

(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
2021 年 3 月 4 日 (04.03.2021)



(10) 国际公布号
WO 2021/036353 A1

(51) 国际专利分类号:
G06T 17/00 (2006.01) **H04N 13/106** (2018.01)

(21) 国际申请号: PCT/CN2020/090469

(22) 国际申请日: 2020 年 5 月 15 日 (15.05.2020)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201910784772.1 2019 年 8 月 23 日 (23.08.2019) CN

(71) 申请人: 上海亦我信息技术有限公司(SHANG HAI
YIWO INFORMATION TECHNOLOGY CO., LTD.)
[CN/CN]; 中国上海市嘉定区陈翔路 768 号 6 幢
B 区 1120 室, Shanghai 201802 (CN)。

(72) 发明人: 赵明(ZHAO, Ming); 中国上海市嘉定区
陈翔路 768 号 6 幢 B 区 1120 室, Shanghai 201802

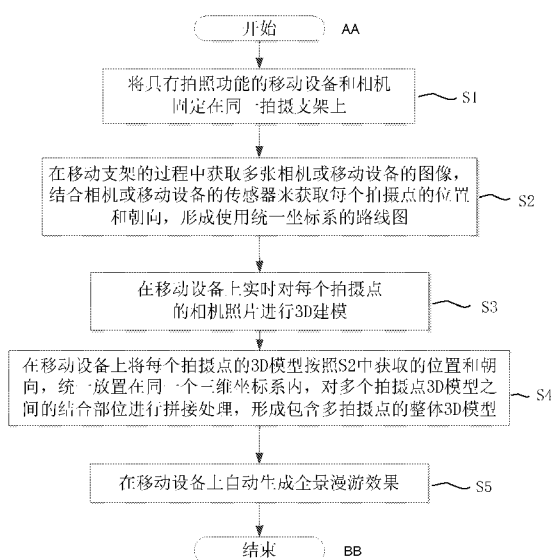
(CN)。 向中正(XIANG, Zhongzheng); 中国上海市嘉定区陈翔路 768 号 6 幢 B 区 1120 室, Shanghai 201802 (CN)。 蔡锴(CAI, Pei); 中国上海市嘉定区陈翔路 768 号 6 幢 B 区 1120 室, Shanghai 201802 (CN)。

(74) 代理人: 北京天达共和律师事务所等(EAST
& CONCORD PARTNERS et al.); 中国北京市朝阳区东三环北路 8 号亮马河大厦写字楼 2 座 19 层, Beijing 100004 (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,
CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB,
GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS,
JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK,
LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX,

(54) Title: PHOTOGRAPHING-BASED 3D MODELING SYSTEM AND METHOD, AND AUTOMATIC 3D MODELING APPARATUS AND METHOD

(54) 发明名称: 基于拍照的 3D 建模系统与方法、自动 3D 建模装置及方法



S1 FIX A MOBILE DEVICE AND A CAMERA WHICH HAVE A PHOTOGRAPHING FUNCTION ON THE SAME PHOTOGRAPHING SUPPORT
S2 OBTAIN A PLURALITY OF IMAGES OF THE CAMERA OR THE MOBILE DEVICE IN THE PROCESS OF MOVING THE SUPPORT, AND COMBINE A SENSOR OF THE CAMERA OR THE MOBILE DEVICE TO OBTAIN THE POSITION AND ORIENTATION OF EACH PHOTOGRAPHING POINT SO AS TO FORM A ROUTE MAP USING A UNIFIED COORDINATE SYSTEM
S3 PERFORM THREE-DIMENSIONAL (3D) MODELING ON A CAMERA PHOTO OF EACH PHOTOGRAPHING POINT IN REAL TIME ON THE MOBILE DEVICE
S4 ON THE MOBILE DEVICE, UNIFORMLY PLACE A 3D MODEL OF EACH PHOTOGRAPHING POINT IN THE SAME 3D COORDINATE SYSTEM ACCORDING TO THE POSITION AND ORIENTATION OBTAINED IN STEP S2, AND SPLICE JOINT PARTS OF THE 3D MODELS OF THE PLURALITY OF PHOTOGRAPHING POINTS TO FORM AN OVERALL 3D MODEL CONTAINING THE PLURALITY OF PHOTOGRAPHING POINTS
S5 AUTOMATICALLY GENERATE A PANORAMIC ROAMING EFFECT ON THE MOBILE DEVICE
AA START
BB END

图 4

(57) Abstract: Disclosed are a photographing-based three-dimensional (3D) modeling system and method, and an automatic 3D modeling apparatus and method. The photographing-based 3D modeling method comprises: S1) fixing a mobile device and a camera on the same photographing support; S2) obtaining a plurality of images for positioning of the camera or the mobile device in the process of moving the support, and obtaining the position and orientation of each photographing point to form a route map using a unified coordinate system; S3) performing 3D modeling on the mobile device or a remote server on the basis of an image for 3D modeling at each photographing point; and S4) uniformly placing a 3D model of each photographing point in the same 3D coordinate system according to the position and orientation obtained in step S2, and splicing joint parts of the 3D models of the plurality of photographing points to form an overall 3D model containing the plurality of photographing points. According to the present invention, various photographing devices can be supported, and the relative position of each photographing point and the lens orientation of the camera can

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, GH, GM, KE, LR, LS, MW, MZ,
NA, RW, 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, 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))。

be automatically and accurately recorded, and a 3D model and 2D planar graph which can roam inside can be automatically generated.

(57) 摘要: 本发明公开了一种基于拍照的3D建模系统及方法、自动3D建模装置及方法, 其中包括: S1) 将移动设备和相机固定在同一拍摄支架上; S2) 在移动支架的过程中获取多张相机或移动设备的定位用图像, 获取每个拍摄点的位置和朝向, 形成使用统一坐标系的路线图; S3) 在所述移动设备上或远程服务器端基于在每个拍摄点的3D建模用图像进行3D建模; S4) 将每个拍摄点的3D模型按照S2中获取的位置和朝向, 统一放置在同一个三维坐标系内, 对多个拍摄点3D模型之间的结合部位进行拼接处理, 形成包含多拍摄点的整体3D模型。本发明能够支持多种拍摄设备, 自动精确记录各个拍摄点的相对位置和相机的镜头朝向, 自动生成可在内部漫游的3D模型以及2D平面图。

基于拍照的 3D 建模系统及方法、自动 3D 建模装置及方法

相关申请的交叉引用

本申请基于申请号为 201910784772.1、申请日为 2019 年 08 月 23 日的中国专利申请提出，并要求该中国专利申请的优先权，该中国专利申请的全部内容在此引入本申请作为参考。

技术领域

本发明涉及一种 3D 建模系统及方法，尤其涉及一种基于拍照的 3D 建模系统及方法、自动 3D 建模装置及方法。

背景技术

本发明所要解决的技术问题是一种基于拍照的三维空间建模方案，可用于单空间或多空间的 3D 建模和/或 2D 平面图生成。

以往，基于拍照的 3D 建模方法主要有以下两种，但是都有明显的缺点：

方法 a) 使用能够识别深度信息的相机，直接生成 3D 模型。这种方法依赖于更加复杂的硬件，导致更高的设备成本，往往需要专业摄影师才能操作，不利于普及；

方法 b) 在一个拍摄点距离较近的地方拍摄两张照片，优选间隔厘米级和分米级并连续进行匹配和定位，然后使用 MVS (Multi View Stereo, 参考 <https://github.com/cdcseacave/openMVS>) 进行建模，优点是整个过程全自动不需要人工干预，但是缺点也很明显：

缺点 1: 计算量较大，不易在移动设备等计算资源有限的设备上建模，通常需要上传到服务器 (云端/PC)，利用更强的计算能力，运行建模算法；

缺点 2: 无法给出具体间隔多远拍摄的规范，如果太密则使用太麻烦，耗时过长；如果只是靠两两可见，靠感觉，可能建模失败，而拍摄时无法给用户提示。

此外，以往还提供过基于拍照重建三维空间场景的方法，但这些方法中基本都

无法基于 3D 建模用的图像自动生成 3D 模型，需要大量的人工操作介入来修正每个空间的 3D 模型。另外，无法对多个空间各自的 3D 模型进行自动拼接处理，需要人工观察各空间的 3D 模型并找出其共同点后，手动进行拼接，耗时耗力。

发明内容

为了克服以往方法的上述缺点中的一个或多个，本发明使用了创新的方法，即使用深度学习和图像处理方法进行单拍摄点建模，既可以在移动设备那样的有限的计算能力下运行，也可以将相关数据上传到云端服务器运行。此外，在利用移动设备实时建模的情况下，为了提高实时性，只对房间轮廓建模，不复原家具、饰品等物体模型；通过构建拍摄点定位系统，把多个拍摄点的模型按照位置和方向放在同一坐标系内；对多个拍摄点的独立模型进行优化处理，处理好结合部分，生成整体 3D 模型、2D 平面图。

本发明支持的拍摄方式广泛，包括但不限于手机鱼镜头、全景相机、带鱼镜头的相机，以及普通手机和普通数码相机等，成本低廉。

普通照片（定义）：使用普通数码相机（包括普通单反、微单、Point&Shoot 相机等）、全景相机、带鱼镜头的相机，普通手机和带鱼镜头的手机，以及摄像头采集的照片。区别于双目视觉，普通照片并不能通过同一拍摄点拍摄的两张照片复原三维信息。普通照片以下简称照片。

使用全景相机时，通常获取的是全景图。部分计算机视觉、图像算法，例如直线检测，需要将全景图转换为不变形的图片。以下使用的照片和图片表述包括全景照片和经转换过的不变形的图片。

本发明提供一种基于拍照的 3D 建模系统及方法、自动 3D 建模装置及方法，能够支持多种拍摄设备，并基于获取的各个拍摄点的相对位置和相机的镜头拍摄方向信息，自动将各个拍摄点的 3D 模型拼接而生成整体 3D 模型。本发明还能生成 2D 平面图。

具体而言，本发明提供一种基于拍照的 3D 建模系统，其包括：拍摄单元，用于对多个空间分别拍摄该空间的第一图像；3D 模型生成单元，分别基于所述拍摄单元针对各个所述空间拍摄的所述第一图像，生成各个所述空间的 3D 模型；拍摄位置获取单元，用于获取所述拍摄单元拍摄各个空间的第一图像时的位置和拍摄方向信息；3D 模型拼接单元，基于所述位置和拍摄方向信息，将所述各个空间的 3D 模型在同一

个三维坐标系内进行拼接处理，形成包括所述各个空间的整体 3D 模型。

进一步地，所述拍摄单元在所述各个空间之间移动的过程中拍摄多张第二图像，所述拍摄位置获取单元通过所述多张第二图像进行特征点匹配，来获得各拍摄点的相对位移和拍摄方向信息，例如可形成包括所有拍摄点的在同一坐标系内的路线图，从而获取所述拍摄单元对所处空间拍摄第一图像时的位置和拍摄方向信息。

进一步地，所述拍摄单元具有定位传感器和方向传感器；所述拍摄位置获取单元基于所述拍摄单元对所处空间拍摄第一图像时提供的定位信息和方向信息，来获取所述拍摄单元对所处空间拍摄第一图像时的位置和拍摄方向信息。

进一步地，所述拍摄单元在所述各个空间之间移动的过程中拍摄多张第二图像；所述拍摄单元具有定位传感器和方向传感器；所述拍摄位置获取单元通过所述拍摄单元拍摄的多张第二图像中的相近拍摄点图像进行特征点匹配，来获得各拍摄点的相对位移和拍摄方向信息，例如可形成包括所有拍摄点的在同一坐标系内的路线图，并通过结合所述拍摄单元对所处空间拍摄第一图像时的定位信息和方向信息对路线图进行校正，从而获取所述拍摄单元对所处空间拍摄第一图像时的位置和拍摄方向信息。

进一步地，所述拍摄位置获取单元还通过所述拍摄单元具备的位移传感器例如可以包括加速度传感器、速度传感器提供的位移信息例如加速度信息和移动速度信息，对所述相对位移即路线图和拍摄方向信息进行校正。

进一步地，所述 3D 模型拼接单元根据所述拍摄位置获取单元获取的拍摄各个房间时的所述位置和拍摄方向信息，例如可利用一个转换矩阵，把单个房间的所述 3D 模型的局部坐标转换为全局的世界坐标，从而获得所有拍摄点的整体 3D 模型。

进一步地，把单个房间的所述 3D 模型的局部坐标转换为全局的世界坐标的方法，包括，使所述拍摄单元移动预定距离，所述拍摄位置获取单元获取所述预定距离的两个端点的坐标，所述两个端点的坐标的差和所述预定距离的比值即为所述局部坐标和所述世界坐标的比值；或者利用所述拍摄位置获取单元识别到的特征点，预估所述空间的底面或顶面所在的平面的高度与所述拍摄单元的高度的比值，即为所述局部坐标和所述世界坐标的比值。

进一步地，在所述拍摄单元进行第一个拍摄点的拍摄之前、或在后续拍摄的移动过程中，移动预定距离以获得预定数量的所述特征点。

进一步地，所述拍摄单元具有双目镜头，在同一个拍摄点，该双目镜头分别拍摄

所述第一图像；所述 3D 模型生成单元通过对所述双目镜头各自拍摄的所述第一图像进行图像比对，确定对应像素，并获得每个所述对应像素的深度信息，以用于生成所述 3D 模型。

进一步地，所述 3D 模型生成单元通过深度学习技术，预测所述第一图像中每个像素的深度，计算或直接运用所述深度学习技术预测所述每个像素的法线方向，以生成各个所述空间的 3D 模型。

进一步地，所述拍摄单元由相机和/或带拍照功能的移动设备例如手机实现；所述 3D 模型生成单元由所述手机实现或者由远程服务器实现，在由所述远程服务器实现的情况下，其通过网络接收所述相机和/或所述带拍照功能的手机拍摄并发送来的第一图像，生成各个所述空间的 3D 模型；所述拍摄位置获取单元由所述相机或所述手机实现；所述 3D 模型拼接单元，由所述手机实现或者由远程服务器实现，在由所述远程服务器实现的情况下，其通过网络接收所述拍摄位置获取单元发送来的各个空间的所述位置和拍摄方向信息，基于该位置和拍摄方向信息完成所述拼接处理，并将生成的所述整体 3D 模型发送给所述手机或其它设备。

进一步地，实现所述拍摄单元的相机和带拍照功能的手机固定在同一拍摄支架上；在移动所述支架的过程中，获取多张所述相机或所述带拍照功能的手机所拍摄的第二图像，从而获取所述相机或所述带拍照功能的手机对所处空间拍摄第一图像时的位置和拍摄方向信息。

进一步地，基于所述相机或所述带拍照功能的手机的定位系统，使用所述相机或所述带拍照功能的手机拍摄的第二图像，通过相近拍摄点的第二图像进行特征点匹配来获得各拍摄点的相对位移和拍摄方向信息，从而提供每个拍摄点的相对位置和方向。

进一步地，在所述拍摄单元对第一个所述空间拍摄所述第一图像之前、或在后续拍摄的移动过程中，通过以下方式的一种或多种，获取所述相机的镜头和所述手机的拍摄方向的夹角：

这里，相机的镜头的拍摄方向可以是指常见的全景相机的前后两个鱼眼镜头中的其中一个镜头的方向；或者是指通过旋转一个镜头来拍摄多张照片的全景相机的拍摄第一张照片时的镜头的方向。

(1) 同时运行基于所述手机的定位系统和基于所述相机的定位系统，并移动所述支架一段距离，此时两套系统各提供一个位移向量，两个向量的夹角即为所述相

机的镜头与所述手机的拍摄方向的夹角；

(2) 通过手工旋转所述相机的预览图或者拍摄的图像，指定与所述手机拍摄方向一致的角度；

(3) 通过图像识别算法匹配所述手机与所述相机的预览图或拍摄的图像，找到所述夹角；

(4) 通过使用额外的标识、包括在支架上增加刻度与所述手机的安装方向形成固定夹角，然后在所述相机的预览图或图像中识别该标识来计算所述相机的镜头和所述手机拍摄方向的夹角；

(5) 通过在所述支架上设置卡槽，保证所述相机与所述手机（移动设备）形成已知的固定夹角。

进一步地，所述空间是房间；所述第一图像是所述房间的室内图像；所述 3D 模型生成单元基于深度学习技术，识别所述第一图像中的地板、天花板、墙壁中的至少一者的图像，并基于图像处理技术，对识别出的图像进行分块，每块近似认为是一个平面，使地板和天花板的图像块位于水平平面，使墙壁的图像块位于垂直平面，通过求解各个平面方程生成所述 3D 模型；其中，对于在所述第一图像中相交的两个平面，使得计算得到的交线与实际观察到的交线的误差最小。

进一步地，所述 3D 模型生成单元还利用计算机视觉技术识别所述室内图像中的墙角，并将墙角连接起来，生成房间的粗略模型。

进一步地，所述 3D 模型拼接单元对所述多个房间各自的 3D 模型进行统一校正，包括使用统计的方法对所有的房间的墙线方向进行统一校正，使得所有房间在一定偏差范围内的墙线平行；所述 3D 模型拼接单元在将所述各个房间的 3D 模型进行拼接处理时，对重叠的部分和/或出现的空洞进行修正。

进一步地，本发明的基于拍照的 3D 建模系统还包括 2D 平面图生成单元，以如下方式生成 2D 平面图：将生成的所述 3D 模型的每个面向垂直于地板的平面投影，并将这些投影合并成为一个多边形；对得到的所述多边形进行校正和简化，包括以下至少一种方式：(1) 仅保留所述多边形的主要顶点，删除小的凹凸，(2) 利用计算机视觉技术检测图片中的直线，进而确定墙壁的走向，并将与墙壁走向近似平行或垂直的边都修正到对应的方向上；基于所述拍摄位置获取单元获取的各个空间的所述位置和拍摄方向信息，将生成的所述各个房间的 2D 平面图在同一个二维坐标系内进行拼接处理，形成由所述各个房间各自的 2D 平面图拼接而成的整体 2D 平面图；

识别并标示出门和/或窗的位置，包括使用深度学习的方法，在所述室内图像上识别出门和/或窗的位置，或者根据所述拍摄单元对属于同一套房子的多个房间分别拍摄所述第一图像时的移动轨迹与该房间轮廓的交点，确定门的位置。

进一步地，所述 2D 平面图生成单元对所述多个房间各自的 2D 平面图进行统一校正，包括使用统计的方法对所有的房间的墙线方向进行统一校正，使得所有房间在一定偏差范围内的墙线平行；所述 2D 平面图生成单元在将所述各个房间的 2D 平面图进行拼接处理时，对重叠的部分和/或出现的空洞进行修正。

进一步地，本发明的基于拍照的 3D 建模系统也可以包括：2D 平面图生成单元，以如下方式生成 2D 平面图：将所述 3D 模型拼接单元生成的所述整体 3D 模型的每个面向垂直于地板的平面投影，并将这些投影合并成为一个多边形；对得到的所述多边形进行校正和简化，包括以下至少一种方式：（1）仅保留所述多边形的主要顶点，删除小的凹凸，（2）利用计算机视觉技术检测图片中的直线，进而确定墙壁的走向，并将与墙壁走向近似平行或垂直的边都修正到对应的方向上；识别并标示出门和/或窗的位置，包括使用深度学习的方法，在所述室内图像上识别出门和/或窗的位置，或者根据所述拍摄单元对属于同一套房子的多个房间分别拍摄所述第一图像时的移动轨迹与该房间轮廓的交点，确定门的位置。

此外，本发明还提供一种自动 3D 建模装置，包括：3D 模型生成单元，基于建模对象所包括的多个空间各自的第一图像，分别生成各个所述空间的 3D 模型；3D 模型拼接单元，基于所述多个空间各自的所述第一图像被拍摄时的位置和拍摄方向信息，将所述 3D 模型生成单元生成的所述各个空间的 3D 模型在同一个三维坐标系内进行拼接处理，形成由所述各个空间各自的 3D 模型拼接而成的整体 3D 模型。

此外，本发明还提供一种自动 3D 建模方法，包括：3D 模型生成步骤，基于建模对象所包括的多个空间各自的第一图像，分别生成各个所述空间的 3D 模型；3D 模型拼接步骤，基于所述多个空间各自的所述第一图像被拍摄时的位置和拍摄方向信息，将所述 3D 模型生成步骤中生成的所述各个空间的 3D 模型在同一个三维坐标系内进行拼接处理，形成由所述各个空间各自的 3D 模型拼接而成的整体 3D 模型。

此外，本发明还提供一种基于拍照的 3D 建模方法，包括如下步骤：S1) 将具有拍照功能的移动设备和相机固定在同一拍摄支架上；S2) 在移动支架的过程中获取多张相机或移动设备的第二图像，结合所述相机或移动设备的传感器来获取每个拍摄点的位置和拍摄方向，形成使用统一坐标系的路线图；S3) 在所述移动设备上或远

程服务器端基于在每个拍摄点拍摄的第一图像进行 3D 建模；S4) 将每个拍摄点的所述 3D 模型按照 S2 中获取的位置和拍摄方向，统一放置在同一个三维坐标系内，对多个拍摄点 3D 模型之间的结合部位进行拼接处理，形成包含多拍摄点的整体 3D 模型。

进一步地，所述步骤 S2 是基于所述移动设备或相机的定位系统，使用所述移动设备或相机拍摄的第二图像，通过相近拍摄点的第二图像进行特征点匹配来找到拍摄点的相对位移和拍摄方向信息，形成包括所有拍摄点的在同一坐标系内的所述路线图，并提供每个拍摄点的位置和方向。

进一步地，所述步骤 S2 还包括通过所述移动设备或相机的传感器，获取包括移动的加速度、速度和方向信息，对所述路线图进行校正。

进一步地，所述步骤 S2 还包括获取所述相机的镜头和所述移动设备的拍摄方向的夹角，在初始化阶段同时运行基于所述移动设备的定位系统和基于所述相机的定位系统，并移动支架一段距离，此时两套系统各提供一个位移向量，两个向量的夹角即为所述相机的镜头与所述移动设备的拍摄方向的夹角；或者通过手工旋转所述相机的预览图或者拍摄的图像，指定与所述移动设备拍摄方向一致的角度；或者通过图像识别算法匹配所述移动设备与所述相机的预览图或拍摄的图像，找到夹角；或者通过使用额外的标识、包括在支架上增加刻度与所述移动设备的安装方向形成固定夹角，然后在所述相机的预览图或图像中识别该标识来计算所述相机的镜头和所述移动设备拍摄方向的夹角。

进一步地，所述步骤 S3 包括：S31) 基于深度学习技术，识别图像中的地板、天花板、墙壁的至少一者的图像；S32) 基于图像处理技术，对识别出的图像进行分块，每块近似认为是一个平面，地板和天花板的图像块位于水平平面，墙壁的图像块位于垂直平面，求解各个平面方程生成所述 3D 模型；对于在图像中相交的两个平面，以这两个平面的一条交线为约束条件，使得计算得到的交线与实际观察到的交线的误差最小。

进一步地，所述步骤 S3 还包括：对于室内图像，利用计算机视觉技术识别图像中的墙角，并将墙角连接起来即为房间的粗略模型。

进一步地，所述步骤 S4 包括：S41) 根据每个拍摄点的位置和拍摄方向，例如可利用一个转换矩阵，把单个拍摄点的 3D 模型的局部坐标转换为成全局的世界坐标，从而获得所有拍摄点的整体 3D 模型；S42) 对多个拍摄点的 3D 模型进行统一校正，

包括使用统计的方法对所有的拍摄点的墙线方向进行统一校正，使得所有房间在一定偏差范围内的墙线平行；S43）在将各个拍摄点的 3D 模型进行拼接处理时，对其重叠的部分和/或出现的空洞进行修正。

本发明对比现有技术，可实现如下有益效果的一种或多种：能够支持多种拍摄设备；3D 建模及拼接等工作既能够在移动设备这样的计算能力有限的设备中运行，又能够在远程服务器端进行；能够基于获取的各个拍摄点的相对位置和相机的镜头拍摄方向信息，自动将各个拍摄点的 3D 模型拼接而生成整体 3D 模型；还能根据需要而生成 2D 平面图。本发明建模成功率高；支持每个房间只拍摄一组图像并合成全景图，效率高，用户体验好；建模效率高，既可在拍摄过程中实时建模，又可在远程服务器端精确建模；所见即所得，用户可以参照实时建模结果选择拍摄点位，防止漏拍；建模可不包括家具等干扰物，有利于生成准确的户型图。

附图说明

图 1 为本发明可以应用于其中的一个示例性系统架构图；

图 2 为本发明的基于拍照的 3D 建模系统的一个实施例的结构示意图；

图 3 为本发明的基于拍照的 3D 建模系统的另一个实施例的结构示意图；

图 4 为本发明的基于拍照的 3D 建模方法的一个实施例的流程示意图；

图 5 为本发明的自动 3D 建模装置的一个实施例的结构示意图；

图 6 为本发明的自动 3D 建模装置的另一个实施例的结构示意图；

图 7 为本发明的自动 3D 建模方法的一个实施例的流程示意图；

图 8 为本发明的电子设备的一个实施例的结构示意图。

结合附图并参考以下具体实施方式，本发明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。

具体实施方式

除非另有定义，本文所使用的所有的技术和科学术语与属于本发明的技术领域的技术人员通常理解的含义相同；本文中在申请的说明书中所使用的术语只是为了描述具体的实施例的目的，不是旨在于限制本发明；本发明的说明书和权利要求书及上述附图说明中的术语“包括”和“具有”以及它们的任何变形，意图在于覆盖不

排他的包含。本发明的说明书和权利要求书或上述附图中的术语“第一”、“第二”等是用于区别不同对象，而不是用于描述特定顺序。

在本文中提及“实施例”意味着，结合实施例描述的特定特征、结构或特性可以包含在本发明的至少一个实施例中。在说明书中的各个位置出现该短语并不一定均是指相同的实施例，也不是与其它实施例互斥的独立的或备选的实施例。本领域技术人员显式地和隐式地理解的是，本文所描述的实施例可以与其它实施例相结合。

为了使本技术领域的人员更好地理解本发明的方案，下面结合附图和实施例对本发明作进一步的描述。

[系统结构]

首先，说明本发明的一个实施例的系统的结构。如图 1 所示，系统结构 100 可以包括移动设备 101、102、103、104，网络 105 和服务器 106。网络 105 用以在终端设备 101、102、103、104 和服务器 106 之间提供通信链路的介质。

在本实施例中，图 1 所示的移动设备 101、102、103 或 104 可以通过网络 105 进行各种信息的传输。网络 105 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。需要指出的是，上述无线连接方式可以包括但不限于 3G/4G/5G 连接、Wi-Fi 连接、蓝牙连接、WiMAX 连接、Zigbee 连接、UWB 连接、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）以及其他现在已知或将来开发的网络连接方式。网络 105 可以利用诸如 HTTP（Hyper Text 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。

用户可以使用移动设备 101、102、103、104 通过网络 105 与服务器 106 交互，以接收或发送消息等。移动设备 101、102、103 或 104 上可以安装有各种客户端应用，例如视频直播与播放类应用、网页浏览器应用、购物类应用、搜索类应用、即时通信工具、邮箱客户端、社交平台软件等。

移动设备 101、102、103 或 104 可以是具有触摸显示屏和/或支持网页浏览的各种电子设备，并且具有拍照功能，包括但不限于智能手机、平板电脑、电子书阅读器、MP3 播放器（动态影像专家压缩标准音频层面 3）、MP4（动态影像专家压缩标准音频层面 4）播放器、头戴式显示设备、笔记本电脑、数字广播接收器、PDA（个人数字助理）、PMP（便携式多媒体播放器）、车载终端等等的移动终端以及诸如数字 TV、台式计算

机等等。

服务器 106 可以是提供各种服务的服务器，例如对移动设备 101、102、103 或 104 上进行 3D 建模提供支持的后台服务器。

应该理解，图 1 中的移动设备、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移动设备、网络和服务器的。

这里，移动设备可以独立或通过与其他电子终端设备例如相机固定在同一个支架（例如三脚架）上，配合运行安卓系统中的应用实现本发明的实施例方法，也可以运行其他操作系统中的应用例如 iOS 系统、Windows 系统、鸿蒙系统等的应用实现本发明的实施例方法。

[基于拍照的 3D 建模系统]

图 2 为本发明基于拍照的 3D 建模系统的一个实施例的结构示意图。如图 2 所示，该实施例中基于拍照的 3D 建模系统包括：拍摄单元 201，用于对多个所处空间分别拍摄该空间的第一图像，这里第一图像例如可以是 3D 建模用图像，可以包括普通照片、全景照片、处理过的全景照片（例如去变形等）。该拍摄单元 201 可以由移动设备中的拍摄模块来实现。

这里，拍摄单元 201 可以在各个空间之间移动的过程中拍摄多张第二图像，这里第二图像例如可以是定位用图像，可以包括普通照片、全景照片、处理过的全景照片（例如去变形等）。这里，第一图像和第二图像可以为相同的图像，也可以是部分相同的图像，也可以是不同的图像，并不做限定。此处的定位用图像还可以是拍摄单元 201 拍摄的照片、预览图、视频帧等，可以存储，也可以不存储而仅用于进行特征点的识别及匹配。

这里，拍摄单元 201 例如具有定位传感器和方向传感器，能够获得在对所处空间拍摄 3D 建模用图像时的定位信息和方向信息，这里，定位传感器例如可以是加速度传感器、陀螺仪、线性加速度传感器、旋转矢量传感器、重力传感器等中的一个或多个；方向传感器例如可以是方向传感器、磁力传感器等中的一个或多个。

3D 模型生成单元 202，分别基于拍摄单元 201 针对各个空间拍摄的 3D 建模用图像，生成各个空间的 3D 模型；

在一个或多个实施例中，例如，拍摄单元 201 具有双目镜头，在同一个拍摄点，该双目镜头分别拍摄 3D 建模用图像；则 3D 模型生成单元 202 通过该双目镜头各自拍摄的 3D 建模用图像进行图像比对，确定对应像素，并获得每个对应像素的深度

信息，以用于生成 3D 模型。

当然，在一个或多个实施例中，例如，3D 模型生成单元 202 还可以通过深度学习技术，预测 3D 建模用图像中每个或一些像素的深度，计算或直接运用深度学习技术预测每个或一些像素的法线方向，以生成各个空间的 3D 模型。

这里，在一个或多个实施例中，深度学习技术预测 3D 建模用图像中每个像素的深度或预测每个像素的法线方向的方法例如可以为一种用于从单个室内 360° 图像中预测密集深度和表面法线以及平面边界的训练平面感知卷积神经网络的方法（比如可参见 Pano Pops: Indoor 3D Reconstruction with a Plane-Aware Network）；或者使用大规模三维数据集以端到端的方式学习从 360° 图像中预测深度的方法（比如可参见 OmniDepth: Dense Depth Estimation for Indoors Spherical Panoramas）。

拍摄位置获取单元 203，用于获取拍摄单元 201 拍摄各个空间的 3D 建模用图像时的位置和拍摄方向信息，当然还可以获取包括镜头的焦距、镜头的扫描区间、以及其他可以影响图像内容的拍摄参数比如焦距、广角或者望远镜头设置等会影响图像内容特征的失真或影响图像内容特征的相对尺寸的拍摄参数信息；

这里，拍摄位置获取单元 203 例如可以通过拍摄单元 201 拍摄的多张定位用图像中的相近拍摄点图像进行特征点匹配，来获得各拍摄点的相对位移和拍摄方向信息，例如可形成包括所有拍摄点的在同一坐标系内的路线图，从而获取拍摄单元 201 对所处空间拍摄 3D 建模用图像时的位置和拍摄方向信息。

这里，拍摄位置获取单元 203 还可以例如基于拍摄单元 201 对所处空间拍摄 3D 建模用图像时提供的定位信息和方向信息，来获取拍摄单元 201 对所处空间拍摄 3D 建模用图像时的位置和拍摄方向信息。

这里，拍摄位置获取单元 203 还通过拍摄单元 201 具备的传感器，包括位移传感器例如加速度传感器、速度传感器等以及陀螺仪、气压传感器或其他动作传感器提供的位移信息例如加速度信息和移动速度信息或其他动作/运动信息等，对相对位移形成的路线图和拍摄方向信息进行校正。

3D 模型拼接单元 204，基于拍摄位置获取单元 203 获取的各个空间的位置和拍摄方向信息，将 3D 模型生成单元 202 生成的各个空间的 3D 模型在同一个三维坐标系内进行拼接处理，形成由各个空间各自的 3D 模型拼接而成的整体 3D 模型。

这里，3D 模型拼接单元 204 还可以根据拍摄位置获取单元 203 获取的拍摄各个房间时的位置和拍摄方向信息，例如利用一个转换矩阵，把单个房间的 3D 模型的局

部坐标转换为全局的世界坐标，从而获得所有拍摄点的整体 3D 模型。

这里，把单个房间的 3D 模型的局部坐标转换为全局的世界坐标的方法，包括，使拍摄单元 201 移动预定距离，拍摄位置获取单元 203 获取预定距离（例如为 1 米）的两个端点的坐标，两个端点的坐标的差和预定距离的比值即为局部坐标和世界坐标的比值；或者，利用拍摄位置获取单元 203 识别到的特征点，预估该空间的底面或顶面所在的平面的高度与拍摄单元 201 的高度的比值，即为局部坐标和世界坐标的比值，其中，在拍摄单元 201 进行第一个拍摄点的拍摄之前、或在后续拍摄的移动过程中，移动预定距离以获得预定数量的特征点。

这里，预估该空间的底面或顶面所在的平面的高度与拍摄单元 201 的高度的比值的方法例如为，把拍摄点垂直投影到底面，再连接特征点（在底面上），这三个点形成一个三角形。假设投影线为 L1，拍摄点到特征点连线为 L2，投影点到特征点的连线为 L3；L1 和 L2 之间的夹角已知（根据全景图的特性），根据 L3 的长度以及以上夹角，使用三角函数可以算出 L1，再根据与实际的相机高度计算出比例尺。

这里，预定距离需要满足获得预定数量的特征点的足够距离。

具体来说，在一个或多个实施例中，例如，拍摄单元 201 只使用相机或者手机摄像头，由于得到的坐标都是相对值，必须将坐标转换为绝对值。也就是说图像比对的算法通常不具有精确的比例尺，坐标是相对的，没有具体尺寸。这也导致由不同图片计算出来的位移、比例尺不统一，导致错位。上述转换坐标的方法具体实现时可以是：

a) 可以让用户移动指定的距离（例如 1 米），获取该移动距离的两个端点的坐标，两端点坐标的差和移动距离的比值即为局部坐标和世界坐标的坐标比值；

b) 利用系统识别到的特征点，预估例如房间地面或顶面所在的平面。假设坐标系中垂直方向的坐标轴为 z 轴，该平面的方程为 $z = a$ ；由于拍摄单元 201 的高度已知（或拍摄单元 201 到房顶的高度已知），为 h，则 a/h 即为局部坐标与世界坐标的比值。这里，由于估算 a 需要找到同一平面（地面或房顶）上的一定数量的特征点，在实现中可以设置初始化过程，即移动足够长的一段距离（例如 2 米以上），才能在不同的环境中都能积累足够多的特征点。该初始化过程可以放在第一个拍摄点之前，如果初始化失败，可以再次进行，不影响后续拍摄；也可以把初始化过程放在后续拍摄点之间的移动过程中。

在该实施例中，拍摄单元 201 例如可以由相机和/或带拍照功能的手机实现；

在一个或多个实施例中，例如可以将实现拍摄单元 201 的相机和带拍照功能的手机固定在同一拍摄支架上；在移动支架的过程中，获取多张相机或带拍照功能的手机所拍摄的定位用图像，从而获取相机或带拍照功能的手机对所处空间拍摄 3D 建模用图像时的位置和拍摄方向信息。

这里，还可以基于相机或所述带拍照功能的手机的定位系统，使用相机或带拍照功能的手机拍摄的定位用图像，通过相近拍摄点的定位用图像进行特征点匹配来获得各拍摄点的相对位移和拍摄方向信息，从而提供每个拍摄点的相对位置和方向。

在一个或多个实施例中，由于拍摄点的位置、方向和路线图通过手机获得。由于相机可以通过旋转螺丝固定到支架上，每次安装相机与手机的夹角可能不同（但是在拍摄一套房源的过程中夹角不变）。单房间的 3D 模型需要转动这个夹角，再根据手机获取的位置和拍摄方向，放入到全局坐标中，形成整体的 3D 模型。

这里，在拍摄单元 201 对第一个空间拍摄 3D 建模用图像之前、或在后续拍摄的移动过程中，可以通过以下方式的一种或多种，获取相机的镜头和手机的拍摄方向的夹角：

这里，相机的镜头的拍摄方向可以是指常见的全景相机的前后两个鱼眼镜头中的其中一个镜头的方向；或者是指通过旋转一个镜头来拍摄多张照片的全景相机的拍摄第一张照片时的镜头的方向。

(1) 同时运行基于手机的定位系统和基于相机的定位系统，并移动支架一段距离，此时两套系统各提供一个位移向量，两个向量的夹角即为相机的镜头与手机的拍摄方向的夹角；

(2) 通过手工旋转相机的预览图或者拍摄的图像，指定与手机拍摄方向一致的角度；

(3) 通过图像识别算法匹配手机与相机的预览图或拍摄的图像，找到夹角，这里，找到夹角可能的实现方法至少可以包括以下方式之一：

计算手机与相机两个设备采集到的图像中的特征值，例如通过尺度不变特征变换 (Scale-invariant feature transform, SIFT)，根据匹配的特征值在两个图像中的位置差异，计算两个镜头拍摄方向的夹角；或

用两个镜头采集到的视频流分别构建基于视觉的同步定位与映射系统 (Visual Simultaneous Localization And Mapping, VSLAM)，两个系统内相机的位移的夹角即为镜头拍摄方向的夹角；

(4) 通过使用额外的标识、包括在支架上增加刻度与手机的安装方向形成固定夹角，然后在相机的预览图或图像中识别该标识来计算相机的镜头和手机拍摄方向的夹角；

(5) 通过在所述支架上设置卡槽，保证所述相机与所述手机（移动设备）形成已知的固定夹角。

当然，这里，拍摄点的位置、方向和路线图也可以由相机图像计算得到，在这种情况下，计算 3D 模型不依赖于相机与手机的夹角，所以手机可以不用固定在支架上。

这里，如果相机也具备方向传感器，可以通过直接获取相机和手机的方向计算夹角。

3D 模型生成单元 202 由手机实现或者由远程服务器实现，在由远程服务器实现的情况下，其通过网络接收相机和/或带拍照功能的手机拍摄并发送来的 3D 建模用图像，定位用图像，和/或动作传感器获得的信息，生成各个所述空间的 3D 模型；

拍摄位置获取单元 203 例如可以由所述相机或所述手机实现；

3D 模型拼接单元 204，例如可以由所述手机实现或者由远程服务器实现，在由所述远程服务器实现的情况下，其通过网络接收拍摄位置获取单元 203 发送来的各个空间的所述位置和拍摄方向信息，基于该位置和拍摄方向信息完成拼接处理，并将生成的整体 3D 模型发送给所述手机或其它设备。图 3 为本发明基于拍照的 3D 建模系统的另一个实施例的结构示意图。如图 3 所示，该实施例的基于拍照的 3D 建模系统中，基于拍照的 3D 建模空间例如是指房间；3D 建模用图像例如是该房间的室内图像，包括：

拍摄单元 301，用于对多个房间分别拍摄该房间的 3D 建模用图像；

这里，拍摄单元 301 可以在各个房间之间移动的过程中拍摄多张定位用图像。

这里，拍摄单元 301 例如具有定位传感器和方向传感器，能够获得在对所处房间拍摄 3D 建模用图像时的定位信息和方向信息。

3D 模型生成单元 302，分别基于拍摄单元 301 针对各个房间拍摄的 3D 建模用图像，生成各个房间的 3D 模型；

这里，3D 模型生成单元 302 基于深度学习技术，识别 3D 建模用图像中的地板、天花板、墙壁中的至少一者的图像，并基于图像处理技术，对识别出的图像进行分块，每块近似认为是一个平面，使地板和天花板的图像块位于水平平面，使墙壁的图像块位于垂直平面，通过求解各个平面方程生成 3D 模型；其中，对于在 3D 建模用

图像中相交的两个平面，使得计算得到的交线与实际观察到的交线的误差最小。

这里，3D 模型生成单元 302 还利用计算机视觉技术识别室内图像中的墙角，并将墙角连接起来，生成房间的粗略模型。

这里，在一个或多个实施例中，识别图像中的墙角的方法例如可以使用兴趣点检测与描述自监督训练框架（例如可参见 SuperPoint: Self-Supervised Interest Point Detection and Description），再将墙角连接起来形成房间的粗略模型，以实现能够捕捉在同一三维空间结构中频繁出现的对象例如墙角之间的几何关系。

拍摄位置获取单元 303，用于获取拍摄单元 301 拍摄各个房间的 3D 建模用图像时的位置和拍摄方向信息；

这里，拍摄位置获取单元 303 例如可以通过拍摄单元 301 拍摄的多张定位用图像中的相近拍摄点图像进行特征点匹配，来获得各拍摄点的相对位移和拍摄方向信息，例如可形成包括所有拍摄点的在同一坐标系内的路线图，从而获取拍摄单元 301 对所处房间拍摄 3D 建模用图像时的位置和拍摄方向信息。

这里，拍摄位置获取单元 303 还可以例如基于拍摄单元 301 对所处房间拍摄 3D 建模用图像时提供的定位信息和方向信息，来获取拍摄单元 301 对所处房间拍摄 3D 建模用图像时的位置和拍摄方向信息。

这里，拍摄位置获取单元 303 还通过拍摄单元 301 具备的加速度传感器、速度传感器提供的加速度信息和移动速度信息，对路线图进行校正。

3D 模型拼接单元 304，基于拍摄位置获取单元 303 获取的各个房间的位置和拍摄方向信息，将 3D 模型生成单元 302 生成的各个房间的 3D 模型在同一个三维坐标系内进行拼接处理，形成由各个房间各自的 3D 模型拼接而成的整体 3D 模型。

这里，3D 模型拼接单元 304 还可以根据拍摄位置获取单元 303 获取的拍摄各个房间时的位置和拍摄方向信息，例如可利用一个转换矩阵，把单个房间的 3D 模型的局部坐标转换为全局的世界坐标，从而获得所有拍摄点的整体 3D 模型。

这里，3D 模型拼接单元 304 还可以对多个房间各自的 3D 模型进行统一校正，包括使用统计的方法对所有的房间的墙线方向进行统一校正，其中，对于室内场景，大部分情况下各个房间的墙满足平行和垂直的关系，通过找出每个房间墙线方向的均值或中值，或用 RANSAC (Random Sample Consensus 随机抽样一致性) 等算法找出最可能的墙线方向，将误差在一定范围内的房间调整为统一的方向，使得所有房间在一定偏差范围内的墙线平行；

这里，3D 模型拼接单元 304 还可以在将各个房间的 3D 模型进行拼接处理时，对重叠的部分和/或出现的空洞进行修正，这里，修正的方式至少可以包括以下方式之一：

假设房间的位置准确，但轮廓识别有误差，则对重叠部分进行裁剪，并填补上空洞部分；

假设房间的轮廓识别较准，但位置有误差，则移动各个房间的位置，以尽量消除重叠和空洞。

当然，实际使用中也可以结合两种方法，反复迭代进行，以接近真实情况。

2D 平面图生成单元 305，以如下方式生成 2D 平面图：

1、将生成的 3D 模型的每个面向垂直于地板的平面投影，并将这些投影合并成为一个多边形；

2、对得到的多边形进行校正和简化，例如可以包括以下方式：

（1）仅保留多边形的主要顶点，删除小的凹凸，例如，小于墙的标准厚度（例如 12cm 或 24cm）的凹凸，可以删除；

（2）利用计算机视觉技术检测图片中的直线，进而确定墙壁的走向，并将与墙壁走向近似平行或垂直的边都修正到对应的方向上；

当然，对得到的多边形进行校正和简化的方式当然也可以是其他方式，并不做限定；

3、基于拍摄位置获取单元 303 获取的各个房间的位置和拍摄方向信息，将生成的各个房间的 2D 平面图在同一个二维坐标系内进行拼接处理，形成由各个房间各自的 2D 平面图拼接而成的整体 2D 平面图；

4、识别并标示出门和/或窗的位置，包括使用深度学习的方法，在室内图像上识别出门和/或窗的位置，或者根据拍摄单元 301 对属于同一套房子的多个房间分别拍摄 3D 建模用图像时的移动轨迹与该房间轮廓的交点，确定门的位置。

这里，在一个或多个实施例中，使用深度学习的方法在室内图像上识别出门和/或窗的位置的方法例如可以通过单阶段的关键点检测模型，通过三个关键点而不是一对关键点检测每个目标物例如门和/或窗，以提高准确率及返回率。

这里，2D 平面图生成单元 305 还可以对多个房间各自的 2D 平面图进行统一校正，包括使用统计的方法对所有的房间的墙线方向进行统一校正，使得所有房间在一定偏差范围内的墙线平行，这里，统一校正的方法可以与前述相同，不再赘述；

这里，2D 平面图生成单元 305 还可以在将各个房间的 2D 平面图进行拼接处理时，对重叠的部分和/或出现的空洞进行修正。

这里，2D 平面图生成单元还可以以如下方式生成 2D 平面图：

1、将 3D 模型拼接单元 304 生成的整体 3D 模型的每个面向垂直于地板的平面投影，并将这些投影合并成为一个多边形；

2、对得到的多边形进行校正和简化，例如可以包括以下方式：

(1) 仅保留多边形的主要顶点，删除小的凹凸，

(2) 利用计算机视觉技术检测图片中的直线，进而确定墙壁的走向，并将与墙壁走向近似平行或垂直的边都修正到对应的方向上；

当然，对得到的多边形进行校正和简化的方式当然也可以是其他方式，并不做限定；

3、识别并标示出门和/或窗的位置，包括使用深度学习的方法，在室内图像上识别出门和/或窗的位置，或者根据拍摄单元 301 对属于同一套房子的多个房间分别拍摄 3D 建模用图像时的移动轨迹与该房间轮廓的交点，确定门的位置。

这里，在一个或多个实施例中，使用深度学习的方法在室内图像上识别出门和/或窗的位置的方法例如可以通过单阶段的关键点检测模型，通过三个关键点而不是一对关键点检测每个目标物例如门和/或窗，以提高准确率及返回率。

[基于拍照的 3D 建模方法]

图 4 为本发明基于拍照的 3D 建模方法流程示意图。

请参见图 4，本发明提供的基于拍照的 3D 建模方法，包括如下步骤：

S1) 将具有拍照功能的移动设备（包括手机、平板电脑等）和/或相机（包括全景、鱼眼和普通数码相机）固定在同一拍摄支架（包括三脚架）上。

S2) 在移动支架的过程中获取多张相机或移动设备的定位用图像，通过图像处理算法，结合相机或移动设备的传感器来获取每个拍摄点的位置和拍摄方向，形成使用统一坐标系的路线图。

这里，步骤 S2 基于移动设备或相机的定位系统，使用移动设备或相机拍摄的定位用图像，通过相近拍摄点的定位用图像进行特征点匹配来找到拍摄点的相对位移和拍摄方向信息，形成包括所有拍摄点的在同一坐标系内的所述路线图，并提供每个拍摄点的位置和方向。

这里，步骤 S2 还包括通过移动设备或相机的传感器，获取包括移动的加速度、

速度和方向信息，对所述路线图进行校正。

这里，步骤 S2 还包括获取相机的镜头和移动设备的拍摄方向的夹角，在初始化阶段同时运行基于移动设备的定位系统和基于相机的定位系统，并移动支架一段距离，此时两套系统各提供一个位移向量，两个向量的夹角即为相机的镜头与移动设备的拍摄方向的夹角；或者通过手工将相机与移动设备调节成朝向一致的角度，比如旋转相机的预览图或者拍摄的图像，指定与移动设备拍摄方向一致的角度；或者通过图像识别算法匹配移动设备与相机的预览图或拍摄的图像，找到夹角；或者通过使用额外的标识、包括在支架上增加刻度与移动设备的安装方向形成固定夹角，然后在相机的预览图或图像中识别该标识来计算相机的镜头和移动设备拍摄方向的夹角。

S3) 在移动设备上或远程服务器端基于在每个拍摄点拍摄的 3D 建模用图像，通过深度学习算法或其他方法进行 3D 建模，获得每个拍摄点的 3D 模型和/或 2D 平面图。

这里，步骤 S3 包括：

S31) 基于深度学习技术，识别图像中的地板、天花板、墙壁的至少一者的图像；

S32) 基于图像处理技术，对识别出的图像进行分块，每块近似认为是一个平面，地板和天花板的图像块位于水平平面，墙壁的图像块位于垂直平面，求解各个平面对应的方程生成 3D 模型；对于在图像中相交的两个平面，以这两个平面的一条交线为约束条件，使得计算得到的交线与实际观察到的交线的误差最小。

这里，步骤 S3 还包括：对于室内图像，利用计算机视觉技术识别图像中的墙角，并将墙角连接起来即为房间的粗略模型。这里，在一个或多个实施例中，识别图像中的墙角的方法例如可以使用兴趣点检测与描述自监督训练框架（例如可参见 SuperPoint: Self-Supervised Interest Point Detection and Description），再将墙角连接起来形成房间的粗略模型，以实现能够捕捉在同一三维空间结构中频繁出现的对象例如墙角之间的几何关系。

S4) 将每个拍摄点的 3D 模型按照 S2 中获取的位置和拍摄方向，统一放置在一个三维坐标系内，对多个拍摄点 3D 模型之间的结合部位进行拼接处理，形成多个拍摄点的 3D 模型和/或 2D 平面图，对所有房间墙的方向进行统一校正并对重叠和空洞情况进行优化处理。正常房型中多数房间由平行的墙构成，而单个拍摄点形成的房间模型中本来平行的墙会有一定偏差（不平行）；通过考虑多个房间墙的方向，找

出一个主流方向，并以此调整所有房间墙的方向。

这里，步骤 S4 包括：

S41) 根据每个拍摄点的位置和拍摄方向，例如可利用一个转换矩阵，把单个拍摄点的 3D 模型的局部坐标转换为成全局的世界坐标，从而获得所有拍摄点的整体 3D 模型；

S42) 对多个拍摄点的 3D 模型进行统一校正，包括使用统计的方法对所有的拍摄点的墙线方向进行统一校正，使得所有房间在一定偏差范围内的墙线平行；

S43) 在将各个拍摄点的 3D 模型进行拼接处理时，对其重叠的部分和/或出现的空洞进行修正。

S5) 在移动设备上自动生成全景漫游效果。

下面，将结合基于拍照的 3D 建模系统说明本实施例的基于拍照的 3D 建模方法的应用。

一、硬件系统

本实施例中将手机和相机都固定在同一个支架（包括例如三脚架等）上。

二、系统初始化

本发明使用以下两种方法之一来获取每个拍摄点的拍摄位置和相机的拍摄方向：

方法一) 基于手机的定位系统，即使用手机的图像（照片、视频或预览图），通过相近拍摄点图像进行特征点匹配来找到拍摄点的位移，以及优选使用移动设备的传感器（包括陀螺仪 Gyroscope、加速度计 Accelerometer、指南针 Compass 等）进行校正，以此生成路线图，并提供拍摄点的位置和方向；

方法二) 基于相机的定位系统，即使用相机的图像（照片、视频或预览图），通过相近拍摄点图像进行特征点匹配来找到拍摄点的位移，优选是间隔厘米级和分米级并连续进行匹配和定位，以及优选使用相机的传感器（包括陀螺仪 Gyroscope、加速度计 Accelerometer、指南针 Compass 等）进行校正，以此生成路线图，并提供拍摄点的位置和方向。

两种方法对比：方法一基于手机系统，由于手机有多种传感器，一般可以提供较为精确的绝对坐标信息，可以测量拍摄点间的绝对距离，但使用前需要额外的初始化过程。

方法二由于相机往往不具备完善的传感器，只能提供拍摄位置的相对坐标，但不需要额外的初始化以对齐路径和单个拍摄点 3D 模型的坐标轴；此外，如果拍摄路

径中存在回环，方法二提供的坐标误差较小。

使用方法一时，手机提供的坐标是基于手机自身的坐标系（一般一个轴指向与地面垂直的方向，另两个轴分别指向前后和左右方向），而基于全景照片生成的 3D 模型的坐标系是基于相机的坐标系，两者的坐标轴不重合，为了解决这个问题，需要对系统进行手动或自动的初始化，可以采用手动或自动的方法：

手动：用户使用额外的测量工具或者在支架等设备上增加刻度，手工输入相机镜头和手机拍摄方向的夹角；

自动：初始化阶段同时运行方法一和方法二，并移动设备一段距离，优选移动 1-3 米，此时两套系统可以各提供一个系统的位移向量，两个向量的夹角即为相机镜头与手机的拍摄方向的夹角。

三、拍摄点位置和拍摄方向的确定

上述系统开始运行后，即可提供拍摄者所在的位置和拍摄方向信息。

四、单个拍摄点 3D 模型生成

传统上，基于拍照的建模方法有以下两种，但是都有明显的缺点：

传统方法 a) 使用能够识别深度信息的相机，直接生成 3D 模型。这种方法依赖于更加复杂的硬件，导致更高的设备成本，往往需要专业摄影师才能操作，不利于普及；

传统方法 b) 在一个拍摄点距离较近的地方拍摄两张照片，优选间隔厘米级和分米级并连续进行匹配和定位，然后使用 MVS (Multi View Stereo, 比如可参考 <https://github.com/cdcseacave/openMVS>) 进行建模，优点是整个过程全自动不需要人工干预，但是缺点也很明显：

缺点 1: 计算量较大，无法在移动设备上实时建模，通常上传到服务器(云端/PC)，利用更强的计算能力，运行建模算法；

缺点 2: 无法给出具体间隔多远拍摄的规范，如果太密则使用太麻烦，耗时过长；如果只是靠两两可见，靠感觉，可能建模失败，而拍摄时无法给用户提示。

为了克服以上的缺点，本发明使用了创新的方法：为了提高实时性，达到所见即所得的效果，通常只对房间轮廓（墙体位置）建模，而不去复原家具、饰品等房间附属物体的模型。即：

i. 基于深度学习技术，识别图像中的地板，天花板，墙壁和屋顶等部分，这些部分所在的平面要么法线方向确定（地板，天花板），要么法线在水平面上（墙壁）；

ii. 基于图像处理技术，对图像进行分块，每块可近似认为是一个平面。对于地板部分的块，平面方程是已知的。假设 y 轴垂直向上，则地板部分的方程为 $y+1 = 0$ 。对于墙壁部分，平面方程为 $Ax+Cz+D=0$ ，天花板部分为 $y+D = 0$ ，其他部分为 $Ax+By+Cz+D = 0$ ，生成 3D 模型的过程即为求解各个平面方程的过程。对于在图片中相交的两个平面，在图片中有一条交线，以此为约束条件，上述求解方程的过程可以变为一个最小化问题，使得对于相交的两个平面，计算得到的交线与实际观察到的交线的误差最小；

iii. 也可使用其他方法对场景进行建模。例如在室内，可以利用计算机视觉技术结合深度学习来识别图中的墙角，将墙角连接起来即为房间的粗略模型。这里，在一个或多个实施例中，识别图像中的墙角的方法例如可以使用兴趣点检测与描述自监督训练框架（例如可参见 SuperPoint: Self-Supervised Interest Point Detection and Description），再将墙角连接起来形成房间的粗略模型，以实现能够捕捉在同一三维空间结构中频繁出现的对象例如墙角之间的几何关系。

iv. 2D 平面图生成，得到每个拍摄点的 3D 模型后，可以进一步生成平面图，尤其是对室内场景的应用，很多时候都需要平面图。方法是：

1. 将 3D 模型的每个面投影到 2D 俯视图平面上；
2. 将这些投影进行合并成一个大的多边形；
3. 对得到的多边形进行校正和简化，例如可以包括以下方式：

a) 得到的多边形往往点的数量较多，可以对多边形进行简化，仅保留 2D 平面图上多边形的顶点，删除小的凹凸；

b) 对于室内场景，可以利用计算机视觉技术检测图片中的直线，进而确定墙壁的走向，将与墙壁走向近似平行或垂直的边都修正到对应的方向上。

4. 门和/或窗的识别。对于室内场景，需要在 2D 平面图上对门和/或窗进行标注，可以使用以下两个方法：

- a) 直接使用深度学习的方法，在全景图上识别门和/或窗的位置和大小；

这里，在一个或多个实施例中，使用深度学习的方法在室内图像上识别出门和/或窗的位置和大小的方法例如可以通过单阶段的关键点检测模型，通过三个关键点而不是一对关键点检测每个目标物例如门和/或窗的位置和大小，以提高准确率及返回率。

- b) 由于基于手机或相机的定位系统不仅给出知道每个拍摄点的位置和拍摄方向，

还具备整个拍摄过程中相机的移动轨迹，这个路径和房间本身的轮廓的交点必然是门的位置。

五、多个拍摄点 3D 模型以及 2D 平面图生成

a) 步骤 4 解决了每个拍摄点的 3D 模型生成，得到的 3D 模型坐标都是相对于拍摄点的相对坐标。为了将这些模型结合起来，从而生成完整的 3D 模型和 2D 平面图。首先，由于已经知道每个拍摄点的位置和拍摄方向，例如可以利用一个转换矩阵，把单个模型的局部坐标转换成全局的世界坐标。

b) 在此基础上，可以进一步对模型和平面图进行校正。

i. 单各拍摄点的模型使用了直线进行校正，一般是存在误差的，拍摄了多个点位后，可以用统计的方法对所有的拍摄点进行统一校正，例如，使用 RANSEC (Random Sample Consensus) 等方法，找到最合理的校正直线，使得所有房间在一定偏差范围内的墙线平行、避免出现小的偏差角度；

ii. 由于建模的误差，多个拍摄点的 3D 模型和 2D 平面图放在一起可能存在重叠，空洞等情况，可以自动去除重叠的部分，在 2D 平面图上对空洞进行填补等。

六、即时展示

上述过程可以完全在手机上自动进行，完成后，可以使用展示软件在手机上即时进行展示，漫游等，并能上传至云端分享给他人。

七、人工编辑

由于定位系统、单拍摄点 3D 建模算法以及多拍摄点 3D 模型/2D 平面图优化各环节均可能存在误差，为了得到更高精度的模型，本发明允许用户手工对拍摄的结果进行编辑，手工编辑可以借助基于软件的审核和编辑工具。

[自动 3D 建模装置]

图 5 为本发明的自动 3D 建模装置的一个实施例的结构示意图；如图 5 所示，包括：

3D 模型生成单元 501，基于建模对象所包括的多个空间各自的 3D 建模用图像，分别生成各个空间的 3D 模型；

3D 模型拼接单元 502，基于多个空间各自的 3D 建模用图像被拍摄时的位置和拍摄方向信息，将 3D 模型生成单元 501 生成的各个空间的 3D 模型在同一个三维坐标系内进行拼接处理，形成由各个空间各自的 3D 模型拼接而成的整体 3D 模型。

这里，3D 模型拼接单元 502 还可以根据位置和拍摄方向信息，例如可利用一个

转换矩阵，把单个空间的所述 3D 模型的局部坐标转换为全局的世界坐标，从而获得所有空间的整体 3D 模型。

图 6 为本发明的自动 3D 建模装置的另一个实施例的结构示意图；在该实施例中，自动 3D 建模空间例如是房间；3D 建模用图像例如是所述房间的室内图像。

如图 6 所示，该实施例包括 3D 模型生成单元 601，基于建模对象所包括的多个房间各自的 3D 建模用图像，分别生成各个房间的 3D 模型；

这里，3D 模型生成单元 601 基于深度学习技术，识别 3D 建模用图像中的地板、天花板、墙壁中的至少一者的图像，并基于图像处理技术，对识别出的图像进行分块，每块近似认为是一个平面，使地板和天花板的图像块位于水平平面，使墙壁的图像块位于垂直平面，通过求解各个平面方程生成所述 3D 模型；其中，对于在 3D 建模用图像中相交的两个平面，使得计算得到的交线与实际观察到的交线的误差最小；

这里，3D 模型生成单元 601 还利用计算机视觉技术识别室内图像中的墙角，并将墙角连接起来，生成房间的粗略模型。这里，在一个或多个实施例中，识别图像中的墙角的方法例如可以使用兴趣点检测与描述自监督训练框架（例如可参见 SuperPoint: Self-Supervised Interest Point Detection and Description），再将墙角连接起来形成房间的粗略模型，以实现能够捕捉在同一三维空间结构中频繁出现的对象例如墙角之间的几何关系。

3D 模型拼接单元 602，基于多个房间各自的 3D 建模用图像被拍摄时的位置和拍摄方向信息，将 3D 模型生成单元 601 生成的各个房间的 3D 模型在同一个三维坐标系内进行拼接处理，形成由各个房间各自的 3D 模型拼接而成的整体 3D 模型。

这里，3D 模型拼接单元 602 还可以根据位置和拍摄方向信息，例如可利用一个转换矩阵，把单个房间的 3D 模型的局部坐标转换为全局的世界坐标，从而获得所有房间的整体 3D 模型。

这里，3D 模型拼接单元 602 还可以对多个房间各自的 3D 模型进行统一校正，包括使用统计的方法对所有的房间的墙线方向进行统一校正，使得所有房间在一定偏差范围内的墙线平行；

这里，3D 模型拼接单元 602 还可以在将各个房间的 3D 模型进行拼接处理时，对重叠的部分和/或出现的空洞进行修正。

2D 平面图生成单元 603，以如下方式生成 2D 平面图：

1、将生成的 3D 模型的每个面向垂直于地板的平面投影，并将这些投影合并成

为一个多边形；

2、对得到的多边形进行校正和简化，例如可以包括以下方式：

(1) 仅保留所述多边形的主要顶点，删除小的凹凸，

(2) 利用计算机视觉技术检测图片中的直线，进而确定墙壁的走向，并将与墙壁走向近似平行或垂直的边都修正到对应的方向上；

当然，对得到的多边形进行校正和简化的方式当然也可以是其他方式，并不做限定；

3、基于位置和拍摄方向信息，将生成的各个房间的 2D 平面图在同一个二维坐标系内进行拼接处理，形成由各个房间各自的 2D 平面图拼接而成的整体 2D 平面图；

4、识别并标示出门和/或窗的位置，包括使用深度学习的方法，在室内图像上识别出门和/或窗的位置，或者根据对属于同一套房子的多个房间分别拍摄 3D 建模用图像时获取的移动轨迹与该房间轮廓的交点，确定门的位置。

这里，在一个或多个实施例，使用深度学习的方法在室内图像上识别出门和/或窗的位置和大小的方法例如可以通过单阶段的关键点检测模型，通过三个关键点而不是一对关键点检测每个目标物例如门和/或窗的位置和大小，以提高准确率及返回率。

这里，2D 平面图生成单元 603 还可以对多个房间各自的 2D 平面图进行统一校正，包括使用统计的方法对所有的房间的墙线方向进行统一校正，使得所有房间在一定偏差范围内的墙线平行；

这里，2D 平面图生成单元 603 还可以在将各个房间的 2D 平面图进行拼接处理时，对重叠的部分和/或出现的空洞进行修正。

这里，2D 平面图生成单元 603 还可以以如下方式生成 2D 平面图：

1、将 3D 模型拼接单元 602 生成的整体 3D 模型的每个面向垂直于地板的平面投影，并将这些投影合并成为一个多边形；

2、对得到的所述多边形进行校正和简化，例如可以包括以下方式：

(1) 仅保留所述多边形的主要顶点，删除小的凹凸，

(2) 利用计算机视觉技术检测图片中的直线，进而确定墙壁的走向，并将与墙壁走向近似平行或垂直的边都修正到对应的方向上；

当然，对得到的多边形进行校正和简化的方式当然也可以是其他方式，并不做限定；

3、识别并标示出门和/或窗的位置，包括使用深度学习的方法，在室内图像上识别出门和/或窗的位置，或者根据对属于同一套房子的多个房间分别拍摄所述 3D 建模用图像时的移动轨迹与该房间轮廓的交点，确定门的位置。具体方法例如通过上述的单阶段的关键点检测模型，这里不再赘述。

[自动 3D 建模方法]

图 7 为本发明的自动 3D 建模方法的一个实施例的流程示意图；如图 7 所示，包括：

3D 模型生成步骤 S71，基于建模对象所包括的多个空间各自的 3D 建模用图像，分别生成各个空间的 3D 模型；

在本实施例中，自动 3D 建模空间例如是房间；3D 建模用图像例如是所述房间的室内图像。

在 3D 模型生成步骤 S71 中，基于深度学习技术，识别 3D 建模用图像中的地板、天花板、墙壁中的至少一者的图像，并基于图像处理技术，对识别出的图像进行分块，每块近似认为是一个平面，使地板和天花板的图像块位于水平平面，使墙壁的图像块位于垂直平面，通过求解各个平面方程生成 3D 模型；其中，对于在 3D 建模用图像中相交的两个平面，使得计算得到的交线与实际观察到的交线的误差最小；

在 3D 模型生成步骤 S71 中还利用计算机视觉技术识别室内图像中的墙角，并将墙角连接起来，生成房间的粗略模型。这里，在一个或多个实施例中，识别图像中的墙角的方法例如可以使用兴趣点检测与描述自监督训练框架（例如可参见 SuperPoint: Self-Supervised Interest Point Detection and Description），再将墙角连接起来形成房间的粗略模型，以实现能够捕捉在同一三维空间结构中频繁出现的对象例如墙角之间的几何关系。

3D 模型拼接步骤 S72，基于多个房间各自的 3D 建模用图像被拍摄时的位置和拍摄方向信息，将 3D 模型生成步骤 S71 中生成的各个房间的 3D 模型在同一个三维坐标系内进行拼接处理，形成由各个房间各自的 3D 模型拼接而成的整体 3D 模型。

这里，在 3D 模型拼接步骤 S72 中，还可以根据位置和拍摄方向信息，例如可利用一个转换矩阵，把单个空间的 3D 模型的局部坐标转换为全局的世界坐标，从而获得所有空间的整体 3D 模型。

2D 平面图生成步骤 S73，以如下方式生成 2D 平面图：

1、将生成的 3D 模型的每个面向垂直于地板的平面投影，并将这些投影合并成

为一个多边形；

2、对得到的所述多边形进行校正和简化，例如可以包括以下方式：

(1) 仅保留所述多边形的主要顶点，删除小的凹凸，

(2) 利用计算机视觉技术检测图片中的直线，进而确定墙壁的走向，并将与墙壁走向近似平行或垂直的边都修正到对应的方向上；

这里，对得到的多边形进行校正和简化的方式当然也可以是其他方式，并不做限定；

3、基于位置和拍摄方向信息，将生成的各个房间的 2D 平面图在同一个二维坐标系内进行拼接处理，形成由各个房间各自的 2D 平面图拼接而成的整体 2D 平面图；

4、识别并标示出门和/或窗的位置，包括使用深度学习的方法，在室内图像上识别出门和/或窗的位置，或者根据对属于同一套房子的多个房间分别拍摄 3D 建模用图像时获取的移动轨迹与该房间轮廓的交点，确定门的位置。具体方法例如通过上述的单阶段的关键点检测模型，这里不再赘述。

这里，2D 平面图生成步骤 S73 还可以对多个房间各自的 2D 平面图进行统一校正，包括使用统计的方法对所有的房间的墙线方向进行统一校正，使得所有房间在一定偏差范围内的墙线平行；

这里，2D 平面图生成步骤 S73 还可以在将各个房间的 2D 平面图进行拼接处理时，对重叠的部分和/或出现的空洞进行修正。

这里，2D 平面图生成步骤 S73 还可以以如下方式生成 2D 平面图：

1、将 3D 模型拼接步骤 S72 生成的整体 3D 模型的每个面向垂直于地板的平面投影，并将这些投影合并成为一个多边形；

2、对得到的所述多边形进行校正和简化，例如可以包括以下方式：

(1) 仅保留所述多边形的主要顶点，删除小的凹凸，

(2) 利用计算机视觉技术检测图片中的直线，进而确定墙壁的走向，并将与墙壁走向近似平行或垂直的边都修正到对应的方向上；

当然，对得到的多边形进行校正和简化的方式当然也可以是其他方式，并不做限定；

3、识别并标示出门和/或窗的位置，包括使用深度学习的方法，在室内图像上识别出门和/或窗的位置，或者根据对属于同一套房子的多个房间分别拍摄 3D 建模用图像时的移动轨迹与该房间轮廓的交点，确定门的位置。具体方法例如通过上述的

单阶段的关键点检测模型，这里不再赘述。

[电子设备]

下面参考图 8，其示出了适于用来实现本发明实施例的电子设备的（例如图 1 中的移动设备或服务器）800 的结构示意图。本发明实施例中的电子设备可以是上述系统中的各种移动设备，并且优选具有拍照功能，以独立或通过与其他电子终端设备例如相机固定在同一个支架（例如三脚架）上，配合运行各类移动操作系统中的应用软件实现本发明的实施例方法。图 8 示出的电子设备仅仅是一个示例，不应对本发明实施例的功能和使用范围带来任何限制。

如图 8 所示，电子设备 800 可以包括处理装置（例如中央处理器、图形处理器等）801，用于控制电子设备的整体操作。处理装置可以包括一个或多个处理器来执行指令，以完成上述的方法的全部或部分步骤。此外，处理装置 801 还可以包括一个或多个模块，用于处理和其他装置或单元之间的交互。

存储装置 802 用于存储各种类型的数据，存储装置 802 可以是包括各种类型的计算机可读存储介质或者它们的组合，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。在本发明中，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。

传感器装置 803，用于感受规定的被测量的信息并按照一定的规律转换成可用输出信号，可以包括一个或多个传感器。例如，其可以包括加速度传感器，陀螺仪传感器，磁传感器，压力传感器或温度传感器等，用于检测电子设备的打开/关闭状态、相对定位、加速/减速、温度、湿度和光线等的变化。

处理装置 801、存储装置 802 以及传感器装置 803 通过总线 804 彼此相连。输入/输出（I/O）接口 805 也连接至总线 804。

多媒体装置 806 可以包括触摸屏、触摸板、键盘、鼠标、摄像头、麦克风等的输入装置用以接收来自用户的输入信号，在各种输入装置可以与上述传感器装置 803 的各种传感器配合完成例如手势操作输入、图像识别输入、距离检测输入等；多媒体装置 806 还可以包括例如液晶显示器（LCD）、扬声器、振动器等的输出装置。

电源装置 807, 用于为电子设备中的各种装置提供电力, 可以包括电源管理系统、一个或多个电源及为其他装置分配电力的组件。

通信装置 808, 可以允许电子设备 800 与其他设备进行无线或有线通信以交换数据。

上述各项装置也均可以连接至 I/O 接口 805 以实现电子设备 800 的应用。

虽然图 8 示出了具有各种装置的电子设备 800, 但是应理解的是, 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。

特别地, 根据本发明的实施例, 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, 本发明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从网络上被下载和安装, 或者从存储装置被安装。在该计算机程序被处理装置执行时, 执行本发明实施例的方法中限定的上述功能。

在本发明的上下文中, 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。

要说明的是, 本发明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。而在本发明中, 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。

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。

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发明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，图示了按照本发明各种实施例的系统、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。也要注意的，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。

描述于本发明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。

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。

以上描述仅为本发明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，本发明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。例如上述特征与本发明中公开的（但不限于）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。

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发明的范围的限制。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。

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 书

1、一种基于拍照的 3D 建模系统，其特征在于，包括：

拍摄单元，用于对多个空间分别拍摄该空间的第一图像；

3D 模型生成单元，分别基于所述拍摄单元针对各个所述空间拍摄的所述第一图像，生成各个所述空间的 3D 模型；

拍摄位置获取单元，用于获取所述拍摄单元拍摄各个空间的第一图像时的位置和拍摄方向信息；

3D 模型拼接单元，基于所述位置和拍摄方向信息，将所述各个空间的 3D 模型在同一个三维坐标系内进行拼接处理，形成包括所述各个空间的整体 3D 模型。

2、如权利要求 1 所述的基于拍照的 3D 建模系统，其特征在于，

所述拍摄单元在所述各个空间之间移动的过程中拍摄多张第二图像，

所述拍摄位置获取单元通过所述多张第二图像进行特征点匹配，来获得各拍摄点的相对位移和拍摄方向信息。

3、如权利要求 1 所述的基于拍照的 3D 建模系统，其特征在于，

所述拍摄单元具有定位传感器和方向传感器；

所述拍摄位置获取单元基于所述拍摄单元对所处空间拍摄第一图像时提供的定位信息和方向信息，来获取所述拍摄单元对所处空间拍摄第一图像时的位置和拍摄方向信息。

4、如权利要求 1 所述的基于拍照的 3D 建模系统，其特征在于，

所述拍摄单元在所述各个空间之间移动的过程中拍摄多张第二图像；

所述拍摄单元具有定位传感器和方向传感器；

所述拍摄位置获取单元通过所述多张第二图像进行特征点匹配，来获得各拍摄点的相对位移和拍摄方向信息，并通过所述拍摄单元对所处空间拍摄第一图像时的定位信息和方向信息对所述相对位移和拍摄方向信息进行校正，从而获取所述拍摄单元对所处空间拍摄第一图像时的位置和拍摄方向信息。

5、如权利要求 4 所述的基于拍照的 3D 建模系统，其特征在于，

所述拍摄位置获取单元基于位移传感器获取的位移信息，对所述相对位移和拍摄方向信息进行校正。

6、如权利要求 1 所述的基于拍照的 3D 建模系统，其特征在于，

所述 3D 模型拼接单元根据所述拍摄位置获取单元获取的拍摄各个空间时的所述位置和拍摄方向信息,把单个空间的所述 3D 模型的局部坐标转换为全局的世界坐标,从而获得所述多个空间的整体 3D 模型。

7、如权利要求 6 所述的基于拍照的 3D 建模系统,其特征在于,
所述把单个空间的所述 3D 模型的局部坐标转换为全局的世界坐标,包括,
使所述拍摄单元移动预定距离,
所述拍摄位置获取单元获取所述预定距离的两个端点的位置信息,
将所述两个端点的位置的差和所述预定距离相比获取所述局部坐标和所述世界坐标的比值。

8、如权利要求 6 所述的基于拍照的 3D 建模系统,其特征在于,
所述把单个空间的所述 3D 模型的局部坐标转换为全局的世界坐标,包括,
所述拍摄位置获取单元在所述第一图像上识别特征点,
基于识别到的所述特征点,预估所述空间的底面或顶面所在的平面的高度,
将所述空间的底面或顶面所在的平面的高度与所述拍摄单元的高度相比获取所述局部坐标和所述世界坐标的比值。

9、如权利要求 8 所述的基于拍照的 3D 建模系统,其特征在于,
在所述拍摄单元进行第一个拍摄点的拍摄之前、或在后续拍摄的移动过程中,移动预定距离以获得预定数量的所述特征点。

10、如权利要求 1 所述的基于拍照的 3D 建模系统,其特征在于,
所述拍摄单元具有双目镜头,在同一个拍摄点,该双目镜头分别拍摄所述第一图像;

所述 3D 模型生成单元通过对所述双目镜头各自拍摄的所述第一图像进行图像比对,确定对应像素,并获得每个所述对应像素的深度信息,以用于生成所述 3D 模型。

11、如权利要求 1 所述的基于拍照的 3D 建模系统,其特征在于,
所述 3D 模型生成单元通过深度学习技术,预测所述第一图像中每个像素的深度,计算或直接运用所述深度学习技术预测所述每个像素的法线方向。

12、如权利要求 1 至 11 的任一项所述的基于拍照的 3D 建模系统,其特征在于,
所述拍摄单元由相机和/或带拍照功能的移动设备实现;

所述 3D 模型生成单元由所述移动设备实现或者由远程服务器实现,在由所述远程服务器实现的情况下,其通过网络接收所述相机和/或所述带拍照功能的移动设备

拍摄并发送来的第一图像，生成各个所述空间的 3D 模型；

所述拍摄位置获取单元由所述相机或所述移动设备实现；

所述 3D 模型拼接单元，由所述移动设备实现或者由远程服务器实现，在由所述远程服务器实现的情况下，其通过网络接收所述拍摄位置获取单元发送来的各个空间的所述位置和拍摄方向信息，基于该位置和拍摄方向信息完成所述拼接处理，并将生成的所述整体 3D 模型发送给所述移动设备或其它设备。

13、如权利要求 12 所述的基于拍照的 3D 建模系统，其特征在于，

实现所述拍摄单元的相机和带拍照功能的移动设备固定在同一拍摄支架上；

在移动所述支架的过程中，获取多张所述相机或所述带拍照功能的移动设备所拍摄的第二图像，从而获取所述相机或所述带拍照功能的移动设备对所处空间拍摄第一图像时的位置和拍摄方向信息。

14、如权利要求 13 所述的基于拍照的 3D 建模系统，其特征在于，

基于所述相机或所述带拍照功能的移动设备的定位系统，使用所述相机或所述带拍照功能的移动设备拍摄的第二图像，通过相近拍摄点的第二图像进行特征点匹配来获得各拍摄点的相对位移和拍摄方向信息，从而提供每个拍摄点的相对位置和方向。

15、如权利要求 13 所述的基于拍照的 3D 建模系统，其特征在于，还包括，

在所述拍摄单元对第一个所述空间拍摄所述第一图像之前、或在后续拍摄的移动过程中，通过以下方式的一种或多种，获取所述相机的镜头和所述移动设备的拍摄方向的夹角：

(1) 同时运行基于所述移动设备的定位系统和基于所述相机的定位系统，并移动所述支架一段距离，此时两套系统各提供一个位移向量，两个向量的夹角即为所述相机的镜头与所述移动设备的拍摄方向的夹角；

(2) 通过手工旋转所述相机的预览图或者拍摄的图像，指定与所述移动设备拍摄方向一致的角度；

(3) 通过图像识别算法匹配所述移动设备与所述相机的预览图或拍摄的图像，找到所述夹角；

(4) 通过使用额外的标识、包括在支架上增加刻度与所述移动设备的安装方向形成固定夹角，然后在所述相机的预览图或图像中识别该标识来计算所述相机的镜头和所述移动设备拍摄方向的夹角；

(5) 通过在所述支架上设置卡槽，保证所述相机与所述移动设备形成已知的固定夹角。

16、如权利要求 1 所述的基于拍照的 3D 建模系统，其特征在于，

所述空间是房间；

所述第一图像是所述房间的室内图像；

所述 3D 模型生成单元基于深度学习技术，识别所述第一图像中的地板、天花板、墙壁中的至少一者的图像，并基于图像处理技术，对识别出的图像进行分块，每块近似认为是一个平面，使地板和天花板的图像块位于水平平面，使墙壁的图像块位于垂直平面，通过求解各个平面方程生成所述 3D 模型；其中，对于在所述第一图像中相交的两个平面，使得计算得到的交线与实际观察到的交线的误差最小。

17、如权利要求 16 所述的基于拍照的 3D 建模系统，其特征在于，

所述 3D 模型生成单元还利用计算机视觉技术识别所述室内图像中的墙角，并将墙角连接起来，生成房间的粗略模型。

18、如权利要求 16 所述的基于拍照的 3D 建模系统，其特征在于，

所述 3D 模型拼接单元对所述多个房间各自的 3D 模型进行统一校正，包括使用统计的方法对所有的房间的墙线方向进行统一校正，使得所有房间在一定偏差范围内的墙线平行；

所述 3D 模型拼接单元在将所述各个房间的 3D 模型进行拼接处理时，对重叠的部分和/或出现的空洞进行修正。

19、如权利要求 16 所述的基于拍照的 3D 建模系统，其特征在于，还包括：

2D 平面图生成单元，以如下方式生成 2D 平面图：

将生成的所述 3D 模型的每个面向垂直于地板的平面投影，并将这些投影合并成为一个多边形；

对得到的所述多边形进行校正和简化，包括以下至少一种方式：

(1) 仅保留所述多边形的主要顶点，删除小的凹凸，

(2) 利用计算机视觉技术检测图片中的直线，进而确定墙壁的走向，并将与墙壁走向近似平行或垂直的边都修正到对应的方向上；

基于所述拍摄位置获取单元获取的各个空间的所述位置和拍摄方向信息，将生成的所述各个房间的 2D 平面图在同一个二维坐标系内进行拼接处理，形成由所述各

个房间各自的 2D 平面图拼接而成的整体 2D 平面图；

识别并标示出门和/或窗的位置，包括使用深度学习的方法，在所述室内图像上识别出门和/或窗的位置，或者根据所述拍摄单元对属于同一套房子的多个房间分别拍摄所述第一图像时的移动轨迹与该房间轮廓的交点，确定门的位置。

20、如权利要求 19 所述的基于拍照的 3D 建模系统，其特征在于，

所述 2D 平面图生成单元对所述多个房间各自的 2D 平面图进行统一校正，包括使用统计的方法对所有的房间的墙线方向进行统一校正，使得所有房间在一定偏差范围内的墙线平行；

所述 2D 平面图生成单元在将所述各个房间的 2D 平面图进行拼接处理时，对重叠的部分和/或出现的空洞进行修正。

21、如权利要求 16 所述的基于拍照的 3D 建模系统，其特征在于，还包括：

2D 平面图生成单元，以如下方式生成 2D 平面图：

将所述 3D 模型拼接单元生成的所述整体 3D 模型的每个面向垂直于地板的平面投影，并将这些投影合并成为一个多边形；

对得到的所述多边形进行校正和简化，包括以下至少一种方式：

(1) 仅保留所述多边形的主要顶点，删除小的凹凸，

(2) 利用计算机视觉技术检测图片中的直线，进而确定墙壁的走向，并将与墙壁走向近似平行或垂直的边都修正到对应的方向上；

识别并标示出门和/或窗的位置，包括使用深度学习的方法，在所述室内图像上识别出门和/或窗的位置，或者根据所述拍摄单元对属于同一套房子的多个房间分别拍摄所述第一图像时的移动轨迹与该房间轮廓的交点，确定门的位置。

22、一种自动 3D 建模装置，其特征在于，包括：

3D 模型生成单元，基于建模对象所包括的多个空间各自的第一图像，分别生成各个所述空间的 3D 模型；

3D 模型拼接单元，基于所述多个空间各自的所述第一图像被拍摄时的位置和拍摄方向信息，将所述 3D 模型生成单元生成的所述各个空间的 3D 模型在同一个三维坐标系内进行拼接处理，形成由所述各个空间各自的 3D 模型拼接而成的整体 3D 模型。

23、如权利要求 22 所述的自动 3D 建模装置，其特征在于，

所述 3D 模型拼接单元根据所述位置和拍摄方向信息，把单个空间的所述 3D 模

型的局部坐标转换为全局的世界坐标，从而获得所有空间的整体 3D 模型。

24、如权利要求 22 或 23 所述的自动 3D 建模装置，其特征在于，

所述空间是房间；

所述第一图像是所述房间的室内图像；

所述 3D 模型生成单元基于深度学习技术，识别所述第一图像中的地板、天花板、墙壁中的至少一者的图像，并基于图像处理技术，对识别出的图像进行分块，每块近似认为是一个平面，使地板和天花板的图像块位于水平平面，使墙壁的图像块位于垂直平面，通过求解各个平面方程生成所述 3D 模型；其中，对于在所述第一图像中相交的两个平面，使得计算得到的交线与实际观察到的交线的误差最小；

所述 3D 模型生成单元还利用计算机视觉技术识别所述室内图像中的墙角，并将墙角连接起来，生成房间的粗略模型。

25、如权利要求 24 所述的自动 3D 建模装置，其特征在于，

所述 3D 模型拼接单元对所述多个房间各自的 3D 模型进行统一校正，包括使用统计的方法对所有的房间的墙线方向进行统一校正，使得所有房间在一定偏差范围内的墙线平行；

所述 3D 模型拼接单元在将所述各个房间的 3D 模型进行拼接处理时，对重叠的部分和/或出现的空洞进行修正。

26、如权利要求 24 所述的自动 3D 建模装置，其特征在于，还包括：

2D 平面图生成单元，以如下方式生成 2D 平面图：

将生成的所述 3D 模型的每个面向垂直于地板的平面投影，并将这些投影合并成为一个多边形；

对得到的所述多边形进行校正和简化，包括以下至少一种方式：

(1) 仅保留所述多边形的主要顶点，删除小的凹凸，

(2) 利用计算机视觉技术检测图片中的直线，进而确定墙壁的走向，并将与墙壁走向近似平行或垂直的边都修正到对应的方向上；

基于所述位置和拍摄方向信息，将生成的所述各个房间的 2D 平面图在同一个二维坐标系内进行拼接处理，形成由所述各个房间各自的 2D 平面图拼接而成的整体 2D 平面图；

识别并标示出门和/或窗的位置，包括使用深度学习的方法，在所述室内图像上识别出门和/或窗的位置，或者根据对属于同一套房子的多个房间分别拍摄所述第一

图像时获取的移动轨迹与该房间轮廓的交点，确定门的位置。

27、如权利要求 26 所述的自动 3D 建模装置，其特征在于，

所述 2D 平面图生成单元对所述多个房间各自的 2D 平面图进行统一校正，包括使用统计的方法对所有的房间的墙线方向进行统一校正，使得所有房间在一定偏差范围内的墙线平行；

所述 2D 平面图生成单元在将所述各个房间的 2D 平面图进行拼接处理时，对重叠的部分和/或出现的空洞进行修正。

28、如权利要求 24 所述的自动 3D 建模装置，其特征在于，还包括：

2D 平面图生成单元，以如下方式生成 2D 平面图：

将所述 3D 模型拼接单元生成的所述整体 3D 模型的每个面向垂直于地板的平面投影，并将这些投影合并成为一个多边形；

对得到的所述多边形进行校正和简化，包括以下至少一种方式：

(1) 仅保留所述多边形的主要顶点，删除小的凹凸，

(2) 利用计算机视觉技术检测图片中的直线，进而确定墙壁的走向，并将与墙壁走向近似平行或垂直的边都修正到对应的方向上；

识别并标示出门和/或窗的位置，包括使用深度学习的方法，在所述室内图像上识别出门和/或窗的位置，或者根据对属于同一套房子的多个房间分别拍摄所述第一图像时的移动轨迹与该房间轮廓的交点，确定门的位置。

29、一种自动 3D 建模方法，其特征在于，包括：

3D 模型生成步骤，基于建模对象所包括的多个空间各自的第一图像，分别生成各个所述空间的 3D 模型；

3D 模型拼接步骤，基于所述多个空间各自的所述第一图像被拍摄时的位置和拍摄方向信息，将所述 3D 模型生成步骤中生成的所述各个空间的 3D 模型在同一个三维坐标系内进行拼接处理，形成由所述各个空间各自的 3D 模型拼接而成的整体 3D 模型。

30、如权利要求 29 所述的自动 3D 建模方法，其特征在于，

在所述 3D 模型拼接步骤中，根据所述位置和拍摄方向信息，把单个空间的所述 3D 模型的局部坐标转换为全局的世界坐标，从而获得所有空间的整体 3D 模型。

31、如权利要求 29 或 30 所述的自动 3D 建模方法，其特征在于，

所述空间是房间；

所述第一图像是所述房间的室内图像；

在所述 3D 模型生成步骤中，基于深度学习技术，识别所述第一图像中的地板、天花板、墙壁中的至少一者的图像，并基于图像处理技术，对识别出的图像进行分块，每块近似认为是一个平面，使地板和天花板的图像块位于水平平面，使墙壁的图像块位于垂直平面，通过求解各个平面方程生成所述 3D 模型；其中，对于在所述第一图像中相交的两个平面，使得计算得到的交线与实际观察到的交线的误差最小；

在所述 3D 模型生成步骤中还利用计算机视觉技术识别所述室内图像中的墙角，并将墙角连接起来，生成房间的粗略模型。

32、如权利要求 31 所述的自动 3D 建模方法，其特征在于，还包括：

2D 平面图生成步骤，以如下方式生成 2D 平面图：

将生成的所述 3D 模型的每个面向垂直于地板的平面投影，并将这些投影合并成为一个多边形；

对得到的所述多边形进行校正和简化，包括以下至少一种方式：

(1) 仅保留所述多边形的主要顶点，删除小的凹凸，

(2) 利用计算机视觉技术检测图片中的直线，进而确定墙壁的走向，并将与墙壁走向近似平行或垂直的边都修正到对应的方向上；

基于所述位置和拍摄方向信息，将生成的所述各个房间的 2D 平面图在同一个二维坐标系内进行拼接处理，形成由所述各个房间各自的 2D 平面图拼接而成的整体 2D 平面图；

识别并标示出门和/或窗的位置，包括使用深度学习的方法，在所述室内图像上识别出门和/或窗的位置，或者根据对属于同一套房子的多个房间分别拍摄所述第一图像时获取的移动轨迹与该房间轮廓的交点，确定门的位置。

33、如权利要求 32 所述的自动 3D 建模方法，其特征在于，

所述 2D 平面图生成步骤对所述多个房间各自的 2D 平面图进行统一校正，包括使用统计的方法对所有的房间的墙线方向进行统一校正，使得所有房间在一定偏差范围内的墙线平行；

所述 2D 平面图生成步骤中，在将所述各个房间的 2D 平面图进行拼接处理时，对重叠的部分和/或出现的空洞进行修正。

34、如权利要求 31 所述的自动 3D 建模方法，其特征在于，还包括：

2D 平面图生成步骤，以如下方式生成 2D 平面图：

将所述 3D 模型拼接步骤中生成的所述整体 3D 模型的每个面向垂直于地板的平面投影，并将这些投影合并成为一个多边形；

对得到的所述多边形进行校正和简化，包括以下至少一种方式：

(1) 仅保留所述多边形的主要顶点，删除小的凹凸，

(2) 利用计算机视觉技术检测图片中的直线，进而确定墙壁的走向，并将与墙壁走向近似平行或垂直的边都修正到对应的方向上；

识别并标示出门和/或窗的位置，包括使用深度学习的方法，在所述室内图像上识别出门和/或窗的位置，或者根据对属于同一套房子的多个房间分别拍摄所述第一图像时的移动轨迹与该房间轮廓的交点，确定门的位置。

35、一种基于拍照的 3D 建模方法，其特征在于，包括如下步骤：

S1) 将具有拍照功能的移动设备和相机固定在同一拍摄支架上；

S2) 在移动支架的过程中获取多张相机或移动设备的第二图像，结合所述相机或移动设备的传感器来获取每个拍摄点的位置和拍摄方向，形成使用统一坐标系的路线图；

S3) 在所述移动设备上或远程服务器端基于在每个拍摄点拍摄的第一图像进行 3D 建模；

S4) 将每个拍摄点的所述 3D 模型按照 S2 中获取的位置和拍摄方向，统一放置在同一个三维坐标系内，对多个拍摄点 3D 模型之间的结合部位进行拼接处理，形成包含多拍摄点的整体 3D 模型。

36、如权利要求 35 所述的基于拍照的 3D 建模方法，其特征在于，

所述步骤 S2 是基于所述移动设备或相机的定位系统，使用所述移动设备或相机拍摄的第二图像，通过相近拍摄点的第二图像进行特征点匹配来找到拍摄点的相对位移和拍摄方向信息，形成包括所有拍摄点的在同一坐标系内的所述路线图，并提供每个拍摄点的位置和方向。

37、如权利要求 35 或 36 所述的基于拍照的 3D 建模方法，其特征在于，

所述步骤 S2 还包括通过所述移动设备或相机的传感器，获取包括移动的加速度、速度和方向信息，对所述路线图进行校正。

38、如权利要求 36 所述的基于拍照的 3D 建模方法，其特征在于，

所述步骤 S2 还包括获取所述相机的镜头和所述移动设备的拍摄方向的夹角，在初始化阶段同时运行基于所述移动设备的定位系统和基于所述相机的定位系统，并

移动支架一段距离，此时两套系统各提供一个位移向量，两个向量的夹角即为所述相机的镜头与所述移动设备的拍摄方向的夹角；或者通过手工旋转所述相机的预览图或者拍摄的图像，指定与所述移动设备拍摄方向一致的角度；或者通过图像识别算法匹配所述移动设备与所述相机的预览图或拍摄的图像，找到夹角；或者通过使用额外的标识、包括在支架上增加刻度与所述移动设备的安装方向形成固定夹角，然后在所述相机的预览图或图像中识别该标识来计算所述相机的镜头和所述移动设备拍摄方向的夹角。

39、如权利要求 35 所述的基于拍照的 3D 建模方法，其特征在于，所述步骤 S3 包括：

S31) 基于深度学习技术，识别图像中的地板、天花板、墙壁的至少一者的图像；

S32) 基于图像处理技术，对识别出的图像进行分块，每块近似认为是一个平面，地板和天花板的图像块位于水平平面，墙壁的图像块位于垂直平面，求解各个平面对应的方程生成所述 3D 模型；对于在图像中相交的两个平面，以这两个平面的一条交线为约束条件，使得计算得到的交线与实际观察到的交线的误差最小。

40、如权利要求 35 所述的基于拍照的 3D 建模方法，其特征在于，

所述步骤 S3 还包括：对于室内图像，利用计算机视觉技术识别图像中的墙角，并将墙角连接起来即为房间的粗略模型。

41、如权利要求 35 所述的基于拍照的 3D 建模方法，其特征在于，所述步骤 S4 包括：

S41) 根据每个拍摄点的位置和拍摄方向，把单个拍摄点的 3D 模型的局部坐标转换为成全局的世界坐标，从而获得所有拍摄点的整体 3D 模型；

S42) 对多个拍摄点的 3D 模型进行统一校正，包括使用统计的方法对所有的拍摄点的墙线方向进行统一校正，使得所有房间在一定偏差范围内的墙线平行；

S43) 在将各个拍摄点的 3D 模型进行拼接处理时，对其重叠的部分和/或出现的空洞进行修正。

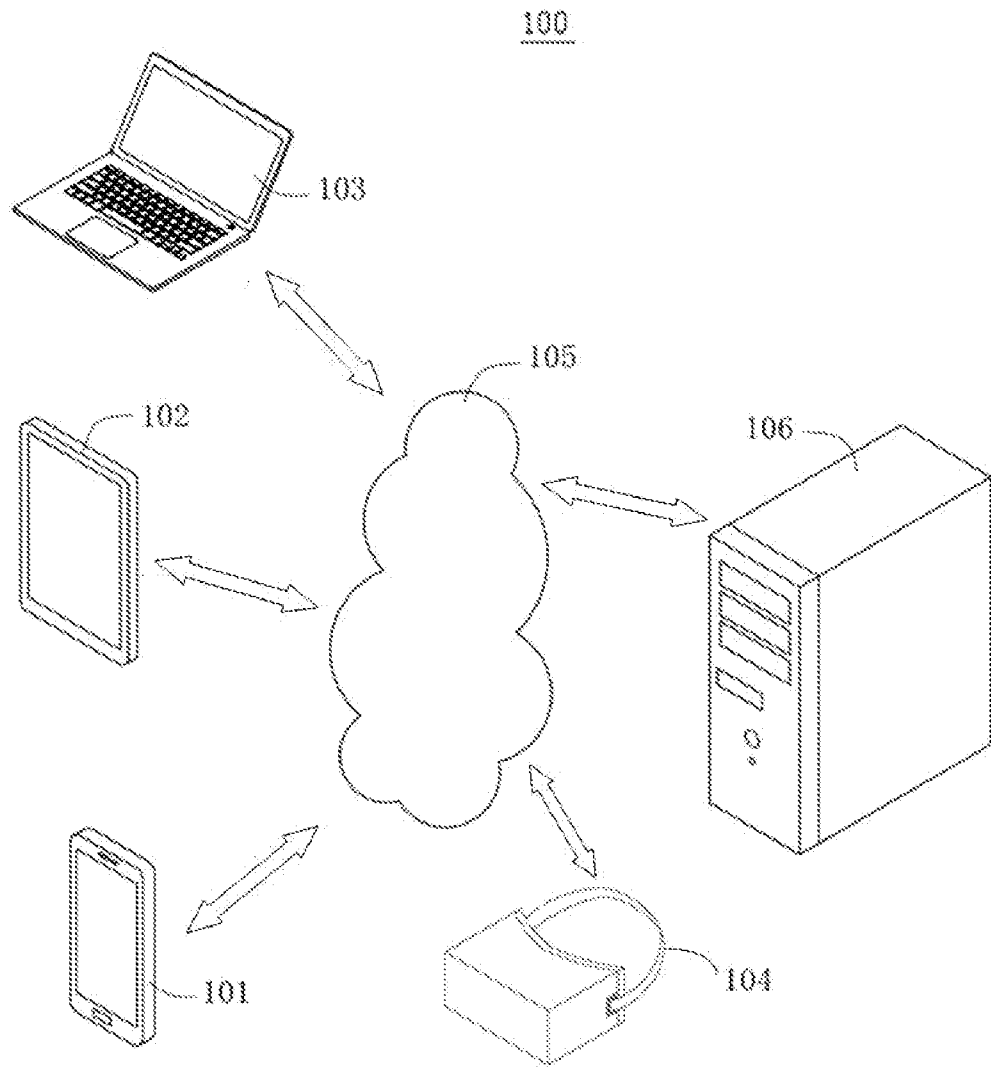


图 1

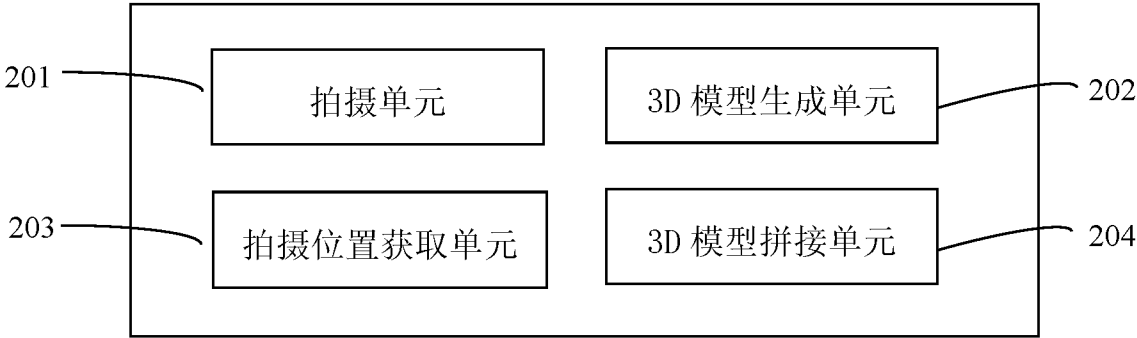


图 2

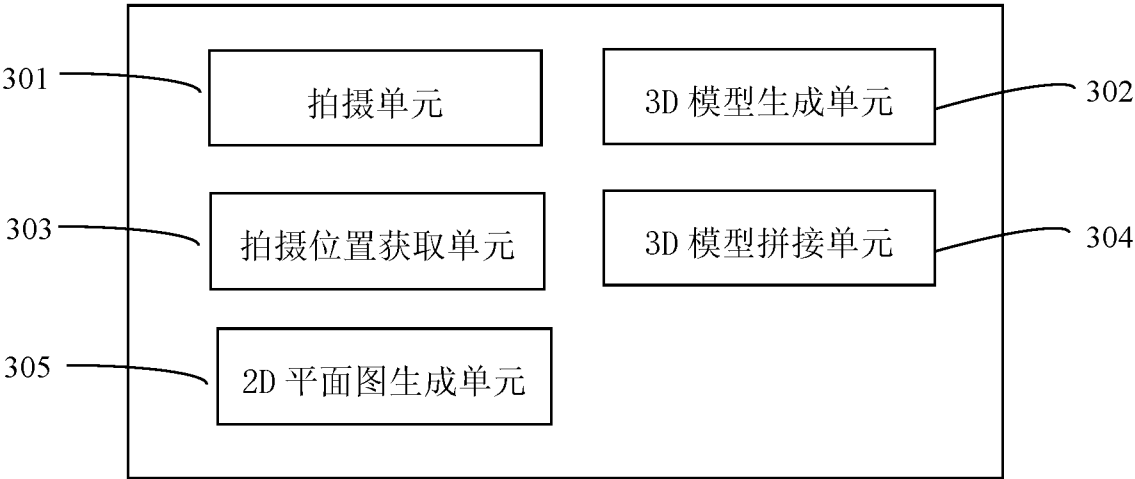


图 3

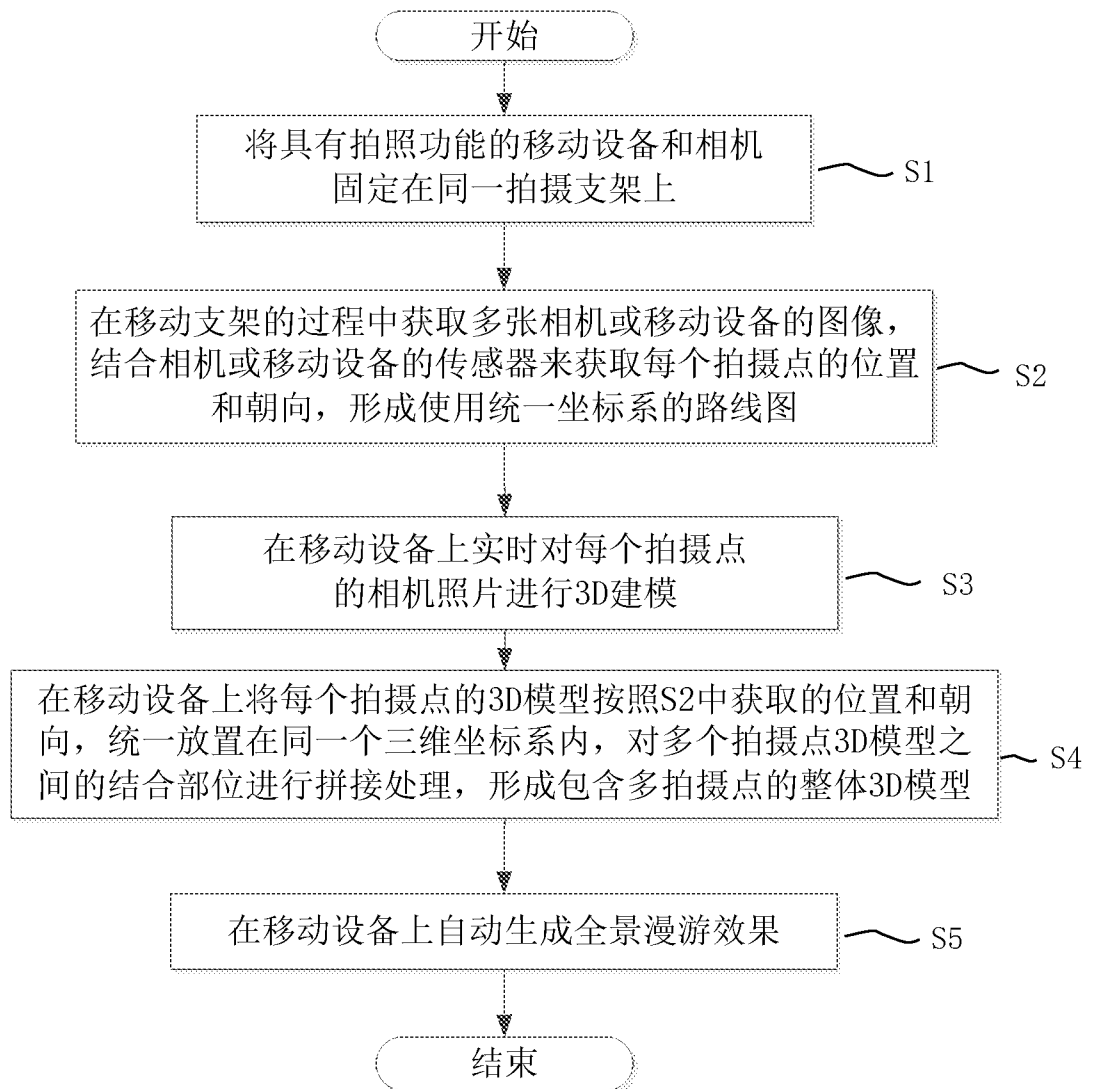


图 4

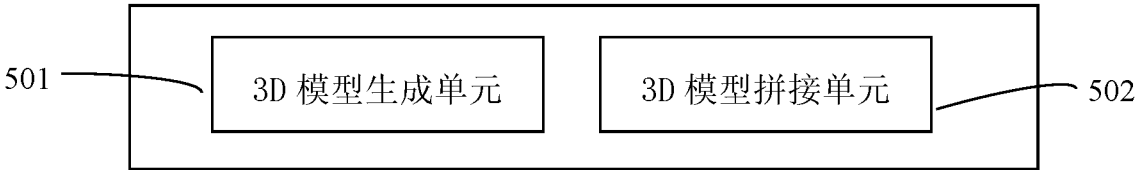


图 5

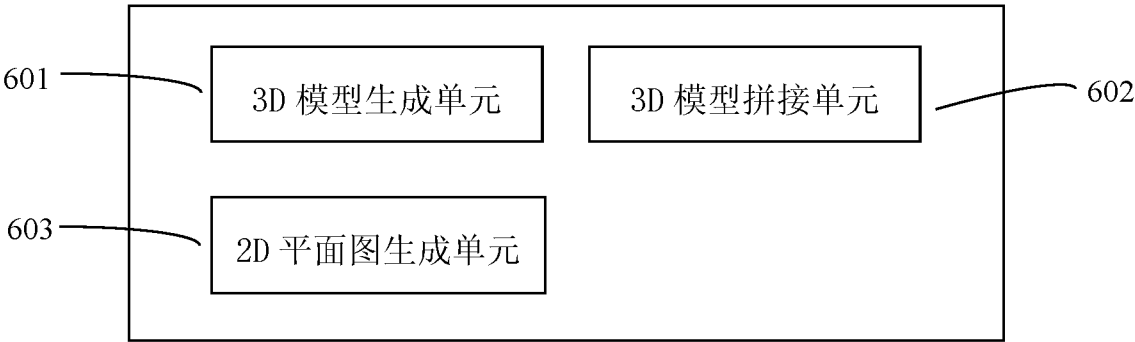


图 6

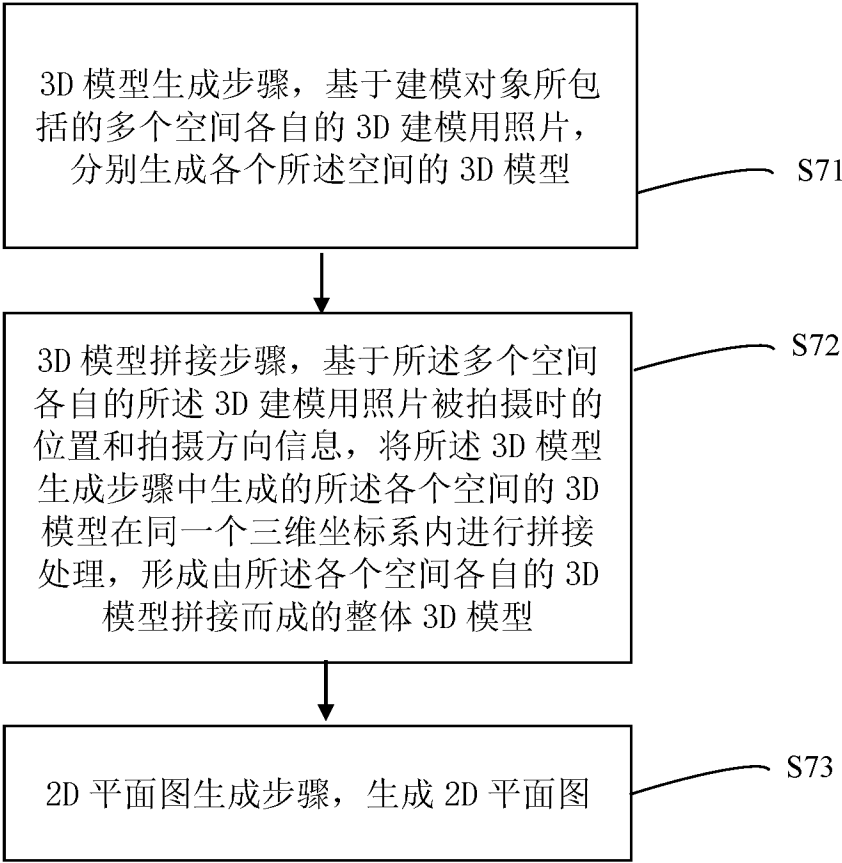


图 7

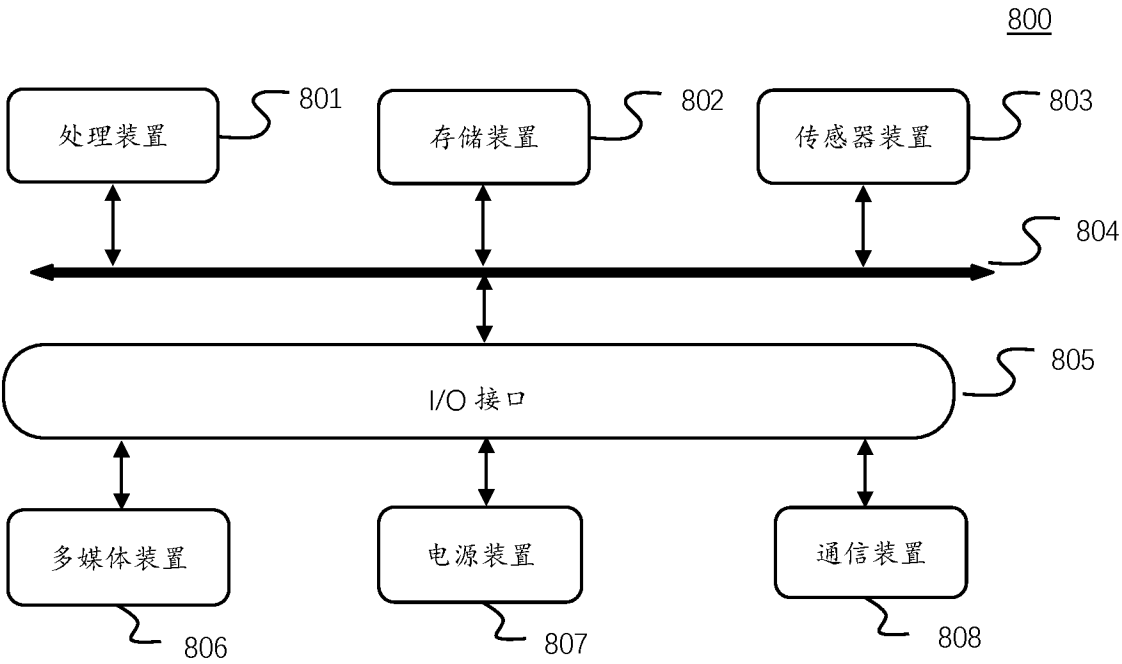


图 8

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2020/090469

A. CLASSIFICATION OF SUBJECT MATTER

G06T 17/00(2006.01)i; H04N 13/106(2018.01)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)

G06T; H04N

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)

CNPAT; CNKI; WPI; EPODOC; IEEE: 房间, 空间, 图像, 位置, 方向, 角度, 拼接, 组合, 合成, 拼合, 坐标系, 3D, 三维, 立体, 相机, 像机, 拍摄, 位移, 移动, 路线, camera, position, combin+, coordinate, three-dimensional, stereo+, direction, space, locat+, route

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 110505463 A (SHANGHAI YIWO INFORMATION TECHNOLOGY CO., LTD.) 26 November 2019 (2019-11-26) entire document	1-41
Y	CN 108961395 A (SHANGHAI YIWO INFORMATION TECHNOLOGY CO., LTD.) 07 December 2018 (2018-12-07) claims 1-11, description pages 4-9	1-41
Y	CN 108198133 A (YUNNAN UNITED VISUAL TECHNOLOGY CO., LTD.) 22 June 2018 (2018-06-22) claims 1-6, description pages 3-4, figure 1	1-41
A	CN 106092056 A (NATIONAL UNIVERSITY OF DEFENSE TECHNOLOGY) 09 November 2016 (2016-11-09) entire document	1-41
A	CN 107424118 A (TIANJIN UNIVERSITY) 01 December 2017 (2017-12-01) entire document	1-41
A	CN 1605830 A (TIANJIN UNIVERSITY) 13 April 2005 (2005-04-13) entire document	1-41



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

“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

29 July 2020

Date of mailing of the international search report

19 August 2020

Name and mailing address of the ISA/CN

China National Intellectual Property Administration (ISA/
CN)
No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing
100088
China

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2020/090469**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 2007286715 A (BROTHER IND. LTD.) 01 November 2007 (2007-11-01) entire document	1-41
A	CN 105488852 A (THE 715TH RESEARCH INSTITUTE OF CHINA SHIPBUILDING INDUSTRY CORPORATION) 13 April 2016 (2016-04-13) entire document	1-41
A	US 2016343166 A1 (TEAMLAB INC.) 24 November 2016 (2016-11-24) entire document	1-41

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2020/090469

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	110505463	A	26 November 2019	CN	110769240	A	07 February 2020
CN	108961395	A	07 December 2018	WO	2020006941	A1	09 January 2020
				US	2020111250	A1	09 April 2020
				CN	108961395	B	30 July 2019
				CN	111247561	A	05 June 2020
CN	108198133	A	22 June 2018	None			
CN	106092056	A	09 November 2016	CN	106092056	B	13 July 2018
CN	107424118	A	01 December 2017	None			
CN	1605830	A	13 April 2005	CN	1233984	C	28 December 2005
JP	2007286715	A	01 November 2007	JP	4835243	B2	14 December 2011
CN	105488852	A	13 April 2016	CN	105488852	B	26 April 2019
US	2016343166	A1	24 November 2016	JP	6340017	B2	06 June 2018
				JP	WO2015098807	A1	23 March 2017
				WO	2015098807	A1	02 July 2015

国际检索报告

国际申请号

PCT/CN2020/090469

A. 主题的分类

G06T 17/00(2006.01)i; H04N 13/106(2018.01)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)

G06T; H04N

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))

CNPAT; CNKI; WPI; EPODOC; IEEE: 房间、空间、图像、位置、方向、角度、拼接、组合、合成、拼合、坐标系、3D、三维、立体、相机、像机、拍摄、位移、移动、路线、camera、position、combin+、coordinate、three-dimensional、stereo+、direction、space、locat+、route

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
PX	CN 110505463 A (上海亦我信息技术有限公司) 2019年 11月 26日 (2019 - 11 - 26) 全文	1-41
Y	CN 108961395 A (上海亦我信息技术有限公司) 2018年 12月 7日 (2018 - 12 - 07) 权利要求1-11, 说明书第4-9页	1-41
Y	CN 108198133 A (云南联合视觉科技有限公司) 2018年 6月 22日 (2018 - 06 - 22) 权利要求1-6, 说明书第3-4页, 附图1	1-41
A	CN 106092056 A (中国人民解放军国防科学技术大学) 2016年 11月 9日 (2016 - 11 - 09) 全文	1-41
A	CN 107424118 A (天津大学) 2017年 12月 1日 (2017 - 12 - 01) 全文	1-41
A	CN 1605830 A (天津大学) 2005年 4月 13日 (2005 - 04 - 13) 全文	1-41
A	JP 2007286715 A (BROTHER IND. LTD.) 2007年 11月 1日 (2007 - 11 - 01) 全文	1-41

☒ 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

“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

2020年 7月 29日

国际检索报告邮寄日期

2020年 8月 19日

ISA/CN的名称和邮寄地址

中国国家知识产权局(ISA/CN)

中国北京市海淀区蓟门桥西土城路6号 100088

传真号 (86-10)62019451

受权官员

吴倩

电话号码 86-(010)-53961822

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
A	CN 105488852 A (中国船舶重工集团公司第七一五研究所) 2016年 4月 13日 (2016 - 04 - 13) 全文	1-41
A	US 2016343166 A1 (TEAMLAB INC.) 2016年 11月 24日 (2016 - 11 - 24) 全文	1-41

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2020/090469

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	110505463	A	2019年 11月 26日	CN	110769240	A	2020年 2月 7日
CN	108961395	A	2018年 12月 7日	WO	2020006941	A1	2020年 1月 9日
				US	2020111250	A1	2020年 4月 9日
				CN	108961395	B	2019年 7月 30日
				CN	111247561	A	2020年 6月 5日
CN	108198133	A	2018年 6月 22日	无			
CN	106092056	A	2016年 11月 9日	CN	106092056	B	2018年 7月 13日
CN	107424118	A	2017年 12月 1日	无			
CN	1605830	A	2005年 4月 13日	CN	1233984	C	2005年 12月 28日
JP	2007286715	A	2007年 11月 1日	JP	4835243	B2	2011年 12月 14日
CN	105488852	A	2016年 4月 13日	CN	105488852	B	2019年 4月 26日
US	2016343166	A1	2016年 11月 24日	JP	6340017	B2	2018年 6月 6日
				JP	W02015098807	A1	2017年 3月 23日
				WO	2015098807	A1	2015年 7月 2日