

(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 2022/257567 A1

(51) 国际专利分类号:
H04N 19/31 (2014.01)

(21) 国际申请号: PCT/CN2022/083960

(22) 国际申请日: 2022 年 3 月 30 日 (30.03.2022)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
202110656768.4 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, 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)。

(54) Title: MEDIA DATA PROCESSING METHOD AND RELATED DEVICE

(54) 发明名称: 一种媒体数据的处理方法及相关设备

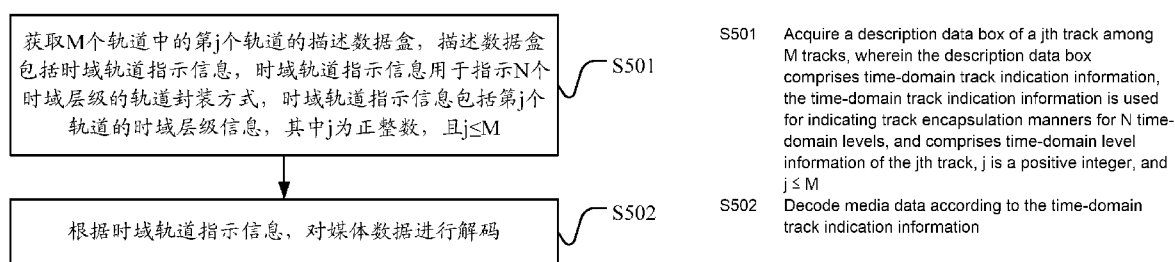


图 5

(57) Abstract: Provided in the embodiments of the present application are a media data processing method and a related device. The method comprises: acquiring a description data box of a jth track among M tracks, wherein the description data box comprises time-domain track indication information, the time-domain track indication information is used for indicating track encapsulation manners for N time-domain levels, and comprises time-domain level information of a time-domain level encapsulated in the jth track, j is a positive integer, and $j \leq M$; and decoding media data according to the time-domain track indication information. The flexibility of performing multi-track encapsulation on different time-domain levels can thus be improved.

(57) 摘要: 本申请实施例提供一种媒体数据的处理方法及相关设备, 其中的方法包括: 获取所述M个轨道中的第j个轨道的描述数据盒, 所述描述数据盒包括时域轨道指示信息, 所述时域轨道指示信息用于指示所述N个时域层级的轨道封装方式, 所述时域轨道指示信息包括所述第j个轨道封装的时域层级的时域层级信息, 其中j为正整数, 且 $j \leq M$; 根据所述时域轨道指示信息, 对所述媒体数据进行解码, 可提升对不同时域层级进行多轨封装时的灵活性。

本国际公布：

— 包括国际检索报告(条约第21条(3))。

一种媒体数据的处理方法及相关设备

本申请要求于2021年06月11日提交中国专利局、申请号为202110656768.4、申请名称为“一种媒体数据的处理方法及相关设备”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

本申请涉及计算机技术领域，尤其涉及媒体数据的处理。

背景技术

10 在相关的视频编码技术中，时域层级划分的技术得到了支持，且在系统层封装技术中，针对轨道内的不同时域层级也进行了指示。然而，一些系统层的封装技术中，仅考虑了使用一个轨道封装视频位流的情景。

发明内容

本申请实施例提供媒体数据的处理方法及相关设备，可提升对不同时域层级进行封装时的灵活性。

15 一方面，本申请实施例提供一种媒体数据的处理方法，所述媒体数据包括多个媒体帧，所述多个媒体帧被划分为N个时域层级，所述多个媒体帧被封装到M个轨道中，其中M和N均为大于1的正整数；包括：

获取所述M个轨道中的第j个轨道的描述数据盒，所述描述数据盒包括时域轨道指示信息，所述时域轨道指示信息用于指示所述N个时域层级的轨道封装方式，所述时域轨道指示信息包括所述第j个轨道封装的时域层级的时域层级信息，其中j为正整数，且 $j \leq M$ ；

20 根据所述时域轨道指示信息，对所述媒体数据进行解码。

一方面，本申请实施例提供一种媒体数据的处理方法，包括：

根据媒体数据包括的每个媒体帧的帧间依赖关系，确定所述每个媒体帧的时域层级，得到N个时域层级的媒体帧；其中，N为大于1的正整数；

25 分别将所述N个时域层级的媒体帧封装到M个轨道中，并生成对应的描述数据盒；所述M个轨道中第j个轨道的描述数据盒包括时域轨道指示信息，所述时域轨道指示信息用于指示所述N个时域层级的轨道封装方式，所述时域轨道指示信息包括所述第j个轨道封装的时域层级的时域层级信息，其中，M为大于1的正整数。

30 一方面，本申请实施例提供一种媒体数据的处理装置，所述媒体数据包括多个媒体帧，所述多个媒体帧被划分为N个时域层级，所述多个媒体帧被封装到M个轨道中，其中M和N均为大于1的正整数；所述装置包括：

获取单元，用于获取所述M个轨道中的第j个轨道的描述数据盒，所述描述数据盒包括时域轨道指示信息，所述时域轨道指示信息用于指示所述N个时域层级的轨道封装方式，所述时域轨道指示信息包括所述第j个轨道封装的时域层级的时域层级信息，其中j为正整数，且 $j \leq M$ ；

35 处理单元，用于根据所述时域轨道指示信息，对所述媒体数据进行解码。

一方面，本申请实施例提供另一种媒体数据的处理装置，包括：

确定单元，用于根据媒体数据包括的每个媒体帧的帧间依赖关系，确定所述每个媒体

帧的时域层级，得到N个时域层级的媒体帧；其中，N为大于1的正整数；

处理单元，用于分别将所述N个时域层级的媒体帧封装到M个轨道中，并生成对应的描述数据盒；所述描述数据盒包括时域轨道指示信息，所述时域轨道指示信息用于指示所述N个时域层级的轨道封装方式，所述时域轨道指示信息包括所述第j个轨道封装的时域层级的时域层级信息，其中，M为大于1的正整数。

一方面，本申请实施例提供一种计算机设备，包括：

处理器，适于实现一条或多条指令；以及，

存储器，存储有一条或多条指令，一条或多条指令适于由处理器加载并执行上述方面的媒体数据的处理方法。

一方面，本申请实施例提供了一种计算机可读存储介质，所述计算机可读存储介质存储有计算机程序，所述计算机程序由于执行上述方面的媒体数据的处理方法。

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媒体数据的处理方法。

在本申请实施例中，支持内容生成设备将多时域层级的媒体帧封装到不同的轨道，并在封装过程中，将在每个轨道的描述数据盒中记录时域轨道指示信息，从而实现关联不同时域层级的轨道，指示轨道内的时域层级信息，标注基准轨道，指示多时域层级轨道相互组合的策略信息。那么，内容消费设备则可根据轨道内的描述数据盒选取合适的时域层级，将不同轨道的样本组合后进行解码呈现，从而保证了多轨道封装的灵活性，并最大化地节省了解码计算资源。

附图说明

图1示出了本申请一个示例性实施例提供的一种视频处理的流程图；

图2示出了本申请一个示例性实施例提供的针对媒体数据的处理方法的示意图；

图3示出了本申请一个示例性实施例提供的编码单元的示意图；

图4示出了本申请一个示例性实施例提供的时域层级的划分的示意图；

图5示出了本申请一个示例性实施例提供的一种媒体数据的处理方法的示意图；

图6a示出了本申请一个示例性实施例提供的一种多时域的多轨道封装方式的示意图；

图6b示出了本申请一个示例性实施例提供的一种多时域的多轨道封装方式的示意图；

图7示出了本申请一个示例性实施例提供的一种媒体数据的处理方法的示意图；

图8示出了本申请一个示例性实施例提供的一种媒体数据的处理装置的结构示意图；

图9示出了本申请一个示例性实施例提供的一种媒体数据的处理装置的结构示意图；

图10示出了本申请一个示例性实施例提供的一种计算机设备的结构示意图。

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，显然，所描述的实施例仅仅是本申请一部分实施例，而不是全部的实施例。基于本申请中的实施例，本领域普通技术人员在没有作出创造性劳动前提下所获得的所有其他实施例，都属于本申请保护的范围。

本申请实施例涉及媒体数据的处理技术。其中，媒体数据（或称之为多媒体数据）是

指由内容上相互关联的文本、图形、图像、声音、动画、活动图像等媒体的数据所形成的复合数据，而在本申请实施例所提及的媒体数据，主要包括由声音构成的音频数据，以及由图像和声音等构成的视频数据，且本申请实施例中，主要以媒体数据为视频数据为例，对媒体数据的数据处理过程进行详细说明，而在媒体数据为音频数据时，可参见本申请实

5 施例。本申请实施例涉及的对媒体数据的处理过程，主要包括媒体数据采集，媒体数据编码，媒体数据文件封装，媒体数据文件传输，媒体数据解码和最终的数据呈现，而在该媒体数据为视频数据时，那么，针对视频数据的完整处理过程可如图1所示，具体可包括：视频采集，视频编码，视频文件封装，视频传输，视频文件解封装，视频解码和最终的视频呈现。

10 视频采集用于将模拟视频转换为数字视频，并按数字视频文件的格式进行保存，也就是说，视频采集可将视频信号转换为二进制数字信息，其中，由视频信号转换为的二进制信息是一种二进制数据流，该二进制信息也可称为该视频信号的码流或者位流(Bitstream)，视频编码则是通过压缩技术，将原始视频格式的文件转换为另一种视频格式文件。在本申请实施例中提及视频媒体内容的生成，包括摄像机采集生成的真实场景，以及计算机生

15 成的屏幕内容场景，而从视频信号的获取方式看，视频信号可以分为摄像机拍摄到的以及计算机生成的两种方式，由于统计特性的不同，其对应的压缩编码方式也可能有所区别，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HEVC/H.265), VVC(versatile video coding, 国际视频编码标准VVC/H.266), 以及中国国家视频编码标准AVS (Audio Video Coding Standard, 中国国家视频编码标准

20 AVS), 或AVS3 (由AVS标准组推出的第三代视频编码标准) 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，具体可如图2所示：

①块划分结构 (block partition structure)：输入图像 (如视频数据中的一个媒体帧) 根据一个的大小，划分成若干个不重叠的处理单元，每个处理单元将进行类似的压缩操作。这个处理单元被称作CTU (Coding Tree Unit, 编码树单元)，或者LCU (Largest Coding Unit, 最大编码单元)。其中，编码树单元一般由最大编码单元开始往下划分，CTU再往下，可以

25 继续进行更加精细的划分，得到一个或多个基本编码的单元，称之为CU (Coding Unit, 编码单元)。每个CU是一个编码环节中最基本的元素。以下描述的是对每一个CU可能采用的各种编码方式，其中，LCU (或CTU) 和CU之间的关系可如图3所示。

②预测编码 (Predictive Coding)：包括了帧内预测和帧间预测等方式，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。编码端需要为当前CU决定在众

30 多可能的预测编码模式中，选择最适合的一种，并告知解码端。

a. 帧内预测 (Intra (picture) Prediction)：预测的信号来自于同一图像内已经编码重建过的区域。

b. 帧间预测 (Inter (picture) Prediction)：预测的信号来自已经编码过的，不同于当前图

35 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DFT(Discrete Fourier Transform, 离散傅里叶变换)，DCT (离散余弦变换，是DFT的一个子集) 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，编码端也需要为当前编码CU选择其中的一种变换，并告知解码端。量化的精细程度通常由量化参数
5 (Quantization Parameter, QP) 来决定，QP取值较大大，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
10 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)，SAO (Sample Adaptive Offset, 样点自适应补偿) 或者ALF (Adaptive Loop Filter, 自适应环路滤波) 等滤波器，可以有效的降低量化所
15 产生的失真程度。由于这些经过滤波后的重建图像，将做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。

图2中展示了视频编码器的基本流程，图2中以第k个CU (标记为 $S_k[x, y]$) 为例进行举例说明，其中，k为大于等于1且小于等于输入的当前图像中的CU的数量的正整数， $S_k[x, y]$ 表示第k个CU中坐标为 $[x, y]$ 的像素点，x表示像素点的横坐标，y表示像素点的中坐标， $S_k[x, y]$ 经过运动补偿或者帧内预测等中的一种较优处理后获得预测信号 $\widehat{S}_k[x, y]$ ， $S_k[x, y]$ 与 $\widehat{S}_k[x, y]$ 相减得到残差信号 $U_k[x, y]$ ，然后对该残差信号 $U_k[x, y]$ 进行变换和量化，量化输出的数据由两个不同的去处：一是送个熵编码器进行熵编码，编码后的码流输出到一个缓冲器 (buffer) 中保存，等待传出去；另一个应用是进行反量化和反变换后，得到信号
25 $U'_k[x, y]$ 。将信号 $U'_k[x, y]$ 与 $\widehat{S}_k[x, y]$ 相加得到新的预测信号 $S^*_k[x, y]$ ，并将 $S^*_k[x, y]$ 送到当前图像的缓冲器中保存。 $S^*_k[x, y]$ 经过帧内一图像预测获得 $f(S^*_k[x, y])$ ， $S^*_k[x, y]$ 经过环路滤波后获得 $S'_k[x, y]$ ，并将 $S'_k[x, y]$ 送至解码图像缓冲器中保存，以用于生成重建视频。 $S'_k[x, y]$ 经过运动一补偿预测后获得 $S'_l[x+m_x, y+m_y]$ ， $S'_l[x+m_x, y+m_y]$ 表示参考块， m_x 和 m_y 分别表示运动矢量的水平和竖直分量。

在对媒体数据进行编码后，则需要对编码后的数据流进行封装并传输给用户，视频文件封装是指按照封装格式 (或容器，或文件容器)，将已经编码压缩好的视频和音频按照一定的格式存放在一个文件中，常见的封装格式包括AVI格式 (Audio Video Interleaved, 音频
35

视频交错格式)或者ISOBMFF(ISO Based Media File Format, 基于ISO(International Standard Organization, 国际标准化组织)标准的媒体文件格式), 其中, ISOBMFF是媒体文件的封装标准, 最典型的ISOBMFF文件即MP4(Moving Picture Experts Group 4, 动态图像专家组4)文件, 其中, 本申请实施例的主要改进点也是针对ISOBMFF数据盒的。在一个实施例中, 将音频码流和视频码流按照如ISOBMFF的文件格式封装在文件容器中形成封装文件, 在该封装文件中, 以样本(sample)为文件封装过程中的封装单位, 一个封装文件由多个样本组成。也就是说, 在媒体文件的封装过程中, 通常将一个媒体帧作为一个样本进行封装, 从而生成封装文件, 而在该媒体数据为视频媒体时, 该媒体帧即为媒体帧, 而在该媒体数据为音频媒体时, 该媒体帧即为音频帧, 也即, 针对视频媒体的封装文件包括多个媒体帧, 而音频媒体的封装文件则包括多个音频帧, 在本申请实施例中, 以媒体数据为视频媒体为了, 并以封装文件中的一个样本为视频媒体中的一个媒体帧为例进行说明。

封装后的文件将通过视频传输到用户终端, 用户终端则可在进行解封装, 解码等逆操作后, 在用户终端中进行最终视频内容的呈现。其中, 封装后的文件可通过传输协议发送到用户终端, 该传输协议例如可以是DASH(dynamic adaptive streaming over HTTP, 基于HTTP的动态自适应流是一种自适应比特率流技术), 采用DASH进行传输可以使高质量流媒体可以通过传统的HTTP网络服务器以互联网传递, 在DASH中, 用MPD(media presentation description, DASH中的媒体演示描述信令)描述媒体片段信息, 且DASH中, 一个或多个媒体成分的组合, 比如某种分辨率的视频文件可以看做一个Representation(代表), 而包含的多个Representation可看做是一个Adaptation Set(一个视频流的集合), 一个DASH可包含一个或多个Adaptation Set。

可以理解, 用户终端的文件解封装的过程与上述的文件封装过程是相逆的, 用户终端可按照封装时的文件格式要求对封装文件进行解封装, 得到音频码流和视频码流。用户终端的解码过程也与编码过程是相逆的, 该用户终端可对音频码流解码, 还原出音频内容。根据上述编码过程可以看出, 在解码端, 对于每一个CU, 解码器获得压缩码流后, 先进行熵解码, 获得各种模式信息及量化后的变换系数。各个系数经过反量化及反变换, 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。

在视频编码技术中, 还涉及一种时域分层技术, 该技术可将不同的媒体帧按照解码时的依赖关系, 划分为不同的时域层级, 具体来说, 采用该时域分层技术进行时域层级的划分, 被划分为低层级的媒体帧, 在解码时无需参考更高层级的媒体帧, 如图4所示, 箭头表示解码时的依赖关系, 从 I_0 帧至 B_1 帧的箭头表示 B_1 帧在解码时需要参考 I_0 帧进行解码, 即 B_1 帧的解码必须依赖 I_0 帧的解码进行, 其余帧之间的关系以此类推。由图4中箭头指示的帧间依赖关系可知, 所有媒体帧根据帧间依赖关系被划分为 $L_0 \sim L_3$ 四个时域层级, 属于每个时域层级的媒体帧在解码时都不依赖更高层级的媒体帧。其中, 需要说明的是, 本申请实施例所提及的时域层级的低和高是一个相对的概念, 如图4中确定的 $L_0 \sim L_3$ 这四个时域层级, 对于 L_0 时域层级而言, $L_1 \sim L_3$ 均为高时域层级, 而对于 L_1 时域层级而言, L_3 时域层级为 L_1

的高时域层级，而L0时域层级则为L1的低时域层级。

如图4所示，媒体帧的类型主要包括I帧（Intra Slice，帧内条带）、B帧和P帧，其中，I帧也被称为关键帧，属于帧内压缩，在解码时仅需参考I帧其本身的信息即可，B帧为双向预测编码帧，在解码时即需要参考前面已有的帧，又需要参考后面待解码的帧，而P帧为前向预测编码帧，即P帧在解码时需要参考前面相关帧的信息才能解码，而在图4中针对I帧、B帧和P帧下添加的阿拉伯数字下标用于表示其各自所处的对应的时域层级。可以理解，基于I帧、P帧和B帧这三类媒体帧在解码时的特性，由于要使得进行时域层级划分后，在属于各时域层级的媒体帧中，属于低时域层级的媒体帧在进行解码时不依赖高时域层级，那么也就可以理解，最低时域层级（如上述的L0时域层级）中的媒体帧在解码时，将不依赖于属于其他任何时域层级的，也即属于最低时域层级的媒体帧可进行独立解码显示，那么，被划分到最低时域层级的媒体帧必然包括I帧。

由于在对媒体帧进行时域层级划分时，属于低时域层级的媒体帧在解码时无需参考高时域层级的媒体帧，如图4所示，假设视频数据中的媒体帧一个包括L0~L3这四个时域层级，且图4中的箭头用于表示各媒体帧在解码时的依赖关系，也就是说，从I₀帧至B₁帧的箭头表示，处于L1时域层级的B₁帧在解码时需要参考处于L0时域层级的I₀帧，处于L1时域层级的B₁帧在解码时需要参考处于L0时域层级的P₀帧，处于L2时域层级的第一个B₂帧在解码时需要参考处于L0时域层级的I₀帧，以及处于L1时域层级的B₁帧，处于L2时域层级的第二个B₂帧在解码时需要参考处于L1时域层级的B₁帧，以及处于L0时域层级的P₀帧，处于L3时域层级的第一个B₃帧在解码时需要参考处于L2时域层级的第一个B₂帧，以及处于L0时域层级的I₀帧，处于L3时域层级的第二个B₃帧在解码时需要参考处于L2时域层级的第一个B₂帧，以及处于L1时域层级的B₁帧，处于L3时域层级的第三个B₃帧在解码时需要参考处于L1时域层级的B₁帧，以及处于L2时域层级的第二个B₂帧，处于L3时域层级的第四个B₃帧在解码时需要参考处于L2时域层级的第二个B₂帧，以及处于L0时域层级的P₀帧。

现有的AVS3视频编码技术，可支持时域层级划分技术（或称之为时域分层技术），同时在AVS3系统层封装技术中，针对轨道内的不同时域层级也进行了指示，采用现有的支持时域分层技术，在对媒体帧进行封装时，可通过封装文件中的时域层数字段（temporal_layer_num）指示封装各媒体帧的轨道对应视频码流中的时域层数，此外，还将通过时域层级标识字段（temporal_layer_id）指示封装目标媒体帧的轨道对应的视频码流中各媒体帧所属的时域层级。其中，轨道是指一系列有时间属性的按照ISO基本媒体文件格式（ISO base media file format, ISO BMFF）的封装方式的样本，比如视频track，视频track是通过将视频编码器编码每一帧后产生的码流按照ISO BMFF的规范封装后得到的。现有的AVS3解码器配置信息（即描述数据盒）记录给出了针对AVS3编码方式的解码器配置信息，该解码器配置信息可采用配置信息1进行表示，该配置信息1具体如下：

```
class Avs3DecoderConfigurationRecord{//AVS3解码器配置记录
    unsigned int(8) configurationVersion; //8位无符号整数的配置版本字段
    unsigned int(8) profile_id; //简介标识符
    unsigned int(8) level_id; //水平标识符
```

bit(6) reserved = '111111'b; //保留字段, 一般字段需要整数个byte, 所以需要保留的
bit (位) 来补足

unsigned int(2) chroma_format; //色度格式

bit(5) reserved = '11111'b; //

5 unsigned int(3) encoding_precision; //编码精度

bit(4) reserved = '1111'b; //

unsigned int(4) frame_rate_code; //编码帧率

bit(6) reserved = '111111'b; //

unsigned int(2) library_indication; //库指示

10 bit(4) reserved = '11111'b; //

unsigned int(3) temporal_layer_num; //时域层数字段

}

由上述可知, 虽然在AVS3视频编码技术中, 时域层级划分的技术得到了支持, 同时在
AVS3系统层封装技术中, 针对轨道内的不同时域层级也进行了指示。然而, 当前系统层的
15 封装技术中, 仅考虑了使用一个轨道封装视频位流的情景。若用户将一个包含不同时域层
级的视频位流封装到多个视频文件轨道, 则现有技术难以提供足够的信息以支持用户有选
择性地通过不同文件轨道重构视频位流并消费。

基于此, 本申请实施例提出的媒体数据的处理方法, 可支持针对AVS3编码中时域层级
划分技术的多轨道文件封装, 在该媒体数据为视频数据时, 具体实施步骤如下:

20 1、在视频编码环节, 内容生成设备可根据视频数据的各媒体帧之间的帧间依赖关系,
为不同媒体帧确定时域层级;

2、根据媒体帧的时域层级, 将视频位流封装为多个不同的轨道, 并在每个文件轨道中
指示该轨道包含的具体时域层级信息, 包括时域层级id、时域层级对应的帧率和码率信息
等;

25 3、对于包含最低时域层级的轨道, 将其标识为基准轨道, 其余高时域层级的轨道通
过'tlrf'索引至该基准轨道。同时, 在基准轨道的文件封装中, 给出多时域层级相互组合的
策略信息;

4、在用户消费侧, 用户所在的内容消费设备可根据自身设备能力以及多个时域层级相
互组合的策略信息, 选择一个或多个所需时域层级对应的轨道, 解封装这些不同的轨道并
30 重构为一个位流进行解码。最终达到灵活选择文件轨道, 节省解码计算资源的目的。

为支持上述步骤, 本申请实施例在系统层添加了若干描述性字段, 以扩展现有
ISOBMFF数据盒的形式举例, 定义了相关的字段以支持AVS3时域层级划分的多轨道文件封
装技术, 下面结合图5, 对本申请实施例提出的媒体数据的处理方法进行详细说明, 其中,
该媒体数据的处理方法可由任意进行媒体内容消费的内容消费设备执行, 可以理解, 该内
容消费设备包括进行媒体内容消费的用户所在的终端设备(或服务器), 其中, 该媒体数据
35 包括多个媒体帧, 所述多个媒体帧被划分为N个时域层级, 所述多个媒体帧被封装到M个轨
道中, M和N均为大于1的正整数。

如图5所示,该方法具体可包括:

S501,获取M个轨道中的第j个轨道的描述数据盒,描述数据盒包括时域轨道指示信息。

其中,时域轨道指示信息用于指示N个时域层级的轨道封装方式,时域轨道指示信息包括第j个轨道封装的时域层级的时域层级信息,其中j为正整数,且 $j \leq M$ 。

5 首先,内容消费设备获取的描述数据盒是基于内容生成设备对媒体数据的编码封装生成的,内容生成设备在对媒体数据进行封装时,可基于各媒体帧所属的时域层级,将属于不同层级时域的媒体帧封装到多个不同的轨道中,其中,该媒体数据包括多个媒体帧,多个媒体帧被划分为N个时域层级,那么,该多个媒体帧可被封装到M个轨道中,其中M和N均为正整数。需要说明的是,封装该多个媒体帧的轨道,根据在轨道中所封装的媒体帧在解码时的特征,可将该M个轨道划分为基准轨道和非基准轨道,其中,基准轨道是指该轨道中所封装的媒体帧能被独立解码,即基准轨道中封装的媒体帧在解码时,不会参考其他任何轨道中的媒体帧,那么可以理解,在媒体数据为视频数据时,该基准轨道中封装的媒体帧必然包括I帧,那么综合上述的,被分到最低时域层级的媒体帧也必然包括I帧,即是说,基准轨道中封装的时域层级必然包括最低时域层级。

15 也就是说,如果媒体数据是视频数据,内容生成设备在需要将视频数据发送到用户侧进行消费显示时,可先在视频编码时,根据各个媒体帧的帧间依赖关系,为不同媒体帧确定时域层级,进而可根据时域层级,将视频位流封装到多个不同的轨道,并通过描述数据盒在每个文件轨道中指示具体的时域层级信息,那么相应的,在用户侧进行消费时,则可基于自身终端设备的解码能力进行时域层级的选取,从而可节省解码计算资源。在一个实施例中,若一个媒体数据包括的多个媒体帧分别属于N个时域层级,那么,内容生成设备在将属于N个时域层级的媒体帧封装到一个或多个轨道后,将在每个轨道生成相应的描述数据盒,以使得内容消费设备(如用户侧的终端设备)可基于该描述数据盒的记录,确定内容生成设备对属于N个时域层级的媒体帧的封装方式,并进一步地选择合适的时域层级的媒体帧进行解码显示。

25 在本申请实施例中,该描述数据盒是通过在现有的ISOBMFF数据盒中添加时域轨道信息,从而实现的支持VAS3时域层级划分的多轨道文件封装技术,可以理解,在该ISOBMFF数据盒中添加时域轨道信息包括在该ISOBMFF数据盒中扩展添加的一个或多个相关字段。在本申请实施例中,以该M个轨道中第j个轨道的描述数据盒为例,对在描述数据盒中扩展的相关字段进行详细说明。其中,在该描述数据盒中扩展的相关字段可如配置信息2所示,该配置信息2具体如下:

```
class Avs3DecoderConfigurationRecord{
    ...
    unsigned bit(1) multi_temporal_track_flag;
    if(temporal_layer_num > 1 || multi_temporal_track_flag == 1){
35 TemporalLayerInfoBox(temporal_layer_num, multi_temporal_track_flag);
    }
}
```

```
class TemporalLayerInfoBox extends FullBox('tlin', temporal_layer_num,
multi_temporal_track_flag) {
for(i=0; i<temporal_layer_num; i++){
5   unsigned int(8) temporal_layer_id[i];
}
if(multi_temporal_track_flag == 1){
unsigned int(1) base_track_flag;
bit(7) reserved;
if(base_track_flag == 1){
10   unsigned int(8) total_temporal_layer_num;
for(i=0; i< total_temporal_layer_num; i++){
unsigned int(3) temporal_layer_id[i];
unsigned int(4) frame_rate_code[i];
unsigned int(18) temporal_bit_rate_lower[i];
15   unsigned int(12) temporal_bit_rate_upper[i];
bit(3) reserved;
}
unsigned int(8) total_temporal_track_num;
for(i=0; i< total_temporal_track_num; i++){
20   unsigned int(32) track_ID [i];
unsigned int(1) is_output_track_flag[i];
unsigned int(1) is_alternative_track_flag [i];
if(is_alternative_track_flag [i]){
unsigned int(32) alternate_track_ID;
25   }
bit(6) reserved;
}
}
30 }
else{
for(i=0; i< temporal_layer_num; i++){
unsigned int(3) temporal_layer_id[i];
unsigned int(4) frame_rate_code[i];
35   unsigned int(18) temporal_bit_rate_lower[i];
unsigned int(12) temporal_bit_rate_upper[i];
bit(3) reserved;
```

```
}  
}  
}
```

其中，该描述数据盒中包括的时域轨道指示信息所包括的字段分别为上述配置信息2
5 中的多时域轨道标识字段（multi_temporal_track_flag）、总时域层数字段
（total_temporal_layer_num）、时域层数字段（时域层数字段）、时域层级标识字段
（temporal_layer_id[i]）、帧率字段（frame_rate_code[i]）、低码率字段
（temporal_bit_rate_lower[i]）、高码率字段（temporal_bit_rate_upper[i]）、基准轨道标识字段
（base_track_flag）、轨道标识字段（track_ID[i]）、优先解码呈现字段（is_output_track_flag[i]）、
10 替代轨道字段（is_alternative_track_flag[i]），以及替代轨道标识字段（alternate_track_ID）。
其中，上述提及的字段中的时域层数字段（temporal_layer_id[i]）、帧率字段
（frame_rate_code[i]）、低码率字段（temporal_bit_rate_lower[i]）以及高码率字段
（temporal_bit_rate_upper[i]）用于指示相应轨道（如上述的第j个轨道）中具体的时域层级
信息。

15 下面，将对配置信息2中涉及的语义和语法进行详细说明：

（1）多时域轨道标识字段（multi_temporal_track_flag）用于指示媒体数据的N个时域
层级的轨道封装方式，该轨道封装方式包括：多轨道封装方式和单轨道封装方式，其中，
在多时域轨道标识字段为第一数值时，多时域轨道标识字段用于指示属于N个时域层级的
多个媒体帧被封装到多个不同的轨道中，而在当多时域轨道标识字段为第二数值时，多时
20 域轨道标识字段用于指示属于N个时域层级的多个媒体帧被封装到单个轨道中，具体地，
该第一数值可以为1，而该第二数值则可以为0。

（2）时域层数字段（temporal_layer_num）用于指示当前轨道（即上述的第j个轨道）
包含的时域层级的数量。如配置信息2所示，在时域层数字段的取值大于1，即该第j个轨道
封装了该时域层数字段的取值数量的多个时域层级，或者，在多时域轨道标识字段的取值
25 为1，即该媒体数据的1个时域层级被封装到第j个轨道，内容消费设备在解码时，可进一步
从该描述数据盒中读取相关字段的取值，从而根据各相关字段的取值进行解码显示，其中，
该描述数据盒是一个为‘tlin’类型的数据盒。

（3）该第j个轨道中的描述数据盒中包括了该第j个轨道具体的时域层级信息，该时域
层级信息包括时域层级标识字段（temporal_layer_id[i]），该时域层级字段用于指示单个时
30 域层级的ID（Identity document，一种唯一的身份标识），其中，N个时域层级中的一个时域
层级对应一个temporal_layer_id，也就是说，该时域层级标识字段可用于指示第j个轨道封装
的时域层级中，第i个时域层级的层级标识。结合配置信息2可知，在该第j个轨道的描述数
据盒中，第j个轨道封装的时域层级中第i个时域层级的层级标识将被记录在
temporal_layer_id[i]中。

35 此外，该第j个轨道中的时域层级信息还包括帧率字段（frame_rate_code[i]）和码率，
该帧率字段用于指示累计到属于第i个时域层级的媒体帧（即时域层级等于
temporal_layer_id[i]）时的帧率，该码率信息用于指示累计到属于第i个时域层级的媒体帧（即

时域层级等于 $\text{temporal_layer_id}[i]$) 时的码率, 该码率信息包括低码率字段 ($\text{temporal_bit_rate_lower}[i]$), 该低码率字段用于指示累计到属于第 i 个时域层级的媒体帧时的码率的低18位, 此外, 该码率信息还包括高码率字段 ($\text{temporal_bit_rate_upper}[i]$), 该高码率字段用于指示累计到属于第 i 个时域层级的媒体帧时的码率的高12位。其中, 累计是指假设 $\text{temporal_layer_id}[i] = 3$, 那么, 在 $\text{temporal_layer_id}[i] = 3$ 时, 其对应的帧率 (和码率) 是 $\text{temporal_layer_id}[i]$ 分别取到小于3的所有帧的帧率 (和码率)。

如上述的配置信息2所示, 如果描述数据盒中多时域轨道标识字段的取值为0, 即指示该媒体数据中属于 N 个时域层级的多个媒体帧被封装到单个轨道时, 内容消费设备可通过读取时域层级标识字段, 帧率字段, 低码率字段, 以及高码率字段中的取值, 确定累计到各时域层级时对应的帧率和码率, 从而该内容消费设备可结合自身的解码性能, 选取属于部分或全部时域层级的媒体帧进行解码显示。

(4) 在多时域轨道标识字段的取值为1时, 说明该媒体数据中属于 N 个时域层级的多个媒体帧被封装到多个不同的轨道中, 在属于 N 个时域层级的多个媒体帧被封装到多个不同的轨道的情况下, 内容消费设备在读取帧率字段, 低码率字段, 以及高码率字段中的取值, 以确定累计到各时域层级对应的帧率和码率之前, 内容消费设备还需要读取一些其他字段的值, 其中, 内容消费设备还需要读取的这些字段包括如下的①-④中提及的部分或全部字段:

①时域轨道指示信息包括的基准轨道标识字段 (base_track_flag)。

基准轨道标识字段用于指示第 j 个轨道是否为基准轨道; 当基准轨道标识字段为第一数值时, 基准轨道标识字段用于指示第 j 个轨道为基准轨道, 当基准轨道标识字段为第二数值时, 基准轨道标识字段用于指示第 j 个轨道为非基准轨道; 其中, 基准轨道中封装的媒体帧被独立解码, 其中, 该第一数值可以为1, 该第二数值可以为0。由上述的配置信息2可知, 在内容生成设备对属于 N 个时域层级的媒体帧进行多轨封装时, 累计到属于各个时域层级的媒体帧时的帧率和码率是被记录在基准轨道中的, 也就是说, 只有在基准轨道的描述数据盒中的帧率字段、低码率字段和高码率字段存在取值, 而在非基准轨道的描述数据盒中, 该帧率字段、低码率字段和高码率字段为空。

在一个实施例中, 在内容消费设备从第 j 个轨道的描述数据盒中, 读取到基准轨道标识字段的取值为1, 那么, 说明该第 j 个轨道为基准轨道, 进而, 内容消费设备还可从该第 j 个轨道的描述数据盒中读取帧率字段、低码率字段和低码率字段的取值, 从而确定出各时域层级被累计时所对应的帧率和码率。

②总时域层数字段 ($\text{total_temporal_layer_num}$)。

总时域层数字段用于指示当前文件包含的所有轨道对应的时域层级总数, 即用于指示在 M 个轨道中所封装的时域层级的总数量。在多个时域层级采用多轨封装的方式时, 如果第 j 个轨道为基准轨道, 那么, 内容生成设备在确定基准轨道的描述数据盒中时域层级标识字段 ($\text{temporal_layer_id}[i]$) 的取值时, 将基于总时域层数字段, 在基准轨道的描述数据盒中记录每个时域层级的标识。

③索引类型标识字段 (tlrf)。

索引类型标识字段用于在多时域层级使用多轨道封装时，定义基准轨道（或基准时域层级轨道）和非基准轨道（或高时域层级轨道）之间的索引关系，其中，基准时域层级轨道为包含最低的时域层级ID的轨道，基准时域层级轨道在一个文件中仅有一个，其余包含各时域层级的轨道均为高时域层级的轨道。

5 高时域层级轨道应通过轨道索引数据盒（TrackReferenceBox）索引至其解码所依赖的基准时域层级轨道。而在该高时域层级轨道的TrackReferenceBox中应添加对应的轨道索引类型数据盒（TrackReferenceTypeBoxes），其中，TrackReferenceTypeBoxes数据盒中通过track_IDs指示当前的基准轨道（或称之为基准时域层级轨道），其中，该非基准轨道和该基准轨道之间的索引通过TrackReferenceTypeBoxes中对应的索引类型标识字段
10 （reference_type）标识，该类型字段定义为：

'tlrf': 被索引的轨道为基准时域层级轨道。

也就是说，如果内容消费设备从第j个轨道的描述数据盒中，读取到基准轨道标识字段的取值为0，则说明该第j个轨道为非基准轨道，而在该第j个轨道为非基准轨道，则第j个轨道还包括轨道索引数据盒，轨道索引数据盒包括轨道索引类型数据盒；轨道索引类型数据盒包括轨道标识字段和索引类型标识字段；轨道标识字段用于存储基准轨道的标识，索引类型标识字段用于指示被索引的轨道为基准轨道。也就可以理解，在第j个轨道为非基准轨道时，由于在非基准轨道的描述数据盒中，帧率字段、低码率字段以及高码率字段的取值均为空，所以，在第j个轨道为非基准轨道的情况下，内容消费设备通过第j个轨道的描述数据盒将无法读取到帧率字段、低码率字段以及高码率字段的取值，进而无法确定出累计到
15 每个时域层级的帧率和码率，那么，在这种情况下，内容消费设备可通过第j个轨道中的轨道索引数据盒所包括的轨道索引类型数据盒中的索引类型标识字段，从非基准轨道（即该第j个轨道）索引到基准轨道中，并基准轨道中读取上述的帧率字段、低码率字段和高码率字段的取值。

④时域轨道指示信息包括的轨道组合策略信息，该轨道组合策略信息包括轨道标识字段（track_ID [i]），优先解码呈现字段（is_output_track_flag[i]），替代轨道字段（is_alternative_track_flag [i]），以及替代轨道标识字段（alternate_track_ID）。

其中，轨道标识字段用于指示包含部分时域层级的轨道的标识（ID），该M个轨道中的一个轨道对应一个track_ID。

优先解码呈现字段用于指示第j个轨道（即当前轨道）是否为优先解码呈现的轨道；当
30 优先解码呈现字段为第一数值时，优先解码呈现字段用于指示第j个轨道为优先解码呈现的轨道；当优先解码呈现字段为第二数值时，优先解码呈现字段用于指示第j个轨道不是优先解码呈现的轨道；其中，该第一数值可以为1，该第二数值可以为0。

替代轨道字段用于指示第j个轨道（即当前轨道）是否为M个轨道中的一个轨道的替代轨道；当替代轨道字段为第一数值时，替代轨道字段用于指示第j个轨道为M个轨道中的一个轨道的替代轨道；当替代轨道字段为第二数值时，替代轨道字段用于指示第j个轨道不是
35 替代轨道；其中，该第一数值可以为1，该第二数值可以为0。

该替代轨道标识字段用于指示所述第j个轨道（即当前轨道）替代的一个轨道的标识。

综合上述对配置信息2中的语法和语义的说明，在媒体数据中包括的多个媒体帧属于N个不同的时域层级时，内容生成设备在对媒体数据所包括的多个媒体帧进行封装时可分为以下两种情况：

第一种情况，可将属于N个时域层级的媒体帧封装到单个轨道中。那么内容消费设备在获取到针对媒体数据的封装文件后，就可从封装分别属于N个时域层级的媒体帧的单个轨道的描述数据盒中，通过读取层级标识字段，帧率字段，低码率字段和高码率字段，确定出该N个时域层级中，第i个时域层级的层级标识，以及相应的帧率和码率，进而使得内容消费设备可结合自身的解码能力，选取部分或全部的时域层级的媒体帧进行解码，即可转而执行S502。

第二种情况，可将属于N个时域层级的媒体帧封装到多个不同的轨道中，内容生成设备在采用多轨封装的方式进行媒体帧的封装时，将在基准轨道中记录各轨道中封装的各时域层级相互组合的策略，以及累计到各时域层级对应的媒体帧的帧率和码率，并将其他轨道通过索引类型标识字段索引到基准轨道中，那么内容消费设备就可通过基准轨道记录的相关信息，并结合自身的解码能力，选择部分或全部的媒体帧进行解码，即转而执行S502。

S502，根据时域轨道指示信息，对媒体数据进行解码。

在内容消费设备获取到第j个轨道中的描述数据盒后，将根据该描述数据盒中的时域轨道指示信息，对媒体数据进行解码，具体地，内容消费设备可根据时域轨道指示信息及解码设备的解码性能，保留N个时域层级中与解码性能匹配的时域层级，并对保留的时域层级的媒体帧进行解码显示。在内容消费设备根据时域轨道指示信息及解码设备的解码性能，保留N个时域层级中与解码性能匹配的时域层级时，在一种实现方式中，由于时域轨道指示信息包括多时域轨道标识字段，时域层级标识字段、帧率字段和码率信息，码流信息包括低码率字段和高码率字段，那么，该内容消费设备则可读取时域轨道指示信息中的多时域轨道标识字段的值，在读取的多时域轨道标识字段为第二数值时，指示N个时域层级的媒体帧被封装到单个轨道，并读取时域层级标识字段的值，帧率字段的值，以及码率信息中低码率字段的值和高码率字段的值；从而使得内容消费设备可根据时域层级标识字段的值，帧率字段的值，以及码率信息中低码率字段的值和高码率字段的值，以及解码设备（即上述的内容消费设备）的解码性能，保留N个时域层级中与解码性能匹配的时域层级。

若该媒体数据为视频数据，则内容生成设备将视频数据（或视频内容）属于N个时域层级的媒体帧封装到单个轨道后，内容消费设备进行解码消费的过程具体如下：

内容生成设备对视频内容A进行编码、封装，假设视频内容A存在3个时域层级L0~L2，属于该三个时域层级的任一时域层级的媒体帧均封装入一个轨道，且累计到各个时域层级时，对应的帧率和码率如下：

L0:20fps, bitrate=1mbps;

L1:30fps, bitrate=1.5mbps;

L2:60fps, bitrate=3mbps.

内容生成设备根据内容消费设备的请求，将视频文件A分别发送给用户1和用户2，用户1和用户2分别收到对应的文件A，根据轨道中各个时域层级对应的帧率和码率信息，解

码消费。具体为：

用户1所在的内容消费设备的解码设备性能较好，选择保留L0~L2的全部媒体帧解码呈现，获得最佳观看效果，而用户2所在的内容消费设备的解码设备性能较差，选择保留L0的全部媒体帧，并可丢弃L1~L2的媒体帧，仅对保留的属于L0的媒体帧解码呈现。

5 在内容消费设备根据时域轨道指示信息及解码设备的解码性能，保留N个时域层级中与解码性能匹配的时域层级时，在另一种实现方式中，如果内容消费设备可读取时域轨道指示信息中的多时域轨道标识字段的值，在内容消费设备读取的多时域轨道标识字段为第一数值时，指示N个时域层级的媒体帧被封装到多个不同的轨道，在各轨道中的时域层级无重叠时，该内容消费设备可从基准轨道中读取时域层级标识字段的值，帧率字段的值，10 以及码率信息中低码率字段的值和高码率字段的值；基准轨道中封装的媒体帧被独立解码，并进而根据从基准轨道读取的时域层级标识字段的值，帧率字段的值，以及码率信息中低码率字段的值和高码率字段的值，以及解码设备的解码性能，保留部分或全部与解码性能匹配的轨道中的时域层级。

若该媒体数据为视频数据，则内容生成设备将视频数据（或视频内容）属于N个时域15 层级的媒体帧封装到多个不同轨道，且各轨道时域无重叠，内容消费设备进行解码消费的过程可如图6a所示，具体如下：

内容生成设备对视频内容A进行编码、封装，假设视频内容A存在3个时域层级L0~L2，且属于这三个时域层级的媒体帧被分别封装至三个不同轨道，其中，rack1为基准轨道，track2和track3以'tlrf'类型索引至track1。在基准轨道中，指示累计到各个时域层级时，对应的20 帧率和码率如下：

L0:20fps, bitrate=1mbps;

L1:30fps, bitrate=1.5mbps;

L2:60fps, bitrate=3mbps。

内容生成设备根据内容消费设备的请求，将视频文件A分别发送给用户1和用户2，用25 户1和用户2分别收到对应的文件A，根据轨道中各个时域层级对应的帧率和码率信息，解码消费。具体为：

用户1所在的内容消费设备的解码设备性能较好，选择保留track1~track3的全部媒体帧30 解码呈现，获得最佳观看效果，而用户2所在的内容消费设备的解码设备性能较差，选择保留track1的全部媒体帧，并可丢弃track2和track3的媒体帧，仅对保留的属于track1的媒体帧解码呈现。

在另一种实现方式中，若内容消费设备读取的多时域轨道标识字段为第一数值，并指示N个时域层级的媒体帧被封装到多个不同的轨道，在各轨道中的时域层级存在重叠时，该内容消费设备则可从基准轨道中读取轨道组合策略信息中的各字段的取值，并根据轨道35 组合策略信息中的各字段的取值，以及解码设备的解码性能，保留部分或全部与解码性能匹配的轨道中的时域层级，其中，时域轨道指示信息还包括轨道组合策略信息，轨道组合策略信息包括轨道标识字段，优先解码呈现字段，替代轨道字段，以及替代轨道标识字段。如图6b所示，若该媒体数据为视频数据，则内容生成设备将视频数据（或视频内容）属于

N个时域层级的媒体帧封装到多个不同轨道，且各轨道时域层级存在重叠，那么内容消费设备进行解码消费的过程可具体如下：

内容生成设备对视频内容A进行编码、封装，假设视频内容A存在3个时域层级L0~L2，且属于这三个时域层级的媒体帧被分别封装至三个不同轨道，其中，rack1为基准轨道，track2和track3以'tlrf'类型索引至track1，Track2和track3各自包含一部分L1和L2的媒体帧，且互不重叠，track2和track3的解码均依赖track1，但track2和track3之间没有依赖关系。在基准轨道中，指示各个轨道在组合时的信息：

track1: is_output_track_flag=1; is_alternative_track_flag=0:

track2: is_output_track_flag=1; is_alternative_track_flag=0:

track3: is_output_track_flag=0; is_alternative_track_flag=1; alternative_track=2。

内容生成设备根据内容消费设备的请求，将视频文件A分别发送给用户1和用户2，用户1和用户2分别收到对应的文件A，根据轨道中各个时域层级对应的帧率和码率信息，解码消费。具体为：

用户1所在的内容消费设备的解码设备性能较好，选择保留track1~track3的全部媒体帧解码呈现，获得最佳观看效果，而用户2所在的内容消费设备的解码设备性能较差，选择保留track1和track2的全部媒体帧，并可弃track3的媒体帧，解码track1中的媒体帧进行呈现。

在一个实施例中，内容消费设备在保留的媒体帧的数量为一个或多个，对保留的时域层级的媒体帧进行解码显示时，则可根据保留的一个或多个媒体帧中每个媒体帧的解码时间，对保留的一个或多个媒体帧按照解码时间重新排序（即重构），进而可对重新排序后的一个或多个媒体帧进行解码显示。也就是说，在组合不同轨道的媒体帧时，内容消费设备根据封装时每个媒体帧对应的解码时间，按照解码时间排列所选的多个轨道中所有媒体帧，重构之后再行解码。

在本申请实施例中，支持内容生成设备将多时域层级的媒体帧封装到不同的轨道，并在封装过程中，将在每个轨道的描述数据盒中记录时域轨道指示信息，从而实现关联不同时域层级的轨道，指示轨道内的时域层级信息，标注基准轨道，指示多时域层级轨道相互组合的策略信息。那么，内容消费设备则可根据轨道内的描述数据盒选取合适的时域层级，将不同轨道的样本组合后进行解码呈现，从而保证了多轨道封装的灵活性，并最大化地节省了解码计算资源。

下面，将结合图7，对内容生成设备对媒体数据的封装过程进行说明，该内容生成设备具体可以是服务器，或者也可以是终端设备，其中，该服务器可以是独立的服务器，也可以是多个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、中间件服务和人工智能平台等基础云计算服务的云服务器。请参见图7，是本申请实施例提供的一种媒体数据的处理方法的示意流程图，如图7所示，该方法可包括：

S701，根据媒体数据包括的每个媒体帧的帧间依赖关系，确定每个媒体帧的时域层级，得到N个时域层级的媒体帧；其中，N为大于1的正整数。

S702，分别将N个时域层级的媒体帧封装到M个轨道中，并生成对应的描述数据盒；M

个轨道中第j个轨道的描述数据盒包括时域轨道指示信息，时域轨道指示信息用于指示N个时域层级的轨道封装方式，时域轨道指示信息包括第j个轨道封装的时域层级的时域层级信息，其中，M为大于1的正整数。

在S701和S702中，媒体数据中各媒体帧的帧间依赖关系可如图4所示，进而可基于该帧间依赖关系确定每个媒体帧的时域层级。其中，时域轨道指示信息包括多时域轨道标识字段，多时域轨道标识字段用于指示N个时域层级的轨道封装方式，那么，内容生成设备在根据媒体帧的封装过程，生成第j个轨道的描述数据盒时，若内容生成设备将N个时域层级的多个媒体帧被封装到多个不同的轨道中，则生成多时域轨道标识字段的取值为第一数值；而如果该内容生成设备将N个时域层级的多个媒体帧被封装到单个轨道中，则生成多时域轨道标识字段的取值为第二数值。此外，该时域轨道指示信息还包括总时域层数字段，那么内容生成设备在根据媒体帧的封装过程，生成第j个轨道的描述数据盒时，还将根据M个轨道封装的时域层级的总数量，生成总时域层数字段的取值。

其中，该第j个轨道的时域层级信息包括时域层数字段，那么，该内容生成设备在根据媒体帧的封装过程，生成第j个轨道的描述数据盒时，将根据第j个轨道封装的时域层级的数量，生成时域层数字段的取值。此外，该第j个轨道的时域层级信息包括时域层级标识字段、帧率字段和码率信息，那么，内容生成设备在根据媒体帧的封装过程，生成第j个轨道的描述数据盒时，可将第j个轨道封装的时域层级中，第i个时域层级的层级标识存储在时域层级标识字段中，将累计到属于第i个时域层级的媒体帧时的帧率存储到帧率字段中，以及将累计到属于第i个时域层级的媒体帧时的码率，作为码率信息。在一个实施例中，码率信息包括低码率字段和高码率字段；那么，内容生成设备在将累计到属于第i个时域层级的媒体帧时的码率，作为码率信息时，可将属于第i个时域层级的媒体帧时的码率的低18位，存储在低码率字段中，并将属于第i个时域层级的媒体帧时的码率的高12位，存储在高码率字段中。

时域轨道指示信息还包括基准轨道标识字段，那么内容生成设备在根据媒体帧的封装过程，生成第j个轨道的描述数据盒时，若第j个轨道为基准轨道，则生成基准轨道标识字段的取值为第一数值，而若第j个轨道为非基准轨道，则生成基准轨道标识字段的取值为第二数值；其中，基准轨道中封装的媒体帧被独立解码。此外，时域轨道指示信息还包括轨道组合策略信息，轨道组合策略信息包括轨道标识字段，优先解码呈现字段，替代轨道字段，以及替代轨道标识字段；那么，内容生成设备在根据媒体帧的封装过程，生成第j个轨道的描述数据盒时，可将包含部分时域层级的轨道的标识存储在轨道标识字段中，并在第j个轨道为优先解码呈现的轨道，生成优先解码呈现字段的取值为第一数值，若第j个轨道不是优先解码呈现的轨道，则生成优先解码呈现字段的取值为第二数值；以及，若第j个轨道为一个轨道的替代轨道，则生成替代轨道字段的取值为第一数值，并将第j个轨道替代的一个轨道的标识存储在替代轨道标识字段中；若第j个轨道不是替代轨道，则生成替代轨道字段的取值为第二数值，其中，该第一数值可以为1，该第二数值可以为0。

在一个实施例中，若第j个轨道为非基准轨道，则内容生成设备将生成第j个轨道的轨道索引数据盒，轨道索引数据盒包括轨道索引类型数据盒；轨道索引类型数据盒包括轨道标识字段和索引类型标识字段，此外，该内容生成设备可将基准轨道的标识存储在轨道标识

字段中,并根据索引类型标识字段,将第j个轨道索引至基准轨道。

在本申请实施例中,内容生成设备通过媒体数据中所包括的媒体帧之间的帧间依赖关系,可确定每个媒体帧的时域层级,进而可将N个时域层级的媒体帧分别封装到M个轨道中,并基于媒体帧的封装过程,生成第j个轨道的描述数据盒,并在描述数据盒中,为时域轨道指示信息包括的各字段设置相应的取值,以通过各字段的取值将不同时域层级的轨道进行关联,指示轨道内的时域层级信息,对基准轨道进行标注,指示多时域层级轨道相互组合的策略,从而实现指示内容消费设备,该内容生成设备的媒体帧封装过程,那么也就使得内容消费设备可根据轨道中描述数据盒中的各字段的取值,选取合适的时域层级进行解码呈现,保证了多轨道封装的灵活性,并可最大化节省解码计算资源。

上述详细阐述了本申请实施例的方法,为了便于更好地实施本申请实施例的上述方案,相应地,下面提供了本申请实施例的装置。

请参见图8,图8示出了本申请一个示例性实施例提供的一种媒体数据的处理装置的结构示意图;该媒体数据的处理装置可以是运行于上述内容消费设备中的一个计算机程序(包括程序代码),例如该媒体数据的处理装置可以是内容消费设备中的一个应用软件。由图8所示,该媒体数据的处理装置可包括:获取单元801和处理单元802。

在一个示例性实施例中,该媒体数据的处理装置可以用于执行图5所示的方法中的相应步骤;所述媒体数据包括多个媒体帧,所述多个媒体帧被划分为N个时域层级,所述多个媒体帧被封装到M个轨道中,其中M和N均为大于1的正整数;则:

获取单元801,用于获取所述M个轨道中的第j个轨道的描述数据盒,所述描述数据盒包括时域轨道指示信息,所述时域轨道指示信息用于指示所述N个时域层级的轨道封装方式,所述时域轨道指示信息包括所述第j个轨道封装的时域层级的时域层级信息,其中j为正整数,且 $j \leq M$;

处理单元802,用于根据所述时域轨道指示信息,对所述媒体数据进行解码。

在一个实施例中,所述时域轨道指示信息包括多时域轨道标识字段,所述多时域轨道标识字段用于指示所述N个时域层级的轨道封装方式;

当所述多时域轨道标识字段为第一数值时,所述多时域轨道标识字段用于指示属于所述N个时域层级的多个媒体帧被封装到多个不同的轨道中;

当所述多时域轨道标识字段为第二数值时,所述多时域轨道标识字段用于指示属于所述N个时域层级的多个媒体帧被封装到单个轨道中。

在一个实施例中,所述时域轨道指示信息包括总时域层数字段;所述总时域层数字段用于指示所述M个轨道封装的时域层级的总数量。

在一个实施例中,所述第j个轨道的时域层级信息包括时域层数字段,所述时域层数字段用于指示所述第j个轨道封装的时域层级的数量。

在一个实施例中,所述第j个轨道的时域层级信息包括时域层级标识字段、帧率字段和码率信息;

所述时域层级标识字段用于指示所述第j个轨道封装的时域层级中,第i个时域层级的层级标识;

所述帧率字段用于指示累计到属于第i个时域层级的媒体帧时的帧率；

所述码率信息用于指示累计到属于第i个时域层级的媒体帧时的码率。

在一个实施例中，所述码率信息包括低码率字段和高码率字段；

所述低码率字段用于指示累计到属于第i个时域层级的媒体帧时的码率的低18位；

5 所述高码率字段用于指示累计到属于第i个时域层级的媒体帧时的码率的高12位。

在一个实施例中，所述时域轨道指示信息包括基准轨道标识字段；所述基准轨道标识字段用于指示所述第j个轨道是否为基准轨道；

10 当所述基准轨道标识字段为第一数值时，所述基准轨道标识字段用于指示所述第j个轨道为基准轨道，当所述基准轨道标识字段为第二数值时，所述基准轨道标识字段用于指示所述第j个轨道为非基准轨道；

其中，所述基准轨道中封装的媒体帧被独立解码。

在一个实施例中，所述时域轨道指示信息还包括轨道组合策略信息，所述轨道组合策略信息包括轨道标识字段，优先解码呈现字段，替代轨道字段，以及替代轨道标识字段；

所述轨道标识字段用于指示包含部分时域层级的轨道的标识；

15 所述优先解码呈现字段用于指示所述第j个轨道是否为优先解码呈现的轨道；当所述优先解码呈现字段为第一数值时，所述优先解码呈现字段用于指示所述第j个轨道为优先解码呈现的轨道；当所述优先解码呈现字段为第二数值时，所述优先解码呈现字段用于指示所述第j个轨道不是优先解码呈现的轨道；

20 所述替代轨道字段用于指示所述第j个轨道是否为所述M个轨道中的一个轨道的替代轨道；当所述替代轨道字段为第一数值时，所述替代轨道字段用于指示所述第j个轨道为所述M个轨道中的一个轨道的替代轨道；当所述替代轨道字段为第二数值时，所述替代轨道字段用于指示所述第j个轨道不是替代轨道；

所述替代轨道标识字段用于指示所述第j个轨道替代的一个轨道的标识。

25 在一个实施例中，若所述第j个轨道为非基准轨道，则所述第j个轨道还包括轨道索引数据盒，所述轨道索引数据盒包括轨道索引类型数据盒；

所述轨道索引类型数据盒包括轨道标识字段和索引类型标识字段；

所述轨道标识字段用于存储基准轨道的标识，所述索引类型标识字段用于指示被索引的轨道为基准轨道。

在一个实施例中，所述处理单元802，具体用于：

30 根据所述时域轨道指示信息及解码设备的解码性能，保留所述N个时域层级中与所述解码性能匹配的时域层级；

对保留的时域层级的媒体帧进行解码显示。

35 在一个实施例中，所述时域轨道指示信息包括多时域轨道标识字段，时域层级标识字段、帧率字段和码率信息，所述码流信息包括低码率字段和高码率字段；所述处理单元802，具体用于：

读取所述时域轨道指示信息中的多时域轨道标识字段的值，在读取的所述多时域轨道标识字段为第二数值时，指示所述N个时域层级的媒体帧被封装到单个轨道，并读取所述

时域层级标识字段的值, 所述帧率字段的值, 以及所述码率信息中低码率字段的值和高码率字段的值;

根据所述时域层级标识字段的值, 所述帧率字段的值, 以及所述码率信息中低码率字段的值和高码率字段的值, 以及所述解码设备的解码性能, 保留所述N个时域层级中与所
5 述解码性能匹配的时域层级。

在一个实施例中, 所述时域轨道指示信息包括多时域轨道标识字段, 时域层级标识字段、帧率字段和码率信息, 所述码流信息包括低码率字段和高码率字段; 所述处理单元802, 具体用于:

10 读取所述时域轨道指示信息中的多时域轨道标识字段的值, 在读取的所述多时域轨道标识字段为第一数值时, 指示所述N个时域层级的媒体帧被封装到多个不同的轨道, 在各轨道中的时域层级无重叠时, 从基准轨道中读取所述时域层级标识字段的值, 所述帧率字段的值, 以及所述码率信息中低码率字段的值和高码率字段的值; 所述基准轨道中封装的媒体帧被独立解码;

15 根据从所述基准轨道读取的所述时域层级标识字段的值, 所述帧率字段的值, 以及所述码率信息中低码率字段的值和高码率字段的值, 以及所述解码设备的解码性能, 保留部分或全部与所述解码性能匹配的轨道中的时域层级。

在一个实施例中, 所述时域轨道指示信息还包括轨道组合策略信息, 所述轨道组合策略信息包括轨道标识字段, 优先解码呈现字段, 替代轨道字段, 以及替代轨道标识字段;

20 所述处理单元802, 还用于在读取的所述多时域轨道标识字段为第一数值时, 指示所述N个时域层级的媒体帧被封装到多个不同的轨道, 在各轨道中的时域层级存在重叠, 从基准轨道中读取所述轨道组合策略信息中的各字段的取值;

所述处理单元802, 还用于根据所述轨道组合策略信息中的各字段的取值, 以及所述解码设备的解码性能, 保留部分或全部与所述解码性能匹配的轨道中的时域层级。

在一个实施例中, 所述处理单元802, 具体用于:

25 根据保留的一个或多个媒体帧中每个媒体帧的解码时间, 对保留的一个或多个媒体帧按照所述解码时间重新排序;

对重新排序后的一个或多个媒体帧进行解码显示。

30 在本申请实施例中, 支持内容生成设备将多时域层级的媒体帧封装到不同的轨道, 并在封装过程中, 将在每个轨道的描述数据盒中记录时域轨道指示信息, 从而实现关联不同时域层级的轨道, 指示轨道内的时域层级信息, 标注基准轨道, 指示多时域层级轨道相互组合的策略信息。那么, 处理单元602则可根据轨道内的描述数据盒选取合适的时域层级, 将不同轨道的样本组合后进行解码呈现, 从而保证了多轨道封装的灵活性, 并最大化地节省了解码计算资源。

35 上述详细阐述了本申请实施例的方法, 为了便于更好地实施本申请实施例的上述方案, 相应地, 下面提供了本申请实施例的装置。

请参见图9, 图9示出了本申请一个示例性实施例提供的一种媒体数据的处理装置的结构示意图; 该媒体数据的处理装置可以是运行于上述内容生成设备中的一个计算机程序(包

括程序代码), 例如该媒体数据的处理装置可以是内容生成设备中的一个应用软件。由图9所示, 该媒体数据的处理装置可包括: 确定单元901和处理单元902。

在一个示例性实施例中, 该媒体数据的处理装置可以用于执行图7所示的方法中的相应步骤; 则:

5 确定单元901, 用于根据媒体数据包括的每个媒体帧的帧间依赖关系, 确定所述每个媒体帧的时域层级, 得到N个时域层级的媒体帧; 其中, N为大于1的正整数;

处理单元902, 用于分别将所述N个时域层级的媒体帧封装到M个轨道中, 并生成对应的描述数据盒; 所述M个轨道中第j个轨道的描述数据盒包括时域轨道指示信息, 所述时域轨道指示信息用于指示所述N个时域层级的轨道封装方式, 所述时域轨道指示信息包括所述第j个轨道封装的时域层级的时域层级信息, 其中, M为大于1的正整数。

10 在一个实施例中, 所述时域轨道指示信息包括多时域轨道标识字段, 所述多时域轨道标识字段用于指示所述N个时域层级的轨道封装方式; 所述处理单元902, 具体用于:

若将所述N个时域层级的多个媒体帧被封装到多个不同的轨道中, 则生成所述多时域轨道标识字段的取值为第一数值;

15 若将所述N个时域层级的多个媒体帧被封装到单个轨道中, 则生成所述多时域轨道标识字段的取值为第二数值。

在一个实施例中, 所述时域轨道指示信息包括总时域层数字段; 所述处理单元902, 具体用于:

根据所述M个轨道封装的时域层级的总数量, 生成所述总时域层数字段的取值。

20 在一个实施例中, 所述第j个轨道的时域层级信息包括时域层数字段; 所述处理单元902, 具体用于:

根据所述第j个轨道封装的时域层级的数量, 生成所述时域层数字段的取值。

在一个实施例中, 所述第j个轨道的时域层级信息包括时域层级标识字段、帧率字段和码率信息; 所述处理单元902, 具体用于:

25 将所述第j个轨道封装的时域层级中, 第i个时域层级的层级标识存储在所述时域层级标识字段中;

将累计到属于第i个时域层级的媒体帧时的帧率存储到所述帧率字段中;

将累计到属于第i个时域层级的媒体帧时的码率, 作为所述码率信息。

30 在一个实施例中, 所述码率信息包括低码率字段和高码率字段; 所述处理单元902, 具体用于:

将属于第i个时域层级的媒体帧时的码率的低18位, 存储在所述低码率字段中;

将属于第i个时域层级的媒体帧时的码率的高12位, 存储在所述高码率字段中。

在一个实施例中, 所述时域轨道指示信息包括基准轨道标识字段; 所述处理单元902, 具体用于:

35 若所述第j个轨道为基准轨道, 则生成所述基准轨道标识字段的取值为第一数值;

若所述第j个轨道为非基准轨道, 则生成所述基准轨道标识字段的取值为第二数值;

其中, 所述基准轨道中封装的媒体帧被独立解码。

在一个实施例中，所述时域轨道指示信息还包括轨道组合策略信息，所述轨道组合策略信息包括轨道标识字段，优先解码呈现字段，替代轨道字段，以及替代轨道标识字段；所述处理单元902，具体用于：

将包含部分时域层级的轨道的标识存储在所述轨道标识字段中；

5 若所述第j个轨道为优先解码呈现的轨道，则生成所述优先解码呈现字段的取值为第一数值，若所述第j个轨道不是优先解码呈现的轨道，则生成所述优先解码呈现字段的取值为第二数值；

10 若所述第j个轨道为一个轨道的替代轨道，则生成所述替代轨道字段的取值为第一数值，并将所述第j个轨道替代的一个轨道的标识存储在所述替代轨道标识字段中；若所述第j个轨道不是替代轨道，则生成所述替代轨道字段的取值为第二数值。

在一个实施例中，所述处理单元902，还用于若所述第j个轨道为非基准轨道，则生成第j个轨道的轨道索引数据盒，所述轨道索引数据盒包括轨道索引类型数据盒；所述轨道索引类型数据盒包括轨道标识字段和索引类型标识字段；

15 所述处理单元902，还用于将所述基准轨道的标识存储在所述轨道标识字段中，并根据所述索引类型标识字段，将所述第j个轨道索引至基准轨道。

在本申请实施例中，处理单元902通过媒体数据中所包括的媒体帧之间的帧间依赖关系，可确定每个媒体帧的时域层级，进而可将N个时域层级的媒体帧分别封装到M个轨道中，并基于媒体帧的封装过程，生成第j个轨道的描述数据盒，并在描述数据盒中，为时域轨道指示信息包括的各字段设置相应的取值，以通过各字段的取值将不同时域层级的轨道进行关
20 联，指示轨道内的时域层级信息，对基准轨道进行标注，指示多时域层级轨道相互组合的策略，从而实现指示内容消费设备，该处理单元902的媒体帧封装过程，那么也就使得内容消费设备可根据轨道中描述数据盒中的各字段的取值，选取合适的时域层级进行解码呈现，保证了多轨道封装的灵活性，并可最大化节省解码计算资源。

25 请参见图10，是本申请实施例提供的一种计算机设备的结构示意图，该计算机设备可以是上述的该内容消费设备，或者也可以是上述的内容生成设备，其中，该计算机设备可以是服务器，也可以是终端设备。如图10所示的本实施例中的计算机设备可包括：一个或多个处理器101；一个或多个输入设备102，一个或多个输出设备103和存储器104。上述处理器101、输入设备102、输出设备103和存储器104通过总线105连接。存储器104用于存储计算机程序，所述计算机程序包括程序指令，处理器101 用于执行所述存储器104存储
30 的程序指令。

所述存储器104可以包括易失性存储器（volatile memory），如随机存取存储器（random-access memory，RAM）；存储器104也可以包括非易失性存储器（non-volatile memory），如快闪存储器（flash memory），固态硬盘（solid-state drive，SSD）等；存储器104还可以包括上述种类的存储器的组合。

35 所述处理器101可以是中央处理器（central processing unit，CPU）。所述处理器101还可以进一步包括硬件芯片。上述硬件芯片可以是专用集成电路（application-specific integrated circuit，ASIC），可编程逻辑器件（programmable logic device，PLD）等。该PLD可以是现

场可编程逻辑门阵列 (field-programmable gate array, FPGA), 通用阵列逻辑 (generic array logic, GAL) 等。所述处理器101也可以为上述结构的组合。

本申请实施例中, 所述存储器104用于存储计算机程序, 所述计算机程序包括程序指令, 处理器101用于执行存储器104存储的程序指令, 用来实现上述如图5中媒体数据的处理方法所涉及的步骤, 其中, 所述媒体数据包括多个媒体帧, 所述多个媒体帧被划分为N个时域层级, 所述多个媒体帧被封装到M个轨道中, 其中M和N均为大于1的正整数。

在一个实施例中, 所述处理器101被配置调用所述程序指令, 用于执行:

获取所述M个轨道中的第j个轨道的描述数据盒, 所述描述数据盒包括时域轨道指示信息, 所述时域轨道指示信息用于指示所述N个时域层级的轨道封装方式, 所述时域轨道指示信息包括所述第j个轨道封装的时域层级的时域层级信息, 其中j为正整数, 且 $j \leq M$;

根据所述时域轨道指示信息, 对所述媒体数据进行解码。

在一个实施例中, 所述存储器104用于存储计算机程序, 所述计算机程序包括程序指令, 处理器101用于执行存储器104存储的程序指令, 还可用来实现上述如图7中相应方法的步骤。

在一个实施例中, 所述处理器101被配置调用所述程序指令, 用于执行:

根据媒体数据包括的每个媒体帧的帧间依赖关系, 确定所述每个媒体帧的时域层级, 得到N个时域层级的媒体帧; 其中, N为大于1的正整数;

分别将所述N个时域层级的媒体帧封装到M个轨道中, 并生成对应的描述数据盒; 所述M个轨道中第j个轨道的描述数据盒包括时域轨道指示信息, 所述时域轨道指示信息用于指示所述N个时域层级的轨道封装方式, 所述时域轨道指示信息包括所述第j个轨道封装的时域层级的时域层级信息, 其中, M为大于1的正整数。

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。

本领域普通技术人员可以理解实现上述实施例方法中的全部或部分流程, 是可以通过计算机程序来指令相关的硬件来完成, 所述的程序可存储于一计算机可读取存储介质中, 该程序在执行时, 可包括如上述各方法的实施例的流程。其中, 所述的存储介质可为磁碟、光盘、只读存储记忆体 (Read-Only Memory, ROM) 或随机存储记忆体 (Random Access Memory, RAM) 等。

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、一种媒体数据的处理方法，所述媒体数据包括多个媒体帧，所述多个媒体帧被划分为N个时域层级，所述多个媒体帧被封装到M个轨道中，其中M和N均为大于1的正整数；所述方法由内容消费设备执行，所述方法包括：

5 获取所述M个轨道中的第j个轨道的描述数据盒，所述描述数据盒包括时域轨道指示信息，所述时域轨道指示信息用于指示所述N个时域层级的轨道封装方式，所述时域轨道指示信息包括所述第j个轨道封装的时域层级的时域层级信息，其中j为正整数，且 $j \leq M$ ；

根据所述时域轨道指示信息，对所述媒体数据进行解码。

10 2、如权利要求1所述的方法，所述时域轨道指示信息包括多时域轨道标识字段，所述多时域轨道标识字段用于指示所述N个时域层级的轨道封装方式；

当所述多时域轨道标识字段为第一数值时，所述多时域轨道标识字段用于指示属于所述N个时域层级的多个媒体帧被封装到多个不同的轨道中；

当所述多时域轨道标识字段为第二数值时，所述多时域轨道标识字段用于指示属于所述N个时域层级的多个媒体帧被封装到单个轨道中。

15 3、如权利要求1所述的方法，所述时域轨道指示信息包括总时域层数字段；所述总时域层数字段用于指示所述M个轨道封装的时域层级的总数量。

4、如权利要求1所述的方法，所述第j个轨道的时域层级信息包括时域层数字段，所述时域层数字段用于指示所述第j个轨道封装的时域层级的数量。

20 5、如权利要求1所述的方法，所述第j个轨道的时域层级信息包括时域层级标识字段、帧率字段和码率信息；

所述时域层级标识字段用于指示所述第j个轨道封装的时域层级中，第i个时域层级的层级标识；

所述帧率字段用于指示累计到属于第i个时域层级的媒体帧时的帧率；

所述码率信息用于指示累计到属于第i个时域层级的媒体帧时的码率。

25 6、如权利要求5所述的方法，所述码率信息包括低码率字段和高码率字段；

所述低码率字段用于指示累计到属于第i个时域层级的媒体帧时的码率的低18位；

所述高码率字段用于指示累计到属于第i个时域层级的媒体帧时的码率的高12位。

7、如权利要求1所述的方法，所述时域轨道指示信息包括基准轨道标识字段；所述基准轨道标识字段用于指示所述第j个轨道是否为基准轨道；

30 当所述基准轨道标识字段为第一数值时，所述基准轨道标识字段用于指示所述第j个轨道为基准轨道，当所述基准轨道标识字段为第二数值时，所述基准轨道标识字段用于指示所述第j个轨道为非基准轨道；

其中，所述基准轨道中封装的媒体帧被独立解码。

35 8、如权利要求1所述的方法，所述时域轨道指示信息还包括轨道组合策略信息，所述轨道组合策略信息包括轨道标识字段，优先解码呈现字段，替代轨道字段，以及替代轨道标识字段；

所述轨道标识字段用于指示包含部分时域层级的轨道的标识；

所述优先解码呈现字段用于指示所述第j个轨道是否为优先解码呈现的轨道；当所述优先解码呈现字段为第一数值时，所述优先解码呈现字段用于指示所述第j个轨道为优先解码呈现的轨道；当所述优先解码呈现字段为第二数值时，所述优先解码呈现字段用于指示所述第j个轨道不是优先解码呈现的轨道；

5 所述替代轨道字段用于指示所述第j个轨道是否为所述M个轨道中的一个轨道的替代轨道；当所述替代轨道字段为第一数值时，所述替代轨道字段用于指示所述第j个轨道为所述M个轨道中的一个轨道的替代轨道；当所述替代轨道字段为第二数值时，所述替代轨道字段用于指示所述第j个轨道不是替代轨道；

所述替代轨道标识字段用于指示所述第j个轨道替代的一个轨道的标识。

10 9、如权利要求1所述的方法，若所述第j个轨道为非基准轨道，则所述第j个轨道还包括轨道索引数据盒，所述轨道索引数据盒包括轨道索引类型数据盒；

所述轨道索引类型数据盒包括轨道标识字段和索引类型标识字段；

所述轨道标识字段用于存储基准轨道的标识，所述索引类型标识字段用于指示被索引的轨道为基准轨道。

15 10、如权利要求1~9任一项所述的方法，所述根据所述时域轨道指示信息，对所述媒体数据进行解码，包括：

根据所述时域轨道指示信息及解码设备的解码性能，保留所述N个时域层级中与所述解码性能匹配的时域层级；

对保留的时域层级的媒体帧进行解码。

20 11、如权利要求10所述的方法，所述时域轨道指示信息包括多时域轨道标识字段，时域层级标识字段、帧率字段和码率信息，所述码流信息包括低码率字段和高码率字段；

所述根据所述时域轨道指示信息及所述解码设备的解码性能，保留所述N个时域层级中与所述解码性能匹配的时域层级，包括：

25 读取所述时域轨道指示信息中的多时域轨道标识字段的值，在读取的所述多时域轨道标识字段为第二数值时，指示所述N个时域层级的媒体帧被封装到单个轨道，并读取所述时域层级标识字段的值，所述帧率字段的值，以及所述码率信息中低码率字段的值和高码率字段的值；

30 根据所述时域层级标识字段的值，所述帧率字段的值，以及所述码率信息中低码率字段的值和高码率字段的值，以及所述解码设备的解码性能，保留所述N个时域层级中与所述解码性能匹配的时域层级。

12、如权利要求10所述的方法，所述时域轨道指示信息包括多时域轨道标识字段，时域层级标识字段、帧率字段和码率信息，所述码流信息包括低码率字段和高码率字段；

所述根据所述时域轨道指示信息及所述解码设备的解码性能，保留所述N个时域层级中与所述解码性能匹配的时域层级，包括：

35 读取所述时域轨道指示信息中的多时域轨道标识字段的值，在读取的所述多时域轨道标识字段为第一数值时，指示所述N个时域层级的媒体帧被封装到多个不同的轨道，在各轨道中的时域层级无重叠时，从基准轨道中读取所述时域层级标识字段的值，所述帧率字

段的值, 以及所述码率信息中低码率字段的值和高码率字段的值; 所述基准轨道中封装的媒体帧被独立解码;

根据从所述基准轨道读取的所述时域层级标识字段的值, 所述帧率字段的值, 以及所述码率信息中低码率字段的值和高码率字段的值, 以及所述解码设备的解码性能, 保留部分或全部与所述解码性能匹配的轨道中的时域层级。

13、如权利要求12所述的方法, 所述时域轨道指示信息还包括轨道组合策略信息, 所述轨道组合策略信息包括轨道标识字段, 优先解码呈现字段, 替代轨道字段, 以及替代轨道标识字段; 所述方法还包括:

在读取的所述多时域轨道标识字段为第一数值时, 指示所述N个时域层级的媒体帧被封装到多个不同的轨道, 在各轨道中的时域层级存在重叠, 从基准轨道中读取所述轨道组合策略信息中的各字段的取值;

根据所述轨道组合策略信息中的各字段的取值, 以及所述解码设备的解码性能, 保留部分或全部与所述解码性能匹配的轨道中的时域层级。

14、如权利要求10所述的方法, 保留的媒体帧的数量为一个或多个, 所述对保留的时域层级的媒体帧进行解码, 包括:

根据保留的一个或多个媒体帧中每个媒体帧的解码时间, 对保留的一个或多个媒体帧按照所述解码时间重新排序;

对重新排序后的一个或多个媒体帧进行解码。

15、一种媒体数据的处理方法, 所述方法由内容生成设备执行, 所述方法包括:

根据媒体数据包括的每个媒体帧的帧间依赖关系, 确定所述每个媒体帧的时域层级, 得到N个时域层级的媒体帧; 其中, N为大于1的正整数;

分别将所述N个时域层级的媒体帧封装到M个轨道中, 并生成对应的描述数据盒; 所述M个轨道中第j个轨道的描述数据盒包括时域轨道指示信息, 所述时域轨道指示信息用于指示所述N个时域层级的轨道封装方式, 所述时域轨道指示信息包括所述第j个轨道封装的时域层级的时域层级信息, 其中, M为大于1的正整数。

16、如权利要求15所述的方法, 所述时域轨道指示信息包括多时域轨道标识字段, 所述多时域轨道标识字段用于指示所述N个时域层级的轨道封装方式;

所述根据媒体帧的封装过程, 生成第j个轨道的描述数据盒, 包括:

若将所述N个时域层级的多个媒体帧被封装到多个不同的轨道中, 则生成所述多时域轨道标识字段的取值为第一数值;

若将所述N个时域层级的多个媒体帧被封装到单个轨道中, 则生成所述多时域轨道标识字段的取值为第二数值。

17、如权利要求15所述的方法, 所述时域轨道指示信息包括总时域层数字段;

所述根据媒体帧的封装过程, 生成第j个轨道的描述数据盒, 包括:

根据所述M个轨道封装的时域层级的总数量, 生成所述总时域层数字段的取值。

18、如权利要求15所述的方法, 所述第j个轨道的时域层级信息包括时域层数字段;

所述根据媒体帧的封装过程, 生成第j个轨道的描述数据盒, 包括:

根据所述第j个轨道封装的时域层级的数量，生成所述时域层数字段的取值。

19、如权利要求15所述的方法，所述第j个轨道的时域层级信息包括时域层级标识字段、帧率字段和码率信息；

所述根据媒体帧的封装过程，生成第j个轨道的描述数据盒，包括：

5 将所述第j个轨道封装的时域层级中，第i个时域层级的层级标识存储在所述时域层级标识字段中；

将累计到属于第i个时域层级的媒体帧时的帧率存储到所述帧率字段中；

将累计到属于第i个时域层级的媒体帧时的码率，作为所述码率信息。

20、如权利要求19所述的方法，所述码率信息包括低码率字段和高码率字段；所述将
10 累计到属于第i个时域层级的媒体帧时的码率，作为所述码率信息，包括：

将属于第i个时域层级的媒体帧时的码率的低18位，存储在所述低码率字段中；

将属于第i个时域层级的媒体帧时的码率的高12位，存储在所述高码率字段中。

21、如权利要求15所述的方法，所述时域轨道指示信息包括基准轨道标识字段；

所述根据媒体帧的封装过程，生成第j个轨道的描述数据盒，包括：

15 若所述第j个轨道为基准轨道，则生成所述基准轨道标识字段的取值为第一数值；

若所述第j个轨道为非基准轨道，则生成所述基准轨道标识字段的取值为第二数值；

其中，所述基准轨道中封装的媒体帧被独立解码。

22、如权利要求15所述的方法，所述时域轨道指示信息还包括轨道组合策略信息，所述
20 轨道组合策略信息包括轨道标识字段，优先解码呈现字段，替代轨道字段，以及替代轨道标识字段；

所述根据媒体帧的封装过程，生成第j个轨道的描述数据盒，包括：

将包含部分时域层级的轨道的标识存储在所述轨道标识字段中；

25 若所述第j个轨道为优先解码呈现的轨道，则生成所述优先解码呈现字段的取值为第一数值，若所述第j个轨道不是优先解码呈现的轨道，则生成所述优先解码呈现字段的取值为第二数值；

若所述第j个轨道为一个轨道的替代轨道，则生成所述替代轨道字段的取值为第一数值，并将所述第j个轨道替代的一个轨道的标识存储在所述替代轨道标识字段中；若所述第j个轨道不是替代轨道，则生成所述替代轨道字段的取值为第二数值。

23、如权利要求15所述的方法，所述方法还包括：

30 若所述第j个轨道为非基准轨道，则生成第j个轨道的轨道索引数据盒，所述轨道索引数据盒包括轨道索引类型数据盒；所述轨道索引类型数据盒包括轨道标识字段和索引类型标识字段；

将所述基准轨道的标识存储在所述轨道标识字段中，并根据所述索引类型标识字段，将所述第j个轨道索引至基准轨道。

35 24、一种媒体数据的处理装置，所述媒体数据包括多个媒体帧，所述多个媒体帧被划分为N个时域层级，所述多个媒体帧被封装到M个轨道中，其中M和N均为大于1的正整数；所述装置包括：

获取单元，用于获取所述M个轨道中的第j个轨道的描述数据盒，所述描述数据盒包括时域轨道指示信息，所述时域轨道指示信息用于指示所述N个时域层级的轨道封装方式，所述时域轨道指示信息包括所述第j个轨道封装的时域层级的时域层级信息，其中j为正整数，且 $j \leq M$;

5 处理单元，用于根据所述时域轨道指示信息，对所述媒体数据进行解码。

25、一种媒体数据的处理装置，包括：

确定单元，用于根据媒体数据包括的每个媒体帧的帧间依赖关系，确定所述每个媒体帧的时域层级，得到N个时域层级的媒体帧；其中，N为大于1的正整数；

10 处理单元，用于分别将所述N个时域层级的媒体帧封装到M个轨道中，并生成对应的描述数据盒；所述M个轨道中第j个轨道的描述数据盒包括时域轨道指示信息，所述时域轨道指示信息用于指示所述N个时域层级的轨道封装方式，所述时域轨道指示信息包括所述第j个轨道封装的时域层级的时域层级信息，其中，M为大于1的正整数。

26、一种计算机设备，包括：

处理器，适于实现一条或多条指令；以及，

15 存储器，存储有一条或多条指令，所述一条或多条指令适于由所述处理器加载并执行如权利要求1~14任一项所述的方法，或者，执行如权利要求15~23任一项所述的方法。

20 27、一种计算机可读存储介质，所述计算机可读存储介质存储有计算机程序，所述计算机程序包括第一程序指令和第二程序指令，所述第一程序指令当被处理器执行时使所述处理器执行如权利要求1~14任一项所述的方法，所述第二程序指令当被处理器执行时使所述处理器执行如权利要求15~23任一项所述的方法。

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1~14任一项所述的方法，或者，执行权利要求15~23任一项所述的方法。

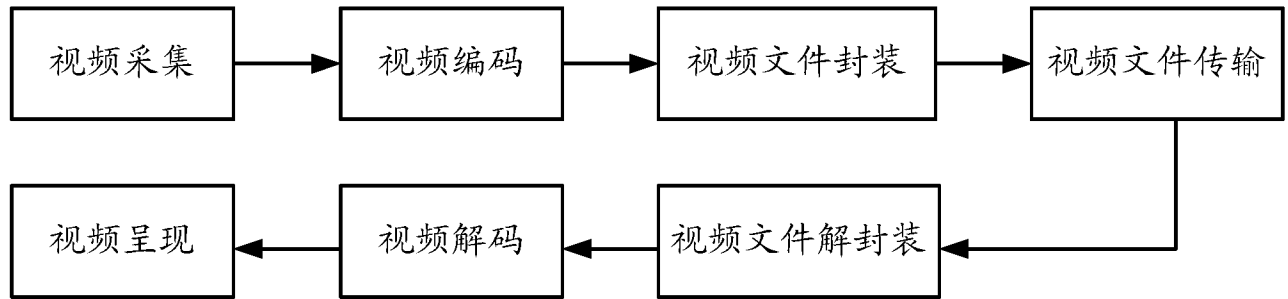


图 1

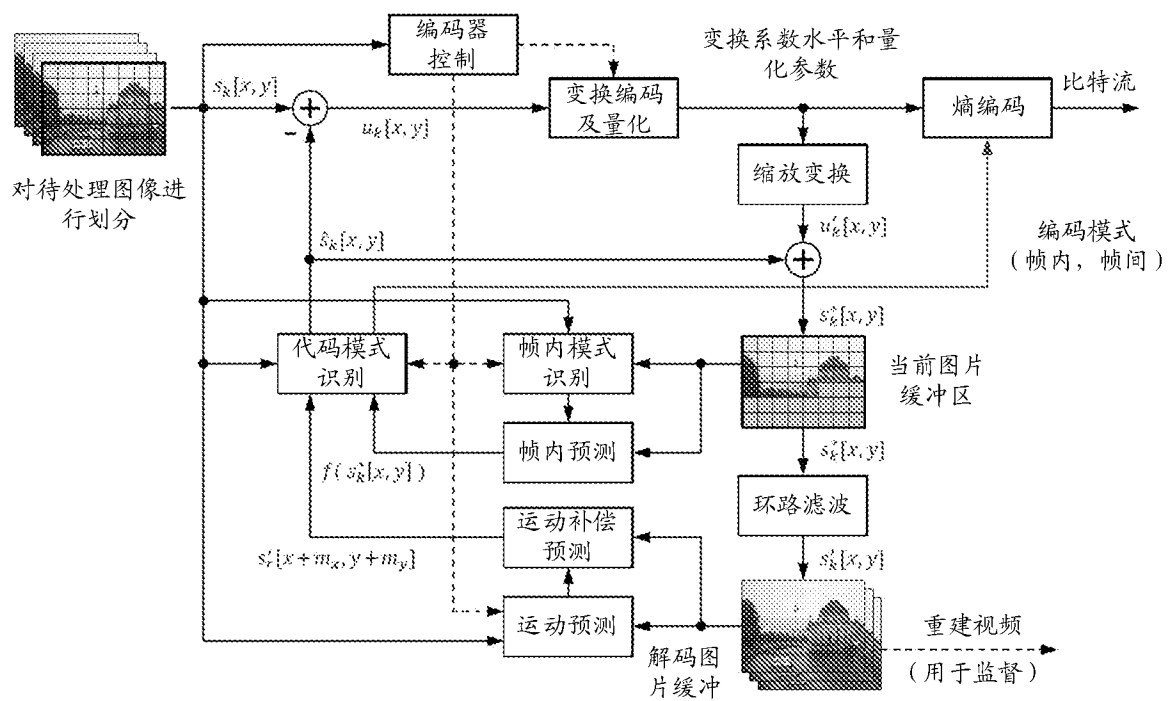
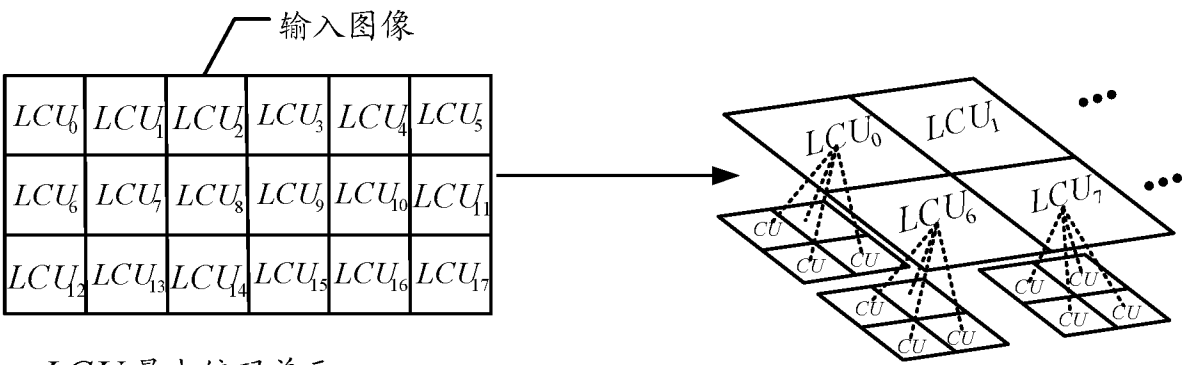


图 2



LCU :最大编码单元
 CU : 编码单元 (图像块)

图 3

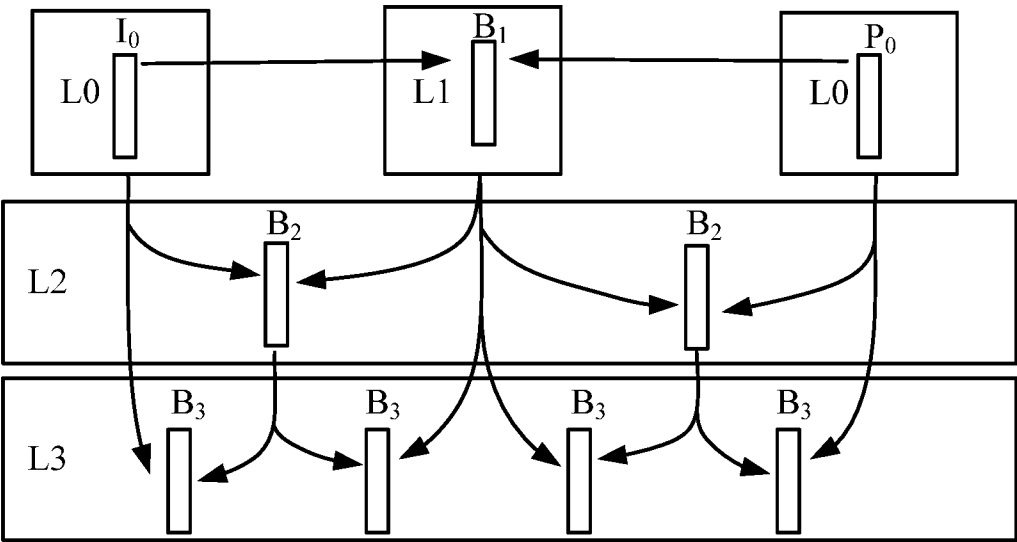


图 4

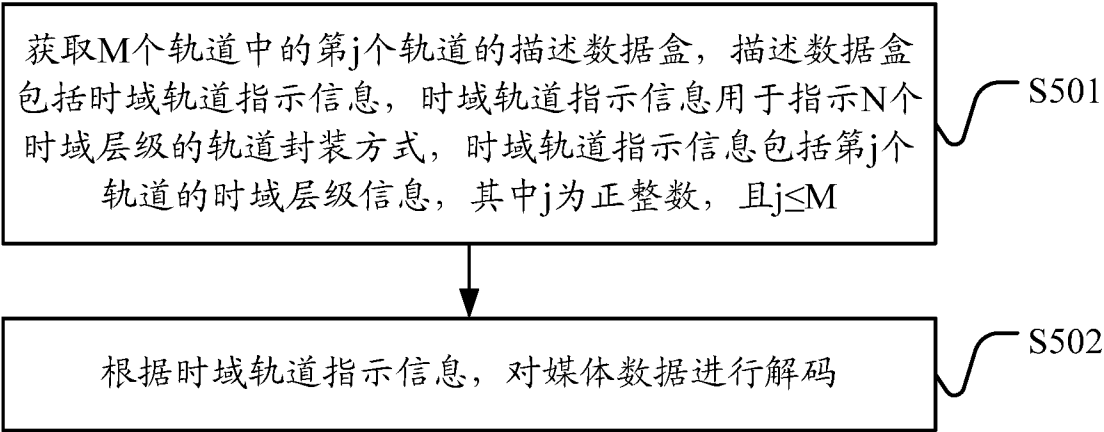


图 5

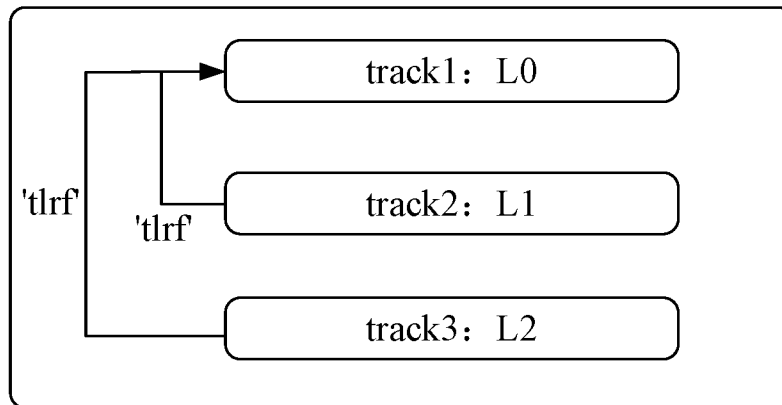


图 6a

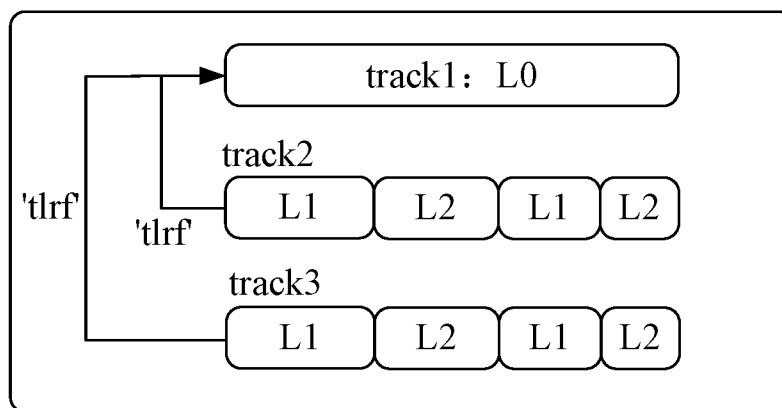


图 6b

根据媒体数据包括的每个媒体帧的帧间依赖关系，确定每个媒体帧的时域层级，得到N个时域层级的媒体帧；其中，N为正整数

S701

分别将N个时域层级的媒体帧封装到M个轨道中，并根据媒体帧的封装过程，生成第j个轨道的描述数据盒；描述数据盒包括时域轨道指示信息，时域轨道指示信息用于指示N个时域层级的轨道封装方式，时域轨道指示信息包括第j个轨道的时域层级信息，其中，M和j均为正整数，且 $j \leq M$

S702

图 7

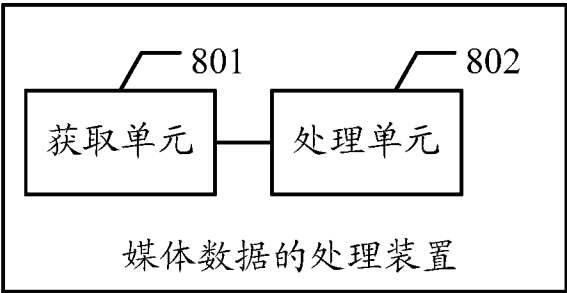


图 8

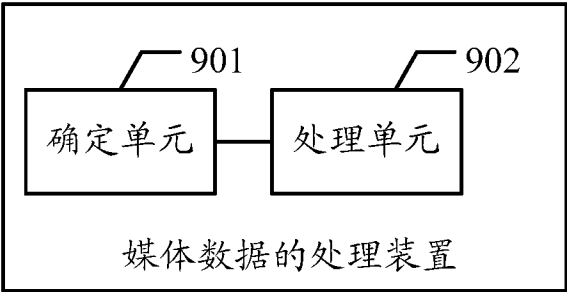


图 9

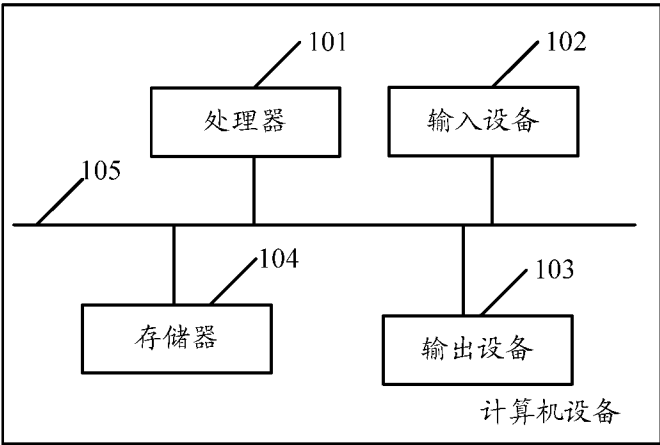


图 10

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2022/083960

A. CLASSIFICATION OF SUBJECT MATTER

H04N 19/31(2014.01)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)

CNABS, CNTXT, CNKI, VEN, USTXT, WOTXT, EPTXT: 编码, 解码, 数据, 帧, 时域, 轨道, encode, decode, data, frame, temporal domain, track, path

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN 112804256 A (TENCENT TECHNOLOGY SHENZHEN CO., LTD.) 14 May 2021 (2021-05-14) entire document	1-28
A	CN 106664446 A (CANON INC.) 10 May 2017 (2017-05-10) entire document	1-28
A	CN 101518087 A (NOKIA CORP.) 26 August 2009 (2009-08-26) entire document	1-28
A	GB 2506911 A (CANON K. K.) 16 April 2014 (2014-04-16) 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

10 June 2022

Date of mailing of the international search report

21 June 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/CN2022/083960

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	106664446	A	10 May 2017	KR	20170012396	A	02 February 2017
				WO	2016001337	A1	07 January 2016
				GB	201411731	D0	13 August 2014
				US	2017171282	A1	15 June 2017
				JP	2020022187	A	06 February 2020
				JP	2017525249	A	31 August 2017
				EP	3164994	A1	10 May 2017
CN	101518087	A	26 August 2009	ZA	200901936	B	31 March 2010
				KR	20090045375	A	07 May 2009
				RU	2009109847	A	27 October 2010
				WO	2008023328	A2	28 February 2008
				EP	2055107	A2	06 May 2009
				AU	2007287222	A1	28 February 2008
				CA	2661578	A1	28 February 2008
GB	2506911	A	16 April 2014	GB	201218370	D0	28 November 2012

国际检索报告

国际申请号

PCT/CN2022/083960

A. 主题的分类

H04N 19/31 (2014.01) 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))

CNABS, CNTXT, CNKI, VEN, USTXT, WOTXT, EPTXT: 编码, 解码, 数据, 帧, 时域, 轨道, encode, decode, data, frame, temporal domain, track, path

C. 相关文件

类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求
A	CN 112804256 A (腾讯科技深圳有限公司) 2021年5月14日 (2021 - 05 - 14) 全文	1-28
A	CN 106664446 A (佳能株式会社) 2017年5月10日 (2017 - 05 - 10) 全文	1-28
A	CN 101518087 A (诺基亚公司) 2009年8月26日 (2009 - 08 - 26) 全文	1-28
A	GB 2506911 A (CANON KK) 2014年4月16日 (2014 - 04 - 16) 全文	1-28

☐ 其余文件在C栏的续页中列出。☒ 见同族专利附件。

* 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

2022年6月10日

国际检索报告邮寄日期

2022年6月21日

ISA/CN的名称和邮寄地址

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

受权官员

飞雁

电话号码 (86-10)62411532

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2022/083960

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	106664446	A	2017年5月10日	KR	20170012396	A	2017年2月2日
				WO	2016001337	A1	2016年1月7日
				GB	201411731	D0	2014年8月13日
				US	2017171282	A1	2017年6月15日
				JP	2020022187	A	2020年2月6日
				JP	2017525249	A	2017年8月31日
				EP	3164994	A1	2017年5月10日
CN	101518087	A	2009年8月26日	ZA	200901936	B	2010年3月31日
				KR	20090045375	A	2009年5月7日
				RU	2009109847	A	2010年10月27日
				WO	2008023328	A2	2008年2月28日
				EP	2055107	A2	2009年5月6日
				AU	2007287222	A1	2008年2月28日
				CA	2661578	A1	2008年2月28日
GB	2506911	A	2014年4月16日	GB	201218370	D0	2012年11月28日