



(12)发明专利

(10)授权公告号 CN 104106262 B

(45)授权公告日 2017.01.18

(21)申请号 201380008735.2

(22)申请日 2013.01.14

(65)同一申请的已公布的文献号

申请公布号 CN 104106262 A

(43)申请公布日 2014.10.15

(30)优先权数据

13/369,168 2012.02.08 US

(85)PCT国际申请进入国家阶段日

2014.08.08

(86)PCT国际申请的申请数据

PCT/US2013/021353 2013.01.14

(87)PCT国际申请的公布数据

W02013/119352 EN 2013.08.15

(73)专利权人 微软技术许可有限责任公司

地址 美国华盛顿州

(72)发明人 Z·刘 Z·张 Z·李

(74)专利代理机构 上海专利商标事务所有限公司 31100

代理人 顾嘉运

(51)Int.Cl.

H04N 13/00(2006.01)

G06T 7/00(2006.01)

审查员 李萍

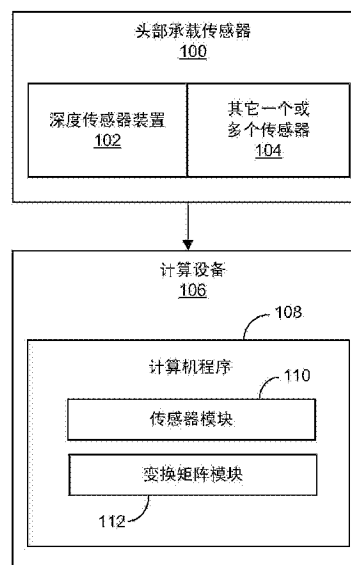
权利要求书6页 说明书18页 附图13页

(54)发明名称

使用深度相机的头部姿态跟踪

(57)摘要

呈现了使用被配置以放置在用户的头部上的一组传感器的头部姿态跟踪技术的实施例。该组传感器包括深度传感器装置,它和至少一个其它类型的传感器被用于标识场景内的特征的三维位置。该组传感器中的每个传感器的输出的数据是周期性输入,且每次输入数据时,它被用于计算变换矩阵,当该变换矩阵被应用于在当输入第一传感器数据时所建立的先前确定的头部姿态位置和朝向时标识了当前的头部姿态位置和朝向。随后,将该变换矩阵应用于先前确定的头部姿态位置和朝向以标识当前的头部姿态位置和朝向。



1. 一种用于头部姿态跟踪的系统,包括:

被配置以放置在用户的头部上的一组传感器,该组传感器包括被用于标识在由该组传感器所感测到的场景内的点的三维位置的深度传感器装置,以及彩色摄像机,所述深度传感器装置和所述彩色摄像机被同步以周期性地以深度帧和彩色图像帧形式产生同时期的场景数据,并被校准以便将每个彩色图像帧中的每个像素映射到同时期产生的深度帧中的对应的三维场景位置;

计算设备;以及

包括由所述计算设备执行的程序模块的计算机程序,其中所述计算设备由所述计算机程序的程序模块来指示以便:

输入每个同时期产生的深度帧和彩色图像帧,

对于在第一之后输入的每个同时期产生的深度帧和彩色图像帧对,

标识最后输入的彩色图像帧和紧接在所述最后输入的彩色图像帧之前产生的彩色图像帧之间的匹配特征,

使用所标识的匹配特征和在所述最后输入的彩色图像帧和紧接在所述最后输入的彩色图像帧之前产生的彩色图像帧这两者中的所述匹配特征的对应的三维位置估计第一变换矩阵,所述第一变换矩阵定义了点从紧接在所述最后输入的彩色图像帧之前产生的彩色图像帧到所述最后输入的彩色图像帧的平移和旋转,

估计定义点从第一输入彩色图像帧到最后输入彩色图像帧的平移和旋转的最终变换矩阵,所述估计通过将定义了点从第一输入彩色图像帧到紧接在所述最后输入的彩色图像帧之前的输入彩色图像帧的平移和旋转的先前计算的变换矩阵与定义了点从所述紧接在所述最后输入的彩色图像帧之前的输入彩色图像帧到该最后输入的彩色图像帧的平移和旋转的第一变换矩阵进行累加来完成,以及

将所述最终的变换矩阵应用于在第一输入彩色图像帧中描述的场景内的先前确定的头部姿态位置和朝向以标识在所述最后输入彩色图像帧中描述的场景内的当前的头部姿态位置和朝向。

2. 如权利要求1所述的系统,其特征在于,所述用于标识匹配特征的程序模块包括用于使用光流技术标识在所述最后输入的彩色图像帧和紧接所述最后输入的彩色图像帧之前产生的彩色图像帧之间的匹配点的子模块。

3. 如权利要求2所述的系统,其特征在于,还包括:在执行所述用于估计所述第一变换矩阵的程序模块之前就执行的深度校正程序模块,所述深度校正程序模块包括用于以下操作的子模块:

定义围绕每个所标识的匹配点的窗口;

标识每个窗口内的离所述彩色摄像机的最接近的点;以及

将所标识的每个窗口内的最接近的点指定为替代与所述窗口相关联的先前所标识的匹配点的匹配点。

4. 如权利要求2所述的系统,其特征在于,还包括:在执行所述用于估计所述第一变换矩阵的程序模块之前就执行的匹配点异常值解除指定程序模块,所述匹配点异常值解除指定程序模块包括将具有落在规定最小等级之下的匹配质量等级的每对所标识的匹配点解除作为匹配点的指定。

5. 如权利要求2所述的系统,其特征在于,还包括:在执行所述用于估计所述第一变换矩阵的程序模块之后且在执行所述用于估计所述最终的变换矩阵的程序模块之前执行的匹配点异常值解除指定程序模块,所述匹配点异常值解除指定程序模块包括用于以下操作的子模块:

(a)将最后估计的变换矩阵应用于在最后输入的彩色图像帧中所标识的匹配点以查明在紧接最后输入的彩色图像帧之前产生的彩色图像帧中与最后输入的彩色图像帧中的匹配点相对应的匹配点;

(b)对于在紧接最后输入的彩色图像帧之前产生的彩色图像帧中的每个先前所标识的匹配点,计算在所述先前所标识的匹配点的坐标和通过将最后估计的变换矩阵应用于在最后输入的彩色图像帧中的匹配点所查明的该匹配点的坐标之间的差异度量,所述在最后输入的彩色图像帧中的匹配点与在紧接最后输入的彩色图像帧之前产生的彩色图像帧中先前所标识的匹配点相对应;

(c)计算剩余的匹配点对的差异度量的平均并确定所计算的平均是否超出规定最小等级;

(d)只要所计算的平均超出规定最小等级,

按量级排序所述差异度量并从最大开始向下以量级标识规定百分比的差异度量,

将与所标识的差异度量之一相关联的每个匹配点对的在紧接最后输入的彩色图像帧之前产生的彩色图像帧中先前所标识的匹配点和其在最后输入的彩色图像帧中的对应的匹配点解除作为匹配点的指定,

使用剩余的匹配点重新估计所述最后估计的变换矩阵;

(e)确定是否已经达到规定的变换矩阵重新估计的迭代的最大数目;以及

(f)只要还没有达到规定的变换矩阵重新估计的迭代的最大数目,重复子模块(a)到(f)的执行。

6. 如权利要求2所述的系统,其特征在于,还包括:在执行所述用于标识在最后输入的彩色图像帧和紧接在最后输入的彩色图像帧之前产生的彩色图像帧之间的匹配特征的程序模块之前就执行的关键帧程序模块,所述关键帧程序模块包括用于以下操作的子模块:

确定是否要建立新的关键帧,以及

只要确定要建立新的关键帧,记录所述最后输入的彩色图像帧并将所记录的帧指定为关键帧。

7. 如权利要求6所述的系统,其特征在于,所述用于确定是否要建立新的关键帧的子模块包括只要最后输入的彩色图像帧是所述第一输入彩色图像帧,以及只要所述最后输入的彩色帧描述少于先前建立的关键帧中所描述的场景的规定部分,就确定要建立新的关键帧。

8. 如权利要求6所述的系统,其特征在于,所述关键帧程序模块还包括用于以下操作的子模块:

只要确定不要建立新的关键帧,就确定是否要发生关键帧匹配;

只要确定要发生关键帧匹配或已经建立了新的关键帧,

标识在所述最后输入的彩色图像帧和所记录的描述了最后输入的彩色图像帧中所描述的场景的至少规定部分的关键帧之间的关键帧匹配点,

使用所标识的关键帧匹配点和在所述最后输入的彩色图像帧和所记录的描述了最后输入的彩色图像帧中所描述的场景的至少规定部分的关键帧两者中的关键帧匹配特征的对应的三维位置来估计第二变换矩阵,所述第二变换矩阵定义了点从所记录的描述了最后输入的彩色图像帧中所描述的场景的至少规定部分的关键帧到所述最后输入的彩色图像帧的平移和旋转,

使用所述第二变换矩阵和定义了点从第一输入彩色图像帧到所记录的描述了最后输入的彩色图像帧中所描述的场景的至少规定部分的关键帧的平移和旋转的先前计算的关键帧变换矩阵来估计第三变换矩阵,所述第三变换矩阵定义了点从所述第一输入彩色图像帧到所述最后输入彩色图像帧的平移和旋转,

只要所述最后输入的彩色图像帧是新建立的关键帧,就将所述第三变换矩阵指定为定义了点从所述第一输入彩色图像帧到新建立的关键帧的平移和旋转的关键帧变换矩阵,

估计定义点从第一输入彩色图像帧到最后输入彩色图像帧的平移和旋转的最终变换矩阵,所述估计通过将定义了点从第一输入彩色图像帧到紧接在所述最后输入的帧之前的输入彩色图像帧的平移和旋转的先前计算的变换矩阵与定义了点从所述紧接最后输入的彩色图像帧之前的帧到所述最后输入的彩色图像帧的平移和旋转的第一变换矩阵进行累加来完成,

将所述第三和最终的变换矩阵合并以产生定义了点从第一输入彩色图像帧到最后输入的彩色图像帧的平移和旋转的单个经组合的变换矩阵,以及将经组合的变换矩阵指定为所述最后输入的彩色图像帧的最终变换矩阵。

9.如权利要求8所述的系统,其特征在于,还包括:

在执行用于标识关键帧匹配点的子模块之后且在执行用于估计第二变换矩阵的子模块之前执行的深度校正子模块,所述深度校正程序子模块包括用于以下操作的子模块:

定义围绕每个所标识的关键帧匹配点的窗口;

标识每个窗口内的离所述彩色摄像机的最接近的点;以及

将所标识的每个窗口内的最接近的点指定为替代与所述窗口相关联的先前所标识的关键帧匹配点的关键帧匹配点,

在执行所述用于深度校正的子模块之后且在执行用于估计第二变换矩阵的子模块之前执行的关键帧匹配点异常值解除指定子模块,所述关键帧匹配点异常值解除指定子模块包括将具有落在规定最小等级之下的匹配质量等级的每对所标识的关键帧匹配点解除作为关键帧匹配点的指定,以及

在执行所述用于估计第二变换矩阵的子模块之后且在执行所述用于估计第三变换矩阵的子模块之前执行的第二关键帧匹配点异常值解除指定子模块,所述第二关键帧匹配点异常值解除指定子模块包括用于以下操作的子模块:

(a)将所述第二变换矩阵应用于在最后输入的彩色图像帧中所标识的关键帧匹配点以查明在所记录的描述了在所述最后输入的彩色图像帧中所描述的场景的至少规定部分的关键帧中与在最后输入的彩色图像帧中的关键帧匹配点相对应的关键帧匹配点;

(b)对于在所记录的描述了在所述最后输入的彩色图像帧中所描述的场景的至少规定部分的关键帧中的每个先前标识的关键帧匹配点,计算在所述先前所标识的匹配点的坐标和通过将所述第二变换矩阵应用于在最后输入的彩色图像帧中与在所记录的描述了在所

述最后输入的彩色图像帧中所描述的场景的至少规定部分的关键帧中的先前标识的关键帧匹配点相对应的关键帧匹配点所查明的关键帧匹配点的坐标之间的差异度量,

(c)计算剩余的关键帧匹配点对的差异度量的平均并确定所计算的平均是否超出规定最小等级;

(d)只要所计算的平均超出规定最小等级,

按量级排序所述差异度量并从最大开始向下以量级标识规定百分比的差异度量,

将与所标识的差异度量之一相关联的每个关键帧匹配点对的在所记录的描述了在所述最后输入的彩色图像帧中所描述的场景的至少规定部分的关键帧中的先前标识的关键帧匹配点和其在最后输入的彩色图像帧中的对应的关键帧匹配点解除作为关键帧匹配点的指定,

使用剩余的关键帧匹配点重新估计所述第二变换矩阵;

(e)确定是否已经达到规定的变换矩阵重新估计的迭代的最大数目;以及

(f)只要还没有达到规定的变换矩阵重新估计的迭代的最大数目,重复子模块(a)到(f)的执行。

10.如权利要求8所述的系统,其特征在于,所述确定是否要发生关键帧匹配的子模块包括:

确定自从最后一次执行关键帧匹配以来是否已经输入了规定数目的帧,

只要确定自从最后一次执行关键帧匹配以来还没有输入规定数目的帧,指定关键帧匹配发生的时机未到,以及

只要确定自从最后一次执行关键帧匹配以来已经输入规定数目的帧,指定关键帧匹配发生的时机已到。

11.一种用于头部姿态跟踪的系统,包括:

被配置以放置在用户的头部上的一组传感器,该组传感器包括被用于标识在由该组传感器所感测到的场景内的点的三维位置的深度传感器装置,彩色摄像机以及一套测量围绕三个轴的角速度和沿所述三个轴的线性加速的惯性传感器,所述深度传感器装置和所述彩色摄像机被同步以周期性地以深度帧和彩色图像帧形式产生同时期的场景数据,并被校准以便将每个彩色图像帧中的每个像素映射到同时期产生的深度帧中的对应的三维场景位置,并且所述惯性传感器以等于或超过提供深度和彩色图像帧的速率提供了角速度和线性加速数据的帧;

计算设备;以及

包括由所述计算设备执行的程序模块的计算机程序,其中所述计算设备可由所述计算机程序的程序模块来指示以便:

输入所产生的每个惯性传感器帧;

对于第一之后的每个输入惯性传感器帧,

使用最后输入的惯性传感器帧来估计当前基于惯性传感器的变换矩阵,所述变换矩阵定义了点从紧接之前的输入惯性传感器帧到最后输入的惯性传感器帧的平移和旋转,

估计定义点从第一输入惯性传感器帧到最后惯性传感器帧的平移和旋转的最终基于惯性传感器的变换矩阵,所述估计通过将定义了点从第一输入惯性传感器帧到紧接在最后输入的惯性传感器帧之前的惯性传感器帧的平移和旋转的先前计算的变换矩阵与定义

了点从所述紧接之前的输入惯性传感器帧到最后输入的惯性传感器帧的平移和旋转的当前的基于惯性传感器的变换矩阵进行累加来完成，

确定是否已经产生了新的彩色视频和深度帧，

只要还没有产生新的彩色视频和深度帧，将该最终的基于惯性传感器的变换矩阵应用于先前确定的与第一输入惯性传感器帧相关联的头部姿态位置和朝向以标识当前的头部姿态位置和朝向，

只要确定已经产生了新的彩色视频和深度帧，

输入所述新的彩色图像和深度帧，

标识在最后输入的彩色图像帧和紧接在所述最后输入的彩色图像帧之前的彩色图像帧之间的匹配特征，

使用所标识的匹配特征和在所述最后输入的彩色图像帧和紧接在所述最后输入的彩色图像帧之前产生的彩色图像帧这两者中的所述匹配特征的对应的三维位置估计当前的基于图像的变换矩阵，所述基于图像的变换矩阵定义了点从紧接在所述最后输入的彩色图像帧之前产生的彩色图像帧到所述最后输入的彩色图像帧的平移和旋转，

估计定义点从第一输入惯性传感器帧到最后输入彩色图像帧的平移和旋转的最终的基于图像的变换矩阵，所述估计通过将定义了点从第一输入惯性传感器帧到紧接在最后输入的彩色图像帧之前的彩色图像帧的平移和旋转的先前计算的变换矩阵与定义了点从所述紧接之前的彩色图像帧到该最后输入的彩色图像帧的平移和旋转的当前的基于图像的变换矩阵进行累加来完成，

将最终的基于图像的变换矩阵和最终的基于惯性传感器的变换矩阵合并以产生定义了点从第一输入惯性传感器帧到最后输入的惯性传感器帧的平移和旋转的单个经组合的变换矩阵，以及

将经组合的变换矩阵应用于先前确定的与第一输入惯性传感器帧相关联的头部姿态位置和朝向以标识当前的头部姿态位置和朝向。

12. 如权利要求11所述的系统，其特征在于，还包括在执行用于输入所述新彩色图像和深度帧的程序模块之后且在执行用于标识在最后输入的彩色图像帧和紧接在所述最后输入的彩色图像帧之前的彩色图像帧之间的匹配特征的程序模块之前来执行的关键帧程序模块，所述关键帧程序模块包括用于以下操作的子模块：

确定是否要建立新的关键帧，以及

只要确定要建立新的关键帧，记录所述最后输入的彩色图像帧并将所记录的帧指定为关键帧。

13. 如权利要求12所述的系统，其特征在于，所述用于确定是否要建立新的关键帧的子模块包括只要最后输入的彩色图像帧是所述第一输入彩色图像帧，且只要所述最后输入的彩色图像帧描述少于先前建立的关键帧中所描述的场景的规定部分，就确定要建立新的关键帧。

14. 如权利要求12所述的系统，其特征在于，所述关键帧程序模块还包括用于以下操作的子模块：

只要确定不要建立新的关键帧，就确定是否要发生关键帧匹配；

只要确定要发生关键帧匹配或已经建立了新的关键帧，

标识在所述最后输入的彩色图像帧和所记录的描述了最后输入的彩色图像帧中所描述的场景的至少规定部分的关键帧之间的关键帧匹配点，

使用所标识的关键帧匹配点和在所述最后输入的彩色图像帧和所记录的描述了最后输入的彩色图像帧中所描述的场景的至少规定部分的关键帧两者中的关键帧匹配特征的对应的三维位置来估计第二变换矩阵，所述第二变换矩阵定义了点从所记录的描述了最后输入的彩色图像帧中所描述的场景的至少规定部分的关键帧到所述最后输入的彩色图像帧的平移和旋转，

使用所述第二变换矩阵和定义了点从第一输入彩色图像帧到所记录的描述了最后输入的彩色图像帧中所描述的场景的至少规定部分的关键帧的平移和旋转的先前计算的关键帧变换矩阵来估计第三变换矩阵，所述第三变换矩阵定义了点从所述第一输入彩色图像帧到所述最后输入彩色图像帧的平移和旋转，

只要所述最后输入的彩色图像帧是新建立的关键帧，就将所述第三变换矩阵指定为定义了点从所述第一输入彩色图像帧到新建立的关键帧的平移和旋转的关键帧变换矩阵，

估计定义点从第一输入惯性传感器帧到最后输入彩色图像帧的平移和旋转的最终基于图像的变换矩阵，所述估计通过将定义了点从第一输入惯性传感器帧到紧接在最后输入的彩色图像帧之前的输入彩色图像帧的平移和旋转的先前计算的变换矩阵与定义了点从所述紧接之前的彩色图像帧到该最后输入的彩色图像帧的平移和旋转的当前的基于图像的变换矩阵进行累加来完成，

将所述第三和最终的基于图像的变换矩阵合并以产生定义了点从第一输入惯性传感器帧到最后输入的彩色图像帧的平移和旋转的单个经组合的基于图像的变换矩阵，以及

将所述经组合的基于图像的变换矩阵指定为所述最终的基于图像的变换矩阵。

使用深度相机的头部姿态跟踪

[0001] 背景

[0002] 头部姿态跟踪通常涉及周期性确定人的头部在空间中的位置和朝向。人的头部在空间中的位置一般被定义为按照为空间预先建立的世界坐标系来表示的3D位置。该位置可以,例如,与头部质心(即所估计的人的头部的中心)相关联。尽管可另选地规定与人的头部相关联的另一个容易查明的点。头部姿态朝向通常按照绕在头部质心(或其它规定的头部点)处具有它们的公共原点的三个正交轴的旋转(俯仰、翻滚、偏航)来定义。典型地,俯仰是头部的上和下的移动,偏航是头部的左和右的移动,而翻滚是头部从一侧到另一侧的移动。

[0003] 头部姿态被用于各种应用中。例如,在增强现实应用的上下文中跟踪头部姿态。在增强现实应用中,用户佩戴具有至少一个半透明显示器的护目镜或类似的设备,这样,用户可以看到被呈现在显示器上的现实世界和虚拟对象这两者。虚拟对象需要如同它们是现实环境的部分那样出现。一种保证虚拟对象如同它们是现实环境的部分那样出现的技术组件是头部姿态跟踪。当用户移动他或她的头部时,呈现在(随用户头部移动的)显示器上的虚拟对象需要相对于现实环境固定地出现。

[0004] 一种实现精确头部姿态跟踪的尝试涉及例如通过在已知的墙或天花板的已知位置处放置具有已知图案的标记来操纵现实环境。由(被安装在用户头部或空间中的其它地方的)传统的摄像机所捕捉的图像随后通过使用计算机视觉技术基于标记在图像中的位置计算用户的头部姿态来进行处理。另一种尝试涉及使用安装在用户头部(例如在头盔上或一副护目镜中)的惯性传感器(例如陀螺仪、加速计和罗盘)来查明用户的头部姿态。另外,已经提议了在其中将惯性传感器和一个或多个传统的摄像机相组合来获得头部姿态估计而无需操纵环境的头部姿态跟踪方案。

发明内容

[0005] 在此所述的头部姿态跟踪技术实施例通常也涉及周期性地跟踪人的头部在空间中的位置和朝向。在一个通用的示例性实施例中,这通过使用被配置以放置在用户的头部上的一组传感器来完成。该组传感器有利地包括深度传感器装置,它被用于标识通过该组传感器和至少一个其它类型的传感器所感测的场景内的点的三维位置。计算设备被用于运行头部姿态跟踪计算机程序。该计算机程序包括用于周期性地输入由该组传感器中的每个传感器输出的数据的模块。它还包括用于在每次从一个或多个传感器输入数据时使用所输入的数据来计算变换矩阵的模块。当被应用于在输入第一传感器数据时先前所确定的头部姿态的位置和朝向时,该变换矩阵标识了当前的头部姿态的位置和朝向。

[0006] 具体而言,在一个实施例中,前述其它类型的传感器是彩色摄像机。深度传感器装置和彩色摄像机被同步,以便周期性地分别产生深度帧和彩色图像帧形式的同时期场景数据。另外,对深度传感器装置和彩色摄像机校准以便将每个彩色图像帧中的每个像素映射到同时期产生的深度帧中相应的三维场景位置(如果可能的话)。前述计算机程序随后涉及第一输入每个同时期产生的深度帧和彩色图像帧。对于这些帧对输入(在所述第一之后)的每个帧对输入,标识在最后输入的彩色图像帧和紧接在该最后输入的彩色图像帧之前所产

生的彩色图像帧(即紧接之前的彩色图像帧)之间的匹配特征。随后使用所标识的匹配特征和在最后输入的彩色图像帧和紧接在该最后输入的彩色图像帧之前所产生的彩色图像帧之间的匹配特征的对应的三维位置来估计第一变换矩阵。该变换矩阵定义了点从一个帧到另一个帧,且具体地从紧接在该最后输入的彩色图像帧之前所产生的彩色图像帧到该最后输入的彩色图像帧,的平移和旋转。接着,估计定义了这些点从第一彩色图像帧输入到最后输入的彩色图像帧的平移和旋转的最终变换矩阵。这通过将定义了从第一彩色图像帧到紧接在该最后输入的彩色图像帧之前的彩色图像帧的点的平移和旋转的先前计算的变换矩阵与定义了从所述紧接之前的彩色图像帧到该最后输入的彩色图像帧的点的平移和旋转的第一变换矩阵进行累加来完成。随后,将该最终的变换矩阵应用于在第一彩色图像帧输入中描述的场景内的先前所确定的姿态位置和朝向以标识当前的头部姿态位置和朝向。

[0007] 在又一个实施例中,前述的其它类型的传感器不仅包括如前所配置的彩色摄像机,而且还包括一套测量围绕三个轴的角速度和沿三个轴的线性加速的惯性传感器。该惯性传感器以等于或超过提供深度和彩色图像帧的速率的速率提供了角速度和线性加速数据的帧。在这种情况下,前述计算机程序涉及第一输入所产生的每个惯性传感器帧。随后,对于每个惯性传感器帧输入(在第一之后的),使用最后输入的惯性传感器帧来估计当前基于惯性传感器的变换矩阵。该基于惯性传感器的变换矩阵定义了点从紧接之前的惯性传感器帧输入到最后输入的惯性传感器帧的平移和旋转。随后估计定义点从第一惯性传感器帧输入到最后惯性传感器帧的平移和旋转的最终的基于惯性传感器的变换矩阵,所述估计通过将定义了点从第一惯性传感器帧到紧接在该最后输入的惯性传感器帧之前的惯性传感器帧(即紧接之前的惯性传感器帧)的平移和旋转的先前计算的变换矩阵与定义了点从所述紧接之前的惯性传感器帧到该最后输入的惯性传感器帧的平移和旋转的当前基于惯性传感器的变换矩阵进行累加来完成。接着,确定是否已经产生了新的彩色视频和深度帧。如果没有,将该最终的基于惯性传感器的变换矩阵应用于先前确定的与第一惯性传感器帧输入相关联的头部姿态位置和朝向以标识当前的头部姿态位置和朝向。然而,如果已经产生了新的彩色视频和深度帧,将它们输入并且标识在最后输入的彩色图像帧和紧接最后输入的彩色图像帧之前所产生的彩色图像帧之间的匹配特征。随后使用所标识的匹配特征和在最后输入的彩色图像帧和紧接在该最后输入的彩色图像帧之前所产生的彩色图像帧之间的匹配特征的对应的三维位置来估计基于图像的变换矩阵。该基于图像的变换矩阵定义了点从紧接在该最后输入的彩色图像帧之前所产生的彩色图像帧到该最后输入的彩色图像帧的平移和旋转。接着估计定义点从第一惯性传感器帧输入到最后输入彩色图像帧的平移和旋转的最终的基于图像的变换矩阵,所述估计通过将定义了点从第一惯性传感器帧到紧接在该最后输入的彩色图像帧之前的彩色图像帧的平移和旋转的先前计算的变换矩阵与定义了点从所述紧接之前的彩色图像帧到该最后输入的彩色图像帧的平移和旋转的当前基于图像的变换矩阵进行累加来完成。该最终的基于图像的变换矩阵和最终的基于惯性传感器的变换矩阵合并以产生定义了点从第一惯性传感器帧输入到最后输入的惯性传感器帧的平移和旋转的单个经组合的变换矩阵。随后将该经组合的变换矩阵应用于先前确定的与第一惯性传感器帧输入相关联的头部姿态位置和朝向以标识当前的头部姿态位置和朝向。

[0008] 应该注意,提供前述发明内容是为了以简化的形式介绍将在以下详细描述中进一

步描述的概念选择。本发明内容不旨在标识出所要求保护的主题的关键特征或必要特征，也不旨在用于帮助确定所要求保护的主题的范围。

[0009] 附图简述

[0010] 参考以下描述、所附权利要求书以及附图，将更好地理解本发明的具体特征、方面和优点，附图中：

[0011] 图1是描述用于实现在此所述的头部姿态跟踪技术实施例的头部姿态跟踪系统的简化框图。

[0012] 图2是在实现在此所述的头部姿态跟踪技术实施例中使用的各种帧和坐标方案的简化附图。

[0013] 图3是描述将深度传感器装置和传统的彩色摄像机一起使用的图1的示例性头部姿态跟踪系统的简化图表。

[0014] 图4A-D是大体概述使用由图3描述的深度传感器装置和彩色摄像机输出的数据的头部跟踪过程的实施例的流程图。

[0015] 图5是大体概述用于确定关键帧匹配是否将发生的过程的一个实现的流程图。

[0016] 图6是描述将深度传感器装置和传统的彩色摄像机以及一套惯性传感器一起使用的图1的示例性头部姿态跟踪系统的简化图表。

[0017] 图7A-E是大体概述使用由图6描述的深度传感器装置、彩色摄像机和该套惯性传感器输出的数据的头部跟踪过程的实施例的流程图。

[0018] 图8是描绘了构成用于实现在此描述的头部姿态跟踪技术的各实施例的示例性系统的通用计算设备的图表。

具体实施方式

[0019] 在以下对头部姿态跟踪技术的各实施例的描述中，对附图进行了参考，附图构成了各实施例的一部分且在其中作为说明示出了可在其中实践该技术的具体实施例。可以理解，可以使用其他实施例并且可以做出结构改变而不背离本发明技术的范围。

[0020] 1.0 头部姿态跟踪

[0021] 通常，在此所述的头部姿态跟踪技术实施例周期性地跟踪人的头部在空间中的位置和朝向。在图1所示的一个通用实现中，该实现涉及一种使用被配置以放置在用户的头部上的一组传感器100的系统。该组头部承载的传感器100包括深度传感器装置102，它被用于标识通过该组传感器和至少一个其它类型的传感器104所感测的场景内的特征的三维位置。另外，该示例性头部姿态跟踪系统具有运行计算机程序108的计算设备106（例如在本说明书后面可找到的计算环境章节中所描述的计算设备），所述计算机程序具有包括传感器模块110和变换矩阵模块112的程序模块。通常，传感器模块周期性地输入由该组传感器中的每个传感器输出的数据，并且变换矩阵模块在每次从一个或多个传感器输入数据时使用所输入的数据来计算变换矩阵。当被应用于在输入第一传感器数据时先前所确定的头部姿态的位置和朝向时，该变换矩阵可被用于标识当前的头部姿态的位置和朝向。

[0022] 在此所述的头部姿态跟踪技术实施例使用了不同的传感器方案。如上所示，这些方案的每个方案包括被用于标识在深度传感器装置和空间内的位置之间的距离的深度传感器装置（它典型地包括某种类型的投影仪和接收器）。在一个示例性实现中，深度传感器

装置与传统的彩色摄像机一起被使用。在另一示例性实现中,深度传感器装置与传统的彩色摄像机和一套惯性传感器一起被使用。这些方案的每个方案还包括先前所示的用于处理由各种传感器输出的传感器信号的计算设备。在下述的章节中,前述的示例性实现的每个实现都将被更加详细地描述,并且每个实现的实现过程也将被同样描述。

[0023] 然而,首先注意,在头部姿态跟踪中,估计头部坐标帧(帧“H”)或传感器帧(帧“S”)在世界帧(帧“W”)中的位置(T_{WH} 或 T_{WS})和朝向(R_{WH} 或 R_{WS})。传感器帧是指彩色相机帧(帧“C”)202和深度传感器装置的接收机帧204(帧“DC”)以及当被包括时的惯性传感器帧(帧“I”)206。这些传感器帧被变换成与头部帧相同,并且假设在跟踪开始处世界帧200覆盖了经变换的传感器帧。具体而言,这三个传感器帧的每个具有三个轴。在这三个传感器帧之间的相对变换不随时间改变。这样,所述相对变换是离线预校准的。实际上,在开始处的彩色相机传感器帧被用作世界帧,并且深度传感器接收机和惯性传感器帧被变换为匹配。前述帧被如图2所示定义,其中世界帧200的坐标被标记为 X_w, Y_w, Z_w ,彩色相机帧202的坐标被标记为 X_c, Y_c, Z_c ,深度传感器接收机帧204的坐标被标记为 X_{DC}, Y_{DC}, Z_{DC} ,而惯性传感器帧206的坐标被标记为 X_I, Y_I, Z_I 。注意,箭头 T_I 和 T_{DC} 表示与深度传感器接收机和惯性传感器帧相关联的前述的相对变换,而虚线0指示在彩色相机帧和世界帧之间的对应性。

[0024] 除了前述帧和坐标之外,在以下说明书中将使用下述符号和定义:

[0025] p^1 : “点p在帧1中的坐标”;

[0026] p^2 : “点p在帧2中的坐标”;

[0027] R_{12} : “从帧1到帧2的旋转”或“在帧1中描述的帧2”;

[0028] T_{12} : “从帧1的原点到在帧1中描述的帧2的原点的平移”;以及

[0029] $p^1 = R_{12}p^2 + T_{12}$: “将点p的坐标从帧2变换到帧1”。

[0030] 通常,所使用的约定是上标指示点在哪个帧中,而下标指示其它信息,例如变换的次序。

[0031] 1.1 使用彩色视频的示例性系统实现和过程

[0032] 相机和深度传感器装置

[0033] 如前所述,一个示例性系统实现通常使用包括深度传感器装置302和传统的彩色摄像机304的头部承载的传感器300以及计算机306,如图3所示。可以理解,可使用各种类型的深度传感器装置。作为示例而非限制,深度传感器装置可以是红外投影仪和红外摄像机形式的红外接收器。红外投影仪和摄像机可以以下述传统方式工作:红外投影仪将红外光发射到环境中,而红外摄像机捕捉从环境中的对象和表面射出的红外光反射。随后,将所反射的红外光用于以传统方式计算离开接收机和环境中的一个点或特征的距离。对于彩色摄像机,这可以是任意合适的可视光摄像机(也称为RGB(红/绿/蓝)摄像机)。

[0034] 注意,深度传感器装置和彩色摄像机被配置以便由用户佩戴在他或她的头部上(安装在头盔上或一副护目镜或眼镜中,或在这两者之上)。可以理解,在深度传感器装置和彩色摄像机之间维护了固定的空间关系。而且,假设深度传感器装置和彩色摄像机的朝向和位置模仿用户头部的朝向和位置。还有,深度传感器装置和彩色摄像机被配置以便提供同步的彩色和深度图像序列。

[0035] 在跟踪之前,彩色摄像机和深度相机相对于彼此被校准以获得在它们之间的变换。这使用传统的方法来完成。使用校准的结果,彩色图像中的像素可以被映射到对应的深

度图像中的像素,如果可能的话。在一个实现中,通过将光流跟踪器应用于彩色图像序列来执行姿态跟踪。深度信息可以用在变换计算中。然而,单纯基于光流跟踪的头部姿态估计可能最终导致漂移。因此,在一个实现中,记录一个或多个关键帧,并且通过比较跟踪结果和这些关键帧来校正所述结果。

[0036] 具体而言,参照图4A-D,在在此所述的头部姿态跟踪技术实施例的一个实现中,由深度传感器装置和彩色摄像机所捕捉的图像数据是以逐帧基础被捕捉的。每个被捕捉的帧依次被输入到前述的计算机中。这样,过程通过输入最后捕捉的彩色视频和深度帧(过程动作400)开始。随后确定是否要建立新的关键帧(过程动作402)。在一个版本中,第一帧输入被建立为关键帧。该第一关键帧可以是为该过程期间建立的唯一关键帧。或者,可以周期性地建立新的关键帧。例如,只要最后输入的帧描述少于在先前建立的关键帧中所描述的场景区的规定部分(例如40%)。这样,可以记录关键帧的集合,以便随时构建周围场景的可视特征全景。稍后,将更加详细地描述建立一个或多个关键帧的目的。

[0037] 只要确定没有新的关键帧要被建立,随后确定关键帧匹配是否发生(过程动作404)。光流跟踪在连续帧中跟踪特征点。在连续帧之间计算的变换最终可能产生漂移效应。在一个实现中,该漂移效应通过使用绝对基准被校正。一种创建这种绝对基准的方式是当处理连续帧时将一个或多个关键帧与它们的变换矩阵存储回第一彩色图像帧输入。只要当前帧描述了也在关键帧中被捕捉的场景的一部分(例如50%),就可能对漂移进行校正,如稍后所述的。注意,以这种方式使用关键帧匹配还具有允许从暂时的跟踪失败中恢复的优点。然而,出于稍后将描述的原因,可以并不针对每个新的彩色图像帧输入都执行关键帧匹配。只要没有对最后输入的彩色图像帧执行关键帧匹配,对于在第一之后的每个新的帧输入,使用前述光流技术来标识在最后输入的彩色图像帧和紧接最后输入的彩色图像帧之前的彩色图像帧之间的匹配特征(过程动作406)。任意合适的传统的光流方法都可被用于此目的,尽管在一个实现中允许实时执行是计算上高效的。假设相机帧速率为大约30HZ,对于正常速度的运动,光流跟踪器可以高效地找出在前帧和当前帧之间的点匹配。所找到的匹配的点的数目依赖于环境和跟踪器中的参数设置。注意,在光流过程中有时是图像特征而非点被匹配。虽然,在此提供的本说明书是指匹配的点,但可以理解其它匹配的特征可以被用于匹配的点的替换或者补充。

[0038] 接着,在过程动作420中,使用光流信息和先前输入深度帧来估计当前的变换矩阵。该变换矩阵定义了点从紧接在最后输入的彩色图像帧之前所产生的帧到该最后输入的彩色图像帧的平移和旋转。注意,光流技术标识在两个帧中的匹配点的2D图像坐标。然而,存在缩放模糊性且精度难以保证。这是对应于这两个彩色图像帧的深度图像发挥作用的地方。深度图像提供了场景的3D信息,并且因此,可以为每个彩色相机帧找到匹配的点的3D位置。给定两个或更多组匹配的点的3D坐标,可以使用标准方法来估计前述变换矩阵。

[0039] 然而,由于匹配的点经常是拐角点的事实,从对应的深度图像中查明在彩色图像中的给定点的深度值可能是困难的。这样,存在一种匹配的点是3D拐角的可能性。在这种情况下,该点在深度映射上的投影可能落在边缘上或空白区域中。这引起了模糊性,因为对于两个匹配的点,一个可以在较近的表面上找到,而另一个可以在更远的表面上找到,或者它们两者之一可以具有无效的深度值。这将减少有用的匹配的数目。为了解决这个问题,在一个(在图4C以虚线框示出的指示动作的可选性质的)实现中,在计算变换矩阵之前,使用深

度校正模块。具体而言,该模块涉及在直接被映射的匹配点周围定义小的搜索窗口(例如3x3像素)(过程动作408)。随后标识在每个窗口内与相机最近的点,并将其指定为经细化的匹配点(过程动作410)。这些经细化的匹配点随后可以被用于在过程动作420中计算变换矩阵。通过选择较近的表面,可以避免前述模糊性。

[0040] 注意,在光流跟踪期间,将很可能引入匹配点异常值。可以通过删除典型地由所使用的匹配算法所标识的具有较低匹配质量的匹配来减少异常值的数目。具体而言,在一个(在图4C以虚线框示出的指示动作的可选性质的)实现中,在计算变换矩阵之前,选择先前未选择的一对匹配点(过程动作412)。随后,在过程动作414确定由所使用来标识所选择的匹配点的匹配算法所标识的匹配质量等级是否落在规定最小等级(例如1毫米)之下。如果是,所选择的匹配点被去除作为匹配点的指定(过程动作416)。一旦所选择的匹配点已经被去除指定,或如果所选择的点的匹配质量等级不落在规定最小等级之下,确定是否存在还未被选择的任意剩余的匹配点(过程动作418)。如果存在,则按需重复过程动作412到418直到已经考虑了所有的匹配点对。

[0041] 虽然前述过程将减少异常匹配点的数目,但由于剩余的异常值还是经常无法保证高性能。因此,在删除了具有(如由它们的匹配质量落在规定最小等级之下所指示的)坏匹配质量的匹配之后,可以使用迭代过程来进一步移除至少最明显的剩余异常值。具体而言,在一个(在图4D以虚线框示出的指示动作的可选性质的)实现中,最后估计的变换矩阵(R_{AR} , T_{AR})(其中A表示在先前帧中的一组点,该组点分别匹配当前帧的一组点B中的一个点)被用于变换组B中的点(过程动作422)。随后将每个经变换的点所得到的点坐标与在组A中其对应的点相比较以建立差异度量(过程动作424)。接着,计算匹配点对的差异度量的平均,并且在过程动作426中确定该平均是否超出规定的最小等级(例如1毫米)。如果该平均并未超过规定的最小等级,那么,按照量级对差异度量排序并从最大开始向下以量级标识规定百分比(例如10%)的差异度量(过程动作428)。随后,将与所标识的差异度量之一相关联的每个匹配对的组B中的匹配点和其在组A的对应匹配点去除作为匹配点对的指定(过程动作430)。接着,使用剩余的匹配点对重新估计变换矩阵(过程动作432)。随后,确定是否已经达到变换矩阵重新估计迭代的规定的最大数目(例如3)(过程动作434)。如果否,则适当重复过程动作422到434。这会继续直到要么先前计算的平均不超过规定的最小等级要么确定已经达到变换矩阵重新估计迭代的规定的最大数目。

[0042] 接着估计定义点从第一彩色图像帧输入到最后帧输入的平移和旋转的最终变换矩阵,所述估计通过将定义了点从第一输入彩色图像帧到紧接在该最后输入的帧之前的彩色图像帧的平移和旋转的先前计算的变换矩阵与定义了点从所述紧接最后输入的彩色图像帧之前的帧到该最后输入的彩色图像帧的平移和旋转的当前变换矩阵进行累加来完成(过程动作438)。还应注意,累加前述变换矩阵以使用传统的方法形成最终的变换矩阵。

[0043] 然而,如果在过程动作402中,确定要建立新的关键帧,那么,记录最后输入的彩色相机帧并将其指定为当前关键帧(过程动作436)。另外,相对于关键帧匹配是否会发生的决策而言,由于即使当前帧和关键帧两者都描述相同的场景部分,它们之间还是可能存在明显的运动,因此,在匹配中典型地需要(例如使用传统的加速稳健特征(Speeded Up Robust Features(SURF))匹配过程可以找到的)强的特征以实现精确结果。不幸地是,这种强力特征检测器和描述器通常在计算上是昂贵的。因此,在一种版本中,仅针对每少量(例如15)帧

来进行关键帧匹配。这是合理的,因为漂移仅在变换错误已经针对少量帧累积之后才会明显起来。鉴于前述,图5概述了一种确定关键帧匹配是否要发生的方式。首先,确定自从最后一次执行关键帧匹配以来是否已经输入了规定数目的帧(过程动作500)。如果规定数目的帧还未被输入,那么指定未到用于新的关键帧匹配的时间(过程动作502)。然而,如果已经输入了规定数目的帧,被认为到了执行关键帧匹配过程的时间(过程动作504)。

[0044] 鉴于前述并再次参见图4A,如果在过程动作404中确定关键帧匹配要发生,或如果最后输入的彩色图像帧已经被指定为当前关键帧,那么标识在最后输入的彩色图像帧和描述最后输入的帧中所描述的场景的至少规定部分的关键帧(此后将被称为经标识的关键帧)之间的匹配点(过程动作440)。如前所述,在一个实现中,出于该目的使用传统的SURF匹配过程。注意,在过程动作408到410以及过程动作412到418中先前描述的可选匹配点细化和异常值消除动作也可以在此实现以细化所标识的关键帧和最后输入帧之间的匹配点位置。

[0045] 一旦在最后输入的彩色图像帧和所标识的关键帧之间的匹配点已经被标识,它们就被用于与来自对应的深度帧的3D位置数据相结合以在过程动作442中估计当前的变换矩阵(即, $(R_{t,k}, T_{t,k})$)。该变换矩阵定义了点从所标识的关键帧(即k)到最后输入帧(即t)的平移和旋转。注意,该变换矩阵估计是以与在过程动作420中估计在连续帧之间的变换矩阵相同的方式来完成的。另外,在过程动作422到434中先前所述的可选的迭代异常值消除动作也可以在此实现以潜在地增加所估计的变换的精确性。随后,使用在所标识的关键帧和最后输入帧之间获得的变换矩阵以及定义了点从第一彩色图像帧输入(即帧1)到所标识的关键帧的平移和旋转的先前计算的关键帧变换矩阵(即 $(R_{1,t}, T_{1,t})$)来估计定义了点从第一彩色图像帧输入到最后输入帧的平移和旋转的关键帧匹配的变换矩阵(过程动作444)。在一个实现中,这是通过将所标识的关键帧和最后输入帧之间估计的变换矩阵与所述第一帧和所标识的关键帧之间的先前计算的变换矩阵相乘来完成的。在最后输入的彩色图像帧已经被指定为当前关键帧的情况下,与该帧相关联的关键帧匹配的变换矩阵随后被指定为用于该帧的关键帧变换矩阵(过程动作445)。这建立了定义点从第一彩色图像帧输入到当前关键帧的平移和旋转的预计算的关键帧变换矩阵,用于在创建下一关键帧时计算关键帧匹配的变换矩阵。接着估计定义点从第一彩色图像帧输入到最后输入帧的平移和旋转的变换矩阵,所述估计通过将定义了点从第一输入彩色图像帧到紧接在该最后输入的帧之前的彩色图像帧输入的平移和旋转的最终变换矩阵与定义了点从所述紧接最后输入的彩色图像帧之前的帧到该最后输入的彩色图像帧的平移和旋转的当前变换矩阵进行累加来完成。如前所述,由于对于光流跟踪来说,以第一开始从每个帧到下一帧所估计的变换(即, $(R_{12}, T_{12}), (R_{23}, T_{23}), \dots, (R_{k-1,k}, T_{k-1,k})$)可以被累加以使用传统的方法产生从第一帧关键帧(即帧1)到最当前帧(即帧k)的变换矩阵,所以,这是可能的。

[0046] 在过程中的该点处,定义点从第一彩色图像帧输入到最后输入的帧的平移和旋转的单独的变换矩阵基于光流跟踪和关键帧匹配这两者存在。这些变换矩阵被合并以产生单个经组合的变换矩阵(过程动作448)。

[0047] 在一个实现中,执行在这两个变换矩阵估计之间的加权内插以取得潜在更好的变换。由于通过深度图像知晓了匹配点的3D位置,就可能使用该3D信息来定义变换估计差错度量。具体而言,从坐标帧1(F1)中考虑图像1(I1)。具体而言,从坐标帧2(F2)中考虑图像2

(I2)。先前估计的变换矩阵(R_{12}, T_{12})描述了从F1到F2的旋转和平移。在来自F2的I2中的图像点通过应用(R_{12}, T_{12})被变换成F1。随后,将F1的原点(在那时的相机的光学中心)与每个经变换的点相连。如果在3D空间中的相连线上存在I1中的点,计算这两个点之间的距离。随后,上述差错度量是对涉及该计算的所有点(即所述点是在所述线上的I1点)上的所计算的距离的平均。为了减少计算量,在一个实现中,将图像下采样到与上述变换过程之前的相同程度(例如50%)。另外,在变换后在I1所描述的场景之外将存在一些没有被考虑的点。

[0048] 一旦计算了变换估计差错度量(此后被称为差错),在为当前帧所计算的两个变换度量之间执行内插。在一个实现中,这通过使用四元组空间中的线性内插来完成。具体而言,上述权重可以使用下述公式计算:

$$[0049] \quad \text{权值} = \frac{1}{1 + (\alpha * \text{差错})^2}$$

[0050] 其中 α 是常数(例如10000)。随后,在使用所得到的权值来在为当前帧所计算的两个变换度量之间进行内插之前对它们进行归一化。在一个实现中,该内插如下实现。使得p1和q1标记第一变换矩阵的位置和朝向(以四元组形式)。使得p2和q2标记第二变换矩阵的位置和朝向(以四元组形式)。使得w1标记对应于第一变换矩阵的差错,而w2标记对应于第二变换矩阵的差错。使得 $a1 = w1 / (w1 + w2)$,而 $a2 = w2 / (w1 + w2)$ 。随后,内插位置和四元组为 $p = a1 * p1 + a2 * p2$, $q = a1 * q1 + a2 * q2$ 。注意,在前述在四元组中描述的朝向的线性内插在当这两个四元组之间的角度小时很有效。对于更大的角度,可以使用球形线性内插。

[0051] 再次参考图4B,一旦生成了单个经组合的变换矩阵,它被指定为最终的变换矩阵(过程动作450)。最终的变换矩阵(在过程动作438或450中生成)随后被应用于先前确定的在第一彩色图像帧输入中所描述的场景内的头部姿态位置和朝向以标识在最后输入的彩色图像帧中所描述的场景中的当前头部姿态位置和朝向(过程动作452)。随后,可将当前头部姿态和变换矩阵用于各种其它任务,包括计算在增强现实应用中的当前帧的虚拟对象的位置。

[0052] 接着,确定是否已经输入新的彩色视频和深度帧(过程动作454)。如果是,则按需重复过程动作400到454。只要有新的帧被捕捉和输入到该计算机,就继续这样。

[0053] 1.2使用彩色摄像机、深度传感器装置和惯性传感器的示例性系统实现和过程

[0054] 如前所述,另一示例性系统实现通常使用包括深度传感器装置602、传统的彩色摄像机604和一套惯性传感器606的头部承载的传感器600以及计算机608,如图6所示。可以理解,在此同样,可以使用各种类型的深度传感器装置,例如如前所述的红外投影仪和红外相机装置。相对于惯性传感器,这通常包括一组使用加速度计和陀螺仪来测量速度、朝向和重力的传感器。在一个实现中,使用了传统的惯性测量单元,它包括陀螺仪和加速度计(例如允许传感器被集成到一个小芯片中的3轴MEM陀螺仪和3轴MEM加速度计)。陀螺仪测量围绕三个轴的角速度,而加速度计测量沿这三个轴的线性加速度。

[0055] 注意,深度传感器装置和彩色摄像机以惯性传感器被配置以便由用户佩戴在他或她的头部上(例如,安装在头盔上或一副护目镜或眼镜中,或在这两者之上)。可以理解,在这些传感器之间维持着固定空间关系,并且它们共享公共的坐标系原点。而且,假设这些传感器的朝向和位置模仿了用户的头部的朝向和位置。在此同样,深度传感器装置和彩色摄像机被配置以便提供同步的彩色和深度图像序列。

[0056] 惯性传感器的加入具有许多优点,特别是当环境不具有视觉丰富的纹理时。具体而言,相机和惯性传感器具有互补特性。例如,相机具有较低的帧速率且在处理中要求更多的计算能力,但在视觉特征丰富的环境中可以提供比惯性传感器明显更加精确的测量。反之,惯性传感器可以达到非常高的帧速率且处理有效得多,并且可以帮助示例性系统克服视觉特征薄弱的时期。而且,基于陀螺仪和加速度计的跟踪将导致几秒钟的显著漂移,但相机可被用于记录陆标并校正该漂移。

[0057] 在一个实现中,使用来自惯性传感器的测量并将这与来自彩色和深度图像的(当可用时的)视觉测量相组合以执行姿态跟踪。此处同样,相对于视觉测量,将光流跟踪器应用于彩色图像序列。深度信息可以用在变换计算中。然而,单纯基于光流跟踪的头部姿态估计可能最终导致漂移。因此,在一个实现中,记录一个或多个关键帧,并且通过比较跟踪结果和这些关键帧来校正所述结果。

[0058] 具体而言,参照图7A-E,在在此所述的头部姿态跟踪技术实施例的一个实现中,由深度传感器装置和彩色摄像机所捕捉的图像数据是以逐帧基础被捕捉的。另外,由惯性传感器所捕捉的惯性数据是以至少等于且典型地快于前述图像数据的速率被周期性地捕捉。为了方便,每组同时期捕捉的惯性传感器读数将被称为惯性传感器的一个帧。每个被捕捉的图像和惯性传感器帧依次被输入到前述的计算机中。过程通过监视传入帧(过程动作700)开始。周期性地确定是否已经输入了新的(在第一之后)惯性传感器帧(过程动作702)。如果没有,监视继续。如果已经输入了新的惯性传感器帧,估计当前的基于惯性传感器的变换矩阵(过程动作704)。该基于惯性传感器的变换矩阵定义了从紧接之前的惯性传感器帧输入到最后输入的惯性传感器帧的平移和旋转,并使用传统的方法来估计。随后,将基于惯性传感器的变换矩阵用于估计最终的基于惯性传感器的变换矩阵,该最终的基于惯性传感器的变换矩阵定义了从第一惯性传感器帧输入到最后输入的惯性传感器帧的平移和旋转(过程动作706)。这通过使用传统的方法将定义了从第一惯性传感器帧到紧接最后输入的惯性传感器帧之前的惯性传感器帧输入的平移和旋转的单个经组合的变换矩阵与定义了从紧接之前的惯性传感器帧输入到最后输入的惯性传感器帧的平移和旋转的当前基于惯性传感器的变换矩阵累加来完成。接着,确定是否已经产生新的彩色视频和深度帧(过程动作708)。如果没有,将该最终的基于惯性传感器的变换矩阵应用于先前确定的与第一惯性传感器帧输入相关联的头部姿态位置和朝向以标识当前的头部姿态位置和朝向(过程动作710)。随后,可将当前头部姿态和最终的基于惯性传感器的变换矩阵用于各种其它任务,包括计算在增强现实应用中的当前帧中的虚拟对象的位置。

[0059] 然而,如果在过程动作708中确定已经产生了新的彩色视频和深度帧,那么以类似于结合图4A-D的方式来处理它们。具体而言,输入新的彩色视频和深度帧(过程动作712),并确定是否要建立新的关键帧(过程动作714)。在一个版本中,第一帧输入被建立为关键帧。该第一关键帧可以是为该过程期间建立的唯一关键帧。或者,可以周期性地建立新的关键帧。例如,只要最后输入的帧描述少于在先前建立的关键帧中所描述的场景的规定部分(例如40%)。这样,可以记录关键帧的集合,以便随时构建周围场景的可视特征全景。稍后,将更加详细地描述建立一个或多个关键帧的目的。

[0060] 只要确定没有新的关键帧要被建立,随后确定关键帧匹配是否发生(过程动作716)。光流跟踪在连续帧中跟踪特征点。在连续帧之间计算的变换最终可能产生漂移效应。

在一个实现中,该漂移效应通过使用绝对基准被校正。一种创建这种绝对基准的方式是当处理连续帧时将一个或多个关键帧与它们的变换矩阵存储回第一彩色图像帧输入。只要当前帧描述了也在关键帧中被捕捉的场景的一部分(例如50%),就可能对漂移进行校正,如稍后所述的。注意,以这种方式使用关键帧匹配还具有允许从暂时的跟踪失败中恢复的优点。然而,出于稍后将描述的原因,可以并不针对每个新的彩色图像帧输入都执行关键帧匹配。只要没有对最后输入的彩色图像帧执行关键帧匹配,对于在第一之后的每个新的帧输入,使用前述光流技术来标识在最后输入的彩色图像帧和紧接最后输入的彩色图像帧之前的彩色图像帧之间的匹配特征(过程动作718)。任意合适的传统的光流方法都可被用于此目的,尽管在一个实现中允许实时执行是计算上高效的。假设相机帧速率为大约30HZ,对于正常速度的运动,光流跟踪器可以高效地找出在前帧和当前帧之间的点匹配。所找到的匹配的点的数目依赖于环境和跟踪器中的参数设置。注意,在光流过程中有时是图像特征而非点被匹配。虽然,在此提供的本说明书是指匹配的点,但可以理解其它匹配的特征可以被用于匹配的点的替换或者补充。

[0061] 接着,在过程动作732中,使用光流信息和先前输入深度帧来估计当前的变换矩阵。该变换矩阵定义了点从紧接在最后输入的彩色图像帧之前所产生的帧到该最后输入的彩色图像帧的平移和旋转。如前所述,光流技术标识在两个帧中的匹配点的2D图像坐标。然而,存在缩放模糊性且精度难以保证。这是对应于这两个彩色图像帧的深度图像发挥作用的地方。深度图像提供了场景的3D信息,并且因此,可以为每个彩色相机帧找到匹配的点的3D位置。给定两个或更多组匹配的点的3D坐标,可以使用标准方法来估计前述变换矩阵。

[0062] 还是如前所述,由于匹配的点经常是拐角点的事实,从对应的深度图像中查明在彩色图像中的给定点的深度值可能是困难的。这样,存在一种匹配的点是3D拐角的可能性。在这种情况下,该点在深度映射上的投影可能落在边缘上或空白区域中。这引起了模糊性,因为对于两个匹配的点,一个可以在较近的表面上找到,而另一个可以在更远的表面上找到,或者它们两者之一可以具有无效的深度值。这将减少有用的匹配的数目。为了解决这个问题,在一个(在图7D以虚线框示出的指示动作的可选性质的)实现中,在计算变换矩阵之前,使用深度校正模块。具体而言,该模块涉及在直接被映射的匹配点周围定义小的搜索窗口(例如3x3像素)(过程动作720)。随后标识在每个窗口内与相机最接近的点,并将其指定为经细化的匹配点(过程动作722)。这些经细化的匹配点随后可以被用于在过程动作732中计算变换矩阵。通过选择较近的表面,可以避免前述模糊性。

[0063] 另外,在光流跟踪期间,将很可能引入匹配点异常值。可以通过删除典型地由所使用的匹配算法所标识的具有较低匹配质量的匹配来减少异常值的数目。具体而言,在一个(在图7D以虚线框示出的指示动作的可选性质的)实现中,在计算变换矩阵之前,选择先前未选择的一对匹配点(过程动作724)。随后,在过程动作726确定由所使用来标识所选择的匹配点的匹配算法所标识的匹配质量等级是否落在规定最小等级(例如1毫米)之下。如果是,所选择的匹配点被去除作为匹配点的指定(过程动作728)。一旦所选择的匹配点已经被去除指定,或如果所选择的点的匹配质量等级不落在规定最小等级之下,确定是否存在还未被选择的任意剩余的匹配点(过程动作730)。如果存在,则按需重复过程动作724到730直到已经考虑了所有的匹配点对。

[0064] 如前所述,虽然前述过程将减少异常匹配点的数目,但由于剩余的异常值还是经

常无法保证高性能。因此,在删除了具有(如由它们的匹配质量落在规定最小等级之下所指示的)坏匹配质量的匹配之后,可以使用迭代过程来进一步移除至少最明显的剩余异常值。具体而言,在一个(在图7E以虚线框示出的指示动作的可选性质的)实现中,最后估计的变换矩阵(R_{AR}, T_{AR})(其中A表示在先前帧中的一组点,该组点分别匹配当前帧的一组点B中的一个点)被用于变换组B中的点(过程动作734)。随后将每个经变换的点所得到的点坐标与在组A中其对应的点相比较以建立差异度量(过程动作736)。接着,计算匹配点对的差异度量的平均,并且在过程动作738中确定该平均是否超出规定的最小等级(例如1毫米)。如果该平均并未超过规定的最小等级,那么,按照量级对差异度量排序并从最大开始向下以量级标识规定百分比(例如10%)的差异度量(过程动作740)。随后,将与所标识的差异度量之一相关联的每个匹配对的组B中的匹配点和其在组A的对应匹配点去除作为匹配点对的指定(过程动作742)。接着,使用剩余的匹配点对重新估计变换矩阵(过程动作744)。随后,确定是否已经达到变换矩阵重新估计迭代的规定的最大数目(例如3)(过程动作746)。如果否,则适当重复过程动作734到746。这会继续直到要么先前计算的平均不超过规定的最小等级要么确定已经达到变换矩阵重新估计迭代的规定的最大数目。

[0065] 接着估计定义点从第一惯性传感器帧输入到最后输入的彩色图像帧的平移和旋转的变换矩阵,所述估计通过将定义了从第一惯性传感器帧到紧接在该最后输入的彩色图像帧之前的彩色图像帧的平移和旋转的单个经组合的变换矩阵与定义了从所述紧接之前的彩色图像帧到该最后输入的彩色图像帧的平移和旋转的当前变换矩阵进行累加来完成(过程动作750)。还应注意,累加前述变换矩阵以使用传统的方法形成该变换矩阵。

[0066] 然而,如果在过程动作714中,确定要建立新的关键帧,那么,记录最后输入的彩色相机帧并将其指定为当前关键帧(过程动作748)。另外,相对于关键帧匹配是否会发生的决策而言,由于即使当前帧和关键帧两者都描述相同的场景部分,它们之间还是可能存在明显的运动,因此,在匹配中通常需要(例如使用传统的加速稳健特征(Speeded Up Robust Features(SURF)))匹配过程可以找到的)强的特征以实现精确结果。不幸地是,这种强力特征检测器和描述器通常在计算上是昂贵的。因此,在一种版本中,仅针对每少量(例如15)帧来进行关键帧匹配,如前所述。这是合理的,因为漂移仅在变换错误已经针对少量帧累积之后才会明显起来。鉴于前述,图5概述了一种确定关键帧匹配是否要发生的方式。首先,确定自从最后一次执行关键帧匹配以来是否已经输入了规定数目的帧(过程动作500)。如果规定数目的帧还未被输入,那么指定未到用于新的关键帧匹配的时间(过程动作502)。然而,如果已经输入了规定数目的帧,被认为到了执行关键帧匹配过程的时间(过程动作504)。

[0067] 鉴于前述并再次参见图7B,如果在过程动作716中确定关键帧匹配要发生,或如果最后输入的彩色图像帧已经被指定为当前关键帧,那么标识在最后输入的彩色图像帧和描述最后输入的帧中所描述的场景的至少规定部分的关键帧(此后将被称为经标识的关键帧)之间的匹配点(过程动作752)。如前所述,在一个实现中,出于该目的使用传统的SURF匹配过程。注意,在过程动作720到722以及过程动作724到730中先前描述的可选匹配点细化和异常值消除动作也可以在此实现以细化所标识的关键帧和最后输入帧之间的匹配点位置。

[0068] 一旦在最后输入的彩色图像帧和所标识的关键帧之间的匹配点已经被标识,它们就被用于与来自对应的深度帧的3D位置数据相结合以在过程动作754中估计当前的变换矩

阵(即, $(R_{t,k}, T_{t,k})$)。该变换矩阵定义了点从所标识的关键帧(即k)到最后输入帧(即t)的平移和旋转。注意, 该变换矩阵估计是以与在过程动作732中估计在连续帧之间的变换矩阵相同的方式完成的。另外, 在过程动作734到746中先前所述的可选的迭代异常值消除动作也可以在此实现以潜在地增加所估计的变换的精确性。随后, 使用在所标识的关键帧和最后输入的彩色图像帧之间获得的当前的变换矩阵以及定义了点从第一彩色图像帧输入(即帧1)到所标识的关键帧的平移和旋转的先前计算的关键帧变换矩阵(即 $(R_{1,t}, T_{1,t})$)来估计定义了点从第一彩色图像帧输入到最后输入彩色图像帧的平移和旋转的关键帧匹配的变换矩阵(过程动作756)。在一个实现中, 这是通过将所标识的关键帧和最后输入帧之间估计的变换矩阵与所述第一帧和所标识的关键帧之间的先前计算的变换矩阵相乘来完成的。在最后输入的彩色图像帧已经被指定为当前关键帧的情况下, 与该帧相关联的关键帧匹配的变换矩阵随后被指定为用于该帧的关键帧变换矩阵(过程动作757)。这建立了定义点从第一彩色图像帧输入到当前关键帧的平移和旋转的预计算的关键帧变换矩阵, 用于在创建下一关键帧时计算关键帧匹配的变换矩阵。另外, 估计定义点从第一惯性传感器帧输入到最后输入的彩色图像帧的平移和旋转的变换矩阵, 所述估计通过将定义了点从第一惯性传感器帧到紧接在该最后输入的彩色图像帧之前的彩色图像帧的平移和旋转的单个经组合的变换矩阵与定义了点从所述紧接之前的彩色图像帧到该最后输入的彩色图像帧的平移和旋转的当前变换矩阵进行累加来完成(过程动作758)。如前所述, 由于对于光流跟踪来说, 以一开始从每个帧到下一帧所估计的变换(即, $(R_{12}, T_{12}), (R_{23}, T_{23}), \dots, (R_{k-1}, T_{k-1,k}), (R_{k-1,k}, T_{k-1,k})$)可以被累加以使用传统方法产生从第一帧关键帧(即帧1)到最当前帧(即帧k)的变换矩阵, 所以, 这是可能的。

[0069] 在过程中的该点处, 要么基于惯性数据和光流跟踪, 要么基于惯性数据、光流跟踪和关键帧匹配, 定义点从第一惯性传感器帧输入到最后输入的惯性传感器帧的平移和旋转的单独的变换矩阵存在。在第一种情况中, 基于光流跟踪的定义了点从第一惯性传感器帧输入到最后输入的彩色图像帧的平移和旋转的变换矩阵与定义了点从第一惯性传感器帧输入到最后输入的惯性传感器帧的平移和旋转的最终基于惯性传感器的变换矩阵被合并以产生定义了点从第一惯性传感器帧输入到最后输入的惯性传感器帧的平移和旋转的单个经组合的变换矩阵(过程动作760)。在后一种情况中, 基于光流跟踪和关键帧匹配两者的定义了点从第一惯性传感器帧输入到最后输入的彩色图像帧的平移和旋转的变换矩阵以及定义了点从第一惯性传感器帧输入到最后输入的惯性传感器帧的平移和旋转的最终基于惯性传感器的变换矩阵被合并以产生定义了点从第一惯性传感器帧输入到最后输入的惯性传感器帧的平移和旋转的单个经组合的变换矩阵(过程动作762)。

[0070] 在一个实现中, 使用递归贝叶斯框架和扩展的卡尔曼(Kalman)滤波器(EKF)来合并变换矩阵。具体而言, 一种经常使用且强大的用于处理这种合并和状态估计问题的工具是递归贝叶斯滤波器。它可以将来自不同源的测量合并到一个统一的框架中, 并通过建立的处理模型和观察模型来估计隐藏状态。它也可以自然地处理不同传感器的不同帧速率。

[0071] 递归贝叶斯滤波器被用于通过建模系统并测量系统输出来估计系统的隐藏状态变量的后验概率分布函数(PDF)。对于具有状态向量x的给定系统, 其中x可能不是可察觉的, 其处理通常可以如下被建模:

[0072] $x_k = f(x_{k-1}, u_k, v_{k-1}),$

[0073] 其中, u_k 是到系统的输入, 而 v_k 是处理噪声。该公式描述了状态变量随输入 u_k 的影响的时间演变。

[0074] 在相同时间, 系统的可测量的输出可以使用下述测量公式与状态变量相关:

$$[0075] \quad z_k = h(x_k, e_k),$$

[0076] 其中 z_k 是观察向量, 而 e_k 是测量噪声。

[0077] 想法是在给定 $z_{1:k}$ 的情况下估计 x_k 的后验分布。众所周知, 可以通过应用贝叶斯定理来获得递归方案, 导致:

$$[0078] \quad p(x_k | z_{1:k}) = \frac{p(z_k | x_k) p(x_k | z_{1:k-1})}{\int p(z_k | x_k) p(x_k | z_{1:k-1}) dx_k}$$

$$[0079] \quad p(x_{k+1} | z_{1:k}) = \int p(x_{k+1} | x_k) p(x_k | z_{1:k}) dx_k.$$

[0080] 在上述公式中的多维积分并不能在所有情况下都被分析上解出, 而是在少数特殊情况下才能解出。通常的情况是当处理公式和测量公式都是线性的且处理噪声和测量噪声可以使用零均值高斯分布被建模时。

$$[0081] \quad x_k = Ax_{k-1} + Bu_k + v_{k-1}$$

$$[0082] \quad z_k = Hx_k + e_k$$

$$[0083] \quad p(w) \sim N(0, Q)$$

$$[0084] \quad p(v) \sim N(0, R),$$

[0085] 其中 Q 是处理噪声协方差, 而 R 是测量噪声协方差。

[0086] 在这种情况下, 所有的密度将是高斯的, 且它对于仅传播均值和协方差来说是足够的。这导致著名的卡尔曼(Kalman)滤波器:

[0087] 时间更新:

$$[0088] \quad \hat{x}_k^- = A\hat{x}_{k-1} + Bu_k$$

$$[0089] \quad P_k^- = AP_{k-1}A^T + Q$$

[0090] 测量更新:

$$[0091] \quad K_k = P_k^- H^T (HP_k^- H^T + R)^{-1}$$

$$[0092] \quad \hat{x}_k = \hat{x}_k^- + K_k(z_k - H\hat{x}_k^-)$$

$$[0093] \quad P_k = (I - K_k H) P_k^-.$$

[0094] 在模型是非线性的情况中, 可以通过计算部分求导来局部地对其线性化。这导致扩展的卡尔曼滤波器(EKF):

$$[0095] \quad A_{[k]} = \frac{\partial f_{[k]}}{\partial x_{[k]}}(\hat{x}_{k-1}, u_k, 0)$$

$$[0096] \quad V_{[k]} = \frac{\partial f_{[k]}}{\partial w_{[k]}}(\hat{x}_{k-1}, u_k, 0)$$

$$[0097] \quad H_{[k]} = \frac{\partial h_{[k]}}{\partial x_{[k]}}(\hat{x}_k, 0)$$

$$[0098] \quad E_{[i,j]} = \frac{\partial h_{[i,j]}}{\partial v_{[i,j]}}(\hat{x}_k, 0)$$

[0099] 时间更新:

$$[0100] \quad \hat{x}_k^- = f(\hat{x}_{k-1}, u_k, 0)$$

$$[0101] \quad P_k^- = A_k P_{k-1} A_k^T + W_k Q_{k-1} W_k^T$$

[0102] 测量更新:

$$[0103] \quad K_k = P_k^- H_k^T (H_k P_k^- H_k^T)^{-1}$$

$$[0104] \quad \hat{x}_k = \hat{x}_k^- + K_k (z_k - h(\hat{x}_k^-, 0))$$

$$[0105] \quad P_k = (I - K_k H_k) P_k^-$$

[0106] 给定前述,可以定义一个系统模型。具体而言,状态变量应该包括头部的位置和朝向。还可包括线速度和加速度以及旋转速度,以便可以假设恒定的线加速和旋转速度。还包括传感器偏置以处理惯性传感器的轻微变化偏置。注意,下面假设惯性传感器是具有公共坐标帧原点的加速度计(a)和陀螺仪(g)。这样,在一种实现中,

$$[0107] \quad x_k = (p_k^W, \dot{p}_k^W, \ddot{p}_k^W, q_k^W, w_k^I, b_{g,k}^I, b_{a,k}^I)$$

[0108] 其中 p_k^W 是帧“W”中所描述的头部位置。 $\dot{p}_k^W$ 和 $\ddot{p}_k^W$ 分别是头部的线速度和加速度。 q_k^W 是帧“W”中使用四元组描述的头部朝向。 w_k^I 是在帧“I”中描述的头部的角速度。偏置 $b_{g,k}^I, b_{a,k}^I$ 也是在惯性传感器坐标帧中。四元组被用于描述朝向,因为它在跟踪空间中是连续的。

[0109] 接着,处理被如下建模:

$$[0110] \quad p_k^W = p_{k-1}^W + T \dot{p}_{k-1}^W + \frac{T^2}{2} \ddot{p}_{k-1}^W$$

$$[0111] \quad \dot{p}_k^W = \dot{p}_{k-1}^W + T \ddot{p}_{k-1}^W$$

$$[0112] \quad \ddot{p}_k^W = \ddot{p}_{k-1}^W + v_{\ddot{p},k-1}^W$$

$$[0113] \quad q_k^W = q_{k-1}^W \otimes \exp\left(\frac{T}{2} w_{k-1}^I\right)$$

$$[0114] \quad w_k^I = w_{k-1}^I + v_{w,k-1}^I$$

$$[0115] \quad b_{g,k}^I = b_{g,k-1}^I + v_{g,k-1}^I$$

$$[0116] \quad b_{a,k}^I = b_{a,k-1}^I + v_{a,k-1}^I$$

[0117] 所述操作被定义为:

$$[0118] \quad w = \begin{pmatrix} w_x \\ w_y \\ w_z \end{pmatrix}$$

$$[0119] \quad \exp(w) \triangleq \begin{pmatrix} \cos\|w\| \\ \frac{w}{\|w\|} \sin\|w\| \end{pmatrix}$$

$$[0120] \quad q_1 = \begin{pmatrix} a_1 \\ x_1 \\ y_1 \\ z_1 \end{pmatrix}$$

$$[0121] \quad q_2 = \begin{pmatrix} a_2 \\ x_2 \\ y_2 \\ z_2 \end{pmatrix}$$

$$[0122] \quad q_1 \odot q_2 \triangleq \begin{pmatrix} a_1 a_2 - \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} \cdot \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix} \\ a_1 \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix} + \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} a_2 + \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} \times \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix} \end{pmatrix}$$

[0123] 对于观察公式,它们需要解决惯性传感器数据以及彩色图像和深度图像数据。对于惯性传感器(即陀螺仪和加速度计),观察公式是:

$$[0124] \quad \tilde{x}_{w,k}^I = (R(q_k^W))^{-1}(\tilde{p}_k^W + \tilde{q}^W) + h_{w,k}^I + e_{w,k}^I$$

$$[0125] \quad \tilde{x}_{g,k}^I = w_k^I + b_{g,k}^I + e_{g,k}^I$$

[0126] 然而,关于如何使用图像数据,存在不同的选项。在一个实现中,为了线性化观察公式,变换估计被用作测量。来自光流跟踪和关键帧匹配(如果可用的话)这两者的估计被如下包括:

[0127]

$$\tilde{x}_{c,q,k-s,k}^{C_{k-s}} = (q_{k-s}^W \odot q_{IC})^{-1} \odot (q_k^W \odot q_{IC}) + e_{c,q,k-s,k}^{C_{k-s}}$$

$$\tilde{x}_{c,T,k-s,k}^{C_{k-s}} = (R(q_{IC}))^{-1} (R(q_{k-s}^W))^{-1} \left((R(q_k^W)T_{IC} + p_k^W) - (R(q_{k-s}^W)T_{IC} + p_{k-s}^W) \right) + e_{c,q,k-s,k}^{C_{k-s}}$$

[0128]

$$\tilde{x}_{c,q,k-t,k}^{C_{k-t}} = (q_{k-t}^W \odot q_{IC})^{-1} \odot (q_k^W \odot q_{IC}) + e_{c,q,k-t,k}^{C_{k-t}}$$

$$\tilde{x}_{c,T,k-t,k}^{C_{k-t}} = (R(q_{IC}))^{-1} (R(q_{k-t}^W))^{-1} \left((R(q_k^W)T_{IC} + p_k^W) - (R(q_{k-t}^W)T_{IC} + p_{k-t}^W) \right) + e_{c,q,k-t,k}^{C_{k-t}}$$

[0129] 其中 $\tilde{x}_{c,q,k-s,k}^{C_{k-s}}$ 是描述在相机坐标帧中在k-s处从光流流水线所获得的时间阶段k-s到当前时间阶梯k的旋转的四元组。 q_{IC} 是从帧“I”到帧“C”的已知旋转。 $\tilde{x}_{c,T,k-s,k}^{C_{k-s}}$ 是在相机

坐标帧中在k-s处从光流流水线所获得的时间阶梯k-s到阶梯k的平移。 $\begin{matrix} \text{---} & \text{---} & \text{---} & \text{---} & \text{---} \\ \text{---} & \text{---} & \text{---} & \text{---} & \text{---} \\ \text{---} & \text{---} & \text{---} & \text{---} & \text{---} \end{matrix}$ 是描述从关键帧流水线所获得的时间阶梯k-t到当前时间阶梯k的旋转的四元组。时间阶梯t是在其处记录了关键帧的帧。 $\begin{matrix} \text{---} & \text{---} & \text{---} & \text{---} & \text{---} \\ \text{---} & \text{---} & \text{---} & \text{---} & \text{---} \\ \text{---} & \text{---} & \text{---} & \text{---} & \text{---} \end{matrix}$ 是从该关键帧流水线的平移。注意,当不存在关键帧匹配变换矩阵且仅有基于惯性数据和光流跟踪的矩阵时,用于关键帧匹配变换矩阵的内插权重被设定为零。

[0130] 注意,如果Q和R参数被微调,EKF将运行得更加有效。如前所述,其中Q是处理噪声协方差,而R是测量噪声协方差。通常,Q越小,处理越精确,且R越小,测量越精确。R可以被离线测量,但Q仅能基于先前知识或移动的假设来微调。

[0131] 一旦已经生成了单个经组合的变换矩阵,将它应用于先前确定的与第一惯性传感器帧输入相关联的头部姿态位置和朝向以标识当前的头部姿态位置和朝向(过程动作764)。随后,可将当前头部姿态和经组合的变换矩阵用于各种其它任务,包括计算在增强现实应用中的当前帧的虚拟对象的位置。

[0132] 随后,只要捕捉到新的惯性传感器、彩色视频和深度帧并将其输入到计算设备中,就重复前述过程。

[0133] 2.0示例性头部姿态跟踪应用

[0134] 如前所述,以上述方式来跟踪用户的头部姿态在包括增强现实的各种应用中是有用的。在增强现实应用中,需要精确且强健的头部姿态跟踪系统来确保虚拟对象的稳定显示。除了上述用于跟踪用户头部姿态的头部承载的传感器之外,这样的增强现实应用还将包括头部佩戴显示器(HMD)。例如,在一个实现中,HMD采用了安置在用户一只眼睛前面的单个半透明眼镜显示器的形式。

[0135] 3.0示例性操作环境:

[0136] 本文描述的头部姿态跟踪技术的各实施例可在多种通用或专用计算机系统环境或配置内操作。图8示出其上可实现本文描述的头部姿态跟踪技术的各实施例的实施方式和元素的通用计算机系统的简化示例。应当注意,图8中由折线或虚线所表示的任何框表示简化计算设备的替换实施方式,并且以下描述的这些替换实施方式中的任一个或全部可以结合贯穿本文所描述的其他替换实施方式来使用。

[0137] 例如,图8示出了总系统图,其示出简化计算设备10。这样的计算设备通常可以在具有至少一些最小计算能力的设备中找到,这些设备包括但不限于个人计算机、服务器计算机、手持式计算设备、膝上型或移动计算机、诸如蜂窝电话和PDA等通信设备、多处理器系统、基于微处理器的系统、机顶盒、可编程消费电子产品、网络PC、小型计算机、大型计算机、视频媒体播放器等。

[0138] 为允许设备实现本文描述的头部姿态跟踪技术的各实施例,该设备应当具有足够的计算能力和系统存储器以启用基本计算操作。具体而言,如图8所示,计算能力一般由一个或多个处理单元12示出,并且还可包括一个或多个GPU14,这两者中的任一个或全部与系统存储器16通信。注意,通用计算设备的处理单元12可以是专用微处理器,如DSP、VLIW、或其他微控制器、或可以是一个或多个处理核的常规CPU,包括多核CPU中的专用的基于GPU核。

[0139] 另外,图8的简化计算设备还可包括其他组件,诸如例如通信接口18。图8的简化计

算设备还可包括一个或多个常规计算机输入设备20(例如,定点设备、键盘、音频输入设备、视频输入设备、触觉输入设备、用于接收有线或无线数据传输的设备等)。图8的简化计算设备还可包括其他光学组件,诸如例如一个或多个常规显示设备24和其他计算机输出设备22(例如,音频输出设备、视频输出设备、用于传送有线或无线数据传输的设备等)。注意,通用计算机的典型的通信接口18、输入设备20、输出设备22、以及存储设备26对本领域技术人员而言是公知的,并且在此不会详细描述。

[0140] 图8的简化计算设备还可包括各种计算机可读介质。计算机可读介质可以是可由计算机10经由存储设备26访问的任何可用介质,并且包括是可移动28和/或不可移动30的易失性和非易失性介质,该介质用于存储诸如计算机可读或计算机可执行指令、数据结构、程序模块或其他数据等信息。作为示例而非限制,计算机可读介质可包括计算机存储介质和通信介质。计算机存储介质包括但不限于:计算机或机器可读介质或存储设备,诸如DVD、CD、软盘、磁带驱动器、硬盘驱动器、光盘驱动器、固态存储器设备、RAM、ROM、EEPROM、闪存或其他存储器技术、磁带盒、磁带、磁盘存储或其他磁存储设备、或可用于存储所需信息并且可由一个或多个计算设备访问的任何其他设备。

[0141] 诸如计算机可读或计算机可执行指令、数据结构、程序模块等信息的保留还可通过使用各种上述通信介质中的任一种来编码一个或多个已调制数据信号或载波或其他传输机制或通信协议来实现,并且包括任何有线或无线信息传递机制。注意,术语“已调制数据信号”或“载波”一般指以对信号中的信息进行编码的方式设置或改变其一个或多个特征的信号。例如,通信介质包括诸如有线网络或直接线连接等携带一个或多个已调制数据信号的有线介质,以及诸如声学、RF、红外线、激光和其他无线介质等用于传送和/或接收一个或多个已调制数据信号或载波的无线介质。上述通信介质的任一组合也应包括在通信介质的范围之内。

[0142] 此外,可以按计算机可执行指令或其他数据结构的形式存储、接收、传送或者从计算机或机器可读介质或存储设备和通信介质的任何所需组合中读取具体化本文描述的头部姿态跟踪技术实施方式中的部分或全部的软件、程序和/或计算机程序产品或其各部分。

[0143] 最终,本文描述的头部姿态跟踪技术的各实施例还可在由计算设备执行的诸如程序模块等计算机可执行指令的一般上下文中描述。一般而言,程序模块包括执行特定任务或实现特定抽象数据类型的例程、程序、对象、组件、数据结构等。本文描述的各实施例还可以在其中一个任务由通过一个或多个通信网络链接的一个或多个远程处理设备执行或者在该一个或多个设备的云中执行的分布式计算环境中实现。在分布式计算环境中,程序模块可以位于包括媒体存储设备在内的本地和远程计算机存储介质中。此外,上述指令可以部分地或整体地作为可以包括或不包括处理器的硬件逻辑电路来实现。

[0144] 其他实施例

[0145] 在又一个示例性硬件系统实现中,深度传感器装置可以与一套惯性传感器一起使用,但不存在彩色相机。该实现使用深度传感器和迭代最接近的点(ICP)过程来计算定义了从第一深度帧输入到最后输入的深度帧的平移和旋转的变换矩阵。该基于深度帧的变换矩阵随后将取代先前所述的最终的基于图像的变换矩阵,所述最终的基于图像的变换矩阵是使用采用彩色摄像机、深度传感器和惯性传感器的先前所述的示例性系统中的彩色图像帧来估计的。这样,基于深度帧的变换矩阵与最终的基于惯性传感器的变换矩阵合并以

产生定义了从第一惯性传感器帧输入到最后输入的惯性传感器帧的平移和旋转的单个经组合的变换矩阵,并且随后将该经组合的变换矩阵应用于先前确定的与第一惯性传感器帧输入相关联的头部姿态位置和朝向以标识当前的头部姿态位置和朝向。

[0146] 还应当注意,可以按所需的任何组合来使用本说明书全文中的上述实施例的任一个或全部以形成另外的混合实施例。另外,尽管用结构特征和/或方法动作专用的语言描述了本主题,但可以理解,所附权利要求书中定义的主题不必限于上述具体特征或动作。更确切而言,上述具体特征和动作是作为实现权利要求的示例形式公开的。

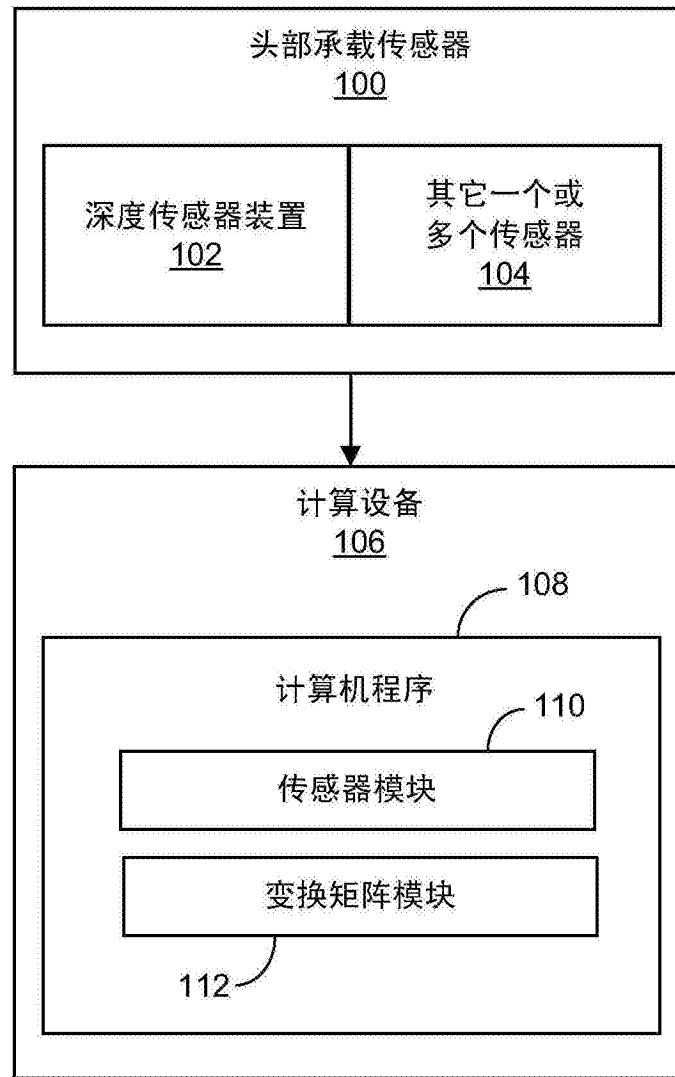


图1

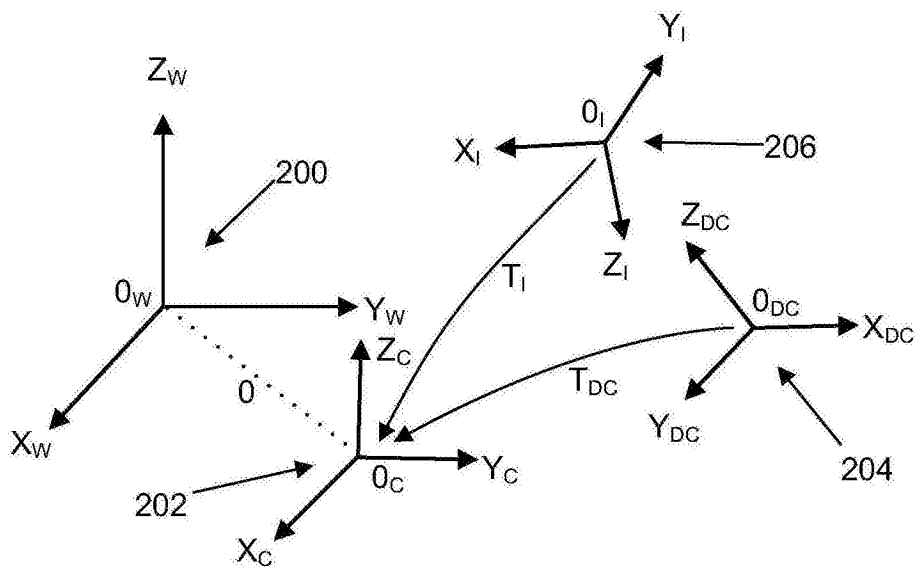


图2

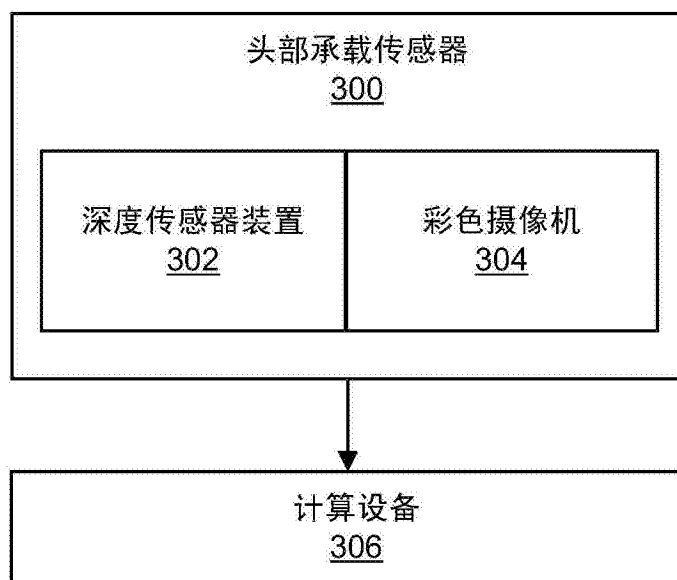


图3

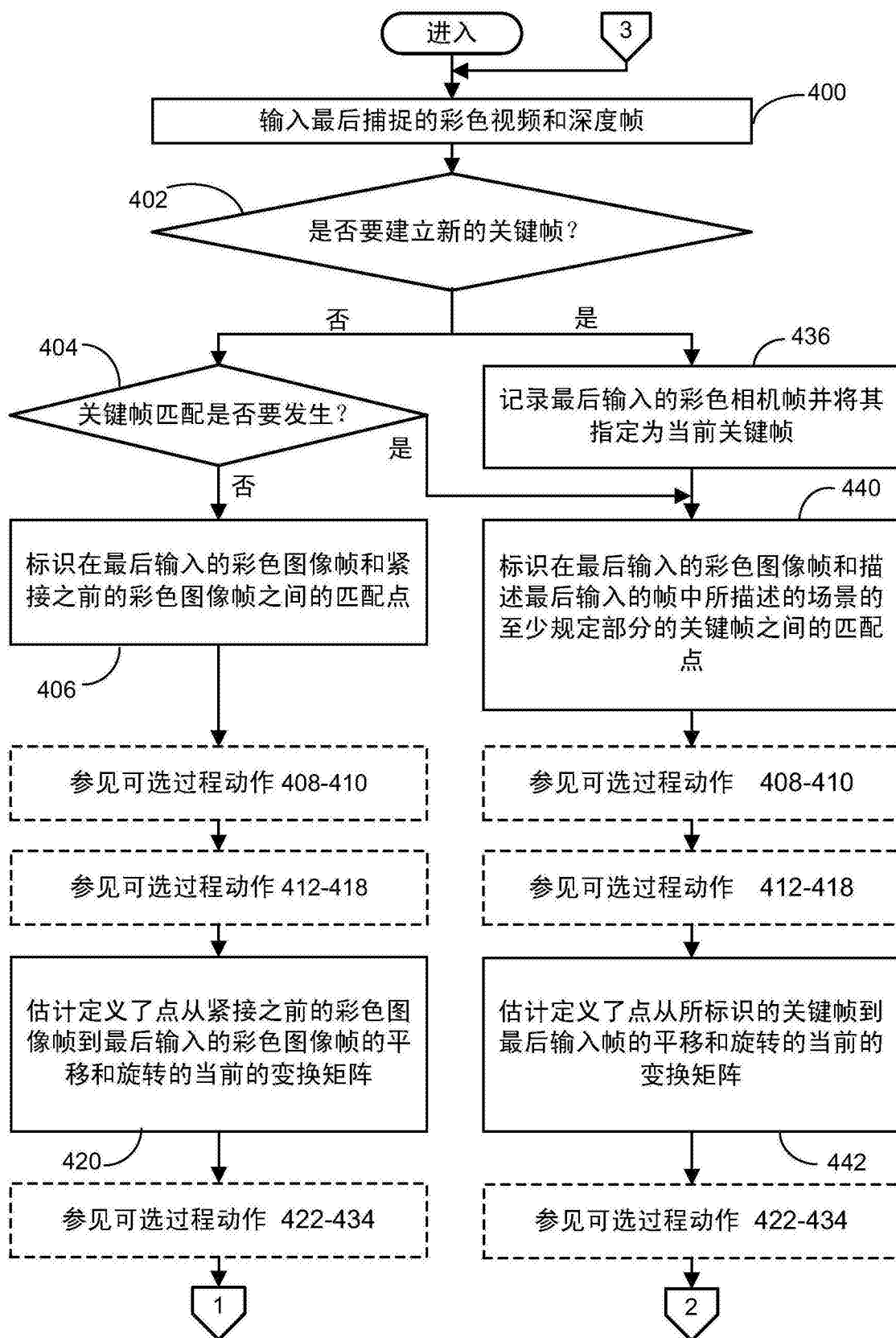


图4A

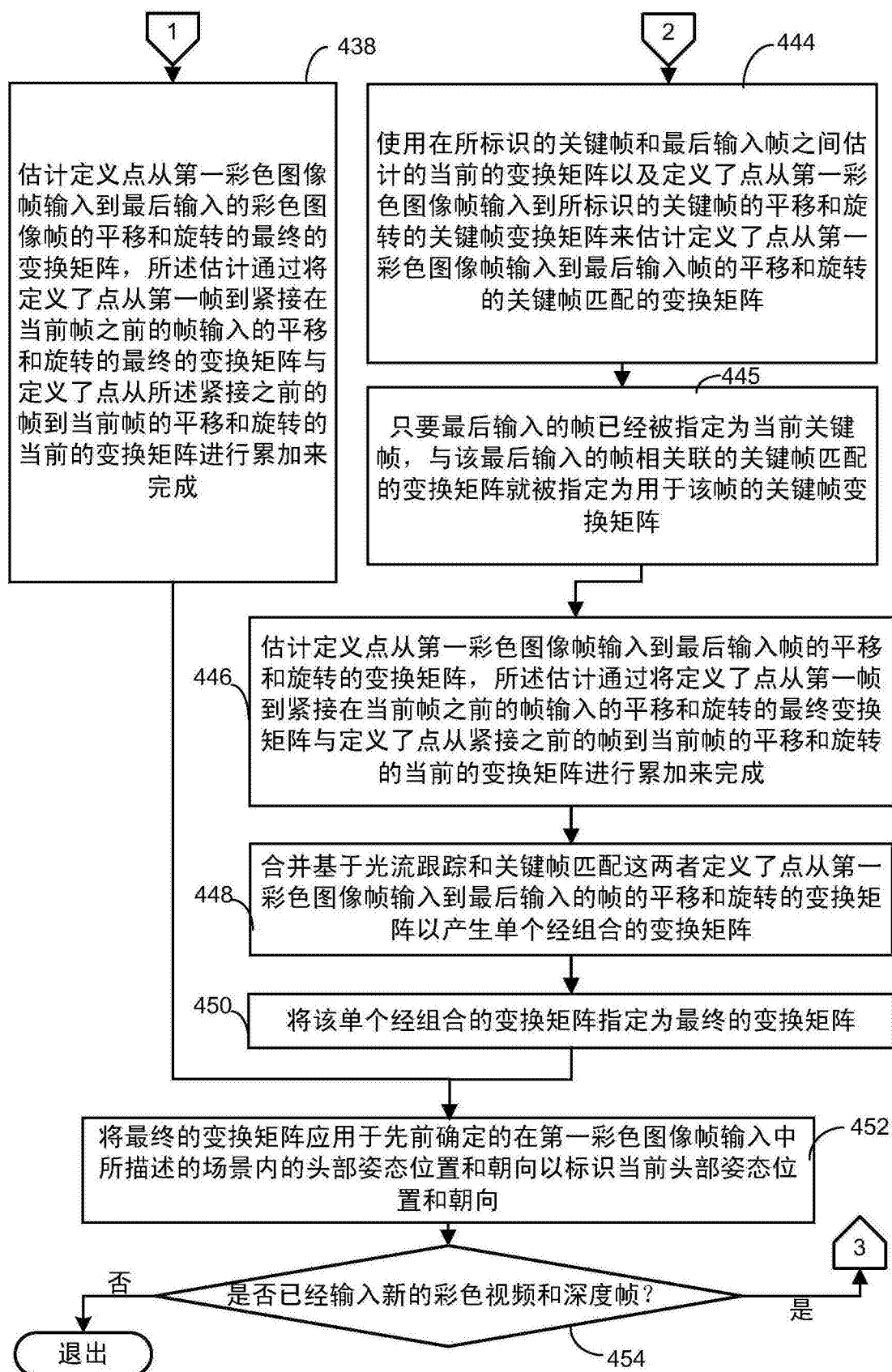


图4B

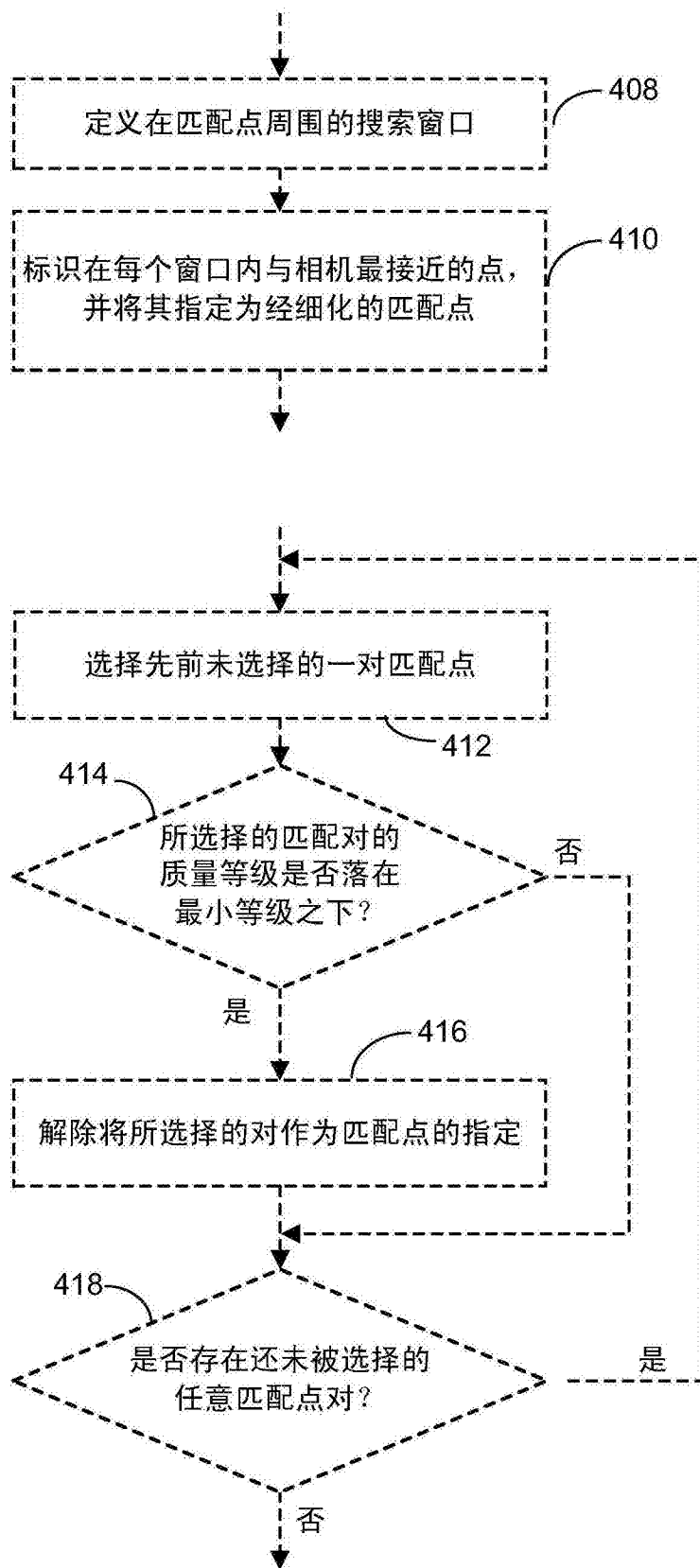


图4C

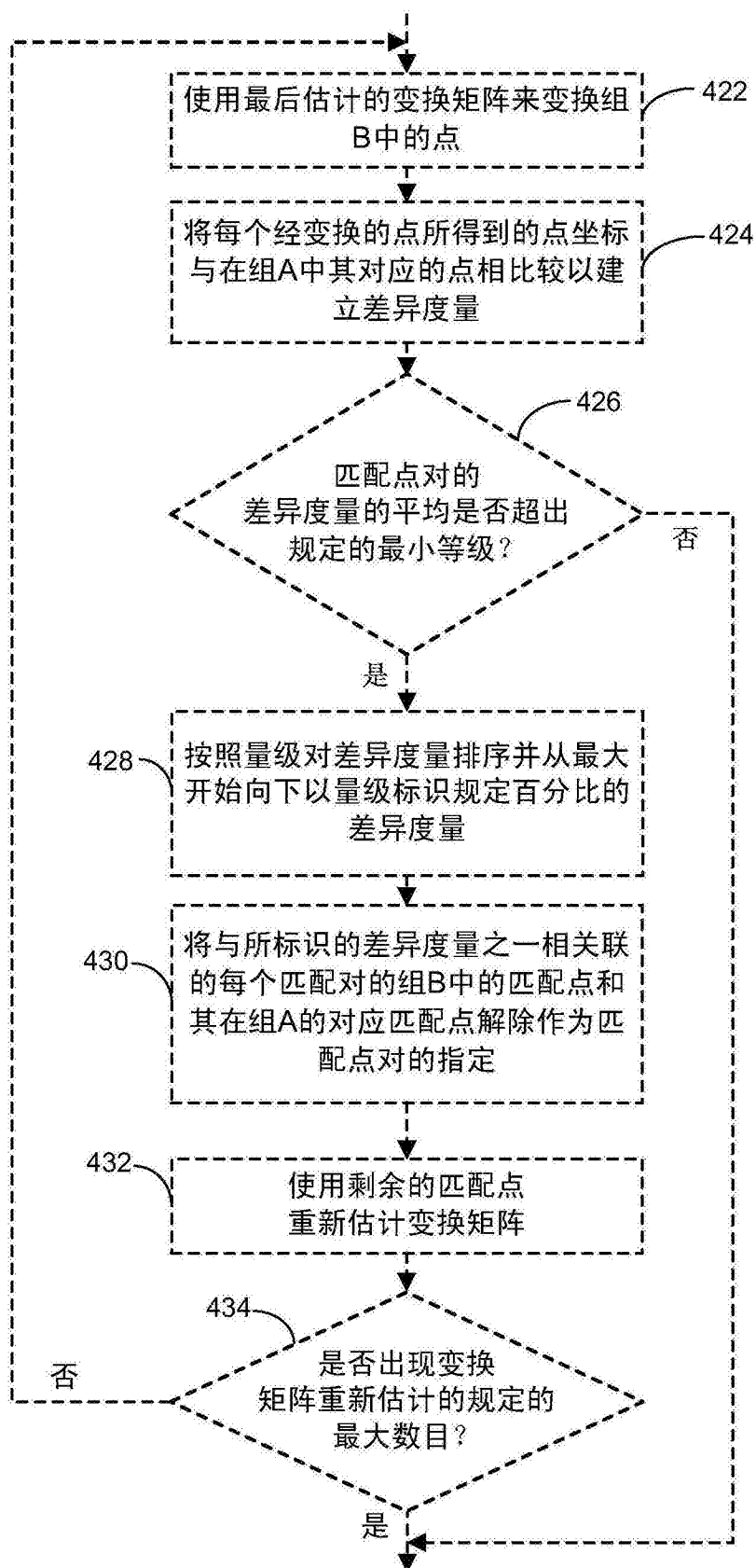


图4D

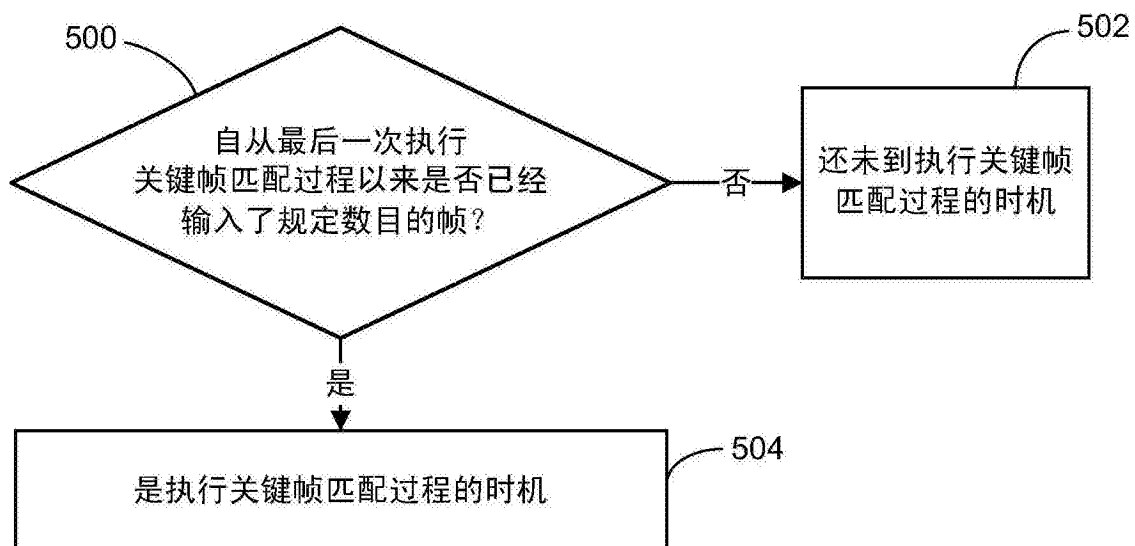


图5

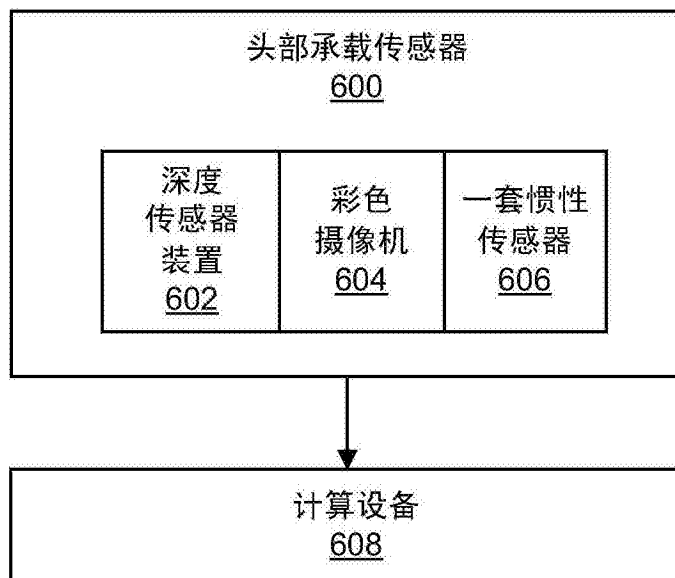


图6

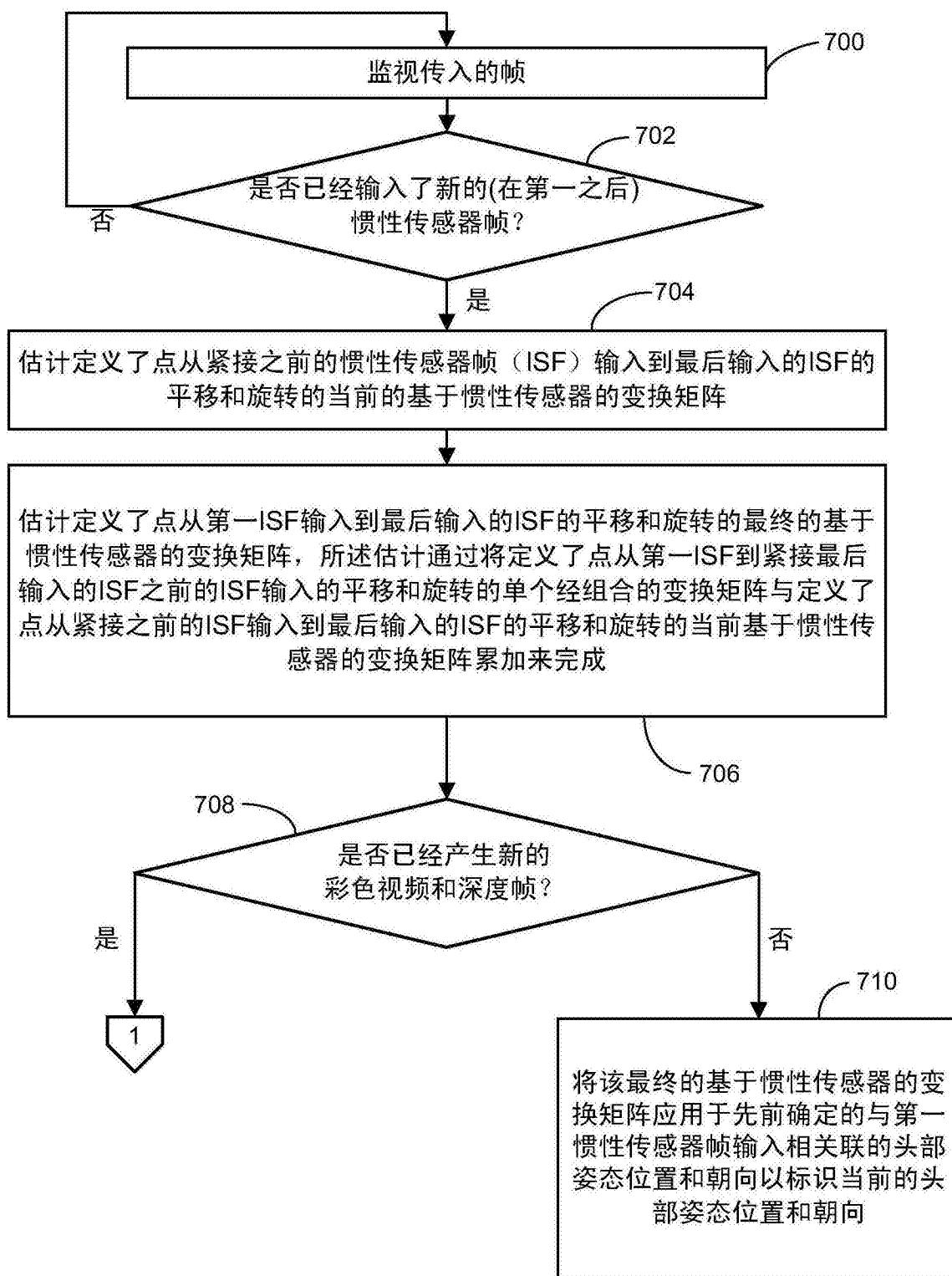


图7A

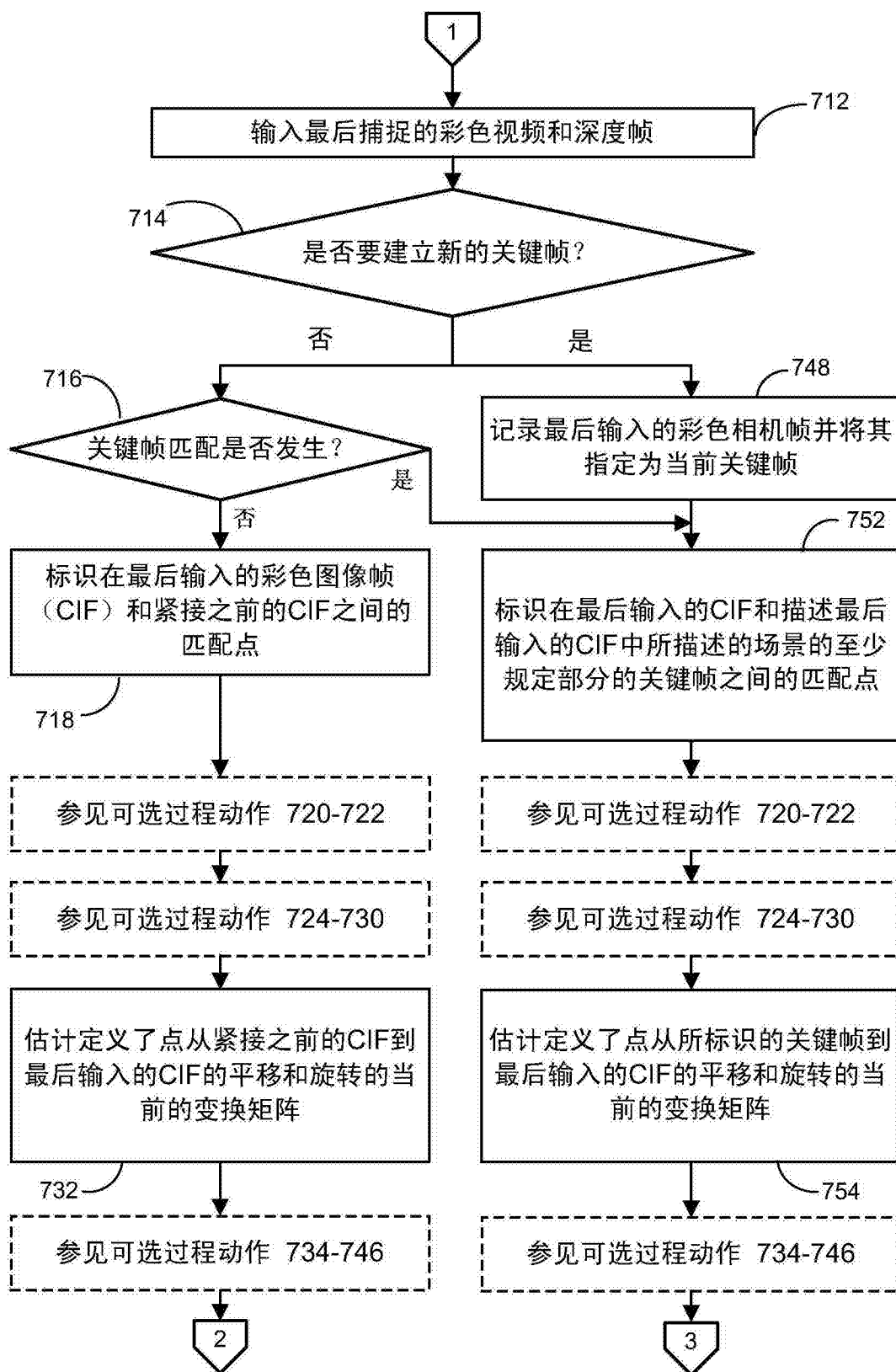


图7B

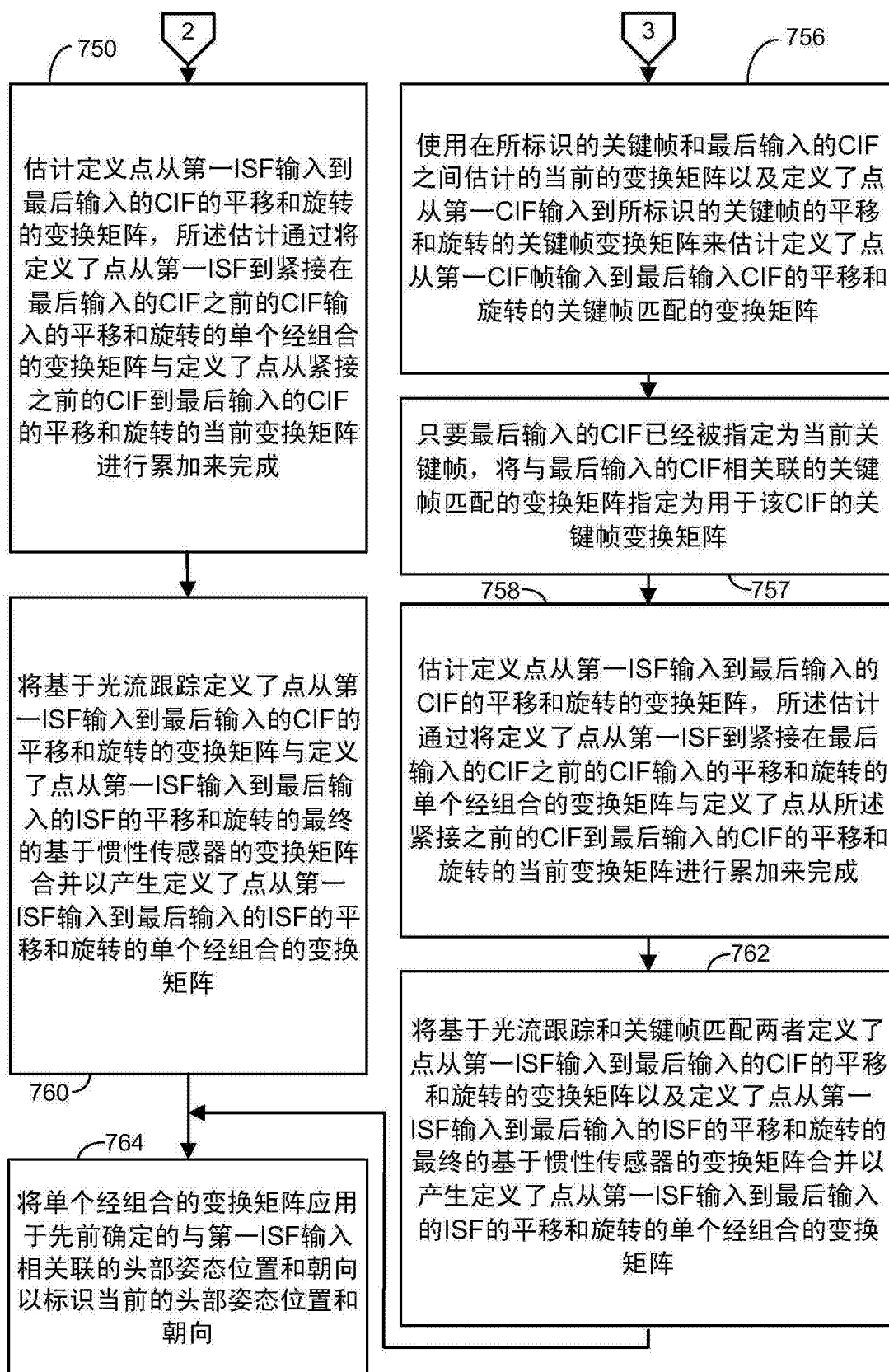


图7C

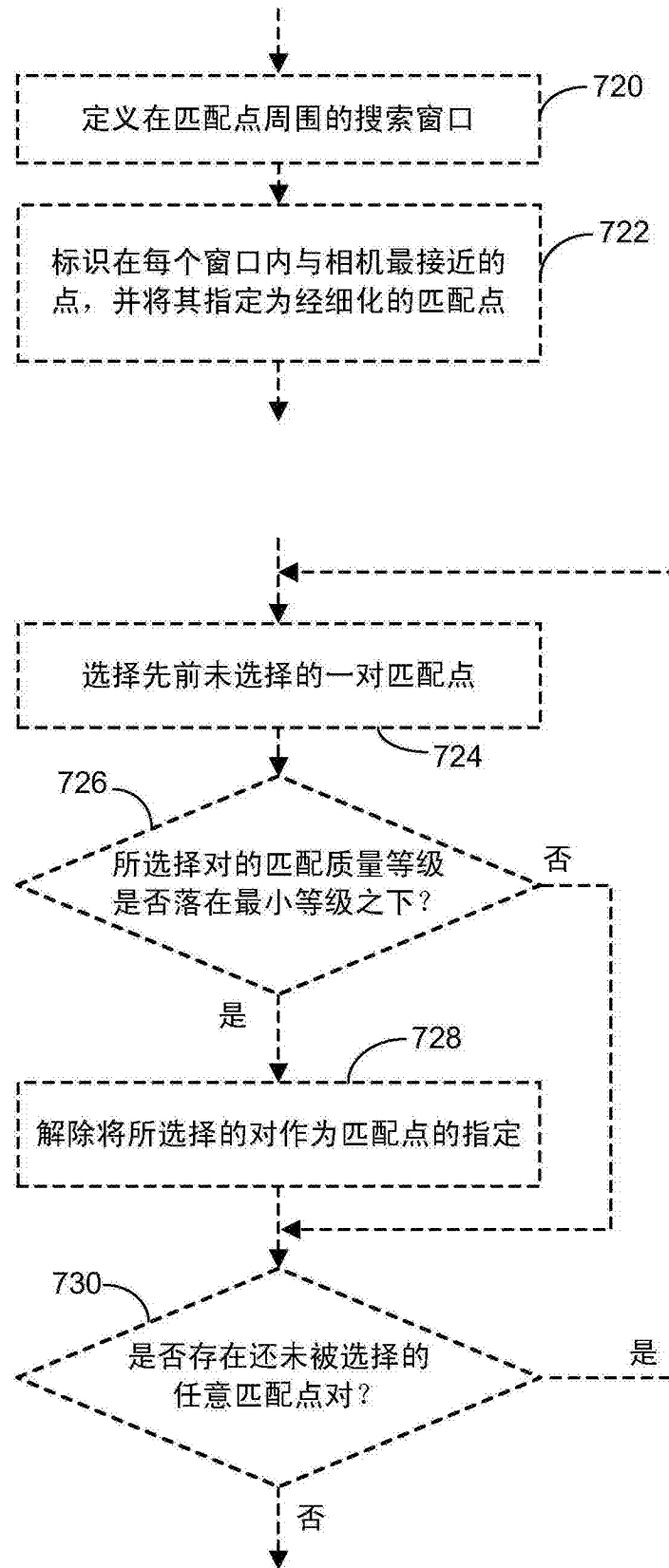


图7D

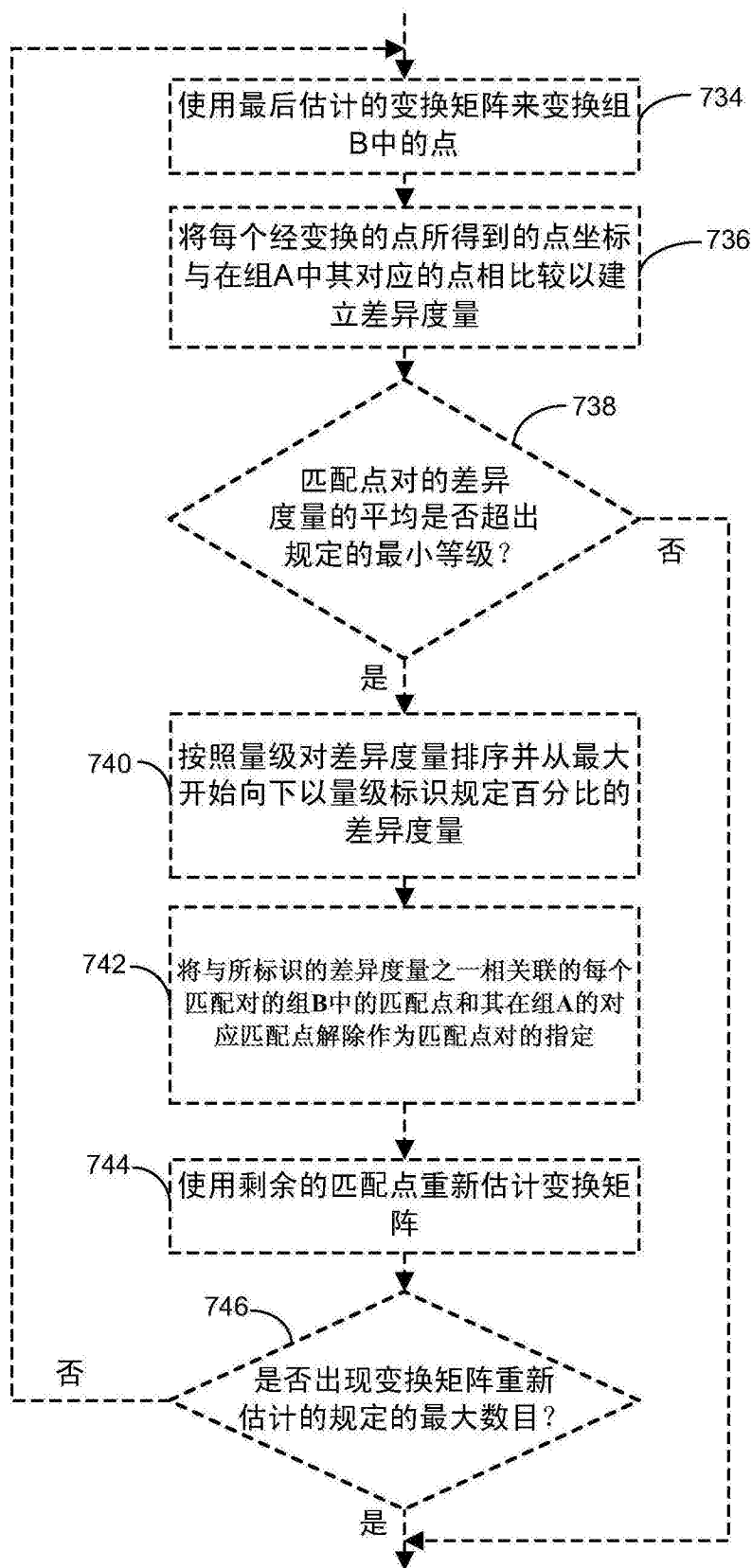


图7E

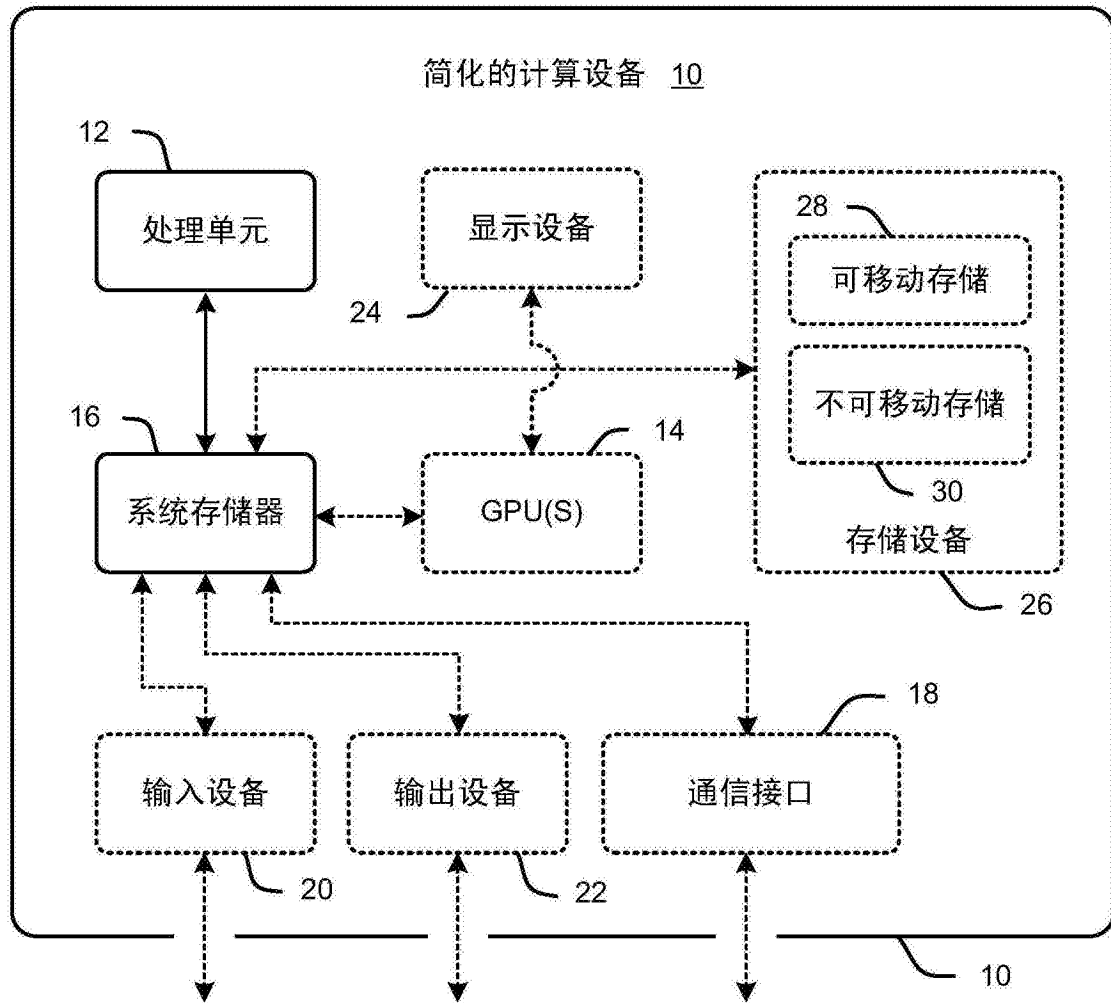


图8