



(12)发明专利申请

(10)申请公布号 CN 112489121 A

(43)申请公布日 2021.03.12

(21)申请号 201910857230.2

(22)申请日 2019.09.11

(71)申请人 丰图科技(深圳)有限公司

地址 518052 广东省深圳市南山区粤海街道深圳市软件产业基地1栋B座4楼

(72)发明人 罗炜 孙玉权

(74)专利代理机构 深圳翼盛智成知识产权事务所(普通合伙) 44300

代理人 彭绪坤

(51) Int. Cl.

G06T 7/73(2017.01)

G06T 7/80(2017.01)

G06T 17/05(2011.01)

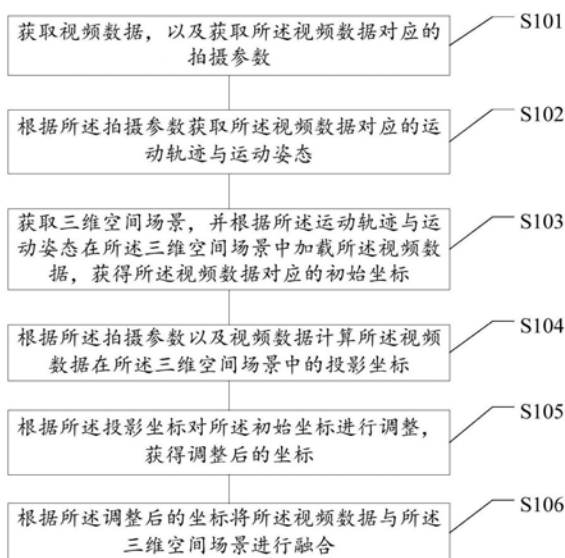
权利要求书2页 说明书13页 附图2页

(54)发明名称

视频融合方法、装置、设备、及存储介质

(57)摘要

本申请实施例公开了一种视频融合方法、装置、设备及计算机可读存储介质,本申请实施例通过获取视频数据,以及获取所述视频数据对应的拍摄参数;根据所述拍摄参数获取所述视频数据对应的运动轨迹与运动姿态;获取三维空间场景,并根据所述运动轨迹与运动姿态在所述三维空间场景中加载所述视频数据,获得所述视频数据对应的初始坐标;根据所述拍摄参数以及视频数据计算所述视频数据在所述三维空间场景中的投影坐标;根据所述投影坐标对所述初始坐标进行调整,获得调整后的坐标;根据所述调整后的坐标将所述视频数据与所述三维空间场景进行融合。实现运动视频与三维场景进行融合,提高了获取视频数据中的具体位置的准确性以及效率。



1. 一种视频融合方法,其特征在于,包括:
  - 获取视频数据,以及获取所述视频数据对应的拍摄参数;
  - 根据所述拍摄参数获取所述视频数据对应的运动轨迹与运动姿态;
  - 获取三维空间场景,并根据所述运动轨迹与运动姿态在所述三维空间场景中加载所述视频数据,获得所述视频数据对应的初始坐标;
  - 根据所述拍摄参数以及视频数据计算所述视频数据在所述三维空间场景中的投影坐标;
  - 根据所述投影坐标对所述初始坐标进行调整,获得调整后的坐标;
  - 根据所述调整后的坐标将所述视频数据与所述三维空间场景进行融合。
2. 根据权利要求1所述的视频融合方法,所述拍摄参数包括角度姿态,其特征在于,获取所述视频数据对应的拍摄参数,包括:
  - 获取初始地理坐标以及初始角度姿态;
  - 获取所述视频数据的序列图像;
  - 提取相邻序列图像的特征点;
  - 根据所述特征点获取相邻序列图像的基本矩阵;
  - 将所述基本矩阵与所述运动轨迹对应的运动矩阵进行融合,获得融合后的矩阵;
  - 通过所述融合后的矩阵对所述初始角度姿态进行调整,获得角度姿态,以及通过所述融合后的矩阵对所述初始地理坐标进行调整,获得调整后的地理坐标。
3. 根据权利要求1所述的视频融合方法,所述拍摄参数还包括相机内部参数与相机外部参数,其特征在于,所述根据所述拍摄参数以及视频数据计算所述视频数据在所述三维空间场景中的投影坐标,包括:
  - 获取所述视频数据的视频帧,以及获取所述视频帧的初始像素坐标;
  - 基于所述视频帧的初始像素坐标对所述视频帧进行畸变矫正,获得视频帧的像素坐标;
  - 基于所述像素坐标与相机内部参数,通过坐标转换获取所述视频数据对应的相机坐标;
  - 根据所述相机外部参数与相机坐标,获取所述视频数据在所述三维空间场景中的投影坐标。
4. 根据权利要求1所述的视频融合方法,所述拍摄参数包括定位时间,其特征在于,所述根据所述调整后的坐标将所述视频数据与所述三维空间场景进行融合之后,还包括:
  - 获取三维渲染帧率、视频数据中的视频帧率;
  - 将所述三维渲染帧率对应的三维空间的渲染时间、视频帧率对应的视频时间与所述定位时间进行匹配,获得匹配后的时间;
  - 根据所述运动轨迹数据、角度姿态与匹配后的时间,移动所述投影坐标;
  - 根据移动后的投影坐标将所述视频数据与所述三维空间场景进行融合。
5. 根据权利要求4所述的视频融合方法,其特征在于,所述将所述三维渲染帧率对应的三维空间的渲染时间、视频帧率对应的视频时间与所述定位时间进行匹配,获得匹配后的时间,包括:
  - 获取三维渲染帧率,将所述三维渲染帧率设置为视频帧率;

根据所述三维渲染帧率获取对应的三维空间的渲染时间；

根据所述视频帧率获取所述视频时间；

根据所述视频时间、三维空间的渲染时间与定位时间将所述视频时间、三维空间的渲染时间与定位时间进行匹配,获得匹配后的视频时间、三维空间的渲染时间与定位时间。

6. 根据权利要求4所述的视频融合方法,其特征在于,所述将所述三维渲染帧率对应的三维空间的渲染时间、视频帧率对应的视频时间与所述定位时间进行匹配,获得匹配后的时间,包括:

将所述三维渲染帧率设置为预设值;

根据所述三维渲染帧率获取对应的三维空间的渲染时间;

根据所述三维空间的渲染时间计算所述视频数据的播放时间;

根据所述三维空间的渲染时间、播放时间与定位时间将所述视频时间、三维空间的渲染时间与定位时间进行匹配,获得匹配后的视频时间、三维空间的渲染时间与定位时间。

7. 根据权利要求1-6任一项所述的视频融合方法,其特征在于,所述根据所述调整后的坐标将所述视频数据与所述三维空间场景进行融合之后,还包括:

接收分屏显示指令;

根据所述分屏显示指令,分别显示所述视频数据与融合后的三维空间场景。

8. 一种视频融合装置,其特征在于,包括:

第一获取单元,用于获取视频数据,以及获取所述视频数据对应的拍摄参数;根据所述拍摄参数获取所述视频数据对应的运动轨迹与运动姿态;获取三维空间场景,并根据所述运动轨迹与运动姿态在所述三维空间场景中加载所述视频数据,获得所述视频数据对应的初始坐标;

计算单元,用于根据所述拍摄参数以及视频数据计算所述视频数据在所述三维空间场景中的投影坐标;

调整单元,用于根据所述投影坐标对所述初始坐标进行调整,获得调整后的坐标;

第一融合单元,用于根据所述调整后的坐标将所述视频数据与所述三维空间场景进行融合。

9. 一种设备,其特征在于,包括处理器和存储器,所述存储器中存储有程序代码,所述处理器调用所述存储器中的程序代码时执行如权利要求1至7任一项所述的视频融合方法。

10. 一种计算机可读存储介质,其特征在于,所述存储介质存储有计算机程序,该程序被处理器加载以执行如权利要求1至7任一项所述的视频融合方法。

## 视频融合方法、装置、设备、及存储介质

### 技术领域

[0001] 本申请涉及数据处理技术领域，具体涉及一种视频融合方法、装置、设备及存储介质。

### 背景技术

[0002] 目前，通常的定位算法研究都是基于不带视频融合的二维或三维空间信息平台，但是在实际应用中不仅需要确定运动目标的坐标，还需要知道运动目标的具体空间进行定位与追踪，而基于不带视频融合的二维或三维空间信息平台，无法确定运动目标的坐标，导致视频和三维空间信息融合偏差大，准确性非常低。

### 发明内容

[0003] 本申请实施例提供一种视频融合方法、装置、设备及存储介质，可以实现确定运动目标的坐标，从而提高融合的准确性与精准性。

[0004] 第一方面，本申请实施例提供了一种视频融合方法，包括：

[0005] 获取视频数据，以及获取所述视频数据对应的拍摄参数；

[0006] 根据所述拍摄参数获取所述视频数据对应的运动轨迹与运动姿态；

[0007] 获取三维空间场景，并根据所述运动轨迹与运动姿态在所述三维空间场景中加载所述视频数据，获得所述视频数据对应的初始坐标；

[0008] 根据所述拍摄参数以及视频数据计算所述视频数据在所述三维空间场景中的投影坐标；

[0009] 根据所述投影坐标对所述初始坐标进行调整，获得调整后的坐标；

[0010] 根据所述调整后的坐标将所述视频数据与所述三维空间场景进行融合。

[0011] 在一些实施方式中，所述拍摄参数包括角度姿态，所述获取所述视频数据对应的拍摄参数，包括：

[0012] 获取初始地理坐标以及初始角度姿态；

[0013] 获取所述视频数据的序列图像；

[0014] 提取相邻序列图像的特征点；

[0015] 根据所述特征点获取相邻序列图像的基本矩阵；

[0016] 将所述基本矩阵与所述运动轨迹对应的运动矩阵进行融合，获得融合后的矩阵；

[0017] 通过所述融合后的矩阵对所述初始角度姿态进行调整，获得角度姿态，以及通过所述融合后的矩阵对所述初始地理坐标进行调整，获得调整后的地理坐标。

[0018] 在一些实施方式中，所述拍摄参数还包括相机内部参数与相机外部参数，所述根据所述拍摄参数以及视频数据计算所述视频数据在所述三维空间场景中的投影坐标，包括：

[0019] 获取所述视频数据的视频帧，以及获取所述视频帧的初始像素坐标；

[0020] 基于所述视频帧的初始像素坐标对所述视频帧进行畸变矫正，获得视频帧的像素

坐标；

[0021] 基于所述像素坐标与相机内部参数,通过坐标转换获取所述视频数据对应的相机坐标；

[0022] 根据所述相机外部参数与相机坐标,获取所述视频数据在所述三维空间场景中的投影坐标。

[0023] 在一些实施方式中,所述拍摄参数包括定位时间,所述根据所述调整后的坐标将所述视频数据与所述三维空间场景进行融合之后,还包括：

[0024] 获取三维渲染帧率、视频数据中的视频帧率；

[0025] 将所述三维渲染帧率对应的三维空间的渲染时间、视频帧率对应的视频时间与所述定位时间进行匹配,获得匹配后的时间；

[0026] 根据所述运动轨迹数据、角度姿态与匹配后的时间,移动所述投影坐标；

[0027] 根据移动后的投影坐标将所述视频数据与所述三维空间场景进行融合。

[0028] 在一些实施方式中,所述将所述三维渲染帧率对应的三维空间的渲染时间、视频帧率对应的视频时间与所述定位时间进行匹配,获得匹配后的时间,包括：

[0029] 获取三维渲染帧率,将所述三维渲染帧率设置为视频帧率；

[0030] 根据所述三维渲染帧率获取对应的三维空间的渲染时间；

[0031] 根据所述视频帧率获取所述视频时间；

[0032] 根据所述视频时间、三维空间的渲染时间与定位时间将所述视频时间、三维空间的渲染时间与定位时间进行匹配,获得匹配后的视频时间、三维空间的渲染时间与定位时间。

[0033] 在一些实施方式中,所述将所述三维渲染帧率对应的三维空间的渲染时间、视频帧率对应的视频时间与所述定位时间进行匹配,获得匹配后的时间,包括：

[0034] 将所述三维渲染帧率设置为预设值；

[0035] 根据所述三维渲染帧率获取对应的三维空间的渲染时间；

[0036] 根据所述三维空间的渲染时间计算所述视频数据的播放时间；

[0037] 根据所述三维空间的渲染时间、播放时间与定位时间将所述视频时间、三维空间的渲染时间与定位时间进行匹配,获得匹配后的视频时间、三维空间的渲染时间与定位时间。

[0038] 在一些实施方式中,所述根据所述调整后的坐标将所述视频数据与所述三维空间场景进行融合之后,还包括：

[0039] 接收分屏显示指令；

[0040] 根据所述分屏显示指令,分别显示所述视频数据与融合后的三维空间场景。

[0041] 第二方面,本申请实施例还提供了一种视频融合装置,包括：

[0042] 第一获取单元,用于获取视频数据,以及获取所述视频数据对应的拍摄参数；根据所述拍摄参数获取所述视频数据对应的运动轨迹与运动姿态；获取三维空间场景,并根据所述运动轨迹与运动姿态在所述三维空间场景中加载所述视频数据,获得所述视频数据对应的初始坐标；

[0043] 计算单元,用于根据所述拍摄参数以及视频数据计算所述视频数据在所述三维空间场景中的投影坐标；

- [0044] 调整单元,用于根据所述投影坐标对所述初始坐标进行调整,获得调整后的坐标;
- [0045] 第一融合单元,用于根据所述调整后的坐标将所述视频数据与所述三维空间场景进行融合。
- [0046] 在一些实施方式中,所述第一获取单元,包括:
- [0047] 第一获取子单元,用于获取初始地理坐标以及初始角度姿态;获取所述视频数据的序列图像;
- [0048] 提取子单元,用于提取相邻序列图像的特征点;
- [0049] 第二获取子单元,用于根据所述特征点获取相邻序列图像的基本矩阵;
- [0050] 融合子单元,用于将所述基本矩阵与所述运动轨迹对应的运动矩阵进行融合,获得融合后的矩阵;
- [0051] 调整子单元,用于通过所述融合后的矩阵对所述初始角度姿态进行调整,获得角度姿态,以及通过所述融合后的矩阵对所述初始地理坐标进行调整,获得调整后的地理坐标。
- [0052] 在一些实施方式中,所述计算单元,包括:
- [0053] 第三获取子单元,用于获取所述视频数据的视频帧,以及获取所述视频帧的初始像素坐标;
- [0054] 矫正子单元,用于基于所述视频帧的初始像素坐标对所述视频帧进行畸变矫正,获得视频帧的像素坐标;
- [0055] 第四获取子单元,用于基于所述像素坐标与相机内部参数,通过坐标转换获取所述视频数据对应的相机坐标;
- [0056] 第五获取子单元,用于根据所述相机外部参数与相机坐标,获取所述视频数据在所述三维空间场景中的投影坐标。
- [0057] 在一些实施方式中,所述视频融合装置,还包括:
- [0058] 第二获取单元,用于获取三维渲染帧率、视频数据中的视频帧率;
- [0059] 匹配单元,用于将所述三维渲染帧率对应的三维空间的渲染时间、视频帧率对应的视频时间与所述定位时间进行匹配,获得匹配后的时间;
- [0060] 移动单元,用于根据所述运动轨迹数据、角度姿态与匹配后的时间,移动所述投影坐标;
- [0061] 第二融合单元,用于根据移动后的投影坐标将所述视频数据与所述三维空间场景进行融合。
- [0062] 在一些实施方式中,所述匹配单元,包括:
- [0063] 第六获取子单元,用于获取三维渲染帧率,将所述三维渲染帧率设置为视频帧率;根据所述三维渲染帧率获取对应的三维空间的渲染时间;根据所述视频帧率获取所述视频时间;
- [0064] 第一匹配子单元,用于根据所述视频时间、三维空间的渲染时间与定位时间将所述视频时间、三维空间的渲染时间与定位时间进行匹配,获得匹配后的视频时间、三维空间的渲染时间与定位时间。
- [0065] 在一些实施方式中,所述匹配单元,包括:
- [0066] 设置子单元,用于将所述三维渲染帧率设置为预设值;

- [0067] 第七获取子单元,用于根据所述三维渲染帧率获取对应的三维空间的渲染时间;
- [0068] 计算子单元,用于根据所述三维空间的渲染时间计算所述视频数据的播放时间;
- [0069] 第二匹配子单元,用于根据所述三维空间的渲染时间、播放时间与定位时间将所述视频时间、三维空间的渲染时间与定位时间进行匹配,获得匹配后的视频时间、三维空间的渲染时间与定位时间。
- [0070] 在一些实施方式中,所述视频融合装置,还包括:
- [0071] 接收单元,用于接收分屏显示指令;
- [0072] 显示单元,用于根据所述分屏显示指令,分别显示所述视频数据与融合后的三维空间场景。
- [0073] 第三方面,本申请实施例还提供了一种设备,所述设备包括处理器和存储器,所述存储器中存储有程序代码,所述处理器调用所述存储器中的程序代码时执行如上所述的视频融合方法。
- [0074] 第四方面,本申请实施例还提供了一种存储介质,所述存储介质存储有计算机程序,该程序被处理器加载以执行如上所述的视频融合方法。
- [0075] 本申请实施例通过获取视频数据,其中,视频数据为运动视频数据,以及获取视频数据对应的拍摄参数;根据拍摄参数获取视频数据对应的运动轨迹与运动姿态;获取三维空间场景,并根据运动轨迹与运动姿态在三维空间场景中加载视频数据,获得视频数据对应的初始坐标,不需要构建三维空间场景,从而提高了加载视频数据的效率;根据拍摄参数以及视频数据计算视频数据在三维空间场景中的投影坐标,从而确定视频数据中运动目标的坐标,根据投影坐标对初始坐标进行调整,获得调整后的坐标;根据调整后的坐标将所述视频数据与所述三维空间场景进行融合,实现将运动视频数据准确地映射到真实的三维场景中,与其他地理空间数据一体化渲染;以及将三维空间场景的地理要素映射到视频数据上,增强视频图像的信息,并且由于提高了加载视频数据的效率,从而提高了视频与三维空间场景的效率。

## 附图说明

- [0076] 为了更清楚地说明本申请实施例中的技术方案,下面将对实施例描述中所需要使用的附图作简单地介绍,显而易见地,下面描述中的附图仅仅是本申请的一些实施例,对于本领域技术人员来讲,在不付出创造性劳动的前提下,还可以根据这些附图获得其他的附图。
- [0077] 图1是本申请实施例提供的视频融合方法的流程示意图;
- [0078] 图2是本申请实施例提供的视频融合装置的结构示意图;
- [0079] 图3是本申请实施例提供的设备的结构示意图。

## 具体实施方式

[0080] 下面将结合本申请实施例中的附图,对本申请实施例中的技术方案进行清楚、完整地描述,显然,所描述的实施例仅仅是本申请一部分实施例,而不是全部的实施例。基于本申请中的实施例,本领域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,都属于本申请保护的范围。

[0081] 请参阅图1,图1是本申请一实施例提供的视频融合方法的流程示意图。该视频融合方法的执行主体可以是本申请实施例提供的视频融合装置,或者集成了该视频融合装置的设备,比如终端或服务器等,该设备可以是安装有摄像头和IMU(Inertial measurement unit,惯性测量模块)的智能手机、平板电脑、掌上电脑、或者笔记本电脑、固定电脑、以及服务器等等。该视频融合方法可以包括:

[0082] S101,获取视频数据,以及获取所述视频数据对应的拍摄参数。

[0083] 具体地,在本实施例中,视频数据可通过安装在交通工具上的拍摄装置拍摄得到,比如通过安装在三轮车、公交车、两轮电动车上的拍摄装置进行拍摄,得到视频数据,或者通过手持拍摄装置拍摄得到,比如通过手机进行拍摄,得到视频数据,具体此处不作限定。然后获取视频数据对应的拍摄参数,其中,视频数据对应的拍摄参数,为视频数据对应的拍摄装置采集得到的参数,包括拍摄装置中的IMU(Inertial measurement unit,惯性测量单元)和GPS(Global Positioning System,全球定位系统)采集得到的定位时间、定位坐标以及角度姿态等。

[0084] 具体地,当拍摄参数包括角度姿态时,获取角度姿态的过程可包括:

[0085] 获取初始地理坐标以及初始角度姿态;

[0086] 获取所述视频数据的序列图像;

[0087] 提取相邻序列图像的特征点;

[0088] 根据所述特征点获取相邻序列图像的基本矩阵;

[0089] 将所述基本矩阵与所述运动轨迹对应的运动矩阵进行融合,获得融合后的矩阵;

[0090] 通过所述融合后的矩阵对所述初始角度姿态进行调整,获得角度姿态,以及通过所述融合后的矩阵对所述初始地理坐标进行调整,获得调整后的地理坐标。

[0091] 获取初始地理坐标以及初始角度姿态,其中,初始角度姿态为通过IMU进行采集得到的角度姿态,初始地理坐标为拍摄模块进行拍摄时的实际定位坐标,然后进一步获取视频数据的序列图像,并提取相邻序列图像的特征点,具体可通过SURF特征提取方法进行特征点提取,根据对极几何原理以及特征点获取相邻序列图像的基本矩阵,将基本矩阵与运动轨迹对应的运动矩阵进行融合,获得融合后的矩阵,其中,运动轨迹对应的运动矩阵为IMU获取的运动矩阵,再通过融合后的矩阵对初始角度姿态进行调整,获得角度姿态,以及通过融合后的矩阵对初始地理坐标进行调整,获得调整后的地理坐标,从而实现地理坐标以及角度姿态的精度修正,获得精度更高的角度姿态,使得后续视频融合的精度更高。

[0092] S102,根据所述拍摄参数获取所述视频数据对应的运动轨迹与运动姿态。

[0093] 具体地,视频数据对应的运动轨迹为视频数据中的运动目标的运动轨迹,根据运动目标每一时刻的定位坐标,获取运动目标在三个方向上的位移、速度以及加速度,其中,三个方向为X轴、Y轴以及Z轴上的方向,从而根据运动目标在三个方向上的位移、速度以及加速度构建运动目标的运动轨迹。运动姿态为IMU采集得到的定位时间、定位坐标以及加速度等。

[0094] S103,获取三维空间场景,并根据所述运动轨迹与运动姿态在所述三维空间场景中加载所述视频数据,获得所述视频数据对应的初始坐标。

[0095] 获取三维空间场景,并加载三维空间场景中的数字地形,航空影像或卫星影像,三维建筑模型,道路模型,矢量点、线、面等,其中,三维空间模型可以为预先建立的三维虚拟



空间场景,具体可根据采集的景象或模型,具体可包括航空影像或卫星影像,三维建筑模型,道路模型,矢量点、线、面等,与三维模型进行融合,得到三维虚拟空间场景。然后根据运动轨迹与运动姿态在三维空间场景中加载视频数据,形成具有初始空间坐标的目标三维空间场景,加载方式可以通过GB28281协议在线实时加载,也可以离线进行加载,在此不做限制。

[0096] S104,根据所述拍摄参数以及视频数据计算所述视频数据在所述三维空间场景中的投影坐标。

[0097] 其中,拍摄参数可包括相机外部参数以及相机内部参数,即具体可根据相机的内部参数与外部参数,以及视频数据中的图像计算视频数据在三维空间场景中的投影坐标。

[0098] 具体地,步骤S104可包括:

[0099] 获取所述视频数据的视频帧,以及获取所述视频帧的初始像素坐标;

[0100] 基于所述视频帧的初始像素坐标对所述视频帧进行畸变矫正,获得视频帧的像素坐标;

[0101] 基于所述像素坐标与相机内部参数,通过坐标转换获取所述视频数据对应的相机坐标;

[0102] 根据所述相机外部参数与相机坐标,获取所述视频数据在所述三维空间场景中的投影坐标。

[0103] 首先获取视频数据中的各个视频帧,具体可通过将视频数据由电信号转化为处于压缩状态的图像信号,进一步对该压缩状态的图像信号进行解压缩,即可获得视频帧。然后以视频帧左上角为原点,以像素为单位建立坐标系,即可获得视频帧中各个像素点的初始像素坐标。由于相机制造工艺偏差,以及入射光线在通过各个透镜时的折射误差和CCD (Charge-coupled Device,电荷耦合器件)点阵位置误差等,实际的光学系统存在着非线性几何失真,从而使目标像点与理论像点之间存在着多种几何畸变,因此需要对视频帧进行畸变矫正,从而获得原始的畸变图,进一步获得原始的畸变图的像素坐标,具体地,获取相机内部参数中的镜头畸变系数以及畸变模型,通过畸变模型与镜头畸变系数对初始像素坐标进行调整,获得像素坐标。然后基于像素坐标与相机内部参数,通过坐标转换获取视频数据对应的相机坐标,具体地,基于相机内部参数,通过对应的转换关系,获取视频数据对应的相机坐标,转换关系具体为: $u-u_0=f_s x/z=f_x x/z$ , $v-v_0=f_s y/z=f_y y/z$ ,其中, $f_s x=f_x$ , $f_s y=f_y$ ,分别定义为X和Y方向的有效焦距, $f_x$ 、 $f_y$ 为相机内部参数, $u_0$ 、 $v_0$ 为像素坐标原点,进一步结合镜头畸变系数(k、s、p),从而获得相机坐标。然后根据相机外部参数与相机坐标,获取视频数据在所述三维空间场景中的投影坐标,具体地,转换关系具体为:

$$\begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} R & T \\ O^T & 1 \end{bmatrix} * \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 0 \end{bmatrix} * \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix}, \text{其中, } T \text{ 是投影坐标系原点在相机坐标系中}$$

的坐标,矩阵R是正交旋转矩阵,R满足约束条件 $r_{11}^2+r_{12}^2+r_{13}^2=1$ , $r_{21}^2+r_{22}^2+r_{23}^2=1$ , $r_{31}^2+r_{32}^2+r_{33}^2=1$ ,R与 $t_x$ , $t_y$ , $t_z$ 为相机外部参数。

[0104] 利用视频帧中的参考点坐标(x,y,z)和图像坐标(u,v)来确定摄像机内部的几何和光学特性(内部参数)以及相机在三维世界中的坐标关系(外部参数)。内部参数包括镜头

畸变系数 ( $k, s, p$ ) 和图像坐标原点 ( $u_0, v_0$ ) 等。外部参数包括摄像机坐标系相对于世界坐标系的正交旋转矩阵  $R$  和平移向量  $t$  等参数。

[0105] S105, 根据所述投影坐标对所述初始坐标进行调整, 获得调整后的坐标。

[0106] 根据计算得到的投影坐标对初始坐标进行调整, 比如将初始坐标的坐标值调整为投影坐标的坐标值, 并相应地移动所述初始坐标的坐标值, 获得调整后的坐标。

[0107] S106, 根据所述调整后的坐标将所述视频数据与所述三维空间场景进行融合。

[0108] 根据调整后的坐标将视频数据与三维空间场景进行场景匹配后进行融合, 从而将视频数据与三维空间场景进行融合。进一步地, 在将视频数据与三维空间场景进行融合之后, 可将融合后的三维空间场景进行展示, 具体展示方式可包括单独展示融合后的三维空间场景, 或者分屏进行展示, 即在屏幕中分别显示视频数据与融合后的三维空间场景, 在分屏展示时, 具体可根据时间进行同步展示, 即显示的视频数据与融合后的三维空间场景之间的具体场景是对应的, 比如假设显示的视频数据的画面为房子, 则显示的融合后的三维空间场景的画面也一样为房子。具体实施过程中, 还可以视频图像动态投影到场景模型表面, 或者视频图像按轨迹在三维场景中移动, 三维场景相机沿摄像头轨迹移动, 并播放视频, 锁定视频场景, 将三维场景反向融合在视频上显示等。

[0109] 进一步地, 由于视频数据可以为实时采集的视频, 因此可以进行实时融合, 具体地, 步骤106之后还包括:

[0110] 获取三维渲染帧率、视频数据中的视频帧率;

[0111] 将所述三维渲染帧率对应的三维空间的渲染时间、视频帧率对应的视频时间与所述定位时间进行匹配, 获得匹配后的时间;

[0112] 根据所述运动轨迹数据、角度姿态与匹配后的时间, 移动所述投影坐标;

[0113] 根据移动后的投影坐标将所述视频数据与所述三维空间场景进行融合。

[0114] 获取三维渲染率, 以及视频数据中的视频帧率, 其中三维渲染率指展示融合后的三维空间场景时, 每秒钟刷新的图片的帧数, 视频帧率即播放视频时, 每秒钟刷新的图片的帧数。将三维渲染率对应的三维空间的渲染时间、视频帧率对应的视频时间与定位时间进行匹配, 获得匹配后的时间; 即将三维空间的渲染时间、视频时间以及定位时间进行同步, 具体可将三维渲染帧率作为基准进行时间匹配, 或者以视频帧率作为基准进行时间匹配。再根据运动轨迹数据、角度姿态与匹配后的时间, 移动投影坐标, 即根据视频的更新, 包括时间更新、运动轨迹更新以及角度姿态的更新, 移动投影坐标, 根据移动后的投影坐标将视频数据与三维空间场景进行融合, 以便实时将视频数据与三维空间场景进行融合, 可实现实时对视频数据中的运动目标进行定位以及追踪。

[0115] 具体地, 当将视频帧率作为基准进行时间匹配时, 将所述三维渲染帧率对应的三维空间的渲染时间、视频帧率对应的视频时间与所述定位时间进行匹配, 获得匹配后的时间的过程包括:

[0116] 获取三维渲染帧率, 将所述三维渲染帧率设置为视频帧率;

[0117] 根据所述三维渲染帧率获取对应的三维空间的渲染时间;

[0118] 根据所述视频帧率获取所述视频时间;

[0119] 根据所述视频时间、三维空间的渲染时间与定位时间将所述视频时间、三维空间的渲染时间与定位时间进行匹配, 获得匹配后的视频时间、三维空间的渲染时间与定位时

间。

[0120] 具体地,获取三维渲染帧率,即获取初始三维渲染帧率,其中,初始三维渲染帧率可以为通用的播放帧率,或者为0,然后将获取到的初始三维渲染帧率设置为视频帧率,即将初始三维渲染帧率设置为视频帧率同样的值,然后通过三维渲染帧率获取对应的三维空间的渲染时间,即每秒播放的图片数量,然后根据视频帧率获取视频时间,同理,视频时间即每秒钟播放的视频图像数量,根据视频时间、三维空间的渲染时间与定位时间将视频时间、三维空间的渲染时间与定位时间进行匹配,获得匹配后的视频时间、三维空间的渲染时间与定位时间,由于将初始三维渲染帧率设置为视频帧率同样的值,因此视频时间、三维空间的渲染时间是一致的,此时,只需要将视频时间与定位时间进行匹配即可,具体地,将开始进行拍摄的定位时间,与视频时间进行匹配,比如假设视频拍摄是从2秒开始,则将定位时间中的2秒与视频时间中的1秒进行匹配,从而实现视频时间、三维空间的渲染时间与定位时间的匹配。

[0121] 具体地,当将三维渲染帧率作为基准进行时间匹配时,将所述三维渲染帧率对应的三维空间的渲染时间、视频帧率对应的视频时间与所述定位时间进行匹配,获得匹配后的时间的过程包括:

[0122] 将所述三维渲染帧率设置为预设值;

[0123] 根据所述三维渲染帧率获取对应的三维空间的渲染时间;

[0124] 根据所述三维空间的渲染时间计算所述视频数据的播放时间;

[0125] 根据所述三维空间的渲染时间、播放时间与定位时间将所述视频时间、三维空间的渲染时间与定位时间进行匹配,获得匹配后的视频时间、三维空间的渲染时间与定位时间。

[0126] 具体地,将三维渲染帧率设置为预设值,具体可根据用户的需求进行设置,比如用户习惯播放的倍速播放为比较高的倍速时,由于倍速播放为比较高的倍速,视频播放的时间更短,此时可将三维渲染帧率设置为较高值,由于是根据设置的三维渲染帧率获取对应的三维空间的渲染时间,根据三维空间的渲染时间计算视频数据的播放时间,从而使得计算的播放时间较短,即与用户设置的三维渲染帧率对应,比如,假设在正常倍速1.0的情况下,视频播放的时间为一小时,而在1.5倍速下,只需要40分钟即可播放完成,因此,在用户习惯播放的倍速播放为比较高的倍速时,需要计算的播放时间比较短,此时,可以将三维渲染帧率设置为较高值,从而使得计算得到的播放时间比较短,与用户的习惯相匹配。具体地,根据接收预设值,并根据接收到的预设值对三维渲染帧率进行设置,然后设置的三维渲染帧率获取对应的三维空间的渲染时间,根据三维空间的渲染时间计算视频数据的播放时间,具体可根据视频数据的大小与三维空间的渲染时间计算视频数据的播放时间,根据三维空间的渲染时间、播放时间与定位时间将视频时间、三维空间的渲染时间与定位时间进行匹配,获得匹配后的视频时间、三维空间的渲染时间与定位时间,具体可参考将视频帧率作为基准进行时间匹配时的匹配过程,在此不再赘述。

[0127] 本实施例通过获取视频数据,其中,视频数据为运动视频数据,以及获取视频数据对应的拍摄参数;根据拍摄参数获取视频数据对应的运动轨迹与运动姿态;获取三维空间场景,并根据运动轨迹与运动姿态在三维空间场景中加载视频数据,获得视频数据对应的初始坐标,不需要构建三维空间场景,从而提高了加载视频数据的效率;根据拍摄参数以及

视频数据计算视频数据在三维空间场景中的投影坐标,从而确定视频数据中运动目标的坐标,根据投影坐标对初始坐标进行调整,获得调整后的坐标;根据调整后的坐标将所述视频数据与所述三维空间场景进行融合,实现将运动视频数据准确地映射到真实的三维场景中,与其他地理空间数据一体化渲染;以及将三维空间场景的地理要素映射到视频数据上,增强视频图像的信息,并且由于提高了加载视频数据的效率,从而提高了视频与三维空间场景的效率。

[0128] 为便于更好的实施本申请实施例提供的视频融合方法,本申请实施例还提供一种基于上述视频融合装置。其中名词的含义与上述视频融合方法中相同,具体实现细节可以参考方法实施例中的说明。

[0129] 请参阅图2,图2为本申请实施例提供的视频融合装置的结构示意图,其中该物视频融合装置可以包括第一获取单元201、计算单元202、调整单元203以及第一融合单元204等。

[0130] 具体地,视频融合装置包括:

[0131] 第一获取单元201,用于获取视频数据,以及获取所述视频数据对应的拍摄参数;根据所述拍摄参数获取所述视频数据对应的运动轨迹与运动姿态;获取三维空间场景,并根据所述运动轨迹与运动姿态在所述三维空间场景中加载所述视频数据,获得所述视频数据对应的初始坐标;

[0132] 计算单元202,用于根据所述拍摄参数以及视频数据计算所述视频数据在所述三维空间场景中的投影坐标;

[0133] 调整单元203,用于根据所述投影坐标对所述初始坐标进行调整,获得调整后的坐标;

[0134] 第一融合单元204,用于根据所述调整后的坐标将所述视频数据与所述三维空间场景进行融合。

[0135] 在一些实施方式中,所述第一获取单元201,包括:

[0136] 第一获取子单元,用于获取初始地理坐标以及初始角度姿态;

[0137] 提取子单元,用于提取相邻序列图像的特征点;

[0138] 第二获取子单元,用于根据所述特征点获取相邻序列图像的基本矩阵;

[0139] 融合子单元,用于将所述基本矩阵与所述运动轨迹对应的运动矩阵进行融合,获得融合后的矩阵;

[0140] 调整子单元,用于通过所述融合后的矩阵对所述初始角度姿态进行调整,获得角度姿态,以及通过所述融合后的矩阵对所述初始地理坐标进行调整,获得调整后的地理坐标。

[0141] 在一些实施方式中,所述计算单元202,包括:

[0142] 第三获取子单元,用于获取所述视频数据的视频帧,以及获取所述视频帧的初始像素坐标;

[0143] 矫正子单元,用于基于所述视频帧的初始像素坐标对所述视频帧进行畸变矫正,获得视频帧的像素坐标;

[0144] 第四获取子单元,用于基于所述像素坐标与相机内部参数,通过坐标转换获取所述视频数据对应的相机坐标;

[0145] 第五获取子单元,用于根据所述相机外部参数与相机坐标,获取所述视频数据在所述三维空间场景中的投影坐标。

[0146] 在一些实施方式中,所述视频融合装置,还包括:

[0147] 第二获取单元,用于获取三维渲染帧率、视频数据中的视频帧率;

[0148] 匹配单元,用于将所述三维渲染帧率对应的三维空间的渲染时间、视频帧率对应的视频时间与所述定位时间进行匹配,获得匹配后的时间;

[0149] 移动单元,用于根据所述运动轨迹数据、角度姿态与匹配后的时间,移动所述投影坐标;

[0150] 第二融合单元,用于根据移动后的投影坐标将所述视频数据与所述三维空间场景进行融合。

[0151] 在一些实施方式中,所述匹配单元,包括:

[0152] 第六获取子单元,用于获取三维渲染帧率,将所述三维渲染帧率设置为视频帧率;根据所述三维渲染帧率获取对应的三维空间的渲染时间;根据所述视频帧率获取所述视频时间;

[0153] 第一匹配子单元,用于根据所述视频时间、三维空间的渲染时间与定位时间将所述视频时间、三维空间的渲染时间与定位时间进行匹配,获得匹配后的视频时间、三维空间的渲染时间与定位时间。

[0154] 在一些实施方式中,所述匹配单元,包括:

[0155] 设置子单元,用于将所述三维渲染帧率设置为预设值;

[0156] 第七获取子单元,用于根据所述三维渲染帧率获取对应的三维空间的渲染时间;

[0157] 计算子单元,用于根据所述三维空间的渲染时间计算所述视频数据的播放时间;

[0158] 第二匹配子单元,用于根据所述三维空间的渲染时间、播放时间与定位时间将所述视频时间、三维空间的渲染时间与定位时间进行匹配,获得匹配后的视频时间、三维空间的渲染时间与定位时间。

[0159] 在一些实施方式中,所述视频融合装置,还包括:

[0160] 接收单元,用于接收分屏显示指令;

[0161] 显示单元,用于根据所述分屏显示指令,分别显示所述视频数据与融合后的三维空间场景。

[0162] 其中,以上各个操作的具体实施可参见前面的第一实施例,在此不再赘述。

[0163] 图3示出了本发明实施例提供的设备的具体结构框图,该设备为视频融合设备,具体用于实施上述实施例中提供的视频融合方法。该设备400可以为智能手机或平板电脑等终端,或者是服务器。

[0164] 如图3所示,设备400可以包括RF (Radio Frequency, 射频) 电路110、包括有一个或一个以上(图中仅示出一个) 计算机可读存储介质的存储器120、输入单元130、显示单元140、传输模块170、包括有一个或者一个以上(图中仅示出一个) 处理核心的处理器180以及电源190等部件。本领域技术人员可以理解,图3中示出的设备400结构并不构成对设备400的限定,可以包括比图示更多或更少的部件,或者组合某些部件,或者不同的部件布置。其中:

[0165] RF电路110用于接收以及发送电磁波,实现电磁波与电信号的相互转换,从而与通

讯网络或者其他设备进行通讯。RF电路110可包括各种现有的用于执行这些功能的电路元件,例如,天线、射频收发器、数字信号处理器、加密/解密芯片、用户身份模块(SIM)卡、存储器等等。RF电路110可与各种网络如互联网、企业内部网、无线网络进行通讯或者通过无线网络与其他设备进行通讯。上述的无线网络可包括蜂窝式电话网、无线局域网或者城域网。上述的无线网络可以使用各种通信标准、协议及技术,包括但不限于全球移动通信系统(Global System for Mobile Communication,GSM)、增强型移动通信技术(Enhanced Data GSM Environment,EDGE)、宽带码分多址技术(Wideband Code Division Multiple Access,WCDMA)、码分多址技术(Code Division Access,CDMA)、时分多址技术(Time Division Multiple Access,TDMA)、无线保真技术(Wireless Fidelity,Wi-Fi)(如美国电气和电子工程师协会标准IEEE 802.11a,IEEE 802.11b,IEEE802.11g和/或IEEE 802.11n)、网络电话(Voice over Internet Protocol,VoIP)、全球微波互联接入(Worldwide Interoperability for Microwave Access,Wi-Max)、其他用于邮件、即时通讯及短消息的协议,以及任何其他合适的通讯协议,甚至可包括那些当前仍未被开发出来的协议。

[0166] 存储器120可用于存储软件程序以及模块,如上述实施例中视频融合方法的程序指令/模块,处理器180通过运行存储在存储器120内的软件程序以及模块,从而执行各种功能应用以及数据处理,即实现物体体积计算的功能。存储器120可包括高速随机存储器,还可包括非易失性存储器,如一个或者多个磁性存储装置、闪存、或者其他非易失性固态存储器。在一些实例中,存储器120可进一步包括相对于处理器180远程设置的存储器,这些远程存储器可以通过网络连接至设备400。上述网络的实例包括但不限于互联网、企业内部网、局域网、移动通信网及其组合。

[0167] 输入单元130可用于接收输入的数字或字符信息,以及产生与用户设置以及功能控制有关的键盘、鼠标、操作杆、光学或者轨迹球信号输入。具体地,输入单元130可包括触敏表面131以及其他输入设备132。触敏表面131,也称为触摸显示屏或者触控板,可收集用户在其上或附近的触摸操作(比如用户使用手指、触笔等任何适合的物体或附件在触敏表面131上或在触敏表面131附近的操作),并根据预先设定的程式驱动相应的连接装置。可选的,触敏表面131可包括触摸检测装置和触摸控制器两个部分。其中,触摸检测装置检测用户的触摸方位,并检测触摸操作带来的信号,将信号传送给触摸控制器;触摸控制器从触摸检测装置上接收触摸信息,并将它转换成触点坐标,再送给处理器180,并能接收处理器180发来的命令并加以执行。此外,可以采用电阻式、电容式、红外线以及表面声波等多种类型实现触敏表面131。除了触敏表面131,输入单元130还可以包括其他输入设备132。具体地,其他输入设备132可以包括但不限于物理键盘、功能键(比如音量控制按键、开关按键等)、轨迹球、鼠标、操作杆等中的一种或多种。

[0168] 显示单元140可用于显示由用户输入的信息或提供给用户的信息以及设备400的各种图形用户接口,这些图形用户接口可以由图形、文本、图标、视频和其任意组合来构成。显示单元140可包括显示面板141,可选的,可以采用LCD(Liquid Crystal Display,液晶显示器)、OLED(Organic Light-Emitting Diode,有机发光二极管)等形式来配置显示面板141。进一步的,触敏表面131可覆盖显示面板141,当触敏表面131检测到在其上或附近的触摸操作后,传送给处理器180以确定触摸事件的类型,随后处理器180根据触摸事件的类型

在显示面板141上提供相应的视觉输出。虽然在图3中,触敏表面131与显示面板141是作为两个独立的部件来实现输入和输出功能,但是在某些实施例中,可以将触敏表面131与显示面板141集成而实现输入和输出功能。

[0169] 设备400通过传输模块170(例如Wi-Fi模块)可以帮助用户收发电子邮件、浏览网页和访问流式媒体等,它为用户提供了无线的宽带互联网访问。虽然图3示出了传输模块170,但是可以理解的是,其并不属于设备400的必须构成,完全可以根据需要在不改变发明的本质的范围内而省略。

[0170] 处理器180是设备400的控制中心,利用各种接口和线路连接整个手机的各个部分,通过运行或执行存储在存储器120内的软件程序和/或模块,以及调用存储在存储器120内的数据,执行设备400的各种功能和处理数据,从而对手机进行整体监控。可选的,处理器180可包括一个或多个处理核心;在一些实施例中,处理器180可集成应用处理器和调制解调处理器,其中,应用处理器主要处理操作系统、用户界面和应用程序等,调制解调处理器主要处理无线通信。可以理解的是,上述调制解调处理器也可以不集成到处理器180中。

[0171] 设备400还包括给各个部件供电的电源190(比如电池),在一些实施例中,电源可以通过电源管理系统与处理器180逻辑相连,从而通过电源管理系统实现管理充电、放电、以及功耗管理等功能。电源190还可以包括一个或一个以上的直流或交流电源、再充电系统、电源故障检测电路、电源转换器或者逆变器、电源状态指示器等任意组件。

[0172] 具体在本实施例中,设备400的显示单元140是触摸屏显示器,设备400还包括有存储器120,以及一个或者一个以上的程序,其中一个或者一个以上程序存储于存储器120中,且经配置以由一个或者一个以上处理器180执行一个或者一个以上程序包含用于进行以下操作的指令:

[0173] 获取视频数据,以及获取所述视频数据对应的拍摄参数;

[0174] 根据所述拍摄参数获取所述视频数据对应的运动轨迹与运动姿态;

[0175] 获取三维空间场景,并根据所述运动轨迹与运动姿态在所述三维空间场景中加载所述视频数据,获得所述视频数据对应的初始坐标;

[0176] 根据所述拍摄参数以及视频数据计算所述视频数据在所述三维空间场景中的投影坐标;

[0177] 根据所述投影坐标对所述初始坐标进行调整,获得调整后的坐标;

[0178] 根据所述调整后的坐标将所述视频数据与所述三维空间场景进行融合。

[0179] 在上述实施例中,对各个实施例的描述都各有侧重,某个实施例中未详述的部分,可以参见上文针对视频融合方法的详细描述,此处不再赘述。

[0180] 本领域普通技术人员可以理解,上述实施例的各种方法中的全部或部分步骤可以通过程序来完成,或通过程序控制相关的硬件来完成,该程序可以存储于一计算机可读存储介质中,并由处理器进行加载和执行。

[0181] 为此,本申请实施例提供一种存储介质,所述存储介质存储有计算机程序,该程序被处理器加载以执行本申请实施例所提供的任一种视频融合方法中的步骤。例如,该计算机程序可以执行如下步骤:

[0182] 获取视频数据,以及获取所述视频数据对应的拍摄参数;

[0183] 根据所述拍摄参数获取所述视频数据对应的运动轨迹与运动姿态;

[0184] 获取三维空间场景,并根据所述运动轨迹与运动姿态在所述三维空间场景中加载所述视频数据,获得所述视频数据对应的初始坐标;

[0185] 根据所述拍摄参数以及视频数据计算所述视频数据在所述三维空间场景中的投影坐标;

[0186] 根据所述投影坐标对所述初始坐标进行调整,获得调整后的坐标;

[0187] 根据所述调整后的坐标将所述视频数据与所述三维空间场景进行融合。

[0188] 以上各个操作的具体实施可参见前面的实施例,在此不再赘述。

[0189] 其中,该存储介质可以包括:只读存储器(ROM,Read Only Memory)、随机存取记忆体(RAM,Random Access Memory)、磁盘或光盘等。

[0190] 由于该存储介质中所存储的计算机程序,可以执行本申请实施例所提供的任一种视频融合方法中的步骤,因此,可以实现本申请实施例所提供的任一种视频融合方法所能实现的有益效果,详见前面的实施例,在此不再赘述。

[0191] 以上对本申请实施例所提供的一种视频融合方法、装置、设备及存储介质进行了详细介绍,本文中应用了具体个例对本申请的原理及实施方式进行了阐述,以上实施例的说明只是用于帮助理解本申请的方法及其核心思想;同时,对于本领域的技术人员,依据本申请的思想,在具体实施方式及应用范围上均会有改变之处,综上所述,本说明书内容不应理解为对本申请的限制。



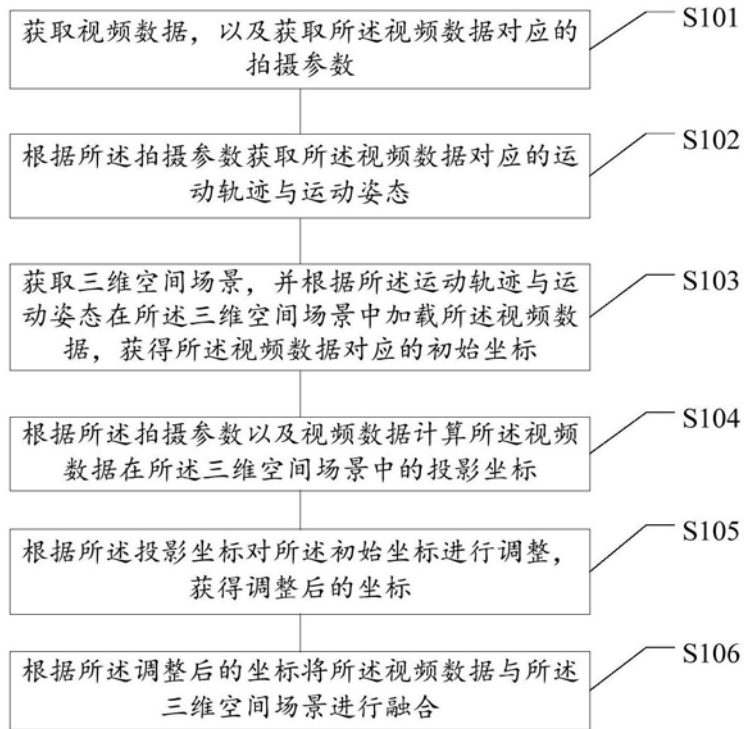


图1



图2

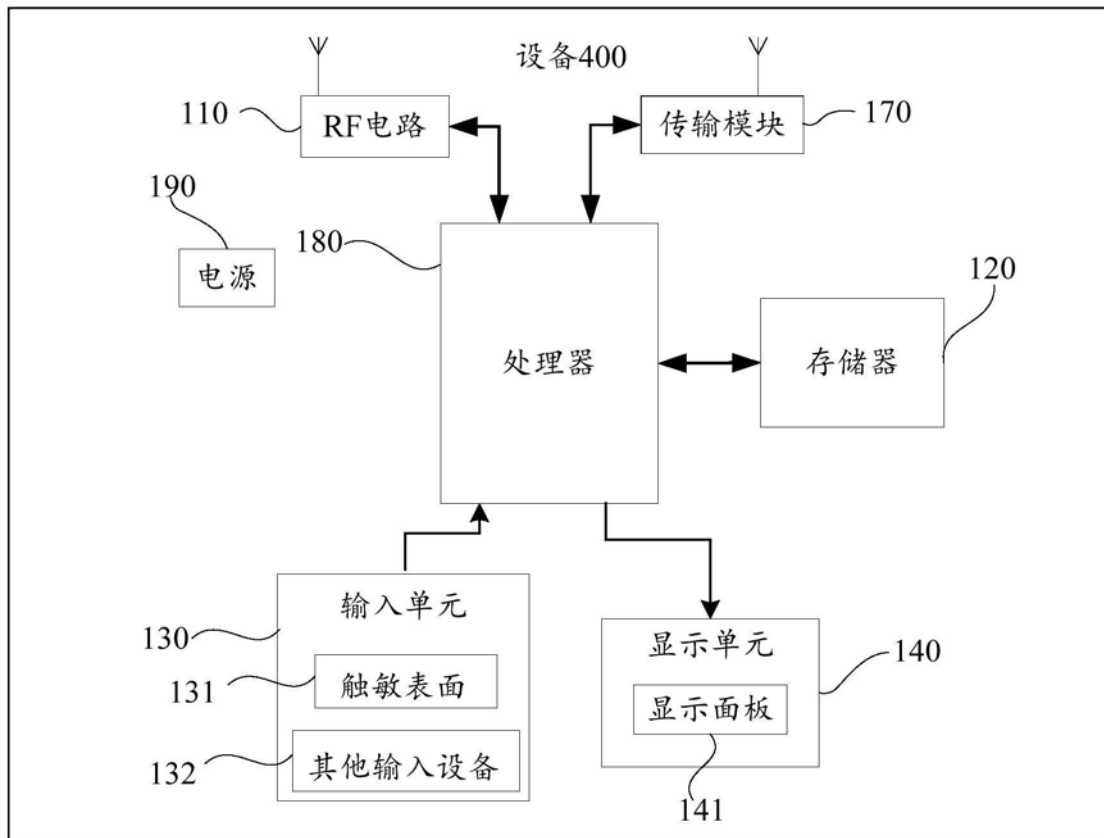


图3