



## (12) 发明专利申请

(10) 申请公布号 CN 113470068 A

(43) 申请公布日 2021. 10. 01

(21) 申请号 202110629411.7

G06N 3/04 (2006.01)

(22) 申请日 2021.06.07

(71) 申请人 北京深睿博联科技有限责任公司

地址 102209 北京市昌平区北七家镇七北路42号院4号楼3层3单元309

申请人 杭州深睿博联科技有限公司

(72) 发明人 吴子丰 冯权洸 哈卿 李一鸣

俞益洲 乔昕

(74) 专利代理机构 北京天方智力知识产权代理

事务所(普通合伙) 11719

代理人 王海龙

(51) Int. Cl.

G06T 7/246 (2017.01)

G06T 5/00 (2006.01)

G06N 3/08 (2006.01)

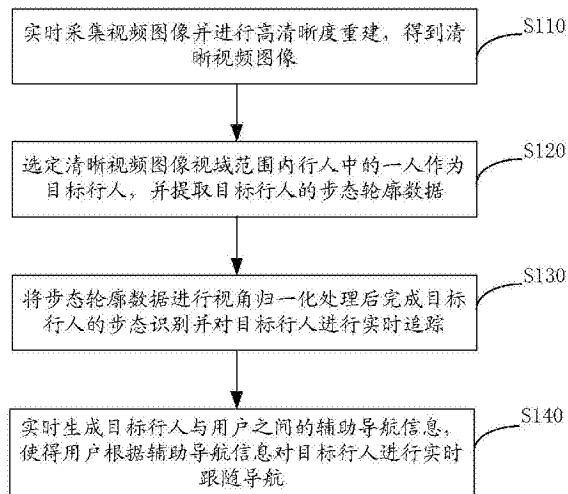
权利要求书2页 说明书10页 附图2页

(54) 发明名称

一种复杂场景中的跟随导航方法和系统

(57) 摘要

本发明提供了一种复杂场景中的跟随导航方法和系统,包括:实时采集视频图像并进行高清晰度重建,得到清晰视频图像;选定清晰视频图像视域范围内行人中的一人作为目标行人,并提取目标行人的步态轮廓数据;将步态轮廓数据进行视角归一化处理后完成目标行人的步态识别并对目标行人进行实时追踪;实时生成目标行人与用户之间的辅助导航信息,使得用户根据辅助导航信息对目标行人进行实时跟随导航。本发明能够实时地自动检测出用户周围环境中的行人,并识别和跟踪特定的目标行人,使具有视觉障碍的用户能够在开放环境中自主地跟随目标行人活动。



1. 一种复杂场景中的跟随导航方法,其特征在于,该方法包括:  
实时采集视频图像并进行高清晰度重建,得到清晰视频图像;  
选定清晰视频图像视域范围内行人中的一人作为目标行人,并提取目标行人的步态轮廓数据;  
将步态轮廓数据进行视角归一化处理后完成目标行人的步态识别并对目标行人进行实时追踪;  
实时生成目标行人与用户之间的辅助导航信息,使得用户根据辅助导航信息对目标行人进行实时跟随导航。
2. 如权利要求1所述的一种复杂场景中的跟随导航方法,其特征在于,所述实时采集视频图像并进行高清晰度重建,得到清晰视频图像包括:  
构建实际场景中拍摄的清晰视频图像数据集和带有噪声的模糊图像数据集;  
根据清晰视频图像数据集和模糊图像数据集对CycleGAN网络进行训练,得到训练好的CycleGAN网络;  
利用训练好的CycleGAN网络对实时采集的视频图像进行高清晰度重建。
3. 如权利要求1所述的一种复杂场景中的跟随导航方法,其特征在于,所述选定清晰视频图像视域范围内行人中的一人作为目标行人,并提取目标行人的步态轮廓数据包括:  
将清晰视频图像的初始帧通过特征金字塔目标检测模型进行实时检测,得到检测结果;  
设定检测结果中的一个行人作为跟随导航的目标行人;  
将清晰视频图像初始帧的目标行人检测框输入至训练好的SiamMask模型中作为视频图像后续帧的跟踪和检索依据对目标行人的步态轮廓进行实时分割和提取,得到步态轮廓数据;  
对步态轮廓数据进行降噪处理,得到完整的步态轮廓数据。
4. 如权利要求3所述的一种复杂场景中的跟随导航方法,其特征在于,还包括:  
对原始步态轮廓数据增加随机的图像掩模,利用图像掩模对原始步态轮廓数据进行点乘操作,得到随机遮挡步态轮廓数据;  
同时将原始步态轮廓数据和随机遮挡步态轮廓数据训练SiamMask模型,得到训练好的SiamMask模型。
5. 如权利要求3所述的一种复杂场景中的跟随导航方法,其特征在于,所述降噪处理包括:  
利用腐蚀及开运算操作去除步态轮廓数据的毛刺噪声;  
利用膨胀及闭运算操作填充步态轮廓数据的孤立像素区域。
6. 如权利要求1所述的一种复杂场景中的跟随导航方法,其特征在于,所述将步态轮廓数据进行视角归一化处理后完成目标行人的步态识别并对目标行人进行实时追踪包括:  
将非正前方视角的步态轮廓数据归一化至正前方步态轮廓数据,视角归一化后的步态轮廓数据;  
将视角归一化后的步态轮廓数据输入至GaitSet模型,利用GaitSet模型对每一帧步态轮廓数据进行提取,通过集合池化运算对所有帧的步态轮廓数据进行融合,通过水平金字塔映射完成特征学习,得到目标行人的步态特征,将步态特征和目标行人进行特征匹配完

成对目标行人的识别和追踪。

7. 如权利要求1所述的一种复杂场景中的跟随导航方法,其特征在于,所述实时生成目标行人与用户之间的辅助导航信息包括:

利用双目立体视觉测量方法以及视差原理生成用户与目标行人的距离与方位。

8. 一种复杂场景中的跟随导航系统,其特征在于,该系统包括:

视频图像重建模块:用于实时采集视频图像并进行高清晰度重建,得到清晰视频图像;

步态轮廓数据提取模块:用于选定清晰视频图像视域范围内行人中的一人作为目标行人,并提取目标行人的步态轮廓数据;

跨视角步态识别模块:用于将步态轮廓数据进行视角归一化处理后完成目标行人的步态识别并对目标行人进行实时追踪;

导航信息生成模块:用于实时生成目标行人与用户之间的辅助导航信息,使得用户根据辅助导航信息对目标行人进行实时跟随导航。

9. 一种电子设备,其特征在于,该设备包括:

处理器、存储器、与网关通信的接口;

存储器用于存储程序和数据,所述处理器调用存储器存储的程序,以执行权利要求1至7任一项所述的一种复杂场景中的跟随导航方法。

10. 一种计算机可读存储介质,其特征在于,所述计算机可读存储介质包括程序,所述程序在被处理器执行时用于执行权利要求1至7任一项所述的一种复杂场景中的跟随导航方法。

## 一种复杂场景中的跟随导航方法和系统

### 技术领域

[0001] 本发明涉及计算机视觉技术领域,具体涉及一种复杂场景中的跟随导航方法和系统。

### 背景技术

[0002] 具有视力功能障碍的人群,在面对复杂、多变的室内及户外场景时,往往无法自行活动,给其生活带来了许多不便。为了提高视力功能障碍人群的行动能力,国内外许多公司设计和生产了可穿戴的智能导航设备(如智能手杖、智能眼镜、语音助手等),可以提供周围环境的感知信息进行辅助导航,进而提高视力功能障碍人群的生活质量。

[0003] 现有产品(如助行眼镜等)的辅助导航方法可以分为两大类,第一类依赖于多传感器集成,通过多模态信号采集的方式进行环境感知,其缺陷是产品成本较高,容易屏蔽掉对价格敏感的潜在用户;第二类方法是基于视觉导航,然而大部分并没有针对复杂场景进行模型的设计和优化,当存在运动模糊、光线变化、物体遮挡时,模型很难对目标行人进行准确跟踪,这将严重影响用户体验。同时当前模型较少考虑目标行人与用户之间视角的动态变化,而复杂场景下的视角变化是经常出现的一种现象,因此现有模型在跨视角的步态识别上错误率较高,在视角变化时不能提供正确的导航信息。

### 发明内容

[0004] 鉴于上述问题,本发明实施例一种复杂场景中的跟随导航方法和系统,解决现有复杂场景识别错误率高导航信息不准的技术问题。

[0005] 为解决上述技术问题,本发明提供以下技术方案:

[0006] 第一方面,本发明提供一种复杂场景中的跟随导航方法,该方法包括:

[0007] 实时采集视频图像并进行高清晰度重建,得到清晰视频图像;

[0008] 选定清晰视频图像视域范围内行人中的一人作为目标行人,并提取目标行人的步态轮廓数据;

[0009] 将步态轮廓数据进行视角归一化处理后完成目标行人的步态识别并对目标行人进行实时追踪;

[0010] 实时生成目标行人与用户之间的辅助导航信息,使得用户根据辅助导航信息对目标行人进行实时跟随导航。

[0011] 在本发明一实施例中,所述实时采集视频图像并进行高清晰度重建,得到清晰视频图像包括:

[0012] 构建实际场景中拍摄的清晰视频图像数据集和带有噪声的模糊图像数据集;

[0013] 根据清晰视频图像数据集和模糊图像数据集对CycleGAN网络进行训练,得到训练好的CycleGAN网络;

[0014] 利用训练好的CycleGAN网络对实时采集的视频图像进行高清晰度重建。

[0015] 在本发明一实施例中,所述选定清晰视频图像视域范围内行人中的一人作为目标

行人,并提取目标行人的步态轮廓数据包括:

[0016] 将清晰视频图像的初始帧通过特征金字塔目标检测模型进行实时检测,得到检测结果;

[0017] 设定检测结果中的一个行人作为跟随导航的目标行人;

[0018] 将清晰视频图像初始帧的目标行人检测框输入至训练好的SiamMask模型中作为视频图像后续帧的跟踪和检索依据对目标行人的步态轮廓进行实时分割和提取,得到步态轮廓数据;

[0019] 对步态轮廓数据进行降噪处理,得到完整的步态轮廓数据。

[0020] 在本发明一实施例中,还包括:

[0021] 对原始步态轮廓数据增加随机的图像掩模,利用图像掩模对原始步态轮廓数据进行点乘操作,得到随机遮挡步态轮廓数据;

[0022] 同时将原始步态轮廓数据和随机遮挡步态轮廓数据训练SiamMask模型,得到训练好的SiamMask模型。

[0023] 在本发明一实施例中,所述降噪处理包括:

[0024] 利用腐蚀及开运算操作去除步态轮廓数据的毛刺噪声;

[0025] 利用膨胀及闭运算操作填充步态轮廓数据的孤立像素区域。

[0026] 在本发明一实施例中,所述将步态轮廓数据进行视角归一化处理后完成目标行人的步态识别并对目标行人进行实时追踪包括:

[0027] 将非正前方视角的步态轮廓数据归一化至正前方步态轮廓数据,视角归一化后的步态轮廓数据;

[0028] 将视角归一化后的步态轮廓数据输入至GaitSet模型,利用GaitSet模型对每一帧步态轮廓数据进行提取,通过集合池化运算对所有帧的步态轮廓数据进行融合,通过水平金字塔映射完成特征学习,得到目标行人的步态特征,将步态特征和目标行人进行特征匹配完成对目标行人的识别和追踪。

[0029] 在本发明一实施例中,所述实时生成目标行人与用户之间的辅助导航信息包括:

[0030] 利用双目立体视觉测量方法以及视差原理生成用户与目标行人的距离与方位。

[0031] 第二方面,本发明提供一种复杂场景中的跟随导航系统,该系统包括:

[0032] 视频图像重建模块:用于实时采集视频图像并进行高清晰度重建,得到清晰视频图像;

[0033] 步态轮廓数据提取模块:用于选定清晰视频图像视域范围内行人中的一人作为目标行人,并提取目标行人的步态轮廓数据;

[0034] 跨视角步态识别模块:用于将步态轮廓数据进行视角归一化处理后完成目标行人的步态识别并对目标行人进行实时追踪;

[0035] 导航信息生成模块:用于实时生成目标行人与用户之间的辅助导航信息,使得用户根据辅助导航信息对目标行人进行实时跟随导航。

[0036] 在本发明一实施例中,所述视频图像重建模块具体用于:

[0037] 构建实际场景中拍摄的清晰视频图像数据集和带有噪声的模糊图像数据集;

[0038] 根据清晰视频图像数据集和模糊图像数据集对CycleGAN网络进行训练,得到训练好的CycleGAN网络;

- [0039] 利用训练好的CycleGAN网络对实时采集的视频图像进行高清晰度重建。
- [0040] 在本发明一实施例中,所述步态轮廓数据提取模块具体用于:
- [0041] 将清晰视频图像的初始帧通过特征金字塔目标检测模型进行实时检测,得到检测结果;
- [0042] 设定检测结果中的一个行人作为跟随导航的目标行人;
- [0043] 将清晰视频图像初始帧的目标行人检测框输入至训练好的SiamMask模型中作为视频图像后续帧的跟踪和检索依据对目标行人的步态轮廓进行实时分割和提取,得到步态轮廓数据;
- [0044] 对步态轮廓数据进行降噪处理,得到完整的步态轮廓数据。
- [0045] 在本发明一实施例中,所述步态轮廓数据提取模块还用于:
- [0046] 对原始步态轮廓数据增加随机的图像掩模,利用图像掩模对原始步态轮廓数据进行点乘操作,得到随机遮挡步态轮廓数据;
- [0047] 同时将原始步态轮廓数据和随机遮挡步态轮廓数据训练SiamMask模型,得到训练好的SiamMask模型。
- [0048] 在本发明一实施例中,所述步态轮廓数据提取模块中的降噪处理包括:
- [0049] 利用腐蚀及开运算操作去除步态轮廓数据的毛刺噪声;
- [0050] 利用膨胀及闭运算操作填充步态轮廓数据的孤立像素区域。
- [0051] 在本发明一实施例中,所述跨视角步态识别模块具体用于:
- [0052] 将非正前方视角的步态轮廓数据归一化至正前方步态轮廓数据,视角归一化后的步态轮廓数据;
- [0053] 将视角归一化后的步态轮廓数据输入至GaitSet模型,利用GaitSet模型对每一帧步态轮廓数据进行提取,通过集合池化运算对所有帧的步态轮廓数据进行融合,通过水平金字塔映射完成特征学习,得到目标行人的步态特征,将步态特征和目标行人进行特征匹配完成对目标行人的识别和追踪。
- [0054] 在本发明一实施例中,所述导航信息生成模块具体用于:
- [0055] 利用双目立体视觉测量方法以及视差原理生成用户与目标行人的距离与方位。
- [0056] 第三方面,本发明提供一种电子设备,该设备包括:
- [0057] 处理器、存储器、与网关通信的接口;
- [0058] 存储器用于存储程序和数据,所述处理器调用存储器存储的程序,以执行第一方面任一项提供的一种复杂场景中的跟随导航方法。
- [0059] 第四方面,本发明提供一种计算机可读存储介质,所述计算机可读存储介质包括程序,所述程序在被处理器执行时用于执行第一方面任一项提供的一种复杂场景中的跟随导航方法。
- [0060] 从上述描述可知,本发明实施例提供一种复杂场景中的跟随导航方法和系统,本发明可以根据在复杂场景中利用实时采集视频图像并进行高清晰度重建,解决复杂场景导致的视频图像退化问题;能够实时地自动检测出用户周围环境中的目标行人,并识别和跟踪特定的目标行人;利用步态轮廓数据视角归一化处理,解决复杂场景导致的视角动态变化问题,对于视角变化有较好的鲁棒性;可以实时生成目标行人与用户之间的辅助导航信息,为视觉功能障碍人群提供高精度的跟随导航信息,从而使具有视觉障碍的用户能够在

开放环境中自主地跟随目标行人进行活动。

## 附图说明

[0061] 图1所示为本发明一种复杂场景中的跟随导航方法的流程示意图；

[0062] 图2所示为本发明一种复杂场景中的跟随导航系统的结构示意图；

[0063] 图3所示为本发明一种电子设备的结构示意图。

## 具体实施方式

[0064] 为使本发明的目的、技术方案及优点更加清楚、明白，以下结合附图及具体实施方式对本发明作进一步说明。显然，所描述的实施例仅仅是本发明一部分实施例，而不是全部的实施例。基于本发明中的实施例，本领域普通技术人员在没有作出创造性劳动前提下所获得的所有其他实施例，都属于本发明保护的范围。

[0065] 首先需要说明的是，本发明的主要应用在佩戴于用户眼部的设备（例如智能眼镜或头盔等）。适用人群为视觉障碍用户，可以理解的是视觉障碍用户并不是指存在完全视觉障碍的盲人，而是存在部分视觉障碍的人，他们无法靠自身双眼进行正常行走，但是他们可以看到佩戴于眼前设备上的辅助导航信息，并跟随辅助导航信息进行跟随导航自主活动。

[0066] 基于现有技术的缺点，本发明实施例提供了一种复杂场景中的跟随导航方法的具体实施方式，如图1所示，该方法具体包括：

[0067] S110:实时采集视频图像并进行高清晰度重建，得到清晰视频图像。

[0068] 具体地，视频图像可以是可见光视频、红外成像视频等，这些视频图像通过佩戴于用户眼前的设备采集。当用户佩戴设备进行室内或室外活动时，常常因运动幅度过大导致设备产生晃动等现象，从而使得采集到的图像数据呈现出边缘细节缺失等运动模糊，这无疑将影响目标行人的识别精度。同时雾、霾、雨、雪等恶劣天气也会降低设备的成像质量。因此需要首先考虑对视频图像进行高清晰度重建，从而提升复杂环境下的成像质量，缓解因运动模糊和恶劣天气造成的图像退化问题。在传统的视频图像恢复研究中，多采用数字滤波算法（如维纳滤波等），需要对滤波器进行设计和优化，其鲁棒性和泛化能力有限。鉴于传统技术存在的问题，本发明通过机器学习中的深度学习训练生成对抗网络。让生成对抗网络学习到清晰图像与模糊图像之间的最优变换，得到内容清晰的视频图像，从而提高复杂环境下的视频成像质量，这样不仅能够为视觉障碍人士提供更加清晰的图像，同时还可以为目标行人的识别提供清晰视频图像。

[0069] S120:选定清晰视频图像视域范围内行人中的一人作为目标行人，并提取目标行人的步态轮廓数据。

[0070] 具体地，经过高清晰度重建得到的清晰视频图像中包含着用户周围的所有行人，可以通过检测模型将这些行人全部检测出来，并将检测出的行人全部作为备选目标行人，由用户自主选择需要跟随的备选目标行人作为最终的目标行人。被选定的目标行人会被检测模型生成的检测框选取，在视频图像的初始帧时选取的目标行人检测框作为后续帧的检索依据，然后将目标行人分割出来，利用一个统一的深度学习框架对目标行人的步态轮廓进行提取，获得目标行人的步态轮廓数据。

[0071] S130:将步态轮廓数据进行视角归一化处理后完成目标行人的步态识别并对目标

行人进行实时追踪。

[0072] 具体地,由于影响行人步态识别精度的因素有很多,其中设备拍摄视角的动态变化是其中非常重要的一个因素。用户在实时导航过程中,由于周围环境复杂多变,其跟随路线并不是简单的直线,这就导致目标行人相对于用户之间的视角在不断发生变化。传统步态识别算法大多只考虑简单场景下的识别问题,对于复杂场景下跨视角的步态识别缺乏针对性的设计。为了解决跨视角识别问题,可用通过构建视角转换模型,将目标行人的步态轮廓数据实时归一化到某一特定视角,从而解决因视角不同导致的识别失败情况的发生。视角转换模型可以由轻量级的生成对抗网络进行实现,从而在保证步态轮廓视角转换的基础上减小计算量,提高模型的实时性。通过对行人步态轮廓特征与目标行人步态轮廓特征进行对比,实现对目标行人的识别与跟踪。

[0073] S140:实时生成目标行人与用户之间的辅助导航信息,使得用户根据辅助导航信息对目标行人进行实时跟随导航。

[0074] 具体地,在经过步态识别后,目标行人即被设备自动识别、定位出来。为了能够为视觉障碍人士提供跟随导航功能,根据识别出来的目标行人检测框的大小、位置等信息,自动计算出目标行人与用户之间的距离和方位,以及目标行人的识别概率,从而为设备跟随导航的可视化提供相关数据。用户根据提供的实时导航信息,即可完成复杂场景下的目标行人跟随。

[0075] 在本实施例中,本实施例可以根据在复杂场景中利用实时采集视频图像并进行高清晰度重建,解决复杂场景导致的视频图像退化问题;能够实时地自动检测出用户周围环境中的目标行人,并识别和跟踪特定的目标行人;利用步态轮廓数据视角归一化处理,解决复杂场景导致的视角动态变化问题,对于视角变化有较好的鲁棒性;可以实时生成目标行人与用户之间的辅助导航信息,为视觉功能障碍人群提供高精度的跟随导航信息,从而使具有视觉障碍的用户能够在开放环境中自主地跟随目标行人进行活动。

[0076] 在上述实施例的基础上,本发明一实施例中,步骤S110具体包括:

[0077] 构建实际场景中拍摄的清晰视频图像数据集和带有噪声的模糊图像数据集;

[0078] 根据清晰视频图像数据集和模糊图像数据集对CycleGAN网络进行训练,得到训练好的CycleGAN网络;

[0079] 利用训练好的CycleGAN网络对实时采集的视频图像进行高清晰度重建。

[0080] 具体地,CycleGAN网络的输入为模糊的图像,输出为去模糊后的清晰图像。为了训练CycleGAN网络,需要构建两个数据集,分别为设备拍摄的清晰图像形成的清晰视频图像数据集和带有噪声的模糊图像数据集。经过训练,CycleGAN网络可以学习到模糊图像到清晰图像之间的最优变换,从而有效提高复杂成像环境下的视频数据质量。此外,相比其他基于深度学习的图像恢复模型对于训练数据的限制(必须同时出现某一场景的模糊图像和清晰图像),CycleGAN网络则无上述要求,从而大大降低了训练数据准备的难度。

[0081] 在本实施例中,利用CycleGAN网络可以更容易的学习到模糊图像与清晰图像之间的最优变换,解决用户因走动、手扶眼睛等动作造成的视频图像运动模糊,以及因雾、霾、雨、雪等恶劣天气条件造成的视频图像退化,可以适应各种复杂的场景,扩大应用场景。

[0082] 在上述实施例的基础上,本发明一实施例中,步骤S120具体包括:

[0083] 将清晰视频图像的初始帧通过特征金字塔目标检测模型进行实时检测,得到检测



结果;

[0084] 具体地,可以理解的是特征金字塔目标检测模型(FPN)设计了自上而下的网络结构和横向连接,以此融合具有高分辨率的浅层特征和具有丰富语义信息的深层特征。先通过下采样的卷积神经网络获得不同尺度的特征。这些顶层特征再通过上采样的网络和低层特征做融合,不同层输出可以得到对不同尺度的敏感的特征,检测不同形态的行人,最后得到的检测结果中包含清晰图像初始帧内所有行人的检测框。

[0085] 设定检测结果中的一个行人作为跟随导航的目标行人;

[0086] 具体地,可以通过以下方法设定目标行人,将头戴设备对准需要跟随的目标行人并保持5秒以上,当停留时间达到阈值时对目标行人进行锁定。被锁定的目标行人的检测框会作为初始检测框送入SiamMask模型中。

[0087] 将清晰视频图像初始帧的目标行人检测框输入至训练好的SiamMask模型中作为视频图像后续帧的跟踪和检索依据对目标行人的步态轮廓进行实时分割和提取,得到步态轮廓数据;

[0088] 具体地,可以理解的是SiamMask模型可以将初始检测框和视频后续帧送入一个双分支的卷积神经网络,分别进行特征提取,并对各自提取的特征进行互相关融合,基于融合后的特征进行多任务学习,分别输出目标行人的步态轮廓数据、目标行人的检测框以及目标行人识别正确的概率。

[0089] 对步态轮廓数据进行降噪处理,得到完整的步态轮廓数据。

[0090] 具体地,因复杂场景下提取到的行人步态轮廓存在噪声污染问题,可以通过数字图像处理领域经典的数学形态学方法对步态轮廓数据进行降噪和标准化处理,对椒盐噪声等进行过滤。利用腐蚀及开运算操作去除步态轮廓数据的毛刺噪声,利用膨胀及闭运算操作填充步态轮廓数据的孤立像素区域。进而使得步态轮廓的完整性更高。

[0091] 在本实施例中,利用特征金字塔目标检测模型可以快速获得SiamMask模型需要的初始帧目标行人检测框,通过SiamMask模型完成目标行人的步态轮廓数据提取,通过将二者有效结合,从而为步态识别提供边缘清晰、时间连续的步态轮廓数据,提高在复杂场景的鲁棒性和适应性。

[0092] 在上述实施例的基础上,本发明一实施例中,步骤S120还包括:

[0093] 对原始步态轮廓数据增加随机的图像掩模,利用图像掩模对原始步态轮廓数据进行点乘操作,得到随机遮挡步态轮廓数据;

[0094] 同时将原始步态轮廓数据和随机遮挡步态轮廓数据训练SiamMask模型,得到训练好的SiamMask模型。

[0095] 具体地,在对SiamMask模型训练时对数据采用随机遮挡的方式进行数据增强。其中,基于随机遮挡方式进行数据增强是指对步态轮廓数据增加随机的图像掩模,利用图像掩模数据对步态轮廓进行点乘操作,模拟出目标行人的步态轮廓被周围物体遮挡的情况,从而增加步态轮廓训练样本的难度和多样性。在SiamMask模型训练过程中,同时使用原始的步态轮廓数据和经过随机遮挡得到随机遮挡步态轮廓数据。

[0096] 在本实施例中,可以有效避免因相互遮挡导致的目标行人步态轮廓不完整的问题,提高步态识别过程中对于未遮挡区域的学习能力,增强模型对于复杂环境的鲁棒性和适应性。

[0097] 在上述实施例的基础上,本发明一实施例中,步骤S130具体包括:

[0098] 将非正前方视角的步态轮廓数据归一化至正前方步态轮廓数据,视角归一化后的步态轮廓数据;

[0099] 具体地,视角的方向是通过目标行人相对于用户的方向定义,正常情况下目标行人是在用户的正前方视角内,当进行变向时目标行人可能跳出用户的正前方视角,通过将非正前方视角的步态轮廓数据归一化至正前方步态轮廓数据,使得在用户视角逐渐转换,目标行人始终出现在用户的正前方视角内。

[0100] 将视角归一化后的步态轮廓数据输入至GaitSet模型,利用GaitSet模型对每一帧步态轮廓数据进行提取,通过集合池化运算对所有帧的步态轮廓数据进行融合,通过水平金字塔映射完成特征学习,得到目标行人的步态特征,将步态特征和目标行人进行特征匹配完成对目标行人的识别和追踪。

[0101] 具体地,特征匹配采用K近邻算法,首先将提取的步态特征在训练特征集中找到与其最邻近的K个实例,若K个实例中的多数属于目标行人,则判定该输入步态特征同属于目标行人,则特征匹配成功并完成了目标行人的识别。

[0102] 在本实施例中,可以实时转换用户的视角,跟随目标行人动态变化,保障目标行人始终出现在用户的正前方视角内。提高在跨视角的步态识别上的正确率,保证在视角变化时也能够提供正确的导航信息,以应对复杂的场景。

[0103] 在上述实施例的基础上,本发明一实施例中,用户视角的转换可以通过视角转换网络实现:

[0104] 视角转换网络采用轻量级的生成对抗网络进行实现,从而在保证步态轮廓视角转换的基础上减小计算量,提高模型的实时性。其中,轻量级的生成对抗网络负责将非正前方视角的步态轮廓转换到正前方视角下,该视角转换网络由生成器和判别器组成,其中生成器由轻量级的全卷积神经网络组成,判别器由轻量级的全连接神经网络组成。视角转换网络的训练需要两组数据:即非正前方视角的步态轮廓数据以及对应的正前方视角数据。在完成视角转换网络训练后,该生成对抗网络的生成器即可实现步态轮廓的视角转换,此时输入一张非正前方的步态轮廓图像,视角转换网络即可输出转换后的正前方的步态轮廓图像。

[0105] 在上述实施例的基础上,本发明一实施例中,步骤S140具体包括:利用双目立体视觉测量方法以及视差原理生成用户与目标行人的距离与方位。

[0106] 具体地,在经过步态识别后,目标行人就会被设备自动识别、定位出来。其中目标行人与用户之间的距离和方位通过经典的双目立体视觉测量得到,此时设备的左右镜片相当于两台摄像机,可以从不同角度同时获取周围景物的两幅数字图像,并基于视差原理即可恢复出目标行人的三维几何信息,通过计算可以得到用户与目标行人的距离与方位。而目标识别的概率可以提示用户是否可以相信设备的识别结果,如果遇到某些极端复杂的场景导致识别概率过低(如小于0.1),用户可以选择放弃目标行人的跟随导航,并进行重置,重新提供目标行人在该场景下的初始检测框,以提高跟随导航的性能。

[0107] 在本实施例中,可以利用识别出来的目标行人检测框的大小、位置等信息,自动计算出目标行人与用户之间的距离和方位,以及目标行人的识别概率,从而为跟随导航的可视化提供相关数据。用户根据提供的实时辅助导航信息,即可完成复杂场景下的目标行人

跟随导航。

[0108] 需要指出的是,本发明的整体实现过程如下:用户在复杂场景下进行行走前,首先通过特征金字塔目标检测模型对目标行人进行检测,从而得到目标行人的检测框,作为后续目标跟踪的初始化数据。在用户行走的过程中,利用CycleGAN网络对设备拍摄的视频图像进行高清晰度重建,从而解决用户因走动、手扶眼镜等动作造成的视频图像运动模糊,以及因雾、霾、雨、雪等恶劣天气条件造成的视频图像退化。接下来,利用SiamMask模型对目标行人的运动轨迹进行实时追踪,并对目标行人的步态轮廓进行自动提取,从而为行人的步态识别提供边缘清晰、时间连续的步态轮廓数据。同时考虑了复杂场景下的行人遮挡问题,利用随机遮挡的方式避免因相互遮挡导致的目标行人步态轮廓不完整的问题。在复杂场景下,目标行人与用户之间的视角存在较大的差异性,将对行人识别结果造成较大误差,因此利用视角转换网络对目标行人的步态轮廓进行归一化处理,并在此基础上利用GaitSet模型完成目标行人的识别。最后将目标行人的检测框、到用户的距离、方位等信息实时显示在设备上,从而为视觉障碍用户提供实时的导航信息,便于其对目标行人进行跟随。

[0109] 基于同一发明构思,本申请实施例还提供了一种复杂场景中的跟随导航系统,可以用于实现上述实施例所描述的一种复杂场景中的跟随导航方法,如下面的实施例所述。由于一种复杂场景中的跟随导航系统解决问题的原理与一种复杂场景中的跟随导航方法相似,因此一种复杂场景中的跟随导航系统的实施可以参见方法实施,重复之处不再赘述。以下所使用的,术语“单元”或者“模块”可以实现预定功能的软件和/或硬件的组合。尽管以下实施例所描述的系统较佳地以软件来实现,但是硬件,或者软件和硬件的组合的实现也是可能并被构想的。

[0110] 本发明提供了一种复杂场景中的跟随导航系统,如图2所示。在图2中,该系统包括:

[0111] 视频图像重建模块210:用于实时采集视频图像并进行高清晰度重建,得到清晰视频图像;

[0112] 步态轮廓数据提取模块220:用于选定清晰视频图像视域范围内行人中的一人作为目标行人,并提取目标行人的步态轮廓数据;

[0113] 跨视角步态识别模块230:用于将步态轮廓数据进行视角归一化处理后完成目标行人的步态识别并对目标行人进行实时追踪;

[0114] 导航信息生成模块240:用于实时生成目标行人与用户之间的辅助导航信息,使得用户根据辅助导航信息对目标行人进行实时跟随导航。

[0115] 在本发明一实施例中,视频图像重建模块210具体用于:

[0116] 构建实际场景中拍摄的清晰视频图像数据集和带有噪声的模糊图像数据集;

[0117] 根据清晰视频图像数据集和模糊图像数据集对CycleGAN网络进行训练,得到训练好的CycleGAN网络;

[0118] 利用训练好的CycleGAN网络对实时采集的视频图像进行高清晰度重建。

[0119] 在本发明一实施例中,步态轮廓数据提取模块220具体用于:

[0120] 将清晰视频图像的初始帧通过特征金字塔目标检测模型进行实时检测,得到检测结果;

[0121] 设定检测结果中的一个行人作为跟随导航的目标行人;

[0122] 将清晰视频图像初始帧的目标行人检测框输入至训练好的SiamMask模型中作为视频图像后续帧的跟踪和检索依据对目标行人的步态轮廓进行实时分割和提取,得到步态轮廓数据;

[0123] 对步态轮廓数据进行降噪处理,得到完整的步态轮廓数据。

[0124] 在本发明一实施例中,步态轮廓数据提取模块220还用于:

[0125] 对原始步态轮廓数据增加随机的图像掩模,利用图像掩模对原始步态轮廓数据进行点乘操作,得到随机遮挡步态轮廓数据;

[0126] 同时将原始步态轮廓数据和随机遮挡步态轮廓数据训练SiamMask模型,得到训练好的SiamMask模型。

[0127] 在本发明一实施例中,步态轮廓数据提取模块220中的降噪处理包括:

[0128] 利用腐蚀及开运算操作去除步态轮廓数据的毛刺噪声;

[0129] 利用膨胀及闭运算操作填充步态轮廓数据的孤立像素区域。

[0130] 在本发明一实施例中,跨视角步态识别模块230具体用于:

[0131] 将非正前方视角的步态轮廓数据归一化至正前方步态轮廓数据,视角归一化后的步态轮廓数据;

[0132] 将视角归一化后的步态轮廓数据输入至GaitSet模型,利用GaitSet模型对每一帧步态轮廓数据进行提取,通过集合池化运算对所有帧的步态轮廓数据进行融合,通过水平金字塔映射完成特征学习,得到目标行人的步态特征,将步态特征和目标行人进行特征匹配完成对目标行人的识别和追踪。

[0133] 在本发明一实施例中,导航信息生成模块240具体用于:

[0134] 利用双目立体视觉测量方法以及视差原理生成用户与目标行人的距离与方位。

[0135] 本申请的实施例还提供能够实现上述实施例中的一种复杂场景中的跟随导航方法中全部步骤的一种电子设备的具体实施方式,参见图3,电子设备300具体包括如下内容:

[0136] 处理器310、存储器320、通信单元330和总线340;

[0137] 其中,处理器310、存储器320、通信单元330通过总线340完成相互间的通信;通信单元330用于实现服务器端设备以及终端设备等相关设备之间的信息传输。

[0138] 处理器310用于调用存储器320中的计算机程序,处理器执行计算机程序时实现上述实施例中的一种复杂场景中的跟随导航方法中的全部步骤。

[0139] 本领域普通技术人员应理解:存储器可以是,但不限于,随机存取存储器(Random Access Memory,简称:RAM),只读存储器(Read Only Memory,简称:ROM),可编程只读存储器(Programmable Read-Only Memory,简称:PROM),可擦除只读存储器(Erasable Programmable Read-Only Memory,简称:EPROM),电可擦除只读存储器(Electric Erasable Programmable Read-Only Memory,简称:EEPROM)等。其中,存储器用于存储程序,处理器在接收到执行指令后,执行程序。进一步地,上述存储器内的软件程序以及模块还可包括操作系统,其可包括各种用于管理系统任务(例如内存管理、存储设备控制、电源管理等)的软件组件和/或驱动,并可与各种硬件或软件组件相互通信,从而提供其他软件组件的运行环境。

[0140] 处理器可以是一种集成电路芯片,具有信号的处理能力。上述的处理器可以是通用处理器,包括中央处理器(Central Processing Unit,简称:CPU)、网络处理器

(NetworkProcessor, 简称:NP) 等。可以实现或者执行本申请实施例中的公开的各方法、步骤及逻辑框图。通用处理器可以是微处理器或者该处理器也可以是任何常规的处理器等。

[0141] 本申请还提供一种计算机可读存储介质, 所述计算机可读存储介质包括程序, 所述程序在被处理器执行时用于执行前述任一方法实施例提供的一种复杂场景中的跟随导航方法。

[0142] 本领域普通技术人员应理解: 实现上述各方法实施例的全部或部分步骤可以通过程序指令相关的硬件来完成。前述的程序可以存储于一计算机可读取存储介质中。该程序在执行时, 执行包括上述各方法实施例的步骤; 而前述的存储介质包括: ROM、RAM、磁碟或者光盘等各种可以存储程序代码的介质, 具体的介质类型本申请不做限制。

[0143] 以上所述, 仅为本发明较佳的具体实施方式, 但本发明的保护范围并不局限于此, 任何熟悉本技术领域的技术人员在本发明揭露的技术范围内, 可轻易想到的变化或替换, 都应涵盖在本发明的保护范围之内。因此, 本发明的保护范围应该以权利要求书的保护范围为准。

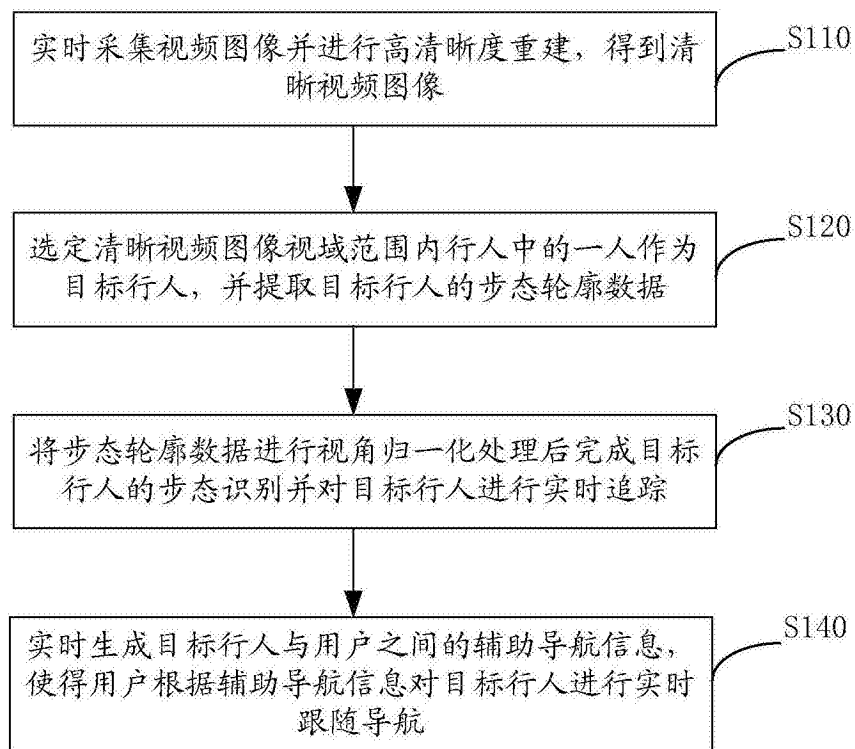


图1

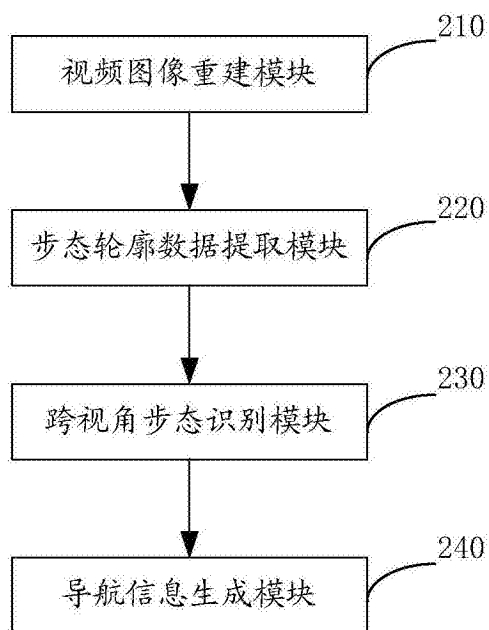


图2

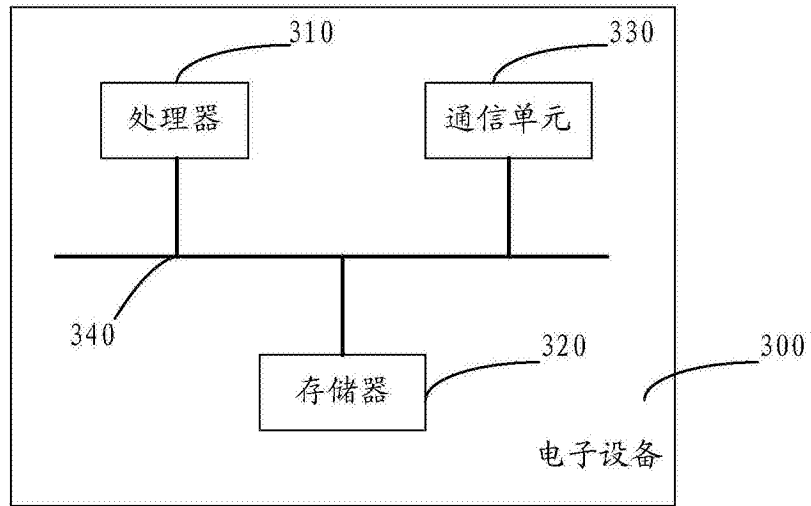


图3