



(12)发明专利

(10)授权公告号 CN 104322052 B

(45)授权公告日 2018.03.20

(21)申请号 201380024596.2
 (22)申请日 2013.05.09
 (65)同一申请的已公布的文献号
 申请公布号 CN 104322052 A
 (43)申请公布日 2015.01.28
 (30)优先权数据
 1208088.3 2012.05.09 GB
 (85)PCT国际申请进入国家阶段日
 2014.11.10
 (86)PCT国际申请的申请数据
 PCT/GB2013/051205 2013.05.09
 (87)PCT国际申请的公布数据
 W02013/167901 EN 2013.11.14
 (73)专利权人 恩卡姆技术有限公司
 地址 英国伦敦市卡莱尔大街
 (72)发明人 塞缪尔·布瓦万 布赖斯·米休德

(74)专利代理机构 北京连和连知识产权代理有限公司 11278

代理人 杨帆

(51)Int.Cl.
 H04N 5/222(2006.01)
 H04N 5/265(2006.01)
 H04N 5/272(2006.01)
 H04N 13/156(2018.01)
 H04N 13/204(2018.01)
 H04N 13/275(2018.01)
 H04N 13/239(2018.01)
 G06T 15/20(2011.01)
 G06T 19/00(2011.01)

(56)对比文件
 CN 101383910 A,2009.03.11,
 CN 1604014 A,2005.04.06,
 US 2008158345 A1,2008.07.03,
 US 2002139920 A1,2002.10.03,

审查员 侯瑜

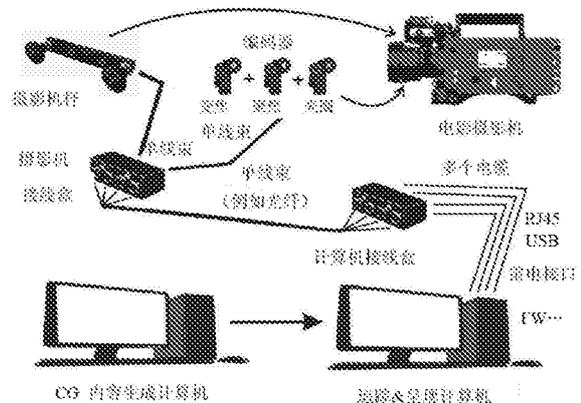
权利要求书3页 说明书28页 附图8页

(54)发明名称

用于实时混合或复合计算机生成的三维对象和电影摄影机馈送视频的系统

(57)摘要

一种实时混合或复合计算机生成的三维对象和电影摄影机馈送视频的方法,其中电影摄影机的机身可以在三维移动,且在摄影机中或连接至摄影机的传感器提供限定摄影机的三维位置和三维方向或使三维位置能够被计算出的实时定位数据。



1. 一种无标记系统,其用于实时混合或复合计算机生成的三维对象和来自摄影机馈送的视频,为电视广播、电影或视频游戏实时产生增强现实视频,其特征在于,其中:

(a) 摄影机的机身可以在三维移动,且在摄影机中或直接或间接连接到摄影机的传感器提供确定摄影机的三维位置和三维方向或使摄影机的三维位置和三维方向能被计算出的实时定位数据,该传感器包括从加速计、陀螺仪或磁力计中选出的非光学传感器;

(b) 一对立体见证摄影机直接或间接固定到摄影机上;

(c) 所述实时定位数据由系统自动使用以创建、调用、呈现或修改计算机生成的三维对象;

(d) 产生的计算机生成的三维对象与来自摄影机的馈送视频混合或复合以提供电视广播、电影或视频游戏的增强现实视频;

以及其中:

(e) 摄影机的三维位置和方向是以参照产生的真实世界的三维地图并至少部分通过使用来自非光学传感器的实时三维定位数据加上光流来进行确定,其中,所述见证摄影机测量场景,在处理器上运行的软件在场景中检测不是手动或人工添加到场景中的自然标记。

2. 根据权利要求1所述的系统,其特征在于,所述传感器包括三维距离传感器,其获取来自摄影机的视频输出中每个像素的深度。

3. 根据权利要求2所述的系统,其特征在于,边缘深度通过在摄影机的高分辨率馈送视频上重新投影三维范围传感器深度来进行改善。

4. 根据权利要求1所述的系统,其特征在于,所述一对立体见证摄影机高速见证摄影机使该系统能够完全初始化,而没有仅仅测量要追踪场景的独立阶段,恰恰相反,在摄影机用于拍摄视频的同时测量连续发生。

5. 根据权利要求4所述的系统,其特征在于,所述高速见证摄影机的速度至少为100fps。

6. 根据权利要求4所述的系统,其特征在于,其中两台高速见证摄影机组成立体系统,其使软件能够处理图像且能够在即使一点也不移动摄影机系统的情况下产生即时三维点云。

7. 根据权利要求6所述的系统,其特征在于,系统可通过使用两个摄影机之间的间隔和对极几何的认知将场景中的大量点与它们在三维空间的位置结合。

8. 根据权利要求4所述的系统,其特征在于,其中三维点云中每个像素的深度使用从每个立体见证摄影机获得的相应二维纹理补丁和极线线性搜索算法得到。

9. 根据权利要求1所述的系统,其特征在于,所述系统运行结合了来自所述立体见证摄影机系统的光流数据与来自硬件传感器的实时定位数据的融合算法。

10. 根据权利要求9所述的系统,其特征在于,其中所述融合算法是根据扩展卡尔曼滤波预测/校正技术整合来自所有的传感器的输出以及重新校准所有的传感器,以确定摄影机的位置和方向,其中所有的传感器可以包括加速计、陀螺仪、磁力计、三维距离传感器。

11. 根据权利要求10所述的系统,其特征在于,其中当确定如何合并每个传感器的数据时,扩展卡尔曼滤波融合算法使用与每个传感器输出有关的置信度数据。

12. 根据权利要求1所述的系统,其特征在于,其中所述立体见证摄影机系统产生的关键帧是视觉追踪过程的一部分且是在立体见证摄影机馈送视频的4个不同分辨率水平计算

的实时图像。

13. 根据权利要求1所述的系统,其特征在于,其中三维距离传感器用于提高与重新构造的三维点有关的深度测量的精度,或拒绝重新构造的三维点。

14. 根据权利要求13所述的系统,其特征在于,其中三维距离传感器用于实时深度键控以实现动态遮挡并抑制绿色阶段的最终使用。

15. 根据权利要求1所述的系统,其特征在于,其中软件尝试在三维映射中生成均匀分布的点云以大大减小追踪损耗且增加追踪精度,其中产生了更多的视差,因此估计的摄影机位置更精确。

16. 根据权利要求1所述的系统,其特征在于,所述系统提供了摄影机追踪系统,当导演/摄影师追踪、摇摄、倾斜连接到摄影机的摄影机追踪系统时,所述摄影机追踪系统结合三维场景的实时测量或即时测量与摄影机的追踪。

17. 根据权利要求1所述的系统,其特征在于,所述系统容许场景的实时连续测量以产生限定场景的更完整的点云。

18. 根据权利要求1所述的系统,其特征在于,所述系统使用附加旋转不变性描述符至场景中检测到的特征点以方便追踪的恢复。

19. 根据权利要求1所述的系统,其特征在于,所述系统使用与传感器提供的信息有关的恒速模型,以通过之前正确计算的或确认的位置预测摄影机的下一个位置。

20. 根据权利要求19所述的系统,其特征在于,所述系统使用所述预测在当前画面上重新投影三维点云,使点匹配算法能够匹配在立体见证摄影机系统的实时馈送视频中识别的点和在新的三维点云中投影的点。

21. 根据权利要求20所述的系统,其特征在于,所述系统为摄影机追踪使用莱文贝格-马夸特最小化方案以使在立体见证摄影机系统的实时馈送视频中识别的点和在新的三维点云中投影的点之间的误差减到最小。

22. 根据权利要求1所述的系统,其特征在于,其中用户可以使用摄影机追踪系统生成的三维点云定义三维蔽光框。

23. 根据权利要求1所述的系统,其特征在于,其中摄影机和立体见证摄影机使用调制光源对画面采集延迟进行校准。

24. 根据权利要求1所述的系统,其特征在于,其中摄影机是以下的任何一种:升降台摄影机、摄影机稳定器、手持式摄影机、移动摄影车式摄影机、三脚架摄影机、智能手机、增强现实眼镜。

25. 一种无标记方法,其用于实时混合或复合计算机生成的三维对象和摄影机的馈送视频,为电视广播、电影或视频游戏产生增强现实视频,其特征在于,其中:

(a) 摄影机的机身可以在三维移动,且在摄影机中或直接或间接连接到摄影机的传感器提供确定摄影机的三维位置和三维方向或使摄影机的三维位置和三维方向能够被计算出的实时定位数据,该传感器包括从加速计、陀螺仪或磁力计中选出的非光学传感器;

(b) 一对立体见证摄影机直接或间接固定到摄影机上;

(c) 所述实时定位数据由系统自动使用以创建、调用、显示或修改计算机生成的三维对象;

(d) 产生的计算机生成的三维对象与摄影机的馈送视频混合或复合以为电视广播、电

影或视频游戏提供增强现实视频；

以及其中：

(e) 摄影机的三维位置和方向是以参照产生的真实世界的三维地图并至少部分通过使用来自非光学传感器的实时三维定位数据加上光流来进行确定，其中，所述见证摄影机测量场景，在处理器上运行的软件在场景中检测不是手动或人工添加到场景中的自然标记。

26. 根据权利要求25所述的方法，其特征在于，其中记录实时定位数据并标出时间码，以为后期制作过程提供匹配移动数据。

27. 根据上述权利要求26所述的方法，其特征在于，所述方法容许下述的一个或多个：

- 摄影棚用摄影机的实时追踪
- 摄影机稳定器的实时追踪
- 升降台摄影机的实时追踪
- 移动摄影车式摄影机的实时追踪
- 摄影机稳定器的实时追踪
- 外部广播OB的实时追踪
- 在智能手机上实施的实时追踪
- 在穿戴式计算设备上实施的实时追踪
- 将实时数据用于二维后期制作
- 将实时数据用于三维立体内容的后期转换
- 将实时数据用于本地三维立体内容
- 三维图形插入
- 用于摄影棚内或现场的广告置入的三维图形插入
- OB的三维图形插入
- 其他赞助的图像的三维图形插入
- 特定观众位置的三维图形插入
- 特定观众的三维图形插入
- 特定时间的三维图形插入
- 填充人群场景的三维图形插入
- 绿屏替换的三维图形插入
- 在博物馆和文化、历史遗址和自然遗产解说中心辅助学习的教育内容的三维图形插入
- 场景中对象的绝对或相对尺寸的测量。

用于实时混合或复合计算机生成的三维对象和电影摄影机馈送视频的系统

[0001] 发明背景

1. 技术领域

[0002] 一种用于实时混合或复合计算机生成的3D对象和电影摄影机(例如摄像机)的馈送视频以为电视广播、电影或视频游戏产生实时增强现实视频的系统。

2. 背景技术

[0003] 在过去的20年中,在这一领域出现了大量的研究和商业活动;可以参考来自以下公司的视频插入或增强现实系统,例如运动视觉有限公司(Sportvision, Inc),该公司已开发出用于体育赛事(例如美式足球)的电视观看增强,增加了观众看到叠加在球场上的虚拟的第一次进攻线。其他已在这一领域开发出系统的公司包括PVI有限公司。已知系统的一个共同特征是它们主要依靠分析来自摄影机的馈送视频的特征以确定摄影机指向真实世界场景中的哪一部分;该系统早期已建立那个场景的三维映射,以便一旦它知道它指向场景中的哪个位置,它能够在馈送视频中以如下方式添加或复合计算机产生的对象(如虚拟的第一次进攻线),通过对象的位置和方向使它看起来好像场景的自然部分。这种方式中仅仅依赖光流的一个缺点是这种系统可能会不可靠。

[0004] 其它系统依靠以纯标记为基础的方法(例如光动力飞行器技术公司(Lightcraft Technologies))。他们需要操作者把真实的物理标记(大小为1米x1米)放在布景上由系统进行检测。它非常低效,因为它需要几小时或几天来建立舞台,这对于影视制作来说是不可能完成的。它也具有许多限制,因为物理标记必须总是保持在他们系统的视野中。也可以参考附录1中引用的论文。

发明内容

[0005] 本发明的一个方面是一种用于实时混合或复合计算机生成的三维对象和电影摄影机(例如摄影机)馈送的视频,以为电视广播、电影或视频游戏产生增强现实视频的系统,其中:

[0006] (a) 电影摄影机的机身可以在三维移动,且在电影摄影机中或直接或间接连到电影摄影机的传感器提供限定电影摄影机的三维位置和三维方向或使电影摄影机的三维位置和三维方向能够计算出的实时定位数据,以及

[0007] (b) 然后,系统自动使用实时定位数据以创建、调用、呈现或修改计算机产生的三维对象,以及

[0008] (c) 然后,混合或复合产生的计算机生成的三维对象和电影摄影机馈送的视频以为电视广播、电影或视频游戏提供增强现实视频。

[0009] 可选特征—一些或全部特征可以彼此结合,包括以下:

[0010] • 实时混合或复合计算机生成的三维对象和电影摄影机的实时馈送视频。

- [0011] • 测试和使用电影摄影机的实时变焦、聚焦以及光圈设置,连同实时定位数据,以便在三维场景的预定位置和方向正确地呈现三维对象。
- [0012] • 传感器包括加速计和陀螺仪(六自由度传感器(6DOF sensor))
- [0013] • 传感器包括测量三维平移加速度的三轴加速计、测量三维角速度的三轴陀螺仪、以及测量三维绝对朝向的磁力计,因此组成了九自由度传感器。
- [0014] • 传感器包括三维距离传感器,例如结构光或飞行时间摄影机。
- [0015] • 三维距离传感器获取摄影机的视频输出中每个像素的深度(depth)。
- [0016] • 通过在电影摄影机的高分辨率馈送视频上重新投影三维距离传感器深度来完善边缘深度。
- [0017] • 传感器形成可以完全固定到电影摄影机上的单元
- [0018] ○该单元包括一个或两个见证摄影机(witness camera)。
- [0019] ○该单元包括三维距离传感器,其获得视频输出中每个像素的深度。
- [0020] ○该单元包括六自由度或九自由度传感器。
- [0021] ○该单元使用可切换的悬挂器(hanger)以允许它被固定到不同尺寸和设计的电影摄影机和杆上。
- [0022] ○该单元可扩展为包括其它形式的传感器。
- [0023] ○该单元可以组成测量设备,其可以用于测量复杂的场景并无线传输限定测量的三维场景的数据至计算机,该计算机然后可以追踪或恢复场景的追踪。
- [0024] ○电影摄影机包括传统编码器以读取摄影机变焦、聚焦和光圈。
- [0025] • 系统包括一个单一的见证摄影机(单视场的情况)或两个见证摄影机(立体视场的情况),配备有镜头,可以是180度鱼镜头。
- [0026] • 见证摄影机偏置于电影摄影机,并且偏移值使用包括第一和第二圆圈组的校准表得到,每个圆圈处于相对一个或多个其他圆圈已知的位置且每个圆圈使用二进制大对象图像算法(blob image algorithm)进行识别。
- [0027] • 使用包括若干圆圈的校准表校准电影摄影机镜头的光学畸变,每个圆圈处于相对一个或多个其他圆圈已知的位置且每个圆圈使用二进制大对象图像算法进行识别。
- [0028] • 电影摄影机是以下的任何一种:升降架摄影机、摄影机稳定器、手持式摄影机、移动摄影车式摄影机、三脚架摄影机、智能手机、增强现实眼镜。
- [0029] • 电影摄影机的三维位置和方向参考部分通过使用来自传感器的实时三维定位数据加上光流生成的真实世界的三维映射确定,其中见证摄影机测量场景且在处理器上运行的软件检测场景中不是手动或人为增加到场景中的自然标记(“特征点”)。
- [0030] • 该系统使用一个或两个高速(例如至少100fps)见证摄影机使该系统能够完全初始化,没有仅仅测量追踪场景的独立阶段(“即时测量”),恰恰相反,在摄影机用于拍摄视频的同时测量连续发生。
- [0031] • 立体视场系统启用软件处理图像且甚至在一点也不用移动摄影机系统的情况下产生即时三维点云(例如使用两个摄影机之间的间隔和对极几何的认知将场景中的大量点与它们在三维空间的位置结合)。
- [0032] • 三维点云中每个像素的深度使用从每个立体见证摄影机获得的相应的二维纹理补丁和极线线性搜索算法(epi-polar line search algorithm)得到。

- [0033] • 系统运行结合了来自见证摄影机系统的光流数据和来自硬件传感器的实时定位数据的融合算法。
- [0034] • 融合算法是根据扩展卡尔曼滤波 (Extended Kalman Filter, EKF) 预测/校正技术整合所有的传感器输出, 以及重新校准所有的传感器, 以确定摄影机的位置和方向, 其中所有的传感器可以包括加速计、陀螺仪、磁力计、三维距离传感器。
- [0035] • 当确定如何合并来自每个传感器的数据时, EKF 融合算法使用与从每个传感器的输出有关的置信度数据。
- [0036] • 见证摄影机系统产生的关键帧是视觉追踪过程的一部分且是在见证摄影机馈送视频的四个不同分辨率水平计算的实时图像。
- [0037] • 系统包括 (a) 内容生成计算机, 其提供三维计算机生成的虚拟图形、对象和地点的动画, 以及 (b) 呈现计算机 (其可以独立于或不独立于内容生成计算机), 其中限定电影摄影机的三维位置的实时定位数据由内容生成计算机和呈现计算机中的二者之一或二者使用以实时引起要生成的计算机生成的三维对象, 该实时计算机生成的三维对象可以实时插入和混合到电影摄影机的馈送视频中以形成馈送视频中所示的场景的自然部分。
- [0038] • 计算机生成的三维对象是动画, 其可以在场景中的任何地方移动且可以以内容生成计算机确定的方式改变他们的形状和外观。
- [0039] • 计算机生成的三维对象是人或生物的动画图形, 当混合在场景中时其以逼真的方式移动 (如跑、跳舞、步行、战斗、飞行、跳跃.....)。
- [0040] • 摄影机定位或追踪数据也可用于后期制作以方便后期制作 CGI (计算机图形影像)。
- [0041] • 三维距离传感器用于提高与重新构造的三维点有关的深度测量的精度, 或拒绝重新构造的三维点。
- [0042] • 三维距离传感器用于实时深度键控以实现动态遮挡 (dynamic occlusion) 并抑制绿色阶段 (green stage) 的最终使用。
- [0043] • 系统使用了小型摄影机注册对象 (camera registration object), 例如已知尺寸和覆盖有已知图案放置在场景中的板, 以便检测到的图案的边缘地方被作为三维点云 (以及因此该世界) 的起点。
- [0044] • 摄影机注册对象包含至少两个已知尺寸、真垂直排布且使用二进制大对象图像识别算法识别的球。
- [0045] • 系统包括表明磁北的磁力计、表明重力方向 (以及因此给出真垂线) 的加速计、表明系统是否仰摄/俯摄或向左方或右方摇摄或围绕光轴旋转的陀螺仪、实现对从初始位置的三维平移的推断的三轴加速计。
- [0046] • 软件尝试在三维映射中生成均匀分布的点云以大大减小追踪损失且增加追踪精度 (产生的视差越多, 因此估计的摄影机位置越精确)。
- [0047] • 摄影机追踪系统可以无线连接到摄影机系统且当产生三维点云时能够因此在布景周围快速移动, 不需要在布景中拖拽电缆, 不像早期的见证摄影机系统。
- [0048] • 当导演/摄影师追踪、摇摄、倾斜连接到电影摄影机的摄影机追踪系统时, 摄影机追踪系统使三维场景的实时测量 (单视场的情况) 或即时测量 (立体显示的情况) 与电影摄影机的追踪结合。

- [0049] • 系统使电影摄影机追踪的所有方面完全自动化,包括旋转、平移、聚焦、光圈、焦距;且使混合到视频中的三维计算机生成的内容的缩放、定位和方向自动化。
- [0050] • 系统实现对场景的实时连续测量以产生限定场景的更完整的点云。
- [0051] • 系统附加旋转不变性描述符(rotation invariant descriptor),例如使用ORB(对象请求代理),至场景中检测到的特征点以促进追踪的恢复。
- [0052] • 系统使用与传感器提供的信息有关的恒速模型,使用之前正确计算的或确认的位置预测电影摄影机的下一个位置。它使用该预测在当前画面上重新投影三维点云,使点匹配算法能够匹配见证摄影机系统在实时馈送视频中识别的点和在新的三维点云中投影的点。
- [0053] • 系统为摄影机追踪使用莱文贝格-马夸特最小化方案(Levenberg-Marquardt minimization scheme)以使在实时馈送视频中见证摄影机系统识别的点和在新的三维点云中投影的点之间的误差减到最小。
- [0054] • 用户可以使用摄影机追踪系统产生的三维点云限定三维蔽光框(mask),例如三维抠像蔽光框(3D Garbage Matting mask)。
- [0055] • 三维对象包括静态对象、动态动画、虚拟世界、虚拟人物、虚拟建筑、虚拟景观、虚拟电影布景以及动画数据库中的任何数据。
- [0056] • 电影摄影机和见证摄影机使用调制光源对画面采集延时进行了校准,例如通过比较与闪光灯(flash LED)有关的光强度曲线。
- [0057] 其他概念—每个可以和上述定义的特征的任何一个或以下定义的任何其他概念结合。
- [0058] 一种实时混合或复合计算机生成的三维对象和电影摄影机的馈送视频的方法。其中电影摄影机的机身可以在三维移动且在摄影机中或连接至摄影机的传感器提供限定摄影机的三维位置和三维方向或使三维位置能够被计算出的实时定位数据。
- [0059] 一种用于实时混合或复合计算机生成的三维对象和电影摄影机(例如摄影机)的馈送视频以为电视广播、电影或视频游戏产生增强现实视频的方法,其中:
- [0060] (a) 电影摄影机的机身可以在三维移动且在电影摄影机中或直接或间接连接到电影摄影机的传感器提供限定电影摄影机的三维位置和三维方向或使电影摄影机的三维位置和三维方向能够被计算出的实时定位数据,以及
- [0061] (b) 然后,实时定位数据由系统自动使用以创建、调用、渲染或修改计算机生成的三维对象以及
- [0062] (c) 然后,产生的计算机生成的三维对象与电影摄影机的馈送视频混合或复合以为电视广播、电影或视频游戏提供增强现实视频。
- [0063] 上述定义和使用的方法或系统使导演(或摄影导演)能够实时将摄影机中前期制作的计算机生成的三维资源(通常前期可视化(previsualisation)或视觉效果资源)加入到摄影机拍摄的实景电影底片或视频图像上,使导演能够探索可能的摄影机视角且随自动混合到导演看得见的视频中的计算机生成的三维资源实时移动。
- [0064] 上述定义的记录实时定位数据并标出时间码的方法或系统为后期制作过程提供了匹配移动数据。
- [0065] 上述定义和使用的方法或系统使虚拟对象或图像能够插入到播放视频流中。

- [0066] 上述定义的方法或系统实现了以下的一个或多个：
- [0067] • 摄影棚用摄影机的实时追踪
 - [0068] • 摄影机稳定器 (Steadicam) 的实时追踪
 - [0069] • 升降架摄影机的实时追踪
 - [0070] • 移动摄影车摄影机的实时追踪
 - [0071] • 摄影机稳定器的实时追踪
 - [0072] • 外部广播 (OB) 的实时追踪
 - [0073] • 将实时数据 (如追踪数据) 用于二维后期制作
 - [0074] • 将实时数据 (如追踪数据) 用于三维立体内容的后期转换
 - [0075] • 将实时数据 (如追踪数据) 用于本地三维立体内容
 - [0076] • 三维图像插入
 - [0077] • 摄影棚或现场广告植入的三维图形插入
 - [0078] • OB的三维图形插入
 - [0079] • 其他赞助图像的三维图形插入
 - [0080] • 特定观众位置的三维图形插入
 - [0081] • 特定观众的三维图形插入
 - [0082] • 特定时间的三维图形插入
 - [0083] • 添加人群场景的三维图形插入
 - [0084] • 绿屏替换 (green screen replacement) 的三维图形插入
 - [0085] • 在博物馆和文化、历史遗址和自然遗产解说中心辅助学习的教育内容的三维图形插入
 - [0086] • 场景中对象的绝对和相对尺寸的测量
- [0087] 上述定义的方法或系统,其中电影摄影机是下述之一：
- [0088] -所有具有标准视角范围的摄影机 (cameras with a standard tether)
 - [0089] -需要战术光纤连接 (tactical optical fibre connection) 的摄影机
 - [0090] -需要RF (射频) /无线连接的摄影机
- [0091] 上述定义的方法或系统,部署在以下市场之一：
- [0092] -电影/电视 (非直播)
 - [0093] -商业广告 (非直播)
 - [0094] -商业广告直播
 - [0095] -广播 (非体育)
 - [0096] -广播OB
 - [0097] -棚内体育节目
 - [0098] -基于OB的体育节目
 - [0099] -电视广告置入直播
 - [0100] -互联网用途 (非直播)
 - [0101] -互联网直播
 - [0102] -基于区域的互联网直播
 - [0103] -互联网广告置入直播

[0104] -博物馆/遗产内容

[0105] -博物馆/遗产广告

[0106] -建筑

[0107] -游戏

[0108] 上述定义的用于实现增强现实图像的方法或系统使能够通过任何显示设备显示,包括智能手机、增强现实眼镜、依靠实时定位数据自动改变的增强现实图像的外观。

[0109] 一种影片、电影、电视节目或视频游戏,其中实时计算机生成的三维对象与摄影机的馈送视频混合,其中摄影机的机身可以在三维移动且在摄影机中或连接到摄影机的传感器提供限定摄影机的三维位置或使三维位置能够被计算出的实时定位数据。

[0110] 一种适合连接到传统摄影机的场景测量和追踪设备,其中摄影机的机身可以在三维移动且在场景测量和追踪设备中的硬件传感器提供限定摄影机的三维位置或使三维位置能够被计算出的实时定位数据。

[0111] 一种手持式或便携式摄影机,包括在摄影机中或连接到摄影机的传感器,其提供限定摄影机相对三维参考系的三维位置和三维方向或使三维位置能够通过分析其他数据(例如光流数据)的系统的整体或一部分被计算出的实时定位数据。

[0112] 一种电影摄影机,其包括或连接到立体见证摄影机系统,见证摄影头系统产生广角(如180度)立体图像,启用软件处理图像并在一点也不追踪和移动摄影机系统的情况下产生即时三维点云。

[0113] 一种用于连接电影摄影机的摄影机追踪系统,其具有产生立体图像的见证摄影机系统,启用软件处理图像并在一点也不移动摄影机系统的情况下产生即时三维点云并提供电影摄影机的实时追踪(位置、方向、变焦、聚焦、以及光圈)。

[0114] 增强现实眼镜,包括在眼镜中或连接到眼镜的传感器,其提供限定眼镜相对于三维参考系的三维位置和三维方向或使三维位置能够通过分析其他数据(例如光流数据)的系统的整体或一部分被计算出的实时定位数据。

[0115] 本实施系统被称为Ncam。Ncam的各种应用包括以下:

[0116] 电影、电视和商业广告的视觉效果

[0117] 前期制作和制作的视觉效果:Ncam旨在帮助填补拍摄过程(制作)期间可视化预览(前期制作)和视觉特效VFX(后期制作)之间的差距。Ncam为那些希望在摄影机VFX取景同时实时拍摄提供了解决方法。使用之前创建的常常来自可视化预览过程的内容,Ncam能够实时在实景拍摄电影底片上复合那些可视化预览资源,一般是三维模型或动画。

[0118] 将精心制作且经过检验的可视化预览置于布景上可以节省大量的时间。电影制作人可以设计可视化预览或VFX,它是否正在绿屏上拍摄和复合虚拟背景或在前景上叠加生物或对象。然后,电影制片人能够撤销探测可能的摄影机视角的可视化预览和VFX的控制且在没有通常约束的情况下匆忙并实时移动。

[0119] 依次记录数据并标出时间码,给出VFX部摄影机匹配移动数据(六自由度追踪)同时为编辑提供最终的VFX的“临时”。

[0120] 通过拍摄摄影机中的“临时”VFX和在早期过程中设计VFX,取出VFX中的一些猜测工作作为后期制作过程是可能的。我们能够减轻VFX设计过程,从而节约了损耗。

[0121] 虚拟广告-现场直播

[0122] Ncam的技术非常适用于具有广播空间的实时虚拟广告。数码产品布置变得越来越流行,其中在拍摄后后期制作期间插入产品,与摄影机中的真实物理产品截然相反。然而,通过增加数码产品布置现场直播的能力,我们能够开发各种可能性。例子可以包括在体育场上的品牌标志或肥皂剧中的罐装饮料。实时根据拍摄环境灯光呈现拟真CG图像的能力是至关重要的。

[0123] 全息游戏

[0124] 将Ncam的技术与全息显示器以及眼镜(例如谷歌眼镜)结合可以提供完全身临其境的游戏体验。将现实世界与虚拟世界和人物进行了混合。可能性是无止境的。

[0125] 博物馆和遗产

[0126] Ncam的技术为博物馆和文物部门提供了大范围的潜在应用。

[0127] 作为探索性工具,它可以用于遗址(例如遗迹)的重建,显示出该地点在它旧日辉煌时是什么样子。

[0128] 在博物馆内,Ncam可以用作教学工具,也许是为了展示恐龙的尺寸和运动,它的内脏或骨架结构。另一个例子可以是探索内燃机的工作,有效观察分解的动画图,但是在完全身临其境的三维空间。

[0129] 虚拟摄影机-VCS

[0130] Ncam技术完全适合于这个应用。虚拟摄影机本质上是显示完整CG三维环境的液晶LCD屏幕。当操作者移动LCD屏幕时,它在运动方面就像物理摄影机。LCD摄影机在XYZ方向平移和旋转上进行追踪且实时显示CG引擎的完整的CG环境。当前有各种VCS(虚拟摄影机系统)解决方案可用,但是所有的在技术方面都有局限性。它们趋向于耗费时间设置、局限于它们操作的空间和环境、以及价格昂贵。以Ncam为基础的VCS很可能呈现平板电脑的形状,结合LCD屏幕本地计算过程。可以根据要求从远程PC或可能本地进行CG计算。这种理念使得能够通过可用性和价格点向多个电影中的之前不能使用这些工具的人的部门开放存取。这对于设计师、建筑师、VFX、游戏公司、CG和动画工作室等是非常有用的。

[0131] 这个VCS也可以形成潜在博物馆和遗产解决方案的支柱。

[0132] 虚拟侦察系统-VSS

[0133] 如果你设想正在拍摄一部电影,且它主要在蓝幕/绿幕上。作为电影制片人或导演,我如何解决哪一个视角或镜头将是最好的?终究,我可能有少量的物理布景和几个演员,但是除了我的想象我不清楚哪个摄影机视角是最好的,更不必说向整个剧组解释我们应该在哪里安装摄影机和做出镜头的有效选择。当前我们可以移动电影摄影机和它的所有设备至布景上的位置,设置Ncam并仅仅查看,即认识到它在另一个位置会更好。要是我们有轻量级便携式系统迅速且高效地做出那些决定就好了。这就是Ncam VSS。

[0134] 像今天这样,VSS是Ncam,但是集成到小型化摄影机-想象是DSLR(数码单反相机)。本质上它是具有Ncam优势的数码取景器。像今天这样,我们结合所有的Ncam元件在电影布景上,而是是在便携式设备中。电视连播的虚拟产品也可以非常受益于此工具,博物馆/遗址以及想要向潜在投资者展示他们的新创作的建筑师同样。

附图说明

[0135] 图1

[0136] 配置2.3的硬件描述。要注意的是,在此绘出的版本合并了追踪工作站和呈现工作站。

[0137] 图2

[0138] 实时获得电影摄影机的位置和旋转的总体追踪算法。

[0139] 图3

[0140] 为我们的系统创建的三维自动调平注册对象。该对象用于真实场景中以在我们的软件中自动计算真实世界的起点、方向和尺寸。平面版本在图5的右侧示出。

[0141] 图4

[0142] 用于自然标记的检测的图像金字塔。该图也示出了一些与标记有关的二维纹理补丁。

[0143] 图5

[0144] Ncam开发和使用的校准表。左:用于校准所有镜头的校准表。中间:用于电影摄影机和见证摄影机之间的对齐计算 (lineup computing) 的校准表。左边的表也可以用于电影摄影机上的焦距足够短的情况 (在布景上通过实验确定)。右:注册对象的平面版本。更容易运用且可以放置在地面上任何位置。当添加虚拟对象至真实场景时,圆圈给出了给定世界方向、比例和起点(0,0,0)的标准正交基。上部左边圆圈的中心是世界的起点。

[0145] 图6

[0146] 安装到具有不同悬挂器尺寸的(15mm和19mm杆)的电影摄影机的摄影机杆。

[0147] 图7

[0148] 无线测量设备。当需要测量时(即例如单视场的情况(参见2.1,或非常复杂的装置(参见2.1.3的末尾))),然后,整个追踪过程使用该测量设备实现。当测量完成时,所有的追踪数据(点云、关键点等)无线传输至追踪/呈现工作站,其即刻恢复追踪并开始流输出追踪数据和最终复合的结果。

[0149] 图8

[0150] 摄影机/计算机接线盒。要注意的是,此处示出了使用多个RJ45电缆的版本。所有这些接线盒也具有变体,其中那些电缆被单个战术光纤电缆替换。

具体实施方式

[0151] 本技术是关于在电影布景、电视广播或视频游戏上混合实时计算机生成的三维对象和来自真实摄影机(real camera)(本发明中其余部分称为电影摄影机)的真实生活馈送视频。本系统的主要目标是:

[0152] • 为电影导演提供虚拟人物和环境的实时取景工具。

[0153] • 实时显示直接融入真实视频中的虚拟对象的照片般真实结果。

[0154] • 为现场拍摄和电影后期制作提供电影摄像机位置的精确估计(称为摄影机追踪)。适用于所有的电影摄影机:升降机摄影机、摄影机稳定器、手持式摄影机、三脚架摄影机等。

[0155] • 提供真实场景的三维几何模型的近似值、真实对象(和/或演员)和虚拟环境(三维对象、人物等)之间的操作遮挡。

[0156] 在此描述的技术首先是强大的传感器融合算法,因此融合从3个(立体视场的情

况)或2个(单视场的情况)不同的光流和4个其他传感器(陀螺仪、加速计、磁力计、三维距离传感器)的物理测量结果计算出的自然标记(即非手动添加到物理场景中的)。它也是完全非入侵式系统,不需要手动添加任何物理标记或在真实场景中直接可见的任何物体做它的工作。

[0157] 1 硬件

[0158] 本技术根据必须流输出的数据的类型具有两种不同可能的硬件配置。

[0159] 这两种配置共同具有以下安装到称为摄影机杆(详情参见图6)的设备上的硬件,以及单独的连接器,其本质上是合并或分开所有不同的电缆至一个或几个独立的绝缘线束(loom)(见图8详细的图)的接线盒。

[0160] • 根据选定的方式(单视场或立体市场)的一台或两台高速摄影机(至少100fps),称为见证摄影机。

[0161] • 每个见证摄影机上安装有180度鱼镜头。

[0162] • 九自由度传感器,包括陀螺仪、磁力计和加速计(或当磁力计不能使用时是六自由度)。

[0163] • 三维距离传感器,其拍摄像素的深度。

[0164] 除了摄影机杆之外,所有的配置具有安装到电影摄影机上的普通镜头编码器以读取变焦、聚焦和光圈值。它们也共同具有无线测量装置(详情参见图7)以测量场景并研究它。然而,应当指出的是,在完全立体视场(见2.5)的情况下,更确切地说即时测量的情况下(见2.5.1),不需要测量设备。

[0165] 除了这些共同的特征,以下描述了这两种配置:

[0166] 配置1:摄影机流

[0167] 1.无线嵌入式计算机

[0168] (a) 嵌入式追踪计算机安装在电影摄影机上。它无线流式输出电影摄影机平移和旋转(RT)以及来自编码器的变焦、聚焦和光圈(ZFI)至外部呈现工作站。

[0169] (b) 外部内容生成工作站。该机器通常运行计算机制图软件以为最终复合提供三维计算机生成的动画和相关的CG内容。它也流式输出动画数据至外部呈现工作站(1-1c)。

[0170] (c) 外部呈现工作站,其一方面从嵌入式计算机(1-1a)接收数据,另一方面也运用来自内容生成工作站(1-1b)的动画数据。这种呈现工作站使用来自1-1a的电影摄影机RT+ZFI显示来自1-1b的三维虚拟环境,且在来自电影摄影机的真实的视频流内混合结果。它也流式输出最终复合的结果回到电影摄影机目镜或无线控制监控器。

[0171] 2.有线嵌入式计算机

[0172] (a) 嵌入式追踪计算机安装到电影摄影机上。它使用战术光纤电缆流式输出电影摄影机平移、旋转(RT)以及编码器数据(ZFI)至外部呈现工作站。

[0173] (b) 类似配置1-1b。

[0174] (c) 类似配置1-1c,除了最终复合的结果通过HD-SDI(高清晰度串行数字接口)电缆发送回摄影机。

[0175] 配置2:总流

[0176] 1.无线嵌入式计算机

[0177] (a) 嵌入式追踪计算机安装到电影摄影机上。在该单元没有实现真实的处理。它只

是无线流式输出所有的传感器数据至外部追踪工作站。流式输出见证摄影机的2×视频流、1×九自由度数据、1×距离传感器数据。数据的总量需要大于142MB/sec 最小的带宽(见证摄影机要求63MB/sec,距离传感器74MB/sec,九自由度传感器4.4MB/sec)。

[0178] (b) 外部追踪工作站,其从嵌入式追踪计算机接收数据。这个单元处理所有的传感器数据以计算电影摄影机的旋转和平移矩阵。它流式输出摄影机RT+ZFI至外部呈现工作站(2-1d)。

[0179] (c) 外部内容生成工作站。该机器通常运行计算机制图软件以为最终复合提供三维计算机生成的动画和相关的CG内容。它也流式输出动画数据至外部呈现工作站(2-1d)。

[0180] (d) 外部呈现工作站,其一方面从嵌入式计算机(2-1b)接收数据,另一方面也运用来自内容生成工作站(2-1c)的动画数据。这种呈现工作站使用来自2-1b的电影摄影机RT+ZFI显示来自2-1c的三维虚拟环境,且在来自电影摄影机的真实的视频流内混合结果。它也流式输出最终复合的结果回到电影摄影机目镜或无线控制监控器。

[0181] 2. 有线嵌入式计算机:

[0182] (a) 嵌入式追踪计算机安装到电影摄影机上。它使用战术光纤电缆或多个Cat-6RJ45电缆流式输出电影摄影机平移和旋转(RT)至外部呈现工作站(2-2c)。编码器再次用于变焦、聚焦和光圈(ZFI)。

[0183] (b) 类似配置2-1c。

[0184] (c) 类似配置2-1d,除了最终复合的结果通过HD-SDI电缆被发送回摄影机。

[0185] 3非嵌入式计算机(参见图1图示说明:所有其他版本可以很容易地从此得出):

[0186] (a) 单一摄影机追踪设备(摄影机杆)被连接到电影摄影机且连接到接线盒。

[0187] (b) 外部追踪工作站,其从嵌入式追踪设备接收数据。这个单元处理所有的传感器数据以计算电影摄影机的旋转和平移矩阵。编码器再次被用于变焦、聚焦和光圈(ZFI)。它流式输出摄影机RT+ZFI至呈现工作站(2-3d)。

[0188] (c) 类似配置2-1c。

[0189] (d) 类似配置2-1d,除了最终复合的结果通过HD-SDI电缆被发送回摄影机。

[0190] 除了之前的配置之外,每个版本也具有两个其他变体。一个具有实际上与外部呈现工作站融合的内容生成工作站,第二个变体具有与追踪工作站融合的外部呈现工作站。

[0191] 2. 软件

[0192] 已创建若干新的科学方法和软件以实现精确的追踪结果和实时增强现实[19]。这些方法在下一部分进行描述。

[0193] 2.1单视场无标记馈送视频追踪(Monoscopic Marker less Tracking from video feed)

[0194] 2.1.1过程概述

[0195] 在该技术中使用的单视场技术在两个单独的通路上建立,这两个通路技术上几乎相同,除了一个需要用户交互然而另一个完全自动。

[0196] 第一通路被称为测量。用户采用了建立在部分1(也参见图7)的第二段描述的硬件上的设备,以扫描真实场景和注册对象。在该阶段,当用户四处扫描环境时,由系统在三维自动获悉和计算自然标记。用户一开始进行测量,他也就必须拍摄注册对象(参见2.1.3.2)以便自动估计世界的比例、起点和方向。一旦这种情况实现,用户可以继续扫描场景的剩余

部分以建立自然标记的可能的最好的三维点云(自然标记是人眼不可见的特征点,其由我们的算法计算)。这个测量实现的越彻底,最终的结果越好。测量部分是所有的工作中最重要的工作,它非常关键。

[0197] 第二通路是电影摄影机追踪(参见2.1.4),其由位于电影摄影机中的追踪设备完成。根据各种情况以及遵循之前题为“硬件”部分的说明,适用不同的配置。这个步骤是全自动的,且通常不需要任何人机交互。

[0198] 2.1.2算法概述

[0199] 在此描述的技术是根据多个传感器融合技术的增强现实构架(参见2.3)。

[0200] 事实上,它不依赖于通常的唯一光流数据来追踪摄影机和插入虚拟对象。在相当多的例子中这种技术已经证明它们的科学和实际限制。例如,如果用户隐藏了用于追踪的摄像机,那么追踪失败和丢失。如果摄影机指向还没有测量的区域(参见2.1.3),会发生完全一样的问题。此外,即使电影摄影机一点不移动,由于由纯视觉追踪算法计算出的所检测到的自然标记的精度,仍然存在一些不可避免的轻微抖动问题。滤除数据部分解决了这个问题,但你决不会得到完全稳定的虚拟对象,当使用长焦距镜头时更是如此。由于我们的技术使用其它传感器来知道摄像机移动与否,因此我们的追踪算法不会有这个问题。

[0201] 普通光流技术像SLAM[9],PTAM[8],RSLAM[13]等,使用由摄影机拍摄的视频中包含的信息粗略估计它的位置和方向。对于它的视觉追踪部分(称为视觉追踪),我们的技术具有类似的方法,除了一方面当它丢失时跟踪的核心并不是简单的关键帧搜索,第二方面它使用7个可用的传感器(一个电影摄影机,两个见证摄影机,陀螺仪,加速计,磁力计和三维距离传感器)的所有数据来计算摄像机的精确位置和方向。

[0202] 我们使用扩展卡尔曼滤波(EKF,参见2.3.2)[7,21]整合以下传感器:陀螺仪,加速计,磁力计传感器和见证摄影机。EKF技术是我们软件的主体且所有的预测/校正算法是基于允许使用每个传感器中最佳的数据合并方法。这为所有其他技术不能使用的简单的和关键的情况提供了无与伦比的稳健性。事实上,当传感器失去追踪时(实际上意味着它的测量结果不再可靠),EKF仍然能够通过合并其他可靠的和剩余的传感器的数据获得位置/方向估计。除此之外,由于每个传感器具有置信度,这影响合并算法以便防止系统使用不准确的信息。例如,普通陀螺仪具有天生的缺陷称为漂移,其往往更改它的方向,甚至当它不动时。运行时间越长,陀螺仪产生的误差越大。陀螺仪的漂移校正通过使用视觉追踪来完成,当它的误差太大时视觉追踪由陀螺仪校正。EKF能够通过使用之前正确的估计值预测电影摄影机的运动从而找到电影摄影机的最佳的位置和旋转,然后根据所有传感器给出的新的测量结果校正它的预测。

[0203] 此外,当摄影机丢失时,大多数纯光学为基础的追踪技术使用关键帧(即快照(snapshot))来定位摄影机。如果你用摄影机指向的当前视野没有关键帧,那么恢复失败。更强大的技术当然是可用的,像SIFT[11]或SURF[1]技术,以解决例如你仅仅具有垂直旋转的摄影机。事实上,由于所有这些系统比较当前帧和最相似的关键帧,使用旋转和缩放不变性描述符获得更好更快的匹配是很有趣的。然而,如果你在摄影机的位置没有任何关键帧,那么你的系统没有办法恢复它的位置。我们的技术集聚了每个技术(需要指出的是我们使用ORB[18]作为关键帧描述符)的优点,并且我们同时运行三个不同的线程来恢复丢失的追踪(详情参见2.1.4.4)。恢复过程中可以看见基于概率和随机方法的总体算法。当追踪丢失

时,立即使用椭圆形方法在最新已知位置周围三维空间采样且将该椭圆形内部所有的关键帧与目前的关键帧进行比较。此外,所有来自剩余传感器(陀螺仪,加速计,磁力计,三维距离传感器)的其他数据仍然合并搜索中,允许我们的软件消除所有看似合理的候选。当然,如果没有关键帧是足够好的候选,那么系统会使用除了视觉追踪之外的所有其他传感器,以计算当前位置的近似值。该技术的显著结果是,我们的技术允许用户在不丢失追踪的情况下,将摄影机指向尚未被测量的位置。然而我们的算法从不停止采样三维空间以寻找匹配目前快照的关键帧(即更好RT近似值)。如果恢复了视觉追踪,使用和更新所有其他传感器数据以除了校正摄影机之外校正它们自己的信息(参见扩展卡尔曼滤波2.3.2)。

[0204] 我们的技术的最新(即普通)版本是立体视场(参见2.5),这意味着它使用了间隔21厘米的两个见证摄影机(见图6)。这种技术不需要任何单应性(homography)的计算或任意单个见证摄影机的手动平移(参见2.1.3)来计算馈送视频中检测到的自然标记的深度信息。该方法在大多数情况下通过测量阶段淘汰简化了我们的主要算法(参见2.5)。如在单视场技术中,视觉追踪本身具有完整的子像素方法,允许系统在图像金字塔的4个不同水平追踪摄影机位置(参见图4),然而所有其他以光流为基础的技术当搜索特征点时只在两个不同的图像分辨率不正当地使用像素技术。

[0205] 2.1.3第一阶段:测量

[0206] 2.1.3.1初始化步骤

[0207] 第一步是用户采用设备完成物理/手动水平平移,以便拍摄两个画面。该算法通过使用称为FASTER[17]的角点检测方法自动检测图像中的自然标记(即特征点)。应用匹配以找到点对之间的对应关系。对于一个图像,所有检测到的点必须保持在相同的三维平面上,以便能够计算出从该平面上到图片正确的单应性。这提供了有关三维平面的摄像机的转动和平移。第二张图片遵循相同的原理,我们得到当前图片的摄像机的第二位置和平移。现在可以计算从一个摄像机到另一个的转换,并获得第一三维点云。

[0208] 让我们考虑用摄影机拍摄到的两个视图(左和右)。对于每个视图,见证摄影机指向平面上存在的点M。 M^L 和 M^R 分别是M在左和右视图中的透视投影。

[0209] 我们可以写成:

$$[0210] \quad M^L = H \cdot M^R \quad (1)$$

$$[0211] \quad H = (D \cdot R + \vec{T} \cdot \vec{N}^*) \quad (2)$$

[0212] 其中:

[0213] • R是转动矩阵,以其 M^L 关于 M^R 旋转。

[0214] • $\vec{N}^*(A, B, C)$ 是到点M属于的平面的法向矢量。

[0215] • D是到平面的距离,该平面例如 $Ax+By+Cz+D=0$ 。

[0216] 等式1中的H的求解为:

$$[0217] \quad H = \begin{pmatrix} x_{M^L} & 0 \\ y_{M^L} & 0 \\ 1 & 0 \\ 0 & x_{M^R} \\ 0 & y_{M^R} \\ 0 & 1 \\ -x_{M^L} * x_{M^R} & -x_{M^L} * y_{M^R} \\ -y_{M^L} * x_{M^R} & -y_{M^L} * y_{M^R} \\ -x_{M^R} & -y_{M^R} \end{pmatrix}^t$$

[0218] 我们的算法需要4个共面点的最小值进行正确的初始化。这是很平常的,我们其实有比这更多的点。这就是为什么对于每个单应性我们使用整个点集中的四个点计算n个可能的单应性。然后,我们使用RANSAC[4]方法实际上可以找到最佳可能的单应性,并建立了一组正常值和异常值的候选集。异常值将被拒绝而正常值将由高斯-牛顿方法(Gauss-Newton method)进行完善,使两个视图中点的重新投影的误差减到最小。从滤除的点集中计算出最好的单应性,现在需要估计左和右视图之间的旋转和平移矩阵。这很容易通过使用两个正交矩阵U,V和对角矩阵T的奇异值分解来实现。矩阵H现在写为H=U Γ V^t

$$[0219] \quad \Gamma = (D' \cdot R' + T' \cdot N'^t) \quad (3)$$

[0220] 其中:

$$[0221] \quad R = \det(U) \cdot \det(V) \cdot U \cdot R' \cdot V^t$$

$$[0222] \quad T = U' \cdot T'$$

$$[0223] \quad N = V \cdot N'$$

$$[0224] \quad D = \det(U) \cdot \det(V) \cdot D'$$

[0225] 当我们不能运用待定的单应性时,等式3有两种可能的解法:

[0226] 解法1:

$$[0227] \quad T' = (\lambda_2 - \lambda_3) \begin{pmatrix} \varepsilon_1 \sqrt{\frac{\lambda_2^2 - \lambda_3^2}{\lambda_1^2 - \lambda_3^2}} \\ 0 \\ -\varepsilon_2 \sqrt{\frac{\lambda_2^2 - \lambda_3^2}{\lambda_1^2 - \lambda_3^2}} \end{pmatrix} \quad (4)$$

$$[0228] \quad R' = \begin{pmatrix} \cos \theta & 0 & -\sin \theta \\ 0 & 1 & 0 \\ \sin \theta & 0 & \cos \theta \end{pmatrix} \quad (5)$$

[0229] 其中:

$$[0230] \quad \begin{cases} \cos \theta = \frac{(\lambda_2^2 + \lambda_1 \lambda_3)}{\Lambda} \\ \sin \theta = (\lambda_1 - \lambda_3) \cdot \varepsilon_1 \varepsilon_3 \frac{\sqrt{(\lambda_1^2 - \lambda_2^2)(\lambda_2^2 - \lambda_3^2)}}{\Lambda} \\ \Lambda = \lambda_2(\lambda_1 + \lambda_3) \end{cases}$$

[0231] λ_1 、 λ_2 和 λ_3 是来自等式(2)的矩阵H的特征值,排序为例如 $\lambda_1, \lambda_2, \lambda_3$ 且 $\lambda_1 \neq \lambda_2 \neq \lambda_3$ 。

[0232] 解法2:

$$[0233] \quad T' = (\lambda_1 + \lambda_3) \begin{pmatrix} \varepsilon_1 \sqrt{\frac{\lambda_2^2 - \lambda_3^2}{\lambda_1^2 - \lambda_3^2}} \\ 0 \\ \varepsilon_3 \sqrt{\frac{\lambda_2^2 - \lambda_3^2}{\lambda_1^2 - \lambda_3^2}} \end{pmatrix} \quad (6)$$

$$[0234] \quad R' = \begin{pmatrix} \cos \varphi & 0 & \sin \varphi \\ 0 & -1 & 0 \\ \sin \varphi & 0 & -\cos \varphi \end{pmatrix} \quad (7)$$

[0235] 其中:

$$[0236] \quad \begin{cases} \cos \varphi = \frac{(\lambda_1 \lambda_3 - \lambda_2^2)}{\Lambda} \\ \sin \varphi = \varepsilon_1 \varepsilon_3 \frac{\sqrt{(\lambda_1^2 - \lambda_2^2)(\lambda_2^2 - \lambda_3^2)}}{\Lambda} \\ \Lambda = \lambda_2(\lambda_1 - \lambda_3) \end{cases}$$

[0237] λ_1 、 λ_2 和 λ_3 是来自等式(2)的矩阵H的特征值,排序为例如 $\lambda_1, \lambda_2, \lambda_3$ 且 $\lambda_1 \neq \lambda_2 \neq \lambda_3$ 。

[0238] 除了特征点集,该算法也拍摄两个关键位置(KP)作为见证摄影机从每一个视角看到的物体的快照。附加点的获悉总是三个组成部分的总和:新检测到的点,关键位置和补丁,这在2.1.3.3进行了描述。

[0239] 最后,请注意,在立体显示的情况下(参见2.5),测量阶段是自动的和即时的。对于极其复杂的情况,手工测量可能仍然使用(在场景中实验决定),但是这是无对照(anecdotal)的。

[0240] 2.1.3.2 三维世界的自动缩放、起点和方向

[0241] 在单视场测量期间,计算世界的精确的原点,缩放和取向是不可能,这个世界要么需要真实对象的先验知识,包括它的精确形状和尺寸,要么需要用于计算单应性的这两个视图之间使用的距离。应当指出的是,在最后情况下,这无论如何不会给你该世界的起点或方向。大多数系统没有把这部分考虑进去。他们经常要求用户通过输入视频中可见的三维对象的尺寸手动进行缩放。其他选项是在拍摄过程中使用户在特定的建模/动画软件中按比例绘制三维虚拟角色,且也适当地将其在真实世界中对准。然而,任何尺寸误差——即使很小——对最终复合的质量也有显著影响,这就是为什么在我们的技术(见下文)中它是

如此仔细地实现。一个不良比例观察到的最常见的结果是，用户感觉该虚拟对象在地面上滑动，而不是完全附属于它。

[0242] 我们提出了一种新的方法，当进行测量时其需要把专有的注册对象放在场景中，由于对象具有非常清楚的形状且所有它的尺寸也是已知的，因此自动检测到该对象。无需用户交互。

[0243] 我们的注册对象是由几个部分组成的，其本质上是球形和支柱，如图3所示。球体是在我们的系统中该对象的关键部分，因为它们允许使用二进制大对象检测算法通过视频识别其图案。然后，计算投影球体(即圆)的重心，并且当我们准确地知道每个球之间的方向和距离，该系统能够给出世界的非常精确的比例。我们还提供了注册对象的平面版本(参见图5的右图)，这常常更方便地从真实舞台上的一个地方移动到另一个地方。

[0244] 一旦这个已经实现，用户能够例如把一个1.80米高的虚拟对象放到真实的视频中，并确保它看起来真的是那么高和正确。其他方法由于任务手动方面无法获得这种质量。

[0245] 自动缩放实际上是它自身测量期间完成的工作，因为它发生在测量之初，把它也认为是初始化步骤是非常重要的。计算的缩放真正用于测量之后的步骤以建立三维标记的精确布景。现在让我们来看看详细的测量本身。

[0246] 2.1.3.3测量过程和点云构造

[0247] 该测量需要用户使用测量设备穿过场景。由于该系统根据它在以前获悉的置信度总是计算新点，总是以见证摄影机从之前的布景中看到足够多的点以添加新候选的方式建立测量。

[0248] 在见证摄影机移动期间，该算法创建由在图像金字塔的四个不同水平(level)获悉的三个数据集组成的映射(参见图4)：

[0249] • 关键位置。关键位置包含见证摄影机看见的物体的四个不同分辨率的快照。金字塔的低水平是见证摄影机的原始分辨率(在本例中是 640×480)。该关键位置在跟踪弄错时也被该算法用于恢复摄影机位置。

[0250] • 由FASTER[17]在关键位置的每张图像中估计的二维特征点集。FASTER是非常著名的算法，其主要特点是角点检测算法。每一个关键点还具有附加的ORB描述符，以确保当丢失追踪时更快的恢复(参见2.1.4.4)。

[0251] • 集中在由前面步骤检测到的每个特征点上的补丁集(16×16 二维纹理)。在新点检测期间，无法从单个视图开始计算它们的深度。这就是补丁的用途。可以通过关键位置的4张图片找到在尽可能接近的关键位置的两个补丁之间的一致性来应用极线搜索(参见图2，点云构建器矩形)。一旦在两个视图中检测到点(即补丁)，那么有可能计算出三维特征点。该三维特征点集被称为映射。明白关键位置A和B之间的搜索是通过穿过金字塔A和B的同等水平以及B的金字塔分支实现的是重要的(见图4)。

[0252] 在映射构建期间，摄像机根据测量操作者的位移移动。此时，我们的软件只知道先前计算(即跟踪)出的摄像机参数。为了计算摄像机的新位置，我们需要当前画面和自然标记的三维点云。一方面，FASTER在图像金字塔的几个水平(当前水平和当前水平+1)计算二维标记集，另一方面，三维点云重新投影到当前画面中。如果我们知道摄像机的位置为从它的视角被重新投射的点，最后步骤才能实现。但是，这正是我们试图计算的。因此，我们的软件使用与由我们的九自由度传感器(见2.3)所提供的信息关联的恒速模型(constant

velocity model)以使用先前正确计算出的位置来预测摄影机的下一个位置。使用该预测,三维点云可以重新投影到当前画面上且可以应用点匹配算法找到由FASTER检测的二维点和三维点云的投影点之间的对应。使用莱文贝格-马夸特[10,12,15]算法(Levenberg-Marquardt[10,12,15]algorithm)使两组标记集之间的误差减到最小,因为它是已知的对于那种问题最好的优化算法。如果匹配点的数目除以投影点的总数大于给定的临界值,该系统能够成功追踪摄影机(追踪良好),并且允许添加新的关键位置。为了后来的用户,由FASTER检测到的在三维点云中没有找到匹配的二维点以及他们相关的二维 16×16 的纹理补丁存储在存储器中。他们将被要求产生新的三维特征点(见下一段)。

[0253] 如果三个条件都满足,添加新的关键位置(以及新的特征点)的加入。首先,正如我们在上一段所说,跟踪必须良好意味着它足够精确或没有丢失。第二,每30个画面(每秒1个)添加关键位置以避免建立过大的数据集。第三,新的关键位置必须在距最靠近关键位置10cm的最小距离。这可以防止系统静止时,获悉附加点。

[0254] 当顺利通过所有这些测试时,该软件可以添加新的快照和新的三维特征点。第一步是再次使用FASTER在画面(快照)中检测新的相关二维点。因为我们有来自追踪的不匹配二维特征点集,我们现在尝试匹配“旧”的二维特征点集和新计算的点集。这是通过使用新点集的二维纹理补丁由传统极线搜索(classical epipolar search)完成的。补丁沿着极线移动(参见图2)尝试匹配“旧”点集中的二维补丁集。由于极线约束,计算来自两个二维点的匹配的三维点的深度是可能的。这实际上类似于单应性阶段的匹配步骤(参见2)。如果新的点被添加到映射上,应用局部光束法平差(local bundle adjustment)。在使用这些新的特征点改善了摄影机的位置和旋转之后,莱文贝格-马夸特最小化算法用于改善所有特征点的三维位置。这是在 $k+1$ 个画面窗口上完成的,意味着最小化考虑了当前快照加上 k 最接近的快照以正确估计摄影机的当前位置和旋转(局部光束法平差)。k的值可以是根据实验确定或者根据最小化算法给定的误差临界值适应性计算。

[0255] 除了局部光束法平差,总体光束法平差也应用于所有现有的关键位置。由于关键位置集随着时间变得越来越大,系统需要越来越多的时间进行总体调整。如果添加了新的关键位置,我们的软件没有足够的时间完成局部和/或总体光束法平差,然后调整就会取消以将优先级留给添加的新的关键位置。新的关键位置的添加一旦成功它们就会被再次应用。

[0256] 一旦现有的点云存在,可以追踪摄影机。这在下一部分进行描述。

[0257] 2.1.4第二阶段:电影摄影机追踪

[0258] 2.1.4.1几何摄影机校准

[0259] 摄影机校准有几个目标。一方面,它给出我们数学关系以描述如何在传感器上创建图像。另一方面,它计算摄影机的图像平面上三维点和它们的投影之间的转换矩阵。

[0260] 测量和追踪过程的这个重要的先决条件意味着我们必须计算两种不同的属性:内在和外在参数。内在参数与摄影机和镜头特性有关,并根据它们被物理构建的方式移动。它们包括图像格式、摄影机的光学中心与图像平面之间的距离、以及主点。图像格式是由将像素和距离关联的两个比例因子组成。主点是光学中心在图像平面的投影(即理想情况下在图像的中间)。除此之外,一些镜头需要计算它们的光学畸变并考虑到用于跟踪。在校准过程中,具有若干圆形图案的平坦表面由摄影机连续拍摄(参见图5)。对于所有的位置,通过

莱文贝格—马夸特最小化算法计算外在参数(即摄影机的位置和旋转)和内在参数(包括光学畸变),其使所有快照之间二次投影误差最小。

[0261] 这个过程是基于德韦尔奈(Devernavy)和福热拉(Faugeras)的工作[3]的非常著名的技术。

[0262] 2.1.4.2电影摄影机偏移

[0263] 这里开发的总体系统需要放置直接连接到目标电影摄影机的追踪设备。使用之前的描述符追踪技术估计见证摄影机的位置和方向。然而,存在通过为电影和见证摄影机使用相同的图案的缩放的问题。事实上,由于我们的见证摄影机采用180度镜头,校准图案中的圆形往往在见证摄影机视野中会出现得非常小从而使他们很难检测。

[0264] 我们已经创建了我们自己的校准表,其实际上是由两种不同的尺寸和对比图案组成的。表的一个部分在纯白色的背景上具有纯黑色圆圈(针对见证摄影机)。表的第二部分是由在纯黑色背景上的纯白色圆圈组成的(电影摄影机)。然后,我们使用非常简单的图案检测器寻找见证和电影摄影机的这两个图案。因为我们知道这两个图案的精确相对位置,通过“传播”计算出的电影摄影机的位置和旋转寻找电影摄影机的位置和旋转,反之亦然。事实上,由于两个摄影机直接连接在一起,我们能够利用对齐算法精确计算他们之间的三维偏移(R,T)。我们已经开发出了基于先前描述的圆形图案的对齐算法。图案的确切大小,形状和位置是已知的,由两个摄影机分别识别图案,从而增加了见证和电影摄影机之间的简单数学关系(旋转和平移)。

[0265] 2.1.4.3追踪阶段

[0266] 从追踪过程中分离测量仅仅是务实的方式,以描述如何在布景上使用系统。事实上,跟踪过程与图2描述的追踪并无不同且在测量期间使用。然而,它使用了一个更简单的设备(参见图6),其与测量设备相当接近,除了由于很明显的的原因它不包括触摸屏。

[0267] 图2的视觉跟踪矩形框描绘了如何找到在新画面中FASTER检测到的二维自然标记和现有的三维点云之间的匹配。与每个自然标记有关且来自两个数据集的二维纹理补丁进行线性比较。一旦所有的特征点都已经被处理,应用高斯牛顿算法找到的最佳可能的匹配,并计算摄影机的位置和方向。这样的最小化算法通常需要大约10次迭代以收敛于该解答。为了增加的标准稳健性达到最小化,我们利用一个统计Tukey M估计量(最大似然型估计量)[6]。这也确保了收敛不被异常值中断。

[0268] 每个追踪在获得摄影机的最终参数之前在金字塔两个不同的水平进行两次。首先,通过以图像金字塔中粗糙水平来计算摄影机的位置和方向的近似值和使用自然标记的子集大大提高了该算法的收敛性。第二,在更大的画面上(约20倍大)实现了相同的计算,并且与先前的近似值合并以获得最终精确的结果。

[0269] 2.1.4.4追踪失败和恢复

[0270] 摄像机在真正布景上追踪时,很多情况都可能发生,其中视觉追踪丢失。这往往发生,例如当待在见证摄影机或杆前面的人们仅仅是被物体挡住时。对于所有我们失去视觉追踪的情况,三种不同的恢复线程同时启动,其中第一个得到恢复的被接受。

[0271] 使用扩展卡尔曼滤波(参见2.3.2)的预测功能,可以首先尝试摄影机位置恢复,从EKF预测参数开始搜索。事实上,由于EKF是预测校正算法,它总是在它实际移动到那个位置之前试图猜测摄影机的下一个位置。当预测正确时,它因此可能得到即时恢复。在实践中,

当见证摄影机隐藏时和当系统使用IMU(惯性测量装置)追踪时预测进行得非常好。然而,当隐藏见证摄影机时电影摄影机也存在显著的平移的情况下,它常常失败。

[0272] 恢复的第二个层次非常不同。在ORB(定向的FAST和旋转的BRIEF)技术[18]之后,每个特征点都有与之关联的旋转不变性描述符。它首先依赖于著名的FAST[16]技术来计算与哈尔斯评分(Harris score)[5]相关联的多尺度特征(即在图像金字塔的不同层次,参见图4)。该描述符的定向使用强度的质心来计算。第二个描述符是rBRIEF,这是在关键点的定向之后旋转的简要的描述符[2]。使用这两个描述符,我们的技术能够从普通视觉追踪技术失败的位置恢复。例如,让我们说,由于任何原因我们失去了视觉追踪。假设我们现在平移摄影机和将摄影机由绕Z轴旋转90度,同时仍然丢失。摄影机看到的当前图像以前从未获悉,但它仍然指向我们添加的关键位置和之前获悉关键点的方向。然而,由于普通恢复过程不是旋转不变的,这些系统无法恢复追踪。在我们的例子中,只要电影摄影机已经指向某个之前的位置(旋转无关紧要),与每个关键点有关的旋转不变性描述符的使用可以快速恢复(使用暴力匹配技术(brute force matching technique))。

[0273] 最后,第三个恢复线程是更加残酷和计算密集型的。围绕最后已知的摄影机位置建立椭球形(三维椭圆)。该椭球具有由计算出的参数协方差组成的大小,但是扩展卡尔曼滤波。然后在EKF的误差协方差之后椭球内产生多个样本。RANSAC算法应用到该集,以便为所搜索的摄影机的位置和方向找到最佳可能的匹配。。

[0274] 2.2动态变焦、聚焦和光圈

[0275] 在总体追踪过程期间,例如导演和摄影导演可以决定变焦或聚焦于场景中的人物/对象。我们的技术集成了若干普通编码器,例如阿莱(Arri)提出的编码器,但也能运用任何其他编码器来实时捕捉变焦、聚焦和光圈值。这允许动态变焦和布景上的景深。

[0276] 2.3传感器融合技术

[0277] 所有纯光学为基础的无标记技术不能在许多情景中追踪。最常见的情况是在见证摄影机的当前视野内没有留下足够的天然标记。在理想的情况下,这种情况很少发生,因为用户集中注意力不使例如他的手挡住镜头。在现实中,这种情况发生了很多,只是因为摄制组经常需要更换镜头,修改摄像机支架,清洁镜头,移动到另一个拍摄地点等。在以普通标记为基础的技术和无标记(光流为基础)的技术中,这防止了追踪系统找到该摄像机的位置和旋转的良好估计。

[0278] 我们创建了传感器融合方法,其使用多个传感器(见证摄影机、陀螺仪、加速计和磁力计)自动校正见证摄影机的位置和旋转。该技术是我们技术的核心。第一步是系统的校准,这意味着相对见证摄影机校准九自由度IMU。这是一个关键的步骤,以确保所有的传感器在相同的基础下工作。九自由度传感器和见证摄影机之间的关系接近一个手眼技术(hand-eye technique)[20],但我们增加了一些有趣的科学贡献,使之适合我们的跟踪算法。

[0279] 2.3.1IMU-见证摄影机校准

[0280] 让我们考虑转换 $R_{IMU \rightarrow C}$,其计算从九自由度传感器(IMU)到两个见证摄影机(C)之一的旋转。 R_{IMU} 是已知处于局部基(local basis)且 R_C 也已知处于它自己的局部基。目标是解出 $R_{IMU \rightarrow C}$ 。 $R_C = R_{IMU \rightarrow C} \cdot R_{IMU}$

[0281] 确保 $R_{IMU \rightarrow C}$ 总是不变的,由于IMU和摄影机之间的关系不随着时间改变(它们都连

接到静态摄影机杆)。

[0282] 因此写为 $t_i, i \in [0, n]$ 是可能的:

$$R_C^{t_0} = R_{IMU \rightarrow C} \cdot R_{IMU}^{t_0}$$

[0283]

$$R_C^{t_1} = R_{IMU \rightarrow C} \cdot R_{IMU}^{t_1}$$

$$R_C^{t_2} = R_{IMU \rightarrow C} \cdot R_{IMU}^{t_2}$$

[0284] 我们可以从之前的等式得出:

$$R_C^{t_1 - t_0} = R_{IMU \rightarrow C} \cdot R_{IMU}^{t_1 - t_0}$$

[0285]

$$R_C^{t_{n-1} - t_{n-2}} = R_{IMU \rightarrow C} \cdot R_{IMU}^{t_{n-1} - t_{n-2}}$$

$$R_C^{t_n - t_{n-1}} = R_{IMU \rightarrow C} \cdot R_{IMU}^{t_n - t_{n-1}}$$

[0286] 具有关系:

$$[0287] \quad R_C^{t_{n-1} - t_{n-2}} = R_C^{t_1} \cdot (R_C^{t_1 - t_0})^{-1}$$

$$[0288] \quad R_{IMU}^{t_n - t_{n-1}} = R_{IMU}^{t_1} \cdot (R_{IMU}^{t_1 - t_0})^{-1}$$

[0289] 由陀螺仪给出的旋转测量结果和视觉追踪都应该是相同的。因此,考虑到 R_{IMU} 和 R_C 描述相同的转动角但是在不同的地点是合乎逻辑的。使用在李代数中定义为 $\text{Log}: SO(3) \rightarrow so(3)$ 的旋转矩阵的指数映射,转换矩阵 R_C 为轴角 (\vec{r}_C, α) 表示是可能的:

$$[0290] \quad \alpha = \arccos\left(\frac{\text{tr}(R_C) - 1}{2}\right)$$

$$[0291] \quad \vec{r}_C = \frac{1}{2 \sin \alpha} \begin{bmatrix} R_c(3,2) & R_c(2,3) \\ R_c(1,3) & R_c(3,1) \\ R_c(2,1) & R_c(1,2) \end{bmatrix}$$

[0292] 其中 $\text{tr}(R_C)$ 是矩 $\text{tr}(R_c) = \sum_{j=1}^3 R_c(j, j)$,

[0293] 我们现在可以写出以下超定的等式系统:

$$[0294] \quad S = \begin{cases} \vec{r}_C^{t_1 - t_0} = R_{IMU \rightarrow C} \cdot \vec{r}_{IMU}^{t_1 - t_0} \\ \vdots \\ \vec{r}_C^{t_{n-1} - t_{n-2}} = R_{IMU \rightarrow C} \cdot \vec{r}_{IMU}^{t_{n-1} - t_{n-2}} \\ \vec{r}_C^{t_n - t_{n-1}} = R_{IMU \rightarrow C} \cdot \vec{r}_{IMU}^{t_n - t_{n-1}} \end{cases} \quad (8)$$

[0295] 至于任何超定的系统,使用奇异值分解(SVD)解出之前的矩形矩阵为 $S = U \sigma V^t$ 是可能的,其中 U 是 $S \cdot S^t$ 的特征矢量的 $m \times n$ 矩阵, V 是 $S^t \cdot S$ 的特征矢量的 $n \times n$ 矩阵, D 是 S 的有序奇异值的 $n \times n$ 对角矩阵。

[0296] 之前的等式解法遭受测量的噪音的影响。因此我们应用莱文贝格-马夸特最小化

方案使误差最小化:

$$[0297] \quad E = \|(R_{IMU \rightarrow C} \times \overrightarrow{r_{IMU}^{t_k - t_{k-1}}}, \overrightarrow{r_C^{t_k - t_{k-1}}} - 1\| \quad (9)$$

[0298] 由于我们有IMU与见证摄影机之间的数学关系,增添所有的测量结果至扩展卡尔曼滤波以便提取最好的每一个传感器和计算电影摄影机的旋转和平移是可能的。

[0299] 2.3.2扩展卡尔曼滤波

[0300] 扩展卡尔曼滤波(EKF)在航空航天和机器人研究领域是非常著名的技术,其合并来自多个不同的传感器的数据。本质上它通过计算估计误差和应用两个阶段,预测和校正,来估计噪音测量的非线性系统的参数。在预测阶段,卡尔曼滤波使用在先前的时步(timestep)估计的参数和误差计算出新的参数,以及在时步相关的误差。修正阶段使用新的噪音测量更新预测。这些新的测量允许从预测的参数和误差计算新的估计参数和误差。如果误差大于给定的临界值,参数被修正。

[0301] 在我们的例子中,以非常特殊的方式开发EKF,由于我们有多与三维旋转有关的测量结果(陀螺仪+见证摄影机)并且他们不能轻易地篡改。

[0302] 我们技术中EKF的想法是把某种置信值放到每个测量结果中(见证摄影机、陀螺仪、磁力计、加速计)且使用获得最佳的置信度分数的传感器迭代地修正当前估计的位置和旋转。在实践中,EKF是比那要复杂一些,可以用数学描述为非线性系统的预测修正算法。考虑到时步t时的状态向量 $\hat{X}(t)$,算法应用于我们的运动方程的当前物理模型来预测矢量 $\hat{X}(t)$ 新的状态为 $\hat{X}(t + \Delta t)$ 。然后修正阶段在时步t+ Δt 产生新的状态矢量 $\hat{X}(t + \Delta t)$ 。

[0303] 在我们的例子中,EKF模型可以写成如下:

$$[0304] \quad \begin{cases} Y(t) = h(X(t), B) \\ \hat{X}(t+\Delta t) = f(\hat{X}(t), \Delta t) \end{cases} \quad (10)$$

[0305] h是我们的可观察量的状态方程, X_t 是总体中平移、旋转、平移速度和角速度的矢量,f是状态矢量的更新函数,B是系统中的总体噪音。

[0306] h的定义

$$[0307] \quad Y(t) = h(X(t), B) \quad (11)$$

[0308] 如果在一方面,我们考虑来自无标记视觉追踪的测量结果(即见证摄影机)为 $h_{witness}$,另一方面来自IMU的测量结果为HMU,我们可以推断h是由两种可观察量组成的, $h_{witness}$ 和 h_{IMU} 。通过定义,h是简单的函数,它的一般式由鉴定给定: $Y_{witness}(t) = X(t)$, $Y_{IMU}(t) = X(t)$ 。根据发送测量结果的传感器的种类,我们现在可以写成两个版本h,因此在视觉追踪的情况下重写等式11为:

$$[0309] \quad Y(t) = \begin{cases} Y_{witness}(t).T = X(t).T \\ Y_{witness}(t).R = X(t).R \\ Y_{witness}(t).\vec{V} = X(t).\vec{V} \\ Y_{witness}(t).\vec{\Omega} = X(t).\vec{\Omega} \end{cases} \quad (12)$$

[0310] $Y_{witness}(t) \cdot \{a \in T, R, V, \Omega\}$ 是视觉追踪在时步t平移(T,单位为米)、旋转(R,单位

为弧度)、速度(\vec{V} ,单位为 $m \cdot s^{-1}$)和角速度(,单位为 $rad \cdot s^{-1}$)的观察值 y ($y \in R^{12}$)。

[0311] 在传感器是陀螺仪的情况下,等式11是:

$$[0312] \quad Y(t) = \begin{cases} Y_{IMU}(t) \cdot R & = X(t) \cdot R \\ Y_{IMU}(t) \cdot \vec{\Omega} & = X(t) \cdot \vec{\Omega} \end{cases} \quad (13)$$

[0313] $Y_{IMU}(t) \cdot \{\alpha \in R, \Omega\}$ 是在时步 t 旋转(R ,单位为弧度)和角速度 $\vec{\Omega}$,单位为 $rad \cdot s^{-1}$) IMU的观察值。使用九自由度传感器实际上没有可计算的平移或速度。

[0314] f 的定义

$$[0315] \quad \hat{X}(t)^{\Delta t} = f(\hat{X}(t), \Delta t) \quad (14)$$

[0316] 其中 $\hat{X}(t)^{\Delta t}$ 是在时步 $t+\Delta t$ 预测状态矢量, $\hat{X}(t)$ 是在时步 t 估计的状态矢量。

[0317] 函数 f 是状态矢量的更新函数,定义为:

$$[0318] \quad \hat{X}(t)^{\Delta t} = \begin{cases} \hat{X}(t) \cdot T + \hat{X}(t) \cdot V \cdot \Delta t \\ \hat{X}(t) \cdot R \times \hat{X}(t) \cdot \Omega \cdot \Delta t \\ \hat{X}(t) \cdot \vec{V} \\ \hat{X}(t) \cdot \vec{\Omega} \end{cases} \quad (15)$$

[0319] 注意的是,卡尔曼滤波的普通更新函数不适用于我们的例子,因为旋转矩阵不可篡改且被写为线性形式 $A \cdot X+B$ 。

[0320] 预测步骤

[0321] 预测步骤使用卡尔曼滤波的定义写成:

$$[0322] \quad \hat{X}(t)^{\Delta t} = f(\hat{X}(t), \Delta t) \quad (16)$$

$$[0323] \quad P_t^{\Delta t} = A_t \cdot P_t \cdot A_t^T + Q \cdot \Delta t \quad (17)$$

[0324] 其中:

[0325] • $P_t^{\Delta t}$ 是在时步 $t+\Delta t$ 预测误差的协方差矩阵,

[0326] • P_t 是在时步 t 误差的协方差矩阵,

[0327] • A 是转移矩阵,例如 $A_t = \frac{\partial f}{\partial X_t} | \hat{X}_{t-1}$, A_t 是 f 的导数矩阵,计算为有限差值的总和: $\frac{f(x+\Delta) - f(x)}{\Delta}$,

[0328] • Q 是状态矢量的噪音的12x12协方差矩阵。平移状态矢量的噪音启发式估计为 $1e^{-2}$,以及旋转状态矢量的噪音为 $1 \cdot e^{-3}$ 。

$$[0329] \quad Q = \begin{bmatrix} \begin{array}{|c|c|c|c|} \hline 0.01 & 0.0 & 0.0 & 0.0 \\ \hline T & & & \\ \hline \end{array} & \begin{array}{|c|c|c|c|} \hline 0.0 & 0.001 & 0.0 & 0.0 \\ \hline R & & & \\ \hline \end{array} & \begin{array}{|c|c|c|c|} \hline 0.0 & 0.0 & 0.0 & 0.0 \\ \hline V & & & \\ \hline \end{array} & \begin{array}{|c|c|c|c|} \hline 0.0 & 0.0 & 0.0 & 0.0 \\ \hline \Omega & & & \\ \hline \end{array} \\ \hline \end{bmatrix}$$

[0330] 校正步骤

[0331] 以下是卡尔曼定义,在我们的例子中校正写成:

$$[0332] \quad K(t + \Delta t) = P_t^{\Delta t} \cdot H^T(t + \Delta t) \cdot (U(t + \Delta t) + H(t + \Delta t) \cdot P_t^{\Delta t} \cdot H^T(t + \Delta t))^{-1} \quad (18)$$

$$[0333] \quad P(t + \Delta t) = (I - K(t + \Delta t) \cdot H(t + \Delta t)) \cdot P_t^{\Delta t} \quad (19)$$

$$[0334] \quad \hat{X}(t + \Delta t) = \begin{cases} \hat{X}(t)^{\Delta t} \cdot T + K(t + \Delta t) \times J_{witness|IMU} \cdot T \\ \hat{X}(t)^{\Delta t} \cdot R + K(t + \Delta t) \times J_{witness|IMU} \cdot R \\ \hat{X}(t)^{\Delta t} \cdot V + K(t + \Delta t) \times J_{witness|IMU} \cdot V \\ \hat{X}(t)^{\Delta t} \cdot \Omega + K(t + \Delta t) \times J_{witness|IMU} \cdot \Omega \end{cases} \quad (20)$$

[0335] 其中:

[0336] • I是单位矩阵,

[0337] • K(t+Δt)在时步t+Δt的卡尔曼增益,

[0338] • H是函数h的导数矩阵,依据当前选定的传感器它是 $h_{witness}$ 或 h_{IMU}

$$[0339] \quad H(t + \Delta t) = \frac{\partial h}{\partial X} \Big|_{\hat{X}_t^{\Delta t}} \quad (21)$$

[0340] 在我们的例子中,雅克比行列式(jacobian)计算为有限差值的总和:

$$[0341] \quad \frac{h_{witness|IMU}(x + \Delta) - h_{witness|IMU}(x)}{\Delta}$$

[0342] • U(t+Δt)是当前选定的传感器(见证摄影机或IMU)的噪音测量结果的协方差矩阵。该值根据测试的质量(置信度)进行修改。在视觉追踪的情况下,它依赖于追踪误差ε:

$$[0343] \quad U(t + \Delta t) = \begin{cases} 0.0001 \text{ if } \varepsilon < 1.0 \text{ pixel} \\ 10.0 \text{ if } \varepsilon \geq 1.0 \text{ pixel} \end{cases} \quad (22)$$

[0344] 为IMU确定U(t+Δt)更简单,因为噪音误差由所有传感器的制造商直接给定。之前的等式直接由这些误差替换。

[0345] • P(t+Δt)是在时步t+Δt误差的协方差矩阵。

[0346] • V和Ω是分别是速度矢量 \vec{v} 和角速度矢量 $\vec{\omega}$ 的矩阵表示。

[0347] • 等式20中的 $J_{\text{witness}}|_{\text{IMU}}$ 也根据传感器的种类写为:

$$[0348] \quad J_{\text{witness}}.T = y(t + \Delta t).T - \hat{Y}_t^{\Delta t}.T$$

$$[0349] \quad J_{\text{witness}}.R = y(t + \Delta t).R \times \hat{Y}_t^{\Delta t}.R^{-1}$$

$$[0350] \quad J_{\text{witness}}.V = y(t + \Delta t).V - \hat{Y}_t^{\Delta t}.V$$

$$[0351] \quad J_{\text{witness}}.\Omega = y(t + \Delta t).\Omega \times \hat{Y}_t^{\Delta t}.\Omega$$

$$[0352] \quad J_{\text{IMU}}.R = y(t + \Delta t).R \times \hat{Y}_t^{\Delta t}.R^T$$

$$[0353] \quad J_{\text{IMU}}.\Omega = y(t + \Delta t).\Omega \times \hat{Y}_t^{\Delta t}.\Omega^{-1}$$

[0354] • $y(t + \Delta t)$ 是选定的传感器给出的当前测量结果。 $y(t + \Delta t) \cdot \{\alpha \in T, R, V, \Omega\}$ 分别是平移、旋转、速度和角速度的测量结果。 $J \cdot \{\alpha \in T, R, V, \Omega\}$ 是卡尔曼创新 (Kalman Innovation)。

[0355] 2.4. 三维重构和深度键控

[0356] 使用集成到我们摄影机杆中的三维距离传感器实现了实时三维重构。该距离传感器可以是结构光设备 (例如微软体感游戏, 华硕Xtion) 或飞行时间摄影机 (例如像先进科学理念 (Advanced Scientific Concept) 的虎眼石 (TigerEye)) 也被称为闪光激光雷达 (Flash Lidar)。这两种技术都有局限性, 我们的技术现在只用短距离传感器华硕Xtion。照常, 需要校准步骤从而知道华硕Xtion相对见证摄影机的位置。我们只需对影像对齐使用与见证摄影机完全相同的技术 (参见2.1.4.2)。当距离传感器具有灰度摄像机时, 我们使用我们通常的校准模式使它在两个摄影机中检测到: 距离传感器的摄影机和我们的见证摄影机之一。这个过程非常简单, 类似于2.1.4.2。

[0357] 在我们的技术中距离传感器用于两个目的。一方面, 它帮助删除例如远离对象后面创建的不良的特征点。特征点在距离传感器的深度映射上的简单的重新投影可以告诉我们特征点是否正确。在立体追踪期间 (参见2.5), 通过对极几何创建的每个点仍然可以近于或远于由核线给出的估计。我们比较重构三维点与它在距离传感器的深度映射上的重新投影的当前值, 以提高它的精度或仅仅拒绝它。

[0358] 距离传感器的第二个用途依赖于它本身的深度映射。当它给定了最大范围为12米的深度时, 我们能够通过虚拟人物/对象将布景中的每个人 (演员) 遮挡在设定的范围内, 反之亦然。这叫做动态遮挡, 在电影和实况转播中提高了视觉效果的另一真实性水平。我们的技术采用的主要电影摄影机供给 (这也可以通过增加附加的高分辨率见证摄影机实现) 通过在电影摄影机上重新投影深度和完善结果 (边缘增强) 增强距离传感器的分辨率。这种技术能够提供实时的深度键控, 从而消除了约束绿屏 (binding green screen) 的使用特别是当使用功能更强大的距离传感器 (闪光激光雷达)。

[0359] 2.5 立体无标记技术

[0360] 前面描述的使用单视场方法的技术在真实情景中其实并不够高效。因为它需要已知的真实对象用作注册来计算世界的大小, 方向和位置, 将它用在电影上常常是很难的, 其中当你尝试注册 (检测) 对象时许多人总是你的路径上。但该单视场方法的主要局限性是它需要测量场景的事实。由于单应性必须被用于该系统的初始化, 它意味着技术人员必须来到布景开始测量它。在实践中, 该操作很少取得成功, 因为在现场工作的人们在最后几秒之

前不会离开这个地方。

[0361] 我们已建立立体显示方法以摆脱该局限性,以及因此带来即时测量和追踪的完全原始概念。

[0362] 2.5.1即时测量

[0363] 嵌入摄像机追踪的立体视场技术与任何其他技术相比是重大的进步,包括我们前面描述的单视场方法。事实上,使用两台预校准的见证摄影机带来了一些好处。第一个是即时测量。不像在单视场方法描述的技术,不需要预先测量场景。由于两个见证摄影机之间的精确距离是完全已知的,因为他们使用内部同步锁定硬件进行同步,我们能够不用移动杆立即得到点云。然后,在段落中所述的单应性不再使用,由于点的三维重建是由两个见证摄影机通过立体显示理论带来的。第二个主要问题是,不需要手动缩放。在这两个视频流拍摄期间,我们运行莱文贝格-马夸特最小化方案,以尽量减少由对极几何估计的三维点的投影和在当前视频中检测到的二维特征点之间像素误差的总和。

[0364] 2.5.2自动注册

[0365] 然而,使用即时测量增加了新的问题,其使用在单视场方法中检测到的注册对象时不可能的,这是使用如在单视场的方式检测到的一个注册对象是不可能的。这一局限性的主要原因来自照相机杆直接连接到电影摄影机的事实,且请求摄影组到处移动他们的摄影机以获悉注册对象是不可能的。这就是为什么我们添加了第二个自动注册,其不需要任何物理注册对象被添加到场景中。然而,它要求用户与点云具有两个不同的交互:点选择(其用于告诉该系统,我们打算用作地面参考的点),以及绕Y(垂直)轴的手动旋转(其用于相对场景正确定向CG内容)。

[0366] 为了能够只具有这两个非常快和非常简单的用户交互,那么我们必须计算所有其他剩余的自由度,这意味着计算具有Y轴完全垂直和正确缩放比例的标准正交基。我们的技术只需在杆中利用嵌入式加速度计以计算重力方向,即Y轴。磁力计给出正北方向,从而提供了一个近似矢量,其允许我们使用普通向量积计算正交基。

[0367] 缩放比例甚至更容易计算。它是完全即时和自动的,不需要任何用户交互。由于精确已知见证摄影机之间的距离,所有的三维重构点直接拥有正确的缩放比例,因此不像在普通纯运动结构算法中那样要检测附加对象。

[0368] 2.5.3即时跟踪

[0369] 跟踪过程与2.1.4.3中描述的单视场方法的略有不同。每个新的点使用对极几何创建,其确保三维特征点的比例和深度总是正确的。这意味着,我们的系统一启动,由于光速准确和即时测量它已经开始跟踪。

[0370] 我们计算所有从左和右摄像机看到的重新投影的三维点的并集,以确保跟踪过程的最大稳定性。仅仅由与特征匹配的ORB描述符结合的核线的三维交集产生点云(参见2.1.4.4)。特征匹配和全体最小化方案仍然类似于在单视场方法中所使用的。

[0371] 为了能够依靠即时测量,我们还高度修改了在点云中建立新的点的方式。除了点的立体重构之外,我们增加了网格离差(grid dispersion)和特征点均匀性的概念。在普通追踪算法中,启动二维特征检测算法,并在三维中重构点,如2.1.3.3和2.1.4.4所述。然而,这些点中的许多是不可靠的(错误的深度,边缘点,移动点,阴影点等),并且该算法趋向于在高对比度区域获悉很多点和在低对比度区域没有点。这会在追踪数据中产生大量的噪

音,给出摄影机抖动的感觉。我们创建的网格离差技术以几乎均匀的方式设法分散系统获悉的自然标记。我们首先通过使用非常高的临界值开始产生特征点的最大值。然后,我们根据用户给定的临界值(例如 16×16 网格,每个单元两个样品)使用哈里斯评分(Harris score)(例如我们选择两个最好的点)减小特征点的数目。因此,每一个点具有用于评价它多可靠的质量标准。如果均匀性不是足够好,该系统也局部改变对比临界值,以确保能够甚至在低对比度区域捕捉到点。我们的算法在整个图像上分散特征点的事实也防止系统轻易丢失跟踪。事实上,像我们的没有网格离差技术的普通单视场方法可能有很多点,例如用于追踪图片的下左部。如果演员进入被隐藏所有这些天然标记的地方,跟踪就会立即失去,由于我们在整个画面上分散特征,这在我们的发明中是不会发生的,。

[0372] 2.6增强现实世界

[0373] 2.6.1摄影机/见证延迟问题

[0374] 由于电影摄影机和见证摄影机以不同的速度运行,它们在画面获取过程可能产生不同的延迟。电影摄影机实际上是创建迄今为止最高的延迟,迫使我们计算它以便适当地增强现实世界。

[0375] 我们再次使用我们自己的校准设备自动计算两个摄像机之间的延迟。在两个摄影机前面使用了以1Hz闪烁的小的高亮度发光二极管。已经开发出由强度阈值技术增强的二进制大对象检测器算法以检测这两个摄像机中的发光二极管。从摄影机中的发光二极管强度获取生成的相关强度曲线使用最小化算法进行了比较,以便找到样本之间的最佳时间一致性。因此最小化几个统计参数,并从在10秒左右的校准中自动提取以毫秒为单位的延迟。

[0376] 2.6.2呈现和增强现实

[0377] 一旦我们让摄影机跟踪工作,我们的软件可以从不同的CG内容生成软件(如实时动画软件(Autodesk MotionBuilder))接收数据。那些数据直接在我们的软件内整合和呈现(在追踪和呈现工作站合并的情况下,参见图1),以与真实背景或前景(如果是绿色级)进行复合。

[0378] 我们的技术固有的另一个有趣的特征是三维消除抠像。消除抠像肯定不是我们的发明,但它在我们的发明中以一种创新的方式进行。事实上,由于我们的追踪算法生成精确的三维点云(由于立体视场的方式),当被重新投影到电影摄影机视频上时所有这些点在三维和二维上呈现。我们已经开发出一种简单的用户界面,其允许用户在三维点云中选择多个点。因此,我们通过对所有选定的点进行三角测量使用听觉三角测量算法(Ear Triangulation algorithm) [14]创建三维对象。最近创建的表面现在可以被激活作为消除抠像对象,这意味着在电影摄影机视频上那个对象的二维投影覆盖的每个像素现在被三维内容生成工作站生成的三维虚拟环境替换(参见1)。需要注意的是在使用远程深度传感器(即至少50米)的情况下消除抠像技术变得无效,由于我们的深度键控方法很容易替换它(参见2.4)。

[0379] 附录1

[0380] 参考文献

[0381] [1]赫伯特·贝,廷尼图特拉,吕克凡古·瑟夫.加速稳健性特征.第9届欧洲计算机视觉国际会议论文集[C].柏林,海德堡,施普林格出版社.2006(6):404-417。(Herbert Bay,Tinne Tuytelaars,and Luc Van Gool.Surf:speeded up robust features.In

Proceedings of the 9th European conference on Computer Vision-Volume Part I, ECCV'06, pages 404-417, Berlin, Heidelberg, 2006. Springer-Verlag)

[0382] [2] 迈克尔·卡隆德,文森特·利普特,克里斯托夫·斯察,帕斯卡·菲阿布里夫. 独立二进制稳健性的基本特点[C]. 第11届欧洲计算机视觉国际会议论文集. 柏林,海德堡,施普林格出版社. 2010(4):778-792. (Michael Calonder, Vincent Lepetit, Christoph Strecha, and Pascal Fua. Brief: binary robust independent elementary features. In Proceedings of the 11th European conference on Computer vision: Part IV, ECCV'10, pages 778-792, Berlin, Heidelberg, 2010. Springer-Verlag.)

[0383] [3] 弗雷德里克·达沃尼,奥维尔·法戈瑞斯. 直线必须是直的[C]. 国际光学工程学会, 2001(2567). (Frederic Devernay and Olivier Faugeras. Straight lines have to be straight. In In SPIE, volume 2567, 2001.)

[0384] [4] 马丁A·费什勒,罗伯特C·博尔斯. 随机样本一致性一个应用到图像分析和自动化制图的范式模型[C]. 美国计算机协会通讯, 1981, 24(6):381-395. (Martin A. Fischler and Robert C. Bolles. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. Communications of the ACM, 24(6):381-395, 1981.)

[0385] [5] 克里斯·哈里斯,麦克·斯蒂芬斯. 一种组合的角落和边缘检测器[C]. 第四阿尔维愿景会议. 1988:147-151. (Chris Harris and Mike Stephens. A combined corner and edge detector. In In Proc. of Fourth Alvey Vision Conference, pages 147-151, 1988.)

[0386] [6] D.C. 霍格林, F. 莫斯特勒, J.W. 图基. 解读鲁棒探索性数据分析, 威力系列概率与数理统计: 应用概率统计[M]. 约翰·威力与桑恩. 1983. (D.C. Hoaglin, F. Mosteller, and J.W. Tukey. Understanding Robust and Exploratory Data Analysis. Wiley series in probability and mathematical statistics: applied probability and statistics. John Wiley and Sons, 1983.)

[0387] [7] 鲁道夫·埃米尔·卡尔曼. 一种新的以线性滤波和预测的问题的方法[J]. 美国机械工程师学会基础工程出版社. 1960, 82(D):35-45. (Rudolph Emil Kalman. A new approach to linear filtering and prediction problems. Transactions of the ASME-Journal of Basic Engineering, 82(Series D):35-45, 1960.)

[0388] [8] 格奥尔格·克莱恩, 大卫·莫瑞. 工作区并行跟踪和映射小型增强现实技术[C]. 美国华盛顿, 第六届电气和电子工程师协会和美国计算机协会国际研讨会上混合和增强现实程序. 美国计算机协会, 计算机社会. 2007, (7):1-10. (Georg Klein and David Murray. Parallel tracking and mapping for small ar workspaces. In Proceedings of the 2007 6th IEEE and ACM International Symposium on Mixed and Augmented Reality, ISMAR'07, pages 1-10, Washington, DC, USA, 2007. IEEE Computer Society.)

[0389] [9] J.J. 伦纳德, H. 朗迪·怀特. 基于追踪几何学的机器人追踪[J]. 电气和电子工程师协会-机器人与自动化. 1991, 7(3). (J.J. Leonard and Durrant H. Whyte. Mobile robot localization by tracking geometric beacons. IEEE Transactions on Robotics and Automation, 7(3), 1991.)

- [0390] [10]K.利文贝格.一种用于一些问题在最小二乘的解决方案[J].夸脱相联平行处理机数学.1944,(2):164-168。(K.Levenberg.A method for the solution of certain problems in least squares.Quart App-1.Math,2:164-168,1944.)
- [0391] [11]G.戴维·劳氏.局部尺度不变特征的目标识别[C].第二届计算机视觉会议论文集.1999,2(99):1150-。(David G.Lowe.Object recognition from local scale-invariant features.In Proceedings of the International Conference on Computer Vision-Volume 2-Volume 2,ICCV'99,pages1150-,Washington,DC,USA,1999.IEEE Computer Society.)
- [0392] [12]D.麦夸特.非线性参数的最小二乘估计的一种算法.美国工业与应用数学学会.1963,(11):431-441。(D.Marquardt.An algorithm for least-squares estimation of nonlinear parameters.SIAM J.Appl.Math,11:431-441,1963.)
- [0393] [13]克里斯多夫·梅,加布·西不利,马克·康明斯,保罗·纽曼,兰·瑞德.瑞斯姆:一种使用大比例尺立体声固定时间的系统[J].国际应用数学和计算机科学.2011.9,94(2):198-214。(Christopher Mei,Gabe Sibley,Mark Cummins,Paul Newman,and Ian Reid.Rslam:A system for large-scale mapping in constant-time using stereo.Int.J.Comput.Vision,94(2):198-214,September 2011.)
- [0394] [14]加里·霍西尔·梅斯特.多边形有耳[J].美国数学月刊.1975,(82):648-651。(Gary Hosier Meisters.Polygons have ears.Amer.Math.Monthly,82:648-651,1975.)
- [0395] [15]J.J.摩尔.数学运算讲稿之理论与实践:利文贝格-马奎特测定法[M].施普林格出版社.1977。(J.J.More'.The Levenberg-Marquardt algorithm:implementation and theory,pages 105-116.Lecture Notes in Mathematics 630.SPRINGER,1977.)
- [0396] [16]爱德华·罗斯滕,汤姆·缀蒙德.高速角点检测的机器学习能力[C].柏林,海德堡,第9届欧洲会议计算机视觉国际会议论文集.2006,(1):430-443。(Edward Rosten and Tom Drummond.Machine learning for high-speed corner detection.In Proceedings of the 9th European conference on Computer Vision-Volume Part I,ECCV'06,pages 430-443,Berlin,Heidelberg,2006.Springer-Verlag.)
- [0397] [17]爱德华·罗斯滕,里德·波特,汤姆·缀蒙德.一种角点检测的机器学习方法[J].电气与电子工程师协会.模式分析与机器智能.2010,(32):105-119。(Edward Rosten, Reid Porter,and Tom Drummond.Faster and better:A machine learning approach to corner detection.IEEE Trans.Pattern Analysis and Machine Intelligence,32:105-119,2010.)
- [0398] [18]伊桑·路布里,文森特·拉波德,库尔特·康农里格,盖理·布拉德斯基.Orb:一种可选择的筛选和浏览算法[C].美国,华盛顿,国际计算机视觉国际会议论文集.2011,(11):2564-2571。(Ethan Rublee.Vincent Rabaud,Kurt Konolige,and Gary Bradski.Orb:An efficient alternative to sift or surf.In Proceedings of the 2011International Conference on Computer Vision,ICCV'11,pages 2564-2571,Washington,DC,USA,2011.IEEE Computer Society.)
- [0399] [19]伊凡.E.萨瑟兰.一种头戴式三维显示器[C].秋季联合计算机会议12月9-11

日论文集.美国,纽约,1968,68(1):757-764。(Ivan E.Sutherland.A head-mounted three dimensional display.In Proceedings of the December 9-11,1968,fall joint computer conference,part I,AFIPS'68(Fall,part I),pages 757-764,New York,NY,USA,1968.ACM.)

[0400] [20] 罗杰·Y·蔡,赖默尔·K·伦茨.一种新的完全自主、高效的3D机器人手眼校准技术[C].第四届机器人研究国际会议论文集.美国,剑桥,马萨诸塞,1988:287-297。(Roger Y.Tsai and Reimer K.Lenz.A new technique for fully autonomous and efficient 3d robotics hand-eye calibration.In Proceedings of the 4th international symposium on Robotics Research,pages 287-297,Cambridge,MA,USA,1988.MIT Press.)

[0401] [21] J.K. 乌尔曼.多目标追踪算法[J].美国科学家.1992,80(2):128-141。(J.K.Uhlmann.Algorithms for multiple target tracking.American Scientist,80(2):128-141,1992.)

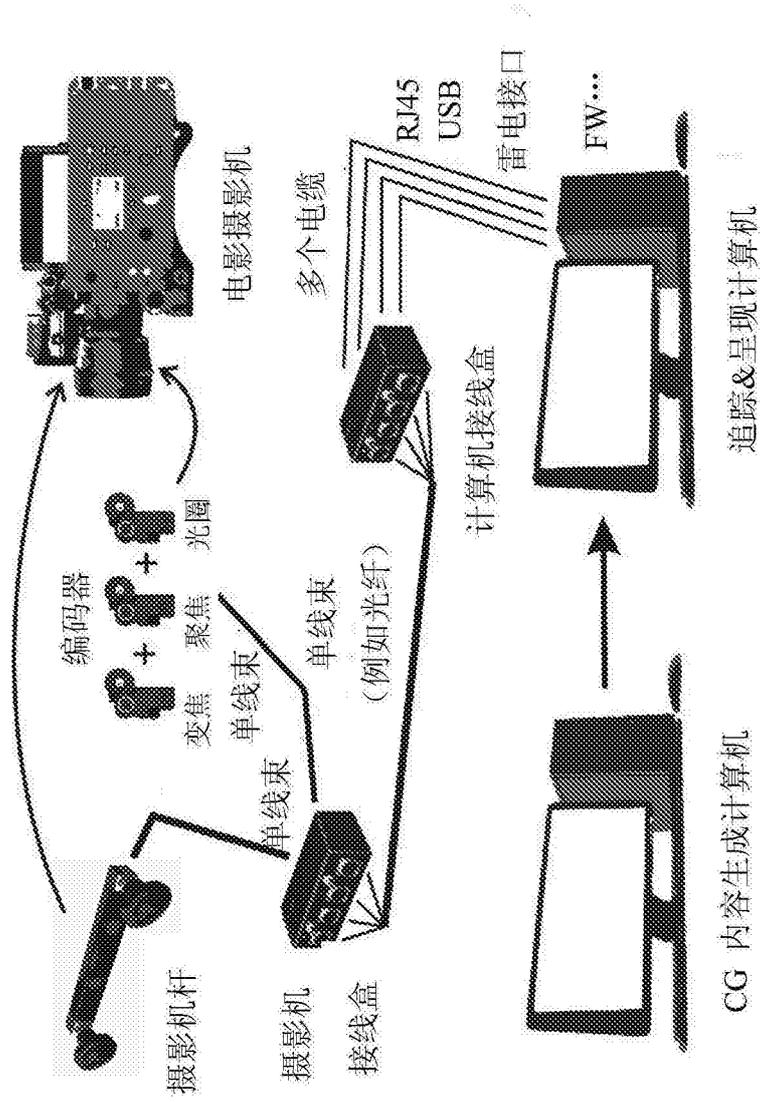


图1

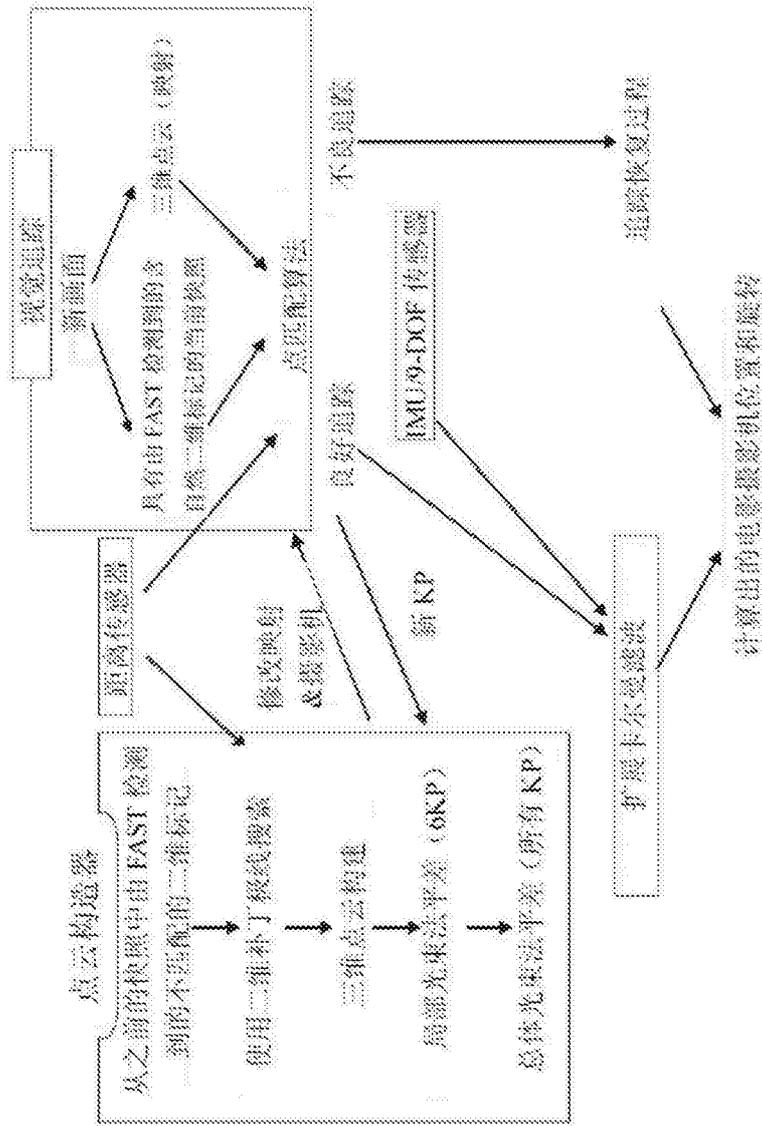


图2

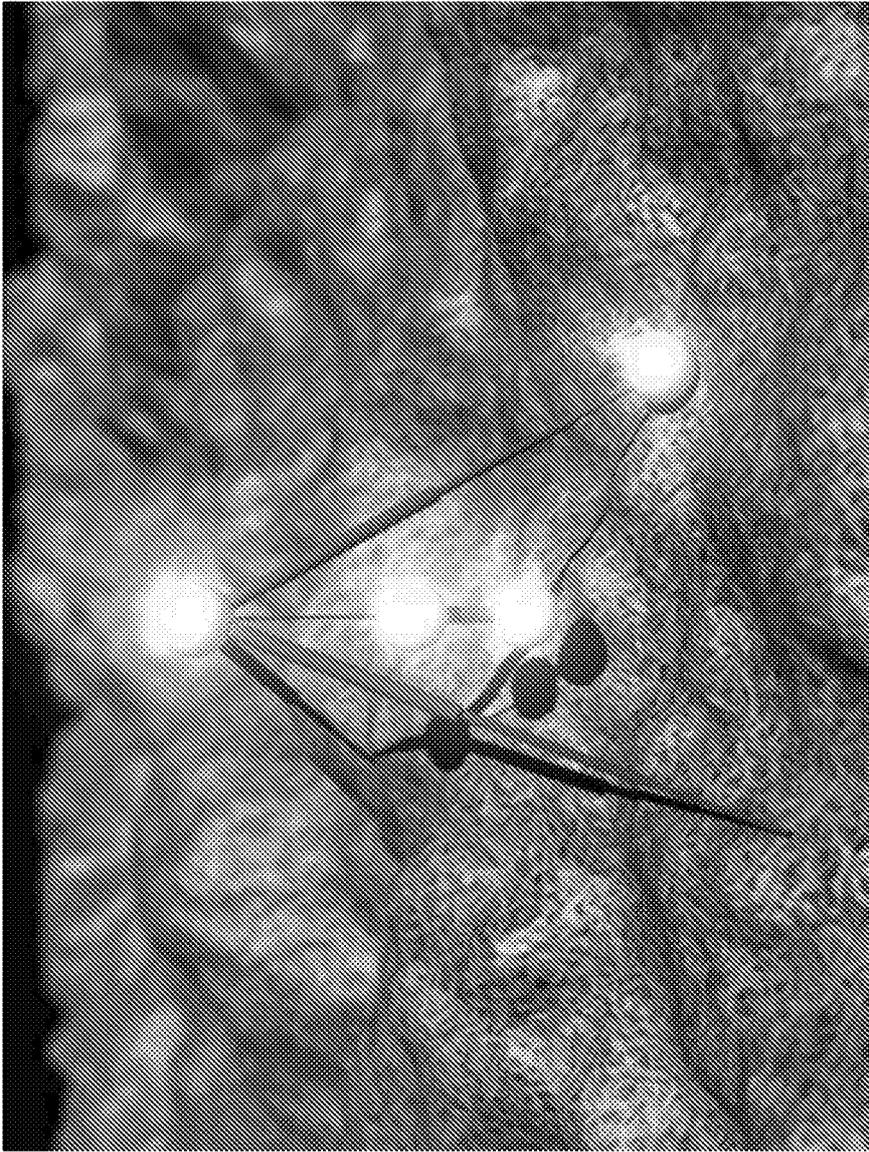


图3

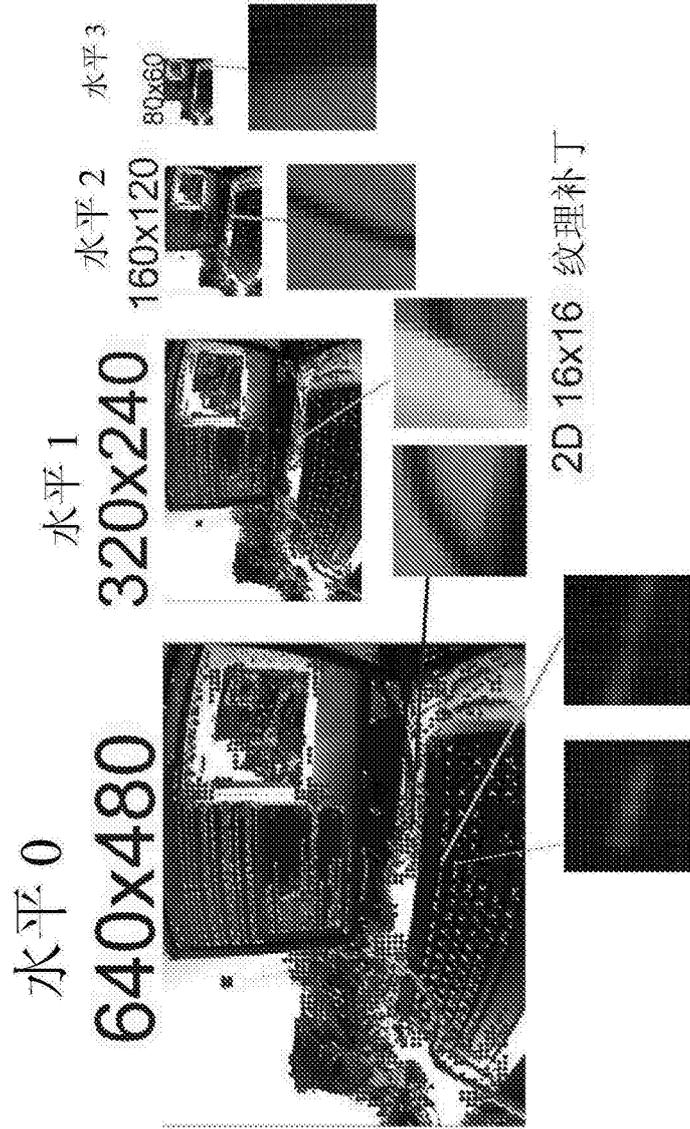


图4

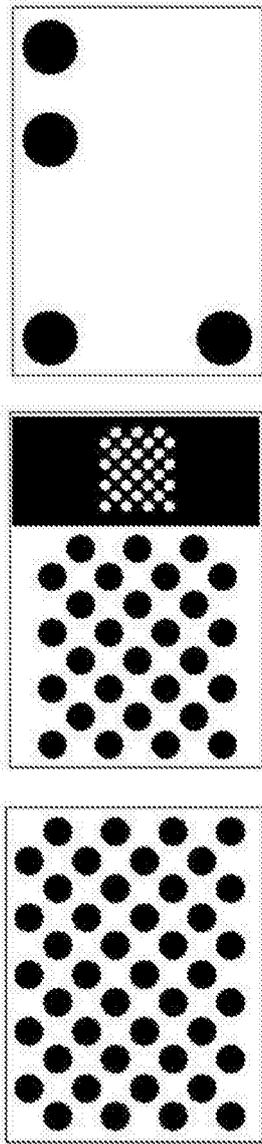


图5

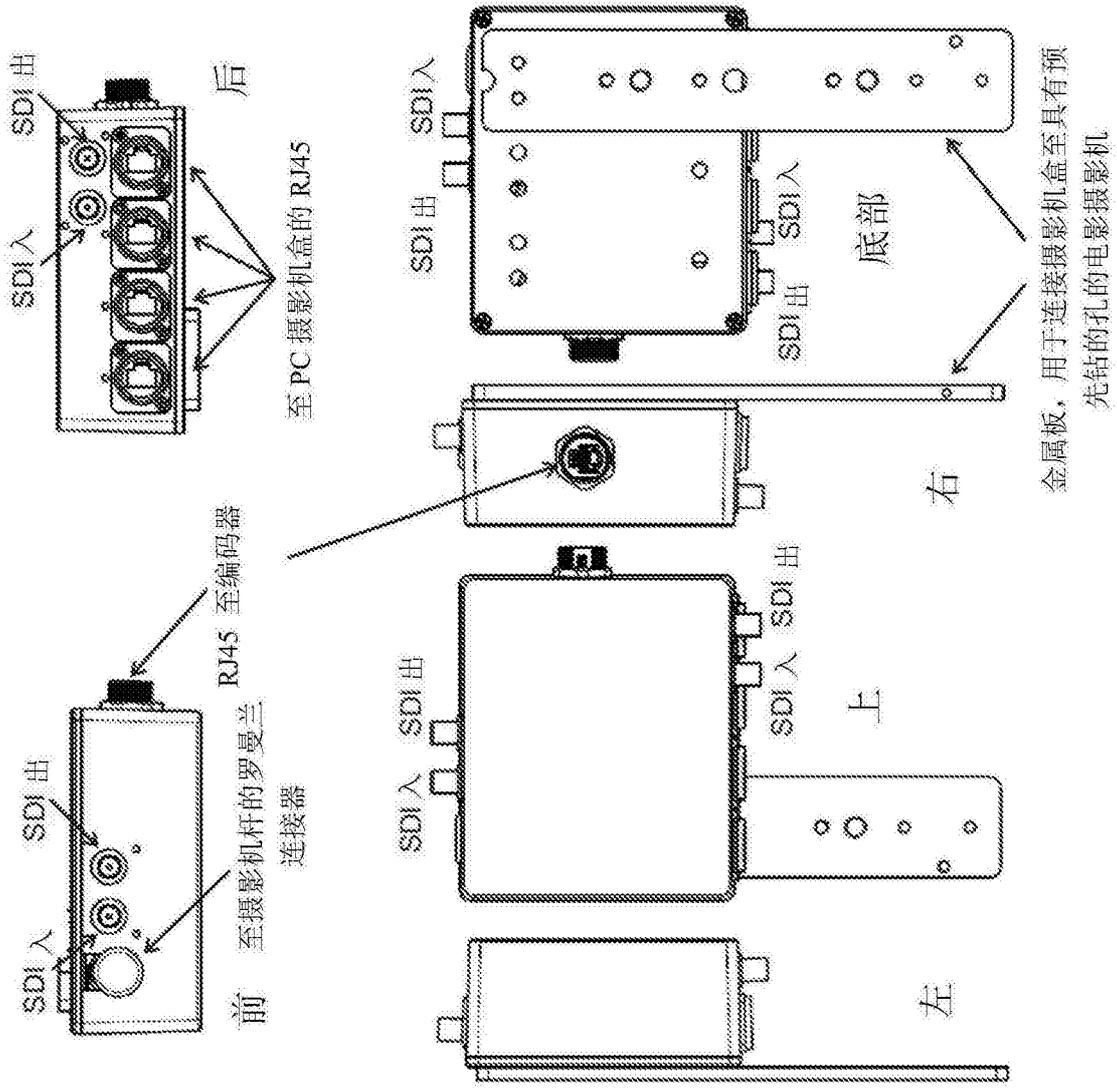


图8