



## (12) 发明专利

(10) 授权公告号 CN 110139108 B

(45) 授权公告日 2023.07.18

(21) 申请号 201910417853.8

(74) 专利代理机构 北京康信知识产权代理有限

(22) 申请日 2012.11.09

责任公司 11240

(65) 同一申请的已公布的文献号

专利代理人 刘丹

申请公布号 CN 110139108 A

(51) Int.CI.

(43) 申请公布日 2019.08.16

H04N 19/597 (2014.01)

(30) 优先权数据

H04N 19/52 (2014.01)

61/558,651 2011.11.11 US

H04N 19/46 (2014.01)

(62) 分案原申请数据

H04N 19/194 (2014.01)

201280066739.1 2012.11.09

(73) 专利权人 GE视频压缩有限责任公司

(56) 对比文件

地址 美国纽约

CN 101690220 A, 2010.03.31

(72) 发明人 海科·施瓦茨 托马斯·维甘德

CN 101867813 A, 2010.10.20

CN 101986716 A, 2011.03.16

JP 2008022549 A, 2008.01.31

WO 2007037645 A1, 2007.04.05

审查员 田秋硕

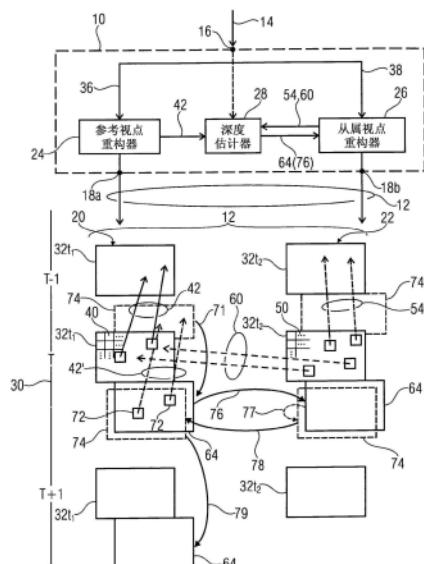
权利要求书5页 说明书35页 附图8页

## (54) 发明名称

用于将多视点信号编码到多视点数据流中的装置及方法

## (57) 摘要

本发明提供一种用于将多视点信号编码到多视点数据流中的装置及方法,该装置包括如下配置的从属视点编码器:针对多视点信号的从属视点中的当前图片的至少一个区块,推导运动矢量预测因子候选者列表;针对从属视点的当前图片的至少一个区块,将指定运动矢量预测因子候选者列表的一个运动矢量预测因子候选者的索引信息插入到多视点数据流中;以及通过使用取决于所指定的运动矢量候选者的运动矢量执行从属视点的当前图片的至少一个区块的运动补偿预测来编码从属视点的当前图片的至少一个区块。



1. 一种用于重构被编码到数据流中的多视点信号的装置,包括:  
从属视点重构器,所述从属视点重构器被配置用于:  
处理与所述多视点信号的从属视点中的当前图片的编码块相关联的多个运动矢量候选者的列表,所述处理包括:  
    基于与所述多视点信号的参考视点的当前图片中的参考块相关联的第二运动矢量估计第一运动矢量,所述参考块与所述从属视点中的当前图片的所述编码块相对应,  
    将所述第一运动矢量添加到所述列表中,以及  
    从所述数据流中提取索引信息,所述索引信息从所述列表中指定至少一个候选者用于运动补偿预测;以及  
    通过基于所述至少一个候选者执行所述运动补偿预测来重构所述从属视点的所述当前图片中的所述编码块,  
    其中,所述从属视点重构器被配置用于:获得关于所述从属视点中的当前图片的所述编码块的视差矢量,其中,所述视差矢量表示所述从属视点的当前图片与所述参考视点的当前图片之间的视差,和基于所述视差矢量识别所述参考视点的当前图片中的所述参考块,  
    所述装置进一步包括深度估计器,所述深度估计器被配置用于:  
    获得与所述参考视点的当前图片相关联的运动数据;  
    将与所述参考视点的当前图片相关联的所述运动数据应用于所述参考视点的先前图片的深度图估计,以生成所述参考视点的当前图片的第一估计深度图;  
    基于所述参考视点的当前图片的所述第一估计深度图,得出所述从属视点的当前图片的第二估计深度图;以及  
    基于所述从属视点的当前图片的所述第二估计深度图,确定所述视差矢量。  
2. 根据权利要求1所述的装置,其中,所述从属视点中的所述编码块的所述视差矢量还基于与所述多视点信号的一个或多个先前解码的编码块相关联的至少一个视差矢量。  
3. 根据权利要求1所述的装置,其中,所述从属视点重构器还被配置用于:  
从所述数据流中提取针对所选择的运动向量的运动向量残差;  
基于所选择的运动矢量和运动矢量残差,获得精细运动矢量;以及  
基于所述精细运动矢量执行运动补偿预测,以预测所述编码块。  
4. 根据权利要求1所述的装置,其中,所述列表包括以下的一项或多项:  
基于与所述从属视点的当前图片中的空间邻居编码块相关联的运动数据确定的第一估计运动矢量,  
基于与所述从属视点的先前图片中的时间相邻编码块相关联的运动数据确定的第二估计运动矢量,  
通过修改所述第一估计运动矢量得出的第一修改估计运动矢量,以及  
通过修改所述第二估计运动矢量得出的第二修改估计运动矢量。  
5. 一种非暂时性机器可读介质,存储有用于重构被编码到数据流中的多视点信号的信息,其中,当机器读取所述信息时,使所述机器执行多个操作,包括:  
    处理与所述多视点信号的从属视点中的当前图片的编码块相关联的多个运动矢量候选者的列表,所述处理包括:

基于与所述多视点信号的参考视点的当前图片中的参考块相关联的第二运动矢量估计第一运动矢量,所述参考块与所述从属视点中的当前图片的所述编码块相对应,

将所述第一运动矢量添加到所述列表中,以及

从所述数据流中提取索引信息,所述索引信息从所述列表中指定至少一个候选者用于运动补偿预测;以及

通过基于所述至少一个候选者执行所述运动补偿预测来重构所述从属视点中的当前图片的所述编码块,

获得关于所述从属视点中的当前图片的所述编码块的视差矢量,其中,所述视差矢量表示所述从属视点的当前图片与所述参考视点的当前图片之间的视差,和基于所述视差矢量识别所述参考视点的当前图片中的所述参考块,

其中,获得所述视差矢量包括:

获得与所述参考视点的当前图片相关联的运动数据;

将与所述参考视点的当前图片相关联的所述运动数据应用于所述参考视点的先前图片的深度图估计,以生成所述参考视点的当前图片的第一估计深度图;

基于所述参考视点的当前图片的所述第一估计深度图,得出所述从属视点的当前图片的第二估计深度图;以及

基于所述从属视点的当前图片的所述第二估计深度图,确定所述视差矢量。

6.根据权利要求5所述的非暂时性机器可读介质,其中,所述从属视点中的所述编码块的所述视差矢量还基于与所述多视点信号的一个或多个先前解码的编码块相关联的至少一个视差矢量。

7.根据权利要求5所述的非暂时性机器可读介质,其中,所述多个操作进一步包括:

从所述数据流中提取针对所选择的运动向量的运动向量残差;

基于所选择的运动矢量和运动矢量残差,获得精细运动矢量;以及

基于所述精细运动矢量执行运动补偿预测,以预测所述编码块。

8.根据权利要求5所述的非暂时性机器可读介质,其中,所述列表包括以下的一项或多项:

基于与所述从属视点的当前图片中的空间邻居编码块相关联的运动数据确定的第一估计运动矢量,

基于与所述从属视点的先前图片中的时间相邻编码块相关联的运动数据确定的第二估计运动矢量,

通过修改所述第一估计运动矢量得出的第一修改估计运动矢量,以及

通过修改所述第二估计运动矢量得出的第二修改估计运动矢量。

9.一种用于将多视点信号编码到数据流中的装置,包括:

从属视点编码器,所述从属视点编码器被配置用于:

处理与所述多视点信号的从属视点中的当前图片的编码块相关联的多个运动矢量候选者的列表,所述处理包括:

基于与所述多视点信号的参考视点的当前图片中的参考块相关联的第二运动矢量估计第一运动矢量,所述参考块与所述从属视点中的当前图片的所述编码块相对应,

将所述第一运动矢量添加到所述列表中,

通过运动补偿预测从所述列表中确定要用于重构所述从属视点的当前图片中的所述编码块的至少一个候选者,以及

生成指定所述至少一个候选者的索引信息;以及

将所述索引信息插入所述数据流,

其中,所述从属视点编码器被配置用于:获得关于所述从属视点中的当前图片的所述编码块的视差矢量,其中,所述视差矢量表示所述从属视点的当前图片与所述参考视点的当前图片之间的视差,和基于所述视差矢量识别所述参考视点的当前图片中的所述参考块,

所述装置进一步包括深度估计器,所述深度估计器被配置用于通过以下步骤估计所述视差矢量:

获得与所述参考视点的当前图片相关联的运动数据;

将与所述参考视点的当前图片相关联的所述运动数据应用于所述参考视点的先前图片的深度图估计,以生成所述参考视点的当前图片的第一估计深度图;

基于所述参考视点的当前图片的所述第一估计深度图,得出所述从属视点的当前图片的第二估计深度图;以及

基于所述从属视点的当前图片的所述第二估计深度图,确定所述视差矢量。

10.根据权利要求9所述的装置,其中,所述从属视点中的所述编码块的所述视差矢量还基于与所述多视点信号的一个或多个先前解码的编码块相关联的至少一个视差矢量。

11.根据权利要求9所述的装置,其中,所述从属视点编码器还被配置用于:

从所述数据流中提取针对所选择的运动向量的运动向量残差;

基于所选择的运动矢量和运动矢量残差,获得精细运动矢量;以及

基于所述精细运动矢量执行运动补偿预测,以预测所述编码块。

12.根据权利要求9所述的装置,其中,所述列表包括以下的一项或多项:

基于与所述从属视点的当前图片中的空间邻居编码块相关联的运动数据确定的第一估计运动矢量,

基于与所述从属视点的先前图片中的时间相邻编码块相关联的运动数据确定的第二估计运动矢量,

通过修改所述第一估计运动矢量得出的第一修改估计运动矢量,以及

通过修改所述第二估计运动矢量得出的第二修改估计运动矢量。

13.一种用于将多视点信号编码到数据流中的方法,包括:

处理与所述多视点信号的从属视点中的当前图片的编码块相关联的多个运动矢量候选者的列表,

基于与所述多视点信号的参考视点的当前图片中的参考块相关联的第二运动矢量估计第一运动矢量,所述参考块与所述从属视点中的当前图片的所述编码块相对应,

将所述第一运动矢量添加到所述列表中,

通过运动补偿预测从所述列表中确定要用于重构所述从属视点的当前图片中的编码块的至少一个候选者,以及

生成指定所述至少一个候选者的索引信息;以及

将所述索引信息插入所述数据流,

获得关于所述从属视点中的当前图片的所述编码块的视差矢量,其中,所述视差矢量表示所述从属视点的当前图片与所述参考视点的当前图片之间的视差,和基于所述视差矢量识别所述参考视点的当前图片中的所述参考块,

其中,获得所述视差矢量包括:

获得与所述参考视点的当前图片相关联的运动数据;

将与所述参考视点的当前图片相关联的所述运动数据应用于所述参考视点的先前图片的深度图估计,以生成所述参考视点的当前图片的第一估计深度图;

基于所述参考视点的当前图片的所述第一估计深度图,得出所述从属视点的当前图片的第二估计深度图;以及

基于所述从属视点的当前图片的所述第二估计深度图,确定所述视差矢量。

14. 一种用于重构被编码到数据流中的多视点信号的方法,

处理与所述多视点信号的从属视点中的当前图片的编码块相关联的多个运动矢量候选者的列表,包括:

基于与所述多视点信号的参考视点的当前图片中的参考块相关联的第二运动矢量估计第一运动矢量,所述参考块与所述从属视点中的当前图片的所述编码块相对应,

将所述第一运动矢量添加到所述列表中,以及

从所述数据流中提取索引信息,所述索引信息从所述列表中指定至少一个候选者用于运动补偿预测;以及

通过基于所述至少一个候选者执行所述运动补偿预测来重构所述从属视点中的当前图片的所述编码块,

获得关于所述从属视点中的当前图片的所述编码块的视差矢量,其中,所述视差矢量表示所述从属视点的当前图片与所述参考视点的当前图片之间的视差,和基于所述视差矢量识别所述参考视点的当前图片中的所述参考块,

其中,获得所述视差矢量包括:

获得与所述参考视点的当前图片相关联的运动数据;

将与所述参考视点的当前图片相关联的所述运动数据应用于所述参考视点的先前图片的深度图估计,以生成所述参考视点的当前图片的第一估计深度图;

基于所述参考视点的当前图片的所述第一估计深度图,得出所述从属视点的当前图片的第二估计深度图;以及

基于所述从属视点的当前图片的所述第二估计深度图,确定所述视差矢量。

15. 一种非暂时性机器可读介质,存储有包括数据流的数据,所述数据包括将多视点信号编码在其中的所述数据流,其中,当所述数据在由机器读取时,使所述机器执行多个操作,所述多个操作包括:

处理与所述多视点信号的从属视点中的当前图片的编码块相关联的多个运动矢量候选者的列表,

基于与所述多视点信号的参考视点的当前图片中的参考块相关联的第二运动矢量估计第一运动矢量,所述参考块与所述从属视点中的当前图片的所述编码块相对应,

将所述第一运动矢量添加到所述列表中,

通过运动补偿预测从所述列表中确定要用于重构所述从属视点的当前图片中的编码

块的至少一个候选者,以及

生成指定所述至少一个候选者的索引信息;以及

将所述索引信息插入所述数据流,

获得关于所述从属视点中的当前图片的所述编码块的视差矢量,其中,所述视差矢量表示所述从属视点的当前图片与所述参考视点的当前图片之间的视差,和基于所述视差矢量识别所述参考视点的当前图片中的所述参考块,

其中,获得所述视差矢量包括:

获得与所述参考视点的当前图片相关联的运动数据;

将与所述参考视点的当前图片相关联的所述运动数据应用于所述参考视点的先前图片的深度图估计,以生成所述参考视点的当前图片的第一估计深度图;

基于所述参考视点的当前图片的所述第一估计深度图,得出所述从属视点的当前图片的第二估计深度图;以及

基于所述从属视点的当前图片的所述第二估计深度图,确定所述视差矢量。

## 用于将多视点信号编码到多视点数据流中的装置及方法

[0001] 本申请为国际申请日为2012年11月9日、国际申请号为PCT/EP2012/072299、发明名称为“使用深度图估计和更新的高效多视点编码”的中国国家阶段申请的分案申请，该中国国家阶段申请的进入国家阶段日为2014年7月10日、申请号为201280066739.1、发明名称为“使用深度图估计和更新的高效多视点编码”。

### 技术领域

[0002] 本发明涉及根据多视点编解码器的多视点编码。

### 背景技术

[0003] 在多视点视频编码中，视频场景的两个或两个以上视点（其同时被多个照相机捕获）被编码在单个的比特流中。多视点视频编码的主要目的是通过提供3-d观看印象为终端用户提供先进的多媒体体验。如果两个视点被编码，则两个重构的视频序列可以被显示在（利用眼镜）传统的立体显示器上。然而，需要使用传统立体显示器的眼镜对用户来说经常是令人厌烦的。能够在没有眼镜的情况下有高质量立体观看印象是当前在研发方面的一个重要主题。用于这些自由立体显示器的有前景的技术基于柱状透镜系统。原则上，圆柱形透镜阵列以视频场景的多个视点同时显示的方式被安装在传统显示器上。每个视点在小圆锥中显示，使得用户的每只眼睛都可以看见不同的图像。这种效果在没有特殊眼镜的情况下产生了立体印象。然而，这些自由立体显示器通常需要相同视频场景的10-30个视点（如果进一步改善技术，则会需要甚至更多的视点）。超过2个视点也可以用于为用户提供交互式地选择视频场景的视角的可能性。但是视频场景的多个视点的编码与传统单视点（2-d）视频相比大大增加了所需比特率。通常，所需比特率利用编码视点的数量增加了近似线性的路径。用于减少自由立体显示器的传输数据的量的构思由仅传输少量视点（可能2-5个视点），但额外传输所谓的深度图组成，该深度图表示一个或多个视点的图像样本的深度（真实世界的物体到照相机的距离）。考虑到少量具有相对应深度图的编码视点，可以通过合适的渲染技术在接收器侧创建高质量中间视点（存在于编码视点之间的虚拟视点）以及在一定程度上至照相机阵列的一个或两个端部的额外的视点。

[0004] 对于立体视频编码和一般的多视点视频编码（具有或不具有深度图）这两者来说，重要的是利用不同视点之间的相互依赖性。由于所有视点都表示相同的视频场景（从稍微不同的角度来说），故在多个视点之间存在大量的相互依赖性。用于设计高效多视点视频编码系统的目的有效地利用了这些相互依赖性。在针对多视点视频编码的传统途径中，例如如与ITU-T Rec.H.264|ISO/IEC 14496-10的多视点视频编码（MVC）扩展一样，只有利用视点相互依赖性的技术才是来自已编码的视点的图像样本的视差补偿预测，其概念上类似于在传统2-d视频编码中使用的运动补偿预测。然而，通常只有从已编码的视点中预测小的图像样本子集，因为时间上的运动补偿预测常常更有效（两个时间上连续的图像之间的相似度在同一时间点大于相邻视点之间的相似度）。为了进一步改善多视点视频编码的有效性，需要将高效的运动补偿预测与视点间预测技术相组合。一种可能是重新使用在一个视点中

编码的运动数据以预测其他视点的运动数据。由于所有视点都表示相同的视频场景,故在一个视点中的运动基于真实世界的场景的几何形状连接至其他视点中的运动,其可以通过深度图和一些照相机参数来表示。

[0005] 在最新的图像和视频编码中,图片或用于图片的样本阵列的特定集合通常被分解成区块,该区块与特定编码参数相关联。图片通常由多个样本阵列(亮度和色度)组成。另外,图片还可以与额外的辅助样本阵列相关联,该额外的辅助样本阵列例如可以指定透明度信息或深度图。每个图片或样本阵列通常被分解成区块。区块(或样本阵列的相对应的区块)通过图片间预测或图片内预测来进行预测。区块可以具有不同尺寸并且可以是方形的或矩形的。将图片划分成区块可以通过语法来固定或可以(至少部分)在比特流内部用信号传递。经常传输针对预定尺寸的区块的信号细分的语法元素。这些语法元素可以指定区块是否被细分成以及如何被细分成更小的区块并且是否是相关联的编码参数(例如,出于预测目的)。对于区块的所有样本(样本阵列的相对应区块)来说,以某一方式指定相关联的编码参数的解码。在该示例中,使用相同组的预测参数来预测区块中的所有样本,预测参数诸如参考索引(识别这组已经编码的图片中的参考图片)、运动参数(指定参考图片和当前图片之间的区块运动的度量)、用于指定内插滤波器,帧内预测模式等的参数。运动参数可以是通过具有水平和垂直分量的位移矢量或通过更高阶运动参数(诸如由六个分量组成的仿射运动参数)来表示。还可以不止一组特定预测参数(诸如参考索引和运动参数)与单个区块相关联。在此情况下,对于这些特定预测参数的每一组,生成针对区块(或样本阵列的相对应区块)的单个的中间预测信号,并通过包括叠加中间预测信号的组合来建立最终预测信号。相对应的加权参数和同样潜在的恒定偏移量(其被添加到加权和中)对于图片或参考图片或一组参考图片来说,可以是固定的或其可以包括在用于相对应区块的预测参数的集合中。原始区块(或样本阵列的相对应的区块)及其预测信号(也被称为残余信号)之间的差通常被变换和量化。通常,二维变换适用于残余信号(或残余区域的相对应的样本阵列)。针对变换编码,区块(或样本阵列的相对应区块)在应用变换之前可以被进一步拆分,预测参数的特定集合已经被用于该区块。变换区块可以等于或小于用于预测的区块。同样可以变换区块包括用于预测的一个以上的区块。不同的变换区块可以具有不同尺寸并且变换区块可以表示方形或矩形区块。在变换之后,由此产生的变换系数被量化并获得所谓的变换系数水平。对变换系数水平和预测参数以及细分信息(如果存在)进行熵编码。

[0006] 最新技术的多视点视频编码以简单的方式扩展2d视频编码技术。从概念上讲,并行编码(或解码)对应于不同视点的两个或两个以上视频序列。或更具体地,对于每个访问点(或时间点),与不同视点相对应的图片按给定的视点顺序进行编码。MVC比特流一直包含基础视点,其可以在不参照任何其他视点的情况下进行解码。这利用基本2-d视频编码标准/方案确保了向后兼容性。比特流通常以这样的方式构造,该方式通过通过丢弃整个比特流的一些包可以以简单的方式提取与基础视点相对应的子比特流(以及另外与编码视点的特定子集相对应的子比特流)。为了利用视点之间的相关性,当前访问单元的已经编码视点的图片可以用于当前视点的区块的预测。该预测经常被称为视差补偿预测或视点间预测。其基本上等同于传统的2-d视频编码中的运动补偿预测,不同之处仅在于参考图片表示当前访问单元内部的不同视点的图片(即,在相同的时间点)而不表示不同时间点的相同视点的图片。为了将视点间预测整合到基本的2-d视频编码方案的设计中,针对每个图片,构造

一个或多个参考图片列表。对于基础视点(可独立解码的视点)来说,仅将传统的时间参考图片插入到参考图片列表中。然而,对于所有的其他视点,视点间参考图片可以被插入到除时间参考图片之外(或代替其)的参考图片列表中。那些图片被插入到参考图片列表中通过视频编码标准/方案来确定和/或在比特流内部(例如,在参数集和/或片报头中)用信号传递。然后通过编码(或推断)参考图片索引来用信号传递针对当前视点的特定区块所选择的时间还是视点间参考图片。即,视点间参考图片以与传统的时间参考图片完全相同的方式使用;仅构造略微扩展的参考图片列表。

[0007] 多视点视频编码的最新技术是ITU-T Rec.H.264|ISO/IEC JTC 1[1][2]的多视点视频编码(MVC)扩展。MVC是朝向多视点视频编码的ITU-T Rec.H.264|ISO/IEC JTC 1的简单扩展。与高层语法的一些扩展相比,只有已经添加的工具才是如上所述的视差补偿预测。然而,应注意,视差补偿预测通常仅被用于小部分的区块。除了由于场景内部的运动而被覆盖或被揭开的区域之外,时间运动补偿预测通常提供比视差补偿预测更好的预测信号,特别是在当前图片和参考图片之间的时间距离较小时。如果运动补偿时间预测可以与合适的视点间预测技术组合,则可以改善整体编码效率。可伸缩视频编码存在概念上类似的问题,其中具有不同分辨率或逼真度的相同视频序列的两个表示被编码在单一比特流中。对于增强层,原则上存在预测样本区块的两种可能性(如果我们忽略空间帧内预测),使用来自自己编码的增强层图片的运动补偿时间预测或来自更低层的层间预测。在可伸缩视频编码(SVC)扩展[3]中,传统的运动补偿时间预测已经与运动参数的层间预测相组合。对于增强层区块来说,其具有重新使用协同定位基础层区块的运动数据,而将其应用到增强层(即,使用具有基础层运动数据的增强层参考图片)的可能。以这种方式,层内部的时间运动补偿预测有效地与运动数据的层间预测相组合。隐藏于该技术的总体思想是可伸缩比特流中的所有的层示出了相同的内容,并因此每个层内部的运动是相同的。这不一定意味着由于以下效果而使用于一个层的最佳运动参数也是下一个层的最佳运动参数:(1)参考图片的量化修改了样本值并且由于不同层以不同方式量化,故给出最小失真的运动参数因不同层而不同。(2)由于层以不同比特率进行编码,故运动参数的特定集合通常对应于速率和失真之间的不同折衷。并且在率失真优化编码(其例如通过最小化失真D和相关联的速率R的拉格朗日函数 $D+\lambda R$ 来实现)中,不同运动参数从率失真的意义上来说针对不同层可以是最佳的(由 $\lambda$ 以及相关联失真或速率给出的操作点可以不同)。尽管如此,基础层和增强层中的(最佳)运动参数通常是类似的。并且重新使用基础层(并因此与小速率R相关联)的运动参数的模式通常很有可能导致比独立于基础层的最佳模式更少的总成本( $D+\lambda R$ )。或换句话说,通过选择具有基础层运动数据的模式代替具有最佳增强运动数据的模式而相关联的失真增加 $\Delta D$ 很可能小于与速率减小相关联的成本( $\Delta D+\lambda \Delta R < 0$ )。

[0008] 概念上,还可以在多视点视频编码中使用针对SVC的类似构思。多台照相机从不同角度捕获相同视频场景。然而,如果真实世界的物体在场景中移动,则不同的捕获视点中的运动参数不是独立的。但与物体的位置在所有层(层只表示相同捕获视频的不同分辨率或不同质量)中都相同的可伸缩编码相反,投影运动的相互关系更复杂并取决于几个照相机参数以及真实世界场景中的3-d关系。但如果给定所有相关照相机参数(诸如焦点长度、照相机的距离和照相机的光学轴线的方向)以及投影物体点(深度图)的距离,则特定视点内部的运动可以基于另一个视点的运动来推导。总体来说,对于编码视频序列或视点,我们不

需要知道物体点的精确运动；相反，诸如样本的区块的运动矢量简单参数是足够的。本着这样的精神，不同视点之间的运动参数的关系还可以在一定程度上被简化。

[0009] 然而，有利地，选择编码多视点信号时的编码顺序使得传达各个视点的纹理的图片在对应深度图之前被编码以便能够在编码/解码深度图时有效地利用从编码/解码图片获知的特征。甚至换句话说，去除深度图和相关联图片之间的冗余在从图片引导至深度图的编码顺序的情况下结果更有效，反之则不成立。然而，在解码器解码不同视点的图片时，服从这个编码顺序会导致在解码侧缺乏可用的深度图信息，因为其深度图仍未被重构。不利的是，参考视点的编码参数不能被有效地利用。在不存在视点的深度图的多视点数据流的情况下，情况更加严重。

## 发明内容

[0010] 本发明的目的是提供一种更高效的多视点编码的构思。

[0011] 该目的通过未决独立权利要求的主题来实现。

[0012] 根据本发明的第一方面，利用一种思想，据此，由于其传输无论如何都不是预期的，或由于纹理/图片及其深度图之间的优选编码顺序，或由于在传输或解码期间预期丢弃来自比特流的深度数据造成的参考视点的当前图片的深度图的缺失可以被充分解决以便通过估计参考视点和从属视点的图片的深度图和使用在多视点数据流内用信号传递的运动和/或视差数据更新所述深度图来减少视点间冗余。具体地，几乎所有多视点数据流都具有其中定义的随机访问点，即，与多视点信号的视点的图片对应的时间点，在没有时间预测和事先编码图片的其他相关性的情况下，而就参考视点被关注而言仅仅使用帧内预测，和就从属视点被关注而言使用帧内预测以及基于视差的预测来编码所述视点。因此，在视点间预测的多视点数据流内用信号传递的视差数据被用来初始化从属视点的深度图估计，并且该主要深度图估计在多视点编码的进一步进展期间使用在多视点数据流内用信号传递的运动数据和/或视差数据来连续更新。由此获得的不断更新的深度图估计允许以比在没有访问该深度图估计的情况下更高效的方式来执行视点间冗余减少的各种相关方法。

[0013] 根据另一个方面，利用以下发现：与从属视点的图片的区块的运动预测因子候选者的放大列表相关联的开销与由运动矢量候选者的添加导致的运动矢量预测质量的增益相比相对较低，其从视差补偿的意义上来说从参考视点的协同定位区块确定。这两个区块之间的视差可以或不可以使用第一方面来确定。

[0014] 根据本发明一方面，提供一种用于将多视点信号编码到多视点数据流中的装置，包括从属视点编码器，从属视点编码器被配置为：针对多视点信号的从属视点中的当前图片的至少一个区块，通过以下各项来推导运动矢量预测因子候选者列表：经由与多视点信号的先前编码部分相关联的运动和视差矢量来确定用于至少一个区块的视差矢量，视差矢量表示在从属视点的当前图片与从属视点的当前区块上的多视点信号的参考视点的当前图片之间的视差；使用所确定的视差矢量来确定参考视点的当前图片内的区块；将运动矢量添加到取决于与所确定的参考视点的图片的区块相关联的运动矢量的运动矢量预测因子候选者列表中；针对从属视点的当前图片的至少一个区块，将指定运动矢量预测因子候选者列表的一个运动矢量预测因子候选者的索引信息插入到多视点数据流中；以及通过使用取决于所指定的运动矢量候选者的运动矢量执行从属视点的当前图片的至少一个区块

的运动补偿预测来编码从属视点的当前图片的至少一个区块。

[0015] 另一方面,提供一种用于将多视点信号编码到多视点数据流中的装置,包括:参考视点编码器,被配置为随着经由多视点数据流传输用于参考视点的运动数据来基于用于参考视点的运动数据使用运动补偿预测来编码多视点信号的参考视点的当前图片;以及深度估计器,被配置为通过以下各项来估计从属视点的当前图片的深度图:通过将用于参考视点的运动数据应用到参考视点的先前图片的深度图估计上来生成参考视点的当前图片的深度图估计;以及将所更新的参考视点的当前图片的深度图估计变形到从属视点中以获得从属视点的当前图片的深度图估计;从属视点编码器,被配置为使用深度图估计将从属视点的当前图片编码到多视点数据流的从属视点部分。

[0016] 根据本发明的一方面,提供了一种用于将多视点信号编码到多视点数据流中的方法,包括:针对多视点信号的从属视点中的当前图片的至少一个区块,通过以下各项推导运动矢量预测因子候选者列表:经由与多视点信号的先前编码部分相关联的运动和视差矢量来确定用于至少一个区块的视差矢量,视差矢量表示从属视点的当前图片与在从属视点的当前区块上的多视点信号的参考视点的当前图片之间的视差;使用所确定的视差矢量来确定参考视点的当前图片内的区块;将运动矢量添加到取决于与所确定的参考视点的图片的区块相关联的运动矢量的运动矢量预测因子候选者列表中,以及针对从属视点的当前图片的至少一个区块,将指定运动矢量预测因子候选者列表的一个运动矢量预测因子候选者的索引信息插入到多视点数据流中;以及通过使用取决于所指定的运动矢量候选者的运动矢量执行从属视点的当前图片的至少一个区块的运动补偿预测来编码从属视点的当前图片的至少一个区块。

[0017] 根据本发明一方面,提供了一种用于将多视点信号编码到多视点数据流中的方法,包括:随着经由多视点数据流传输用于参考视点的运动数据,基于用于参考视点的运动数据使用运动补偿预测来编码多视点信号的参考视点的当前图片;以及通过以下各项来估计从属视点的当前图片的深度图:

[0018] 通过将用于参考视点的运动数据应用到参考视点的先前图片的深度图估计上来生成参考视点的当前图片的深度图估计;以及

[0019] 将所更新的参考视点的当前图片的深度图估计变形到从属视点中以获得从属视点的当前图片的深度图估计;

[0020] 使用深度图估计将从属视点的当前图片编码到多视点数据流的从属视点部分中。

[0021] 根据本发明一方面,提供一种计算机程序,具有程序代码,当在计算机上运行程序代码时,程序代码用于执行上述方法。

[0022] 本发明的优选的实施是从属权利要求的主题。

## 附图说明

[0023] 下面参照附图更详细地阐述本发明的优选实施方式,在附图中:

[0024] 图1示出了根据实施方式的用于重构多视点信号的装置的框图;

[0025] 图2示出了根据实施方式的用于编码适合于图1的装置的多视点信号的装置的框图;

[0026] 图3示出了不同视点和时间实例的图片中的投影物体点、时间运动矢量和视差矢

量之间的总体关系；

[0027] 图4示出了(利用当前区块内部的特定样本位置)用于推导给出了当前图片的参考视点和深度图估计中的运动的当前区块的运动矢量的基本过程；

[0028] 图5示出了映射针对一个视点至另一个视点给出的深度图的基本过程：(左)给出了用于视点的深度图，其中灰色区域表示背景，白色区域表示前景物体；(中)通过将样本移位对应于深度值的视差矢量并且保持一个以上的样本被投影至此的位置的前景物体而获得的转换深度图，黑色区域表示样本不投影至此的去遮挡(disoccluded)区域；(右)在利用背景的深度值填充去遮挡区域之后的转换深度图；

[0029] 图6示出了(利用视差矢量)用于随机访问单元的深度图的生成；

[0030] 图7示出了利用在基础视点中编码的运动参数进行的估计深度图的时间预测；以及

[0031] 图8示出了利用实际编码运动和视差矢量进行的深度图更新。

## 具体实施方式

[0032] 图1示出了用于重构被编码到多视点数据流14中的多视点信号12的装置10的实施方式。装置10包括用于多视点数据流14的输入端16和分别用于参考视点信号20和从属视点信号22的两个输出端18a,18b。

[0033] 此外，装置10包括连接在输入端16与输出端18a之间的参考视点重构器24以及连接在输入端16与输出端18b之间的从属视点重构器26。装置10的深度图估计器28连接在参考视点重构器24的参数输出端与从属视点重构器26的参考输入端之间，并连接至从属视点重构器26的参数输出端。

[0034] 如下面将更详细的描述，图1的装置或解码器10通过遵循根据哪个参考信号20在从属视点22之前被处理的编码/解码顺序重构来自多视点数据流14的多视点信号12。如图1中所示，多视点信号12不仅可以表示与相应的视点20和22相关联的视点或来自不同视点方向的一个共同场景的空间采样，而且还可以表示如在图1中通过沿时间轴线30示出的三个时间点T-1、T和T+1而示例性地示出的该场景的时间采样。对于每个时间点，视点20和22均包括图片32t<sub>1</sub>和32t<sub>2</sub>，其中，图片32t<sub>1,2</sub>均表示相应的纹理图(texture map)。

[0035] 应注意，图1假设了两个视点20和21具有按时间排列的图片32t<sub>1,2</sub>。然而，视点20与视点22之间的时间分辨率可以不同。自然地，这同样适用于深度图和图片的空间分辨率。

[0036] 此外，解码器10被配置为依次及时地处理多视点信号12。更确切地说，解码器10被配置为在继续处理后续时间实例T的图片和深度图之前重构某个时间实例(诸如T-1)的图片32t<sub>1,2</sub>。在这点上，应注意，多视点信号12的时间实例之间的时间编码顺序可以分别等于图片和深度图的呈现时间顺序，或可以与此不同。

[0037] 参考视点重构器24被配置为重构来自多视点数据流14的参考视点部分36参考视点20，同时从属视点重构器26被配置为基于多视点数据流14的从属视点部分38来重构从属视点22。实际上，参考视点重构器24和从属视点重构器26可以被配置为以类似的方式操作。例如，参考重构器24和从属视点重构器26可以逐区块进行操作。两者例如可以分别被配置作为混合式视频解码器。参考视点重构器24例如通过向区块40分配可用编码模式的相应编码模式中的一个来重构当前时间点T的图片32t<sub>1</sub>，该图片被细分成区块40。将图片32t<sub>1</sub>细分

成区块可以在默认情况下被预定义或可以在多视点数据流14内用信号传递。细分可以以规则的方式将图片32t<sub>1</sub>细分成相同尺寸的区块或不同尺寸的区块。甚至此外,可以进行多树细分,使得区块40的区块尺寸可以局部适配于图片内容。可用的编码模式可根据通过从按图片32t<sub>1</sub>的区块之间定义的解码顺序在当前区块之前已经重构的区块的已经重构样本来预测哪个参考视点重构器24填充相应的区块40而包括一个或多个帧内预测模式;以及根据哪个参考视点重构器24通过使用诸如运动矢量、参考图片索引等的运动数据的运动补偿和/或预测来重构相应的区块而包括一个或多个帧间预测模式。例如,出于说明目的,两个区块被示例性地示出以通过帧间预测来重构。用于这些帧间预测区块的运动数据42可以包括被参考视点重构器24用来复制通过运动参考索引所索引的参考图片32t<sub>1</sub>的重构版本的相应部分的运动矢量,该运动参考索引也被运动数据42所包括。运动数据42被多视点数据流14的参考视点部分36包括。

[0038] 从属视点重构器26的操作与具有从属视点重构器26的参考视点重构器24完全相同,然而从属视点重构器26被配置为从从属视点部分38重构从属视点22。因此,在重构当前时间点T的当前图片32t<sub>2</sub>时,从属视点重构器26还可以使用区块式处理,其利用细分可以在多视点数据流14内用信号传递或被固定的区块50。可替换地,如下面更详细描述的基于深度图的视点间预测细分区块50可以被从属视点重构器26使用以便从细分视点20的区块40推导细分视点22的区块50。就编码模式被关注而言,从属视点重构器26可以支撑编码模式,如针对参考视点重构器24描述的一样。因此,说明性地,两个区块50被示例性地示出以使用运动数据54分别进行帧间预测,以便分别从事先重构图片32t<sub>2</sub>的重构版本的相应部分适当地复制。同时,该运动数据58表示视点22的当前图片或当前时间实例的运动数据。然而,除了这些编码模式之外,从属视点重构器26具有支持一个或多个视点间预测模式使用视差补偿预测的能力以便从相同时间实例的视点20的部分复制相应区块,但发生空间移位,如一些视差数据所定义。在图1中,示例性地示出了图片32t<sub>2</sub>中的一个视差预测区块与相对应视差数据60。视差数据60例如可以包括沿视点20和22之间的视点偏移方向的视差矢量或至少视差分量,以及任选指示参考视点的视点索引,从属视点22的相应区块50取决于所述参考视点,该索引就如图1中示例性所示的两个以上视点共同存在而言可以是有利的。

[0039] 即,参考视点重构器24和从属视点重构器26以一种方式操作以便沿时间轴线30并在视点间方向上尽可能地减少视点20与22之间的冗余。例如,对于预测诸如运动数据和视差数据以及编码模式的边带信息(side information)以及上面提及的细分信息这也是可靠的。所有的信息示出了彼此之间在时间方向上以及两个视点之间的冗余。

[0040] 然而,如果从属视点重构器26访问了当前解码图片32t<sub>2</sub>的深度图,则从属视点重构器26可以更高效地利用视点20和22之间的冗余。因此,深度估计器28被配置为以下面更详细描述的方式提供深度图估计64作为用于当前时间点T的当前图片32t<sub>2</sub>的深度图的估计,并且从属视点重构器26被配置为使用该深度图估计64重构来自多视点数据流14的从属视点部分38的从属视点22的当前时间点的当前图片32t<sub>2</sub>。例如,即将具有深度图估计64的从属视点重构器26能够基于当前视点22的深度图估计64和用于参考视点20的当前图片的运动数据42预测从属视点22的当前图片的运动数据54并且使基于预测到运动数据利用运动补偿预测来重构从属视点22的当前图片。例如,当前视点重构器24可以被配置为在预测

运动数据54时,使用从属视点22的深度数据估计64来定位参考视点20的当前图片中的相应位置并使用相应位置处的参考视点20的当前图片的运动数据42来预测从属视点22的当前图片的运动数据54。在以下描述中,在下面更详细地描述如何使用深度数据估计64的完成空间查找的可能方式。具体地,在以下描述中,更详细地促进运动数据42形成针对运动数据54的良好预测的事实。自然地,提纯数据以便用信号传递针对运动数据54的预测残差可以被从属视点部分38包括。具体地,如下面更详细阐述,从属视点重构器26可以被配置为应用从用于从属视点22的当前时间点的图片32t<sub>2</sub>的当前区块50内的一个或多个预定样本位置的深度数据估计64得到的视差矢量,并且使用这些视差矢量以便使用包含一个或多个变形(wrap)位置的一个或多个区块40的运动数据42作为针对当前区块50的运动数据54的预测因子来定位视点20的相同时间点的图片32中的相应的或变形的位置。就当前区块50内的一个以上的此参考样本位置而言,由目标区块或区块40的一个或多个参考运动数据产生的平均值或中间值可以被用作预测因子。

[0041] 此外,从属视点重构器26可被配置为基于从属视点22的深度数据估计64来预测从属视点22的针对当前图片的视差数据60,并且基于预测的当前视差数据利用视差补偿预测来重构从属视点22的当前图片。再者,提纯可以在多视点数据流14的从属视点部分38内用信号传递并被从属视点重构器26用来提纯所预测的当前视差数据。此外,如上文所描述的,理论上,区块50的视差数据60也可以相同的方式进行预测。如下面更详细阐述,从属视点重构器26可被配置为通过将深度数据转换成视差矢量并使用这些视差矢量分别直接作为用于视差数据60和62内的视差矢量的预测因子基于当前视点的深度数据估计64来预测视差数据60和62。

[0042] 自然地,从属视点重构器26可支持刚才提及的可能性的任何组合以便使用深度数据估计以减少视点20和22之间的视点间冗余。

[0043] 为了得到刚才提及的深度数据估计64,深度估计器28作用如下。

[0044] 具体地,根据图1的实施方式,深度估计器28确保每个图片32t<sub>1,2</sub>具有与此相关联的深度图估计64,在一连串更新中其连续地从彼此推导出。如下面更详细阐述,深度估计器28被配置为在视点20与22之间以乒乓球的方式不断更新深度图估计64,主要目的是提供具有此深度图估计64的从属视点22的每个图片32t<sub>2</sub>以便用作用于上述改善视点间冗余减少的基础。

[0045] 初步地,假设深度估计器28已访问用于参考视点20的一个或多个先前图片32t<sub>1</sub>的估计这样的深度估计,诸如时间实例T-1。在下面进行进一步描述深度估计器28如何获得访问用于参考视点20的先前的解码图片32t<sub>1</sub>的这样的深度图估计74的方式。然而,应注意,这样的深度图数据可用于在所谓的随机访问单元内的参考视点20的第一图片32t<sub>1</sub>的多视点数据流14内间歇性地明确地进行信号传递,即在不参照信号12的任何先前部分的情况下可解码的多组图片32t<sub>1</sub>。为了说明该可能性,虚线连接深度估计器28与输入端16。在以下描述中,呈现了一种可能,其中这种起始深度图的额外传输不是必须的。相反,用于随机访问单元内按编码顺序的从属视点22的第一图片32t<sub>2</sub>的数据流部分38内的视差数据被用来构造随机访问单元内的按编码顺序的参考视点20的第一图片32t<sub>1</sub>的起始深度图。

[0046] 具体地,深度估计器28被配置为例如通过将针对当前时间实例T的参考视点20的当前图片32t<sub>1</sub>的运动数据42应用到时间点T-1时的参考视点20的任何先前图片32t<sub>1</sub>的深度

图估计74上来生成从属视点22的当前图片32t<sub>2</sub>的深度图64。正如上面已经提到的,参考视点重构器24基于在用于参考视点20的多视点数据流14内用信号传递的运动数据42使用运动补偿预测来重构参考视点20的当前图片32t<sub>1</sub>。深度估计器28访问该运动数据42并将该运动数据42用于一连串更新的所提及更新之一,即从先前时间点T-1的参考图片32t<sub>1</sub>的深度图估计74至当前时间点T的当前图片32t<sub>1</sub>的深度图估计64的转变。下面将更详细地阐述这可以如何执行的方式。初步地,应该充分注意,将运动数据42的71应用到先前时间实例T-1的深度图74上可意味着协同定位区块72,即,当前图片32t<sub>1</sub>的深度图估计64内的协同定位至区块40的部分(对此运动数据42已经在流部分36中用信号传递),利用所参考的深度图估计(即,由等于运动数据42的运动数据42'指出的所参考的深度图估计74内的部分处的先前时间实例T-1的图片32t<sub>1</sub>的深度图估计74)的内容来更新,即,从所述内容复制。剩余的孔可以通过插值法和/或外推法利用当前图片32t<sub>1</sub>的区块40之间的帧内编码区块提供的额外信息来填充。结果,深度图估计64已经更新(或通过从T-1转变至T而生成)。

[0047] 再者,深度估计器28只执行该更新/转变71以便实行下面进一步描述的一连串更新以充当用于推导相同时间点T的从属视点22的当前图片32t<sub>2</sub>的深度图估计64的依据。为了完成推导,深度估计器28将参考视点20的当前图片32t<sub>1</sub>的更新深度图估计64变形到从属视点22中以便获得从属视点22的当前图片32t<sub>2</sub>的深度图估计64。也就是说,因为运动数据42仅按区块粒度定义,故变形78产生的视点22的更新/转变71和由此产生的深度图估计64表示很粗略的深度估计,但如下所示,此粗略估计是充分的以便大大增加在执行视点间冗余减少时,的效率。

[0048] 虽然同样在下面进一步描述了关于变形76的可能细节,但简要地说,从属视点重构器26可以被配置为通过从当前图片32t<sub>1</sub>的深度图估计64推导视差矢量并将所推导的视差矢量应用到深度图估计64本身上来执行变形78,以便获得从属视点22的当前图片32t<sub>2</sub>的变形深度图估计64。

[0049] 因此,深度估计器28已为从属视点重构器26提供变形76的结果,即,视点22的当前时间点T的深度图估计64,从属视点重构器26能够使用该深度图估计64来执行上述视点间冗余减少,在下面更详细地阐述针对其的可能的实施。

[0050] 然而,深度估计器28继续更新(77)该深度图估计64以便获得用于参考视点22的当前图片32t<sub>2</sub>的更新深度图估计74并由此保持导致针对下一时间实例T+1的估计的一连串更新。因此,从属视点重构器26被配置为以与上文针对更新步骤71描述的类似的方式使用从属视点22的视差和/或运动数据54和60(至少针对运动数据54)来更新(77)当前时间实例T的从属视点22的当前图片32t<sub>2</sub>的深度图估计64。也就是说,从属视点重构器26使用流部分38内的时间实例T的图片32t<sub>2</sub>的视差/运动数据来重构该图片32t<sub>2</sub>。就视差数据60被关注而言,深度估计器28可以轻易将视差数据54内包含的视差矢量转换成深度值并基于这些深度值将更新的深度值分配给从属视点22的当前图片32t<sub>2</sub>的更新深度图估计79b的样本,其被协同定位至图片32t<sub>2</sub>中的相应视差预测区块50。可使用运动数据54以便将由等于运动数据54的运动数据指出的其部分处的从属视点22的参考的先前时间实例T-1的图片32t<sub>2</sub>的深度图估计74的内容复制到当前图片32t<sub>2</sub>的更新深度图估计74内的协同定位至区块50的部分中,该运动数据42在流部分36中已经针对所述区块50进行信号传递。剩余的孔可以通过插值法和/或外推法利用当前图片32t<sub>1</sub>的区块40之间的帧内编码区块提供的额外信息来填

充。结果,当前图片 $32t_2$ 的更新深度图估计74已经更新完成(或通过从T-1转变至T而生成)。下面将进一步描述在随机访问单元的开始可以如何推导从属视点22的所参考的先前时间实例T-1的图片 $32t_2$ 的深度图估计74的可能性。然而,在此随机访问单元的开始的上述提及的视点20的可能明确传输用于视点22的深度图可以被变形到视点22中以可替换地获得从属视点22的所参考的先前时间实例T-1的图片 $32t_2$ 的深度图估计74。

[0051] 为了减少区块效应,更新71和77可通过使用加权函数来执行,从而降低区块边界处各个区块的更新的影响。

[0052] 也就是说,基于如通过变形76获得的深度图估计64,从属视点重构器26基于用于被多视点数据流14的从属视点部分38包括的从属视点22的视差和/或运动数据54和60使用视差和/或运动补偿预测来重构从属视点22的当前图片 $32t_2$ ,并且通过这样做,从属视点重构器26为深度估计器28提供视差和/或运动数据54,60,然后被深度估计器68用来执行更新77。

[0053] 在该更新77之后,深度估计器28能够将从属视点22的当前图片 $32t_2$ 的更新深度图估计74变形78回到参考视点20以便获得时间实例T的参考视点20的当前图片 $32t_1$ 的更新深度图估计74,这然后可以充当转变/更新79至下一时间实例T+1的依据/参考,以此类推。

[0054] 从那时起,深度估计器28只迭代重复过程71、76、77和78(其中,步骤79对应于步骤71)以便沿时间轴线30模仿深度图估计以便不断利用深度图估计64来支持从属视点重构器26。

[0055] 将在下面进行更详细的描述关于所有这些步骤71、76、77、78和79的另外的细节。所有这些另外的细节应分别适用于针对图1提出的描述。

[0056] 在描述关于上述构思的另外的细节之前,参照图2描述了符合图1的解码器的编码器的实施方式。图2示出了用于将多视点信号12编码成多视点数据流14的装置并且为此包括参考视点编码器80、从属视点编码器82和深度估计器84,其中编码器通常使用参考符号90示出。参考视点编码器80被配置为将多视点信号12的参考视点20编码成数据流14的参考视点部分36,同时从属视点编码器82负责将多视点信号12的从属视点22编码成多视点数据流14的从属视点部分38。参考视点编码器80和从属视点编码器82可以以向后预测方式操作并且深度估计器84可以被配置为以上面针对解码器10描述的方式通过使用可从参考视点编码器80和从属视点编码器82获得的相同信息来执行深度图估计及其连续更新。也就是说,深度估计器84被配置为通过将已经用于以运动补偿方式预测参考视点的当前图片的参考视点的运动数据42应用到参考视点20的先前图片 $32t_1$ 的深度图估计上并将由此获得的参考视点20的当前图片 $32t_1$ 的深度图估计64变形76到从属视点22中以便获得从属视点22的当前图片 $32t_2$ 的深度图估计64来生成从属视点22的当前图片 $32t_2$ 的深度图估计64。类似地,深度估计器84还执行后续更新步骤77和以下向后变形步骤78。为此,参考视点编码器80和从属视点编码器82可以并联连接在编码器90的端与输出端之间,同时深度估计器84可以连接在参考视点编码器80的参数输出与从属视点编码器82的参考输入端之间并连接至从属视点编码器82的参数输出端。参考视点编码器80的重构输出端可以连接至参考视点编码器80的预测参数输出(诸如内部预测区块的输出)。

[0057] 从属视点编码器82可以以上文参照图1所描述的方式使用深度图估计64来编码从属视点22的当前图片或当前时间点,即,在生成或没有生成针对相应运动或视差数据的预

测残余数据的情况下用于预测运动数据58或至少54,或预测视差数据60和62或至少60或至少这些项的一部分,以便形成从属视点部分38的一部分。

[0058] 在下文中,呈现出了更详细的实施方式,当与使用区块合并或诸如HEVC中的规则排列树根区块的多树区块划分的混合式编码类型组合时,所述实施方式尤其是有优势的。

[0059] 已经全部基于ITU-T Rec.H.264|ISO/IEC 14496-10的MVC扩展开发了针对采用参考视点的运动数据以高效编码另外的视点的最新的技术构思。ITU-T和ISO/IEC JTC 1/WG 11的新视频编码标准化项目(也被称为HEVC)展现出传统2-d视频编码技术方面非常有前景的改进。HEVC的现有工作草案与ITU-T Rec.H.264|ISO/IEC 14496-10相比提供了显著的编码增益。为了实现这些增益,与ITU-T Rec.H.264|ISO/IEC 14496-10相比扩展了几个构思。运动参数编码和运动补偿预测领域的主要改进包括以下各项:

[0060] • 尽管用于ITU-T Rec.H.264|ISO/IEC 14496-10中的运动补偿预测的区块尺寸在 $4 \times 4$ 至 $16 \times 16$ 亮度样本的范围内,但是在 $4 \times 4$ 至 $64 \times 64$ 亮度样本的范围内的HEVC中支持更大的各种区块尺寸。另外,基本编码单元不由固定宏区块和子宏区块给出,而是被自适应地选择。最大的编码单元通常为 $64 \times 64$ 亮度样本的区块,但最大区块尺寸实际上可以在比特流内侧用信号传递。将区块分割成子区块可以建立4个或4个以上级别的细分层次。

[0061] • 运动矢量不是通过使用固定运动矢量预测器进行编码。相反,存在运动矢量预测器候选者的列表,并且基于区块自适应选择这些预测因子之一。所选择的预测因子在比特流内侧用信号传递。

[0062] • ITU-T Rec.H.264|ISO/IEC 14496-10提供SKIP和DIRECT,运动参数(假设的次数、参考索引、运动矢量)对此完全源自于已经编码的信息,而无需编码任何额外的参数(残余信息除外)。HEVC提供所谓的合并模式。针对该模式,建立由空间和时间相邻区块的运动参数给出的运动参数候选者的列表。为以合并模式编码的区块选择的运动参数(包括假设的次数、参考索引、运动矢量)通过将索引传输至候选列表中而被用信号传递。

[0063] 以下描述将描述采用已经编码视点的运动数据以及当前视点的已经编码图片的视差数据来编码多视点视频编码中的当前视点的当前图片,该构思表示上述实施方式的可能的实施。此外,将更详细地解释由上述和以下实施方式产生的优势。通过采用已经编码的运动和视差信息来预测当前视点的时间运动(或视差),可以显著减小当前视点的运动数据速率,这会导致节省用于多视点视频序列编码的总比特率。所描述的构思提供直接推导区块(或一组普通的样本)的所有运动数据的可能性,在这种情况下,不针对区块传输进一步的运动信息。并且其还提供推导被添加到候选运动矢量预测因子的列表中的运动矢量预测因子的可能性。对于后者的可能性,针对区块传输至运动矢量预测因子的列表的索引以及运动矢量差,其指定用于运动补偿预测的最终运动矢量。在本发明的具体实施方式中,基于已经编码的运动和视差信息不但可以推导区块的运动参数,而且还可以推导区块的划分信息(其可以将区块分割成更小的区块并将分开的运动参数分配给每个子区块)。构思适用于一般的基于区块的混合式编码方法,而无需假设任何特定宏区块或子宏区块结构。不修改一般的基于区块的运动补偿,而只修改运动参数的编码,使得该构思可以被整合到复杂度增加极小的一般的基于区块的混合式视频编码方案中。它也可以直接扩展至更一般的构思,其中对唯一运动参数来说不是矩形区块而是其他组样本是相关联的。该构思在具有和没有额外深度图的情况下均适用于多视点编码。用于计算运动参数的视差信息可以基于编

码深度图(其基于编码视差矢量)来推导。

[0064] 以下描述将描述采用已经编码视点的运动数据以及当前视点的已经编码图片的视差数据来编码多视点视频编码中的当前视点的当前图片,该构思表示上述实施方式的可能的实施。此外,将更详细地解释由上述和以下实施方式产生的优势。通过采用已经编码的运动和视差信息来预测当前视点的时间运动(或视差),可以显著减小当前视点的运动数据速率,这会导致节省用于多视点视频序列编码的总比特率。本发明提供直接推导区块(或一组普通的样本)的所有运动数据的可能性,在这种情况下,不针对区块传输另外的运动信息。并且其还提供推导被添加到候选运动矢量预测因子的列表中的运动矢量预测因子的可能性。对于后者的可能性,针对区块传输至运动矢量预测因子的列表的索引以及运动矢量差,其指定用于运动补偿预测的最终运动矢量。在本发明的具体实施方式中,基于已经编码的运动和视差信息不但可以推导区块的运动参数,而且还可以推导区块的划分信息(其可以将区块分割成更小的区块并将分开的运动参数分配给每个子区块)。该构思适用于一般的基于区块的混合式编码方法,而无需假设任何特定宏区块或子宏区块结构。在不修改一般的基于区块的运动补偿,而只修改运动参数的编码情况下,使得该构思可以被整合到复杂度增加极小的一般的基于区块的混合式视频编码方案中。它也可以直接扩展至更一般的构思,其中对唯一运动参数来说不是矩形区块而是其他组样本是相关联的。该构思适用于使用额外深度图的多视点编码。用于计算运动参数的视差信息可以基于编码深度图来推导。

[0065] 与采用已经编码的视点的运动数据的传统技术相比,现在呈现的构思的一个优势是完全基于编码运动和视差/深度信息来推导运动/视差预测因子,而无需假设视差场的任何特定结构。在任何时候,假设视差场不可以通过对图像的不断平移或仿射参数进行很好的估算,相反,实际编码的视差信息用于访问已经编码的视点的运动。此外,假设宏区块的视差不类似于相邻区块的视差,该假设是不安全的。通过使用实际编码的深度/视差信息,该构思为物体边界处的区块提供了合适的视差估计。此外,由于没有做出当前区块的运动类似于相邻区块的运动的假设,因此提供了物体边界处的改进的运动参数预测因子。此外,该构思不要求视差校正的任何传输,此外,该构思不要求修改混合式视频编码设计的实际运动/视差补偿过程以便被构建成此。只改变运动参数和/或运动参数预测因子的推导,使得其可以在没有任何大修改的情况下被包括在传统的视频编码设计中并具有较小的复杂度。另外,应注意,在具有和没有深度图的情况下,该构思都适用于编码。深度图不需要被编码成比特流的一部分。相反,编码视差矢量可以用于推导视差。

[0066] 下文中描述的构思可以被分解成以下步骤:

[0067] • 推导当前视点的当前图片的深度/视差数据。

[0068] • 基于推导的深度/视差数据推导当前区块的候选运动或视差数据。

[0069] • 编码当前区块的运动或视差数据。

[0070] 在下文中,更详细地描述包括优选实施方式的这些步骤。针对基于区块的运动补偿利用平移运动矢量来描述所有步骤。然而,该构思也适用于更一般的方案,其中一组通用的样本(例如,区块的非矩形部分,或任何其他形状)与一组唯一的运动参数相关联,并且其还适用于编码方案,其中运动补偿使用更高阶的运动模型(例如,仿射运动模型或其他N参数运动模型)来执行。

[0071] 在描述构思的细节之前,我们简要描述对于上述的更通用的实施方式同样有效的优势和基本思想。在图3中示出了不同视点中的真实世界的物体点的投影和不同时间实例的投影之间的基本关系。假设我们知道视点内侧的真实运动以及视点之间的真实视差,给出对应的运动和视差矢量如下:

[0072] • 当前视点的运动矢量由当前视点的参考图片和当前视点的当前图片中的投影物体点的位置的差给出,  $MV_C(x_{C,t}) = x_{C,t-1} - x_{C,t}$

[0073] • 参考视点的运动矢量由参考视点的参考图片和参考视点的当前图片中的投影物体点的位置的差给出,  $MV_R(x_{R,t}) = x_{R,t-1} - x_{R,t}$

[0074] • 当前时间点的视差矢量由参考视点的当前图片和当前视点的当前图片中的投影物体点的位置的差给出,  $DV_t(x_{C,t}) = x_{R,t} - x_{C,t}$

[0075] • 参考时间点的视差矢量由参考视点的参考图片和当前视点的参考图片中的投影物体点的位置的差给出,  $DV_{t-1}(x_{C,t-1}) = x_{R,t-1} - x_{C,t-1}$

[0076] 因此,我们得到运动和视差矢量之间的以下关系:

[0077]  $MV_C(x_{C,t}) + DV_{t-1}(x_{C,t-1}) - MV_R(x_{R,t}) - DV_t(x_{C,t}) = 0$

[0078] 如果给出运动和视差矢量中的三个,则第四个矢量可以通过简单的加法来计算。具体地,可以根据下式来推导当前视点的时间运动矢量:

[0079]  $MV_C(x_{C,t}) = -MV_R(x_{R,t}) + DV_t(x_{C,t}) - DV_{t-1}(x_{C,t-1})$

[0080] 条件是给出参考图片中的相同物体点的运动矢量以及两个时间点的视差矢量。在多数情况下,给出了参考视点的运动矢量(或运动参数),因为该视点已经使用运动补偿预测进行编码。但通常不给出视差,其只可以被估计。然而,通过使用两个估计的值,最终结果的准确性可能很不准确且不适于推导足够准确的时间运动矢量预测因子。但总体说来,人们假设物体点的深度(距照相机的真实世界的物体点的距离)在对应于当前和参考图片的时间实例之间几乎是不变的(两个连续图片之间的至或自照相机的物体运动通常要比物体至照相机的距离小得多)。然后,视差同样几乎是不变的并且运动矢量之间的关系简化为:

[0081]  $MV_C(x_{C,t}) \approx MV_R(x_{R,t}) = MV_R(x_{C,t} + DV_t(x_{C,t}))$

[0082] 应注意,我们仍然需要估计当前访问单元(当时时间点)中的视差以便基于参考视点内侧的运动预测当前视点内侧的运动。然而,视差估计的准确性不那么重要,因为其只用于评估参考视点中的运动数据。运动补偿操作以及运动数据的编码基于多个样本的区块来进行,并且此外,相邻样本或区块的运动经常非常类似。尽管如此,视差的准确估计总体改善了运动参数的预测。所估计的视差矢量  $DV_t(x_{C,t})$  还可以被用作视差补偿预测的视差矢量(即,使用参考视点中的当前访问单元的图片作为参考图片),其可以表示运动的特殊模式和运动参数的基于视差的预测。

[0083] 候选运动或视差数据的推导

[0084] 在下文中,我们描述了诸如图1中的视点22的特定视点(其不是向后兼容基础视点)中的当前图片的给定区块的运动数据,给出了已经编码的参考视点或已经编码的参考视点集合的运动数据的基本推导(诸如图1中的20)。针对该描述,我们假设给出当前图片的深度数据的估计(诸如图1中的64)。稍后,我们描述了该深度估计可以如何推导并且所推导的运动数据可以如何用于高效编码当前视点。当前图片的深度数据64由像素式或区块式深度图给出。如果给出像素式深度图,则深度可以指定相关联图片的每个样本(或每个亮度样

本)的深度值。如果给出区块式深度图,则深度图指定相关联图片的样本(或亮度样本)的M×N区块的深度值。例如,指定可用于运动补偿的最小的可能区块尺寸(例如,4×4或8×8区块)的每个区块的深度值。从概念上讲,由深度图的样本给出的深度值指定真实世界的深度z的函数,该深度是相关联的真实世界的物体点(真实世界的物体点的投影是给定位置的图像样本)与照相机之间的距离:

$$[0085] \quad d = f_{dz}(z)$$

[0086] 深度值按特定的精度给出(此外,深度值经常被估计,因为实际深度通常是不已知的)。在多数情况下,深度值由整数给出。考虑了深度值和特定照相机参数(诸如焦距、照相机之间的距离、最小和最大深度值、或这些参数的函数),深度值d可以被转换成视差矢量 $v = [v_x, v_y]^T$ :

$$[0087] \quad v(x) = f_{vd}(d(x), x)$$

[0088] 其中 $f_{vd}$ 指定将样本位置 $x = [x, y]^T$ 处的深度值d映射至视差矢量的函数。尤其重要的设定为一维并行照相机配置,其特征在于以下特性:

[0089] • 照相机阵列的所有照相机都是相同类型的并具有相同的焦距

[0090] • 所有照相机的光轴都平行并位于相同的平面内侧

[0091] • 图像传感器的扫描线平行于包含光轴的该平面

[0092] 在这种情况下,视差矢量的垂直分量总是零, $v = [v, 0]^T$ 。每个真实世界的物体点在所有视点中都具有相同的垂直位置。其水平位置取决于物体点的深度。水平位置之间的差通过视差给出:

$$[0093] \quad v = f_{vd}(d)$$

[0094] 在重要的情况下,以获得视差v与深度值d之间的线性关系 $v = m_{vd} \cdot d + n_{vd}$ 的方式给出真实世界的深度z与深度值d之间的关系,

[0095] 其中, $m_{vd}$ 和 $n_{vd}$ 由照相机参数给出。深度值d通常以整数值的形式给出。并且针对内部计算,如果所获得的视差值是整数值,则其通常也是优选的。例如,视差v可以在相同的单元中表达出来,用于运动/视差补偿预测中的运动/视差矢量(例如,二分之一、四分之一或半分之一样本准确性)。在这种情况下,视差的整数值可以通过整数方程来获得

$$[0096] \quad v = (m_{vd}^* \cdot d + n_{vd}^*) \gg u_{vd}$$

[0097] 其中,“ $\gg$ ”指定向右的比特移位(在二进制补码运算中), $m_{vd}^*$ 和 $n_{vd}^*$ 分别是 $m_{vd}$ 和 $n_{vd}$ 的标度(和取整)版本。

[0098] 使用所描述的给定深度值和实际视差之间的基本关系,我们描述了使用运动信息的优选实施方式,其基于一个或多个参考视点中的已经编码的运动信息和给出的估计深度值来推导。

[0099] 方法1:切换运动/视差矢量预测

[0100] 在本发明的优选实施方式中,诸如针对模块24、26、80和82的基本的多视点视频编码方案包括以下参数被作为比特流21的一部分传输的模式:

[0101] • 指定参考图片的给定列表的特定(时间或视点间)参考图片的参考图片索引。如果参考图片的给定列表由单一元素组成,则该索引不被传输,但可以在解码器侧推断。参考图片包括时间和/或视点间参考图片。

[0102] • 指定运动/视差矢量预测因子候选者的给定列表的运动/视差矢量预测因子的运动/视差矢量预测因子索引。如果运动/视差矢量预测因子候选者的列表由单一元素组成，则该索引不被传输，但可以在解码器侧推断。对于图片的至少一个区块，运动/视差矢量预测因子候选者的列表包括基于已经编码的视点中的给定深度/视差信息和运动信息来推导的运动/视差矢量。

[0103] 另外，指定用于运动/视差补偿预测的运动/视差矢量和所选择预测因子（由至运动/视差矢量预测因子候选列表的传输索引指示）之间的运动/视差矢量差可以被作为比特流的一部分传输。在一个实施方式中，该运动/视差矢量差可以独立于参考索引和所选择预测因子来编码。在本发明的另一个实施方式中，运动/视差矢量差根据所传输的参考索引和/或所选择预测因子来编码。例如，如果选择特定运动/视差预测因子，则只可以编码运动/视差矢量差。

[0104] 参考图片列表和运动/视差矢量预测因子候选列表在编码器和解码器侧以相同的方式进行推导。在特定配置中，一个或多个参数在比特流中传输，以指定参考图片列表和/或运动/视差矢量预测因子候选列表如何被推导。对于本发明的优选实施方式，针对诸如22的从属视点中的图片的区块中的至少一个，运动/视差矢量预测因子候选者的列表包含基于给定（估计）深度值或基于给定（估计）深度值和已经编码的视点的运动参数来推导的运动或视差矢量预测因子候选者。在基于给定深度值和已经编码的视点的运动参数推导的运动/视差矢量预测因子附近，运动/视差矢量预测因子的候选列表可以包含空间预测运动矢量（例如，直接相邻区块（左边或上方区块）的运动/视差矢量、基于直接相邻区块的运动/视差矢量来推导的运动/视差矢量）和/或时间预测运动/视差矢量（例如，基于相同视点的已经编码图片中的协同定位区块的运动/视差矢量来推导的运动/视差矢量）。如下文中所述，可以执行运动/视差矢量候选者的推导，运动/视差矢量候选者通过使用给定深度数据64和其他视点诸如20的已经编码运动参数诸如42来获得。

[0105] 基于针对当前区块的代表深度的推导的推导

[0106] 在本发明的第一优选实施方式中，首先基于给定的基于样本或基于区块的深度图来获得给定区块50的代表深度值d。在一个优选实施方式中，考虑给定区块50的特定样本位置x，其可以是左上样本、右下样本、中间样本、或任何其他特定样本。与样本相关联的深度值 $d = d(x)$ （如通过给定区块式或样本式深度图64给出）被用作代表深度值。在另一个优选实施方式中，给定区块（例如，拐角样本或所有样本）的两个或两个以上样本位置 $x_i$ 被考虑并且是基于相关联深度值 $d_i = d(x_i)$ 的，代表深度值d被计算为深度值 $d_i$ 的函数。代表深度值可以通过深度值 $d_i$ 的集合的任意函数来获得。可能的函数为深度值 $d_i$ 的平均值、深度值 $d_i$ 的中值、深度值 $d_i$ 的最小值、深度值 $d_i$ 的最大值或任何其他函数。在获得给定区块的代表深度值d之后，运动/视差矢量预测展开如下：

[0107] • 如果针对区块50编码的参考索引指的是视点间参考图片（即，在与当前图片相同的时间实例的编码图片，但在已经编码视点诸如20中），则代表深度值基于如上所述的给定照相机或转换参数来被转换成视差矢量v， $v = f_{vd}(d)$ ，并且运动/视差矢量预测因子被设定为等于该视差矢量v。

[0108] • 否则（参考索引指的是时间参考图片（即，相同视点（诸如22）的已经编码图片），运动矢量预测因子基于给定参考视点或参考视点诸如20的集合来推导。参考视点或参考视

点的几何通过特定算法来确定或在比特流14中用信号传递。作为示例,参考视点在相同的时间点可以是事先编码的视点,或其可以是具有至当前视点的最小距离的已经编码视点(在相同的时间点),或由特定算法确定的任何其他的已经编码的视点。已经编码的视点的集合可以是当前时间点的已经编码视点的集合或该集合的任何子集。

[0109] 如果使用单一参考视点,则运动矢量预测因子推导如下。基于当前视点22和参考视点20的照相机参数或对应转换参数,代表深度d被转换成视差矢量 $v=f_{vd}(d)$ 。然后,假设视差矢量v,确定参考视点20中的样本位置 $x_r$ 。因此,考虑了当前区块50的特定样本位置 $x_r$ ,其可以是区块的左上样本、右下样本、中间样本或区块的任何其他样本。参考样本位置 $x_r$ 通过将视差矢量v添加到当前区块内侧的特定样本位置x中来获得。如果视差矢量v通过子样本准确性给出,则在被添加到样本位置中之前其被取整到样本准确性。假设参考样本位置 $x_r$ ,确定覆盖参考样本位置 $x_r$ 的参考视点20的图片32 $t_1$ (在当前时间点作为当前图片)中的区块40(区块是与唯一预测参数相关联的样本集合)。如果该参考区块40在帧间编码模式(即,采用运动补偿预测的模式,包括SKIP或MERGE模式)下编码,则调查用于预测该区块的一个参考图片或多个参考图片。令 $t_{c,R}$ 是被称为针对当前区块50编码的参考索引的参考图片(在当前视点中)的时间点。并且令 $t_{R,R}^i$ 是用于预测参考视点20中的参考区块(其覆盖样本位置 $x_r$ )的参考图片的时间点。如果用于预测参考区块40的一个或多个参考图片是与由当前区块50的参考索引给出的参考图片相同时间点的图片(即,如果 $t_{c,R}$ 等于值 $t_{R,R}^i$ 任何一个),则对应的运动矢量42用于推导当前区块50的运动矢量预测因子。如果时间点 $t_{R,R}^i$ 确切之一等于 $t_{c,R}$ ,则当前区块50的运动矢量预测因子被设为等于与 $t_{R,R}^i$ 的对应值相关联的参考区块40的运动矢量42。如果时间点 $t_{R,R}^i$ 的两个或两个以上等于 $t_{c,R}$ ,则运动矢量预测因子被设为等于参考区块40的相关联运动矢量的给定函数。可能的函数使用第一运动矢量(按任何特定顺序,例如通过使用 $t_{R,R}^i = t_{c,R}$ 的第一假设),另一个可能的函数使用运动矢量的平均值,又一可能的函数使用候选运动矢量的中值,或使用于推导运动矢量预测因子的所有分量的运动矢量分量的中值。如果相关联的参考图片都不具有等于 $t_{c,R}$ 的时间点 $t_{R,R}^i$ 。则运动矢量预测因子被标记成不可用。在本发明的优选实施方式中,无效的运动矢量预测因子没有包括在运动/视差矢量预测因子候选者的列表中。在本发明的另一个优选实施方式中,无效的运动矢量预测因子包括在运动/视差矢量预测因子候选者的列表中以增加解析算法的鲁棒性,但其不可以由编码器选择。在本发明的另一优选实施方式中,无效的运动矢量预测因子利用另一个定义的运动矢量来替换,所述定义的运动矢量例如可以是零矢量或使用相邻区块的运动矢量推导的运动矢量。而不是时间点( $t_{R,R}^i, t_{c,R}$ ),指定类似度量的其他参数可以用于确定运动参数集(由参考图片索引和运动矢量组成)是否可以用于推导运动矢量预测因子。例如,可以使用图片顺序计数(类似地如H.264中定义)或可以使用参考索引。

[0110] 如果使用两个或两个以上参考视点的集合,则运动矢量预测因子还可以基于所有参考视点中的信息来推导。在优选实施方式中,参考视点按特定顺序排序。作为示例,参考视点可以按其编码的顺序进行排序。或作为另一示例,参考视点按增加当前视点的距离的顺序排序。然后,排序集合的第一参考视点被调查并推导对应的运动矢量预测因子。如果该

运动矢量预测因子被标记为可用(即,其不被标记为不可用),则使用运动矢量预测因子。否则,如果运动矢量预测因子被标记为不可用,则给定排序集合中的下一个参考视点被调查,等等。在本发明的另一个优选实施方式中,给定集合的所有参考视点都被调查并且运动矢量的候选集合由与相关联时间点 $t_{R,R}^i$ 等于 $t_{C,R}$ 的参考图片相关联的所有对应运动矢量组成。最终运动矢量预测因子然后通过候选运动矢量的集合的函数来推导。可能的函数使用第一运动矢量(按任何特定顺序),另一个可能的函数使用运动矢量的平均值,又一可能的函数使用候选运动矢量的中值,或使用于推导运动矢量预测因子的所有分量的运动矢量分量的中值。

[0111] 为了进一步说明,在图4中使用与图1中的类似参考符号来描述推导在参考视点20中给出运动的当前区块 $50_C$ 的运动矢量和当前图片 $32t_2(T)$ 的深度图估计(使用当前区块 $50_C$ 内侧的特定样本位置)的基本过程以便简化图4的描述到图1的映射,以便充当可能实现的更详细的解释的可能源。假设当前区块 $50_C$ 中的样本位置 $x$ 和该样本位置的深度值 $d$ (其由深度图的估计64给出)推导视差矢量102,并基于该视差矢量102,推导参考视点20中的参考样本位置 $x_R$ 。然后,参考视点图片 $32t_1(T)$ 中的覆盖参考样本位置 $x_R$ 的区块 $40_R$ 的运动参数 $42_R$ 被用作当前视点22中的当前区块 $50_C$ 的运动参数的候选者。或可选地,参考区块的运动参数的子集用于当前区块 $50_C$ 。如果给出当前区块 $50_T$ 的参考索引,则只有参考区块 $40_R$ 的指相同时间点T(或图片顺序计数或参考索引)的运动参数 $42_R$ 才是当前区块 $50_C$ 的给定参考索引或被考虑。

[0112] 基于给定区块的多个深度值的推导

[0113] 在本发明的第二优选实施方式中,当前区块 $50_C$ 不用代表深度表示,但区块内侧的不同样本位置的不同深度值被推导出来并用于推导候选运动矢量预测因子的集合。假设当前区块,样本位置 $x^i$ 集合被考虑。样本位置集合可以包括区块的左上样本、右上样本、右下样本、左下样本、或中间样本。对于样本位置 $x^i$ 的每一个,深度值 $d^i$ 通过给定深度图分配。根据给定参考索引是否是指时间或视点间参考,以下项适用。

[0114] • 如果针对区块 $50_C$ 编码的参考索引指的是视点间参考图片(即,在与当前图片相同的时间实例的编码图片,但在已经编码视点中),则深度值 $d^i$ 基于如上所述的给定照相机或转换参数来被转换成视差矢量 $v$ , $v^i = f_{vd}(d^i)$ 。然后,运动/视差矢量预测因子根据这些视差矢量 $v^i$ 推导。运动/视差矢量预测因子被设定为等于最经常出现的视差矢量 $v^i$ ,或其可以被设为视差矢量 $v^i$ 的中值(或分量式中值),或其可以被设为视差矢量 $v^i$ 的平均值,或其可以通过视差矢量 $v^i$ 的任何其他函数来确定。

[0115] • 否则(参考索引指的是时间参考图片(即,相同视点的已经编码图片)),运动矢量预测因子基于给定参考视点或参考视点的集合来推导。对于每个样本位置 $x^i$ ,深度值 $d^i$ 被推导并被映射至视差矢量 $v^i$ 。然后,对于每个视差矢量 $v^i$ (被标记为可用),运动矢量 $m^i$ 通过上文指定的任何算法来推导(针对第一优选实施方式)。然后,最终运动矢量预测因子通过运动矢量 $m^i$ 的函数给出。运动矢量预测因子可以被设为等于最经常出现的运动矢量 $m^i$ ,或其可以被设为运动矢量 $m^i$ 的中值(或分量式中值),或其可以被设为运动矢量 $m^i$ 的平均值,或其可以通过运动矢量 $m^i$ 的任何其他函数来确定。

[0116] 方法2:针对其推导所有相关联的运动参数的模式

[0117] 在本发明的另一个优选实施方式中,多视点视频编码方案包括编码模式,其中所

有运动参数(包括假设次数、参考索引、运动矢量)都基于给定深度图64和已经编码的视点20的运动参数42来推导。在本发明的特定实施方式中,该模式可以被编码成候选运动参数的列表的候选者(因为其在当前HEVC工作草案中的合并语法中使用)。这意味着编码器和解码器以相同的方式推导区块的运动参数候选者的列表,其中运动参数候选者之一是基于已经编码的视点20的运动推导的运动参数。然后,索引被编码,其向解码器用信号传递使用这些运动参数候选者中的哪一个。在合并语法的上下文中,可以认为当前区块与参考视点中的“协同定位”(本着表示类似内容的精神)区块合并。在另一个实施方式中,特定语法元素用信号传递新编码模式的使用。在略有修改的版本中,用于生成预测信号的运动假设的次数可以在比特流内侧明确地进行信号传递,并且只推导参考索引和相关联运动矢量。在另一个修改的版本中,运动矢量差可以被额外传输以便提纯所推导的运动参数。

[0118] 基于多个潜在参考索引的推导

[0119] 在本发明的第一优选实施方式中,用于当前区块50<sub>c</sub>的运动参数的推导使用上文针对方法1描述的任何构思并考虑一个以上的潜在参考索引。在下文中,我们首先描述可以如何推导特定运动假设的参考索引(和参考列表)以及相关联运动矢量。作为第一步骤,确定给定参考列表的参考索引的排序集合。这例如可以仅是单一参考索引(例如,参考列表的第一索引或表示时间参考图片的第一索引),或其可以由参考列表的前两个索引组成,或其可以由参考列表的所有参考索引组成,或其可以由表示时间参考图片的第一参考索引和不等于表示时间参考图片(即,列表中的第二时间参考图片或第一视点间参考图片)的第一参考索引的第一参考索引组成。可能存在参考索引的任何其他定义集合。假设参考索引的排序集合,第一参考索引被考虑并且该参考索引的运动矢量通过上文针对方法1描述的任何实施方式来推导。如果推导的运动矢量被标记为不可用,则下一参考索引被考虑并推导对应的运动矢量。继续该过程直至返回可用的运动矢量或测试列表的所有参考索引。如果没有发现可用的运动矢量,则最终运动参数被标记为不可用。在一个配置中,不可用的运动参数不被插入到运动参数的候选列表中。在第二配置中,不可用的运动参数被插入到运动参数的候选列表中(针对解析鲁棒性),但不允许编码器选择不可用的运动参数。在第三配置中,不可用的运动参数用特定运动参数替换,例如其可以是零参考索引和零运动矢量或基于当前区块的空间邻域中的运动参数来推导的参考索引和运动矢量。如果新编码模式通过特定语法元素用信号传递并且推导的运动参数不可用,则不传输对应的语法元素(并且不使用编码模式)或不允许编码器选择指定新编码模式的使用的语法元素的值或不可用的运动参数用特定运动参数替换(同上)。

[0120] 如果运动假设的次数或使用的参考列表的次数被明确编码,则针对如上所指定的每个运动假设或参考列表确定由参考索引和运动矢量组成的运动参数的集合。

[0121] 如果运动假设的次数或使用的参考列表的次数未被明确编码,则还基于参考视点中的实际编码运动参数来推导运动假设次数或所采用的参考列表。假设可以使用的运动假设的最大次数或参考列表的最大集合,针对每个运动假设(参考列表),推导运动参数的集合,如上所述。然后,运动假设的次数(所使用的参考图片列表的集合)通过所推导的运动参数被标记为可用的假设(参考列表)来给出。作为示例,如果我们具有两个潜在运动假设并针对这两个运动假设推导运动参数(参考索引和运动矢量)的有效集合,则新编码模式利用所推导的运动参数指定双向预测。然而,如果只针对假设(参考列表)之一推导运动参数的

有效集合,则新编码模式利用有效运动参数的集合指定单向预测(一个假设)。如果不针对任何运动假设(参考列表)推导运动参数的有效集合,则运动参数的完整集合被标记为不可用。在这种情况下,运动参数的集合不添加到候选运动参数的列表中,或者其被添加(针对解析鲁棒性)但不被编码器使用,或其用运动参数的特定定义集合替换(例如,利用一个运动假设,等于0的参考索引和等于0的运动矢量)。还可以针对一个或多个潜在运动假设来检查参考索引的另一个集合。

[0122] 基于单个代表深度值的推导

[0123] 在本发明的第二优选实施方式中,首先推导参考视点中的参考区块,然后该区块的运动参数被用作当前区块的运动参数候选者。这里,运动假设导电次数以及参考索引和运动矢量从参考视点中的参考区块复制。该实施方式的基本构思在图2中示出并在上文进行了简要描述。首先,代表深度值d,和基于该深度值的视差矢量v,和参考样本位置 $x_R$ 通过针对方法1描述的任意算法来推导。然后,考虑了覆盖参考样本位置 $x_R$ 的参考视点中的区块(也被称为参考区块)。当前区块的运动参数(或运动参数的一个候选者)被设为等于推导参考区块的运动参数。例如在以下场景中还可以修改运动参数(尤其是参考索引和运动假设的次数):

[0124] • 如果参考视点的参考图片列表以与当前视点的参考图片列表不同的方式构造(即,特定参考索引不总是指两个列表的相同访问单元),则当前区块的参考索引可以以其指的是与参考视点中的对应参考图片相同的时间点的图片(或具有相同的图片顺序计数的图片)的方式来修改。

[0125] • 如果参考视点中的参考索引指的是视点间参考图片,则当前视点的参考索引可以以其指的是所选视点间参考图片(例如,与当前视点相同的视点间参考图片或用参考视点表示的参考图片)的方式来修改。在这种情况下,运动矢量还必须用视差矢量替换,该视差矢量可以通过将代表深度d转换为视差矢量而获得。

[0126] • 如果不针对参考区块中使用的所有参考图片,对应的图片(相同的时间点或参考索引的图片顺序计数)在当前区块的参考列表中可用,则指的是不可用于当前区块的参考图片的运动假设可以被视为不存在。

[0127] • 如果对参考区块进行帧内编码,则运动参数可以用视差补偿预测的运动参数替换。例如,参考索引可以以其指的是参考视点图片的方式设定并且运动参数可以被设为等于通过将代表深度d转换为视差矢量而获得的视差矢量。作为替代,运动参数可被标记为不可用。

[0128] 与方法1组合

[0129] 在实施方式中,除了由方法2的实施方式描述的编码模式(推导包括运动假设的次数、参考索引和运动矢量或运动矢量预测因子的所有运动参数)之外,还可以支持由方法1的实施方式描述的编码模式(编码参考索引、推导运动矢量或运动矢量预测因子)。

[0130] 方法3:针对其推导所有相关联运动参数以及区块划分的模式

[0131] 在本发明的另一个优选实施方式中,多视点视频编码方案包括编码模式,其中给定区块的子区块的不同运动参数基于已经编码的视点(诸如20)和估计深度图64的运动参数来推导。或者换句话说,多视点视频编码方案包括区块的编码模式,其中将区块50<sub>c</sub>划分为更新的子区块以及与子区块相关联的运动参数基于已经编码的参考视点和估计深度图

中的运动参数来推导。

[0132] 针对该编码模式,定义最小区块尺寸,其可以等于针对运动/视差补偿预测支持的最小区块尺寸或可以是针对运动/视差补偿预测支持的最小区块尺寸的倍数。如果给定当前区块 $50_c$ 小于或等于所定义的最小区块尺寸,则出于运动/视差补偿目的当前区块 $50_c$ 不被分割成更小的区块并推导相关联运动参数,如上文针对方法2的任何实施方式所述。如果给定当前区块大于所定义的最小区块尺寸,则其被分割成具有等于所定义的最小区块尺寸的尺寸的子区块。对于这些子区块的每一个,使用上述方法2的任何实施方式来推导运动参数的集合。

[0133] 如果任何子区块的运动参数被标记为不可用(例如,因为对应参考区块 $40_R$ 在帧内模式下编码或其仅使用视点间预测),则其可以用运动参数可用的任何相邻子区块的运动参数来替换。这样的算法可以以相邻区块按特定定义顺序(其可以取决于要替换的子区块的位置)测试的方式操作,并且要替换的子区块的运动参数被设为等于按指定顺序的具有有效运动参数的第一子区块的运动参数。

[0134] 在本发明的具体实施方式中,具有给定最小区块尺寸的所获得的子区块指定给定当前区块 $50_c$ 的划分。在本发明的另一个实施方式中,所获得的子区块可以基于相关联运动参数进行组合以便形成用于运动/视差补偿预测的更大区块。组合子区块可以以分层的方式展开。因此,在第一分层阶段,可以考虑四个相邻区块的集合。如果运动假设的次数和相关联参考图片和运动矢量对所有这四个子区块来说都相同,则这四个子区块被汇总为更大的区块(具有等于原始子区块的运动参数的运动参数)。在下一个分层阶段,考虑下一个层级的四个区块(由4个原始子区块组成)。如果所有这四个区块在前一分层阶段都被汇总为更大的区块并且运动假设的次数和相关联参考图片和运动矢量对所有这四个区块来说都相同,则这四个区块再次被汇总为更大的区块(具有等于原始子区块的运动参数的运动参数)。该算法一直持续到给定当前区块的最高可能层级。在极端情况下(如果所有子区块的运动参数都相同),整个当前区块不被分割但与运动参数的唯一集合相关联。在略有修改的版本中,如果运动矢量不完全相同,则4个区块也被汇总为更大的区块,但运动矢量之间的差(其可以被定义为最大分量差或矢量差的绝对值)小于或等于定义阈值(运动假设的次数和所采用的参考图片仍然必须相同)。在这种情况下,与更大的区块相关联的运动矢量根据4个子区块的运动参数来确定。可能的函数为运动矢量的平均值、运动矢量的中值(或分量式中值)、任何特定子区块的运动矢量、或在四个子区块中最经常出现的运动矢量。

[0135] 在本发明的实施方式中,除了由方法3的实施方式描述的编码模式(推导区块分割以及包括运动假设的次数、参考索引和运动矢量或运动矢量预测因子的所有运动参数)之外,还可以支持由方法1的实施方式描述的编码模式(编码参考索引、推导运动矢量或运动矢量预测因子)。另外,可以支持根据方法2的任何实施方式的编码模式。

[0136] 编码运动和视差数据

[0137] 如上文已经描述的,使用表示本发明的实施方式的编码模式必须向解码器发出信号。这可以以不同的方式来实现。在一个版本中,特定语法元素(其可以是标志)可以被插入到语法中,其用信号传递是否使用传统推导的运动矢量预测因子(或运动矢量或运动参数集)或是否使用已经使用已经编码的视点的深度图估计和运动参数推导的运动矢量预测因子(或运动矢量或运动参数集)。在另一个版本中,已经使用已经编码的视点的深度图估计

和运动参数推导的运动矢量预测因子(或运动矢量或运动参数集)可以被插入到传统推导的运动矢量预测因子(或运动矢量或运动参数集)的候选列表中并传输索引,所述索引用信号传递使用哪一个运动矢量预测因子(或运动矢量或运动参数集)。特定语法元素或至候选列表的索引可以使用固定长度编码、可变长度编码、算术编码(包括基于上下文的自适应二进制算术编码)或PIPE编码进行传输。如果使用上下文自适应编码,则上下文可以基于相邻区块的参数(例如,特定语法元素或至候选列表的索引)来推导。

[0138] 在本发明的优选实施方式中,多视点视频编码方案包括一个或多个运动假设通过传输参考图片索引、运动矢量预测因子索引和每个运动假设的运动矢量差而用信号传递的编码模式。针对该编码模式,候选运动矢量预测因子的列表基于所传输的参考图片索引和所传输的所有信号来推导,运动矢量候选者之一被使用。通过使用该实施方式,运动矢量候选者之一(针对至少一个区块)基于已经编码的视点的深度图估计和运动参数来推导(参照上述方法1)。在略有修改的版本中,运动矢量差不被传输但被推断等于0(针对所有运动矢量候选者或只针对基于已经编码的视点的深度图估计和运动参数来推导的运动矢量候选者)。

[0139] 在本发明的另一个优选实施方式中,多视点视频编码方案包括一个或多个运动假设通过传输运动参数索引(或合并索引)而用信号传递的编码模式。针对该编码模式,推导运动参数(包括运动假设的次数、参考索引、运动矢量)的候选集合的列表。通过使用该实施方式,运动参数的候选集合之一(针对至少一个区块)基于已经编码的视点的深度图估计和运动参数来推导(参见上述方法2和3)。在该实施方式的特定版本中,基于已经编码的视点的深度图估计和运动参数来推导的运动参数的候选集合(针对至少一个区块)包括划分当前区块的信息(参见上述方法3)。在该实施方式的略有修改的版本中,运动矢量差可以另外进行传输(可能取决于运动参数的所选集合)。

[0140] 在本发明的另一个实施方式中,多视点视频编码方案包括对应于方法2或3的编码模式并且语法包括指定是否使用该编码模式的标志。

[0141] 推导用于当前图片的深度图估计

[0142] 基于已经编码的视点(诸如20)的运动参数推导当前图片50C的区块的运动参数(如迄今所述)要求当前图片的深度图的估计64是可用的。如上所述,该深度图估计64可以指定样本式深度图(针对当前图片的每个样本指定了深度值)或区块式深度图(针对样本的区块指定了深度值)。深度图估计64可以基于已经编码的参数(诸如深度图或视差矢量和运动参数)来推导。原则上,推导当前图片的深度图估计64的可能性可以被分为两类。针对一个类别,深度图估计基于实际编码的深度图来推导。下面描述的对应可能性涉及编码深度图存在于比特流中(在其被使用之前)。第二类别的构思不要求深度图被编码成比特流的一部分。相反,深度图估计基于编码视差矢量来推导。可以应用第二类程序,与深度图是否被编码成比特流的一部分无关。这是上文针对图1和图2讨论的情况,以下描述到目前为止提供了关于各个方面的个别可转换细节。还应注意,当编码深度图时,这两类方法都可以被应用。还可以针对不同帧选择不同方法。在下文中,描述了推导深度图估计(具有和没有编码深度图)的基本构思和优选实施方式。

[0143] 类别1:基于编码深度图的推导

[0144] 如果与当前图片 $32t_2(T)$ 相关联的深度图在当前图片之前被编码,则重构的深度

图可直接被用作当前图片的真实深度图的估计。还可以预处理编码深度图(例如,通过应用a对其进行滤波处理)并且还可以使用预滤波处理的结果作为用于推导运动参数的深度图的估计。

[0145] 在多数配置中,与特定图片相关联的深度图在图片 $32t_2(T)$ 之后(经常直接在相关联图片之后)被编码。此配置允许被传输以编码传统视频图片的编码参数(诸如运动参数)可以用于预测用于编码深度图的编码参数,这提高了整体编码效率。但在此配置中,与图片相关联的深度图在推导运动参数54时不能被用作深度图的估计。然而,已经编码的视点(同一访问单元的)诸如20的深度图通常是可用的并且可以用于推导当前图片的深度图的估计。基础视点(独立视点)20的深度图至少在编码任何从属视点22之前是可用的。由于任何视点的深度图都可以一定程度表示投影视频场景的几何形状(结合照相机参数诸如焦距和照相机之间的距离),因此其可以被映射至另一个视点。因此,如果当前图片 $32t_2(T)$ 的深度图是不可用的,则相同访问单元20的已经编码的视点的编码深度图被映射至当前视点并且该映射结果被用作深度图估计。

[0146] 在下文中,我们描述了实现该映射的特定算法。如上所述,每个深度值d都对应于两个给定视点之间的位移矢量v。假设所传输的照相机或转换参数,深度值d可以通过映射转换为位移矢量 $v = f_{vd}(d)$ 。因此,假设参考深度图(已经编码)中的特定样本位置 $x_R$ 处的深度值d。当前深度图中的相同深度值的样本位置 $x_C$ 通过将视差矢量添加到 $x_R$ 来获得, $x_C = x_R + v$ 。因此,参数深度图的每个深度值都可以被映射至当前深度图的样本位置以便获得当前图片的深度图估计64。然而,由于在一个视点中可见的物体的部分在另一个视点中不可见,因此在分配有一个以上深度值的当前视点22的深度图中存在样本位置,并且在没有分配深度值的当前视点的深度图中也存在样本位置。这些样本位置可以进行如下处理:

[0147] • 如果向特定样本位置分配一个以上深度值,则意味着前景物体在背景物体前面发生位移。因此,向此样本位置分配表示至照相机的最小距离的深度值d。

[0148] • 如果不向特定样本位置分配深度值,则意味着前景物体移动并且事先覆盖的背景是可见的。针对这些区域的最好的办法是假设去遮挡背景具有与相邻背景样本相同的深度。因此,没有分配深度值的区域用表示至照相机的最大距离的周围样本的深度至来填充。

[0149] 该算法在下文中被更详细地指出。为了简化以下描述,我们假设与更小的深度值相比,更大的深度值表示至照相机的更小距离(但该算法可以针对相反的假设被轻易修改):

[0150] 1.当前图片的深度图(估计)的所有样本都被设为未定义的深度值(例如,-1)。

[0151] 2.对于参考深度图的每个样本位置 $x_R$ ,以下项适用:

[0152] a.样本位置 $x_R$ 处的深度值d适用给定的照相机或转换参数被转换为视差矢量v,视差矢量v被取整为样本准确性(若适用的话),并且当前图片内侧的样本位置通过 $x_C = x_R + v = x_R + \text{round}(f_{vd}(d))$ 来推导。

[0153] b.如果当前图片中的样本位置 $x_R$ 处的深度值具有未定义值,则样本位置处的深度值被设为等于深度值d。

[0154] c.否则,如果当前图片中的样本位置 $x_R$ 处的深度值具有定义值 $d_x$ (其中 $d_x < d$ ),则样本位置处的深度值被修改并被设为等于深度值d。

[0155] 3.当前深度图中具有未定义深度值的区域通过特定孔填充算法来填充。针对此孔

填充算法,在当前视点中未被覆盖的背景的深度值基于周围深度值的样本来推导。作为示例,可以分配周围样本的最小深度图值。但是可能存在更复杂的孔填充算法。

[0156] 基于非常简单的示例在图5中进一步示出了将给定视点的深度图映射至不同视点的算法。图5示出了将针对一个视点20给出的诸如 $32t_1(T)$ 的深度图映射至另一个视点22的可能的过程。在左手边,示出了参考视点的给定深度图,其中阴影区域表示背景,白色区域表示前景物体。在图5的中间,示出了中间的转换深度图,其通过移位具有对应于深度值的视差矢量的给定图的样本并保持一个以上样本投影至此的位置的前景物体来获得。黑色区域表示样本不投影至此的去遮挡区域。图5的右边示出了在利用背景的深度值填充去遮挡区域(即,背景填充)之后的转换深度图。

[0157] 在本发明的具体实施方式中,孔填充可以通过分开处理转换深度图的线的尤其简单的算法来实现。针对每个线段(其由连续的未定义深度值组成),两个周围值被考虑,并且线段的所有深度样本都用这两个深度值(背景深度)的更小深度值来替换。如果线段只具有一个周围深度值(因为其位于图像边界处),则线段的深度样本用该值替换。如果完整的线在该处理之后具有未定义值,则相同的过程适用于深度图的列。

[0158] 虽然上述算法已经针对样本式深度图进行了描述,但其还可以适用于区块式深度图(导致更低的复杂度)或参考视点的给定样本式深度图可以首先被转换成区块式深度图(通过下采样)并且然后该算法可以适用于区块式深度图。

[0159] 类别2:基于编码视差和运动矢量的推导

[0160] 如果深度图没有被编码成比特流的一部分,则深度图的估计可以通过使用编码运动和视差矢量来生成。以下构思的基本思想可以总结如下。解码(多视点)视频序列通常开始于随机访问单元。随机访问单元中的基础视点的图片在不参考任何其他图片的情况下进行帧内编码。随机访问单元中的从属视点的图片可以使用视差补偿预测进行帧内编码或预测。通常,多数区块通过视差补偿预测来编码,因为其通常给出比帧内预测更好的预测结果。由于编码视差矢量(其用于视差补偿预测)可以被转换为深度值(适用反函数 $f_{vd}^{-1}$ ),因此视差矢量可以直接用于生成与随机访问单元中的从属视点相关联的基于区块的深度图(帧内编码区块的深度可以基于周围视差补偿区块的深度来估计)。然后,所获得的深度图可以被映射至基础视点。基础视点的下一图片通常使用大部分运动补偿的预测来编码。这里,可以假设深度数据的运动与纹理信息(深度和相关联纹理样本属于相同的物体点)的运动相同。考虑到该假设,基础视点中的第一图片的估计深度数据可以进行运动补偿以获得当前访问单元中的基础视点的深度图的估计。然后,基础视点的(运动补偿)深度图估计可以被映射至从属视点以获得当前图片的深度图估计(在当前视点中)。如果两个以上的视点被编码,则可以简化第三视点、第四视点等的深度图估计的创建,因为我们还具有访问单元的前两个视点的深度图估计。这些深度图估计之一(优选基础视点)可以被映射至第三、第四或任何以下视点以便生成该视点的深度图估计。

[0161] 进一步通过一些示意图示出生成深度图估计的思想(示出了利用两个视点的多视点编码的处理步骤,如由深度估计器28执行的一样)。编码/解码开始于随机访问单元,基础视点图片 $32t_1(0)$ 针对其进行帧内编码并且非基础视点图片 $32t_2(0)$ 只使用视点内和视点间预测(但不使用运动补偿预测)来编码。在编码随机访问单元“0”中的第二视点22之后,使用该视点22的编码视差矢量122来生成120该第二视点22的基于区块的深度图估计,如图6所

示。第二视点22的深度图估计 $64_2(0)$ 然后被映射至第一视点(基础视点)20并获得第一视点20的深度图估计 $64_1(0)$ 。应注意,对于随机访问单元的第二视点22,可以使用基于基础视点的运动参数和视差估计推导运动/视差参数,因为深度图的估计在编码随机访问单元的第二视点22时是不可用的。

[0162] 如果编码第三视点,则前两个视点的任意一个(优选第二视点)的深度图估计可以被映射至第三视点,由此得到第三视点的深度图估计,其可以用于推导第三视点的运动参数。在编码第三视点之后,基于区块的深度图可以使用第三视点的编码视差矢量(其稍后可以用于生成任何以下视点的深度图估计)来生成。对于任何以下视点,基本上可以使用与第三视点相同的处理。

[0163] 非随机访问单元中的基础视点的图片通常主要通过运动补偿预测来编码,因为运动补偿预测通常给出比帧内编码更好的编码效率。在编码基础视点的图片之后,使用图片 $32t_1(1)$ 的运动参数42(1)来生成140(在图1中比较71)该图片的深度图估计,如图7所示。因此,新深度图估计 $64_1(1)$ 的每个区块通过运动补偿一个对应参考图片或多个参考图片的深度图估计 $64_1(0)$ (在图1中比较74)来创建140。所使用的参考图片和对应运动矢量42(1)是在相关联图片的数据流中编码的参考图片和运动矢量。帧内编码区块的深度样本可以通过空间预测来获得。基础视点的深度图估计然后被映射至142(在图1中比较76)第二视点的坐标系以便获得可以用于推导运动参数的第二视点的深度图估计 $64_2(1)$ ,即以执行视点间冗余减少。

[0164] 对于任何另外的编码视点,可以通过将任何已经编码的视点(基础视点、第二视点等)的深度图估计映射至对应视点来生成深度图估计。

[0165] 在实际编码第二视点(或任何以下视点)的图片之后,可以使用实际编码运动和视差矢量来更新160(在图1中比较77)相关联的深度图估计,如图8中所示。针对使用视差补偿编码的区块,深度图可以通过将编码视差矢量60转换为162如上所述的深度值来获得。针对使用运动补偿模式编码的区块,深度样本可以通过运动补偿参考帧 $32t_2(0)$ 的深度图估计来获得。或可选地,被添加到当前深度图估计 $64_2(1)$ 的深度校正值可以基于当前视点和参考视点的编码运动参数42(1)和54(1)来推导。帧内编码区块的深度样本可以使用空间预测或使用相邻区块的运动参数来预测。在生成第二视点的更新深度图估计74之后,该深度图估计74被映射至164(在图1中比较78)基础视点20以获得基础视点20的深度图更新 $64'_1(1)$ (在图1中比较74)。

[0166] 如果编码两个以上视点,则用于这些视点的深度图更新过程与第二视点的相同。然而,基础视点深度图仅在编码第二视点之后被更新。

[0167] 针对深度图的运动补偿操作或者可以使用编码子样本的准确运动矢量来执行。然而,如果深度图的运动补偿操作利用样本(或甚至区块)准确性来执行,则其是优选的(从复杂度以及视点的编码效率点来说)。因此,实际编码的运动矢量被取整到样本或区块准确性并且这些取整矢量用于执行运动补偿。此外,所描述的构思可以适用于样本式以及区块式深度图估计。使用基于区块的深度图的优点是复杂度较低,并且所有处理步骤的存储需求也较低。利用基于区块的深度图,每个深度样本表示相关联图片的样本的区块(例如, $4 \times 4$ 区块或 $8 \times 8$ 区块)的深度。可以针对基于区块的深度图以简单的方式(即,通过简单考虑深度图的较低分辨率——一个深度样本只表示多个纹理样本而不是一个纹理样本)来执行所有

描述的操作。

[0168] 除了将给定深度图从一个视点映射至另一个视点之外(如上所述),该算法还包含以下基本步骤:

[0169] • 基于随机访问单元的图片的视差矢量来创建深度图。

[0170] • 使用相关联图片的运动参数进行基础视点深度图的时间预测。

[0171] • 使用相关联图片的实际编码运动和视差矢量来更新深度图估计。

[0172] 这些算法步骤的具体实施方式描述如下。

[0173] 创建随机访问单元中的图片的深度图

[0174] 在本发明的具体实施方式中,创建随机访问单元中的从属视点的图片的深度图展开如下。总体来说,此图片包含使用视差补偿预测编码的区块以及进行帧内编码的区块。首先考虑使用视差补偿预测编码的区块。视差矢量被转换为深度值并将这些深度值分配给深度图的对应样本。如果使用两个或两个以上运动假设,则选择其中一个假设或最终深度值被设为等于各个运动假设的深度值的函数(例如,平均值、中值、最大值或最小值)。在分配所有视差补偿区块的深度值之后,帧内编码区块的深度值通过空间帧内预测来获得。在一个版本中,这些样本可以通过使用于相关联纹理图片的相同帧内预测模式来获得。在另一个版本中,帧内区块的深度可以设为等于通过周围样本(或区块)的加权平均值获得的深度值,其中加权因子可以基于所使用的帧内预测模式来确定。在另一个版本中,帧内区块的深度可以通过将深度样本设为等于由周围帧内样本的特定函数(例如,平均值、中值、最大值或最小值)给出的值来获得。还可能存在其他控件预测算法。帧内编码区块的深度赋值还可以在单个回路内侧在图像中的区块上进行。这意味着按特定顺序(例如,编码顺序)来处理区块,并且针对视差补偿和帧内区块这两者,按该顺序生成深度值(即,帧内编码区块的深度赋值直到处理所有视差补偿区块才需要等待)。

[0175] 基础视点深度图的时间预测

[0176] 总体来说,基础视点的图片包含运动补偿区块和帧内编码区块。运动补偿区块的深度值通过对对应参考图片的深度图估计的运动补偿预测来推导。如果纹理图片的区块使用单一运动假设来编码,则该区块的深度样本可以通过使参考图片(由用信号传递的参考索引)的深度图估计的深度样本移位所传输的(或推断的)运动矢量来获得。该运动补偿操作可以利用传输运动矢量的准确性(其通常是子样本准确性)或利用样本准确或区块准确的运动矢量来执行。如果运动补偿利用子样本准确性来执行,则内插滤波器适用于在子样本位置生成样本。如果运动补偿利用样本或区块准确性来执行,则在使用之前对所传输的运动矢量取整。如果相关联图片的区块利用两个以上的运动假设来编码,则针对深度图的运动补偿可以选择假设之一,或所有运动假设通过生成区块的最终深度图估计而被用作各个运动假设的深度预测信号的加权和。此外,规定尺寸的区块的深度样本可以被设为等于代表深度。该代表深度可以通过选择区块内侧的特定位置并使用运动补偿推导该位置的深度值来获得,或其可以通过定义该区块的运动补偿深度值的函数来获得。这样的函数可以是深度样本的平均值、或深度样本的中值、或深度样本的最小值或最大值、或最经常出现在区块中的深度图样本值。

[0177] 使用编码运动和视差参数来更新深度图

[0178] 如上所述,图片的深度图估计在编码该图片之后可以被更新。在下文中,我们描述

了此更新算法的一些实施方式。图片的区块按特定顺序(例如,编码顺序)处理,针对每个区块,以下适用:

[0179] • 如果区块已经进行了帧内编码(在相关联纹理图片中),则该区块的深度样本值可以通过空间预测使用相邻区块的样本来推导。上面已经描述了此空间预测技术的一些示例。深度样本值还可以通过运动补偿使用相邻区块的运动参数来获得(参见以下运动补偿区块的描述)。如果在处理所有运动和视差补偿区块之后对帧内区块进行了处理,则有时是优选的,因为更多相邻区块是可用的并且可以用于深度样本或运动参数的空间预测。

[0180] • 否则,如果区块使用一个或多个视差矢量(视差假设)来编码,则深度样本通过将视差矢量转换为深度值来推导到。如果只有一个视差矢量(视差假设)用于该区块,则深度值由对应视差矢量给出。如果两个或两个以上视差矢量用于该区块,则视差矢量之一可以选择来推导深度值,或针对每个视差矢量,可以推导深度值,并且最终分配的深度值通过应用单独推导的深度值的函数来获得。可能的函数可以尤其是各个深度值的最小值或最大值、深度值的中值、深度值的平均值或最经常出现的深度值。

[0181] 注意,使用视差矢量编码的区块可以另外与时间运动矢量相关联。在这种情况下,可以忽略时间运动矢量。或推导的深度值可以与针对时间运动假设(见下文)以任何特定方式(例如,通过对着两个信号取平均值)推导的深度值组合。

[0182] • 否则,该区块只使用时间运动假设来编码并且时间运动假设用于更新区块的深度样本。

[0183] 在本发明的第一实施方式中,深度图样本通过简单的运动补偿预测使用与当前视点的参考图片相关联的深度图估计来推导。该运动补偿操作可以由上述基础视点深度图的时间预测的任何实施方式来实现。

[0184] 在本发明的第二实施方式中,深度图样本不进行简单的运动补偿,相反深度校正值基于针对当前视点编码的运动矢量和针对参考视点中的对应区块编码的运动矢量来推导,该深度校正值被添加到所估计的深度图中以便获得更新的深度图估计。此方法的优点是可以考虑两个实例之间的深度变化。

[0185] 令 $d_{prd}$ 为针对样本或区块的当前深度估计,令 $m_{curr}$ 为实际用于当前区块的运动补偿的运动矢量,并令 $m_{ref}$ 为参考视点中的参考区块的运动矢量(参考区块通过使用由如上所述的深度估计 $d_{prd}$ 给出的视差矢量来推导)。此外,令 $v_{t-1}$ 为参考访问单元中的视差矢量并令 $v_t$ 为当前访问单元中的视差矢量。得出我们在描述开始时推导的运动和视差矢量之间的相互关系的基本方程,

$$m_{curr} + v_{t-1} - m_{ref} - v_t = 0,$$

[0187] 如果我们假设参考图片中的编码运动参数和视差是正确的,则当前视差矢量可以被设为等于

$$v_t = v_{t-1} + (m_{curr} - m_{ref})$$

[0189] 视差矢量可以被转换为深度,反之亦然。因此,当前深度可以表示为

$$d = d_{t-1} + f_{vd}^{-1}(m_{curr} - m_{ref})$$

[0191] 深度 $d_{t-1}$ 是参考图像中给出的深度值,其可以使用当前区块的运动矢量来访问。还可以用深度 $d_{prd}$ 的当前估计来替换参考图像中的深度,因为该深度已经使用运动补偿从参考图像中的深度推导出来。

[0192] 我们使用当前和参考区块的运动矢量示出了用于获得更新深度值的基本构思。该基本构思可以在各种实际实现中使用。例如,如果一个以上的运动矢量被用于当前区块和/或参考区块,则可以选择单个运动矢量或不同深度值可以针对不同运动矢量来推导并且最终深度通过使用单独推导的深度值的平均值(或任何其他函数)来获得。此外,应该只考虑当前和参考区块中的运动假设,其使用相同的参考访问单元(或参考时间点,或参考图片顺序计数,或参考索引)。如果不存在这些运动假设,则可以使用上述传统的运动补偿过程,或深度样本可以被标记为不可用并且稍后通过合适的后处理步骤来替换。此外,新深度值可以基于区块的单个样本、或区块的所有样本、或区块的样本的所选子集来推导。当推导不同样本的不同深度值时,可以为当前区块的不同区域(例如,为当前区块内侧的所有 $4 \times 4$ 区块)分配独立深度值,或不同深度值用于例如通过使用平均值、中值、最小值、最大值或最经常出现的深度值来推导整个区块的代表深度。深度图更新过程还可以与一些预处理步骤组合。例如,不可用的深度值可以用相邻样本或区块的深度来替换。

[0193] 针对方法1-3的前述描述,应注意,如果与确定用于识别参考视点的相应区块 $40_R$ 的视差矢量的其他类型一起使用,则通过添加从区块,诸如参考视点中的区块 $40_R$ 获得的相应信息来给运动或运动/视差矢量/参数候选列表润色的这些不同替代方案也是有利的。根据该实施方式,图1的解码器中的深度估计器28是任选的,就像图2的编码器中的深度估计器84一样。

[0194] 具体地,根据后一个方面,方法1-3的上述描述还揭露了一种用于将多视点信号重构为多视点数据流的装置,该装置包括从属视点重构器26,其被配置为进行以下操作一遍重构从属视点22。参照图1和图4的多视点信号的示例性说明,以便描述功能。具体地,从属视点重构器26针对当前图片 $32t_2(T)$ 的区块50c和从属视点22通过首先确定当前区块50c的视差矢量102来推导运动矢量预测因子候选者的列表,该视差矢量102表示从属视点22的当前图片 $32t_2(T)$ 和区块50c处的多视点信号的参考视点20的当前图片 $32t_1(T)$ 之间的视差。为了这样做,从属视点重构器26使用与多视点信号的事先解码部分相关联的运动和视差矢量诸如图片 $32t_1(T)$ 和 $32t_2(T-1)$ 的运动/视差矢量。在上述的其他实施方式中,与参考视点的当前图片 $32t_1(T)$ 相关联的估计深度图被用作确定视差矢量102的依据,其中使用从属视点以及上述参考视点的事先编码图片的运动和视差矢量来进行深度图的估计和更新,并且在这方面,上述描述也应该并入当前实施方式,但原则上,还存在其他可能性。例如,从属视点重构器26可从空间上/时间上预测当前区块50c的视差矢量并使用该预测视差矢量作为视差矢量102。

[0195] 然后,从属视点重构器26使用所确定的视差矢量102来确定参考视点的当前图片内的区块 $40_R$ ,并将运动矢量添加到运动矢量预测因子候选者的列表中,这取决于与所确定的区块 $40_R$ 相关联的运动矢量,即,运动矢量 $42_R$ 。

[0196] 如上所述,在推导运动矢量预测因子候选者的列表时,从属视点重构器还可被配置为从从属视点22的空间和/或时间相邻区块即空间和/或时间相邻的当前区块50c从空间上和/或时间上预测一个或多个另外的运动矢量。一个或多个另外的运动矢量或由此推导的版本然后通过从属视点重构器26被添加到运动矢量预测因子候选者的列表中。

[0197] 从属视点重构器针对区块50c从多视点数据流提取指定运动矢量预测因子候选者的列表之一的索引信息,为了更具体,从其从属视点部分22提取。因为假设当前区块50c进

行运动补偿预测,即,与时间预测模式相关联,所以从属视点重构器26通过使用运动矢量执行区块50<sub>c</sub>的运动补偿预测来重构区块50<sub>c</sub>,该运动矢量等于或至少取决于指定的运动矢量候选者,即由索引信息进行索引或指定的运动矢量候选者。与运动矢量候选者的放大列表相关联的开销与运动矢量预测质量的增益相比相对较低,该运动矢量预测质量由添加从参考视点确定的运动矢量候选者42<sub>R</sub>导致。

[0198] 如上文已经描述的,用于当前区块50<sub>c</sub>的由从属视点重构器26提取的运动信息不一定局限于索引信息。相反,从属视点重构器26可进一步被配置为针对区块50<sub>c</sub>提取关于指定运动矢量候选者的运动矢量差并执行区块50<sub>c</sub>的重构,以使得所使用的运动矢量进一步取决于运动矢量差和指定运动矢量候选者(即,由运动矢量预测因子候选者的列表中的索引信息指定的)的和。

[0199] 在上文中,严格区分了运动和视差补偿预测。然而,如果相同的模式用于用信号传递这两者,则这两者之间的差可以成为零,其中这两者之间的差可从对图片进行索引的所有推导出来,相对于该图片执行运动补偿预测。然后,解码器从属视点重构器的刚才提及的功能可通过用“运动/视差”替换“运动”来重写,因为没有什么区别。自然地,应该采取措施,以便实际指定的矢量候选者指的是相同类型的事先编码图片,即,时间上在前面的或视点方向上在前面的,或者甚至添加到列表也要受条件限制并据此执行。

[0200] 根据上述方法1,从属视点重构器26被配置为针对区块50<sub>c</sub>提取指定参考图片(包括参考视点20的当前图片32t<sub>1</sub>(T)和从属视点22的已经解码图片32t<sub>2</sub>(t<T))的列表的参考图片的另外的参考图片索引,并且从属视点重构器26可以被配置为利用作为从属视点22的已经解码图片之一的参考图片使用由参考图片索引指定的从属视点的一个已经解码图片作为参考来执行运动补偿预测,并且入口参考图片是参考视点的当前图片32t<sub>1</sub>(T),则将所确定的视差矢量102或从所确定的视差矢量102推导的修改视差矢量添加到视差矢量预测候选者的列表中,从多视点数据流中提取指定视差矢量预测因子候选者的列表的视差矢量预测因子候选者的索引信息并通过使用取决于指定视差矢量预测候选者的视差矢量执行区块50C的视差补偿预测使用参考视点20的当前图片32t<sub>1</sub>(T)作为参考来重构区块50C。再者,可消除运动补偿和视差补偿预测之间的差。相同的预测模式可针对区块50C用信号传递。关于运动补偿或视差补偿预测是否实际由从属视点重构器26执行,可由参考图片索引定义,该参考图片索引索引到缓冲器或包含时间前导子(即,从属视点的事先解码图片)以及视点前导子(即,其他视点的事先解码图片)这两者的参考图片列表中。

[0201] 如从方法2变得清楚的,从属视点重构器26还可被配置为经由运动/视差矢量预测因子候选者的列表来执行运动矢量预测因子候选者的列表的推导,所述运动/视差矢量预测因子候选者的列表是运动/视差参数候选者的列表,各自包括数个假设,按照假设,运动/视差运动矢量和参考索引指定刚才提及的参考图片的共同列表中的参考图片。从属视点重构器然后可以被配置为将运动/视差参数添加到取决于与确定区块40<sub>R</sub>相关联的运动/视差参数的运动/视差参数候选者的列表,并通过使用取决于由索引信息指定的运动/视差参数候选者的运动/视差参数对区块50<sub>c</sub>执行运动/视差补偿预测来重构区块50<sub>c</sub>。运动参数如上所述可同时确定假设的次数、参考索引、和每次假设的运动/视差矢量差。同样如上文所述,假设的次数事先诸如通过图片的类型来确定。

[0202] 并且如方法3中所述,从属视点重构器另外可以被配置为就运动/视差预测被关注

而言另外采用划分来自区块 $50_R$ 的区块 $50_C$ 。

[0203] 根据后一个方面的图2的编码器会被配置为因此在将多视点信号编码成多视点数据流的方法中起作用。具体地,从属视点重构器26针对区块 $50_C$ 以相同的方式来推导运动矢量预测因子候选者的列表。也就是说,经由与多视点信号的事先编码部分相关联的运动和视差矢量来确定区块 $50_C$ 的视差矢量,其表示从属视点22的当前图片和从属视点22的当前图片处的多视点信号的参考视点20的当前图片之间的视差。然后,使用所确定的视差矢量来确定参考视点的当前图片内的区块 $50_R$ ,并将运动矢量添加到取决于与参考视点的图片的确定区块相关联的运动矢量的运动矢量预测因子候选者的列表中。从属视点重构器针对区块 $50_C$ 将指定运动矢量预测因子候选者的列表的运动矢量预测因子候选者的索引信息插入到多视点数据流中,并通过使用取决于指定运动矢量候选者的运动矢量执行区块 $50_C$ 的运动补偿预测来编码区块 $50_C$ 。

[0204] 虽然在设备的上下文中已经描述了一些方面,但是很显然这些方面还表示相应方法的描述,其中,块或装置与方法步骤或方法步骤的特征对应。同样,在方法步骤的上下文中描述的方面还表示对应块、对应设备的项目或特征的描述。一些或所有方法步骤可由(或利用)硬件设备(例如,微处理器、可编程计算机或电子电路)执行。在一些实施方式中,一些或所有最重要的方法步骤可由这种设备执行。

[0205] 根据某些实施要求,本发明的实施方式可以以硬件或以软件来实施。可使用数字存储介质,例如,具有存储在其上的电子可读控制信号的软盘、DVD、蓝光、CD、ROM、PROM、EPROM、EEPROM或FLASH存储器,执行该实施,这些控制信号与可编程计算机系统配合(或者能够与其配合)以便执行各个方法。因此,数字存储介质可以是计算机可读的。

[0206] 根据本发明的一些实施方式包括数据载体,该载体具有电子可读控制信号,这些控制信号能够与可编程计算机系统配合,这样使得执行在本文中描述的方法。

[0207] 通常,本发明的实施方式可实施为具有程序代码的计算机程序产品,在计算机程序产品运行在计算机上时,该程序代码被操作用于执行一种方法。例如,程序代码可存储在机器可读载体上。

[0208] 其他实施方式包括用于执行在本文中描述的一种方法的计算机程序,该计算机程序存储在机器可读载体上。

[0209] 换言之,因此,本发明方法的实施方式是具有程序代码的计算机程序,程序代码用于在计算机程序运行在计算机上时执行在本文中描述的一种方法。

[0210] 因此,本发明方法的另一个实施方式是数据载体(或数字存储介质或计算机可读介质),该数据载体包括在其上记录的计算机程序,该计算机程序用于执行在本文中描述的一种方法。数据载体、数字存储介质或记录介质一般是有形和/或非暂存的。

[0211] 因此,本发明方法的另一个实施方式是表示计算机程序的数据流或信号序列,用于执行在本文中描述的一种方法。例如,该数据流或信号序列可被配置为通过数据通信连接(例如,通过互联网)来传输。

[0212] 另一个实施方式包括处理器件,例如,计算机或可编程逻辑器件,该处理器件被配置为或者适配为执行在本文中描述的一种方法。

[0213] 另一个实施方式包括在其上安装计算机程序的计算机,用于执行在本文中描述的一种方法。

[0214] 根据本发明的另一个实施方式包括设备或系统,被配置为将用于执行在本文中描述的一种方法的计算机程序(例如,电或光学)传输给接收器。例如,接收器可为计算机、移动装置、存储装置等。该设备或系统例如可包括文件服务器,用于将计算机程序传输给接收器。

[0215] 在一些实施方式中,可编程逻辑装置(例如,现场可编程门阵列)可用于执行在本文中描述的方法的一些或所有功能。在一些实施方式中,现场可编程门阵列可与微处理器配合,以便执行在本文中描述的一种方法。通常,这些方法优选地由任何硬件设备执行。

[0216] 根据本发明的实施方式,包括:

[0217] 1. 一种用于重构被编码到多视点数据流中的多视点信号的装置,包括从属视点重构器(26),所述从属视点重构器(26)被配置为:

[0218] 针对所述多视点信号的从属视点(22)中的当前图片的至少一个区块,通过以下各项来推导运动矢量预测因子候选者的列表:

[0219] 经由与所述多视点信号的先前解码部分相关联的运动和视差矢量来确定用于所述至少一个区块的视差矢量,所述视差矢量表示所述从属视点(22)的所述当前图片与在所述从属视点(22)的所述当前图片的所述至少一个区块上的所述多视点信号的参考视点的当前图片之间的视差;

[0220] 使用所确定的视差矢量来确定所述参考视点的所述当前图片内的区块;以及

[0221] 将运动矢量添加到取决于与所确定的所述参考视点的图片的区块相关联的运动矢量的运动矢量预测因子候选者列表中,以及

[0222] 针对所述从属视点(22)的所述当前图片的所述至少一个区块,从所述多视点数据流提取指定所述运动矢量预测因子候选者列表的一个运动矢量预测因子候选者的索引信息;以及

[0223] 通过使用取决于所指定的运动矢量候选者的运动矢量执行所述从属视点(22)的所述当前图片的所述至少一个区块的运动补偿预测来重构所述从属视点(22)的所述当前图片的所述至少一个区块。

[0224] 2. 根据项1所述的装置,进一步包括:

[0225] 参考视点重构器(24),被配置为基于在用于所述参考视点(20)的所述多视点数据流(14)内传输的运动数据(42)使用运动补偿预测来重构所述多视点信号(12)的所述参考视点(20)的所述当前图片(32t1);以及

[0226] 深度估计器(28),被配置为通过以下各项来估计所述从属视点(22)的所述当前图片(32t2)的深度图:

[0227] 通过将用于所述参考视点(20)的所述运动数据(42)应用到所述参考视点(20)的先前图片的深度图估计(74)上来生成(71)所述参考视点(20)的所述当前图片(32t1)的深度图估计(64);以及

[0228] 将所更新的所述参考视点(20)的所述当前图片(32t1)的深度图估计(74)变形(76)到所述从属视点(22)中以获得所述从属视点(22)的所述当前图片(32t2)的所述深度图估计(64),

[0229] 其中,所述从属视点重构器(26)被配置为在确定用于所述至少一个区块的所述视差矢量时,在所述至少一个区块上的深度数据估计经受深度至视差的转换以获得所确定的

视差矢量。

[0230] 3. 根据项1或2所述的装置,其中,所述从属视点重构器(26)被配置为针对所述从属视点(22)的所述当前图片的所述至少一个区块提取关于所指定的运动矢量候选者的另外的运动矢量差并且执行所述当前图片的所述至少一个区块的所述重构,使得所使用的运动矢量进一步取决于所述运动矢量差和所指定的运动矢量候选者的和。

[0231] 4. 根据前述项中任一项所述的装置,其中,所述从属视点重构器被配置为针对所述从属视点(22)的所述当前图片的所述至少一个区块提取指定包括所述参考视点的当前图片和所述从属视点(22)的已解码的图片的参考图片列表的参考图片的另外的参考图片索引,其中,所述从属视点重构器被配置为如果所述参考图片是所述从属视点(22)的已解码的图片中的一个,则使用所述从属视点(22)的一个已解码的图片作为参考来执行所述运动补偿预测,并且如果所述参考图片是所述参考视点的所述当前图片,则将所确定的视差矢量或从所确定的视差矢量推导出的修改的视差矢量添加到视差矢量预测候选者列表中,从所述多视点数据流提取指定所述视差矢量预测因子候选者列表的一个视差矢量预测因子候选者的索引信息并且通过使用取决于使用所述参考视点的所述当前图片作为参考的所指定的视差矢量候选者的视差矢量对所述从属视点(22)的所述当前图片的所述至少一个区块执行视差补偿预测来重构所述从属视点(22)的所述当前图片的所述至少一个区块。

[0232] 5. 根据前述项中任一项所述的装置,其中,所述从属视点重构器进一步被配置为在推导所述运动矢量预测因子候选者列表时,从所述从属视点(22)的空间上和/或时间上相邻区块在空间上和/或时间上预测另外的运动矢量并将所述另外的运动矢量或从所述另外的运动矢量推导出的版本添加到运动矢量可预测候选者列表中。

[0233] 6. 根据前述项中任一项所述的装置,其中,所述从属视点重构器被配置为经由运动/视差矢量预测因子候选者列表来执行所述运动矢量预测因子候选者列表的推导,所述运动/视差矢量预测因子候选者列表是运动/视差参数候选者列表,所述运动/视差参数候选者列表均包括多个假设,通过假设,运动/视差运动矢量和参考索引指定包括参考视点的当前图片和所述从属视点(22)的先前解码的图片的参考图片列表中的参考图片,其中,所述从属视点重构器被配置为将运动/视差参数添加到取决于与所确定的所述参考视点的所述当前图片的区块相关联的运动/视差参数的所述运动/视差参数候选者列表,并且通过使用取决于由所述索引信息所指定的运动/视差参数候选者的运动/视差参数对所述从属视点(22)的所述当前图片的所述至少一个区块执行运动/视差补偿预测来重构所述从属视点(22)的所述当前图片的所述至少一个区块。

[0234] 7. 一种用于将多视点信号编码到多视点数据流中的装置,包括从属视点编码器,所述从属视点编码器被配置为:

[0235] 针对所述多视点信号的从属视点(22)中的当前图片的至少一个区块,通过以下各项来推导运动矢量预测因子候选者列表:

[0236] 经由与所述多视点信号的先前编码部分相关联的运动和视差矢量来确定用于所述至少一个区块的视差矢量,所述视差矢量表示在所述从属视点(22)的所述当前图片与所述从属视点(22)的当前区块上的所述多视点信号的参考视点的当前图片之间的视差;

[0237] 使用所确定的视差矢量来确定所述参考视点的所述当前图片内的区块;

[0238] 将运动矢量添加到取决于与所确定的所述参考视点的图片的区块相关联的运动

矢量的所述运动矢量预测因子候选者列表中；

[0239] 针对所述从属视点(22)的所述当前图片的所述至少一个区块，将指定所述运动矢量预测因子候选者列表的一个运动矢量预测因子候选者的索引信息插入到所述多视点数据流中；以及

[0240] 通过使用取决于所指定的运动矢量候选者的运动矢量执行所述从属视点(22)的所述当前图片的所述至少一个区块的运动补偿预测来编码所述从属视点(22)的所述当前图片的所述至少一个区块。

[0241] 8.一种用于重构被编码到多视点数据流中的多视点信号的装置，包括：

[0242] 参考视点重构器(24)，被配置为基于在用于参考视点(20)的多视点数据流(14)内传输的运动数据(42)使用运动补偿预测来重构所述多视点信号(12)的所述参考视点(20)的当前图片( $32t_1$ )；以及

[0243] 深度估计器(28)，被配置为通过以下各项来估计从属视点(22)的当前图片( $32t_2$ )的深度图：

[0244] 通过将用于所述参考视点(20)的所述运动数据(42)应用到所述参考视点(20)的先前图片的深度图估计(74)上来生成(71)所述参考视点(20)的所述当前图片( $32t_1$ )的深度图估计(64)；以及

[0245] 将所更新的所述参考视点(20)的所述当前图片( $32t_1$ )的深度图估计(74)变形(76)到所述从属视点(22)中以获得所述从属视点(22)的所述当前图片( $32t_2$ )的所述深度图估计(64)，

[0246] 从属视点重构器(26)，被配置为使用所述深度图估计从所述多视点数据流的从属视点部分重构所述从属视点的所述当前图片。

[0247] 9.根据项8所述的装置，其中，所述运动数据(42)按区块粒度进行定义。

[0248] 10.根据项8或9所述的装置，其中，所述从属视点重构器(26)被配置为基于用于由所述多视点数据流(14)的从属视点部分(38)组成的所述从属视点(22)的视差和/或运动数据(54,60)使用视差和/或运动补偿预测来重构所述从属视点(22)的所述当前图片( $32t_2$ )，

[0249] 其中，所述深度估计器被配置为使用用于所述从属视点(22)的所述视差和/或运动数据(54,60)来更新(77)所述从属视点(22)的所述当前图片的所述深度图估计(64)。

[0250] 11.根据项10所述的装置，其中，所述深度估计器被配置为将所更新的所述从属视点(22)的所述当前图片( $32t_2$ )的深度图估计(74)变形(78)到所述参考视点(20)中以获得所更新的所述参考视点(20)的所述当前图片( $32t_1$ )的深度图估计(74)，其中，所述深度估计器(28)被配置为通过以下各项来估计所述从属视点(22)的下一个图片( $32t_2$ )的所述深度图：

[0251] 通过将所述参考视点(20)的所述下一个图片( $32t_1$ )的所述运动数据(42)应用到所更新的所述参考视点(20)的所述当前图片的深度图估计(74)上来生成(79)所述参考视点(20)的所述下一个图片( $32t_1$ )的深度图估计(64)；以及

[0252] 将所更新的所述参考视点(20)的所述下一个图片( $32t_1$ )的深度图估计(74)变形到所述从属视点(22)中以获得所述从属视点(22)的下一个当前图片( $32t_2$ )的所述深度图估计(64)。

[0253] 12.根据项8至11中任一项所述的装置，其中，所述从属视点重构器被配置为基于

所述从属视点的所述当前图片的所述深度数据估计和所述参考视点的所述当前图片的运动数据来预测所述从属视点的所述当前图片的运动数据,并且基于所预测的运动数据使用运动补偿预测来重构所述从属视点的所述当前图片。

[0254] 13.根据项12所述的装置,其中,当前视点重构器被配置为在预测所述从属视点的所述当前图片的所述运动数据时,使用所述从属视点的所述当前视点的所述深度数据估计来定位所述参考视点的所述当前图片中的对应位置,并且使用所述对应位置处的所述参考视点的所述当前图片的所述运动数据来预测所述从属视点的所述当前图片的所述运动数据。

[0255] 14.根据项8至13中任一项所述的装置,其中,所述从属视点重构器被配置为基于所述从属视点的所述当前图片的所述深度数据估计来针对所述从属视点的所述当前图片预测视差数据,并且基于所预测的视差数据使用视差补偿预测来重构所述从属视点的所述当前图片。

[0256] 15.根据项14所述的装置,其中,所述从属视点重构器被配置为通过将所述深度数据转换成视差矢量,基于所述从属视点的所述当前图片的所述深度数据估计针对所述从属视点的所述当前图片来预测所述视差数据。

[0257] 16.根据项15所述的装置,其中,所述从属视点重构器被配置为使用所述深度图估计作为预测参考来从所述多视点数据流的从属视点部分的从属视点深度图部分重构所述从属视点的所述当前图片的深度图。

[0258] 17.一种用于将多视点信号编码到多视点数据流中的装置,包括:

[0259] 参考视点编码器,被配置为随着经由所述多视点数据流传输用于参考视点的运动数据来基于用于所述参考视点的运动数据使用运动补偿预测来编码所述多视点信号的所述参考视点的当前图片;以及

[0260] 深度估计器,被配置为通过以下各项来估计所述从属视点(22)的所述当前图片( $32t_2$ )的所述深度图:

[0261] 通过将用于所述参考视点(20)的所述运动数据(42)应用到所述参考视点(20)的先前图片的深度图估计(74)上来生成(71)所述参考视点(20)的当前图片( $32t_1$ )的深度图估计(64);以及

[0262] 将所更新的所述参考视点(20)的所述当前图片( $32t_1$ )的深度图估计(74)变形(76)到所述从属视点(22)中以获得所述从属视点(22)的所述当前图片( $32t_2$ )的所述深度图估计(64);

[0263] 从属视点编码器,被配置为使用所述深度图估计将所述从属视点的当前图片编码到所述多视点数据流的从属视点部分。

[0264] 18.一种用于重构被编码到多视点数据流中的多视点信号的方法,包括:

[0265] 针对所述多视点信号的从属视点(22)中的当前图片的至少一个区块,通过以下各项来推导运动矢量预测因子候选者列表:

[0266] 经由与所述多视点信号的先前解码部分相关联的运动和视差矢量来确定用于所述至少一个区块的视差矢量,所述视差矢量表示所述从属视点(22)的所述当前图片与在所述从属视点(22)的所述当前图片的区块中的至少一个区块上的所述多视点信号的参考视点的当前图片之间的视差;

- [0267] 使用所确定的视差矢量来确定所述参考视点的所述当前图片内的区块；
- [0268] 将运动矢量添加到取决于与所述参考视点的图片的所确定的区块相关联的运动矢量的所述运动矢量预测因子候选者列表；以及
- [0269] 针对所述从属视点(22)的所述当前图片的所述至少一个区块，从所述多视点数据流提取指定所述运动矢量预测因子候选者列表的一个运动矢量预测因子候选者的索引信息；以及
- [0270] 通过使用取决于所指定的运动矢量候选者的运动矢量执行所述从属视点(22)的所述当前图片的所述至少一个区块的运动补偿预测来重构所述从属视点(22)的所述当前图片的所述至少一个区块。
- [0271] 19.一种用于将多视点信号编码到多视点数据流中的方法，包括：
- [0272] 针对所述多视点信号的从属视点(22)中的当前图片的至少一个区块，通过以下各项推导运动矢量预测因子候选者列表：
- [0273] 经由与所述多视点信号的先前编码部分相关联的运动和视差矢量来确定用于所述至少一个区块的视差矢量，所述视差矢量表示所述从属视点(22)的所述当前图片与在所述从属视点(22)的所述当前区块上的所述多视点信号的参考视点的当前图片之间的视差；
- [0274] 使用所确定的视差矢量来确定所述参考视点的所述当前图片内的区块；
- [0275] 将运动矢量添加到取决于与所确定的所述参考视点的图片的区块相关联的运动矢量的所述运动矢量预测因子候选者列表中，以及
- [0276] 针对所述从属视点(22)的所述当前图片的所述至少一个区块，将指定所述运动矢量预测因子候选者列表的一个运动矢量预测因子候选者的索引信息插入到所述多视点数据流中；以及
- [0277] 通过使用取决于所指定的运动矢量候选者的运动矢量执行所述从属视点(22)的所述当前图片的所述至少一个区块的运动补偿预测来编码所述从属视点(22)的所述当前图片的所述至少一个区块。
- [0278] 20.一种用于重构被编码到多视点数据流中的多视点信号的方法，包括：
- [0279] 基于在用于参考视点(20)的所述多视点数据流(14)内传输的运动数据(42)使用运动补偿预测来重构所述多视点信号(12)的所述参考视点(20)的当前图片( $32t_1$ )；以及
- [0280] 通过以下各项来估计从属视点(22)的当前图片( $32t_2$ )的深度图：
- [0281] 通过将用于所述参考视点(20)的所述运动数据(42)应用到所述参考视点(20)的先前图片的深度图估计(74)上来生成(71)所述参考视点(20)的所述当前图片( $32t_1$ )的深度图估计(64)；以及
- [0282] 将所更新的所述参考视点(20)的所述当前图片( $32t_1$ )的深度图估计(74)变形(76)到所述从属视点(22)中以获得所述从属视点(22)的所述当前图片( $32t_2$ )的深度图估计(64)，
- [0283] 使用所述深度图估计重构来自所述多视点数据流的从属视点部分的所述从属视点的所述当前图片。
- [0284] 21.一种用于将多视点信号编码到多视点数据流中的方法，包括：
- [0285] 随着经由所述多视点数据流传输用于参考视点的运动数据，基于用于所述参考视点的所述运动数据使用运动补偿预测来编码所述多视点信号的参考视点的当前图片；以及

[0286] 通过以下各项来估计从属视点(22)的当前图片( $32t_2$ )的深度图：

[0287] 通过将用于所述参考视点(20)的所述运动数据(42)应用到所述参考视点(20)的先前图片的深度图估计(74)上来生成(71)所述参考视点(20)的所述当前图片( $32t_1$ )的深度图估计(64)；以及

[0288] 将所更新的所述参考视点(20)的所述当前图片( $32t_1$ )的深度图估计(74)变形(76)到所述从属视点(22)中以获得所述从属视点(22)的所述当前图片( $32t_2$ )的所述深度图估计(64)；

[0289] 使用所述深度图估计将所述从属视点的所述当前图片编码到所述多视点数据流的从属视点部分中。

[0290] 22.一种计算机程序,具有程序代码,当在计算机上运行所述程序代码时,所述程序代码用于执行根据项18或21所述的方法。

[0291] 上述实施方式仅仅用于说明本发明的原理。要理解的是,在本文中描述的配置和细节的修改和变形对于本领域的技术人员是显而易见的。因此,其目的在于,仅仅由未决专利权利要求的范围限制,而不由本文实施方式的描述和说明的方式所呈现的特定细节通来限制。

[0292] 参考文献

[0293] [1] ITU-T and ISO/IEC JTC 1, "Advanced video coding for generic audiovisual services," ITU-T Recommendation H.264 and ISO/IEC 14496-10 (MPEG-4 AVC), 2010.

[0294] [2] A. Vetro, T. Wiegand, G. J. Sullivan, "Overview of the Stereo and Multiview Video Coding Extension of the H.264/PEG-4 AVC Standard", Proceedings of IEEE, vol. 99, no. 4, pp. 626-642, Apr. 2011.

[0295] [3] H. Schwarz, D. Marpe, T. Wiegand, "Overview of the Scalable Video Coding Extension of the H.264/AVC Standard", IEEE Transactions on Circuits and Systems for Video Technology, vol. 17, no. 9, pp. 1103-1120, Sep. 2007.

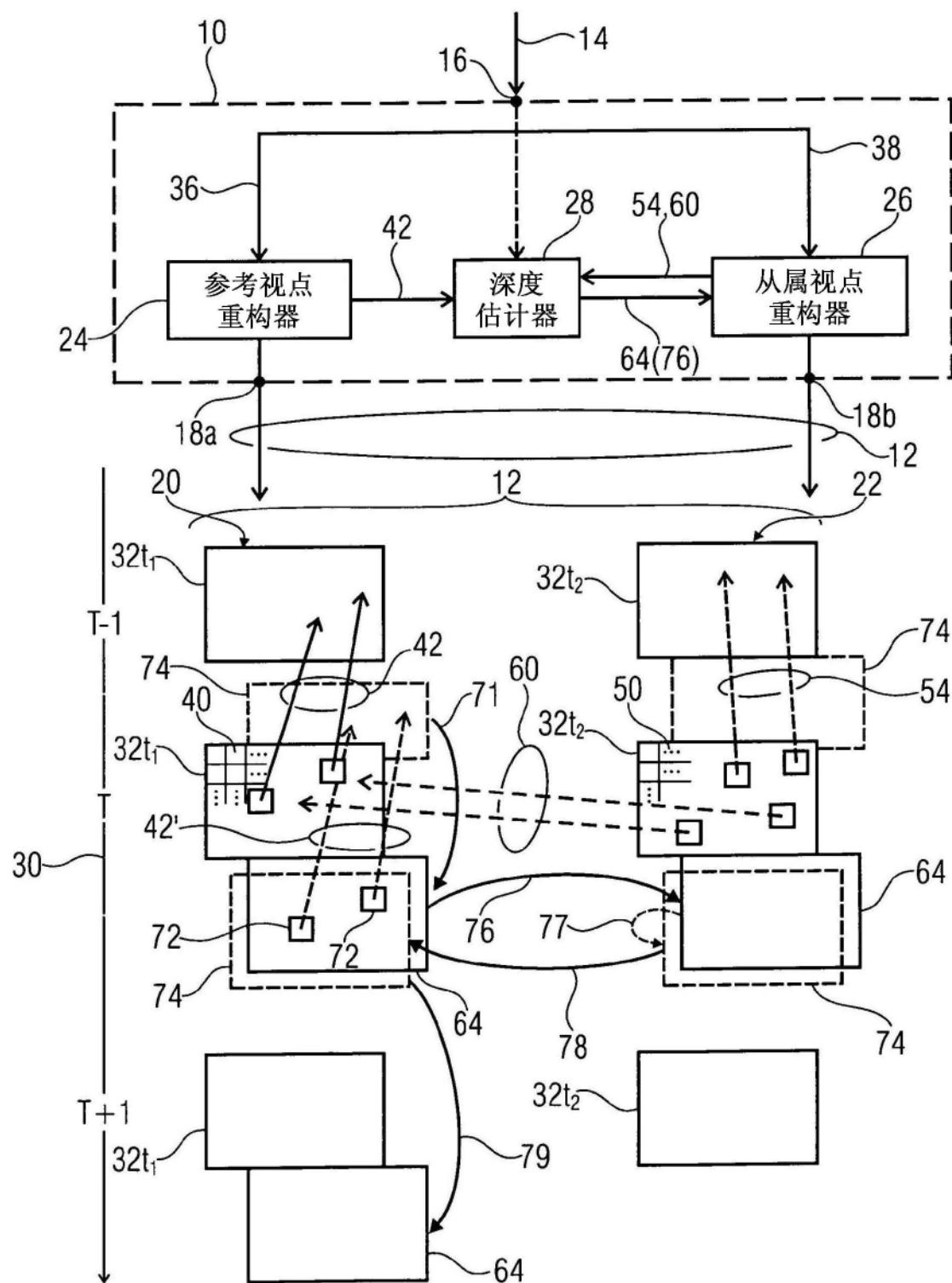


图1

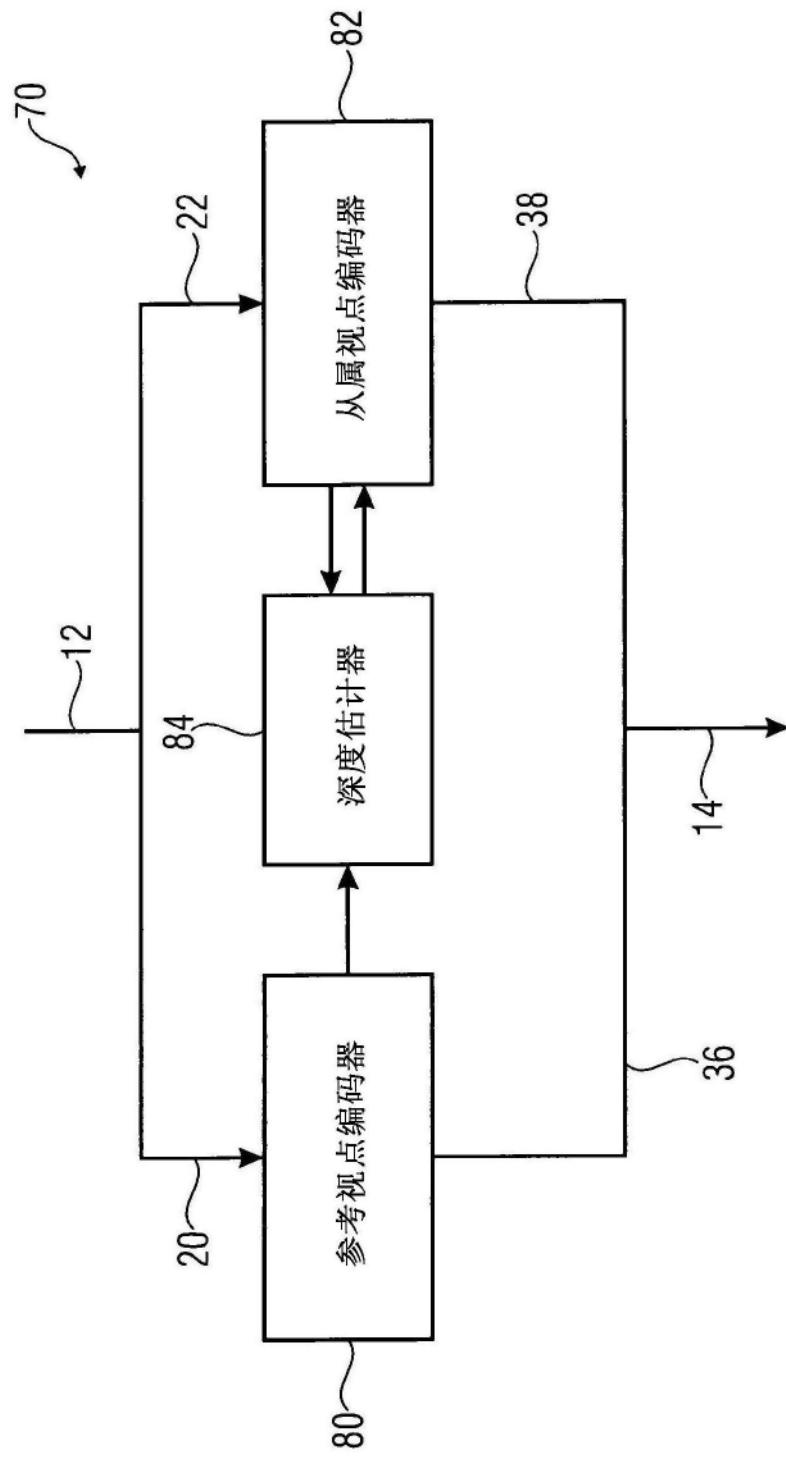


图2

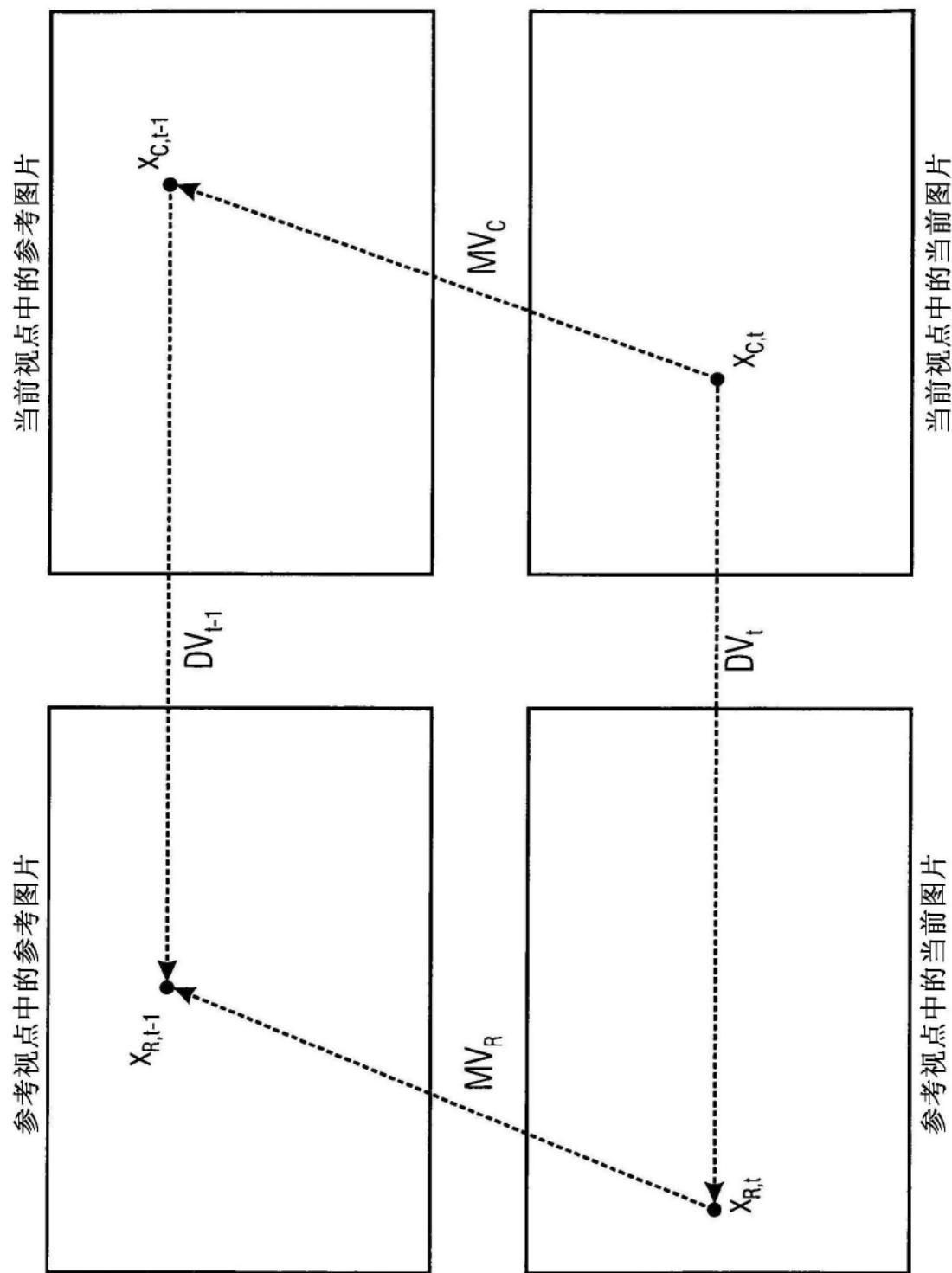


图3

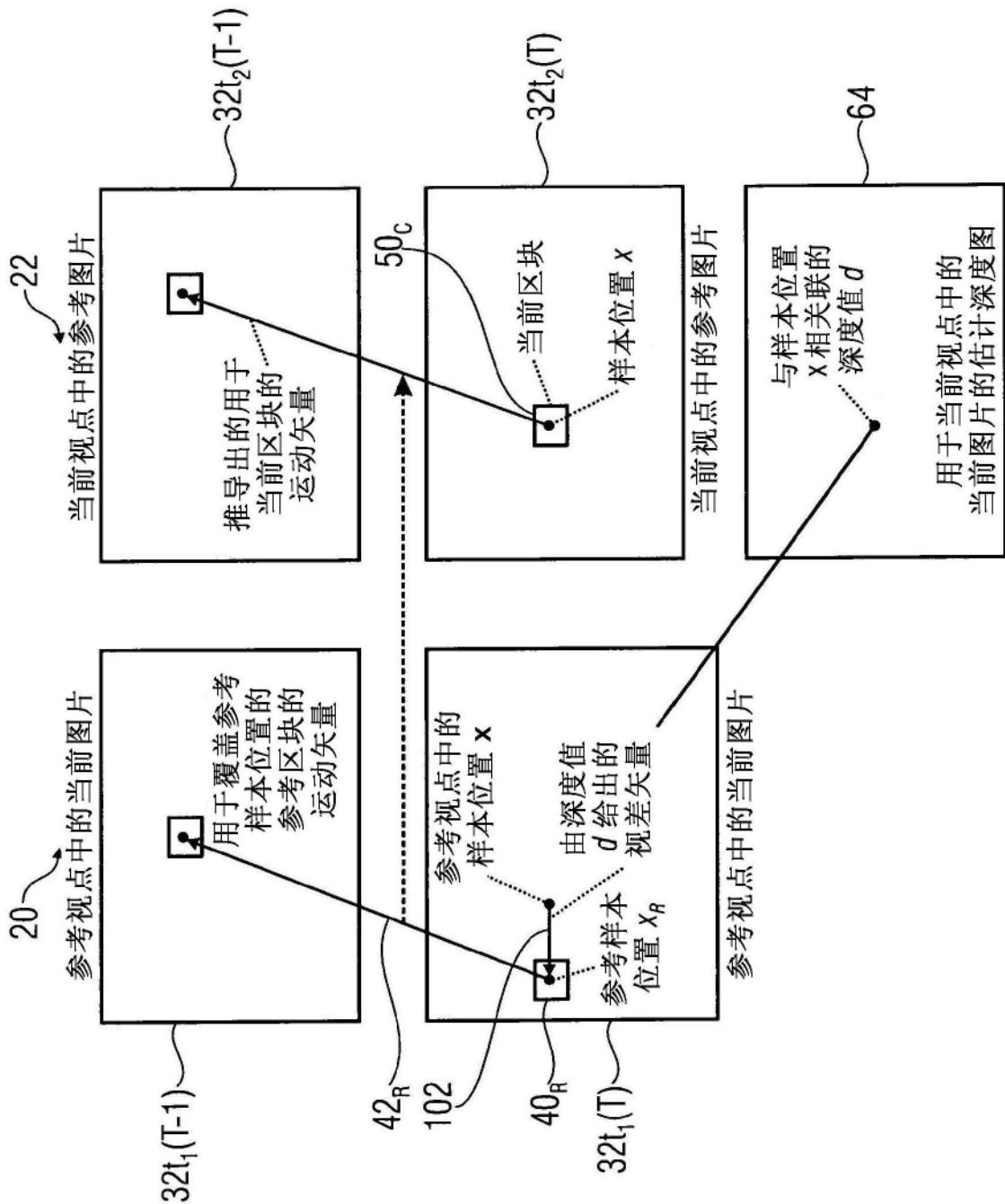


图4

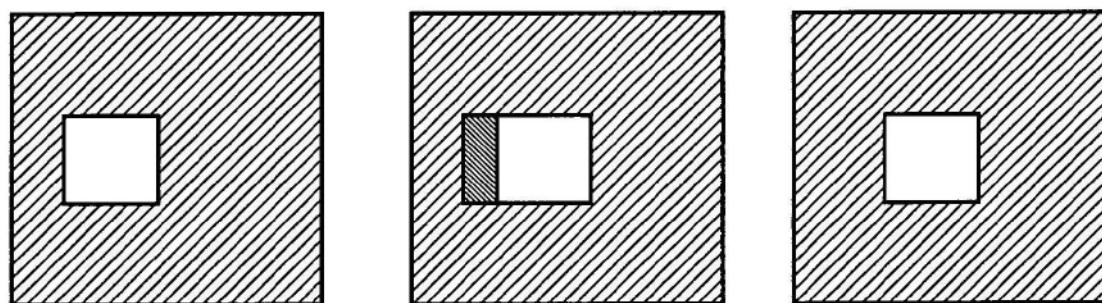


图5

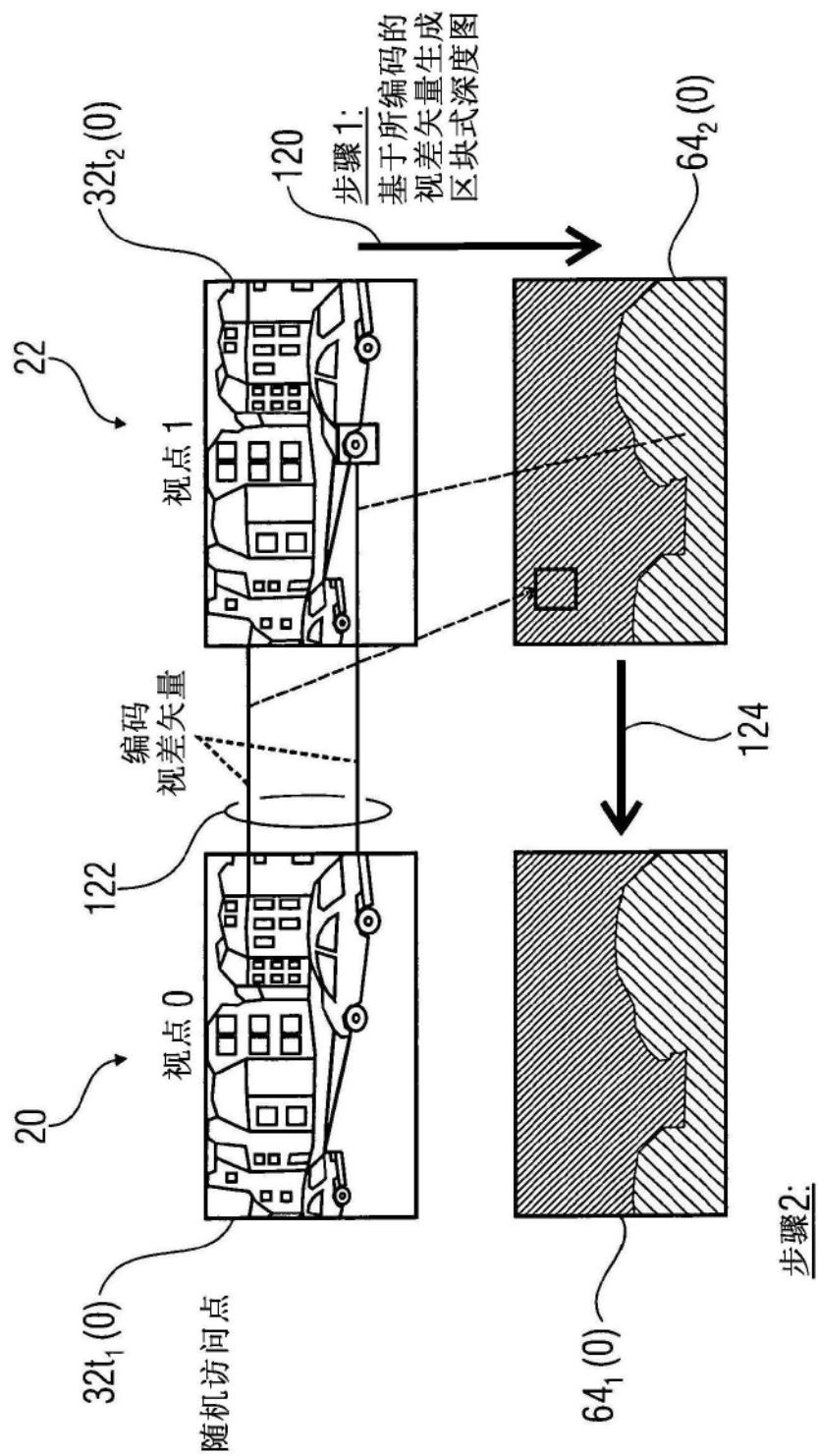


图6

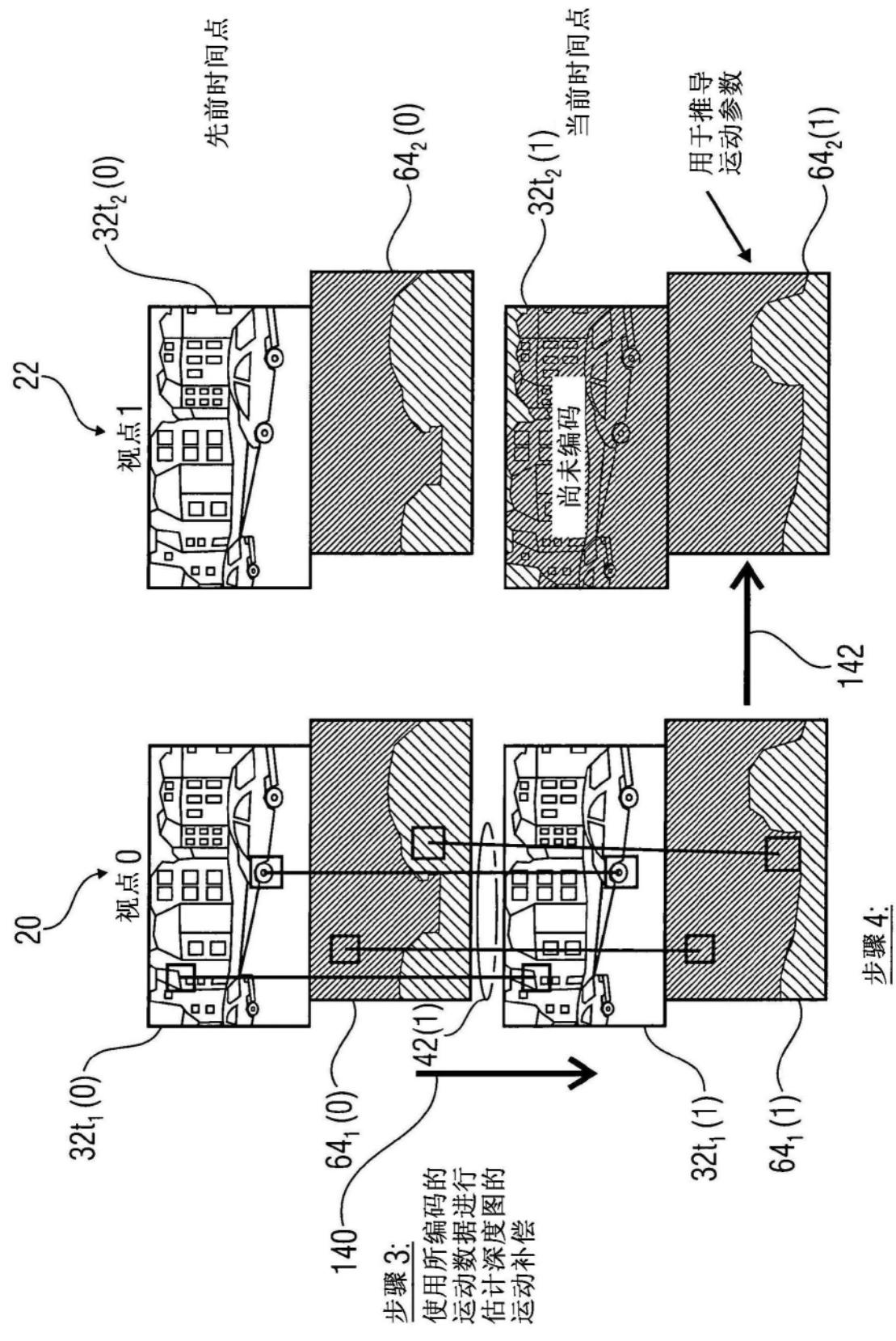


图7

