



# (12)发明专利

(10)授权公告号 CN 106716490 B

(45)授权公告日 2020.03.13

(21)申请号 201580047818.1

(22)申请日 2015.09.04

(65)同一申请的已公布的文献号  
申请公布号 CN 106716490 A

(43)申请公布日 2017.05.24

(30)优先权数据  
62/048,041 2014.09.09 US  
14/845,076 2015.09.03 US

(85)PCT国际申请进入国家阶段日  
2017.03.06

(86)PCT国际申请的申请数据  
PCT/US2015/048569 2015.09.04

(87)PCT国际申请的公布数据  
W02016/040161 EN 2016.03.17

(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 菲利普·格拉斯米格  
迪特尔·施马尔施蒂格  
格哈德·赖特迈尔

(74)专利代理机构 北京律盟知识产权代理有限公司  
代理人 宋献涛

(51)Int.Cl.

G06T 9/00(2006.01)

H04N 19/30(2014.01)

H04N 19/23(2014.01)

H04N 19/27(2014.01)

(56)对比文件

CN 1402852 A,2003.03.12,说明书第28,54-55,80页.

US 2006132482 A1,2006.06.22,说明书第21,74段.

WO 2013130208 A1,2013.09.06,全文.

Christian Feller等.Model-based video compression for real world data.《2013 IEEE Third international conference on consumer electronics-Berlin》.2013,210-211.

H.G. Musmann等.Object-oriented analysis-synthesis coding of moving images.《Signal Processing: Image Communication》.1989,第1卷(第2期),118,124-131.

J. Ahlberg等.Face tracking for model-based coding and face animation.《International Journal of Imaging Systems and Technology》.2003,第13卷(第1期),11-12.

审查员 严开元

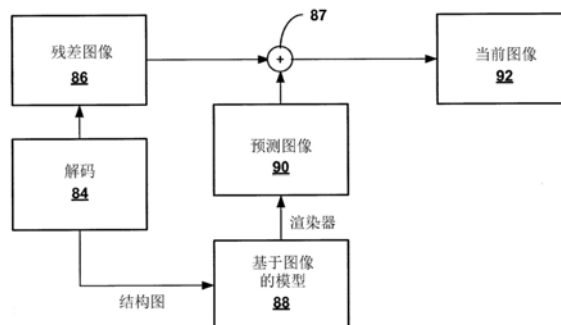
权利要求书4页 说明书20页 附图9页

(54)发明名称

用于视频译码的同时定位与地图构建

(57)摘要

本发明描述视频编码及解码技术,其中预测图像是由将复合图像纹理映射到代理几何结构形成,所述代理几何结构提供当前图像或先前经编码或解码图像的三维结构的近似值。所述预测图像与所述当前图像之间的残差用于编码或解码所述当前图像。



1. 一种解码视频数据的方法,所述方法包括:

基于复合图像及场景结构图而产生合成图像,其中所述复合图像是由先前经解码的一或多个图像构造,其中所述场景结构图包括所述视频数据的当前图像的场景结构图或先前经解码的所述视频数据的图像的场景结构图,其中所述场景结构图包含所述当前图像或先前经解码的所述图像内的三维点的坐标值,所述三维点的坐标值指示所述点的位置和相对深度,其中产生所述合成图像包括利用所述当前图像的摄像机位置及定向信息来渲染所述合成图像,使得所述合成图像和所述当前图像的摄像机位置及定向相同,且其中产生所述合成图像进一步包括:

互连所述场景结构图的点以形成代理几何结构;

将所述复合图像纹理映射到所述代理几何结构以形成基于图像的模型;及

渲染所述基于图像的模型以产生所述合成图像;

确定残差图像,其中所述残差图像指示所述当前图像与所述合成图像之间的差异,且其中确定所述残差图像包括基于包括背景静态部分的所述当前图像的一或多个部分来确定所述残差图像;及

基于所述合成图像及所述残差图像而重构所述当前图像。

2. 根据权利要求1所述的方法,其中确定所述残差图像包括:接收所述当前图像的所述残差图像。

3. 根据权利要求1所述的方法,其中纹理映射包括:基于一或多个先前经解码图像的摄像机位置及定向信息而将所述复合图像纹理映射到所述代理几何结构。

4. 根据权利要求1所述的方法,其进一步包括:

接收所述当前图像的所述场景结构图的信息或指示所述当前图像的所述场景结构图与先前经解码的所述图像的所述场景结构图之间的差异的信息中的一者;及

基于所述所接收信息而产生所述场景结构图,

其中产生所述合成图像包括:基于所述复合图像及所述所产生场景结构图而产生所述合成图像。

5. 根据权利要求1所述的方法,其进一步包括:

接收关于用于构造所述复合图像的所述一或多个先前经解码图像及所述当前图像的摄像机位置及摄像机定向中的一者或两者,

其中产生所述合成图像包括:基于所述一或多个先前经解码图像及所述当前图像的所述摄像机位置及所述摄像机定向中的一者或两者而产生所述合成图像。

6. 一种编码视频数据的方法,所述方法包括:

基于复合图像及场景结构图而产生合成图像,其中所述复合图像是由先前经编码的一或多个图像构造,其中所述场景结构图包括所述视频数据的当前图像的场景结构图或先前经编码的所述视频数据的图像的场景结构图,其中所述场景结构图包含所述当前图像或先前经编码的所述图像内的三维点的坐标值,所述三维点的坐标值指示所述点的位置和相对深度,其中产生所述合成图像包括利用所述当前图像的摄像机位置及定向信息来渲染所述合成图像,使得所述合成图像和所述当前图像的摄像机位置及定向相同,且其中产生所述合成图像进一步包括:

互连所述场景结构图的点以形成代理几何结构;

将所述复合图像纹理映射到所述代理几何结构以形成基于图像的模型；及  
渲染所述基于图像的模型以产生所述合成图像；

基于所述合成图像及所述当前图像而确定残差图像，其中所述残差图像指示所述当前图像与所述合成图像之间的差异，且其中确定所述残差图像包括基于包括背景静态部分的所述当前图像的一或多个部分来确定所述残差图像；及

输出指示所述残差图像的信息以编码所述视频数据的所述当前图像。

7. 根据权利要求6所述的方法，其进一步包括：

确定关于用于构造所述复合图像的所述一或多个先前经编码图像的摄像机位置及定向信息及关于所述当前图像的摄像机位置及定向信息，

其中纹理映射包括：基于所述一或多个先前经编码图像的所述摄像机位置及定向信息而将所述复合图像纹理映射到所述代理几何结构。

8. 根据权利要求7所述的方法，其进一步包括：

利用同时定位与地图构建SLAM技术确定所述当前图像的所述摄像机位置，

其中确定关于所述一或多个先前经编码图像的所述摄像机位置包括：利用所述SLAM技术来确定关于所述一或多个先前经编码图像的所述摄像机位置。

9. 根据权利要求6所述的方法，其进一步包括：

利用同时定位与地图构建SLAM技术来产生所述场景结构图。

10. 根据权利要求6所述的方法，其进一步包括：

确定所述当前图像的所述场景结构图的信息或指示所述当前图像的所述场景结构图与先前经编码的所述图像的所述场景结构图之间的差异的信息中的一者；及

输出所述经确定信息。

11. 根据权利要求6所述的方法，其中所述合成图像包括第一合成图像，所述方法进一步包括：

输出关于用于构造所述复合图像的所述一或多个先前经编码图像及关于所述当前图像的摄像机位置及摄像机定向中的一者或两者，其中所述摄像机位置及所述摄像机定向中的所述一者或两者由解码器处理器用于产生与所述第一合成图像相同的第二合成图像。

12. 根据权利要求6所述的方法，其进一步包括：

确定所述当前图像的前景非静态部分及所述当前图像的背景静态部分，

其中输出所述残差图像包括：输出第一层中的所述残差图像，所述第一层不同于包含关于所述当前图像的所述前景非静态部分的残差数据的第二层。

13. 一种用于译码视频数据的装置，所述装置包括：

视频存储器，其经配置以存储先前经译码且用于构造复合图像的一或多个图像；及

译码器处理器，其经配置以：

基于所述复合图像及场景结构图而产生合成图像，其中所述场景结构图包括所述视频数据的当前图像的场景结构图或先前经译码的所述视频数据的图像的场景结构图，其中所述场景结构图包含所述当前图像或先前经译码的所述图像内的三维点的坐标值，所述三维点的坐标值指示所述点的位置和相对深度，其中为了产生所述合成图像，所述译码器处理器经配置以利用所述当前图像的摄像机位置及定向信息来渲染所述合成图像，使得所述合成图像和所述当前图像的摄像机位置及定向相同，且其中为了产生所述合成图像，所述译

码器处理器进一步经配置以：

互连所述场景结构图的点以形成代理几何结构；

将所述复合图像纹理映射到所述代理几何结构以形成基于图像的模型；及

渲染所述基于图像的模型以产生所述合成图像；及

基于所述当前图像的残差图像而译码所述当前图像，其中所述残差图像指示所述当前图像与所述合成图像之间的差异，且其中所述残差图像是基于包括背景静态部分的所述当前图像的一或多个部分。

14. 根据权利要求13所述的装置，其中所述译码器处理器包括解码器处理器，且其中所述解码器处理器经配置以接收所述当前图像的所述残差图像，且其中为译码所述当前图像，所述解码器处理器经配置以通过基于所述合成图像及所述残差图像而重构所述当前图像而解码所述当前图像。

15. 根据权利要求14所述的装置，其中所述解码器处理器经配置以：

接收所述当前图像的所述场景结构图的信息或指示所述当前图像的所述场景结构图与所述先前经译码图像的所述场景结构图之间的差异的信息中的一者；及

基于所述所接收信息而产生所述场景结构图，

其中为产生所述合成图像，所述解码器处理器经配置以基于所述复合图像及所述所产生场景结构图而产生所述合成图像。

16. 根据权利要求14所述的装置，其中所述解码器处理器经配置以：

接收关于用于构造所述复合图像的所述一或多个先前经译码图像及所述当前图像的摄像机位置及摄像机定向中的一者或两者，

其中为产生所述合成图像，所述解码器处理器经配置以基于所述一或多个先前经译码图像及所述当前图像的所述摄像机位置及所述摄像机定向中的一者或两者而产生所述合成图像。

17. 根据权利要求13所述的装置，其中所述译码器处理器包括编码器处理器，其中为了译码所述当前图像，所述编码器处理器经配置以：

基于所述合成图像及所述当前图像而确定所述残差图像；及

输出指示所述残差图像的信息以编码所述视频数据的所述当前图像。

18. 根据权利要求17所述的装置，其中所述编码器处理器经配置以利用同时定位与地图构建SLAM技术来产生所述场景结构图。

19. 根据权利要求17所述的装置，其中所述编码器处理器经配置以：

确定所述当前图像的所述场景结构图的信息或指示所述当前图像的所述场景结构图与所述先前经译码图像的所述场景结构图之间的差异的信息中的一者；及

输出所述经确定信息。

20. 根据权利要求17所述的装置，其中所述合成图像包括第一合成图像，且其中所述编码器处理器经配置以：

利用同时定位与地图构建SLAM技术来确定关于所述一或多个先前经译码图像的摄像机位置；及

输出关于用于构造所述复合图像的一或多个先前经译码图像及关于所述当前图像的所述摄像机位置及摄像机定向中的一者或两者，其中所述摄像机位置及所述摄像机定向中

的所述一者或两者由解码器处理器用于产生与所述第一合成图像相同的第二合成图像。

21. 根据权利要求13所述的装置, 其中为产生所述合成图像, 所述译码器处理器包括图形处理单元GPU, 其中所述GPU经配置以产生所述合成图像。

22. 根据权利要求13所述的装置, 其中为纹理映射, 所述译码器处理器经配置以基于一个或多个先前经译码图像的摄像机位置及定向信息而将所述复合图像纹理映射到所述代理几何结构。

23. 一种非暂时性计算机可读存储媒体, 其上存储有指令, 所述指令在被执行时致使用于译码视频数据的装置的一或多个处理器进行以下操作:

基于复合图像及场景结构图而产生合成图像, 其中所述复合图像是由先前经译码的一个或多个图像构造, 其中所述场景结构图包括所述视频数据的当前图像的场景结构图或先前经译码的所述视频数据的图像的场景结构图, 其中所述场景结构图包含所述当前图像或先前经译码的所述图像内的三维点的坐标值, 所述三维点的坐标值指示所述点的位置和相对深度, 其中致使所述一或多个处理器产生所述合成图像的指令包括致使所述一或多个处理器利用所述当前图像的摄像机位置及定向信息来渲染所述合成图像, 使得所述合成图像和所述当前图像的摄像机位置及定向相同的指令, 且其中致使所述一或多个处理器产生所述合成图像的指令进一步包括致使所述一或多个处理器进行以下操作的指令:

互连所述场景结构图的点以形成代理几何结构;

将所述复合图像纹理映射到所述代理几何结构以形成基于图像的模型; 及

渲染所述基于图像的模型以产生所述合成图像; 及

基于所述当前图像的残差图像而译码所述当前图像, 其中所述残差图像指示所述当前图像与所述合成图像之间的差异, 且其中所述残差图像是基于包括背景静态部分的所述当前图像的一或多个部分。

## 用于视频译码的同时定位与地图构建

[0001] 本申请案主张在2014年9月9日提出申请的美国临时申请案62/048,041的权益,所述美国临时申请案的全部内容以引用的方式并入本文中。

### 技术领域

[0002] 本发明涉及视频译码。

### 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、平板计算机、电子书阅读器、数字摄像机、数字记录装置、数字媒体播放器、视频游戏装置、视频游戏控制台、蜂窝式或卫星无线电电话、所谓“智能电话”、视频电话会议装置、视频流式处理装置及其类似者。数字视频装置实施视频译码技术,例如由MPEG-2、MPEG-4、ITU-T H.263、ITU-T H.264/MPEG-4第10部分、高级视频译码(AVC)、目前在研发中的高效率视频译码(HEVC)标准及这些标准的扩展所定义的标准中所描述之那些技术。视频装置可通过实施这些视频译码技术而更高效地发射、接收、编码、解码及/或存储数字视频信息。

### 发明内容

[0004] 本发明描述利用当前图像及场景结构图的合成图像来编码或解码所述当前图像的用于视频译码的技术。所述场景结构图可包含图像的一组稀少的点,所述点使用多个多边形(例如,三角形)互连以形成代理几何结构。所述场景结构图可包含当前图像或先前经译码(例如,经编码或经解码)图像的一组稀少的点。所述合成图像为经映射到代理几何结构的一或多个先前经译码图像(例如,叠加在代理几何结构上的一或多个先前经译码图像)。

[0005] 编码器处理器确定合成图像与当前图像之间的差异,且发信通知所得的残差图像(例如,指示残差图像的输出信息)。另外,编码器处理器间或地发信通知场景结构图的信息且可发信通知关于每一图像的摄像机位置及定向信息(摄像机姿势信息)。发信通知摄像机姿势信息可并非在所有情况中为必需的。解码器处理器接收场景结构图的信息。所述解码器处理器重构所述合成图像,使用用以形成所述代理几何结构的所述场景结构图及一或多个先前经解码的图像。解码器处理器还确定残差数据(例如,基于来自编码器处理器的指示残差图像的信息)、可能摄像机姿势信息并基于所述残差图像、可能摄像机姿势信息及合成图像而重构所述当前图像。在一些实例中,所述编码器处理器可不针对每一图像产生并发信通信所述场景结构图。在这些实例中,所述编码器处理器及所述解码器处理器重新使用先前场景结构图。所述编码器处理器可间或地发信通知所述场景结构图的更新。

[0006] 在一个实例中,本发明描述一种解码视频数据的方法,所述方法包括:基于复合图像及场景结构图而产生合成图像,其中所述复合图像是由先前经解码的一或多个图像构造,其中所述场景结构图包括所述视频数据的当前图像的场景结构图或先前经解码的所述

视频数据的图像的的场景结构图,且其中所述场景结构图包含所述当前图像或先前经解码的所述图像内的三维点的坐标值;确定所述视频数据的所述当前图像的残差图像,其中所述残差图像指示所述当前图像与所述合成图像之间的差异;及基于所述合成图像及所述残差图像而重构所述当前图像。

[0007] 在一个实例中,本发明描述一种编码视频数据的方法,所述方法包括:基于复合图像及场景结构图而产生合成图像,其中所述复合图像是由先前经编码的一或多个图像构造,其中所述场景结构图包括所述视频数据的当前图像的场景结构图或先前经编码的所述视频数据的图像的的场景结构图,且其中所述场景结构图包含所述当前图像或先前经编码的所述图像内的三维点的坐标值;基于所述合成图像及所述当前图像而确定残差图像,其中所述残差图像指示所述当前图像与所述合成图像之间的差异;及输出指示所述残差图像的信息以编码所述视频数据的所述当前图像。

[0008] 在一个实例中,本发明描述一种用于编码视频数据的装置,所述装置包括:视频存储器,所述视频存储器经配置以存储先前经译码且用于构造复合图像的一或多个图像;及译码器处理器。所述译码器处理器经配置以基于所述复合图像及场景结构图而产生合成图像,其中所述场景结构图包括所述视频数据的当前图像的场景结构图或先前经译码的所述视频数据的图像的的场景结构图,且其中所述场景结构图包含所述当前图像或先前经译码的所述图像内的三维点的坐标值,及基于所述当前图像的残差图像而译码所述当前图像,其中所述残差图像指示所述当前图像与所述合成图像之间的差异。

[0009] 在一个实例中,本发明描述一种计算机可读存储媒体,其上存储有指令,所述指令在被执行时致使装置的一或多个处理器对视频数据进行译码以基于复合图像及场景结构图而产生合成图像的,其中所述复合图像是由先前经译码的一或多个图像构造,其中所述场景结构图包括所述视频数据的当前图像的场景结构图或先前经译码的所述视频数据的图像的的场景结构图,且其中所述场景结构图包含所述当前图像或先前经译码的所述图像内的三维点的坐标值,及基于所述当前图像的残差图像而译码所述当前图像,其中所述残差图像指示所述当前图像与所述合成图像之间的差异。

[0010] 在一个实例中,本发明描述一种用于对视频数据进行译码的装置,所述装置包括用于基于复合图像及场景结构图而产生合成图像的装置,其中所述复合图像是由先前经译码的一或多个图像构造,其中所述场景结构图包括所述视频数据的当前图像的场景结构图或先前经译码的视频数据的图像的的场景结构图,且其中所述场景结构图包含所述当前图像或先前经译码的所述图像内的三维点的坐标值,及用于基于所述当前图像的残差图像而对所述当前图像进行译码的装置,其中所述残差图像指示所述当前图像与所述合成图像之间的差异。

[0011] 下文的随附图式及描述中阐明一或多个实例的细节。根据所述描述、图式以及权利要求书将明了其它特征、目的及优点。

## 附图说明

[0012] 图1为说明经配置以实施或以其它方式利用本发明中所描述的一或多个实例视频译码技术的实例视频编码及解码系统的框图。

[0013] 图2为说明经配置以实施或以其它方式利用本发明中所描述的一或多个实例视频

编码技术的编码器处理器的实例的框图。

[0014] 图3A及3B为说明场景结构图的实例的图形图。

[0015] 图4为说明场景结构图的点的互连的实例的概念图。

[0016] 图5为说明根据本发明的一或多个实例视频编码技术的数据流程图。

[0017] 图6为说明根据本发明中所描述的技术的视频编码的实例的概念图。

[0018] 图7为说明经配置以实施或以其它方式利用本发明中所描述的一或多个实例视频解码技术的解码器处理器的实例的框图。

[0019] 图8为说明根据本发明的一或多个实例视频解码技术的数据流程图。

[0020] 图9为说明视频编码的实例方法的流程图。

[0021] 图10为说明视频解码的实例方法的流程图。

### 具体实施方式

[0022] 本发明描述用于基于合成图像及场景结构图对当前图像进行视频译码(例如,编码或解码)的各种技术。以此方式,至少针对当前图像的部分,视频译码可能并不需要常规运动补偿。如本发明中别处更详细描述,构造场景结构图所需要的信息量可相对低,且合成图像可为使用标准硬件用场景结构图及先前图像构造。因此,本发明中所描述的技术可相对于常规运动补偿技术减少视频译码所需要的带宽。另外,由于技术可在标准硬件上实施,因此实施本发明中描述的技术可不需要任何额外组件。

[0023] 场景结构图包含图像内的一些点(例如,关键点)的坐标值。例如,图像可为三维(3D)图像。在此状况下,3D图像并非立体图像,而确切地说为其内存在对象(例如,尽管图像囊括2D空间,但对象似乎相对于另一对象进一步在后面)的相对深度的图像。

[0024] 场景结构图可包含关于每一点的x、y及z坐标。场景结构图的点可界定多边形(例如,三角形)的顶点,所述多边形在经连接在一起时形成界定图像的结构网格。在此情景中,场景结构图的互连可被视为图像的代理几何结构。在本发明中,用于形成代理几何结构的场景结构图可为当前图像的场景结构图或先前经编码或解码图像的场景结构图,如下文更详细所描述。

[0025] 可存在产生场景结构图的各种方式。产生场景结构图的一个实例方式为使用同时定位与地图构建(SLAM)技术。然而,本发明中所描述的技术可扩展成产生场景结构图的其它方式。出于说明的目的,关于SLAM技术描述用于产生场景结构图的技术。举例来说,从产生场景结构图的SLAM处理器(例如,经配置以实施SLAM技术的处理器)的角度描述本发明中所描述的技术。

[0026] 在一些实例中,SLAM处理器可基于当前图像或基于先前经编码或解码图像而产生用于编码或解码当前图像的场景结构图。举例来说,执行视频编码的装置可包含SLAM处理器。在此实例中,SLAM处理器可基于当前图像而产生当前图像的场景结构图。编码器处理器可基于所产生场景结构图而编码当前图像并输出所产生场景结构图,如在本发明中别处更详细描述。执行视频解码的装置可接收场景结构图且基于场景结构图而解码当前图像,也如本发明中别处更详细地描述。

[0027] 在一些实例中,SLAM处理器可不需要基于当前图像而产生场景结构图。在这些实例中,编码器处理器可基于先前所产生场景结构图(例如,编码器处理器先前已产生的场景



结构图)而编码当前图像。类似地,解码器处理器可基于先前所接收场景结构图(例如,解码器处理器先前已接收的场景结构图)而解码当前图像。

[0028] 除场景结构图外,编码器处理器及解码器处理器使用合成图像来分别编码及解码当前图像。在执行视频编码的装置或执行视频解码的装置内可存在经配置以产生合成图像的组件。出于说明的目的,用产生合成图像的图形处理单元(GPU)描述以下内容。

[0029] 用于视频编码的装置及用于视频解码的装置两者包含相应GPU。GPU中的每一者执行基本上相同功能以产生合成图像。因此,对GPU产生合成图像的方式的以下描述适于用于视频编码的装置及用于视频解码的装置两者。

[0030] GPU可基于一或多个图像(例如,先前经解码图像)及场景结构图而产生合成图像。此外,场景结构图可为由当前图像产生的场景结构图或由先前经编码或解码图像产生的场景结构图。在其中GPU使用多于一个图像来产生合成图像的实例中,GPU或某一其它组件可执行混合操作以混合图像(例如,先前经解码或编码图片)中的两者或多于两者以产生复合图像。在其中GPU仅使用一个图像的实例中,图像可被视为复合图像。

[0031] GPU包含纹理引擎,所述纹理引擎的功能为将纹理图叠加在多边形的互连上。根据本发明中所描述的技术,GPU使用(例如)顶点着色器或输入混编器来互连场景结构图的稀疏填充点。下文提供稀疏填充点的实例。举例来说,场景结构图的点可被视为GPU互连以形成网格的三角形的顶点。GPU可互连场景结构图的点以形成代理几何结构。如上文所描述,场景结构图的点是由 $x$ 、 $y$ 及 $z$ 坐标界定,且因此场景结构图经界定在3D空间中。

[0032] 复合图像(例如,一个图像或两个或多于两个图像的混合)可被视为纹理图。在此实例中,复合图像的内容可为3D,但每一像素是由 $x$ 及 $y$ 坐标识别。为避免与场景结构图的 $x$ 、 $y$ 及 $z$ 坐标混淆,复合图像(例如,纹理图)的像素的 $x$ 及 $y$ 坐标被分别称作 $u$ 及 $v$ 坐标。

[0033] GPU的纹理引擎可将复合图像的每一像素映射到3D场景结构图以将复合图像叠加在场景结构图上。换句话说,纹理引擎将由 $(u,v)$ 坐标界定的像素中的每一者映射到场景结构图的 $(x,y,z)$ 坐标上。纹理引擎的结果可被视为基于图像的模型(IBM)。此IBM模型可为其像素由 $(x,y,z)$ 坐标界定的图形构造。

[0034] GPU可渲染IBM以产生合成图像。举例来说,GPU执行标准图形处理任务以处理IBM以渲染合成图像。合成图像可被视为最终可视图像(例如,其中像素由显示器上的 $(x,y)$ 像素坐标界定)。然而,编码器处理器及解码器处理器可利用合成图像作为预测图像而非输出合成图像供观看。

[0035] 编码器处理器可确定当前图像与合成图像之间的差异(被称作残差图像)并输出指示残差图像的信息。解码器处理器确定残差图像(例如,基于指示残差图像的所接收信息),且将残差图像与局部产生合成图像汇总在一起以重构当前图像。

[0036] 在其中SLAM处理器基于当前图像而产生场景结构图的实例中,编码器处理器可另外输出场景结构图的信息(例如,场景结构图的点的坐标)。在这些实例中,解码器处理器接收场景结构图的信息且解码器处理器侧上的GPU基于所接收场景结构图而产生合成图像。在其中SLAM处理器并未基于当前图像产生场景结构图且GPU替代地重新使用先前场景结构图来产生合成图像的实例中,编码器处理器可不输出场景结构图的信息。解码器处理器侧的GPU可重新使用相同先前场景结构图来产生合成图像。

[0037] 在其中SLAM处理器基于编码器处理器侧上的GPU用于产生合成图像的当前图像而

产生场景结构图的一些实例中,编码器处理器未必需要输出场景结构图的点的坐标。确切地说,编码器处理器可相对于先前场景结构图输出场景结构图的点的渐进式更新。解码器处理器接收场景结构图的这些更新值且GPU更新先前场景结构图以产生GPU用于产生用于解码当前图像的合成图像的场景结构图。在本发明中,场景结构图的信息包含场景结构图的点的坐标或相对于先前场景结构图更新点的坐标的值。

[0038] 本发明中所描述的技术可大体适用于视频编码及解码。作为一个实例,技术可用于其中通过移动摄像机观察场景且其中需要实时视频编码及解码的实例中。举例来说,在视频电话或视频会议中,可实时需要视频编码及解码且本发明中所描述的技术提供此实时视频编码及解码。举例来说,GPU可相对快速地产生合成图像,从而促进实时视频编码及解码,且输出及接收场景结构图所需要的带宽量可相对低。

[0039] 尽管上文经描述关于实时视频编码及解码且例如借助移动摄像机,但应理解技术并不限于此。在一些实例中,视频编码及视频解码技术适于离线视频编码及解码技术(例如,非实时)且适于其中摄像机为静止的视频编码及解码技术。

[0040] 此外,在其中摄像机移动的一些实例中,编码器处理器可将指示摄像机的定向及位置(被称作摄像机姿势)的信息输出到解码器处理器。解码器处理器出于解码(例如,重构)当前图像的目的利用摄像机姿势。举例来说,摄像机姿势信息指示拍摄图像的“位置”。在其中摄像机移动的实例中,关于当前图像的摄像机姿势可不同于关于先前经译码图像的摄像机姿势。因此,为渲染合成图像,相应GPU可利用当前图像的摄像机姿势信息使得关于合成图像及当前图像的摄像机姿势相同。此外,相应GPU可在将复合图像映射到场景结构图时利用关于用于构造复合图像的一或多个先前经译码图像的摄像机姿势信息。

[0041] 在上述实例中,相应GPU使用关于当前图像的摄像机姿势信息来渲染合成图像且使用先前经编码或解码图像中的一或多者的摄像机姿势信息(例如,基于GPU处于编码器侧还是解码器侧)来将复合图像映射到场景结构图。因此,在一些实例中,编码器处理器可使用摄像机姿势信息来编码当前图像且可发信通信关于每一图像的摄像机姿势信息。解码器处理器可接收关于每一图像的摄像机姿势信息且可使用所述摄像机姿势信息来解码当前图像。在其中摄像机并未移动的实例中,摄像机姿势信息可经隐式导出或可经设置成恒定值。在此些实例中,编码器处理器未必需要发信通知关于每一图像的摄像机姿势信息,且解码器处理器未必需要接收关于每一图像的摄像机姿势信息。

[0042] 在上述实例中,解码器处理器可经配置以将相反过程实施为编码器处理器的过程。在本发明中,术语“译码器处理器”可用于广泛地指代解码器处理器或编码器处理器。举例来说,译码器处理器可基于复合图像及场景结构图而产生合成图像。场景结构图包括当前图像的场景结构图或先前经译码(视情况,经编码或经解码)的图像的场景结构图,且场景结构图包含关于当前图像或先前经译码图像内的三维点的坐标值。在此状况下,解码器处理器及编码器处理器两者可经配置以按基本上类似方式产生合成图像,且因此此实例经描述为由译码器处理器执行。

[0043] 译码器处理器可基于当前图像的残差图像而译码(例如,视情况,编码或解码)当前图像。残差图像指示当前图像与合成图像之间的差异。

[0044] 举例来说,针对其中译码器处理器指代解码器处理器的状况,解码器处理器可基于接收残差图像而确定残差图像,解码当前图像,视频解码器可基于合成图像及残差图像

而重构当前图像。针对其中译码器处理器指编码器处理器的状况，编码器处理器可基于合成图像及当前图像而确定残差图像，并输出指示残差图像（由解码器处理器用于解码当前图像的残差图像）的信息以编码当前图像。

[0045] 在本发明中，先前经译码（例如，经编码或经解码）图像可指先前所显示的图像，但技术并不限于此。举例来说，先前经译码图像可为在显示次序中在时间上较早的图片。然而，在一些实例中，先前经译码图像可为在时间上与当前图像（例如，立体图像）同时或非常快地显示的图像。

[0046] 作为实例，针对立体视图，彼此同时或非常快地显示两个或多于两个图像。针对立体视图，图像中的一者可用于译码（例如，编码或解码）立体视图的另一图像。针对此些实例，用于译码的图像为先前经译码图像的实例，这是因为这些图像在当前图像之前经编码或解码，所述当前图像使用这些经译码图像来确定残差。

[0047] 在本发明中，有时，使用术语先前图像。应理解，先前图像并不限于仅其中先前图像比当前图像早显示（但此为一种可能性）的状况。确切地说，先前图像指相对于当前图像先前经编码或解码的图像。先前经编码或解码的图像可在当前图像之前显示，可与当前图像几乎同时地显示（例如，针对立体或3D视图），或可能在当前图像之后显示。此外，在本发明中，术语“图像”及“图片”可交互地使用。

[0048] 图1为说明可经配置以实施或以其它方式利用本发明中所描述的一或多个实例视频译码技术的实例视频编码及解码系统10的框图。在图1的实例中，源装置12包含视频源18、编码器处理器20及输出接口22。目的地装置14包含输入接口28、解码器处理器30及显示装置32。根据本发明，源装置12的编码器处理器20及目的地装置14的解码器处理器30可经配置以实施本发明中所描述的实例技术。在一些实例中，源装置及目的地装置可包含其它组件或布置。举例来说，源装置12可从外部视频源18（例如，外部摄像机）接收视频数据。同样地，目的地装置14可与外部显示装置介接，而非包含整合式显示装置。

[0049] 如图1中所展示，系统10包含提供由目的地装置14解码的经编码视频数据的源装置12。特定来说，源装置12经由链路16将视频数据提供到目的地装置14。源装置12及目的地装置14可包括广泛范围的装置中的任何者，包含桌上型计算机、笔记本（即，膝上型）计算机、平板计算机、机顶盒、例如所谓“智能”电话的电话手机、所谓“智能”板、电视机、摄像机、显示装置、数字媒体播放器、视频游戏主机、视频流式装置或其类似者。在一些状况中，源装置12及目的地装置14可经配备以用于无线通信。

[0050] 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可包括能够使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可包括用以使得源装置12能够将经编码视频数据实时直接地发射到目的地装置14的通信媒体（有线或无线媒体）。经编码视频数据可根据通信标准（例如无线通信协议）进行调制，且被发射到目的地装置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。通信媒体可形成基于数据包的网络的部分，例如局域网、广域网或例如因特网的全球网。通信媒体可包含路由器、交换机、基站或可用于促进从源装置12到目的地装置14的通信的任何其它设备。

[0051] 在一些实例中，经编码数据可从输出接口22输出到存储装置，例如，存储装置31。类似地，可通过输入接口28从存储装置31存取经编码数据。存储装置31可包含各种分布式

或局部存取数据存储媒体中的任一者,例如,硬盘驱动、蓝光盘、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适合数字存储媒体。在又一实例中,存储装置31可对应于文件服务器或可存储由源装置12产生的经编码视频的另一中间存储装置。目的地装置14可经由流式处理或下载来从存储装置存取所存储视频数据。文件服务器可为能够存储经编码视频数据且将所述经编码视频数据发射到目的地装置14的任何类型的服务器。实例性文件服务器包含web服务器(例如,用于网站)、FTP服务器、网络附加存储(NAS)装置或本地磁盘驱动器。目的地装置14可通过任何标准数据连接(包含因特网连接)来存取经编码视频数据。此可包含无线信道(例如,Wi-Fi连接)、有线连接(例如,DSL、电缆调制解调器等等),或适于存取存储于文件服务器上的经编码视频数据的两者的组合。来自存储装置的经编码视频数据的发射可为流式发射、下载发射或其组合。

[0052] 本发明的技术未必限于无线应用或设置。所述技术可应用于支持各种多媒体应用中的任何者的视频编码,例如无线电视广播、有线电视发射、卫星电视发射、因特网流式视频发射,例如经由HTTP的动态自适应流式处理(DASH)、经编码到数据存储媒体上的数字视频、对存储于数据存储媒体上的数字视频的解码或其它应用。在一些实例中,系统10可经配置以支持单向或双向视频发射以支持例如视频流化、视频播放、视频广播及/或视频电话的应用。

[0053] 图1的所说明系统10仅为一个实例,且本发明中所描述的技术可由任何数字视频编码及/或解码装置执行。尽管本发明的技术通常由视频编码装置执行,但技术也可由编码器处理器/解码器处理器(通常称作“CODEC”)执行。此外,本发明的技术也可由视频预处理器执行。源装置12及目的地装置14仅为这些编码装置的实例,其中源装置12产生经译码视频数据以供发射到目的地装置14。在一些实例中,装置12、14可以大体上对称方式操作使得装置12、14中的每一者包含视频编码及解码组件。因此,系统10可支持装置12、14之间的单向或双向视频发射,例如,用于视频流式处理、视频播放、视频广播或视频电话。

[0054] 源装置12的视频源18可包含视频捕获装置,例如视频摄像机、含有先前所捕获视频的视频存档,及/或用以从视频内容提供者接收视频的视频馈送接口。作为又一替代方案,视频源18可产生基于计算机图形数据作为源视频,或直播视频、经存档视频及计算机产生视频的组合。然而,如上文所提及,本发明中所描述的技术通常可应用于视频译码,且可应用于无线及/或有线应用。在每一状况下,所捕获、预捕获或计算机产生视频可由编码器处理器20进行编码。输出接口22可接着将经编码视频信息输出到链路16上或到存储装置31。

[0055] 目的地装置14的输入接口28从链路16及/或存储装置31接收信息。所接收信息可包含由编码器处理器20产生的语法信息,所述语法信息也由解码器处理器30使用,包含描述图像的特性及/或处理的语法元素。显示装置32向用户显示经解码视频数据,且可包括各种显示装置中的任一者,例如,阴极射线管(CRT)、液晶显示器(LCD)、等离子显示器、有机发光二极管(OLED)显示器或另一类型的显示装置。

[0056] 尽管图1中未展示,但在一些方面中,编码器处理器20及解码器处理器30可各自与音频编码器及解码器整合,且可包含适当MUX-DEMUX单元或其它硬件及软件以在共用数据流或单独数据流中处置音频及视频两者的编码。如果适用,MUX-DEMUX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)的其它协议。

[0057] 编码器处理器20及解码器处理器30各自可实施为各种适合电路中的任一者,例如一或多个微处理器、数字信号处理器(DSP)、专用集成电路(ASIC)、场可编程门阵列(FPGA)、离散逻辑、集成电路(IC)、软件、硬件、固件或其任一组合。包含编码器处理器20及/或解码器处理器30的装置可包括集成电路、微处理器及/或无线通信装置(例如,蜂窝式电话)。当技术部分地以软件实施时,装置可将用于软件的指令存储在适合计算机可读媒体中且可在硬件中使用一或多个处理器来执行指令以执行本发明的技术。

[0058] 根据本发明中所描述的技术,编码器处理器20从视频源18接收视频数据。视频数据包含关于多个图像的视频数据,所述多个图像一起形成视频,包含关于当前图像的视频数据。关于当前图像的视频数据包含关于当前图像的像素的像素值(例如,颜色值)及坐标。为编码当前图像,编码器处理器20利用场景结构图。在一些实例中,编码器处理器20基于当前图像的视频数据而产生场景结构图。在一些实例中,编码器处理器20重新使用先前产生的场景结构图(例如,针对先前经编码图像而产生的场景结构图)。

[0059] 场景结构图包含关于当前图像或先前图像(例如,先前经编码图像)的稀少数目个点的三维(3D)坐标。举例来说,场景结构图中界定的点的数目小于当前图像中的像素的总数目。在本发明中所描述的技术中,发射场景结构图中的点的3D坐标所需要的字节量可基本上小于发射当前图像所需要的字节量。在一些状况中,与当前图像相比,发射场景结构图需要小于1%(例如,0.26%)的字节。此外,在一些实例中,并非输出场景结构图的所有点,且仅输出场景结构图的更新,从而致使与图像相比发射关于场景结构图的信息需要更少字节。

[0060] 应理解,在本发明中所描述的技术中,当前图像囊括2D空间。然而,当前图像内的对象似乎在前景、背景或中间的某处中。换句话说,虽然当前图像中的像素由2D坐标界定,但当前图像中似乎存在对象的相对深度。编码器处理器20利用当前图像中的对象的此相对深度来形成场景结构图的3D坐标。举例来说,编码器处理器20可使用同时定位与地图构建(SLAM)技术来形成当前图像的稀疏填充点的3D坐标以形成场景结构图。在一些实例中,编码器处理器20未必需要形成当前图像的场景结构图,且可将先前构造场景结构图重新用于编码当前图像。在这些实例中,编码器处理器20可已使用类似技术(例如,SLAM技术)由先前经编码图像来构造场景结构图。

[0061] 下文更详细地描述SLAM技术。然而,SLAM技术仅经描述为构造场景结构图的一个方式。构造场景结构图的其它方式是可能的,且技术并不限于使用SLAM技术来构造场景结构图。

[0062] 场景结构图可被视为图像的代理几何结构(例如,当前图像或先前经编码图像)。举例来说,由场景结构图界定的点可被视为多边形(例如,三角形)的顶点。如果点互连(例如,如果编码器处理器20或解码器处理器30互连所述点),那么结果为图像的三维结构,因此为术语“场景结构图”。

[0063] 场景结构图不应与多视图视频译码中所使用的深度图混淆。在多视图视频译码中,在基本上类似时间呈现多个图像,且观看者的右眼观看图像中的一者,而观看者的左眼观看图像中的另一者。每一眼观看不同图像的结果导致观看者体验包括三维视体(viewing volume)的图像。

[0064] 在多视图视频译码中,当前图像的深度图用于各种视频译码技术。深度图自身经

看作像图像,其中深度图中的像素的像素值指示纹理图(即,包含实际内容的图像)中的对应像素的相对深度。

[0065] 如本发明中所描述,场景结构图并非图像,与深度图不同。场景结构图包含关于图像中的稀少数目个点的坐标。另一方面,深度图提供远比图像中的稀少数目个点多的相对深度值。在一些状况下,图像中的点的数目可如此小使得当输出场景结构图时不需要压缩,与使用习用视频译码技术压缩的深度图不同。此外,如更详细所描述,编码器处理器20利用场景结构图来形成作用于编码当前图像的预测图像的合成图像。多视图视频译码中的深度图并不用于这些用途。换句话说,深度图的形式可使深度图不可用于本发明中所描述的技术,与场景结构图不同,且与场景结构图相比,深度图的数据量可需要额外宽度。

[0066] 编码器处理器20利用场景结构图来形成合成图像。举例来说,编码器处理器20检索先前图像或多个先前图像(例如,先前经编码图片)且形成复合图像(例如,通过混合先前经编码图像,或在不使用多个图像的条件下,复合图像可为单个图像)。编码器处理器20可接着互连场景结构图的点以形成网格,且使复合图像重叠在网格上。举例来说,复合图像可被视为纹理图,且编码器处理器20使用图形处理技术将纹理图叠加在场景结构图的网格上。结果为基于图像的模型(IBM),所述基于图像的模型为场景结构图的网格上的复合图像的三维图形构造。编码器处理器20渲染三维图形构造以形成二维图像。此二维图像为合成图像。

[0067] 编码器处理器20利用合成图像作为预测图像,且确定当前图像与合成图像之间的差异(被称作残差图像)。编码器处理器20可对残差图像执行额外(任选)编码,并输出指示残差图像的信息(例如,可依据其确定残差图像的信息)。在一些实例中,编码器处理器20还可输出场景结构图或相对于先前场景结构图的场景结构图的渐进式改变。场景结构图的信息量或渐进式改变可相对小,且可不需要此信息的进一步压缩,但可能进一步压缩此信息。针对每一实例可不需要输出场景结构图或渐进式改变。举例来说,编码器处理器20可重新使用先前场景结构图来构造合成图像,在所述状况下,编码器处理器20可不输出场景结构图或场景结构图的渐进式改变。

[0068] 在一些实例中,编码器处理器20可从视频源18接收摄像机位置及/或摄像机定位信息(被称作摄像机姿势信息)。在一些实例中,编码器处理器20可利用SLAM技术来确定摄像机位置。然而,可利用用以确定摄像机位置及/或摄像机姿势的其它技术,例如外部追踪系统。编码器处理器20可利用摄像机姿势信息来编码当前图像。举例来说,关于图像的摄像机姿势信息可被认为拍摄图像的定向及位置。

[0069] 编码器处理器20可利用关于用于构造复合图像的一或多个先前经编码图像的摄像机姿势信息来确定将复合图像叠加在场景结构图上的方式。举例来说,基于由一或多个先前经编码图图像的摄像机姿势信息指示的视角,编码器处理器20可确定应将特定像素置于场景结构图上以形成基于图像的模型(IBM)的位置。举例来说,如果摄像机姿势信息指示关于一或多个先前经编码图像的摄像机的位置及定向为笔直的,那么编码器处理器20将把像素映射到场景结构图的位置将与在摄像机姿势信息指示摄像机的位置及定位成角度的情况下不同。这是因为包含像素的对象将基于视角出现在不同相对位置处。

[0070] 编码器处理器20还可将摄像机姿势信息用于当前图像。举例来说,如上文所描述,合成图像形成预测图像。为了减少残差数据(即,为了使合成图像与当前图像之间的差最小

化),编码器处理器20可基于当前图像的摄像机姿势信息而渲染基于图像的模型(IBM)使得合成图像的摄像机姿势与当前图像的摄像机姿势相同。

[0071] 在一些实例中,编码器处理器20可输出关于每一图像的摄像机姿势信息使得解码器处理器30可以与编码器处理器20类似方式利用摄像机姿势信息,但用于解码当前图像。在一些实例中,编码器处理器20在输出之前任选地压缩摄像机姿势信息以减少需要输出的信息量。

[0072] 摄像机姿势信息可用于其中摄像机移动的实例中(即,关于视频序列中的一或多个图像的摄像机姿势信息不同),例如在视频电话或视频会议中。然而,在一些状况下,摄像机姿势可恒定。在这些实例中,编码器处理器20未必需要输出关于每一图像的摄像机姿势信息,且可输出摄像机姿势信息一次。在一些实例中,编码器处理器20及解码器处理器30可各自经预先配置有特定摄像机姿势信息,在所述状况下,编码器处理器20可不输出摄像机姿势信息。应理解,甚至在其中摄像机姿势信息未改变的实例中,编码器处理器20可仍发射关于图像中的一或多者的摄像机姿势信息。

[0073] 为进一步压缩视频数据(例如,残差数据、场景结构图信息及/或摄像机姿势信息),编码器处理器20可利用一或多个实例技术。作为一个实例,残差数据可被视为图像。编码器处理器20可使用视频压缩技术(例如,基于块的编码)来压缩残差数据。作为另一实例,编码器处理器20可将熵编码技术作为基于块编码的部分且用于场景结构图信息及/或摄像机姿势信息,例如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。熵编码技术仅经描述为一个实例。其它编码技术是可能的。

[0074] 在一些实例中,编码器处理器20可分割当前图像的视频数据。举例来说,编码器处理器20可确定当前图像的哪些部分相对于先前图像(例如,先前经编码图像)为静态的,且可确定当前图像的哪些部分正改变(例如,非静态)。编码器处理器20可将当前图像的非静态部分与静态部分分开(例如,非静态部分为不同于包含静态部分的层的层)。编码器处理器20可将本发明中所描述的技术用于静态部分,而将常规视频译码技术用于非静态部分。然而,编码器处理器20可能将本发明中所描述的技术用于静态部分及非静态部分两者。

[0075] 可存在借以确定当前图像的哪些部分为静态及哪些部分为非静态的各种方式。作为一个实例,视频源18可首先捕获场景,且编码器处理器20可创建背景的基于图像的模型(IBM)。视频源18可接着捕获实际场景,例如人移动穿过先前记录的房间。编码器处理器20可确定两个视频序列之间(例如,背景视频序列与具有移动对象(在此实例中为人)的序列之间)的差异。编码器处理器20可接着确定非静态部分及静态部分。在此实例中,类似于上文,编码器处理器20可在用于从相应IBM渲染合成图像的非静态及静态图像的记录过程(视频捕获过程)中从视频源18接收摄像机姿势信息。

[0076] 解码器处理器30执行编码器处理器20的相反过程。然而,解码器处理器30不需要构造场景结构图。确切地说,解码器处理器30接收当前图像的场景结构图或接收关于当前图像的场景结构图相对于先前场景结构图的渐进式改变。在一些实例中,解码器处理器30未接收场景结构图的任何信息,且重新使用先前场景结构图(例如,针对先前经解码图像所产生的场景结构图)。

[0077] 作为实例,解码器处理器30可解码残差数据。举例来说,残差数据为基于框经编



码,解码器处理器30可基于框地解码残差数据。残差数据也可以其它方式编码。在一些实例中,残差数据可未经编码,在所述状况下,解码器处理器30未解码残差数据。在任何情况下,解码器处理器30接收残差数据(例如,经编码残差数据或未经编码残差数据)。

[0078] 另外,解码器处理器30可接收关于当前图像的摄像机姿势信息且在此信息经编码的情况下解码所述摄像机姿势信息。如上文所描述,解码器处理器30还可接收场景结构图或当关于当前图像的场景结构图相对于先前场景结构图的渐进式改变。在一些实例中,场景结构图的信息或渐进式改变可未经编码或以其它方式经压缩,且解码器处理器30可不需要解码关于场景结构图或渐进式改变的信息。

[0079] 类似于编码器处理器20,解码器处理器30利用场景结构图及复合图像来产生合成图像,其中合成图像为预测图像。在其中解码器处理器30接收摄像机姿势信息的实例中,解码器处理器30可使用摄像机姿势信息来产生合成图像。举例来说,如同编码器处理器20,解码器处理器30可混合一或多个先前图像(例如,先前经解码图像)以形成复合图像,或单个先前经解码图像可为复合图像。解码器处理器30可将复合图像映射到场景结构图(且任选地使用摄像机姿势信息)以形成IBM,且可渲染IBM以产生合成图像。解码器处理器30可将残差数据添加到合成图像以重构当前图像。

[0080] 编码器处理器20及解码器处理器30可广义地被称作译码器处理器。举例来说,编码器处理器20及解码器处理器30可经配置以产生形成关于当前图像的预测图像的合成图像。为产生合成图像,编码器处理器20及解码器处理器30可经配置以执行类似功能,且为简便起见,这些功能经描述为由译码器处理器执行(其实例包含编码器处理器20及解码器处理器30)。

[0081] 举例来说,译码器处理器可经配置以基于先前复合图像(经存储在视频存储器中)及场景结构图而产生。场景结构图包含当前图像的场景结构图或先前经译码图像的场景结构图,且场景结构图包含关于当前图像或先前经译码图像内的三维点的坐标值。

[0082] 译码器处理器可经配置以基于所述当前图像的残差图像而译码所述当前图像,其中所述残差图像指示当前图像与合成图像之间的差异。举例来说,在译码器处理器为解码器处理器30的情况下,解码器处理器30可通过基于所接收残差图像及合成图像而重构当前图像来解码当前图像。在译码器处理器为编码器处理器20的情况下,编码器处理器20可通过确定残差图像及输出残差图像的信息来编码当前图像。

[0083] 一些其它技术已提出使用SLAM技术来用于视频译码用途。举例来说,在这些其它技术中,译码器处理器(不同于编码器处理器20及解码器处理器30)确定非常详细的场景结构图,而非基于稀少的点。这些其它技术的详细场景结构图可为将先前图像叠加在详细结构图产生与当前图像相同的图像的此细节。

[0084] 在这些其它技术中,输出并接收关于详细的场景结构图的点坐标,但不发送或接收任何残差数据。这是因为将先前图像叠加在详细场景结构图上产生与当前图像基本上相同的图像。因此,这些其它技术不依赖于任何残差数据。在这些其它技术中,合成图像并非预测图像,但替代地,被认为是当前图像的复本。

[0085] 这些其它技术可存在某些缺点。举例来说,输出及接收关于详细的场景结构图的坐标可为带宽数量大的。此外,所得图像(例如,在将先前图像叠加在详细的场景结构图上的情况下)可并非与当前图像良好匹配的图像。因此,图像质量可其它视频译码技术更



糟。

[0086] 在本发明中所描述的技术中,译码器处理器基于当前图像或先前经译码图像中的较少点产生合成图像,且此合成图像形成为预测图像。因此,可能需要输出及接收较少点,且合成图像与当前图像之间的差异可足够小使得需要输出及接收有限数据(例如,残差图像具有相对小的数据量)。以此方式,可存在需要发射及接收的数据量的减少,提升带宽效率。此外,由于合成图像为预测图像且发射及接收残差图像,因此经重构的当前图像可具有比其中不发送任何残差图像的实例更好的质量。

[0087] 图2为说明经配置以实施或以其它方式利用本发明中所描述的一或多个实例视频编码技术的编码器处理器20的实例的框图。如所说明,编码器处理器20包含同时定位与地图构建(SLAM)处理器34,图形处理单元(GPU)36、视频存储器38及编码器48。在一些实例中,SLAM处理器34、GPU 36及编码器48可一起形成在单芯片上以形成单片系统(SoC)。在一些实例中,SLAM处理器34、GPU 36及编码器48可形成在单独芯片上。

[0088] 编码器处理器20各自可实施为各种适合电路中的任一者,例如一或多个微处理器、数字信号处理器(DSP)、专用集成电路(ASIC)、场可编程门阵列(FPGA)、离散逻辑、软件、硬件、固件或其任一组合。举例来说,本发明中所描述的技术可从设备或装置的角度描述。作为一个实例,设备或装置可包含编码器处理器20(例如,源装置12作为无线通信装置的部分),且编码器处理器20可包含经配置以实施本发明中所描述的技术的一或多个处理器。作为另一实例,设备或装置可包含微处理器或包含编码器处理器20的集成电路(IC),且微处理器或IC可为源装置12或另一类型的装置的部分。

[0089] 编码器处理器20的组件可为源装置12的现有实例的部分。举例来说,作为标准硬件的部分,源装置(例如,源装置12)的大部分实例包含GPU(例如,GPU 36)、编码器(例如,编码器48)及视频存储器(例如,视频存储器38)。即使装置不包含SLAM处理器34,装置的中央处理单元(CPU)可经配置以充当SLAM处理器34。举例来说,通过对装置的CPU或主机处理器执行硬件或固件,CPU或主机处理器可经配置以实现SLAM处理器34的功能。

[0090] 另外,GPU 36经说明为经配置以实施本发明中所描述的实例技术的一个实例组件。GPU 36的最快平行处理能力使得GPU 36成为用于实施本发明中所描述的技术的适合选项。然而,技术并不限于此。可能利用除GPU 36外的组件来实施包含GPU的实例技术。此外,一些GPU包含处理能力以充当通用GPU(GPGPU)。在一些实例中,GPU 36可为GPGPU且可经配置以执行SLAM处理器34的功能(即,SLAM处理器34及GPU 36可为相同GPGPU的部分)。

[0091] 视频存储器38可为形成编码器处理器20的芯片的部分。在一些实例中,视频存储器38可在编码器处理器20外部,且可为源装置12的系统存储器的部分。在一些实例中,视频存储器38可为在编码器处理器20内部的存储器及系统存储器的组合。

[0092] 视频存储器38可存储用于编码当前图像的数据。举例来说,如所说明,视频存储器存储场景结构图40(在由SLAM处理器34针对当前图像产生的条件下),存储先前场景结构图42(例如,SLAM处理器34针对先前经编码图像产生的场景结构图)、一或多个图像44(例如,先前经编码图像),及GPU 36产生的合成图像46。视频存储器38可由各种存储器装置中的任一者形成,例如,动态随机存取存储器(DRAM),包含同步DRAM(SDRAM)、磁阻式RAM(MRAM)、电阻式RAM(RRAM)或其它类型的存储器装置。

[0093] 如图2中所说明,SLAM处理器34接收当前图像并产生场景结构图40。SLAM处理器34

经配置以从当前图像提取稀少的3D点云重构,且在一些状况下可确保经重构点在屏幕空间中的均衡分布。

[0094] SLAM技术通常在机器人及自主车辆中通常用于创建几何周围地图,所述几何周围地图依赖于用以创建几何周围地图的传感器的反馈。SLAM处理器34可经配置以利用视觉SLAM (VSLAM) 感测技术来产生几何周围地图。举例来说,SLAM处理器34可从视频源18接收所捕获图像并执行SLAM技术来产生几何周围地图。

[0095] 在产生几何地图中,SLAM处理器34可追踪视频源18的3D摄像机姿势。举例来说,在例如其中用户拿着移动装置走动的视频电话或其中控制摄像机位置及定位的视频会议的实例中,摄像机往往移动,但场景大部分为静态。由于摄像机的位置改变,捕获图像的角度也改变。通过追踪摄像机的3D位置及定向(姿势),SLAM处理器34可较佳地能够构造几何周围地图。在一些实例中,SLAM处理器34可利用SLAM技术来确定摄像机位置。

[0096] 在一些状况下,可使用SLAM技术来增强现实(AR)。举例来说,图形视频或图像经叠加在背景区域上。SLAM处理器34使用SLAM技术映射场景结构,且AR视频展示经修改视频流(例如,其中AR视频叠加)。用于AR的SLAM的经典实例为来自电影“星球大战:第四集—新希望”的其中机器人R2D2投射莱亚(Leia)公主的影片记录的场景。

[0097] 为产生几何周围地图,SLAM处理器34构造场景结构图。如上文所描述,场景结构图包含关于当前图像中的稀少数目个点的(x,y,z)坐标。关于当前图像中的点的x、y、z坐标可基于摄像机姿势,且针对给定摄像机姿势指示点中的每一者的相对位置。举例来说,z坐标的值可指示特定点的相对深度,所述相对深度指示点观看者将感知所述点的距离。此外,当前图像囊括2D区域,但图像的内容包含点的相对深度(其由z坐标指示)。

[0098] 图3A及3B为说明场景结构图的实例的图形图。举例来说,图3A及3B说明不同图像。图3A及3B中所说明的点形成这些图像的场景结构图。举例来说,图3A及3B说明针对y坐标的特定值在x方向及z方向延伸的方格。存在于此方格上的点由其x、y及z坐标界定。

[0099] SLAM处理器34将所产生场景结构图存储在视频存储器38中(例如,将场景结构图40存储在视频存储器38中)。然而,SLAM处理器34未必需要产生每一图像的场景结构图。如所说明,视频存储器38还可存储先前场景结构图42。SLAM处理器34可已产生先前图像(例如,在当前图像之前经编码的图像)的场景结构图。在产生先前图像时,SLAM处理器34可已将所得场景结构图存储在视频存储器38中。接着,针对当前图像,先前产生场景结构图变成先前场景结构图42。

[0100] 在根据本发明中所描述的技术,GPU 36可检索场景结构图40(如果可用)或先前场景结构图42(在场景结构图40不可用的情况下)中的一者。GPU 36可互连场景结构图的点以形成代理几何结构。举例来说,场景结构图的点可被视为多边形(例如,三角形)的顶点。GPU 36可执行顶点着色器或可利用硬连线输入汇编器来互连场景结构图的点。作为一个实例,GPU 36可利用渐进式狄洛尼(Delauney)三角测量技术来互连所述点。

[0101] 点的互连的结果为表示用于产生场景结构图的图像的结构代理几何结构。概念上,代理几何结构中的点的互连形成三角网格,所述三角网格形成用于产生场景结构图的图像的相对三维结构的粗略估计。举例来说,点的互连提供深度近似值且比实际深度图稀少得多。然而,点的互连以形成三角网格可提供用以叠加复合图像以产生预测图像的足够细节,如下文更详细描述。

[0102] 图4为说明场景结构图的点的互连的实例的概念图。举例来说,如所说明,图4包含形成代理几何结构的多个三角形,且三角形的顶点由场景结构图界定。顶点中的每一者可由x、y及z坐标界定。举例来说,所说明代理几何结构包含一些3D点及3D三角形。在一些实例中,GPU 36可经配置以产生代理几何结构(例如,经由输入汇编器或顶点着色器)。GPU 36可渐进地建构代理几何结构且可暂时地存储代理几何结构(例如,仅用于产生合成图像,如下文所描述)。

[0103] 除了存储先前场景结构图42及/或场景结构图40外,视频存储器38还可存储一或多个图像44(例如,先前经编码图像)及其对应摄像机姿势信息,例如在其中摄像机移动的实例中。一或多个图像44为编码器处理器20先前经编码的图像。视频存储器38存储编码器处理器20先前存储的原始图像,或编码器48包含用以重构经编码图像的反馈路径,且视频存储器38将这些经重构图像存储为一或多个图像44。在一些实例中,可比当前图像早地显示一或多个图像44,但比当前图像早显示的一或多个图像44并非要求且在一些实例中可与当前图像几乎同时地(例如,用于立体观看)或在当前图像之后显示一或多个图像44。

[0104] 在一些实例中,一或多个图像44(例如,先前经编码图像)可为关键帧,也被称作I帧。举例来说,在一些实例中,为避免传播编码错误,编码器处理器20可间或地编码图像而无需参考另一预测图像。举例来说,编码器48可使用图像内的像素的样本值且非依据来自任何其它图像的值来编码图像。仅参考图像内的样本编码的图像被称作I帧或关键帧。从其它图像预测的图像被称作P帧。

[0105] 一或多个图像44未必需要在每一实例中为关键帧。然而,出于实例及说明的目的,关于一或多个图像44所描述的技术为关键帧。

[0106] 另外,本发明中所描述的技术不需要利用多个关键帧(例如,多个图像44),且可利用单个关键帧(例如,单个图像44)。在其中利用多个图像44的实例中,GPU 36可对两个或多于两个图像44执行混合操作以形成复合图像。举例来说,GPU 36包含在其图形管线中包含混合单元,所述混合单元经具体配置以混合像素值。多个图像44可为到混合单元的输入,且混合单元的输出可为经混合图像。在任何额外图形处理之后,如果需要,那么GPU 36的输出为复合图像。尽管为说明,但GPU 36将复合图像存储在视频存储器38中。在其中仅利用一个图像44的实例中,GPU 36可未混合,且单个图像44为复合图像。

[0107] 根据本发明中所描述的技术,复合图像为纹理图且代理几何结构(例如,通过互连场景结构图的点形成的网格)为纹理图被映射到的对象。举例来说,在纹理映射中,纹理图是以二维坐标(u,v)界定且纹理图被映射到的对象是以三维界定。GPU 36的功能中的一者可为执行用于图形混合的纹理映射,且技术利用GPU 36的能力来以快速且高效的方式执行纹理映射以产生作用于视频编码及解码的预测图像的合成图像。

[0108] 举例来说,GPU 36可执行纹理引擎或可预配置有硬件以执行纹理功能。在任一实例中,GPU 36将来自纹理图的二维点映射到图形对象,本质上,将纹理图叠加在对象上。纹理化的结构为高度详细的图形对象,所述高度详细的图形对象看上去比纯图形构造更真实。作为实例,对象为球体且纹理图为世界的2D图像。通过将世界的2D图像包裹到三维球体作为纹理映射的部分,与构造地球仪的图形相比,GPU 36可渲染更详细且视觉愉悦的地球仪。

[0109] 在图2中所说明的实例中,GPU 36将来自复合图像的像素映射到代理几何结构上

的三维位置。如上文所描述,在一些实例中,GPU 36可利用一或多个图像44(用于构造复合图像)的摄像机姿势信息来将来自复合图像的像素映射到代理几何结构上的三维位置(例如,摄像机姿势信息影响像素将映射到的三维位置)。尽管复合图像的内容指示相对深度,但复合图像的全部像素是由二维(u,v)坐标界定。由于代理几何结构为三维的,但纹理映射的结果为三维对象,被称作基于图像模型(IBM)。

[0110] GPU 36对IBM执行进一步图形处理。举例来说,GPU 36可渲染IBM。作为渲染过程的部分,GPU 36将IBM的三维坐标转换成形成合成图像46二维屏幕坐标。在一些实例中,GPU 36可基于当前摄像机位置而渲染IBM。举例来说,由于合成图像46形成预测图像,GPU 36可基于当前摄像机姿势信息而渲染IBM使得合成图像46与当前图像之间的比较有效且残差数据最小化。

[0111] GPU 36将渲染为合成图像46的结果存储在视频存储器38中。合成图像46可为具有针对每一坐标的像素值的二维图像。在本发明中所描述的技术中,合成图像46形成预测图像。

[0112] 举例来说,编码器48接收合成图像46及当前图像,且确定当前图像与合成图像之间的残差图像(例如,残差数据)。编码器48可将位流中的所得残差图像输出作为经编码视频数据。在一些实例中,在输出残差图像之前,编码器48可执行对残差图像的一些额外编码,例如熵编码或基于块的编码。以此方式,无论编码器48执行对残差图像的额外编码与否,编码器48可输出指示残差图像的信息(例如,可依据其确定残差图像的信息)。

[0113] 在一些实例中,编码器48可经配置以实施本发明中所描述的技术且还可经配置以实施常规视频编码过程。举例来说,编码器48可经配置以仅基于图像内的样本而编码图像(例如,将图像编码成I帧)。编码器48可利用常规视频编码执行此视频编码技术。

[0114] 作为另一实例,在一些实例中,在使用其它视频编码技术编码当前图像中的非静态前景对象且使用本发明中所描述的技术编码静态背景对象的情况下,可实现编码增益。举例来说,编码器48可包含模式选择单元,所述模式选择单元在使用本发明中所描述的技术编码图像的部分与使用其它技术编码图像的其它部分之间选择。使用本发明中所描述的技术编码的部分可形成一个层,且使用其它技术编码的部分可形成另一层。应理解,可能使用本发明中所描述的技术编码整个当前图像。

[0115] 在其中前景非静态部分形成一个层且背景静态部分形成另一层的实例中,SLAM处理器34对前景执行SLAM技术。在此些实例中,编码器48可单独地编码前景层及背景层。

[0116] 在一些实例(其在图2中以虚线说明)中,SLAM处理器34将摄像机姿势信息输出到编码器48。编码器48编码摄像机姿势信息以供输出到位流中(例如,作为经编码视频数据的部分)。此外,在一些实例中,SLAM处理器34可将场景结构图40的信息输出到编码器48,且编码器48编码场景结构图40的信息。在一些实例中,当编码器处理器20利用场景结构图40时,编码器48可确定场景结构图40与先前场景结构图42之间的差异,且可输出所述差异(例如,其表示场景结构图中的渐进式改变),而非场景结构图40的实际值。如上文所描述,编码器处理器20可不在每一实例中利用场景结构图40。在这些实例中,编码器48可不输出场景结构图40或场景结构图40与先前场景结构图42之间的差异的信息。在这些实例中的一些中,编码器48可输出语法元素,所述语法元素指示解码器处理器30应利用本地存储的先前场景结构图。

[0117] 图5为说明根据本发明的一或多个实例视频编码技术的数据流程图。如图5的实例中所说明,编码器处理器20接收当前图像50。SLAM处理器34实施SLAM技术以产生场景结构图,被称作SLAM 52。GPU 36利用复合图像及场景结构图来形成基于图像的模型54。GPU 36渲染基于图像的模型54以产生合成图像46,其为预测图像56。

[0118] 减法单元53将预测图像56从当前图像50减去。减法的结果为残差图像58。如上文所描述,编码器48可执行额外编码(在图5中经说明为编码60)。在一些实例中,为将预测图像56从当前图像50减去,减法单元53使预测图像56的值乘以负一并当前图像50加到结果。为简洁起见,本发明将减法单元53描述为将预测图像56从当前图像50减去,且可通过直接减法,通过确定图像的负数(例如,使值与负一相乘)并将其加到另一值,或通过某一其它技术来执行此运算。

[0119] 编码器48输出所得经编码图像作为经编码视频数据位流62的部分。另外,位流62可包含关于当前图像50的摄像机姿势信息。位流62还可包含场景结构图的信息或指示场景结构图与先前场景结构图之间的差(例如,指示场景结构图相对于先前场景结构图的渐进式改变)的信息。然而,位流62未必需要在每一实例中包含场景结构图或场景结构图与先前场景结构图之间的差异。

[0120] 图6为说明根据本发明中所描述的技术的视频编码的实例的概念图。在图6中,图像64为当前图像且图像66为合成图像。图像68说明残差图像(例如,图像64减去图像66)。

[0121] 图7为说明经配置以实施或以其它方式利用本发明中所描述的一或多个实例视频解码技术的解码器处理器30的实例的框图。如所说明,解码器处理器30包含GPU 70、解码器80及视频存储器82。在一些实例中,GPU 70及解码器80可一起形成在单芯片上以形成单频系统(SoC)。在一些实例中,GPU 70及解码器80可形成在单独芯片上。

[0122] 解码器处理器30可实施为各种适合电路中的任一者,例如一或多个微处理器、数字信号处理器(DSP)、专用集成电路(ASIC)、场可编程门阵列(FPGA)、离散逻辑、软件、硬件、固件或其任一组合。举例来说,本发明中所描述的技术可从设备或装置的角度描述。作为一个实例,设备或装置可包含解码器处理器30(例如,目的地装置14作为无线通信装置的部分),且解码器处理器30可包含经配置以实施本发明中所描述的技术的一或多个处理器。作为另一实例,设备或装置可包含微处理器或包含解码器处理器30的集成电路(IC),且微处理器或IC可为目的地装置14或另一类型装置的部分。

[0123] 解码器处理器30的组件可为目的地装置14的现有实例的部分。举例来说,作为标准硬件的部分,目的地装置14的大部分实例包含GPU(例如,GPU 70)、解码器(例如,解码器80)及视频存储器(视频存储器82)。GPU 70经说明为经配置以实施本发明中所描述的实例技术的一个实例组件。GPU 70的最快平行处理能力可使得GPU 70成为用于实施本发明中所描述的技术的适合选项。然而,技术并不限于此。可能利用除GPU 70外的组件来实施包含目的地装置14的CPU的实例技术。另外,一些GPU包含充当通用GPU(GPGPU)的处理能力。在一些实例中,GPU 70可为GPGPU且可经配置以执行解码器80的功能(即,解码器80及GPU 70可为相同GPGPU的部分)。

[0124] 视频存储器82可为形成解码器处理器30的芯片的部分。在一些实例中,视频存储器82可在解码器处理器30外部,且可为目的地装置14的系统存储器的部分。在一些实例中,视频存储器82可为在解码器处理器30内部的存储器及系统存储器的组合。

[0125] 视频存储器82可存储用于解码当前图像的数据。举例来说,如所说明,视频存储器82在经接收作为经编码视频数据位流的部分的情况下存储场景结构图72,存储先前场景结构图74、一或多个图像76及GPU 70产生的合成图像78。视频存储器82可由各种存储器装置中的任一者形成,例如,动态随机存取存储器(DRAM),包含同步DRAM(SDRAM)、磁阻式RAM(MRAM)、电阻式RAM(RRAM)或其它类型的存储器装置。

[0126] 解码器80接收包含关于当前图像的残差图像的信息的视频数据的位流(例如,残差图像是由将合成图像从当前图像减法而产生)。解码器80还接收摄像机姿势信息。在一些实例中,解码器80接收当前图像的场景结构图的信息,解码器80将其作为场景结构图72存储在视频存储器82中。在一些实例中,解码器80接收指示当前图像的场景结构图与先前场景结构图(例如,先前场景结构图74)之间的差异的信息。解码器80将差异值与先前场景结构图74汇总在一起,并存储所得值作为场景结构图72。

[0127] GPU 70接收场景结构图72或先前场景结构图74中的一者并互连点以使用(例如)顶点着色器或输入汇编器形成代理几何结构。另外,GPU 70混合图像76中的两者或多于两者以形成复合图像或使用图像76中的单个者作为复合图像。

[0128] 解码器处理器30的GPU 70及编码器处理器20的GPU 36可经配置以利用完全相同的图像集(例如,先前经解码图片)。举例来说,如同GPU 36,GPU 70可利用一或多个关键帧(例如,I帧),且一或多个图像76可全部为关键帧。以此方式,GPU 70形成的复合图像与GPU 36形成的复合图像基本上相同。

[0129] GPU 70可使用复合图像作为纹理图且使用代理几何结构作为纹理图被映射到的对象来执行纹理映射。GPU 70可接着渲染结果以创建合成图像78。GPU 70可实施GPU 36实施以产生合成图像46的基本上相同的过程以产生合成图像78。举例来说,由于GPU 70利用与GPU 36相同的场景结构图来形成代理几何结构,且利用与GPU 36相同的图像来形成复合图像,合成图像78及合成图像46可为基本上类似。另外,GPU 70可类似地将用于构造复合图像的一或多个图像76的摄像机姿势信息用于纹理映射以产生IBM,并利用当前图像的摄像机姿势信息来渲染合成图像。

[0130] 解码器80可将残差图像加到合成图像78。加法的结果为所重构的当前图像。

[0131] 此外,类似于编码器48,解码器80可经配置以实施本发明中所描述的技术且还可经配置以实施常规视频解码过程。举例来说,解码器80可经配置以仅基于图像内的样本而解码图像(例如,将图像解码成I帧)。解码器80可利用常规视频解码执行此视频解码技术。

[0132] 作为另一实例,在一些实例中,在使用其它视频解码技术解码当前图像中的非静态前景对象且使用本发明中所描述的技术解码静态背景对象的情况下,那么可实现解码增益。举例来说,解码器80可包含模式选择单元,所述模式选择单元在使用本发明中所描述的技术解码图像的部分与使用常规技术解码图像的其它部分之间选择。举例来说,解码器80可接收具有使用本发明中所描述的技术编码的当前图像的部分的一个层及具有使用其它技术编码的当前图像的部分的另一层。

[0133] 在实例中,在前景非静态部分形成一个层且背景静态部分形成另一层的情况下,解码器80可单独地解码前景层及背景层。解码器处理器30可单独地播放前景非静态层及背景静态层。

[0134] 解码器80还可经配置以解码提供解码信息的语法元素。举例来说,在其中GPU 70

将使用先前场景结构图74来产生代理几何结构的实例中,解码器80可解码指示GPU 70应使用先前场景结构图74的语法元素。

[0135] 图8为说明根据本发明的一或多个实例视频解码技术的数据流程图。举例来说,解码器80执行位流(例如,位流62)的解码,经说明为解码84。解码器80基于指示编码器处理器20输出的残差图像的信息而确定残差图像86。在一些实例中,解码器80输出场景结构图(例如,直接地或通过指示场景结构图与先前场景结构图之间的差异的信息汇总到先前场景结构图的值)。

[0136] GPU 70由所接收的场景结构图形成代理几何结构。在一些实例中,GPU 70可利用先前场景结构图。举例来说,解码84可不涉及解码场景结构图或场景结构图与先前场景结构图之间的差异。在此些实例中,GPU 70重新使用先前场景结构图。

[0137] 在本发明中所描述的技术中,GPU 70基于场景结构图及复合图像而产生代理几何结构,并执行纹理映射,其中复合图像为纹理图且代理几何结构为纹理图经映射到的对象。纹理映射的结果为基于图像的模型88。GPU 70渲染基于图像的模型88以产生合成图像,由预测图像90表示。

[0138] 汇总单元87将残差图像86加到预测图像90。加法的结果为当前图像92。

[0139] 以所述方式,本发明中所描述的实例技术提供用于视频编码及解码的实例技术。在一些实例中,可不需要常规运动补偿。举例来说,可能不需要关于运动矢量的信息、参考图片列表及常规运动补偿中所使用的其它此信息,从而允许所需要带宽量的减少。虽然不需要运动补偿信息,但可需要输出摄像机姿势及场景结构图(场景结构图的改变)信息。然而,针对摄像机矩阵及场景结构图的3D点每图像需要仅少数字节,从而产生微不足道的带宽要求。

[0140] 此外,本发明中所描述的技术可良好地适于实时(on-the-fly)视频编码及解码。举例来说,编码器处理器20及解码器处理器30能够用相应GPU实时产生相应预测图像(例如,合成图像)。此可被视为与在解码期间由GPU上的纹理映射及深度缓冲硬件进行的运动补偿类似。

[0141] 此外,技术还可用于立体显示。举例来说,合成图像可替代当前图像的摄像机姿势在理论上从任一视点(例如,可用于产生用于立体视图的图像)渲染。在此状况下,所得合成图像可能对于非立体视图中的预测用途并非理想的,但对于小改变,合成图像可为足够的。此些技术可被称作有限自由视点。

[0142] 图9为说明编码视频数据的实例方法的流程图。如图9的实例中所说明,编码器处理器20的GPU 36可基于复合图像及当前图像的场景结构图或先前经编码图像的场景结构图而产生合成图像(100)。复合图像由先前经编码的一或多个图像构造。举例来说,复合图像可为两个或多于两个先前经编码关键帧的混合或单个先前经编码关键帧。场景结构图包含关于视频数据的当前图像或先前经编码的视频数据的图像内的三维点的坐标值。举例来说,编码器处理器20的SLAM处理器34可利用同时定位与地图构建(SLAM)技术来产生场景结构图。

[0143] 在一些实例中,为产生合成图像,GPU 36可互连场景结构图的点以形成代理几何结构。GPU 36可将复合图像纹理映射到代理几何结构以形成基于图像的模型。在一些实例中,GPU 36可基于用于构造复合图像的一或多个图像的摄像机姿势信息将所述复合图像纹



理映射到代理几何结构以形成基于图像的模型。GPU 36可渲染基于图像的模型以产生合成图像。在一些实例中，GPU 36可基于当前图像的摄像机姿势信息而渲染基于图像的模型以产生合成图像。

[0144] 编码器处理器20的编码器48可基于合成图像及当前图像而确定残差图像(102)。残差图像指示当前图像与合成图像之间的差异。编码器48可输出残差图像(例如，输出指示残差图像的信息)以编码视频数据的当前图像(104)。在一些实例中，编码器48还可输出当前图像的场景结构图的信息或指示当前图像的场景结构图与先前经编码图像的场景结构图之间的差异的信息中的一者，且可输出摄像机位置与摄像机定向中的一者或两者。

[0145] 在一些实例中，编码器处理器20或源装置12的某一其它单元可确定当前图像的前景非静态部分及当前图像的背景静态部分。在这些实例中，为确定残差图像，编码器48可经配置以基于合成图像及当前图像的背景静态部分而确定残差图像。此外，在这些实例中，为输出残差数据，编码器48可经配置以输出不同于第二层的第一层中的残差数据，所述第二层包含关于当前图像的前景非静态部分的残差数据。

[0146] 图10为说明解码视频数据的实例方法的流程图。解码器处理器30的解码器80可确定视频数据的当前图像的残差图像(200)。残差图像指示当前图像与合成图像之间的差异。举例来说，解码器80可接收指示编码器处理器20输出的残差图像的信息。指示残差图像的此信息可为经编码信息或未经编码信息。在任一实例中，解码器80可基于指示编码器处理器20输出的残差图像的信息而确定残差图像(例如，重构残差图像)。

[0147] 解码器处理器30的GPU 70可基于复合图像及当前图像的场景结构图或先前经解码图像的场景结构图而产生合成图像(202)。复合图像由先前经解码的一或多个图像构造。场景结构图包括视频数据的当前图像的场景结构图或先前解码的视频数据的图像的场景结构图，且场景结构图包含关于当前图像或先前经解码的图像内的三维点的坐标值。解码器80可基于合成图像及残差图像而重构当前图像(204)。

[0148] 举例来说，如同GPU 36，GPU 70可互连场景结构图的点以形成代理几何结构。GPU 70可将复合图像纹理映射到代理几何结构以形成基于图像的模型，且可能基于先前经解码的一或多个图像的摄像机姿势信息。GPU 70可渲染基于图像的模型以产生合成图像，且可能基于当前图像的摄像机姿势信息。

[0149] 在一些实例中，复合图像可为两个或多于两个先前经解码关键帧的混合或单个先前经解码关键帧。场景结构图包含关于当前图像或先前经解码图像内的三维点的坐标值。

[0150] 解码器80可接收当前图像的场景结构图的信息或指示当前图像的场景结构图与先前经解码图像的场景结构图之间的差异的信息中的一者。在一些实例中，解码器80还可接收摄像机位置及摄像机定向中的一者或两者。

[0151] 在一些实例中，解码器80可接收当前图像的背景静态部分的残差图像。在这些实例中，解码器80可基于合成图像及关于当前图像的背景静态部分的残差数据而重构当前图像。

[0152] 应认识到，取决于实例，本文中所描述的诸技术中的任一者的特定动作或事件可以不同顺序执行，可经添加、合并或完全省去(例如，并非所有所描述动作或事件为实践技术所必需的)。此外，在某些实例中，可(例如)经由多线程处理、中断处理或多处理器同时地而非依序地执行动作或事件。



[0153] 在一或多个实例中,所描述的功能可以硬件、软件、固件或其任何组合实施。如果以软件实施,那么所述功能可作为一或多个指令或代码存储于计算机可读媒体上或者经由计算机可读媒体发射且由基于硬件的处理单元执行。计算机可读媒体可包含计算机可读存储媒体,其对应于例如数据存储媒体的有形媒体,或包含促进(例如)根据通信协议将计算机程序自一个位置传送到另一位置的任一媒体的通信媒体。以此方式,计算机可读媒体通常可对应于(1)非暂时性的有形计算机可读存储媒体或(2)例如信号或载波的通信媒体。数据存储媒体可为可由一或多个计算机或一或多个处理器存取以检索用于实施本发明中所描述的技术的指令、代码及/或数据结构的任何可用媒体。计算机程序产品可包含计算机可读媒体。

[0154] 通过实例的方式且非限制性,此些计算机可读存储媒体可包括RAM、ROM、EEPROM、CD-ROM或其它光学磁盘存储装置、磁盘存储装置或其它磁性存储装置、快闪存储器或可用于以指令或数据结构的形式存储所要代码且可由计算机存取的其它媒体。此外,任何连接恰当地称为计算机可读媒体。举例来说,如果使用同轴电缆、光纤电缆、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发射指令,那么所述同轴电缆、光纤电缆、双绞线、DSL或例如红外线、无线电及微波等无线技术皆包含于媒体的定义中。然而,应理解计算机可读存储媒体及数据存储媒体不包含连接、载波、信号或其它暂时性媒体,但替代地是针对非暂时性、有形存储媒体。如本文中所使用,磁盘及光盘包含光盘(CD)、激光光盘、光学光盘、数字多功能光盘(DVD)、软盘及蓝光光盘,其中磁盘通常以磁性方式再现数据,而光盘通过激光以光学方式再现数据。上述各项的组合还应包括在计算机可读媒体的范围内。

[0155] 指令可由一或多个处理器执行,例如一或多个数字信号处理器(DSP)、通用微处理器、专用集成电路(ASIC)、场可编程逻辑阵列(FPGA)或其它等效集成或离散逻辑电路。因此,如本文中所使用的术语“处理器”可指前述结构或适于实施本文中所描述的技术的任何其它结构中的任一者。另外,在一些方面中,本文中所描述的功能性可提供在经配置以用于编码及解码或并入于组合式编解码器中的专用硬件及/或软件模块内。此外,技术可以一或多个电路或逻辑元件来完全实施。

[0156] 本发明的技术可以广泛各种装置或装备(包含无线手机、集成电路(IC)或IC组(例如,芯片组))实施。各种组件、模块或单元在本发明中经描述以强调经配置以执行所揭示技术的装置的功能方面,而未必需要由不同硬件单元实现。确切来说,如上文所描述,各种单元可以编解码硬件单元组合或通过交互操作硬件单元(包含如上文所描述的一或多个处理器)的集合结合适合软件及/或固件提供。

[0157] 已描述各种实例。这些及其它实例在以下权利要求书的范围内。

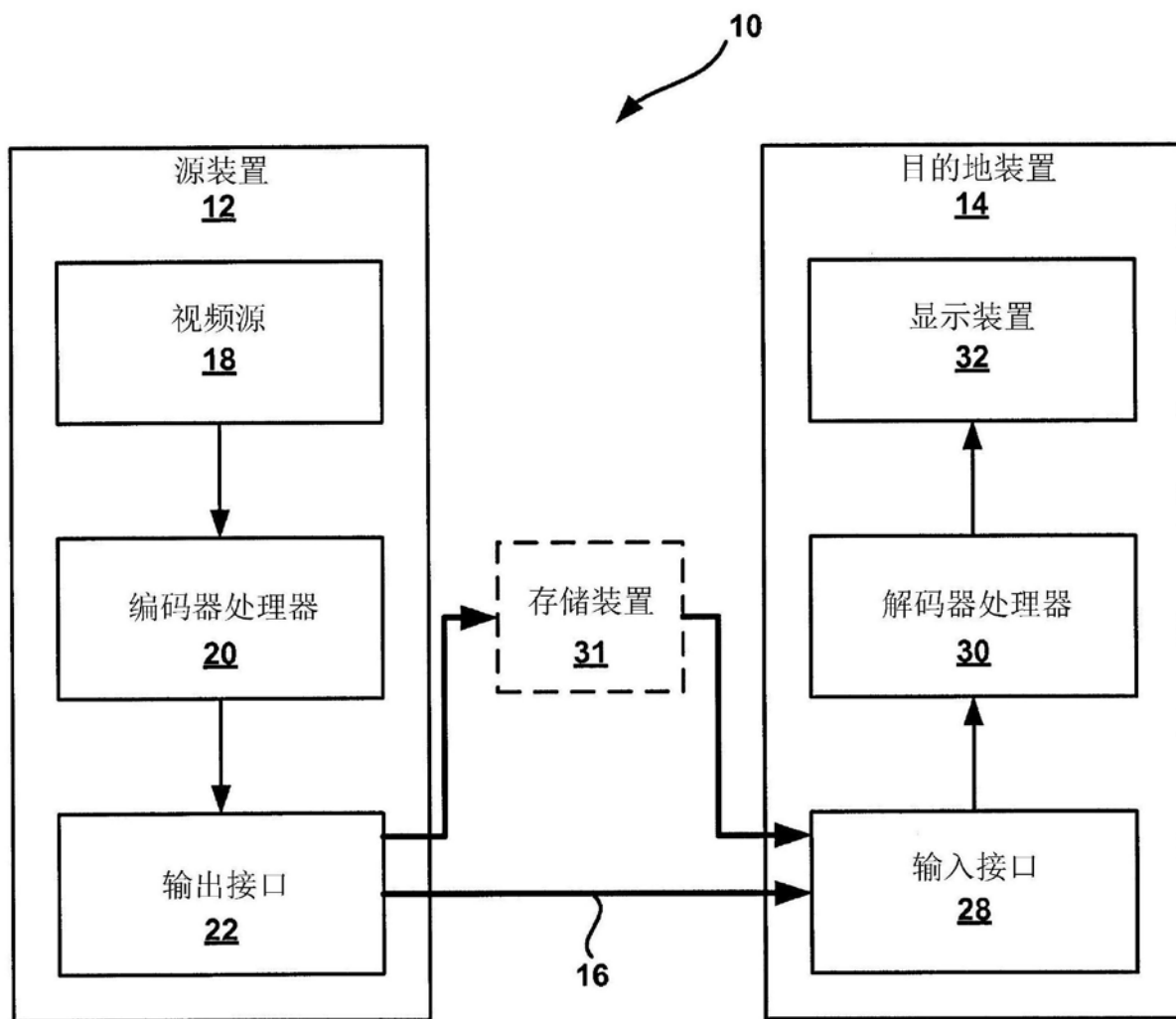


图1

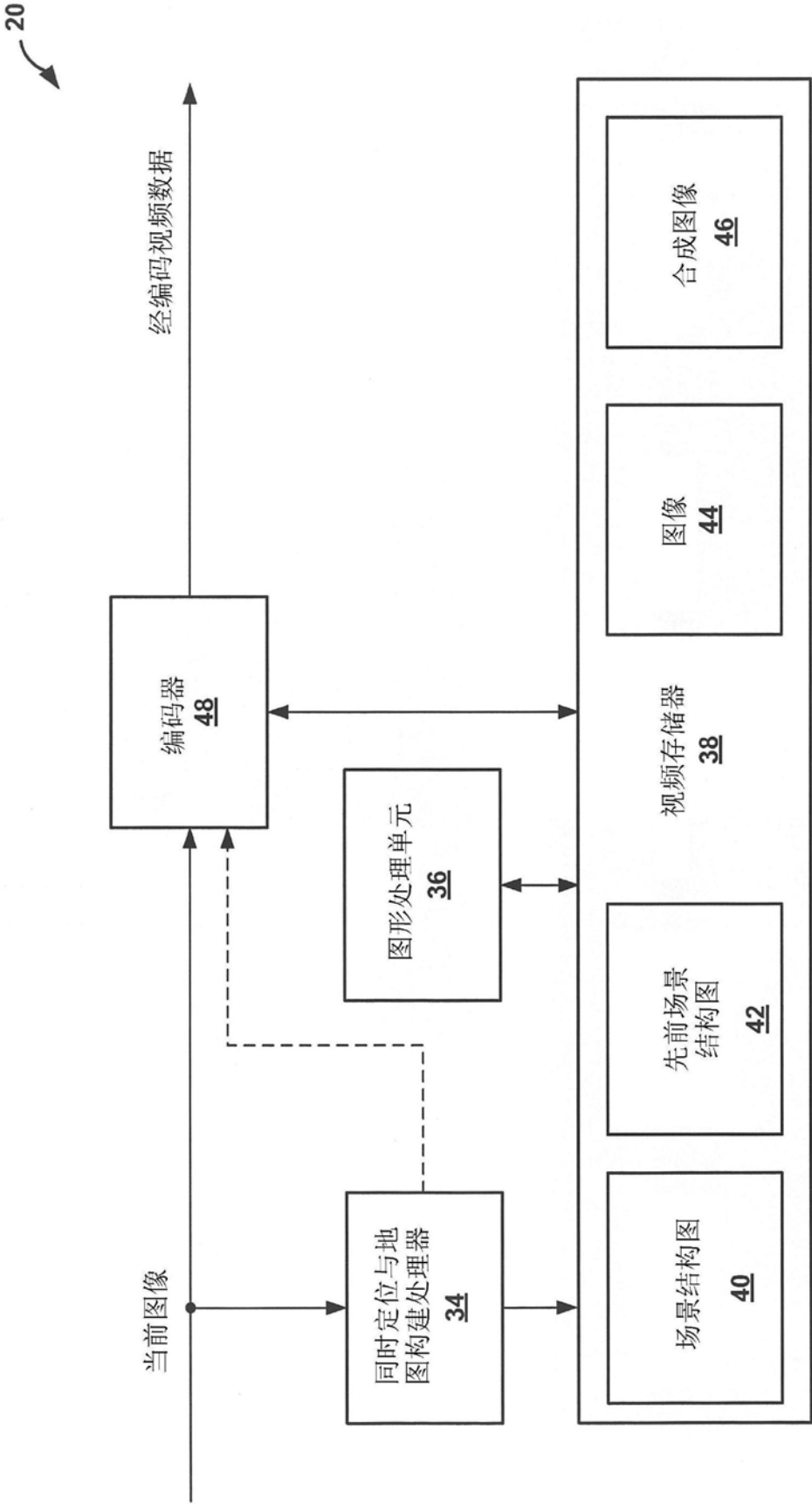


图2

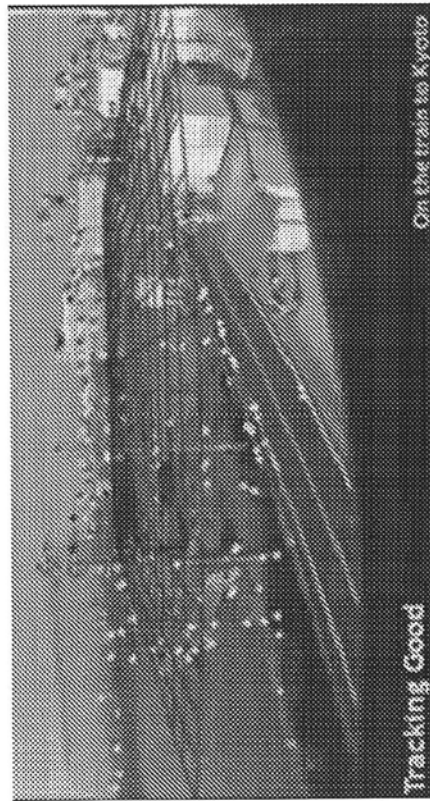


图3A

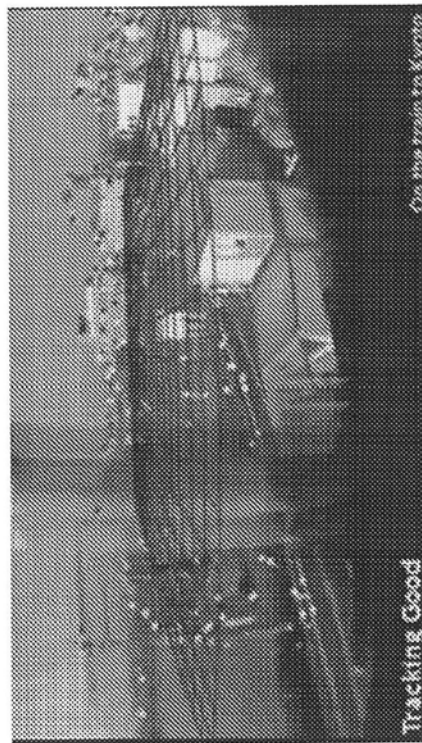


图3B

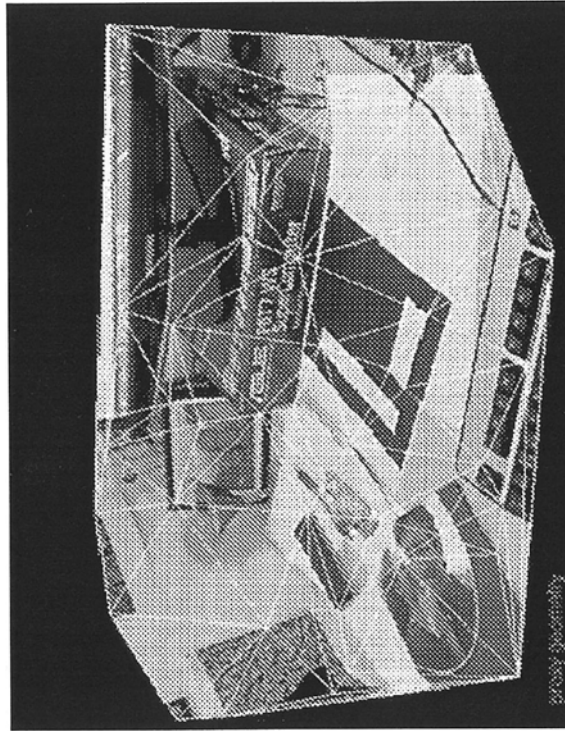


图4

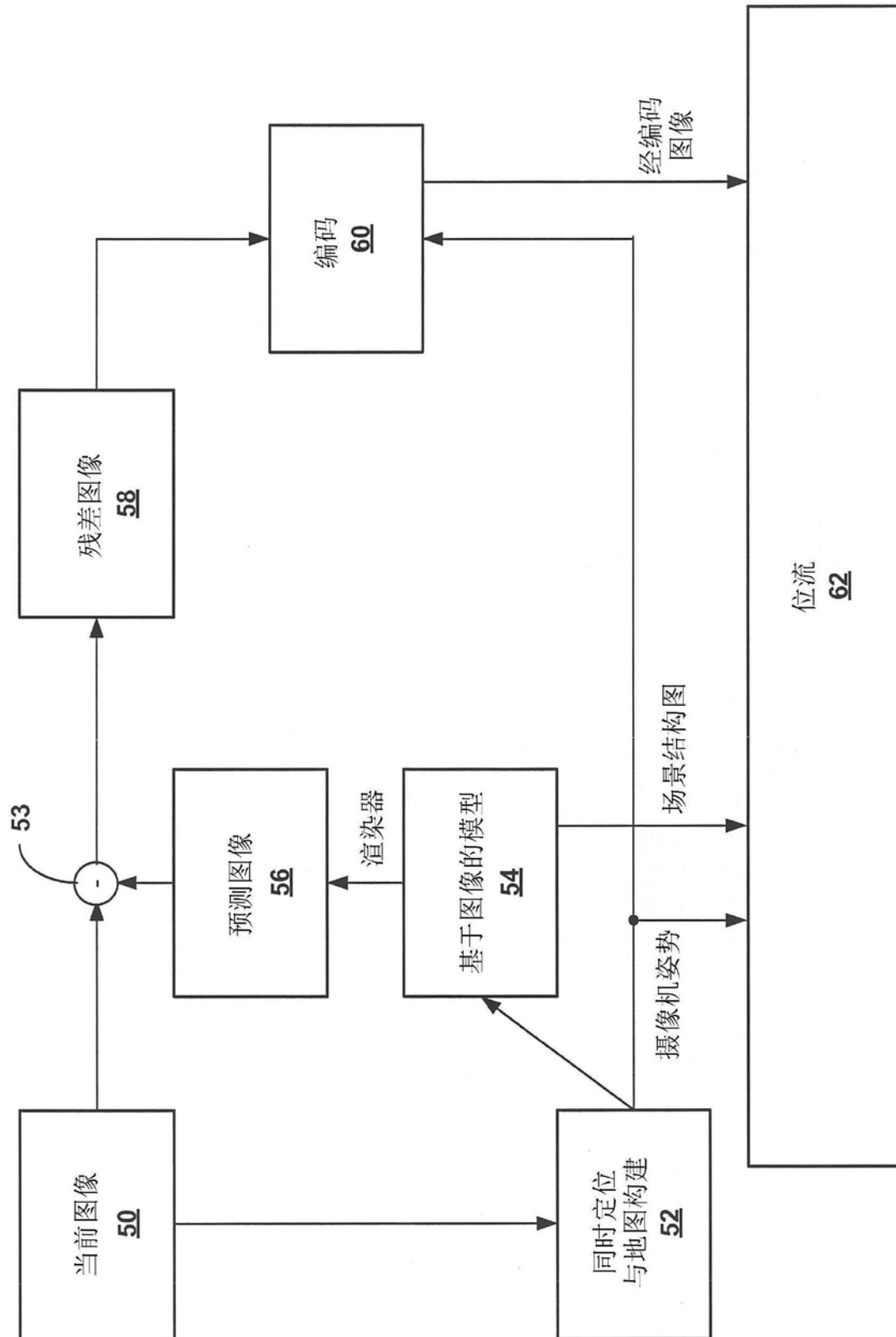


图5



图6

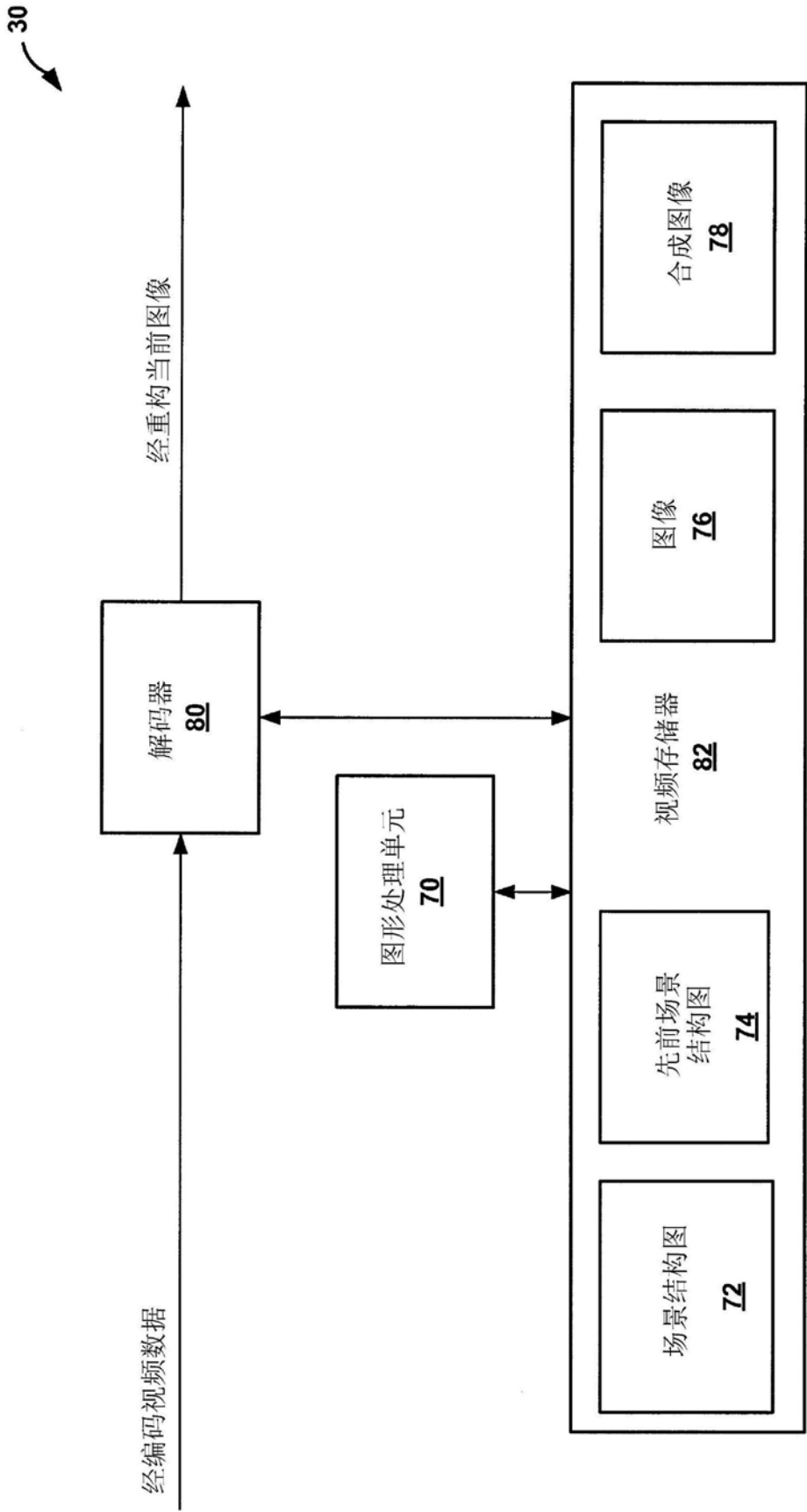


图7



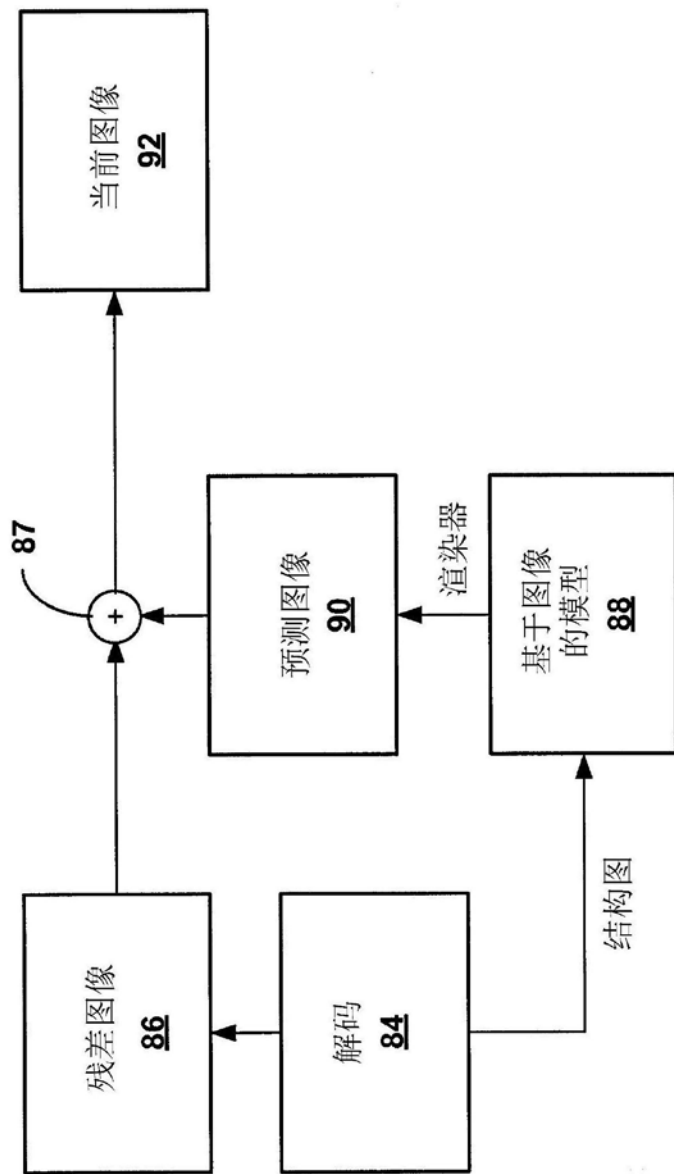


图8

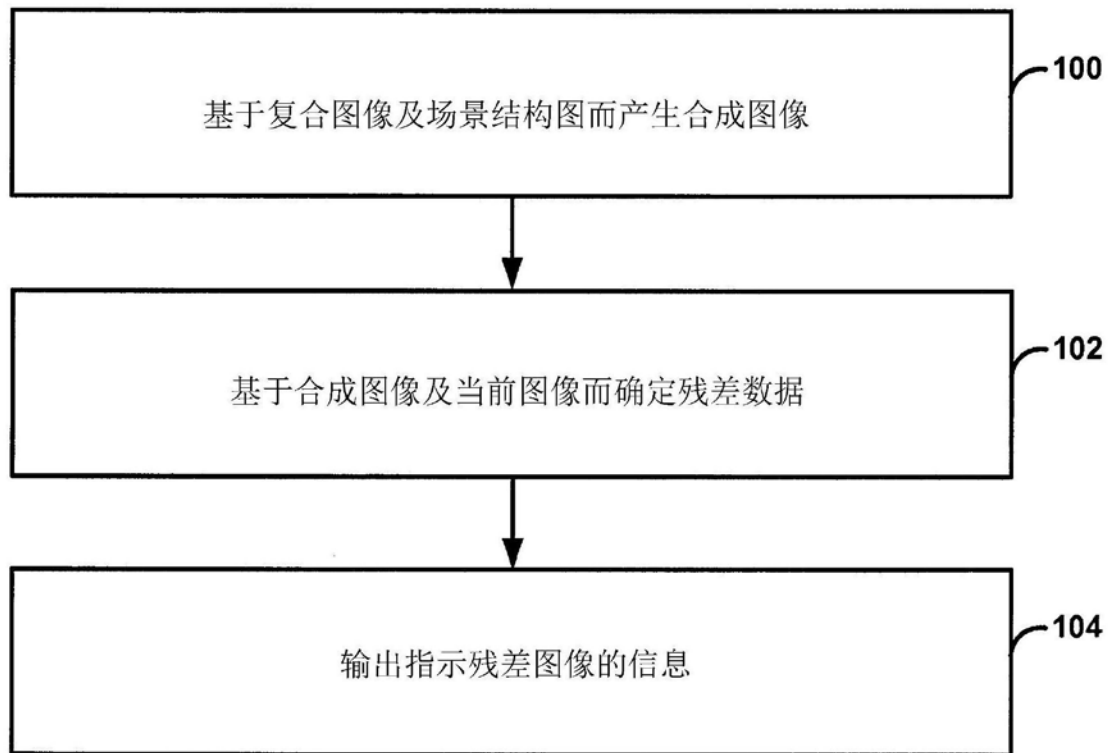


图9

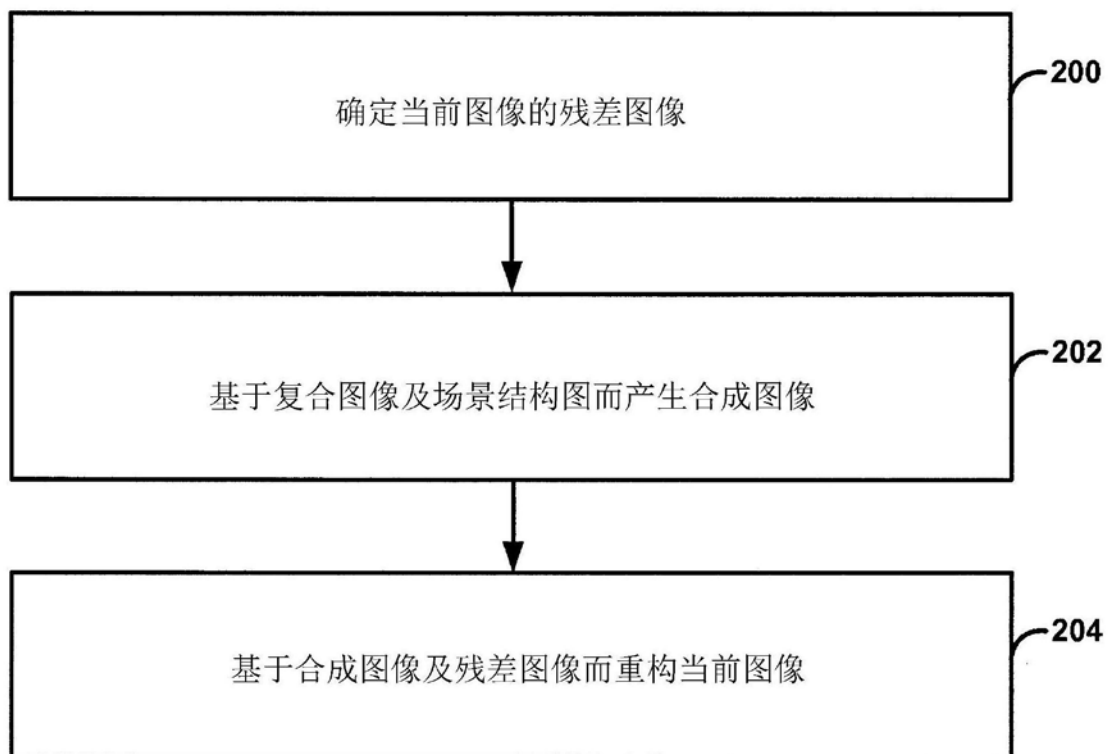


图10