



(12) 发明专利

(10) 授权公告号 CN 108141597 B

(45) 授权公告日 2022. 03. 08

(21) 申请号 201680056197.8

(22) 申请日 2016.09.21

(65) 同一申请的已公布的文献号
申请公布号 CN 108141597 A

(43) 申请公布日 2018.06.08

(30) 优先权数据
62/234,644 2015.09.29 US
62/294,897 2016.02.12 US
62/323,496 2016.04.15 US
62/337,736 2016.05.17 US
15/270,507 2016.09.20 US

(85) PCT国际申请进入国家阶段日
2018.03.27

(86) PCT国际申请的申请数据
PCT/US2016/052935 2016.09.21

(87) PCT国际申请的公布数据
W02017/058615 EN 2017.04.06

(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 赵欣 陈建乐 阿米尔·赛义德
马尔塔·卡切维奇
瓦迪姆·谢廖金

(74) 专利代理机构 北京市柳沈律师事务所
11105
代理人 安之斐

(51) Int.Cl.
H04N 19/159 (2006.01)
H04N 19/176 (2006.01)
H04N 19/61 (2006.01)
H04N 19/12 (2006.01)
H04N 19/18 (2006.01)

(56) 对比文件
CN 102474268 A, 2012.05.23
CN 102986215 A, 2013.03.20
WO 2012014461 A1, 2012.02.02
审查员 于典

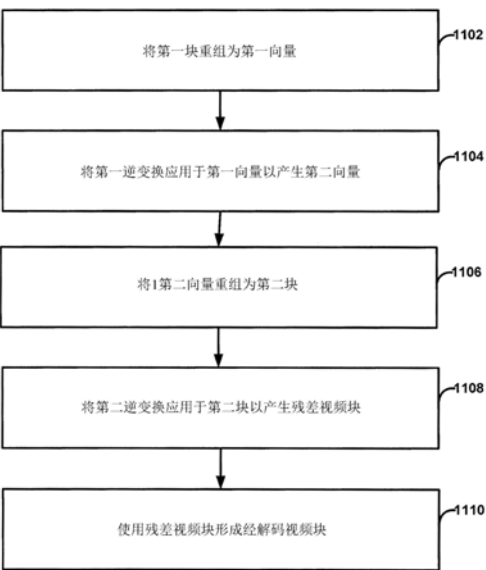
权利要求书6页 说明书40页 附图17页

(54) 发明名称

用于具有重组的视频译码的非可分离二次变换

(57) 摘要

本发明描述其中解码器经配置以根据系数扫描次序将第一2维系数块重组为第一1维系数向量并将第一逆变换应用于所述第一1维系数向量以产生第二1维系数向量的技术。所述第一逆变换为非可分离变换。所述解码器经进一步配置以将所述第一1维系数向量重组为第二2维系数块并将第二逆变换应用于所述第二2维系数块以产生残差视频块。所述第二逆变换将所述第二2维系数块从频域转换到像素域。所述解码器经进一步配置以形成经解码视频块,其中形成所述经解码视频块包括将所述残差视频块与一或多个预测性块求和。



1. 一种用于解码视频数据的方法,所述方法包括:

根据系数扫描次序将第一2维系数块重组为第一1维系数向量,其中所述第一2维系数块为第一 4×4 系数群,其中所述第一1维系数向量为第一 16×1 系数向量,并且其中将所述第一2维系数块重组为所述第一1维系数向量包括将所述第一 4×4 系数群重组为所述第一 16×1 系数向量;

将第一逆变换应用于所述第一1维系数向量以产生第二1维系数向量,所述第一逆变换为非可分离变换;

基于帧内预测模式,将所述第二1维系数向量重组为第二2维系数块,其中所述第二1维系数向量是第二 16×1 系数向量,其中所述第二2维系数块是第二 4×4 系数群,并且其中将所述第二1维系数向量重组为第二2维系数块包括将所述第二 16×1 系数向量重组为所述第二 4×4 系数群;

将第二逆变换应用于所述第二2维系数块以产生残差视频块,所述第二逆变换将所述第二2维系数块从频域转换到像素域;

通过基于所述帧内预测模式和来自当前图片的一或多个先前解码的块的数据执行帧内预测,来生成一或多个预测性块;及

形成所述当前图片的经解码视频块,其中形成所述经解码视频块包括将所述残差视频块与所述一或多个预测性块求和。

2. 根据权利要求1所述的方法,其中:

所述方法进一步包括确定用于所述第一 4×4 系数群的所述系数扫描次序,及

将所述第一2维系数块重组为所述第一1维系数向量包括基于所述经确定系数扫描次序将所述第一2维系数块重组为所述第一1维系数向量。

3. 根据权利要求1所述的方法,包括:

根据所述系数扫描次序重组所述第一2维系数块,其中在所述第一2维系数块中的具有较小索引值的系数放置在所述第一1维系数向量中的较小扫描索引位置处。

4. 根据权利要求1所述的方法,其中将所述第一逆变换应用于所述第一1维系数向量是基于所述第二2维系数块的非零系数的量不大于阈值。

5. 根据权利要求4所述的方法,其中所述阈值针对所述视频数据的不同块大小、不同帧内预测模式或其一组合是不同的。

6. 根据权利要求1所述的方法,其进一步包括:

基于共享所述第一逆变换的索引的多个色彩分量的非零系数的总数的和、总绝对和或平方值的和不大于是于阈值而确定所述第一逆变换的所述索引是在所述视频数据的位流中传信;

解码包括所述第一逆变换的所述索引的块层级语法元素,其中所述索引用于所述多个色彩分量中的每一者;及

将所述第一逆变换应用于所述第一1维系数向量包括基于所述第一逆变换的所述索引对所述第一1维系数向量执行所述第一逆变换。

7. 根据权利要求1所述的方法,其进一步包括:

基于在变换跳过模式中译码变换单元而确定所述第一逆变换的索引是在所述视频数据的位流中传信;以及

从所述位流解码所述第一逆变换的所述索引，

其中将所述第一逆变换应用于所述第一1维系数向量包括基于所述第一逆变换的所述索引对所述第一1维系数向量执行所述第一逆变换。

8. 根据权利要求1所述的方法，其进一步包括：

将 $M \times N$ 个系数块分成多个非重叠 4×4 子块，所述多个非重叠 4×4 子块包含所述第一2维系数块，其中 M 及 N 各自大于4， M 及 N 各自为4的倍数，且 M 不等于 N ；及

针对所述多个非重叠 4×4 子块中的每一 4×4 子块执行所述第一逆变换以产生多个2维系数块，所述多个2维系数块包含所述第二2维系数块。

9. 根据权利要求1所述的方法，其中应用所述第一逆变换包括将2维变换矩阵乘以所述第一1维系数向量。

10. 根据权利要求9所述的方法，其中所述2维变换矩阵为通过明度帧内预测模式及译码单元层级索引选择的 16×16 非可分离变换矩阵。

11. 根据权利要求1所述的方法，其进一步包括：

逆量化所述第一2维系数块。

12. 根据权利要求1所述的方法，其中将所述第二1维系数向量重组为第二2维系数块基于选择垂直明度预测模式还是水平明度预测模式。

13. 一种用于编码视频数据的方法，所述方法包括：

通过基于帧内预测模式和来自当前图片的一或多个先前编码的块的数据执行帧内预测，来生成一或多个预测性块；

形成所述当前图片的残差视频块，其中形成所述残差视频块包括从正被编码的视频块减去所述一或多个预测性块；

将第一变换应用于所述残差视频块的至少部分以产生第一2维系数块，所述第一变换将所述残差视频块从像素域转换到频域；

基于所述帧内预测模式，将所述第一2维系数块重组为第一1维系数向量，其中所述第一2维系数块为第一 4×4 系数群，其中所述第一1维系数向量为第一 16×1 系数向量，并且其中将所述第一2维系数块重组为所述第一1维系数向量包括将所述第一 4×4 系数群重组为所述第一 16×1 系数向量；

将第二变换应用于所述第一1维系数向量的至少部分以产生第二1维系数向量，所述第二变换为非可分离变换；及

根据系数扫描次序将所述第二1维系数向量重组为第二2维系数块，其中所述第二1维系数向量是第二 16×1 系数向量，其中所述第二2维系数块是第二 4×4 系数群，并且其中将所述第二1维系数向量重组为第二2维系数块包括将所述第二 16×1 系数向量重组为所述第二 4×4 系数群。

14. 一种用于解码视频数据的设备，其包括：

存储器，其经配置以存储所述视频数据；及

视频解码器，其经配置以：

根据系数扫描次序将第一2维系数块重组为第一1维系数向量，其中所述第一2维系数块为第一 4×4 系数群，其中所述第一1维系数向量为第一 16×1 系数向量，并且其中为了将所述第一2维系数块重组为所述第一1维系数向量，所述视频解码器被配置为将所述第一4

$\times 4$ 系数群重组为所述第一 16×1 系数向量;

将第一逆变换应用于所述第一1维系数向量以产生第二1维系数向量,所述第一逆变换为非可分离变换;

基于帧内预测模式,将所述第二1维系数向量重组为第二2维系数块,其中所述第二1维系数向量是第二 16×1 系数向量,其中所述第二2维系数块是第二 4×4 系数群,并且其中为了将所述第二1维系数向量重组为第二2维系数块,所述视频解码器被配置为将所述第二 16×1 系数向量重组为所述第二 4×4 系数群;

将第二逆变换应用于所述第二2维系数块以产生残差视频块,所述第二逆变换将所述第二2维系数块从频域转换到像素域;

通过基于所述帧内预测模式和来自当前图片的一或多个先前解码的块的数据执行帧内预测,来生成一或多个预测性块;

将所述残差视频块与所述一或多个预测性块求和;及

基于所述残差视频块与所述一或多个预测性块的所述和形成所述当前图片的经解码视频块。

15. 根据权利要求14所述的设备,其中所述视频解码器经进一步配置以:

确定用于所述第一 4×4 系数群的系数扫描次序;及

其中,为了将所述第一2维系数块重组为所述第一1维系数向量,所述视频解码器被配置为基于所述经确定系数扫描次序将所述第一2维系数块重组为所述第一1维系数向量。

16. 根据权利要求14所述的设备,其中,为了重组所述第一2维系数块,所述视频解码器经进一步配置以:

根据所述系数扫描次序重组所述第一2维系数块,其中在所述第一2维系数块中的具有较小索引值的系数放置在所述第一1维系数向量中的较小扫描索引位置处。

17. 根据权利要求14所述的设备,其中所述视频解码器经进一步配置以:

基于所述第二2维系数块的非零系数的量不大于阈值而将所述第一逆变换应用于所述第一1维系数向量。

18. 根据权利要求17所述的设备,其中所述阈值针对所述视频数据的不同块大小、不同帧内预测模式或其一组合而不同。

19. 根据权利要求14所述的设备,其中所述视频解码器经进一步配置以:

基于共享所述第一逆变换的索引的多个色彩分量的非零系数的总数的和、总绝对和或平方值的和不大于阈值而确定所述第一逆变换的索引是在所述视频数据的位流中传信;

解码包括所述第一逆变换的所述索引的块层级语法元素,其中所述索引用于所述多个色彩分量中的每一者;及

基于所述第一逆变换的所述索引将所述第一逆变换应用于所述第一1维系数向量。

20. 根据权利要求14所述的设备,其中所述视频解码器经进一步配置以:

基于变换单元是在变换跳过模式中译码而确定所述第一逆变换的索引是在所述视频数据的位流中传信;

从所述位流解码所述第一逆变换的所述索引;及

将所述第一逆变换应用于所述第一1维系数向量,包括基于所述第一逆变换的所述索引对所述第一1维系数向量执行所述第一逆变换。

21. 根据权利要求14所述的设备,其中所述视频解码器经进一步配置以:

将 $M \times N$ 系数块分成多个非重叠 4×4 子块,所述多个非重叠 4×4 子块包含所述第一2维系数块, M 及 N 各自大于4, M 及 N 各自为4的倍数,且 M 不等于 N ;及

对于所述多个非重叠 4×4 子块中的每一 4×4 子块执行所述第一逆变换以产生多个2维系数块,所述多个2维系数块包含所述第二2维系数块。

22. 根据权利要求14所述的设备,其中应用所述第一逆变换包括将2维变换矩阵乘以所述第一1维系数向量。

23. 根据权利要求22所述的设备,其中所述2维变换矩阵为通过明度帧内预测模式及译码单元层级索引选择的 16×16 非可分离变换矩阵。

24. 根据权利要求14所述的设备,其中所述视频解码器经进一步配置以:

逆量化所述第一2维系数块。

25. 根据权利要求14所述的设备,其中所述设备包括经配置以显示包括所述经解码视频块的图片的显示器。

26. 根据权利要求14所述的设备,其中所述设备包括摄像机、计算机、移动装置、广播接收器装置或机顶盒中的一或多个者。

27. 一种用于编码视频数据的设备,其包括:

存储器,其经配置以存储所述视频数据;及

视频编码器,其经配置以:

通过基于帧内预测模式和来自当前图片的一或多个先前编码的块的数据执行帧内预测,来生成一或多个预测性块;

形成所述当前图片的残差视频块,其中形成所述残差视频块包括从正被编码的视频块减去所述一或多个预测性块;

将第一变换应用于所述残差视频块的至少部分以产生第一2维系数块,所述第一变换将所述残差视频块从像素域转换到频域;

基于所述帧内预测模式,重组所述第一2维系数块以产生第一1维系数向量,其中所述第一2维系数块为第一 4×4 系数群,其中所述第一1维系数向量为第一 16×1 系数向量,并且其中为了将所述第一2维系数块重组为所述第一1维系数向量,所述视频编码器被配置为将所述第一 4×4 系数群重组为所述第一 16×1 系数向量;

将第二变换应用于所述第一1维系数向量的至少部分以产生第二1维系数向量,所述第二变换为非可分离变换;及

根据系数扫描次序将所述第二1维系数向量重组为第二2维系数块,其中所述第二1维系数向量是第二 16×1 系数向量,其中所述第二2维系数块是第二 4×4 系数群,并且其中为了重组所述第二1维系数向量,所述视频编码器被配置为将所述第二 16×1 系数向量重组为所述第二 4×4 系数群。

28. 一种用于解码视频数据的设备,其包括:

用于根据系数扫描次序将第一2维系数块重组为第一1维系数向量的装置,其中所述第一2维系数块为第一 4×4 系数群,其中所述第一1维系数向量为第一 16×1 系数向量,并且其中用于重组所述第一2维系数块的装置包括用于将所述第一 4×4 系数群重组为所述第一 16×1 系数向量的装置;

用于将第一逆变换应用于所述第一1维系数向量以产生第二1维系数向量的装置,所述第一逆变换为非可分离变换;

用于基于帧内预测模式,将所述第二1维系数向量重组为第二2维系数块的装置,其中所述第二1维系数向量是第二 16×1 系数向量,其中所述第二2维系数块是第二 4×4 系数群,并且其中用于将所述第二1维系数向量重组为第二2维系数块的装置包括用于将所述第二 16×1 系数向量重组为所述第二 4×4 系数群的装置;

用于将第二逆变换应用于所述第二2维系数块以产生残差视频块的装置,所述第二逆变换将所述第二2维系数块从频域转换到像素域;

用于通过基于所述帧内预测模式和来自当前图片的一或多个先前解码的块的数据执行帧内预测,来生成一或多个预测性块的装置;及

用于形成所述当前图片的经解码视频块的装置,其中用于形成所述经解码视频块的所述装置包括用于将所述残差视频块与所述一或多个预测性块求和的装置。

29. 一种用于编码视频数据的设备,其包括:

用于通过基于帧内预测模式和来自当前图片的一或多个先前编码的块的数据执行帧内预测,来生成一或多个预测性块的装置;

用于形成所述当前图片的残差视频块的装置,其中形成所述残差视频块包括从正被编码的视频块减去所述一或多个预测性块;

用于将第一变换应用于所述残差视频块的至少部分以产生第一2维系数块的装置,所述第一变换将所述残差视频块从像素域转换到频域;

用于基于所述帧内预测模式,重组所述第一2维系数块以产生第一1维系数向量的装置,其中所述第一2维系数块为第一 4×4 系数群,其中所述第一1维系数向量为第一 16×1 系数向量,并且其中用于将所述第一2维系数块重组为所述第一1维系数向量的装置包括用于将所述第一 4×4 系数群重组为所述第一 16×1 系数向量的装置;

用于将第二变换应用于所述第一1维系数向量的至少部分以产生第二1维系数向量的装置,所述第二变换为非可分离变换;及

用于根据系数扫描次序将所述第二1维系数向量重组为第二2维系数块的装置,其中所述第二1维系数向量是第二 16×1 系数向量,其中所述第二2维系数块是第二 4×4 系数群,并且其中用于重组所述第二1维系数向量的装置包括用于将所述第二 16×1 系数向量重组为所述第二 4×4 系数群的装置。

30. 一种存储指令的非暂时性计算机可读存储媒体,所述指令在执行时使得用于解码视频数据的装置的一或多个处理器进行以下操作:

根据系数扫描次序将第一2维系数块重组为第一1维系数向量,其中所述第一2维系数块为第一 4×4 系数群,其中所述第一1维系数向量为第一 16×1 系数向量,并且其中为了重组所述第一2维系数块重组,所述指令还使得所述一或多个处理器将所述第一 4×4 系数群重组为所述第一 16×1 系数向量;

将第一逆变换应用于所述第一1维系数向量以产生第二1维系数向量,所述第一逆变换为非可分离变换;

基于帧内预测模式,将所述第二1维系数向量重组为第二2维系数块,其中所述第二1维系数向量是第二 16×1 系数向量,其中所述第二2维系数块是第二 4×4 系数群,并且其中为

了将所述第二1维系数向量重组为第二2维系数块,所述指令还使得所述一或多个处理器将所述第二 16×1 系数向量重组为所述第二 4×4 系数群;

将第二逆变换应用于所述第二2维系数块以产生残差视频块,所述第二逆变换将所述第二2维系数块从频域转换到像素域;

通过基于所述帧内预测模式和来自当前图片的一或多个先前解码的块的数据执行帧内预测,来生成一或多个预测性块;及

形成所述当前图片的经解码视频块,其中形成所述经解码视频块包括将所述残差视频块与所述一或多个预测性块求和。

用于具有重组的视频译码的非可分离二次变换

[0001] 相关申请案

[0002] 本申请案主张2015年9月29日申请的第62/234,644号美国临时申请案,2016年2月12日申请的第62/294,897号美国临时申请案,2016年4月15日申请的第62/323,496号美国临时申请案,及2016年5月17日申请的第62/337,736号美国临时申请案的权利,所述案中的每一者的全部内容以引用的方式并入本文中。

技术领域

[0003] 本发明涉及视频译码。

背景技术

[0004] 数字视频能力可并入到广泛范围的装置中,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)、膝上型或台式计算机、平板计算机、电子书阅读器、数字摄像机、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“智能型电话”)、视频电话会议装置、视频流装置及其类似者。数字视频装置实施视频译码技术,例如包含由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、ITU-T H.264/MPEG-4高级视频译码(AVC)第10部分、及ITU-T H.265高效率视频译码(HEVC)所界定的各种标准及这些标准的扩展中所描述的那些技术。视频装置可通过实施这些视频译码技术来更有效率地发射、接收、编码、解码及/或存储数字视频信息。

[0005] 视频译码技术包含空间(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,可将视频切片(例如,视频帧或视频帧的一部分)分割成视频块(其也可被称作树型块)、译码单元(CU)及/或译码节点。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来编码图片的经帧内译码(I)切片中的视频块。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。图片可被称作帧,且参考图片可被称作参考帧。

[0006] 空间或时间预测导致待译码块的预测性块。残差数据表示待译码的原始块与预测性块之间的像素差。根据指向形成预测性块的参考样本的块的运动向量及指示经译码块与预测性块之间的差的残差数据来编码经帧间译码块。根据帧内译码模式及残差数据编码经帧内译码块。为进行进一步压缩,可将残差数据从像素域变换到变换域,从而产生残差变换系数,可随后量化所述残差变换系数。可扫描最初布置成二维阵列的经量化的变换系数以便产生变换系数的一维向量,且可应用熵译码以达成甚至更多压缩。

发明内容

[0007] 一般来说,本发明描述视频译码中的与应用于帧内或帧间预测残差数据的二次变换设计相关的技术。变换可用于高级视频编码解码器的情形中,例如HEVC的扩展或下一代

视频译码标准。

[0008] 在一个实例中,本发明描述一种方法,所述方法包括:逆量化第一系数块;将第一逆变换应用于经逆量化第一系数块的至少部分以产生第二系数块,第一逆变换为非可分离变换;将第二逆变换应用于第二系数块以产生残差视频块,第二逆变换将第二系数块从频域转换到像素域;及形成经解码视频块,其中形成经解码视频块包括将残差视频块与一或多个预测性块求和。

[0009] 在一个实例中,本发明描述一种方法,所述方法包括:形成残差视频块,其中形成残差视频块包括从正被译码的视频块减去一或多个预测性块;将第一变换应用于残差视频块以产生第一系数块,第一变换将残差视频块从像素域转换到频域;将第二变换应用于第一系数块的至少部分以产生第二系数块,第二变换为非可分离变换;及量化第二系数块。

[0010] 在一个实例中,本发明描述一种用于解码视频数据的设备,所述设备包括:存储器,其经配置以存储视频数据;及视频解码器,其经配置以:逆量化第一系数块;将第一逆变换应用于经逆量化第一系数块的至少部分以产生第二系数块,第一逆变换为非可分离变换;将第二逆变换应用于第二系数块以产生残差视频块,第二逆变换将第二系数块从频域转换到像素域;将残差视频块与一或多个预测性块求和;及基于残差视频块与一或多个预测性块的和而形成经解码视频块。

[0011] 在一个实例中,本发明描述一种用于编码视频数据的设备,所述设备包括:存储器,其经配置以存储视频数据;及视频编码器,其经配置以:形成残差视频块,其中形成残差视频块包括从正被译码的视频块减去一或多个预测性块;将第一变换应用于残差视频块以产生第一系数块,第一变换将残差视频块从像素域转换到频域;将第二变换应用于第一系数块的至少部分以产生第二系数块,第二变换为非可分离变换;及量化第二系数块。

[0012] 在一个实例中,本发明描述一种用于解码视频数据的设备,所述设备包括:用于逆量化第一系数块的装置;用于将第一逆变换应用于经逆量化第一系数块的至少部分以产生第二系数块的装置,第一逆变换为非可分离变换;用于将第二逆变换应用于第二系数块以产生残差视频块的装置,第二逆变换将第二系数块从频域转换到像素域;及用于形成经解码视频块的装置,其中用于形成经解码视频块的所述装置包括用于将残差视频块与一或多个预测性块求和的装置。

[0013] 在一个实例中,本发明描述一种用于编码视频数据的设备,所述设备包括:用于形成残差视频块的装置,其中形成残差视频块包括从正被译码的视频块减去一或多个预测性块;用于将第一变换应用于残差视频块以产生第一系数块的装置,第一变换将残差视频块从像素域转换到频域;用于将第二变换应用于第一系数块的至少部分以产生第二系数块的装置,第二变换为非可分离变换;及用于量化第二系数块的装置。

[0014] 在一个实例中,本发明描述一种存储指令的非暂时性计算机可读存储媒体,所述指令当经执行时使得用于解码视频数据的装置的一个或多个处理器执行以下操作:逆量化第一系数块;将第一逆变换应用于经逆量化第一系数块的至少部分以产生第二系数块,第一逆变换为非可分离变换;将第二逆变换应用于第二系数块以产生残差视频块,第二逆变换将第二系数块从频域转换到像素域;将残差视频块与一或多个预测性块求和;及基于残差视频块与一或多个预测性块的和而形成经解码视频块。

[0015] 在一个实例中,本发明描述一种存储指令的非暂时性计算机可读存储媒体,所述

指令当经执行时使得用于编码视频数据的装置的一个或多个处理器执行以下操作:形成残差视频块,其中形成残差视频块包括从正被译码的视频块减去一或多个预测性块;将第一变换应用于残差视频块以产生第一系数块,第一变换将残差视频块从像素域转换到频域;将第二变换应用于第一系数块的至少部分以产生第二系数块,第二变换为非可分离变换;及量化第二系数块。

[0016] 在一个实例中,本发明描述一种用于解码视频数据的方法,所述方法包括:根据系数扫描次序将第一2维系数块重组为第一1维系数向量;将第一逆变换应用于第一1维系数向量以产生第二1维系数向量,第一逆变换为非可分离变换;将第一1维系数向量重组为第二2维系数块;将第二逆变换应用于第二2维系数块以产生残差视频块,第二逆变换将第二2维系数块从频域转换到像素域;及形成经解码视频块,其中形成经解码视频块包括将残差视频块与一或多个预测性块求和。

[0017] 在一个实例中,本发明描述一种用于编码视频数据的方法,所述方法包括:形成残差视频块,其中形成残差视频块包括从正被译码的视频块减去一或多个预测性块;将第一变换应用于残差视频块的至少部分以产生第一2维系数块,第一变换将残差视频块从像素域转换到频域;将第一2维系数块重组为第一1维系数向量;将第二变换应用于第一1维系数向量的至少部分以产生第二1维系数向量,第二变换为非可分离变换;及根据系数扫描次序将第二1维系数向量重组为第二2维系数块。

[0018] 在一个实例中,本发明描述一种用于解码视频数据的设备,所述设备包括:存储器,其经配置以存储视频数据;及视频解码器,其经配置以:根据系数扫描次序将第一2维系数块重组为第一1维系数向量;将第一逆变换应用于第一1维系数向量以产生第二1维系数向量,第一逆变换为非可分离变换;将第一1维系数向量重组为第二2维系数块;将第二逆变换应用于第二2维系数块以产生残差视频块,第二逆变换将第二2维系数块从频域转换到像素域;将残差视频块与一或多个预测性块求和;将残差视频块与一或多个预测性块求和;及基于残差视频块与一或多个预测性块的和而形成经解码视频块。

[0019] 在一个实例中,本发明描述一种用于编码视频数据的设备,所述设备包括:存储器,其经配置以存储视频数据;及视频编码器,其经配置以:形成残差视频块,其中形成残差视频块包括从正被译码的视频块减去一或多个预测性块;将第一变换应用于残差视频块的至少部分以产生第一2维系数块,第一变换将残差视频块从像素域转换到频域;重组第一2维系数块以产生第一1维系数向量;将第二变换应用于第一1维系数向量的至少部分以产生第二1维系数向量,第二变换为非可分离变换;及根据系数扫描次序将第二1维系数向量重组为第二2维系数块。

[0020] 在一个实例中,本发明描述一种用于解码视频数据的设备,所述设备包括:用于根据系数扫描次序将第一2维系数块重组为第一1维系数向量的装置;用于将第一逆变换应用于第一1维系数向量以产生第二1维系数向量的装置,第一逆变换为非可分离变换;用于将第一1维系数向量重组为第二2维系数块的装置;用于将第二逆变换应用于第二2维系数块以产生残差视频块的装置,第二逆变换将第二2维系数块从频域转换到像素域;及用于形成经解码视频块的装置,其中用于形成经解码视频块的所述装置包括用于将残差视频块与一或多个预测性块求和的装置。

[0021] 在一个实例中,本发明描述一种用于编码视频数据的设备,所述设备包括:用于形

成残差视频块的装置,其中用于形成残差视频块的所述装置包括用于从正被译码的视频块减去一或多个预测性块的装置;用于将第一变换应用于残差视频块的至少部分以产生第一2维系数块的装置,第一变换将残差视频块从像素域转换到频域;用于重组第一2维系数块以产生第一1维系数向量的装置;用于将第二变换应用于第一1维系数向量的至少部分以产生第二1维系数向量的装置,第二变换为非可分离变换;及用于根据系数扫描次序将第二1维系数向量重组为第二2维系数块的装置。

[0022] 在一个实例中,本发明描述一种存储指令的非暂时性计算机可读存储媒体,所述指令当经执行时使得用于解码视频数据的装置的一个或多个处理器执行以下操作:根据系数扫描次序将第一2维系数块重组为第一1维系数向量;将第一逆变换应用于第一1维系数向量以产生第二1维系数向量,第一逆变换为非可分离变换;将第一1维系数向量重组为第二2维系数块;将第二逆变换应用于第二2维系数块以产生残差视频块,第二逆变换将第二2维系数块从频域转换到像素域;及形成经解码视频块,其中形成经解码视频块包括将残差视频块与一或多个预测性块求和。

[0023] 在一个实例中,本发明描述一种存储指令的非暂时性计算机可读存储媒体,所述指令当经执行时使得用于编码视频数据的装置的一个或多个处理器执行以下操作:形成残差视频块,包括从正被译码的视频块减去一或多个预测性块;将第一变换应用于残差视频块的至少部分以产生第一2维系数块,第一变换将残差视频块从像素域转换到频域;重组第一2维系数块以产生第一1维系数向量;将第二变换应用于第一1维系数向量的至少部分以产生第二1维系数向量,第二变换为非可分离变换;及根据系数扫描次序将第二1维系数向量重组为第二2维系数块。

[0024] 在以下随附图式及描述中阐述一或多个实例的细节。其它特征、目标及优势将从描述、图式及权利要求书显而易见。

附图说明

[0025] 图1为说明可利用本发明中所描述的技术的实例视频编码及解码系统的框图。

[0026] 图2为可利用本发明中所描述的技术的基于HEVC中的残差四分树的变换方案的说明。

[0027] 图3A为视频编码器处的变换过程的说明。

[0028] 图3B为视频解码器处的变换过程的说明。

[0029] 图4为通过视频解码器应用的实例旋转变换的说明。

[0030] 图5为说明包含二次变换系数重组过程的实例编码过程的框图。

[0031] 图6为说明包含二次变换系数重组过程的实例解码过程的框图。

[0032] 图7A为从明度帧内模式到变换集合索引的实例映射的说明。

[0033] 图7B为从0到66的帧内预测模式(IPM)范围的说明。

[0034] 图8为说明可实施本发明中所描述的技术的视频编码器的实例的框图。

[0035] 图9为说明可实施本发明中所描述的技术的视频解码器的实例的框图。

[0036] 图10A为根据本发明中所描述的一或多个技术的通过视频编码器进行的实例变换处理的说明。

[0037] 图10B为根据本发明中所描述的一或多个技术的通过视频解码器进行的实例逆变

换处理的说明。

[0038] 图11为说明可实施本发明中所描述的技术的视频数据的第一实例解码的流程图。

[0039] 图12为说明可实施本发明中所描述的技术的视频数据的第一实例编码的流程图。

[0040] 图13为说明可实施本发明中所描述的技术的视频数据的第二实例解码的流程图。

[0041] 图14为说明可实施本发明中所描述的技术的视频数据的第二实例编码的流程图。

具体实施方式

[0042] 一般来说,本发明描述视频译码中的与应用于帧内或帧间预测残差数据的二次变换设计相关的技术。变换可用于高级视频编码解码器的情形中,例如HEVC的扩展或下一代视频译码标准。

[0043] 视频编码器可以适合于从视频编码器输出到视频解码器的形式表示视频数据的残差块。需要减少用以表示残差块的数据量以使得减少从视频编码器发射到视频解码器的数据量。在视频译码中,相对于非可分离变换更多地应用可分离变换,这是由于可分离变换与非可分离变换相比可使用更少的运算(例如,加法、乘法)。可分离变换可经写成两个或大于两个滤波的乘积的滤波。对比来说,非可分离滤波不可写成两个或大于两个滤波的乘积。

[0044] 视频编码器也可应用二次变换以增加系数块的能量精简性,而非仅依赖于将像素域中的残差块转换到频域中的系数块的变换。举例来说,二次变换可将系数块的非零系数集中为更接近系数块的DC系数。因此,在系数块的DC系数与系数块的最后有效(即,非零)系数之间可存在较少系数,从而导致用以表示残差块的数据的减少数量。类似地,视频解码器可应用逆一次变换以将系数块变换成残差块。以此方式,可减少用以表示残差块的数据,借此减少针对视频数据的频宽及/或存储要求,且可能减少视频解码器及视频编码器的能量使用。

[0045] 现有二次变换设计将二次变换应用为可分离变换。在一些情况下,已应用可分离变换而不是非可分离变换,这是由于可分离变换相较于非可分离变换具有较低复杂度。因此,对于视频译码方案来说,可分离变换可能已变得比非可分离变换更有利。然而,非可分离变换用作二次变换可提供改进的变换效率。举例来说,与可分离二次变换相比,装置可能更快或以更少复杂度应用非可分离二次变换。因此,根据本发明的实例,视频解码器可将第一逆变换应用于经逆量化第一系数块的至少部分以产生第二系数块,第一逆变换为非可分离变换。另外,在此实例中,视频解码器可将第二逆变换应用于第二系数块以产生残差视频块。第二逆变换可将第二系数块从频域转换到像素域。在此实例中,视频解码器可形成经解码视频块。在一些实例中,作为形成经解码视频块的部分,视频解码器可将残差视频块与一或多个预测性块求和。

[0046] 视频数据压缩(例如,不同帧内预测模式)的差异可改变非可分离二次变换的有效性。举例来说,当使用第一帧内预测模式时,第一非可分离二次变换与第二非可分离二次变换相比可具有更高能量精简性。然而,在实例中,当使用第二帧内预测模式时,第二非可分离二次变换与第一非可分离二次变换相比可具有更高能量精简性。因此,可能需要准许从候选非可分离二次变换的集合中选择候选非可分离二次变换的技术。即,视频编码器可从非可分离二次变换的集合中选择候选非可分离二次变换,而非使用相同非可分离二次变换用于不同视频数据。举例来说,视频编码器可基于帧内预测模式调配非可分离二次变换的

子集。类似地,在一些实例中,视频解码器可从非可分离二次变换的集合中选择候选非可分离二次变换。举例来说,视频解码器可基于视频数据的已经解码信息调配非可分离二次变换的子集。以此方式,更合适的非可分离二次变换可经选择以应用于残差块,与使用仅一个非可分离二次变换候选者的实例相比较,此可导致用以表示残差块的数据量减少。

[0047] 在一些实例中,视频编码器可将变换系数重组成二维系数块,而非必定以向量格式量化变换系数。更特定来说,在一些情况下,量化单元可经配置以量化二维系数块,二维系数块可为通过可分离变换输出的格式。因此,相同量化单元可用于其中视频编码器仅仅使用可分离变换的情形,以及其中视频编码器使用可分离变换及非可分离变换的情形中。

[0048] 图1为说明可利用本发明中所描述的技术的实例视频编码及解码系统10的框图。如图1中所展示,系统10包含源装置12,其提供待在稍后时间由目的地装置14解码的经编码视频数据。详细来说,源装置12经由计算机可读媒体16将视频数据提供到目的地装置14。源装置12及目的地装置14可包括广泛范围的装置中的任一者,包含台式计算机、笔记型(即,膝上型)计算机、平板计算机、机顶盒、例如所谓的“智能型”电话的电话手机、所谓的“智能型”板、电视、摄像机、显示装置、数字媒体播放器、视频游戏控制台、视频流装置或类似者。在一些情况下,可装备源装置12及目的地装置14以用于无线通信。

[0049] 目的地装置14可经由计算机可读媒体16接收待解码的经编码视频数据。计算机可读媒体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的任一类型的媒体或装置。在一个实例中,计算机可读媒体16可包括通信媒体以使源装置12能够即时地将经编码视频数据直接发射到目的地装置14。可根据通信标准(例如,无线通信协议)调制经编码视频数据,且将其发射到目的地装置14。通信媒体可包括任何无线或有线通信媒体,例如射频(RF)频谱或一或多个实体发射线。通信媒体可形成基于分组的网络(例如,区域网、广域网或例如因特网的全球网络)的部分。通信媒体可包含路由器、交换器、基站或任何其它可用于促进从源装置12到目的地装置14的通信的设备。

[0050] 在一些实例中,可从输出接口22将经编码数据输出到存储装置。类似地,可通过输入接口从存储装置存取经编码数据。存储装置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,硬盘机、蓝光光盘、DVD、CD-ROM、快闪存储器、易失性或非易失性存储器或用于存储经编码视频数据的任何其它合适的数字存储媒体。在再一实例中,存储装置可对应于文件服务器或可存储由源装置12产生的经编码视频的另一中间存储装置。目的地装置14可经由流或下载从存储装置存取存储的视频数据。文件服务器可为能够存储经编码视频数据并将经编码视频数据发射到目的地装置14的任何类型的服务器。实例文件服务器包含网页服务器(例如,用于网站)、FTP服务器、网络附接存储(NAS)装置或本地磁盘机。目的地装置14可经由任何标准数据连接(包含因特网连接)而存取经编码视频数据。此数据连接可包含适合于存取存储于文件服务器上的经编码的视频数据的无线信道(例如,Wi-Fi连接)、有线连接(例如,DSL、缆线调制解调器,等等),或两者的组合。从存储装置的经编码视频数据的发射可为流发射、下载发射或其组合。

[0051] 本发明的技术未必限于无线应用或设定。所述技术可应用于支持多种多媒体应用中的任一者的视频译码,例如,空中电视广播、有线电视发射、卫星电视发射、因特网流视频发射(例如,经由HTTP的动态自适应流(DASH))、经编码到数据存储媒体上的数字视频、存储于数据存储媒体上的数字视频的解码或其它应用。在一些实例中,系统10可经配置以支持

单向或双向视频发射以支持例如视频流、视频播放、视频广播及/或视频电话的应用。

[0052] 在图1的实例中,源装置12包含视频源18、视频编码器20及输出接口22。目的地装置14包含输入接口28、视频解码器30及显示装置32。根据本发明,源装置12的视频编码器20可经配置以应用本发明中描述的技术,例如与二次变换有关的技术。根据本发明,目的地装置14的视频解码器30可经配置以应用本发明中描述的技术,例如与二次变换有关的技术。在其它实例中,源装置及目的地装置可包含其它组件或布置。举例来说,源装置12可从外部视频源18(例如外部摄像机)接收视频数据。同样地,目的地装置14可与外部显示装置介接,而非包含集成显示装置。

[0053] 如上文所提及,源装置12包含输出接口22且目的地装置14包含输入接口28。在一些实例中,输出接口22表示发射器且输入接口28表示接收器。在其它实例中,输出接口22及输入接口28表示收发器的实例(即,能够无线地发射及接收数据信号的接口)。收发器可经配置以在无线信号中发送及接收视频数据。举例来说,输出接口22当实施为收发器时可发送包含经编码视频数据的数据信号(例如,计算机可读媒体16),而输入接口28当实施为收发器时可接收包含经编码视频数据的数据信号(例如,计算机可读媒体16)。如上文所论述,视频编码器20可将经编码视频数据提供到输出接口22,而输入接口28可将经编码视频数据提供到视频解码器30。

[0054] 图1的所说明的系统10仅为一个实例。本发明中所描述的技术可由任何数字视频编码及/或解码装置执行。尽管通常本发明的技术由视频编码装置和视频解码装置来执行,但所述技术也可由视频编码器/解码器(通常被称作“编码解码器”)执行。源装置12及目的地装置14仅为其中源装置12产生经译码视频数据以供用于发射到目的地装置14的这些译码装置的实例。在一些实例中,装置12、14可以大体上对称的方式操作,使得装置12、14中的每一者包含视频编码及解码组件。因此,系统10可支持源装置12与目的地装置14之间的单向或双向视频发射,例如用于视频流、视频播放、视频广播或视频电话。

[0055] 源装置12的视频源18可包含视频捕获装置,例如视频摄像机、含有先前捕获的视频的视频存档及/或用于从视频内容提供者接收视频的视频馈入接口。作为另一替代,视频源18可产生基于计算机图形的数据作为源视频,或实况视频、存档视频及计算机产生的视频的组合。在一些情况下,源装置12及目的地装置14可形成所谓的摄像机电话或视频电话。然而,如上文所提及,本发明中描述的技术一般适用于视频译码,且可适用于无线及/或有线应用。在每一情况下,捕获、预先捕获或计算机产生的视频可由视频编码器20编码。经编码视频信息可接着由输出接口22输出到计算机可读媒体16上。

[0056] 计算机可读媒体16可包含暂态媒体,例如无线广播或有线网络发射。在一些实例中,网络服务器(未图示)可从源装置12接收经编码视频数据,且(例如)经由网络发射将经编码视频数据提供到目的地装置14。因此,在各种实例中,计算机可读媒体16可理解为包含各种形式的一或多个计算机可读媒体。

[0057] 目的地装置14的输入接口28从计算机可读媒体16接收信息。计算机可读媒体16的信息可包含由视频编码器20定义且也由视频解码器30使用的语法信息,其包含描述块及其它经译码单元的特性及/或处理的语法元素。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,例如阴极射线管(CRT)、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。

[0058] 视频编码器20及视频解码器30各自可实施为多种合适的编码器电路中的任一者,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。当所述技术部分地以软件实施时,装置可将用于软件的指令存储于合适的非暂时性计算机可读媒体中,且在硬件中使用一或多个处理器执行指令以执行本发明的技术。视频编码器20及视频解码器30中的每一者可包含于一或多个编码器或解码器中,编码器或解码器中的任一者可集成为各别装置中的组合式编码器/解码器(编码解码器)的部分。包含视频编码器20及/或视频解码器30的装置可包括集成电路、微处理器及/或无线通信装置(例如蜂窝式电话)。

[0059] 视频编码器20及视频解码器30可根据视频译码标准(例如上文所描述的标准)且在一些实例中根据高效视频译码(HEVC)标准(也称作ITU-T H.265或所述HEVC标准的扩展)或根据下一代视频译码标准操作。然而,本发明的技术并不限于任何特定译码标准。视频译码标准的其它实例包含MPEG-2及ITU-T H.263。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MUX-DEMUX单元,或其它硬件及软件,以处置在共同数据流或单独数据流中的音频及视频两者的编码。如果适用,则MUX-DEMUX单元可符合协议(例如ITU H.223多路复用器协议)或其它协议(例如,用户数据报协议(UDP))。

[0060] 大体上关于ITU-T H.265(也称作高效视频译码(HEVC),其描述于2013年4月的“系列H:音频视觉及多媒体系统,音频视觉服务的基础架构-移动视频的译码(SERIES H: AUDIOVISUAL AND MULTIMEDIA SYSTEMS, Infrastructure of audiovisual services—Coding of moving video)”高效视频译码,ITU-T H.265中)描述本发明的技术。然而,这些技术可应用于其它视频译码标准,包含HEVC的扩展及其它标准的扩展。其它视频译码标准的实例包含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/MPEG-4第10部分高级视频译码(AVC),包含其可调式视频译码(SVC)及多视图视频译码(MVC)扩展。

[0061] H.265标准最近由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)完成。在下文中被称作HEVC WD的HEVC草案规范可从http://phenix.int-evry.fr/jct/doc_end_user/documents/14_Vienna/wg11/JCTVC-N1003-v1.zip获得。

[0062] HEVC的多视图扩展(MV-HEVC)也已由JCT-3V开发。MV-HEVC的工作草案(WD)的实例(被称作MV-HEVC WD8)可从phenix.it-sudparis.eu/jct2/doc_end_user/documents/8_Valencia/wg11/JCT3V-H1002-v5.zip获得。HEVC的可调式扩展(名为SHVC)也已由JCT-VC开发。被称作SHVC WD6的SHVC的工作草案(WD)的实例可从phenix.it-sudparis.eu/jct/doc_end_user/documents/17_Valencia/wg11/JCTVC-Q1008-v2.zip获得。

[0063] 在HEVC及其它视频编解码器中,视频帧或图片可分成可包含明度及色度样本两者的树型块的序列,树型块也称为最大译码单元(LCU)。位流内的语法数据可定义LCU的大小,LCU为就像素的数目来说的最大译码单元。切片包含按译码次序的许多连续树型块。视频帧或图片可分割成一或多个切片。

[0064] 在HEVC中为产生经编码CTU(即为编码CTU),视频编码器20可对CTU的译码树型块递归地执行四分树分割,以将译码树型块划分成译码块,因此命名“译码树单元”。译码块为

样本的 $N \times N$ 块。CU可为具有明度样本阵列、Cb样本阵列及Cr样本阵列的图像的明度样本的译码块及色度样本的两个对应译码块,及用以译码译码块的样本的语法结构。在单色图片或具有三个单独色彩平面的图片中,CU可包括单一译码块及用以译码所述译码块的样本的语法结构。

[0065] 可根据四分树数据结构将每一树型块分裂成CU。一般来说,四分树数据结构包含每CU一个节点,其中根节点对应于树型块。如果CU分裂成四个子CU,则对应于所述CU的节点包含四个叶节点,所述四个叶节点中的每一者对应于所述子CU中的一者。四分树数据结构中的每一节点可提供对应CU的语法数据。举例来说,所述四分树中的节点可包含分裂旗标,从而指示对应于所述节点的CU是否分裂成子CU。针对CU的语法元素可经递回地定义,且可取决于所述CU是否分裂成子CU。如果CU不被进一步分裂,则其被称作叶CU。在本发明中,叶CU的四个子CU也称作叶CU,即使不存在原始叶CU的明确分裂。举例来说,如果 16×16 大小的CU未经进一步分裂,则尽管所述 16×16 CU从未经分裂,但4个 8×8 子CU也被称作叶CU。

[0066] 除CU不具有大小区别外,CU具有与H.264标准的宏块类似的用途。举例来说,树型块可分裂成四个子节点(也称作子CU),且每一子节点又可为上代节点且可被分裂成另外四个子节点。被称作四分树的叶节点的最终的未分裂子节点包括译码节点,所述译码节点也被称作叶CU。与经译码位流相关联的语法数据可定义可分裂树型块的最大次数(其被称作最大CU深度),且也可定义所述译码节点的最小大小。因此,位流也可定义最小译码单元(SCU)。CU的大小对应于译码节点的大小,且大体上为正方形形状。CU的大小可在 8×8 像素达到具有最大大小(例如 64×64 像素或更大)的树型块的大小的范围内。

[0067] CU包含译码节点及与所述译码节点相关联的预测单元(PU)及变换单元(TU)。每一CU可含有一或多个PU及一或多个TU。与CU相关联的语法数据可描述(例如)将CU分割成一或多个PU。作为编码CU的部分,视频编码器20可将CU的译码块分割成一或多个预测块。预测块可为经应用相同预测的样本的矩形(即,正方形或非正方形)块。CU的预测单元(PU)可为图片的明度样本的预测块、色度样本的两个对应预测块,及用以预测预测块样本的语法结构。视频编码器20可产生CU的每一PU的明度、Cb及Cr预测块的预测性明度、Cb及Cr块。在单色图片或具有三个单独色彩平面的图片中,PU可包括单一预测块及用以预测所述预测块的语法结构。

[0068] 视频编码器20可使用帧内预测或帧间预测来产生PU的预测性块。每一CU是以帧内预测模式或帧间预测模式中的一者译码。如果视频编码器20使用帧内预测来产生PU的预测性块,则视频编码器20可基于与PU相关联的图像的经解码样本来产生PU的预测性块。如果视频编码器20使用帧间预测来产生PU的预测性块,则视频编码器20可基于不同于与PU相关联的图片的一或多个图片的经解码样本而产生PU的预测性块。

[0069] 分割模式可在CU是经跳过或直接模式编码、经帧内预测模式编码抑或经帧间预测模式编码之间而不同。PU可分割成非正方形。

[0070] 此外,视频编码器20可将CU的每一残差块分解为一或多个变换块。变换块可为经应用相同变换的样本的矩形(正方形或非正方形)块。CU的变换单元(TU)可为明度样本的变换块、色度样本的两个对应的变换块,及用于对变换块样本进行变换的语法结构。因此,CU的每一TU可与明度变换块、Cb变换块及Cr变换块相关联。与TU相关联的明度变换块可为CU的明度残差块的子块。Cb变换块可为CU的Cb残差块的子块。Cr变换块可为CU的Cr残差块的

子块。在单色图片或具有三个单独色彩平面的图片中, TU可包括单一变换块及用于对所述变换块的样本进行变换的语法结构。在一些实例中, 明度及色度分量的相同CU的残差块可以不同方式分割。

[0071] 与CU相关联的语法数据也可描述(例如)根据四分树将CU分割成一或多个TU。TU可为正方形或非正方形(例如, 矩形)。本发明可使用术语“块”来指在HEVC的上下文中的CU、PU或TU中的任一者, 或在其它标准的上下文中的类似数据结构(例如, H. 264/AVC中的宏块及其子块)。

[0072] HEVC标准允许根据TU进行变换, 所述变换对于不同CU可不同。TU的大小通常是基于针对经分割LCU定义的给定CU内的PU的大小而设定, 但此可并非总是如此情况。TU的大小通常与PU相同或比PU小。在一些实例中, 可使用被称为“残差四分树”(RQT)的四分树结构而将对应于CU的残差样本再分为较小单元。可将RQT的叶节点称作变换单元(TU)。与TU相关联的像素差值可经变换以产生可经量化的变换系数。

[0073] 叶CU可包含一或多个PU。大体来说, PU表示对应于所述对应CU的全部或一部分的空间区域, 且PU可包含用于检索及/或产生用于PU的参考样本的数据。此外, PU包含与预测有关的数据。举例来说, 当PU经帧内模式编码时, PU的数据可包含于RQT中, 所述RQT可包含描述用于对应于PU的TU的帧内预测模式的数据。RQT也可被称作变换树。在一些实例中, 可在叶CU语法, 而不是RQT中传信帧内预测模式。作为实例, 当PU经帧间模式编码时, PU可包含定义PU的运动信息(例如一或多个运动向量)的数据。定义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)。

[0074] 具有一或多个PU的叶CU也可包含一或多个TU。如上文所论述, 可使用RQT(也称作TU四分树结构)来指定变换单元。举例来说, 分裂旗标可指示叶CU是否分裂成四个变换单元。接着, 可将每一变换单元进一步分裂为其它子TU。当TU未进一步分裂时, 可将其称作叶TU。大体来说, 对于帧内译码来说, 属于叶CU的所有叶TU共享相同的帧内预测模式。即, 一般应用相同帧内预测模式来计算叶CU的所有TU的预测值。对于帧内译码, 视频编码器可使用帧内预测模式将每一叶TU的残差值计算为在CU的对应于所述TU的部分与原始块之间的差。TU不必限于PU的大小。因此, TU可大于或小于PU。对于帧内译码, PU可与用于同一CU的对应叶TU共置。在一些实例中, 叶TU的最大大小可对应于对应叶CU的大小。

[0075] 此外, 叶CU的TU也可与各别四分树数据结构(称作RQT或如上文所提的变换树)相关联。即, 叶CU可包含指示所述叶CU如何被分割成TU的四分树。TU四分树的根节点大体对应于叶CU, 而CU四分树的根节点大体对应于树型块。将RQT的未被分裂的TU称作叶TU。一般来说, 除非另有指示, 否则本发明分别使用术语CU及TU来指叶CU及叶TU。

[0076] 视频序列通常包含视频帧或图片系列。图片的每一切片可包含描述所述各别切片的编码模式的切片语法数据。视频编码器20通常对个别视频切片内的视频块进行操作, 以便编码视频数据。视频块可对应于CU内的译码节点。视频块可具有固定或变化的大小, 且可根据指定译码标准而在大小方面不同。

[0077] 在本发明中, “N×N”与“N乘N”可被互换地使用以指在垂直及水平尺寸方面的视频块的像素尺寸, 例如, 16×16像素或16乘16像素。一般来说, 16×16块在垂直方向上具有16

个像素 ($y=16$) 且在水平方向上具有16个像素 ($x=16$)。同样地, $N \times N$ 块通常在垂直方向上具有 N 个像素且在水平方向上具有 N 个像素, 其中 N 表示非负整数值。可按行及列来排列块中的像素。此外, 块未必需要在水平方向上与垂直方向上具有相同数目个像素。举例来说, 块可包括 $N \times M$ 个像素, 其中 M 未必等于 N 。

[0078] 在使用CU的PU的帧内预测性或帧间预测性译码之后, 视频编码器20可计算CU的TU的残差数据。PU可包括描述在空间域 (也被称作像素域) 中产生预测性像素数据的技术或模式的语法数据, 且TU可包括在对残差视频数据应用变换 (例如离散余弦变换 (DCT)、整数变换、小波变换或概念上类似的变换) 之后变换域中的系数。残差数据可对应于未经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可形成包含表示CU的残差数据的经量化变换系数的TU。即, 视频编码器20可计算残差数据 (以残差块的形式)、变换残差块以产生变换系数的块, 且接着量化变换系数以形成经量化变换系数。视频编码器20可形成包含经量化变换系数的TU, 以及其它语法信息 (例如, TU的分裂信息)。

[0079] 如上文所提, 在应用任何变换以产生变换系数之后, 视频编码器20可执行变换系数的量化。量化通常指将变换系数量化以可能地减少用以表示所述系数的数据的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或全部相关联的位深度。举例来说, 可在量化期间将 n 位值降值舍位到 m 位值, 其中 n 大于 m 。

[0080] 在量化之后, 视频编码器20可扫描变换系数, 从而从包含经量化变换系数的二维矩阵产生一维向量。扫描可经设计以将较高能量 (且因此较低频率) 系数置于阵列前部, 及将较低能量 (且因此较高频率) 系数置于阵列后部。在一些实例中, 视频编码器20可利用预定义扫描次序来扫描经量化变换系数以产生可被熵编码的序列化向量。在其它实例中, 视频编码器20可执行适应性扫描。在扫描经量化变换系数以形成一维向量后, 视频编码器20可 (例如) 根据上下文自适应性可变长度译码 (CAVLC)、上下文自适应性二进制算术译码 (CABAC)、基于语法的上下文自适应性二进制算术译码 (SBAC)、机率区间分割熵 (PIPE) 译码或另一熵编码方法熵编码表示一维向量中的变换系数的特定语法元素。视频编码器20也可熵编码用于由视频解码器30在解码视频数据时使用的与经编码视频数据相关联的语法元素。

[0081] 一般来说, 视频解码器30执行尽管与由视频编码器20执行的过程互逆但类似的过程, 以解码经编码数据。举例来说, 视频解码器30逆量化且逆变换所接收TU的变换系数以再生残差块。视频解码器30可使用传信的预测模式 (帧内预测或帧间预测) 以形成经预测块。视频解码器30可 (在逐像素基础上) 使经预测块与残差块组合以再生原始块。视频解码器30可执行额外处理, 例如执行解块过程以减少沿块边界的视觉伪影。此外, 视频解码器30可以尽管与视频编码器20的CABAC编码过程互逆但与其大体上类似的方式使用CABAC解码语法元素。

[0082] 如前所提及, 视频编码器20可将DCT、整数变换、小波变换或概念地类似变换应用于残差视频数据。这些可分离变换可指示导出输入信号的替代表示的过程。给定 N 点向量 $x = [x_0, x_1, \dots, x_{N-1}]^T$ 及给定向量的集合 $\{\phi_0, \phi_1, \dots, \phi_{M-1}\}$, 可使用 $\phi_0, \phi_1, \dots, \phi_{M-1}$ 的线性组合估计或恰好地表示 x , 其可公式化如下,

$$[0083] \quad \hat{x} = \sum_{i=0}^{M-1} f_i \cdot \phi_i$$

[0084] 其中 \hat{x} 可为 x 的近似值或等值,向量 $f=[f_1, f_2, \dots, f_{M-1}]$ 被称作变换系数向量且 $\{\phi_0, \phi_1, \dots, \phi_{M-1}\}$ 为变换基础向量。

[0085] 在视频译码的情形中,变换系数可大致不相关且稀疏。举例来说,输入向量 x 的能量可精简于仅仅几个变换系数上,且剩余大部分变换系数通常接近于零。

[0086] 给定特定输入数据,在能量精简性方面的最佳变换可为所谓的卡忽南-拉维(Karhunen-Loeve)变换(KLT)。KLT使用输入数据的协方差矩阵的特征向量作为变换基础向量。因此,KLT实际上为数据相依的变换且不具有通用数学公式。然而,在某些假定(例如输入数据形成一阶静止马尔可夫过程)下,文献中已证实对应KLT实际上为单一变换的正弦族中的成员,在IEEE模式分析与机器智能汇刊(IEEE Trans.on Pattern Analysis and Machine Intelligence),1,356(1979)杰恩(Jain),A.K.的“单一变换的正弦族”中描述所述单一变换。单一变换的正弦族可使用经公式化如下的变换基础向量指示变换:

$$[0087] \quad \phi_m(k) = A \cdot e^{ik\theta} + B \cdot e^{-ik\theta}$$

[0088] 在上述等式中, e 为大致等于2.71828的自然对数的底, A 、 B 及 θ 一般为复数,且取决于 m 的值。另外,在上述等式中, ϕ_m 为向量, $\phi_m(k)$ 指示向量 ϕ_m 的第 k 个分量,且 i 指示复数的虚部。

[0089] 若干熟知变换(包含离散傅立叶变换、余弦变换、正弦变换及KLT(用于一阶静止马尔可夫过程))为单一变换的此正弦族的成员。如IEEE信号处理汇刊(IEEE Trans.Sig.Processing)SP-42,1038-1051(1994年),S.A.马尔图齐(Martucci)的“对称卷积以及离散正弦及余弦变换(Symmetric convolution and the discrete sine and cosine transforms)”中所描述,完整DCT及离散正弦变换(DST)族可基于不同类型包含总共16个变换,且不同类型DCT及DST的完整定义在下文中给出。不同类型可对应于 A 、 B 及 θ 的不同值。

[0090] 假定输入 N 点向量表示为 $x=[x_0, x_1, \dots, x_{N-1}]^T$,且 N 点向量通过乘以矩阵而变换到表示为 $y=[y_0, y_1, \dots, y_{N-1}]^T$ 的另一 N 点变换系数向量。将 x 变换到 y 的过程可根据以下变换公式中的一者进一步说明,其中 k 在0到 $N-1$ (包含0及 $N-1$)范围内:

[0091] DCT类型I (DCT-1):

$$[0092] \quad y_k = \sum_{n=0}^{N-1} \sqrt{\frac{2}{N-1}} \cos\left(\frac{\pi \cdot n \cdot k}{N-1}\right) \cdot w_0 \cdot w_1 \cdot x_n,$$

$$[0093] \quad \text{其中 } w_0 = \begin{cases} \frac{1}{\sqrt{2}}, & \text{如果 } n = 0 \text{ 或 } n = N-1 \\ 1, & \text{否则} \end{cases}, \quad w_1 = \begin{cases} \frac{1}{\sqrt{2}}, & \text{如果 } k = 0 \text{ 或 } k = N-1 \\ 1, & \text{否则} \end{cases}$$

[0094] DCT类型II (DCT-2):

$$[0095] \quad y_k = \sum_{n=0}^{N-1} \sqrt{\frac{2}{N}} \cos\left(\frac{\pi \cdot (n+0.5) \cdot k}{N-1}\right) \cdot w_0 \cdot x_n$$

[0096] 其中 $w_0 = \begin{cases} \frac{1}{\sqrt{2}}, & \text{如果 } k = 0 \\ 1, & \text{否则} \end{cases}$

[0097] DCT类型III (DCT-3) :

[0098]
$$y_k = \sum_{n=0}^{N-1} \sqrt{\frac{2}{N}} \cos\left(\frac{\pi \cdot n \cdot (k + 0.5)}{N}\right) \cdot w_0 \cdot x_n$$

[0099] 其中 $w_0 = \begin{cases} \frac{1}{\sqrt{2}}, & \text{如果 } n = 0 \\ 1, & \text{否则} \end{cases}$

[0100] DCT类型IV (DCT-4) :

[0101]
$$y_k = \sum_{n=0}^{N-1} \sqrt{\frac{2}{N}} \cos\left(\frac{\pi \cdot (n+0.5) \cdot (k+0.5)}{N}\right) \cdot x_n,$$

[0102] DCT类型V (DCT-5) :

[0103]
$$y_k = \sum_{n=0}^{N-1} \sqrt{\frac{2}{N-0.5}} \cos\left(\frac{\pi \cdot n \cdot k}{N-0.5}\right) \cdot w_0 \cdot w_1 \cdot x_n,$$

[0104] 其中 $w_0 = \begin{cases} \frac{1}{\sqrt{2}}, & \text{如果 } n = 0 \\ 1, & \text{否则} \end{cases}, w_1 = \begin{cases} \frac{1}{\sqrt{2}}, & \text{如果 } k = 0 \\ 1, & \text{否则} \end{cases}$

[0105] DCT类型VI (DCT-6) :

[0106]
$$y_k = \sum_{n=0}^{N-1} \sqrt{\frac{2}{N-0.5}} \cos\left(\frac{\pi \cdot (n+0.5) \cdot k}{N-0.5}\right) \cdot w_0 \cdot w_1 \cdot x_n$$

[0107] 其中 $w_0 = \begin{cases} \frac{1}{\sqrt{2}}, & \text{如果 } n = N-1 \\ 1, & \text{否则} \end{cases}, w_1 = \begin{cases} \frac{1}{\sqrt{2}}, & \text{如果 } k = 0 \\ 1, & \text{否则} \end{cases}$

[0108] DCT类型VII (DCT-7) :

[0109]
$$y_k = \sum_{n=0}^{N-1} \sqrt{\frac{2}{N-0.5}} \cos\left(\frac{\pi \cdot n \cdot (k+0.5)}{N-0.5}\right) \cdot w_0 \cdot w_1 \cdot x_n$$

[0110] 其中 $w_0 = \begin{cases} \frac{1}{\sqrt{2}}, & \text{如果 } n = 0 \\ 1, & \text{否则} \end{cases}, w_1 = \begin{cases} \frac{1}{\sqrt{2}}, & \text{若 } k = N-1 \\ 1, & \text{否则} \end{cases}$

[0111] DCT类型VIII (DCT-8) :

[0112]
$$y_k = \sum_{n=0}^{N-1} \sqrt{\frac{2}{N+0.5}} \cos\left(\frac{\pi \cdot (n+0.5) \cdot (k+0.5)}{N+0.5}\right) \cdot x_n,$$

[0113] DST类型I (DST-1) :

[0114]
$$y_k = \sum_{n=0}^{N-1} \sqrt{\frac{2}{N+1}} \sin\left(\frac{\pi \cdot (n+1) \cdot (k+1)}{N+1}\right) \cdot x_n,$$

[0115] DST类型II (DST-2) :

$$[0116] \quad y_k = \sum_{n=0}^{N-1} \sqrt{\frac{2}{N}} \sin\left(\frac{\pi \cdot (n+0.5) \cdot (k+1)}{N}\right) \cdot w_0 \cdot x_n$$

$$[0117] \quad \text{其中 } w_0 = \begin{cases} \frac{1}{\sqrt{2}}, & \text{如果 } k = N-1 \\ 1, & \text{否则} \end{cases}$$

[0118] DST类型III (DST-3) :

$$[0119] \quad y_k = \sum_{n=0}^{N-1} \sqrt{\frac{2}{N}} \sin\left(\frac{\pi \cdot (n+1) \cdot (k+0.5)}{N}\right) \cdot w_0 \cdot x_n$$

$$[0120] \quad \text{其中 } w_0 = \begin{cases} \frac{1}{\sqrt{2}}, & \text{如果 } n = N-1 \\ 1, & \text{否则} \end{cases}$$

[0121] DST类型IV (DST-4) :

$$[0122] \quad y_k = \sum_{n=0}^{N-1} \sqrt{\frac{2}{N}} \sin\left(\frac{\pi \cdot (n+0.5) \cdot (k+0.5)}{N}\right) \cdot x_n,$$

[0123] DST类型V (DST-5) :

$$[0124] \quad y_k = \sum_{n=0}^{N-1} \sqrt{\frac{2}{N+0.5}} \sin\left(\frac{\pi \cdot (n+1) \cdot (k+1)}{N+0.5}\right) \cdot x_n,$$

[0125] DST类型VI (DST-6) :

$$[0126] \quad y_k = \sum_{n=0}^{N-1} \sqrt{\frac{2}{N+0.5}} \sin\left(\frac{\pi \cdot (n+0.5) \cdot (k+1)}{N+0.5}\right) \cdot x_n,$$

[0127] DST类型VII (DST-7) :

$$[0128] \quad y_k = \sum_{n=0}^{N-1} \sqrt{\frac{2}{N+0.5}} \sin\left(\frac{\pi \cdot (n+1) \cdot (k+0.5)}{N+0.5}\right) \cdot x_n,$$

[0129] DST类型VIII (DST-8) :

$$[0130] \quad y_k = \sum_{n=0}^{N-1} \sqrt{\frac{2}{N-0.5}} \cos\left(\frac{\pi \cdot (n+0.5) \cdot (k+0.5)}{N-0.5}\right) \cdot w_0 \cdot w_1 \cdot x_n$$

$$[0131] \quad \text{其中 } w_0 = \begin{cases} \frac{1}{\sqrt{2}}, & \text{如果 } n = N-1 \\ 1, & \text{否则} \end{cases}, \quad w_1 = \begin{cases} \frac{1}{\sqrt{2}}, & \text{如果 } k = N-1 \\ 1, & \text{否则} \end{cases}$$

[0132] 变换类型 (例如DST类型) 可由变换基底函数的数学公式指定。举例来说, 4点DST-VII及8-点DST-VII具有相同变换类型, 不管N的值。

[0133] 在不损失一般性的情况下, 所有上文变换类型可使用下文的通用公式来表示:

$$[0134] \quad y_m = \sum_{n=0}^{N-1} T_{m,n} \cdot x_n.$$

[0135] 在上文等式中, T为可由一个某种变换 (例如, DCT类型-I到DCT类型-VIII, 或DST类型-I到DST类型-VIII) 的定义指定的变换矩阵, 且T的列向量 (例如, $[T_{i,0}, T_{i,1}, T_{i,2}, \dots, T_{i,N-1}]$) 为第i个变换基础向量。应用于N点输入向量的变换可被称作N点变换。

[0136] 还应注意, 应用于1-D输入数据x的以上变换公式可以如下文的矩阵乘法形式来表

示

$$[0137] \quad y = T \cdot x$$

[0138] 在上述等式中, T 指示变换矩阵, x 指示输入数据向量, 且 y 指示输出变换系数向量。

[0139] 先前章节中介绍的变换可应用于1-D输入数据, 且变换也可经扩展用于2-D输入数据源。在以下论述中, X 为输入 $M \times N$ 数据阵列。于2-D输入数据上应用变换的技术可包含可分离的及非可分离的2-D变换。

[0140] 可分离的2-D变换顺序对 X 的水平向量及垂直向量应用1-D变换, 公式如下:

$$[0141] \quad Y = C \cdot X \cdot R^T$$

[0142] 在上述等式中, C 及 R 分别表示给定 $M \times M$ 及 $N \times N$ 变换矩阵。从所述公式, 可见 C 针对 X 的列向量应用1-D变换, 而 R 针对 X 的行向量应用1-D变换。在本文件的稍后部分中, 为简单起见, C 及 R 表示为左 (垂直) 及右 (水平) 变换, 且其都形成变换对。存在 C 等于 R 且为正交矩阵的情况。在此情况下, 可分离的2-D变换可由仅一个变换矩阵确定。

[0143] 在一些实例中, 非可分离的2-D变换可首先通过应用以下数学映射将 X 的所有元素重组成单个向量, 即 X' :

$$[0144] \quad X'_{(i \cdot N + j)} = X_{i,j}$$

[0145] 接着, 1-D变换 T' 可如下应用于 X' :

$$[0146] \quad Y = T' \cdot X'$$

[0147] 在上述等式中, T' 为 $(M \cdot N) \times (M \cdot N)$ 变换矩阵。

[0148] 在视频译码中, 可应用可分离2-D变换, 此是由于2-D变换与1-D变换相比可使用较少操作计数 (例如, 加法、乘法)。

[0149] 在常规视频编解码器 (例如 H.264/AVC) 中, 4点及8点 DCT 类型 II 的整数近似值应用于帧内及帧间预测残差两者。换句话说, 视频译码器可将4点及8点 DCT 类型 II 的整数近似值应用于使用帧内或帧间预测产生的残差块。为更好地适应残差样本的各种统计数据, 将除 DCT 类型 II 外的更灵活类型的变换用于新一代视频编解码器中。举例来说, 在 HEVC 中, 4点类型 VII DST 的整数近似值可用于帧内预测残差。如 IEEE 声学、语音和信号处理国际会议 (IEEE International Conference on Acoustics, Speech and Signal Processing; ICASSP), 2010年3月, 第726到729页, J. 汉 (Han)、A. 萨克西纳 (Saxena) 及 K. 罗斯 (Rose) 的“针对视频/图像译码中联合最佳的空间预测及自适应性变换 (Towards jointly optimal spatial prediction and adaptive transform in video/image coding)”中所描述, 已在理论上证实并以实验方式验证用于 HEVC 的 DST 类型 VII 与 DCT 类型 II 相比对于沿帧内预测方向产生的残差向量更高效。举例来说, DST 类型 VII 与 DCT 类型 II 相比对于由水平帧内预测方向产生的行残差向量更高效。在 HEVC 中, 4点 DST 类型 VII 的整数近似值仅仅应用于 4×4 明度帧内预测残差块。用于 HEVC 中的 4点 DST-VII 在下文展示:

[0150] 4×4 DST-VII:

[0151] {29, 55, 74, 84}

[0152] {74, 74, 0, -74}

[0153] {84, -29, -74, 55}

[0154] {55, -84, 74, -29}

[0155] 在 HEVC 中, 对于不为 4×4 明度帧内预测残差块的残差块, 也可应用4点、8点、16点

及32点DCT类型II的整数近似值,如下文所示:

[0156] 4点DCT-II:

[0157] {64,64,64,64}

[0158] {83,36,-36,-83}

[0159] {64,-64,-64,64}

[0160] {36,-83,83,-36}

[0161] 8点DCT-II:

[0162] {64,64,64,64,64,64,64,64}

[0163] {89,75,50,18,-18,-50,-75,-89}

[0164] {83,36,-36,-83,-83,-36,36,83}

[0165] {75,-18,-89,-50,50,89,18,-75}

[0166] {64,-64,-64,64,64,-64,-64,64}

[0167] {50,-89,18,75,-75,-18,89,-50}

[0168] {36,-83,83,-36,-36,83,-83,36}

[0169] {18,-50,75,-89,89,-75,50,-18}

[0170] 16点DCT-II:

[0171] {64,64,64,64,64,64,64,64,64,64,64,64,64,64,64,64}

[0172] {90,87,80,70,57,43,25,9,-9,-25,-43,-57,-70,-80,-87,-90}

[0173] {89,75,50,18,-18,-50,-75,-89,-89,-75,-50,-18,18,50,75,89}

[0174] {87,57,9,-43,-80,-90,-70,-25,25,70,90,80,43,-9,-57,-87}

[0175] {83,36,-36,-83,-83,-36,36,83,83,36,-36,-83,-83,-36,36,83}

[0176] {80,9,-70,-87,-25,57,90,43,-43,-90,-57,25,87,70,-9,-80}

[0177] {75,-18,-89,-50,50,89,18,-75,-75,18,89,50,-50,-89,-18,75}

[0178] {70,-43,-87,9,90,25,-80,-57,57,80,-25,-90,-9,87,43,-70}

[0179] {64,-64,-64,64,64,-64,-64,64,64,-64,-64,64,64,-64,-64,64}

[0180] {57,-80,-25,90,-9,-87,43,70,-70,-43,87,9,-90,25,80,-57}

[0181] {50,-89,18,75,-75,-18,89,-50,-50,89,-18,-75,75,18,-89,50}

[0182] {43,-90,57,25,-87,70,9,-80,80,-9,-70,87,-25,-57,90,-43}

[0183] {36,-83,83,-36,-36,83,-83,36,36,-83,83,-36,-36,83,-83,36}

[0184] {25,-70,90,-80,43,9,-57,87,-87,57,-9,-43,80,-90,70,-25}

[0185] {18,-50,75,-89,89,-75,50,-18,-18,50,-75,89,-89,75,-50,18}

[0186] {9,-25,43,-57,70,-80,87,-90,90,-87,80,-70,57,-43,25,-9}

[0187] 32点DCT-II:

[0188] {64,64}

[0189] {90,90,88,85,82,78,73,67,61,54,46,38,31,22,13,4,-4,-13,-22,-31,-38,-46,-54,-61,-67,-73,-78,-82,-85,-88,-90,-90}

[0190] {90,87,80,70,57,43,25,9,-9,-25,-43,-57,-70,-80,-87,-90,-90,-87,-80,-70,-57,-43,-25,-9,9,25,43,57,70,80,87,90}

- [0191] {90,82,67,46,22,-4,-31,-54,-73,-85,-90,-88,-78,-61,-38,-13,13,38,61,78,88,90,85,73,54,31,4,-22,-46,-67,-82,-90}
- [0192] {89,75,50,18,-18,-50,-75,-89,-89,-75,-50,-18,18,50,75,89,89,75,50,18,-18,-50,-75,-89,-89,-75,-50,-18,18,50,75,89}
- [0193] {88,67,31,-13,-54,-82,-90,-78,-46,-4,38,73,90,85,61,22,-22,-61,-85,-90,-73,-38,4,46,78,90,82,54,13,-31,-67,-88}
- [0194] {87,57,9,-43,-80,-90,-70,-25,25,70,90,80,43,-9,-57,-87,-87,-57,-9,43,80,90,70,25,-25,-70,-90,-80,-43,9,57,87}
- [0195] {85,46,-13,-67,-90,-73,-22,38,82,88,54,-4,-61,-90,-78,-31,31,78,90,61,4,-54,-88,-82,-38,22,73,90,67,13,-46,-85}
- [0196] {83,36,-36,-83,-83,-36,36,83,83,36,-36,-83,-83,-36,36,83,83,36,-36,-83,-83,-36,36,83,83,36,-36,-83,-83,-36,36,83}
- [0197] {82,22,-54,-90,-61,13,78,85,31,-46,-90,-67,4,73,88,38,-38,-88,-73,-4,67,90,46,-31,-85,-78,-13,61,90,54,-22,-82}
- [0198] {80,9,-70,-87,-25,57,90,43,-43,-90,-57,25,87,70,-9,-80,-80,-9,70,87,25,-57,-90,-43,43,90,57,-25,-87,-70,9,80}
- [0199] {78,-4,-82,-73,13,85,67,-22,-88,-61,31,90,54,-38,-90,-46,46,90,38,-54,-90,-31,61,88,22,-67,-85,-13,73,82,4,-78}
- [0200] {75,-18,-89,-50,50,89,18,-75,-75,18,89,50,-50,-89,-18,75,75,-18,-89,-50,50,89,18,-75,-75,18,89,50,-50,-89,-18,75}
- [0201] {73,-31,-90,-22,78,67,-38,-90,-13,82,61,-46,-88,-4,85,54,-54,-85,4,88,46,-61,-82,13,90,38,-67,-78,22,90,31,-73}
- [0202] {70,-43,-87,9,90,25,-80,-57,57,80,-25,-90,-9,87,43,-70,-70,43,87,-9,-90,-25,80,57,-57,-80,25,90,9,-87,-43,70}
- [0203] {67,-54,-78,38,85,-22,-90,4,90,13,-88,-31,82,46,-73,-61,61,73,-46,-82,31,88,-13,-90,-4,90,22,-85,-38,78,54,-67}
- [0204] {64,-64,-64,64,64,-64,-64,64,64,-64,-64,64,64,-64,-64,64,64,-64,-64,64,64,-64,-64,64,64,-64,-64,64}
- [0205] {61,-73,-46,82,31,-88,-13,90,-4,-90,22,85,-38,-78,54,67,-67,-54,78,38,-85,-22,90,4,-90,13,88,-31,-82,46,73,-61}
- [0206] {57,-80,-25,90,-9,-87,43,70,-70,-43,87,9,-90,25,80,-57,-57,80,25,-90,9,87,-43,-70,70,43,-87,-9,90,-25,-80,57}
- [0207] {54,-85,-4,88,-46,-61,82,13,-90,38,67,-78,-22,90,-31,-73,73,31,-90,22,78,-67,-38,90,-13,-82,61,46,-88,4,85,-54}
- [0208] {50,-89,18,75,-75,-18,89,-50,-50,89,-18,-75,75,18,-89,50,50,-89,18,75,-75,-18,89,-50,-50,89,-18,-75,75,18,-89,50}
- [0209] {46,-90,38,54,-90,31,61,-88,22,67,-85,13,73,-82,4,78,-78,-4,82,-73,-13,85,-67,-22,88,-61,-31,90,-54,-38,90,-46}
- [0210] {43,-90,57,25,-87,70,9,-80,80,-9,-70,87,-25,-57,90,-43,-43,90,-57,-

25,87,-70,-9,80,-80,9,70,-87,25,57,-90,43}

[0211] {38,-88,73,-4,-67,90,-46,-31,85,-78,13,61,-90,54,22,-82,82,-22,-54,90,-61,-13,78,-85,31,46,-90,67,4,-73,88,-38}

[0212] {36,-83,83,-36,-36,83,-83,36,36,-83,83,-36,-36,83,-83,36,36,-83,83,-36,-36,83,-83,36}

[0213] {31,-78,90,-61,4,54,-88,82,-38,-22,73,-90,67,-13,-46,85,-85,46,13,-67,90,-73,22,38,-82,88,-54,-4,61,-90,78,-31}

[0214] {25,-70,90,-80,43,9,-57,87,-87,57,-9,-43,80,-90,70,-25,-25,70,-90,80,-43,-9,57,-87,87,-57,9,43,-80,90,-70,25}

[0215] {22,-61,85,-90,73,-38,-4,46,-78,90,-82,54,-13,-31,67,-88,88,-67,31,13,-54,82,-90,78,-46,4,38,-73,90,-85,61,-22}

[0216] {18,-50,75,-89,89,-75,50,-18,-18,50,-75,89,-89,75,-50,18,18,-50,75,-89,89,-75,50,-18,-18,50,-75,89,-89,75,-50,18}

[0217] {13,-38,61,-78,88,-90,85,-73,54,-31,4,22,-46,67,-82,90,-90,82,-67,46,-22,-4,31,-54,73,-85,90,-88,78,-61,38,-13}

[0218] {9,-25,43,-57,70,-80,87,-90,90,-87,80,-70,57,-43,25,-9,-9,25,-43,57,-70,80,-87,90,-90,87,-80,70,-57,43,-25,9}

[0219] {4,-13,22,-31,38,-46,54,-61,67,-73,78,-82,85,-88,90,-90,90,-90,88,-85,82,-78,73,-67,61,-54,46,-38,31,-22,13,-4}

[0220] 如<http://www.hhi.fraunhofer.de/fields-of-competence/image-processing/research-groups/image-video-coding/hevc-high-efficiency-video-coding/transform-coding-using-the-residual-quadtree-rqt.html>中所描述,为了适应残差块的各种特性,使用残差四分树(RQT)的变换译码结构可应用于HEVC中。如上文简单描述,每一图片可分成多个译码树单元(CTU)。可针对特定图片块或切片以光栅扫描次序译码CTU。CTU可为正方形块且可表示四分树(例如,译码树)的根。CTU大小可在 8×8 到 64×64 明度样本范围内,但通常使用 64×64 。每一CTU可经进一步分裂成称为CU的较小正方形块。在CTU经递归地分裂成CU之后,每一CU可经进一步分成一或多个PU及一或多个TU。CU分割成TU可基于四分树方法递归地实施。因此,每一CU的残差信号是通过树状结构(即,残差四分树(RQT))来译码。

[0221] RQT允许从 4×4 达 32×32 明度样本的TU大小。图2展示其中CU包含10个TU(标记有字母a到j)及对应块分割的实例。RQT的每一节点实际上可为TU。可以深度优先树遍历次序处理个别TU,深度优先树遍历次序在图2中说明为字母次序,其遵循使用深度优先遍历的递归Z扫描。四分树方法可使得能够使变换适应于残差信号的变化空间频率特性。通常,具有更大空间支持的更大变换块大小提供更佳频率分辨率。然而,具有更小空间支持的更小变换块大小提供更佳空间分辨率。空间及频率分辨率两者之间的折衷可通过编码器模式决策加以选择。举例来说,更大或更小变换块大小的选择可基于速率-失真最佳化技术。速率-失真最佳化技术可针对每一译码模式(例如,特定RQT分裂结构)计算译码位与重建失真的加权和(例如,速率-失真成本)。另外,速率-失真最佳化技术可选择具有最少速率-失真成本的译码模式作为最佳模式。

[0222] 在RQT中可定义三个参数:RQT的最大深度、RQT的最小允许变换大小及最大允许变换大小。最小及最大变换大小可在 4×4 到 32×32 样本范围内变化,其可对应于先前段落中提及的所支持块变换。RQT的最大允许深度可限制TU的数目。最大深度等于零可意谓若每一所包含变换块(TB)达到最大允许的变换大小(例如 32×32),则译码块(CB)无法经更进一步分裂。

[0223] 三个参数中的每一者相互作用并影响RQT结构。考虑根CB大小为 64×64 ,最大深度等于零且最大变换大小等于 32×32 的情况。在此情况下,CB可被分割至少一次,这是由于否则的话其将产生可不被允许的 64×64 TB。RQT参数(例如,最大RQT深度、最小及最大变换大小)可在序列参数集层级处在位流中发射。关于RQT深度,可针对经帧内及帧间译码CU指定且传信不同值。

[0224] 可将四分树变换应用于帧内残差块及帧间残差块两者。与当前残差四分树分割相同大小的DCT-II变换可经应用于残差块。然而,如果当前残差四分树块为 4×4 并由帧内预测产生,则可应用上述 4×4 DST-VII变换。在HEVC中,更大大小的变换(例如, 64×64 变换)主要归因于其受限的益处及对于相对较小分辨率视频的相对高的复杂度而不被采用。

[0225] 尽管相较于常规DCT类型II,DST类型VII可有效改进帧内译码效率,但变换效率可能由于预测残差呈现各种统计数据而相对受限,且DCT类型II及DST类型VII的固定使用可不会有效适应所有可能的情况。

[0226] 如2009年8月,光学工程(Opt.Eng.),第48卷,第8号,第087004-1到087004-14页,里姆(Lim)等人的“速率-失真最佳化自适应性变换译码(Rate-distortion optimized adaptive transform coding)”中所描述,变换方案适应性地使用DCT或DST的整数版本以用于预测残差,对于每一块,传信DCT或DST变换是否用于预测残差。如2008年10月,IEEE第15次国际会议图像处理的学报(Proc.15th IEEE Int.Conf.Image Process.),第2116到2119页,Y.叶(Ye)及M.卡兹维茨(Karczewicz)的“基于双向帧内预测、定向变换及自适应性系数扫描的改进的H.264帧内译码(Improved H.264intra coding based on bidirectional intra prediction,directional transform,and adaptive coefficient scanning)”中所描述,每一帧内预测模式可映射到唯一变换(例如,C及R)对、预定义为KLT对,使得模式相依变换(MDDT)适用。以此方式,不同KLT可用于不同帧内预测模式。然而,使用哪一变换经预定义并取决于帧内预测模式。

[0227] 然而,如2012年1月IEEE电路系统视频技术学报(IEEE Trans.Circuits Syst.Video Technol.)第22卷,第1号,第138到151页,X.赵(Zhao)等人的“用速率-失真最佳化变换的视频译码(Video coding with rate-distortion optimized transform)”中所描述,可使用更多变换且这些实例可明确地传信来自从离线训练过程导出的变换候选者的预定义集合的变换的索引(例如,编码指示索引的数据)。类似于MDDT,在这些实例中,每一帧内预测方向可具有其变换对的唯一集合。可传信索引以指定从所述集合选择哪个变换对。举例来说,对于最小块大小 4×4 ,可存在多达四个垂直KLT及多达四个水平KLT。因此,在此实例中,可选择16个组合。对于较大块大小,可使用较少数目个组合。在本发明中提出的技术可应用于帧内及帧间预测残差两者。在本发明中,帧内预测残差指使用帧内预测产生的残差数据。此外,在本发明中,帧间预测残差指使用帧间预测产生的残差数据。用于帧间预测残差,可选择KLT的多达16个组合,且可针对每一块传信所述组合(对于 4×4 为四且对

于 8×8 为十六)中的一者的索引。

[0228] 如IEEE图像处理学报 (IEEE Trans. Image Processing), 萨克西纳等人的“用于图像/视频译码中的帧内预测的基于DCT/DST的变换译码 (DCT/DST-based transform coding for intra prediction in image/video coding)”, 及2012年IEEE电路系统视频技术学报 (IEEE Trans. Circuits Syst. Video Technol.) 第22卷, 第4号, 第545到554页, 易欧 (Yeo) 等人的“用于译码双向帧内预测残差值的模式相依变换 (Mode-dependent transforms for coding directional intra prediction residuals)”中所描述, 可使用多个变换。然而, 在这些实例中, 替代使用KLT (其可经训练), DCT (DCT-II) 或DST (DST-VII) 可用于变换单元 (其中左变换及右变换两者相同的), 且待使用哪一者由所传信旗标确定。

[0229] 此外, 如2013年11月, IEEE的信号处理中所选择主题的杂志 (IEEE Journal of Selected Topics in Signal Processing), 第7卷, 第6期, 周 (Zou) 等人的“基于针对帧内块译码的Lloyd型算法的速率-失真最佳化变换 (Rate-Distortion Optimized Transforms Based on the Lloyd-Type Algorithm for Intra Block Coding)”中所描述, 使用若干预定义KLT对, 可对于译码单元传信 (而不是导出) 变换对的索引, 以使得CU的每一TU使用同一对变换。

[0230] 如JCTVC-G281, 安 (An) 等人的“非-CE7: 用于帧间预测残差的边界相依变换 (Non-CE7: Boundary-Dependent Transform for Inter-Predicted Residue)”中所描述, 可根据TU在CU内的位置而将多个变换选择用于TU的帧间预测残差。C及R变换两者可选自DST-VII及DST-VII的倒转版本。因此, 对于CU内的TU, 多达四个组合可为可能的。然而, 由于组合可完全由PU的位置确定, 因此可能不需要传信正在使用哪个组合。

[0231] 在2015年3月23日申请的第62/137,038号美国临时申请案、2015年1月26日申请的第62/107,996号美国临时申请案及2016年1月25日申请的第15/005,736号美国专利申请案中, 增强型多变换 (EMT) 技术经提出用于帧内及帧间预测残差两者。在EMT中, CU层级旗标可经传信以指示是仅仅使用常规DCT-2变换抑或使用其它非DCT2类型变换。如果CU层级经传信为1, 则可进一步对于当前CU内部的每一TU传信两位TU层级索引以指示将来自变换子集的哪个水平/垂直变换用于当前TU。变换子集可含有从DST-VII、DCT-VIII、DCT-V及DST-I中选择的一个变换, 且选择可基于帧内预测模式及其为水平变换子集抑或垂直变换子集。

[0232] 图3A为视频编码器 (例如视频编码器20) 处的实例变换过程的说明。图3B为视频解码器 (例如视频解码器30) 处的实例变换过程的说明。在图3A的实例中, 视频编码器20应用前向一次变换40 (fwdT), 紧接着二次变换41 (fwdR), 紧接着前向量化42 (fwdQua)。在图3B的实例中, 视频解码器30应用逆量化43 (invQ), 紧接着逆二次变换44 (invR), 紧接着逆一次变换45 (invTran)。前向一次变换40可将残差样本从像素域转换成频域中的变换系数。逆一次变换43可将频域中的变换系数转换成像素域中的残差样本。

[0233] 二次变换41可以用于变换系数的更好能量精简。如图3A及3B中所说明, 二次变换41可将另一变换应用于从第一变换过程导出的变换系数上。

[0234] 如2011年3月16日到23日, 瑞士日内瓦, ITU-T SG16WP3及ISO/IEC JTC1/SC29/WG11的视频译码联合合作小组 (JCT-VC), JCTVC-E380, E. 阿史那 (Alshina), A. 阿施恩 (Alshin), F. 费尔南德斯 (Fernandes), A. 萨克西纳, V. 赛瑞珍 (Seregin), Z. 马 (Ma), W. - J. Han (三星) 的“CE7: 由三星进行的ROT的实验结果 (CE7: Experimental results of ROT

by Samsung) ”中所描述,二次变换可为旋转变换(ROT)。ROT不替换一次变换。然而,ROT仅对于变换系数矩阵的低频部分被应用为二次变换。使用如上文中所描述的ROT的情况下,对于每一帧内译码TU,传信指示应用来自四个预定义变换候选者中的哪个ROT的索引。 4×4 二次变换应用于 4×4 帧内译码TU,而 8×8 二次变换应用于 8×8 及更大的帧内译码TU。举例来说,在实例中,二次变换可指定如下。

$$\begin{aligned}
 & R_{vertical}(\alpha_1, \alpha_2, \alpha_3) = \\
 [0235] \quad & = \begin{pmatrix} \cos \alpha_1 \cos \alpha_3 - \sin \alpha_1 \cos \alpha_2 \sin \alpha_3 & -\sin \alpha_1 \cos \alpha_3 - \cos \alpha_1 \cos \alpha_2 \sin \alpha_3 & \sin \alpha_2 \sin \alpha_3 & 0 \\ \cos \alpha_1 \sin \alpha_3 + \sin \alpha_1 \cos \alpha_2 \cos \alpha_3 & -\sin \alpha_1 \sin \alpha_3 + \cos \alpha_1 \cos \alpha_2 \cos \alpha_3 & -\sin \alpha_2 \cos \alpha_3 & 0 \\ \sin \alpha_1 \sin \alpha_2 & \cos \alpha_1 \sin \alpha_2 & \cos \alpha_2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \\
 & R_{horizontal}(\alpha_4, \alpha_5, \alpha_6) = \\
 [0236] \quad & = \begin{pmatrix} \cos \alpha_4 \cos \alpha_6 - \sin \alpha_4 \cos \alpha_5 \sin \alpha_6 & -\sin \alpha_4 \cos \alpha_6 - \cos \alpha_4 \cos \alpha_5 \sin \alpha_6 & \sin \alpha_5 \sin \alpha_6 & 0 \\ \cos \alpha_4 \sin \alpha_6 + \sin \alpha_4 \cos \alpha_5 \cos \alpha_6 & -\sin \alpha_4 \sin \alpha_6 + \cos \alpha_4 \cos \alpha_5 \cos \alpha_6 & -\sin \alpha_5 \cos \alpha_6 & 0 \\ \sin \alpha_4 \sin \alpha_5 & \cos \alpha_4 \sin \alpha_5 & \cos \alpha_5 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}
 \end{aligned}$$

[0237] 在上文所说明的四个变换候选者中,一个候选者对应于不应用二次变换,且其它三个候选者对应于由 α_1 、 α_2 、 \dots 、 α_6 的三个不同设定产生的 $R_{vertical}$ 及 $R_{horizontal}$ 。更多详细描述可在2010年4月15日到23日,德国德累斯顿,ITU-T SG16WP3及ISO/IEC JTC1/SC29/WG11的视频译码联合合作小组(JCT-VC),JCTVC-A124,K.麦肯(McCann),W.-J.汉,I.-K.基姆(Kim),J.-H.米恩(Min),E.阿史那,A.阿施恩,T.李,J.陈,V.赛瑞珍,S.李,Y.-M.匡(Hong),M.-S.钟(Cheon),N.石亚科沃(Shlyakhov)的“由三星(及BBC)进行的视频译码技术提议(Video coding technology proposal by Samsung(and BBC))”中发现。

[0238] 图4为通过视频解码器(例如视频解码器30)应用的实例ROT的说明。在图4的实例中,视频解码器对大小在 4×4 到 32×32 范围内的变换系数块62执行逆量化60(Inv Q)。此外,在图4的实例中,视频解码器仅对于变换系数矩阵的低频部分应用逆ROT 64(Inv ROT)作为二次变换。举例来说,如图4的实例中所示, 4×4 及 8×8 变换系数矩阵的所有变换系数可被认为在变换系数矩阵的低频部分中。然而,如图4中所示,仅 16×16 及 32×32 变换系数矩阵的左上 8×8 子块中的变换系数被考虑为在变换系数矩阵的低频部分中。此外,在图4的实例中,视频解码器将逆DCT变换66(Inv DCT变换)应用于变换系数矩阵以将变换系数矩阵转换成残差块。在图4的实例中,在将逆DCT变换66应用于 16×16 及 32×32 变换系数矩阵之前,视频解码器可使用由应用逆ROT 64产生的对应 8×8 变换系数矩阵替换 16×16 及 32×32 变换系数矩阵的左上 8×8 子块。

[0239] 如2015年6月,ITU-T SG16Doc.VCEG-AZ05,E.阿史那,A.阿施恩,J.-H.米恩,K.蔡(Choi),A.萨克西纳,M.布达嘉威(Budagavi)的“针对下一代视频译码的已知工具性能调查(Known tools performance investigation for next generation video coding)”中所描述,ROT可经扩展到整个TU,不仅仅低频部分。更特定来说,三个 4×4 可分离二次变换候选者可经预定义,且所选择候选者可在CU层级处以2位索引明确地传信。2位索引可在本文中称为ROT索引。在一个实例中,当ROT索引为0时,不应用二次变换。然而,在此实例中,当ROT索引为1、2及3时,对应于三个预定义二次变换候选者中的一者的二次变换可应用于当前CU

内部的每一TU。给定所选择二次变换,对于当前TU的每一 4×4 子块,可应用二次 4×4 可分离二次变换。

[0240] 二次变换的设计可将可分离变换应用为二次变换。然而,由于非可分离变换可在2-D图像块上提供优良译码增益,因此可改进二次变换效率。换句话说,在视频译码的情形中,变换系数可大致不相关且稀疏,使得输入向量x的能量可仅仅精简在几个变换系数上,且剩余大部分变换系数可通常接近于0。在实例中,KLT(其为非可分离变换)可具有最佳能量精简性,此是由于其使用输入数据的协方差矩阵的特征向量作为变换基础向量而非仅将系数从像素域变换到频域。

[0241] 更特定来说,可分离变换可捕获沿水平及/或垂直方向的像素相关性。对比来说,非可分离变换可捕获2维图像块中的任一对两个像素之间的相关性。因此,非可分离变换可具有比可分离变换更多的减少数据相关性的灵活性。举例来说,对于展示沿非水平或非垂直方向(例如,45度边缘纹理)的相关性的图像块,可分离变换可能不会高效地减少沿45度方向的像素之间的相关性,而非可分离(变换)可有效减少相关性。

[0242] 鉴于上文所论述问题,提出以下技术。应理解下文所描述的本发明的技术可结合用于自适应性交叉分量残差预测的技术(例如在第15/006,994号美国专利申请案及第62/108,491号美国临时申请案中描述的那些技术)一起使用。下文描述在本发明中提出的详细列举技术。视频编码器20及/或视频解码器30可个别地应用详细列举技术。替代地,视频编码器20及/或视频解码器30可应用详细列举技术的任何组合。

[0243] 根据第一技术,视频编码器20及/或视频解码器30经配置以在量化与变换过程之间应用非可分离二次变换。举例来说,应用于平坦模式的非可分离变换集合(包含三个非可分离变换)可为如下。

[0244]

246	-36	-35	-2	-41	9	6	0	-27	6	6	-1	-2	0	-1	0
-42	-220	39	26	-100	20	18	-3	28	38	-9	-6	5	-2	-2	1
-18	106	3	-17	-222	-16	51	4	24	-19	-6	3	23	1	-7	0
-30	-36	-150	17	-11	-181	6	30	-59	17	45	-4	9	25	-2	-5
-30	-3	-192	13	-8	150	2	-30	50	4	29	0	5	-24	-2	6
-31	-7	18	7	-24	76	13	-15	-230	-33	49	4	26	-7	-6	3
3	22	10	16	47	8	208	1	-6	127	2	-28	13	-6	-43	1

[0245]

-2	38	11	-11	-28	11	-129	-6	-23	204	10	-42	46	-3	10	3
11	21	8	215	6	-2	-18	5	11	-31	-41	-37	116	9	-23	3
-4	25	4	117	-27	1	-5	3	-21	28	4	-34	-209	-38	54	7
-9	-7	-48	-33	-2	-13	0	-29	-55	6	-233	-8	-7	-55	-5	17
4	-4	14	-3	3	-36	-7	5	13	-11	54	11	32	-238	-27	49
3	1	10	7	0	-44	5	-246	8	-5	27	-7	1	16	7	42
-2	0	0	-2	-8	0	-42	-9	0	-8	4	-30	-61	16	-241	-22
1	8	1	48	-3	-1	-7	-8	-5	51	-13	243	-12	7	-28	2
0	0	0	-5	0	7	-2	44	1	-1	2	-3	-8	51	-17	246
-242	48	31	1	51	-7	-8	-1	26	-10	-3	1	1	1	1	0
-60	-155	43	12	-173	30	37	-2	45	34	-14	-4	15	-7	-3	1
-4	-187	13	28	161	-21	-30	2	-25	32	0	-6	-23	6	6	-1
-31	-10	-63	12	-49	-189	37	26	-120	30	51	-7	25	34	-12	-5
22	25	225	-16	-1	-10	20	-3	-111	-1	-20	2	1	11	-4	-1
38	12	68	-12	14	-154	4	27	175	6	-52	3	-24	18	1	-6
6	41	-6	10	44	31	118	-12	18	198	-23	-42	50	-21	-46	8
2	-19	-9	26	36	9	212	-4	10	-125	10	11	-42	1	-22	3
10	-15	7	41	24	-6	-10	3	16	-71	-34	9	232	0	-50	5
-7	-33	-16	-241	17	-2	24	-27	1	-1	12	53	37	6	-17	1
-10	-4	-44	-10	-9	-32	1	-33	-59	-16	-212	5	-26	-100	14	38
5	-2	27	-3	5	-30	-2	27	15	-5	104	5	13	-224	16	40
4	1	9	24	-4	-37	-11	-243	17	-6	41	-18	0	4	-3	45
3	4	0	2	9	1	44	-6	3	18	-1	19	58	20	243	6
1	7	0	54	0	0	2	-13	0	43	2	245	-10	1	-22	5
0	0	2	-4	0	8	-4	47	3	0	9	-3	-5	53	-7	245
-251	21	33	5	18	-7	1	-1	27	-1	-6	0	6	-1	0	0
23	194	-14	-21	152	40	-28	-9	-4	-32	-5	4	-14	-8	4	2
0	-150	-31	19	192	-10	-49	0	26	33	-2	-4	-20	2	7	0
-19	20	-120	-2	-7	-187	-56	27	-76	-62	29	16	-3	22	14	-3
26	-5	203	19	47	-102	41	26	-73	-44	-26	4	-19	8	-2	-4
33	34	24	-5	-10	-105	-23	17	212	19	-60	-3	34	23	-10	-4
1	45	-6	1	22	-53	79	15	-32	205	74	-34	23	58	-4	-16
0	15	59	-8	-41	10	-210	-34	-49	89	-46	-29	4	18	37	2
3	24	-23	230	-3	23	-10	71	-12	-5	-41	-41	45	6	-17	-13
1	-18	1	-50	28	15	2	-21	-41	-30	-34	9	226	58	-57	-8
8	-2	53	33	-3	25	-55	25	51	-48	203	45	45	49	69	-1
-4	-6	-14	-4	-3	34	26	-16	0	-40	-55	-18	-62	223	55	-41
5	4	2	65	3	-41	24	-221	12	-26	34	-84	8	-7	8	34
-1	0	13	-1	-7	6	-50	-8	8	-16	60	-18	-60	44	-223	-57
1	8	-2	58	-2	-6	2	-67	-7	43	-28	219	-15	32	-34	65
-1	-1	1	-14	0	7	-6	52	0	-8	10	-60	-10	49	-40	234

[0246] 在此实例的一个方面中,非可分离二次变换可为KLT,其可从离线训练导出,或使用假定图像相关模型以分析形式离线导出。在一些实例中,图像相关模型为函数 $f(x, y)$,其量测随机向量的第 i 与第 j 个元素之间的协方差。随机向量可为具有多个维度的随机变数。换句话说,图像相关函数可为对称的: $f(x, y) = f(y, x)$,且协方差矩阵 $\text{Cov}(x, y) = f(x, y)$ 可为半正定。相关模型的实例为 $f\{x*y\} = \rho^{|x-y|}$,其中 $0 \leq \rho \leq 1$ 。

[0247] 在此实例的一个方面中,非可分离二次变换可为KLT,其可从离线训练导出,或使用在编码/解码过程期间在线导出的KLT以分析形式离线导出。在一些情况下,在编码/解码过程期间,视频编码器20及/或视频解码器30经配置以收集变换系数的统计数据,且更新(其中应用二次变换)变换系数的相关矩阵。基于经更新相关矩阵,视频编码器20及/或视频解码器30可经配置以提取本征向量,并将KLT变换核心作为由按其对应本征值的次序布置

的所有本征向量组成的矩阵导出。

[0248] 根据第一技术的实例,视频编码器20可形成残差视频块。在一些实例中,作为形成残差视频块的部分,视频编码器20从正被译码的视频块减去一或多个预测性块。另外,视频编码器20可将第一变换应用于残差视频块以产生第一系数块。在此实例中,第一变换将残差视频块从像素域转换到频域。如本文所使用,像素域或空间域可指其中像素的值的变化的对应于那个像素的明度及/或色彩的变化域。然而,如本文所使用,频域可指其中像素的值的变化的对应于像素的像素值在像素域中改变的速率的变化域。此外,视频编码器20可将二次变换应用于第一系数块的至少部分以产生第二系数块。根据本发明的第一技术,第二变换为非可分离变换,例如KLT。视频编码器20接着可量化第二系数块。

[0249] 类似地,根据第一技术的实例,视频解码器30可逆量化第一系数块。另外,在此实例中,视频解码器30可将第一逆变换(即,逆二次变换)应用于经逆量化第一系数块的至少部分以产生第二系数块。根据本发明的第一技术,第一逆变换为非可分离变换,例如KLT。此外,在此实例中,在应用第一逆变换以产生第二系数块之后,视频解码器30可将第二逆变换应用于第二系数块以产生残差视频块。在此实例中,第二逆变换将第二系数块从频域转换到像素域。另外,在此实例中,视频解码器30可形成经解码视频块。在一些实例中,作为形成经解码视频块的部分,视频解码器30将残差视频块与一或多个预测性块求和。举例来说,视频解码器30可将残差视频块与一或多个预测性块求和,并基于残差视频块与一或多个预测性块的和形成经解码视频块。

[0250] 根据第二技术,视频编码器20及/或视频解码器30执行从三个或三个以上非可分离变换候选者的预先选择以调配非可分离变换的子集。在一些实例中,非可分离变换的子集可指KLT,其通过离线训练而导出,且针对编码器20及/或解码器30两者,变换核心经硬译码为固定数目。视频编码器20及/或视频解码器30可从变换的子集中选择用于当前TU的最终变换。

[0251] 举例来说,视频编码器20可调配包含三个或三个以上非可分离变换候选者的非可分离变换的集合的子集。在此实例中,视频编码器20可从非可分离变换的集合的子集中选择一个候选者作为第一逆变换。在此实例的一些方面中,非可分离变换的集合包含十二个变换子集,十二个变换子集包含非可分离变换的集合的子集。在此实例的一些方面中,非可分离变换的集合是由视频数据的明度帧内预测模式指定。举例来说,视频编码器20可指定视频数据的第一明度帧内预测模式与非可分离的变换的第一集合一起使用,并可指定视频数据的第二明度帧内预测模式与非可分离的变换的第二集合一起使用。更特定来说,视频编码器20可通过传信的索引(例如,指示索引的数据)及所选择变换子集指定变换候选者。

[0252] 类似地,视频解码器30可调配包含三个或三个以上非可分离变换候选者的非可分离变换的集合的子集。在此实例中,视频解码器30可从非可分离变换的集合的子集中选择一个候选者作为第二变换。在此实例的一些方面中,非可分离变换的集合包含十二个变换子集,十二个变换子集包含非可分离变换的集合的子集。在此实例的一些方面中,非可分离变换的集合是由视频数据的明度帧内预测模式指定。

[0253] 在实例中,用以调配变换的子集的预先选择可由已经解码信息(例如帧内预测模式、EMT CU层级及/或TU层级索引、RQT深度、经量化系数、参考预测块、PU内部的当前TU的相对位置(TU定位于当前PU的哪一边界)、块大小、块形状(其为正方形抑或非正方形块,及高

度对宽度的比率))确定。换句话说,视频解码器30可根据经解码信息而非依赖于单独或额外传信预先选择变换的子集。举例来说,视频编码器20可基于经解码信息确定子集,解码信息包括帧内预测模式、CU层级EMT索引、TU层级EMT索引、残差四分树深度、经量化系数、参考预测块及PU内部的当前TU的相对位置中的一或多个。在此实例的一些方面中,应用十二个变换子集,且指定从帧内预测模式到变换子集选择的映射的查找表经预定义,给定帧内预测模式及此查找表,首先选择变换子集(其包含三个不同非可分离变换),接着进一步通过经解码信息指定三个非可分离变换中的一者并将其用以执行第一变换。类似地,在此实例的一些方面中,视频解码器30可基于以下各者中的一或多个确定子集:帧内预测模式、CU层级EMT索引、TU层级EMT索引、残差四分树深度、经量化系数、参考预测块、单元PU内部的当前TU的相对位置、块大小,及块形状(其为正方形抑或非正方形块,及高度对宽度的比率)。

[0254] 根据第二技术的一些实例,变换子集的数目可限于小整数,例如,1、2、3、4、12或34。另外,在此实例的一些方面中,不同变换子集可含有不同类型的变换。举例来说,视频编码器20可预先选择具有一或多个左变换的第一子集、具有一或多个右变换的第二集合或具有左及右变换的第三集合作为预先选择的子集。接着,在实例中,视频编码器20可从预先选择的子集选择一个变换候选者作为所选择变换。举例来说,视频编码器20可使用编码器模式决策选择变换,且视频编码器20编码指示变换的索引的数据。类似地,视频解码器30可从位流解码索引并基于经解码索引从预先选择的变换的集合中选择变换。

[0255] 在一些实例中,对应于不同帧内预测模式的非可分离变换的集合的变换子集中的变换候选者的总数是不同的。举例来说,视频编码器20及/或视频解码器30可使用用于第一帧内预测模式的第一子集及用于第二帧内预测模式的第二子集。

[0256] 在实例中,存在总共12个变换子集。在此实例中,12个变换子集中的每一各别变换子集含有三个非可分离变换候选者。变换子集可通过明度帧内预测模式指定,且对于不同帧内模式,可应用同一变换集合。如本文所使用,帧内预测模式可指用于正被预测的块的各种预测方向。帧内预测模式的实例可包含(但不限于)垂直、水平、对角线左下、对角线右下、垂直右、水平下、垂直左、水平上及用于正被预测块的其它预测方向。换句话说,在此实例的一些方面中,非可分离的变换的集合通过视频数据的明度帧内预测模式指定。举例来说,视频解码器30可选择12个变换子集中的一个变换子集用于特定明度帧内预测模式。在实例的方面中,视频解码器30可选择所选择子集中所含有的三个非可分离变换候选者中的第一变换以用于第一帧内预测模式,并可选择所选择子集中所含有的三个非可分离变换候选者中的第二变换以用于第二帧内预测模式。

[0257] 在实例中,存在总共12个变换子集,且每一变换子集可含有五个非可分离变换候选者。变换集合可通过明度帧内预测模式及EMT CU层级及TU层级索引指定。对于不同帧内模式,可应用同一变换集合。

[0258] 在实例中,变换子集可含有仅仅一个变换。对于不同TU大小,变换子集中的变换的数目可不同,典型数目可为(但不限于)2、3及4。对于不同帧内预测模式,变换子集中的变换的数目可不同,典型数目可为(但不限于)2、3及4。在实例的方面中,对于平坦或DC预测模式,变换子集中的变换的数目为2,而对于其它预测模式,变换子集中的变换的数目为3。

[0259] 图5及图6涉及本发明的第三技术。图5为说明包含二次变换系数重组过程的实例编码过程的一部分的框图。图6为说明包含二次变换系数重组过程的实例解码过程的一部

分的框图。在图5中,视频编码器20将一次变换100应用于残差块101以产生2维系数块102。此外,图5的编码过程包含非可分离二次变换106。非可分离二次变换106可采取1维系数向量156作为输入并产生1维系数向量108作为输出。因此,在图5的实例中,一次变换100产生2维系数块102,但非可分离二次变换106采用1维系数向量作为输入。因此,视频编码器20可执行重组过程104以将2维系数块102变换成1维系数向量105。在一些实例中,重组过程104可认为是非可分离二次变换106的部分。

[0260] 此外,量化过程114可采取2维系数块112作为输入并可输出经量化2维系数块116。因此,根据本发明的第三技术,视频编码器20可在视频编码器20在编码期间应用二次变换106之后应用二次变换系数重组过程110以用于将通过非可分离二次变换106导出的1维系数向量108重组为2维系数块112。应用重组过程110可排除修改量化过程106以接纳1维系数向量的需要。

[0261] 在图6中,视频解码器30可将逆量化150应用于经量化2维系数块151以导出2维系数块152。此外,图6的解码过程包含逆二次变换158。逆二次变换158可采取1维系数向量156作为输入并输出1维系数向量160。举例来说,逆二次变换158可采取 16×1 系数向量作为输入并可输出 16×1 系数向量。然而,逆量化150可输出2维系数块152,且一次逆变换166可采取2维系数块164作为输入。因此,根据本发明的第三技术,在视频解码器30在解码过程期间应用二次逆变换158之前,视频解码器30可应用二次变换系数重组过程154以用于将通过逆量化150导出的2维系数块152重组为1维系数向量156。在应用逆二次变换158之前应用重组过程154可排除修改逆量化150以产生1维系数向量的需要。此外,视频解码器30可应用二次变换系数重组过程162以用于将1维系数向量160重组为2维系数块164。视频解码器30接着可将一次逆变换166应用于2维系数块164以产生残差块168。在应用一次逆变换166之前应用重组过程162可排除修改一次逆变换166的需要。在一些实例中,重组过程162被认为二次逆变换158的一部分。

[0262] 因此,根据第三技术,视频编码器20可形成残差视频块。作为形成残差视频块的部分,视频编码器20可从正被译码的视频块减去一或多个预测性块。另外,视频编码器20可将第一变换(例如,一次变换100)应用于残差视频块的至少部分以产生第一2维系数块(例如,2维系数块102)。在此实例中,第一变换将残差视频块从像素域转换到频域。另外,视频编码器20可将第一2维系数块重组为第一1维系数向量。在实例中,视频编码器20可将第二变换应用于第一1维系数向量的至少部分以产生第二1维系数向量。在此实例中,第二变换为非可分离变换。在实例中,第二变换为非可分离变换。在此实例中,视频编码器20可根据系数扫描次序将第二1维系数向量重组为第二2维系数块。

[0263] 类似地,视频解码器30可根据系数扫描次序将第一2维系数块(例如,2维系数块152)重组为第一1维系数向量(例如,1维系数向量156)。在此实例中,视频解码器30可通过将非可分离变换矩阵乘以第一1维系数向量以产生第二1维系数向量(例如,1维系数块160)而应用第一逆变换(例如,非可分离二次逆变换158)。在此实例中,第一逆变换为非可分离变换。在此实例中,视频解码器30可将第二1维系数向量重组为2维系数块(例如,2维系数块164)。视频解码器30可将第二逆变换(例如,一次逆变换166)应用于第二2维系数块以产生残差视频块(例如,残差块168)。在此实例中,第二逆变换将第二2维系数块从频域转换到像素域。在此实例中,视频解码器30可形成经解码视频块。作为形成经解码视频块的部分,视

频解码器30可将残差视频块与一或多个预测性块求和。

[0264] 在第三技术的一些实例中,应用第二变换包括将2维变换矩阵乘以第一1维系数向量。在实例的一些情况下,第一2维系数块为 4×4 块且2维变换矩阵为通过明度帧内预测模式及译码单元层级索引选择的 16×16 非可分离变换矩阵。举例来说,视频编码器20可通过将 16×16 非可分离变换矩阵乘以第一1维系数向量(例如,1维系数向量105)而应用第二变换(例如,非可分离二次变换106),借此产生第二1维系数向量(例如,1维系数向量108)。类似地,视频解码器30可通过将 16×16 非可分离变换矩阵乘以第一1维系数向量(例如,1维系数向量156)而应用第二变换(例如,非可分离二次变换158),借此产生第二1维系数向量(例如,1维系数向量160)。

[0265] 在第三技术的一些实例中,对 4×4 系数群执行的二次变换系数重组过程可取决于 4×4 系数群的系数扫描次序。举例来说,如果视频编码器20在熵编码期间使用Z形扫描次序以处理表示 4×4 系数群的系数的语法元素,则视频编码器20可在作为应用非可分离二次变换的部分将 4×4 系数群重组为 16×1 系数向量时使用相同Z形扫描次序。

[0266] 此外,在第三技术的一些实例中,视频编码器20可确定用于 4×4 系数群的系数扫描次序。在此实例中,视频编码器20可基于经确定系数扫描次序将第一1维系数向量重组为第一2维系数块(例如, 4×4 系数群)。在此实例的一些方面中,视频编码器20可将第一1维系数向量重组为第一2维系数块以使得第一2维系数块具有对应于(例如,匹配)经确定系数扫描次序的系数扫描次序。

[0267] 类似地,视频解码器30可确定用于 4×4 系数群的系数扫描次序。在此实例中,视频解码器30可基于经确定系数扫描次序将第二2维系数块(例如, 4×4 系数群)重组为第二1维系数向量。在此实例的一些方面中,视频解码器30可将第二2维系数块重组为第二1维系数向量,以使得第二1维系数向量具有对应于(例如,匹配)经确定系数扫描次序的系数扫描次序。

[0268] 在第三技术的一些实例中,对 4×4 系数群执行的二次变换系数重组过程可取决于已编码/解码信息,例如帧内预测模式、EMT CU层级及/或TU层级索引、RQT深度及/或经量化系数。换句话说,视频编码器20的重组过程110可基于帧内预测模式、EMT CU层级及/或TU层级索引、RQT深度及/或经量化系数将1维系数向量108重组为2维系数块112。举例来说,视频编码器20可基于选择垂直明度预测模式抑或水平明度预测模式而将1维系数向量108重组为2维系数块112,以使得1维系数向量具有对应(例如,匹配)经确定系数扫描次序的系数扫描次序。类似地,视频解码器30的重组过程154可基于帧内预测模式、EMT CU层级及/或TU层级索引、RQT深度及/或经量化系数重组2维系数块152以产生1维系数向量156。

[0269] 在第三技术的一个实例中,在将非可分离二次变换(例如,二次变换102)应用于特定 4×4 系数块之后,视频编码器20可导出 16×1 系数向量(例如,1维系数向量108)并可根据(即,沿)用于特定 4×4 块的系数扫描次序将所得16个系数重组为 4×4 块(例如,2维系数块112)。视频编码器20可以具有较小索引的系数放置在 4×4 系数块中的具有较小扫描索引的位置处的方式执行16个系数的重组。举例来说,当逐行从 4×4 系数块的顶部到 4×4 系数块的底部扫描 4×4 系数块(例如,根据帧内预测模式)时,视频编码器20可将系数块的顶部列的四个系数重组为1维系数向量的前四个系数,将系数块的次顶部行的四个系数重组为1维系数向量的接下来四个系数,等等。

[0270] 在对TU系数执行解量化之后,视频解码器30可对于当前TU的每一 4×4 子块导出 4×4 经解量化二次变换系数块(例如,2维系数块152)。此外,对于当前TU的每一各别 4×4 子块,视频解码器30可基于用于 4×4 块的系数扫描次序将各别 4×4 经解量化块中的16个系数重组为各别 16×1 向量(例如,1维系数向量156)。视频解码器30可以具有较小扫描索引的系数放置在 16×1 系数向量中的具有较小索引的位置处的方式执行 4×4 系数的重组。

[0271] 在实例中,视频编码器20及/或视频解码器30可基于固定 4×4 系数扫描次序(例如,水平、垂直、对角线或Z形扫描次序)执行重组过程。以此方式, 16×1 系数向量中的具有较小索引的16个二次变换系数可根据固定预定义 4×4 系数扫描次序放置在 4×4 系数块中。

[0272] 根据第三技术的实例,在变换过程(例如,一次变换100)之后,视频编码器20可通过首先使用 $\vec{b}(i \cdot N + j) = B(i, j)$ 将B表示为向量 \vec{b} 且接着执行以下计算而将二次变换(例如,二次变换102)应用于当前 $M \times N$ TU的每一各别 4×4 变换系数子块B:

$$[0273] \quad \vec{f} = T \cdot \vec{b}.$$

[0274] 在上述等式中,T为基于如上文所描述的明度帧内预测模式及传信的CU层级旗标而选择的 16×16 非可分离变换矩阵。此外,在图5的实例中, \vec{b} 可对应于1维系数向量105。由于应用上述等式,因此对于每一各别 4×4 子块,视频编码器20导出各别 16×1 二次变换系数向量 \vec{f} 。在图5的实例中, \vec{f} 对应于1维系数向量108。视频编码器20可通过执行以下操作将各别 16×1 二次变换系数向量重组为各别 4×4 系数块:

$$[0275] \quad F(\text{scan_x}(j), \text{scan_y}(j)) = \vec{f}(j).$$

[0276] 在上述等式中, $j=0,1,\dots,15$,scan_x及scan_y可分别为如已在HEVC中定义的x及y座标的扫描表。以此方式,各别 4×4 变换系数子块B可由各别导出的二次 4×4 变换系数块F替换。在图5的实例中,F对应于2维系数块112。在视频编码器20已执行用于每一 4×4 变换系数子块的二次变换之后,视频编码器20可进一步量化并熵编码经更新 $M \times N$ 个系数块。

[0277] 对于每一各别 4×4 系数子块F,在量化过程(例如,逆量化150)之后,视频解码器30可通过将各别 4×4 系数子块F重组为 16×1 向量而导出 16×1 二次变换系数向量 \vec{f} 。

$$[0278] \quad \vec{f}(j) = F(\text{scan_x}(j), \text{scan_y}(j)).$$

[0279] 在此上述等式中, $j=0,1,\dots,15$,scan_x及scan_y分别为如已在HEVC中定义的 4×4 系数子块F的x及y座标的扫描表。在图6的实例中, \vec{f} 对应于1维系数向量156且F对应于2维系数块152。视频解码器30接着可通过执行以下计算对 \vec{f} 应用逆二次变换(例如,二次逆变换158):

$$[0280] \quad \vec{b} = T' \cdot \vec{f},$$

[0281] 在上述等式中,T'为可通过如上文中所描述的明度帧内预测模式及传信的CU层级旗标选择的 16×16 非可分离变换矩阵。由于上述等式的矩阵乘法,对于每一各别 4×4 子块i,视频解码器30导出各别 16×1 二次变换系数向量 \vec{b} 。在图6的实例中, \vec{b} 对应于1维系数向量160。视频解码器30通过计算 $B(i, j) = \vec{b}(i \cdot N + j)$ 而将 16×1 二次变换系数向量重组为 $4 \times$

4系数块。在图6的实例中,B对应于2维系数块164。视频解码器30可通过导出的二次 4×4 变换系数块B替换 4×4 变换系数子块F。在视频解码器30对于 4×4 系数子块中的每一者执行逆二次变换之后,视频解码器30可输入经更新 $M\times N$ 个系数块到逆变换(例如,一次逆变换166)。

[0282] 如上文所指示,在一些实例中,视频编码器20可选择非可分离二次变换。根据第四技术,视频编码器20可传信(例如,编码)使用适用于TU、PU、CU或任何图像译码块单元的索引对非可分离二次变换的选择的指示且视频解码器30可接收所述指示。换句话说,视频编码器20可针对TU、PU、CU或其它图像译码块单元编码指示出指示所选择非可分离二次变换的索引的数据。举例来说,可通过用于TU、PU、CU或其一组合的索引传信对来自非可分离变换的集合的第一逆变换的选择。类似地,视频解码器30可从视频数据的位流获得指示TU、PU、CU或其一组合的索引的数据,所述索引指示所选择非可分离二次变换。为易于解释,本发明可将指示所选择非可分离二次变换的索引称为非可分离二次变换(NSST)索引。

[0283] 在一些实例中,视频编码器20可熵编码NSST索引。作为熵编码NSST索引的部分,视频编码器20可二进制化NSST索引。为了熵译码NSST索引,固定长度码可用于二进制化,或可变长度码(例如截短一元码或指数哥伦布码)可用于二进制化。

[0284] 根据第五技术,如果残差信号的能量受到限制,则视频编码器20可对于某些TU跳过NSST索引的传信,且可不应用二次变换。举例来说,如果不存在针对当前TU所发射的非零系数,则视频编码器20可针对当前TU跳过NSST索引的传信。换句话说,如果不存在针对当前TU编码的非零系数,则视频编码器20可跳过指示当前TU的NSST索引的数据的编码。类似地,在其它实例中,NSST索引传信的跳过可适用于LCU、CU、PU或任何其它块层级语法元素。因此,根据第五技术,视频编码器20可在视频数据的编码位流中产生指示适用于TU、PU、CU或其一组合的索引的数据,所述索引指示所选择候选者为第二变换。

[0285] 类似地,视频解码器30可从视频数据的位流获得指示适用于TU、PU、CU或其一组合的索引的语法元素。在此实例中,视频解码器30可从非可分离变换的子集中选择候选者,包括选择通过索引指示的候选者作为第一逆变换。

[0286] 在一些实例中,如果在某一块层级处发射的非零系数的总数或总绝对和或平方值的和不大于给定阈值,则跳过在那个某一块层级处的NSST索引且不应用二次变换。举例来说,当在某一LCU处所发射的非零系数的总数不大于10时,视频编码器20可跳过NSST索引的传信且不应用二次变换。在一些实例中,当在某一CU处所发射的非零系数的总绝对和不大于100时,视频编码器20可跳过NSST索引的传信且不应用二次变换。在一个实例中,当在某一PU处所发射的非零系数的平方值的和不大于100时,视频编码器20可跳过NSST索引的传信且不应用二次变换。

[0287] 因此,根据第五技术,视频编码器20可确定是否在位流中编码二次变换(例如,非可分离二次变换106)的索引。二次变换的索引可从多个可用二次变换中识别二次变换。在此实例中,视频编码器20可基于在特定块层级(例如,在LCU、CU、PU、TU中)处所发射的非零系数的总数的和、总绝对和或平方值和不大于给定阈值进行不在位流中编码二次变换的索引的确定。在此实例中,视频编码器20可在特定块层级处传信二次变换的索引。在此实例中,视频编码器20可将其索引在位流中传信的二次变换应用于2维系数块的至少部分。

[0288] 类似地,视频解码器30可确定是否在位流中传信二次变换(例如,非可分离逆变换

158)的索引。在此实例中,视频解码器30可基于在特定块层级处的非零系数的总数的和、总绝对和或平方值和不大与给定阈值进行不在位流中编码二次变换的索引的确定。在此实例中,二次变换的索引系在特定块层级处传信。在此实例中,视频解码器30可将为由在位流中传信的索引指示的二次变换的逆的二次逆变换应用于1维系数向量。

[0289] 在一个实例中,视频解码器30可响应于视频编码器20确定2维系数块的非零系数的量不大与阈值而将二次逆变换应用于1维系数向量。在此实例中,针对视频数据的不同块大小、不同帧内预测模式或其一组合,阈值可不同。

[0290] 类似地,视频编码器20可确定2维系数块的非零系数的量是否大于阈值。在此实例中,视频编码器20可响应于确定2维系数块的非零系数的量不大与阈值而将二次变换应用于2维系数块的至少部分。在此实例中,针对视频数据的不同块大小、不同帧内预测模式或其一组合,阈值可不同。

[0291] 在一些实例中,非零系数的总数的阈值为1。在实例中,非零系数的总数的阈值针对不同块大小或不同帧内预测模式可为不同的。

[0292] 在一些实例中,视频编码器20及视频解码器30可将二次变换及二次逆变换应用于所有色彩分量(例如,Y、Cb及Cr)的系数值。在其它实例中,视频编码器20及视频解码器30可将二次变换及二次逆变换应用于一些但并非所有色彩分量。举例来说,视频编码器20及视频解码器30可将二次变换及二次逆变换应用于明度(即,Y)分量的系数值,但并不应用于色度(例如,Cb及Cr)分量的系数值。在视频编码器20及视频解码器30将二次变换及二次逆变换应用于两个或大于两个色彩分量的实例中,两个或大于两个色彩分量可共享NSST索引。

[0293] 根据第五技术的实例,如果NSST索引在某一块层级处被传信并共享于若干分量(例如,Y、Cb、Cr)且如果来自共享相同NSST索引的所有分量的在那个某一块层级处发射的非零系数的总数的和或总绝对和或平方值的和不大与给定阈值,则NSST索引可被跳过且不应用二次变换。在实例中,非零系数的总数的阈值为1、2或3。在一些实例中,非零系数的总数的阈值大于3。

[0294] 在实例中,如果NSST索引是在某一块层级处传信并经共享于Y、Cb及Cr分量,且如果Y、Cb及Cr分量的非零系数的总数和小于给定阈值,则NSST索引可被跳过且可不应用二次变换。举例来说,当Y、Cb及Cr分量的组合的在某一CU处发射的非零系数的绝对和不大与100时,视频编码器20可跳过NSST索引的传信且不应用二次变换。

[0295] 在实例中,如果NSST索引是在某一块层级处传信并经共享于Cb及Cr分量,且如果Cb及Cr分量的非零系数的总数和小于给定阈值,则视频编码器20可跳过NSST索引的传信且不应用二次变换。举例来说,当某一CU的Cb及Cr分量的非零系数的总数不大与3时,视频编码器20可跳过NSST索引的传信且不应用二次变换。

[0296] 根据第五技术的一些实例,非零系数的总数的阈值针对不同块大小或不同帧内预测模式可为不同的。举例来说,LCU可具有比CU及PU大的阈值。类似地,视频编码器20可将第一阈值用于使用垂直帧内预测模式译码的块,且可将第二不同阈值用于使用水平帧内预测模式译码的块。

[0297] 在第五技术的实例中,如果NSST索引是在某一块层级处传信且经共享于Y、Cb及Cr分量,且如果仅仅Y分量的非零系数的总数小于给定阈值,则可跳过NSST索引且可不应用二次变换。举例来说,当仅仅Y分量的在某一CU处所发射的非零系数的总数不大与1、2、3或另

一阈值时,视频编码器20可跳过NSST索引的传信且不应用二次变换。

[0298] 在实例中,如果NSST索引是在某一块层级处传信且经共享于Cb及Cr分量,且如果仅仅Cb分量的非零系数的总数小于给定阈值,则可跳过NSST索引且可不应用二次变换。举例来说,当仅仅Cb分量的在某一CU处所发射的非零系数的总数不大于1、2、3或另一阈值时,视频编码器20可跳过NSST索引的传信且不应用二次变换。类似地,当仅仅Cr分量的在某一CU处所发射的非零系数的总数不大于1、2、3或另一阈值时,视频编码器20可跳过NSST索引的传信且不应用二次变换。

[0299] 另外或替代地,在第五技术的一些实例中,非零系数的总数的阈值针对不同块大小或不同帧内预测模式可系不同的。换句话说,LCU可具有比CU及PU大的阈值。类似地,16×16色度预测模式可具有比4×4色度预测模式大的阈值。举例来说,视频编码器20可将第一阈值用于使用垂直帧内预测模式译码的块,且可将第二不同阈值用于使用水平帧内预测模式译码的块。

[0300] 在第五技术的实例中,当LCU、CU、PU或其它的类型块的大小大于或小于预定义阈值,或在给定阈值范围内时,可跳过NSST索引的传信且不应用二次变换。举例来说,当CU的大小小于8×8的预定义值时,视频编码器20可跳过在经编码视频位流中产生指示NSST索引的数据且不应用二次变换。

[0301] 在第五技术的实例中,当使用变换跳过模式译码TU时,可跳过NSST索引的传信且不应用二次变换。换句话说,当使用变换跳过模式译码TU时,视频编码器20可跳过在经编码视频位流中产生指示NSST索引的数据且不应用二次变换。

[0302] 举例来说,视频编码器20可基于正在变换跳过模式中译码TU确定是否在位流中传信二次变换的索引(例如,编码指示所述索引的数据)。在此实例中,视频编码器20在经编码视频位流中产生指示位流中的二次变换的索引的数据,且视频编码器20可基于正在位流中编码指示第二变换的索引的数据而将二次变换应用于第一2维系数块的至少部分。类似地,视频解码器30可基于正在变换跳过模式中译码TU确定二次变换的索引是否在位流中传信。换句话说,视频解码器30可通过二次变换的索引指示的二次逆变换应用于1维系数向量。

[0303] 此外,在第五技术的一些实例中,当使用二次变换译码块(例如,TU)时,可不传信变换跳过模式。换句话说,视频编码器20可在经编码视频位流中产生指示变换跳过模式及二次变换中的一者的数据。举例来说,视频编码器20可在经编码视频位流中产生指示变换跳过模式的数据,并省略传信二次变换。类似地,视频编码器20可省略在经编码视频位流中产生指示变换跳过模式的数据,并在经编码视频位流中产生指示二次变换的数据。

[0304] 根据本发明的第六技术,对于特定译码模式,二次变换可被停用及/或并不被传信,且反之亦然。举例来说,这些译码模式可包含(但未必限于)变换跳过模式、线性方法(LM)模式、交叉分量残差预测模式等。因此,如果应用二次变换,则特定译码模式可被停用及/或并不被传信。举例来说,在一些情况下,当应用二次变换时,变换跳过模式、LM模式及交叉分量残差预测模式可被停用及/或变换跳过模式、LM模式及交叉分量残差预测模式的索引不被编码。类似地,二次变换可在使用变换跳过模式、LM模式及交叉分量残差预测模式中的一或多者时被停用及/或不被传信。一般来说,自适应性交叉分量残差预测可包含从同一块的另一一色彩分量的残差预测一个色彩分量(例如,Y、Cr或Cb)的残差。

[0305] 在第六技术的实例中,当启用二次变换时,特定模式可始终被停用。举例来说,基

于二次变换被启用,变换跳过模式被停用。

[0306] 在第六技术的实例中,当二次变换经启用时,特定模式可在一些条件下停用但在其它条件下启用。条件可包含(但不限于)块大小、非零变换系数的数目、译码针对明度抑或色度分量、相邻预测模式及其它条件,且反之亦然。

[0307] 在第六技术的实例中,特定预测模式可在二次变换经应用时仍被启用,但用于传信此特定预测模式的上下文模型化可经修改以使得上下文模型化取决于所传信二次索引/旗标。换句话说,在二次变换被应用时,可准许应用特定预测模式,但用于特定预测模式的上下文模型化被修改。

[0308] 在实例中,二次变换仍经启用用于特定预测模式,但用于传信二次变换的上下文模型化可经修改以使得上下文模型化取决于特定预测模式是否经应用。换句话说,在应用二次变换时可准许使用特定预测模式,而非停用不通过视频编码器20及/或视频解码器30应用的预测模式,但用于特定预测模式的上下文模型化经修改。

[0309] 根据本发明的第七实例,当将二次变换应用于其它非明度分量(例如,色度)上时及当二次变换选择对于某一译码信息有某一依赖性时,NSST索引值可从同置的明度分量继承(例如,再使用)且不经编码用于非明度分量。此译码信息的实例可包含(但不限于)帧内预测模式。如本文所使用,同置可指相同块的分量。在此实例的方面中,以使用非明度分量的译码信息而不是继承同置明度分量的译码信息的方式实现对于某一译码信息的依赖性。

[0310] 在实例中,当二次变换经执行用于色度分量(例如,Cr分量及/或Cb分量),且视频解码器30基于所传信NSST索引及帧内预测模式选择二次变换时,则针对色度分量,视频解码器30可基于同置明度分量的NSST索引及当前色度分量的帧内预测模式选择二次变换。

[0311] 当色度分量的帧内预测模式从同置明度分量(例如,Y分量)继承时,则同置明度分量的帧内预测模式可用于选择二次变换。举例来说,视频解码器30可基于同置明度分量的NSST索引及同置明度分量的帧内预测模式从子集中选择二次变换候选者。

[0312] 在实例中,当色度分量的帧内预测模式为LM模式(例如,如第15/006,994号美国专利申请案中所描述)时,二次变换可不应用于色度分量。举例来说,视频解码器30可将二次变换候选者应用于同置明度分量,且避免将二次变换候选者应用于色度分量。

[0313] 在一个实例中,当二次变换经应用时,LM模式可不被应用或传信。在此情况下,可因此修改帧内模式译码以使得LM模式并非为待传信的候选帧内模式。此帧内模式译码修改可包含(但不限于):减少待译码的可用帧内模式的数目,或以其它帧内预测模式替代LM模式使得可用帧内模式的总数相同。换句话说,当应用二次变换时,帧内模式的传信可经修改以不再传信LM模式,此是由于当应用二次变换时LM模式可经假定为并非候选帧内模式。

[0314] 另外或替代地,当色度分量的帧内预测模式为LM模式时,可使用继承的NSST索引及色度分量的平坦/DC模式选择二次变换。举例来说,视频解码器30可基于同置明度分量的NSST索引及色度分量的平坦/DC模式从子集中选择二次变换候选者。

[0315] 在实例中,同置明度分量的NSST索引及帧内预测模式两者可经继承用于当前非明度分量。举例来说,视频解码器30可基于同置非明度分量的NSST索引及同置非明度分量的帧内预测模式从子集中选择二次变换候选者。

[0316] 图7A为从明度帧内模式到变换集合索引的实例映射的说明。可通过视频编码器20及/或视频解码器30执行下文所描述的实例。如图7A中所说明,下文所描述的实例可实施于

67个明度帧内预测模式上,包含平坦(0)、DC(1)及65个有角度的明度帧内预测模式(2到66)。为选择变换矩阵以用于执行非可分离变换,可预定义总共12个非可分离 16×16 变换集合。每一变换集合含有三个 16×16 变换候选者。在实例中,变换集合可通过明度帧内预测模式指定。举例来说,如图7A中所示,明度帧内预测模式‘0’指定变换集合索引‘0’,明度帧内预测模式‘1’指定变换集合索引‘0’,明度帧内预测模式‘2’指定变换集合索引‘1’,等等。

[0317] 在图7A的实例中,变换集合可经指定,且在每一变换集合内,来自3个候选者的所选择变换可通过经明确传信的CU层级ROT索引而指定。举例来说,ROT索引可为位流中的每一CU传信的索引值,且ROT索引值可在0到3范围内。在一些情况下,ROT索引0指定不应用二次变换,且ROT索引1到3指定来自3个不同ROT的一个所选择变换。在图7A中定义从明度预测模式到变换集合索引的映射。

[0318] 另外,根据图7A,对于大于34的明度帧内预测模式(IPM),可应用帧内模式68-IPM的相同变换集合索引。然而,为了利用帧内预测模式IPM与68-IPM之间的对称性,在编码器/解码器处,可在进行二次变换之前/之后转置变换系数块。更特定来说,在图7B的实例中,帧内预测模式(IPM)在0到66范围内。如图7B中所说明,帧内模式IPM及68-IPM为对称的。举例来说,帧内模式18(水平预测)及50(垂直预测)为对称的。由于IPM及68-IPM为对称的,应用于这些两个模式上的非可分离变换具有某一联系。举例来说,如果我们转置从模式50(垂直预测)预测的残差块,则残差统计数据应非常类似于从模式18(水平预测)预测的残差块。因此,用于帧内模式IPM的非可分离变换应类似于用于帧内模式IPM的非可分离变换加转置操作。为了利用此对称性及减少非可分离变换集合的总数,对于大于34的帧内模式IPM,可使用68-IPM的相同变换集合且可在二次变换之后转置变换系数块。在JVET-C1001的章节2.2.1:http://phenix.it-sudparis.eu/jvet/doc_end_user/documents/3_Geneva/wg11/JVET-C1001-v3.zip中,描述66明度帧内预测模式。

[0319] 在本发明的第八技术中,视频编码器20及/或视频解码器30可经配置以将NSST的不同大小及不同形状(除正方形或非正方形以外)应用于不同块大小。在实例中, 4×4 NSST可应用于 4×4 块,且 8×8 NSST可应用于 8×8 及更大块的左上 8×8 变换系数上。举例来说,视频编码器20可将 4×4 二次变换应用于 4×4 块,可将 8×8 二次变换应用于 8×8 块,且可将 8×8 二次变换应用于大于 8×8 的块的左上 8×8 系数。举例来说,视频编码器20可将 8×8 二次变换应用于 16×16 块的左上 8×8 系数。

[0320] 在第八技术的实例中,当 4×4 NSST应用于 $M \times N$ 个块(其中 $M \geq 4$ 、 $N \geq 4$ 、 M 及 N 为4的倍数、 $M \neq N$)时, $M \times N$ 个块可由多个非重叠 4×4 子块分隔。对于每一 4×4 子块,可应用 4×4 NSST。举例来说,视频编码器20可将第一 4×4 二次变换应用于 8×4 块的左上 4×4 系数,并将第二 4×4 二次变换应用于 8×4 块的在左上 4×4 系数下方的 4×4 系数,以使得视频编码器20将NSST应用于 8×4 块的两个 4×4 块中的每一者。

[0321] 举例来说,视频编码器20可将 $M \times N$ 个系数块分成多个非重叠 4×4 子块。在此实例中,多个非重叠 4×4 子块包含第一2维系数块及第二2维系数块, M 及 N 各自大于4, M 及 N 各自为4的倍数,且 M 并不等于 N 。在此实例中,对于多个非重叠 4×4 子块的每一各别 4×4 子块,视频编码器20执行二次变换以产生各别系数块。

[0322] 类似地,视频解码器30可将 $M \times N$ 个系数块分成多个非重叠 4×4 子块,多个非重叠 4×4 子块包含第一2维系数块, M 及 N 各自大于4, M 及 N 各自为4的倍数,且 M 不等于 N 。在此实例

中,视频解码器30可对于多个非重叠 4×4 子块的每一 4×4 子块执行 4×4 逆NSST。在此实例中,对于多个 4×4 子块的每一各别 4×4 子块,视频解码器30应用逆二次变换以产生各别2维系数块。

[0323] 在第八技术的实例中,当 8×8 NSST应用于 $M\times N$ 个块(其中 $M\geq 8$, $N\geq 8$, M 及 N 为8的倍数)时, $M\times N$ 个块可分成多个非重叠 8×8 子块,且对于每一 8×8 子块,可应用 8×8 NSST。举例来说,视频编码器20可将第一 8×8 二次变换应用于 8×64 块的左上 8×8 系数,并将第二 8×8 二次变换应用于 8×64 块的在左上 8×8 系数下方的 8×8 系数,等等,以使得视频编码器20将NSST应用于 8×64 块的八个 8×8 块中的每一者。

[0324] 在第八技术的实例中,替代正方形NSST,NSST可应用于位于系数块的任意区域中的变换系数的群组。举例来说,视频编码器20可将第一 4×16 二次变换应用于 8×64 块的中心 4×16 系数。在实例中,呈系数扫描次序的最初 K 个变换系数可经组织成单一向量,且可对此变换系数向量执行非可分离变换。在此实例中,系数扫描次序可为(但不限于)对角线扫描、水平扫描或垂直扫描。

[0325] 在第八技术的实例中,对于 $N\times N$ 块,可应用归零NSST,以使得仅仅计算前 M 个系数,且剩余 $N-M$ 个系数可被认为是零。举例来说,对于包含 N 数目个系数的1维系数向量,视频编码器20可应用仅计算1维系数向量的 N 个系数中的前 M 个系数并将1维系数向量的剩余(即, $N-M$)系数设定为零的NSST。

[0326] 如果一群置零系数形成系数群(CG)(例如,如HEVC中所定义的 4×4 子块),则当熵译码此CG时,旗标(其传信在此CG中是否存在至少一个非零系数)可经跳过且不经译码。在一些实例中,替代不译码置零系数的语法元素,置零系数的语法元素可仍被译码但使用单独上下文。

[0327] 此外,在一些实例中,如果一群置零系数形成系数群(CG)(例如,如HEVC中所定义的 4×4 子块),则当熵译码此CG时,最后位置译码可经修改以包含最后位置不可呈现在这些置零系数置中的约束。

[0328] 图8为说明可实施本发明中所描述的技术的视频编码器20的实例的框图。根据本发明,视频编码器20可经配置以应用在本发明中描述的技术,例如与二次变换有关的技术。视频编码器20可执行视频切片内的视频块的帧内译码及帧间译码。帧内译码依赖于空间预测以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以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)可指若干基于空间的译码模式中的任一者。帧间模式(例如单向预测(P模式)或双向预测(B模式))可指若干基于时间的译码模式中的任一者。

[0329] 如图8中所展示,视频编码器20接收待编码的视频帧内的当前视频块。在图8的实例中,视频编码器20包含模式选择单元240、参考图片存储器264(其也可被称作经解码图片缓冲器(DPB))、求和器250、变换处理单元252、量化单元254及熵编码单元256。模式选择单元240又包含运动补偿单元244、运动估计单元242、帧内预测单元246及分割单元248。为了视频块重建,视频编码器20还包含逆量化单元258、逆变换单元260及求和器262。在一些实例中,还包含解块滤波器(图8中未展示)以对块边界滤波以从重建的视频移除方块效应伪影。解块滤波器将求和器262的输出滤波。除了解块滤波器外,也可使用额外滤波器(回路中或回路后)。为简洁起见未展示此类滤波器,但如果需要,则可对求和器250的输出进行滤波

(作为回路中滤波器)。

[0330] 在编码过程期间,视频编码器20接收待译码的视频帧或切片。可将所述帧或切片分成多个视频块。运动估计单元242及运动补偿单元244执行所接收视频块相对于一或多个参考帧中的一或多个块的帧间预测性编码以提供时间预测。替代地,帧内预测单元246可执行所接收视频块相对于与待译码块相同的帧或切片中的一或多个相邻块的帧内预测性编码以提供空间预测。视频编码器20可执行多个译码遍次,(例如)以选择用于每一视频数据块的适当译码模式。

[0331] 此外,分割单元248可基于对先前译码遍次中的先前分割方案的评估而将视频数据的块分割成子块。举例来说,分割单元248可最初将帧或切片分割成LCU,且基于速率失真分析(例如,速率失真最佳化)来将所述LCU中的每一者分割成子CU。模式选择单元240可进一步产生指示将LCU分割为子CU的四分树数据结构。四分树的叶节点CU可包含一或多个PU及一或多个TU。

[0332] 模式选择单元240可例如基于误差结果选择预测模式(帧内或帧间)中的一者,且提供所得预测块到求和器250以产生残差数据并提供到求和器262以重建经编码块以用作参考帧。模式选择单元240还将语法元素(例如运动向量、帧内模式指示符、分割信息及其它这些语法信息)提供到熵编码单元256。

[0333] 运动估计单元242及运动补偿单元244可高度集成,但为概念目的而单独说明。由运动估计单元242执行的运动估计为产生运动向量的过程,运动向量估计视频块的运动。举例来说,运动向量可指示在当前视频帧或图片内的视频块的PU相对于在参考帧(或其它经译码单元)内的预测性块(其相对于在所述当前帧(或其它经译码单元)内正经译码的当前块)的位移。预测性块为就像素差来说被发现紧密地匹配待译码块的块,所述像素差可通过绝对差总和(SAD)、平方差总和(SSD)或其它差量度确定。在一些实例中,视频编码器20可计算存储于参考图片存储器264中的参考图片的次整数像素位置的值。举例来说,视频编码器20可内插参考图片的四分之一像素位置、八分之一像素位置或其它分数像素位置的值。因此,运动估计单元242可执行关于全像素位置及分数像素位置的运动搜寻且输出具有分数像素精度的运动向量。

[0334] 运动估计单元242通过比较PU的位置与参考图片的预测性块的位置而计算帧间译码切片中的视频块的PU的运动向量。参考图片可是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于参考图片存储器264中的一或多个参考图片。运动估计单元242将所计算运动向量发送到熵编码单元256及运动补偿单元244。

[0335] 由运动补偿单元244执行的运动补偿可涉及基于由运动估计单元242确定的运动向量来提取或产生预测性块。再次,在一些实例中,运动估计单元242与运动补偿单元244可在功能上集成。在接收当前视频块的PU的运动向量之后,运动补偿单元244即可在参考图片列表中的一者中定位运动向量所指向的预测性块。求和器250形成残差视频块。作为形成残差视频块的部分,求和器250可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,如下文所论述。一般来说,运动估计单元242相对于明度分量执行运动估计,且运动补偿单元244将基于明度分量计算的运动向量用于色度分量与明度分量两者。模式选择单元240也可产生与视频块及视频切片相关联的语法元素以供视频解码器30在解码视频切片的视频块时使用。

[0336] 如上文所描述,作为由运动估计单元242及运动补偿单元244执行的帧间预测的替代方案,帧内预测单元246可对当前块进行帧内预测。详细来说,帧内预测单元246可确定待用以编码当前块的帧内预测模式。在一些实例中,帧内预测单元246可(例如)在单独编码遍次期间使用各种帧内预测模式编码当前块,且帧内预测单元246(或在一些实例中,模式选择单元240)可从经测试模式中选择适当帧内预测模式来使用。

[0337] 举例来说,帧内预测单元246可针对各种经测试帧内预测模式使用速率失真分析计算速率失真值,且从所述经测试模式当中选择具有最佳速率失真特性的帧内预测模式。速率失真分析大体上确定经编码块与原始、未经编码块(其经编码以产生经编码块)之间的失真(或误差)量,以及用以产生经编码块的位率(即,位的数目)。帧内预测单元246可根据各种经编码块的失真及速率来计算比率以确定哪一帧内预测模式展现所述块的最佳速率失真值。

[0338] 在选择用于块的帧内预测模式后,帧内预测单元246可将指示用于块的所选帧内预测模式的信息提供到熵编码单元256。熵编码单元256可编码指示所选帧内预测模式的信息。视频编码器20可在所发射的位流中包含配置数据,其可包含多个帧内预测模式索引表及多个经修改帧内预测模式索引表(也称作码字映射表);各种块的编码上下文的定义;及待用于所述上下文中的每一者的最有可能的帧内预测模式、帧内预测模式索引表及经修改帧内预测模式索引表的指示。

[0339] 视频编码器20形成残差视频块。作为形成残差视频块的部分,视频编码器20可从正被译码的原始视频块减去来自模式选择单元240的预测数据。求和器250表示执行此减法运算的所述或所述组件。变换处理单元252将变换(例如离散余弦变换(DCT)或概念上类似的变换)应用于残差块,从而产生包括变换系数值的视频块。可使用小波变换、整数变换、子带变换、DST或其它类型的变换而不是DCT。在任何情况下,变换处理单元252将变换应用于残差块,从而产生变换系数的块。变换可将残差信息从像素域转换到变换域,例如频域。

[0340] 另外,如上文所论述,变换处理单元252可经配置以应用非可分离的二次变换。在一些实例中,变换处理单元252可经配置以根据系数扫描次序将通过非可分离的二次变换形成的1维系数向量重组为第二2维系数块以供量化单元254使用。变换处理单元252可将所得变换系数发送到量化单元254。

[0341] 量化单元254量化变换系数以进一步减小位率。量化过程可减小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。

[0342] 在量化之后,熵编码单元256扫描并熵编码指示经量化变换系数的语法元素。举例来说,熵编码单元256可执行CAVLC、CABAC、SBAC、PIPE译码或另一熵译码技术。在基于上下文的熵译码的情况下,上下文可基于相邻块。在由熵编码单元256进行熵译码之后,可将经编码位流发射到另一装置(例如,视频解码器30),或加以存档以供稍后发射或检索。

[0343] 逆量化单元258及逆变换单元260分别应用逆量化及逆变换以重建像素域中的残差块。详细来说,求和器262将经重建残差块添加到由运动补偿单元244或帧内预测单元246较早产生的运动补偿预测块,以产生用于存储于参考图片存储器264中的经重建的视频块。所述经重建的视频块可由运动估计单元242及运动补偿单元244使用作为参考块以对随后视频帧中的块进行帧间译码。

[0344] 视频编码器20一般使用上文所论述过程来对经译码视频序列中的每一图片的每

一块进行编码。另外,在一些实例中,视频编码器20可确定向其指派图片中的每一者的时间层。此外,视频编码器20可经配置以对其它层(例如,其它视图、可调式视频译码层等)的图片进行编码。在任何情况下,视频编码器20可进一步针对(例如,各种视频维度的)一或多个层编码指示每一图片属于的层的数据。

[0345] 图9为说明可实施本发明中所描述的技术的视频解码器30的实例的框图。根据本发明,视频解码器30可经配置以应用在本发明中描述的技术,例如与二次变换有关的技术。在图9的实例中,视频解码器30包含熵解码单元270、运动补偿单元272、帧内预测单元274、逆量化单元276、逆变换单元278、参考图片存储器282及求和器280。在一些实例中,视频解码器30可执行大体上与关于视频编码器20(图8)所描述的编码遍次互逆的解码遍次。运动补偿单元272可基于从熵解码单元270接收的运动向量产生预测数据,而帧内预测单元274可基于从熵解码单元270接收的帧内预测模式指示符产生预测数据。

[0346] 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。视频解码器30的熵解码单元270熵解码位流以产生经量化系数、运动向量或帧内预测模式指示符及其它语法元素。熵解码单元270将运动向量及其它语法元素转递到运动补偿单元272。视频解码器30可在视频切片层级及/或视频块层级接收语法元素。

[0347] 当视频切片经译码为经帧内译码(I)切片时,帧内预测单元274可基于所传信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而产生当前视频切片的视频块的预测数据。在视频帧经译码为经帧间译码(即,B或P或一般化P及B图片(GPB))切片时,运动补偿单元272基于从熵解码单元270所接收的运动向量及其它语法元素产生当前视频切片的视频块的预测性块。可从参考图片列表中的一者内的参考图片中的一者产生预测性块。视频解码器30可基于存储于参考图片存储器282中的参考图片使用预设建构技术建构参考图片列表(列表0及列表1)。运动补偿单元272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,且使用预测信息产生用于正被解码的当前视频块的预测性块。举例来说,运动补偿单元272使用所接收语法元素中的一些以确定用以译码视频切片的视频块的预测模式(例如,帧内或帧间预测)、帧间预测切片类型(例如,B切片、P切片或GPB切片)、用于所述切片的参考图片列表中的一或多者的建构信息、用于所述切片的每一经帧间编码视频块的运动向量、用于所述切片的每一经帧间译码视频块的帧间预测状态、及用以解码当前视频切片中的视频块的其它信息。

[0348] 运动补偿单元272也可执行基于内插滤波器的内插。运动补偿单元272可使用如由视频编码器20在视频块的编码期间使用的内插滤波器,以计算参考块的次整数像素的内插值。在此情况下,运动补偿单元272可根据接收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。

[0349] 逆量化单元276逆量化(即,解量化)位流中所提供并由熵解码单元270解码的经量化的变换系数。逆量化过程可包含使用由视频解码器30针对视频切片中的每一视频块计算的量化参数 QP_V 以确定应应用的量化程度及(同样地)逆量化程度。

[0350] 逆变换单元278将逆变换(例如,逆DCT、逆整数变换或概念地类似的逆变换过程)应用于变换系数,以便在像素域中产生残差块。另外,根据本发明的各种技术,逆变换单元278可经配置以应用非可分离的逆二次变换。在一些实例中,逆变换单元278可经配置以根

据系数扫描次序将1维系数向量重组为第二2维系数块。

[0351] 在运动补偿单元272基于运动向量及其它语法元素产生当前视频块的预测性块之后,视频解码器30形成经解码视频块。作为形成经解码视频块的部分,视频解码器30可将来自逆变换单元278的残差块与由运动补偿单元272产生的对应预测性块求和。求和器280表示执行此求和运算的所述或所述组件。如果需要,也可应用解块滤波器来对经解码块滤波以便移除块效应伪影。也可使用其它回路滤波器(在译码回路中或在译码回路之后)以使像素移转平滑,或以其它方式改进视频品质。接着,将给定帧或图片中的经解码视频块存储于存储用于随后运动补偿的参考图片的参考图片存储器282中。参考图片存储器282还存储经解码视频用于稍后在显示装置(例如图1的显示装置32)上呈现。

[0352] 视频解码器30一般使用上文所论述的过程来对经译码视频序列中的每个图片的每一块进行解码。另外,在一些实例中,视频解码器30可对指示向其指派图片的时间层的数据进行解码。此外,视频解码器30可经配置以对其它层(例如,其它视图、可调式视频译码层等)的图片进行解码。在任何情况下,视频解码器30可进一步针对(例如,各种视频维度的)一或多个层解码指示每一图片属于的层的数据。

[0353] 在下文更详细地描述可通过视频编码器20或视频解码器30执行的本发明的技术。

[0354] 图10A为根据本发明中所描述的一或多个技术的通过视频编码器20进行的实例变换处理的说明。在图10A的实例中,视频编码器20的变换处理单元252包括一次变换单元802、二次变换单元804及重组单元806。一次变换单元802将一次变换(例如DST或DCT)应用于残差块上以产生2维系数块。二次变换单元804将非可分离变换(例如KLT)应用于2维系数块以产生1维向量。重组单元806将1维向量重组为2维系数块。量化单元254量化2维块。

[0355] 图10B为根据本发明中所描述的一或多个技术的通过视频解码器30进行的实例变换处理的说明。在图10B的实例中,视频解码器30的逆变换单元278包括重组单元824、二次逆变换单元826及一次逆变换单元828。逆量化单元276逆量化提供于位流中且由熵解码单元270解码的经量化变换系数。重组单元824将变换系数从2维系数块重组到1维向量。二次逆变换单元826将非可分离逆变换(例如KLT)应用于1维向量上以产生2维系数块。一次逆变换单元828将可分离逆变换应用于2维系数块上以产生用于视频数据的残差系数块。

[0356] 图11为说明可实施本发明中所描述的技术的视频数据的第一实例解码的流程图。可通过视频解码器30执行图11的实例技术。在图11的实例中,解码器(例如,视频解码器30)逆量化第一系数块(902)。解码器将第一逆变换应用于经逆量化第一系数块的至少部分以产生第二系数块(904)。举例来说,第一逆变换为非可分离变换。举例来说,解码器可将逆KLT应用于经逆量化第一系数块上。解码器可将第二逆变换应用于第二系数块以产生残差视频块(906)。所述第二逆变换将所述第二系数块从频域转换到像素域。举例来说,解码器将逆DCT或DST应用于第二系数块上。解码器形成经解码视频块(908)。作为形成经解码视频块的部分,解码器可将残差视频块与一或多个预测性块求和。

[0357] 图12为说明可实施本发明中所描述的技术的视频数据的第一实例编码的流程图。如所描述,图12的实例技术可通过编码器20执行。在图12的实例中,编码器(例如,视频编码器20)形成残差视频块(1002)。作为形成残差视频块的部分,编码器可从正被译码的视频块减去一或多个预测性块。编码器将第一变换应用于残差视频块以产生第一系数块(1004)。举例来说,第一变换将残差视频块从像素域转换到频域。举例来说,编码器可将DCT或DST应

用于残差视频块上。编码器将第二变换应用于第一系数块的至少部分以产生第二系数块(1006)。举例来说,第二变换为非可分离变换。举例来说,编码器可将KLT应用于第二系数块上。接下来,编码器量化第二系数块以用于熵编码(1008)。

[0358] 图13为说明可实施本发明中所描述的技术的视频数据的第二实例解码的流程图。如所描述,图13的实例技术可通过视频解码器30执行。在图13的实例中,解码器(例如,视频解码器30)根据系数扫描次序将第一2维系数块重组为第一1维系数向量(1102)。举例来说,解码器根据系数扫描次序重组第一2维系数块,其中具有较小索引值的第一2维系数块中的系数放置于第一1维系数向量中的较小扫描索引位置处。接下来,解码器将第一逆变换应用于第一1维系数向量以产生第二1维系数块(1104)。第一逆变换为非可分离变换。举例来说,解码器将逆KLT应用于1维系数向量上。决定器根据系数扫描次序将第二1维系数块重组为第二2维系数向量(1106)。接下来,解码器将第二逆变换应用于第二2维系数块以产生残差视频块(1108)。举例来说,第二逆变换将第二2维系数块从频域转换到像素域。举例来说,解码器将逆DCT或DST应用于第二2维系数块上。解码器形成经解码视频块(1110)。作为形成经解码块的部分,解码器可将残差视频块与一或多个预测性块求和。

[0359] 图14为说明可实施本发明中所描述的技术的视频数据的第二实例编码的流程图。如所描述,图14的实例技术可通过视频编码器20执行。在图14的实例中,编码器(例如,视频编码器20)形成残差视频块(1202)。作为形成残差视频块的部分,编码器可从正被译码的视频块减去一或多个预测性块。编码器将第一变换应用于残差视频块的至少部分以产生第一2维系数块(1204)。举例来说,第一变换将残差视频块从像素域转换到频域。举例来说,编码器将DCT或DST应用于残差视频块上。编码器将第一2维系数重组为第一1维系数向量(1206)。接下来,编码器将第二变换应用于第一1维系数块的至少部分以产生第二1维系数向量(1208)。举例来说,第二变换为非可分离变换。举例来说,编码器将KLT应用于第一1维系数块上。接下来,编码器根据系数扫描次序将第二1维系数块重组为第二2维系数块(1210)。举例来说,解码器根据系数扫描次序重组1维系数向量,其中具有较小索引值的第二1维系数向量中的系数放置在第二2维系数块中的较小扫描索引位置处以用于量化及熵编码。

[0360] 应认识到,取决于实例,本文中所描述的技术中的任一者的某些动作或事件可以不同序列被执行、可被添加、合并或完全省去(例如,并非所有所描述动作或事件为实践所述技术所必要)。此外,在某些实例中,可例如经由多线程处理、中断处理或多个处理器同时而非依序执移动作或事件。

[0361] 在一或多个实例中,所描述的功能可以硬件、软件、固件或其任何组合来实施。如果以软件实施,则所述功能可作为一或多个指令或程序代码而存储于计算机可读媒体上或经由计算机可读媒体进行发射,且通过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,其对应于有形媒体(例如数据存储媒体),或包含有助于将计算机程式从一处传送到另一处(例如,根据通信协议)的任何媒体的通信媒体。以此方式,计算机可读媒体大体可对应于(1)为非暂时性的有形计算机可读存储媒体,或(2)例如信号或载波的通信媒体。数据存储媒体可为可由一或多个计算机或一或多个处理器存取以检索指令、程序代码及/或数据结构以用于实施本发明中所描述的技术的任何可用媒体。计算机程式产品可包含计算机可读媒体。

[0362] 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,或可用以存储呈指令或数据结构形式的所要程序代码且可由计算机存取的任何其它媒体。此外,任何连接适当地称为计算机可读媒体。举例来说,如果使用同轴电缆、光纤缆线、双绞线、数字订户线(DSL)或例如红外线、无线电及微波的无线技术从网站、服务器或其它远程源发射指令,则同轴电缆、光纤缆线、双绞线、DSL或例如红外线、无线电及微波的无线技术包含于媒体的定义中。然而,应理解,计算机可读存储媒体及数据存储媒体不包含连接、载波、信号或其它暂时性媒体,而取而代之,是有关非暂时性的有形存储媒体。如本文所使用,磁盘及光盘包含紧密光盘(CD)、激光光盘、光学光盘、数字多功能光盘(DVD)、软盘及蓝光光盘,其中磁盘通常以磁性方式再生数据,而光盘用激光以光学方式再生数据。以上各者的组合也应包含于计算机可读媒体的范围内。

[0363] 指令可由一或多个处理器执行,所述一或多个处理器例如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)、现场可编程门阵列(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”可指上述结构或适用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,在一些方面中,本文中所描述的功能性可提供于经配置用于编码及解码的专用硬件及/或软件模块内,或并入于组合式编码解码器中。另外,技术可以全部实施于一或多个电路或逻辑元件中。

[0364] 本发明的技术可在多种装置或设备中实施,所述装置或设备包含无线手机、集成电路(IC)或IC集合(例如,芯片组)。在本发明中描述各种组件、模块或单元以强调经配置以执行所揭示技术的装置的功能方面,但未必需要通过不同硬件单元来实现。确切来说,如上文所描述,各种单元可组合于编码解码器硬件单元中或由结合合适的软件及/或固件的包含如上文所描述的一或多个处理器的互操作性硬件单元的集合提供。已描述各种实例。这些及其它实例是在以下权利要求书的范围内。

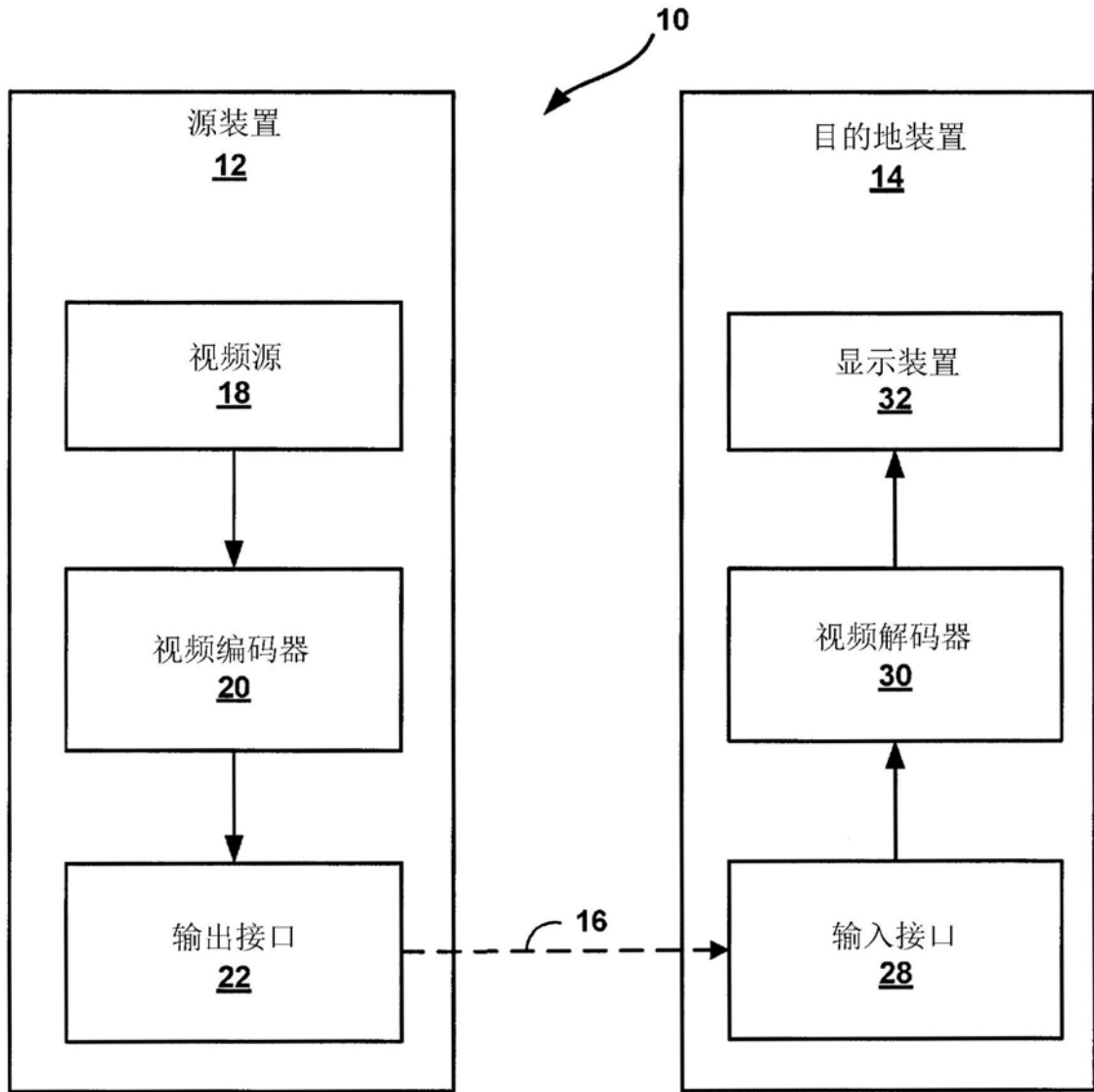


图1

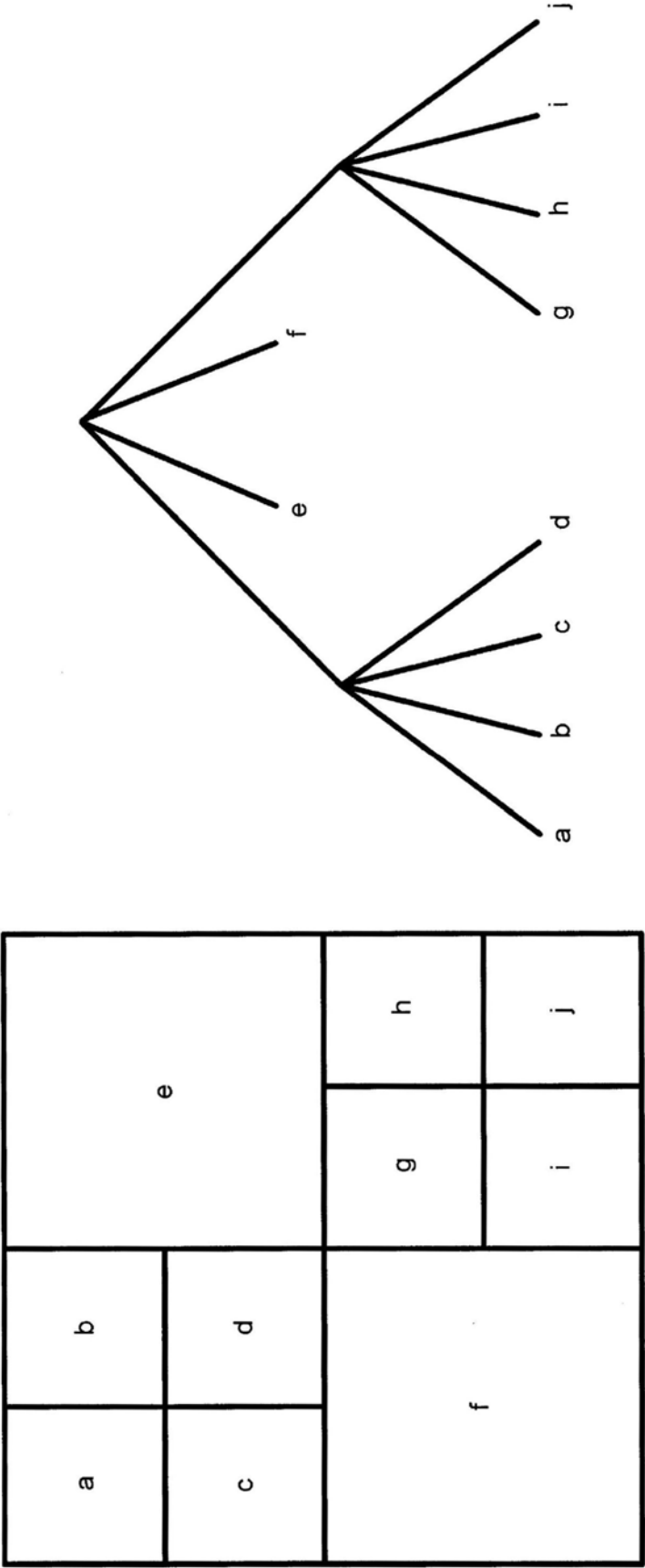


图2

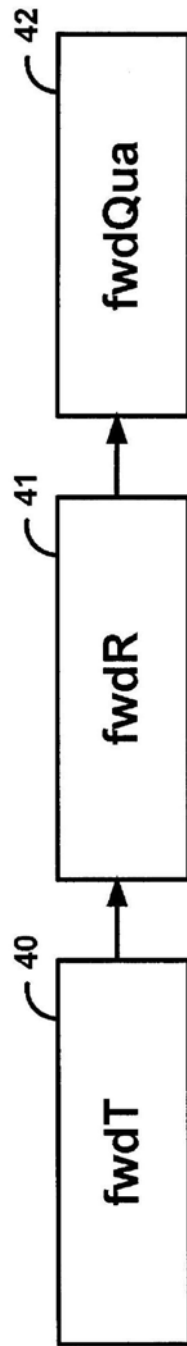


图3A



图3B

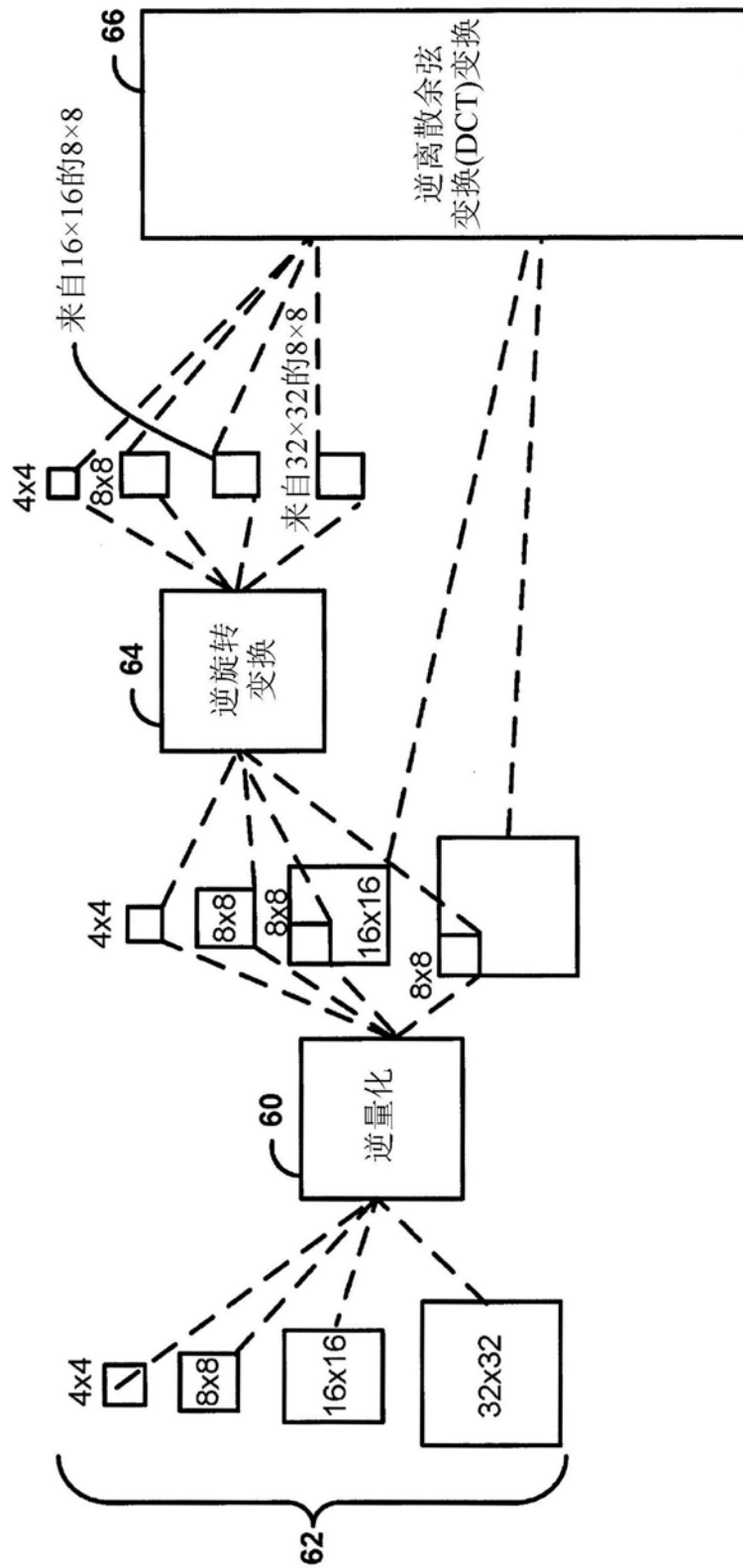


图4

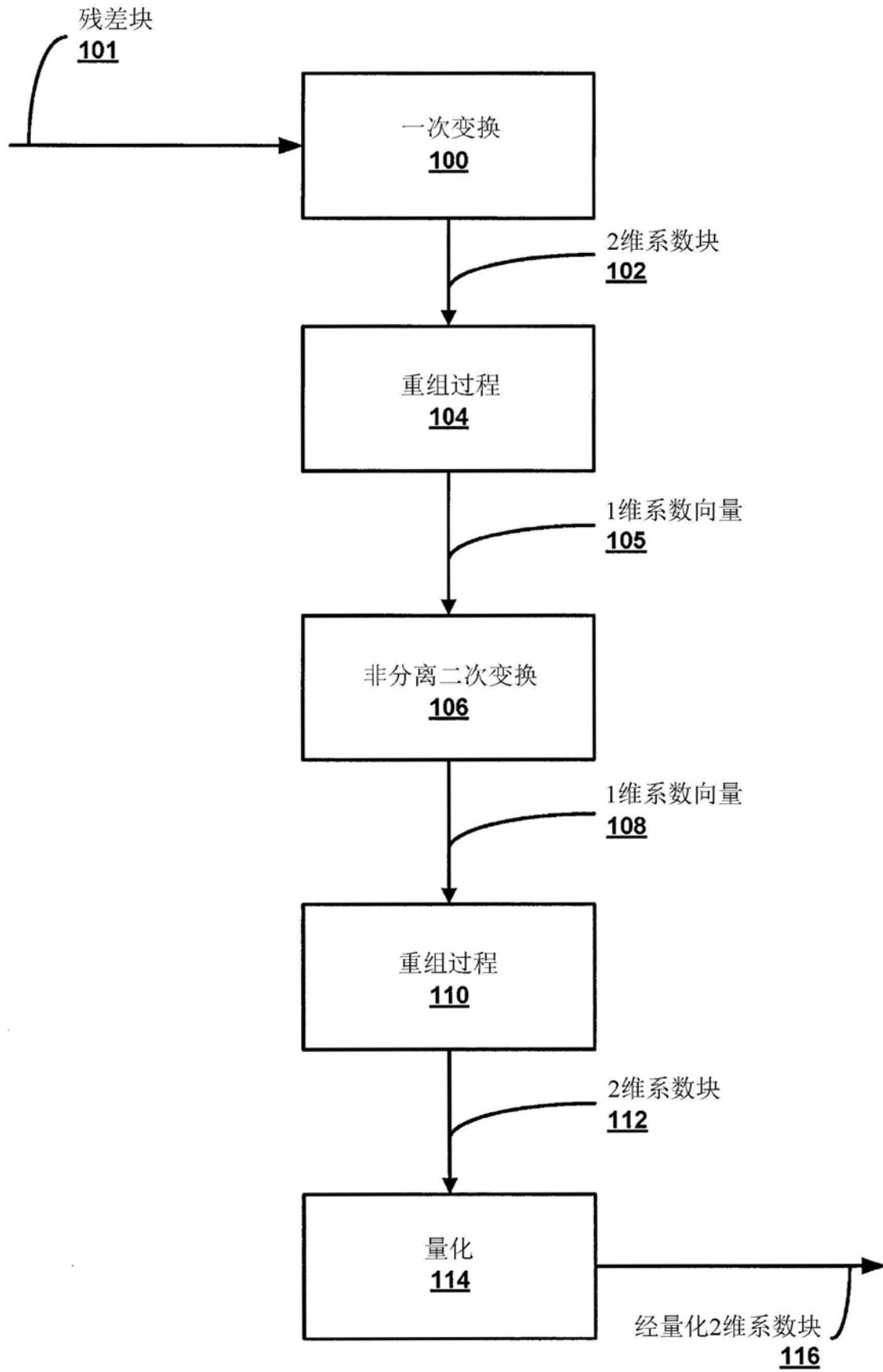


图5

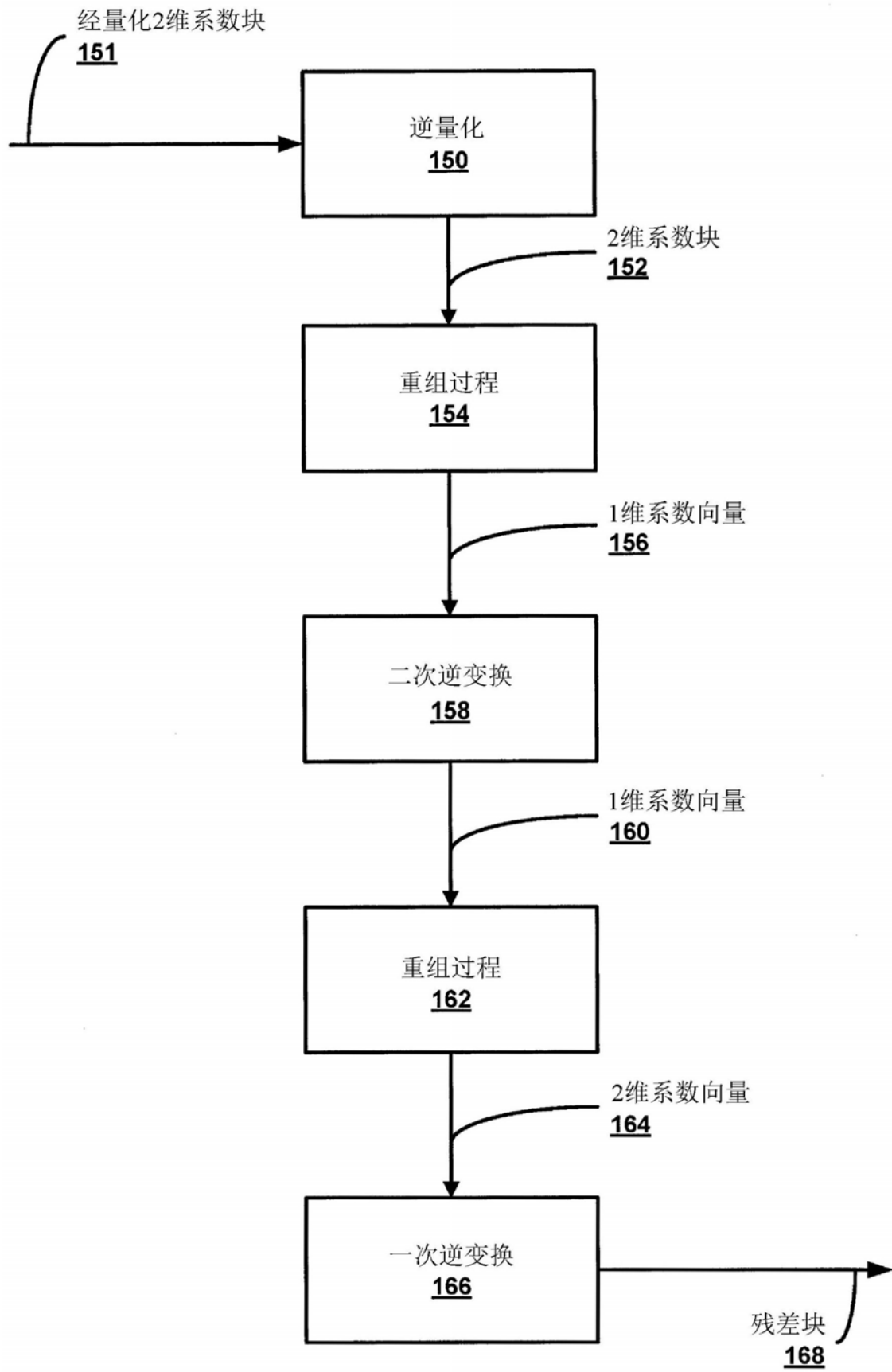


图6

明度帧内模式	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33
变换集合索引	0	0	1	2	1	2	1	2	3	4	3	4	3	4	5	5	5	6	6	6	7	7	7	8	9	8	9	8	9	10	11	10	11	10
明度帧内模式	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	
变换集合索引	11	10	11	10	11	10	9	8	9	8	9	8	7	7	7	6	6	6	5	5	5	4	3	4	3	4	3	2	1	2	1	2	1	

图7A

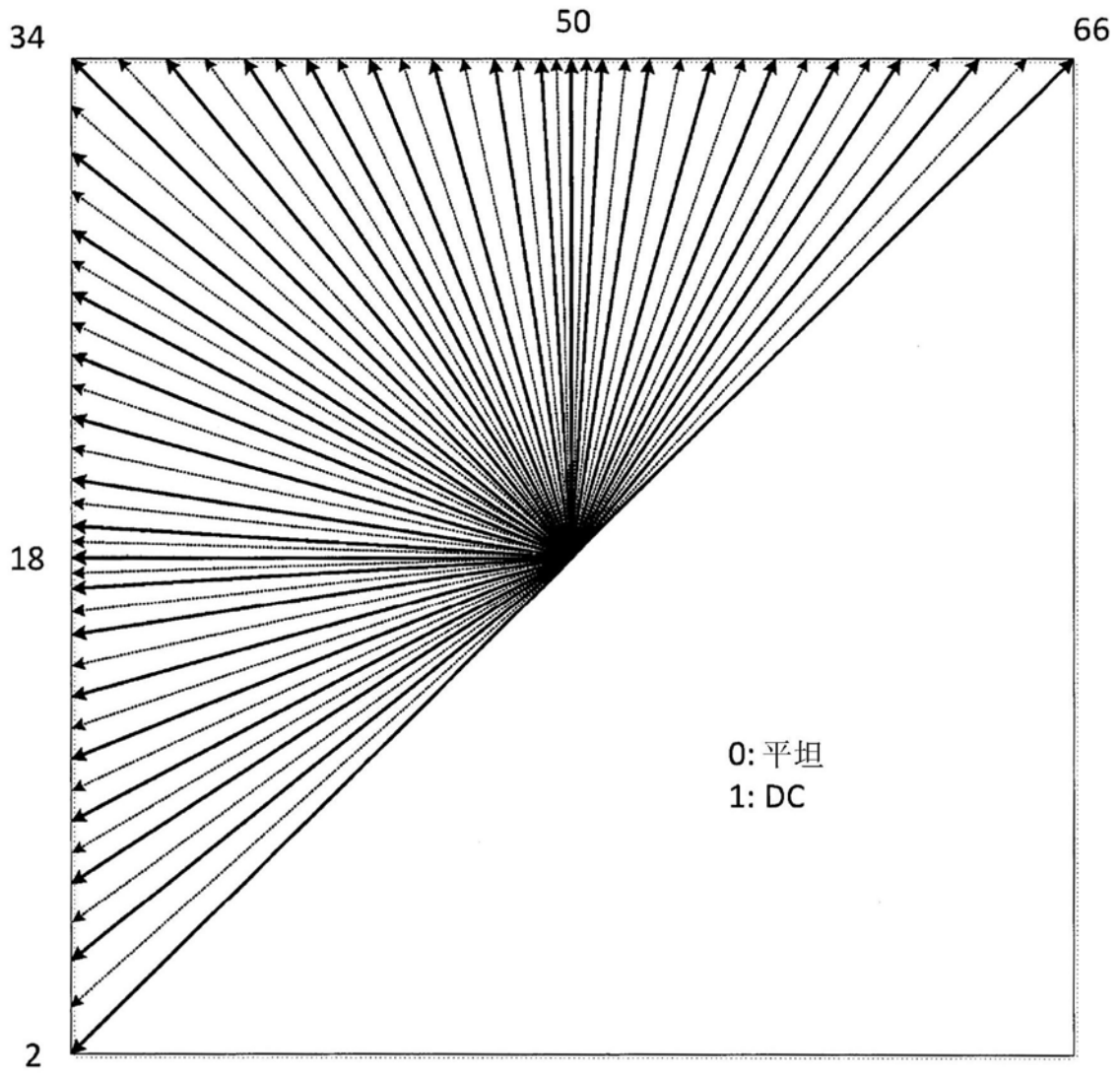


图7B

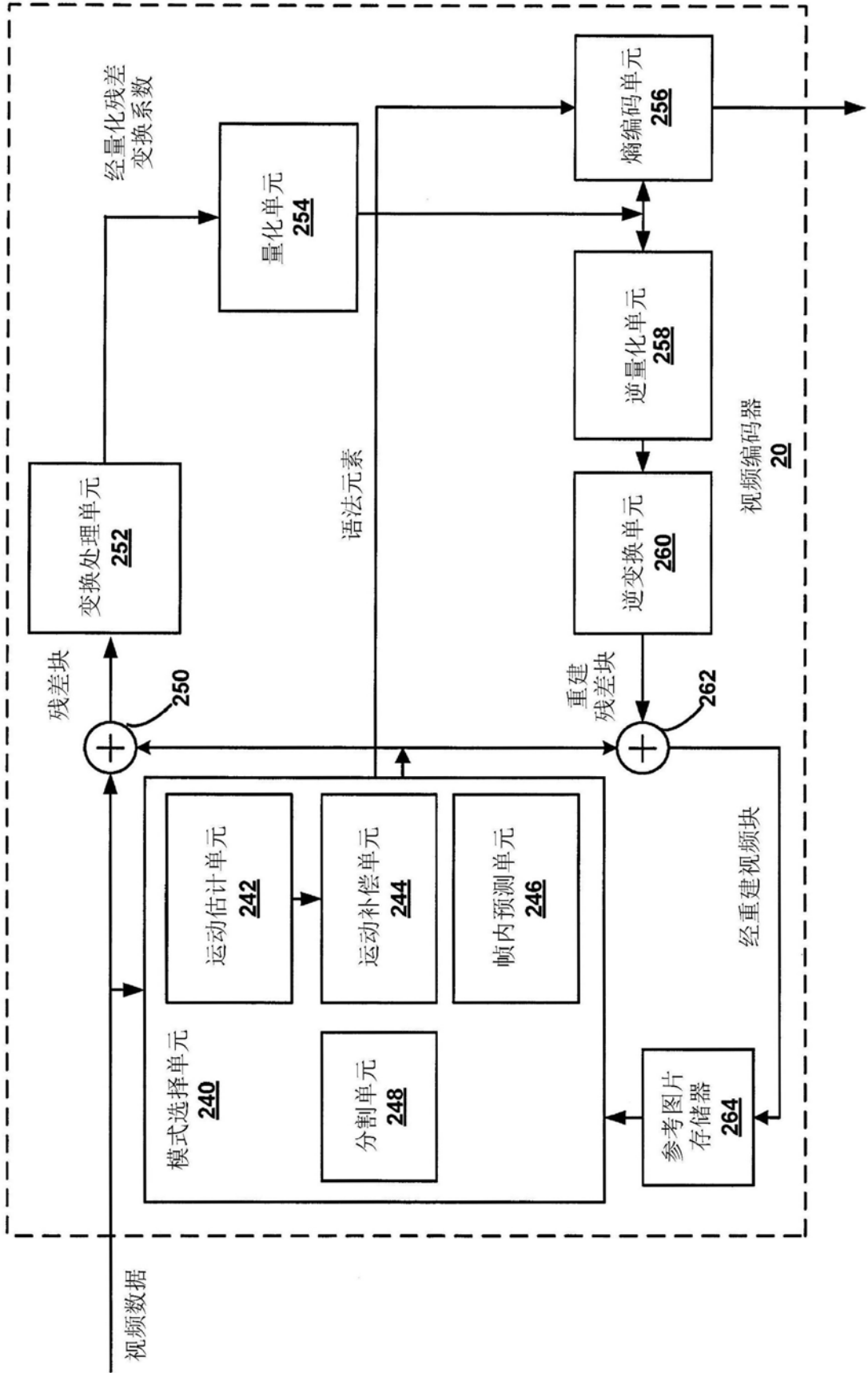


图8

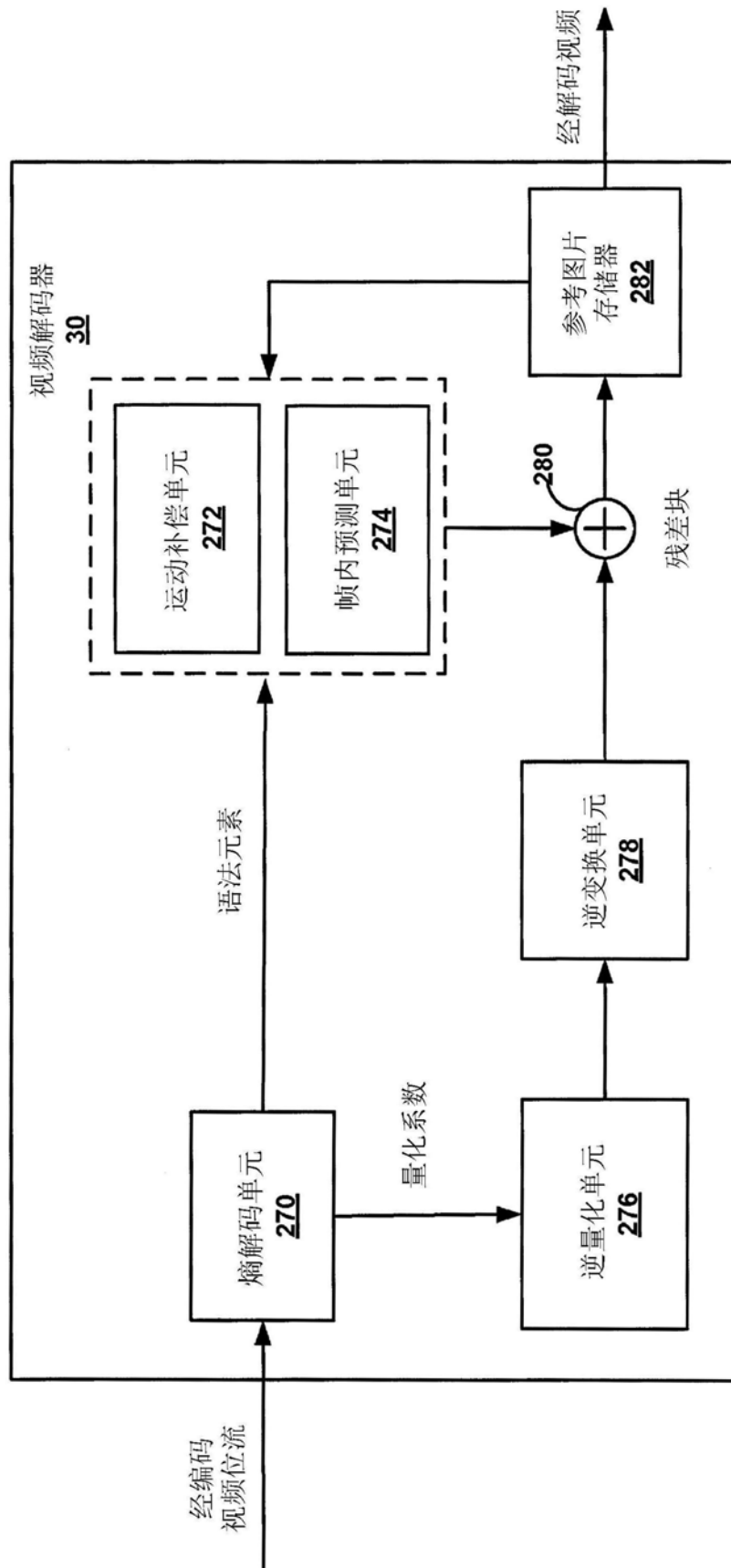


图9

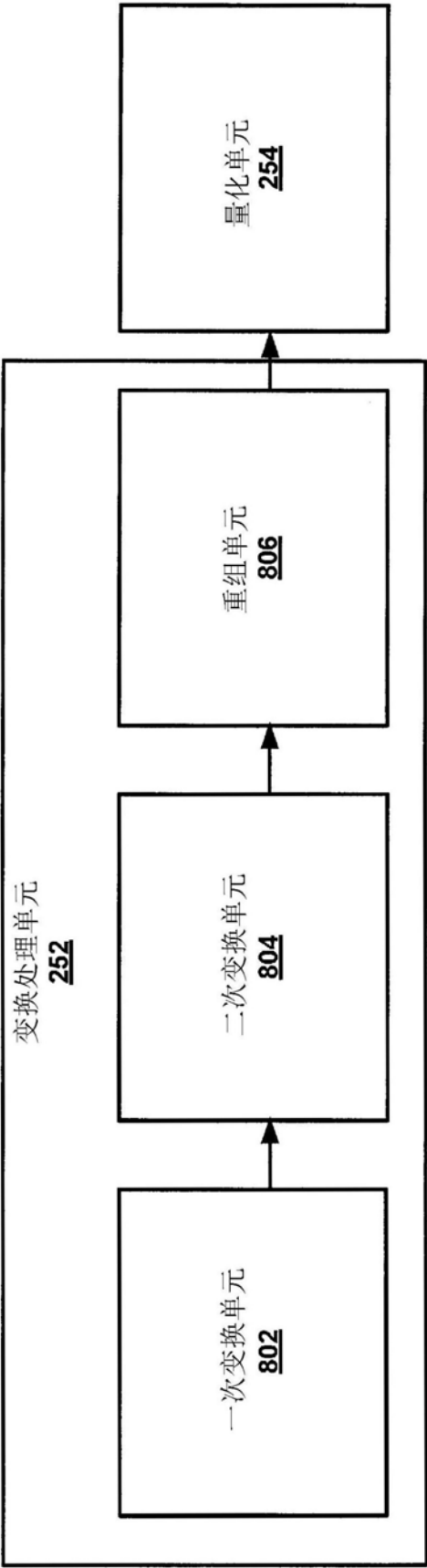


图10A

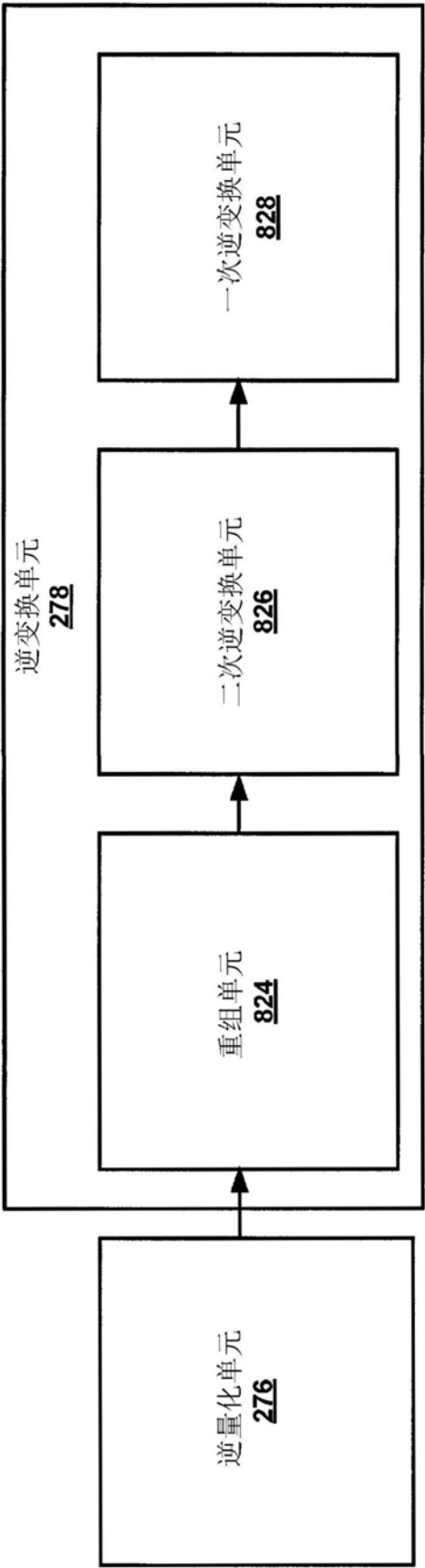


图10B

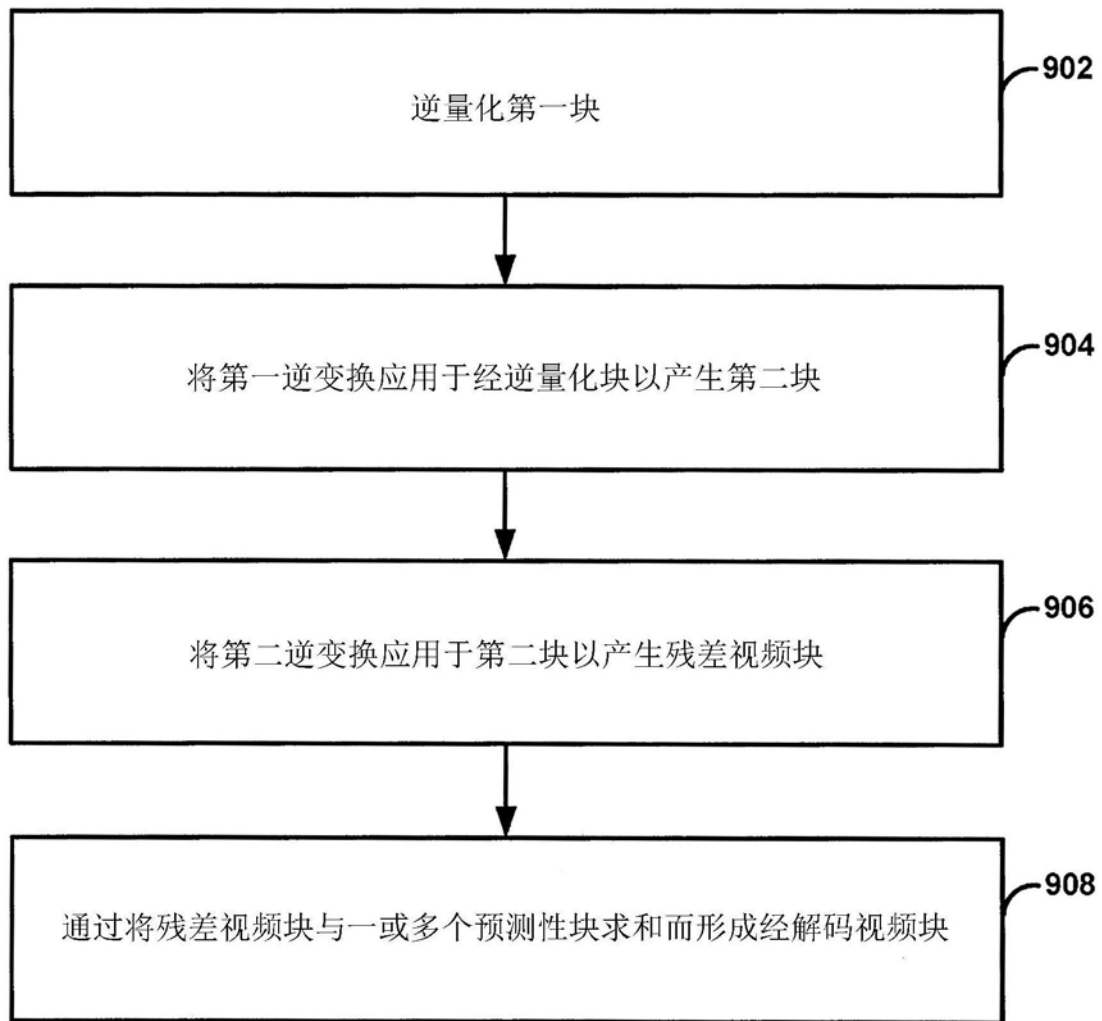


图11

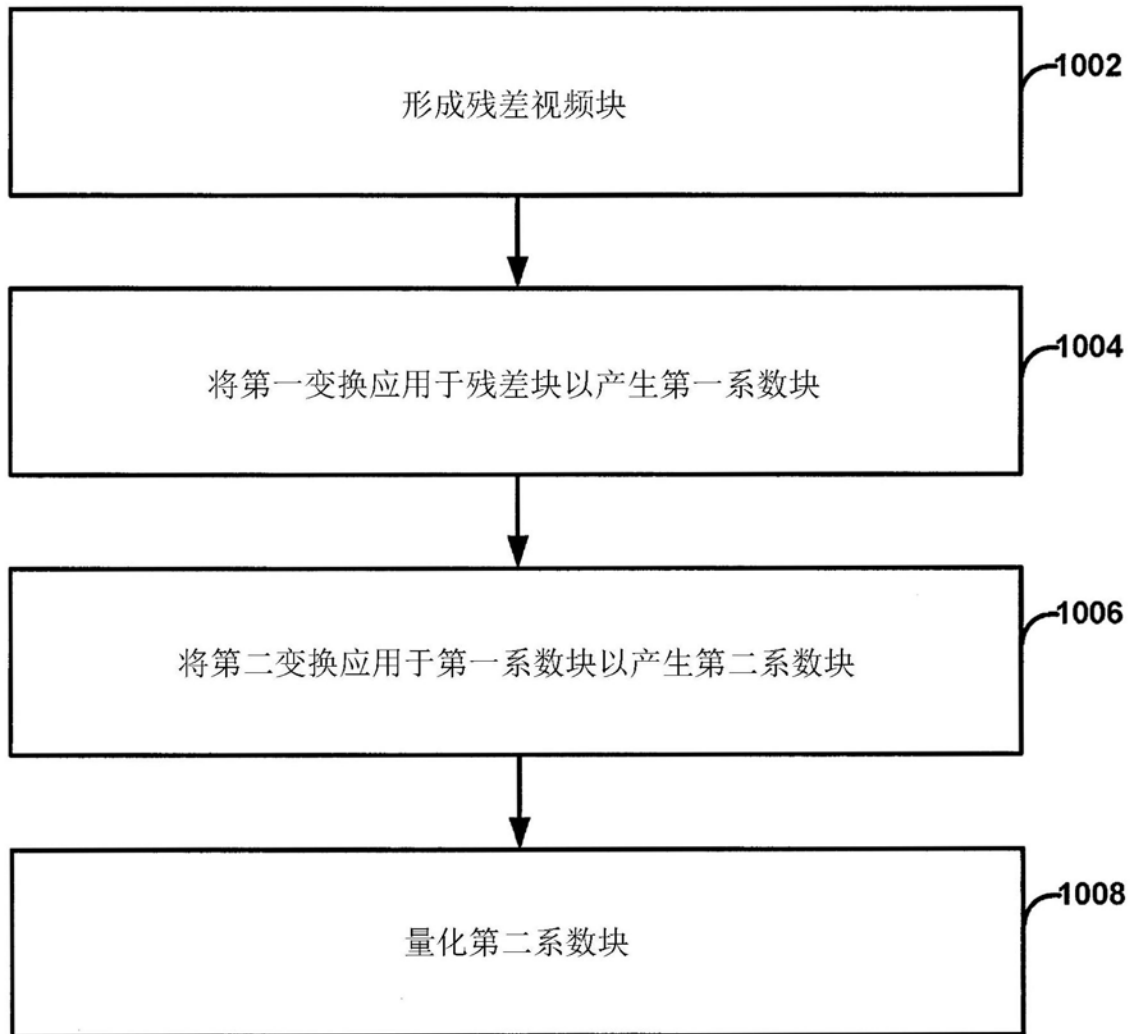


图12

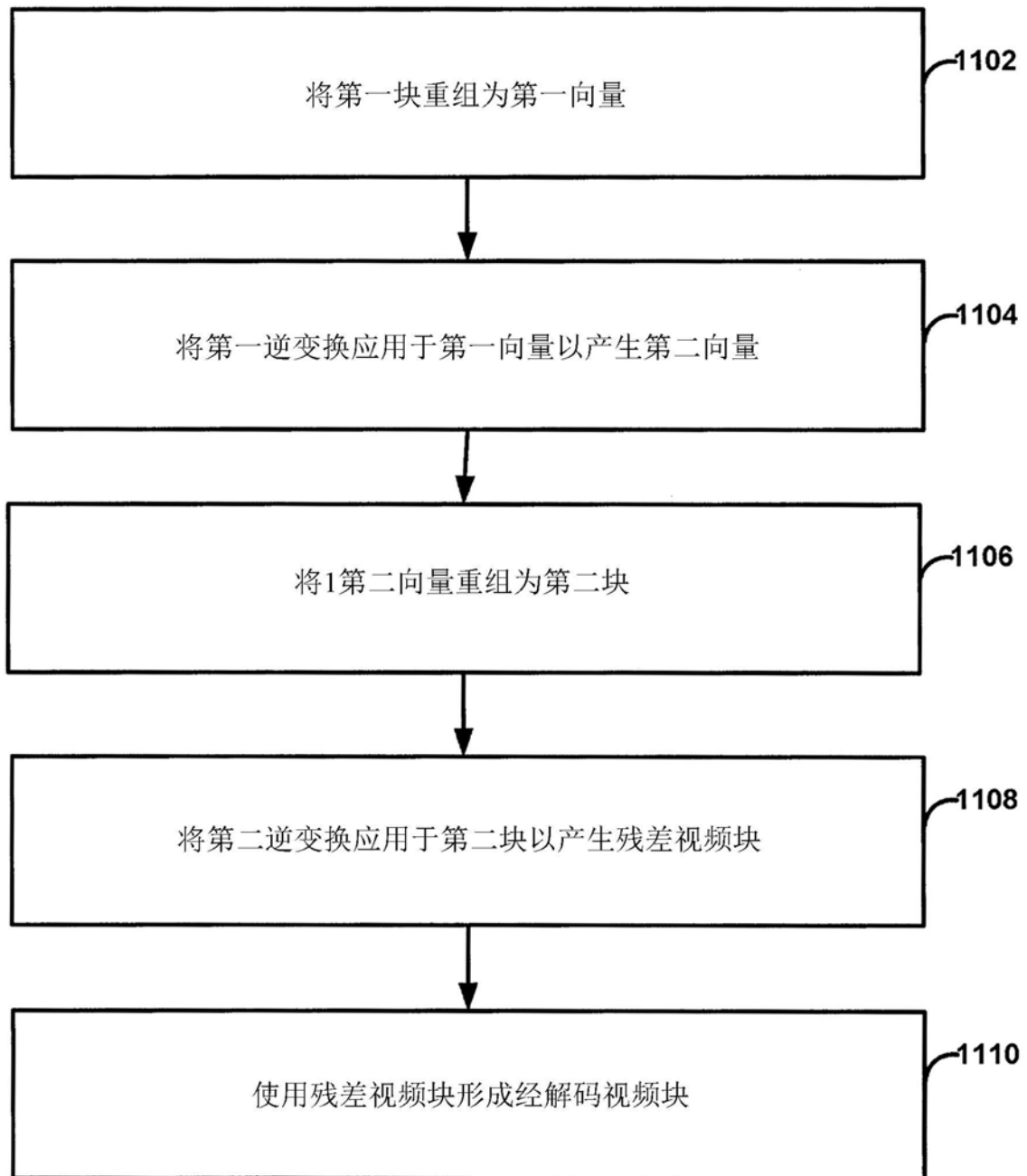


图13

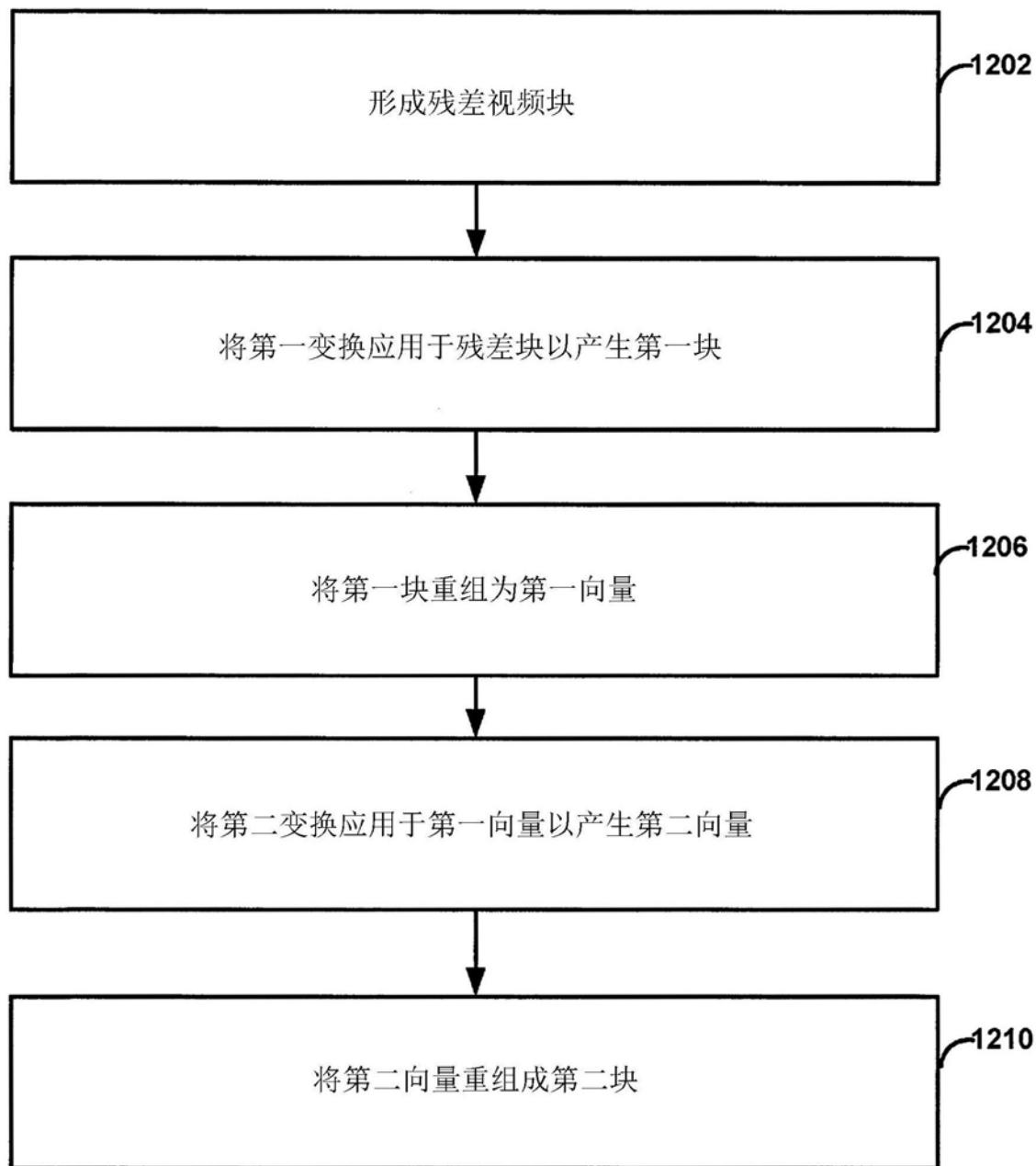


图14