



(12) 发明专利申请

(10) 申请公布号 CN 113793255 A

(43) 申请公布日 2021.12.14

(21) 申请号 202111056671.6

(22) 申请日 2021.09.09

(71) 申请人 百度在线网络技术(北京)有限公司
地址 100080 北京市海淀区上地十街10号

(72) 发明人 李朋辉 徐静涛 范学峰 崔家华
张柳清 仲亮亮 李国洪 高菲

(74) 专利代理机构 北京市金杜律师事务所
11256

代理人 赵林琳

(51) Int. Cl.

G06T 3/00 (2006.01)

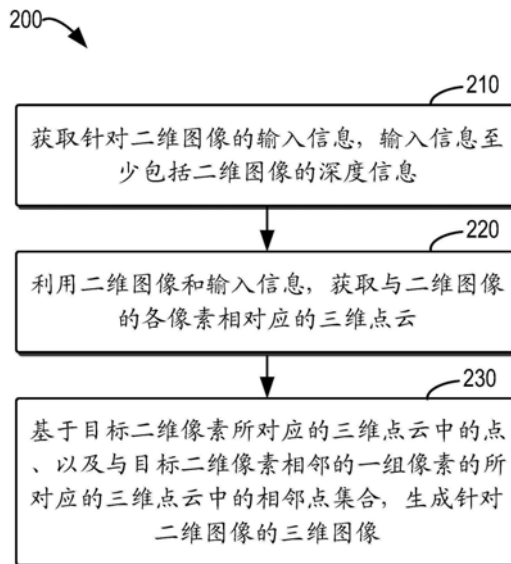
权利要求书3页 说明书9页 附图6页

(54) 发明名称

用于图像处理的方法、装置、设备、存储介质和程序产品

(57) 摘要

根据本公开的示例实施例,提供了用于图像处理的方法、装置、设备、存储介质和程序产品。用于图像处理的方法包括获取针对二维图像的输入信息,输入信息至少包括二维图像的深度信息;利用二维图像和输入信息,获取与二维图像的各像素相对应的三维点云;以及基于目标二维像素所对应的三维点云中的点、以及与目标二维像素相邻的一组像素的所对应的三维点云中的相邻点集合,生成针对二维图像的三维图像。以此方式,能够通过二维照片实现三维模型的构建,具有良好的普遍性和通用性,由此能够大幅提升用户的沉浸式和交互式体验。



1. 一种图像处理方法,包括:
 - 获取针对二维图像的输入信息,所述输入信息至少包括所述二维图像的深度信息;
 - 利用所述二维图像和所述输入信息,获取与所述二维图像的各像素相对应的三维点云;以及
 - 基于目标二维像素所对应的所述三维点云中的点、以及与目标二维像素相邻的一组像素的所对应的所述三维点云中的相邻点集合,生成针对所述二维图像的三维图像。
2. 根据权利要求1所述的方法,其中生成针对所述二维图像的三维图像包括:
 - 基于所述目标二维像素所对应的所述三维点云中的点、以及所述相邻点集合中的至少两个点,获取针对所述目标像素和所述一组像素的平面网格;以及
 - 基于所获取的平面网格,生成针对所述二维图像的所述三维图像。
3. 根据权利要求2所述的方法,其中所述相邻点集合中的所述至少两个点所对应的像素相邻。
4. 根据权利要求1-3中任一项所述的方法,其中所述输入信息还包括以下中的至少一项:
 - 所述二维图像的前景掩码;
 - 所述二维图像的背景掩码;以及
 - 所述二维图像的经修补图像信息。
5. 根据权利要求1-3中任一项所述的方法,其中获取与所述二维图像的各像素相对应的三维点云包括:
 - 对所述二维图像和所述输入信息进行归一化处理,以得到所述三维点云。
6. 根据权利要求1-3中任一项所述的方法,其中获取与所述二维图像的各像素相对应的三维点云包括:
 - 将所述二维图像的像素所在的平面坐标系转换成三维坐标系;以及
 - 基于所述二维图像的像素,生成所述三维坐标系中的所述三维点云。
7. 根据权利要求6所述的方法,其中所述平面坐标系包括像素坐标系或图像坐标系中的至少一种,并且所述三维坐标系包括相机坐标系或世界坐标系中的至少一种。
8. 根据权利要求2或3所述的方法,其中所述平面网格包括三角网格。
9. 根据权利要求1-3或7中任一项所述的方法,其中所述二维图像包括纹理信息或色彩信息中的至少一种信息,并且其中生成针对所述二维图像的三维图像还包括:
 - 基于所述二维图像的像素与所述三维点云中点的对应关系,将所述纹理信息或所述色彩信息中的至少一种信息绘制在所述平面网格上;
 - 利用所绘制的所述平面网格,呈现针对所述二维图像的三维图像。
10. 一种视频处理方法,包括:
 - 根据权利要求1-9中任一项所述的方法,针对视频流中的各帧的二维图像生成对应的三维图像;以及
 - 利用所生成的三维图像,生成三维视频流。
11. 一种图像处理的装置,包括:
 - 输入信息获取模块,被配置为获取针对二维图像的输入信息,所述输入信息至少包括所述二维图像的深度信息;

三维点云获取模块,被配置为利用所述二维图像和所述输入信息,获取与所述二维图像的各像素相对应的三维点云;以及

三维图像生成模块,被配置为基于目标二维像素所对应的所述三维点云中的点、以及与目标二维像素相邻的一组像素的所对应的所述三维点云中的相邻点集合,生成针对所述二维图像的三维图像。

12. 根据权利要求11所述的装置,其中所述三维图像生成模块包括:

平面网格获取模块,被配置为基于所述目标二维像素所对应的所述三维点云中的点、以及所述相邻点集合中的至少两个点,获取针对所述目标像素和所述一组像素的平面网格;以及

三维图像生成子模块,被配置为基于所获取的平面网格,生成针对所述二维图像的所述三维图像。

13. 根据权利要求12所述的装置,其中所述相邻点集合中的所述至少两个点所对应的像素相邻。

14. 根据权利要求11-13中任一项所述的装置,其中所述输入信息还包括以下中的至少一项:

所述二维图像的前景掩码;

所述二维图像的背景掩码;以及

所述二维图像的经修补图像信息。

15. 根据权利要求11-13中任一项所述的装置,其中所述三维点云获取模块包括:

归一化处理模块,被配置为对所述二维图像和所述输入信息进行归一化处理,以得到所述三维点云。

16. 根据权利要求11-13中任一项所述的装置,其中所述三维点云获取模块包括:

三维坐标系转换模块,被配置为将所述二维图像的像素所在的平面坐标系转换成三维坐标系;以及

三维点云生成模块,被配置为基于所述二维图像的像素,生成所述三维坐标系中的所述三维点云。

17. 根据权利要求16所述的装置,其中所述平面坐标系包括像素坐标系或图像坐标系中的至少一种,并且所述三维坐标系包括相机坐标系或世界坐标系中的至少一种。

18. 根据权利要求12或13所述的装置,其中所述平面网格包括三角网格。

19. 根据权利要求11-13或17中任一项所述的装置,其中所述二维图像包括纹理信息或色彩信息中的至少一种信息,并且其中所述三维图像生成模块还包括:

平面网格绘制模块,被配置为基于所述二维图像的像素与所述三维点云中点的对应关系,将所述纹理信息或所述色彩信息中的至少一种信息绘制在所述平面网格上;

三维图像呈现模块,被配置为利用所绘制的所述平面网格,呈现针对所述二维图像的三维图像。

20. 一种视频处理装置,包括:

第二三维图像生成模块,被配置为根据权利要求10中任一项所述方法,针对视频流中的各帧的二维图像生成对应的三维图像;以及

三维视频流生成模块,被配置为利用所生成的三维图像,生成三维视频流。

21. 一种电子设备,所述设备包括:

一个或多个处理器;以及

存储装置,用于存储一个或多个程序,当所述一个或多个程序被所述一个或多个处理器执行,使得所述一个或多个处理器实现根据权利要求1-10中任一项所述的方法。

22. 一种计算机可读存储介质,其上存储有计算机程序,所述程序被处理器执行时实现根据权利要求1-10中任一项所述的方法。

23. 一种计算机程序产品,包括计算机程序,所述计算机程序在被处理器执行时,执行权利要求1-10中任一项所述的方法。

用于图像处理的方法、装置、设备、存储介质和程序产品

技术领域

[0001] 本公开的实施例主要涉及计算机领域,并且更具体地,涉及用于图像处理的方法、装置、设备、存储介质和程序产品。

背景技术

[0002] 图片的二维显示是当前最主要的图片展示形式。二维图像一般是指平面图像。二维图像只有左右、上下四个方向,不存在前后。因此,二维图像只有面积,没有体积。一般来说,二维图像可以是RGB图像或者灰度图像。当用户需要更好的沉浸式或交互式体验时,则需要一种将二维图像转换为三维图像的方法。

发明内容

[0003] 根据本公开的实施例,提供了一种图像处理的方案。

[0004] 在本公开的第一方面中,提供了一种图像处理方法,包括:获取针对二维图像的输入信息,输入信息至少包括二维图像的深度信息;利用二维图像和输入信息,获取与二维图像的各像素相对应的三维点云;以及基于目标二维像素所对应的三维点云中的点、以及与目标二维像素相邻的一组像素的所对应的三维点云中的相邻点集合,生成针对二维图像的三维图像。

[0005] 在本公开的第二方面中,提供了一种视频处理方法,包括根据本公开的第一方面的方法,针对视频流中的各帧的二维图像生成对应的三维图像;以及利用所生成的三维图像,生成三维视频流。

[0006] 在本公开的第三方面中,提供了一种图像处理的装置,包括:输入信息获取模块,被配置为获取针对二维图像的输入信息,输入信息至少包括二维图像的深度信息;三维点云获取模块,被配置为利用二维图像和输入信息,获取与二维图像的各像素相对应的三维点云;以及三维图像生成模块,被配置为基于目标二维像素所对应的三维点云中的点、以及与目标二维像素相邻的一组像素的所对应的三维点云中的相邻点集合,生成针对二维图像的三维图像。

[0007] 在本公开的第四方面中,提供了一种视频处理装置,包括:第二三维图像生成模块,被配置为根据本公开第二方面的方法,针对视频流中的各帧的二维图像生成对应的三维图像;以及三维视频流生成模块,被配置为利用所生成的三维图像,生成三维视频流。

[0008] 在本公开的第五方面中,提供了一种电子设备,设备包括:一个或多个处理器;以及存储装置,用于存储一个或多个程序,当一个或多个程序被一个或多个处理器执行,使得一个或多个处理器实现根据本公开第一方面或第二方面的方法。

[0009] 在本公开的第六方面中,提供了一种计算机可读存储介质,其上存储有计算机程序,程序被处理器执行时实现根据本公开第一方面或第二方面的方法。

[0010] 在本公开的第七方面中,提供了一种计算机程序产品,包括计算机程序,计算机程序在被处理器执行时,执行根据本公开第一方面或第二方面的方法。

[0011] 应当理解,发明内容部分中所描述的内容并非旨在限定本公开的实施例的关键或重要特征,亦非用于限制本公开的范围。本公开的其它特征将通过以下的描述变得容易理解。

附图说明

[0012] 结合附图并参考以下详细说明,本公开各实施例的上述和其他特征、优点及方面将变得更加明显。在附图中,相同或相似的附图标注表示相同或相似的元素,其中:

[0013] 图1示出了本公开的多个实施例能够在其中实现的示例环境的示意图;

[0014] 图2示出了根据本公开的实施例的生成三维图像的过程的流程图;

[0015] 图3示出了根据本公开的一些实施例的生成三维图像过程中基于点云生成三维图像的过程的示意图;

[0016] 图4示出了根据本公开的一些实施例的生成三维视频流的过程的流程图;

[0017] 图5示出了根据本公开的一些实施例的生成三维图像的装置的示意图;

[0018] 图6示出了根据本公开的实施例的生成三维视频流的装置的示意性框图;以及

[0019] 图7示出了能够实施本公开的多个实施例的计算设备的框图。

具体实施方式

[0020] 下面将参照附图更详细地描述本公开的实施例。虽然附图中显示了本公开的某些实施例,然而应当理解的是,本公开可以通过各种形式来实现,而且不应该被解释为限于这里阐述的实施例,相反提供这些实施例是为了更加透彻和完整地理解本公开。应当理解的是,本公开的附图及实施例仅用于示例性作用,并非用于限制本公开的保护范围。

[0021] 在本公开的实施例的描述中,术语“包括”及其类似用语应当理解为开放性包含,即“包括但不限于”。术语“基于”应当理解为“至少部分地基于”。术语“一个实施例”或“该实施例”应当理解为“至少一个实施例”。术语“第一”、“第二”等等可以指代不同的或相同的对象。下文还可能包括其他明确的和隐含的定义。

[0022] 另外,在本文中,应当理解,术语“3D”可以等同于“三维”,术语“2D”可以等同于“二维”,并且“三维图像”可以等同于“三维图像模型”。

[0023] 如前文所简要提及的,需要一种将二维图像转换为三维图像的方法,以满足更好的沉浸式或交互式体验。传统上,主要有两种解决方案。在一种方案中,可以基于三维软件进行设计,用户可以直接利用例如Blender等三维模型来将二维图像生成三维图像。然而,这种方式需要依赖已经存在的三维模型,而且其应用场景只能存在于已知的三维信息的场景中,因此,该方案的应用范围比较有限。

[0024] 在另一种传统方案中,通常基于三角测量原理的硬件扫描方式来将二维图像转变为三维图像。然而,这种方案需要依赖于相应的硬件(基于结构光、tof等)。针对特定尺寸的待测物,需要扫描多次,才能实现三维模型重建。另外,对于特定的扫描硬件,待重建物体尺寸越大,扫描次数越多,相应的计算量也就越大;而且,由于该方案是基于硬件实现的,必须对三维模型重建的物体进行现场扫描,因此该技术方案在使用上具有很大的局限性。

[0025] 为了至少部分地解决上述问题和其他的潜在问题,本文提出一种基于二维图像生成三维图像的方案。在该方案中,能够利用二维照片实现三维模型的构建,具有良好的普遍

性和通用性,无需依赖额外硬件、无需进行现场扫描。而且,还可以利用三维点云中的点和二维图像的像素存在的一一对应关系,将三维图像的色彩信息和纹理信息绘制到三维图像模型上。这样一来,利用该方案生成的三维图像的色彩信息和纹理信息不会丢失,从而产生高质量的3D模型,进而大幅度提升用户的沉浸式和交互式体验。

[0026] 以下将参照附图来具体描述本公开的实施例。

[0027] 图1示出了本公开的多个实施例能够在其中实现的示例环境100的示意图。在该示例环境100中,可以将二维图像110输入到计算设备120中,在计算设备120进行相应计算(例如将在下文介绍的深度计算和图像修复等),生成三维图像130。

[0028] 在一些实施例中,二维图像110可以是如前所述的平面图像。需要说明的是,该二维图像110可以是具有前景和背景的图像,也可以是不具有前景或者不明显具有前景的图像。

[0029] 需要说明,二维图像中的前景和背景是图像技术领域常规的概念。一般来说,图像的前景是看起来最接近观看者的视觉平面,而图像中的背景是距离观看者较远的合成中的平面。例如,在一副人物二维图像中,人物通常是图像的前景,而图像中其他的部分往往被称为背景。而对于一些图像,例如蓝天白云图像,则其没有前景或没有明显的前景,这也属于二维图像的一种。

[0030] 在本公开的不同实施例中,可以对于不同种类的二维图像110通过计算设备120进行不同的处理。这将在下文中进行详细介绍。

[0031] 相应地,三维图像130(也可以称为“三维图像模型”)通常是指具有高度、宽度和纵深的立体图像。在本公开的实施例中,三维图像130可以是基于二维图像110而获得的能够让用户更加沉浸式和交互式体验的图像。

[0032] 在一些实施例中,计算设备120可以包括台式电脑、平板电脑、个人数字助理(PDA)、服务器、主机等、或能够进行有线数据通信或无线数据通信的任何其他处理器使能设备,本公开对此不做限制。

[0033] 为了更清楚地理解本公开的实施例所提供的生成针对二维图像的三维图像的方案,将参照图2来进一步描述本公开的实施例。图2示出了根据本公开的实施例的生成三维图像的过程200的流程图。过程200可以由图1的计算设备120来实现。为便于讨论,将结合图1来描述过程200。

[0034] 在框210,获取针对二维图像110的输入信息,输入信息至少包括二维图像110的深度信息。深度信息可以包括二维图像110的深度图像,深度图像可以通过将二维图像110输入深度模型而获取。应当理解,深度图像的分辨与与二维图像110的分辨率是一致的。

[0035] 在一些实施例中,深度模型可以部署在计算设备120中,也可以部署在不同于计算设备120中的其他计算设备中,只要能够获得所需要的深度信息作为框220的的输入即可,本公开对此不做限制。

[0036] 在一些实施例中,输入信息还可以包括二维图像110的前景掩码、背景掩码(前景掩码和背景掩码也可以统称为“前景背景掩码”)和经修补图像信息等。

[0037] 在这样的实施例中,二维图像110通常具有明显的前景和背景。可以通过分割模型获取前景掩码和背景掩码,并且通过图像修补(inpainting)模型来获取经修补图像(即经修补图像信息)。这样,输入信息可以包括深度图像、经修补图像、前景掩码和背景掩码。

[0038] 需要说明,分割模型和图像修补模型可以部署在计算设备120中,也可以部署在不同于计算设备120中的其他计算设备中,只要能够获得所需要的深度信息作为框220的输入即可,本公开对此不做限制。

[0039] 在框220,利用二维图像110和输入信息,获取与二维图像110的各像素相对应的三维点云。可以看出,在框220,二维图像110也可以被视为一种输入信息。

[0040] 在一些实施例中,可以对二维图像110和输入信息进行归一化处理,以得到三维点云。归一化是一种简化计算的方式,一种是把数变为(0,1)之间的小数,或者将有量纲的表达式,经过变换,化为无量纲的表达式,成为纯量。以此方式,可以方便数据处理,减少计算量。

[0041] 在一些实施例中,可以将二维图像110的像素所在的平面坐标系转换成三维坐标系,并且基于二维图像110的像素,生成三维坐标系中的三维点云。平面坐标系可以包括像素坐标系或图像坐标系中的至少一种,并且三维坐标系包括相机坐标系或世界坐标系中的至少一种。以此方式,可以以快捷的方式实现二维图像110到三维图像的转变。

[0042] 在上文没有前景或没有明显前景的二维图像110的实施例中,可以直接基于深度图像进行三维变换(例如数学投影变换)来生成归一化处理后的三维点云。需要说明的是,上述三维变换的方式仅仅是示例性的,任何能够实现二维到三维变换的方式均是可行的,本公开对此不做限制。

[0043] 在具有前景和背景的二维图像110中的实施例中,如前所述,输入信息还可以包括经修补图像、前景掩码和背景掩码。下面,将以平面坐标系为像素坐标系、三维坐标系为相机坐标系为例,详细介绍通过如下等式(1)-(6)数学投影变换的方式获取与二维图像110的各个像素相对应的三维点云的过程。

$$[0044] \quad Z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{d_x} & 0 & u_0 \\ 0 & \frac{1}{d_y} & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} \quad (1)$$

$$[0045] \quad Z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & u_0 & 0 \\ 0 & f_y & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} \quad (2)$$

$$[0046] \quad f_x = \frac{f}{d_x} \quad (3)$$

$$[0047] \quad f_y = \frac{f}{d_y} \quad (4)$$

$$[0048] \quad X_c = \frac{u - u_0}{f_x} Z_c \quad (5)$$

$$[0049] \quad Y_c = \frac{v - v_0}{f_y} Z_c \quad (6)$$

[0050] 其中, Z_c 是深度图像中的像素的高度信息, u 和 v 分别为二维像素坐标系中的坐标位置, u_0 、 v_0 为图像中心, f 为虚拟摄像机的焦距,单位是mm; d_x 、 d_y 为像元尺寸。

[0051] 在该实施例中,通过等式(1)和等式(2)的数学变换,可以得到等式(3)和等式(4)所示的x轴和y轴上的归一化焦距 f_x 和 f_y 。在得到归一化焦距 f_x 和 f_y 之后,可以基于小孔成像

原理,通过等式 (5) 和等式 (6) 得到基于小孔成像的相机坐标系中像素的X轴和Y轴的坐标 X_c 和 Y_c 。

[0052] 需要说明,等式 (1) 和等式 (2) 的数学变换方式仅仅是示例性的,本领域技术人员可以根据实际需要进行调整。并且坐标 X_c 和 Y_c 的确定方式也是示意性的,本领域技术人员可以根据实际需要调整等式 (3) 和等式 (4) 来以其他方式确定归一化焦距,本公开对此不做限制。

[0053] 还需要说明,上述等式 (5) 和等式 (6) 是以 u_0 、 v_0 为图像中心来得到X轴和Y轴的坐标 X_c 和 Y_c 的。本领域技术人员还可以采用其他的任意像素点的位置来计算相关像素的坐标,本公开对此不做限制。

[0054] 进一步,可以设虚拟相机的水平和垂直视场角为 φ_h 和 φ_v ,则可以得到:

$$[0055] \quad f_x = \frac{f}{d_x} = \frac{\frac{w}{2}}{\tan \frac{\varphi_h}{2}} \quad (7)$$

$$[0056] \quad f_y = \frac{f}{d_y} = \frac{\frac{h}{2}}{\tan \frac{\varphi_v}{2}} \quad (8)$$

[0057] 结合等式 (7) 和 (8),可以将图像像素坐标系转换到基于小孔成像的相机坐标系,即设置特定的渲染视场角,即可由二维像素坐标系中的像素,生成三维相机坐标系中的点云。由于深度图范围为 $0 \sim 2^{16} - 1$,为了保证普遍性,因此将深度图范围映射到 $0 \sim 1$ 。这样一来,可以通过上述简便的步骤即可基于二维图像110得到三维图像。和基于硬件扫描获取的点云的方式相比,由于硬件不可避免的存在偏差,点云容易存在丢失(NAN值),而根据深度学习获取的点云不存在点云丢失问题,因此可以呈现更高质量的三维图像模型。

[0058] 需要说明,上述二维图像110变换为三维图像的方式仅仅是示例性的,本领域技术人员还可以采用任何合适的方式实现上述变换,或者对上述提及的参数进行相应调整来实现上述变换,本公开对此不做限制。

[0059] 在框230,基于目标二维像素所对应的三维点云中的点、以及与目标二维像素相邻的一组像素的所对应的三维点云中的相邻点集合,生成针对二维图像110的三维图像。

[0060] 在该实施例中,目标二维像素可以是二维图像110中的任意像素,例如可以是如图3所示的像素X。一组像素是与目标二维像素相邻的像素的集合,例如可以是如图3所示的像素A-H的集合。由于二维像素和三维点云中的点有一一对应的关系,可以理解,目标二维像素和相邻的一组像素对应的三维点云中的点在空间上具有一定的位置关系。基于目标二维像素所对应的三维点云中的点以及相邻的一组像素所对应的三维点云中的相邻点集合,生成针对二维图像110的三维图像。

[0061] 下面将结合图3介绍框230进一步实现的示例性步骤。图3示出了根据本公开的一些实施例的基于点云生成三维图像的过程300的示意图。在图3中,如上文所述,目标二维像素可以是X,与目标二维像素相邻的一组像素可以像素A-H的集合。对应地,相邻点集合可以是该一组像素A-H在三维点云中对应的点的集合。

[0062] 在一些实施例中,可以基于目标二维像素X所对应的三维点云中的点、以及相邻点集合中的至少两个点,获取针对所述目标像素和所述一组像素的平面301,并且基于所获取的平面网格301,生成针对所述二维图像110的所述三维图像。

[0063] 在该实施例中,生成平面网格301的过程可以包括编码和面片化过程。也就是说,

可以基于像素与点云的对应关系,对三维坐标系中的点云进行编码和面片化,从而生成针对二维图像的三维图像模型。以此方式,可以以平面网格301的方式实现三维图像建模,即以一张二维图像即可完成三维建模,不必依赖额外硬件,更不必进行现场扫描,具备较好的实用价值。

[0064] 在该实施例中,参考图3,由于通常确定平面仅仅需要三个点即可实现,因此基于目标二维像素X所对应的三维点云中的点、以及一组像素中所对应的三维点云中的任意两个点就可以实现面片化(即平面化)。在这种情形下,面片可以是三角面片。对应地,平面网格301可以包括至少一个三角网格。

[0065] 在一个实施例中,更具体地,参考图3,在实现三角面片的过程中,所选取的相邻点集合中的两个点所对应的二维图像110中的像素是相邻的。例如,在确定目标二维像素X之后,可以选取相邻的一组像素中像素A和像素B所对应的3D点云中的点,并且基于像素和点云的对应关系,将像素A、像素X和像素B在3D点云中进行面片化,得到三角面片310。

[0066] 同样地,可以将像素X、像素B和像素C进行三角面片化得到三角面片320,将像素X、像素E和像素C三角面片化得到三角面片330,将像素X、像素E和像素H进行三角面片化得到三角面片340,将像素X、像素G和像素H进行三角面片化得到三角面片350,以此类推,得到完整的由8个三角面片组成的平面网格301。进一步地,将上述方法拓展到整个二维图像110的各个像素,即可得到三维图像模型。以此方式,可以得到无损的三维图像,从而大大提升用户的交互体验和沉浸体验。

[0067] 需要说明,上述三角面片化的过程仅仅是示例性的,本领域技术人员还可以通过其他任意合适的方式来实现三维图像建模,本公开对此不做限制。例如,在精度要求相对较低的情况下,还可以不必采用三角面片的方式来进行上述建模。也就是说,在选择一组像素中时,可以不必采取相邻的像素,这样形成的平面网格301则可以由相对较少的面片组成。以此方式,在牺牲部分精度,但是计算量会大大降低,可以适用于一些低精度要求的三维图像模型中。

[0068] 在一些实施例中,继续参考图3,二维图像110通常包括纹理信息或色彩信息中的至少一种信息。在这样的实施例中,可以基于二维图像110的像素与三维点云中点的对应关系,将纹理信息或所述色彩信息中的至少一种信息绘制在平面网格301上,并且可以利用所绘制的平面网格301,呈现针对二维图像110的三维图像。纹理信息是一种反映图像中同质现象的视觉特征,纹理信息体现了物体表面的具有缓慢变化或者周期性变化的表面结构组织排列属性。不同于灰度、颜色等图像特征,纹理通过像素及其周围空间邻域的灰度分布来表现。而色彩信息则对应于图像的灰度、颜色等图像特征。

[0069] 在该实施例中,基于二维图像110的深度图像本身与二维图像110存在严格的一一对应关系,也就是意味着三维点云与纹理信息和色彩信息存在一一对应关系。因此,可以对已经得到三维图像模型进行纹理信息或色彩信息中的至少一种信息进行渲染,从而得到完整的三维图像模型。事实上,该步骤可以理解为将纹理信息和色彩信息中的至少一种信息贴到平面网格301的对应面片上的过程,也即对三维图像模型进行上色和渲染的过程。

[0070] 在基于硬件扫描获取点云的方案中,由于硬件不可避免的存在偏差,点云与纹理信息和色彩信息存在一对多的对应关系,纹理信息与点云信息不可能完全一一对应,最终会引起纹理质量的下降。而在该实施例中,由于这种空间强相关关系,使得纹理信息和色彩

信息不会丢失,从而得到高质量的、能完整体现二维图像110立体化信息的三维图像模型。

[0071] 图4示出了根据本公开的一些实施例的生成三维视频流的过程400的流程图。过程400可以在如图1所示的计算设备120中实现,也可以在其他任意合适的计算设备中实现。为便于讨论,将结合图1来描述过程400。

[0072] 在框410,针对视频流中的各帧的二维图像110生成对应的三维图像。在一些实施例中,将二维图像110生成三维图像的过程可以根据如图2所示的方法来实现或根据其他任意合适的方法实现。应当理解,在需要对视频流的部分帧生成三维图像时,也可以选取部分特定帧进行三维图像建模,本公开对此不做限制。

[0073] 在框420,利用所生成的三维图像,生成三维视频流。以此方式,可以进一步在生成三维图像的基础上得到三维视频流,进一步提升用户的沉浸式体验和交互体验。

[0074] 图5示出了根据本公开的一些实施例的生成三维图像的装置500的示意图。装置500包括输入信息获取模块510、三维点云获取模块520和三维图像生成模块530。

[0075] 输入信息获取模块510被配置为获取针对二维图像110的输入信息,所述输入信息至少包括二维图像110的深度信息。如前所述,输入信息还可以包括以下中的至少一项:二维图像的前景掩码;二维图像的背景掩码;以及二维图像的经修补图像信息。

[0076] 三维点云获取模块520被配置为利用二维图像110和输入信息,获取与二维图像110的各像素相对应的三维点云。

[0077] 在一些实施例中,三维点云获取模块520可以包括归一化处理模块,归一化处理模块被配置为对二维图像和输入信息进行归一化处理,以得到三维点云。

[0078] 在一些实施例中,三维点云获取模块520还可以包括三维坐标系转换模块和三维点云生成模块,三维坐标系转换模块被配置为将二维图像110的像素所在的平面坐标系转换成三维坐标系,并且三维点云生成模块被配置为基于二维图像110的像素,生成三维坐标系中的三维点云。

[0079] 在一些实施例中,平面坐标系可以包括像素坐标系或图像坐标系中的至少一种,并且三维坐标系可以包括相机坐标系或世界坐标系中的至少一种。

[0080] 三维图像生成模块530被配置为基于目标二维像素所对应的三维点云中的点、以及与目标二维像素相邻的一组像素的所对应的三维点云中的相邻点集合,生成针对二维图像110的三维图像。

[0081] 在一些实施例中,三维图像生成模块530还可以包括平面网格获取模块和三维图像生成子模块,平面网格获取模块被配置为基于目标二维像素所对应的三维点云中的点、以及相邻点集合中的至少两个点,获取针对目标像素和一组像素的平面网格,并且三维图像生成子模块被配置为基于所获取的平面网格,生成针对二维图像110的三维图像。

[0082] 在一些实施例中,相邻点集合中的至少两个点所对应的像素可以是相邻的。

[0083] 在一些实施例中,其中三维图像生成模块530还包括平面网格绘制模块和三维图像呈现模块,平面网格绘制模块被配置为基于二维图像110的像素与三维点云中点的对应关系,将纹理信息或色彩信息中的至少一种信息绘制在平面网格上,并且三维图像呈现模块被配置为利用所绘制的平面网格,呈现针对二维图像110的三维图像。

[0084] 图6示出了根据本公开的实施例的生成三维视频流的装置600的示意性框图。装置600包括第二三维图像生成模块610和三维视频流生成模块620。装置600可以在如图1所示

的计算设备120中实现,也可以在其他任意合适的设备中实现。为便于讨论,将结合图1来描述过程600。

[0085] 第二三维图像生成模块610被配置为针对视频流中的各帧的二维图像110生成对应的三维图像。生成三维图像的步骤可以利用如前所述的装置500来实现。

[0086] 三维视频流生成模块620被配置为利用所生成的三维图像,生成三维视频流。以此方式,可以进一步在生成三维图像的基础上得到三维视频流,进一步提升用户的沉浸式体验和交互体验。

[0087] 图7示出了能够实施本公开的多个实施例的计算设备700的框图。设备700可以用于实现图1的计算设备120。如图所示,设备700包括中央处理单元(CPU)701,其可以根据存储在只读存储器(ROM)702中的计算机程序指令或者从存储单元708加载到随机访问存储器(RAM)703中的计算机程序指令,来执行各种适当的动作和处理。在RAM 703中,还可存储设备700操作所需的各种程序和数据。CPU701、ROM 702以及RAM 703通过总线704彼此相连。输入/输出(I/O)接口705也连接至总线704。

[0088] 设备700中的多个部件连接至I/O接口705,包括:输入单元706,例如键盘、鼠标等;输出单元707,例如各种类型的显示器、扬声器等;存储单元708,例如磁盘、光盘等;以及通信单元709,例如网卡、调制解调器、无线通信收发机等。通信单元709允许设备700通过诸如因特网的计算机网络和/或各种电信网络与其他设备交换信息/数据。

[0089] 处理单元701执行上文所描述的各个方法和处理,例如过程200和过程400中的一个或多个。例如,在一些实施例中,过程200和过程400中的一个或多个可被实现为计算机软件程序,其被有形地包含于机器可读介质,例如存储单元708。在一些实施例中,计算机程序的部分或者全部可以经由ROM 702和/或通信单元709而被载入和/或安装到设备700上。当计算机程序加载到RAM 703并由CPU 701执行时,可以执行上文描述的过程200和过程400中的一个或多个的一个或多个步骤。备选地,在其他实施例中,CPU 701可以通过其他任何适当的方式(例如,借助于固件)而被配置为执行过程200和过程400中的一个或多个。

[0090] 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:场可编程门阵列(FPGA)、专用集成电路(ASIC)、专用标准产品(ASSP)、芯片上系统的系统(SOC)、负载可编程逻辑设备(CPLD)等等。

[0091] 用于实施本公开的方法的程序代码可以采用一个或多个编程语言的任何组合来编写。这些程序代码可以提供给通用计算机、专用计算机或其他可编程数据处理装置的处理单元或控制器,使得程序代码当由处理单元或控制器执行时使流程图和/或框图中所规定的功能/操作被实施。程序代码可以完全在机器上执行、部分地在机器上执行,作为独立软件包部分地在机器上执行且部分地在远程机器上执行或完全在远程机器或服务器上执行。

[0092] 在本公开的上下文中,机器可读介质可以是有形的介质,其可以包含或存储以供指令执行系统、装置或设备使用或与指令执行系统、装置或设备结合地使用的程序。机器可读介质可以是机器可读信号介质或机器可读储存介质。机器可读介质可以包括但不限于电子的、磁性的、光学的、电磁的、红外的、或半导体系统、装置或设备,或者上述内容的任何合适组合。机器可读存储介质的更具体示例会包括基于一个或多个线的电气连接、便携式计算机盘、硬盘、随机存取存储器(RAM)、只读存储器(ROM)、可擦除可编程只读存储器(EPROM)

或快闪存储器)、光纤、便捷式紧凑盘只读存储器(CD-ROM)、光学储存设备、磁储存设备、或上述内容的任何合适组合。

[0093] 此外,虽然采用特定次序描绘了各操作,但是这应当理解为要求这样操作以所示出的特定次序或以顺序次序执行,或者要求所有图示的操作应被执行以取得期望的结果。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,但是这些不应当被解释为对本公开的范围的限制。在单独的实施例的上下文中描述的某些特征还可以组合地实现在单个实现中。相反地,在单个实现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现中。

[0094] 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。

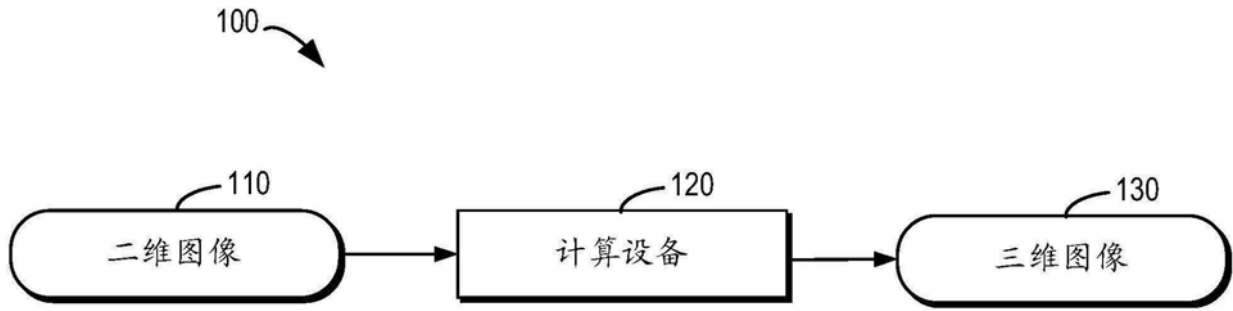


图1

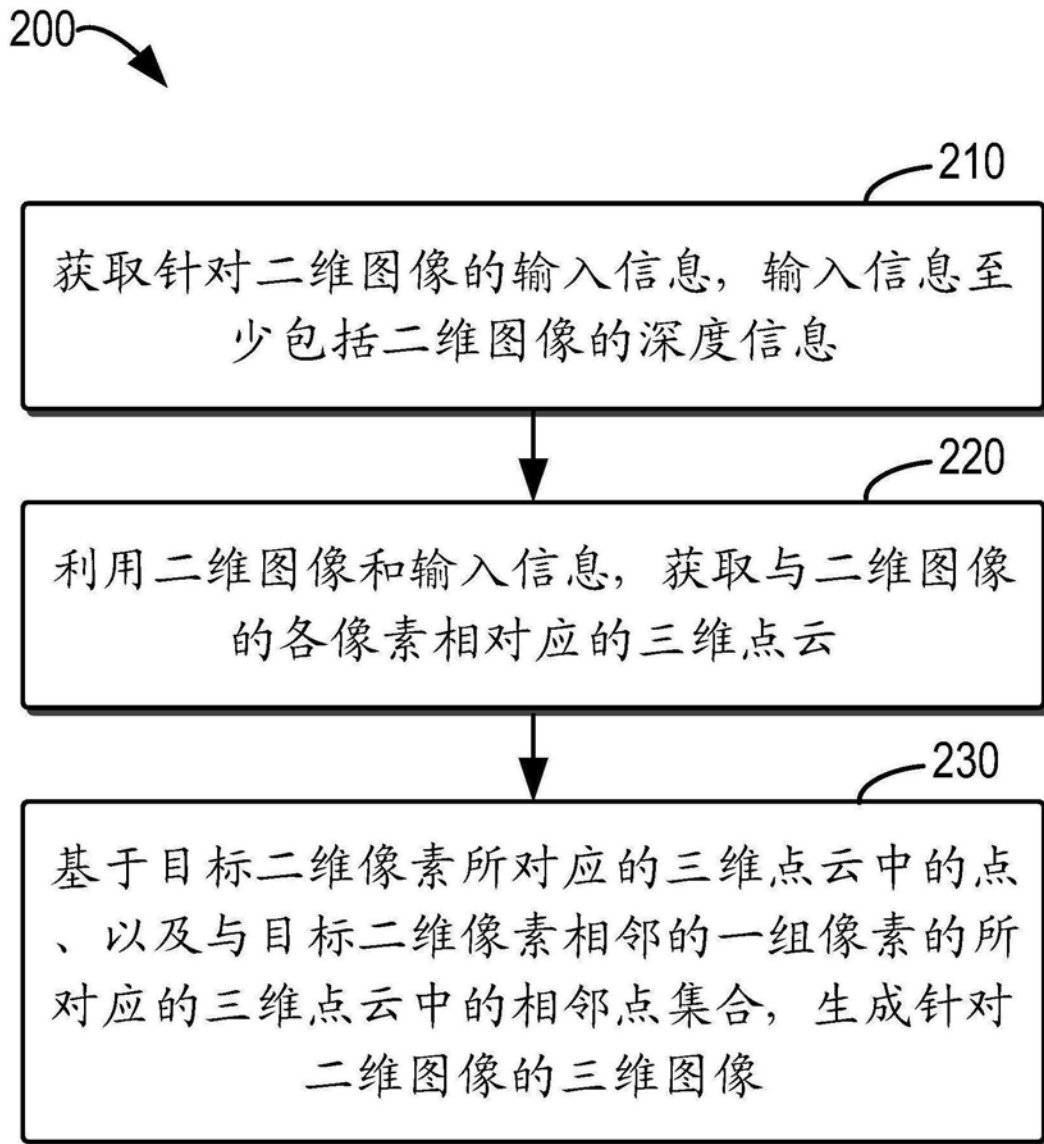


图2

300

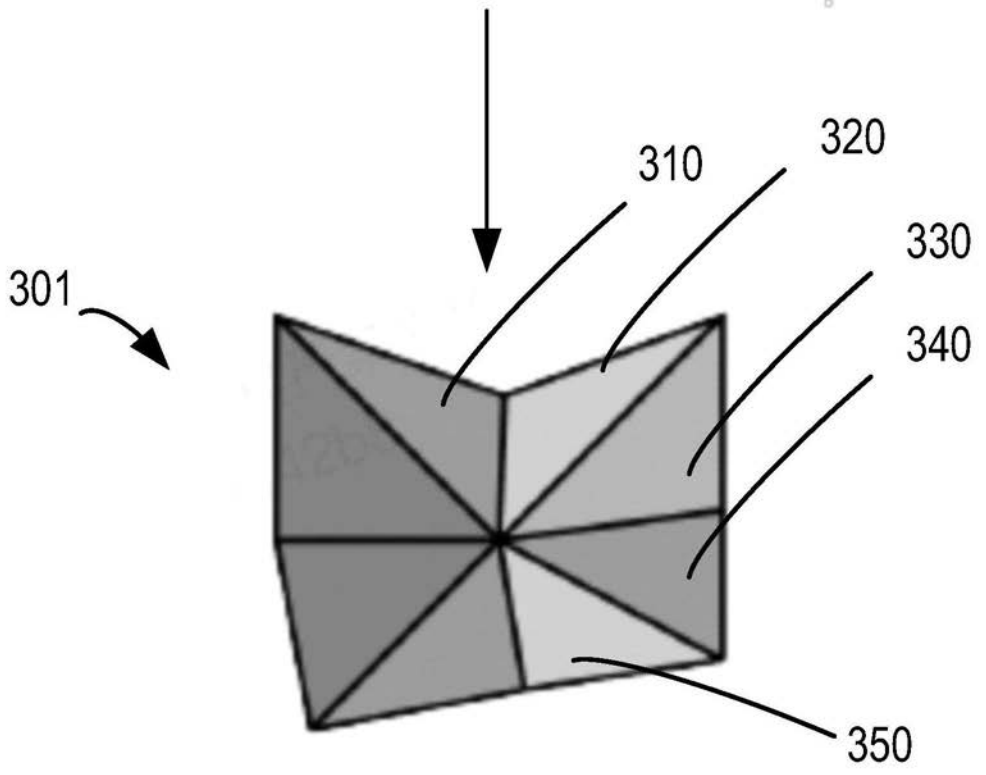
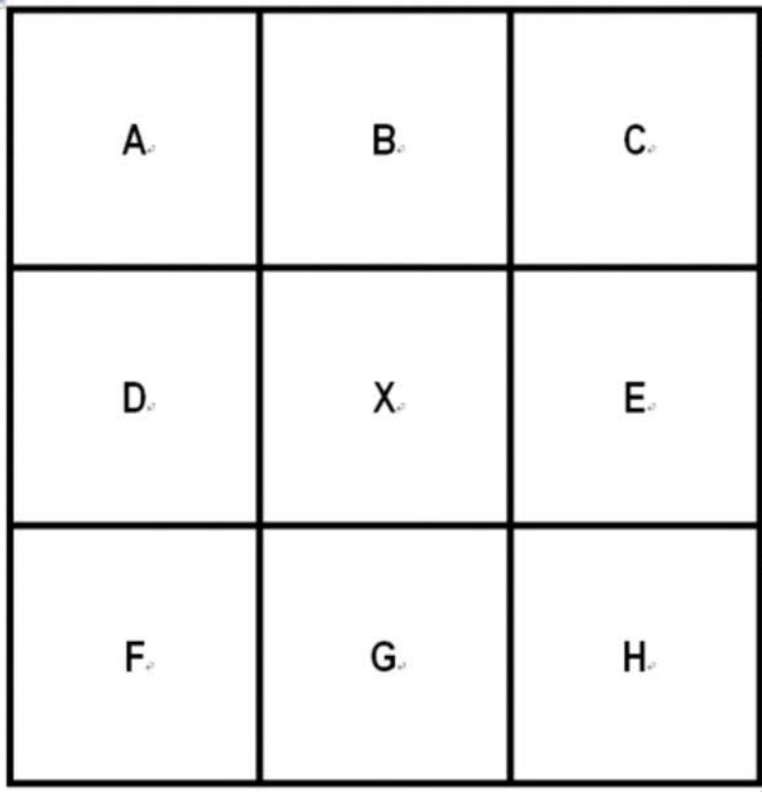


图3

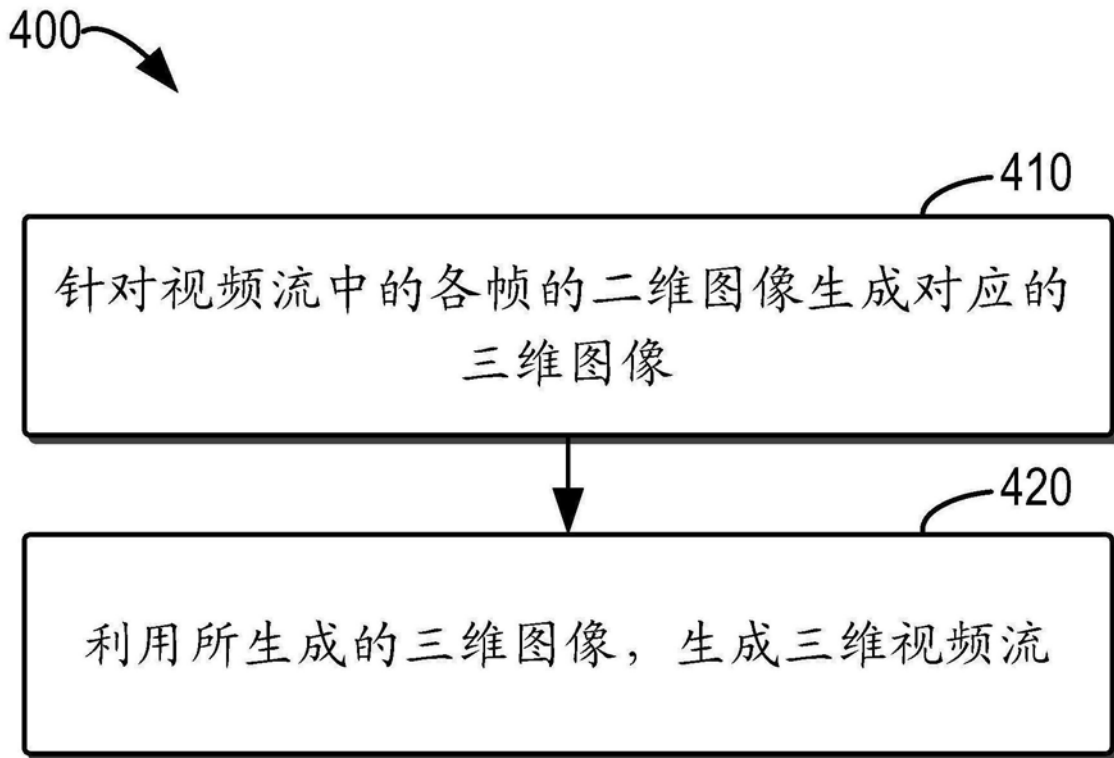


图4

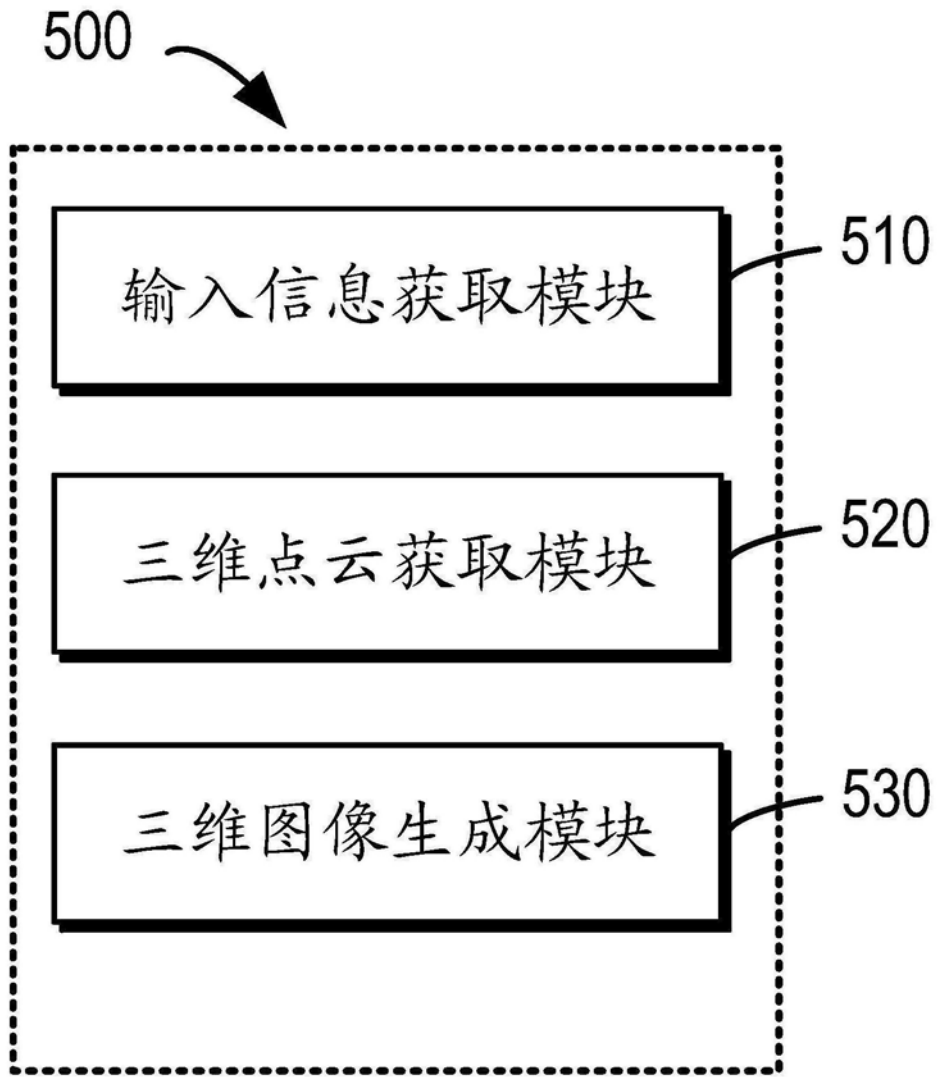


图5

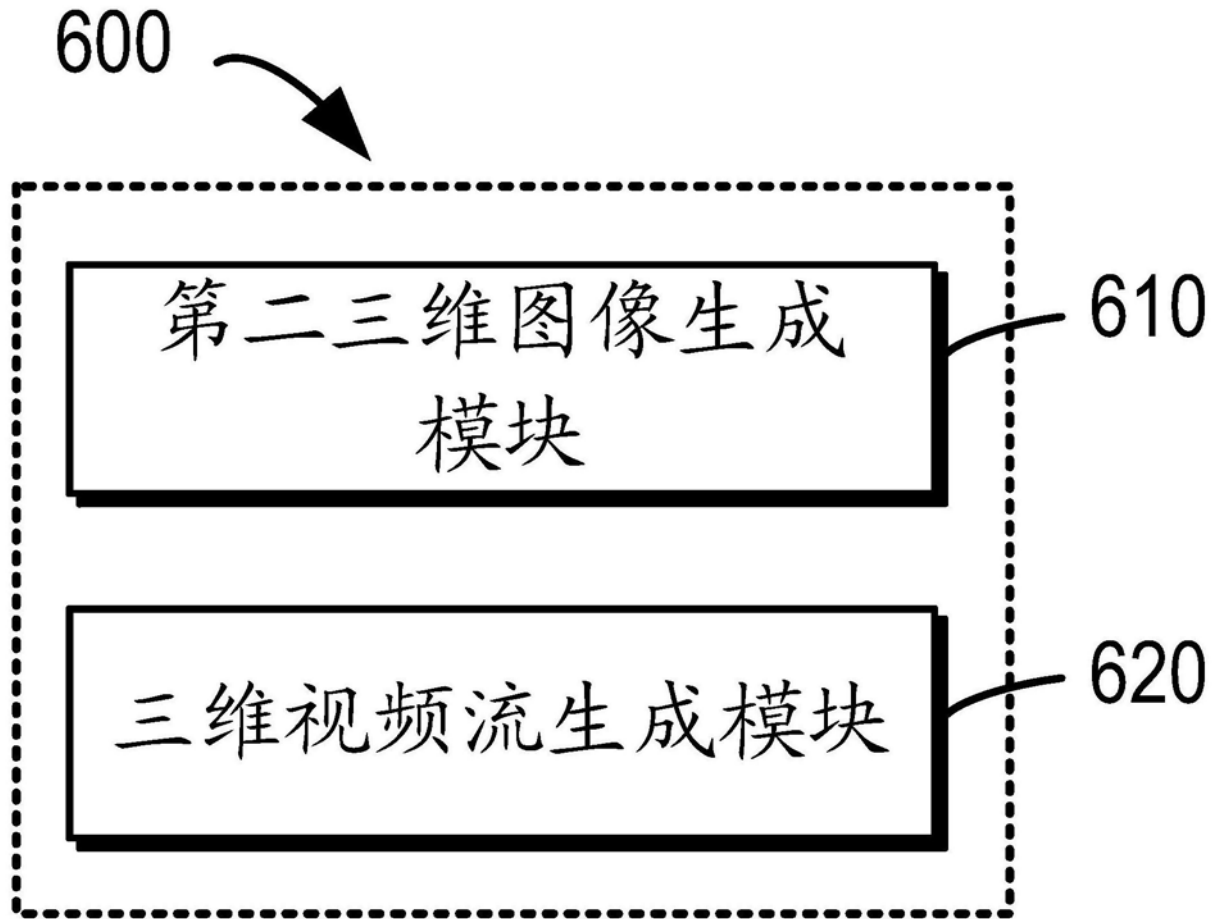


图6

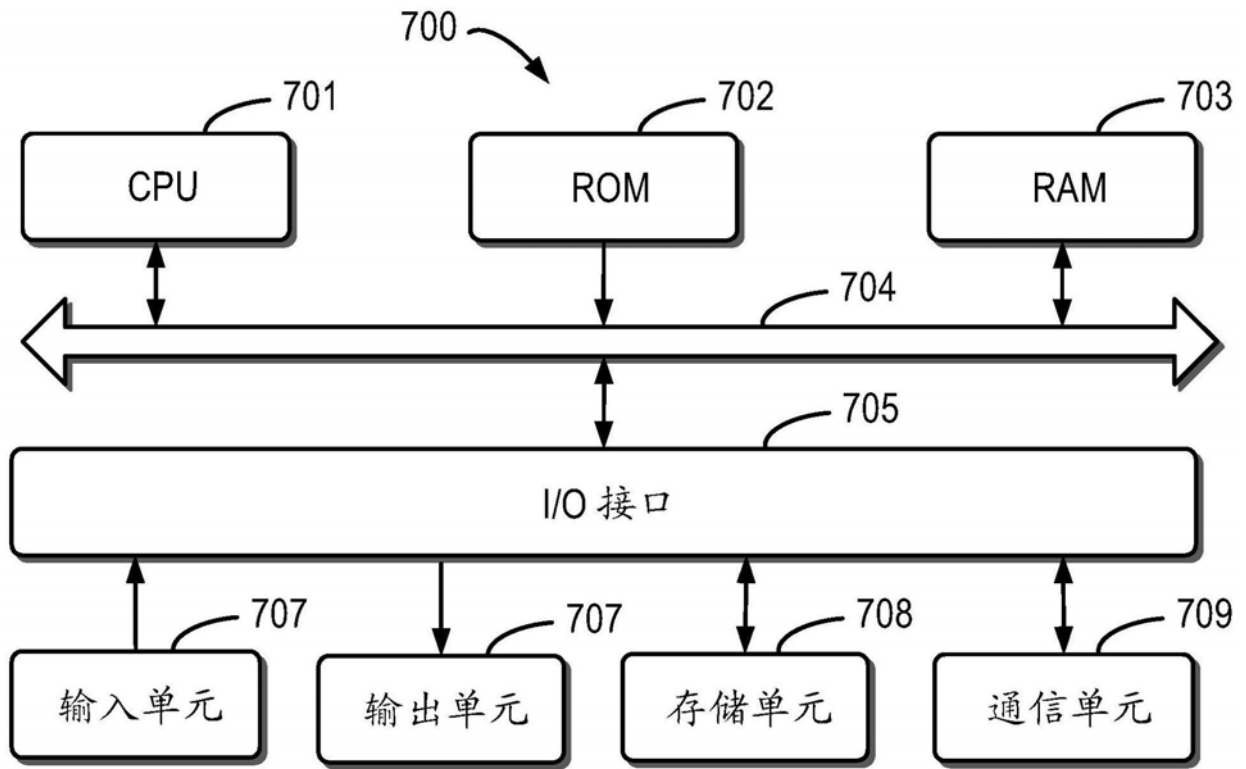


图7