

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2025 年 1 月 9 日 (09.01.2025)



(10) 国际公布号
WO 2025/007702 A1

- (51) 国际专利分类号:
G06F 3/01 (2006.01) **G06T 17/05** (2011.01)
- (21) 国际申请号: PCT/CN2024/098014
- (22) 国际申请日: 2024 年 6 月 7 日 (07.06.2024)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
202310822219.9 2023年7月5日 (05.07.2023) CN
- (71) 申请人: 北京字跳网络技术有限公司
(**BEIJING ZITIAO NETWORK TECHNOLOGY CO., LTD.**) [CN/CN]; 中国北京市海淀区紫金数码园4号楼2层0207, Beijing 100190 (CN)。
- (72) 发明人: 范帝楷 (**FAN, Dikai**); 中国北京市朝阳区七圣中街12号院融中心B1小邮局, Beijing 100028 (CN)。
- (74) 代理人: 中国贸促会专利商标事务所有限公司(**CCPIT PATENT AND TRADEMARK LAW OFFICE**); 中国北京市西城区复兴门内大街158号远洋大厦10层, Beijing 100031 (CN)。

- (81) 指定国(除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW。
- (84) 指定国(除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, ME, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

(54) **Title:** MESH CONSTRUCTION METHOD AND APPARATUS, AND ELECTRONIC DEVICE, MEDIUM AND PROGRAM PRODUCT

(54) 发明名称: 网格构建方法、装置、电子设备、介质和程序产品

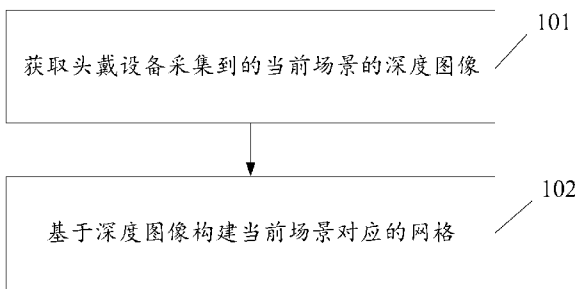


图 1

- 101 Acquire a depth image of the current scene collected by a head-mounted device
- 102 On the basis of the depth image, construct a mesh corresponding to the current scene

(57) **Abstract:** Disclosed in the embodiments of the present disclosure are a mesh construction method and apparatus, and an electronic device, a medium and a program product. The mesh construction method is applied to a head-mounted device. The mesh construction method comprises: acquiring a depth image of the current scene collected by a head-mounted device; and on the basis of the depth image, constructing a mesh corresponding to the current scene, wherein the depth image comprises a target pixel point, the actual depth of a target object corresponding to the target pixel point meets a preset condition in the current scene, and during the construction of the mesh, a depth value of the target pixel point is a preset depth value.

(57) 摘要: 本公开实施例公开了网格构建方法、装置、电子设备、介质和程序产品。网格构建方法, 应用于头戴设备, 所述网格构建方法包括: 获取头戴设备采集到的当前场景的深度图像; 基于所述深度图像构建所述当前场景对应的网格; 其中, 所述深度图像包括目标像素点, 所述目标像素点对应的目标对象的实际深度在所述当前场景中符合预设条件; 在构建所述网格的过程中, 所述目标像素点的深度值采用预设深度值。

本国际公布:

— 包括国际检索报告(条约第21条(3))。

网格构建方法、装置、电子设备、介质和程序产品

相关技术的交叉引用

本申请是以申请号为 202310822219.9、申请日为 2023 年 7 月 5 日的中国专利申请
5 为基础，并主张其优先权，该中国专利申请的公开内容在此作为整体引入本申请中。

技术领域

本公开涉及三维重建技术领域，尤其涉及一种网格构建方法、装置和电子设备。

10 背景技术

随着扩展现实技术的发展，各种扩展现实设备也随之发展。对于扩展现实设备来说，具备透视功能。透视功能可以实现通过扩展现实设备看到真实的世界，因此，在开启透视功能时，会对用户当前所能看到的场景进行三维重建，以得到该场景对应的网格，即三维场景模型。

15

发明内容

提供该公开内容部分以便以简要的形式介绍构思，这些构思将在后面的具体实施方式部分被详细描述。该公开内容部分并不旨在标识要求保护的技术方案的关键特征或必要特征，也不旨在用于限制所要求的保护的技术方案的范围。

20

本公开实施例提供了一种网格构建方法、装置和电子设备。

第一方面，本公开实施例提供了一种网格构建方法，应用于头戴设备，所述网格构建方法包括：获取头戴设备采集到的当前场景的深度图像；基于所述深度图像构建所述当前场景对应的网格；其中，所述深度图像包括目标像素点，所述目标像素点对应的目标对象的实际深度在所述当前场景中符合预设条件；在构建所述网格的过程中，所述目标像素点的深度值采用预设深度值。

25

第二方面，本公开实施例提供了一种网格构建装置，应用于头戴设备，所述网格构建装置包括：获取单元，用于获取头戴设备采集到的当前场景的深度图像；构建单元，用于基于所述深度图像构建所述当前场景对应的网格；其中，所述深度图像包括目标像素点，所述目标像素点对应的目标对象的实际深度在所述当前场景中符合预设条件；在构建所述网格的过程中，所述目标像素点的深度值采用预设深度值。

30

第三方面，本公开实施例提供了一种电子设备，包括：一个或多个处理器；存储装置，

用于存储一个或多个程序，当所述一个或多个程序被所述一个或多个处理器执行，使得所述一个或多个处理器实现如第一方面所述的网格构建方法。

第四方面，本公开实施例提供了一种计算机可读介质，其上存储有计算机程序，该程序被处理器执行时实现如第一方面所述的网格构建方法。

5

附图说明

结合附图并参考以下具体实施方式，本公开各实施例的上述和其他特征、优点及方面将变得更加明显。贯穿附图中，相同或相似的附图标记表示相同或相似的元素。应当理解附图是示意性的，原件和元素不一定按照比例绘制。

10

图 1 是根据本公开的网格构建方法的一个实施例的流程图；

图 2 是根据本公开的网格构建方法的另一个实施例的流程图；

图 3 是根据本公开的网格构建装置的一个实施例的结构示意图；

图 4 是本公开的一个实施例的网格构建方法可以应用于其中的示例性系统架构；

图 5 是根据本公开实施例提供的电子设备的基本结构的示意图。

15

具体实施方式

下面将参照附图更详细地描述本公开的实施例。虽然附图中显示了本公开的某些实施例，然而应当理解的是，本公开可以通过各种形式来实现，而且不应该被解释为限于这里阐述的实施例，相反提供这些实施例是为了更加透彻和完整地理解本公开。应当理解的是，

20

本公开的附图及实施例仅用于示例性作用，并非用于限制本公开的保护范围。

应当理解，本公开的方法实施方式中记载的各个步骤可以按照不同的顺序执行，和/或并行执行。此外，方法实施方式可以包括附加的步骤和/或省略执行示出的步骤。本公开的范围在此方面不受限制。

25

本文使用的术语“包括”及其变形是开放性包括，即“包括但不限于”。术语“基于”是“至少部分地基于”。术语“一个实施例”表示“至少一个实施例”；术语“另一实施例”表示“至少一个另外的实施例”；术语“一些实施例”表示“至少一些实施例”。其他术语的相关定义将在下文描述中给出。

需要注意，本公开中提及的“第一”、“第二”等概念仅用于对不同的装置、模块或单元进行区分，并非用于限定这些装置、模块或单元所执行的功能的顺序或者相互依存关系。

30

需要注意，本公开中提及的“一个”、“多个”的修饰是示意性而非限制性的，本领域技术人员应当理解，除非在上下文另有明确指出，否则应该理解为“一个或多个”。

本公开实施方式中的多个装置之间所交互的消息或者信息的名称仅用于说明性的目的，而并不是用于对这些消息或信息的范围进行限制。

本公开实施方式的技术方案可以应用于各种扩展现实场景中，或者应用于其他涉及到透视功能的应用场景中。

5 在这些应用场景中，当开启透视功能时，对用户当前所能看到的场景进行三维重建，生成三维场景模型，生成的三维场景模型在本公开实施例中称之为网格。

随着头戴设备和扩展现实技术的发展，越来越多的用户希望能在室外场景中使用头戴设备，做一些户外娱乐，例如：真人户外游戏；也有不少的商家希望头戴设备开发更多的扩展现实效果，例如：在游乐场中显示虚拟形象。即，目前，存在着在户外场景中构建环境网格的需求。

10 但是，由于户外场景面临尺度较大的问题，可能无法像室内一样实时构建当前视角下的环境网格，以满足透视使用。

基于此，本公开实施例提供一种网格构建方案，该网格构建方案考虑到尺度较大的环境的深度准确性差的问题，该部分若直接进行网格的构建，由于其深度的不准确性会导致网格构建的不准确性。因此，利用预设深度值实现该部分像素点的网格构建，能够解决深度的不准确性带来的问题，保证环境网格构建的准确性；从而使得构建的网格可以满足透视使用。进而，该网格构建方案可以应用于基于户外场景的网格构建。

请参考图 1，其示出了根据本公开的网格构建方法的一个实施例的流程。该网格构建方法可以应用于终端设备。如图 1 所示，该网格构建方法包括以下步骤：

20 步骤 101，获取头戴设备采集到的当前场景的深度图像。

在一些实施例中，头戴设备包括深度传感器，该深度传感器可用于实现深度图像的获取。深度传感器，例如：TOF（Time of flight，飞行时间）相机。

在一些实施例中，头戴设备包括双目相机，该双目相机可基于双目识别实现深度图像的获取。

25 在一些实施例中，当前场景为户外场景；在另一些实施例中，当前场景也可以为非户外场景，但是当前场景中存在深度难以估计的对象。

因此，在一些实施例中，深度图像包括目标像素点，该目标像素点对应的目标对象的实际深度在当前场景中符合预设条件。

预设条件例如，实际深度难以预测、实际深度大于或远大于预设深度等。

30 目标对象例如，天空、玻璃，或者其他深度难以预测的对象，在此不作限定。

在一些实施例中，由于目标像素点是目标对象对应的像素点，所以需要先对目标对象

进行检测，才能确定目标像素点。

因此，作为一种可选的实施方式，该网格构建方法还包括：确定深度图像对应的目标对象检测结果；根据目标对象检测结果确定目标像素点。

可以理解，由于目标对象的深度难以预测，其深度值的准确性难以得到保证，如直接
5 基于深度图像对目标对象进行检测，检测结果的准确性可能也不够高。

因此，作为一种可选的实施方式，确定深度图像对应的目标对象检测结果，包括：获取深度图像对应的环境图像；该环境图像包括与深度图像的像素点一一对应的像素点；基于环境图像，确定目标对象检测结果。

在一些实施例中，环境图像可以是彩色图像、灰度图像等，区别于深度图像的图像。

10 在一些实施例中，环境图像和深度图像均可以通过双目相机获取；在另一些实施例中，环境图像和深度图像可以采用不同的相机获取；例如，深度图像通过双目相机获取，环境图像通过灰度相机获取；或者环境图像通过彩色相机获取等。

在一些实施例中，环境图像包括与深度图像的像素点一一对应的像素点，则，灰度图像与深度图像之间具有像素点对应关系，从而，可以基于环境图像的目标对象检测结果，
15 确定深度图像的目标对象检测结果。

在一些实施例中，基于环境图像和预先训练的神经网络模型，实现目标对象检测结果
的确定。

其中，预先训练的神经网络模型对应的训练数据集可以是标注有目标对象的图像，利用这些图像对神经网络模型进行训练，使得训练的神经网络模型可用于检测目标对象。

20 在一些实施例中，在神经网络模型的训练过程中，可以采用一些能够提高神经网络模型的精度的实施方式，例如：预设训练次数，在达到训练次数后，视为完成模型的训练。再例如：预设测试数据集，利用测试数据集对训练的神经网络模型的精度进行测试，若精度未达到标准，则继续对模型进行训练，或者改进模型等。

在一些实施例中，目标对象检测结果既可以用于表征环境图像中是否包括目标对象，
25 还可以在包括目标对象的情况下，指示目标对象所在的区域，或者说目标对象对应的目标像素点。

从而，基于环境图像的目标对象检测结果，可以确定深度图像中的目标对象对应的目标像素点。

步骤 102，基于深度图像构建当前场景对应的网格。

30 其中，在构建网格的过程中，目标像素点的深度值采用预设深度值。

在一些实施例中，由于目标像素点的深度值无法预测，而在基于深度图像构建当前对

应的网格的过程中，需要应用到目标像素点的深度值。因此，在本公开实施例中，采用预设深度值作为目标像素点的深度值。

在一些实施例中，结合不同的目标对象，预设深度值可以不同。例如，对于天空这种对象，预设深度值可以为 10，或者其他的预设的最远深度。

5 在一些实施例中，预设深度值还可以结合人眼瞳距选取，在此不对具体值作限定。

在一些实施例中，为了避免后续基于深度图像构建网格时出现错误，在深度图像中，可以将目标像素点对应的深度值标记为无效深度值，后续在应用时，将这些无效深度值替换为预设深度值即可。无效深度值，例如-1。

10 作为一种可选的实施方式，步骤 102 包括：基于深度图像中各个像素点的深度值和各个像素点的像素坐标，确定各个像素点的球面坐标；基于各个像素点的球面坐标，构建当前场景对应的网格。

在这种实施方式中，先基于各个像素点的深度值和各个像素点的像素坐标，将各个像素点的像素坐标转换为球面坐标；从而，再根据各个像素点的球面坐标，构建当前场景对应的网格。可以理解，对于目标像素点的深度值，采用预设深度值。

15 在一些实施例中，假设网格所在坐标系为 I，图像所在坐标系为 S，用 R_{IS}, t_{IS} 表征将图像坐标系下的像素点转换到网格坐标系下，用 $\pi^{-1}(\cdot)$ 表征将图像的像素坐标投影到图像坐标系下的归一化坐标 ($z=1$) 上，则，基于深度图像，将深度图中的点投影到网格坐标系下，可表示为： $P_I = R_{IS}\pi^{-1}(px)D(px) + t_{IS}, if D(px) > 0$ 。

20 进一步地，用 $D(px)$ 表征深度图像上像素坐标为 px 的深度值，对于目标像素点，即 $D(px) = -1$ 的像素点（即无效的像素点的深度在深度图像中置为-1），对应的转换关系为： $P_I = R_{IS}normalize(\pi^{-1}(px)) * 10$ 。此处，预设深度取值为 10，以及， $normalize(\cdot)$ 表征把向量单位化。

进一步地，再将网格所在坐标系下的点转换为球面坐标： $\begin{pmatrix} \tilde{x}_I \\ \tilde{y}_I \\ \tilde{z}_I \end{pmatrix} = normalize \left(\begin{pmatrix} x_I \\ y_I \\ z_i \end{pmatrix} \right)$,

25 $\begin{pmatrix} \phi \\ \theta \\ \rho \end{pmatrix} = \begin{pmatrix} \arccos \tilde{z}_I \\ \arctan \frac{\tilde{x}_I}{\tilde{y}_I} \\ 1.0/|P_I| \end{pmatrix}$ 。其中， $|P_I|$ 表征向量 P_I 的模长。

在一些实施例中，基于各个像素点的球面坐标，可实现当前场景对应的网格的构建，但是，此时构建的网格可能存在着畸变的可能性。因此，为了解决网格的畸变问题，在本公开实施方式中，可通过球面采样的方式，构建环境网格。

作为一种可选的实施方式，基于各个像素点的球面坐标，构建当前场景对应的网格，包括：基于各个像素点的球面坐标，生成各个像素点对应的球面；基于球面进行采样，确定多个采样点的网格坐标；基于多个采样点的网格坐标，构建当前场景对应的网格。

在一些实施例中，假设 P_i 表示球面网格上的第 i 个采样点，则其球面坐标可以表示为：

$$5 \quad \begin{pmatrix} \phi_i \\ \theta_i \\ \rho_i \end{pmatrix}。$$

进一步地，确定多个采样点的网格坐标，包括：确定多个采样点分别对应的渲染逆深度；其中，该渲染逆深度为基于邻域采样点的深度值确定的深度值；基于多个采样点分别对应的渲染逆深度和多个采样点分别对应的球面坐标，确定多个采样点分别对应的网格坐标。

10 在一些实施例中，对于各个采样点来说，需要利用渲染逆深度重新求解网格坐标，该渲染逆深度需要结合采样点的邻域采样点的深度值确定。

在一些实施例中，针对多个采样点中的采样点，基于该采样点的邻域采样点的深度值确定该采样点的平均逆深度值和权重影响值；基于该采样点的平均逆深度值、该采样点的权重影响值、初始渲染逆深度和预设的能量关系，确定该采样点的渲染逆深度。

15 在一些实施例中，邻域采样点，例如：3*3 范围内的邻域采样点。

在一些实施例中，预设的能量关系，用于表征基于球面网格构建的能量方程。

在一些实施例中，该采样点的平均逆深度值，可以理解为基于权重影响值，对邻域采样点的深度值求平均值获得的深度值。

20 在一些实施例中，若邻域采样点对应的像素点为目标像素点，则该邻域采样点的深度值为预设深度值的倒数。若邻域采样点对应的像素点不是目标像素点，则深度值采用深度图像中的深度值。

在一些实施例中，权重影响值基于该采样点的邻域采样点的数量确定。

25 在一些实施例中，平均逆深度值表示为： $\bar{\rho}_i = \frac{\sum(\rho > 0.1? \rho : 0.1)}{count}$ ，权重影响值表示为： $w_i = count$ 。其中， $count$ 表示这个点的邻域采样点的数量，表示如果邻域采样点的深度大于 0.1（即预设深度的倒数），则用原来的深度值，否则取 0.1。即，针对目标像素点对应的邻域采样点，其深度值采用 0.1，而针对其他像素点对应的邻域采样点，其深度值采用本来的深度值。

在一些实施例中，预设的能量关系可表示为： $E = \sum_i w_i (\rho_i - \bar{\rho}_i)^2 + \sum_i \sum_{j \in N(i)} (\rho_i - \rho_j)^2$ ，其中， $N(i)$ 表示和网格采样点 i 直接相连的采样点的集合，也即邻域采样点集合， ρ_i 表示渲染

染逆深度，其可以对应一个初始值，然后基于该初始值进行优化。

基于该预设的能量关系，对能量进行优化，优化完成的能量对应的渲染逆深度，即为最终确定的渲染逆深度。

从而，各个采样点对应的最终网格坐标表示为：
$$\begin{pmatrix} x_i \\ y_i \\ z_i \end{pmatrix} = \rho_i \begin{pmatrix} \sin \phi \sin \theta \\ \sin \phi \cos \theta \\ \cos \phi \end{pmatrix}$$
。基于该网格

5 采样点的三维坐标，最终可得到用于渲染的网格模型。

在一些实施例中，在对能量进行优化时，可预设优化条件，例如：前后两次确定的能量之间的差值小于预设差值，即能量基本不再变化；再例如，最终确定的能量值符合预设的能量值条件等，在此不作限定。

10 进一步参考图 2，本公开提供一种三维建模的应用流程图，在图 2 中，先基于灰度图实现天空分割，然后基于与灰度图对应的深度图实现天空深度的剔除，接着基于天空剔除的结果进行网格投影与求解，最终完成三维模型的渲染。其中，天空深度的剔除，本质上是将天空部分的深度值进行标识，以在网格投影与求解的过程中，对这部分深度值采用特殊的处理方式，以避免对最终的网格构建结果的影响。

15 从而，基于头戴设备采集到的当前场景的深度图像构建当前场景对应的网格；其中，该深度图像包括实际深度在当前场景中符合预设条件的目标对象，该部分若直接进行网格的构建，由于其深度的不准确性会导致网格构建的不准确性。因此，利用预设深度值实现该部分像素点的网格构建，能够解决深度的不准确性带来的问题，保证环境网格构建的准确性；从而使得构建的网格可以满足透视使用。

20 以及，在基于深度图像进行网格构建的过程中，利用邻域采样的方式，解决网格的畸变问题，使得最终得到的网格平滑性较好，提高最终的渲染模型的效果。

进一步参考图 3，作为对上述各图所示方法的实现，本公开提供了一种网格构建装置的一个实施例，该装置实施例与图 1 所示的网格构建方法实施例相对应，该装置具体可以应用于各种电子设备中。

25 如图 3 所示，本实施例的网格构建装置包括：获取单元 301，用于获取头戴设备采集到的当前场景的深度图像；所述深度图像包括目标像素点，所述目标像素点对应的目标对象的实际深度在所述当前场景中符合预设条件；构建单元 302，用于基于所述深度图像构建所述当前场景对应的网格；其中，在构建所述网格的过程中，所述目标像素点的深度值采用预设深度值。

30 在一些实施例中，构建单元 302 还用于确定所述深度图像对应的目标对象检测结果；

根据所述目标对象检测结果确定所述目标像素点。

在一些实施例中，构建单元 302 进一步用于获取所述深度图像对应的环境图像；所述环境图像包括与所述深度图像的像素点一一对应的像素点；基于所述环境图像，确定所述目标对象检测结果。

5 在一些实施例中，构建单元 302 进一步用于基于所述深度图像中各个像素点的深度值和各个像素点的像素坐标，确定所述各个像素点的球面坐标；基于所述各个像素点的球面坐标，构建所述当前场景对应的网格。

在一些实施例中，构建单元 302 进一步用于基于所述各个像素点的球面坐标，生成所述各个像素点对应的球面；基于所述球面进行采样，确定多个采样点的网格坐标；

10 基于所述多个采样点的网格坐标，构建所述当前场景对应的网格。

在一些实施例中，构建单元 302 进一步用于确定所述多个采样点分别对应的渲染逆深度；其中，该渲染逆深度为基于邻域采样点的深度值确定的深度值；基于所述多个采样点分别对应的渲染逆深度和所述多个采样点分别对应的球面坐标，确定所述多个采样点分别对应的网格坐标。

15 在一些实施例中，构建单元 302 还用于：针对所述多个采样点中的采样点，基于该采样点的邻域采样点的深度值确定该采样点的平均逆深度值和权重影响值；基于该采样点的平均逆深度值、该采样点的权重影响值、初始渲染逆深度和预设的能量关系，确定该采样点的渲染逆深度。

20 在一些实施例中，若邻域采样点对应的像素点为所述目标像素点，则该邻域采样点的深度值为所述预设深度值的倒数。

在一些实施例中，所述权重影响值基于该采样点的邻域采样点的数量确定。

请参考图 4，图 4 示出了本公开的一个实施例的网格构建方法可以应用于其中的示例性系统架构。

25 如图 4 所示，系统架构可以包括终端设备 401、402、403，网络 404，服务器 405。网络 404 可以用以在终端设备 401、402、403 和服务器 405 之间提供通信链路的介质。网络 404 可以包括各种连接类型，例如有线、无线通信链路或者光纤电缆等等。

终端设备 401、402、403 可以通过网络 404 与服务器 405 交互，以接收或发送消息等。终端设备 401、402、403 上可以安装有各种客户端应用，例如网页浏览器应用、
30 搜索类应用、新闻资讯类应用。终端设备 401、402、403 中的客户端应用可以接收用户的指令，并根据用户的指令完成相应的功能，例如根据用户的指令在信息中添加相

应信息。

终端设备 401、402、403 可以是硬件，也可以是软件。当终端设备 401、402、403 为硬件时，可以是具有显示屏并且支持网页浏览的各种电子设备，包括但不限于智能手机、平板电脑、电子书阅读器、MP3 播放器（Moving Picture Experts Group Audio Layer III，动态影像专家压缩标准音频层面 3）、MP4（Moving Picture Experts Group Audio Layer IV，动态影像专家压缩标准音频层面 4）播放器、膝上型便携计算机和台式计算机等等。当终端设备 401、402、403 为软件时，可以安装在上述所列举的电子设备中。其可以实现成多个软件或软件模块（例如用来提供分布式服务的软件或软件模块），也可以实现成单个软件或软件模块。在此不做具体限定。

服务器 405 可以是提供各种服务的服务器，例如接收终端设备 401、402、403 发送的信息获取请求，根据信息获取请求通过各种方式获取信息获取请求对应的展示信息。并展示信息的相关数据发送给终端设备 401、402、403。

需要说明的是，本公开实施例所提供的网格构建方法可以由终端设备执行，相应地，网格构建装置可以设置在终端设备 401、402、403 中。此外，本公开实施例所提供的网格构建方法还可以由服务器 405 执行，相应地，网格构建装置可以设置于服务器 405 中。

应该理解，图 4 中的终端设备、网络和服务器的数目仅仅是示意性的。根据实现需要，可以具有任意数目的终端设备、网络和服务器的。

下面参考图 5，其示出了适于用来实现本公开实施例的电子设备的结构示意图。本公开实施例中的终端设备可以包括但不限于诸如移动电话、笔记本电脑、数字广播接收器、PDA（个人数字助理）、PAD（平板电脑）、PMP（便携式多媒体播放器）、车载终端（例如车载导航终端）等等的移动终端以及诸如数字 TV、台式计算机等等的固定终端。图 5 示出的电子设备仅仅是一个示例，不应对本公开实施例的功能和使用范围带来任何限制。

如图 5 所示，电子设备可以包括处理装置（例如中央处理器、图形处理器等）501，其可以根据存储在只读存储器（ROM）502 中的程序或者从存储装置 508 加载到随机访问存储器（RAM）503 中的程序而执行各种适当的动作和处理。在 RAM 503 中，还存储有电子设备 500 操作所需的各种程序和数据。处理装置 501、ROM 502 以及 RAM 503 通过总线 504 彼此相连。输入/输出（I/O）接口 505 也连接至总线 504。

通常，以下装置可以连接至 I/O 接口 505：包括例如触摸屏、触摸板、键盘、鼠

标、摄像头、麦克风、加速度计、陀螺仪等的输入装置 506；包括例如液晶显示器(LCD)、扬声器、振动器等的输出装置 507；包括例如磁带、硬盘等的存储装置 508；以及通信装置 509。通信装置 509 可以允许电子设备与其他设备进行无线或有线通信以交换数据。虽然图 5 示出了具有各种装置的电子设备，但是应理解的是，并不要求实施或具

5 备所有示出的装置。可以替代地实施或具备更多或更少的装置。

特别地，根据本公开的实施例，上文参考流程图描述的过程可以被实现为计算机软件程序。例如，本公开的实施例包括一种计算机程序产品，其包括承载在非暂态计算机可读介质上的计算机程序，该计算机程序包含用于执行流程图所示的方法的程序代码。在这样的实施例中，该计算机程序可以通过通信装置 509 从网络上被下载和安

10 装，或者从存储装置 508 被安装，或者从 ROM502 被安装。在该计算机程序被处理装置 501 执行时，执行本公开实施例的方法中限定的上述功能。

需要说明的是，本公开上述的计算机可读介质可以是计算机可读信号介质或者计算机可读存储介质或者是上述两者的任意组合。计算机可读存储介质例如可以是——但不限于——电、磁、光、电磁、红外线、或半导体的系统、装置或器件，或者任意以

15 上的组合。计算机可读存储介质的更具体的例子可以包括但不限于：具有一个或多个导线的电连接、便携式计算机磁盘、硬盘、随机访问存储器(RAM)、只读存储器(ROM)、可擦式可编程只读存储器(EPROM 或闪存)、光纤、便携式紧凑磁盘只读存储器(CD-ROM)、光存储器件、磁存储器件、或者上述的任意合适的组合。在本公开中，计算机可读存储介质可以是任何包含或存储程序的有形介质，该程序可以被指令执行系统、

20 装置或者器件使用或者与其结合使用。而在本公开中，计算机可读信号介质可以包括在基带中或者作为载波一部分传播的数据信号，其中承载了计算机可读的程序代码。这种传播的数据信号可以采用多种形式，包括但不限于电磁信号、光信号或上述的任意合适的组合。计算机可读信号介质还可以是计算机可读存储介质以外的任何计算机可读介质，该计算机可读信号介质可以发送、传播或者传输用于由指令执行系统、装

25 置或者器件使用或者与其结合使用的程序。计算机可读介质上包含的程序代码可以用任何适当的介质传输，包括但不限于：电线、光缆、RF（射频）等等，或者上述的任意合适的组合。

在一些实施方式中，客户端、服务器可以利用诸如 HTTP（HyperText Transfer Protocol，超文本传输协议）之类的任何当前已知或未来研发的网络协议进行通信，并

30 且可以与任意形式或介质的数字数据通信（例如，通信网络）互连。通信网络的示例包括局域网（“LAN”），广域网（“WAN”），网际网（例如，互联网）以及端对端网络

(例如, ad hoc 端对端网络), 以及任何当前已知或未来研发的网络。

上述计算机可读介质可以是上述电子设备中所包含的; 也可以是单独存在, 而未装配入该电子设备中。

上述计算机可读介质承载有一个或者多个程序, 当上述一个或者多个程序被该电子设备执行时, 使得该电子设备: 获取头戴设备采集到的当前场景的深度图像; 所述深度图像包括目标像素点, 所述目标像素点对应的目标对象的实际深度在所述当前场景中符合预设条件; 基于所述深度图像构建所述当前场景对应的网格; 其中, 在构建所述网格的过程中, 所述目标像素点的深度值采用预设深度值。

可以以一种或多种程序设计语言或其组合来编写用于执行本公开的操作的计算机程序代码, 上述程序设计语言包括但不限于面向对象的程序设计语言—诸如 Java、Smalltalk、C++, 还包括常规的过程式程序设计语言—诸如“C”语言或类似的设计语言。程序代码可以完全地在用户计算机上执行、部分地在用户计算机上执行、作为一个独立的软件包执行、部分在用户计算机上部分在远程计算机上执行、或者完全在远程计算机或服务器上执行。在涉及远程计算机的情形中, 远程计算机可以通过任意种类的网络——包括局域网(LAN)或广域网(WAN)—连接到用户计算机, 或者, 可以连接到外部计算机 (例如利用因特网服务提供商来通过因特网连接)。

附图中的流程图和框图, 图示了按照本公开各种实施例的系统、方法和计算机程序产品的可能实现的体系架构、功能和操作。在这点上, 流程图或框图中的每个方框可以代表一个模块、程序段、或代码的一部分, 该模块、程序段、或代码的一部分包含一个或多个用于实现规定的逻辑功能的可执行指令。也应当注意, 在有些作为替换的实现中, 方框中所标注的功能也可以以不同于附图中所标注的顺序发生。例如, 两个接连地表示的方框实际上可以基本并行地执行, 它们有时也可以按相反的顺序执行, 这依所涉及的功能而定。也要注意, 框图和/或流程图中的每个方框、以及框图和/或流程图中的方框的组合, 可以用执行规定的功能或操作的专用的基于硬件的系统来实现, 或者可以用专用硬件与计算机指令的组合来实现。

描述于本公开实施例中所涉及到的单元可以通过软件的方式实现, 也可以通过硬件的方式来实现。其中, 单元的名称在某种情况下并不构成对该单元本身的限定, 例如, 获取单元 301 还可以被描述为“获取头戴设备采集到的当前场景的深度图像的单元”。

本文中以上描述的功能可以至少部分地由一个或多个硬件逻辑部件来执行。例如, 非限制性地, 可以使用的示范类型的硬件逻辑部件包括: 现场可编程门阵列(FPGA)、专用集成电路(ASIC)、专用标准产品(ASSP)、片上系统(SOC)、复杂可编程逻辑

设备（CPLD）等等。

在本公开的上下文中，机器可读介质可以是有形的介质，其可以包含或存储以供指令执行系统、装置或设备使用或与指令执行系统、装置或设备结合地使用的程序。机器可读介质可以是机器可读信号介质或机器可读储存介质。机器可读介质可以包括
5 但不限于电子的、磁性的、光学的、电磁的、红外的、或半导体系统、装置或设备，或者上述内容的任何合适组合。机器可读存储介质的更具体示例会包括基于一个或多个线的电气连接、便携式计算机盘、硬盘、随机存取存储器(RAM)、只读存储器(ROM)、可擦除可编程只读存储器（EPROM 或快闪存储器）、光纤、便捷式紧凑盘只读存储器（CD-ROM）、光学储存设备、磁储存设备、或上述内容的任何合适组合。

10

以上描述仅为本公开的较佳实施例以及对所运用技术原理的说明。本领域技术人员应当理解，本公开中所涉及的公开范围，并不限于上述技术特征的特定组合而成的技术方案，同时也应涵盖在不脱离上述公开构思的情况下，由上述技术特征或其等同特征进行任意组合而形成的其它技术方案。例如上述特征与本公开中公开的（但不限
15 于）具有类似功能的技术特征进行互相替换而形成的技术方案。

15

此外，虽然采用特定次序描绘了各操作，但是这不应理解为要求这些操作以所示出的特定次序或以顺序次序执行来执行。在一定环境下，多任务和并行处理可能是有利的。同样地，虽然在上面论述中包含了若干具体实现细节，但是这些不应被解释为对本公开的范围的限制。在单独的实施例的上下文中描述的某些特征还可以组合
20 地实现在单个实施例中。相反地，在单个实施例的上下文中描述的各种特征也可以单独地或以任何合适的子组合的方式实现在多个实施例中。

20

尽管已经采用特定于结构特征和/或方法逻辑动作的语言描述了本主题，但是应当理解所附权利要求书中所限定的主题未必局限于上面描述的特定特征或动作。相反，上面所描述的特定特征和动作仅仅是实现权利要求书的示例形式。

权利要求

1、一种网格构建方法，应用于头戴设备，所述网格构建方法包括：

获取所述头戴设备采集到的当前场景的深度图像，所述深度图像包括目标像素点，所述目标像素点对应的目标对象的实际深度在所述当前场景中符合预设条件；以及

基于所述深度图像构建所述当前场景对应的网格，其中，在构建所述网格的过程中，所述目标像素点的深度值采用预设深度值。

2、根据权利要求 1 所述的网格构建方法，还包括：

确定所述深度图像对应的目标对象检测结果；以及

根据所述目标对象检测结果确定所述目标像素点。

3、根据权利要求 2 所述的网格构建方法，所述确定深度图像对应的目标对象检测结果，包括：

获取所述深度图像对应的环境图像，所述环境图像包括与所述深度图像的像素点一一对应的像素点；以及

基于所述环境图像，确定所述目标对象检测结果。

4、根据权利要求 3 所述的网格构建方法，其中，所述环境图像包括彩色图像或灰度图像。

5、根据权利要求 3 或 4 所述的网格构建方法，其中，所述基于所述环境图像，确定所述目标对象检测结果，包括：

基于所述环境图像和预先训练的神经网络模型，确定所述目标对象检测结果。

6、根据权利要求 1 至 5 中任一项所述的网格构建方法，其中，对于不同的目标对象，采用不同的预设深度值。

7、根据权利要求 1 所述的网格构建方法，其中，所述基于所述深度图像构建所述当前场景对应的网格，包括：

基于所述深度图像中各个像素点的深度值和各个像素点的像素坐标，确定所述各个像素点的球面坐标；以及

基于所述各个像素点的球面坐标，构建所述当前场景对应的网格。

8、根据权利要求 7 所述的网格构建方法，其中，所述基于所述各个像素点的球面坐标，构建所述当前场景对应的网格，包括：

基于所述各个像素点的球面坐标，生成所述各个像素点对应的球面；

基于所述球面进行采样，确定多个采样点的网格坐标；以及

基于所述多个采样点的网格坐标，构建所述当前场景对应的网格。

9、根据权利要求 8 所述的网格构建方法，其中，所述确定所述多个采样点的网格坐标，包括：

确定所述多个采样点分别对应的渲染逆深度，其中，所述渲染逆深度为基于邻域采样点的深度值确定的深度值；以及

基于所述多个采样点分别对应的渲染逆深度和所述多个采样点分别对应的球面坐标，确定所述多个采样点分别对应的网格坐标。

10、根据权利要求 9 所述的网格构建方法，还包括：

针对所述多个采样点中的采样点，基于所述采样点的邻域采样点的深度值确定所述采样点的平均逆深度值和所述采样点的权重影响值；以及

基于所述采样点的平均逆深度值、所述采样点的权重影响值、初始渲染逆深度和预设的能量关系，确定所述采样点的渲染逆深度。

11、根据权利要求 8 或者 9 所述的网格构建方法，其中，所述邻域采样点对应的像素点为所述目标像素点，所述邻域采样点的深度值为所述预设深度值的倒数。

12、根据权利要求 8 或 9 所述的网格构建方法，其中，所述邻域采样点对应的像素点不为所述目标像素点，所述邻域采样点的深度值为所述深度图像中所述邻域采样点的像素点对应的深度值。

13、根据权利要求 10 所述的网格构建方法，其中，所述权重影响值基于所述采样点的邻域采样点的数量确定。

14、一种网格构建装置，应用于头戴设备，所述网格构建装置包括：

获取单元，用于获取头戴设备采集到的当前场景的深度图像；所述深度图像中包括目

标像素点，所述目标像素点对应的目标对象的实际深度在所述当前场景中符合预设条件；
以及

构建单元，用于基于所述深度图像构建所述当前场景对应的网格；其中，在构建所述网格的过程中，所述目标像素点的深度值采用预设深度值。

15、一种电子设备，包括：

一个或多个处理器；以及

存储装置，用于存储一个或多个程序，

当所述一个或多个程序被所述一个或多个处理器执行，使得所述一个或多个处理器实现如权利要求 1-13 中任一所述的方法。

16、一种计算机可读介质，其上存储有计算机程序，所述计算机程序被处理器执行时实现如权利要求 1-13 中任一所述的方法。

17、一种计算机程序产品，包括承载在非暂态计算机可读介质上的计算机程序，所述计算机程序在被处理器执行时实现如权利要求 1-13 中任一所述的方法。

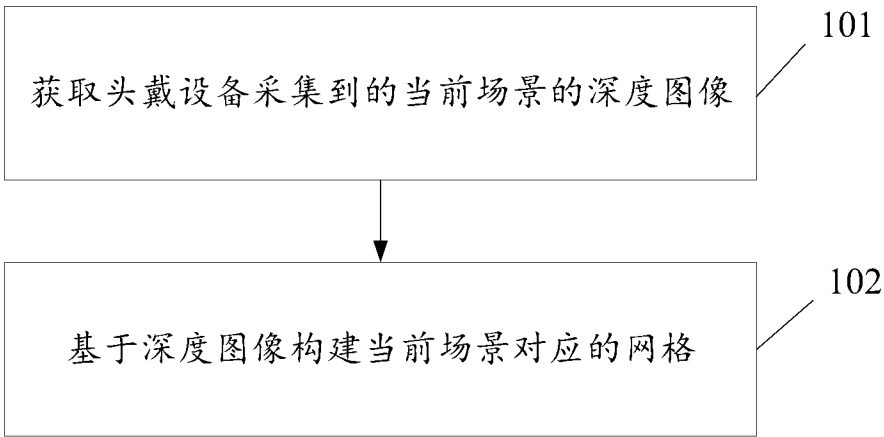


图 1

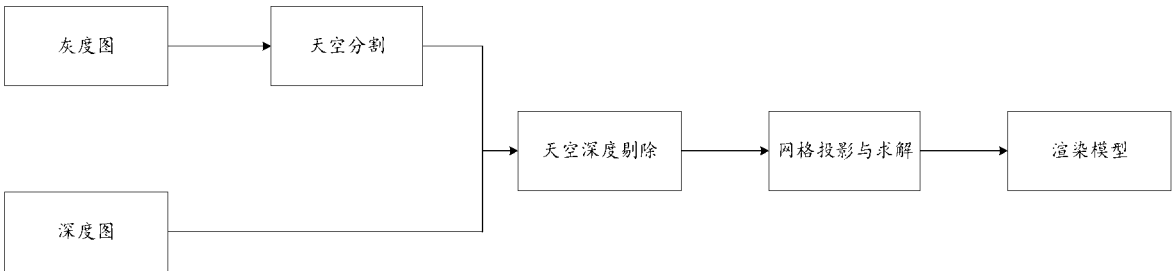


图 2



图 3

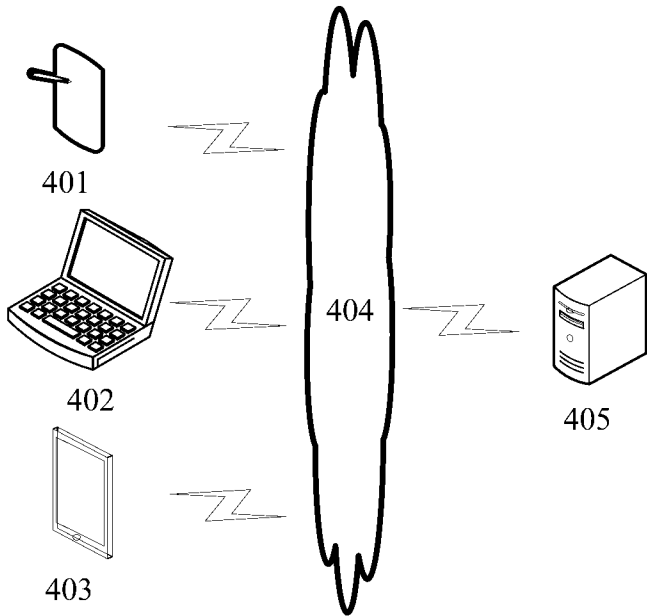


图 4

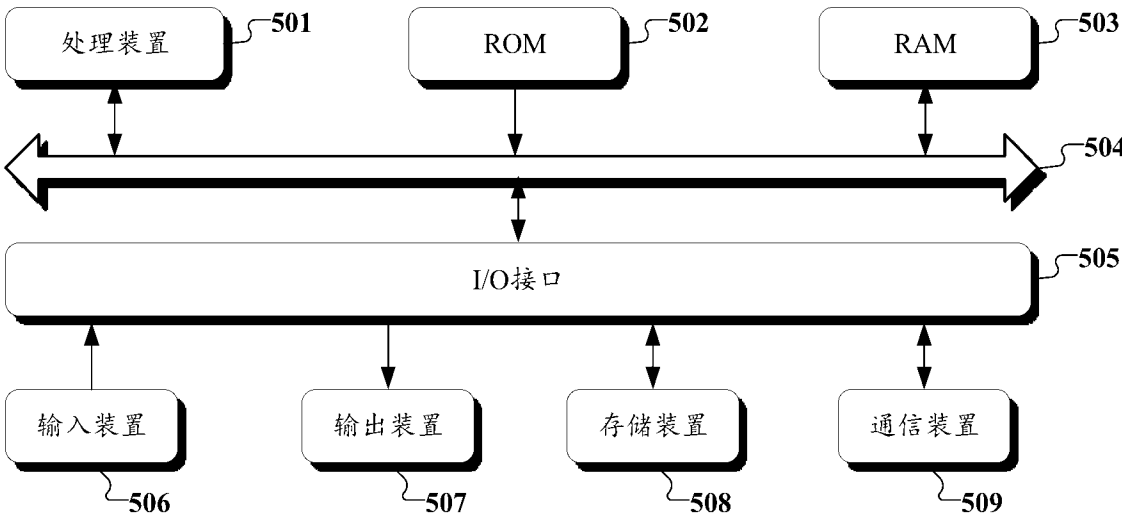


图 5

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2024/098014

A. CLASSIFICATION OF SUBJECT MATTER G06F 3/01(2006.01)i; G06T 17/05(2011.01)i According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) G06F; G06T Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) VEN, CNABS, CNTXT, WOTXT, USTXT, CNKI, IEEE: 网格, 三维场景, 构建, 头戴设备, 采集, 渲染, 深度图像, 像素点, 预设, 深度值, mesh, three-dimensional scene, build, headset, collect, render, depth image, pixel point, preset, depth value		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	CN 114860070 A (BEIJING GLOBAL CROWN TECHNOLOGY DEVELOPMENT CO., LTD.) 05 August 2022 (2022-08-05) description, paragraphs [0062]-[0108] and [0113]-[0115], and figures 1-6	1-17
Y	CN 111476087 A (SHENZHEN SENSETIME TECHNOLOGY CO., LTD.) 31 July 2020 (2020-07-31) description, paragraphs [0072]-[0088]	1-17
A	CN 113711165 A (MICROSOFT TECHNOLOGY LICENSING LLC.) 26 November 2021 (2021-11-26) entire document	1-17
A	CN 114742703 A (YINGSHI INNOVATION TECHNOLOGY CO., LTD.) 12 July 2022 (2022-07-12) entire document	1-17
A	US 2011109617 A1 (MICROSOFT CORP.) 12 May 2011 (2011-05-12) entire document	1-17
☐ Further documents are listed in the continuation of Box C. ☒ See patent family annex.		
* Special categories of cited documents: “A” document defining the general state of the art which is not considered to be of particular relevance “D” document cited by the applicant in the international application “E” earlier application or patent but published on or after the international filing date “L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) “O” document referring to an oral disclosure, use, exhibition or other means “P” document published prior to the international filing date but later than the priority date claimed “T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention “X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone “Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art “&” document member of the same patent family		
Date of the actual completion of the international search 07 August 2024		Date of mailing of the international search report 13 August 2024
Name and mailing address of the ISA/CN China National Intellectual Property Administration (ISA/CN) China No. 6, Xitucheng Road, Jimenqiao, Haidian District, Beijing 100088		Authorized officer Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2024/098014

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	114860070	A	05 August 2022	None			
CN	111476087	A	31 July 2020	None			
CN	113711165	A	26 November 2021	AU	2020260956	A1	07 October 2021
				CA	3132796	A1	29 October 2020
				WO	2020219180	A1	29 October 2020
				MX	2021012844	A	25 November 2021
				US	2021241536	A1	05 August 2021
				US	11580704	B2	14 February 2023
				US	2020334908	A1	22 October 2020
				US	11004269	B2	11 May 2021
				EP	3959584	A1	02 March 2022
CN	114742703	A	12 July 2022	None			
US	2011109617	A1	12 May 2011	None			

A. 主题的分类

G06F 3/01(2006.01)i; G06T 17/05(2011.01)i

按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类

B. 检索领域

检索的最低限度文献(标明分类系统和分类号)

G06F; G06T

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用))

VEN, CNABS, CNTXT, WOTXT, USTXT, CNKI, IEEE; 网格, 三维场景, 构建, 头戴设备, 采集, 渲染, 深度图像, 像素点, 预设, 深度值, mesh, three-dimensional scene, build, headset, collect, render, depth image, pixel point, preset, depth value

C. 相关文件

类型*	引用文件, 必要时, 指明相关段落	相关的权利要求
Y	CN 114860070 A (北京世冠金洋科技发展有限公司) 2022年8月5日 (2022 - 08 - 05) 说明书第[0062]-[0108], [0113]-[0115]段、附图1-6	1-17
Y	CN 111476087 A (深圳市商汤科技有限公司) 2020年7月31日 (2020 - 07 - 31) 说明书第[0072]-[0088]段	1-17
A	CN 113711165 A (微软技术许可有限责任公司) 2021年11月26日 (2021 - 11 - 26) 全文	1-17
A	CN 114742703 A (影石创新科技股份有限公司) 2022年7月12日 (2022 - 07 - 12) 全文	1-17
A	US 2011109617 A1 (MICROSOFT CORPORATION) 2011年5月12日 (2011 - 05 - 12) 全文	1-17

☐ 其余文件在C栏的续页中列出。

☒ 见同族专利附件。

* 引用文件的具体类型:

“A” 认为不特别相关的表示了现有技术一般状态的文件

“D” 申请人在国际申请中引证的文件

“E” 在国际申请日的当天或之后公布的在先申请或专利

“L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件(如具体说明的)

“O” 涉及口头公开、使用、展览或其他方式公开的文件

“P” 公布日先于国际申请日但迟于所要求的优先权日的文件

“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件

“X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性

“Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性

“&” 同族专利的文件

国际检索实际完成的日期

2024年8月7日

国际检索报告邮寄日期

2024年8月13日

ISA/CN的名称和邮寄地址

中国国家知识产权局
中国北京市海淀区蓟门桥西土城路6号 100088

授权官员

吴卿

电话号码 (+86) 010-53961436

PCT/ISA/210 表(第2页) (2022年7月)

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2024/098014

检索报告引用的专利文件			公布日 (年/月/日)	同族专利			公布日 (年/月/日)
CN	114860070	A	2022年8月5日	无			
CN	111476087	A	2020年7月31日	无			
CN	113711165	A	2021年11月26日	AU	2020260956	A1	2021年10月7日
				CA	3132796	A1	2020年10月29日
				WO	2020219180	A1	2020年10月29日
				MX	2021012844	A	2021年11月25日
				US	2021241536	A1	2021年8月5日
				US	11580704	B2	2023年2月14日
				US	2020334908	A1	2020年10月22日
				US	11004269	B2	2021年5月11日
				EP	3959584	A1	2022年3月2日
CN	114742703	A	2022年7月12日	无			
US	2011109617	A1	2011年5月12日	无			