



(45)授权公告日 2017.03.08

迪特尔·施马尔施蒂格

1. 一种用于姿势估计的方法,其包括:

接收以相机俘获的场景的查询图像;

使用当俘获所述查询图像时所述相机的近似位置来从数据库获得所述场景的多个参考图像,每一参考图像具有相关联位置估计,其中每一参考图像的所述相关联位置估计并不包含相关联定向;

基于所述查询图像和所述多个参考图像产生描述所述查询图像的相机中心的未知的位置与每一参考图像的参考相机中心之间的方向的方向向量;以及

使用所述方向向量和每一参考图像的所述相关联位置估计来确定具有六个自由度的所述相机的姿势。

2. 根据权利要求1所述的方法,其中产生描述所述查询图像的所述相机中心与每一参考图像的所述参考相机中心之间的方向的方向向量包括使用单应性、本质矩阵和基本矩阵中的至少一者。

3. 根据权利要求1所述的方法,其中通过使用所述方向向量和每一参考图像的所述相关联位置估计解决三点姿势问题来确定具有六个自由度的所述相机的所述姿势。

4. 根据权利要求1所述的方法,其中使用当俘获所述查询图像时所述相机的所述近似位置来从所述数据库获得所述场景的所述多个参考图像包括:

使用当俘获所述查询图像时所述相机的所述近似位置来产生参考图像的第一集合;以及

通过将所述查询图像中的特征匹配于每一参考图像中的特征来过滤参考图像的所述第一集合以产生所述多个参考图像。

5. 根据权利要求1所述的方法,其中每一参考图像具有相关联相机校准,其中产生所述方向向量使用每一参考图像的所述相关联相机校准。

6. 一种用于姿势估计的设备,其包括:

第一接口,其能够接收以相机俘获的场景的查询图像;

第二接口,其能够使用当俘获所述查询图像时所述相机的近似位置来从数据库获得所述场景的多个参考图像,每一参考图像具有相关联位置估计,其中每一参考图像的所述相关联位置估计并不包含相关联定向;以及

处理器,其经耦合以从所述第一接口接收所述查询图像且从所述第二接口接收所述多个参考图像,所述处理器经配置以基于所述查询图像和所述多个参考图像产生描述所述查询图像的相机中心的未知的位置与每一参考图像的参考相机中心之间的方向的方向向量,且使用所述方向向量和每一参考图像的所述相关联位置估计来确定具有六个自由度的所述相机的姿势。

7. 根据权利要求6所述的设备,其中所述设备进一步包括所述相机,并且其中所述第一接口是所述相机与所述处理器之间的接口且所述第二接口是无线接口。

8. 根据权利要求6所述的设备,其中所述第一接口是无线接口且所述第二接口是所述数据库与所述处理器之间的接口。

9. 根据权利要求6所述的设备,其中所述处理器经配置以使用单应性、本质矩阵和基本矩阵中的至少一者产生描述所述查询图像的所述相机中心与每一参考图像的所述参考相机中心之间的方向的方向向量。

10. 根据权利要求6所述的设备,其中所述处理器经配置以通过使用所述方向向量和每一参考图像的所述相关联位置估计解决三点姿势问题来确定具有六个自由度的所述相机的所述姿势。

11. 根据权利要求6所述的设备,其中所述处理器进一步经配置以通过将所述查询图像中的特征匹配于每一参考图像中的特征来过滤所述多个参考图像。

12. 根据权利要求6所述的设备,其中每一参考图像具有相关联相机校准,其中产生所述方向向量使用每一参考图像的所述相关联相机校准。

13. 一种用于姿势估计的设备,其包括:

用于接收以相机俘获的场景的查询图像的装置;

用于使用当俘获所述查询图像时所述相机的近似位置来从数据库获得所述场景的多个参考图像的装置,每一参考图像具有相关联位置估计,其中每一参考图像的所述相关联位置估计并不包含相关联定向;

用于基于所述查询图像和所述多个参考图像产生描述所述查询图像的相机中心的未知的位置与每一参考图像的参考相机中心之间的方向的方向向量的装置;以及

用于使用所述方向向量和每一参考图像的所述相关联位置估计来确定具有六个自由度的所述相机的姿势的装置。

14. 根据权利要求13所述的设备,其中所述用于产生描述所述查询图像的所述相机中心与每一参考图像的所述参考相机中心之间的方向的方向向量的装置使用单应性、本质矩阵和基本矩阵中的至少一者。

15. 根据权利要求13所述的设备,其中通过使用所述方向向量和每一参考图像的所述相关联位置估计解决三点姿势问题来确定具有六个自由度的所述相机的所述姿势。

16. 根据权利要求13所述的设备,其中所述用于使用当俘获所述查询图像时所述相机的所述近似位置来从所述数据库获得所述场景的所述多个参考图像的装置使用当俘获所述查询图像时所述相机的所述近似位置来产生参考图像的第一集合;以及通过将所述查询图像中的特征匹配于每一参考图像中的特征来过滤参考图像的所述第一集合以产生所述多个参考图像。

17. 根据权利要求13所述的设备,其中每一参考图像具有相关联相机校准,其中所述用于产生所述方向向量的装置使用每一参考图像的所述相关联相机校准。

基于场景结构的自我姿势估计

[0001] 待决临时申请案的交叉参考

[0002] 本申请案主张2012年7月12日申请且名称为“基于场景结构的自我姿势估计”的第13/547,987号美国申请案的优先权,所述美国申请案又在35USC119下主张2012年3月2日申请且名称为“基于场景结构的自我姿势估计”的第61/606,209号美国临时申请案的优先权,以上两者转让给本受让人且以引用的方式并入本文中。

技术领域

[0003] 本文所述的标的物的实施例大体上涉及位置和跟踪,且更具体地说涉及基于视觉的跟踪。

背景技术

[0004] 姿势估计描述确定相机相对于环境的外部校准的任务,其可为已知或未知的。最近基于图像的姿势估计方法已经证实为用于解决实现非常精确的6自由度(6DOF)姿势估计的任务的良好备选择。基于图像的姿势估计传统地需要场景的某种重构或3D模型。稀疏3D模型的使用就当其进行在线姿势估计时的准确性和速度而言是有利的,然而,这些模型的建造通常是复杂的离线过程。因此,所需的3D模型的可扩展性和可维护性仍然是未解决的问题。

发明内容

[0005] 使用在查询图像中俘获的场景结构和来自数据库的参考图像执行姿势估计。所述参考图像中的每一者具有相关联位置估计。基于所述查询图像和所述多个参考图像产生描述所述查询图像的相机中心的未知的位置与每一参考图像的参考相机中心之间的方向的方向向量。可使用例如单应性、本质矩阵或基本矩阵来产生所述方向向量。使用所述方向向量和每一参考图像的所述相关联位置估计来确定具有六个自由度的姿势估计。例如可通过使用所述方向向量和每一参考图像的所述相关联位置估计解决三点姿势问题来确定所述姿势估计。

[0006] 在一个实施方案中,一种方法包含:接收以相机俘获的场景的查询图像;从数据库接收所述场景的多个参考图像,每一参考图像具有相关联位置估计;基于所述查询图像和所述多个参考图像产生描述所述查询图像的相机中心的未知的位置与每一参考图像的参考相机中心之间的方向的方向向量;以及使用所述方向向量和每一参考图像的所述相关联位置估计来确定具有六个自由度的相机的姿势。

[0007] 在一个实施方案中,一种设备包含:第一接口,其能够接收以相机俘获的场景的查询图像;第二接口,其能够从数据库接收所述场景的多个参考图像,每一参考图像具有相关联位置估计;以及处理器,其经耦合以从所述第一接口接收所述查询图像且从所述第二接口接收所述多个参考图像,所述处理器经配置以基于所述查询图像和所述多个参考图像产生描述所述查询图像的相机中心的未知的位置与每一参考图像的参考相机中心之间的方

向的方向向量,且使用所述方向向量和每一参考图像的所述相关联位置估计来确定具有六个自由度的所述相机的姿势。

[0008] 在一个实施方案中,一种设备包含:用于接收以相机俘获的场景的查询图像的装置;用于从数据库接收所述场景的多个参考图像的装置,每一参考图像具有相关联位置估计;用于基于所述查询图像和所述多个参考图像产生描述所述查询图像的相机中心的未知的位置与每一参考图像的参考相机中心之间的方向的方向向量的装置用于;以及用于使用所述方向向量和每一参考图像的所述相关联位置估计来确定具有六个自由度的所述相机的姿势的装置。

[0009] 在一个实施方案中,一种包含存储于其上的程序代码的非暂时性计算机可读媒体,包含:用以接收以相机俘获的场景的查询图像的程序代码;用以从数据库接收所述场景的多个参考图像的程序代码,每一参考图像具有相关联位置估计;用以基于所述查询图像和所述多个参考图像产生描述所述查询图像的相机中心的未知的位置与每一参考图像的参考相机中心之间的方向的方向向量的程序代码;以及用以使用所述方向向量和每一参考图像的所述相关联位置估计来确定具有六个自由度的所述相机的姿势的程序代码。

附图说明

[0010] 图1A和1B在下文分别图解说明能够使用基于场景结构的姿势估计的移动装置的前侧和背侧。

[0011] 图2图解说明展示其中移动装置接入数据库以获得参考图像用于基于场景结构的姿势估计的系统的框图。

[0012] 图3是说明使用图像的数据库确定6DOF姿势的方法的流程图。

[0013] 图4示意性地图解说明使用图像的数据库确定6DOF姿势。

[0014] 图5示意性地图解说明产生参考图像的数据库。

[0015] 图6是能够使用基于场景结构的姿势估计的移动装置的框图。

[0016] 图7图解说明能够使用图像数据库确定6DOF姿势的服务器的框图。

具体实施方式

[0017] 以下图1A和1B分别图解说明移动装置100的前侧和背侧,所述移动装置100能够使用基于场景结构的姿势估计来确定6DOF姿势而无需3D重构。如本文所描述,姿势估计是使用在查询图像中俘获的场景结构和来自数据库的参考图像执行的。所述参考图像中的每一者具有相关联位置估计。举例来说,对于由所述查询图像和这些参考图像中的每一者所形成的对,使用例如SIFT(比例尺-不变量-特征-变换)等图像特征来估计极几何。从已知的极几何,可获得核点且连同所述参考图像的已知位置一起可处理为所述查询图像的单个2D-3D对应。基于至少三个参考图像,可使用标准三点姿势问题。查询相机的完整6DOF姿势可例如通过在稳健求解器中应用基于RANSAC(随机样本共识)的方案而在全局上下文中估计。此方法和从真实世界点及其图像观察估计相机的姿势大不相同。将已知参考相机位置处理为世界点且将相机相对于查询相机的放置处理为所述观察。

[0018] 如图1所说明,移动装置100包含外壳101、可为触摸屏显示器的显示器102以及扬声器104和麦克风106。移动装置100进一步包含面朝前的相机110以对环境成像。移动装置

100捕获场景的图像,例如作为视频帧,且使用所述图像与来自数据库的地理标记的参考图像以确定移动装置100的位置。移动装置100使用所述地理标记的参考图像与所述俘获的查询图像以基于多个单应的、本质的或基本的矩阵来确定6DOF姿势估计。

[0019] 如本文中所使用,移动装置指代任何便携式电子装置,例如蜂窝式或其它无线通信装置、个人通信系统(PCS)装置、个人导航装置(PND)、个人信息管理器(PIM)、个人数字助理(PDA)或其它合适的移动装置。移动装置可能接收无线通信和/或导航信号,例如导航定位信号。术语“移动装置”也既定包含(例如)通过短程无线、红外线、有线连接或其它连接而与个人导航装置(PND)通信的装置,而不管装置处或PND处是否发生卫星信号接收、辅助数据接收和/或定位相关处理。并且,“移动装置”既定包含能够俘获其环境的图像(或视频)的所有电子装置,包含无线通信装置、计算机、膝上型计算机、平板计算机等。

[0020] 图2图解说明展示其中移动装置100可经由无线网络120使用远程服务器130接入数据库135的系统的框图。数据库135可包含地理标记的参考图像。地理标记的参考图像的相关集可位于数据库135中,例如基于成像匹配或使用查询图像的粗略位置估计。查询图像的粗略的位置估计可使用常规定位技术从移动装置100获得,例如使用卫星定位系统(SPS)122的粗略的GPS测量或使用例如接入点124或蜂窝式塔126等无线源的三边测量。发射器的SPS系统122经定位以使得实体能够至少部分地基于从发射器接收的信号确定其在地球上或上方的位置。在特定实例中,此类发射器可位于地球轨道人造卫星(SV)上,例如,例如全球定位系统(GPS)、伽利略、Glonass或Compass等全球导航卫星系统(GNSS)或其它非全球系统的群集中。因此,如本文所使用,SPS可包含一或多个全球和/或地区性导航卫星系统和/或扩增系统的任何组合,且SPS信号可包含SPS、SPS状和/或与此类一或多个SPS相关联的其它信号。

[0021] 移动装置100还可将查询图像(或来自查询图像的特征集合)提供到服务器130,且将在服务器130上执行发现地理标记的参考图像的相关集合的工作负荷以及任何其它所希望的在此所述的过程。可替代地,服务器130可提供一组参考图像,或每一参考图像的预先计算的特征集合到移动装置100,且移动装置100可执行在此所述的过程。另外,参考图像的数据库或参考图像的预先计算的特征集合可存储在所述移动装置100上,进而避免通过网络120接入远程服务器130的需要。

[0022] 图3是说明使用图像的数据库确定6DOF姿势的方法的流程图。如所说明,接收由相机俘获的场景的查询图像I(202)。所述查询图像可为例如视频帧。举例来说,在服务器130执行姿势估计过程的情况下,所述查询图像I可由服务器130经由网络120从移动装置100接收。如果移动装置100执行姿势估计过程,那么所述查询图像I可例如由移动装置100中的处理器从相机110接收。所述查询图像可具有相关联位置估计,即,俘获所述查询图像的相机中心的大致位置估计,其可在俘获所述查询图像的时间期间或接近所述时间获得。与查询图像相关联的相关联位置估计可使用已知定位技术获得,包含GPS测量或使用无线源的三边测量,且因此是限于定位技术的准确性的大致位置,例如用于GPS类型测量的10到50英尺。

[0023] 从数据库接收场景的多个参考图像I',每一参考图像具有相关联位置估计(204)。举例来说,在服务器130执行姿势估计过程的情况下,所述多个参考图像I'可由服务器130中的处理器从数据库135接收。如果移动装置100执行姿势估计过程,那么所述多个参考图

像 I' 可由移动装置100经由服务器130和网络120从数据库135接收。每一参考图像的相关联位置估计可不包含相关联定向。因此,举例来说,基于查询图像的相关联位置估计,可通过过滤数据库而收集 N 个相关图像的第一集合 S 。第一集合 S 中的每一参考图像具有接近与查询图像相关联的位置估计的相关联位置估计。查询图像的位置估计与每一参考图像的位置估计之间的可接受距离可取决于若干因素,例如位置,例如对于大峡谷可使用比城市大的距离,以及参考图像的可用性,例如如果存在很少参考图像或场景结构远离则较大距离可为可接受的。举例来说,场景结构如果是通常从大距离成像的则可为远离的,例如,自由女神像通常是从大距离的船上拍照的。在第一集合 S 中的每一参考图像靠近拍摄查询图像的位置定位的情况下,很可能这些参考图像中的一些将描绘同一场景。参考图像集合可经精炼以仅包含描绘与查询图像相同场景的参考图像。举例来说,通过将查询图像中的特征与参考图像的第一集合中的每一参考图像中的特征进行比较且几何学上检验匹配特征,相关图像的第一集合 S 可经过滤以产生相关图像的第二集合 S' 。借助于实例,对于通过查询图像和参考图像所形成的所有图像对,可使用图像特征估计极几何。由于由特征匹配引起的可能错误而可使用用于离群值移除的稳健方案。

[0024] 基于查询图像 I 和多个参考图像 I' 产生三维(3D)方向向量,其描述针对查询图像 I 的相机中心的未知的位置与针对每一参考图像 I' 的参考相机中心之间的方向(206)。换句话说,确定查询相机 C 与参考相机 C' 之间的相对平移。可使用例如基本矩阵、本质矩阵或单应性,基于从查询图像 I 和每一参考图像 I' 提取的极几何产生方向向量。使用所述方向向量和每一参考图像的相关联位置估计来确定具有六个自由度的相机的姿势(208)。举例来说,可通过使用所述方向向量和每一参考图像的相关联位置估计解决三点姿势问题来确定具有六个自由度的相机的姿势。

[0025] 图4示意性地图解说明使用图像的数据库确定6DOF姿势。如所说明,俘获查询图像 I 且使用查询图像 I 的粗略的位置估计304来从数据库135获得参考图像 I' 306的集合 S 。

[0026] 图5借助于实例示意性地图解说明数据库135的产生。图5图解说明包含建筑物310、312的场景309的俯视图,其中存在通过场景309中的正方形识别许多图像 I'_A 、 I'_B 、 I'_C 和 I'_D (有时统称为图像 I')的位置。所述图像 I' 以及当拍摄图像时相机的相关联位置估计(pos)提供到数据库135且存储在其中。必要时,也可以将例如内部相机校准等与图像 I' 相关联的额外信息存储在数据库中。必要时,数据库135可仅存储从图像 I' 提取的特征,而不是整个图像自身。

[0027] 提供查询图像 I 和参考图像 I' 的集合 S 用于几何恢复316。如上文所论述,可使用查询图像 I 的粗略的位置估计产生的参考图像的初始集合可通过从使用例如SIFT、SURF(加速稳健特征)等提取的图像 I 和 I' 的图像特征的匹配318而精炼。可在计算图像对之间的极几何同时执行匹配318过程。模型估计320产生每一图像对 I 与 I' 之间的相对姿势作为例如基本矩阵 F 、本质矩阵 E 或单应性 H 。

[0028] 一般来说,如果没有关于相机的信息是已知的,那么可确定基本矩阵 F ,而如果内部校准 K 对于相机是已知的,那么可确定本质矩阵 E 。如果另外可假定所述场景中存在主要平面,那么可确定引起的单应性 H 。另外,所需的最小数目的对应随着额外信息减少,例如,基本矩阵 F 需要至少7个,本质矩阵需要5个,且单应性 H 需要仅4个。

[0029] 通过基本矩阵 F 描述的例如查询图像 I 和参考图像 I' 中的每一者的图像对之间的

极几何可使用例如SIFT (比例尺不变量特征变换)、SURF (加速稳健特征) 等图像特征来估计。对于两个图像I和I', 将任意场景特征点X描绘为I中的观察x和I'中的观察x'。通过匹配所述观察x和x'的描述符, 建立图像I与I'之间的对应。对于大量图像特征, 可以此方式获得对应性的整个集合。可从所述集合S丢弃具有小于匹配所述查询图像的特征的阈值的任何参考图像。所述阈值可取决于许多因素, 包含: a) 所描绘的场景, b) 所采用特征的估计性能, 和c) 初始匹配的数目。良好的试探是接受具有至少算法所需的对应性数目的对以较晚地估计基本矩阵F、本质矩阵E或单应性H, 例如如果采用标准8点算法, 那么单应性可使用4+1, 同时基本矩阵F、本质矩阵E可使用8+1。

[0030] 基本矩阵F包封查询相机C和参考相机C'的完整固有投影几何, 且可使用非线性优化从查询图像I与参考图像I'之间的至少6个对应性来估计。可以线性最小平方方式使用8个对应性。基本矩阵F是3x3矩阵, 以使得对于图像I和I'中的单个场景点X_i的观察x_i和x'_i

$$[0031] \quad \mathbf{x}_i^T \mathbf{F} \mathbf{x}_i = 0. \quad \text{等式1}$$

[0032] 可替代地, 如果查询相机C和参考相机C'两者的内部相机校准K和K'是已知的, 那么可将相机校准应用于所述点以获得其归一化的座标为

$$[0033] \quad \hat{\mathbf{x}}_i = \mathbf{K}^{-1} \mathbf{x}_i \text{ 和 } \hat{\mathbf{x}}'_i = \mathbf{K}'^{-1} \mathbf{x}'_i. \quad \text{等式2}$$

[0034] 从所述归一化的坐标, 本质矩阵E可估计为基本矩阵F的专门的形式, 以使得

$$[0035] \quad \hat{\mathbf{x}}_i^T \mathbf{E} \hat{\mathbf{x}}_i = 0. \quad \text{等式3}$$

[0036] 本质矩阵E可从查询图像I与参考图像I'之间的至少5个对应性估计, 且如下与基本矩阵F有关

$$[0037] \quad \mathbf{E} = \mathbf{K}^T \mathbf{F} \mathbf{K} \quad \text{等式4}$$

[0038] 由于图像I与I'之间的对应性集合中错误的存在, 可采用RANSAC稳健地执行基本矩阵F或本质矩阵E的估计。

[0039] 基于查询图像I与参考图像I'之间的极几何, 可确定查询相机C与参考相机C'之间的相对平移, 进而提供图4中的3D方向向量作为基本矩阵、本质矩阵。举例来说, 关于C相对于C'的相对平移的信息即方向向量可如下提取。

[0040] F的右零向量是核点e, 其基本上是图像I中的参考相机C'的观察。实际上, 所述核点e描述从查询相机C见到参考相机C'的方向。在通过查询相机C的内部校准K标准化核点e之后,

$$[0041] \quad \hat{\mathbf{e}} = \mathbf{K}^{-1} \mathbf{e} \quad \text{等式5}$$

[0042] 描述查询相机C的坐标系中参考相机C'必然驻留的方向。等效地, 如果有可能估计图像I和I'的本质矩阵E, 那么核点 $\hat{\mathbf{e}}$ 可直接计算为E的右零向量。在分解之后存在符号不明确性, 且因此, 基本矩阵F和本质矩阵E提供无方向向量。

[0043] 除使用基本矩阵或本质矩阵导出方向向量324之外, 可使用基于单应性的模型估计320来产生查询图像的未知的位置与同每一参考图像相关联的估计位置之间的方向向量324。在场景具有主要平面结构 (人造环境中常见) 且查询相机C和参考相机C'两者的内部校准K和K'已知的情况下, 可采用单应性而不是极几何来估计相机相对于彼此的相对放置。

[0044] 给定单个物理平面IT, I与I'之间的最少4个对应性足以用于2D单应性H的估计。单应性H包封查询图像I与参考图像I'之间的2D投影变换, 其对于单一底层平面II有效。对于

II上的任意场景点 X_i , 3×3 矩阵H将I中的点的观察 $\hat{x}_i$ 变换到其在I'中的观察 $\hat{x}'_i$, 以使得

$$[0045] \quad H\hat{x}_i = \hat{x}'_i. \quad \text{等式6}$$

[0046] 类似于极几何, 单应性H对查询相机C与参考相机C'之间的相对姿势改变P进行编码:

$$[0047] \quad P = [R | t], \quad \text{等式7}$$

[0048] 其中R和t分别表示所述姿势改变的旋转和平移部分。让H表示查询图像I与I'之间的单应性, 随后

$$[0049] \quad H = R + \frac{1}{d} \mathbf{t} \mathbf{n}^T \quad \text{等式8}$$

[0050] 其中d是从查询相机C到给定平面II的距离, 且n是II相对于查询相机C的法向量, 且R和t是参考相机C'相对于查询相机C的相对旋转和平移。执行矩阵H从等式8的分解323, 其中为了数值稳定性, 在分解之前归一化H以使得 $\det(H) = 1$ 。因此, 计算用于R、t、d和n的两个物理上可能的解。

[0051] 无法从H完全推断查询相机C与参考相机C'之间的真实平移。然而, 考虑特定参考相机C'的位置, t界定查询相机C必须驻留的方向。平移t可经变换到查询相机C的坐标系中, 获得

$$[0052] \quad \mathbf{v} = -R^T \mathbf{t} \quad \text{等式9}$$

[0053] 其中v表示参考相机C'相对于查询相机C的坐标系的方向向量324。类似于上述使用基本矩阵或本质矩阵获得的方向向量324, 可随后使用使用单应性产生的方向向量324连同参考图像I'的位置估计326用于具有6个自由度oC查询相机C的稳健姿势估计328。

[0054] 在描述查询图像I的未知的位置与参考图像I'的每一相关联位置估计之间的方向的方向向量324和与参考图像I'中的每一者相关联的位置估计326的情况下, 可使用三点姿势问题执行查询相机C的稳健姿势估计328。对于针孔相机模型, 已知相机校准K允许图像测量 x_i 到光线 $\mathbf{v}_i$ 的转换且可测量其成对角度 $\angle(\mathbf{v}_i; \mathbf{v}_j)$ 。在针孔相机模型中, 三个已知3D点 X_i 及其对应图像测量 x_i 引起三个成对角测量。这些足以计算用于相机位置和定向的有限数目的解。为了确定方案的似真性, 使用3D点的第四观察。

[0055] 然而, 不同于针孔相机模型, 本情况并不在像空间中操作。换句话说, 光线 $\mathbf{v}_i$ 是参考相机C'的估计方向向量 $\hat{\mathbf{e}}_i$, 其被获得作为来自基本矩阵F或本质矩阵E的核点或作为从单应性H获得的平移方向。参考相机C'具有已知3D点 X_i 。可计算所述成对角度测量 $\angle(\hat{\mathbf{e}}_i; \hat{\mathbf{e}}_j)$, 其产生与针孔相机情况中相同的等式系统。对于三个已知参考相机位置C' i , 可计算成对3D点距离 $l_{i,j}$ 。此外, 角度 $\theta_{i,j}$ 对于对应方向向量 $\hat{\mathbf{e}}_i$ 和 $\hat{\mathbf{e}}_j$ 是已知的。未知量是查询相机C的中心(在针孔情况下投影的中心的等效物)与相机中心C' i 之间的距离 d_i :

$$[0056] \quad l_{i,j} = ||C'_i - C'_j||$$

$$[0057] \quad \theta_{i,j} = \angle(\hat{\mathbf{e}}_i, \hat{\mathbf{e}}_j) \quad \text{等式10}$$

$$[0058] \quad d_i = ||C - C'_i||.$$

[0059] 使用余弦定理, 所述三点对中的每一者给出一个等式:

[0060] $l_{i,j}^2 = C_i'^2 + C_j'^2 - 2C_i'C_j'\cos(\theta_{i,j})$ 等式11

[0061] 等式11是如在常规针孔相机模型的情况下的相同多项式系统且可以相同技术解决,其中差异是在针孔相机模型情况中相机校准矩阵K用以将图像测量转换到向量且因此成对欧几里得角度测量,而在本情况中,光线由不同的姿势的平移部分界定。

[0062] 在查询相机C见到参考相机C'的方向上存在符号不明确性。标准三点姿势算法的实施使用导引光线,其忽略位于相机中心后方的3D点的可能性。解决此问题的一种方法是相应修改实施方案,从而考虑四次式的负根作为有效解。这将随后可能产生部分地点镜像的姿势。那些可通过检查旋转分量的负行列式而证实。可替代地,可使用所属领域的技术人员已知的广义三点姿势算法,因为其暗含地以可忽略的额外成本对未经引导光线起作用且返回相同解。

[0063] 需要至少三个点-光线对来解决所述问题,同时使用第四对来验证所述解的物理似真性。因此,理论上可使用最少4个参考图像I针对查询相机C计算完整6DOF姿势,然而实际上,可能需要使用10个或10个以上参考图像I'解决所述问题。此外,采用RANSAC的稳健估计方案可以用作通过单应性递送的方向估计,本质矩阵和基本矩阵提供两个可能解。

[0064] 因此,无需具有当前场景的明确重构且无需明确重构场景便解决完整6DOF姿势估计的问题。解的准确性直接取决于参考图像I'的数目和几何估计的质量。因此用于几何恢复316(即单应性、本质或基本矩阵估计)的计算努力可针对不同准确性水平来缩放。

[0065] 图6是能够使用图像数据库确定6DOF姿势的移动装置100的框图。移动装置100包含用于俘获图像的相机110。移动装置100可进一步包含用于从网络120(图2中所示)接收无线信号的无线接口170。无线接口170可使用各种无线通信网络,例如无线广域网(WWAN)、无线局域网(WLAN)、无线个域网(WPAN)等。术语“网络”与“系统”经常可互换使用。WWAN可为码分多址(CDMA)网络、时分多址(TDMA)网络、频分多址(FDMA)网络、正交频分多址(OFDMA)网络、单载波频分多址(SC-FDMA)网络、长期演进(LTE)等。CDMA网络可以实施一或多种无线电接入技术(RAT),例如cdma2000、宽带CDMA(W-CDMA)等。Cdma2000包含IS-95、IS-2000和IS-856标准。TDMA网络可以实施全球移动通信系统(GSM)、数字高级移动电话系统(D-AMPS)或某种其它RAT。来自名为“第三代合作伙伴计划”(3GPP)的协会的文献中描述了GSM和W-CDMA。来自名为“第三代合作伙伴计划2”(3GPP2)的协会的文献中描述了cdma2000。3GPP和3GPP2文献是公众可获得的。WLAN可为IEEE 802.11x网络,且WPAN可为蓝牙网络、IEEE802.15x或某种其它类型的网络。此外,可使用WWAN、WLAN和/或WPAN的任何组合。

[0066] 移动装置100可任选地包含非视觉导航传感器171,例如运动或位置传感器,例如加速度计、陀螺仪、电子罗盘或其它类似运动感测元件。导航传感器171的使用可帮助上文描述的方法的多个动作。

[0067] 移动装置100可进一步包含用户接口103,其包含显示器102、小键盘105或其它输入装置,用户可通过其将信息输入到移动装置100中。必要时,可通过将虚拟小键盘集成到具有触摸传感器的显示器102中而排除小键盘105。用户接口103还可包含麦克风106和扬声器104,例如如果移动装置100是例如蜂窝式电话等移动装置。当然,移动装置100可包含与本发明不相关的其它元件。

[0068] 图7图解说明服务器130的框图,其可类似于移动装置100且能够使用图像数据库

确定6DOF姿势。服务器130包含外部接口131,其用以经由网络120与移动装置100通信。外部接口131可为到路由器(未图示)的有线接口或用于任何各种无线通信网络中的无线接口。服务器130可进一步包含用户接口133,其可包含例如显示器以及小键盘或其它输入装置,通过其用户可将信息输入到服务器130中。服务器130进一步图解说明为包含数据库135,但应理解,数据库135可在服务器130外部。

[0069] 移动装置100和服务器130两者图解说明为包含控制单元,其经连接以从移动装置中的相机110和服务器130中的数据库135接收查询图像。如图6中所说明,控制单元180耦合到相机110、SPS接收器150、无线接口170和导航传感器171(如果包含)。在移动装置100中,控制单元180经由第一接口111接受且处理来自相机110的数据,以及经由第二接口即无线接口170接受且处理从数据库135接收的参考图像I',如上文所论述。如图7中所说明,控制单元耦合到外部接口131、用户接口133和数据库135。因此,在服务器130中控制单元180经由第一接口111即外部接口131接受且处理来自相机110的数据,以及经由第二接口137接受且处理来自数据库135的参考图像I',如上文所论述。

[0070] 移动装置100或服务器130中的控制单元180可由总线180b、处理器181和相关联存储器184、硬件182、软件185和固件183提供。控制单元180可包含提取且匹配查询图像I与参考图像I'之间的特征的图像匹配单元192,和使用例如基本矩阵、本质矩阵或单应性产生查询图像I与参考图像I'之间的方向向量的模型估计单元194,如上文所论述。另外,稳健姿势估计单元196,其使用所述方向向量和与参考图像I'相关联的位置估计确定6DOF姿势。图像匹配单元192、模型估计单元194和稳健姿势估计单元196为了清楚起见而单独地且与处理器181分开图解说明,但可为单个单元和/或基于在处理器181中运行的软件185中的指令在处理器181中实施。将理解如本文所使用,处理器181以及图像匹配单元192、模型估计单元194和稳健姿势估计单元196中的一或多者可以但不必包含一或多个微处理器、嵌入处理器、控制器、专用集成电路(ASIC)、数字信号处理器(DSP)及类似者。术语处理器既定描述由系统而非具体硬件所实施的功能。此外,如本文中所使用,术语“存储器”指代任何类型的计算机存储媒体,其包含与移动平台相关联的长期、短期或其它存储器,且并不限于任何特定类型的存储器或特定数目的存储器,或特定类型的其上存储有存储器的媒体。

[0071] 因此,移动装置100和服务器130两者包含用于接收以相机俘获的场景的查询图像的装置,其可为例如移动装置100中的接口111或服务器130中的外部接口131。用于从数据库接收场景的多个参考图像(每一参考图像具有相关联位置估计)的装置可为移动装置100中的无线接口170或服务器130中的接口137。用于基于查询图像和多个参考图像产生描述查询图像的相机中心的未知的位置与每一参考图像的参考相机中心之间的方向的方向向量的装置可为例如模型估计单元194或执行从软件185接收的指令的处理器181。用于使用每一参考图像的所述方向向量和相关联位置估计来确定具有6个自由度的相机的姿势的装置可为稳健姿势估计单元196或执行从软件185接收的指令的处理器181。

[0072] 取决于应用,可通过各种装置实施本文中所描述的方法。举例来说,这些方法可在硬件182、固件183、软件185或其任何组合中实施。对于硬件实施方案,处理单元可实施于一或多个专用集成电路(ASIC)、数字信号处理器(DSP)、数字信号处理装置(DSPD)、可编程逻辑装置(PLD)、现场可编程门阵列(FPGA)、处理器、控制器、微控制器、微处理器、电子装置、经设计以执行本文中所描述的功能的其它电子单元,或其组合内。

[0073] 对于固件和/或软件实施方案,可用执行本文中所描述的功能的模块(例如,程序、功能等等)实施方法。在实施本文所述的方法时,可以使用任何有形地体现指令的机器可读媒体。举例来说,软件代码可存储在存储器184中且由处理器181执行。存储器可在处理器181内或外部实施。

[0074] 如果以固件和/或软件实施,那么可将功能作为一个或一个以上指令或代码存储在计算机可读媒体上。实例包含经编码有数据结构的非暂时性计算机可读媒体,和经编码有计算机程序的计算机可读媒体。计算机可读媒体包括物理计算机存储媒体。存储媒体可为可由计算机存取的任何可用媒体。借助于实例而非限制,此类计算机可读媒体可包括RAM、ROM、EEPROM、CD-ROM或其它光盘存储装置、磁盘存储装置或其它磁性存储装置、或可用以以指令或数据结构形式存储所要程序代码且可由计算机存取的任何其它媒体;如本文所使用,磁盘和光盘包含压缩光盘(CD)、激光光盘、光学光盘、数字多功能光盘(DVD)、软磁盘和蓝光光盘,其中磁盘通常以磁性方式再现数据,而光盘用激光以光学方式再现数据。以上的组合也应当包含在计算机可读媒体的范围内。

[0075] 尽管出于指导性目的,结合具体实施例来说明本发明,但本发明并不限于这些具体实施例。在不脱离本发明的范围的情况下可作出各种适应和修改。因此,不应将所附权利要求书的精神和范围限于前述描述。

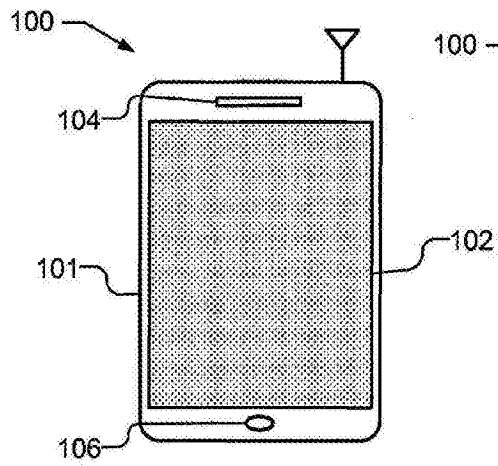


图1A

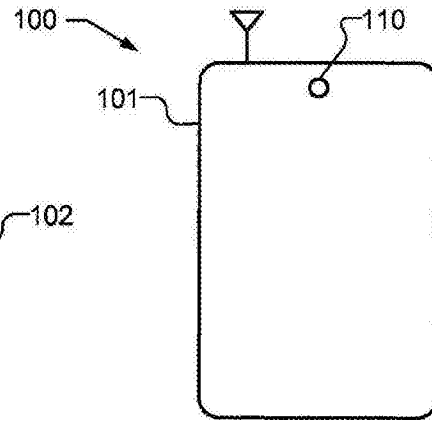


图1B

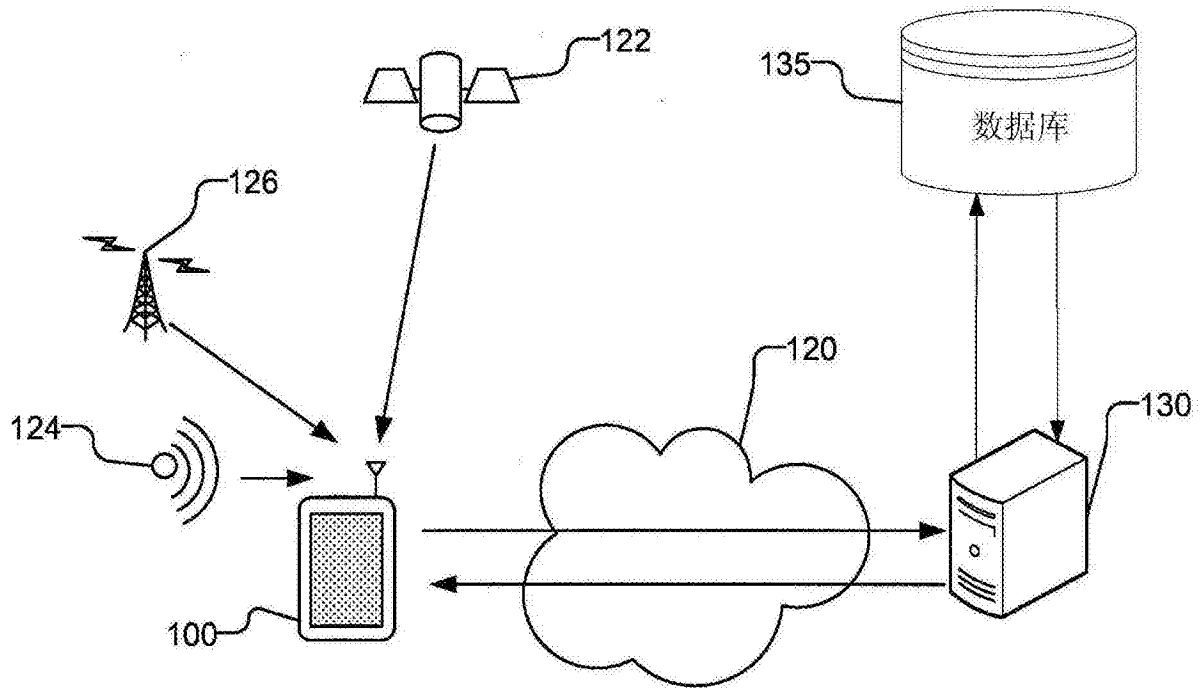


图2

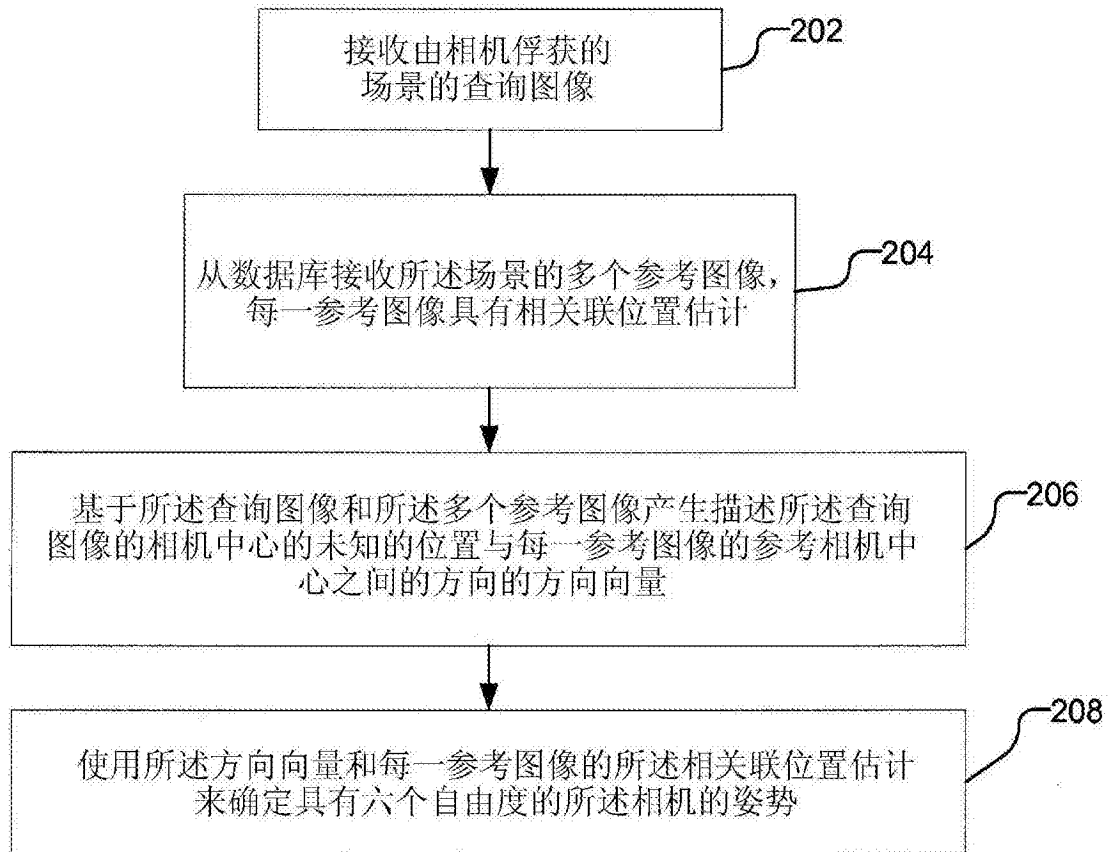


图3

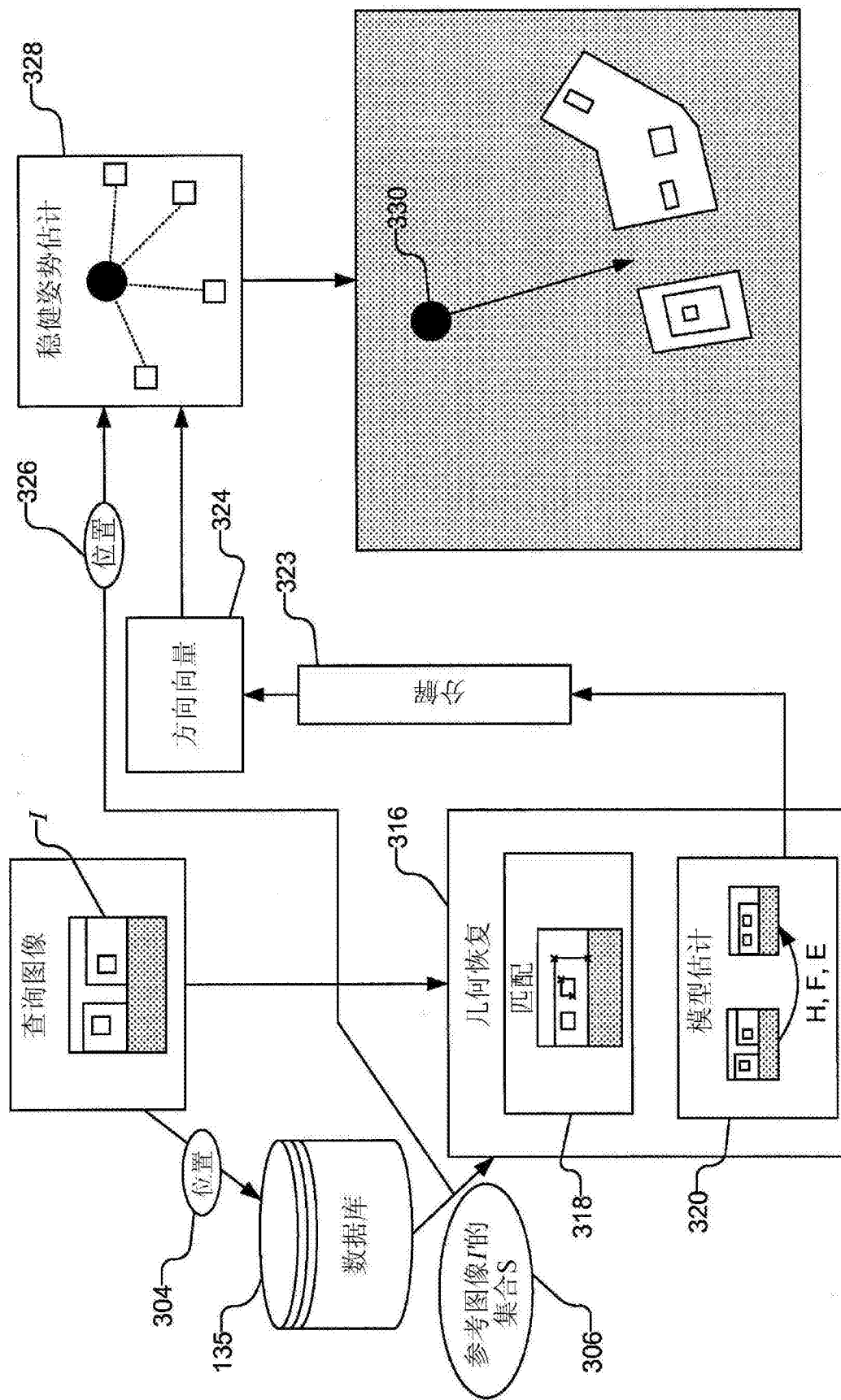


图4

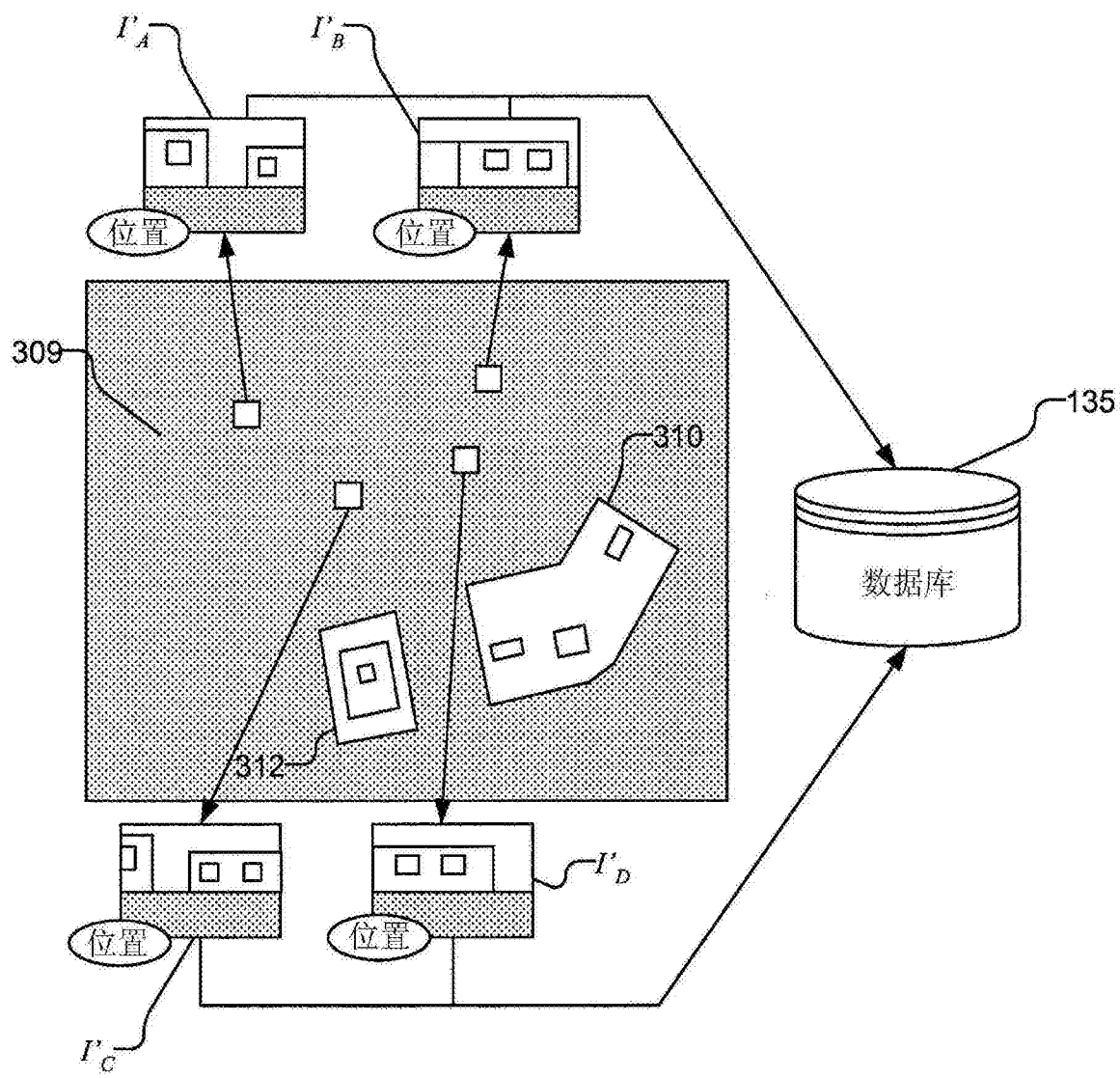


图5

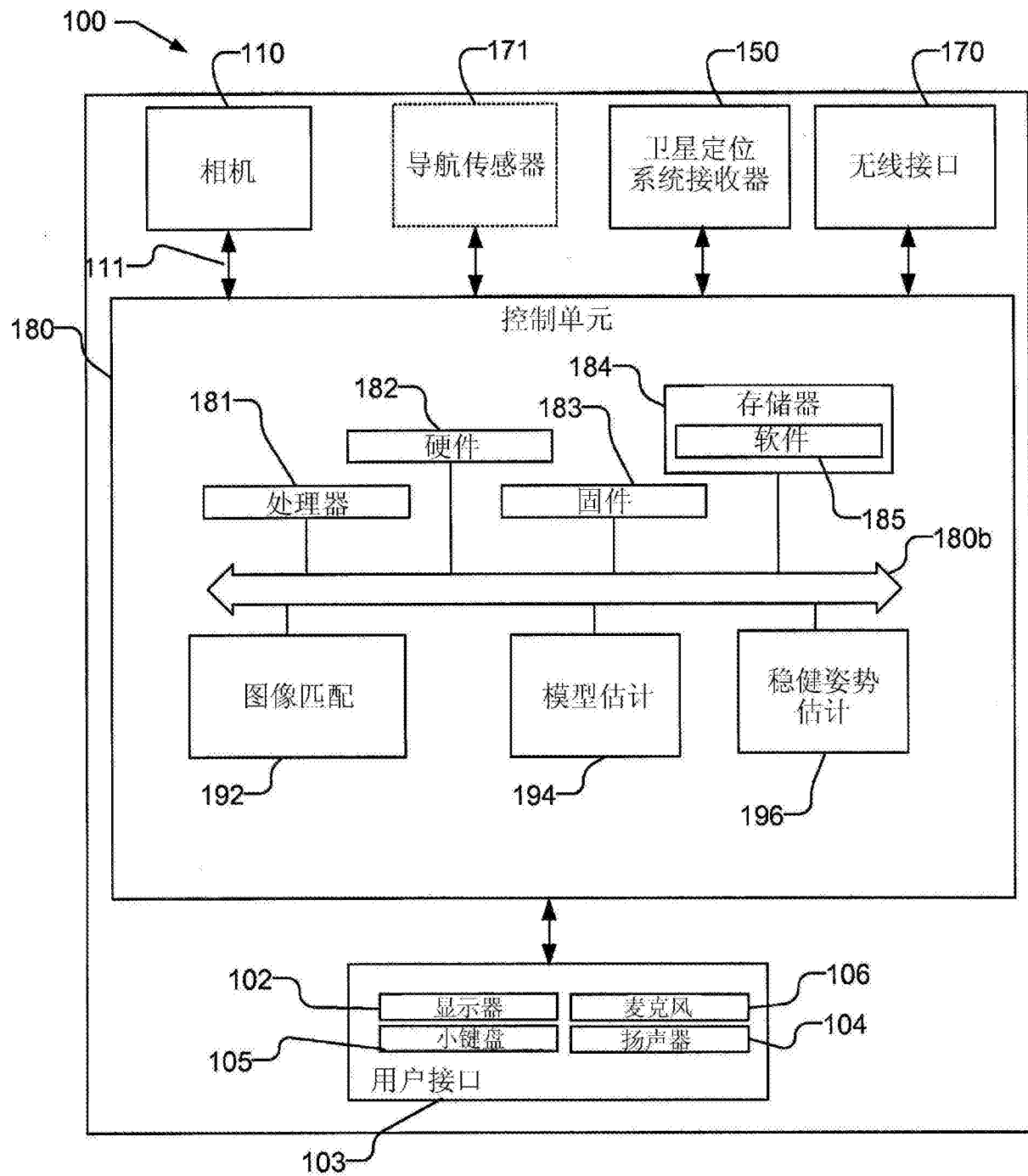


图6

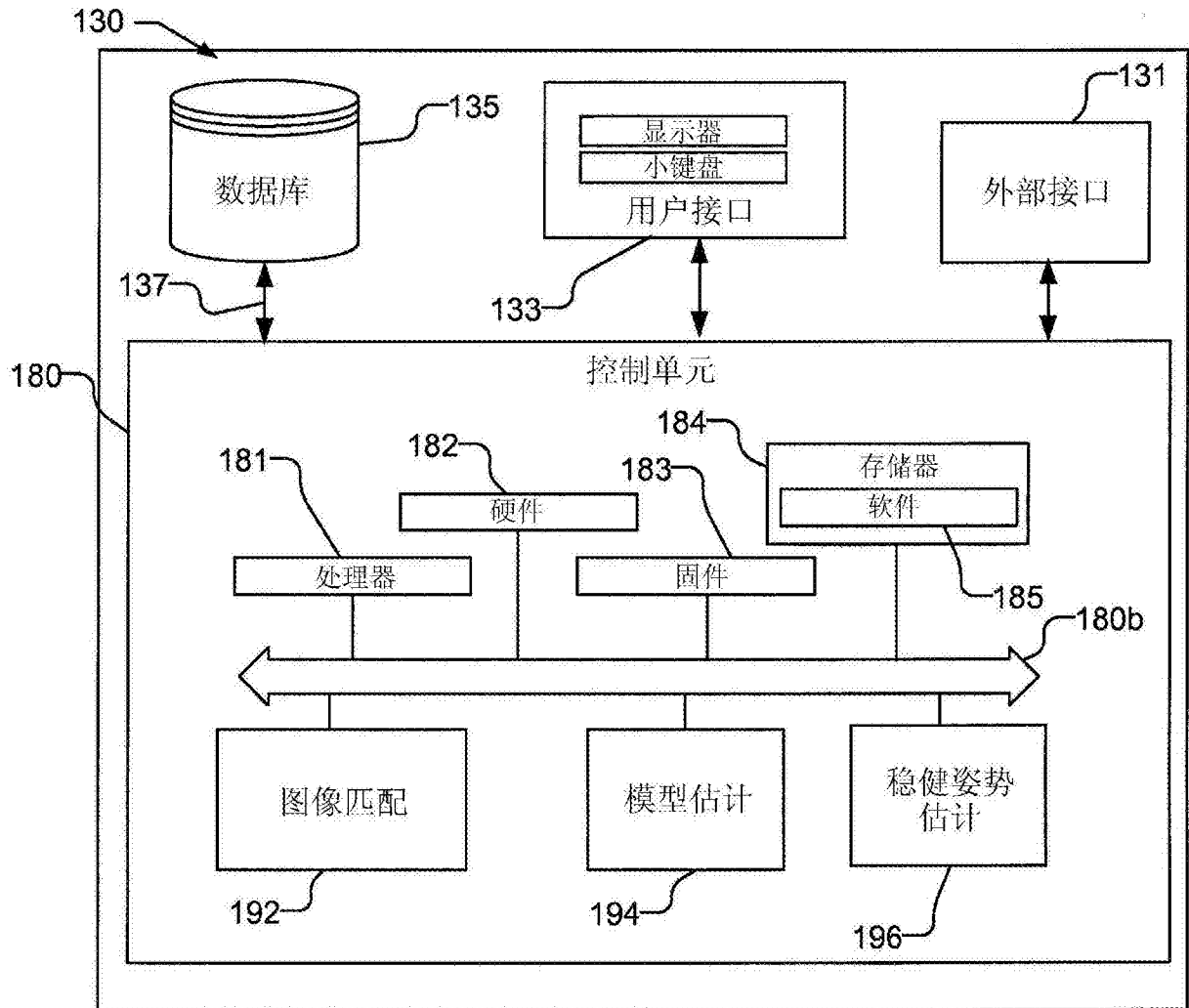


图7