



(12)发明专利

(10)授权公告号 CN 105052150 B

(45)授权公告日 2018.06.08

(21)申请号 201480017438.9

(22)申请日 2014.03.28

(65)同一申请的已公布的文献号
申请公布号 CN 105052150 A

(43)申请公布日 2015.11.11

(30)优先权数据
61/806,705 2013.03.29 US
14/228,139 2014.03.27 US

(85)PCT国际申请进入国家阶段日
2015.09.22

(86)PCT国际申请的申请数据
PCT/US2014/032235 2014.03.28

(87)PCT国际申请的公布数据
W02014/160971 EN 2014.10.02

(73)专利权人 高通股份有限公司
地址 美国加利福尼亚州

(72)发明人 王益魁

(74)专利代理机构 北京律盟知识产权代理有限
责任公司 11287

代理人 宋献涛

(51)Int.Cl.
H04N 19/70(2006.01)
H04N 19/88(2006.01)
H04N 21/44(2006.01)

(56)对比文件
CN 1934865 A,2007.03.21,
CN 101390399 A,2009.03.18,
Y.-K. Wang et al.RTP Payload Format
for H.264 Video.《Internet Engineering
Task Force (IETF)》.2011,第4.1、5.5、8.1节.
T. Schierl et al.RTP Payload Format
for High Efficiency Video Coding draft-
schierl-payload-rtp-h265-00.txt.《Internet
Engineering Task Force (IETF)》.2012,全文.

审查员 冯雪飞

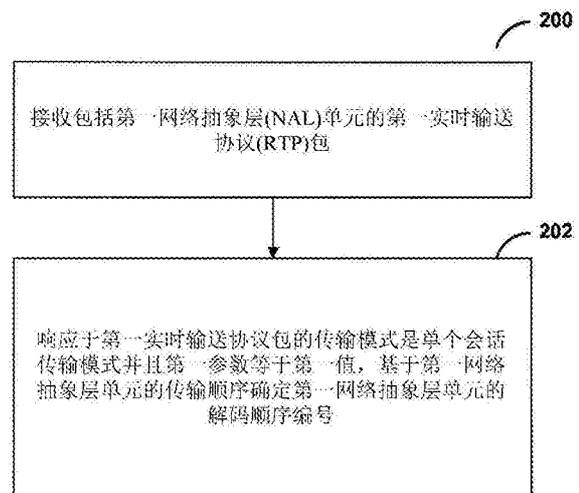
权利要求书3页 说明书20页 附图10页

(54)发明名称

改进的有效负载格式设计

(57)摘要

一种用于处理视频数据的装置包含:存储器;经配置以接收实时输送协议RTP包的接收器;以及一或多个处理器,其经配置以:接收第一实时输送协议RTP包,其包括第一网络抽象层NAL单元;以及响应于所述第一RTP包的传输模式是单个会话传输模式并且第一参数等于第一值,基于所述第一NAL单元的传输顺序确定所述第一NAL单元的解码顺序编号。



1. 一种处理视频数据的方法,所述方法包括:

接收第一实时传输协议RTP包,其包括第一网络抽象层NAL单元;

接收传输模式参数,其指示所述RTP包的传输模式;

确定与所述第一RTP包相关联的第一参数是否等于第一值,其中所述第一参数指明按照传输顺序在解包缓冲器中在所述第一NAL单元之前并且按照解码顺序在所述第一NAL单元之后的NAL单元的最大数目;以及

响应于所述第一RTP包的所述传输模式是单个会话传输模式并且所述第一参数等于所述第一值,基于所述第一NAL单元的传输顺序确定所述第一NAL单元的解码顺序编号。

2. 根据权利要求1所述的方法,其中所述第一值等于零。

3. 根据权利要求1所述的方法,其中所述最大数目等于32767。

4. 根据权利要求1所述的方法,其中所述解码顺序编号等于传输顺序编号。

5. 根据权利要求1所述的方法,其进一步包括:

接收第二RTP包,其包括第二NAL单元;

响应于所述第二RTP包的传输模式是多会话传输模式,并且响应于所述第二NAL单元按照传输顺序是第一个NAL单元,确定所述第二NAL单元按照解码顺序是第一个。

6. 根据权利要求1所述的方法,其进一步包括:

接收第二RTP包,其包括第二NAL单元;

响应于所述第二NAL单元的解码顺序编号等于所述第一NAL单元的所述解码顺序编号,在对所述第一NAL单元进行解码之前对所述第二NAL单元进行解码。

7. 根据权利要求1所述的方法,其中所述RTP包包括单个NAL单元包。

8. 一种用于处理视频数据的装置,所述装置包括:

经配置以存储所述视频数据的存储器;

经配置以接收实时传输协议RTP包的接收器;

一或多个处理器,其经配置以:

接收第一实时传输协议RTP包,其包括所述视频数据的第一网络抽象层NAL单元;

接收传输模式参数,其指示所述RTP包的传输模式;

确定与所述第一RTP包相关联的第一参数是否等于第一值,其中所述第一参数指明按照传输顺序在解包缓冲器中在所述第一NAL单元之前并且按照解码顺序在所述第一NAL单元之后的NAL单元的最大数目;以及

响应于所述第一RTP包的传输模式是单个会话传输模式并且第一参数等于所述第一值,基于所述第一NAL单元的传输顺序确定所述第一NAL单元的解码顺序编号。

9. 根据权利要求8所述的装置,其中所述第一值等于零。

10. 根据权利要求8所述的装置,其中所述最大数目等于32767。

11. 根据权利要求8所述的装置,其中所述解码顺序编号等于传输顺序编号。

12. 根据权利要求8所述的装置,其中所述一或多个处理器进一步经配置以:

接收第二RTP包,其包括第二NAL单元;

响应于所述第二RTP包的传输模式是多会话传输模式,并且响应于所述第二NAL单元按照传输顺序是第一个NAL单元,确定所述第二NAL单元按照解码顺序是第一个。

13. 根据权利要求8所述的装置,其中所述一或多个处理器进一步经配置以:

接收第二RTP包,其包括第二NAL单元;

响应于所述第二NAL单元的解码顺序编号等于所述第一NAL单元的所述解码顺序编号,在对所述第一NAL单元进行解码之前对所述第二NAL单元进行解码。

14. 根据权利要求8所述的装置,其中所述RTP包包括单个NAL单元包。

15. 根据权利要求8所述的装置,其中所述装置包括以下各项中的至少一者:

集成电路;

微处理器;或

无线通信装置,其包括视频译码器。

16. 一种存储指令的非易失性计算机可读存储介质,所述指令在由一或多个处理器执行时使所述一或多个处理器:

接收第一实时传送协议RTP包,其包括第一网络抽象层NAL单元;

接收传输模式参数,其指示所述RTP包的传输模式;

确定与所述第一RTP包相关联的第一参数是否等于第一值,其中所述第一参数指明按照传输顺序在解包缓冲器中在所述第一NAL单元之前并且按照解码顺序在所述第一NAL单元之后的NAL单元的最大数目;以及

响应于所述第一RTP包的传输模式是单个会话传输模式并且第一参数等于所述第一值,基于所述第一NAL单元的传输顺序确定所述第一NAL单元的解码顺序编号。

17. 根据权利要求16所述的非易失性计算机可读存储介质,其中所述第一值等于零。

18. 根据权利要求16所述的非易失性计算机可读存储介质,其中所述最大数目等于32767。

19. 根据权利要求16所述的非易失性计算机可读存储介质,其中所述解码顺序编号等于传输顺序编号。

20. 根据权利要求16所述的非易失性计算机可读存储介质,其存储进一步的指令,所述指令在由所述一或多个处理器执行时使所述一或多个处理器:

接收第二RTP包,其包括第二NAL单元;

响应于所述第二RTP包的传输模式是多会话传输模式,并且响应于所述第二NAL单元按照传输顺序是第一个NAL单元,确定所述第二NAL单元按照解码顺序是第一个。

21. 根据权利要求16所述的非易失性计算机可读存储介质,其存储进一步的指令,所述指令在由所述一或多个处理器执行时使所述一或多个处理器:

接收第二RTP包,其包括第二NAL单元;

响应于所述第二NAL单元的解码顺序编号等于所述第一NAL单元的所述解码顺序编号,在对所述第一NAL单元进行解码之前对所述第二NAL单元进行解码。

22. 根据权利要求16所述的非易失性计算机可读存储介质,其中所述RTP包包括单个NAL单元包。

23. 一种处理视频数据的方法,所述方法包括:

产生第一实时传送协议RTP包,其包括第一网络抽象层NAL单元;

产生传输模式参数,其指示所述RTP包的传输模式;

确定与所述第一RTP包相关联的第一参数,其中所述第一参数指明按照接收顺序在解包缓冲器中在所述第一NAL单元之前并且按照解码顺序在所述第一NAL单元之后的NAL单元

的最大数目;以及

响应于所述第一RTP包的传输模式是单个会话传输模式并且所述第一参数等于第一值,基于所述第一NAL单元的解码顺序设置所述第一NAL单元的传输顺序。

24. 根据权利要求23所述的方法,其中所述第一值等于零。

25. 根据权利要求23所述的方法,其中所述最大数目等于32767。

26. 根据权利要求23所述的方法,其中所述RTP包包括单个NAL单元包。

改进的有效负载格式设计

[0001] 本申请案主张2013年3月29日提交的美国临时申请61/806,705的益处,所述申请的整个内容以引用的方式并入本文中。

技术领域

[0002] 本发明涉及视频数据的处理。

背景技术

[0003] 可将数字视频能力并入到多种多样的装置中,所述装置包含数字电视、数字直播系统、无线广播系统、个人数字助理(PDA)、膝上型或桌上型计算机、平板计算机、电子图书阅读器、数码相机、数字记录装置、数字媒体播放器、视频游戏装置、视频游戏控制台、蜂窝式或卫星无线电电话(所谓的“智能电话”)、视频电话会议装置、视频流式传输装置及其类似者。数字视频装置实施视频压缩技术,例如,由MPEG-2、MPEG-4、ITU-T H.263、ITU-T H.264/MPEG-4第10部分、高级视频译码(AVC)定义的标准、目前正在开发的高效率视频译码(HEVC)标准及此类标准的扩展中所描述的视频压缩技术。视频装置可以通过实施此类视频压缩技术来更有效率地传输、接收、编码、解码及/或存储数字视频信息。

[0004] 视频压缩技术执行空间(图片内)预测及/或时间(图片间)预测来减少或移除视频序列中固有的冗余。对于基于块的视频译码来说,视频切片(即,视频帧或视频帧的一部分)可分割成视频块,视频块也可称作树块、译码单元(CU)及/或译码节点。使用相对于同一图片中的相邻块中的参考样本的空间预测对图片的经帧内译码(I)切片中的视频块进行编码。图片的经帧间译码(P或B)切片中的视频块可使用相对于同一图片中的相邻块中的参考样本的空间预测或相对于其它参考图片中的参考样本的时间预测。图片可被称为帧,且参考图片可被称为参考帧。

[0005] 空间或时间预测产生用于待译码的块的预测性块。残余数据表示待译码原始块与预测性块之间的像素差。经帧间译码块是根据指向形成预测性块的参考样本块的运动向量及指示经译码块与预测性块之间的差的残余数据编码的。经帧内译码块是根据帧内译码模式及残余数据编码的。为了进一步压缩,可将残余数据从像素域变换到变换域,从而产生残余变换系数,接着可以对残余变换系数进行量化。可扫描起初按二维阵列布置的经量化变换系数,以便产生变换系数的一维向量,且可应用熵译码以实现更多压缩。

发明内容

[0006] 一般来说,本发明描述用于处理视频数据的技术。具体来说,本发明描述改进的实时传输协议(RTP)有效负载格式设计。

[0007] 在一个实例中,一种处理视频数据的方法包含:接收第一实时传输协议(RTP)包,其包括第一网络抽象层(NAL)单元;以及响应于所述第一RTP包的传输模式是单个会话传输模式并且第一参数等于第一值,基于所述第一NAL单元的传输顺序确定所述第一NAL单元的解码顺序编号。

[0008] 在另一个实例中,一种用于处理视频数据的装置包含:存储器;经配置以接收实时传输协议(RTP)包的接收器;以及一或多个处理器,其经配置以:接收第一实时传输协议(RTP)包,其包括第一网络抽象层(NAL)单元;以及响应于所述第一RTP包的传输模式是单个会话传输模式并且第一参数等于第一值,基于所述第一NAL单元的传输顺序确定所述第一NAL单元的解码顺序编号。

[0009] 在另一实例中,一种计算机可读存储媒体存储指令,所述指令在由一或多个处理器执行时使所述一或多个处理器:接收第一实时传输协议(RTP)包,其包括第一网络抽象层(NAL)单元;以及响应于所述第一RTP包的传输模式是单个会话传输模式并且第一参数等于第一值,基于所述第一NAL单元的传输顺序确定所述第一NAL单元的解码顺序编号。

[0010] 在另一个实例中,一种用于处理视频数据的设备包含:用于接收第一实时传输协议(RTP)包的装置,所述RTP包包括第一网络抽象层(NAL)单元;用于响应于所述第一RTP包的传输模式是单个会话传输模式并且第一参数等于第一值,基于所述第一NAL单元的传输顺序确定所述第一NAL单元的解码顺序编号的装置。

[0011] 在另一个实例中,一种处理视频数据的方法包含:产生第一实时传输协议(RTP)包,其包括第一网络抽象层(NAL)单元;以及响应于所述第一RTP包的传输模式是单个会话传输模式并且第一参数等于第一值,基于所述第一NAL单元的解码顺序设置所述第一NAL单元的传输顺序。

[0012] 随附图式及以下描述中陈述一或多个实例的细节。其它特征、目标及优点将从所述描述及图式以及权利要求书而显而易见。

附图说明

[0013] 图1是说明可利用本发明中描述的技术的实例视频编码及解码系统的框图。

[0014] 图2展示聚集包结构的视觉表示。

[0015] 图3是说明可实施本发明中描述的技术的实例视频编码器的框图。

[0016] 图4是说明可实施本发明中描述的技术的实例视频解码器的框图。

[0017] 图5是说明形成网络的一部分的一组实例装置的框图。

[0018] 图6展示根据本发明的技术对NAL单元进行解包的实例方法。

[0019] 图7展示根据本发明的技术对NAL单元进行包化的实例方法。

[0020] 图8展示根据本发明的技术对NAL单元进行解包的实例方法。

[0021] 图9展示根据本发明的技术对NAL单元进行包化的实例方法。

[0022] 图10展示根据本发明的技术对NAL单元进行解包的实例方法。

[0023] 图11展示根据本发明的技术对NAL单元进行包化的实例方法。

具体实施方式

[0024] 本发明介绍用于经由RTP传输经编码视频数据的实时传输协议(RTP)有效负载格式的改进的设计的多种技术。RTP是一种传输协议,如IETF RFC 3550中所规定,IETF RFC3550截至2013年3月29日为止可以从<http://www.ietf.org/rfc/rfc3550.txt>获得,并且其全文以引用的方式并入本文中。根据IETF RFC 3550,以向具有实时特性的数据(例如交互式音频和视频)提供端到端传递服务的打算研发了RTP。根据RTP传输的数据被包化成

RTP包。RTP包是包含RTP标头和有效负载数据的数据包。RTP包的有效负载数据可以是经编码视频数据。经编码视频数据可以例如采用一或多个网络抽象层(NAL)单元的形式。

[0025] 为了经由RTP传输根据视频编解码器编码的视频数据,可能需要指定用于所述视频编解码器的RTP有效负载格式。举例来说,RFC 6184(其截至2013年3月29日为止在<http://www.ietf.org/rfc/rfc6184.txt>可以获得)指明用于H.264视频的RTP有效负载格式,并且RFC 6190(其截至2013年3月29日为止在<http://www.ietf.org/rfc/rfc6190.txt>可以获得)指明用于SVC视频的RTP有效负载格式,这两份文件的全文以引用的方式并入本文中。用于HEVC视频的RTP有效负载格式的最近草案截至2013年3月29日为止从<http://tools.ietf.org/html/draft-schierl-payload-rtp-h265-01>可以获得,并且其全文以引用的方式并入本文中。这些各种标准描述了如何将经译码视频数据(例如经译码NAL单元)包化成RTP包。

[0026] 根据HEVC规范,将NAL单元定义为这样一种语法结构,其含有后面要跟着的数据的类型的指示以及视需要穿插有防模拟字节的含有原始字节序列有效负载(RBSP)形式的所述数据的字节。VCL NAL单元包含视频译码层数据,而非VCL NAL单元可以包含关于视频译码层数据的一些其它数据。根据HEVC,存取单元被定义为一组NAL单元,其根据指定分类规则彼此相关联,按照解码顺序是连续的,并且含有恰好一个经译码图片。除了含有经译码图片的VCL NAL单元之外,存取单元还可含有非VCL NAL单元。存取单元的解码始终得到经解码图片。RTP包是用于传输NAL单元的包。

[0027] RFC 6184和RFC 6190中的RTP有效负载格式和用于HEVC的现有草案的RTP有效负载格式的设计与几个可能的问题或缺点相关联。作为一个实例,指定多个包化模式,并且指定许多类型的包,从而导致可能难以选择要使用的包化模式和包类型。作为另一实例,只有通过使用RFC 6184和RFC 6190中定义的多时间聚集包(MTAP)才能交错一个存取单元的网络抽象层(NAL)单元。然而,当一个RTP包中仅仅聚集一个存取单元的NAL单元时,这些RTP包全部具有相同时戳。因此,简单地依赖于RTP包的RTP时戳虽然是足够的,但是按照RFC 6184和RFC 6190的要求发送额外时间信息可能会浪费带宽。交错一个存取单元的NAL单元允许在不同包中传输一个图片的交错的经译码切片,因而当一个包丢失时,可以利用接收到的邻近切片以便更好地隐藏。

[0028] 为了解决上文介绍的可能的问题和缺点,本发明介绍了用于改进的RTP有效负载格式设计的几种技术。根据一种技术,不区分包化模式,从而使得非交错和交错包化都是可能的,单个会话传输和多个会话传输都是可能的,并且通用解包过程是基于NAL单元的绝对解码顺序编号值指定的,所述解码顺序编号值可以从在包有效负载中用信号表示的任选的信息导出的。

[0029] 根据另一技术,聚集包的设计允许用于交错一个存取单元的NAL单元,而不需要发送冗余时间信息。如本发明所描述,聚集包可以在输送多个较小切片时改进视频译码。根据本发明的技术允许交错NAL单元可以改进总经重构图像质量。举例来说,如果聚集包包含交错NAL单元并且聚集包丢失,则交错NAL单元很可能对应于分散的视频块群组而不是邻近视频块。错误隐藏技术通常对于较小丢失区域更有效,并且因此,与隐藏邻近视频块群组的丢失相比,可能更有效隐藏分散视频块群组的丢失。

[0030] 图1是说明可以结合本发明中描述的技术使用的实例视频处理系统10的框图。系

统10可以使用本发明中描述的RTP技术例如产生、处理和传输视频数据。如图1中所示,系统10包含源装置12,其提供稍后将由目的地装置14解码的经编码视频数据。经编码视频数据可以通过媒体感知网络元件(MANE) 29从源装置12路由到目的地装置14。源装置12和目的地装置14可包括多种多样的装置中的任一者,包含桌上型计算机、笔记型(即,膝上型)计算机、平板计算机、机顶盒、电话手持机(例如所谓的“智能”电话)、所谓的“智能”平板电脑、电视机、相机、显示装置、数字媒体播放器、视频游戏控制台、视频流式传输装置或类似者。在一些情况下,可装备源装置12和目的地装置14以用于无线通信。

[0031] 系统10可根据不同视频译码标准、专属标准或技术或多视图译码的任何其它方式操作。举例来说,视频编码器20及视频解码器30可根据视频压缩标准操作,所述视频压缩标准例如包含ITU-T H.261、ISO/IEC MPEG-1Visual、ITU-T H.262或ISO/IEC MPEG-2Visual、ITU-T H.263、ISO/IEC MPEG-4Visual及ITU-T H.264(也被称作ISO/IEC MPEG-4AVC),包含其可缩放视频译码(SVC)及多视图视频译码(MVC)扩展。MVC扩展的最近公开可用的联合草案描述于2010年3月的“用于通用视听服务的高级视频译码”,ITU-T建议H.264中。MVC扩展的更为新近的公开可用的联合草案描述于2011年6月的“用于通用视听服务的高级视频译码”,ITU-T建议H.264中。MVC扩展的当前联合草案已从2012年1月起获得审批。

[0032] 此外,存在一种新的视频译码标准,即高效率视频译码(HEVC)标准,其目前正由ITU-T视频译码专家组(VCEG)及ISO/IEC动画专家组(MPEG)的视频译码联合合作小组(JCT-VC)进行开发。截至2013年3月15日,被称作HEVC WD9的HEVC的新近工作草案(WD)可从http://phenix.int-evry.fr/jct/doc_end_user/documents/11_Shanghai/wg11/JCTVC-K1003-v10.zip获得。出于描述的目的,在HEVC或H.264标准及此些标准的扩展的上下文中描述视频编码器20及视频解码器30。然而,本发明的技术不限于任何特定译码标准。视频压缩标准的其它实例包含MPEG-2及ITU-T H.263。例如称为On2VP6/VP7/VP8的那些译码技术的专属译码技术也可实施本文所述的技术中的一或多个者。本发明的技术潜在地适用于几种视频译码标准,包含HEVC等。

[0033] 目的地装置14可经由链路16接收待解码的经编码视频数据。链路16可包括能够将经编码视频数据从源装置12移动到目的地装置14的任何类型的媒体或装置。在一个实例中,链路16可包括使得源装置12能够实时地将经编码视频数据直接传输到目的地装置14的通信媒体。经编码视频数据可根据通信标准(例如,无线通信协议)来调制,且被传输到目的地装置14。通信媒体可包括任何无线或有线通信媒体,例如射频(RF)频谱或一或多个物理传输线路。通信媒体可能形成基于包的的网络(例如,局域网、广域网或全球网络,例如因特网)的一部分。通信媒体可包含路由器、交换器、基站或任何其它可以用于促进从源装置12到目的地装置14的通信的设备。链路16可以包含一或多个MANE,例如MANE 27,其将视频数据从源装置12路由到目的地装置14。

[0034] 或者,可将经编码数据从输出接口22输出到存储装置25。类似地,可由输入接口从存储装置25存取经编码数据。存储装置25可包含多种分布式或本地存取的数据存储媒体中的任一者,例如硬盘驱动器、蓝光光盘、DVD、CD-ROM、快闪存储器、易失性或非易失性存储器或任何其它用于存储经编码视频数据的合适的数字存储媒体。在另一实例中,存储装置25可对应于文件服务器或可保持由源装置12产生的经编码视频的另一中间存储装置。目的地装置14可经由流式传输或下载从存储装置25存取所存储的视频数据。文件服务器可以是任

何类型的能够存储经编码的视频数据并且将经编码的视频数据传输到目的地装置14的服务器。实例文件服务器包含网络服务器(例如,用于网站)、FTP服务器、网络附接存储(NAS)装置或本地磁盘驱动器。目的地装置14可以通过任何标准数据连接(包含因特网连接)来存取经编码视频数据。此连接可包含适合于存取存储于文件服务器上的经编码视频数据的无线信道(例如,Wi-Fi连接)、有线连接(例如,DSL、电缆调制解调器等)或两者的组合。经编码视频数据从存储装置25的传输可为流式传输、下载传输或两者的组合。从存储装置25检索的视频数据可以使用一或多个MANE(例如MANE 27)路由到目的地装置14。

[0035] 本发明的技术未必限于无线应用或设置。所述技术可应用于视频译码以支持多种多媒体应用,例如,空中电视广播、有线电视传输、卫星电视传输、流式视频传输(例如,经由因特网)、编码数字视频以用于存储于数据存储媒体上、解码存储于数据存储媒体上的数字视频或其它应用。在一些实例中,系统10可经配置以支持单向或双向视频传输,以支持例如视频流式传输、视频重放、视频广播及/或视频电话等应用。

[0036] 在图1的实例中,源装置12包含视频源18、视频编码器20、包化器21和输出接口22。在一些情况下,输出接口22可包含调制器/解调器(调制解调器)及/或发射器。在源装置12中,视频源18可包含例如视频俘获装置(例如,摄像机)、含有先前俘获的视频的视频存档、用于从视频内容提供者接收视频的视频馈入接口及/或用于产生计算机图形数据作为源视频的计算机图形系统或此类源的组合等源。作为一个实例,如果视频源18是摄像机,那么源装置12和目的地装置14可以形成所谓的相机电话或视频电话。然而,本发明中所描述的技术一般可适用于视频译码,且可应用于无线及/或有线应用。

[0037] 可由视频编码器20编码所俘获视频、预俘获视频或计算机产生的视频。经编码视频数据可经由源装置12的输出接口22直接传输到目的地装置14。经编码视频数据还可(或替代地)存储到存储装置25上以供稍后由目的地装置14或其它装置存取以用于解码及/或回放。

[0038] 目的地装置14包含输入接口28、解包器29、视频解码器30和显示装置32。在一些情况下,输入接口28可包含接收器及/或调制解调器。目的地装置14的输入接口28经由链路16接收经编码视频数据。经由链路16传送或在存储装置25上提供的经编码视频数据可包含由视频编码器20产生的多种语法元素以供由例如视频解码器30等视频解码器用于对视频数据进行解码。这些语法元素可与在通信媒体上传输、存储在存储媒体上或存储在文件服务器中的经编码视频数据包含在一起。

[0039] 显示装置32可与目的地装置14集成或在目的地装置14外部。在一些实例中,目的地装置14可包含集成显示装置,且还经配置以与外部显示装置介接。在其它实例中,目的地装置14可为显示装置。一般来说,显示装置32将经解码视频数据显示给用户,且可包括多种显示装置中的任一者,例如液晶显示器(LCD)、等离子显示器、有机发光二极管(OLED)显示器或另一类型的显示装置。

[0040] 尽管图1中未展示,但在一些方面中,视频编码器20及视频解码器30可各自与音频编码器及解码器集成,且可包含适当多路复用器-多路分用器单元或其它硬件及软件,以处置对共用数据流或单独数据流中的音频及视频两者的编码。在一些实例中,如果适用的话,多路复用器-多路分用器单元可以符合ITU H.223多路复用器协议,或例如用户数据报协议(UDP)等其它协议。

[0041] 视频编码器20和视频解码器30各自可实施为例如一或多个微处理器、数字信号处理器(DSP)、专用集成电路(ASIC)、现场可编程门阵列(FPGA)、离散逻辑、软件、硬件、固件或其任何组合的多种合适编码器电路中的任一者。当部分地用软件实施所述技术时,装置可将用于所述软件的指令存储于合适的非暂时性计算机可读媒体中且使用一或多个处理器用硬件执行所述指令以执行本发明的技术。视频编码器20和视频解码器30中的每一者可包含在一或多个编码器或解码器中,所述编码器或解码器中的任一者可集成为相应装置中的组合编码器/解码器(编解码器)的一部分。

[0042] JCT-VC正在致力于开发HEVC标准。HEVC标准化努力是基于被称作HEVC测试模型(HM)的视频译码装置的演进模型。HM假设视频译码装置根据例如ITU-T H.264/AVC相对于现有装置的几项额外能力。举例来说,虽然H.264提供九种帧内预测编码模式,但HM可提供多达三十三种帧内预测编码模式。

[0043] 总的来说,HM的工作模型描述视频帧或图片可以划分成一连串树块或最大译码单元(LCU),其包含亮度和色度样本的两种块。树块具有与H.264标准的宏块类似的目的。切片包含按译码顺序的多个连续树块。视频帧或图片可被分割成一或多个切片。每一树块可根据二叉树分裂成译码单元(CU)。举例来说,作为二叉树的根节点的树块可分裂成四个子节点,且每一子节点又可为父节点并分裂成另外四个子节点。最后的未经分裂的子节点(作为二叉树的叶节点)包括译码节点,即,经译码视频块。与经译码位流相关联的语法数据可限定树块可分裂的最大次数,且还可限定译码节点的最小大小。树块的分裂可以发生在亮度域中,并且可以在色度域中得到仿效,可能使用对叶节点的进一步子取样。

[0044] CU包含译码节点以及与所述译码节点相关联的预测单元(PU)及变换单元(TU)。CU的大小对应于译码节点的大小并且形状必须是正方形。CU的大小范围可从 8×8 像素到具有最大 64×64 像素或更大的树块的大小。每一CU可含有一或多个PU及一或多个TU。举例来说,与CU相关联的语法数据可描述将CU分割成一或多个PU。分割模式可以在CU被跳过或经直接模式编码、经帧内预测模式编码或经帧间预测模式编码之间有区别。PU可以分割成非正方形形状。举例来说,与CU相关联的语法数据还可描述根据二叉树将CU分割成一或多个TU。TU可为正方形或非正方形形状。

[0045] HEVC标准允许根据TU变换,TU可针对不同CU而有所不同。TU的大小通常是基于针对经分割LCU定义的给定CU内的PU的大小而确定,但是情况可能并不总是如此。TU通常与PU大小相同或小于PU。在一些实例中,可以使用被称为“残余二叉树”(RQT)的二叉树结构将对应于CU的残余样本细分成较小单元。RQT的叶节点可被称为变换单元(TU)。可以变换与TU相关联的像素差值以产生变换系数,所述变换系数可以经量化。

[0046] 一般来说,PU包含与预测过程有关的数据。举例来说,当PU经帧内模式编码时,PU可包含描述PU的帧内预测模式的数据。作为另一实例,当PU经帧间模式编码时,PU可包括限定PU的运动向量的数据。限定PU的运动向量的数据可描述(例如)运动向量的水平分量、运动向量的垂直分量、运动向量的分辨率(例如,四分之一像素精度或八分之一像素精度)、运动向量所指向的参考图片,及/或运动向量的参考图片列表(例如,列表0、列表1或列表C)。

[0047] 一般来说,TU用于变换和量化过程。具有一或多个PU的给定CU还可包含一或多个变换单元(TU)。在预测之后,视频编码器20可计算对应于PU的残余值。残余值包括像素差值,所述像素差值可使用TU变换成变换系数、经量化且扫描以产生串行化变换系数以用于

熵译码。本发明通常使用术语“视频块”来指CU的译码节点。在一些特定情况下,本发明还可使用术语“视频块”来指包含译码节点以及PU及TU的树块,即,LCU或CU。

[0048] 视频序列通常包含一系列视频帧或图片。图片群组(GOP)一般包括一系列的一或多个所述视频图片。GOP可包含GOP的标头、图片中的一者或一者以上的标头或其它地方中的语法数据,其描述GOP中包含的图片的数目。图片的每一切片可包含切片语法数据,其描述用于相应切片的编码模式。视频编码器20通常对个别视频切片内的视频块进行操作以便对视频数据进行编码。视频块可对应于CU内的译码节点。视频块可以具有固定或变化的大小,并且可以根据指定译码标准而有不同大小。

[0049] 作为实例,HM支持各种PU大小的预测。假定特定CU的大小为 $2N \times 2N$,那么HM支持 $2N \times 2N$ 或 $N \times N$ 的PU大小的帧内预测,和 $2N \times 2N$ 、 $2N \times N$ 、 $N \times 2N$ 或 $N \times N$ 的对称PU大小的帧间预测。HM还支持用于 $2N \times nU$ 、 $2N \times nD$ 、 $nL \times 2N$ 和 $nR \times 2N$ 的PU大小的帧间预测的不对称分割。在不对称分割中,不分割CU的一个方向,但是另一方向分割成25%及75%。CU的对应于25%分割区的部分指示成“n”,接着是用“上”、“下”、“左”或“右”指示。因而,举例来说,“ $2N \times nU$ ”是指水平地分割的 $2N \times 2N$ CU,其中上方有 $2N \times 0.5N$ PU,而下方有 $2N \times 1.5N$ PU。

[0050] 在本发明中,“ $N \times N$ ”与“N乘N”可互换地使用以依据垂直及水平尺寸来指代视频块的像素尺寸,例如, 16×16 像素或16乘16像素。总的来说, 16×16 块将在垂直方向上具有16个像素($y=16$),且在水平方向上具有16个像素($x=16$)。同样, $N \times N$ 块总体上在垂直方向上具有N个像素,并且在水平方向上具有N个像素,其中N表示非负整数值。一块中的像素可布置成若干行和若干列。此外,块未必需要在水平方向与垂直方向上具有相同数目的像素。举例来说,块可包括 $N \times M$ 像素,其中M未必等于N。

[0051] 在使用CU的PU进行帧内预测性或帧间预测性译码之后,视频编码器20可以计算用于CU的TU的残余数据。PU可包括空间域(还被称作像素域)中的像素数据,且在将变换应用到残余视频数据之后,TU可包括变换域中的系数,所述变换例如是离散余弦变换(DCT)、整数变换、小波变换或概念上类似的变换。所述残余数据可对应于未编码图片的像素与对应于PU的预测值之间的像素差。视频编码器20可以形成包含用于CU的残余数据的TU,并且接着变换TU以产生用于CU的变换系数。

[0052] 在用于产生变换系数的任何变换之后,视频编码器20可执行变换系数的量化。量化通常是指将变换系数量化以可能减少用以表示系数的数据量从而提供进一步压缩的过程。量化过程可减少与系数中的一些系数或全部相关联的位深度。举例来说,n位值可在量化期间被舍去成m位值,其中n大于m。

[0053] 在一些实例中,视频编码器20可利用预定义扫描顺序来扫描经量化变换系数以产生可被熵编码的串行化向量。在其它实例中,视频编码器20可执行自适应性扫描。在扫描经量化变换系数以形成一维向量之后,视频编码器20可例如根据上下文自适应可变长度译码(CAVLC)、上下文自适应二进制算术译码(CABAC)、基于语法的上下文自适应二进制算术译码(SBAC)、概率区间分割熵(PIPE)译码或另一熵编码方法对所述一维向量进行熵编码。视频编码器还20可对与经编码视频数据相关联的语法元素进行熵编码以供视频解码器30在对视频数据进行解码时使用。

[0054] 为了执行CABAC,视频编码器20可以向待传输的符号指派上下文模型内的一个上下文。上下文可涉及(例如)符号的相邻值是否为非零。为了执行CAVLC,视频编码器20可以

选择用于待传输的符号的可变长度码。VLC中的码字可经构造使得相对较短的代码对应于更有可能的符号,而较长的代码对应于不太可能的符号。以此方式,相对于(例如)针对待传输的每一符号使用相等长度的码字,使用VLC可实现位节省。概率确定可基于指派给符号的上下文。

[0055] 本发明中描述的技术可以独立应用或联合应用。这些技术的方面可以通过包化器21和解包器29执行。在一些情况下,包化器21可被称为RTP发送方,或简单地称为发送方,而解包器29可被称为RTP接收方或简单地称为接收方。这些技术的方面概括如下:

[0056] -用信号表示在片段单元(FU)中传送的经译码拼片的第一译码树单元(CTU)的地址

[0057] ○在FU有效负载之前,在FU结构中用信号表示拼片ID(或两个拼片ID值的差量)。这个信令指明或指示拼片中的第一CTU的拼片扫描中的CTU地址(以及光栅扫描中的地址)。

[0058] ○替代地,在FU有效负载之前,在FU结构中用信号表示在FU中传送的经译码拼片中的第一CTU的拼片扫描中的CTU地址(或两个此类值的差量)。

[0059] ○替代地,在FU有效负载之前,在FU结构中用信号表示在FU中传送的经译码拼片中的第一CTU的光栅扫描中的CTU地址(或两个此类值的差量)。

[0060] ○替代地,仅仅当存在指示存在信令的指示(例如媒体类型参数)时才存在所述信令(用以上形式中的任一者)。此媒体类型参数可以简单地指示存在以上信令,或指示使用拼片(并且如果指示未使用拼片,则以上信令不存在)。

[0061] -当在多个FU中传送拼片时:

[0062] ○在FU标头中使用/添加S旗标以指示分断式拼片的开头。

[0063] ■对于这个,存在任何上述信令以导出拼片中的第一CTU的CTU地址的(进一步)条件是S旗标等于0。

[0064] ○在FU标头中使用/添加E旗标以指示分断式拼片的结尾。

[0065] -在RTP包有效负载标头中使用/添加旗标以指示是否包中的所有NAL单元都含有相依性切片片段。

[0066] ○替代地,RTP包标头中的两个位用以指示以下之一

[0067] ■包中的所有NAL单元都是相依性切片片段。

[0068] ■包中的NAL单元中的至少一者是相依性切片片段,对于所述相依性切片片段,对应的独立切片片段不在相同包中。

[0069] ■包中的NAL单元中的至少一者是独立切片片段。

[0070] ■包中的所有NAL单元都是独立切片片段。

[0071] ○在仅仅含有一个NAL单元的包中,仅仅需要一个位来指示NAL单元是否含有相依性切片片段。

[0072] ○替代地,仅仅当存在指示存在信令的指示(例如媒体类型参数)时才存在所述信令(用以上形式中的任一者)。此媒体类型参数可以简单地指示以上信令的存在,或指示相依性切片片段的使用(并且如果指示未使用相依性切片片段,则以上信令不存在)。

[0073] 现在将描述有效负载结构的方面。这些有效负载结构可以由包化器21产生并且由解包器29剖析。RTP包的有效负载的前两个字节可以限定有效负载标头。有效负载标头可以由与HEVC NAL单元标头相同的字段构成(F,Type,LayerId和TID,其对应于语法元素

forbidden_zero_bit, nal_unit_type, nuh_layer_id, 和 nuh_temporal_id_plus1, 如 HEVC WD 10 的 7.3.1.2 部分中规定的), 这与有效负载结构的类型无关。

[0074] 指定了三种不同类型的 RTP 包有效负载结构。接收方可以通过有效负载标头中的 Type 字段来识别 RTP 包有效负载的类型。接收方可以是包含视频解码器的装置的解包器, 或者可以形成 MANE 或其它网络实体的一部分。这三种不同有效负载结构如下:

[0075] 单个 NAL 单元包: 有效负载中含有单个 NAL 单元, 并且 NAL 单元的 NAL 单元标头也充当有效负载标头。当 tx 模式等于 “MST” 或 sprop-depack-buf-nalus 大于 0 时, 不得使用单个 NAL 单元包。

[0076] 聚集包 (AP): 一个存取单元内含有一或多个 NAL 单元。见下文。

[0077] 分断单元 (FU): 含有单个 NAL 单元的子组。见下文。

[0078] 现在将描述包化器 21 和解包器 29 支持的传输模式。本发明的技术可以使得能够经由单个 RTP 会话或多个 RTP 会话传输 HEVC 位流。所述概念和工作原理符合 RFC6190, 并且遵循类似的但可能更简单的设计。如果仅仅有一个 RTP 会话用于传输 HEVC 位流, 则传输模式被称作单个会话传输 (SST); 否则 (有一个以上 RTP 会话用于传输 HEVC 位流), 传输模式被称作多会话传输 (MST)。

[0079] SST 应当用于一点到一点单播场景, 而 MST 应当用于一点到多点多播场景, 其中不同接收方需要相同 HEVC 位流的不同操作点来改进带宽利用效率。

[0080] 如果 tx 模式等于 “SST”, 则必须使用 SST。否则 (tx 模式等于 “MST”), 则必须使用 MST。

[0081] 现在将描述解码顺序编号的方面。对于每一 NAL 单元, 导出可变 AbsDon, 其表示指示 NAL 单元解码顺序的解码顺序编号。

[0082] 假设 NAL 单元 n 是 RTP 会话内按传输顺序的第 n 个 NAL 单元。

[0083] 如果 tx 模式等于 “SST” 并且 sprop-depack-buf-nalus 等于 0, 则导出 AbsDon[n] (NAL 单元 n 的 AbsDon 的值) 等于 n。

[0084] 否则 (tx 模式等于 “MST” 或 sprop-depack-buf-nalus 大于 0), 则如下导出 AbsDon[n], 其中 DON[n] 是 NAL 单元 n 的变量 DON 的值:

[0085] - 如果 n 等于 0 (即 NAL 单元 n 就是按传输顺序的第一个 NAL 单元), 则 AbsDon[0] 设置成等于 DON[0]。

[0086] - 否则 (n 大于 0), 则以下适用于 AbsDon[n] 的导出:

[0087] If $DON[n] == DON[n-1]$,

[0088] $AbsDon[n] = AbsDon[n-1]$

[0089] If $(DON[n] > DON[n-1] \text{ and } DON[n] - DON[n-1] < 32768)$,

[0090] $AbsDon[n] = AbsDon[n-1] + DON[n] - DON[n-1]$

[0091] If $(DON[n] < DON[n-1] \text{ and } DON[n-1] - DON[n] \geq 32768)$,

[0092] $AbsDon[n] = AbsDon[n-1] + 65536 - DON[n-1] + DON[n]$

[0093] If $(DON[n] > DON[n-1] \text{ and } DON[n] - DON[n-1] \geq 32768)$,

[0094] $AbsDon[n] = AbsDon[n-1] - (DON[n-1] + 65536 - DON[n])$

[0095] If $(DON[n] < DON[n-1] \text{ and } DON[n-1] - DON[n] < 32768)$,

[0096] $AbsDon[n] = AbsDon[n-1] - (DON[n-1] - DON[n])$

[0097] 对于任何两个NAL单元m和n,以下适用:

[0098] -AbsDon[n]大于AbsDon[m]指示按NAL单元解码顺序NAL单元n在NAL单元m之后。

[0099] -当AbsDon[n]等于AbsDon[m]时,两个NAL单元的NAL单元解码顺序可以采用任何一种顺序。

[0100] -AbsDon[n]小于AbsDon[m]指示按照解码顺序NAL单元n在NAL单元m前面。

[0101] 当按NAL单元解码顺序的两个连续NAL单元具有不同AbsDon值时,按照解码顺序的第二NAL单元的AbsDon值必须大于第一NAL单元的AbsDon值,并且两个AbsDon值之间的绝对差值可以大于或等于1。

[0102] 现在将描述聚集包(AP)。图2展示聚集包结构的视觉表示。聚集包120包含有效负载标头122(标示为PayloadHdr),后面是有效负载数据124。有效负载数据包含一或多个聚集单元,图2中展示为聚集包0到N。每一聚集单元可以包含一个NAL单元。举例来说,聚集单元0包含NAL单元0,聚集单元1包含NAL单元1,并且聚集单元N包含NAL单元N。图2还展示有效负载标头的前16个位,其包含F位、TYPE字段、R字段(有时也被称作LayerId字段)和TID字段。

[0103] 引入AP以使得能够降低较小NAL单元的包化开销,例如大多数非VCL NAL单元,其大小通常仅仅为几个八位字节。AP在一个存取单元内聚集多个NAL单元。一个AP中有待承载的每一NAL单元包封在一个聚集单元中。一个AP中聚集的NAL单元是采用NAL单元解码顺序。AP可以由有效负载标头(标示为PayloadHdr)后面是一或多个聚集单元组成。

[0104] 有效负载标头中的字段如下设置。所述每一聚集NAL单元的F位等于零,则F位必须等于0;否则,F位必须等于1。Type字段必须等于48。LayerId值必须等于所有聚集NAL单元的LayerId的最低值。TID值必须是所有聚集NAL单元的TID的最低值。

[0105] AP可以承载需要的数量的聚集单元;然而,AP中的总数据量显然必须能放入IP包中,并且大小应当选择成使得所得IP包小于MTU大小以避免IP层分断。AP不得含有FU。AP不得嵌套;即,一个AP不得含有另一AP。

[0106] AP中的第一聚集单元可以由以下各项组成:任选的16位DONL字段(按照网络字节顺序),后面是16位不带正负号的大小信息(按照网络字节顺序),其指示以字节为单位的NAL单元大小(不包含这两个八位字节,但是包含NAL单元标头),后面是NAL单元本身,包含其NAL单元标头。

[0107] DONL字段当存在时指明聚集NAL单元的解码顺序编号的16个最低有效位的值。

[0108] 如果tx模式等于“MST”或sprop-depack-buf-nalus大于0,则作为AP中的第一聚集单元的一个聚集单元中必须存在DONL字段,并且导出聚集NAL单元的变量DON等于DONL字段的值。否则(tx模式等于“SST”并且sprop-depack-buf-nalus等于0),作为AP中的第一聚集单元的一个聚集单元中不得出现DONL字段。

[0109] 并非AP中的第一聚集单元的一个聚集单元可以由以下各项组成:任选的8位DOND字段,后面是16位不带正负号的大小信息(按照网络字节顺序),其指示以字节为单位的NAL单元大小(不包含这两个八位字节,但是包含NAL单元标头),后面是NAL单元本身,包含其NAL单元标头。

[0110] 当存在时,DOND字段加1可以指明当前聚集NAL单元的解码顺序编号值与相同AP中的前一个聚集NAL单元之间的差值。与要求按照NAL单元在RTP包中出现的顺序对NAL单元进

行解码的有效负载结构不同的是,使用本发明中描述的DOND和DONL参数可以允许指定特定的解码顺序。

[0111] 如果tx模式等于“MST”或sprop-depack-buf-nalus大于0,则并非AP中的第一聚集单元的一个聚集单元中必须存在DOND字段,并且导出聚集NAL单元的变量DON等于相同AP中的前一个聚集NAL单元的DON加DOND字段值加1,以65536取模。否则(tx模式等于“SST”并且sprop-depack-buf-nalus等于0),并非AP中的第一聚集单元的一个聚集单元中不得存在DOND字段。

[0112] 在替代方案中,DOND字段可以具有不同长度,例如4位。在另一个替代方案中,两个非第一聚集单元共享一个8位字段,4位用于每一聚集单元以用信号表示DOND值。在又一替代方案中,DOND字段的长度通过媒体类型参数用信号表示,并且所述参数的值等于0意味着不存在DOND字段。

[0113] 现在将描述分断单元(FU)。引入分断单元(FU)以使得能够将单个NAL单元分断成多个RTP包。NAL单元的一个片断可以由NAL单元的整数数目的连续八位字节组成。相同NAL单元的片断必须用连续顺序以升序的RTP序列号发送(第一片断与最后片断之间不发送相同RTP包流内的其它RTP包)。

[0114] 当NAL单元被分断并且在FU内传达时,其被称作分断式NAL单元。AP不得分断。FU不得嵌套;即,一个FU不得含有另一FU。

[0115] 承载FU的RTP包的RTP时戳设置成分断式NAL单元的NALU时间。

[0116] FU可以由以下各项组成:有效负载标头(标示为PayloadHdr),一个八位字节的FU标头,任选的16位DONL字段(按照网络字节顺序),和FU有效负载。

[0117] 有效负载标头中的字段如下设置。Type字段必须等于49。字段F、LayerId和TID必须分别等于分断式NAL单元的字段F、LayerId和TID。

[0118] FU标头可以由S位、E位和6位Type字段组成。

[0119] 在这个实例中,FU标头字段的语义如下:

[0120] S:1位

[0121] 当设置成一时,S位指示分断式NAL单元的开头,即,FU有效负载的第一字节也是分断式NAL单元的有效负载的第一字节。当FU有效负载不是分断式NAL单元有效负载的开头时,S位必须设置成零。

[0122] E:1位

[0123] 当设置成一时,E位指示分断式NAL单元的结尾,即,有效负载的最后字节也是分断式NAL单元的最后字节。当FU有效负载不是分断式NAL单元的最后片断时,E位必须设置成零。

[0124] Type:6位

[0125] 字段Type必须等于分断式NAL单元的字段Type。

[0126] DONL字段当存在时可以指明分断式NAL单元的解码顺序编号的16个最低有效位的值。

[0127] 如果tx模式等于“MST”或sprop-depack-buf-nalus大于0,并且S位等于1,则FU中必须存在DONL字段,并且导出分断式NAL单元的变量DON等于DONL字段的值。否则(tx模式等于“SST”并且sprop-depack-buf-nalus等于0,或S位等于0),则FU中不得存在DONL字段。

[0128] 不得在一个FU中传输非分断式NAL单元;即,在相同FU标头中开头位和结尾位不得都设置成一。

[0129] FU有效负载可以由分断式NAL单元的有效负载的片断组成,使得如果连续FU的FU有效负载(以S位等于1的FU开始,并且以E位等于1的FU结束)是依次串接的,则分断式NAL单元的有效负载可以重构。不是像FU有效负载中一样包含分断式NAL单元的NAL单元标头,而是在FU的FU有效负载标头的F、LayerId和TID字段和FU的FU标头的Type字段中传达分断式NAL单元的NAL单元标头的信息。FU有效负载可以具有任何数目的八位字节并且可以是空的。

[0130] 如果FU丢失,则接收方应当抛弃对应于相同分断式NAL单元的按照传输顺序的所有接下来的分断单元,除非已知接收方中的解码器准备好得体地处理不完整的NAL单元。

[0131] 端点中或MANE中的接收方可以将NAL单元的第一n-1片断聚合到(不完整的)NAL单元,即使未接收到所述NAL单元的片断n。在这种情况下,必须将NAL单元的forbidden_zero_bit设置成一以指示语法违规。

[0132] 现在将论述包化规则。以下包化规则适用:

[0133] -如果对于一个RTP会话,tx模式等于“MST”或sprop-depack-buf-nalus大于0,则所述RTP会话中承载的NAL单元的传输顺序可以不同于NAL单元解码顺序。否则(对于一个RTP会话,tx模式等于“SST”且sprop-depack-buf-nalus等于0),则RTP会话中承载的NAL单元的传输顺序必须与NAL单元解码顺序相同。

[0134] -当tx模式等于“MST”或sprop-depack-buf-nalus大于0时,无法使用单个NAL单元包。在这种情况下,可以使用AP将单个NAL单元包封在一个RTP包中。

[0135] -应当将小尺寸的NAL单元连同一或多个其它NAL单元一起包封在聚集包中,以便避开较小NAL单元的不必要的包开销。举例来说,非VCL NAL单元(例如存取单元定界符、参数集或SEI NAL单元)通常是较小的。

[0136] -应当将每一非VCL NAL单元连同其相关联的VCL NAL单元一起包封在聚集包中,因为如果相关联的VCL NAL单元不可用,则通常非VCL NAL单元将是无意义的。

[0137] -TID值指示RTP包的相对重要性。TID的较低值指示较高重要性。与不太重要的NAL单元相比,更重要的NAL单元可以受到更好的保护以防传输损失。

[0138] 现在将描述解包过程。解包背后的总体概念是在一个RTP会话和所有相依性RTP会话(如果存在的话)中从RTP包中获得NAL单元,并且将其按照NAL单元解码顺序传递到解码器。

[0139] 解包过程取决于实施方案。因此,以下描述应当视为合适的实施方案的一个实例。也可以使用其它方案,只要针对相同输入的输出与下述过程相同即可。输出的含义与NAL单元组相同,并且其两者的顺序都是相同的。相对于所描述的算法可以实现优化。

[0140] 与缓冲器管理有关的所有一般RTP机制都适用。具体来说,重复的或过时的RTP包(通过RTP序列号和RTP时戳指示)被移除。为了确定解码的确切时间,必须考虑到例如可能故意延迟以容许适当的流间同步的因素。

[0141] 只有NAL单元类型值在0到47范围内(包含0和47)的NAL单元可以被传递到解码器。NAL单元类型值在48到63范围内(包含48和63)的NAL单元类结构不得传递到解码器。

[0142] 接收方包含接收方缓冲器,其用于补偿传输延迟抖动,将NAL单元从传输顺序重新

排序成NAL单元解码顺序,并且在适用时在MST中恢复NAL单元解码顺序。在这个部分中,是在不存在传输延迟抖动的假设下描述接收方的操作。为了与也用于补偿传输延迟抖动的实际接收方缓冲器区分,这个部分中,下文中将接收方缓冲器称为解包缓冲器。接收器还应当为传输延迟抖动做准备;即,为传输延迟抖动缓冲和解包缓冲预留分开的缓冲器,或者对传输延迟抖动和解包两者使用一个接收方缓冲器。此外,接收器应当在缓冲操作中考虑到传输延迟抖动;例如,通过在起始解码和回放之前的额外初始缓冲。

[0143] 在接收方中存在两种缓冲状态:初始缓冲和播放时缓冲。初始缓冲在接收初始化时开始。在初始缓冲之后,解码和回放开始,并且使用播放时缓冲模式。

[0144] 接收方按照接收顺序将传入包存储到接收方缓冲器中,并且将每一会话的RTP包中的NAL单元按照RTP序列号顺序传递到再多路复用缓冲器。计算CS-DON值并且将其针对每一NAL单元存储在再多路复用缓冲器中。

[0145] 不管缓冲状态如何,接收方都将传入NAL单元按照接收顺序存储到解包缓冲器中。单个NAL单元包、AP和FU中承载的NAL单元单独地存储于解包缓冲器中,并且计算AbsDon值且针对每一NAL单元存储AbsDon值。

[0146] 初始缓冲一直持续到条件A(解包缓冲器中的NAL单元的数目大于最高RTP会话的sprop-depack-buf-nalus值)为真为止。在初始缓冲之后,每当条件A为真时,反复地应用以下操作直到条件A变成假为止:从解包缓冲器中移除解包缓冲器中具有最小AbsDon值的NAL单元,并且将其传递到解码器。

[0147] 当不再有NAL单元流入解包缓冲器中时,将所有保留在解包缓冲器中的NAL单元从缓冲器中移除并且按照AbsDon值增序传递到解码器。

[0148] 现在将论述媒体类型注册。从IETF树分派用于HEVC编解码器的媒体亚型。

[0149] 接收方必须忽略任何未指定的参数。

[0150] 媒体类型名称:视频

[0151] 媒体亚型名称:H265

[0152] 必需参数:无

[0153] 任意的参数:

[0154] tx模式:

[0155] 这个参数指示传输模式是SST还是MST。这个参数可以是适用于一个特定会话中的所有包的媒体类型参数。换句话说,所述值可能对于所述会话的所有包是固定的。

[0156] tx模式的值必须等于“MST”或“SST”。当不存在时,推断tx模式值等于“SST”。

[0157] 如果所述值等于“MST”,则必须使用MST。否则(所述值等于“SST”),必须使用SST。

[0158] 对于MST中的所有RTP会话,tx模式值必须等于“MST”。

[0159] sprop-depack-buf-nalus:

[0160] 这个参数可以指明按照接收顺序在解包缓冲器中在NAL单元前面的NAL单元的最大数目,并且按照解码顺序在NAL单元之后。这个参数可以是适用于一个特定会话中的所有包的媒体类型参数。换句话说,所述值可能对于所述会话的所有包是固定的。

[0161] sprop-depack-buf-nalus的值必须是0到32767范围(包含0和32767)内的整数。

[0162] 当不存在时,推断sprop-depack-buf-nalus值等于0。

[0163] 当RTP会话取决于一或多个其它RTP会话(在这种情况下tx模式必须等于“MST”)

时, sprop-depack-buf-nalus值必须大于0。

[0164] sprop-depack-buf-bytes:

[0165] 这个参数用信号表示以字节为单位的解包缓冲器的必需大小。如同部分6中的规定, 参数值必须大于或等于解包缓冲器的最大缓冲器占用率(以字节为单位)。

[0166] sprop-depack-buf-bytes的值必须是0到4294967295范围(包含0和4294967295)内的整数。

[0167] depack-buf-cap:

[0168] 这个参数用信号表示接收方实施方案的能力, 并且以字节为单位指示接收方具有的可用于重构NAL单元解码顺序的解包缓冲器空间量。接收方能够处理对其sprop-depack-buf-bytes参数值小于或等于这个参数的任何流。

[0169] 当不存在时, 推断depack-buf-req值等于0。depack-buf-cap值必须是0到4294967295范围(包含0和4294967295)内的整数。

[0170] 图3是说明可实施本发明中描述的技术的实例视频编码器20的框图。视频编码器20可执行视频切片内的视频块的帧内及帧间译码。帧内译码依赖于空间预测来减少或移除给定视频帧或图片内的视频中的空间冗余。帧间译码依赖于时间预测来减少或移除视频序列的邻近帧或图片内的视频中的时间冗余。帧内模式(I模式)可以指若干基于空间的压缩模式中的任一者。例如单向预测(P模式)或双向预测(B模式)等帧间模式可以指代若干基于时间的压缩模式中的任一者。

[0171] 在图3的实例中, 视频编码器20包含分割单元35、预测处理单元41、滤波单元63、图片存储器64、求和器50、变换处理单元52、量化单元54, 及熵编码单元56。预测处理单元41包含运动估计单元42、运动补偿单元44及帧内预测处理单元46。为了视频块重构, 视频编码器20还包含逆量化单元58、逆变换处理单元60, 及求和器62。滤波单元63既定表示一或多个环路滤波器, 例如解块滤波器、自适应环路滤波器(ALF)及样本自适应偏移(SAO)滤波器。尽管在图3中将滤波单元63展示为环内滤波器, 但在其它配置中, 可将滤波单元63实施为环路后滤波器。图3还展示后处理装置57, 其可以对视频编码器20产生的经编码视频数据执行额外处理。本发明的技术在一些情况下可以通过视频编码器20实施。然而, 在其它情况下本发明的技术可以通过后处理装置57实施。举例来说, 在一些情况下, 相对于图1的包化器21描述的技术可以通过后处理装置57的包化器执行。

[0172] 如图3中所展示, 视频编码器20接收视频数据, 且分割单元35将数据分割成视频块。此分割还可包含分割成切片、拼片或其它较大单元, 以及例如根据LCU及CU的四叉树结构的视频块分割。视频编码器20总体上说明对待编码视频切片内的视频块进行编码的组件。所述切片可以划分成多个视频块(且可能划分成被称作拼片的多组视频块)。预测处理单元41可基于错误结果(例如, 译码速率及失真等级)针对当前视频块选择多种可能译码模式中的一者, 例如多种帧内译码模式中的一者或多种帧间译码模式中的一者。预测处理单元41可将所得经帧内或帧间译码块提供到求和器50以产生残余块数据, 并提供到求和器62以重构经编码块以用作参考图片。

[0173] 预测处理单元41内的帧内预测处理单元46可相对于与待译码当前块在相同帧或切片中的一或多个相邻块执行当前视频块的帧内预测性译码, 以提供空间压缩。预测处理单元41内的运动估计单元42及运动补偿单元44相对于一或多个参考图片中的一或多个预

测性块执行当前视频块的帧间预测译码以提供时间压缩。

[0174] 运动估计单元42可经配置以根据用于视频序列的预定模式来确定用于视频切片的帧间预测模式。预定模式可将序列中的视频切片指明为P切片、B切片或GPB切片。运动估计单元42及运动补偿单元44可以高度集成,但出于概念目的分别加以说明。运动估计单元42所执行的运动估计是产生运动向量的过程,所述过程估计视频块的运动。举例来说,运动向量可以指示当前视频帧或图片内的视频块的PU相对于参考图片内的预测性块的位移。

[0175] 预测性块为发现在像素差值方面紧密匹配待译码视频块的PU的块,所述像素差值可由绝对差总和(SAD)、平方差总和(SSD)或其它差度量来确定。在一些实例中,视频编码器20可计算存储在图片存储器64中的参考图片的子整数像素位置的值。举例来说,视频编码器20可以内插四分之一像素位置、八分之一像素位置或参考图片的其它分数像素位置的值。因此,运动估计单元42可以相对于整数像素位置和分数像素位置执行运动搜索并且输出具有分数像素精确度的运动向量。

[0176] 运动估计单元42通过比较PU的位置与参考图片的预测性块的位置来计算用于经帧间译码切片中的视频块的PU的运动向量。参考图片可以选自第一参考图片列表(列表0)或第二参考图片列表(列表1),其中的每一者识别存储在图片存储器64中的一或多个参考图片。运动估计单元42将计算出来的运动向量发送到熵编码单元56及运动补偿单元44。

[0177] 通过运动补偿单元44执行的运动补偿可以涉及基于通过运动估计(可能执行子像素精度的内插)确定的运动向量提取或产生预测性块。在接收到当前视频块的PU的运动向量后,运动补偿单元44即刻可以在参考图片列表中的一者中定位所述运动向量指向的预测性块。视频编码器20通过从正被译码的当前视频块的像素值减去预测性块的像素值从而形成像素差值来形成残余视频块。像素差值形成用于所述块的残余数据,并且可包含明度和色度差分量两者。求和器50表示可执行此减法运算的一或多个组件。运动补偿单元44还可产生与视频块和视频切片相关联的供视频解码器30在对视频切片的视频块进行解码时使用的语法元素。

[0178] 作为如上文所描述由运动估计单元42和运动补偿单元44执行的帧间预测的替代方案,帧内预测处理单元46可以对当前块进行帧内预测。确切地说,帧内预测处理单元46可以确定用来对当前块进行编码的帧内预测模式。在一些实例中,帧内预测处理单元46可(例如)在不同的编码编次期间使用各种帧内预测模式对当前块进行编码,且帧内预测处理单元46(或在一些实例中为模式选择单元40)可从测试模式中选择适当帧内预测模式来使用。举例来说,帧内预测处理单元46可使用速率-失真分析计算用于各种所测试的帧内预测模式的速率-失真值,且从所述所测试的模式当中选择具有最佳速率失真特性的帧内预测模式。速率失真分析一般确定经编码块与经编码以产生所述经编码块的原始的未经编码块之间的失真(或误差)的量,以及用于产生经编码块的位速率(也就是说,位数目)。帧内预测处理单元46可根据用于各种经编码块的失真和速率计算比率,以确定哪个帧内预测模式对于所述块展现最佳速率失真值。

[0179] 在任何情况下,在选择用于块的帧内预测模式之后,帧内预测处理单元46可将指示块的选定帧内预测模式的信息提供到熵编码单元56。熵编码单元56可以根据本发明的技术对指示所述选定帧内预测模式的信息进行编码。视频编码器20在传输的位流中可包含配置数据,其可包含多个帧内预测模式索引表及多个经修改的帧内预测模式索引表(也称为

码字映射表),对用于各种块的编码上下文的定义,及对最可能帧内预测模式、帧内预测模式索引表及经修改的帧内预测模式索引表的指示以用于所述上下文中的每一者。

[0180] 在预测处理单元41经由帧间预测或帧内预测产生当前视频块的预测性块之后,视频编码器20通过从当前视频块减去预测性块而形成残余视频块。残余块中的残余视频数据可包含在一或多个TU中且应用到变换处理单元52。变换处理单元52使用例如离散余弦变换(DCT)或概念上类似的变换等变换将残余视频数据变换成残余变换系数。变换处理单元52可将残余视频数据从像素域转换到变换域,例如频域。

[0181] 变换处理单元52可将所得变换系数发送到量化单元54。量化单元54将变换系数量化以进一步降低位速率。量化过程可降低与系数中的一些或全部相关联的位深度。可通过调整量化参数来修改量化程度。在一些实例中,量化单元54可接着执行对包含经量化变换系数的矩阵的扫描。替代地,熵编码单元56可以执行所述扫描。

[0182] 在量化之后,熵编码单元56对经量化变换系数进行熵编码。举例来说,熵编码单元56可执行上下文自适应可变长度译码(CAVLC)、上下文自适应二进制算术译码(CABAC)、基于语法的上下文自适应二进制算术译码(SBAC)、概率区间分割熵(PIPE)译码或另一熵译码方法或技术。在由熵编码单元56进行熵编码后,可将经编码位流传输到视频解码器30,或将经编码位流存档以供稍后由视频解码器30传输或检索。熵编码单元56还可对正被译码的当前视频切片的运动向量及其它语法元素进行熵编码。

[0183] 逆量化单元58和逆变换处理单元60分别应用逆量化和逆变换以在像素域中重构残余块,以供稍后用作参考图片的参考块。运动补偿单元44可以通过将残余块加到参考图片列表中的一者内的参考图片中的一者的预测性块中来计算参考块。运动补偿单元44还可将一或多个内插滤波器应用于经重构残余块以计算子整数像素值用于运动估计。求和器62将经重构残余块加到由运动补偿单元44产生的经运动补偿的预测块以产生参考块以用于存储在图片存储器64中。参考块可供运动估计单元42及运动补偿单元44用作参考块以对后续视频帧或图片中的块进行帧间预测。

[0184] 图4是说明可以实施本发明中描述的技术的实例网络实体79和视频解码器30的框图。在图4的实例中,视频解码器30包含熵解码单元80、预测处理单元81、逆量化单元86、逆变换处理单元88、求和器90、滤波单元91和图片存储器92。预测处理单元81包含运动补偿单元82及帧内预测处理单元84。在一些实例中,视频解码器30可执行大体上与相对于来自图3的视频编码器20描述的编码遍次互逆的解码遍次。

[0185] 在解码过程期间,视频解码器30从视频编码器20接收表示经编码视频切片的视频块和相关联语法元素的经编码视频位流。视频解码器30可从网络实体79接收经编码视频位流。网络实体79可例如为服务器、MANE、视频编辑器/剪接器,或经配置以实施上文所描述的技术中的一或多者的其它此装置。网络实体79可包含或不包含视频编码器20。如上文所描述,本发明中所描述的技术中的一些可在网络实体79将经编码视频位流传输到视频解码器30之前由网络实体79实施。在一些视频解码系统中,网络实体79及视频解码器30可为分开的装置的部分,而在其它情况下,关于网络实体79描述的功能性可由包括视频解码器30的相同装置执行。虽然图1将解包器29展示为目的装置14的一部分,但是上文关于解包器29描述的技术也可由网络实体79内的解包器执行。

[0186] 在解码过程期间,视频解码器30从视频编码器20接收表示经编码视频切片的视频

块及相关联语法元素的经编码视频位流。视频块可以例如经由一或多个MANE (例如图1中的MANE 27或图4中的网络实体79) 从视频编码器20路由到视频解码器30。视频解码器30的熵解码单元80对位流进行熵解码以产生经量化系数、运动向量及其它语法元素。熵解码单元80将运动向量及其它语法元素转发到预测处理单元81。视频解码器30可以接收视频切片层级及/或视频块层级的语法元素。

[0187] 当视频切片经译码为经帧内译码(I)切片时,预测处理单元81的帧内预测处理单元84可基于用信号表示的帧内预测模式及来自当前帧或图片的先前经解码块的数据产生用于当前视频切片的视频块的预测数据。当将视频切片译码为经帧间译码(即,B、P或GPB)切片时,预测处理单元81的运动补偿单元82基于从熵解码单元80接收的运动向量和其它语法元素而产生当前视频切片的视频块的预测性块。预测性块可以从参考图片列表中的一者内的参考图片中的一者产生。视频解码器30可以基于存储在图片存储器92中的参考图片使用默认建构技术建构参考帧列表—列表0和列表1。

[0188] 运动补偿单元82通过剖析运动向量及其它语法元素确定用于当前视频切片的视频块的预测信息,并且使用所述预测信息产生用于经解码的当前视频块的预测性块。举例来说,运动补偿单元82使用一些接收到的语法元素确定用于对视频切片的视频块进行译码的预测模式(例如,帧内预测或帧间预测)、帧间预测切片类型(例如,B切片、P切片或GPB切片)、切片的参考图片列表中的一或多者的建构信息、切片的每一经帧间编码的视频块的运动向量、切片的每一经帧间译码的视频块的帧间预测状态和用以对当前视频切片中的视频块进行解码的其它信息。

[0189] 运动补偿单元82还可基于内插滤波器执行内插。运动补偿单元82可使用由视频编码器20在视频块的编码期间使用的内插滤波器来计算参考块的子整数像素的内插值。在此状况下,运动补偿单元82可从接收到的语法元素确定由视频编码器20使用的内插滤波器且使用所述内插滤波器来产生预测性块。

[0190] 逆量化单元86将在位流中提供且由熵解码单元80解码的经量化变换系数逆量化,即,解量化。逆量化过程可包含使用由视频编码器20针对视频切片中的每一视频块所计算的量化参数以确定应当应用的量化程度且同样确定应当应用的逆量化程度。逆变换处理单元88将逆变换应用于变换系数,例如逆DCT、逆整数变换或概念上类似的逆变换过程,以便产生像素域中的残余块。

[0191] 在运动补偿单元82基于运动向量和其它语法元素产生了当前视频块的预测性块之后,视频解码器30通过将来自逆变换处理单元88的残余块与运动补偿单元82产生的对应预测性块求和来形成经解码视频块。求和器90表示可执行此求和运算的组件。在需要时,还可使用环路滤波器(在译码环路中或在译码环路之后)来使像素转变平滑或者以其它方式改进视频质量。滤波单元91既定表示一或多个环路滤波器,例如解块滤波器、自适应环路滤波器(ALF)及样本自适应偏移(SAO)滤波器。尽管在图4中将滤波单元91展示为环路内滤波器,但在其它配置中,可将滤波单元91实施为环路后滤波器。接着将给定帧或图片中的经解码视频块存储于图片存储器92中,所述图片存储器存储用于后续运动补偿的参考图片。图片存储器92还存储经解码视频用于以后在显示装置(例如图1的显示装置32)上呈现。

[0192] 图5是说明形成网络150的一部分的一组实例装置的框图。在这个实例中,网络150包含路由装置154A、154B(路由装置154)和转码装置156。路由装置154及转码装置156既定

表示可形成网络150的部分的少数装置。例如交换器、集线器、网关、防火墙、桥接器及其它此些装置的其它网络装置也可包含在网络150内。另外,可沿着服务器装置152与客户端装置158之间的网络路径提供额外的网络装置。在一些实例中,服务器装置152可对应于源装置12(图1),而客户端装置158可对应于目的地装置14(图1)。路由装置154可以例如是经配置以路由媒体数据的MANE。

[0193] 一般来说,路由装置154实施一或多个路由协议以通过网络150交换网络数据。一般来说,路由装置154执行路由协议以通过网络150发现路线。通过执行此些路由协议,路由装置154B可发现从自身经由路由装置154A到服务器装置152的网络路线。图5的多个装置表示可实施本发明的技术且可经配置以根据本发明的技术处理RTP数据的装置的实例。

[0194] 图6展示根据本发明的技术处理视频数据的方法的实例。图6的技术可以例如由目的地装置14的装置执行,并且更具体地说,可以由目的地装置14的解包器29执行。解包器29根据RTP协议接收第一聚集包(160)。第一聚集包可以例如包含一个有效负载标头和一或多个聚集单元。解包器29可以剖析第一聚集单元以确定第一参数的值(162)。第一参数可以例如对应于上文所论述的DONL参数并且可以指明解码顺序编号。解包器29可以剖析第二聚集单元以确定第二参数的值(164)。第二聚集单元可以在第一聚集单元之后,并且第二参数可以例如对应于上文所论述的DOND参数。基于第一参数和第二参数,解包器29确定第二聚集单元的解码顺序。

[0195] 图7展示根据本发明的技术处理视频数据的方法的实例。图7的技术可以例如由源装置12的装置执行,并且更具体地说可以由源装置12的包化器21执行。包化器21根据RTP协议接收一或多个NAL单元,并且将所述一或多个NAL单元包化成第一聚集包(170)。第一聚集包可以例如包含一个有效负载标头和一或多个聚集单元。包化器21基于包含于第一聚集单元中的NAL单元的解码顺序编号来设置第一聚集单元的第一参数的值(172)。第一参数可以例如对应于上文所论述的DONL参数并且可以指明解码顺序编号。第一参数可以例如指明解码顺序编号的多个最低有效位的值。基于包含于第二聚集单元中的NAL单元的解码顺序与包含于第一聚集单元中的NAL单元的解码顺序编号之间的差值,包化器21可以设置第二聚集单元的第二参数的值(174)。第二聚集单元可以在第一聚集单元之后,并且第二参数可以例如对应于上文所论述的DOND参数。第二参数可以例如识别第一参数与解码顺序编号之间的差值。

[0196] 图8展示根据本发明的技术处理视频数据的方法的实例。图8的技术可以例如由目的地装置14的装置执行,并且更具体地说,可以由目的地装置14的解包器29执行。解包器29接收包含分断式NAL单元的子组的第一分断单元(180)。解包器29剖析分断单元的起始位以确定第一分断单元是否包含分断式NAL单元的开头(182)。起始位可以例如是如上所述的S位。响应于第一分断单元包含分断式NAL单元的开头,并且第一分断单元的传输模式是多会话传输模式和第一参数大于第一值这两个条件中的一或两者,解包器29剖析第二参数以确定分断式NAL单元的解码顺序。第一参数可以例如是如上所述的sprop-depack-buf-nalus参数,并且第一值可以是零。第二参数可以例如是如上所述的DONL参数。目的地装置14可以基于所确定的解码顺序对分断式NAL单元进行解码(186)。

[0197] 图9展示根据本发明的技术处理视频数据的方法的实例。图9的技术可以例如由源装置12的装置执行,并且更具体地说可以由源装置12的包化器21执行。包化器21产生

包括分断式NAL单元的子组的第一分断单元(190)。第一分断单元例如包含分断式NAL单元的开头。包化器21设置分断单元的起始位以指示第一分断单元包含分断式NAL单元的开头(192)。起始位可以例如是如上所述的S位。响应于第一分断单元包含分断式NAL单元的开头,并且第一分断单元的传输模式是多会话传输模式和第一参数大于第一值这两个条件中的一或两者,包化器21设置第二参数以指示分断式NAL单元的解码顺序。第一参数可以例如是如上所述的sprop-depack-buf-nalus参数,并且第一值可以是零。第二参数可以例如是如上所述的DONL参数。包化器21可以传输分断式NAL单元(196)。第一参数可以例如指明在解包缓冲器中按照接收顺序在第一NAL单元之前并且按照解码顺序在第一NAL单元之后的NAL单元的最大数目,并且第二参数可以指明解码顺序编号的最低有效位的数目的值。

[0198] 图10展示根据本发明的技术处理视频数据的方法的实例。图10的技术可以例如由例如目的地装置14的装置执行,并且更具体地说,可以由目的地装置14的解包器29执行。解包器29接收包括第一NAL单元的第一RTP包(200)。响应于第一RTP包的传输模式是单个会话传输模式并且第一参数等于第一值,解包器29基于第一NAL单元的传输顺序确定第一NAL单元的解码顺序编号(202)。第一参数可以例如是如上所述的sprop-depack-buf-nalus参数,并且所述值可以等于零。

[0199] 图11展示根据本发明的技术处理视频数据的方法的实例。图11的技术可以例如由例如源装置12的装置执行,并且更具体地说可以由源装置12的包化器21执行。包化器21产生包括第一NAL单元的RTP包(210)。响应于第一RTP包的传输模式是单个会话传输模式并且第一参数等于第一值,基于第一NAL单元的解码顺序设置第一NAL单元的传输顺序(212)。第一参数可以例如是如上所述的sprop-depack-buf-nalus参数。

[0200] 在一或多个实例中,所描述功能可以用硬件、软件、固件或其任何组合来实施。如果用软件实施,则所述功能可作为一或多个指令或代码在计算机可读媒体上存储或传输,且由基于硬件的处理单元执行。计算机可读媒体可包含计算机可读存储媒体,其对应于例如数据存储媒体或通信媒体等有形媒体,通信媒体包含促进将计算机程序从一处传送到另一处的任何媒体(例如,根据通信协议)。以此方式,计算机可读媒体通常可以对应于(1)有形计算机可读存储媒体,其是非暂时性的,或(2)通信媒体,例如信号或载波。数据存储媒体可以是可由一或多个计算机或一个或多个处理器存取以检索用于实施本发明中描述的技术的指令、代码及/或数据结构的任何可用媒体。计算机程序产品可包含计算机可读媒体。

[0201] 借助于实例而非限制,此类计算机可读存储媒体可包括RAM、ROM、EEPROM、CD-ROM或其它光盘存储装置、磁盘存储装置或其它磁性存储装置、快闪存储器或任何其它用来存储呈指令或数据结构的形式的所要程序代码并且可由计算机存取的媒体。并且,可恰当地将任何连接称作计算机可读媒体。举例来说,如果使用同轴电缆、光缆、双绞线、数字订户线(DSL)或例如红外线、无线电和微波等无线技术从网站、服务器或其它远程源传输指令,那么同轴电缆、光缆、双绞线、DSL或例如红外线、无线电和微波等无线技术包含在媒体的定义中。然而,应理解,所述计算机可读存储媒体和数据存储媒体并不包含连接、载波、信号或其它瞬时性媒体,而是实际上针对于非瞬时性有形存储媒体。如本文所使用,磁盘及光盘包含压缩光盘(CD)、激光光盘、光学光盘、数字多功能光盘(DVD)、软性磁盘及蓝光光盘,其中磁盘通常以磁性方式再现数据,而光盘用激光以光学方式再现数据。上述各项的组合也应包含在计算机可读媒体的范围内。

[0202] 指令可以由一或多个处理器执行,所述一或多个处理器例如是一或多个数字信号处理器(DSP)、通用微处理器、专用集成电路(ASIC)、现场可编程逻辑阵列(FPGA)或其它等效的集成或离散逻辑电路。因此,如本文中所使用的术语“处理器”可指代上述结构或适合于实施本文中所描述的技术的任何其它结构中的任一者。另外,在一些方面,本文所述的功能性可以在经配置用于编码及解码的专用硬件及/或软件模块内提供,或者并入在组合式编解码器中。并且,所述技术可以完全实施于一或多个电路或逻辑元件中。

[0203] 本发明的技术可在多种多样的装置或设备中实施,包含无线手持机、集成电路(IC)或一组IC(例如,芯片组)。本发明中描述各种组件、模块或单元来强调经配置以执行所揭示的技术的装置的功能方面,但未必需要通过不同硬件单元实现。实际上,如上文所描述,各种单元可以配合合适的软件及/或固件组合在编解码器硬件单元中,或者通过互操作硬件单元的聚集来提供,所述硬件单元包含如上文所描述的一或多个处理器。

[0204] 已描述各种实例。这些和其它实例在所附权利要求书的范围内。

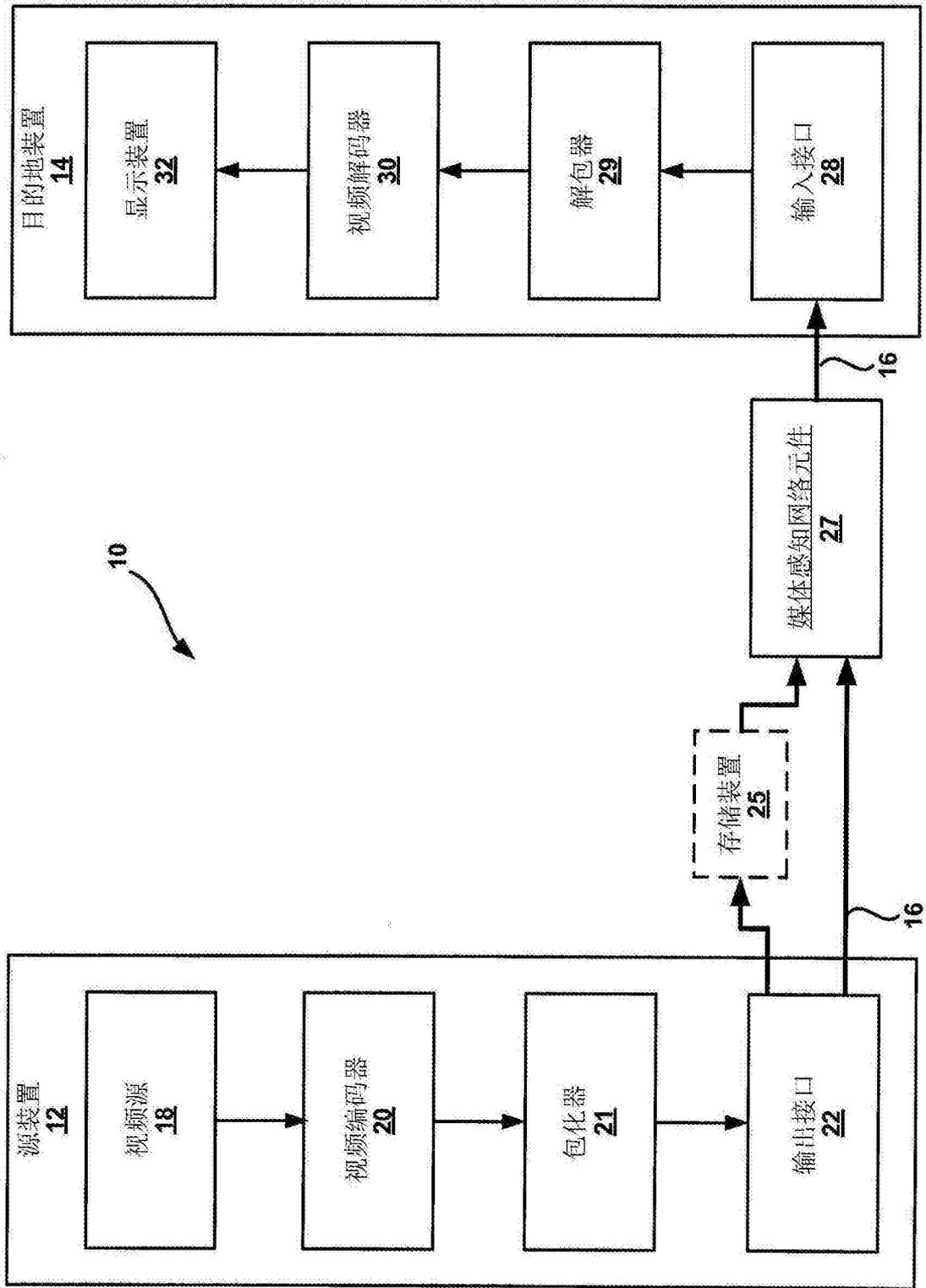


图1

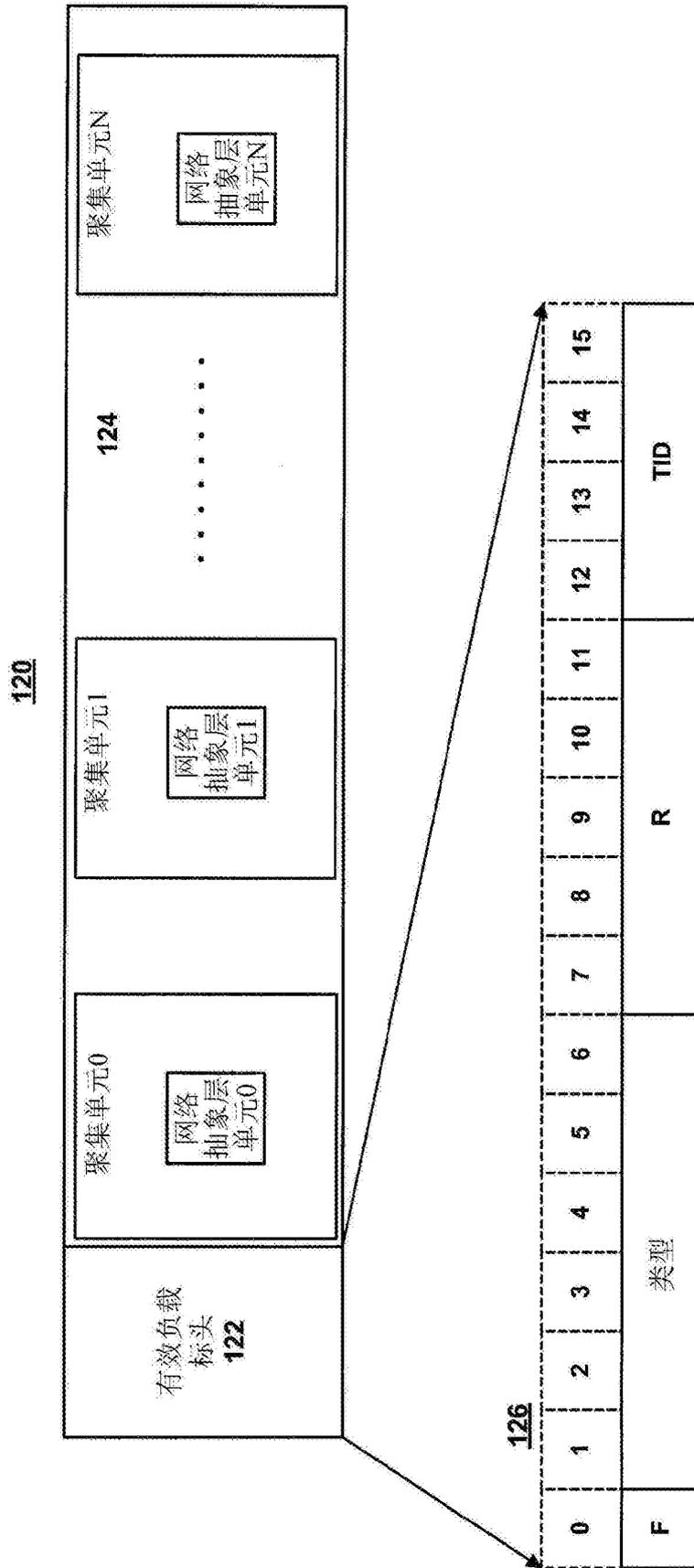


图2

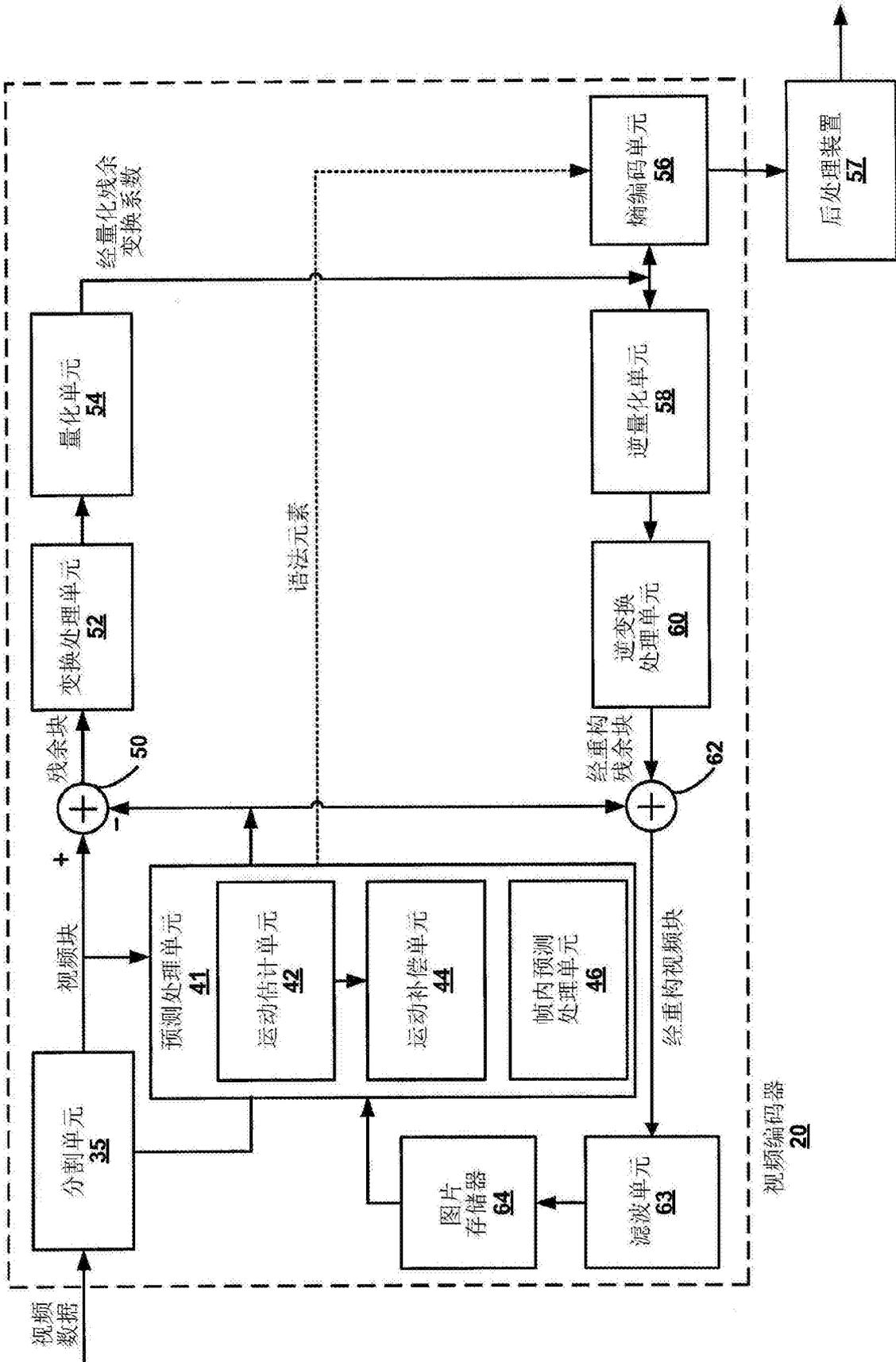


图3

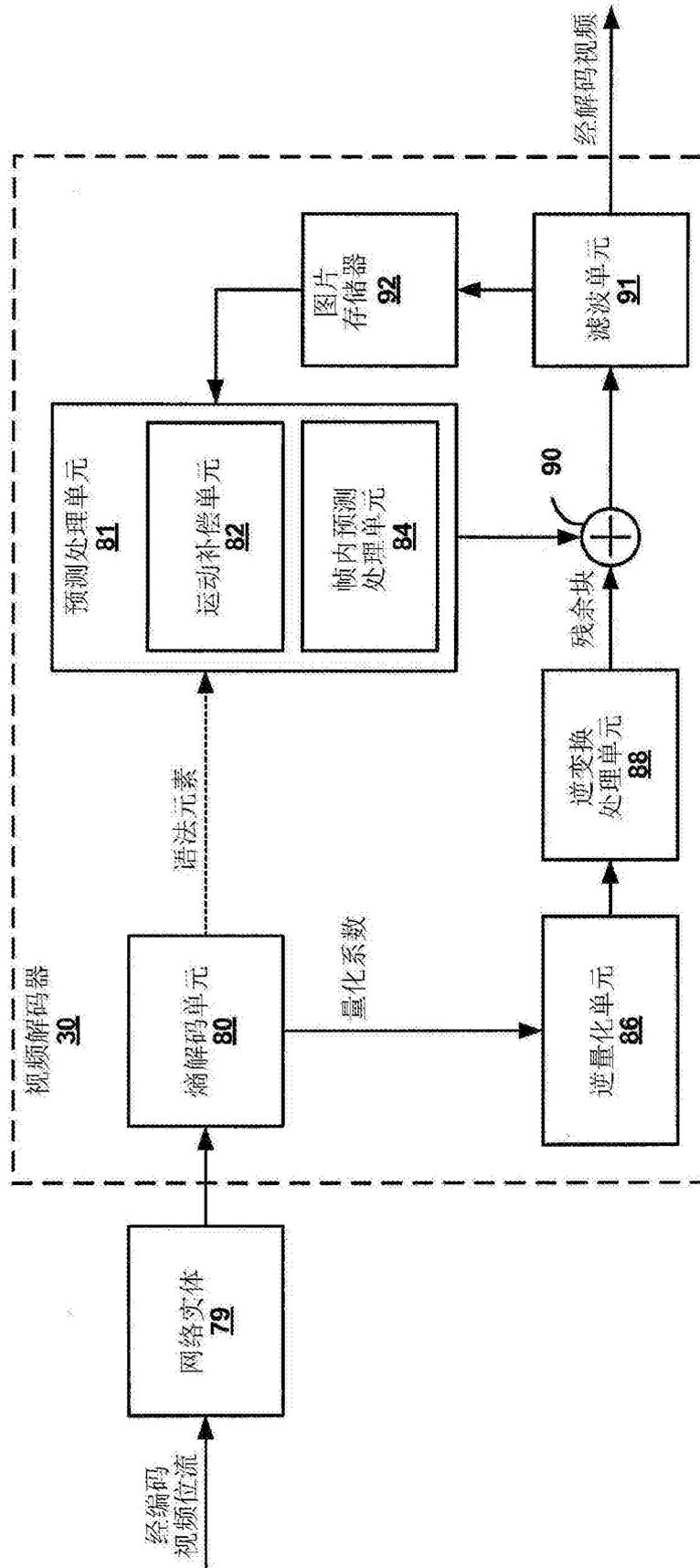


图4

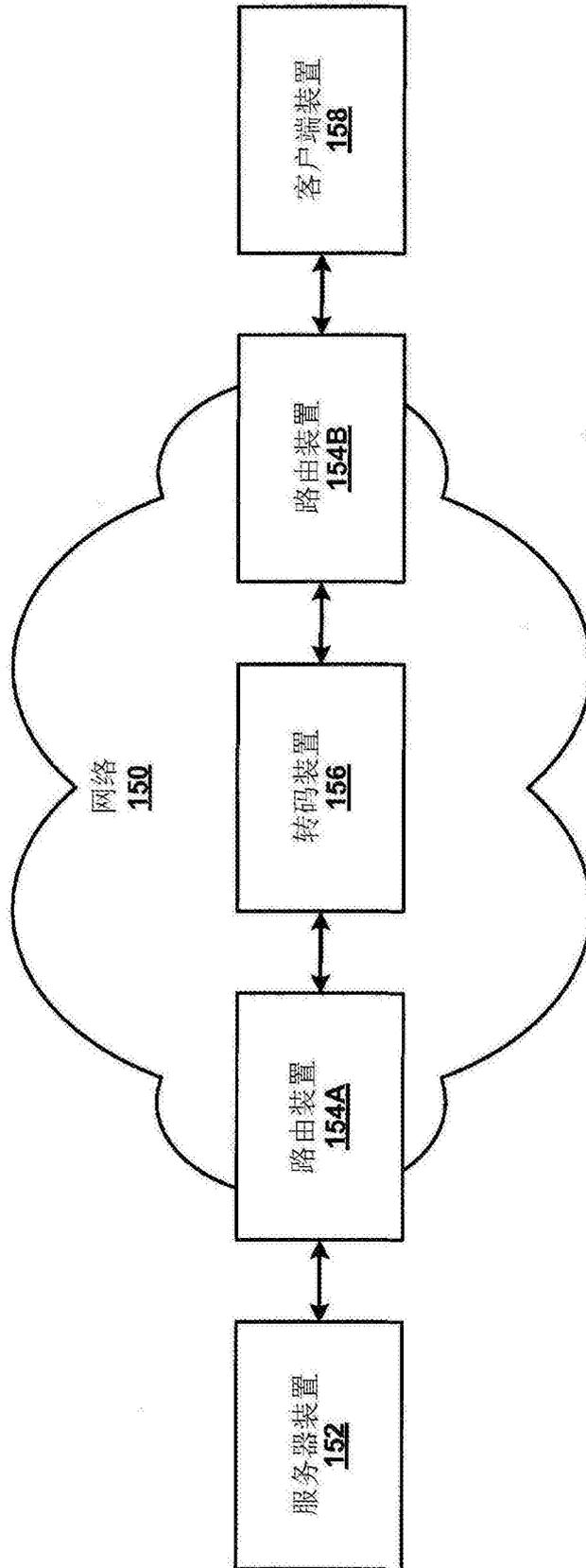


图5

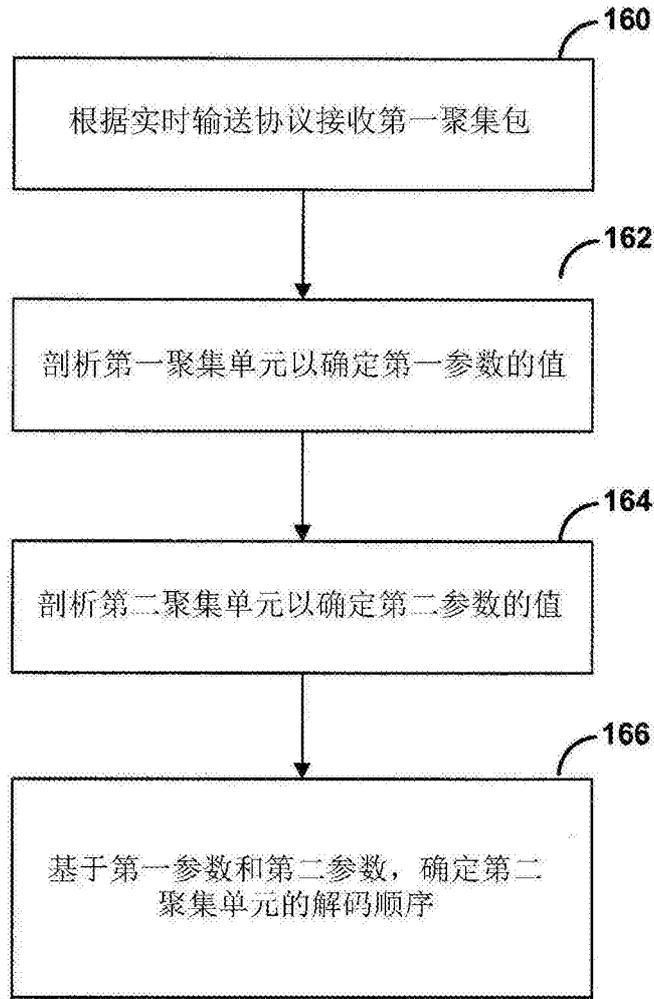


图6

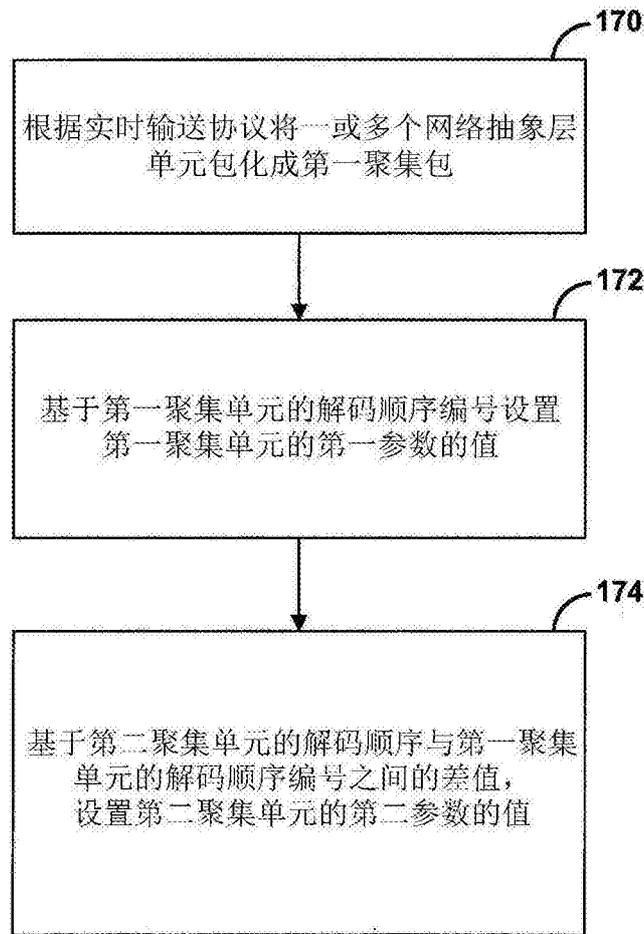


图7

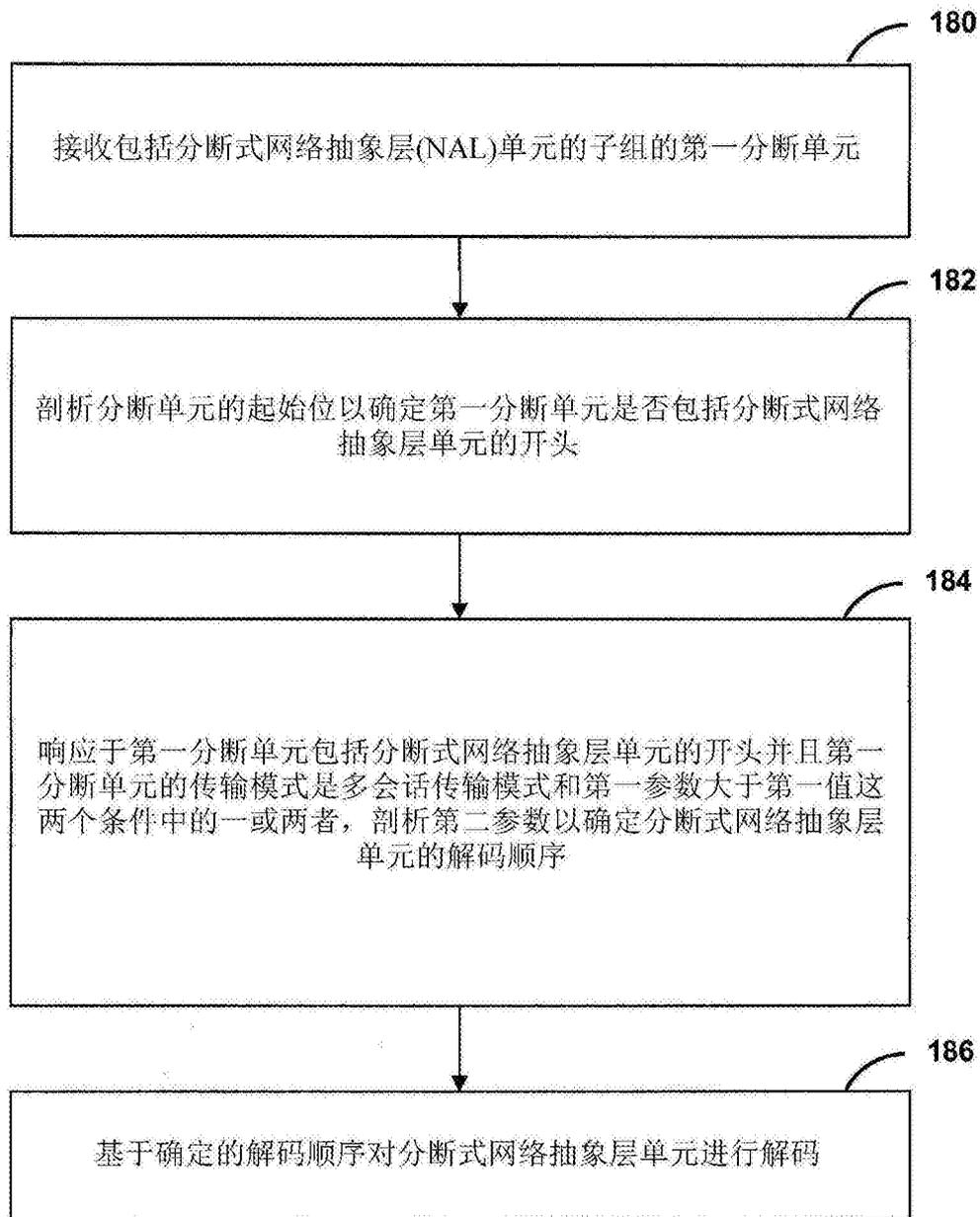


图8

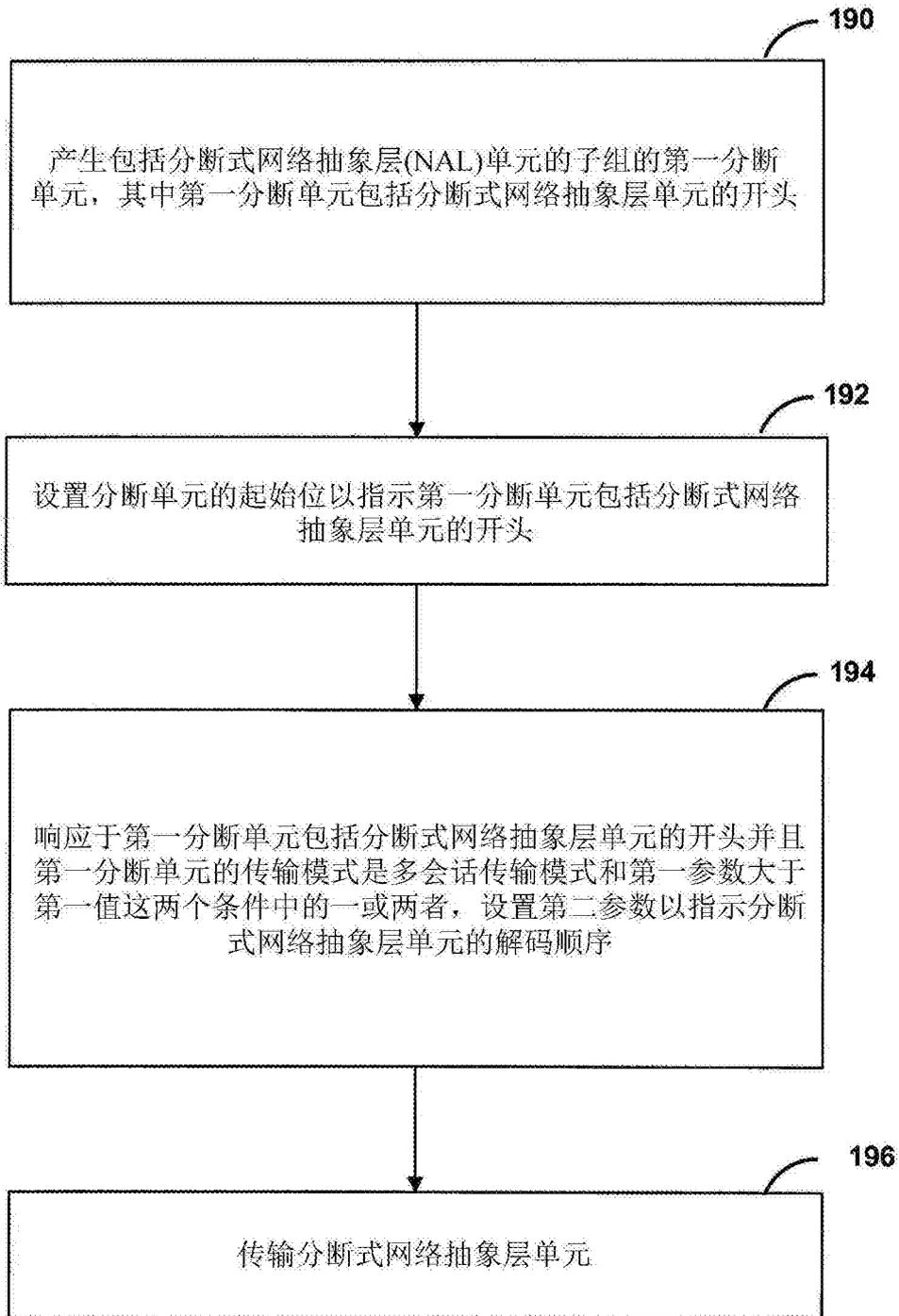


图9

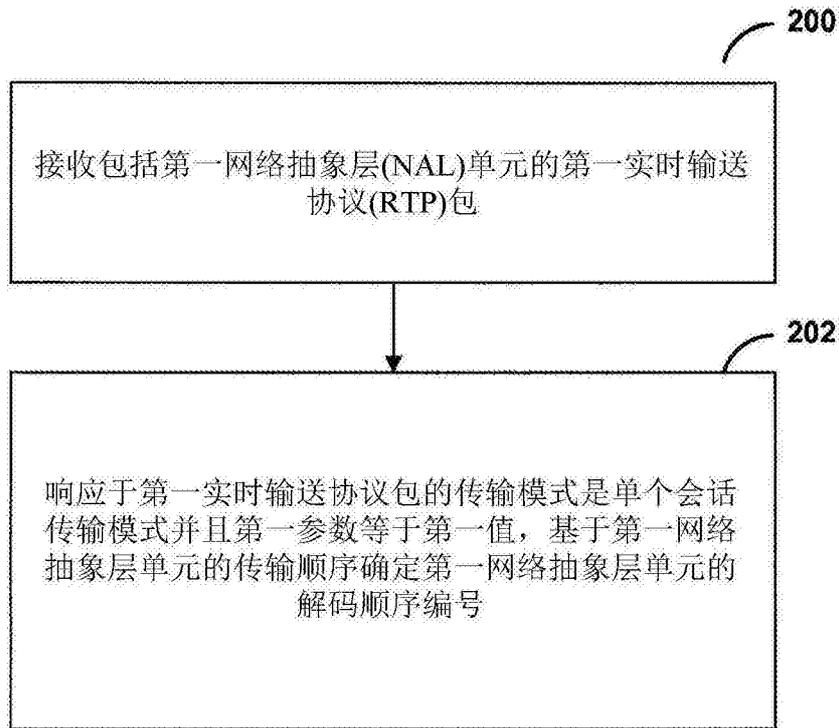


图10

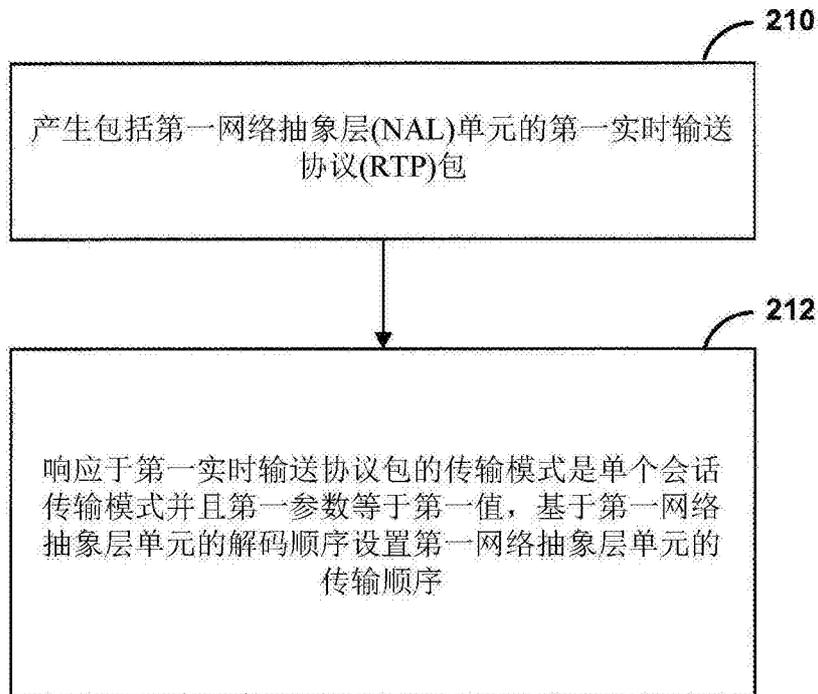


图11