



(12)发明专利

(10)授权公告号 CN 104969225 B

(45)授权公告日 2019.03.19

(21)申请号 201480007660.0

(22)申请日 2014.02.05

(65)同一申请的已公布的文献号

申请公布号 CN 104969225 A

(43)申请公布日 2015.10.07

(30)优先权数据

13/772,640 2013.02.21 US

(85)PCT国际申请进入国家阶段日

2015.08.06

(86)PCT国际申请的申请数据

PCT/US2014/014763 2014.02.05

(87)PCT国际申请的公布数据

W02014/130237 EN 2014.08.28

(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 丹尼尔·瓦格纳 潘琪

(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 17/00(2006.01)

G06F 16/20(2019.01)

(56)对比文件

US 6173070 B1,2001.01.09,

CN 101271469 A,2008.09.24,

US 2013216135 A1,2013.08.22,

审查员 刘雅洁

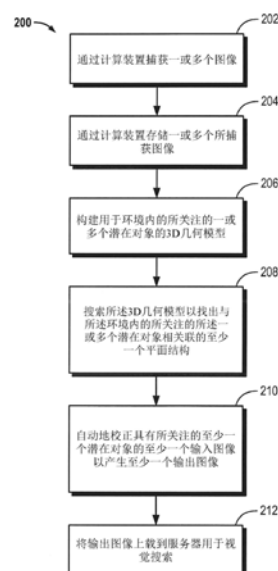
权利要求书5页 说明书8页 附图5页

(54)发明名称

用于视觉搜索的自动图像校正

(57)摘要

本发明揭示一种可以执行用于视觉搜索的自动图像校正的计算装置。在计算装置处实施的方法包含从图像捕获装置接收一或多个图像;通过所述计算装置存储所述一或多个图像;基于所述一或多个图像中的至少一个图像构建用于环境内的所关注的一或多个潜在对象的三维3D几何模型;及自动地产生具有所关注的至少一个潜在对象的至少一个经校正图像用于视觉搜索。



1. 一种在计算装置处实施的方法,其包括:

从图像捕获装置接收一或多个图像,所述一或多个图像包含环境及所述环境内的所关注的一或多个潜在对象的至少一部分;

通过所述计算装置存储所述一或多个图像;

构建包含在所述一或多个图像中的所述环境的至少部分三维3D几何模型,所述至少部分3D几何模型包含所述环境内的所关注的所述一或多个潜在对象的至少一部分,其中所述至少部分3D几何模型使用所述一或多个图像中的至少一个图像构建;及

使用所述至少部分3D几何模型自动地产生具有所关注的至少一个潜在对象的至少一个经校正图像用于视觉搜索,所述至少一个经校正图像包含所关注的所述至少一个潜在对象的经修改正视图,其中所述经修改正视图相对于所述一或多个图像中的所关注的所述至少一个潜在对象是正视的或正交的。

2. 根据权利要求1所述的方法,其进一步包括:

将所述至少一个经校正图像自动地上载到服务器用于所述视觉搜索。

3. 根据权利要求1所述的方法,其进一步包括:

将所述至少一个经校正图像上载到服务器用于所述视觉搜索。

4. 根据权利要求1所述的方法,其中所述一或多个图像通过与所述计算装置相关联的图像捕获装置捕获,其中在不接收用户输入的情况下发生自动地产生具有所关注的至少一个潜在对象的至少一个经校正图像用于所述视觉搜索。

5. 根据权利要求1所述的方法,其中在后台操作期间所述一或多个图像通过与所述计算装置相关联的图像捕获装置自动地捕获而不接收用户输入。

6. 根据权利要求1所述的方法,其中在某一时间段期间所述一或多个图像通过与所述计算装置相关联的图像捕获装置捕获,其中所述图像捕获装置在所述时间段的至少一部分内移动。

7. 根据权利要求1所述的方法,其中所述一或多个图像通过与所述计算装置相关联的图像捕获装置捕获,其中所述图像捕获装置能够提供所关注的所述至少一个潜在对象的深度信息且在所述图像捕获装置不需要移动时捕获所述一或多个图像。

8. 根据权利要求1所述的方法,其进一步包括:

搜索所述环境的所述至少部分3D几何模型以找出与所述环境内的所关注的所述一或多个潜在对象的至少一部分相关联的至少一个平面结构。

9. 根据权利要求1所述的方法,其中所述环境的所述至少部分3D几何模型使用结构运动系统构建。

10. 根据权利要求1所述的方法,其进一步包括:

搜索所述环境的所述至少部分3D几何模型以找出与所述环境内的所关注的所述一或多个潜在对象的至少一部分相关联的至少两个平面结构。

11. 根据权利要求1所述的方法,其中自动地产生具有所关注的至少一个潜在对象的至少一个经校正图像用于所述视觉搜索包括将一个平面结构展开成每一经校正图像的正视图。

12. 根据权利要求1所述的方法,其中自动地产生具有所关注的至少一个潜在对象的至少一个经校正图像包括校正来自多个输入图像的所关注的所述至少一个潜在对象的部分

以产生所述至少一个经校正图像。

13. 一种计算装置,其包括:

存储媒体,其用以存储一或多个所接收图像,所述一或多个所接收图像包含环境及所述环境内的所关注的一或多个潜在对象的至少一部分;及

处理电路,其耦合到所述存储媒体,所述处理电路经配置以执行用以进行以下操作的指令:构建包含在所述一或多个图像中的所述环境的至少部分三维3D几何模型,所述至少部分3D几何模型包含所述环境内的所关注的所述一或多个潜在对象的至少一部分,其中所述至少部分3D几何模型使用所述一或多个所存储图像中的至少一个图像构建;使用所述至少部分3D几何模型自动地产生具有所关注的至少一个潜在对象的至少一个经校正图像用于视觉搜索,所述至少一个经校正图像包含所关注的所述至少一个潜在对象的经修改正视图,其中所述经修改正视图相对于所述一或多个图像中的所关注的所述至少一个潜在对象是正视的或正交的。

14. 根据权利要求13所述的计算装置,其中所述处理电路进一步经配置以执行用以将所述至少一个经校正图像自动地上载到服务器用于所述视觉搜索的指令。

15. 根据权利要求13所述的计算装置,其中所述处理电路进一步经配置以执行用以将所述至少一个经校正图像上载到服务器用于所述视觉搜索的指令。

16. 根据权利要求13所述的计算装置,其进一步包括:

图像捕获装置,其耦合到所述存储媒体,所述图像捕获装置用以捕获所述一或多个图像,其中所述存储媒体用以从所述图像捕获装置接收所述一或多个图像。

17. 根据权利要求16所述的计算装置,其中所述处理电路经配置以执行用以自动地产生具有所关注的至少一个潜在对象的至少一个经校正图像用于所述视觉搜索而不接收用户输入的指令。

18. 根据权利要求16所述的计算装置,其中在后台操作期间所述图像通过所述图像捕获装置自动地捕获而不接收用户输入。

19. 根据权利要求16所述的计算装置,其中在某一时间段期间所述一或多个图像通过所述图像捕获装置捕获,其中所述图像捕获装置在所述时间段的至少一部分内移动。

20. 根据权利要求16所述的计算装置,其中所述图像捕获装置能够提供所关注的所述至少一个潜在对象的深度信息且在所述图像捕获装置不需要移动时捕获所述一或多个图像。

21. 根据权利要求13所述的计算装置,其中所述处理电路进一步经配置以执行用以搜索所述环境的所述至少部分3D几何模型以找出与所述环境内的所关注的所述一或多个潜在对象的至少一部分相关联的至少一个平面结构的指令。

22. 根据权利要求13所述的计算装置,其中所述环境的所述至少部分3D几何模型使用结构运动系统构建。

23. 根据权利要求13所述的计算装置,其中所述处理电路进一步经配置以执行用以搜索所述环境的所述至少部分3D几何模型以找出与所述环境内的所关注的所述一或多个潜在对象的至少一部分相关联的至少两个平面结构的指令。

24. 根据权利要求13所述的计算装置,其中自动地产生具有所关注的至少一个潜在对象的至少一个经校正图像用于所述视觉搜索包括将一个平面结构展开成每一经校正图像

的正视图。

25. 根据权利要求13所述的计算装置,其中自动地产生具有所关注的至少一个潜在对象的至少一个经校正图像包括校正来自多个输入图像的所关注的所述至少一个潜在对象的部分以产生所述至少一个经校正图像。

26. 一种在计算装置处执行的计算机程序产品,其包括:

计算机可读媒体,其包括用于进行以下操作的代码:

通过所述计算装置存储一或多个图像,所述一或多个图像包含环境及所述环境内的所关注的一或多个潜在对象的至少一部分;

通过所述计算装置构建包含在所述一或多个图像中的所述环境的至少部分三维3D几何模型,所述至少部分3D几何模型包含所述环境内的所关注的所述一或多个潜在对象的至少一部分,其中所述至少部分3D几何模型使用所述一或多个图像中的至少一个图像构建;及

使用所述至少部分3D几何模型自动地产生具有所关注的至少一个潜在对象的至少一个经校正图像用于视觉搜索,所述至少一个经校正图像包含所关注的所述至少一个潜在对象的经修改正视图,其中所述经修改正视图相对于所述一或多个图像中的所关注的所述至少一个潜在对象是正视的或正交的。

27. 根据权利要求26所述的计算机程序产品,其进一步包括用于将所述至少一个经校正图像自动地上载到服务器用于所述视觉搜索的代码。

28. 根据权利要求26所述的计算机程序产品,其进一步包括用于将所述至少一个经校正图像上载到服务器用于所述视觉搜索的代码。

29. 根据权利要求26所述的计算机程序产品,其中所述一或多个图像通过与所述计算装置相关联的图像捕获装置捕获,其中在不接收用户输入的情况下发生自动地产生具有所关注的至少一个潜在对象的至少一个经校正图像用于所述视觉搜索。

30. 根据权利要求26所述的计算机程序产品,其中在后台操作期间所述一或多个图像通过与所述计算装置相关联的图像捕获装置捕获而不接收用户输入。

31. 根据权利要求26所述的计算机程序产品,其中在某一时间段期间所述一或多个图像通过与所述计算装置相关联的图像捕获装置捕获,其中所述图像捕获装置在所述时间段的至少一部分内移动。

32. 根据权利要求26所述的计算机程序产品,其中所述一或多个图像通过与所述计算装置相关联的图像捕获装置捕获,其中所述图像捕获装置能够提供所关注的所述至少一个潜在对象的深度信息且在所述图像捕获装置不需要移动时捕获所述一或多个图像。

33. 根据权利要求26所述的计算机程序产品,其进一步包括用于搜索所述环境的所述至少部分3D几何模型以找出与所述环境内的所关注的所述一或多个潜在对象的至少一部分相关联的至少一个平面结构的代码。

34. 根据权利要求26所述的计算机程序产品,其中所述环境的所述至少部分3D几何模型使用结构运动系统构建。

35. 根据权利要求26所述的计算机程序产品,其进一步包括用于搜索所述环境的所述至少部分3D几何模型以找出与所述环境内的所关注的所述一或多个潜在对象的至少一部分相关联的至少两个平面结构的代码。

36. 根据权利要求26所述的计算机程序产品,其中自动地产生具有所关注的至少一个潜在对象的至少一个经校正图像用于所述视觉搜索包括将一个平面结构展开成每一经校正图像的正视图。

37. 根据权利要求26所述的计算机程序产品,其中自动地产生具有所关注的至少一个潜在对象的至少一个经校正图像包括校正来自多个输入图像的所关注的所述至少一个潜在对象的部分以产生所述至少一个经校正图像。

38. 一种图像校正设备,其包括:

用于存储一或多个所接收图像的装置,所述一或多个所接收图像包含环境及所述环境内的所关注的一或多个潜在对象的至少一部分;

用于构建包含在所述一或多个所接收图像中的所述环境的至少部分三维3D几何模型的装置,所述至少部分3D几何模型包含所述环境内的所关注的所述一或多个潜在对象的至少一部分,其中所述至少部分3D几何模型使用所述一或多个所接收图像中的至少一个图像构建;及

用于使用所述至少部分3D几何模型自动地产生具有所关注的至少一个潜在对象的至少一个经校正图像用于视觉搜索的装置,所述至少一个经校正图像包含所关注的所述至少一个潜在对象的经修改正视图,其中所述经修改正视图相对于所述一或多个图像中的所关注的所述至少一个潜在对象是正视的或正交的。

39. 根据权利要求38所述的图像校正设备,其进一步包括:

用于将所述至少一个经校正图像自动地上载到服务器用于所述视觉搜索的装置。

40. 根据权利要求38所述的图像校正设备,其进一步包括:

用于将所述至少一个经校正图像上载到服务器用于所述视觉搜索的装置。

41. 根据权利要求38所述的图像校正设备,进一步包括:

用于捕获所述一或多个图像的装置,其中所述用于存储的装置用以从所述用于捕获的装置接收所述一或多个图像。

42. 根据权利要求41所述的图像校正设备,其中在不接收用户输入的情况下发生自动地产生具有所关注的至少一个潜在对象的至少一个经校正图像用于所述视觉搜索。

43. 根据权利要求41所述的图像校正设备,其中在后台操作期间自动地捕获所述一或多个图像而不接收用户输入。

44. 根据权利要求41所述的图像校正设备,其中在某一时间段期间捕获所述一或多个图像,其中所述用于捕获所述一或多个图像的装置在所述时间段的至少一部分内移动。

45. 根据权利要求41所述的图像校正设备,其中所述用于捕获的装置能够提供所关注的所述至少一个潜在对象的深度信息且在所述用于捕获的装置不需要移动时捕获所述一或多个图像。

46. 根据权利要求38所述的图像校正设备,其进一步包括:

用于搜索所述环境的所述至少部分3D几何模型以找出与所述环境内的所关注的所述一或多个潜在对象的至少一部分相关联的至少一个平面结构的装置。

47. 根据权利要求38所述的图像校正设备,其进一步包括:

用于搜索所述环境的所述至少部分3D几何模型以找出与所述环境内的所关注的所述一或多个潜在对象的至少一部分相关联的至少两个平面结构的装置。

48. 根据权利要求38所述的图像校正设备, 其中所述用于自动地产生具有所关注的至少一个潜在对象的至少一个经校正图像用于所述视觉搜索的装置包括将一个平面结构展开成每一经校正图像的正视图。

49. 根据权利要求38所述的图像校正设备, 其中所述用于自动地产生具有所关注的至少一个潜在对象的至少一个经校正图像的装置包括校正来自多个输入图像的所关注的所述至少一个潜在对象的部分以产生所述至少一个经校正图像。

50. 根据权利要求38所述的图像校正设备, 其中所述三维3D几何模型使用结构运动系统构建。

用于视觉搜索的自动图像校正

技术领域

[0001] 本发明大体上涉及一种能够进行用于视觉搜索的自动图像校正的计算装置。

背景技术

[0002] 用户正从各种不同计算装置(例如,移动装置及非移动装置两者)访问例如视觉搜索服务等不同服务。举例来说,这些不同计算装置包含家用计算机、工作计算机、移动电话、移动装置、平板计算机等。视觉搜索已变成受欢迎的服务。用户将图像上载到相对于存储在其数据库中的其它图像与所述图像匹配的服务器,并且最终返回关于经上载图像的信息。相对于数据库图像与查询图像匹配的算法通常经设计,使得它们可以处理一定量的变形(即,转换、尺度、旋转及视角效果)。此外,这些方法可以检测图片中的含有额外不相关细节(即,杂波)的对象。然而,即使算法能够处理这些难题,所述算法也需要更多时间及处理资源来处理这些难题。相机图像中的不相关杂波不仅使服务器更难以找到所关注的对象,而且还增加发送到服务器的图像的大小。

发明内容

[0003] 本发明的各方面可以涉及一种可以执行用于视觉搜索的自动图像校正的计算装置。所述计算装置可以包含存储媒体,其存储一或多个图像;及处理电路,其经配置以执行用以基于所述一或多个图像中的至少一个图像构建用于环境内的所关注的一或多个潜在对象的三维(3D)几何模型的指令。所述处理电路还经配置以执行用以自动地产生具有所关注的至少一个潜在对象的至少一个经校正图像用于视觉搜索的指令。

[0004] 本发明的各方面还可以涉及一种在计算装置处实施的方法。所述方法包含从图像捕获装置接收一或多个图像;通过所述计算装置存储所述一或多个图像;基于所述一或多个图像中的至少一个图像构建用于环境内的所关注的一或多个潜在对象的三维(3D)几何模型;及自动地产生具有所关注的至少一个潜在对象的至少一个经校正图像用于视觉搜索。

[0005] 本发明的各方面还可以涉及一种在计算装置处执行的计算机程序产品。所述计算机程序产品包含计算机可读媒体,所述计算机可读媒体包含用于执行以下操作的代码:通过所述计算装置存储一或多个图像;基于所述一或多个图像中的至少一个图像通过所述计算装置构建用于环境内的所关注的一或多个潜在对象的三维(3D)几何模型;及自动地产生具有所关注的至少一个潜在对象的至少一个经校正图像用于自动搜索。

[0006] 本发明的各方面还可以涉及一种设备,所述设备包含用于存储一或多个所接收图像的装置;用于基于所述一或多个所存储图像中的至少一个图像构建用于环境内的所关注的一或多个潜在对象的三维(3D)几何模型的装置;及用于自动地产生具有所关注的至少一个潜在对象的至少一个经校正图像用于视觉搜索的装置。

[0007] 本发明的各方面还可以涉及一种用于执行视觉搜索的服务器。所述服务器可以包含存储媒体,其存储图像;及处理电路,其经配置以执行用以从用于视觉搜索的计算装置中

接收具有所关注的至少一个潜在对象的至少一个经校正图像且提取表示所述至少一个经校正图像的特征的描述符的指令。所述至少一个经校正图像的所述经提取描述符可以设计成旋转、尺度及光照不变,而不需要视角或仿射失真不变。所述处理电路可以进一步经配置以执行用以将所述至少一个经校正图像的所述经提取描述符与存储在数据库中的图像的描述符匹配的指令。

[0008] 本发明的各方面还可以涉及一种在服务器处实施的方法。所述方法包含通过所述服务器存储多个图像;从用于视觉搜索的计算装置接收具有所关注的至少一个潜在对象的至少一个经校正图像;及提取表示所述至少一个经校正图像的特征的描述符。所述至少一个经校正图像的所述经提取描述符可以设计成旋转、尺度及光照不变,而不需要视角或仿射失真不变。

[0009] 本发明的各方面还可以涉及一种在服务器处执行的计算机程序产品。所述计算机程序产品包含计算机可读媒体,所述计算机可读媒体包含用于执行以下操作的代码:存储多个图像;从用于视觉搜索的计算装置接收具有所关注的至少一个潜在对象的至少一个经校正图像;及提取表示所述至少一个经校正图像的特征的描述符。所述至少一个经校正图像的所述经提取描述符可以设计成旋转、尺度及光照不变,而不需要视角或仿射失真不变。

附图说明

[0010] 图1是具有用于自动地产生至少一个经校正图像的计算装置的系统的方框图。

[0011] 图2是用以说明用于自动地校正通过计算装置捕获的图像的过程的流程图。

[0012] 图3是如通过计算装置捕获的广告牌的图像。

[0013] 图4是在所述图像经校正之后的广告牌的图像的正视图。

[0014] 图5说明用于执行视觉搜索的服务器。

[0015] 图6是用以说明用于执行视觉搜索的在服务器处实施的过程的流程图。

具体实施方式

[0016] 词语“示例性”或“实例”在本文中用于表示“充当实例、例子或说明”。本文中描述为“示例性”或描述为“实例”的任何方面或实施例未必应被解释为比其它方面或实施例优选或有利。

[0017] 图1是具有用于自动地产生至少一个经校正图像的计算装置的系统的方框图。具体而言,系统100说明计算装置101,其可以自动地产生至少一个经校正图像,使得视觉搜索服务具有正确地检测用户的所关注的一或多个对象的更高机会。计算装置101可以包含处理电路110、用以存储指令120及图像122的存储媒体112、电源装置114、显示装置116、用户接口118、收发器119及用于捕获图像(例如,数字静态图像、形成视频的图像序列)的图像捕获装置144(例如,相机、摄像机等)。在另一实施例中,所述图像捕获装置位于计算装置外部。所述图像捕获装置可以与计算装置相关联且以通信方式耦合到所述计算装置。举例来说,计算装置(例如,移动装置)可处于用户的口袋中且无线地连接到安装在用户的一副眼镜上的图像捕获装置(例如,相机)。示例性存储媒体(例如,计算机可读媒体)耦合到处理电路,使得处理器可以从存储媒体读取信息及将信息写入到存储媒体。在替代方案中,存储媒体可以与处理电路成一体式。应了解,显示装置116可以是计算装置101上的典型显示装置,

所述计算装置例如,移动装置、蜂窝电话、个人数字助理、移动计算机、平板计算机等。用户接口118可以是键盘、触摸屏或另一种类型的用户接口输入装置。此外,电源装置114可以是用以为计算装置101供电的电池装置。收发器119可以用于通过无线链路130将呼叫及数据传输到无线网络131/从无线网络131接收呼叫及数据。

[0018] 具体而言,计算装置101可以包含处理电路110,其经配置以执行用以基于一或多个所捕获图像中的至少一个图像构建用于环境内的所关注的一或多个潜在对象的三维(3D)几何模型且自动地产生具有所关注的至少一个潜在对象的至少一个经校正图像用于视觉搜索的指令120。处理电路110可以进一步经配置以执行用以将至少一个经校正图像自动地上载到服务器用于视觉搜索的指令。举例来说,计算装置的用户可以不提供用于将经校正图像上载到服务器的用户输入。或者,用户可以意识到经校正图像正上载到服务器且用户可以手动地选择用于将经校正图像上载到服务器用于视觉搜索的选项。

[0019] 在一个实施例中,图像捕获装置在后台操作期间捕获图像并且位于所捕获图像中的所关注的潜在对象的至少一个经校正图像自动地被产生用于视觉搜索而不接收用户输入。图像捕获装置可以在后台操作期间捕获图像,而不从用户接收输入用于捕获图像。

[0020] 在一些方面中,图像捕获装置在某一时间段期间捕获图像,其中计算装置在所述时间段的至少一部分内移动。在其它方面中,图像捕获装置能够提供所关注的至少一个潜在对象的深度信息且在计算装置由于所述深度信息而不需要移动时捕获图像。或者,对于位于计算装置外部的图像捕获装置,图像捕获装置在某一时间段期间捕获图像,其中图像捕获装置在所述时间段的至少一部分内移动。在其它替代方面中,图像捕获装置提供所关注的至少一个潜在对象的深度信息且在图像捕获装置由于所述深度信息而不需要移动时捕获图像。

[0021] 应了解,如下文将描述的本发明的各方面可以结合由计算装置101的处理电路110和/或计算装置101和/或其它装置的其它电路执行的指令来实施。具体地,计算装置101的电路包含但不限于处理电路110,其可以在程序、例程的控制下或在指令的执行下运行以执行根据本发明的实施例的方法或过程。举例来说,此程序可以在固件或软件中加以实施(例如,存储在存储媒体112和/或其它位置中)且可以由处理器来实施,例如,处理电路110和/或计算装置101的其它电路。此外,应了解,术语“处理电路”、“处理器”、“微处理器”、“电路”、“控制器”等是指任何类型的逻辑或能够执行逻辑、命令、指令、软件、固件、功能性等的电路。

[0022] 此外,计算装置101可以通过无线网络131经由一或多个无线通信链路130通信,所述无线网络基于或以其它方式支持任何合适的无线通信技术。举例来说,在一些方面中,计算装置101可以与包含无线网络131的网络相关联。在一些方面中,网络可以包括人体局域网络或个人局域网络(例如,超宽带网络)。在一些方面中,网络可以包括局域网或广域网。无线装置可以支持或以其它方式使用多种无线通信技术、协议或标准(例如,CDMA、TDMA、OFDM、OFDMA、WiMAX及Wi-Fi等)中的一或多个者。类似地,无线装置可以支持或以其它方式使用多种对应调制或多路复用方案中的一或多个者。无线装置因此可以包含适当组件(例如,空中接口)以使用以上或其它无线通信技术建立一或多个无线通信链路及经由一或多个无线通信链路通信。举例来说,装置可以包含具有相关联发射器及接收器组件(例如,发射器及接收器)的无线收发器,所述无线收发器可以包含促进在无线媒体上的通信的多个组件(例

如,信号产生器及信号处理器)。众所周知,计算装置101因此可以无线方式与其它移动装置、蜂窝电话、其它有线及无线计算机、因特网网站等通信。

[0023] 另外参考图2,展示流程图以说明用于自动地产生经校正图像的过程200。在一个实施例中,此过程通过将经校正图像自动地产生到正视图中使得视觉搜索服务具有正确地检测所关注的对象的更高机会来改进视觉搜索匹配结果。另外,此过程可以将来自多个图片的对象的部分自动地组合成单个图像。用于这些改进的基础是除了仅使用对于图片的光度测量之外收集关于场景的3D几何知识的系统。此系统构建环境的3D几何模型(例如,环境的密集3D点云映射)。

[0024] 在块202处,与计算装置相关联的图像捕获装置捕获一或多个图像。图像捕获装置可以与计算装置一体化或位于计算装置外部。在块204处,计算装置存储一或多个所捕获图像(例如,将图像存储在存储媒体中)。在块206处,所述过程基于一或多个所捕获图像中的至少一者构建用于环境内的所关注的一或多个潜在对象的3D几何模型。对于单目相机,运动结构系统提供3D几何模型的此类几何信息。在一些情况下,计算装置在某一时间段期间捕获图像,其中计算装置在所述时间段的至少一部分内移动。举例来说,用户可以平移环境内的计算装置以捕获环境的图像。可以存在于深度相机(例如,RGB-D相机)上的其它传感器可以用于直接提供深度而不需要运动。在这些情况下,计算装置提供所关注的至少一个潜在对象的深度信息且在计算装置不需要移动时捕获图像。深度相机提供所捕获像素的深度信息。深度相机可以从每一对象的表面感测反射光。出于实际目的,需要实时(例如,立即、几乎立即)重构场景的几何结构的实时系统。近来,同时定位与映射(SLAM)系统已变得足够高效和稳固以实际用于包含移动电话的计算装置上。通过SLAM系统,用户指向在所关注的对象处的相机且开始移动。在用户与相机一起移动时,SLAM系统追踪相机图像中的细节且构建环境的几何模型。在配备有RGB-D相机的装置的情况下,单个图像足以构建可以通过运动伸长的初始几何模型。所关注的对象不必完全可见,用户也不必位于所述所关注的对象的正前方。大部分对象在相机图像中的一些点处可见是足够的。SLAM系统还获取其需要用于其内部目的的图片(例如,关键帧),同时SLAM系统构建环境的几何模型。可以周期性地(例如,每隔2至5秒)获取关键帧以确定环境中的相机位置。

[0025] 在一个实施例中,计算装置捕获一或多个图像且自动地产生位于一或多个所捕获图像中的所关注的潜在对象的至少一个经校正图像用于视觉搜索,而不接收用户输入。举例来说,用户将相机朝向所关注的对象对准且相机捕获所关注的对象及其它所关注的潜在对象的图像。计算装置随后自动地校正所关注的对象及其它所关注的潜在对象并且经校正图像可以用于视觉搜索。计算装置的用户接口可以不指示经自动校正的图像将用于视觉搜索。

[0026] 在一些情况下,计算装置在后台操作期间捕获图像而不接收用于捕获图像的用户输入。在后台操作期间,计算装置的用户接口可以不指示图像被捕获或经校正图像将用于视觉搜索。在块208处,在可能在短时间段中(例如,至多几秒)出现的基于3D几何模型至少部分完成环境的重构后,所述过程搜索3D几何模型以找出与环境内的所关注的一或多个潜在对象相关联的至少一个平面结构(例如,主要平面结构)。这些所关注的潜在对象可以包含基于用户将相机朝向对象对准而捕获的对象及此外位于环境内的其它对象。以此方式,与可能已由用户既定或完全看到的所关注的潜在对象的数目相比,所述过程搜索用于平面

结构的更广泛数目的所关注的潜在对象。部分完成的重构可足够用于搜索且随后稍后的搜索可以搜索稍后已完成的环境的其它部分。计算装置可以具有用于每一平面结构的可信度量。在块210处,所述过程以正视图或正交视图产生(例如,自动地产生)具有所关注的至少一个潜在对象的至少一个经校正图像用于视觉搜索。产生经校正输出图像可以包含通过SLAM系统将平面结构的输入图像从之前获取的图片展开,以便产生平面结构的经校正视图。用户可能已既定捕获一些经校正输出图像,而其它经校正输出图像可能尚未既定被捕获。用户可能不知道经校正输出图像中的任一者或一些将用于视觉搜索或计算装置可能不接收用于视觉搜索的用户输入。

[0027] 在已产生经校正图像之后,在块212处,系统可以向用户呈现所述经校正图像以进行确认或立即将所述经校正图像自动地上载到视觉搜索服务器。在现今带宽通常可用于移动电话的情况下,上载图像及接收搜索结果通常仅花费几秒。因此,自动方法通常将是优选的。如结合图5及图6所论述,使用经校正图像可以潜在地加速在服务器侧上所需的描述符提取及匹配。

[0028] 图3及图4说明校正输入图像300以产生输出图像400。图3是如通过计算装置捕获的广告牌302的图像300。在图像300的捕获期间,计算装置相对于广告牌302不正交或垂直。相反地,计算装置以某一角度倾斜向上以便捕获图像300。图4是在计算装置已校正图像300之后广告牌402的图像400的正视图。对于此任务,广告牌的平面在单个图像中是否完全可见或广告牌的平面是否跨越多个图像是无关紧要的。计算装置可以将图像400发送到服务器用于视觉搜索。

[0029] 在实施例中,用户构建整个房间的3D几何模型。举例来说,房间可以包含多个小平面对象(例如,图片、海报、产品盒子、杂志等)。构建3D几何模型的计算装置还定位小平面对象的平面结构、校正这些平面结构且将所述平面结构发送到服务器用于视觉搜索而用户不必识别及选择对象。

[0030] 在另一实施例中,用户不能够获取覆盖足够大部分的所关注的对象的单个图片。这在相机不具有宽视野(与移动电话相机的情况一样)且不存在用以足够远离对象移动使得其在相机图像中完全可见的足够空间时发生。举例来说,计算装置可以重构包含具有大型绘画的窄走廊的环境,使得计算装置(例如,移动电话)的相机无法在单次相机拍摄(不具有太陡峭的角度)中捕获整个绘画。计算装置检测覆盖整个绘画的大平面、将来自多个所捕获输入图像的绘画校正成一个经校正输出图像且将所述经校正输出图像发送到服务器用于视觉搜索。

[0031] 在一个实施例中,自动地产生具有所关注的至少一个潜在对象的至少一个经校正图像用于视觉搜索包含将一个平面结构展开成每一输出图像的正视图。在另一个实施例中,自动地产生具有所关注的至少一个潜在对象的至少一个经校正图像包含校正来自多个输入图像的所关注的至少一个潜在对象的部分以产生至少一个输出图像。

[0032] 图5说明用于执行视觉搜索的服务器500。服务器500包含存储媒体510,其具有数据库512以存储图像;及处理电路520,其经配置以执行用以经由网络接口540接收至少一个经校正图像的指令522。网络接口可以通过有线或无线链路耦合到局域网或广域网。至少一个经校正图像包含如通过计算装置捕获用于视觉搜索的所关注的至少一个潜在对象。处理电路520经配置以执行用以提取表示至少一个经校正图像的特征(例如,高对比度区域)的

描述符的指令。一系列数字(例如,128个数字)可以表示特征。至少一个经校正图像的经提取描述符设计成旋转、尺度及光照不变,而不需要视角或仿射失真不变。处理电路520进一步经配置以执行用以将至少一个经校正图像的经提取描述符与存储在数据库中的图像中的任一者的描述符匹配的指令522。处理电路520进一步经配置以执行用以传输与存储在数据库中的具有与经提取描述符匹配的描述符的至少一个图像相关联的信息的指令522。在一个实施例中,至少一个经校正图像的经提取描述符在视角或仿射失真上是变化的。示例性存储媒体(例如,具有软件代码或指令的计算机可读媒体)耦合到处理电路,使得处理电路可以从存储媒体读取信息及将信息写入到存储媒体。在替代方案中,存储媒体可以与处理电路成一体式。

[0033] 另外参考图6,展示流程图以说明在用于视频搜索的服务器处实施的过程600。过程600包含在块602处将多个图像存储在服务器的存储媒体中。存储媒体可以包含用于存储图像的数据库。过程600包含在块604处从用于视觉搜索的计算装置接收具有所关注的至少一个潜在对象的至少一个经校正图像。在块606处,所述过程提取表示至少一个经校正图像的特征的描述符。至少一个经校正图像的经提取描述符设计成旋转、尺度及光照不变,而不需要视角或仿射失真不变。在块608处,所述过程将至少一个经校正图像的经提取描述符与存储在数据库中的图像的描述符匹配。在块610处,所述过程传输与存储在数据库中的具有与经提取描述符匹配或基本上匹配的描述符的至少一个图像相关联的信息。

[0034] 表示图像上的特征的描述符通常设计成旋转、尺度、光照及一定程度的视角失真不变。如按照定义,此不变性以损失辨别力为代价,这些变量的变化应对计算出的描述符具有极小影响。使用经校正图像潜在地允许使用不太恒定但更可辨别的描述符,此外所述描述符提取及匹配起来更快且更便宜。描述符将仍需旋转、尺度及光照不变,但最难处理的不变性(即,视角失真)将不再要求。更可辨别的描述符将通过使用例如尺度不变特征变换(SIFT)允许消歧否则可能类似的描述符而允许大量对象进行检测。在一个实施例中,至少一个经校正图像的经提取描述符在视角失真或仿射失真上是变化的。

[0035] 应了解,当计算装置或服务器为移动或无线装置时,其可以通过基于或以其它方式支持任何合适的无线通信技术的无线网络经由一或多个无线通信链路通信。举例来说,在一些方面中,计算装置或服务器可以与包含无线网络的网络相关联。在一些方面中,网络可以包括人体局域网络或个人局域网络(例如,超宽带网络)。在一些方面中,网络可以包括局域网或广域网。无线装置可以支持或以其它方式使用多种无线通信技术、协议或标准(例如,CDMA、TDMA、OFDM、OFDMA、WiMAX及Wi-Fi等)中的一或多者。类似地,无线装置可以支持或以其它方式使用多种对应调制或多路复用方案中的一或多者。无线装置因此可以包含适当组件(例如,空中接口)以使用以上或其它无线通信技术建立一或多个无线通信链路及经由一或多个无线通信链路通信。举例来说,装置可以包括具有相关联的发射器及接收器组件(例如,发射器及接收器)的无线收发器,所述无线收发器可以包含促进在无线媒体上的通信的多个组件(例如,信号产生器及信号处理器)。众所周知,移动无线装置因此可以无线方式与其它移动装置、蜂窝电话、其它有线及无线计算机、因特网网站等通信。

[0036] 本文中所描述的技术可以用于多种无线通信系统,例如码分多址(CDMA)、时分多址(TDMA)、频分多址(FDMA)、正交频分多址(OFDMA)、单载波FDMA(SC-FDMA)及其它系统。术语“系统”与“网络”通常可互换地使用。CDMA系统可以实施例如通用陆地无线接入(UTRA)、

CDMA2000等无线电技术。UTRA包含宽带CDMA(W-CDMA)和CDMA的其它变体。CDMA2000涵盖过渡标准(IS)-2000、IS-95及IS-856标准。TDMA系统可以实施例如全球移动通信系统(GSM)的无线电技术。OFDMA系统可以实施例如以下各者等无线电技术:演进型通用陆地无线接入(演进型UTRA或E-UTRA)、超移动宽带(UMB)、电气电子工程师学会(IEEE)802.11(Wi-Fi)、IEEE 802.16(WiMAX)、IEEE 802.20、闪存-OFDM.RTM.等。通用陆地无线接入(UTRA)及E-UTRA是通用移动通信系统(UMTS)的一部分。3GPP长期演进(LTE)是UMTS的使用E-UTRA的即将到来的版本,其在下行链路上采用OFDMA且在上行链路上采用SC-FDMA。UTRA、E-UTRA、UMTS、LTE及GSM描述于来自名为“第三代合作伙伴计划”(3GPP)的组织的文献中。CDMA2000及UMB描述于来自名为“第三代合作伙伴计划2”(3GPP2)的组织的文献中。

[0037] 本文中的教示可以并入到多种设备(例如,装置)中(例如,在其内实施或由其执行)。举例来说,本文中教示的一或多个方面可以并入到电话(例如,蜂窝电话)、个人数据助理(“PDA”)、平板计算机、移动计算机、膝上型计算机、平板计算机、娱乐装置(例如,音乐或视频装置)、耳机(例如,头戴式耳机、听筒等)、医疗装置(例如,生物计量传感器、心率监测仪、计步器、EKG装置等)、用户I/O装置、计算机、服务器、销售点装置、娱乐装置、机顶盒或任何其它合适的装置中。这些装置可以具有不同的功率及数据要求。

[0038] 在一些方面中,无线装置可以包括用于通信系统的接入装置(例如,Wi-Fi接入点)。此接入装置可以提供(例如)经由有线或无线通信链路到另一网络(例如,广域网,例如因特网或蜂窝式网络)的连接性。因此,接入装置可以使得另一装置(例如,Wi-Fi站)能够接入另一网络或一些其它功能性。另外,应了解,所述装置中的一或两者可以为便携式的,或在一些情况下,相对非便携式的。

[0039] 所属领域的技术人员将理解,可以使用多种不同技术及技艺中的任一者来表示信息及信号。举例来说,可以通过电压、电流、电磁波、磁场或磁粒子、光场或光粒子或其任何组合来表示在以上描述中可以始终参考的数据、指令、命令、信息、信号、位、符号及码片。

[0040] 所属领域的技术人员将进一步了解,可以将结合本文中所揭示的实施例而描述的各种说明性逻辑块、模块、电路及算法步骤实施为电子硬件、计算机软件或两者的组合。为清楚说明硬件与软件的此互换性,上文已大致关于其功能性而描述了各种说明性组件、块、模块、电路及步骤。此功能性是实施为硬件还是软件取决于特定应用及强加于整个系统的设计约束。所属领域的技术人员可以针对每一特定应用以不同方式来实施所描述的功能性,但此类实施方案决策不应被解释为会导致脱离本发明的范围。

[0041] 可以使用经设计以执行本文所描述的功能的通用处理器、数字信号处理器(DSP)、专用集成电路(ASIC)、现场可编程门阵列(FPGA)或其它可编程逻辑装置、离散门或晶体管逻辑、离散硬件组件或其任何组合来实施或执行结合本文中所揭示的实施例而描述的各种说明性逻辑块、模块及电路。通用处理器可以为微处理器,但在替代方案中,处理器可以为任何常规的理器、控制器、微控制器或状态机。处理器还可实施为计算装置的组合,例如,DSP与微处理器的组合、多个微处理器、结合DSP芯体的一或多个微处理器,或任何其它此类配置。

[0042] 结合本文中所揭示的实施例而描述的方法或算法的步骤可以直接体现于硬件、由处理器执行的软件模块或其两者的组合中。软件模块可以驻留在RAM存储器、快闪存储器、ROM存储器、EPROM存储器、EEPROM存储器、寄存器、硬盘、可移除式磁盘、CD-ROM,或此项技术

中已知的任何其它形式的存储媒体中。示例性存储媒体耦合到处理器,使得处理器可以从存储媒体读取信息及将信息写入到存储媒体。在替代方案中,存储媒体可以与处理器成一体式。处理器及存储媒体可以驻留在ASIC中。ASIC可以驻留在用户终端中。在替代方案中,处理器及存储媒体可以作为离散组件驻留在用户终端中。

[0043] 在一或多个示例性实施例中,所描述的功能可以在硬件、软件、固件或其任何组合中实施。如果在软件中实施为计算机程序产品,那么可以将功能作为一或多个指令或代码存储于计算机可读媒体上或经由计算机可读媒体传输。计算机可读媒体包含计算机存储媒体及包含促进计算机程序从一处传递到另一处的任何媒体的通信媒体两者。存储媒体可以是可由计算机访问的任何可用媒体。借助于实例而非限制,此类计算机可读媒体可以包括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,或可以用于运载或存储呈指令或数据结构形式的所需程序代码且可以由计算机接入的任何其它媒体。而且,任何连接被适当地称为计算机可读媒体。举例来说,如果使用同轴缆线、光纤缆线、双绞线、数字订户线(DSL)或例如红外线、无线电及微波的无线技术从网站、服务器或其它远程源传输软件,那么同轴缆线、光纤缆线、双绞线、DSL或例如红外线、无线电及微波的无线技术包含于媒体的定义中。如本文中所使用,磁盘及光盘包含压缩光盘(CD)、激光光盘、光学光盘、数字多功能光盘(DVD)、软性磁盘及蓝光光盘,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。

[0044] 提供对所揭示的实施例的先前描述以使所属领域的技术人员能够制作或使用本发明。所属领域的技术人员将容易了解对这些实施例的各种修改,且可以在不脱离本发明的精神或范围的情况下将本文定义的一般原理应用到其它实施例。因此,本发明并不希望限于本文所展示的实施例,而应符合与本文所揭示的原理及新颖特征相一致的最广泛范围。

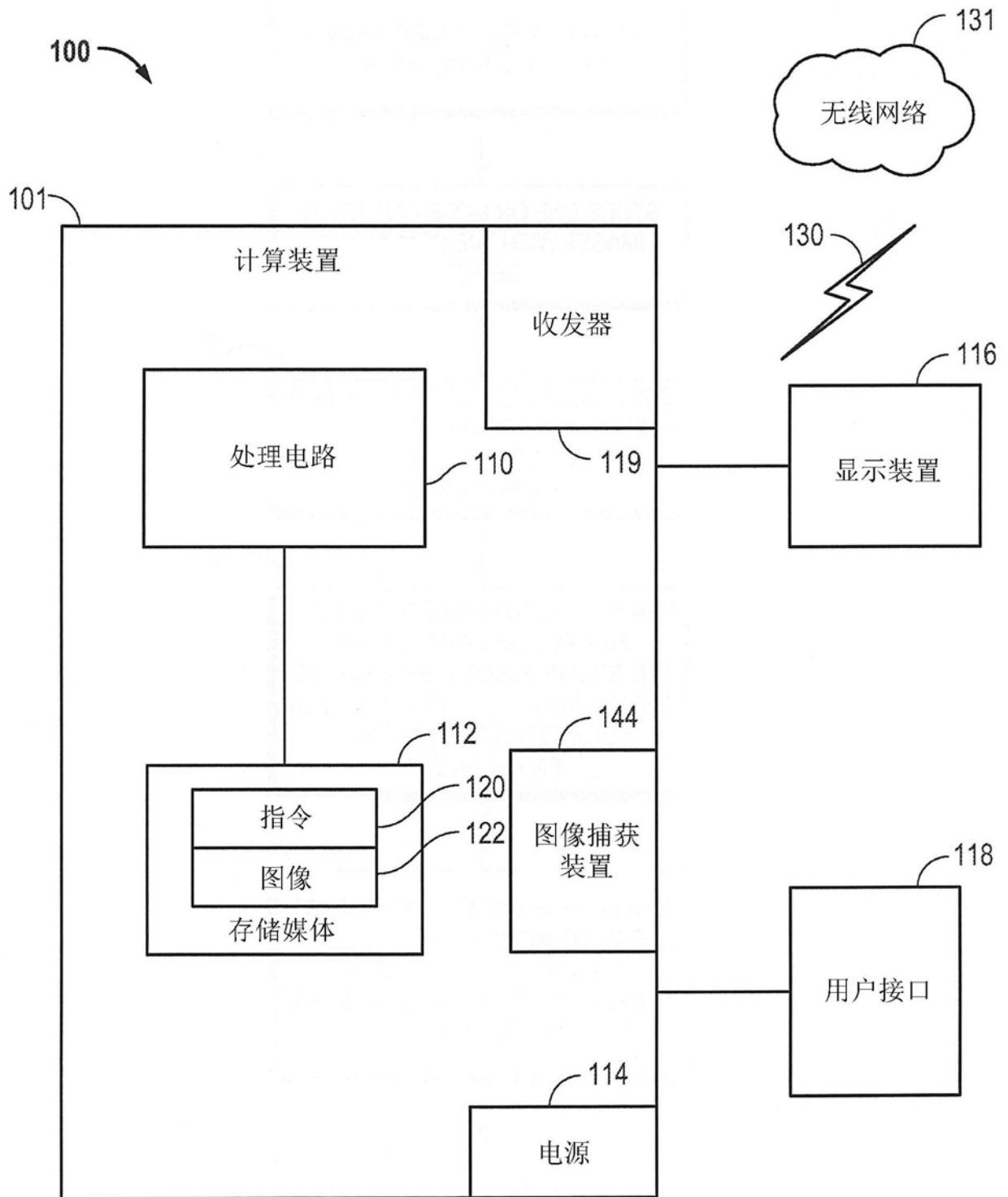


图1

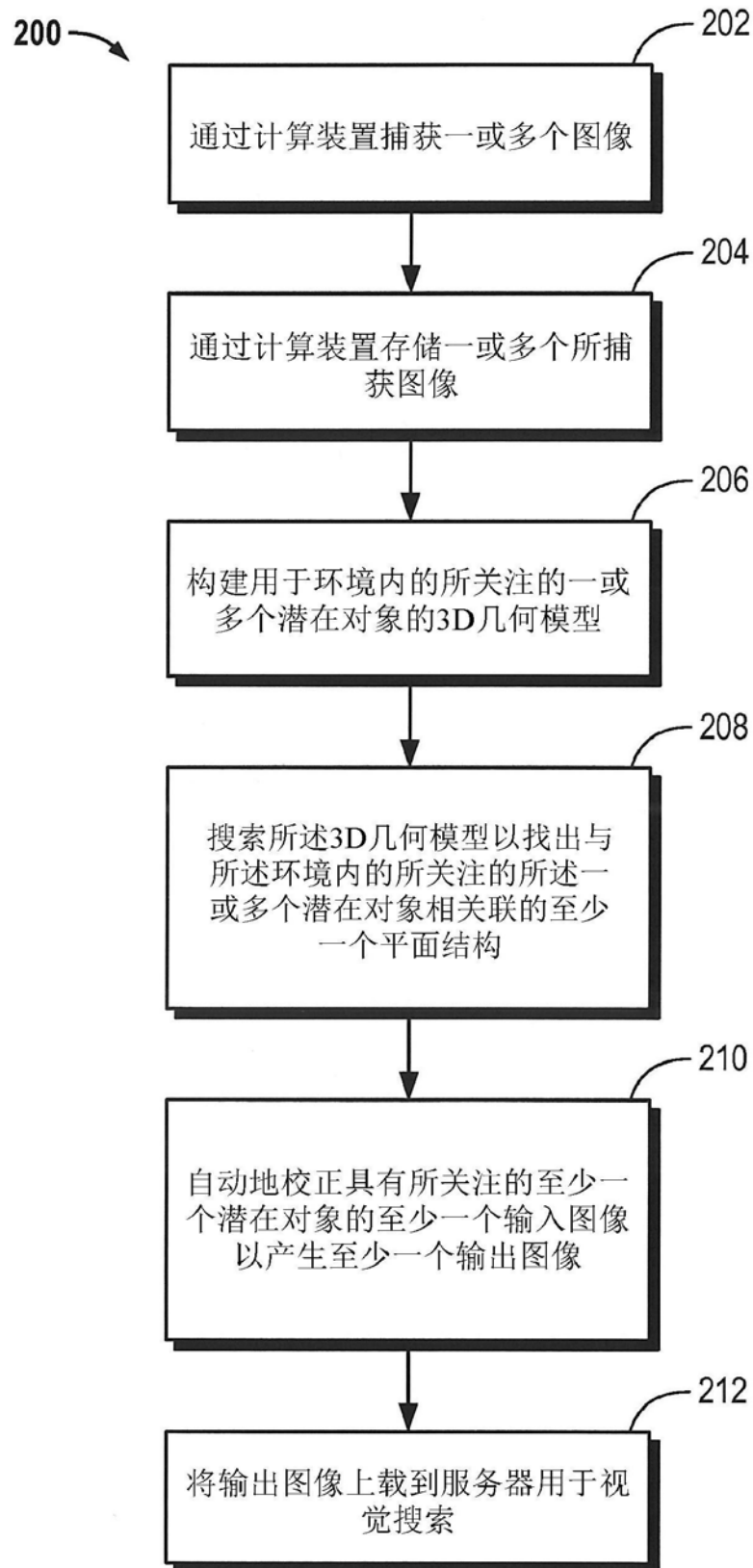


图2

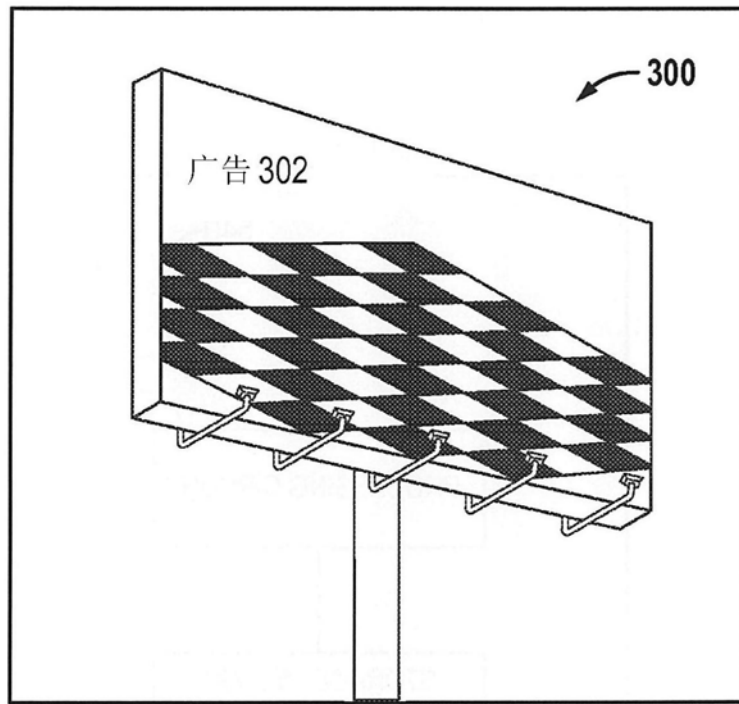


图3

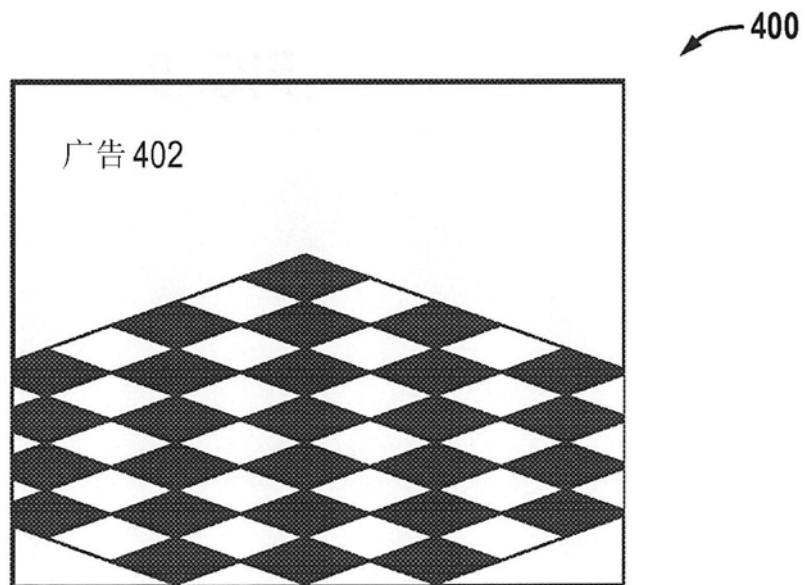


图4

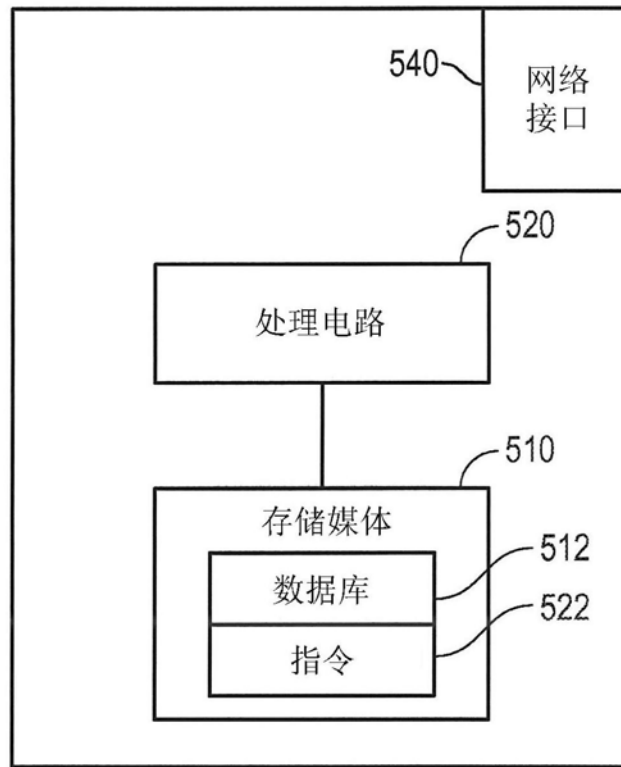


图5

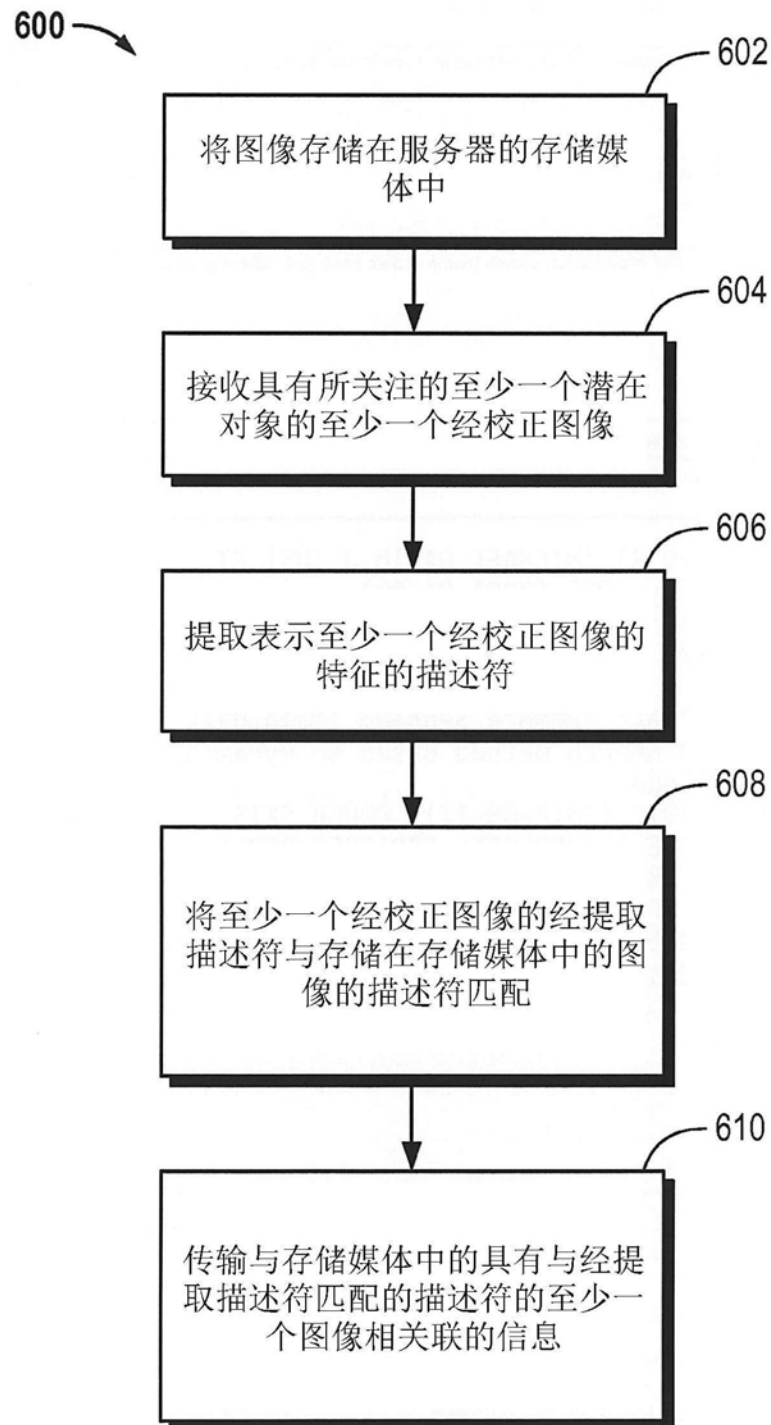


图6