



(12) 发明专利

(10) 授权公告号 CN 113661692 B

(45) 授权公告日 2023.04.18

(21) 申请号 202080027558.2

(22) 申请日 2020.09.28

(65) 同一申请的已公布的文献号
申请公布号 CN 113661692 A

(43) 申请公布日 2021.11.16

(30) 优先权数据
62/908,238 2019.09.30 US
17/031,227 2020.09.24 US

(85) PCT国际申请进入国家阶段日
2021.10.09

(86) PCT国际申请的申请数据
PCT/US2020/053031 2020.09.28

(87) PCT国际申请的公布数据
W02021/067171 EN 2021.04.08

(73) 专利权人 腾讯美国有限责任公司
地址 美国加利福尼亚州帕洛阿尔托公园大道2747号

(72) 发明人 伊拉吉·索达加

(74) 专利代理机构 北京德琦知识产权代理有限公司 11018
专利代理师 陈美娥 王琦

(51) Int.Cl.
H04L 65/61 (2022.01)
H04L 65/60 (2022.01)
H04L 65/65 (2022.01)
H04L 67/02 (2022.01)

(56) 对比文件
US 2019281100 A1, 2019.09.12
CN 104185996 A, 2014.12.03
CN 107534793 A, 2018.01.02
CN 106034262 A, 2016.10.19
审查员 朱晓旭

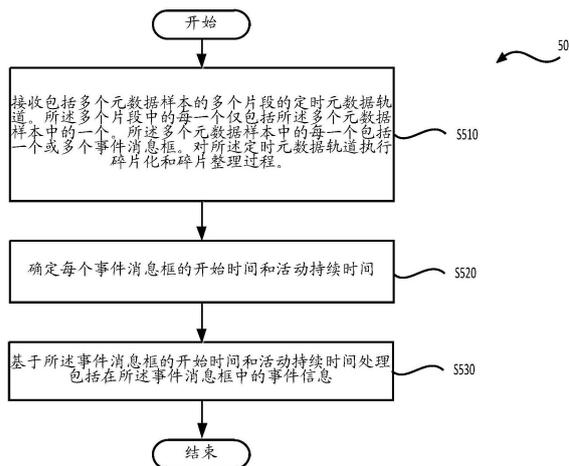
权利要求书2页 说明书12页 附图4页

(54) 发明名称

接收媒体数据的方法、装置和非易失性计算机可读存储介质

(57) 摘要

本公开的各方面提供了接收媒体数据的方法、装置和非易失性计算机可读存储介质。一种装置包括处理电路,所述处理电路接收包括多个元数据样本的多个片段的定时元数据轨道,所述多个片段中的每一个仅包括所述多个元数据样本中的一个,所述多个元数据样本中的每一个包括一个或多个事件消息框。对所述定时元数据轨道执行碎片化和碎片整理过程。所述处理电路确定每个事件消息框的开始时间和活动持续时间。所述处理电路基于所述事件消息框的开始时间和活动持续时间,处理包括在所述事件消息框中的事件信息。



1. 一种接收媒体数据的方法,其特征在于,所述方法包括:

接收包括多个元数据样本的多个片段的定时元数据轨道,所述多个片段中的每一个仅包括所述多个元数据样本中的一个,所述多个元数据样本中的每一个包括一个或多个事件消息框;

接收所述多个元数据样本中的每一个的开始呈现时间,所述开始呈现时间等于包括相应元数据样本的所述多个片段中的不同片段的开始呈现时间;

确定每个事件消息框的开始时间和活动持续时间;以及

基于所述事件消息框的开始时间和活动持续时间,处理包括在所述事件消息框中的事件信息,其中

对所述定时元数据轨道执行碎片化和碎片整理过程。

2. 根据权利要求1所述的方法,其特征在于,每个事件消息框的时标等于所述定时元数据轨道的时标。

3. 根据权利要求1所述的方法,其特征在于,所述多个元数据样本中的每一个的持续时间等于包括相应元数据样本的所述多个片段中的不同片段的持续时间。

4. 根据权利要求1所述的方法,其特征在于,所述多个元数据样本中的每一个的开始呈现时间和持续时间等于包括相应元数据样本的所述多个片段中的一个的最早呈现时间和持续时间。

5. 根据权利要求1所述的方法,其特征在于,所述多个元数据样本中的每一个的开始呈现时间是包括在所述相应元数据样本中的事件消息框的开始呈现时间值和开始呈现时间增量值中的一个的锚点。

6. 根据权利要求1-5任一项所述的方法,其特征在于,每个事件消息框的开始时间和活动持续时间的总和受所述定时元数据轨道的结束呈现时间的限制。

7. 根据权利要求1-5任一项所述的方法,其特征在于,所述多个片段中的每一个是通用媒体应用格式CMAF碎片或基于超文本传输协议的动态自适应流传输DASH片段。

8. 根据权利要求1-5任一项所述的方法,其特征在于,包括在所述多个元数据样本中的一个中的每个事件消息框包括不同的方案标识符。

9. 根据权利要求1-5任一项所述的方法,其特征在于,所述碎片化和碎片整理过程基于国际标准化组织ISO/国际电工技术委员会IEC基本媒体文件格式ISOBMFF碎片化和碎片整理过程。

10. 一种接收媒体数据的装置,其特征在于,所述装置包括处理电路,被配置为执行权利要求1-9任一项所述的方法。

11. 一种接收媒体数据的装置,其特征在于,所述装置包括:

接收模块,用于接收包括多个元数据样本的多个片段的定时元数据轨道,所述多个片段中的每一个仅包括所述多个元数据样本中的一个,所述多个元数据样本中的每一个包括一个或多个事件消息框;接收所述多个元数据样本中的每一个的开始呈现时间,所述开始呈现时间等于包括相应元数据样本的所述多个片段中的不同片段的开始呈现时间;

确定模块,用于确定每个事件消息框的开始时间和活动持续时间;以及

处理模块,用于基于所述事件消息框的开始时间和活动持续时间,处理包括在所述事件消息框中的事件信息,其中

对所述定时元数据轨道执行碎片化和碎片整理过程。

12. 一种非易失性计算机可读存储介质,其特征在於,用于存储指令,当所述指令由用于接收媒体数据的计算机执行时,使所述计算机执行权利要求1-9任一项所述的方法。

接收媒体数据的方法、装置和非易失性计算机可读存储介质

[0001] 通过引用并入本文

[0002] 本申请要求于2020年9月24日提交的、申请号为17/031,227、名称为“基于HTTP的动态自适应流传输的方法和装置”的美国专利申请的优先权,其要求于2019年9月30日提交的、申请号为62/908,238、名称为“具有嵌入事件消息框的单样本稀疏定时元数据DASH片段和CMAF片段”的美国临时申请的优先权。在先申请的全部公开内容通过引用整体并入本文。

技术领域

[0003] 本公开描述了总体上涉及基于超文本传输协议的动态自适应流传输的方法和装置的实施例。

背景技术

[0004] 本文所提供的背景描述旨在总体上呈现本公开的背景。在背景技术部分以及本说明书的各个方面中所描述的目前已署名的发明人的工作所进行的程度,并不表明其在本公开提交时作为现有技术,且从未明示或暗示其被承认为本公开的现有技术。

[0005] 基于超文本传输协议的运动图像专家组(MPEG)动态自适应流传输(DASH)提供了通过IP网络的流传输多媒体内容的标准。DASH标准允许在媒体段中携带事件消息框。

发明内容

[0006] 本公开的各方面提供了接收媒体数据的装置。一种装置包括处理电路,所述处理电路接收包括多个元数据样本的多个片段的定时元数据轨道,所述多个片段中的每一个仅包括所述多个元数据样本中的一个,所述多个元数据样本中的每一个包括一个或多个事件消息框。所述处理电路确定每个事件消息框的开始时间和活动持续时间。所述处理电路基于所述事件消息框的开始时间和活动持续时间,处理包括在所述事件消息框中的事件信息。对所述定时元数据轨道执行碎片化和碎片整理过程。

[0007] 在实施例中,每个事件消息框的时标等于所述定时元数据轨道的时标。

[0008] 在实施例中,所述多个元数据样本中的每一个的呈现时间和持续时间等于包括相应元数据样本的所述多个片段中的一个的最早呈现时间和持续时间。

[0009] 在实施例中,所述多个元数据样本中的每一个的呈现时间是包括在所述相应元数据样本中的事件消息框的呈现时间值和呈现时间增量值中的一个的锚点。

[0010] 在实施例中,每个事件消息框的开始时间和活动持续时间的总和受所述定时元数据轨道的结束呈现时间的限制。

[0011] 在实施例中,所述多个片段中的每一个是通用媒体应用格式(CMAF)碎片和基于超文本传输协议的动态自适应流传输(DASH)片段中的一个。

[0012] 在实施例中,包括在所述多个元数据样本中的一个中的每个事件消息框包括不同的方案标识符。

[0013] 在实施例中,所述碎片化和碎片整理过程基于国际标准化组织(ISO)/国际电工技

术委员会 (IEC) 基本媒体文件格式 (ISOBMFF) 碎片化和碎片整理过程。

[0014] 本公开的各方面提供了接收媒体数据的方法。在一种方法中,接收包括多个元数据样本的多个片段的定时元数据轨道,所述多个片段中的每一个仅包括所述多个元数据样本中的一个,所述多个元数据样本中的每一个包括一个或多个事件消息框。确定每个事件消息框的开始时间和活动持续时间。基于所述事件消息框的开始时间和活动持续时间处理包括在所述事件消息框中的事件信息。对所述定时元数据轨道执行碎片化和碎片整理过程。

[0015] 本公开的各方面还提供了非易失性计算机可读存储介质,用于存储指令,当所述指令由用于接收媒体数据的计算机执行时,使所述计算机执行所述接收媒体数据的方法的任意一种或组合。

附图说明

[0016] 根据以下详细描述和附图,所公开的主题的其他特征、性质和各种优点将进一步明确,其中:

[0017] 图1示出了根据本公开实施例的示例性的基于超文本传输协议的动态自适应流传输 (DASH) 系统。

[0018] 图2示出了根据本公开实施例的示例性DASH客户端架构。

[0019] 图3示出了根据本公开实施例的示例性带内事件定时模型。

[0020] 图4示出了根据本公开实施例的可包括在内容流中的媒体轨道和定时元数据轨道的示例。

[0021] 图5示出了根据一些实施例的概述过程示例的流程图。

[0022] 图6是根据实施例的计算机系统的示意图。

具体实施方式

[0023] I. 基于超文本传输协议的动态自适应流传输 (DASH) 和媒体呈现描述 (MPD, Media Presentation Description)

[0024] 基于超文本传输协议的动态自适应流传输 (DASH) 是一种自适应比特率流传输技术,它能够使用超文本传输协议 (HTTP) 基础设施,例如web服务器、内容分发网络 (CDN)、各种代理和缓存等,来实现媒体内容的流传输。DASH支持从DASH服务器到DASH客户端的点播和直播,并允许DASH客户端控制流传输会话,使得DASH服务器不需要在大规模部署中处理额外的流适配管理负载。DASH还允许DASH客户端从各种DASH服务器选择流传输,从而进一步实现网络的负载平衡,以利于DASH客户端。DASH提供不同媒体轨道之间的动态切换,例如,通过改变比特率来适应网络条件。

[0025] 在DASH中,媒体呈现描述 (MPD) 文件为DASH客户端提供信息,以便通过从DASH服务器下载媒体片段来自适应地流传输媒体内容。MPD文件可以碎片化并部分交付,以减少会话启动延迟。MPD文件还可以在流传输会话期间更新。在一些示例中,MPD文件支持内容可访问性特征、评级和相机视图的表示。DASH还支持多视图和可缩放编码内容的交付。

[0026] MPD文件可以包含一个或多个时段 (period) 的序列。该一个或多个时段中的每一个可以由MPD文件中的时段元素定义。MPD文件可以包括MPD的availableStartTime属性和

每个时段的开始 (start) 属性。对于具有动态类型的媒体呈现 (例如, 用于直播服务), 时段的开始属性和MPD属性availableStartTime的和以及媒体片段的持续时间可以以协调世界时 (UTC) 格式指示该时段的可用时间, 特别是对应时段中的每个呈现的第一个媒体片段。对于具有静态类型的媒体呈现 (例如, 用于点播服务), 第一个时段的开始属性可以为0。对于其它任何时段, 开始属性可以指定对应时段的开始时间相对于第一个时段的开始时间之间的时间偏移。每个时段可以延长到下一时段的开始, 或者如果是最后一个时段, 则可以延长到媒体呈现的结束。时段开始时间可以是精确的, 并反映因为播放先前所有时段的媒体而产生的实际定时。

[0027] 每个时段可以包含一个或多个适配集, 并且每个适配集可以包含相同媒体内容的一个或多个呈现。呈现可以是音频或视频数据的多个可选已编码版本中的一个。多个呈现可以因编码类型而不同, 例如因比特率、分辨率、和/或视频数据的编解码器以及比特率、和/或音频数据的编解码器而不同。术语呈现可用于指代与多媒体内容的特定时段相对应并以特定方式编码的一段已编码音频或视频数据。

[0028] 可以将特定时段的适配集分配给由MPD文件中的组属性指示的组。通常认为同一组中的适配集是相互可替代的。例如, 可以将特定时段的视频数据的每个适配集分配给同一组, 使得可以选择任何一个适配集进行解码, 以显示对应时段的多媒体内容的视频数据。在一些示例中, 一个时段内的媒体内容可以由来自组0的一个适配集 (如果存在) 来呈现, 或由来自每个非零组的至多一个适配集的组合来呈现。时段的每个呈现的定时数据可以相对于时段的开始时间来表示。

[0029] 一个呈现可以包括一个或多个片段。每个呈现可以包括初始化片段, 或者呈现的每个片段可以是自初始化的。当存在时, 初始化片段可以包含用于访问该呈现的初始化信息。在一些情况下, 初始化片段不包含媒体数据。一个片段可以由标识符唯一地引用, 例如统一资源定位符 (URL)、统一资源名称 (URN) 或统一资源标识符 (URI)。MPD文件可以为每个片段提供标识符。在一些示例中, MPD文件还可以以范围属性的形式提供字节范围, 该字节范围可以对应于URL、URN或URI可访问的文件内的片段的数据。

[0030] 每个呈现还可以包括一个或多个媒体组件, 其中每个媒体组件可以对应于一个单独媒体类型的已编码版本, 例如音频、视频或定时文本 (例如, 用于隐藏式字幕)。媒体组件可以是跨一个呈现内的连续媒体片段的边界时间连续的。

[0031] 在一些实施例中, DASH客户端可以从DASH服务器访问和下载MPD文件。也就是说, DASH客户端可以检索MPD文件, 以用于发起实时会话。基于MPD文件, 对于每个选定的呈现, DASH客户端可以做出多个决定, 包括确定服务器上可用的最新片段, 确定下一片段和可能的未来片段的片段可用性开始时间, 确定何时开始播放该片段以及从该片段中的哪个时间线开始播放, 以及确定何时获取/取得新的MPD文件。一旦播放了服务, 客户端就可以跟踪直播服务与其自己的播放之间的漂移, 该漂移需要检测和补偿。

[0032] II. 事件消息框

[0033] 国际标准化组织 (ISO) / 国际电工技术委员会 (IEC) 23009-1 DASH标准引入了事件消息框, 用于携带具有媒体片段的事件。ISO/IEC 23000-19通用媒体应用格式 (CMAF) 允许在每个CMAF块的开头包含事件消息框。已经讨论了作为CMAF稀疏元数据轨道 (sparse metadata track) 的一部分的事件消息的传输。然而, 出现了一个关于这种轨道是否满足

ISO/IEC 14496-12 ISO/IEC基本媒体文件格式 (ISOBMFF) 轨道的碎片化/碎片整理要求的问题。

[0034] 本公开包括针对事件信息 (例如包括在定时元数据轨道中的事件信息) 的信令和处理的实施例。事件信息可以对应于与媒体呈现 (例如, 连续的音频和/或视频呈现) 中的时间点或时段相关联的媒体定时事件。例如, 事件信息可用于动态内容替换、广告插入、与音频和/或视频一起呈现补充内容、对网页进行改变、以及执行在媒体呈现 (例如, 音频和/或视频媒体流) 的媒体时间线上的特定点处触发的应用代码。此外, 可以通过不同的方法提供事件信息。

[0035] 媒体定时事件可用于携带与媒体流同步的信息。例如, 事件信息可以包括描述媒体呈现内容的元数据 (或定时元数据), 例如节目或章节标题、或地理位置信息。此外, 事件信息可以包括媒体播放器的控制消息, 该控制信息与媒体呈现的回放期间的特定时间相关联, 例如广告插入提示。

[0036] 本公开的实施例可以在MPEG-DASH中实现。定时元数据轨道可以包括携带事件信息的嵌入式事件消息框 (emsg)。定时元数据轨道可用于携带与上述MPD事件和带内事件类似的信息。事件信息可以包括与其它轨道中提供的信息时间同步的元数据。由于这种同步, 可以不规则地或不连续地提供定时元数据轨道中的元数据, 并且可以将这种时间元数据轨道称为稀疏定时元数据轨道。

[0037] 当定时元数据轨道携带事件消息框作为元数据样本的一部分时, 每个事件消息框可以包括其自己的定时, 该定时使用通用媒体应用格式 (CMAF) 块或轨道 (包括事件消息框) 的最早呈现时间作为锚点。然而, ISO/IEC基本媒体文件格式 (ISOBMFF) 要求, 如果ISOBMFF轨道经过任何碎片化和碎片整理过程, 则保留数据的定时。由于碎片化和碎片整理过程中块或轨道边界丢失, 携带嵌入式事件的稀疏定时元数据轨道可能无法支持该要求。

[0038] 本公开提出了用于携带嵌入式事件的稀疏定时元数据轨道的方法, 以支持任何任意碎片化和碎片整理过程, 从而成为有效的ISOBMFF轨道。

[0039] 图1示出了根据本公开实施例的示例性DASH系统 (100)。在DASH系统 (100) 中, MPD文件从DASH服务器 (101) (例如, 内容服务器) 发送到DASH客户端 (102)。DASH客户端 (102) 可以基于MPD文件从DASH服务器 (101) 接收媒体片段。DASH客户端 (102) 可以向DASH服务器 (101) 发送更新MPD文件的请求。DASH服务器 (101) 可以提供内容流, 包括主要内容 (例如, 主节目) 和一个或多个定时元数据轨道。

[0040] 图2示出了根据本公开实施例的示例性DASH客户端架构。DASH客户端 (或DASH播放器) 可以被配置为与应用 (212) 通信并处理各种类型的事件, 包括 (i) MPD事件、(ii) 带内事件、以及 (iii) 定时元数据事件。

[0041] 清单解析器 (210) 可以解析清单 (例如, MPD)。例如, 清单可以由DASH服务器 (101) 提供。清单解析器 (210) 可以提取关于嵌入在定时元数据轨道中的MPD事件、带内事件和定时元数据事件的事件信息。可以将提取的事件信息提供给DASH逻辑 (211) (例如, DASH播放器控制、选择和启发式逻辑)。DASH逻辑 (211) 可以基于事件信息将清单中用信号通知的事件方案通知给应用 (212)。

[0042] 事件信息可以包括用于区分不同事件流的事件方案信息。应用 (212) 可以使用事件方案信息来订阅感兴趣的事件方案。应用 (212) 还可以通过一个或多个订阅应用程序接

口 (API) 为每个订阅方案指示期望的调度模式。例如,应用 (212) 可以向DASH客户端发送订阅请求,该订阅请求标识一个或多个感兴趣的事件方案和任何期望的对应调度模式。

[0043] 如果应用 (212) 订阅了作为一个或多个定时元数据轨道的一部分而被交付的一个或多个事件方案,则带内事件和‘moof’解析器 (203) 可以将一个或多个定时元数据轨道流传输到定时元数据轨道解析器 (204)。例如,带内事件和‘moof’解析器 (203) 解析电影碎片框(‘moof’),并随后基于来自DASH逻辑 (211) 的控制信息解析定时元数据轨道。

[0044] 定时元数据轨道解析器 (204) 可以提取嵌入在定时元数据轨道中的事件消息。提取的事件消息可以存储在事件和定时元数据缓冲器 (206) 中。同步器/调度器模块 (208) (例如,事件和定时元数据同步器和调度器) 可以将订阅的事件调度 (或发送) 到应用 (212)。

[0045] MPD中描述的MPD事件可以由清单解析器 (210) 解析并存储在事件和定时元数据缓冲器 (206) 中。例如,清单解析器 (210) 解析MPD的每个事件流元素,并解析每个事件流元素中描述的每个事件。对于在MPD中用信号通知的每个事件,诸如呈现时间和事件持续时间之类的事件信息可以存储在与事件相关联的事件和定时元数据缓冲器 (206) 中。

[0046] 带内事件和‘moof’解析器 (203) 可以解析媒体片段,以提取带内事件消息。任何这种识别的带内事件和相关联的呈现时间和持续时间都可以存储在事件和定时元数据缓冲器 (206) 中。

[0047] 因此,事件和定时元数据缓冲器 (206) 可以将MPD事件、带内事件和/或定时元数据事件存储其中。事件和定时元数据缓冲器 (206) 可以是例如先进先出 (FIFO) 缓冲器。事件和定时元数据缓冲器 (206) 可以对应于媒体缓冲器 (207) 进行管理。例如,只要媒体片段存在于媒体缓冲器 (207) 中,与该媒体片段相对应的任何事件或定时元数据都可以存储在事件和定时元数据缓冲器 (206) 中。

[0048] DASH访问API (202) 可以通过HTTP协议栈 (201) 管理包括媒体内容和各种元数据的内容流 (或数据流) 的获取和接收。DASH访问API (202) 可以将接收的内容流分成不同的数据流。提供给带内事件和‘moof’解析器 (203) 的数据流可以包括媒体片段、一个或多个定时元数据轨道、以及包括在媒体片段中的带内事件信令。在实施例中,提供给清单解析器 (210) 的数据流可以包括MPD。

[0049] DASH访问API (202) 可以将清单转发给清单解析器 (210)。除了描述事件之外,清单还可以向DASH逻辑 (211) 提供关于媒体片段的信息,该DASH逻辑 (211) 可以与应用 (212) 以及带内事件和‘moof’解析器 (203) 通信。应用 (212) 可以与DASH客户端处理的媒体内容相关联。在应用 (212)、DASH逻辑 (211)、清单解析器 (210) 和DASH访问API (202) 之间交换的控制/同步信号可以基于清单中提供的关于媒体片段的信息控制从HTTP协议栈 (201) 获取媒体片段。

[0050] 带内事件和‘moof’解析器 (203) 可以将媒体数据流解析为媒体片段,包括媒体内容、定时元数据轨道中的定时元数据、以及媒体片段中的任何用信号通知的带内事件。包括媒体内容的媒体片段可以由文件格式解析器 (205) 解析,并存储在媒体缓冲器 (207) 中。

[0051] 存储在事件和定时元数据缓冲器 (206) 中的事件可以允许同步器/调度器 (208) 通过事件/元数据API向应用 (212) 传送与应用 (212) 相关的可用事件 (或感兴趣的事件)。应用 (212) 可以被配置为处理该可用事件 (例如,MPD事件、带内事件或定时元数据事件),并通过通知同步器/调度器 (208) 来订阅特定事件或定时元数据。存储在事件和定时元数据缓冲器

(206)中的与应用(212)不相关但与DASH客户端本身相关的任何事件都可以由同步器/调度器(208)转发到DASH逻辑(211)进行进一步处理。

[0052] 响应于应用(212)订阅特定事件,同步器/调度器(208)可以向应用(212)传送与应用(212)已订阅的事件方案相对应的事件实例(或定时元数据样本)。事件实例可以根据订阅请求指示的调度模式(例如,针对特定事件方案)或默认调度模式来传送。例如,在接收(on-receive)调度模式中,事件实例可以在事件和定时元数据缓冲器(206)中接收时被发送到应用(212)。另一方面,在启动(on-start)调度模式中,事件实例可以在其相关联的呈现时间(例如与来自媒体解码器(209)的定时信号同步)被发送到应用(212)。

[0053] 应注意,在DASH客户端架构中,粗数据流线表示媒体数据流,窄数据流线表示事件和定时元数据数据流,虚线数据流线表示控制和同步。此外,相同的处理模型可用于CMAF事件。

[0054] 图3示出了根据本公开实施例的事件消息框的示例性定时模型。在定时模型中,每个事件消息框可以由媒体时间线上的三个定时参数来描述:(i)事件到达时间(AT),它是包括事件消息框的片段的最早呈现时间;(ii)事件呈现/开始时间(ST),它是媒体(MPD)时间线中事件变为活动状态的时刻;以及(iii)事件持续时间(DU),在该事件持续时间期间事件处于活动状态。

[0055] 事件消息框可以插入在媒体片段的开头。因此,携带事件消息框的媒体片段的最早呈现时间可以认为是事件消息框在媒体时间线上的位置。DASH客户端可以在事件消息框的AT之前或AT处获取和解析媒体片段。

[0056] 事件消息框的ST可以与轨道中携带事件消息框的媒体片段(例如,DASH片段或CMAF块)的位置有偏移。ST的锚点可以根据事件消息框的版本而不同。对于版本0的事件消息框,锚点可以是携带事件消息框的CMAF碎片/DASH片段的最早呈现时间。对于版本1的事件消息框,锚点可以是携带事件消息框的CMAF轨道/DASH时段的最早呈现时间。

[0057] 根据ISO/IEC 23000-19,事件消息框的时标(timescale)可以等于CMAF轨道的MediaHeaderBox中的时标。事件消息框可以插入在CMAF块、CMAF碎片或CMAF片段的开头。CMAF碎片中的版本0的事件消息框的锚点是该片段的最早呈现时间。此外,根据ISO/IEC 23009-1,事件消息框可以放置在片段的第一个‘moof’框之前,或者可以放置在任何媒体数据框(‘mdat’)和‘moof’框之间。在后一种情况下,具有相同id值的等效‘emsg’应该出现在任何片段的第一个‘moof’框之前。

[0058] III. 碎片化/碎片整理过程

[0059] ISOBMFF的碎片化/碎片整理过程可以对具有任意大小碎片的ISOBMFF轨道进行碎片化和碎片整理。在任一过程中,生成的非碎片化或碎片化轨道应该是有效的ISOBMFF轨道。在这种情况下,包括在新轨道中的所有事件消息框都必须保持正确的定时。

[0060] 根据本公开的各方面,可以以元数据样本为单位对元数据轨道执行碎片化/碎片整理过程。因此,在碎片化/碎片整理过程期间,可以保持单个元数据样本的正确定时。然而,在片段包括多个元数据样本的情况下,在碎片化/碎片整理过程期间可能无法保持该片段的正确定时。此外,如果该片段包括版本0的事件消息框,并且在碎片化/碎片整理过程中没有保持该片段的正确定时,则由于版本0的事件消息框的ST的锚点是该片段的最早呈现时间,因此可能无法保持事件消息框的正确定时。

[0061] CMAF片段/DASH片段在任何任意碎片化和碎片整理下的生存性可以分析如下。

[0062] 如果DASH片段/CMAF碎片包括版本0的事件消息框,则由于事件消息框的锚点是DASH片段/CMAF片段的最早呈现时间并且在DASH/CMAF媒体轨道的碎片化/碎片整理过程期间可能丢失,因此碎片化/碎片整理过程可能失败。

[0063] 如果DASH片段/CMAF碎片包括版本1的事件消息框,则由于事件消息框的锚点是DASH/CMAF媒体轨道的最早呈现时间并且在DASH/CMAF媒体轨道的碎片化/碎片整理过程期间被保留,因此碎片化/碎片整理过程可以通过。

[0064] 表1总结了碎片化/碎片整理过程的DASH和CMAF轨道的状态。

[0065] 表1

轨道	碎片化	碎片整理
DASH w/emsg V0	X	X
DASH w/emsg V1	OK	OK
CMAF w/emsg V0	X	X
CMAF w/emsg V1	OK	OK

[0067] 图4示出了根据本公开实施例的可包括在内容流中的媒体轨道和定时元数据轨道的示例。元数据轨道可以包括事件消息框。事件消息框可用于为与媒体呈现时间相关的一般事件提供信令。在一些实施例中,如果DASH客户端检测到具有未在MPD中定义的方案的事件消息框,则DASH客户端忽略该事件消息框。

[0068] 事件消息框可以包括消息方案标识信息(例如,scheme_id_uri)、包含在事件消息框中的事件的可选项、定时信息和事件数据。定时信息可以指示其他时间信息的时标(例如,以每秒的滴答数为单位),例如事件的媒体呈现时间相对于参考呈现时间(例如,片段或元数据样本的开始)的媒体呈现时间增量、事件的媒体呈现时间、以及事件持续时间(例如,媒体呈现时间中)。

[0069] 元数据轨道可以携带事件消息框,作为包括在元数据轨道中的元数据样本的一部分。此外,元数据样本可以包括一个或多个事件消息框。每个事件消息框可以属于一个事件方案,该事件方案由方案URI id和相应事件消息框的可选项定义。由于来自多个方案的事件实例可以包括在一个或多个元数据样本中,因此需要在DASH清单中标识事件方案,以便DASH客户端发现这些方案。

[0070] DASH包括两个元素,可用于描述MPD中的事件方案。这两个元素是用于MPD事件的事件流元素(例如,EventStream)和用于带内事件的带内事件流元素(例如,InbandEventStream)。这两个事件方案可以使用相同的构造。

[0071] IV. 单样本稀疏定时元数据

[0072] 根据本公开的各方面,包括嵌入式事件消息框的元数据片段/碎片可以是单样本DASH片段/CMAF碎片或多样本DASH片段/CMAF碎片。单样本DASH片段/CMAF碎片只能包括一个元数据样本,并且元数据样本的持续时间等于DASH片段/CMAF碎片的持续时间。多样本DASH片段/CMAF碎片可以包括多个元数据样本。

[0073] 如果单样本DASH片段/CMAF碎片包括版本0的事件消息框,则碎片化/碎片整理过程可以通过。由于片段/碎片的最早呈现时间与包括在片段/碎片中的唯一一个元数据样本的呈现时间相同,因此如果事件消息框的锚点被认为是包括事件消息框的元数据样本的呈

现时间,则可以在碎片化/碎片整理过程期间保留事件消息框的定时。

[0074] 如果单样本DASH片段/CMAF碎片包括版本1的事件消息框,则由于轨道的最早呈现时间是事件消息框的锚点,因此碎片化/碎片整理过程可以通过。

[0075] 如果多样本DASH片段/CMAF碎片包括版本0的事件消息框,则由于事件消息框的锚点是片段/碎片的最早呈现时间并且在碎片化/碎片整理过程期间可能丢失,因此碎片化/碎片整理过程可能失败。

[0076] 如果多样本DASH片段/CMAF碎片包括版本1的事件消息框,则由于轨道的最早呈现时间是事件消息框的锚点,因此碎片化/碎片整理过程可以通过。

[0077] 表2总结了碎片化/碎片整理过程中具有嵌入式事件消息框的单样本和多样本元数据片段/碎片的状态。

[0078] 表2

	轨道	碎片化	碎片整理
[0079]	每片段单样本 w/ emsg V0	OK*	OK*
	每片段单样本 w/ emsg V1	OK	OK
	每片段多样本 w/ emsg V0	X	X
	每片段多样本 w/ emsg V1	OK	OK
*在锚点参考为携带事件的样本的呈现时间的情况下。			

[0080] 根据本公开的各方面,可以在CMAF稀疏元数据轨道上应用以下约束,以满足ISOBMFF碎片化/碎片整理过程:(i) 每个CMAF碎片/DASH片段仅包括一个元数据样本(或限于一个元数据样本),并且元数据样本的持续时间是碎片/片段的持续时间;(ii) CMAF碎片/DASH片段的最早呈现时间是元数据样本的呈现时间;(iii) 在非碎片化轨道中,版本0的每个事件消息框可以使用元数据样本(包括各个事件消息框)的呈现时间作为相应事件消息框的呈现时间偏移参数(例如,presentation_time_delta)的锚点;(iv) 在非碎片化轨道中,版本1的每个事件消息框可以使用轨道的最早呈现时间作为相应事件消息框的呈现时间参数(例如,presentation_time)的锚点;(v) 在所有情况下(例如,DASH或CMAF,碎片化或非碎片化),每个事件消息框(例如,版本0或版本1)的时标可以等于轨道的时标;以及(vi) 事件消息框的结束时间(例如,事件开始时间加上事件持续时间)不超过轨道的结束(或最后)呈现时间,即使事件持续时间的值表示事件消息框的结束时间超出轨道的结束呈现时间。

[0081] 因为携带事件消息框的元数据样本的呈现时间在任意碎片化/碎片整理过程期间不会改变,因此约束(i)和(ii)可以允许轨道的碎片化/碎片整理,而不改变版本0的事件消息框的定时。

[0082] 约束(iii)和(iv)是对非碎片化轨道的约束,允许维护碎片化/分段的单样本稀疏元数据轨道和非碎片化轨道之间的关系。

[0083] 约束(v)可以确保事件的定时与轨道样本定时对齐,以便在文件格式解析器(205)和媒体解码器(209)使用整数算法时不存在分数漂移。

[0084] 约束(vi)可以将事件持续时间限制为最大轨道持续时间,因此可以简化文件格式解析器(205)中的操作,尤其是因为应用事件必须传递给不一定具有轨道持续时间的应用(212)。因此,如果将每个事件消息框的持续时间设置为使得事件消息框在轨道之前或在轨

道结束时结束,则如果在将事件传递给应用(212)之前事件持续时间超过轨道持续时间,则文件格式解析器(205)不需要截断事件持续时间。

[0085] 单样本稀疏定时元数据轨道的一些优点可以列出如下。因为每个碎片/片段仅包括一个(同步)元数据样本,因此单样本稀疏定时元数据轨道的结构很简单。每个元数据样本可以包括一个或多个具有各种方案/子方案的事件消息框。如果碎片化/碎片整理过程保持单样本碎片/片段约束,则碎片化轨道可以通过碎片化/碎片整理过程。对于事件消息框在非碎片化轨道中的锚点的约束非常简单且易于维护。可以保持事件消息框的属性,以便使用与媒体轨道分离的轨道来传递,从而传递不依赖于任何特定的媒体轨道。

[0086] 根据本公开的各方面,可以生成稀疏定时元数据轨道,其中每个片段/碎片仅包括一个元数据样本。元数据样本可以包括一个或多个事件消息框。每个事件消息框可以包括不同的方案标识符和携带消息数据负载的相关子方案标识符的不同值。事件消息框的时标等于轨道的时标。元数据样本的呈现时间和持续时间等于包括元数据样本的片段/片段的最新呈现时间和持续时间。在碎片整理的情况下,事件消息框的呈现时间(例如,presentation_time)和/或呈现时间偏移(例如,presentation_time_delta)的锚点是包括事件消息框的元数据样本的呈现时间。

[0087] V. 流程图

[0088] 图5示出了根据本公开实施例的概述过程(500)的流程图。在各种实施例中,过程(500)由处理电路执行,例如DASH客户端(102)中的处理电路。在一些实施例中,过程(500)以软件指令实现,因此当处理电路执行该软件指令时,处理电路执行过程(500)。过程(500)开始于(S510),其中过程(500)接收包括多个元数据样本的多个片段的定时元数据轨道。所述多个片段中的每一个可以仅包括所述多个元数据样本中的一个。所述多个元数据样本中的每一个包括一个或多个事件消息框。然后,过程(500)进行到步骤(S520)。

[0089] 在步骤(S520),过程(500)确定每个事件消息框的开始时间和活动持续时间。然后,过程(500)进行到步骤(S530)。

[0090] 在步骤(S530),过程(500)基于所述事件消息框的开始时间和活动持续时间处理包括在所述事件消息框中的事件信息。然后,过程(500)结束。

[0091] 在实施例中,每个事件消息框的时标等于所述定时元数据轨道的时标。

[0092] 在实施例中,所述多个元数据样本中的每一个的呈现时间和持续时间等于包括相应元数据样本的所述多个片段中的一个的最早呈现时间和持续时间,对所述定时元数据轨道执行碎片化和碎片整理过程。

[0093] 在实施例中,所述多个元数据样本中的每一个的呈现时间是包括在所述相应元数据样本中的事件消息框的呈现时间值和呈现时间增量值中的一个的锚点。

[0094] 在实施例中,每个事件消息框的开始时间和活动持续时间的总和受所述定时元数据轨道的结束呈现时间的限制。

[0095] 在实施例中,所述多个片段中的每一个是CMAF碎片和DASH片段中的一个。

[0096] 在实施例中,包括在所述多个元数据样本中的一个中的每个事件消息框包括不同的方案标识符。

[0097] 在实施例中,所述碎片化和碎片整理过程基于ISO/IEC ISOBMFF碎片化和碎片整理过程。

[0098] VI. 计算机系统

[0099] 上述技术可以通过计算机可读指令实现为计算机软件,并且物理地存储在一个或多个计算机可读介质中。例如,图6示出了计算机系统(600),其适于实现所公开主题的某些实施例。

[0100] 所述计算机软件可通过任何合适的机器代码或计算机语言进行编码,通过汇编、编译、链接等机制创建包括指令的代码,所述指令可由一个或多个计算机中央处理单元(CPU),图形处理单元(GPU)等直接执行或通过译码、微代码等方式执行。

[0101] 所述指令可以在各种类型的计算机或其组件上执行,包括例如个人计算机、平板电脑、服务器、智能手机、游戏设备、物联网设备等。

[0102] 图6所示的用于计算机系统(600)的组件本质上是示例性的,并不用于对实现本公开实施例的计算机软件的使用范围或功能进行任何限制。也不应将组件的配置解释为与计算机系统(600)的示例性实施例中所示的任一组件或其组合具有任何依赖性 or 要求。

[0103] 计算机系统(600)可以包括某些人机界面输入设备。这种人机界面输入设备可以通过触觉输入(如:键盘输入、滑动、数据手套移动)、音频输入(如:声音、掌声)、视觉输入(如:手势)、嗅觉输入(未示出),对一个或多个人类用户的输入做出响应。所述人机界面设备还可用于捕获某些媒体,气与人类有意识的输入不必直接相关,如音频(例如:语音、音乐、环境声音)、图像(例如:扫描图像、从静止影像相机获得的摄影图像)、视频(例如二维视频、包括立体视频的三维视频)。

[0104] 人机界面输入设备可包括以下中的一个或多个(仅绘出其中一个):键盘(601)、鼠标(602)、触控板(603)、触摸屏(610)、数据手套(未示出)、操纵杆(605)、麦克风(606)、扫描仪(607)和照相机(608)。

[0105] 计算机系统(600)还可以包括某些人机界面输出设备。这种人机界面输出设备可以通过例如触觉输出、声音、光和嗅觉/味觉来刺激一个或多个人类用户的感受。这样的人机界面输出设备可包括触觉输出设备(例如通过触摸屏(610)、数据手套(未示出)或操纵杆(605)的触觉反馈,但也可以有不用作输入设备的触觉反馈设备)、音频输出设备(例如,扬声器(609)、耳机(未示出))、视觉输出设备(例如,包括阴极射线管屏幕、液晶屏幕、等离子屏幕、有机发光二极管屏的屏幕(610),其中每一个都具有或没有触摸屏输入功能、每一个都具有或没有触觉反馈功能——其中一些可通过诸如立体画面输出的手段输出二维视觉输出或三维以上的输出;虚拟现实眼镜(未示出)、全息显示器和放烟箱(未示出))以及打印机(未示出)。这些视觉输出设备(例如触摸屏(610))可以通过图形适配器(650)连接到系统总线(648)。

[0106] 计算机系统(600)还可以包括人可访问的存储设备及其相关介质,如包括具有CD/DVD的高密度只读/可重写式光盘(CD/DVD ROM/RW)(620)或类似介质(621)的光学介质、拇指驱动器(622)、可移动硬盘驱动器或固体状态驱动器(623),诸如磁带和软盘(未示出)的传统磁介质,诸如安全软件保护器(未示出)等的基于ROM/ASIC/PLD的专用设备,等等。

[0107] 本领域技术人员还应当理解,结合所公开的主题使用的术语“计算机可读介质”不包括传输介质、载波或其它瞬时信号。

[0108] 计算机系统(600)还可以包括通往一个或多个通信网络(655)的网络接口(654)。例如,该一个或多个通信网络(655)可以是无线的、有线的、光学的。该一个或多个通信网络

(655) 还可为局域网、广域网、城域网、车载网络和工业网络、实时网络、延迟容忍网络等等。该一个或多个通信网络 (655) 的示例还包括以太网、无线局域网、蜂窝网络 (GSM、3G、4G、5G、LTE等) 等局域网、电视有线或无线广域数字网络 (包括有线电视、卫星电视、和地面广播电视)、车载和工业网络 (包括CANBus) 等等。某些网络通常需要外部网络接口适配器, 用于连接到某些通用数据端口或外围总线 (649) (例如, 计算机系统 (600) 的USB端口); 其它系统通常通过连接到如下所述的系统总线集成到计算机系统 (600) 的核心 (例如, 以太网接口集成到PC计算机系统或蜂窝网络接口集成到智能电话计算机系统)。通过使用这些网络中的任何一个, 计算机系统 (600) 可以与其它实体进行通信。所述通信可以是单向的, 仅用于接收 (例如, 无线电视), 单向的仅用于发送 (例如CAN总线到某些CAN总线设备), 或双向的, 例如通过局域或广域数字网络到其它计算机系统。上述的每个网络和网络接口可使用某些协议和协议栈。

[0109] 上述的人机界面设备、人可访问的存储设备以及网络接口可以连接到计算机系统 (600) 的核心 (640)。

[0110] 核心 (640) 可包括一个或多个中央处理单元 (CPU) (641)、图形处理单元 (GPU) (642)、以现场可编程门阵列 (FPGA) (643) 形式的专用可编程处理单元、用于特定任务的硬件加速器 (644) 等。这些设备以及只读存储器 (ROM) (645)、随机存取存储器 (646)、内部大容量存储器 (例如内部非用户可存取硬盘驱动器、固态硬盘等) (647) 等可通过系统总线 (648) 进行连接。在某些计算机系统中, 可以以一个或多个物理插头的形式访问系统总线 (648), 以便可通过额外的中央处理单元、图形处理单元等进行扩展。外围装置可直接附接到核心的系统总线 (648), 或通过外围总线 (649) 进行连接。外围总线的体系结构包括外部外围组件互联PCI、通用串行总线USB等。

[0111] CPU (641)、GPU (642)、FPGA (643) 和加速器 (644) 可以执行某些指令, 这些指令组合起来可以构成上述计算机代码。该计算机代码可以存储在ROM (645) 或RAM (646) 中。过渡数据也可以存储在RAM (646) 中, 而永久数据可以存储在例如内部大容量存储器 (647) 中。通过使用高速缓冲存储器可实现对任何存储器设备的快速存储和检索, 高速缓冲存储器可与一个或多个CPU (641)、GPU (642)、大容量存储器 (647)、ROM (645)、RAM (646) 等紧密关联。

[0112] 所述计算机可读介质上可具有计算机代码, 用于执行各种计算机实现的操作。介质和计算机代码可以是为本公开的目的而特别设计和构造的, 也可以是计算机软件领域的技术人员所熟知和可用的介质和代码。

[0113] 作为实施例而非限制, 具有体系结构 (600) 的计算机系统, 特别是核心 (640), 可以作为处理器 (包括CPU、GPU、FPGA、加速器等) 提供执行包含在一个或多个有形的计算机可读介质中的软件的功能。这种计算机可读介质可以是与上述的用户可访问的大容量存储器相关联的介质, 以及具有非易失性的核心 (640) 的特定存储器, 例如核心内部大容量存储器 (647) 或ROM (645)。实现本公开的各种实施例的软件可以存储在这种设备中并且由核心 (640) 执行。根据特定需要, 计算机可读介质可包括一个或一个以上存储设备或芯片。该软件可以使得核心 (640) 特别是其中的处理器 (包括CPU、GPU、FPGA等) 执行本文所述的特定过程或特定过程的特定部分, 包括定义存储在RAM (646) 中的数据结构以及根据软件定义的过程来修改这种数据结构。另外或作为替代, 计算机系统可以提供逻辑硬连线或以其它方式包含在电路 (例如, 加速器 (644)) 中的功能, 该电路可以代替软件或与软件一起运行以执行

本文所述的特定过程或特定过程的特定部分。在适当的情况下,对软件的引用可以包括逻辑,反之亦然。在适当的情况下,对计算机可读介质的引用可包括存储执行软件的电路(如集成电路(IC)),包含执行逻辑的电路,或两者兼备。本公开包括任何合适的硬件和软件组合。

[0114] 虽然本公开已对多个示例性实施例进行了描述,但实施例的各种变更、排列和各种等同替换均属于本公开的范围之内。因此应理解,本领域技术人员能够设计多种系统和方法,所述系统和方法虽然未在本文中明确示出或描述,但其体现了本公开的原则,因此属于本公开的精神和范围之内。

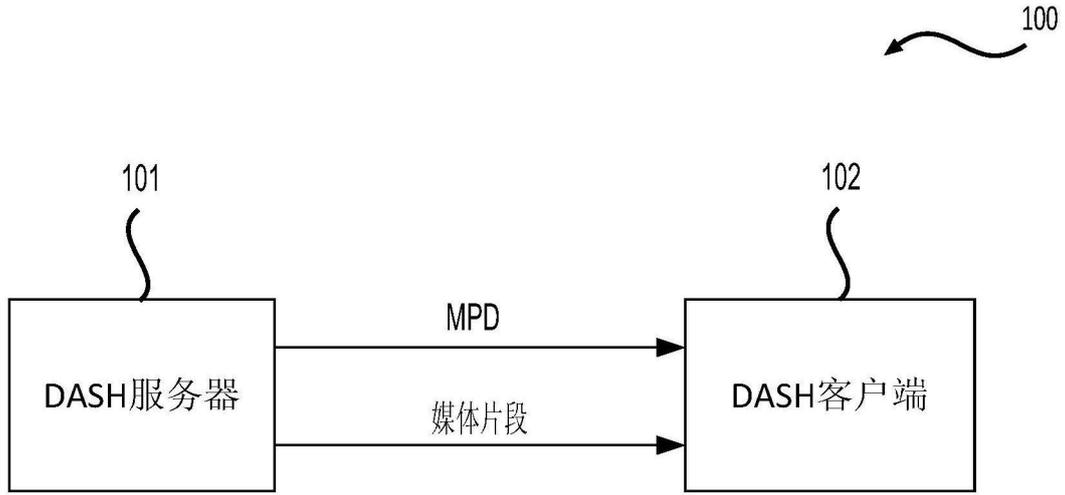


图1

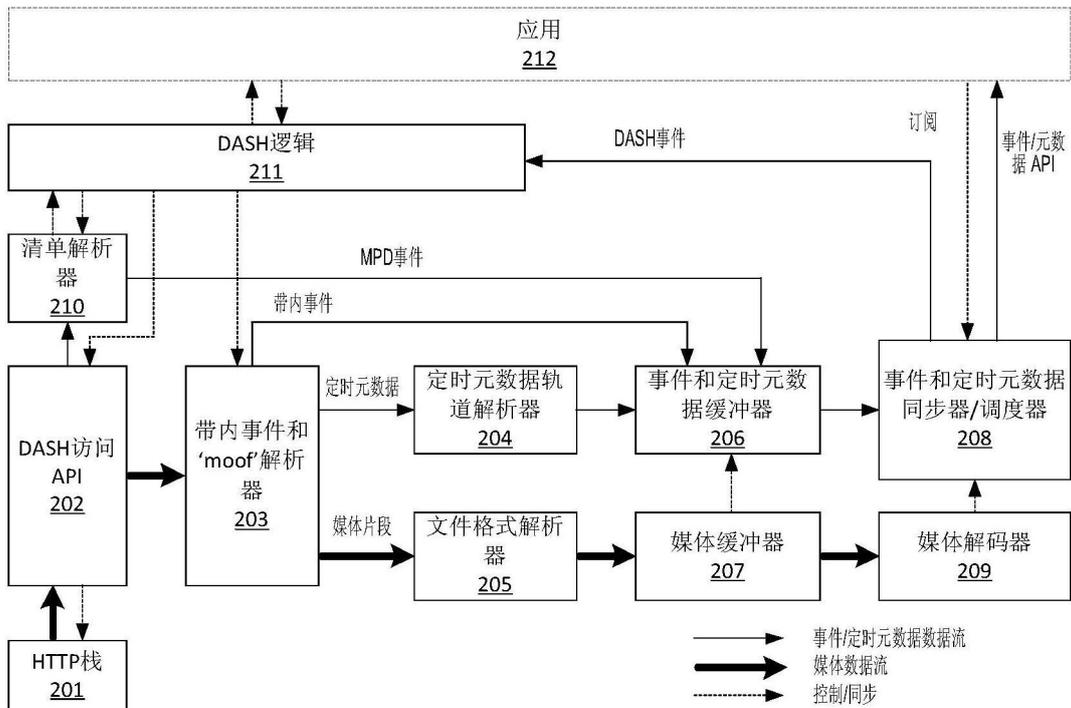


图2

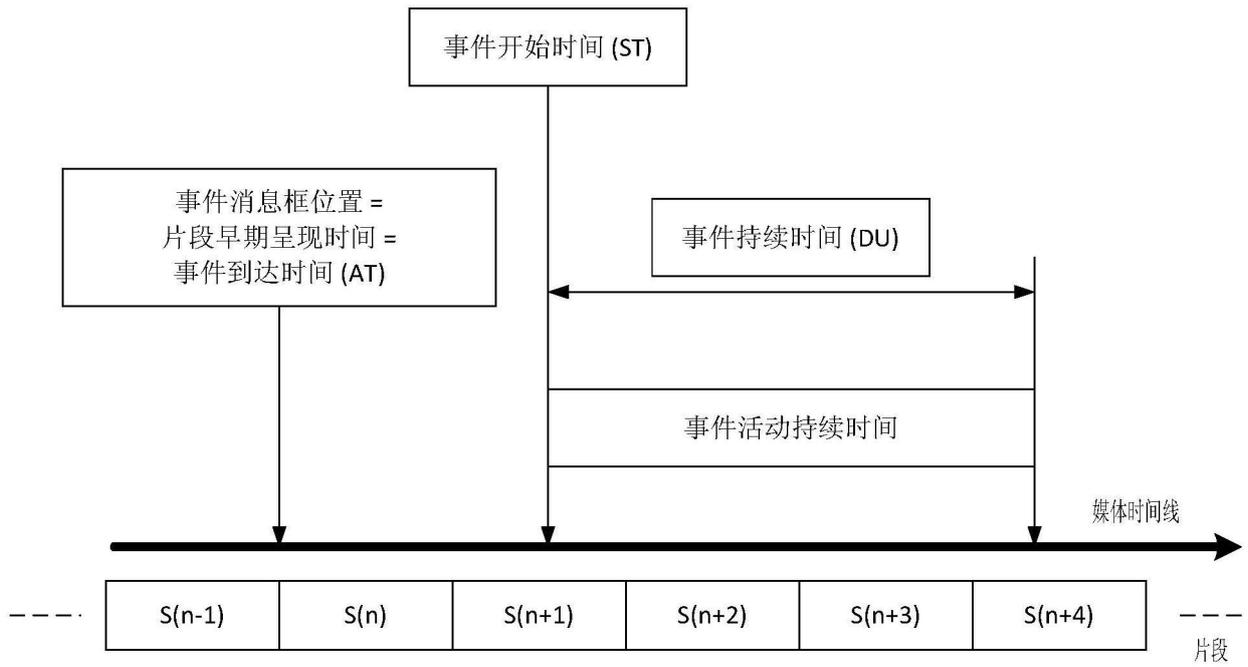


图3

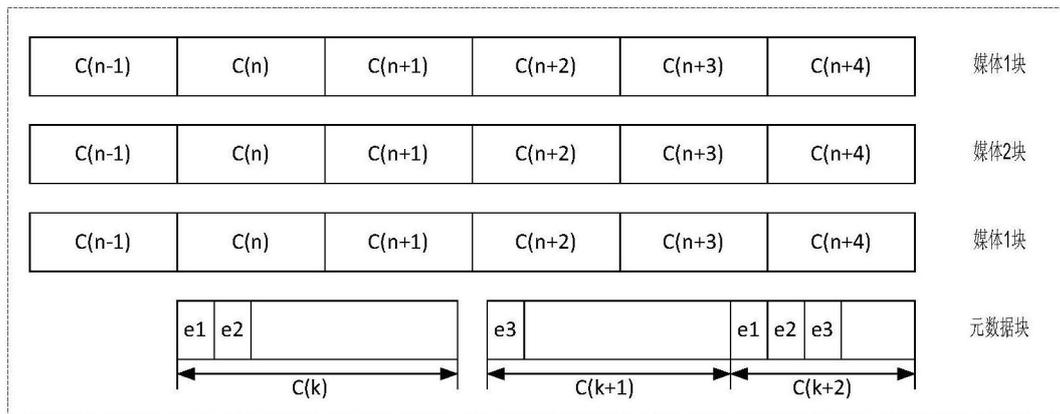


图4

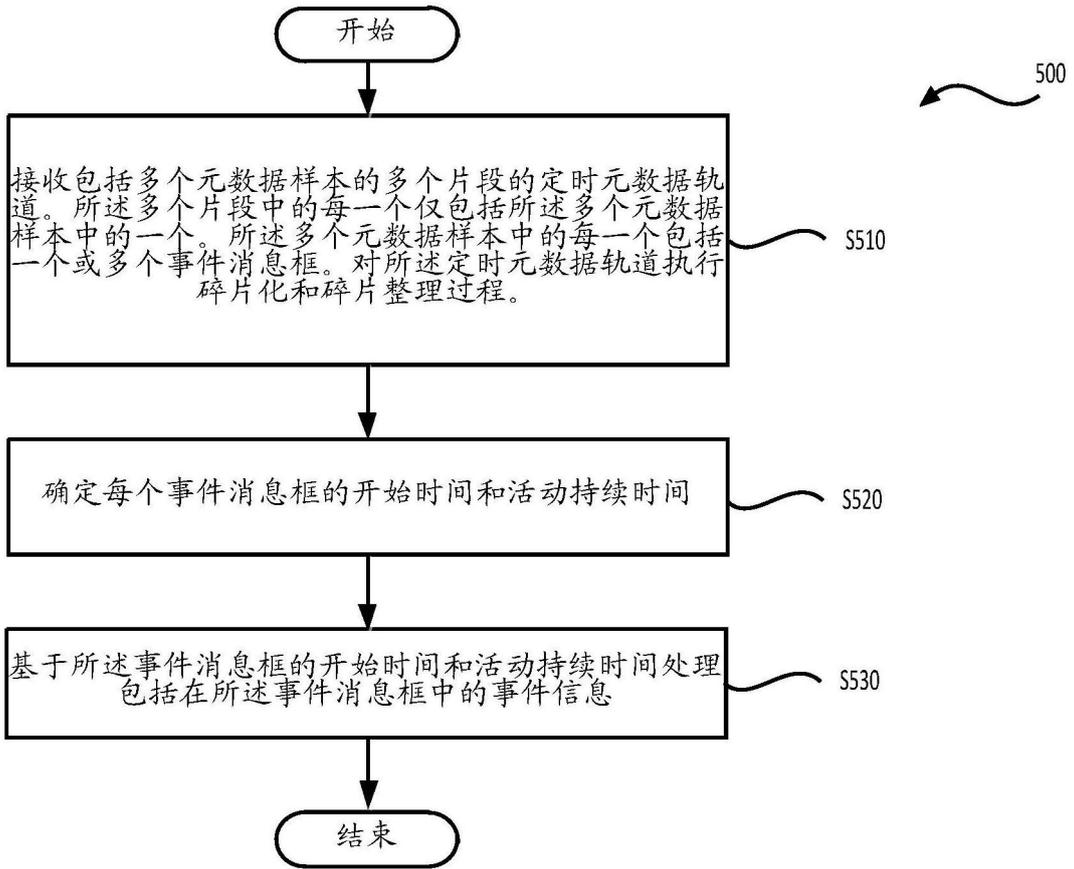


图5

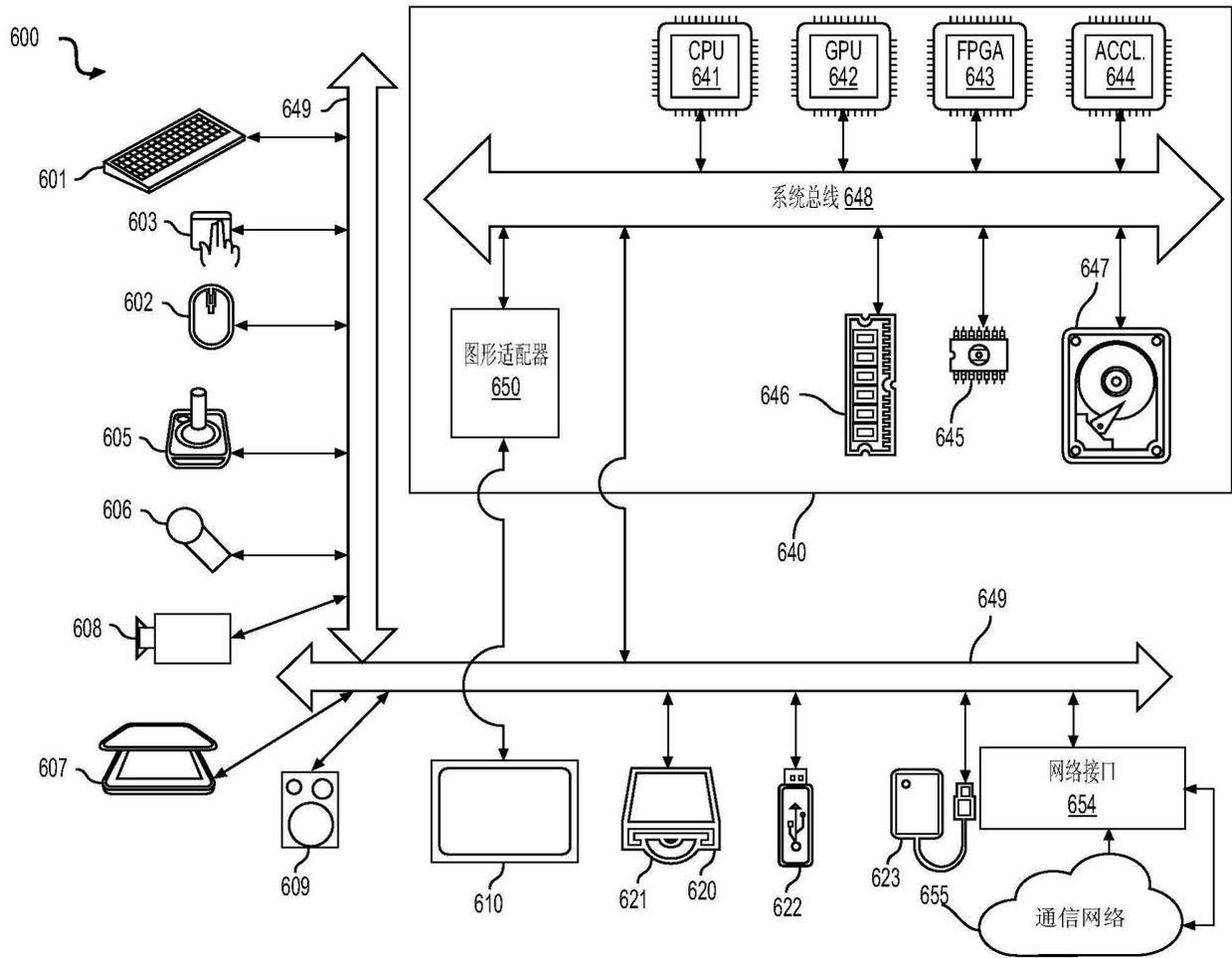


图6