

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2022 年 12 月 15 日 (15.12.2022)



(10) 国际公布号
WO 2022/257518 A1

(51) 国际专利分类号:
H04N 13/117 (2018.01)

(21) 国际申请号: PCT/CN2022/080257

(22) 国际申请日: 2022 年 3 月 11 日 (11.03.2022)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
202110659190.8 2021年6月11日 (11.06.2021) CN

(71) 申请人: 腾讯科技(深圳)有限公司 (TENCENT TECHNOLOGY (SHENZHEN) COMPANY LIMITED) [CN/CN]; 中国广东省深圳市南山区高新区科技中一路腾讯大厦 35 层, Guangdong 518057 (CN)。

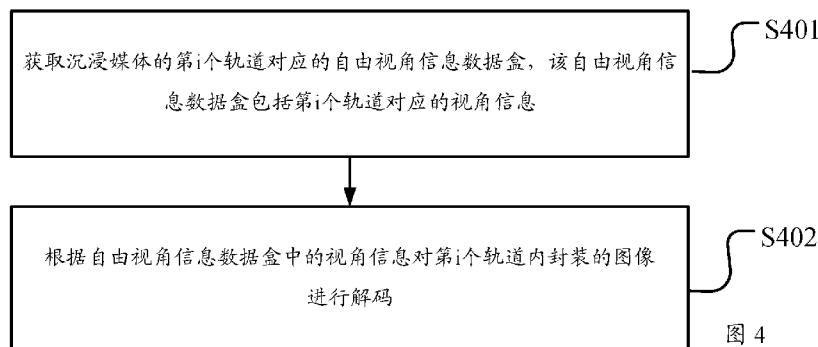
(72) 发明人: 胡颖 (HU, Ying); 中国广东省深圳市南山区高新区科技中一路腾讯大厦 35 层, Guangdong 518057 (CN)。

(74) 代理人: 深圳市深佳知识产权代理事务所(普通合伙) (SHENPAT INTELLECTUAL PROPERTY AGENCY); 中国广东省深圳市罗湖区南湖街道春风路庐山大厦 B 座 18C2、18D、18E、18E2, Guangdong 518001 (CN)。

(81) 指定国(除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH,

(54) **Title:** DATA PROCESSING METHOD AND APPARATUS FOR IMMERSIVE MEDIA, AND RELATED DEVICE AND STORAGE MEDIUM

(54) 发明名称: 沉浸媒体的数据处理方法、装置、相关设备及存储介质



S401 Obtain a free viewing angle information data box corresponding to an ith tack of immersive media, the free viewing angle information data box comprising viewing angle information corresponding to the ith track
S402 Decode, according to the viewing angle information in the free viewing angle information data box, an image encapsulated in the ith track

图 4

(57) **Abstract:** Embodiments of the present application disclose a data processing method and apparatus for immersive media, and a device and a storage medium. The immersive media is composed of images captured by N cameras at different viewing angles; the immersive media is encapsulated in M tracks; an image from at least one camera is encapsulated in one track, and N and M are both integers greater than 1. The method comprises: obtaining a free viewing angle information data box comprising viewing angle information corresponding to the ith tack of the immersive media, the free viewing angle information data box comprising viewing angle information corresponding to the ith track (S401), and i being an integer greater than or equal to 1 and less than or equal to M; and decoding, according to the viewing angle

PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK,
SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG,
US, UZ, VC, VN, WS, ZA, ZM, ZW。

(84) 指定国 (除另有指明, 要求每一种可提供的地区
保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ,
NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM,
AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG,
CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU,
IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT,
RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI,
CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

— 包括国际检索报告 (条约第21条(3))。

information in the free viewing angle information data box, an image encapsulated in the i th track (S402). By using the embodiments of the present application, a content consumption device may select an appropriate image to perform decoding consumption according to the viewing angle information in each track and the current user position.

(57) 摘要: 本申请实施例公开了一种沉浸媒体的数据处理方法、装置、设备及存储介质, 其中沉浸媒体是由处于不同视角的 N 个相机拍摄的图像组成, 沉浸媒体被封装到 M 个轨道, 一个轨道中封装来自至少一个相机的图像, N 和 M 均为大于1的整数; 包括: 获取沉浸媒体的第 i 个轨道对应的自由视角信息数据盒, 自由视角信息数据盒包括第 i 个轨道对应的视角信息 (S401), i 为大于或等于1且小于或等于 M 的整数; 根据自由视角信息数据盒中的视角信息对第 i 个轨道内封装的图像进行解码 (S402)。采用本申请实施例可以使得内容消费设备根据各个轨道中的视角信息以及用户当前位置选择合适的图像进行解码消费。

沉浸媒体的数据处理方法、装置、相关设备及存储介质

本申请要求于2021年06月11日提交中国专利局、申请号为202110659190.8、申请名称为“沉浸媒体的数据处理方法、装置、相关设备及存储介质”的中国专利申请的优先权，其全部内容通过引用结合在本申请中。

5 技术领域

本申请涉及音视频领域，尤其涉及媒体的数据处理。

背景技术

10 沉浸媒体是指能为消费者带来沉浸式体验的媒体内容，沉浸媒体又可以称为自由视角视频，自由视角视频通常是由相机阵列从多个角度对同一个三维场景进行拍摄，得到不同视角的深度图像和/或纹理图像，这些深度图像和/或纹理图像组成了自由视角视频。内容消费设备可以根据用户当前所在位置以及各个图像来源的相机视角，选择解码某些图像进行消费。

15 目前在自由视角视频制作过程中一般是采用大规模图集信息数据盒指示自由视角视频相关的参数信息（比如相机采集的深度图以及纹理图的分辨力宽度与高度、每个视角对应的相机标识符等等），从而省略图集轨道中的其余图集信息。

发明内容

20 本申请实施例提供了一种沉浸媒体的数据处理方法、装置、设备及存储介质，可以将一沉浸媒体的不同视角相机采集的图像封装到多个不同轨道，并采用每个轨道对应的自由视角信息数据盒来指示每个轨道中图像来源相机的视角信息，以便于内容消费设备根据各个轨道中的视角信息以及用户当前位置选择合适的图像进行解码消费。

一方面，本申请实施例提供了一种沉浸媒体的数据处理方法，该沉浸媒体是由处于不同视角的N个相机拍摄的图像组成，所述沉浸媒体被封装到M个轨道，一个轨道中封装来自至少一个相机的图像，N和M均为大于1的整数，数据处理方法包括：

25 获取第i个轨道对应的自由视角信息数据盒，所述自由视角信息数据盒子，包括所述第i个轨道对应的视角信息，i为大于等于1且小于等于M的整数；

根据所述自由视角信息数据盒中的视角信息对所述第i个轨道内封装的图像进行解码。

一方面，本申请实施例提供了另一种沉浸媒体的处理方法，包括：

30 将沉浸媒体封装到M个轨道中，所述沉浸媒体是由处于不同视角的N个相机拍摄的图像组成的，一个轨道中封装来自至少一个相机的图像，所述M个轨道属于同一个轨道组，N和M均为大于或等于1的整数；

根据第i个轨道中w个图像的封装过程，生成第i个轨道对应的自由视角信息数据盒，所述自由视角信息数据盒包括所述第i个轨道对应的视角信息； $1 \leq i \leq M$ ， $w \geq 1$ 。

一方面，本申请实施例提供了一种沉浸媒体的数据处理装置，包括：

35 获取单元，用于获取沉浸媒体的第i个轨道对应的自由视角信息数据盒，所述自由视角信息数据盒子，包括所述第i个轨道对应的视角信息，沉浸媒体是由处于不同视角的N个相机拍摄的图像组成，所述沉浸媒体被封装到M个轨道，一个轨道中封装来自至少一个相机的图像，N和M均为大于1的整数；i为大于或等于1且小于或等于M的整数；

-2-

解码单元，用于根据所述自由视角信息数据盒中的视角信息对所述第*i*个轨道内封装的图像进行解码。

一方面，本申请实施例提供了另一种沉浸媒体的数据处理装置，包括：

封装单元，用于将沉浸媒体封装到*M*个轨道中，所述沉浸媒体是由处于不同视角的*N*个相机拍摄的图像组成的，一个轨道中封装来自至少一个相机的图像，所述*M*个轨道属于同一个轨道组，*N*和*M*均为大于或等于1的整数；

生成单元，用于根据第*i*个轨道中*w*个图像的封装过程，生成第*i*个轨道对应的自由视角信息数据盒，所述自由视角信息数据盒包括所述第*i*个轨道对应的视角信息； $1 \leq i \leq M$ ， $w \geq 1$ 。

一方面，本申请实施例提供了一种内容消费设备，包括：

处理器，适于实现一条或多条计算机程序；以及

计算机存储介质，所述计算机存储介质存储有一条或多条计算机程序，所述一条或计算机程序程序适于由处理器加载并执行：

获取沉浸媒体的第*i*个轨道对应的自由视角信息数据盒，所述自由视角信息数据盒子，包括所述第*i*个轨道对应的视角信息；该沉浸媒体是由处于不同视角的*N*个相机拍摄的图像组成，所述沉浸媒体被封装到*M*个轨道，一个轨道中封装来自至少一个相机的图像，*N*和*M*均为大于1的整数，*i*为大于或等于1且小于或等于*M*的整数；

根据所述自由视角信息数据盒中的视角信息对所述第*i*个轨道内封装的图像进行解码。

一方面，本申请实施例提供了一种内容制作设备，包括：

处理器，适于实现一条或多条计算机程序；以及

计算机存储介质，所述计算机存储介质存储有一条或多条计算机程序，所述一条或计算机程序程序适于由处理器加载并执行：

将沉浸媒体封装到*M*个轨道到*M*个轨道中，所述沉浸媒体是由处于不同视角的*N*个相机拍摄的图像组成的，一个轨道中封装来自至少一个相机的图像，所述*M*个轨道属于同一个轨道组，*N*和*M*均为大于或等于1的整数；

根据第*i*个轨道中*w*个图像的封装过程，生成第*i*个轨道对应的自由视角信息数据盒，所述自由视角信息数据盒包括所述第*i*个轨道对应的视角信息； $1 \leq i \leq M$ ， $w \geq 1$ 。

又一方面，本申请实施例提供一种存储介质，所述存储介质用于存储计算机程序，所述计算机程序用于执行以上方面的方法。

又一方面，本申请实施例提供了一种包括指令的计算机程序产品，当其在计算机上运行时，使得所述计算机执行以上方面的方法。

本申请实施例中将一沉浸媒体封装至*M*个轨道，该沉浸媒体是由处于不同视角的*N*个相机拍摄的图像组成的，一个轨道中可以包括来自至少一个相机的图像，这*M*个轨道是属于同一个轨道组的，实现了一个沉浸视频封装到多个轨道的场景；另外，内容制作设备为每个轨道生成一个自由视角信息数据盒，通过第*i*个轨道对应的自由视角信息数据盒指示第*i*个轨道对应的视角信息，比如相机所在的具体视角位置，进而内容消费设备在根据第*i*个轨道对应的视角信息进行该轨道内图像的解码时，可以保证解码的视频与用户所在位置更加

匹配，提高沉浸视频的呈现效果。

附图说明

- 图1a是本申请实施例提供的一种用户消费3DoF沉浸媒体的示意图；
图1b是本申请实施例提供的一种用户消费3DoF+沉浸媒体的示意图；
5 图1c是本申请实施例提供的一种用户消费6DoF沉浸视频的示意图；
图2a是本申请实施例提供的一种沉浸媒体系统的架构图；
图2b是本申请实施例提供的一种沉浸媒体的传输方案的示意图；
图3a是本申请实施例提供的一种视频编码基本框图；
图3b是本申请实施例通过的一种输入图像划分的示意图；
10 图4是本申请实施例提供的一种沉浸媒体的数据处理方法的流程示意图；
图5是本申请实施例提供的另一种沉浸媒体的数据处理方法的流程示意图；
图6是本申请实施例提供的一种沉浸媒体的数据处理装置的结构示意图；
图7是本申请实施例提供的另一种沉浸媒体的数据处理装置的结构示意图；
图8是本申请实施例提供的一种内容消费设备的结构示意图；
15 图9是本申请实施例提供的一种内容制作设备的结构示意图。

具体实施方式

下面将结合本申请实施例中的附图，对本申请实施例中的技术方案进行清楚、完整地描述。

- 本申请实施例涉及到沉浸媒体的数据处理技术，所谓沉浸媒体是指能够提供沉浸式的
20 媒体内容，使沉浸于该媒体内容中的用户能够获得现实世界中视觉、听觉等感官体验的媒体文件。具体的，沉浸媒体可以是三自由度（3Degree of Freedom, 3DoF）沉浸媒体，3DoF+沉浸媒体或者六自由度（6Degree of Freedom, 6DoF）沉浸媒体。

- 参见图1a，为本申请实施例提供的一种用户消费3DoF沉浸媒体的示意图，图1a所示的3DoF沉浸媒体是指用户在一个三维空间的中心点固定，用户头部沿着X轴、Y轴和Z轴旋转来观看沉浸媒体的媒体内容提供的画面。参见图1b，为本申请实施例提供的一种用户消费3DoF+沉浸媒体的示意图，3DoF+是指当沉浸媒体提供的虚拟场景具有一定的深度信息，用户头部可以基于3DoF在一个有限的空间内移动来观看媒体内容提供的画面。参见图1c，为本申请实施例提供的一种用户消费6DoF沉浸视频的示意图，6DoF分为窗口6DoF、全方向6DoF和6DoF，其中，窗口6DoF是指用户在X轴、Y轴的旋转移动受限，以及在Z轴的平移受限；例如，用户不能够看到窗户框架外的景象，以及用户无法穿过窗户。全方向6DoF是指用户在X轴、Y轴和Z轴的旋转移动受限，例如，用户在受限的移动区域中不能自由的穿过三维的360度VR内容。6DoF是指用户可以沿着X轴、Y轴、Z轴自由平移，例如，用户可以在三维的360度VR内容中自由的走动。简单来讲，6DoF沉浸视频不仅可以允许用户沿着X轴、Y轴以及Z轴旋转消费媒体内容，还可以沿着X轴、Y轴以及Z轴自由运动来消费媒体内容。
35

沉浸媒体内容包括以各种形式在三维（3-Dimension, 3D）空间中表示的视频内容，例如以球面形式表示的三维视频内容。具体地，沉浸媒体内容可以是VR（Virtual Reality，虚

—4—

拟现实) 视频内容、多视角视频内容、全景视频内容、球面视频内容或360度视频内容; 所以, 沉浸媒体又可称为VR视频、自由视角视频、全景视频、球面视频或360度视频。另外, 沉浸媒体内容还包括与三维空间中表示的视频内容相同步的音频内容。

参见图2a, 为本申请实施例提供的一种沉浸媒体系统的架构图。在图2a所示的沉浸媒体系统中包括内容制作设备和内容消费设备, 内容制作设备可以指沉浸媒体的提供者(例如沉浸内容的内容制作者)所使用的计算机设备, 该计算机设备可以是终端, 比如智能手机、平板电脑、笔记本电脑、台式计算机、智能音箱、智能手表、智能车载等; 该计算机设备也可以是服务器, 比如独立的物理服务器, 也可以是多个物理服务器构成的服务器集群或者分布式系统, 还可以是提供云服务、云数据库、云计算、云函数、云存储、网络服务、云通信、中间件服务、域名服务、安全服务、CDN、以及大数据和人工智能平台等基础云计算服务的云服务器。

内容消费设备可以指沉浸媒体的使用者(例如用户)所使用的计算机设备, 该计算机设备可以是终端, 比如个人计算机、智能移动设备如智能手机、VR设备(例如VR头盔、VR眼镜等)。沉浸媒体的数据处理过程包括在内容制作设备侧的数据处理过程以及在内容消费设备侧的数据处理过程。

在内容制作设备侧的数据处理过程主要包括: (1) 沉浸媒体的媒体内容的获取与制作过程; (2) 沉浸媒体的编码及封装的过程。在内容消费侧的数据处理过程主要包括: (3) 沉浸媒体的解封装及解码的过程; (4) 沉浸媒体的渲染过程。另外, 内容制作设备与内容消费设备之间涉及沉浸媒体的传输过程, 该传输过程可以基于各种传输协议来进行, 此处的传输协议可以包括但不限于: DASH (Dynamic Adaptive Streaming over HTTP, 动态自适应流媒体传输) 协议、HLS (HTTP Live Streaming, 动态码率自适应传输) 协议、SMTP (Smart Media Transport Protocol, 智能媒体传输协议)、TCP (Transmission Control Protocol, 传输控制协议) 等。

参见图2b, 为本申请实施例提供的一种沉浸媒体的传输方案的示意图。如图2b所示, 为了解决沉浸媒体自身数据量过大带来的传输带宽负荷问题, 在沉浸媒体的处理过程中, 通常选择将原始的沉浸媒体在空间上切分为多个分块视频后, 分别编码后进行封装, 再传输给客户端消费。下面结合图2b, 分别对沉浸媒体的数据处理过程进行详细介绍。首先介绍内容制作设备侧的数据处理过程:

(1) 沉浸媒体的媒体内容的获取与制作过程:

S1: 沉浸媒体的媒体内容的获取过程。

沉浸媒体的媒体内容是通过捕获设备采集现实世界的声音-视觉场景获得的。在一个实施例中, 捕获设备可以是指设于内容制作设备中的硬件组件, 例如捕获设备是指终端的麦克风、摄像头以及传感器等等。在其他实施例中, 该捕获设备也可以是独立与内容制作设备但与内容制作设备相连接的硬件装置, 例如与服务器相连接的摄像头。该捕获设备可以包括但不限于: 音频设备、摄像设备及传感设备。其中, 音频设备可以包括音频传感器、麦克风等。摄像设备可以包括普通摄像头、立体摄像头、光场摄像头等。传感设备可以包括激光设备、雷达设备等。捕获设备的数据可以为多个, 这些捕获设备可以被部署在现实

空间中的一些特定视角以同时捕获该空间内不同视角的音频内容以及视频内容，捕获的音频内容和视频内容在时间和空间上均保持同步。举例来说，3DoF沉浸内容的媒体内容是由一组摄像机或一个带有多个摄像头和传感器的摄像设备录制的，6DoF沉浸媒体的媒体内容主要由相机阵列拍摄得到的点云、光场等形式的内容制作而成。

5 S2: 沉浸媒体的媒体内容的制作过程。

捕获到的音频内容本身就是适合被执行沉浸媒体的音频编码的内容，因此无需对捕获到的音频内容进行其他处理。而捕获到的视频内容需要进行一系列制作流程后才可以称为适合被执行沉浸媒体的视频编码的内容，该制作流程具体可以包括：

10 ①拼接，由于捕获到的沉浸媒体的视频内容是捕获设备在不同视角下拍摄得到的，拼接就是指对这些各个视角拍摄的视频内容拼接成一个完整的、能够反映现实空间360度视觉全景的视频，即拼接后的视频是一个在三维空间表示的全景视频。

②投影，投影就是指将拼接形成的一个三维视频映射到一个二维（2-Dimension，2D）图像上的过程，投影形成的2D图像称为投影图像；投影的方式可包括但不限于：经纬图投影、正六面体投影。

15 需要说明的是，由于采用捕获设备只能捕获到全景视频，这样的视频经内容制作设备处理并传输至内容消费设备进行相应的数据处理后，内容消费设备侧的用户只能通过一些特定动作（如头部旋转）来观看360度的视频信息，而执行非特定动作（如移动头部）并不能获得相应的视频变化，VR体验不佳，因此需要额外提供与全景视频相匹配的深度信息，来使用户获得更优的沉浸度和更佳的VR体验，这就涉及多种制作技术，常见的制作技术包括6DoF制作技术、3DoF制作技术以及3DoF+制作技术。

20 采用6DoF制作技术和3DoF+制作技术得到的沉浸媒体可以包括自由视角视频。自由视角视频作为一种常见的3DoF+和6DoF沉浸式媒体，其是由多个相机采集的、包含不同视角的，支持用户3DoF+或6DoF交互的沉浸式媒体视频。具体的，3DoF+沉浸式媒体是由一组相机或一个带有多个摄像头和传感器的相机录制而成，相机通常可以获取在设备中心周围所有方向的内容。6DoF沉浸式媒体主要由相机阵列拍摄得到的点云、光场等形式的内容制作而成。

（3）沉浸媒体的媒体内容的编码过程。

25 投影图像可以被直接进行编码，也可以对投影图像进行区域封装之后再进行编码。参见图3a，为本申请实施例提供的一种视频编码基本框图。现代主流视频编码技术，以国际视频编码标准HEVC（High Efficiency Video Coding），国际视频编码标准VVC（Versatile Video Coding），以及中国国家视频编码标准AVS（Audio Video Coding Standard）为例，采用了混合编码框架，对输入的原始视频信号，进行了如下一系列的操作和处理：

35 1）块划分结构（block partition structure）：根据处理单元的大小将输入图像划分成若干个不重叠的处理单元，对每个处理单元进行类似的压缩操作。这个处理单元被称为编码树单元（Coding Tree Unit，CTU），或者最大编码单元（Largest Coding Unit，LCU）。CTU可以继续更加精细的划分，得到一个或多个基本编码的单元，称为编码单元（Coding Unit，CU）。每个CU是一个编码缓解中最基本的元素。参见图3b，为本申请实施例通过的一种输

—6—

入图像划分的示意图。以下描述的是对每一个CU可能采用的各种编码方式。

2) 预测编码 (Predictive Coding): 包括帧内预测 (Intra (picture) Prediction) 和帧间预测 (Inter (picture) Prediction)。沉浸媒体的原始视频信号经过选定的已重建视频信号的预测后, 得到残差视频信号。内容制作设备需要为当前CU决定在众多可能的预测编码模式中, 选择最合适的一种, 并告知内容消费设备。其中, 帧内预测所预测的信号来自于同一个图像内已经过编码重建过的区域, 帧间预测所预测的信号来自已经编码过的, 不同于当前图像的其他图像 (称之为参考图像)。

3) 变换编码及量化 (Transform & Quantization): 残差视频信号经过离散傅里叶变换 (Discrete Fourier Transform, DFT), 离散余弦变换 (Discrete Cosine Transform, DCT) 等变换操作, 将信号转换到变换域中, 称之为变换系数。在变换域中的信号, 进一步的进行有损的量化操作, 丢失掉一定的信息, 使得量化后的信号有利于压缩表达。在一些视频编码标准中, 可能有多于一种变换方式可以选择, 因此, 内容制作设备也需要为当前编码CU选择其中的一种变换, 并告知内容播放设备。量化的精细程度通常由量化参数 (Quantization Parameter, QP) 来决定, QP取值较大, 表示更大取值范围的系数将被量化为同一个输出, 因此通常会带来更大的失真, 及较低的码率; 相反, QP取值较小, 表示较小取值范围的系数将被量化为同一个输出, 因此通常会带来较小的失真, 同时对应较高的码率。

4) 熵编码 (Entropy Coding) 或统计编码: 量化后的变换域信号, 将根据各个值出现的频率, 进行统计压缩编码, 最后输出二值化 (0或者1) 的压缩码流。同时, 编码产生其他信息, 例如选择的模式, 运动矢量等, 也需要进行熵编码以降低码率。统计编码是一种无损编码方式, 可以有效的降低表达同样的信号所需要的码率。常见的统计编码方式有变长编码 (VLC, Variable Length Coding) 或者基于上下文的二值化算术编码 (CABAC, Content Adaptive Binary Arithmetic Coding)。

5) 环路滤波 (Loop Filtering): 已经编码过的图像, 经过反量化, 反变换及预测补偿的操作 (上述2~4的反向操作), 可获得重建的解码图像。重建图像与原始图像相比, 由于存在量化的影响, 部分信息与原始图像有所不同, 产生失真 (Distortion)。对重建图像进行滤波操作, 例如去块效应滤波 (deblocking), 取样自适应偏移 (Sample Adaptive Offset, SAO) 滤波器或者自适应环路滤波器 (Adaptive Loop Filter, ALF) 等, 可以有效的降低量化所产生的失真程度。由于这些经过滤波后的重建图像, 将作为后续编码图像的参考, 用于对将来的信号进行预测, 所以上述的滤波操作也被称为环路滤波, 及在编码环路内的滤波操作。

此处需要说明的是, 如果采用6DoF (Six Degrees of Freedom, 六自由度) 制作技术 (用户可以在模拟的场景中较自由的移动时, 称为6DoF), 在视频编码过程中需要采用特定的编码方式 (如点云编码) 进行编码。

(4) 沉浸媒体的封装过程。

将音频码流和视频码流按照沉浸媒体的文件格式 (如ISO/BMFF (ISO Base Media File Format, ISO基媒体文件格式)) 封装到文件容器 (轨道) 中形成沉浸媒体的媒体资源文件, 该媒体资源文件可以是媒体文件或者媒体片段形成的沉浸媒体的媒体文件, 并按照沉浸媒体的文件格式要求采用媒体呈现描述信息 (Media presentation description, MPD) 记录该沉

浸媒体的媒体文件资源的元数据，此处的元数据是对于沉浸媒体的呈现有关的信息的总称，该元数据可以包括对媒体内容的描述信息、对视窗的描述信息以及对媒体内容呈现相关的信令信息等等。如图2a所示，内容制作设备会存储经过数据处理过程之后形成的媒体呈现描述信息和媒体文件资源。

5 下面介绍内容消费设备侧的数据处理过程：

(1) 沉浸媒体的解封以及解码的过程：

内容消费设备可以通过内容制作设备的推荐或者按照内容消费设备侧用户需求自适应动态从内容制作设备获得沉浸媒体的媒体文件资源和相应媒体呈现描述信息，例如内容消费设备可以根据用户的头部/眼睛/身体的跟踪信息确定用户的朝向和位置，再基于确定的朝向和位置动态向内容制作设备请求获得相应的媒体文件资源。媒体文件资源和媒体呈现描述信息通过传输机制（如DASH、SMT）由内容制作设备传输给内容消费设备。内容消费设备的解封装过程与内容制作设备的封装过程是相逆的，内容消费设备按照沉浸媒体的文件格式要求对获取到的媒体文件资源进行解封装，得到音频码流和视频码流。内容消费设备的解码过程与内容制作设备的编码过程是相逆的，内容消费设备对音频码流进行音频解
10 码，还原出音频内容，以及内容消费设备对视频码流进行解码，得到视频内容。其中，内容消费设备对视频码流的解码过程可以包括如下：①对视频码流进行解码，得到平面的投影图像。②根据媒体呈现描述信息将投影图像进行重建处理以转换为3D图像，此处的重建处理是指二维的投影图像重新投影至3D空间中的处理。

根据上述编码过程可以看出，在内容消费设备侧，对于每一个CU，内容消费设备获得
20 压缩码流后，先进行熵解码，获得各种模式信息以及量化后的变换系数。各个系数经过反复量化以及变换，得到残差信号。另一方面，根据已知的编码模式信息，可获得该CU对应的预测信号，两者相加之后，即可得到重建信号。最后解码图像的重建值，需要经过环路滤波的操作，产生最终的输出信号。

(2) 沉浸媒体的渲染过程。

内容消费设备根据媒体呈现描述信息中与渲染、视窗相关的元数据对音频解码得到的音频内容以及视频解码得到的3D图像进行渲染，渲染完成即实现了对该3D图像的播放输出。特别地，如果采用3DoF和3DoF+的制作技术，内容消费设备主要基于当前视点、视差、深度信息等对3D图像进行渲染，如果采用6DoF的制作技术，内容消费设备主要基于当前视点对视窗内的3D图像进行渲染。其中，视点指用户的观看位置点，视差是指用户的双目产生的
30 的视线差或由于运动产生的视线差，视窗是指观看区域。

上述描述的沉浸媒体系统支持数据盒(Box)，数据盒是指包括元数据的数据块或对象，即数据盒子中包括了相应媒体内容的元数据。由上述沉浸媒体的数据处理过程可知，在对沉浸媒体进行编码后，需要对编码后的沉浸媒体进行封装并传输给用户。本申请实施例中沉浸媒体主要指自由视角视频，现有技术中考虑到自由视角视频在制作过程中，图集信息
35 仅有相机参数即可获取，纹理图像和深度图像在平面帧中的位置也较为固定时，可使用大规模图集信息数据盒指示相关的参数信息，从而省略图集轨道中的其余图集信息。

具体实现中，大规模图集信息数据盒的语法可参见下述代码段1所示：

—8—

```

aligned(8) class V3CLargeScaleAtlasBox extends FullBox('vlsa',0,0) {
  unsigned int(8) camera_count;
    unsigned int(8) padding_size_depth;
    unsigned int(8) padding_size_texture;
5    for(i=0; i<camera_count; i++){
      unsigned int(8) camera_id;
      float(32) camera_resolution_x;
      float(32) camera_resolution_y;
      unsigned int(8) depth_downsample_factor;
10     unsigned int(32) texture_vetex_x;
      unsigned int(32) texture_vetex_y;
      unsigned int(32) depth_vetex_x;
      unsigned int(32) depth_vetex_y;
      unsigned int(32) camera_para_length;
15     for (i=0; i < camera_para_length; i++) {
        bit(8) camera_parameter;
    }
  }
}

```

上述代码段1所示语法的语义如下: camera_count表示采集沉浸媒体的所有相机的个数; padding_size_depth表示在对深度图像进行编码时采用的保护带宽度; padding_size_texture表示对纹理图像进行编码时采用的保护带宽度; camera_id表示处于一个视角相机的相机标识符, camera_resolution_x表示一个相机采集的纹理图像、深度图像的分辨率宽度, camera_resolution_y表示一个相机采集的纹理图像、深度图像的分辨率高度; depth_downsample_factor表示深度图像降采样的倍数因子, 深度图像的实际分辨率宽度与高度为相机采集分辨率宽度与高度的1/2 depth_downsample_factor; depth_vetex_x表示深度图像左上顶点相对于平面帧原点(平面帧的左上顶点)偏移的横轴分量, depth_vetex_y表示深度图像左上顶点相对于平面帧原点偏移的纵轴分量; texture_vetex_x表示纹理图像的左上顶点相对于平面帧原点偏移的横轴分量, texture_vetex_y表示纹理图像的左上顶点相对于平面帧原点偏移的纵轴分量; camera_para_length表示容积视频重构时所需的相机参数的长度, 以字节为单位; camera_parameter表示容积视频重构时所需的相机参数。

从上面的大规模图集信息数据盒的语法中可以看出, 虽然大规模图集信息数据盒中指示了自由视角视频帧中纹理图像和深度图像的布局信息, 并给出了相关的相机参数比如 camera_resolution_y以及camera_resolution_x等, 但是上述只考虑了自由视角视频封装在单轨道的场景, 没有考虑到自由视角视频被封装到多轨道的场景。并且, 上述大规模图集信息数据盒中指示了自由视角视频内纹理图和深度图的排布信息以及相关的相机参数, 但是只考虑了将自由视角视频封装到单轨道的情况, 并未考虑多轨道封装的场景, 另外大规模

图集信息数据盒中指示的相机参数，无法作为内容消费设备选择不同视角的图像进行解码消费的依据，也就是说根据上述大规模图集信息数据盒中记载的相机参数，内容消费设备无法知道哪个图像是适合当前用户位置信息的，从而给内容消费设备解码带来不便。

基于此，本申请实施例提供了一种沉浸媒体的数据处理方案，在该数据处理方案中将沉浸媒体封装到M个轨道中，该沉浸媒体是由处于不同视角的N个相机拍摄的图像组成的，这M个轨道是属于同一个轨道组的，实现了一个沉浸视频封装到多个轨道的场景；另外，内容制作设备为每个轨道生成一个自由视角信息数据盒，通过第i个轨道对应的自由视角信息数据盒指示第i个轨道对应的视角信息，比如相机所在的具体视角位置，进而内容消费设备在根据第i个轨道对应的视角信息进行该轨道内图像的解码显示时，可以保证解码显示的视频与用户所在位置更加匹配，提高沉浸视频的呈现效果。

基于上述描述，本申请实施例提供了一种沉浸媒体的数据处理方法，参见图4，为本申请实施例提供的一种沉浸媒体的数据处理方法的流程示意图。图4所述的数据处理方法可由内容消费设备执行，具体可由内容消费设备的处理器执行。图4所示的数据处理方法可包括如下步骤：

S401、获取沉浸媒体的第i个轨道对应的自由视角信息数据盒，该自由视角信息数据盒包括第i个轨道对应的视角信息。

其中，沉浸媒体是由处于不同视角的N个相机拍摄的图像组成的，沉浸媒体被封装到M个轨道中，一个轨道中封装来自至少一个相机拍摄的图像，这M个轨道属于同一个轨道组。第i个轨道是指M个轨道中被选择的一个，后面具体介绍如何从M个轨道中选择第i个轨道。

封装了沉浸视频的这M个轨道通过每个轨道内封装的自由视角轨道组信息关联的。下面以第i个轨道为例，第i个轨道内封装有第i个轨道对应的自由视角轨道组信息，该自由视角轨道组信息用于指示第i个轨道与封装了沉浸视频的其他轨道属于同一个轨道组。第i个轨道内封装的自由视角轨道组信息可以通过扩展轨道组数据盒得到，第i个轨道内封装的自由视角轨道组信息的语法可以表示为如下代码段2所示：

```
aligned(8) class AvsFreeViewGroupBox extends TrackGroupTypeBox('afvg') {  
    // track_group_id is inherited from TrackGroupTypeBox;  
    unsigned int(8) camera_count;  
    for(i=0; i<camera_count; i++){  
        unsigned int(32) camera_id;  
    }  
    unsigned int(2) depth_texture_type;  
    bit(6) reserved;  
}
```

在上述代码段2中，自由视角轨道组信息是通过扩展轨道组数据盒得到，以“a3fg”轨道组类型标识。在所有包含“afvg”类型TrackGroupTypeBox（轨道组数据盒）的轨道中，组ID相同的轨道属于同一个轨道组。

在代码段2中，假设第i个轨道内封装了前述N个相机中k个相机采集的图像，任一相机

- 10 -

采集的图像可以是纹理图像或深度图像中的至少一种。代码段2中，自由视角轨道组信息的语法的含义可以如下：camera_count表示第i个轨道组内包含纹理图像或深度图像来源的相机的数目，例如camera_count=k，表示第i个轨道内的封装了来自k个相机的纹理图像和/或深度图像。一个相机对应一个标识信息，每个标识信息存储在自由视角轨道组信息的第一相机标识字段camera_id中，应当理解的，一个相机对应一个标识信息，第i个轨道内包括k个相机，因此，自由轨道组信息中包括k个第一相机标识字段camera_id。

自由轨道组信息中还包括图像类型字段depth_texture_type，该图像类型字段用于指示第j个相机采集图像所属的图像类型，j大于0且小于k，也就是说一个图像类型字段depth_texture_type用于指示一个相机采集图像所属的图像类型中，由于第i个轨道中的纹理图像或者深度图像来源于k个相机，因此，自由轨道组信息中包括k个图像类型字段。

指示第j个相机采集的图像所属图像类型的图像类型字段depth_texture_type具体可参见下述表1所示：

表1

depth_texture_type	含义
0	保留
1	表明包含对应相机拍摄的纹理图像
2	表明包含对应相机拍摄的深度图像
3	表明包含对应相机拍摄的纹理图像和深度图像

具体来讲，当图像类型字段depth_texture_type为1时，表明第j个相机采集的图像所属图像类型为纹理图像；当图像类型字段depth_texture_type为2时，表明第j个相机采集的图像所属图像类型为深度图像；当图像类型字段depth_texture_type为3时指示第j个相机采集的图像所属图像类型为纹理图像和深度图像。

由S401可知，第i个轨道对应的自由视角信息数据盒AvsFreeViewInfoBox中包括了第i个轨道对应的视角信息，参见下代码段3，为第i个轨道对应的自由视角信息数据盒的语法表示，下面结合代码段3具体介绍第i个轨道对应的自由视角信息数据盒以及自由视角信息数据盒中包括的视角信息。其中，代码段3具体如下：

```

aligned(8) class AvsFreeViewInfoBox extends FullBox('afvi'){
    unsigned int(8) stitching_layout;
    unsigned int(8) texture_padding_size;
    unsigned int(8) depth_padding_size;
    unsigned int(8) camera_model;
    unsigned int(8) camera_count;
    for(i=0; i<camera_count; i++){
        unsigned int(32) camera_id;
        signed int(32) camera_pos_x;
        signed int(32) camera_pos_y;
        signed int(32) camera_pos_z;
    }
}

```

— 11 —

```

signed int(32) camera_focal_length_x;
signed int(32) camera_focal_length_y;
unsigned int(32) camera_resolution_x;
unsigned int(32) camera_resolution_y;
5 unsigned int(8) depth_downsample_factor;
unsigned int(32) texture_vetex_x;
unsigned int(32) texture_vetex_y;
unsigned int(32) depth_vetex_x;
unsigned int(32) depth_vetex_y;
10 ExtendedParaStruct();
}
}
aligned(8) class ExtendedParaStruct t() {
unsigned int(8) para_num;
15 for(i=0; i<para_num; i++){
    unsigned int(8) para_type;
    unsigned int(16) para_length;
    for (i=0; i < para_length; i++) {
        bit(8) camera_parameter;
20     }
    }
}

```

①第i个轨道对应的视角信息可以包括视频拼接布局指示字段stitching_layout，该视频拼接布局字段主要用于指示第i个轨道内包括的纹理图像和深度图像是否拼接编码；具体地，
 25 当视频拼接布局指示字段stitching_layout为第一数据值时，指示第i个轨道内包括的纹理图像和深度图像是拼接编码的；当视频拼接布局指示字段stitching_layout为第二数值时，指示第i个轨道内包括的纹理图像和深度图像是分别编码的。举例来说，假设沉浸媒体是指6DoF视频，第一数值为0，第二数值为1，则在第i个轨道内视频拼接布局指示字段stitching_layout可以参见下述表2所示：

30

表2

stitching_layout	6DoF视频拼接布局
0	深度图和纹理图拼接编码
1	深度图和纹理图分别编码
其他	保留

②第i个轨道对应的视角信息还包括相机模型字段camera_model，该字段用于指示第i个轨道内k个相机的相机模型，由前述可知，k表示第i个轨道内深度图像和/或纹理图像来源的相机数量，在第i个轨道对应的视角信息中还可以包括相机数量字段camera_count，该相

- 12 -

机数量字段用于存储第i个轨道内深度图像和/或纹理图像来源的相机数量，假设表示为k。

具体地，当相机模型字段camera_model为第三数值时，指示第j个相机所属相机模型为第一模型；当相机模型字段camera_model为第四数值时，指示第j个相机所属相机模型为第二模型。其中，第一模型可以指针孔模型，第二模型可以指鱼眼模型，假设第三数值为0，第四数值为1，沉浸媒体为6DoF视频，第i个轨道对应的视角信息中相机模型字段camera_model可参见下述表3所示：

表3

Camera model	6DoF视频相机模型
0	针孔模型
1	鱼眼模型
其他	保留

③第i个轨道对应的视角信息还包括纹理图像的保护带宽度字段texture_padding_size，以及深度图像的保护带宽度字段depth_padding_size。纹理图像的保护带宽度字段用于存储第i个轨道内对纹理图像编码时采用的保护带宽度，深度图像的保护带宽度用于存储第i个轨道内对深度图像编码时采用的保护带宽度。

④第i个轨道对应的视角信息还包括第二相机标识字段camera_id，第二相机标识字段camera_id用于存储第i个轨道内第j个相机的标识信息，由前述描述可知，第i个轨道内包括的纹理图像和深度图像来源这k个相机，j的取值是大于等于0，且小于k。也就是说一个第二相机标识字段camera_id存储k个相机中任意一个相机的标识信息，因此，需要k个相机标识字段来存储k个相机的标识信息。需要说明的是，这里的第二相机标识字段camera_id和前述的自由视角轨道组信息中的第一相机标识字段camera_id的作用相同，均是用于存储第i个轨道内第j个相机的标识信息。

⑤第i个轨道对应的视角信息还包括相机属性信息字段，相机属性信息字段用于存储第j个相机的相机属性信息，第j个相机的相机属性信息可以包括第j个相机位置的横轴分量取值、纵轴分量取值以及竖轴分量取值，第j个相机焦距的横轴分量取值和纵轴分量取值，以及第j个相机采集图像的分辨率宽度与高度。因此，相机属性信息字段可具体包括：1）相机位置的横轴分量字段camera_pos_x，用于存储第j个相机位置的横轴分量取值（也称为x分量取值）；2）相机位置的纵轴分量字段camera_pos_y，用于存储第j个相机位置的纵轴分量取值（也称为y分量取值）；3）相机位置的竖轴分量字段camera_pos_z，用于存储第j个相机位置的竖轴分量取值（也称为z分量取值）；4）相机焦距的横轴分量字段focal_length_x，用于存储第j个相机焦距的横轴分量取值（也称为x分量取值）；5）相机焦距的纵轴分量字段focal_length_y，用于存储第j个相机焦距的纵轴分量取值（也称为y分量取值）；6）相机采集图像的分辨率宽度字段camera_resolution_x，用于存储第j个相机采集图像的分辨率宽度；7）相机采集图像的分辨率宽度字段camera_resolution_y，用于存储第j个相机采集图像的分辨率高度。

需要说明的是，由于第i个轨道内的纹理图像和/深度图像来源于k个相机，一个相机属性信息字段用于存储一个相机的相机属性信息，因此，第i个轨道对应的视角信息中包括k

- 13 -

个相机属性信息字段用于存储k个相机的相机属性信息。相应的，一个相机属性字段包括上述1)-7)，k个相机属性信息便包括k个上述1)-7)。

⑥第i个轨道对应的视角信息还包括图像信息字段，该图像信息字段用于存储第j个相机采集图像的图像信息。图像信息可以包括以下至少一种：深度图像降采样的倍数因子、深度图像左上顶点相对于平面帧原点的偏移量或纹理图像左上顶点相对于平面帧原点的偏移量。基于此，图像信息字段可以具体包括：1)深度图像降采样的倍数因子字段depth_downsample_factor，用于存储深度图像降采样的倍数因子；2)纹理图像左上顶点横轴偏移字段texture_vetex_x，用于存储纹理图像左上顶点相对于平面帧原点的偏移量横轴分量；3)纹理图像左上顶点纵轴偏移字段texture_vetex_y，用于存储纹理图像左上顶点相对于平面帧原点的偏移量纵轴分量；4)深度图像左上顶点横轴偏移字段depth_vetex_x，用于存储深度图像左上顶点相对于平面帧原点的横轴偏移量；5)深度图左上顶点纵轴偏移字段depth_vetex_y，用于存储深度图像左上顶点相对于平面帧原点的纵轴偏移量。

需要说明的是，一个图像信息字段用于存储一个相机采集图像的图像信息，第i个轨道内包括k个相机采集的深度图像和/或纹理图像，因此第i个轨道对应的自由视角信息中包括k个图像信息字段。

⑦第i个轨道对应的视角信息还包括自定义相机参数字段camera_parameter，该自定义相机参数字段用于存储第j个相机的第f个自定义相机参数，f为大于等于0且小于h的整数，h表示第j个相机的自定义相机参数的数量，第j个相机的自定义相机参数的数量可以存储在视角信息的自定义相机参数数量字段para_num中。

需要说明的是，由于第j个相机的自定义相机参数的数量为h个，一个自定义相机参数字段用于存储一个自定义相机参数，因此，第j个相机对应的自定义相机参数字段的数量可以为h个。又因为，每个相机对应h个自定义相机参数，第i个轨道内包括k个相机，因此，第i个轨道对应的视角信息中包括k*h个自定义相机参数字段。

⑧第i个轨道对应的视角信息还包括自定义相机参数类型字段para_type，自定义相机参数类型字段para_type用于存储第j个相机的第f个自定义相机参数所属类型。需要说明的是，第j个相机的一个自定义相机参数所属类型存储在一个自定义相机参数类型字段中，由于第j个相机对应h个自定义相机参数，因此自定义相机参数类型字段的数量为h。与⑦同理的，第i个轨道对应的视角信息中包括k*h个自定义相机参数类型字段。

⑨第i个轨道对应的视角信息还包括自定义相机参数长度字段para_length，该自定义相机参数长度字段para_length用于存储第j个相机的第f个自定义相机参数的长度。由于第j个相机的一个自定义相机参数的长度存储在一个自定义相机参数长度字段中，因为第j个相机包括h个自定义相机参数，因此自定义相机参数长度字段的数量为h个。

本申请实施例中，S401中获取第i个轨道对应的自由视角信息数据盒是依据内容制作设备发送的信令描述文件实现的。具体地，在执行S401之前，获取沉浸媒体对应的信令描述文件，该信令描述文件包括沉浸媒体对应的自由视角相机描述子，自由视角相机描述子用于记录每个轨道内视频片段对应的相机属性信息，任一轨道内的视频片段是由任一轨道内的纹理图像和/或深度图像组成的；该与自由视角相机描述子被封装于所述媒体数据的媒体

呈现描述文件的自适应集层级中, 或者所述信令描述文件被封装于所述媒体呈现描述文件的表示层级中。

- 5 本申请实施例中, 自由视角相机描述子可以表示为 AvsFreeViewCamInfo, 其为 SupplementalProperty 元素, 其 @schemeIdUri 属性为 "urn:avs:ims:2018:av31"。当自由视角相机描述子位于表示层级中时, 可用于描述该表示 representation 层级对应的轨道内视频片段对应的相机属性信息; 当自由视角相机描述子位于自适应层级 adaptation set 时, 可以用于描述自适应层级中多个轨道内视频片段对应的相机属性信息。

其中, 自由视角相机描述子中各个元素以及属性可如下表4所示:

表4

Elements and Attributes for AvsFreeViewCamInfo descriptor	作用	数据类型	解释
AvsFreeViewCam	0..N	avs:ims:2018:av31	若自由视角相机描述子存在于 representation 层级中, 该元素用于指示 representation 层级对应的轨道图像对应的相机属性; 若自由视角相机描述子存在于 adaptation set 层级中, 该元素指示 adaptation set 层级中多个轨道中图像对应的相机的属性信息。
AvsFreeViewCam @cameraId	M	xs:unsignedInt	指示每个相机对应的标识信息, 该标识信息与轨道的 AvsFreeViewInfoBox 中的标识信息对应。
AvsFreeViewCam @ camera_pos_x	M	xs:float	指示相机位置的x分量取值。
AvsFreeViewCam @ camera_pos_y	M	xs:float	指示相机位置的y分量取值。
AvsFreeViewCam @ camera_pos_z	M	xs:float	指示相机位置的z分量取值。
AvsFreeViewCam @ focal_length_x	M	xs:float	指示相机焦距的x分量取值。
AvsFreeViewCam @ focal_length_y	M	xs:float	指示相机焦距的y分量取值。

- 10 进一步的, 内容消费设备基于信令描述文件获取第i个轨道对应的自由视角信息数据盒。在一个实施例中, 获取第i个轨道对应的自由视角信息数据盒, 包括: 基于自由视角相机描

述子中记录的每个轨道内图像对应的相机属性信息，从N个相机中选择与用户所在位置信息匹配的候选相机；向内容制作设备发送获取候选相机拍摄的图像的第一资源请求，第一资源请求用于指示内容制作设备根据M个轨道中每个轨道对应的自由视角信息数据盒中视角信息，从M个轨道中选择第i个轨道并返回第i个轨道对应的自由视角信息数据盒，第i个轨道中封装候选相机拍摄的图像；接收内容制作设备返回的第i个轨道对应的自由视角信息数据盒。在这种方式下，内容消费设备只需要获取所需的轨道对应的自由视角信息数据盒，不必获取所有轨道对应的自由视角信息数据盒，可节省传输资源。下面举例说明：

(1) 假设内容制作设备生成自由视角视频并将自由视角视频封装为多个轨道，每个轨道中可以包括一个视角的纹理图像和深度图像，也就是说一个轨道内封装了来自一个相机的纹理图像和深度图像，一个轨道内的纹理图像和深度图像组成了一个视频片段，这样一来，可以理解为一个轨道内封装了来自处于一个视角的相机的视频片段。假设自由视角视频被封装到3个轨道，每个轨道内视频的纹理图像和深度图像组成一个视频片段，因此自由视角视频包括3个视频片段，分别表示为Representation1、Representation2以及Representation3；

(2) 内容制作设备在信令生成环节，根据每个轨道对应的自由视角信息数据盒中的视角信息生成该自由视角视频对应的信令描述文件，信令描述文件中可携带自由视角相机描述子。假设自由视角视频自由视角相机描述子记录了3个视频片段对应的相机属性信息如下：

Representation1: {Camera1: ID=1;Pos=(100,0,100);Focal=(10,20)};

Representation2: {Camera2: ID=2;Pos=(100,100,100);Focal=(10,20)};

Representation1: {Camera3: ID=3;Pos=(0,0,100);Focal=(10,20)};

(3) 内容制作设备将信令描述文件发送至内容消费设备；

(4) 内容消费设备根据信令描述文件和用户带宽，依据用户所在位置信息和信令描述文件中的相机属性信息，选取来自Camera2和Camera3的视频片段并向内容制作设备请求；

(5) 内容制作设备将来自Camera2和Camera3的视频片段的轨道的自由视角信息数据盒发送至内容消费设备，内容消费设备根据获取到的轨道对应的自由视角信息数据盒初始化解码器，解码对应的视频片段并消费。

另一个实施例中，获取第i个轨道对应的自由视角信息数据盒，包括：基于信令描述文件和用户带宽向内容制作设备发送第二资源请求，第二资源请求用于指示内容制作设备返回M个轨道的M个自由视角信息数据盒，一个轨道对应一个自由视角信息数据盒；根据每个轨道对应的自由视角信息数据盒中的视角信息和用户所在位置信息，从M个自由视角信息数据盒中获取第i个轨道对应的自由视角信息数据盒。在这种方式下，内容消费设备虽然获取到所有轨道的自由视角信息数据盒，但是不对所有轨道中图像进行解码消费，只解码与用户当前位置匹配的第i个轨道中图像，节省了解码资源。下面举例来说：

(1) 假设内容制作设备生成自由视角视频并将自由视角视频封装为多个轨道，每个轨道中可以包括一个视角的纹理图像和深度图像，也就是说每个轨道封装来自一个视角的相机的视频片段，假设自由视角视频被封装到3个轨道的视频片段分别表示为Representation1、Representation2以及Representation3；

(2) 内容制作设备根据每个轨道对应的视角信息，生成该自由视角视频的信令描述文

件，如下所示：

Representation1: {Camera1: ID=1;Pos= (100,0,100) ;Focal= (10,20) };

Representation2: {Camera2: ID=2;Pos= (100,100,100) ;Focal= (10,20) };

Representation1: {Camera3: ID=3;Pos= (0,0,100) ;Focal= (10,20) };

5 (3) 内容制作设备将(2)中的信令描述文件发送至内容消费设备;

(4) 内容消费设备根据信令描述文件和用户带宽向内容制作设备请求所有轨道对应的自由视角信息数据盒;假设自由视角视频被封装到3个轨道中,分别为Track1、Track2、以及Track3每个轨道中封装的视频片段对应的相机属性信息可以如下:

Track1: {Camera1: ID=1; Pos= (100,0,100) ;Focal=(10,20)};

10 Track2: {Camera2: ID=2; Pos= (100,100,100) ;Focal=(10,20)};

Track3: {Camera3: ID=3; Pos= (0,0,100) ;Focal=(10,20)};

(5) 内容消费设备根据获取到的所有轨道对应的自由视角信息数据盒中的视角信息和用户当前观看的位置信息,选择track2和track3的轨道中封装的视频片段进行解码消费。

S402、根据自由视角信息数据盒中的视角信息对第i个轨道内封装的图像进行解码。

15 具体实现中,根据自由视角信息数据盒中的视角信息对第i个轨道内封装的图像进行解码显示,可以包括:根据第i个轨道对应的视角信息初始化解码器;再根据视角信息中与编码相关的指示信息对第i个轨道内封装的图像进行解码处理。

本申请实施例中将一沉浸媒体封装至M个轨道,该沉浸媒体是由处于不同视角的N个相机拍摄的图像组成的,一个轨道中可以包括来自至少一个相机的图像,这M个轨道是属于
20 同一个轨道组的,实现了一个沉浸视频封装到多个轨道的场景;另外,内容制作设备为每个轨道生成一个自由视角信息数据盒,通过第i个轨道对应的自由视角信息数据盒指示第i个轨道对应的视角信息,比如相机所在的具体视角位置,进而内容消费设备在根据第i个轨道对应的视角信息进行该轨道内图像的解码显示时,可以保证解码显示的视频与用户所在位置更加匹配,提高沉浸视频的呈现效果。

25 基于上述的沉浸媒体的数据处理方法,本申请实施例提供了另一种沉浸媒体的数据处理方法。参见图5,为本申请实施例提供的一种沉浸媒体的数据处理方法的流程示意图,图5所示的沉浸媒体的数据处理方法可由内容制作设备执行,具体可由内容制作设备的处理器执行。图5所示的沉浸媒体的数据处理方法可包括如下步骤:

30 S501、将沉浸媒体封装到M个轨道中,该沉浸媒体是由处于不同视角的相机拍摄的图像组成的,一个轨道中封装了来自至少一个相机的图像,这M个轨道属于同一个轨道组。

S502、根据第i个轨道中w个图像的封装过程生成第i个轨道对应的自由视角信息数据盒。

其中,w为大于或等于1的整数,第i个轨道内封装了w个图像,这w个图像可以是来自处于某一个或某几个视角的相机拍摄的深度图像和纹理图像中的至少一种。结合图4实施例中的代码段3,具体介绍根据w个图像在第i个轨道的封装过程生成第i个轨道对应的自由视角
35 信息数据盒,具体可以包括:

①第i个轨道对应的自由视角信息数据盒中包括自由视角轨道组信息,该自由视角轨道组信息可以指示第i个轨道与其他封装了沉浸媒体的轨道属于同一个轨道组。该自由视角轨

道组信息包括第一相机标识字段camera_id和图像类型字段depth_texture_type, S502中根据第i个轨道内w个图像的封装过程生成所述第i个轨道对应的自由视角信息数据盒, 包括: 确定w个图像来源于k个相机, k为大于1的整数; 将k个相机中的第j个相机的标识信息存储在第一相机标识字段camera_id中, j大于0且小于k; 根据所述第j个相机采集的图像所属图像类型确定所述图像类型字段depth_texture_type, 所述图像类型包括深度图像和纹理图像中任意一个或多个。应当理解的, 一个相机的标识信息存储在一个第一相机标识字段camera_id, 如果第i个轨道内包含的图像来源于k个相机, 那么自由视角轨道组信息中可以包括k个第一相机标识字段camera_id。同理的, 自由视角轨道组包括k个图像类型字段depth_texture_type。

具体地, 根据第j个相机采集的图像所属图像类型确定图像类型字段, 包括: 当第j个相机采集的图像所属图像类型为纹理图像, 则将图像类型字段设置为第二数值; 当第j个相机采集的图像所属图像类型为深度图像, 则将图像类型字段设置为第三数值; 当第j个相机采集的图像所属图像类型为纹理图像和深度图像, 则将图像类型字段设置为第四数值。其中, 第二数值可以为1, 第三数值可以为2, 第四数值可以为3, 图像类型字段可参见图4实施例中表2所示。

②第i个轨道对应的视角信息包括视频拼接布局指示字段stitching_layout, S502中根据第i个轨道内w个图像的封装过程生成第i个轨道对应的自由视角信息数据盒, 包括: 如果第i个轨道内包含的纹理图像和深度图像是拼接编码的, 则将视频布局指示字段stitching_layout设置为第一数值; 如果第i个轨道内包含的纹理图像和深度图像是分别编码的, 则将视频布局指示字段stitching_layout设置为第二数值。其中, 第一数值可以为0, 第二数值可以为1, 视频拼接布局指示字段可以参见图4实施例中表3所示。

③第i个轨道对应的视角信息还包括相机模型字段camera_model; S502中根据第i个轨道内w个图像的封装过程生成第i个轨道对应的自由视角信息数据盒, 包括: 当采集w个图像的相机所属的相机模型为第一模型时, 将相机模型字段camera_model设置为第一数值; 当采集w个图像的相机所属的相机模型为第二模型时, 将相机模型字段camera_model设置为第二数值。同上述, 第一数值可以为0, 第二数值可以为1, 相机模型字段可以参见图4实施例中表4所示。

④第i个轨道对应的视角信息还包括纹理图像的保护带宽字段texture_padding_size和深度图像的保护带宽字段depth_padding_size。S502中根据第i个轨道内w个图像的封装过程生成第i个轨道对应的自由视角信息数据盒, 包括: 获取第i个轨道内对纹理图像编码时采用的纹理图像的保护带宽度, 并将纹理图像的保护带宽度存储在纹理图像的保护带宽度字段texture_padding_size; 获取第i个轨道内对深度图像编码时采用的深度图像的保护带宽度, 并将深度图像的保护带宽度存储在深度图像的保护带宽度字段depth_padding_size。

⑤第i个轨道对应的视角信息还包括第二相机标识字段camera_id和相机属性信息字段, S502中根据第i个轨道内w个图像的封装过程生成第i个轨道对应的自由视角信息数据盒, 包括: 将第j个相机的标识信息存储在第二相机标识字段camera_id, j大于或等于0且小于k, k表示w个图像来源相机的数量; 获取第j个相机的相机属性信息, 并将获取到的相机属性信息存储在相机属性信息字段。其中, 所述第j个相机的相机属性信息包括以下任意一种或多

种：第j个相机位置的横轴分量取值、纵轴分量取值以及竖轴分量取值，第j个相机焦距的横轴分量取值和纵轴分量取值，以及第j个相机采集图像的分辨率宽度与高度。

基于此，相机属性信息字段可具体包括：1) 相机位置的横轴分量字段camera_pos_x，用于存储第j个相机位置的横轴分量取值（也称为x分量取值）；2) 相机位置的纵轴分量字段camera_pos_y，用于存储第j个相机位置的纵轴分量取值（也称为y分量取值）；3) 相机位置的竖轴分量字段camera_pos_z，用于存储第j个相机位置的竖轴分量取值（也称为z分量取值）；4) 相机焦距的横轴分量字段focal_length_x，用于存储第j个相机焦距的横轴分量取值（也称为x分量取值）；5) 相机焦距的纵轴分量字段focal_length_y，用于存储第j个相机焦距的纵轴分量取值（也称为y分量取值）；6) 相机采集图像的分辨率宽度字段camera_resolution_x，用于存储第j个相机采集图像的分辨率宽度；7) 相机采集图像的分辨率高度字段camera_resolution_y，用于存储第j个相机采集图像的分辨率高度。

需要说明的是，由于第i个轨道内的纹理图像和/深度图像来源于k个相机，一个相机属性信息字段用于存储一个相机的相机属性信息，因此，第i个轨道对应的视角信息中包括k个相机属性信息字段用于存储k个相机的相机属性信息。相应的，一个相机属性字段包括上述1) -7)，k个相机属性信息便包括k个上述1) -7)。

⑥第i个轨道对应的视角信息还包括图像信息字段，S502中根据w个图像的封装过程生成第i个轨道对应的自由视角信息数据盒，包括：获取所述第j个相机采集图像的图像信息，并将获取到的图像信息存储在图像信息字段；其中，所述图像信息包括如下一个或多个：深度图像降采样的倍数因子、深度图像左上顶点相对于平面帧原点的偏移量、纹理图像左上顶点相对于平面帧原点的偏移量。

基于此，图像信息字段可以具体包括：1) 深度图像降采样的倍数因子字段depth_downsample_factor，用于存储深度图像降采样的倍数因子；2) 纹理图像左上顶点横轴偏移字段texture_vetex_x，用于存储纹理图像左上顶点相对于平面帧原点的偏移量横轴分量；3) 纹理图像左上顶点纵轴偏移字段texture_vetex_y，用于存储纹理图像左上顶点相对于平面帧原点的偏移量纵轴分量；4) 深度图像左上顶点横轴偏移字段depth_vetex_x，用于存储深度图像左上顶点相对于平面帧原点的横轴偏移量；5) 深度图左上顶点纵轴偏移字段depth_vetex_y，用于存储深度图像左上顶点相对于平面帧原点的纵轴偏移量。

需要说明的是，一个图像信息字段用于存储一个相机采集图像的图像信息，第i个轨道内包括k个相机采集的深度图像和/或纹理图像，因此第i个轨道对应的自由视角信息中包括k个图像信息字段。

⑦第i个轨道对应的视角信息还包括自定义相机参数字段、自定义相机参数类型字段以及自定义相机参数长度字段；根据第i个轨道内w个图像的封装过程生成第i个轨道对应的自由视角信息数据盒，包括：获取第f个自定义相机参数，并将第f个自定义相机参数存储在自定义相机参数字段，f为大于等于0且小于h的整数，h表示第j个相机的自定义相机参数的数量；确定第f个自定义相机参数所属的参数类型以及第f个自定义相机参数的长度，将第f个自定义相机参数所属的参数类型存储在自定义相机参数类型字段，以及将第f个自定义相机参数的长度存储在自定义相机参数长度字段。

需要说明的是，由于第j个相机的自定义相机参数的数量为h个，一个自定义相机参数
字段用于存储一个自定义相机参数，因此，第j个相机对应的自定义相机参数字段的数量可以
为h个。又因为，每个相机对应h个自定义相机参数，第i个轨道内包括k个相机，因此，第i
个轨道对应的视角信息中包括k*h个自定义相机参数字段。需要说明的是，由于第j个相机
5 的自定义相机参数的数量为h个，一个自定义相机参数字段用于存储一个自定义相机参数，
因此，第j个相机对应的自定义相机参数字段的数量可以为h个。又因为，每个相机对应h个自
定义相机参数，第i个轨道内包括k个相机，因此，第i个轨道对应的视角信息中包括k*h个自
定义相机参数字段。由于第j个相机的一个自定义相机参数的长度存储在一个自定义相机参
数长度字段中，因为第j个相机包括h个自定义相机参数，因此自定义相机参数长度字段的
10 数量为h个。

另外，内容制作设备还可以根据M个轨道对应的自由视角信息数据盒中的视角信息生
成信令描述文件，信令描述文件包括沉浸媒体对应的自由视角相机描述子，自由视角相机
描述子用于记录每个轨道内图像对应的相机属性信息；自由视角相机描述子被封装于沉浸
媒体的媒体呈现描述文件的自适应集层级中，或者信令描述文件被封装于媒体呈现描述文
15 件的表示层级中。

进一步的，内容制作设备可以将信令描述文件发送至内容消费设备，以使内容消费设
备根据信令描述文件获取第i个轨道对应的自由视角信息数据盒。

作为一种可选的实施方式，内容制作设备将信令描述文件发送至内容消费设备，以指
示内容消费设备基于自由视角相机描述子中记录的每个轨道内图像对应的相机属性信息，
20 从N个相机中选择与用户所在位置匹配的候选相机，以及发送获取来源于候选相机的图像
的第一资源请求；响应于第一资源请求，根据M个轨道中每个轨道对应的自由视角信息数
据盒中的视角信息从M个轨道中选择第i个轨道并将i个轨道对应的自由视角信息数据盒发
送至内容消费设备，第i个轨道内包括来源于候选相机的图像。

作为另一种可选的实施方式，内容制作设备将信令描述文件发送至内容消费设备，以
25 指示内容消费设备根据信令描述文件和用户带宽发送第二资源请求；响应于第二资源请求，
将M个轨道对应的M个自由视角信息数据盒发送至内容消费设备，以指示内容消费设备根
据每个轨道对应的自由视角信息数据盒中的视角信息和用户所在位置信息，从M个自由视
角信息数据盒中获取第i个轨道对应的自由视角信息数据盒。

本申请实施例中将一个沉浸媒体封装至M个轨道，该沉浸视频是由处于不同视角的相
30 机采集的图像组成的，这M个轨道是属于同一个轨道组的，一个轨道中封装来自至少一个
相机的图像，实现了一个沉浸视频封装到多个轨道的场景；另外，内容制作设备为根据每
个轨道中图像的封装过程为每个轨道生成一个自由视角信息数据盒，通过第i个轨道对应
的自由视角信息数据盒指示第i个轨道对应的视角信息，比如相机所在的具体视角位置，进
而内容消费设备在根据第i个轨道对应的视角信息进行该轨道内图像的解码显示时，可以保
35 证解码显示的视频与用户所在位置更加匹配，提高沉浸视频的呈现效果。

基于上述的沉浸媒体的数据处理方法实施例，本申请实施例提供了一种沉浸媒体的数
据处理装置，该沉浸媒体的数据处理装置可以是运行于内容消费设备中的一个计算机程序

(包括程序代码),例如该沉浸媒体的数据处理装置可以是内容消费设备中的一个应用软件。参见图6,为本申请实施例提供的一种沉浸媒体的数据处理装置的结构示意图。图6所示的数据处理装置可运行如下单元:

获取单元601,用于获取所述沉浸媒体的第*i*个轨道对应的自由视角信息数据盒,所述自由视角信息数据盒包括所述第*i*个轨道对应的视角信息,*i*为大于或等于1且小于或等于*M*的整数;该沉浸媒体是由处于不同视角的*N*个相机拍摄的图像组成的,所述沉浸媒体被封装到*M*个轨道中,一个轨道中封装来自至少一个相机的图像,所述*M*个轨道属于同一个轨道组,*N*和*M*均为大于1的整数;

处理单元602,用于根据所述自由视角信息数据盒中的视角信息对所述第*i*个轨道内封装的图像进行解码。

在一个实施例中,所述第*i*个轨道内封装了来自所述*N*个相机中*k*个相机的图像,*k*为大于0的整数;所述第*i*个轨道内还封装了所述第*i*个轨道对应的自由视角轨道组信息,所述自由视角轨道组信息用于指示所述第*i*个轨道与封装所述沉浸媒体的其他轨道属于同一个轨道组;所述自由视角轨道组信息包括第一相机标识字段和图像类型字段;

所述第一相机标识字段用于存储所述*k*个相机中第*j*个相机的标识信息,所述图像类型字段用于指示所述第*j*个相机采集的图像所属的图像类型,所述图像类型包括纹理图像或深度图像中的至少一个。

在一个实施例中,若所述第*i*个轨道内包括纹理图像和深度图像,所述第*i*个轨道对应的视角信息包括视频拼接布局指示字段;当所述视频拼接布局指示字段为第一数值时,指示所述第*i*个轨道内包括的纹理图像和深度图像是拼接编码的;当所述视频拼接布局指示字段为第二数值时,指示所述第*i*个轨道内包括的纹理图像和深度图像是分别编码的。

在一个实施例中,所述第*i*个轨道对应的视角信息还包括相机模型字段;当所述相机模型字段为第三数值时,指示所述*j*个相机所属相机模型为第一模型;当所述相机模型字段为第四数值时,指示所述*j*个相机所属相机模型为第二模型。

在一个实施例中,所述第*i*个轨道对应的视角信息还包括纹理图像的保护带宽度字段和深度图像的保护带宽度字段,所述纹理图像的保护带宽度字段用于存储所述第*i*个轨道内纹理图像进行编码时采用的保护带宽度,所述深度图像的保护带宽度字段用于存储所述第*i*个轨道内深度图像进行编码时采用的保护带宽度。

在一个实施例中,所述第*i*个轨道对应的视角信息还包括第二相机标识字段和相机属性信息字段;所述第二相机标识字段用于存储第*j*个相机的标识信息;

所述相机属性信息字段用于存储所述第*j*个相机的相机属性信息,所述第*j*个相机的相机属性信息包括以下至少一种:所述第*j*个相机位置的横轴分量取值、纵轴分量取值以及竖轴分量取值,所述第*j*个相机焦距的横轴分量取值和纵轴分量取值,以及第*j*个相机采集图像的分辨率宽度与高度。

在一个实施例中,所述第*i*个轨道对应的视角信息还包括图像信息字段,所述图像信息字段用于存储所述第*j*个相机采集图像的图像信息,所述图像信息包括以下任意一个:深度图像降采样的倍数因子、深度图像左上顶点相对于平面帧原点的偏移量或纹理图像左上顶

点相对于平面帧原点的偏移量。

在一个实施例中，所述第*i*个轨道对应的视角信息还包括自定义相机参数字段、自定义相机参数类型字段以及自定义相机参数长度字段；所述自定义相机参数字段用于存储所述第*j*个相机的第*f*个自定义相机参数，*f*为大于等于0且小于*h*的整数，*h*表示所述第*j*个相机的自定义相机参数的数量；所述自定义相机参数类型字段用于存储所述第*f*个自定义相机参数所属的参数类型；所述自定义相机参数长度字段用于存储所述第*f*个自定义相机参数的长度。

在一个实施例中，所述获取单元601还用于：

获取所述沉浸媒体对应的信令描述文件，所述信令描述文件包括所述沉浸媒体对应的自由视角相机描述子，所述自由视角相机描述子用于记录每个轨道内视频片段对应的相机属性信息，所述轨道内视频片段是由所述轨道内包括的纹理图像和深度图像组成的；所述自由视角相机描述子被封装于所述沉浸媒体的媒体呈现描述文件的自适应集层级中，或者所述信令描述文件被封装于所述媒体呈现描述文件的表示层级中。

在一个实施例中，所述获取单元601在获取所述沉浸媒体的第*i*个轨道对应的自由视角信息数据盒时，执行如下操作：

基于所述自由视角相机描述子中记录的每个轨道内视频片段对应的相机属性信息，从所述*N*个相机中选择与用户所在位置信息匹配的候选相机；

向内容制作设备发送获取来源于所述候选相机的分块视频的第一资源请求，所述第一资源请求用于指示所述内容制作设备根据所述*M*个轨道中每个轨道对应的自由视角信息数据盒中的视角信息，从所述*M*个轨道中选择第*i*个轨道并返回所述第*i*个轨道对应的自由视角信息数据盒，所述第*i*个轨道中封装的纹理图像和深度图像中的至少一种来自所述候选相机；接收所述内容制作设备返回的所述第*i*个轨道对应的自由视角信息数据盒。

在一个实施例中，所述获取单元601在获取所述沉浸媒体的第*i*个轨道对应的自由视角信息数据盒时，执行如下操作：

基于所述信令描述文件和用户带宽向内容制作设备发送第二资源请求，所述第二资源请求用于指示所述内容制作设备返回所述*M*个轨道的*M*个自由视角信息数据盒，一个轨道对应一个自由视角信息数据盒；

根据每个轨道对应的自由视角信息数据盒中的视角信息和用户所在位置信息，从所述*M*个自由视角信息数据盒中获取第*i*个轨道对应的自由视角信息数据盒。

根据本申请的一个实施例，图4所示的沉浸媒体的数据处理方法所涉及各个步骤可以由图6所示的沉浸媒体的数据处理装置中的各个单元来执行的。例如，图4所述的S401可由图6中所述的数据处理装置中的获取单元601来执行，S402可由图6所示的数据处理装置中的处理单元602来执行。

根据本申请的另一个实施例，图6所示的沉浸媒体的数据处理装置中的各个单元可以分别或全部合并为一个或若干个另外的单元来构成，或者其中的某个（些）单元还可以再拆分为功能上更小的多个单元来构成，这可以实现同样的操作，而不影响本申请的实施例的技术效果的实现。上述单元是基于逻辑功能划分的，在实际应用中，一个单元的功能也可以由多个单元来实现，或者多个单元的功能由一个单元实现。在本申请的其它实施例中，

基于沉浸媒体的数据处理装置也可以包括其它单元，在实际应用中，这些功能也可以由其它单元协助实现，并且可以由多个单元协作实现。

根据本申请的另一个实施例，可以通过在包括中央处理单元（CPU）、随机存取存储介质（RAM）、只读存储介质（ROM）等处理元件和存储元件的例如计算机的通用计算设备上运行能够执行如图4的相应方法所涉及的各步骤的计算机程序（包括程序代码），来构造
5 如图6中所示的沉浸媒体的数据处理装置，以及来实现本申请实施例沉浸媒体的数据处理方法。所述计算机程序可以记载于例如计算机可读存储介质上，并通过计算机可读存储介质装载于上述计算设备中，并在其中运行。

本申请实施例中将一沉浸媒体封装至M个轨道，该沉浸媒体是由处于不同视角的N个相机拍摄的图像组成的，一个轨道中可以包括来自至少一个相机的图像，这M个轨道是属于
10 同一个轨道组的，实现了一个沉浸视频封装到多个轨道的场景；另外，内容制作设备为每个轨道生成一个自由视角信息数据盒，通过第i个轨道对应的自由视角信息数据盒指示第i个轨道对应的视角信息，比如相机所在的具体视角位置，进而内容消费设备在根据第i个轨道对应的视角信息进行该轨道内图像的解码显示时，可以保证解码显示的视频与用户所在位置更加匹配，提高沉浸视频的呈现效果。
15

基于上述的沉浸媒体的数据处理方法以及数据处理装置实施例，本申请实施例提供了另一种沉浸媒体的数据处理装置，该沉浸媒体的数据处理装置可以是运行于内容制作设备中的一个计算机程序（包括程序代码），例如该沉浸媒体的数据处理装置可以是内容制作设备中的一个应用软件。参见图7，为本申请实施例提供的另一种沉浸媒体的数据处理装置的结构示意图。图7所示的数据处理装置可运行如下单元：
20

封装单元701，用于将沉浸媒体封装到M个轨道中，所述沉浸媒体是由处于不同视角的N个相机拍摄的图像组成的，一个轨道中封装来自至少一个相机的图像，所述M个轨道属于同一个轨道组，N和M均为大于或等于1的整数；

生成单元702，用于根据第i个轨道中w个图像的封装过程，生成第i个轨道对应的自由
25 视角信息数据盒，所述自由视角信息数据盒包括所述第i个轨道对应的视角信息； $1 \leq i \leq M$ ， $w \geq 1$ 。

在一个实施例中，所述第i个轨道对应的自由视角信息数据盒包括自由视角轨道组信息，所述自由视角轨道组信息包括第一相机标识字段和图像类型字段，所述封装单元701在根据第i个轨道内w个图像的封装过程生成所述第i个轨道对应的自由视角信息数据盒时，执行如下步骤：
30

确定所述w个图像来源于k个相机，k为大于1的整数；将所述k个相机中的第j个相机的标识信息存储在所述第一相机标识字段中；根据所述第j个相机采集的图像所属图像类型确定所述图像类型字段，所述图像类型包括深度图像或纹理图像中的至少一个。

在一个实施例中，所述封装单元701在根据所述第j个相机采集的图像所属图像类型确定所述图像类型字段时，执行如下步骤：
35

当所述第j个相机采集的图像所属图像类型为纹理图像，则将所述图像类型字段设置为第二数值；当所述第j个相机采集的图像所属图像类型为深度图像，则将所述图像类型字段

设置为第三数值；当所述第j个相机采集的图像所属图像类型为纹理图像和深度图像，则将所述图像类型字段设置为第四数值。

在一个实施例中，所述第i个轨道对应的视角信息包括视频拼接布局指示字段，所述w个图像包括纹理图像和深度图像，所述封装单元 701在根据第i个轨道内w个图像的封装过程生成第i个轨道对应的自由视角信息数据盒时，执行如下步骤：

如果所述第i个轨道内包含的纹理图像和深度图像是拼接编码的，则将所述视频布局指示字段设置为第一数值；如果所述i个轨道内包含的的纹理图像和深度图像是分别编码的，则将所述视频布局指示字段设置为第二数值。

在一个实施例中，所述第i个轨道对应的视角信息还包括相机模型字段；所述封装单元 701在根据第i个轨道内的w个图像的封装过程生成第i个轨道对应的自由视角信息数据盒时，执行如下步骤：

当采集所述w个图像的相机所属的相机模型为第一模型时，将所述相机模型字段设置为第一数值；当采集所述w个图像的相机所属的相机模型为第二模型时，将所述相机模型字段设置为第二数值。

在一个实施例中，所述第i个轨道对应的视角信息还包括纹理图像的保护带宽字段和深度图像的保护带宽字段，所述封装单元701在根据第i个轨道内的w个图像的封装过程生成第i个轨道对应的自由视角信息数据盒时，执行如下步骤：

获取所述第i个轨道内在对纹理图像进行编码时采用的纹理图像的保护带宽度，并将所述纹理图像的保护带宽度存储在所述纹理图像的保护带宽度字段；

获取所述第i个轨道内在对深度图像进行编码时采用的深度图像的保护带宽度，并将所述深度图像的保护带宽度存储在所述深度图像的保护带宽度字段。

在一个实施例中，所述第i个轨道对应的视角信息还包括第二相机标识字段和相机属性信息字段，所述封装单元701在根据第i个轨道内的w个图像的封装过程生成第i个轨道对应的自由视角信息数据盒时，执行如下步骤：

将k个相机中第j个相机的标识信息存储在所述第二相机标识字段，k表示所述w个图像来源相机的数量；

获取所述第j个相机的相机属性信息，并将获取到的相机属性信息存储在所述相机属性信息字段；

所述第j个相机的相机属性信息包括以下任意一种或多种：所述第j个相机位置的横轴分量取值、纵轴分量取值以及竖轴分量取值，所述第j个相机焦距的横轴分量取值和纵轴分量取值，以及所述第j个相机采集图像的分辨率宽度与高度。

在一个实施例中，所述第i个轨道对应的视角信息还包括图像信息字段，所述封装单元 701在根据第i个轨道内的w个图像的封装过程生成第i个轨道对应的自由视角信息数据盒时，执行如下步骤：

获取所述第j个相机采集的图像的图像信息，并将获取到的图像信息存储在所述图像信息字段；其中，所述图像信息包括如下一个或多个：深度图像降采样的倍数因子、深度图像左上顶点相对于平面帧原点的偏移量、纹理图像左上顶点相对于平面帧原点的偏移量。

在一个实施例中，所述第i个轨道对应的视角信息还包括自定义相机参数字段、自定义相机参数类型字段以及自定义相机参数长度字段；所述封装单元701在根据第i个轨道内的w个图像的封装过程生成第i个轨道对应的自由视角信息数据盒时，执行如下步骤：

获取第f个自定义相机参数，并将所述第f个自定义相机参数存储在所述自定义相机参数字段，f为大于等于0且小于h的整数，h表示所述第j个相机的自定义相机参数的数量；

确定所述第f个自定义相机参数所属的参数类型以及所述第f个自定义相机参数的长度，将所述第f个自定义相机参数所属的参数类型存储在所述自定义相机参数类型字段，以及将所述第f个自定义相机参数的长度存储在所述自定义相机参数长度字段。

在一个实施例中，所述生成单元702还用于：生成所述沉浸媒体对应的信令描述文件，所述信令描述文件包括所述沉浸媒体对应的自由视角相机描述子，所述自由视角相机描述子用于记录每个轨道内视频片段对应的相机属性信息，任一轨道内视频片段是由所述任一轨道内封装的图像组成的；所述自由视角相机描述子被封装于所述沉浸媒体的媒体呈现描述文件的自适应集层级中，或者所述信令描述文件被封装于所述媒体呈现描述文件的表示层级中。

在一个实施例中，所述沉浸媒体的数据处理装置还包括发送单元703，用于将所述信令描述文件发送至内容消费设备，以指示所述内容消费设备基于所述自由视角相机描述子中记录的每个轨道内视频片段对应的相机属性信息，从所述多个相机中选择与用户所在位置匹配的候选相机，以及发送获取来源于所述候选相机的分块视频的第一资源请求；响应于所述第一资源请求，根据所述M个轨道中每个轨道对应的自由视角信息数据盒中的视角信息从所述M个轨道中选择第i个轨道并将所述i个轨道对应的自由视角信息数据盒发送至所述内容消费设备。

在一个实施例中，所述发送单元703还用于：将所述信令描述文件发送至内容消费设备，以指示所述内容消费设备根据所述信令描述文件和用户带宽发送第二资源请求；响应于所述第二资源请求，将所述M个轨道对应的M个自由视角信息数据盒发送至所述内容消费设备，以指示所述内容消费设备根据每个轨道对应的自由视角信息数据盒中的视角信息和用户所在位置信息，从所述M个自由视角信息数据盒中获取第i个轨道对应的自由视角信息数据盒。

根据本申请的一个实施例，图5所示的沉浸媒体的数据处理方法所涉及各个步骤可以由图7所示的沉浸媒体的数据处理装置中的各个单元来执行的。例如，图5所述的S501可由图7中所述的数据处理装置中的封装单元702来执行，S502可由图7所述的数据处理装置中的生成单元702来执行。

根据本申请的另一个实施例，图7所示的沉浸媒体的数据处理装置中的各个单元可以分别或全部合并为一个或若干个另外的单元来构成，或者其中的某个（些）单元还可以再拆分为功能上更小的多个单元来构成，这可以实现同样的操作，而不影响本申请的实施例的技术效果的实现。上述单元是基于逻辑功能划分的，在实际应用中，一个单元的功能也可以由多个单元来实现，或者多个单元的功能由一个单元实现。在本申请的其它实施例中，基于沉浸媒体的数据处理装置也可以包括其它单元，在实际应用中，这些功能也可以由其

它单元协助实现，并且可以由多个单元协作实现。

根据本申请的另一个实施例，可以通过在包括中央处理单元（CPU）、随机存取存储介质（RAM）、只读存储介质（ROM）等处理元件和存储元件的例如计算机的通用计算设备上运行能够执行如图5的相应方法所涉及的各步骤的计算机程序（包括程序代码），来构造
5 如图7中所示的沉浸媒体的数据处理装置，以及来实现本申请实施例沉浸媒体的数据处理方法。所述计算机程序可以记载于例如计算机可读存储介质上，并通过计算机可读存储介质装载于上述计算设备中，并在其中运行。

本申请实施例中将一沉浸媒体封装至M个轨道，该沉浸媒体是由处于不同视角的N个相机拍摄的图像组成的，一个轨道中可以包括来自至少一个相机的图像，这M个轨道是属于
10 同一个轨道组的，实现了一个沉浸视频封装到多个轨道的场景；另外，内容制作设备为每个轨道生成一个自由视角信息数据盒，通过第i个轨道对应的自由视角信息数据盒指示第i个轨道对应的视角信息，比如相机所在的具体视角位置，进而内容消费设备在根据第i个轨道对应的视角信息进行该轨道内图像的解码显示时，可以保证解码显示的视频与用户所在位置更加匹配，提高沉浸视频的呈现效果。

15 参见图8，为本申请实施例提供的一种内容消费设备的结构示意图。图8所示的内容消费设备可以指沉浸媒体的使用者所使用的计算机设备，该计算机设备可以是终端。图8所示的内容消费设备可以包括接收器801、处理器802、存储器803以及显示/播放装置804。其中：

接收器801用于实现解码与其他设备的传输交互，具体用于实现内容制作设备与内容消费设备之间关于沉浸媒体的传输。即内容消费设备通过接收器901接收内容制作设备传输
20 沉浸媒体的相关媒体资源。

处理器802或称CPU（Central Processing Unit，中央处理器）是内容制作设备的处理核心，该处理器802适于实现一条或多条计算机程序，具体适于加载并执行一条或多条计算机程序从而实现图4所示的沉浸媒体的数据处理方法的流程。

存储器803是内容消费设备中的记忆设备，用于存储计算机程序和媒体资源。可以理解的，此处的存储器803既可以包括内容消费设备中的内置存储介质，当然也可以包括内容消费设备所支持的扩展存储介质。需要说明的是，存储器803可以是高速RAM存储器，也可以是非不稳定的存储器（non-volatile memory），例如至少一个磁盘存储器；可选的还可以是至少一个位于远离前述处理器的存储器。存储器803提供存储空间，该存储空间用于存储内容消费设备的操作系统。并且，在该存储空间中还用于存储计算机程序，该计算机程序
25 适于被处理器调用并执行，以用来执行沉浸媒体的数据处理方法的各步骤。另外，存储器803还可用于存储经处理器处理后形成的沉浸媒体的三维图像、三维图像对应的音频内容及该三维图像和音频内容渲染所需的信息等。

显示/播放装置804用于输出渲染得到的声音和三维图像。

在一个实施例中，处理器802可包括解析器821、解码器822、转换器823以及渲染器824；

35 其中：

解析器821用于对来自内容制作设备的渲染媒体的封装文件进行解封装，具体是按照沉浸媒体的文件格式要对媒体文件资源进行解封装，得到音频码流和视频码流；并将该音频

码流和视频码流提高给解码器822;

解码器822对音频码流进行音频解码,得到音频内容并提供给渲染器824进行音频渲染。另外,解码器822对视频码流进行解码得到2D图像。根据媒体呈现描述信息提供的元数据,如果该元数据指示沉浸媒体执行过区域封装过程,该2D图像是指封装图像;如果该元数据指示沉浸媒体未执行过区域封装过程,则该平面图像是指投影图像。

转换器823用于将2D图像转换为3D图像。如果沉浸媒体执行过区域封装过程,转换器923还会先将封装图像进行区域解封装得到投影图像。再对投影图像进行重建处理得到3D图像。如果渲染媒体未执行过区域封装过程,转换器923会直接将投影图像重建得到3D图像。

渲染器824用于对沉浸媒体的音频内容和3D图像进行渲染。具体根据媒体呈现描述信息中与渲染、视窗相关的元数据对音频内容及3D图像进行渲染,渲染完成交由显示/播放装置进行输出。

在一个实施例中,处理器802通过调用存储器中的一条或多条计算机程序执行图4所示的沉浸媒体的数据处理方法的各个步骤。具体地,存储器存储一条或多条计算机程序,该一条或多条计算机程序适于由处理器802加载并执行如下步骤:

获取第*i*个轨道对应的自由视角信息数据盒,所述自由视角信息数据盒子,包括所述第*i*个轨道对应的视角信息,*i*为大于等于1且小于等于*M*的整数;该沉浸媒体是由处于不同视角的*N*个相机拍摄的图像组成,所述沉浸媒体被封装到*M*个轨道,一个轨道中封装来自至少一个相机的图像,*N*和*M*均为大于1的整数。

根据所述自由视角信息数据盒中的视角信息对所述第*i*个轨道内封装的图像进行解码显示。

本申请实施例中将一沉浸媒体封装至*M*个轨道,该沉浸媒体是由处于不同视角的*N*个相机拍摄的图像组成的,一个轨道中可以包括来自至少一个相机的图像,这*M*个轨道是属于同一个轨道组的,实现了一个沉浸视频封装到多个轨道的场景;另外,内容制作设备为每个轨道生成一个自由视角信息数据盒,通过第*i*个轨道对应的自由视角信息数据盒指示第*i*个轨道对应的视角信息,比如相机所在的具体视角位置,进而内容消费设备在根据第*i*个轨道对应的视角信息进行该轨道内图像的解码显示时,可以保证解码显示的视频与用户所在位置更加匹配,提高沉浸视频的呈现效果。

参见图9,为本申请实施例提供的一种内容制作设备的结构示意图;该内容制作设备可以是指沉浸媒体的提供者所使用的计算机设备,该计算机设备可以是终端或服务器。如图9所示,该内容制作设备可以包括捕获设备901、处理器902、存储器903以及发射器904。其中:

捕获设备901用于采集现实世界的声音-视觉场景获得沉浸媒体的原始数据(包括在时间和空间上保持同步的音频内容和视频内容)。该捕获设备801可以包括但不限于:音频设备、摄像设备及传感设备。其中,音频设备可以包括音频传感器、麦克风等。摄像设备可以包括普通摄像头、立体摄像头、光场摄像头等。传感设备可以包括激光设备、雷达设备等。

处理器902（或称CPU（Central Processing Unit，中央处理器））是内容制作设备的处理核心，该处理器902适于实现一条或多条计算机程序，具体适于加载并执行一条或多条计算机程序从而实现图4所示的沉浸媒体的数据处理方法的流程。

5 存储器903是内容制作设备中的记忆设备，用于存放程序和媒体资源。可以理解的是，此处的存储器903既可以包括内容制作设备中的内置存储介质，当然也可以包括内容制作设备所支持的扩展存储介质。需要说明的是，存储器可以是高速RAM存储器，也可以是非不稳定的存储器（non-volatile memory），例如至少一个磁盘存储器；可选的还可以是至少一个位于远离前述处理器的存储器。存储器提供存储空间，该存储空间用于存储内容制作设备的操作系统。并且，在该存储空间中还用于存储计算机程序，该计算机程序包括程序指令，且该程序指令适于被处理器调用并执行，以用来执行沉浸媒体的数据处理方法的各步骤。另外，存储器903还可用于存储经处理器处理后形成的沉浸媒体文件，该沉浸媒体文件包括媒体文件资源和媒体呈现描述信息。

15 发射器904用于实现内容制作设备与其他设备的传输交互，具体用于实现内容制作设备与内容消费设备之间关于进行沉浸媒体的传输。即内容制作设备通过发射器904来向内容消费设备传输沉浸媒体的相关媒体资源。

再参见图9可知，处理器902可以包括转换器921、编码器922和封装器923。其中：

20 转换器921用于对捕获到的视频内容进行一系列转换处理，使视频内容成为适合被执行沉浸媒体的视频编码的内容。转换处理可包括：拼接和投影，可选地，转换处理还包括区域封装。转换器921可以将捕获到的3D视频内容转换为2D图像，并提供给编码器进行视频编码。

编码器922用于对捕获到的音频内容进行音频编码形成沉浸媒体的音频码流。还用于对转换器921转换得到的2D图像进行视频编码，得到视频码流。

25 封装器923用于将音频码流和视频码流按照沉浸媒体的文件格式（如ISOBMFF）封装在文件容器中形成沉浸媒体的媒体文件资源，该媒体文件资源可以是媒体文件或媒体片段形成沉浸媒体的媒体文件；并按照沉浸媒体的文件格式要求采用媒体呈现描述信息记录该沉浸媒体的媒体文件资源的元数据。封装器处理得到的沉浸媒体的封装文件会保存在存储器中，并按需提供给内容消费设备进行沉浸媒体的呈现。

30 在一个实施例中，处理器902通过调用存储器中的一条或多条指令来执行图5所示的沉浸媒体的数据处理方法的各个步骤。具体地，存储器803存储有一条或多条计算机程序，该计算机程序适于由处理器902加载并执行但不限于如下步骤：

将所述沉浸媒体封装到M个轨道中，所述沉浸媒体是由处于不同视角的N个相机拍摄的图像组成的，一个轨道中封装来自至少一个相机的图像，所述M个轨道属于同一个轨道组，N和M均为大于或等于1的整数；

35 根据第i个轨道中w个图像的封装过程，生成第i个轨道对应的自由视角信息数据盒，所述自由视角信息数据盒包括所述第i个轨道对应的视角信息； $1 \leq i \leq M$ ， $w \geq 1$ 。

本申请实施例中将一沉浸媒体封装至M个轨道，该沉浸媒体是由处于不同视角的N个相机拍摄的图像组成的，一个轨道中可以包括来自至少一个相机的图像，这M个轨道是属于

—28—

同一个轨道组的，实现了一个沉浸视频封装到多个轨道的场景；另外，内容制作设备为每个轨道生成一个自由视角信息数据盒，通过第i个轨道对应的自由视角信息数据盒指示第i个轨道对应的视角信息，比如相机所在的具体视角位置，进而内容消费设备在根据第i个轨道对应的视角信息进行该轨道内图像的解码显示时，可以保证解码显示的视频与用户所在位置更加匹配，提高沉浸视频的呈现效果。

另外，本申请实施例还提供了一种存储介质，所述存储介质用于存储计算机程序，所述计算机程序用于执行上述实施例提供的方法。

本申请实施例还提供了一种包括指令的计算机程序产品，当其在计算机上运行时，使得计算机执行上述实施例提供的方法。

以上所揭露的仅为本申请较佳实施例而已，当然不能以此来限定本申请之权利范围，因此依本申请权利要求所作的等同变化，仍属本申请所涵盖的范围。

权 利 要 求

1、一种沉浸媒体的数据处理方法，所述方法由内容消费设备执行，所述沉浸媒体是由处于不同视角的N个相机拍摄的图像组成，所述沉浸媒体被封装到M个轨道，一个轨道中封装来自至少一个相机的图像，N和M均为大于1的整数；所述方法包括：

5 获取所述沉浸媒体的第i个轨道对应的自由视角信息数据盒，所述自由视角信息数据盒包括所述第i个轨道对应的视角信息，i为大于或等于1且小于或等于M的整数；

根据所述自由视角信息数据盒中的视角信息对所述第i个轨道内封装的图像进行解码。

2、如权利要求1所述的方法，所述第i个轨道内封装了来自所述N个相机中k个相机的图像，k为大于0的整数；

10 所述第i个轨道内还封装了所述第i个轨道对应的自由视角轨道组信息，所述自由视角轨道组信息用于指示所述第i个轨道与封装所述沉浸媒体的其他轨道属于同一个轨道组；所述自由视角轨道组信息包括第一相机标识字段和图像类型字段；

所述第一相机标识字段用于存储所述k个相机中第j个相机的标识信息，所述图像类型字段用于指示所述第j个相机采集的图像所属的图像类型，所述图像类型包括纹理图像或深度图像中的至少一个。

15 3、如权利要求2所述的方法，若所述第i个轨道内包括纹理图像和深度图像，所述第i个轨道对应的视角信息包括视频拼接布局指示字段；

当所述视频拼接布局指示字段为第一数值时，指示所述第i个轨道内包括的纹理图像和深度图像是拼接编码的；

20 当所述视频拼接布局指示字段为第二数值时，指示所述第i个轨道内包括的纹理图像和深度图像是分别编码的。

4、如权利要求2所述的方法，所述第i个轨道对应的视角信息还包括相机模型字段；当所述相机模型字段为第三数值时，指示所述j个相机所属相机模型为第一模型；当所述相机模型字段为第四数值时，指示所述j个相机所属相机模型为第二模型，所述第二模型不同于所述第一模型。

25 5、如权利要求2所述的方法，所述第i个轨道对应的视角信息还包括纹理图像的保护带宽度字段和深度图像的保护带宽度字段，所述纹理图像的保护带宽度字段用于存储所述第i个轨道内纹理图像进行编码时采用的保护带宽度，所述深度图像的保护带宽度字段用于存储所述第i个轨道内深度图像进行编码时采用的保护带宽度。

30 6、如权利要求2所述的方法，所述第i个轨道对应的视角信息还包括第二相机标识字段和相机属性信息字段；

所述第二相机标识字段用于存储第j个相机的标识信息；

所述相机属性信息字段用于存储所述第j个相机的相机属性信息，所述第j个相机的相机属性信息包括以下至少一种：所述第j个相机位置的横轴分量取值、纵轴分量取值以及竖轴分量取值，所述第j个相机焦距的横轴分量取值和纵轴分量取值，以及第j个相机采集图像的分辨率宽度与高度。

7、如权利要求2所述的方法，所述第i个轨道对应的视角信息还包括图像信息字段，所

述图像信息字段用于存储所述第j个相机采集图像的图像信息，所述图像信息包括以下至少一个：深度图像降采样的倍数因子、深度图像左上顶点相对于平面帧原点的偏移量或纹理图像左上顶点相对于平面帧原点的偏移量。

8、如权利要求2所述的方法，所述第i个轨道对应的视角信息还包括自定义相机参数字段、自定义相机参数类型字段以及自定义相机参数长度字段；所述自定义相机参数字段用于存储所述第j个相机的第f个自定义相机参数，f为大于等于0且小于h的整数，h表示所述第j个相机的自定义相机参数的数量；所述自定义相机参数类型字段用于存储所述第f个自定义相机参数所属的参数类型；所述自定义相机参数长度字段用于存储所述第f个自定义相机参数的长度。

9、如权利要求1~8任一项所述的方法，所述方法还包括：

获取所述沉浸媒体对应的信令描述文件，所述信令描述文件包括所述沉浸媒体对应的自由视角相机描述子，所述自由视角相机描述子用于记录每个轨道内视频片段对应的相机属性信息，所述轨道内视频片段是由所述轨道内包括的纹理图像和深度图像组成的；

所述自由视角相机描述子被封装于所述沉浸媒体的媒体呈现描述文件的自适应集层级中，或者所述信令描述文件被封装于所述媒体呈现描述文件的表示层级中。

10、如权利要求9所述的方法，所述获取所述沉浸媒体的第i个轨道对应的自由视角信息数据盒，包括：

基于所述自由视角相机描述子中记录的每个轨道内视频片段对应的相机属性信息，从所述N个相机中选择与用户所在位置信息匹配的候选相机；

向内容制作设备发送获取来源于所述候选相机的分块视频的第一资源请求，所述第一资源请求用于指示所述内容制作设备根据所述M个轨道中每个轨道对应的自由视角信息数据盒中的视角信息，从所述M个轨道中选择第i个轨道并返回所述第i个轨道对应的自由视角信息数据盒，所述第i个轨道中封装的纹理图像和深度图像中的至少一种来自所述候选相机；

接收所述内容制作设备返回的所述第i个轨道对应的自由视角信息数据盒。

11、如权利要求9所述的方法，所述获取所述沉浸媒体的第i个轨道对应的自由视角信息数据盒，包括：

基于所述信令描述文件和用户带宽向内容制作设备发送第二资源请求，所述第二资源请求用于指示所述内容制作设备返回所述M个轨道的M个自由视角信息数据盒，一个轨道对应一个自由视角信息数据盒；

根据每个轨道对应的自由视角信息数据盒中的视角信息和用户所在位置信息，从所述M个自由视角信息数据盒中获取第i个轨道对应的自由视角信息数据盒。

12、一种媒体数据处理方法，所述方法由内容制作设备执行，所述方法包括：

将沉浸媒体封装到M个轨道中，所述沉浸媒体是由处于不同视角的N个相机拍摄的图像组成的，一个轨道中封装来自至少一个相机的图像，所述M个轨道属于同一个轨道组，N和M均为大于或等于1的整数；

根据第i个轨道中w个图像的封装过程，生成第i个轨道对应的自由视角信息数据盒，所述自由视角信息数据盒包括所述第i个轨道对应的视角信息； $1 \leq i \leq M$ ， $w \geq 1$ 。

13、如权利要求12所述的方法，所述第i个轨道对应的自由视角信息数据盒包括自由视角轨道组信息，所述自由视角轨道组信息包括第一相机标识字段和图像类型字段，所述根据第i个轨道内w个图像的封装过程生成所述第i个轨道对应的自由视角信息数据盒，包括：

确定所述w个图像来源于k个相机，k为大于1的整数；

5 将所述k个相机中的第j个相机的标识信息存储在所述第一相机标识字段中；

根据所述第j个相机采集的图像所属图像类型确定所述图像类型字段，所述图像类型包括深度图像或纹理图像中的至少一个。

14、如权利要求13所述的方法，所述根据所述第j个相机采集的图像所属图像类型确定所述图像类型字段，包括：

10 当所述第j个相机采集的图像所属图像类型为纹理图像，则将所述图像类型字段设置为第二数值；

当所述第j个相机采集的图像所属图像类型为深度图像，则将所述图像类型字段设置为第三数值；

15 当所述第j个相机采集的图像所属图像类型为纹理图像和深度图像，则将所述图像类型字段设置为第四数值。

15、如权利要求12所述的方法，所述第i个轨道对应的视角信息包括视频拼接布局指示字段，所述w个图像包括纹理图像和深度图像，所述根据第i个轨道内w个图像的封装过程生成第i个轨道对应的自由视角信息数据盒，包括：

20 如果所述第i个轨道内包含的纹理图像和深度图像是拼接编码的，则将所述视频布局指示字段设置为第一数值；

如果所述i个轨道内包含的的纹理图像和深度图像是分别编码的，则将所述视频布局指示字段设置为第二数值。

25 16、如权利要求15所述的方法，所述第i个轨道对应的视角信息还包括相机模型字段；所述根据第i个轨道内的w个图像的封装过程生成第i个轨道对应的自由视角信息数据盒，包括：

当采集所述w个图像的相机所属的相机模型为第一模型时，将所述相机模型字段设置为第一数值；

当采集所述w个图像的相机所属的相机模型为第二模型时，将所述相机模型字段设置为第二数值。

30 17、如权利要求16所述的方法，所述第i个轨道对应的视角信息还包括纹理图像的保护带宽字段和深度图像的保护带宽字段，所述根据第i个轨道内w个图像的封装过程生成第i个轨道对应的自由视角信息数据盒，包括：

获取所述第i个轨道内在对纹理图像进行编码时采用的纹理图像的保护带宽度，并将所述纹理图像的保护带宽度存储在所述纹理图像的保护带宽度字段；

35 获取所述第i个轨道内在对深度图像进行编码时采用的深度图像的保护带宽度，并将所述深度图像的保护带宽度存储在所述深度图像的保护带宽度字段。

18、如权利要求17所述的方法，所述第i个轨道对应的视角信息还包括第二相机标识字

段和相机属性信息字段, 所述根据第i个轨道内的w个图像的封装过程生成第i个轨道对应的自由视角信息数据盒, 包括:

将k个相机中第j个相机的标识信息存储在所述第二相机标识字段, k表示所述w个图像来源相机的数量;

5 获取所述第j个相机的相机属性信息, 并将获取到的相机属性信息存储在所述相机属性信息字段;

所述第j个相机的相机属性信息包括以下任意一种或多种: 所述第j个相机位置的横轴分量取值、纵轴分量取值以及竖轴分量取值, 所述第j个相机焦距的横轴分量取值和纵轴分量取值, 以及所述第j个相机采集图像的分辨率宽度与高度。

10 19、如权利要求18所述的方法, 所述第i个轨道对应的视角信息还包括图像信息字段, 所述根据第i个轨道内的w个图像的封装过程生成第i个轨道对应的自由视角信息数据盒, 包括:

获取所述第j个相机采集的图像的图像信息, 并将获取到的图像信息存储在所述图像信息字段; 其中, 所述图像信息包括如下一个或多个: 深度图像降采样的倍数因子、深度图像左上顶点相对于平面帧原点的偏移量、纹理图像左上顶点相对于平面帧原点的偏移量。

15 20、如权利要求19所述的方法, 所述第i个轨道对应的视角信息还包括自定义相机参数字段、自定义相机参数类型字段以及自定义相机参数长度字段; 所述根据第i个轨道内的w个图像的封装过程生成第i个轨道对应的自由视角信息数据盒, 包括:

20 获取第f个自定义相机参数, 并将所述第f个自定义相机参数存储在所述自定义相机参数字段, f为大于等于0且小于h的整数, h表示所述第j个相机的自定义相机参数的数量;

确定所述第f个自定义相机参数所属的参数类型以及所述第f个自定义相机参数的长度, 将所述第f个自定义相机参数所属的参数类型存储在所述自定义相机参数类型字段, 以及将所述第f个自定义相机参数的长度存储在所述自定义相机参数长度字段。

21、如权利要求12~20任一项所述的方法, 所述方法还包括:

25 生成所述沉浸媒体对应的信令描述文件, 所述信令描述文件包括所述沉浸媒体对应的自由视角相机描述子, 所述自由视角相机描述子用于记录每个轨道内视频片段对应的相机属性信息, 任一轨道内视频片段是由所述任一轨道内封装的图像组成的;

所述自由视角相机描述子被封装于所述沉浸媒体的媒体呈现描述文件的自适应集层级中, 或者所述信令描述文件被封装于所述媒体呈现描述文件的表示层级中。

30 22、如权利要求21所述的方法, 所述方法还包括:

将所述信令描述文件发送至内容消费设备, 以指示所述内容消费设备基于所述自由视角相机描述子中记录的每个轨道内视频片段对应的相机属性信息, 从所述多个相机中选择与用户所在位置匹配的候选相机, 以及发送获取来源于所述候选相机的分块视频的第一资源请求;

35 响应于所述第一资源请求, 根据所述M个轨道中每个轨道对应的自由视角信息数据盒中的视角信息从所述M个轨道中选择第i个轨道并将所述i个轨道对应的自由视角信息数据盒发送至所述内容消费设备。

23、如权利要求21所述的方法，所述方法还包括：

将所述信令描述文件发送至内容消费设备，以指示所述内容消费设备根据所述信令描述文件和用户带宽发送第二资源请求；

5 响应于所述第二资源请求，将所述M个轨道对应的M个自由视角信息数据盒发送至所述内容消费设备，以指示所述内容消费设备根据每个轨道对应的自由视角信息数据盒中的视角信息和用户所在位置信息，从所述M个自由视角信息数据盒中获取第i个轨道对应的自由视角信息数据盒。

10 24、一种沉浸媒体的数据处理装置，所述沉浸媒体是由处于不同视角的N个相机拍摄的图像组成，所述沉浸媒体被封装到M个轨道，一个轨道中封装来自至少一个相机的图像，N和M均为大于1的整数；所述装置包括：

获取单元，用于获取所述沉浸媒体的第i个轨道对应的自由视角信息数据盒，所述自由视角信息数据盒包括所述第i个轨道对应的视角信息，i为大于或等于1，且小于或等于M的整数；

15 处理单元，用于根据所述自由视角信息数据盒中的视角信息对所述第i个轨道内封装的图像进行解码。

25、一种沉浸媒体的数据处理装置，包括：

封装单元，用于将所述沉浸媒体封装到M个轨道中，所述沉浸媒体是由处于不同视角的N个相机拍摄的图像组成的，一个轨道中封装来自至少一个相机的图像，所述M个轨道属于同一个轨道组，N和M均为大于或等于1的整数；

20 生成单元，用于根据第i个轨道中w个图像的封装过程，生成第i个轨道对应的自由视角信息数据盒，所述自由视角信息数据盒包括所述第i个轨道对应的视角信息； $1 \leq i \leq M$ ， $w \geq 1$ 。

26、一种计算机设备，包括：

处理器，适于实现一条或多条计算机程序；以及

25 计算机存储介质，所述计算机存储介质存储有一条或多条计算机程序，所述一条或计算机程序适于由处理器加载并执行如权利要求1~23任一项所述的方法。

30 27、一种计算机存储介质，所述计算机存储介质中存储有第一计算机程序程序和第二计算机程序，所述第一计算机程序被处理器执行时，用于执行如权利要求1-11任一项所述的沉浸媒体的数据处理方法；所述第二计算机程序被处理器执行时，用于执行如权利要求12-23任一项所述的沉浸媒体的数据处理方法。

28、一种包括指令的计算机程序产品，当其在计算机上运行时，使得所述计算机执行权利要求1~23任一项所述的方法。

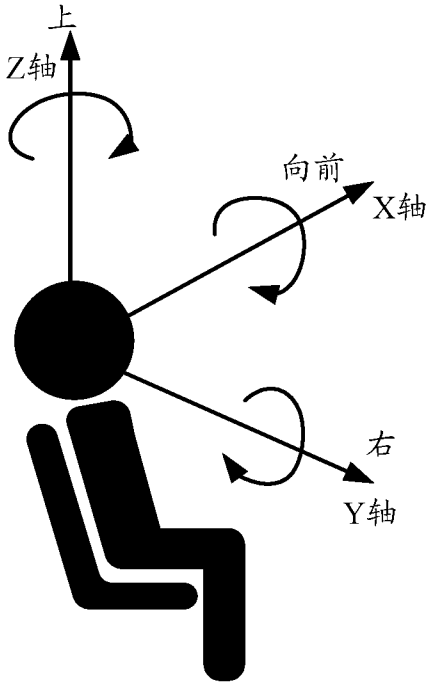


图 1a

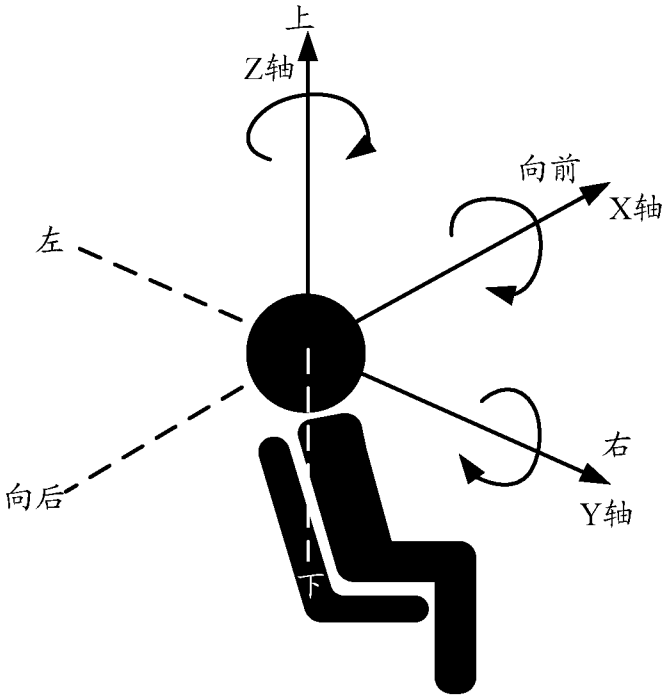


图 1b

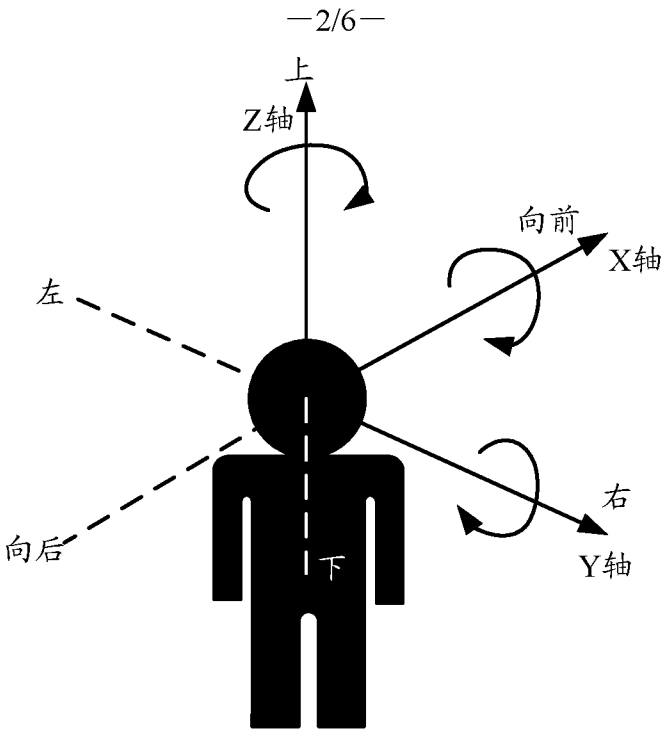


图 1c

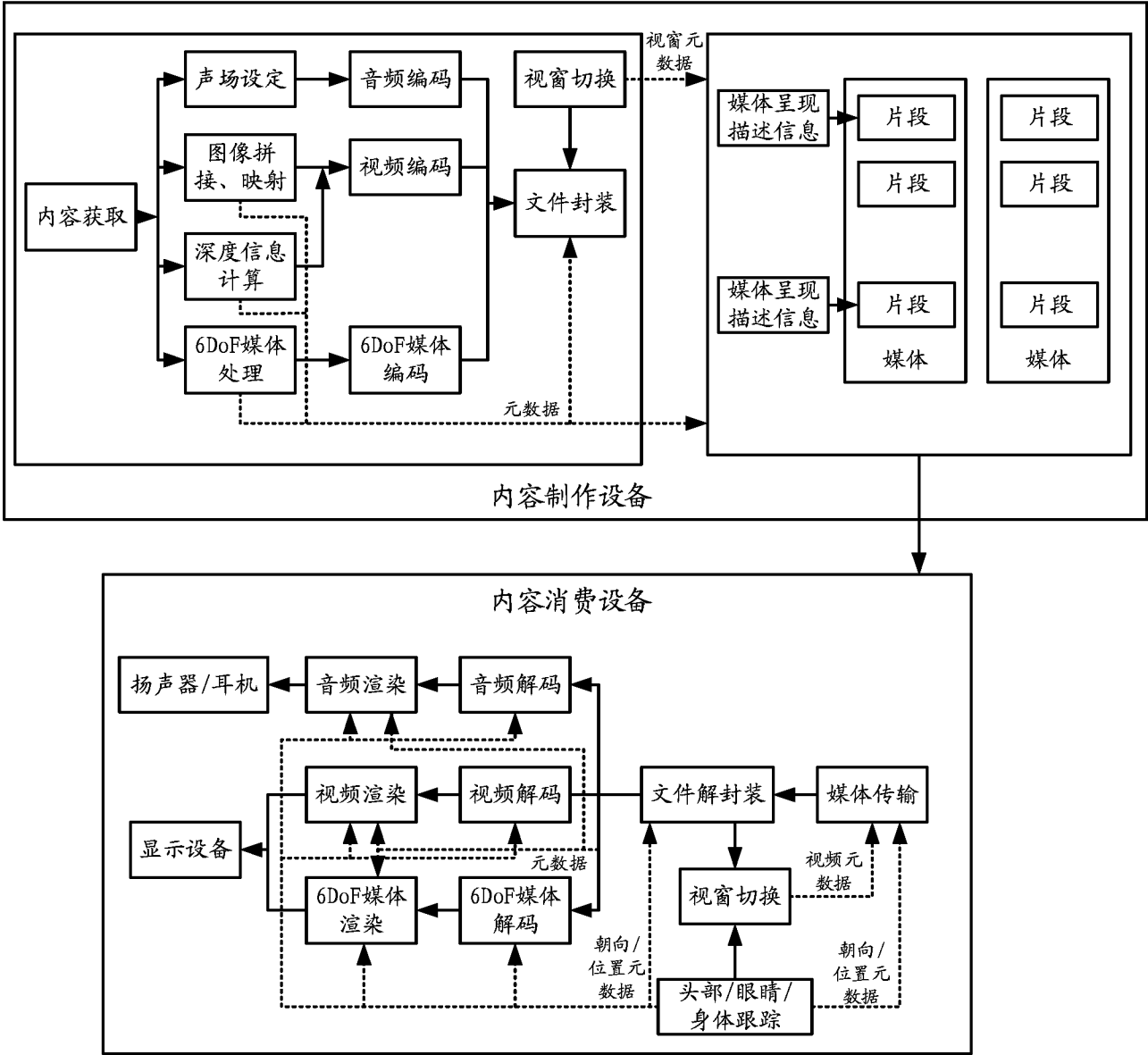


图 2a

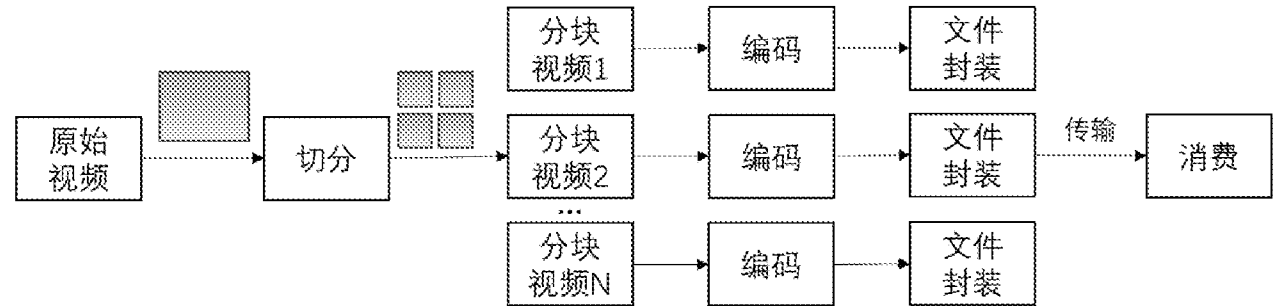


图 2b

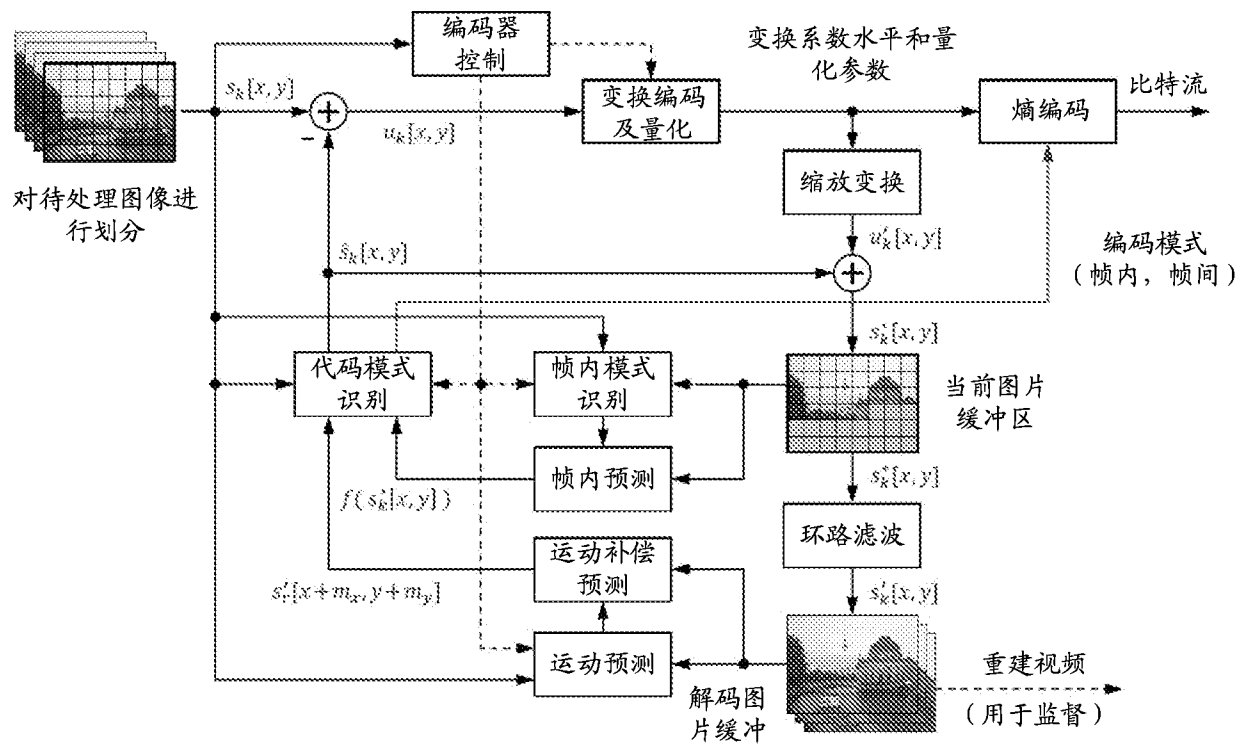
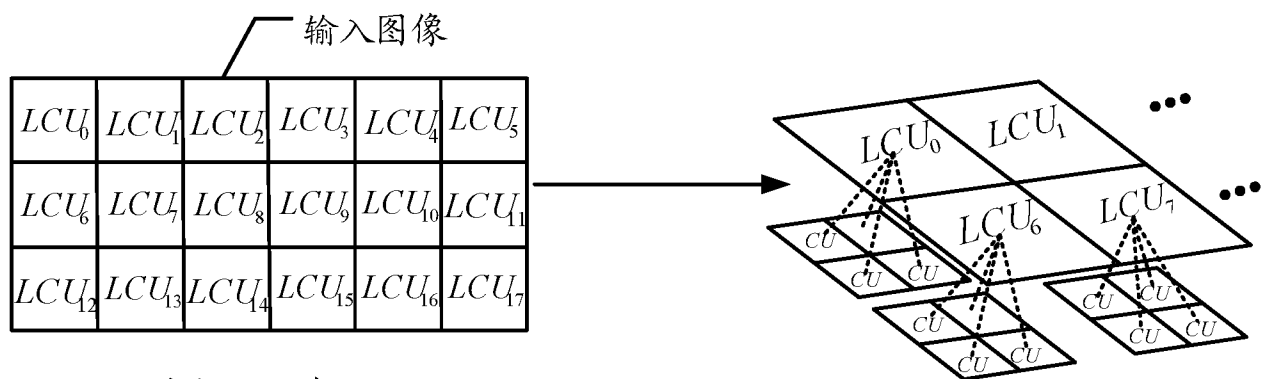


图 3a



LCU:最大编码单元

CU: 编码单元 (图像块)

图 3b

— 5/6 —

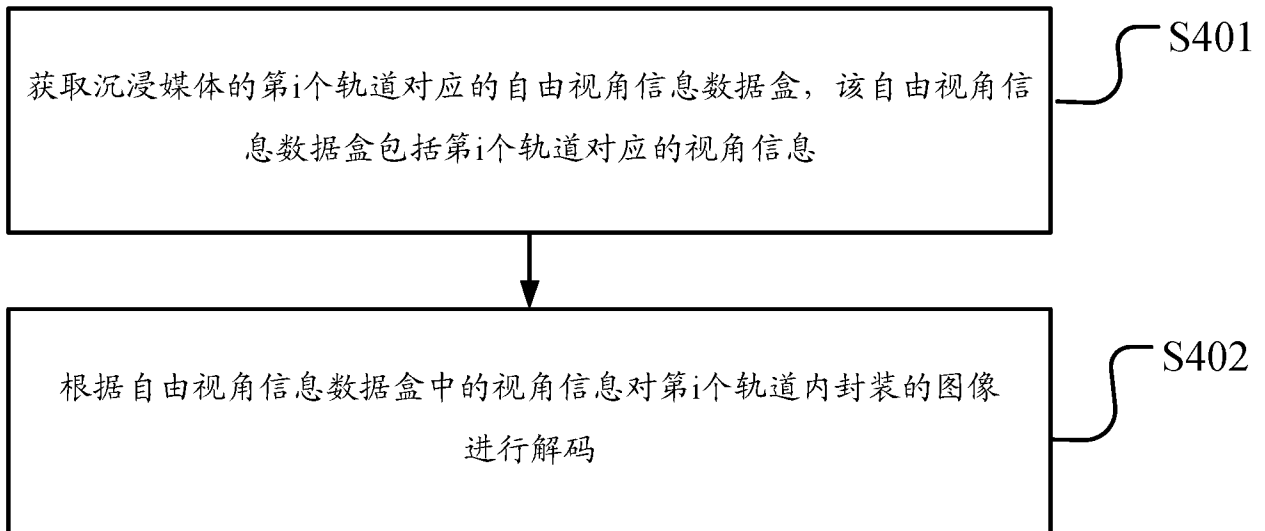


图 4

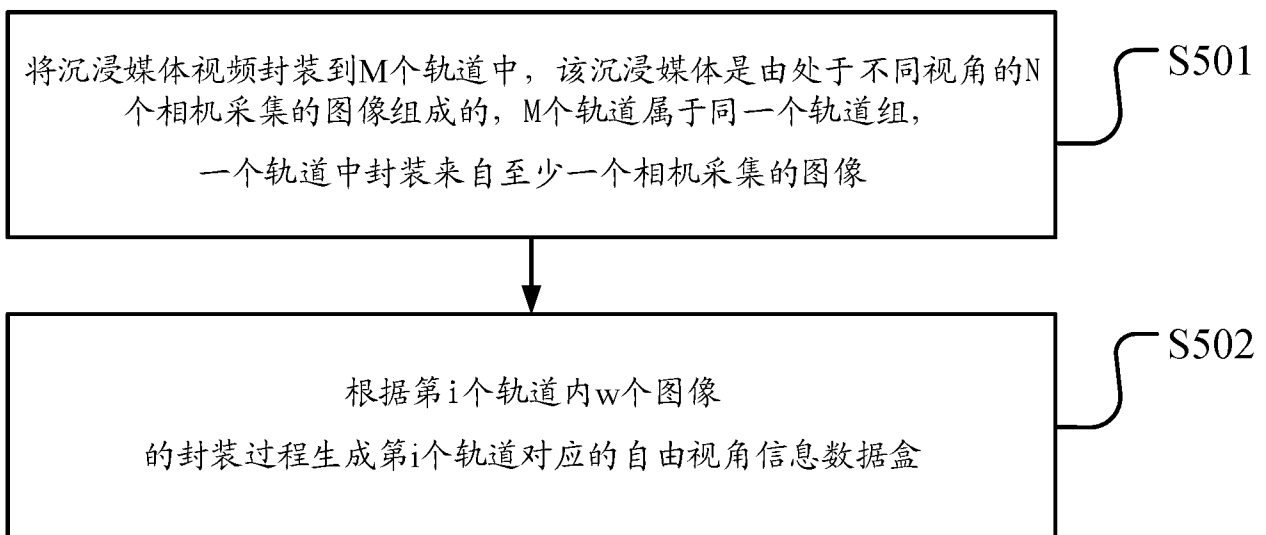


图 5

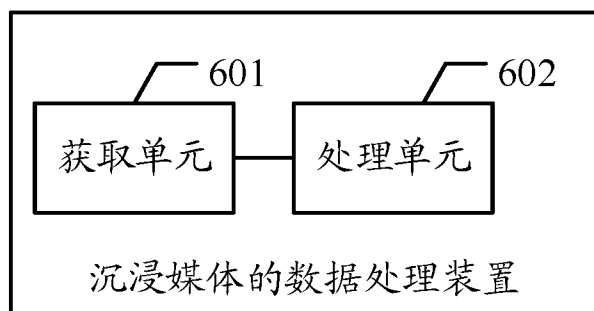


图 6

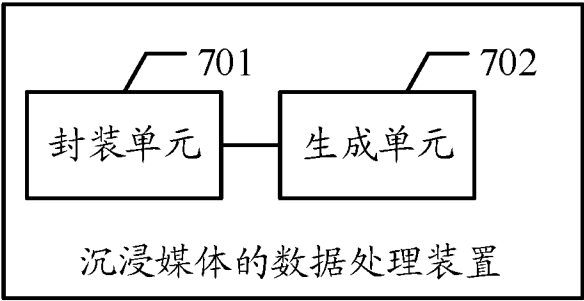


图 7

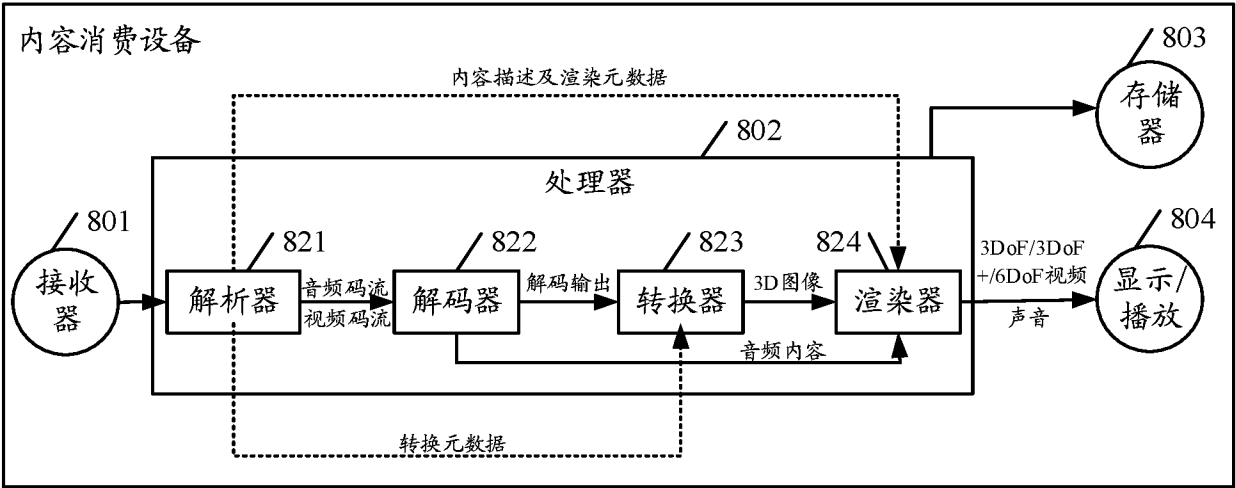


图 8

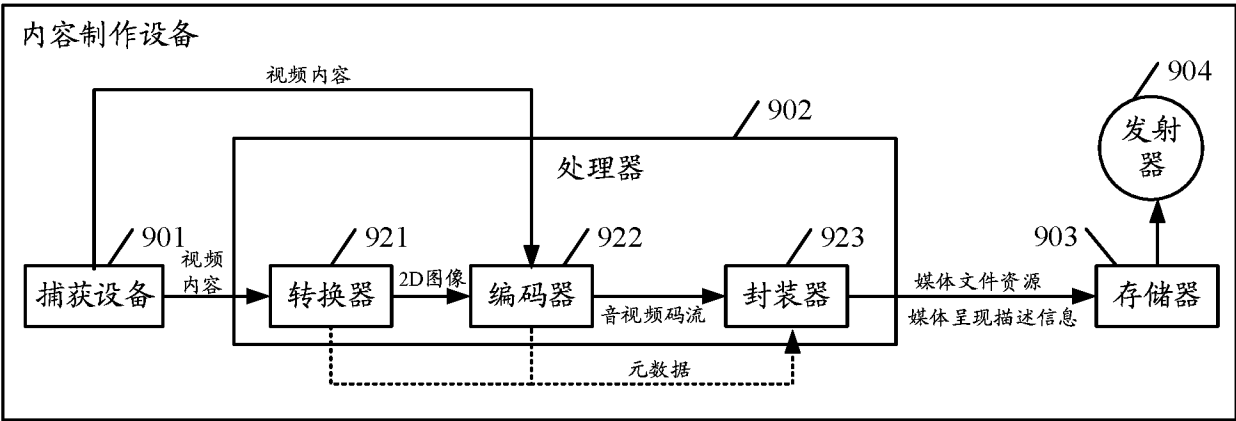


图 9

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2022/080257

A. CLASSIFICATION OF SUBJECT MATTER

H04N 13/117(2018.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

H04N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNTXT; ENTXT; ENTXTC; CJFD; CNKI: 沉浸, 视频, 媒体, 三维, 立体, 虚拟现实, 视点, 视角, 轨道, 数据盒, 深度, 纹理, immers, video, media, 3D, three, dimension+, stereo+, vr, virtual, reality, viewpoint, view, angle, track, databox, depth, texture

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 112804256 A (TENCENT TECHNOLOGY SHENZHEN CO., LTD.) 14 May 2021 (2021-05-14) description, paragraphs [0044]-[0107], and figures 2-8	1-28
A	CN 112492289 A (ZTE CORP.) 12 March 2021 (2021-03-12) entire document	1-28
A	CN 111264058 A (INTERDIGITAL VC HOLDINGS, INC.) 09 June 2020 (2020-06-09) entire document	1-28
A	US 2010128112 A1 (SAMSUNG ELECTRONICS CO., LTD.) 27 May 2010 (2010-05-27) entire document	1-28
A	WO 2019066191 A1 (LG ELECTRONICS INC.) 04 April 2019 (2019-04-04) entire document	1-28



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“E” earlier application or patent but published on or after the international filing date

“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

“O” document referring to an oral disclosure, use, exhibition or other means

“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

“&” document member of the same patent family

Date of the actual completion of the international search

26 April 2022

Date of mailing of the international search report

17 May 2022

Name and mailing address of the ISA/CN

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

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2022/080257

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	112804256	A	14 May 2021	None			
CN	112492289	A	12 March 2021	WO	2021259054	A1	30 December 2021
CN	111264058	A	09 June 2020	EP	3682632	A1	22 July 2020
				US	2020228777	A1	16 July 2020
				WO	2019055389	A1	21 March 2019
				KR	20200051784	A	13 May 2020
				VN	71033	A	25 June 2020
US	2010128112	A1	27 May 2010	KR	20110102365	A	16 September 2011
				EP	2356540	A2	17 August 2011
				WO	2010062117	A2	03 June 2010
				WO	2010062117	A3	30 June 2011
				EP	2356540	A4	17 September 2014
WO	2019066191	A1	04 April 2019	KR	20200032230	A	25 March 2020
				EP	3668092	A1	17 June 2020
				CN	111164969	A	15 May 2020
				EP	36680920	A4	07 October 2020
				CN	111164969	B	02 November 2021
				KR	102322508	B1	05 November 2021

国际检索报告

国际申请号

PCT/CN2022/080257

A. 主题的分类 H04N 13/117(2018.01) i 按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类																				
B. 检索领域 检索的最低限度文献(标明分类系统和分类号) H04N 包含在检索领域中的除最低限度文献以外的检索文献 在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用)) CNTXT;ENTXT;ENTXTC;CJFD;CNKI: 沉浸, 视频, 媒体, 三维, 立体, 虚拟现实, 视点, 视角, 轨道, 数据盒, 深度, 纹理, immers, video, media, 3D, three, dimension+, stereo+, vr, virtual, reality, viewpoint, view, angle, track, databox, depth, texture																				
C. 相关文件 <table border="1"> <thead> <tr> <th>类 型*</th> <th>引用文件, 必要时, 指明相关段落</th> <th>相关的权利要求</th> </tr> </thead> <tbody> <tr> <td>X</td> <td>CN 112804256 A (腾讯科技深圳有限公司) 2021年5月14日 (2021 - 05 - 14) 说明书第[0044]-[0107]段, 附图2-8</td> <td>1-28</td> </tr> <tr> <td>A</td> <td>CN 112492289 A (中兴通讯股份有限公司) 2021年3月12日 (2021 - 03 - 12) 全文</td> <td>1-28</td> </tr> <tr> <td>A</td> <td>CN 111264058 A (交互数字VC控股公司) 2020年6月9日 (2020 - 06 - 09) 全文</td> <td>1-28</td> </tr> <tr> <td>A</td> <td>US 2010128112 A1 (SAMSUNG ELECTRONICS CO LTD) 2010年5月27日 (2010 - 05 - 27) 全文</td> <td>1-28</td> </tr> <tr> <td>A</td> <td>WO 2019066191 A1 (LG ELECTRONICS INC) 2019年4月4日 (2019 - 04 - 04) 全文</td> <td>1-28</td> </tr> </tbody> </table>			类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求	X	CN 112804256 A (腾讯科技深圳有限公司) 2021年5月14日 (2021 - 05 - 14) 说明书第[0044]-[0107]段, 附图2-8	1-28	A	CN 112492289 A (中兴通讯股份有限公司) 2021年3月12日 (2021 - 03 - 12) 全文	1-28	A	CN 111264058 A (交互数字VC控股公司) 2020年6月9日 (2020 - 06 - 09) 全文	1-28	A	US 2010128112 A1 (SAMSUNG ELECTRONICS CO LTD) 2010年5月27日 (2010 - 05 - 27) 全文	1-28	A	WO 2019066191 A1 (LG ELECTRONICS INC) 2019年4月4日 (2019 - 04 - 04) 全文	1-28
类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求																		
X	CN 112804256 A (腾讯科技深圳有限公司) 2021年5月14日 (2021 - 05 - 14) 说明书第[0044]-[0107]段, 附图2-8	1-28																		
A	CN 112492289 A (中兴通讯股份有限公司) 2021年3月12日 (2021 - 03 - 12) 全文	1-28																		
A	CN 111264058 A (交互数字VC控股公司) 2020年6月9日 (2020 - 06 - 09) 全文	1-28																		
A	US 2010128112 A1 (SAMSUNG ELECTRONICS CO LTD) 2010年5月27日 (2010 - 05 - 27) 全文	1-28																		
A	WO 2019066191 A1 (LG ELECTRONICS INC) 2019年4月4日 (2019 - 04 - 04) 全文	1-28																		
☐ 其余文件在C栏的续页中列出。 ☒ 见同族专利附件。																				
<table border="0"> <tr> <td> * 引用文件的具体类型: “A” 认为不特别相关的表示了现有技术一般状态的文件 “E” 在国际申请日的当天或之后公布的在先申请或专利 “L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件(如具体说明的) “O” 涉及口头公开、使用、展览或其他方式公开的文件 “P” 公布日先于国际申请日但迟于所要求的优先权日的文件 </td> <td> “T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件 “X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性 “Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性 “&” 同族专利的文件 </td> </tr> </table>			* 引用文件的具体类型: “A” 认为不特别相关的表示了现有技术一般状态的文件 “E” 在国际申请日的当天或之后公布的在先申请或专利 “L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件(如具体说明的) “O” 涉及口头公开、使用、展览或其他方式公开的文件 “P” 公布日先于国际申请日但迟于所要求的优先权日的文件	“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件 “X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性 “Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性 “&” 同族专利的文件																
* 引用文件的具体类型: “A” 认为不特别相关的表示了现有技术一般状态的文件 “E” 在国际申请日的当天或之后公布的在先申请或专利 “L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件(如具体说明的) “O” 涉及口头公开、使用、展览或其他方式公开的文件 “P” 公布日先于国际申请日但迟于所要求的优先权日的文件	“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件 “X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性 “Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性 “&” 同族专利的文件																			
国际检索实际完成的日期		国际检索报告邮寄日期																		
2022年4月26日		2022年5月17日																		
ISA/CN的名称和邮寄地址		受权官员																		
中国国家知识产权局(ISA/CN) 中国北京市海淀区蓟门桥西土城路6号 100088 传真号 (86-10)62019451		李颖 电话号码 (86-27)59371623																		

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2022/080257

检索报告引用的专利文件			公布日 (年/月/日)	同族专利			公布日 (年/月/日)
CN	112804256	A	2021年5月14日	无			
CN	112492289	A	2021年3月12日	WO	2021259054	A1	2021年12月30日
CN	111264058	A	2020年6月9日	EP	3682632	A1	2020年7月22日
				US	2020228777	A1	2020年7月16日
				WO	2019055389	A1	2019年3月21日
				KR	20200051784	A	2020年5月13日
				VN	71033	A	2020年6月25日
US	2010128112	A1	2010年5月27日	KR	20110102365	A	2011年9月16日
				EP	2356540	A2	2011年8月17日
				WO	2010062117	A2	2010年6月3日
				WO	2010062117	A3	2011年6月30日
				EP	2356540	A4	2014年9月17日
WO	2019066191	A1	2019年4月4日	KR	20200032230	A	2020年3月25日
				EP	3668092	A1	2020年6月17日
				CN	111164969	A	2020年5月15日
				EP	36680920	A4	2020年10月7日
				CN	111164969	B	2021年11月2日
				KR	102322508	B1	2021年11月5日