



(12)发明专利

(10)授权公告号 CN 105052146 B

(45)授权公告日 2019.04.05

(21)申请号 201480013342.5

(22)申请日 2014.03.18

(65)同一申请的已公布的文献号
申请公布号 CN 105052146 A

(43)申请公布日 2015.11.11

(30)优先权数据

61/803,100 2013.03.18 US

61/807,281 2013.04.01 US

14/216,232 2014.03.17 US

(85)PCT国际申请进入国家阶段日
2015.09.09

(86)PCT国际申请的申请数据
PCT/US2014/031029 2014.03.18

(87)PCT国际申请的公布数据
W02014/153340 EN 2014.09.25

(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 维贾伊拉加哈万·提拉鲁马莱
张莉 陈颖 马尔塔·卡切维奇

(74)专利代理机构 北京律盟知识产权代理有限公司 11287

代理人 宋献涛

(51)Int.Cl.

H04N 19/597(2006.01)

(56)对比文件

CN 102946535 A, 2013.02.27,

WO 2012171442 A1, 2012.12.20,

US 2012224634 A1, 2012.09.06,

US 2012189060 A1, 2012.07.26,

审查员 胡瞿

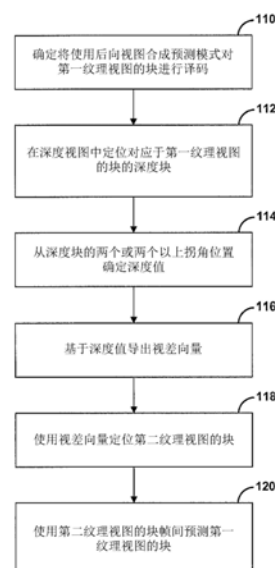
权利要求书4页 说明书29页 附图11页

(54)发明名称

对3D视频译码中的视差向量导出和运动向量预测的简化

(57)摘要

一种用于对三维视频数据进行译码的装置包含视频译码器,其经配置以:确定将使用基于块的视图合成模式对第一纹理视图的第一块进行译码;在深度视图中定位对应于所述第一纹理视图的所述第一块的第一深度块;确定所述第一深度块的两个或两个以上拐角位置的深度值;基于所述深度值导出所述第一块的视差向量;使用所述视差向量定位第二纹理视图的第一块;以及使用所述第二纹理视图的所述第一块帧间预测所述第一纹理视图的所述第一块。



1. 一种对三维3D视频数据进行解码的方法,所述方法包括:
确定将使用基于块的视图合成模式对第一纹理视图的第一块进行解码;
基于相对于所述第一纹理视图的所述第一块位于同一地点的第一深度块的两个或更多个拐角位置的深度值确定所述第一纹理视图的所述第一块的第一视差向量;
对于所述第一纹理视图的所述第一块的第一子块,在第二纹理视图使用针对所述第一纹理视图的所述第一块而确定的所述第一视差向量定位第一预测性块;
对于所述第一块的第二子块,在所述第二纹理视图使用针对所述第一纹理视图的所述第一块而确定的所述第一视差向量定位第二预测性块;
基于所述第一预测性块和所述第二预测性块,解码所述第一纹理视图的所述第一块;
以及
输出包含所述第一纹理视图的所述第一块的经解码视频数据的图片。
2. 根据权利要求1所述的方法,其进一步包括:
确定将使用选自由跳过模式、直接模式、合并模式和帧间预测模式组成的群组的模式对所述第一纹理视图的第二块进行解码;
在深度视图中定位对应于所述第一纹理视图的所述第二块的第二深度块;
确定所述第二深度块的两个或更多个拐角位置的深度值,其中所述第二深度块的所述两个或更多个拐角位置包括所述第一深度块的相同的两个或更多个拐角位置。
3. 根据权利要求1所述的方法,其中所述第一纹理视图包括非基础纹理视图且其中所述第二纹理视图包括基础纹理视图。
4. 根据权利要求1所述的方法,其中所述第一纹理视图的所述第一块包括 $K \times K$ 块,且其中对应于所述第一纹理视图的所述块的所述第一深度块包括 $(K/2) \times (K/2)$ 块或 $K \times K$ 块中的一者。
5. 根据权利要求1所述的方法,其中所述第一纹理视图的所述第一块是根据3D-AVC视频译码标准编码,且所述第二纹理视图是根据H.264/AVC视频译码标准编码。
6. 根据权利要求5所述的方法,其中所述第一纹理视图的所述第一块包含在宏块分区内。
7. 根据权利要求1所述的方法,其中所述第一块是根据3D高效率视频译码HEVC标准编码,且所述第二纹理视图是根据所述HEVC标准编码。
8. 根据权利要求7所述的方法,其中所述第一纹理视图的所述第一块包含在预测单元内。
9. 根据权利要求1所述的方法,其中基于相对于所述第一纹理视图的所述第一块位于同一地点的所述第一深度块的所述两个或更多个拐角位置的所述深度值确定所述第一纹理视图的所述第一块的所述第一视差向量包括从所述第一深度块的两个或更多个拐角位置的所述深度值确定最大深度值。
10. 根据权利要求1所述的方法,其中确定是否将使用所述基于块的视图合成模式对所述第一纹理视图的所述第一块进行解码包括接收语法元素,其中所述语法元素指示将使用所述基于块的视图合成模式对所述第一纹理视图的所述第一块进行解码。
11. 根据权利要求1所述的方法,其进一步包括:在无线通信装置的接收器处接收所述3D视频数据;将所述3D视频数据存储于所述无线通信装置的存储器上;及在所述无线通信

装置的一或多个处理器上处理所述3D视频数据。

12. 根据权利要求11所述的方法, 其中所述无线通信装置包括电话手持机且其中在所述无线通信装置的所述接收器处接收所述3D视频数据包括根据无线通信标准来解调包含所述3D视频数据的信号。

13. 一种用于对三维3D视频数据进行解码的装置, 所述装置包括:

存储器, 其经配置以存储视频数据; 以及

一个或多个处理器, 其经配置以:

确定将使用基于块的视图合成模式对第一纹理视图的第一块进行解码;

基于相对于所述第一纹理视图的所述第一块位于同一地点的第一深度块的两个或更多个拐角位置的深度值确定所述第一纹理视图的所述第一块的第一视差向量;

对于所述第一纹理视图的所述第一块的第一子块, 在第二纹理视图中使用针对所述第一纹理视图的所述第一块而确定的所述第一视差向量定位第一预测性块;

对于所述第一块的第二子块, 在所述第二纹理视图中使用针对所述第一纹理视图的所述第一块而确定的所述第一视差向量定位第二预测性块;

基于所述第一预测性块和所述第二预测性块, 解码所述第一纹理视图的所述第一块; 以及

输出包含所述第一纹理视图的所述第一块的经解码视频数据的图片。

14. 根据权利要求13所述的装置, 其中所述一个或多个处理器进一步经配置以: 确定将使用选自自由跳过模式、直接模式、合并模式和帧间预测模式组成的群组的模式对所述第一纹理视图的第二块进行解码; 在深度视图中定位对应于所述第一纹理视图的所述第二块的第二深度块; 以及确定所述第二深度块的两个或更多个拐角位置的深度值, 其中所述第二深度块的所述两个或更多个拐角位置包括所述第一深度块的相同的两个或更多个拐角位置。

15. 根据权利要求13所述的装置, 其中所述第一纹理视图包括非基础纹理视图且其中所述第二纹理视图包括基础纹理视图。

16. 根据权利要求13所述的装置, 其中所述第一纹理视图的所述第一块包括 $K \times K$ 块, 且其中对应于所述第一纹理视图的所述块的所述第一深度块包括 $(K/2) \times (K/2)$ 块或 $K \times K$ 块中的一者。

17. 根据权利要求13所述的装置, 其中所述第一纹理视图的所述第一块包含在宏块分区内。

18. 根据权利要求13所述的装置, 其中所述第一纹理视图的所述第一块包含在预测单元内。

19. 根据权利要求13所述的装置, 其中为了基于相对于所述第一纹理视图的所述第一块位于同一地点的所述第一深度块的所述两个或更多个拐角位置的所述深度值确定所述第一纹理视图的所述第一块的所述第一视差向量, 所述一个或多个处理器经配置以从所述深度块的两个或更多个拐角位置的所述深度值确定最大深度值。

20. 根据权利要求13所述的装置, 其中所述一个或多个处理器进一步经配置以通过接收语法元素而确定是否将使用所述基于块的视图合成模式对所述第一纹理视图的所述第一块进行解码, 其中所述语法元素指示将使用所述基于块的视图合成模式对所述第一纹理

视图的所述第一块进行解码。

21. 根据权利要求13所述的装置,其中所述装置包括无线通信装置,所述无线通信装置进一步包括经配置以接收所述3D视频数据的接收器。

22. 根据权利要求21所述的装置,其中所述无线通信装置包括电话手持机且其中所述接收器经配置以根据无线通信标准解调包含所述3D视频数据的信号。

23. 一种用于对三维3D视频数据进行解码的装置,所述装置包括:

用于确定将使用基于块的视图合成模式对第一纹理视图的第一块进行解码的装置;

用于基于相对于所述第一纹理视图的所述第一块位于同一地点的第一深度块的两个或更多个拐角位置的深度值确定所述第一纹理视图的所述第一块的第一视差向量的装置;

用于对于所述第一纹理视图的所述第一块的第一子块在第二纹理视图中使用针对所述第一纹理视图的所述第一块而确定的所述第一视差向量定位第一预测性块的装置;

用于对于所述第一块的第二子块在所述第二纹理视图中使用针对所述第一纹理视图的所述第一块而确定的所述第一视差向量定位第二预测性块的装置;

用于基于所述第一预测性块和所述第二预测性块,解码所述第一纹理视图的所述第一块的装置;以及

用于输出包含所述第一纹理视图的所述第一块的经解码视频数据的图片的装置。

24. 根据权利要求23所述的装置,其进一步包括:

用于确定将使用选自由跳过模式、直接模式、合并模式和帧间预测模式组成的群组的模式对所述第一纹理视图的第二块进行解码的装置;

用于在深度视图中定位对应于所述第一纹理视图的所述第二块的第二深度块的装置;

用于确定所述第二深度块的两个或更多个拐角位置的深度值的装置,其中所述第二深度块的所述两个或更多个拐角位置包括所述第一深度块的相同的两个或更多个拐角位置。

25. 根据权利要求24所述的装置,其中所述第一纹理视图包括非基础纹理视图且其中所述第二纹理视图包括基础纹理视图。

26. 根据权利要求24所述的装置,其中所述第一纹理视图的所述第一块是根据3D-AVC视频译码标准编码,且所述第二纹理视图是根据H.264/AVC视频译码标准编码,且其中所述第一纹理视图的所述第一块包含在宏块分区内。

27. 根据权利要求24所述的装置,其中用于基于相对于所述第一纹理视图的所述第一块位于同一地点的所述第一深度块的所述两个或更多个拐角位置的所述深度值确定所述第一纹理视图的所述第一块的所述第一视差向量的装置包括用于从所述深度块的两个或更多个拐角位置的所述深度值确定最大深度值的装置。

28. 一种非易失性存储指令的计算机可读存储媒体,所述指令在由一或多个处理器执行时致使所述一或多个处理器:

确定将使用基于块的视图合成模式对第一纹理视图的第一块进行解码;

基于相对于所述第一纹理视图的所述第一块位于同一地点的第一深度块的两个或更多个拐角位置的深度值确定所述第一纹理视图的所述第一块的第一视差向量;

对于所述第一纹理视图的所述第一块的第一子块,在第二纹理视图中使用针对所述第一纹理视图的所述第一块而确定的所述第一视差向量定位第一预测性块;

对于所述第一块的第二子块,在所述第二纹理视图中使用针对所述第一纹理视图的所

述第一块而确定的所述第一视差向量定位第二预测性块；

基于所述第一预测性块和所述第二预测性块，解码所述第一纹理视图的所述第一块；
以及

输出包含所述第一纹理视图的所述第一块的经解码视频数据的图片。

对3D视频译码中的视差向量导出和运动向量预测的简化

[0001] 本申请案主张2013年3月18日申请的第61/803,100号美国临时申请案、2013年4月1日申请的第61/807,281号美国临时申请案的权益,以上申请案中的每一者的整个内容以引用的方式并入本文中。

技术领域

[0002] 本发明涉及视频编码和解码。

背景技术

[0003] 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)、膝上型或桌上型计算机、数码相机、数字记录装置、数字媒体播放器、视频游戏装置、视频游戏控制台、蜂窝式或卫星无线电电话、视频电话会议装置及其类似者。数字视频装置实施视频压缩技术,例如通过MPEG-2、MPEG-4、ITU-T H.263、ITU-T H.264/MPEG-4第10部分高级视频译码(AVC)、目前正在开发的高效率视频译码(HEVC)标准和此些标准的扩展部分定义的标准中描述的技术,以便更有效地发射、接收和存储数字视频信息。

[0004] 前述标准中的一些的扩展(包含H.264/AVC)可提供用于进行多视图视频译码以便产生立体或三维(“3D”)视频的技术。确切地说,已经提议用于多视图译码的技术以在AVC中使用,与可缩放视频译码(SVC)标准(为对H.264/AVC的可缩放扩展)一起使用,以及与多视图视频译码(MVC)标准(已变为对H.264/AVC的多视图扩展)一起使用。

[0005] 通常,使用两个视图(例如,左视图及右视图)实现立体视频。左视图的图片实质上可与右视图的图片同时显示以实现三维视频效果。举例来说,用户可佩戴偏光被动式眼镜,其将左视图从右视图过滤掉。替代性地,可快速连续展示两个视图的图片,且用户可佩戴主动式眼镜,其以相同频率但具有90度的相位移快速遮光左眼及右眼。

发明内容

[0006] 一般来说,本发明描述用于3D视频译码的技术。确切地说,本发明涉及用于导出视差向量的技术。

[0007] 在一个实例中,一种对3D视频数据进行译码的方法包含:确定将使用基于块的视图合成模式对第一纹理视图的第一块进行译码;在深度视图中定位对应于所述第一纹理视图的所述第一块的第一深度块;确定所述第一深度块的两个或两个以上拐角位置的深度值;基于所述深度值导出所述第一块的视差向量;使用所述视差向量定位第二纹理视图的第一块;以及使用所述第二纹理视图的所述第一块帧间预测所述第一纹理视图的所述第一块。

[0008] 在另一个实例中,一种用于对3D视频数据进行译码的装置包含视频译码器,其经配置以:确定将使用基于块的视图合成模式对第一纹理视图的第一块进行译码;在深度视图中定位对应于所述第一纹理视图的所述第一块的第一深度块;确定所述第一深度块的两

个或两个以上拐角位置的深度值；基于所述深度值导出所述第一块的视差向量；使用所述视差向量定位第二纹理视图的第一块；以及使用所述第二纹理视图的所述第一块帧间预测所述第一纹理视图的所述第一块。

[0009] 在另一个实例中，一种用于对3D视频数据进行译码的装置包含：用于确定将使用基于块的视图合成模式对第一纹理视图的第一块进行译码的装置；用于在深度视图中定位对应于所述第一纹理视图的所述第一块的第一深度块的装置；用于确定所述第一深度块的两个或两个以上拐角位置的深度值的装置；用于基于所述深度值导出所述第一块的视差向量的装置；用于使用所述视差向量定位第二纹理视图的第一块的装置；用于使用所述第二纹理视图的所述第一块帧间预测所述第一纹理视图的所述第一块的装置。

[0010] 在另一个实例中，一种存储指令的计算机可读存储媒体，所述指令在由一或多个处理器执行时致使所述一或多个处理器：确定将使用基于块的视图合成模式对第一纹理视图的第一块进行译码；在深度视图中定位对应于所述第一纹理视图的所述第一块的第一深度块；确定所述第一深度块的两个或两个以上拐角位置的深度值；基于所述深度值导出所述第一块的视差向量；使用所述视差向量定位第二纹理视图的第一块；以及使用所述第二纹理视图的所述第一块帧间预测所述第一纹理视图的所述第一块。

[0011] 随附图式和以下描述中陈述一或多个实例的细节。其它特征、目标及优点将从所述描述及图式以及权利要求书而显而易见。

附图说明

[0012] 图1是说明可利用本发明中描述的技术的实例视频编码及解码系统的框图。

[0013] 图2为说明实例多视图解码次序的概念图。

[0014] 图3为说明用于多视图译码的实例预测结构的概念图。

[0015] 图4是基于后向扭曲的基于块的视图合成预测的概念可视化。

[0016] 图5是展示用于相邻块视差向量导出的空间相邻块的概念图。

[0017] 图6是说明涉及使用相邻块的后向视图合成预测(BVSP)的技术的概念图。

[0018] 图7展示可用于根据本发明的技术导出视差向量的深度块的实例。

[0019] 图8是说明可实施本发明中描述的技术的实例视频编码器的框图。

[0020] 图9是说明可实施本发明中描述的技术的实例视频解码器的框图。

[0021] 图10是说明根据本发明中描述的技术的视差向量导出和帧间预测译码的实例操作的流程图。

[0022] 图11是说明根据本发明中描述的技术的视差向量导出和帧间预测译码的另一实例操作的流程图。

具体实施方式

[0023] 本发明描述涉及视差向量导出的技术，且更具体地说，本发明描述其中视频译码器(例如，视频编码器或视频解码器)导出用于在三维(3D)视频译码中使用的当前视图的当前图片内的当前块的视差向量的技术。视频译码器可使用视差向量来定位不同视图中的对应块。以此方式，视差向量可表示两个不同视图中的两个相似视频块之间的视差。如下文将更详细地解释，视频译码器可使用视差向量作为视差运动向量，且可基于通过视差运动向

量定位的块来预测当前块。视频译码器还可使用视差向量用于其它目的。作为一个实例,视频译码器可使用视差向量来定位另一视图中的对应块,且随后使用所定位块的运动信息用于确定当前块的运动信息。此外,可存在视差向量的再其它使用。在本发明中,术语“当前”一般用于指当前正译码的视图、图片或块。因此,当前块大体上表示正译码的视频数据块,与已经译码的块相反或与尚待译码的块相反。

[0024] 当前图片的当前块的视差向量是指向在与当前图片不同的视图中的对应图片中的对应块的向量。因此,使用视差向量,视频译码器可在对应图片中定位对应于当前图片的当前块的块。在此情况下,对应图片是与当前图片为相同的时间实例但在不同视图中的图片。对应图片中的对应块和当前图片中的当前块可包含相似视频内容;然而,当前图片中的当前块的位置与对应图片中的对应块的位置之间存在至少水平视差。当前块的视差向量提供对应图片中的块与当前图片中的当前块之间的此水平视差的量度。在一些情况下,对应图片内的块的位置与当前图片内的当前块的位置之间也可以存在垂直视差;然而,在许多情况下所述垂直视差将为零。当前块的视差向量还可提供对应图片中的块与当前图片中的当前块之间的此垂直视差的量度。视差向量含有两个分量(x分量和y分量),但在许多情况下垂直分量将等于零。当前视图的当前图片和不同视图的对应图片所显示的时间可为相同的,也就是说当前图片和对应图片是同一时间实例的图片。

[0025] 在2D视频译码中,仅由一个视图分量表示帧,有时称为纹理视图分量,或简称为纹理。在一些类型的3D视频译码中,存在两个视图分量:纹理视图分量和深度视图分量,或简称为纹理和深度。举例来说,每一视图可包含纹理视图和深度视图,其中视图包含多个视图分量,例如纹理视图包含多个纹理视图分量,且深度视图包含多个深度视图分量。每一纹理视图分量与深度视图分量相关联以形成视图的视图分量。深度视图分量表示纹理视图分量中的对象的相对深度。深度视图分量和纹理视图分量可单独地可解码。

[0026] 本发明描述用于导出视差向量的技术。用于导出视差向量的一种此类技术可与后向视图合成预测(BVSP)模式结合使用。视频译码器可确定将使用BVSP模式译码的第一纹理视图的块。视频译码器可在深度视图中定位对应于第一纹理视图的块的深度块且确定深度块的两个或两个以上拐角位置的深度值。基于深度值,视频译码器可导出块的视差向量,且使用视差向量定位第二纹理视图的块。视频译码器可随后使用第二纹理视图的块帧间预测第一纹理视图的块。在此方面,且如将在下文更详细地阐释,本发明的技术可通过使用对应深度块的仅拐角样本确定用于BVSP模式的视差向量而简化视差向量导出过程。

[0027] 在另一实例技术中,对于第一视图的块,视频译码器可定位深度视图中对应于第一纹理视图的块的深度块,且基于深度块的至少一个深度值,导出第一纹理视图的块的视差向量。视频译码器可随后基于所导出的视差向量对所述块的第一子块进行译码且基于同一个所导出的视差向量对宏块的第二子块进行译码。在此方面且如将在下文更详细地阐释,本发明的技术可通过导出块的一个视差向量且使用所述视差用于块的两个或两个以上子块而简化视差向量导出过程。取决于对特定子块指定的译码模式,视频译码器可使用所导出的视差向量作为视差运动向量或可使用视差向量以识别不同视图中的对应块,且根据所述对应块确定用于预测子块的运动信息。

[0028] 图1是说明可经配置以执行本发明中描述的用于视差向量导出的技术的实例视频编码和解码系统10的框图。如图1中所展示,系统10包含源装置12,其产生稍后由目的地装

置14解码的经编码视频数据。源装置12和目的地装置14可包括广泛范围的装置中的任一者,包含桌上型计算机、笔记型(即,膝上型)计算机、平板计算机、机顶盒、电话手持机(例如所谓的“智能”电话)、所谓的“智能”平板计算机、电视机、相机、显示装置、数字媒体播放器、视频游戏控制台、视频流式传输装置或类似者。在一些情况下,可装备源装置12和目的地装置14以用于无线通信。

[0029] 系统10可根据不同视频译码标准、专属标准或多视图译码的任何其它方式操作。以下描述视频译码标准的几个实例,且不应被视为具限制性。视频译码标准包含ITU-T H.261、ISO/IEC MPEG-1Visual、ITU-T H.262或ISO/IEC MPEG-2Visual、ITU-T H.263、ISO/IEC MPEG-4Visual及ITU-T H.264(也被称为ISO/IEC MPEG-4AVC),包含其可缩放视频译码(SVC)及多视图视频译码(MVC)扩展。MVC的最新联合草案在“用于通用视听服务的高级视频译码”(2010年3月的ITU-T建议H.264)中描述,其整个内容以引入的方式并入本文中。MVC的另一联合草案在“用于通用视听服务的高级视频译码(Advanced video coding for generic audiovisual services)”(ITU-T建议H.264,2011年6月)中描述,其整个内容以引入的方式并入本文中。一些额外视频译码标准包含基于AVC的MVC+D和3D-AVC。另外,已由ITU-T视频译码专家组(VCEG)和ISO/IEC动画专家组(MPEG)的视频译码联合合作小组(JCT-VC)开发新视频译码标准,即,高效率视频译码(HEVC)。

[0030] 仅出于说明的目的,以根据例如3D-AVC的H.264标准的实例来描述本发明中描述的技术。然而,本发明中描述的技术不应被视为限于这些实例性标准,且可扩展到用于多视图译码或3D视频译码(例如,3D-HEVC)的其它视频译码标准,或可扩展到与不一定基于特定视频译码标准的多视图译码或3D视频译码相关的技术。举例来说,本发明中描述的技术由用于多视图译码的视频编码器/解码器(编解码器)实施,其中多视图译码包含两个或两个以上视图的译码。

[0031] 目的地装置14可经由链路16接收待解码的经编码视频数据。链路16可包括能够将经编码视频数据从源装置12移动到目的地装置14的任何类型的媒体或装置。在一个实例中,链路16可包括使得源装置12能够实时地将经编码视频数据直接发射到目的地装置14的通信媒体。经编码视频数据可根据通信标准(例如,无线通信协议)加以调制,并发射到目的地装置14。通信媒体可包括任何无线或有线通信媒体,例如射频(RF)频谱或一或多个物理发射线路。通信媒体可形成分组网络(例如,局域网、广域网或全球网络,例如因特网)的部分。通信媒体可包含路由器、交换器、基站或可以用于促进从源装置12到目的地装置14的通信的任何其它装备。

[0032] 替代地,可将经编码数据从输出接口22输出到存储装置34。类似地,可通过输入接口从存储装置34存取经编码数据。存储装置34可包含多种分布式或本地存取的数据存储媒体中的任一者,例如硬盘驱动器、蓝光光盘、DVD、CD-ROM、快闪存储器、易失性或非易失性存储器或用于存储经编码视频数据的任何其它合适数字存储媒体。在另一实例中,存储装置34可对应于可保持由源装置12产生的经编码视频的文件服务器或另一中间存储装置。目的地装置14可从存储装置34经由流式传输或下载来存取所存储的视频数据。文件服务器可为能够存储经编码视频数据且将经编码视频数据传输到目的器件14的任何类型的服务器。实例文件服务器包含网络服务器(例如,用于网站)、FTP服务器、网络附接存储(NAS)器件或本地磁盘驱动器。目的地装置14可经由任何标准数据连接(包含因特网连接)来存取经编码的

视频数据。此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、电缆调制解调器等)或两者的组合。经编码的视频数据从存储装置34的发射可为流式发射、下载发射或两者的组合。

[0033] 本发明的用于视差向量导出的技术不必限于无线应用或设定。所述技术可应用于视频译码以支持多种多媒体应用中的任一者,例如空中电视广播、有线电视发射、卫星电视发射、串流视频发射(例如,经由因特网)、编码数字视频以存储于数据存储媒体上、解码存储于数据存储媒体上的数字视频,或其它应用。在一些实例中,系统10可经配置以支持单向或双向视频发射,以支持例如视频流式传输、视频回放、视频广播和/或视频电话等应用。

[0034] 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。在一些情况下,输出接口22可包含调制器/解调器(调制解调器)及/或发射器。在源装置12中,视频源18可包含例如视频俘获装置(例如,摄像机)、含有先前俘获的视频的视频存档、用于从视频内容提供者接收视频的视频馈入接口及/或用于产生计算机图形数据作为源视频的计算机图形系统,或此类源的组合等源。作为一个实例,如果视频源18是摄像机,那么源装置12和目的地装置14可以形成所谓的摄像机电话或视频电话。然而,本发明中所描述的技术一般来说可适用于视频译码,且可应用于无线和/或有线应用。

[0035] 可由视频编码器20编码所俘获视频、预俘获的视频或计算机产生的视频。经编码视频数据可经由源装置12的输出接口22直接发射到目的地装置14。经编码视频数据还可(或替代地)存储到存储装置34上以供稍后由目的地装置14或其它装置存取以用于解码及/或回放。

[0036] 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。在一些情况下,输入接口28可包含接收器及/或调制解调器。目的地装置14的输入接口28经由链路16接收经编码视频数据。经由链路16传送或在存储装置34上提供的经编码视频数据可包含由视频编码器20所产生的多种语法元素以供由例如视频解码器30的视频解码器用于解码视频数据。此类语法元素可与在通信媒体上发射、存储在存储媒体上或存储在文件服务器中的经编码视频数据包含在一起。

[0037] 显示装置32可与目的地装置14一起集成或在目的地装置外部。在一些实例中,目的地装置14可包含集成式显示装置,且还经配置以与外部显示装置介接。在其它实例中,目的地装置14可为显示装置。一般来说,显示装置32将经解码视频数据显示给用户,且可包括多种显示装置中的任一者,例如液晶显示器(LCD)、等离子显示器、有机发光二极管(OLED)显示器或另一类型的显示装置。

[0038] 尽管图1中未展示,但在一些方面中,视频编码器20和视频解码器30可各自与音频编码器和解码器集成,且可包含适当的多路复用器-多路分用器(MUX-DEMUX)单元或其它硬件和软件,以处置对共同数据流或单独数据流中的音频和视频两者的编码。在一些实例中,如果适用,多路复用器-多路分用器单元可符合ITU H.223多路复用器协议,或例如用户数据报协议(UDP)等其它协议。

[0039] 视频编码器20和视频解码器30各自可经实施为例如一或多个微处理器、数字信号处理器(DSP)、专用集成电路(ASIC)、现场可编程门阵列(FPGA)、离散逻辑、软件、硬件、固件或其任何组合的多种合适编码器电路中的任一者。举例来说,可从设备或装置的角度描述本发明中描述的技术。作为一个实例,所述设备或装置可包含视频解码器30(例如,作为无

线通信装置的的目的地装置14),且视频解码器30可包含经配置以实施本发明中描述的技术(例如,根据本发明中描述的技术解码视频数据)的一或多个处理器。作为另一实例,所述设备或装置可包含包含视频解码器30的微处理器或集成电路(IC),且所述微处理器或IC可为目的地装置14或另一种类型的装置的部分。相同情况可适用于视频编码器20(即,类似于源装置12和/或微控制器或IC的设备或装置包含视频编码器20,其中视频编码器20经配置以根据本发明中描述的技术编码视频数据)。

[0040] 当部分地用软件实施所述技术时,装置可将用于所述软件的指令存储于合适的非暂时性计算机可读媒体中且使用一或多个处理器用硬件执行所述指令以执行本发明的技术。视频编码器20和视频解码器30中的每一者可以包含在一或多个编码器或解码器中,所述编码器或解码器中的任一者可以集成为相应装置中的组合编码器/解码器(编解码器)的部分。

[0041] 视频序列通常包含来自视图的一系列视频图片。图片群组(GOP)通常包括一系列一或多个视频图片。GOP可包含GOP的标头、GOP的一或多个图片的标头或其它地方中的语法数据,其描述GOP中包含的图片的数目。每一图片可包含图片语法数据,其描述用于相应图片的编码模式。视频编码器20通常对个别视频图片内的视频块进行操作以便编码视频数据。如H.264标准中界定,视频块可对应于宏块、宏块的分区且可能对应于分区的子块。视频块可以具有固定或变化的大小,并且可以根据指定译码标准而有不同大小。每一视频图片可包含多个切片。每一切片可包含多个块。

[0042] 作为实例,ITU-T H.264标准支持各种块大小的帧内预测(例如用于明度分量的16乘16、8乘8或4乘4及用于色度分量的8x8),以及各种块大小的帧间预测,例如用于明度分量的16x16、16x8、8x16、8x8、8x4、4x8及4x4及用于色度分量的对应缩放大小。在本发明中,“NxN”及“N乘N”可互换地使用以指代块在垂直及水平尺寸方面的像素尺寸(例如,16x16像素或16乘16像素)。一般来说,16x16块将在垂直方向上具有16个像素($y=16$),且在水平方向上具有16个像素($x=16$)。同样地,NxN块一般在垂直方向上具有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。一块中的像素可布置成若干行和若干列。此外,块未必需要在水平方向与垂直方向上具有相同数目的像素。举例来说,块可包括NxM像素,其中M未必等于N。

[0043] 在对块进行帧内模式编码(例如,帧内预测)时,块可包含描述块的帧内预测模式的数据。作为另一实例,当块经帧间模式编码(例如,经帧间预测)时,所述块可包含界定块的运动向量的信息。此运动向量参考同一视图中的参考图片(例如,时间运动向量),或参考另一视图中的参考图片(例如,视差运动向量)。界定块的运动向量的数据描述(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)。另外,在进行帧间预测时,块可包含参考索引信息,例如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,RefPicList0或RefPicList1)。

[0044] 在H.264标准中,在帧内预测性或帧间预测性译码后,视频编码器20计算宏块的残余数据。残余数据可对应于未经编码图片的像素与H.264中的宏块的预测值之间的像素差。

[0045] 在一些实例中,在进行任何变换以产生变换系数之后,视频编码器20执行变换系数的量化。量化一般是指将变换系数量化以可能减少用以表示系数的数据量从而提供进一步压缩的过程。量化过程减少与系数中的一些或全部相关联的位深度。举例来说,n位值在

量化期间被下舍入到 m 位值,其中 n 大于 m 。

[0046] 在一些实例中,视频编码器20利用预定义扫描次序扫描经量化变换系数以产生可经熵编码的串行化向量。在其它实例中,视频编码器20执行自适应扫描。在扫描经量化变换系数以形成一维向量之后,在一些实例中,视频编码器20根据上下文自适应可变长度译码(CAVLC)、上下文自适应二进制算术译码(CABAC)、基于语法的上下文自适应二进制算术译码(SBAC)、概率区间分割熵(PIPE)译码或另一熵编码方法(作为几个实例)对所述一维向量进行熵编码。视频编码器20还对与经编码视频数据相关联的语法元素进行熵编码以供视频解码器30在对视频数据解码时使用。

[0047] 为了执行CABAC,视频编码器20可将上下文模型内的上下文指派给待发射的符号。举例来说,所述上下文可以涉及符号的相邻值是否为非零。为了执行CAVLC,视频编码器20可选择用于待发射的符号的可变长度码。VLC中的码字可经构造使得相对较短码对应于更有可能符号,而较长码对应于不太可能符号。以此方式,使用VLC可以实现(例如)优于对待发射的每一符号使用等长码字的位节省。概率确定可基于指派给符号的上下文。

[0048] 视频解码器30实施视频编码器20的技术的逆技术。举例来说,视频解码器30对经编码视频位流进行解码且通过逆量化和逆变换确定残余块。视频解码器30将残余块与先前经解码图片的块进行求和以确定图片内的块的像素值。

[0049] 本发明中描述的某些技术可由视频编码器20和视频解码器30两者执行。作为一个实例,视频编码器20可导出视差向量,作为确定如何对视频数据块进行编码的部分。视频编码器20还可导出视差向量,作为用以产生参考图片的解码环路的部分。视频解码器30可执行由视频编码器20执行的同一视差向量导出技术,作为对视频块进行解码的部分。本发明可时常参考由视频解码器30执行的技术,但除非另外规定,否则应假定参考视频解码器30描述的技术也可以由视频编码器20执行。

[0050] 如上文所描述,本发明中描述的技术是针对3D视频译码。为了更好地理解所述技术,以下描述一些H.264/AVC译码技术、从H.264/MVC扩展和高效率视频译码(HEVC)标准的角度的多视图视频译码以及3D-AVC技术。

[0051] 对于H.264/高级视频译码(AVC),对宏块实施视频编码或解码(例如,译码),其中宏块表示经帧间预测或经帧内预测(即,经帧间预测编码或解码或者经帧内预测编码或解码)的帧的一部分。举例来说,在H.264/AVC中,每一帧间宏块(MB)(例如,经帧间预测宏块)可以四个不同方式分割:一个 16×16 MB分区,两个 16×8 MB分区,两个 8×16 MB分区,或四个 8×8 MB分区。一个MB中的不同MB分区可具有用于每一方向的不同参考索引值(即,RefPicList0或RefPicList1)。当MB未分割成多个(1个以上)MB分区时,其在每一方向中针对整个MB分区具有仅一个运动向量。

[0052] 作为视频译码(编码或解码)的部分,视频解码器30可经配置以构造一个或两个参考图片列表,称为RefPicList0和RefPicList1。参考图片列表识别可用以对帧或切片的宏块进行帧间预测的参考图片。举例来说,视频编码器20可用信号表示参考索引和参考图片列表识别符。视频解码器30可接收参考索引和参考图片列表识别符,且根据参考索引和参考图片列表识别符确定将用于对当前宏块进行帧间预测解码的参考图片。

[0053] 当MB分割成四个 8×8 MB分区时,每一 8×8 MB分区可进一步分割成子块。存在四个不同方式来从 8×8 MB分区得到H.264/AVC子块:一个 8×8 子块,两个 8×4 子块,两个 4×8 子块,或四

个4x4子块。每一H.264/AVC子块可具有每一方向中的不同运动向量,但针对每一方向共享同一参考图片索引。8x8MB分区分割成子块的方式称为子块分割。

[0054] 本发明将大体上使用术语块来指代任何视频数据块。举例来说,在H.264译码及其扩展的上下文中,块可指代宏块、宏块分区、子块或任何其它类型的块中的任一者。在HEVC及其扩展的上下文中,块可指代PU、TU、CU或任何其它类型的块中的任一者。如本发明中所使用的子块大体上指代较大块的任何部分。子块还可自身简称为块。当具体地说提及如H.264/AVC中界定的子块时,可使用术语H.264/AVC子块。否则,如在本发明中所使用的子块大体上既定为涵盖H.264/AVC子块以及上述其它类型的子块的通用术语。

[0055] 对于多视图视频译码,存在多个不同视频译码标准。为了避免混淆,当本发明一般地描述多视图视频译码时,本发明使用短语“多视图视频译码”。一般来说,在多视图视频译码中,存在基础视图和一或多个非基础或相依视图。基础视图可在不参考相依视图中的任一者的情况下完全解码(即,基础视图仅以时间运动向量进行帧间预测)。这允许未经配置以用于多视图视频译码的编解码器仍接收完全可解码的至少一个视图(即,可提取出基础视图且丢弃其它视图,从而允许未经配置以用于多视图视频译码的解码器仍对视频内容进行解码,但无3D体验)。所述一或多个相依视图可相对于基础视图或相对于另一相依视图(即,视差补偿预测)或相对于同一视图中的其它图片(即,运动补偿预测)而经帧间预测。

[0056] 鉴于一般使用“多视图视频译码”,首字母缩写MVC与H.264/AVC的扩展相关联。因此,当本发明使用首字母缩写MVC时,本发明具体地参考对H.264/AVC视频译码标准的扩展。H.264/AVC的MVC扩展除时间运动向量之外还依赖于视差运动向量作为另一种类型的运动向量。JCT-3V和MPEG还已经开发称为MVC加深度(MVC+D)的另一视频译码标准。MVC+D应用与用于纹理和深度两者的MVC的那些工具相同的低层级译码工具,其中深度的解码独立于纹理的解码且反之亦然。举例来说,在MVC中,帧仅由一个视图分量表示,所述视图分量称为纹理视图分量或简称为纹理。在MVC+D中,存在两个视图分量:纹理视图分量和深度视图分量,或简称为纹理和深度。举例来说,在MVC+D中,每一视图包含纹理视图和深度视图,其中视图包含多个视图分量,纹理视图包含多个纹理视图分量,且深度视图包含多个深度视图分量。

[0057] 每一纹理视图分量与深度视图分量相关联以形成视图的视图分量。深度视图分量表示纹理视图分量中的对象的相对深度。在MVC+D中,深度视图分量和纹理视图分量是单独可解码的。举例来说,视频解码器30可实施MVC编解码器的两个实例,其中第一编解码器解码纹理视图分量且第二编解码器解码深度视图分量。这两个编解码器可独立于彼此而执行,因为纹理视图分量和深度视图分量是单独地经编码。

[0058] 在MVC+D中,深度视图分量始终紧接着跟随相关联(例如,对应)的纹理视图分量。以此方式,MVC+D支持纹理优先译码,其中纹理视图分量是在深度视图分量之前经解码。

[0059] 纹理视图分量及其相关联(例如,对应)深度视图分量可包含相同的图片次序计数(POC)值和view_id(即,纹理视图分量及其相关联深度视图分量的POC值和view_id是相同的)。POC值指示纹理视图分量的显示次序且view_id指示纹理视图分量和深度视图分量属于的视图。

[0060] 图2展示典型MVC解码次序(即,位流次序)。解码次序布置被称为时间优先译码。应注意,存取单元的解码次序可并不相同于输出或显示次序。在图2中,S0到S7各自是指多视图视频的不同视图。T0到T8各自表示一个输出时间实例。存取单元可包含针对一个输出时

间实例的所有视图的经译码图片。举例来说,第一存取单元可包含针对时间实例T0的所有视图S0到S7,第二存取单元可包含针对时间实例T1的所有视图S0到S7,等等。

[0061] 出于简洁目的,本发明可使用以下定义:

[0062] 视图分量:单个存取单元中的视图的经译码表示。当视图包含经译码纹理和深度表示两者时,视图分量可包含纹理视图分量和深度视图分量。

[0063] 纹理视图分量:单一存取单元中的视图的纹理的经译码表示。

[0064] 深度视图分量:单一存取单元中的视图的深度的经译码表示。

[0065] 如上文所论述,在本发明的上下文中,视图分量、纹理视图分量和深度视图分量可总体上被称作层。在图2中,视图中的每一者包含若干图片集合。举例来说,视图S0包含图片集合0、8、16、24、32、40、48、56及64,视图S1包含图片集合1、9、17、25、33、41、49、57及65,等等。每一集合包含两个图片:一个图片称为纹理视图分量,且另一图片称为深度视图分量。视图的图片集合内的纹理视图分量及深度视图分量可认为是对应于彼此。举例来说,视图的图片集合内的纹理视图分量被视为对应于视图的图片的所述组内的深度视图分量,且反过来也一样(即,深度视图分量对应于其在所述组中的纹理视图分量,且反过来也一样)。如本发明中所使用,对应于深度视图分量的纹理视图分量可被视为纹理视图分量及深度视图分量为单个存取单元的另一视图的部分。

[0066] 纹理视图分量包含所显示的实际图像内容。举例来说,纹理视图分量可包含明度(Y)和色度(Cb和Cr)分量。深度视图分量可指示其对应纹理视图分量中的像素的相对深度。作为一个实例类似,深度视图分量类似于仅包含明度值的灰阶图像。换句话说,深度视图分量可并不传达任何图像内容,而是提供纹理视图分量中的像素的相对深度的量度。

[0067] 举例来说,深度视图分量中的纯白色像素指示对应纹理视图分量中的其对应像素较接近于观察者的视角,且深度视图分量中的纯黑色像素指示对应纹理视图分量中的其对应像素距观察者的视角较远。黑色与白色之间的各种灰度梯度指示不同深度等级。举例来说,深度视图分量中的深灰色像素指示其在纹理视图分量中的对应像素比深度视图分量中的浅灰色像素更远。因为仅需要灰阶来识别像素的深度,所以深度视图分量无需包含色度分量,这是由于深度视图分量的色值可并不服务于任何目的。以上阐释既定为用于使深度图像相关到纹理图像的类似。深度图像中的深度值并不实际上表示灰度梯度,而是实际上表示8位或其它位大小的深度值。

[0068] 仅使用明度值(例如,强度值)来识别深度的深度视图分量是出于说明的目的而提供,且不应被视为限制性的。在其它实例中,可利用任何技术来指示纹理视图分量中的像素的相对深度。

[0069] 图3展示用于多视图视频译码的典型MVC预测结构(包含每一视图内的图片间预测和视图之间的视图间预测两者)。预测方向由箭头指示,箭头指向的对象使用箭头出发的对象作为预测参考。在MVC中,由视差运动补偿支持视图间预测,所述视差运动补偿使用H.264/AVC运动补偿的语法但允许将不同视图中的图片用作参考图片。

[0070] 在图3的实例中,说明八个视图(具有视图ID“S0”到“S7”),且对于每一视图说明十二个时间位置(“T0”到“T11”)。即,图3中的每一行对应于视图,而每一列指示时间位置。

[0071] 尽管MVC具有可由H.264/AVC解码器解码的所谓的基础视图,且MVC还可支持立体视图对,但MVC的优点在于其可支持使用两个以上视图作为3D视频输入且解码通过多个视

图表示的此3D视频的实例。具有MVC解码器的客户端的再现器可预期具有多个视图的3D视频内容。

[0072] 在每一行及每一列的交叉点处指示图3中的图片。H.264/AVC标准可使用术语帧来表示视频的一部分。本发明可互换地使用术语图片与帧。

[0073] 使用包含字母的块来说明图3中的图片，字母指明对应图片是经帧内译码(也就是说，I图片)，还是在一个方向上经帧间译码(也就是说，作为P图片)，或是在多个方向上经帧间译码(也就是说，作为B图片)。大体来说，预测由箭头指示，其中箭头指向的图片使用箭头出发的图片用于预测参考。举例来说，时间位置T0处的视图S2的P图片是从时间位置T0处的视图S0的I图片预测的。

[0074] 如同单视图视频编码，可相对于不同时间位置处的图片预测性地编码多视图视频译码视频序列的图片。举例来说，时间位置T1处的视图S0的b图片具有从时间位置T0处的视图S0的I图片指向其的箭头，从而指示所述b图片是从所述I图片预测的。然而，另外，在多视图视频编码的情况下，图片可经视图间预测。即，视图分量可使用其它视图中的视图分量用于参考。举例来说，在MVC中，如同另一视图中的视图分量为帧间预测参考般实现视图间预测。潜在视图间参考在序列参数集(SPS) MVC扩展中用信号通知且可通过参考图片列表构造过程加以修改，所述参考图片列表构造过程实现帧间预测或视图间预测参考的灵活排序。视图间预测也是包含3D-HEVC(多视图加上深度)的HEVC的所提出多视图扩展的特征。

[0075] 图3提供视图间预测的各种实例。在图3的实例中，视图S1的图片说明为是从视图S1的不同时间位置处的图片预测，以及是从相同时间位置处的视图S0和S2的图片经视图间预测。举例来说，时间位置T1处的视图S1的b图片是从时间位置T0及T2处的视图S1的B图片中的每一者以及时间位置T1处的视图S0及S2的b图片预测。

[0076] 在一些实例中，图3可被视为说明纹理视图分量。举例来说，图2中所说明的I、P、B和b图片可被视为视图中的每一者的纹理视图分量。根据本发明中描述的技术，对于图3中说明的纹理视图分量中的每一者，存在对应深度视图分量。在一些实例中，可以类似于图3中针对对应纹理视图分量所说明的方式的方式预测深度视图分量。

[0077] 也可由MVC支持两个视图的译码。MVC的优势中的一者为MVC编码器可将两个以上视图当作3D视频输入且MVC解码器可解码此多视图表示。因而，具有MVC解码器的任何再现器可解码具有两个以上视图的3D视频内容。

[0078] 如上文所论述，在MVC中，允许在相同存取单元中(在一些情况下意味着具有相同时间实例)的图片当中进行视图间预测。当译码非基础视图中的一者中的图片时，如果图片处于不同视图中但在相同时间实例内，那么可将所述图片添加到参考图片列表中。可将视图间预测参考图片放置在参考图片列表的任何位置中，正如任何帧间预测参考图片一般。如图3中所示，视图分量可出于参考目的使用其它视图中的视图分量。在MVC中，实现视图间预测，如同另一视图中的视图分量为帧间预测参考一般。

[0079] 在MVC中，允许在相同存取单元(即，具有相同时间实例)中的图片当中进行视图间预测。当对非基础视图中的一者中的图片进行译码时，如果图片处于不同视图中但具有相同时间实例，那么可将所述图片添加到参考图片列表中。视图间预测参考图片可放置在参考图像列表的任何位置中，正如任何帧间预测参考图片一般。

[0080] 如图3中所示，视图分量可出于参考目的使用其它视图中的视图分量。此情况被称

为视图间预测。在MVC中,如同另一视图中的视图分量为帧间预测参考般实现视图间预测。

[0081] 在多视图视频译码的上下文中,存在两个种类的运动向量。一个是指向时间参考图片的正常运动向量。对应的时间帧间预测是运动补偿预测(MCP)。另一类型的运动向量是指向不同视图中的图片(即,视图间参考图片)的视差运动向量。对应的帧间预测是视差补偿预测(DCP)。

[0082] 下一部分将论述基于AVC的3D视频译码标准。当前,VCEG和MPEG的3D视频译码联合合作小组(JCT-3V)正在开发基于H.264/AVC的3DV标准,即,3D-AVC。对于3D-AVC,已经包含且支持除MVC中的视图间预测以外的新译码工具。用于3D-AVC的最新软件3D-ATM可从以下链接下载:[3D-ATM版本6.2]:<http://mpeg3dv.research.nokia.com/svn/mpeg3dv/tags/3DV-ATMv6.2/>

[0083] 基于AVC的3D视频(3D-AVC)译码标准当前在由JCT-3V开发,且3D-AVC的最新版本现在公开可用:M·M·汉努克塞拉、Y·陈、T·铃木、J.-R.欧姆、G·J·萨利文,“3D-AVC草案文本5”,JCT3V-C1002,瑞士日内瓦,2013年1月。其从2014年3月14日起从以下链接可用且特此以引用的方式并入本文:http://phenix.it-sudparis.eu/jct2/doc_end_user/documents/3_Geneva/wg11/JCT3V-C1002-v3.zip。

[0084] 现将论述3D-AVC中的视图分量的译码次序。3D-AVC以基础视图的纹理部分可由H.264/AVC解码器完全解码的方式兼容于H.264/AVC。举例来说,基础视图的视图分量中的纹理视图分量可仅以同一基础视图中的其它纹理视图分量进行帧间预测。基础视图中的纹理视图分量不可进行视图间预测。并且,基础视图中的纹理视图分量可不需要对应深度视图分量用于解码目的。

[0085] 对于3D-AVC中的增强视图分量,在一些其它实例技术中,深度可在纹理之前经译码,且纹理视图分量可基于来自深度视图分量的信息而经译码,其也被称作深度优先译码。然而,例如在上述MVC+D中,每一纹理视图分量在纹理优先译码次序中在相应深度视图分量之前经译码。换句话说,在一些其它实例技术中,在3D-AVC中,首先对基础视图的纹理视图分量进行译码,接着是基础视图的相关联深度视图分量,接着是第一增强或相依视图的深度视图分量,接着是第一增强或相依视图的相关联纹理视图分量,接着是第二增强或相依视图的深度视图分量,接着是第二增强或相依视图的相关联纹理视图分量,等。

[0086] 举例来说,如下例示3D-AVC中的纹理和深度视图分量的译码次序。在以下实例中,T0和D0分别指代基础视图的纹理和深度视图分量,且Ti和Di分别指代第i相依视图的纹理和深度视图分量。在以下实例中,考虑三个视图。

[0087] 在第一实例中,考虑的视图是T0、D0、D1、D2、T1和T2。在此实例中,基础视图(T0和D0)是以纹理优先译码次序译码,而相依视图是以深度优先译码次序译码。当前在3D-AVC的共同测试条件中使用混合译码次序。在另一实例中,译码的次序是T0、D0、T1、D1、T2和D2。也就是说,所有视图分量是以纹理优先译码次序译码。

[0088] 如果针对Ti启用视图间预测,那么将参考纹理视图界定为包含视图间参考图片的视图且将对应深度视图界定为具有与参考纹理视图的视图次序索引相同的视图次序索引的参考深度视图。

[0089] 现将论述经由深度图的3D-AVC视差向量导出。用于导出视差向量的技术可随着每一低层级译码工具而变化,但通常,相依视图的深度数据用于视差向量导出以用于纹理视

图分量译码。这是因为由于深度优先译码次序,相依视图的深度视图是可用的。在3D-AVC中使用的低层级译码工具是环路内基于块的视图合成视图间预测(BVSP)和基于深度的运动向量预测(D-MVP)。视频译码器(例如,视频解码器30)可使用从相依视图(有时称为相依帧)中的深度视图(有时称为深度图)的深度值转换的视差向量。在3D-AVC参考软件中,通常,从实际深度图值到特定视图的视差的转换过程的结果与相机参数一起存储在查找表中。

[0090] 现将论述用于视差向量导出的从四个拐角的最大深度导出。为了导出深度值,视频解码器30首先识别深度视图分量的参考深度块。参考深度块与当前MB/分区/子块位于同一地点/对应。在所识别参考深度块中,视频解码器30存取对应于左上、右上、左下和右下深度样本的四个拐角样本。随后通过取所述四个拐角深度样本中的最大者来计算深度值。

[0091] 最后,视频解码器30使用计算的深度值且将视差向量的垂直分量设定为0来从查找表推断视差向量的水平分量。在此方法中,当MB划分成分区或子块时存取的深度样本的数目增加。举例来说,当16x16MB分割成四个8x8分区时,待存取的深度样本的数目是16,且当16x16MB分割成十六个4x4分区时,待存取的深度样本的数目是64。

[0092] 3D-AVC中的BVSP是由3D-AVC以及其它译码标准支持的译码模式。BVSP最初在W·苏等人的“3DV-CE1.a:用于3DV-ATM的基于块的视图合成预测”(JCT3V-A0107)中提出,以上文献可从以下链接下载且以引用的方式并入本文:http://phenix.it-sudparis.eu/jct2/doc_end_user/documents/1_Stockholm/wg11/JCT3V-A0107-v1.zip。

[0093] 图4是基于后向扭曲的BVSP的概念图。参见图4,假定利用以下译码次序,(T0,D0,D1,T1),其中T指代纹理视图且D指代深度视图。纹理分量T0是基础视图,且T1是以VSP译码的相依视图。深度图分量D0(图4中未图示)和D1是与T0和T1相关联的相应深度图。

[0094] 在相依视图T1中,当前经译码块Cb的样本值是使用VSP预测从由基础视图T0的样本值组成的参考区域R(Cb)预测。待译码的当前样本(即Cb)与参考样本(即R(Cb))之间的位移向量(Disp_vec)表示为从与当前经译码纹理样本相关联的深度图值导出的T1与T0之间的视差向量。

[0095] 从深度值到视差向量的转换的过程可例如用以下等式执行:

$$Z(Cb(j,i)) = \frac{1}{\frac{d(Cb(j,i))}{255} \cdot \left(\frac{1}{Z_{近}} - \frac{1}{Z_{远}} \right) + \frac{1}{Z_{远}}}}; \quad (1)$$

$$D(Cb(j,i)) = \frac{f \cdot b}{Z(Cb(j,i))}; \quad (2)$$

[0098] 其中j和i是Cb内的局部空间坐标,dCb(j,i)是视图1的深度图图像中的深度图值,Z是实际对应深度值,且D是到特定视图0的所导出视差向量的水平分量。参数f、b、Z近和Z远是指定相机设置的参数;即,使用的焦距(f)、视图#1与视图#0之间的相机分离(b),以及表示深度图转换的参数的深度范围(Z近,Z远)。

[0099] 应注意,在一些实例中,将所导出的视差向量的垂直分量设定成0。并且,在一些3DV-ATM实施方案中,等式(1)和(2)已经针对每个深度图值(0...255)预先计算且存储为查找表。

[0100] 在图3的实例中,Cb表示当前正译码的块。视频解码器30可使用BVSP模式对Cb进行

译码。如果将使用BVSP模式对Cb进行译码,那么视频解码器30识别对应于Cb的深度块,图4中的d(Cb)。在此方面,对应意味着位于同一地点。深度块d(Cb)包含多个深度值。基于d(Cb)的那些深度值中的一或多者,视频解码器30确定视差值。所述视差值可用作视差向量的x分量,其中y分量设定成0。使用视差向量,视频解码器30可识别不同视图中的参考块,在图4的实例中为R(cb),且基于所述参考块帧间预测块Cb。

[0101] 深度视图分辨率和纹理视图分辨率可为相同的,或它们可为不同的。如果深度视图分辨率和纹理视图分辨率是不同的,那么视频解码器30可执行转换以找到位于同一地点的深度块。视频解码器30可如下执行转换。假设(x,y)表示块Cb的左上位置。深度视图中的位于同一地点块的左上拐角位置由(x>>reduced_resolution_flag,y>>reduced_resolution_flag)表示。等于1的语法元素“reduced_resolution_flag”指定视图分量对的深度视图分量具有比同一视图分量对的纹理视图分量的明度分量低的空间分辨率,且深度视图分量的宽度和高度均为所有纹理视图分量的宽度和高度的二分之一。

[0102] 本发明的下一部分将论述BVSP的若干实施方案问题。一个问题涉及BVSP块的指示,即将使用BVSP模式译码何种块。如下指示BVSP块:

[0103] -使用MB层级的一个旗标来用信号表示当前MB是否以常规跳过/直接模式译码或其是否以跳过/直接模式译码但从合成参考分量预测。在此上下文中,“常规跳过/直接模式”指代在H.264/AVC译码标准中使用的跳过/直接模式的扩展版本,且合成参考分量指代从视图间块产生的参考块。

[0104] -对于每一MB分区(从16x16到8x8),使用每一参考图片列表中的参考索引(或旗标,如用于3D-AVC的一些提议中)来用信号表示参考图片。当在BVSP模式中对分区进行译码时,不用信号表示运动向量差,因为不存在BVSP经译码块的运动向量。

[0105] 当旗标或参考索引指示合成参考分量时,调用如以下项目中描述的一个分区的预测。换句话说,位可指示MB或MB分区是否使用常规跳过/直接模式或BVSP模式经译码。在BVSP模式中,视频解码器30通过如下方式对使用BVSP模式(读者称为合成跳过/直接模式)译码的MB或MB分区的部分进行解码:将MB或MB分区划分为KxK块,且对于每一KxK块,识别位于同一地点的深度块,从深度块获得视差值,且从由视差向量指向的视图间参考图片预测MB或MB分区。

[0106] 在3D-AVC中,对于常规跳过/直接模式,视频解码器30如下导出运动向量和参考索引。参考视图中(由视差向量指向的)对应块的运动向量(如果可用)设定成等于当前块的运动向量,且因此导出当前块的参考索引。如果视图间运动向量不可用,即,由视差向量指向的基础视图中的参考块经帧内译码,那么通过将参考索引设定为零而使用常规基于中值的运动向量预测方案。

[0107] BVSP的另一实施方案问题涉及预测导出过程。对于其大小由NxM指示(其中N或M将为8或16)的每一MB分区,如果MB分区以BVSP模式译码,那么将当前MB分区进一步分割成具有等于KxK的大小的若干子区(其中K可为8x8(如3D-AVC的一些提议中)、4x4、2x2或1x1)。对于每一子区,导出单独的视差向量且从通过视图间参考图片中的所导出的视差向量定位的一个块(例如,图4中的R(cb))预测每一子区。在一些实例性共同测试条件中,K经界定为4。应注意,针对BVSP经译码块不存储导出的视差向量,因为不存在使用此些向量的译码工具。

[0108] BVSP的另一实施方案问题涉及视差向量导出过程。当应用深度优先译码次序时,

通过转换来自对应非基础深度视图中的对应深度块的深度值而导出视差向量,如图4的实例中示出。通过首先从与KxK子区位于同一地点/对应的深度块存取四个拐角深度样本且随后取所述四个存取的深度样本的最大值而计算单个深度值。稍后使用等式(1)和(2)将计算的深度值转换为视差向量。当应用纹理优先译码次序时,停用BVSP模式,因为当解码非基础纹理视图时对应非基础深度视图不可用。

[0109] 现将论述用于正常帧间预测模式的3D-AVC中的基于深度的运动向量预测(D-MVP)。D-MVP指代使用由于深度优先译码次序而可用的当前视图的相关联深度图数据的运动向量预测方法。所述方法是以相依视图中的纹理视图分量而应用。

[0110] 在3D-AVC中,运动向量预测仍利用相邻块。相邻块按次序包含当前块的左边块、上方块、右上方块和左上方块。仅当其它三个相邻块中的一者不含运动向量因此视为不可用时使用左上方块中的运动向量。

[0111] 来自相邻块的运动向量在其具有与待预测的当前运动向量不同的类型的情况下视为不可用。运动向量的类型取决于对应参考索引。即,如果参考索引对应于视图间参考图片,那么运动向量是视差运动向量且类型是“视差”,且如果参考索引对应于时间参考图片(同一视图中),那么运动向量是时间运动向量且类型是“时间”。

[0112] 在3D-AVC中,如果三个相邻块可用,那么采用所述三个相邻块中的运动向量用于当前块的运动向量预测。在时间预测中,如果其运动向量全部具有相同类型且具有相同参考索引,那么如H.264/AVC中直接使用中值滤波器,否则(如果它们属于不同类型且具有不同参考索引),进一步导出运动向量。在当前参考图片是视图间参考图片时,检查运动向量类型及其在相邻块位置中的参考索引,并且,如果它们全部具有相同类型和相同参考索引,那么应用中值滤波器。在两个情况中,如果少于三个相邻块可用,那么进一步导出不可用的块的运动向量以使得三个相邻块变为可用的。

[0113] 如果空间相邻块不含可用时间运动向量,那么将针对当前块预测时间运动向量。通过如上文相对于基于块的视图合成预测所描述而导出的视差向量来识别参考视图图片中的当前块的参考块。针对当前空间相邻块导出含有参考块的中心位置的块的运动向量(如果其是时间运动向量)。如果时间运动向量被视为不可用(帧内块或不指向与当前参考图片对准的参考视图中的参考图片),那么将所导出的运动向量设定成零。

[0114] 如果空间相邻块不含可用视差运动向量且针对当前块将预测视差运动向量,那么如上文相对于基于块的视图合成预测所描述将针对当前块导出的视差向量转换成视差运动向量。

[0115] 在3D-AVC中,D-MVP方法并入到H.264/AVC中的基于常规中值函数的运动向量预测中。因此当空间相邻块中的运动向量可用时,(或最初不可用但通过上文所提及的方法而可用),中值函数仍可应用于三个运动向量,然而其全部应属于同一类型。

[0116] 现将论述用于跳过和直接模式的3D-AVC中的视图间运动预测。3D-AVC中的视图间运动预测是在P-跳过、B-跳过、B-16x16直接和B-8x8直接模式中执行。在这些模式中,用以识别参考块的视差向量是来自相邻块或来自从与当前MB相关联的对应深度块导出的视差向量(见关于基于块的视图合成预测的上文)。针对其可用性检查候选相邻块A、B和C看其是否具有视差运动向量。如果一个空间相邻块可用,即且其含有视差运动向量,那么此视差运动向量变成视差向量。

[0117] 如果空间相邻块A、B和C的视差向量不可用,那么从与当前MB相关联的深度块导出的视差向量(见关于基于块的视图合成预测的上文)用于不可用块。然后,应用中值滤波器以得到视差向量。

[0118] 从以上过程得到的此视差向量用以得到参考视图图片中的参考块。在参考块内,将运动向量(即视图间运动向量)(如果可用)设定成等于当前块的运动向量且因此导出当前块的参考索引。

[0119] 如果视图间运动向量不可用,即,由视差向量指向的基础视图中的参考块经帧内译码,那么使用常规的基于中值的运动向量预测方案。在此情况下,首先导出参考索引,且使用章节1.4.5中描述的D-MVP方案导出当前MB的运动向量预测。

[0120] 现将论述用于视差向量导出的改进。在JCT3V-C0122中,已提出简化视差向量导出方法。在所提出的方法中,在当前MB是以并非跳过或直接的帧间预测模式译码时,具有当前MB的所有分区/子块共享从当前MB的相关联深度块的右下深度样本计算的所导出的视差向量。然而,在当前MB是以跳过或直接模式译码时,利用不同视差向量导出过程,即,存取参考深度块的四个拐角样本。此外,在BVSP模式中,仍需要存取MB分区的每一KxK子区的参考深度块的四个拐角样本。

[0121] 在JCT3V-C0134中,提出当以帧间模式译码时,同一MB内的所有分区块共享来自与当前MB位于同一地点/对应的同一参考深度块的四个拐角样本的最大深度值导出单个视差向量。然而,对于BVSP模式,仍需要存取MB分区的每一KxK子区的参考深度块的四个拐角样本。

[0122] 本发明现将论述3D-HEVC的方面。基于相邻块的视差向量导出(NBDV)可用作3D-HEVC中的视差向量导出方法,其针对所有视图使用纹理优先译码次序。在当前3D-HEVC设计中,通过从参考视图的深度图检索深度数据可进一步精炼从NBDV导出的视差向量。在NBDV中,视差向量(DV)用作两个视图之间的位移的估计器。因为相邻块在视频译码中几乎共享相同运动/视差信息,所以当前块可使用相邻块中的运动向量信息作为良好预测符。遵循此想法,NBDV使用相邻视差信息用于估计不同视图中的视差向量。

[0123] 首先界定若干空间和时间相邻块。随后以通过当前块与候选块之间的相关的优先级确定的预定义次序检查相邻块中的每一者。一旦在候选者中找到视差运动向量(即,指向视图间参考图片的运动向量),便将所述视差运动向量转换成视差向量。利用相邻块的两个集合。一个集合是来自空间相邻块且另一集合是来自时间相邻块。

[0124] 3D-HEVC首先采用JCT3V-A0097,3D-CE5.h:视差向量产生结果(L·张,Y·陈,M·卡塞维茨(高通))中提出的(基于)相邻块视差向量(NBDV)方法。JCTVC-A0126,3D-CE5.h:基于HEVC的3D视频译码的视差向量导出的简化(J·孙,M·古,S·叶(LG))中与简化NBDV一起包含隐式视差向量。基于此,在JCT3V-B0047,3D-CE5.h相关:视差向量导出的改进(J·康,Y·陈,L·张,M·卡塞维茨(高通))中,通过移除存储在经解码图片缓冲器中的隐式视差向量进一步简化NBDV,而且以RAP图片选择改善译码增益。

[0125] 在当前(从本发明的时间起)NBDV中,五个空间相邻块用于视差向量导出。它们是当前块(例如,当前预测单元(PU))的左下方、左边、右上方、上方和左上方块,如图5中由A0、A1、B0、B1或B2指示。应注意这些相邻块与在HEVC中的合并模式中使用的相邻块相同。因此,不需要额外存储器存取。

[0126] 为了检查时间相邻块,视频译码器可首先执行候选图片列表的构造过程。在一个实例中,来自当前视图的多达两个参考图片可视为候选图片。视频译码器可首先将位于同一地点的参考图片插入到候选图片列表中,接着按参考索引的升序插入候选图片的其余部分。当两个参考图片列表中具有相同参考索引的参考图片可用时,位于同一地点的图片的同一参考图片列表中的一个参考图片可先于另一参考图片。对于候选图片列表中的每一候选图片,可确定三个候选区以用于导出时间相邻块。

[0127] 当块是以视图间运动预测译码的时,可导出视差向量以用于选择不同视图中的对应块。隐式视差向量(IDV,也被称作所导出的视差向量)称为在视图间运动预测中导出的视差向量。即使块是以运动预测译码的,为了对随后的块进行译码的目的也不会丢弃所导出的视差向量。

[0128] 在3D-HTM 6.0的当前设计中,视频译码器根据NBDV过程按次序检查时间相邻块中的视差运动向量、空间相邻块中的视差运动向量以及随后的IDV。一旦找到视差运动向量或IDV,过程便终止。

[0129] 视频译码器可通过存取深度信息而进一步精炼使用NBDV导出的视差向量。当从NBDV过程导出一个视差向量时,视频译码器可通过从参考视图的深度图检索深度数据而进一步精炼视差向量。精炼过程可包含两个步骤:

[0130] a) 在例如基础视图等经先前译码参考深度视图中通过所导出的视差向量定位对应深度块;对应深度块的大小与当前PU的大小相同。

[0131] b) 从对应深度块的四个拐角像素选择一个深度值且将其转换为经精炼视差向量的水平分量。视差向量的垂直分量不变。

[0132] 视频译码器可使用经精炼视差向量用于视图间运动预测,同时未精炼的视差向量用于视图间残余预测。另外,经精炼视差向量可存储为一个PU的运动向量(如果所述PU是以后向VSP模式译码)。

[0133] 根据本发明的技术,空间相邻块中的一者可对应于经BVSP译码块,且空间相邻块中的另一者可对应于非经BVSP译码块。举例来说,块A1可对应于经BVSP译码块且块B1可对应于非经BVSP译码块。然而,当对当前块的运动信息进行译码时,视频译码器可使用同一逻辑功能存取块A1和块B1的运动信息。假定经BVSP译码块(即,上文陈述的实例中的块A1)的运动信息包含识别参考图片的参考索引。因此,视频译码器中不需要提供用于存取块A1的运动信息的单独逻辑功能。

[0134] 图6是说明涉及使用相邻块的后向视图合成预测(BVSP)的技术的概念图。已提出并采用BVSP作为用于3D-HEVC的技术。如JCT3V-C0152中提出的后向扭曲VSP方法在第3次JCT-3V会议中采用。此后向扭曲VSP的基本想法与3D-AVC中的基于块的VSP相同。这两个技术均使用后向扭曲和基于块的VSP以避免发射运动向量差且使用更精确的运动向量。实施方案细节由于不同平台而不同。本发明大体上使用首字母缩略词BVSP来指代3D-HEVC中的后向视图合成预测,但BVSP也可指代3D-AVC的基于块的视图合成预测。

[0135] 在3D-HTM中,在共同测试条件中应用纹理优先译码。因此,当对一个非基础纹理视图进行解码时对应非基础深度视图不可用。因此,估计深度信息且用以执行BVSP。为了估计块的深度信息,提出首先从相邻块导出视差向量,且随后使用所导出的视差向量从参考视图获得深度块。

[0136] 在HTM 5.1测试模型中,存在导出视差向量预测符的过程,称为NBDV(相邻块视差向量)。假设 (dv_x, dv_y) 表示从NBDV功能识别的视差向量,且当前块位置是(块 x ,块 y)。提出获取参考视图的深度图像中的(块 $x+dv_x$,块 $y+dv_y$)处的深度块。所获取的深度块将具有当前预测单元(PU)的相同大小,且其随后将用以进行当前PU的后向扭曲。图6说明用于从参考视图定位深度块且随后使用所述深度块用于BVSP预测的步骤。

[0137] 在图6的实例中,深度图片150和纹理图片154对应于同一视图,而纹理图片152对应于不同视图。确切地说,纹理图片152包含相对于充当参考图片的纹理图片154正译码的当前块160。视频译码器可参考与当前块160相邻的相邻块162。相邻块162包含先前所确定的视差向量166。视差向量166可如同当前块160的视差向量164那样导出。因此,视差向量164参考参考视图的深度图片150中的深度块156。

[0138] 视频译码器可随后使用深度块156的像素(即,深度值)来确定当前块160的像素的视差值168(即,纹理值),用于执行后向扭曲。视频译码器可随后从由视差值168识别的像素合成当前块160的经预测块(即,BVSP参考块)的值。视频译码器可随后使用此经预测块预测当前块160。举例来说,在视频编码器20的视频编码期间,视频编码器20可计算经预测块与当前块160之间的逐像素差以产生残余值,视频编码器20可随后对所述残余值进行变换、量化和熵编码。另一方面,在视频解码器30的视频解码期间,视频解码器30可对残余数据进行熵解码、逆量化和逆变换,随后组合所述残余数据(基于逐像素)与经预测块以再生当前块160。

[0139] JCT3V-C0152提出对3D-HEVC的BVSP技术的改变,如下所述。确切地说,斜体文字表示添加到3D-HEVC的文字,而前面是“移除”的带括号文字表示从3D-HEVC的删除:

[0140] 如果在序列中启用BVSP,那么用于视图间运动预测的NBDV过程改变且在以下段落中突出显示差异:

[0141] ●对于时间相邻块中的每一者,如果其使用视差运动向量,那么返回视差运动向量作为视差向量且以3D-HEVC的章节1.6.1.3中描述的方法进一步精炼所述视差向量。

[0142] ●对于空间相邻块中的每一者,以下适用:

[0143] ○按次序对于参考图片列表0和参考图片列表1:

[0144] ■如果其使用视差运动向量,那么返回视差运动向量作为视差向量且以章节1.6.1.3中描述的方法进一步精炼所述视差向量。

[0145] ■否则,如果其使用BVSP模式,那么返回相关联运动向量作为视差向量。以与章节1.6.1.3中描述类似的方式进一步精炼所述视差向量。然而,最大深度值是选自对应深度块的所有像素而不是四个拐角像素,且经精炼视差向量的垂直分量设定成0。

[0146] ●对于空间相邻块中的每一者,如果其使用IDV,那么返回IDV作为视差向量且以章节1.6.1.3中描述的方法进一步精炼所述视差向量。

[0147] 所介绍的BVSP模式视为特殊帧间译码模式,且针对每一PU可维持指示BVSP模式的使用的旗标。将新合并候选者(BVSP合并候选者)添加到合并候选者列表,而不是在位流中用信号表示所述旗标。所述旗标取决于经解码合并候选者索引是否对应于BVSP合并候选者。如下按JCT3V-C0152定义BVSP合并候选者:

[0148] ●每一参考图片列表的参考图片索引:-1

[0149] ●每一参考图片列表的运动向量:经精炼视差向量

[0150] 在JCT3V-C0152中,BVSP合并候选者的插入位置取决于空间相邻块,如下文所述:

[0151] ●如果五个空间相邻块(图5中所示的A0、A1、B0、B1或B2)中的任一者是以BVSP模式译码,即,维持相邻块的旗标等于1,那么将BVSP合并候选者视为对应空间合并候选者且插入到合并候选者列表。BVSP合并候选者将仅插入到合并候选者列表一次。

[0152] ●否则(五个空间相邻块均不是以BVSP模式译码),将BVSP合并候选者插入到合并候选者列表仅处于时间合并候选者之前。

[0153] 应注意,在组合的双向预测合并候选者导出过程期间,可检查额外条件以避免包含BVSP合并候选者。

[0154] JCT3V-J0152进一步提供,具有由 $N \times M$ 表示的大小的每一经BVSP译码的PU进一步分割成具有等于 $K \times K$ (其中 K 可为4或2)的大小的若干子区。对于每一子区,导出单独的视差运动向量,且从通过视图间参考图片中的导出的视差运动向量定位的一个块预测每一子区。换句话说,用于经BVSP译码PU的运动补偿单元的大小设定成 $K \times K$ 。在共同测试条件中, K 设定成4。

[0155] JCT3V-J0152进一步提供,对于以BVSP模式译码的一个PU内的每一子区(4×4 块),首先在具有上文提到的经精炼视差向量的参考深度视图中定位对应 4×4 深度块。第二,选择对应深度块中的十六个深度像素中的最大值。第三,将所述最大值转换成视差运动向量的水平分量。将视差运动向量的垂直分量设定成0。

[0156] 在3D-HEVC中,当应用纹理优先译码次序时,对于每一预测单元(PU),可在考虑/不考虑参考深度视图中的深度值的情况下从NBDV导出视差向量。在获得视差向量之后,针对一个PU的每一 4×4 子区将其进一步精炼(如果所述PU是以BVSP模式译码)。

[0157] 3D-HEVC将精炼过程描述为包含两个步骤:1)从通过所导出的视差向量定位的参考深度视图中的 4×4 深度块选择一个最大深度值;2)将所述深度值转换为经精炼视差向量的水平分量,同时保持经精炼视差向量的垂直分量为0。在针对一个PU的一个 4×4 子区精炼视差向量之后,3D-HEVC提供使用经精炼视差向量来定位参考纹理视图中的一个块用于运动补偿。

[0158] 现将论述3D-AVC中的NBDV。根据一些3D译码技术,MB层级NBDV可用以导出当前MB的视差向量且进一步用于运动向量预测。一旦识别出视差运动向量,即,时间或空间相邻块中的一者使用视图间参考图片,那么返回其作为当前MB的视差向量。

[0159] 根据一些3D译码技术,可通过存取对应深度块而进一步精炼从NBDV导出的视差向量。举例来说,可使用(通过视差向量识别的)参考视图的深度块的四个拐角的深度值,且选择最大深度值且转换成视差向量。根据其它3D译码技术,以针对每一 4×4 或 8×8 块使用对应深度块产生视差运动向量的方式利用BVSP。类似于视差向量导出过程,可使用(通过经精炼视差向量识别的)深度块的四个拐角,且选择最大深度值以转换成视差运动向量。

[0160] 本发明的技术可解决若干潜在问题。作为一个实例,对于帧间模式中的D-MVP,需要多次执行深度到DV转换以导出一个MB中的各种分区块的DV,其在当前MB划分成分区/子块时增加将从存储器存取的深度样本的数目。作为另一实例,对于跳过/直接模式中的D-MVP,通过取相关联深度块的四个拐角样本的最大深度值而针对MB导出视差向量;这要求高存储器存取带宽。作为另一实例,在BVSP模式中,通过取当前子区的相关联深度块的四个拐

角深度样本的最大值而针对大小 $K \times K$ 的子区导出视差向量;这是以高存储器存取带宽为代价完成的。作为另一实例,对于跳过/直接MB,当视图间运动向量不可用时,使用常规基于中值的运动向量预测。作为另一实例,在3D-HEVC中,当视差向量是从进行视图间预测相邻块导出时精炼当前块的视差向量仅要求存取参考深度块的四个拐角。当视差向量是从经BVSP译码相邻块导出时精炼当前块的视差向量且针对以BVSP模式译码的PU产生 4×4 或 8×8 子区的视差运动向量需要稍微不同的设计,其要求存取参考块的所有样本。

[0161] 本发明介绍了当视差向量是通过存取深度图而导出时可简化视差(运动)向量导出过程的技术。这些技术的阐释可主要集中于3D-AVC,但应理解相似想法可适用于3D-HEVC。更具体来说,本发明提供当利用深度到视差转换方案时用于各种帧间预测模式(包含跳过/直接,帧间)以及BVSP模式的简化且通用的视差向量导出方案。可简化计算和存储器存取。

[0162] 图7展示 8×8 深度块170的实例。深度块170的四个拐角样本标记为172A-172D。位流中的“reduced_resolution_flag”可指示深度视图是否具有相对于纹理视图减少的分辨率或深度视图是否具有与纹理视图相同的分辨率。如果reduced_resolution_flag等于1,那么纹理视图中的 16×16 宏块将具有对应纹理视图中的对应(即,位于同一地点) 8×8 深度块,例如 8×8 深度块170。如果reduced_resolution_flag等于0,那么纹理视图中的 16×16 宏块将具有对应纹理视图中的对应(即,位于同一地点) 16×16 深度块。下文将通过参考深度块170来阐释若干实例。

[0163] 根据本发明的一个技术,当对非基础纹理视图分量进行译码时,视频解码器30可通过存取对应于含有当前纹理块的宏块的深度块而导出整个宏块的仅一个视差向量。不管当前MB是否是以跳过/直接模式或其它帧间模式译码,每当针对MB内的任何块需要视差向量时,视频解码器30可从同一MB层级视差向量导出导出相同的唯一视差向量一次。视频解码器30可(例如)通过存取拐角样本172A-172D的深度值且从所述四个拐角样本确定最大深度值而导出视差向量。视频解码器30可随后使用(例如)查找表或以上等式(1)和(2)将最大深度值转换为视差值。在一些实施方案中,视频解码器30可通过存取除所述四个拐角样本外的深度值和/或识别除最大值外的值而导出视差向量。

[0164] 作为一个实例,对于非基础纹理视图的宏块,视频解码器30可定位对应于所述宏块的深度块,且基于所述深度块的至少一个深度值,导出宏块的视差向量。视频解码器30可基于所导出的视差向量对宏块的第一子块进行译码且基于所导出的视差向量对宏块的第二子块进行译码。深度块和宏块可位于同一地点。视频解码器30可通过确定包含深度块的两个或两个以上拐角样本的深度值的深度值的集合而导出宏块的视差向量,且从深度值集合识别最大深度值。视频解码器30可(例如)使用跳过模式和直接模式中的一者对第一子块进行译码且使用除跳过模式或直接模式外的帧间预测模式对第二子块进行译码。

[0165] 跳过/直接模式与帧间预测之间的一个潜在差异是随后当使用跳过/直接模式对宏块进行译码时,将不存在任何分割,即, 16×16 的MB大小经完全译码。当MB是使用除跳过/直接外的模式(即,帧间模式)译码时,可存在MB分割。并且,每一MB分区可进一步分割成子块(即H.264/AVC子块)。根据本发明的技术,针对MB导出的同一视差向量可以用于所有模式,例如跳过/直接和帧间模式。

[0166] 根据本发明的另一技术,当从同一视图内的深度视图分量的对应深度块导出视差

向量时,仅存取当前MB的对应深度块的左下(172A)和右下(172B)拐角深度样本。如图7中示出,通常MB对应于3D-AVC中的同一视图中的8x8参考块。深度块通常具有较低空间分辨率(水平和垂直向下取样二分之一)。因此,16x16宏块对应于8x8深度块。根据本发明的一个技术,仅存取拐角样本172A和172B以导出整个MB的视差向量。此外,存取的深度样本的最大深度值用于将其转换为视差向量。或者,仅存取参考深度块的左上(172C)和右下(172B)拐角样本。或者,仅存取参考深度块的左上(172C)和右上(172D)拐角样本。或者,仅存取参考深度块的左下(172A)和右下(172B)拐角样本。或者,仅存取参考深度块的左下(172A)和右上(172D)拐角样本。或者,仅存取参考深度块的右下(172B)和右上(172D)拐角样本。或者,存取位于对应深度块内的任何其它两个样本,例如一个中心像素和一个拐角像素或两个中心像素等。或者,当存取两个样本时,可使用这两个深度样本的平均/最小值而不是最大值。

[0167] 根据本发明的另一技术,可在BVSP模式中使用相似的深度到视差转换方法以便针对MB分区内的大小KxK的每一子区导出视差向量。首先针对KxK子区识别参考深度块,且存取参考深度块的具有相对坐标中的位置的相同拐角样本(例如,与KxK子区对应/位于同一地点的参考深度块的右下和左下拐角),类似于MB层级视差向量导出中的那些样本。或者,针对BVSP可存取参考深度块的较少拐角样本。

[0168] 根据本发明的另一技术,在当前3D-AVC的跳过/直接模式中,当视图间运动向量不可用时,将运动向量预测符设定成第一可用空间相邻者的运动向量(包含参考索引),而不是使用三个相邻块的运动向量的中值。按次序检查空间相邻者中的任一者是否具有大于或相等0的参考图片索引。如果为真,那么将当前MB的运动向量和参考图片索引分别设定为等于空间相邻块的运动向量和参考索引,且检查过程终止。

[0169] 根据本发明的另一技术,当NBDV用于3D-AVC且参考视图的深度图用以精炼视差向量时,相同的两个拐角样本可用以得到一个最佳深度值以转换成经精炼视差向量。

[0170] 根据本发明的另一技术,当NBDV用于3D-AVC且参考视图的深度图用以在BVSP模式中产生视差运动向量时,相同的两个拐角样本可用以得到一个最佳深度值以转换成视差运动向量。

[0171] 根据本发明的另一技术,在3D-HEVC中,类似地,可以仅检查相邻块的拐角样本的方式对准于在BVSP模式中存取视差运动向量的参考块。换句话说,对于每一KxK子区,视频解码器30仅检查通过含有KxK子区的预测单元的视差向量识别的KxK子区的参考深度块的拐角样本。此外,通过将精炼应用于PU或CU而不是MB块,如上文所描述的简化可适用于3D-HEVC。

[0172] 现将更详细地阐释实施上述技术的方面,从跳过/直接模式中的视差向量导出开始。对于此实例,当前MB的左上样本相对于当前图片的位置由(x,y)指示。从参考深度视图中的左下和右下拐角像素选择用于当前MB的一个深度值(D)为

[0173] $D = \max(D_0, D_1)$,

[0174] 其中函数 $\max(\cdot)$ 返回 D_i (i为0到1)的最大值且 D_i 表示在第i像素位置的深度值,其被计算为

[0175] $i = 0: (x \gg \text{reduced_resolution_flag}, (y \gg \text{reduced_resolution_flag}) + \text{SMD_POS})$,

[0176] $i = 1: ((x \gg \text{reduced_resolution_flag}) + \text{SMD_POS}, (y \gg \text{reduced_resolution_flag}) + \text{SMD_POS})$

flag)+SMD_POS),

[0177] 此处,当“reduced_resolution_flag”分别等于1和0时,SMD_POS等于7和15。等于1的“reduced_resolution_flag”指定视图分量对的深度视图分量具有比同一视图分量对的纹理视图分量的明度分量低的空间分辨率,且深度视图分量的宽度和高度均为所有纹理视图分量的宽度和高度的二分之一。等于0的“reduced_resolution_flag”指定当深度视图分量和纹理视图分量两者存在时,它们具有相同空间分辨率。最后,使用等式(1)和(2)从所述选定深度值D计算视差向量的水平分量且视差向量的垂直分量始终设定成0。

[0178] 现将描述帧间模式中的视差向量导出。在此实例中,相对于当前图片的当前MB(当前分区/子块位于该处)的左上位置由(x,y)表示。从参考深度视图中的左下和右下拐角像素选择用于当前MB的一个深度值(D)为

[0179] $D = \max(D_0, D_1)$,

[0180] 其中函数 $\max(\cdot)$ 返回 D_i (i为0到1)的最大值且 D_i 表示在第i像素位置的深度值,其被计算为

[0181] $i=0: (x \gg \text{reduced_resolution_flag}, (y \gg \text{reduced_resolution_flag}) + \text{SMD_POS})$,

[0182] $i=1: ((x \gg \text{reduced_resolution_flag}) + \text{SMD_POS}, (y \gg \text{reduced_resolution_flag}) + \text{SMD_POS})$,

[0183] 此处,SMD_POS和reduced_resolution_flag如上所述起作用。最后,使用等式(1)和(2)或查找表从所述选定深度值D计算视差向量的水平分量,且视差向量的垂直分量可设定成0。

[0184] 现将描述BVSP模式中的视差向量导出。由 $K \times K$ (其中K可为8或4)的子区的大小,相对于当前图片的当前MB分区内的一个子区的左上位置由(x,y)表示。从参考深度视图中的左下和右下拐角像素针对大小 $K \times K$ 的每一子区选择一个深度值(D)如

[0185] $D = \max(D_0, D_1)$,

[0186] 其中函数 $\max(\cdot)$ 返回 D_i (i为0到1)的最大值且 D_i 表示在第i深度像素位置:

[0187] $i=0: ((x \gg \text{reduced_resolution_flag}), (y \gg \text{reduced_resolution_flag}) + \text{VSP_S})$

[0188] $i=1: ((x \gg \text{reduced_resolution_flag}) + \text{VSP_S}, (y \gg \text{reduced_resolution_flag}) + \text{VSP_S})$

[0189] 此处,VSP_S计算为:

[0190] $\text{VSP_S} = (K \gg \text{reduced_resolution_flag}) - 1$,

[0191] 且reduced_resolution_flag如上文所定义起作用。最后,可使用等式(1)和(2)或查找表从所述选定深度值D计算视差向量的水平分量,且视差向量的垂直分量可设定成0。

[0192] 现将论述当视图间运动不可用时跳过/直接模式中的MVP。在此实例中,对当前MB的空间相邻者A、B和C分别表示运动向量 $Mvc_X[i]$ 和参考索引 $Ric_X[i]$ (X等于0和1),其中i等于0、1和2。当在帧内模式中对对应于当前跳过/直接MB的基础视图中的参考块进行编码时,如下计算当前MB的断言的运动向量($Mvp0, Mvp1$)和参考索引($Ri0, Ri1$):

[0193] 1.将 $Mvp0, Mvp1$ 两者初始化为等于零运动向量[0,0]且参考索引 $Ri0, Ri1$ 设定成等于-1。

[0194] 2.当所有空间候选者相邻者具有等于-1的参考索引(即,对于所有 $i=0,1,2$, $Ric_0[i]=-1$ 且 $Ric_1[i]=-1$)时,则将当前MB的运动向量和参考索引设定成零。

[0195] 3.对于 i 为0到2,以下适用:

[0196] ○如果第 i ($i=0,1,2$) 空间相邻者具有大于或等于零的参考索引,那么将预测运动向量Mvp0和Mvp1分别设定为等于运动向量MvC_0[i]和MvC_1[i]。此外,将参考索引Ri0和Ri1分别设定成等于参考索引Ric_0[i]和Ric_1[i]。

[0197] 图8是说明可实施本发明中描述的技术的视频编码器的实例的框图。举例来说,图8说明可表示3D-AVC顺应或3D-HEVC顺应视频编码器的视频编码器20。将使用例如PU、TU和CU等某些HEVC术语描述视频编码器20,但应理解参考视频编码器20描述的技术也可以关于根据H.264标准译码的视频而执行。

[0198] 视频编码器20可执行视频切片内的视频块的帧内及帧间译码。举例来说,视频编码器20可执行帧间预测编码或帧内预测编码。帧内译码依赖于空间预测来减少或去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或视图间预测来减少或移除视频序列的邻近帧或图片内的时间冗余或不同视图中的图片之间的冗余。帧内模式(I模式)可以指若干基于空间压缩模式中的任一者。例如单向预测(P模式)或双向预测(B模式)等帧间模式可指代若干基于时间的压缩模式中的任一者。

[0199] 在图8的实例中,视频编码器20包含视频数据存储器40、预测处理单元42、参考图片存储器64、求和器50、变换处理单元52、量化处理单元54和熵编码单元56。预测处理单元42包含运动和视差估计单元44、运动和视差补偿单元46以及帧内预测单元48。为了视频块重构,视频编码器20还包含逆量化处理单元58、逆变换处理单元60,及求和器62。还可包含解块滤波器(图8中未图示)以便对块边界进行滤波,以从经重构视频移除成块效应假象。必要时,解块滤波器将通常对求和器62的输出进行滤波。除了解块滤波器之外,还可使用额外环路过滤器(环路内或环路后)。

[0200] 视频数据存储器40可存储待由视频编码器20的组件编码的视频数据。可(例如)从视频源18获得存储在视频数据存储器40中的视频数据。参考图片存储器64是解码图片缓冲器(DPB的一个实例,其存储参考视频数据供视频编码器20用于编码视频数据(例如,在帧内或帧间译码模式中,也被称作帧内或帧间预测译码模式)。视频数据存储器40和参考图片存储器64可由多种存储器装置中的任一者形成,例如包含同步DRAM(SDRAM)的动态随机存取存储器(DRAM)、磁阻式RAM(MRAM)、电阻式RAM(RRAM)或其它类型的存储器装置。视频数据存储器40和参考图片存储器64可由同一存储器装置或单独的存储器装置提供。在各种实例中,视频数据存储器40可与视频编码器20的其它组件一起在芯片上,或相对于那些组件在芯片外。

[0201] 视频编码器20接收视频数据,且分割单元(未图示)将所述数据分割为视频块。此分割还可包含分割为切片、瓦片或其它较大单元,以及视频块分割(例如,宏块分区和分区的子块)。视频编码器20总体上说明对待编码视频切片内的视频块进行编码的组件。所述切片可以分成多个视频块(并且可能分成被称作瓦片的视频块集合)。预测处理单元42可基于错误结果(例如,译码速率和失真层级)选择用于当前视频块的多个可能的译码模式中的一者,例如多个帧内译码模式(帧内预测译码模式)中的一者或多个帧间译码模式(帧间预测译码模式)中的一者。预测处理单元42可(例如)选择BVSP模式对当前块进行译码。预测处理

单元42可将所得经帧内或帧间译码块提供到求和器50以产生残余块数据,且提供到求和器62以重构经编码块以用作参考图片。

[0202] 预测处理单元42内的帧内预测单元48可相对于与待译码当前块在相同的帧或切片中的一或多个相邻块执行当前视频块的帧内预测性译码,以提供空间压缩。预测处理单元42内的运动和视差估计单元44以及运动和视差补偿单元46相对于一或多个参考图片中的一或多个预测性块执行对当前视频块的帧间预测译码以提供时间压缩。

[0203] 运动和视差估计单元44可经配置以根据用于视频序列的预定模式为视频切片确定帧间预测模式。预定模式可将序列中的视频切片指定为P切片或B切片。运动和视差估计单元44及运动和视差补偿单元46可高度集成,但出于概念目的单独地说明。由运动和视差估计单元44所执行的运动估计为产生估计视频块的运动的运动向量的过程。举例来说,运动向量可指示当前视频帧或图片内的视频块相对于参考图片内的预测性块的位移。

[0204] 预测性块是被发现在像素差方面与待译码视频块密切匹配的块,像素差可通过绝对差总和(SAD)、平方差总和(SSD)或其它差异度量来确定。在一些实例中,视频编码器20可以计算存储在参考图片存储器64中的参考图片的子整数像素位置的值。举例来说,视频编码器20可以内插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,运动和视差估计单元44可执行相对于全像素位置及分数像素位置的运动搜索,且输出具有分数像素精度的运动向量。

[0205] 运动和视差估计单元44通过将视频块的位置与参考图片的预测性块的位置进行比较来计算经帧间译码(经帧间预测译码)切片中的视频块的运动向量。参考图片可以选自第一参考图片列表(RefPicList0)或第二参考图片列表(RefPicList1),其中的每一者识别存储在参考图片存储器64中的一或多个参考图片。运动和视差估计单元44将计算的运动向量发送到熵编码单元56以及运动和视差补偿单元46。

[0206] 通过运动和视差补偿单元46执行的运动补偿可以涉及基于通过运动估计(可能执行对子像素精度的内插)确定的运动向量提取或产生预测性块。在接收到当前视频块的PU的运动向量后,运动和视差补偿单元46即刻可以在参考图片列表中的一者中定位所述运动向量指向的预测性块。视频编码器20通过从正被译码的当前视频块的像素值减去预测性块的像素值从而形成像素差值来形成残余视频块。像素差值形成用于所述块的残余数据,且可包含明度及色度差分量两者。求和器50表示可执行此减法运算的一或多个组件。运动和视差补偿单元46还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在对视频切片的视频块进行解码时使用。

[0207] 在一些实例中,运动和视差补偿单元46可执行用于当前视图中的当前块的BVSP。即,运动和视差补偿单元46可确定第一视图中的参考图片存储器64的图片。如上文更详细地阐释,运动和视差补偿单元46可通过存取深度视图中的对应块而确定当前块的视差向量。随后,使用深度块的深度值,运动和视差补偿单元46可扭曲相对于当前块中的像素的位置所确定的第一视图中的图片的像素值,以使得在不同于第一视图和当前视图的第二视图中形成经预测块。运动和视差补偿单元46可将此经预测块提供到求和器50和求和器62,分别用于计算残余和再生当前块。同样,根据本发明的技术,视频编码器20可对界定当前块的运动信息的语法数据进行编码以使得所述运动信息包含具有识别第一视图中从其合成经预测块(即,BVSP参考块)的图片的值的参考索引。

[0208] 作为由运动和视差估计单元44及运动和视差补偿单元46执行的帧间预测(如上文所描述)的替代方案,帧内预测单元48可帧内预测当前块。明确地说,帧内预测单元48可以确定用来对当前块进行编码的帧内预测模式。在一些实例中,帧内预测单元48可以(例如)在单独的编码编次期间使用各种帧内预测模式来编码当前块,并且帧内预测单元48(或在一些实例中为模式选择单元)可以从测试模式中选择适当帧内预测模式来使用。举例来说,帧内预测单元48可以使用速率失真分析计算针对各种经测试帧内预测模式的速率失真值,并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。速率失真分析一般确定经编码块与经编码以产生所述经编码块的原始的未经编码块之间的失真(或误差)的量,以及用于产生经编码块的位速率(也就是说,位数目)。帧内预测单元48可以根据用于各种经编码块的失真和速率计算比率,以确定哪个帧内预测模式对于所述块展现最佳速率失真值。

[0209] 在任何情况下,在选择用于块的帧内预测模式之后,帧内预测单元48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56。熵编码单元56可根据本发明的技术编码指示所选帧内预测模式的信息。视频编码器20在发射的位流中可包含配置数据,其可包含多个帧内预测模式索引表及多个经修改的帧内预测模式索引表(也称为码字映射表),对用于各种块的上下文进行编码的定义,及对最可能帧内预测模式、帧内预测模式索引表及经修改的帧内预测模式索引表的指示以用于所述上下文中的每一者。

[0210] 在预测处理单元42经由帧间预测或帧内预测产生当前视频块的预测块之后,视频编码器20通过从当前视频块减去预测块而形成残余视频块。残余块中的残余视频数据可应用于变换处理单元52。变换处理单元52使用例如离散余弦变换(DCT)或概念上类似的变换等变换将残余视频数据变换成残余变换系数。变换处理单元52可将残余视频数据从像素值域转换到变换域,例如频域。

[0211] 变换处理单元52可将所得变换系数发送到量化处理单元54。量化处理单元54可量化所述变换系数以进一步减小位速率。量化过程可减少与变换系数中的一些或全部相关联的位深度。可通过调整量化参数来修改量化程度。在一些实例中,量化处理单元54可接着执行对包含经量化变换系数的矩阵的扫描。替代地,熵编码单元56可以执行所述扫描。

[0212] 在量化之后,熵编码单元56对经量化变换系数进行熵编码。举例来说,熵编码单元56可执行上下文自适应可变长度译码(CAVLC)、上下文自适应二进制算术译码(CABAC)、基于语法的上下文自适应二进制算术译码(SBAC)、概率区间分割熵(PIPE)译码或另一熵编码方法或技术。在由熵编码单元56进行熵编码后,可将经编码位流发射到视频解码器30,或将经编码位流存档以供稍后由视频解码器30发射或检索。熵编码单元56还可熵编码正被译码的当前视频切片的运动向量及其它语法元素。

[0213] 逆量化处理单元58和逆变换处理单元60分别应用逆量化和逆变换以在像素域中重构残余块,例如以供稍后用作参考图片的参考块。运动和视差补偿单元46可通过将残余块加到参考图片列表中的一者内的参考图片中的一者的预测性块来计算参考块。运动和视差补偿单元46还可将一或多个内插滤波器应用于经重构残余块以计算用于运动估计的子整数像素值。求和器62将经重构的残余块加到由运动和视差补偿单元46产生的经运动补偿的预测块以产生参考块以用于存储在参考图片存储器64中。参考块可由运动和视差估计单元44以及运动和视差补偿单元46用作参考块以帧间预测后续视频帧或图片中的块。

[0214] 以此方式,视频编码器20是可经配置以实施本发明中描述的一或多个实例技术的视频编码器的实例。举例来说,视频数据存储器40存储视频数据。视频数据可包含相依视图的纹理视图分量和对应于纹理视图分量的深度视图分量,视频编码器20将在3D-AVC顺应或3D-HEVC顺应视频译码过程中对其中的每一者进行编码。

[0215] 在本发明中描述的技术中,视频编码器20可包含经配置以在3D-AVC顺应或3D-HEVC顺应视频译码过程中对视频数据的相依视图的纹理视图分量进行编码的一或多个处理器。如上文所描述,3D-AVC中的每一视图包含纹理视图分量和深度视图分量。3D-AVC中存在一个基础视图以及一或多个增强或相依视图,其中一或多个增强或相依视图的纹理视图分量可进行视图间预测。

[0216] 为了对纹理视图分量进行编码,视频编码器20可经配置以评估纹理视图分量中的当前块的一或多个相邻块的运动信息以确定至少一个相邻块是否是以参考除相依视图外的视图中的视图间参考图片的视差运动向量进行视图间预测。视频编码器20可基于相邻块中的一者的视差运动向量导出当前块的视差向量。对于纹理优先译码,视频编码器20可在编码纹理视图分量之后编码对应于纹理视图分量的视频数据的深度视图分量。

[0217] 在一些实例中,视频编码器20的预测处理单元42可为经配置以实施本发明中针对视差向量导出和BVSP译码描述的实例的处理器中的一个实例。在一些实例中,除预测处理单元42外的单元(例如,一或多个处理器)可实施上述实例。在一些实例中,预测处理单元42与视频编码器20的一或多个其它单元结合可实施上述实例。在一些实例中,视频编码器20的处理器(图8中未图示)可单独或与视频编码器20的其它处理器结合实施上述实例。

[0218] 图9是说明可实施本发明中描述的技术的视频解码器的实例的框图。图9是说明可实施本发明中描述的技术的视频解码器的实例的框图。举例来说,图9说明可表示3D-AVC顺应或3D-HEVC顺应视频解码器的视频解码器30。将使用例如PU、TU和CU等某些HEVC术语描述视频解码器30,但应理解参考视频解码器30描述的技术也可以关于根据H.264标准译码的视频而执行。

[0219] 视频解码器30可执行帧间预测解码或帧内预测解码。图9说明视频解码器30。在图9的实例中,视频解码器30包含视频数据存储器69、熵解码单元70、预测处理单元71、逆量化处理单元76、逆变换处理单元78、求和器80和参考图片存储器82。预测处理单元71包含运动和视差补偿单元72和帧内预测单元74。在一些实例中,视频解码器30可执行一般与关于图8的视频编码器20描述的编码遍次互逆的解码遍次。

[0220] 视频数据存储器69可存储待由视频解码器30的组件解码的视频数据,例如经编码视频位流。存储在视频数据存储器69中的视频数据可(例如)从存储装置34、从例如相机等本地视频源、经由视频数据的有线或无线网络通信或者通过存取物理数据存储媒体而获得。视频数据存储器69可形成存储来自经编码视频位流的经编码视频数据的经译码图片缓冲器(CPB)。

[0221] 参考图片存储器82是存储参考视频数据供视频解码器30解码视频数据(例如,在帧内或帧间译码模式中)的经解码图片缓冲器(DPB)的一个实例。视频数据存储器69和参考图片存储器82可由多种存储器装置中的任一者形成,例如包含同步DRAM(SDRAM)的动态随机存取存储器(DRAM)、磁阻式RAM(MRAM)、电阻式RAM(RRAM)或其它类型的存储器装置。视频数据存储器69和参考图片存储器82可由同一存储器装置或单独的存储器装置提供。在各种

实例中,视频数据存储单元69可与视频解码器30的其它组件一起在芯片上,或相对于那些组件在芯片外。

[0222] 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的经编码视频位流和相关联的语法元素。视频解码器30的熵解码单元70对位流进行熵解码以产生经量化系数、运动向量和其它语法元素。熵解码单元70将运动向量及其它语法元素转发到预测处理单元71。视频解码器30可在视频切片层级和/或视频块层级接收语法元素。

[0223] 当视频切片经译码为经帧内译码(I)切片时,预测处理单元71的帧内预测单元74可以基于用信号表示的帧内预测模式和来自当前帧或图片的先前经解码块的数据产生用于当前视频切片的视频块的预测数据。当视频帧经译码为经帧间译码(即,B或P)切片时,预测处理单元71的运动和视差补偿单元72基于从熵解码单元70接收的运动向量及其它语法元素产生用于当前视频切片的视频块的预测性块。预测性块可以从参考图片列表中的一者内的参考图片中的一者产生。视频解码器30可基于存储在参考图片存储器82中的参考图片使用默认构造技术构造参考图片列表(RefPicList0和RefPicList1)。

[0224] 运动和视差补偿单元72通过剖析运动向量和其它语法元素而确定当前视频切片的视频块的预测信息,且使用所述预测信息来产生正被解码的当前视频块的预测性块。举例来说,运动和视差补偿单元72使用所接收语法元素中的一些语法元素确定用于译码视频切片的视频块的预测模式(例如,帧内预测或帧间预测)、帧间预测切片类型(例如,B切片或P切片)、切片的参考图片列表中的一或多者的构造信息、切片的每一经帧间编码的视频块的运动向量、切片的每一经帧间译码的视频块的帧间预测状态,及用以解码当前视频切片中的视频块的其它信息。

[0225] 运动和视差补偿单元72还可执行基于内插滤波器的内插。运动和视差补偿单元72可在编码视频块期间使用如由视频编码器20使用的内插滤波器来计算用于参考块的子整数像素的内插值。在此情况下,运动和视差补偿单元72可从所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产生预测性块。

[0226] 运动和视差补偿单元72可执行用于当前视图中的当前块的后向视图合成预测。即,运动和视差补偿单元72可确定第一视图中的参考图片存储器82的图片。如上文更详细地阐释,运动和视差补偿单元72可确定对应于当前块的深度块,且使用所述深度块的深度值,运动和视差补偿单元72可扭曲相对于当前块中的像素的位置所确定的第一视图中的图片的像素值,以使得在不同于第一视图和当前视图的第二视图中形成经预测块(即,BVSP参考块)。运动和视差补偿单元72可将此经预测块提供到求和器50和求和器80,分别用于计算残余和再生当前块。同样,根据本发明的技术,视频解码器30可对界定当前块的运动信息的语法数据进行解码以使得所述运动信息包含具有识别第一视图中从其合成经预测块的图片的值的参考索引。

[0227] 逆量化处理单元76将在位流中提供且由熵解码单元70解码的经量化变换系数逆量化(即,解量化)。逆量化过程可包含使用由视频编码器20针对视频切片中的每一视频块计算的量化参数以确定应应用的量化程度及同样确定应应用的逆量化程度。逆变换处理单元78将逆变换应用于变换系数(例如,逆DCT、逆整数变换或概念上类似的逆变换过程),以便产生像素域中的残余块。

[0228] 在运动和视差补偿单元72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后,视频解码器30通过对来自逆变换处理单元78的残余块与由运动和视差补偿单元72产生的对应预测性块求和而形成经解码的视频块。求和器80表示执行此求和运算的一或多个组件。必要时,解块滤波器还可应用于对经解码块进行滤波以便移除成块假象。其它环路滤波器(再译码环路中或在译码环路之后)也可用于使像素转变平滑或者以其它方式改善视频质量。给定图片中的经解码视频块接着存储于参考图片存储器82中,所述参考图片存储器存储用于后续运动补偿的参考图片。参考图片存储器82还存储经解码视频以用于稍后在显示装置(例如,图1的显示装置32)上呈现。

[0229] 以此方式,视频解码器30是可经配置以实施本发明中描述的一或多个实例技术的视频解码器的实例。举例来说,视频数据存储器69存储视频数据。视频数据可包含视频解码器30可从其解码相依视图的纹理视图分量和对应于纹理视图分量的深度视图分量的信息,视频编码器20在3D-AVC顺应或3D-HEVC顺应视频译码过程中对其中的每一者进行编码。

[0230] 在本发明中描述的技术中,视频解码器30可包含经配置以在3D-AVC顺应或3D-HEVC顺应视频译码过程中解码视频数据的相依视图的纹理视图分量的一或多个处理器。为了对纹理视图分量进行解码,视频解码器30可经配置以评估纹理视图分量中的当前块的一或多个相邻块的运动信息以确定至少一个相邻块是否是以参考除相依视图外的视图中的视图间参考图片的视差运动向量进行视图间预测。视频解码器30可基于相邻块中的一者的视差运动向量导出当前块的视差向量。对于纹理优先译码,视频解码器30可在解码纹理视图分量之后解码对应于纹理视图分量的视频数据的深度视图分量。

[0231] 在一些实例中,视频解码器30的预测处理单元71可为经配置以实施本发明中针对视差向量导出和BVSP译码描述的实例的处理器。在一些实例中,除预测处理单元71外的单元(例如,一或多个处理器)可实施上述实例。在一些实例中,预测处理单元71与视频解码器30的一或多个其它单元结合可实施上述实例。在又一些其它实例中,视频解码器30的处理器(图9中未图示)可单独或与视频解码器30的其它处理器结合实施上述实例。

[0232] 图10是展示根据本发明的技术的实例性3D视频译码过程的流程图。图10的技术可适用于3D-AVC顺应视频或3D-HEVC顺应视频。将相对于通用视频译码器描述图10的技术,所述通用视频译码器可例如为例如视频编码器20的视频编码器或例如视频解码器的视频解码器。视频译码器确定将使用BVSP模式对第一纹理视图的块进行译码(110)。当执行视频解码时,视频解码器30可(例如)通过接收指示将使用基于块的视图合成模式对第一纹理视图的块进行译码的语法元素而确定将使用BVSP模式对第一纹理视图的块进行解码。当执行视频编码时,视频编码器20可(例如)通过执行若干译码遍次且将BVSP模式识别为产生所要速率失真折中的模式而确定将使用BVSP模式对第一纹理视图的块进行译码。

[0233] 视频译码器在深度视图中定位对应于第一纹理视图的块的深度块(112)。在3D-AVC中,块在此实例中可指代宏块或宏块分区的 $K \times K$ 子区。在3D-HEVC中,块还可指代 $K \times K$ 子区。对应深度块可(例如)为3D-AVC中的位于同一地点的深度块,或可为3D-HEVC中通过使用NBDV产生的视差向量识别的参考基础视图(即,第二视图)中的深度块。视频译码器确定深度块的两个或两个以上拐角位置的深度值(114)。基于所述深度值,视频译码器导出所述块的视差向量(116)。使用视差向量,视频译码器定位第二纹理视图的块(118)。视频译码器使用第二纹理视图的块帧间预测第一纹理视图的块(120)。第一纹理视图可(例如)为非基础

纹理视图,且第二纹理视图可(例如)为基础纹理视图。当图10的技术由3D-AVC顺应视频译码器实施时,第一纹理视图的块可(例如)为宏块分区的子块。当图10的技术由3D-HEVC顺应译码器实施时,第一纹理视图的块可(例如)为预测单元。

[0234] 图11是展示根据本发明的技术的实例性3D视频译码过程的流程图。将使用3D-AVC术语描述图11的技术,但所述技术可潜在地扩展到例如3D-HEVC等其它视频译码标准。将相对于通用视频译码器描述图11的技术,所述通用视频译码器可例如为例如视频编码器20的视频编码器或例如视频解码器的视频解码器。视频译码器执行用于第一纹理视图、第一深度视图、第二纹理视图和第二深度视图的纹理优先译码(122)。第一纹理视图和第一基础深度视图可(例如)为基础视图,而第二纹理视图和第二深度视图是非基础视图。

[0235] 对于第二纹理视图的宏块,视频译码器在第一深度视图中定位对应于所述宏块的深度块(124)。基于深度块的至少一个深度值,视频译码器导出宏块的视差向量(126)。视频译码器可例如通过确定包含深度块的两个或两个以上拐角样本的深度值的深度值的集合而导出视差向量,且从深度值集合识别最大深度值。随后可基于将深度值转换为视差值的转换表或使用某种其它技术将最大深度值转换为视差向量。视频译码器基于所导出的视差向量对宏块的第一子块进行译码(128)。视频译码器基于所导出的视差向量对宏块的第二子块进行译码(130)。视频译码器可(例如)使用跳过模式和直接模式中的一者对第一子块进行译码且使用除跳过模式或直接模式外的帧间预测模式对第二子块进行译码。在图11的实例中,子块可例如为H.264/AVC子块或可为宏块分区。

[0236] 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。如果用软件实施,那么所述功能可以作为一或多个指令或代码在计算机可读媒体上存储或传输,并且由基于硬件的处理单元执行。计算机可读媒体可包含计算机可读存储媒体,其对应于有形媒体,例如数据存储媒体,或包含促进将计算机程序从一处传送到另一处的任何媒体(例如,根据通信协议)的通信媒体。以此方式,计算机可读媒体总体上可对应于(1)非暂时性的有形计算机可读存储媒体,或(2)通信媒体,例如信号或载波。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用于实施本发明中描述的技术的指令、代码和/或数据结构的任何可用媒体。计算机程序产品可包含计算机可读媒体。

[0237] 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、快闪存储器或可以用来存储指令或数据结构的形式的期望程序代码并且可以由计算机存取的任何其它媒体。并且,可恰当地将任何连接称作计算机可读媒体。举例来说,如果使用同轴电缆、光纤缆线、双绞线、数字订户线(DSL)或例如红外线、无线电及微波等无线技术从网站、服务器或其它远程源传输指令,则同轴电缆、光纤缆线、双绞线、DSL或例如红外线、无线电及微波等无线技术包括在媒体的定义中。但是,应理解,所述计算机可读存储媒体和数据存储媒体并不包括连接、载波、信号或其它暂时媒体,而是实际上针对于非暂时性有形存储媒体。如本文所使用,磁盘及光盘包含压缩光盘(CD)、激光光盘、光学光盘、数字多功能光盘(DVD)、软性磁盘及蓝光光盘,其中磁盘通常以磁性方式再现数据,而光盘用激光以光学方式再现数据。上述各项的组合也应包含在计算机可读媒体的范围内。

[0238] 指令可由一个或多个处理器执行,所述一个或多个处理器例如是一或多个数字信号处理器(DSP)、通用微处理器、专用集成电路(ASIC)、现场可编程逻辑阵列(FPGA),或其它

等效的集成或离散逻辑电路。因此,如本文中所使用的术语“处理器”可指上述结构或适合于实施本文中所描述的技术的任何其它结构中的任一者。另外,在一些方面中,本文所描述的功能性可以提供于经配置以用于编码及解码的专用硬件及/或软件模块内,或者并入于组合编解码器中。并且,可将所述技术完全实施于一或多个电路或逻辑元件中。

[0239] 本发明的技术可在广泛多种装置或设备中实施,包含无线手持机、集成电路(IC)或一组IC(例如,芯片组)。本发明中描述各种组件、模块或单元是为了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实现。实际上,如上文所描述,各种单元可以结合合适的软件和/或固件组合在编解码器硬件单元中,或者通过互操作硬件单元的集合来提供,所述硬件单元包括如上文所描述的一或多个处理器。

[0240] 已描述各种实例。这些及其它实例在所附权利要求书的范围内。

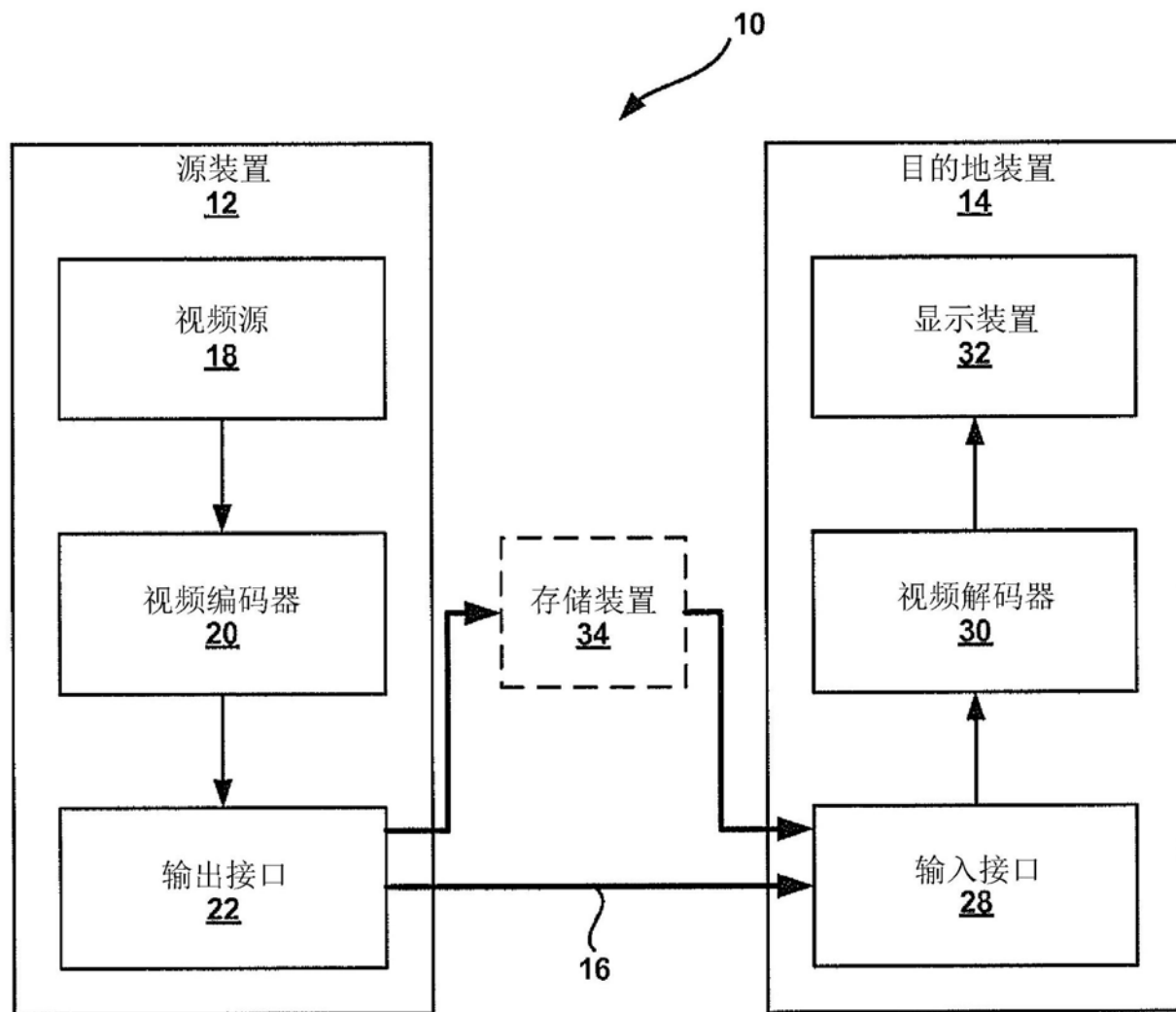


图1

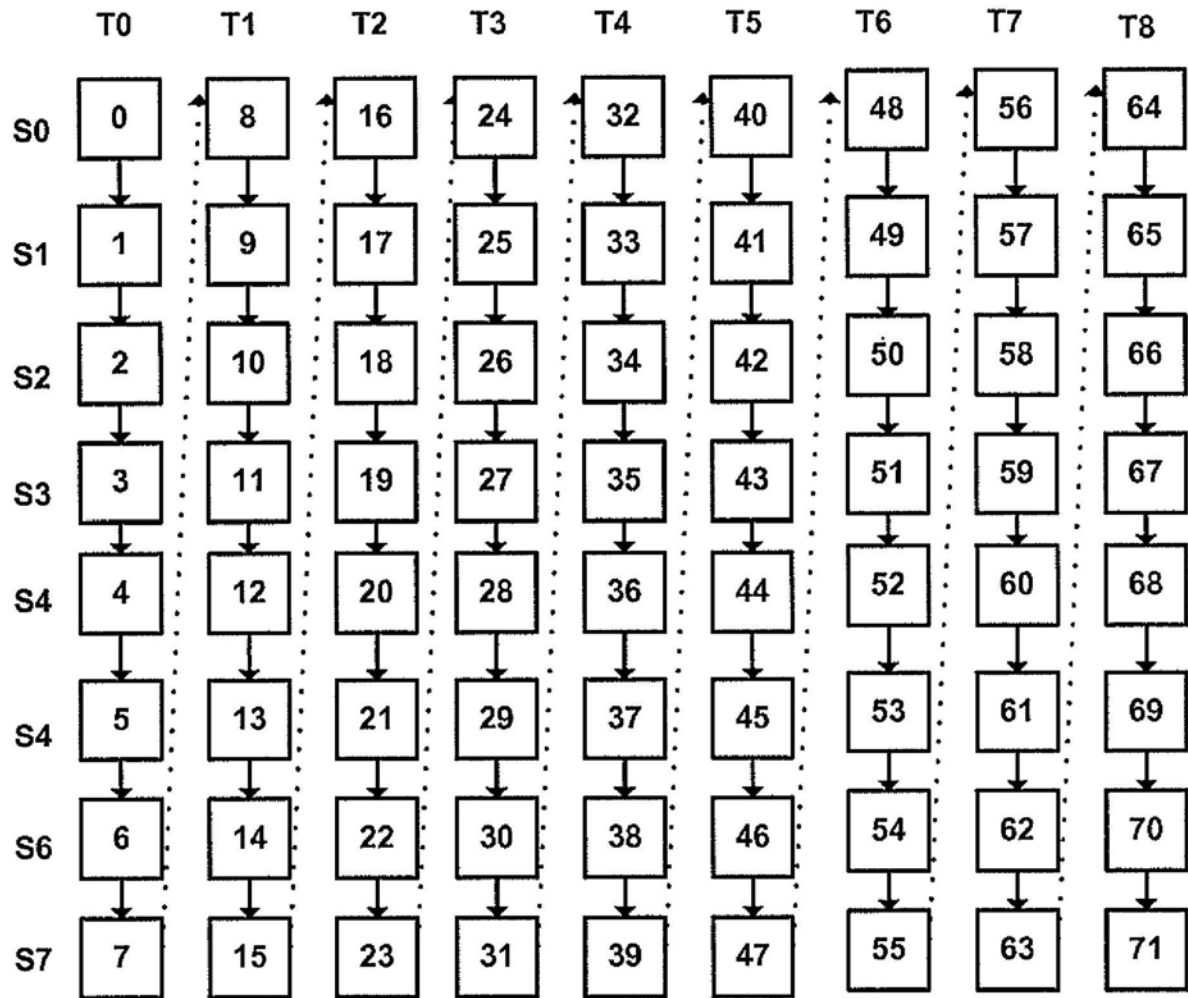


图2

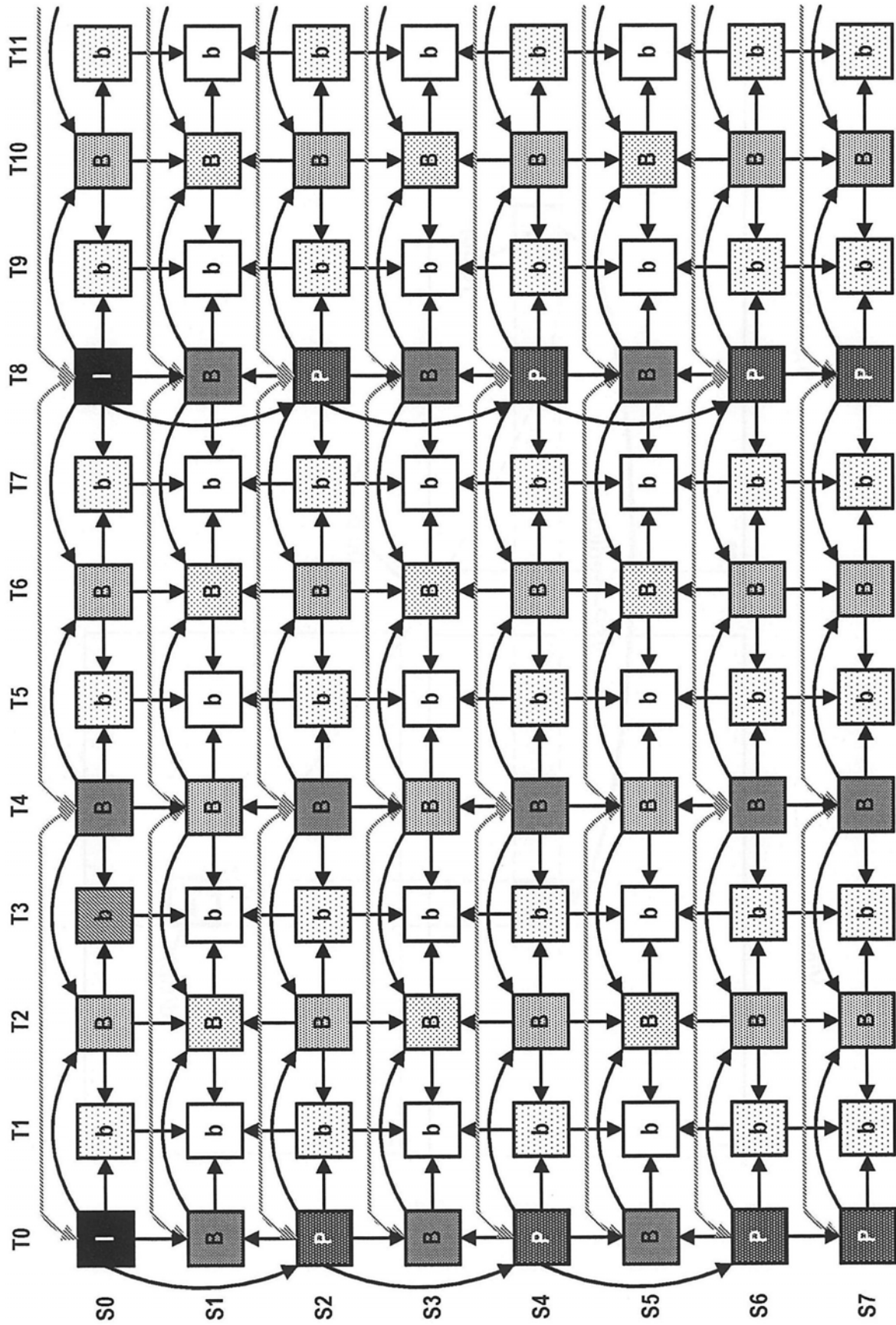


图3

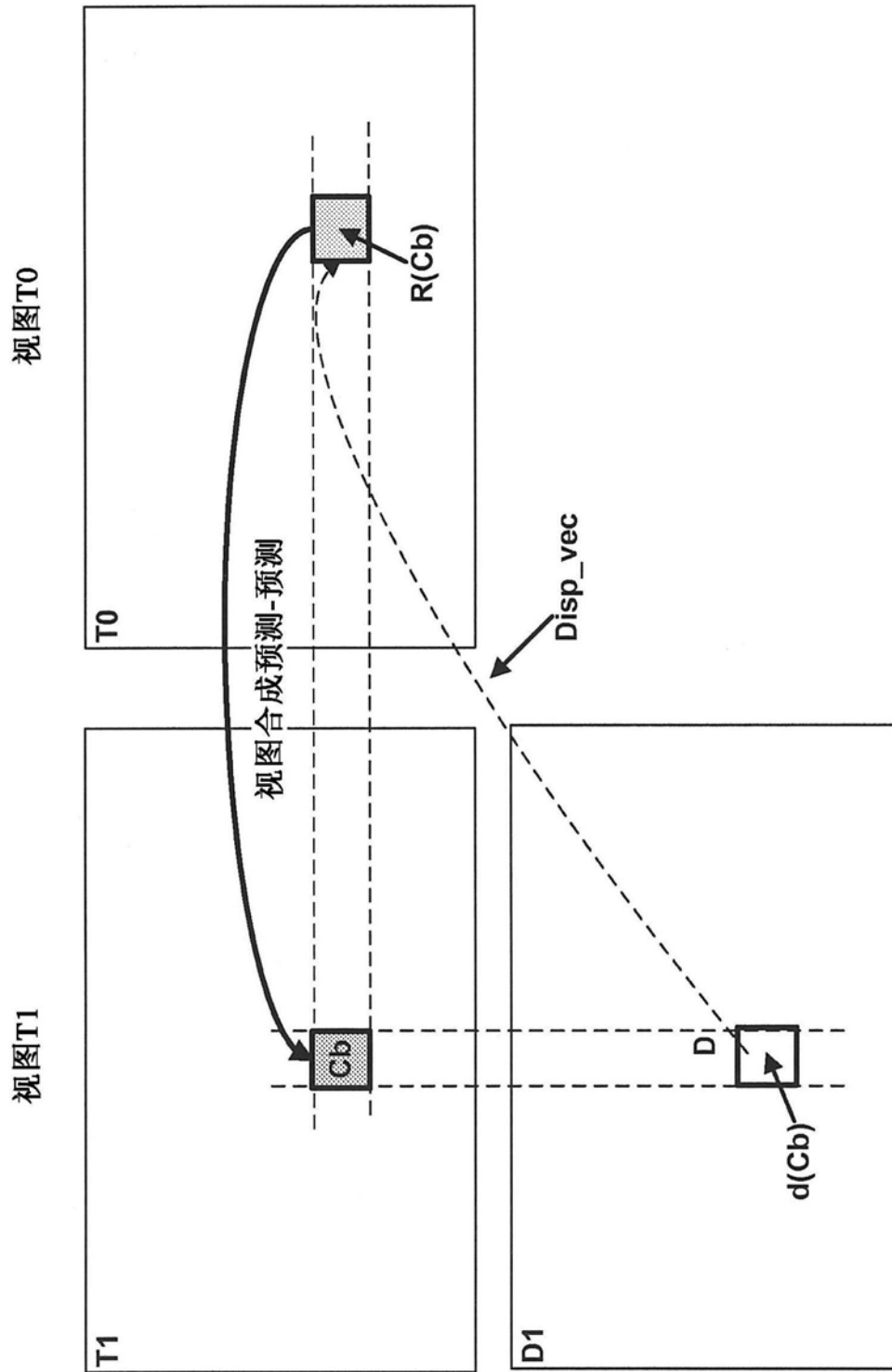


图4

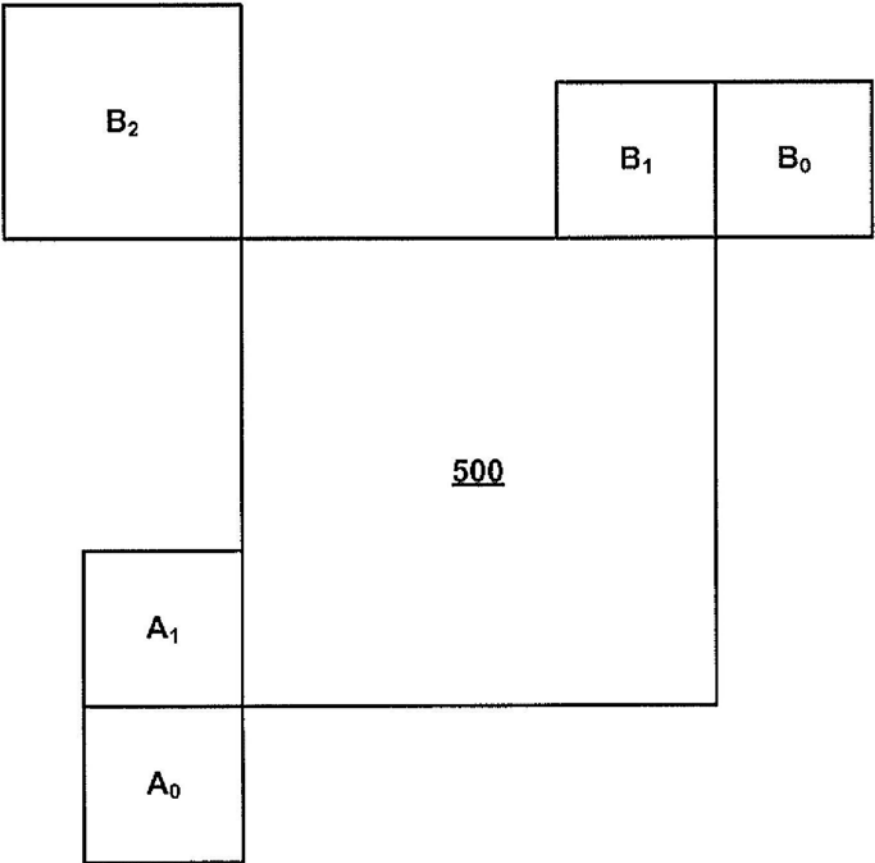


图5

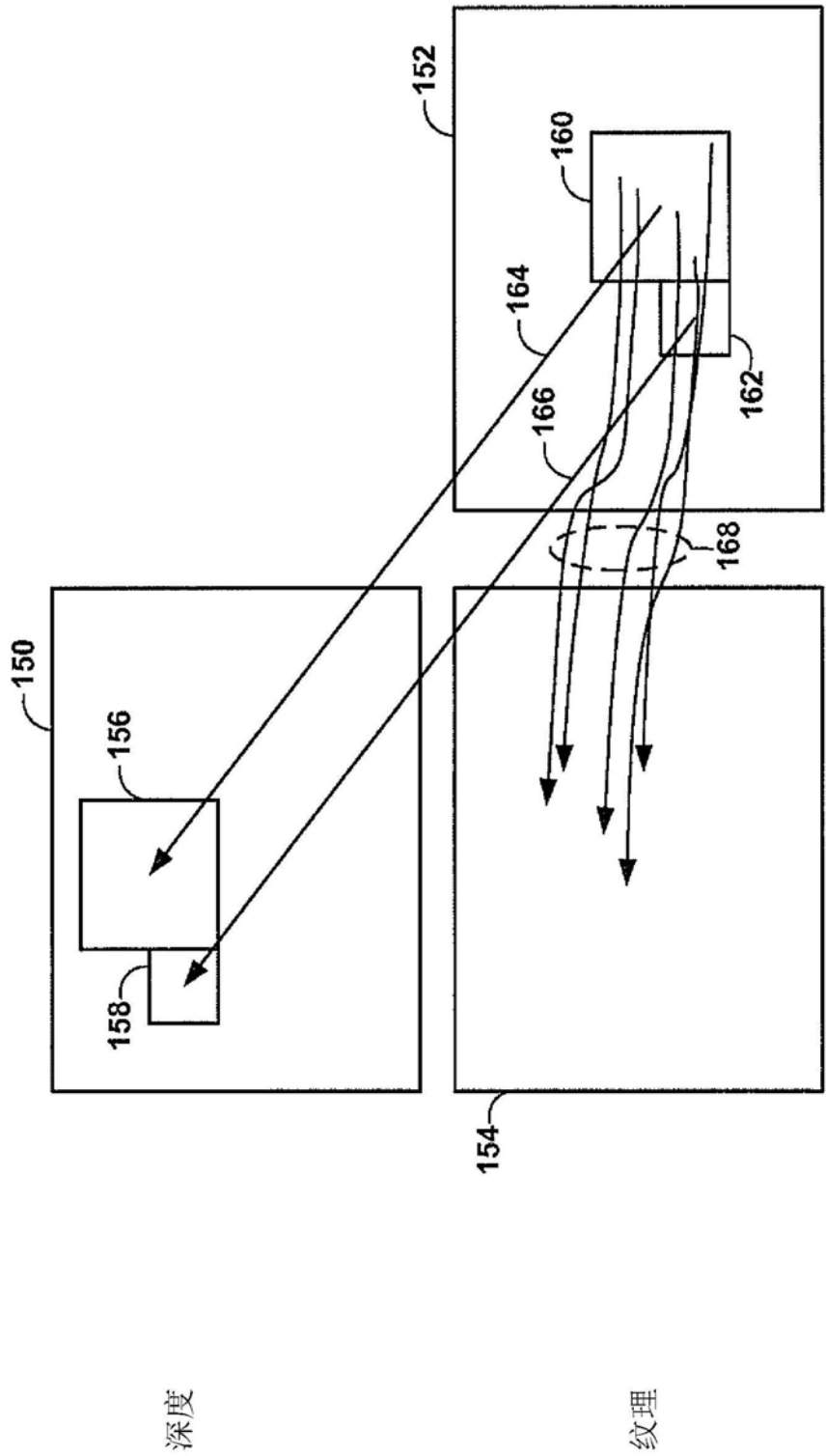


图6

170

172C							172D
172A							172B

图7

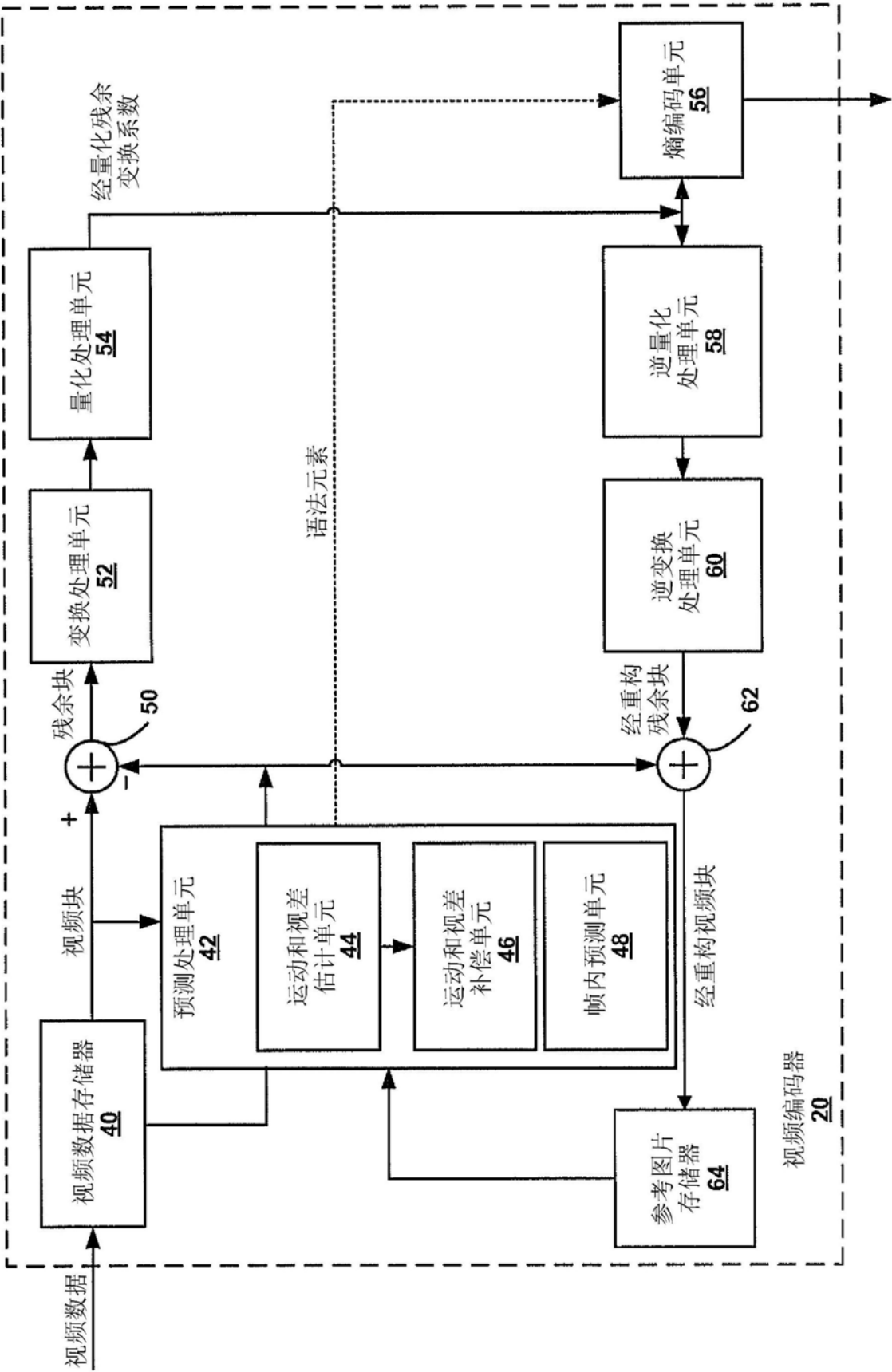


图8

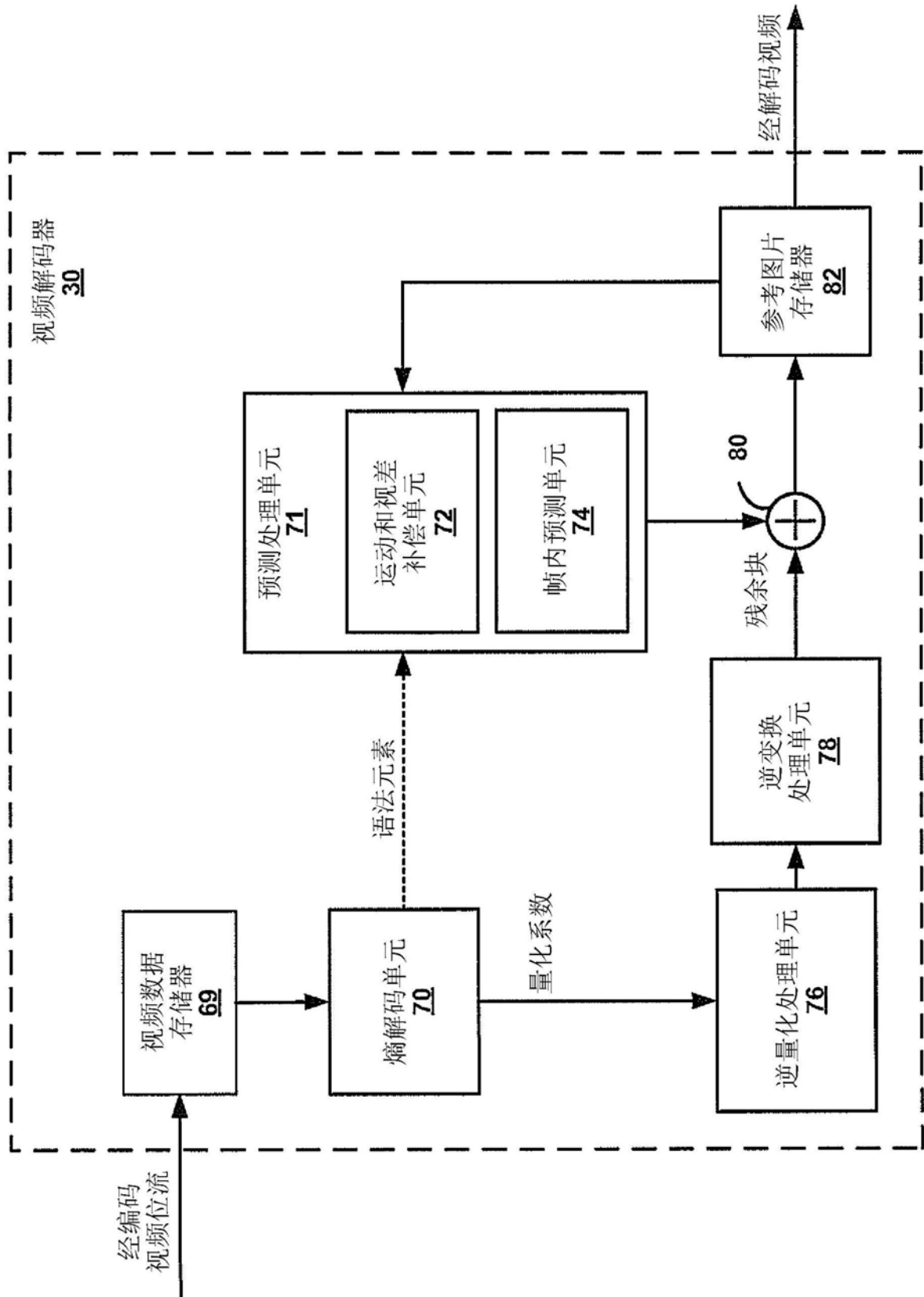


图9

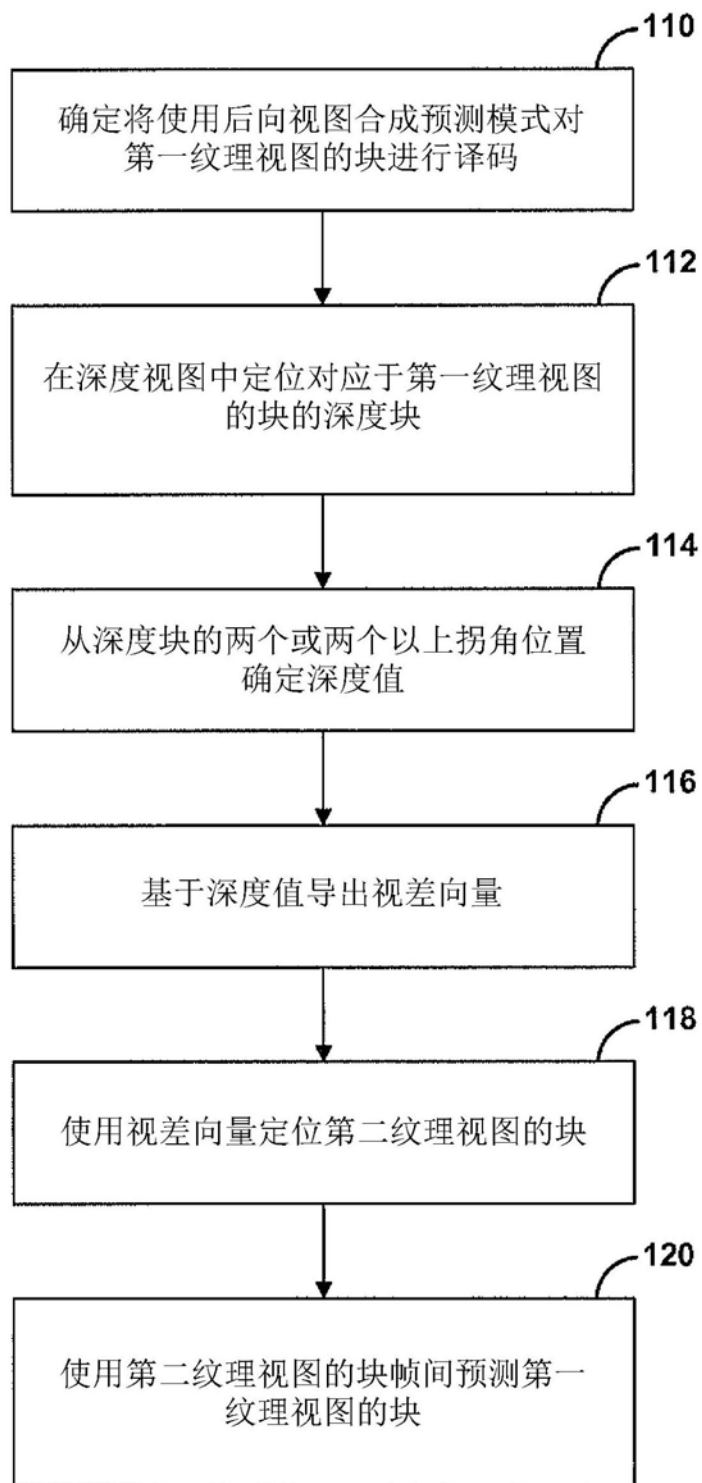


图10

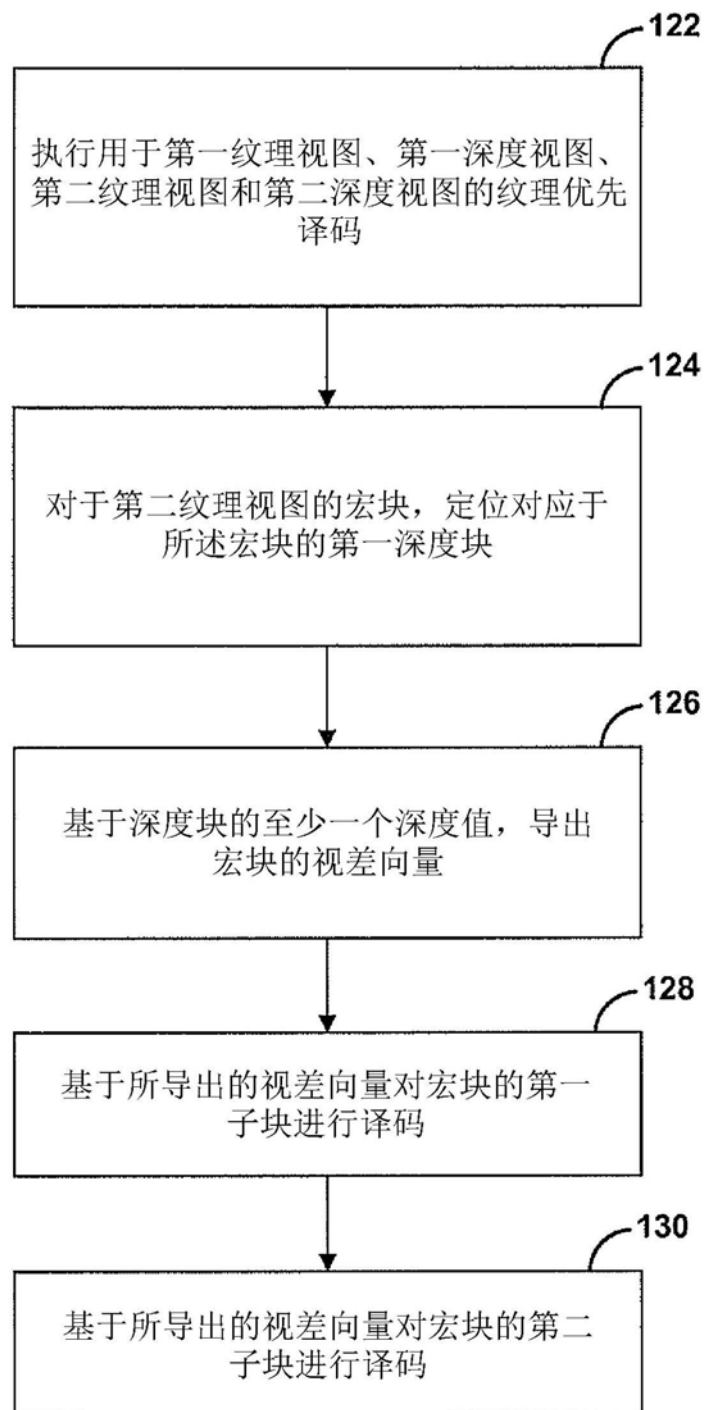


图11