可以可被存储在保存其它程序或数据的文件的一部分,例如,存储在超文本标记语言(HTML,Hyper Text Markup Language)文档中的一个或多个脚本中,存储在专用于所讨论的程序的单个文件中,或者,存储在多个协同文件(例如,存储一个或多个模块、子程序或代码部分的文件)中。

[0226] 作为示例,可执行指令可被部署为在一个计算设备上执行,或者在位于一个地点的多个计算设备上执行,又或者,在分布在多个地点且通过通信网络互连的多个计算设备

上执行。

[0227] 综上所述,通过本发明实施例,由于用于确定待识别图像中每个人体信息的姿态的目标关键点,是由待识别图像的至少两种尺度下的关键点融合得到的;因此,目标关键点的准确度高,从而当根据目标关键点实现人体姿态识别时,识别结果的准确度高,也就提升了人体姿态识别的准确度。

[0228] 以上所述,仅为本发明的实施例而已,并非用于限定本发明的保护范围。凡在本发明的精神和范围之内所作的任何修改、等同替换和改进等,均包含在本发明的保护范围之内。

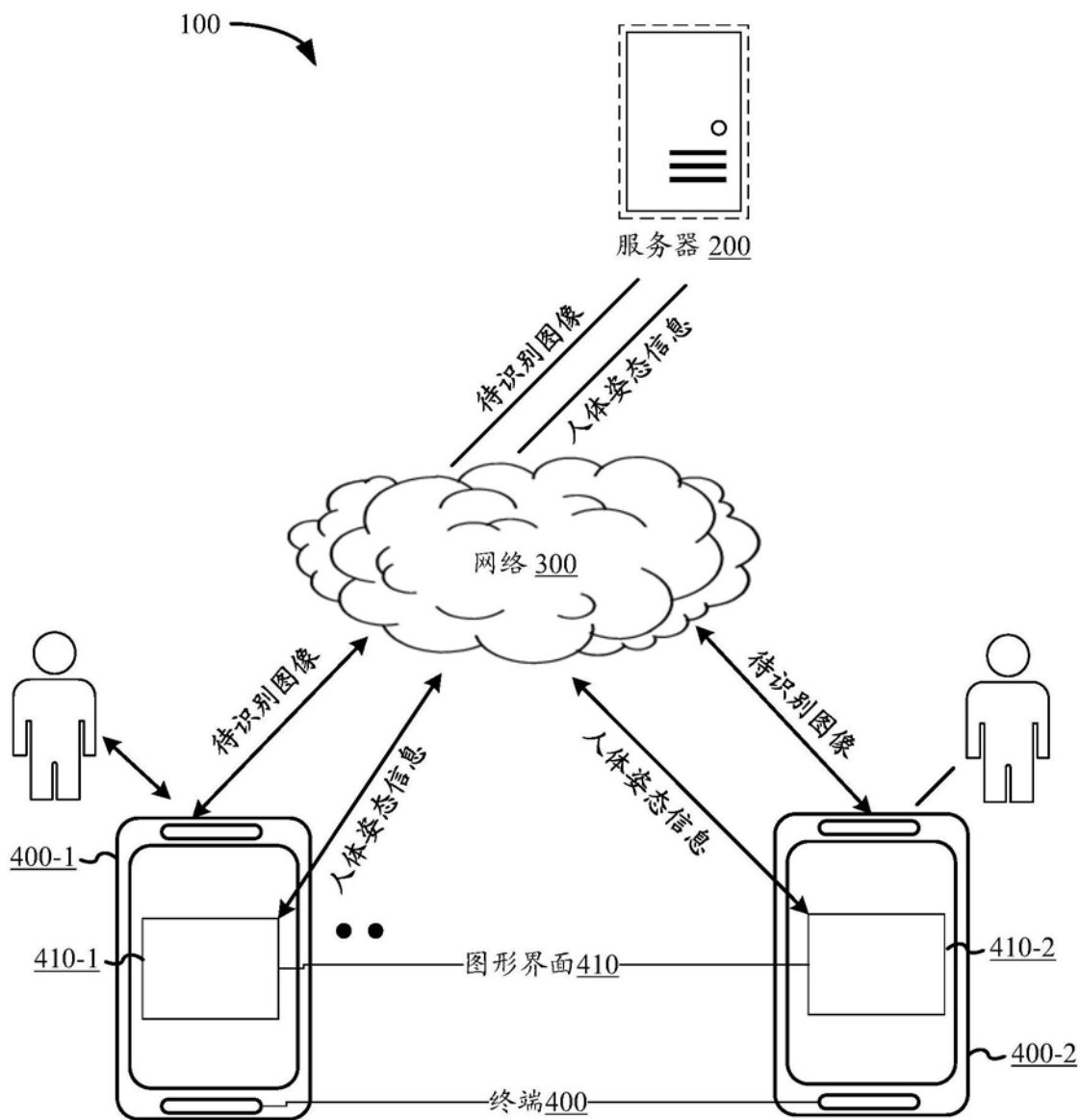


图1

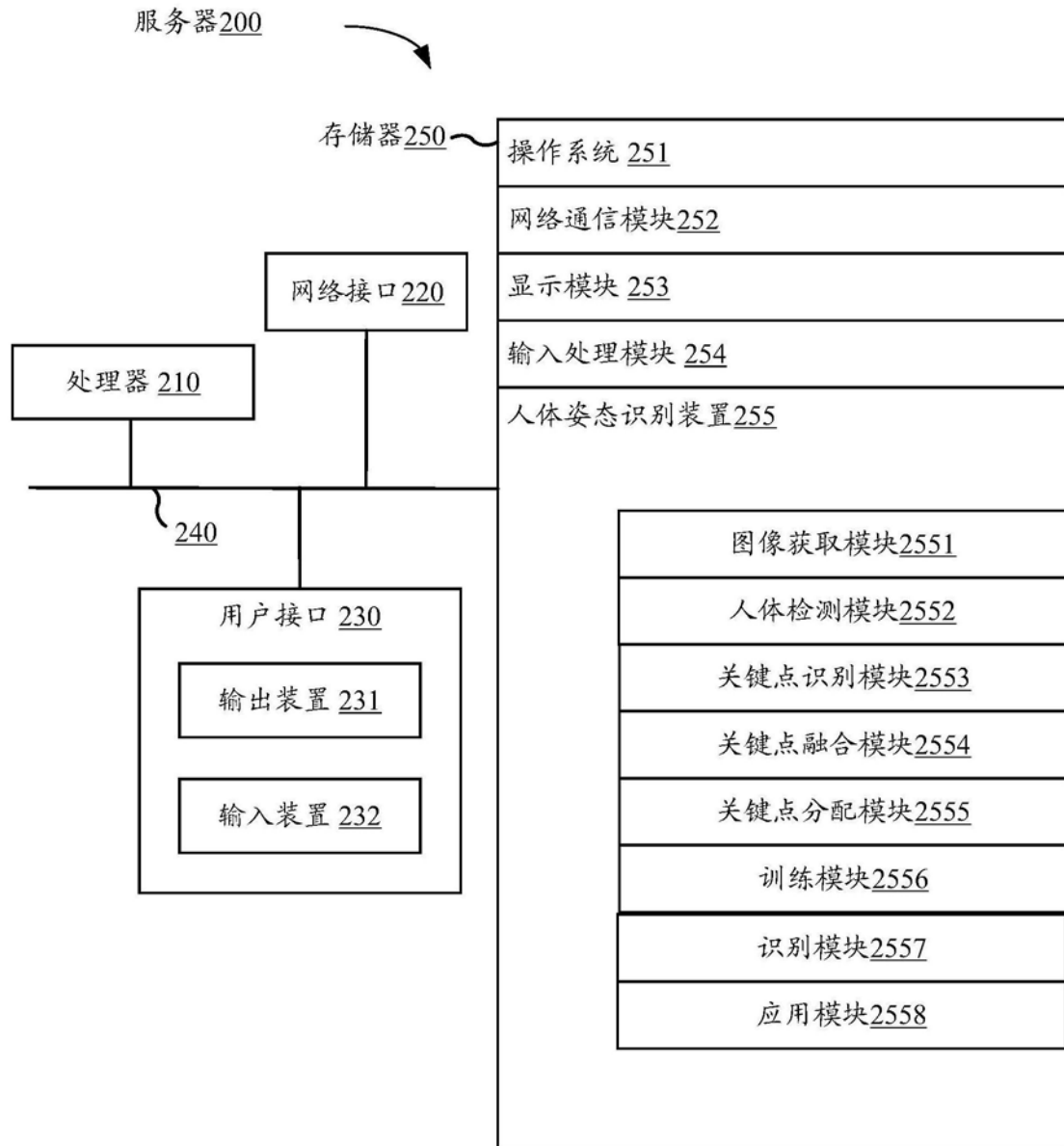


图2

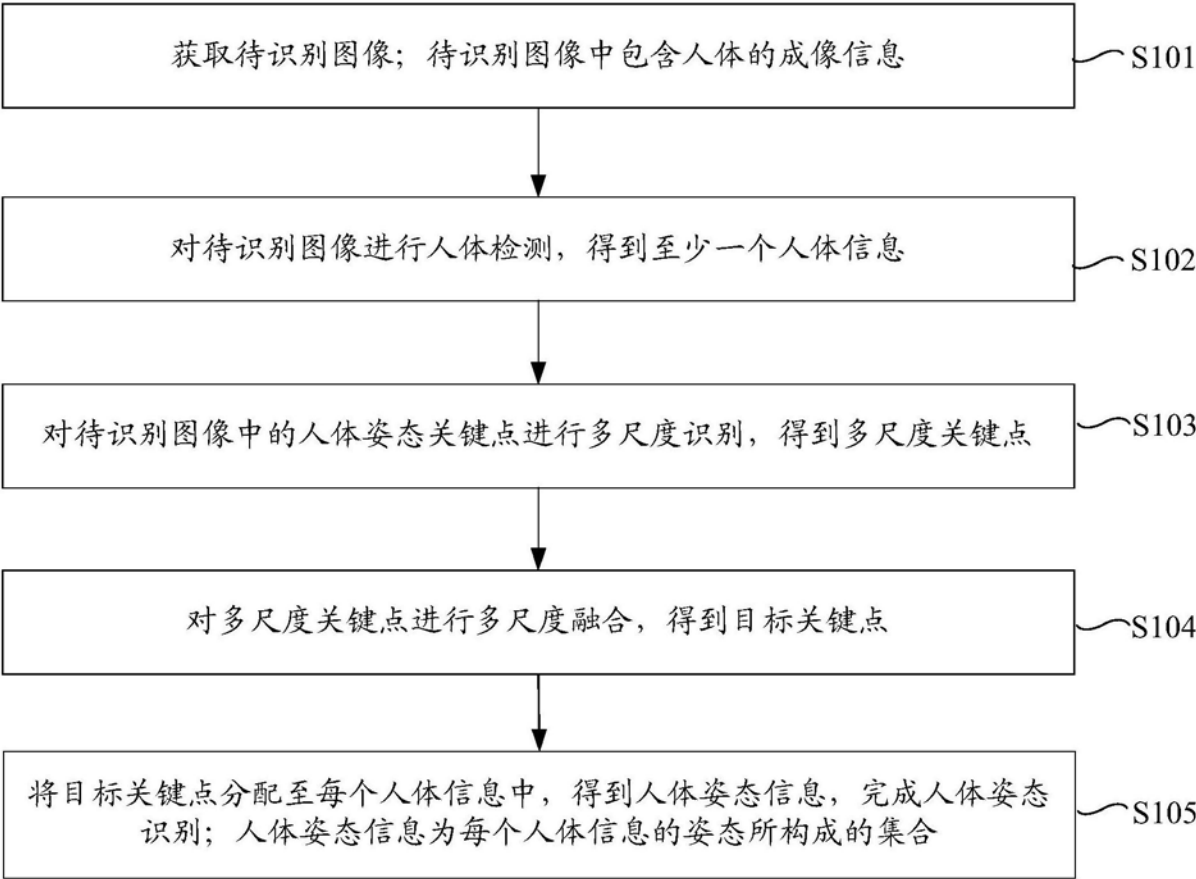


图3

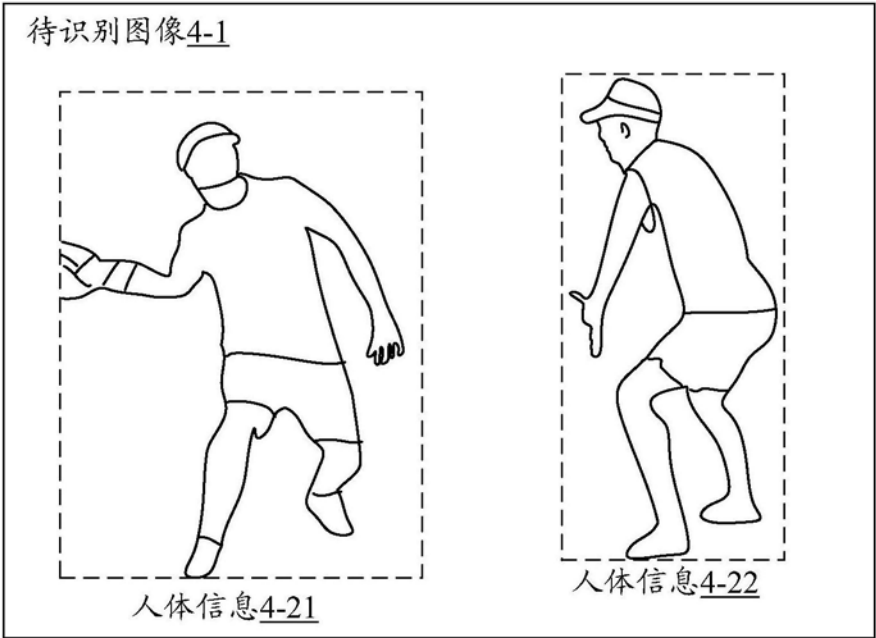


图4

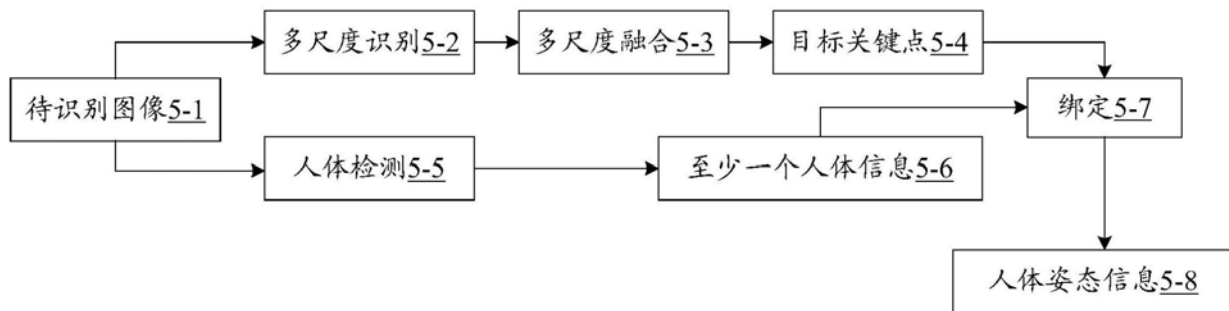


图5

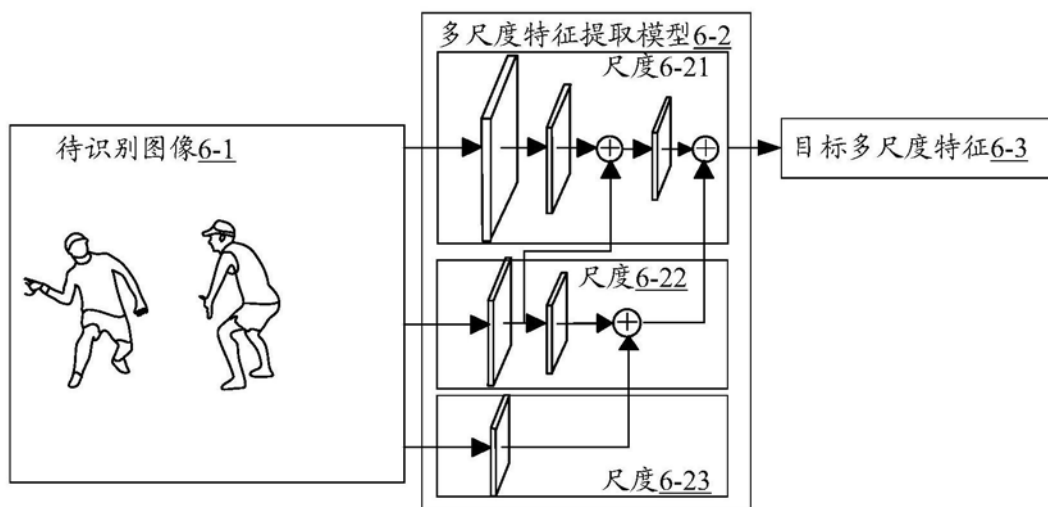


图6

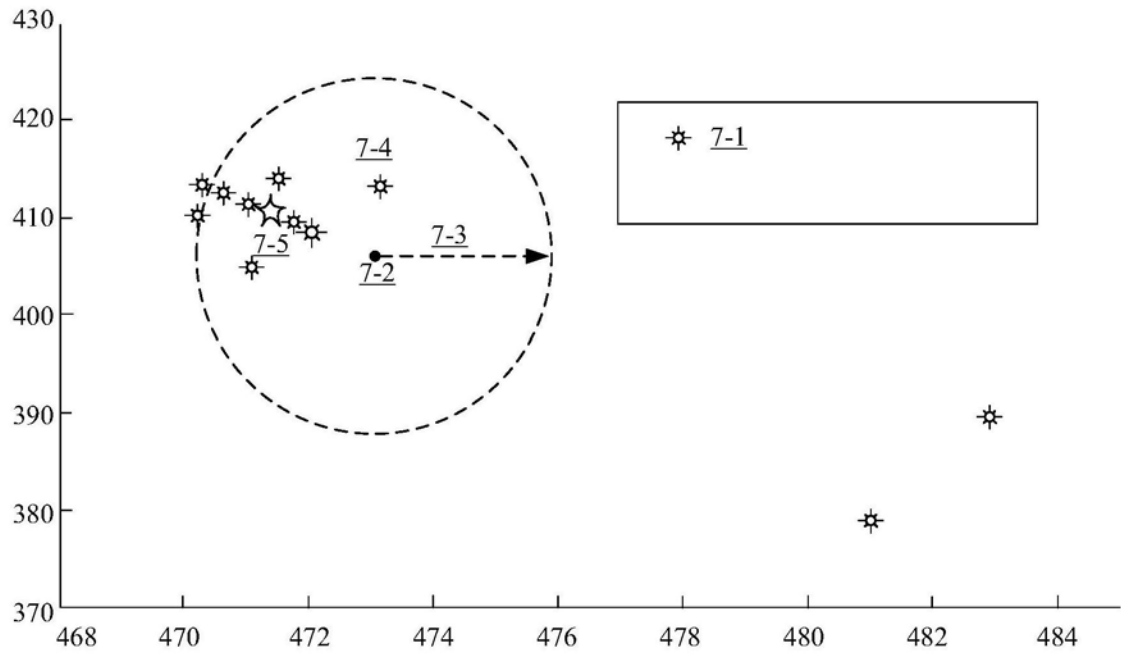


图7

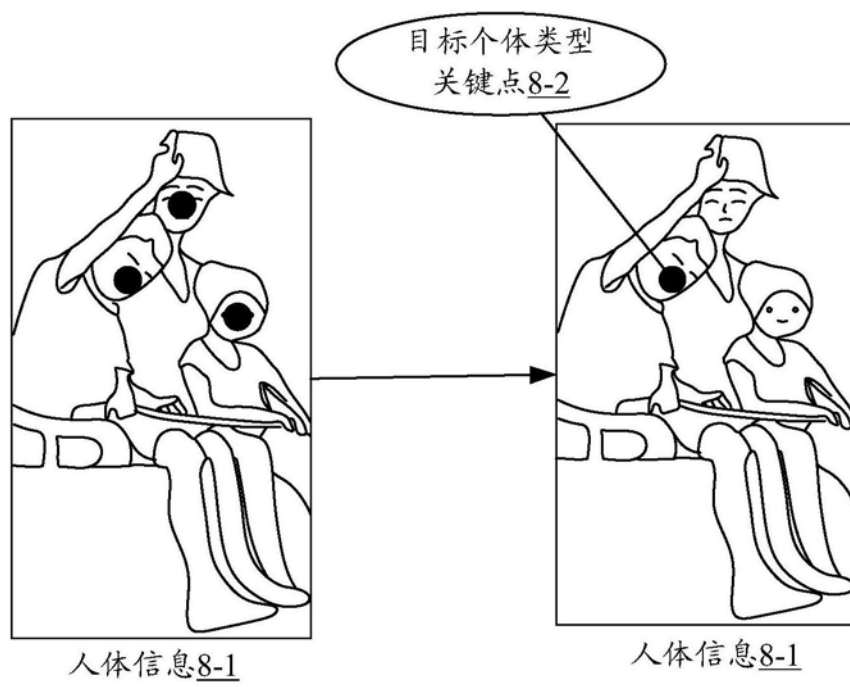


图8

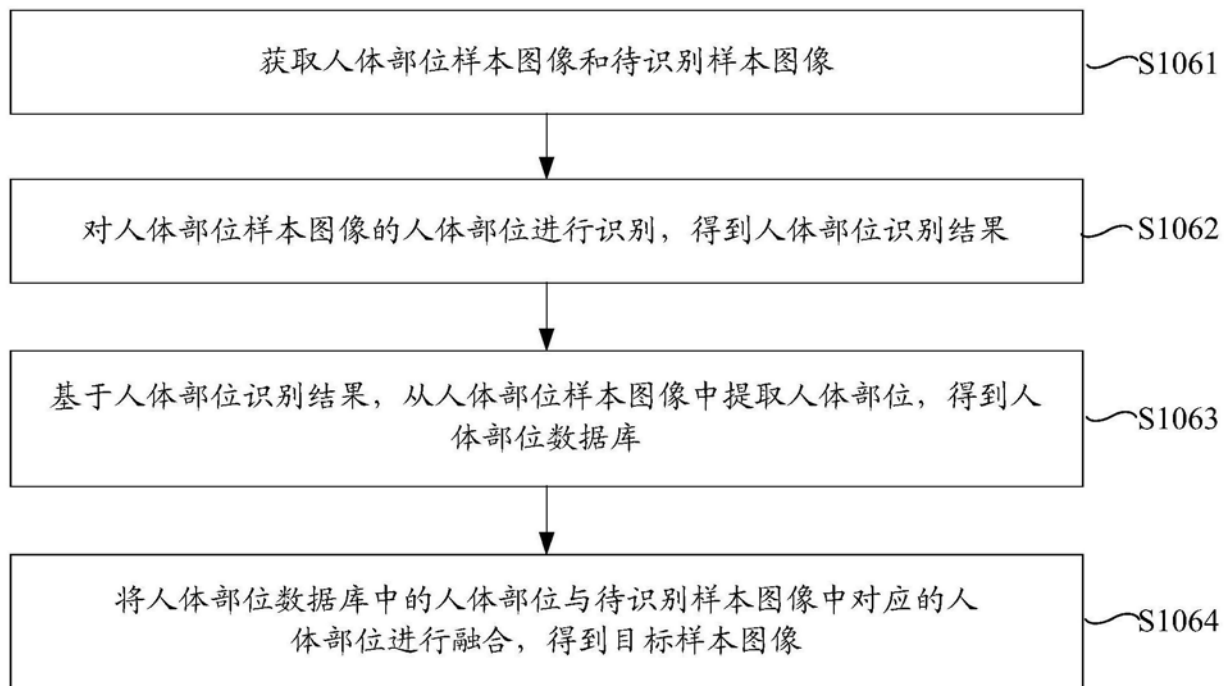
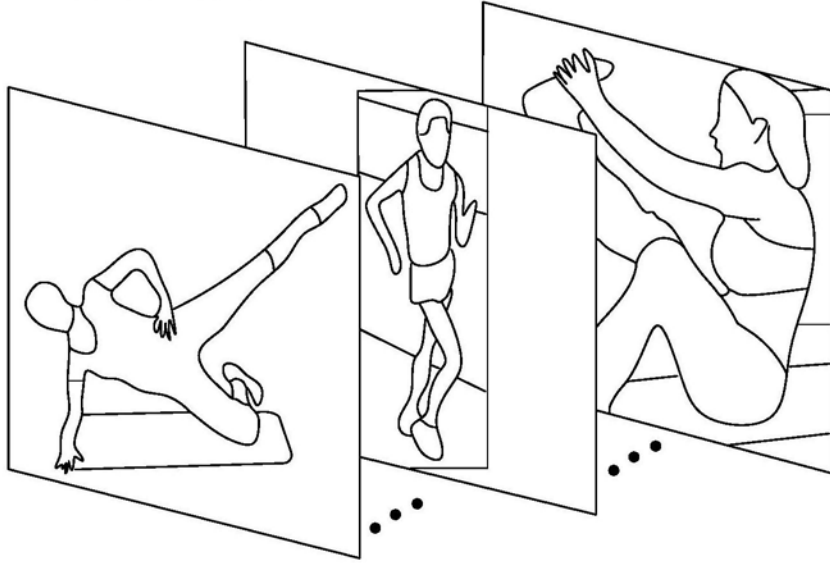


图9

人体部位样本图像10-1



人体部位识别结果10-2

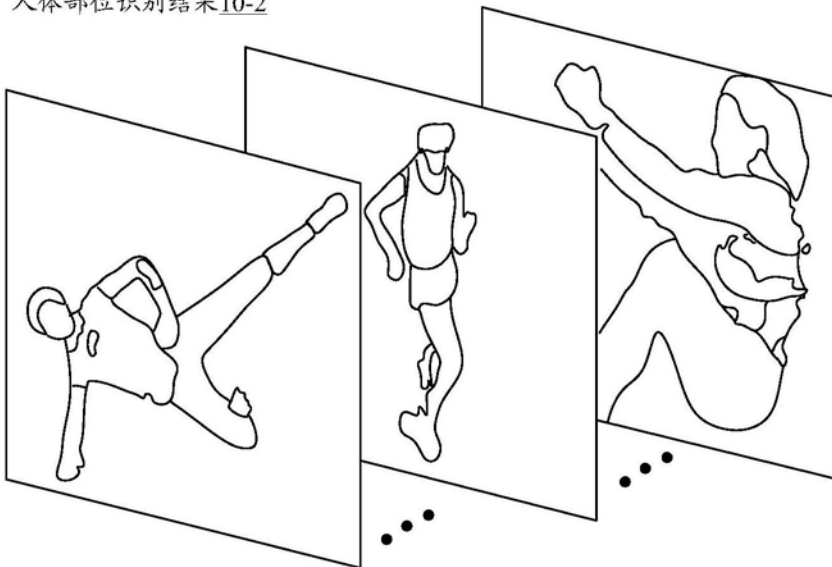


图10

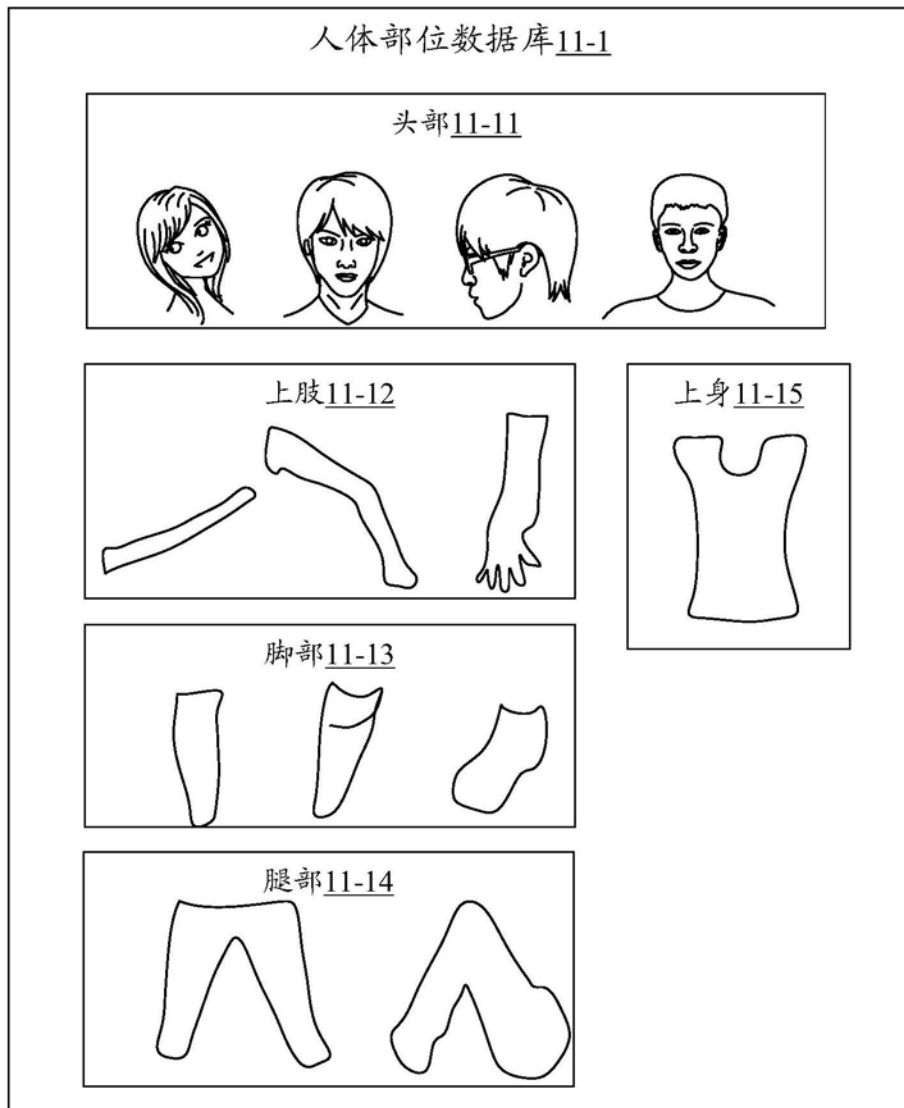


图11

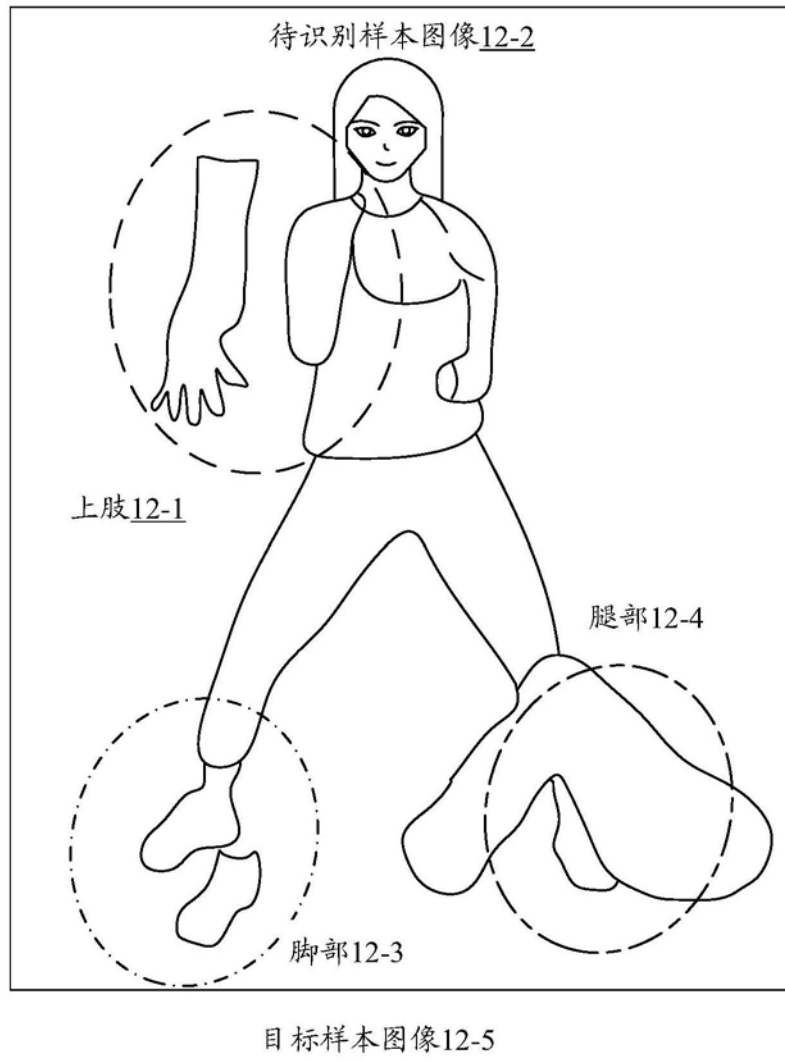


图12

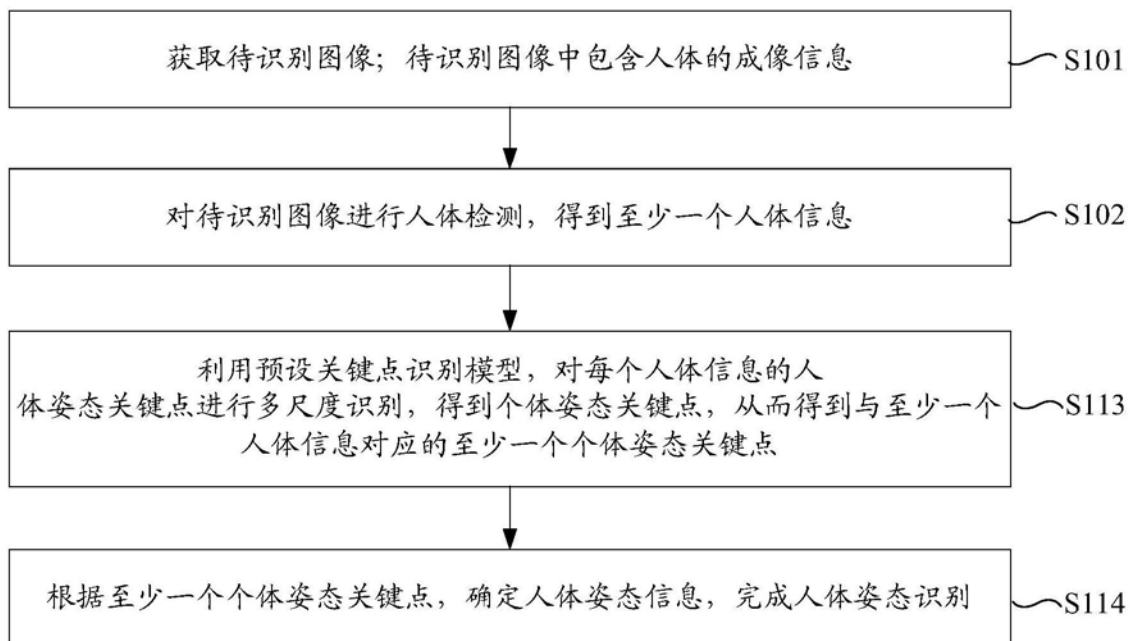


图13

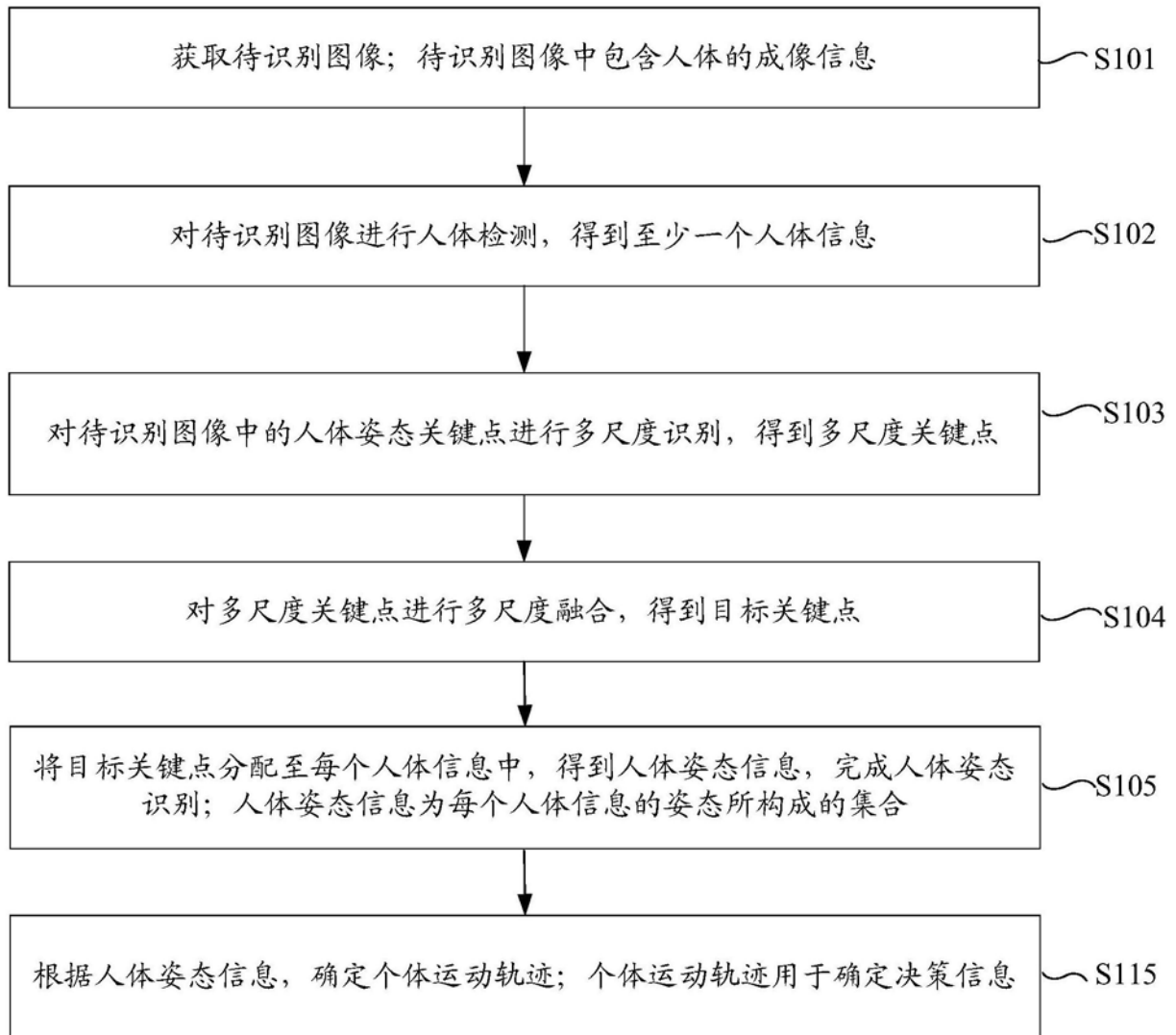


图14

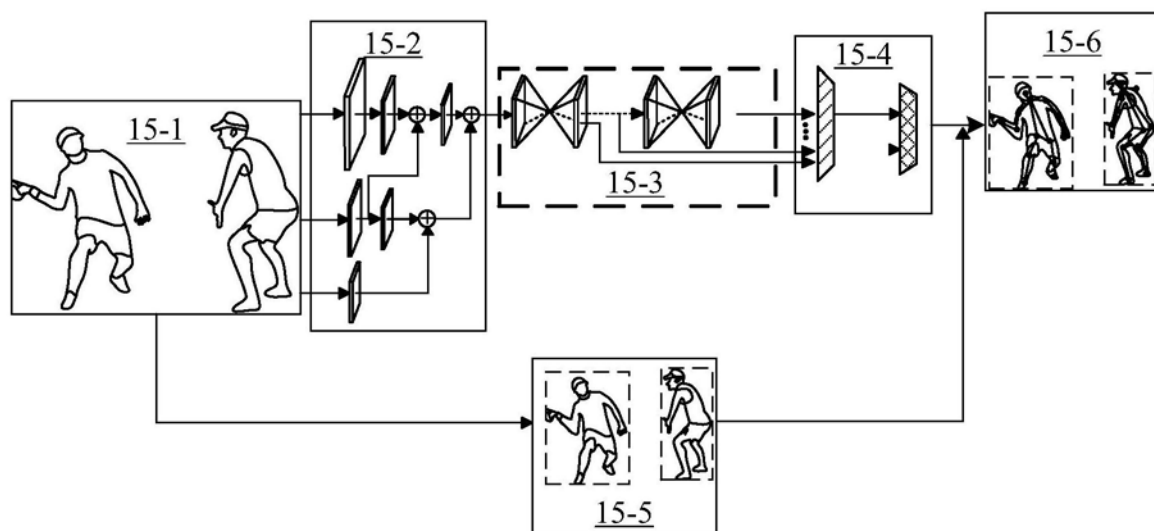


图15

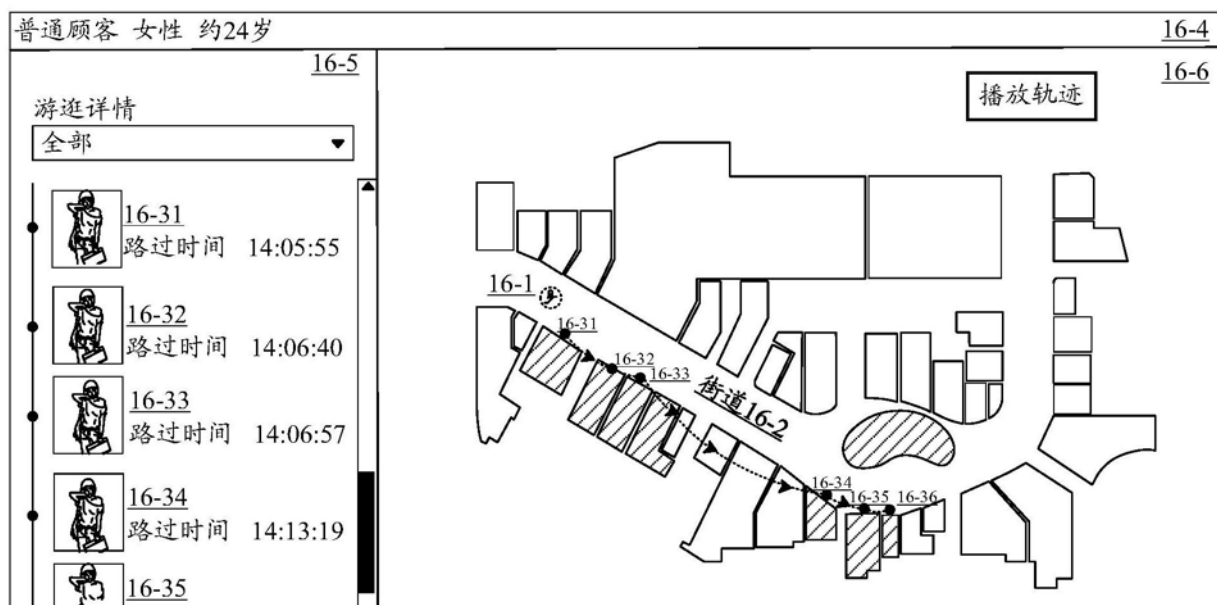


图16

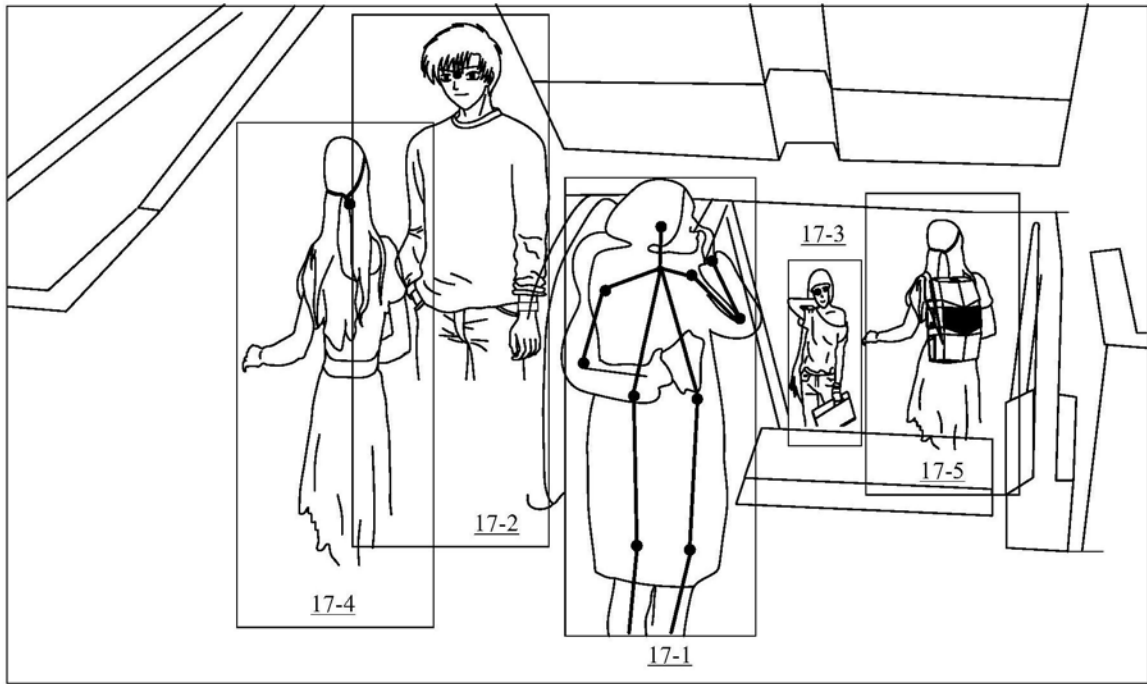


图17