



(12)发明专利

(10)授权公告号 CN 105164728 B

(45)授权公告日 2018.02.06

(21)申请号 201480024088.9

(51)Int.Cl.

(22)申请日 2014.03.26

G06K 9/00(2006.01)

(65)同一申请的已公布的文献号

G06T 7/26(2017.01)

申请公布号 CN 105164728 A

G06T 19/00(2011.01)

(43)申请公布日 2015.12.16

G06T 15/00(2011.01)

(30)优先权数据

61/817,792 2013.04.30 US

(56)对比文件

14/225,157 2014.03.25 US

谭雪婷.移动增强现实中的跟踪技术研究.

(85)PCT国际申请进入国家阶段日

《中国优秀硕士学位论文全文数据库 信息科技辑》.2010,第2010年卷(第12期),

2015.10.28

邹承明等.基于SLAM的虚实配准方法研究.

《计算机与数字工程》.2010,第38卷(第1期),

(86)PCT国际申请的申请数据

ZOKAI S等.Multiview Paraperspective Projection Model for Diminished Reality.

PCT/US2014/031926 2014.03.26

《PROCEEDINGS/THE SECOND IEEE AND ACM INTERNATIONAL SYMPOSIUM ON MIXED AND AUGMENTED REALITY:7-10 OCTOBER 2003,

(87)PCT国际申请的公布数据

TOKYO, JAPAN, IEEE COMPUTER SOCIETY, LOS ALAMITOS, CALIF.[U.A.]》.2003,

W02014/178966 EN 2014.11.06

(续)

(73)专利权人 高通股份有限公司

ALAMITOS, CALIF.[U.A.]》.2003,

地址 美国加利福尼亚州

(72)发明人 迪特尔·施马尔施蒂格

审查员 宿渊源

(74)专利代理机构 北京律盟知识产权代理有限

权利要求书4页 说明书16页 附图7页

责任公司 11287

代理人 宋献涛

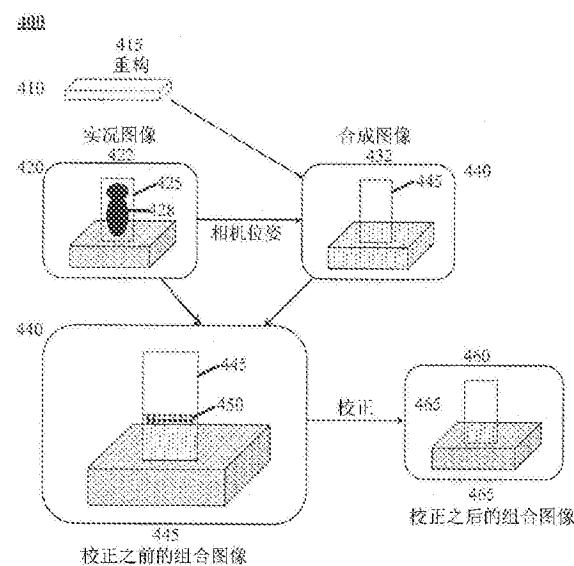
(54)发明名称

用于混合实境的设备和方法

(57)摘要

所揭示的实施例是关于用于混合实境的设备、系统及方法。在一些实施例中，可确定相对于实况图像中的被跟踪对象的相机位姿且使用所述相机位姿从3D模型中的关键帧再现不具有所述被跟踪对象的合成图像。可确定第一掩模区中的像素相对于所述合成图像的子集的光流量值，且可使用所述光流量值确定所述合成图像子集中的每一者中的对应于所述第一掩模中的像素的像素。对于所述第一掩模中的每一像素，可将对应的替换像素确定为随所述合成图像子集中对应于所述第一掩模中的所述对应像素的像素而变。

CN 105164728 B



[转续页]

[接上页]

(56)对比文件

PRAVIN BAT等.Using Photographs to Enhance Videos of a Static Scene.
《RENDERING TECHNIQUES 2007:18TH EUROGRAPHICS SYMPOSIUMON RENDERING》.2007,
Yasuyuki Matsushita等.Full-frame

Video Stabilization.《PROCEEDINGS/2005 IEEE COMPUTER SOCIETY CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION, CVPR 2005:[20-25 JUNE 2005,SAN DIEGO,CA], IEEE,PISCATAWAY,NJ,USA》.2005,第1卷

1. 一种用于混合实境的方法,其包括:

获得多个实况图像,所述实况图像包括被跟踪对象;

针对所述多个实况图像中的一实况图像,确定相对于所述被跟踪对象的相机位姿;

基于所述相机位姿而再现不具有所述被跟踪对象的多个合成图像,每一合成图像是从3D模型中的相异关键帧获得;

比较所述实况及合成图像以确定与所述被跟踪对象相关联的第一掩模区;

计算与所述被跟踪对象相关联的第一掩模区中的多个像素相对于所述多个合成图像中的合成图像子集的光流量值;

针对所述第一掩模中的所述多个像素中的每一者,部分基于所述光流量值确定所述合成图像子集中的每一者中的对应像素;及

确定所述第一掩模中的所述多个像素中的每一者的替换像素,每一替换像素对应于所述第一掩模中的相异像素,其中将每一替换像素确定为随所述合成图像子集中对应于所述第一掩模中的像素的所述对应像素而变。

2. 根据权利要求1所述的方法,其进一步包括再现减弱实境图像,所述减弱实境图像是至少部分基于所述替换像素而获得。

3. 根据权利要求2所述的方法,其进一步包括在所述图像中由所述第一掩模指示的区中再现虚拟对象,所述虚拟对象是基于当前相机位姿而再现。

4. 根据权利要求3所述的方法,其中在所述图像中由第一掩模指示的所述区中再现所述虚拟对象包括:

将所述图像中的至少一个已知实际对象的虚拟表示再现到虚拟z缓冲器中,所述实际对象至少部分在所述第一掩模中;

分辨所述至少一个已知实际对象与所述虚拟z缓冲器中的所述虚拟对象之间的遮挡;及

扭曲所述虚拟z缓冲器以匹配所述图像,所述扭曲部分基于所述至少一个已知实际对象的所述虚拟表示与所述图像中的所述至少一个实际对象之间的光流。

5. 根据权利要求1所述的方法,其中确定所述合成图像子集中的每一者中的对应像素包括:

基于所述光流获得所述第一掩模中的所述多个像素中的每一者的位移及强度校正。

6. 根据权利要求1所述的方法,其中:

将每一替换像素确定为所述合成图像子集中的对应像素的加权平均,其中与所述子集中的合成图像中的对应像素相关联的权重部分基于所述对应像素的所述光流量值及所述实况图像与所述合成图像之间的与所述对应像素相关联的角度位移。

7. 根据权利要求6所述的方法,其中所述第一掩模M'区包括围绕所述被跟踪对象的第二掩模区M,且其中通过从边界M'-M进行分散数据内插而确定M中的每一替换像素的所述光流量值。

8. 根据权利要求1所述的方法,其中所述多个合成图像及所述多个合成图像的所述子集各自包括三个合成图像。

9. 根据权利要求1所述的方法,其中使用同时定位与地图创建SLAM跟踪技术确定所述相机位姿。

10. 一种移动台MS, 其包括:

相机, 其经配置以俘获包括被跟踪对象的多个实况图像,

存储器, 所述存储器用以存储包括多个关键帧的3D模型, 及

处理器, 其耦合到所述相机及所述存储器, 所述处理器经配置以:

针对所述多个图像中的一实况图像, 确定相对于所述被跟踪对象的相机位姿;

基于所述相机位姿而再现不具有所述被跟踪对象的多个合成图像, 每一合成图像是从相异关键帧获得;

比较所述实况及合成图像以确定与所述被跟踪对象相关联的第一掩模区;

计算与所述被跟踪对象相关联的第一掩模区中的多个像素相对于所述多个合成图像中的合成图像子集的光流量值;

针对所述第一掩模中的所述多个像素中的每一者, 部分基于所述光流量值确定所述合成图像子集中的每一者中的对应像素; 及

确定所述第一掩模中的所述多个像素中的每一者的替换像素, 每一替换像素对应于所述第一掩模中的相异像素, 其中将每一替换像素确定为随所述合成图像子集中对应于所述第一掩模中的所述像素的所述对应像素而变。

11. 根据权利要求10所述的MS, 其中所述处理器进一步经配置以:

再现减弱实境图像, 所述减弱实境图像是至少部分基于所述替换像素而获得。

12. 根据权利要求11所述的MS, 其中所述处理器进一步经配置以:

在所述图像中由所述第一掩模指示的区中再现虚拟对象, 所述虚拟对象是基于当前相机位姿而再现。

13. 根据权利要求12所述的MS, 其中为了在所述图像中由第一掩模指示的所述区中再现所述虚拟对象, 所述处理器经配置以:

将所述图像中的至少一个已知实际对象的虚拟表示再现到虚拟z缓冲器中, 所述实际对象至少部分在所述第一掩模中;

分辨所述至少一个已知实际对象与所述虚拟z缓冲器中的所述虚拟对象之间的遮挡; 及

扭曲所述虚拟z缓冲器以匹配所述图像, 所述扭曲部分基于所述至少一个已知实际对象的所述虚拟表示与所述图像中的所述至少一个实际对象之间的光流。

14. 根据权利要求10所述的MS, 其中为了确定所述合成图像子集中的每一者中的对应像素, 所述处理器经配置以:

基于所述光流获得所述第一掩模中的所述多个像素中的每一者的位移及强度校正。

15. 根据权利要求10所述的MS, 其中:

将每一替换像素确定为所述合成图像子集中的对应像素的加权平均, 其中与所述子集中的合成图像中的对应像素相关联的权重部分基于所述对应像素的所述光流量值及所述实况图像与所述合成图像之间的与所述对应像素相关联的角度位移。

16. 根据权利要求15所述的MS, 其中:

所述第一掩模M'区包括围绕所述被跟踪对象的第二掩模区M, 且其中通过从边界M'-M进行分散数据内插而确定M中的每一替换像素的所述光流量值。

17. 根据权利要求10所述的MS, 其中所述多个合成图像及所述多个合成图像的所述子

集各自包括三个合成图像。

18. 根据权利要求10所述的MS，其中所述处理器经配置以使用同时定位与地图创建SLAM跟踪技术确定所述相机位姿。

19. 一种用于混合实境的设备，其包括：

成像装置，所述成像装置用以俘获包括被跟踪对象的多个图像，

存储装置，所述存储装置用以存储包括多个关键帧的3D模型，及

处理装置，所述处理装置耦合到所述成像装置及所述存储装置，所述处理装置进一步包括：

用于针对所述多个实况图像中的一实况图像而确定相对于所述被跟踪对象的成像装置位姿的装置；

用于基于所述成像装置位姿而再现不具有所述被跟踪对象的多个合成图像的装置，每一合成图像是从相异关键帧获得；

用于比较所述实况及合成图像以确定与所述被跟踪对象相关联的第一掩模区的装置；

用于计算与所述被跟踪对象相关联的第一掩模区中的多个像素相对于所述多个合成图像中的合成图像子集的光流量值的装置；

用于针对所述第一掩模中的所述多个像素中的每一者部分基于所述光流量值确定所述合成图像子集中的每一者中的对应像素的装置；及

用于确定所述第一掩模中的所述多个像素中的每一者的替换像素的装置，每一替换像素对应于所述第一掩模中的相异像素，其中将每一替换像素确定为随所述合成图像子集中对应于所述第一掩模中的所述像素的所述对应像素而变。

20. 根据权利要求19所述的设备，其中所述处理装置进一步包括：

用于再现减弱实境图像的装置，所述减弱实境图像是至少部分基于所述替换像素而获得。

21. 根据权利要求20所述的设备，其中所述处理装置进一步包括：

用于在所述图像中由所述第一掩模指示的区中再现虚拟对象的装置，所述虚拟对象是基于当前相机位姿而再现。

22. 根据权利要求21所述的设备，其中用于在所述图像中由第一掩模指示的所述区中再现所述虚拟对象的装置包括：

用于将所述图像中的至少一个已知实际对象的虚拟表示再现到虚拟z缓冲器中的装置，所述实际对象至少部分在所述第一掩模中；

用于分辨所述至少一个已知实际对象与所述虚拟z缓冲器中的所述虚拟对象之间的遮挡的装置；及

用于扭曲所述虚拟z缓冲器以匹配所述图像的装置，所述扭曲部分基于所述至少一个已知实际对象的所述虚拟表示与所述图像中的所述至少一个实际对象之间的光流。

23. 根据权利要求19所述的设备，其中：

将每一替换像素确定为所述合成图像子集中的对应像素的加权平均，其中与所述子集中的合成图像中的对应像素相关联的权重部分基于所述对应像素的所述光流量值及所述实况图像与所述合成图像之间的与所述对应像素相关联的角度位移。

24. 一种非暂时性计算机可读媒体，其包括指令，所述指令在由处理器执行时执行方法

中的步骤,所述步骤包括:

获得多个实况图像,所述实况图像包括被跟踪对象;

针对所述多个实况图像中的一实况图像,确定相对于所述被跟踪对象的相机位姿;

基于所述相机位姿而再现不具有所述被跟踪对象的多个合成图像,每一合成图像是从基于所述相机位姿的3D模型中的相异关键帧获得;

比较所述实况及合成图像以确定与所述被跟踪对象相关联的第一掩模区;

计算与所述被跟踪对象相关联的第一掩模区中的多个像素相对于所述多个合成图像中的合成图像子集的光流量值;

针对所述第一掩模中的所述多个像素中的每一者,部分基于所述光流量值确定所述合成图像子集中的每一者中的对应像素;及

确定所述第一掩模中的所述多个像素中的每一者的替换像素,每一替换像素对应于所述第一掩模中的相异像素,其中将每一替换像素确定为随所述合成图像子集中对应于所述第一掩模中的所述像素的所述对应像素而变。

25.根据权利要求24所述的计算机可读媒体,其进一步包括再现减弱实境图像,所述减弱实境图像是至少部分基于所述替换像素而获得。

26.根据权利要求25所述的计算机可读媒体,其进一步包括在所述图像中由所述第一掩模指示的区中再现虚拟对象,所述虚拟对象是基于当前相机位姿而再现。

27.根据权利要求26所述的计算机可读媒体,其中在所述图像中由第一掩模指示的所述区中再现所述虚拟对象包括:

将所述图像中的至少一个已知实际对象的虚拟表示再现到虚拟z缓冲器中,所述实际对象至少部分在所述第一掩模中;

分辨所述至少一个已知实际对象与所述虚拟z缓冲器中的所述虚拟对象之间的遮挡;及

扭曲所述虚拟z缓冲器以匹配所述图像,所述扭曲部分基于所述至少一个已知实际对象的所述虚拟表示与所述图像中的所述至少一个实际对象之间的光流。

28.根据权利要求24所述的计算机可读媒体,其中确定所述合成图像子集中的每一者中的对应像素包括:

基于所述光流获得所述第一掩模中的所述多个像素中的每一者的位移及强度校正。

29.根据权利要求24所述的计算机可读媒体,其中:

将每一替换像素确定为所述合成图像子集中的对应像素的加权平均,其中与所述子集中的合成图像中的对应像素相关联的权重部分基于所述对应像素的所述光流量值及所述实况图像与所述合成图像之间的与所述对应像素相关联的角度位移。

30.根据权利要求29所述的计算机可读媒体,其中所述第一掩模M'区包括围绕所述被跟踪对象的第二掩模区M,且其中通过从边界M'-M进行分散数据内插而确定M中的每一替换像素的所述光流量值。

用于混合实境的设备和方法

[0001] 相关申请案的交叉参考

[0002] 本申请案主张2014年3月25日申请的题为“来自重构的减弱及中介实境效果 (Diminished and Mediated Reality Effects from Reconstruction)”的第14/225,157号美国申请案的权益及优先权,所述美国申请案又主张2013年4月30日申请的题为“来自重构的减弱及中介实境效果 (Diminished and Mediated Reality Effects from Reconstruction)”的第61/817,792号美国临时申请案的权益及优先权,所述两个申请案以其全文引用的方式并入本文中。

技术领域

[0003] 本发明大体上涉及用于混合实境的设备、系统及方法。

背景技术

[0004] 在计算机视觉及计算机图形中,3维(“3D”)重构是确定实际物体及/或环境的形状及/或外观的过程。一般来说,在本文中使用术语3D模型以指代3D环境的表示被装置模型化。3D重构可基于从包含相机的各种类型的传感器获得的对象的数据及/或图像。例如,手持式相机可用于获取关于小3D场景的信息且产生所述场景的近似虚拟模型。

[0005] 扩增实境(AR)及混合实境(MR)应用经常结合3D重构而使用。在AR中,可处理实际图像以将虚拟对象添加到所述图像。在混合实境(MR)效果中,可基于现有场景的3D模型实施例如减弱实境或中介实境,以移除及/或更改图像中的对象。在减弱实境中,从场景无缝地移除实际对象,而在中介实境中,实际对象被虚拟对象无缝地取代。常规的实时或准实时AR/MR方法经常不能最佳地表现,因为所述方法可使用图像修复技术,其中当前图像的部分被复制在将被移除的对象上。图像修复可导致假影,其可由于场景重构及实时相机跟踪中的准确度的缺乏及/或其它误差而加剧。

[0006] 因此,需要增强所再现的MR图像的质量以实现更佳的用户体验的图像处理方法。

发明内容

[0007] 根据一些方面,揭示用于混合实境应用的方法。在一些实施例中,所述方法可包括:获得多个实况图像,所述实况图像包括被跟踪对象;及针对所述多个实况图像中的一实况图像,确定相对于所述被跟踪对象的相机位姿。可再现不具有所述被跟踪对象的多个合成图像,其中每一合成图像是从基于所述相机位姿的3D模型中的相异关键帧获得。可比较所述实况及合成图像以确定与所述被跟踪对象相关联的第一掩模区,且可计算与所述被跟踪对象相关联的第一掩模区中的多个像素相对于所述多个合成图像中的合成图像子集的光流量值。对于所述第一掩模中的所述多个像素中的每一者,可部分基于所述光流量值确定所述合成图像子集中的每一者中的对应像素;及可确定所述第一掩模中的所述多个像素中的每一者的替换像素,其中每一替换像素对应于所述第一掩模中的相异像素,且将每一替换像素确定为随所述合成图像子集中对应于所述第一掩模中的所述像素的对应像素而

变。

[0008] 在另一方面中，一种移动台 (MS) 可包括：相机，其经配置以俘获包括被跟踪对象的多个实况图像；存储器，所述存储器用以存储包括多个关键帧的3D模型；及处理器，其耦合到所述相机及所述存储器。在一些实施例中，所述处理器可经配置以：针对所述多个实况图像中的一实况图像，确定相对于所述被跟踪对象的相机位姿；再现不具有所述被跟踪对象的多个合成图像，每一合成图像是从基于所述相机位姿的3D模型中的相异关键帧获得；比较所述实况及合成图像以确定与所述被跟踪对象相关联的第一掩模区；计算与所述被跟踪对象相关联的第一掩模区中的多个像素相对于所述多个合成图像中的合成图像子集的光流量值；针对所述第一掩模中的所述多个像素中的每一者，部分基于所述光流量值确定所述合成图像子集中的每一者中的对应像素；及确定所述第一掩模中的所述多个像素中的每一者的替换像素，每一替换像素对应于所述第一掩模中的相异像素，其中将每一替换像素确定为随所述合成图像子集中对应于所述第一掩模中的所述像素的对应像素而变。

[0009] 在另一方面中，一种设备可包括：成像装置，所述成像装置用以俘获包括被跟踪对象的多个图像；存储装置，所述存储装置用以存储包括多个关键帧的3D模型；及处理装置，所述处理装置耦合到所述成像装置及所述存储装置。在一些实施例中，所述处理装置可进一步包括：用于针对所述多个实况图像中的一实况图像，确定相对于所述被跟踪对象的相机位姿的装置；用于再现不具有所述被跟踪对象的多个合成图像的装置，每一合成图像是从基于所述相机位姿的3D模型中的相异关键帧获得；用于比较所述实况及合成图像以确定与所述被跟踪对象相关联的第一掩模区的装置；用于计算与所述被跟踪对象相关联的第一掩模区中的多个像素相对于所述多个合成图像中的合成图像子集的光流量值的装置；用于针对所述第一掩模中的所述多个像素中的每一者部分基于所述光流量值确定所述合成图像子集中的每一者中的对应像素的装置；及用于确定所述第一掩模中的所述多个像素中的每一者的替换像素的装置，每一替换像素对应于所述第一掩模中的相异像素，其中将每一替换像素确定为随所述合成图像子集中对应于所述第一掩模中的所述像素的对应像素而变。

[0010] 在另一方面中，一种计算机可读媒体可包括在由处理器执行时可执行一种方法中的步骤的指令，所述方法可包括：获得多个实况图像，所述实况图像包括被跟踪对象；针对所述多个实况图像中的一实况图像，确定相对于所述被跟踪对象的相机位姿；基于所述相机位姿再现不具有所述被跟踪对象的多个合成图像，每一合成图像是从基于所述相机位姿的3D模型中的相异关键帧获得；比较所述实况及合成图像以确定与所述被跟踪对象相关联的第一掩模区；计算与所述被跟踪对象相关联的第一掩模区中的多个像素相对于所述多个合成图像中的合成图像子集的光流量值；针对所述第一掩模中的所述多个像素中的每一者，部分基于所述光流量值确定所述合成图像子集中的每一者中的对应像素；及确定所述第一掩模中的所述多个像素中的每一者的替换像素，每一替换像素对应于所述第一掩模中的相异像素，其中将每一替换像素确定为随所述合成图像子集中对应于所述第一掩模中的所述像素的对应像素而变。

[0011] 所揭示的实施例还涉及由处理器使用计算机可读媒体或计算机可读存储器创建、存储、存取或修改的软件、固件及程序指令。所描述的方法可在处理器及各种移动装置上执行。

[0012] 下文关于以下各图进一步解释这些及其它实施例。应理解,对于所属领域的技术人员来说,其它方面将从以下详细描述变得容易显而易见,其中借助于说明展示及描述各个方面。图式和详细描述被视为本质上是说明性的而非限制性的。

附图说明

- [0013] 将参考图式仅通过举例描述本发明的实施例。
- [0014] 图1展示能够以与所揭示的实施例一致的方式实施减弱及中介实境效果的示范性移动装置的框图。
- [0015] 图2A展示示范性图像192,其可由混合实境应用处理。
- [0016] 图2B展示减弱实境图像194,其中对象193已被移除。
- [0017] 图2C展示中介实境图像196,其中对象193已被虚拟对象197替换。
- [0018] 图3A展示用于以与所揭示的实施例一致的方式促进来自重构的MR效果(例如减弱及中介实境效果)的示范性方法的流程图。
- [0019] 图3B展示示范性子例程以与所揭示的实施例一致的方式计算与至少一个被跟踪对象相关联的掩模的替换像素以促进来自重构的MR效果的流程图。
- [0020] 图4说明示范性减弱实境应用及单一合成图像的示范性工作流。
- [0021] 图5展示用于以与所揭示的实施例一致的方式促进来自重构的MR效果(例如减弱及中介实境效果)的示范性方法的流程图。
- [0022] 图6展示说明经启用以与所揭示的实施例一致的方式促进减弱及中介实境效果的服务器600的示意框图。

具体实施方式

[0023] 下文结合附图陈述的详细描述希望作为对本发明的各个方面的描述,而不希望表示其中可实践本发明的仅有方面。提供本发明中所描述的每一方面仅作为本发明的实例或说明,且其不一定解释为比其它方面优选或有利。详细描述包含用于提供对本发明的透彻理解的目的的特定细节。然而,所属领域的技术人员将明白,可在没有这些具体细节的情况下实践本发明。在一些情况下,以框图的形式示出众所周知的结构和装置以便避免混淆本发明的概念。首字母缩写词及其它描述性术语可仅出于便利及清晰的目的而使用,且不希望限制本发明的范围。

[0024] 在计算上昂贵的常规的3D重构中,通常与其它感官信息一起以分批模式离线处理一组数字图像,且可通常在较长的处理延迟之后获得环境的3D模型。因此,使用3D重构的实用的实时应用迄今已经受到限制。

[0025] 最近,常规的实时或准实时3D重构归因于增加的处理能力的可用性、高级算法以及新形式的输入数据而已具吸引力。用户现在可准实时地获得对3D重构的反馈,因为所俘获图片被计算装置(包含移动装置)快速处理,进而促进实时或准实时的AR及/或MR应用。

[0026] 可实时交互的AR/MR应用通常组合实际及虚拟图像且在3-D中执行所俘获图像与对象之间的对准。因此,确定什么对象存在于实际图像中以及那些对象的位置可促进许多AR及/或MR系统的有效操作,且可以用于辅助虚拟对象布局、移除、遮挡及其它效果。

[0027] 在计算机视觉中,检测是指在所俘获图像帧中定位目标对象及计算相对于参考框

架的相机位姿的过程。跟踪是指图像帧的时间序列上的相机位姿估计。

[0028] 常规的准实时技术通常使用图像修复技术。图像修复是指使用图像内插技术使用背景或虚拟像素替换像素。常规的图像修复技术遭受场景重构期间缺乏准确度的问题，其被来自实况相机跟踪的误差复合。因此，图像修复技术当结合AR/MR应用使用时经常产生不好的用户体验。准确度的缺乏可导致再现质量的降级及各种图像假影，例如重像及模糊，其贬低了AR/MR体验。因此，本文中揭示的一些实施例应用计算机视觉及图像处理技术来改进重构及跟踪准确度，进而提高用户AR/MR体验。

[0029] 下文关于以下各图进一步解释这些及其它实施例。应理解，对于所属领域的技术人员来说，其它方面将从以下详细描述变得容易显而易见，其中借助于说明展示及描述各个方面。图式和详细描述被视为本质上是说明性的而非限制性的。

[0030] 图1展示能够以与所揭示的实施例一致的方式实施计算机视觉应用(包含减弱及中介实境效果)的示范性移动装置或移动台(MS)100的框图。在一些实施例中，MS 100可能基于3D环境的现有模型实施例如中介实境等MR方法。在一些实施例中，可以与所揭示的实施例一致的方式实时或准实时地实施所述MR方法。

[0031] 如本文中所使用，移动装置或移动台(MS)100可呈以下形式：蜂窝式电话、移动电话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)或个人数字助理(PDA)、膝上型计算机、平板计算机、笔记本及/或手持式计算机。术语移动装置或移动台在本文中可互换地使用。在一些实施例中，MS 100可能够接收无线通信及/或导航信号。

[0032] 此外，术语“移动台”还既定包含例如通过短程无线、红外、有线连接或其它连接与个人导航装置(PND)通信的装置，及/或在所述装置处或所述PND处发生位置相关处理。而且，“移动台”既定包含所有装置，包含能够与服务器(例如图6中的服务器600)通信的各种无线通信装置，不管无线信号接收、辅助数据接收及/或相关处理是否发生在所述装置处、服务器处或与网络相关联的另一装置处。以上各者的任何可操作组合也被视为“移动台”。

[0033] 术语“移动台”还既定包含可未经配置以无线地或经由有线连接连接到网络或以其它方式与另一装置通信的游戏或其它装置。例如，“移动台”可省略通信元件及/或联网功能性。例如，本文中描述的实施例可在未被配置成连接以用于与另一装置进行有线或无线联网的独立装置中实施。

[0034] 如图1中所示，MS 100可包含相机110、传感器组或传感器130、处理器150、存储器160、显示器180及/或收发器170，其可操作性地彼此耦合且耦合到MS 100上的其它功能单元(未图示)。例如，相机110、处理器150、存储器160、显示器180、收发器170及/或其它功能单元可通过总线、线、纤维、链路等或其某一组合操作性地耦合。

[0035] 收发器170可(例如)包含经启用以经由一或多种类型的无线通信网络发射一或多个信号的发射器，及用以接收经由所述一或多种类型的无线通信网络发射的一或多个信号的接收器。收发器170可准许基于多种技术与无线网络通信，所述技术例如(但不限于)毫微微小区、Wi-Fi网络或无线局域网(WLAN)(其可基于IEEE 802.11族标准)、无线个域网(WPAN)、例如蓝牙、近场通信(NFC)、基于IEEE 802.15x族标准等的网络及/或无线广域网(WWAN)(例如LTE、WiMAX等)。在一些实施例中，收发器170可能够从全球导航卫星系统(GNSS)接收信号，所述信号可以用于确定MS 100的位置。移动装置还可包含用于经由有线

网络进行通信的一或多个端口。在一些实施例中，可省略收发器170及/或移动装置100上的一或多个其它端口。

[0036] 在一些实施例中，MS 100可包括例如CCD或CMOS传感器等图像传感器及/或相机110。其在下文称为“相机110”。相机110可将光学图像转换为电子或数字图像且可将所俘获图像发送到处理器150。在一些实施例中，相机110可收容在可穿戴显示器中，其可操作性地耦合到处理器150及/或MS 100中的其它功能单元但与其分开收容。

[0037] 在一些实施例中，相机110可包含多个相机(前向及/或后向相机、广角相机、高分辨率相机)，且还可并入CCD、CMOS及/或其它传感器。相机110(其可为静态及/或视频相机)可俘获环境的一连串静态及/或视频图像帧且将所俘获图像帧发送到处理器150。在一些实施例中，相机110可为可穿戴相机或外部相机，其可操作性地耦合到MS 100中的其它功能单元但与其它功能单元分开收容。在一个实施例中，由摄像机110俘获的图像可呈原始的未经压缩格式且可在经处理及/或存储在存储器160中之前进行压缩。在一些实施例中，可由处理器150使用无损或有损压缩技术来执行图像压缩。在一些实施例中，相机110可为能够俘获3D图像的立体相机。在另一实施例中，相机110可包含能够估计深度信息的深度传感器。

[0038] 处理器150还可执行软件以处理由相机110俘获的图像帧。例如，处理器150可能够处理从相机110接收的一或多个图像帧以确定相机110的位姿、实施各种计算机视觉及图像处理算法、运行AR/MR应用及/或执行对应于从相机110接收的图像的环境的3D重构。相机110的姿态是指相机110相对于参考框架的位置及定向。在一些实施例中，可针对6个自由度(6DOF)确定相机位姿，所述6个自由度其是指三个平移分量(其可由参考帧的坐标的X、Y、Z坐标给定)及三个角度分量(例如，相对于相同参考框架的横滚、俯仰及偏航)。

[0039] 在一些实施例中，可由处理器150基于从相机110接收的图像帧使用视觉跟踪解决方案来确定及/或跟踪相机110及/或MS 100的位姿。在一些实施例中，可使用专用电路实施计算机视觉(CV)模块155，所述专用电路例如为专用集成电路(ASIC)、数字信号处理器(DSP)及/或专用处理器(例如处理器150)。在一些实施例中，CV模块155可包含与MS 100上的一或多个其它处理器通信的功能性。在一些实施例中，CV模块155可实施各种计算机视觉及/或图像处理方法，例如3D重构、AR/MR、图像压缩及滤波。CV模块155还可实施基于计算机视觉的跟踪、基于模型的跟踪、地图创建同时定位与地图创建(SLAM)等。在一些实施例中，由CV模块155实施的方法可基于相机110俘获的彩色或灰度级图像数据，其可用于产生相机的6DOF位姿测量值的估计。

[0040] SLAM是指一类技术，其中在创建环境的地图(例如正由MS 100模型化的环境的地图)的同时跟踪MS 100相对于所述地图的位姿。SLAM技术包含视觉SLAM(VLSAM)，其中由单一相机(例如MS 100上的相机110)俘获的图像可以用于在创建环境的地图的同时跟踪所述相机相对于所述地图的位姿。VSLAM可因此涉及在跟踪相机的6DOF位姿的同时还确定周围环境的3-D结构。例如，在一些实施例中，VSLAM技术可检测一或多个所俘获图像帧中的突出特征小片且将所俘获的成像帧存储为关键帧或参考帧。在基于关键帧的SLAM中，可随后例如通过将当前所俘获图像帧与一或多个关键帧进行比较而确定相机的位姿。

[0041] 在一些实施例中，CV模块155可包括3D重构模块158，其可使用相机位姿及每像素地图信息来创建环境的3D模型或表示。在一些实施例中，可使用专用电路实施3D重构模块158，所述专用电路例如为专用集成电路(ASIC)、数字信号处理器(DSP)及/或专用处理器

(例如处理器150)。在一个实施例中,处理器150可通过使用单眼VSLAM技术建立MS 100周围的环境的粗糙地图来跟踪相机110的位置,以用于对相机110的准确及稳健的6DOF跟踪。术语单眼是指使用单一非立体相机俘获图像或不具有深度信息而俘获的图像。

[0042] 在一些实施例中,3D模型可呈纹理化的3D网格、体积数据集、CAD模型等的形式,其可以用于再现正被模型化的3D环境。例如,在其中使用3D网格的实施例中,可使用VSLAM技术中的关键帧来获取环境的稀疏点云表示。术语稀疏点云是指坐标系统(例如具有X、Y以及Z坐标的3D坐标系统)中的一组分散数据点。随后可使用适当的分散数据内插方法将点云表示转换为3D网格。

[0043] 在一些实施例中,处理器150还可从传感器130接收输入,所述传感器可包含多种其它传感器,例如环境光传感器、声音传感器、惯性测量单元(IMU)等。可包括3轴加速度计、3轴陀螺仪和/或磁力计的所述IMU可向处理器150提供速度、定向和/或其它位置相关信息。在一些实施例中,IMU可与由相机110俘获每一图像帧同步地输出所测量的信息。在一些实施例中,IMU的输出可部分由处理器150使用以确定相机110及/或MS 100的位姿。

[0044] 可使用硬件、固件及软件的组合来实施处理器150。处理器150可表示可配置以执行与3D重构、SLAM、跟踪、模型化、图像处理等相关的计算程序或过程的至少一部分且可从存储器160检索指令及/或数据的一或多个电路。可使用一或多个专用集成电路(ASIC)、中央及/或图形处理单元(CPU及/或GPU)、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)、控制器、微控制器、微处理器、嵌入式处理器芯、电子装置、经设计以执行本文中描述的功能的其它电子单元或其组合来实施处理器150。

[0045] 可在处理器150内及/或在处理器150外部实施存储器160。如本文所使用,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,且不限于任何特定类型的存储器或任何特定数目个存储器或上面存储有存储器的任何特定类型的物理媒体。在一些实施例中,存储器160可保持促进图像处理、SLAM、跟踪、模型化、3D重构、AR/MR应用及由处理器150上的CV模块155执行的其它任务的程序代码。例如,存储器160可保持数据、所俘获的静态图像、深度信息、视频帧、程序结果、3D模型、关键帧以及由各种传感器提供的数据。一般来说,存储器160可表示任何数据存储机构。存储器160可包含(例如)主存储器及/或辅助存储器。主存储器可包含(例如)随机存取存储器、只读存储器等。虽然在图1中说明为与处理器150分离,但应理解,主存储器的全部或部分可提供于处理器150内或另外与处理器150位于同一地点及/或耦合到处理器150。

[0046] 辅助存储器可包含例如与主存储器相同或类似类型的存储器及/或一或多个数据存储装置或系统,例如(举例来说)快闪/USB存储器驱动器、存储卡驱动器、磁盘驱动器、光盘驱动器、磁带驱动器、固态驱动器、混合驱动器等。在某些实施方案中,辅助存储器可操作地接纳非暂时性计算机可读媒体或另外可配置以在耦合到移动装置100的可装卸媒体驱动器(未图示)中耦合到非暂时性计算机可读媒体。在一些实施例中,非暂时性计算机可读媒体可形成存储器160及/或处理器150的部分。

[0047] 此外,MS 100可包含能够再现彩色图像(包含3D图像)的屏幕或显示器180。在一些实施例中,显示器180可用于显示由相机110俘获的实况图像、AR图像、图形用户接口(GUI)、程序输出等。在一些实施例中,显示器180可包括及/或收容有触摸屏,以准许用户经由虚拟

键盘、图标、菜单或其它图形用户接口 (GUI)、用户手势及/或输入装置 (例如触控笔及其它写入工具) 的某一组合输入数据。在一些实施例中, 可使用液晶显示器 (LCD) 显示器或发光二极管 (LED) 显示器 (例如有机LED (OLED) 显示器) 实施显示器180。在其它实施例中, 显示器180可为可穿戴显示器, 其可操作性地耦合到MS 100中的其它功能单元但与其分开收容。在一些实施例中, MS 100可包括用以准许通过耦合到MS 100的单独监视器显示所述MR图像的端口。

[0048] 不是移动装置100中包括的所有模块都已经在图1中展示。还可以与本发明一致的方式以各种方式修改示范性移动装置100, 例如通过添加、组合或省略所展示的功能块中的一或更多者。例如, 在一些配置中, 移动装置100可不包含收发器170。在一个实施例中, MS 100可呈游戏或其它装置的形式, 其可未经配置以连接到网络或另外无线地或经由有线连接与另一装置通信。因此, 可在例如不要求与另一装置通信的移动台中的独立AR系统/装置中使用本文中揭示的实施例。

[0049] 此外, 在某些实例实施方案中, 移动装置100可包含IMU, 所述IMU可包括3轴陀螺仪及/或磁力计。IMU可将速度、定向及/或其它位置相关信息提供给处理器150。在一些实施例中, IMU可与由相机110俘获每一图像帧同步地输出所测量的信息。在一些实施例中, IMU 130的输出可部分由处理器150使用以确定相机110及/或MS 100的位姿。在一些实施例中, IMU的输出可用于确定校正及/或以其它方式调整所估计的位姿。此外, 在一些实施例中, 从相机110接收的图像还可用于重新校准或执行IMU的偏差调整。在一些实施例中, MS 100可包括定位系统 (其可使用GNSS及/或无线 (例如陆地定位信号、Wi-Fi、WWAN、WPAN) 信号) 以确定MS 100的位置信息。在一些实施例中, MS 100可包括多种其它传感器, 例如环境光传感器、麦克风、声音传感器、超声波传感器、激光测距仪等。在一些实施例中, 移动装置100的部分可呈一或多个芯片组及/或其类似者的形式。

[0050] 图2A展示示范性图像192, 其可由混合实境应用处理。在一些实施例中, MS 100可能够执行例如减弱实境等MR, 其中从所俘获图像无缝地移除实际对象。例如, 如图2B中所展示, 在减弱实境图像194中, 已经从图像194移除对象193 (图像192中的一堆书)。

[0051] 在一些实施例中, MS 100可能够执行例如中介实境等MR, 其中可通过虚拟对象无缝地替换实际对象。例如, 如图2C中所展示, 在中介实境图像196中, 对象193 (图像192中的一堆书) 已经被图像196中的虚拟盆栽植物对象197替换。

[0052] 作为另一实例, 在实施中介实境的一个示范性方法中, 可从用户的客厅移除现有的沙发 (减弱实境) 且使用用户可打算购买的沙发替换 (中介实境)。

[0053] 在一些实施例中, 可使用环境的现有3D模型执行MR。在一些实施例中, 可由MS 100上的处理器150使用基于关键帧的SLAM技术 (例如并行跟踪及地图创建 (PTAM)) 处理从相机110接收的图像且获得环境的3D模型。各种众所周知的技术可用于获得环境的3D模型。在一些实施例中, 可使用及/或更新环境的预先存在及/或离线的3D模型。在一些实施例中, 各种交互式3D重构方法 (例如基于关键帧的SLAM) 可用于获得环境的3D模型。

[0054] 作为另一中介实境实例, 置于表面上的玩具动作图可被动画虚拟动作图替换。在一些实施例中, 所述玩具动作图可实际上在被虚拟动作图替换 (中介实境) 之前从图像移除 (减弱实境)。因为实际玩具动作图已首先从图像移除, 所以可基于3D模型在移除之后在所述图像中任意移动替换实际动作图的虚拟动作图。因此, 本文中揭示的实施例当实施中介

实境时不需要使用对实际对象的逐像素“复绘”或“图像修复”。

[0055] 在一些实施例中,通过获得不具有前景的背景的3D模型,可结合前景的再现来处理背景的虚拟再现。在以上实例中,实际玩具动作图可被视为前景,且图像的剩余部分被视为背景。因此,对于以上实例,通过获得不具有实际玩具动作图的背景的3D模型,背景可用于替换前景像素。随后,虚拟玩具动作图可在减弱实境图像中无缝地动画化及再现。

[0056] 在一些实施例中,通过重构3D背景,可由MS 100实时地实施若干MR效果且可在耦合到MS 100的显示器上查看。例如,可从图像移除已知的实际对象或前景对象(减弱实境)。当已知的实际对象在重构期间不存在时及/或当前景对象在重构期间存在时可应用本文中揭示的方法。在一些实施例中,本文中揭示的方法还可促进对在重构期间不存在的先前未知对象的移除。例如,可从图像移除用户的手及/或其它非所要的假影。在一些实施例中,所揭示的方法可促进虚拟对象与实际对象之间的精确实时遮挡效果。

[0057] 图3A展示用于从重构使用促进MR效果(例如减弱及中介实境效果)的示范性方法300的流程图。在一些实施例中,可由MS 100上的处理器150使用由相机110俘获及/或存储在存储器160中的图像帧执行方法300。在一些实施例中,可由处理器150结合一或多个其它功能单元执行方法300。在一些实施例中,可由CV模块155执行方法300。

[0058] 在一些实施例中,实况输入视频图像帧305可包括图像帧序列。在一些实施例中,在步骤315中,可获得环境的3D模型。

[0059] 例如,在一些实施例中,可使用任何离线重构方法或通过手动3D模型化而产生3D模型。在一些实施例中,3D模型(例如环境的CAD模型)可为可用的且从耦合到MS 100的服务器(例如图6中的服务器600)获得。3D模型可包含关于对象的大小及形状的信息、关于对象的有色色彩及外观信息及可应用于对象的表面的纹理。

[0060] 或者,在一些实施例中,可使用各种交互式3D重构方法获得3D模型。在一些实施例中,可使用基于的关键帧地图创建同时定位与地图创建(SLAM)获得环境的3D模型。在一些实施例中,3D模型可不包含已知的被跟踪对象。例如,一或多个前景物体可未包含在3D模型中。

[0061] 基于SLAM的技术准许在产生未知环境的地图的同时定位相机110及/或MS 110的位置。在SLAM中,由相机(例如相机110)获得的图像可用于以相对低的计算开销将未知环境模型化,进而促进实时及/或准实时的模型化。

[0062] 在基于关键帧的SLAM中,一组关键帧(其可为由相机110在各个时间点处俘获的快照或图像帧)可用于产生环境的地图或模型。在一些实施例中,每一关键帧可与以相机为中心的坐标框架相关联,且可包括不同分辨率的图像的角锥。例如,可对所俘获图像关键帧子取样以获得与所述关键帧相关联的不同分辨率的图像。

[0063] 此外,关键帧中的每一特征点可与以下各者相关联:(i)其源关键帧;(ii)与源关键帧相关联的经子取样图像中的一者;及(iii)经子取样图像内的像素位置。每一特征点还可与小片相关联,所述小片是指(经子取样)图像的对应于一区的一部分,其可呈围绕(经子取样)图像中的特征点的多边形的形式。在一些实施例中,在步骤315中,基于关键帧的SLAM可用于处理输入视频图像帧305并获取稀疏3D特征点且获得场景的稀疏点云重构。在一些实施例中,在步骤315中,可周期性地执行SLAM地图创建更新。

[0064] 在步骤310中,可处理实况输入视频图像帧305中的图像帧以获得图像帧的当前相

机视角或相机位姿。在一些实施例中,SLAM跟踪可用于获得当前视角。例如,在一个实施例中,基于的关键帧SLAM在其先前已用于重构的情况下可以用于跟踪。在一些实施例中,还可跟踪将从实况输入视频图像帧305移除(例如在减弱实境应用中)的一或多个已知对象。跟踪是指图像帧的时间序列上的相机位姿估计。相机位姿是指相机相对于参考框架的位置及对准。

[0065] 在一些实施例中,可使用MS 100上的多个处理器150、相同处理器150上的不同处理器核心及/或通过处理器150上的不同进程线程并行地执行步骤310及315。在一些实施例中,可在步骤310中针对每个传入的图像帧跟踪相机110的位姿,同时在步骤315中可周期性地执行SLAM地图创建更新。

[0066] 在步骤320中,可将从SLAM地图创建步骤315输入的稀疏3D点云转换为3D网格。在一些实施例中,可使用分散数据内插方法将稀疏点云转换为网格。给定那些点的一组分散点及一组值,分散数据内插技术旨在获得函数 f ,其确定点(x)是否处于由分散数据点表示的表面上。例如,可界定函数 $f(x)$ 以便在 x 处于表面上且 $f(x) \neq 1$ 的情况下在 x 不处于表面上的情况下产生值1($f(x) = 1$)。

[0067] 在一个实施例中,可应用分层分解技术且可使用图像角锥中的经下取样图像以构建点集合的粗糙到精细阶层,其可随后从最粗糙层级开始经内插。因此,可通过补偿在先前层级处计算的内插函数对所述阶层的点集合进行内插。在一些实施例中,内插可使用紧凑地支持的基底函数。例如,在一个实施例中,内插可使用紧凑地支持的径向基底函数(RBF)。

[0068] RBF是其值仅取决于距称为“中心”的点的距离的实值函数。例如,在N维度欧几里得空间中, $N > 0$,在m个相异点 x_i 的情况下, $1 \leq i \leq m$,函数 $y(x)$ 的近似可表示为N个径向基底函数的总和,每一径向基底函数与不同的中心 x_i 相关联,且通过适当的系数 w_i 进行加权。对于紧凑地支持的RBF,在围绕“中心”的有限范围的周围中, $f(x) \neq 0$ 。

[0069] 在步骤325中,在一些实施例中,可通过基于当前相机位姿将关键帧的纹理地图创建投影到代理几何形状上而获得多个视图,例如纹理化视图S1、S2及S3。术语“代理几何形状”用于指关于场景的近似几何信息。在投影式纹理地图创建中,将相机位姿用作投影的中心将相机位姿及具有纹理的虚拟图像平面投射到几何模型上。

[0070] 在一些实施例中,可基于在步骤310中获得的当前视角或当前相机位姿再现在步骤320中获得的网格,且可通过将SLAM关键帧用作纹理而应用视图相依纹理地图创建(VDTM)。所述网格进而变为基于图像的再现方法中的代理几何形状。

[0071] VDTM是用于从先前所俘获图像的稀疏集合产生具有大致已知几何形状的场景的新视图的技术。在VDTM中,使用几何模型执行此内插以确定每一输入图像中的对应于输出图像中的所要光线的像素。对在角度上接近及/或角度方面相对靠近所要光线的对应光线进行加权以对经内插结果作出最大贡献。

[0072] 在一个实施例中,可确定其相关联的相机位姿在角度移位上相对靠近或在角度上接近于当前视角的多个关键帧,且可使用所述多个关键帧的子集产生合成再现。例如,可确定其相关联的相机位姿在角度空间中最靠近当前相机位姿或视角的三个关键帧,且可使用所述三个关键帧产生三个合成再现S1、S2及S3。

[0073] 在一些实施例中,可使用非结构化的流明图再现(ULR)技术从多个关键帧选择关键帧子集且产生对应的再现。ULR是能够通过使用输入图像的非结构化集合、与输入图像中

的每一者相关联的相机位姿估计及场景的近似几何代理来再现视图的实时再现技术。

[0074] 例如,可使用ULR选择三个关键帧且产生三个合成再现S1、S2及S3,每一者对应于选定关键帧中的一者。例如,可针对当前视角在图像平面中的一组顶点处评估“混合字段”,且可在整个图像上内插所述混合字段。所述混合字段描述当重构给定像素时如何对每一源关键帧图像进行加权。权重可部分基于多种因素,包含(例如)当前视角与关键帧图像的视角之间的角度差异。

[0075] 在步骤328中,可比较实况及合成图像且可从实况图像移除已知跟踪对象中的一或多个。例如,可将在步骤325中获得的合成图像S1、S2及S3与一或多个实况输入图像帧305进行比较。

[0076] 接下来,在步骤330中,可将已知对象的二值图再现到掩模M中,所述掩模标记实况图像中的将被替换的像素。在一些实施例中,掩模M可生长某一量以产生掩模M'。例如,掩模M可通过包含围绕M的区中的像素而生长到M'以便适应跟踪及模型化误差。在一些实施例中,可通过将过滤器操作应用到所述图像而生长掩模。

[0077] 例如,在一个实施例中,过滤器可确定M的轮廓且随后向每个轮廓像素应用圆盘形内核或2D高斯内核。用于图像卷积的内核可表示为数字矩阵且是针对围绕像素的区而界定的。过滤器操作通过将所述矩阵中的值与内核中的对应像素的值进行卷积而获得M中的每一输入像素的输出值。卷积下的结果取决于矩阵中的数字的量值及分布且确定内核内的每一像素对输出像素的贡献。

[0078] 在另一实施例中,过滤器可基于计算到M的距离变换,且随后将阈值应用于距离变换。此过滤器可给予范围0到1中的值,其可被解译为用于混合实际图像及合成图像的阿尔法掩模。

[0079] 在子例程335中,可随后计算M'内的像素的替换像素,如图3B中所描述。在一些实施例中,可使用所计算的替换像素再现减弱实境图像340。所述方法随后通过处理接下来的输入视频图像帧305而开始另一迭代。在一些实施例中,方法300可帮助减少或消除渗色、重像或归因于跟踪或所再现的图像中的其它误差而引起的其它假影。

[0080] 例如,考虑围绕前景对象的掩模的边界附近的少量像素可确保跨越掩模边界的连续性。例如,在其中来自实况图像的像素邻近于来自合成图像的像素的情形中,可通过光流计算的结果校正任何误差(例如来自跟踪)。

[0081] 如早先所描述,常规的技术还可经常导致DR图像的可未显得可靠的填充部分。例如,使用常规技术的DR背景可不同于被移除对象后方的实际背景。当背景复杂或非均匀时,这些不准确及不一致尤其明显。当使用常规技术时,来自用于图像修复非均匀背景法信息显得不一致,这是因为可能将视觉结构复制到它们不配合的位置。相比而言,部分基于使用背景场景的基于图像的模型化,所揭示实施例使用准确背景位置处的实际信息填入被正移除的前景对象遮挡的像素中。因此,所揭示实施例相对不含此假影。

[0082] 在一些实施例中,例如对于其中实际对象被虚拟对象替换的中介实境应用,可修改方法300以在其中实际对象被移除的区域中另外再现所要的虚拟对象。在一些实施例中,方法300可应用于一或多个被跟踪对象。

[0083] 在一些实施例中,还可以其它方式修改方法300。例如,在一个实施例中,当3D模型包含已知的被跟踪目标时,随后,对应于将被移除的对象的掩模可应用于实况输入视频图

像帧305(如在步骤330中),且可修改关于重构的步骤315、320及325以考虑未被掩模的像素。

[0084] 在一些实施例中,可将实况输入视频图像帧305分段为前景对象及背景场景,其在一些情况下可通过标记前景对象来实现。对于减弱实境应用,前景对象可为将被移除的对象。在一个实施例中,可起初使用例如GrabCut等已知方法将前景和背景分段。可随后在实况输入视频图像帧305中跟踪前景对象的轮廓且所述轮廓可用于创建将被移除的前景对象的掩模。可修改关于重构的步骤315、320及325以考虑未被掩模的像素。

[0085] 在一些实施例中,方法300可以用于标记移除。例如,如果用户使用基准标志器(例如白底黑色基准标志器)与场景交互,那么可实际上移除这些标志器。

[0086] 图3B展示示范性子例程335以与所揭示的实施例一致的方式计算与至少一个被跟踪对象相关联的掩模的替换像素以促进来自重构的MR效果的流程图。

[0087] 在步骤344中,如果合成图像中的任一者与实况图像减去(或不具有)已知/被跟踪对象确切地或在某一预定义阈值量度内匹配(在步骤344中的“Y”),那么在步骤346中,可使用来自合成图像的像素替换实况图像中的掩模内的所有像素。在步骤344中,在实况图像减去已知对象与合成图像中的任一者之间不存在匹配或失配程度超过预定义阈值量度的情况下(在步骤344中的“N”),那么所述方法前进到步骤348。

[0088] 在步骤348中,可使用光流方法计算掩模M'内的实际及合成图像像素之间的关系。光流是指图像序列中的强度图案的明显运动。3D表面上的每一移动点当投射到图像平面上时产生二维(2D)路径。术语“2D运动场”用于指所有可见表面点的2D速度。可使用光流估计基于时变的图像强度来计算2D运动场的近似。

[0089] 在一些实施例中,光流计算可限于由掩模M'覆盖的区域。例如,在一个实施例中,可针对实况输入视频图像帧305中的掩模M'到合成图像S1、S2及S3内的每个像素执行光流计算。

[0090] 在步骤350中,可基于光流计算确定M'-M内的每个像素的移位及色彩/强度校正,使得可使来自合成图像的像素与实况图像的在掩模M外部及掩模M'内的部分匹配。

[0091] 在步骤352中,对于M'-M内的像素,可基于在步骤350中确定的光流校正来确定三个对应像素,各自来自S1、S2及S3。

[0092] 接下来,在步骤354中,在一些实施例中,可将M'-M内的每个像素的值计算为三个对应像素的加权平均,各自来自S1、S2及S3。在一些实施例中,像素的权重可分别为所述像素的光流的量值及实况输入视频图像帧305及S1、S2、S3之间的角度移位的函数。

[0093] 在一些实施例中,可将M'-M中的光流计算约束为仿射运动,进而限制所估计的未知数的数目。对于仿射运动,存在六个未知数。对于小区域,仿射变换可提供平滑表面的图像运动的准确近似。

[0094] 在一些实施例中,在步骤356中,可通过从边界M'-M进行分散数据内插而计算掩模M内的像素的光流。可使用经内插光流值将M中的像素的值计算为三个对应像素的加权平均,各自来自S1、S2及S3。此外,可通过基于所述光流混合来自S1、S2及S3的像素而获得M内的像素值。

[0095] 在一些实施例中,在将M中的像素的阿尔法(透明度)值计算为计算M'的部分的情况下,随后可使用像素的所述阿尔法值将所述像素的新计算的值与实况输入视频图像帧

305中的对应像素的原始值混合。在一些实施例中,可(i)省略步骤354,且(ii)通过修改步骤356,使得可计算M'中的每个像素的新值,而修改例程335。

[0096] 在一些实施例中,例如对于中介实境应用,在实际对象可被虚拟对象替换的情况下,可修改例程335以基于从调用的程序接收的参数而在其中实际对象被移除的区域中另外再现所要的虚拟对象。

[0097] 在一些实施例中,可使用方法300及/或335从已知场景移除未知或未预期的对象。在一个实施例中,可通过光流及分段的联合估计而执行存在于先前经重构场景中的未知对象的移除。例如,可修改例程335以包含以下步骤:比较整个合成及实况图像且确定不匹配的区。可随后使用先前描述的技术移除这些区中的任一者。例如,如果用户的手或手指无意地处于图像帧中,那么可从所述场景移除这些。

[0098] 在另一实施例中,可在适当修改的情况下使用方法300及/或335以递送虚拟及实际对象之间的精确遮挡。例如,如果已知(被跟踪或静止)对象存在于实况输入视频图像帧305中且被跟踪对象可遮挡虚拟对象,那么可在虚拟z缓冲器中再现实际已知对象的幻影或虚拟表示。可使用所得的幻影z缓冲器分辨虚拟对象的遮挡。

[0099] 为考虑跟踪误差,可计算幻影z缓冲器图像与实况视频中的实际对象的对应图像之间的光流,且将其用于扭曲幻影z缓冲器图像以匹配实况视频。

[0100] 因此,在一些实施例中,可修改步骤330以再现掩模M'',其不进一步生长,从而对应于虚拟对象的占据面积。在一些实施例中,可随后修改例程335以计算实况视频中的实际对象的虚拟表示M''及对应图像之间的光流以扭曲虚拟z缓冲器而匹配实况视频。

[0101] 对于虚拟对象的每个像素,可使用从扭曲的虚拟z缓冲器的深度测试以确定所述像素是否被绘制。例如,例如在其中虚拟人物可隐藏在实际对象后方的游戏期间可使用其中实际对象遮挡虚拟对象的遮挡。应注意,在上文描述的实施例中,没有移除对象。替代地,针对新虚拟对象的每个像素的从扭曲z缓冲器的深度测试确定所述像素是否被绘制。

[0102] 在一些实施例中,方法300的部分可由MS 100及一或多个服务器或通过收发器170无线地耦合到MS 100的其它计算机的某一组合执行。例如,MS可将数据发送到服务器,且方法300中的一或多个步骤可由服务器执行且结果可传回到MS 100。

[0103] 图4说明用于示范性减弱实境应用及用于单一合成图像的示范性工作流400。在图4中,在步骤410中,例如,可从对应于环境的3D模型获得经重构场景415。在一些实施例中,经重构场景415可不包含前景对象428。

[0104] 在步骤420中,可识别用于移除的前景对象428且可创建掩模425。在一些实施例中,在步骤420中,可从实况图像422确定相机位姿。

[0105] 在步骤430中,使用经重构场景415再现在步骤420中确定的基于相机位姿的合成图像432。如图4中所展示,合成图像432不包含区435中的对应于掩模425的对象425。

[0106] 在步骤440中,合成图像432可与实况图像422组合且与掩模425相关联的像素可被区435中的对应像素替换。在一些实施例中,步骤440中展示的掩模445(其可对应于M')可大于掩模425(其可对应于M)。如图4中所展示,在校正之前,组合图像345中可能存在渗色、重像或归因于跟踪或其它误差而引起的其它假影450。

[0107] 在步骤460中,通过计算实况图像422与合成图像432(及一或多个额外合成图像)之间的光流,可将校正应用于组合图像345。例如,在一个实施例中,所述光流计算可限于由

掩模445覆盖的区域。例如,在一个实施例中,可针对掩模445到一或多个合成图像内的每个像素执行光流计算。光流计算可产生移位及色彩/强度校正,使得可使来自合成图像的像素与掩模M中的实况图像中的像素匹配。可应用所述校正以获得校正之后的组合图像465。

[0108] 图5展示用于以与所揭示的实施例一致的方式促进来自重构的MR效果(例如减弱及中介实境效果)的示范性方法的流程图500。在一些实施例中,方法500可由MS 100执行。在一些实施例中,作为初始化过程的部分,可将第一图像帧分段为前景,其包括被跟踪对象及背景。例如,可使用不具有被跟踪对象的3D模型或例如GrabCut等方法或其它分段技术。

[0109] 在步骤505中,可获得多个实况图像帧中的第一或下一包括被跟踪对象的实况输入视频图像帧305。接下来,在步骤510中,针对所述多个实况图像中的一实况图像,可处理实况图像305以确定相对于被跟踪对象的相机位姿。在一些实施例中,可使用SLAM跟踪技术确定包括被跟踪对象的实况输入视频图像帧305的相机位姿。一般来说,可使用任何实时跟踪技术跨越实况图像帧305来跟踪对象。

[0110] 在步骤520中,可使用相机位姿再现不具有被跟踪对象的多个合成图像,其中每一合成图像可从3D模型505中的相异关键帧获得。例如,可再现不具有被跟踪对象的基于相机位姿在角度接近度上最接近当前实况图像帧的三个合成图像。在一些实施例中,可基于相异关键帧相对于相机位姿的它们的角度移位接近度而选择相异关键帧。在一些实施例中,3D模型可不包含被跟踪对象。

[0111] 在一些实施例中,可离线产生所述3D模型。在一些实施例中,可使用VSLAM技术获得所述3D模型。例如,可使用VSLAM技术获取关键帧且创建环境的稀疏点云表示。在一些实施例中,可从稀疏云表示获得纹理化的3D网格。

[0112] 在步骤525中,可比较所述实况及合成图像以确定与被跟踪对象相关联的第一掩模区。接下来,在步骤530中,可计算与被跟踪对象相关联的第一掩模区中的多个像素相对于多个合成图像中的合成图像子集的光流量值。

[0113] 在步骤530中,对于第一掩模中的所述多个像素中的每一者,可部分基于光流量值而确定合成图像子集中的每一者中的对应像素。在以上实例中,可在三个合成图像中的每一者中确定对应于第一掩模中的像素的合成图像像素。例如,基于所述光流,可获得用于第一掩模中的所述多个像素中的每一者的移位及色彩及/或强度校正。可使用用于掩模内部的像素的移位及色彩/强度校正从合成图像确定与第一掩模内的实况图像的部分匹配的对应的像素。

[0114] 在步骤540中,可确定第一掩模中的所述多个像素中的每一者的替换像素,其中每一替换像素可对应于第一掩模中的相异像素,且其中将每一替换像素确定为随所述合成图像子集中对应于与所述第一掩模中的所述替换像素对应的像素的像素而变。在一些实施例中,可将每一替换像素确定为合成图像子集中的对应像素的加权平均,其中与所述子集中的合成图像中的对应像素相关联的权重部分基于所述对应像素的光流量值及实况图像及合成图像之间的与所述对应像素相关联的角度移位。在一些实施例中,第一掩模(M')区可包括围绕被跟踪对象的第二掩模区(M),且可通过从边界M'-M进行分散数据内插而确定M中的每一替换像素的光流量值。

[0115] 在一些实施例中,可随后再现减弱实境图像,其中可至少部分基于替换像素而获得减弱实境图像。在一些实施例中,还可在图像中的由第一掩模指示的区中再现虚拟对象,

其中基于当前相机位姿而再现所述虚拟对象。在一些实施例中，可通过以下操作在所述图像中的由第一掩模指示的区中再现所述虚拟对象：将所述图像中的至少一个已知实际对象的虚拟表示再现到虚拟z缓冲器中，其中所述实际对象至少部分在第一掩模中；分辨至少一个已知实际对象与虚拟z缓冲器中的虚拟对象之间的遮挡；及扭曲虚拟z缓冲器以匹配所述图像，所述扭曲部分基于至少一个已知实际对象的虚拟表示与图像中的所述至少一个实际对象之间的光流。

[0116] 在一些实施例中，计算第一掩模(M')中的与被跟踪对象相关联的像素的光流量值可包括将所述图像中的被跟踪对象的像素再现到第二掩模(M)中，其中所述第二掩模指示所述图像中的包括像素的区将被替换。另外，可相对于合成图像子集计算图像中的第一和第二掩模($M'-M$)之间的区中的像素的光流量值。此外，可使用分散数据内插计算第二掩模(M)中的像素的光流量值。过程可随后返回到步骤510以开始下一迭代。

[0117] 在一些实施例中，可部分基于替换像素而再现减弱实境图像。此外，可以各种方式修改方法500。例如，在混合实境应用中，可在由替换像素指示的区中再现虚拟对象。在一些实施例中，在由替换像素指示的区中再现虚拟对象可包括将所述图像中的已知/被跟踪的实际对象的虚拟表示再现到虚拟z缓冲器中，所述虚拟z缓冲器可用于分辨已知实际对象与虚拟对象之间的遮挡。此外，可扭曲所述虚拟z缓冲器以匹配所述图像，所述扭曲可基于至少一个已知实际对象的虚拟表示与图像中的所述至少一个实际对象之间的光流。

[0118] 现在参考图6，其为说明经启用而以与所揭示的实施例一致的方式促进减弱及中介实境效果的服务器600的示意框图。在一些实施例中，服务器600可执行方法300、335及/或500的部分。在一些实施例中，方法300、335及/或500可由处理单元650及/或计算机视觉(CV)模块656执行。例如，可由处理单元650及/或CV模块656结合服务器600上的一或多个功能单元及/或结合MS 100而整体或部分地执行以上方法。

[0119] 在一些实施例中，服务器600可经由无线网络(未图示)(其可为WWAN、WLAN或WPAN中的一者)无线地耦合到一或多个移动台100。在一些实施例中，服务器600可包含(例如)一或多个处理单元650、存储器680、存储装置660及(在适用时)通信接口690(例如，有线或无线网络接口)，其可操作性地与一或多个连接620(例如，总线、线、光纤、链路等)耦合。在某些实例实施方案中，服务器600的某一部分可呈芯片组和/或类似物的形式。

[0120] 通信接口690可包含支持有线传输及/或接收并且在需要时可另外或替代地支持经由一或多种类型的无线通信网络对一或多个信号的传输及接收的多种有线及无线连接。通信接口690可包含用于与MS 100及/或各种其它计算机及外围设备通信的接口。例如，在一个实施例中，通信接口690可包括实施由服务器600执行的通信功能中的一或更多的网络接口卡、输入-输出卡、芯片及/或ASIC。在一些实施例中，通信接口690还可与MS 100介接以发送环境的3D模型信息，及/或接收与方法300及/或500相关的数据及/或指令。

[0121] 处理单元650可使用所接收的信息中的一些或全部来执行所请求的计算及/或经由通信接口690将所请求的信息及/或结果发送到MS 100。在一些实施例中，可使用硬件、固件及软件的组合来实施处理单元650。在一些实施例中，处理单元650可包含计算机视觉(CV)模块656，其可实施及执行计算机视觉方法，包含AR/MR程序、SLAM地图产生等。在一些实施例中，CV模块656可包括3D重构模块658，其可执行3D重构及/或提供/更新环境的3D模型。在一些实施例中，处理单元650可表示可配置以执行数据信号计算程序或与服务器600

的操作相关的过程的至少一部分的一或多个电路。

[0122] 本文在流程图和消息流程中描述的方法可依据应用由各种装置实施。举例来说，这些方法可在硬件、固件、软件或其任何组合中实施。对于硬件实施方案，处理单元650可实施于一或多个专用集成电路(ASIC)、数字信号处理器(DSP)、数字信号处理装置(DSPD)、可编程逻辑装置(PLD)、现场可编程门阵列(FPGA)、处理器、控制器、微控制器、微处理器、电子装置、经设计以执行本文中所描述功能的其它电子单元，或其组合内。

[0123] 对于固件和/或软件实施方案，可用执行本文所描述的功能的模块(例如，程序、功能等等)来实施所述方法。在实施本文中所描述的方法的过程中，可使用任何有形地体现指令的机器可读媒体。例如，软件可存储在可装卸媒体驱动器670中，所述可装卸媒体驱动器可支持非暂时性计算机可读媒体676(包含可装卸媒体)的使用。程序代码可驻留在非暂时性计算机可读媒体676或存储器680上且可由处理单元650读取及执行。存储器可在处理单元650内或处理单元650的外部实施。如本文中所使用，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，且不应限于任何特定类型的存储器或任何特定数目个存储器或存储存储器的媒体的类型。

[0124] 如果在固件及/或软件中实施，则所述功能可作为一或多个指令或代码存储在非暂时性计算机可读媒体676及/或存储器680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。例如，包含在其上存储的程序代码的非暂时性计算机可读媒体676可包含以与所揭示的实施例一致的方式促进来自重构的MR效果(例如减弱及中介实境效果)的程序代码。

[0125] 非暂时性计算机可读媒体可包含多种物理计算机存储媒体。存储媒体可以是可通过计算机存取的任何可用媒体。借助于实例而非限制，此类非暂时性计算机可读媒体可包括RAM、ROM、EEPROM、CD-ROM或其它光盘存储装置，磁盘存储装置或其它磁性存储装置，或任何其它可用于存储呈指令或数据结构的形式的所要程序代码且可由计算机存取的媒体；如本文所使用，磁盘和光盘包含压缩光盘(CD)、激光光盘、光学光盘、数字多功能光盘(DVD)，软性磁盘和蓝光光盘，其中磁盘通常以磁性方式重现数据，而光盘用激光以光学方式重现数据。非暂时性计算机可读媒体的其它实施例包含快闪驱动器、USB驱动器、固态驱动器、存储器卡等。以上各者的组合也应包含在计算机可读媒体的范围内。

[0126] 除了存储在计算机可读媒体上之外，可将指令及/或数据作为传输媒体上的信号提供到通信接口690，所述通信接口可将指令/数据存储在存储器680、存储装置660中，及/或将所述指令/数据中继到处理单元650以用于执行。例如，通信接口690可接收指示指令及数据的无线或网络信号。所述指令和数据经配置以致使一或多个处理器实施权利要求书中概述的功能。即，通信设备包含具有指示用以执行所揭示功能的信息的信号的传输媒体。

[0127] 存储器680可表示任何数据存储机构。存储器680可包含(例如)主存储器及/或辅助存储器。主存储器可包含(例如)随机存取存储器、只读存储器、非易失性RAM等。虽然在此实例中说明为与处理单元650分开，但应理解，主存储器的全部或一部分可提供在处理单元650内或以其它方式与处理单元650位于同一地点/耦合。辅助存储器可包含(例如)与主存储器及/或存储装置660(例如包含(例如)硬盘驱动器、光学光盘驱动器、磁带机、固态存储器驱动器等的一或多个数据存储装置660)相同或类似类型的存储器。

[0128] 在一些实施例中，存储装置660可包括可保持关于环境的信息(包含3D模型、关键

帧、关于虚拟对象的信息等)的一或多个数据库。在一些实施例中,可由处理单元650在各种计算期间读取、使用及/或更新所述数据库中的信息。

[0129] 在某些实施方案中,辅助存储器可操作性地接纳非暂时性计算机可读媒体676或另外可配置以耦合到非暂时性计算机可读媒体676。由此,在某些实例实施方案中,可使用非暂时性计算机可读媒体676整体或部分地实施本文中呈现的方法及/或设备,非暂时性计算机可读媒体676可包含存储在其上的计算机可实施的指令,所述计算机可实施的指令在由至少一个处理单元650执行的情况下可操作性地被启用以执行如本文中所描述的实例操作的全部或部分。在一些实施例中,计算机可读媒体676可使用可装卸媒体驱动器670来读取及/或可形成存储器680的部分。

[0130] 本文中揭示的实施例促进将SLAM系统的基础结构用作计算上高效的混合实境应用(例如对象移除或替换)的基础。在一些实施例中,可通过考虑围绕前景对象的掩模的边界附近的少量像素而获得MR效率及/或性能,所述考虑足以确保跨越掩模边界的连续性,其中来自实况图像的像素邻近于来自合成图像的通过光流计算的结果校正的像素。本文中描述的方法可依据应用由各种装置实施。例如,对于固件和/或软件实施方案,可用执行本文中所描述的功能的模块(例如,程序、功能等)实施所述方法。在实施本文中所描述的方法的过程中,可使用任何有形地体现指令的机器可读媒体。举例来说,软件代码可存储在存储器中,并且由处理器单元来执行。在一些实施例中,可将所述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。

[0131] 提供对所揭示方面的先前描述以使得所属领域的任何技术人员都能够制造或使用本发明。所属领域的技术人员将容易明白对这些方面的各种修改,且在不脱离本发明的精神或范围的情况下,本文中所界定的一般原理可应用于其它方面。

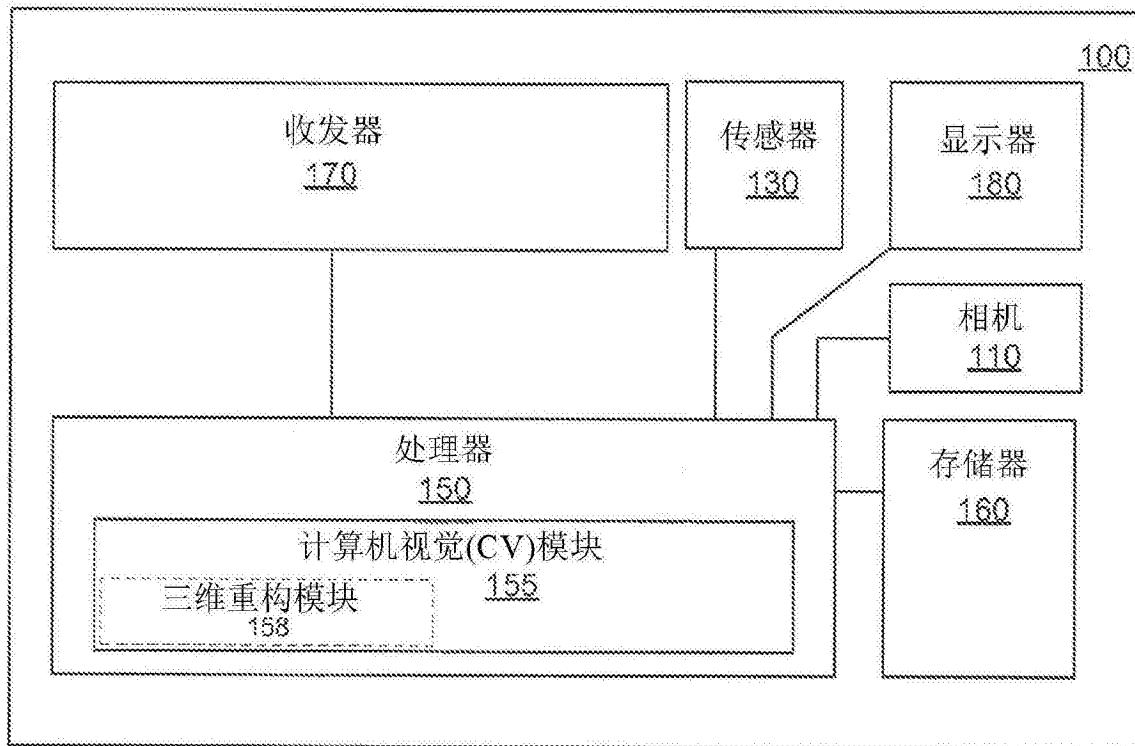


图1

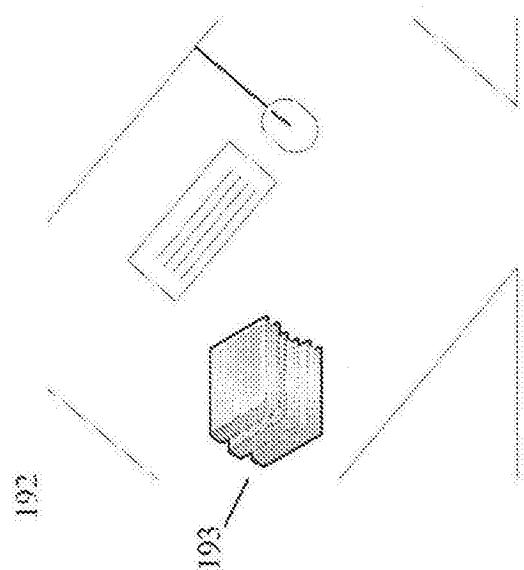


图2A

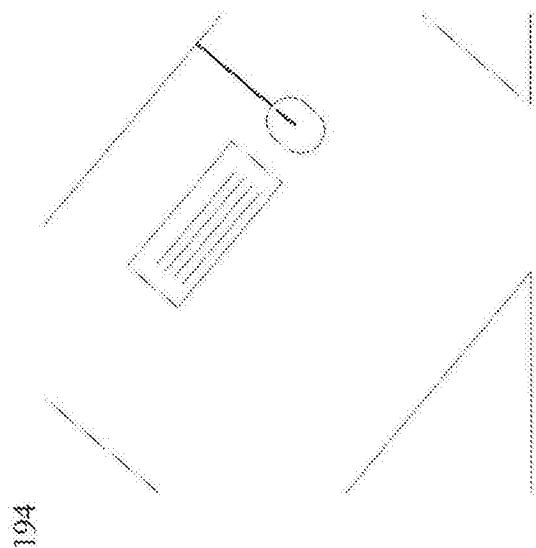


图2B

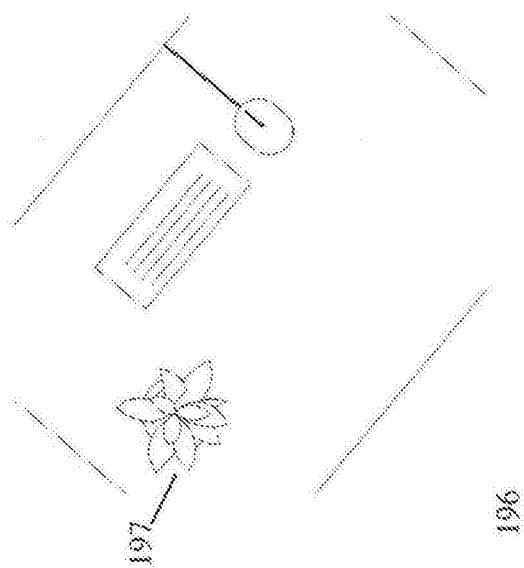


图2C

300

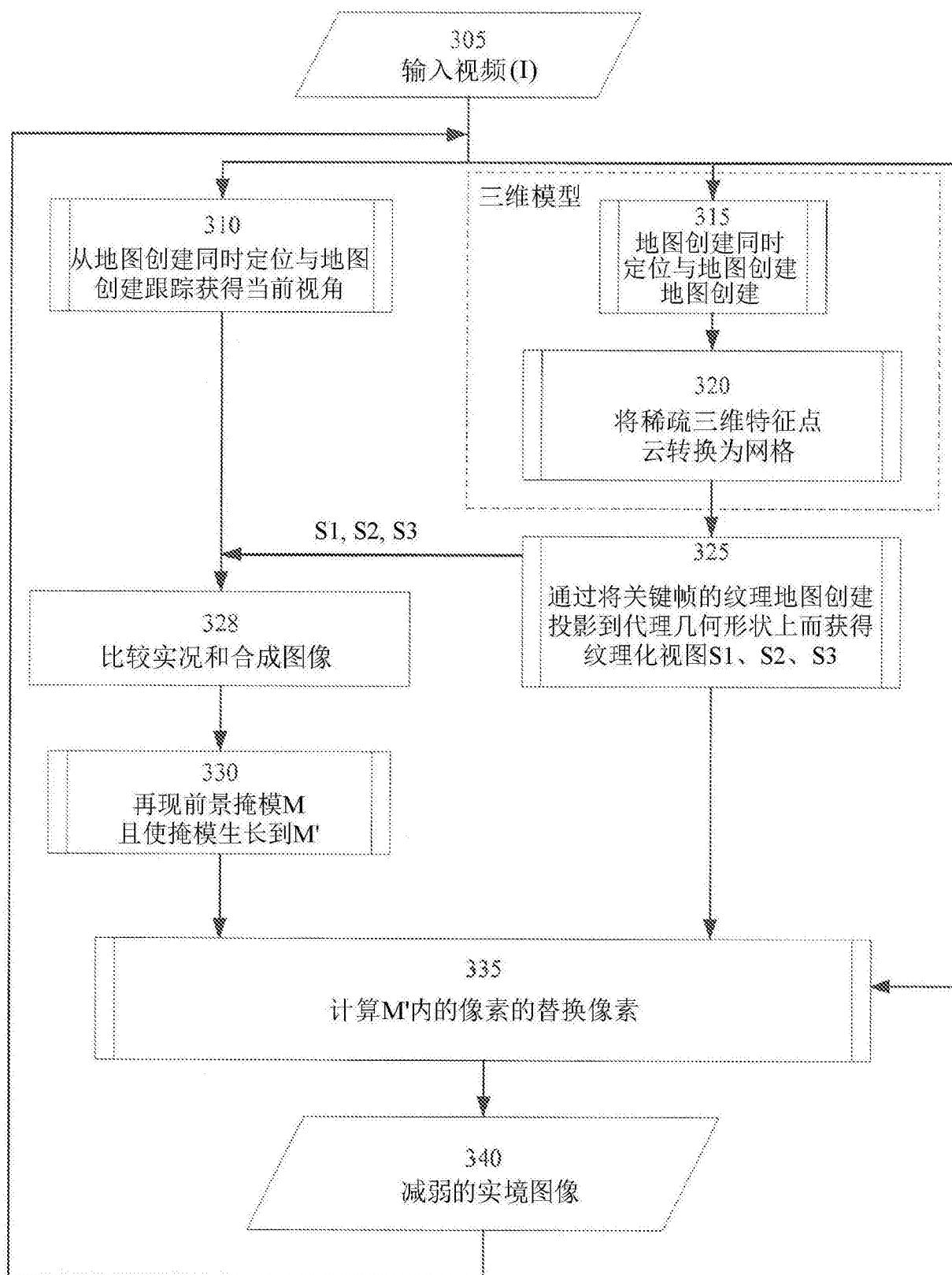


图3A

335

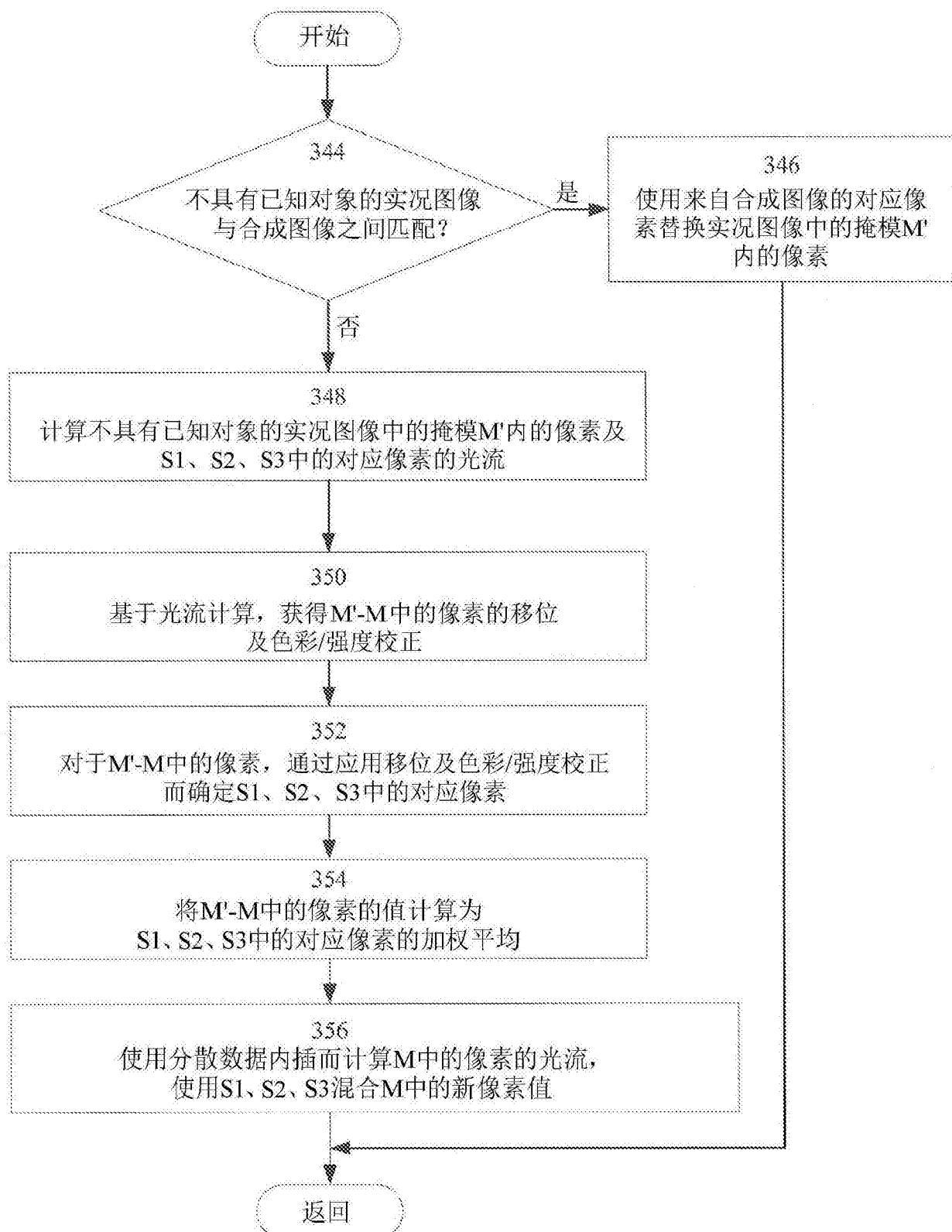


图3B

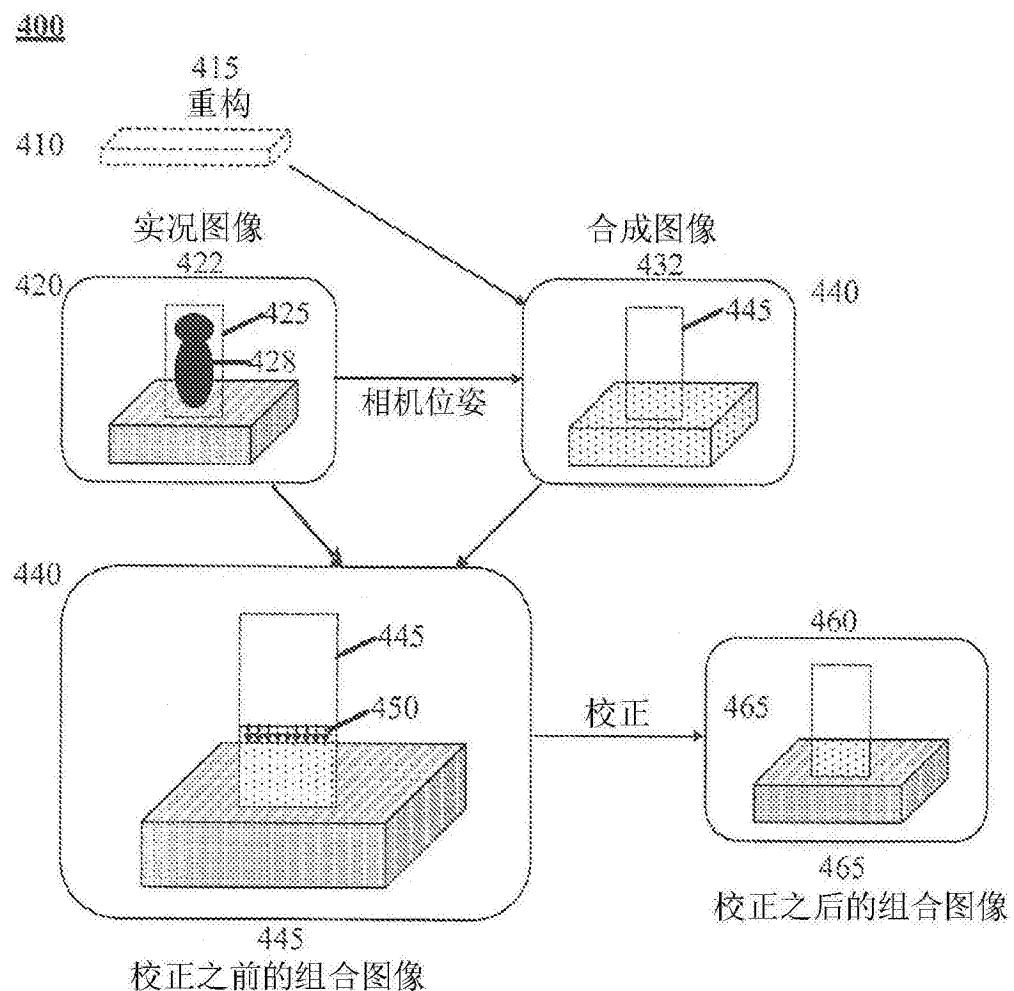


图4

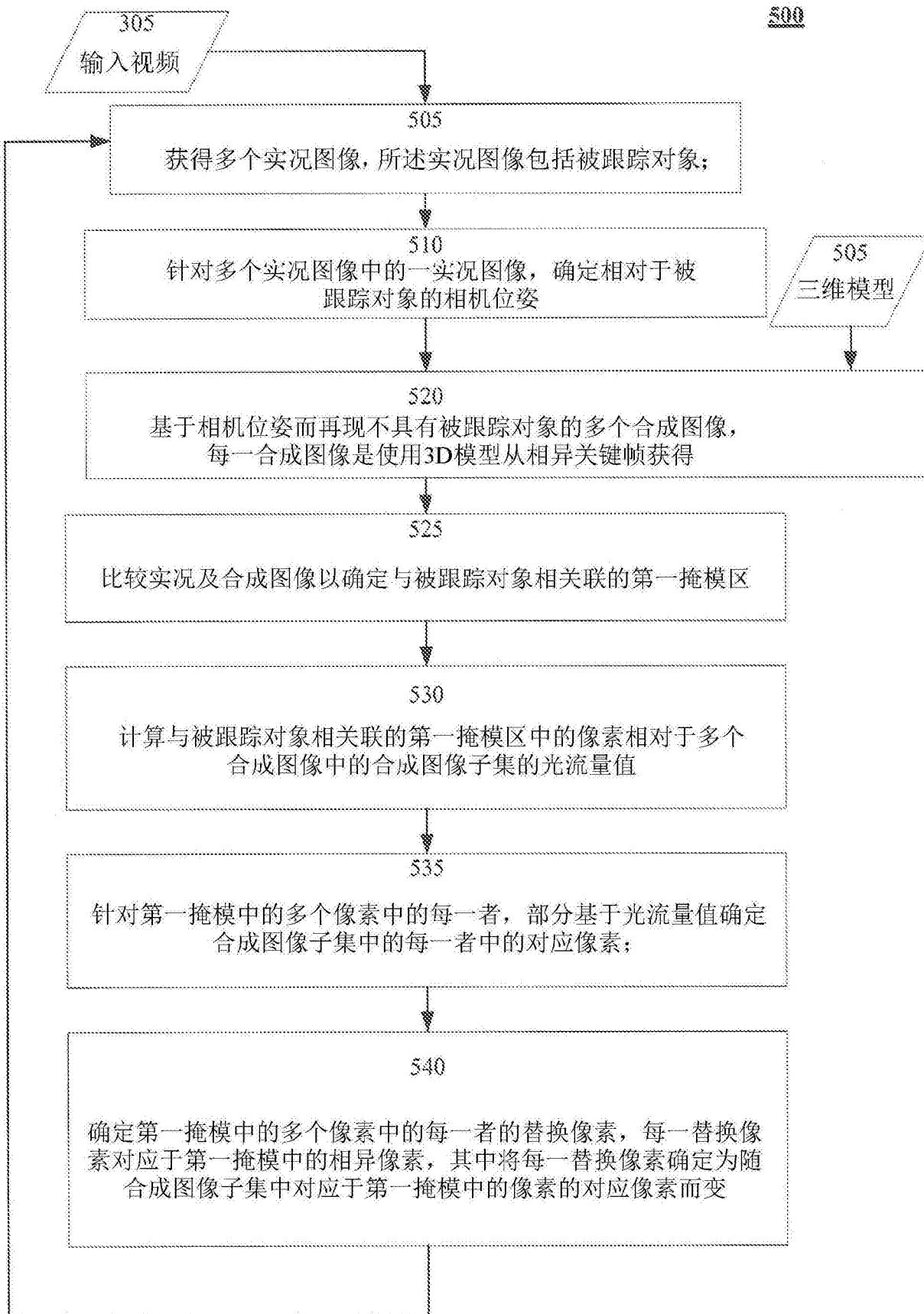


图5

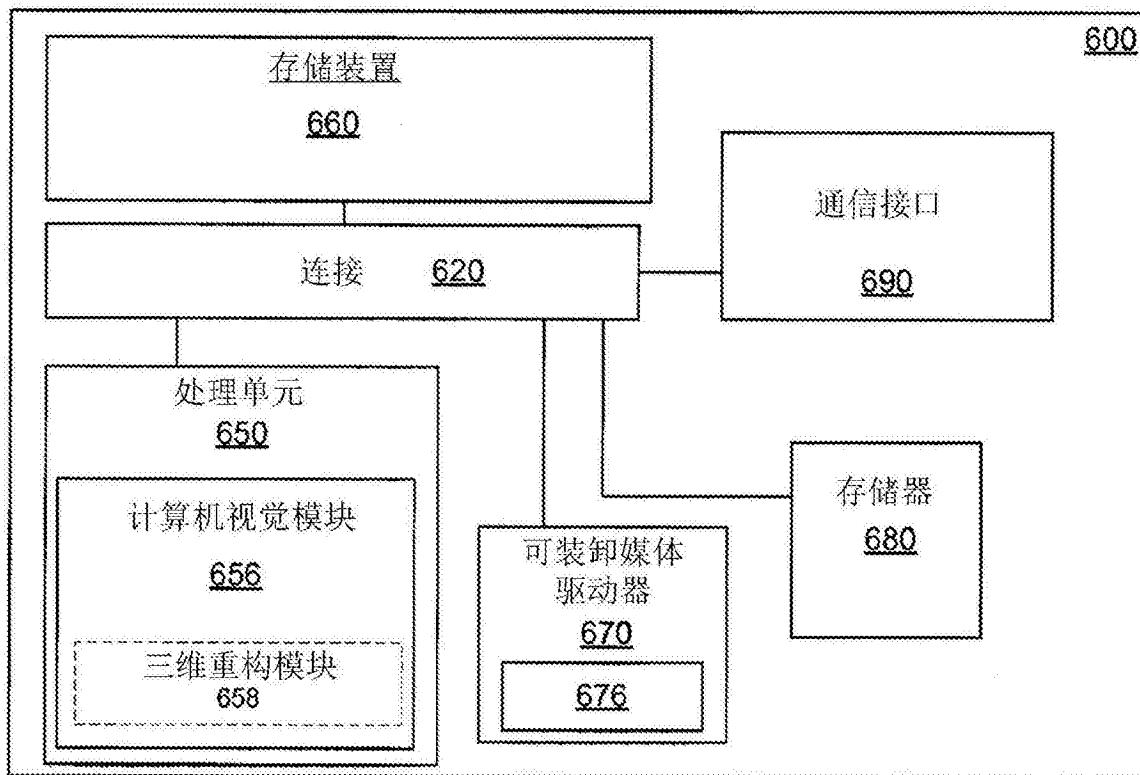


图6