



(12) 发明专利申请

(10) 申请公布号 CN 103180883 A

(43) 申请公布日 2013. 06. 26

(21) 申请号 201180048808. 1

(51) Int. Cl.

(22) 申请日 2011. 10. 07

G06T 17/20 (2006. 01)

G06T 15/20 (2006. 01)

(30) 优先权数据

61/391, 069 2010. 10. 07 US

(85) PCT申请进入国家阶段日

2013. 04. 08

(86) PCT申请的申请数据

PCT/US2011/055489 2011. 10. 07

(87) PCT申请的公布数据

W02012/048304 EN 2012. 04. 12

(71) 申请人 桑格威迪公司

地址 美国加利福尼亚州

(72) 发明人 亚当·普赖尔

(74) 专利代理机构 广州华进联合专利商标代理

有限公司 44224

代理人 王程 黎艳

权利要求书2页 说明书13页 附图19页

(54) 发明名称

快速 3D 模型

(57) 摘要

本发明提供一种用于快速、有效地对真实世界 3D 对象进行 3D 建模的系统和方法。基于感兴趣对象的少到两张照片来生成 3D 模型。两张照片中的每一张照片均可用传统的针孔摄像设备获得。根据本发明一实施例的系统包括新型的相机建模器和用于修正相机参数中的误差的有效方法。本发明的其它应用包括用于动画和现实生活运动照片和视频游戏的快速 3D 建模，以及用于建筑和医疗应用。



1. 用于生成真实世界对象的 3D 模型的系统,包括 :

相机建模器,其包括 :接收相机参数的第一输入 ;接收第一和第二点集的第二输入,该第一和第二点集分别与第一对象的第一和第二图像上的点相对应,所述相机建模器根据相机参数提供投射到 3D 空间中的第一和第二点集的投影 ;

对象建模器,其包括 :接收所述投影的输入 ;基于所述投影提供所述第一对象的 3D 模型的第一输出 ;提供投影误差评估的第二输出 ;

所述系统根据所述投影误差评估调整至少一个相机参数,

相机建模单元基于至少一个调整后的相机模型参数对第一和第二点集进行投影,从而使得对象建模器能够提供所述第一对象的误差修正 3D 模型。

2. 根据权利要求 1 所述的系统,其特征在于 :还包括渲染单元,该渲染单元包含用于接收所述误差修正 3D 模型的输入,所述渲染单元基于所述误差修正 3D 模型来提供第一对象的 2D 表征。

3. 根据权利要求 2 所述的系统,其特征在于还包括 :

2D 显示装置,其包括接收所述误差修正 3D 模型的 2D 表征的输入,所述显示装置显示所述第一对象的 2D 表征 ;

操作者控制装置,其连接至所述显示装置,以使得操作者能够与第一对象的 2D 表征进行交互,以测量所述对象的尺寸。

4. 根据权利要求 1 所述的系统,其特征在于还包括 :

显示装置,用于显示真实世界 3D 对象的第一和第二 2D 图像 ;

操作者输入装置,其连接至所述显示装置,使得操作者能够与显示的 2D 图像进行交互,以定义第一和第二点集。

5. 根据权利要求 4 所述的系统,其特征在于 :所述 2D 显示装置还用于显示第二对象的至少一个图像,且操作者控制装置用于使操作者能够将第二对象的所述图像定位在以下其中一种图像之内 :显示的第一图像,显示的第二图像,显示的、基于误差修正 3D 模型渲染过的图像。

5、用于生成对象的 3D 模型的方法,包括 :

用第一和第二初始相机参数对相机建模器进行初始化 ;

通过相机建模器接收第一和第二 2D 点集,该第一和第二 2D 点集对应于出现在对象的第一和第二 2D 图像中的对象上的点 ;

通过相机建模器将第一和第二 2D 点集投影到 3D 模型空间中 ;

基于所述投影确定 3D 坐标,以组建所述对象的 3D 模型 ;

确定与投影的第一和第二 2D 点集相关的误差 ;

根据所述误差调整至少一个初始相机参数,使得第一和第二 2D 点集按照修正的相机参数被再投影 ;

基于再投影的第一和第二 2D 点集确定 3D 坐标,以组建所述对象的 3D 模型。

6. 根据权利要求 5 所述的方法,其特征在于 :重复进行产生投影的步骤、确定 3D 坐标的步骤、确定误差的步骤以及调整相机参数的步骤,直到确定的误差小于或等于预定误差。

7. 根据权利要求 6 所述的方法,其特征在于 :通过不断改变至少一个相机参数以优化收敛到预定误差上的时间,从而进行所述重复步骤和确定误差的步骤。

8. 根据权利要求 5 所述的方法,其特征在于还包括对误差修正 3D 模型进行渲染、以供显示在显示装置上的步骤。

9. 根据权利要求 5 所述的方法,其特征在于还包括以下步骤:

接收第三点集,该第三点集表征出现在第三图像中的第二对象;

通过在 3D 模型空间中的第三点集上进行操作,调整被表征的第二对象的比例和方位,以与第一对象的比例和方位相匹配;

借助被显示的第一对象来显示第二对象。

快速 3D 模型

[0001] 相关申请的交叉引用

[0002] 本申请要求申请号为 61/391,069、申请日为 2010 年 10 月 7 日、名称为“Rapid 3D Modeling”美国临时专利申请（该申请的发明人与本申请的发明人相同）的优先权，其全部内容在此以引用的方式并入本申请中。

背景技术

[0003] 三维（3D）模型将真实世界中的对象的三个维度表征为存储的几何数据。这些模型可用于提供真实世界中的对象的二维（2D）图形图像。通过对存储在对象的 3D 模型中的维度数据进行计算，使得与所提供的显示在显示设备上的对象的 2D 图像的交互作用模拟与真实世界对象的交互作用。当与真实世界中的对象的物理交互不可能发生、存在危险、不现实或者因其他原因而不可取时，模拟的与对象的交互作用是有用的。

[0004] 产生一个对象的 3D 模型的常规方法包括：由艺术家或工程师利用 3D 模型工具在计算机上创造所述模型。这种方法耗费时间，并且需要熟练工来实施。也可通过将来自真实世界对象的模型扫描到计算机上来形成 3D 模型。典型的 3D 扫描仪收集其视场内与对象的表面有关的距离信息。3D 扫描仪产生的“照片”描述了在该照片的每个点处与表面的距离。这使得照片上每个点的三维位置可识别。该技术通常需要从多个不同方向进行多重扫描，以获得与对象的所有侧面有关的信息。这些技术可用在许多应用中。

[0005] 更进一步，各种各样的应用均将受益于不需工程专业知识、不依靠昂贵且费时的扫描设备就能够快速产生 3D 模型的系统和方法。一个例子发现于太阳能系统安装领域。为了选择用于安装到一个建筑物（例如房屋的屋顶）上的合适的太阳能面板，有必要知道屋顶尺寸。在传统的安装中，技术工人被分派到安装地点，对安装区域进行物理考察和测量，以确定其尺寸。实地考察费时、费钱。在有些情形下实地考察还不切实际。例如，恶劣天气会造成一定时间的拖延。安装地点可能会位于与距离最近的技术工人都相距甚远的位置，或者可能难以到达。让建筑物测量可从显示在显示屏上的 3D 模型处获得、而不是造访真实世界中的建筑物并对其进行物理测量的系统和方法是有用的。

[0006] 有些消费者由于不确定屋顶安装了太阳能面板后呈现的美学效果而不愿意在他们的家里安装太阳能系统。有些消费者由于其他原因，比如对障碍物的关注，而喜欢参与决定面板的安装位置。这些关注会对太阳能的吸收构成障碍。人们需要能够快速提供特定太阳能部件安装在给定房屋上时呈现的逼真视觉效果的系统和方法。

[0007] 本发明的各种实施例能够快速生成 3D 模型，这些模型实现了远程测量和现实世界中 3D 对象的逼真呈现的 3D 图形图像的可视化、操作和交互作用。

发明内容

[0008] 本发明提供了一种快速、有效地对真实世界中的 3D 对象进行 3D 建模的系统和方法。基于感兴趣的对象的少到两张照片来产生 3D 模型。这两张照片中的每一张照片均可使用传统的针孔摄像设备来获得。根据本发明一个实施例的系统包括一新型照相机建模器和

用于纠正照相机参数中的错误的有效方法。本发明的其它应用包括用于动画和真实生活运动图片及视频游戏、以及用于建筑和医疗应用的快速 3D 建模。

附图说明

[0009] 本发明的这些以及其它目的、特征和优点将会随着下文中结合附图所进行的具体描述而变得明晰；在这些附图中：

[0010] 图 1 展示了本发明的 3D 建模系统的一个实施例的示例性部署；

[0011] 图 2 为本发明一个实施例的方法的流程图；

[0012] 图 3 展示了一张示例性的第一图像，该第一图像包含适用于本发明的示例性实施例中的、包括房屋屋顶的对象的俯视图；

[0013] 图 4 展示了示例性的第二图像，该第二图像包含房屋（其屋顶如图 3 所示）的前部正视图，其适合用于本发明的一些示例性实施例中；

[0014] 图 5 是一张表，表中包含与图 3 和图 4 所示的第一和第二图像中的示例性 3D 点相对应的 2D 点集；

[0015] 图 6 展示了包含选自图 3 和图 4 所示的示例性第一和第二图像的直角的示例性 3D 点列表；

[0016] 图 7 展示了包含选自图 3 和图 4 所示的示例性第一和第二图像的地平面的示例性 3D 点；

[0017] 图 8 为根据本发明一个实施例的、用于产生 3D 点的方法的流程图；

[0018] 图 9 为根据本发明一个实施例的、估计误差的方法的流程图；

[0019] 图 10 为一种示例性相机参数发生器的功能的概念展示图，根据本发明的一个实施例，所述机参数发生器适于为相机建模器提供相机参数；

[0020] 图 11 为根据本发明一个实施例的、为相机建模器产生初始的第一相机参数的方法的各步骤流程图；

[0021] 图 12 为根据本发明一个实施例的、用于为相机建模器产生第二相机参数的方法的各步骤流程图；

[0022] 图 13 展示了根据本发明一个实施例的、一个对象的示例性图像，所述对象的图像显示在显示装置提供的示例性图形用户界面（GUI）上，并使操作者能够产生用于所述对象的点集；

[0023] 图 14 展示了根据本发明一个实施例的、用于提供对象的误差纠正 3D 模型的各步骤；

[0024] 图 15 展示了根据本发明一替代性实施例的、用于提供对象的误差纠正 3D 模型的各步骤；

[0025] 图 16 为根据本发明一个实施例的、展示示例性 3D 模型发生器的示意图，所述 3D 模型发生器提供来自第一和第二图像的点集的 3D 模型基投影；

[0026] 图 17 展示了根据本发明一实施例的、由示例性第一和第二相机限定的示例性 3D 模型空间，其中，第一和第二相机的其中之一根据俯视图进行初始化；

[0027] 图 18 展示并描述了根据本发明一实施例的、用于提供纠正的相机参数的方法的各步骤；

[0028] 图 19 为根据本发明一实施例的、展示第一和第二图像、相机建模器和相机发生器之间的关系的示意图；

[0029] 图 20 展示了根据本发明一实施例的、用于产生和存储 3D 模型的方法的各步骤；

[0030] 图 21 展示了根据本发明一实施例的 3D 模型产生系统；

[0031] 图 22 为根据本发明一实施例的、展示用于调整相机参数的方法的流程图；

[0032] 图 23 为根据本发明一实施例的示例性 3D 建模系统的框图；

[0033] 图 24 为根据本发明一实施例的、配备有辅助性对象丈量系统的示例性 3D 建模系统的框图。

具体实施方式

[0034] 图 1

[0035] 图 1 展示了部署在建筑物测量系统中的本发明的一个实施例。图像源 10 包括摄影图像，该摄影图像包括真实世界 3D 住宅建筑物 1 的图像。在本发明的一些实施例中，合适的 2D 图像源包括以图片格式比如 JPEG, TIFF, GIF, RAW 和其它图像存储格式存储的 3D 图像的集合。本发明的有些实施例接收至少一种图像，包括建筑物的鸟瞰图。鸟瞰图提供从四个角度拍得的航拍照片。

[0036] 在本发明的一些实施例中，合适的 2D 图像包括航拍和卫星图像。在一个实施例中，2D 图像源为可由系统 200 通过互联网访问的在线数据库。合适的在线 2D 图像源的例子包括、但不限于：United States Geographical Survey (USGS), The Maryland Global Land Cover Facility and TerraServer-USA(最近改名为 Microsoft Research Maps (MSR))。这些数据库存储有地图和航拍照片。

[0037] 在本发明的一些实施例中，图像为地理参考图像。地理参考图像本身、或在其附加文件（例如世界文件）中包含着信息，该信息指示 GIS 系统如何将该图像与其它数据对齐。适合地理参考图像的格式包括 GeoTiff、jp2 和 MrSid。其它图像可在伴随文件（在 ArcGIS 中为世界文件）中载有地理参考信息，所述伴随文件通常为具有与图像文件相同的名称和后缀的小的文本文件。在本发明的一些实施例中，手动地对图像进行地理参考，以供使用。高分辨率的图像可从订阅数据库如 Google Earth ProTM 处获得。MapquestTM 也适用于本发明的一些实施例。在本发明的一些实施例中，所接收的地理参考图像包含地理信息系统 (GIS) 信息。

[0038] 建筑物 1 的图像被捕获，例如被飞机 5 捕获，用机载图像采集设备例如航拍相机 4 拍摄建筑物 1 的航拍照片。相机 4 拍摄的示例性照片 107 为住宅建筑物 1 的屋顶 106 的俯视图。由相机 4 获得的示例性照片 107 为住宅建筑物 1 的俯视平面图。但是，本发明不限于俯视图。相机 4 也可拍摄建筑物 1 的正投影视图和斜视图，以及其它视图。

[0039] 包括图像源 10 的图像不必限于航拍照片。例如，可通过第二相机（例如地面基相机 9）在地面上拍摄建筑物 1 的附加图像。地面基图像包括但不限于建筑物 1 的主视图、侧视图和后视图。图 1 展示了建筑物 1 的第二照片 108。在该图中，照片 108 展现了建筑物 1 的主视图。

[0040] 根据本发明的实施例，一个对象的第一和第二视图不需要借助任何特定类型的图像采集装置来获得。不同的采集装置在不同时间、出于不同目的所获得的图像均适用于本

发明的各个不同实施例。用于获得第一和第二图像的图像采集装置不需要使任何特定的内在或外在相机具有共同属性。本发明不依赖于用来获得第一和第二图像的实际相机所属的任何内在或外在相机的知识。

[0041] 一旦图像存储在图像源 10 中之后,这些图像就可供选择,并可下载到系统 100 中。在一个用途例子中,操作者 113 从顾客处获得街道地址。操作者 113 可利用图像管理单元 103 来访问图像源 10,例如,通过互联网进行访问。操作者 113 可通过提供街道地址而获得图像。图像源 10 通过提供位于给定的街道地址处的房屋的多张视图来作出反应。本发明各实施例中所用的合适视图包括俯视平面图、正视图、透视图、正投影图、斜视图和其它类型的图像与视图。

[0042] 在本例中,第一图像 107 展示了房屋 1 的第一视图。该第一视图为房屋 1 的屋顶的俯视平面图。第二视图 108 展示了同一房屋 1 的第二视图。该第二视图从与第一视图的视点不同的视点展现了所述屋顶。因此,第一图像 107 包含处于 2D 空间中的第一方位的对象 1 的图像,而第二图像 108 包含处于 2D 空间中的第二方位的相同对象 1 的图像。在本发明的一些实施例中,至少一个图像包括对象的俯视平面图。第一图像 107 和第二图像 108 可在尺寸、宽高比和图像所展示的对象 1 的其它特征上彼此不同。

[0043] 当想要测量建筑物 1 的尺寸时,从图像源 10 获得该建筑物的第一和第二图像。重要的是要注意:与提供所述第一和第二图像的相机 4 和 9 有关的信息不一定要存储在图像源 10 中,也不一定要随着检索到图像一起提供。在许多情形下,从任何源中均不能获得与用来拍摄第一和第二照片的相机有关的信息。本发明的实施例中,不管与实际的第一和第二相机有关的信息是否可以获得,均能够基于第一和第二图像来确定与第一和第二相机有关的信息。

[0044] 在一实施例中,房屋的第一和第二图像由系统 100 接收并向操作者 113 展示。操作者 113 与所述图像交互,以产生点集(控制点)提供给 3D 模型发生器 950。模型发生器 950 提供所述对象的 3D 模型。该 3D 模型被渲染引擎渲染后显示在 2D 显示装置 103 上。操作者 113 用测量应用来测量显示在显示器 103 上的对象的尺寸,以便与显示的对象进行交互。基于与第一和第二图像的规格有关的信息,将模型测量值转化为真实世界测量值。因此真实世界对象的测量是在不需要实地探访的情形下进行的。本发明的实施例能够基于所述对象的至少两张摄影图像来生成建筑物的 3D 模型。

[0045] 图 2

[0046] 图 2 展示并描述了根据本发明一个实施例的、基于所述对象的 3D 模型来测量真实世界对象的方法。

[0047] 在步骤 203,产生要被测量的建筑物的 3D 模型。在步骤 205,模型被渲染在显示装置上,从而使得操作者能够与显示的图像进行交互,以便测量该图像的尺寸。在步骤 207,接收测量值。在步骤 209,将测量值从图像测量值转换为真实世界测量值。此时,该测量值适用于用来为建筑物提供太阳能系统。

[0048] 为了执行步骤 203,本发明的模型发生器接收匹配点并产生 3D 模型。通过运用新的优化技术,将该 3D 模型精细化为经重建的 3D 结构。精细化过的 3D 模型代表真实世界建筑物,其具有足够的精度,以使得建筑物的可用测量值能够通过测量所述精细化过的 3D 模型而获得。

[0049] 为了做到这一点,3D 模型被渲染在显示装置 103 上。对所显示的模型的尺寸进行测量。测量值被转换为真实世界测量值。该真实世界测量值由太阳能供应系统使用,以提供具有太阳能面板的建筑物。

[0050] 图 3 和图 4

[0051] 图 3 和图 4 展示了合适的第一和第二图像的例子。图 3 展示了第一图像 107,该第一图像包括房屋屋顶的俯视平面图。例如,第一图像 107 为位于建筑物屋顶上方、从而能够获得屋顶的俯视平面图的相机拍得的照片。在这个最简单的实施例中,二维的第一图像 107 假定通过传统的将三维对象(在本例中为房屋)投影到二维图像平面上的投影方法获得。

[0052] 图 4 展示了第二图像 108,该第二图像包括图 3 所示房屋(其包含图 3 所示的屋顶)的前部正视图。重要的是要注意,所述第一和第二图像不一定要是立体图像。此外,第一和第二图像也不一定要是扫描图像。在本发明的一个实施例中,第一和第二摄影图像由图像采集装置如相机获得。

[0053] 在本说明书中,“相片”(photograph)是指由落在感光表面上的光创建而得的图像。光敏表面包括胶卷和电子成像仪,例如电荷耦合器件(CCD)或者互补金属氧化物半导体(CMOS)成像设备。在本说明书中,相片是用相机创建而成的。相机(camera)是指包括镜头的设备,所述镜头将场景的可见波长光聚焦到人眼可见的再现中。

[0054] 在本发明的一个实施例中,第一图像 107 包括待测量的真实世界对象的正投影。通常图像采集装置,例如相机或传感器,由车辆或平台(如飞机或卫星)承载,瞄准天低点(nadir point),该天低点位于平台正下方和/或从平台垂直向下。图像中与所述天低点相对应的点或像素为与所述图像采集装置正交的点/像素。图像中的所有其它点或像素相对于所述图像采集装置而言是倾斜的。随着这些点或像素距天低点越来越远,它们相对于图像采集装置也越来越倾斜。同样地,地面采样距离(即:每一像素对应的、或每一像素覆盖的表面积)也越来越大。正交图像中的这种倾斜导致图像中的特征被扭曲,特别是距天低点相对较远的图像。

[0055] 为了用平行于 y 轴的正投影(侧视图)将来自真实世界图像的 3D 点 a_x, a_y, a_z 投影到相应的 2D 点 b_x, b_y 上,可用下面的示例性关系式来描述相应的相机模型:

$$[0056] b_x = s_x a_x + c_x$$

$$[0057] b_y = s_z a_z + c_z$$

[0058] 其中,s 为任意的比例因子,c 为任意偏移量。在本发明的一些实施例中,这些常量用于对齐第一相机模型视口,以匹配第一图像 105 中呈现的视图。用矩阵乘法,所述等式变为:

$$[0059] \begin{bmatrix} b_x \\ b_y \end{bmatrix} = \begin{bmatrix} s_x & 0 & 0 \\ 0 & 0 & s_z \end{bmatrix} \begin{bmatrix} a_x \\ a_y \\ a_z \end{bmatrix} + \begin{bmatrix} c_x \\ c_z \end{bmatrix}$$

[0060] 在本发明的一个实施例中,对正交图像进行失真校正。例如,通过正射纠正(ortho-rectification)对失真进行去除,或者补偿,所述正射纠正通过使正交图像的每个像素适配或弯曲到正高网格或坐标系上,从根本上除去了来自正交图像的斜度。这一正射纠正过程创建一图像,在该图像中,所有像素均具有相同的底面采样距离,并面向北方。因

此,只要图像比例已知,正射纠正图像上的任何点都可用 X、Y 坐标系来定位,地面特征的长度和宽度,以及这些特征之间的相对距离,均可计算出来。

[0061] 在本发明的一个实施例中,第一和第二图像的其中之一包括倾斜图像。倾斜图像可用图像采集装置获得,所述图像采集装置瞄准或对着载有该图像采集装置的平台的侧面或下方。倾斜图像不像正交图像,其显示的是地面特征(如房屋、建筑物和 / 或山脉)的侧面和顶部。倾斜图像的前景中的每个像素均对应于所描绘的表面或对象的较小的面积(即:每个前景像素均具有较小的地面采样距离),然而背景中的每个像素均对应于所描绘的表面或对象的较大的面积(即:每个背景像素均具有较大的地面采样距离)。倾斜图像捕获所述表面或对象的梯形区域或视图,且梯形的前景具有远小于梯形背景的地面采样距离(即:具有更高的分辨率)。

[0062] 图 5

[0063] 一旦第一和第二图像被选定,则选择显示点集(控制点)。在本发明的一些实施例中,点集的选择是手动进行的,例如由操作者手动进行。在本发明的其它实施例中,控制点也可自动选择,例如通过机器视觉特征匹配技术自动选择。对于手动实施方案,操作者在第一图像中选择一个点,并在第二图像中选择一个对应点,其中,这两个点代表真实世界 3D 建筑物中的相同点。

[0064] 为了识别和指示匹配点,操作者 113 与第一和第二显示图像进行交互,以指出显示的第一和第二图像上的相应点。在图 3 和 4 所示的实施例中,真实世界 3D 建筑物 1 的点 A 是指屋顶 1 的右转角。点 A 出现在第一图像 107 和第二图像 108 中,虽然在这两张图像中出现的位置不同。

[0065] 为了指出第一和第二图像中的对应点,操作者将显示标记放置在每个第一和第二图像 105、107 中对象的对应点的上方。例如,将标记放置在第一图像 105 中的对象 102 的点 A 的上方,然后再将标记放置在第二图像 107 中的对象 102 的点 A 的上方。在每个点处,操作者例如通过鼠标点击右或左、或通过其它选择机制来指示出点的选择。本发明实施例中也可使用其它装置,比如追踪球、键盘、光笔、触摸屏、操纵杆,等等。因此,操作者与第一和第二图像进行交互,以产生如图 5 所示的控制点对。

[0066] 在本发明的一个实施例中,可以使用触摸屏显示器。此时,操作者通过触摸屏幕,在显示的图像上选择点或感兴趣的其它区域。像素坐标从显示屏坐标描述转化为,例如与含有感测触摸像素的图像相对应的坐标系统描述。在本发明的其它实施例中,操作者用鼠标将标记,或者其它指示器,放置在图像上要选择的点的上方。点击鼠标,记录所放置的标记的像素坐标。系统 100 将该像素坐标转化为对应的图像坐标。

[0067] 将控制点提供给本发明的 3D 建模系统的 3D 模型发生器 950。通过寻找用于每一点对的极线的交点来实现成像结构的重建。

[0068] 图 6 和图 7

[0069] 图 7 展示了界定地平面的点。在本发明的一些实施例中,生成的 3D 模型参照地面平行进行精细化。图 7 展示了根据图 5 所示示例性控制点列表的示例性控制点列表,其中,图 7 中的控制点包括根据本发明实施例的地平面平行线。

[0070] 图 6 展示了界定与所述对象相关联的直角的点。类似地平面,本发明的一些实施例中可使用直角来精细化 3D 模型。

[0071] 图 8

[0072] 图 8 展示了本发明的系统。如关于图 1-7 的说明所述,操作者从显示在显示装置 803 上的第一和第二图像上选择第一和第二图像点集。第一相机矩阵(相机 1)接收来自第一图像的点集。第二相机矩阵(相机 2)接收来自第二图像的点集。通过提供用于相机 1 和相机 2 矩阵的初始参数,对产生的模型进行初始化。

[0073] 在本发明的一个实施例中,相机参数包括以下内在参数:

[0074] a) (u_0, v_0): 图像中心的像素坐标,其为相机中心在视网膜上的投影。

[0075] b) (a_u, a_v): 图像的比例因子。

[0076] c) ($\text{dim}_x, \text{dim}_y$): 图像的像素大小。

[0077] 外部参数定义如下:

[0078] a) R: 相机轴在参考坐标系中的旋转。

[0079] b) T: 相机中心在参考坐标系中的姿势(pose),以 mm 为单位。

[0080] 相机参数建模单元 815 用于提供与第一和第二图像相对应的相机模型(矩阵)。所述相机模型是对用于获取第一和第二图像的相机的描述。本发明的相机参数模型将第一和第二相机矩阵建模为包括相机约束。本发明的参数模型解释了不大可能出现的或者无效的参数,例如相机位置,其可指出处于远离图像中对象的方向上的镜头。因此,这些参数值在计算测试参数时不需要考虑。

[0081] 所述相机参数建模单元基于、至少部分基于所选的第一和第二图像的属性,对包括第一和第二参数集的参数之间的关系和约束描述关系进行建模。

[0082] 本发明的相机参数模型 1000 包含与第一和第二相机上的位置约束有关的足够多的信息,以防止选择到相机参数的无效或不太可能的子组合。因此,产生 3D 模型的计算时间小于当用于例如不可能或无效或可能性不大的相机位置的参数值被包括进测试参数时所需的计算时间。

[0083] 在一些实施例中,为了描述第一和第二相机在三维欧氏空间中的方位,采用了 3 个参数。本发明的不同实施例代表不同方式下的相机方位。例如,在一个实施例中,相机参数模型用欧拉角(Euler angle)表示相机位置。欧拉角为描述刚性体的方位的 3 个角度。在那些实施例中,用于 3D 模型空间的坐标系将相机位置描述为,是否存在界定包括欧拉角的相机角的真正的万向节。

[0084] 欧拉角也表示将参考(相机)框移动到参考(3D 模型)框的三个组分旋转。因此,通过将三个元素旋转(绕单个轴的旋转)进行组合,可以表征任何方位,且任何旋转矩阵都可被重新组合为三个元素旋转矩阵的产物。

[0085] L

[0086] 对于“点对”中的每一点,模型单元 303 通过对称的假设相机(其获得包含所述点的图像),对视线(或射线)进行投影。通过第一图像极点的线和通过第二图像极点的线会在理想条件下相交,所述理想条件例如:相机模型精确地代表了用于获取所述图像的实际相机,不存在噪音,以及,点对的识别很精确,且在第一和第二照片之间保持一致。

[0087] 在本发明的一个实施例中,3D 模型单元 303 利用三角测量技术,确定投影穿过所述第一和第二相机模型的射线的交点。通常,三角技术是通过测量固定基线的任意末端处已知点与测量点之间的角度,来确定测量点的位置,而不是直接测量点与点之间的距离。接

着,所述点可被固定为三角形的第三点,所述三角形具有一条已知的边和两个已知的角。距一个点的坐标和距离可通过计算三角形的一条边的长度、给定的角度测量值和由所述点与另外两个已知的参考点形成的三角形的其它两条边的测量值而得到。在无误差实施例中,交点的坐标包括 3D 模型空间中所述点的三维位置。

[0088] 根据本发明的一些实施例,3D 模型包括真实世界建筑物的三维表征,其中,所述表征包括参考一坐标系统(如笛卡儿坐标系统)的几何数据。在本发明的一些实施例中,3D 模型包括图形数据文件。所述 3D 表征存储在存储器或处理器(未图示)中,用于进行计算和测量。

[0089] 3D 模型可通过 3D 渲染过程直观显示为二维图像。本发明的系统产生 3D 模型之后,渲染引擎 995 将模型的渲染过的 2D 图像显示在显示装置 103 上。传统的渲染技术适用于本发明。除了渲染之外,3D 模型还可用于图形或非图形计算机模拟和计算。随后可将渲染过的 2D 图像存储起来用于查看。但是,本发明实施例的描述使得,随着操作者 113 指出控制点对,渲染过的 2D 图像能够被近实时地显示在显示器 103 上。

[0090] 包括 3D 模型的 3D 坐标限定了 3D 真实世界空间中建筑物点的位置。与此相反,图像坐标限定了胶卷或电子成像装置上的建筑物图像点的位置。

[0091] 点坐标在 3D 图像坐标与 3D 模型坐标之间进行转换。例如,位于平行于摄影图像平面的平面上的两个点之间的距离,可通过测定它们在图像上的距离来确定,如果图像的比例因子为已知的话。将测得的距离乘以 $1/s$ 。在本发明的一些实施例中,第一和第二图像中的任一个图像或全部两个图像的比例信息是已知的,例如,通过接收作为与下载的图像一起存在的元数据而获知所述比例信息。该比例信息存储起来,以供测量单元 119 使用。因此,测量单元 119 使得操作者 113 能够通过测量显示在显示装置 103 上的模型而对真实世界 3D 对象进行测量。

[0092] 操作者 61 选择至少两个图像,用于下载到系统 100。在本发明的一个实施例中,第一选定图像为房屋的俯视平面图。第二选定图像为房屋的透视图。操作者 61 将两张图像都显示在显示装置 70 上。操作者 61 利用鼠标或其它合适的输入装置,在第一和第二图像上选取点集。对于在第一图像上选取的每一个点,相应地在第二图像上选取对应点。如上所述,系统 100 使得操作者 109 可与显示在 2D 显示装置 103 上的二维图像进行交互,并可操纵该二维图像。在图 1 所示的简化实施例中,至少一个二维图像,例如第一摄影图像 105,是通过处理器 112 从图像源 10 处获得的。在本发明的其它实施例中,合适的 2D 图像源存储在处理器 112 中,可供操作者 109 选择,以显示在显示装置 103 上。本发明不限定图像源的数目和类型。反而,可以采用多种多样的图像源 10,这些图像源包含 2D 图像,用于采集和显示在显示装置 103 上。

[0093] 例如,以上描述的本发明实施例用于基于建筑物的图像,远程测量住宅建筑物的尺寸。在这些实施例中,商业地理图像数据库,例如 MicrosoftTM 支持的数据库,可适用为 2D 图像源。本发明的有些实施例依赖于一个以上的 2D 图像源。例如,第一图像 105 选自第一图像源,而第二图像 107 选自第二不相关图像源。通过消费级成像设备(例如一次性相机、视频摄像机等)获得的图像适用于本发明。类似地,由卫星、地理调查成像设备和各种各样的提供真实世界对象的商业级 2D 图像的其它成像设备获得的专业图像适用于本发明的各种不同实施例。

[0094] 根据本发明的一个替代性实施例，第一和第二图像是用连接至处理器 112 的本地扫描仪扫描而得的。用于每张扫描图像的扫描数据被提供给处理器 112。该扫描图像被显示在显示装置 103 上展示给操作者 109。在另一替代性实施例中，图像采集装置所在的地点为真实世界的房屋所在的地点。在这种情形下，图像采集装置通过互联网将图像提供给处理器 112。图像可以实时提供，也可以存储起来，在未来提供。另一图像源为通过数据网络连接至处理器 112 的图像归档和通信系统。大量的能够产生或发送图像的方法和装置适用于本发明的各个实施例。

[0095] 对模型进行精细化

[0096] 在实践中，极几何并不完全体现在真实照片中。来自第一和第二图像的控制点的 2D 坐标不能以任意精度测量。各种类型的噪音，例如来自镜头畸变或兴趣点检测误差的几何噪音，导致控制点坐标不准确。此外，第一和第二相机的几何形状也并不完全已知。因此，当进行三角技术测量时，由 3D 模型发生器从相应控制点、通过第一和第二相机矩阵投影的线不会总是与 3D 空间相交。在这种情形下，基于对 3D 模型发生器投影的线的相对线位置的评估来估计 3D 坐标。在本发明的一个实施例中，通过识别 3D 模型空间中的点来确定所估计的 3D 点，所述 3D 模型空间中的点代表第一控制点投影与第二控制点投影的最近近端关系。

[0097] 所估计的 3D 点将具有的误差与其偏离真实世界建筑物上的相同点的偏差成正比，对真实世界建筑物上的该点进行直接、无误差的测量。在本发明的一些实施例中，估计误差代表了估计点与 3D 点之间的偏差，所述 3D 点是控制点对的无噪音、无失真、无误差投影的结果。在本发明的其它实施例中，估计误差代表了估计点与 3D 点之间的偏差，其中所述 3D 点代表了在 3D 模型的产生过程中基于外部定义标准定义的（例如由操作者定义的）、真实世界 3D 点的“最佳估计”。

[0098] 再投影误差是与投影点和测量点之间的图像距离相对应的几何误差。再投影过程量化了 3D 点 $\hat{\mathbf{x}}$ 的估计是如何紧密地重新创造了该点的真实投影 \mathbf{x} 。更精确地，使 \mathbf{P} 为相机的投影矩阵，而 $\hat{\mathbf{x}}$ 为 $\hat{\mathbf{x}}$ 的图像投影，即： $\hat{\mathbf{x}} = \mathbf{P} \hat{\mathbf{x}}$ 。 $\hat{\mathbf{x}}$ 的再投影由 $d(\mathbf{x}, \hat{\mathbf{x}})$ 给出，其中 $d(\mathbf{x}, \hat{\mathbf{x}})$ 表示由矢量 \mathbf{x} 和 $\hat{\mathbf{x}}$ 代表的图像点之间的欧氏距离。

[0099] 为了生成尽可能地代表被模仿的 3D 真实世界建筑物的 3D 模型，希望将再投影误差降到最小。因此，为了生成对尺寸测量（例如用于太阳能面板安装目的时）而言精度足够的 3D 模型，本发明的实施例调整第一和第二相机的描述参数，以使投影线尽可能地接近交点，同时确保估计 3D 点位于相机参数模型的约束之内。

[0100] 在本发明的一个实施例中，对如上述产生的 3D 模型坐标进行精细化处理。给定若干个 3D 点，这些 3D 点包含 3D 模型，该 3D 模型通过用相机模型对控制点对进行投影而生成，调整相机参数和包含所述模型的 3D 点，直到所述 3D 模型满足最优标准时为止，所述最优标准涉及所有点的相应图像投影。它相当于 3D 图像和观看参数（即：相机姿势和可能的内在校准和径向失真）的最优化问题，是为了在参数模型约束下获得最佳的重建。本发明的技术有效地将观察到的图像位置与预测图像点之间的再投影误差降到最小，该再投影误差被表示为大量非线性、实值函数的平方之和。这种类型的最小化通常是用非线性最小二乘算法获得的。在这些方法中经常采用的是 Levenberg–Marquardt 算法。

[0101] Levenberg–Marquardt 算法在当前估计值附近对要被最小化的函数进行迭代线性化。这种算法涉及被称为正规方程组的线性系统方案。这种方法虽然有效,但是即使 Levenberg–Marquardt 算法的稀疏变量 (sparse variant) 明确利用正规方程组零模式,避免了存储和操作零元素,这种算法在实际应用于本发明所适用的场合时也还需要花费大量的时间来进行计算。

[0102] 图 8

[0103] 图 8 为根据本发明一个实施例的、用于基于至少两张 2D 图像来生成对象的 3D 模型的方法的各步骤流程图。

[0104] 在步骤 805 处,接收操作者选定的控制点。例如,操作者从含有房屋的第一图像上选定房屋的一部分 A。操作者再在含有相同房屋的第二图像上选定同一房屋的相同部分 A。将第一和第二图像上描绘的被操作者选定的房屋部分的显示坐标提供给处理器。在步骤 807,接收初始相机参数,例如,从操作者处接收。在步骤 809,基于、至少部分基于相机参数模型来计算剩余的相机参数。剩下的步骤 811 至 825 按图 8 所描述的进行。

[0105] 图 9

[0106] 图 9 展示和描述了根据本发明一个实施例的、用于使生成的 3D 模型中的误差最小化的方法。

[0107] 图 10

[0108] 在本发明的一个实施例中,第一和第二相机中的每一个均被建模为安装在相机支撑平台上的相机,所述相机支撑平台位于 3D 模型空间 (915 916) 内。所述平台反过来连接至“相机万向节”。因此不可能的相机位置体现为“万向节锁定”位置。万向节锁定是三维空间内的一个自由度损失,当三个万向节中的其中两个万向节的轴被推入到平行配置结构时,会发生这种“万向节锁定”,将所述系统“锁定”到二维空间内的旋转。

[0109] 图 10 的模型代表根据本发明一个实施例的、用于快速确定最佳第一和第二相机矩阵的有利配置和方法,所述相机矩阵用于将 2D 控制点投影到模型空间。根据所述模型,获得用于第一和第二相机矩阵的初始参数,假定相应的假设相机 (915, 916) 的孔被排列为指向球 905 的中心。此外,一个相机 916 被建模为相对于球 901 定位于坐标 x0, y1, z0 处,位于坐标轴 1009 上,即 :位于所述球的上半球的顶上,相机孔向下瞄准所述球的中心。

[0110] 此外,可能的位置的范围局限于所述球的表面上的位置,进一步地局限于所述球的上半球。此外,相机 915 的 x 轴位置被设置为保持 x = 0。因此,遵循以上约束的相机 915 呈现的位置的 z 轴坐标将处于 z = 1 至 z = -1 之间,其中,相机 915 相对于 y 轴的位置由 z 轴位置决定。相机 915 和 916 中的每一个均可绕其各自光轴自由旋转。

[0111] 图 10 所示的布置提供了相机矩阵初始化参数,其便于 3D 点收敛评估,从初始评估到评估满足定义的收敛准则。

[0112] 获得用于内在相机参数的初始值,在图 11 和 12 所示的方法的初始化步骤中建立所述初始值。这些初始值在所述方法的执行过程中不变。另一方面,通过使一个相机的位置沿两个轴固定,同时使另一相机的位置沿一个轴固定,从而使得用于本发明的连续迭代模拟方法的外在参数的排列减少。

[0113] 图 11 参数方法

[0114] 图 11 展示和描述了基于图 10 中参数模型所给定的 C1 初始参数,来确定相机

1(C1) 俯仰、偏航和滚动的方法。

[0115] 图 12- 参数方法

[0116] 类似地,图 12 展示和描述了基于图 10 中参数模型所给定的 C2 初始参数,来确定相机 2(C2) 俯仰、偏航和滚动的方法。

[0117] 图 13- 示例性 GUI 截屏

[0118] 图 13 是根据本发明实施例的图形用户界面的截屏图,所述图形用户界面使得操作者能够与显示的第一和第二图像进行交互。

[0119] 图 14- 模拟方法 - 使用最低误差输出

[0120] 图 14 为流程图,其展示和描述了用于产生 3D 模型、同时使产生的 3D 模型中的误差最小化的方法的各步骤。

[0121] 图 15- 相机参数和模拟方法

[0122] 图 15 为流程图,其展示和描述了根据本发明实施例的、用于产生 3D 模型的方法的各步骤。

[0123] 图 16

[0124] 图 16 为示意图,其展示了根据本发明一个实施例的、基于来自第一和第二图像的投影点集来提供 3D 模型的 3D 模型发生器的例子。图 16 描绘了 3D 模型的位于 1,2 和 3 处的 3D 点,这些点与第一和第二图像中的 2D 点相对应。3D 模型发生器在控制点对上进行操作,以便为每一控制点对提供相应的 3D 点。对于第一和第二图像上的第一和第二图像点(对应于相同的三维点),所述图像点和所述三维点以及光学中心共平面。

[0125] 3D 空间内的对象可通过设备的取景器映射到对象在图像 2D 空间内的图像上,所述设备的取景器通过透视投影变换技术采集图像。以下参数有时用于描述这种变换:

[0126] $a_{x,y,z}$ - 要被投影的、真实世界 3D 空间中的点。

[0127] $c_{x,y,z}$ - 相机在实际真实世界中的位置。

[0128] $\theta_{x,y,z}$ - 真实世界相机的旋转。当 $c_{x,y,z} = \langle 0,0,0 \rangle$ 且 $\theta_{x,y,z} = \langle 0,0,0 \rangle$ 时,3D 矢量 $\langle 1,2,0 \rangle$ 被投影为 2D 矢量 $\langle 1,2 \rangle$ 。

[0129] $e_{x,y,z}$ - 观察者相对于真实世界显示器表面的位置。

[0130] 这会导致:

[0131] $b_{x,y}$ - a 的 2D 投影。

[0132] 本发明采用上述的反向变换。换句话说,如通过采集图像的设备的取景器所观看到的,本发明在 2D 空间的对象图像上映射一点。为了实现以上目的,本发明提供相机 1 矩阵 731 和相机 2 矩阵 732,以通过将点对投影到 3D 模型空间 760 上而以模型形式重建所述 3D 真实世界对象。

[0133] 相机矩阵 1 和 2 由相机参数限定。相机参数可包括“内在参数”和“外在参数”。外在参数限定相机的外部方位,例如在空间中的位置和观察方向。内在参数限定成像过程的几何参数。主要是镜头的焦距,但也可包括对镜头畸变的描述。

[0134] 因此,第一相机模型(或矩阵)包括采集第一图像的相机的假设描述。第二相机模型(或矩阵)包括采集第二图像的相机的假设描述。在本发明的一些实施例中,相机矩阵 731 和 732 用相机检校技术构建而成。相机检校是寻找相机的产生给定照片或视频的真实参数的过程。相机参数用包括相机矩阵 1 和 2 的 3×4 矩阵来表示。

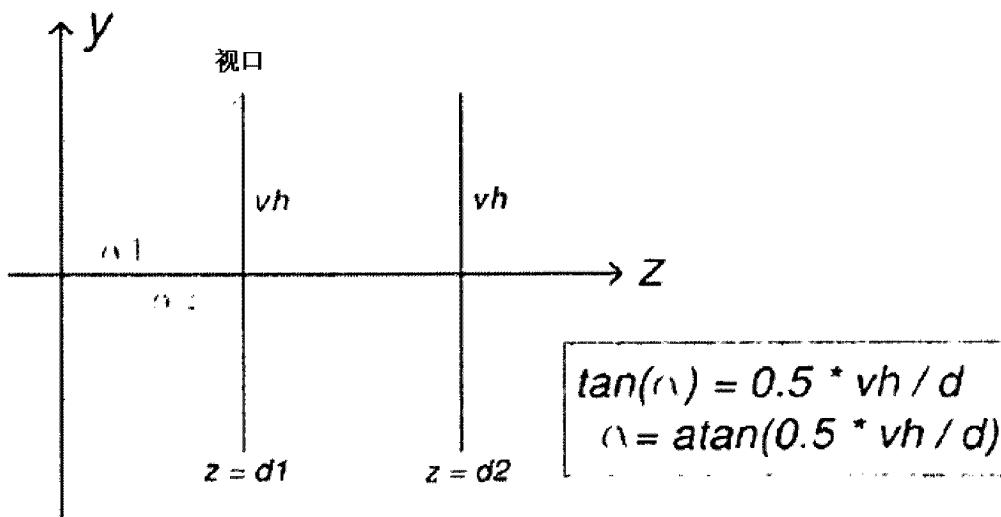
[0135] 图 17- 相机矩阵和模型空间

[0136] 图 17 展示了一个 3D 模型空间, 控制点由第一和第二相机模型投影到所述 3D 模型空间中。

[0137] 这里使用的术语“相机模型”是指 3×4 矩阵, 该矩阵描述了含有真实世界对象的 3D 点通过针孔摄像头式相机映射为所述对象的 2D 图像上的 2D 点。在这种情形下, 所述 2D 场景, 或者摄影框, 就称为视口。

[0138] 相机与投影平面之间的距离为 d ; 视口的尺寸为 vw 和 vh 。这些值一起决定了投影的视场, 也即 : 投影图像中可见的角度 :

[0139]



[0140] 投影仪

[0141] 第一和第二相机矩阵通过假设相机从第一和第二图像投影出自每个 2D 控制点开始的射线, 所述假设相机根据相机模型配置, 所述射线射入提供有 3D 模型的 3D 图像空间中。

[0142] 因此, 每个相机矩阵均根据其自身的相机矩阵参数设置投射射线。由于用于提供第一和第二图像的相机的实际相机参数是未知的, 一种解决途径是估计所述相机参数。

[0143] 已知要通过第一和第二相机矩阵进行投射的 2D 点的给定点集对应于投影到 3D 模型中的理想投影的相同点。基于这方面的知识, 根据本发明的原理的相机参数评估包括以下步骤 : 手动提供估计的初始值, 进行收敛测试, 和基于收敛测试的结果调整相机矩阵。

[0144] 图 18- 图像配准方法

[0145] 图 18 展示和描述了根据本发明一个实施例的、用于使第一和第二图像相对于彼此进行配准的方法的各步骤。

[0146] 图 20 用于生成 3D 模型的方法

[0147] 图 20 展示和描述了根据本发明一个实施例的、用于捆绑调整的方法的各步骤。

[0148] 图 21 模型发生器

[0149] 图 21 为根据本发明一个实施例的 3D 模型发生器的框图。

[0150] 图 22 模型生成方法概述

[0151] 图 22 为一流程图, 其展示和描述了根据本发明一个实施例的、用于捆绑调整的方

法的各步骤。

[0152] 图 23 模型发生器实施例

[0153] 图 23 为根据本发明一个实施例的相机建模单元的框图。

[0154] 部件组成系统 100 可用作独立单元，并可替代性地集成到各种组合中。各部件可用于各种各样的硬件和软件组合中。

[0155] 虽然本发明描述了其优选的设计，但本发明在其公开的精神和原则的范围内还可进行进一步地修改。因此本公开意在涵盖上文所述的结构和元件的任何等同物。此外，本公开还意在涵盖使用本文所公开的基本原理的、本发明的任何变形例、用途或者适应例。另外，本公开也意在涵盖那些虽然偏离本发明的主题、但仍在本领域的已知或惯用做法的范围之内、且落入后附权利要求的限定范围内的方案。虽然已经用特定实施例对本发明进行了展示和描述，但本发明不限于具体实施例。读完本文后，很多种修改、改变和改进对于读者而言都变得显而易见。

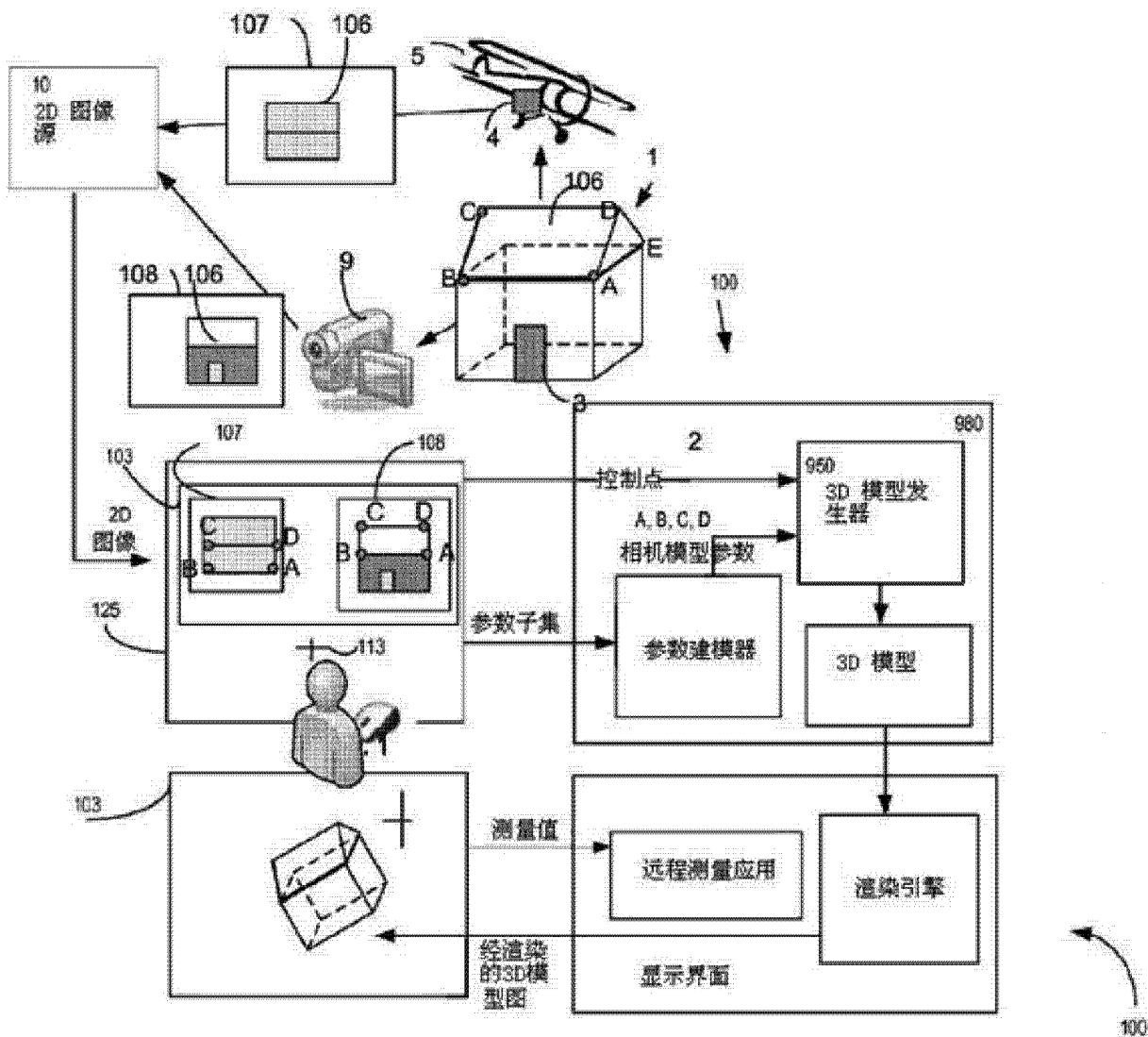
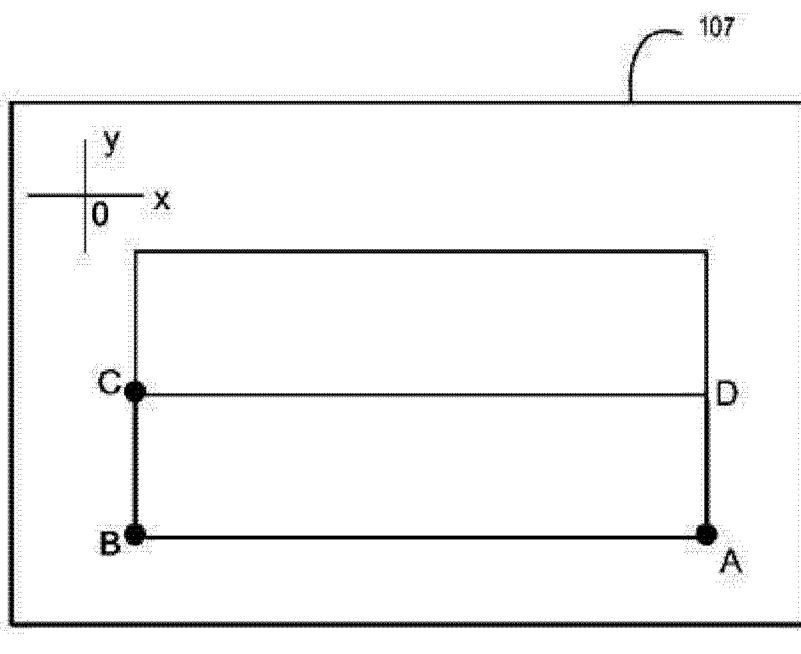
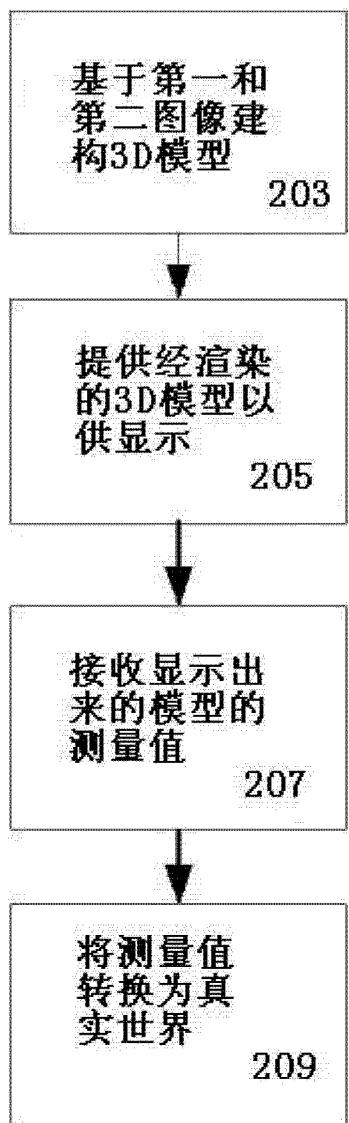


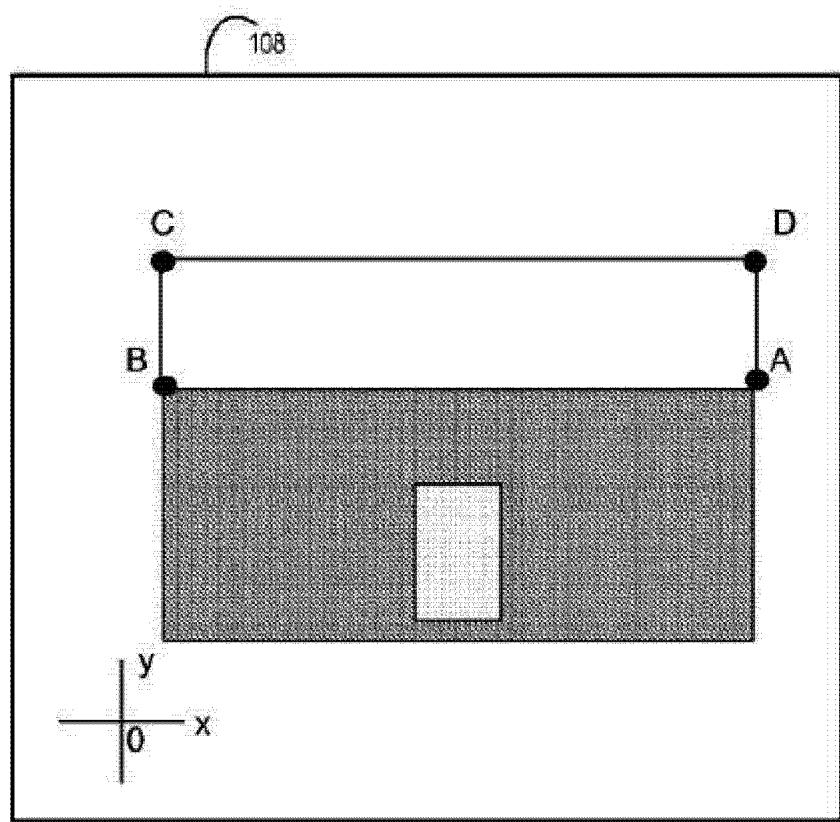
图 1



第一图像

图 3

图 2



第二图像

图 4

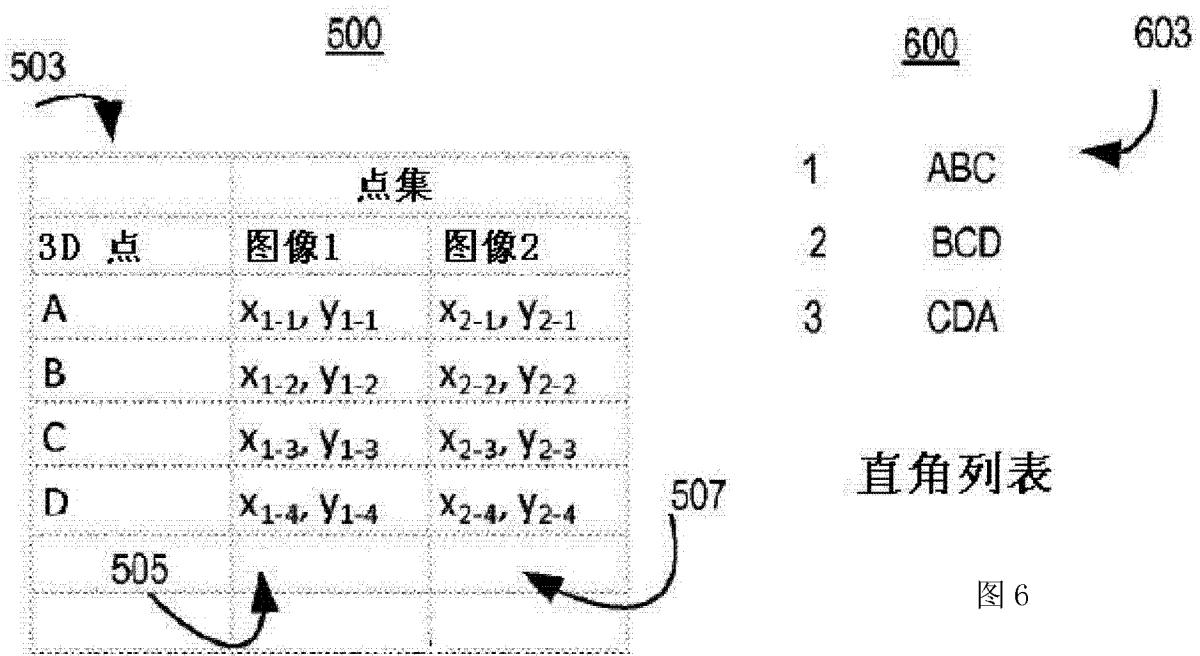
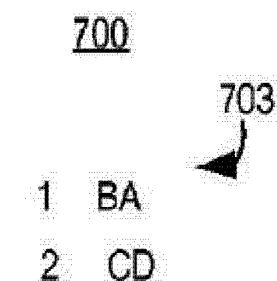


图 5



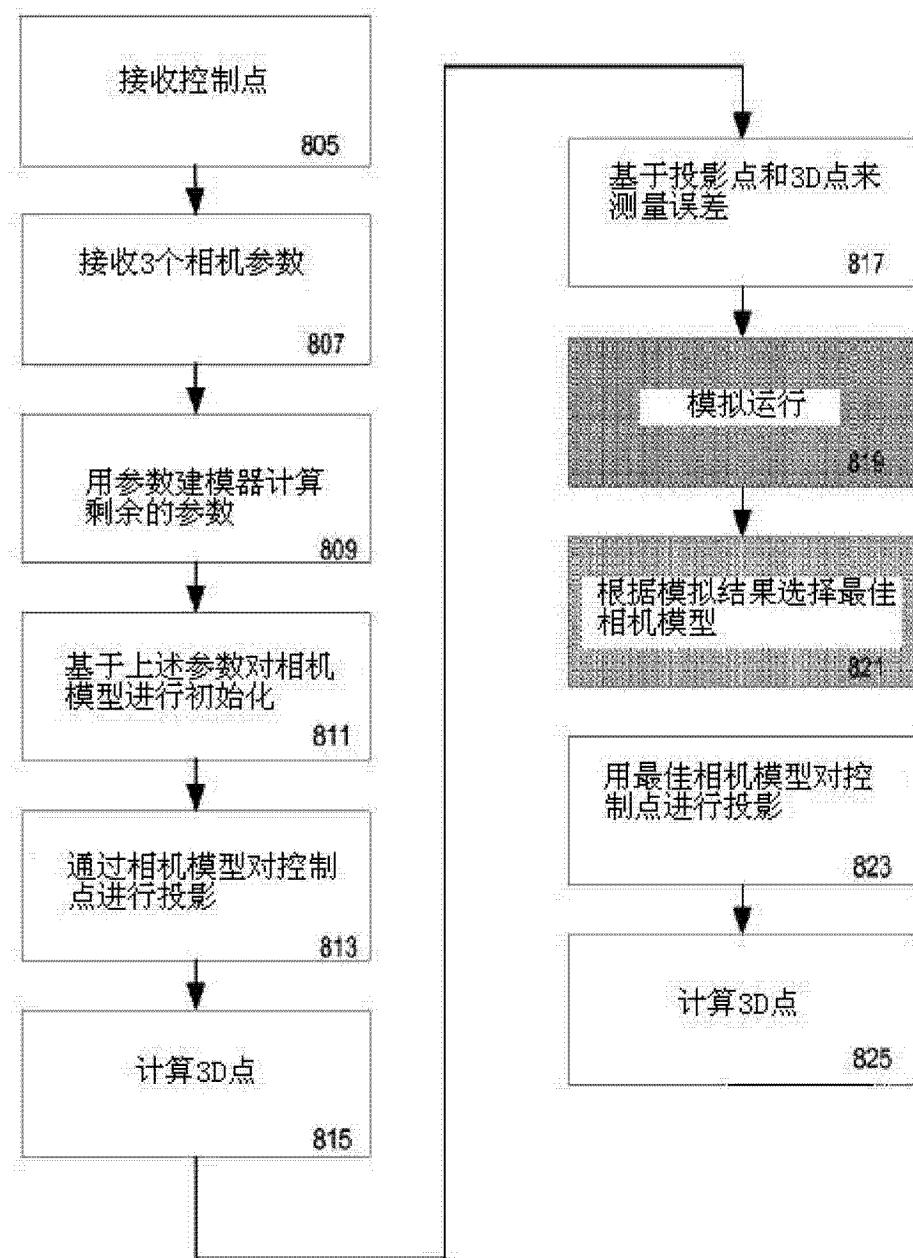


图 8

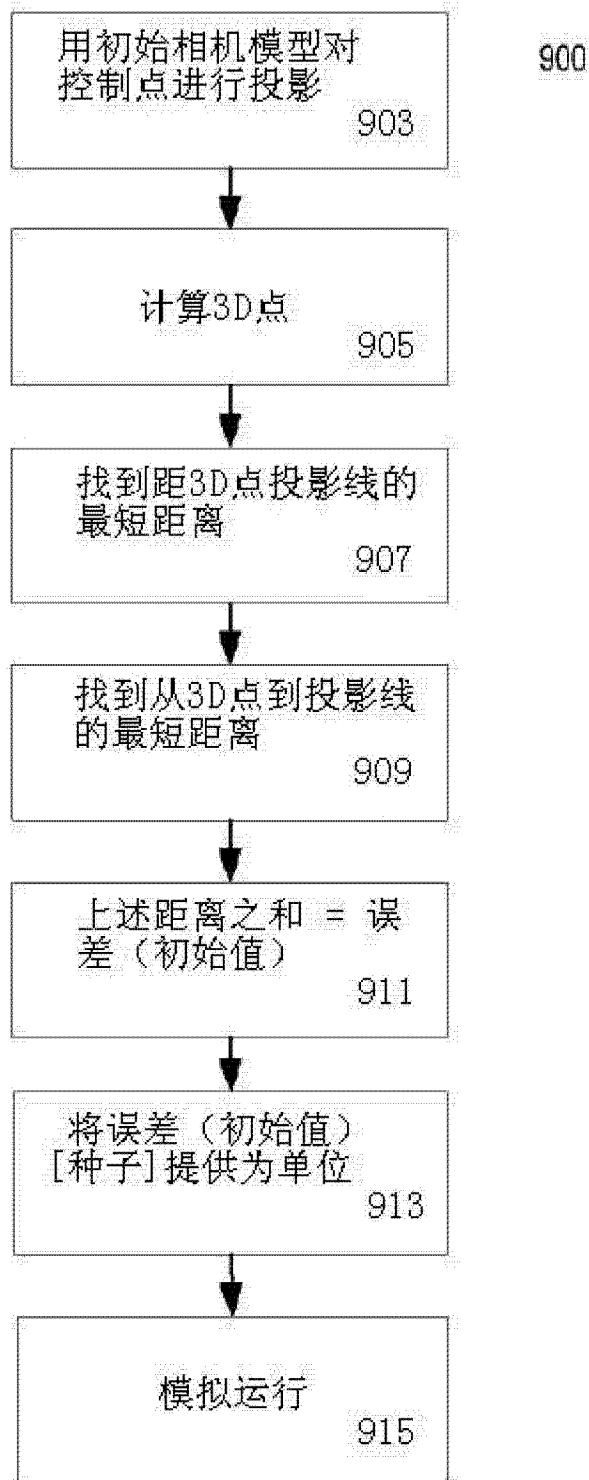


图 9

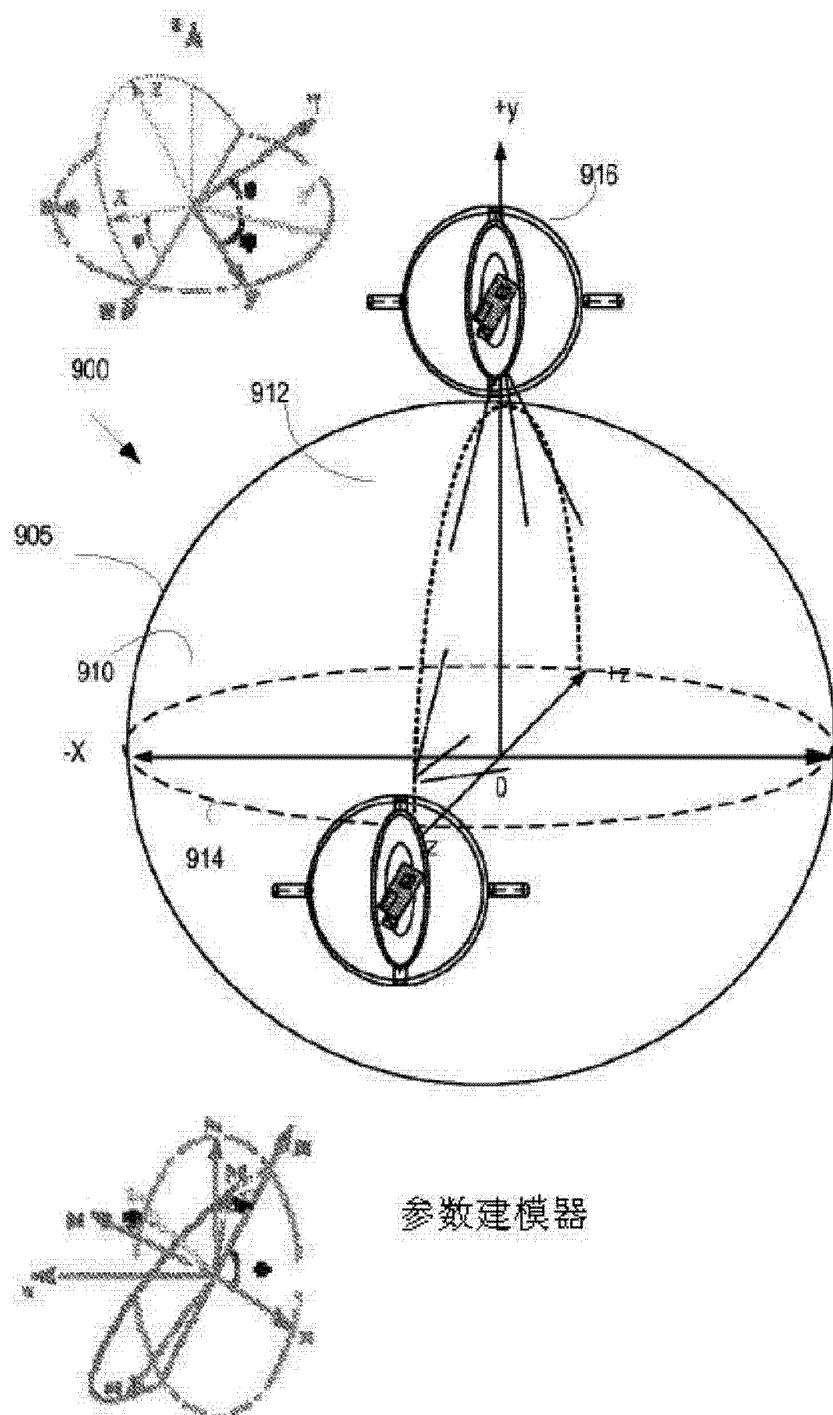


图 10

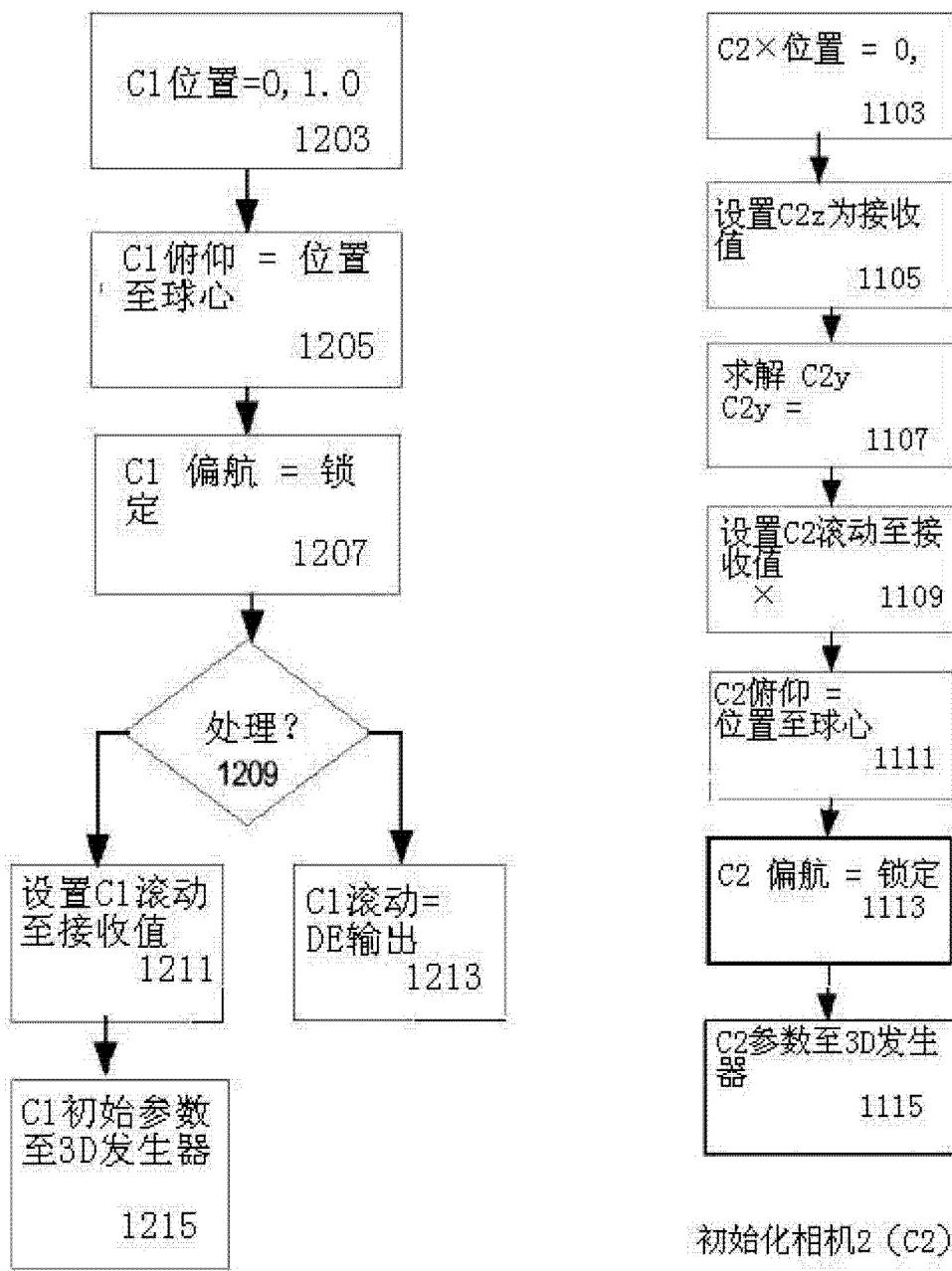


图 11

图 12

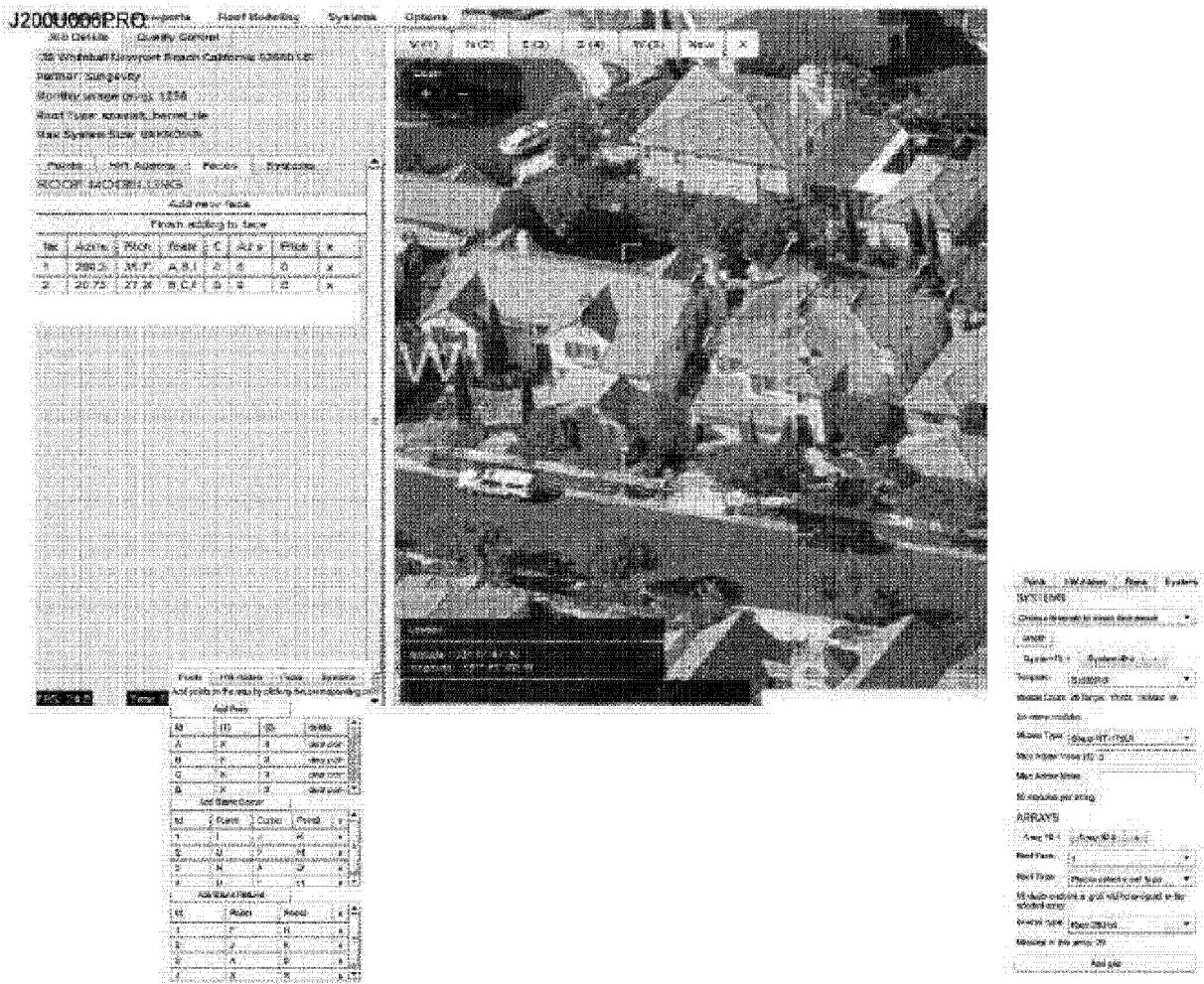


图 13

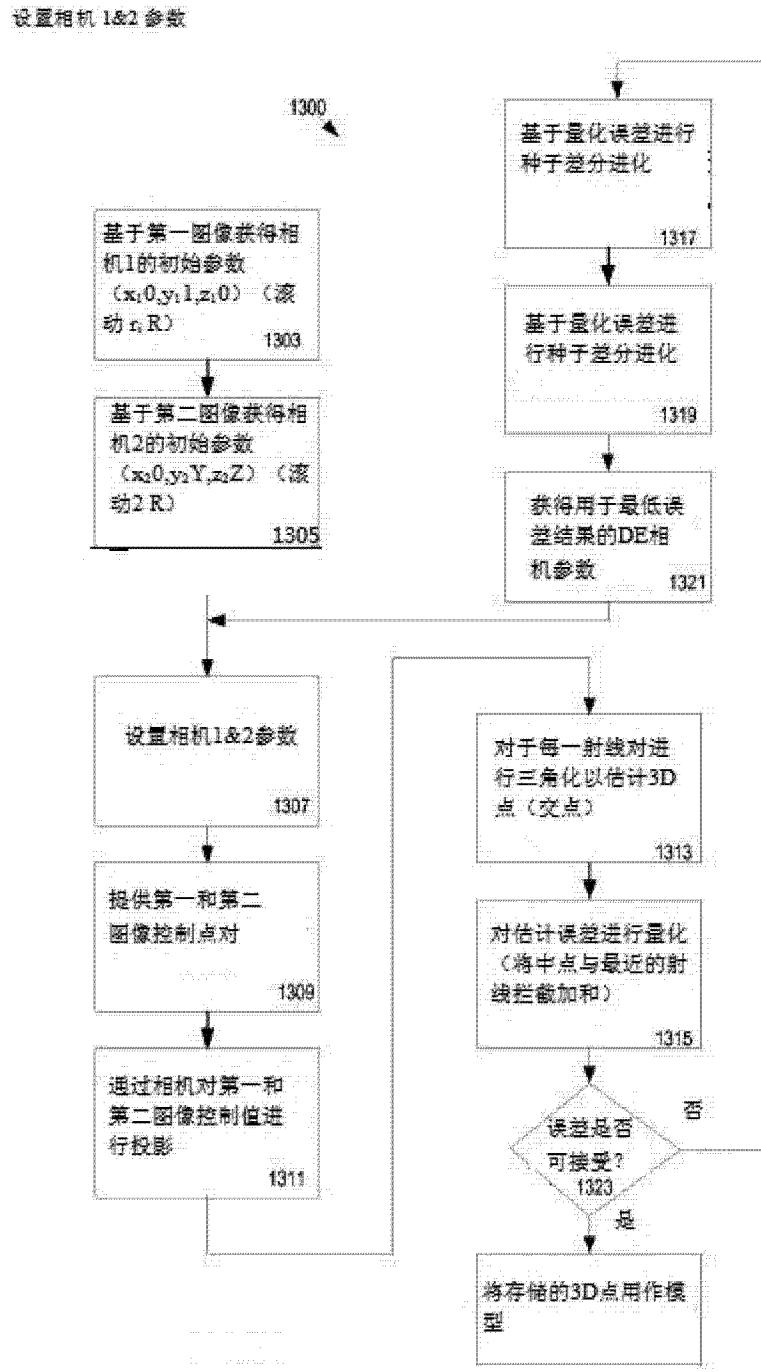


图 14

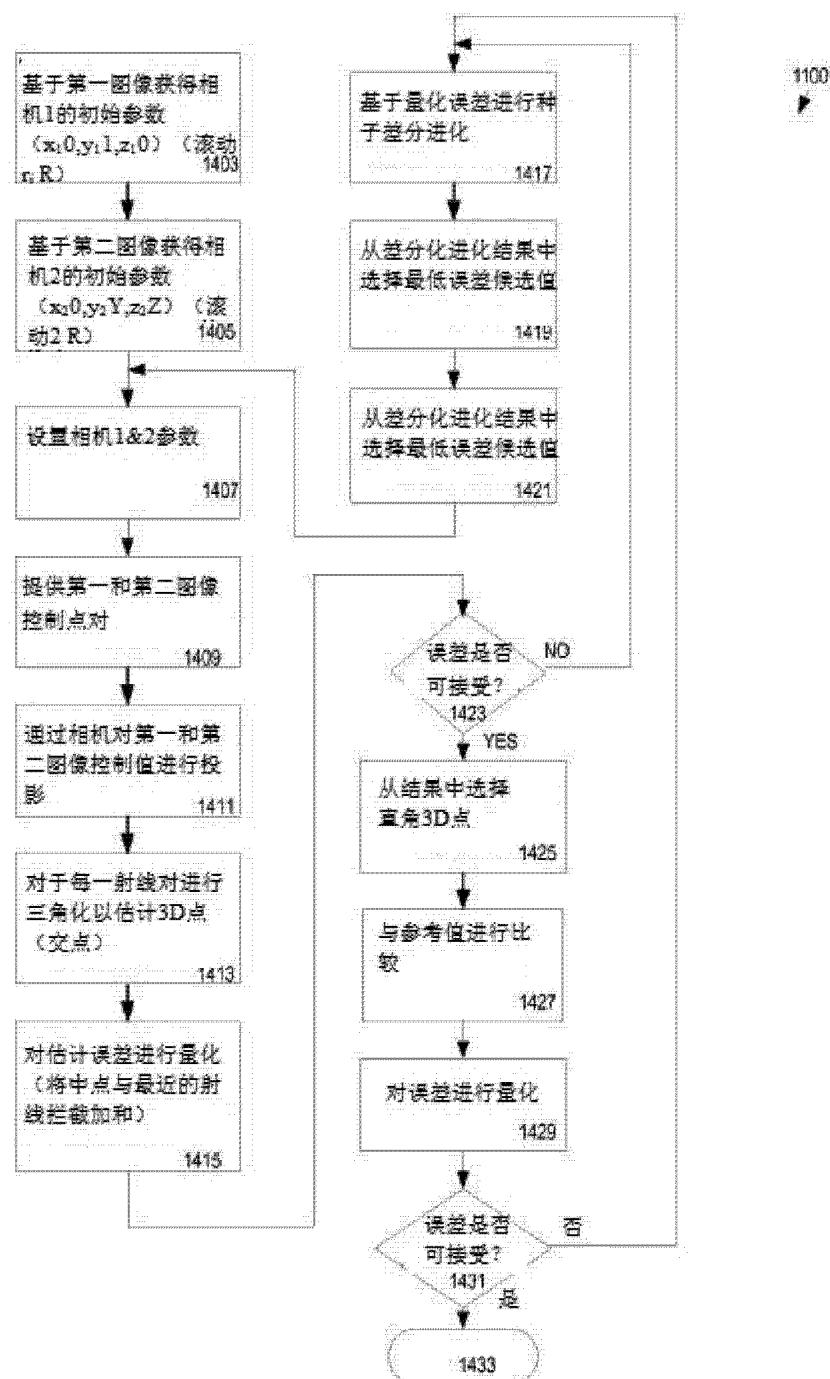
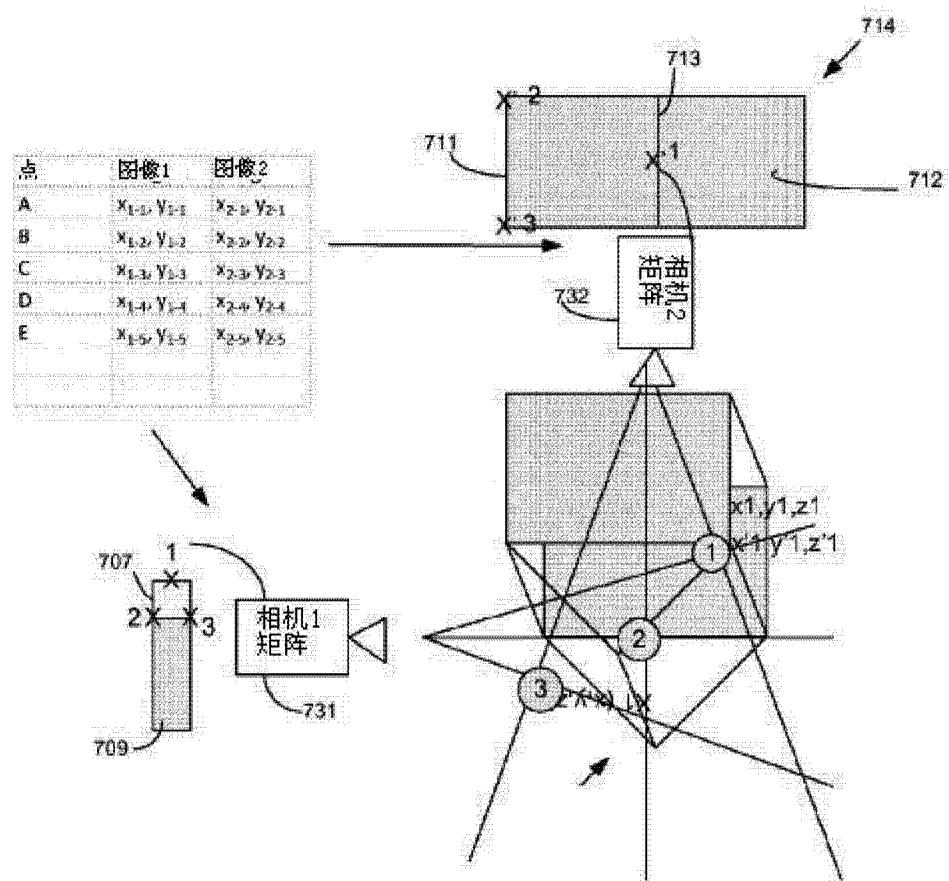


图 15



理想映射

图 16

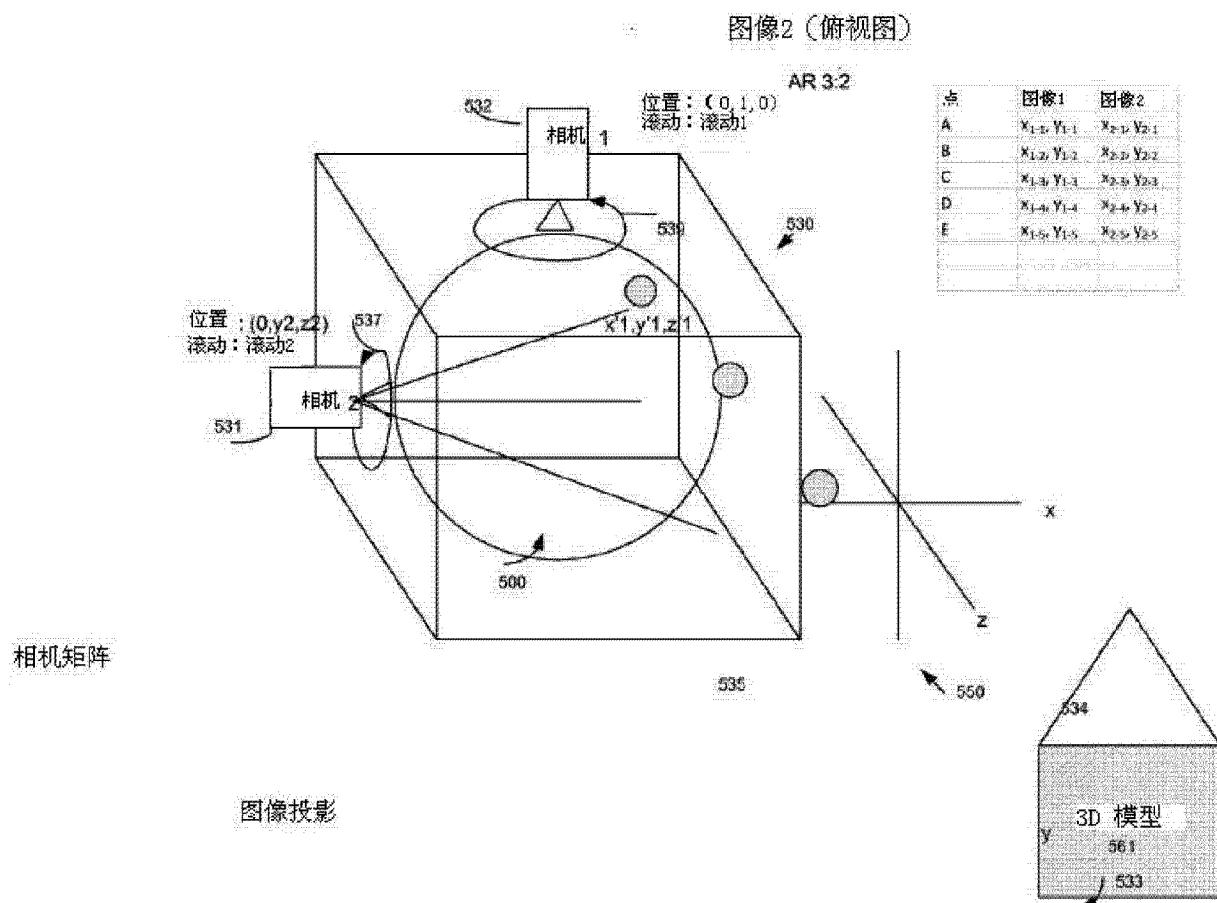


图 17

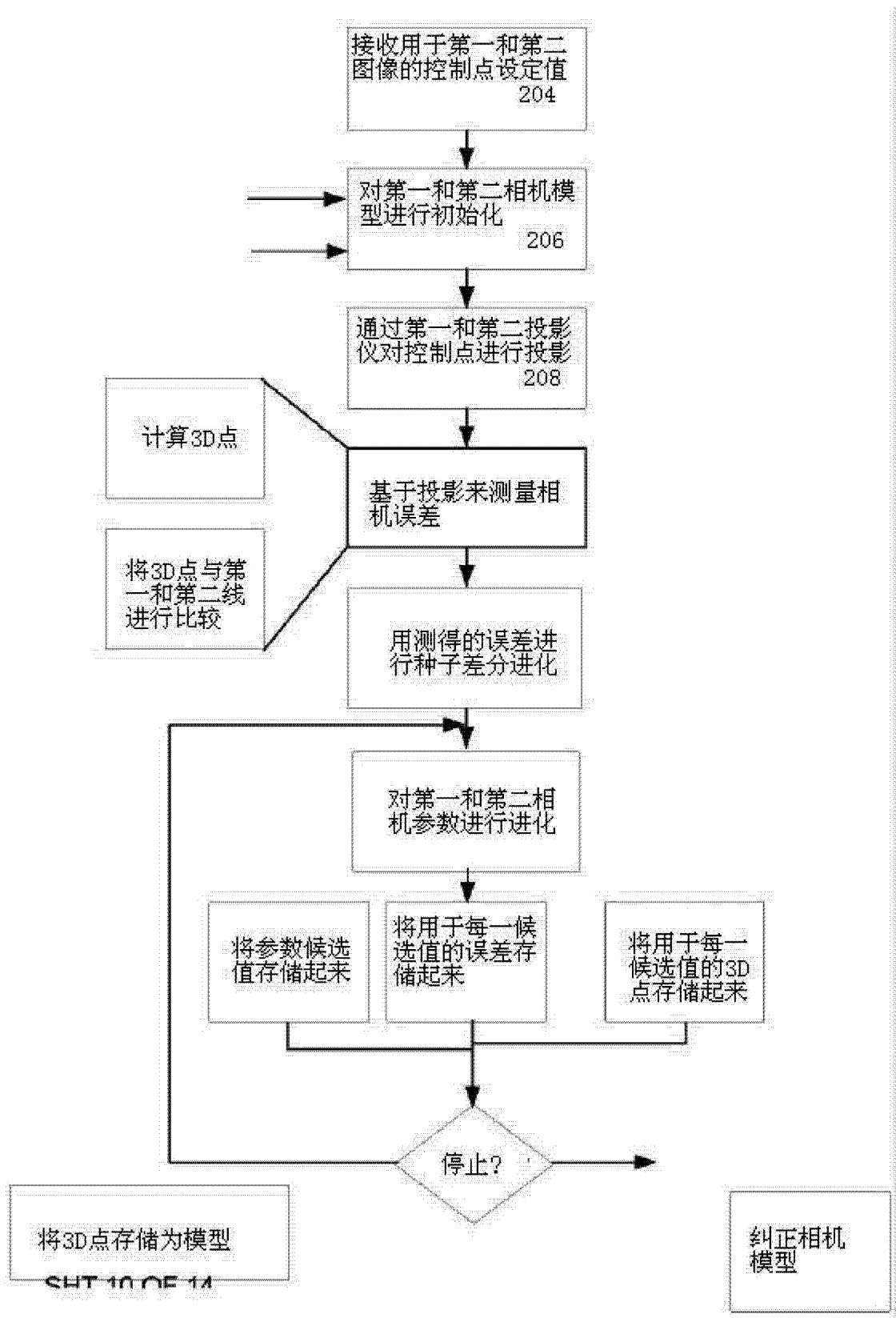


图 18

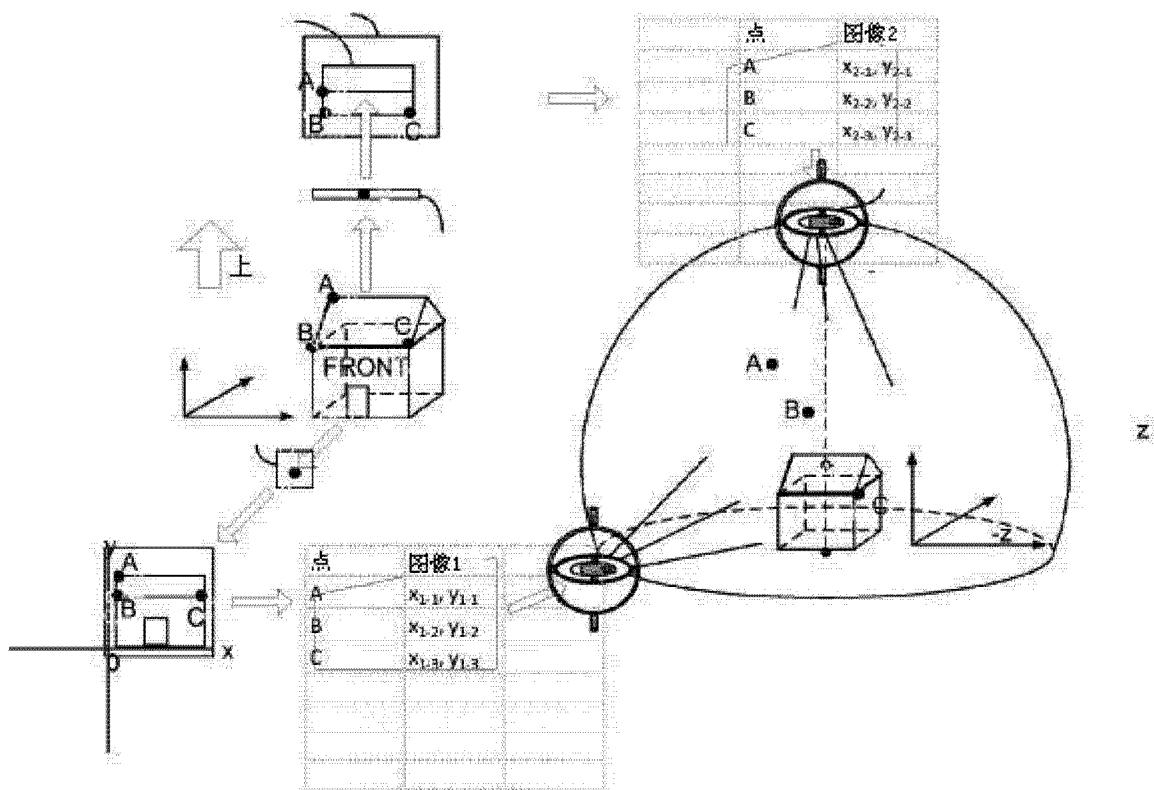


图 19

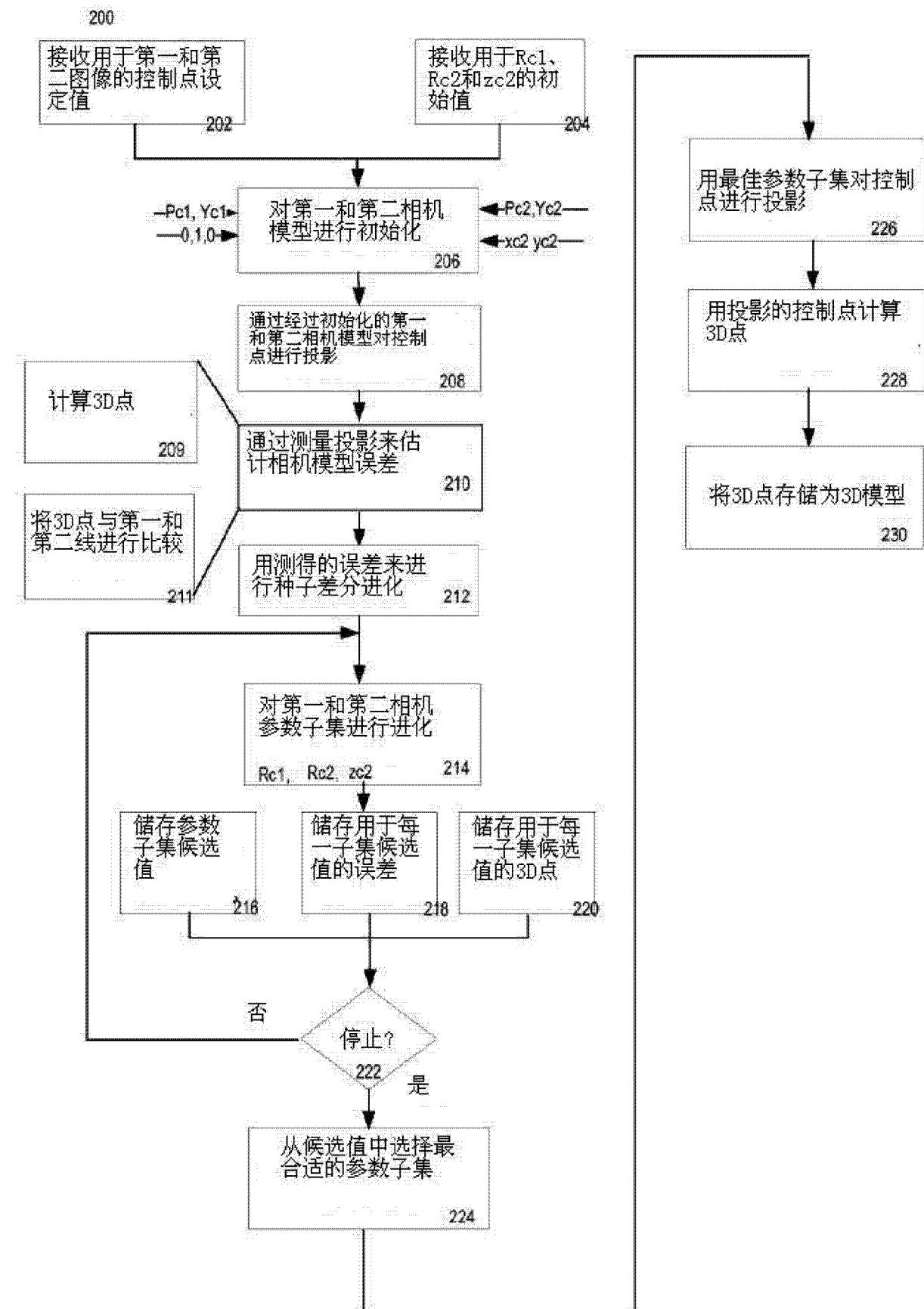


图 20

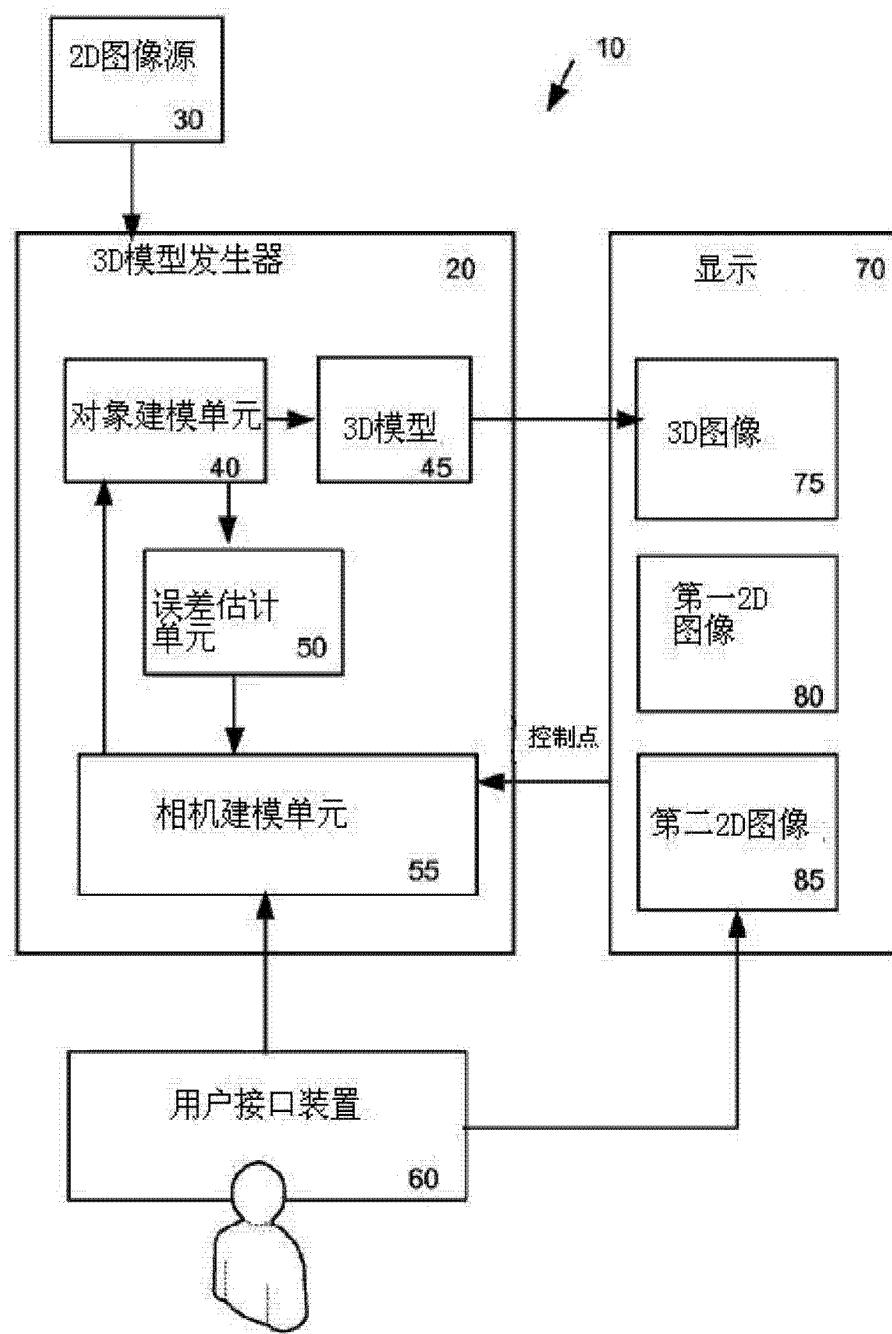


图 21

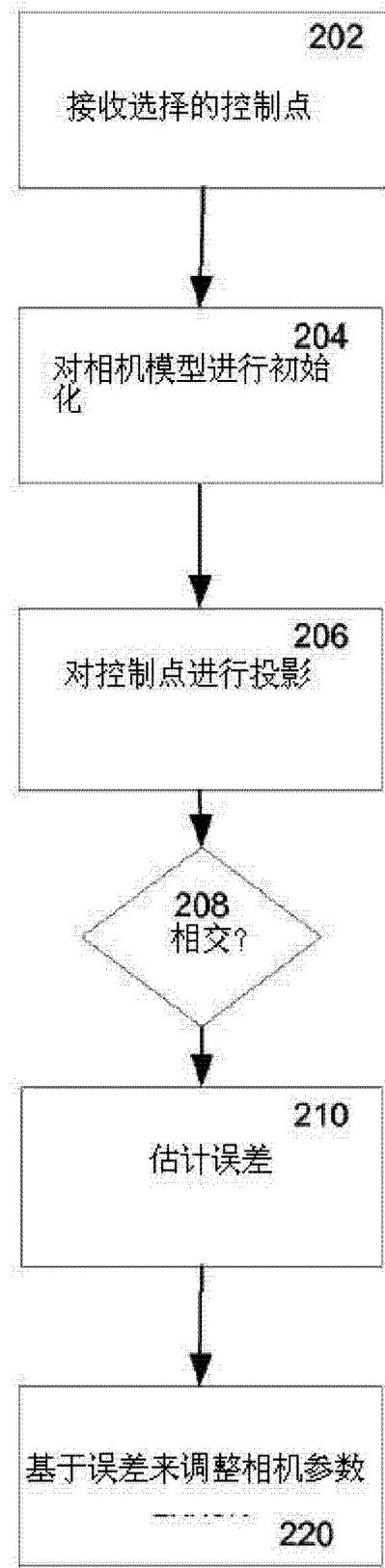


图 22

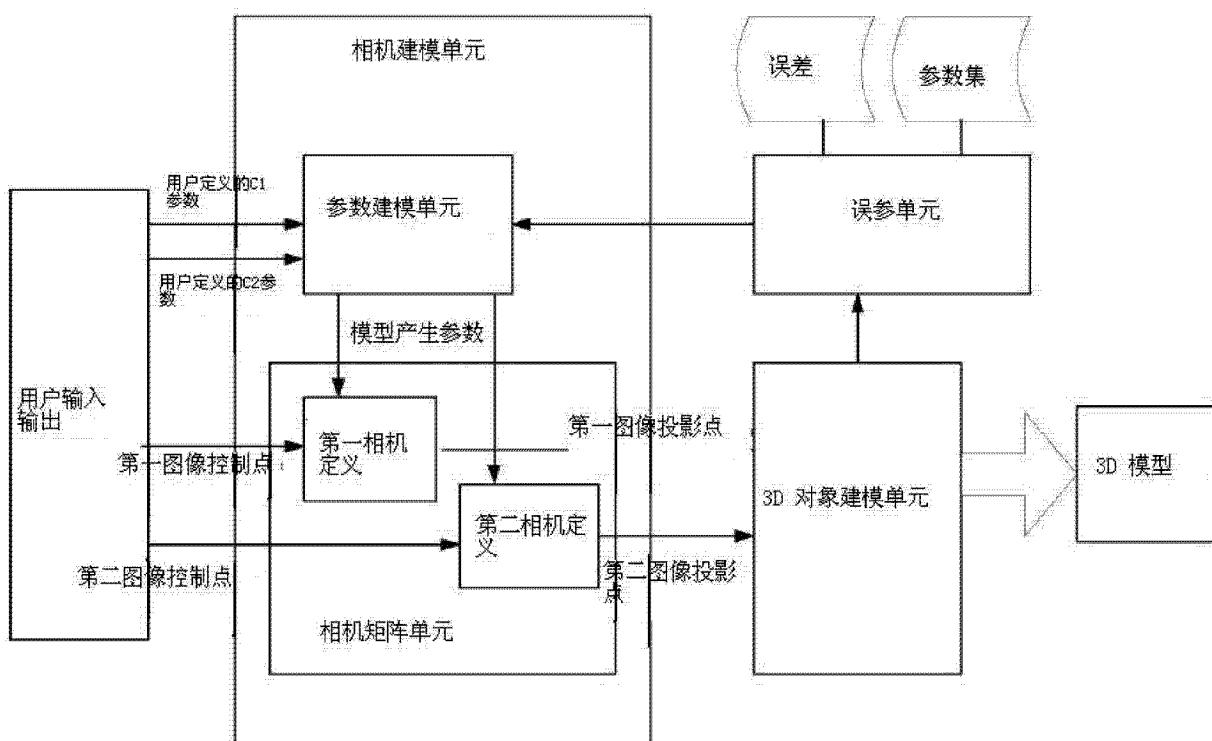


图 23