



(12) 发明专利申请

(10) 申请公布号 CN 116349226 A

(43) 申请公布日 2023.06.27

(21) 申请号 202180064854.4

D·基安法尔 A·J·维格斯

(22) 申请日 2021.09.24

(74) 专利代理机构 北京市柳沈律师事务所

(30) 优先权数据

11105

63/085,092 2020.09.29 US

专利代理师 安之斐

17/448,658 2021.09.23 US

(51) Int.Cl.

(85) PCT国际申请进入国家阶段日

H04N 19/176 (2006.01)

2023.03.22

(86) PCT国际申请的申请数据

PCT/US2021/052008 2021.09.24

(87) PCT国际申请的公布数据

W02022/072245 EN 2022.04.07

(71) 申请人 高通股份有限公司

地址 美国加利福尼亚州

(72) 发明人 王洪涛 V·M·S·A·科特拉

J·陈 M·卡尔切维茨

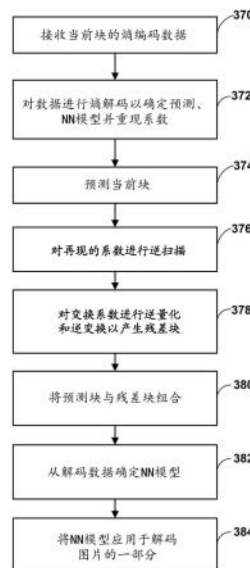
权利要求书4页 说明书33页 附图9页

(54) 发明名称

用于视频编解码期间进行滤波的多神经网络模型

(57) 摘要

一种用于对解码视频数据进行滤波的示例设备包括:存储器,被配置为存储视频数据;以及一个或多个处理器,其在电路中实施并且被配置为:对视频数据的图片进行解码;对表示要用于对解码图片的一部分进行滤波的神经网络模型的语法元素的值进行编解码,该值表示对预定义神经网络模型集合的索引,该索引对应于预定义神经网络模型集合中的神经网络模型;以及使用对应于索引的神经网络模型来对解码图片的部分进行滤波。



1. 一种对解码视频数据进行滤波的方法,所述方法包括:  
对视频数据的图片进行解码;  
对表示要用于对解码图片的一部分进行滤波的神经网络模型的语法元素的值进行编解码,所述值表示对预定义神经网络模型集合的索引,所述索引对应于所述预定义神经网络模型集合中的所述神经网络模型;以及  
使用对应于所述索引的所述神经网络模型来对所述解码图片的所述部分进行滤波。
2. 根据权利要求1所述的方法,还包括对表示所述预定义神经网络模型集合的子集的值进行编解码,所述神经网络模型是所述预定义神经网络模型集合的所述子集中的一个,所述子集小于所述预定义神经网络模型集合。
3. 根据权利要求1所述的方法,其中,表示所述神经网络模型的所述语法元素的所述值是表示多个神经网络模型的值。
4. 根据权利要求3所述的方法,其中,滤波包括:  
将由所述值表示的所述多个神经网络模型中的每一个单独地应用于所述部分以形成不同的结果;以及  
组合每个所述结果以形成最终经滤波的部分。
5. 根据权利要求1所述的方法,其中,对所述语法元素的所述值进行编解码包括在以下中的一个中对所述语法元素的所述值进行编解码:视频参数集VPS、序列参数集SPS、图片参数集PPS、自适应参数集APS、时段内级别、图片组GOP级别、GOP中的时间层级别、图片级别、条带级别、编解码树单元CTU级别、或所述图片的网格的网格级别。
6. 根据权利要求1所述的方法,还包括根据网格来分割所述图片,其中,所述部分包括所述图片的所述网格的元素。
7. 根据权利要求6所述的方法,还包括确定所述网格的元素的数量。
8. 根据权利要求7所述的方法,其中,所述语法元素包括第一语法元素,并且其中,确定所述网格的元素的数量包括对以下中的至少一个的第二语法元素进行解码:视频参数集VPS、序列参数集SPS、图片参数集PPS、图片头、条带头、自适应参数集APS、时段内级别、图片组GOP级别、GOP中的时间层级别、图片级别、条带级别、编解码树单元CTU级别、或网格尺寸级别。
9. 根据权利要求7所述的方法,其中,所述语法元素包括第一语法元素,并且其中,确定所述网格中的元素的数量包括将表示对可能的网格尺寸的集合的索引的第二语法元素的值进行解码。
10. 根据权利要求1所述的方法,还包括在确定所述神经网络模型之前确定要应用所述神经网络模型。
11. 根据权利要求10所述的方法,其中,所述语法元素包括第一语法元素,并且其中,确定要应用所述神经网络模型包括对指示要应用所述神经网络模型的第二语法元素的值进行解码。
12. 根据权利要求11所述的方法,其中,所述第二语法元素是以下中的至少一个:视频参数集VPS、序列参数集SPS、图片参数集PPS、图片头、条带头、自适应参数集APS、时段内级别、图片组GOP级别、GOP中的时间层级别、图片级别、条带级别、编解码树单元CTU级别、或网格尺寸级别。

13. 根据权利要求1所述的方法,其中,所述解码图片的所述部分包括所述解码图片的颜色分量的一部分,所述颜色分量包括亮度分量、蓝色色调度分量或红色色调度分量之一。

14. 根据权利要求13所述的方法,其中,所述语法元素包括第一语法元素,所述方法还包括对联合表示使用所述神经网络模型对所述解码图片的所述颜色分量中的每一个进行滤波的第二语法元素进行编解码。

15. 根据权利要求1所述的方法,其中,所述语法元素的所述值包括所述图片的所述部分的量化参数QP。

16. 根据权利要求1所述的方法,还包括在对所述图片进行解码之前对所述图片进行编码,其中对所述语法元素的所述值进行编解码包括对所述语法元素的所述值进行编码。

17. 根据权利要求16所述的方法,还包括根据率失真计算来确定所述神经网络模型。

18. 一种用于对解码视频数据进行滤波的设备,所述设备包括:

存储器,被配置为存储视频数据;以及

一个或多个处理器,在电路中实施并且被配置为:

对视频数据的图片进行解码;

对表示要用于对解码图片的一部分进行滤波的神经网络模型的语法元素的值进行编解码,所述值表示对预定义神经网络模型集合的索引,所述索引对应于所述预定义神经网络模型集合中的所述神经网络模型;以及

使用对应于所述索引的所述神经网络模型来对所述解码图片的所述部分进行滤波。

19. 根据权利要求18所述的设备,其中,所述一个或多个处理器还被配置为对表示所述预定义神经网络模型集合的子集的值进行编解码,所述神经网络模型是所述预定义神经网络模型集合的所述子集中的一个,所述子集小于所述预定义神经网络模型集合。

20. 根据权利要求18所述的设备,其中,表示所述神经网络模型的所述语法元素的所述值是表示多个神经网络模型的值,并且其中,为了对所述解码图片的所述部分进行滤波,所述一个或多个处理器被配置为:

将由所述值表示的所述多个神经网络模型中的每一个单独地应用于所述部分以形成不同的结果;以及

组合每个所述结果以形成最终经滤波的部分。

21. 根据权利要求18所述的设备,其中,所述一个或多个处理器被配置为在以下中的一个中对所述语法元素的所述值进行编解码:视频参数集VPS、序列参数集SPS、图片参数集PPS、自适应参数集APS、时段内级别、图片组GOP级别、GOP中的时间层级、图片级别、条带级别、编解码树单元CTU级别、或所述图片的网格的网格级别。

22. 根据权利要求18所述的设备,其中,所述一个或多个处理器被配置为根据网格来分割所述图片,其中,所述部分包括所述图片的所述网格的元素。

23. 根据权利要求22所述的设备,其中,所述语法元素包括第一语法元素,并且其中,所述一个或多个处理器还被配置为对表示所述网格的元素的数量的第二语法元素的值进行解码,所述语法元素包括在以下中的至少一个中:视频参数集VPS、序列参数集SPS、图片参数集PPS、图片头、条带头、自适应参数集APS、时段内级别、图片组GOP级别、GOP中的时间层级、图片级别、条带级别、编解码树单元CTU级别、或网格尺寸级别。

24. 根据权利要求18所述的设备,其中,所述语法元素包括第一语法元素,并且其中,所述一个或多个处理器被配置为在确定所述神经网络模型之前确定要应用所述神经网络模型,其中,为了确定要应用所述神经网络模型,所述一个或多个处理器被配置为对指示要应用所述神经网络模型的第二语法元素的值进行解码。

25. 根据权利要求18所述的设备,其中,所述解码图片的所述部分包括所述解码图片的颜色分量的一部分,所述颜色分量包括亮度分量、蓝色色调度分量或红色色调度分量之一。

26. 根据权利要求18所述的设备,其中,所述语法元素的所述值包括所述图片的所述部分的量化参数QP。

27. 根据权利要求18所述的设备,其中,所述一个或多个处理器还被配置为在对所述图片进行解码之前对所述图片进行编码,并且其中,为了对所述语法元素的所述值进行编解码,所述一个或多个处理器被配置为对所述语法元素的所述值进行编码。

28. 根据权利要求27所述的设备,其中,所述一个或多个处理器还被配置为根据率失真计算来确定所述神经网络模型。

29. 根据权利要求18所述的设备,还包括被配置为显示所述解码视频数据的显示器。

30. 根据权利要求18所述的设备,其中,所述设备包括相机、计算机、移动设备、广播接收器设备或机顶盒中的一个或多个。

31. 一种计算机可读存储介质,其上存储有指令,所述指令在被执行时使处理器:

对视频数据的图片进行解码;

对表示要用于对解码图片的一部分进行滤波的神经网络模型的语法元素的值进行编解码,所述值表示对预定义神经网络模型集合的索引,所述索引对应于所述预定义神经网络模型集合中的所述神经网络模型;以及

使用对应于所述索引的所述神经网络模型来对所述解码图片的所述部分进行滤波。

32. 根据权利要求31所述的计算机可读存储介质,还包括用于使所述处理器对表示所述预定义神经网络模型集合的子集的值进行编解码的指令,所述神经网络模型是所述预定义神经网络模型集合的所述子集中的一个,所述子集小于所述预定义神经网络模型集合。

33. 根据权利要求31所述的计算机可读存储介质,其中,表示所述神经网络模型的所述语法元素的所述值是表示多个神经网络模型的值,并且其中,使所述处理器对所述解码图片的所述部分进行滤波的所述指令包括用于使所述处理器进行以下操作的指令:

将由所述值表示的所述多个神经网络模型中的每一个单独地应用于所述部分以形成不同的结果;以及

组合每个所述结果以形成最终经滤波的部分。

34. 根据权利要求31所述的计算机可读存储介质,其中,用于使所述处理器对所述语法元素的所述值进行编解码的所述指令包括用于使所述处理器在以下中的一个中对所述语法元素的所述值进行编解码的指令:视频参数集VPS、序列参数集SPS、图片参数集PPS、自适应参数集APS、时段内级别、图片组GOP级别、GOP中的时间层级别、图片级别、条带级别、编解码树单元CTU级别、或所述图片的网格的网格级别。

35. 根据权利要求31所述的计算机可读存储介质,还包括用于使所述处理器根据网格来分割所述图片的指令,其中,所述部分包括所述图片的所述网格的元素。

36. 根据权利要求35所述的计算机可读存储介质,其中,所述语法元素包括第一语法元素,还包括用于使所述处理器对表示所述网格的元素的数量的第二语法元素的值进行解码的指令,所述语法元素包括在以下中的至少一个中:视频参数集VPS、序列参数集SPS、图片参数集PPS、图片头、条带头、自适应参数集APS、时段内级别、图片组GOP级别、GOP中的时间级别、图片级别、条带级别、编解码树单元CTU级别、或网格尺寸级别。

37. 根据权利要求31所述的计算机可读存储介质,其中,所述语法元素包括第一语法元素,还包括用于使所述处理器在确定所述神经网络模型之前确定要应用所述神经网络模型的指令,其中,用于使所述处理器确定要应用所述神经网络模型的所述指令包括用于使所述处理器对指示要应用所述神经网络模型的第二语法元素的值进行解码的指令。

38. 根据权利要求31所述的计算机可读存储介质,其中,所述解码图片的所述部分包括所述解码图片的颜色分量的一部分,所述颜色分量包括亮度分量、蓝色色调度分量或红色色调度分量之一。

39. 根据权利要求31所述的计算机可读存储介质,其中,所述语法元素的所述值包括所述图片的所述部分的量化参数QP。

40. 根据权利要求31所述的计算机可读存储介质,还包括使所述处理器在对所述图片进行解码之前对所述图片进行编码的指令,并且其中,使所述处理器对所述语法元素的所述值进行编解码的所述指令包括使所述处理器对所述语法元素的所述值进行编码的指令。

41. 根据权利要求40所述的计算机可读存储介质,还包括用于使所述处理器根据率失真计算来确定所述神经网络模型的指令。

42. 一种用于对解码视频数据进行滤波的设备,所述设备包括:

用于对视频数据的图片进行解码的部件;

用于对表示要用于对解码图片的一部分进行滤波的神经网络模型的语法元素的值进行编解码的部件,所述值表示对预定义神经网络模型集合的索引,所述索引对应于所述预定义神经网络模型集合中的所述神经网络模型;以及

用于使用对应于所述索引的所述神经网络模型来对所述解码图片的所述部分进行滤波的部件。

## 用于视频编解码期间进行滤波的多神经网络模型

[0001] 相关申请的引用

[0002] 本申请要求于2021年9月23日提交的美国申请No.17/448,658,以及于2020年9月29日提交的美国临时专利申请No.63/085,092的优先权,这两个申请的全部内容以引用方式并入本文。于2021年9月23日提交的美国申请No.17/448,658要求于2020年9月29日提交的美国临时专利申请No.63/085,092的权益。

### 技术领域

[0003] 本公开涉及视频编解码(coding),包括视频编码和视频解码。

### 背景技术

[0004] 数字视频功能能够集成到多种设备中,包括数字电视、数字直接广播系统、无线广播系统、个人数字助理(PDA)、膝上型或台式计算机、平板计算机、电子书阅读器、数码相机、数字记录设备、数字媒体播放器、视频游戏设备、视频游戏控制台、蜂窝式或卫星无线电话、所谓的“智能电话”、视频电话会议设备、视频流式传输设备等。数字视频设备实施视频编解码技术,诸如那些在由MPEG-2、MPEG-4、ITU-T H.263、ITU-T H.264/MPEG-4第10部分、高级视频编解码(AVC)、ITU-T H.265/高效视频编解码(HEVC)所定义的标准以及此类标准的扩展中所描述的技术。通过实施此类视频编解码技术,视频设备可以更高效地发送、接收、编码、解码和/或存储数字视频信息。

[0005] 视频编解码技术包括空域(图片内)预测和/或时域(图片间)预测来减少或消除视频序列中固有的冗余。对于基于块的视频编解码,可以将视频条带(例如,视频图片或视频图片的一部分)分割为视频块,其也可以称为编解码树单元(CTU)、编解码单元(CU)和/或编解码节点。对于图片的帧内编解码(I)条带中的视频块,可以使用相对于同一图片中相邻块中的参考样点的空域预测进行编码。图片的帧间编解码(P或B)条带中的视频块可以使用相对于同一图片中相邻块中的参考样点的空域预测或者相对于其他参考图片中的参考样点的时域预测。图片可以称为帧,并且参考图片可以称为参考帧。

### 发明内容

[0006] 一般地,本公开描述了用于对可能失真的解码(例如,重建)图片进行滤波的技术。滤波过程可以基于神经网络技术。该滤波过程可以在诸如ITU-T H.266/通用视频编解码(VVC)的扩展或者后续代的视频编解码标准的高级视频编解码器以及任何其他视频编解码器的上下文中使用。具体地,视频编码器可以为解码图片(the decoded picture)的一部分选择神经网络模型,以用于对解码图片的该部分进行滤波。例如,视频编码器可以执行率失真优化(RDO)技术来确定神经网络模型。可替代地,视频编码器可以确定解码图片的该部分的量化参数(QP),并确定该QP被映射到的神经网络模型。视频编码器可以使用QP本身或表示对可用神经网络模型的集合(或集合的子集)的索引的语法元素的单独值来信令通知所确定的神经网络模型。

[0007] 在一个示例中,一种对解码视频数据进行滤波的方法包括:对视频数据的图片进行解码;对表示要用于对解码图片的一部分进行滤波的神经网络模型的语法元素的值进行编解码,该值表示对预定义神经网络模型集合的索引,该索引对应于预定义神经网络模型集合中的神经网络模型;以及使用对应于索引的神经网络模型来对解码图片的该部分进行滤波。

[0008] 在另一示例中,一种用于对解码视频数据进行滤波的设备包括:存储器,被配置为存储视频数据;以及一个或多个处理器,其在电路中实现并且被配置为:对视频数据的图片进行解码;对表示要用于对解码图片的一部分进行滤波的神经网络模型的语法元素的值进行编解码,该值表示对预定义神经网络模型集合的索引,该索引对应于预定义神经网络模型集合中的神经网络模型;以及使用对应于索引的神经网络模型来对解码图片的该部分进行滤波。

[0009] 在另一示例中,一种计算机可读存储介质,其上存储有指令,该指令在被运行时使处理器:对视频数据的图片进行解码;对表示要用于对解码图片的一部分进行滤波的神经网络模型的语法元素的值进行编解码,该值表示对预定义神经网络模型集合的索引,该索引对应于预定义神经网络模型集合中的神经网络模型;以及使用对应于索引的神经网络模型来对解码图片的该部分进行滤波。

[0010] 在另一示例中,一种用于对解码视频数据进行滤波的设备包括:用于对视频数据的图片进行解码的部件;用于对表示要用于对解码图片的一部分进行滤波的神经网络模型的语法元素的值进行编解码的部件,该值表示对预定义神经网络模型集合的索引,该索引对应于预定义神经网络模型集合中的神经网络模型;以及用于使用对应于索引的神经网络模型来对解码图片的该部分进行滤波的部件。

[0011] 在附图和以下描述中阐述一个或多个示例的细节。根据说明书、附图和权利要求书,其他特征、目的和优势将显而易见。

## 附图说明

[0012] 图1是示出可以执行本公开的技术的示例视频编码和解码系统的框图。

[0013] 图2A和图2B是示出示例二叉树(QTBT)结构和对应的编解码树单元(CTU)的概念图。

[0014] 图3是示出混合视频编解码框架的概念图。

[0015] 图4是示出使用图片组(GOP)尺寸为16的分层预测结构的概念图。

[0016] 图5是示出具有四层的基于神经网络的滤波器的概念图。

[0017] 图6是示出可以执行本公开的技术的示例视频编码器的框图。

[0018] 图7是示出可以执行本公开的技术的示例视频解码器的框图。

[0019] 图8是示出根据本公开的技术用于对当前块进行编码的示例方法的流程图。

[0020] 图9是示出根据本公开的技术用于对当前块进行解码的示例方法的流程图。

## 具体实施方式

[0021] 视频编解码标准包括ITU-T H.261、ISO/IEC MPEG-1可视化、ITU-T H.262或ISO/IEC MPEG-2可视化、ITU-T H.263、ISO/IEC MPEG-4可视化和ITU-T H.264(也称为ISO/IEC

MPEG-4AVC)、高效视频编解码(HEVC)或ITU-T H.265(包括其范围扩展)、多视图扩展(MV-HEVC)和可缩放扩展(SHVC)。另一示例视频编解码标准是通用视频编解码(VVC)或ITU-T H.266,其由ITU-T视频编码专家组(VCEG)和ISO/IEC运动图像专家组(MPEG)的联合视频专家小组(JVET)开发。VVC规范的版本1,以下称为“VVC FDIS”,可从[http://phenix.int-evry.fr/jvet/doc\\_end\\_user/documents/19\\_Teleconference/WG\\_11/jvet-s\\_2001-v17.zip](http://phenix.int-evry.fr/jvet/doc_end_user/documents/19_Teleconference/WG_11/jvet-s_2001-v17.zip)获得。

[0023] 本公开的技术通常针对使用基于神经网络的滤波器的滤波技术。典型地,这样的滤波器在为一般视频数据提供良好结果的大数据集上被训练,但是对于视频数据的特定序列可能不是最佳的。相比之下,本公开描述了基于多模型神经网络的滤波,其可以对于特定视频序列产生更好的结果(例如,在比特率和失真方面)。

[0024] 图1是示出可以执行本公开的技术的示例视频编码和解码系统100的框图。本公开的技术通常针对对视频数据进行编解码(编码和/或解码)。一般地,视频数据包括用于处理视频的任何数据。因此,视频数据可以包括原始的未编码的视频、编码的视频、解码的(例如,重建的)视频以及视频元数据(诸如,信令通知的数据)。

[0025] 如图1所示,在此示例中,系统100包括提供将由目的地设备116进行解码和显示的编码的视频数据的源设备102。具体地,源设备102经由计算机可读介质110将视频数据提供给目的地设备116。源设备102和目的地设备116可以包括多种设备中的任何一种,包括台式计算机、笔记本(即,膝上型)计算机、移动设备、平板计算机、机顶盒、手持电话(诸如智能手机)、电视、相机、显示设备、数字媒体播放器、视频游戏控制台、视频流式传输设备等。在一些情况下,源设备102和目的地设备116可以被配备用于无线通信,并且因此可以称为无线通信设备。

[0026] 在图1的示例中,源设备102包括视频源104、存储器106、视频编码器200和输出接口108。目的地设备116包括输入接口122、视频解码器300、存储器120和显示设备118。根据本公开,源设备102的视频编码器200和目的地设备116的视频解码器300可以被配置为应用于使用多个神经网络模型进行滤波的技术。因此,源设备102表示视频编码设备的示例,而目的地设备116表示视频解码设备的示例。在其他示例中,源设备和目的地设备可以包括其他组件或布置。例如,源设备102可以从诸如外部照相机的外部视频源接收视频数据。同样,目的地设备116可以与外部显示设备接口,而不包括集成的显示设备。

[0027] 如图1所示的系统100仅是一个示例。一般地,任何数字视频编码和/或解码设备都可以使用多个神经网络模型执行用于滤波的技术。源设备102和目的地设备116仅仅是此类编解码设备的示例,其中,源设备102生成用于传输到目的地设备116的编解码视频数据。本公开将“编解码”设备表示为执行数据编解码(编码和/或解码)的设备。因此,视频编码器200和视频解码器300表示编解码设备的示例,具体地,分别是视频编码器和视频解码器。在一些示例中,源设备102和目的地设备116可以以基本上对称的方式操作,使得源设备102和目的地设备116中的每一个均包括视频编码和解码组件。因此,系统100可以支持视频源设备102与目的地设备116之间的单向或双向视频传输,例如,用于视频流式传输、视频回放、视频广播或者视频电话。

[0028] 一般地,视频源104表示视频数据源(即,原始的、未编码的视频数据),并将视频数据的连续的图片(也称为“帧”)序列提供给视频编码器200,其对图片的数据进行编码。源设

备102的视频源104可以包括视频捕获设备,诸如摄像机、包含先前捕获的原始视频的视频档案和/或从视频内容提供商接收视频的视频馈送接口。作为进一步的替代方案,视频源104可以生成基于计算机图形的数据作为源视频,或者实况视频、存档视频和计算机生成视频的组合。在每种情况下,视频编码器200对捕获的、预捕获的或计算机生成的视频数据进行编码。视频编码器200可以将图片从接收顺序(有时称为“显示顺序”)重新排列为用于编解码的编解码顺序。视频编码器200可以生成包括编码的视频数据的比特流。然后,源设备102可以经由输出接口108将编码的视频数据输出到计算机可读介质110上,以供例如目的地设备116的输入接口122进行接收和/或检索。

[0029] 源设备102的存储器106和目的地设备116的存储器120表示通用存储器。在一些示例中,存储器106、120可以存储原始视频数据,例如,来自视频源104的原始视频和来自视频解码器300的原始的解码的视频数据。附加地或可替代地,存储器106、120可以分别存储可由例如视频编码器200和视频解码器300执行的软件指令。尽管在此示例中存储器106和存储器120与视频编码器200和视频解码器300分开示出,但是应当理解的是,视频编码器200和视频解码器300还可以包括实现功能上相似或等效目的的内部存储器。此外,存储器106、120可以存储例如从视频编码器200输出并输入到视频解码器300的编码的视频数据。在一些示例中,存储器106、120的一些部分可以分配为一个或多个视频缓冲区,例如,用来存储原始的解码和/或编码视频数据。

[0030] 计算机可读介质110可以表示能够将编码的视频数据从源设备102传输到目的地设备116的任何类型的介质或设备。在一个示例中,计算机可读介质110表示通信介质以使源设备102能够例如经由射频网络或基于计算机的网络将编码的视频数据实时地直接传输到目的地设备116。根据诸如无线通信协议的通信标准,输出接口108可以对包括编码的视频数据的传输信号进行调制,并且输入接口122可以对接收到的传输信号进行解调。通信介质可以包括任何无线或有线通信介质,诸如射频(RF)频谱或者一条或多条物理传输线。通信介质可以形成基于分组的网络的一部分,诸如局域网、广域网或诸如因特网的全球网络。通信介质可以包括路由器、交换机、基站或有助于从源设备102到目的地设备116的通信的任何其他装备。

[0031] 在一些示例中,源设备102可以将编码数据从输出接口108输出到存储设备112。类似地,目的地设备116可以经由输入接口122访问来自存储设备112的编码的数据。存储设备112可以包括各种分布式或本地访问的数据存储介质中的任何一种,诸如硬盘、蓝光光盘、DVD、CD-ROM、闪存、易失性或非易失性存储器,或者用于存储编码的视频数据的任何其他合适的数字存储介质。

[0032] 在一些示例中,源设备102可以向文件服务器114或可以存储由源设备102生成的编码的视频数据的另一中间存储设备输出编码的视频数据。目的地设备116可以经由流式传输或下载来访问来自文件服务器114的存储的视频数据。

[0033] 文件服务器114可以是能够存储编码的视频数据并将编码的视频数据发送到目的地设备116的任何类型的服务器设备。文件服务器114可以表示web服务器(例如,用于网站)、被配置为提供文件传输协议服务(诸如文件传输协议(FTP)或单向传输文件传递(FLUTE)协议)的服务器、内容分发网络(CDN)设备、超文本传输协议(HTTP)服务器、多媒体广播多播服务(MBMS)或增强型MBMS(eMBMS)服务器、和/或网络连接存储(NAS)设备。附加地

或可替代地,文件服务器114可以实施一个或多个HTTP流式传输协议,诸如基于HTTP的动态自适应流式传输(DASH)、HTTP实时流式传输(HLS)、实时流式传输协议(RTSP)、HTTP动态流式传输等。

[0034] 目的地设备116可以通过包括因特网连接的任何标准数据连接来访问来自文件服务器114的编码的视频数据。这可以包括无线信道(例如,Wi-Fi连接)、有线连接(例如,数字订户线路(DSL)、电缆调制解调器等)或者适合访问存储在文件服务器114上的编码的视频数据的两者的组合。输入接口122可以被配置为根据上面讨论的用于从文件服务器114检索或接收媒体数据的各种协议中的任何一个或多个来操作,或者根据用于检索媒体数据的其他此类协议来操作。

[0035] 输出接口108和输入接口122可以表示无线发送器/接收器、调制解调器、有线联网组件(例如,以太网卡)、根据各种IEEE 802.11标准中的任何一种进行操作的无线通信组件,或者其他物理组件。在输出接口108和输入接口122包括无线组件的示例中,输出接口108和输入接口122可以被配置为根据诸如4G、4G-LTE(长期演进)、LTE高级、5G或类似标准的蜂窝通信标准来传输诸如编码的视频数据的数据。在输出接口108包括无线发送器的某些示例中,输出接口108和输入接口122可以被配置为根据其他无线标准,诸如IEEE 802.11规范、IEEE 802.15规范(例如,ZigBee™)、蓝牙™标准等来传输诸如编码的视频数据的数据。在一些示例中,源设备102和/或目的地设备116可以包括相应的片上系统(SoC)设备。例如,源设备102可以包括SoC设备来执行归于视频编码器200和/或输出接口108的功能,并且目的地设备116可以包括SoC设备来执行归于视频解码器300和/或输入接口122的功能。

[0036] 本公开的技术可以应用于支持各种多媒体应用中的任何一种的视频编解码,多媒体应用诸如空中电视广播、有线电视传输、卫星电视传输、诸如基于HTTP的动态自适应流式传输(DASH)的互联网流式视频传输、编码到数据存储介质上的数字视频、对存储在数据存储介质上的数字视频的解码或者其他应用。

[0037] 目的地设备116的输入接口122从计算机可读介质110(例如,通信介质、存储设备112、文件服务器114等)接收编码的视频比特流。编码的视频比特流可以包括由视频编码器200定义的(其也由视频解码器300使用的)信令信息,诸如语法元素,该语法元素具有描述视频块或其他编解码单元(例如,条带、图片、图片组、序列等)的特性和/或处理的值。显示设备118向用户显示解码视频数据的解码图片。显示设备118可以表示各种显示设备中的任何一种,诸如液晶显示器(LCD)、等离子显示器、有机发光二极管(OLED)显示器或另一类显示设备。

[0038] 尽管未在图1中示出,但在一些示例中,视频编码器200和视频解码器300可以各自与音频编码器和/或音频解码器集成在一起,并且可以包括适当的MUX-DEMUX单元或其他硬件和/或软件,来处理公共数据流中包括音频和视频两者的多路复用流。如果适用,MUX-DEMUX单元可以符合ITU H.223多路复用器协议或诸如用户数据报协议(UDP)的其他协议。

[0039] 视频编码器200和视频解码器300可以各自被实施为各种合适的编码器和/或解码器电路中的任何一种,诸如一个或多个微处理器、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。当该技术部分地以软件来实施时,设备可以将用于软件的指令存储在合适的非暂时计算机可读介质中,并使用一个或多个处理器在硬件中执行该指令来执行本公开的技术。视频编码器200和

视频解码器300中的每一个可以被包括在一个或多个编码器或解码器中,这两者任一个都可以集成为各自设备中组合编码器/解码器(CODEC)的一部分。包括视频编码器200和/或视频解码器300的设备可以包括集成电路、微处理器和/或诸如蜂窝电话的无线通信设备。

[0040] 视频编码器200和视频解码器300可以根据视频编解码标准(诸如ITU-T H.265,也称为高效视频编解码(HEVC))或其扩展(诸如多视图和/或可缩放的视频编解码扩展)来操作。可替代地,视频编码器200和视频解码器300可以根据其他专有或工业标准,诸如多功能视频编解码(VVC)。在以下中描述了VVC标准的草案:Bross等人的“多功能视频编码(草案9)”,ITU-TSG 16WP3和ISO/IEC JTC 1/SC 29/WG 11联合视频专家组(JVET),第18次会议:4月15日至24日,JVET-R2001-v8(以下称为“VVC草案9”)。但是,本公开的技术不限于任何特定的编解码标准。

[0041] 一般地,视频编码器200和视频解码器300可以执行图片的基于块的编解码。术语“块”一般是指包括待处理的(例如,待编码、待解码或以其他方式在编码和/或解码过程中使用的)数据的结构。例如,块可以包括亮度和/或色度数据的二维样点矩阵。一般地,视频编码器200和视频解码器300可以对以YUV(例如,Y,Cb,Cr)格式表示的视频数据进行编解码。也就是说,视频编码器200和视频解码器300可以对亮度和色度分量进行编解码,其中色度分量可以包括红色和蓝色色调分量两者,而不是对图片的样点的红色、绿色和蓝色(RGB)数据进行编解码。在一些示例中,视频编码器200在编码之前将接收到的RGB格式的数据转换成YUV表示,并且视频解码器300将YUV表示转换成RGB格式。可替代地,预处理和后处理单元(未示出)可以执行这些转换。

[0042] 本公开通常涉及图片的编解码(例如,编码和解码),以包括对图片的数据进行编码或解码的过程。类似地,本公开可以涉及对图片的块进行编解码以包括对块的数据进行编码或解码的过程,例如,预测和/或残差编解码。编码的视频比特流一般包括用于表示编解码决策(例如,编解码模式)以及图片分割为块的语法元素的一系列值。因此,对图片或块进行编解码的引用一般应理解为对形成图片或块的语法元素的值进行编解码。

[0043] HEVC定义各种的块,包括编解码单元(CU)、预测单元(PU)和变换单元(TU)。根据HEVC,视频编解码器(诸如视频编码器200)根据二叉树结构将编解码树单元(CTU)分割为CU。也就是说,视频编解码器将CTU和CU分割为四个相等的非重叠正方形,并且二叉树的每一个节点具有零个或四个子代节点。没有子代节点的节点可以称为“叶节点”,并且这种叶节点的CU可以包括一个或多个PU和/或一个或多个TU。视频编解码器可以进一步分割PU和TU。例如,在HEVC中,残差二叉树(RQT)表示TU的分割。在HEVC中,PU表示帧间预测数据,而TU表示残差数据。帧内预测的CU包括帧内预测信息,诸如帧内模式指示。

[0044] 作为另一示例,视频编码器200和视频解码器300可以被配置为根据VVC进行操作。根据VVC,视频编解码器(诸如视频编码器200)将图片分割为多个编解码树单元(CTU)。视频编码器200可以根据诸如二叉树-二叉树(QTBT)结构或多类型树(MTT)结构的树结构来对CTU进行分割。QTBT结构移除了多个分割类型的概念,诸如HEVC的CU、PU和TU之间的分离。QTBT结构包括两个级别:根据二叉树分割而分割的第一级别,以及根据二叉树分割而分割的第二级别。QTBT结构的根节点对应于CTU。二叉树的叶节点对应于编解码单元(CU)。

[0045] 在MTT分割结构中,可以使用二叉树(QT)分割、二叉树(BT)分割以及一种或多种类型的三叉树(TT)(也称为三元树(TT))分割来对块进行分割。三叉树或三元树分割是将块划

分为三个子块的分割。在一些示例中,二叉树或三元树分割将块划分为三个子块,而不通过中心来划分原始块。MTT中的分割类型(例如,QT、BT和TT)可以是对称的或不对称的。

[0046] 在一些示例中,视频编码器200和视频解码器300可以使用单个QTBT或MTT结构来表示亮度和色度分量中的每一个,而在其他示例中,视频编码器200和视频解码器300可以使用两个或更多个QTBT或MTT结构,诸如用于亮度分量的一个QTBT/MTT结构和用于两个色度分量的另一QTBT/MTT结构(或者用于相应色度分量的两个QTBT/MTT结构)。

[0047] 视频编码器200和视频解码器300可以被配置为使用依照HEVC的四叉树分割、QTBT分割、MTT分割或其他分割结构。出于说明的目的,关于QTBT分割呈现了本公开的技术的描述。然而,应该理解的是,本公开的技术还可以应用于被配置使用四叉树分割或其他类型的分割的视频编解码器。

[0048] 在一些示例中,CTU包括亮度样点的编解码树块(CTB)、具有三个样点阵列的图像的色度样点的两个对应CTB,或者单色图像或者使用三个单独的颜色平面和用于对样点进行编解码的语法结构来进行编解码的图像的样点的CTB。对于某个N值,CTB可以是 $N \times N$ 样点块,使得将分量划分成CTB是一种分割。该分量可以是来自4:2:0、4:2:2或4:4:4颜色格式的图像的三个阵列(亮度和两个色度)之一的阵列或单个样点,或者单色格式的图像的阵列或单个样点。在一些示例中,对于M和N的一些值,编码块是 $M \times N$ 样点块,使得将CTB划分成编码块是一种分割。

[0049] 可以在图片中以各种方式将块(例如,CTU或CU)分组。作为一个示例,砖块(brick)可以指图片中特定片(tile)中的CTU行的矩形区域。片可以是图片中的特定片列和特定片行中的CTU的矩形区域。片列指的是具有等于图片高度的高度和由语法元素(例如,诸如在图片参数集中)指定的宽度的CTU的矩形区域。片行指的是具有由语法元素(例如,诸如在图片参数集中)指定的高度和等于图片宽度的宽度的CTU的矩形区域。

[0050] 在一些示例中,可以将片分割为多个砖块,每个砖块可以包括该片中的一个或多个CTU行。没有被分割为多个砖块的片也可以称为砖块。但是,作为片的真子集的砖块不能称为片。

[0051] 图片中的砖块也可以在条带中排列。条带可以是图片的整数数量的砖块,其可以被排他地包含在单个网络抽象层(NAL)单元中。在一些示例中,条带包括多个完整片或仅包括一个片的连续序列的完整砖块。

[0052] 本公开可以互换地使用“ $N \times N$ ”和“ $N$ 乘 $N$ ”来指块(诸如CU或其他视频块)在垂直和水平维度方面的样点维度,例如,16×16个样点或16乘16个样点。一般地,16×16的CU在垂直方向上将有16个样点( $y=16$ ),并且在水平方向上将有16个样点( $x=16$ )。同样地, $N \times N$ 的CU一般地在垂直方向上具有N个样点,并且在水平方向上具有N个样点,其中N表示非负整数值。CU中的样点可以按行和列来排列。此外,CU在水平方向上不必具有与垂直方向上相同数量的样点。例如,CU可以包含 $N \times M$ 个样点,其中M不一定等于N。

[0053] 视频编码器200对表示预测和/或残差信息以及其他信息的CU的视频数据进行编码。预测信息指示将如何预测CU以便形成用于CU的预测块。残差信息一般表示编码前CU的样点与预测块之间的逐样点差异。

[0054] 为了预测CU,视频编码器200一般地可以通过帧间预测或帧内预测来形成CU的预测块。帧间预测一般是指从先前编解码的图像的数据中预测CU,而帧内预测一般是指从同

一图片的先前编解码的数据中预测CU。为了执行帧间预测,视频编码器200可以使用一个或多个运动矢量来生成预测块。视频编码器200通常可以执行运动搜索以识别例如在CU和参考块之间的差异的方面与CU紧密地匹配的参考块。视频编码器200可以使用绝对差之和(SAD)、平方差之和(SSD)、平均绝对差(MAD)、均方差(MSD)或其他此类差值计算来计算差值度量以确定参考块是否与当前CU紧密地匹配。在某些示例中,视频编码器200可以使用单向预测或双向预测来预测当前CU。

[0055] VVC的一些示例还提供仿射运动补偿模式,其可以视为帧间预测模式。在仿射运动补偿模式中,视频编码器200可以确定表示非平移运动的两个或更多个运动矢量,诸如放大或缩小、旋转、透视运动或其他不规则运动类型。

[0056] 为了执行帧内预测,视频编码器200可以选择帧内预测模式来生成预测块。VVC的一些示例提供了六十七种帧内预测模式,包括各种方向模式以及平面模式和DC模式。一般地,视频编码器200选择帧内预测模式,其描述当前块(例如,CU的块)的相邻样点,从该相邻样点预测当前块的预测样点。假设视频编码器200以光栅扫描顺序(从左到右、从上到下)对CTU和CU进行编解码,则此类样点通常可以在与当前块相同的图片中在当前块的上方、上左侧或左侧。

[0057] 视频编码器200对表示当前块的预测模式的数据进行编码。例如,对于帧间预测模式,视频编码器200可以对数据进行编码,该数据表示使用各种可用帧间预测模式中的哪一种以及对应模式的运动信息。对于单向或双向帧间预测,例如,视频编码器200可以使用高级运动矢量预测(AMVP)或merge模式来对运动矢量进行编码。视频编码器200可以使用类似模式来对仿射运动补偿模式的运动矢量进行编码。

[0058] 在预测(诸如块的帧内预测或帧间预测)之后,视频编码器200可以计算块的残差数据。残差数据(诸如残差块)表示块与该块的预测块之间的逐样点差异,预测块是使用对应预测模式形成。视频编码器200可以将一个或多个变换应用于残差块以在变换域而非样点域中产生变换数据。例如,视频编码器200可以将离散余弦变换(DCT)、整数变换、小波变换或概念上类似的变换应用于残差视频数据。另外,视频编码器200可以在一次变换之后应用二次变换,诸如模式依赖的不可分二次变换(MDNSSST)、信号依赖的变换、Karhunen-Loeve变换(KLT)等。视频编码器200在应用一个或多个变换之后产生变换系数。

[0059] 如上所述,在进行任何变换以产生变换系数之后,视频编码器200可以对变换系数执行量化。量化通常是指对变换系数进行量化来可能地减少用于表示变换系数的数据量,从而提供进一步的压缩的过程。通过执行量化过程,视频编码器200可以减小与系数中的一些或所有变换系数相关联的比特深度。例如,视频编码器200可以在量化期间将n比特值向下舍入到m比特值,其中n大于m。在一些示例中,为了执行量化,视频编码器200可以执行要量化的值的按比特右移。

[0060] 量化之后,视频编码器200可以扫描变换系数,从而从包括量化变换系数的二维矩阵中产生一维矢量。可以将扫描设计为将较高能量(因此频率较低)的系数放在矢量的前面,并将较低能量(因此频率较高)的变换系数放在矢量的后面。在一些示例中,视频编码器200可以利用预定义的扫描顺序来对量化变换系数进行扫描以产生序列化的矢量,然后对矢量的量化变换系数进行熵编码。在其他示例中,视频编码器200可以执行自适应扫描。在对量化变换系数进行扫描以形成一维矢量之后,视频编码器200可以例如根据上下文自适

应二进制算术编解码 (CABAC) 来对一维矢量进行熵编码。视频编码器200还可以对语法元素的值进行熵编码,语法元素描述与视频解码器300在解码视频数据中使用的编码的视频数据相关联的元数据。

[0061] 为了执行CABAC,视频编码器200可以将上下文模型内的上下文分配给待传输的符号。例如,上下文可以涉及符号的相邻值是否是零值。概率确定可以基于分配给符号的上下文。

[0062] 视频编码器200可以进一步地生成例如在图片标头、块标头、条带标头中到视频解码器300的语法数据,诸如基于块的语法数据、基于图片的语法数据以及基于序列的语法数据,或者生成其他语法数据,诸如序列参数集 (SPS)、图片参数集 (PPS) 或视频参数集 (VPS)。视频解码器300可以类似地对此类语法数据进行解码以确定如何解码对应的视频数据。

[0063] 以此方式,视频编码器200可以生成包括编码的视频数据的比特流,例如,描述将图片分割成块 (例如, CU) 以及块的预测和/或残差信息的语法元素。最终,视频解码器300可以接收比特流并且解码编码的视频数据。

[0064] 一般地,视频解码器300执行由视频编码器200执行的相反过程以对比特流的编码的视频数据进行解码。例如,视频解码器300可以使用CABAC,以与视频编码器200的CABAC编码过程基本相似 (尽管与之相反) 的方式,对比特流的语法元素的值进行解码。语法元素可以定义分割信息,该分割信息用于将图片分割为CTU和根据比如诸如QTBT结构的对应分割结构对每一个CTU进行分割以定义CTU的CU。语法元素可以进一步定义视频数据的块 (例如, CU) 的预测和残差信息。

[0065] 例如,可以由量化变换系数表示残差信息。视频解码器300可以对块的量化变换系数进行逆量化和逆变换以再现该块的残差块。视频解码器300使用信令通知的预测模式 (帧内或帧间预测) 和相关的预测信息 (例如,用于帧间预测的运动信息) 来形成块的预测块。然后,视频解码器300可以 (在逐样点的基础上) 组合预测块和残差块以再现原始块。视频解码器300可以执行附加处理 (诸如执行去方块过程) 来减少沿块边界的视觉伪影。

[0066] 一般地,本公开可以涉及“信令通知”某些信息,诸如语法元素。术语“信令通知”通常可以指对于语法元素和/或用于对编码的视频数据进行解码的其他数据的值的通信。也就是说,视频编码器200可以信令通知比特流中的语法元素的值。一般地,信令通知是指在比特流中生成值。如上所述,源设备102可以基本上实时地 (或非实时地,诸如可能在将语法元素存储到存储设备112以供稍后由目的地设备116检索时发生) 将比特流传送到目的地设备116。

[0067] 图2A和图2B是示出示例四叉树二叉树 (QTBT) 结构130和对应编解码树单元 (CTU) 132的概念图。实线代表四叉树划分,并且虚线指示二叉树划分。在二叉树的每一个划分 (即,非叶) 节点中,信令通知一个标志以指示使用了哪种划分类型 (即,水平或垂直),其中在此示例中,0指示水平划分,并且1指示垂直划分。对于四叉树划分,由于四叉树节点将块水平地和垂直地划分为具有相等尺寸的4个子块,因此无需指示划分类型。相应地,视频编码器200可以编码QTBT结构130的区域树级别 (即,实线) 的语法元素 (例如,划分信息) 和QTBT结构130的预测树级别 (即,虚线) 的语法元素 (例如,划分信息),并且视频解码器300可以对以上进行解码。对于QTBT结构130的终端叶节点代表的CU,视频编码器200可以编码视频数据 (诸如预测和变换数据),并且视频解码器300可以对以上进行解码。

[0068] 一般地,图2B的CTU 132可以与参数相关联,这些参数定义与第一和第二级别的QTBT结构130的节点相对应的块的尺寸。这些参数可以包括CTU尺寸(代表样点中CTU 132的尺寸)、最小四叉树尺寸(MinQTSize,代表最小允许的四叉树叶节点尺寸)、最大二叉树尺寸(MaxBTSIZE,代表最大允许的二叉树根节点尺寸)、最大二叉树深度(MaxBTDepth,代表最大允许的二叉树深度)和最小二叉树尺寸(MinBTSIZE,代表最小允许的二叉树叶节点尺寸)。

[0069] 与CTU相对应的QTBT结构的根节点在QTBT结构的第一级别可以具有四个子代节点,每个子代节点可以根据四叉树分割来进行分割。也就是说,第一级别的节点是叶节点(没有子代节点)或具有四个子代节点。QTBT结构130的示例代表这样的节点,其包括具有实线分支的子代节点和父节点。如果第一级别的节点不大于最大允许的二叉树根节点尺寸(MaxBTSIZE),则能够通过相应的二叉树进一步对节点进行分割。能够迭代一个节点的二叉树划分,直到划分生成的节点达到最小允许的二叉树叶节点尺寸(MinBTSIZE)或最大允许的二叉树深度(MaxBTDepth)。QTBT结构130的示例将这样的节点代表为具有虚线分支。二叉树叶节点表示为编码单元(CU),其用于预测(例如,图片内或图片间预测)和变换,而无需任何进一步分割。如上所述,CU也可以表示为“视频块”或“块”。

[0070] 在QTBT分割结构的一个示例中,CTU尺寸设置为 $128 \times 128$ (亮度样点和两个对应的 $64 \times 64$ 色度样点),MinQTSIZE设置为 $16 \times 16$ ,MaxBTSIZE设置为 $64 \times 64$ ,MinBTSIZE(用于宽度和高度)设置为4,并且MaxBTDepth设置为4。首先,将四叉树分割应用于CTU来生成四叉树叶节点。四叉树叶节点可以具有从 $16 \times 16$ (即,MinQTSIZE)到 $128 \times 128$ (即,CTU尺寸)的尺寸。如果四叉树叶节点为 $128 \times 128$ ,则由于其尺寸超过MaxBTSIZE(在此示例中为 $64 \times 64$ ),因此该四叉树叶节点不会进一步被二叉树划分。否则,四叉树叶节点可以进一步进行二叉树分割。因此,四叉树叶节点也是二叉树的根节点,并且具有为0的二叉树深度。当二叉树深度达到MaxBTDepth(在此示例中为4)时,不允许进一步划分。宽度等于MinBTSIZE(在此示例中为4)的二叉树节点意味着该二叉树节点不允许进一步的垂直划分(即宽度的划分)。类似地,具有等于MinBTSIZE的高度的二叉树节点意味着该二叉树节点不允许进一步的水平划分(即,高度的划分)。如上所述,二叉树的叶节点称为CU并且根据预测和变换对其进行进一步处理而无需进一步分割。

[0071] 图3是示出混合视频编解码框架140的概念图。自H.261以来的视频编解码标准已经基于所谓的混合视频编解码原理,其在图3中示出。术语混合指的是用于减少视频信号中的冗余的两种手段(means)的组合,即,预测142和利用预测残差的量化144进行变换编解码。而预测和变换通过去相关来减少视频信号中的冗余,量化通过降低它们的精度,理想地通过仅去除不相关的细节来减少变换系数表示的数据。这种混合视频编解码设计原则也被用于最近的两个标准,ITU-T H.265/HEVC和ITU-T H.266/VVC。如图3所示,现代混合视频编解码器包括块分割、包括运动补偿或图片间预测和图片内预测的预测142、包括变换和量化的变换/量化144、熵编解码146和后/环内滤波148。

[0072] 块分割被用于将图片划分成更小的块,用于预测和变换过程的操作。早期的视频编解码标准使用固定的块尺寸,通常为 $16 \times 16$ 样点。最近的标准(诸如HEVC和VVC)采用基于树的分割结构来提供灵活的分割。

[0073] 运动补偿或图片间预测利用了视频序列的图片之间(因此是“帧间”)存在的冗余。根据在所有现代视频编解码器中使用的基于块的运动补偿,从一个或多个先前解码的图片

(即参考图片)获得预测。生成帧间预测的对应区域由包括运动矢量和参考图像索引的运动信息指示。

[0074] 图4是示出使用图片组(GOP)尺寸为16的分层预测结构150的概念图。在最近的视频编解码器中,应用图片组(GOP)内的分层预测结构来提高编解码效率。

[0075] 再次参考图3,图片内预测通过从已经编码/解码的空间相邻(参考)样点导出对块的预测,利用存在于图片内(因此是“帧内”)的空间冗余。方向角预测、DC预测和平面或平面预测被用于最近的视频编解码器,包括AVC、HEVC和VVC。

[0076] 变换:混合视频编解码标准将块变换应用于预测残差(不管它是来自图片间预测还是图片内预测)。在早期的标准中(包括H.261、H.262和H.263),采用离散余弦变换(DCT)。在HEVC和VVC中,除了DCT之外,还应用了更多的变换内核,以解决特定视频信号中的不同统计问题。

[0077] 量化旨在降低输入值或输入值集合的精度,以减少表示这些值所需的数据量。在混合视频编解码中,量化通常应用于单个变换后的残差样点(即,变换系数),从而得到整数系数级别。在最近的视频编解码标准中,步长是从控制保真度和比特率的所谓量化参数(QP)导出的。较大的步长降低了比特率,但也恶化了质量,这例如导致视频图片表现出阻塞伪像和模糊的细节。

[0078] 上下文自适应二进制算术编解码(CABAC)是一种熵编解码形式,由于其高效率而被用于最近的视频编解码器,例如,AVC、HEVC和VVC。

[0079] 环后/环内滤波是应用于重建图片以减少编码伪像的滤波过程(或这种过程的组合)。滤波过程的输入通常是重建的图片,它是重建的残差信号(其包括量化误差)和预测的组合。如图3所示,在环内滤波之后的重建图片被存储并用作后续图片的图片间预测的参考。编解码伪像主要由QP决定,因此QP信息通常用于滤波过程的设计。在HEVC中,环路滤波器包括去块滤波和采样自适应偏移(SAO)滤波。在VVC标准中,引入了自适应环路滤波器(ALF)作为第三滤波器。ALF的滤波过程如下示出:

$$[0080] \quad R'(i, j) = R(i, j) + \left( \left( \sum_{k \neq 0} \sum_{l \neq 0} f(k, l) \times K(R(i+k, j+l) - R(i, j), c(k, l)) + 64 \right) \gg 7 \right)$$

[0081] 其中 $R(i, j)$ 是滤波过程前的样点集, $R'(i, j)$ 是滤波过程后的样点值。 $f(k, l)$ 表示滤波器系数, $K(x, y)$ 是裁剪函数并且 $c(k, l)$ 表示裁剪参数。变量 $k$ 和 $l$ 在 $-\frac{L}{2}$ 与 $\frac{L}{2}$ 之间变化,其中 $L$ 表示滤波器长度。裁剪函数 $K(x, y) = \min(y, \max(-y, x))$ ,其对应于函数 $\text{Clip3}(-y, y, x)$ 。裁剪操作引入非线性,以通过减少与当前样点值相差太大的相邻样点值的影响,使ALF更有效。在VVC中,滤波参数可以在比特流中被信令通知,它可以从预定义的滤波器集合中选择。ALF滤波过程还可以概括为以下等式:

$$[0082] \quad R'(i, j) = R(i, j) + \text{ALF\_residual\_output}(R)$$

[0083] 图5是示出具有四层的基于神经网络的滤波器160的概念图。各种研究表明,在混合视频编解码框架中嵌入神经网络(NN)可以提高压缩效率。为了提高预测效率,已经在帧内预测和帧间预测模块中引入了神经网络。基于NN的环路滤波也是近年来一个突出的研究课题。在某些情况下,基于NN的滤波过程被应用为环后滤波。在这种情况下,滤波过程仅应用于输出图片,并且未滤波的图片被用作参考图片。

[0084] 除了现有滤波器之外,还可以应用基于NN的滤波器,诸如去块滤波器、样本自适应偏移(SAO)和/或自适应环路滤波(ALF)。基于NN的滤波器也可以排他地应用,其中基于NN的滤波器被设计成替换所有现有的滤波器。附加地或可替代地,基于NN的滤波器可以被设计成补充、增强或替换任何或所有其他滤波器。

[0085] 如图5所示,基于NN的滤波处理可以将重建的样点作为输入,并且中间输出是残差样点,其被添加回输入以细化输入样点。NN滤波器可以使用所有颜色分量(例如,Y、U和V,或Y、Cb和Cr,即,亮度、蓝色色调度和红色色调度)作为输入,以利用交叉分量相关性。不同的颜色分量可以共享相同的滤波器(包括网络结构和模型参数),或者每个颜色分量可以具有其自己的特定滤波器。

[0086] 滤波过程也可以概括如下: $R'(i, j) = R(i, j) + \text{NN\_filter\_residual\_output}(R)$ 。(多个)基于NN的滤波器的模型结构和模型参数可以预定义并且存储在编码器和解码器中。滤波器也可以在比特流中信令通知。

[0087] 本公开认识到,在一些情况下,基于视频和图像数据库大集合来训练预定义滤波器(例如,基于NN的滤波器或ALF)。滤波器通常可以是最优的,但是对于特定的失真序列可能不是最优的。本公开还认识到,由于预定义的训练图像/视频数据库可能不构成所有可能类型的视频特性,因此将在该给定数据库上训练的滤波器应用于具有不同视频特性的序列可能不会带来任何客观或主观益处,而是可能损害客观或主观质量。

[0088] 根据本公开的技术,图1的视频编码器200和视频解码器300可以被配置为单独或以任何组合来执行以下技术中的任何或所有技术。通常,可以使用多个滤波模型。对于输入图片的每个目标区域,视频编码器200和/或视频解码器300可以选择多个滤波模型中的一个或多个来执行滤波。视频编码器200和视频解码器300可以基于比特流中的信息(例如,基于量化参数(QP))隐式地导出NN模型的选择。可替代地,视频解码器300可以使用由视频编码器200在比特流中显式信令通知的索引来导出NN模型的选择。在一些示例中,显式信令和隐式导出可以组合。

[0089] 在一些示例中,视频编码器200可以在比特流中为所选择的模型选择一个或多个模型和信号索引。当选择一个模型时,该模型被用于对输入图片的对应目标区域进行滤波。当选择多个模型时,联合使用这些模型来对输入图片的目标区域进行滤波。作为示例,视频编码器200和视频解码器300可以将模型分别应用于目标区域,并且所涉及的模型的输出可以被组合为目标区域的最终输出。

[0090] 关闭(switching off)滤波过程可以被用作候选。当对目标区域选择“关闭”时,视频编码器200和视频解码器300不应用任何滤波器,并且输出信号与输入信号相同。

[0091] 选择和信令通知(多个)模型的粒度可以在不同的级别上被设计。信令通知滤波器模型索引的可能级别包括视频参数集/序列参数集/图片参数集(VPS/SPS/PPS)级别、帧内时段(intra period)级别、图片组(GOP)级别、GOP中的时间层级别、图片级别、条带级别、CTU级别、或专门为滤波器信令而设计的网格尺寸 $N \times N$ 。滤波器模型信令的级别的选择可以是固定的;视频编码器200可以将该选择作为语法元素在比特流中向视频解码器300信令通知;或者视频解码器300可以基于比特流中的信息(例如,图片分辨率、QP等)隐式地导出选择。

[0092] 在一个示例实施方式中,存在N个预定义的滤波器模型。在滤波器模型信令的级

别,视频编码器200选择一个模型并信令通知视频解码器300的对应索引,以确定要使用哪个滤波器模型。信令通知该元素的可能级别包括VPS/SPS/PPS级别、帧内时段级别、图片组(GOP)级别、图片级别、条带级别、CTU级别、或专门为滤波器信令设计的网格尺寸 $N \times N$ 。滤波器模型信令的级别的选择可以是固定的、作为语法元素在比特流中被信令通知的、或者基于比特流中的信息(例如,图片分辨率、QP等)隐式导出的。

[0093] 作为另一示例,视频编码器200可以导出预定义模型的子集,并在比特流中将该子集作为序列参数集(SPS)、条带头、图片头、自适应参数集(APS)中的语法元素或任何其他高级语法元素主体信令通知给视频解码器300。子集的尺寸在这里被称为M。M可以是在较低级别(例如,条带头、图片头、CTU级别、网格级别等)预定义或在比特流中信令通知的任何值。如果 $M > 1$ ,则选择M个候选中的一个,并在比特流中作为语法元素被信令通知。

[0094] 作为另一示例,每个滤波器模型可以与QP值相关联,并且对于每个图片,视频编码器200和视频解码器300可以导出“模型选择QP”,并且选择具有最接近当前图片的“模型选择QP”的QP值的模型。在这种情况下,不需要信令通知附加的信息。用于导出“模型选择QP”的信息可以包括:(多个)当前图片QP、(多个)参考图片QP、同一GOP内的图片的(多个)QP、同一帧内帧内时段的图片的(多个)QP等。可替代地,可以将条带中的一个或多个块的QP或块级QP提供给NN模型以对当前条带或当前块进行滤波。

[0095] 作为另一示例,类似于上述示例,每个滤波器模型可以与QP值相关联,并且视频编码器200和视频解码器300可以为每个帧导出“模型选择QP”。视频编码器200和视频解码器300可以基于“模型选择QP”导出所有模型的子集。视频编码器200可以确定具有最接近“模型选择QP”的模型QP的模型,选择这些模型中的一个,并信令通知视频解码器300的附加索引,以从QP导出的子集中选择模型。

[0096] 如上所述,在上面讨论的各种示例的任何或所有示例中,关闭基于NN的建模可以用作候选选择。关闭可以被认为滤波器模型集的附加规则候选,并且“关闭”情况的信令可以与其他滤波器模型的信令统一。作为另一示例,“关闭”可以被认为是一个特殊的候选,并且信令与其他滤波器模型是分开的。

[0097] 视频编码器200和视频解码器300可以被配置为应用用于如上讨论的基于多模型的滤波的开/关控制。也就是说,视频编码器200和视频解码器300将仅当给定控制操作(例如,标志)断言应当应用滤波时才对残差进行滤波并将结果添加到输入样点。该控制信号可以表示为:

[0098] 如果(ApplyFilter为真)  $\{R'(i, j) = R(i, j) + (\text{filter\_residual\_ouput}(R))\}$ ,

[0099] 否则  $\{R'(i, j) = R(i, j)\}$

[0100] 其中ApplyFilter是可以由视频编码器200确定的控制操作,视频编码器200可以在比特流中将表示决策的数据(例如,标志或其他语法元素)信令通知给视频解码器300。

[0101] 作为示例,视频编码器200可以计算应用滤波器的率失真(RD)成本,并将其与不应用滤波器的RD成本进行比较。基于该比较的结果,视频编码器200可以确定控制操作(例如,将给定标志的值设置为0或1状态,其中一个状态表示要应用的滤波器,而另一状态表示不应用滤波器)。视频编码器200然后可以在比特流中信令通知给定标志。视频解码器300可以解析给定标志,并且基于其值,应用滤波器,或者不应用滤波器。由于训练的滤波器对于整个视频序列或对于序列中的给定帧或序列中的给定区域可能不是最佳的,所以控制操作

(标志)被信令通知的粒度也变得重要。

[0102] 在一些示例中,视频编码器200可以导出滤波器开/关的值,并在比特流中将该值作为序列参数集(SPS)、条带头、图片参数集(PPS)、图片头、自适应参数集(APS)或任何其他高级语法元素体中的语法元素来信令通知。视频解码器300可以使用该语法元素的值来确定是否使用滤波器。

[0103] 在一些示例中,视频编码器200可以导出滤波器开/关标志的值,并在比特流中将该值作为块级别(例如,CTU级别)的语法元素来信令通知。视频解码器300可以使用该语法元素的值来确定是否使用滤波器。

[0104] 在一些示例中,视频编码器200可以确定对应网格处的网格尺寸(用于信令)和滤波器开/关标志的值,并在比特流中将每个网格的值作为相应的语法元素来信令通知。视频解码器300可以使用这些语法元素的值来确定是否对网格的每个元素以及网格本身使用滤波器。

[0105] 代替使用固定的网格尺寸用于信令(例如,其中网格尺寸总是固定为CTU尺寸),视频编码器200可以基于RD成本为给定条带选择性地选择不同的网格尺寸。例如,对于给定的条带,视频编码器200可以应用不同的网格尺寸(例如, $M \times N$ 的尺寸,其中M和N可以取范围[4,8,16,...,帧尺寸]中的值)。视频编码器200可以计算滤波器开启和关闭情况下的对应RD成本。基于最佳RD成本,视频编码器200可以选择网格尺寸和对应的滤波器开/关标志并且在比特流中信令通知表示网格尺寸和滤波器开/关标志的数据。可以在视频编码器200和视频解码器300处预定义网格尺寸,例如,如下面的表1所示:

[0106] 表1

	索引	用于信令的网格尺寸( $M \times N$ )
[0107]	0	$8 \times 8$
	1	$16 \times 16$
[0108]	2	$32 \times 32$
	3	$64 \times 64$
	4	$128 \times 128$
	等等...	

[0109] 视频编码器200可以使用一元码、二进制码、截断二进制码或可变长度码来在比特流中信令通知网格的索引值。

[0110] 视频解码器300可以基于输入序列的分辨率来确定用于信令的网格尺寸。例如,对于较低分辨率的序列,使用较细粒度的网格(比如 $8 \times 8$ ),而对于较高分辨率的序列,使用较粗粒度的网格(比如 $128 \times 128$ )。

[0111] 在一些示例中,视频编码器200可以针对每个颜色分量分别确定滤波器开/关标志。因此,亮度(Y)、色度(Cb)和色度(Cr)具有它们自己的单独控制标志。

[0112] 在一些示例中,滤波器开/关标志仅针对单个分量被信令通知。其他分量可以共享相同的标志。例如,视频编码器200可以仅针对亮度分量信令通知控制标志,然后Cb和Cr分

量可以共享相同的标志。在其他示例中，视频编码器200可以信令通知亮度分量的控制标志和色度分量的联合标志。

[0113] 视频编码器200和视频解码器300可以应用单独的CABAC上下文来信令通知针对单独分量的标志(例如，Y、Cb和Cr可以具有单独的上下文)。而且，视频编码器200和视频解码器300可以使用空间/时间相邻块标志开/关信息来确定用于信令通知针对当前块的标志的上下文。相邻块可以直接在当前块的上方和/或左侧，或在一定数量的块内。

[0114] 用于计算用于确定开/关标志的RD成本的失真度量可以是SAD(绝对差值之和)或SATD(绝对变换差值之和)或任何其他失真度量。

[0115] 图6是示出可以执行本公开的技术的示例视频编码器200的框图。提供图6是为了解释的目的并且不应认为是对本公开中广泛示例和描述的技术的约束。出于说明的目的，本公开在诸如开发中的ITU-T H.265/HEVC视频编解码标准和VVC视频编解码标准的视频编解码标准的上下文中描述了视频编码器200。然而，本公开的技术不限于这些视频编解码标准并且通常可应用于其他视频编码和解码标准。

[0116] 在图6的示例中，视频编码器200包括视频数据存储器230、模式选择单元202、残差生成单元204、变换处理单元206、量化单元208、逆量化单元210、逆变换处理单元212、重建单元214、滤波器单元216、解码图片缓冲器(DPB) 218以及熵编码单元220。视频数据存储器230、模式选择单元202、残差生成单元204、变换处理单元206、量化单元208、逆量化单元210、逆变换处理单元212、重建单元214、滤波器单元216、DPB 218和熵编码单元220中的任何一个或全部可以在一个或多个处理器中或在处理电路中实施。例如，视频编码器200的单元可以被实施为一个或多个电路或逻辑元件，作为硬件电路的一部分，或者作为处理器、ASIC或FPGA的一部分。此外，视频编码器200可以包括附加的或替代的处理器或处理电路，以执行这些和其他功能。

[0117] 视频数据存储器230可以存储将由视频编码器200的组件进行编码的视频数据。视频编码器200可以从例如视频源104(图1)接收存储在视频数据存储器230中的视频数据。DPB 218可以用作参考图片存储器，参考图片存储器存储参考视频数据，供视频编码器200用于预测后续视频数据时使用。视频数据存储器230和DPB 218可以由多种存储设备中的任何一种形成，诸如动态随机存取存储器(DRAM)，包括同步DRAM(SDRAM)、磁阻RAM(MRAM)、阻变RAM(RRAM)或其他类型的存储设备。视频数据存储器230和DPB 218可以由相同的存储设备或单独的存储器设备提供。在各种示例中，视频数据存储器230可以与视频编码器200的其他组件一起置于片上，如图所示，或者相对于那些组件置于片外。

[0118] 在本公开中，对视频数据存储器230的引用不应解释为限于视频编码器200内部的存储器(除非特别说明如此)或者视频编码器200外部的存储器(除非特别说明如此)。当然，对视频数据存储器230的引用应理解为存储视频编码器200接收用于编码的视频数据(例如待编码的当前块的视频数据)的参考存储器。图1的存储器106还可以对来自视频编码器200各种单元的输出提供临时存储。

[0119] 示出的图6的各个单元用来帮助理解由视频编码器200执行的操作。这些单元可以被实施为固定功能电路、可编程电路或其组合。固定功能电路是指提供特定功能并预设能够在能够执行的操作上的电路。可编程电路是指可以被编程以执行各种任务并且在可以执行的操作中提供灵活功能的电路。例如，可编程电路可以执行软件或固件，软件或固件使可编程

电路以由软件或固件的指令所定义的方式操作。固定功能电路可以执行软件指令(例如,以接收参数或输出参数),但是固定功能电路执行的操作类型通常是不可变的。在一些示例中,一个或多个单元可以是不同的电路块(固定功能或可编程),并且在一些示例中,一个或多个单元可以是集成电路。

[0120] 视频编码器200可以包括由可编程电路形成的算术逻辑单元(ALU)、基本功能单元(EFU)、数字电路、模拟电路和/或可编程内核。在使用由可编程电路执行的软件来执行视频编码器200的操作的示例中,存储器106(图1)可以存储视频编码器200接收和执行的软件的指令(例如,目标代码),或视频编码器200内另一存储器(未示出)可以存储这样的指令。

[0121] 视频数据存储器230被配置为存储接收的视频数据。视频编码器200可以从视频数据存储器230中检索视频数据的图片,并将视频数据提供给残差生成单元204和模式选择单元202。视频数据存储器230中的视频数据可以是待编码的原始视频数据。

[0122] 模式选择单元202包括运动估计单元222、运动补偿单元224和帧内预测单元226。模式选择单元202可以包括附加功能单元,以根据其他预测模式来执行视频预测。作为示例,模式选择单元202可以包括调色板单元、帧内块复制单元(其可以是运动估计单元222和/或运动补偿单元224的一部分)、仿射单元、线性模型(LM)单元等。

[0123] 模式选择单元202通常协调多个编码遍次(pass)来测试编码参数的组合以及得出这种组合的率失真值。编码参数可以包括CTU分割为CU、CU的预测模式、CU的残差数据的变换类型、CU的残差数据的量化参数等。模式选择单元202可以最终选择编码参数的组合,其具有比其他测试的组合更佳的率失真值。

[0124] 视频编码器200可以将从视频数据存储器230中检索的图片分割为一系列CTU,并将一个或多个CTU封装在条带内。模式选择单元202可以根据树结构(诸如上述HEVC的QTBT结构或二叉树结构)来对图片的CTU进行分割。如上所述,视频编码器200可以根据树结构通过对CTU进行分割来形成一个或多个CU。这样的CU通常也可以称为“视频块”或“块”。

[0125] 一般地,模式选择单元202还控制其组件(例如,运动估计单元222、运动补偿单元224和帧内预测单元226)以生成当前块(例如,当前CU或HEVC中PU和TU的重叠部分)的预测块。为了对当前块的帧间预测,运动估计单元222可以执行运动搜索来识别在一个或多个参考图片(例如,存储在DPB 218中的一个或多个先前编解码的图片)中的一个或多个紧密匹配的参考块。具体地,运动估计单元222可以根据例如绝对差之和(SAD)、差值平方和(SSD)、平均绝对差(MAD)、均方差(MSD)等来计算表示潜在参考块与当前块有多相似的值。运动估计单元222通常可以使用当前块与考虑中的参考块之间的逐样点差异来执行这些计算。运动估计单元222可以识别具有这些计算生成的最低值的参考块,这指示与当前块最紧密匹配的参考块。

[0126] 运动估计单元222可以形成一个或多个运动矢量(MV),运动矢量定义参考图片中的参考块相对于当前图片中的当前块的位置。然后,运动估计单元222可以将运动矢量提供给运动补偿单元224。例如,对于单向帧间预测,运动估计单元222可以提供单个运动矢量,而对于双向帧间预测,运动估计单元222可以提供两个运动矢量。然后,运动补偿单元224可以使用运动矢量来生成预测块。例如,运动补偿单元224可以使用运动矢量来检索参考块的数据。作为另一示例,如果运动矢量具有分数样点精度,则运动补偿单元224可以根据一个或多个插值滤波器来对预测块进行插值。此外,对于双向帧间预测,运动补偿单元224可以

检索由各自运动矢量标识的两个参考块的数据并且(例如,通过逐样点平均或加权平均)组合所检索的数据。

[0127] 作为另一示例,对于帧内预测或帧内预测编解码,帧内预测单元226可以根据与当前块相邻的样点来生成预测块。例如,对于方向模式,帧内预测单元226通常可以数学地组合相邻样点的值,并且在当前块上沿定义的方向填充这些计算值来产生预测块。作为另一示例,对于DC模式,帧内预测单元226可以计算当前块的相邻样点的平均值,并且生成预测块以包括用于预测块的每一个样点的所得到的平均值。

[0128] 模式选择单元202将预测块提供给残差生成单元204。残差生成单元204从视频数据存储单元230接收当前块的原始的、未编码的版本,并从模式选择单元202接收预测块。残差生成单元204计算当前块和预测块之间的逐样点差异。得到的逐样点差异定义当前块的残差块。在一些示例中,残差生成单元204还可以使用残差差分脉冲编解码调制(RDPCM)来确定残差块中样点值之间的差异以生成残差块。在某些示例中,可以使用执行二进制减法的一个或多个减法器电路来形成残差生成单元204。

[0129] 在模式选择单元202将CU分割为PU的示例中,每个PU可以与亮度预测单元和对应的色度预测单元相关联。视频编码器200和视频解码器300可以支持具有各种尺寸的PU。如上所述,CU的尺寸可以指CU的亮度编解码块的尺寸,并且PU的尺寸可以指PU的亮度预测单元的尺寸。假设特定CU的尺寸为 $2N \times 2N$ ,则视频编码器200可以支持 $2N \times 2N$ 或 $N \times N$ 的PU尺寸用于帧内预测,以及 $2N \times 2N$ 、 $2N \times N$ 、 $N \times 2N$ 、 $N \times N$ 或类似的对称PU尺寸用于帧间预测。视频编码器200和视频解码器300还可以支持 $2N \times nU$ 、 $2N \times nD$ 、 $nL \times 2N$ 和 $nR \times 2N$ 的PU尺寸的非对称分割用于帧间预测。

[0130] 在模式选择单元202不将CU进一步分割为PU的示例中,每一个CU可以与亮度编解码块和对应的色度编解码块相关联。如上所述,CU的尺寸可以指CU的亮度编解码块的尺寸。视频编码器200和视频解码器300可以支持 $2N \times 2N$ 、 $2N \times N$ 或 $N \times 2N$ 的CU尺寸。

[0131] 对于其他视频编解码技术,诸如作为一些示例的帧内块复制模式编解码、仿射模式编解码和线性模型(LM)模式编解码,模式选择单元202经由与编解码技术相关联的各自单元来生成正在被编码的当前块的预测块。在一些示例中,诸如调色板模式编解码,模式选择单元202可以不生成预测块,而是可以生成语法元素,该语法元素指示基于所选择的调色板重建块的方式。在这样的模式中,模式选择单元202可以将这些语法元素提供给熵编码单元220,以对其进行编码。

[0132] 如上所述,残差生成单元204接收当前块和对应的预测块的视频数据。然后,残差生成单元204生成当前块的残差块。为了生成残差块,残差生成单元204计算预测块和当前块之间的逐样点差异。

[0133] 变换处理单元206将一个或多个变换应用于残差块以生成变换系数的块(在本文中称为“变换系数块”)。变换处理单元206可以将各种变换应用于残差块以形成变换系数块。例如,变换处理单元206可以将离散余弦变换(DCT)、方向变换、Karhunen-Loeve变换(KLT)或概念上类似的变换应用于残差块。在一些示例中,变换处理单元206可以对残差块执行多次变换,例如,初次变换和诸如旋转变换的二次变换。在一些示例中,变换处理单元206不将变换应用于残差块。

[0134] 量化单元208可以对变换系数块中的变换系数进行量化以产生经量化的变换系数

块。量化单元208可以根据与当前块相关联的量化参数(QP)值来对变换系数块的变换系数进行量化。视频编码器200(例如,经由模式选择单元202)可以通过调整与CU相关联的QP值来调整应用于与当前块相关联的变换系数块的量化程度。量化可能会引入信息损失,并且因此,经量化的变换系数可能比变换处理单元206产生的原始变换系数具有更低的精度。

[0135] 逆量化单元210和逆变换处理单元212可以分别对经量化的变换系数块应用逆量化和逆变换,以从变换系数块重建残差块。重建单元214可以基于重建的残差块和由模式选择单元202生成的预测块来产生对应于当前块的重建块(尽管潜在地具有一些程度的失真)。例如,重建单元214可以将重建的残差块的样点添加到来自模式选择单元202生成的预测块的对应样点以产生重建块。

[0136] 滤波器单元216可以对重建块执行一个或多个滤波器操作。例如,滤波器单元216可以执行去方块操作来减少沿着CU的边缘的块效应伪影。在一些示例中,可以跳过滤波器单元216的操作。滤波器单元216可以被配置为执行本公开的各种技术,例如,确定将用于对解码图片进行滤波的一个或多个神经网络模型(NN模型)232和/或是否应用NN模型滤波。模式选择单元202可以使用经滤波和未滤波的图片两者来执行RD计算,以确定RD成本以确定是否执行NN模型滤波,然后向熵编码单元220提供表示例如是否要执行NN模型滤波、要用于当前图片或其一部分的一个或多个NN模型232等的的数据。

[0137] 具体地,根据本公开的技术,滤波器单元216可以确定NN模型232包括可以应用于解码图片的一部分的可用NN模型的集合。在一些示例中,滤波器单元216可以确定只有NN模型232的子集可用于解码图片的该部分(其中子集包括比NN模型232的完整集合更少的NN模型)。在这样的示例中,滤波器单元216可以向熵编码单元220提供定义可用于解码图片的该部分的NN模型232的子集的数据。熵编码单元220可以例如在SPS、PPS、APS、条带头、图片头或其他高级语法元素中信令通知指示子集的数据。

[0138] 在一些示例中,滤波器单元216可以为解码图片的一部分选择NN个模型232中的一个。为了选择NN个模型232中的一个,滤波器单元216可以确定解码图片的该部分的量化参数。可替代地,模式选择单元202可以执行率失真优化(RDO)过程以确定NN个模型232中的哪一个产生最佳RDO性能,并选择NN个模型232中产生最佳RDO性能的一个。模式选择单元202然后可以提供表示对与所确定的产生最佳RDO性能的NN模型相对应的可用NN模型232的集合(或子集)的索引的值。

[0139] 在一些示例中,模式选择单元202还可以确定是启用还是禁用针对例如解码图片的特定区域、整个解码图片、解码图片序列等的NN模型232的使用。模式选择单元202可以向熵编码单元220提供指示NN模型232的使用是被启用还是被禁用的数据。熵编码单元220还可以对表示是启用还是禁用使用NN模型232的数据进行编码,例如在以下一个或多个中:视频参数集(VPS)、序列参数集(SPS)、图片参数集(PPS)、图片头、条带头、自适应参数集(APS)、帧内时段级别、图片组(GOP)级别、GOP中的时间级别、图片级别、条带级别、编解码树单元(CTU)级别或网格尺寸级别。

[0140] 在一些示例中,解码视频数据的部分可以是网格的元素。也就是说,视频编码器200可以将解码图片划分为多个元素的网格,例如,由网格形成的解码图片的行和列中的元素。熵编码单元220还可以对表示网格的数据进行编码,例如,表示行和列的数量和/或网格的元素的数量的语法元素的值。

[0141] 视频编码器200将重建(并且在一些情况下经滤波的)块存储在DPB 218中。例如,在不需滤波器单元216的操作的示例中,重建单元214可以将重建块存储到DPB 218。在需滤波器单元216的操作的示例中,滤波器单元216可以将经滤波的重建块存储到DPB 218。运动估计单元222和运动补偿单元224可以从DPB 218中检索参考图片,该参考图片是由重建(并且潜在地滤波的)块形成的,来对随后编码的图片的块进行帧间预测。另外,帧内预测单元226可以使用当前图片的DPB 218中的重建块来对当前图片中的其他块进行帧内预测。

[0142] 一般地,熵编码单元220可以对从视频编码器200的其他功能组件接收的语法元素进行熵编码。例如,熵编码单元220可以对来自量化单元208的量化变换系数块进行熵编码。作为另一示例,熵编码单元220可以对来自模式选择单元202的预测语法元素(例如,用于帧间预测的运动信息或用于帧内预测的帧内模式信息)进行熵编码。熵编码单元220可以对视频数据的另一示例的语法元素执行一个或多个熵编码操作来生成熵编码的数据。例如,熵编码单元220可以对数据执行上下文自适应可变长度编解码(CAVLC)操作、CABAC操作、可变到可变(V2V)长度编解码操作、基于语法的上下文自适应二进制算术编解码(SBAC)操作、概率区间分割熵(PIPE)编解码操作、指数-Golomb编解码操作或另一类型的熵编码操作。在一些示例中,熵编码单元220可以以旁路模式操作,在旁路模式中语法元素不被熵编码。

[0143] 视频编码器200可以输出比特流,其包括对条带或图片的块进行重建所需的熵编码的语法元素。具体地,熵编码单元220可以输出该比特流。

[0144] 上述操作是针对块来描述的。这样的描述应该理解为是用于亮度编解码块和/或色度编解码块的操作。如上所述,在一些示例中,亮度编解码块和色度编解码块是CU的亮度和色度分量。在一些示例中,亮度编解码块和色度编解码块是PU的亮度和色度分量。

[0145] 在一些示例中,不需要针对色度编解码块重复关于亮度编解码块执行的操作。作为一个示例,不需要重复用于识别亮度编解码块的运动矢量(MV)和参考图片的操作来识别用于色度块的MV和参考图片。相反,可以缩放亮度编解码块的MV来确定色度块的MV,并且参考图片可以是相同的。作为另一示例,针对亮度编解码块和色度编解码块,帧内预测过程可以是相同的。

[0146] 以此方式,图6的视频编码器200表示用于对解码视频数据进行滤波的示例设备的示例,该设备包括:存储器,被配置为存储视频数据;以及一个或多个处理器,其在电路中实施并且被配置为:对视频数据的图片进行解码;对表示要用于对解码图片的一部分进行滤波的神经网络模型的语法元素的值进行编解码,该值表示对预定义神经网络模型集合的索引,该索引对应于预定义神经网络模型集合中的神经网络模型;以及使用对应于索引的神经网络模型来对解码图片的该部分进行滤波。

[0147] 图7是示出可以执行本公开的技术的示例视频解码器300的框图。提供图7是出于解释的目的,而非限制本公开中广泛示例和描述的技术。为了说明的目的,本公开描述了根据VCC和HEVC(ITU-T H.265)的技术的视频解码器300。然而,本公开的技术可以由被配置为其他视频编解码标准的视频编解码设备来执行。

[0148] 在图7的示例中,视频解码器300包括编解码图片缓冲器(CPB)存储器320、熵解码单元302、预测处理单元304、逆量化单元306、逆变换处理单元308、重建单元310、滤波器单元312和解码图片缓冲器(DPB)314。CPB存储器320、熵解码单元302、预测处理单元304、逆量化单元306、逆变换处理单元308、重建单元310、滤波器单元312和DPB 314中的任何一个或

全部可以在一个或多个处理器或处理电路中实施。例如,视频解码器300的单元可以被实施为一个或多个电路或逻辑元件,作为硬件电路的一部分,或者作为处理器、ASIC或FPGA的一部分。此外,视频解码器300可以包括附加的或替代的处理器或处理电路,以执行这些和其他功能。

[0149] 预测处理单元304包括运动补偿单元316和帧内预测单元318。预测处理单元304可以包括附加单元以按照其他预测模式来执行预测。作为示例,预测处理单元304可以包括调色板单元、帧内块复制单元(其可以形成运动补偿单元316的一部分)、仿射单元、线性模型(LM)单元等。在其他示例中,视频解码器300可以包括更多、更少或不同的功能组件。

[0150] CPB存储器320可以存储将由视频解码器300的组件进行解码的视频数据,诸如编码的视频比特流。例如,可以从计算机可读介质110(图1)获得存储在CPB存储器320中的视频数据。CPB存储器320可以包括存储来自编码的视频比特流的编码的视频数据(例如,语法元素)的CPB。而且,CPB存储器320可以存储除了编解码图片的语法元素之外的视频数据,诸如表示来自视频解码器300的各种单元的输出的临时数据。DPB 314通常存储解码的图片,当对编码的视频比特流的后续数据或图片进行解码时,视频解码器300可以输出该解码的图片和/或将其用作参考视频数据。CPB存储器320和DPB 314可以由多种存储器设备中的任何一种形成,诸如动态随机存取存储器(DRAM),包括同步DRAM(SDRAM)、磁阻RAM(MRAM)、阻变RAM(RRAM)或其他类型的存储器设备。CPB存储器320和DPB 314可以由相同的存储器设备或独立的存储器设备提供。在各种示例中,CPB存储器320可以与视频解码器300的其他组件置于片上,或者相对于那些组件置于片外。

[0151] 附加地或替代地,在一些示例中,视频解码器300可以从存储器120(图1)中检索编解码的视频数据。即,存储器120可以如上述讨论与CPB存储器320一起存储数据。同样,当视频解码器300的某些或全部功能在将由视频解码器300的处理电路执行的软件中实施时,存储器120可以存储将由视频解码器300执行的指令。

[0152] 示出图7所示的各种单元以帮助理解由视频解码器300执行的操作。这些单元可以被实施为固定功能电路、可编程电路或其组合。类似于图6,固定功能电路是指提供特定功能并且在可以执行的操作上预设的电路。可编程电路是指可以被编程以执行各种任务并且在可以执行的操作中提供灵活功能的电路。例如,可编程电路可以执行软件或固件,软件或固件使可编程电路以由软件或固件的指令所定义的方式操作。固定功能电路可以执行软件指令(例如,以接收参数或输出参数),但是固定功能电路执行的操作类型通常是不可变的。在一些示例中,一个或多个单元可以是不同的电路块(固定功能或可编程),并且在一些示例中,一个或多个单元可以是集成电路。

[0153] 视频解码器300可以包括ALU、EFU、数字电路、模拟电路和/或由可编程电路形成的可编程核心。在视频解码器300的操作由在可编程电路上执行的软件执行的示例中,片上或片外存储器可以存储视频解码器300接收并执行的软件的指令(例如,目标代码)。

[0154] 熵解码单元302可以从CPB接收编码的视频数据,并且对视频数据进行熵解码以再现语法元素。预测处理单元304、逆量化单元306、逆变换处理单元308、重建单元310和滤波器单元312可以基于从比特流中提取的语法元素来生成解码的视频数据。

[0155] 一般地,视频解码器300在逐块的基础上重建图片。视频解码器300可单独地对每一个块执行重建操作(其中当前正在进行重建(即,解码)的块可以称为“当前块”)。

[0156] 熵解码单元302可以对定义经量化的变换系数块的经量化的变换系数的语法元素以及诸如量化参数(QP)和/或(多个)变换模式指示的变换信息进行熵解码。逆量化单元306可以使用与经量化的变换系数块相关联的QP来确定量化程度,并且同样地,确定逆量化程度以供逆量化单元306应用。逆量化单元306可以(例如,执行按比特左移操作)对经量化的变换系数进行逆量化。逆量化单元306从而可以形成包括变换系数的变换系数块。

[0157] 在逆量化单元306形成变换系数块之后,逆变换处理单元308可以将一个或多个逆变换应用于变换系数块来生成与当前块相关联的残差块。例如,逆变换处理单元308可以将逆DCT、逆整数变换、逆Karhunen-Loeve变换(KLT)、逆旋转变换、逆方向变换或另一逆变换应用于变换系数块。

[0158] 进一步来说,预测处理单元304根据由熵解码单元302熵解码的预测信息语法元素来生成预测块。例如,如果预测信息语法元素指示当前块是帧间预测的,则运动补偿单元316可以生成预测块。在这种情况下,预测信息语法元素可以指示DPB 314中的参考图片,从该参考图片中检索参考块,以及指示识别参考图片中的参考块相对于当前图片中的当前块的位置的运动矢量。运动补偿单元316通常可以以与针对运动补偿单元224(图6)所描述的方式基本上相似的方式来执行帧间预测过程。

[0159] 作为另一示例,如果预测信息语法元素指示当前块是帧内预测的,则帧内预测单元318可以根据由预测信息语法元素指示的帧内预测模式来生成预测块。再次,帧内预测单元318通常可以以与针对帧内预测单元226(图6)所描述的方式基本上相似的方式来执行帧内预测过程。帧内预测单元318可以从DPB 314中检索当前块的相邻样点的数据。

[0160] 重建单元310可以使用预测块和残差块来重建当前块。例如,重建单元310可以将残差块的样点添加到预测块的对应样点来重建当前块。

[0161] 滤波器单元312可以对重建块执行一个或多个滤波器操作。例如,滤波器单元312可以执行去方块操作来减少沿着重建块的边缘的块效应伪影。不一定在所有示例中都执行滤波器单元312的操作。例如,视频解码器300可以显式或隐式地确定是否要使用NN模型322来执行神经网络模型滤波,例如,使用本文讨论的各种技术中的任何或所有技术。此外,视频解码器300可以显式或隐式地确定NN个模型322中的一个或多个和/或待解码和滤波的当前图片的网格尺寸。因此,当滤波被打开(switched on)时,滤波器单元312可以使用NN个模型322中的一个或多个来对当前解码图片的一部分进行滤波。

[0162] 滤波器单元312可以对重建块执行一个或多个滤波器操作。例如,滤波器单元312可以执行去方块操作来减少沿着CU/TU的边缘的块效应伪影。在一些示例中,可以跳过滤波器单元312的操作。滤波器单元312可以被配置为执行本公开的各种技术,例如,以确定将用于对解码图片进行滤波的一个或多个神经网络模型(NN模型) 232和/或是否应用NN模型滤波。模式选择单元202可以使用经滤波和未滤波的图片两者来执行RD计算,以确定RD成本以确定是否执行NN模型滤波,然后向熵编码单元220提供表示例如是否要执行NN模型滤波、要用于当前图片或其一部分的一个或多个NN模型322等的的数据。

[0163] 具体地,根据本公开的技术,滤波器单元312可以确定NN模型322包括可以应用于解码图片的一部分的可用NN模型的集合。在一些示例中,滤波器单元312可以确定只有NN模型322的子集可用于解码图片的该部分(其中子集包括比NN模型322的完整集合更少的NN模型)。在这样的示例中,滤波器单元312可以从熵解码单元302接收定义可用于解码图片的该

部分的NN模型322的子集的解码数据。熵解码单元302可以例如在SPS、PPS、APS、条带头、图片头或其他高级语法元素中对指示子集的数据进行解码。

[0164] 在一些示例中,滤波器单元312可以为解码图片的一部分选择NN个模型322中的一个。为了选择NN个模型322中的一个,滤波器单元312可以确定解码图片的该部分的量化参数(QP)并且确定对应于QP的NN模型322中的一个。可替代地,滤波单元312可以接收表示对与所确定的要使用的NN模型相对应的可用NN模型322的集合(或子集)的索引的值。

[0165] 在一些示例中,滤波器单元312还可以确定是启用还是禁用针对例如解码图片的特定区域、整个解码图片、解码图片序列等的NN模型322的使用。滤波器单元312可以从熵解码单元302接收指示NN模型322的使用是被启用还是被禁用的数据。熵解码单元302可以对表示是启用还是禁用使用NN模型322的数据进行解码,例如,在以下中的一个或多个中:视频参数集(VPS)、序列参数集(SPS)、图片参数集(PPS)、图片头、条带头、自适应参数集(APS)、帧内时段级别、图片组(GOP)级别、GOP中的时间层级别、图片级别、条带级别、编解码树单元(CTU)级别或网格尺寸级别。

[0166] 在一些示例中,解码视频数据的部分可以是网格的元素。也就是说,视频解码器300可以将解码图片划分为多个元素的网格,例如,由网格形成的解码图片的行和列中的元素。熵解码单元302可以对表示网格的数据进行解码,例如,表示行和列的数量和/或网格的元素的数量的语法元素的值。

[0167] 视频解码器300可以将重建块存储在DPB 314中。例如,在不执行滤波器单元312的操作的示例中,重建单元310可以将重建块存储到DPB 314。在执行滤波器单元312的操作的示例中,滤波器单元312可以将经滤波的重建块存储到DPB 314。如上所述,DPB 314可以向预测处理单元304提供参考信息,诸如用于帧内预测的当前图片的样点以及用于随后的运动补偿的先前解码的图片。此外,视频解码器300可以从DPB 314输出解码的图片,用于随后呈现在诸如图1的显示设备118的显示设备上。

[0168] 以此方式,图7的视频解码器300表示用于对解码视频数据进行滤波的设备的示例,该设备包括:存储器,被配置为存储视频数据;以及一个或多个处理器,其在电路中实施并且被配置为:对视频数据的图片进行解码;对表示要用于对解码图片的一部分进行滤波的神经网络模型的语法元素的值进行编解码,该值表示对预定义神经网络模型集合的索引,该索引对应于预定义神经网络模型集合中的神经网络模型;以及使用对应于索引的神经网络模型来对解码图片的部分进行滤波。

[0169] 图8是示出根据本公开的技术用于对当前块进行编码的示例方法的流程图。当前块可以包括当前CU。尽管是关于视频编码器200(图1和图3)描述的,但应当理解的是,其他设备可以被配置为执行类似于图8的方法。

[0170] 在此示例中,视频编码器200最初对当前块进行预测(350)。例如,视频编码器200可以形成用于当前块的预测块。然后,视频编码器200可以计算当前块的残差块(352)。为了计算残差块,视频编码器200可以计算原始的、未编码的块与当前块的预测块之间的差。然后,视频编码器200可以对残差块的系数进行变换和量化(354)。接下来,视频编码器200可以扫描残差块的经量化的变换系数(356)。在扫描期间或在扫描之后,视频编码器200可以对系数进行熵编码(358)。例如,视频编码器200可以使用CAVLC或CABAC对系数进行编码。然后,视频编码器200可以输出块的经熵编解码的数据(360)。

[0171] 视频编码器200还可以在对当前块进行编码之后对当前块进行解码,以将当前块的解码版本用作后续编码的数据的参考数据(例如,在帧间或帧内预测模式中)。因此,视频编码器200可以对系数进行逆量化和逆变换以再现残差块(362)。视频编码器200可以将残差块与预测块组合以形成解码块(364)。

[0172] 根据本公开的技术,在以上述方式对图片的所有块进行编码和解码之后,视频编码器200可以确定用于包括当前块的解码图片的一部分的神经网络(NN)模型,以应用于解码图片的该部分(366)。在一个示例中,视频编码器200可以确定解码图片的该部分的量化参数(QP),并确定对应于QP的NN模型。在另一示例中,视频编码器200可以执行率失真优化(RDO)过程以例如从可用的NN模型集合中选择NN模型。也就是说,视频编码器200可以将各种NN模型应用于解码图片的该部分,然后计算各种NN模型中的每一个的RDO值,并为解码图片的该部分选择产生最佳测试RDO值的NN模型。

[0173] 然后,视频编码器200可以对表示NN模型的值进行编码(368)。当视频编码器200基于QP选择NN模型时,视频编码器200可以简单地对QP进行编码以表示NN模型。也就是说,对于块编码的QP(例如,作为TU信息的一部分)可以被提供给NN模型以选择适当的基于NN的滤波器。可替代地,当视频编码器200从可用NN模型的集合(或集合的子集)中选择NN模型时,视频编码器200可以对表示对集合(或子集)的索引的语法元素的值进行编码,其中该索引对应于NN模型在该集合或子集中的位置。语法元素可以形成以下中的一部分:视频参数集(VPS)、序列参数集(SPS)、图片参数集(PPS)、自适应参数集(APS)、帧内时段级别、图片组(GOP)级别、GOP中的时间层级、图片级别、条带级别、编解码树单元(CTU)级别、或图片的网格的网格级别。然后,视频编码器200可以将所确定的NN模型应用于解码图片的该部分(369),以对解码图片的该部分进行滤波并将解码图片存储在DPB 218中。

[0174] 在一些示例中,视频编码器200可以为解码图片的该部分确定多个NN模型,而不是确定单个NN模型。视频编码器200可以将每个NN模型分别应用于解码图片的该部分,以形成不同的滤波结果。视频编码器200然后可以组合每个滤波结果以形成解码图片的最终经滤波的部分。

[0175] 解码图片的该部分可以对应于网格的元素。也就是说,视频编码器200可以将解码图片划分成具有多个元素(例如,由相应数量的行和列形成的区域)的网格。视频编码器200可以在以下至少一个中信令通知表示诸如网格的行和列的数量或者元素的数量数据:视频参数集(VPS)、序列参数集(SPS)、图片参数集(PPS)、图片头、条带头、自适应参数集(APS)、帧内时段级别、图片组(GOP)级别、GOP中的时间层级、图片级别、条带级别、编解码树单元(CTU)级别或网格尺寸级别。在一些示例中,视频编码器200可以信令通知表示对可能的网格尺寸集合的索引的语法元素的值。

[0176] 以此方式,图8的方法表示用于对解码视频数据进行滤波的方法的示例,包括:对视频数据的图片进行解码;对表示要用于对解码图片的一部分进行滤波的神经网络模型的语法元素的值进行编解码,该值表示对预定义神经网络模型集合的索引,该索引对应于预定义神经网络模型集合中的神经网络模型;以及使用对应于索引的神经网络模型来对解码图片的该部分进行滤波。

[0177] 图9是示出根据本公开的技术用于对当前块进行解码的示例方法的流程图。当前块可以包括当前CU。尽管是关于视频解码器300(图1和图4)描述的,但是应当理解,其他设

备可以被配置为执行类似于图9的方法。

[0178] 视频解码器300可以接收当前块的经熵编码的数据,诸如经熵编码的预测信息和与当前块相对应的残差块的系数的经熵编码的数据(370)。视频解码器300可以对熵编码的数据进行熵解码,以确定当前块的预测信息、用于包括当前块的图片的一部分的神经网络(NN)模型,并再现残差块的系数(372)。视频解码器300可以对当前块进行预测(374),例如,使用用于当前块的预测信息所指示的帧内或帧间预测模式,以计算用于当前块的预测块。然后,视频解码器300可以对再现的系数进行逆扫描(376),以创建经量化的变换系数的块。然后,视频解码器300可以对经量化的变换系数进行逆量化和逆变换以产生残差块(378)。视频解码器300可以通过将预测块和残差块组合来最终解码当前块(380)。

[0179] 视频解码器300还可以从解码数据(382)确定要应用于包括当前块的解码图片的一部分的NN模型。例如,视频解码器300可以对用于解码图片的该部分(例如,对于该部分中的一个或多个块)的量化参数(QP)进行解码,并确定该QP所对应的NN模型。作为另一示例,视频解码器300可以对表示对可用NN模型的集合(或集合的子集)的索引的语法元素的值进行解码,并且确定该可用NN模型的集合或子集中该索引所对应的NN模型。然后,视频解码器300可以将NN模型应用于解码图片的该部分(384),以对解码图片的该部分进行滤波。在其中滤波器是环内滤波器的一个示例中,视频解码器300将包括经滤波的部分的解码图片存储到解码图片缓冲器314中。

[0180] 以此方式,图9的方法表示用于对解码视频数据进行滤波的方法的示例,包括:对视频数据的图片进行解码;对表示要用于对解码图片的一部分进行滤波的神经网络模型的语法元素的值进行编解码,该值表示对预定义神经网络模型集合的索引,该索引对应于预定义神经网络模型集合中的神经网络模型;以及使用对应于索引的神经网络模型来对解码图片的该部分进行滤波。

[0181] 以下条款概括了本公开的各种技术:

[0182] 条款1:一种对解码视频数据进行滤波的方法,该方法包括:确定要被用于对视频数据的解码图片的一部分进行滤波的一个或多个神经网络模型;以及使用一个或多个神经网络模型来对解码图片的该部分进行滤波。

[0183] 条款2:根据条款1的方法,其中,确定一个或多个神经网络模型包括确定不多于单个神经网络模型。

[0184] 条款3:根据条款1的方法,其中,确定一个或多个神经网络模型包括确定多个神经网络模型。

[0185] 条款4:根据条款3的方法,其中,滤波包括:将多个神经网络模型中的每一个分别应用于该部分以形成不同的结果;以及组合每个结果以形成最终经滤波的部分。

[0186] 条款5:根据条款1-条款4的任一项的方法,其中,确定包括使用在包括视频数据的比特流中信令通知的索引来确定。

[0187] 条款6:根据条款1-条款5的任一项的方法,其中,该部分包括网格的元素。

[0188] 条款7:根据条款1-条款6的任一项的方法,还包括确定网格的元素的数量。

[0189] 条款8:根据条款7的方法,其中,确定网格的元素的数量包括对以下中的至少一个的语法元素进行解码:视频参数集(VPS)、序列参数集(SPS)、图片参数集(PPS)、图片头、条带头、自适应参数集(APS)、帧内时段级别、图片组(GOP)级别、GOP中的时间层级别、图片级

别、条带级别、编解码树单元 (CTU) 级别、或网格尺寸级别。

[0190] 条款9:根据条款7和条款8的任一项的方法,其中,确定网格中的元素数量包括将表示对可能的网格尺寸的集合的索引的语法元素的值进行解码。

[0191] 条款10:根据条款1-条款9的任一项的方法,还包括在确定一个或多个神经网络模型之前确定要应用一个或多个神经网络模型。

[0192] 条款11:根据条款10的方法,其中,确定要应用一个或多个神经网络模型包括对具有指示要应用该一个或多个神经网络模型的值的语法元素进行解码。

[0193] 条款12:根据条款11的方法,其中,该语法元素是以下中的至少一个:视频参数集 (VPS)、序列参数集 (SPS)、图片参数集 (PPS)、图片头、条带头、自适应参数集 (APS)、帧内时段级别、图片组 (GOP) 级别、GOP中的时间层级别、图片级别、条带级别、编解码树单元 (CTU) 级别、或网格尺寸级别。

[0194] 条款13:根据条款1-条款12的任一项的方法,其中,该部分包括解码图片的颜色分量的一部分,颜色分量包括亮度分量、蓝色色调度分量或红色色调度分量之一。

[0195] 条款14:根据条款13的方法,还包括对联合表示使用一个或多个神经网络模型对解码图片的颜色分量中的每一个进行滤波的语法元素进行解码。

[0196] 条款15:根据条款1-条款14的任一项的方法,还包括:对当前图片进行编码;以及对当前图片进行解码以形成当前图片。

[0197] 条款16:根据条款15的方法,其中,确定包括根据率失真计算来确定。

[0198] 条款17:一种用于对视频数据进行解码的设备,该设备包括用于执行条款1-条款16的任一项的方法的一个或多个部件。

[0199] 条款18:根据条款17的设备,其中,一个或多个部件包括在电路中实施的一个或多个处理器。

[0200] 条款19:根据条款17的设备,还包括被配置为显示解码视频数据的显示器。

[0201] 条款20:根据条款17的设备,其中,该设备包括相机、计算机、移动设备、广播接收器设备或机顶盒中的一个或多个。

[0202] 条款21:根据条款17的设备,还包括被配置为存储视频数据的存储器。

[0203] 条款22:一种计算机可读存储介质,其上存储有指令,该指令在被执行时使处理器执行条款1-条款16的任一项的方法。

[0204] 条款23:一种对解码视频数据进行滤波的方法,该方法包括:对视频数据的图片进行解码;对表示要用于对解码图片的一部分进行滤波的神经网络模型的语法元素的值进行编解码,该值表示对预定义神经网络模型集合的索引,该索引对应于预定义神经网络模型集合中的神经网络模型;以及使用对应于索引的神经网络模型来对解码图片的该部分进行滤波。

[0205] 条款24:根据条款23的方法,还包括对表示预定义神经网络模型集合的子集的值进行编解码,神经网络模型是预定义神经网络模型集合的子集中的一个,该子集小于预定义神经网络模型集合。

[0206] 条款25:根据条款23的方法,其中,表示神经网络模型的语法元素的值是表示多个神经网络模型的值。

[0207] 条款26:根据条款25的方法,其中,滤波包括:将由该值表示的多个神经网络模型

中的每一个分别应用于该部分以形成不同的结果;以及组合每个结果以形成最终经滤波的部分。

[0208] 条款27:根据条款23的方法,其中,对语法元素的值进行编解码包括在以下中的一个中对该语法元素的值进行编解码:视频参数集(VPS)、序列参数集(SPS)、图片参数集(PPS)、自适应参数集(APS)、帧内时段级别、图片组(GOP)级别、GOP中的时间层级别、图片级别、条带级别、编解码树单元(CTU)级别、或图片的网格的网格级别。

[0209] 条款28:根据条款23的方法,还包括根据网格来分割图片,其中该部分包括图片的网格的元素。

[0210] 条款29:根据条款28的方法,还包括确定网格的元素的数量。

[0211] 条款30:根据条款29的方法,其中,确定网格的元素的数量包括对以下中的至少一个的语法元素进行解码:视频参数集(VPS)、序列参数集(SPS)、图片参数集(PPS)、图片头、条带头、自适应参数集(APS)、帧内时段级别、图片组(GOP)级别、GOP中的时间层级别、图片级别、条带级别、编解码树单元(CTU)级别、或网格尺寸级别。

[0212] 条款31:根据条款29的方法,其中,确定网格中的元素的数量包括将表示对可能的网格尺寸的集合的索引的语法元素的值进行解码。

[0213] 条款32:根据条款23的方法,还包括在确定神经网络模型之前确定要应用神经网络模型。

[0214] 条款33:根据条款32的方法,其中,确定要应用神经网络模型包括对指示要应用神经网络模型的语法元素的值进行解码。

[0215] 条款34:根据条款33的方法,其中,该语法元素是以下中的至少一个:视频参数集(VPS)、序列参数集(SPS)、图片参数集(PPS)、图片头、条带头、自适应参数集(APS)、帧内时段级别、图片组(GOP)级别、GOP中的时间层级别、图片级别、条带级别、编解码树单元(CTU)级别、或网格尺寸级别。

[0216] 条款35:根据条款23的方法,其中,解码图片的该部分包括解码图片的颜色分量的一部分,颜色分量包括亮度分量、蓝色色调度分量或红色色调度分量之一。

[0217] 条款36:根据条款35的方法,还包括对联合表示使用神经网络模型对解码图片的颜色分量中的每一个进行滤波的语法元素进行编解码。

[0218] 条款37:根据条款23的方法,其中,语法元素的值包括图片的部分的量化参数(QP)。

[0219] 条款38:根据条款23的方法,还包括在对图片进行解码之前对图片进行编码,其中对语法元素的值进行编解码包括对语法元素的值进行编码。

[0220] 条款39:根据条款38的方法,还包括根据率失真计算来确定神经网络模型。

[0221] 条款40:一种用于对解码视频数据进行滤波的设备,该设备包括:存储器,被配置为存储视频数据;以及一个或多个处理器,其在电路中实施并且被配置为:对视频数据的图片进行解码;对表示要用于对解码图片的一部分进行滤波的神经网络模型的语法元素的值进行编解码,该值表示对预定义神经网络模型集合的索引,该索引对应于预定义神经网络模型集合中的神经网络模型;以及使用对应于索引的神经网络模型来对解码图片的该部分进行滤波。

[0222] 条款41:根据条款40的设备,其中,一个或多个处理器还被配置为对表示预定义神

神经网络模型集合的子集的值进行编解码,神经网络模型是预定义神经网络模型集合的子集中的一个,该子集小于预定义神经网络模型集合。

[0223] 条款42:根据条款40的设备,其中,表示神经网络模型的语法元素的值是表示多个神经网络模型的值,并且其中,为了对解码图片的该部分进行滤波,一个或多个处理器被配置为:将由该值表示的多个神经网络模型中的每一个分别应用于该部分以形成不同的结果;以及组合每个结果以形成最终经滤波的部分。

[0224] 条款43:根据条款40的设备,其中,一个或多个处理器被配置为在以下中的一个中对语法元素的值进行编解码:视频参数集(VPS)、序列参数集(SPS)、图片参数集(PPS)、自适应参数集(APS)、帧内时段级别、图片组(GOP)级别、GOP中的时间层级别、图片级别、条带级别、编解码树单元(CTU)级别、或图片的网格的网格级别。

[0225] 条款44:根据条款40的设备,其中,一个或多个处理器被配置为根据网格来分割图片,其中,该部分包括图片的网格的元素。

[0226] 条款45:根据条款44的设备,其中,一个或多个处理器还被配置为对表示网格的元素数量的语法元素的值进行解码,该语法元素包括在以下中的至少一个中:视频参数集(VPS)、序列参数集(SPS)、图片参数集(PPS)、图片头、条带头、自适应参数集(APS)、帧内时段级别、图片组(GOP)级别、GOP中的时间层级别、图片级别、条带级别、编解码树单元(CTU)级别、或网格尺寸级别。

[0227] 条款46:根据条款40的设备,其中,一个或多个处理器被配置为在确定神经网络模型之前确定要应用神经网络模型,其中,为了确定要应用神经网络模型,一个或多个处理器被配置为对指示要应用神经网络模型的语法元素的值进行解码。

[0228] 条款47:根据条款40的设备,其中,解码图片的该部分包括解码图片的颜色分量的一部分,颜色分量包括亮度分量、蓝色色调度分量或红色色调度分量之一。

[0229] 条款48:根据条款40的设备,其中,语法元素的值包括图片的部分的量化参数(QP)。

[0230] 条款49:根据条款40的设备,其中,一个或多个处理器还被配置为在对图片进行解码之前对图片进行编码,并且其中,为了对语法元素的值进行编解码,一个或多个处理器被配置为对语法元素的值进行编码。

[0231] 条款50:根据条款27的设备,其中,一个或多个处理器还被配置为根据率失真计算来确定神经网络模型。

[0232] 条款51:根据条款40的设备,还包括被配置为显示解码视频数据的显示器。

[0233] 条款52:根据条款40的设备,其中,该设备包括相机、计算机、移动设备、广播接收器设备或机顶盒中的一个或多个。

[0234] 条款53:一种计算机可读存储介质,其上存储有指令,该指令在被执行时使处理器:对视频数据的图片进行解码;对表示要用于对解码图片的一部分进行滤波的神经网络模型的语法元素的值进行编解码,该值表示对预定义神经网络模型集合的索引,该索引对应于预定义神经网络模型集合中的神经网络模型;以及使用对应于索引的神经网络模型来对解码图片的该部分进行滤波。

[0235] 条款54:根据条款53的计算机可读存储介质,还包括用于使处理器对表示预定义神经网络模型集合的子集的值进行编解码的指令,该神经网络模型是预定义神经网络模型

集合的子集中的一个,该子集小于预定义神经网络模型集合。

[0236] 条款55:根据条款53的计算机可读存储介质,其中,表示神经网络模型的语法元素的值是表示多个神经网络模型的值,并且其中,使处理器对解码图片的该部分进行滤波的指令包括用于使处理器进行以下操作的指令:将由该值表示的多个神经网络模型中的每一个分别应用于该部分以形成不同的结果;以及组合每个结果以形成最终经滤波的部分。

[0237] 条款56:根据条款53的计算机可读存储介质,其中,用于使处理器对语法元素的值进行编解码的指令包括用于使处理器在以下中的一个中对语法元素的值进行编解码的指令:视频参数集(VPS)、序列参数集(SPS)、图片参数集(PPS)、自适应参数集(APS)、帧内时段级别、图片组(GOP)级别、GOP中的时间层级别、图片级别、条带级别、编解码树单元(CTU)级别、或图片的网格的网格级别。

[0238] 条款57:根据条款53的计算机可读存储介质,还包括用于使处理器根据网格来分割图片的指令,其中,该部分包括图片的网格的元素。

[0239] 条款58:根据条款57的计算机可读存储介质,还包括用于使处理器对表示网格的元素数量的语法元素的值进行解码的指令,该语法元素包括在以下中的至少一个中:视频参数集(VPS)、序列参数集(SPS)、图片参数集(PPS)、图片头、条带头、自适应参数集(APS)、帧内时段级别、图片组(GOP)级别、GOP中的时间层级别、图片级别、条带级别、编解码树单元(CTU)级别、或网格尺寸级别。

[0240] 条款59:根据条款53的计算机可读存储介质,还包括用于使处理器在确定神经网络模型之前确定要应用神经网络模型的指令,其中,用于使处理器确定要应用神经网络模型的指令包括用于使处理器对指示要应用神经网络模型的语法元素的值进行解码的指令。

[0241] 条款60:根据条款53的计算机可读存储介质,其中,解码图片的该部分包括解码图片的颜色分量的一部分,颜色分量包括亮度分量、蓝色色调度分量或红色色调度分量之一。

[0242] 条款61:根据条款53的计算机可读存储介质,其中,语法元素的值包括图片的部分的量化参数(QP)。

[0243] 条款62:根据条款53的计算机可读存储介质,还包括使处理器在对图片进行解码之前对图片进行编码的指令,并且其中,使处理器对语法元素的值进行编解码的指令包括使处理器对语法元素的值进行编码的指令。

[0244] 条款63:根据条款62的计算机可读存储介质,还包括用于使处理器根据率失真计算来确定神经网络模型的指令。

[0245] 条款64:一种用于对解码视频数据进行滤波的设备,该设备包括:用于对视频数据的图片进行解码的部件;用于对表示要用于对解码图片的一部分进行滤波的神经网络模型的语法元素的值进行编解码的部件,该值表示对预定义神经网络模型集合的索引,该索引对应于预定义神经网络模型集合中的神经网络模型;以及用于使用对应于索引的神经网络模型来对解码图片的该部分进行滤波的部件。

[0246] 条款65:一种对解码视频数据进行滤波的方法,该方法包括:对视频数据的图片进行解码;对表示要用于对解码图片的一部分进行滤波的神经网络模型的语法元素的值进行编解码,该值表示对预定义神经网络模型集合中的索引,该索引对应于预定义神经网络模型集合中的神经网络模型;以及使用对应于索引的神经网络模型来对解码图片的该部分进

行滤波。

[0247] 条款66:根据条款65的方法,还包括对表示预定义神经网络模型集合的子集的值进行编解码,神经网络模型是预定义神经网络模型集合的子集中的一个,该子集小于预定义神经网络模型集合。

[0248] 条款67:根据条款65和条款66的任一项的方法,其中,表示神经网络模型的语法元素的值是表示多个神经网络模型的值。

[0249] 条款68:根据条款67的方法,其中,滤波包括:将由该值表示的多个神经网络模型中的每一个分别应用于该部分以形成不同的结果;以及组合每个结果以形成最终经滤波的部分。

[0250] 条款69:根据条款65-条款68的任一项的方法,其中,对语法元素的值进行编解码包括在以下中的一个中对该语法元素的值进行编解码:视频参数集(VPS)、序列参数集(SPS)、图片参数集(PPS)、自适应参数集(APS)、帧内时段级别、图片组(GOP)级别、GOP中的时间层级别、图片级别、条带级别、编解码树单元(CTU)级别、或图片的网格的网格级别。

[0251] 条款70:根据条款65-条款69的任一项的方法,还包括根据网格来分割图片,其中,该部分包括图片的网格的元素。

[0252] 条款71:根据条款70的方法,还包括确定网格的元素的数目。

[0253] 条款72:根据条款71的方法,其中,确定网格的元素的数目包括对以下中的至少一个的语法元素进行解码:视频参数集(VPS)、序列参数集(SPS)、图片参数集(PPS)、图片头、条带头、自适应参数集(APS)、帧内时段级别、图片组(GOP)级别、GOP中的时间层级别、图片级别、条带级别、编解码树单元(CTU)级别、或网格尺寸级别。

[0254] 条款73:根据条款71的方法,其中,确定网格中的元素的数目包括将表示对可能的网格尺寸的集合的索引的语法元素的值进行解码。

[0255] 条款74:根据条款65-条款73的任一项的方法,还包括在确定神经网络模型之前确定要应用神经网络模型。

[0256] 条款75:根据条款74的方法,其中,确定要应用神经网络模型包括对指示要应用神经网络模型的语法元素的值进行解码。

[0257] 条款76:根据条款75的方法,其中,该语法元素是以下中的至少一个:视频参数集(VPS)、序列参数集(SPS)、图片参数集(PPS)、图片头、条带头、自适应参数集(APS)、帧内时段级别、图片组(GOP)级别、GOP中的时间层级别、图片级别、条带级别、编解码树单元(CTU)级别、或网格尺寸级别。

[0258] 条款77:根据条款65-条款76的任一项的方法,其中,解码图片的该部分包括解码图片的颜色分量的一部分,颜色分量包括亮度分量、蓝色色调度分量或红色色调度分量之一。

[0259] 条款78:根据条款77的方法,还包括对联合表示使用神经网络模型对解码图片的颜色分量中的每一个进行滤波的语法元素进行编解码。

[0260] 条款79:根据条款65-条款78的任一项的方法,其中,语法元素的值包括图片的部分的量化参数(QP)。

[0261] 条款80:根据条款65的方法,还包括在对图片进行解码之前对图片进行编码,其中,对语法元素的值进行编解码包括对语法元素的值进行编码。

[0262] 条款81:根据条款80的方法,还包括根据率失真计算来确定神经网络模型。

[0263] 条款82:一种用于对解码视频数据进行滤波的设备,该设备包括:存储器,被配置为存储视频数据;以及一个或多个处理器,其在电路中实施并且被配置为:对视频数据的图片进行解码;对表示要用于对解码图片的一部分进行滤波的神经网络模型的语法元素的值进行编解码,该值表示对预定义神经网络模型集合的索引,该索引对应于预定义神经网络模型集合中的神经网络模型;以及使用对应于索引的神经网络模型来对解码图片的该部分进行滤波。

[0264] 条款83:根据条款82的设备,其中,一个或多个处理器还被配置为对表示预定义神经网络模型集合的子集的值进行编解码,该神经网络模型是预定义神经网络模型集合的子集中的一个,该子集小于预定义神经网络模型集合。

[0265] 条款84:根据条款82和条款83的任一项的设备,其中,表示神经网络模型的语法元素的值是表示多个神经网络模型的值,并且其中,为了对解码图片的该部分进行滤波,一个或多个处理器被配置为:将由该值表示的多个神经网络模型中的每一个分别应用于该部分以形成不同的结果;以及组合每个结果以形成最终经滤波的部分。

[0266] 条款85:根据条款82-条款84的任一项的方法,其中,一个或多个处理器被配置为在以下中的一个中对语法元素的值进行编解码:视频参数集(VPS)、序列参数集(SPS)、图片参数集(PPS)、自适应参数集(APS)、帧内时段级别、图片组(GOP)级别、GOP中的时间层级别、图片级别、条带级别、编解码树单元(CTU)级别、或图片的网格的网格级别。

[0267] 条款86:根据条款82-条款85的任一项的设备,其中,一个或多个处理器被配置为根据网格来分割图片,其中,该部分包括图片的网格的元素。

[0268] 条款87:根据条款86的设备,其中,一个或多个处理器还被配置为对表示网格的元素的数量的语法元素的值进行解码,该语法元素包括在以下中的至少一个中:视频参数集(VPS)、序列参数集(SPS)、图片参数集(PPS)、图片头、条带头、自适应参数集(APS)、帧内时段级别、图片组(GOP)级别、GOP中的时间层级别、图片级别、条带级别、编解码树单元(CTU)级别、或网格尺寸级别。

[0269] 条款88:根据条款82-条款87的任一项的设备,其中,一个或多个处理器被配置为在确定神经网络模型之前确定要应用神经网络模型,其中,为了确定要应用神经网络模型,一个或多个处理器被配置为对指示要应用神经网络模型的语法元素的值进行解码。

[0270] 条款89:根据条款82-条款88的任一项的设备,其中,解码图片的该部分包括解码图片的颜色分量的一部分,颜色分量包括亮度分量、蓝色色调度分量或红色色调度分量之一。

[0271] 条款90:根据条款82-条款89的任一项的设备,其中,语法元素的值包括图片的部分的量化参数(QP)。

[0272] 条款91:根据条款82的设备,其中,一个或多个处理器还被配置为在对图片进行解码之前对图片进行编码,并且其中,为了对语法元素的值进行编解码,一个或多个处理器被配置为对语法元素的值进行编码。

[0273] 条款92:根据条款91的设备,其中,一个或多个处理器还被配置为根据率失真计算来确定神经网络模型。

[0274] 条款93:根据条款82-条款92的任一项的设备,还包括被配置为显示解码视频数据

的显示器。

[0275] 条款94:根据条款82-条款93的任一项的设备,其中,该设备包括相机、计算机、移动设备、广播接收器设备或机顶盒中的一个或多个。

[0276] 条款95:一种计算机可读存储介质,其上存储有指令,该指令在被执行时使处理器:对视频数据的图片进行解码;对表示要用于对解码图片的一部分进行滤波的神经网络模型的语法元素的值进行编解码,该值表示对预定义神经网络模型集合的索引,该索引对应于预定义神经网络模型集合的神经网络模型;以及使用对应于索引的神经网络模型来对解码图片的该部分进行滤波。

[0277] 条款96:根据条款95的计算机可读存储介质,还包括用于使处理器对表示预定义神经网络模型集合的子集的值进行编解码的指令,该神经网络模型是预定义神经网络模型集合的子集中的一个,该子集小于预定义神经网络模型集合。

[0278] 条款97:根据条款95和条款96的任一项的计算机可读存储介质,其中,表示神经网络模型的语法元素的值是表示多个神经网络模型的值,并且其中,使处理器对解码图片的该部分进行滤波的指令包括用于使处理器进行以下操作的指令:将由该值表示的多个神经网络模型中的每一个分别应用于该部分以形成不同的结果;以及组合每个结果以形成最终经滤波的部分。

[0279] 条款98:根据条款95-条款97的任一项的计算机可读存储介质,其中,用于使处理器对语法元素的值进行编解码的指令包括用于使处理器在以下中的一个中对语法元素的值进行编解码的指令:视频参数集(VPS)、序列参数集(SPS)、图片参数集(PPS)、自适应参数集(APS)、帧内时段级别、图片组(GOP)级别、GOP中的时间层级别、图片级别、条带级别、编解码树单元(CTU)级别、或图片的网格的网格级别。

[0280] 条款99:根据条款95-条款98的任一项的计算机可读存储介质,还包括用于使处理器根据网格来分割图片的指令,其中,该部分包括图片的网格的元素。

[0281] 条款100:根据条款99的计算机可读存储介质,还包括用于使处理器对表示网格的元素的数量的语法元素的值进行解码的指令,该语法元素包括在以下中的至少一个中:视频参数集(VPS)、序列参数集(SPS)、图片参数集(PPS)、图片头、条带头、自适应参数集(APS)、帧内时段级别、图片组(GOP)级别、GOP中的时间层级别、图片级别、条带级别、编解码树单元(CTU)级别、或网格尺寸级别。

[0282] 条款101:根据条款95-条款100的任一项的计算机可读存储介质,还包括用于使处理器在确定神经网络模型之前确定要应用神经网络模型的指令,其中,用于使处理器确定要应用神经网络模型的指令包括用于使处理器对指示要应用神经网络模型的语法元素的值进行解码的指令。

[0283] 条款102:根据条款95-条款101的任一项的计算机可读存储介质,其中,解码图片的该部分包括解码图片的颜色分量的一部分,颜色分量包括亮度分量、蓝色色调度分量或红色色调度分量之一。

[0284] 条款103:根据条款95-条款102的任一项的计算机可读存储介质,其中,语法元素的值包括图片的部分的量化参数(QP)。

[0285] 条款104:根据条款95-条款103的任一项的计算机可读存储介质,还包括使处理器在对图片进行解码之前对图片进行编码的指令,并且其中,使处理器对语法元素的值进行

编解码的指令包括使处理器对语法元素的值进行编码的指令。

[0286] 条款105:根据条款104的计算机可读存储介质,还包括用于使处理器根据率失真计算来确定神经网络模型的指令。

[0287] 条款106:一种用于对解码视频数据进行滤波的设备,该设备包括:用于对视频数据的图片进行解码的部件;用于对表示要用于对解码图片的一部分进行滤波的神经网络模型的语法元素的值进行编解码的部件,该值表示对预定义神经网络模型集合的索引,该索引对应于预定义神经网络模型集合中的神经网络模型;以及用于使用对应于索引的神经网络模型来对解码图片的该部分进行滤波的部件。

[0288] 应该认识到,根据示例,本文中描述的任何技术的某些动作或事件可以以不同的序列执行,可以被一起添加、合并或省去(例如,不是所有描述的动作或事件是技术实践所必须的)。此外,在某些示例中,动作或事件可以,例如通过多线程处理、中断处理或多个处理器并发地执行而不是顺序地执行。

[0289] 在一个或多个示例中,可以以硬件、软件、固件或其任意组合来实施所描述的功能。如果以软件来实施,则功能可以作为一个或多个指令或代码存储在计算机可读介质上或通过计算机可读介质传输,并由基于硬件的处理单元执行。计算机可读介质可以包括计算机可读存储介质,其对应于诸如数据存储介质之类的有形介质,或者通信介质,包括例如根据通信协议来促进将计算机程序从一个地方转移到另一地方的任何介质。以这种方式,计算机可读介质通常可以对应于(1)非暂时性的有形计算机可读存储介质,或者(2)诸如信号或载波之类的通信介质。数据存储介质是可以由一个或多个计算机或一个或多个处理器访问以检索指令、代码和/或数据结构以实施本公开中描述的技术的任何可用介质。计算机程序产品可以包括计算机可读介质。

[0290] 作为示例而非限制,这种计算机可读存储介质可以包括RAM、ROM、EEPROM、CD-ROM或其他光盘存储、磁盘存储或其他磁性存储设备、闪存或可以用于以指令或数据结构形式存储所需程序代码并且可以由计算机访问的任何其他介质。而且,任何连接都适当地称为计算机可读介质。例如,如果使用同轴电缆、光纤电缆、双绞线、数字订户线(DSL)或无线技术(诸如红外、无线电和微波)从网站、服务器或其他远程源发送指令,则介质的定义包括同轴电缆、光纤电缆、双绞线、DSL或无线技术(诸如红外、无线电和微波)。然而,应当理解,计算机可读存储介质和数据存储介质不包括连接、载波、信号或其他暂时性介质,而是针对非暂时性有形存储介质。如本文使用的,磁盘和光盘包括压缩光盘(CD)、激光光盘、光学光盘、数字多功能光盘(DVD)、软盘和蓝光光盘,其中,磁盘通常以磁性方式再现数据,而光盘用激光光学地再现数据。上述的组合也应包括在计算机可读介质的范围内。

[0291] 指令可以由一个或多个处理器执行,诸如一个或多个数字信号处理器(DSP)、通用微处理器、专用集成电路(ASIC)、现场可编程门阵列(FPGA)或其他等效集成的或分立逻辑电路。因此,如本申请中所使用的术语“处理器”和“处理电路”可以是指任何前述结构或适合于实施本申请中描述的技术的任何其他结构。另外,在一些方面,本申请中描述的功能可以在被配置用于编码和解码的专用硬件和/或软件模块内提供,或合并到组合编解码器中。同样,该技术可以在一个或多个电路或逻辑元件中完全实施。

[0292] 本公开的技术可以在包括无线手机、集成电路(IC)或IC集合(例如,芯片集)的多种设备或装置中实施。在本公开中描述各种组件、模块或单元以强调被配置为执行所公开

技术的设备的功能方面,但不一定需要由不同硬件单元来实现。相反,如上所述,各种单元可以组合在编解码器硬件单元中,或者由互操作硬件单元的集合来提供,包括与合适的软件和/或固件结合的如上所述的一个或多个处理器。

[0293] 已经描述了各种示例。这些和其他示例在所附权利要求的范围内。

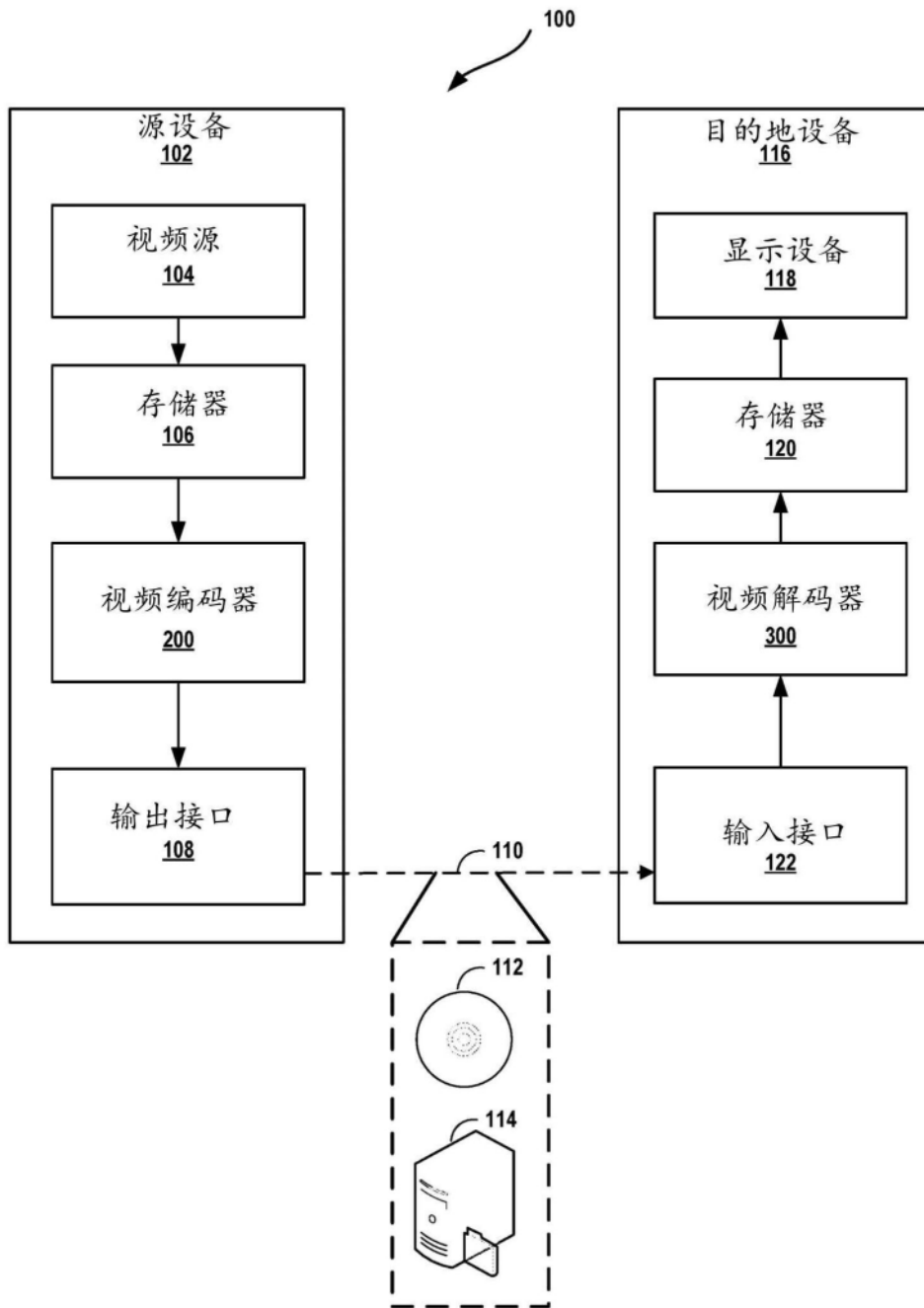


图1

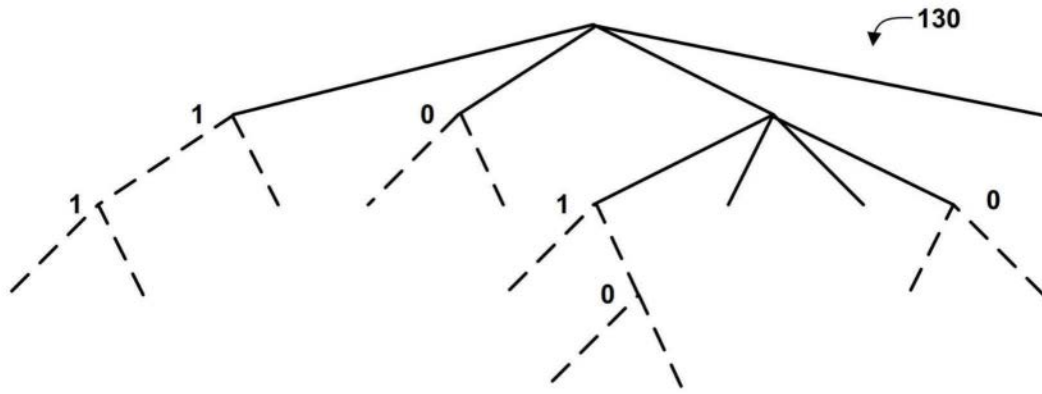


图2A

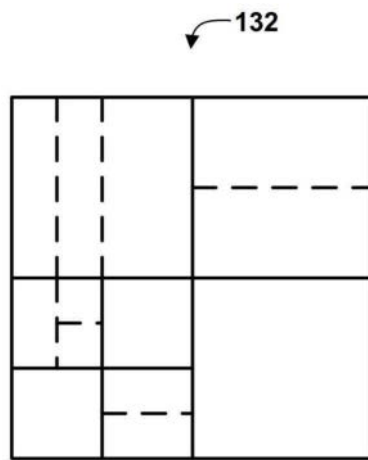


图2B

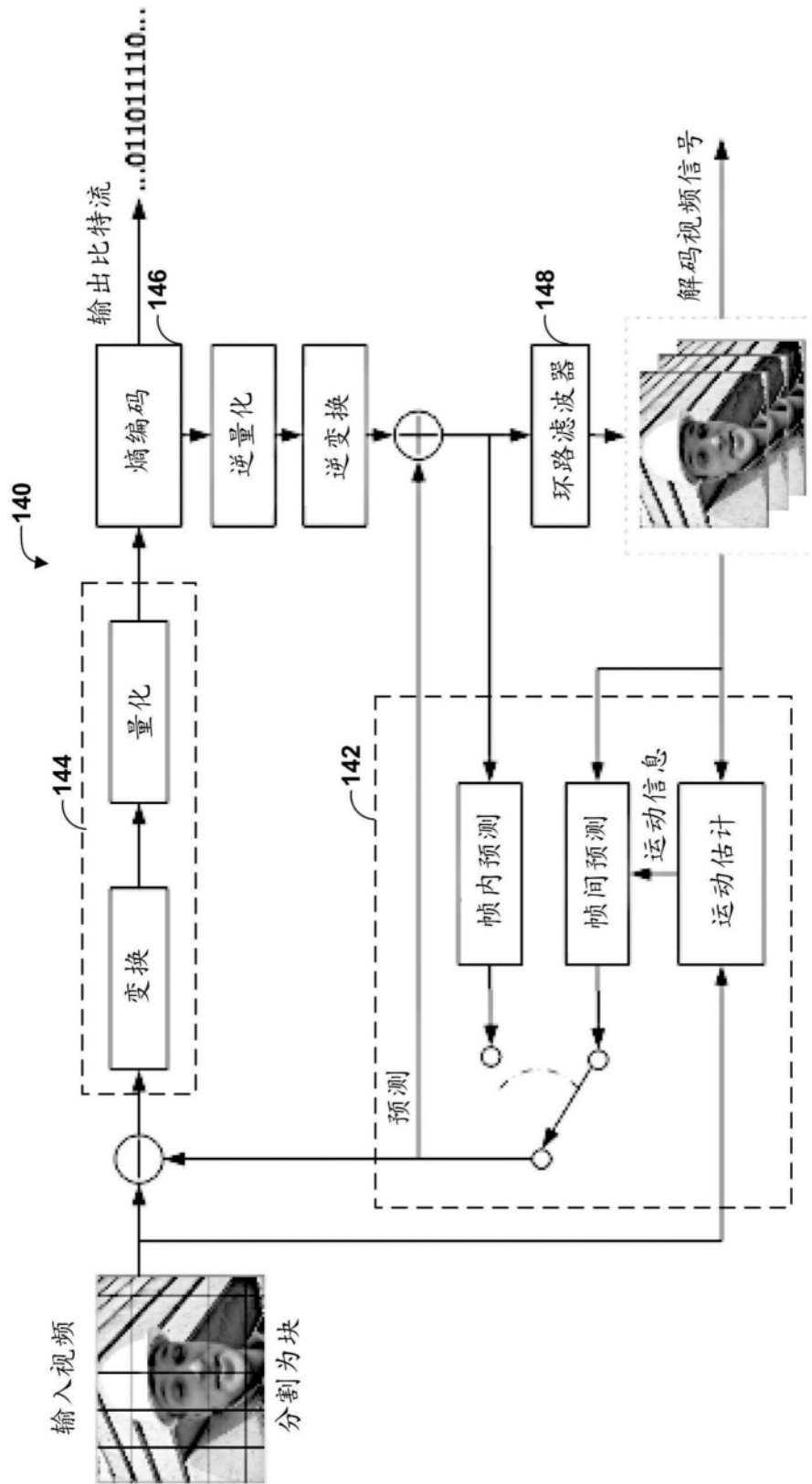


图3

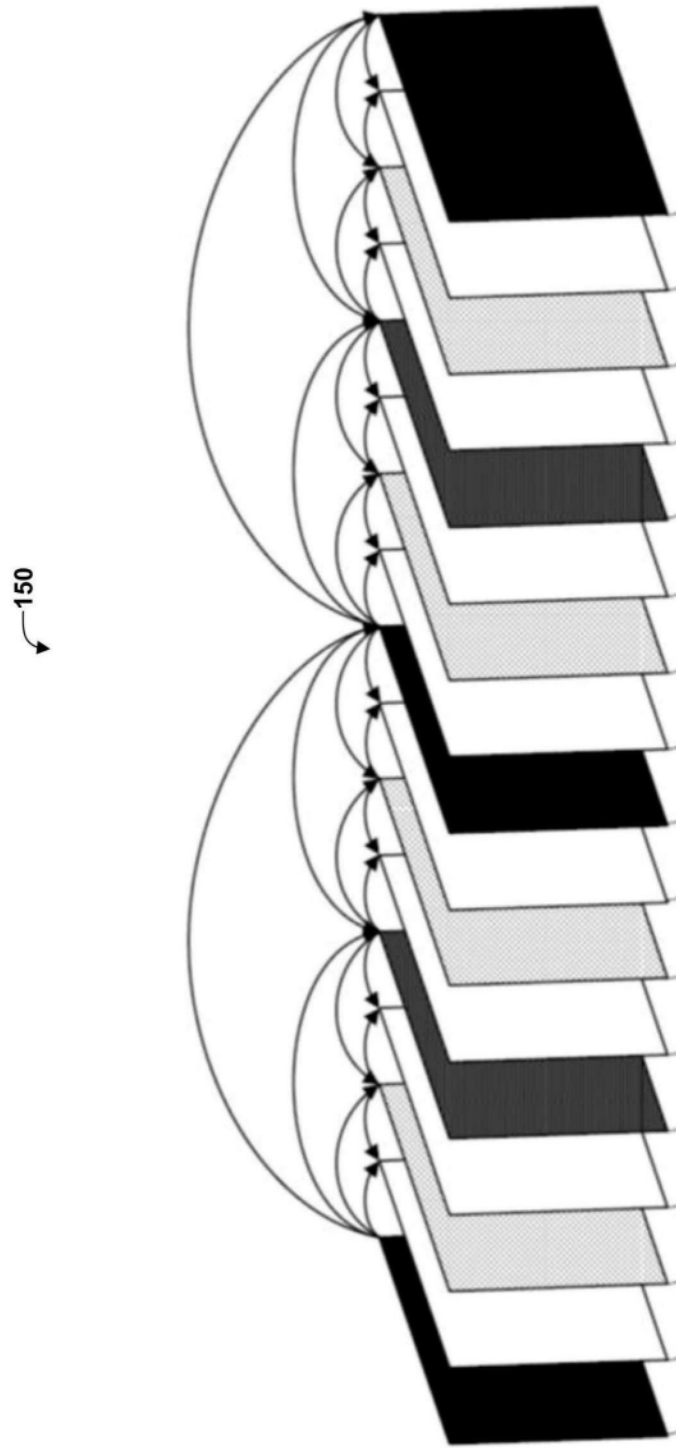


图4



图5

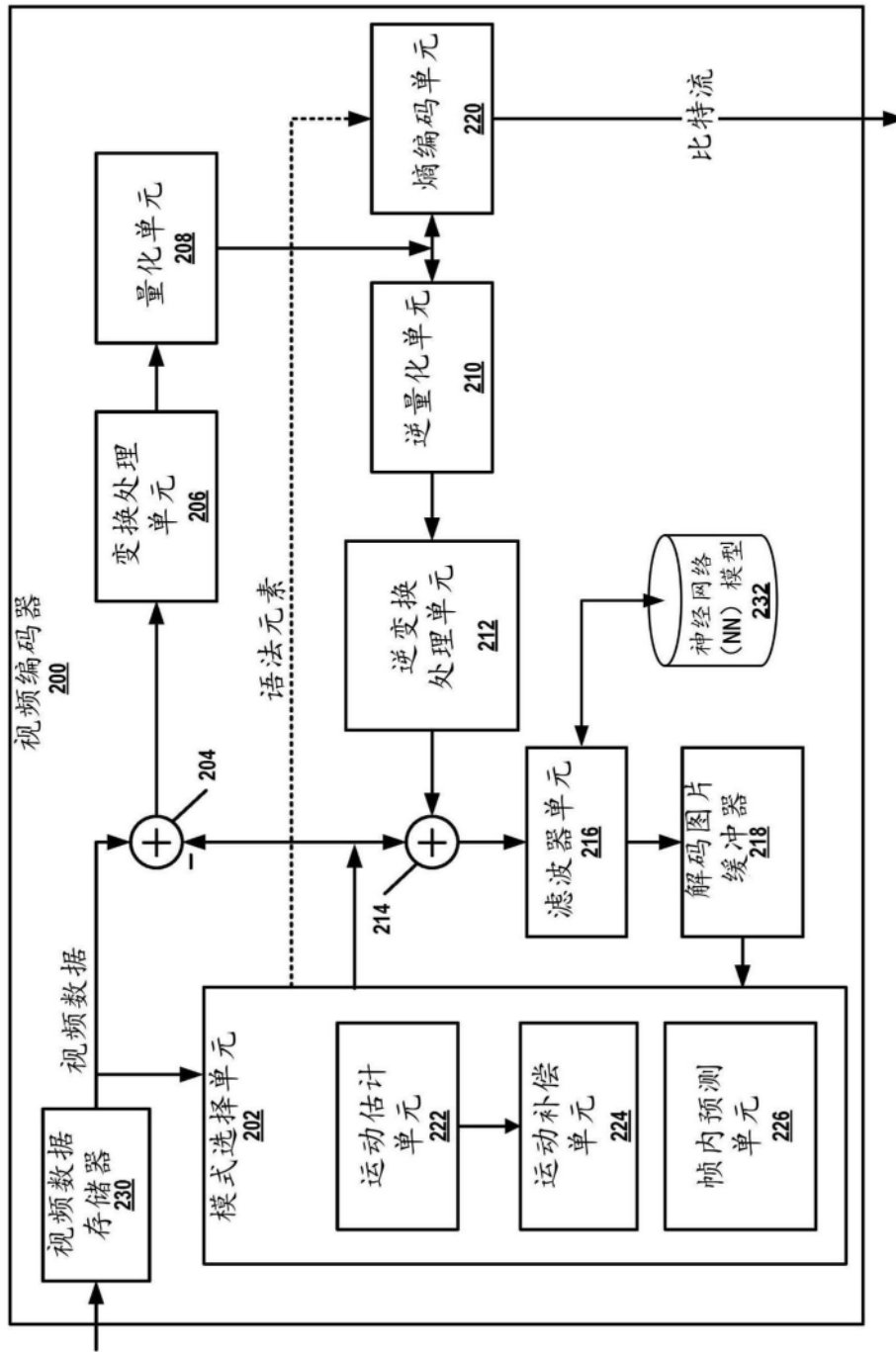


图6

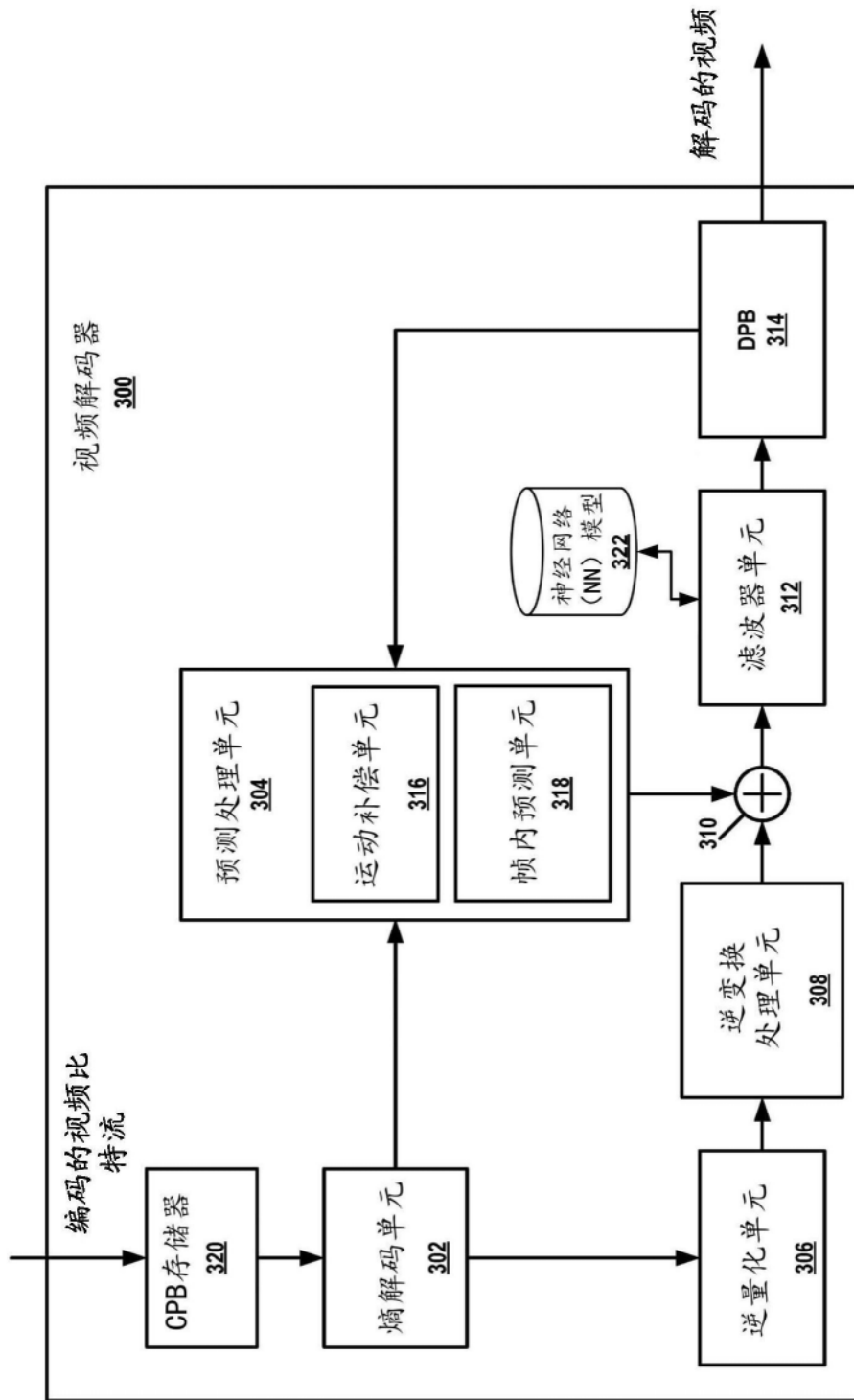


图7

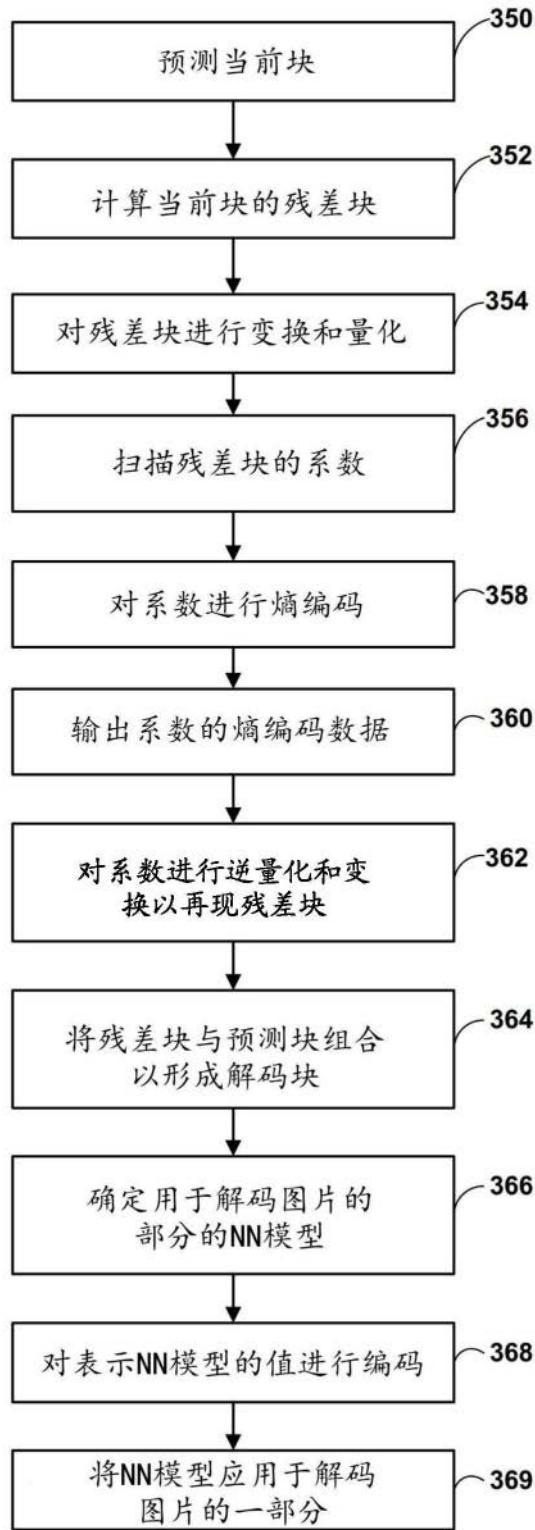


图8

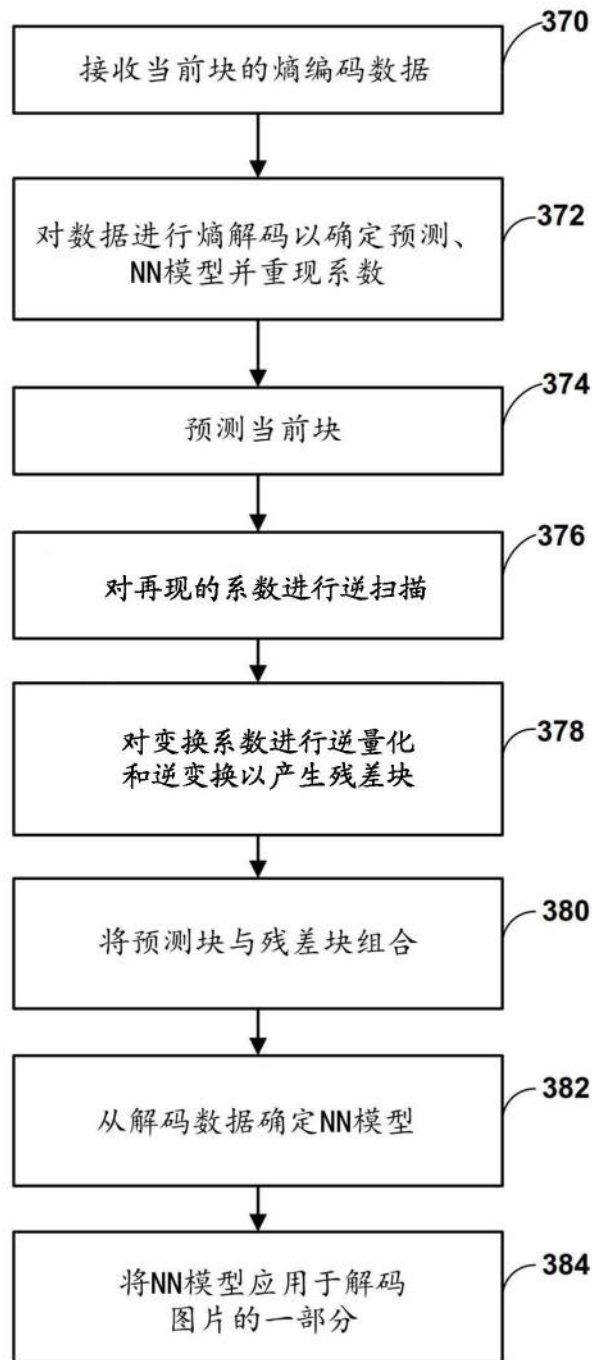


图9