



(12) 发明专利

(10) 授权公告号 CN 110679146 B

(45) 授权公告日 2022.05.03

(21) 申请号 201880034993.0

(73) 专利权人 皇家飞利浦有限公司

(22) 申请日 2018.03.02

地址 荷兰艾恩德霍芬

(65) 同一申请的已公布的文献号

(72) 发明人 C·韦雷坎普 B·克龙

申请公布号 CN 110679146 A

(74) 专利代理机构 永新专利商标代理有限公司
72002

(43) 申请公布日 2020.01.10

代理人 刘兆君

(30) 优先权数据

(51) Int.CI.

17163509.7 2017.03.29 EP

H04N 13/117 (2018.01)

(85) PCT国际申请进入国家阶段日

审查员 裴暑云

2019.11.27

(86) PCT国际申请的申请数据

PCT/EP2018/055183 2018.03.02

(87) PCT国际申请的公布数据

W02018/177681 EN 2018.10.04

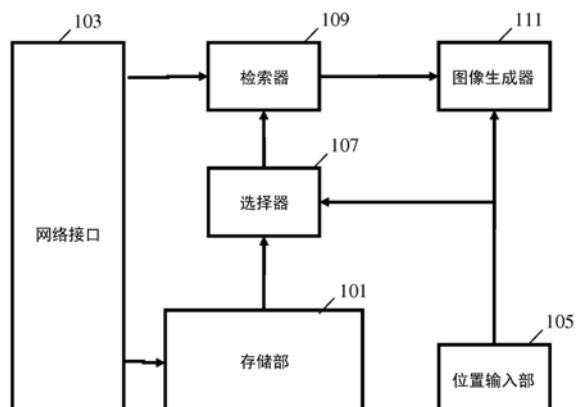
权利要求书2页 说明书19页 附图6页

(54) 发明名称

根据视频进行的图像生成

(57) 摘要

一种装置包括：存储部(101)，其用于存储N维空间中的路线集合，其中，所述路线集合中的每条路线与包括帧的视频项相关联，所述帧包括图像信息和深度信息两者；输入部(105)，其接收观看者位置指示；选择器(107)，其响应于取决于距离度量的选择标准而选择所述路线集合中的第一路线，所述距离度量取决于所述观看者位置指示和所述路线集合中的所述路线的位置；检索器(103、109)，其从视频源(203)中检索与所述第一路线相关联的第一视频项；图像生成器(111)，其根据来自所述第一视频项的第一帧集合来生成针对所述观看者位置指示的至少一幅视图图像。在系统中，相对于所述路线集合中的其他路线，所述选择标准偏向于当前选择的路线。



201

1. 一种用于生成至少一个视图图像的装置,包括:

存储部(101),其用于存储N维空间中的表示空间姿态的路线集合,所述路线集合包括多条路线,并且所述路线集合中的每条路线描述穿过所述N维空间的相机姿态轨迹并且被链接到包括帧的视频项,所述帧包括针对所述路线的相机姿态的图像信息和深度信息两者;

输入部(105),其用于接收观看者位置指示;

选择器(107),其用于响应于取决于距离度量的选择标准而选择所述路线集合中的第一路线,所述距离度量取决于所述观看者位置指示和所述路线集合中的所述路线的位置;

检索器(103、109),其用于从视频源(203)中检索与所述第一路线相关联的第一视频项;

图像生成器(111),其用于根据来自所述第一视频项的第一帧集合来生成针对所述观看者位置指示的至少一幅视图图像;

其中,相对于所述路线集合中的其他路线,所述选择标准偏向于当前选择的路线。

2. 根据权利要求1所述的装置,其中,所述图像生成器(111)被布置为:响应于针对相对于所述观看者位置指示的所述第一视频项的帧的路线位置数据而选择所述第一帧。

3. 根据权利要求1或2所述的装置,其中,所述图像生成器(111)被布置为:如果所述第一路线被选择为所述当前选择的路线,则对所述第一帧集合的选择偏向于到所述第一视频项的当前选择的第一帧集合具有较短时间距离的帧。

4. 根据权利要求1或2所述的装置,其中,所述图像生成器(111)被布置为:如果所述第一路线被选择为所述当前选择的路线,则基于帧选择度量来执行对所述第一帧集合的所述选择,所述帧选择度量随着到所述第一视频项的当前选择的帧的时间距离的减小而增大。

5. 根据权利要求1所述的装置,其中,所述存储部(101)还被布置为:存储针对视频序列的路线的交叉位置;并且其中,所述选择标准取决于所述交叉位置。

6. 根据权利要求5所述的装置,其中,所述选择器(107)被布置为:针对与交叉位置相对应的观看者位置指示而降低朝向所述当前选择的路线的所述偏向。

7. 根据权利要求5所述的装置,其中,朝向所述当前选择的路线的所述偏向取决于所述观看者位置指示与所述交叉位置之间的距离。

8. 根据权利要求5至7中的任一项所述的装置,其中,所述存储部(101)被布置为:存储具有作为节点的交叉以及由边反映的路线的图表。

9. 根据权利要求5-7中的任一项所述的装置,其中,与针对至少一些视频序列的交叉位置相关联的帧是帧内编码帧。

10. 根据权利要求5-7中的任一项所述的装置,其中,所述图像生成器被布置为:至少根据所述第一帧集合中的第一帧和来自被链接到先前选择的第二路线的视频项的第二帧来生成所述视图图像。

11. 根据权利要求10所述的装置,其中,所述第二帧是与所述第一路线和所述先前选择的第二路线之间的交叉相关联的帧。

12. 根据权利要求1所述的装置,其中,所述存储部(101)还被布置为:存储指示所述路线集合中的各条路线之间的空间关系的关系数据,并且对所述第一路线的所述选择响应于所述关系数据。

13. 根据权利要求12所述的装置,其中,所述关系数据可以指示两条路线之间的来自以下组中的至少一种空间关系:

所述两条路线的相交;

所述两条路线的接合;

所述两条路线中的第一路线与所述两条路线中的第二路线的相反移动方向相对应;以及

所述两条路线是平行的。

14. 根据权利要求1-2和5-7中的任一项所述的装置,其中,所述视频项是包括帧内编码帧和帧间编码帧两者的编码视频项,并且所述选择器被布置为:相对于与帧间编码帧相对应的观看者位置指示,针对与帧内编码帧相对应的观看者位置指示而降低朝向所述当前选择的路线的所述偏向。

15. 一种用于生成至少一个视图图像的方法,包括:

存储N维空间中的表示空间姿态的路线集合,所述路线集合包括多条路线,并且所述路线集合中的每条路线描述穿过所述N维空间的相机姿态轨迹并且被链接到包括帧的视频项,所述帧包括针对所述路线的相机姿态的图像信息和深度信息两者;

接收观看者位置指示;

响应于取决于距离度量的选择标准而选择所述路线集合中的第一路线,所述距离度量取决于所述观看者位置指示和所述路线集合中的所述路线的位置;

从视频源中检索与所述第一路线相关联的第一视频项;

根据来自所述第一视频项的第一帧集合来生成针对所述观看者位置指示的至少一幅视图图像;

其中,相对于所述路线集合中的其他路线,所述选择标准偏向于当前选择的路线。

根据视频进行的图像生成

技术领域

[0001] 本发明涉及根据视频项进行的图像生成,并且特别涉及但非排他性地涉及根据存储的可以源自真实世界环境的视频捕获的视频项进行的针对虚拟现实环境的视图图像生成。

背景技术

[0002] 近年来,随着利用和消费视频的新的服务和方式的不断开发和推出,图像和视频应用的种类和范围已经大幅增加。

[0003] 例如,一种日益流行的服务是以使得观看者能够与系统主动且动态地交互以改变绘制的参数的方式来提供图像序列。在许多应用中极具吸引力的特征是改变观看者的有效观看位置和观看方向的能力,例如,允许观看者在正被呈现的场景中移动和“环视”。

[0004] 这种特征能够特别允许给用户提供虚拟现实体验。这可以允许用户在虚拟环境中(相对)自由地四处移动且动态地改变其位置和其正在观看的地方。通常,这种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请求的视图。对于计算机和控制台而言,该方法在例如游戏应用中(例如在第一人称射击游戏类别中)是熟知的。

[0005] 特别是对于虚拟现实应用而言,也期望所呈现的图像是三维图像。实际上,为了优化观看者的身临其境的感觉,通常优选使用户体验的呈现场景为三维场景。实际上,虚拟现实体验应当优选允许用户选择他/她自身的位置、相机视点和相对于虚拟世界的时刻。

[0006] 通常,虚拟现实应用会固有地受到限制,这是因为虚拟现实应用基于预定的场景模型并且通常基于人造的虚拟世界模型。非常期望虚拟现实体验可以基于真实世界捕获来提供。然而,在许多情况下,这种方法极受限制或者趋向于要求根据真实世界捕获而构建的真实世界的虚拟模型。然后通过评价该模型来生成虚拟现实体验。

[0007] 然而,当前方法趋向于次优的用户体验,并且通常趋向于具有高计算要求或高通信资源要求和/或提供具有例如降低质量或有限自由度的次优的用户体验。

[0008] 作为应用的示例,虚拟现实眼镜当前正在进入市场。这些眼镜允许观看者体验所捕获的360度(全景)视频。这些360度视频通常是使用相机装备预先捕获的,其中,个体图像被拼接在一起而成为单个球形映射。360度视频的常见立体格式是上/下和左/右。类似于非全景立体视频,左眼图片和右眼图片被压缩为单个H.264视频流的部分。在解码单个帧之后,观看者旋转他/她的头部来观看他/她周围的世界。当前示例是由谷歌制作的对卡内基音乐厅中的音乐会的记录[https://performingarts.withgoogle.tabhumble.com/en_us/performances/carnegie-hall]。在该示例中,观看者能够体验360度环视效应,并且能够以离散方式在从不同位置记录的三个视频流之间进行切换。在切换时,加载另一视频流,这样会中断体验。

[0009] 立体全景视频方法的一个缺点是观看者无法改变在虚拟世界中的位置。除了全景立体视频以外,全景深度图的编码和传输可以允许补偿观看者在客户端侧的小幅平移运动,但是这种补偿将固有地限于小幅变化和移动并且将无法提供身临其境且自由的虚拟现

实体验。

[0010] 相关技术是其中具有深度图的多个视点在单个视频流中进行编码和传输的自由视点视频。除了众所周知的时间预测方案以外,也可以通过利用视点之间的角度依赖性来减小视频流的比特率。然而,该方法仍然要求高比特率并且在能够生成的图像方面是受限制的。该方法实际上无法提供在三维虚拟现实世界中完全自由移动的体验。

[0011] 遗憾的是,在现有技术中没有一种技术能够递送理想体验,而是通常趋向于在位置和观看方向的变化自由度方面受到限制。另外,这些技术趋向于要求极高的数据速率并且提供包括比生成个体图像/视图所需的数据更多的数据的数据流。

[0012] 另外,现有技术的方法趋向于要求大存储容量,以便存储大量的例如单独捕获的图像。而且,现有技术的方法趋向于要求相对复杂的处理并且具有高资源计算需求。除此之外,现有技术的方法还趋向于次优地生成视图图像,并且特别地,通常给移动于虚拟世界中的用户提供不连贯且相对不一致的体验。

[0013] 例如,KR 20170007098 A公开了一种用于生成针对如下的给定的用户姿态的图像的系统,该给定的用户姿态可以不同于针对其已经捕获了场景的图像的捕获姿态。该方法可以选择最近的捕获姿态并根据来自最近的捕获姿态的图像来生成针对该用户姿态的视图图像。然而,该方法趋向于具有上文所述的所有缺点,并且特别要求高计算资源,非常复杂,基于复杂且困难的捕获操作,并且趋向于生成不是最优质量的且在用户移动时趋向于产生不一致的视图图像。

[0014] 因此,改进的方法将是有利的。特别地,允许改进操作,提高灵活性,增大用于生成与不同的观看位置和方向相对应的图像的范围,改善虚拟现实体验,减小数据速率,便于分布,减小复杂性,便于实施,降低存储要求和/或提高性能和/或操作的方法将是有利的。

发明内容

[0015] 因此,本发明寻求单独地或以任意组合来优选减轻、缓解或消除上文所述的缺点中的一个或多个缺点。

[0016] 根据本发明的一个方面,提供了一种装置,所述装置包括:存储部,其用于存储N维空间中的表示空间姿态的路线集合,所述路线集合包括多条路线,并且所述路线集合中的每条路线描述穿过所述N维空间的相机姿态轨迹并且被链接到包括帧的视频项,所述帧包括针对所述路线的相机姿态的图像信息和深度信息两者;输入部,其用于接收观看者位置指示;选择器,其用于响应于取决于距离度量的选择标准而选择所述路线集合中的第一路线,所述距离度量取决于所述观看者位置指示和所述路线集合中的所述路线的位置;检索器,其用于从视频源中检索与所述第一路线相关联的第一视频项;图像生成器,其用于根据来自所述第一视频项的第一帧集合来生成针对所述观看者位置指示的至少一幅视图图像;其中,相对于所述路线集合中的其他路线,所述选择标准偏向于当前选择的路线。

[0017] 在许多场景中并且对于许多应用,该方法可以提供改善的性能和/或操作。特别地,在许多场景中,该方法可以提供改善的虚拟现实体验。特别地,通常能够实现具有低延迟、高质量、低通信资源要求、低存储资源要求和/或低计算资源要求的高效方法。

[0018] 例如,该方法可以允许降低对客户端-服务器布置的通信或计算资源要求,该客户端-服务器布置用于基于存储视频项的服务器而在客户端设备处提供虚拟现实体验,并且

可能也提供路线数据。

[0019] 例如,该方法可以用于基于与路线数据相关联(被链接到路线数据)的视频项而不是基于个体图像和相关联位置来表示虚拟环境。这例如可以允许使用高效视频编码技术,从而降低存储和通信资源要求。

[0020] 在许多实施例中,该方法可以允许根据真实世界捕获来方便地生成表示虚拟环境的数据。例如,能够基于在真实世界环境中四处移动时所捕获的视频项来有效地生成与真实世界环境相对应的虚拟环境中的虚拟体验。

[0021] 特别地,N维空间可以包括与在N维空间中的位置和/或视图取向有关的维度。N是整数,特别地,N可以表示反映虚拟环境中的位置的1-3个空间维度和反映虚拟环境中的视图取向的1-3个空间维度。

[0022] 特别地,观看者位置指示可以是N维空间中的位置的M维(其中,M通常等于N)指示。特别地,观看者位置指示可以表示反映虚拟环境中的位置的1-3个空间维度和反映虚拟环境中的视图取向的1-3个空间维度。

[0023] 视频项包括图像信息和深度信息两者,从而允许根据可以与通过视频项捕获的视点不同的视点进行的视图图像生成。这样可以提供改善的体验,例如允许用户相对于视频项的图像的相机方向进行转弯。

[0024] 每条路线可以被链接到一个视频项,其中,该路线描述穿过N维空间的与该视频项相对应的(可能为虚拟的)相机轨迹。沿着该路线的位置可以与视频项的帧的位置相关联。

[0025] 选择器可以被布置为:增大响应于针对给定路线的距离度量指示观看者位置指示与给定路线之间的距离减小而选择给定路线的概率。距离度量可以例如基于适合的范数。

[0026] 视图图像可以被生成为对应于观看者在由观看者位置指示给定的N维空间中的位置处将感知到的视图。

[0027] 输入部可以接收观看者位置指示的序列,并且该装置可以被布置为:生成对应的视图图像的序列。该装置的操作可以动态迭代,从而提供动态用户体验。例如,观看者位置指示可以描述穿过N维空间的路径或路线,并且作为响应,该装置可以生成与观看者沿着该路径或路线移动将体验的视图相对应的图像序列。

[0028] 当前选择的路线可以是针对先前观看者位置指示而选择的路线,并且特别地,可以是针对最近接收的观看者位置指示而选择的路线。特别地,观看者位置指示可以是观看者位置指示的序列中的观看者位置指示,并且当前选择的路线可以是针对在当前的观看者位置指示之前所接收的最近的观看者位置指示而选择的路线。

[0029] 视频项可以是编码视频项。视频项可以包括帧内编码帧和帧间编码帧。帧内编码帧是独立于其他帧而编码的。帧间编码帧是相对于至少一个其他帧而编码的。视频项可以包括I帧以及B编码帧和/或P编码帧。

[0030] 可以通过与针对该路线集合中的其他路线不同地针对当前选择的路线的选择标准来实施朝向当前选择的路线的偏向。该选择标准相对于当前选择的路线和该路线集合中的其他路线可以是非对称的。特别地,相对于针对该路线集合中的其他路线的距离度量,选择标准相对于针对当前选择的路线的距离度量可以是非对称的。

[0031] 选择标准可以被布置为:引起针对观看者位置指示的不同值而选择不同路线。偏向可以使得将选择当前选择的路线,除非到另一路线的距离度量比到当前路线的距离度量

短一个因子以上。

[0032] 第一帧集合可以是一个或多个帧。在一些实施例中，图像生成器可以被布置为：响应于来自一个以上的视频项的帧而进一步生成视图图像。例如，如果选择新的路线，则可以根据先前选择的路线和新选择的路线两者的帧来生成视图图像。

[0033] 根据本发明的任选特征，所述图像生成器被布置为：响应于针对相对于所述观看者位置指示的所述第一视频项的帧的路线位置数据而选择所述第一帧。

[0034] 在许多实施例中，这样可以提供改善的性能和/或操作。

[0035] 根据本发明的任选特征，所述图像生成器被布置为：如果所述第一路线被选择为所述当前选择的路线，则对所述第一帧集合的选择偏向于到所述第一视频项的当前选择的第一帧集合具有较短时间距离的帧。

[0036] 在许多实施例中，这样可以提供改善的性能和/或操作。特别地，这样可以提供改善的时间一致性并且可以减轻在时间上变化的伪影。

[0037] 根据本发明的任选特征，所述图像生成器被布置为：如果所述第一路线被选择为所述当前选择的路线，则基于帧选择度量来执行对所述第一帧集合的所述选择，所述帧选择度量随着到所述第一视频项的当前选择的帧的时间距离的减小而增大。

[0038] 在许多实施例中，这样可以提供改善的性能和/或操作。特别地，这样可以提供改善的时间一致性并且可以减轻在时间上变化的伪影。

[0039] 根据本发明的任选特征，所述存储部还被布置为：存储针对视频序列的路线的交叉位置；并且所述选择标准取决于所述交叉位置。

[0040] 在许多实施例中，这样可以提供改善的性能，并且特别地，可以提供对合适的视频项的改善的选择。例如，在一些实施例中，这可以例如通过允许该装置检索针对潜在的与当前路线交叉的未来路线的视频项来降低延迟。

[0041] 交叉位置可以指示路线之间的交叉，并且可以例如以N维空间的维度（中的至少一些维度）的坐标的形式来提供，或者例如相对于路线数据或视频项（例如，帧编号）可以是给定的。

[0042] 如果两条路线的位置处于低于阈值的距离处，则可以被视为存在交叉。

[0043] 交叉可以是任何接合，包括T型接合或一条路线的结束与另一路线的开始之间的接合。

[0044] 根据本发明的任选特征，所述选择器被布置为：针对与交叉位置相对应的观看者位置指示而降低朝向所述当前选择的路线的所述偏向。

[0045] 这样可以提供改善的性能，并且特别地，可以通过使得与其他地方相比更可能在交叉处发生不同视频项之间的转变来改善的这种转变。在一些实施例中，可以仅在交叉处允许到新的路线的转变。

[0046] 根据本发明的任选特征，其中，朝向所述当前选择的路线的所述偏向取决于所述观看者位置指示与所述交叉位置之间的距离。

[0047] 这样可以提供改善的性能，并且特别地，可以通过使得与其他地方相比更可能在交叉处发生不同视频项之间的转变来改善的这种转变。在一些实施例中，可以仅在交叉处允许到新的路线的转变。

[0048] 根据本发明的任选特征，所述存储部被布置为：存储具有作为节点的交叉以及由

边反映的路线的图表。

[0049] 在许多实施例中,这样可以提供特别高效的操作。

[0050] 根据本发明的任选特征,与针对至少一些视频序列的交叉位置相关联的帧是帧内编码帧。

[0051] 在许多实施例中,这样可以便于和/或改善操作。例如,这样可以便于不同路线/视频项之间的转变,这是因为这样可以减少需要被解码的帧的数量。

[0052] 根据本发明的任选特征,所述图像生成器被布置为:至少根据所述第一帧集合中的第一帧和来自先前选择的第一路线的第二帧来生成所述视图图像。

[0053] 在许多实施例中,这样可以便于和/或改善操作。

[0054] 第一帧可以被选择为帧内编码帧,并且/或者第二帧可以被选择为帧内编码帧。特别地,第一帧可以是交叉帧。视图图像的生成可以包括根据来自先前选择的路线的帧内编码(I)帧和来自第一路线的帧内编码(I)帧(第一帧)来对预测帧(通常为双向(B)帧)进行解码。(先前选择的路线可以对应于在选择第一路线之前的当前选择的路线,特别是在选择器将第一路线选择为与当前选择的路线(在选择第一路线之前)不同的路线时)。这通常可以是当该装置从一条路线切换到另一路线时的场景。

[0055] 根据本发明的任选特征,所述第二帧是与所述第一路线和所述先前选择的第二路线之间的交叉相关联的帧。

[0056] 在许多实施例中,这样可以便于和/或改善操作。

[0057] 根据本发明的任选特征,所述存储部还被布置为:存储指示所述路线集合中的各条路线之间的空间关系的关系数据,并且对所述第一路线的所述选择响应于所述关系数据。

[0058] 在许多实施例中,这样可以提供改善的性能和/或操作。

[0059] 根据本发明的任选特征,所述关系数据可以指示两条路线之间的来自以下组中的至少一种空间关系:所述两条路线的相交;所述两条路线的接合;所述两条路线中的第一路线与所述两条路线中的第二路线的相反移动方向相对应;以及所述两条路线是平行的。

[0060] 在许多实施例中,这样可以提供改善的性能和/或操作。

[0061] 根据本发明的任选特征,所述视频项是包括帧内编码帧和帧间编码帧两者的编码视频项,并且所述选择器被布置为:相对于与帧间编码帧相对应的观看者位置指示,针对与帧内编码帧相对应的观看者位置指示而降低朝向所述当前选择的路线的所述偏向。

[0062] 这样可以降低计算资源使用情况,并且特别地,可以降低用于对适合于生成视图图像的帧进行解码的复杂性和资源使用情况。

[0063] 根据本发明的一个方面,提供一种方法,所述方法包括:存储N维空间中的表示空间姿态的路线集合,所述路线集合包括多条路线,并且所述路线集合中的每条路线描述穿过所述N维空间的相机姿态轨迹并且被链接到包括帧的视频项,所述帧包括针对所述路线的相机姿态的图像信息和深度信息两者;接收观看者位置指示;响应于取决于距离度量的选择标准而选择所述路线集合中的第一路线,所述距离度量取决于所述观看者位置指示和所述路线集合中的所述路线的位置;从视频源中检索与所述第一路线相关联的第一视频项;根据来自所述第一视频项的第一帧集合来生成针对所述观看者位置指示的至少一幅视图图像;其中,相对于所述路线集合中的其他路线,所述选择标准偏向于当前选择的路线。

[0064] 参考下文描述的(一个或多个)实施例,本发明的这些方面、特征和优点以及其他方面、特征和优点将是明显的并且得到阐明。

附图说明

[0065] 将仅通过示例的方式参考附图来描述本发明的实施例,在附图中:

[0066] 图1图示了根据本发明的一些实施例的用于响应于观察者位置指示而生成视图图像的设备的元件的示例;

[0067] 图2图示了包括如客户端设备的图1的设备的客户端服务器布置的示例;

[0068] 图3图示了用于图2的客户端服务器布置的服务器的示例;

[0069] 图4图示了用于捕获环境的捕获路线的示例;

[0070] 图5图示了在虚拟环境中相对于捕获路线的观察者路线的示例;

[0071] 图6图示了用于虚拟环境的捕获路线的示例;

[0072] 图7图示了表示虚拟环境的捕获路线/视频项的图表的示例;并且

[0073] 图8图示了表示虚拟环境的捕获路线/视频项之间的关系的图表的示例。

具体实施方式

[0074] 允许用户在虚拟世界中四处移动的虚拟体验变得日益流行,并且满足这种需求的服务正在进行开发。然而,提供高效的虚拟现实服务是极具挑战性的,特别是在体验要基于真实世界环境的捕获而非基于完全虚拟生成的人造世界的情况下。

[0075] 用于提供虚拟现实服务的方法是存储大量图像,其中,每幅图像被链接到特定位置并通常被链接到视图取向/方向。在虚拟世界中给定用户位置的情况下,能够搜索所存储的位置以检索针对该位置(或针对该位置附近)的图像,并且这能够被适当地绘制给用户。例如,服务器可以接收用户的位置,并且可以从本地图像存储部中提取相关图像并生成与用户穿过虚拟世界的路径相对应的专用视频序列。

[0076] 然而,这种方法将要求存储大量数据以确保充分覆盖虚拟环境。所要求的图像的密度以及因此要求存储的图像的总数可以通过连同图像一起存储深度信息而减小。这可以允许针对没有与所存储的位置和方向精确对准的视图位置和方向的视图图像的本地生成。然而,通常仍然需要大量的存储图像。另外,个体地存储图像和相关联的位置所要求的数据结构通常不允许优化的存储效率。

[0077] 另外,该方法引起通常具有相对较高的计算和通信资源使用需求的相对较高的复杂性。特别地,该方法不能很好地扩展到例如通过互联网访问共同的服务器的越来越多的用户。在这种情况下,服务器将需要针对每个用户生成专用视频信号并因此针对每个用户个体地执行(通常复杂的)视频编码。

[0078] 图1图示了可以在许多场景中提供改善的虚拟现实用户体验的装置。在一些实施例中,该装置可以是客户端设备201,该客户端设备201如图2所示例如经由网络205(例如,互联网)与远程服务器203建立联系。服务器203可以被布置为:同时支持潜在大量的客户端设备201。

[0079] 该装置被布置为:生成与针对给定的观察者观看姿态的场景的视图相对应的图像。在许多实施例中,该装置被布置为:生成针对给定的观察者姿态的立体图像,或者等效

地,可以生成针对与观看者的两眼之间的偏移相对应的稍微不同的观察者姿态的两幅图像。

[0080] 在该领域中,术语放置和姿态用作针对位置和/或方向/取向的共用术语。例如对象、相机或视图的位置与方向/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/分量通常描述对应的对象的位置/定位或取向/方向的个体属性。当然,在许多情况下,例如,如果一个或多个分量被视为是固定的或不相关的,则放置或姿态可以被视为具有较少分量或用较少分量来表示(例如,如果所有对象都被视为处于相同高度且具有水平取向,则四个分量可以提供对对象的姿态的完整表示)。在下文中,术语姿态用于指代可以由(与可能的最大自由度相对应的)一个至六个值表示的位置和/或取向。说明将集中于姿态具有最大自由度(即,位置和取向中的每个具有三个自由度得到总共六个自由度)的实施例和示例。因此,姿态可以由表示这六个自由度的六个值的集合或向量来表示,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,姿态也可以由较少的值来表示。

[0081] 该装置包括存储部101,该存储部101被布置为:存储针对N维空间(其中,N是整数)中的路线集合的路线数据。该N维空间可以是虚拟世界的空间,并且特别地,可以是与用户在该空间中的位置相对应的三维空间。在一些实施例中,这些维度中的一个或多个维度可以与虚拟空间中的视图取向相对应。因此,在一些实施例中,用户可以由三个自由度(3DoF)来表示,并且在其他实施例中,用户可以由例如与三维位置和三维取向相对应的六个自由度(6DoF)来表示。该N维空间表示空间姿态,并且特别地,表示在表示虚拟世界中的空间维度的N维空间中的姿态。

[0082] 因此,存储部101存储N维空间中的路线(例如,路径或轨迹)。这些路线中的每条路线提供穿过虚拟现实世界的路径或轨迹。另外,已经存储的路线集合中的每条路线与视频项相关联/被链接到视频项。因此,每条路线被链接到包括与路线的位置和/或取向相对应的图像序列的视频项。等效地,每个视频项可以提供与沿着路线的视图相对应的图像序列。因此,针对给定的路线的视频项包括与在沿着路线移动时的视图相对应的图像序列。因此,针对特定路线的视频项提供在相机沿着路线移动时将由相机捕获的或者等效地将由沿着路线移动的虚拟用户体验的图像,而不是与静态图像、位置或视图相对应的视频项。每个视频项包括针对经链接的路线的相机姿态的图像信息和深度信息两者的帧,并且每条路线描述穿过N维空间的对应的相机姿态轨迹(相机姿态)。

[0083] 每个视频项可以包括与沿着路线的位置序列相对应的图像序列。播放视频项将提供沿着路线移动的人或相机的视图。这些路线中的每条路线可以由沿着路线的位置序列来表示。在一些实施例中,可以仅考虑明确定义的位置(例如,针对每个帧可以包括一个位置),而在其他实施例中,可以通过内插来确定存储部中存储的明确值之间的位置。在其他实施例中,每条路线可以例如通过提供例如作为时间变量的函数的三个位置值的数学函数来定义。时间变量可以例如被量化成与针对对应的视频项的帧时间相对应的时间段。

[0084] 在该示例中,存储部101被耦合到将客户端设备201接口连接到网络205(在以下说明中将其视为互联网)的网络接口103。特别地,网络接口103可以为客户端设备201提供连接到服务器203的功能。

[0085] 在该示例中,服务器203是能够同时服务多个用户的共用服务器。图3图示了服务

器203的一些元件。在该示例中,服务器203包括服务器网络接口301,该服务器网络接口301被布置为:将服务器203接口连接到网络205,从而允许服务器203与客户端设备201通信。服务器网络接口301被耦合到服务器控制器303,该服务器控制器303被布置为:控制和执行服务器203的许多操作。服务器控制器303被耦合到视频存储部305和路线存储部307。

[0086] 在该示例中,路线存储部307存储针对该路线集合的路线数据,并且视频存储部305存储对应的链接的视频项。特别地,服务器203可以通过如前所述的具有相关联的视频项的路线集合来表示虚拟现实世界。例如,虚拟环境可以由穿过虚拟环境的N维空间的(可能相对较大的)路线数来表示。针对这些路线中的每条路线,存储表示沿着这些路线的视图的视频项,即,可以生成用于表示在相机或用户沿着路线移动时相机或用户的视图的视频项。

[0087] 在许多实施例中,虚拟环境可以是真实环境的表示,并且视频项和路线可以通过真实环境中的记录/捕获来生成。例如,可以携带位置感知视频相机穿过真实环境,并且所捕获的视频可以被存储为视频项,并且对应的位置信息可以被存储为路线。在一些实施例中,可以例如通过使虚拟相机移动穿过虚拟环境来根据虚拟环境的虚拟模型生成视频项和路线数据。

[0088] 所存储的视频项包括表示沿着路线的视图的帧序列/图像序列。该方法的特定优点是能够使用视频编码技术来编码帧/图像,从而生成视频项。这允许非常高效的编码,得到针对给定的图像质量的较低的数据速率,并且由此通常允许比在编码个体图像时更加高效的编码。特别地,能够使用诸如运动估计和(通常双向的)预测性视频编码之类的视频编码技术来减少数据量。在特定的实施例中,将帧编码为I帧(帧内——在不依赖于任何其他帧的情况下进行编码)或编码为B帧(双向帧——特别基于根据I帧的运动估计的相对于I帧进行编码的剩余帧)。另外,除了降低的存储要求和通信要求以外,视频编码的使用也允许使用高效的现成的编码和解码功能(包括专用硬件)。可以使用的视频编码技术例如包括H.264或MPEG2编码。

[0089] 视频项包括图像信息和深度信息两者。特别地,视频项可以是3D视频项,其中,光强度图像/帧与针对帧的深度信息相关联。应当理解,可以使用不同方法来提供深度信息,包括立体图像、明确深度或视差图等。深度信息允许灵活的视图生成,并且特别地允许、便于或改善与没有与路线/视频项提供图像的位置精确对齐的位置(/取向)相对应的视图图像的生成。

[0090] 在该示例中,虚拟环境因此由例如通过使相机移动穿过真实世界环境而捕获的通常相对稀疏的空间路线集合来表示。每条路线被链接到视频项以及呈以路线数据的形式的元数据,该路线数据提供针对视频项的位置信息,并且特别地可以提供针对视频项的每个帧的位置指示。图4图示了二维环境可以如何被多个(可能相当任意的)路线覆盖的示例。

[0091] 在该示例中,客户端设备201被布置为:从服务器203请求路线数据和视频项,并且作为响应,服务器203将接着提供所请求的数据。

[0092] 然而,在该示例中,客户端设备201不是仅仅下载所有数据并在本地处理这些数据,而是首先仅检索路线数据,然后使用该路线数据来动态地确定哪条(哪些)路线靠近所需要的位置。然后从服务器203中选择性地检索针对所选择的路线的视频项并使用该视频项来生成适合的视图。

[0093] 因此,在特定的示例中,当客户端设备201启动虚拟现实应用时,客户端设备201从服务器203请求路线数据。作为响应,服务器203从路线存储部307中提取路线数据并将该路线数据传输到客户端设备201。然而,最初并不提供视频数据,而是仅通过网络205来传输路线数据。

[0094] 更详细地,当初始化新的虚拟现实应用时,网络接口103将用于传输路线存储部307中存储的路线数据的请求传输到服务器203。作为响应,服务器203传输该路线数据。该路线数据由网络接口103接收并且被存储在存储部101中。因此,在初始下载之后,存储部101存储针对N维空间中的路线集合的路线数据的本地复制品。虽然这些路线中的每条路线与表示路线的视图的对应的视频项相关联(/链接到该对应的视频项),但是并不下载这些视频项,而是将这些视频项保留在服务器处。

[0095] 客户端设备201还包括被布置为:接收观看者位置指示的位置输入部105。观看者位置指示对观看者在虚拟环境的N维空间中的位置进行指示,并且因此对虚拟环境的所需视图进行指示。在一些实施例中,观看者位置指示可以响应于特定的用户输入而被生成,例如通过用户直接提供位置或者例如控制操纵杆或鼠标而生成。在其他实施例中,观看者位置指示可以通过例如也在客户端设备201上运行的应用而生成。例如,观看者位置指示可以通过游戏应用而生成。

[0096] 观看者位置指示可以是提供针对观看者的三维位置和三维取向的完整的六维姿态指示。在其他实施例中,例如可以仅提供二维位置以及二维取向(例如当观看者始终被认为处于相同的高度且路线实质上是二维路线时)。

[0097] 客户端设备201还包括被耦合到存储部101和位置输入部105的选择器107。选择器107被布置为:基于所接收的观看者位置指示来从所存储路线集合中选择路线。特别地,针对给定的观看者位置指示,选择器107被布置为:评价取决于距离度量的选择标准,该距离度量取决于观看者位置指示和该路线集合中的路线的位置。因此,基于观看者位置指示和针对所存储的路线的位置数据,评价距离度量并且在合适的情况下选择一条或多条路线来生成针对当前的观看者位置指示的视图图像。

[0098] 客户端设备201还包括被耦合到选择器107和网络接口103的检索器109。检索器109接收关于哪个(哪些)视频项已经被选择的信息,并且作为响应,经由网络接口103而与服务器203通信以检索(一个或多个)选择性视频项。

[0099] 客户端设备201还包括图像生成器111,图像生成器111用于根据来自第一视频项的第一帧集合来生成针对观看者位置指示的至少一幅视图图像。所检索的视频项被转发到图像生成器111,该图像生成器接着生成与观看者在由观看者位置指示所指示的虚拟环境中的位置处将看到的图像相对应的图像。用于生成视图图像的帧的选择也可以基于观看者位置指示和空间距离度量。实际上,由于视频项中的帧具有沿着路线的位置,因此可以选择(例如经受其他参数或考虑因素)与观看者位置指示的位置最紧密对应的(一个或多个)帧。

[0100] 因此,针对给定的观看者位置指示,客户端设备201可以接着选择该视频片段的(一个或多个)适当视频项和(一个或多个)适当帧并且使用所选择的帧来生成视图图像。在一些实施例和场景中,所选择的帧可以直接充分紧密对应于针对观看者位置指示的视图图像,并且所生成的视图图像可以被直接生成为所选择的帧的复制品。

[0101] 然而,在大部分实施例和场景中,(一个或多个)帧的位置(包括取向)与观看者位

置指示之间的差异能够通过处理帧来进行补偿。特别地,该处理可以是基于深度信息的视图变换。应当理解,许多用于这种视图变换的方法对于本领域技术人员将是已知的,并且能够在不减损本发明的情况下使用。

[0102] 因此,客户端设备201被布置为:基于所接收的观看者位置指示来生成视图图像。实际上,所描述的过程是迭代的,并且因此客户端设备201能够响应于观看者位置指示的序列而生成视图图像的序列。特别地,客户端设备201能够生成针对沿着路线移动穿过虚拟环境的观看者的视图序列。

[0103] 所生成的图像能够提供观看者在遵循由观看者位置指示的序列所指示的轨迹移动穿过虚拟环境时将看到的视图。然而,在该示例中,这是根据基于视频项和路线的方法来实现的,而非通过依赖于个体图像来实现的。这提供更加高效的方法,并且允许高效编码,便于捕获,减小复杂性,降低存储要求等。此外,该系统可以允许实质上降低客户端设备201中的通信要求和存储要求,这是因为仅需要检索和下载所要求的视频项。因此,该方法可以区分对路线数据和视频项数据的处理,并且特别是区分对路线数据和视频项数据的通信。由于路线数据通常将比视频数据低至少一个数量级,因此这可以提供对下载通信要求的实质降低。

[0104] 对适当路线的选择基于距离度量,该距离度量反映(该)观看者位置指示(的位置)与路线的位置之间的距离。特别地,距离度量可以反映观看者位置指示与针对当前选择的视频项/路线的当前帧的位置之间的距离,以及观看者位置指示与针对当前未选择的路线的路线位置之间的最小距离。该距离可以根据任何适合的范数或例如通过适合的范数的单调增加函数来确定。

[0105] 这种距离度量可以通过指示哪条路线具有最靠近所需要的位置的位置来提供对适合的路线的良好指示。然而,在图1的客户端设备201中,选择标准不仅仅基于距离度量并且不仅仅选择最靠近所需要的视点的路线,而是,选择器107被布置为:使用相对于其他路线偏向于当前选择的路线的选择标准。因此,该选择标准可以偏向于停留在当前路线上,并且因此基于相同的视频项来生成视图图像而非跳转到另一视频项。

[0106] 在图1的客户端设备201中,选择器107的选择标准相对于针对当前选择的路线(以及因此视频项)的距离度量与针对当前未选择的其他路线(以及因此视频项)的距离度量是非对称的。特别地,在将选择其他路线之前(即,在客户端设备201将切换到另一路线/视频项之前),所使用的选择标准将要求针对另一路线的距离度量比针对当前路线的距离度量小特定裕量或阈值。

[0107] 可以提供参考图5的特定的示例,图5图示了两条平行路线RT1和RT2,其中,RT1是具有基于对应的视频项的当前生成的视图图像的当前选择的路线。观看者穿过环境的路线RT3位于两条所捕获的路线RT1、RT2之间并且距这两条路线的距离是变化的。

[0108] 针对该示例的选择标准的示例可以是:

$$[0109] r_{\text{选择}} = \begin{cases} \text{argmin}_r (d_r) & \text{如果 } d_r > \Delta \\ r_{\text{当前}} & \text{其他情况} \end{cases}$$

[0110] 其中, $r_{\text{当前}}$ 是当前选择的路线, d_r 是到路线 r 的最短距离,并且 $r_{\text{选择}}$ 是所选择的路线。该选择标准具有以下结果:切换到不同的路线和视频仅发生在偏离当前路线超过阈值距离

Δ 时。因此,只有在到当前选择的路线的距离超过特定阈值 Δ 时,该路线才被选择为到观看者位置指示的距离最短的路线。图5图示了正在应用的该选择标准。在该示例中,虽然RT2在一些位置处更靠近观看者位置参数向量,但是所选择的路线始终是RT1。

[0111] 选择标准偏向于当前选择的路线在实践中可以提供实质益处。首先,其趋向于提供具有降低的通信要求的更高效的方法,这是因为需要检索和下载的视频项的数量通常能够被大大减少。

[0112] 另外,其可以提供具有较少的可感知伪影的实质上更一致的输出。实际上,发现提供了更稳定且时间上和空间上更一致的虚拟环境体验。

[0113] 在一些实施例中,路线数据不限于针对路线的位置信息,并且也包括提供关于路线如何彼此相关的信息的数据。

[0114] 特别地,路线数据还可以包括与路线之间的交叉有关的信息。例如,如图4中所例示的,在许多实施例中,路线可以基于相对随机的路线,并且这样可以得到路线之间的许多交叉和重叠。实际上,交叉通常也可以发生在有意地生成特定路线以提供对环境的良好覆盖时。例如,可以通过有意地走过多条平行和正交的路线以生成网格路线系统来覆盖给定的空间。在这种示例中,不同的正交路线之间将存在许多交叉。

[0115] 在不同的实施例中,可以使用不同的标准来考虑路线之间存在的交叉。当两条路线具有共同位置时,可以认为这两条路线之间发生交叉。交叉包括两条路线之间的任何接合(例如,T型接合)或两条以上的路线之间的接合。在一些实施例中,当一条路线上的位置与另一路线上的位置之间的距离低于阈值时,可以认为存在交叉/接合。该阈值通常非常低,并且因此,仅在两条路线至少在一个位置处实质重叠时,才通常认为存在交叉。

[0116] 在许多实施例中,选择标准还取决于所存储的交叉位置。特别地,选择器107可以被布置为:针对与交叉位置相对应的观看者位置指示而降低朝向当前选择的路线的偏向。在当前路线上的最佳匹配当前的观看者位置指示(例如,最短距离)的位置是当前路线与另一路线交叉的交叉点时,观看者位置指示可以对应于交叉位置。

[0117] 因此,在一些实施例中,选择标准将使得在新的路线与当前路线交叉时(特别地,在观看者位置指示到当前路线上的映射是交叉点时)比新的路线与当前路线不交叉时更容易选择新的路线。

[0118] 实际上,在一些实施例中,仅在观看者位置指示对应于交叉点/位置时才可能选择新的路线。在这种实施例中,观看者位置指示可以例如根据穿过虚拟环境的所需路径或轨迹而改变。该观看者位置指示将被投影到当前路线上,并且当前选择的视频的对应的帧可以被提取并且被用于生成与观看者位置指示相对应的视图图像。然而,针对当前选择的路线上的不处于交叉点处的投影位置,选择器107可以不选择新的路线,即,在这种情况下,朝向当前选择的路线的偏向是极端的,这是因为其并不允许选择任何新的路线。然而,当观看者位置指示被改变到其中当前选择的路线上的投影位置或对应位置匹配交叉点的情况下的值时,选择标准可以允许选择在该交叉处的任何其他路线。因此,朝向当前选择的路线的极端偏向被降低并且可能被完全消除,例如,在一些实施例中,选择器107可以仅将路线选择为最靠近观看者位置指示的交叉的路线。

[0119] 因此,这种方法可以允许客户端设备201“锁定”到可用路线并且仅在交叉点处“跳转”到另一路线。虽然这样可以引起基于在空间上比(来自其他视频项的)其他可用帧更远

离观看者位置指示的帧而生成视图图像,但是在许多场景中,该方法可以提供优选的用户体验。特别地,该方法允许提高所生成的输出视图图像序列的一致性并且可以降低错误或伪影的数量和显著性。

[0120] 在许多实施例中,对其他路线的选择不限于交叉点,而是执行偏向的更为渐变的变化。在许多实施例中,朝向当前选择的路线的偏向可以取决于指示观看者位置指示与交叉位置之间的距离的距离度量。在许多实施例中,距离度量可以是合适的范数,或者在一些实施例中,可以沿着路线并且特别地沿着当前选择的路线来测量距离。因此,观看者位置指示可以被投影到当前选择的路线上(例如,路线上的到观看者位置指示具有最短距离的位置),并且可以确定沿着路线从该投影位置到交叉点的距离。

[0121] 在大部分实施例中,将减小朝向当前选择的路线的偏向以降低到交叉点的距离。在这种示例中,客户端设备201仍然将寻求“锁定”到当前路线,但是可以允许在不同路线之间“跳转”,即使不在交叉点处也是如此。该方法可以允许更灵活的方法,这样例如可以允许在适当时在路线之间进行某种“切角”。

[0122] 对用于生成针对给定的观看者位置指示的视图图像的(一个或多个)帧的选择可以是从基于不同的考虑因素所选择的视频项(对应于所选择的路线)中选择的。然而,帧通常是响应于针对相对于观看者位置指示的第一视频项的帧的路线位置数据而由图像生成器111选择的。特别地,将所选择的视频项中的帧的位置与观看者位置指示进行比较,并且可以确定距离度量。在一些实施例中,一个或多个帧可以仅被选择为在空间上最靠近观看者位置指示的帧。

[0123] 在一些实施例中,对(一个或多个)帧的选择还可以考虑时间方面。特别地,当用于生成视图图像的帧是来自当前选择的路线/视频的帧时(即,当系统停留在相同路线上并且尚未跳转到新的路线时),该选择可以使帧选择偏向于(视频项中)在时间上接近于当前选择的帧的帧。因此,在不改变正使用的视频项的情况下选择新的帧时,将新的帧选择为使得存在朝向视频项中的较长时间距离的偏向。例如,如果两个帧集合到观看者位置指示在空间上等距,则图像生成器111将接着选择最靠近先前选择的帧的帧。

[0124] 特别地,可以使用针对到视频项的当前选择的帧的时间距离减小而增大的帧选择度量。特别地,视频项中的帧编号之间的差异越小,帧选择度量可以越大。

[0125] 在许多实施例中,该方法可以提供改善的性能。特别地,在许多场景中,该方法可以提供改善的操作,其中,当前路线实际上与其自身交叉。图6提供这种路线的示例。在该示例中,该路线自身重叠,从而创建交叉点。如果帧的选择仅基于空间距离,则停留在交叉点附近且可能略微(例如,随机)移动的观看者位置指示可以引起用于生成视图图像的帧集合中的频繁移位,即,它可以在第一次到达交叉点的帧到第二次到达交叉点的帧之间切换。因此,所选择的帧将随机地并且可能频繁地在视频项中的非常不同的时间处的帧之间切换。这将引起基于不同帧的图像生成,并且这可能引起相当大的时间噪声。

[0126] 然而,通过适当偏向来抵御在时间上跳转到相同视频项的不同部分,能够避免上述不利情况并且能够生成更加一致且稳定的输出。例如,针对图6的示例,如果用户从图片的底部朝向与交叉点相对应的区移动,则当用户处于该区中时,视图图像将基于第一次到达交叉点的帧并且将如此继续进行。然而,如果用户从图片的左边接近与交叉点相对应的区时,则当用户处于该区中时,视图图像将基于第二次到达交叉点的帧,并且将如此继续进

行。

[0127] 所描述的系统基于使用与穿过虚拟空间的路线 (/路径/轨迹) 相对应的一个 (或多个) 视频项来生成合适的视图图像。这种系统中的关键考虑因素是如何选择将使用哪个视频项，并且特别是何时切换路线/视频项。

[0128] 在客户端设备201中，该方法寻求绘制来自相同视频的视图图像并且进行偏向以抵御切换到新的视频项。在类似的脉络中，当选择在视频项内将使用哪些帧时，客户端设备201进行偏向以抵御时间跳转。

[0129] 这样提供了多个优点并且基于各种考虑因素。例如：

[0130] 1、通常，使用立体相机来估计用于视图合成的视差/深度，并且使用时间深度滤波来加强时间稳定性。虽然深度值可能不正确，但是深度值从一个帧到下一帧通常是一致的，这意味着通常避免了时间伪影。

[0131] 2、与单个视频中的相邻帧相对应的虚拟相机位置通常紧挨在一起。例如，在每秒30帧下，虚拟相机可以相隔3mm至1cm。来自不同视频的虚拟相机位置可以间隔远至20cm或更大，这意味着将参考视点突然移位这样的大距离可能会产生伪影。

[0132] 3、频繁切换将增大编码的额外负担，这是因为单个视频流内的每次大跳转或视频之间的每次跳转将要求对帧内编码帧进行下载和解码。

[0133] 因此，该方法涉及表示环境，使得能够针对在环境中不限于特定姿态的用户的不同观看姿态而生成图像。

[0134] 在该领域中并且针对这种应用的主要挑战之一是如何有效表示和捕获环境而使得允许针对不同姿态生成高质量图像，同时仍然能够在不要求过多存储器的情况下存储高质量图像；以允许在没有过多复杂性和不要求过多计算资源的情况下合适地处理和生成图像的格式，能够降低客户端/服务器应用中的通信要求等。另一非常大的挑战是实现优选不要求太昂贵的或专门的电路且没有繁重过程或苛刻要求的方式的实际捕获操作。

[0135] 常规方法是仅从通常预定的观看姿态拍摄场景的个体图片 (图像和深度) 并然后存储这些图像和深度图片。通常，捕获姿态是预定的并且用户必须将相机手动移动到新的姿态，拍摄图片，将相机移动到下一姿态，拍摄另一图片等。这是非常耗时且程序繁琐的过程。替代地，能够使用大量相机，但是这样不仅昂贵而且也非常繁琐，这是因为所有相机都需要以正确姿态小心地定位。另外，要求大量图片以用于准确操作，并且因此需要以能存取的方式存储大量个体图片。

[0136] 在常规方法中，然后通过使用众所周知的视点移位算法之一从可用的和存储的图片进行视点切换来生成针对给定的观看姿态的视图图像。

[0137] 在所描述的方法中，场景不是由个体图片表示的，而是由多条路线和链接的视频项表示的。例如，该方法允许捕获过程，其中，一个或多个人能够仅在给定的环境中四处移动，从而捕获具有相关联的位置的视频。例如，一群人可以仅在场景中四处走动，从而随着这群人沿着例如随机路线走动而捕获视频。然后能够编译具有相关联的路线数据的视频项以提供具有相关联的视频项的多条路线的集合，从而提供对场景的表示。

[0138] 因此，该方法可以允许更加简化的捕获操作，并且可以例如允许通过随视频捕获大量随机路线来简单地执行捕获。因此，使用路线和视频流允许捕获更大量的信息，从而常常不需要在有利位置处进行严格受控的图像捕获，而不是要求使用个体图像进行细致且有

计划的捕获以生成对场景的充分准确表示。

[0139] 特别地,通过使场景表示以视频项(和相关联的路线信息)为基础而不是仅仅以个体图像为基础,能够大大降低对存储器/存储设备的要求,并且/或者能够存储的信息量可以更高。对视频进行编码比对个体图像进行编码更加高效。视频编码不仅利用图像内的冗余来减少所要求的数据量,而且还(并且常常更有效地)利用不同帧之间的冗余来(例如使用运动估计)大大减少针对每个帧的数据量。对于例如虚拟现实应用的许多典型场景(例如,静态场景),不同帧之间的相关性高至足以实现数据的极大减少。

[0140] 因此,与使用离散图像相比,使用具有相关联的路线的视频项的一个效应是其允许更加高效的压缩,在实践中,这种更加高效的压缩允许执行完全不同的捕获操作,其中,表示并非基于来自仔细选择的位置的图像,而是基于通常可以是相当随机的多条(并且通常是相对较大量数的)路线。该方法通常降低存储要求并且极大地减少捕获表示的复杂性和工作量。

[0141] 然而,所描述的方法不仅仅与通过多条链接的路线和视频项得到的环境表示有关,而是也与如何根据这种表示来生成图像有关。

[0142] 在所描述的方法中,对生成视图图像所基于的合适内容的确定依赖于基于路线的选择过程。选择标准用于选择路线,并且图像是根据链接的视频的至少一个帧来生成的。选择基于距观看者位置的距离,而选择标准还包括朝向当前选择的路线的偏向。

[0143] 因此,选择标准还包括,相对于其他路线,朝向当前选择的路线的偏向,而不是仅选择最近的路线。因此,系统不是仅选择最近的帧,而是可以偏向于继续基于在时间上一致选择属于相同路线的帧来创建图像。系统可以有效地“锁定”到路线上并停留在该路线上(作为用于生成图像的基础),直到偏向不再足以使用户保持锁定到该路线为止。例如,选择标准可以通过要求以下操作来偏向于当前路线:切换到不同路线要求到该其他路线的距离小于到当前路线的距离的一半。

[0144] 该方法在很大程度上取决于通过多条路线和链接的视频项来表示场景的特定方法。所描述的方法的图像(帧)和捕获位置并非是独立于和孤立于其他图像和捕获位置,而是视频项包括紧密相互关联的帧并且这些帧与路线紧密相互关联。实际上,视频项的帧具有空间关系且彼此紧密相关,即,视频项的连续帧通常具有高度相关性。类似地,路线定义与视频项直接相关的多个位置,因此,路线不仅提供关于空间中的位置的信息,而且还提供时间信息,例如,关于位置针对视频链接中的帧(其固有地是时间表示)如何改变的信息。

[0145] 发明人已经认识到这些差异并且已经认识到仅选择最接近当前视图的图像信息的方法不太适合于以下的表示和捕获数据,该表示和捕获数据不是仅仅针对仅由位置定义的孤立捕获,而且提供具有紧密相关属性(特别地,在视频项中具有时间关系)的紧密相关的捕获的序列。

[0146] 特别地,发明人已经认识到,虽然仅选择最近的图片(并且因此仅选择基于视频项的捕获中的帧)确实是可能的,但是这在计算上将是非常复杂的,并且因帧与捕获位置之间的相关性被忽略而趋向于产生错误。

[0147] 特别地,发明人已经认识到,视频项的帧并非作为个体帧而被编码,而是参考视频项中的其他帧而被编码。通常,视频项包括一些帧内编码帧,其具有通常在实质上更大数量的、参考其他帧而被编码的帧间编码帧,并且具有比帧内编码帧更高的效率。

[0148] 特别地,这使得难以提取个体图像,这是因为提取个体图像将要求提取两个相关的帧内编码帧,对这两个相关的帧内编码帧进行解码,提取预测所需帧所根据的任何预测帧,使用帧内编码帧对预测帧进行解码,提取所需帧,对所需帧进行解码以及最后进行所有这些操作的组合。这具有非常高的计算负担。此外,编码错误、噪声和伪影将不可避免,并且这些问题会降低所得图像的质量。

[0149] 发明人还已经认识到,通过不仅考虑位置而且也考虑观看者位置的历史,能够适应和利用视频项中存在的时间和帧间相关性。这通过基于路线而非仅考虑个体位置做出选择并且通过使该选择进一步偏向于“固守”于相同路线上来完成。

[0150] 这特别引起对于视频编码的属性更加高效的方法,并且这特别解决了根据视频项生成图像的问题。特别地,在实践中已经发现,这极大地减小了计算负担。基于路线的选择和从相同视频项/路线选择(通常为连续的)帧的一致性增大引起帧内编码帧的相当多的重复使用,帧内编码帧针对基于其来编码的所有帧间编码帧仅要求被解码一次。类似地,如果已经使用中间帧来生成先前图像,则针对另一帧生成这种中间预测帧的要求将不会引入额外负担。

[0151] 另外,不可避免的编码/解码错误和伪影在该方法中针对连续图像将趋向于更加一致。例如,针对帧内编码帧而发生的编码/解码伪影在基于来自相同视频项/路线的使用该帧内编码帧编码的帧所生成的所有图像中将被等同地表示。这对于可能不太明显的错误或伪影通常可以是优选的,然而错误或伪影在连续生成的图像之间不太一致。

[0152] 因此,用于对路线/视频项的基于偏向路线的选择的方法提供实质改善的性能,并且特别允许降低计算复杂性和资源使用情况,降低存储要求,便于视图图像生成,改善图像质量。该方法也可以降低例如对基于客户端服务器的应用的通信要求。

[0153] 在一些实施例中,路线数据可以包括通过使交叉作为节点并且使路线作为边来表示的图表。图7提供了与房屋相对应的虚拟环境的示例。

[0154] 在该示例中,每个节点因此表示路线之间的交叉,包括T型接合(n2)。在这种图表中,交叉还可以反映一条路线结束并且另一路线开始的位置。在该示例中,每条路线因此由两个节点来定界,其中,节点针对多条路线通常是共用的。该方法特别适合于其中路线和视频项通过专用的和有计划的捕获而生成的或其中捕获路线和视频项已经被划分成较小的路线和视频项的场景。该图表可以例如被编码为二元部分或文本(例如,xml)并且在应用初始化时可以(至少部分地)从服务器203被传输到客户端设备201。

[0155] 图表的使用在许多实施例中可以是高效的并且可以提供高效操作。例如,图表结构允许在给定的路线的结尾处进行确定,多个(0、1、……)个定义的其他路线可以开始,并且因此可以使对对应的视频的检索和解码提前初始化,从而减小可能的迟滞和中断。

[0156] 在一些实施例中,路线数据还可以包括指示该路线集合中的各条路线之间的空间关系的关系数据。然后,关系数据可以用于对当前的观看者位置指示的路线的选择。

[0157] 作为特定的示例,关系数据可以针对每条路线提供对该路线与其他路线相交的指示。在一些实施例中,选择器107然后可以针对当前选择的路线检查关系数据以查看当前选择的路线与哪些其他路线相交且具有交叉。然后,选择器107在确定是否要选择不同路线时可以接着仅考虑例如这些特定路线。

[0158] 在一些实施例中,关系数据可以提供对哪些路线彼此形成接合的指示。这种数据

可以由选择器107用于选择用于视图合成的适当视频和帧。例如,当接近T型接合时,能够基于位置数据提前预测用户在该接合处是倾向于右转还是左转。

[0159] 在一些实施例中,关系数据可以提供对一些路线与针对相同路径的相反移动方向相对应的指示。当观看者在其路径上转向(例如180度)时,选择器107可以使用这种数据来切换方向。

[0160] 在一些实施例中,关系数据可以提供对彼此基本平行的路线的指示。选择器107可以使用这种数据来决定切换到更好地表示用户前进的方向的另一路径。

[0161] 例如,关系可以以图表的形式来提供,该图表具有作为节点的路线以及指示对应的两个节点之间的关系的边。图8图示了这种图表的示例。

[0162] 如前所述,在不减损本发明的情况下可以使用视频项的任何合适的视频编码。在许多实施例中,视频项可以使用不考虑任何其他帧而编码的帧内编码帧和相对于至少一个其他帧而编码的预测帧两者来编码。预测帧可以相对于一个或多个帧内编码帧和/或一个或多个预测帧来编码。预测帧可以是仅相对于一个或多个较早帧(即,视频项中的时间上较早的帧)来编码的帧,或者可以例如是相对于一个或多个较早帧且相对于一个或多个较晚帧(即,相对于在当前帧之前和之后的帧)来编码的双向预测帧。视频编码可以是H.264或H.265视频编码。

[0163] 在许多实施例中,对于编码过程而言,仅使用I帧(帧内编码帧图片)以及B帧(双向预测帧)可以是特别有利的。在许多实施例中,这可以引起视图图像的低延迟生成。

[0164] 因此,视频项可以是编码视频项,其包括帧内编码帧和预测/帧间编码帧。在一些实施例中,预测帧可以仅是双向预测/编码帧。因此,视图图像的生成可以包括图像生成器111对用于生成视图图像的(一个或多个)帧进行解码。

[0165] 特别地,所描述的方法寻求基于多条路线的集合来提供对虚拟环境的视图的高效捕获、存储和回放。对视频项而非个体图像的使用降低了存储要求、通信带宽要求等。这得以实现是因为视频编码趋向于比对单独图像进行编码的情况更加高效。特别地,对预测帧的使用可以非常显著地减小表示帧所要求的数据量。然而,对视频捕获和编码的使用对于表示其中用户可以自由移动的虚拟环境是违反直觉的。例如,在这种情况下,用户可以在与视频被捕获的方向的相反方向上移动,从而不仅不遵循视频内容的空间和时间关系,甚至与视频内容的空间和时间关系正好相反。然而,发明人已经认识到,该方法可以提供非常高效的操作。

[0166] 所描述的具有朝向停留在相同路线上的偏向的基于路线的选择的方法提供改善的操作。特别地,发明人已经认识到,当用户继续被链接到相同路线而非切换到另一路线时(即,当路线的选择从一条路线改变为另一路线时),由于对与用户视点相对应的帧进行帧间编码(预测),因此这种偏向(除了得到时间上更稳定的视图合成质量以外)会引起带宽要求的极大减少(用于检索/传达帧的数据速率减小)。

[0167] 在许多实施例中,该选择在选择路线时可以将视频编码考虑在内。特别地,在许多实施例中,选择器107可以被布置为:针对与帧内编码帧相对应的观看者位置指示,降低朝向当前选择的路线的偏向。选择器107可以被布置为:相对于与帧间编码(预测)帧相对应的观看者位置指示,针对与帧内编码帧相对应的观看者位置指示,降低朝向当前选择的路线的偏向。因此,当观看者位置对应于帧间编码/预测帧时,选择器更可能选择新的路线而非

继续留在相同路线上。

[0168] 在一些实施例中,所考虑的帧(其影响偏向)可以是当前/先前选择的路线的帧,或者可以是新的候选路线的帧。因此,该帧可以是将用于生成针对当前/先前选择的路线(如果该路线被选择的话)的视图图像的帧,或者可以是将用于生成针对候选路线(如果该路线被选择的话)的视图图像的帧。因此,该帧可以是当前选择的路线的或第一路线(所选择的路线)的帧。在许多实施例中,针对作为帧内编码帧的目标/候选帧和针对作为帧内编码帧的当前/先前选择的路线帧两者,可以降低偏向。

[0169] 如果选择给定的路线,则针对观看者位置指示的(一个或多个)对应的帧可以是该给定的路线的将用于生成针对该观看者位置指示的视图图像的(一个或多个)帧。对应的帧通常可以是给定的路线中最靠近由观看者位置指示(通常使用也用于选择路线的距离度量)所指示的姿态/位置的帧。

[0170] 在许多实施例中,在从一条路线切换到另一路线之后,第一帧可以选择为帧内编码帧。当切换到新的路线时,图像生成器111因此可以接着首先选择帧内编码帧并且使用该帧内编码帧来生成视图图像。

[0171] 在许多实施例中,这些方法可以用于在不同路线之间建立链路。例如,图9图示了具有与帧内编码帧相关联的多个位置(由小圆圈指示)的两条路线901、903。在这种实施例中,针对帧内编码帧可以降低对停留在相同路线上的偏向,并且针对新的路线所选择的帧可以对应于另一帧内编码帧。特别地,针对发生在两个帧内编码帧之间的路线切换可以(通常极大地)降低朝向当前路线的偏向。因此,如由图9中的箭头所指示的,“跳转”将趋向于发生在帧内编码帧之间。

[0172] 在实践中,如前所述,这可以通过在表示路线的图表中包括这种考虑因素而实现。实际上,在这种图表中,可以接近相交/交叉来创建节点。此外,节点可以选择为对应于帧内编码帧,其中,帧内编码帧之间的区段对应于帧间编码帧。然后可以提供链接这种节点的另外的数据,这些另外的数据例如指示节点之间的关系(例如,“相交”或“返回路径”)。

[0173] 如后文所述,系统可以(特别地)(例如通过对视频项进行转码)在路线之间的交叉处生成帧内编码帧。

[0174] 在许多实施例中,还可以优选在给定的路线的I帧之间使用相对较大的量B帧。该方法非常高效,这是因为B帧(其是在补偿预测之后的剩余帧)通常比I帧压缩得更多。由于静态场景中的相邻相机位置之间的高相干性,因此对大量B帧的使用通常是可接受的。在I帧之间具有大量B帧对于视频回放也具有吸引力,这是因为针对给定的用户移动要求对较少I帧进行解码。

[0175] 然而,为了减轻在因选择新的路线而改变视频项时使用大量B帧的潜在缺点,该方法可以使用对I帧的选择性分配。

[0176] 在许多实施例中,可以生成I帧并且将I帧包括在针对交叉的视频项中。因此,与交叉相关联的帧可以被编码为I帧。在许多实施例中,这可以提供特别有吸引力的性能,这是因为其可以便于路线切换。实际上,使交叉帧被编码为I帧将趋向于引起在交叉处执行切换,从而提供以I帧开始和/或结束的相关联的视频项,并且因此避免为了编码B帧而必须编码额外的I帧。

[0177] 实际上,如果交叉帧被编码为I帧,则相同的I帧可能被用于先前选择的路线和新

选择的路线两者。因此,来自先前选择的路线的I帧可以与用于新选择的路线的I帧一起使用,以便对新选择的路线的B帧进行解码。在一些实施例中,这可以使用先前路线的I帧来实现,该I帧并不精确地处于交叉点处,但是可能靠近交叉点并因此提供合理的估计。然而,在优选方法中,先前选择的路线的I帧可以是交叉处的帧,而实际上,I帧甚至可以是用于与先前选择的路线相关联的视频项和用于与新选择的路线相关联的视频项的共用帧。

[0178] 在许多实施例中,可以实施B帧和I帧的有利的组织和结构。例如,服务器203可以首先提供合适的I帧。即使不直接与当前的观看者位置指示相关联的,这种I帧也可以由客户端设备201进行解码。然而,当图像生成器111选择适当的B帧以用于生成视图图像时,解码的I帧然后可以提供允许对适当的B帧快速解码的参考帧。

[0179] 实际上,在许多实施例中,可以使用层级式视频编码方法,其中,B帧被布置在层级式结构中,使得一些B帧还可以被用作用于其他B帧的参考帧。在这种情况下,适当的B帧还可以被解码以准备用作用于其他B帧的参考帧。

[0180] 因此,在一些实施例中,检索器109因此可以包括解码器,该解码器对多个I帧并且可能对B帧进行解码,以准备将这些帧用于对已经相对于这些帧进行编码的其他B帧进行解码。

[0181] 因此,在一些实施例中,可以采取B帧的层级式排序,从而允许从视频中随机存取和双向回放,同时通过避免大部分I帧而维持高编码效率。例如,在任何时间,解码器都可以具有至少两个I帧以及准备好(解密的和解码的)用作对B帧进行解码的参考的多个(0个或多个)B帧。

[0182] 诸如H.264/MPEG-4AVC和HEVC的较新的视频编码方案可以支持这种方法。

[0183] 应当理解,虽然以上说明集中于如图2所示的客户端服务器方法,但是该构思不限于此,而是同样合适于许多其他功能性布置和分布。例如,在一些实施例中,存储视频项的视频存储部可以是本地存储部,并且检索器109可以从本地存储部检索视频。在一些实施例中,该方法可以用于存储视频项和路线数据两者且响应于接收的观看者位置指示而生成视图图像的单个独立设备。

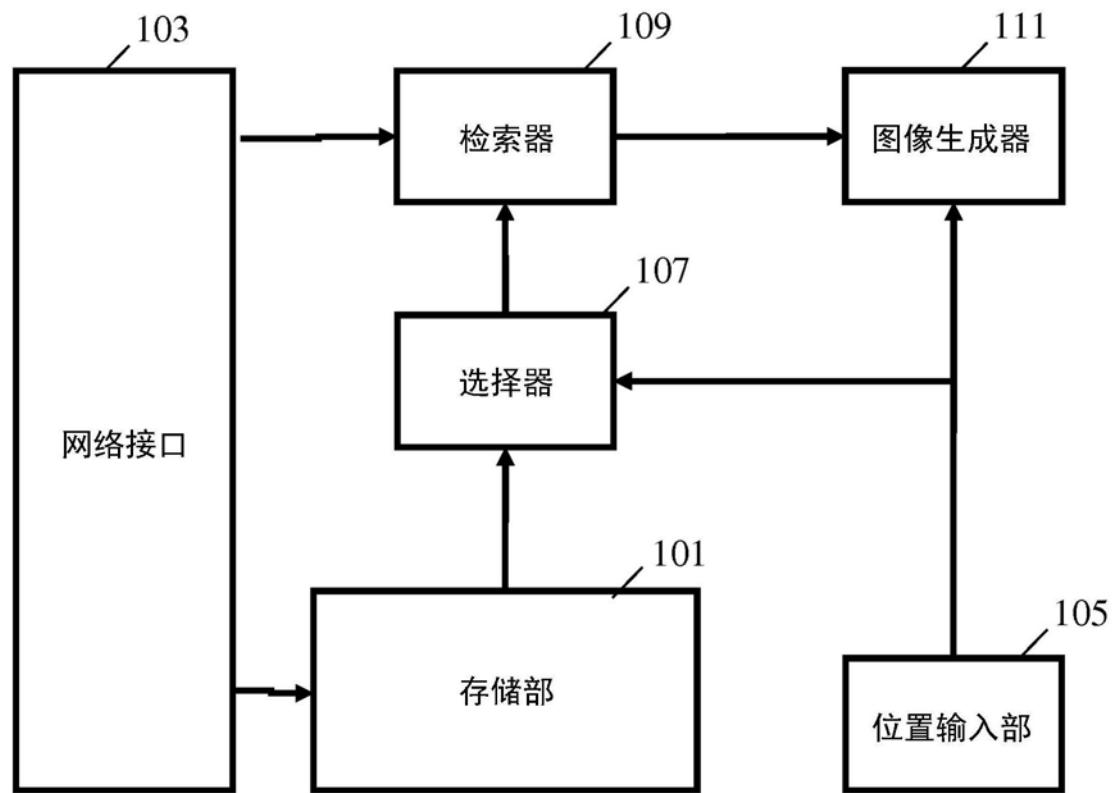
[0184] 应当理解,为了清楚起见,以上描述已经参考了不同的功能电路、单元和处理器来描述本发明的实施例。然而,应当理解,可以在不脱离本发明的情况下使用不同的功能电路、单元或处理器之间的任何合适的功能分布。例如,被图示为由单独的单元、处理器或控制器执行的功能可以由相同的处理器或控制器来执行。因此,对特定功能单元或电路的引用仅被视为对用于提供所描述的功能的合适单元的引用,而不是指示严格的逻辑或物理结构或组织。

[0185] 本发明能够以包括硬件、软件、固件或这些项目的任意组合的任意适合形式来实施。本发明也可以任选地被至少部分地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部件可以以任何合适的方式被物理地、功能地和逻辑地实施。实际上,功能可以被实施在单个单元中,在多个单元中或作为其他功能单元的部分。正因如此,本发明可以被实施在单个单元中,或者可以被物理地或功能地分布在不同的单元、电路和处理器之间。

[0186] 虽然已经结合一些实施例描述了本发明,但是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定。另外,虽然可能看起来结合特定的实

施例描述了特征,但是本领域技术人员将认识到,也可以根据本发明组合所描述的实施例的各种特征。在权利要求中,术语“包括”并不排除存在其他元件或步骤的存在。

[0187] 此外,虽然多个模块、元件、电路或方法步骤被单独列出,但是多个模块、元件、电路或方法步骤也可以通过例如单个电路、单元或处理器来实施。另外,虽然各个特征可以被包括在不同的权利要求中,但是这些特征也可以被有利地组合,并且被包括在不同的权利要求中并不意味着特征的组合是不可行的和/或不利的。在一种类型的权利要求中包括特征并不意味着对该类型的限制,而是指示该特征在适当时同样适用于其他权利要求类型。此外,权利要求中的特征的顺序并不意味着该特征必须按任何特定顺序进行工作,并且特别地,方法权利要求中的各个步骤的顺序并不意味着必须以该顺序执行这些步骤。而是,可以以任何合适的顺序来执行这些步骤。另外,单数引用并不排除多个。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为使示例清楚,而不应被解释为以任何方式限制权利要求的范围。



201

图1

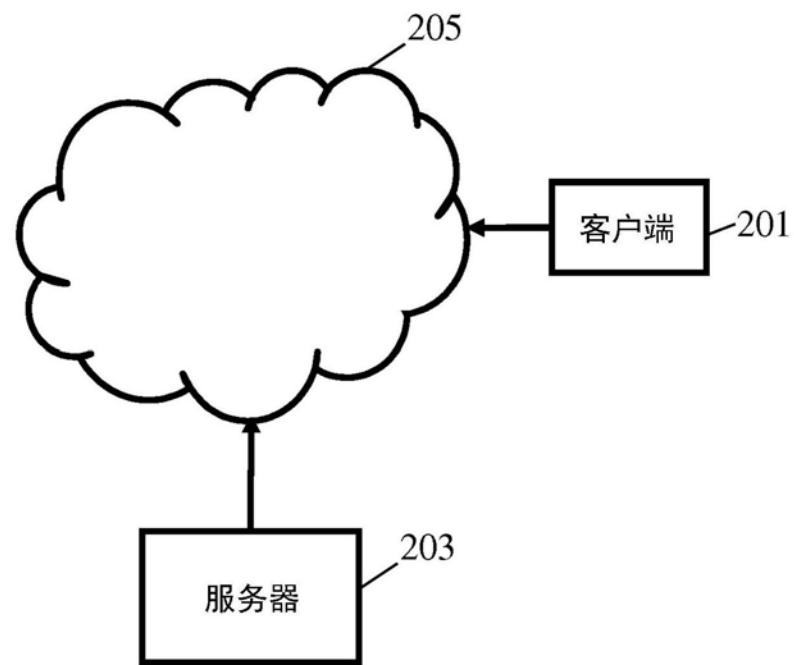


图2

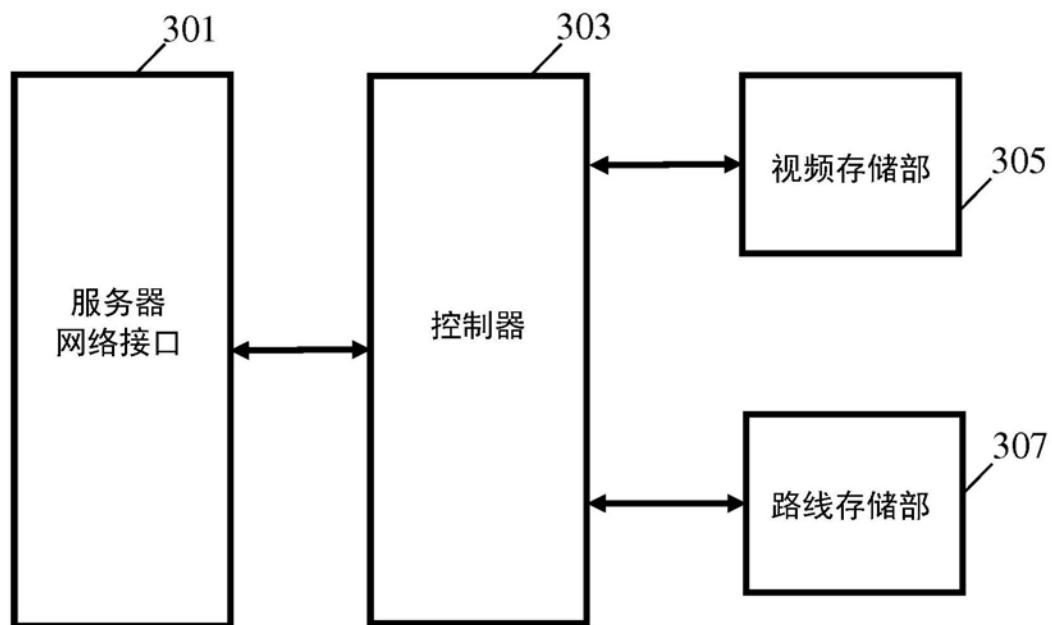
101

图3

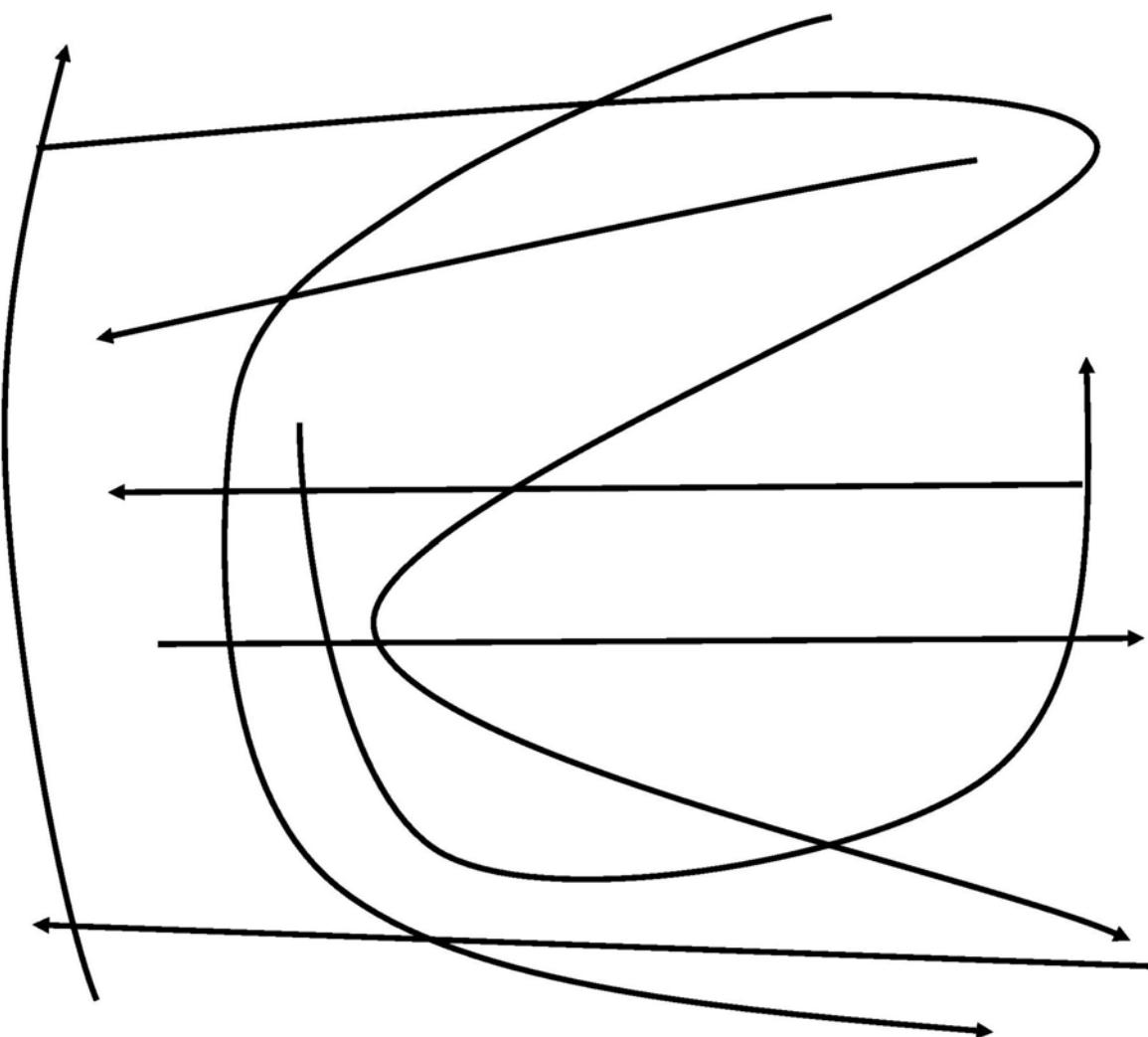


图4

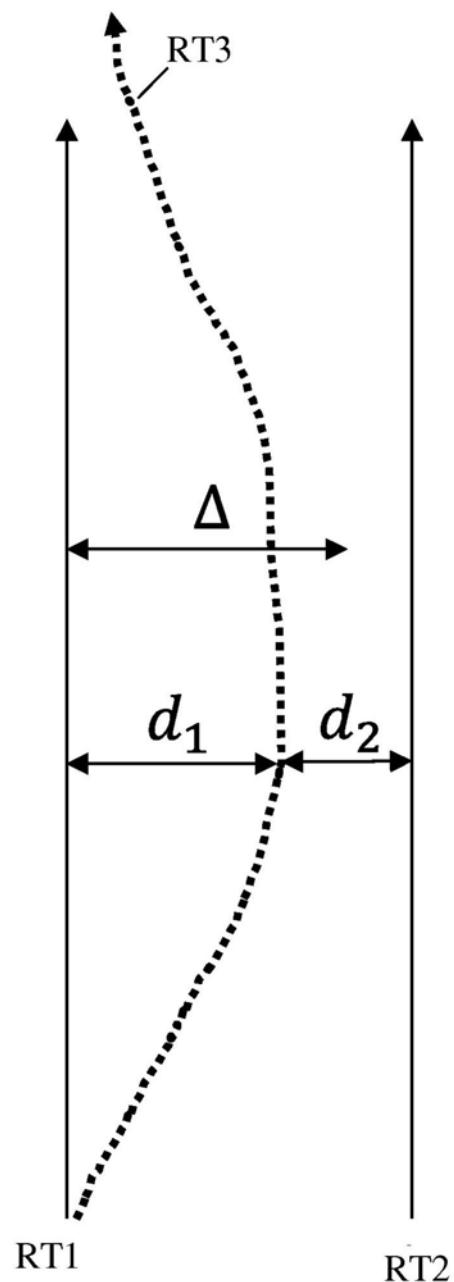


图5

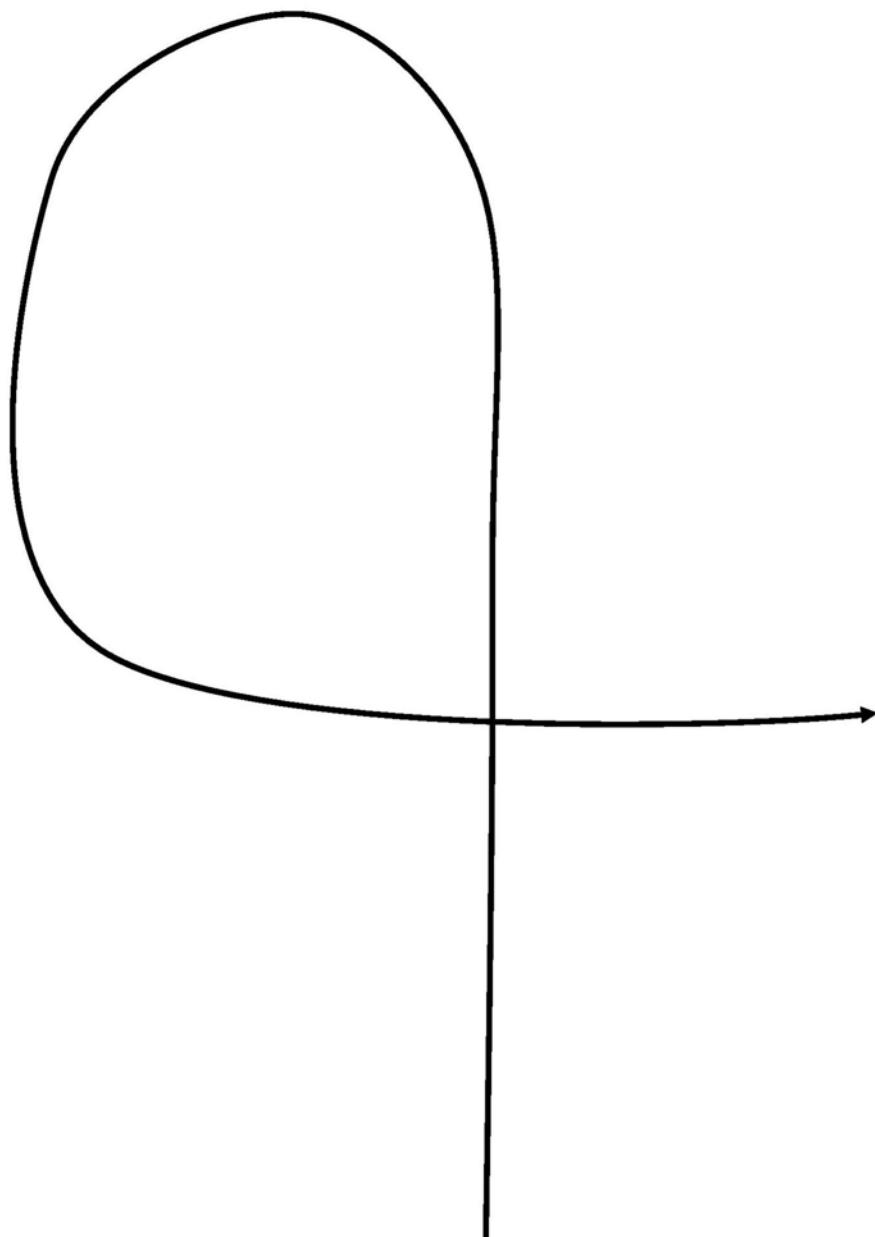
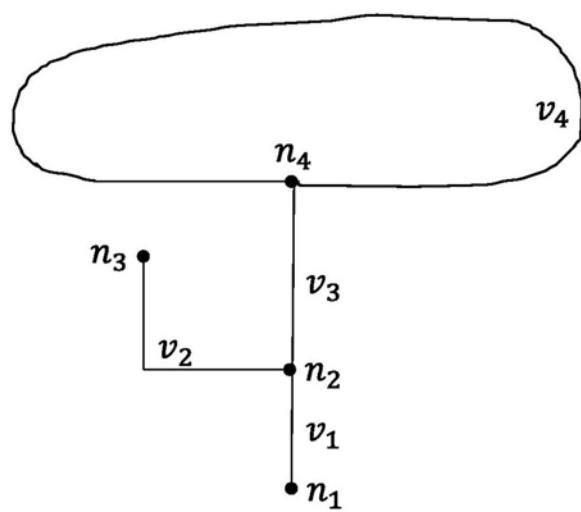


图6

路线图表



每个视频的元数据

视频	开始节点	结束节点	环境
v_1	n_1	n_2	过道
v_2	n_2	n_3	楼梯
v_3	n_2	n_4	过道
v_4	n_4	n_4	起居室

每个帧的元数据
针对每个帧k的 R_k 、 t_k

图7

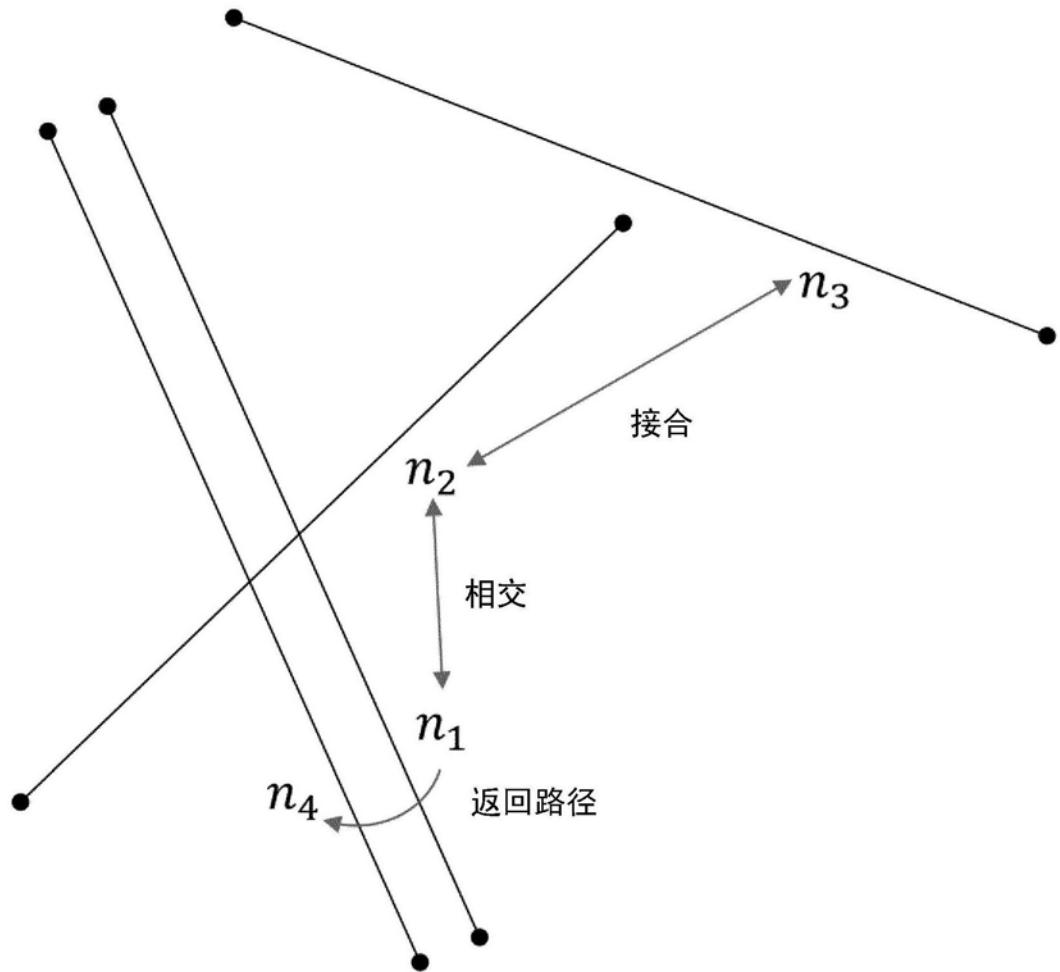


图8