(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
2024 年 3 月 28 日 (28.03.2024)

WIPO | PCT

(10) 国际公布号
WO 2024/060923 A1

(51) 国际专利分类号:
G06T 7/50 (2017.01) *G06T 7/55* (2017.01)

(21) 国际申请号: PCT/CN2023/114570

(22) 国际申请日: 2023 年 8 月 24 日 (24.08.2023)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
202211160924.9 2022年9月22日 (22.09.2022) 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: 温佳伟 (**WEN, Jiawei**); 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)。 宋小东 (**SONG, Xiaodong**); 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)。 郭亨凯 (**GUO, Hengkai**); 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: 北京品源专利代理有限公司 (**BEYOND ATTORNEYS AT LAW**); 中国北京市海淀区莲花池东路39号西金大厦6层, Beijing 100036 (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,

(54) **Title:** DEPTH ESTIMATION METHOD AND APPARATUS FOR MOVING OBJECT, AND ELECTRONIC DEVICE AND STORAGE MEDIUM

(54) 发明名称: 移动物体的深度估计方法、装置、电子设备及存储介质

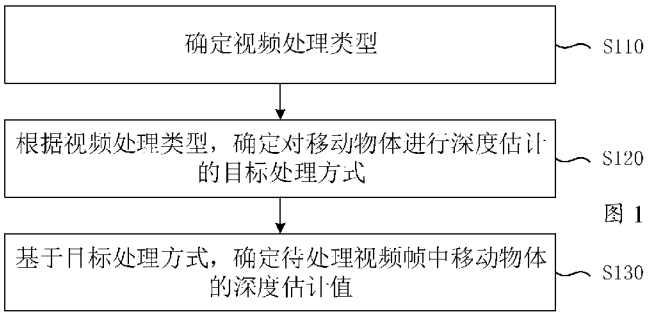


图 1

S110 Determine a video processing type
S120 According to the video processing type, determine a target processing mode for performing depth estimation on a moving object
S130 On the basis of the target processing mode, determine an estimated depth value for a moving object in a video frame to be processed

(57) **Abstract:** Provided are a depth estimation method and apparatus for a moving object, and an electronic device and a storage medium. The depth estimation method for a moving object comprises: determining a video processing type (S110); according to the video processing type, determining a target processing mode for performing depth estimation on a moving object (S120); and on the basis of the target processing mode, determining an estimated depth value for a moving object in a video frame to be processed (S130).

(57) 摘要: 移动物体的深度估计方法、装置、电子设备及存储介质。移动物体的深度估计方法包括: 确定视频处理类型(S110)；根据视频处理类型，确定对移动物体进行深度估计的目标处理方式(S120)；基于目标处理方式，确定待处理视频帧中移动物体的深度估计值(S130)。

WO 2024/060923 A1

[见续页]

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)。

本国际公布:

— 包括国际检索报告(条约第21条(3))。

移动物体的深度估计方法、装置、电子设备及存储介质

本申请要求在2022年09月22日提交中国专利局、申请号为202211160924.9的中国专利申请的优先权，该申请的全部内容通过引用结合在本申请中。

技术领域

本公开涉及图像处理技术领域，例如涉及移动物体的深度估计方法、装置、电子设备及存储介质。

背景技术

随着计算机视觉技术的发展，同步定位和建图（Simultaneous Localization and Mapping，SLAM）算法被广泛应用于增强现实、虚拟现实、自动驾驶以及机器人或无人机的定位导航等领域。

相关技术中，将图像输入至SLAM系统中，利用SLAM系统对图像中的场景深度信息进行提取，以基于场景深度信息对图像中的物体深度进行估计，然而，这种深度估计方式仅适用于静态物体，对于视频中的动态物体，很难实现有效的深度估计。

发明内容

本公开提供移动物体的深度估计方法、装置、电子设备及存储介质，以实现对视频中的移动物体深度信息进行准确估计的效果。

第一方面，本公开提供了一种移动物体的深度估计方法，该方法包括：

确定视频处理类型；

根据所述视频处理类型，确定对移动物体进行深度估计的目标处理方式；

基于所述目标处理方式，确定待处理视频帧中移动物体的深度估计值。

第二方面，本公开还提供了一种移动物体的深度估计装置，该装置包括：

视频处理类型确定模块，设置为确定视频处理类型；

目标处理方式确定模块，设置为根据所述视频处理类型，确定对移动物体进行深度估计的目标处理方式；

深度估计值确定模块，设置为基于所述目标处理方式，确定待处理视频帧中移动物体的深度估计值。

第三方面，本公开还提供了一种电子设备，所述电子设备包括:

一个或多个处理器;

存储装置，设置为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上述的移动物体的深度估计方法。

第四方面，本公开还提供了一种包含计算机可执行指令的存储介质，所述计算机可执行指令在由计算机处理器执行时用于执行上述的移动物体的深度估计方法。

第五方面，本公开还提供了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，所述计算机程序包含用于执行上述的移动物体的深度估计方法的程序代码。

附图说明

图 1 是本公开实施例所提供的一种移动物体的深度估计方法的流程示意图;

图 2 是本公开实施例所提供的另一种移动物体的深度估计方法的流程示意图;

图 3 是本公开实施例所提供的一种移动物体的深度估计装置的结构示意图;

图 4 是本公开实施例所提供的一种电子设备的结构示意图。

具体实施方式

下面将参照附图描述本公开的实施例。虽然附图中显示了本公开的一些实施例，然而本公开可以通过多种形式来实现，提供这些实施例是为了理解本公开。本公开的附图及实施例仅用于示例性作用。

本公开的方法实施方式中记载的多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。

本文使用的术语“包括”及其变形是开放性包括，即“包括”。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。

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。

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。

在使用本公开实施例公开的技术方案之前，均应当依据相关法律法规通过恰当的方式对本公开所涉及个人信息的类型、使用范围、使用场景等告知用户并获得用户的授权。

例如，在响应于接收到用户的主动请求时，向用户发送提示信息，以明确地提示用户，其请求执行的操作将需要获取和使用到用户的个人信息。从而，使得用户可以根据提示信息来自主地选择是否向执行本公开技术方案的操作的电子设备、应用程序、服务器或存储介质等软件或硬件提供个人信息。

作为一种实现方式，响应于接收到用户的主动请求，向用户发送提示信息的方式例如可以是弹窗的方式，弹窗中可以以文字的方式呈现提示信息。此外，弹窗中还可以承载供用户选择“同意”或者“不同意”向电子设备提供个人信息的选择控件。

上述通知和获取用户授权过程仅是示意性的，不对本公开的实现方式构成限定，其它满足相关法律法规的方式也可应用于本公开的实现方式中。

本技术方案所涉及的数据（包括数据本身、数据的获取或使用）应当遵循相应法律法规及相关规定的要求。

在介绍本技术方案之前，可以先对应用场景进行示例性说明。示例性的，当用户利用移动终端的摄像装置拍摄视频，并将拍摄后的视频上传至基于 SLAM 算法的系统，或者，在数据库中选择目标视频，并将视频主动上传至基于 SLAM 算法的系统后，系统即可对视频中的场景深度信息进行解析，以基于场景深度信息对视频帧中所包含的物体深度信息进行估计，然而，基于当前的深度信息估计方法仅可以对视频帧中的静态物体的深度信息进行估计，无法对视频帧中的动态物体进行准确的深度估计。此时，基于本公开实施例的方案，可以利用 SLAM 系统所提供的场景深度信息以及三维空间信息对视频帧中的移动物体的深度信息进行估计，从而实现了对视频帧中的动态物体的深度信息进行准确估计的效果。

图 1 是本公开实施例所提供的一种移动物体的深度估计方法的流程示意图，本公开实施例适用于对视频帧中的移动物体的深度信息进行估计的情形，该方法可以由移动物体的深度估计装置来执行，该装置可以通过软件和/或硬件的形

式实现，例如，通过电子设备来实现，该电子设备可以是移动终端、个人电脑（Personal Computer，PC）端或服务器等。

如图 1 所示，所述方法包括:

S110、确定视频处理类型。

在本实施例中，执行本公开实施例提供的移动物体的深度估计方法的装置，可以集成在支持特效视频处理功能的应用软件中，且该软件可以安装至电子设备中，例如，电子设备可以是移动终端或者 PC 端等。应用软件可以是对图像/视频处理的一类软件，其应用软件在此不再一一赘述，只要可以实现图像/视频处理即可。还可以是专门研发的应用程序，来实现添加特效并将特效进行展示的软件中，亦或是集成在相应的页面中，用户可以通过 PC 端中集成的页面来实现对特效视频的处理。

本实施例的技术方案可以在基于移动终端实时摄像的过程中执行，也可以在系统接收到用户主动上传的视频数据后执行，同时，本公开实施例的方案可以应用于增强现实（Augmented Reality，AR）、虚拟现实（Virtual Reality，VR）以及自动驾驶等多种应用场景内。

在本实施例中，视频处理类型可以为基于用户对待处理视频的上传方式所确定的视频处理方式。视频处理类型包括实时处理类型和后处理类型。在实际应用中，若待处理视频为用户基于移动终端摄像装置实时拍摄得到的，并且，基于移动终端对待处理视频中所包含的移动物体进行深度估计时，可以将当前待处理视频的视频处理类型确定为实时处理类型；若待处理视频为已拍摄完成的视频，并且，由用户主动上传至系统，此时，对接收到的待处理视频中所包含的移动物体进行深度估计，则待处理视频的视频处理类型可以为后处理类型。

在本实施例中，若系统接收到的视频数据为基于移动终端摄像装置实时拍摄得到的，则可以将视频处理类型确定为实时处理类型；若系统接收到视频数据时视频数据是已经拍摄完成的完整视频数据，则可以将视频处理类型确定为后处理类型。这样设置的好处在于：可以增强移动物体深度估计处理方式的多样性，以使待处理视频帧中的移动物体深度估计既可以基于移动终端实时进行，还可以对完整视频中的移动物体进行深度估计，提高了视频处理的多样性，满足了用户的个性化需求。

S120、根据视频处理类型，确定对移动物体进行深度估计的目标处理方式。

在本实施例中，当检测到用户触发特效操作时，移动终端的摄像装置可以是实时面向用户以实现待处理视频的采集，并根据预先编写的程序对待处理视频进行解析，即可得到多个待处理视频帧，此时，可以将视频处理类型确定为实时

处理类型。相应的，待处理视频帧中可以包括移动物体。移动物体可以为入镜画面中任何姿态或位置信息发生变化的对象，例如可以是用户或者动物等。

深度估计可以是计算机视觉领域中的一个子任务，其目的是获取物体与拍摄点之间的距离，可以为三维重建、距离感知、SLAM、视觉里程估计、视频插帧、图像重建等一系列任务提供深度信息。移动物体的深度信息可以是最终所呈现的画面中与移动物体相对应的像素点与拍摄点之间的距离，也可以通过像素点在相机坐标系下的位置坐标进行表示。

在本实施例中，当确定视频处理类型为实时处理类型时，则可以确定对视频帧中移动物体进行深度估计的目标处理方式可以为与实时处理类型相对应的深度均值估计方式。其中，深度均值估计方式可以为确定与移动物体相关联的一些像素点的深度值，并对这些深度值取均值，从而可以将最终得到的平均深度值作为移动物体的深度信息。

S130、基于目标处理方式，确定待处理视频帧中移动物体的深度估计值。

在本实施例中，用户可以基于移动终端的摄像装置实时对移动物体进行视频拍摄，并实时上传至移动终端，因此，系统所获取的实时拍摄的视频即为待处理视频，基于预先编写的程序对待处理视频进行解析，即可得到多个待处理视频帧。深度估计值可以为与移动物体相对应的至少一个像素点与拍摄点之间的距离，也可以为与移动物体相对应的至少一个像素点在相机坐标系下的坐标值。

在本实施例中，目标处理方式可以为深度均值估计方式，在基于目标处理方式确定移动物体的深度估计值时，可以首先确定移动物体中满足深度均值估计条件的目标像素点，进而可以基于这些目标像素点的深度值确定深度均值，从而可以将最终得到的深度均值作为移动物体的深度估计值。

基于目标处理方式，确定待处理视频帧中移动物体的深度估计值，可以包括：确定与待处理视频帧相对应的拍摄参数以及移动物体的像素点参数；基于拍摄参数、像素点参数以及约束条件，确定目标像素点；基于目标像素点的点云数据，确定移动物体的深度估计值。

在本实施例中，拍摄参数可以是待处理视频帧在进行位姿优化之后的相机位姿参数。可以基于与待处理视频帧相对应的摄像装置中的陀螺仪以及惯性测量单元，获取相机位置信息和旋转信息，以基于相机位置信息和旋转信息，确定待处理视频帧的初始位姿，基于光束平差法（Bundle Adjustment，BA）对初始位姿进行优化，并将优化后的位姿作为与待处理视频帧相对应的拍摄参数。这样设置的好处在于：可以为同步定位与建图系统提供较高的 BA 速度，从而保证了系统对视频帧进行处理的实时性。像素点参数可以为待处理视频帧中用于构成移

动物体的至少一个像素点的像素坐标。在对移动物体进行拍摄，以得到多个待处理视频帧时，待处理视频帧中不仅包含移动物体，还可以包含移动物体所在的场景，因此，在确定移动物体的像素点参数时，可以首先确定移动物体的掩膜图像，从而可以基于该掩膜图像确定构成移动物体的至少一个像素点的像素坐标。

在本实施例中，约束条件可以为空间几何信息约束条件，即在特定位置观测一像素点时，确定该像素点的状态是否与该特定位置相对应，若该像素点的状态与其观测位置相对应，则可以确定该像素点满足约束条件；若该像素点的状态与其观测位置不对应，则可以确定该像素点不满足约束条件。

在本实施例中，在得到待处理视频帧后，可以基于与待处理视频帧相对应的摄像装置的传感器的参数，确定待处理视频帧的初始位姿，再基于位姿优化方法对初始位姿进行优化，将优化后的位姿作为与待处理视频帧相对应的拍摄参数，同时，确定待处理视频帧中移动物体的像素点坐标，以作为像素点参数，基于拍摄参数、像素点参数以及约束条件，确定目标像素点，从而可以基于目标像素点的点云数据，确定移动物体的深度估计值。这样设置的好处在于：可以基于约束条件将移动物体的多个像素点划分为动态像素点和静态像素点，并将动态像素点筛选出来作为移动物体的跟踪像素点，提高了移动物体的深度估计值的准确率，提升了待处理视频帧中移动物体的定位效果。

在实际应用中，可以首先确定待处理视频帧的初始位姿，并基于位姿优化方法对初始位姿进行优化，得到与待处理视频帧相对应的拍摄参数，同时，确定与移动物体相对应的至少一个像素点的像素坐标，以得到像素点参数，基于拍摄参数、像素点参数以及约束条件，确定与移动物体相对应的多个像素点中满足约束条件的像素点，并将这些像素点作为目标像素点。

基于拍摄参数、像素点参数以及约束条件，确定目标像素点，包括：对拍摄参数和像素点参数进行三角化处理，以得到像素点参数所对应的点云数据；基于点云数据和约束条件，确定反投影像素参数；基于像素点参数和反投影像素参数，确定所述目标像素点。

在本实施例中，三角化处理可以为基于角点检测算法确定相应的点云数据。其中，角点检测算法可以为 KLT 角点检测法，也称为 KLT 光流追踪法。KLT 角点检测法通过在多个关键帧中确定一个适合跟踪的参考关键帧，并确定该参考关键帧的特征点，从而基于该特征点确定相应的点云数据（Point Cloud Data，PCD）。点云数据通常用于逆向工程中，是一种以点的形式记录的数据，这些点既可以表示三维空间中的坐标，也可以表示颜色或者光照强度等信息，在实际应用过程中，点云数据一般还包括点坐标精度、空间分辨率和表面法向量等内容，一般以 PCD 格式进行保存，在这种格式下，点云数据的可操作性较强，能够在

后续过程中提高点云配准和融合的速度，本公开实施例对此不再赘述。

在实际应用中，在确定拍摄参数和像素点参数后，即可基于角点检测算法对拍摄参数和像素点参数进行三角化处理，从而可以得到与像素点参数相对应的三维点云数据。根据点云数据和约束条件，确定这些点云数据在相机坐标系下的参数，即，将三维点云数据转换为二维坐标的形式，可以将转换之后的二维坐标参数作为反投影像素参数，由于点云数据是基于像素点参数确定的，反投影像素参数是基于点云数据确定的，并且，像素点参数和反投影像素参数均为二维坐标参数，因此，可以通过确定像素点参数与相应的反投影像素参数是否一致，来确定目标像素点，即，将像素点参数与相应的反投影像素参数不一致的像素点作为目标像素点。移动物体的像素点是基于掩膜图像确定的，在实际应用过程中，通常采用部署在移动终端的模型对待处理视频帧进行处理，以得到与移动物体相对应的掩膜图像，一般情况下，为了提高移动终端的处理效率以及减少模型在移动终端的内存占用率，部署在移动终端的模型通常为模型结构简单、处理速度快的模型，在应用该模型对待处理视频帧进行移动物体掩膜图像处理时，可能会使得到的掩膜图像大于移动物体的实际尺寸，从而会将不属于移动物体的静态背景点也划分进来。对于静态像素点来说，一般是可以满足约束条件的，对于动态像素点来说，是不满足约束条件的，因此，可以通过确定与移动物体相对应的像素点是否满足约束条件，将动态像素点和静态像素点区分出来，以便可以针对不同的像素点采取不同的处理方式，从而可以最终得到移动物体的深度估计值。这样设置的好处在于：可以更加精准地确定移动物体的像素点，进而可以针对不同像素点采取不同的处理方式，提高了移动物体深度估计值的准确率。

示例性的，基于点云数据和约束条件，确定反投影像素参数可以基于如下公式进行确定：

$$s_i \begin{bmatrix} u_i \\ v_i \\ 1 \end{bmatrix} = K exp\left(\xi^{\wedge}\right) \begin{bmatrix} X_i \\ Y_i \\ Z_i \\ 1 \end{bmatrix}$$

其中，s_i可以表示任一像素点的深度值，（u_i，v_i）可以表示任一像素点的像素坐标，K可以表示相机内参，$exp\left(\xi^{\wedge}\right)$可以表示相机姿态，即 R，T 矩阵，（$X_i$，$Y_i$，$Z_i$）可以表示任一像素点的三维点云坐标。

在确定目标像素点后，即可根据目标像素点的点云数据，确定移动物体的深度估计值。

基于目标像素点的点云数据，确定移动物体的深度估计值，包括：根据目标像素点的点云数据，确定目标像素点所属的至少两个待使用视频帧；根据目标像

素点在至少两个待使用视频帧的深度值，确定移动物体的深度估计值。

在本实施例中，在得到目标像素点后，可以对这些目标像素点进行三角化处理，得到与目标像素点相对应的点云数据，在包含移动物体的多个待处理视频帧中对点云数据进行观测，并将可以观测到点云数据的至少两个待处理视频帧作为待使用视频帧。

在实际应用中，在确定目标像素点所属的至少两个待使用视频帧后，即可确定目标像素点在相机坐标系下的深度值，并对这些深度值作均值处理，可以将最终得到的深度值均值作为移动物体的深度估计值。这样设置的好处在于：可以在移动终端实现移动物体深度信息的粗略估计，提高了移动物体的深度估计的效率。

若移动物体是处于静止状态时，则基于掩膜图像所确定的移动物体的多个像素点均满足约束条件，即，多个像素点的像素点参数和反投影像素参数相一致，此时，可以对这些像素点进行三角化处理，得到与这些像素点相对应的点云数据，并将这些点云数据存放至 SLAM 系统中，以便可以通过 SLAM 系统确定移动物体的深度估计值。

本实施例是对视频处理类型为实时处理类型的待处理视频帧中的移动物体的深度估计值进行确定，在本实施例的基础上，当视频处理类型为后处理类型时，其对应的目标处理方式也会相应发生变化，下面可以对后处理类型进行详细阐述。

本公开实施例的技术方案，通过确定视频处理类型，根据视频处理类型，确定对移动物体进行深度估计的目标处理方式，最后，基于目标处理方式，确定待处理视频帧中移动物体的深度估计值，解决了相关技术中仅可以对静态物体的深度信息进行估计的问题，实现了对视频帧中的移动物体的深度信息进行准确估计的效果，并且，提高了深度估计的适用范围，满足了用户的个性化需求，提升了用户体验。

图 2 是本公开实施例所提供的另一种移动物体的深度估计方法的流程示意图，在前述实施例的基础上，当视频处理类型为后处理类型时，其对应目标处理方式可以为逆深度估计方式，进而可以基于逆深度估计方式，确定移动物体的深度估计值。其实施方式可以参见本实施例技术方案。其中，与上述实施例相同或者相应的技术术语在此不再赘述。

如图 2 所示，该方法包括如下步骤：

S210、确定视频处理类型为后处理类型。

上述实施例是对视频处理类型为实时处理类型的待处理视频帧中的移动物体的深度估计值进行确定，在上述实施例的基础上，当视频处理类型为后处理类型时，其对应的目标处理方式也会相应发生变化，下面可以对后处理类型进行阐述。

在本实施例中，可以预先开发视频上传控件，当检测到用户对应用内的视频上传控件的触发操作时，可以对用户主动上传的视频进行接收，并将该视频作为待处理视频，基于预先编写的程序对待处理视频进行解析，即可得到多个待处理视频帧。相应的，待处理视频帧中包含移动物体，移动物体可以为用户、动物或者入镜画面中任何姿态或位置信息发生变化的物体。当接收到完整的待处理视频时，则可以将包含移动物体的视频帧作为待处理视频帧，并对这些视频帧进行特效处理，以得到相应的特效视频帧，可以将此种视频处理方式作为后处理类型。

S220、根据后处理类型，确定对移动物体进行深度估计的目标处理方式为逆深度估计方式。

在本实施例中，在接收到待处理视频并确定视频处理类型为后处理类型后，即可将对待处理视频帧中移动物体进行深度估计的目标处理方式确定为逆深度估计方式。其中，逆深度估计方式可以为基于与移动物体相对应的至少一个像素点的逆深度值，对移动物体的深度估计值进行确定。

当视频处理方式为后处理类型时，即，是对完整的视频数据进行移动物体的深度估计，此时，与实时处理类型不同的是，在接收到完成的视频数据后，可以确定视频数据中每个待处理视频帧中每个像素点的深度信息，进而基于这些深度信息对移动物体的深度信息进行估计，然而，每一待处理视频帧中不同像素点的深度信息的分布范围较大，并且深度分布形式不稳定，因此，可以确定与这些深度信息相对应的逆深度信息，以基于这些逆深度信息对移动物体的深度估计值进行确定。这样设置的好处在于：逆深度分布形式更加符合高斯分布的分布形式，会更加稳定，从而在确定移动物体的深度估计值也会更加准确。

在每一待处理视频帧中包括远景像素点和近景像素点，对于远景像素点来说，由于这些远景像素点与拍摄点之间的距离较远，导致这些像素点的视差较小，在确定与这些远景像素点相对应的点云数据时，点云数据的精度也会较低，因此，可以采用逆深度的方式以减弱远景像素点对计算过程的影响，将远景像素点和近景像素点的深度值转换为逆深度值，进而可以基于这些逆深度值进行后续的计算，从而可以达到提高计算精度的效果。

S230、基于逆深度估计方式，确定待处理视频帧中移动物体的深度估计值。

在本实施例中，在确定目标处理方式为逆深度估计方式后，即可确定待处理

视频帧中每个像素点的逆深度值，从而可以基于这些逆深度值对移动物体的深度估计值进行确定。

基于逆深度估计方式，确定待处理视频帧中移动物体的深度估计值，包括：对目标视频中每个待处理视频帧进行三角化处理，得到每个待处理视频帧中每个像素点的逆深度值；通过对同一待处理视频帧中的多个逆深度值进行聚类处理，确定移动物体的深度估计值。

在本实施例中，目标视频可以为用户主动上传的，需要对视频中的移动物体的深度信息进行确定的视频。在实际应用过程中，在接收到目标视频中多个待处理视频帧时，可以基于角点检测算法对每个待处理视频帧进行三角化处理，以得到与每个待处理视频帧相对应的点云数据，可以根据平移旋转矩阵将与每个待处理视频帧相对应的点云数据转换至相机坐标系下，以得到每个像素点在相机坐标下的深度值，然后，对这些深度值进行求逆处理，即，分别确定这些深度值的负一次方，即可得到每个像素点的逆深度值，从而可以通过对同一待处理视频帧中的多个逆深度值进行聚类处理，确定移动物体的深度估计值。这样设置的好处在于：基于每个像素点的逆深度值进行移动物体的深度估计，可以减弱远景像素点对深度估计的影响，从而可以提高深度估计值的准确率，提升目标视频中移动物体在不同时间戳下定格点的显示效果。

聚类处理可以为对多个逆深度值进行分类处理，可以为二分类，即，将多个逆深度值划分为两大类。

通过对同一待处理视频帧中的多个逆深度值进行聚类处理，确定移动物体的深度估计值，包括：依据多个逆深度值的大小进行排序后，确定相邻两个逆深度值之间的深度差值；获取深度差值最大的两个目标逆深度值，并基于大于目标逆深度值的多个逆深度值确定移动物体的深度估计值。

在实际应用中，对于同一待处理视频帧中的多个逆深度值来说，可以首先确定每个逆深度值的大小，并依据多个逆深度值的大小进行排序，然后，确定相邻两个逆深度值之间差值作为深度差值，并确定与最大深度差值相对应的相邻两个逆深度值，将这两个逆深度值作为目标逆深度值，可以基于这两个目标逆深度值将多个逆深度值划分为两大类，一类为大于目标逆深度值的多个逆深度值，另一类为小于目标逆深度值的多个逆深度值，最后，可以基于大于目标逆深度值的多个逆深度值来对移动物体的深度估计值进行确定。这样设置的好处在于：可以基于多个逆深度值，实现近景像素点和远景像素点的分类，从而可以基于近景像素点的深度信息确定移动物体的深度信息。

本实施例中，由于多个逆深度值是按照从大到小的顺序排序的，故两个目标逆深度值为多个逆深度值中数值相邻的两个值。因此在基于大于目标逆深度值

的多个逆深度值确定移动物体的深度估计值时，所使用的目标逆深度值可以为两个目标逆深度值中的任意一个，均可以达到对多个逆深度值进行分类的效果。

在基于目标逆深度值对多个逆深度值进行分类时，若任一类中的逆深度值的数量小于预设阈值时，则可以认为这一类中的逆深度值可能是存在一定误差的，为了提高移动物体深度估计值的准确率，可以将这些逆深度值删除，并对剩余的多个逆深度值再次执行排序分类的操作，以在再次分类结束后，基于再次分类结果中大于目标逆深度值的多个逆深度值，确定移动物体的深度估计值。

基于此，在基于大于目标逆深度值的多个逆深度值确定移动物体的深度估计值之前，还包括：若大于或小于目标逆深度值的逆深度值的数量与逆深度值总数量之间的比值小于预设比例，则删除该大于或小于目标逆深度值的逆深度值，并重新执行确定所述目标逆深度值的操作。

在本实施例中，预设比例可以为任意值，预设比例可以为 5%。

在实际应用中，在基于目标逆深度值将多个逆深度值划分为大于目标逆深度值的多个逆深度值，以及小于目标逆深度值的多个逆深度值之后，可以确定这两类中逆深度值的数量与当前待处理视频帧中逆深度值的总数量之间的比值，若其中任意一类对应的比值小于预设比例，则可以将这一类中的逆深度值删除，并基于剩余的多个逆深度值的大小进行重新排序，然后，确定相邻两个逆深度之间的差值，并将差值最大的两个逆深度值作为目标逆深度值，基于目标逆深度值对剩余的多个逆深度值进行分类，从而可以最终基于大于目标逆深度值的多个逆深度值，确定移动物体的深度估计值。这样设置的好处在于：可以将存在较大误差的逆深度值筛选出来并删除，以达到提高移动物体深度估计值的准确率的效果。

基于大于目标逆深度值的多个逆深度值确定移动物体的深度估计值，包括：对大于目标逆深度值的多个逆深度值进行均值处理得到逆深度均值，根据逆深度均值确定移动物体的深度估计值。

在得到大于目标逆深度值的多个逆深度值之后，由于这些逆深度值所对应的像素点为待处理视频帧的近景像素点，在基于近景像素点进行计算时，可以得到较为准确的计算结果，并且，对于移动物体来说，一般会处于待处理视频帧中的前景部分，因此，在确定移动物体的深度估计值时，基于大于目标逆深度值的多个逆深度值进行计算，可以得到更加精确的深度估计结果。

在实际应用中，可以对大于目标逆深度值的多个逆深度值作均值处理，并对得到的逆深度均值进行再次求逆处理，得到与逆深度均值相对应的深度均值，可以将该深度均值作为移动物体的深度估计值。这样设置的好处在于：基于近景像

素点的深度信息确定移动物体的深度信息，可以达到提高深度估计准确率的效果。

对于目标视频中的任一待处理视频帧，均可以采用上述技术方法确定视频帧中移动物体的深度估计值，进而，在得到每一待处理视频帧中移动物体的深度估计值之后，可以将多个待处理视频帧拼接起来，即可得到完整目标视频中移动物体的深度估计值。

本公开实施例的技术方案，通过确定视频处理类型为后处理类型，根据后处理类型，确定对移动物体进行深度估计的目标处理方式为逆深度估计方式，最后，基于逆深度估计方式，确定待处理视频帧中移动物体的深度估计值，解决了相关技术中仅可以对静态物体的深度信息进行估计的问题，实现了对视频帧中的移动物体的深度信息进行准确估计的效果，并且，提高了深度估计的适用范围，满足了用户的个性化需求，提升了用户体验。

图 3 为本公开实施例所提供的一种移动物体的深度估计装置的结构示意图，如图 3 所示，所述装置包括：视频处理类型确定模块 310、目标处理方式确定模块 320 以及深度估计值确定模块 330。

视频处理类型确定模块 310，设置为确定视频处理类型；目标处理方式确定模块 320，设置为根据所述视频处理类型，确定对移动物体进行深度估计的目标处理方式；深度估计值确定模块 330，设置为基于所述目标处理方式，确定待处理视频帧中移动物体的深度估计值。

在上述技术方案的基础上，所述视频处理类型包括实时处理类型和后处理类型。

在上述技术方案的基础上，所述目标处理方式包括与所述实时处理类型相对应的深度均值估计方式，或与所述后处理类型所对应的逆深度估计方式。

在上述技术方案的基础上，所述目标处理方式包括深度均值估计方式，深度估计值确定模块 330 包括：拍摄参数确定子模块、目标像素点确定子模块以及深度估计值确定子模块。

拍摄参数确定子模块，设置为确定与所述待处理视频帧相对应的拍摄参数以及所述移动物体的像素点参数；目标像素点确定子模块，设置为基于所述拍摄参数、像素点参数以及约束条件，确定目标像素点；深度估计值确定子模块，设置为基于所述目标像素点的点云数据，确定所述移动物体的深度估计值。

在上述技术方案的基础上，目标像素点确定子模块包括：点云数据确定单元、反投影像素参数确定单元以及目标像素点确定单元。

点云数据确定单元，设置为对所述拍摄参数和所述像素点参数进行三角化处理，以得到所述像素点参数所对应的点云数据；反投影像素参数确定单元，设置为基于所述点云数据和所述约束条件，确定反投影像素参数；目标像素点确定单元，设置为基于所述像素点参数和所述反投影像素参数，确定所述目标像素点。

在上述技术方案的基础上，深度估计值确定子模块包括：待使用视频帧确定单元和深度估计值确定单元。

待使用视频帧确定单元，设置为根据所述目标像素点的点云数据，确定所述目标像素点所属的至少两个待使用视频帧；深度估计值确定单元，设置为根据所述目标像素点在所述至少两个待使用视频帧的深度值，确定所述移动物体的深度估计值。

在上述技术方案的基础上，所述目标处理方式包括逆深度估计方式，深度估计值确定模块 330 还包括：逆深度值确定子模块和深度估计值确定子模块。

逆深度值确定子模块，设置为对目标视频中每个待处理视频帧进行三角化处理，得到每个待处理视频帧中每个像素点的逆深度值；深度估计值确定子模块，设置为通过对同一待处理视频帧中的多个逆深度值进行聚类处理，确定所述移动物体的深度估计值。

在上述技术方案的基础上，深度估计值确定子模块包括：深度差值确定单元和深度估计值确定单元。

深度差值确定单元，设置为依据多个逆深度值的大小进行排序后，确定相邻两个逆深度值之间的深度差值；深度估计值确定单元，设置为获取深度差值最大的两个目标逆深度值，并基于大于所述目标逆深度值的多个逆深度值确定所述移动物体的深度估计值。

在上述技术方案的基础上，所述装置还包括：逆深度值删除模块。

逆深度值删除模块，设置为在所述基于大于所述目标逆深度值的多个逆深度值确定所述移动物体的深度估计值之前，若大于或小于所述目标逆深度值的逆深度值的数量与逆深度值总数量之间的比值小于预设比例，则删除所述大于或小于所述目标逆深度值的逆深度值，并重新执行确定所述目标逆深度值的操作。

在上述技术方案的基础上，深度估计值确定单元，设置为对大于所述目标逆深度值的多个逆深度值进行均值处理得到逆深度均值，根据所述逆深度均值确定所述移动物体的深度估计值。

本公开实施例的技术方案，通过确定视频处理类型，根据视频处理类型，确定对移动物体进行深度估计的目标处理方式，最后，基于目标处理方式，确定待

处理视频帧中移动物体的深度估计值，解决了相关技术中仅可以对静态物体的深度信息进行估计的问题，实现了对视频帧中的移动物体的深度信息进行准确估计的效果，并且，提高了深度估计的适用范围，满足了用户的个性化需求，提升了用户体验。

本公开实施例所提供的移动物体的深度估计装置可执行本公开任意实施例所提供的移动物体的深度估计方法，具备执行方法相应的功能模块和效果。

上述装置所包括的多个单元和模块只是按照功能逻辑进行划分的，但并不局限于上述的划分，只要能够实现相应的功能即可；另外，多个功能单元的名称也只是为了便于相互区分，并不用于限制本公开实施例的保护范围。

图 4 为本公开实施例所提供的一种电子设备的结构示意图。下面参考图 4，其示出了适于用来实现本公开实施例的电子设备（例如图 4 中的终端设备或服务器）500 的结构示意图。本公开实施例中的终端设备可以包括诸如移动电话、笔记本电脑、数字广播接收器、个人数字助理（Personal Digital Assistant，PDA）、平板电脑（Portable Android Device，PAD）、便携式多媒体播放器（Portable Media Player，PMP）、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电视（Television，TV）、台式计算机等等的固定终端。图 4 示出的电子设备 500 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 4 所示，电子设备 500 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（Read-Only Memory，ROM）502 中的程序或者从存储装置 508 加载到随机访问存储器（Random Access Memory，RAM）503 中的程序而执行多种适当的动作和处理。在 RAM 503 中，还存储有电子设备 500 操作所需的多种程序和数据。处理装置 501、ROM 502 以及 RAM 503 通过总线 504 彼此相连。输入/输出（Input/Output，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（Liquid Crystal Display，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 500 与其他设备进行无线或有线通信以交换数据。虽然图 4 示出了具有多种装置的电子设备 500，并不要求实施或具备所有示出的装置。可以替代地实施或具备更多或更少的装置。

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装，或者从存储装置 508 被安装，或者从 ROM 502 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。

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。

本公开实施例提供的电子设备与上述实施例提供的移动物体的深度估计方法属于同一构思，未在本实施例中详尽描述的技术细节可参见上述实施例，并且本实施例与上述实施例具有相同的效果。

本公开实施例提供了一种计算机存储介质，其上存储有计算机程序，该程序被处理器执行时实现上述实施例所提供的移动物体的深度估计方法。

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，或者任意以上的组合。计算机可读存储介质的例子可以包括：具有一个或多个导线的电连接、便携式计算机磁盘、硬盘、RAM、ROM、可擦式可编程只读存储器（Erasable Programmable Read-Only Memory，EPROM 或闪存）、光纤、便携式紧凑磁盘只读存储器（Compact Disc Read-Only Memory，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、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：电线、光缆、射频（Radio Frequency，RF）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超文本传输协议（HyperText Transfer Protocol，HTTP）之类的任何当前已知或未来研发的网络协议进行通信，并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（Local Area Network，LAN），广域网（Wide Area Network，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：确定视频处理类型；根据所述视频处理类型，确定对移动物体进行深度估计的目标处理方式；基于所述目标处理方式，确定待处理视频帧中移动物体的深度估计值。

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”语言或类似的程序设计语言。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 LAN 或 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，图示了按照本公开多种实施例的系统、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。也要注意的是，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，单元的名称在一种情况下并不构成对该单元本身的限定，例如，第一获取单元还可以被描述为“获取至少两个网际协议地址的单元”。

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（Field Programmable Gate Array，FPGA）、专用集成电路（Application Specific

Integrated Circuit，ASIC）、专用标准产品（Application Specific Standard Parts，ASSP）、片上系统（System on Chip，SOC）、复杂可编程逻辑设备（Complex Programming Logic Device，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、RAM、ROM、EPROM 或快闪存储器、光纤、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。

此外，虽然采用特定次序描绘了多个操作，但是这不应当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多个实现细节，但是这些不应当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一些特征还可以组合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多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。

1、一种移动物体的深度估计方法，包括:

确定视频处理类型;

根据所述视频处理类型，确定对移动物体进行深度估计的目标处理方式;

基于所述目标处理方式，确定待处理视频帧中移动物体的深度估计值。

2、根据权利要求 1 所述的方法，其中，所述视频处理类型包括实时处理类型和后处理类型。

3、根据权利要求 2 所述的方法，其中，所述目标处理方式包括与所述实时处理类型相对应的深度均值估计方式，或与所述后处理类型所对应的逆深度估计方式。

4、根据权利要求 1 所述的方法，其中，所述目标处理方式包括深度均值估计方式，所述基于所述目标处理方式，确定待处理视频帧中移动物体的深度估计值，包括:

确定与所述待处理视频帧相对应的拍摄参数以及所述移动物体的像素点参数;

基于所述拍摄参数、所述像素点参数以及约束条件，确定目标像素点;

基于所述目标像素点的点云数据，确定所述移动物体的深度估计值。

5、根据权利要求 4 所述的方法，其中，所述基于所述拍摄参数、所述像素点参数以及约束条件，确定目标像素点，包括:

对所述拍摄参数和所述像素点参数进行三角化处理，以得到所述像素点参数所对应的点云数据;

基于所述点云数据和所述约束条件，确定反投影像素参数;

基于所述像素点参数和所述反投影像素参数，确定所述目标像素点。

6、根据权利要求 4 所述的方法，其中，所述基于所述目标像素点的点云数据，确定移动物体的深度估计值，包括:

根据所述目标像素点的点云数据，确定所述目标像素点所属的至少两个待使用视频帧;

根据所述目标像素点在所述至少两个待使用视频帧的深度值，确定所述移动物体的深度估计值。

7、根据权利要求 1 所述的方法，其中，所述目标处理方式包括逆深度估计方式，所述基于所述目标处理方式，确定待处理视频帧中移动物体的深度估计值，

包括:

对目标视频中每个待处理视频帧进行三角化处理，得到每个待处理视频帧中每个像素点的逆深度值;

通过对同一待处理视频帧中的多个逆深度值进行聚类处理，确定所述移动物体的深度估计值。

8、根据权利要求 7 所述的方法，其中，所述通过对同一待处理视频帧中的多个逆深度值进行聚类处理，确定所述移动物体的深度估计值，包括:

依据所述多个逆深度值的大小进行排序后，确定相邻两个逆深度值之间的深度差值;

获取深度差值最大的两个目标逆深度值，并基于大于所述目标逆深度值的多个逆深度值确定所述移动物体的深度估计值。

9、根据权利要求 8 所述的方法，在所述基于大于所述目标逆深度值的多个逆深度值确定所述移动物体的深度估计值之前，还包括:

在大于或小于所述目标逆深度值的逆深度值的数量与逆深度值总数量之间的比值小于预设比例的情况下，删除所述大于或小于所述目标逆深度值的逆深度值，并重新执行确定所述目标逆深度值的操作。

10、根据权利要求 8 所述的方法，其中，所述基于大于所述目标逆深度值的多个逆深度值确定所述移动物体的深度估计值，包括:

对大于所述目标逆深度值多个逆深度值进行均值处理得到逆深度均值，根据所述逆深度均值确定所述移动物体的深度估计值。

11、一种移动物体的深度估计装置，包括:

视频处理类型确定模块，设置为确定视频处理类型;

目标处理方式确定模块，设置为根据所述视频处理类型，确定对移动物体进行深度估计的目标处理方式;

深度估计值确定模块，设置为基于所述目标处理方式，确定待处理视频帧中移动物体的深度估计值。

12、一种电子设备，包括:

至少一个处理器;

存储装置，设置为存储至少一个程序;

当所述至少一个程序被所述至少一个处理器执行，使得所述至少一个处理

器实现如权利要求 1-10 中任一所述的移动物体的深度估计方法。

13、一种包含计算机可执行指令的存储介质，所述计算机可执行指令在由计算机处理器执行时用于执行如权利要求 1-10 中任一所述的移动物体的深度估计方法。

14、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，所述计算机程序包含用于执行如权利要求 1-10 中任一所述的移动物体的深度估计方法的程序代码。

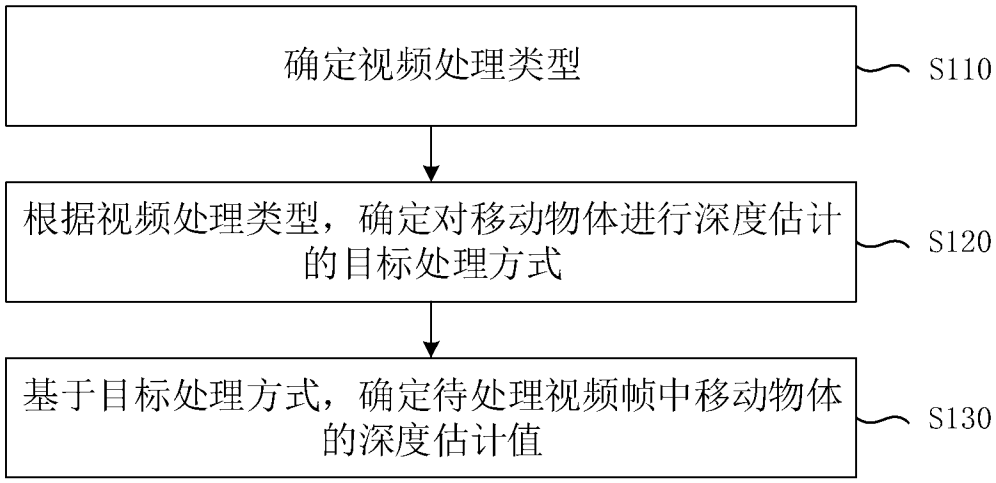

确定视频处理类型
S110
根据视频处理类型，确定对移动物体进行深度估计的目标处理方式
S120
基于目标处理方式，确定待处理视频帧中移动物体的深度估计值
S130

图 1

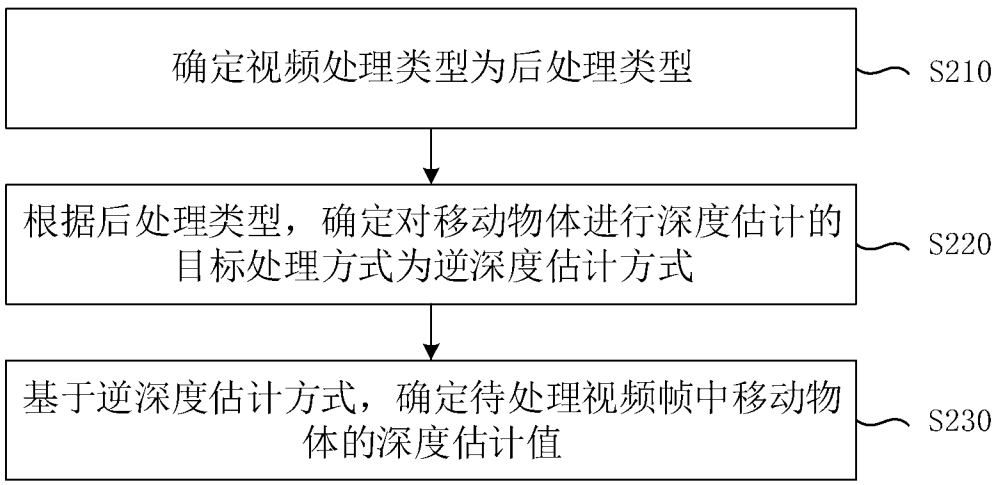

确定视频处理类型为后处理类型
S210
根据后处理类型，确定对移动物体进行深度估计的目标处理方式为逆深度估计方式
S220
基于逆深度估计方式，确定待处理视频帧中移动物体的深度估计值
S230

图 2

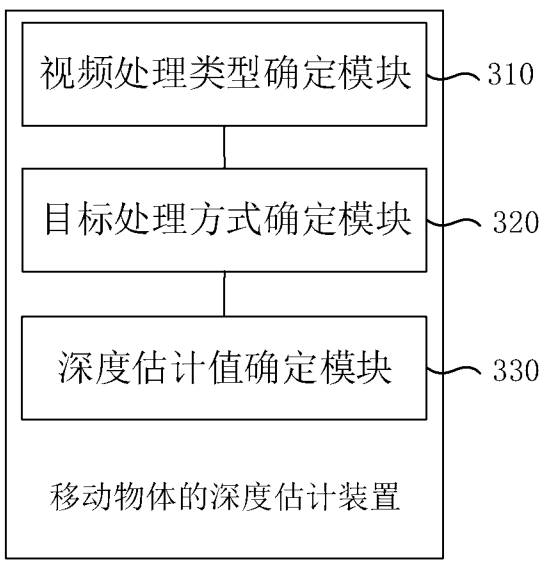

视频处理类型确定模块
310
目标处理方式确定模块
320
深度估计值确定模块
330
移动物体的深度估计装置

图 3

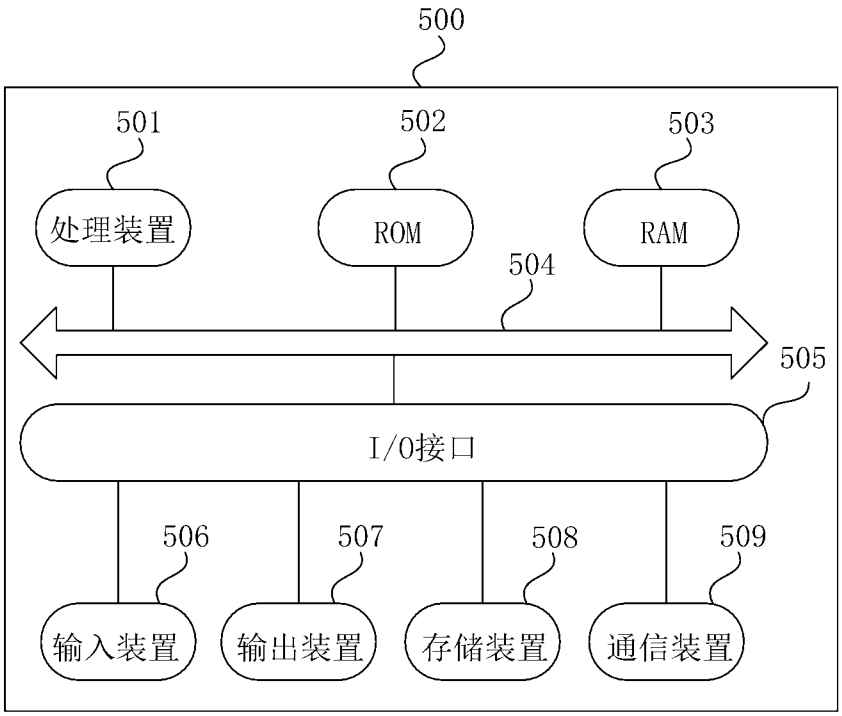

500
501
502
503
处理装置
ROM
RAM
504
505
I/O接口
506
507
508
509
输入装置
输出装置
存储装置
通信装置

图 4

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2023/114570

A. CLASSIFICATION OF SUBJECT MATTER

G06T 7/50(2017.01)i; G06T7/55(2017.01)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)

IPC: 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)

CNTXT, USTXT, VEN, CNKI, IEEE, BING, BAIDU: 深度, 逆深度, 均值, 实时, 上传视频, 动态, 三角化, 反投影, dynamic, depth, real time, upload, video, inverse depth, triangulation, back projection, average

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 105100771 A (SHANDONG UNIVERSITY) 25 November 2015 (2015-11-25) description, paragraphs [0013]-[0016]	1, 11-14
Y	CN 105100771 A (SHANDONG UNIVERSITY) 25 November 2015 (2015-11-25) description, paragraphs [0013]-[0016]	2-6
Y	CN 113643342 A (BEIJING DAJIA INTERNET INFORMATION TECHNOLOGY CO., LTD.) 12 November 2021 (2021-11-12) description, paragraphs [0006]-[0008], [0111], and [0179]	2-6
A	CN 113989717 A (BEIJING BYTEDANCE NETWORK TECHNOLOGY CO., LTD.) 28 January 2022 (2022-01-28) entire document	1-14
A	US 2013314400 A1 (KABUSHIKI KAISHA TOSHIBA) 28 November 2013 (2013-11-28) entire document	1-14

☐ 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	Date of mailing of the international search report
07 November 2023	**13 November 2023**

Name and mailing address of the ISA/CN	Authorized officer
China National Intellectual Property Administration (ISA/CN) **China No. 6, Xitucheng Road, Jimenqiao, Haidian District, Beijing 100088**	Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2023/114570

Patent document cited in search report	Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
CN 105100771 A	25 November 2015	None	
CN 113643342 A	12 November 2021	None	
CN 113989717 A	28 January 2022	None	
US 2013314400 A1	28 November 2013	JP 2013247577 A	09 December 2013

国际检索报告

国际申请号
PCT/CN2023/114570

A. 主题的分类

G06T 7/50(2017.01)i；G06T7/55(2017.01)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)

IPC：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））

CNTXT,USTXT,VEN,CNKI,IEEE,BING,BAIDU：深度，逆深度，均值，实时，上传视频，动态，三角化，反投影，dynamic，depth，real time，upload，video，inverse depth，triangulation，back projection，average

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
X	CN 105100771 A (山东大学) 2015年11月25日 (2015 - 11 - 25) 说明书第[0013]-[0016]段	1、11-14
Y	CN 105100771 A (山东大学) 2015年11月25日 (2015 - 11 - 25) 说明书第[0013]-[0016]段	2-6
Y	CN 113643342 A (北京达佳互联信息技术有限公司) 2021年11月12日 (2021 - 11 - 12) 说明书第[0006]-[0008]、[0111]、[0179]段	2-6
A	CN 113989717 A (北京字节跳动网络技术有限公司) 2022年1月28日 (2022 - 01 - 28) 全文	1-14
A	US 2013314400 A1 (KABUSHIKI KAISHA TOSHIBA) 2013年11月28日 (2013 - 11 - 28) 全文	1-14

☐ 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	国际检索报告邮寄日期
2023年11月7日	2023年11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	受权官员 王琦瑶 电话号码 (+86) 010-53961363

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2023/114570

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 105100771 A	2015年11月25日	无	
CN 113643342 A	2021年11月12日	无	
CN 113989717 A	2022年1月28日	无	
US 2013314400 A1	2013年11月28日	JP 2013247577 A	2013年12月9日