



(12) 发明专利

(10) 授权公告号 CN 113366825 B

(45) 授权公告日 2024.06.14

(21) 申请号 202080011498.5

B·克龙

(22) 申请日 2020.01.17

(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

72002

申请公布号 CN 113366825 A

专利代理师 刘兆君

(43) 申请公布日 2021.09.07

(51) Int.Cl.

(30) 优先权数据

H04N 13/117 (2018.01)

19154195.2 2019.01.29 EP

H04N 21/218 (2011.01)

(85) PCT国际申请进入国家阶段日

H04N 19/597 (2014.01)

2021.07.29

H04N 13/243 (2018.01)

(86) PCT国际申请的申请数据

H04N 13/282 (2018.01)

PCT/EP2020/051075 2020.01.17

H04N 13/194 (2018.01)

H04N 13/178 (2018.01)

(87) PCT国际申请的公布数据

(56) 对比文件

W02020/156827 EN 2020.08.06

EP 3419301 A1, 2018.12.26

(73) 专利权人 皇家飞利浦有限公司

US 2017244948 A1, 2017.08.24

地址 荷兰艾恩德霍芬

US 2018316908 A1, 2018.11.01

(72) 发明人 W·H·A·布鲁斯 C·韦雷坎普

审查员 杨勤之

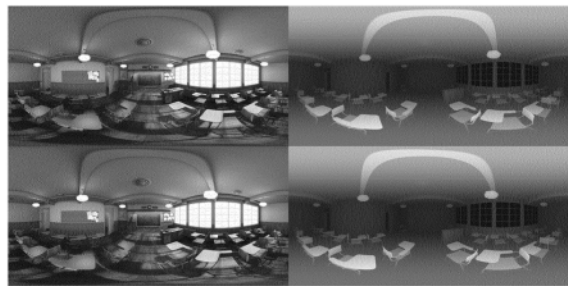
权利要求书2页 说明书17页 附图7页

(54) 发明名称

用于生成表示场景的图像信号的装置和方法

(57) 摘要

一种装置包括接收表示场景的图像信号的接收器(301)。所述图像信号包括图像数据,所述图像数据包括多幅图像,其中,每幅图像包括像素,所述像素表示沿着具有从射线原点出发的射线方向的射线的所述场景的图像属性。所述射线原点对于至少一些像素来说处于不同的位置。所述图像信号还包括多个参数,所述多个参数将针对像素的所述射线原点和/或所述射线方向的变化描述为像素图像位置的函数。绘制器(303)根据所述多幅图像并基于所述多个参数来绘制图像。



1. 一种用于根据表示场景的图像信号来绘制图像的装置,所述装置包括:
接收器(301、307),其用于接收所述图像信号,所述图像信号包括图像数据和元数据,
所述图像数据包括多幅图像,每幅图像包括像素,其中,每个像素表示沿着具有从射线
原点出发的射线方向的射线的所述场景的图像属性,其中,每个像素与射线的射线原点和
射线方向相链接,所述射线原点对于至少一些像素来说处于不同的位置,其中,针对一像素
的所述射线原点和所述射线方向指示针对该像素的从场景坐标到图像位置的投影,并且
所述元数据包括多个参数,所述多个参数将针对像素的所述射线原点和所述射线方向
中的至少一项的变化描述为像素图像位置的函数;以及
绘制器(303),其用于根据所述多幅图像并响应于所述多个参数来绘制图像。
2. 根据权利要求1所述的装置,其中,所述多幅图像中的至少一幅图像的水平像素行的
射线位置被映射到所述场景的水平平面中的曲线,所述多个参数描述所述曲线的属性。
3. 根据权利要求2所述的装置,其中,所述曲线是闭合曲线。
4. 根据权利要求3所述的装置,其中,所述闭合曲线是卵形,并且所述多个参数描述所
述卵形的属性。
5. 根据权利要求3所述的装置,其中,所述曲线是椭圆,并且所述多个参数描述所述椭
圆的属性。
6. 根据任一前述权利要求所述的装置,其中,所述多个参数由被链接到所述多幅图像
中的图像并具有比所述图像的分辨率更低的分辨率的映射图来提供,所述映射图具有指示
针对与所述映射图中的像素值的位置相对应的所述图像中的位置的射线原点和射线方向
中的至少一项的像素值;并且其中,所述绘制器(303)被布置为通过从所述映射图的所述像
素值进行内插来确定针对所述图像中的一些位置的射线原点和射线方向中的至少一项。
7. 根据任一前述权利要求所述的装置,其中,从图像位置到所述射线原点和所述射线
方向中的至少一项的映射是连续函数。
8. 根据任一前述权利要求所述的装置,其中,所述图像数据包括多幅图像,并且所述多
个参数描述用于将像素图像位置映射到针对所述多幅图像中的至少两幅图像的所述射线
原点和所述射线方向中的一项的不同函数。
9. 根据任一前述权利要求所述的装置,其中,所述多幅图像中的至少两幅图像表示部
分视场,所述部分视场对于所述至少两幅图像来说是不同的。
10. 根据任一前述权利要求所述的装置,其中,所述射线原点和所述射线方向中的至少
一项的所述变化是所述射线方向的变化。
11. 根据任一前述权利要求所述的装置,其中,所述射线原点和所述射线方向中的至少
一项的所述变化是所述射线原点的变化。
12. 根据任一前述权利要求所述的装置,其中,所述多幅图像包括第一光强度图像和针
对所述第一光强度图像的深度值图像,所述第一深度值图像包括针对所述第一光强度图像
的像素的深度值,针对所述第一光强度图像的第一像素的深度值指示沿着所述射线方向从
针对所述第一像素的射线原点到由所述第一像素表示的对象的距离;并且
所述绘制器(303)被布置为响应于所述第一光强度图像和所述第一深度值图像而绘制
图像。
13. 根据权利要求12所述的装置,其中,所述绘制器(303)被布置为:响应于所述深度值

图像而确定针对所述第一光强度图像的图像对象的场景位置,并且响应于所述场景位置而绘制图像。

14. 一种用于生成表示场景的图像信号的装置,所述装置包括:

第一生成器(201),其用于生成包括多幅图像的图像数据,每幅图像包括像素,其中,每个像素表示沿着具有从射线原点出发的射线方向的射线的所述场景的图像属性,其中,每个像素与射线的射线原点和射线方向相链接,所述射线原点对于至少一些像素来说处于不同的位置,其中,针对一像素的所述射线原点和所述射线方向指示针对该像素的从场景坐标到图像位置的投影;以及

第二生成器(205),其用于生成元数据,所述元数据包括多个参数,所述多个参数将针对像素的所述射线原点和所述射线方向中的至少一项的变化描述为像素图像位置的函数;以及

信号生成器(203),其用于生成所述图像信号,所述信号生成器被布置为在所述图像信号中包括所述图像数据和所述元数据。

15. 一种根据表示场景的图像信号来绘制图像的方法,所述方法包括:

接收所述图像信号,所述图像信号包括图像数据和元数据,

所述图像数据包括多幅图像,每幅图像包括像素,其中,每个像素表示沿着具有从射线原点出发的射线方向的射线的所述场景的图像属性,其中,每个像素与射线的射线原点和射线方向相链接,所述射线原点对于至少一些像素来说处于不同的位置,其中,针对一像素的所述射线原点和所述射线方向指示针对该像素的从场景坐标到图像位置的投影,并且

所述元数据包括多个参数,所述多个参数将针对像素的所述射线原点和所述射线方向中的至少一项的变化描述为像素图像位置的函数;并且

根据所述多幅图像并响应于所述多个参数来绘制图像。

16. 一种生成表示场景的图像信号的方法,所述方法包括:

生成包括多幅图像的图像数据,每幅图像包括像素,其中,每个像素表示沿着具有从射线原点出发的射线方向的射线的所述场景的图像属性,其中,每个像素与射线的射线原点和射线方向相链接,所述射线原点对于至少一些像素来说处于不同的位置,其中,针对一像素的所述射线原点和所述射线方向指示针对该像素的从场景坐标到图像位置的投影;并且

生成元数据,所述元数据包括多个参数,所述多个参数将针对像素的所述射线原点和所述射线方向中的至少一项的变化描述为像素图像位置的函数;并且

生成所述图像信号,信号生成器被布置为在所述图像信号中包括所述图像数据和所述元数据。

17. 一种存储有计算机程序的计算机可读介质,所述计算机程序包括计算机程序代码模块,当所述计算机程序在计算机上运行时,所述计算机程序代码模块适于执行权利要求15或16的所有步骤。

用于生成表示场景的图像信号的装置和方法

技术领域

[0001] 本发明涉及表示场景的图像信号,并且特别地但非排他性地涉及生成表示场景的图像信号以及作为虚拟现实应用的部分的根据该图像信号来绘制图像。

背景技术

[0002] 近年来,随着新服务以及利用和消费视频的方式不断得到开发和引入,图像和视频应用的种类和范围显著增加。

[0003] 例如,一种越来越流行的服务是以观看者能够主动和动态地与系统交互以改变绘制参数的方式提供图像序列。在许多应用中,一个非常吸引人的特征是能够改变观看者的有效观看位置和观看方向,从而例如允许观看者在所呈现的场景中移动和“环顾四周”。

[0004] 这样的特征能够特别允许向用户提供虚拟现实体验。这可以允许用户例如(相对)自由地在虚拟环境中移动并动态地改变他的位置和他正在看的地方。通常,这样的虚拟现实应用基于场景的三维模型,其中,该模型被动态地评价以提供特定的请求视图。这种方法是众所周知的,例如从用于计算机和控制台的游戏应用(例如,第一人称射击游戏)可知。

[0005] 还希望所呈现的图像是三维图像,特别是对于虚拟现实应用。实际上,为了优化观看者的沉浸感,用户通常更喜欢体验作为3D场景的呈现场景。实际上,虚拟现实体验应当优选允许用户选择他/她自己的位置、相机视点以及相对于虚拟世界的时刻。

[0006] 通常,虚拟现实应用固有地受到限制,因为它们基于场景的预定模型,并且通常基于虚拟世界的人工模型。通常希望基于现实世界捕获来提供虚拟现实体验。然而,在许多情况下,这样的方法受限于或往往要求根据现实世界捕获来构建现实世界的虚拟模型。然后通过评价该模型来生成虚拟现实体验。

[0007] 然而,当前的方法往往是次优的,并且常常具有较高的计算或通信资源要求和/或提供次优的用户体验(例如因质量下降或自由受限)。

[0008] 在许多例如虚拟现实应用中,场景可以通过图像表示来表示,例如通过一个或多个表示场景的特定观看姿态的图像来表示。在一些情况下,这样的图像可以提供场景的广角视图并且可以覆盖例如完整的360°视图或覆盖完整的视球。

[0009] 已经提出要提供基于360°视频流的虚拟现实体验,其中,场景的完整的360°视图由服务器针对给定的观看者位置来提供,从而允许客户端生成不同方向的视图。特别地,有前途的虚拟现实(VR)应用之一是全向视频(例如VR360或VR180)。该方法往往会引起高数据速率,因此提供完整的360°视球的视点数量通常被限制为较低的数量。

[0010] 举个具体的示例,虚拟现实眼镜已经进入市场。这些眼镜让观看者可以体验捕获到的360°(全景)视频。这些360°视频通常是使用相机绑定预先捕获的,其中,个体图像被拼接在一起而成为单个球面映射。在一些这样的实施例中,可以生成表示来自给定视点的完整球面视图的图像并将其传输给驱动器,该驱动器被布置为为眼镜生成与用户的当前视图相对应的图像。

[0011] 在许多系统中,可以提供场景的图像表示,其中,图像表示包括场景中的一个或多

个捕获点/视点的图像并且通常还包括该捕获点/视点的深度。在许多这样的系统中,绘制器可以被布置为动态地生成匹配当前的本地观看者姿态的视图。例如,可以动态地确定观看者姿态,并且动态地生成视图以匹配该观看者姿态。

[0012] 在传送图像表示以允许本地绘制器动态地合成针对不同观看姿态的视图图像的系统中,图像表示的特定属性和属性对于优化操作至关重要。希望图像表示能够提供允许绘制图像具有高图像质量并同时维持足够低的数据速率的信息。另外,图像表示的生成和使用都希望低复杂性和低资源使用。

[0013] 已经提出了许多用于通过图像数据来表示场景的不同格式,并且其中的许多格式已经被各种标准机构标准化。一种支持完整360°图像的特定格式被称为全向立体。在这种格式中,为右眼提供一幅图像,并且为左眼提供另一幅图像,其中,每幅图像都包括当围绕眼睛的中心点旋转360°时沿着与观看者的眼睛相对应的视圆的切线的视图。

[0014] 然而,虽然许多常规的图像表示和格式可以在许多应用和服务中提供良好的性能,但是它们往往至少在某些情况下是次优的。

[0015] 因此,改进的用于处理和生成包括场景的图像表示的图像信号的方法将是有利的。特别地,允许改进的操作、提高的灵活性、改善的虚拟现实体验、降低的数据速率、提高效率、促进的分布、降低的复杂性、促进的实施、减少的存储要求、提高的图像质量、改善的绘制、改善的用户体验和/或改善的性能和/或操作的系统和/或方法将是有利的。

发明内容

[0016] 因此,本发明寻求优选单独地或以任何组合方式减轻、减缓或消除上述缺点中的一个或多个缺点。

[0017] 根据本发明的一个方面,提供了一种用于根据表示场景的图像信号来绘制图像的装置,所述装置包括:接收器,其用于接收所述图像信号,所述图像信号包括图像数据和元数据,所述图像数据包括多幅图像,每幅图像包括像素,其中,每个像素表示沿着具有从射线原点出发的射线方向的射线的所述场景的图像属性,所述射线原点对于至少一些像素来说处于不同的位置,并且所述元数据包括多个参数,所述多个参数将针对像素的所述射线原点和所述射线方向中的至少一项的变化描述为像素图像位置的函数;以及绘制器,其用于根据所述多幅图像并响应于所述多个参数来绘制图像。

[0018] 本发明可以提供改善的场景表示并且可以在许多实施例和场景中提供相对于图像信号的数据速率的绘制图像的改进的图像质量。在许多实施例中,能够提供更高效的场景表示,例如允许以降低的数据速率实现给定的质量。该方法可以提供更灵活和更高效的用于绘制场景图像的方法,并且可以允许针对例如场景属性的改善的调整。

[0019] 在许多实施例中,该方法可以采用适合用于灵活、高效和高性能的虚拟现实(VR)应用的场景的图像表示。在许多实施例中,它可以允许或实现图像质量与数据速率之间的折衷得到显著改善的VR应用。在许多实施例中,它可以允许提高的感知图像质量和/或降低的数据速率。

[0020] 该方法可以适用于例如支持在接收端针对移动和头部旋转进行调整的广播视频服务。

[0021] 多幅图像可以特别是光强度图像、深度映射图和/或透明度映射图。图像属性可以

是深度属性、透明度属性或光强度属性(例如,颜色通道值)。

[0022] 在许多实施例中,每个像素可以具有不同的原点和方向,这可以经由查找表或经由参数描述的(线性、余弦或正弦等)函数来提供。

[0023] 图像数据包括多幅图像,即,一幅或多幅图像/至少一幅图像。元数据包括多个参数,即,两个或更多个参数/至少两个参数。

[0024] 根据本发明的任选特征,所述多幅图像中的至少一幅图像的水平像素行的射线位置被映射到所述场景的水平平面中的曲线,所述多个参数描述所述曲线的属性。

[0025] 这在许多实施例中可以提供特别高效的表示。它可以降低复杂性并且在许多示例中可以减少需要传送的参数信息的量。它可以相应地减少为此要求的元数据的量,从而减少开销。

[0026] 根据本发明的任选特征,所述曲线是椭圆,并且所述多个参数描述所述椭圆的属性。

[0027] 这在许多实施例中可以是特别高效的方法,并且可以在复杂性、数据速率、对特定条件的调整性、自由度和绘制图像的潜在图像质量之间提供特别合意的折衷。

[0028] 属性可以具体是大小和偏心率的指示。

[0029] 根据本发明的任选特征,所述曲线是闭合曲线。

[0030] 这在许多实施例中可以是特别有利的。

[0031] 根据本发明的任选特征,所述闭合曲线是卵形,并且所述多个参数描述所述卵形的属性。

[0032] 这在许多实施例中可以是特别高效的方法,并且可以在复杂性、数据速率、对特定条件的调整性、自由度和绘制图像的潜在图像质量之间提供特别合意的折衷。

[0033] 根据本发明的任选特征,所述多个参数由被链接到所述多幅图像中的图像并具有比所述图像的分辨率更低的分辨率的映射图来提供,所述映射图具有指示针对与所述映射图中的像素值的位置相对应的所述图像中的位置的射线原点和射线方向中的至少一项的像素值;并且其中,所述绘制器被布置为通过从所述映射图的所述像素值进行内插来确定针对所述图像中的一些位置的射线原点和射线方向中的至少一项。

[0034] 这在许多实施例中可以是特别高效的方法,并且可以在复杂性、数据速率、对特定条件的调整性、自由度和绘制图像的潜在图像质量之间提供特别合意的折衷。

[0035] 该方法可以特别允许针对射线位置和/或方向的高度灵活性和调整性,同时维持低程度的传送该信息所要求的元数据的开销。

[0036] 根据本发明的任选特征,从图像位置到所述射线原点和所述射线方向中的至少一项的映射是连续函数。

[0037] 这在许多实施例中可以是特别有利的。

[0038] 根据本发明的任选特征,所述图像数据包括多幅图像,并且所述多个参数描述用于将像素图像位置映射到针对所述多幅图像中的至少两幅图像的所述射线原点和所述射线方向中的一项的不同函数。

[0039] 这可以允许在调整射线原点/方向方面具有更高的自由度,并且可以相应地允许提高的图像质量。图像数据中包括的多幅图像可以是多幅图像,并且多个参数可以描述针对这些图像中的至少两幅图像的不同函数。

- [0040] 在一些实施例中,多个参数描述针对至少两幅图像的射线原点之间的不同偏移。
- [0041] 根据本发明的任选特征,所述多幅图像中的至少两幅图像表示部分视场,所述部分视场对于所述至少两幅图像来说是不同的。
- [0042] 这在许多实施例中可以是特别有利的。
- [0043] 根据本发明的任选特征,所述射线原点和所述射线方向中的至少一项的所述变化是所述射线方向(但可能不是所述射线原点)的变化。
- [0044] 根据本发明的任选特征,所述射线原点和所述射线方向中的至少一项的所述变化是所述射线原点(但可能不是所述射线方向)的变化。
- [0045] 在一些实施例中,多个参数将针对像素的射线方向的变化描述为像素图像位置的函数。
- [0046] 在一些实施例中,多个参数将针对像素的射线原点的变化描述为像素图像位置的函数。
- [0047] 根据本发明的任选特征,所述多幅图像包括第一光强度图像和针对所述第一光强度图像的深度值图像,所述第一深度值图像包括针对所述第一光强度图像的像素的深度值,针对所述第一光强度图像的第一像素的深度值指示沿着所述射线方向从针对所述第一像素的射线原点到由所述第一像素表示的对象的距离;并且所述绘制器被布置为响应于所述第一光强度图像和所述第一深度值图像而绘制图像。
- [0048] 根据本发明的任选特征,所述绘制器被布置为:响应于所述深度值图像而确定针对所述第一光强度图像的图像对象的场景位置,并且响应于所述场景位置而绘制图像。
- [0049] 根据本发明的一个方面,提供了一种用于生成表示场景的图像信号的装置,所述装置包括:第一生成器,其用于生成包括多幅图像的图像数据,每幅图像包括像素,其中,每个像素表示沿着具有从射线原点出发的射线方向的射线的所述场景的图像属性,所述射线原点对于至少一些像素来说处于不同的位置;以及第二生成器,其用于生成元数据,所述元数据包括多个参数,所述多个参数将针对像素的所述射线原点和所述射线方向中的至少一项的变化描述为像素图像位置的函数;以及信号生成器,其用于生成所述图像信号,所述信号生成器被布置为在所述图像信号中包括所述图像数据和所述元数据。
- [0050] 图像信号还可以包括针对多幅图像中的至少第一图像的至少第一深度映射图,第一深度映射图包括针对第一图像的像素的深度值,针对第一图像的第一像素的深度值指示沿着射线方向从针对第一像素的射线原点到由第一像素表示的对象的距离。
- [0051] 多幅图像可以包括第一光强度图像和针对第一光强度图像的深度值图像,第一深度值图像包括针对第一光强度图像的像素的深度值,针对第一光强度图像的第一像素的深度值指示沿着射线方向从针对第一像素的射线原点到由第一像素表示的对象的距离。
- [0052] 根据本发明的一个方面,提供了一种根据表示场景的图像信号来绘制图像的方法,所述方法包括:接收所述图像信号,所述图像信号包括图像数据和元数据,所述图像数据包括多幅图像,每幅图像包括像素,其中,每个像素表示沿着具有从射线原点出发的射线方向的射线的所述场景的图像属性,所述射线原点对于至少一些像素来说处于不同的位置,并且所述元数据包括多个参数,所述多个参数将针对像素的所述射线原点和所述射线方向中的至少一项的变化描述为像素图像位置的函数;并且根据所述多幅图像并响应于所述多个参数来绘制图像。

[0053] 根据本发明的一个方面,提供了一种生成表示场景的图像信号的方法,所述方法包括:生成包括多幅图像的图像数据,每幅图像包括像素,其中,每个像素表示沿着具有从射线原点出发的射线方向的射线的所述场景的图像属性,所述射线原点对于至少一些像素来说处于不同的位置;并且生成元数据,所述元数据包括多个参数,所述多个参数将针对像素的所述射线原点和所述射线方向中的至少一项的变化描述为像素图像位置的函数;并且生成所述图像信号,信号生成器被布置为在所述图像信号中包括所述图像数据和所述元数据。

[0054] 根据本发明的一个方面,提供了一种图像信号,包括:图像数据,其包括多幅图像,每幅图像包括像素,其中,每个像素表示沿着具有从射线原点出发的射线方向的射线的所述场景的图像属性,所述射线原点对于至少一些像素来说处于不同的位置;以及元数据,其包括多个参数,所述多个参数将针对像素的所述射线原点和所述射线方向中的至少一项的变化描述为像素图像位置的函数。

[0055] 参考下文描述的(一个或多个)实施例,本发明的这些和其他方面、特征和优点将是显而易见的并且得到阐明。

附图说明

[0056] 将仅通过示例的方式并参考附图来描述本发明的实施例,在附图中:

[0057] 图1图示了用于提供虚拟现实体验的布置的示例;

[0058] 图2图示了根据本发明的一些实施例的装置的元件的示例;

[0059] 图3图示了根据本发明的一些实施例的装置的元件的示例;并且

[0060] 图4图示了场景的全向立体图像表示的示例;

[0061] 图5图示了场景的全向立体图像表示的示例;

[0062] 图6图示了具有深度映射图的全向立体图像的示例;

[0063] 图7图示了根据本发明的一些实施例的用于图像表示的射线原点和方向的示例;

[0064] 图8图示了根据本发明的一些实施例的用于图像表示的射线原点和方向的示例;

[0065] 图9图示了根据本发明的一些实施例的用于图像表示的射线原点和方向的示例;

并且

[0066] 图10图示了根据本发明的一些实施例的用于图像表示的射线原点和方向的示例。

具体实施方式

[0067] 允许用户在虚拟世界中四处走动的虚拟体验正变得越来越流行,并且正在开发服务以满足这样的需求。然而,提供高效的虚拟现实服务非常具有挑战性,特别是如果体验是基于对现实世界环境的捕获而不是基于完全虚拟生成的人工世界。

[0068] 在许多虚拟现实应用中,观看者姿态输入被确定为反映虚拟观看者在该场景中的姿态。虚拟现实装置/系统/应用然后为与观看者姿态相对应的观看者生成与场景的视图和视口相对应的一幅或多幅图像。

[0069] 通常,虚拟现实应用生成针对左眼和右眼的单独视图图像形式的三维输出。然后通过合适的方式(例如通常是VR头戴式设备的独立的左眼显示器和右眼显示器)将这些输出呈现给用户。在其他实施例中,图像可以例如被呈现在自动立体显示器上(在这种情

况下,可以针对观看者姿态生成大量视图图像),或者实际上在一些实施例中可以(例如使用常规的二维显示器)仅生成单个二维图像。

[0070] 在不同应用中可以以不同方式确定观看者姿态输入。在许多实施例中,可以直接跟踪用户的身体移动。例如,调查用户区的相机可以检测和跟踪用户的头部(甚至眼睛)。在许多实施例中,用户可以佩戴能够通过外部方式和/或内部方式跟踪的VR头戴式设备。例如,头戴式设备可以包括加速度计和陀螺仪,它们提供关于头戴式设备以及因此头部的移动和旋转的信息。在一些示例中,VR头戴式设备可以传输信号或包括(例如视觉)标识符,这使得外部传感器能够确定VR头戴式设备的移动。

[0071] 在一些系统中,观看者姿态可以通过手动方式来提供,例如通过用户手动控制操纵杆或类似的手动输入来提供。例如,用户可以通过用一只手控制第一模拟操纵杆来手动移动虚拟观看者在场景中四处走动,并且通过用另一只手手动移动第二模拟操纵杆来手动控制虚拟观看者正在看的方向。

[0072] 在一些应用中,可以使用手动方法与自动方法的组合来生成输入的观看者姿态。例如,头戴式设备可以跟踪头部的取向并且观看者在场景中的移动/位置可以通过用户使用操纵杆来控制。

[0073] 图像的生成基于虚拟世界/环境/场景的合适表示。在一些应用中,可以为场景提供完整的三维模型,并且能够通过评价该模型来确定从特定观看者姿态看到的场景的视图。

[0074] 在许多实际系统中,场景可以由包括图像数据的图像表示来表示。图像数据通常可以包括与一个或多个捕获或锚定姿态相关联的一幅或多幅图像,并且特别地可以包括针对一个或多个视口的图像,其中,每个视口对应于特定姿态。可以使用包括一幅或多幅图像的图像表示,其中,每幅图像表示针对给定观看姿态的给定视口的视图。针对其提供图像数据的此类观看姿态或位置常常被称为锚定姿态或位置或捕获姿态或位置(因为图像数据通常可以对应于(将)由被定位在场景中的具有与捕获姿态相对应的位置和取向的相机所捕获的图像)。

[0075] 许多典型的VR应用可以在这样的图像表示的基础上继续提供与针对当前观看者姿态的场景的视口相对应的视图图像,其中,图像被动态地更新以反映观看者姿态的变化,并且图像是基于表示(可能)虚拟场景/环境/世界的图像数据来生成的。该应用可以通过执行本领域技术人员已知的视图合成和视图移位算法来做到这一点。

[0076] 在该领域中,术语放置和姿态被用作针对位置和/或方向/取向的通用术语。例如对象、相机、头部或视图的位置和方向/取向的组合可以被称为姿态或放置。因此,放置或姿态指示可以包括六个值/分量/自由度,其中,每个值/分量通常描述对应对象的位置/定位或取向/方向的独立属性。当然,在许多情况下,可能会以较少的分量考虑或表示放置或姿态,例如在一个或多个分量被认为是固定的或不相关的情况下(例如在所有对象都被认为处于相同的高度并且具有水平取向的情况下,四个分量可以提供对象姿态的完整表示)。在下文中,术语姿态用于指代可以由一到六个值(对应于最大可能的自由度)表示的位置和/或取向。

[0077] 许多VR应用基于具有最大自由度的姿态,即,位置和取向中的每个的三个自由度会得到总共六个自由度。姿态因此可以由表示六个自由度的六个值的集合或向量来表示,

并且因此姿态向量可以提供三维位置和/或三维方向指示。然而,应当理解,在其他实施例中,姿态可以由更少的值来表示。

[0078] 姿态可以是取向和位置中的至少一个。姿态值可以指示取向值和位置值中的至少一个。

[0079] 基于为观看者提供最大自由度的系统或实体通常被称为具有6个自由度(6DoF)。许多系统和实体仅提供取向或位置,并且这些系统和实体通常被称为具有3个自由度(3DoF)。

[0080] 在一些系统中,可以通过例如不使用(甚至根本无法访问)任何远程VR数据或处理的独立设备在本地向观看者提供VR应用。例如,诸如游戏控制台之类的设备可以包括用于存储场景数据的存储装置、用于接收/生成观看者姿态的输入部,以及用于根据场景数据来生成对应图像的处理部。

[0081] 在其他系统中,可以在远离观看者的地方实施和执行VR应用。例如,用户本地的设备可以检测/接收移动/姿态数据,该移动/姿态数据被传输到远程设备,该远程设备处理该数据以生成观看者姿态。远程设备然后可以基于描述场景的场景数据来生成针对观看者姿态的合适的视图图像。然后将视图图像传输到观看者本地的设备,在该设备处呈现视图图像。例如,远程设备可以直接生成由本地设备直接呈现的视频流(通常是立体/3D视频流)。因此,在这样的示例中,除了传输移动数据和呈现接收到的视频数据之外,本地设备可以不执行任何VR处理。

[0082] 在许多系统中,功能可以被分布在本地设备和远程设备上。例如,本地设备可以处理接收到的输入和传感器数据以生成被连续传输到远程VR设备的观看者姿态。远程VR设备然后可以生成对应的视图图像并且将这些视图图像传输到本地设备以供呈现。在其他系统中,远程VR设备可能不直接生成视图图像,而是可以选择相关的场景数据并将其传输到本地设备,然后该本地设备可以生成所呈现的视图图像。例如,远程VR设备可以识别最接近的捕获点并提取对应的场景数据(例如来自该捕获点的球面图像和深度数据)并将其传输到本地设备。本地设备然后可以处理接收到的场景数据以生成针对特定的当前观看姿态的图像。观看姿态通常将对应于头部姿态,并且对观看姿态的引用通常可以被等效地认为对应于对头部姿态的引用。

[0083] 在许多应用中,尤其是针对广播服务,源可以以独立于观看者姿态的场景的图像(包括视频)表示的形式传输场景数据。例如,针对单个捕获位置的单个视球的图像表示可以被传输到多个客户端。个体客户端然后可以在本地合成对应于当前观看者姿态的视图图像。

[0084] 引起特别关注的应用是以下应用:其支持有限量的移动,使得所呈现的视图被更新以跟随与基本上静态的观看者相对应的小的移动和旋转,该观看者仅进行小的头部移动和头部旋转。例如,坐着的观看者能够转动他的头部并稍微移动他的头部,其中,所呈现的视图/图像被调整以跟随这些姿态变化。这样的方法可以提供高度具有沉浸感的例如视频体验。例如,观看体育赛事的观看者可能会觉得他出现在竞技场中的特定地点。

[0085] 这样的有限自由度应用具有以下优点:它能够提供改善体验,同时又不要求从许多不同位置看到的场景的准确表示,从而显著降低了捕获要求。类似地,需要提供给绘制器的数据量能够大大减少。实际上,在许多场景中,只需要向能够从中生成所希望的视图的本

地绘制器提供针对单个视点的图像并且通常还提供深度数据。

[0086] 该方法可能特别适合用于需要通过带宽受限的通信信道将数据从源传送到目的地的应用,例如,广播或客户端服务器应用。

[0087] 图1图示了VR系统的这样一个示例,其中,远程VR客户端设备101与VR服务器103例如经由网络105(例如,互联网)进行联络。服务器103可以被布置为同时支持潜在大量的客户端设备101。

[0088] VR服务器103可以例如通过传输包括图像数据形式的图像表示的图像信号来支持广播体验,客户端设备能够使用该图像信号来在本地合成与适当姿态相对应的视图图像。

[0089] 图2图示了VR服务器103的示例性实施方式的示例元件。

[0090] 该装置包括第一生成器201,该第一生成器201被布置为以一幅或多幅图像的形式生成场景的图像表示。图像可以例如是基于对场景模型的评价或者基于(例如通过潜在的大量相机)现实世界的捕获来生成的。

[0091] 该装置还包括输出处理器203,该输出处理器203生成包括图像表示的图像信号,因此图像信号具体包括一幅或多幅图像的图像数据。在许多实施例中,输出处理器207可以被布置为对图像进行编码并且将它们包括在合适的数据流(例如,根据合适的标准生成的数据流)中。

[0092] 输出处理器207还可以被布置为向远程客户端/设备发送或广播图像信号,并且特别地,可以将图像信号传送到客户端设备101。

[0093] 图3图示了根据本发明的一些实施例的用于绘制图像的装置的一些元件的示例。该装置将在图1的系统的背景中进行描述,其中,装置具体是客户端设备101。

[0094] 客户端设备101包括数据接收器301,该数据接收器301被布置为接收来自服务器103的图像信号。应当理解,可以使用任何合适的用于通信的方法和格式,而不会偏离本发明。数据接收器可以相应地接收针对多幅图像(即,一幅或多幅图像)的图像数据。

[0095] 数据接收器301被耦合到绘制器303,该绘制器303被布置为生成针对不同视口/观看者姿态的视图图像。

[0096] 客户端设备101还包括观看姿态确定器305,该观看姿态确定器305被布置为动态地确定当前观看者姿态。特别地,观看姿态确定器305可以接收来自头戴式设备的反映头戴式设备的移动的数据。观看姿态确定器305可以被布置为基于接收到的数据来确定观看姿态。在一些实施例中,观看姿态确定器305可以接收例如传感器信息(例如,加速度计和陀螺仪数据)并由此确定观看姿态。在其他实施例中,头戴式设备可以直接提供观看姿态数据。

[0097] 观看姿态被馈送到绘制器303,该绘制器303继续生成与从处于当前观看者姿态的观看者的两只眼睛看到的场景的视图相对应的视图图像。使用任何合适的图像生成和合成算法根据接收到的图像数据来生成视图图像。具体算法将取决于具体的图像表示以及各个实施例的偏好和要求。

[0098] 应当理解,虽然该方法可以特别地用于动态地生成与检测到的观看者运动相对应的视图图像,但是该方法也可以以其他方式进行使用。例如,可以将预定的姿态集合本地存储在绘制器303中,并且可以按顺序向观看者提供对应的视图,从而提供“脚本式”体验。

[0099] 绘制器303被布置为基于接收到的图像表示来生成针对当前观看姿态的视图图像。特别地,可以为立体显示器(例如,头戴式设备)生成右眼图像和左眼图像,或者可以为

自动立体显示器的视图生成多幅视图图像。应当理解,已知许多不同的算法和技术能够用于根据所提供的场景图像来生成视图图像,并且可以根据特定实施例来使用任何合适的算法。

[0100] 现有的3D图像格式被称为全向立体(ODS)。对于ODS,为观看者的左眼和右眼提供图像。然而,与表示来自单个左眼位置的视口的左眼图像和表示来自单个右眼位置的视口的右眼图像不同,像素仅表示眼睛正前方的方向上的场景,并且几乎没有视场。当眼睛围绕眼睛位置的中心点旋转(对应于观看者围绕该点旋转他的头部)时,图像由表示正前方视图的像素来形成。因此,围绕眼睛之间的中心点的旋转固有地形成一个圆,眼睛留在这个圆上。生成ODS的图像以反映针对这个圆上的不同位置的视图。特别地,生成针对圆上的给定位置的像素值以反映在该点处沿着圆的切线方向的视图。

[0101] 因此,对于ODS,创建针对左眼图像和右眼图像的射线,使得这些射线的原点位于其直径通常等于瞳孔距离(例如~6.3厘米)的圆上。对于ODS,针对与视圆的切线相对应的相反方向并以围绕视圆的规则角距离来捕获窄角度图像部分(参见图4)。

[0102] 因此,对于ODS,为左眼生成图像,其中,每个像素列对应于单位圆上的一个位置,并且反映在该位置处在与ODS视圆的切线方向上的射线。ODS视圆上的位置对于每一列都是不同的,并且通常对ODS视圆定义相对大量的等距位置,从而覆盖整个360°视场,其中,每一列对应于一个位置。因此,单个ODS图像捕获完整的360°视场,其中,每一列对应于ODS视圆上的不同位置并且对应于不同射线方向。

[0103] ODS包括右眼图像和左眼图像。如图6所示,对于这些图像中的给定列,左眼图像和右眼图像将反映ODS视圆上的相对位置的射线。因此,ODS图像格式提供360°视图以及仅基于两幅图像的立体信息。

[0104] 对于给定的取向(视角),可以通过组合针对在与给定取向的视口内的观看方向相匹配的方向的窄角度图像部分来生成图像。因此,通过组合与在不同方向上的捕获相对应的窄角度图像部分(但不同的窄角度图像部分来自圆上的不同位置)来形成给定的视图图像。因此,视图图像包括来自视圆上的不同位置的捕获而不是仅来自单个视点的捕获。然而,如果ODS表示的视圆足够小(相对于场景的内容),则能够将其影响降低到可接受的水平。另外,由于沿着给定方向的捕获能够重复用于多个不同的观看取向,因此使得所要求的图像数据量显著减少。针对观看者的两只眼睛的视图图像通常是通过在针对适当切线的相反方向上的捕获来生成的。

[0105] 图5图示了能够由ODS支持的理想的头部旋转的示例。在该示例中,头部旋转而使得两只眼睛沿着一个直径等于瞳孔距离的圆移动。假设这对应于ODS视圆的宽度,则能够通过选择与不同视点取向相对应的适当的窄角度图像部分来简单地确定针对不同取向的视图图像。

[0106] 然而,对于标准ODS,观察者将感知立体视觉而不是运动视差。即使观看者的运动很小(大约几厘米),没有运动视差也往往会提供不愉快的体验。例如,如果观察者移动而使得眼睛不再正好落在ODS视圆上,那么基于简单地选择和组合适当的窄角度图像部分来生成视图图像将导致所生成的视图图像与用户的眼睛仍然停留在视圆上的情况一样,因此不会表现出用户移动他的头部时应产生的视差,这将导致感知无法相对于现实世界移动。

[0107] 为了解决这个问题并允许基于ODS数据来生成运动视差,ODS格式可以被扩展为包

括深度信息。可以为每个窄角度图像部分添加一个窄角度深度映射图部分。图6图示了具有相关联的深度映射图的ODS图像的示例。该深度信息可以用于执行视点移位,使得所生成的图像对应于视圆外部(或内部)的新位置(例如,可以使用已知图像和基于深度的视点移位算法来处理每幅视图图像或窄角度图像部分)。例如,可以为每只眼睛创建3D网格,并且能够使用基于针对左眼和右眼的网格和纹理进行的ODS数据绘制来引入运动视差。

[0108] 然而,图像表示是否基于例如针对不同的捕获姿态或ODS数据的多幅图像,生成针对与提供图像数据的锚定姿态不同的姿态的视图图像往往会引入伪影和错误,从而导致潜在的图像质量下降。

[0109] 在图1-3的系统的的方法中,使用了不同的方法,并且特别是使用了不同的图像表示。该方法不是基于眼睛的旋转,并且确实打破了基本常规假设,即,人类立体感知和相貌与图像表示之间必须存在密切的相关性。替代地,该方法提供了非常灵活且调整性强的方法,该方法可以在许多情况下提供显著更好的性能,并且特别是显著改善的图像质量与数据速率的折衷。

[0110] 该方法基于图像表示,其中,图像包括像素,每个像素表示沿着具有从射线原点出发的射线方向的射线的场景的图像属性。因此,每个像素都与作为射线/直线的原点的位置相链接。每个像素还与作为从原点出发的射线/直线方向的方向相链接。因此,每个像素与由位置/原点和来自该位置/原点的方向所定义的射线/直线相链接。像素值由针对场景的在像素的射线与场景对象(包括背景)的第一个交叉点处的适当属性给出。因此,像素值表示场景的在起源于射线原点位置并具有与像素相关联的射线方向的射线/直线的远端处的属性。

[0111] 在许多情况下,图像属性可以是光强度属性,而像素值可以是针对场景对象(或背景)的在从原点位置出发的射线方向上的光强度。特别地,像素值可以是针对场景的从射线原点发出的射线方向上的光强度值。在这种情况下,像素值可以是在原点位置处从由射线方向指示的方向接收的射线的光强度的度量。光强度可以是颜色通道中的光强度,或者是例如在有限的带宽中的光强度。

[0112] 如果图像属性例如是深度属性,则像素值可以指示在射线方向上从射线原点到第一图像对象的距离。对于图像属性为透明度的示例,像素值可以反映图像对象在从射线原点出发的射线方向的方向上的透明度。

[0113] 在许多实施例中,可以提供光强度图像与深度值图像(通常被称为深度映射图)的匹配对。在这种情况下,对于给定的像素(在光强度图像和深度映射图中的相同位置处),光强度图像值可以指示针对场景对象(或背景)的在从射线原点出发的射线方向上的光强度,并且深度值可以指示在射线方向上从射线原点到场景/图像对象的距离。因此,对于光强度图像中的每个像素,深度映射图可以包括指示到由像素表示的对象的距离的深度值。可以使用适合用于指示该距离的任何参数/量度,例如可以以距离、视差、Z值、1/Z值等来给出深度值。

[0114] 另外,与所有像素表示从相同位置看到的视图的常规图像相比,所描述的方法的图像表示包括(至少一些)表示不同原点/位置的像素。通常,(至少一些)像素也可以表示不同的方向。

[0115] 该方法基于允许灵活且调整性地选择射线原点和/或射线方向。特别地,在许多实

施例中,可以(例如根据场景特性)针对给定条件对针对给定图像的像素的射线方向和射线原点进行调整和至少部分优化。这种调整和优化可以在源处(即,特别是在VR服务器103处)执行,并且元数据被传送到接收器(特别是客户端设备101)并且在呈现过程中使用。

[0116] 例如,对于给定的图像,可以确定位置集合,其中,图像的像素值表示具有处于这些位置的原点并例如具有相对于曲线的预定方向(例如垂直于该位置处的曲线的梯度)的射线。图7中提供了这样的示例,其中,箭头指示从曲线上的给定射线原点出发的射线方向。可以生成图像,其中,每一列提供一个箭头的像素值,即,每一列表示针对给定水平射线方向和射线原点的垂直视图。因此,图像表示整个曲线的视图信息,其中,每一列表示沿着曲线在不同方向上和从不同原点看到的非常窄的视图。

[0117] 作为另一示例,对于给定的图像,可以确定方向集合,其中,图像的像素值表示具有预定原点(例如是一条线上的等距点)的射线。图8提供了这样的示例,其中,箭头指示从给定的预定射线原点出发的射线方向。可以生成图像,其中,每一列提供一个箭头的像素值,即,每一列表示针对给定水平射线方向和射线原点的垂直视图。因此,图像表示来自不同的射线方向和原点的集合的场景视图信息,其中,每一列表示沿着线在不同方向上和从不同原点看到的非常窄的视图。

[0118] 在一些情况下,调整既可以包括射线方向又可以包括射线原点,例如可以根据场景而灵活地选择射线方向和射线原点这两者。

[0119] 在许多实施例中,该方法可以允许提高图像质量,并且可以例如允许射线聚焦在场景对象或特别重要的区域(例如,面部)上或者聚焦在更难以捕获的对象或区域(例如,高度详细的对象)上。然而,这种改善的场景表示能够在仍然保持低数据速率的同时实现,并且特别是通过仍然表示例如单一图像(或例如少量的单一图像或立体图像,例如,2-6幅单一图像或1-3幅立体图像)中的信息来实现。因此,能够生成非常高效的图像信号,从而允许针对给定数据速率得到高图像质量的绘制图像。

[0120] 为了支持灵活的动态调整和优化,VR服务器101还包括元数据生成器205,该元数据生成器205被布置为生成元数据,该元数据包括两个或更多参数(自由度),这两个或更多参数(自由度)将针对像素的射线原点和/或射线方向的变化描述为像素图像位置的函数。参数因此可以提供对图像表示的(一幅或多幅)图像的像素与对应的射线原点和射线方向之间的映射或链接的描述。

[0121] 元数据生成器205被耦合到输出处理器203,该输出处理器203被布置为包括描述图像信号中的参数的元数据。应当理解,可以使用用于通过元数据表示参数以及用于将该元数据编码和包含在图像信号中的任何合适的方法,而不会偏离本发明。

[0122] 客户端设备101还包括元数据处理器307,该元数据处理器307被馈送有来自数据接收器301的图像信号的元数据。元数据处理器307被布置为从元数据中提取参数并将这些参数馈送到绘制器。绘制器303使用这些参数来确定针对接收到的(一幅或多幅)图像的给定像素的射线原点和射线方向。因此,绘制器可以被布置为响应于图像和参数来执行绘制。

[0123] 绘制器303可以被布置为:确定像素位置与射线原点和射线方向中的至少一项之间的映射,并且基于该映射来确定针对接收到的多幅图像的像素的射线原点和射线方向。绘制器303然后可以基于接收到的(一幅或多幅)图像和针对接收到的(一幅或多幅)图像的像素确定的所确定的射线方向和射线原点来合成一幅或多幅输出图像。

[0124] 应当理解,可以使用任何合适的绘制方法和算法。在一些实施例中,绘制器303可以通过为对应于图像的视口选择具有合适的射线原点和方向的像素来简单地生成图像。例如,可以通过内插/外推来填充任何间隙或孔洞。

[0125] 在许多实施例中,绘制可以基于光强度图像和相关联的深度这两者。特别地,如前所述,图像数据信号可以包括光强度图像和深度值图像这两者。通常,图像可以包括光强度图像以及相关联的深度值图像,针对光强度图像中的每个像素,深度值图像提供像素表示的到场景对象的深度的指示。因此,对于光强度像素和对应的深度值像素,射线原点和射线方向可以是相同的,并且特别地,光强度值可以指示在射线方向上从射线原点看到的对象的光强度,而深度值可以指示沿着射线方向从射线位置到对象的距离。

[0126] 在这样的场景中,绘制器303可以基于光强度值和深度值这两者来绘制图像。这可以例如利用从视图移位已知的技术(例如基于深度和射线原点和方向),可以使用基本几何来计算与生成图像的视口相对应的图像中的给定像素的位置。结果得到的图像然后可以(例如使用选择组合来)组合具有重叠位置的值,其中,确定的值最接近生成图像的观看者姿态。类似地,可以例如通过内插来填充任何间隙。

[0127] 在许多实施例中,绘制器303可以被布置为使用深度值图像(深度映射图)来将像素投影到世界/场景位置。特别地,可以确定这样的像素:其具有等于射线原点加上射线方向(以单位向量给出)乘以由针对像素的深度值指示的距离的世界/场景位置。可以对所有像素完成这项操作,从而构建3D模型。实际上,该方法可以用于生成网格,例如,以像素为顶点的网格。然后,针对对应像素的光强度值可以提供针对模型的视觉表示。

[0128] 当生成针对特定观看姿态的图像时,可以由绘制器303来评价这样生成的模型。这样的评价/处理可以包括视图移位,并且该方法可以允许针对一定范围的观看姿态有效地生成图像。该方法可以专门支持能够绘制视差以匹配观看者运动的应用和服务。

[0129] 作为特定示例,绘制器303可以通过以下操作来合成针对特定观看姿态的图像:计算浮点图像坐标图,利用作为参考系的输入相机将图像坐标解投影(unprojection)到世界坐标,应用单一仿射变换 $x \rightarrow Rx+t$ 来使虚拟相机成为参考系,将世界坐标投影到虚拟图像上,并且使用三角形光栅化根据结果得到的映射图来扭曲图像。解投影操作计算射线的原点和方向,并且最后计算射线末端的场景点。投影操作是逆操作。在给定投影类型的情况下,投影器找到对应于场景点的图像位置。关于更多的细节,请参阅参考视图合成器(RVS)手册(ISO/IEC JTC1/SC29/WG11 MPEG/N18068,2018年10月,中国澳门特别行政区)(例如可从<https://mpeg.chiariglione.org/standards/exploration/immersive-video/reference-view-synthesizer-rvs-manual>中获得)。

[0130] RVS软件是支持从/到常规的等距柱状投影和透视投影的合成的软件的示例。在许多实施例中,可以对射线方向或射线原点施加一些限定或限制。例如,在许多实施例中,射线方向可以与射线原点具有预定关系,使得射线原点的信息固有地描述射线方向。例如,如前所述,射线可以被垂直地导向定义射线原点的曲线的梯度/切线。

[0131] 图9图示了针对水平图像轴线变化的射线原点和射线方向的特定示例。该示例包括两个发散射线方向区域901、903和一个会聚射线方向区域905。图9还示出了可以如何使用射线来为虚拟/合成视图生成目标射线907。该图图示了如何将具有(由不同的球指定的)不同原点的射线组合成单幅图像的示例。

[0132] 应当注意,图像可以是场景的表示,并且可以表示从场景坐标到图像坐标的投影。射线原点和方向可以反映该投影,即,从场景坐标到图像坐标的投影可以引起每个像素表示/对应于/是特定场景坐标(其可以是三维的)的投影。射线原点和射线方向可以反映/指示这个场景坐标,因此针对像素的射线原点和射线方向指示由该像素表示的场景坐标以及针对该像素的从场景坐标到图像位置的投影。

[0133] 更特别地,连续图像位置可以由 $u=(u, v)$ 来表示,例如,对于行索引 i 和列索引 j ,像素中心在 $(0.5+j, 0.5+i)$ 处。到图像位置/坐标的投影可以是能够被参数化的投影。投影参数集可以用 Θ 来表示。该投影和参数集可以提供关于射线原点和方向的信息(因为射线原点和方向表示场景坐标)。射线原点可以用 $r_0=(x, y, z)$ 来表示,它是场景坐标中的3D向量,而射线方向可以用 $\hat{r}=(d_x, d_y, d_z)$ 来表示,它也可以是场景坐标中的3D向量(并且特别地,它可以是单位向量)。

[0134] 表示投影的射线角度/射线原点映射由函数 $f: (u; \Theta) \rightarrow (r_0, \hat{r})$ 给出,其中,该函数由图像投影给出。

[0135] 特别地,函数 f 可以是连续函数和/或是在图像内不具有任何不连续性的函数。

[0136] 因此,对于每个像素,可以提供反映针对像素的场景坐标与图像坐标之间的投影的射线原点和射线方向。在包括常规的全向视频应用在内的常规方法中,从场景坐标进行的投影基于针对图像的单个视点(因此是单个射线原点)以及预定和固定的射线方向(例如,全向图像可以基于在围绕中心点的视球上的投影)。当前方法的优点是它允许灵活和调整性的变化,这意味着能够使用的投影具有高度灵活性。因此,该方法可以允许显著改善的调整,尤其是质量优化。例如,可以将更多像素分配给被认为特别感兴趣或可能对失真特别敏感的场景区域。

[0137] 因此,每个提供的图像可以表示经调整或优化的投影。这些对于所有图像可以相同也可以不同。每幅图像可以是表示场景或对象的像素的(通常为矩形)阵列。对于每幅图像,投影可以使得阵列/图像内的接近度指示场景中的空间接近度。相邻像素通常可以具有相似的射线原点和方向。特别地,相邻像素在场景中通常具有相似的空间位置(通常除了对象的边缘等(即,存在深度跳跃的地方)以外)。

[0138] 图像中的像素可以共享(广义)投影,这意味着存在将像素(阵列)位置映射到射线角度和原点的参数化。

[0139] 在许多实施例中,每幅图像可以表示从场景坐标到图像坐标的投影,该投影是连续函数。

[0140] 在许多实施例中,每幅图像可以表示从场景坐标到图像坐标的连续投影。

[0141] 在许多实施例中,每幅图像可以表示从场景坐标到图像坐标的投影,该投影不包括任何不连续性。

[0142] 在许多实施例中,每幅图像可以表示从场景坐标到图像坐标的投影,该投影使得每个场景坐标仅投影到一个图像坐标。

[0143] 在许多实施例中,每幅图像可以表示从场景坐标到图像坐标的投影,该投影是单射函数。

[0144] 每幅图像可以表示从场景坐标到图像坐标的投影,在许多实施例中,每幅图像都

是一对一的函数。

[0145] 在许多实施例中,每幅图像可以表示从场景坐标到图像坐标的投影,该投影使得图像中的接近度指示场景中的接近度。

[0146] 在许多实施例中,每幅图像可以表示从场景坐标到图像坐标的投影,并且针对第一像素的射线原点和射线方向指示被投影到第一像素的场景坐标。针对第一像素的射线原点、射线方向和深度值可以指示场景中的三维位置。

[0147] 在许多实施例中,像素位置到射线原点的映射可以是连续映射。例如,函数可以定义从图像中的x位置(水平位置)到场景的水平平面中的二维位置的映射。对于二维位置的分量中的每个分量,该函数可以是连续函数。图7图示了这样的方法的示例。

[0148] 因此,在一些示例中,水平像素行的射线位置可以被映射到场景的水平平面中的曲线。映射可以例如通过针对部件的单独功能来实现,例如,水平场景平面中的位置的两个分量x、y可以被定义为:

$$[0149] \quad x = f_x(x_i)$$

$$[0150] \quad y = f_y(y_i)$$

[0151] 其中, x_i 表示图像中的像素的x位置,并且 $f_x(x_i)$ 和 $f_y(y_i)$ 是合适的平滑函数,这意味着函数域内没有大的波动(导数值)。

[0152] 由于像素位置是离散的,因此函数也可以被认为是离散的。在这种情况下,如果相邻值之间的绝对差与函数的范围相比很小,则可以认为该函数是连续的。更具体地,绝对差应小于具有相同范围的线性函数的绝对差的十分之一。

[0153] 事实上,离散函数能够通过利用内插函数进行变换来扩展为连续函数。相邻值之间的较大差异会导致较大的导数值。

[0154] 在许多实施例中,VR服务器103和客户端设备101都可以知道曲线的形状属性或限制。例如,可以预先确定曲线的一般形状,并且参数可以描述该曲线的一些可变参数,例如,大小、曲率量、特定事件的位置(例如,方向改变)等。这在许多实施例中可以显著减少描述曲线所要求的元数据的量。

[0155] 未闭合的参数曲线的示例是螺旋。螺旋能够从例如4厘米的半径处开始并在12厘米的半径处停止。与ODS格式类似,射线能够被存储在与螺旋位置相切的方向上。与ODS格式相比,螺旋将使用户能够稍微向侧面移动他/她的头部以感知运动视差,从而从螺旋中选择不同的图像部分。

[0156] 在许多实施例中,曲线可以是闭合曲线。在这样的实施例中,最左边像素位置的射线原点可以与最右边像素位置的射线原点相邻。闭合曲线在许多实施例中可以是卵形,或者特别地在水平场景平面中可以是椭圆。已经发现这样的形状在许多实施例中特别有利,因为它们提供了高度的灵活性和调整性,这样可以允许改善的图像表示,同时还允许低复杂性的处理并且只要求很少的参数来描述像素位置与射线原点之间的映射。

[0157] 图10图示了曲线为闭合曲线并且特别地曲线为椭圆的示例。在该示例中,射线原点被确定为围绕椭圆的等距位置,并且射线方向沿着椭圆在射线原点处的切线。该图示出了八条射线,但是应当理解,在大多数实施例中,定义了更多数量的射线。例如,对应于1024个像素的图像的水平分辨率的1024条射线对于许多实施例可能是典型的。

[0158] 椭圆可以由两个(或更多)参数来定义,这些参数描述椭圆的大小和形状。例如,如

果客户端设备101知道射线原点被等距定位在椭圆上并且射线方向沿着椭圆的切线,则唯一要求的参数是例如椭圆的宽度dx和长度dy。因此,能够实现仅具有最小开销的非常高效的通信。因此,值dx和dy可以指示椭圆的形状(例如,偏心率)和大小。

[0159] 应当理解,在许多实施例中,可以使用除了椭圆之外的其他闭合曲线。例如,更一般地,曲线可以是卵形,并且特别地,曲线可以是笛卡尔卵形。

[0160] 在大多数实施例中,曲线是简单的(不是自相交的)、凸的和/或可微的。曲线通常可以是以一条或两条对称轴对称的闭合曲线。

[0161] 在许多实施例中,曲线将是平滑的并且不会具有小于例如1厘米、2厘米或5厘米的半径的曲率。

[0162] 在许多实施例中,合适曲线的其他示例包括卡西尼卵形、莫斯蛋曲线、超椭圆和/或体育场形。

[0163] 在一些实施例中,客户端设备101可以预先知道多条可能的曲线,并且元数据的参数中的一个参数可以指示这些曲线中的一条曲线,而其他参数指示所选择的曲线的属性。

[0164] 例如,曲线在极坐标 (r, θ) 中可以由函数 $f: \theta \rightarrow r$ 来表示,其中,f的傅立叶变换在低次谐波($\cos 2\theta$ 、 $\sin 4\theta$ 等)中具有最大能量。

[0165] 例如,椭圆的公式为 $r(\theta) = 1/(1 - e \cos \theta)$,其中,e是偏心率,并且1是半正焦弦。对于 $e=0$,这将简化为半径为1的圆。在这样的情况下,e和1的参数可以由元数据来提供。

[0166] 在一些实施例中,用于提供多个参数的特别有利的方法是使用被链接到图像中的一幅或多幅图像的映射图。映射图的分辨率可以低于图像的分辨率。特别地,可以提供参数映射图,其中,为图像的像素的子集提供参数值。特别地,参数值可以是射线原点的指示,例如,水平面上的二维坐标。替代地或额外地,参数值可以是射线方向(例如,作为角度指示)。

[0167] 对于在映射图中提供了参数值的图像中的像素,绘制器403可以直接将对应的射线原点和/或射线方向确定为存储在参数映射图中的值。然而,对于参数映射图不包括其值的图像像素,绘制器403可以在映射图中存在的值之间进行内插。例如,仅考虑水平像素位置,如果参数映射图包括像素10和像素20的射线原点,则能够通过内插在像素10和像素20的射线原点之间进行内插来找到像素11到19的射线原点。应当理解,能够使用不同的内插,但是已经发现,在许多实施例中,简单的线性内插将足以生成得到高图像质量的射线原点。

[0168] 在许多实施例中,该方法可以特别灵活高效地提供能够高效地描述任何希望的映射的参数数据。该方法不要求在客户端设备101处对映射的任何预定知识。该方法还特别适合用于提供多方向的映射图信息。

[0169] 实际上,虽然前面的描述集中于映射仅依赖于x(水平)像素位置的实施例并且因此仅依赖于将相同的映射应用于所有像素行的实施例,但是在一些实施例中,映射也可以依赖于y(垂直)像素位置。例如,与图像中心的行相比,对于朝向图像的上边缘或下边缘的像素行,可以使用不同的映射或闭合曲线。

[0170] 这样的灵活且可变的映射可以使用较低分辨率的二维映射图来有效地传送,其中,射线原点/方向被存储为值。在这样的情况下,可以在水平方向和垂直方向上都应用内插。

[0171] 因此,在一些实施例中,可以将参数提供为作为值阵列的下采样映射图(可能该阵列可以是一维的)。然后可以使用内插来为中间像素生成射线原点/方向。

[0172] 例如,对于360°视频(全向图像),内插可以沿着水平图像边缘循环进行。(因此当图像行被循环时,除了相同的循环平移之外,内插结果将是相同的。)

[0173] 内插通常被简单地认为是在给定的值网格的情况下确定中间值的任务。然而,当内插函数是连续的时,对离散映射图进行内插的行为会生成一个连续的表面。当可微时能够是可微等。

[0174] 可能的内插的示例可以包括:

[0175] ●作为非连续函数的零阶(矩形)内插。

[0176] ●连续但不可微的一阶(双线性)内插。

[0177] ●可微但对于二阶不可微的二阶(双三次)。

[0178] ●其他函数,例如,平滑的Lanczos。

[0179] 在许多实施例中,图像数据可以包括多幅图像。在一些情况下,对于不同的图像,可以使用从图像位置到射线原点/方向的相同映射。例如,在一些实施例中,这对于利用被应用于左眼图像和右眼图像这两者的相同函数来发送立体图像的场景来说是有用的。

[0180] 在一些实施例中,可以对多幅图像应用相同的映射,但是对结果位置应用偏移。例如,如果图像是从相距一米但在图像与捕获方向之间具有相同关系的场景捕获位置提供的,则能够使用相同的函数,随后添加一米的偏移。在这样的实施例中,偏移可以例如为在客户端设备101处预先确定和已知的,或者偏移可以作为元数据的部分进行传送。因此,在一些实施例中,参数可以描述至少两幅图像的射线原点之间的不同偏移。

[0181] 在一些实施例中,用于将像素图像位置映射到射线原点和/或方向的不同函数可以用于图像中的至少两幅图像。这些不同的函数/映射可以由包含在元数据中的不同参数来表示。

[0182] 因此,先前针对被应用于一幅图像的一条曲线描述的处理可以被独立地应用于多条曲线和图像。

[0183] 这样的方法可以提供额外的灵活性和调整性,并且可以提高质量。

[0184] 在许多实施例中,例如在上述特定示例中,图像可以是全向图像并且可以是完整的360°图像。然而,在一些实施例中,图像中的至少两幅图像可能仅表示部分视场,因此可能不是完全全向的。在这样的情况下,这两幅图像可以特定地表示不同的视场。这可以再次提高灵活性和调整性。

[0185] 应当理解,可以使用用于确定像素位置与射线原点/射线之间的合适映射的任何合适方法,而不会偏离本发明。在一些实施例中,VR服务器103可以被简单地布置为应用经常使用的常量映射(但是客户端设备101可能不知道,因为不同的服务器可能使用不同的映射)。例如,VR服务器103的设计者可能已经考虑到具有给定的大小和偏心率的椭圆可以适合用于VR服务器103所支持的大多数场景,并且它可以使用这样的映射来生成图像。

[0186] 在其他实施例中,VR服务器可以包括用于根据场景特性在不同映射之间进行选择的功能。例如,VR服务器可以被布置用于为例如与足球场相对应的场景选择一条曲线,而为与音乐厅相对应的场景选择不同的曲线。

[0187] 还应当理解,可以使用通过参数描述为像素图像位置的函数的射线原点/射线方向的变化任何合适的表示,例如,提供原点/方向或定义特定(例如,一维或二维)函数的参数的映射图。确切的(一种或多种)关系和描述(一种或多种)关系的参数将取决于个体实

施例的偏好和要求。

[0188] 在一些实施例中,VR服务器103可以例如被布置为执行优化过程以确定映射并因此确定参数。例如,对于给定的场景模型和视点,可以为具有候选参数值的多条可能的候选曲线生成图像表示。然后,该算法可以基于这些图像来合成针对不同视点的视图图像并将它们与通过评价模型直接生成的这样的视图图像进行比较。可以选择得到最小差异的候选曲线和参数。

[0189] 应当理解,为了清楚起见,以上描述已经参考不同的功能电路、单元和处理器描述了本发明的实施例。然而,显然可以使用不同的功能电路、单元或处理器之间的任何合适的功能分布,而不会偏离本发明。例如,被图示为由单独的处理器或控制器执行的功能可以由相同的处理器或控制器来执行。因此,对特定的功能单元或电路的引用仅被视为对用于提供所描述的功能的合适模块的引用,而不是指示严格的逻辑或物理结构或组织。

[0190] 本发明能够以任何合适的形式来实施,包括硬件、软件、固件或这些项目的任何组合。本发明可以任选地至少部分地被实施为在一个或多个数据处理器和/或数字信号处理器上运行的计算机软件。本发明的实施例的元件和分量可以以任何合适的方式在物理上,在功能上和逻辑上实施。实际上,该功能可以在单个单元中,在多个单元中或作为其他功能单元的部分来实施。正因如此,本发明可以在单个单元中实施或者可以在物理上和功能上分布在不同的单元、电路和处理器之间。

[0191] 虽然已经结合一些实施例描述了本发明,但是这并不旨在将本发明限于本文阐述的特定形式。而是,本发明的范围仅由权利要求来限制。另外,虽然似乎结合特定实施例描述了特征,但是本领域技术人员将认识到,根据本发明可以组合所描述的实施例的各种特征。在权利要求中,术语包括并不排除其他元件或步骤的存在。

[0192] 此外,虽然被单独列出,但是多个模块、元件、电路或方法步骤可以通过例如单个电路、单元或处理器来实施。另外,虽然单独的特征可能被包括在不同的权利要求中,但是这些特征可能被有利地组合,并且这些特征被包括在不同的权利要求中并不意味着这些特征的组合是不可行的和/或不是有利的。而且,在一个权利要求类别中包括特征并不意味着对这一类别的限制,而是表明该特征在适当时同样适用于其他权利要求类别。此外,权利要求中的特征的顺序并不意味着这些特征必须以任何特定的顺序工作,特别是方法权利要求中的各个步骤的顺序并不意味着这些步骤必须按该顺序执行。而是,这些步骤可以以任何合适的顺序执行。另外,单数引用并不排除复数。因此,对“一”、“一个”、“第一”、“第二”等的引用并不排除多个。权利要求中的附图标记仅被提供用于清楚地说明示例,而不应被解释为以任何方式限制权利要求的范围。

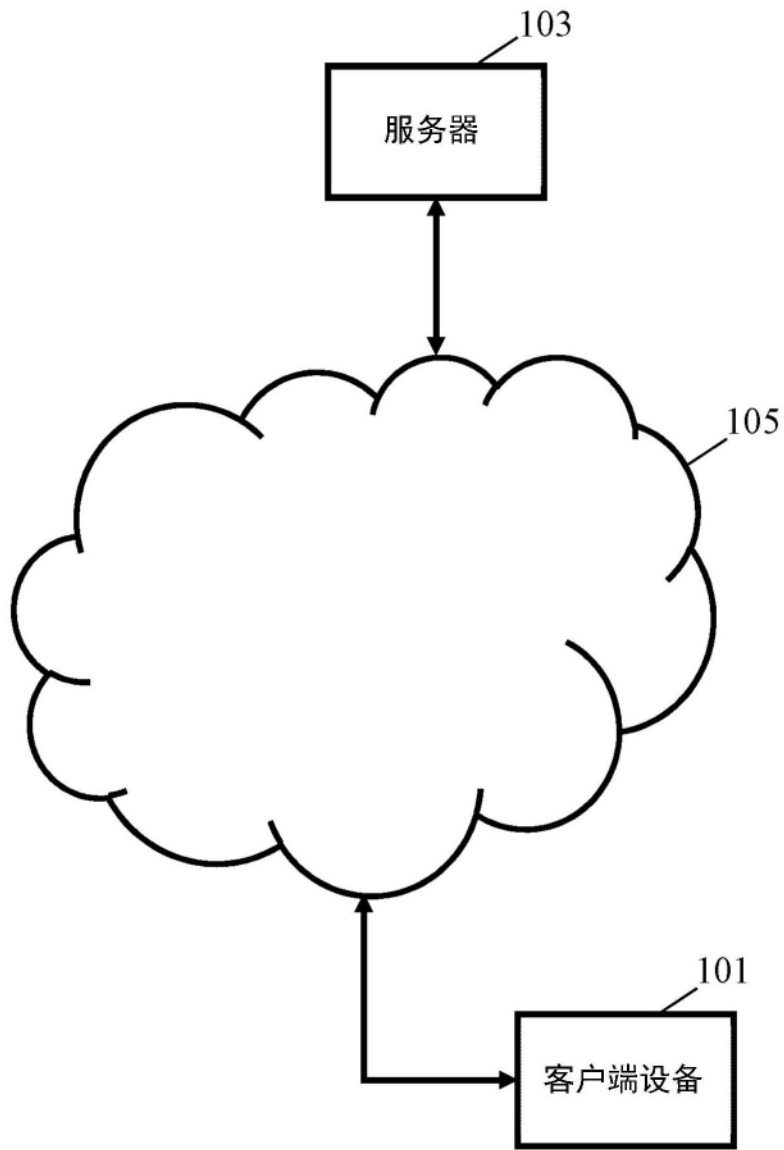


图1

103

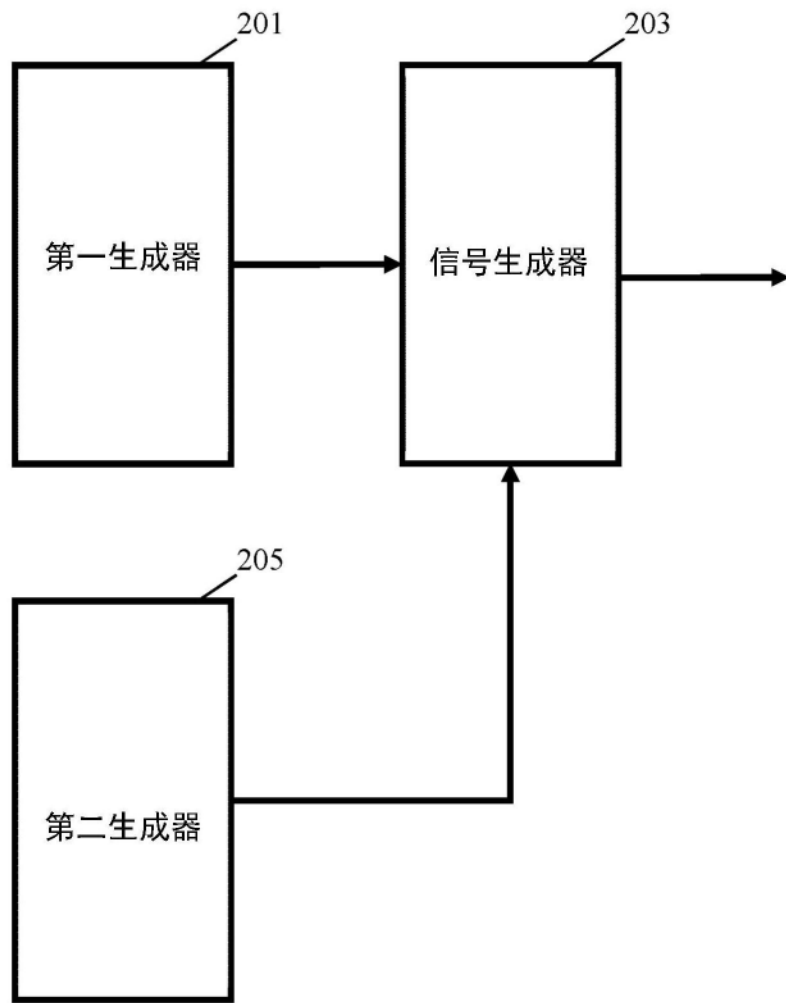


图2

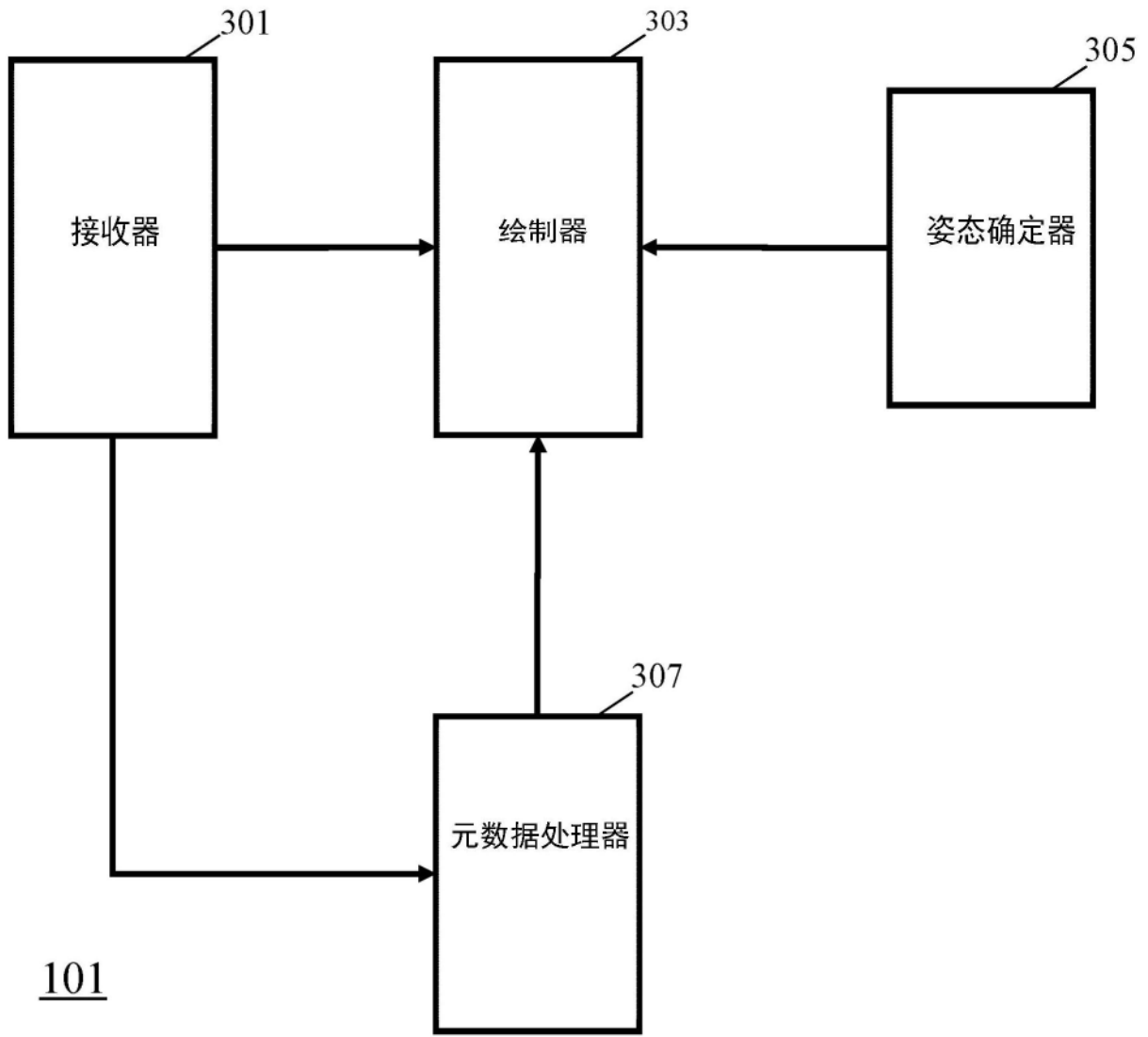


图3

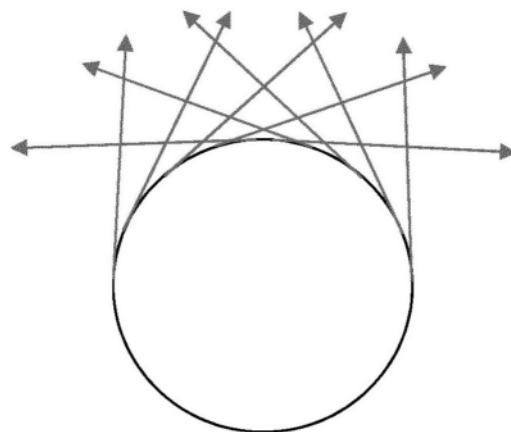


图4

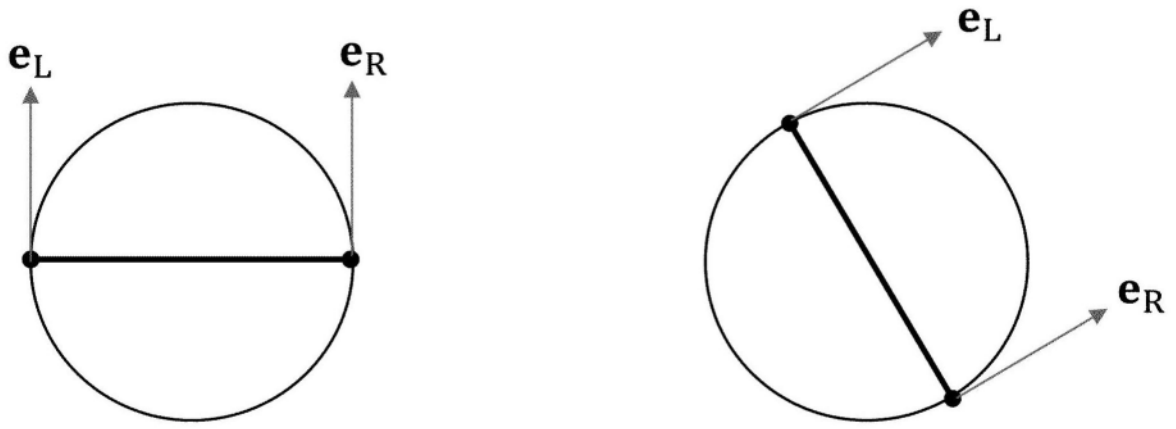


图5

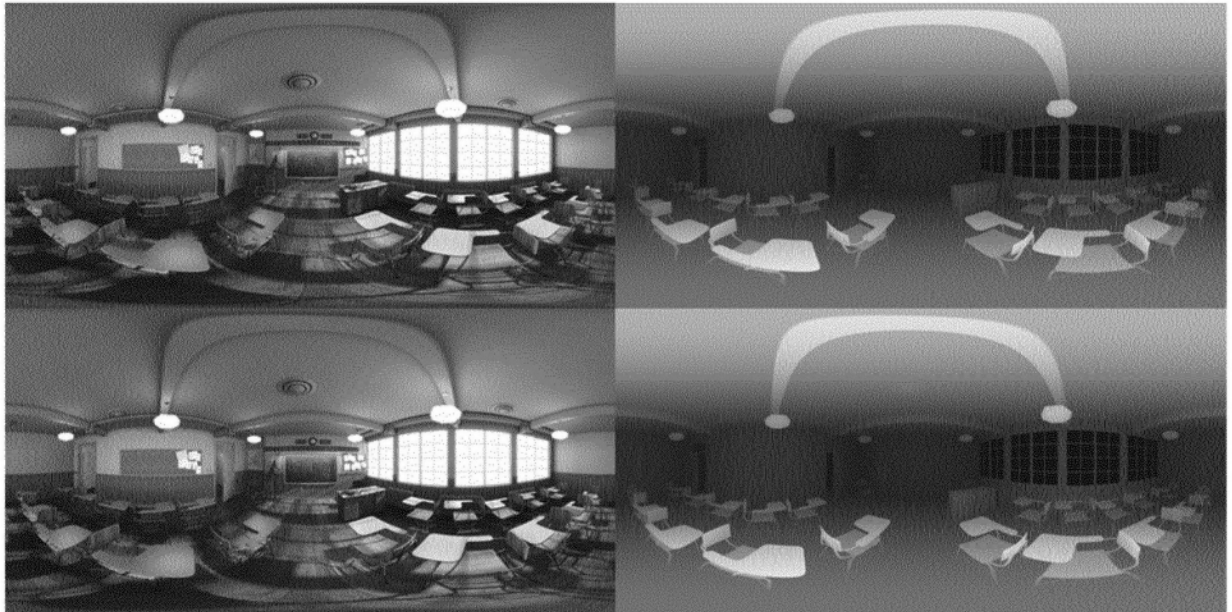


图6

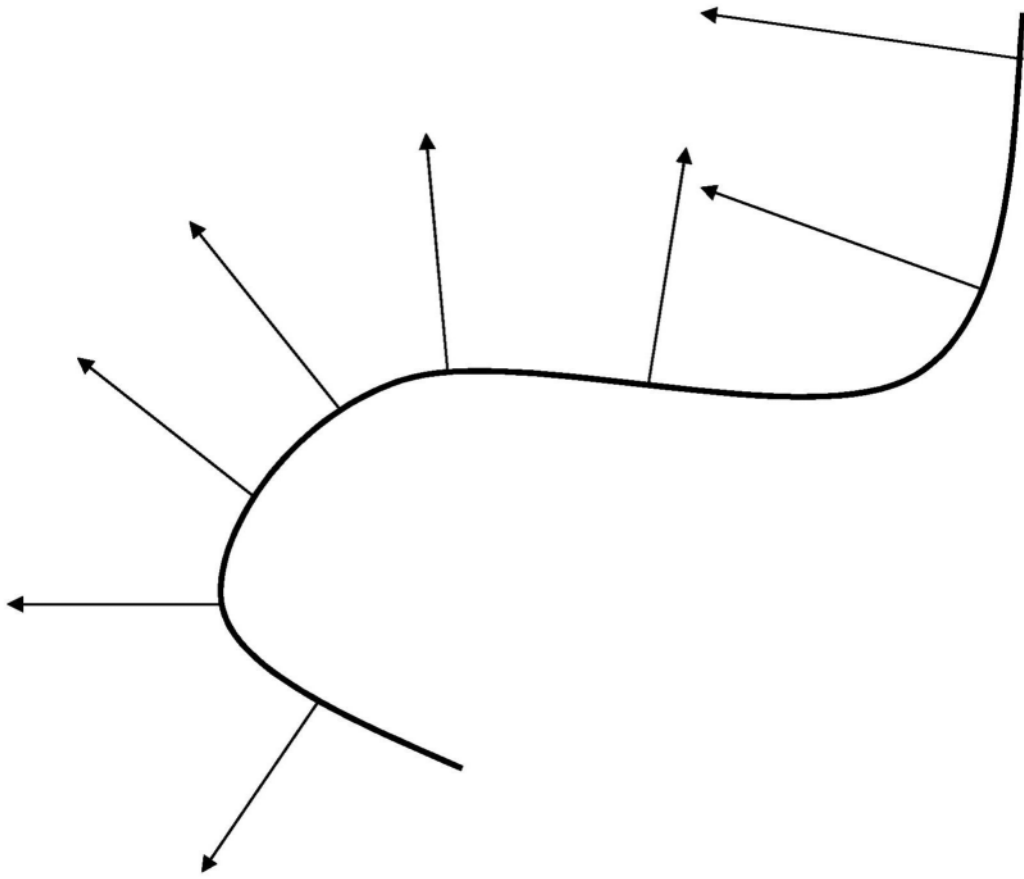


图7

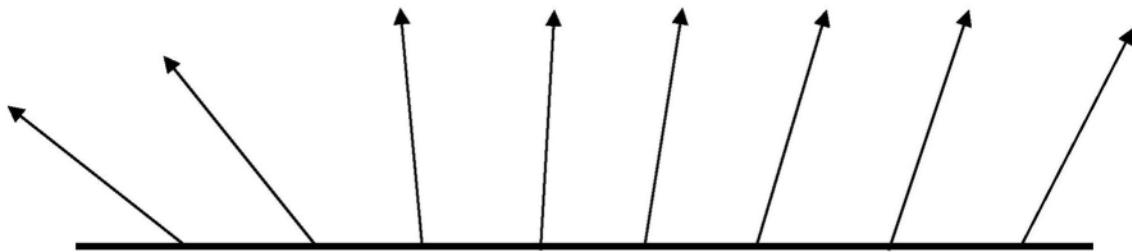


图8

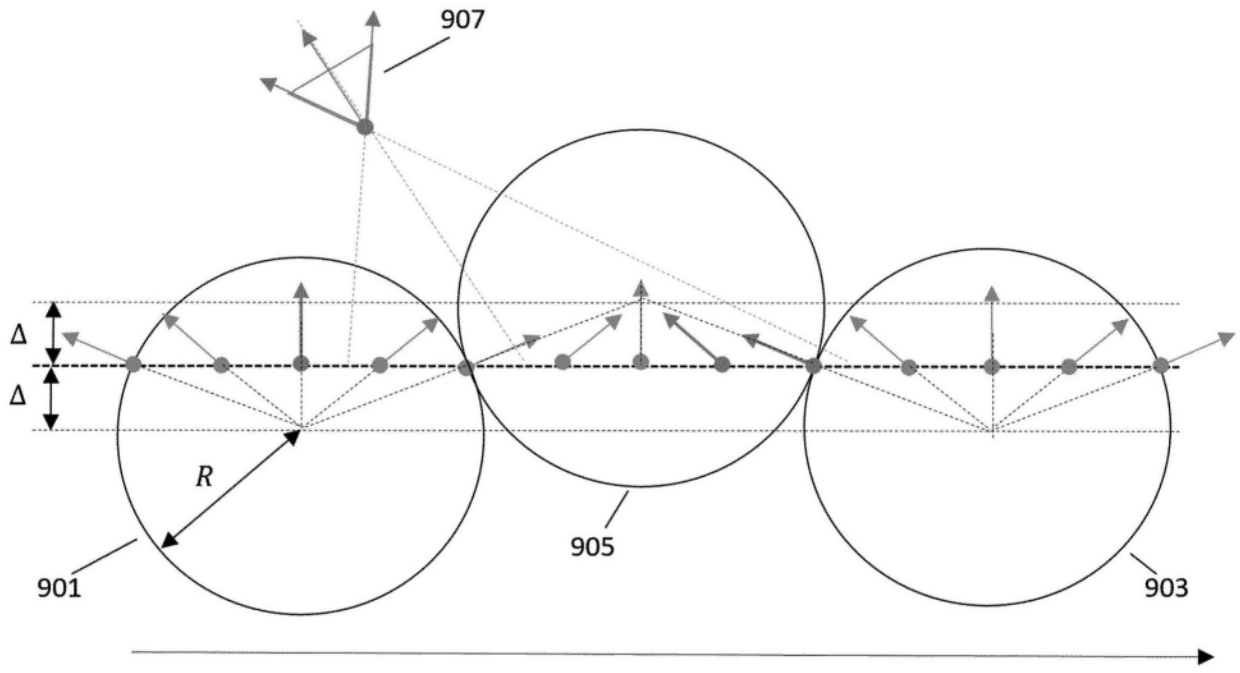


图9

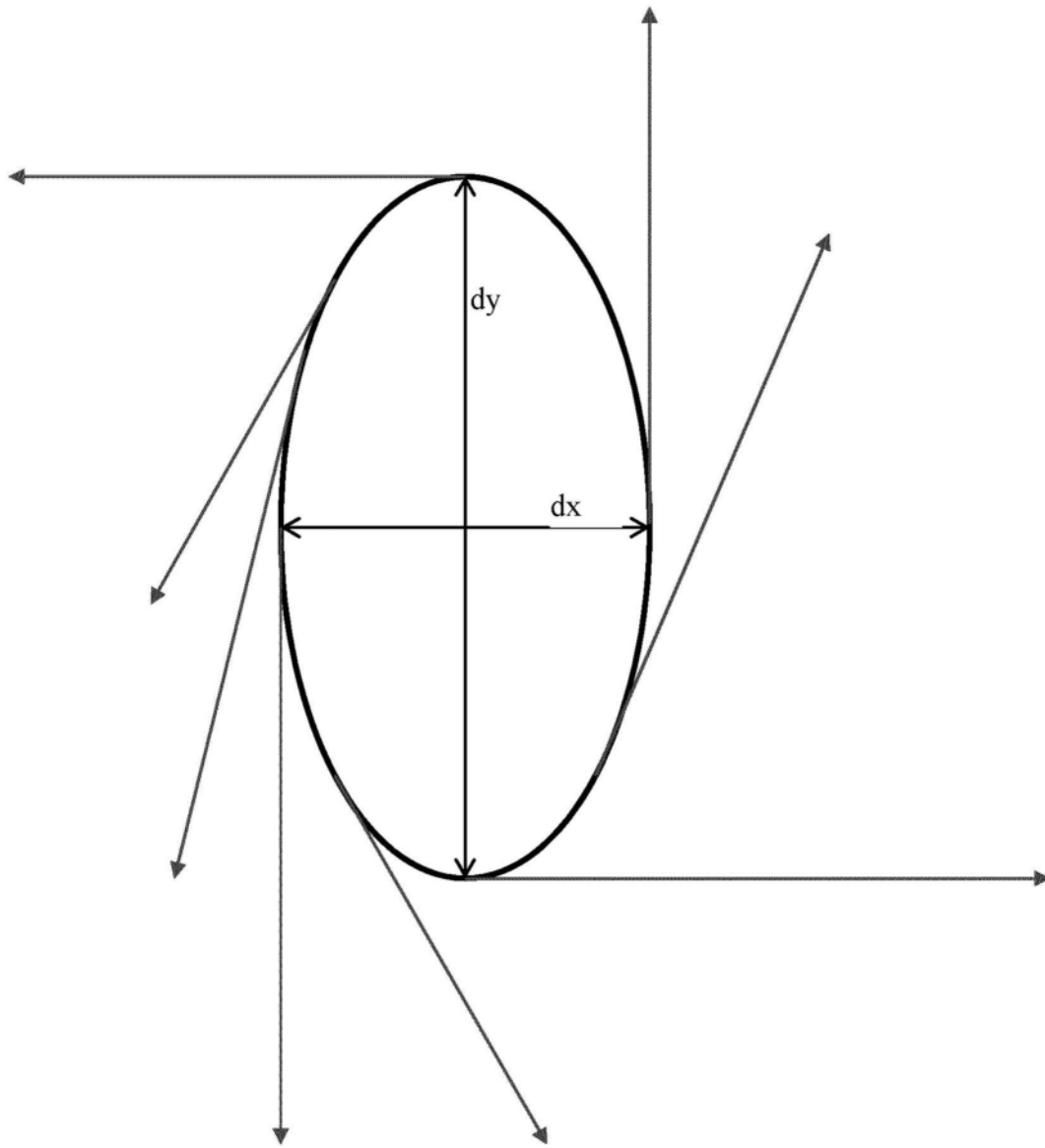


图10