



(12)发明专利

(10)授权公告号 CN 104769643 B

(45)授权公告日 2017.08.11

(21)申请号 201380057461.6

(22)申请日 2013.10.16

(65)同一申请的已公布的文献号
申请公布号 CN 104769643 A

(43)申请公布日 2015.07.08

(30)优先权数据
61/722,089 2012.11.02 US
13/789,577 2013.03.07 US

(85)PCT国际申请进入国家阶段日
2015.04.30

(86)PCT国际申请的申请数据
PCT/US2013/065249 2013.10.16

(87)PCT国际申请的公布数据
W02014/070457 EN 2014.05.08

(73)专利权人 高通股份有限公司

地址 美国加利福尼亚州

(72)发明人 玛荷许·瑞玛屈安卓恩
丹尼尔·瓦格纳
穆拉利·拉马斯瓦米·查里

(74)专利代理机构 北京律盟知识产权代理有限公司 11287

代理人 宋献涛

(51)Int.Cl.
G06T 7/246(2017.01)

审查员 马晋涛

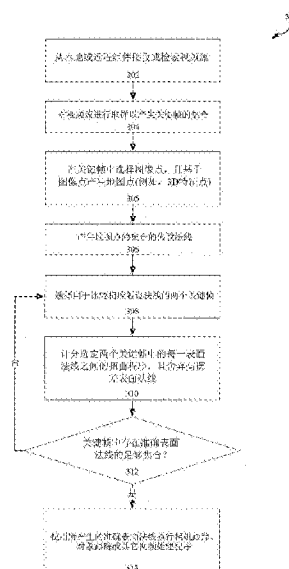
权利要求书3页 说明书10页 附图7页

(54)发明名称

用于使用可平行化架构中的图像来初始化
及解算面元的局部几何形状或表面法线的方法

(57)摘要

本文中描述用于解算视频流中所观测的场景中的对象的表面法线的系统及方法。所述系统及方法可包含：对所述视频流进行取样以产生关键帧的集合；产生所述关键帧中的每一者中的地图点的集合的假设表面法线；通过从所述假设表面法线中的每一者计算的扭曲矩阵使第一关键帧中的对应地图点的贴片扭曲到第二关键帧的视点；计算所述两个关键帧中的每一假设表面法线之间的扭曲误差；及舍弃在所述第一与第二关键帧之间具有高扭曲误差的假设表面法线。



1. 一种用于解算视频流中的对象的表面法线的方法,其包括:
对所述视频流进行取样以产生关键帧的集合;
选择关键帧的所述集合中的多个图像点;
使用关键帧的所述集合中的所述多个图像点产生由关键帧的所述集合成像的场景中的所述多个图像点的对应的3D特征点;
产生对应于所述3D特征点中的每一者的假设表面法线;
通过从所述假设表面法线中的每一者计算的扭曲矩阵使对应于第一关键帧中的3D特征点的图像像素贴片扭曲到第二关键帧的视点;
计分所述第一关键帧及所述第二关键帧之间的由每一假设表面法线产生的扭曲误差;
及
舍弃产生高于预定水平的所述第一关键帧及所述第二关键帧之间的扭曲误差得分的假设表面法线。
2. 根据权利要求1所述的方法,其进一步包括:
将未舍弃的假设表面法线标记为准确估计。
3. 根据权利要求2所述的方法,其进一步包括:
确定经标记表面法线的数目是否大于或等于表面法线的预定义数目;及
在确定经标记表面法线的所述数目小于表面法线的所述预定义数目后,即刻对一对新关键帧重复扭曲、计分及舍弃操作。
4. 根据权利要求3所述的方法,其进一步包括:
确定具有最低扭曲误差的假设表面法线。
5. 根据权利要求4所述的方法,其进一步包括:
使用具有所述最低扭曲误差的所述所确定的假设表面法线执行所述视频流中的相机跟踪及目标跟踪中的一者。
6. 根据权利要求4所述的方法,其进一步包括:
使用所述经标记表面法线执行所述视频流中的相机跟踪及目标跟踪中的一者。
7. 根据权利要求4所述的方法,其进一步包括:
对多个关键帧的所述扭曲误差求和。
8. 根据权利要求1所述的方法,其中所述扭曲误差是基于归一化交叉相关性得分。
9. 根据权利要求1所述的方法,其进一步包括:
选择具有在预定义分离范围内的基线的用于扭曲的所述第一及第二关键帧。
10. 根据权利要求1所述的方法,其中所述第一及第二关键帧为所述视频流中的非顺序关键帧。
11. 一种用于解算视频流中的对象的表面法线的计算装置,其包括:
用于对所述视频流进行取样以产生关键帧的集合的取样装置;
用于选择关键帧的所述集合中的多个图像点、使用关键帧的所述集合中的所述多个图像点产生由关键帧的所述集合成像的场景中的所述多个图像点的对应的3D特征点,及产生对应于所述3D特征点中每一者的假设表面法线的假设装置;
用于通过从所述假设表面法线中的每一者计算的扭曲矩阵使对应于第一关键帧中的3D特征点的图像像素贴片扭曲到第二关键帧的视点的扭曲装置;

用于计分所述第一关键帧及所述第二关键帧之间的由每一假设表面法线产生的扭曲误差的计分装置;及

用于舍弃产生高于预定水平的所述第一关键帧及所述第二关键帧之间的扭曲误差得分的假设表面法线的消除装置。

12. 根据权利要求11所述的计算装置,其进一步包括:

用于将未舍弃的假设表面法线标记为准确估计的标记装置。

13. 根据权利要求12所述的计算装置,其进一步包括:

用于确定经标记表面法线的数目是否大于或等于表面法线的预定义数目的确定装置,其中在确定经标记表面法线的所述数目小于表面法线的所述预定义数目后,(1)所述扭曲装置使一对新关键帧中的对应3D特征点的图像像素贴片扭曲,(2)所述计分装置计分所述对新关键帧中的每一假设表面法线之间的扭曲误差,及(3)所述消除装置舍弃具有高扭曲误差的所述一对新关键帧中的假设表面法线。

14. 根据权利要求13所述的计算装置,其中所述标记装置进一步确定具有最低扭曲误差的假设表面法线。

15. 根据权利要求14所述的计算装置,其中经标记假设表面法线用于执行所述视频流中的相机跟踪及对象跟踪中的一者。

16. 根据权利要求14所述的计算装置,其中具有所述最低扭曲误差的所述所确定的假设表面法线用于执行所述视频流中的相机跟踪及目标跟踪中的一者。

17. 根据权利要求14所述的计算装置,其中所述计分装置对多个关键帧的所述扭曲误差求和。

18. 根据权利要求11所述的计算装置,其中所述扭曲误差是基于归一化交叉相关性得分。

19. 根据权利要求11所述的计算装置,其进一步包括:

用于选择具有在预定义分离范围内的基线的用于扭曲的所述第一及第二关键帧的选择装置。

20. 根据权利要求11所述的计算装置,其中所述第一及第二关键帧为所述视频流中的非顺序关键帧。

21. 一种具有用于解算视频流中的对象的表面法线的处理器的计算装置,其包括:

用于俘获所述视频流的相机;

表面法线解算器,其包括:

用于对所述视频流进行取样以产生关键帧的集合的取样单元;

用于选择关键帧的所述集合中的多个图像点、使用关键帧的所述集合中的所述多个图像点产生由关键帧的所述集合成像的场景中的所述多个图像点的对应的3D特征点,及产生对应于所述3D特征点的每一者的假设表面法线的假设单元;

用于通过从所述假设表面法线中的每一者计算的扭曲矩阵使对应于第一关键帧中的3D特征点的图像像素贴片扭曲到第二关键帧的视点的扭曲单元;

用于计分所述第一关键帧及所述第二关键帧之间的由每一假设表面法线产生的扭曲误差的计分单元;及

用于舍弃产生高于预定水平的所述第一关键帧及所述第二关键帧之间的扭曲误差得

分的假设表面法线的消除单元。

22. 根据权利要求21所述的计算装置,其中所述表面法线解算器进一步包括:

用于将未舍弃的假设表面法线标记为准确估计的标记单元。

23. 根据权利要求22所述的计算装置,其中所述表面法线解算器进一步包括:

用于确定经标记表面法线的数目是否大于或等于表面法线的预定义数目的确定单元,其中在确定经标记表面法线的所述数目小于表面法线的所述预定义数目后,(1)所述扭曲单元使一对新关键帧中的对应3D特征点的图像像素贴片扭曲,(2)所述计分单元计分所述对新关键帧中的每一假设表面法线之间的扭曲误差,及(3)所述消除单元舍弃具有高扭曲误差的所述一对新关键帧中的假设表面法线。

24. 根据权利要求23所述的计算装置,其中所述标记单元进一步确定具有最低扭曲误差的假设表面法线。

25. 根据权利要求24所述的计算装置,其中经标记假设表面法线用于执行所述视频流中的相机跟踪及对象跟踪中的一者。

26. 根据权利要求24所述的计算装置,其中具有所述最低扭曲误差的所述所确定的假设表面法线用于执行所述视频流中的相机跟踪及目标跟踪中的一者。

27. 根据权利要求24所述的计算装置,其中所述计分单元对多个关键帧的所述扭曲误差求和。

28. 根据权利要求21所述的计算装置,其中所述扭曲误差是基于归一化交叉相关性得分。

29. 根据权利要求21所述的计算装置,其中所述表面法线解算器进一步包括:

用于选择具有在预定义分离范围内的基线的用于扭曲的所述第一及第二关键帧的选择单元。

30. 根据权利要求21所述的计算装置,其中所述第一及第二关键帧为所述视频流中的非顺序关键帧。

用于使用可平行化架构中的图像来初始化及解算面元的局部 几何形状或表面法线的方法

技术领域

[0001] 本文中所揭示的标的物通常涉及用于例如计算机视觉及扩增实境 (AR) 应用程序等用于映射及跟踪技术的应用程序中的系统、方法及设备。具体来说,系统及过程用于通过比较可平行化架构中的关键帧的集合中的假设法线来计算表面法线。

背景技术

[0002] 相关背景

[0003] 例如游戏、计算机视觉及扩增实境系统等应用程序依赖于对象的准确及有效检测及跟踪。在未知环境中跟踪相机,或跟踪在相机前方移动的未知对象涉及首先建造环境或对象的外观及结构表示。随后可通过匹配来自传入图像的图像特征与可存储为关键帧的特征贴片的所有存储外观表示来继续进行跟踪。

[0004] 为了将来自过去关键帧(第一帧)的特征匹配到当前帧(第二帧)上,针对相似度,比较环绕过去关键帧(第一帧)中的特征点位置的局部外观或像素强度与其中预期特征所位于的当前帧(第二帧)中的候选区中的局部外观或像素强度。量化或测量相似度的一个常用方法为经由使用过去关键帧中的像素的窗口与当前关键帧中的像素的窗口之间的归一化交叉相关(NCC)得分。在两个帧/图像中的特征的校正位置处,来自两个帧/图像的像素的窗口之间的NCC得分很高,从而指示良好相似性。然而,在两个图像的相机姿态相当不同且存在两个图像中的特征的外观的改变时,无法直接比较特征窗口,且NCC得分很低。在此情况下,环绕第一帧中的特征位置的图像区域必须扭曲到第二帧的图像区域的视点,从而看起来类似于如从第二视点在第二帧中所见的特征的外观。环境的局部几何形状或局部表面法线对于此匹配来说是重要的,从而使外观从一个视图扭曲到另一视图,以用于从两个不同视角的宽基线匹配。场景的几何形状通常近似为中心定在3D特征点位置处且具有3D的任意平面法线向量的小平坦贴片。通过局部几何形状的此表示,需要中心定在每一3D特征点位置处的平坦贴片的法向量以便使围绕图像中的特征位置的窗口的外观从一个视点扭曲到另一视点。另外,对于以动态方式与环境交互的扩增,需要估计表面法线或贴片法线。举例来说,如果从其弹开的表面的法线为已知的,那么可使弹力球扩增为真实的;及在扩增实境的应用程序中,这些是中心定在3D特征点处的贴片的表面法线。传统地,使用最小化经扭曲贴片之间的平方差和的迭代梯度下降算法来估计这些法线。举例来说,Nicholas Moltom、Andrew Davison及Ian Reid在其标题为“来自运动的实时结构的局部平坦贴片特征(Locally Planar Patch Features for Real-Time Structure from Motion)”的论文中描述此估计,所述论文呈现于2004英国机器视觉会议。此方法为本身迭代算法,所述算法难以平行化。另外,此过程需要相对准确的初始解来收敛到全局最佳值。此需要计算梯度,这需要较高表示精度。归因于计算限制,对照明改变的最传统解也不是恒定的。

发明内容

[0005] 本发明的实施例是针对通过解算所俘获视频流中的对象的局部几何形状或表面法线来执行相机及目标跟踪。在一个实施例中,产生在视频流中可见的场景中的3D特征点的假设法线。假设法线表示相对于由移动相机所看见的对象的表面法线的近似。因为视频流从第一关键帧摇动到第二关键帧中所表示的对象的不同视图,所以针对特定3D特征点假设的所有法线当中的校正表面法线为具有以下特性的法线:在校正法线用于使中心定在第一关键帧中的图像特征位置周围的像素的窗口/贴片扭曲到第二关键帧的相机视图中时,来自第一帧的像素的经扭曲窗口具有与第二关键帧中的像素的对应窗口的高相关性。用于使像素从第一关键帧扭曲到第二关键帧的不正确地假设的法线将产生第一帧中的像素的经扭曲窗口与第二关键帧中的像素的窗口之间的低相关性得分。在一个实施例中,像素的窗口/贴片可为中心定在预定位置处的正方形或矩形像素网格或贴片(例如,8x8或10x12像素网格)。在一个实施例中,如果存在两个以上关键帧,那么校正法线可在此扭曲程序之后产生任何两个关键帧之间的高相关性得分。在准确表面法线的计算/估计之后,可正好匹配及跟踪视频流中的关键帧之间的对应对象或点。

[0006] 在一个实施例中,在可平行化架构中分析用于关键帧的集合的假设法线以确定其准确性,其中并行评估法线的集合以用于确定在扭曲之后在若干对关键帧之间的高相关性得分。假设可能表面法线及使用多对关键帧确定其准确性提供与使用梯度下降或类似迭代方法的替代方案相比较对图像噪声稳健的解。另外,能够以并行方式实施计算导致与迭代或类似方法相比较更快速的解。

[0007] 另一实施例是针对用于解算视频流中的对象的表面法线的方法,其包括:取样视频流以产生关键帧的集合;选择关键帧的集合中的对应图像点;使用两个或两个以上关键帧中的对应图像点产生由关键帧的集合成像的场景中的3D特征点;产生3D特征点的假设表面法线;通过从假设表面法线中的每一者计算的扭曲矩阵使对应于第一关键帧中的3D特征点的图像像素贴片扭曲到第二关键帧的视点;计分第一及第二关键帧之间的由每一假设表面法线产生的扭曲误差;及舍弃产生第一及第二关键帧之间的高扭曲误差的假设表面法线。

[0008] 另一实施例是针对包括指令的非暂时性机器可读媒体,所述指令在由机器执行时致使所述机器执行用于解算视频流中的对象的表面法线的操作,所述指令包括:取样视频流以产生关键帧的集合;选择关键帧的集合中的对应图像点;使用两个或两个以上关键帧中的对应图像点产生由关键帧的集合成像的场景中的3D特征点;产生3D特征点的集合的假设表面法线;通过从假设表面法线中的每一者计算的扭曲矩阵使对应于第一关键帧中的3D特征点的图像像素贴片扭曲到第二关键帧的视点;计分第一及第二关键帧之间的由每一假设表面法线产生的扭曲误差;及舍弃产生第一及第二关键帧之间的高扭曲误差的假设表面法线。

[0009] 另一实施例是针对用于解算视频流中的对象的表面法线的计算装置,其包括:用于取样视频流以产生关键帧的集合的取样装置;用于选择关键帧的集合中的对应图像点、使用两个或两个以上关键帧中的对应图像点产生由关键帧的集合成像的场景中的3D特征点及产生3D特征点的假设表面法线的假设装置;用于通过从假设表面法线中的每一者计算的扭曲矩阵使对应于第一关键帧中的3D特征点的像素图像贴片扭曲到第二关键帧的视点的扭曲装置;用于计分第一及第二关键帧之间的由每一假设表面法线产生的扭曲误差的计

分装置；及用于舍弃产生第一及第二关键帧之间的高扭曲误差的假设表面法线的消除装置。

[0010] 另一实施例是针对具有用于解算视频流中的对象的表面法线的处理器的计算装置，其包括：用于俘获视频流的相机；表面法线解算器，其包括：用于取样视频流以产生关键帧的集合的取样单元；用于选择关键帧的集合中的对应图像点、使用两个或两个以上关键帧中的对应图像点产生由关键帧的集合成像的场景中的3D特征点及产生3D特征点的假设表面法线的假设单元；用于通过从假设表面法线中的每一者计算的扭曲矩阵使对应于第一关键帧中的3D特征点的图像像素贴片扭曲到第二关键帧的视点的扭曲单元；用于计分第一及第二关键帧之间的由每一假设表面法线产生的扭曲误差的计分单元；及用于舍弃产生第一及第二关键帧之间的高扭曲误差的假设表面法线的消除单元。

[0011] 上文概述不包含本发明的所有方面的详尽列表。预期本发明包含可从上文概括的各种方面的所有合适组合以及在下文详细描述中所揭示及尤其在本申请案申请的权利要求书中所指出的那些方面实践的所有系统及方法。这些组合具有上文概述中并未特定叙述的特定优势。

附图说明

[0012] 本发明的特征、性质和优点在结合附图阅读时将从下文陈述的详细描述变得更清楚，在附图中相同参考字符始终对应地进行识别，且其中：

[0013] 图1展示可在未知环境中跟踪相机及/或可使用经估计表面法线跟踪在相机前方移动的未知对象的计算装置的组成图；

[0014] 图2展示视频流的第一关键帧中的书上的地图点的表面法线；

[0015] 图3展示检视平坦贴片的两相机；

[0016] 图4展示用于解算可平行化架构中的对象的表面法线的基于假设的过程；

[0017] 图5展示书4的具有假设法线的集合的实例关键帧；

[0018] 图6展示用于通过解算所俘获视频流中的对象的局部几何形状或表面法线来执行相机及目标跟踪的表面法线解算器；及

[0019] 图7展示针对图中的3D特征点产生的假设法线方向。

具体实施方式

[0020] 图1展示可在未知环境中跟踪相机及/或可使用经估计表面法线跟踪在相机前方移动的未知对象的计算装置100的组成图；计算装置100可为桌上型计算机、膝上型计算机、移动电话、蜂窝式电话、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)、平板或其它类似装置。

[0021] 计算装置100可包含存储器单元102及用于执行所俘获视频流中的相机及/或目标跟踪的操作的处理器104。在一个实施例中，计算装置100可包含集成到装置100中或经由有线或无线外围设备连接而连接到装置100的一或多个视频相机106a、b。视频相机106a、b可俘获将用以执行相机及/或目标跟踪的视频流，如下文将更完整地描述。相机106a、b可使用能够俘获视频流的任何类型的图像传感器，包含电荷耦合装置 (CCD) 有源像素传感器及互补金属氧化物半导体 (CMOS) 有源像素传感器。

[0022] 在一个实施例中,计算装置100可包含用于与一或多个有线或无线网络通信的网络接口单元110。网络接口单元110可提供计算装置100内的存取存储于远程装置/服务器上的数据的组件。举例来说,网络接口单元110可提供存取存储于远程服务器上的视频流的处理器104,所述视频流可经由因特网存取。网络接口单元110可能能够使用各种协议与远程装置/服务器通信。举例来说,网络接口单元110可经编程以与码分多址(CDMA)协议、时分多址(TDMA)协议、频分多址(FDMA)协议、3GPP长期演进(LTE)协议、传输控制协议及因特网协议(TCP/IP)、IEEE 802.X等中的一或多个者一起操作。举例来说,网络接口单元110可使用无线网络信道存取及检索远程服务器上的视频流。

[0023] 在一个实施例中,计算装置100中的表面法线解算器108有助于通过解算所俘获视频流中的对象的局部几何形状或表面法线来执行相机及目标跟踪。表面法线解算器108可处理视频流,且结合处理器104、存储器单元102、集成式数字信号处理器112及计算装置100的其它组件解算表面法线。在准确表面法线的计算/估计之后,可精确地匹配及跟踪视频流中的帧之间的对应对象或点。

[0024] 图2展示视频流的第一关键帧中的书204上的地图点(例如,3D特征点)的表面法线202。如本文中所使用,表面法线为正交于关键帧中的对象上的指定地图点的向量或平面。在一个实施例中,关键帧中的地图点的精确法线最初是未知的。然而,可产生地图点的假设法线。假设法线表示相对于关键帧中所见的对象的表面法线的近似。因为视频流摇动到第二关键帧中的对象(例如,书204)的不同视图,所以可通过使用假设表面法线的集合中的每一法线使图像/特征贴片从第一关键帧扭曲到第二关键帧中的图像/特征贴片及接着计算第二关键帧中的对应图像/特征贴片的相关性得分或扭曲误差(相关性得分及扭曲误差具有反比关系)来计分针对地图点产生的假设表面法线。如果作为使用特定假设表面法线的此扭曲程序(上文也描述)结果的两个帧的特征贴片之间的相关性得分低于设定的阈值(例如0.9或0.8),那么可舍弃所述特定假设法线。此不准确性可由多个关键帧中的相同地图点的贴片之间的在扭曲程序期间的高检测误差得分(即低归一化交叉相关性(NCC)得分)展示。准确表面法线对于扩增来说是重要的,从而以物理上合理的方式与表面交互。具体来说,对于一些扩增实境应用程序,应在大约5°到10°以内估计每一法线。然而,对于其它应用程序,准确性要求可能不同(更小或更大)。因此,舍弃不准确的假设表面法线会消除较晚处理程序中的潜在误差。

[0025] NCC匹配解决了找到两个图像之间的相似性的问题。经由使用NCC得分,对于均差及标准差,匹配是恒定的(线性效果)。对于图像数据,这表示对于亮度偏移及对比度改变,NCC是恒定的。在比较用不同相机或在当天的不同时间拍摄的图像时,NCC得分的这些特性会有更合乎需要的效果。

[0026] 在经由搜索窗口计算NCC得分时,考虑参考贴片以找到搜索窗口内的图像/特征贴片最佳匹配的位置。下式描述此操作:

$$[0027] \quad \gamma(u, v) = \frac{\sum_{x,y} [f(x, y) - \bar{f}_{u,v}] [t(x-u, y-v) - \bar{t}]}{\left\{ \sum_{x,y} [f(x, y) - \bar{f}_{u,v}]^2 \sum_{x,y} [t(x-u, y-v) - \bar{t}]^2 \right\}^{0.5}}$$

[0028] 其中

[0029] • x、y为贴片内的像素位置

- [0030] • u, v 为搜索窗口内的像素位置
- [0031] • $f()$ 存取贴片数据
- [0032] • $t()$ 存取搜索窗口内的图像数据
- [0033] • $\bar{f}_{u,v}$ (也被写成 f_{avg}) 为贴片中的所有像素的平均值
- [0034] • $\bar{t}$ (也被写成 t_{avg}) 为当前与贴片进行比较的区域的所有像素的平均值
- [0035] • $\sum_{x,y} [f(x,y) - \bar{f}_{u,v}]^2$ 为贴片的方差 (平方标准差)
- [0036] • $\sum_{x,y} [t(x-u, y-v) - \bar{t}]^2$ 为当前进行比较的区域的方差 (标准差)。

[0037] 在上文等式中, 一对最优像素位置 $\langle u, v \rangle$ 寻求获得最佳匹配。

[0038] Nicholas Molton、Andrew Davison 及 Ian Reid 的标题为“来自运动的实时结构的局部平坦贴片特征 (Locally Planar Patch Features for Real-Time Structure from Motion)”的论文中的章节 2.1 (其呈现于 2004 英国机器视觉会议) 描述用于执行贴片扭曲的方法。方法通过如图 3 中所示分析检视平坦贴片的两个相机的几何形状而开始。确切地说, 图 3 展示含有在具有零旋转的起点处的第一相机 (相机 0)、在从相机 0 旋转 R 的位置 t 处的第二相机 (相机 1) 及在位置 x_p (其在具有法线 n 的局部平坦表面上) 处的点的场景。相机 0 的图像平面中的在位置 (u_0, v_0) 处的点向后投影到光线 ρu_0 , 其中 $u_0 = (u_0, v_0, 1)^T$ 。在平面 p 上的点 x 满足等式 $n^T x = n^T x_p$ 。在光线与平面的相交处, $x = \rho u_0$, 因此 $\rho n^T u_0 = n^T x_p$, 且 $\rho = \frac{n^T x_p}{n^T u_0}$ 。相交

具有齐次坐标 $(\rho v_0, 1)^T = (n^T x_p u_0, n^T u_0)^T$ 。将此投影到第二图像中给出如下等式:

$$[0039] \quad u_1 = R \left[I | -t \right] \begin{pmatrix} n^T x_p u_0 \\ n^T u_0 \end{pmatrix} = R \left[n^T x_p I - t n^T \right] u_0$$

[0040] 其中 I 为 3×3 单位矩阵。因此, 两个相机图像中的理想化图像位置由单应性相关:

$$[0041] \quad H_1 = R \left[n^T x_p I - t n^T \right]$$

[0042] 如果相机具有内部矩阵 C_0 及 C_1 , 那么像素坐标中的单应性为:

$$[0043] \quad H = C_1 R \left[n^T x_p I - t n^T \right] C_0^{-1}$$

[0044] 此等式可用以预测其模板图像与当前相机图像之间的特征的视觉外观的变换。找到此预测与特征的实际外观之间的差异, 且使用图像配准方法进行表面法线估计的相关校正。此方法提供反向组合调整 (即, 用调整 H_i 更新原始图像变换 H_{n-1}), 从而得到变换 H_n 的新估计, 以使得:

$$[0045] \quad H_n = H_{n-1} H_i^{-1}$$

[0046] 可在表面法线 n 的改变方面参数化 H_i 。通过添加向量 d 来调整法线 n , 其中 $d = \alpha \bar{x}_p^1 + \beta \bar{x}_p^2$ 。向量 $\bar{x}_p^1$ 及 $\bar{x}_p^2$ 可经选择为垂直于彼此及 x_p 的任何单位向量, 且 α 及 β 参数化调整的量。相加调整 d 足以使表面法线在平面将在相机 0 中可见的定向的范围内变化。使用上文等式, 调整变换 H_i (即, 扭曲矩阵) 变为:

$$[0047] \quad H_i = H_n^{-1} H_{n-1}$$

$$H_i = \left[C_1 R \left[(n+d)^T x_p I - t(n+d)^T \right] C_0^{-1} \right]^{-1} \left[C_1 R \left[n^T x_p I - t n^T \right] C_0^{-1} \right]$$

$$[0048] \quad H_i = C_0 \left[n^T x_p I - t n^T + t d^T \right]^{-1} \left[n^T x_p I - t n^T \right] C_0^{-1}$$

$$H_i = C_0 \left[n^T x_p I - t n^T + t \left(\alpha \bar{x}_p^1 + \beta \bar{x}_p^2 \right)^T \right]^{-1} \left[n^T x_p I - t n^T \right] C_0^{-1}$$

[0049] 在一个实施例中,在可平行化架构中分析关键帧的集合的假设法线以确定其准确性,其中并行评估法线的集合以用于确定若干对关键帧之间的在扭曲之后的高相关性得分或低扭曲误差。假设可能表面法线及使用多对关键帧确定其准确性提供与使用梯度下降或类似迭代方法的替代方案相比较对图像噪声稳健的解。举例来说,此假设驱动系统及过程不会遭受取决于初始化而陷于局部最小值的问题。此外,梯度下降方法取决于图像梯度的准确估计,但在存在图像噪声的情况下,图像梯度估计变得不准确。因此,在存在图像像素噪声的情况下,假设表面法线计算更稳健。

[0050] 图4展示用于解算可平行化架构中的对象的表面法线的基于假设的过程300。过程300可由图2中所示的计算装置100的组件中的一或多者执行。举例来说,法线解算器108可结合处理器104、存储器单元102及计算装置100的其它组件执行下文描述的操作。

[0051] 过程300开始于操作302,其中接收或检索视频流。视频流可从相机106a、b中的一者获得,或从局部或远程存储器源获得。举例来说,视频流可存储于存储器单元102中及可由表面法线解算器108及处理器104存取。在另一情况下,视频流存储于远程服务器上且可由使用网络接口单元110由表面法线解算器108存取。视频流可使用任何译码格式或技术编码。举例来说,视频流可使用高级视频编解码器高清晰度(AVCHD)、音频视频交错(AVI)、Adobe Flash视频格式(FLY)、Google视频文件格式(GVI)、H.264及移动图片专家组(MPEG)系列标准编码。

[0052] 在操作302处进行接收或检索后,即刻在操作304处取样视频流以产生一系列关键帧。从视频流获得的关键帧的数目可取决于视频流的长度及指定的取样速率。举例来说,三十秒视频流可以5Hz的速率进行取样以输出150个关键帧。在其它实施例中,可使用不同取样速率。在操作304处通过取样产生的每一关键帧为在特定时刻在视频流中的场景的完整表示(例如,场景的静态图像)。选择取样速率以输出关键帧,所述关键帧显示视频流内的对象的显著差异或移动,且提供有用的基线。在一个实施例中,集成到计算装置100中的专用视频取样单元对视频流进行取样以产生关键帧的集合。

[0053] 在操作305处,在关键帧的集合中选择图像点,且图像点用以产生由关键帧的集合观测的场景中的地图点(例如,3D特征点)。图像点为对应于由关键帧观测的场景中的三维元素(即,3D特征点)的关键帧中的二维位置。

[0054] 在操作306处,针对地图点的集合产生假设法线。在一个实施例中,假设法线对应于每一关键帧中所表示的场景的单位球面中的不同离散方向。举例来说,可相对于x轴以45度引导在第一地图点处的第一假设法线,而可相对于x轴以90度引导在第二地图点处的第二假设法线。这些假设法线中的每一者表示由关键帧成像的对象的表面法线的估计。在一个实施例中,针对每一地图点产生假设法线的集合。举例来说,可产生60个单独地图点的60个假设表面法线。在其它情况下,可产生20到100个假设表面法线。举例来说,在一个实施例中,可产生30个单独地图点的30个假设表面法线,而在另一实施例中,可产生50个单独地图

点的50个假设表面法线。尽管这些假设表面法线为粗略的估计,但将经由比较及计分来确定其准确性,如下文所更详细描述。对于使具有大基线的视图之间的关键帧中的贴片的外观扭曲,估计及确定准确表面法线是重要的。因此,舍弃不准确的假设表面法线消除较晚处理程序中的潜在误差。

[0055] 在操作308处,选择两个关键帧以用于比较相应假设法线。选定关键帧可为视频流中的顺序/邻近帧或视频流中的非顺序/非邻近帧。在一个实施例中,选定关键帧经选择为时间上充分分离以确保由关键帧表示的场景中的对象的移动。举例来说,可设定下分离界限以指示关键帧之间的必要分离度。关键帧的分离提供可充分区分的足够宽基线以用于计分/比较假设/经估计法线。尽管关键帧经选择为相对分离,但在一个实施例中,分离不应超出上分离界限,以使得关键帧实质上并不相异。举例来说,在50个关键帧的连续集合中,下分离界限可为3个关键帧,而上分离界限可为5个关键帧。在此实例中,在操作308处选择的关键帧可为帧5及9(即,分开4个关键帧)。在其它实施例中,下分离界限可在1到5个关键帧的范围内,而上分离界限可在5到10个关键帧的范围内。尽管在关键帧方面表示上及下分离界限,但在其它实施例中,可在原始视频流中的关键帧之间的时间差方面表示分离界限。通过选择具有相对相似性及良好基线的关键帧,可较精确地执行扭曲及比较操作。

[0056] 在选择两个关键帧之后,操作310使用对应于图像像素贴片的地图点的假设表面法线中的每一者使所述贴片的外观从一个选定关键帧扭曲到另一选定关键帧。在一个实施例中,像素的窗口/贴片可为中心定在关键帧中所表示的图像中的预定位置处的正方形或矩形像素网格或贴片(例如,8x8或10x12像素网格)。可使用经由上文所描述的过程计算的扭曲矩阵执行扭曲。计分每一假设表面法线以确定准确性(即,误差度)。在一个实施例中,得分是基于来自第一关键帧的经扭曲贴片与第二关键帧中的目标贴片之间的NCC(归一化交叉相关性)得分,其中假设法线用以计算扭曲。产生高于预定义相关性得分水平的得分的假设法线经确定为准确假设,而低于预定义相关性得分水平的得分经确定为不准确的且被舍弃。舍弃低相关性得分假设会消除不太可能的解,同时保留更大可能的估计。在一个实施例中,可对用于多个关键帧及扭曲程序的扭曲误差求和来改进跨越多个关键帧的误差估计的稳健性。

[0057] 图5展示书404的具有假设法线402的集合的实例关键帧。经着色法线402a、b表示不准确的法线,其在操作310处计算的得分低于预定义的相关性得分水平,而未着色法线402c、d、e表示其得分高于预定义的相关性得分水平的准确法线。如上所指出,舍弃不准确的法线402a、b(例如,经着色表面法线),而保留准确法线402c、d、e以供相机及目标跟踪操作使用。

[0058] 在一个实施例中,对于多对关键帧重复操作308及310。举例来说,在关键帧5及9已进行比较及扭曲从而展现两个帧中的准确及不准确的法线之后,可以类似方式处理关键帧5以及关键帧2。因为重复比较帧,所以计算的数目归因于假设修剪随时间的流逝而减少(即,早点消除不太可能的假设而不用使用所有若干对帧对其进行测试)。

[0059] 通过使用多对关键帧,增加估计及表面法线解的稳健性。在足够数目的关键帧对已进行比较/扭曲且在操作312处确定呈现为准确的一或多个表面法线的清洁集合之后,操作308及310因法线估计已收敛而终止。其后,经修剪表面法线的最终集合可用以在操作314处执行相机跟踪、目标跟踪或其它视频处理程序。在一个实施例中,操作312可确定具有最

高相关性得分水平(即,最低扭曲误差)的表面法线。具有最高相关性得分的此经确定表面法线其后可在操作314处用以执行相机跟踪、目标跟踪或其它视频处理程序。

[0060] 上文所描述的用于解算表面法线的过程可并行化,因为表面假设法线为独立的,且可单独地执行对应比较/扭曲操作。可平行化及可独立地操作的此能力提供采用高速缓存位置的潜在硬件加速的可用性。

[0061] 在一个实施例中,使用定向直方图以用于初始化及优化每一关键帧中的假设法线。举例来说,对于关键帧中的每一地图点,针对每一关键帧计算观测此地图点的从地图点到相机的中心的方向。地图点的近似法线估计为每一地图点的这些向量的平均值。

[0062] 图6展示根据一个实施例的表面法线解算器108的功能单元框图及一些构成硬件组件。图6中所示的组件表示包含于表面法线解算器108中的元件,且不应解释为排除其它组件。表面法线解算器108的组件可用以执行用于解算上文所描述的对象表面法线的基于假设的过程300的操作中的一或多个。

[0063] 在一个实施例中,表面法线解算器108包含取样装置/单元502。取样装置502对视频流进行取样以产生一系列关键帧。从视频流获得的关键帧的数目可取决于视频流的长度及指定的取样速率。选择取样速率以输出显示视频流内的对象的显著差异或移动及提供有用基线的关键帧。在一个实施例中,取样装置502可为DSP、通用硬件处理器、滤波器的集合及/或任何类似装置。

[0064] 在一个实施例中,表面法线解算器108包含假设装置/单元504。假设装置504产生由取样装置502产生的关键帧中的每一者中的地图点的集合。在一个实施例中,假设法线对应于相对于每一关键帧中所表示的场景的单元球面的不同离散方向。举例来说,图7展示针对图中的3D特征点而产生的由从球面突出的线701表示的假设法线方向。假设装置504可为DSP、通用硬件处理器、滤波器的集合及/或任何类似装置。在一个实施例中,假设装置504可选择关键帧中的图像点。两个或两个以上关键帧中的对应图像点可用以使用多视图几何形状中的熟知技术产生由关键帧成像的场景中的地图点(例如,3D特征点)。举例来说, Richard Hartley和Andrew Zisserman的计算机视觉中书多视图几何形状描述可用以产生地图点(例如,3D特征点)的技术。图像点为对应于由关键帧成像的场景中的三维元素(即,3D特征点)的关键帧中的二维位置。

[0065] 在一个实施例中,表面法线解算器108包含选择装置/单元506。选择装置506从由取样装置502产生的关键帧的集合选择两个关键帧。在一个实施例中,选定关键帧经选择为时间上充分分离的,从而确保由关键帧表示的场景中的对象的移动。选择装置506可为DSP、通用硬件处理器、滤波器的集合及/或任何类似装置。

[0066] 在一个实施例中,表面法线解算器108包含扭曲装置/单元508。扭曲装置508使用对应于图像贴片的地图点的假设表面法线中的每一者使所述贴片的外观从一个选定关键帧扭曲到另一选定关键帧。扭曲装置508可为DSP、通用硬件处理器、滤波器的集合及/或任何类似装置。

[0067] 在一个实施例中,表面法线解算器108包含计分装置/单元510。计分装置510计分每一表面法线以确定准确性(即,误差度)。在一个实施例中,得分为来自第一关键帧的经扭曲贴片与第二关键帧中的目标贴片之间的NCC(归一化交叉相关性)得分,其中假设法线用以计算扭曲。在一个实施例中,计分装置510对多个关键帧的扭曲误差求和来改进跨越多个

关键帧的误差估计。计分装置510可为DSP、通用硬件处理器、滤波器的集合及/或任何类似装置。

[0068] 在一个实施例中,表面法线解算器108包含消除装置/单元512及标记装置/单元514。消除装置512舍弃/消除得分低于预定义相关性得分水平的假设法线,而标记装置514标记得分高于预定义相关性得分水平的假设法线。得分高于预定义相关性得分水平的假设法线经确定为准确假设,而低于预定义相关性得分水平的得分经确定为不准确的且被舍弃。舍弃低相关性得分假设消除不太可能的解,同时保留更大可能的估计。消除装置512及标记装置514可各自为DSP、通用硬件处理器、滤波器的集合及/或任何类似装置。

[0069] 在一个实施例中,表面法线解算器108包含确定装置/单元516。确定装置516确定经标记表面法线的集合是否大于或等于表面法线的预定义数目。在经标记表面法线的数目低于预定义数目时,确定装置516致使扭曲装置、计分装置及消除装置处理一对新关键帧。确定装置516可为DSP、通用硬件处理器、滤波器的集合及/或任何类似装置。在一个实施例中,确定装置516可确定具有最高相关性得分水平(即,最低扭曲误差)的表面法线。具有最高相关性得分的此经确定表面法线其后可用以执行相机跟踪、目标跟踪或其它视频处理程序。

[0070] 取决于应用,可由各种装置来实施本文中所描述的方法。举例来说,这些方法可以硬件、固件、软件或其任何组合来实施。对于涉及硬件的实施方案,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、经设计以执行本文所描述的功能的其它电子单元,或其组合内。

[0071] 对于涉及固件及/或软件的实施方案,可用执行本文中所描述的功能的模块(例如,程序、功能等等)来实施方法。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,软件代码可以存储在存储器中并且通过处理单元执行。存储器可以实施在处理单元内或处理单元外部。在本文中使用时,术语“存储器”是指任何类型的长期、短期、易失性、非易失性或其它存储器,并且不应限于任何特定类型的存储器或任何特定类型数目的存储器或存储存储器的媒体的类型。

[0072] 如果以固件及/或软件来实施,那么可将功能作为一或多个指令或代码存储在计算机可读媒体上。实例包含编码有数据结构的计算机可读媒体和编码有计算机程序的计算机可读媒体。计算机可读媒体包含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含RAM、ROM、EEPROM、CD-ROM或其它光盘存储装置,磁盘存储装置、半导体存储装置或其它存储装置,或任何其它可用于存储指令或数据结构的形式所要程序代码并且可由计算机存取的媒体;在本文中使用,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),软性磁盘和蓝光光盘,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。上述各项的组合也应包含在计算机可读媒体的范围内。

[0073] 除了存储在计算机可读媒体上之外,还可将指令和/或数据提供为通信设备中包含的传输媒体上的信号。举例来说,通信设备可包含具有指示指令及数据的信号的收发器。所述指令和数据经配置以致使一或多个处理单元实施权利要求书中概述的功能。即,通信设备包含具有指示执行所揭示的功能的信息的信号的传输媒体。在第一时间,通信设备中

包含的传输媒体可包含用以执行所揭示功能的信息的第一部分,而在第二时间,通信设备中包含的传输媒体可包含用以执行所揭示功能的信息的第二部分。

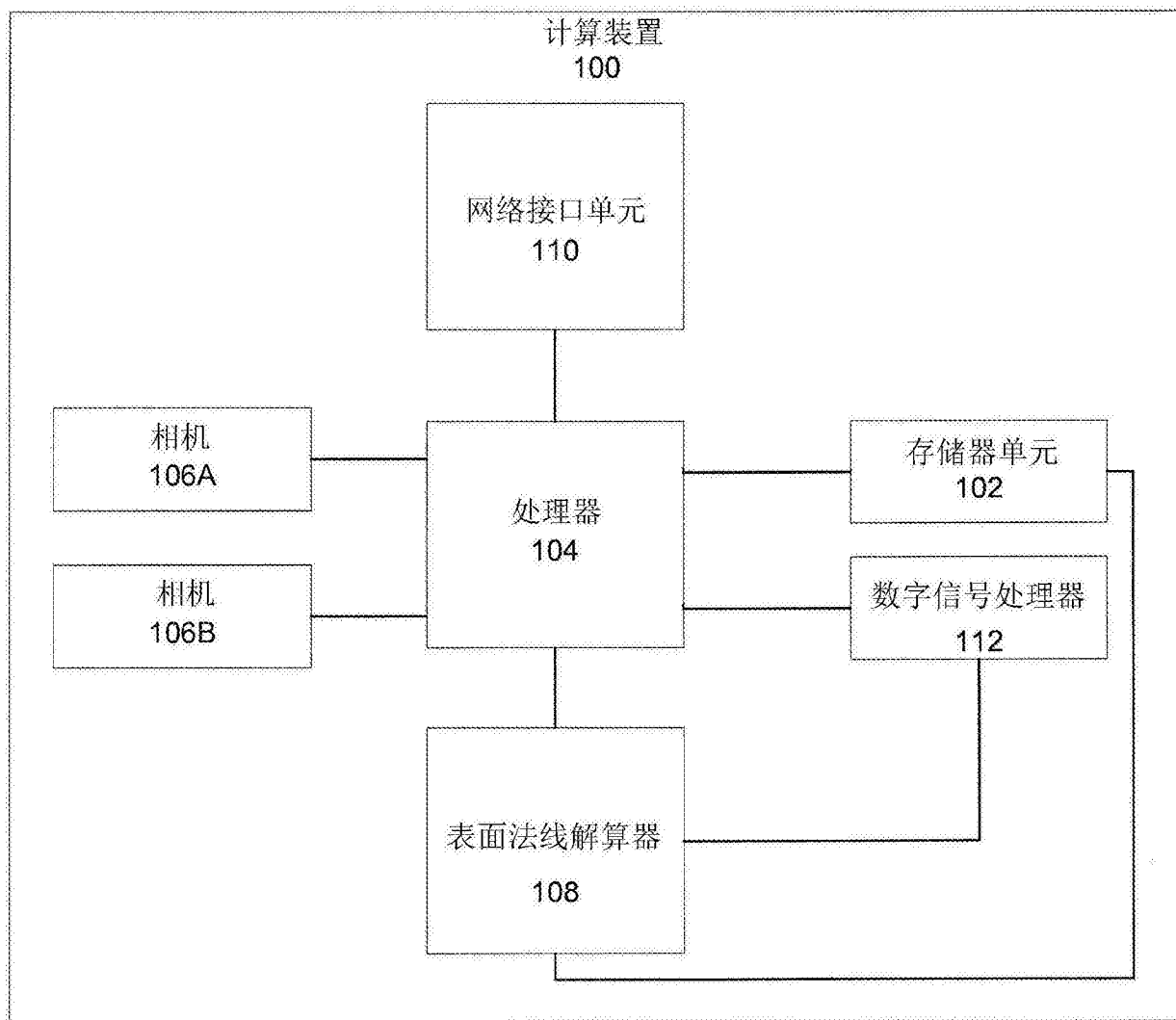


图1

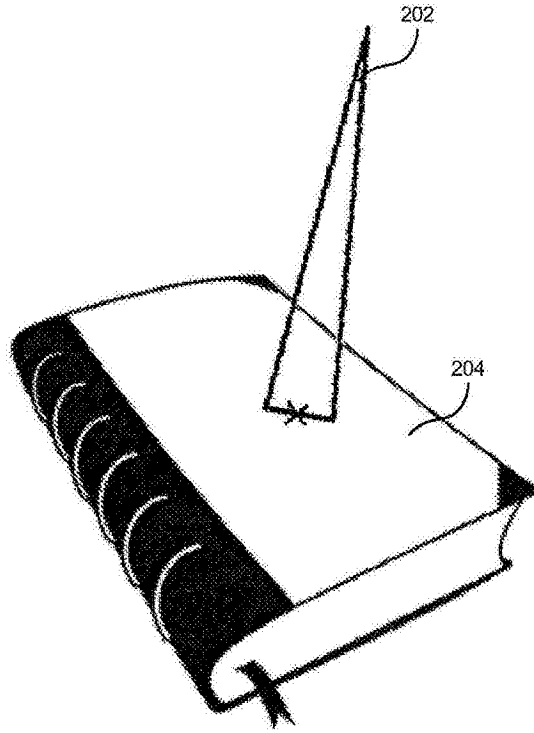


图2

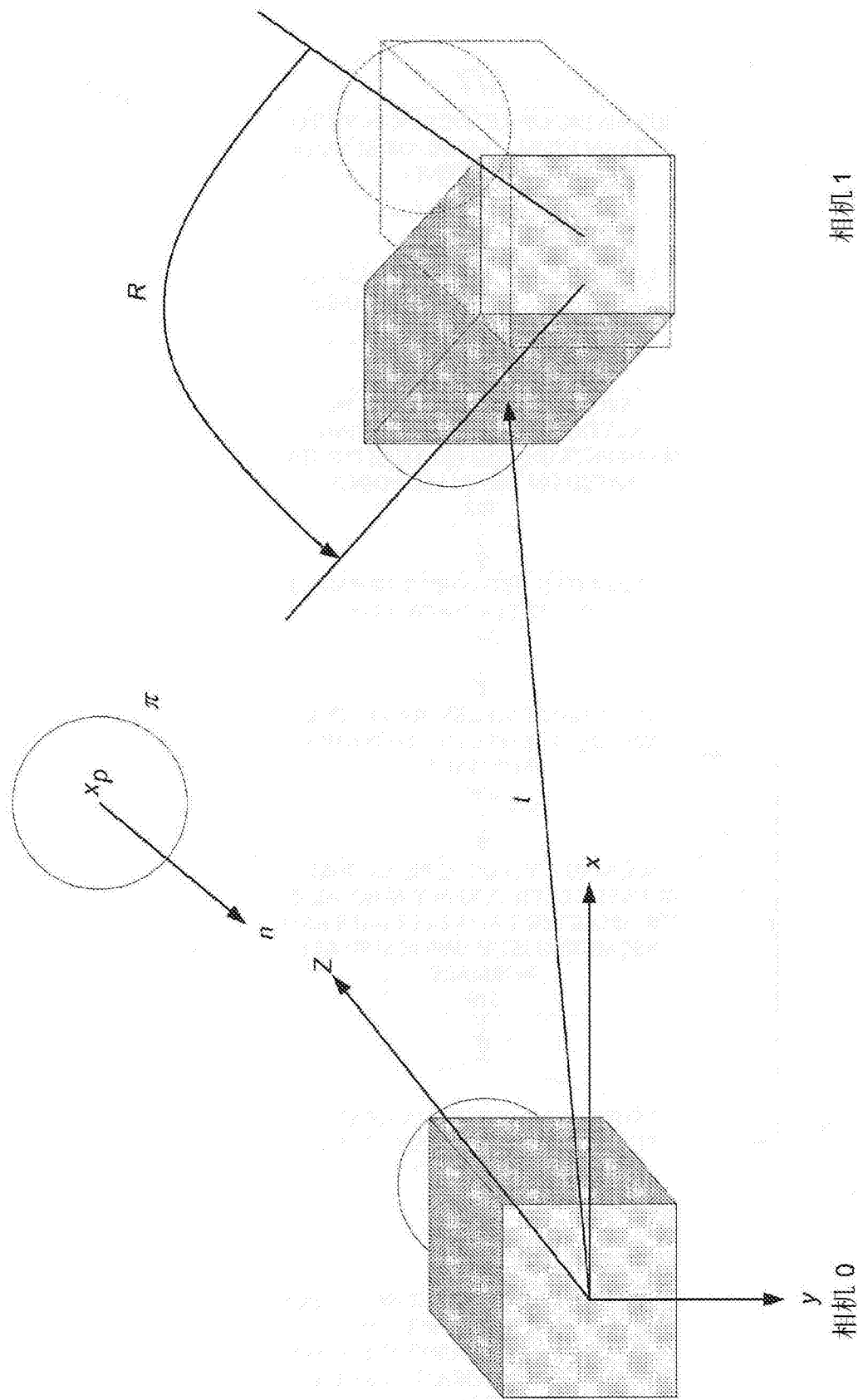


图3

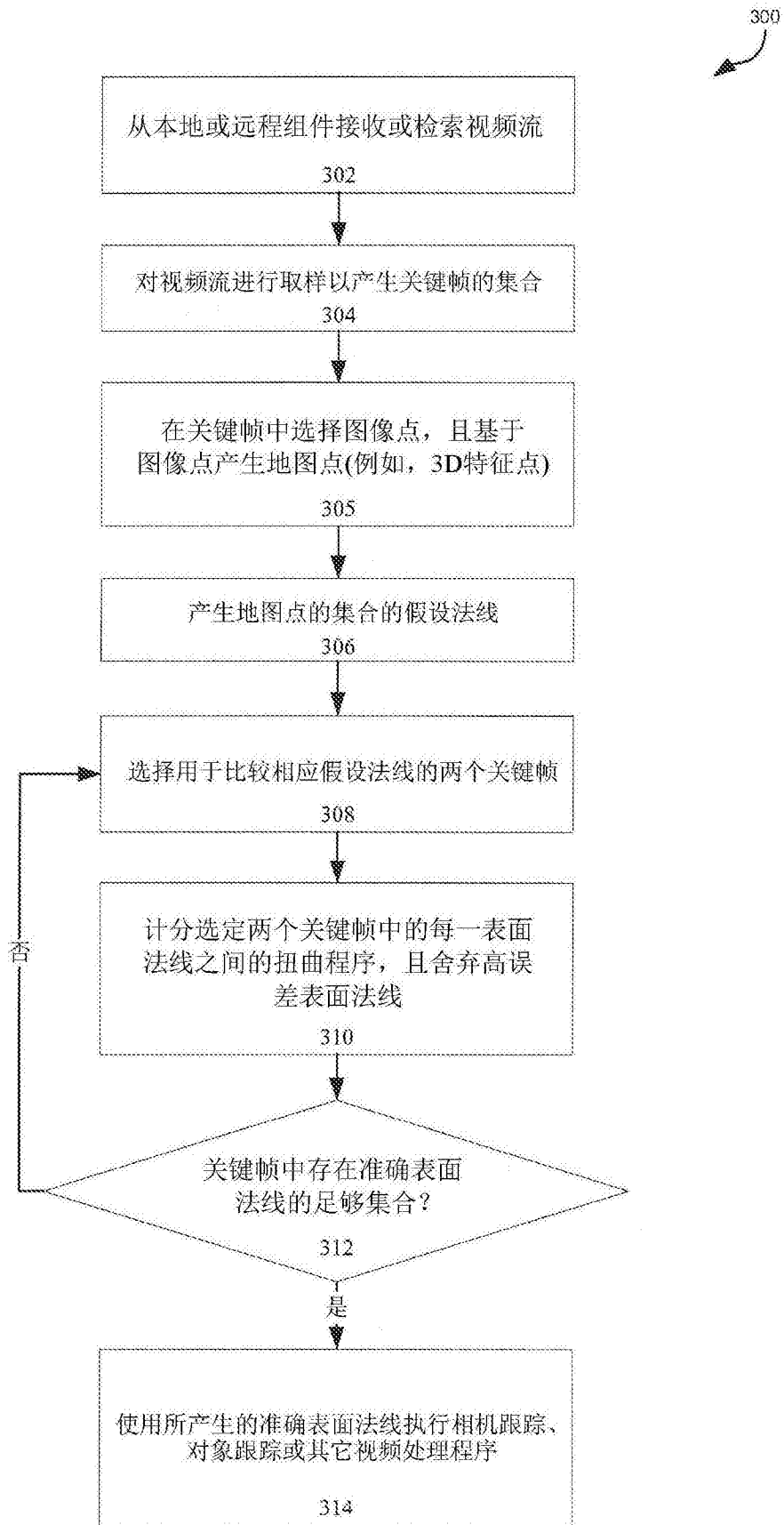


图4

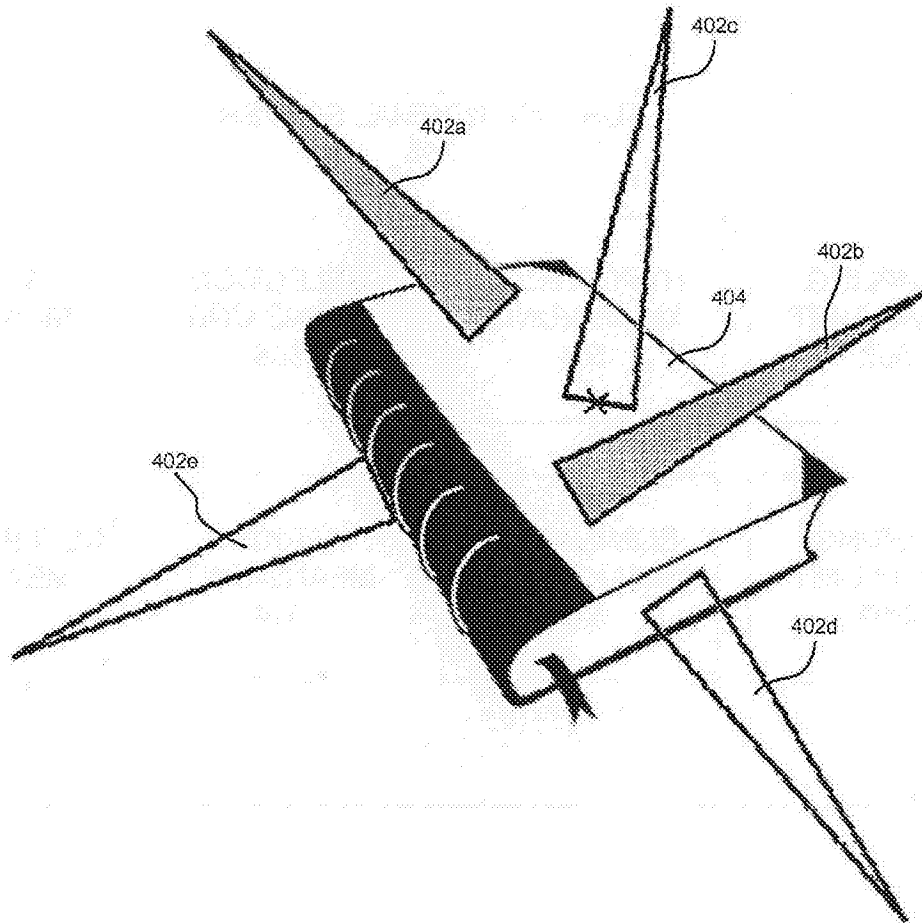


图5



图6

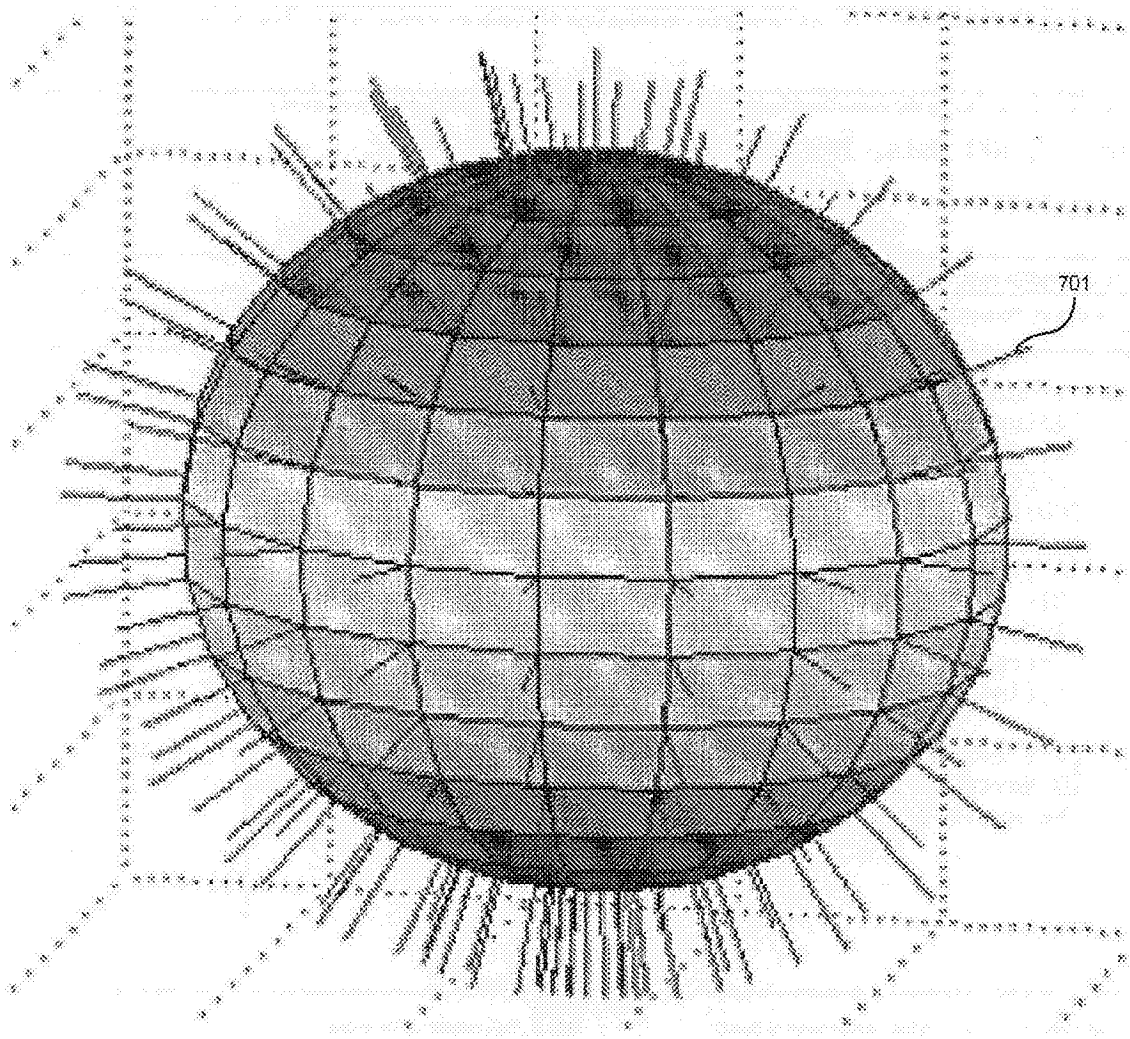


图7