



(12) 发明专利

(10) 授权公告号 CN 111742548 B

(45) 授权公告日 2023.06.27

(21) 申请号 201980014202.2

(22) 申请日 2019.01.04

(65) 同一申请的已公布的文献号
申请公布号 CN 111742548 A

(43) 申请公布日 2020.10.02

(30) 优先权数据
18305044.2 2018.01.19 EP

(85) PCT国际申请进入国家阶段日
2020.08.19

(86) PCT国际申请的申请数据
PCT/US2019/012272 2019.01.04

(87) PCT国际申请的公布数据
W02019/143486 EN 2019.07.25

(73) 专利权人 交互数字VC控股公司
地址 美国特拉华州

(72) 发明人 J.弗勒罗 R.多尔 T.塔皮

(74) 专利代理机构 北京市柳沈律师事务所
11105
专利代理师 于小宁

(51) Int.Cl.
H04N 13/161 (2006.01)
H04N 13/178 (2006.01)
H04N 13/268 (2006.01)
H04N 19/00 (2006.01)

(56) 对比文件
US 2015245063 A1, 2015.08.27
WO 2016003340 A1, 2016.01.07
CN 111742549 A, 2020.10.02
审查员 刘柳群

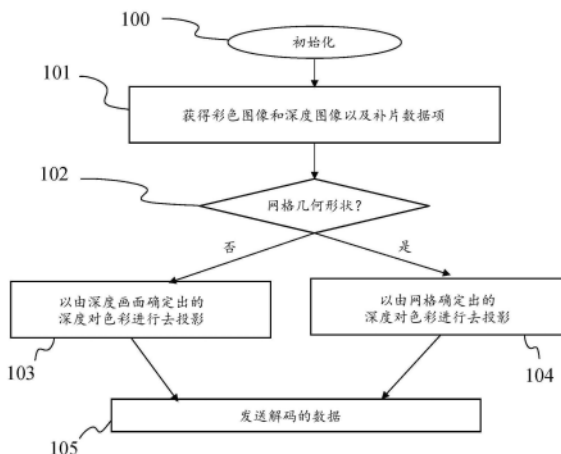
权利要求书3页 说明书11页 附图5页

(54) 发明名称

在数据流中编码及解码来自数据流的三维场景的方法和装置

(57) 摘要

提供用于编码和解码承载表示三维场景的数据的数据流的方法和装置,该数据流包括打包在色彩图像中的色彩画面;打包在深度图像中的深度画面;以及包括去投影数据的一组补片数据项;用于在色彩图像取回色彩画面的数据和几何数据。两种类型的几何数据是可能的。第一种类型的数据描述如何在深度图像中取回深度画面。第二种类型的数据包括3D网格的标识符。该网格的顶点坐标和各个面用于取回去投影的场景中各个点的位置。



1. 一种从数据流中解码三维场景的方法,该方法包括:
 - 从所述数据流中获得:
 - 打包在色彩图像中的色彩画面;
 - 打包在深度图像中的深度画面;和
 - 元数据,对于所述色彩图像的每个色彩画面,所述元数据包括:
 - 去投影数据;
 - 色彩数据,包括色彩画面在色彩图像中的位置的描述;和
 - 属于类型组的类型的几何数据,包括:第一种几何类型,包括深度画面在深度图像中的位置的描述;和第二种几何类型,包括被表示为三维网格的三维对象的标识符;
 - 对于所述色彩图像的色彩画面:
 - 在几何数据是第一种几何类型的情况下,根据由几何数据描述的深度画面的像素确定出的深度值,使用去投影数据去投影由色彩数据描述的色彩画面的像素的颜色;
 - 在几何数据是第二种几何类型的情况下,根据三维网格的顶点和面确定出的深度值,使用去投影数据去投影由色彩数据描述的色彩画面的像素的颜色。
2. 根据权利要求1所述的方法,其中,从流中获得第二种类型的几何数据的几何数据中识别的三维网格的顶点和面。
3. 根据权利要求1所述的方法,色彩数据还包括色彩图像中的色彩画面的大小和/或形状的描述。
4. 根据权利要求1所述的方法,其中,所述流包括色彩图像、深度图像和相关联的元数据的序列,深度图像和有关元数据与色彩图像相关联;色彩图像由时间信息进行结构化。
5. 一种用于在数据流中编码三维场景的方法,该方法包括:
 - 获得一组补片,所述一组补片包括:
 - 去投影数据;
 - 色彩画面;和
 - 包括深度画面的第一种几何类型的几何数据,和包括被表示为三维网格的三维对象的标识符的第二种几何类型的几何数据;
 - 通过打包所述补片的色彩画面和色彩数据生成色彩图像,色彩数据包括打包在色彩图像中的色彩画面的位置的描述;
 - 通过打包具有第一种几何类型的几何数据的补片的深度画面并通过深度图像中打包的深度画面的位置的描述来替换对应补片的几何数据中的深度画面来生成深度图像;
 - 在流中编码:
 - 生成的色彩图像;
 - 生成的深度图像;
 - 元数据,其包括每个补片的几何数据、色彩数据和去投影数据。
6. 根据权利要求5所述的方法,其中,在流的单独的语法元素中进一步编码第二种类型的几何数据的几何数据中识别的三维网格的顶点和面。
7. 根据权利要求5所述的方法,其中,色彩数据还包括色彩图像中的色彩画面的大小和/或形状的描述。
8. 根据权利要求5所述的方法,其中,所述流包括色彩图像、深度图像和相关联的元数

据的序列,深度图像和有关元数据与色彩图像相关联;色彩图像由时间信息进行结构化。

9. 一种用于从数据流中解码三维场景的设备,该设备包括与处理器相关联的存储器,该处理器被配置为:

-从所述数据流中获得:

- 打包在色彩图像中的色彩画面;
- 打包在深度图像中的深度画面;和
- 元数据,对于所述色彩图像的每个色彩画面,所述元数据包括:

-去投影数据;

-色彩数据,包括色彩画面在色彩图像中的位置的描述;和

-属于类型组的类型的几何数据,包括:第一种几何类型,包括深度画面在深度图像中的位置的描述;和第二种几何类型,包括被表示为三维网格的三维对象的标识符;

-对于所述色彩图像的色彩画面:

• 在几何数据是第一种几何类型的情况下,根据由几何数据描述的深度画面的像素确定出的深度值,使用去投影数据去投影由色彩数据描述的色彩画面的像素的颜色;

• 在几何数据是第二种几何类型的情况下,根据三维网格的顶点和面确定出的深度值,使用去投影数据去投影由色彩数据描述的色彩画面的像素的颜色。

10. 根据权利要求9所述的设备,其中,从流中获得第二种类型的几何数据的几何数据中识别的三维网格的顶点和面。

11. 根据权利要求9所述的设备,其中,色彩数据还包括色彩图像中的色彩画面的大小和/或形状的描述。

12. 根据权利要求9所述的设备,其中,所述流包括色彩图像、深度图像和相关联的元数据的序列,深度图像和有关元数据与色彩图像相关联;色彩图像由时间信息进行结构化。

13. 一种用于在数据流中编码三维场景的设备,该设备包括与处理器关联的存储器,该处理器配置为:

-获得一组补片,所述一组补片包括:

- 去投影数据;
- 色彩画面;和
- 第一种几何类型的几何数据,其包括深度画面;和第二种几何类型的几何数据,其包括被表示为三维网格的三维对象的标识符;

-通过打包所述补片的色彩画面和色彩数据来生成色彩图像,色彩数据包括打包在色彩图像中的色彩画面的位置的描述;

-通过打包具有第一种几何类型的几何数据的补片的深度画面并通过深度图像中打包的所述补片的深度画面的位置的描述来替换对应补片的几何数据中的深度画面来生成深度图像;

-在流中编码:

- 生成的色彩图像;
- 生成的深度图像;
- 元数据,其包括每个补片的几何数据、色彩数据和去投影数据。

14. 根据权利要求13所述的设备,其中,在流的单独的语法元素中进一步编码第二种类

型的几何数据的几何数据中识别的三维网格的顶点和面。

15. 根据权利要求13所述的设备,其中,色彩数据还包括色彩图像中的色彩画面的大小和/或形状的描述。

16. 根据权利要求13所述的设备,其中,所述流包括色彩图像、深度图像和相关联的元数据的序列,深度图像和有关元数据与色彩图像相关联;色彩图像由时间信息进行结构化。

在数据流中编码及解码来自数据流的三维场景的方法和装置

技术领域

[0001] 本原理一般涉及在数据流中编码以及解码来自数据流的三维场景或三维场景的序列的编码和解码。具体地,但不排他地,本原理的技术领域涉及对场景的纹理的色彩图像和深度图像或者场景的几何形状的3D网格的编码/解码。

背景技术

[0002] 本部分旨在向读者介绍本领域的各个方面,其可能与以下描述和/或要求保护的本原理的各个方面有关。认为该讨论有助于向读者提供背景信息,以有助于更好地理解本原理的各个方面。因此,应该理解的是,要鉴于此地阅读这些陈述,而不作为对现有技术的承认。

[0003] 沉浸式视频(也称为360°平面视频)允许用户通过围绕静止视点的头部旋转来观看自己周围的一切。旋转仅允许3自由度(3DoF)体验。即使3DoF视频足以用于首次全向视频体验(例如使用头戴式显示设备(HMD)),对于将期望更大自由度的观看者来说,例如通过体验视差,3DoF视频也可能很快变得令人沮丧。另外,3DoF也可能引起头晕,因为用户绝不是仅仅旋转其头部,而是还在三个方向上平移头部,平移在3DoF视频体验中不被再现。

[0004] 体积式视频(也称为6自由度(6DoF)视频)是3DoF视频的替代。当观看6DoF视频时,除了旋转之外,用户还可以在观看的内容内平移其头部以及甚至身体,并体验视差以及甚至体积。这样的视频极大地增加沉浸的感觉和场景深度的感知,并且也通过在头部平移期间提供一致的视觉反馈而防止头晕。内容通过专用传感器创建,允许同时记录感兴趣场景的色彩和深度。结合摄影机技术使用的一套色彩相机是进行这种记录的一种常用方式。

[0005] 体积式视频是3D场景的序列。编码体积式视频的一种解决方案是将3D场景的序列中的每个3D场景投影到被聚类为色彩画面和深度画面(称为补片)的投影图上。补片以存储在视频流的视频轨道中的色彩图像和深度图像打包。这种编码具有使用标准的图像和视频处理标准的优点。在解码时,色彩图像的像素以由存储在相关联的深度画面中的信息确定的深度被去投影。这样的解决方案是有效的。然而,将这种大量数据编码为视频流的视频轨道中的图像产生问题。比特流的大小引起有关存储空间,网络传输和解码性能的比特率技术问题。

发明内容

[0006] 以下提出本原理的简要概述,以便提供对本原理的一些方面的基本理解。本概述不是对本原理的广泛综述。不意图标识本原理的关键或重要要素。以下概述仅以简化形式提出本原理的一些方面,作为下面提供的更详细描述的前言。

[0007] 本原理涉及一种在流中编码3D场景的方法。本公开涉及一种在流中编码三维场景的方法。该方法包括:

[0008] -获得一组补片,补片包括:

[0009] • 去投影数据;

[0010] • 色彩图像;和

[0011] • 第一种几何类型的几何数据包括深度画面;或第二种几何类型的几何数据包括被表示为网格的三维对象的标识符;

[0012] -通过打包所述补片的色彩画面并将色彩数据存储在对应的补片中来生成色彩图像,色彩数据包括打包在色彩图像中的所述补片的色彩画面的位置的描述;

[0013] -通过打包具有第一种几何类型的几何数据的补片的深度画面并通过深度图像中打包的所述补片的深度画面的位置的描述来替换对应补片的几何数据中的深度画面来生成深度图像;

[0014] -在流中编码;

[0015] • 生成的色彩图像;

[0016] • 生成的深度图像;

[0017] • 一组补片数据项,补片数据项包括去投影数据,补片的色彩数据和所述补片的几何数据。

[0018] 本公开还涉及一种适于在流中编码三维场景的设备。该设备包括与处理器相关联的存储器,该处理器被配置为执行在流中编码三维场景的本方法。

[0019] 本公开还涉及从数据流解码三维场景的方法。该方法包括:

[0020] -从所述数据流中获得:

[0021] • 色彩图像,包括打包在所述彩色图像中的色彩画面;

[0022] • 深度图像,包括打包在所述深度图像中的深度画面;和

[0023] • 一组补片数据项,补片数据项包括:

[0024] -去投影数据;

[0025] -色彩数据,包括色彩画面在色彩图像中的位置的描述;和

[0026] -第一种几何类型的几何数据包括深度画面在深度图像中的位置的描述;或第二种几何类型的几何数据包括被表示为网格的三维对象的标识符;

[0027] -对于该组补片数据项:

[0028] • 如果几何数据是第一种几何类型,则根据由几何数据描述的深度画面的像素确定出的深度值,使用去投影数据去投影由色彩数据描述的色彩画面的像素的颜色;

[0029] • 如果几何数据是第二种几何类型,则根据网格的顶点和面确定出的深度值,使用去投影数据去投影由色彩数据描述的色彩画面的像素的颜色。

[0030] 本公开还涉及适于从数据流解码三维场景的设备。该设备包括与处理器相关联的存储器,该处理器被配置为执行从数据流中解码三维场景的本方法。

[0031] 本公开还涉及一种承载表示三维场景的数据的数据流。该数据流包括:

[0032] -至少一个色彩图像,包括打包在所述彩色图像中的色彩画面;

[0033] -至少一个深度图像,包括打包在所述深度图像中的深度画面,深度图像与第一色彩图像相关联;和

[0034] -与所述第一色彩图像相关联的一组补片数据项,补片数据项包括:

[0035] • 去投影数据;

[0036] • 色彩数据,其包括色彩画面在第一色彩图像中的位置的描述;和

[0037] • 第一种几何类型的几何数据包括深度画面在与第一色彩图像相关联的深度图

像中的位置的描述;或第二种几何类型的几何数据包括被表示为网格的三维对象的标识符。

附图说明

[0038] 在阅读以下描述时,将更好地理解本公开,并且将出现其他特定特征和优点,该描述参考附图,在附图中:

[0039] -图1示出根据本原理的非限制性实施例的对象的三维(3D)模型和与3D模型对应的点云的点;

[0040] -图2示出根据本原理的非限制性实施例的表示三维场景的图像,三维场景包括被捕获为点云的若干个对象的表面表示;

[0041] -图3例示根据本原理的非限制性实施例的3D场景的序列的编码,传输和解码的示例;

[0042] -图4例示根据本原理的非限制性实施例的相对于图2的3D场景的两个补片,每个补片包括色彩画面,深度画面和去投影数据。

[0043] -图5示出根据本原理的非限制性实施例的由被表示为网格的3D场景的一部分的投影所获得的补片;

[0044] -图6例示根据本原理的非限制性实施例的将获得的补片打包在色彩图像和深度图像中;

[0045] -图7示出根据本原理的非限制性实施例的,当通过基于分组的传输协议来传输数据时的流的语法的实施例的示例;

[0046] -图8例示根据本原理的非限制性实施例的,可以被配置为实现结合图9和/或图10描述的方法的设备的架构的示例;

[0047] -图9例示根据本原理的非限制性实施例的,在被配置为图3的编码器31的图8的设备中,用于在流中编码点云的方法;

[0048] -图10例示根据本原理的非限制性实施例的,在被配置为图3的设备33的图8的设备中,用于从流解码三维场景的方法。

具体实施方式

[0049] 在下文中将参考附图更全面地描述本原理,附图中示出本原理的示例。然而,本原理可以以许多替代形式实施,并且不应该被解释为限于这里阐述的示例。因此,虽然本原理易于进行各种修改和替换形式,但是其具体示例通过附图中的示例示出并且将在本文中详细描述。然而,应该理解,不意图将本原理限制于公开的特定形式,相反,本公开要覆盖落入权利要求定义的本原理的精神和范围内的所有修改,等同物和替代物。

[0050] 本文使用的术语仅用于描述特定示例的目的,并不旨在限制本原理。如本文使用的,单数形式“一”,“一个”和“该”旨在也包括复数形式,除非上下文另有明确说明。将进一步理解,当在本说明书中使用术语“包括”,“包含”,“包括”和/或“包括”指定陈述的特征,整数,步骤,操作,要素和/或组件的存在,但是不排除存在或附加一个或多个其他特征,整数,步骤,操作,要素,组件和/或其分组。此外,当元件被称为“响应”或“连接”到另一个元件时,它可以直接响应或连接到另一个元件,或者可以存在中间元件。相反,当元件被称为

“直接响应”或“直接连接”到其他元件时,不存在中间元件。如本文使用,术语“和/或”包括一个或多个相关所列项目的任何和所有组合,并且可以缩写为“/”。

[0051] 将理解,尽管本文可以使用术语第一,第二等来描述各种要素,但是这些元件不应受这些术语的限制。这些术语仅用于区分一个要素与另一个要素。例如,第一要素可以被称为第二要素,并且类似地,在不脱离本原理的教导的情况下,第二要素可以被称为第一要素。

[0052] 尽管一些图包括通信路径上的箭头以示出通信的主要方向,但是要理解,通信可以在与所示箭头相反的方向上发生。

[0053] 关于框图和操作流程图描述一些示例,其中每个框表示用于实现(多个)指定逻辑功能的电路元件,模块或包括一个或多个可执行指令的代码部分。还应注意,在其他实现方式中,框中标注的(多个)功能可以不按所述顺序发生。例如,连续示出的两个框实际上可以基本上同时执行,或者这些框有时取决于所涉及的功能可以以相反的顺序执行。

[0054] 本文对“根据示例”或“在示例中”的引用意味着结合该示例描述的特定特征,结构或特性可以包括在本原理的至少一个实现方式中。在说明书中各处出现的短语“根据示例”或“在示例中”不一定都指代相同的示例,也不一定是与其他示例相互排斥的单独或替代示例。

[0055] 权利要求中出现的附图标记仅是示例性的,并且对权利要求的范围没有限制作用。虽然没有明确描述,但是这些示例和变型可以以任何组合或子组合使用。

[0056] 描述了本原理用于在数据流中编码/解码来自数据流的三维(3D)场景或三维场景的序列。

[0057] 图1示出对象10的三维(3D)模型以及与3D模型10对应的点云11的点。模型11可以是3D网格表示,并且点云10的点可以是网格的顶点。点10也可以是散布在网格的各个面的表面上的点。模型10也可以表示为云11的点的喷溅版本;通过喷溅点云11的点来创建模型10的表面。模型10可以由许多不同的表示(诸如,体素或样条)来表示。图1例示出总是有可能从3D对象的表面表示中定义点云。反之,总是有可能从云的点创建3D对象的表面表示。如本文使用,将3D对象的点(通过3D场景的扩展点)投影到图像上等同于将该3D对象的任何表示投影到图像上。

[0058] 图2示出表示三维场景的图像20,该三维场景包括被捕获为点云的若干个对象的表面表示。图像20从与获取视点不同的视点生成。例如,图像20右侧的人物不完整,其左臂和后背的点不可用(例如,其没有被捕获)来完成表面表示。3D场景可以包括不同类型的表示。例如,场景中的人物表示为点云,而雕像表示为带纹理的网格。总是从视点呈现3D场景。不是场景的所有点都从视点可见,场景的某些部分被可见部分遮挡。体积式视频的观看者可以在定义一组可能的视点的区域中的3D场景内移动。从该组可能的视点可见的3D场景的每个点都必须被编码。

[0059] 图3例示3D场景的序列的编码,传输和解码的示例。根据本编码方法的原理,由编码器31在流32中编码至少3D场景30的序列。解码器33从源获得流32。例如,源属于包含以下的集合:

[0060] -本地存储器,例如,视频存储器或RAM(随机存取存储器),闪存,ROM(只读存储器),硬盘;

[0061] -存储接口,例如,具有大容量存储器,RAM,闪存,ROM,光盘或磁性载体的接口;
[0062] -通信接口,例如,有线接口(例如,总线接口,广域网接口,局域网接口)或无线接口(诸如,IEEE 802.11接口或Bluetooth®接口);

[0063] -用户界面,诸如,图形用户界面,使用户能够输入数据。

[0064] 解码器33根据本解码方法解码来自流32的3D场景34的序列。根据本原理,3D场景34的序列尽可能类似于3D场景30的序列。可以由呈现的35从源获得3D场景34的序列。呈现器35计算要显示的图像用于6DoF视频观看体验。

[0065] 图4例示相对于图2的3D场景的两个补片,每个补片包括色彩画面,深度画面和去投影数据。可以使用不同的投影过程来生成补片。例如,可以通过剥离投影过程来生成补片。另一种可能性是使用立方体映射投影的八叉树。可以使用任何合适的投影过程来生成补片。适用于本原理的投影过程生成一组补片,这些补片包括色彩画面,深度画面(或存储两者像素数据的唯一画面)和去投影数据,以允许解码器从两个画面中解码三维场景。在等矩形投影映射的示例中,去投影数据可以包括投影的中心的坐标,补片所对应的参考系中的立体角(例如, $[\theta_{\min}, \theta_{\max}]$, $[\phi_{\min}, \phi_{\max}]$)和用于编码深度画面中的深度的深度值的范围。通过投影过程去投影存储在两个画面中的信息所需的每个数据都包括在去投影数据中。这些数据允许解码器取回3D坐标和在两个投影图中投影的点的色彩。在图4的示例中,生成包括色彩画面401,深度画面402和一组投影数据403的补片40。该补片编码表示图2的3D场景的人物之一的正面的数据。色彩画面401是场景的这一部分的点在场景3D空间中定位的平面上的投影结果。深度画面402的像素存储该平面与3D场景的该部分之间的距离。在图4的示例中,深度像素越亮,场景的点越靠近平面。黑色深度像素意味着在色彩画面中的这些坐标处没有色彩信息。在图4上提供补片41的第二示例。补片41与图2的3D场景的第二人物的一部分对应。补片41包括色彩画面411,深度画面412和去投影数据413。去投影数据413是不同的id去投影数据403。在等矩形投影映射的示例中,投影的中心可能不同,在这两个画面中编码的3D空间的立体角不同并且深度画面的深度编码的动态性也可能不同。补片可以具有不同的大小和不同的形状。在本文档的示例中,每个补片是矩形。这仅是示例,并且根据投影过程,补片可以具有任何形状(例如,椭圆形或梯形)。如果未被缺省地设置,则此信息包含在去投影数据中。

[0066] 根据本原理,根据任何合适的方法获得表示要编码的3D场景的一组补片。然后将这些补片打包到色彩图像和深度图像中,以便例如根据ISO BMFF标准添加到数据流的视频轨道中。3D场景的补片的色彩画面和深度画面的累积大小可能很可观。

[0067] 图5示出通过投影被表示为网格的3D场景的一部分而获得的补片。3D场景由3D信息组成,其可以根据各种表示来结构化。3D场景的一部分可能由点云,样条曲线的另一部分,还有彩色体素的另一部分组成。点云是体积式视频最常用的表示,尤其是在现实世界获取中。但是,在电影制作的后期制作阶段,经常在3D场景中添加一些对象(例如,小说人物,装饰特征,墙壁,地板)。这些附加对象通常由计算机图形方法创建的3D网格表示。在图5的示例中,3D网格54属于场景。3D网格包括顶点和各个面。顶点包括有关如何在网格的表面上管理光的一组分量。网格与称为纹理的画面相关联,并且顶点包括纹理坐标。这样的补片具有两种几何类型的组中的第一种类型的几何数据(深度画面)。

[0068] 根据本原理,在投影过程中,表示为3D网格的3D场景的点被投影到单独的补片中。例如,3D网格54的投影生成包括色彩画面51,深度画面52和去投影数据53的补片50。所以,深度画面52表示根据投影条件的3D网格54的几何形状(也就是,仅顶点坐标和各个面)。根据投影条件,有关光和纹理坐标的信息被编码在色彩图像51中。用于在唯一的色彩图像中编码纹理特征和光特征的字节的数量小于编码顶点分量加上纹理画面的相同的信息所需要的字节的数量,这是因为这些数据可能在3D场景序列中改变,所以在每次修改时都需要在流中进行编码。但是,顶点和面列表需要几个字节,并且随时间改变的可能性小。将几何形状编码为深度画面一般要求大量空间。因此,根据本原理,通过投影表示为3D网格的3D场景的一部分的点而获得的补片包括该网格的标识符。该补片保留了其已经从中生成的3D网格的指针。在变型中,补片50不包括深度画面52,而仅包括标识3D网格54的数据。这种补片具有两种几何类型的组中的第二种类型的几何数据。

[0069] 根据本原理,获得关于图4和图5描述的一组补片。该组补片承载按照由一组可能的视点确定的必须在数据流中编码的场景中每个点的数据。

[0070] 图6例示将获得的补片打包在色彩图像和深度图像中。对为3D场景获得的每个色彩画面执行打包操作,以生成色彩图像61。存在许多启发式算法来解决将矩形单元打包到矩形箱(即,要生成的图像)中的NP完全问题。例如Jukka Jylänki的“A Thousand Ways to Pack the Bin”中描述的算法或“MaxRects”算法(即,单箱-最佳短边第一配置),其以相当低的计算成本提供良好的填充率。在打包操作结束时,图像中色彩画面的位置(x,y)(例如,左下角位置),形状(例如,矩形,椭圆形)和大小,并且根据打包算法,指示补片是否已旋转的布尔值被分配给补片数据项列表的每个补片。可以通过在视频编码器的编码单元上对齐色彩画面来优化打包,以改进解码阶段。图6示出其中图4和5的打包后的色彩画面的画面。

[0071] 对为具有第一类型的几何数据的补片而获得的深度画面执行打包操作,以生成深度图像62。仅将未被表示为3D网格的3D场景的各部分进行投影而获得的深度画面存储在打包后的深度图像中。

[0072] 图7示出当通过基于分组的传输协议来传输数据时的流的语法的实施例的示例。图7示出体积式视频流的示例结构70。该结构包含容器,该容器以独立的语法元素来组织流。该结构可以包括首部部分71,该首部部分是该流的每个语法元素共用的一组数据。例如,首部部分包括有关语法元素的元数据,以描述它们中每个的性质和作用。首部部分还可以包括用于编码6DoF呈现的第一色彩图像的一组可能视点的坐标,以及有关画面的大小和分辨率的信息。该结构包括有效载荷,也称为视频轨道,包括色彩画面72和深度画面73。图像可能已经根据视频压缩方法被压缩。

[0073] 视频轨道的第三语法元素74包括一组补片数据项。补片数据项是包括被编码为字节的值的数据结构。可以使用任何合适的方法来压缩该数据结构。补片数据项包括去投影数据,色彩画面在色彩图像中的位置(以及可选的形状和大小)以及几何数据。有两种几何类型。第一种几何类型的几何数据包括深度画面在深度图像中的位置(以及可选地,形状和大小)。第二种几何类型的几何数据包括3D场景的3D网格的标识符。

[0074] 流的这三个语法元素相关联在一起。根据本原理,流可以包括这些数据的三元组的序列。对于流的第一语法元素中的每个色彩画面,深度画面与第二语法元素中的色彩画面相关联,并且一组补片数据项与流的第三语法元素中的色彩画面相关联。这些三元组被

排序,并与时间信息(诸如,用于动态视频呈现的时间戳)相关联。

[0075] 出于说明目的,在ISOBMFF文件格式标准的上下文中,将典型在MOOV类型的框内的ISOBMFF轨道中参考色彩图,深度图和元数据,而色彩图和深度图数据本身嵌入在mdat类型的媒体数据框中。

[0076] 图8示出可以被配置为实现关于图9和/或10描述的方法的设备80的示例架构。设备80可以被配置为图3的编码器31或解码器33。

[0077] 设备80包含以下元件,这些元件通过数据和地址总线81链接在一起:

[0078] -微处理器82(或CPU),例如是DSP(或数字信号处理器);

[0079] -ROM(或只读存储器)83;

[0080] -RAM(或随机存取存储器)84;

[0081] -存储接口85;

[0082] -I/O接口86,用于接收来自应用的要发送的数据;和

[0083] -电源,例如,电池。

[0084] 根据示例,电源在设备外部。在每个提到的存储器中,说明书中使用的“寄存器”一词可以对应于小容量的区域(一些比特)或非常大的区域(例如,整个程序或接收的或解码的大量数据)。ROM 83至少包括程序和参数。ROM 83可以存储用于执行根据本原理的技术的算法和指令。当接通时,CPU 82将程序上载到RAM中并执行相应指令。。

[0085] RAM 84在寄存器中包括由CPU 82执行并在设备80接通之后上载的程序,寄存器中的输入数据,寄存器中方法的不同状态的中间数据,以及寄存器中用于执行方法的其他变量。

[0086] 本文描述的实现方式可以在例如方法或处理,装置,软件程序产品,数据流或信号中实现。即使仅在单个实现方式的形式上下文中讨论(例如,仅作为方法或设备讨论),讨论的特征的实现方式也可以以其他形式(例如,程序)实现。装置可以在例如适当的硬件,软件和固件中实现。方法可以在例如诸如例如处理器之类的装置中实现,该处理器通常涉及处理设备,一般包括例如计算机,微处理器,集成电路或可编程逻辑设备。处理器还包括通信设备,诸如例如计算机,蜂窝电话,便携式/个人数字助理(“PDA”),以及便于终端用户之间的信息通信的其他设备。

[0087] 根据图3的编码或编码器31的示例,三维场景30从源获得。例如,源属于包含以下的集合:

[0088] -本地存储器(83或84),例如,视频存储器或RAM(随机存取存储器),闪存,ROM(只读存储器),硬盘;

[0089] -存储接口(85),例如,具有大容量存储器,RAM,闪存,ROM,光盘或磁性载体的接口;

[0090] -通信接口(86),例如,有线接口(例如,总线接口,广域网接口,局域网接口)或无线接口(诸如,IEEE 802.11接口或Bluetooth®接口);

[0091] -用户界面,诸如,图形用户界面,使用户能够输入数据。

[0092] 根据图3的解码器或(多个)解码器33的示例,流被发送到目的地;具体而言,目的地属于包含以下的集合:

[0093] -本地存储器(83或84),例如,视频存储器或RAM,闪存,硬盘;

[0094] -存储接口 (85), 例如, 具有大容量存储器, RAM, 闪存, ROM, 光盘或磁性载体的接口; 和

[0095] -通信接口 (86), 例如, 有线接口 (例如, 总线接口 (例如USB (或通用串行总线)), 广域网接口, 局域网接口, HDMI (高清多媒体接口) 接口) 或无线接口 (诸如, IEEE 802.11接口, WiFi®或 Bluetooth®接口)。

[0096] 根据编码或编码器的示例, 包括表示体积式场景的数据的比特流被发送到目的地。作为示例, 比特流存储在本地或远程存储器中, 例如, 视频存储器 (84) 或RAM (84), 硬盘 (83)。在变型中, 比特流被发送到存储接口 (85), 例如, 具有大容量存储器, 闪存, ROM, 光盘或磁性载体的接口和/或通过通信接口 (85) 传输, 例如, 到点对点链路, 通信总线, 点对多点链路或广播网络的接口。

[0097] 根据图3的解码或解码器或呈现器33的示例, 从源获得比特流。示例性地, 从本地存储器读取比特流, 例如, 视频存储器 (84), RAM (84), ROM (83), 闪存 (83) 或硬盘 (83)。在变型中, 比特流从存储接口 (85) 接收, 例如, 具有大容量存储器, RAM, ROM, 闪存, 光盘或磁性载体的接口和/或从通信接口 (85) 接收, 例如, 到点对点链路, 总线, 点对多点链路或广播网络的接口。

[0098] 根据示例, 设备80被配置为实现上面关于图9或10描述的方法, 并且属于包括以下的集合:

[0099] -移动设备;

[0100] -通信设备;

[0101] -游戏设备;

[0102] -平板电脑 (或平板计算机);

[0103] -膝上型计算机;

[0104] -照相机;

[0105] -摄像机;

[0106] -编码芯片;

[0107] -服务器 (例如, 广播服务器, 视频点播服务器或网络服务器)。

[0108] 图9例示根据本原理的非限制性实施例的, 在被配置为编码器的设备80 (参照图8描述) 中用于在流中编码点云的方法。

[0109] 在步骤90中, 设备80的不同参数被更新。特别地, 从源获得3D场景, 在3D场景的空间中确定一组可能的视点, 初始化投影映射, 并且确定投影图, 画面和图像的大小和分辨率。

[0110] 在步骤91中, 对3D场景执行投影处理, 以获得表示从该组可能的视点的每个视点可见的场景部分的一组彩色画面和深度画面。可以使用任何合适的方法来获得该组彩色画面和深度画面。彩色画面和深度画面存储在称为补片的数据结构中。执行彩色画面和深度画面的去投影所需的数据从投影过程参数进行计算, 并且存储在补片数据结构中。存在用于补片的两种类型的几何数据。在缺省情况下, 补片作为包含深度画面的几何数据。这种几何数据是第一种类型几何数据。由3D网格表示的3D场景的对象被投影在单独的补片中, 其包括彩色画面和几何数据, 该几何数据包括指向已从中生成补片的3D网格的标识符。这种几何数据是第二种类型的几何数据。在变型中, 第二种类型的几何数据也可以包括深度画

面。

[0111] 在步骤92中,对获得的该组补片的每个色彩画面执行打包过程,以生成色彩图像。在该步骤可以使用任何合适的方法来将色彩画面有效地打包在色彩图像中。色彩图像也称为色彩补片图集。根据色彩画面的点的投影将在色彩补片图集中占据的大小以给定的角分辨率(例如,每个像素3秒或每个像素5秒),将色彩画面布置在色彩图像中。色彩画面在色彩图像中的位置存储在称为色彩数据的数据结构中。如果色彩画面的形状没有按缺省设置(例如,矩形),则用于色彩图像中该色彩画面的形状被存储在色彩数据结构中。如果打包在色彩图像中的色彩画面的分辨率不是常数(即,对于每个色彩画面不是相同的),则色彩图像中色彩图像的大小被存储在色彩数据中。

[0112] 在步骤93中,将第一类型的几何形状的补片的深度画面打包在深度图像中。仅打包第一类型的补片的深度画面(这是为什么可以从第二类型的几何形状的补片的几何数据中移除深度画面的原因)。由于深度图像明显小于色彩图像,并且明显小于其中将打包每个深度画面的深度图像,因此这具有节省流中的很多比特率的优点。指向3D网格的标识符需要流中少量的字节,而编码为画面的相同信息需要多得多的字节。

[0113] 在步骤94中,根据关于图7描述的结构生成表示3D场景的流。由第二种类型的几何数据的标识符指向的3D网格可以被添加到流的第四语法元素中,或者可以通过不同的方式(例如本地存储)被发送到解码器。

[0114] 图10例示根据本原理的非限制性实施例的,在被配置为图3的设备33的设备80(关于图8描述)中,用于从流解码三维场景的方法。

[0115] 在步骤100中,设备80的不同参数被更新。具体地,从源获得流,在3D场景的空间中确定一组可能的视点。在变型中,从流中解码该组可能的视点。3D网格的列表从源获得并且存储在存储器中。3D网格仅包括顶点坐标和面列表。在变型中,从可以被同步的流的第四语法元素取回3D网格的列表。

[0116] 在步骤101中,从流获得表示3D场景的数据。这些数据包括色彩图像,深度图像和称为补片数据项的一组数据结构。补片数据项包括去投影数据,色彩画面在色彩图像中的位置(以及可选的形状和大小)以及几何数据。存在两种几何类型。第一种几何类型的几何数据包括深度画面在深度图像中的位置(以及可选地,形状和大小)。第二种几何类型的几何数据包括存储在3D网格的列表中的3D网格的标识符。

[0117] 然后对于至少一个补片数据项,优选地对于每个补片数据项,重复本方法的后续步骤。

[0118] 在步骤102中,执行测试。如果补片的几何数据是第一几何类型,则执行步骤103。否则,如果补片的几何数据是第二几何类型,则执行步骤104。

[0119] 在步骤103中,使用色彩数据结构的数据从色彩图像中取回色彩画面。使用补片的几何数据的数据从深度图像中取回深度画面。根据深度画面中的对应像素计算出的深度通过使用去投影数据来对色彩画面的像素进行去投影。因此,场景的一部分被解码。

[0120] 在步骤104中,使用色彩数据结构的数据从色彩图像中取回色彩画面;根据补片的几何数据中包含的标识符,在存储在存储器中的3D网格的列表中取回3D网格。根据取回的3D网格的表面计算出的深度,通过使用去投影数据对色彩画面的像素进行去投影。因为3D网格的顶点坐标和各个面已知且本地化在3D场景的3D空间中,所以去投影过程能够根据网

格的表面上的去投影数据来计算去投影的像素的坐标。因此,场景的一部分被解码。

[0121] 当已经对每个补片数据项执行步骤103或步骤104时,整个3D场景已经被解码。在步骤105中,将解码的3D场景发送到另一模块,例如图3的呈现器35。

[0122] 自然地,本公开不限于先前描述的实施例。特别地,本公开不限于用于编码/解码承载表示三维场景(或三维场景的序列)的数据的流的方法和和设备,而是还扩展到以3个自由度方式或以体积方式(即,3DoF+或6DoF)呈现三维场景的方法或者实现这些方法的任何设备以及尤其是包括至少一个CPU和/或至少一个GPU的任何设备。

[0123] 本公开还涉及一种用于显示从包括表示三维场景的信息的数据流呈现的图像的方法(和配置的设备)。

[0124] 本公开还涉及一种用于发送和/或接收根据本原理编码的流的方法(和配置的设备)。

[0125] 本文描述的实现方式可以在例如方法或处理,装置,软件程序产品,数据流或信号中实现。即使仅在单个实现方式的形式上下文中讨论(例如,仅作为方法或设备讨论),讨论的特征的实现方式也可以以其他形式(例如,程序)实现。装置可以在例如适当的硬件,软件和固件中实现。方法可以在例如诸如处理器之类的装置中实现,该处理器通常涉及处理设备,一般包括例如计算机,微处理器,集成电路或可编程逻辑设备。处理器还包括通信设备,诸如例如智能电话、平板电脑、计算机,移动电话,便携式/个人数字助理(“PDA”),以及便于终端用户之间的信息通信的其他设备。

[0126] 本文描述的各种处理和特征的实现方式可以体现在各种不同的设备或应用中,特别是例如,与数据编码、数据解码、视图生成、纹理处理以及图像和相关纹理信息和/或深度信息的其他处理相关联的设备或应用中。这样的设备的示例包括编码器,解码器,处理来自解码器的输出的后处理器,提供输入给编码器的预处理器,视频编码器,视频解码器,视频编解码器,网络服务器,机顶盒,膝上型计算机,个人计算机,蜂窝电话,PDA以及其他通信设备。应该清楚的是,该设备可以是移动的,甚至可以安装在移动车辆中。

[0127] 另外,可以通过由处理器执行的指令来实现方法,并且可以将这样的指令(和/或由实现方式产生的数据值)存储在诸如例如集成电路,软件载体之类的处理器可读介质上,或其他存储设备上,诸如例如硬盘,致密盘(CD),光盘(诸如,例如DVD,通常称为数字多功能盘或数字视频盘),随机存取存储器(“RAM”)或只读存储器(“ROM”)。指令可以形成有形地体现在处理器可读介质上的应用程序。指令可以在例如硬件,固件,软件或其组合中。指令可以在例如操作系统,单独的应用或两者的组合中找到。因此,处理器的特征可以在于例如被配置为实施处理的设备和包括具有用于实施处理的指令的处理器可读介质的设备(诸如,存储设备)。此外,除了或代替指令,处理器可读介质可以存储由实现方式产生的数据值。

[0128] 对于本领域技术人员明显的是,实现方式可以产生各种信号,这些信号被格式化以承载可以例如被存储或传输的信息。该信息可以包括例如用于执行方法的指令,或者由描述的实现方式之一产生的数据。例如,信号可以被格式化以承载用于写入或读取描述的实施例的语法的规则作为数据,或者承载由描述的实施例写入的实际语法值作为数据。这样的信号可以被格式化,例如,作为电磁波(例如,使用频谱的射频部分)或者作为基带信号。格式化可以包括,例如,编码数据流和用编码的数据流调制载波。信号承载的信息可以是例如模拟或数字信息。如已知的,信号可以通过各种不同的有线或无线链路传输。信号可

以存储在处理器可读介质上。

[0129] 已经描述多个实现方式。然而,将理解,可以进行各种修改。例如,可以组合,补充,修改或删除不同实现方式的元素以产生其他实现方式。另外,普通技术人员将理解,其他结构和处理可以替代公开的那些,并且作为结果的实现方式将以至少基本相同的(多个)方式执行至少基本相同的(多个)功能,以至少实现与公开的实现方式基本相同的(多个)结果。因此,本申请考虑这些和其他实现方式。

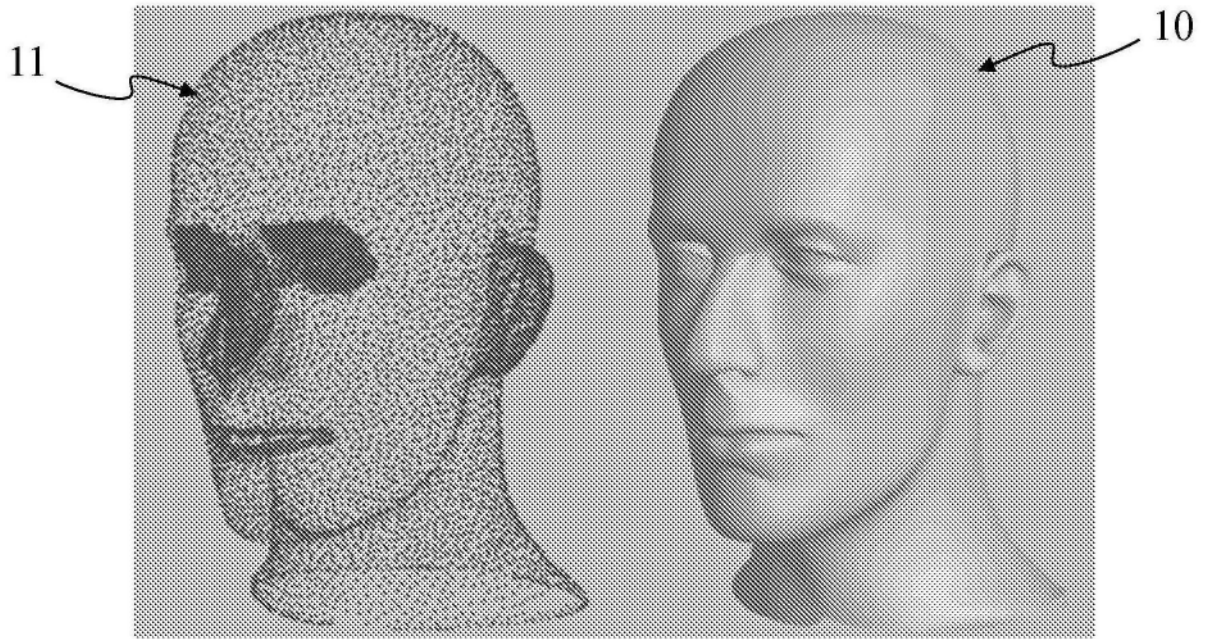


图1



图2

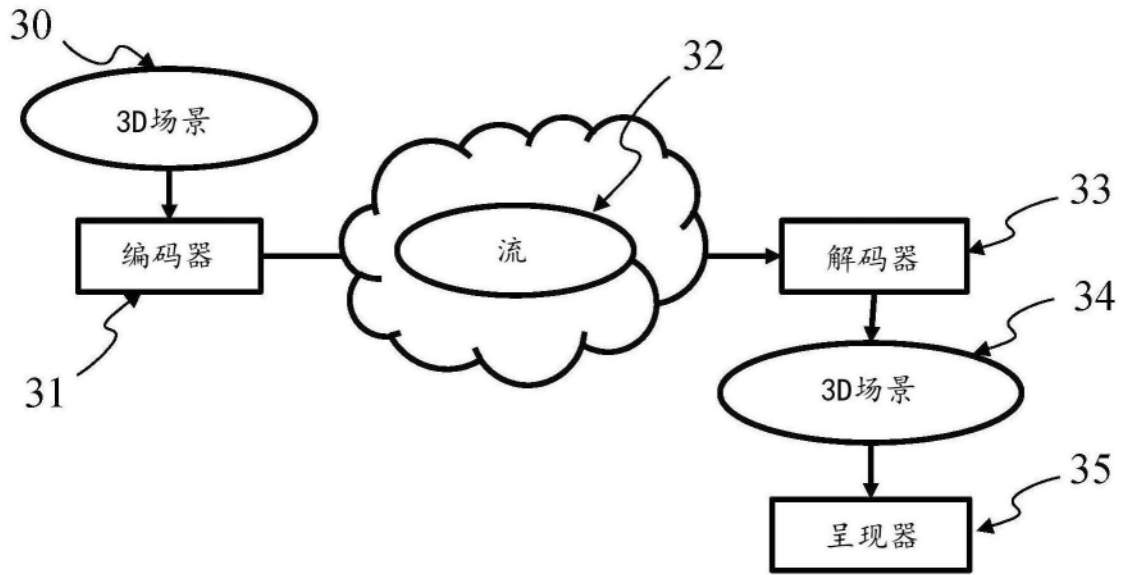


图3

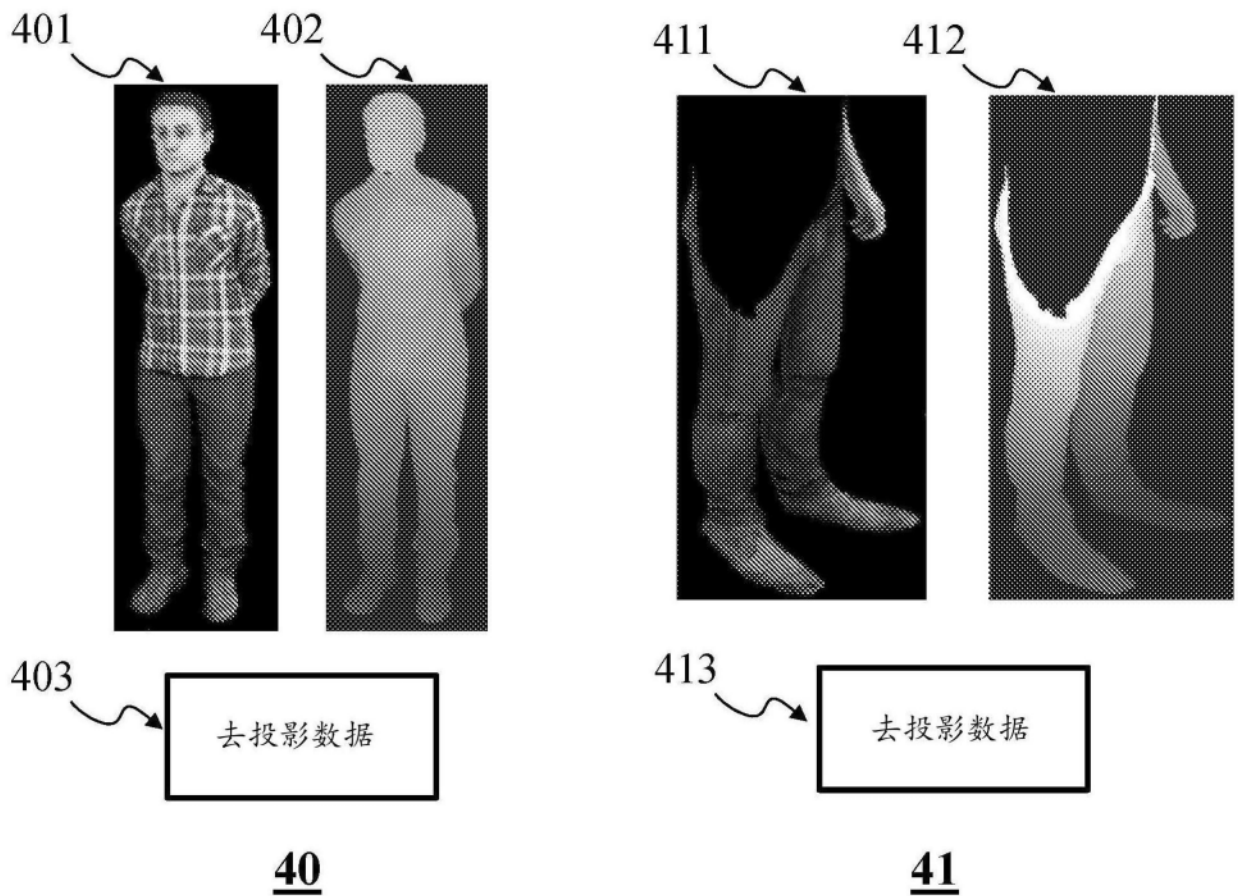


图4

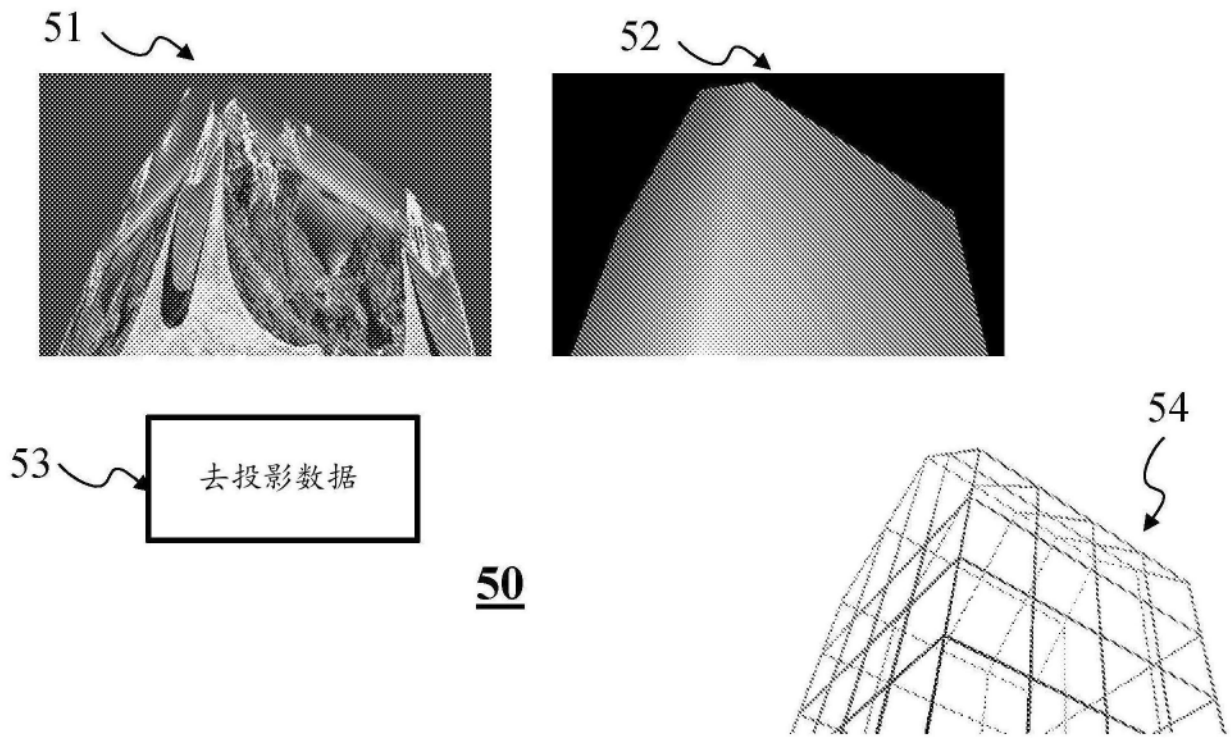


图5

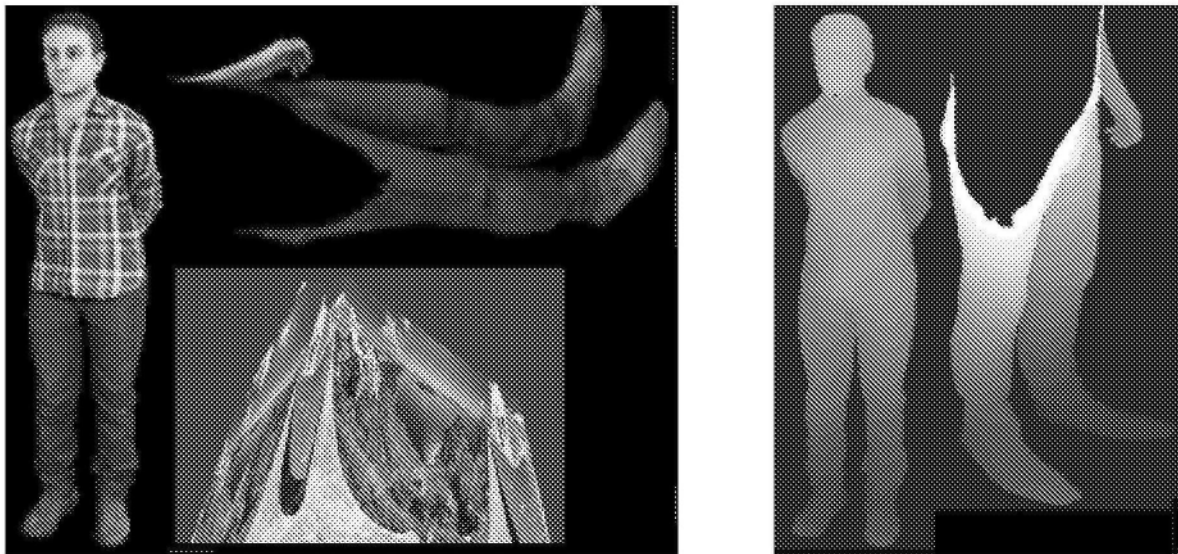


图6



图7

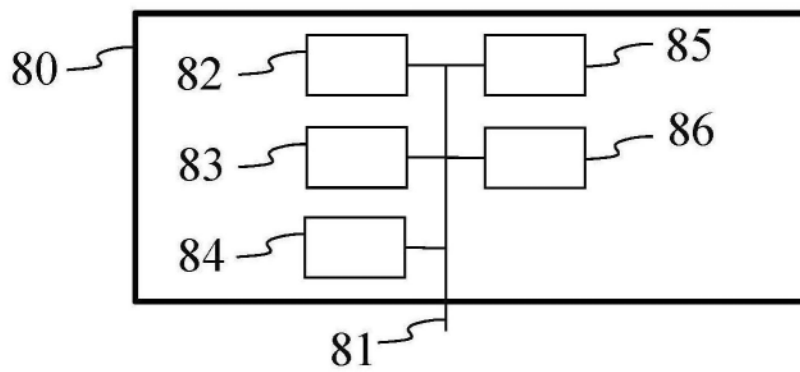


图8

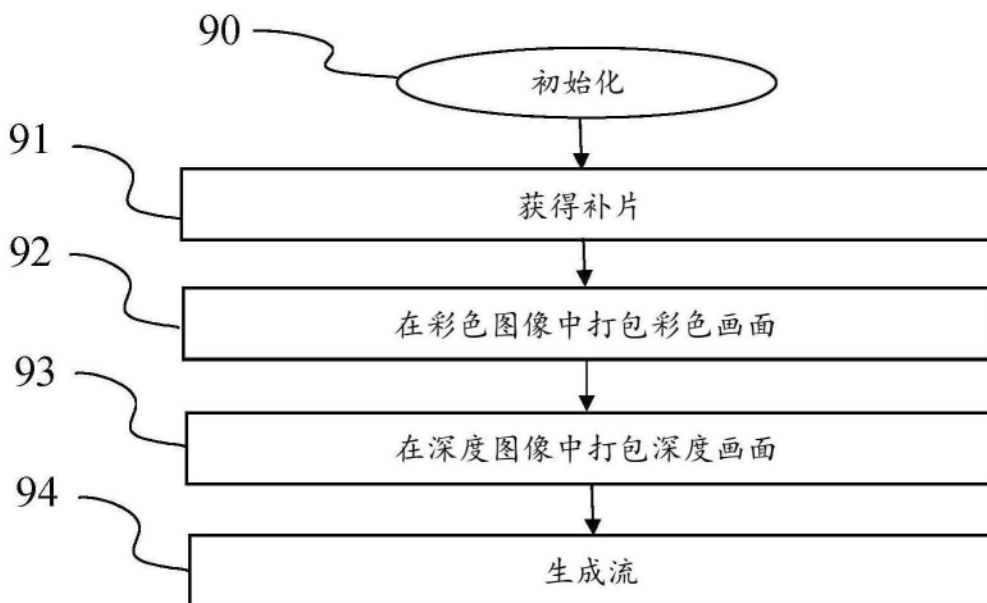


图9

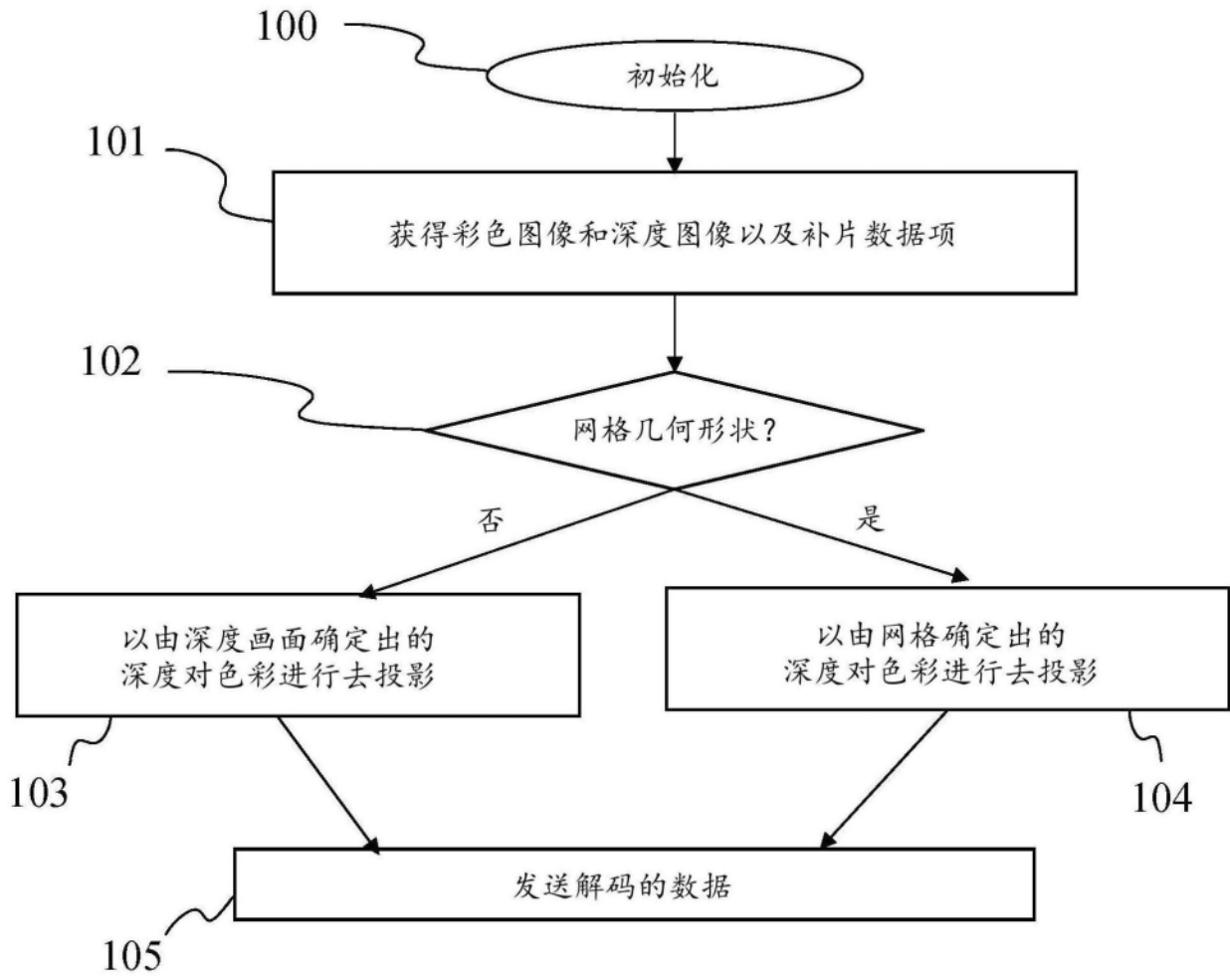


图10