



(12) 发明专利

(10) 授权公告号 CN 109155056 B

(45) 授权公告日 2023.07.21

(21) 申请号 201780028837.9

M·卡切维奇

(22) 申请日 2017.04.25

(74) 专利代理机构 永新专利商标代理有限公司

(65) 同一申请的已公布的文献号  
申请公布号 CN 109155056 A

72002

(43) 申请公布日 2019.01.04

专利代理人 赵腾飞

(30) 优先权数据

(51) Int.CI.

G06T 3/00 (2006.01)

62/340,914 2016.05.24 US  
15/495,709 2017.04.24 US

(56) 对比文件

US 6002430 A, 1999.12.14  
US 6002430 A, 1999.12.14  
US 7016544 B2, 2006.03.21  
US 2005062869 A1, 2005.03.24  
US 2006023105 A1, 2006.02.02  
US 2002063802 A1, 2002.05.30  
US 2010309286 A1, 2010.12.09  
CN 1209933 A, 1999.03.03  
CN 101627635 A, 2010.01.13

(85) PCT国际申请进入国家阶段日

审查员 原野

2018.11.09

(86) PCT国际申请的申请数据

权利要求书3页 说明书22页 附图12页

PCT/US2017/029441 2017.04.25

(87) PCT国际申请的公布数据

W02017/204971 EN 2017.11.30

(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 G·范德奥维拉 M·科班

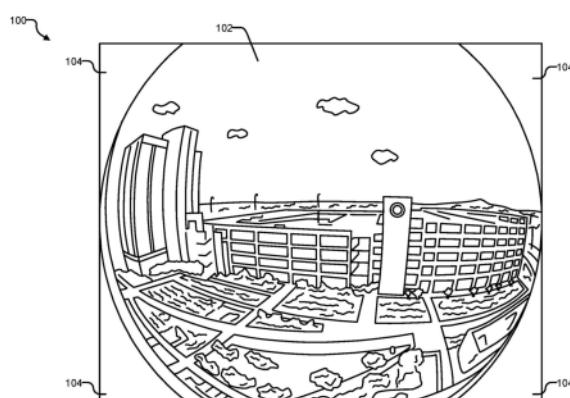
(54) 发明名称

具有用于360度视频的透镜失真校正的鱼眼

呈现

(57) 摘要

在各种实施方案中，提供用于校正鱼眼图像中存在的失真，且呈现所述图像以用于显示为360度视频的系统及方法。在各种实施方案中，计算装置可接收由全向相机捕捉的2维视频数据。所述计算装置可将来自每一视频帧的图像映射至3维半球形表示。在各种实施方案中，可使用多项式模型执行此映射。所述3维半球形表示随后可用于360度视频展示中以提供虚拟现实体验。



1.一种处理360度视频数据的方法,所述方法包括:

获得由全向相机捕捉的2维视频数据,其中所述2维视频数据包含场景的图像,且其中,在所述图像中,所述场景已变形成所述图像的圆形区域;

通过使用经调整的径向值确定2维图像中的像素的坐标来将所述图像映射至3维半球形表示,其中映射所述图像包含将所述图像的所述圆形区域中的像素映射至所述3维半球形表示上的对应位置,其中映射所述图像校正所述场景的所述变形,且其中所述经调整的径向值对应于所述变形的非线性度并且将经确定的坐标用于所述映射中;

其中,所述视频数据是从经编码的位流获得,其中所述经编码的位流包含描述所述变形的一或多个参数,且其中映射所述图像包含使用所述一或多个参数;及

将所述3维半球形表示用于360度视频展示中。

2.根据权利要求1所述的方法,其中映射所述图像包含使用多项式方程式以将所述图像中的一点投影至在所述3维半球形表示上的对应点,且其中所述图像中的所述点提供将映射至所述3维半球形表示上的所述点的像素。

3.根据权利要求1所述的方法,其中所述图像包含大于180度的视场,且所述方法进一步包括:

使用已根据所述视场按比例调整的径向值调整所述2维图像中的像素的坐标;及将所述经调整的坐标用于所述映射中。

4.根据权利要求1所述的方法,其中所述图像包含至少180度的视场。

5.根据权利要求1所述的方法,其中所述全向相机包含鱼眼透镜。

6.根据权利要求1所述的方法,其中所述一或多个参数经编码于所述经编码的位流中的补充增强信息SEI消息中。

7.根据权利要求1所述的方法,其中所述经编码的位流包含于根据ISO基本媒体文件格式而格式化的文件中,且其中所述一或多个参数以所述文件中的结构编码。

8.根据权利要求1所述的方法,其中所述一或多个参数包含多项式阶、多项式系数、多项式缩放参数或映射缩放参数中的至少一或多者。

9.一种用于处理经编码的360度视频数据的设备,其包括:

存储器,其经配置以存储由全向相机捕捉的2维视频数据;及处理器,其经配置以:

获得所述2维视频数据,其中所述2维视频数据包含场景的图像,且其中,在所述图像中,所述场景已变形成所述图像的圆形区域;

通过使用经调整的径向值确定2维图像中的像素的坐标来将所述图像映射至3维半球形表示,其中映射所述图像包含将所述图像的所述圆形区域中的像素映射至所述3维半球形表示上的对应位置,其中映射所述图像校正所述场景的所述变形,且其中所述经调整的径向值对应于所述变形的非线性度并且将经确定的坐标用于所述映射中;

其中,所述视频数据是从经编码的位流获得,其中所述经编码的位流包含描述所述变形的一或多个参数,且其中映射所述图像包含使用所述一或多个参数;及

将所述3维半球形表示用于360度视频展示中。

10.根据权利要求9所述的设备,其中映射所述图像包含使用一多项式方程式以将所述图像中的一点投影至所述3维半球形表示上的一对应点,且其中所述图像中的所述点提供将映射至所述3维半球形表示上的所述点的像素。

11. 根据权利要求9所述的设备,其中所述图像包含大于180度的视场,且其中所述处理器经进一步配置以:使用已根据所述视场按比例调整的径向值调整所述2维图像中的像素的坐标;及将所述经调整的坐标用于所述映射中。

12. 根据权利要求9所述的设备,其中所述图像包含至少180度的视场。

13. 根据权利要求9所述的设备,其中所述全向相机包含鱼眼透镜。

14. 根据权利要求9所述的设备,其中所述一或多个参数经编码于所述经编码的位流中的补充增强信息SEI消息中。

15. 根据权利要求9所述的设备,其中所述经编码的位流包含于根据ISO基本媒体文件格式而格式化的文件中,且其中所述一或多个参数以所述文件中的结构编码。

16. 根据权利要求9所述的设备,其中所述一或多个参数包含多项式阶、多项式系数、多项式缩放参数或映射缩放参数中的至少一或多者。

17. 根据权利要求9所述的设备,其进一步包括:移动装置,其中所述移动装置包含所述全向相机。

18. 一种用于解码视频的设备,其包括:

存储器,其经配置以存储由全向相机捕捉的2维视频数据;及处理器,其经配置以:

获得所述2维视频数据,其中所述2维视频数据包含场景的图像,且其中,在所述图像中,所述场景已变形成所述图像的圆形区域;

通过使用经调整的径向值确定2维图像中的像素的坐标来将所述图像映射至3维半球形表示,其中映射所述图像包含将所述图像的所述圆形区域中的像素映射至所述3维半球形表示上的对应位置,其中映射所述图像校正所述场景的所述变形,且其中所述经调整的径向值对应于所述变形的非线性度并且将经确定的坐标用于所述映射中;

其中,所述视频数据是从经编码的位流获得,其中所述经编码的位流包含描述所述变形的一或多个参数,且其中映射所述图像包含使用所述一或多个参数;及

将所述3维半球形表示用于360度视频展示中。

19. 根据权利要求18所述的设备,其中映射所述图像包含使用多项式方程式以将所述图像中的一点投影至所述3维半球形表示上的对应点,且其中所述图像中的所述点提供将映射至所述3维半球形表示上的所述点的像素。

20. 根据权利要求18所述的设备,其中所述图像包含大于180度的视场,且其中所述处理器经进一步配置以:

调整使用已根据所述视场按比例调整的径向调整的所述2维图像中的像素的坐标;及将所述经调整的坐标用于所述映射中。

21. 根据权利要求18所述的设备,其中所述图像包含至少180度的视场。

22. 根据权利要求18所述的设备,其中所述全向相机包含鱼眼透镜。

23. 根据权利要求18所述的设备,其中所述一或多个参数经编码于所述经编码的位流中的补充增强信息SEI消息中。

24. 根据权利要求18所述的设备,其中所述经编码的位流包含于根据ISO基本媒体文件格式而格式化的文件中,且其中所述一或多个参数以所述文件中的结构编码。

25. 根据权利要求18所述的设备,其中所述一或多个参数包含多项式阶、多项式系数、多项式缩放参数或映射缩放参数中的至少一或多者。

26. 根据权利要求18所述的设备,其进一步包括:

移动装置,其中所述移动装置包含所述全向相机。

27. 根据权利要求18所述的设备,其进一步包括:

移动装置,其中所述移动装置包含用于显示所述360度视频展示的显示器。

28. 一种上面存储有指令的非暂时性计算机可读媒体,所述指令在由一或多个处理器执行时使得所述一或多个处理器以:

获得由全向相机捕捉的2维视频数据,其中所述2维视频数据包含场景的图像,且其中,在所述图像中,所述场景已变形成所述图像的圆形区域;

通过使用经调整的径向值确定2维图像中的像素的坐标来将所述图像映射至3维半球形表示,其中映射所述图像包含将所述图像的所述圆形区域中的像素映射至所述3维半球形表示上的对应位置,其中映射所述图像校正所述场景的所述变形,且其中所述经调整的径向值对应于所述变形的非线性度并且将经确定的坐标用于所述映射中;

其中,所述视频数据是从经编码的位流获得,其中所述经编码的位流包含描述所述变形的一或多个参数,且其中映射所述图像包含使用所述一或多个参数;及

将所述3维半球形表示用于360度视频展示中。

29. 一种设备,其包括:

用于获得2维视频数据的装置,其中所述2维视频数据包含场景的图像,且其中,在所述图像中,所述场景已变形成所述图像的圆形区域;

用于通过使用经调整的径向值确定2维图像中的像素的坐标来将所述图像映射至3维半球形表示的装置,其中映射所述图像包含将所述图像的所述圆形区域中的像素映射至所述3维半球形表示上的对应位置,其中映射所述图像校正所述场景的所述变形,且其中所述经调整的径向值对应于所述变形的非线性度并且将经确定的坐标用于所述映射中;

其中,所述视频数据是从经编码的位流获得,其中所述经编码的位流包含描述所述变形的一或多个参数,且其中映射所述图像包含使用所述一或多个参数;及

用于将所述3维半球形表示用于360度视频展示中的装置。

## 具有用于360度视频的透镜失真校正的鱼眼呈现

### 发明内容

[0001] 360度视频可为观察者提供一种沉浸式体验。举例来说，360度视频可为观察者提供一种虚拟现实体验，将所述观察者置于虚拟不同时间及/或场所中。作为另一实例，360度视频可为观察者提供由远程装置（例如，无人操纵的空中载具或其它装置）捕捉的视频内容的第一人称视图。用于捕捉360度视频的一种方法为使用全向相机。全向相机可通过仅几个透镜捕捉较宽视场。所得图像展现鱼眼效应。

[0002] 在各种实施方案中，提供用于呈现360度视频展示的鱼眼图像，同时校正鱼眼图像中的固有失真的系统、方法及计算机可读媒体。360度视频可通过使用折反射式系统及/或鱼眼透镜的全向相机捕捉。此类相机可捕捉少至一个或两个图像中的场景。虽然这些图像可捕捉大量信息，诸如180度或甚至360度视场，但所述场景经压缩成所述图像且因此对于人眼呈现失真。此外，所述图像为2维的，且需要映射至3维表示以便通过360度视频系统显示。

[0003] 根据至少一个实例，提供用于处理360度视频数据的方法。所述方法包含获得由全向相机捕捉的2维视频数据。所述2维视频数据可包含场景的图像。在所述图像中，场景已变形成所述图像的圆形区域。所述方法进一步包含将图像映射至3维半球形表示。映射图像可包含将图像的圆形区域中的像素映射至3维半球形表示上的对应位置。映射图像可校正场景的变形。所述方法进一步包含将3维半球形表示用于360度视频展示中。

[0004] 在另一实例中，提供一种设备，所述设备包含存储器及经配置以处理360度视频数据的处理器。所述处理器经配置以获得由全向相机捕捉的2维视频数据且可获得由全向相机捕捉的2维视频数据。所述2维视频数据可包含场景的图像。在所述图像中，场景已变形成所述图像的圆形区域。处理器经进一步配置以将所述图像映射至3维半球形表示且可将所述图像映射至3维半球形表示。映射图像可包含将图像的圆形区域中的像素映射至3维半球形表示上的对应位置。映射图像可校正场景的变形。处理器经进一步配置以将所述3维半球形表示用于360度视频展示中且可将所述3维半球形表示用于360度视频展示中。

[0005] 在另一实例中，提供在其上存储有指令的计算机可读取媒体，在所述指令由处理器执行时执行一种方法，所述方法包含获得由全向相机捕捉的2维视频数据。所述2维视频数据可包含场景的图像。在所述图像中，场景已变形成所述图像的圆形区域。所述方法进一步包含将图像映射至3维半球形表示。映射图像可包含将图像的圆形区域中的像素映射至3维半球形表示上的对应位置。映射图像可校正场景的变形。所述方法进一步包含将3维半球形表示用于360度视频展示中。

[0006] 在另一实例中，提供一种设备，所述设备包含用于获得由全向相机捕捉的2维视频数据的装置。所述2维视频数据可包含场景的图像。在所述图像中，场景已变形成所述图像的圆形区域。所述设备进一步包含用于将图像映射至3维半球形表示的装置。映射图像可包含将图像的圆形区域中的像素映射至3维半球形表示上的对应位置。映射图像可校正场景的变形。所述设备进一步包含用于将3维半球形表示用于360度视频展示中的装置。

[0007] 在一些方面中，映射图像包含使用多项式方程式将图像中的点投影至3维半球形

表示上的对应点。所述图像中的所述点可提供将映射至3维半球形表示上的所述点的像素。

[0008] 在一些方面中,变形为非线性的。在这些方面中,上文所描述的方法、设备及计算机可读媒体进一步包括确定2维图像中的像素的坐标。可使用经调整的径向值确定所述坐标,其中所述经调整的径向值对应于变形的非线性度。这些方面进一步包含将所述标准化坐标用于映射中。

[0009] 在一些方面中,图像包含大于180度的视场。在这些方面中,上文所描述的方法、设备及计算机可读媒体进一步包括调整2维图像中的像素的坐标,其中使用已根据视场按比例调整的径向值来调整所述坐标。这些方面进一步包含将所述经调整的坐标用于映射中。

[0010] 在一些方面中,图像包含至少180度的视场。

[0011] 在一些方面中,全向相机包含鱼眼透镜。

[0012] 在一些方面中,视频数据是从经编码的位流获得,其中所述经编码的位流包含描述所述变形的一或多个参数。在这些方面中,映射图像包含使用所述一或多个参数。在一些方面中,所述一或多个参数经编码于经编码位流中的补充增强信息(SEI)消息中。在一些方面中,所述经编码的位流包含于根据ISO基本媒体文件格式而格式化的文件中,且其中所述一或多个参数以所述文件中的结构来编码。在一些方面中,所述一或多个参数包含多项式阶、多项式系数、多项式缩放参数或映射缩放参数中的至少一或者者。

[0013] 在一些方面中,诸如上文所描述的设备可包含移动装置。在一些实施方案中,移动装置包含全向相机。在一些实施方案中,移动装置包含用于显示360度视频展示的显示器。

[0014] 此发明内容并不意在识别所要求保护主题的关键或基本特征,也不意在单独用以确定所要求保护主题的范围。应参考此专利的整个说明书的适当部分、任何或全部图式及每一权利要求来理解主题。

[0015] 在参考以下说明书、权利要求书及随附图式时,前述内容连同其它特征及实施例将变得更显而易见。

## 附图说明

[0016] 下文参考以下图式详细描述本申请案的说明性实施例:

[0017] 图1示出使用鱼眼透镜捕捉的图像的实例。

[0018] 图2A及图2B示出本文所论述的技术的实例应用。

[0019] 图3示出半球形表示的实例用途。

[0020] 图4示出180度角鱼眼透镜的横截面的实例,及用于所述透镜的对应图像平面。

[0021] 图5示出半球形鱼眼透镜的实例,及用于所述半球形鱼眼透镜的对应图像平面。

[0022] 图6A示出可由全向相机捕捉的图像的实例。

[0023] 图6B示出待将图6A的图像中所捕捉的场景映射至的半球形表示的实例。

[0024] 图7A及图7B示出在存在及不存在校正的情况下所呈现的图像的实例。

[0025] 图8示出用于将2维鱼眼图像映射至3维半球形表示以用于在360度视频展示中显示的过程的实例。

[0026] 图9为示出包含编码装置及解码装置的系统的实例的框图。

[0027] 图10为示出实例编码装置的框图。

[0028] 图11为示出实例解码装置的框图。

## 具体实施方式

[0029] 下文提供本发明的某些方面及实施例。这些方面及实施例中的一些可被独立地应用且其中的一些可以将对所属领域的技术人员显而易见的组合来加以应用。在以下描述中,出于解释的目的,阐述特定细节以便提供对本发明的实施例的透彻理解。然而,各种实施例可在无这些特定细节的情况下得以实践将为显而易见的。图式及描述不意在为限定性的。

[0030] 随后描述仅提供示范性实施例,且并不意在限制本发明的范围、适用性或配置。确切来说,示范性实施例的随后描述将为所属领域的技术人员提供能够实施示范性实施例的描述。应理解,可在不背离如所附权利要求书中所阐述的本发明的精神及范围的情况下对元件的功能及布置做出各种改变。

[0031] 在以下描述中给出特定细节以提供对实施例的透彻理解。然而,所属领域的一般技术人员中的一者应理解,实施例可在无需这些具体细节的情况下实践。举例来说,电路、系统、网络、过程及其它组件可在框图形式中展示为组件以便在不必要细节下不与实施例混淆。在其它情况下,可在无不必要细节的情况下展示熟知电路、过程、算法、结构及技术以便避免混淆实施例。

[0032] 另外,应注意,个别实施例可描述为被描绘为流程图、流图、数据流图、结构图或框图的过程。尽管流程图可将操作描述为依序过程,但许多操作可并行地或同时执行。另外,可重新布置操作的次序。过程在其操作完成时终止,但可具有不包含于图中的额外步骤。过程可对应于方法、函数、程序、子例程、子程序等。当过程对应于函数时,其终止可对应于函数传回至主叫函数或主函数。

[0033] 术语“计算机可读媒体”包含但不限于便携式或非便携式存储装置、光学存储装置及能够存储、含有或携载指令及/或数据的各种其它媒体。计算机可读媒体可包含非暂时性媒体,非暂时性媒体中可存储数据且不包含载波及/或无线地或经由有线连接传播的暂时电子信号。非暂时性媒体的实例可包含(但不限于)磁盘或磁带、诸如紧密光盘(CD)或数字化通用光盘(DVD)的光学存储媒体、快闪存储器、存储器或存储器装置。计算机可读媒体可具有存储于其上的代码及/或机器可执行指令,所述代码及/或机器可执行指令可表示程序、函数、子程序、程序、例程、子例程、模块、软件包、种类或指令、数据结构或程序语句的任何组合。一个码段可通过传递及/或接收信息、数据、论证、参数或存储器内容耦合至另一码段或硬件电路。信息、引数、参数、数据等可经由包含存储器共享、消息传递、令牌传递、网络发射或其类似者的任何合适方式传递、转发或发射。

[0034] 此外,可由硬件、软件、固件、中间软件、微码、硬件描述语言或其任何组合实施实施例。当以软件、固件、中间软件或微码加以实施时,用以执行必要任务的程序代码或码段(例如,计算机程序产品)可存储于计算机可读或机器可读媒体中。处理器可执行必要任务。

[0035] 虚拟现实描述可以似乎真实或物理方式相互作用的三维环境。在一些情况下,体验虚拟现实环境的用户使用诸如头戴式显示器(HMD)及也视情况选用的衣物(例如,配备有传感器的手套)的电子装备以与虚拟环境相互作用。随着用户在真实世界中移动,在虚拟环境中呈现的图像也改变,从而给予用户其正在虚拟环境内移动的感知。在一些情况下,虚拟环境包含与用户的移动相关的声,从而给予用户声音来源于特定方向或来源的视听。虚拟现实视频可以极高质量被捕捉且呈现,从而潜在地提供真实沉浸式虚拟现实体验。虚拟

现实应用包含游戏、训练、教育、体育视频及线上购物以及其它。

[0036] 360度视频为经捕捉以在虚拟现实环境中显示的视频。在一些应用中，来自真实世界的视频可用于虚拟现实环境的展示中，相对于计算机产生的图形，诸如可在游戏及虚拟世界中找到。在这些申请案中，用户可以与用户可体验其存在位置的相同方式体验另一位置。举例来说，用户可在使用定位于旧金山(San Francisco)中的360度视频系统时体验柏林(Berlin)的步行旅行。

[0037] 360度视频系统通常包含视频捕捉装置及视频显示装置，且可能也包含诸如服务器、数据存储器及数据发射装备的其它中间装置。视频捕捉装置可包含相机集合，即多个相机的集合，每一相机定向于不同方向且捕捉不同视图。在一些应用中，六个相机可用以捕捉以相机集合的位置为中心的全360度视图。一些视频捕捉装置可使用较少相机，诸如主要捕捉侧至侧视图或使用具有宽视场的透镜的视频捕捉装置。视频大体包含帧，其中帧为场景的电子译码静态图像。相机每秒捕捉某数目个帧，其通常被称作相机的帧速率。

[0038] 在一些情况下，为获得无缝360度视图，可对由相机集合中的相机中的每一者捕捉的视频执行图像缝合。在360度视频产生的情况下的图像缝合涉及组合或合并来自视频帧重叠或将以其它方式连接所在的区域中的邻接相机的视频帧。结果将为大体球形帧，但类似于麦卡托投影(Mercator projection)，经合并的数据通常以平面方式表示。举例来说，合并视频帧中的像素可经映射至立方体形状或一些其它三维平面形状(例如，角锥形、八面体、十面体等)的平面上。视频捕捉及视频显示装置可以意味着将视频帧处理为像素的栅格的光栅原理操作，在此情况下，正方形平面、矩形平面或其它合适成形的平面可用于表示球状环境。

[0039] 映射至平面表示的360度视频帧可经编码及/或压缩以供存储及/或发射。编码及/或压缩可使用视频编解码器(例如，H.265/HEVC顺应式编解码器、H.264/AVC顺应式编解码器或其它合适的编解码器)完成且产生经压缩视频位流(或经编码视频位流)或位流组。下文进一步详细地描述使用视频编解码器对视频数据进行编码。

[0040] 在一些实施方案中，所述经编码视频位流可存储及/或囊封于媒体格式或文件格式中。所述所存储位流可例如经由网络发射至接收器装置，所述接收器装置可解码且呈现视频以供显示。此接收器装置在本文中可被称作视频显示装置。举例来说，360度视频系统可从经编码视频数据产生经囊封文件(例如，使用国际标准组织(ISO)基本媒体文件格式及/或衍生的文件格式)。举例来说，视频编解码器可编码视频数据，且囊封引擎可通过将视频数据囊封于一或多个ISO格式媒体文件中而产生媒体文件。替代地或另外，所存储的位流可直接从存储媒体提供至接收器装置。

[0041] 接收器装置也可实施编解码器以解码及/或解压缩经编码视频位流。在经编码视频位流存储及/或囊封于媒体格式或文件格式中的情况下，接收器装置可支持用以将视频位流封装入文件中的媒体或文件格式，且可提取视频(及也可能音频)数据以产生经编码视频数据。举例来说，接收器装置用经囊封视频数据剖析媒体文件以产生经编码视频数据，且接收器装置中的编解码器可解码经编码视频数据。

[0042] 接收器装置随后可将经解码视频信号传送至呈现装置(例如，视频显示装置、播放器装置，或其它合适的呈现装置)。呈现装置包含例如头戴式显示器、虚拟现实电视机及其他180度或360度显示装置。一般来说，头戴式显示器能够追踪穿戴者的头部的移动及/或穿

戴者的眼睛的移动。头戴式显示器可使用追踪信息以呈现对应于穿戴者正观察的方向的360度视频的部分,使得穿戴者以其将体验真实世界的相同方式体验虚拟环境。呈现装置可以捕捉视频的相同帧速率或以不同帧速率呈现视频。

[0043] 用于捕捉360度视频的相机集合可包含各种全向相机、折反射式相机(使用透镜及弧形反射镜的相机),及/或配备有鱼眼透镜的相机。全向相机的一个实例为使用在相对方向上聚焦的两个鱼眼透镜的Ricohθ-S。

[0044] 全向相机(诸如折反射式相机及具有鱼眼透镜的相机)通常捕捉具有大量失真的图像。图1示出使用鱼眼透镜捕捉的图像100的实例。鱼眼透镜为可具有高至180度或更大的视场的广角透镜。配备有两个此类透镜的背对背置放的相机可因此捕捉一起提供360度视图(或更大)的两个图像。

[0045] 然而,实现此宽视场的透镜的极端曲率也使得图像变失真。如图1的实例中所示,在图像100中捕捉的场景102在形状上为圆形,且根据所述透镜的曲率变形。由于相机传感器通常为矩形,所以图像100为矩形且无像素经捕捉的角104通常为空白或黑色。在此实例中,已裁剪场景102的顶部及底部。可因各种原因裁剪场景102的顶部及底部。在一个说明性实例中,由于透镜的形状及/或相机传感器的形状,可裁剪场景102的顶部及底部。

[0046] 实例图像100捕捉大量信息,在此情况下,180度视场。其它图像可捕捉超过180度,诸如270度视图。然而,场景102的失真并不为观察者提供极逼真的展示。另外,图像100为3维场景的平坦、2维表示。为展示为360度视频,图像100需要经映射至球状、3维(3-D)表示。

[0047] 在各种实施方案中,提供用于校正鱼眼图像中存在的失真,且呈现用于显示为360度视频的图像的系统及方法。在各种实施方案中,计算装置可接收由全向相机捕捉的2维视频数据。所述计算装置可将来自每一视频帧的图像映射至3维半球形表示。在各种实施方案中,可使用多项式模型执行此映射。3维半球形表示随后可用于360度视频展示中以提供虚拟现实体验。

[0048] 图2A及图2B示出本文所论述的技术的实例应用。图2A的实例示出可由全向相机捕捉的图像200。如上文所讨论,图像200在形状上为矩形,其中所述矩形形状对应于相机的图像捕捉传感器的大小及形状。图像200包含圆形区域202,所述圆形区域包含捕捉场景的像素。其中图像200的无像素经捕捉的角204可保持空白或可包含黑色像素(例如,具有0或255的像素值)。

[0049] 在圆形区域202中捕捉的像素包含某一视场,其中所述视场通过透镜及/或相机决定。举例来说,视相机的透镜、反射镜及/或传感器的结构而定,圆形区域202可包含90度视场、180度视场、270度视场,或一些其它度的视场。为将视场拟合于图像200中,以线性或非线性方式将所述像素变形成圆形区域202,如下文进一步论述。

[0050] 在各种实施方案中,下文所描述的技术将圆形区域202中的像素映射至半球形表示210,其实例展示于图2B中。半球形表示210随后可用于使用虚拟现实呈现装置向观察者展示图像200。半球形表示210可为球体的一半(其表示180度视图)、小于球体的一半(例如,在图像200捕捉小于180度的视图时),或大于球体的一半(例如,在图像200捕捉超过180度的视图时)。

[0051] 在各种实施方案中,下文所论述的技术具有将圆形区域202中的像素拉伸成半球形形状的效应。举例来说,圆形区域202的中心202可对应于半球形表示210的中心222或顶

点。作为另一实例,圆形区域202的最顶部点214a可对应于半球形表示210的最顶部点224a(例如北极),且圆形区域202的最底部点214b可对应于半球形表示210的最底部点224b(例如南极)。类似地,圆形区域202的最右端214c及最左端214d可对应于在此实例中半球形表示210的最右端224c及不可见的最左端224d。在圆形区域202的中心212与边缘之间的像素可进一步均匀分布于半球形表示210的表面上方。

[0052] 所得半球形表示210可用于将平坦图像200展示在360度显示器中。图3示出半球形表示310的实例用途。使用360度显示装置,可呈现半球形表示310以使得给予观察者320其在由半球形表示310捕捉的像素所表示的场景的内部的视听。在一些情况下,半球形表示310的顶点可定向在观察者的视场的中心处。在各种实施方案中,可针对将在观察者320后面的视图提供额外半球形表示。

[0053] 存在各种类型的鱼眼透镜,所述鱼眼透镜中的每一者以不同方式将广角视场映射至图像平面。一个实例为角鱼眼透镜,其也称作 $f\text{-}\theta$ 透镜。在角鱼眼透镜下,距图像的中心的距离与相机视图方向的角度成比例。因此,解析度在整个图像中大致相等。角鱼眼透镜可用于一直至全360度的角。

[0054] 图4示出180度角鱼眼透镜420的横截面的实例,及用于所述透镜420的对应图像平面402。所述图示描绘从与透镜的图像平面402垂直的侧面观察的透镜420。经示出的视图也可为透镜420的俯视图,或距与所述透镜的图像平面402垂直的任何其它角度的视图,然而出于此实例的目的,将假定图4示出侧视图。在图4的左侧示出所述透镜的图像平面402的正视图。

[0055] 角透镜的构造为使得,对于距零的任一角度(其中零为透镜420的中心)(其也视为相机位置406),在彼角度处捕捉的空间中的点线性对应于图像平面402中的点 $\rho$ 。也就是说,在例如 $\alpha=$ 距零的45度,进入透镜420的光将在图像平面402的中心与图像平面402的顶部的中间的点处经捕捉。图像平面402的顶部对应于 $\alpha=90$ 度乘以缩放或放大因数。类似地,在 $\alpha=$ 距零的22.5度处,进入所述透镜的光将在 $\rho=$ 零与在45度捕捉的像素的中间的图像平面402中经捕捉,且在距零的67.5度处,光将在45度与图像平面402的顶部的中间的图像平面402中经捕捉。由于假定在此实例中从侧面观察到透镜420,所以距图像平面402的中心的距离 $\rho$ 将仅在垂直方向上改变。

[0056] 可将由图4的实例透镜420产生的图像平面402映射至半球形,其中所述半球形可展示于使用3维网格的计算装置中。举例来说,可使用图形处理单元(GPU)呈现来自图像平面402的圆形图像。将图像映射至半球形可矫正透镜420产生的自然失真。在下文进一步论述用于此映射的技术。

[0057] 其它类型的鱼眼透镜产生从所述透镜的中心至图像的外边缘的非线性失真。图5示出半球形鱼眼透镜520的实例。透镜520的侧视图示出于图式的右侧,且用于所述透镜的图像平面502的正视图示出于图式的左侧。

[0058] 半球形透镜产生半球体至图像平面502上的平行投影。在此实例及其它实例中,图像变得径向压缩,且朝向透镜的外边缘发生更多压缩。也就是说,在角度 $\alpha$ 从零增加时,其中零为透镜520的中心(其也被视为相机位置506),在图像平面502中捕捉的图像变得以非线性方式逐渐经压缩。由这些类型的透镜捕捉的图像需要校正失真的非线性度。

[0059] 在各种实施方案中,下文所论述的技术可用于校正鱼眼投影可引入至经捕捉图像

的非线性失真。在各种实施方案中,这些技术包含使用采用多项式模型的全向相机校准技术。全向相机校准技术论述于大卫·史卡拉穆札(Davide Scaramuzza)等人,“用于精确全向相机校准的灵活技术及运动结构(A Flexible Technique for Accurate Omnidirectional Camera Calibration and Structure from Motion)”中,其以全文引用的方式并入本文中。在全向相机(包含折反射式及折射相机)已根据此技术经校准后,可针对由相机的传感器捕捉的每一像素确定真实世界中的点。

[0060] 图6A示出可由全向相机捕捉的图像602的实例。如上文所论述,图像602可具有鱼眼失真,其中在所述图像中捕捉的场景已变形成圆形形状。出于将图像602映射至半球形表示610的目的,所述图像的水平轴线已标记为u且垂直轴线已标记为v。对于水平及垂直维度两者,零位于图像602的中心处。在此实例中,ρ经定义为沿通过图像602所形成的圆形的半径的点。此点604可具有坐标(u,v)。

[0061] 图6B示出在图像602中所捕捉的场景将映射至的半球形表示610的实例。出于此映射的目的,已界定正交X轴及Y轴。在各种实施方案中,X轴可对应于(例如)真实世界中的上及下,且Y轴可对应于左及右。Z轴已经界定为与相机的图像传感器垂直。可使用坐标(x,y,z)描述在半球形表示610的表面上的点614。点614也可具有描述在X轴与Y轴之间的点的旋转的对应角度值Φ。点614也可具有角度值r,所述角度值描述在Z轴与由X轴及Y轴所形成的平面之间的点的旋转。

[0062] 在各种实施方案中,上文所提及的全向相机校准方法使用以下多项式以使图像602中的点604的位置(u,v)与相对于传感器轴线(也就是说,相对于图6B中所示的Z轴)的点614相关:

$$f(\rho) = a_0 + a_1 \rho + a_2 \rho^2 + \cdots + a_N \rho^N$$

[0064] 在以上方程式中,系数 $a_i$ ( $i=0,1,2,\dots,N$ ),且多项式幂N为通过校准确定的模型参数,且如上文所论述的ρ为距传感器轴线的距离,其可被称为径向值。

[0065] 相机校准大体上包含确定特定相机的固有参数及非固有参数。相机的固有参数包含相机的焦距及光学中心。在以上方程式中,系数 $a_i$ 为描述图像602的形状的固有参数。非固有参数可包含相对于地面旋转(例如,倾斜、偏离及/或滚动)的度数,及相机的3维真实世界位置(称作相机的“翻译”)。所述非固有参数大体上从相机的传感器的中心量测。相机校准大体上涉及确定相机的固有及非固有参数。在各种实施方案中,此可以自动化方式,及/或在捕捉校准图案的图像的辅助下完成。

[0066] 以上函数 $f(\rho)$ 假定图像602可均一旋转,意味着对于从图像602的中心的给定距离,在所述距离处的图像602中的全部点捕捉与相机传感器等距的空间中的点。此为合理的假定,这是由于相机透镜通常经制造为精确地对称。所述函数因此可用于将圆形鱼眼图像602映射至半球形表示610上。如上文所提及,所述函数使图像602中的坐标与相对于半球形表示610的Z轴的对应点相关。因此,对于半球形表示610上的任一点614,所述函数可用于在图像602中找到对应的点604。随后可从图像602中的点604获得像素,且将所述像素置放于对应点614处的半球形表示610中。

[0067] 当图6A中的图像602包含半球形表示610中的点614(具有旋转坐标(r,Φ)的坐标(x,y,z))的线性失真(诸如图4的实例中所描述)时,可使用以下方程式确定对应点604(在图像平面602中具有坐标(u,v)):

[0068]  $u = \frac{r \cos \Phi + 1.0}{2.0}$

[0069]  $v = \frac{(r \sin \Phi + 1.0)}{2.0}$

[0070] 在以上方程式中, u及v经标准化, 意味着u及v各自在0与1之间变化。

[0071] 可使用函数atan2确定旋转坐标( $r, \Phi$ )。所述atan2函数为具有两个自变数的反正切函数。对于任何实数自变数x及y均不等于零时, atan2(y, x)在弧度上返回平面的正向X轴与通过坐标(x, y)给定的平面上的点之间的角度。对于逆时针角来说所述角度为正值(例如, 平面的上半部分, 其中 $y > 0$ ), 且对于顺时针角来说为负值(例如, 平面的下半部分, 其中 $y < 0$ )。使用atan2, 可按以下计算r及 $\Phi$ :

[0072]  $r = \frac{\text{atan}2(\sqrt{x^2 + y^2}, z)}{\pi/2}$

[0073]  $\Phi = \text{atan}2(y, x)$

[0074] 在以上方程式中, r将介于0至1的范围内, 且 $\Phi$ 将介于0至 $2\pi$ 的范围内。

[0075] 使用以上方程式, 可将图像602的像素映射至半球形表示610。特定言的, 对于具有坐标(x, y, z)的给定点614, 可计算图像602中的具有坐标(u, v)的对应点604。随后可从图像602中的点604获得像素且将所述像素置放于(x, y, z)处的半球形表示中。

[0076] 当图像602包含非线性失真(诸如图5的实例中所描述)时, 以下额外及/或替代性计算可用于将图像602(如图6A中所示出)映射至半球形表示610。这些方程式可校正非线性失真。在以下论述中, 对于包含非线性失真的图像602中的点,  $u'$ 及 $v'$ (可被称为经校正标准化坐标)将用以描述这些点的坐标。

[0077] 如上文所提及, 距图像602的中心的距离 $\rho$ 非线性地变化。 $\rho$ 的值可表示为:

[0078]  $\rho = k_0 r$

[0079] 在以上方程式中, 可如上文所描述计算角度值r。

[0080] 函数 $f(\rho)$ 可进一步经如下修改:

[0081]  $f(\rho) = k_1(a_0 + a_1\rho + a_2\rho^2 + \dots + a_N\rho^N)$

[0082] 在以上两个方程式中,  $k_0$ 及 $k_1$ 为缩放参数, 且 $a_i$ 为通过如上文所论述的相机校准所确定的系数。函数 $f(\rho)$ 提供具有坐标(x, y, z)的点614至传感器轴线Z上的投影。因此函数 $f(\rho)$ 的结果提供z坐标。

[0083] 参数 $k_0$ 将r缩放成 $\rho$ ; 也就是说,  $k_0$ 根据 $\rho$ 的非线性变化调整r。参数 $k_0$ 因此可用以调整用于鱼眼图像中的非线性压缩的多项式方程式, 诸如上文关于图5所论述。参数 $k_0$ 的值可从r及函数 $f(\rho)$ 导出。特定言的, 可确定 $k_0$ 以使得, 当 $r=1$ 时, 则 $f(\rho)=0$ 。如图6B中所示出,  $r=1$ 为图像602将映射至的半球形表示610的极端边缘。在半球形表示610的边缘处, z坐标为零, 因此 $f(\rho)$ (其提供z坐标)也为零。

[0084] 参数 $k_1$ 将 $f(\rho)$ 缩放成z。用以导出函数 $f(\rho)$ 的相机校准方法假定半球形投影, 其中 $f(\rho) < 0$ 。因此,  $k_1$ 可设定为负值(诸如-1)以使得 $f(\rho)$ 产生正值。替代地, 系数 $a_i$ 的正负号可改变。替代地, 在一些情况下, 所述正负号改变可包含于下文所论述的公式中。

[0085] 可使用以下方程式确定图像平面602中的点( $u'$ ,  $v'$ )的坐标:

[0086]  $u' = \frac{r' \cos \Phi + 1.0}{2.0}$

[0087]  $v' = \frac{r' \sin \Phi + 1.0}{2.0}$

[0088] 在以上方程式中,  $u'$  及  $v'$  经标准化, 且在0与1之间变化。

[0089] 对于以上方程式,  $r'$  可按以下计算:

[0090]  $r' = \frac{2r - \text{atan}2(\rho, f(\rho))}{\pi/2}$

[0091] 使用以上方程式, 径向失真-其中图像602随着距图像的中心的距离增加时经受越来越多压缩-可在将图像映射至半球形表示610时经校正。特定言的, 对于半球形表示610上的点614(x, y, z), 可确定图像中的具有坐标( $u'$ ,  $v'$ )的点。随后可将点( $u'$ ,  $v'$ )的像素置于点614(x, y, z)处。

[0092] 在一些情况下, 可需要额外的缩放因数。举例来说, 当图像包含大于或小于180度的视场及/或视场的部分已经裁剪时, 可引入缩放因数 $\alpha$ 及 $\beta$ 以适应视场及/或裁剪(对于经裁剪鱼眼图像的实例, 参见图1)。当视场对称时, 诸如在全部方向上220度, 则 $\alpha$ 及 $\beta$ 可相同(例如, 两者等于 $180/220$ )。当视场不对称时,  $\alpha$ 及/或 $\beta$ 可设定为适应不对称性的值。举例来说, 在图1的实例中, 假定视场为180度, 则 $\beta$ 可设定为1, 且 $\alpha$ 可设定为 $180/n$ , 其中n为从图像的中心至帧的经裁剪上部或下部边缘的距离。

[0093] 通过 $\alpha$ 及 $\beta$ , 可使用以下方程式确定确定图像602中的点(表示为( $u''$ ,  $v''$ ))的方程式:

[0094]  $u'' = \frac{\alpha r' \cos \Phi + 1.0}{2.0}$

[0095]  $v'' = \frac{\beta r' \sin \Phi + 1.0}{2.0}$

[0096] 在以上方程式中, 可如上文所论述确定 $r'$ 。

[0097] 图7A及图7B示出在存在及不存在校正的情况下所呈现的图像的实例。在图7A的实例中, 图像702呈现为由全向相机所捕捉, 而对于由所述相机造成的失真无任何校正。如由此实例所示, 图像702变形以使得场景中的架构及路径朝向消失点弯曲, 而非直线。

[0098] 在图7B的实例中, 通过使用如上文所论述的技术呈现图像710。在此实例图像710中, 架构的边缘及所述架构前方的路径为直线, 如其将在真实世界中呈现。为了清楚且易于理解起见, 实例图像710在此处示出为2维且在形状上为矩形。如上文所示出, 实例图像710呈现成3维、半球形表示且在此处获得所示出的校正。

[0099] 在各种实施方案中, 与如上文所论述的多项式模型相关联的参数中的一些或全部可通过使用全向相机所捕捉的图像数据传送。这些参数包含多项式系数 $a_i$ ( $i=0, 1, 2, \dots, N$ )、多项式阶N、缩放参数 $k_0$ 、 $k_1$ , 及/或缩放参数 $\alpha$ 及 $\beta$ 以及其它者。在与视频数据包含在一起时, 这些参数随后可用于将视频帧中的鱼眼图像呈现于半球形表示中, 以用于360度视频展示。在一些实施方案中, 所述参数可在视频捕捉装置处使用以将由相机捕捉的2维图像映射至3维表示。3维表示随后可经编码用于存储及/或发射。在一些实施方案中, 所述2维图像可经编码用于存储及/或发射, 且接收器装置可解码所述图像且使用所述参数以呈现所述图

像用于360度视频展示。

[0100] 各种方法可用于传送所述模型参数。举例来说,使用位流的语法结构可使所述参数包含于视频位流中。举例来说,可使所述参数包含于一或多个补充增强信息(SEI)消息中。作为另一实例,可将所述参数并入至用以存储360度视频的文件格式中。作为另一实例,可将所述参数并入至特定网络协议扩展中。

[0101] 在各种实施方案中,上文所论述的多项式模型的替代物可用于得到相同的结果。举例来说,也可使用替代参数利用多项式模型(诸如分段线性模型)的近似值。

[0102] 图8为示出用于将2维鱼眼图像映射至3维半球形表示以用于在360度视频展示中显示的过程800的实例的流程图。在802处,过程800包含获得由全向相机所捕捉的2维视频数据,其中所述2维视频数据包含场景的图像,且其中在所述图像中,所述场景已变形成所述图像的圆形区域。在一些实施方案中,变形可为径向线性的;也就是说,所述图像均一地从圆形区域的中心至外边缘变形。在一些情况下,变形可为非线性,以使得图像包含比在中心存在更多的朝向圆形区域的外边缘的压缩。

[0103] 在一些实施方案中,视频数据是从经编码位流获得。在这些实施方案中,经编码位流可包含描述图像的变形的参数。这些参数可包含例如多项式阶、多项式系数、多项式缩放参数,或映射缩放参数以及其他者。在一些实施方案中,所述参数可从经编码位流中的SEI消息获得。在一些实施方案中,所述参数可包含于含有经编码位流的文件中。在这些实施方案中,文件可包含用于存储所述参数的数据结构。

[0104] 在804处,过程800包含将图像映射至3维半球形表示,其中映射所述图像包含将图像的圆形区域中的像素映射至3维半球形表示上的对应位置,其中映射所述图像校正场景的所述变形。在一些实施方案中,映射所述图像包含使用多项式方程式将图像中的点投影至3维半球形表示上的对应点。图像中的点可提供将映射至3维半球形表示上的点的像素。

[0105] 在一些实施方案中,图像变形成圆形区域可为非线性的。在这些实施方案中,将图像映射至3维半球形表示可包含使用经调整的径向值确定2维图像中的像素的坐标。经调整的径向值可对应于变形的非线性度。举例来说,经调整的径向值从图像的中心非线性地增加至外边缘。使用经调整的径向值确定的坐标随后可用以将图像的像素映射至半球形表示。

[0106] 在一些实施方案中,图像包含大于180度的视场。在这些实施方案中,过程800进一步包含使用已根据视场按比例调整的径向值调整2维图像中的像素的坐标。所述坐标随后可用以将来自图像的像素映射至半球形表示。在一些实施方案中,图像包含180度或更小的视场。

[0107] 在806处,过程包含将3维半球形表示用于360度视频展示中。举例来说,可呈现3维半球形表示以用于通过视频显示装置显示。替代地或另外,3维半球形表示可经编码用于存储及/或发射,其中所述经编码数据可随后经解码且显示。

[0108] 在一些实施方案中,过程800可在移动装置(诸如,智能电话、平板计算机、膝上型计算机、个人数字助理或任何其它种类的可无线连接至网络及/或足够小且轻以易于运输的计算装置)中实施。在这些实施方案中,移动装置可包含用于捕捉360度视频的全向相机。在一些实施方案中,移动装置可包含用于显示360度视频的显示器。

[0109] 在一些实例中,过程800可通过计算装置或设备(诸如下文相对于图9所论述的系

统)执行。举例来说,过程800可通过图9中展示的系统900及/或存储器908或输出端910执行。在一些情况下,计算装置或设备可包含处理器、微处理器、微计算机或经配置以实施图8的过程800的步骤的装置的其它组件。在一些实例中,计算装置或设备可包含经配置以捕捉包含视频帧的视频数据(例如,视频序列)的相机。举例来说,计算装置可包含相机装置(例如,全向相机或其它类型的相机装置),所述相机装置可包含视频编解码器。在一些实例中,捕捉视频数据的相机或其它捕捉装置与计算装置分离,在此情况下,计算装置接收所捕捉视频数据。计算装置可进一步包含经配置以传达视频数据的网络接口。网络接口可经配置以传达基于因特网协议(IP)的数据或其它合适类型的数据。

[0110] 过程800经示出为逻辑流程图,其操作表示可在硬件、计算机指令或其组合中实施的操作的序列。在计算机指令的上下文中,操作表示存储于一或多个计算机可读存储媒体上的计算机可执行指令,在由一或多个处理器执行时,所述计算机可执行指令执行所叙述操作。一般来说,计算机可执行指令包含执行特定功能或实施特定数据类型的例程、程序、对象、组件、数据结构及其类似者。描述操作的次序并不意在被理解为限制,且任何数目个经描述操作可按任何次序及/或与同时实施过程组合。

[0111] 另外,过程800可在经配置有可执行指令的一或多个计算机系统的控制下执行,且可被实施为共同在一或多个处理器上通过硬件或其组合执行的程序代码(例如,可执行指令、一或多个计算机程序或一或多个应用程序)。如上文所提及,程序代码可存储于计算机可读或机器可读存储媒体上,例如,呈包括可由一或多个处理器执行的多个指令的计算机程序的形式。计算机可读或机器可读存储媒体可为非暂时性的。

[0112] 图9为示出包含编码装置904及解码装置912的系统900的实例的框图。编码装置904可为源装置的部分,且解码装置912可为接收装置的部分。源装置及/或接收装置可包含电子装置,诸如,移动或静止电话手持机(例如,智能电话、蜂窝电话或类似者)、桌上型计算机、膝上型计算机或笔记本计算机、平板计算机、机顶盒、电视机、相机、显示装置、数字媒体播放器、视频游戏控制台、视频流式发射装置或任何其它合适的电子装置。在一些实例中,源装置及接收装置可包含用于无线通信的一或多个无线收发器。本文所描述的译码技术适用于各种多媒体应用中的视频译码,包含流式视频发射(例如,经由因特网)、电视广播或发射、编码数字视频以供存储于数据存储媒体上、解码存储于数据存储媒体上的数字视频或其它应用。在一些实例中,系统900可支持单向或双向视频发射以支持诸如视频会议、视频流式发射、视频播放、视频广播、游戏及/或视频电话的应用。

[0113] 通过使用视频译码标准或协议以产生经编码视频位流,编码装置904(或编码器)可用以编码视频数据,包含虚拟现实视频数据。视频译码标准包含ITU-T H.261、ISO/IEC MPEG-1Visual、ITU-T H.262或ISO/IEC MPEG-2Visual、ITU-T H.263、ISO/IEC MPEG-4Visual及ITU-T H.264(也称为ISO/IEC MPEG-4AVC),包含其可调式视频译码及多视图视频译码扩展(分别称为SVC及MVC)。已由ITU-T视频译码专家组(VCEG)及ISO/IEC动画专家组(MPEG)的视频译码联合协作小组(JCT-VC)完成较近期的视频译码标准(高效率视频译码(HEVC))。对HEVC的各种扩展处理多层视频译码,且也正由JCT-VC开发,包含对HEVC的多视图扩展(称作MV-HEVC)及对HEVC的可调式扩展(称作SHVC)或任何其它合适的译码协议。

[0114] 本文中所描述的实施方案使用HEVC标准或其扩展来描述实例。然而,本文中所描述的技术及系统也可适用于其它译码标准,诸如AVC、MPEG、其扩展或已可用或尚不可用或

尚未开发的其它合适的译码标准。因此,虽然可参考特定视频译码标准描述本文中所描述的技术及系统,但所属领域的一般技术人员将了解,描述不应解译为仅适用于彼特定标准。

[0115] 视频源902可提供视频数据至编码装置904。视频源902可为源装置的部分,或可为除源装置以外的装置的部分。视频源902可包含视频捕捉装置(例如,视频相机、相机电话、视频电话或其类似者)、含有经存储视频的视频存档、提供视频数据的视频服务器或内容提供者、从视频服务器或内容提供者接收视频的视频馈入接口、用于产生计算机图形视频数据的计算机图形系统、这些源的组合或任何其它合适的视频源。视频源902的一个实例可包含因特网协议相机(IP相机)。IP相机为可用于监控、家庭安全或其它合适应用的数字视频相机的类型。不同于类比闭路电视(CCTV)相机,IP相机可经由计算机网络及因特网发送及接收数据。

[0116] 来自视频源902的视频数据可包含一或多个输入图像或帧。图像或帧为作为视频的部分的静态图像。编码装置904的编码器引擎906(或编码器)编码视频数据以产生经编码视频位流。在一些实例中,经编码视频位流(或“视频位流”或“位流”)为一系列的一或多个经译码视频序列。经译码视频序列(CVS)包含一系列存取单元(AU),其始于具有在基础层中且具有某些性质的随机存取点图像的AU,直至且不包含具有在基础层中且具有某些性质的随机存取点图像的下一AU。举例来说,开始CVS的随机存取点图像的某些属性可包含等于1的RASL旗标(例如,NoRas1OutputFlag)。否则,随机存取点图像(具有等于0的RASL旗标)并不开始CVS。存取单元(AU)包含一或多个经译码图像以及对应于共享相同输出时间的经译码图像的控制信息。图像的经译码切片以位流级别囊封至称作网络抽象层(NAL)单元的数据单元中。举例来说,HEVC视频位流可包含一或多个CVS,所述一或多个CVS包含NAL单元。两种类别的NAL单元以HEVC标准存在,包含视频译码层(VCL)NAL单元及非VCL NAL单元。VCL NAL单元包含经译码图像数据的一个切片或切片片段(下文描述),且非VCL NAL单元包含关于一或多个经译码图像的控制信息。

[0117] NAL单元可含有形成视频数据的经译码表示(诸如,视频中的图像的经译码表示)的位序列(例如,经编码视频位流、位流的CVS或其类似者)。编码器引擎906通过将每一图像分割成多个切片而产生图像的经译码表示。切片随后分割成明度样本及色度样本的译码树型块(CTB)。明度样本的CTB及色度样本的一或多个CTB连同样本的语法被称为译码树型单元(CTU)。CTU为用于HEVC编码的基本处理单元。CTU可分割成具有不同大小的多个译码单元(CU)。CU含有被称作译码块(CB)的明度及色度样本阵列。

[0118] 明度及色度CB可进一步分割成预测块(PB)。PB为使用用于帧间预测的相同运动参数的明度或色度分量的样本的块。所述明度PB及一或多个色度PB连同相关联语法形成预测单元(PU)。在位流中针对每一PU发信号通知运动参数集合,且所述运动参数集合用于明度PB及一或多个色度PB的帧间预测。CB也可被分割成一或多个变换块(TB)。TB表示色彩分量的样本的正方形块,对所述正方形块应用同一二维变换以用于译码预测残余信号。变换单元(TU)表示明度及色度样本的TB以及对应语法元素。

[0119] CU的大小对应于译码节点的大小,且可为正方形形状。举例来说,CU的大小可为 $8 \times 8$ 样本、 $16 \times 16$ 样本、 $32 \times 32$ 样本、 $64 \times 64$ 样本或达至对应CTU的大小的任何其它适当大小。片语“N×N”在本文中用于指代就垂直及水平尺寸来说视频块的像素尺寸(例如,8像素×8像素)。可按列及行来排列块中的像素。在一些实施例中,块在水平方向上可不具有与在

垂直方向上相同的像素数目。与CU相关联的语法数据可描述例如将CU分割成一或多个PU。分割模式可在CU经帧内预测模式编码或经帧间预测模式编码之间有所不同。PU可分割成非正方形形状。与CU相关联的语法数据也可描述(例如)根据CTU将CU分割成一或多个TU。TU可为正方形或非正方形形状。

[0120] 根据HEVC标准,可使用变换单元(TU)来执行变换。TU可针对不同CU而变化。可基于给定CU内的PU的大小而对TU设定大小。TU可与PU大小相同或小于PU。在一些实例中,可使用被称为残余四分树(RQT)的四分树结构将对应于CU的残余样本再分成较小单元。RQT的叶节点可对应于TU。可变换与TU相关联的像素差值以产生变换系数。变换系数可随后由编码器引擎906量化。

[0121] 一旦视频数据的图像被分割成CU,编码器引擎906即使用预测模式来预测每一PU。随后从原始视频数据中减去预测以得到残余(下文描述)。对于每一CU,可使用语法数据在位流内部发信预测模式。预测模式可包含帧内预测(或图像内预测)或帧间预测(或图像间预测)。使用帧内预测时,使用(例如)用以发现PU的平均值的DC预测、用以使平面表面拟合于PU的平面预测、用以从相邻数据外插的方向预测或任何其它合适类型的预测从同一图像中的相邻图像数据预测每一PU。使用帧间预测时,使用运动补偿预测从一或多个参考图像(按输出次序在当前图像之前或之后)中的图像数据预测每一PU。可(例如)以CU级别作出使用图像间预测或图像内预测来译码图像区域的决策。在一些实例中,图像的一或多个切片被指派有切片类型。切片类型包含I切片、P切片及B切片。I切片(帧内,可独立地解码)为仅仅通过帧内预测译码的图像的切片,且因此可独立地解码,这是由于I切片仅仅需要帧内的数据来预测切片的任一块。P切片(单向预测帧)为可使用帧内预测及单向帧间预测译码的图像的切片。P切片内的每一块系使用帧内预测或帧间预测而译码。当帧间预测应用时,块仅仅通过一个参考图像而预测,且因此参考样本仅仅来自一个帧的一个参考区。B切片(双向预测性帧)为可使用帧内预测及帧间预测译码的图像的切片。B切片的块可从两个参考图像双向预测,其中每一图像贡献一个参考区且两个参考区的样本集经加权(例如,使用相等权重)以产生经双向预测块的预测信号。如上文所解释,一个图像的切片经独立译码。在一些情况下,图像可仅作为一个切片而被译码。

[0122] PU可包含与预测过程有关的数据。举例来说,当使用帧内预测编码PU时,PU可包含描述用于PU的帧内预测模式的数据。作为另一实例,当使用帧间预测编码PU时,PU可包含定义PU的运动向量的数据。定义PU的运动向量的数据可描述(例如)运动向量的水平分量、运动向量的垂直分量、运动向量的解析度(例如,四分之一像素精度或八分之一像素精度)、运动向量所指向的参考图像,及/或运动向量的参考图像列表(例如,列表0、列表1或列表C)。

[0123] 编码装置904随后可执行变换及量化。举例来说,在预测之后,编码器引擎906可计算对应于PU的残余值。残余值可包括像素差值。在预测执行之后可能剩余的任何残余数据系使用块变换进行变换,所述块变换可基于离散余弦变换、离散正弦变换、整数变换、小波变换或其它合适的变换函数。在一些情况下,一或多个块变换(例如,大小 $32 \times 32$ 、 $16 \times 16$ 、 $8 \times 8$ 、 $4 \times 4$ ,或类似者)可应用于每一CU中的残余数据。在一些实施例中,TU可用于由编码器引擎906实施的变换及量化过程。具有一或多个PU的给定CU也可包含一或多个TU。如下文进一步详细描述,可使用块变换将残余值转换成变换系数,且随后可使用TU来量化及扫描残余值以产生用于熵译码的串列化变换系数。

[0124] 在一些实施例中，在使用CU的PU进行帧内预测性或帧间预测性译码之后，编码器引擎906可计算CU的TU的残余数据。PU可包括空间域(或像素域)中的像素数据。在应用块变换之后，TU可包括变换域中的系数。如先前所提及，残余数据可对应于未经编码图像的像素与对应于PU的预测值之间的像素差值。编码器引擎906可形成包含CU的残余数据的TU，且可随后变换TU以产生CU的变换系数。

[0125] 编码器引擎906可执行变换系数的量化。量化通过量化变换系数以减少用于表示系数的数据的量而提供进一步压缩。举例来说，量化可减少与系数中的一些或全部相关联的位深度。在一个实例中，具有n位值的系数可在量化期间下舍入至m位值，其中n大于m。

[0126] 一旦执行量化，则经译码视频位流包含经量化变换系数、预测信息(例如，预测模式、运动向量或其类似者)、分割信息及任何其它合适的数据(诸如其它语法数据)。随后可通过编码器引擎906熵编码经译码视频位流的不同元素。在一些实例中，编码器引擎906可利用预定义扫描次序来扫描经量化变换系数以产生可经熵编码的串列化向量。在一些实例中，编码器引擎906可执行自适应扫描。在扫描经量化变换系数以形成向量(例如，一维向量)之后，编码器引擎906可熵编码向量。举例来说，编码器引擎906可使用上下文自适应性可变长度译码、上下文自适应性二进位算术编码、基于语法的上下文自适应性二进位算术编码、机率区间分割熵译码或另一合适的熵编码技术。

[0127] 编码装置904的输出端910可经由通信链路920将组成经编码视频位流数据的NAL单元发送至接收装置的解码装置912。解码装置912的输入端914可接收NAL单元。通信链路920可包含由无线网络、有线网络或有线与无线网络的组合提供的通道。无线网络可包含任何无线接口或无线接口的组合，且可包含任何合适的无线网络(例如，因特网或其它广域网、分组网络、WiFi<sup>TM</sup>、射频(RF)、UWB、WiFi直连(WiFi-Direct)、蜂窝、长期演进(LTE)、WiMax<sup>TM</sup>或其类似者)。有线网络可包含任何有线接口(例如，光纤、以太网、电力线以太网、经由同轴电缆的以太网、数字信号线(DSL)或其类似者)。可使用各种设备来实施有线及/或无线网络，所述设备诸如基站、路由器、存取点、桥接器、闸道器、交换器或其类似者。可根据通信标准(诸如，无线通信协议)调制经编码视频位流数据，且将其发射至接收装置。

[0128] 在一些实例中，编码装置904可将经编码视频位流数据存储于存储器908中。输出端910可从编码器引擎906或从存储器908检索经编码视频位流数据。存储器908可包含多种分布式或本地存取数据存储媒体中的任一者。举例来说，存储器908可包含硬盘驱动器、存储光盘、快闪存储器、易失性或非易失性存储器或用于存储经编码视频数据的任何其它合适的数字存储媒体。

[0129] 解码装置912的输入端914接收经编码视频位流数据，且可将视频位流数据提供至解码器引擎916，或提供至存储器918以供稍后由解码器引擎916使用。解码器引擎916可通过熵解码(例如，使用熵解码器)及提取组成经编码视频数据的一或多个经译码视频序列的元素而解码经编码视频位流数据。解码器引擎916可随后重新按比例调整经编码视频位流数据且对经编码视频位流数据执行反变换。残余数据随后传递至解码器引擎916的预测阶段。解码器引擎916随后预测像素的块(例如，PU)。在一些实例中，预测被添加至反变换的输出(残余数据)。

[0130] 解码装置912可将经解码视频输出至视频目的地装置922，视频目的地装置可包含用于将经解码视频数据显示给内容的消费者的显示器或其它输出装置。在一些方面中，视

频目的地装置922可为包含解码装置912的接收装置的部分。在一些方面中,视频目的地装置922可为除接收装置外的分离装置的部分。

[0131] 补充增强信息(SEI)消息可包含于视频位流中。举例来说,SEI消息可用于载送不必要的信息(例如,后设数据)以便通过解码装置912解码位流。此信息适用于改进经解码输出的显示或处理(例如,此类信息可由解码器侧实体使用以改进内容的可视性)。

[0132] 在一些实施例中,视频编码装置904及/或视频解码装置912可分别与音频编码装置及音频解码装置集成。视频编码装置904及/或视频解码装置912也可包含实施上文所描述的译码技术所必需的其它硬件或软件,诸如一或多个微处理器、数字信号处理器(DSP)、专用集成电路(ASIC)、现场可编程门阵列(FPGA)、离散逻辑、软件、硬件、固件或其任何组合。视频编码装置904及视频解码装置912可集成为相应装置中的组合式编码器/解码器(编码解码器)的部分。

[0133] 对HEVC标准的扩展包含多视图视频译码扩展(被称作MV-HEVC)及可调式视频译码扩展(被称作SHVC)。MV-HEVC及SHVC扩展共享分层译码的概念,其中不同层包含于经编码视频位流中。经译码视频序列中的每一层由唯一层识别符(ID)定址。层ID可存在于NAL单元的标头中以识别NAL单元所相关联的层。在MV-HEVC中,不同层可表示视频位流中的同一场景的不同视图。在SHVC中,提供以不同空间解析度(或图像解析度)或不同重构建保真度表示视频位流的不同可调式层。可调式层可包含基础层(层ID=0)及一或多个增强层(层ID=1、2、…n)。基础层可符合HEVC的第一版本的设定档,且表示位流中的最低可用层。与基础层相比,增强层具有增加的空间解析度、时间解析度或帧速率及/或重构建保真度(或质量)。增强层经阶层式组织,且可(或可不)取决于较低层。在一些实例中,可使用单一标准编解码器来译码不同层(例如,使用HEVC、SHVC或其它译码标准编码全部层)。在一些实例中,可使用多标准编码解码器来译码不同层。举例来说,可使用AVC来译码基础层,而可使用对HEVC标准的SHVC及/或MV-HEVC扩展来译码一或多个增强层。一般来说,层包含VCL NAL单元集合及对应的非VCL NAL单元集合。NAL单元被指派特定层ID值。在层可取决于较低层的意义上,层可为阶层式的。

[0134] 一般来说,层包含VCL NAL单元集合及对应的非VCL NAL单元集合。NAL单元被指派特定层ID值。在层可取决于较低层的意义上,层可为阶层式的。层集合系指表示在位流内的独立的层集合,意味着在解码过程中层集合内的层可取决于层集合中的其它层,但并不取决于任何其它层来进行解码。因此,层集合中的层可形成可表示视频内容的独立位流。可通过子位流提取过程的操作从另一位流获得层集合中的层的集合。层集合可对应于待在解码器希望根据某些参数操作时被解码的层集合。

[0135] 由相机(例如,鱼眼相机或其它全向相机)捕捉的视频数据可经译码以减小用于发射及存储所需的数据的量。译码技术可实施于实例视频编码及解码系统(例如,系统900)中。在一些实例中,系统包含提供稍后由目的地装置解码的经编码视频数据的源装置。确切地说,源装置经由计算机可读媒体将视频数据提供至目的地装置。源装置及目的地装置可包括广泛范围的装置中的任一者,包含桌上型计算机、笔记型(也就是说,膝上型)计算机、平板计算机、机顶盒、电话手持机(诸如,所谓的“智能”电话)、所谓的“智能”板、电视机、相机、显示装置、数字媒体播放器、视频游戏控制台、视频流式发射装置或其类似者。在一些情况下,源装置及目的地装置可经装备以用于无线通信。

[0136] 编码装置1004及解码装置1112的特定细节分别展示于图10及图11中。图10为示出可实施本发明中所描述的技术中的一或更多的实例编码装置1004的框图。编码装置1004可(例如)产生本文中所描述的语法结构(例如,VPS、SPS、PPS或其它语法元素的语法结构)。编码装置1004可执行视频切片内的视频块的帧内预测及帧间预测译码。如先前所描述,帧内译码至少部分地依赖于空间预测以减少或移除给定视频帧或图像内的空间冗余。帧间译码至少部分地依赖于时间预测以减少或移除视频序列的邻近或周围帧内的时间冗余。帧内模式(I模式)可指代若干基于空间的压缩模式中的任一者。帧间模式(诸如,单向预测(P模式)或双向预测(B模式))可指代若干基于时间的压缩模式中的任一者。

[0137] 编码装置1004包含分割单元35、预测处理单元41、滤波器单元63、图像存储器64、求和器50、变换处理单元52、量化单元54及熵编码单元56。预测处理单元41包含运动估计单元42、运动补偿单元44及帧内预测处理单元46。对于视频块重构建,编码装置1004也包含反量化单元58、反变换处理单元60及求和器62。滤波器单元63意在表示一或多个回路滤波器(诸如,解块滤波器、自适应回路滤波器(ALF)及样本自适应偏移(SAO)滤波器)。虽然滤波器单元63在图10中展示为回路滤波器,但在其它配置中,滤波器单元63可实施为后回路滤波器。后处理装置57可对由编码装置1004产生的经编码视频数据执行额外处理。在一些情况下,本发明的技术可通过编码装置1004实施。然而,在其它情况下,本发明的技术中的一或更多者可通过后处理装置57实施。

[0138] 如图10中所展示,编码装置1004接收视频数据,且分割单元35将所述数据分割成视频块。分割也可包含(例如)根据LCU及CU的四分树结构分割成切片、切片片段、图像块或其它较大单元,以及视频块分割。编码装置1004通常示出编码待编码的视频切片内的视频块的组件。可将切片划分成多个视频块(且可能划分成被称作图像块的视频块集合)。预测处理单元41可基于错误结果(例如,译码速率及失真等级,或其类似者)选择多个可能的译码模式中的一者(诸如,多个帧内预测译码模式中的一者或多个帧间预测译码模式中的一者)以用于当前视频块。预测处理单元41可将所得经帧内或帧间译码块提供至求和器50以产生残余块数据且提供至求和器62以重构经编码块以用作参考图像。

[0139] 预测处理单元41内的帧内预测处理单元46可执行当前视频块相对于与待译码的当前块在相同的帧或切片中的一或多个相邻块的帧内预测译码,以提供空间压缩。预测处理单元41内的运动估计单元42及运动补偿单元44执行当前视频块相对于一或多个参考图像中的一或多个预测性块的帧间预测性译码,以提供时间压缩。

[0140] 运动估计单元42可经配置以根据视频序列的预定图案来确定用于视频切片的帧间预测模式。预定图案可将序列中的视频切片指定为P切片、B切片或GPB切片。运动估计单元42及运动补偿单元44可高度集成,但出于概念目的而分开说明。由运动估计单元42执行的运动估计为产生运动向量的过程,所述运动向量估计视频块的运动。运动向量(例如)可指示当前视频帧或图像内的视频块的预测单元(PU)相对于参考图像内的预测性块的移位。

[0141] 预测性块为被发现在像素差方面紧密地匹配待译码视频块的PU的块,所述像素差可通过绝对差总和(SAD)、平方差总和(SSD)或其它差量度确定。在一些实例中,编码装置1004可计算存储于图像存储器64中的参考图像的次整数像素位置的值。举例来说,编码装置1004可内插所述参考图像的四分之一像素位置、八分之一像素位置或其它分率像素位置的值。因此,运动估计单元42可执行相对于全像素位置及分数像素位置的运动搜寻且输出

具有分数像素精确度的运动向量。

[0142] 运动估计单元42通过将PU的位置与参考图像的预测性块的位置比较而计算经帧间译码切片中的视频块的PU的运动向量。参考图像可选自第一参考图像列表(列表0)或第二参考图像列表(列表1),所述参考图像列表中的每一者识别存储在图像存储器64中的一或多个参考图像。运动估计单元42将所计算运动向量发送至熵编码单元56及运动补偿单元44。

[0143] 由运动补偿单元44所执行的运动补偿可涉及基于由运动估计所确定的运动向量而提取或产生预测性块,可能执行子像素精确度的内插。在接收到当前视频块的PU的运动向量之后,运动补偿单元44可在参考图像列表中定位运动向量指向的预测性块。编码装置1004通过从正经译码的当前视频块的像素值减去预测性块的像素值从而形成像素差值来形成残余视频块。像素差值形成用于块的残余数据,且可包含明度差分量及色度差分量两者。求和器50表示执行此减法运算的组件。运动补偿单元44也可产生与视频块及视频切片相关联的语法元素,以供图11的解码装置1112用于解码视频切片的视频块。

[0144] 图10的帧内预测处理单元46可对当前块进行帧内预测,作为如上文所描述的通过运动估计单元42及运动补偿单元44执行的帧间预测的替代方案。确切地说,帧内预测处理单元46可确定帧内预测模式以用以编码当前块。在一些实例中,帧内预测处理单元46可(例如)在分开的编码遍次期间使用各种帧内预测模式来编码当前块,且帧内预测处理单元46(或在一些实例中,模式选择单元40)可从所测试模式选择适当帧内预测模式来使用。举例来说,帧内预测处理单元46可使用对各种所测试帧内预测模式的速率-失真分析来计算速率-失真值,且可在所测试模式间选择具有最佳速率-失真特性的帧内预测模式。速率-失真分析大体上确定经编码块与原始、未编码块(其经编码以产生经编码块)之间的失真(或误差)量,以及用以产生经编码块的位速率(也就是说,位的数目)。帧内预测处理单元46可根据各种经编码块的失真及速率计算比率以确定哪一帧内预测模式展现所述块的最佳速率-失真值。

[0145] 在任何情况下,在选择用于块的帧内预测模式之后,帧内预测处理单元46可将指示用于块的选定帧内预测模式的信息提供至熵编码单元56。熵编码单元56可编码指示选定帧内预测模式的信息。编码装置1004可将各种块的编码上下文的定义以及待用于上下文中的每一者的最可能的帧内预测模式、帧内预测模式索引表及经修改帧内预测模式索引表的指示包含于经发射位流配置数据中。位流配置数据可包含多个帧内预测模式索引表及多个经修改帧内预测模式索引表(也被称作码字映射表)。

[0146] 在预测处理单元41经由帧间预测或帧内预测产生当前视频块的预测性块之后,编码装置1004通过从当前视频块减去预测性块而形成残余视频块。残余块中的残余视频数据可包含于一或多个TU中且应用于变换处理单元52。变换处理单元52使用诸如离散余弦变换(DCT)或概念上类似的变换的变换将残余视频数据变换为残余变换系数。变换处理单元52可将残余视频数据从像素域变换至变换域(诸如,频域)。

[0147] 变换处理单元52可将所得变换系数发送至量化单元54。量化单元54量化变换系数以进一步减小位速率。量化过程可减小与所述系数中的一些或全部相关联的位深度。可通过调整量化参数来修改量化程度。在一些实例中,量化单元54随后可执行对包含经量化的变换系数的矩阵的扫描。替代地,熵编码单元56可执行所述扫描。

[0148] 在量化之后,熵编码单元56对经量化变换系数进行熵编码。举例来说,熵编码单元56可执行上下文自适应性可变长度译码(CAVLC)、上下文自适应性二进位算术译码(CABAC)、基于语法的上下文自适应性二进位算术译码(SBAC)、机率区间分割熵(PIPE)译码或另一熵编码技术。在由熵编码单元56进行熵编码之后,可将经编码位流发射至解码装置1112,或经存档以供稍后由解码装置1112发射或检索。熵编码单元56也可熵编码正经译码的当前视频切片的运动向量及其它语法元素。

[0149] 反量化单元58及反变换处理单元60分别应用反量化及反变换以重建像素域中的残余块以供稍后用作参考图像的参考块。运动补偿单元44可通过将残余块添加至参考图像列表内的参考图像中的一者的预测性块来计算参考块。运动补偿单元44也可将一或多个内插滤波器应用于经重构残余块以计算次整数像素值以用于运动估计。求和器62将经重构的残余块添加至由运动补偿单元44产生的运动补偿预测块以产生参考块以用于存储于图像存储器64中。参考块可由运动估计单元42及运动补偿单元44用作为参考块以对后续视频帧或图像中的块进行帧间预测。

[0150] 以此方式,图10的编码装置1004表示经配置以产生经编码视频位流的语法的视频编码器的实例。编码装置1004可(例如)产生如上文所描述的VPS、SPS及PPS参数集。编码装置1004可执行本文中所描述的技术中的任一者,包含上文关于图10及图11描述的过程。本发明的技术已大体上关于编码装置1004加以描述,但如上文所提及,本发明的技术中的一些也可通过后处理装置57实施。

[0151] 图11为示出实例解码装置1112的框图。解码装置1112包含熵解码单元80、预测处理单元81、反量化单元86、反变换处理单元88、求和器90、滤波器单元91及图像存储器92。预测处理单元81包含运动补偿单元82及帧内预测处理单元84。在一些实例中,解码装置1112可执行大体上与关于来自图10的编码装置1004描述的编码遍次互逆的解码遍次。

[0152] 在解码过程期间,图11的解码装置1112接收表示由图10的编码装置1004发送的经编码视频切片及相关联语法元素的视频块的经编码视频位流。在一些实施例中,图11的解码装置1112可从编码装置1004接收经编码视频位流。在一些实施例中,解码装置1112可从诸如以下的网络实体79接收经编码视频位流:服务器、媒体感知网络元件(MANE)、视频编辑器/剪接器或经配置以实施上文所描述的技术中的一或者者的其它此装置。网络实体79可包含或可不包含编码装置1004。本发明中所描述的技术中的一些可在网络实体79将经编码视频位流发射至解码装置1112之前由网络实体79实施。在一些视频解码系统中,网络实体79及解码装置1112可为单独装置的部分,而在其它情况下,关于网络实体79所描述的功能性可通过包括解码装置1112的相同装置进行执行。

[0153] 解码装置1112的熵解码单元80对位流进行熵解码以产生经量化系数、运动向量及其它语法元素。熵解码单元80将运动向量及其它语法元素转发至预测处理单元81。解码装置1112可接收视频切片级别及/或视频块级别的语法元素。熵解码单元80可处理并解析诸如VPS、SPS及PPS集合的一或多个参数集中的固定长度语法元素及可变长度语法元素两者。

[0154] 当将视频切片译码为经帧内译码(I)切片时,预测处理单元81的帧内预测处理单元84可基于发信号通知的帧内预测模式及来自当前帧或图像的先前经解码块的数据而产生当前视频切片的视频块的预测数据。当视频帧经译码为经帧间译码(也就是说,B、P或GPB)切片时,预测处理单元81的运动补偿单元82基于从熵解码单元80接收的运动向量及其

它语法元素而产生当前视频切片的视频块的预测性块。可从参考图像列表内的参考图像中的一者产生预测性块。解码装置1112可基于存储于图像存储器92中的参考图像使用默认构建技术来构建参考帧列表(列表0及列表1)。

[0155] 运动补偿单元82通过解析运动向量及其它语法元素来确定用于当前视频切片的视频块的预测信息,且使用所述预测信息以产生正经解码的当前视频块的预测性块。举例来说,运动补偿单元82可使用参数集中的一或多个语法元素来确定用于译码视频切片的视频块的预测模式(例如,帧内或帧间预测)、帧间预测切片类型(例如,B切片、P切片或GPB切片)、切片的一或多个参考图像列表的构建信息、切片的每一经帧间编码视频块的运动向量、切片的每一经帧间译码视频块的帧间预测状态及用以解码当前视频切片中的视频块的其它信息。

[0156] 运动补偿单元82也可执行基于内插滤波器的内插。运动补偿单元82可使用如在编码视频块期间由编码装置1004使用的内插滤波器来计算参考块的次整数像素的内插值。在此情况下,运动补偿单元82可从所接收语法元素确定由编码装置1004使用的内插滤波器,且可使用内插滤波器来产生预测性块。

[0157] 反量化单元86反量化(或解量化)位流中所提供,且由熵解码单元80解码的经量化的变换系数。反量化过程可包含使用由编码装置1004针对视频切片中的每一视频块计算的量化参数以确定应应用的量化程度及(同样地)反量化程度。反变换处理单元88将反变换(例如,反DCT或其它合适的反变换)、反整数变换或概念上类似的反变换过程应用于变换系数以便在像素域中产生残余块。

[0158] 在运动补偿单元82基于运动向量及其它语法元素而产生当前视频块的预测性块之后,解码装置1112通过将来自反转换处理单元88的残余块与由运动补偿单元82所产生的相应预测性块求和,从而形成经解码视频块。求和器90表示执行此求和运算的一或多个组件。若需要,也可使用回路滤波器(在译码回路中或在译码回路后)以使像素转变平滑,或以其它方式改进视频质量。滤波器单元91意在表示一或多个回路滤波器,诸如,解块滤波器、自适应回路滤波器(ALF)及样本自适应偏移(SAO)滤波器。尽管滤波器单元91在图11中展示为回路内滤波器,但在其它配置中,滤波器单元91可实施为回路后滤波器。给定帧或图像中的经解码视频块随后存储于图像存储器92中,所述图像存储器存储用于后续运动补偿的参考图像。图像存储器92也存储经解码视频以供稍后呈现于显示装置(诸如,图9中所展示的视频目的地装置922)上。

[0159] 在前述描述中,申请案的方面参考其特定实施例而描述,但所属领域的技术人员将认识到本发明不限于此。因此,尽管本文中已详细描述申请案的说明性实施例,但应理解,本发明概念可以其它方式不同地体现并使用,且所附权利要求书意在解释为包含除先前技术所限制外的这些变化。上文所描述的发明的各种特征及方面可单独地或联合地使用。另外,在不脱离本说明书的更广精神及范围的情况下,实施例可用于超出本文所描述的所述环境及应用的任何数目个环境及应用。因此,本说明书及图式被视为说明性而非限定性。出于说明的目的,以特定次序描述方法。应了解,在替代实施例中,可以与所描述的次序不同的次序执行所述方法。

[0160] 在组件被描述为“经配置以”执行某些操作的情况下,可(例如)通过设计用以执行操作的电子电路或其它硬件、通过编程用以执行操作的可编程电子电路(例如,微处理器或

其它合适的电子电路)或其任何组合来实现这些配置。

[0161] 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、电路及算法步骤可被实施为电子硬件、计算机软件或两者的组合。为了清楚地说明硬件与软件的此可互换性,各种说明性组件、块、模块、电路及步骤已在上文大体按其功能性加以描述。将此功能实施为硬件还是软件取决于施加于整个系统上的特定应用及设计约束。所属领域的技术人员可针对每一特定应用而以不同方式来实施所描述功能性,但这些实施决策不应被解译为导致脱离本发明的范围。

[0162] 本文所描述的技术可以硬件、软件、固件或其任何组合来实施。此技术可以诸如以下各者的多种装置中的任一者实施:通用计算机、无线通信装置手持机,或具有包含无线通信装置手持机及其它装置中的应用的多个用途的集成电路装置。可将描述为模块或组件的任何特征一起实施于集成式逻辑装置中或分开来实施为离散但可互操作的逻辑装置。若以软件实施,则所述技术可至少部分由包括包含当经执行时执行上文所描述方法中的一或多个者的指令的程序代码的计算机可读数据存储媒体实现。计算机可读数据存储媒体可形成计算机程序产品的部分,计算机程序产品可包含封装材料。计算机可读媒体可包括存储器或数据存储媒体,诸如,随机存取存储器(RAM),诸如,同步动态随机存取存储器(SDRAM)、只读存储器(ROM)、非易失性随机存取存储器(NVRAM)、电可擦除可编程只读存储器 EEPROM)、快闪存储器、磁性或光学数据存储媒体及类似者。所述技术另外或替代地可至少部分由计算机可读通信媒体实现,所述计算机可读通信媒体携载或传达呈指令或数据结构的形式且可由计算机存取、读取及/或执行的程序代码,诸如,传播的信号或波。

[0163] 程序代码可由可包含一或多个处理器的处理器执行,诸如,一或多个数字信号处理器(DSP)、通用微处理器、专用集成电路(ASIC)、现场可编程逻辑阵列(FPGA)或其它等效集成式或离散逻辑电路。此处理器可经配置以执行本发明中所描述的技术中的任一者。通用处理器可为微处理器;但在替代方案中,处理器可为任何常规处理器、控制器、微控制器或状态机。处理器也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、一或多个微处理器结合DSP核心或任何其它此配置。因此,如本文中所使用的术语“处理器”可指代前述结构、前述结构的任何组合或适合于实施本文中描述的技术的任何其它结构或设备中的任一者。此外,在一些方面中,本文中描述的功能性可提供于经配置用于编码及解码的专用软件模块或硬件模块,或并入于组合式视频编码器-解码器(编码解码器)中。

[0164] 本文中所论述的译码技术可体现于实例视频编码及解码系统中。系统包含通过目的地装置提供待稍后经解码的经编码视频数据的源装置。确切地说,源装置经由计算机可读媒体将视频数据提供至目的地装置。源装置及目的地装置可包括广泛范围的装置中的任一者,包含桌上型计算机、笔记型(即,膝上型)计算机、平板计算机、机顶盒、电话手持机(诸如所谓的“智能”电话)、所谓的“智能”板、电视机、相机、显示装置、数字媒体播放器、视频游戏控制台、视频流式发射装置或类似者。在一些情况下,源装置及目的地装置可经装备以用于无线通信。

[0165] 目的地装置可经由计算机可读媒体接收待解码的经编码视频数据。计算机可读媒体可包括能够将经编码视频数据从源装置移动至目的地装置的任何类型的媒体或装置。在一个实例中,计算机可读媒体可包括通信媒体以使得源装置能够实时地将经编码视频数据直接发射至目的地装置。可根据通信标准(诸如,无线通信协议)调制经编码视频数据,且将

经编码视频数据发射至目的地装置。通信媒体可包括任何无线或有线通信媒体，诸如，射频(RF)频谱或一或多个物理发射线。通信媒体可形成分组网络(诸如，局域网、广域网或诸如因特网的全域网)的部分。通信媒体可包含路由器、交换器、基站或可适用于促进从源装置至目的地装置的通信的任何其它设备。

[0166] 在一些实例中，经编码数据可从输出接口输出至存储装置。类似地，可由输入接口从存储装置存取经编码数据。存储装置可包含多种分布式或本地存取数据存储媒体中的任一者，诸如硬盘驱动器、蓝光光盘、DVD、CD-ROM、快闪存储器、易失性或非易失性存储器或用于存储经编码视频数据的任何其它合适的数字存储媒体。在另一实例中，存储装置可对应于文件服务器或可存储由源装置产生的经编码视频的另一中间存储装置。目的地装置可经由流式发射或下载从存储装置存取所存储的视频数据。文件服务器可为能够存储经编码视频数据且将所述经编码视频数据发射至目的地装置的任何类型的服务器。实例文件服务器包含网页服务器(例如，用于网站)、FTP服务器、网络附加存储(NAS)装置或本地磁盘驱动器。目的地装置可经由任何标准数据连接(包含因特网连接)来存取经编码视频数据。此数据连接可包含适于存取存储于文件服务器上的经编码视频数据的无线通道(例如，Wi-Fi连接)、有线连接(例如，DSL、缆线数据机等)，或两者的结合。来自存储装置的经编码视频数据的发射可为流式发射、下载发射或其组合。

[0167] 本发明的技术不必限于无线应用或设定。所述技术可应用于支持多种多媒体应用中的任一者的视频译码，诸如，空中电视广播、有线电视发射、卫星电视发射、因特网流式视频发射(诸如，经由HTTP的动态自适应流式发射(DASH))、经编码至数据存储媒体上的数字视频、存储在数据存储媒体上的数字视频的解码或其它应用。在一些实例中，系统可经配置以支持单向或双向视频发射以支持诸如视频流式发射、视频播放、视频广播及/或视频电话的应用。

[0168] 在一个实例中，源装置包括视频源、视频编码器及输出接口。目的地装置可包含输入接口、视频解码器及显示装置。源装置的视频编码器可经配置以应用本文中所揭示的技术。在其它实例中，源装置及目的地装置可包含其它组件或布置。举例来说，源装置可从外部视频源(诸如，外部相机)接收视频数据。同样地，目的地装置可与外部显示装置介接，而不包含集成式显示装置。

[0169] 以上实例系统仅为一个实例。用于并行地处理视频数据的技术可由任何数字视频编码及/或解码装置执行。尽管本发明的技术通常由视频编码装置执行，但所述技术也可由视频编码器/解码器(通常被称作“编解码器(CODEC)”)执行。此外，本发明的技术也可由视频预处理器执行。源装置及目的地装置仅为源装置产生经译码视频数据以供发射至目的地装置的此类译码装置的实例。在一些实例中，源装置及目的地装置可以大体上对称的方式操作，使得所述装置中的每一者包含视频编码及解码组件。因此，实例系统可支持视频装置之间的单向或双向视频发射，例如用于视频流式发射、视频播放、视频广播或视频电话。

[0170] 视频源可包含视频捕捉装置，诸如视频相机、含有先前捕捉的视频的视频存档及/或用以从视频内容提供者接收视频的视频馈入接口。作为另一替代方案，视频源可产生基于计算机图形的数据作为源视频，或实况视频、存档视频及计算机产生的视频的组合。在一些情况下，若视频源为视频相机，则源装置及目的地装置可形成所谓的相机电话或视频电话。然而，如上文所提及，本发明中描述的技术通常可适用于视频译码，且可适用于无线及/

或有线应用。在每一情况下,可由视频编码器编码所捕捉、经预捕捉或计算机产生的视频。经编码视频信息可随后由输出接口输出至计算机可读媒体上。

[0171] 如所提及,计算机可读媒体可包含暂态媒体,诸如无线广播或有线网络发射;或存储媒体(也就是说,非暂时性存储媒体),诸如硬盘、随身碟、紧密光盘、数字视频光盘、蓝光光盘或其它计算机可读媒体。在一些实例中,网络服务器(图中未展示)可从源装置接收经编码视频数据,且(例如)经由网络发射将经编码视频数据提供至目的地装置。类似地,媒体生产设施(诸如,光盘冲压设施)的计算装置可从源装置接收经编码视频数据且生产含有经编码视频数据的光盘。因此,在各种实例中,计算机可读媒体可理解为包含多种形式的一或多个计算机可读媒体。

[0172] 目的地装置的输入接口从计算机可读媒体接收信息。计算机可读媒体的信息可包含由视频编码器定义的语法信息(其也由视频解码器使用),所述语法信息包含描述块及其它经译码单元(例如,图像群组(GOP))的特性及/或处理的语法元素。显示装置将经解码视频数据显示给用户,且可包括多种显示装置中的任一者,诸如阴极射线管(CRT)、液晶显示器(LCD)、等离子体显示器、有机发光二极管(OLED)显示器或另一类型的显示装置。本申请案的各种实施例已经描述。

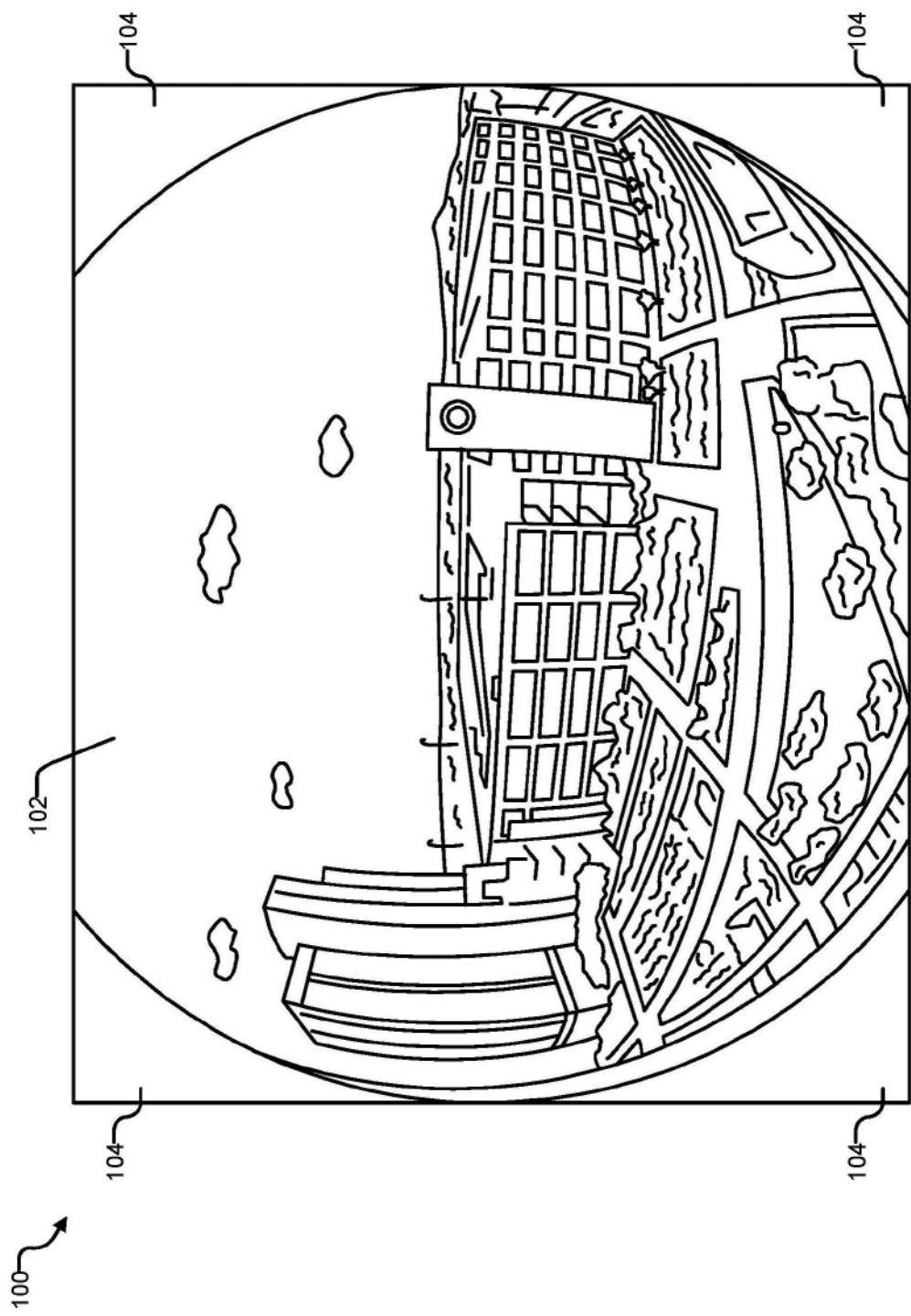


图1

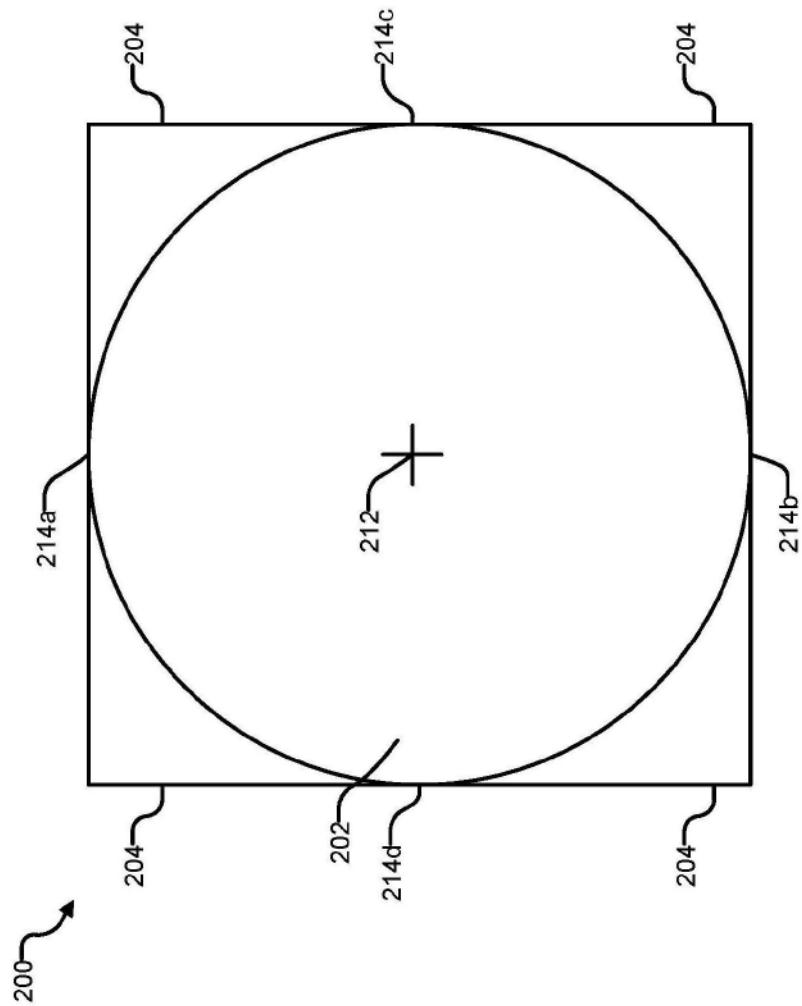


图2A

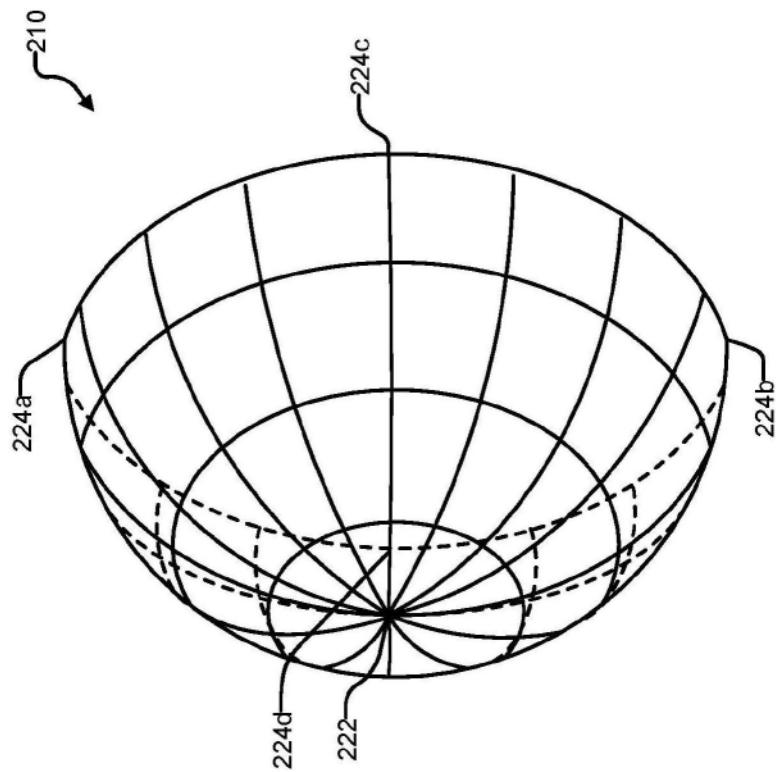


图2B

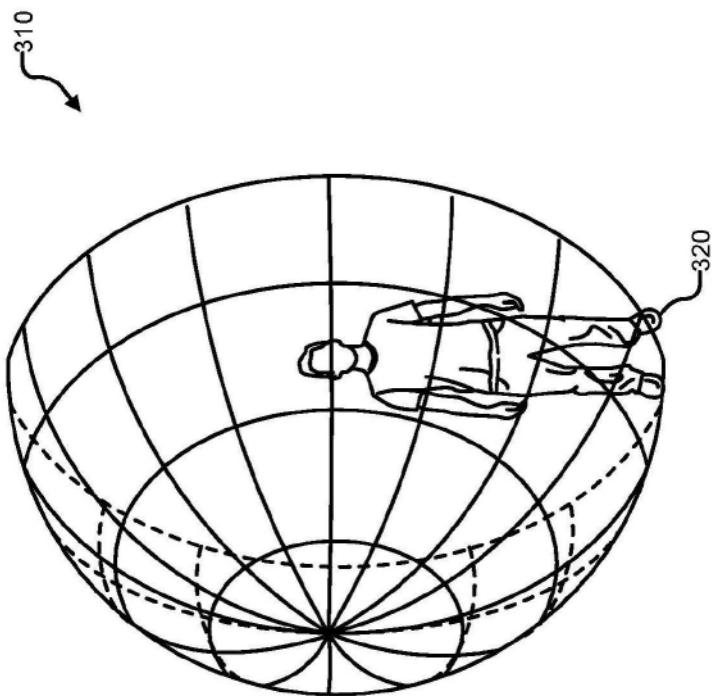


图3

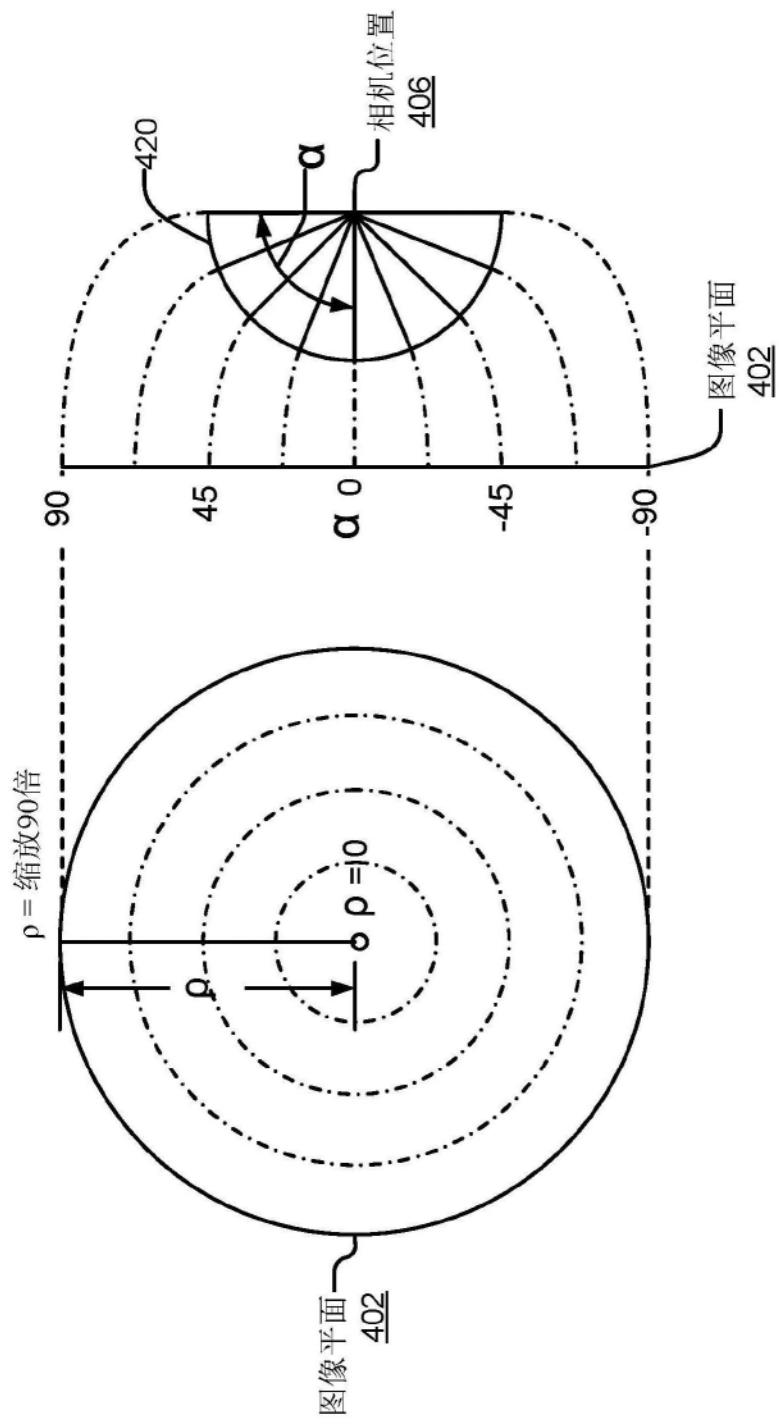


图4

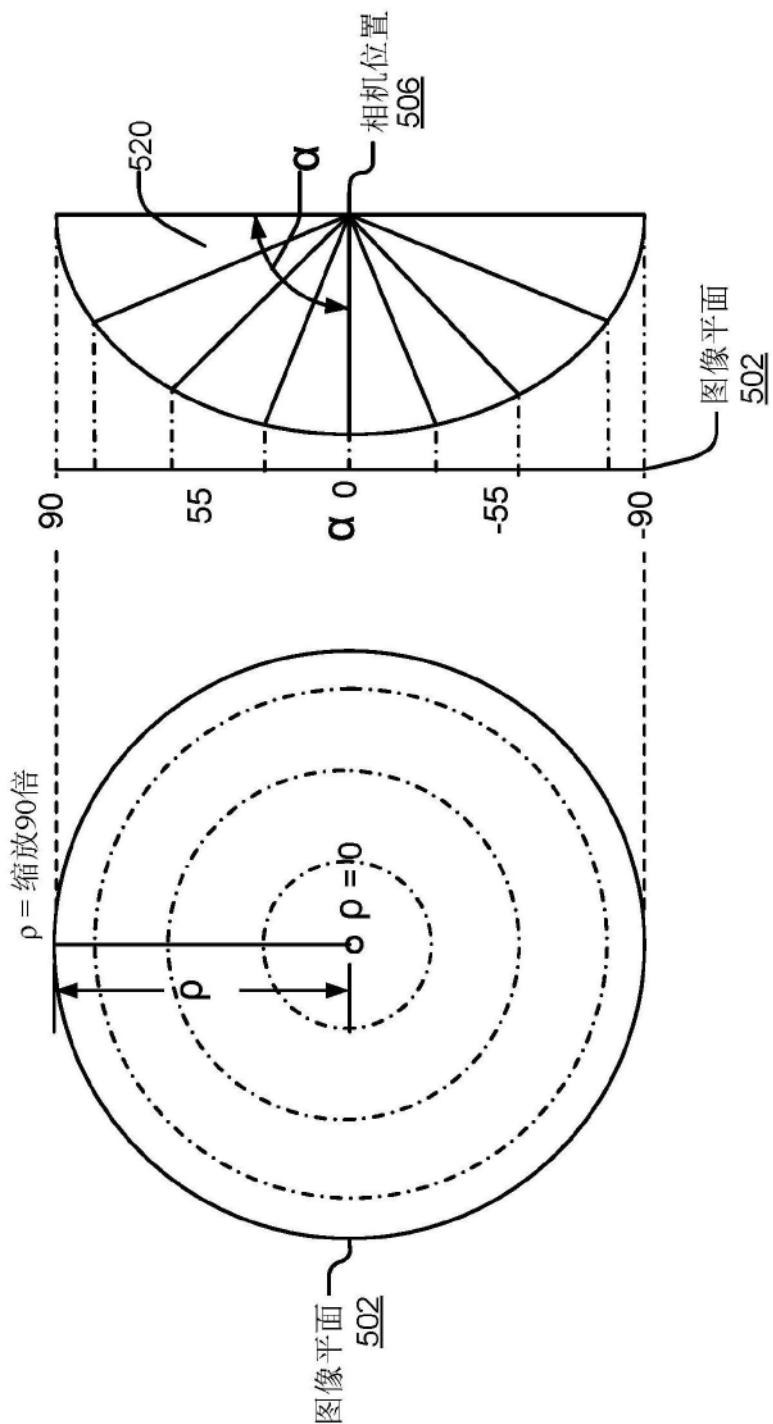


图5

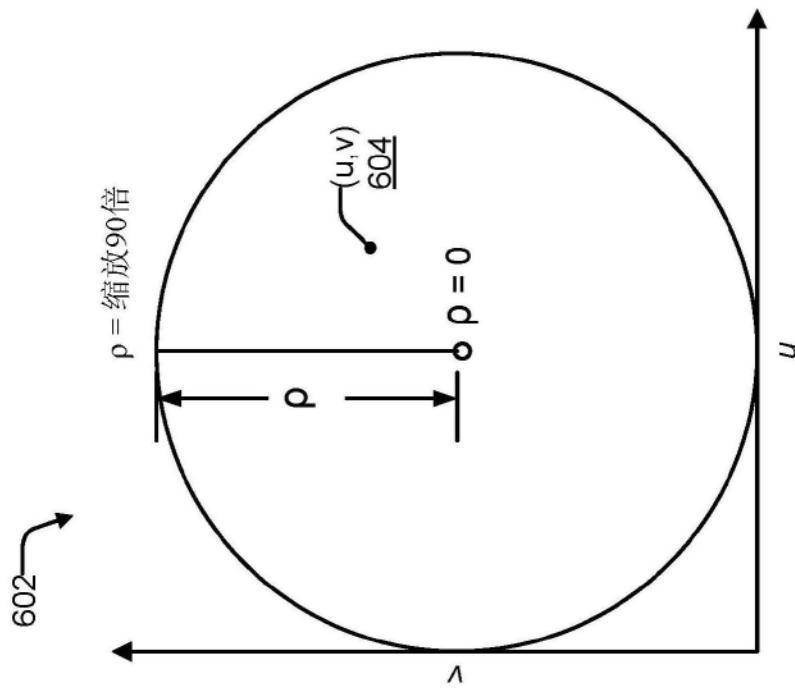


图6A

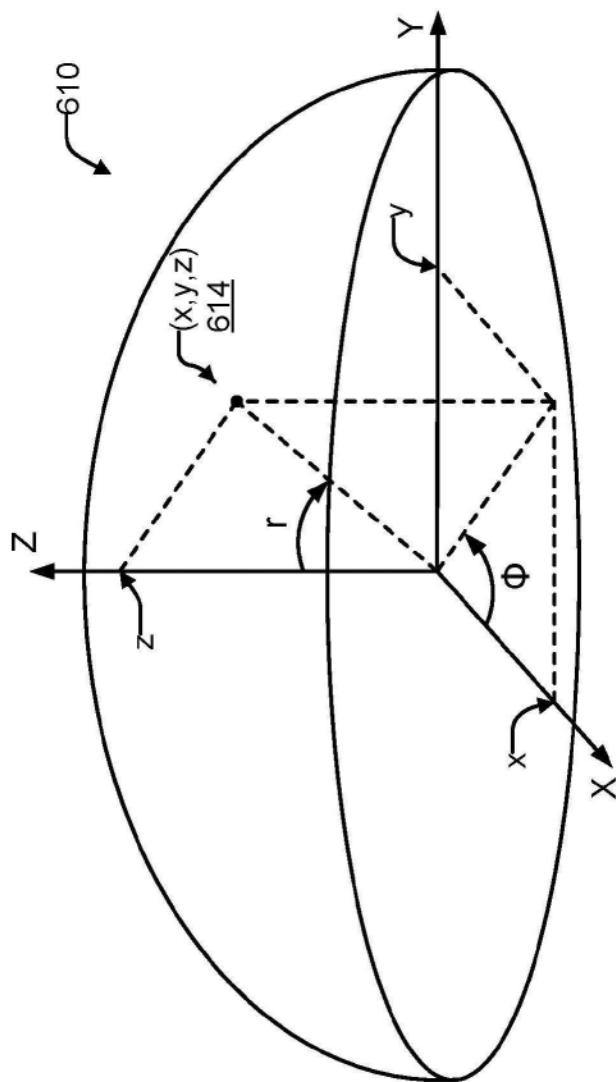


图6B

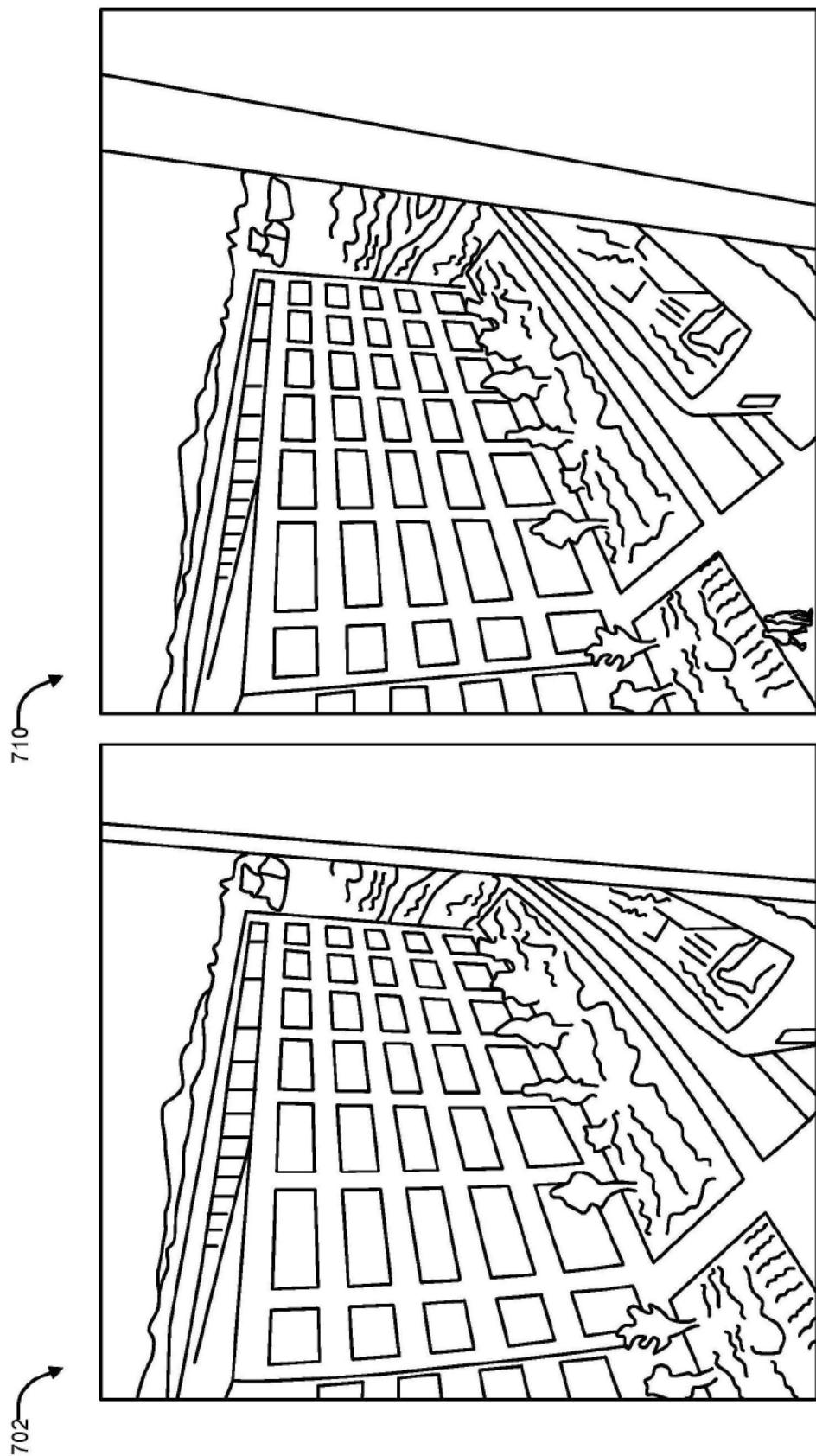


图7B

图7A

800

获得由全向相机捕捉的2维视频数据，其中所述2维视频数据包含场景的图像，其中在所述图像中，所述场景已变形成图像的圆形区域

802



将所述图像映射至3维半球形表示，其中映射图像包含将所述图像的圆形区域中的图元映射至3维半球形表示上的对应位置，其中映射图像校正所述场景的变形。

804



将所述3维半球形表示用于360度视频展示中

806

图8

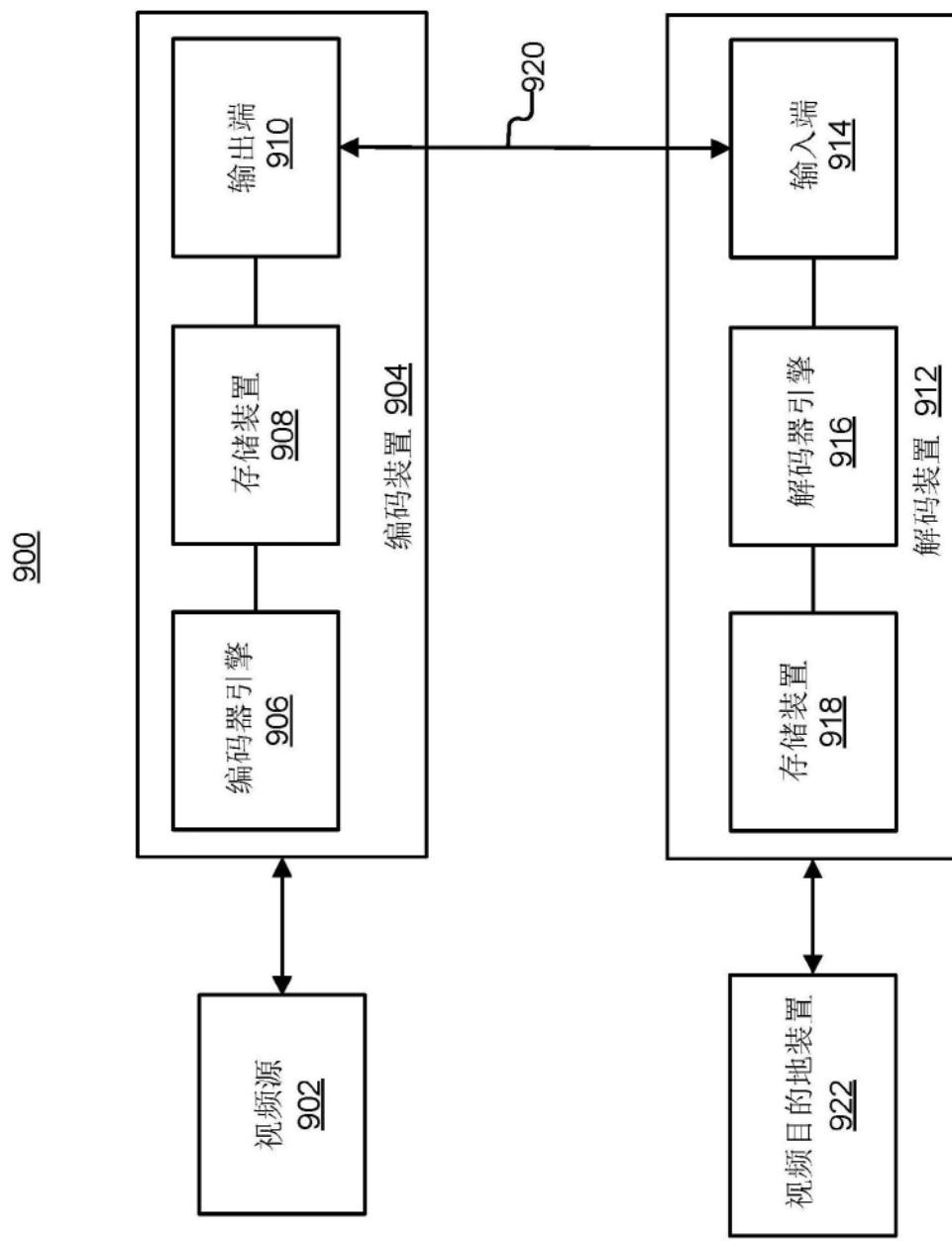


图9

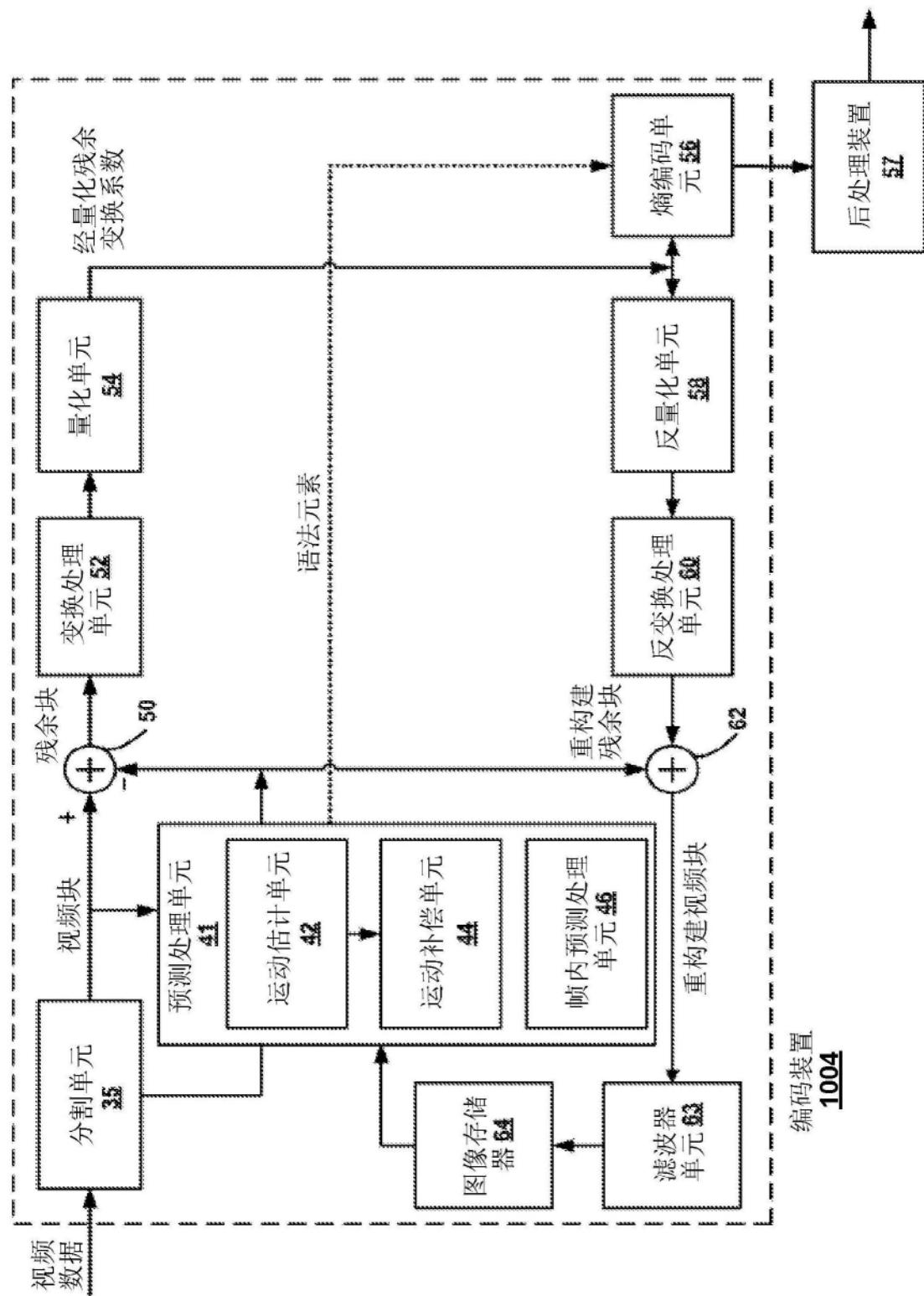


图10

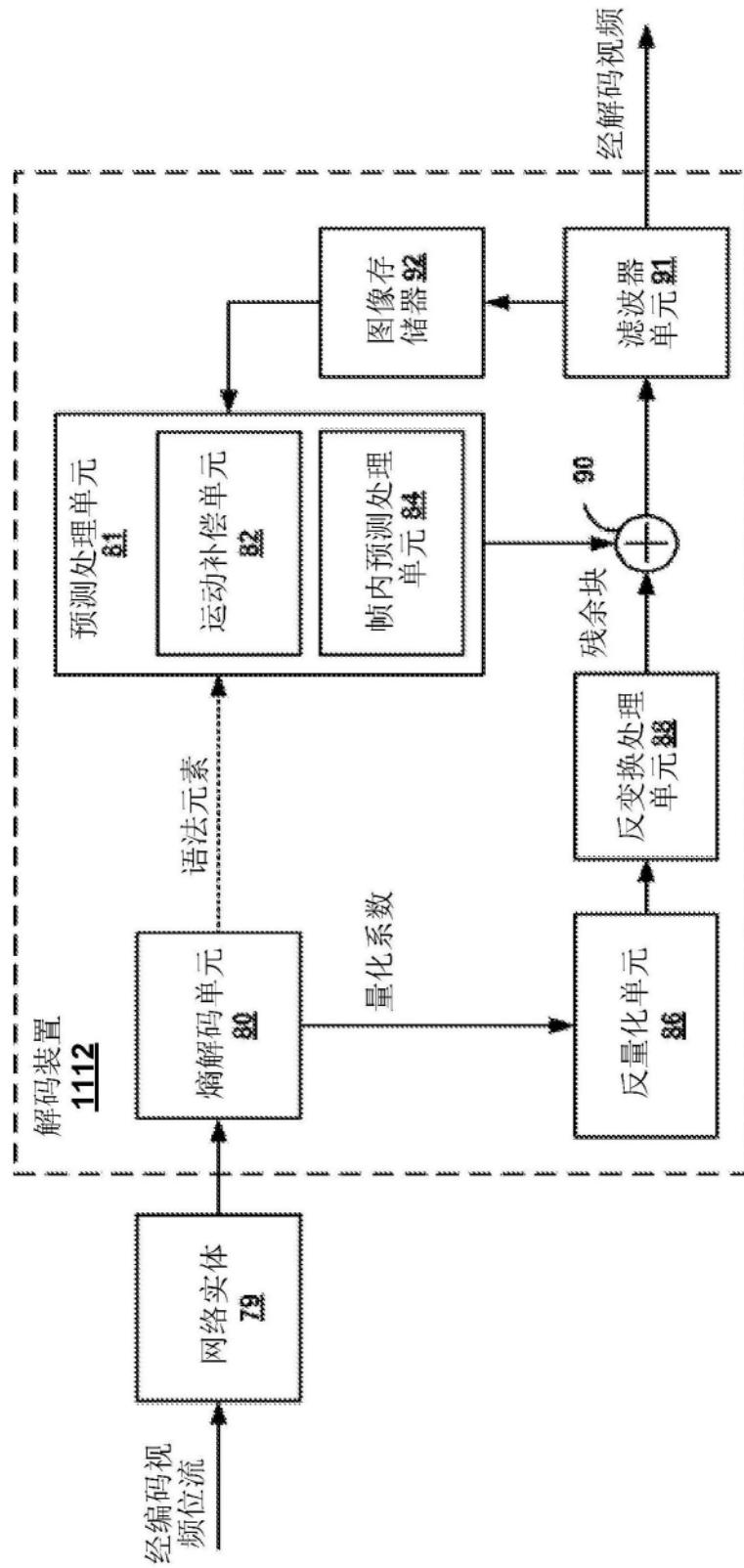


图11