



(12) 发明专利申请

(10) 申请公布号 CN 112102389 A

(43) 申请公布日 2020.12.18

(21) 申请号 202010963915.8

(22) 申请日 2014.11.21

(62) 分案原申请数据

201480083523.5 2014.11.21

(71) 申请人 苹果公司

地址 美国加利福尼亚

(72) 发明人 S·克诺尔 D·库尔茨

(74) 专利代理机构 中国贸促会专利商标事务所  
有限公司 11038

代理人 罗亚男

(51) Int. Cl.

G06T 7/579 (2017.01)

G06T 7/246 (2017.01)

G06T 7/277 (2017.01)

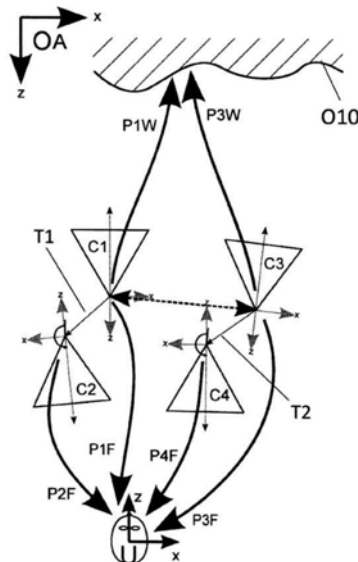
权利要求书2页 说明书28页 附图13页

(54) 发明名称

确定实物至少一部分的3D重构件的空间坐标的方法和系统

(57) 摘要

本发明提供了确定实物至少一部分的3D重构件的空间坐标的方法和系统。该方法包括：从第一相机系统获取包括第一实物的至少一部分的第一图像；从第二相机系统获取第二图像，第二图像包括与已知的几何性质相关联的第二实物的至少一部分，其中，第一相机系统和第二相机系统具有已知的空间关系；基于第二图像和第二实物的所述至少一部分的已知的几何性质来确定第二实物的至少一部分的比例；和基于所述已知的空间关系和第二实物的所述至少一部分的所述比例来确定第一实物处于绝对比例的特性。



1. 一种确定第一实物的至少一部分的3D重构件的空间坐标的方法,包括:  
从第一相机系统获取包括第一实物的至少一部分的第一图像;  
从第二相机系统获取第二图像,第二图像包括与已知的几何性质相关联的第二实物的至少一部分,其中,第一相机系统和第二相机系统具有已知的空间关系;  
基于第二图像和第二实物的所述至少一部分的所述已知的几何性质来确定第二实物的至少一部分的比例;和  
基于所述已知的空间关系和第二实物的所述至少一部分的所述比例来确定第一实物处于绝对比例的特性。
2. 如权利要求1所述的方法,其中,第二实物包括人脸的至少一部分。
3. 如权利要求2所述的方法,其中,所述人脸的至少一部分的所述比例基于通用人脸模型而确定。
4. 如权利要求2所述的方法,其中,所述人脸属于特定用户,并且其中,所述人脸的所述至少一部分的所述比例基于用户特定的人脸模型而确定。
5. 如权利要求2-4中任一项所述的方法,其中,第一相机系统和第二相机系统包括在电子设备中。
6. 如权利要求1-4中任一项所述的方法,其中,第一相机系统的视锥和第二相机系统的视锥不重叠。
7. 如权利要求1-4中任一项所述的方法,还包括:  
根据所述第一实物处于绝对比例的所述特性来以与第一实物一致的比例呈现虚拟物体。
8. 一种非瞬态计算机可读介质,包括用于确定第一实物的至少一部分的3D重构件的空间坐标的计算机可读代码,所述计算机可读代码可由处理器执行来:  
从第一相机系统获取包括第一实物的至少一部分的第一图像;  
从第二相机系统获取第二图像,第二图像包括与已知的几何性质相关联的第二实物的至少一部分,其中,第一相机系统和第二相机系统具有已知的空间关系;  
基于第二图像和第二实物的所述至少一部分的所述已知的几何性质来确定第二实物的至少一部分的比例;和  
基于所述已知的空间关系和第二实物的所述至少一部分的所述比例来确定第一实物处于绝对比例的特性。
9. 如权利要求8所述的非瞬态计算机可读介质,其中,第二实物包括人脸的至少一部分。
10. 如权利要求9所述的非瞬态计算机可读介质,其中,所述人脸的至少一部分的所述比例基于通用人脸模型而确定。
11. 如权利要求9所述的非瞬态计算机可读介质,其中,所述人脸属于特定用户,并且其中,所述人脸的所述至少一部分的所述比例基于用户特定的人脸模型而确定。
12. 如权利要求8-11中任一项所述的非瞬态计算机可读介质,其中,第一相机系统和第二相机系统包括在电子设备中。
13. 如权利要求8-11中任一项所述的非瞬态计算机可读介质,其中,第一相机系统的视锥和第二相机系统的视锥不重叠。

14. 如权利要求8-11中任一项所述的非瞬态计算机可读介质,还包括用于执行以下操作的计算机可读代码:

根据第一实物处于绝对比例的所述特性来以与第一实物一致的比例呈现虚拟物体。

15. 一种用于确定第一实物的至少一部分的3D重构件的空间坐标的系统,包括:

第一相机系统;

第二相机系统,其与第一相机系统具有已知的空间关系;

一个或多个处理器;和

计算机可读介质,包括计算机可读代码,所述计算机可读代码可被所述一个或多个处理器执行来:

从第一相机系统获取包括第一实物的至少一部分的第一图像;

从第二相机系统获取第二图像,第二图像包括与已知的几何性质相关联的第二实物的至少一部分;

基于第二图像和第二实物的所述至少一部分的所述已知的几何性质来确定第二实物的至少一部分的比例;和

基于所述已知的空间关系和第二实物的所述至少一部分的所述比例来确定第一实物处于绝对比例的特性。

16. 如权利要求15所述的系统,其中,第二实物包括人脸的至少一部分。

17. 如权利要求16所述的系统,其中,所述人脸的至少一部分的所述比例基于通用人脸模型而确定。

18. 如权利要求16所述的系统,其中,所述人脸属于特定用户,并且其中,所述人脸的所述至少一部分的所述比例基于用户特定的人脸模型而确定。

19. 如权利要求15-18中任一项所述的系统,其中,第一相机系统的视锥和第二相机系统的视锥不重叠。

20. 如权利要求15-18中任一项所述的系统,还包括用于执行以下操作的计算机可读代码:

根据所述第一实物处于绝对比例的所述特性来以与第一实物一致的比例呈现虚拟物体。

## 确定实物至少一部分的3D重构件的空间坐标的方法和系统

[0001] 本申请是申请日为2014年11月21日申请号为201480083523.5发明名称为“用于确定实物的至少一部分的处于绝对空间比例的3D重构件的空间坐标的方法和系统”的发明专利申请的分案申请。

### 技术领域

[0002] 本公开涉及一种方法和系统,用于确定实物的至少一部分的处于绝对空间比例的3D重构件的空间坐标。

### 背景技术

[0003] 涉及图像分析的计算机视觉方法常常用于,例如,导航、物体识别、3D重构、和增强现实应用程序等。图像可由单个相机或不同的相机来捕获。图像特征(诸如角和边)检测与图像特征提取是各种计算机视觉方法或算法中的常见步骤,诸如,图像型识别、图像型追踪、图像型重构、图像型分类和图像扭曲。例如,视觉型同步定位与地图构建(SLAM)是一种众所周知的计算机视觉方法,该方法通过一个或多个相机来重构真实环境并对此一个或多个相机进行追踪。假设一个或多个相机捕获到了至少两张图像,典型的SLAM方法包括:特征检测、说明、匹配、三角测量、以及(全局)图优化。

[0004] 根据单眼捕获设备捕获的一组图像来确定实物结构,这种方式形成了依据比例而定的空间(或几何)结构的重构件。即,重构使用对其而言绝对空间单位(诸如,单位米)的比例因子未知的空间单位。在许多应用中,需要获得例如具有绝对单位的重构件,也称作“处于绝对比例”的重构件。这常常要求了解处于绝对比例的至少一个距离,例如,实物多个部分之间的距离、或拍摄用于进行重构的相应图像时相机相对于实物的位置之间的距离。

[0005] 因此,各种SLAM和SfM系统的常见问题是:真实环境的几何模型是依据作为不确定因子的比例来重构。如果实物位置且相机拍摄重构图像的位姿也未知,则无法确定场景的绝对空间比例。例如,根据图2a所示的两张图像——一张从前方I(W1)进行拍摄,一张从右边I(W2)进行拍摄——无法断定它是实物大小的真车还是逼真的小型迷你车。因此,无法断定拍摄这两张图像的相机彼此间隔多少米(在汽车是实物大小的情况下)或仅仅间隔几厘米(在汽车是迷你型的情况下)。然而,如果相机绝对空间比例(例如,两个相机相隔2.34米)或物体的多个部分的绝对空间比例(例如,汽车头灯相隔3.45米)有关的其它信息已知,能够以绝对比例进行重构。

[0006] 在场景的绝对空间比例无法确定的情况下,SLAM系统可随机分配一个比例,例如,通过从图像空间的像素视差测量值中确定初始关键帧,并且针对两个相应相机位姿之间的基线来假设某通用的实际距离。因此,重构的3D特征在几何模型相关坐标系中具有坐标,当绝对坐标处于真实世界时,该模型相对于绝对坐标的比例因子未知,例如,毫米、厘米、米、或英寸。此外,根据重新获得的几何模型所计算的相机位置也是依据比例而定,见参考文献[4]。

[0007] 比例因子不确定,这为在(例如)机器人系统或车辆的视觉型导航中确定相机处于

绝对比例的实际移动带来挑战,并且在增强现实应用程序中为正确地覆盖相对于真实环境的相机图像的虚拟视觉信息带来挑战。作为一个示例,视觉型导航应用程序能够确定相机运动的形状(例如,相机在圆形路径上移动),但是它无法确定处于绝对比例的平移部分(例如,距离或位置),例如,如果圆形的半径为1米或10米。作为另一示例,考虑增强现实应用程序,该应用程序将一件虚拟家具进行叠加,该家具在环境直播视频上进行了空间配准。如果相机以随机(即,任意)比例在坐标系中进行追踪,则叠加的虚拟家具也将具有任意比例。高2米的虚拟碗柜可能看起来是1米高的桌子的三倍,或者其可能看起来是桌子高度的一半,这取决于重构期间所选择的任意比例。很显然,这不合需要。相反,高2米的虚拟碗柜应当看起来是它旁边高1米的真实桌子的二倍。相机中通过叠加而增强的实物和虚物的比例应当一致。为了实现这一点,需要了解真实环境的几何模型的(正确)绝对比例。

[0008] 而且,在多个真实物体的多个几何模型已经分别通过相同的用于同时追踪多个真实物体的视觉型SLAM系统创建出来的情形中,如参考文献[8],比例因子不确定的问题相当明显。通常来说,针对多个几何模型中的每个模型应用随机比例值。如果SLAM系统在多个几何模型中转换,比例可能发生改变,因此,严重影响了像增强现实等计算机视觉应用程序中的用户体验。

[0009] 已经提出了各种方法用于确定正确的比例因子,该比例因子可定义其处于真实世界时真实环境的重构几何模型的实际大小。

[0010] 例如,参考文献[1]中Davison等人提出,将绝对空间尺寸已知的校准物体引入SLAM系统的绝对比例确定场景中。由此,他们需要改变场景外观,因为他们使用相同的相机来捕获校准物体以及捕获用以重构SLAM的场景。用户同样需要获得校准物体。

[0011] 参考文献[5]中Lemaire等人提出,使用立体摄像系统(即,两个通过重叠的相机视锥进行位移的相机)来解决SLAM系统中绝对比例的确定问题。然而,使用立体相机仅解决了部分问题,这是因为两个相机之间的位移相对于环境距离必须很显著,以便可靠地计算出环境深度。也需要了解两个相机之间处于绝对比例的位移,即,以诸如毫米、厘米、米或英寸等为单位的位移。

[0012] 参考文献[14]也公开了以下方案:通过具有重叠相机视锥的多相机装置来估算绝对比例。然而,两个相机之间的位移相对于环境距离必须很显著,以便可靠地计算出环境深度。

[0013] 在参考文献[6]中Lieberknecht等人通过采用RGB-D相机将深度信息集成到单眼视觉型SLAM中,以实现准确缩放的几何模型的重构,RGB-D相机可提供与图像像素相关的绝对深度信息。可根据已知的处于绝对比例的深度信息来确定绝对比例。然而,与常见的单眼RGB相机相比,RGB-D相机设备在手持式设备(例如,移动手机、平板电脑、或PDA)中并不常用。同样地,如果存在大量红外环境光线,和白天时室外环境一样,基于将红外线投射到场景中的主动式立体深度相机无法可靠地运行。

[0014] 参考文献[7]中Klein等人通过以下方式解决了比例估算的问题:当相机捕获3D三角测量所需的两张图像时,手动定义相机两个位置之间的基线(即,处于绝对比例的距离),3D三角测量用于进行环境重构。

[0015] 也可使用传感器与惯性测量单元(IMU)的融合件来估算绝对比例,如参考文献[9]所公开。该方案的一个问题是:传感器值不准确导致比例估算值也不准确。昂贵(即,计算密

集型)技术,像“卡尔曼滤波”或“光束平差”等,用于解决该问题,但是与现成装置(诸如,移动手机)集成的IMU的精确度通常不足以准确地估算绝对比例。

[0016] 因此,需要提供一种方法和系统,用于确定实物的至少一部分的处于绝对空间比例的3D重构件的空间坐标,该方法和系统能够重构处于绝对比例的真实物体,或者确定将处于任意比例的重构件的坐标映射成绝对比例的比例因子。

## 发明内容

[0017] 根据第一方面,本发明公开了一种确定实物的至少一部分的处于绝对空间比例的3D重构件的空间坐标的方法,包括以下步骤:

[0018] a) 接收第一图像的图像信息,第一图像包括利用第一相机捕获的第一实物的至少一部分

[0019] b) 接收第三图像的图像信息,第三图像包括利用第三相机捕获的第一实物的该至少一部分,

[0020] c) 接收第二图像的图像信息,第二图像包括利用第二相机捕获的第二实物的至少第一部分,

[0021] d) 接收第四图像的图像信息,第四图像包括利用第四相机捕获的第二实物的至少第二部分,

[0022] 其中第一相机的视锥和第二相机的视锥不重叠,并且其中第三相机的视锥和第四相机的视锥不重叠,

[0023] e) 提供第一相机与第二相机之间的第一空间变换,并且提供第三相机与第四相机之间的第二空间变换,

[0024] f) 提供第一比例信息和第二比例信息,第一比例信息指示第二实物的至少第一部分的绝对空间比例,第二比例信息指示第二实物的至少第二部分的绝对空间比例,

[0025] g) 根据第二图像、第一比例信息、第四图像和第二比例信息,确定第二相机的至少一部分的位姿以及第四相机的至少一部分的位姿,其中第二相机的至少一部分的位姿和第四相机的至少一部分的位姿被限定在第二共同坐标系中,

[0026] h) 根据第二相机的至少一部分的位姿和第一空间变换,确定第一相机的至少一部分的位姿,根据第四相机的至少一部分的位姿和第二空间变换,确定第三相机的至少一部分的位姿,其中第一相机的至少一部分的位姿和第三相机的至少一部分的位姿被限定在第二共同坐标系中,

[0027] i) 根据第一图像、第三图像和第一相机的至少一部分的位姿以及第三相机的至少一部分的位姿,确定第一实物的至少一部分的处于绝对空间比例的3D重构件的空间坐标。

[0028] 根据一个实施方案,第二相机的至少一部分的位姿、第四相机的至少一部分的位姿、第一相机的至少一部分的位姿和第三相机的至少一部分的位姿均包括平移信息。

[0029] 根据一个实施方案,步骤i)包括:根据第一图像和第三图像,确定第一相机的位姿和第三相机的位姿,其中第一相机的位姿和第三图像的位姿被限定在第一共同坐标系中;根据第一共同坐标系中第一相机的位姿的平移信息和第三相机的位姿的平移信息,确定指示至少一个距离的第一平移信息;根据第二共同坐标系中第一相机的至少一部分的位姿和第三相机的至少一部分的位姿,确定指示至少一个距离的第二平移信息;根据第一平移信

息和第二平移信息,确定比例因子;根据第一图像、第三图像和第一相机的位姿以及第三相机的位姿,确定第一共同坐标系中第一实物的至少一部分的处于任意空间比例的3D重构件的空间坐标;根据比例因子,将所确定的3D重构件的空间坐标从第一共同坐标系的任意空间比例变换为第二共同坐标系的绝对空间比例。

[0030] 根据一个实施方案,步骤i)进一步包括:根据第一图像和第三图像,确定第一相机的位姿和第三相机的位姿,其中第一相机的位姿和第三相机的位姿被限定在第一共同坐标系中;提供被限定在第二共同坐标系中的第一相机的至少一部分的位姿和第三相机的至少一部分的位姿;根据第二共同坐标系中第一相机的至少一部分的位姿以及第三相机的至少一部分的位姿,更新第一共同坐标系中与第一相机的位姿相关的平移信息以及与第三相机的位姿相关的平移信息;根据第一图像、第三图像、第一相机的更新的位姿和第三相机的更新的位姿,确定第一实物的至少一部分的处于绝对空间比例的3D重构件的空间坐标。

[0031] 例如,第二相机的至少一部分的位姿包括3DoF平移信息和3DoF旋转信息,第四相机的至少一部分的位姿包括3DoF平移信息和3DoF旋转信息,第一相机的至少一部分的位姿包括3DoF平移信息和3DoF旋转信息,第三相机的至少一部分的位姿包括3DoF平移信息和3DoF旋转信息。

[0032] 根据第二方面,本发明公开了一种方法,用于确定第一实物的至少一部分的处于绝对空间比例的3D重构件的空间坐标,包括以下步骤:

[0033] a) 接收第一图像的图像信息,第一图像包括利用第一相机捕获的第一实物的至少一部分,

[0034] b) 接收第三图像的图像信息,第三图像包括利用第三相机捕获的第一实物的至少一部分,

[0035] c) 接收第二图像的图像信息,第二图像包括利用第二相机捕获的第二实物的至少一部分,

[0036] d) 接收第四图像的图像信息,第四图像包括利用第四相机捕获的第二实物的至少第二部分,

[0037] 其中第一相机的视锥和第二相机的视锥不重叠,并且其中第三相机的视锥和第四相机的视锥不重叠,

[0038] e) 提供第一比例信息和第二比例信息,第一比例信息指示第二实物的至少第一部分的绝对空间比例,第二比例信息指示第二实物的至少第二部分的绝对空间比例,

[0039] f) 根据第二图像、第一比例信息、第四图像和第二比例信息,确定第二相机的至少一部分的位姿以及第四相机的至少一部分的位姿,其中第二相机的至少一部分的位姿和第四相机的至少一部分的位姿被限定在第二共同坐标系中,其中第二相机的至少一部分的位姿包括平移信息,第四相机的至少一部分的位姿包括平移信息,

[0040] g) 根据第一图像和第三图像,确定第一相机的位姿和第三相机的位姿,第一相机的位姿和第三相机的位姿被限定在第一共同坐标系中,

[0041] i) 该方法还包括

[0042] i0) 根据第二相机的至少一部分的位姿以及第四相机的至少一部分的位姿,确定指示至少一个距离的第二平移信息,

[0043] i1) 根据第一相机位姿的平移以及第三相机位姿的平移,确定指示至少一个距离

的第一平移信息，

[0044] i2) 根据第一平移信息和第二平移信息，确定比例因子，

[0045] i3) 根据第一图像、第三图像、第一相机的位姿和第三相机的位姿，确定第一实物的至少一部分的处于任意比例的3D重构件的空间坐标，

[0046] i4) 根据比例因子，将所确定的3D重构件的空间坐标从任意空间比例变换为绝对空间比例，

[0047] 或者

[0048] k) 替代步骤i0)-i4)，该方法进一步包括

[0049] k1) 根据第二相机的至少一部分的位姿以及第四相机的至少一部分的位姿，将与第一相机的位姿相关的平移信息以及与第三相机的位姿相关的平移信息更新为处于绝对空间比例下，

[0050] k2) 根据第一图像、第三图像、第一相机的更新的位姿、和第三相机的更新的位姿，确定第一实物的至少一部分的处于绝对空间比例的3D重构件的空间坐标。

[0051] 以下实施方案同样地可以根据第一方面或第二方面来实施。

[0052] 根据一个实施方案，第二共同坐标系与第二实物相关，步骤g) (第一方面) 或步骤f) (第二方面) 分别包括：根据第二图像和第一比例信息，确定第二相机的至少一部分的位姿；根据第四图像和第二比例信息，确定第四相机的至少一部分的位姿。

[0053] 根据一个实施方案，第二共同坐标系与所选相机相关，所选相机为第二相机和第四相机中一者，其中第二相机和第四相机中另一者为未选相机，步骤g) (第一方面) 或步骤f) (第二方面) 分别包括：提供第二共同坐标系中与所选相机相关的位姿；根据第二图像、第一比例信息、第四图像、和第二比例信息，确定与未选相机相关的位姿。

[0054] 根据一个实施方案，第一共同坐标系与第一实物或第一相机与第三相机中一者相关。

[0055] 根据一个实施方案，第二实物的至少第一部分和第二实物的至少第二部分是相同部分，第一比例信息和第二比例信息相同，或者第二实物的至少第一部分和第二实物的至少第二部分是不同部分，第一比例信息和第二比例信息不同。

[0056] 根据一个实施方案，第一相机和第二相机面朝相反的方向，并且第三相机和第四相机面朝相反的方向。

[0057] 根据一个实施方案，第二实物为人脸，步骤g) (第一方面) 或步骤f) (第二方面) 分别通过脸部追踪方法来实施，其中第一比例信息和/或第二比例信息被提供为至少两个脸部特征的位置之间的处于绝对比例的至少一个距离。

[0058] 根据一个实施方案，第二实物为人脸，步骤g) (第一方面) 或步骤f) (第二方面) 分别通过脸部追踪方法来实施，并且其中第一比例信息和/或第二比例信息被提供为第二图像和第四图像中的至少一者中的第二实物的至少一部分的位置的处于绝对比例的至少一种深度信息。

[0059] 根据一个实施方案，第二相机和第四相机中至少一者提供处于绝对比例的深度信息，步骤g) (第一方面) 或f) (第二方面) 是分别以绝对比例利用视觉测程法来实施的，第一比例信息和/或第二比例信息是通过深度信息来提供的。

[0060] 根据一个实施方案，第一相机和第三相机是处于不同时间点的相同相机，并且第



二相机和第四相机是处于不同时间点的相同相机。

[0061] 根据一个实施方案,第二实物为人脸。

[0062] 根据一个实施方案,第一比例信息和/或第二比例信息包括瞳孔间距。

[0063] 根据一个实施方案,该方法还包括:提供第二相机的至少一部分的固有相机参数以及第四相机的至少一部分的固有相机参数。

[0064] 优选地,携带第一相机至第四相机的用户是静止的。例如,该方法还包括确定用户何时是静止的步骤。

[0065] 根据一个实施方案,第二相机和第四相机均包括红外相机。

[0066] 根据一个实施方案,比例因子根据一组位姿对N来确定,其中 $N > 2$ 。

[0067] 根据一个实施方案,该方法还使用关于第二实物的通用人脸模型。

[0068] 根据一个实施方案,该方法还使用关于第二实物的基于概率分布的人脸模型。

[0069] 根据一个实施方案,该方法还使用关于第二实物的被校准的、具体地自适应的、重构的或手动测量的人脸模型。

[0070] 根据一个实施方案,该方法还包括脸部识别或分类以选择与第二物体相关的模型。

[0071] 根据一个实施方案,该方法还包括:检测用于开始所述方法的用户输入,并向所述用户提供利用所述第二相机来执行某运动的指令,所述运动是基于利用所述第二相机捕获的图像进行测量的;接收用户交互,所述用户交互触发所述3D重构件的空间坐标的确定。

[0072] 根据一个实施方案,在捕获第一图像、第二图像、第三图像和第四图像时将第二实物视作相对于第一实物静止。

[0073] 根据一个实施方案,第一图像、第二图像、第三图像和第四图像是基于确定在捕获所述第一图像、所述第二图像、所述第三图像和所述第四图像的同时所述第二实物相对于所述第一实物静止的方法来选择的。

[0074] 根据一个实施方案,根据第二相机的至少一部分的位姿以及第四相机的至少一部分的位姿来确定指示至少一个距离的第二平移信息仅考虑第二实物与第二相机之间的距离以及第二实物与第四相机之间的距离。

[0075] 根据一个实施方案,对多组第一图像、第二图像、第三图像和第四图像重复执行第二方面中的步骤a)至i2),其中不同图像组能够或能够不相互重叠,从而导致步骤i2)中的多个比例因子,该方法还包括从多个比例因子中确定单个比例因子并在步骤i4)使用单个比例因子来变换空间坐标。

[0076] 根据一个实施方案,对多组第一图像、第二图像、第三图像和第四图像重复执行第二方面中的步骤a)至i1),其中不同图像组能够或能够不相互重叠,从而导致步骤i1)中的多种第一平移信息以及步骤i0)中产生多种第二平移信息,其中步骤i2)根据步骤i1)中多种第一平移信息以及步骤i0)中多种第二平移信息来确定比例因子。

[0077] 根据一个实施方案,根据通用脸部模型,例如,基于概率分布的脸部模型,将所述第一比例信息和/或所述第二比例信息提供为至少两个脸部特征的位置之间的处于对比例的至少一个距离。

[0078] 根据一个实施方案,根据对特定脸部的至少两个脸部特征的位置之间的处于绝对比例的至少一个距离的之前的校准,例如通过自适应拟合、重构、手动测量,将所述第一比

例信息和/或所述第二比例信息提供为至少两个脸部特征的位置之间的处于绝对比例的至少一个距离。

[0079] 根据一个实施方案,根据基于视觉脸部识别所选择的模型,将所述第一比例信息和/或所述第二比例信息提供为至少两个脸部特征的位置之间的处于绝对比例的至少一个距离。

[0080] 根据一个实施方案,根据基于视觉脸部分类所选的用于从针对不同脸部类别的通用脸部模型的数据集确定脸部特性诸如年龄、性别、种族、体重、或身高的模型,将所述第一比例信息和/或所述第二比例信息提供为至少两个脸部特征的位置之间的处于绝对比例的至少一个距离。

[0081] 根据一个实施方案,所述第一实物的处于绝对比例的所述3D重构用于确定处于绝对比例的相机位姿。例如,处于绝对比例的相机位姿用于叠加增强现实应用程序中的数字化信息。

[0082] 根据一个实施方案,该方法由用户输入来触发。根据另一实施方案,该方法自动触发。

[0083] 根据另一方面,本发明公开了一种用于确定第一实物的至少一部分的处于绝对空间比例来3D重构件的空间坐标的系统,包括处理系统,用于执行上文所公开的多个方面和实施方案中所列的步骤。

[0084] 根据优选实施方案,当具有可捕获人脸和实物的捕获装置时,我们由此使用脸部来确定处于绝对比例的距离,这可用于实现本发明的目的。捕获设备可以是单个相机或一组刚性连接的相机,例如,常用的移动手机中的单个相机或一组刚性连接的相机。对于具有前置相机和后置相机的此类移动手机,前置相机可捕获用户脸部,而后置相机捕获(第一)实物。

[0085] 一般来说,本发明涉及如何确定第一实物的处于绝对比例的3D重构件的空间坐标的问题,该问题可通过以下方式解决:从处于绝对比例的第二实物至少两个点的空间坐标已知的至少两个视角中每个视角,用成像装置观察第一实物和第二实物,这实现了对至少两个视角之间处于绝对空间比例的距离的确定,并且实现了对用于第一实物进行空间坐标重构的绝对空间比例的确定。

[0086] 此外,发明者发现:在用后置相机捕获物体或环境的同时,可通过前置相机来捕获用户脸部。其还包括:意识到人脸的绝对比例变化有限,因此为以绝对比例进行测量提供了良好的比例限制条件。如果特定用户脸部的空间特性未知,可使用通用人脸模型来确定绝对空间比例。由此引入了误差,该误差取决于形成估算值的人脸空间特性的变化程度。即使所呈现的空间特性不同于实际特性,这从某种程度上而言形成了更加不准确的绝对比例估算值,多个比例估算值使用相同的特性并且相同的脸部将总是形成相似的绝对比例。由此,所估算的比例可能不精确但是是准确的,即,一致的。如果使用经过校准(即,测量)的特定脸部,则该脸部的比例将以绝对比例被精确掌握。

[0087] 此外,已经发现,场景中不需要任何另外的已知物体。通过采用现代化手持装置的前置相机和后置相机,提供了绝对比例相关信息的人脸不需要是前置相机所捕获与重构的场景部分的一部分,并且相反可以在用户操作应用程序的同时进行捕获。作为结果,用户脸部不会成为场景重构的一部分,正好与将标识物或已知物体添加到场景中,然后标识物或

已知物体作为场景的一部分进行重构的方案相反。

[0088] 本发明所述的方法可以,例如,结合视觉型同步定位与地图构建(SLAM)来使用,诸如参考文献[1]所公开的SLAM,这是一项众所周知的用于在不需要对环境有任何事先了解的情况下通过一个或多个相机来创建真实环境几何模型的技术。相同技术的另一常见术语为运动恢复结构(SfM)。具有至少深度信息的几何模型也称为真实环境的3D图。环境几何模型的创建也称作真实环境的(3D)重构。所创建的(通常称作重构的)几何结构可以由多个3D特征来表示,诸如,点特征或边特征。3D特征说明了真实环境的物理3D特征(也称为结构)。真实环境也可称作真实场景、实物,或者可以理解为包括一个或多个实物。

[0089] 重构的几何模型可用于不同应用程序中。例如,可虚拟化呈现在显示屏上,或者用三维打印机打印出来。也可作为测量实物上多点之间的空间距离的基础。

[0090] 重构的几何模型也可用于,根据相机所捕获的当前实物图像,确定相机相对于实物的位姿(即,位置和/或取向)。通过将所提取的当前相机图像的2D特征与几何模型中存在的3D特征进行匹配,例如,通过本地图像特征描述符(参考文献[20]),可建立起多个2D-3D对应关系。然后可根据这些对应关系来计算几何模型坐标系中的相机位置和取向。该过程称为相机位姿评估,有时也称为相机追踪。相对于实物对相机进行追踪的问题也可表述为相对于相机追踪实物的问题。如果解决了两个问题中一个问题,第二个问题的解决方案就是第一个问题的解决方案的反面。因此,当谈论总体概念时,相机追踪和物体追踪可互换使用。

[0091] 视觉型SLAM同时执行相机追踪和环境重构。视觉型SLAM可促成许多应用,诸如,机器人系统或车辆的视觉型导航。具体地,视觉型SLAM是未知真实环境中支持增强现实(AR)系统或应用程序(见参考文献[3])的一项前景广阔的技术。

[0092] 本发明的目的是确定重构的真实环境几何模型的绝对比例,使得几何模型坐标系处于绝对比例下,即存在已知的比例因子,该比例因子将重构几何模型的坐标系单位映射为实际中的绝对空间单位。例如,可以将模型按比例缩小至毫米,使得模型的模型单位对应于实物的毫米。在此情况下,如果模型中两点间隔56个单位,则实物上的对应点彼此相距56mm。可实施以绝对比例定义模型,使得模型坐标系中的单位对应于任何实际距离(例如,12.34米),只要该距离已知。

[0093] 因此,可确定至少一个比例因子,可用于按比例缩放坐标系中定义的多个3D特征,坐标系与描述实物的几何模型相关。

[0094] 在一个实施方案中,本发明确定比例因子,该比例因子将以任意比例定义的现有实物模型的坐标按比例缩放至以绝对比例定义的坐标系。

[0095] 在另一实施方案中,本发明确定至少两个相机之间处于绝对比例的距离,这一点则实现了:根据两个相机的图像,以绝对比例重构相机中可见实物的3D模型。

[0096] 我们可确定空间坐标系之间的比例因子,在空间坐标系中,容貌的空间特性按实际度量单位(例如,厘米)进行定义。如果定义了特征位置的坐标系已经给出,比例因子可用于将坐标系(和特征的空间坐标分别)按比例进行缩放以相对于实际度量单位(像1单位=1mm)具有一对一的比例。如果定义了特征位置的坐标系还没有确定,我们可使用比例因子直接将坐标系(和特征的空间坐标分别)进行初始化以相对于实际度量单位(像1单位=1mm)具有一对一的比例。

[0097] 此外,坐标系也可维持原状。对于要求绝对比例(即,表示实际比例)的操作而言,所确定的空间比例因子则可用于提取实际空间距离,或者将位姿或特征的空间坐标变换为相对于实际度量单位具有一对一比例的坐标系。

[0098] 在另一实施方案中,为第一相机的两张图像提供了两个处于绝对比例的位姿,例如,两个时间点上的后置相机,该相机用于通过三角测量进行场景结构重构。可根据面朝用户的不同相机所捕获的图像,通过脸部追踪法来提供处于绝对比例的两个位姿,在下文中称为前置相机,具有相对于第一相机的已知空间变换。

[0099] 在许多可能情景中,用户手持配备了后置相机的移动设备。前置相机可用于捕获周围环境图像,例如,用以将周围环境图像与增强现实应用程序中的虚物进行叠加或用于其它图像处理应用程序中。在此类应用程序中,常常要求重构周围环境中的实物以及/或者估算相机的位姿或相机相对于实物或环境的运动。

[0100] 在诸如下图6所示的我们具有另外的前置(也称为面朝用户的)相机的情景中,我们可使用用户脸部图像,该用户操作应用程序并因此已经出现在真实环境中,用于估算根据面朝世界的相机所捕获的图像所创建的图(即,3D重构图)的绝对比例。

[0101] 这具有多种优势:可获得用户的脸部,因此无需添加任何另外的几何结构或物体,用户脸部由面朝用户的相机进行捕获,因此不需要任何繁琐装置。由于用户通常为了体验该应用程序而看着屏幕,面朝用户的相机总是可以捕获用户脸部,而面朝世界的相机可捕获真实环境的视图。由于只要用户面朝或看着手持设备的显示屏就总是可获得用户脸部,可支持自动更新或重新进行比例估算。人脸的几何结构也仅限于几何结构的变化范围内,由此可实现对多数人的脸部特征的尺寸和比例进行有效呈现与限制。即,仅仅通过用户脸部和捕获设备,比例估算可由任何人完成,无需另外的已知物体。特定用户也可对其脸部进行特殊校准,从而实现了更高的精度。也可并入可识别众多人的脸部识别程序,例如,参考文献[19]中的脸部识别程序,用以识别出现在面朝用户的相机的图像中的用户,然后从预设置数据库中选择用户脸部相应的正确的绝对尺寸。通过网络连接可访问,脸部识别程序可在本地运行或远程执行。通过网络连接可访问,可本地或远程提供包含正确的用户脸部绝对尺寸的预设置数据库。

[0102] 另一实施方案使用虚拟脸部分类法,例如,如参考文献[12]所公开,用以确定诸如人的年龄、性别、种族、体重、身高等特性,然后使用针对特定人体类别的通用脸部模型。

[0103] 可结合以下假设来使用两个相机(例如,面朝世界的相机和面朝用户的相机):两个相机的坐标系之间的空间关系已知,例如,刚体变换。面朝世界的相机可用于确定:在与真实环境相关的坐标系中,面朝世界的相机相对于真实环境的位姿;和/或与位于真实环境中的实物相关的物体坐标系。这将实现所需的虚体与实物之间的校准,虚体可叠加在相机图像上,实物为真实环境中由面朝世界的相机所捕获的图像中的实物。假设两个相机的坐标系之间的变换已知,面朝用户的相机所捕获的用户脸部的绝对比例信息可变换为真实环境坐标系和/或物体坐标系。这将通过面朝世界的相机实现绝对比例下的SLAM重构。

[0104] 对于包括面朝用户的相机和面朝世界的刚性连接相机的相机装置的每个视角而言,我们可确定由两个位姿组成的对:面朝用户的相机相对于用户脸部处于绝对空间比例的位姿、以及面朝世界的相机相对于第一实物处于任意比例的位姿。给出面朝用户的相机与面朝世界的相机之间的空间变换,我们可以通过以下方式确定面朝世界的相机相对于用

户脸部处于绝对空间比例的位姿：用面朝用户的相机F与面朝世界的相机之间的空间变换，变换面朝用户的相机相对于用户脸部处于绝对比例的位姿。

[0105] 给出从双相机装置的两个不同视角获得的两个此类变换位姿，我们可确定面朝世界的相机的两个处于绝对比例的位姿之间的平移距离 $D_{abs}$ 。通过面朝世界的相机相对于以任意比例定义的第一实物的两个位姿，我们可以确定这两个位姿之间处于任意比例的平移距离 $D_{arb}$ 。

[0106] 最后，从相对于实物的坐标系任意比例变换到绝对比例的比例因子可确定为 $D_{abs}$ 与 $D_{arb}$ 之比。 $S=D_{abs}/D_{arb}$ 。

[0107] 不参见附图，通过以下方式确定第一相机WC(捕获至少一部分实际物体的图像用于进行SLAM重构)的两个相机位姿PW1和PW2之间的绝对空间距离，第一相机WC属于捕获装置C：通过相对于至少两种图像的脸部进行图像型相机位姿估算，观察属于相同的步骤装置C的第二相机FC(捕获至少一部分人脸的图像)的空间平移和旋转，其中至少一张脸部图像I(F1)在相机FC的相机位姿PF1处进行拍摄，即，相机WC在此时间点处于相机位姿PW1，另一脸部图像I(F2)在相机FC的相机位姿PF2处进行拍摄，即，相机WC在此时间点处于相机位姿PW2。至少一部分实际物体可以由相机WC分别在相机位姿PW1和PW2处的图像I(W1)和I(W2)中捕获到。图像I(W1)和I(W2)可用于实物重构或相机位姿估算应用程序中(例如，SLAM)。对于以绝对比例确定实物的空间特性的方法，两个相机位姿PW1与PW2之间的绝对空间距离应当不为零。

[0108] 在一个实施方案中，比例估算不仅根据两对相应的位姿PW和PF(即，PW1和PF1、以及PW2和PF2)进行，这些位姿通过在这些位姿处所捕获的相应四张图像进行确定(基本如图2所示)，也根据多对位姿PW和PF进行，每对位姿通过所捕获的图像进行确定。可通过(例如)模型拟合法，诸如中值、平均值或RANSAC来合并均基于两对位姿( $W_i$ 和 $F_i$ 、以及 $W_j$ 和 $F_j$ )的多个比例估算值。例如根据位姿之间的最小距离或位姿测量值的不确定性评分与质量评分，模型拟合法可另外地考虑某些位姿对是否适合进行比例估算。面朝用户的相机旋转时 $F_i$ 至 $F_j$ 的差别与面朝世界的相机旋转时 $W_i$ 至 $W_j$ 的差别之间的一致性也可用作(例如)位姿测量值的不确定性评分和/或质量评分以及作为第二实物是否相对于第一实物移动的指标。面朝用户的相机的两个位姿之间的旋转部分也可用于确定这些位姿是否适合于进行比例估算。忽略面朝用户的相机和面朝世界的相机之间的真实变换并且假设它们的原点相同，面朝用户的相机的两个位姿之间的旋转部分可引入比例估算误差(见图8)，因此优选仅具有可忽略的/很小的旋转度。也可使用两条完整的位姿轨迹(面朝用户的相机一条轨迹，面朝世界的相机一条)来估算脸部在捕获期间没有移动的可能性。这可(例如)通过如Umeyama公开的以下方法(参考文献[10])进行估算：校正这两条轨迹，计算配准之后的剩余误差。如果剩余误差在特定阈值之上，这可能表示头部相对于实物移动了。在此情况下，可丢弃所确定的比例因子，并且可重新开始校准。也可使用面朝用户的相机的旋转率与面朝世界的相机的旋转率之间的一致性来支持校正两条轨迹，并且在计算剩余误差时考虑旋转率的一致性。

[0109] 现代化手持移动设备，诸如，移动手机、平板电脑、或平板计算机可配备有两个相机(例如面朝用户的相机和面朝世界的相机)，指向两个相对的方向。移动设备的显示屏通常面朝与面朝用户的相机相同的方向。

[0110] 本发明的一个可能的实施方案为：通过面朝用户的相机所捕获的用户脸部图像来估算绝对比例。然后将绝对比例用于重构以及以绝对比例通过另一相机（指向与面朝用户的相机相反方向且通常具有相对于面朝用户的相机已知的空间变换的面朝世界的相机）进行实物追踪。可以将这两个相机附接到手持设备或移动设备上，例如，移动手机、平板电脑、或平板计算机。此外，显示设备，例如，LCD屏，可以附接到移动设备上。

[0111] 移动设备的两个相机可具有固定的空间关系，例如，刚体变换，这可通过校准程序进行确定，例如，手眼校准，通过使用至少一个已知标识物或另外的追踪系统。

[0112] 常用的先进方案要求特殊相机（具有基于主动式立体或被动式立体或渡越时间的深度传感器）或或另外的装置，以便估算实物的绝对比例。这绝对限制了这些方案的适用性。

[0113] 真实场景SLAM图的绝对比例的一个估算方案为：直接在用作SLAM方法输入的相机图像中检测已知物体（如参考文献[1]所述）；使用已知物体的已知绝对比例来推断出该图的绝对比例。该方案的一个问题是：需要获得已知物体以及另外的设置步骤，其中在场景中添加另外的已知物体。这改变了原有场景并且要求相机朝向已知物体。

[0114] 与使用像平面标识物等必须放到室内明显之处并由SLAM相机捕获用于进行比例估算的特殊物体的最先进技术（像参考文献[1]所用的技术）相比，用户脸部具有极大优势：不需要特别注意在整个重构过程中将物体保持在面朝用户的相机的视野范围内。

[0115] 与在待重构场景中添加具有已知绝对空间特性的物体的方案（参考文献[1]）相比，本发明的另一明显不同在于，本发明不依赖于用相同相机捕获已知物体用于重构实物或场景，而是相反地使用第二相机来捕获脸部。由此，脸部不成为重构件的一部分，这与向场景中添加实物相反。

[0116] 与诸如参考文献[1]中在场景中添加已知物体并由此需要配备有相机的计算机、用户、待重构实物、和另外的用于进行校准的用户将需要随身携带的特殊已知物体的方案相反，本发明仅需要配备有相机的计算机、用户、和待重构实物。

[0117] 根据本发明的多个实施方案，一个明显的确定SLAM重构件的绝对比例的优势为：明确使用两个或多个脸部特征或基准点之间的绝对距离（例如，脸部两只眼睛之间的距离、或脸部上眼睛与嘴巴之间的距离、或左眼角与右眼角之间的距离）。当用户观察显示设备时，这些距离可以在面朝用户的相机（即，指向用户的相机、或指向由一个或多个用于捕获用户脸部的光学仪器（例如镜子或光学透镜）所反射的用户图像的相机）所捕获的用户图像中进行识别。这实现了关于人脸及其绝对空间特性的先验知识的应用。脸部图像通常由面朝用户的相机进行捕获，并且用于估算绝对比例，而不是依赖场景中几何图形已知的另外物体进行比例估算。通过明确使用在用户观察显示屏的整个期间通常或总是可获得的脸部（例如，使用脸部特定特点），可随时进行比例估算而不会对场景产生任何影响。此外，脸部在所有人之间的几何结构的变化范围有限，通过聚焦脸部，可采用专业算法用于通过用户脸部进行比例估算。尤其适合于进行比例估算的脸部区域可进行预习得和/或预定义。可通过已经建立的u位姿追踪算法在实时追踪时配准脸部区域。可考虑会对比例估算产生不良影响的脸部区域，例如，由于不同的人这些区域在大小和形状方面具有明显不同，并且将其从比例估算中排除（像鼻子大小或耳朵大小）。

[0118] 在一个实施方案中，移动设备的显示设备的法线和面朝用户的相机的光轴优选具

有相同的方向。在此情况下,当用户观察显示设备上的视觉化信息(例如,增强场景)时,面朝用户的相机通常或总是可以捕获到用户脸部。因此,总是可以根据脸部图像来估算绝对比例。

[0119] 例如,根据本发明的处理系统至少部分包括在移动设备(诸如,移动手机、可佩戴式计算机、平板计算机、移动计算机,通常称作笔记本电脑、或头戴式显示器,诸如用于光学透视式增强现实应用程序)中和/或经适配以与移动设备进行通信的服务器计算机中。处理系统可仅仅包括在这些装置中一种装置内,例如,在移动设备中或在服务器计算机中,或者可以是分布式系统,其中一个或多个处理任务由处理系统中一个或多个处理设备进行分配与处理,这些处理装置分散各处,并通过点到点通信方式或通过网络进行通信。

[0120] 根据一个实施方案,系统包括移动设备,移动设备包括一个或多个相机以及,例如,显示屏。

[0121] 本文所述的与该方法有关的任何步骤、实施方案、方面和实施例可以同样地或类似地由用于执行相应步骤、实施方案、方面或实施例的处理系统(软件和/或硬件)来实施。处理系统内所用的任何处理设备可按照这种方式通过软件和/或硬件进行配置,并且可通过通信网络,例如,通过服务器计算机或点到点通信方式与一个或多个相机、显示器和/或任何其它部件进行通信。

[0122] 根据另一方面,本发明还涉及包括软件代码段的计算机程序产品,所述软件代码段被适配为执行根据本发明所述的方法。具体地,软件代码段包含在非暂态计算机可读介质上。软件代码段可被加载到本文所述的一个或多个处理设备(诸如,微处理器)的存储器中。任何所使用的处理设备(诸如,一个或多个微处理器)可经由通信网络例如经由本文所述的服务器计算机或点到点通信方式进行通信。

## 附图说明

[0123] 现将相对于附图来描述本发明的方面和实施方案,其中:

[0124] 图1示出了根据本发明的实施方案的方法的流程图,

[0125] 图2示出了本发明的一个可能的实施方案以及使用单眼SLAM或SfM时比例模糊的问题,

[0126] 图3示出了本发明的另一实施方案,

[0127] 图4示出了根据本发明的实施方案的所涉及的坐标系和变换。

[0128] 图5示出了根据本发明的实施方案的捕获装置,捕获装置包括处于不同位姿上的面朝用户的相机和面朝世界的刚性附接相机,

[0129] 图6示出了本发明的用手持设备来实施的示例性实施方案,

[0130] 图7示出了指导用户进行示例性比例估算过程的图形用户界面的实施例。

[0131] 图8示出了第一相机与第二相机之间的空间变换的影响,

[0132] 图9示出了本发明的另一实施方案,

[0133] 图10示出了本发明的另一实施方案,

[0134] 图11示出了本发明的另一实施方案,

[0135] 图12示出了本发明的另一实施方案。

## 具体实施方式

[0136] 根据单眼捕获设备所捕获到的一组图像来确定实物结构,这种方式形成了依据比例而定的空间(或几何)结构的重构件。这意味着重构使用对其而言绝对空间单位(诸如,单位米)未知的空间单位。在许多应用中,需要获得具有绝对单位的重构件,也称作具有“绝对比例”的重构件。为此,可使用关于处于绝对比例的至少一个距离的知识:拍摄用于进行重构的相应图像时实物多个部分之间的距离、或相机相对于实物的位置之间的距离。处于绝对比例的这一距离可以(例如)是具体某人的眼睛距离或通用眼睛距离或脸部基准点的任何其它空间特性。当具有捕获脸部和实物的捕获装置时,本文所公开的实施方案使用脸部来确定处于绝对比例的距离。捕获设备可以是单个相机或一组刚性连接的相机,例如,手机中的单个相机或一组刚性连接的相机。前置相机通常捕获用户脸部,而后置相机捕获实物。

[0137] 有利地,本发明利用用户脸部(用户脸部并不一致但多数人具有相似的特性),这一点可优选地用于手持增强现实程序中。

[0138] 本发明实现了以绝对空间比例(以下也简称为绝对比例)重构实物或环境的结构。这(例如)实现了以绝对比例进行相机位姿追踪,相机位姿追踪实现了虚物叠加,这些虚物以绝对比例进行定义以与重构的实物保持一致的比例。相机位姿追踪也实现了真实空间测量,从而实现了基于绝对空间比例的计算过程,像物理模拟(例如,重力加速度)或虚物与实物之间的冲突检测(例如,物体在空间上是否符合实际)。

[0139] 代替使用需要添加到场景中的已知物体(即,标识物),本发明的实施方案使用总是存在的用户脸部。与其它方案相反,本发明不需要任何用户输入,本发明不需要惯性传感器,提供比消费级惯性传感器更精确的结果。

[0140] 图1示出了根据本发明的实施方案的方法的流程图。在第一步骤S1中,提供了:包括利用第一相机捕获的第一实物的至少一部分的第一图像、包括利用第二相机捕获的第二实物的至少一部分的第二图像、包括利用第三相机捕获的第一实物的至少一部分的第三图像、以及包括利用第四相机捕获的第二实物的第四图像。在第二步骤S2中,提供了:第一相机和第二相机之间处于绝对比例的空间变换、以及第三相机与第四相机之间处于绝对比例的空间变换。在第三步骤S3中,提供第二实物的至少一部分的绝对空间比例。在第四步骤S4中,根据第二图像以及第二实物的至少一部分的绝对比例信息,确定第二相机的处于绝对比例的位姿的至少一部分,根据第四图像以及第二实物的至少一部分的绝对比例信息,确定第四相机的至少一部分处于绝对比例的位姿。

[0141] 在另一实施方案中,也可能没有确定与第二实物(例如,人脸)有关的两个完全位姿,即捕获第二图像的第二相机位姿以及捕获第四图像的第四相机位姿,但是仅仅确定两个位姿之间的差别,换句话说,仅仅确定第四相机相对于第二相机的位姿。

[0142] 在另一实施方案中,也可能仅仅确定两个位姿之间的平移距离,即,仅仅确定第四相机位姿与第二相机位姿之间的距离。

[0143] 在下一个步骤S5中,根据第二相机的位姿以及第一相机与第二相机之间的空间变换,确定第一相机的至少一部分处于绝对比例的位姿,根据第四相机的位姿以及第三相机与第四相机之间的空间变换,确定第三相机的至少一部分的处于绝对比例的位姿。

[0144] 在另一实施方案中,也可能没有确定与第一实物有关的两个完全位姿,即捕获第



一图像的第一相机位姿以及捕获第三图像的第三相机位姿,但是仅仅确定两个位姿之间的差别,换句话说,仅仅确定第三相机相对于第一相机的位姿。

[0145] 在另一实施方案中,也可能仅仅确定两个位姿之间的平移距离,即,仅仅确定第三相机位姿与第一相机位姿之间的距离。

[0146] 然后可以将该距离和所估算出的第二相机与第四相机之间的距离一起用于确定第一相机(例如,所谓的面朝世界的相机或后置相机)的绝对比例。例如,比例因子可以定义真实环境的重构几何模型的真实大小,或者用于将处于任意比例的重构件坐标映射成绝对比例。

[0147] 在步骤S6中,处于绝对比例的第一实物的3D重构件(也称作几何模型)的空间坐标可通过第一相机的至少一部分的位姿、第三相机的至少一部分的位姿、第一图像和第三图像进行确定,或者通过两个位姿之间的差别进行确定,或者通过两个位姿之间的距离进行确定。

[0148] 在另一实施方案中,根据(第一共同坐标系中)第一相机和第三相机之间的距离以及(第二共同坐标系中)第二相机与第四相机之间的距离,计算第一相机(例如,面朝世界的相机)的绝对比例因子。

[0149] 图2示出了本发明的一个可能的实施方案以及使用单眼SLAM或SfM时比例模糊的问题。图2a示出了自上而下的场景图,该场景包括大型车O1(例如,用于驾驶的真实汽车)、四个相机W1、F1、W2和F2(分别对应第一相机、第二相机、第三相机和第四相机)、和用户U1。图2a还以四幅小图示出了分别由四个相机W1、F1、W2和F2拍摄的图像I(W1)、I(F1)、I(W2)和I(F2)。图2b示出了自上而下的场景图,该场景包括小型车O2(例如,孩童型玩具车)、四个相机W3、F3、W4和F4(分别对应第一相机、第二相机、第三相机和第四相机)、和用户U1。图2b还以四幅小图示出了分别由四个相机W3、F3、W4和F4拍摄的图像I(W3)、I(F3)、I(W4)和I(F4)。

[0150] 假设,尽管大型车O1和小型车O2的大小明显不同,图像I(W1)和I(W3)以及图像I(W2)和I(W4)基本完全相同。这是因为相机W3和W4的位姿以与O2相对于O1按比例缩放相同的方式而相对于W1和W2按比例缩放。这显示了比例模糊的问题。仅仅根据一张或多张图像确定汽车的绝对大小是不可能的,因此,根据图像I(W1)、I(W2)、I(W3)或I(W4)来区别大型车O1和小型车O2是不可能的。因此,如果汽车的大小未知,根据图像来确定W1和W2或者W3和W4的相机位置之间处于绝对比例的距离也是不可能的。然而,如果W1和W2的相机位置之间的绝对距离已知,例如,为1米,确定汽车的绝对大小也是不可能的。确定实际物体的绝对大小在许多应用程序中将是有利的。因此,根据本发明,确定实际物体的绝对大小并以绝对比例将其进行重构是有利的。

[0151] 除了汽车以外,用户U1(即,人)也位于图2a和图2b所示的场景中。根据本发明的实施方案,我们捕获到了用户U1的包括其脸部的图像I(F1)、I(F2)以及I(F3)、I(F4)。在一个实施方案中,F1和W1、F2和W2、F3和W3的相机位姿之间、以及F4和W4之间的空间变换已知,并且可能相同,如果图像由两个刚性连接的相机进行捕获,即W1、W2、W3和W4是处于不同时间点上的相同物理相机,F1、F2、F3和F4是处于不同时间点上的相同物理相机。在另一实施方案中,W1、W2、W3和W4是不同相机,F1、F2、F3和F4是不同相机。

[0152] 由于用户U1在图2a和图2b中的大小没有改变,对应的图像I(F1)和I(F2)分别不同于图像I(F3)和I(F4)。I(F1)与I(F2)之间的脸部位置比I(F3)与I(F4)之间的脸部位置差别

更大,这表示,相机W1与W2之间的运动比W3与W4之间的运动更大。现在假设,用户脸部的至少一个脸部基准点的至少一个空间特性按绝对比例给出,例如,已知瞳孔间距为63毫米(已知通用脸部模型的眼睛距离为63mm)。假设该基准点(例如,眼睛中心位置)可通过脸部或眼睛检测算法的方式自动在图像I (F1) 和I (F2) 中进行定位。也可使用其它脸部点,诸如以下一项或多项:位置(角落位置、中心位置、或边界区)、大小、形状、轮廓、区域、比例、比率、左眼与右眼之间的距离以及左眼和右眼外观(瞳孔、虹膜、角膜、巩膜、内眦、外眦、中心、上下眼睑、睫毛,……)、鼻梁、鼻子(鼻尖、鼻背、鼻翼、鼻孔鼻小柱,……)、人中、嘴唇、左右耳、左右眼眉、牙齿、左右脸颊、下颌、脖子、喉结。然后可根据图像I (F1) 和I (F2) 中脸部基准点的位置以绝对比例确定相机F1的位置与相机F2的位置之间的距离,这是因为这些基准点具有处于绝对比例的已知空间特性。在相机W1与F1之间的空间变换以及W2与F2之间的空间变换已知的情况下,也可以绝对比例计算相机W1的位置与相机W2的位置之间的距离。这一点同样适用于W3、F3、W4和F4。

[0153] 相机图像I (W1) 和I (W2) 和拍摄这些图像的相机W1与W2之间的已知距离一起最终能够以绝对比例重构物体O1,或者确定比例因子。将O1的处于任意比例的现有重构件映射成绝对比例,即映射成与实际空间度量单位(诸如,米或英寸)具有已知关系的坐标系。

[0154] 作为结果,O1的重构件在比例上不同于O2的重构件,这实现了对O1和O2的区分。

[0155] 图3示出了本发明的另一实施方案。示出了处于不同时间点上的自上而下的场景图,附图上部示出了第一时间点的场景,附图下部示出了第二时间点的场景。附图左部分的小图示出了相应相机所捕获的图像。相机W5用于捕获汽车O1的图像I (W5),同时相对于相机W5具有已知空间变换的相机F5捕获用户U1和其脸部的图像I (F5)。然后两个相机朝用户远离汽车,从而形成相机位姿F6和W6以及对应的图像I (F6) 和I (W6)。同样地,仅仅给出相机图像I (W5) 和I (W6) 不可能确定汽车的绝对大小,即,汽车在实际中的大小。我们假设相机F5、W5、F6、W6的固有参数(尤其是焦距)已知。我们进一步假设,在用于眼睛检测或脸部检测的特定脸部模型或通用脸部模型中,用户U1的具有绝对单位(例如,厘米)瞳孔间距P已知。而且,用户瞳孔之间的距离(P5和P6)可以在像素坐标中的图像(F5) 和I (F6) 中进行确定。在相机F5的成像平面到(左眼和右眼)两个瞳孔的距离相同的特殊情况下,距离P5与F5的固有相机参数一起足以计算出瞳孔间的点与F5相机中心之间处于绝对比例的距离D5,即,具有绝对实际单位的距离,诸如米,因为 $D5 = (f * P) / P5$ ,其中f是以像素为单位的相机F5的焦距(如果脸部以相机的主点为中心,则采用该公式)。类似地,当相机F6的成像平面到(左眼和右眼)两个瞳孔的距离相同,可根据F6的焦距、图像I (F6) 中的瞳孔距离P6、以及用户U1处于绝对比例的瞳孔间距P来确定用户U1与相机F6之间的距离D6。如果在该场景中用户U1相对于汽车O1没有移动,则距离D5与D6之差可用于确定相机W5的位置与相机W6的位置之间处于绝对比例的距离。这实现了在考虑相机图像I (W5) 和I (W6) 的情况下对汽车O1的绝对比例的确

[0156] 图4示出了所涉及的坐标系和变换过程的实施方案。用相对于实际距离已知的比例因子(诸如,米)以绝对单位来定义物体坐标系O。坐标系W与面朝世界的相机(诸如W1、W2等)相关,坐标系F与面朝用户的相机(诸如F1、F2等)相关。当对面朝世界的相机的图像进行SfM或SLAM时,则相机的坐标系W以任意比例进行定义,如上所述。假设W与F之间的变换(即,(处于绝对比例的)旋转R和平移T)已知。该变换可以是静止的以及下线后经过校准的(即,

在用于场景中之前),例如,如果两个相机是移动手机或平板计算机的面朝用户的相机和面朝世界的相机,坐标系F以绝对比例进行定义,这是因为其位姿是通过用户脸部坐标系U中的脸部特征进行估算,其以绝对比例进行定义,并且脸部特征也是以绝对比例进行定义。本发明的理念之一就是现在利用以下事实:F可以绝对比例进行定义,F与W之间的刚体变换已知。作为结果,在本实施方案中,同样地W可以绝对比例进行定义,这一点在没有坐标系F和U的情况下,即在没有面朝用户的相机以及用户脸部的情况下,是不可能的。

[0157] 图5示出了捕获装置的实施方案,捕获装置包括处于不同位姿上的面朝用户的相机和面朝世界的刚性附接相机。面朝世界的相机和面朝用户的相机可以是至少两个不同的相机或单个相机,例如,捕获视野为360°的全方位相机。在此情况下,相应相机的相机图像可定义为整体图像的不同两部分。对于单个相机,位姿F和W可相同。

[0158] 位姿F8和W8处所用的相机可以与不同时间点上的位姿F7和W7处所用的相机相同,或者它们可以是同时或在不同时间点上使用的不同相机。面朝用户的相机的图像实现了对相机位置F7与F8之间处于绝对比例的空间距离信息(标量值)的确定、以及对相机W7的位置与相机W8的位置之间处于绝对比例的空间距离d推断信息(也是标量值)的确定,这是因为F7与W7之间的变换以及F8与W8之间的变换已知。

[0159] 图6示出了本发明的用手持设备H9来实施的示例性实施方案,诸如移动手机、智能手机、平板手机、或平板计算机,包括可捕获用户脸部U9的面朝用户的相机F9以及可捕获实物O9的面朝世界的相机W9。此外,手持设备H9可具有处理设备P9和无线网络单元。本文所公开的任何实施方案的任何步骤可以由处理设备P9在手持设备H9中进行本地执行,或者通过无线网络单元发送至远端服务器计算机S9或另一移动设备。手持设备的面朝用户的相机、面朝世界的相机、和显示屏之间可具有已知的空间关系。

[0160] 根据一个实施方案,本文所述的处理系统可至少部分包括在手持设备H9中,并且/或者经适配与手持设备H9进行(例如,无线)通信的服务器计算机S9中。处理系统可仅仅包括在这些装置中的一者中,例如,手持设备H9中或服务器计算机S9中,或者可以是分布式系统,其中一个或多个处理任务(执行一个或多个方法步骤)由一个或多个处理设备(诸如,微处理器)进行下发和处理,一个或多个处理设备在空间上分散,彼此进行通信。

[0161] 有了此类系统装置,用户U9可通过以下方式重构实物O9,在此实施例中,实物为处于绝对比例的椅子:将手持设备H9移动到至少两个不同的视角下,同时头部U9相对于实物O9不移动。这实现了,将O9重构件用作相机位姿估算的参考项时,对处于绝对比例的U9的位姿的确定。这同样地实现了:以绝对比例(见图像I92)而不是以错误的任意比例(见图像I93)将虚物(例如,虚拟枕头)叠加在面朝世界的相机W9的相机图像I91上。这使增强物更加真实,在图像I92中具有与实际一致的比例,而在图像I93中虚拟枕头相对于真实椅子太小。

[0162] 图7示出了根据本发明的实施方案的指导用户进行比例估算的图形用户界面的实施例。在此实施方案中,该方法在移动手机D2中进行,移动手机D2包括显示设备D1(在此情况下为触摸屏)、面朝用户的相机F10以及在附图中不可见的在手机背部的面朝世界的相机。在此实施方案中,面朝世界的相机反复捕获放在移动手机D2后方的实物(汽车)的图像。这些图像在显示设备D1上显示。在初始状态S11下,用户可通过触摸显示屏上的按钮G1来开启比例估算过程。此后面朝用户的相机F10用于确定面朝用户的相机相对于用户脸部处于绝对比例的位姿。此外,此后面朝世界的相机用于确定面朝世界的相机相对于实物(汽车)

处于任意比例的位姿。然后,在阶段S12,按钮外观发生改变,提示其去激活(见空的按钮G2)。另外的图形用户界面元件G3和G4出现,为相机运动提供指令以执行该过程并且将该过程的进度可视化。此外,可将标签G5显示在显示设备上,提供文本指令或信息。当用户将手机(以及由此面朝世界的相机)移动至状态S13之后,进度指示器G6相应地进行调整并显示已经取得了进展。一旦相机运动足够大以进行比例估算(S14),进度指示器G7显示已经实现目标。此外,标签G8可通知用户校准成功或失败。此时,面朝用户的相机以及面朝世界的相机可停止捕获和/或追踪。此外,应用程序可再次处于初始状态S11,除了现在可估算实物(汽车)的绝对比例,并且面朝世界的相机可确定面朝世界的相机相对于实物(汽车)处于绝对比例的位姿。通过再次触摸按钮G1,用户可触发另一比例估算流程。

[0163] 图8示出了第一相机与第二相机之间的空间变换的影响。第一相机和第三相机可以由相同的物理相机来实施(例如,面朝世界的相机,例如,移动手机的面朝世界的相机),第二相机和第四相机可以由相同的物理相机来实施(例如,面朝用户的相机,例如,移动手机的面朝用户的相机)。经发现,这两个相机(例如,面朝世界的相机和面朝用户的相机)一起可形成双相机,其中两个相机之间的相机视锥不重叠。

[0164] 图8a、8b和8c的实施方案示出了此类双相机的三种不同装置,其不同在于第一相机与第二相机之间(以及由此第三相机与第四相机之间)的空间平移长度。每幅子图包含了自上而下的场景图,场景包括四个相机,其中每幅子图显示了处于两个位姿下的双相机,其中在所有子图中第一相机和第三相机具有相同的位姿。相机的原点以OR1、OR2、OR3和OR4示出。在每幅自上而下的视图中,标出了相机原点OR1与OR3(上面)之间的绝对空间距离以及相机原点OR2和OR4(下面)之间的绝对空间距离。相机原点之间的绝对空间距离值得注意,这是因为其包含绝对比例信息。

[0165] 第一相机(指数“1”)与第三相机(指数“3”)之间的旋转运动引起第二相机(指数“2”)与第四相机(指数“4”)之间的平移运动,这取决于第一相机与第二相机之间发生空间平移的长度并引起以下差异:OR1与OR3之间的绝对空间距离长度与OR2与OR4之间的绝对空间距离长度相比的差异。

[0166] 在图8a中,其中第一相机与第二相机之间的空间平移长度为零,相机原点OR1与OR3(上面)之间的绝对空间距离以及相机原点OR2与OR4(下面)之间的绝对空间距离相同。

[0167] 在图8b中,其中第一相机与第二相机之间的空间平移长度与第二相机与第四相机之间的平移运动长度相比相当小,相机原点OR1与OR3(上面)之间的绝对空间距离以及相机原点OR2与OR4(下面)之间的绝对空间距离相似但并不完全相同,并且当可以引入很小的误差时可视为相等。

[0168] 在图8c中,其中第一相机与第二相机之间的空间平移长度与第二相机与第四相机之间的平移运动长度相比并不小,相机原点OR1与OR3(上面)之间的绝对空间距离以及相机原点OR2与OR4(下面)之间的绝对空间距离相当不同,并且在不可以引入很大误差的情况下不可视为相等。

[0169] 根据所提供的第一相机与第二相机(参见上述步骤e)之间的第一空间变换,确定第一相机的至少一部分相对于位姿P2F的位姿(P1F)(参见步骤h)。同样地根据所提供的第三相机与第四相机(参见步骤e)之间的第二空间变换,确定第三相机的至少一部分相对于位姿P4F的位姿(P3F)(参见步骤h)。

[0170] 根据第二相机的位姿、第一相机与第二相机之间的空间变换 $T_1$ 、第四相机的位姿、和第三相机与第四相机之间的空间变换 $T_2$ ，确定第一相机的位姿与第三相机的位姿之间的绝对空间距离，其中 $T_1$ 和 $T_2$ 可以相同。此变换可能是6D刚体变换，可包括坐标系中第一相机与第二相机之间处于绝对比例的相对空间位置和取向的校准过程。这可以通过诸如参考文献[11]所公开的方法进行确定。在此情况下，考虑了由相机1与相机3之间的旋转运动而引起的相机2与相机4之间的平移运动，第一相机(相机1)的位姿与第三相机(相机3)的位姿之间的绝对空间距离的计算过程是精确的。

[0171] 根据另一实施方案(参见上文所述的本发明的第二方面)，该方法将相机F(第二相机)与相机W(第一相机)之间的空间距离视作零。由此，忽略由第二相机的位姿与第四相机的位姿之间的运动而引起的第三相机的位姿与第一相机的位姿之间的平移运动。对于第一相机计算位姿与第三相机计算位姿的平移部分，所引起的误差小于或等于第一相机与第二相机之间的实际距离。对于 $OR_2$ 与 $OR_4$ 之间的绝对空间距离，所引起的误差小于或等于第一相机与第二相机之间的实际距离的两倍。所引起的误差也取决于第二相机的位姿与第四相机的位姿之间的旋转量。 $180^\circ$ 旋转引起更高的误差。第二相机的位姿与第四相机的位姿之间仅仅发生基本平移的运动，即，仅仅具有可忽略的旋转部分，这只会引起可忽略的误差。

[0172] 这表示，对于第一相机与第二相机之间的距离跟第二相机与第四相机之间的平移运动相比较小的相机装置而言，第一相机与第二相机之前的空间变换的平移部分可视为恒等变换。这也表示，对于仅仅具有(对于第一相机与第二相机)与第二相机与第四相机之间的平移运动相比可忽略的旋转部分的相机运动而言，第一相机与第二相机之前的空间变换的平移部分可视为恒等变换。

[0173] 图9示出了本发明的另一实施方案，其中第一相机的至少一部分(面朝世界的相机)的位姿用作3D重构方法的输入，使得该方法可以绝对比例重构物体 $O_9$ 。可通过脸部追踪算法以绝对比例来确定两个(例如)面朝用户的相机 $C_2$ (第二相机)和 $C_4$ (第四相机)相对于用户脸部 $U$ 的坐标系的位姿 $P_{2F}$ 和 $P_{4F}$ 。给出相机 $C_1$ (第一相机)与 $C_2$ (第二相机)之间的空间变换 $T_1$ 以及相机 $C_3$ (第三相机)和 $C_4$ (第四相机)之间的空间变换 $T_2$ ，可通过级联法在用户脸部 $U$ 的坐标系中以绝对比例来确定(例如)面朝世界的相机 $C_1$ 的位姿 $P_{1F}$ 以及(例如)面朝用户的相机 $C_3$ 的位姿 $P_{3F}$ 。本实施方案然后利用共同坐标系中以绝对比例确定的两个面朝世界的相机 $C_1$ 和 $C_3$ 的位姿，以通过以下方式以绝对比例创建实物 $O_9$ 的3D重构件：建立相机 $C_1$ 和 $C_3$ 的相机图像的对应关系以及对其深度进行三角测量。在本实施方案中，3D重构方法并不估算相机位姿，而是使用所提供的处于绝对比例的相机位姿。

[0174] 在另一实施方案中，3D重构方法估算场景的3D结构和相机位姿，相机位姿的解空间仅限于平移距离等于两个相机位姿 $C_1$ 与 $C_3$ 之间的已知平移距离的位姿对，该已知平移距离是上述过程的结果。

[0175] 图10示出了本发明的另一实施方案，其中确定了以任意比例定义的第一坐标系与以绝对比例定义的第二坐标系之间的比例因子。在本实施方案中，可通过脸部追踪算法以绝对比例来确定两个面朝用户的相机 $C_2$ 和 $C_4$ 相对于用户脸部 $U$ 的坐标系的位姿 $P_{2F}$ 和 $P_{4F}$ 。给出 $C_1$ 与 $C_2$ 之间的空间变换 $T_1$ 以及 $C_3$ 和 $C_4$ 之间的空间变换 $T_2$ ，可通过级联法在用户脸部 $U$ 的坐标系中以绝对比例来确定面朝世界的相机 $C_1$ 的位姿 $P_{1F}$ 以及面朝用户的相机 $C_3$ 的位姿 $P_{3F}$ 。根据这些，相机 $C_1$ 相对于 $P_{1F}$ 的位置与相机 $C_3$ 相对于 $P_{3F}$ 的位置之间的平移距离可以绝

对比例确定为 $D_{abs}$ 。面朝世界的相机 $C1$ 和 $C3$ 的相机图像可用于：根据实物 $O10$ 的处于任意比例的3D重构件，以任意比例确定与实物 $O10$ 相关的坐标系 $0A$ 中相机 $C1$ 的位姿 $P1W$ ，以及以任意比例确定与实物 $O10$ 相关的坐标系中相机 $C3$ 的位姿 $P3W$ 。根据这两个位姿，相机 $C1$ 相对于 $P1W$ 的位置与相机 $C3$ 相对于 $P3W$ 的位置之间的平移距离可以实物 $O10$ 的任意重构比例确定为 $D_{arb}$ 。最后，根据以绝对比例进行的脸部追踪所确定的相机位姿，将实物重构件的坐标系中相机 $C1$ 的位置与相机 $C3$ 的位置之间的平移距离以任意比例确定为 $D_{arb}$ 、以及将相机 $C1$ 与 $C3$ 之间的平移距离以绝对比例确定为 $D_{abs}$ 之后，可以将实物重构件的坐标系的任意比例与绝对比例之间的比例因子确定为 $D_{abs}/D_{arb}$ 。比例因子可用于按比例缩放实物 $O10$ 的3D重构件的坐标，使得其在按比例缩放后以决定比例进行定义。

[0176] 在另一实施方案中，假设空间变换 $T1$ 和 $T2$ 具有非常小的处理为零的平移部分，因此，计算两个面朝用户的相机 $C2$ 与 $C4$ 之间相对于用户坐标系的处于绝对比例的平移距离 $D_{abs}$ ，而非计算面朝世界的相机 $C1$ 与 $C3$ 相对于用户坐标系的位姿之间处于绝对比例的 $D_{abs}$ 。

[0177] 图11示出了本发明的另一实施方案，并且根据本发明的实施方案示出了如何将四幅以上图像中的输入数据合并成估算出来的最终比例因子。

[0178] 根据本发明的实施方案，方框1101对应于根据所提供的四幅图像 $I(C1)$ 、 $I(C2)$ 、 $I(C3)$ 、 $I(C4)$ 以及所提供的两次变换 $T1$ 和 $T2$ 所进行的比例因子的一次单独估算过程。

[0179] 图像 $I(C1)$ 是 $C1$ 第一相机所捕获的图像，示出了第一实物的一部分。图像 $I(C2)$ 是 $C2$ 第二相机所捕获的图像，示出了第二实物的一部分。 $T1$ 是相机 $C1$ 的位姿与相机 $C2$ 的位姿之间的已知变换。根据示出了第一实物的一部分的图像 $I(C1)$ ，确定处于任意比例的第一共同坐标系中相机 $C1$ 的位姿 $P1W$ 。根据示出了第二实物的一部分的图像 $I(C2)$ ，确定处于任意比例的第二共同坐标系中相机 $C2$ 的位姿 $P2F$ 。通过所提供的变换 $T1$ 由位姿 $P2F$ 的变换来确定第二共同坐标系中相机 $C1$ 的位姿 $P1F$ 。该变换是任选的，即可以将 $T1$ 的平移部分视作恒等变换（即，平移部分为零或可以忽略），从而使 $P1F$ 的平移部分等于 $P2F$ 的平移部分。

[0180] 图像 $I(C3)$ 是 $C3$ 第三相机所捕获的图像，示出了第一实物的一部分。图像 $I(C4)$ 是 $C4$ 第四相机所捕获的图像，示出了第二实物的一部分。 $T2$ 是相机 $C3$ 的位姿与相机 $C4$ 的位姿之间的已知变换。根据示出了第一实物的一部分的图像 $I(C3)$ ，确定处于任意比例的第一共同坐标系中相机 $C3$ 的位姿 $P3W$ 。根据示出了第二实物的一部分的图像 $I(C4)$ ，确定处于任意比例的第二共同坐标系中相机 $C4$ 的位姿 $P4F$ 。通过所提供的变换 $T2$ 由位姿 $P4F$ 的变换来确定第二共同坐标系中相机 $C3$ 的位姿 $P3F$ 。该变换是任选的，即也可以将 $T2$ 的平移部分视作恒等变换（即，平移部分为零或可以忽略），从而使 $P3F$ 的平移部分等于 $P4F$ 的平移部分。

[0181] 根据第一共同坐标系中第一相机 $C1$ 的位姿 $P1W$ 的平移部分、第二共同坐标系中第一相机 $C1$ 的位姿 $P1F$ 的平移部分、第一共同坐标系中第三相机 $C3$ 的位姿 $P3W$ 的平移部分、以及第二共同坐标系中第三相机 $C3$ 的位姿 $P3F$ 的平移部分，可确定将第一共同坐标系从任意比例按比例缩放至绝对比例的比例因子 $S1$ 。

[0182] 根据本发明的实施方案，方框1102对应于根据所提供的四幅图像 $I(C5)$ 、 $I(C6)$ 、 $I(C7)$ 、 $I(C8)$ 以及所提供的两次变换 $T3$ 和 $T4$ 所进行的比例因子的另一单独估算过程。由此图像可以是不同的图像或部分相同的图像，如方框1101所用。例如， $I(C5)$ 和 $I(C6)$ 可分别与 $I(C3)$ 和 $I(C4)$ 相同。在另一实施例中， $I(C7)$ 和 $I(C8)$ 可分别与 $I(C3)$ 和 $I(C4)$ 相同。

[0183] 图像I (C5) 是相机C5所捕获的图像, 示出了第一实物的一部分。图像I (C6) 是相机C6所捕获的图像, 示出了第二实物的一部分。T3是相机C5的位姿与相机C6的位姿之间的已知变换。根据示出了第一实物的一部分的图像I (C5), 确定处于任意比例的第一共同坐标系中相机C5的位姿P5W。根据示出了第二实物的一部分的图像I (C6), 确定处于任意比例的第二共同坐标系中相机C6的位姿P6F。通过所提供的变换T3由位姿P6F的变换来确定第二共同坐标系中相机C5的位姿P5F。该变换是任选的, 即也可以将T3的平移部分视作恒等变换, 从而使P5F的平移部分等于P6F的平移部分。

[0184] 图像I (C7) 是相机C3所捕获的图像, 示出了第一实物的一部分。图像I (C8) 是相机C8所捕获的图像, 示出了第二实物的一部分。T4是相机C7的位姿与相机C8的位姿之间的已知变换。根据示出了第一实物的一部分的图像I (C7), 确定处于任意比例的第一共同坐标系中相机C7的位姿P7W。根据示出了第二实物的一部分的图像I (C8), 确定处于任意比例的第二共同坐标系中相机C8的位姿P8F。通过所提供的变换T4由位姿P8F的变换来确定第二共同坐标系中相机C7的位姿P7F。该变换是任选的, 即也可以将T4的平移部分视作恒等变换, 从而使P7F的平移部分等于P8F的平移部分。

[0185] 根据第一共同坐标系中相机C5的位姿P5W、第二共同坐标系中相机C5的位姿P5F、第一共同坐标系中相机C7的位姿P7W、以及第二共同坐标系中相机C7的位姿P7F, 可确定将第一共同坐标系从任意比例按比例缩放至绝对比例的比例因子S2。

[0186] 最后可通过如下方式将单独的比例估算值S1和S2合并成最终的比例估算值S: 通过(例如) 像平均数、平均值、中值、概率最大化或RANSC等一个或多个方法进行数学模型拟合。

[0187] 上述示例包括两个单独的比例估算值S1和S2, 并且将其合并成最终的比例估算值。当然, 类似地可进行两次以上单独的计算过程, 例如, 3、4、5或100等, 并合并全部估算值。

[0188] 图12示出了本发明的另一实施方案, 并且根据本发明的实施方案示出了如何将四幅以上图像中的输入数据合并成估算出来的最终比例因子。

[0189] 根据本发明的实施方案, 该实施例是基于所提供的六幅图像I (C1)、I (C2)、I (C3)、I (C4)、I (C5)、I (C6) 以及所提供的三次变换T1、T2和T3。

[0190] 图像I (C1) 是C1第一相机所捕获的图像, 示出了第一实物的一部分。图像I (C2) 是C2第二相机所捕获的图像, 示出了第二实物的一部分。T1是相机C1的位姿与相机C2的位姿之间的已知变换。根据示出了第一实物的一部分的图像I (C1), 确定处于任意比例的第一共同坐标系中相机C1的位姿P1W。根据示出了第二实物的一部分的图像I (C2), 确定处于任意比例的第二共同坐标系中相机C2的位姿P2F。通过所提供的变换T1由位姿P2F的变换来确定第二共同坐标系中相机C1的位姿P1F。该变换是任选的, 即也可以将T1的平移部分视作恒等变换, 从而使P1F的平移部分等于P2F的平移部分。

[0191] 图像I (C3) 是C3第三相机所捕获的图像, 示出了第一实物的一部分。图像I (C4) 是C4第四相机所捕获的图像, 示出了第二实物的一部分。T2是相机C3的位姿与相机C4的位姿之间的已知变换。根据示出了第一实物的一部分的图像I (C3), 确定处于任意比例的第一共同坐标系中相机C3的位姿P3W。根据示出了第二实物的一部分的图像I (C4), 确定处于任意比例的第二共同坐标系中相机C4的位姿P4F。通过所提供的变换T2由位姿P4F的变换来确定

第二共同坐标系中相机C3的位姿P3F。该变换是任选的,即也可以将T2的平移部分视作恒等变换,从而使P3F的平移部分等于P4F的平移部分。

[0192] 图像I (C5) 是相机C5所捕获的图像,示出了第一实物的一部分。图像I (C6) 是相机C6所捕获的图像,示出了第二实物的一部分。T3是相机C5的位姿与相机C6的位姿之间的已知变换。根据示出了第一实物的一部分的图像I (C5), 确定处于任意比例的第一共同坐标系中相机C5的位姿P5W。根据示出了第二物体的一部分的图像I (C6), 确定处于任意比例的第二共同坐标系中相机C6的位姿P6F。通过所提供的变换T3由位姿P6F的变换来确定第二共同坐标系中相机C5的位姿P5F。该变换是任选的,即也可以将T3的平移部分视作恒等变换,从而使P5F的平移部分等于P6F的平移部分。

[0193] 根据位姿P1W、P3W和P5W,即第一共同坐标系中相机C1、C3和C5的位姿、以及位姿P1F、P3F和P5F,即第二共同坐标系中相机C1、C3和C5的位姿,将第一共同坐标系从任意比例按比例缩放至绝对比例的比例因子S可通过 ([0193]) 一种或多种方法,像迭代最近点 (ICP)、Umeyama方法、或Kabsch方法、或其它最小二乘法和/或RANSAC模型拟合、基于点集的配准法,进行数学模型拟合。

[0194] 图12示出了使用3个相机捕获第一物体,3个相机捕获第二物体。当然也可以利用第一共同坐标系中甚至3个以上的相机位姿以及第二共同坐标系中相应相机所对应的3个位姿,例如,第一和第二共同坐标系中4个、5个、6个或100个等位姿对。

[0195] 下面进一步公开了本发明的其它实施方案而没有明确地参考绘图或附图。

[0196] 根据其它实施方案,方法假设,在捕获有利于进行比例估算的图像的同时,用户脸部相对于待追踪或待重构实物是静止定位的。

[0197] 根据其它实施方案,方法检测到用户脸部相对于实物静止定位的时候,然后比例估算时仅仅使用当用户脸部相对于实物静止定位时所捕获的图像。这可以 ([0197]) 通过以下方式完成:将面朝用户的相机的两个位姿的对极几何图与面朝世界的相机中的特征移动进行比较或反过来。确定头部是否相对于第一实物移动的另一方案是基于一组面朝用户的相机和面朝世界的相机的对应位姿。这些位姿可转化为表示相机位置的3D点。用以确定二组点之间的相似变换的算法 ([0197]) 由Umeyama [10] 公开。配准之后可计算剩余误差。如果剩余误差超过特定阈值,两组对应的位姿则视作与刚体变换无关。这表示头部 (即,脸部) 已经相对于第一实物发生移动。

[0198] 根据另外的实施方案,方法通过以下方式处理并补偿脸部相对于第一实物的运动:通过视觉追踪来估算头部相对于实物的运动,以估算出面朝用户的相机的相机图像中脸部的运动以及面朝世界的相机的相机图像中第一实物的运动。

[0199] 根据另外的实施方案,方法通过以下方式处理并补偿脸部相对于实物的运动:通过视觉追踪来估算头部相对于实物的运动,以分别估算出面朝世界的相机的相机图像中脸部的运动和背景的运动。

[0200] 根据另外的实施方案,通过相机F1和F2的位姿来计算相机W1与W2之间的绝对空间距离可包括:校准处于绝对比例的坐标系中坐标系F与W之间的相对空间位置与取向,例如,通过诸如参考文献 [11] 公开的方法。

[0201] 根据另一实施方案,坐标系F和W之间的变换为6 DoF (DoF: 自由度) 刚体变换,包括3D旋转和3D平移。



[0202] 根据另外的实施方案,该方法提供并考虑了坐标系F与W之间的空间距离。

[0203] 根据另外的实施方案,方法将坐标系F与W之间的空间距离视为零,从而忽略了由坐标系(相机)W的旋转而引起的坐标系(相机)F的位姿的平移移动以及反之亦可,这导致,对应于相机W1和W2的位姿,相机F1和F2的位姿出现小于或等于坐标系(相机)F与(相机)W之间的实际距离的误差。

[0204] 根据另外的实施方案,方法将坐标系F与W之间的取向差别视为180度,即相应相机的光轴平行且相机面朝相反的方向。

[0205] 根据一个实施方案,方法使用像参考文献[15]公开的脸部基准点追踪器进行脸部校正并确定脸部特征的2D位置以及图像中的基准点。

[0206] 根据一个实施方案,方法使用脸部追踪方法,传送与脸部有关的面朝用户相机的全6D位姿,像参考文献[16]所公开,其中将统计型人体测量3D刚性模型用作人体头部的近似体。随后将所传送的6D位姿以可能方式根据特定脸部特征(像,眼睛距离)的尺寸的其它规格进行比例与平移修改。

[0207] 根据一个实施方案,方法使用脸部追踪方法,根据用户特定脸部的预获得、拟合或配置模型,传送与处于绝对比例的脸部有关的面朝用户相机的全6D位姿。

[0208] 根据一个实施方案,方法使用脸部追踪方法,传送包含脸部相对于面朝用户相机的取向的3D位姿(像参考文献[17]所公开),并一起使用3D位姿与检测到的脸部特征来推断另外的位姿平移信息。对应于脸部特征的两个位置之间的视角可(例如)与这两个位置之间补偿头部旋转所需的实际距离一起使用,以弥补从相机到脸部的距离。不同位姿之间的距离变化可用于推断相机移动的绝对比例信息。

[0209] 根据一个实施方案,方法使用脸部追踪方法,传送包含脸部相对于面朝用户相机的左右取向(即,偏转取向)的1D位姿(像参考文献[17]所公开),并一起使用该1D位姿与检测到的脸部特征来推断另外的位姿平移信息,假设这些脸部特征分布于脸部的水平线上,像眼睛中心。两个眼睛中心之间的视角可以与补偿头部1D左右旋转(偏转)所需的实际眼睛距离一起使用,以弥补从相机到脸部/眼睛的距离。不同位姿之间的距离变化可用于推断相机移动的绝对比例信息。

[0210] 根据另一实施方案,该方法假设,相机F1和F2的位姿仅限于在脸部前方,其中在朝向或背离脸部(几乎)只进行平移时,F2的位姿不同于F1的位姿,使用检测到的眼睛位置、眼睛之间的相关视角、以及所需的实际眼睛距离来弥补从相机到脸部/眼睛的距离,眼睛位置通过(例如)参考文献[18]公开的方法在所捕获的图像中检测出来。这也在图3中示出。两只眼睛之间的视角可通过相应视向的给出了视角余弦的点积进行计算。然后可如下计算到达脸部的距离:用一半的眼睛距离除以一半视角的切角。

[0211] 相机F1与F2以及W1与W2的相应位姿之间的距离变化分别可用于推断相机W1和W2的坐标系的绝对比例信息。

[0212] 根据一个实施方案,至少对于一个脸部特征,以绝对空间单位提供了空间特性,例如,人体瞳孔间距,据此该特性可作为单个值或概率分布来给出。该特性可以单独针对特定用户进行确定或多人通用。同样地,可以针对不同的人群(性别、种族、年龄……)定义多个值/概率分布,有关群体可通过用户手动输入或其它自动标记或分类程序(例如,如参考文献[12]所公开的程序)进行选择。

[0213] 根据一个实施方案,特定脸部可根据以下内容自动进行校准:

[0214] 提供给第二相机(例如,面朝用户相机)的绝对比例信息。绝对比例信息可(例如)通过第二实物的深度信息提供,第二实物的深度信息是通过以下方法获得的深度:散焦、渡越时间、结构光、主动照明法、亮度法、拍摄光、激光测距仪、多频相移、干扰量度法或被动式立体法。迹线较小的立体法,和移动手机的情况一样,可以更加可靠地作用于第二相机(例如,面朝用户的相机),其中第二实物为靠近相机的脸部,通常靠近程度不超过50cm,这与第一相机(例如,面朝世界的相机)的立体法相反,其中第一实物位于距离脸部较远的任意位置。这是因为,对于处于所捕获物体的深度中的某深度“分辨率”,所需基线的大小取决于所捕获的物体到相机的距离。

[0215] 根据一个实施方案,特定脸部可手动进行校准或者使用通用模型。由此,统计模型也可用以确定比例估算时的不确定性,该不确定性指出了不同人体中脸部空间特性的不同程度。

[0216] 瞳孔间距的手动校准可,例如,通过镜子或尺子来完成。面朝镜子,保持头部直立,居于镜子前方,将尺子水平放置到脸部前方,尽可能近地放到眼睛下方,在镜子中可看见测量标记。在不移动头部或尺子的情况下应当进行以下测量:闭上一只眼睛,另一只眼睛打开,可读取瞳孔中心下面尺子上的测量值。可以用另一只眼睛重复该过程(闭上之前打开的眼睛,并且打开之前闭上的眼睛)。两个读数之差即是瞳孔间距。以绝对比例进行瞳孔间距或其它脸部特征的半自动校准可(例如)通过以下方式进行:使用双相机设置,通过后置相机捕获的图像以绝对比例进行相机位姿估算(例如,以绝对比例进行的基于标识物的追踪、物体追踪或SLAM)。同时,待校准面部特征在面朝用户的相机中进行追踪当用户脸部相对于用于后置相机进行追踪的实际物体而静止定位时,绝对比例可通过采用相机位姿之间的已知变换而转化为脸部特征。

[0217] 根据如参考文献[13]公开的统计数据,通用模型可(例如)包含脸部特征的绝对空间特性的均值和概率分布,诸如,瞳孔间距。此类模型也可包括某数值的非确定性信息。通用模型也可包括多个测量值的多个(联合)概率分布,使得绝对比例根据联合概率分布中最大概率的参数空间中的位置进行确定,联合空间分布是根据所观察到的脸部特征的特性。

[0218] 根据一个实施方案,该方法合并不同来源(包括根据至少一个用户脸部图像所进行至少一次绝对比例估算)以及以下任一者的绝对比例估算值:IMU、GPS、场景中的已知物体、散焦深度、手动输入、被动式立体法、渡越时间、结构光、主动式照明法、亮度法、投影缩减式深度估算、以前的物体重构件的比例估算值的历史数据,比例估算值的历史数据可能与物体分类组合以仅仅考虑以前物体的以前物体重构件。

[0219] 根据本发明的实施方案,比例估算可作为迭代过程通过多个位姿对来进行,每个位姿对包括两个时间点上的两个位姿。可检测出不同测量值之间的不一致,并且比例的最佳一致值可通过以(例如)取平均值、取中值、取直方图中最大值的方式合并不同的测量值,并且可能根据(例如)各个比例估算过程中的不确定性、年龄或测量值之间的不一致性进行加权。合并不同测量值也可通过模型拟合法进行,诸如,实施像卡尔曼滤波器一样的贝叶斯滤波器以推断出绝对比例。

[0220] 同样地,对于许多用例,可假设用户脸部与面朝用户相机之间的距离的变化有限,例如,其中用户靠近装置,这是因为用户手持装置,或因为用户在显示设备旁边以体验AR应

用程序,由此也接近面朝用户相机。距离变化有限使情景更加稳健地用于标准相机的比例估算过程。也实现了:采用面朝用户的相机,仅仅估算小范围内的深度,不能估算较远物体的深度,诸如,隔壁房屋。

[0221] 本发明的一个可能实施方案包括:与面朝用户的深度相机的组合,该相机实现了通过用户脸部外观进行比例估算,以依赖于更精确的关于实际脸部模型几何结构和决定空间尺寸的信息。这实现了更多细节的提取,并且提高了以下需求:依赖来自标准脸部基础模型中的统计数据,或拟合与扭曲某通用脸部模型,或配置用户专用比例值。使用深度相机的另一优点是:可以在非常黑暗或非常明亮的环境中或量度变化非常强烈和/或频繁的环境中进行脸部检测与位姿估算。在此类环境中,标准动态范围小的单眼相机最可能将无法检测脸部并估算脸部的位姿。

[0222] 本发明的另一可能实施方案包括第二相机和/或第四相机,第二相机和/或第四相机为深度相机(例如,面朝用户的深度相机),深度相机实现了,采用基于深度信息的视觉视觉测程法,通过任何出现在深度相机视锥中的实物进行比例估算。这形成了处于绝对比例相机位姿,即使在没有利用任何脸部特性的情况下。

[0223] 在另一实施方案中,第二相机和/或第四相机(例如,面朝用户的相机)为红外相机,红外相机非常适合于脸部检测与追踪,或者以下列波段中至少一者进行成像:极紫外、近紫外、近红外、中红外、长波红外、或远红外。

[0224] 在另一实施方案中,第二相机和/或第四相机(例如,面朝用户的相机)由多个相机组成,诸如被动式立体相机或两个或多个相机的任何其它组合,可能对不同波段的光谱进行成像,诸如可见光、极紫外、近紫外、近红外、中红外、长波红外、或远红外。

[0225] 根据一个实施方案,该方法可还包括:检测第一用户输入。

[0226] 根据一个实施方案,要求用户用第二相机(例如,面朝用户的相机)进行某运动,例如,30cm的平移,这根据第二相机(例如,面朝用户的相机)捕获的图像进行测量。应用程序可以在执行该过程的时候将该过程可视化,如图7所示。

[0227] 用户交互可触发绝对比例估算过程。用户交互可以是按下按钮、触摸屏幕、语音识别和/或动作识别。

[0228] 在另一实施方案中,该方法在不与用户界面进行任何交互的情况下执行。在此实施方案中,该方法在后台执行并且在相机运动适合于进行比例估算时确定比例估算值。

[0229] 此外,本发明涉及一种在计算机上实施的用户交互方法,如本文所述,使用户与根据本发明的方法进行交互。

[0230] 用户输入可包括一次或多次用户交互。用户交互可以是以下一者:说话、奔跑、跳跃、眨眼睛、和/或移动用户身体任何部位。用户交互也可以是按下按钮、触摸屏幕、对麦克风说话、注视、或打手势。用户交互也可以是:将实物放入相机的视野中,使得根据相机捕获的实物的至少一部分的图像可识别实物。

[0231] 用户输入则也可以是在按住按钮不放的同时可进行的特定相机运动。此类特定运动可以是:面朝与背离脸部移动相机、或上下移动相机、或左右移动相机。

[0232] 移动手机,如本文可用,包含至少一个相机,用于捕获图像。移动手机还包括处理设备,可如本文所述用于执行根据本发明的多个步骤中的任意步骤。移动设备也包括触摸屏,触摸屏可显示图形用户界面,使得用户可触摸或按下物理按钮或图形用户界面上所显

示的按钮。

[0233] 本发明的实施方案在本文参考使用移动设备或手持设备的情况下进行描述,诸如,移动手机,但是原则上,本发明可以与任何处理设备一起应用,诸如在包括一个或多个微处理器的常用计算机设备中实施,用于执行相应步骤(诸如可佩带式计算机、平板计算机、移动计算机、所谓的笔记本电脑、或头戴式显示器,诸如用于光学透视式增强现实应用程序的计算机设备)。本文所述的这些步骤也可以由处理装置网络(诸如计算机网络)和/或与服务器计算机进行通信的移动设备网络来执行。

[0234] 本发明的实施方案可用于各种应用程序中,包括:增强现实应用程序,实现了在真实环境中放置与显示虚物;导航应用程序,使用相机来估算用户的位置和/或运动;模拟所捕获物体在空间上与其它物体或环境的关系的程序;或测量应用程序,旨在测量处于绝对比例的重构物体上的多个点之间的距离。

[0235] 例如,本发明的实施方案可用于增强现实应用程序中。处于绝对比例的实物的重构件作为基础条件,用于根据在相机图像与实物重构模型之间建立2D-3D对应关系,确定相机相对于物体的位姿。此类增强现实应用程序可以在实时视图中叠加虚拟3D物体,诸如虚拟椅子,使得物体看上去相对于实物静止。这需要掌握相对于实物的相机位姿(即位置和取向)。由于根据本发明的实施方案,实物的重构件以绝对比例进行的定义,同样地位姿也可以绝对比例进行估算,这实现了以绝对比例对虚物进行叠加。由此,如果实物为桌子且虚物为椅子,则将虚拟椅子放到真实桌子旁边时,虚拟椅子将以与真实桌子同等的大小出现。相反,当使用处于任意比例的重构件时,椅子可以是桌子高度的5倍或桌子高度的五分之一,这两种情况都导致了不合实际的外观。在应当为用户提供机会以通过叠加虚拟模型来评估某物体是否在空间上适合真实环境的增强现实应用程序中,具有处于绝对比例的实物或环境的重构件尤为重要。例如,此类应用可用于从视觉上评估沙发是否适合某起居室或其是否太大。在不了解如发明所提供的绝对比例的情况下,这是不可能的。

[0236] 在另一实施例中,本发明的实施方案可用于测量应用程序中。通过本发明的实施方案以绝对比例进行实物重构。软件应用程序则可为用户提供以下选项:选择重构件上的点,例如,通过点击设备(诸如,鼠标)点击这些点的方式。应用程序则可计算两个所选点之间的(欧几里德)距离并向用户显示该距离。如果此类应用程序作用于处于绝对比例的重构件,如本发明所提供,则计算出的距离也处于绝对比例,例如,毫米。反之,当使用处于任意比例的重构件时,重构件上所测得的距离只可用于比较该物体上彼此的相对距离,但是所有测量值将与绝对单位无关,诸如毫米或英寸。

[0237] 在另一实施例中,本发明的实施方案可用于视觉型导航中,用于根据视觉视觉测程法估算用户的运动,以定位用户并提供反馈。为了更新相对于坐标系的位置,其需要估算处于绝对比例的运动,在该坐标系存储了地图和导航信息。没有绝对比例,则只能确定所覆盖的轨迹,不能确定实际所覆盖的真实距离信息。同样地,由于比例的变化,一段时间后该形状可严重变形。室内导航由此需要测量值保持长期稳健和一致。同样地,速度也很重要,由于速度是指距离除以时间,其直接取决于重构件的比例。没有处于绝对比例的重构件,通过运动估算所获得的速度信息也不处于绝对比例。

[0238] 即使相对于实际距离的绝对关系未知,其它应用程序也可从可重复的重构件比例中受益。该可重复性通过所提出的方法来实现。当(例如)单独为场景的若干部分创建重构

件时,需要各个图以相同的比例进行定义。这使合并场景的各个部分更加容易。可重复比例也实现了:克服更长时间的重构过程中可能发生的比例变化问题。

[0239] 一般来说,下面给出了术语的进一步解释,以下其它方面和实施方案可结合本发明的多个方面进行应用。

[0240] 如本文所用,相机是用于捕获实物的图像信息的图像捕获设备。可提供或校准至少一个相机或多个相机的固有相机参数。可在如下情况下应用本发明:从提供实物图像的任何相机接收图像信息。并不限于用于提供RGB格式的彩色图像的相机。也可适用于任何其它颜色格式并且也适用于单色图像,例如,适用于提供灰度格式图像的相机。所用的相机还可提供具有深度数据的图像。无需在与(彩色/灰度)图像相同的分辨率下提供深度数据。提供具有深度数据的图像的相机通常被称为RGB-D相机。RGB-D相机系统可以是渡越时间(TOF)相机系统或使用结构光的相机系统。至少一个相机或多个相机也可捕获人眼不可见的光。例如,至少一个相机或多个相机可以是捕获红外线的红外相机。

[0241] 实物可以是在实际中具有物理几何形状的任何物体。实物还可包括一个或多个实物。例如,实物可以是房间或车辆的内部。实物也可以使,例如,椅子,大楼、山峰、树木或桌子。房间的内部(作为实物)还可包括椅子和/或桌子。在本文,实物也可称为真实环境或真实场景。实物也可以是多个实物的排列。例如,室内物体可以是椅子、桌子和沙发的排列。

[0242] 真实环境的几何模型(3D图形或也称作3D重构件)可通过2D观察值的三角测量进行创建,2D观察值存在于一个或多个相机捕获的多张图像中。三角测量是3D重构程序中常用的方法,其基于相机图像,也称作运动恢复结构(SfM),见参考文献[2]。

[0243] 相机位姿说明了特定位置上的相机与参考坐标系之间的空间关系或刚性变换。参考坐标系可以与实物或与另一位置上的相机相关。空间关系或刚体变换说明了至少一次平移、或至少一次旋转、或它们在3D空间中的组合、或至少一个距离。

[0244] 3D特征表示或说明了相应实物或至少一部分相应实物的物理3D特征。3D特征,例如,不限于:点、边、线、段、角和/或任何其它几何形状。

[0245] 为了说明物体的几何结构,点的位置、其它几何元素、和它们的像(例如)方向等其它特性在坐标系中进行唯一确定。特定元素(例如,位置)由有序数字或坐标元组进行定义。

[0246] 我们在下面称为笛卡尔坐标系,记住:其它坐标系中的几何图形的表示方式(像,例如,极坐标或齐次坐标)可通过坐标变换转化为笛卡尔坐标系,反之亦可,坐标变换说明了坐标系之间的关系并且为通过其它坐标系中的对应坐标表示一个坐标系中的特定坐标系提供了公式。

[0247] 在笛卡尔坐标系中,坐标元组的每个元素对应于点到相应超平面的符号距离。两个坐标元组之间的距离本身可定义为欧几里德距离,说明了连接两个坐标元组的线段长度。该距离本身由此也作为1D坐标给出。

[0248] 应当定义沿着坐标轴的单位的意思。该测量单位是数量,作为表示该特性(例如,长度)的现有数量的因子。当定义了测量值的坐标系单位与实际空间参考单位之间的关系已知时,我们对测量值(例如,定义测量值位置的空间坐标、以及分别定义空间距离以及空间坐标之间的长度的坐标)进行表征,为其指定绝对空间单位,也称为绝对空间比例。实际空间参考单位可以(例如)是米(米制)或任何其它具有转化成米的固定已知转化率的单位。

[0249] 测量值可通过以下方式处于绝对比例下:设定坐标系的空间基本单位,其中测量

值直接以实际空间参考单位进行定义;或者指定坐标系相对于实际空间参考单位的空间比例。

[0250] 不具有绝对空间单位但是处于任意比例的坐标系中的测量值的示例(例如)通过重构3D点给出,3D点重构通过三角测量两张图像中对应的点特征而进行,两张图像在空间中的不同相机位姿处进行捕获,其中这两个相机位姿之间的基线长度未知。尽管坐标轴单位相同,但是单位本身并未确定。即,尽管坐标系内两段距离之比是正确的,但是整个坐标系相对于实际的绝对比例未知。坐标系以及重构件本身据此认为是合乎比例的。绝对比例并不是指物体的绝对平移补偿值,这取决于坐标系的原点,但是是指绝对大小。

[0251] 脸部特征和基座点可包括:位置(角落位置、中心位置、或边界区)、大小、形状、轮廓、区域、比例、比率、左眼与右眼之间的距离以及左眼和右眼外观(瞳孔、虹膜、角膜、巩膜、内眦、外眦、中心、上下眼睑、睫毛,……)、鼻梁、鼻子(鼻尖、鼻背、鼻翼、鼻孔鼻小柱,……)(大小、形状)、人中、嘴唇、左右耳、左右眼眉、牙齿、左右脸颊、下颌、脖子、喉结、皮肤结构和坚固性(像毛孔)、面部和头部毛发等。

[0252] 视觉测程法是指通过分析相关相机图像来确定相机位置与取向的方法。当提供了与一张相机图像中至少一个像素相关的具有绝对比例的深度信息(例如,成像表面上一个像素的距离,单位为毫米)时,则视觉测程法可确定处于绝对空间比例的相机位姿(即,位置和取向)。术语视觉测程法常常和SLAM互换使用。

[0253] 参考文献:

[0254] 1.Davison,Andrew J.等人“MonoSLAM:Real-time single camera SLAM.”*Pattern Analysis and Machine Intelligence,IEEE Transactions on* 29.6(2007):第1052页-1067页。

[0255] 2.Hartley,Richard和Andrew Zisserman*Multiple view geometry in computer vision.*第2卷.Cambridge,2000.

[0256] 3.Azuma,Ronald等人“Recent advances in augmented reality.”*Computer Graphics and Applications,IEEE* 21.6(2001):第34-47页。

[0257] 4.Strasdat,Hauke,J.M.M.Montiel,和Andrew J.Davison.“Scale drift-aware large scale monocular SLAM.”*Proceedings of Robotics:Science and Systems(RSS).*第2卷.No.3.2010.

[0258] 5.Lemaire,Thomas等人“Vision-based slam:Stereo and monocular approaches.”*International Journal of Computer Vision* 74.3(2007):第343-364页。

[0259] 6.Lieberknecht,Sebastian等人“RGB-D camera-based parallel tracking and meshing.”*Mixed and Augmented Reality(ISMAR),2011 10th IEEE International Symposium on.IEEE,2011.*

[0260] 7.Klein,Georg,和David Murray.“Parallel tracking and mapping for small AR workspaces.”*Mixed and Augmented Reality,2007.WACV 2007.6th IEEE and ACM International Symposium on.IEEE,2007.*

[0261] 8.Castle,Robert,Georg Klein,和David W.Murray.“Video-rate localization in multiple maps for wearable augmented reality.”*Wearable Computers,2008.ISWC 2008.12th IEEE International Symposium on.IEEE,2008.*

- [0262] 9.Nützi,Gabriel等人“Fusion of IMU and vision for absolute scale estimation in monocular SLAM.”Journal of intelligent&robotic systems 61.1-4 (2011):第287-299页。
- [0263] 10.Umeyama,Shinji.“Least-squares estimation of transformation parameters between two point patterns.”Pattern Analysis and Machine Intelligence,IEEE Transactions on 13.4(1991):第376-380页。
- [0264] 11.Esquivel,Sandro,Felix Woelk和Reinhard Koch.“Calibration of a multi-camera rig from non-overlapping views.”Pattern Recognition.Springer Berlin Heidelberg,2007.第82-91页。
- [0265] 12.Han,Hu等人“Demographic Estimation from Face Images:Human vs.Machine Performance.”
- [0266] 13.Dodgson,Neil A.“Variation and extrema of human interpupillary distance.”Proceedings of SPIE.第5291卷.2004.
- [0267] 14.Clipp,Brian等人“Robust 6dofmotion estimation for non-overlapping, multi-camera systems.”Applications of Computer Vision,2008.WACV 2008.IEEE Workshop on.IEEE,2008
- [0268] 15.Ren,Shaoqing等人“Face Alignment at 3000FPS via Regressing Local Binary Features.”
- [0269] 16.Martins,Pedro和Jorge Batista.“Accurate single view model-based head pose estimation.”Automatic Face&Gesture Recognition,2008.FG’08.8th IEEE International Conference on.IEEE,2008.
- [0270] 17.Asthana,Akshay等人“Incremental Face Alignment in the Wild.”Computer Vision and Pattern Recognition(CVPR),2014IEEE Conference on.IEEE, 2014.
- [0271] 18.Wang,Peng等人“Automatic eye detection and its validation.”Computer Vision and Pattern Recognition-Workshops,2005.CVPR工作室.IEEE Computer Society Conference on.IEEE,2005.
- [0272] 19.Turk,Matthew A.和Alex P.Pentland.“Face recognition using eigenfaces.”Computer Vision and Pattern Recognition,1991.Proceedings CVPR’ 91.,IEEE Computer Society Conference on.IEEE,1991.
- [0273] 20.Daniel Kurz,Selim Behimane“Method of providing a descriptor for at least one feature of an image and method of matching features”US 20120219188 A1.

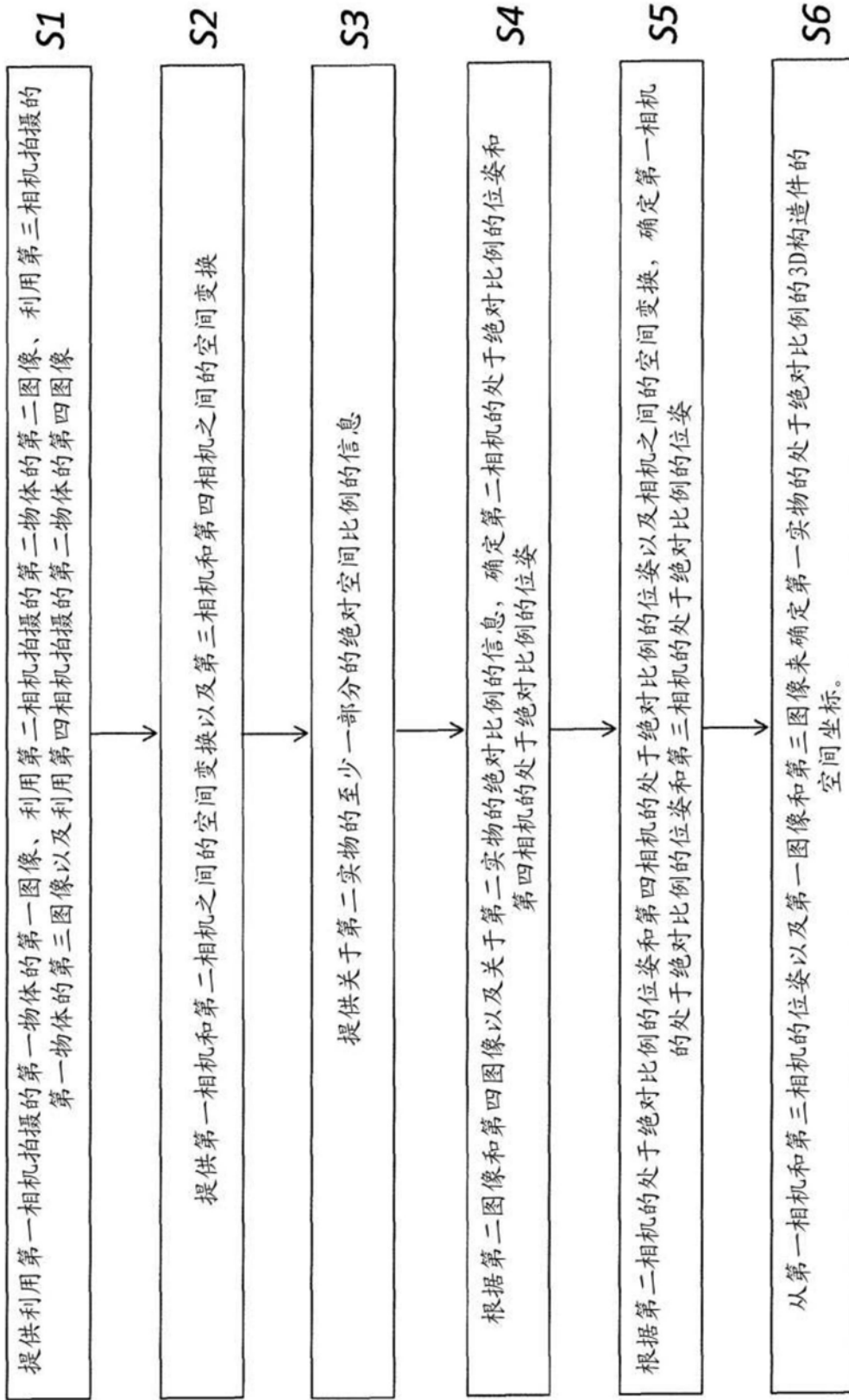


图1



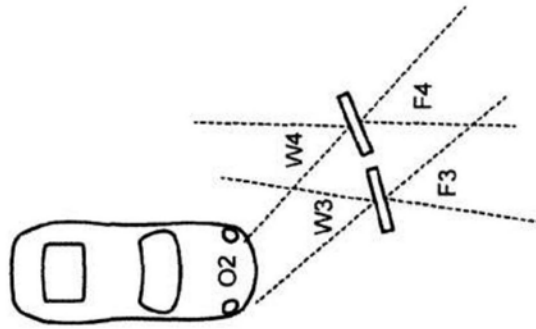
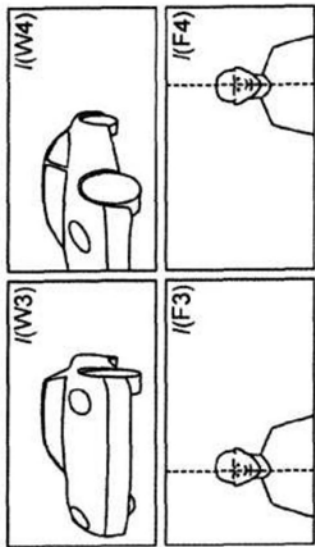


图 2b

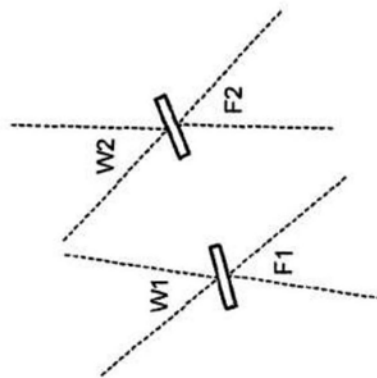
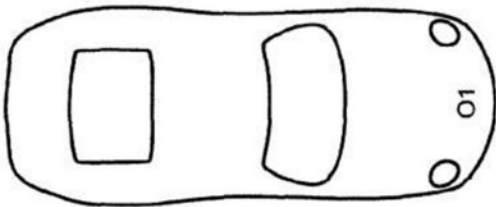
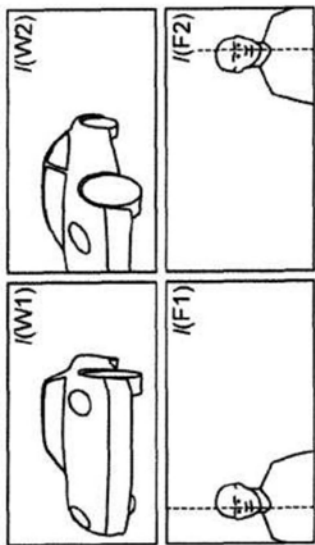


图 2a

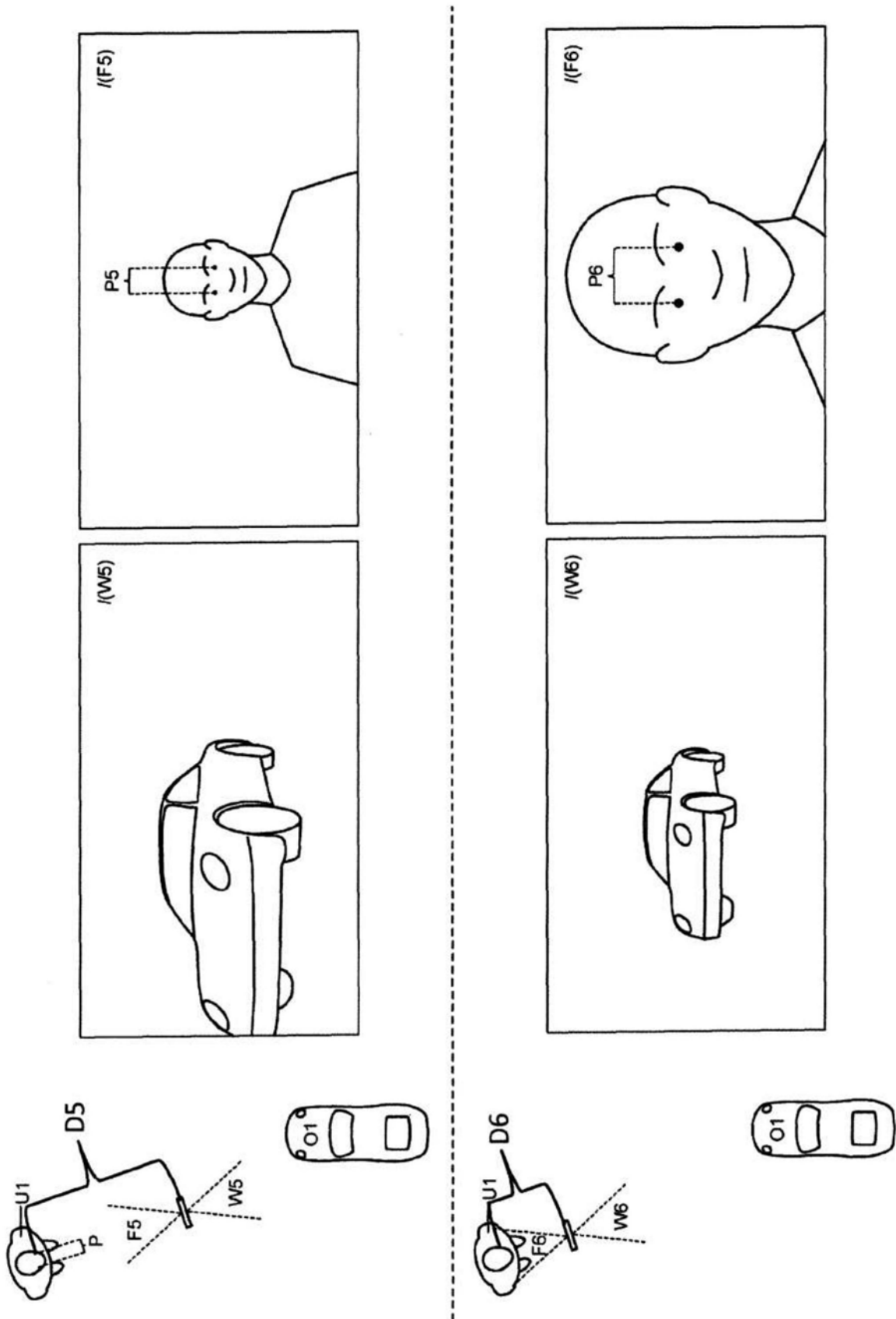


图3

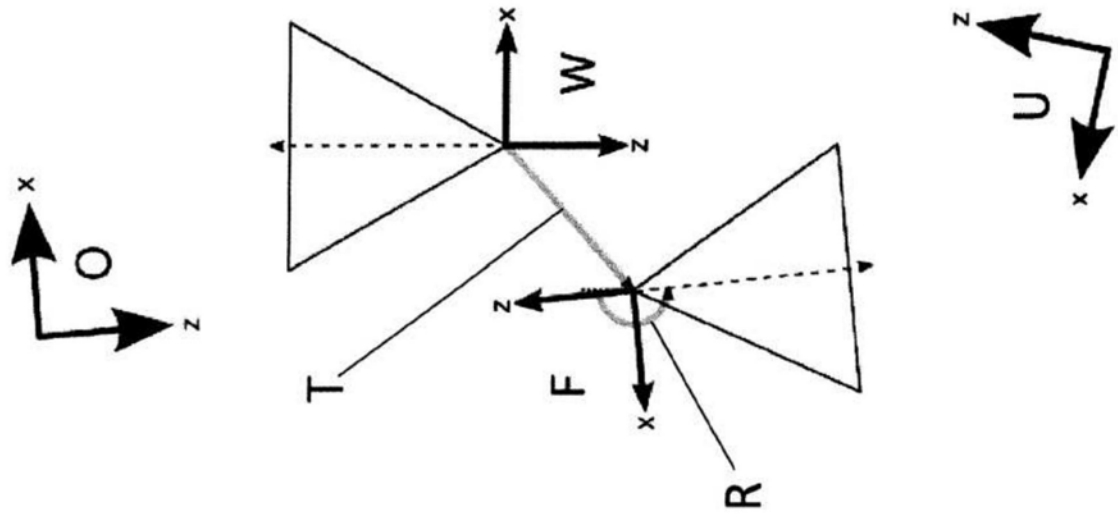


图4

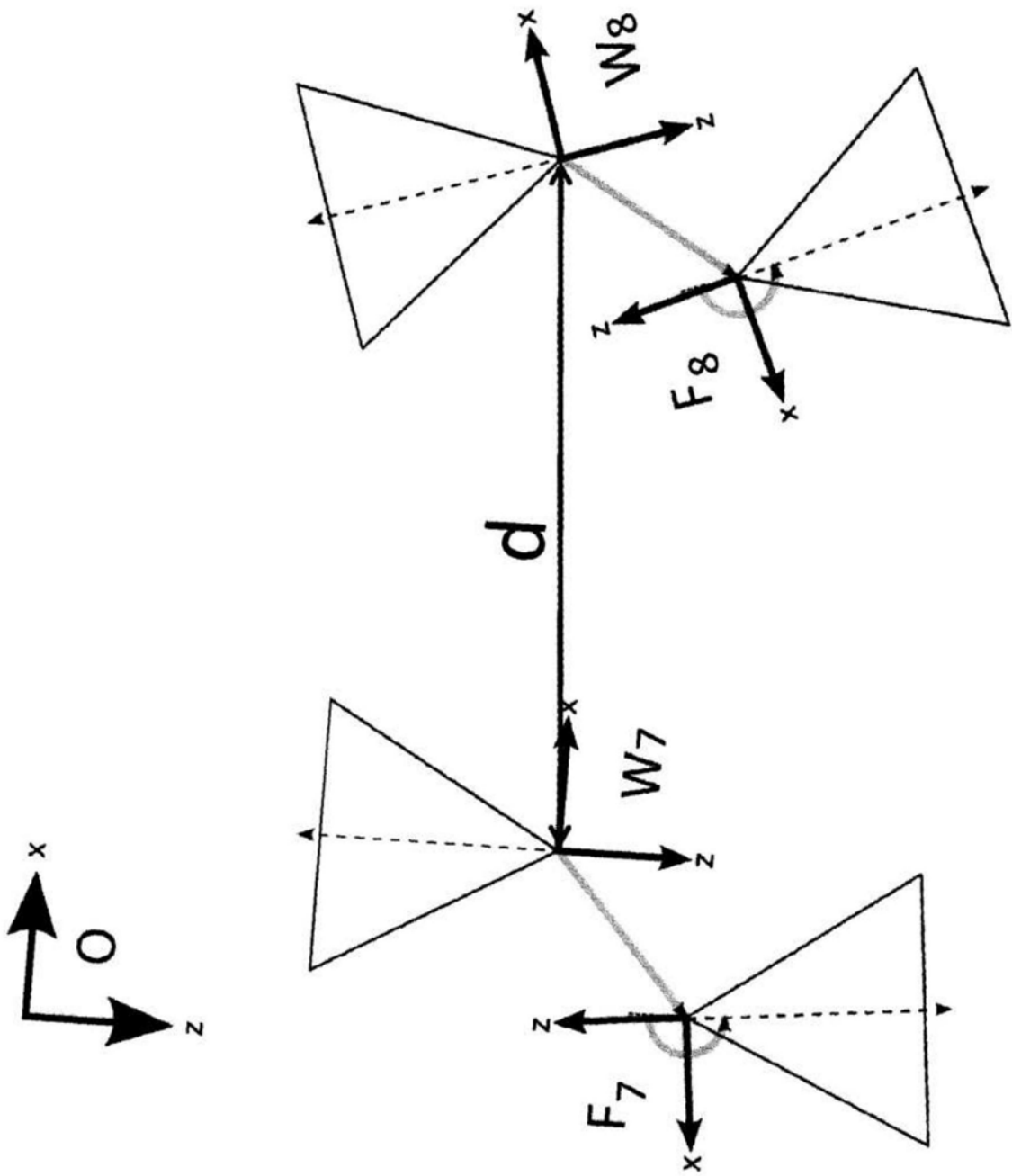


图5

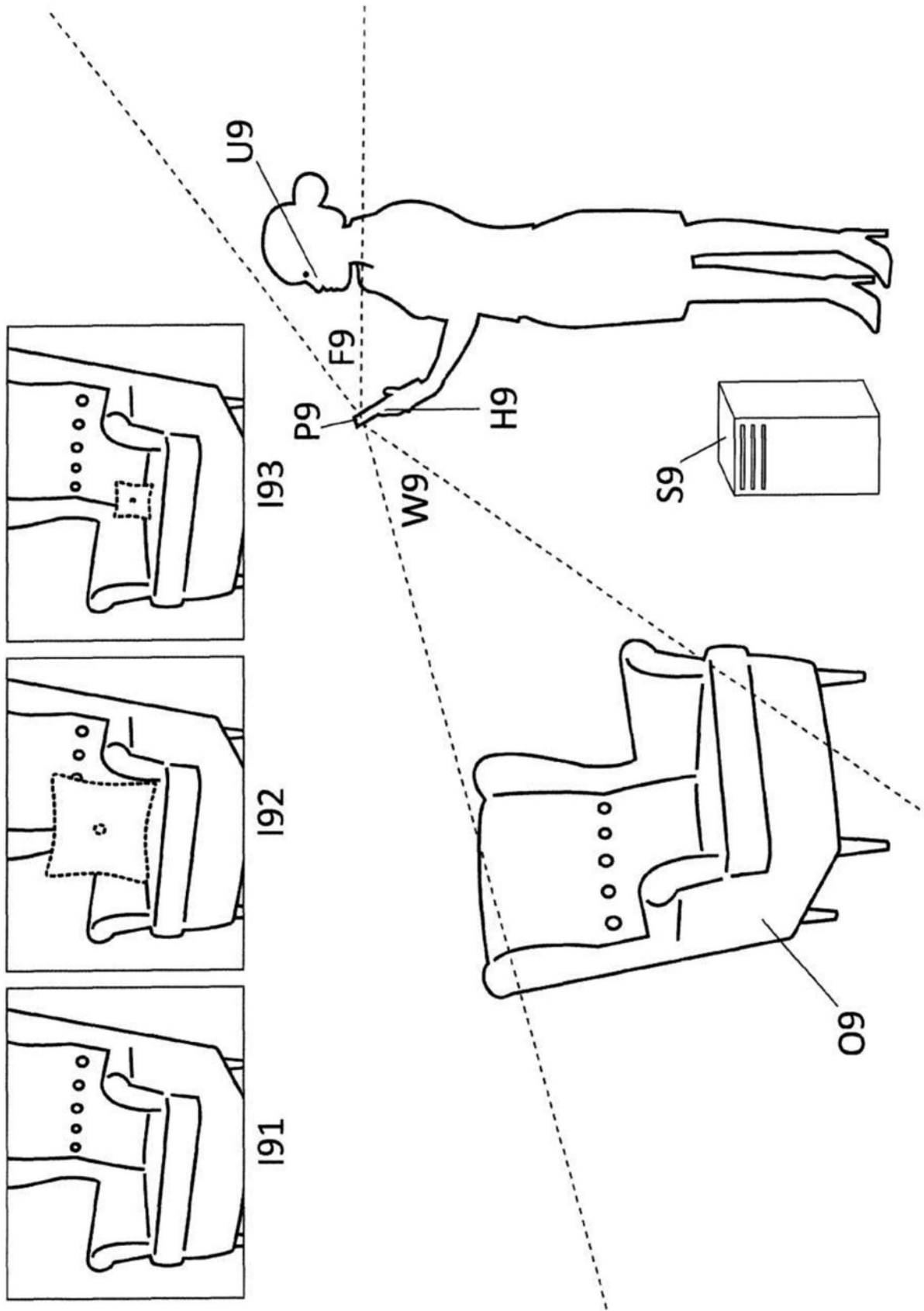


图6

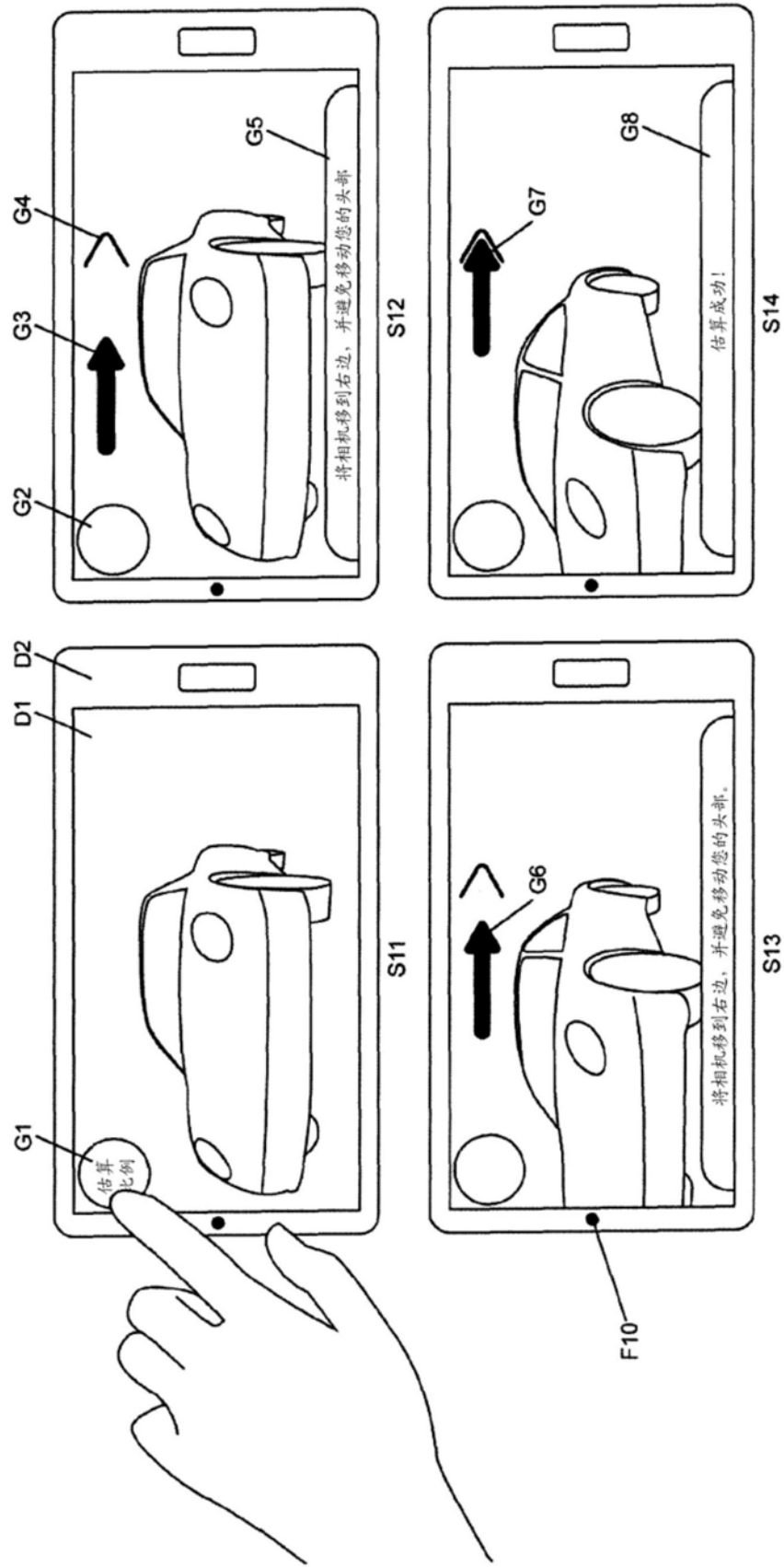


图7

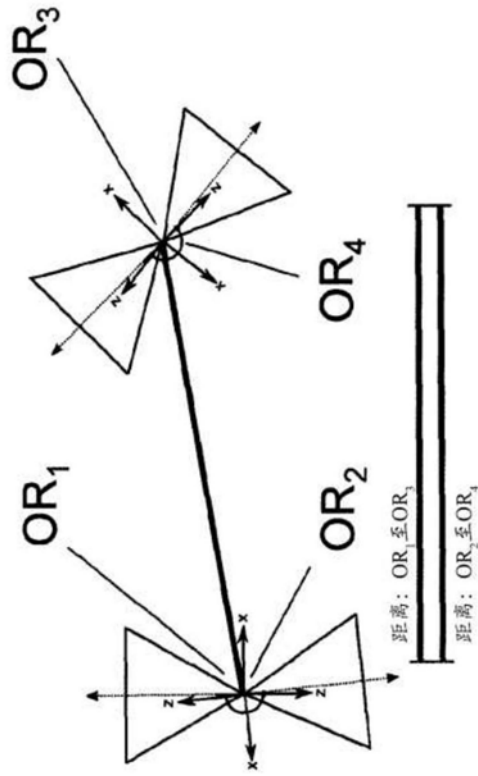


图8a

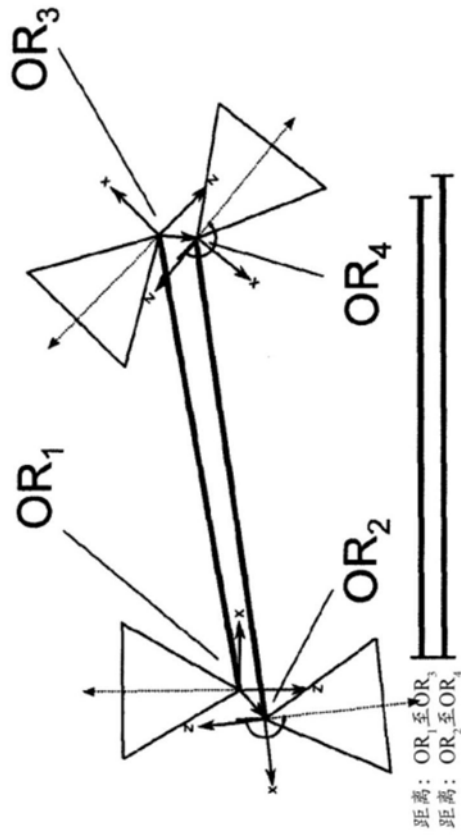


图8b



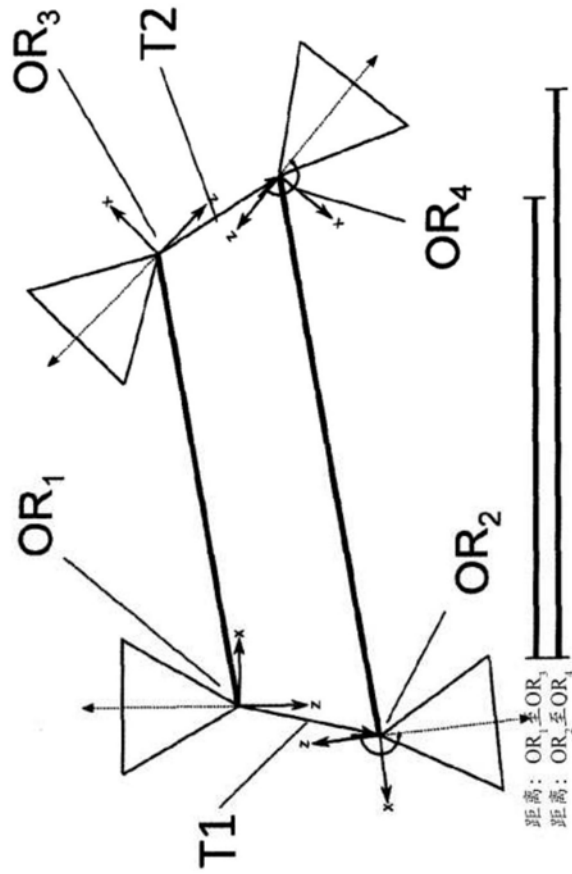


图8c

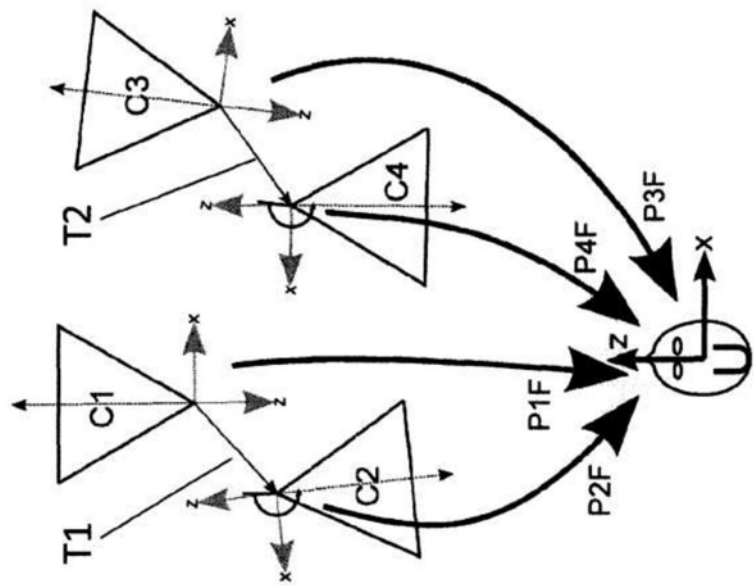
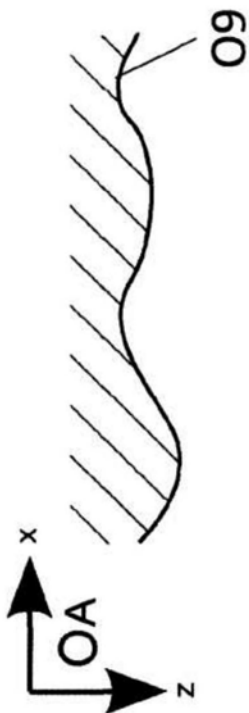


图9

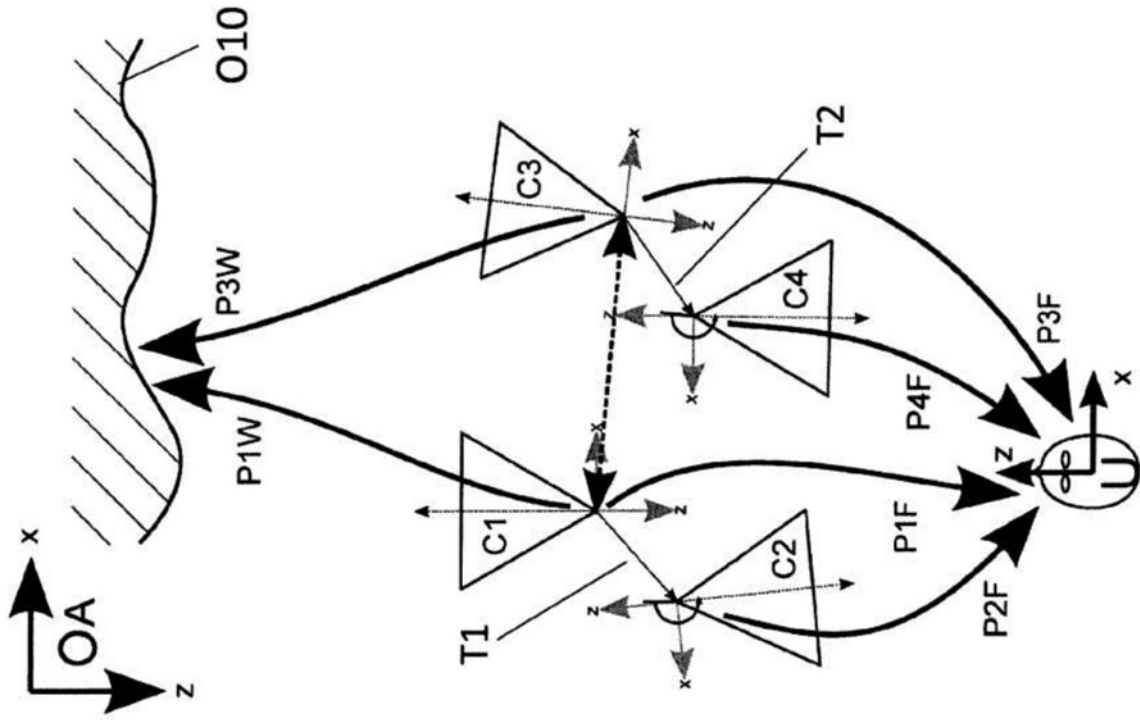


图10

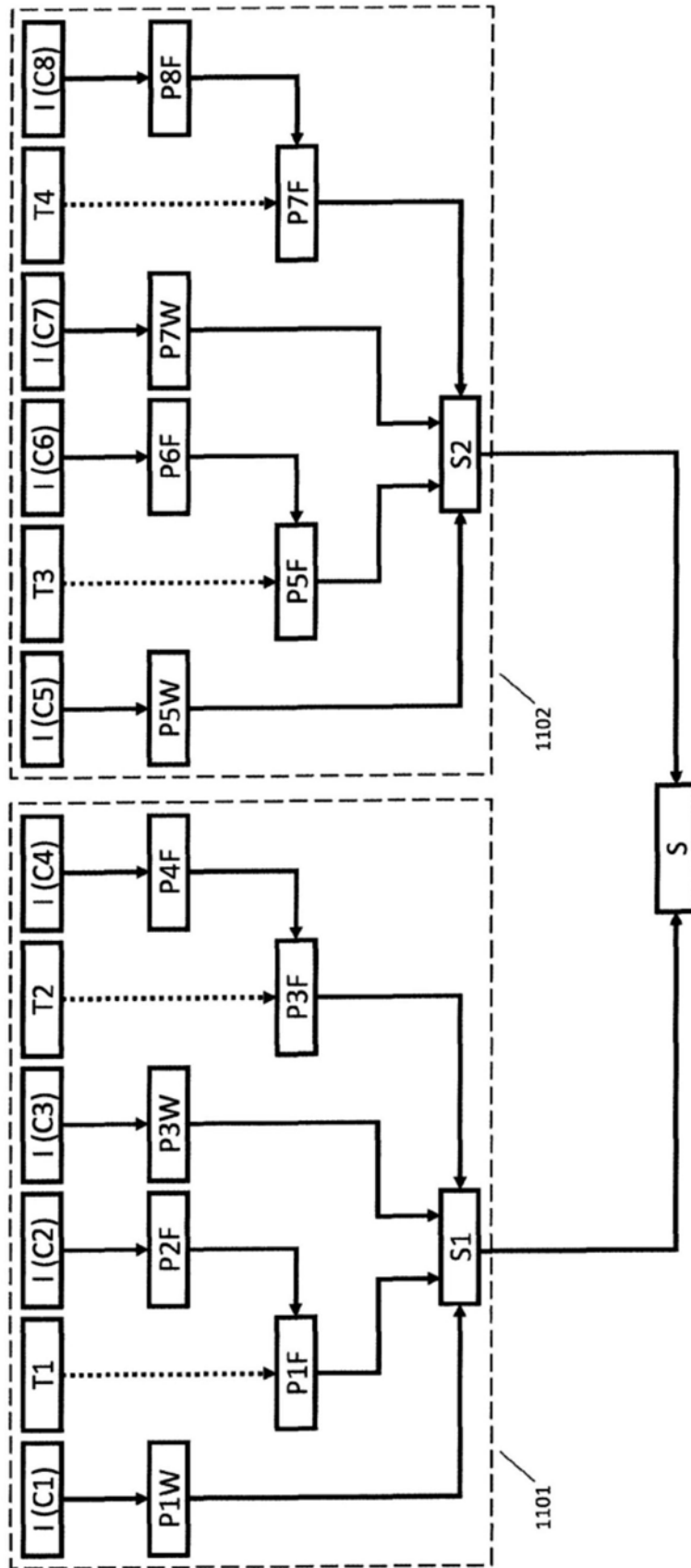


图11

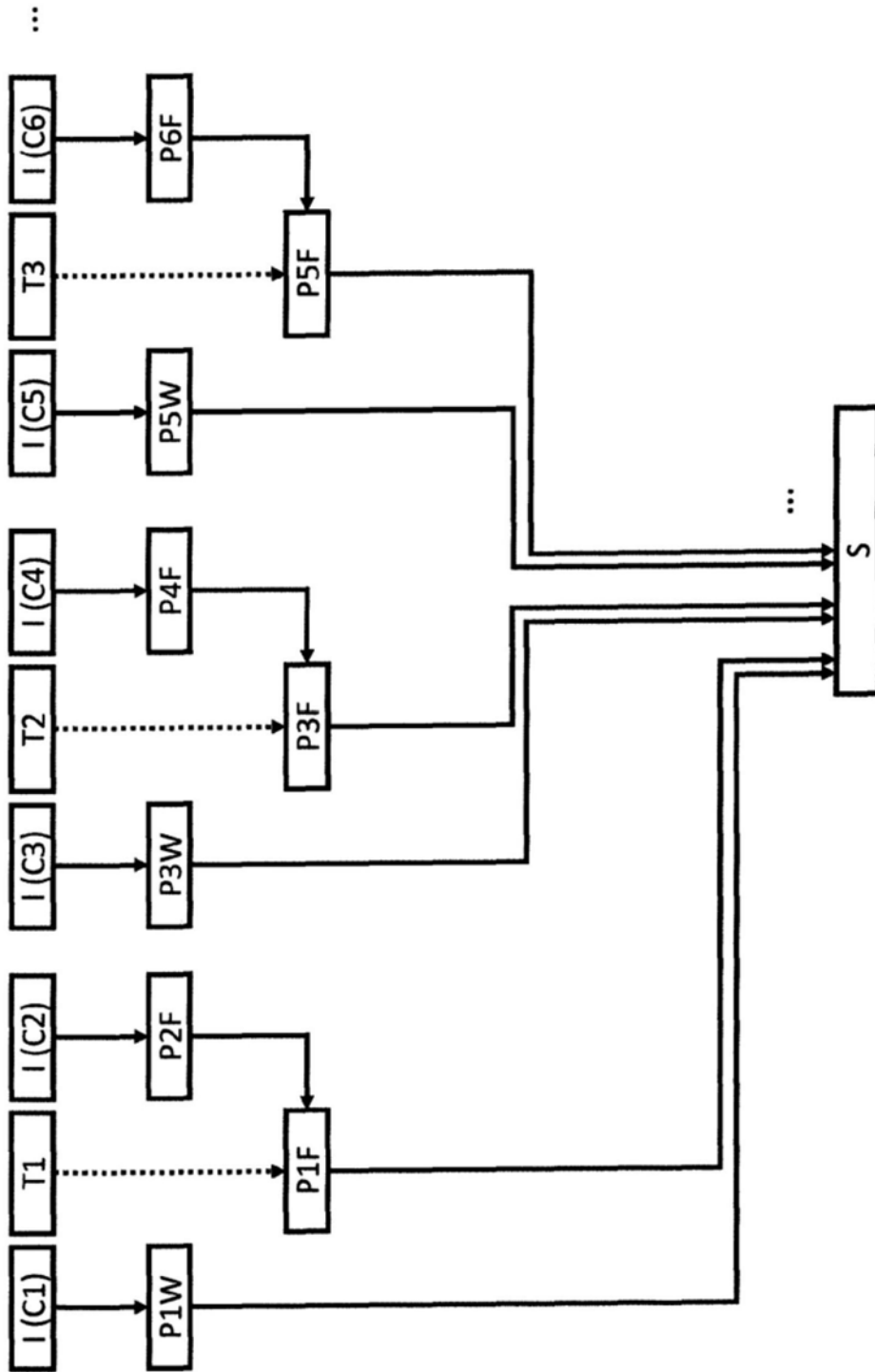


图12