



(12) 发明专利

(10) 授权公告号 CN 109074161 B

(45) 授权公告日 2021.10.01

(21) 申请号 201780024291.X

(22) 申请日 2017.03.10

(65) 同一申请的已公布的文献号
申请公布号 CN 109074161 A

(43) 申请公布日 2018.12.21

(30) 优先权数据
62/332,918 2016.05.06 US
15/269,157 2016.09.19 US

(85) PCT国际申请进入国家阶段日
2018.10.18

(86) PCT国际申请的申请数据
PCT/US2017/021914 2017.03.10

(87) PCT国际申请的公布数据
W02017/192207 EN 2017.11.09

(73) 专利权人 高通股份有限公司
地址 美国加利福尼亚州

(72) 发明人 维贾雅拉克希米·拉贾孙达拉姆·

拉韦恩德拉恩

米纳·艾曼·萨利赫·扬尼·马卡尔

(74) 专利代理机构 永新专利商标代理有限公司
72002

代理人 赵腾飞

(51) Int.Cl.
G06F 3/01 (2006.01)
G06T 9/00 (2006.01)
H04N 19/70 (2014.01)

(56) 对比文件
US 2003095131 A1, 2003.05.22
US 2016012855 A1, 2016.01.14
US 2016104510 A1, 2016.04.14
US 2014098185 A1, 2014.04.10
CN 1625246 A, 2005.06.08

审查员 刘瑞

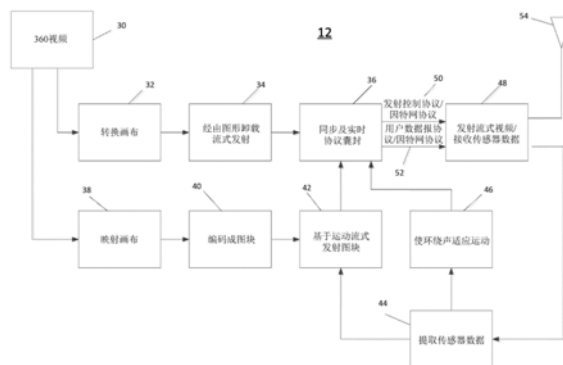
权利要求书3页 说明书11页 附图5页

(54) 发明名称

用于360度视频的混合图形及像素域架构

(57) 摘要

在用于处理视频数据的方法和设备中,一或多个处理器经配置以在像素域中编码所存储视频数据的一部分以产生像素域视频数据,第一图形处理单元经配置以在图形域中处理视频数据以产生图形域视频数据,且接口发射所述图形域视频数据及所述像素域视频数据。一或多个处理器经配置以将所述视频数据剖析成图形流及音频-视频流且对所述视频数据进行解码,传感器感测用户的运动适应性,且第二图形处理单元经配置以利用从所述图形流接收的纹理信息在球面表面上产生画布,并基于所述感测到的所述用户的运动适应性而显现视野。



1. 一种经配置以处理视频数据的设备,所述设备包括:

存储器,其经配置以存储所述视频数据;

一或多个处理器,其经配置以在像素域中编码所存储的视频数据的一部分以产生像素域视频数据,其中,所述像素域视频数据包括各自具有360视频图像的一部分的多个图块;

第一图形处理单元,其经配置以在图形域中处理所述视频数据以产生图形域视频数据,其中所述图形域视频数据包括具有所述360视频图像的全部的纹理;以及

接口,其用以发射包含具有所述360视频图像的所述全部的所述纹理的所述图形域视频数据及包含所述多个图块的子集的所述像素域视频数据的子集,

其中,所述一或多个处理器进一步经配置以将所述视频数据拼接在一起以形成等矩形画布,且其中,所述第一图形处理单元进一步经配置以将所述等矩形画布转换为所述纹理且在球体内部显现所述纹理,

其中,所述第一图形处理单元经配置以经由所述接口以第一帧率发射所述纹理,且其中所述一或多个处理器进一步经配置以经由所述接口以大于所述第一帧率的第二帧率发射所述像素域视频数据的所述子集,并且

其中,所述第一图形处理单元经配置以经由所述接口以第一分辨率发射所述纹理,且其中,所述一或多个处理器进一步经配置以经由所述接口以大于所述第一分辨率的第二分辨率发射所述像素域视频数据的所述子集。

2. 根据权利要求1所述的设备,其中,所述一或多个处理器经配置以将所述等矩形画布映射到立方图或金字塔形投影中的一者,并以多种分辨率编码所述多个图块,且其中,所述多个图块的所述子集在用户视野内。

3. 根据权利要求2所述的设备,其中,所述一或多个处理器经配置以确定所述用户的运动适应性,并基于所确定的所述用户的运动适应性来确定所述视野。

4. 根据权利要求2所述的设备,其中,所述一或多个处理器经配置以确定所述用户的运动适应性,基于所确定的所述用户的运动适应性来确定所述视野,且基于经确定的视野来调适环绕声视频。

5. 根据权利要求3所述的设备,其中,所述一或多个处理器经配置以在第三分辨率下基于所述运动适应性发射所述图块子集中处于所述视野的中心的图块,且在低于所述第三分辨率的第四分辨率下发射所述图块子集中在所述视野内但不在所述视野的所述中心的图块。

6. 一种经配置以处理视频数据的设备,所述设备包括:

存储器,其经配置以存储包括像素域视频数据及图形域视频数据的视频流;

一或多个处理器,其经配置以将所存储的视频流剖析成图形流及音频-视频流,并对所剖析的视频流进行解码,其中,所解码的视频流包括像素域数据,所述像素域数据包含各自具有360视频图像的一部分的多个图块,且其中,所述多个图块总共包含少于所述360视频图像的全部;

传感器,其用以感测用户的运动适应性;及

图形处理单元,其经配置以利用来自所述图形流的纹理信息在球面表面上产生画布,并基于所感测的所述用户的运动适应性来显现视野,其中所述纹理信息包含所述360视频图像的所述全部,

其中,在所述图形流中的所述纹理信息是以第一帧率存在于所述图形流中的,并且所述像素域数据是以大于所述第一帧率的第二帧率存在于所解码的视频流中的,

其中,所述纹理信息是采用第一分辨率的,且所述像素域数据是采用大于所述第一分辨率的第二分辨率的。

7.根据权利要求6所述的设备,其中,所述一或多个处理器经配置以对处于所述用户的所述视野内的所述多个图块进行解码,对所述多个图块中是低分辨率图块的图块和所述多个图块中是高分辨率图块的图块进行上采样,并组合所述低分辨率图块和所述高分辨率图块以形成单一图像覆盖。

8.根据权利要求7所述的设备,其中,所述一或多个处理器经配置以基于所述用户的经感测运动适应性来形成所述单一图像覆盖。

9.根据权利要求8所述的设备,其中,所述一或多个处理器经配置以基于确定所述用户的所述经感测运动适应性导致经上采样图块在所述视野外而不形成所述单一图像覆盖。

10.根据权利要求9所述的设备,其中,所述用户的所述经感测运动适应性包括头部运动轨道。

11.根据权利要求9所述的设备,其中,所述一或多个处理器经配置以将所述用户的头部运动范围与可视性窗口进行比较。

12.一种处理视频数据的方法,其包括:

存储所述视频数据;

在像素域中编码所述视频数据的一部分以产生像素域视频数据,其中,所述像素域视频数据包括各自具有360视频图像的一部分的多个图块;

在图形域中处理所述视频数据以产生图形域视频数据,其中,所述图形域视频数据包括具有所述360视频图像的全部的纹理;

将所述视频数据拼接在一起以形成等矩形画布;

将所述等矩形画布转换为所述纹理;

在球体内部显现所述纹理;以及

发射包含具有所述360视频图像的所述全部的所述纹理的所述图形域视频数据及包含所述多个图块的子集的所述像素域视频数据的子集,其中,所述发射包括:

以第一帧率和第一分辨率来发射所述纹理;以及

以大于所述第一帧率的第二帧率和大于所述第一分辨率的第二分辨率来发射所述像素域视频数据的所述子集。

13.根据权利要求12所述的方法,其进一步包括:

将所述等矩形画布映射到立方图或金字塔形投影中的一者;及

以多种分辨率编码所述多个图块,其中所述多个图块中的所述子集在用户的视野内。

14.根据权利要求13所述的方法,其进一步包括:

确定所述用户的运动适应性;及

基于所述用户的经确定运动适应性来确定所述视野。

15.根据权利要求13所述的方法,其进一步包括:

确定所述用户的运动适应性;

基于所确定的所述用户的运动适应性来确定所述视野;及

基于所确定的视野来调适环绕声视频。

16. 根据权利要求14所述的方法,其进一步包括:

基于所确定的运动适应性,以第三分辨率发射所述图块子集中处于所述视野的中心的图块;及

以低于所述第三分辨率的第四分辨率发射所述图块子集中在所述视野内但不在所述视野的所述中心的图块。

17. 根据权利要求12所述的方法,其进一步包括:

解码所述视频数据及将所述视频数据剖析成图形流及音频-视频流;

感测用户的运动适应性;

利用从所述图形流接收的纹理信息在球面表面上产生画布;及

基于所感测的所述用户的运动适应性来显现视野。

18. 一种处理视频数据的方法,其包括:

存储包括像素域视频数据及图形域视频数据的视频流;

将所存储的视频流剖析成图形流及音频-视频流并解码所述视频流,其中,所解码的视频流包括像素域数据,所述像素域数据包含各自具有360视频图像的一部分的多个图块,且其中所述多个图块总共包含少于所述360视频图像的全部;

感测用户的运动适应性;

利用来自所述图形流的纹理信息在球面表面上产生画布,其中,所述纹理信息包含所述360视频图像的所述全部;及

基于所感测的所述用户的运动适应性来显现视野,

其中,在所述图形流中的所述纹理信息是以第一帧率存在于所述图形流中的,并且所述像素域数据是以大于所述第一帧率的第二帧率存在于所解码的视频流中的,

其中,所述纹理信息是采用第一分辨率的,且所述像素域数据是采用大于所述第一分辨率的第二分辨率的。

19. 根据权利要求18所述的方法,其进一步包括:

解码在所述用户的所述视野内的所述多个图块;

对所述多个图块中是低分辨率图块的图块和所述多个图块中是高分辨率图块的图块进行上采样;及

组合所述低分辨率图块和所述高分辨率图块以形成单一图像覆盖。

20. 根据权利要求19所述的方法,其进一步包括:基于所感测的所述用户的运动适应性来形成所述单一图像覆盖。

21. 根据权利要求20所述的方法,其进一步包括:基于确定所感测的所述用户的运动适应性导致经上采样的图块在所述视野外而不形成所述单一图像覆盖。

22. 根据权利要求21所述的方法,其中,所感测的所述用户的运动适应性包括头部运动轨道。

23. 根据权利要求21所述的方法,其进一步包括:将所述用户的头部运动范围与可视性窗口进行比较。

用于360度视频的混合图形及像素域架构

技术领域

[0001] 本申请案要求保护2016年5月6日申请的美国临时申请案第62/332,918号的权益,其每一者的全部内容以引用的方式并入本文中。

[0002] 本发明涉及360度视频的译码及发射。

背景技术

[0003] 由于当前360度视频技术,用户所体验的视频环境已变得与视频本身的主题一样重要。此类360视频技术可涉及从360度摄像机或网站到实时视频显示器(例如虚拟现实(VR)头戴式显示器(HMD))的360视频图形的实时流式发射及/或360视频的实时流式发射。VR HMD允许用户通过转头改变视角来体验在其周围发生的事情。为了创建360度视频,可使用特殊的相机集合来同时记录所有360度场景,或多个视图(例如视频及/或计算机产生的图像)可拼接在一起以形成图像。

发明内容

[0004] 大体来说,本发明描述用于360视频数据的流式发射的技术。在一个实例中,本发明提出视频编码器,其经配置以在像素域中编码360视频数据的一部分以产生像素域视频数据;第一图形处理单元,其用以在图形域中处理360视频数据以产生图形域视频数据;及接口,其用以发射图形域视频数据及像素域视频数据以用于流式发射360视频数据。

[0005] 在一个实例中,本发明提出一种经配置以处理视频数据的设备。存储器经配置以存储视频数据,一或多个处理器经配置以在像素域中编码所存储视频数据的一部分以产生像素域视频数据,且第一图形处理单元经配置以在图形域中处理视频数据以产生图形域视频数据。接口经配置以发射图形域视频数据及像素域视频数据作为视频流。

[0006] 在另一个实例中,本发明进一步提出一种经配置以处理视频数据的设备。存储器经配置以存储包括像素域视频数据及图形域视频数据的视频流,且一或多个处理器经配置以将所存储视频流剖析成图形流及音频-视频流,并对所述视频流进行解码。传感器感测用户的运动适应性,且图形处理单元经配置以利用来自图形流的纹理信息在球面表面上产生画布(canvas),并基于用户的经感测运动适应性而显现视野。

[0007] 在另一实例中,处理视频数据的方法包括:存储视频数据;在像素域中编码视频数据的一部分来产生像素域视频数据;在图形域中处理视频数据以产生图形域视频数据;及发射图形域视频数据及像素域视频数据。在另一实例中,处理视频数据的方法包括:存储包括像素域视频数据及图形域视频数据的视频流;将所存储视频流剖析成图形流及音频-视频流并对所述视频流进行解码;感测用户的运动适应性;利用来自图形流的纹理信息在球面表面上产生画布;及基于用户的经感测运动适应性而显现视野。

[0008] 在一个实例中,本发明提出流式发射360视频数据的方法,所述方法包括:在图形域中流式发射360视频数据;在像素域中编码360视频数据的至少一个部分;及发射图形域中的流式360视频数据及像素域中的360视频数据的至少一个部分。

[0009] 在另一个实例中,本发明进一步提出提取传感器数据及处理所提取的传感器数据来确定用户的运动适应性。在另一实例中,本发明进一步提出,在图形域中流式发射360视频数据包括将图形域中的360视频数据转换为纹理以供图形处理器执行,以便在球体内部显现360视频数据,施加用于同步的时戳及实时协议囊封视频数据以供传送。在另一实例中,本发明进一步提出,在像素域中流式发射360视频数据的至少一部分包括将360视频数据映射到立方图、等矩形及金字塔形投影中的一者,在各种分辨率下将360视频数据的至少一个部分编码成多个图块,及基于用户的经确定运动适应性来流式发射多个图块的一部分。

[0010] 在另一实例中,本发明提出接收流式360视频数据,将流式360视频数据剖析成图形流及音频-视频流,在球面表面上产生画布,感测用户的运动,及基于所产生的画布及所感测的运动而显现视野。在另一实例中,本发明提出一种源装置,所述源装置包括:处理器,其经配置以在图形域中产生并流式发射360视频数据,在像素域中产生并流式发射360视频数据,发射流式360视频数据;目的地装置,其用以接收所发射的流式360视频且包括处理器,所述处理器经配置以将流式360视频数据剖析成图形流及音频-视频流,在球面表面上产生画布,感测用户的动作,且基于所产生的画布及所感测的运动显现视野。

[0011] 在另一实例中,本发明提出将图形域中的360视频数据转换成纹理以供由图形处理器执行,以用于在球体内部显现360视频数据,施加用于同步的时戳及实时协议囊封视频数据以供传送。

[0012] 在另一实例中,本发明进一步提出存储指令的非暂时性计算机可读储存媒体,所述指令在经执行时致使一或多个处理器执行一种方法,所述方法包括流式发射360视频数据的方法的任何组合。

[0013] 在附图和以下描述中阐述本发明的一或多个方面的细节。本发明中所描述的技术的其它特性、目标和优点将从描述、图式且从权利要求书显而易见。

附图说明

[0014] 图1为可利用本发明中所描述的一或多种技术的用于处理视频数据的设备的框图。

[0015] 图2为说明可实施本发明中所描述的一或多种技术的实例源装置的框图。

[0016] 图3为说明可实施本发明中所描述的一或多种技术的实例装置的框图。

[0017] 图4为根据本发明的实例的用于处理视频数据的方法的流程图。

[0018] 图5为根据本发明的实例的用于处理视频数据的方法的流程图。

具体实施方式

[0019] 与视频有关的数据的流式发射有利地快速发生,以避免用户在视野(FoV)变化期间体验到的滞后时间效果,而不损害视频质量或分辨率。确切地说,由于必需的编码及解码开销,用于360视频的纯像素域架构(例如仅由视频编码/解码技术组成的架构)可施加延时负担,而如果希望发射(例如实时无线发射)全部360度视频,那么纯图形域架构可具有过高纹理开销。

[0020] 本发明描述利用混合图形及像素域方法的用于视频数据的流视频系统的技术,其

中由图形域管线提供基线质量及立即反应,且由像素域管线提供高分辨率视频。另外,流视频系统还解决用于定向音频及视频同步的VR HMD应用程序的头部运动适应性。因此,本发明中所描述的技术及装置充分利用实时应用程序的像素及图形域流的强度,并结合可视性限制的窗口考虑人类视野特性。

[0021] 图1为可利用本发明中所描述的一或多种技术的用于处理视频数据的设备的框图。如图1中所示,用于流式发射视频的系统10包含视频源装置12,其提供稍后将由目的地装置14解码的经编码视频数据。确切地说,源装置12经由计算机可读媒体11将视频数据提供到目的地装置14。源装置12和目的地装置14可包括广泛范围的装置中的任一者,包含台式计算机、笔记型(即,膝上型)计算机、平板计算机、机顶盒、电话手持机(例如,所谓的“智能”电话、所谓的“智能”板)、电视机、相机、显示装置、数字媒体播放器、视频游戏控制台、视频流式发射装置或类似者。在一些情况下,可装备源装置12和目的地装置14以用于无线通信。在一个实例中,目的地装置14可为虚拟现实(VR)头戴式显示器(HMD),且视频源装置12可产生360视频且将360视频流式发射到目的地装置。

[0022] 目的地装置14可经由计算机可读媒体11接收待解码的经编码视频数据且将经编码视频数据存储在存储器29中。计算机可读媒体11可包括能够将经编码视频数据从源装置12移动到目的地装置14的任何类型的媒体或装置。在一个实例中,计算机可读媒体11可包括使得源装置12能够实时将经编码视频数据直接发射到目的地装置14的通信媒体。经编码视频数据可根据通信标准(例如,无线通信协议)来调制,且发射到目的地装置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传输线。通信媒体可形成基于包的的网络(例如,局域网、广域网或全球网络,例如因特网)的一部分。通信媒体可包含路由器、交换器、基站或任何其它可以用于促进从源装置12到目的地装置14的通信的设备。

[0023] 目的地装置14可经由包含因特网连接的任何标准数据连接来存取经编码的视频数据。这可包含无线通道(例如Wi-Fi连接)、有线连接(例如DSL、电缆调制解调器等)或两者的适合于存取存储在文件服务器上的经编码视频数据的组合。经编码视频数据从存储装置的发射可为流式发射、下载发射或其组合。

[0024] 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多种多媒体应用中的任一者的视频译码,例如空中电视广播、有线电视发射、卫星电视发射、因特网流视频发射(例如,HTTP动态自适应流式传输(DASH))、编码到数据存储媒体上的数字视频,存储在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,以支持例如视频流式发射、视频重放、视频广播和/或视频电话的应用。

[0025] 如图1的实例系统10中所说明,视频源装置12包含视频编码器16,其使用如下文详细描述的本发明的经组合图形域及像素域技术编码来自视频数据源17的视频数据(例如360视频数据)且经由发射输入/输出接口18发射经编码视频。在包含沿目的地装置14定位的一或多个传感器24的实例中,视频编码器16可经由输入/输出接口18从目的地装置14接收传感器数据,以便确定用户的运动,如下文所描述。视频数据也由将经处理视频数据发射到输入/输出接口18的图形处理单元(GPU)19从视频源17接收。此外,处理器15从视频源17接收视频数据且处理视频数据以供视频编码器16及/或GPU 19使用。源装置17可包含能够编码环绕声音频的麦克风(图中未展示)。

[0026] 目的地装置14包含视频解码器20及图形处理单元(GPU) 23。经编码视频数据可从视频源装置12无线发射且经由发射输入/输出接口22在目的地装置14处接收,并由处理器21处理。所得经处理的视频数据随后由解码器20解码,及/或由图形处理单元(GPU) 23用来经由显示处理器25及显示器27产生或显现视野(FoV),如下文详细描述。如果目的地装置14为虚拟现实(VR)头戴式显示器(HMD),那么还可包含例如一或多个传感器24(例如一或多个运动传感器)来感测用户的定向和运动。显示处理器25从视频解码器20接收经解码视频数据,同时从GPU 23接收所显现视野,且处理接收到的数据,以使360视频能够产生,从而供用户在目的地装置14的显示器27上查看。

[0027] 图1的所说明的系统10仅为一个实例。根据本发明的技术的用于译码及发射视频数据的技术可通过任何数字视频编码及/或解码装置执行。虽然大体上本发明的技术通过视频编码或解码装置执行,但是所述技术还可以通过视频编解码器执行。此外,本发明的技术还可通过视频处理器执行。源装置12和目的地装置14仅为此类译码装置的实例,其中源装置12产生用于发射到目的地装置14的经译码视频数据。

[0028] 源装置12的视频源17可包含视频捕捉装置(例如360度摄像机系统、摄像机)、含有先前捕捉的视频的视频存档及/或用以从视频内容提供者接收视频的视频馈入接口。作为另一个实例,视频数据源17可产生基于计算机图形的数据作为源视频,或实况视频、存档视频和计算机产生的视频的组。在一些情况下,如果视频数据源17为摄像机,那么源装置12和目的地装置14可形成所谓的相机电话或视频电话。然而,如上文所提及,本发明中所描述的技术一般可适用于视频译码,且可应用于无线和/或有线应用。在每一情况下,可由视频编码器16编码所捕捉、预先捕捉或计算机产生的视频。经编码视频信息可接着通过输出接口18输出到计算机可读媒体11上。

[0029] 计算机可读媒体11可包含瞬时媒体,例如无线广播或有线网络发射,或存储媒体(即,非暂时性存储媒体),例如硬盘、快闪驱动器、压缩光盘、数字视频光盘、蓝光光盘或其它计算机可读媒体。在一些实例中,网络服务器(图中未展示)可从源装置12接收经编码视频数据且例如经由网络发射将经编码视频数据提供到目的地装置14。类似地,媒体生产设施(例如,光盘冲压设施)的计算装置可从源装置12接收经编码视频数据且生产含有经编码视频数据的光盘。因此,在各种实例中,计算机可读媒体11可理解为包含各种形式的一或多个计算机可读媒体。

[0030] 目的地装置14的输入接口22从计算机可读媒体11接收信息。计算机可读媒体11的信息可包含由视频编码器16定义的语法信息,所述语法信息还供视频解码器20使用,所述语法信息包含描述块及其它经译码单元的特性及/或处理的语法元素。显示装置27向用户显示经解码视频数据且可包含多种显示装置中的任一者,例如360度视频显示器及VR HMD、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。

[0031] 视频编码器16和视频解码器20各自可实施为可适用的多种合适的编码器或解码器电路中的任一者,例如,一或多个微处理器、数字信号处理器(DSP)、专用集成电路(ASIC)、现场可编程门阵列(FPGA)、离散逻辑电路、软件、硬件、固件或其任何组合。当部分地以软件实施所述技术时,装置可将用于所述软件的指令存储于合适的非暂时性计算机可读媒体中且使用一或多个处理器以硬件执行所述指令以执行本发明的技术。视频编码器16

和视频解码器20中的每一者可包含在一或多个编码器或解码器中,所述编码器或解码器中的任一者可集成为组合式视频编码器/解码器(编解码器)的部分。包含视频编码器16和/或视频解码器20的装置可包括集成电路、微处理器和/或无线通信装置,例如,蜂窝式电话。

[0032] 显示处理器25可为经配置以对图像数据执行2D操作(例如上采样、下采样、共混、复合、缩放、旋转及其它像素处理)的处理器。显示处理器25还可经配置以驱动显示器27。显示处理器25从存储器29(例如,GPU将呈像素表示形式的图形数据输出到的帧缓冲器及/或其它存储器)预取或提取多个图像层(例如像素域视频数据及图形域视频数据)并将这些层复合为单一图像以供显示。显示处理器25可处理来自多个层的像素。举例来说,显示处理器25可共混来自多个层的像素,并将经共混像素以图块格式写回到存储器中。接着,以光栅格式从存储器读取经共混像素,并将其发送到显示器27以供呈现。

[0033] 视频编码器16和视频解码器20可根据视频压缩标准(例如,H.264/AVC或HEVC标准)操作。然而,本发明的技术不限于任何特定译码标准,且可使用任何视频译码技术。

[0034] 在HEVC中,视频图片可分成包含亮度和色度样本两者的树块或最大译码单元(LCU)的序列。位流内的语法数据可界定LCU(就像素数目来说,其为最大译码单位)的大小。切片包含多个连续译码树单元(CTU)。CTU中的每一个可包括明度样本的译码树块、色度样本的两个对应的译码树块,以及用于对译码树块的样本进行译码的语法结构。在黑白图片或具有三个单独颜色平面的图片中,CTU可以包括单个译码树块和用于对所述译码树块的样本进行译码的语法结构。

[0035] 视频图片可分割成一或多个切片。每一树块可以根据四叉树分裂成译码单元(CU)。一般来说,四叉树数据结构包含每个CU一个节点,其中根节点对应于树块。如果CU分裂成四个子CU,那么对应于CU的节点包含四个叶节点,其中叶节点中的每一者对应于所述子CU中的一者。CU可包括具有亮度样本阵列、Cb样本阵列和Cr样本阵列的图像的亮度样本的译码块和色度样本的两个对应译码块,以及用于对译码块的样本进行译码的语法结构。在单色图片或具有三个单独颜色平面的图片中,CU可包括单个译码块及用于对所述译码块的样本进行译码的语法结构。译码块是样本的 $N \times N$ 块。

[0036] 四叉树数据结构的每一节点可提供对应的CU的语法数据。举例来说,四叉树中的节点可包含分裂标志,其表明对应于所述节点的所述CU是否分裂成子CU。CU的语法元素可以递归地定义,且可以取决于所述CU是否分成若干子CU。如果CU未进一步分裂,那么将其称作叶CU。在本发明中,叶CU的四个子CU也将被称作叶CU,即使不存在原始叶CU的明确分裂时也是如此。举例来说,如果 16×16 大小的CU不进一步分裂,那么尽管 16×16 CU从未分裂,四个 8×8 子CU也将被称作叶CU。

[0037] CU具有与H.264标准的宏块类似的目的,除了CU不具有大小区别之外。举例来说,树块可分裂成四个子节点(还称为子CU),并且每一子节点又可为父节点并且可分裂成另外四个子节点。最后未分裂的子节点(被称作四叉树的叶节点)包括译码节点,也被称作叶CU。与经译码位流相关联的语法数据可定义树块可分裂的最大次数,被称作最大CU深度,且还可定义译码节点的最小大小。因此,位流还可定义最小译码单元(SCU)。本发明使用术语“块”来指代HEVC的背景下的CU、PU或TU中的任一者或其它标准的背景下的类似数据结构(例如H.264/AVC下的宏块及其子块)。

[0038] CU包含译码节点和与所述译码节点相关联的预测单元(PU)和变换单元(TU)。CU的

大小对应于译码节点的大小并且形状必须是正方形。CU的大小范围可从 8×8 像素到具有最大 64×64 像素或更大的树块的大小。每一CU可含有一或多个PU和一或多个TU。

[0039] 一般来说,PU表示对应于相对应的CU的全部或一部分的空间区域,并且可包含用于检索PU的参考样本的数据。此外,PU包含与预测相关的数据。举例来说,当PU经帧内模式编码时,用于PU的数据可以包含在残余四叉树(RQT)中,所述残余四叉树可包含描述用于对应于PU的TU的帧内预测模式的数据。作为另一实例,当PU经帧间模式编码时,PU可包含定义PU的一或多个运动向量的数据。预测块可为其上应用相同预测的样本的矩形(即,正方形或非正方形)块。CU的PU可包含图片的亮度样本的预测块、图片的色度样本的两个对应的预测块,以及用以对预测块样本进行预测的语法结构。在黑白图片或具有三个单独彩色平面的图片中,PU可以包括单个预测块和用于对预测块样本进行预测的语法结构。

[0040] 在将变换(例如,离散余弦变换(DCT)、整数变换、小波变换或概念上类似的变换)应用于残余视频数据之后,TU可包含变换域中的系数。所述残余数据可对应于未经编码图片的像素与对应于PU的预测值之间的像素差。视频编码器20可形成包含用于CU的残余数据的TU,并且随后变换TU以产生用于CU的变换系数。变换块可为应用相同变换的样本的矩形块。CU的变换单元(TU)可包括明度样本的变换块、色度样本的两个对应变换块和用以对变换块样本进行变换的语法结构。在黑白图片或具有三个单独彩色平面的图片中,TU可以包括单个变换块和用于对变换块样本进行变换的语法结构。

[0041] 在变换之后,视频编码器16可以执行变换系数的量化。量化一般是指变换系数经量化以可能地减少用于表示系数的数据量从而提供进一步压缩的过程。量化过程可减小与系数中的一些或全部相关联的位深度。举例来说,可在量化期间将n位值向下舍入到m位值,其中n大于m。

[0042] 视频编码器16可以扫描变换系数,从而从包含经量化变换系数的二维矩阵产生一维向量。所述扫描可经设计以将较高能量(并且因此较低频率)的系数放置在阵列的前面,并且将较低能量(并且因此较高频率)的系数放置在阵列的后面。在一些实例中,视频编码器16可利用预定义扫描顺序来扫描经量化变换系数以产生可被熵编码的串行化向量。在其它实例中,视频编码器16可以执行自适应扫描。

[0043] 在扫描经量化变换系数以形成一维向量之后,视频编码器16可例如根据上下文自适应可变长度译码(CAVLC)、上下文自适应二进制算术译码(CABAC)、基于语法的上下文自适应二进制算术译码(SBAC)、概率区间分割熵(PIPE)译码或另一熵编码方法对一维向量进行熵编码。视频编码器16还可以对与经编码视频数据相关联的语法元素进行熵编码以供视频解码器20在解码视频数据时使用。

[0044] 视频编码器16可进一步向视频解码器20发送语法数据,例如基于块的语法数据、基于图片的语法数据及基于图片群组(GOP)的语法数据,例如在图片标头、块标头、切片标头或GOP标头中。GOP语法数据可描述相应GOP中的数个图片,且图片语法数据可指示用以对对应图片进行编码的编码/预测模式。

[0045] 在获得经译码视频数据后,视频解码器20可以执行通常与相对于视频编码器16描述的编码遍次互逆的解码遍次。举例来说,视频解码器20可获得表示来自视频编码器16的经编码视频切片的视频块及相关联的语法元素的经编码视频位流。视频解码器20可使用位流中所含有的数据来重构原始的未经编码视频序列。

[0046] 在HEVC中,对于每一块,运动信息集合可为可用的。运动信息集合含有用于前向和后向预测方向的运动信息。如本文所描述,前向和后向预测方向是双向预测模式的两个预测方向,且术语“前向”和“后向”不一定暗示几何方向。实情为,如本文中所使用,“前向”及“后向”分别对应于当前图片的参考图片列表0 (RefPicList0) 及参考图片列表1 (RefPicList1)。在仅一个参考图片列表可用于图片或切片的情况下,视频编码器16及/或视频解码器20可确定仅RefPicList0。因此,在仅一个参考图片列表可用于当前图片或切片的情况下,视频编码器16及/或视频解码器20可确定当前图片/切片的每一块的运动信息始终是“前向的”。

[0047] 对于每个预测方向,运动信息含有参考索引和运动向量。在一些情况下,为简单起见,视频编码器16可编码运动向量,使得所述运动向量本身可以假设其具有相关联参考索引的方式而被引用。举例来说,视频解码器20可重构运动向量,且基于所述运动向量,视频解码器20可使特定参考索引与运动向量相关联。更具体地说,视频编码器16及/或视频解码器20可使用参考索引来识别相对于对应的运动向量的当前参考图片列表(例如RefPicList0或RefPicList1)中的特定参考图片。运动向量具有水平和垂直分量。

[0048] 图片次序计数(POC)广泛用于视频译码标准以识别图片的显示次序。在一些实例中,尽管视频编码器16可将一个经译码视频序列内的两个图片编码成具有相同POC值,但通常,单一经译码视频序列可不包含具有相同POC值的多个图片。在多个经译码视频序列存在于位流中的情况下,就解码次序来说,具有相同POC值(但处于不同视频序列中)的图片可彼此相对接近。视频编码器16及/或视频解码器20通常可将图片的POC值用于参考图片列表构建、如在HEVC中的参考图片集合导出及运动向量缩放。

[0049] 图2为说明可实施本发明中所描述的一或多种技术的实例360视频编码器的框图。如图2中所示,例如可由360摄像机系统捕捉的视频数据通过处理器15拼接在一起以形成等矩形画布(30)。处理器15将画布转换为纹理(32)以供由图形处理器(例如图1的GPU 19)(例如用于在球体内部显现360视频纹理的OpenGL VR应用程序)执行(32)。举例来说,在框32中由GPU 19进行的可包含低或高分辨率视频数据的纹理转换可为每个刷新周期(例如每隔一秒)更新。经由图形卸载利用具有图1的的GPU 19的可缩放流式发射功能执行纹理的流式发射(34),其可包含无损压缩。GPU 19可施加用于同步的时戳且流数据经实时协议(RTP)封装以供传送(36)。由GPU 19流式发射的经发射纹理可以相对低的帧率(例如每秒2帧(fps))发送。在本发明的一个实例中,由GPU 19流式发射的纹理使得显示器(例如VR HMD)的用户能够使用较低分辨率的整个360视频图像。可实时调适使用GPU 19显现的视频的FoV,是因为利用呈现用户的头部运动的本地传感器数据在目的地装置14处应用自适应。另外,不会对应于FoV变化来改变或刷新所需纹理。

[0050] 除了通过块32至36中所示的GPU 19在图形域中流式发射视频数据之外,块30的经拼接视频还由视频编码器16在像素域中编码且发射。像素域中的经编码视频可以较高帧率(例如60fps)及较高分辨率发送。通过这种方式,360视频中的运动对象可经编码并以较好平滑度更详细地显现。

[0051] 举例来说,处理器15可将画布映射到立方图或金字塔形投影(38),且视频编码器16可在各种分辨率下将流视频编码于图块中(40)。视频编码器16可经配置以使用任何视频压缩技术编码经拼接画布的视频数据的图块。举例来说,视频编码器16可使用混合视频编

码器(例如H.264/AVC或H.265(HEVC))来编码图块。视频编码器16可经配置以基于用户的经感测头部运动来发射图块的子集(42)。实际发射的图块的子集可为当前处于用户的视野(FoV)内的那些图块。为了基于用户的头部运动流式发射子集图块(42),视频编码器16还从检测用户的头部运动的一或多个传感器24提取传感器数据,并使用感测到的数据来处理用户的运动(44),有助于基于用户的头部运动来流式发射子集图块(42)。在本发明的一个实例中,由源装置12发送的图块的子集包含处于根据传感器数据确定的FoV的中心的高分辨率图块,且还发送对应于周边视觉区域中的图像的在空间上环绕高分辨率图块的低分辨率图块。用户的眼睛很可能仅感测到屏幕中心处的较高分辨率。由于用户的眼睛未聚焦到用户视野周界上的图块上,所以以较高分辨率发送这些图块没有益处。因此,可使用较低分辨率,从而节省带宽。这可使得更容易保持视野中心处的分辨率较高,同时还维持高帧率。使用对应于用户的头部运动的传感器数据在源装置12及目的地装置14处进行像素域数据的FoV自适应。在源装置12处,待使用的图块的分辨率可基于FoV自适应确定,以使得最高质量(分辨率)是在FoV中心处,其中质量朝向FoV外围减少。源装置12处的FoV自适应应为大规模自适应且可引发相比目的地装置14处的头部运动的时序的延迟。在目的地装置14处,FoV自适应可包含与头部运动的时序相关联的校正以解决头部运动的较大移位发生在传感器数据捕捉到头部运动的时间与同所捕捉头部运动相关联的传感器数据到达源装置12的时间之间发生的情况。

[0052] 另外,在框44中提取及处理的传感器数据可由视频编码器16利用,以使环绕音频适应用户运动(例如头部运动),以便使运动与图形背景同步(46)。举例来说,一旦视频编码器16感测到用户的头部运动,那么相应地通过确定哪些像素在视角内来改变显示器的视角。在另一实施例中,环绕音频可全部经流式发射,且定向音频可基于目的地装置14处(即,虚拟现实(VR)头戴式显示器(HMD))处的头部运动增强。在施加用于同步的时戳及用于传送的实时协议(RTP)囊封(36)期间包含流式发射图块(42)及使环绕音频适应用户运动(例如头部运动)(46)两者。举例来说,图形帧、(图块的)经编码视频帧及音频帧可针对相同参考时钟加上时戳,且每一者的采样率可不同。然而,所施加的时戳对应于在从360视频源和环绕音频源捕捉时的图像和音频。

[0053] 经由天线54,视频编码器16使用发射控制协议/因特网协议(TCP/IP)50或用户数据报协议/因特网协议(UDP/IP)52通过WiFi/WiGig发射来自框36的所得带时戳的流式视频(框48)。

[0054] 通过这种方式,流视频系统10利用混合图形及像素域方法,其中由图形域管线提供基线质量及立即反应(框32至36),且由像素域管线提供高分辨率视频(框38至46)。另外,流视频系统10还解决头部运动适应性及定向音频和视频同步。因此,本发明充分利用实时应用程序的像素及图形域流的强度,并结合可视性限制的窗口考虑人类视野特性。图形域的益处是,即使有频繁头部运动,视频的基线质量也连续。像素域覆盖为当前FoV提供高质量帧率及分辨率。

[0055] 图3为说明可实施本发明中所描述的一或多种技术的实例目的地装置14的框图。如图3中所示,目的地装置14接收如上文所描述经由天线60从源装置12无线发射的所得带时戳的流式视频(36),且使用发射控制协议/互联网协议(TCP/IP)或用户数据报协议/因特网协议(UDP/IP)接收WiFi/WiGig信号(62)。目的地装置14将接收到的视频流剖析成图形流

和音频-视频流(64)。GPU 23利用接收到的纹理信息在球面表面上产生或创建整个画布,(66)且基于从一或多个传感器24接收的用户的最近经感测头部运动(74)来显现视野(FoV),即,视角(68)。

[0056] 另外,目的地装置14的视频解码器20对经剖析视频流进行解码(70),且对接收到的视野的图块进行上采样及合并(72)。举例来说,目的地装置14可对以较低分辨率发送的图块中任一者上采样,以使得此类图块可与高分辨率图块组合以形成单一图像覆盖。可以每个刷新周期(例如每隔一秒)更新图块组织(即,待显示的FoV)。然而,可使用不同刷新速率。举例来说,刷新频率可取决于在发射链路上的处理功率和服务功率(QoS)。举例来说,刷新速率可从每帧一秒、零点五秒到100毫秒变化。

[0057] 在上采样及合并视野的图块(72)期间,目的地装置14通过拦截传感器数据和传送与虚拟现实(VR)头戴式显示器(HMD)装置78相关联的传感器框架(76)来确定头部位置和移动数据(74)。通过这种方式,在视野的图块的上采样及合并(72)期间及在视野显现(68)期间利用基于最近头部运动的预测性适应。在一个实例中,覆盖在基线表面上的图像(即,图形域纹理)的自适应是基于最近的头部运动。如果头部运动将导致显现图块以基于最近头部运动离开视野,那么目的地装置14并未覆盖这些图块。自适应可包含基于头部运动轨道推断经译码运动信息。在另一实例中,可针对可视性窗口测试头部运动范围(例如,度/秒),以在用户移动头部时确保图像的无缝过渡。

[0058] 除了在上采样(72)及显现(68)期间利用以外,经由天线60无线发射头部位置及运动传感器数据(框74)以供编码器12使用,如上文所描述。根据呈现时戳复合及显现(80)经上采样及合并(框72)的视野的图块和所显现视野(框68),以使用低分辨率视频图形覆盖高分辨率视频图形,且接着通过虚拟现实(VR)头戴式显示器(HMD)装置78经由显示器27输出到用户,以及经由扬声器(图中未展示)将适应于根据来自框64的剖析音频流确定的最近头部运动的环绕音频(框82)输出到用户。

[0059] 图4为根据本发明的实例的用于处理视频数据的方法的流程图。如图4中所说明,根据一个实例,源装置12接收视频数据(100),且源装置12的一或多个处理器15及16经配置以在像素域中编码视频数据来产生像素域视频数据(102)。此外,源装置12的图形处理单元19经配置以在图形域中处理视频数据以产生图形域视频数据(104)。图形域视频数据及像素域视频数据通过用于流式发射视频数据的接口18发射(106)。

[0060] 在编码以产生像素数据(102)期间,一或多个处理器15及16可将360视频数据拼接在一起以形成等矩形画布以产生图形域视频数据(104),图形处理单元19可将画布转换为纹理,以用于在球体内部显现纹理。在一个实例中,可使用除了等矩形投影之外的投影。举例来说,可在利用适当变换的整个处理管线期间使用立方体投影或金字塔形投影。

[0061] 可以第一帧率经由接口18发射纹理(106)且以大于第一帧率的第二帧率经由接口18发射像素域视频数据(106)。根据一个实例,可以第一分辨率经由接口18发射纹理(106)且可以大于第一分辨率的第二分辨率经由接口发射像素域视频数据(106)。根据一个实例,在编码以产生像素数据(102)期间,一或多个处理器15及16可将画布映射到立方图或金字塔形投影中的一者,在多种分辨率下将流数据的编码成多个图块,及发射(106)处于用户视野内的多个图块中的一或多个图块。根据一个实例,一或多个处理器15及16经配置以确定用户的运动适应性并基于用户的经确定运动适应性来确定视野。根据一个实例,一或多个

处理器15及16经配置以确定用户的运动适应性,基于用户的经确定运动适应性确定视野且基于所确定的视野调适环绕声视频。根据一个实例,一或多个处理器15及16经配置以在第一分辨率下基于运动适应性发射视野中心处的图块且在低于第一分辨率的第二分辨率下发射在视野内但不在视野中心的图块。

[0062] 图5为根据本发明的实例的处理视频数据的方法的流程图。如图5中所说明,根据一个实例,目的地装置14从视频解码器20接收流式视频且将所述流式视频存储在存储器29中(108),一或多个处理器20及21将接收到的视频流剖析成图形流及音频-视频流并对经剖析视频流进行解码(110)。目的地装置14的一或多个传感器24感测用户的运动适应性(112),且图形处理单元19利用来自图形流的纹理信息在球面表面上生成画布(114),并基于用户的经感测运动适应性来显现视野(116)。

[0063] 根据一个实例,一或多个处理器20及21可对处于用户视野内的多个图块进行解码,对多个图块中是低分辨率图块的图块和多个图块中是高分辨率图块的图块进行上采样,并组合低分辨率图块和高分辨率图块以形成单一图像覆盖。根据一个实例,单一图像覆盖可基于经由一或多个传感器24感测的用户的所感测运动适应性而形成。根据一个实例,一或多个处理器20及21可确定用户的经感测运动适应性是否导致经上采样的图块在视野外,且不会基于确定用户的经感测运动适应性导致经上采样的图块在视野外而形成单一图像覆盖。

[0064] 根据一个实例,用户的经感测运动适应性(112)可包含头部运动轨道。根据一个实例,可将用户的头部运动范围与可视性窗口进行比较,且可作出关于用户的经感测运动是否导致经上采样的图块在视野外的判定,且若如此,视频解码器则不会形成单一图像覆盖。

[0065] 通过这种方式,根据本发明的一个实例,视频流系统10利用上述技术来流式发射360度视频。流视频系统10利用混合图形及像素域方法,其中由图形域管线提供基线质量及立即反应,且由像素域管线提供高分辨率纹理。另外,流视频系统10还解决头部运动适应性及定向音频及视频同步。因此,本发明充分利用实时应用程序的像素及图形域流的强度,并结合可视性限制的窗口考虑人类视野特性。

[0066] 根据一个实例,本发明提出:译码图形域中的360视频数据以创建图形域视频数据;在像素域中编码360视频数据以创建像素域视频数据;及发射图形域及像素域视频数据。

[0067] 在另一实例中,本发明进一步揭示:提取传感器数据;处理所提取的传感器数据以确定用户的运动适应性;及基于所提取的传感器发射像素域视频数据的图块的子集。在另一实例中,本发明进一步揭示:将图形域中的360视频数据转换成纹理以供由图形处理器执行,以用于在球体内部显现360视频数据,施加用于同步的时戳,及实时协议囊封视频数据以供传送。在另一实例中,本发明进一步揭示:将360视频数据映射到立方图和金字塔形投影中的一者;在各种分辨率下将360视频数据编码成多个图块;及基于用户的经确定运动适应性流式发射多个图块的子集。在另一实例中,本发明进一步揭示:使环绕音频适应用户的经确定运动适应性且使所确定的运动与图形背景同步;及在施加用于同步的时戳及实时协议囊封视频数据以供传送期间利用多个图块中的流式预定图块及经调适的环绕音频。

[0068] 在另一实例中,本发明进一步揭示:在图形域中译码360视频数据以创建图形域视频数据;在像素域中译码360视频数据以创建像素域视频数据;发射图形域及像素域视频数

据;接收所发射的图形域及像素域视频数据并将其剖析成图形流及音频-视频流;在球面表面上生成画布;感测用户的运动;及基于生成的画布及感测到的运动显现视野。

[0069] 在另一实例中,本发明揭示一种流式360视频系统,其包括源装置,所述源装置包括源处理器,其经配置以在图形域中译码360视频数据以创建图形域视频数据,在像素域中译码360视频数据以创建像素域视频数据,且发射图形域及像素域视频数据;及目的地装置,其用以接收所发射的图形域及像素域视频数据且包括目的地处理器,所述目的地处理器经配置以将所发射的图形域及像素域视频数据剖析成图形流及音频-视频流,在球面表面上产生画布,感测用户的运动,且基于所产生的画布及所感测的运动显现视野。

[0070] 在一或多个实例中,所描述功能可以硬件、软件、固件或其任何组合来实施。如果以软件实施,则所述功能可以作为一或多个指令或代码在计算机可读媒体上存储或传输,并且由基于硬件的处理单元执行。计算机可读媒体可包含计算机可读存储媒体,其对应于有形媒体,例如数据存储媒体,或通信媒体,其包含有助于例如根据通信协议将计算机程序从一处传送到另一处的任何媒体。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)通信媒体,例如,信号或载波。数据存储媒体可以是可由一或多个计算机或一或多个处理器存取以检索用于实施本发明中描述的技术的指令、代码和/或数据结构的任何可用媒体。计算机程序产品可包含计算机可读媒体。

[0071] 借助于实例而非限制,此类计算机可读存储媒体可包括RAM、ROM、EEPROM、CD-ROM或其它光盘存储装置、磁盘存储装置或其它磁性存储装置、快闪存储器,或可用以存储呈指令或数据结构形式的所要程序代码且可由计算机存取的任何其它媒体。并且,适当地将任何连接称作计算机可读媒体。举例来说,如果使用同轴电缆、光纤缆线、双绞线、数字订户线(DSL)或例如红外线、无线电及微波等无线技术从网站、服务器或其它远程源发射指令,那么同轴电缆、光纤缆线、双绞线、DSL或例如红外线、无线电及微波等无线技术包含在媒体的定义中。然而,应理解,计算机可读存储媒体和数据存储媒体并不包含连接、载波、信号或其它暂时性媒体,而实际上是针对非暂时性有形存储媒体。如本文中所使用,磁盘和光盘包含压缩光盘(CD)、激光光盘、光学光盘、数字多功能光盘(DVD)、软盘及蓝光光盘,其中磁盘通常以磁性方式显现数据,而光盘用激光以光学方式显现数据。以上各项的组合也应包含于计算机可读媒体的范围内。

[0072] 指令可以由一或多个处理器执行,所述一或多个处理器例如是一或多个数字信号处理器(DSP)、通用微处理器、专用集成电路(ASIC)、现场可编程逻辑阵列(FPGA)或其它等效的集成或离散逻辑电路。因此,如本文中所使用的术语“处理器”可指代前述结构或适用于实施本文中所描述的技术的任何其它结构中的任一者。另外,在一些方面中,本文中所描述的功能性可在经配置以用于编码和解码的专用硬件和/或软件模块内提供,或并入在组合编解码器中。并且,所述技术可完全实施于一或多个电路或逻辑元件中。

[0073] 本发明的技术可实施于广泛多种装置或设备中,包含无线手持机、集成电路(IC)或一组IC(例如芯片组)。本发明中描述各种组件、模块或单元是为了强调经配置以执行所揭示的技术的装置的功能方面,但未必需要通过不同硬件单元实现。确切地,如上文所描述,各种单元可结合合适的软件和/或固件组合在编解码器硬件单元中,或由互操作硬件单元的集合来提供,所述硬件单元包含如上文所描述的一或多个处理器。

[0074] 已描述了各种实例。这些及其它实例在以下权利要求书的范围内。

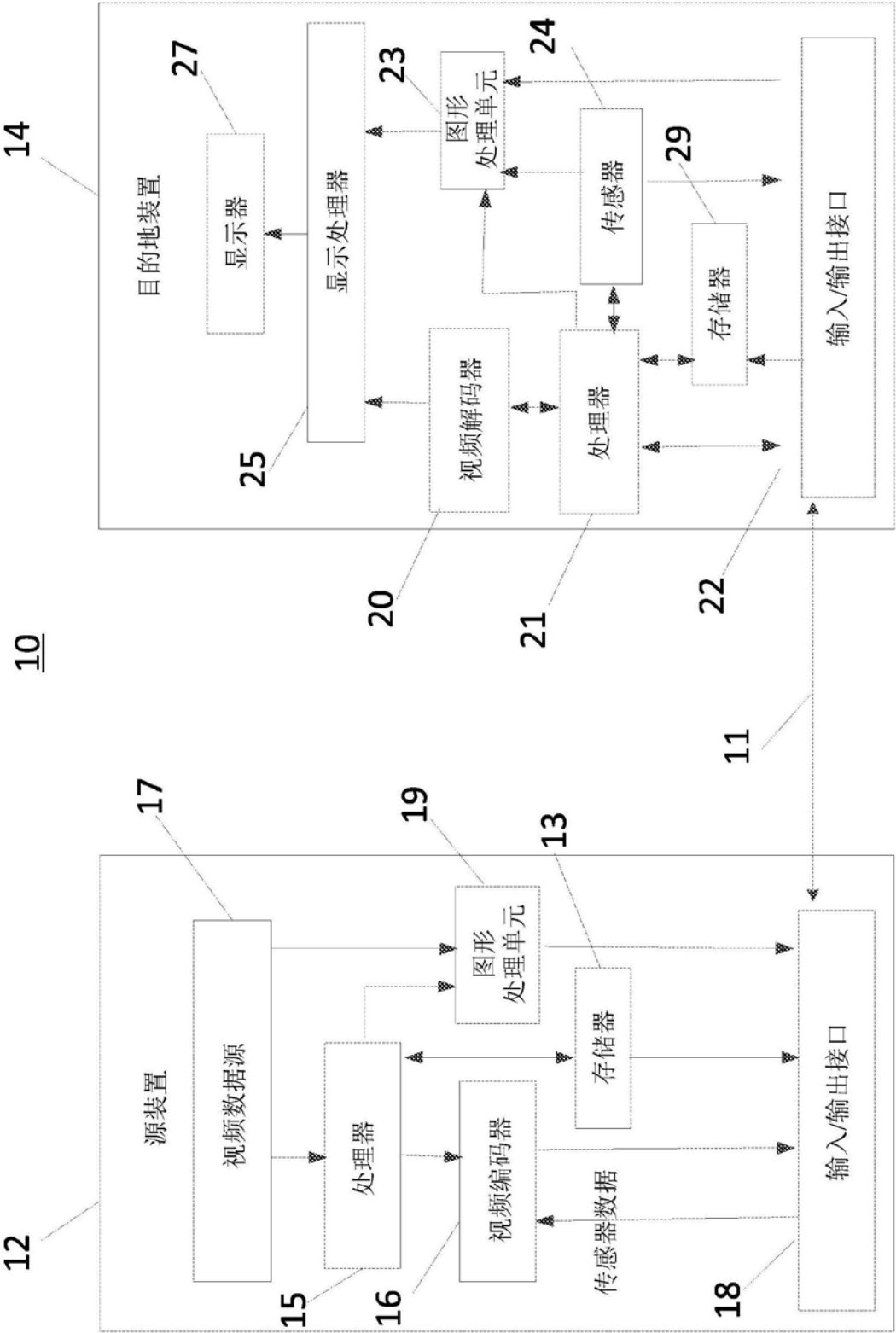


图1

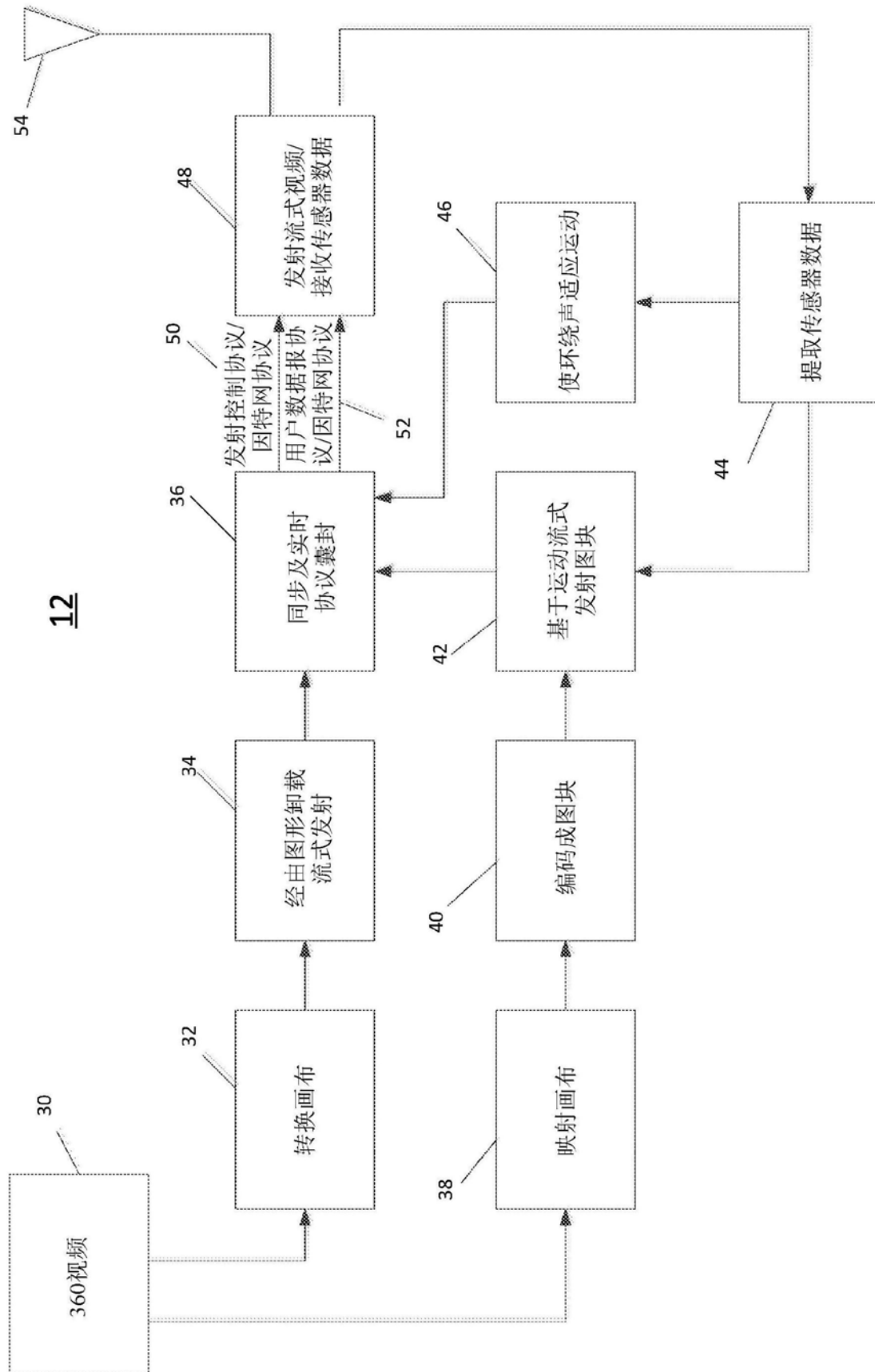


图2

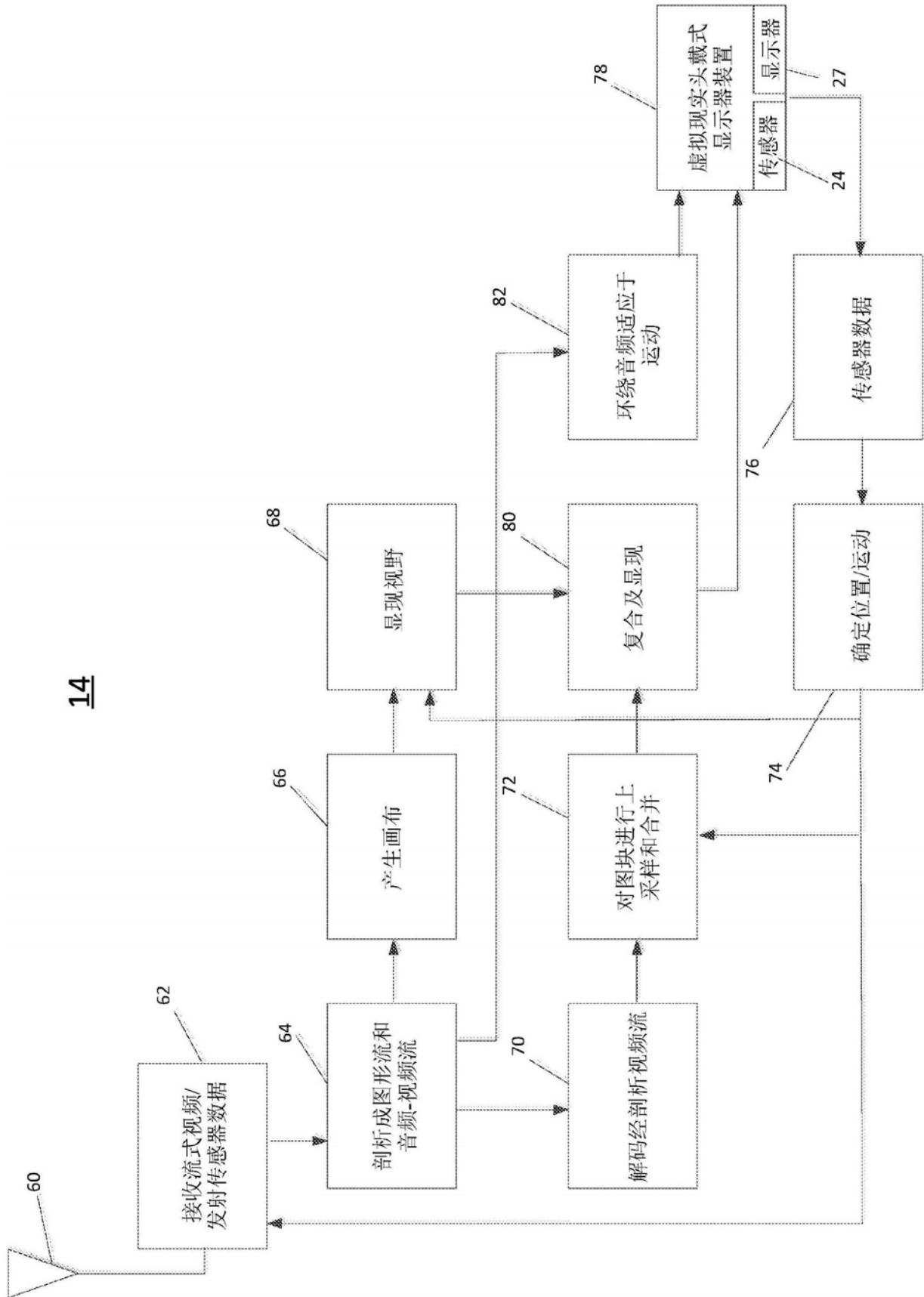


图3

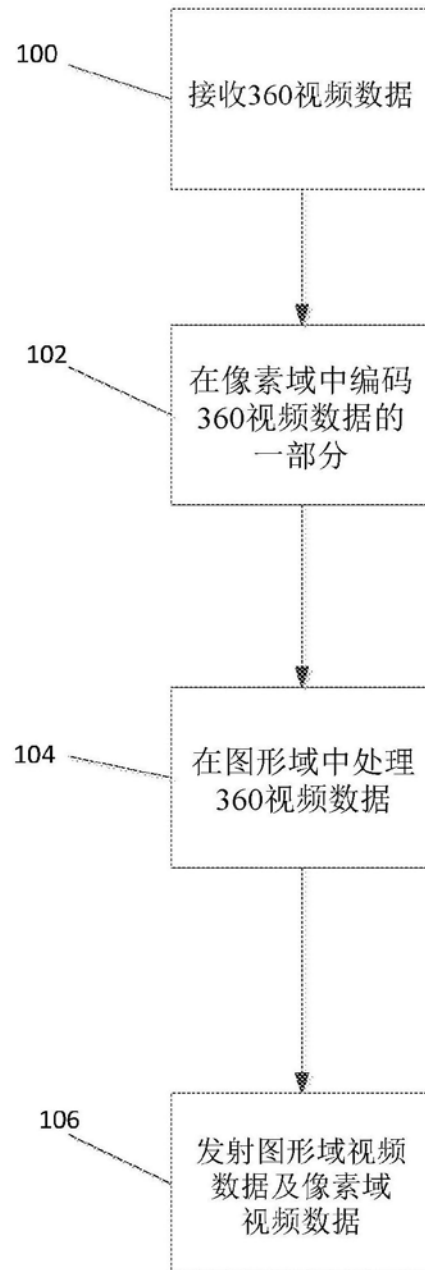


图4

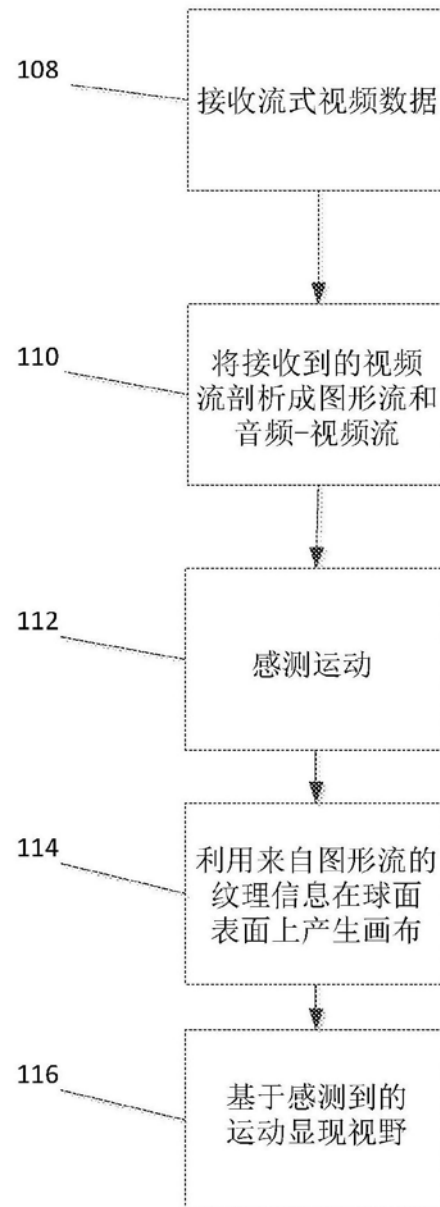


图5