

三维对象建模方法、图像处理方法、图像处理装置

本申请要求于 2019 年 04 月 12 日递交的中国专利申请第 201910296077.0 号、第 201910296079.X 号和第 201910295558.X 号的优先权，在此全文引用上述中国专利申请公开的内容以作为本申请的一部分。

技术领域

本发明涉及图像处理和三维对象建模领域，尤其涉及三维对象建模方法与设备、图像处理方法、图像处理装置及介质。

背景技术

在对象建模领域，如何使得生成的对象模型具有高分辨率和/或高准确度是业界极力追求的目标。

对象建模可以实现用户足不出户（例如通过网络）即可浏览三维对象的 2D 和/或 3D 结构，并且，对象的 3D 建模可实现如同身临其境的效果，这在虚拟现实领域是一项非常重要的应用。

目前主要有两种方式进行三维对象建模。

一种是采用普通图像采集设备从不同的角度来拍摄多个图像，然后将从不同角度拍摄的多个图像组合/拼接起来，构成三维对象的三维模型。但是，这种方式在拍摄多个图像时，拍摄角度和位置一般需要是固定的，这种拍摄方式有很大的缺陷，往往无法拍摄不能移动的大的三维对象，尤其是要对多个这样的各自位置固定、空间密闭的、而且很难全方位多角度拍摄的三维对象（例如室内房屋或室外复杂空间）进行建模时；而且，这种方式需要复杂的图像拼接，很难得到高准确度的三维目标。

另一种方式是直接使用三维扫描设备获取三维对象的三维点云，然后进行三维点云的拼接，生成三维模型。但是，这种三维扫描设备的图像采集设备精度不高，导致捕获的图像分辨率不高，致使生成的三维模型的分辨率不够高。

因此，如何得到高分辨率的采集图像，如何高效地处理采集的图像以便为对象建模提供高精度和高分辨率的建模准备数据、甚至如何使得提供的数据能够简化后续的模型生成过程，如何有效地提高所生成的对象模型的分辨率和/或准确性，如何准确地拼接图像以得到逼真的三维模型是本发明考

虑要解决的技术问题。

发明内容

为了解决以上问题之一，本发明提供了一种三维对象建模方法与设备、
5 图像处理方法、图像处理装置及介质。

根据本发明的一个实施例，提供一种三维对象建模方法，包括：平面轮廓提取步骤，其中，针对为待处理的至少一个三维对象的每一个三维对象拍摄的至少一幅全景图像，提取所述每一个三维对象的至少一幅全景图像的在三维空间中的平面轮廓；尺度归一化步骤，其中，基于相机位置，将所述
10 每一个三维对象的至少一幅全景图像的在三维空间中的平面轮廓的尺度进行归一化，得到经过归一化的各全景图像的在三维空间中的平面轮廓；以及多对象拼接步骤，其中，基于相机位置，对经过尺度归一化的各全景图像的在三维空间中的平面轮廓的三维点坐标进行旋转和平移操作，以将所述三维点坐标统一到同一个坐标系中，从而将各三维对象的平面轮廓拼接成多对象平面轮廓。
15 廓。

可选地，所述平面轮廓提取步骤包括：相机位置获取步骤，其中，利用针对为所述待处理的至少一个三维对象中的每一个三维对象拍摄的至少两幅全景图像的几何关系，获取在拍摄所述每一个三维对象的至少两幅全景图像的每幅全景图像时的全景相机的所述相机位置、以及所述每幅全景图
20 像上的匹配特征点的三维点坐标；单图像平面轮廓生成步骤，其中，对于所述每一个三维对象的至少两幅全景图像的每幅全景图像，基于所述全景图像上的、轮廓特征属于特定类别的像素点中的边缘像素点所围成的轮廓，生成所述全景图像的在所述三维空间中的平面轮廓，其中，所述特定类别至少包括所述全景图像中的三维物体的顶部、底部、支撑部。

25 可选地，所述单图像平面轮廓生成步骤包括：基于所述每一个三维对象的至少两幅全景图像中的每幅全景图像上的像素点之间的特征相似度，确定所述每一个三维对象的至少两幅全景图像中的每幅全景图像上的、轮廓特征属于所述特定类别的像素点中的所述边缘像素点，其中，两个像素点的特征相似度为所述两个像素点的特征之差的绝对值，每个所述像素点的特
30 征包括灰度、颜色。

可选地，所述平面轮廓提取步骤包括：相机位置获取步骤，其中，利用

针对为所述待处理的至少一个三维对象中的每一个三维对象拍摄的至少两幅全景图像的几何关系，获取在拍摄所述每一个三维对象的至少两幅全景图像的每幅全景图像时的全景相机的所述相机位置、以及所述每幅全景图像上的匹配特征点的三维点坐标；单图像平面轮廓生成步骤，其中，对于所述每一个三维对象的至少两幅全景图像的每幅全景图像，通过用于提取图像轮廓的深度学习方法，提取所述全景图像的在所述三维空间中的平面轮廓。

可选地，通过下述方式来训练所述深度学习模型：通过人工标注作为训练图像的多个全景图像上的被关注类型的三维对象的轮廓特征，生成所述被关注类型的三维对象的轮廓特征训练数据集；利用所述被关注类型的三维对象的所述轮廓特征训练数据集来训练所述深度学习模型，由此得到经过训练的深度学习模型，其中，所述深度学习模型的输出包含所述被关注类型的三维对象的轮廓特征类别。

可选地，所述单图像平面轮廓生成步骤包括：将所述每一个三维对象的至少两幅全景图像输入到经过训练的所述深度学习模型中，得到所述每一个三维对象的至少两幅全景图像中每个像素点对应的轮廓特征的类别；提取出所述每一个三维对象的至少两幅全景图像的每幅全景图像中，轮廓特征属于特定类别的像素点中的处于边缘的边缘像素点作为特定类别边缘像素点；假定所述每一个三维对象的至少两幅全景图像的每幅全景图像上的所有特定类别边缘像素点具有同一高度 h_c ，作为特定类别的轮廓的假定轮廓高度，并将所述每幅全景图像上的所述特定类别边缘像素点投影到三维平面上，得到与所述每幅全景图像对应的特定类别三维点，然后基于与所述每幅全景图像对应的所述特定类别三维点，形成各全景图像的在所述三维空间中的平面轮廓；其中，所述特定类别包括所述全景图像中三维对象的顶部。

可选地，所述相机位置是在所述相机位置获取步骤中得到的，所述相机位置获取步骤包括：获取所述每一个三维对象的至少一幅全景图像对应的视频流和时间戳；基于所述时间戳确定所述视频流中与所述每一个三维对象的至少一幅全景图像对应的至少一个视频帧；根据所述至少一个视频帧的位置，确定拍摄所述每一个三维对象的至少一幅全景图像时的全景相机的相机位置。

可选地，所述相机位置是预先确定的或者是在所述相机位置获取步骤

中得到的，相机位置获取步骤包括：利用针对为所述待处理的至少一个三维对象的每一个三维对象拍摄的至少两幅全景图像的几何关系，进行所述每一个三维对象的至少两幅全景图像之间的特征点匹配，并记录所述每一个三维对象的至少两幅全景图像中互相匹配的特征点作为匹配特征点；通过
5 对于所述每一个三维对象的至少两幅全景图像的每幅全景图像，减小所述每幅全景图像上的匹配特征点的重投影误差，来得到拍摄所述每幅全景图像时的全景相机的所述相机位置、以及所述每幅全景图像上的匹配特征点的三维点坐标。

可选地，尺度归一化步骤包括：在所述相机位置是在所述相机位置获取
10 步骤中得到的情况下，将在获取所述相机位置时得到的所述每一个三维对象的至少一幅全景图像上的所有三维点坐标中的高度值从小到大排序，取排序靠前的高度值的中值或者均值作为特定类别的轮廓的提取的轮廓高度 h_c' ；以及利用特定类别的轮廓的假定轮廓高度 h_c 与特定类别的轮廓的提取的轮廓高度 h_c' 的比值，从所述每一个三维对象的至少两幅全景图像的在三维空间中的
15 平面轮廓生成经过归一化的所述每一个三维对象的至少两幅全景图像的在三维空间中的平面轮廓，其中，特定类别的轮廓的假定轮廓高度 h_c 是任意假定的一个高度。

可选地，尺度归一化步骤包括：在相机位置是预先确定的相机确定高度 h_c' 的情况下，利用假定的相机假定高度 h_c 与相机确定高度 h_c' 的比值，从所述
20 每一个三维对象的至少两幅全景图像的在三维空间中的平面轮廓生成经过归一化的所述每一个三维对象的至少两幅全景图像的在三维空间中的平面轮廓，其中，相机假定高度 h_c 是任意假定的一个高度。

可选地，在尺度归一化步骤之后，三维对象建模方法还包括单对象平面轮廓生成步骤，包括：对于针对为所述待处理的至少一个三维对象拍摄的至
25 少两幅全景图像，通过以下方式、逐个地确定其中是否有多幅全景图像属于同一个三维对象：如果两幅全景图像之间有多于特定比例的匹配特征点，则确定这两幅全景图像属于同一个三维对象；如果确定多幅全景图像属于同一个三维对象，则对于由所述多幅全景图像得到的同一个三维对象的各平面轮廓，取所述同一个三维对象的所有平面轮廓的并集，作为所述三维对象的在三维空间中的
30 平面轮廓。

可选地，在所述多对象拼接步骤，还能够基于各单个三维对象的在三维

空间中的平面轮廓，拼接得到在三维空间中的多对象平面轮廓。

可选地，多对象拼接步骤包括：假设所有的全景图像的在三维空间的平面轮廓共有 N 个，第 n 个平面轮廓的第 p 个三维点表示为 X_p^n ，所述第 n 个平面轮廓对应的全景图像被拍摄时的相机位置表示为 $\{R_n, t_n\}$ ，其中， R_n 为用于表示所述相机位置的旋转参数的旋转矩阵， t_n 为用于表示所述相机位置的平移参数的平移向量，其中， N 为大于 1 的整数， n 为大于等于 1 的整数，选定第 i 个平面轮廓对应的全景图像被拍摄时的相机位置作为参考坐标系，可通过如下的公式将其它平面轮廓的三维点统一到所述参考坐标系下：

$$X_p^{n \rightarrow i} = R_i R_n^T (X_p^n - t_n) + t_i,$$

10 将除第 i 个平面轮廓外的所有的经过尺度归一化的平面轮廓的三维点通过所述公式进行转换，以将所有平面轮廓的三维点统一到同一坐标系下，从而将各三维对象的平面轮廓拼接成多对象平面轮廓。

可选地，所述三维对象建模方法还包括：多对象轮廓优化步骤，其中，计算多对象轮廓中，两个单对象轮廓的相邻边之间的距离，若所述距离非零且
15 小于特定阈值，则将所述两个单对象轮廓进行偏移，以使所述相邻边的距离变为零。

可选地，所述三维对象建模方法还包括：3D 模型生成步骤，其中，在多对象拼接步骤之后，将拼接得到的在三维空间中的多对象平面轮廓转化为多对象 3D 模型。

20 根据本发明的一个示例性实施例，提供一种三维对象建模设备，包括：平面轮廓提取装置，被配置用于针对为待处理的至少一个三维对象的每一个三维对象拍摄的至少一幅全景图像，提取所述每一个三维对象的至少两幅全景图像的在三维空间中的平面轮廓；尺度归一化装置，被配置用于基于相机位置，将所述每一个三维对象的至少两幅全景图像的在三维空间中的平面轮廓的尺度进行归一化，得到经过归一化的各全景图像的在三维空间
25 中的平面轮廓；以及多对象拼接装置，被配置用于基于相机位置，对经过尺度归一化的各全景图像的在三维空间中的平面轮廓的三维点坐标进行旋转和平移操作，以将所述三维点坐标统一到同一个坐标系中，从而将各三维对象的平面轮廓拼接成多对象平面轮廓。

30 可选地，所述相机位置是预先确定的或者是通过以下方式得到的：利用针对为所述待处理的至少一个三维对象的每一个三维对象拍摄的至少两幅

全景图像的几何关系，进行所述每一个三维对象的至少两幅全景图像之间的特征点匹配，并记录所述每一个三维对象的至少两幅全景图像中互相匹配的特征点作为匹配特征点；以及通过对于所述每一个三维对象的至少两幅全景图像的每幅全景图像，减小所述每幅全景图像上的匹配特征点的重投影误差，来得到拍摄所述每幅全景图像时的全景相机的所述相机位置、以及所述每幅全景图像上的匹配特征点的三维点坐标。

可选地，尺度归一化装置还被配置用于：在相机位置是提取到的情况下，将在获取所述相机位置时得到的所述每一个三维对象的至少一幅全景图像上的所有三维点坐标中的高度值从小到大排序，取排序靠前的高度值的中值或者均值作为特定类别的轮廓的提取的轮廓高度 h_c' ；以及利用特定类别的轮廓的假定轮廓高度 h_c 与特定类别的轮廓的提取的轮廓高度 h_c' 的比值，从所述每一个三维对象的至少两幅全景图像的在三维空间中的平面轮廓生成经过归一化的所述每一个三维对象的至少两幅全景图像的在三维空间中的平面轮廓，其中，特定类别的轮廓的假定轮廓高度 h_c 是任意假定的一个高度。

可选地，尺度归一化装置还被配置用于：在相机位置是预先确定的相机确定高度 h_c' 的情况下，利用假定的相机假定高度 h_c 与相机确定高度 h_c' 的比值，从所述每一个三维对象的至少两幅全景图像的在三维空间中的平面轮廓生成经过归一化的所述每一个三维对象的至少两幅全景图像的在三维空间中的平面轮廓，其中，相机假定高度 h_c 是任意假定的一个高度。

可选地，单对象平面轮廓生成装置还被配置用于：对于针对为所述待处理的至少一个三维对象拍摄的至少两幅全景图像，通过以下方式、逐个地确定其中是否有多幅全景图像属于同一个三维对象：如果两幅全景图像之间有多于特定比例的匹配特征点，则确定这两幅全景图像属于同一个三维对象；以及如果确定多幅全景图像属于同一个三维对象，则对于由所述多幅全景图像得到的同一个三维对象的各平面轮廓，取所述同一个三维对象的所有平面轮廓的并集，作为所述三维对象的在三维空间中的平面轮廓。

可选地，所述多对象拼接装置还被配置用于能够基于各单个三维对象的在三维空间中的平面轮廓，拼接得到在三维空间中的多对象平面轮廓。

可选地，多对象拼接装置还被配置用于：假设所有的全景图像的在三维空间的平面轮廓共有 N 个，第 n 个平面轮廓的第 p 个三维点表示为 X_p^n ，所述第 n 个平面轮廓对应的全景图像被拍摄时的相机位置表示为 $\{R_n, t_n\}$ ，其中， R_n 为

用于表示所述相机位置的旋转参数的旋转矩阵， t_n 为用于表示所述相机位置的平移参数的平移向量，其中， N 为大于 1 的整数， n 为大于等于 1 的整数，选定第 i 个平面轮廓对应的全景图像被拍摄时的相机位置作为参考坐标系，可通过如下的公式将其它平面轮廓的三维点统一到所述参考坐标系下：

5
$$X_p^{n-i} = R_i R_n^T (X_p^n - t_n) + t_i,$$

将除第 i 个平面轮廓外的所有的经过尺度归一化的平面轮廓的三维点通过所述公式进行转换，以将所有平面轮廓的三维点统一到同一坐标系下，从而将各三维对象的平面轮廓拼接成多对象平面轮廓。

可选地，所述三维对象建模设备还包括：多对象轮廓优化装置，被配置用于计算多对象轮廓中，两个单对象轮廓的相邻边之间的距离，若所述距离非零且小于特定阈值，则将所述两个单对象轮廓进行偏移，以使所述相邻边的距离变为零。

可选地，所述三维对象建模设备还包括：3D 模型生成装置，被配置用于将拼接得到的在三维空间中的多对象平面轮廓转化为多对象 3D 模型。

15 根据本发明的一个实施例，提供一种图像处理方法，该图像处理方法包括：相机位置获取步骤，其中，利用针对为待处理的至少一个三维对象中的每一个三维对象拍摄的至少两幅全景图像的几何关系，获取在拍摄所述每一个三维对象的至少两幅全景图像时的全景相机的位置、以及所述每一个三维对象的至少两幅全景图像上的匹配特征点的三维点坐标；单图像平面轮廓生成步骤，其中，对于所述每一个三维对象的至少两幅全景图像的每幅全景图像，基于所述全景图像上的、轮廓特征属于特定类别的像素点中的边缘像素点所围成的轮廓，生成所述全景图像的在三维空间中的平面轮廓；以及尺度归一化步骤，其中，将所获取的在拍摄所述每一个三维对象的至少两幅全景图像时的全景相机的位置的尺度和所述每一个三维对象的至少两幅全景图像的在
20 三维空间中的平面轮廓的尺度进行归一化，得到经过归一化的各全景图像的在
25 三维空间中的平面轮廓。

可选地，所述特定类别至少包括全景图像中的三维物体的顶部、底部、支撑部。

30 可选地，相机位置获取步骤包括：利用拍摄的所述每一个三维对象的至少两幅全景图像的几何关系，进行所述每一个三维对象的至少两幅全景图像之间的特征点匹配，并记录所述每一个三维对象的至少两幅全景图像

中互相匹配的特征点作为匹配特征点；以及通过对于所述每一个三维对象的至少两幅全景图像的每幅全景图像，减小所述每幅全景图像上的匹配特征点的重投影误差，来得到拍摄所述每幅全景图像时的全景相机的所述相机位置、以及所述每幅全景图像上的匹配特征点的三维点坐标。

- 5 可选地，单图像平面轮廓生成步骤包括：基于所述每一个三维对象的至少两幅全景图像中的每幅全景图像上的像素点之间的特征相似度，确定所述每一个三维对象的至少两幅全景图像中的每幅全景图像上的、轮廓特征属于特定类别的像素点中的所述边缘像素点，其中，两个像素点的特征相似度为所述两个像素点的特征之差的绝对值，所述像素点的特征包括灰度、颜色。

- 10 可选地，尺度归一化步骤包括：将在相机位置获取步骤得到的所述每一个三维对象的至少一幅全景图像上的所有三维点坐标中的高度值从小到大排序，取排序靠前的高度值的中值或者均值作为特定类别的轮廓的提取的轮廓高度 h_c' ；以及利用特定类别的轮廓的假定轮廓高度 h_c 与特定类别的轮廓的提取的轮廓高度 h_c' 的比值，从所述每一个三维对象的至少两幅全景图像的在三维空间中的平面轮廓，生成经过归一化的各全景图像的在三维空间中的平面轮廓，其中，所述特定类别的轮廓的假定轮廓高度 h_c 是任意假定的一个高度。

- 20 根据本发明的一个示例性实施例，提供一种图像处理方法，包括：相机位置获取步骤，其中，利用针对待处理的至少一个三维对象中的每一个三维对象拍摄的至少两幅全景图像的几何关系，获取在拍摄所述每一个三维对象的至少两幅全景图像的每幅全景图像时的全景相机的所述相机位置、以及所述每幅全景图像上的匹配特征点的三维点坐标；单图像平面轮廓生成步骤，其中，对于所述每一个三维对象的至少两幅全景图像的每幅全景图像，通过用于提取图像轮廓的深度学习模型，提取所述全景图像的在所述三维空间中的平面轮廓；以及尺度归一化步骤，其中，将所获取的在拍摄所述每一个三维对象的至少两幅全景图像时的全景相机的位置的尺度和所述每一个三维对象的至少两幅全景图像的在三维空间中的平面轮廓的尺度进行归一化，得到经过归一化的各全景图像的在三维空间中的平面轮廓。

- 30 可选地，通过下述方式来训练所述深度学习模型：通过人工标注作为训练图像的多个全景图像上的被关注类型的三维对象的轮廓特征，生成所述被关注类型的三维对象的轮廓特征训练数据集；利用所述被关注类型的三

维对象的所述轮廓特征训练数据集来训练所述深度学习模型，由此得到经过训练的深度学习模型，其中，所述深度学习模型的输出包含所述被关注类型的三维对象的轮廓特征类别。

5 可选地，单图像平面轮廓生成步骤包括：将所述每一个三维对象的至少两幅全景图像输入到经过训练的所述深度学习模型中，得到所述每一个三维对象的至少两幅全景图像中每个像素点对应的轮廓特征的类别；提取出所述每一个三维对象的至少两幅全景图像的每幅全景图像中，轮廓特征属于特定类别的像素点中的处于边缘的边缘像素点作为特定类别边缘像素点；假定所述每一个三维对象的至少两幅全景图像的每幅全景图像上的所有特
10 定类别边缘像素点具有同一高度 h_c ，作为特定类别的轮廓的假定轮廓高度，并将所述每幅全景图像上的所述特定类别边缘像素点投影到三维平面上，得到与所述每幅全景图像对应的特定类别三维点，然后基于与所述每幅全景图像对应的所述特定类别三维点，形成各全景图像的在所述三维空间中的平面轮廓。

15 可选地，特定类别是所述全景图像中三维对象的顶部。

可选地，相机位置获取步骤包括：利用针对为所述待处理的至少一个三维对象的每一个三维对象拍摄的至少两幅全景图像的几何关系，进行所述每一个三维对象的至少两幅全景图像之间的特征点匹配，并记录所述每一个三维对象的至少两幅全景图像中互相匹配的特征点作为匹配特征点；
20 以及通过对于所述每一个三维对象的至少两幅全景图像的每幅全景图像，减小所述每幅全景图像上的匹配特征点的重投影误差，来得到拍摄所述每幅全景图像时的全景相机的所述相机位置、以及所述每幅全景图像上的匹配特征点的三维点坐标。

可选地，尺度归一化步骤包括：将在获取所述相机位置时得到的所述每一个三维对象的至少一幅全景图像上的所有三维点坐标中的高度值从小到大排序，取排序靠前的高度值的中值或者均值作为特定类别的轮廓的提取的轮廓高度 h_c' ；以及利用特定类别的轮廓的假定轮廓高度 h_c 与特定类别的轮廓的提取的轮廓高度 h_c' 的比值，从所述每一个三维对象的至少两幅全景图像的在三维空间中的平面轮廓生成经过归一化的所述每一个三维对象的至少两幅全景图像的在三维空间中的平面轮廓，其中，特定类别的轮廓的假定轮廓高度 h_c 是任意假定的一个高度。
30

根据本发明的一个实施例，提供一种三维对象建模方法，该三维对象建模方法包括：图像处理步骤，其中，使用如上所述的图像处理方法之一，对所述每一个三维对象的至少一幅全景图像进行图像处理，以得到经过归一化的各全景图像的在三维空间中的平面轮廓；以及多对象拼接步骤，其中，基于所述经过归一化的各全景图像的在三维空间中的平面轮廓，拼接得到多对象平面轮廓。

可选地，该三维对象建模方法还包括：单对象平面轮廓生成步骤，其中，基于所述图像处理步骤中得到的经过归一化的各全景图像的平面轮廓，得到各单个三维对象的在三维空间中的平面轮廓。

10 可选地，单对象平面轮廓生成步骤包括：对于针对为待处理的至少一个三维对象拍摄的至少两幅全景图像，通过以下方式、逐个地确定其中是否有多幅全景图像属于同一个三维对象：如果两幅全景图像之间有多于特定比例的匹配特征点，则确定这两幅全景图像属于同一个三维对象；以及如果确定多幅全景图像属于同一个三维对象，则对于由所述多幅全景图像得到的同一个三维对
15 象的各平面轮廓，取这些平面轮廓的并集，作为所述三维对象的平面轮廓。

可选地，在所述多对象拼接步骤，还能够基于各全景图像的在三维空间中的平面轮廓，拼接得到在三维空间中的多对象平面轮廓。

可选地，该三维对象建模方法还包括：3D 模型生成步骤，其中，在多对象拼接步骤之后，将拼接得到的在三维空间中的多对象平面轮廓转化为对象
20 3D 模型。

可选地，3D 模型生成步骤包括：将拼接得到的所述多对象平面轮廓中的顶部平面轮廓在内部进行三维点插值，并将得到的各顶部平面轮廓上的所有三维点坐标投影到相应的全景图像坐标系中，以获取顶部纹理；将拼接得到的所述多对象平面轮廓中的底部平面轮廓在内部进行三维点插值，并将得到的各
25 底部平面轮廓上的所有三维点坐标投影到相应的全景图像坐标系中，以获取底部纹理；连接顶部轮廓和底部轮廓之间相同平面位置上的三维顶点，构成支撑部的平面轮廓，并对于所述支撑部的平面轮廓内部进行三维点插值，并将得到的各支撑部的平面轮廓的所有三维点坐标投影到相应的全景图像坐标系中，以便获取支撑部纹理；基于所述顶部纹理、底部纹理、支撑部纹理，生成整个三
30 维对象的 3D 纹理模型。

可选地，在所述 3D 模型生成步骤中，将得到的各三维对象的顶部平面轮

廓上的所有三维点坐标中的作为所述特定类别的轮廓的提取的轮廓高度 h_c' 的、相机距离相应三维对象顶部的提取高度的高度值，替换为相机距离相应三维对象底部的提取高度 h_f' ，并且使得到的各三维对象顶部平面轮廓上的所有三维点坐标中的长度和宽度值保持不变，得到相应的各三维对象的底部平面轮廓，

5 其中，所述相机距离相应三维对象顶部的所述特定类别的轮廓的提取的轮廓高度 h_c' 通过如下方式得到：将在获取所述相机位置时得到的所述每一个三维对象的至少一幅全景图像上的所有三维点坐标中的高度值从小到大排序，取排序靠前的高度值的中值或者均值作为所述相机距离相应三维对象顶部的所述特定类别的轮廓的提取的轮廓高度 h_c' ，以及所述相机距离相应三维对象底部的提取高度 h_f' 通过如下方式得到：将在获取所述相机位置时得到的所述每一个

10 三维对象的至少一幅全景图像上的所有三维点坐标中的高度值从小到大排序，取排序靠后的高度值的中值或者均值作为相机距离相应三维对象底部的提取高度 h_f' 。

根据本发明的一个实施例，提供一种图像处理设备，该图像处理设备包

15 括：相机位置获取装置，被配置用于利用针对为待处理的至少一个三维对象中的每一个三维对象拍摄的至少两幅全景图像的几何关系，获取在拍摄所述每一个三维对象的至少两幅全景图像时的全景相机的位置、以及所述每一个三维对象的至少两幅全景图像上的匹配特征点的三维点坐标；单图像平面轮廓生成装置，被配置用于对于所述每一个三维对象的至少两幅全景图像的每

20 幅全景图像，基于所述全景图像上的、轮廓特征属于特定类别的像素点中的边缘像素点所围成的轮廓，生成所述全景图像的在三维空间中的平面轮廓；以及尺度归一化装置，被配置用于将所获取的在拍摄所述每一个三维对象的至少两幅全景图像时的全景相机的位置的尺度和所述每一个三维对象的至少两幅全景图像的在三维空间中的平面轮廓的尺度进行归一化，得到经过归一化的各

25 全景图像的在三维空间中的平面轮廓。

可选地，特定类别至少包括全景图像中的三维物体的顶部、底部、支撑部。

可选地，单图像平面轮廓生成装置还被配置用于：基于所述每一个三维对象的至少两幅全景图像中的每幅全景图像上的像素点之间的特征相似

30 度，确定所述每一个三维对象的至少两幅全景图像中的每幅全景图像上的、轮廓特征属于所述特定类别的像素点中的所述边缘像素点，其中，两

个像素点的特征相似度为所述两个像素点的特征之差的绝对值，每个所述像素点的特征包括灰度、颜色。

可选地，尺度归一化装置还被配置用于：将由相机位置获取装置得到的每一个三维对象的至少一幅全景图像上的所有三维点坐标中的高度值从小到大排序，取排序靠前的高度值的中值或者均值作为用作特定类别的轮廓的提取的轮廓高度 h_c' 的、相机距离相应三维对象顶部的提取高度 h_c' ；以及利用所述相机距离相应三维对象顶部的假定高度 h_c 与所述相机距离相应三维对象顶部的提取高度 h_c' 的比值，从所述每一个三维对象的至少两幅全景图像的在三维空间中的平面轮廓生成经过归一化的所述每一个三维对象的至少两幅全景图像的平面轮廓，其中，相机距离相应三维对象顶部的假定高度 h_c 是任意假定的一个高度。

根据本发明的一个实施例，提供一种对象建模设备，该对象建模设备还包括：如上所述的图像处理设备之一，被配置用于对每一个三维对象的至少一幅全景图像进行图像处理，以得到经过归一化的各全景图像的在三维空间中的平面轮廓；以及多对象拼接装置，被配置用于基于所述经过归一化的各全景图像的在三维空间中的平面轮廓，拼接得到在三维空间中的多对象平面轮廓。

可选地，该对象建模设备还包括：单对象平面轮廓生成装置，被配置用于基于所述经过归一化的各全景图像的在三维空间中的平面轮廓，得到各单个三维对象的在三维空间中的平面轮廓。

可选地，单对象平面轮廓生成装置还被配置用于：对于至少两幅全景图像，通过以下方式、逐个地确定是否有多幅全景图像属于同一个三维对象：如果两幅全景图像之间有多于特定比例的匹配特征点，则可以判定这两幅全景图像属于同一个三维对象；以及如果确定多幅全景图像属于同一个三维对象，则对于由所述多幅全景图像得到的同一个三维对象的各平面轮廓，取所述平面轮廓的并集，作为所述三维对象的在三维空间中的平面轮廓。

可选地，多对象拼接装置还被配置用于基于各单个三维对象的在三维空间中的平面轮廓，拼接得到在三维空间中的多对象平面轮廓。

可选地，该对象建模设备还包括：3D 模型生成装置，被配置用于将拼接得到的在三维空间中的多对象平面轮廓转化为三维对象 3D 模型。

根据本发明的一个示例性实施例，提供一种图像处理设备，包括：相机

位置获取装置，被配置用于利用针对待处理的至少一个三维对象中的每一个三维对象拍摄的至少两幅全景图像的几何关系，获取在拍摄所述每一个三维对象的至少两幅全景图像时的全景相机的位置、以及所述每一个三维对象的至少两幅全景图像上的匹配特征点的三维点坐标；单图像平面轮廓生成装置，
5 被配置用于对于所述每一个三维对象的至少两幅全景图像的每幅全景图像，通过用于提取图像轮廓的深度学习模型，提取所述每幅全景图像的在三维空间中的平面轮廓；以及尺度归一化装置，被配置用于将所获取的在拍摄所述每一个三维对象的至少两幅全景图像时的全景相机的位置的尺度和所述每一个三维对象的至少两幅全景图像的在三维空间中的平面轮廓的尺度进行归一化，得
10 到经过归一化的各全景图像的在三维空间中的平面轮廓。

可选地，通过下述方式来训练所述深度学习模型：通过人工标注作为训练图像的多个全景图像上的被关注类型的三维对象的轮廓特征，生成被关注类型的三维对象的轮廓特征训练数据集；利用被关注类型的三维对象的轮廓特征训练数据集来训练所述深度学习模型，由此得到经过训练的深度学习模型，其中，深度学习模型的输出包含所述被关注类型的三维对象的轮廓特征类别。

可选地，单图像平面轮廓生成装置还被配置用于：将针对为所述待处理的至少一个三维对象的每一个三维对象拍摄的至少两幅全景图像输入到经过训练的所述深度学习模型中，得到所述每一个三维对象的至少两幅全景
20 图像中每个像素点对应的轮廓特征的类别；提取出每一个三维对象的至少两幅全景图像中，轮廓特征属于特定类别的像素点中的处于边缘的边缘像素点作为特定类别边缘像素点；假定所述每一个三维对象的至少两幅全景图像的每幅全景图像上的所有特定类别边缘像素点具有同一高度 h_c ，作为特定类别的轮廓的假定轮廓高度，并将所述每幅全景图像上的所述特定类别
25 边缘像素点投影到三维平面上，得到与所述每幅全景图像对应的特定类别三维点，然后基于与所述每幅全景图像对应的所述特定类别三维点，形成各全景图像的在所述三维空间中的平面轮廓。

可选地，所述特定类别包括所述全景图像中三维对象的顶部。

可选地，相机位置获取装置还被配置用于：利用拍摄的每一个三维对象的至少两幅全景图像的几何关系，进行每一个三维对象的至少两幅全景
30 图像之间的特征点匹配，并记录所述每一个三维对象的至少两幅全景图像

中互相匹配的特征点作为匹配特征点；以及通过对于所述每一个三维对象的至少两幅全景图像的每幅全景图像，减小所述每幅全景图像上的匹配特征点的重投影误差，来得到拍摄所述每幅全景图像时的全景相机的所述相机位置、以及所述每幅全景图像上的匹配特征点的三维点坐标。

- 5 可选地，尺度归一化装置还被配置用于：将由相机位置获取装置得到的每一个三维对象的至少一幅全景图像上的所有三维点坐标中的高度值从小到大排序，取排序靠前的高度值的中值或者均值作为特定类别的轮廓的提取的轮廓高度 h_c' ；以及利用特定类别的轮廓的假定轮廓高度 h_c 与特定类别的轮廓的提取的轮廓高度 h_c' 的比值，从所述每一个三维对象的至少两幅全景图像的在
- 10 三维空间中的平面轮廓生成经过归一化的每一个三维对象的至少两幅全景图像的在三维空间中的平面轮廓，其中，特定类别的轮廓的假定轮廓高度 h_c 是任意假定的一个高度。

可选地，所述特定类别是所述全景图像中三维对象的顶部。

- 根据本发明的一个示例性实施例，提供一种三维对象建模设备，包
- 15 括：图像处理装置，被配置用于使用如上所述的图像处理设备之一，对每一个三维对象的至少一幅全景图像进行图像处理，以得到经过归一化的各全景图像的在三维空间中的平面轮廓；以及多对象拼接装置，被配置用于基于所述经过归一化的各全景图像的在三维空间中的平面轮廓，拼接得到多对象平面轮廓。

- 20 可选地，所述三维对象建模设备还包括：单对象平面轮廓生成装置，被配置用于基于由所述图像处理装置得到的经过归一化的各全景图像的平面轮廓，得到各单个三维对象的在三维空间中的平面轮廓。

- 可选地，单对象平面轮廓生成装置还被配置用于：对于至少两幅全景图像，通过以下方式、逐个地确定其中是否有多幅全景图像属于同一个三维对象：
- 25 如果两幅全景图像之间有多于特定比例的匹配特征点，则确定这两幅全景图像属于同一个三维对象；以及如果确定多幅全景图像属于同一个三维对象，则对于由所述多幅全景图像得到的同一个三维对象的各平面轮廓，取各平面轮廓的并集，作为所述三维对象的平面轮廓。

- 可选地，所述多对象拼接装置还能够基于各单个三维对象的在三维空间
- 30 中的平面轮廓，拼接得到在三维空间中的多对象平面轮廓。

可选地，所述三维对象建模设备还包括：3D 模型生成装置，被配置用于

将拼接得到的在三维空间中的多对象平面轮廓转化为多对象 3D 模型。

根据本发明的再一个实施例，提供一种图像处理装置，包括：处理器；以及存储器，其上存储有可执行代码，当所述可执行代码被所述处理器执行时，使所述处理器执行上面描述的图像处理方法之一。

- 5 根据本发明的又一个实施例，提供一种非暂时性机器可读存储介质，其上存储有可执行代码，当所述可执行代码被处理器执行时，使所述处理器执行上面描述的图像处理方法之一。

- 根据本发明的再一个实施例，提供一种图像处理装置，包括：处理器；以及存储器，其上存储有可执行代码，当所述可执行代码被所述处理器执行
10 时，使所述处理器执行上面描述的三维对象建模方法之一。

根据本发明的又一个实施例，提供一种非暂时性机器可读存储介质，其上存储有可执行代码，当所述可执行代码被处理器执行时，使所述处理器执行上面描述的三维对象建模方法之一。

- 本发明基于使用全景相机拍摄的待处理的三维对象的多张全景图像，
15 进行三维对象的 2D 建模和 3D 建模，克服了现有技术中使用 3D 扫描设备生成对象模型导致的模型分辨率不高的缺陷。

在本发明中，通过使用全景相机拍摄至少一个对象（例如，房间）的全景图像，为多对象建模（例如房屋或车辆之类的建模等）提供高分辨率的采集图像。

- 20 进一步地，在本发明中，采用了高效的三维对象建模方法和图像处理方法，为多对象建模（例如房屋或车辆之类的建模）提供了高分辨率的建模所需数据，而且，所提供的建模所需数据能够简化后续的模型生成过程。

再进一步地，通过本发明的建模方法，能够有效地提高所生成的模型（例如房屋或车辆等的 2D 和/或 3D 模型）的分辨率和/或准确度。

- 25 本发明适用于单对象的 2D 建模和 3D 建模，也适用于多对象的 2D 建模和 3D 建模，能够基于每个三维对象的全景图像来进行 2D 建模和 3D 建模，也就是说，本发明不限制建模的应用场景，例如，本发明可以适用于房屋建模、车辆建模等等各种基于图像来进行建模的场景，本发明实际上是提供
30 了一种创新的综合性的图像处理方案，能够应用于各种基于全景图像来进行对象建模的很多 VR（虚拟现实）场景，例如房屋建模（VR 看房）、车辆建模（VR 看车）、购物场所建模（VR 购物）等等。

附图说明

通过结合附图对本公开示例性实施方式进行更详细的描述，本公开的上述以及其它目的、特征和优势将变得更加明显，其中，在本公开示例性实施方式中，相同的附图标记通常代表相同部件。

图 1A 给出了根据本发明的一个示例性实施例的三维对象建模方法的示意性流程图。

图 1B 给出了根据本发明的一个示例性实施例的三维对象的多个全景图像的在三维空间中的平面轮廓的示意图。

10 图 1C 给出了根据本发明的一个示例性实施例的三维对象的在三维空间中的平面轮廓的示意图。

图 2 给出了根据本发明的另一个示例性实施例的三维对象建模的总体过程的示意性流程图。

15 图 3 给出了根据本发明的一个示例性实施例的三维对象建模设备的示意性框图。

图 4A 给出了根据本发明的一个示例性实施例的图像处理方法的示意性流程图。

图 4B 给出了根据本发明的一个示例性实施例的图像处理与对象建模的总体过程的示意性流程图。

20 图 4C 给出了根据本发明的一个示例性实施例的图像处理设备的示意性框图。

图 4D 给出了根据本发明的一个示例性实施例的对象建模设备的示意性框图。

25 图 5A 给出了根据本发明的另一个示例性实施例的图像处理方法的示意性流程图。

图 5B 给出了根据本发明的一个示例性实施例的作为语义分割模型的深度学习模型的训练方法的示意性流程图。

图 5C 给出了根据本发明的一个示例性实施例的基于语义分割模型来提取全景图像的平面轮廓过程的示意性流程图。

30 图 5D 给出了根据本发明的一个示例性实施例的图像处理与建模的总体过程的示意性流程图。

图 5E 给出了根据本发明的另一个示例性实施例的图像处理设备的示意性框图。

图 5F 给出了根据本发明的另一个示例性实施例的三维对象建模设备的示意性框图。

5 图 6 给出了根据本发明的一个示例性实施例的图像处理装置的示意性框图。

具体实施方式

下面将参照附图更详细地描述本公开的优选实施方式。虽然附图中显示了本公开的优选实施方式，然而应该理解，可以以各种形式实现本公开而不应被这里阐述的实施方式所限制。相反，提供这些实施方式是为了使本公开更加透彻和完整，并且能够将本公开的范围完整地传达给本领域的技术人员。这里需要说明的是，本申请中的数字、序号以及附图标记仅是为了方便描述而出现的，对于本发明的步骤、顺序等等均不构成任何限制，除非在说明书中明确指出了步骤的执行有特定的先后顺序。

在对象建模领域，尤其是 2D 和 3D 建模领域，国内外的技术方案主要分为两大类：手动制作和自动化建模。

对于手动制作的方法，需要依赖大量的人工操作，对于对象的三维结构进行标识，并手动进行多个对象模型之间的拼接。手动制作一套三维对象的 3D 模型需要较长时间，因此大量的三维对象数据需要众多的人员来手动制作，人员成本过高，很难实际应用。

对于自动化 3D 建模的方法，目前大多数采用专业的 3D 扫描设备，可以直接获取单个对象的三维点云，然后进行三维点云的拼接，生成 3D 模型。但是，这种专业的 3D 扫描设备的图像采集设备精度不高，导致捕获的图像分辨率不高，致使生成的三维模型的分辨率不够高。而且，这种 3D 扫描设备通常价格昂贵，很难满足消费级应用的需求。

本发明提供了三维对象建模方法、三维对象建模设备、图像处理方法、图像处理设备、以及图像处理装置和计算机介质。

首先，在本发明中，采用普通的全景相机来拍摄每个三维对象，得到三维对象的高分辨率全景图像，克服了背景技术中描述的 3D 扫描相机捕获的图像分辨率不高的缺陷。

然后，利用拍摄的多张全景图像，可以提取单个全景图像的在三维空间中的平面轮廓（可称为“单图像平面轮廓”）。

进一步地，通过尺度归一化，可以实现单图像平面轮廓的尺度与相机位置的尺度之间的统一，生成经过归一化的各单图像平面轮廓，为后续进行三维对象建模提供高分辨率的并且充分的数据准备，减少后续处理工作的难度。

再进一步地，可以通过对属于同一个三维对象的各单图像平面轮廓进行融合，获得准确的单对象平面轮廓。

再进一步地，可以在三维空间中拼接各单对象平面轮廓，得到多对象模型（此时为 2D 模型）。

另外，可以对多对象模型进行修正，以得到更准确的模型，使得模型展示效果更好。

最后，通过 3D 模型生成，得到完整的、高分辨率的并且准确的 3D 模型。

下面，为了便于理解和描述，将以房屋建模为例，来详细描述本发明的各个处理过程。

这里，首先简单介绍全景相机。全景相机与普通相机的不同在于，普通相机通常只用一个镜头来拍摄，而全景相机用两个或者两个以上的镜头来拍摄，所以全景相机可以实现 360 度的拍摄。

图 1A 给出了根据本发明的一个示例性实施例的三维对象建模方法的示意性流程图。

在根据本发明的一个示例性实施例的三维对象建模方法中，基于针对一个三维对象（例如在一个房间内）拍摄的至少一幅全景图像（一个全景图像只对应一个房间（三维对象），但是一个房间内可能拍摄多个全景图像，即一个房间可能对应多个全景图像），提取该三维对象的每幅全景图像的在三维空间中的平面轮廓，然后对提取的平面轮廓进行归一化，以得到建模所需平面轮廓，然后，通过坐标变换，来实现多平面轮廓的拼接，由此得到多对象平面轮廓（2D 模型），接着，可以再从该多对象平面轮廓生成三维对象的 3D 模型。

如图 1A 所示，在步骤 S110，针对为待处理的至少一个三维对象中的每一个三维对象拍摄的至少一幅全景图像，提取每一个三维对象的至少一

幅全景图像的在三维空间中的平面轮廓。

需要说明的是，一个三维对象（比如室内场景的某个房间）可以对应多张全景图像，该三维对象可以认为由底面（即底部）、顶面（即顶部）和墙面（即支撑部）构成。那么，该全景图像在三维空间中的平面轮廓指的是该全景图像对应的三维对象的底面的轮廓，即全景图像在三维空间中的平面轮廓为二维轮廓。如图 1B 所示，在不同位置（即 A 位置、B 位置和 C 位置）针对某个三维对象拍摄了三张全景图像，每张全景图像的平面轮廓对应不同的矩形框（也可以是其他样式的轮廓），例如，在 A 位置拍摄得到的全景图像所对应的矩形的平面轮廓最大，在 C 位置拍摄得到的全景图像所对应的矩形的平面轮廓最小。需要说明的是，图 1B 中的三张全景图像中的三个黑点表示拍摄全景图像时的相机位置。

需要说明的是，在每个相机位置，可以为同一个三维对象拍摄至少一幅全景图像。

在本步骤中，平面轮廓的提取可以通过多种方式来实现。以下将简要地给出一个例子来说明平面轮廓的提取方法。

以至少一个三维对象为室内房屋为例，则每个房间可以看作一个三维对象，为每个房间拍摄至少一幅全景图像，所以一幅全景图像对应于一个房间，但每个房间可能对应于多个全景图像。

在这种情况下，由于天花板一定是在相机的上方，所以，在拍摄的全景图像中，最上方的像素点一定是属于天花板。再次，由于属于天花板上的像素点大多数具有相似特征，因此，可以根据像素点的特征相似度，最终得到属于天花板的所有像素点。

例如，首先将全景图像第一行的所有像素点均看作是属于天花板的像素点；对于第二行的每个像素点，计算其与第一行中属于同一列的像素点的特征相似度（这里的特征可采用颜色、灰度等特征；两个像素点的特征相似度例如可以为两个像素点的特征之差（例如灰度之差或颜色之差等）的绝对值等）。如果特征相似度在某个阈值之内（若采用 0-255 灰度值，阈值例如可设置为 10），则该像素也属于天花板，再继续计算该列上第三行与第二行的相似度、第四行与第三行的相似度...，直到特征相似度超过该阈值，此时的像素点位置即是天花板的边缘像素点。

这些天花板的边缘像素点形成了天花板的边缘，因此，将这些边缘像素

点投影到三维空间可形成天花板的平面轮廓。

下面将描述像素点到三维空间的投影。

- 假定一个全景图像的宽为 W ，高为 H ，并且假定得到的天花板的边缘像素点 c 在图像坐标系下的坐标为 (p_c, q_c) 。由于全景图像是由球面投影得到的，因此，
- 5 在球面坐标系下将其表示为 (θ_c, ϕ_c) ，其中 $\theta_c \in [-\pi, \pi]$ 为经度， $\phi_c \in [-\pi/2, \pi/2]$ 为纬度。

其中球面坐标系与图像坐标系的关系可通过如下公式1得到：

$$\begin{cases} \theta_c = \left(q_c - \frac{(W-1)}{2} \right) / (W-1) * 2\pi \\ \phi_c = - \left(p_c - \frac{(H-1)}{2} \right) / (H-1) * \pi \end{cases} \quad (1)。$$

- 由于天花板可认为是平面，因此可以认为天花板边缘的像素点距离相机具有统一的高度 h_c （这里的 h_c 可以假定为任意一个值，比如100），可以称为“天花板距离相机的假定高度” h_c 。为了避免误解，这里需要说明的是，天花板距离相机的假定高度 h_c 不是图像坐标系下的纵坐标，而是三维坐标系下的高度（即，三维坐标系下 y 轴的值）。
- 10 天花板距离相机的假定高度” h_c 。为了避免误解，这里需要说明的是，天花板距离相机的假定高度 h_c 不是图像坐标系下的纵坐标，而是三维坐标系下的高度（即，三维坐标系下 y 轴的值）。

如下公式2是用于将天花板边缘的像素点 c 在球面坐标系下的坐标 (θ_c, ϕ_c) 投影到三维平面上的三维点坐标 (x_c, y_c, z_c) ：

$$\begin{cases} x_c = -h_c / \tan(\phi_c) * \sin(\theta_c) \\ y_c = h_c \\ z_c = -h_c / \tan(\phi_c) * \cos(\theta_c) \end{cases} \quad (2)。$$

15

在本文中，术语“图像坐标系”指图像像素点所在的坐标系，主要用来描述图像中像素点所在的位置。由此，全景图像坐标系指全景图像的像素点所在的坐标系，主要用来描述全景图像中像素点所在的位置。

- 请注意，上面仅给出了基于全景图像上的天花板特征点的相似性来生成该全景图像的三维空间中的平面轮廓的一种例子，本发明可以使用的方法不限于该例子。
- 20 生成该全景图像的三维空间中的平面轮廓的一种例子，本发明可以使用的方法不限于该例子。

- 由于天花板可认为是平面，而且，由于全景相机一般使用三角架支撑，在拍摄房屋时，其高度一般为固定的，因此可以认为由所处理的全景图像得到的天花板边缘的各像素点距离相机具有统一的高度，可称为“相机距离天花板的高度”。
- 25 高度”。

这里，由于全景相机一般使用三角架支撑，其高度为固定的，所以也可认为相机距离天花板的高度、相机距离地板的高度是固定的。

对于本步骤得到的三维空间中的平面轮廓（天花板平面轮廓），可以给该轮廓上的各三维点假定一个高度值，比如假定相机距离天花板的高度为 h_c （可称为“相机距离天花板的假定高度” h_c ），并且该假定的高度可以为任意的值，比如 100（可以通过后续的处理来求出真实的相机距离天花板的实际高度，后续所获取的该高度可称为“相机距离天花板的提取高度”）。为了避免混淆，下文将这里假定的相机距离天花板的高度 h_c 称为“相机距离天花板的假定高度” h_c 。

在上面的实施例中，可以基于全景图像，自动化地得到该图像的平面轮廓，而无需人工参与制作，并且无需使用昂贵的 3D 扫描设备。

10 例如，在一些实施例中，在步骤 S110 中，平面轮廓提取步骤包括：相机位置获取步骤，其中，利用针对为待处理的至少一个三维对象的每一个三维对象拍摄的至少两幅全景图像的几何关系，获取在拍摄每一个三维对象的至少两幅全景图像时的全景相机的相机位置、以及每一个三维对象的至少两幅全景图像上的匹配特征点的三维点坐标；单图像平面轮廓生成步骤，其中，
15 对于每一个三维对象的至少两幅全景图像的每幅全景图像，基于所述全景图像上的、轮廓特征属于特定类别的像素点中的边缘像素点所围成的轮廓，生成所述全景图像的在所述三维空间中的平面轮廓。

例如，特定类别至少包括全景图像中的三维物体的顶部、底部、支撑部。

例如，单图像平面轮廓生成步骤包括：基于每一个三维对象的至少两幅
20 全景图像中的每幅全景图像上的像素点之间的特征相似度，确定每一个三维对象的至少两幅全景图像中的每幅全景图像上的、轮廓特征属于特定类别的像素点中的边缘像素点。这里，两个像素点的特征相似度为两个像素点的特征之差的绝对值，每个像素点的特征包括灰度、颜色等等。

例如，在另一些实施例中，在步骤 S110 中，平面轮廓提取步骤包括：相机位置获取步骤，其中，利用针对为所述待处理的至少一个三维对象中的每一个三维对象拍摄的至少两幅全景图像的几何关系，获取在拍摄所述每一个三维对象的至少两幅全景图像的每幅全景图像时的全景相机的所述相机位置、以及所述每幅全景图像上的匹配特征点的三维点坐标；单图像平面轮廓生成步骤，其中，对于所述每一个三维对象的至少两幅全景图像的每幅全景
25 全景图像，通过用于提取图像轮廓的深度学习模型，提取所述全景图像的在所述三维空间中的平面轮廓。

例如，在一些示例中，通过下述方式来训练深度学习模型：通过人工标注作为训练图像的多个全景图像上的被关注类型的三维对象的轮廓特征，生成所述被关注类型的三维对象的轮廓特征训练数据集；利用所述被关注类型的三维对象的所述轮廓特征训练数据集来训练所述深度学习模型，由此得到经过训练的深度学习模型。

例如，深度学习模型的输出包含被关注类型的三维对象的轮廓特征类别。

例如，在一些示例中，单图像平面轮廓生成步骤包括：将所述每一个三维对象的至少两幅全景图像输入到经过训练的所述深度学习模型中，得到所述每一个三维对象的至少两幅全景图像中每个像素点对应的轮廓特征的类别；提取出所述每一个三维对象的至少两幅全景图像的每幅全景图像中，轮廓特征属于特定类别的像素点中的处于边缘的边缘像素点作为特定类别边缘像素点；假定所述每一个三维对象的至少两幅全景图像的每幅全景图像上的所有特定类别边缘像素点具有同一高度 h_c ，作为特定类别的轮廓的假定轮廓高度，并将所述每幅全景图像上的所述特定类别边缘像素点投影到三维平面上，得到与所述每幅全景图像对应的特定类别三维点，然后基于与所述每幅全景图像对应的所述特定类别三维点，形成各全景图像的在所述三维空间中的平面轮廓。

例如，特定类别包括全景图像中三维对象的顶部。

如图 1A 所示，在步骤 S120，基于相机位置，将每一个三维对象的至少一幅全景图像的在三维空间中的平面轮廓的尺度进行归一化，得到经过归一化的各全景图像的在所述三维空间中的平面轮廓。也就是说，将相机位置的尺度和在步骤 S110 得到的全景图像的三维空间平面轮廓的尺度进行归一化（尺度归一化）。

需要说明的是，尺度归一化的目的是为了统一各个全景图像的相机位置的尺度和平面轮廓的尺度，该相机位置的尺度和平面轮廓的尺度可以是真实的尺度，比如米。也就是说，相机位置的坐标和平面轮廓的大小要在统一的度量单位下。

例如，相机位置包括相机的坐标和方向。

由于在步骤 S110 中得到房间的三维空间平面轮廓时采用的是相机距离天花板的假定高度 h_c （当时并不确定相机高度），因此，相机位置的尺度（相机

的实际高度)和三维对象的三维空间平面轮廓的尺度不统一,给后续的房间轮廓拼接造成一定的困难。

在步骤 S110 中,将拍摄每一个三维对象的至少一幅全景图像时的相机位置的尺度和每一个三维对象的至少一幅全景图像的在三维空间中的平面轮廓的尺度进行归一化,以使得能够执行后续的多对象拼接处理。

在本归一化步骤中,可能有两种情况:第一种情况是具有确定的相机位置(真实的相机高度),第二种情况是仍然不确定相机位置。

在相机位置是预先确定的相机确定高度 h_c' 的情况下,可以利用在平面轮廓提取步骤 S110 中假定的相机距离天花板的假定高度 h_c 与相机确定高度 h_c' 的比值,从每一个三维对象的至少两幅全景图像的在三维空间中的平面轮廓生成经过归一化的每一个三维对象的至少两幅全景图像的在三维空间中的平面轮廓。这里,特定类别的轮廓的假定轮廓高度 h_c 可以是任意假定的一个高度,例如坐标值为 100 或者任意一个其它值。

对于相机位置是在相机位置获取步骤中得到的情况,首先,可以通过相机位置获取步骤来获取相机位置,相机位置获取步骤包括:利用针对为待处理的至少一个三维对象的每一个三维对象拍摄的至少两幅全景图像的几何关系,进行每一个三维对象的至少两幅全景图像之间的特征点匹配,并记录每一个三维对象的至少两幅全景图像中互相匹配的特征点作为匹配特征点;以及通过对于所述每一个三维对象的至少两幅全景图像的每幅全景图像,减小所述每幅全景图像上的匹配特征点的重投影误差,来得到拍摄所述每幅全景图像时的全景相机的所述相机位置、以及所述每幅全景图像上的匹配特征点的三维点坐标。

然后,在步骤 S120 中,在相机位置是在相机位置获取步骤中得到的情况下,可以通过以下方式来实现尺度归一化:将获取相机位置时得到的每一个三维对象的至少一幅全景图像上的所有三维点坐标中的高度值从小到大排序,取排序靠前的高度值的中值或者均值作为特定类别的轮廓的提取的轮廓高度 h_c' ;以及利用特定类别的轮廓的假定轮廓高度 h_c 与特定类别的轮廓的提取的轮廓高度 h_c' 的比值,从每一个三维对象的至少两幅全景图像的在三维空间中的平面轮廓生成经过归一化的每一个三维对象的至少两幅全景图像的在三维空间中的平面轮廓。同样,这里的特定类别的轮廓的假定轮廓高度 h_c 也是任意假定的一个高度,例如 100 或者任意一个其它值。

例如，在另一些实施例中，相机位置可以基于拍摄全景图像时的视频流实时确定，例如，相机位置是在相机位置获取步骤中得到的，相机位置获取步骤包括：获取每一个三维对象的至少一幅全景图像对应的视频流和时间戳；基于时间戳确定视频流中与每一个三维对象的至少一幅全景图像对应的至少一个视频帧；根据至少一个视频帧的位置，确定拍摄每一个三维对象的至少一幅全景图像时的全景相机的相机位置。视频流中的每一视频帧均具有确定的位置信息，拍摄全景图像时的全景相机的相机位置即为该全景图像对应的视频帧对应的位置。

在本发明中，可以采用基于多视图几何的方法来获取（估计）相机位置，限于篇幅，在此不再详细描述。

图 2 给出了根据本发明的另一个示例性实施例的三维对象建模的总体过程的示意性流程图。

可选地，如图 2 所示，在尺度归一化步骤之后，三维对象建模方法还包括单对象平面轮廓生成步骤 S125，步骤 S125 包括：对于针对为待处理的至少一个三维对象拍摄的至少两幅全景图像，通过以下方式、逐个地确定其中是否有多幅全景图像属于同一个三维对象：如果两幅全景图像之间有多于特定比例的匹配特征点，则确定这两幅全景图像属于同一个三维对象；以及如果确定多幅全景图像属于同一个三维对象，则对于由所述多幅全景图像得到的同一个三维对象的各平面轮廓，取同一个三维对象的所有平面轮廓的并集，作为三维对象的在三维空间中的平面轮廓。

例如，在步骤 S125，可以基于经过归一化的各全景图像的在三维空间中的平面轮廓，得到各单个对象的在三维空间中的平面轮廓。

在本发明中，从一个全景图像得到一个相应的在三维空间中的平面轮廓，可称为“单对象平面轮廓”。

例如，以每个房间作为被关注类型的一个三维对象为例，如上所述，由于拍摄的全景图像中可能包含同一房间的多张全景图像，在这种情况下，同一房间将对应多个在三维空间中的平面轮廓，导致在后续的多房间拼接处理得到的多房间平面轮廓中，可能出现某个或多个房间的各自的不同全景图像得到的平面轮廓不重合、导致拼接的轮廓出现重叠或混乱的现象。因此，可以考虑先进行同一个房间（对象）轮廓的融合（可称为“单对象融合”），以避免这种现象。而且，单对象融合（例如，同一房间轮廓的融合）也可消除单对象轮廓不完整

的现象。

针对上述的需要单对象融合的情形，下面将以一个房间作为一个三维对象为例，给出下述的示例性方法。

首先，判断两幅全景图像是否属于同一房间。

- 5 这里可以采用基于特征点匹配的方式，如果两幅全景图像之间有多于某个比例（特定比例，例如 50%等）的匹配特征点，则可以判定这两幅全景图像属于同一房间。

10 然后，如果多幅全景图像属于同一房间，即，对于由不同全景图像得到的同一房间的平面轮廓，取这些平面轮廓的并集，作为三维空间中的单房间平面轮廓（一个房间一个轮廓，避免一个房间多个单图像轮廓的情形），由此实现在三维空间中同一房间轮廓的融合。

其中，上述的匹配特征点的比例可以采用以下方式设置：假设图像 1 有 n_1 个特征点，图像 2 有 n_2 个特征点，两幅图像的匹配特征点数为 n 个。那么匹配特征点的比例可以为 $n/\min(n_1, n_2)$ 。

- 15 可选地，可以设定为如果两幅全景图像之间的匹配特征点的比例大于例如 50%，则认为两幅图像是同一个房间。

这里，匹配特征点的比例的设置和比例的实际大小可根据实际情况来试验或凭借经验确定，本发明对此不作额外限制。

- 20 如上所述，在本发明中，对于上述的至少两幅全景图像，可以通过如下的单房间融合的方式来确定是否有多幅全景图像属于同一房间：如果两幅全景图像之间有多于特定比例的匹配特征点，则可以判定这两幅全景图像属于同一房间。

如果确定多幅全景图像属于同一房间，则对于由所述多幅全景图像得到的同一房间的各平面轮廓，取这些平面轮廓的并集，作为该房间的平面轮廓。

- 25 另外，在同一房间的轮廓融合之后，由于获取的轮廓边缘可能存在噪声，例如可能表现为边线不直以及边与边不垂直的现象。所以本发明还可以进一步对于每一房间的轮廓进行直角多边形拟合，得到更合理的房间平面轮廓。

- 30 通过上述的专门针对单对象进行的优化处理，例如单对象融合和/或直角多边形拟合等等，可以得到更准确的单对象平面轮廓，利于后续的 2D 和 3D 模型的生成，提高模型的分辨率和准确度。

请注意，该步骤并非三维对象的二维或三维建模必须的步骤，而是可以提

高模型的准确度的优选处理方式。

如图 1A 所示，在步骤 S140，可以利用相机位置，对经过尺度归一化的各全景图像的在三维空间中的平面轮廓的三维点坐标进行旋转和平移操作，以将这些三维点坐标统一到同一个坐标系中，从而将各三维对象的平面轮廓拼接成多对象平面轮廓。

在本步骤 S140 中，要实现经过尺度归一化的各三维对象的平面轮廓的拼接，以将其拼接成多对象平面轮廓，可以采用手动的方式来实现，另外，也可以采用自动的方式来实现多房间拼接。在本发明中采用自动的方式来实现多对象拼接。

例如，在多对象拼接步骤 S140，还能够基于各单个三维对象的在三维空间中的平面轮廓，拼接得到在三维空间中的多对象平面轮廓。

需要说明的是，一个三维对象（例如，室内场景的房间）的平面轮廓指的是该三维对象的平面轮廓，该三维对象的平面轮廓可以是由多张全景图像的平面轮廓拼接而成。拼接过程包括对经过尺度归一化的属于该三维对象的全景图像的平面轮廓的三维点坐标进行旋转和平移操作之后，统一到同一个坐标系中，从而得到拼接后的该对象的平面轮廓。如图 1C 所示，将图 1B 所示的针对同一个三维对象对应的不同位置（即 A 位置、B 位置和 C 位置）的三张全景图像，对应的三个平面轮廓进行拼接，则得到该三维对象的在三维空间中的整体的平面轮廓，该三维对象的整体的平面轮廓也为二维轮廓。然后，将多个三维对象的平面轮廓进行拼接（例如，对各个三维对象的平面轮廓的三维点坐标进行尺度归一化、旋转和平移、拼接等操作），从而得到多对象平面轮廓。

以下将给出一个本发明提出的自动化多对象拼接方案。

下面将以一个房间作为一个三维对象为例，详细描述其具体操作。

假设有 N 个房间的轮廓，第 n 个房间轮廓的第 p 个三维点表示为 X_p^n ，该房间的相机位置表示为 $\{R_n, t_n\}$ ，其中， R_n 为用于表示相机位置的旋转参数的旋转矩阵， t_n 为用于表示相机位置的平移参数的平移向量。

此时，可以选定第一个房间的相机位置作为参考坐标系，这是因为目前得到的房间轮廓是在各自坐标系下的轮廓位置，需要统一到一个坐标系下，所以需要选定一个参考坐标系。准确地说，可以选定第一个房间的相机位置所在的坐标系作为参考坐标系。然后，可通过如下公式 3 将其它房间的轮廓三维点统

一到该坐标系下：

$$X_p^{n \rightarrow 1} = R_1 R_n^{-1} (X_p^n - t_n) + t_1 \quad (3)。$$

将除第一个房间外的所有的经过尺度归一化的轮廓三维点（例如，天花板边缘、墙面边缘、地板边缘上三维点）通过公式 3 进行转换，即可将所有房间的三维点统一到同一坐标系（即，第一个房间的参考坐标系）下，由此便能够实现多房间平面轮廓的拼接。

这里，可以选任意一个房间的坐标系作为参考坐标系（比如第 i 个房间的坐标系作为参考坐标系，此时公式 3 中的 $n \rightarrow 1$ 变为 $n \rightarrow i$ ， R_1 变为 R_i ， t_1 变为 t_i ），本发明对此不作任何限制，因为本发明中需要的是相对位置关系，不是绝对位置关系。例如，当选定第 i 个平面轮廓对应的全景图像被拍摄时的相机位置作为参考坐标系的情况下，可通过如下的公式将其它平面轮廓的三维点统一到参考坐标系下：

$$X_p^{n \rightarrow i} = R_i R_n^{-1} (X_p^n - t_n) + t_i \quad ,$$

将除第 i 个平面轮廓外的所有的经过尺度归一化的平面轮廓的三维点通过上述公式进行转换，以将所有平面轮廓的三维点统一到同一坐标系下，从而将各三维对象的平面轮廓拼接成多对象平面轮廓。

这里，在该步骤的多对象拼接之后得到的多对象平面轮廓，可以作为所述至少一个（包括多个）三维对象的 2D 模型（例如 2D 户型图）输出。

根据本发明的另一个示例性实施例，可选地，本发明的三维对象建模方法还可以在多对象轮廓优化步骤 S145，对多对象轮廓进行修正，即优化多对象轮廓。例如，计算多对象轮廓中，两个单对象轮廓的相邻边之间的距离，若距离非零且小于特定阈值，则将两个单对象轮廓进行偏移，以使相邻边的距离变为零。

请注意，本步骤也并非三维对象的二维或三维建模必须的步骤，而是可以提高模型的准确度的优选处理方式。

本发明中，在对多对象轮廓进行拼接之后，还可以进一步对多对象轮廓进行修正，以得到更加准确的多对象轮廓。

以一个房间作为被关注类型的一个三维对象为例，由于受到单图像平面轮廓提取精度以及相机位置提取精度的影响，相邻多维对象（比如一套室内房屋）的轮廓在拼接后可能存在重叠区域或者出现缝隙的情况，因此，可以进一步地针对这两种情况进行轮廓的校正。

校正方法例如可以如下。首先，计算两两轮廓的相邻边（这些相邻边理论上应当是重合的，即，理论上应当作为多房间轮廓的一个重合边）之间的距离，若该距离小于某个阈值，则可以判定这两条边是相邻的关系，此时可以将该轮廓进行相应的偏移，使相邻边的距离变为 0（即使得相邻边变得彼此重合，成为重合边），由此修正相邻边之间的重叠或者缝隙。

对于上述的该阈值，举一个例子来说，比如可以计算这些应当作为一个重合边的相邻边的平均长度 L ，并以该平均长度的某个比例作为阈值，比如可以使 $0.2 * L$ 作为该距离阈值。

请注意，上面仅仅是为了便于理解起见而给出的例示的阈值，实际上，本发明对于该阈值的不作额外的限制，其可以通过实验和经验来确定。

由此，经过上述单房间轮廓融合和多房间轮廓修正后的多房间轮廓可作为该套房屋的一个完整而准确的 2D 户型图（房屋的 2D 模型）。

进一步地，在步骤 S150，3D 模型生成，即将生成的多对象平面轮廓转化为至少一个三维对象的 3D 模型。在多对象拼接步骤之后，将拼接得到的在三维空间中的多对象平面轮廓转化为多对象 3D 模型。

例如，3D 模型生成步骤包括：

- 1) 将拼接得到的多对象平面轮廓中的顶部平面轮廓在内部进行三维点插值，并将得到的各顶部平面轮廓上的所有三维点坐标投影到相应的全景图像坐标系中，以获取顶部纹理；
- 2) 将拼接得到的多对象平面轮廓中的底部平面轮廓在内部进行三维点插值，并将得到的各底部平面轮廓上的所有三维点坐标投影到相应的全景图像坐标系中，以获取底部纹理；
- 3) 连接顶部轮廓和底部轮廓之间相同平面位置上的三维顶点，构成支撑部的平面轮廓，并对于支撑部的平面轮廓内部进行三维点插值，并将得到的各支撑部的平面轮廓的所有三维点坐标投影到相应的全景图像坐标系中，以便获取支撑部纹理；
- 4) 基于顶部纹理、底部纹理、支撑部纹理，生成整个三维对象的 3D 纹理模型。

另外，可选地，在 3D 模型生成步骤 S150 中，将得到的各三维对象的顶部平面轮廓上的所有三维点坐标中的作为特定类别的轮廓的提取的轮廓高度 h_c' 的、相机距离相应三维对象顶部的提取高度的高度值，替换为相机距离相

应三维对象底部的提取高度 h_f' ，并且使得到的各三维对象顶部平面轮廓上的所有三维点坐标中的长度和宽度值保持不变，得到相应的各三维对象的底部平面轮廓。

例如，上述的“相机距离相应三维对象顶部的特定类别的轮廓的提取的轮廓高度 h_c' ”通过如下方式得到：将在获取相机位置时得到的至少一幅全景图像上的所有三维点坐标中的高度值从小到大排序，取排序靠前的高度值的中值或者均值作为相机距离相应三维对象顶部的特定类别的轮廓的提取的轮廓高度 h_c' 。

另外，类似地，上述的“相机距离相应三维对象底部的提取高度 h_f' ”通过如下方式得到：将在获取相机位置时得到的至少一幅全景图像上的所有三维点坐标中的高度值从小到大排序，取排序靠后的高度值的中值或者均值作为相机距离相应三维对象底部的提取高度 h_f' 。

为了便于理解和描述，下面将以房屋建模为例来说明。

对于在之前的步骤中得到的多对象平面轮廓（例如多房间平面轮廓），在内部进行三维点插值，然后将所有的三维点坐标投影到相应的全景图像中，以便获取天花板纹理（颜色值）。

这里，将对三维点插值的方法进行举例说明。比如，假设得到的多房间平面轮廓的天花板轮廓是个矩形，假设其长为 H 、宽为 W ，则可以将其长与宽分别划分为 N 个间隔，可得到总共 $N*N$ 个插值点。然后，可选择该矩形的某个顶点（假设该顶点的三维点坐标为 (x,y,z) ）作为原点，则可以得到这 $N*N$ 个点的坐标依次为 $(x+H/N,y,z), (x+2*H/N,y,z), \dots (x,y+W/N,z), (x,y+2*W/N,z), \dots (x+H/N,y+W/N,z) \dots$ 。由此可见，在三维点插值之后，得到轮廓内部的密集三维点坐标。

请注意，上面是为了便于理解起见而给出了一个三维点插值的具体例子，实际上，本发明可用的三维点插值的方法可以有很多，并不限于该例子。

另外，例如，具体的投影方法可以如下。假定插值后的三维点坐标为 (x_i, y_i, z_i) ，其投影到全景图像上的经纬度为 (θ_i, ϕ_i) ，则该投影公式可以用如下的公式 4 表示：

$$\begin{cases} \theta_i = \arctan(x_i, z_i) \\ \phi_i = \arctan\left(-y_i, \sqrt{x_i^2 + z_i^2}\right) \end{cases} \quad (4)。$$

由该公式得到经纬度后,再根据公式 1 即可得到该三维点在全景图像平面的坐标,该点处的颜色值即可作为该三维点的纹理。

对于大多数场景来说,天花板的轮廓与地板的轮廓可假定为平行且相同。因此,利用上述获取的校正后的每个房间的天花板平面轮廓,再加上在上面得到的相机距离地板的提取高度 h_f' ,同样通过公式 2 即可生成多房间地板平面轮廓的三维点。

这里,地板的平面轮廓的形状假定是跟天花板一样的,也就是水平面的三维坐标 x 和 z 是一样的,不同之处在于高度,即垂直方向的 y 值(比如天花板的平面轮廓是在相机的上方,地板是在相机的下方,所以高度不同)。因此,只需将在上述得到的天花板轮廓的三维点坐标中的 y 值(相机距离天花板的提取高度 h_c')替换为相机距离地板的提取高度 h_f' 即可。

与天花板平面轮廓的三维点插值类似地,对于地板平面轮廓,在内部进行三维点插值,然后利用公式 4 投影到相应全景图像中,以便获取地板纹理。

然后,连接天花板轮廓和地板轮廓之间相同平面位置上的三维顶点,可构成多个墙面的平面轮廓,同样地,对于这些平面轮廓内部进行三维点插值,然后利用公式 4 投影到相应全景图像中,以便获取墙面的纹理。

由此,可生成完整房屋的 3D 纹理模型。

通过本发明的上述三维对象建模方法,能够有效提高所生成的模型的分辨率和准确性。

而且,需要指出,上面这些是为了便于理解和描述起见,以房屋建模为例对本发明的基于图像来进行建模的方法进行了说明,实际上,本发明不应仅限于房屋建模的应用场景,而是能够适用于各种基于图像来进行建模的场景,例如对车辆建模以便实现 VR(虚拟现实)看车等等场景,本发明实际上是提供了一种创新的综合性的图像处理方案。

图 3 给出了根据本发明的一个示例性实施例的三维对象建模设备的示意性框图。

如图 3 所示,根据本发明的一个示例性实施例的三维对象建模设备 100 可以包括平面轮廓提取装置 110、尺度归一化装置 120、多轮廓拼接装置 130。

其中,平面轮廓提取装置 110 可以被配置用于针对为待处理的至少一个三维对象的每一个三维对象拍摄的至少一幅全景图像,提取每一个三维对象的至少一幅全景图像的在三维空间中的平面轮廓。

这里，如上所述，至少一幅全景图像是针对一个三维对象拍摄的，每个三维对象对应于一幅或多幅全景图像。

尺度归一化装置 120 可以被配置用于将相机位置的尺度与每一个三维对象的至少一幅全景图像的在三维空间中的平面轮廓的尺度进行归一化，
5 得到经过归一化的各全景图像的在三维空间中的平面轮廓。

多对象拼接装置 130 可以被配置用于基于相机位置，对经过尺度归一化的各全景图像的在三维空间中的平面轮廓的三维点坐标进行旋转和平移操作，以将这些三维点坐标统一到同一个坐标系中，从而将各全景图像的在三维空间中的平面轮廓拼接成多对象平面轮廓。

10 其中，可选地，上述的相机位置可以是预先确定的。

或者，上述的相机位置也可以是通过以下方式得到的：

- 1) 利用为待处理的至少一个三维对象的每一个三维对象拍摄的至少两幅全景图像的几何关系，进行这些全景图像之间的特征点匹配，并记录这些全景图像中互相匹配的特征点作为匹配特征点；以及
- 15 2) 通过对于每一个三维对象的至少两幅全景图像中的每幅全景图像，减小每幅全景图像上的匹配特征点的重投影误差，来得到拍摄每幅全景图像时的全景相机的相机位置、以及每幅全景图像上的匹配特征点的三维点坐标。

这里，在相机位置是提取得到的情况下，可以将获取相机位置时得到的每一个三维对象的至少一幅全景图像上的所有三维点坐标中的高度值从小到大排序，取排序靠前的高度值的中值或者均值作为特定类别的轮廓的提取的轮廓高度 h_c' ；以及利用特定类别的轮廓的假定轮廓高度 h_c 与特定类别的轮廓的提取的轮廓高度 h_c' 的比值，从每一个三维对象的至少两幅全景图像的在三维空间中的平面轮廓生成经过归一化的每一个三维对象的至少两幅全景图像的
20 在三维空间中的平面轮廓。

其中，这里的相机假定高度 h_c （即，相机距离天花板的假定高度）是任意假定的一个高度。

另外，在相机位置是预先确定的相机确定高度 h_c' 的情况下，可以利用在平面轮廓提取步骤 S110 中假定的相机假定高度 h_c 与相机确定高度 h_c' 的比值，从所述每一个三维对象的至少两幅全景图像的在三维空间中的平面轮廓生成经过归一化的各全景图像的在三维空间中的平面轮廓。同样，这里的相
30

机假定高度 h_c (即, 相机距离天花板的假定高度) 也是任意假定的一个高度。需要说明的是, 在本申请中, 相机假定高度和特定类别的轮廓的假定轮廓高度是相同的, 类似地, 相机确定高度和特定类别的轮廓的提取的轮廓高度是相同的。

- 5 可选地, 该三维对象建模设备 100 还包括单对象平面轮廓生成装置 125, 其被配置用于:

对于至少两幅全景图像, 通过以下方式、逐个地确定其中是否有多幅全景图像属于同一个三维对象: 如果两幅全景图像之间有多于特定比例的匹配特征点, 则确定这两幅全景图像属于同一个三维对象; 以及

- 10 如果确定多幅全景图像属于同一个三维对象, 则取该多幅全景图像的在三维空间中的平面轮廓的并集, 作为该三维对象的在三维空间中的平面轮廓。

另外, 可选地, 所述多对象拼接装置 130 还被配置用于能够基于各单个三维对象的在三维空间中的平面轮廓, 拼接得到在三维空间中的多对象平面轮廓。

- 15 另外, 可选地, 所述多对象拼接装置 130 可以通过如下的坐标变换, 来实现多平面轮廓的拼接:

- 假设所有的全景图像的在三维空间的平面轮廓共有 N 个, 第 n 个平面轮廓的第 p 个三维点表示为 X_p^n , 该第 n 个平面轮廓对应的全景图像被拍摄时的相机位置表示为 $\{R_n, t_n\}$, 其中, R_n 为用于表示该相机位置的旋转参数的旋转矩阵, t_n 为用于表示该相机位置的平移参数的平移向量, 其中 N 为大于 1 的整数, n 为大于等于 1 的整数。此时, 可以选定第 i 个平面轮廓对应的全景图像被拍摄时的相机位置作为参考坐标系, 可通过如下的公式将其它平面轮廓的三维点统一到该坐标系下:

$$X_p^{n-i} = R_i R_n^T (X_p^n - t_n) + t_i, \quad ,$$

- 25 将除第 i 个平面轮廓外的所有的经过尺度归一化的平面轮廓的三维点通过该公式进行转换, 以将所有平面轮廓的三维点统一到同一坐标系下, 从而将各全景图像的在三维空间中的平面轮廓拼接成多对象平面轮廓, 即将各三维对象的平面轮廓拼接成多对象平面轮廓。

- 30 另外, 可选地, 该三维对象建模设备 100 还可以包括多对象轮廓优化装置 135, 其被配置用于计算多对象平面轮廓中, 两个单对象平面轮廓的相邻边之间的距离, 若该距离非零且小于特定阈值, 则将该两个单对象平面轮廓进行

偏移，以使该相邻边的距离变为零。

另外，可选地，该三维对象建模设备 100 还可以包括 3D 模型生成装置 140，其被配置用于将拼接得到的在三维空间中的多对象平面轮廓转化为多对象 3D 模型。

5 这里，上述的三维对象建模设备 100 的各装置 110、120、125、140、145、150 等分别与在上面详细描述的步骤 S110、120、125、140、145、150 等相对应，在此不再赘述。

由此，通过本发明的上述三维对象建模设备，能够有效提高所生成的模型的分辨率和准确性。

10 而且，需要指出，上面这些是为了便于理解和描述起见，以三维对象（例如，房屋）建模为例对本发明的基于图像来进行建模的技术方案进行了说明，实际上，本发明不应仅限于三维对象（例如，房屋）建模的应用场景，而是能够适用于各种基于图像来进行三维对象建模的场景，例如，车辆建模等等，本发明实际上是提供了一种创新的综合性的图像处理方案。

15 图 4A 给出了根据本发明的一个示例性实施例的图像处理方法的示意性流程图。

下面将先参照图 4A 来介绍根据本发明的一个示例性实施例的图像处理方法，来为后续的建模处理做充分的数据准备，简化后续的处理过程。如图 4A 所示，图像处理过程包括相机位置获取 S410、单图像平面轮廓生成
20 S420、尺度归一化 S430 三个步骤，建模过程可以包括多个后续步骤。

根据本发明的一个示例性实施例的图像处理方法可以包括以下步骤。

在步骤 S410，利用针对为待处理的至少一个三维对象的每一个三维对象的拍摄至少两幅全景图像的几何关系，获取在拍摄每一个三维对象的至少两幅全景图像时的全景相机的位置、以及每一个三维对象的至少两幅全景图像
25 上的匹配特征点的三维点坐标。

例如，至少两幅全景图像是针对一个三维对象拍摄的，每个三维对象对应于一幅或多幅全景图像。

在步骤 S420，对于每一个三维对象的至少两幅全景图像的每幅全景图像，基于全景图像上的、轮廓特征属于特定类别的像素点中的边缘像素点所围
30 成的轮廓，生成全景图像的在三维空间中的平面轮廓。

由此，在本发明中，可以基于全景图像，自动化地得到该图像的平面轮廓，

而无需人工参与制作，并且无需使用昂贵的 3D 扫描设备。

在步骤 S430，将所获取的在拍摄每一个三维对象的至少两幅全景图像时的全景相机的位置的尺度和每一个三维对象的至少两幅全景图像的在三维空间中的平面轮廓的尺度进行归一化，得到经过归一化的各全景图像的在三维空间中的平面轮廓。

这里，可选地，步骤 S410 可以参考上述三维对象建模方法的实施例关于相机位置获取步骤的描述，步骤 S420 可以参考上述三维对象建模方法的实施例关于单图像平面轮廓生成步骤的描述，步骤 S430 可以参考上述三维对象建模方法的实施例关于尺度归一化步骤的描述，重复之处不再赘述。

经过上述的图像处理过程，为后续模型生成提供了高分辨率的基础。而且，通过上述的图像处理过程，为后续模型生成提供在三维空间中的并且经过归一化的各全景图像的平面轮廓，简化了后续模型生成工作，减少了处理时间，提高了处理效率。

经过上述的图像处理，为对象建模提供了模型生成所需的平面轮廓数据。

图 4B 给出了根据本发明的一个示例性实施例的图像处理与对象建模的总体过程的示意性流程图。

图 4B 包括上面描述的图像处理部分与接下来将描述的对象建模部分。

下面，将参照图 4B 的对象建模部分来描述根据本发明的一个示例性实施例的三维对象建模方法。

根据本发明的一个示例性实施例的对象建模过程可以包括如下步骤。

在图像处理步骤，使用如上所述的任意一种图像处理方法，对每一个三维对象的至少一幅全景图像进行图像处理，以得到经过归一化的各全景图像的在三维空间中的平面轮廓。

如图 4B 所示，在多对象拼接步骤 S440，基于上述的经过归一化的各全景图像的在三维空间中的平面轮廓，拼接得到多对象平面轮廓。

另外，可选地，上述的对象建模方法还可以包括单对象平面轮廓生成步骤 S435，用于基于上述图像处理步骤中得到的、经过归一化的各全景图像的平面轮廓，得到各单个三维对象的在三维空间中的平面轮廓。单对象平面轮廓生成步骤 S435 可以参考上述三维对象建模方法的实施例关于单对象平面轮廓生成步骤 S125 的描述，重复之处不再赘述。

另外, 可选地, 在多对象拼接步骤 S440, 还能够基于各单个三维对象的在三维空间中的平面轮廓, 拼接得到在三维空间中的多对象平面轮廓。

即, 在本发明中, 既可以直接基于各全景图像的在三维空间中的平面轮廓, 来拼接得到在三维空间中的多对象平面轮廓; 也可以先得到单个三维对象的在三维空间中的平面轮廓, 然后基于各单个三维对象的在三维空间中的平面轮廓, 来拼接得到在三维空间中的多对象平面轮廓。

另外, 可选地, 上述的述对象建模方法还可以包括 3D 模型生成步骤 S450, 用于在多对象拼接步骤 S440 之后, 将拼接得到的在三维空间中的多对象平面轮廓转化为对象 3D 模型。3D 模型生成步骤 S450 可以参考上述三维对象建模方法的实施例关于 3D 模型生成步骤 S150 的描述, 重复之处不再赘述。

下面, 为了便于理解和描述, 将以房屋图像处理和房屋建模为例, 来进一步地详细描述本发明的各个处理过程。

在这里给出的示例中, 上述的房屋可以包括多个房间, 每个房间可以视为一个三维对象, 房屋的建模可以视为多个三维对象的建模。例如, 可以通过拼接房屋里的作为多个三维对象的多个房间, 来构造房屋的整体模型。

对于上述的总体图像处理和对象建模方法中的术语, 例如, 其中的“顶部”在本示例中可以是房间的“天花板”, “底部”在本例中可以是房间的“地板”, “支撑部”在本例中可以是“墙面”。另外, “轮廓特征属于特定类别的像素点中的边缘像素点所围成的轮廓”在本例中可以指“由属于天花板的像素点中的边缘像素点所围成的天花板轮廓”。“特定类别的轮廓的提取的轮廓高度”在本例中可以是“相机到天花板的提取高度”, 类似地, “特定类别的轮廓的假定轮廓高度”在本例中可以是“相机到天花板的假定高度”。

在根据本发明的一个示例性实施例的图像处理方法中, 基于在房屋内拍摄的至少一幅全景图像(一个全景图像只对应一个房间, 但是一个房间内可能拍摄多个全景图像, 即一个房间可能对应多个全景图像), 获取拍摄这些图像的全景相机的位置, 再基于获取的相机位置, 来提取全景图像的平面轮廓, 然后对提取的平面轮廓进行归一化, 以得到建模所需平面轮廓。

因此, 如图 4A 所示, 在步骤 S410, 利用在一个房间内拍摄的至少两幅全景图像的几何关系, 获取拍摄这些全景图像的全景相机的位置。

在本发明中, 可选地, 可以采用基于多视图几何的方法来解决这一问题。具体地, 步骤 S410 例如可以包括以下操作:

1) 进行这些全景图像之间的特征点匹配, 并记录这些图像中互相匹配的特征点; 以及

2) 通过对于每张全景图像, 减小该全景图像上的匹配特征点的重投影误差, 求出每张全景图像的相机位置和该全景图像上的匹配特征点的三维点坐标。

对于上述的步骤 1), 其中, 在图像处理技术中, 图像特征点指的是图像灰度值发生剧烈变化的点或者在图像边缘上曲率较大的点(即两个边缘的交点)。图像特征点能够反映图像本质特征, 能够标识图像中的目标物体。

如何高效且准确地匹配出两个不同视角的图像中的同一个物体, 是许多计算机视觉应用中的第一步。虽然图像在计算机中是以灰度矩阵的形式存在的, 但是利用图像的灰度并不能准确地找出两幅图像中的同一个物体。这是由于灰度受光照的影响, 并且当图像视角变化后, 同一个物体的灰度值也会随之变化。所以, 需要找出一种能够在相机进行移动和旋转(视角发生变化)时, 仍然能够保持不变的特征, 利用这些不变的特征来找出不同视角的图像中的同一个物体。

因此, 为了能够更好地进行图像匹配, 需要在图像中选择具有代表性的区域, 例如: 图像中的角点、边缘和一些区块。其中, 角点的辨识度最高。在很多计算机视觉处理中, 通常提取角点作为特征点, 对图像进行匹配, 可以使用的方法例如有 SFM (Structure Form Motion, 运动恢复结构)、SLAM (Simultaneous Localization and Mapping, 视觉即时定位与地图构建) 等。

但是, 单纯的角点并不能很好地满足需求, 例如: 相机从远处得到的是角点, 但是在近处就可能不是角点; 或者, 当相机旋转后, 角点就发生了变化。为此, 计算机视觉的研究者们设计了许多更为稳定的特征点, 这些特征点不会随着相机的移动、旋转或者光照的变化而变化, 可以使用的方法例如有 SIFT (Scale-Invariant Feature Transform, 尺度不变特征变换)、SURF(Speed Up Robust Features, 加速稳健特征)等。

一个图像的特征点由两部分构成: 关键点 (Keypoint) 和描述子 (Descriptor)。关键点指的是该特征点在图像中的位置, 有些还具有方向、尺度信息; 描述子通常是一个向量, 描述关键点周围像素的信息。通常, 在匹配时, 只要两个特征点的描述子在向量空间的距离相近, 就可以认为它们是同一个特征点。

特征点的匹配通常需要以下三个步骤：1) 提取图像中的关键点；2) 根据得到的关键点位置，计算特征点的描述子；3) 根据特征点的描述子，进行匹配。

5 可选地，本步骤中的特征点匹配的相关处理可以使用例如开源计算机视觉库 OpenCV 来实现。为了简洁和不模糊本发明的主题，本文对于这部分的更详细的处理过程不再赘述。

在进行这些全景图像之间的特征点匹配之后，要记录这些全景图像中互相匹配的特征点（也称为“匹配特征点”），可以例如按照如下的方式进行匹配特征点的记录。

10 比如，图像 1 上的特征点 a 与图像 2 上的特征点 b 是匹配的，图像 2 上的特征点 b 又与图像 3 上的特征点 c 是匹配的，图像 3 上的特征点 c 又与图像 4 上的特征点 d 也是匹配的，则可以记录一条特征点匹配数据 (a, b, c, d)（也可称为“特征点跟踪轨迹”）。由此，得到输入的这些全景图像中互相匹配的特征点有关的记录数据。

15 对于上述的步骤 2)，其中，图像重投影是指通过对任意视点的参考图像进行投影来产生新的图像，即，图像重投影能够改变已生成的图像的视线方向。

具体而言，在本发明中，图像重投影指的是将图像 1 上一个特征点 p1 对应的三维点坐标，通过当前相机参数投影到另外一幅图像 2 中，在这张
20 图像 2 上的得到的投影点 q2 与图像 1 上的特征点 p1 在该图像 2 中的匹配特征点 p2 之间的位置差，构成重投影误差 (Reprojection Error)。其中，图像 2 中的匹配特征点 p2 是实际的位置，而经过重投影得到的投影点 q2 是估计的位置，通过使得投影点 q2 与匹配特征点 p2 之间的位置差尽量最小化，即，使得投影点 q2 与匹配特征点 p2 尽可能地重合，求解出相机位置。

25 其中，优化（减小）重投影误差的目标函数中包含的变量包括相机位置和特征点的三维坐标，在逐步减小（优化）重投影误差的过程中，得到相机位置和特征点的三维坐标。

可选地，在本发明中，可以通过结合梯度下降算法和德罗内三角算法 (Delaunay Triangulation)，来减小重投影误差，达到优化的目的。

30 其中，在使用梯度下降算法时，将匹配特征点的三维点坐标当作常量，将相机位置当作变量，反之，在使用德罗内三角算法时，将匹配特征点的三

维点坐标当作变量，将相机位置当作常量。

可选地，在本发明中，可以使用渐进式求解，来提高求解的相机位置和三维点坐标的精度，即，在求解过程中，通过每次增加一幅图像的方式来求解它的相机位置和匹配特征点的三维点坐标。其中，渐进式求解的方法包括
5 例如 ISFM (Incremental SFM)。

另外，进一步可选地，可以采用集束优化 (bundle adjustment) 来进一步减小重投影误差。具体地，可以在每张全景图像都执行完减小重投影误差以获得相机位置和三维点坐标的处理之后，最后统一使用集束优化来同时优化所有的相机位置以及所有的三维点坐标。也可以在减小重投影误差以
10 获得相机位置和三维点坐标的处理过程中，在针对任意张全景图像的相机位置和三维点坐标获取之后，添加集束优化的处理，以对获取的相机位置和三维点坐标进行优化。

这里，集束优化指的是同时优化所有相机的位置以及所有的三维点坐标，区别于渐进式求解中分别只优化当前相机位置以及当前图像上的三维
15 点坐标的方式。

另外，除了上面描述的渐进式求解之外，还可以采用全局式求解的方法。

在步骤 S420，基于每一个三维对象的至少两幅全景图像的每幅全景图像上的属于天花板的边缘像素点所围成的轮廓，生成每一个三维对象的至少两幅全景图像的每幅全景图像在三维空间中的平面轮廓。需要说明的是，
20 关于确定边缘像素点的说明可以参考上述实施例中的相关描述。

在步骤 S430，将在步骤 S410 得到的拍摄每一个三维对象的至少两幅全景图像时的相机位置的尺度和在步骤 S420 得到的每一个三维对象的至少两幅全景图像的三维空间平面轮廓的尺度进行归一化。

一方面，由于步骤 S410 中相机位置获取中的尺度不确定性，因此无法确定相机到天花板轮廓的真实高度。另一方面，步骤 S420 中得到房间的三维空间平面轮廓时采用的是相机距离天花板的假定高度 h_c ，因此，得到的相机位置的尺度和房间的三维空间平面轮廓的尺度不统一，给后续的房间轮廓拼接造成一定的困难。

在本步骤，将拍摄每两幅全景图像时的相机位置的尺度和各全景图像的
30 在三维空间中的平面轮廓的尺度进行归一化，以使得能够执行后续的多对象拼接处理。

可选地，在本步骤中，将在步骤 S410 得到的所有三维点的高度坐标值(Y轴上的值)进行从小到大排序，取排序靠前的 a 个高度坐标值的中值(或者均值等等其它合理的取值方式)作为相机距离天花板的提取高度 h_c' 。

最后，利用相机距离天花板的假定高度 h_c 与提取高度 h_c' 的比值，重新生成尺度归一化的单房间平面轮廓。

例如，可以通过将假定高度 h_c 与提取高度 h_c' 的比值与在步骤 S420 得到的平面轮廓上的边界点的坐标相乘，得到尺度归一化的平面轮廓上的边界点的坐标，由此得到尺度归一化的平面轮廓。

另一方面，也可以取排序靠后的 b 个高度坐标值的中值(或者均值等等其它合理的取值方式)作为相机距离地板的提取高度 h_f' (该提取高度将在后续的 3D 模型生成步骤等中被使用)。

经过上述的图像处理为建模提供了模型生成所需的平面轮廓数据，接下来将继续举例来描述房屋建模方法。

可选地，如图 4B 所示，在步骤 S435，可以基于经过归一化的各全景图像的在三维空间中的平面轮廓，得到各单个房间(即单个对象)的在三维空间中的平面轮廓。关于步骤 S435 的说明可以参考上述步骤 S125 的相关描述。

在步骤 S440，基于在步骤 S410 获取的相机位置以及在步骤 S430 得到的经过尺度归一化的各房间平面轮廓，拼接多个房间(即多个对象)的轮廓。

可选地，在本步骤 S440 中，可以利用获取的相机位置，对经过尺度归一化的每个房间轮廓的三维点坐标进行旋转和平移操作，以将各房间的三维点坐标统一到同一个坐标系中，从而实现多房间平面轮廓的拼接。

可选地，如图 4B 所示，上述的对象建模方法还可以包括轮廓优化步骤 S445，即对多对象(例如，多房间)的轮廓进行修正。轮廓优化步骤 S445 可以参考上述三维对象建模方法的实施例关于多对象轮廓优化步骤 S145 的描述，重复之处不再赘述。

可选地，还可以进一步在步骤 S450，将生成的多房间平面轮廓转化为房屋 3D 模型。

首先对于在之前的步骤中得到的多对象平面轮廓(例如多房间平面轮廓)中的特定类别平面轮廓(例如房间的天花板平面轮廓)，在内部进行三维点插值，然后将所有的三维点坐标投影到相应的全景图像中，以便获取天花板纹理(颜色值)。关于三维点插值的方式可以参考上述关于三维点插值的相

关说明。

然后，连接天花板轮廓和地板轮廓之间相同平面位置上的三维顶点，可构成多个墙面的平面轮廓，同样地，对于这些平面轮廓内部进行三维点插值，然后利用上述的公式 4 投影到相应全景图像中，以便获取墙面的纹理。

5 由此，可生成完整房屋的 3D 纹理模型。

图 4C 给出了根据本发明的一个示例性实施例的图像处理设备的示意性框图。

如图 4C 所示，根据本发明的一个示例性实施例的图像处理设备 400 可以包括相机位置获取装置 410、单图像平面轮廓生成装置 420、尺度归一化
10 装置 430。

相机位置获取装置 410 可以被配置用于利用针对为待处理的至少一个三维对象的每一个三维对象拍摄的至少两幅全景图像的几何关系，获取在拍摄每一个三维对象的至少两幅全景图像时的全景相机的位置、以及每一个三维对象的至少两幅全景图像上的匹配特征点的三维点坐标。

15 单图像平面轮廓生成装置 420 可以被配置用于对于每一个三维对象的至少两幅全景图像的每幅全景图像，基于全景图像上的、轮廓特征属于特定类别的像素点中的边缘像素点所围成的轮廓，生成全景图像的在三维空间中的平面轮廓。

例如，特定类别至少包括全景图像中的三维对象的顶部、底部、支撑部
20 等等。

尺度归一化装置 430 可以被配置用于将所获取的在拍摄每一个三维对象的至少两幅全景图像时的全景相机的位置的尺度和每一个三维对象的至少两幅全景图像的在三维空间中的平面轮廓的尺度进行归一化，得到经过归一化的各全景图像的在三维空间中的平面轮廓。

25 可选地，单图像平面轮廓生成装置 420 还可以被配置用于基于每一个三维对象的至少两幅全景图像中的每幅全景图像上的像素点之间的特征相似度，确定每一个三维对象的至少两幅全景图像中的每幅全景图像上的、轮廓特征属于特定类别的像素点中的边缘像素。

其中，两个像素点的特征相似度为这两个像素点的特征之差的绝对值。
30 这里，像素点的特征可以包括灰度、颜色等等。

可选地，尺度归一化装置 430 还可以被配置用于将由相机位置获取装

置得到的每一个三维对象的至少一幅全景图像上的所有三维点坐标中的高度值,从小到大排序,取排序靠前的高度值的中值或者均值作为用作特定类别的轮廓的提取的轮廓高度 h_c' 的、相机距离相应三维对象顶部的提取高度 h_c' ; 以及利用所述相机距离相应三维对象顶部的假定高度 h_c 与所述相机距离相应三维对象顶部的提取高度 h_c' 的比值,从每一个三维对象的至少两幅全景图像的在三维空间中的平面轮廓,生成经过归一化的每一个三维对象的至少两幅全景图像的平面轮廓。

相机距离相应三维对象顶部的假定高度 h_c 是任意假定的一个高度。

在本发明中,通过使用全景相机拍摄房间的全景图像,为三维对象建模(例如房屋建模等)提供高分辨率的采集图像。

进一步地,在本发明中,采用了高效的图像处理设备,为三维对象建模(例如房屋建模)提供了高分辨率的建模准备数据,而且,所提供的建模准备数据能够简化后续的模式生成过程。

图 4D 给出了根据本发明的一个示例性实施例的对象建模设备的示意性框图。

如图 4D 所示,该对象建模设备 4000 可以包括图 4C 所示的图像处理设备 400 以及多对象拼接装置 440。

图像处理设备 400 可以被配置用于处理每一个三维对象的至少一幅全景图像,生成经过归一化的各全景图像的在三维空间中的平面轮廓。

多对象拼接装置 440 可以被配置用于基于经过归一化的各全景图像的平面轮廓,拼接得到多对象平面轮廓。

可选地,该三维对象建模设备 4000 还可以包括:单对象平面轮廓生成装置 435,其可以被配置用于基于经过归一化的各全景图像的平面轮廓,得到各单个三维对象(例如,房间)的平面轮廓。

可选地,单对象平面轮廓生成装置 435 还可以被配置用于对于至少两幅全景图像,通过以下方式、逐个地确定是否有多幅全景图像属于同一个三维对象:如果两幅全景图像之间有多于特定比例的匹配特征点,则可以判定这两幅全景图像属于同一个三维对象;以及如果确定多幅全景图像属于同一个三维对象,则对于由多幅全景图像得到的同一个三维对象的各平面轮廓,取同一个三维对象的所有平面轮廓的并集,作为三维对象的在三维空间中的平面轮廓。

另外，可选地，多对象拼接装置 440 还可以被配置用于基于由单对象平面轮廓生成装置 435 生成的、各单个三维对象的在三维空间中的平面轮廓，拼接得到在三维空间中的多对象平面轮廓。

另外，可选地，三维对象建模设备 4000 还可以包括多对象轮廓优化装置 445，其可以被配置用于对多对象拼接装置 440 得到的多对象平面轮廓进行轮廓修正。

可选地，该三维对象建模设备 4000 还可以包括 3D 模型生成装置 450，其可以被配置用于将拼接得到的在三维空间中的多对象平面轮廓转化为三维对象 3D 模型。

这里，上述的对象建模设备 4000 的各装置 410、420、430、435、440、445、450 等分别与在上面描述的步骤 S410、420、430、435、440、445、450 等相对应，在此不再赘述。

由此，通过本发明的上述对象建模设备，能够有效提高所生成的模型的分辨率和准确性。

图 5A 给出了根据本发明的另一个示例性实施例的图像处理方法的示意性流程图。

下面，为了便于理解和描述，将以房屋图像处理和房屋建模作为本发明的图像处理方法和三维对象建模的实例，来详细描述本发明的各个处理过程。

如图 5A 所示，根据本发明的一个示例性实施例的图像处理方法可以包括相机位置获取 S510、单图像平面轮廓生成 S520、尺度归一化 S530 三个步骤，建模过程可以包括后续的多个步骤，但是请注意，后续的这些步骤并非都是必须的。

下面将先介绍上述的根据本发明的一个示例性实施例的图像处理方法，来为后续的建模处理做充分的数据准备，简化后续的处理过程。

如图 5A 所示，在步骤 S510，利用针对待处理的至少一个三维对象的每一个三维对象拍摄的至少两幅全景图像的几何关系，获取在拍摄每一个三维对象的至少两幅全景图像的每幅全景图像时的全景相机的相机位置、以及每幅全景图像上的匹配特征点的三维点坐标。

在本发明中，可选地，可以采用基于多视图几何的方法来解决这一问题。步骤 S510 可以参考上述实施例中关于步骤 S410 的描述，重复之处不

再赘述。

在步骤 S520, 对于每一个三维对象的至少两幅全景图像的每幅全景图像, 通过用于提取图像轮廓的深度学习模型, 提取每一个三维对象的至少两幅全景图像的每幅全景图像的在三维空间中的平面轮廓, 即对于每一个三维对象的至少两幅全景图像, 通过作为语义分割模型的深度学习模型来生成每一个三维对象的至少两幅全景图像的在三维空间中的平面轮廓。

在本步骤中, 可以通过例如图 5B 所示的下述方式来训练所述深度学习模型:

1) 在步骤 S521, 标注全景图像上的分类数据, 作为训练数据集, 即通过人工标注作为训练图像的多个全景图像上的被关注类型的三维对象的轮廓特征, 生成被关注类型的三维对象的轮廓特征训练数据集。

这里, 以室内房屋中的每个房间作为一个被关注类型的三维对象为例, 那么房间里的家具则不是被关注类型的三维对象。

另外, 上述的轮廓特征可以是比如体现三维对象的整体构成的轮廓或者说边界, 比如房间的天花板轮廓、地板轮廓等。

2) 在步骤 S522, 训练深度学习模型, 即利用被关注类型的三维对象的轮廓特征训练数据集来训练深度学习模型, 由此得到经过训练的深度学习模型。

例如, 深度学习模型的输出包含被关注类型的三维对象的轮廓特征类别。这里以房屋建模为例, 被关注类型的三维对象的轮廓特征类别可以是每个房间的天花板、墙面、地板这三类。

然后, 基于上述的语义分割模型(即深度学习模型), 生成各全景图像的在三维空间中的平面轮廓的步骤包括以下操作(参见图 5C 的步骤 S523~S525)。

1) 在步骤 S523, 利用深度学习模型划分全景图像的每个像素点所属的类别, 即将针对为待处理的至少一个三维对象的每一个三维对象拍摄的至少一幅全景图像输入到经过训练的深度学习模型中, 得到全景图像中每个像素点对应的轮廓特征的类别。

以房间作为被关注类型的三维对象为例, 被关注类型的三维对象的轮廓特征类别可以是每个房间的天花板、墙面、地板这三类。

2) 在步骤 S524, 提取轮廓特征属于特定类别的像素点中的边缘像素点,

即提取出每一个三维对象的至少两幅全景图像的每幅全景图像中，轮廓特征属于特定类别的像素点中的处于边缘的边缘像素点作为特定类别边缘像素点。

5 以房间作为被关注类型的三维对象为例，特定类别的像素点可以是天花板上的像素点，作为特定类别边缘像素点的处于边缘的那些特定类别的像素点可以是天花板边缘像素点。

3) 在步骤 S525，将特定类别边缘像素点投影到三维空间，生成各全景图像的在三维空间中的平面轮廓，即假定每一个三维对象的至少一幅全景图像上的所有特定类别边缘像素点具有同一高度 h_c ，作为特定类别的轮廓的假定轮廓高度，并将每幅全景图像上的特定类别边缘像素点投影到三维平面上，得到与每幅全景图像对应的特定类别三维点，然后基于与每幅全景图像对应的特定类别三维点，形成各全景图像的在所述三维空间中的平面轮廓。

15 例如，特定类别包括全景图像中三维对象的顶部。以房间作为被关注类型的三维对象为例，特定类别的轮廓的假定轮廓高度可以是相机距离天花板的高度。

下面将以室内房屋为例，详细描述本步骤的具体内容。

例如，对于绝大多数场景来说，房间的天花板都是平面，可以用来表示该房间的平面图，因此本发明中采用通过作为语义分割模型的深度学习模型来提取天花板轮廓的方法来得到全景图像的平面轮廓。

20 这里，语义分割指的是对于一幅图像，对其中的每个像素点进行类别的划分。因此，本发明的语义分割模型可以被看作用来对图像上的像素点进行类别划分的深度学习模型。

本领域技术人员可知，机器学习可分为浅层学习 (Shallow Learning) 和深度学习 (Deep Learning)。其中，浅层学习的隐层节点的层次通常较少，深度学习的隐层节点的层次通常多一些，比如，深度学习模型通常有 5 层、6 层，甚至可以有 10 多层或者更多层的隐层节点。

30 在语义分割模型中，通常需要事先定义像素点的类别。比如，举例来说，对于室内房屋场景，一般可将像素点的类别定义为天花板、地板、墙面、门、柜子、沙发等等。对于室外场景，例如可以将像素点的类别定义为例如天空、路面、树木、建筑等等。

传统的语义分割技术大多采用分类器加图模型的方法。常用的传统分

分类器包括支持向量机 (SVM)、随机森林 (Random Forest) 等分类算法。分类器的输入通常为人工设计的局部特征,常用的特征有 RGB、灰度、SIFT 等。分类器对图像中的每个像素点逐一进行类别判断。常用的图模型技术包括马尔可夫随机场 (Markov random fields)、条件随机场 (Conditional random fields), 其作用在于加强相邻像素点类别的一致性。

随着深度学习技术在语义分割中的应用,深度学习的方法已经大大超越传统的语义分割技术。

常用的语义分割的深度学习模型主要是基于 CNN (卷积神经网络) 的框架。由于语义分割需要输出每个像素点的类别(如果输入图像大小为 $H*W$, 则输出也为 $H*W$), 因此在传统 CNN 基础上, 需要引入上采样的方法来增大最终输出的分辨率(最简单的上采样方式为最近邻采样)。由此, 根据不同的上采样方式, 常见的语义分割模型有 DeepLab、UperNet、PSPNet 等等。

这些技术通过采集大量普通相机拍摄的图像, 对其中的每个像素点进行人工的语义标注, 比如室外场景标注为天空、路面、树木、建筑等。在训练深度网络时, 将这些样本送入深度语义分割模型, 输出得到估计的概率矩阵, 并采用交叉熵损失 (Cross Entropy Loss) 的目标函数减小估计值与真实标注值之间的误差, 直到最终误差不再变化, 则模型训练完成。

对于输入的待处理图像, 将其输入到训练好的深度学习模型, 得到输出的概率矩阵, 计算每个位置上的最大概率输出值所对应的维度, 以此作为该像素对应的类别值。比如, 输入的图像大小为 $H*W$, 模型输出的概率矩阵的大小为 $H*W*C$, 其中 C 表示类别的个数。图像中每个像素点都对应一个 C 维的概率向量 (概率向量之和为 1), 其最大值的位置即为该像素点对应的类别标签。

在本发明中, 通过对全景数据的标注, 来实现全景图像的语义分割, 这不同于传统的利用普通图像进行语义分割的方法。

在本发明中, 具体地, 例如, 首先可以在室内房屋的全景图像上, 通过人工标注天花板与墙面的边界线以及地板与墙面的边界线的方法, 来生成包含天花板、墙面和地板的训练数据 (由于在后续的模式生成过程中, 将自动生成天花板的边界线、墙面的边界线、地板的边界线, 所以在此不需要人工标注这些信息)。

然后, 利用这些训练数据来训练其输出包含天花板、地板、墙面这三个

类别的深度学习模型。即，对于每幅全景图像，输入到经过训练的该深度学习模型中，即可从其输出全景图像中每个像素点对应的类别，其类别即上述的天花板、地板、墙面这三个类别中的一种。

接着，提取出其类别属于“天花板”的像素点中处于边缘的那些像素点
5 （可称为“天花板边缘的像素点”）。

假定天花板边缘的这些像素点具有同一高度信息（由于同属于天花板），然后，将这些像素点投影到三维平面上，得到对应的三维点，基于这些三维点形成天花板在三维空间中的平面轮廓。

具体地，本步骤中将像素点（即，天花板边缘的像素点）投影到三维平
10 面上的方法例如可以参考上述像素点到三维空间的投影的相关描述。

请注意，上面仅给出了通过深度学习模型来生成该全景图像的三维空间中的平面轮廓的一种例子，本发明可以使用的方法不限于该例子。

在步骤 S530，将所获取的在拍摄每一个三维对象的至少两幅全景图像时的全景相机的位置的尺度和每一个三维对象的至少两幅全景图像的在三维空
15 间中的平面轮廓的尺度进行归一化，得到经过归一化的各全景图像的在三维空间中的平面轮廓。也就是说，在步骤 S530 中，将在步骤 S510 得到的拍摄每一个三维对象的至少两幅全景图像时的相机位置的尺度和在步骤 S520 得到的每一个三维对象的至少两幅全景图像的三维空间平面轮廓的尺度进行归一化。

20 这里，可选地，步骤 S530 可以参考上述实施例中关于步骤 S430 的描述，重复之处不再赘述。

下面将以房间作为被关注类型的三维对象为例，来详细地描述具体过程。

首先，将在步骤 S510 得到的稀疏三维点，一一对应地投影到全景图像坐标系上，得到全景图像坐标系上的投影点。

25 然后，从上述的全景图像坐标系上的投影点中，基于全景图像的语义分割结果（通过将全景图像输入到语义分割模型，经过该模型的语义分割处理，已经将全景图像上的像素点分类标识为天花板、地板、墙面三类），找到这些投影点中的与在语义分割结果中被标识为“天花板”的像素点一致（或者说位置相同）的投影点，并将与这些被标识为“天花板”的像素点一致的投影点所对应的
30 稀疏三维点的高度（可以认为是三维坐标中的 y 轴上的值）进行统计，排序后取中值或者直接取这些三维点的高度的均值等，作为相机距离天花板的提取高

度 h_c' 。

最后，利用相机距离天花板的假定高度 h_c 与提取高度 h_c' 的比值，重新生成尺度归一化的单房间平面轮廓。

例如，可以通过将假定高度 h_c 与提取高度 h_c' 的比值与在步骤 S520 得到的平面轮廓上的边界点的坐标相乘，得到尺度归一化的平面轮廓上的边界点的坐标，由此得到尺度归一化的平面轮廓。

与上面得到相机距离天花板的提取高度 h_c' 类似地，可以得到相机距离地板的提取高度 h_f' （该提取高度将在后续的步骤中被使用）。

请注意，上述的尺度归一化处理过程仅是一个例子，本发明不应当仅限于该例子。

经过上述的图像处理过程，为后续的模式生成提供了高分辨率的基础。而且，通过上述的图像处理过程，为后续的模式生成提供在三维空间中的并且经过归一化的各全景图像的平面轮廓，简化了后续的模式生成工作，减少了处理时间，提高了处理效率。

经过上述的图像处理为建模提供了模型生成所需的平面轮廓数据，接下来将描述根据本发明的一个示例性实施例的建模方法（为了便于理解和描述起见，以房屋建模为例）。

图 5D 给出了根据本发明的一个示例性实施例的图像处理与对象建模的总体过程的示意性流程图。图 5D 包括上述的图像处理部分与接下来将描述的对象建模部分。

可选地，如图 5D 所示，在步骤 S535，可以基于经过归一化的各全景图像的平面轮廓，得到各单个对象的平面轮廓。

在步骤 S540，基于在步骤 S510 获取的相机位置以及在步骤 S530 得到的经过尺度归一化的各对象平面轮廓，拼接多个对象（例如，房间）的轮廓，以得到多对象平面轮廓。

可选地，在本步骤 S540 中，可以利用获取的相机位置，对经过尺度归一化的每个对象轮廓的三维点坐标进行旋转和平移操作，以将各房间的三维点坐标统一到同一个坐标系中，从而实现多对象平面轮廓的拼接。

可选地，还可以在步骤 S545，对多对象轮廓进行修正。请注意，本步骤也并非三维对象的二维或三维建模必须的步骤，而是可以提高模型的准确度的优选处理方式。

本发明中，在对多对象轮廓进行拼接之后，还可以进一步对多对象轮廓进行修正，以得到更加准确的多对象轮廓。

可选地，还可以进一步在步骤 S550，将生成的多对象平面轮廓转化为对象（例如，房屋）3D 模型。3D 模型生成步骤 S550 可以参考上述三维对象建模方法的实施例中关于 3D 模型生成步骤 S150 和 S450 的描述，重复之处不再赘述。

图 5E 给出了根据本发明的一个示例性实施例的图像处理设备的示意性框图。

如图 5E 所示，根据本发明的一个示例性实施例的图像处理设备 500 可以包括相机位置获取装置 510、单图像平面轮廓生成装置 520、尺度归一化装置 530。

相机位置获取装置 510 可以被配置用于利用针对为待处理的至少一个三维对象的每一个三维对象拍摄的至少两幅全景图像的几何关系，获取在拍摄这些全景图像中的每一个时的全景相机的位置、以及该全景图像上的匹配特征点的三维点坐标。

单图像平面轮廓生成装置 520 可以被配置用于被配置用于对于每个全景图像，通过用于提取图像轮廓的深度学习模型，提取该全景图像的在三维空间中的平面轮廓。

尺度归一化装置 530 可以被配置用于将所获取的在拍摄每一个三维对象的至少两幅全景图像时的全景相机的位置的尺度和每一个三维对象的至少两幅全景图像的在三维空间中的平面轮廓的尺度进行归一化，得到经过归一化的各全景图像的平面轮廓。

可选地，通过下述方式来训练所述深度学习模型：通过人工标注作为训练图像的多个全景图像上的被关注类型的三维对象的轮廓特征，生成被关注类型的三维对象的轮廓特征训练数据集；利用被关注类型的三维对象的所述轮廓特征训练数据集来训练深度学习模型，由此得到经过训练的深度学习模型，其中，深度学习模型的输出包含被关注类型的三维对象的轮廓特征类别。

可选地，单图像平面轮廓生成装置 520 还可以被配置用于：将至少一幅全景图像输入到经过训练的该深度学习模型中，得到该全景图像中每个像素点对应的轮廓特征的类别；提取出每幅全景图像中，轮廓特征属于特定类

别的像素点中的处于边缘的那些像素点作为特定类别边缘像素点；假定所述至少一幅全景图像上的所有特定类别边缘像素点具有同一高度 h_c ，作为特定类别的轮廓的假定轮廓高度，并将每幅全景图像上的特定类别边缘像素点投影到三维平面上，得到与每幅全景图像对应的特定类别三维点，然后基于与每幅全景图像对应的特定类别三维点，形成各全景图像的在三维空间中的平面轮廓。

可选地，相机位置获取装置 110 还可以被配置用于：利用拍摄的每一个三维对象的至少两幅全景图像的几何关系，进行每一个三维对象的至少两幅全景图像之间的特征点匹配，并记录每一个三维对象的至少两幅全景图像中互相匹配的特征点作为匹配特征点；以及通过对于每一个三维对象的至少两幅全景图像的每幅全景图像，减小每幅全景图像上的匹配特征点的重投影误差，来得到拍摄每幅全景图像时的全景相机的相机位置、以及每幅全景图像上的匹配特征点的三维点坐标。

可选地，尺度归一化装置 530 还可以被配置用于：将由相机位置获取装置得到的每一个三维对象的至少一幅全景图像上的所有三维点坐标中的高度值从小到大排序，取排序靠前的高度值的中值或者均值作为特定类别的轮廓的提取的轮廓高度 h_c' ；以及利用特定类别的轮廓的假定轮廓高度 h_c 与特定类别的轮廓的提取的轮廓高度 h_c' 的比值，从每一个三维对象的至少两幅全景图像的在三维空间中的平面轮廓生成经过归一化的每一个三维对象的至少两幅全景图像的在三维空间中的平面轮廓，其中，特定类别的轮廓的假定轮廓高度 h_c 是任意假定的一个高度。

在本发明中，通过使用全景相机拍摄房间的全景图像，为三维对象的二维或三维建模（例如房屋建模等）提供高分辨率的采集图像。

进一步地，在本发明中，采用了高效的图像处理设备，为建模（例如房屋建模）提供了高分辨率的建模准备数据，而且，所提供的建模准备数据能够简化后续模型生成过程。

图 5F 给出了根据本发明的一个示例性实施例的三维对象建模设备的示意性框图。

如图 5F 所示，该三维对象建模设备 5000 可以包括图 5E 所示的图像处理设备 500 以及多对象拼接装置 540。

其中，图像处理设备 500 可以被配置用于处理至少一幅全景图像，生

成经过归一化的各全景图像的在三维空间中的平面轮廓。

多对象拼接装置 540 可以被配置用于基于经过归一化的各全景图像的平面轮廓，拼接得到多对象平面轮廓。

可选地，该三维对象建模设备 5000 还可以包括：单对象平面轮廓生成装置 535，其可以被配置用于基于经过归一化的各全景图像的平面轮廓，得到各单个三维对象（例如，房间）的平面轮廓。

可选地，单对象平面轮廓生成装置 520 还可以被配置用于：对于至少两幅全景图像，通过以下方式、逐个地确定其中是否有多幅全景图像属于同一个三维对象：如果两幅全景图像之间有多于特定比例的匹配特征点，则确定这两幅全景图像属于同一个三维对象；以及如果确定多幅全景图像属于同一个三维对象，则对于由该多幅全景图像得到的同一个三维对象的各平面轮廓，取这些平面轮廓的并集，作为该三维对象的平面轮廓。

另外，可选地，多对象拼接装置 540 还可以被配置用于基于单对象平面轮廓生成装置 535 生成的各单个三维对象的在三维空间中的平面轮廓，来拼接得到多对象平面轮廓。

另外，可选地，三维对象建模设备 5000 还可以包括多对象轮廓优化装置 545，其可以被配置用于对多对象拼接装置 540 得到的多对象平面轮廓进行轮廓修正。

可选地，该三维对象建模设备 5000 还可以包括 3D 模型生成装置 550，其可以被配置用于将拼接得到的多对象平面轮廓转化为房屋 3D 模型。

这里，上述的房屋建模设备 5000 的各装置 510、520、530、535、540、545、550 等分别与在上面详细描述的步骤 S510、520、530、535、540、545、550 等相对应，在此不再赘述。

由此，通过本发明的上述房屋建模设备，能够有效提高所生成的模型的分辨率和准确性。

图 6 给出了根据本发明的一个示例性实施例的图像处理装置的示意性框图。

参见图 6，该图像处理装置 1 包括存储器 10 和处理器 20。

处理器 20 可以是一个多核的处理器，也可以包含多个处理器。在一些实施例中，处理器 20 可以包含一个通用的主处理器以及一个或多个特殊的协处理器，例如图形处理器（GPU）、数字信号处理器（DSP）等等。在一

些实施例中，处理器 20 可以使用定制的电路实现，例如特定用途集成电路（ASIC，Application Specific Integrated Circuit）或者现场可编程逻辑门阵列（FPGA，Field Programmable Gate Arrays）。

存储器 10 上存储有可执行代码，当所述可执行代码被所述处理器 20 执行时，使所述处理器 20 执行上面描述的三维对象建模方法之一，又例如，当所述可执行代码被所述处理器 20 执行时，使所述处理器 20 执行上面描述的图像处理方法之一。其中，存储器 10 可以包括各种类型的存储单元，例如系统内存、只读存储器（ROM），和永久存储装置。其中，ROM 可以存储处理器 20 或者计算机的其他模块需要的静态数据或者指令。永久存储装置可以是可读写的存储装置。永久存储装置可以是即使计算机断电后也不会失去存储的指令和数据的非易失性存储设备。在一些实施方式中，永久性存储装置采用大容量存储装置（例如磁或光盘、闪存）作为永久存储装置。另外一些实施方式中，永久性存储装置可以是可移除的存储设备（例如软盘、光驱）。系统内存可以是可读写存储设备或者易失性可读写存储设备，例如动态随机访问内存。系统内存可以存储一些或者所有处理器在运行时需要的指令和数据。此外，存储器 10 可以包括任意计算机可读存储媒介的组合，包括各种类型的半导体存储芯片（DRAM，SRAM，SDRAM，闪存，可编程只读存储器），磁盘和/或光盘也可以采用。在一些实施方式中，存储器 10 可以包括可读和/或写的可移除的存储设备，例如激光唱片（CD）、只读数字多功能光盘（例如 DVD-ROM，双层 DVD-ROM）、只读蓝光光盘、超密度光盘、闪存卡（例如 SD 卡、min SD 卡、Micro-SD 卡等等）、磁性软盘等等。计算机可读存储媒介不包含载波和通过无线或有线传输的瞬间电子信号。

此外，根据本发明的方法还可以实现为一种计算机程序或计算机程序产品，该计算机程序或计算机程序产品包括用于执行本发明的上述方法（三维对象建模方法和/或图像处理方法）中限定的上述各步骤的计算机程序代码指令。

或者，本发明还可以实施为一种非暂时性机器可读存储介质（或计算机可读存储介质、或机器可读存储介质），其上存储有可执行代码（或计算机程序、或计算机指令代码），当所述可执行代码（或计算机程序、或计算机指令代码）被电子设备（或计算设备、服务器等）的处理器执行时，使所述

处理器执行根据本发明的上述方法的各个步骤。

本领域技术人员还将明白的是，结合这里的公开所描述的各种示例性逻辑块、模块、电路和算法步骤可以被实现为电子硬件、计算机软件或两者的组合。

- 5 附图中的流程图和框图等等显示了根据本发明的多个实施例的系统和方法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，所述模块、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，方框中所标记的功能也可以以不同于附
- 10 图中标记的顺序发生。例如，两个连续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。
- 15 以上已经描述了本发明的各实施例，上述说明是示例性的，并非穷尽性的，并且也不限于所披露的各实施例。在不偏离所说明的各实施例的范围和精神的情况下，对于本技术领域的普通技术人员来说许多修改和变更都是显而易见的。本文中所用术语的选择，旨在最好地解释各实施例的原理、实际应用或对市场中的技术的改进，或者使本技术领域的其它普通技术人员
- 20 能理解本文披露的各实施例。

权利要求书

1. 一种三维对象建模方法，包括：

5 平面轮廓提取步骤，其中，针对为待处理的至少一个三维对象中的每一个三维对象拍摄的至少一幅全景图像，提取所述每一个三维对象的至少一幅全景图像的在三维空间中的平面轮廓；

尺度归一化步骤，其中，基于相机位置，将所述每一个三维对象的至少一幅全景图像的在三维空间中的平面轮廓的尺度进行归一化，得到经过归一化的各全景图像的在所述三维空间中的平面轮廓；以及

10 多对象拼接步骤，其中，基于所述相机位置，对经过尺度归一化的各全景图像的在所述三维空间中的平面轮廓的三维点坐标进行旋转和平移操作，以将所述三维点坐标统一到同一个坐标系中，从而将各三维对象的平面轮廓拼接成多对象平面轮廓。

15 2. 如权利要求 1 所述的三维对象建模方法，其中，所述平面轮廓提取步骤包括：

相机位置获取步骤，其中，利用针对为所述待处理的至少一个三维对象中的每一个三维对象拍摄的至少两幅全景图像的几何关系，获取在拍摄所述每一个三维对象的至少两幅全景图像的每幅全景图像时的全景相机的所述相机位置、以及所述每幅全景图像上的匹配特征点的三维点坐标；

20 单图像平面轮廓生成步骤，其中，对于所述每一个三维对象的至少两幅全景图像的每幅全景图像，基于所述全景图像上的、轮廓特征属于特定类别的像素点中的边缘像素点所围成的轮廓，生成所述全景图像的在所述三维空间中的平面轮廓，

25 其中，所述特定类别至少包括所述全景图像中的三维物体的顶部、底部、支撑部。

3. 如权利要求 2 所述的三维对象建模方法，其中，所述单图像平面轮廓生成步骤包括：

30 基于所述每一个三维对象的至少两幅全景图像中的每幅全景图像上的像素点之间的特征相似度，确定所述每一个三维对象的至少两幅全景图像

中的每幅全景图像上的、轮廓特征属于所述特定类别的像素点中的所述边缘像素点，

其中，两个像素点的特征相似度为所述两个像素点的特征之差的绝对值，每个所述像素点的特征包括灰度、颜色。

5

4. 如权利要求 1 所述的三维对象建模方法，其中，所述平面轮廓提取步骤包括：

相机位置获取步骤，其中，利用针对为所述待处理的至少一个三维对象中的每一个三维对象拍摄的至少两幅全景图像的几何关系，获取在拍摄所述每一个三维对象的至少两幅全景图像的每幅全景图像时的全景相机的所述相机位置、以及所述每幅全景图像上的匹配特征点的三维点坐标；

单图像平面轮廓生成步骤，其中，对于所述每一个三维对象的至少两幅全景图像的每幅全景图像，通过用于提取图像轮廓的深度学习模型，提取所述全景图像的在所述三维空间中的平面轮廓。

15

5. 如权利要求 4 所述的三维对象建模方法，其中，通过下述方式来训练所述深度学习模型：

通过人工标注作为训练图像的多个全景图像上的被关注类型的三维对象的轮廓特征，生成所述被关注类型的三维对象的轮廓特征训练数据集；

利用所述被关注类型的三维对象的所述轮廓特征训练数据集来训练所述深度学习模型，由此得到经过训练的深度学习模型，

其中，所述深度学习模型的输出包含所述被关注类型的三维对象的轮廓特征类别。

6. 如权利要求 5 所述的三维对象建模方法，其中，所述单图像平面轮廓生成步骤包括：

将所述每一个三维对象的至少两幅全景图像输入到经过训练的所述深度学习模型中，得到所述每一个三维对象的至少两幅全景图像中每个像素点对应的轮廓特征的类别；

提取出所述每一个三维对象的至少两幅全景图像的每幅全景图像中，轮廓特征属于特定类别的像素点中的处于边缘的边缘像素点作为特定类别

边缘像素点；

假定所述每一个三维对象的至少两幅全景图像的每幅全景图像上的所有特定类别边缘像素点具有同一高度 h_c ，作为特定类别的轮廓的假定轮廓高度，并将所述每幅全景图像上的所述特定类别边缘像素点投影到三维平面上，得到与所述每幅全景图像对应的特定类别三维点，然后基于与所述每幅全景图像对应的所述特定类别三维点，形成各全景图像的在所述三维空间中的平面轮廓；

其中，所述特定类别包括所述全景图像中三维对象的顶部。

10 7. 如权利要求 2-6 任一项所述的三维对象建模方法，其中，所述相机位置是在所述相机位置获取步骤中得到的，

所述相机位置获取步骤包括：

获取所述每一个三维对象的至少一幅全景图像对应的视频流和时间戳；

15 基于所述时间戳确定所述视频流中与所述每一个三维对象的至少一幅全景图像对应的至少一个视频帧；

根据所述至少一个视频帧的位置，确定拍摄所述每一个三维对象的至少一幅全景图像时的全景相机的相机位置。

20 8. 如权利要求 2-6 任一项所述的三维对象建模方法，其中，所述相机位置是预先确定的或者是在所述相机位置获取步骤中得到的，

所述相机位置获取步骤包括：

25 利用针对为所述待处理的至少一个三维对象的每一个三维对象拍摄的至少两幅全景图像的几何关系，进行所述每一个三维对象的至少两幅全景图像之间的特征点匹配，并记录所述每一个三维对象的至少两幅全景图像中互相匹配的特征点作为匹配特征点；以及

通过对于所述每一个三维对象的至少两幅全景图像的每幅全景图像，减小所述每幅全景图像上的匹配特征点的重投影误差，来得到拍摄所述每幅全景图像时的全景相机的所述相机位置、以及所述每幅全景图像上的匹配特征点的三维点坐标。

30

9. 如权利要求 8 所述的三维对象建模方法，其中，所述尺度归一化步

骤包括：

在所述相机位置是在所述相机位置获取步骤中得到的情况下，

将在获取所述相机位置时得到的所述每一个三维对象的至少一幅全景图像上的所有三维点坐标中的高度值从小到大排序，取排序靠前的高度值的中值或者均值作为特定类别的轮廓的提取的轮廓高度 h_c' ；以及利用特定类别的轮廓的假定轮廓高度 h_c 与特定类别的轮廓的提取的轮廓高度 h_c' 的比值，从所述每一个三维对象的至少两幅全景图像的在三维空间中的平面轮廓生成经过归一化的所述每一个三维对象的至少两幅全景图像的在三维空间中的平面轮廓，

10 在所述相机位置是预先确定的相机确定高度 h_c' 的情况下，

利用假定的相机假定高度 h_c 与相机确定高度 h_c' 的比值，从所述每一个三维对象的至少两幅全景图像的在三维空间中的平面轮廓生成经过归一化的所述每一个三维对象的至少两幅全景图像的在三维空间中的平面轮廓，

其中，所述特定类别的轮廓的假定轮廓高度 h_c 是任意假定的一个高度。

15

10. 如权利要求 1~9 任一项所述的三维对象建模方法，其中，在所述尺度归一化步骤之后，还包括单对象平面轮廓生成步骤，包括：

对于针对为所述待处理的至少一个三维对象拍摄的至少两幅全景图像，通过以下方式、逐个地确定其中是否有多幅全景图像属于同一个三维对象：如果两幅全景图像之间有多于特定比例的匹配特征点，则确定这两幅全景图像属于同一个三维对象；以及

20 如果确定多幅全景图像属于同一个三维对象，则对于由所述多幅全景图像得到的同一个三维对象的各平面轮廓，取所述同一个三维对象的所有平面轮廓的并集，作为所述三维对象的在三维空间中的平面轮廓。

25

11. 如权利要求 10 所述的三维对象建模方法，其中，在所述多对象拼接步骤，还能够基于各单个三维对象的在三维空间中的平面轮廓，拼接得到在三维空间中的多对象平面轮廓。

30 12. 如权利要求 1~11 任一项所述的三维对象建模方法，其中，所述多对象拼接步骤包括：

假设所有的全景图像的在三维空间的平面轮廓共有 N 个, 第 n 个平面轮廓的第 p 个三维点表示为 X_p^n , 所述第 n 个平面轮廓对应的全景图像被拍摄时的相机位置表示为 $\{R_n, t_n\}$, 其中, R_n 为用于表示所述相机位置的旋转参数的旋转矩阵, t_n 为用于表示所述相机位置的平移参数的平移向量, 其中, N 为大于 1 的整数, n 为大于等于 1 的整数,

选定第 i 个平面轮廓对应的全景图像被拍摄时的相机位置作为参考坐标系, 可通过如下的公式将其它平面轮廓的三维点统一到所述参考坐标系下:

$$X_p^{n \rightarrow i} = R_i R_n^{-1} (X_p^n - t_n) + t_i, \quad ,$$

将除所述第 i 个平面轮廓外的所有的经过尺度归一化的平面轮廓的三维点通过所述公式进行转换, 以将所有平面轮廓的三维点统一到同一坐标系下, 从而将各三维对象的平面轮廓拼接成所述多对象平面轮廓。

13. 如权利要求 12 所述的三维对象建模方法, 其中, 所述三维对象建模方法还包括:

15 多对象轮廓优化步骤, 其中, 计算多对象轮廓中, 两个单对象轮廓的相邻边之间的距离, 若所述距离非零且小于特定阈值, 则将所述两个单对象轮廓进行偏移, 以使所述相邻边的距离变为零。

14. 如权利要求 1~13 中的任何一项所述的三维对象建模方法, 还包括:
20 3D 模型生成步骤, 其中, 在所述多对象拼接步骤之后, 将拼接得到的在三维空间中的所述多对象平面轮廓转化为多对象 3D 模型。

15. 如权利要求 14 所述的三维对象建模方法, 其中, 所述 3D 模型生成步骤包括:

25 将拼接得到的所述多对象平面轮廓中的顶部平面轮廓在内部进行三维点插值, 并将得到的各顶部平面轮廓上的所有三维点坐标投影到相应的全景图像坐标系中, 以获取顶部纹理;

将拼接得到的所述多对象平面轮廓中的底部平面轮廓在内部进行三维点插值, 并将得到的各底部平面轮廓上的所有三维点坐标投影到相应的全景图像坐标系中, 以获取底部纹理;

连接顶部轮廓和底部轮廓之间相同平面位置上的三维顶点, 构成支撑部的

平面轮廓，并对于所述支撑部的平面轮廓内部进行三维点插值，并将得到的各支撑部的平面轮廓的所有三维点坐标投影到相应的全景图像坐标系中，以便获取支撑部纹理；

5 基于所述顶部纹理、底部纹理、支撑部纹理，生成整个三维对象的 3D 纹理模型。

16. 如权利要求 15 所述的三维对象建模方法，其中，在所述 3D 模型生成步骤中，将得到的各三维对象的顶部平面轮廓上的所有三维点坐标中的作为所述特定类别的轮廓的提取的轮廓高度 h_c' 的、相机距离相应三维对象顶部的提取高度的高度值，替换为相机距离相应三维对象底部的提取高度 h_f' ，并且使
10 得到的各三维对象顶部平面轮廓上的所有三维点坐标中的长度和宽度值保持不变，得到相应的各三维对象的底部平面轮廓，其中，

所述相机距离相应三维对象顶部的所述特定类别的轮廓的提取的轮廓高度 h_c' 通过如下方式得到：将在获取所述相机位置时得到的所述每一个三维对象的至少一幅全景图像上的所有三维点坐标中的高度值从小到大排序，取排序靠前的高度值的中值或者均值作为所述相机距离相应三维对象顶部的所述特定类别的轮廓的提取的轮廓高度 h_c' ，以及

所述相机距离相应三维对象底部的提取高度 h_f' 通过如下方式得到：将在获取所述相机位置时得到的所述每一个三维对象的至少一幅全景图像上的所有三维点坐标中的高度值从小到大排序，取排序靠后的高度值的中值或者均值作为相机距离相应三维对象底部的提取高度 h_f' 。
20

17. 一种图像处理方法，包括：

相机位置获取步骤，其中，利用针对为待处理的至少一个三维对象中的每一个三维对象拍摄的至少两幅全景图像的几何关系，获取在拍摄所述每一个三维对象的至少两幅全景图像时的全景相机的位置、以及所述每一个三维对象的至少两幅全景图像上的匹配特征点的三维点坐标；
25

单图像平面轮廓生成步骤，其中，对于所述每一个三维对象的至少两幅全景图像的每幅全景图像，基于所述全景图像上的、轮廓特征属于特定类别的像素点中的边缘像素点所围成的轮廓，生成所述全景图像的在三维空间中的平面轮廓；以及
30

尺度归一化步骤, 其中, 将所获取的在拍摄所述每一个三维对象的至少两幅全景图像时的全景相机的位置的尺度和所述每一个三维对象的至少两幅全景图像的在三维空间中的平面轮廓的尺度进行归一化, 得到经过归一化的各全景图像的在三维空间中的平面轮廓。

5

18. 一种图像处理方法, 包括:

相机位置获取步骤, 其中, 利用针对待处理的至少一个三维对象中的每一个三维对象拍摄的至少两幅全景图像的几何关系, 获取在拍摄所述每一个三维对象的至少两幅全景图像时的全景相机的位置、以及所述每一个三维对象的至少两幅全景图像上的匹配特征点的三维点坐标;

单图像平面轮廓生成步骤, 其中, 对于所述每一个三维对象的至少两幅全景图像的每幅全景图像, 通过用于提取图像轮廓的深度学习模型, 提取所述每全景图像的在三维空间中的平面轮廓; 以及

尺度归一化步骤, 其中, 将所获取的在拍摄所述每一个三维对象的至少两幅全景图像时的全景相机的位置的尺度和所述每一个三维对象的至少两幅全景图像的在三维空间中的平面轮廓的尺度进行归一化, 得到经过归一化的各全景图像的在三维空间中的平面轮廓。

19. 一种图像处理装置, 包括:

处理器; 以及

存储器, 其上存储有可执行代码, 当所述可执行代码被所述处理器执行时, 使所述处理器执行如权利要求 1~16 中任何一项所述的方法。

20. 一种三维对象建模方法, 包括:

图像处理步骤, 其中, 使用如权利要求 17 或 18 所述的图像处理方法, 对所述每一个三维对象的至少一幅全景图像进行图像处理, 以得到经过归一化的各全景图像的在三维空间中的平面轮廓; 以及

多对象拼接步骤, 其中, 基于所述经过归一化的各全景图像的在三维空间中的平面轮廓, 拼接得到多对象平面轮廓。

30

21. 如权利要求 20 所述的三维对象建模方法, 还包括:

单对象平面轮廓生成步骤，其中，基于所述图像处理步骤中得到的经过归一化的各全景图像的平面轮廓，得到各单个三维对象的在所述三维空间中的平面轮廓。