

基于场景重建生成视频的方法、存储介质和系统

交叉援引

本公开要求于 2023 年 04 月 25 日提交中国专利局、申请号 202310467335.3、申请名称“基于场景重建生成视频的方法、存储介质和系统”的中国专利申请的优先权，其全部内容通过引用结合在本公开中。

技术领域

本公开涉及计算机技术、人工智能技术以及虚拟现实/增强现实技术领域，具体而言，涉及一种基于场景重建生成视频的方法、存储介质和系统。

背景技术

近年来，基于图像数据（如航拍数据、遥感数据等）进行三维虚拟场景重建的技术构思在倾斜摄影、地形测绘、文旅宣传、景区/园区/建筑重建和虚拟现实漫游等领域具有越来越大的应用价值。然而，现有技术提供的基于图像数据的场景重建方法通常采用显示网格（Mesh）模型（也称显示建模）的重建构思，这种方法得到的网格模型存在较多的网格缺陷（如几何错误、破洞等问题），难以提供良好的自由视角场景漫游的效果，即使经过人工优化基于该网格模型渲染的场景漫游视频的真实感仍然较差。

针对上述的问题，目前尚未提出有效的解决方案。

发明内容

本公开实施例提供了一种基于场景重建生成视频的方法、存储介质和系统，以至少解决由于现有技术采用显示建模方式进行场景重建导致的建模结果缺陷多、场景漫游视频的真实感差的技术问题。

根据本公开实施例的一个方面，提供了一种基于场景重建生成视频的方法，包括：获取待重建图像数据，其中，待重建图像数据被设置为记录待重建的预设三维场景的影像；对待重建图像数据进行稀疏重建，得到重建结果；基于待重建图像数据和重建结果进行隐式建模，得到建模结果；利用建模结果生成目标视频，其中，目标视频被设置为对预设三维场景进行场景漫游。

根据本公开实施例的另一方面，还提供了一种基于场景重建生成视频的方法，包括：响应作用于操作界面上的输入指令，在操作界面上显示待重建图像数据，其中，待重建图像数据被设置为记录待重建的预设三维场景的影像；响应作用于操作界面上的重建指令，在操作界面上显示目标视频，其中，目标视频被设置为对预设三维场景进行场景漫游，目标视频利用建模结果生成，建模结果通过对待重建图像数据和重建结果进行隐式建模得到，重建结果通过对待重建图像数据进行稀疏重建得到。

根据本公开实施例的另一方面，还提供了一种基于场景重建生成视频的方法，包括：在虚拟现实 VR 设备或增强现实 AR 设备的呈现画面上展示初始默认画面；驱动 VR 设备或 AR 设备从展示初始默认画面跳转至展示目标视频，其中，目标视频被设置为对待重建的预设三维场景进行场景漫游，目标视频利用建模结果生成，建模结果通过对待重建图像数据和重建结果进行隐式建模得到，重建结果通过对待重建图像数据进行稀疏重建得到，待重建图像数据被设置为记录预设三维场景的影像。

根据本公开实施例的另一方面，还提供了一种计算机可读存储介质，上述计算机可读存储介质包括存储的程序，其中，在上述程序运行时控制上述计算机可读存储介质所在设备执行任意一项上述的基于场景重建生成视频的方法。

根据本公开实施例的另一方面，还提供了一种基于场景重建生成视频的系统，包括：处理器；以及存储器，与上述处理器连接，被设置为为上述处理器提供处理以下处理步骤的指令：获取待重建图像数据，其中，待重建图像数据被设置为记录待重建的预设三维场景的影像；对待重建图像数据进行稀疏重建，得到重建结果；基于待重建图像数据和重建结果进行隐式建模，得到建模结果；利用建模结果生成目标视频，其中，目标视频被设置为对预设三维场景进行场景漫游。

在本公开实施例中，通过获取待重建图像数据，其中，待重建图像数据被设置为记录待重建的预设三维场景的影像；进一步对待重建图像数据进行稀疏重建，得到重建结果；采用基于待重建图像数据和重建结果进行隐式建模的方式得到建模结果；进而利用建模结果生成目标视频，其中，目标视频被设置为对预设三维场景进行场景漫游。由此，本公开实施例达到了通过稀疏重建和隐式建模的方式基于待重建图像数据进行场景建模以渲染得到场景漫游视频的目的，从而实现了提升场景建模结果的准确度和场景漫游视频的真实感的技术效果，进而解决了由于现有技术采用显示建模方式进行场景重建导致的建模结果缺陷多、场景漫游视频的真实感差的技术问题。

容易注意到的是，上面的通用描述和后面的详细描述是为了对本公开进行举例和解释，并不构成对本公开的限定。

附图说明

此处所说明的附图用来提供对本公开的进一步理解，构成本公开的一部分，本公开的示意性实施例及其说明被设置为解释本公开，并不构成对本公开的不当限定。在附图中：

图1示出了一种被设置为实现基于场景重建生成视频的方法的计算机终端（或移动设备）的硬件结构框图；

图2是根据本公开实施例1的一种基于场景重建生成视频的方法的虚拟现实设备的硬件环境的示意图；

图3是根据本公开实施例1的一种基于场景重建生成视频的方法的计算环境的结构框图；

图4是根据本公开实施例1的一种基于场景重建生成视频的方法的流程图；

图5是根据本公开实施例1的一种可选的场景重建过程的示意图；

图6是根据本公开实施例1的一种可选的隐式建模过程的示意图；

图7是根据本公开实施例2的一种基于场景重建生成视频的方法的流程图；

图8是根据本公开实施例3的一种基于场景重建生成视频的方法的流程图；

图9是根据本公开实施例4的一种基于场景重建生成视频的装置的结构示意图；

图10是根据本公开实施例4的一种可选的基于场景重建生成视频的装置的结构示意图；

图11是根据本公开实施例4的另一种可选的基于场景重建生成视频的装置的结构示意图；

图12是根据本公开实施例5的一种计算机终端的结构框图。

具体实施方式

为了使本技术领域的人员更好地理解本公开方案，下面将结合本公开实施例中的附图，对本公开实施例中的技术方案进行清楚、完整地描述，显然，所描述的实施例是本公开一部分的实施例，而不是全部的实施例。基于本公开中的实施例，本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例，都应当属于本公开保护的范围。

需要说明的是，本公开的说明书和权利要求书及上述附图中的术语“第一”、“第二”等是用于区别类似的对象，而不必用于描述特定的顺序或先后次序。应该理解这样使用的数据在适当情况下可以互换，以便这里描述的本公开的实施例能够以除了在

这里图示或描述的那些以外的顺序实施。此外，术语“包括”和“具有”以及他们的任何变形，意图在于覆盖不排他的包含，例如，包含了一系列步骤或单元的过程、方法、系统、产品或设备不必限于清楚地列出的那些步骤或单元，而是可包括没有清楚地列出的或对于这些过程、方法、产品或设备固有的其它步骤或单元。

首先，在对本公开实施例进行描述的过程中出现的部分名词或术语适用于如下解释：

神经辐射场（Neural Radiance Fields, NeRF）：是指利用深度学习技术将图像和三维空间的形变进行整合，以生成高质量的三维场景的一种方法。在本公开中，NeRF指基于多张图像和隐式网格模型对三维场景进行重建的方法。

运动恢复机构（Structure from Motion, SfM）：是指一种基于多张二维图像序列估计三维场景结构的技术。

多层感知机（Multi-layer Perceptron, MLP）：是指一种包括多层全连接层的神经网络。

近年来，基于图像数据（如航拍数据、遥感数据等）进行三维虚拟场景重建的技术构思在倾斜摄影、地形测绘、文旅宣传、景区/园区/建筑重建和虚拟现实漫游等领域具有越来越大的应用价值。然而，现有技术提供的基于图像数据的基于场景重建生成视频的方法通常采用显示网格（Mesh）模型的重建构思，这种方法得到的网格模型存在较多的网格缺陷（如几何错误、破洞等问题），难以提供良好的自由视角场景漫游的效果，即使经过人工优化基于该网格模型渲染的场景漫游视频的真实感仍然较差。

在上述现有技术的基础上，相关技术领域的技术人员提出了 NeRF 算法，在进行场景重建的过程中使用 NeRF 算法，以隐式建模的方式得到建模结果，进而基于该建模结果渲染得到场景图像。然而，在本公开之前的现有 NeRF 算法针对物体模型重建或小尺寸场景重建，难以适用于大规模场景（如航拍数据对应的园区、城市等场景，或者遥感数据对应的地形、地块等场景）的重建任务；并且，基于现有 NeRF 算法所得到的重建结果难以渲染得到真实感较强的自由视角场景漫游视频。

实施例 1

根据本公开实施例，还提供了一种基于场景重建生成视频的方法实施例，需要说明的是，在附图的流程图示出的步骤可以在诸如一组计算机可执行指令的计算机系统中执行，并且，虽然在流程图中示出了逻辑顺序，但是在某些情况下，可以以不同于此处的顺序执行所示出或描述的步骤。

本公开实施例一所提供的方法实施例可以在移动终端、计算机终端或者类似的运算装置中执行。图 1 示出了一种被设置为实现基于场景重建生成视频的方法的计算机终端（或移动设备）的硬件结构框图。如图 1 所示，计算机终端 10（或移动设备 10）可以包括一个或多个（图中采用 102a, 102b, ……，102n 来示出）处理器 102（处理器 102 可以包括但不限于微处理器（Microcontroller Unit, MCU）或可编程逻辑器件（Field Programmable Gate Array, FPGA）等的处理装置）、被设置为存储数据的存储器 104、以及被设置为通信功能的传输装置 106。除此以外，计算机终端 10 还可以包括：显示器、输入/输出接口（I/O 接口）、通用串行总线（Universal Serial Bus, USB）端口（可以作为 BUS 总线的端口中的一个端口被包括）、网络接口、光标控制设备（如鼠标、触控板等）、键盘、电源和/或相机。

本领域普通技术人员可以理解，图 1 所示的结构为示意，其并不对上述电子装置的结构造成限定。例如，计算机终端 10 还可包括比图 1 中所示更多或者更少的组件，或者具有与图 1 所示不同的配置。

应当注意到的是上述一个或多个处理器 102 和/或其他数据处理电路在本文中通常可以被称为“数据处理电路”。该数据处理电路可以全部或部分的体现为软件、硬件、固件或其他任意组合。此外，数据处理电路可为单个独立的处理模块，或全部或

部分的结合到计算机终端 10 (或移动设备) 中的其他元件中的任意一个内。如本公开实施例中所涉及到的, 该数据处理电路作为一种处理器控制 (例如与接口连接的可变电阻终端路径的选择)。

存储器 104 可用于存储应用程序的软件程序以及模块, 如本公开实施例中的基于场景重建生成视频的方法对应的程序指令/数据存储装置, 处理器 102 通过运行存储在存储器 104 内的软件程序以及模块, 从而执行各种功能应用以及数据处理, 即实现上述的基于场景重建生成视频的方法。存储器 104 可包括高速随机存储器, 还可包括非易失性存储器, 如一个或者多个磁性存储装置、闪存、或者其他非易失性固态存储器。在一些实例中, 存储器 104 可进一步包括相对于处理器 102 远程设置的存储器, 这些远程存储器可以通过网络连接至计算机终端 10。上述网络的实例包括但不限于互联网、企业内部网、局域网、移动通信网及其组合。

传输装置 106 被设置为经由一个网络接口与网络连接以接收或者发送数据。上述的网络具体实例可包括计算机终端 10 的通信供应商提供的有线和/或无线网络。在一个实例中, 传输装置 106 包括一个网络适配器 (Network Interface Controller, NIC), 其可通过基站与其他网络设备相连从而可与互联网进行通讯。在一个实例中, 传输装置 106 可以为射频 (Radio Frequency, RF) 模块, 其被设置为通过无线方式与互联网进行通讯。

如图 1 所示的显示器可以例如触摸屏式的液晶显示器 (LCD), 该液晶显示器可使得用户能够与计算机终端 10 (或移动设备) 的用户界面进行交互。

此处需要说明的是, 在一些可选实施例中, 上述图 1 所示的计算机设备 (或移动设备) 可以包括硬件元件 (包括电路)、软件元件 (包括存储在计算机可读介质上的计算机代码)、或硬件元件和软件元件两者的结合。应当指出的是, 图 1 为特定具体实例的一个实例, 并且旨在示出可存在于上述计算机设备 (或移动设备) 中的部件的类型。

此处需要说明的是, 在一些实施例中, 上述图 1 所示的计算机设备 (或移动设备) 具有触摸显示器 (也被称为“触摸屏”或“触摸显示屏”)。在一些实施例中, 上述图 1 所示的计算机设备 (或移动设备) 具有图像用户界面 (GUI), 用户可以通过触摸敏感表面上的手指接触和/或手势来与 GUI 进行人机交互, 此处的人机交互功能可选的包括如下交互: 创建网页、绘图、文字处理、制作电子文档、游戏、视频会议、即时通信、收发电子邮件、通话界面、播放数字视频、播放数字音乐和/或网络浏览等、被设置为执行上述人机交互功能的可执行指令被配置/存储在一个或多个处理器可执行的计算机程序产品或可读存储介质中。

本公开实施例一所提供的方法实施例可以在虚拟现实 (Virtual Reality, VR) 设备或增强现实 (Augmented Reality, AR) 设备中执行。以 VR 设备为例, 图 2 是根据本公开实施例 1 的一种基于场景重建生成视频的方法的虚拟现实设备的硬件环境的示意图。如图 2 所示, 虚拟现实设备 204 与终端 206 相连接, 终端 206 与服务器 202 通过网络进行连接, 上述虚拟现实设备 204 并不限定于: 虚拟现实头盔、虚拟现实眼镜、虚拟现实一体机等, 上述终端 204 并不限定于 PC、手机、平板电脑等, 服务器 202 可以为媒体文件运营商对应的服务器, 上述网络包括但不限于: 广域网、城域网或局域网。

可选地, 本公开实施例的虚拟现实设备 204 包括: 存储器、处理器和传输装置 (图 2 中未示出)。存储器被设置为存储应用程序, 该应用程序可以被设置为执行本公开实施例提供的基于场景重建生成视频的方法。处理器可以调用并执行存储器中存储的应用程序, 以通过本公开实施例提供的基于场景重建生成视频的方法得到待识别部位对应的目标网格点。

可选地, 本公开实施例的终端 206 也可以被设置为执行基于场景重建生成视频的

方法, 并通过 VR 设备或 AR 设备的显示画面展示待识别部位对应的目标网格点, 或者通过 VR 设备或 AR 设备关联的图像输出设备显示执行基于场景重建生成视频的方法生成的待识别部位对应的目标网格点。

图 2 示出的硬件结构框图, 不仅可以作为上述 AR/VR 设备 (或移动设备) 的硬件环境的示例性框图, 还可以作为上述服务器的示例性框图。

图 3 是根据本公开实施例 1 的一种基于场景重建生成视频的方法的计算环境的结构框图, 图 3 以框图示出了使用上述图 2 所示的 VR 设备 (或移动设备) 作为计算环境 301 中计算节点的一种实施例。

仍然如图 3 所示, 计算环境 301 包括运行在分布式网络上的多个 (图 3 中采用 310-1, 310-2, ..., 来示出) 计算节点 (如服务器)。每个计算节点都包含本地处理和内存资源, 终端用户 302 可以在计算环境 301 中远程运行应用程序或存储数据。应用程序可以作为计算环境 301 中的多个服务 320-1 (代表服务 A), 320-2 (代表服务 D), 320-3 (代表服务 E) 和 320-4 (代表服务 H) 进行提供。

仍然如图 3 所示, 终端用户 302 可以通过客户端上的 web 浏览器或其他软件应用程序提供和访问服务, 在一些实施例中, 可以将终端用户 302 的供应和/或请求提供给入口网关 330。入口网关 330 可以包括一个相应的代理来处理针对服务 320 (计算环境 301 中提供的一个或多个服务) 的供应和/或请求。

仍然如图 3 所示, 服务 320 是根据计算环境 301 支持的各种虚拟化技术来提供或部署的。在一些实施例中, 可以根据基于虚拟机 (Virtual Machine, VM) 的虚拟化、基于容器的虚拟化和/或类似的方式提供服务 320。基于虚拟机的虚拟化可以通过初始化虚拟机来模拟真实的计算机, 在不直接接触任何实际硬件资源的情况下执行程序 and 应用程序。在虚拟机虚拟化机器的同时, 根据基于容器的虚拟化, 可以启动容器来虚拟化整个操作系统 (Operating System, OS), 以便多个工作负载可以在单个操作系统实例上运行。

在基于容器虚拟化的一个实施例中, 服务 320 的若干容器可以被组装成一个 POD (例如, Kubernetes POD)。举例来说, 如图 3 所示, 服务 320-2 可以配备一个或多个 POD 340-1, 340-2, ..., 340-N (统称为 POD 340)。每个 POD 340 可以包括代理 345 和一个或多个容器 342-1, 342-2, ..., 342-M (统称为容器 342)。POD 340 中一个或多个容器 342 处理与服务的一个或多个相应功能相关的请求, 代理 345 通常控制与服务相关的网络功能, 如路由、负载均衡等。其他服务 320 也可以类似于 POD 340 的 POD。

在操作过程中, 执行来自终端用户 302 的用户请求可能需要调用计算环境 301 中的一个或多个服务 320, 执行一个服务 320 的一个或多个功能可能需要调用另一个服务 320 的一个或多个功能。如图 3 所示, 服务 320-1 (代表服务 A) 从入口网关 330 接收终端用户 302 的用户请求, 服务 320-1 (代表服务 A) 可以调用服务 320-2 (代表服务 D), 服务 320-2 (代表服务 D) 可以请求服务 320-3 (代表服务 E) 执行一个或多个功能。

上述的计算环境可以是云计算环境, 资源的分配由云服务提供上管理, 允许功能的开发无需考虑实现、调整或扩展服务器。该计算环境允许开发人员在不构建或维护复杂基础设施的情况下执行响应事件的代码。服务可以被分割完成一组可以自动独立伸缩的功能, 而不是扩展单个硬件设备来处理潜在的负载。

在上述运行环境下, 本公开提供了如图 4 所示的基于场景重建生成视频的方法。图 4 是根据本公开实施例 1 的一种基于场景重建生成视频的方法的流程图, 如图 4 所示, 该基于场景重建生成视频的方法包括:

步骤 S41, 获取待重建图像数据, 其中, 待重建图像数据被设置为记录待重建的预设三维场景的影像;

步骤 S42, 对待重建图像数据进行稀疏重建, 得到重建结果;

步骤 S43, 基于待重建图像数据和重建结果进行隐式建模, 得到建模结果;

步骤 S44, 利用建模结果生成目标视频, 其中, 目标视频被设置为对预设三维场景进行场景漫游。

上述基于场景重建生成视频的方法可以但不限于应用于以下场景: 视觉导航场景 (例如, 基于地图构建三维虚拟场景用于对机器人或无人机进行自主导航), VR/AR 场景制作场景 (例如, VR 展厅制作、VR/AR 游戏场景设计等), 三维打印场景 (例如, 基于图像重建虚拟模型用于实体生产), 三维影视创作及后期制作场景 (例如, 基于给定图像重建三维场景或三维场景中的虚拟物体), 文旅宣传场景 (例如, 景区、园区、机场、车站、道路等区域的三维虚拟场景重建, 文旅宣传视频制作等)。

作为一种示例性的实施例, 上述稀疏重建为基于 SfM 算法的三维场景重建操作, 上述重建结果包括待重建图像数据对应的相机位姿。上述隐式建模为基于 NeRF 算法的三维场景建模优化操作。上述建模结果包括待重建的预设三维场景对应的场景模型。上述目标视频为待重建的预设三维场景对应的自由视角场景漫游视频, 目标视频为利用渲染工具 (如游戏引擎、图像编辑软件等) 渲染得到的视频。

需要说明的是, 上述稀疏重建所采用的重建算法不限于 SfM 算法。例如, 重建算法还可以是: 基于贪心算法的稀疏表示, 如匹配追踪 (Matching Pursuit, MP) 算法、正交匹配追踪 (Orthogonal Matching Pursuit, OMP) 算法等; 基于迭代优化的稀疏表示, 如基于 L1 范数的最小绝对值收缩和选择算子 (Least Absolute Shrinkage and Selection Operator, Lasso) 算法、基于 L0 范数的反向传播 (Backpropagation, BP) 算法等; 基于压缩感知理论的稀疏表示, 如基于多项式收缩算法的压缩采样匹配追踪 (Compressive Sampling Matching Pursuit, CoSaMP) 算法、基于单正交匹配追踪的连续投影 (Successive Projection, SP) 算法等; 基于字典学习的稀疏表示, 如 K 奇异值分解 (K-Singular Value Decomposition, K-SVD) 算法、正交字典匹配追踪 (Matching Pursuit Algorithm with Orthogonal Dictionary, MOD) 算法等; 基于卷积神经网络的稀疏表示, 如基于稀疏编码的卷积神经网络 (Convolutional Neural Network, CNN) 算法等; 基于贝叶斯方法的稀疏表示, 如基于变分自编码器的贝叶斯稀疏编码算法等。上述隐式建模所采用的建模算法不限于 NeRF 算法。例如, 建模算法还可以是: 隐马尔可夫模型 (Hidden Markov Model, HMM) 算法, 决策树 (Decision Tree) 算法, 随机森林 (Random Forest) 算法, 支持向量机 (Support Vector Machine, SVM) 算法和神经网络 (Neural Network) 算法等。本公开并不对上述重建算法和建模算法进行限定。

本公开实施例提供的上述基于场景重建生成视频的方法可以运行于上述应用场景对应的客户端, 客户端从预设数据库或者用户实时输入的数据中确定应用场景中待重建图像数据, 然后对待重建图像数据进行稀疏重建, 得到重建结果, 基于待重建图像数据和重建结果进行隐式建模, 得到建模结果; 利用建模结果生成目标视频, 进一步可选地, 客户端还可以将目标视频显示在对应的图形用户界面上, 通过输出设备 (如显示屏、VR 眼镜等) 提供给用户。

本公开实施例提供的上述基于场景重建生成视频的方法可以运行于上述应用场景对应的服务端。上述服务端可以是独立的服务器或者服务器集群, 根据客户端给定的待重建图像数据, 通过上述稀疏重建和隐式建模方式生成应用场景对应的目标视频。上述服务端还可以是云服务器, 通过软件即服务 (Software as a Service, SaaS) 的方式与客户端进行实时交互, 根据客户端给定的待重建图像数据, 对待重建图像数据进行稀疏重建, 得到重建结果, 基于待重建图像数据和重建结果进行隐式建模, 得到建模结果; 利用建模结果生成目标视频, 进而将目标视频返回给客户端以提供给用户。

在本公开实施例中, 通过获取待重建图像数据, 其中, 待重建图像数据被设置为

记录待重建的预设三维场景的影像；进一步对待重建图像数据进行稀疏重建，得到重建结果；采用基于待重建图像数据和重建结果进行隐式建模的方式得到建模结果；进而利用建模结果生成目标视频，其中，目标视频被设置为对预设三维场景进行场景漫游。由此，本公开实施例达到了通过稀疏重建和隐式建模的方式基于待重建图像数据进行场景建模以渲染得到场景漫游视频的目的，从而实现了提升场景建模结果的准确度和场景漫游视频的真实感的技术效果，进而解决了由于现有技术采用显示建模方式进行场景重建导致的建模结果缺陷多、场景漫游视频的真实感差的技术问题。

根据本公开实施例提供的基于场景重建生成视频的方法，通过神经隐式网格模型进行基于图像数据的场景重建，从而渲染得到真实感较强的场景漫游视频。特别是，上述基于场景重建生成视频的方法尤其适被设置为基于航拍数据的大规模场景重建。

图 5 是根据本公开实施例 1 的一种可选的场景重建过程的示意图，如图 5 所示，以基于园区航拍数据对园区进行场景重建为例，场景重建过程包括以下四个阶段：第一阶段，航拍数据采集阶段；第二阶段，运动恢复 (SfM) 稀疏重建阶段；第三阶段，神经隐式优化建模阶段；第四阶段，漫游视频渲染阶段。

作为一种示例性的实施例，预设三维场景为园区场景，待重建图像数据为园区场景对应的航拍数据。如图 5 所示，在第一阶段，通过航拍设备实时采集得到航拍数据，其中，航拍数据包括至少一张航拍影像；在第二阶段，对航拍数据进行 SfM 稀疏重建，得到至少一张航拍影像中每张航拍影像对应的相机位姿（相当于上述重建结果）；在第三阶段，基于航拍数据和至少一张航拍影像对应的相机位姿进行隐式建模，也即，对重建结果进行建模优化，得到建模结果；在第四阶段，基于建模结果，利用视频编辑器生成目标视频，其中，目标视频为自由视角下的园区漫游视频。

作为一种示例性的实施例，基于建模结果生成目标视频包括：获取自由视角对应的视角轨迹；根据视角轨迹基于建模结果渲染得到目标视频。

在一种可选的实施例中，在执行基于待重建图像数据和重建结果进行隐式建模，得到建模结果的过程中，可以包括如下执行步骤：

首先，基于待重建图像数据和重建结果，在多个尺度特征上构建预设三维场景对应的栅格，得到构建结果，其中，构建结果被设置为确定预设三维场景在多个尺度特征上的三维空间信息与光场信息；

其次，利用构建结果进行隐式建模，得到建模结果。

在上述可选的实施例中，多个尺度特征为根据多个不同特征尺度划分的多个不同数据特征。对应地，上述构建结果包括多个尺度特征中每个尺度特征对应的预设三维场景的栅格，每个尺度特征对应的栅格中包含的栅格尺寸相同，多个尺度特征对应于多种栅格尺寸。预设三维场景对应的栅格中的每个栅格存储有对应的三维空间信息与光场信息，其中，三维空间信息包括预设三维场景的几何形状数据和空间位置数据，光场信息为预设三维场景的光学表现数据（例如，多种光学参数，包括：不透明度、亮度、饱和度等）。进一步，基于优化的 NeRF 算法，利用构建结果进行隐式建模，得到上述建模结果。

图 6 是根据本公开实施例 1 的一种可选的隐式建模过程的示意图，如图 6 所示，在图 5 中的第三阶段进行隐式建模的过程中：首先，构建预设三维场景对应的场景占用栅格；然后，基于场景占用栅格确定多尺度特征的哈希表，其中，哈希表中存储有多尺度特征的编码数据；进一步，基于多尺度特征哈希表进行神经体渲染优化，得到建模结果，其中，神经体渲染优化所采用的隐式建模算法包括大规模场景采样策略、稀疏约束和深度约束。

在一种可选的实施例中，在执行基于待重建图像数据和重建结果，在多个尺度特征上构建预设三维场景对应的栅格，得到构建结果的过程中，还可以包括如下执行步骤：

首先，基于待重建图像数据和重建结果确定预设三维场景对应的包围边界框；
其次，对包围边界框内的三维空间进行多尺度栅格化划分，得到划分结果；
最后，利用划分结果对应的光场信息，获取预设三维场景对应的栅格，得到构建结果。

仍然以基于园区航拍数据对园区进行场景重建为例，基于园区的航拍数据和重建结果，确定园区对应的包围边界框。示例性地，包围边界框可以是多边形柱体，多边形底面的尺寸和形状由园区的占地区域确定，多边形高度为预设的场景高度。例如，对于占地长500米、宽400米的园区，该园区对应的航拍最大高度设置为200米，则上述园区的包围边界框为500米×400米×200米的长方体边界框。

作为一种示例性的实施例，基于待重建图像数据和重建结果确定预设三维场景对应的包围边界框包括：基于待重建图像数据确定待重建的预设三维场景的影像对应的相机位姿信息；基于重建结果，确定预设三维场景在多个尺度特征上的三维空间信息与光场信息，其中，三维空间信息包括预设三维场景的场景尺寸数据；根据相机位姿信息、三维空间信息和光场信息，确定预设三维场景对应的包围边界框。

作为一种示例性的实施例，划分结果包括多个子划分结果，对包围边界框内的三维空间进行多尺度栅格化划分，得到划分结果包括：确定多个划分尺度，根据多个划分尺度对包围边界框内的三维空间进行划分，得到多个子划分结果。仍然以基于园区航拍数据对园区进行场景重建为例，确定五个划分尺度分别为：1米、2米、5米、10米和50米；基于五个划分尺度中的每个划分尺度，分别对500米×400米×200米的长方体边界框进行划分，得到五个子划分结果。以划分尺度为1米时为例，将500米×400米×200米的长方体边界框划分为500×400×200个1米×1米×1米的栅格，得到对应的第一子划分结果。以划分尺度为2米时为例，将500米×400米×200米的长方体边界框划分为250×200×100个2米×2米×2米的栅格，得到对应的第二子划分结果。

需要说明的是，上述划分结果中，每个栅格均对应存储有光场信息，例如，第一子划分结果中，某个1米×1米×1米的栅格对应存储有该栅格内包含的多个采样点的光场信息；第二子划分结果中，某个2米×2米×2米的栅格对应存储有该栅格内包含的多个采样点的光场信息。上述光场信息至少包括以下之一：不透明度、亮度、饱和度等。

作为一种示例性的实施例，光场信息包括不透明度，利用划分结果对应的光场信息，获取预设三维场景对应的栅格，得到构建结果包括：基于划分结果中多个栅格中每个栅格对应的不透明度，从多个栅格中部分或全部栅格作为目标栅格；将该目标栅格作为预设三维场景对应的栅格，得到构建结果。基于划分结果中多个栅格中每个栅格对应的不透明度，从多个栅格中选取目标栅格包括：基于多个栅格中每个栅格对应的不透明度依次进行判定；响应于当前栅格对应的不透明度满足预设条件，确定当前栅格为目标栅格，可选地，预设条件为不透明度高于预设阈值，预设条件被设置为确定当前栅格是否为未被预设三维场景占用的空栅格。

作为一种示例性的实施例，上述基于场景重建生成视频的方法还包括：根据构建结果对划分结果对应的光场信息进行优化，得到更新后的光场信息。具体地，根据构建结果对划分结果中多个栅格的每个栅格对应的不透明度进行多轮迭代优化调整，直至预设三维场景对应的栅格的不透明度为1或者处于预设范围（如0.9至1）。多轮迭代优化调整中每轮优化调整包括：将构建结果中预设三维场景对应的栅格的不透明度调高，将划分结果中除预设三维场景对应的栅格之外的其他栅格的不透明度调低。

容易理解的是，确定预设三维场景对应的栅格（也即场景占用栅格）后，在场景重建过程中，对上述场景占用栅格进行采样，跳过划分结果中除了场景占用栅格之外的其他栅格，从而节省了场景重建过程中的数据采样开销，提高了场景重建效率和目

标视频的渲染效率。

通过上述可选的实施例提供的执行步骤，使得构建结果中所确定的预设三维场景对应的栅格的准确度较高，较少或不会出现将划分结果的多个栅格中预设三维场景未占用的栅格确定为预设三维场景对应的栅格的情况，进而保证后续场景重建结果的准确度。

在一种可选的实施例中，在执行利用构建结果进行隐式建模，得到建模结果的过程中，还可以包括如下执行步骤：

首先，根据多个尺度特征，将构建结果中的三维空间信息与光场信息划分至多个分区，得到划分结果，其中，划分结果中的不同分区被设置为分别存储不同尺度特征的三维空间信息与光场信息；

其次，按照预设采样策略，基于航拍图像采集设备的采集视角和目标渲染位置对构建结果进行采样，得到采样结果，其中，预设采样策略被设置为确定目标渲染位置在采样过程中待使用的采样次数与采样步长，以及从目标渲染位置发射的光线数目；

然后，利用采样结果从划分结果中获取第一多尺度特征；

最后，采用第一多尺度特征进行隐式建模，得到建模结果。

作为一种示例性的实施例，根据多个尺度特征，将构建结果中的三维空间信息与光场信息划分至多个分区，得到划分结果包括：基于构建结果对应的多个划分尺度，确定多个分区；基于多个分区，将构建结果中的多个尺度特征上的三维空间信息与光场信息划分至多个分区，得到划分结果。第一多尺度特征包括待重建图像数据对应的多组特征，多组特征与上述多个划分尺度相对应，例如，每个划分尺度对应第一多尺度特征中的一组特征。上述多个分区可以是与上述多个尺度特征对应的多个存储区域。

仍然以基于园区航拍数据对园区进行场景重建为例，确定五个划分尺度分别为：1米、2米、5米、10米和50米，对应地，确定五个分区。当划分尺度为1米时，将500米×400米×200米的长方体边界框划分为500×400×200个1米×1米×1米的栅格，对应的第一分区的维度为500×400×200。当划分尺度为2米时，将500米×400米×200米的长方体边界框划分为250×200×100个2米×2米×2米的栅格，对应的第二分区的维度为250×200×100。也即，对于包含 $N_1 \times N_2 \times N_3$ 个栅格的子划分结果，对应的分区的维度为 $N_1 \times N_2 \times N_3$ 。

作为一种示例性的实施例，上述多个分区为多尺度特征哈希表，例如，每个分区对应一个特征尺度的哈希表。如图6所示，第一多尺度特征为多尺度特征哈希表，多尺度特征哈希表包括多个不同维度的特征表。如图6所示的特征表1、特征表2、……特征表n对应于n个栅格划分尺度。

容易理解的是，通过利用多尺度特征哈希表，将构建结果中的多个尺度特征上的三维空间信息与光场信息划分至多个分区，能够在场景重建过程中利用多尺度特征哈希表方便灵活地存取不同尺度特征上的三维空间信息与光场信息，增强了场景重建过程中特征编码的表征能力。

作为一种示例性的实施例，根据应用场景需求，获取预设采样策略；按照预设采样策略，确定目标渲染位置在采样过程中待使用的采样次数与采样步长，以及从目标渲染位置发射的光线数目；基于航拍图像采集设备的采集视角和目标渲染位置，按照采样此处和采样步长，沿着目标渲染位置发射的多条光线中的每条光线的方向，对构建结果进行采样，得到采样结果。上述目标渲染位置发射的光线的方向由目标渲染位置和航拍图像采集设备的采集视角确定。例如，上述航拍图像采集设备为航拍相机。上述目标采样位置为渲染过程中的当前渲染像素点。

需要说明的是，上述对构建结果进行采样包括：在构建结果中预设三维场景占用的栅格中上述每条光线经过的栅格进行采样。每条光线对应的采样结果包括多个采样点。

作为一种示例性的实施例，利用采样结果从划分结果中获取第一多尺度特征包括：利用采样结果中多条光线中的每条光线对应的多个采样点，根据采样点的位置信息从划分结果中获取第一多尺度特征，具体地，划分结果为多尺度特征哈希表，根据每个采样点的位置信息，从多尺度特征哈希表中获取多个尺度的特征值，然后将多个尺度的特征值组合得到该采样点的第一多尺度特征。进一步地，采用第一多尺度特征进行隐式建模，得到建模结果。

容易理解的是，本公开实施例中，在利用构建结果进行隐式建模的过程中，采用上述分区存储方式、基于预设采样策略的采样方式以及获取第一多尺度特征的方式，实现基于第一多尺度特征的隐式建模，与现有技术中的基于场景重建生成视频的方法相比，本公开实施例的场景特征采样准确度高，隐式建模的优化效果强，能够得到更加准确和真实的场景重建结果。

在一种可选的实施例中，在执行采用第一多尺度特征进行隐式建模，得到建模结果的过程中，还可以包括如下执行步骤：

首先，采用目标全连接神经网络模型对第一多尺度特征进行像素值预测，得到目标像素值，其中，目标全连接神经网络模型采用多组数据通过深度学习训练得到，多组数据包括：训练图像和训练图像对应的第二多尺度特征；

其次，基于目标像素值生成建模结果。

上述第二多尺度特征包括训练图像对应的多组特征，多组特征与上述多个划分尺度相对应，例如，每个划分尺度对应第二多尺度特征中的一组特征。

在上述可选的实施例中，上述目标全连接神经网络模型为多层感知机（MLP）。MLP 采用多组数据深度学习训练得到，多组数据包括：训练图像和训练图像对应的第二多尺度特征。在训练 MLP 的过程中，上述第二多尺度特征的获取方式可以参照上述利用待重建图像数据对应的构建结果进行隐式建模，得到建模结果的执行步骤。

如图 6 所示，基于多尺度特征哈希表进行神经体渲染优化，得到建模结果。具体地，针对大规模场景（如园区场景、城市场景、地块场景等），确定大规模采样策略；进一步地，按照大规模采样策略对构建结果对应的场景占用栅格进行采样，得到采样结果，并根据采样结果从多尺度特征哈希表中获取第一多尺度特征；然后，采用 MLP 对第一多尺度特征进行像素值预测得到目标像素值，目标像素值可以包括像素点的 RGB 属性值；基于目标像素值生成建模结果。

在一种可选的实施例中，基于场景重建生成视频的方法还包括如下方法步骤：

步骤 S451，采用初始全连接神经网络模型对第二多尺度特征进行像素值预测，得到训练像素值；

步骤 S452，通过训练像素值与训练图像对应的真实像素值确定第一损失；

步骤 S453，基于第一损失对初始全连接神经网络模型的模型参数进行更新，得到目标全连接神经网络模型。

例如，上述步骤 S451 至步骤 S453 的过程还可以具体地如图 6 所示，在神经体渲染优化过程中，对 MLP 进行训练时，采用训练像素值与真实像素值确定的第一损失，对初始 MLP 进行更新得到目标 MLP，其中，训练像素值为初始 MLP 对训练图像对应的第二多尺度特征进行像素值预测所得到的像素值，真实像素值为训练图像对应的真实像素值。示例性地，上述第一损失的计算方式为：训练像素值与真实像素值之差的绝对值。

根据本公开实施例的步骤 S451 至步骤 S453 提供的方法，在场景重建过程中对待使用的目标 MLP 进行训练，能够增强目标 MLP 的模型性能，进而提升对第一多尺度特征进行像素值预测的预测准确度。

在一种可选的实施例中，基于场景重建生成视频的方法还包括如下方法步骤：

步骤 S461，采用初始全连接神经网络模型对第二多尺度特征进行深度值预测，得

到训练深度值；

步骤 S462, 通过训练深度值与训练图像对应的真实深度值确定第二损失；

步骤 S463, 采用采集视角和目标渲染位置确定采样射线；

步骤 S464, 通过采样结果中的采样点在采样射线上的采样权重确定第三损失；

步骤 S465, 基于第二损失与第三损失中至少之一对目标全连接神经网络模型的模型参数进行更新, 得到更新后的目标全连接神经网络模型。

例如, 上述步骤 S461 至步骤 S465 的过程还可以具体地如图 6 所示, 在对初始 MLP 进行训练的过程中, 基于深度约束, 利用训练深度值和真实深度值确定第二损失; 基于稀疏约束, 利用第二多尺度特征对应的采样射线、采样结果和采样权重确定第三损失; 基于第二损失与第三损失中至少之一对初始 MLP 进行模型参数调整, 得到目标 MLP, 或者, 基于第二损失与第三损失中至少之一对目标 MLP 进行模型参数更新, 得到更新后的目标 MLP。

需要说明的是, 上述深度约束和上述稀疏约束为针对大规模场景(如园区场景、城市场景、地块场景等)预设的约束规则。深度约束被设置为剪除部分栅格(如高空中与场景重建无关的栅格), 从而减少场景重建结果中可能存在的漂浮物错误的问题。稀疏约束被设置为确定每条采样射线上多个采样点的采样权重集中在指定范围内, 从而进一步减少场景重建结果中可能存在的漂浮物错误的问题。

示例性地, 深度约束为基于预设深度图的深度图约束。基于深度约束, 利用训练深度值和真实深度值确定第二损失包括: 基于预设深度图、初始 MLP 对第二多尺度特征进行深度值预测得到的训练深度值以及训练图像对应的真实深度值, 确定第二损失。第二损失的计算方式可以是: 在满足预设深度图的深度图约束的情况下, 训练深度值与真实深度值之差的绝对值。

示例性地, 稀疏约束为权重阈值约束。第三损失的计算方式可以是: 针对每条采样射线, 在该采样射线上多个采样点的采样权重满足权重阈值约束的情况下, 计算多个采样点的采样权重的熵损失。

根据本公开实施例的步骤 S461 至步骤 S465 提供的方法, 基于针对大规模场景设计的深度约束和稀疏约束, 在场景重建过程中对待使用的目标 MLP 进行训练, 能够增强目标 MLP 的模型性能, 进而提升对第一多尺度特征进行像素值预测的预测准确度。

在一种可选的实施例中, 在步骤 S41 中, 获取待重建图像数据, 包括如下方法步骤:

步骤 S411, 在预设条件下, 利用航拍图像采集设备对预设三维场景进行航拍数据采集, 得到待重建图像数据, 其中, 预设条件被设置为确定航拍图像采集设备在航拍数据采集过程中的天气条件和飞行条件。

在上述可选的实施例中, 预设条件还被设置为确定采用神经隐式建模方法进行场景建模。示例性地, 在基于航拍图像数据对园区进行场景重建的应用场景中, 上述预设条件至少包括: 拍摄天气为阴天; 航拍图像采集设备的飞行高度为 110 米; 航拍图像采集设备的旁向重叠率为 80%; 航拍图像采集设备的航向重叠率为 85%; 航拍图像采集设备的云台倾斜角度为 -50 度; 当航拍图像采集设备对预设的场景重点区域进行拍摄时, 采用一预设的近距离拍摄。

在一种可选的实施例中, 在步骤 S42 中, 对待重建图像数据进行稀疏重建, 得到重建结果, 包括如下方法步骤:

步骤 S421, 对待重建图像数据进行特征提取, 得到提取结果;

步骤 S422, 对提取结果进行特征匹配, 得到匹配结果;

步骤 S423, 对匹配结果进行稀疏重建, 得到重建结果, 其中, 重建结果被设置为确定待重建图像数据对应的航拍图像采集设备的位姿信息。

仍然如图 5 所示, 在第二阶段, 进行运动恢复 (SfM) 稀疏重建时, 基于所采集的航拍数据, 执行如下流程: 特征提取、特征匹配和稀疏重建, 进而达到基于航拍数据求解航拍图像采集设备的位姿信息的目的。

作为上述执行步骤 S421 的一种可选的实施方式, 对待重建图像数据进行特征提取, 得到提取结果, 包括如下方法步骤:

步骤 S4211, 对待重建图像数据进行预处理, 得到预处理结果;

步骤 S4212, 对预处理结果进行特征提取, 得到提取结果。

作为一种示例性的实施例, 对待重建图像数据进行预处理, 得到预处理结果包括: 根据待重建图像数据中的影像的分辨率进行自适应调整, 得到调整结果; 将待重建图像数据中的影像对应的拍摄元数据 (Exchangeable Image File Format, EXIF) 信息写入对应的调整结果, 得到预处理结果, 其中, EXIF 信息包括: 设备出厂商、设备型号、软件版本号、光圈、快门速度、焦距、感光度、测光模式等。

作为一种示例性的实施例, 提取结果包括图像点特征和图像描述子, 其中, 图像描述子被设置为确定图像的局部特征。对预处理结果进行特征提取, 得到提取结果包括: 对预处理结果中的每张图像进行点特征提取, 得到图像点特征; 对预处理结果中的每张图像进行局部特征提取, 得到图像描述子。

作为上述执行步骤 S422 的一种可选的实施方式, 对提取结果进行特征匹配, 得到匹配结果, 包括如下方法步骤:

步骤 S4221, 对提取结果进行图像配对, 得到配对结果;

步骤 S4222, 对配对结果进行特征匹配, 得到匹配结果。

作为一种示例性的实施例, 对提取结果进行图像配对, 得到配对结果包括: 获取预设三维场景对应的卫星定位数据; 基于卫星定位数据对提取结果进行图像匹配, 得到配对结果。具体地, 基于卫星定位数据, 根据提取结果中的图像点特征和图像描述子, 对提取结果对应的当前图像与该当前图像距离较近的其他场景图像 (可以从预处理结果中选取) 进行配对。由此, 对提取结果对应的多张图像进行两两配对后, 针对配对的两张图像进行图像间约束计算。

作为一种示例性的实施例, 对配对结果进行特征匹配, 得到匹配结果包括: 对配对结果中配对的两张图像进行基于图像描述子的图像特征点匹配, 得到匹配结果。匹配结果被设置为计算配对的两张图像之间的相对旋转关系和相对平移关系。

作为上述执行步骤 S423 的一种可选的实施方式, 对匹配结果进行稀疏重建, 得到重建结果, 包括如下方法步骤:

步骤 S4231, 对匹配结果进行几何过滤, 得到过滤结果;

步骤 S4232, 对过滤结果进行全局三维结构估计, 得到估计结果;

步骤 S4233, 基于估计结果确定重建结果。

作为一种示例性的实施例, 对匹配结果进行几何过滤, 得到过滤结果包括: 基于预设几何约束, 对匹配结果进行匹配过滤以滤除匹配结果中误匹配的图像特征点, 得到过滤结果。例如, 上述预设几何约束包括本征矩阵约束或者基础矩阵约束。

作为一种示例性的实施例, 对过滤结果进行全局三维结构估计, 得到估计结果包括: 基于过滤结果, 计算配对的两张图像之间的相对旋转关系和相对平移关系; 根据相对旋转关系, 确定配对的两张图像对应的旋转平均信息; 根据相对平移关系, 确定配对的两张图像对应的平移平均信息; 基于旋转平均信息和平移平均信息, 采用全局 SfM 算法计算得到全局三维结构的估计结果, 其中, 估计结果包括航拍图像采集设备的内部参数和航拍图像采集设备对应的采集点位姿。上述全局 SfM 算法是对传统的 SfM 算法进行优化得到的, 上述全局 SfM 算法为基于旋转平均信息和平移平均信息的算法。

示例性地, 根据相对旋转关系, 确定配对的两张图像对应的旋转平均信息包括:

利用相对旋转关系对应的旋转约束条件，结合向量范数（本例中采用 L2 范数）构造旋转优化问题，进而利用旋转优化问题求解得到配对的两张图像对应的旋转平均信息。

示例性地，根据相对平移关系，确定配对的两张图像对应的平移平均信息包括：利用相对平移关系对应的平移约束条件，结合向量范数（本例中采用 L1 范数）构造平移优化问题，进而利用平移优化问题求解得到配对的两张图像对应的平移平均信息。

为了进一步提高重建结果的准确度，作为上述执行步骤 S4233 的一种可选的实施方式，基于估计结果确定重建结果，包括如下方法步骤：

步骤 S4234，对估计结果进行地平线矫正，得到第一矫正结果；

步骤 S4235，对第一矫正结果进行航向角矫正，得到第二矫正结果；

步骤 S4236，对第二矫正结果进行点云分块处理，得到分块结果；

步骤 S4237，对分块结果进行数据格式调整，得到重建结果。

作为一种示例性的实施例，对估计结果进行地平线矫正，得到第一矫正结果包括：基于估计结果中的航拍图像采集设备的内部参数进行鲁棒平面拟合，得到目标鲁棒平面；利用目标鲁棒平面的法向方向进行旋转估计，得到旋转矩阵，其中，旋转矩阵被设置为确定航拍图像采集设备与真实地平线之间的旋转偏角；根据旋转矩阵对估计结果进行地平线矫正，得到第一矫正结果。

作为一种示例性的实施例，对第一矫正结果进行航向角矫正，得到第二矫正结果包括：基于估计结果中的航拍图像采集设备对应的采集点位置，对第一矫正结果进行航向角矫正，得到第二矫正结果。通过上述航向角矫正，保证后续过程确定较小的预设三维场景的包围边界框，提高场景重建结果的准确度。第二矫正结果包括经过航向角矫正后的预设三维场景中分布的点云，点云被设置为存储点云位置信息、航拍图像采集设备的内部参数和航拍图像采集设备对应的采集点位置，点云中的每个点对应一个航拍图像采集设备。

作为一种示例性的实施例，对第二矫正结果进行点云分块处理，得到分块结果包括：对第二矫正结果中的点云进行可重叠分块，得到分块结果。上述可重叠分块是指：点云中的某个点可以被划分在至少一个点云块中。通过对第二矫正结果进行点云分块处理，进一步提升了本公开实施例对大规模场景的重建能力。

作为一种示例性的实施例，对分块结果进行数据格式调整，得到重建结果包括：按照预设数据格式对分块结果中多个点云块对应的点云数据进行调整，得到重建结果，上述预设数据格式为点云位置信息、航拍图像采集设备的内部参数和航拍图像采集设备对应的采集点位置对应的统一输出格式。

在一种可选的实施例中，在步骤 S44 中，利用建模结果生成目标视频，包括如下方法步骤：

步骤 S441，获取航拍图像采集设备对应的轨迹信息；

步骤 S442，按照轨迹信息对建模结果进行视频渲染，得到目标视频。

作为一种示例性的实施例，通过图像编辑工具利用建模结果生成目标视频。在图像编辑工具中确定航拍图像采集设备对应的图像渲染轨迹，按照图像渲染轨迹对建模结果进行视频渲染，得到目标视频。具体地，在图像编辑工具中确定航拍图像采集设备对应的图像渲染轨迹包括：获取预先绘制的图像渲染轨迹；基于图像渲染轨迹进行视频渲染预览，得到预览结果；根据预览结果确定图像渲染轨迹是否可用。预先绘制的图像渲染轨迹可以由用户自由绘制，也可以是随机生成的自由漫游轨迹，由此，按照轨迹信息对建模结果进行视频渲染所得到的目标视频为预设三维场景在自由视角下的场景漫游视频。

容易理解的是，根据本公开实施例提供的基于场景重建生成视频的方法，通过改进的神经隐式网格模型进行基于图像数据的场景重建，从而渲染得到真实感较强的场景漫游视频。特别是，上述基于场景重建生成视频的方法尤其适用于基于航拍数据的

大规模场景重建。

需要说明的是，本公开所涉及的用户信息（包括但不限于用户设备信息、用户个人信息等）和数据（包括但不限于用于分析的数据、存储的数据、展示的数据等），均为经用户授权或者经过各方充分授权的信息和数据，并且相关数据的收集、使用和处理需要遵守相关国家和地区的相关法律法规和标准，并提供有相应的操作入口，供用户选择授权或者拒绝。

需要说明的是，对于前述的各方法实施例，为了简单描述，故将其都表述为一系列的动作组合，但是本领域技术人员应该知悉，本公开并不受所描述的动作顺序的限制，因为依据本公开，某些步骤可以采用其他顺序或者同时进行。其次，本领域技术人员也应该知悉，说明书中所描述的实施例均属于优选实施例，所涉及的动作和模块并不一定是本公开所必须的。

通过以上的实施方式的描述，本领域的技术人员可以清楚地了解到根据上述实施例的方法可借助软件加必需的通用硬件平台的方式来实现，当然也可以通过硬件。基于这样的理解，本公开的技术方案本质上或者说对现有技术做出贡献的部分可以以软件产品的形式体现出来，该计算机软件产品存储在一个存储介质（如只读存储器（Read-Only Memory, ROM）、随机存取器（Random Access Memory, RAM）、磁碟、光盘）中，包括若干指令用以使得一台终端设备（可以是手机，计算机，服务器，或者网络设备等）执行本公开各个实施例所述的方法。

实施例 2

在如实施例 1 中的运行环境下，本公开提供了如图 7 所示的另一种基于场景重建生成视频的方法。图 7 是根据本公开实施例 2 的一种基于场景重建生成视频的方法的流程图，如图 7 所示，该基于场景重建生成视频的方法包括：

步骤 S71，响应作用于操作界面上的输入指令，在操作界面上显示待重建图像数据，其中，待重建图像数据被设置为记录待重建的预设三维场景的影像；

步骤 S72，响应作用于操作界面上的重建指令，在操作界面上显示目标视频，其中，目标视频被设置为对预设三维场景进行场景漫游，目标视频利用建模结果生成，建模结果通过对待重建图像数据和重建结果进行隐式建模得到，重建结果通过对待重建图像数据进行稀疏重建得到。

上述操作界面可以是执行上述基于场景重建生成视频的方法的客户端的图形用户界面。上述输入指令可以根据作用于操作界面上的输入操作生成，输入操作可以是触控操作或者通过键盘、鼠标、摇杆等设备实现的操作。当检测到作用于操作界面上的输入指令时，在操作界面上显示待重建图像数据。上述重建指令可以根据作用于操作界面上的重建操作生成，重建操作可以是触控操作或者通过键盘、鼠标、摇杆等设备实现的操作（如点击确定按钮等）。当检测到作用于操作界面上的重建指令时，在操作界面上显示目标视频。上述目标视频可以由上述客户端根据本公开实施例 1 中提供的基于场景重建生成视频的方法的具体实施方式生成。上述目标视频还可以由上述客户端请求对应的服务端生成，该服务端根据本公开实施例 1 中提供的基于场景重建生成视频的方法的具体实施方式生成目标视频。

上述基于场景重建生成视频的方法可以但不限于应用于以下场景：视觉导航场景（例如，基于地图构建三维虚拟场景用于对机器人或无人机进行自主导航），VR/AR 场景制作场景（例如，VR 展厅制作、VR/AR 游戏场景设计等），三维打印场景（例如，基于图像重建虚拟模型用于实体生产），三维影视创作及后期制作场景（例如，基于给定图像重建三维场景或三维场景中的虚拟物体），文旅宣传场景（例如，景区、园区、机场、车站、道路等区域的三维虚拟场景重建，文旅宣传视频制作等）。上述待重建图像数据用于记录上述应用场景中待重建的预设三维场景的影像（如航拍影像）。上述目标视频为上述应用场景中待重建的预设三维场景对应的自由视角下的场

景漫游视频。

由此，根据本公开实施例 2，用户能够通过对客户端的操作界面执行控制操作以生成输入指令和/或重建指令，指定待重建图像数据，并触发客户端执行如下步骤：对待重建图像数据进行稀疏重建得到重建结果；对待重建图像数据和重建结果进行隐式建模得到建模结果；利用建模结果生成目标视频；在操作界面上显示目标视频。上述过程中，用户具有较高的自由度和灵活度，能够根据应用场景需求进行个性化的三维场景重建对应的目标视频的定制，本公开实施例提供的上述方法有利于实际场景中的应用。

需要说明的是，对待重建图像数据进行稀疏重建得到重建结果，对待重建图像数据和重建结果进行隐式建模得到建模结果，以及利用建模结果生成目标视频还包括其他方法步骤，可以参见本公开实施例 1 中的相关描述，不予赘述。

实施例 3

在如实施例 1 中的运行环境下，本公开提供了如图 8 所示的另一种基于场景重建生成视频的方法。图 8 是根据本公开实施例 3 的一种基于场景重建生成视频的方法的流程图，如图 8 所示，该基于场景重建生成视频的方法包括：

步骤 S81，在虚拟现实 VR 设备或增强现实 AR 设备的呈现画面上展示初始默认画面；

步骤 S82，驱动 VR 设备或 AR 设备从展示初始默认画面跳转至展示目标视频，其中，目标视频被设置为对待重建的预设三维场景进行场景漫游，目标视频利用建模结果生成，建模结果通过对待重建图像数据和重建结果进行隐式建模得到，重建结果通过对待重建图像数据进行稀疏重建得到，待重建图像数据被设置为记录预设三维场景的影像。

上述基于场景重建生成视频的方法可以但不限于应用于以下场景：VR/AR 视觉导航场景（例如，基于地图构建三维虚拟场景，并基于三维虚拟场景对机器人或无人机进行 VR/AR 自主导航），VR/AR 场景制作场景（例如，VR 展厅制作、VR/AR 游戏场景设计等），VR/AR 三维影视创作及后期制作场景（例如，基于给定图像重建三维场景或三维场景中的虚拟物体，将重建结果通过 VR/AR 设备提供给观众），VR/AR 文旅宣传场景（例如，景区、园区、机场、车站、道路等区域的三维虚拟场景重建，将重建结果通过 VR/AR 设备提供给旅客）。上述待重建图像数据被设置为记录上述应用场景中待重建的预设三维场景的影像（如航拍影像）。上述目标视频为上述应用场景中待重建的预设三维场景对应的自由视角下的场景漫游视频。

上述初始默认画面为虚拟现实 VR 设备或增强现实 AR 设备在当前时刻显示的场景画面。上述当前时刻显示的场景画面可以是对预设三维场景进行场景重建生成的历史目标视频。当检测到切换指令时，驱动 VR 设备或 AR 设备从展示初始默认画面跳转至展示目标视频，上述切换指令被设置为确定将待展示的目标视频展示在 VR 设备或 AR 设备的呈现画面上。上述目标视频可以由上述虚拟现实 VR 设备或增强现实 AR 设备关联的处理器能够执行如下步骤得到：对待重建图像数据进行稀疏重建得到重建结果；对待重建图像数据和重建结果进行隐式建模得到建模结果；利用建模结果生成目标视频；在操作界面上显示目标视频。

通过本公开实施例的虚拟模型的展示方法，在虚拟现实或增强现实的场景下，生成准确度较高场景重建结果，并生成真实感较强的场景漫游视频，进而在虚拟现实 VR 设备或增强现实 AR 设备上动态且准确地显示预设三维场景对应的自由视角下的场景漫游视频。特别是，上述基于场景重建生成视频的方法尤其适用于基于航拍数据的大规模场景重建。

需要说明的是，本实施例的优选实施方式可以参见实施例 1 或实施例 2 中的相关描述，此处不再赘述。

实施例 4

根据本公开实施例，还提供了一种被设置为实施上述基于场景重建生成视频的方法的装置实施例。图 9 是根据本公开实施例 4 的一种基于场景重建生成视频的装置的结构示意图，如图 9 所示，该装置包括：获取模块 901，被设置为获取待重建图像数据，其中，待重建图像数据被设置为记录待重建的预设三维场景的影像；重建模块 902，被设置为对待重建图像数据进行稀疏重建，得到重建结果；建模模块 903，被设置为基于待重建图像数据和重建结果进行隐式建模，得到建模结果；生成模块 904，被设置为利用建模结果生成目标视频，其中，目标视频被设置为对预设三维场景进行场景漫游。

可选地，上述建模模块 903 还被设置为：基于待重建图像数据和重建结果，在多个尺度特征上构建预设三维场景对应的栅格，得到构建结果，其中，构建结果被设置为确定预设三维场景在多个尺度特征上的三维空间信息与光场信息；利用构建结果进行隐式建模，得到建模结果。

可选地，上述建模模块 903 还被设置为：基于待重建图像数据和重建结果确定预设三维场景对应的包围边界框；对包围边界框内的三维空间进行多尺度栅格化划分，得到划分结果；利用划分结果对应的光场信息，获取预设三维场景对应的栅格，得到构建结果。

可选地，上述建模模块 903 还被设置为：根据多个尺度特征，将构建结果中的三维空间信息与光场信息划分至多个分区，得到划分结果，其中，划分结果中的不同分区被设置为分别存储不同尺度特征的三维空间信息与光场信息；按照预设采样策略，基于航拍图像采集设备的采集视角和目标渲染位置对构建结果进行采样，得到采样结果，其中，预设采样策略被设置为确定目标渲染位置在采样过程中待使用的采样次数与采样步长，以及从目标渲染位置发射的光线数目；利用采样结果从划分结果中获取第一多尺度特征；采用第一多尺度特征进行隐式建模，得到建模结果。

可选地，上述建模模块 903 还被设置为：采用目标全连接神经网络模型对第一多尺度特征进行像素值预测，得到目标像素值，其中，目标全连接神经网络模型采用多组数据通过深度学习训练得到，多组数据包括：训练图像和训练图像对应的第二多尺度特征；基于目标像素值生成建模结果。

可选地，图 10 是根据本公开实施例 4 的一种可选的 Ozi 装置的结构示意图，如图 10 所示，该装置除包括图 9 所示的所有模块外，还包括：第一更新模块 905，被设置为采用初始全连接神经网络模型对第二多尺度特征进行像素值预测，得到训练像素值；通过训练像素值与训练图像对应的真实像素值确定第一损失；基于第一损失对初始全连接神经网络模型的模型参数进行更新，得到目标全连接神经网络模型。

可选地，图 11 是根据本公开实施例 4 的另一种可选的 Ozi 装置的结构示意图，如图 11 所示，该装置除包括图 10 所示的所有模块外，还包括：第二更新模块 906，被设置为采用初始全连接神经网络模型对第二多尺度特征进行深度值预测，得到训练深度值；通过训练深度值与训练图像对应的真实深度值确定第二损失；采用采集视角和目标渲染位置确定采样射线；通过采样结果中的采样点在采样射线上的采样权重确定第三损失；基于第二损失与第三损失中至少之一对目标全连接神经网络模型的模型参数进行更新，得到更新后的目标全连接神经网络模型。

可选地，上述获取模块 901 还被设置为：在预设条件下，利用航拍图像采集设备对预设三维场景进行航拍数据采集，得到待重建图像数据，其中，预设条件被设置为确定航拍图像采集设备在航拍数据采集过程中的天气条件和飞行条件。

可选地，上述重建模块 902 还被设置为：对待重建图像数据进行特征提取，得到提取结果；对提取结果进行特征匹配，得到匹配结果；对匹配结果进行稀疏重建，得到重建结果，其中，重建结果被设置为确定待重建图像数据对应的航拍图像采集设备

的位姿信息。

可选地,上述重建模块 902 还被设置为:对待重建图像数据进行预处理,得到预处理结果;对预处理结果进行特征提取,得到提取结果。

可选地,上述重建模块还被设置为:对提取结果进行图像配对,得到配对结果;对配对结果进行特征匹配,得到匹配结果。

可选地,上述重建模块 902 还被设置为:对匹配结果进行几何过滤,得到过滤结果;对过滤结果进行全局三维结构估计,得到估计结果;基于估计结果确定重建结果。

可选地,上述重建模块 902 还被设置为:对估计结果进行地平线矫正,得到第一矫正结果;对第一矫正结果进行航向角矫正,得到第二矫正结果;对第二矫正结果进行点云分块处理,得到分块结果;对分块结果进行数据格式调整,得到重建结果。

可选地,上述生成模块 904 还被设置为:获取航拍图像采集设备对应的轨迹信息;按照轨迹信息对建模结果进行视频渲染,得到目标视频。

在本公开实施例中,通过获取待重建图像数据,其中,待重建图像数据被设置为记录待重建的预设三维场景的影像;进一步对待重建图像数据进行稀疏重建,得到重建结果;采用基于待重建图像数据和重建结果进行隐式建模的方式得到建模结果;进而利用建模结果生成目标视频,其中,目标视频被设置为对预设三维场景进行场景漫游。由此,本公开实施例达到了通过稀疏重建和隐式建模的方式基于待重建图像数据进行场景建模以渲染得到场景漫游视频的目的,从而实现了提升场景建模结果的准确度和场景漫游视频的真实感的技术效果,进而解决了[关键词]的技术问题。

此处需要说明的是,上述获取模块 901、重建模块 902、建模模块 903 和生成模块 904 对应于实施例 1 中的步骤 S41 至步骤 S44,四个模块与对应的步骤所实现的实例和应用场景相同,但不限于上述实施例 1 所公开的内容。需要说明的是,上述模块或单元可以是存储在存储器(例如,存储器 104)中并由一个或多个处理器(例如,处理器 102a, 102b, ..., 102n)处理的硬件组件或软件组件,上述模块也可以作为装置的一部分可以运行在实施例 1 提供的计算机终端 10 中。

需要说明的是,本实施例的优选实施方式可以参见实施例 1 中的相关描述,此处不再赘述。

实施例 5

根据本公开实施例,还提供了一种计算机终端,该计算机终端可以是计算机终端群中的任意一个计算机终端设备。可选地,在本实施例中,上述计算机终端也可以替换为移动终端等终端设备。

可选地,在本实施例中,上述计算机终端可以位于计算机网络的多个网络设备中的至少一个网络设备。

在本实施例中,上述计算机终端可以执行基于场景重建生成视频的方法中以下步骤的程序代码:获取待重建图像数据,其中,待重建图像数据被设置为记录待重建的预设三维场景的影像;对待重建图像数据进行稀疏重建,得到重建结果;基于待重建图像数据和重建结果进行隐式建模,得到建模结果;利用建模结果生成目标视频,其中,目标视频被设置为对预设三维场景进行场景漫游。

可选地,图 12 是根据本公开实施例 5 的一种计算机终端的结构框图,如图 12 所示,该计算机终端 120 可以包括:一个或多个(图中示出一个)处理器 1202、存储器 1204、存储控制器 1206、以及外设接口 1208,其中,外设接口 1208 与射频模块、音频模块和显示器连接。

其中,存储器 1204 可被设置为存储软件程序以及模块,如本公开实施例中的基于场景重建生成视频的方法和装置对应的程序指令/模块,处理器通过运行存储在存储器内的软件程序以及模块,从而执行各种功能应用以及数据处理,即实现上述的基于场景重建生成视频的方法。存储器 1204 可包括高速随机存储器,还可以包括非易失

性存储器，如一个或者多个磁性存储装置、闪存、或者其他非易失性固态存储器。在一些实例中，存储器 1204 可进一步包括相对于处理器远程设置的存储器，这些远程存储器可以通过网络连接至计算机终端 120。上述网络的实例包括但不限于互联网、企业内部网、局域网、移动通信网及其组合。

处理器 1202 可以通过传输装置调用存储器存储的信息及应用程序，以执行下述步骤：获取待重建图像数据，其中，待重建图像数据被设置为记录待重建的预设三维场景的影像；对待重建图像数据进行稀疏重建，得到重建结果；基于待重建图像数据和重建结果进行隐式建模，得到建模结果；利用建模结果生成目标视频，其中，目标视频被设置为对预设三维场景进行场景漫游。

可选的，上述处理器 1202 还可以执行如下步骤的程序代码：基于待重建图像数据和重建结果，在多个尺度特征上构建预设三维场景对应的栅格，得到构建结果，其中，构建结果被设置为确定预设三维场景在多个尺度特征上的三维空间信息与光场信息；利用构建结果进行隐式建模，得到建模结果。

可选的，上述处理器 1202 还可以执行如下步骤的程序代码：基于待重建图像数据和重建结果确定预设三维场景对应的包围边界框；对包围边界框内的三维空间进行多尺度栅格化划分，得到划分结果；利用划分结果对应的光场信息，获取预设三维场景对应的栅格，得到构建结果。

可选的，上述处理器 1202 还可以执行如下步骤的程序代码：根据多个尺度特征，将构建结果中的三维空间信息与光场信息划分至多个分区，得到划分结果，其中，划分结果中的不同分区被设置为分别存储不同尺度特征的三维空间信息与光场信息；按照预设采样策略，基于航拍图像采集设备的采集视角和目标渲染位置对构建结果进行采样，得到采样结果，其中，预设采样策略被设置为确定目标渲染位置在采样过程中待使用的采样次数与采样步长，以及从目标渲染位置发射的光线数目；利用采样结果从划分结果中获取第一多尺度特征；采用第一多尺度特征进行隐式建模，得到建模结果。

可选的，上述处理器 1202 还可以执行如下步骤的程序代码：采用目标全连接神经网络模型对第一多尺度特征进行像素值预测，得到目标像素值，其中，目标全连接神经网络模型采用多组数据通过深度学习训练得到，多组数据包括：训练图像和训练图像对应的第二多尺度特征；基于目标像素值生成建模结果。

可选的，上述处理器 1202 还可以执行如下步骤的程序代码：采用初始全连接神经网络模型对第二多尺度特征进行像素值预测，得到训练像素值；通过训练像素值与训练图像对应的真实像素值确定第一损失；基于第一损失对初始全连接神经网络模型的模型参数进行更新，得到目标全连接神经网络模型。

可选的，上述处理器 1202 还可以执行如下步骤的程序代码：采用初始全连接神经网络模型对第二多尺度特征进行深度值预测，得到训练深度值；通过训练深度值与训练图像对应的真实深度值确定第二损失；采用采集视角和目标渲染位置确定采样射线；通过采样结果中的采样点在采样射线上的采样权重确定第三损失；基于第二损失与第三损失中至少之一对目标全连接神经网络模型的模型参数进行更新，得到更新后的目标全连接神经网络模型。

可选的，上述处理器 1202 还可以执行如下步骤的程序代码：在预设条件下，利用航拍图像采集设备对预设三维场景进行航拍数据采集，得到待重建图像数据，其中，预设条件被设置为确定航拍图像采集设备在航拍数据采集过程中的天气条件和飞行条件。

可选的，上述处理器 1202 还可以执行如下步骤的程序代码：对待重建图像数据进行特征提取，得到提取结果；对提取结果进行特征匹配，得到匹配结果；对匹配结果进行稀疏重建，得到重建结果，其中，重建结果被设置为确定待重建图像数据对应

的航拍图像采集设备的位姿信息。

可选的，上述处理器 1202 还可以执行如下步骤的程序代码：对待重建图像数据进行预处理，得到预处理结果；对预处理结果进行特征提取，得到提取结果。

可选的，上述处理器 1202 还可以执行如下步骤的程序代码：对提取结果进行图像配对，得到配对结果；对配对结果进行特征匹配，得到匹配结果。

可选的，上述处理器 1202 还可以执行如下步骤的程序代码：对匹配结果进行几何过滤，得到过滤结果；对过滤结果进行全局三维结构估计，得到估计结果；基于估计结果确定重建结果。

可选的，上述处理器 1202 还可以执行如下步骤的程序代码：对估计结果进行地平线矫正，得到第一矫正结果；对第一矫正结果进行航向角矫正，得到第二矫正结果；对第二矫正结果进行点云分块处理，得到分块结果；对分块结果进行数据格式调整，得到重建结果。

可选的，上述处理器 1202 还可以执行如下步骤的程序代码：获取航拍图像采集设备对应的轨迹信息；按照轨迹信息对建模结果进行视频渲染，得到目标视频。

处理器 1202 可以通过传输装置调用存储器存储的信息及应用程序，以执行下述步骤：响应作用于操作界面上的输入指令，在操作界面上显示待重建图像数据，其中，待重建图像数据被设置为记录待重建的预设三维场景的影像；响应作用于操作界面上的重建指令，在操作界面上显示目标视频，其中，目标视频被设置为对预设三维场景进行场景漫游，目标视频利用建模结果生成，建模结果通过对待重建图像数据和重建结果进行隐式建模得到，重建结果通过对待重建图像数据进行稀疏重建得到。

处理器 1202 可以通过传输装置调用存储器存储的信息及应用程序，以执行下述步骤：在虚拟现实 VR 设备或增强现实 AR 设备的呈现画面上展示初始默认画面；驱动 VR 设备或 AR 设备从展示初始默认画面跳转至展示目标视频，其中，目标视频被设置为对待重建的预设三维场景进行场景漫游，目标视频利用建模结果生成，建模结果通过对待重建图像数据和重建结果进行隐式建模得到，重建结果通过对待重建图像数据进行稀疏重建得到，待重建图像数据被设置为记录预设三维场景的影像。

采用本公开实施例，提供了一种场景重建的方案。通过获取待重建图像数据，其中，待重建图像数据被设置为记录待重建的预设三维场景的影像；进一步对待重建图像数据进行稀疏重建，得到重建结果；采用基于待重建图像数据和重建结果进行隐式建模的方式得到建模结果；进而利用建模结果生成目标视频，其中，目标视频被设置为对预设三维场景进行场景漫游。由此，本公开实施例达到了通过稀疏重建和隐式建模的方式基于待重建图像数据进行场景建模以渲染得到场景漫游视频的目的，从而实现了提升场景建模结果的准确度和场景漫游视频的真实感的技术效果，进而解决了由于现有技术采用显示建模方式进行场景重建导致的建模结果缺陷多、场景漫游视频的真实感差的技术问题。

本领域普通技术人员可以理解，图 12 所示的结构为示意，计算机终端也可以是智能手机（如 Android 手机、iOS 手机等）、平板电脑、掌上电脑以及移动互联网设备（Mobile Internet Devices, MID）等终端设备。图 12 其并不对上述计算机终端的结构造成限定。例如，计算机终端 120 还可包括比图 12 中所示更多或者更少的组件（如网络接口、显示装置等），或者具有与图 12 所示不同的配置。

本领域普通技术人员可以理解上述实施例的各种方法中的全部或部分步骤是可以通程序来指令终端设备相关的硬件来完成，该程序可以存储于一计算机可读存储介质中，存储介质可以包括：闪存盘、ROM、RAM、磁盘或光盘等。

实施例 6

根据本公开实施例，还提供了一种计算机可读存储介质。可选地，在本实施例中，上述存储介质可以被设置为保存上述实施例 1、实施例 2 或实施例 3 所提供的基于场

景重建生成视频的方法所执行的程序代码。

可选地，在本实施例中，上述存储介质可以位于计算机网络中计算机终端群中的任意一个计算机终端中，或者位于移动终端群中的任意一个移动终端中。

可选地，在本实施例中，计算机可读存储介质被设置为存储被设置为执行以下步骤的程序代码：获取待重建图像数据，其中，待重建图像数据被设置为记录待重建的预设三维场景的影像；对待重建图像数据进行稀疏重建，得到重建结果；基于待重建图像数据和重建结果进行隐式建模，得到建模结果；利用建模结果生成目标视频，其中，目标视频被设置为对预设三维场景进行场景漫游。

可选地，在本实施例中，计算机可读存储介质被设置为存储被设置为执行以下步骤的程序代码：基于待重建图像数据和重建结果，在多个尺度特征上构建预设三维场景对应的栅格，得到构建结果，其中，构建结果被设置为确定预设三维场景在多个尺度特征上的三维空间信息与光场信息；利用构建结果进行隐式建模，得到建模结果。

可选地，在本实施例中，计算机可读存储介质被设置为存储被设置为执行以下步骤的程序代码：基于待重建图像数据和重建结果确定预设三维场景对应的包围边界框；对包围边界框内的三维空间进行多尺度栅格化划分，得到划分结果；利用划分结果对应的光场信息，获取预设三维场景对应的栅格，得到构建结果。

可选地，在本实施例中，计算机可读存储介质被设置为存储被设置为执行以下步骤的程序代码：根据多个尺度特征，将构建结果中的三维空间信息与光场信息划分至多个分区，得到划分结果，其中，划分结果中的不同分区被设置为分别存储不同尺度特征的三维空间信息与光场信息；按照预设采样策略，基于航拍图像采集设备的采集视角和目标渲染位置对构建结果进行采样，得到采样结果，其中，预设采样策略被设置为确定目标渲染位置在采样过程中待使用的采样次数与采样步长，以及从目标渲染位置发射的光线数目；利用采样结果从划分结果中获取第一多尺度特征；采用第一多尺度特征进行隐式建模，得到建模结果。

可选地，在本实施例中，计算机可读存储介质被设置为存储被设置为执行以下步骤的程序代码：采用目标全连接神经网络模型对第一多尺度特征进行像素值预测，得到目标像素值，其中，目标全连接神经网络模型采用多组数据通过深度学习训练得到，多组数据包括：训练图像和训练图像对应的第二多尺度特征；基于目标像素值生成建模结果。

可选地，在本实施例中，计算机可读存储介质被设置为存储被设置为执行以下步骤的程序代码：采用初始全连接神经网络模型对第二多尺度特征进行像素值预测，得到训练像素值；通过训练像素值与训练图像对应的真实像素值确定第一损失；基于第一损失对初始全连接神经网络模型的模型参数进行更新，得到目标全连接神经网络模型。

可选地，在本实施例中，计算机可读存储介质被设置为存储被设置为执行以下步骤的程序代码：采用初始全连接神经网络模型对第二多尺度特征进行深度值预测，得到训练深度值；通过训练深度值与训练图像对应的真实深度值确定第二损失；采用采集视角和目标渲染位置确定采样射线；通过采样结果中的采样点在采样射线上的采样权重确定第三损失；基于第二损失与第三损失中至少之一对目标全连接神经网络模型的模型参数进行更新，得到更新后的目标全连接神经网络模型。

可选地，在本实施例中，计算机可读存储介质被设置为存储被设置为执行以下步骤的程序代码：在预设条件下，利用航拍图像采集设备对预设三维场景进行航拍数据采集，得到待重建图像数据，其中，预设条件被设置为确定航拍图像采集设备在航拍数据采集过程中的天气条件和飞行条件。

可选地，在本实施例中，计算机可读存储介质被设置为存储被设置为执行以下步骤的程序代码：对待重建图像数据进行特征提取，得到提取结果；对提取结果进行特

征匹配,得到匹配结果;对匹配结果进行稀疏重建,得到重建结果,其中,重建结果被设置为确定待重建图像数据对应的航拍图像采集设备的位姿信息。

可选地,在本实施例中,计算机可读存储介质被设置为存储被设置为执行以下步骤的程序代码:对待重建图像数据进行预处理,得到预处理结果;对预处理结果进行特征提取,得到提取结果。

可选地,在本实施例中,计算机可读存储介质被设置为存储被设置为执行以下步骤的程序代码:对提取结果进行图像配对,得到配对结果;对配对结果进行特征匹配,得到匹配结果。

可选地,在本实施例中,计算机可读存储介质被设置为存储被设置为执行以下步骤的程序代码:对匹配结果进行几何过滤,得到过滤结果;对过滤结果进行全局三维结构估计,得到估计结果;基于估计结果确定重建结果。

可选地,在本实施例中,计算机可读存储介质被设置为存储被设置为执行以下步骤的程序代码:对估计结果进行地平线矫正,得到第一矫正结果;对第一矫正结果进行航向角矫正,得到第二矫正结果;对第二矫正结果进行点云分块处理,得到分块结果;对分块结果进行数据格式调整,得到重建结果。

可选地,在本实施例中,计算机可读存储介质被设置为存储被设置为执行以下步骤的程序代码:获取航拍图像采集设备对应的轨迹信息;按照轨迹信息对建模结果进行视频渲染,得到目标视频。

可选地,在本实施例中,计算机可读存储介质被设置为存储被设置为执行以下步骤的程序代码:响应作用于操作界面上的输入指令,在操作界面上显示待重建图像数据,其中,待重建图像数据被设置为记录待重建的预设三维场景的影像;响应作用于操作界面上的重建指令,在操作界面上显示目标视频,其中,目标视频被设置为对预设三维场景进行场景漫游,目标视频利用建模结果生成,建模结果通过对待重建图像数据和重建结果进行隐式建模得到,重建结果通过对待重建图像数据进行稀疏重建得到。

可选地,在本实施例中,计算机可读存储介质被设置为存储被设置为执行以下步骤的程序代码:在虚拟现实 VR 设备或增强现实 AR 设备的呈现画面上展示初始默认画面;驱动 VR 设备或 AR 设备从展示初始默认画面跳转至展示目标视频,其中,目标视频被设置为对待重建的预设三维场景进行场景漫游,目标视频利用建模结果生成,建模结果通过对待重建图像数据和重建结果进行隐式建模得到,重建结果通过对待重建图像数据进行稀疏重建得到,待重建图像数据被设置为记录预设三维场景的影像。

上述本公开实施例序号为了描述,不代表实施例的优劣。

在本公开的上述实施例中,对各个实施例的描述都各有侧重,某个实施例中沒有详述的部分,可以参见其他实施例的相关描述。

在本公开所提供的几个实施例中,应该理解到,所揭露的技术内容,可通过其它的方式实现。其中,以上所描述的装置实施例是示意性的,例如所述单元的划分,为一种逻辑功能划分,实际实现时可以有另外的划分方式,例如多个单元或组件可以结合或者可以集成到另一个系统,或一些特征可以忽略,或不执行。另一点,所显示或讨论的相互之间的耦合或直接耦合或通信连接可以是通过一些接口,单元或模块的间接耦合或通信连接,可以是电性或其它的形式。

所述作为分离部件说明的单元可以是或者也可以不是物理上分开的,作为单元显示的部件可以是或者也可以不是物理单元,即可以位于一个地方,或者也可以分布到多个网络单元上。可以根据实际的需要选择其中的部分或者全部单元来实现本实施例方案的目的。

另外,在本公开各个实施例中的各功能单元可以集成在一个处理单元中,也可以是各个单元单独物理存在,也可以两个或两个以上单元集成在一个单元中。上述集成

的单元既可以采用硬件的形式实现，也可以采用软件功能单元的形式实现。

所述集成的单元如果以软件功能单元的形式实现并作为独立的产品销售或使用，可以存储在一个计算机可读取存储介质中。基于这样的理解，本公开的技术方案本质上或者说对现有技术做出贡献的部分或者该技术方案的全部或部分可以以软件产品的形式体现出来，该计算机软件产品存储在一个存储介质中，包括若干指令用以使得一台计算机设备（可为个人计算机、服务器或者网络设备等）执行本公开各个实施例所述方法的全部或部分步骤。而前述的存储介质包括：U 盘、ROM、RAM、移动硬盘、磁碟或者光盘等各种可以存储程序代码的介质。

以上所述是本公开的优选实施方式，应当指出，对于本技术领域的普通技术人员来说，在不脱离本公开原理的前提下，还可以做出若干改进和润饰，这些改进和润饰也应视为本公开的保护范围。

工业实用性

本公开实施例所提供的基于场景重建生成视频的方法、存储介质和系统，通过获取待重建图像数据，其中，待重建图像数据被设置为记录待重建的预设三维场景的影像；进一步对待重建图像数据进行稀疏重建，得到重建结果；采用基于待重建图像数据和重建结果进行隐式建模的方式得到建模结果；进而利用建模结果生成目标视频，其中，目标视频被设置为对预设三维场景进行场景漫游。由此，本公开实施例达到了通过稀疏重建和隐式建模的方式基于待重建图像数据进行场景建模以渲染得到场景漫游视频的目的，从而实现了提升场景建模结果的准确度和场景漫游视频的真实感的技术效果，进而解决了由于现有技术采用显示建模方式进行场景重建导致的建模结果缺陷多、场景漫游视频的真实感差的技术问题。

权 利 要 求 书

1. 一种基于场景重建生成视频的方法，包括：
 获取待重建图像数据，其中，所述待重建图像数据用于记录待重建的预设三维场景的影像；
 对所述待重建图像数据进行稀疏重建，得到重建结果；
 基于所述待重建图像数据和所述重建结果进行隐式建模，得到建模结果；
 利用所述建模结果生成目标视频，其中，所述目标视频用于对所述预设三维场景进行场景漫游。
2. 根据权利要求1所述的方法，其中，基于所述待重建图像数据和所述重建结果进行隐式建模，得到所述建模结果包括：
 基于所述待重建图像数据和所述重建结果，在多个尺度特征上构建所述预设三维场景对应的栅格，得到构建结果，其中，所述构建结果被设置为确定所述预设三维场景在所述多个尺度特征上的三维空间信息与光场信息；
 利用所述构建结果进行隐式建模，得到所述建模结果。
3. 根据权利要求2所述的方法，其中，基于所述待重建图像数据和所述重建结果，在所述多个尺度特征上构建所述预设三维场景对应的栅格，得到所述构建结果包括：
 基于所述待重建图像数据和所述重建结果确定所述预设三维场景对应的包围边界框；
 对所述包围边界框内的三维空间进行多尺度栅格化划分，得到划分结果；
 利用所述划分结果对应的光场信息，获取所述预设三维场景对应的栅格，得到所述构建结果。
4. 根据权利要求2所述的方法，其中，利用所述构建结果进行隐式建模，得到所述建模结果包括：
 根据所述多个尺度特征，将所述构建结果中的三维空间信息与光场信息划分至多个分区，得到划分结果，其中，所述划分结果中的不同分区被设置为分别存储不同尺度特征的三维空间信息与光场信息；
 按照预设采样策略，基于航拍图像采集设备的采集视角和目标渲染位置对所述构建结果进行采样，得到采样结果，其中，所述预设采样策略被设置为确定所述目标渲染位置在采样过程中待使用的采样次数与采样步长，以及从所述目标渲染位置发射的光线数目；
 利用所述采样结果从所述划分结果中获取第一多尺度特征；
 采用所述第一多尺度特征进行隐式建模，得到所述建模结果。
5. 根据权利要求4所述的方法，其中，采用所述第一多尺度特征进行隐式建模，得到所述建模结果包括：
 采用目标全连接神经网络模型对所述第一多尺度特征进行像素值预测，得到目标像素值，其中，所述目标全连接神经网络模型采用多组数据通过深度学习训练得到，所述多组数据包括：训练图像和所述训练图像对应的第二多尺度特征；
 基于所述目标像素值生成所述建模结果。
6. 根据权利要求5所述的方法，其中，所述方法还包括：
 采用初始全连接神经网络模型对所述第二多尺度特征进行像素值预测，得到训练像素值；
 通过所述训练像素值与所述训练图像对应的真实像素值确定第一损失；
 基于所述第一损失对所述初始全连接神经网络模型的模型参数进行更新，得到所述目标全连接神经网络模型。
7. 根据权利要求6所述的方法，其中，所述场景重建方法还包括：

- 采用所述初始全连接神经网络模型对所述第二多尺度特征进行深度值预测,得到训练深度值;
- 通过所述训练深度值与所述训练图像对应的真实深度值确定第二损失;
- 采用所述采集视角和所述目标渲染位置确定采样射线;
- 通过所述采样结果中的采样点在所述采样射线上的采样权重确定第三损失;
- 基于所述第二损失与所述第三损失中至少之一对所述目标全连接神经网络模型的模型参数进行更新,得到更新后的所述目标全连接神经网络模型。
8. 根据权利要求1所述的方法,其中,获取所述待重建图像数据包括:
- 在预设条件下,利用航拍图像采集设备对所述预设三维场景进行航拍数据采集,得到所述待重建图像数据,其中,所述预设条件被设置为确定所述航拍图像采集设备在航拍数据采集过程中的天气条件和飞行条件。
9. 根据权利要求8所述的方法,其中,对所述待重建图像数据进行稀疏重建,得到所述重建结果包括:
- 对所述待重建图像数据进行特征提取,得到提取结果;
- 对所述提取结果进行特征匹配,得到匹配结果;
- 对所述匹配结果进行稀疏重建,得到所述重建结果,其中,所述重建结果被设置为确定所述待重建图像数据对应的所述航拍图像采集设备的位姿信息。
10. 根据权利要求8所述的基于场景重建生成视频的方法,其中,利用所述建模结果生成所述目标视频包括:
- 获取所述航拍图像采集设备对应的轨迹信息;
- 按照所述轨迹信息对所述建模结果进行视频渲染,得到所述目标视频。
11. 一种基于场景重建生成视频的基于场景重建生成视频的方法,包括:
- 响应作用于操作界面上的输入指令,在所述操作界面上显示待重建图像数据,其中,所述待重建图像数据被设置为记录待重建的预设三维场景的影像;
- 响应作用于所述操作界面上的重建指令,在所述操作界面上显示目标视频,其中,所述目标视频被设置为对所述预设三维场景进行场景漫游,所述目标视频利用建模结果生成,所述建模结果通过对所述待重建图像数据和重建结果进行隐式建模得到,所述重建结果通过对所述待重建图像数据进行稀疏重建得到。
12. 一种基于场景重建生成视频的基于场景重建生成视频的方法,包括:
- 在虚拟现实VR设备或增强现实AR设备的呈现画面上展示初始默认画面;
- 驱动所述VR设备或所述AR设备从展示所述初始默认画面跳转至展示目标视频,其中,所述目标视频被设置为对待重建的预设三维场景进行场景漫游,所述目标视频利用建模结果生成,所述建模结果通过对待重建图像数据和重建结果进行隐式建模得到,所述重建结果通过对所述待重建图像数据进行稀疏重建得到,所述待重建图像数据被设置为记录所述预设三维场景的影像。
13. 一种计算机可读存储介质,其中,所述计算机可读存储介质包括存储的可执行程序,其中,在所述可执行程序运行时控制所述计算机可读存储介质所在设备执行权利要求1至12中任意一项所述的基于场景重建生成视频的基于场景重建生成视频的方法。
14. 一种基于场景重建生成视频的系统,包括:
- 处理器;
- 存储器,与所述处理器相连接,被设置为为所述处理器提供处理以下处理步骤的指令:
- 获取待重建图像数据,其中,所述待重建图像数据被设置为记录待重建的预设三维场景的影像;
- 对所述待重建图像数据进行稀疏重建,得到重建结果;
- 基于所述待重建图像数据和所述重建结果进行隐式建模,得到建模结果;

利用所述建模结果生成目标视频，其中，所述目标视频被设置为对所述预设三维场景进行场景漫游。