



(12)发明专利申请

(10)申请公布号 CN 110121885 A

(43)申请公布日 2019.08.13

(21)申请号 201780081557.4

(74)专利代理机构 北京市柳沈律师事务所
11105

(22)申请日 2017.12.13

代理人 张晓明

(30)优先权数据

62/440,198 2016.12.29 US

(51)Int.Cl.

H04N 19/162(2006.01)

(85)PCT国际申请进入国家阶段日

H04N 19/167(2006.01)

2019.06.28

H04N 19/17(2006.01)

(86)PCT国际申请的申请数据

PCT/US2017/066205 2017.12.13

(87)PCT国际申请的公布数据

W02018/125579 EN 2018.07.05

(71)申请人 索尼互动娱乐股份有限公司

地址 日本东京都

(72)发明人 H-J.李 J.王 R.克里什曼

D.艾利亚 E.陈

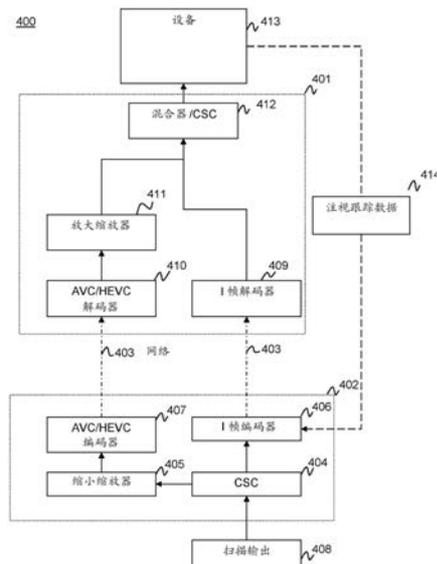
权利要求书3页 说明书17页 附图14页

(54)发明名称

用于利用注视跟踪的VR、低等待时间无线HMD视频流传输的有凹视频链接

(57)摘要

分析注视跟踪数据以确定视频流的图像内的一个或多个感兴趣区域。选择性地缩放视频流数据,使得感兴趣区域内的区段保持高分辨率,同时缩小缩放不在所述感兴趣区域内的区域,以减小传输的带宽成本。还要求保护通过减小高分辨率区域的大小来减少晕动病的方案。



1. 一种方法,其包括:
 - a) 获得表示视频帧的一个或多个感兴趣区域的注视跟踪信息;
 - b) 用低压缩编码器压缩对应于所述一个或多个感兴趣区域的所述视频帧的一个或多个区段以产生低压缩区段数据;
 - c) 缩小缩放所述视频帧并用高压缩编码器压缩得到的缩小缩放的视频帧以产生高压缩帧数据;
 - d) 通过网络传输所述高压缩帧数据和所述低压缩区段数据。
2. 根据权利要求1所述的方法,其中b)还包括将掩码应用于对应于所述一个或多个感兴趣区域的所述一个或多个区段的边缘以产生分辨率劣化的梯度。
3. 根据权利要求2所述的方法,其中应用所述掩码产生分辨率劣化的线性梯度。
4. 根据权利要求2所述的方法,其中应用所述掩码产生分辨率劣化的非线性梯度。
5. 根据权利要求1所述的方法,其中压缩对应于所述一个或多个感兴趣区域的所述视频帧的一个或多个子区段包括响应于晕动病相关信息而调整所述一个或多个感兴趣区域的大小。
6. 根据权利要求5所述的方法,其中所述晕动病相关反应是视觉移动阈值。
7. 根据权利要求5所述的方法,其中所述晕动病相关反应由一个或多个传感器检测。
8. 根据权利要求1所述的方法,其中通过网络传输所述高压缩帧数据和所述低压缩区段数据还包括传输所述一个或多个低压缩子区段的在所述视频流内的位置。
9. 根据权利要求8所述的方法,其中所述一个或多个低压缩子区段的在所述视频流内的所述位置被放置在所述低压缩子区段的帧报头中。
10. 一种系统,其包括:

处理器;

耦接到所述处理器包含处理器可执行指令的存储器,所述指令在被执行时致使所述处理器实现一种方法,所述方法包括:
 - a) 获得表示视频帧的一个或多个感兴趣区域的注视跟踪信息;
 - b) 用低压缩编码器压缩对应于所述一个或多个感兴趣区域的所述视频帧的一个或多个区段以产生低压缩区段数据;
 - c) 缩小缩放所述视频帧并用高压缩编码器压缩得到的缩小缩放的视频帧以产生高压缩帧数据;
 - d) 通过网络传输所述高压缩帧数据和所述低压缩区段数据。
11. 根据权利要求10所述的系统,其中b)还包括将掩码应用于对应于所述一个或多个感兴趣区域的所述一个或多个区段的边缘以产生分辨率劣化的梯度。
12. 根据权利要求11所述的系统,其中应用所述掩码产生分辨率劣化的线性梯度。
13. 根据权利要求11所述的系统,其中应用所述掩码产生分辨率劣化的非线性梯度。
14. 根据权利要求10所述的系统,其中压缩对应于所述一个或多个感兴趣区域的所述视频帧的一个或多个子区段包括响应于晕动病相关信息而调整所述一个或多个感兴趣区域的大小。
15. 根据权利要求14所述的系统,其中所述晕动病相关反应是视觉移动阈值。
16. 根据权利要求14所述的系统,其中所述晕动病相关反应由一个或多个传感器检测。

17. 根据权利要求10所述的系统,其中通过网络传输所述高压压缩帧数据和所述低压缩区段数据还包括传输所述一个或多个低压缩子区段的在所述视频流内的位置。

18. 根据权利要求17所述的方法,其中所述一个或多个低压缩子区段的在所述视频流内的所述位置被放置在所述低压缩子区段的帧报头中。

19. 一种具有计算机可执行指令的非暂时性计算机可读介质,所述指令在被执行时致使计算机实现一种方法,所述方法包括:

- a) 获得表示视频帧的一个或多个感兴趣区域的注视跟踪信息;
- b) 用低压缩编码器压缩对应于所述一个或多个感兴趣区域的所述视频帧的一个或多个区段以产生低压缩区段数据;
- c) 缩小缩放所述视频帧并用高压压缩编码器压缩得到的缩小缩放的视频帧以产生高压压缩帧数据;
- d) 通过网络传输所述高压压缩帧数据和所述低压缩区段数据。

20. 一种方法,其包括:

- a) 将对应于缩小缩放的视频帧的高压缩帧数据解压缩以生成缩小缩放的视频帧数据;
- b) 将所述缩小缩放的视频帧数据放大缩放以生成放大缩放的视频帧数据;
- c) 将对应于一个或多个感兴趣区域的所述视频帧的一个或多个区段的低压缩数据解压缩以生成感兴趣区域数据;
- d) 将所述放大缩放的视频帧数据与所述感兴趣区域数据组合以生成组合帧;
- e) 在显示器上显示所述组合帧。

21. 根据20所述的方法,其中d) 包括确定在所述组合帧中放置对应于所述一个或多个感兴趣区域的所述视频帧的所述一个或多个区段的位置。

22. 根据20所述的方法,其中e) 包括使用注视跟踪信息来确定在所述组合帧中放置对应于所述一个或多个感兴趣区域的所述视频帧的所述一个或多个区段的位置。

23. 根据权利要求20所述的方法,其中e) 包括在所述一个或多个感兴趣区域的边缘处将掩码应用于对应于所述一个或多个感兴趣区域的所述视频帧的所述一个或多个区段,以在所述组合帧的对应于所述放大缩放的视频帧和对应于所述一个或多个感兴趣区域的所述视频帧的所述一个或多个区段的部分之间生成分辨率劣化梯度。

24. 根据权利要求23所述的方法,其还包括将第二掩码应用于所述放大缩放的视频帧数据以对所述放大缩放的视频帧调光。

25. 根据权利要求23所述的方法,其中所述掩码生成线性劣化梯度。

26. 根据权利要求23所述的方法,其中所述掩码生成非线性劣化梯度。

27. 根据权利要求20所述的方法,其还包括基于晕动病相关反应来调整所述一个或多个感兴趣区域的大小。

28. 根据权利要求27所述的方法,其中所述晕动病相关反应是皮电反应。

29. 根据权利要求28所述的方法,其中所述晕动病相关反应是所述视频帧内的运动向量的用户确定阈值。

30. 一种系统,其包括:

处理器;

耦接到所述处理器包含处理器可执行指令的存储器,所述指令在被执行时致使所述处

理器实现一种方法,所述方法包括:

- a) 将对应于缩小缩放的视频帧的高压缩帧数据解压缩以生成缩小缩放的视频帧数据;
- b) 将所述缩小缩放的视频帧数据放大缩放以生成放大缩放的视频帧数据;
- c) 将对应于一个或多个感兴趣区域的所述视频帧的一个或多个区段的低压缩数据解压缩以生成感兴趣区域数据;
- d) 将所述放大缩放的视频帧数据与所述感兴趣区域数据组合以生成组合帧;
- e) 在显示器上显示所述组合帧。

31. 根据权利要求30所述的系统,其中d)包括确定在所述组合帧中放置对应于所述一个或多个感兴趣区域的所述视频帧的所述一个或多个区段的位置。

32. 根据权利要求30所述的系统,其中e)包括使用注视跟踪信息来确定在所述组合帧中放置对应于所述一个或多个感兴趣区域的所述视频帧的所述一个或多个区段的位置。

33. 根据权利要求30所述的系统,其中e)包括在所述一个或多个感兴趣区域的边缘处将掩码应用于对应于所述一个或多个感兴趣区域的所述视频帧的所述一个或多个区段,以在所述组合帧的对应于所述放大缩放的视频帧和对应于所述一个或多个感兴趣区域的所述视频帧的所述一个或多个区段的部分之间生成分辨率劣化梯度。

34. 根据权利要求33所述的系统,其还包括将第二掩码应用于所述放大缩放的视频帧数据以对所述放大缩放的视频帧调光。

35. 根据权利要求33所述的系统,其中所述掩码生成线性劣化梯度。

36. 根据权利要求33所述的系统,其中所述掩码生成非线性劣化梯度。

37. 根据权利要求30所述的系统,其还包括基于晕动病相关反应来调整所述一个或多个感兴趣区域的大小。

38. 根据权利要求37所述的系统,其中所述晕动病相关反应是皮电反应。

39. 根据权利要求38所述的系统,其中所述晕动病相关反应是所述视频帧内的运动向量的用户确定阈值。

40. 一种具有计算机可执行指令的非暂时性计算机可读介质,所述指令在被执行时致使计算机实现一种方法,所述方法包括:

- a) 将对应于缩小缩放的视频帧的高压缩帧数据解压缩以生成缩小缩放的视频帧数据;
- b) 将所述缩小缩放的视频帧数据放大缩放以生成放大缩放的视频帧数据;
- c) 将对应于一个或多个感兴趣区域的所述视频帧的一个或多个区段的低压缩数据解压缩以生成感兴趣区域数据;
- d) 将所述放大缩放的视频帧数据与所述感兴趣区域数据组合以生成组合帧;
- e) 在显示器上显示所述组合帧。

用于利用注视跟踪的VR、低等待时间无线HMD视频流传输的有凹视频链接

[0001] 优先权要求

[0002] 本申请要求2016年12月29日提交的美国临时专利申请号62/440,198的优先权,该申请的全部内容通过引用并入本文。

技术领域

[0003] 本公开的各方面涉及视频压缩。特别地,本公开涉及视频图像内的差分编码质量。

背景技术

[0004] 已经开发了具有宽视野(FOV)的图形显示设备。此类设备包括头戴式显示器(HMD)设备。在HMD设备中,小型显示设备佩戴在用户的头部上。显示设备具有位于一只眼睛(单眼HMD)或每只眼睛(双目HMD)前面的显示光学器件。HMD设备通常包括可感测设备定向并在用户头部移动时改变显示光学器件所示出的场景的传感器。传统上,在宽FOV显示器上显示的大多数视频是使用单个压缩算法压缩的,其中屏幕的所有部分具有相同的分辨率。

[0005] 然而,显示用于虚拟现实(VR)程序(经常与HMD设备结合执行)的视频需要比常规平板显示器更高的帧速率以防止用户经历晕动病。在带宽方面,以非常高的分辨率和高帧速率显示视频是非常昂贵的。用于VR的HMD具有用于在宽FOV中显示视频以获得身临其境的体验的光学系统。虽然主注视点周围的屏幕区域(有时称为有凹区域)需要高分辨率,但主注视点之外的区域仅通过周边视觉观察并因此可以以较低分辨率显示。这种视频有时被称为有凹视频。

[0006] 在此背景下出现了本公开。

附图说明

[0007] 图1A-1B是示出在本公开的方面的上下文内的注视跟踪的示意图。

[0008] 图2A-2B是示出有凹视频的示例的图示。

[0009] 图2C是示出在本公开的方面的上下文内的视频帧的一种可能划分的示意图。

[0010] 图3是描绘根据本公开的方面的系统的框图。

[0011] 图4是根据本公开的方面的总体系统的流程图。

[0012] 图5A-5B是描绘根据本公开的方面的编码器组侧的流程图。

[0013] 图6A-6B是描绘根据本公开的方面的解码器组侧的流程图。

[0014] 图7是示出了根据本公开的方面的感兴趣区域的示例的屏幕空间的示意图

[0015] 图8是描绘根据本公开的方面的晕动病减少方案的流程图。

[0016] 图9A-9H是示出结合本发明的实施方案使用眼睛注视和面部跟踪的示例的示意图。

[0017] 图10A-10D是示出根据本公开的方面的面部定向特性跟踪设置的示意图。

[0018] 图10E是示出根据本公开的方面的可利用面部定向跟踪的便携式设备的示意图。

发明内容

[0019] 与现有技术相关联的缺点被涉及有凹视频的编码和传输的本公开的各方面克服。

[0020] 一种方法,包括:接收视频图像;获得表示一个或多个感兴趣区域的注视跟踪信息;用低压缩编码器压缩对应于一个或多个感兴趣区域的视频图像的一个或多个子区段以产生低压缩视频数据;缩小缩放视频图像并用高压缩编码器压缩视频图像以生成高压缩视频数据;传输高压缩视频数据和一个或多个低压缩视频数据。

[0021] 一种方法,包括:接收高压缩视频数据;接收对应于表示从注视跟踪信息获得的一个或多个感兴趣区域的视频图像的一个或多个低压缩子区段的低压缩视频数据;解压缩高压缩视频图像;放大缩放解压缩的视频图像;解压缩视频图像的一个或多个低压缩子区段;将放大缩放的视频图像与感兴趣区域中的视频图像的一个或多个子区段组合;以及显示组合的放大缩放视频图像和一个或多个子区段。

具体实施方式

[0022] 尽管为了说明的目的,以下详细描述包含许多具体细节,但本领域普通技术人员将理解,对以下细节的许多变化和改变都在本发明的范围内。因此,阐述了下面描述的本公开的说明性实现方式而不失一般性,并且不对所要求保护的发明施加限制。

[0023] 引言

[0024] 眼睛注视跟踪已经用于广泛的应用,包括医学研究、汽车技术、计算机娱乐和视频游戏程序、控制输入设备、增强现实眼镜等。有许多种用于眼睛跟踪(也称为注视跟踪)的技术。用于眼睛注视跟踪和选择性渲染压缩的技术在共同未决申请2015/087,471中描述,该申请的内容通过引用并入本文。这些技术中的一些技术根据用户眼睛的瞳孔的定向来确定用户的注视方向。一些已知的眼睛注视跟踪技术涉及通过从一个或多个光源发射光来照射眼睛以及用传感器检测发射光从角膜的反射。通常,这是使用红外范围内的不可见光源以及通过红外敏感相机捕获被照射眼睛的图像数据(例如,图像或视频)来实现的。然后使用图像处理算法来分析图像数据以确定眼睛注视方向。

[0025] 通常,眼睛跟踪图像分析利用与光从眼睛反射的方式不同的特性,以根据图像来确定眼睛注视方向。例如,可以分析图像以基于图像数据中的角膜反射来识别眼睛位置,并且可以进一步分析图像以基于瞳孔在图像中的相对位置来确定注视方向。

[0026] 用于基于瞳孔位置确定眼睛注视方向的两种常见注视跟踪技术被称为亮瞳孔跟踪和暗瞳孔跟踪。亮瞳孔跟踪涉及用基本上与相机光轴一致的光源照射眼睛,从而致使发射光从视网膜反射并通过瞳孔返回相机。瞳孔在图像中呈现为瞳孔位置处的可识别亮点,类似于在常规闪光摄影期间在图像中出现的红眼效应。在这种注视跟踪方法中,如果瞳孔和虹膜之间的对比不够,则从瞳孔本身的明亮反射有助于系统定位瞳孔。

[0027] 暗瞳孔跟踪涉及用与相机光轴基本上离线的光源进行照射,从而致使通过瞳孔引导的光被反射远离相机光轴,进而在图像中的瞳孔位置处产生的可识别暗点。在替代性暗瞳孔跟踪系统中,针对眼睛的红外光源和相机可以观察角膜反射。此类基于相机的系统跟踪瞳孔的位置,并且由于不同反射深度而提供视差的角膜反射提供附加的准确性。

[0028] 图1A描绘了可以在本公开的上下文中使用的暗瞳孔注视跟踪系统100的示例。注视跟踪系统跟踪用户眼睛E相对于上面呈现可见图像的显示屏101的定向。虽然在图1A的示

例性系统中使用显示屏,但某些替代性实施方案可以利用能够将图像直接投射到用户眼睛中的图像投射系统。在这些实施方案中,将相对于投射到用户眼睛中的图像来跟踪用户的眼睛E。在图1A的示例中,眼睛E通过可变虹膜I从屏幕101收集光,并且晶状体L在视网膜R上投射图像。虹膜中的开口被称为瞳孔。肌肉响应于来自大脑的神经脉冲而控制眼睛E的旋转。上眼睑肌肉ULM和下眼睑肌肉LLM分别响应于其他神经脉冲而控制上眼睑UL和下眼睑LL。

[0029] 视网膜R上的光敏细胞产生电脉冲,其通过视神经ON来发送到用户大脑(未示出)。大脑的视觉皮层解释脉冲。并非视网膜R的所有部分对光同等敏感。具体而言,光敏细胞集中在称为中央凹的区域。

[0030] 所示的图像跟踪系统包括一个或多个红外光源102,例如,将不可见光(例如,红外光)引向眼睛E的发光二极管(LED)。不可见光的部分从眼睛的角膜C反射并且部分从虹膜反射。通过波长选择镜106将反射的不可见光导向合适的传感器104(例如,红外相机)。镜透射来自屏幕101的可见光,但反射从眼睛反射的不可见光。

[0031] 传感器104优选是图像传感器,例如,可以产生眼睛E的图像的数码相机,该图像可以被分析以确定瞳孔的相对位置来确定注视方向GD。该图像可以利用本地处理器120或经由将获得的注视跟踪数据传输到远程计算设备160来产生。可以根据众所周知的架构来配置本地处理器120,例如,单核、双核、四核、多核、处理器-协处理器、cell处理器等。图像跟踪数据可以经由有线连接(未示出)在传感器104与远程计算设备160之间传输,或者无线地在眼睛跟踪设备110中包括的无线收发器125和远程计算设备160中包括的第二无线收发器126之间传输。无线收发器可以被配置成经由合适网络协议来实现局域网(LAN)或个域网(PAN),例如,蓝牙,用于PAN。

[0032] 注视跟踪系统100还可以包括上传感器108和下传感器109,其被配置成例如分别放置在眼睛E的上方和下方。传感器108和109可以是独立部件,或者可以替代地是佩戴在用户头部上的部件110的部分,其可以包括但不限于下面描述的传感器104、本地处理器120或惯性传感器115的任何组合。在图1A所示的示例性系统中,传感器108和109能够从眼睛E周围的那些区域收集关于神经系统的电脉冲和/或肌肉系统的移动和/或振动的数据。该数据可以包括例如由上传感器108和下传感器109监测的眼睛E周围的肌肉和/或神经的电生理和/或振动信息。由传感器108和109收集的电生理信息可以包括例如脑电图(EEG)、肌电图(EMG)或由于眼睛E周围的区域中的神经功能而收集的诱发电位信息。由于检测到眼睛E周围肌肉的肌肉振动或抽搐,传感器108和109还可以能够收集例如肌电图或表面肌电图信息。传感器108还可以能够收集与晕动病反应有关的信息,包括例如心率数据、心电图(ECG)或皮电反应数据。如上所述,由传感器108和109收集的数据可以与图像跟踪数据一起递送到本地处理器120和/或远程计算设备160。

[0033] 注视跟踪系统100还可以能够跟踪用户的头部。头部跟踪可以由惯性传感器115执行,该惯性传感器能够响应于用户头部的位置、运动、定向或定向改变而产生信号。该数据可以被发送到本地处理器120和/或传输到远程计算设备160。惯性传感器115可以是独立的部件,或者可替代地是佩戴在用户头部上的部件110的一部分,其可以包括但不限于上述传感器104、本地处理器120或传感器108和109的任何组合。在替代实施方案中,可以经由跟踪部件110上的光源来执行头部跟踪。注视跟踪系统100还可以包括一个或多个存储器单元

177, (例如,随机存取存储器 (RAM)、动态随机存取存储器 (DRAM)、只读存储器 (ROM) 等。

[0034] 本地处理器120可以被配置成从网络连接125接收编码数据。本地处理器120可以操作地耦接到一个或多个存储器单元177,并且被配置成执行存储在存储器单元177上的一个或多个程序。这些程序的执行可以致使系统对来自远程计算设备160的视频流进行解码,并且生成用于在显示器101上显示的有凹视频。作为示例,并且不作为限制,程序可以包括混合器/转换空间配置程序179、缩放器程序176和解码器程序。180。

[0035] 远程计算设备160可以被配置成与眼睛跟踪设备110和显示屏101协同操作,以便根据本公开的方面执行眼睛注视跟踪并确定照明状况。计算设备160可以包括可根据众所周知的架构来配置的一个或多个处理器单元170,例如,单核、双核、四核、多核、处理器-协处理器、cell处理器等。计算设备160还可以包括一个或多个存储器单元172, (例如,随机存取存储器 (RAM)、动态随机存取存储器 (DRAM)、只读存储器 (ROM) 等)。

[0036] 处理器单元170可以执行一个或多个程序,其部分可以存储在存储器172中,并且处理器170可以操作地耦接到存储器172,例如通过经由数据总线178访问存储器。程序可以被配置成执行眼睛注视跟踪并确定系统100的照明状况。作为示例,并且不作为限制,程序可以包括:注视跟踪程序173,其执行可以致使系统100跟踪用户的注视,例如,如上所述;转换空间配置程序 (CSC) 174,其将视频帧流转换为可由显示设备呈现的形式;编码器程序175;和视频流缩放器程序176,其执行对低分辨率视频帧流和视频帧的所选高分辨率区段进行编码以发送到显示器。编码器程序175可以使用观看者晕动病反应来调整要以高分辨率在显示器上呈现的感兴趣区域 (ROI) 的大小,以便例如通过减小ROI的大小来减少在观看者中发生晕动病的机会。

[0037] 作为示例,并且不作为限制,注视跟踪程序173可以包括处理器可执行指令,其致使系统100根据在从光源102发射光的同时用图像传感器104收集的眼睛跟踪数据以及分别从上传感器108和下传感器109收集的眼睛移动数据来确定系统100的一个或多个注视跟踪参数。注视跟踪程序173还可以包括分析利用图像传感器104收集的图像以便检测照明状况的改变的存在的指令。

[0038] 如图1B所示,可以分析示出用户头部H的图像181以根据瞳孔的相对位置确定注视方向GD。例如,图像分析可以确定图像中的瞳孔P从眼睛E的中心的2维偏移。通过基于眼球的已知大小和形状的三维向量的直接几何计算,可以将瞳孔相对于中心的位置转换为相对于屏幕101的注视方向。所确定的注视方向GD能够在眼睛相对于屏幕101移动时显示眼睛E的旋转和加速度。

[0039] 还如图1B所示,图像还可以包括分别来自角膜C和晶状体L的不可见光的反射187和188。由于角膜和晶状体处于不同的深度,因此可以使用反射之间的视差和折射率来提供确定注视方向GD的附加准确度。这种类型的眼睛跟踪系统的示例是双Purkinje跟踪器,其中角膜反射是第一Purkinje图像并且晶状体反射是第4Purkinje图像。如果用户戴眼镜,则也可能存在来自用户眼镜193的反射190。

[0040] 当前的HMD面板取决于制造商以90或120赫兹 (Hz) 的恒定速率刷新。高刷新率增加了面板的功耗和传输介质的带宽要求以发送帧更新。

[0041] 图3描绘了示例性系统300以进一步示出本公开的各个方面。系统300可以包括耦接到眼睛跟踪显示系统301的计算设备360。眼睛跟踪显示设备301包括本地处理器303、本

地存储器317、众所周知的支持电路305、网络接口316、眼睛跟踪设备302和显示设备304以便根据本公开的方面执行眼睛注视跟踪和/或对眼睛跟踪的校准。显示设备304可以呈阴极射线管(CRT)、平板屏幕、触摸屏、或显示文本、数字、图形符号或其他视觉对象的其他设备的形式。可以根据众所周知的架构来配置本地处理器303,例如,单核、双核、四核、多核、处理器-协处理器、cell处理器等。眼睛跟踪显示系统301还可以包括一个或多个存储器单元317,(例如,随机存取存储器(RAM)、动态随机存取存储器(DRAM)、只读存储器(ROM)等)。

[0042] 本地处理器单元303可以执行一个或多个程序,其部分可以存储在存储器317中,并且处理器303可以操作地耦接到存储器317,例如通过经由数据总线318访问存储器。程序可以被配置成为眼睛跟踪显示系统301创建有凹视频。作为示例,并且不作为限制,程序可以包括视频流混合程序和CSC 313、视频流缩放程序314和解码器程序315。作为示例,并且不作为限制,混合器/CSC 313可以包括处理器可执行指令,其致使系统301组合从缩放器程序314接收的视频流和从解码器程序315接收的子区段流以创建用于在显示设备304上显示的有凹视频。缩放器314可以包含当被执行时致使本地处理器放大从解码器315接收的视频流的指令。解码器程序315可以包含在由本地处理器执行时致使系统从网络接口316接收编码的视频流数据并对其进行解码的指令。可交替地,解码器程序可以被实现为通过例如主总线318通信地耦接到本地处理器的离散逻辑单元(未示出)。根据本公开的方面,眼睛跟踪显示设备301可以是嵌入式系统、移动电话、个人计算机、平板计算机、便携式游戏设备、工作站、游戏控制台、头戴式显示设备等。此外,计算设备360还可以是嵌入式系统、移动电话、个人计算机、平板计算机、便携式游戏设备、工作站、游戏控制台等。

[0043] 眼睛跟踪显示设备301可以耦接到计算设备360,并且可以包括类似于图1A-1B的光源110的动态光源310。作为示例,并且不作为限制,光源310可以是以一个或多个红外LED形式的不可见光源,其可以被配置成照亮用户的眼睛以便利用传感器312收集眼睛跟踪数据。眼睛跟踪设备的传感器312可以是对从光源310发射的光敏感的检测器。例如,传感器312可以是对光源敏感的相机(诸如红外相机),并且相机312可以相对于眼睛跟踪设备和光源定位,使得它可以捕获由光源310照射的区域的图像。

[0044] 计算设备360可以被配置成与眼睛跟踪显示系统301协同操作,以便根据本公开的方面执行眼睛注视跟踪并确定照明状况。计算设备360可以包括可根据众所周知的架构来配置的一个或多个处理器单元370,例如,单核、双核、四核、多核、处理器-协处理器、cell处理器等。计算设备360还可以包括一个或多个存储器单元372,(例如,随机存取存储器(RAM)、动态随机存取存储器(DRAM)、只读存储器(ROM)等)。

[0045] 处理器单元370可以执行一个或多个程序,其部分可以存储在存储器372中,并且处理器370可以操作地耦接到存储器372,例如通过经由数据总线376访问存储器。程序可以被配置成执行眼睛注视跟踪并确定系统300的照明状况。作为示例,并且不作为限制,程序可以包括注视跟踪程序373,其执行可以致使系统300跟踪用户的注视,作为示例,并且不作为限制,注视跟踪程序373可以包括处理器可执行指令,其致使系统300根据在从动态光源310发射光的同时用相机312收集的眼睛跟踪数据来确定系统300的一个或多个注视跟踪参数。注视跟踪程序373还可以包括分析用相机312收集的图像的指令,例如,如上面参考图1B所述。可交替地,编码器程序可以被实现为通过例如主总线318通信地耦接到本地处理器的离散逻辑单元(未示出)。

[0046] 在一些实现方式中,注视跟踪程序373可以分析注视跟踪信息以预测用户的视觉感知被遮挡的时段(例如,在眨眼期间),或者不活动的时段(例如,在扫视期间)。预测这些时段的开始可用于减少不必要的渲染计算、功耗和网络带宽使用。在2016年3月31日提交的共同转让的美国专利申请号15/086,953中描述了此类技术的示例,该申请的全部内容通过引用并入本文。

[0047] 计算设备360和眼睛跟踪显示设备301还可以包括众所周知的支持电路378305,诸如输入/输出(I/O)电路379 306、电源(P/S)380 309、时钟(CLK)381 308、和高速缓存382 307,其可以分别例如经由总线376 318与系统的其他部件通信。计算设备360可以包括网络接口390,以促进与眼睛跟踪显示设备301上的类似配置的网络接口316的通信。处理器单元370 303和网络接口390 316可以被配置成经由合适网络协议来实现局域网(LAN)或个域网(PAN),例如,蓝牙,用于PAN。计算设备360可以可选地包括大容量存储设备384,诸如磁盘驱动器、CD-ROM驱动器、磁带驱动器、闪存等,并且大容量存储设备384可以存储程序和/或数据。计算设备360还可以包括用户接口388以促进系统300和用户之间的交互。用户接口388可以包括键盘、鼠标、光笔、游戏控制板、触摸接口或其他设备。

[0048] 系统300还可以包括控制器(未示出),其与眼睛跟踪显示设备301接口连接以便与由处理器单元370执行的程序交互。系统300还可以执行一个或多个通用计算机应用(未示出),诸如视频游戏或视频流,其可以包括由跟踪设备302感测并由跟踪程序373、CSC 376、缩放器374(其将视频帧数据转换为可由显示设备呈现的形式)、以及视频流编码器375处理的眼睛注视跟踪的各方面。

[0049] 计算设备360可以包括网络接口390,其被配置成能够使用Wi-Fi、以太网端口或其他通信方法。网络接口390可以包含合适的硬件、软件、固件或其某种组合,以促进经由电信网络的通信。网络接口390可以被配置成通过局域网和诸如因特网的广域网实现有线或无线通信。网络接口390还可以包括前述无线收发器,其有助于与眼睛跟踪设备302和显示设备379的无线通信。计算设备360可以通过网络经由一个或多个数据分组399发送和接收文件的数据和/或请求。

[0050] 有凹视频

[0051] 作为本公开的方面,有凹视频可以减小高分辨率视频处理的传输带宽要求和计算负荷,同时仍保留由显示器呈现的图像中的感兴趣区域的必要细节。有凹视频通过以下方式减少计算:对所显示图像的聚焦于中央凹的感兴趣区域(ROI)中的高分辨率视频帧子区段执行低损耗压缩或无压缩,并且在该区域外的视频帧上执行缩小和高压缩。为了利用有凹视频,诸如头戴式显示器(HMD)的图像显示设备将使用眼睛注视跟踪技术来确定用户在屏幕上聚焦的位置。

[0052] 有凹视频可以被配置成在由注视跟踪系统确定的观看者的ROI中显示高分辨率图像。图2A和图2B描绘了根据本公开的方面的有凹视频图像201。在图像201中,ROI 203可以是由眼睛注视跟踪确定为直接位于用户的注视202的中心(有时称为中央凹)的前方的区域。在注视中心的20度范围内,观看者的视敏度最高。因此,以高分辨率显示图像201的该部分是有利的。作为本公开的方面,ROI 203可以是被选择为自观看者的注视202的中心20度之内的屏幕的区域。ROI可以是任何形状,包括但不限于两个重叠的圆形区域或两个相邻的正方形。通常,ROI将由彼此相邻的两个区域组成,但其他形状也将起作用,包括但不限于单

个矩形区域或单个卵形区域或两个非相邻的圆形区域。

[0053] 根据本公开的方面,可以对与图像201相对应的视频数据进行编码,使得ROI 203内的图像的区段205以高分辨率显示,并且ROI 203外部的区段204以降低的分辨率显示。对观看者的ROI 203外部的区段204使用降低的分辨率允许在视频流的传输期间减少带宽使用并减少等待时间,同时对流进行编码而不会损失观看者感知的视频质量。在一些实现方式中,在ROI 203的边缘处可存在过渡区域,其中低和高分辨率区段混合以产生高分辨率与低分辨率之间的逐渐过渡。

[0054] ROI 203可以在视频流内逐帧改变位置。观看者对视频图像内的区域的关注可能随着视频的进展而改变。根据本公开的方面,如图2B中所见的,注视跟踪系统可以检测观看者的左眼和右眼的主注视点206、207中的移动,并且作为响应,系统沿注视点移动向量移动ROI 210、211以对应于新注视点。可以响应于主注视点的位置的改变来获取新ROI 210和211内的高分辨率视频流区段208。可替代地,感兴趣的注视点和区域位于视频流的特定区域中的概率可以通过本领域已知的算法来预测,诸如为可在http://www.kyb.mpg.de/fileadmin/user_upload/files/publications/attachments/50_4486%5b0%5d.pdf获得的Kienzle、Wolfe等人的“*How to find interesting locations in video: a spatiotemporal interest point detector learned from human eye movements*”中描述的算法,其内容通过引用并入本文。基于可能的注视点信息,系统可以获取若干高分辨率区域并将它们发送到显示器,其中所显示的最终图像由注视跟踪器所确定的实际注视点确定。应当注意,ROI移动后的高分辨率区域的显示之间的等待时间不那么重要,因为通常快速的眼睛运动模仿扫视事件并且运动使第一ROI和下一个之间的图像区段模糊。

[0055] 在本公开的另一个实施方案中,注视跟踪系统可以预测ROI从视频流的一个区段到视频的另一区段的移动概率。这种预测可以基于观看者的注视点的移动速度向量。系统还可以使用如上所述的可能的注视点信息来进一步细化ROI位置的预测。

[0056] 视频编码

[0057] 根据以下讨论,可以更好地理解与视频编码有关的本公开的某些方面。作为示例,并且不作为限制,如图2C所示,单个数字图片220(例如,数字视频帧)可以被分解为一个或多个区段。如本文所使用,术语“区段”可以指图片220内的一组一个或多个亮度或色度样本。区段的范围可以从图片中的单个亮度或色度样本到整个图片。区段的非限制性示例包括切片(例如,宏块行)222、宏块224、子宏块226、块228和单独像素230。如图2C所示,每个切片222包含一行或多行宏块224或一个或多个这样行的部分。一行中的宏块的数量取决于宏块的大小以及图片220的大小和分辨率。例如,如果每个宏块包含十六乘十六个色度或亮度样本,则可以通过将图片220的宽度(以色度或亮度样本计)除以十六来确定每行中的宏块的数量。每个宏块224可以被分解为多个子宏块226。每个子宏块226可以被分解为多个块228,并且每个块可以包含多个色度或亮度样本230。作为示例,但不限于本发明,在公共视频编码方案中,每个宏块224可以被分解为四个子宏块226。每个子宏块可以分解成四个块228,并且每个块可以包含十六个色度或亮度样本230的四乘四布置。一些编解码器(诸如H.265)允许将给定图片分解成两个或更多个不同大小的区段用于编码。具体地,H.265标准引入了划分图片的“图块”概念。图块是用一些共享报头信息编码的图像的独立可解码区域。图块还可以用于空间随机访问视频图像的局部区域。图片的典型图块配置包括将图片

分割成矩形区域,每个图块中具有大致相等数量的编码单元(CU)。编码单元类似于H.264标准中的宏块(MB)。然而,CU的大小可以由编码器设置,并且可以大于宏块。CU的大小可以灵活并且适应于视频内容,以便最佳地划分图片。

[0058] 应当注意,每个图片可以是帧或场。帧指的是完整图像。场是用于有助于在某些类型的显示设备上显示图像的图像部分。通常,图像中的色度或亮度样本按行布置。为了有助于显示,有时可以通过将交替的像素行放入两个不同场来分割图像。然后可以交织两个场中的色度或亮度样本行以形成完整图像。对于某些显示设备,诸如阴极射线管(CRT)显示器,这两个场可以简单地一个接一个地快速连续显示。用于照射显示器中的像素的磷光体或其他发光元件的余辉与视觉的持久性相结合导致了两个场被感知为连续图像。对于某些显示设备,诸如液晶显示器,可能需要在显示之前将两个场交织成单个图片。表示编码图像的流数据通常包括指示图像是场还是帧的信息。这样的信息可以包括在图像的报头中。

[0059] 现代视频编码器/解码器(编解码器),例如MPEG2、MPEG4和H.264,通常将视频帧编码为三种基本类型之一,这些基本类型被称为内帧、预测帧和双向预测帧,其通常分别称为I帧、P帧和B帧。

[0060] I帧是不参考除自身之外的任何图片而编码的图片。I帧用于随机访问,并不用作其他P帧或B帧的解码的参考。I帧可以由编码器生成以创建随机访问点(以允许解码器在给定图片位置从头开始正确地解码)。当区分图像细节禁止生成有效的P或B帧时,可以生成I帧。因为I帧包含完整图片,所以I帧通常比P帧或B帧需要更多的位来编码。当在输入视频中检测到场景变化时,视频帧通常被编码为I帧。

[0061] P帧需要一些其他图片的先前解码以便被解码。与I帧相比,P帧通常需要更少的位来进行编码。P帧包含关于以解码顺序相对于先前I帧的变化的编码信息。P帧通常参考图片组(GoP)中的前一个I帧。P帧可以包含图像数据和运动向量位移以及两者的组合。在一些标准编解码器(诸如MPEG-2)中,P帧在解码期间仅使用一个先前解码图片作为参考,并且要求该图片以显示顺序也在P帧之前。在H.264中,P帧可以在解码期间使用多个先前解码图片作为参考,并且可以具有相对于用于其预测的图片的任意显示顺序关系。

[0062] B帧需要I帧或P帧的先前解码以便被解码。与P帧一样,B帧可以包含图像数据和运动向量位移和/或两者的组合。B帧可以包括一些预测模式,其通过对使用两个不同的先前解码的参考区域而获得的预测求平均来形成运动区域(例如,诸如宏块或更小区域的帧片段)的预测。在一些编解码器(诸如MPEG-2)中,B帧从不用作预测其他图片的参考。因此,较低质量的编码(导致使用比其他方式使用的位更少的位)可以用于此类B图片,因为细节的丢失不会损害后续图片的预测质量。在其他编解码器(诸如H.264)中,B帧可以用作或可以不用作其他图片的解码的参考(由编码器决定)。一些编解码器(诸如MPEG-2)在解码期间恰好使用两个先前解码的图片作为参考,并且要求这些图片中的一个以显示顺序在B帧图片之前并且另一个图片跟随它。在其他编解码器(诸如H.264)中,B帧可以在解码期间使用一个、两个或多于两个先前解码的图片作为参考,并且可以具有相对于用于其预测的图片的任意显示顺序关系。与I帧或P帧相比,B帧通常需要更少的位来进行编码。

[0063] 如本文所使用,术语I帧、B帧和P帧可以应用于具有与I帧、B帧和P帧类似的属性的任何流式数据单元,例如,如上面关于流式视频的上下文所描述的。

[0064] 通常存在用于P帧和B帧的两种预测模式,其通常称为帧间预测和帧内预测。帧间

预测基于与正被编码的帧不同的帧的区段。帧内预测基于正被编码的帧内的区段。许多编解码器包括预测模式决策过程,其在最终确定用于对帧内的特定帧或区段编码的模式之前比较两种模式的位率使用。I帧的编码不需要运动搜索或预测模式决定。

[0065] 为了对数字视频图片进行编码,编码器接收多个数字图像并对每个图像进行编码。数字图像的编码可以逐个区段地进行。每个区段的编码过程可以可选地涉及填充、图像压缩和运动补偿。如本文所使用,图像压缩是指将数据压缩应用于数字图像。图像压缩的目的是减少给定图像的图像数据的冗余,以便能够以压缩数据的有效形式存储或传输该图像的数据。

[0066] 熵编码是一种编码方案,其将代码分配给信号以便将代码长度与信号的概率相匹配。通常,熵编码器用于通过将由等长码表示的符号替换为由与概率的负对数成比例的码表示的符号来压缩数据。

[0067] CABAC是用于H.264/MPEG-4 AVC和高效视频编码(HEVC)标准的熵编码的形式。CABAC以提供比视频编码中使用的大多数其他熵编码算法更好的压缩而著称,并且它是为H.264/AVC编码方案提供比其前辈更好的压缩能力的关键元素之一。然而,应当注意到CABAC使用算术编码,其可能需要更大量的处理来解码。

[0068] 上下文自适应可变长度编码(CAVLC)是在H.264/MPEG-4 AVC视频编码中使用的熵编码的形式。在H.264/MPEG-4 AVC中,它用于对残余的Z字形顺序的变换系数块进行编码。它是CABAC的替代方案。CAVLC使用表查找方法并因此需要比CABAC少得多的解码处理,但它不那么有效地压缩数据。由于CABAC倾向于提供更好的压缩效率(比CAVLC多约10%的压缩),因此CABAC在生成编码位流时受到许多视频编码器的青睐。

[0069] 晕动病的减少

[0070] 根据本公开的方面,系统可以能够通过减小高分辨率ROI的大小来减少观看者发生晕动病的机会。系统可以通过本领域已知的方式来检测观看者经历运动相关的疾病,包括但不限于心率、体温、皮电反应或心电图(ECG)。有关检测晕动病的更多信息可以在可在<http://scitecin.isr.uc.pt/Proceedings/Papers/EPCGI/17.pdf>获得的Patroa、Bruno等人的“*How to Deal with Motion Sickness in Virtual reality*”中找到,其内容通过引用并入本文。在本公开的另一个实施方案中,系统通过监测视频流的数据速率来检测晕动病发生的可能性。视频流的数据速率与屏幕上显示的数据量有关,因此高数据速率指示晕动病的可能性增加,因为将存在大量新显示的信息。

[0071] 根据本公开的附加方面,系统还可以采用观看者反馈来减少晕动病。在该系统的一个实施方案中,采用来自观看者的触觉反应来启动晕动病减少方案。系统还可以收集来自观看者的反应并使用反应的统计分析来生成视频流内的潜在晕动病的模型。在另一个实施方案中,系统可以使用用户定义的晕动病阈值,当例如视频流数据速率或视频内的运动向量的变量超过阈值时,启动晕动病减少方案。

[0072] 在检测到与晕动病相关的反应时,系统可以缩小ROI的显示区域以减少向观看者显示的高分辨率信息的量。系统可以将显示区域的较大数量高分辨率区段与低分辨率视频流混合以生成较小的高分辨率区段,其在ROI边缘周围的质量大幅但逐渐劣化。

[0073] 质量混合和周边调光

[0074] 从ROI到图像其余部分的图像分辨率的急剧劣化可能会使观看者感到不安。这样,

根据本公开的方面,高分辨率区域的边缘可以与低分辨率区域混合以在区域之间产生质量劣化梯度。这种劣化梯度可以是从小分辨率到高分辨率的线性或非线性过渡。举例来说,非线性过渡可以是相对于与感兴趣区域相距的距离在高分辨率与低分辨率之间的S形(“S”形)函数过渡。通常,任何平滑、正、“凹凸形”函数的积分都是S形的。S形函数的示例包括但不限于逻辑函数、广义逻辑函数,S形函数包括普通反正切、双曲正切、Gudermannian函数和

误差函数 $\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$)、互补误差函数 $(1 - \text{erf}(x))$ 和代数函数如 $f(x) = \frac{x}{\sqrt{1+x^2}}$ 。

[0075] 逻辑函数具有形式 $f(x) = \frac{L}{1 + e^{-k(x-x_0)}}$, 其中:

[0076] x_0 = S形中点的x值,

[0077] L = 曲线的最大值, 和

[0078] k = 曲线的陡度。

[0079] 附加地,系统可以在围绕ROI边缘使用多个分辨率的视频流,以在高分辨率区域和显示器的其余部分之间生成更平滑的质量劣化。

[0080] 在本公开的其他方面中,ROI周围的区域可以调光以节省电力并产生类似剧院的效果。调光功能可以由观看者选择的功能或嵌入在视频流中的功能。

[0081] 实现方式

[0082] 图4示出了根据本公开的方面的系统400。所示系统400包括两个计算组,即编码组402和解码组401。编码组402经由网络403耦接到解码组401,该网络例如可以是无线网络、有线局域网连接(LAN)、USB连接、串行连接或本领域已知的其他设备连接线。编码组401包括:色彩空间转换(CSC)404、缩小缩放器405、感兴趣区域(ROI) I帧编码器406和缩放图像AVC/HEC编码器407。解码器组包括ROI I帧解码器409、AVC/HEV解码器410、放大缩放器411和视频流混合器/CSC 412。附加地,解码器组耦接到显示设备。注视跟踪单元可以耦接到显示设备413,并且被配置成将注视跟踪信息414传输到编码器组402中的ROI I帧编码器406。

[0083] 编码组可以从扫描输出单元408接收未编码的视频数据流。未编码的视频数据流可以包括来自视频游戏的视频数据或编码的数据流,其必须在被重新编码之前被解码以供当前公开的系统(未示出)使用。然后将初始未修改的视频流发送到CSC 404,其将视频数据流转换为适合于编码的形式。例如,图像的颜色空间可以从RGB(其对于显示器更典型)转换为YUV颜色空间(其对于编码器来说更典型)。然后,CSC 404将初始视频流传递给ROI I帧编码器406和缩小缩放器405。

[0084] 图5A示出了编码器组的ROI编码器侧。ROI I帧编码器406接收从注视跟踪器接收的注视跟踪信息414。根据本公开的方面,ROI I帧编码器406可以使用注视跟踪信息414来确定初始视频流内的ROI。ROI I帧编码器仅对位于ROI中的初始视频流图像的子区段进行编码。

[0085] 对应于ROI的编码子区段的形状可以被编码为正方形,或者该形状可以是圆形、椭圆形、矩形或其他规则或不规则形状。可以存在由ROI I帧编码器编码的两个ROI区段,其对应于观看者的两只眼睛。注视跟踪系统可以检测眼睛的存在或不存在,并且ROI I帧编码器可以调整编码的区段的数量以解决眼睛的不存在。可替代地,ROI I帧编码器可以具有被配置成在针对双眼的ROI中显示单个宽区段的逻辑。

[0086] 根据本公开的方面,构成ROI的图像的区段可以大于观看者的实际ROI,以解决视觉半径的任何微小变化和突然眼睛移动。构成ROI的图像区段可以具有任何合适的大小,例如20个宏块宽×20个宏块长。ROI区段的大小可以在放大操作期间减小或者在缩小操作中增加以减少晕动病。如在先前区段中所讨论的,编码的子区段的大小也可以减小,以在视频流的显示期间减少晕动病。

[0087] ROI I帧编码器406仅将ROI中的视频流中的图像的子区段编码为内帧(I帧),以便减少等待时间和帧级延迟。通过将ROI编码为I帧,省略了某些耗时的操作(例如,运动搜索和编码模式决定),由此减少了等待时间。因为与帧的大小相比ROI相对较小,所以与包含整个图像的普通I帧相比,I帧编码的ROI需要更少的位。ROI I帧编码器可以使用本领域中已知的任何低损耗编码器。适用于此目的的低损耗编码器的示例非限制地是X.264、H.264无损编码器。ROI I帧编码器不需要对ROI外部的视频图像的区段进行编码或发送。视频图像的低损耗编码区段(下文称为低压缩区段)可以使其在原始视频帧内的位置写入低压缩区段的帧报头,使得解码器随后可以在帧内的正确位置重建ROI。一旦已经处理了低压缩子区段,它们就通过网络发送到解码器侧。

[0088] 图5B描绘了编码器组的高压缩侧。缩小缩放器405可以从CSC 404接收初始视频帧。缩小缩放器405可以将大小减小算法应用于视频帧,以降低传输视频流所需的数据速率和编码所需的处理能力。如本领域普通技术人员所知,任何缩放算法都可以用于该过程。合适的算法的示例非限制地可以是双线性插值、三次插值或Lanczos重采样。

[0089] 然后,AVC/HEV编码器407可以从缩小缩放器405接收缩小缩放的视频帧。然后使用有损编码器对缩小缩放的视频流进行编码以减少编码时间和带宽要求。适用于本申请人的编码器可以是但不限于H.264/AVC、或H.265/HEVC。在对视频帧进行编码之后,“高压缩视频帧”通过网络被发送到解码器侧。

[0090] 图6A示出了解码器组的I帧解码器侧。I帧解码器409通过网络从I帧编码器接收低压缩区段数据。然后,低压缩区段数据由I帧解码器409解码,并且解码的区段被发送到混合器/CSC 412。已知的低损耗解码器适用于本公开的应用,包括但不限于FFmpeg。

[0091] 图6B描绘了解码器组的高压缩侧。根据本公开的方面,AVC/HEVC解码器410通过网络从编码器组接收高压缩视频。使用合适的解码器对高压缩视频流进行解码,并将其发送到放大缩放器411。

[0092] 放大缩放器411从AVC/HEVC解码器410接收缩小缩放的视频流。然后使用适当的放大缩放算法对缩小缩放的视频流进行放大缩放,以匹配编码器组中应用的算法。然后将放大缩放的视频流发送到混合器/CSC 412。

[0093] 混合器/CSC 412从放大缩放器411接收视频流,并且从I帧解码器409接收子区段流。根据本公开的方面,混合器/CSC 412可以读取子区段流的帧报头以确定视频流内的放置子区段的位置。可替代地,混合器/CSC可以与ROI I帧编码器同步,以使用注视跟踪数据414来确定子区段流(未示出)的放置位置。根据本公开的附加方面,混合器/CSC 412可以通过诸如单独的有线或无线传输(未示出)之类的其他信道接收子区段流的位置。

[0094] 可以在混合器/CSC 412处将子区段流与视频流混合。可以通过将掩码应用于子区段流和视频流来完成混合。掩码将为子区段流和视频流的图像中的像素值设置权重。掩码可以定义在高分辨率子区段流与低分辨率视频流之间的线性分辨率劣化梯度。根据本公开

的替代方面,掩码可以定义非线性分辨率劣化梯度。然后,混合器/CSC 412将通过以下方式来组合子区段流和视频流:在适当位置将子区段流叠加视频流顶部上并在图像上应用设置的像素值,使得根据像素值权重混合高分辨率图像和低分辨率图像。

[0095] 应用于图像的掩码可以确定所显示的ROI子区段的形状。这样,所显示的ROI子区段的大小和形状可以在混合器/CSC 412处变化。ROI的大小和形状可以在逻辑中固定或变化。混合器/CSC 412可以从ROI I帧编码器获取ROI形状并进一步细化ROI的边缘。掩码还可以用于定义视频流的像素值以实现外围调光。

[0096] 混合器/CSC 412还可以采用具有多于一个潜在ROI子区段的子区段流,并根据注视跟踪数据确定观看者的正确ROI。因此,混合器/CSC 412将掩码应用于潜在的ROI子区段,从而隐藏它们而不隐藏正确的ROI子区段。

[0097] 在替代实施方案中,在混合器/CSC 412处将单个掩码应用于ROI子区段流,并且然后根据应用于ROI子区段流的像素权重通过混合将ROI子区段流叠加在视频流上。该混合方案也可以在ROI I帧编码器处应用。在这样的实施方案中,在ROI I帧编码器处应用掩码像素值,然后在混合器/CSC 412处根据那些掩码像素值将ROI子区段流与视频流混合。

[0098] 根据本公开的附加方面,混合器/CSC 412可以通过基于晕动病反应改变掩码大小或形状来实现如上所述的晕动病减少方案。混合器/CSC 412还可以改变分辨率梯度作为晕动病减少方案的一部分。

[0099] 图7描绘了根据本公开的方面的掩码有凹视频。有凹显示可以具有围绕观看者的中央注视点或ROI 703的一个或多个高分辨率子区段701。系统可以在高分辨率的边缘702周围应用掩码,以实现高分辨率子区段701与低分辨率背景700之间的分辨率劣化梯度。系统可以使用注视跟踪数据来预测ROI的运动向量704。在其他实施方案中,掩码区段702的区域可以响应于晕动病反应而被扩大。可替代地,编码的高分辨率区域705可以由于晕动病反应而减少。

[0100] 图8描绘了晕动病减少方案的框图。运动减少方案800可以在图4的有凹显示系统的编码器侧402或解码器侧401上运行。有凹视频的观看者可以设置移动的阈值水平,在高于该阈值水平时观看者经历晕动病804。可替代地,阈值水平804可以由来自观看者的平均晕动病反应或经验确定的由于晕动病引起的生理反应来确定。系统连续测量视频流或传感器数据内的变量以确定反应是否已超过设置的阈值水平802。监测可以在图3所示的眼睛跟踪显示设备301或计算设备360处进行。根据本公开的附加方面,阈值可以由观看者手动触发。如果设备确定已经超过晕动病阈值,则系统将向ROI I帧编码器406或混合器/CSC 412通知801减少显示的高分辨率区域701。

[0101] 其他显示实现方式

[0102] 尽管本文关于头戴式显示器(HMD)应用描述了示例,但是本公开的方面不限于这些实现方式。HMD实现方式表示相对简单的实现方式,因为用户眼睛和显示屏的相对位置保持或多或少固定。然而,原则上,所公开的系统和方法可以适用于可与注视跟踪一起工作的任何图像显示系统。可以修改注视跟踪系统以在用户的头部和眼睛相对于显示屏不固定的实现方式中跟踪用户的头部和眼睛相对于显示屏的位置和定向。

[0103] 图9A-9H示出了结合本公开的方面的使用面部定向和眼睛注视方向的示例。如在图9A所见,用户的面部920可以出现在通过用户训练的相机获得的图像922_A中。此类相机是

诸如膝上型计算机、智能电话和平板计算机的设备的共同特征。图像分析软件可以识别面部920上的参考点。该软件可以表征这些参考点中的某些,例如,位于嘴部的拐角924_M、鼻部的桥接部924_N、头发中的部分924_H、以及眉毛的顶部924_E,其相对于面部920基本固定。软件还可以将用户眼睛的瞳孔926和拐角928识别为参考点,并确定瞳孔相对于眼角的位置。在一些实现方式中,可以根据眼睛的瞳孔926和拐角928的位置估计用户眼睛的中心。然后,可以估计眼睛的中心,并且可以将瞳孔的位置与中心的估计位置进行比较。在一些实现方式中,可以使用面部对称属性。

[0104] 软件可以通过分析参考点和瞳孔926的相对位置来确定用户的面部特性,例如头部倾斜角度和眼睛注视角度。例如,软件可以通过让用户直视相机来初始化参考点924_E、924_H、924_M、924_N、928,并且将参考点和瞳孔926的位置记录为初始值。然后,软件可以针对这些初始值将头部倾斜角度和眼睛注视角度初始化为零。随后,每当用户直视相机时,如图9A以及图9B所示的对应顶视图,参考点924_E、924_H、924_M、924_N、928和瞳孔926应当处于或接近其初始值。

[0105] 作为示例,并且不作为限制,可以使用五个参考点来估计用户头部的姿势,即每只眼睛的外拐角928、嘴部的外拐角924_M和鼻尖(未示出)。可以通过连接眼睛的中点(例如,眼睛外拐角928之间的中间点)和嘴的中点(例如,在嘴部外拐角924_M之间的中间点)之间的线来找到面部对称轴。可以从鼻子的3D角度在弱透视几何下确定面部方向。可替代地,可以使用相同的五个点来根据到平面的法线确定头部姿势,该法线可以从平面偏斜对称和鼻子位置的粗略估计中找到。头部姿势估计的进一步细节可以在例如IEEE关于模式分析和机器智能的学报,第31卷,第4期,2009年4月,第607-626页中的Erik Murphy的“Head Pose Estimation in Computer Vision:A Survey”中找到,其内容通过引用并入本文。可以与本发明的实施方案结合使用的头部姿势估计的其他示例在图案识别,第33卷(2000年7月7日),第1783-1791页中的Athanasios Nikolaidis的“Facial feature extraction and pose determination”中描述,其全部内容通过引用并入本文。可以结合本发明的实施方案使用的头部姿势估计的附加示例在于FG'00第四届IEEE自动面部和手势识别国际会议论文集,2000年,第499-505页中的Yoshio Matsumoto和Alexander Zelinsky“An Algorithm for Real-time Stereo Vision Implementation of Head Pose and Gaze Direction Measurement”中描述,其全部内容通过引用并入本文。可以结合本发明的实施方案使用的头部姿势估计的另外示例在图像和视觉计算,第20卷,第7期,2002年2月20日,第499-511页的Qiang Ji和Ruong Hu的“3D Face Pose Estimation from a Monocular Camera”中描述,其全部内容通过引用并入本文。

[0106] 当用户倾斜他的头部时,图像中的参考点之间的相对距离可以取决于倾斜角度而改变。例如,如果用户将他的头部围绕垂直轴线Z向右或向左枢转,则眼睛的拐角928之间的水平距离 x_1 可以减少,如图9C所示的图像922_C所示。取决于所使用的特定头部姿势估计算法,其他参考点也可以工作或更容易检测。距离的量变化可以与枢转角 θ_H 相关,如图1E中的对应顶视图所示。应当注意,如果枢轴纯粹围绕Z轴,则例如鼻部的桥接部924_N处的参考点与嘴部的拐角924_M处的参考点之间的垂直距离 y_1 将预期不会显著改变。然而,如果用户将头部向上或向下倾斜,则可以合理地预期该距离 y_1 会改变。还应注意,当确定瞳926相对于眼睛的拐角928的位置以用于注视方向估计时,软件可以考虑头部枢转角 θ_H 。可替代地,软件

可以在确定头部枢转角 θ_H 时考虑瞳孔926相对于眼睛拐角928的位置。如果注视预测更容易,例如通过手持设备上的红外光源,瞳孔可以相对容易地定位,则这种实现可能是有利的。在该示例中,如图9C和图9D所示,用户的眼睛注视角度 θ_E 或多或少地与用户的头部倾斜角度对准。然而,由于用户头部的枢转和眼球形状的三维性质,与瞳孔在初始图像922A中的位置相比,瞳孔926的位置在图像922D中将出现为略微移位。

[0107] 在一些情况下,用户可能正面向相机,但用户的眼睛注视被引导到其他地方,例如,如图9E和图9F中的对应顶视图所示。在该示例中,用户的头部是倾斜角度 θ_H 是零,但眼睛注视角度 θ_E 不是零。相反,用户的眼球逆时针旋转,如图9F所见。因此,参考点924E、924H、924M、124N、928如图9A中那样布置,但瞳孔126在图像922E中向左移位。

[0108] 应当注意,用户的头部可以在一个方向上枢转,并且用户的眼球可以在另一个方向上枢转。例如,如图9H和图9I所示,用户101可以顺时针枢转他的头部并逆时针旋转他的眼球。因此,参考点924E、924H、924M、924N、928如图9D中那样移位,但瞳孔926在图9G所示的图像922G中向右移位。注视跟踪系统100,如图1A-1B所述,可以在确定用户眼睛E的注视方向GD时考虑该配置或上述任何配置。

[0109] 从前面的讨论可以看出,可以仅使用相机来跟踪某些用户的面部定向特性。然而,也可以使用面部定向特性跟踪设置的许多替代形式。图10A-10E示出了根据本发明的实施方案可以实现的五个面部定向特性跟踪系统的示例,以及其他可能的系统。

[0110] 在图10A中,用户1001面向安装在视觉显示器1003的顶部上的相机1005和红外光传感器1007。为了跟踪用户的头部倾斜角度,相机1005可以被配置成执行对象分割(即,跟踪用户的单独的身体部位),并且然后根据所获得的信息估计用户的头部倾斜角度。相机1005和红外光传感器1007耦接到运行软件1013的处理器1013,其可以如上所述地配置。作为示例,并且不作为限制,可以使用运动模型来完成对象分割,以描述目标的图像会如何根据对象的不同可能移动而改变。应当注意,本发明的实施方案可以使用多于一个的相机,例如,一些实现方式可以使用两个相机。一个相机可以提供视野的缩小图像以定位用户,并且第二相机可以放大并聚焦在用户脸部上以提供特写图像以更好地进行头部和注视方向估计。

[0111] 还可以使用该设置来获取用户的眼睛注视方向。作为示例,并且不作为限制,红外光可以最初从红外光传感器1007朝向用户的眼睛引导并由相机1005捕获反射。从反射的红外光提取的信息将允许耦接到相机1005的处理器确定用户的眼睛旋转量。基于视频的眼睛跟踪器通常使用角膜反射和瞳孔中心作为随时间跟踪的特征。

[0112] 因此,图10A示出了根据本发明的实施方案的面部定向特性跟踪设置,其被配置成跟踪用户的头部倾斜角度和眼睛注视方向。应当注意,出于示例的目的,已经假设用户直接在显示器和相机对面。然而,即使用户没有直接在显示器1003和/或相机1005对面,也可以实现本发明的实施方案。例如,用户1001可以是显示器右侧/左侧+45°或-45°。只要用户1001在相机205的视野内,就可以估计头部角度 θ_H 和眼睛注视 θ_E 。然后,可以根据用户1001相对于显示器1003和/或相机1005的位置(例如,如图10A所示的身体角度 θ_B)、头部角度 θ_H 和眼睛注视 θ_E 来计算归一化角度。作为示例,并且不作为限制,如果用户1001被定位成使得身体角度 θ_B 是+45°,并且如果头部以-45°的角度 θ_H 转动,则用户1001通过转动他的头部来固定身体与显示器1003的偏差,并且这几乎和让人直视显示器一样好。具体地,如果例如用户的注视角

度 θ_E 为零(即,用户的瞳孔居中),则归一化角度(例如, $\theta_B+\theta_H+\theta_E$)为零。

[0113] 图10B提供另一种面部定向特性跟踪设置。在图10B中,用户1001面向安装在视觉显示器1003顶部上的相机1005。用户1001同时佩戴一副眼镜1009(例如,一对3D快门眼镜),其具有一对间隔开的红外(IR)光源1011(例如,眼镜1009的每个镜片上的一个IR LED)。相机1005可以被配置成捕获从光源1011发出的红外光,并且然后根据所获得的信息三角测量用户的头部倾斜角度。因为光源1011的位置相对于其在用户面部上的位置不会显著变化,所以该设置将提供对用户头部倾斜角度的相对准确的估计。

[0114] 眼镜1009可以附加地包括相机1010,该相机可以向处理器1013提供图像,其可以与软件1012结合使用以找到视觉显示器1003的位置或估计视觉显示器203的大小。作为示例,并且不作为限制,视觉显示器是具有已知的垂直和水平屏幕尺寸的已知类型。可以显示相对于屏幕的已知大小的测试图像。可以由相机获得测试图像的图像,并且对其进行分析以确定由相机1010获得的图像中的测试图像的定向和尺寸。收集此信息允许系统归一化用户的面部定向特性数据,使得这些特性的计算与显示器1003和用户1001的绝对位置两者无关。此外,添加相机将使系统更准确地估计可见范围。因此,图2B示出了根据本公开的方面的用于确定用户的头部倾斜角度的替代设置。在一些实施方案中,可以将单独的相机安装到眼镜1009的面向用户眼睛的每个镜片,以便通过获得示出瞳孔相对于眼睛的中心或拐角的相对位置的图像来促进注视跟踪,例如,如以上所讨论的。眼镜1009相对于用户眼睛的相对固定位置有助于跟踪用户的眼睛注视角度 θ_E ,而与跟踪用户头部定向 θ_H 无关。

[0115] 图10C提供了第三面部定向特性跟踪设置。在图10C中,用户1001面向安装在视觉显示器1003顶部上的相机1005。用户还握持控制器1015,其具有被配置成促进用户1001与视觉显示器1003上的内容之间的交互的一个或多个相机1017(例如,每侧上有一个)。

[0116] 可以例如使用如以上示例中的显示的测试图像来分析来自相机1017的图像以确定视觉显示器1003的位置,或者估计视觉显示器1003的大小。收集此信息允许系统归一化用户的面部定向特性数据,使得这些特性的计算与显示器1003和用户1001的绝对位置两者无关。此外,将相机1017添加到控制器1015允许系统更准确地估计可见范围。

[0117] 重要的是应当注意,图10C中的设置可以进一步与图10A中的设置(图10C中未示出)组合,以便除了跟踪用户的头部倾斜角度之外还跟踪用户的眼睛注视方向,同时使系统独立于显示器大小和位置。因为在该设置中用户的眼睛是无遮挡的,所以可以通过上面讨论的红外光反射和捕获过程获得他的眼睛注视方向。

[0118] 图10D提供了又一种替代的面部定向特性跟踪设置。在图10D中,用户1001面向安装在视觉显示器1003顶部上的相机1005。用户1001还佩戴具有红外光源1021(例如,每个目镜上有一个)和传声器1023的耳机1019,耳机1019被配置成促进用户1001与视觉显示器1003上的内容之间的交互。很像图10B中的设置,相机1005可以捕获从耳机1019上的光源1021发出的红外光的图像,并且可以从分析所获得的图像对用户的头部倾斜角度进行三角测量。因为耳机1019的位置相对于其在用户面部上的位置趋于不会显著变化,所以该设置可以提供对用户头部倾斜角度的相对准确的估计。

[0119] 除了使用红外光传感器1021来跟踪用户的头部倾斜角度之外,用户头部相对于指定目标的位置也可以由不作为耳机1019的一部分的单独传声器阵列1027跟踪。传声器阵列1027可以被配置成便于确定用户语音的大小和定向,例如,使用在处理器1013上运行的适

当配置的软件1012。例如,在共同转让的美国专利7,783,061、共同转让的美国专利7,809,145和共同转让的美国专利申请公布号2006/0239471中描述了此类方法的示例,所有这三个的全部内容通过引用并入本文。

[0120] 可以在Ruxin Chen和Steven Osman于2010年9月23日提交的题为“BLOW TRACKING USER INTERFACE SYSTEM AND METHOD”的美国专利申请号12/889,347(代理人案卷号SCEA10042US00-I)中找到使用热成像信息对用户语音进行定向跟踪的详细说明,该专利申请通过引用并入本文。作为示例,并且不作为限制,可以使用热成像相机来确定用户语音的定向以检测用户嘴周围的空气中的振动模式,其对应于语音期间用户语音的声音。可以分析振动模式的时间演变以确定对应于用户语音的广义方向的向量。

[0121] 使用传声器阵列1027相对于相机1005的位置和用户关于传声器阵列1027的语音的方向两者,可以计算用户头部相对于指定目标(例如,显示器)的位置。为了在建立用户的头部倾斜角度方面实现更高的准确性,可以组合用于确定头部倾斜角度的红外反射和方向跟踪方法。替代实施方案可以附加地包括惯性传感器1027,如以上相对于图1A描述的。

[0122] 耳机1019可以附加地包括被配置成获得视觉显示器1003的图像的相机1025,可以对这些图像进行分析以找到显示器的位置和/或估计视觉显示器1003的大小。收集此信息允许系统归一化用户的面部定向特性数据,使得这些特性的计算与显示器1003和用户1001的绝对位置两者无关。此外,添加相机将使系统更准确地估计可见范围。在一些实施方案中,一个或多个相机1025可以被安装到耳机1019并面向用户眼睛,以便通过获得示出瞳孔相对于眼睛的中心或拐角的相对位置的眼睛的图像来促进注视跟踪,例如,如以上所讨论的。耳机1019(以及因此相机1025)相对于用户眼睛的相对固定位置有助于跟踪用户的眼睛注视角度 θ_E ,而与跟踪用户头部定向 θ_H 无关。

[0123] 重要的是应当注意,图10D中的设置可以与图10A中的设置(图10D中未示出)组合,以便除了跟踪用户的头部倾斜角度之外还跟踪用户的眼睛注视方向。因为在该设置中用户的眼睛是无遮挡的,所以可以通过上面讨论的红外光反射和捕获过程获得他的眼睛注视方向。

[0124] 本发明的实施方案还可以在手持设备中实现,诸如手机、平板计算机、个人数字助理、便携式互联网设备、或便携式游戏设备、以及其他示例。图10E示出了在手持设备1030的上下文中确定眼睛注视方向的一个可能示例。设备1030通常包括处理器1039,其可以用合适的软件编程,例如,如上所述。设备1030可以包括耦接到处理器1039的显示屏1031和相机1035。一个或多个传声器1033和控制开关1037还可以可选地耦接到处理器1039。传声器1033可以是传声器阵列的一部分。控制开关1037可以是通常与特定类型的手持设备一起使用的任何类型。例如,如果设备1030是蜂窝电话,则控制开关237可以包括数字键盘或字母数字键盘、触摸屏或触模板,如这些设备中常用的。可替代地,如果设备1030是便携式游戏单元,则控制开关1037可以包括数字或模拟操纵杆、数字控制开关、触发器等。在一些实施方案中,显示屏1031可以是触摸屏接口,并且控制开关1037的功能可以通过触摸屏结合合适的软件、硬件或固件来实现。相机1035可以被配置成当用户注视显示屏1031时面向用户1001。处理器1039可以用软件编程以实现头部姿势跟踪和/或眼睛注视跟踪。处理器还可以被配置成利用头部姿势跟踪和/或眼睛注视跟踪信息,例如,如以上所讨论的。

[0125] 应当注意,显示屏1031、传声器1033、相机1035、控制开关1037和处理器1039可以

安装到可容易地握在用户的一只或两只手中的壳体。在一些实施方案中,设备1030可以与一副专用眼镜结合操作,该眼镜可具有与图10B所示并在上文中描述的眼镜1009相同的特征。这种眼镜可以通过无线或有线连接与处理器通信,例如个人区域网络连接,诸如蓝牙网络连接。在一些实施方案中,设备1030可以与耳机结合使用,该耳机可具有与图10D所示并在上文中描述的耳机1019相同的特征。这种耳机可以通过无线或有线连接与处理器通信,例如个人区域网络连接,诸如蓝牙网络连接。设备1030可以包括合适的天线和收发器以促进无线网络连接。

[0126] 应当注意,图10A-10E中描绘的示例仅是根据本公开的方面的可用于跟踪用户面部定向特性的许多设置的几个示例。类似地,除了上述头部倾斜角度和眼睛注视方向之外,还可以跟踪各种身体和其他面部定向特性,以有助于调整在显示器上渲染的感兴趣区域。

[0127] 附加方面

[0128] 本公开的各方面还包括具有在其中实现的计算机可执行指令的非暂时性计算机可读介质,这些指令在被执行时根据上述方面实现视频处理,例如,如以上相对于图1A、图3、图4、图5A-5B、图6A-6C和图8所描述的。合适的计算机可读介质的示例包括但不限于RAM、ROM、硬盘、闪存驱动器、SDRAM、CD-ROM、蓝光、磁带和软盘。

[0129] 结论

[0130] 应当注意,已经参考使用红外光源的眼睛跟踪设备描述了本公开的各方面。然而,应当理解,其他实现方式是可能的。例如,在本公开的实现方式中,其他不可见光源是可能的,诸如紫外光。作为另一示例,在本公开的实现方式中,可见光源对于眼睛照明是可能的,尽管可能期望使用不可见光源以避免分散用户的注意力。

[0131] 虽然以上是对本发明优选实施方案的完整描述,但是可以使用各种替换、修改和等同物。因此,本发明的范围不应参考以上描述来确定,而是应该参考所附权利要求及其等同物的全部范围来确定。本文描述的任何特征(无论是否优选)可以与本文描述的任何其他特征(无论是否优选)组合。在随后的权利要求中,不定冠词“一(a)”或“一(an)”是指该冠词后面的项的一个或多个的数量,除非另有明确说明。所附权利要求不应被解释为包括手段加功能限制,除非在给定权利要求中使用短语“用于…的手段”明确地叙述了这种限制。

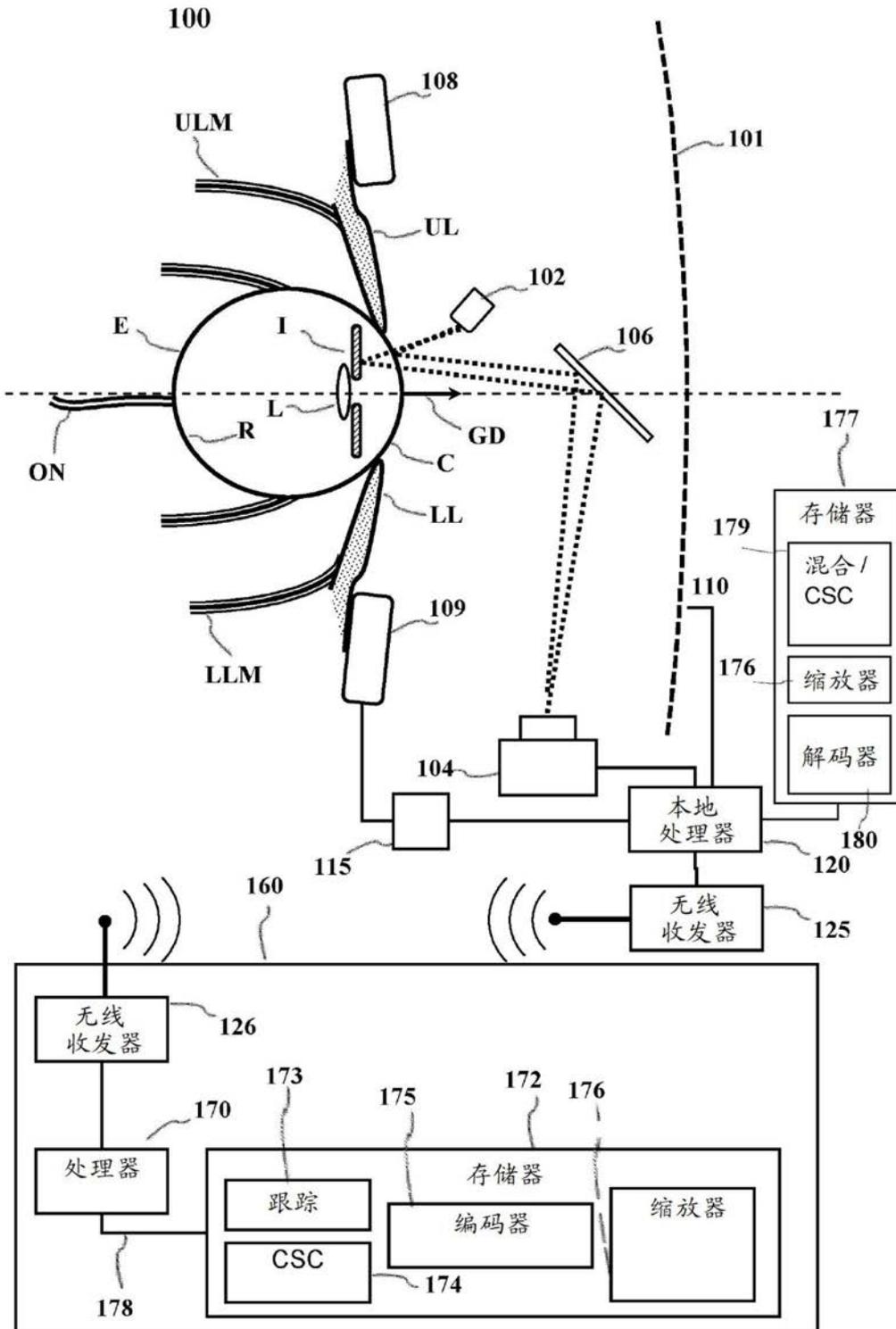


图1A

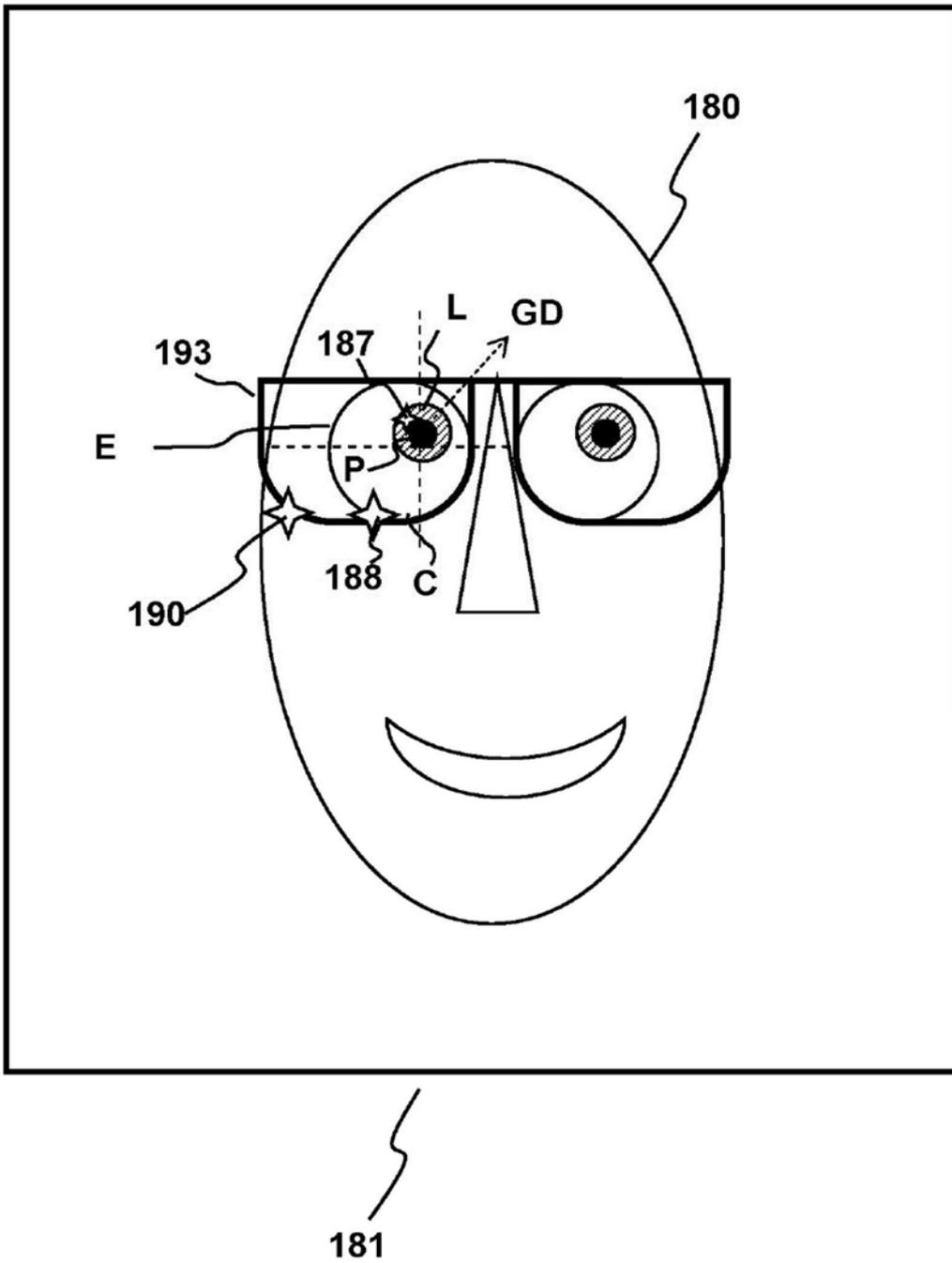


图1B

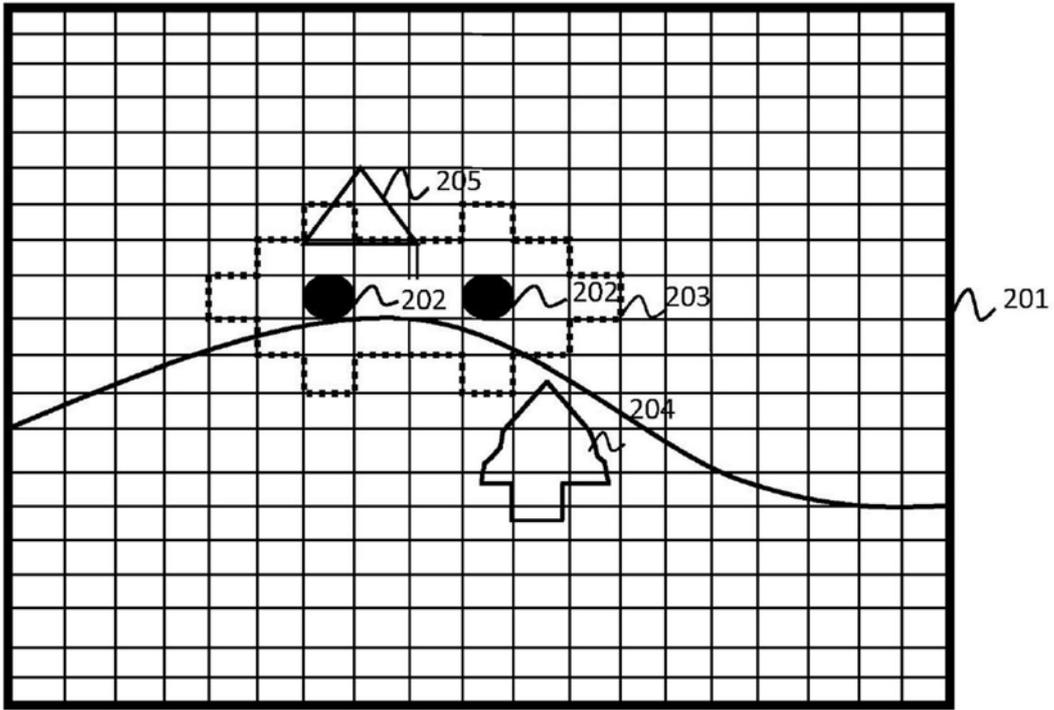


图2A

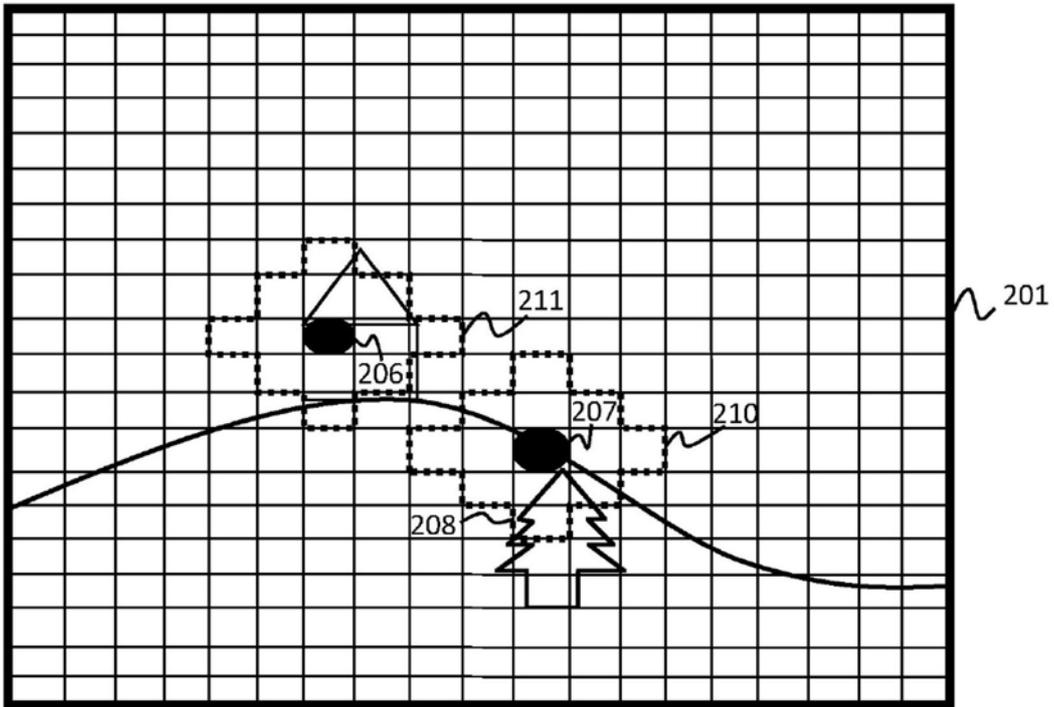


图2B

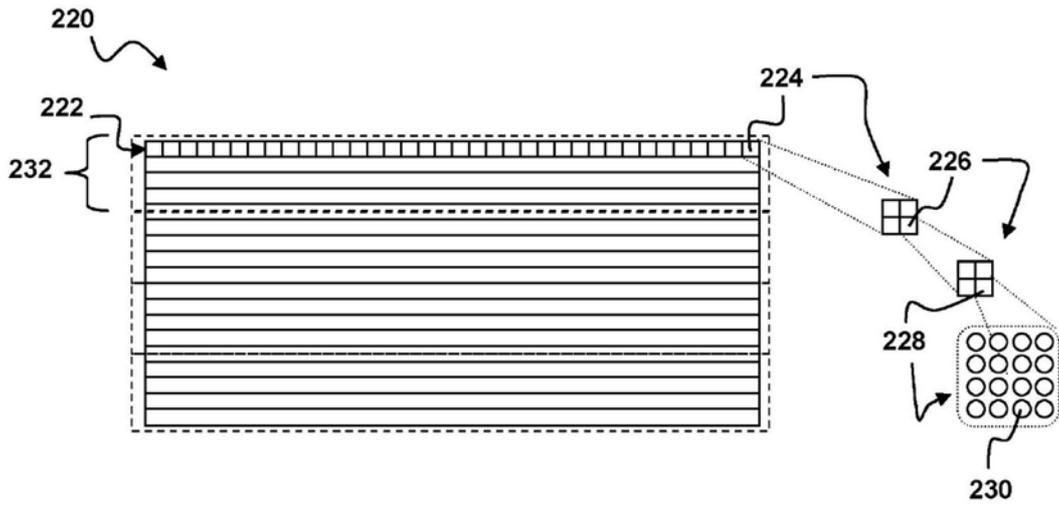


图2C

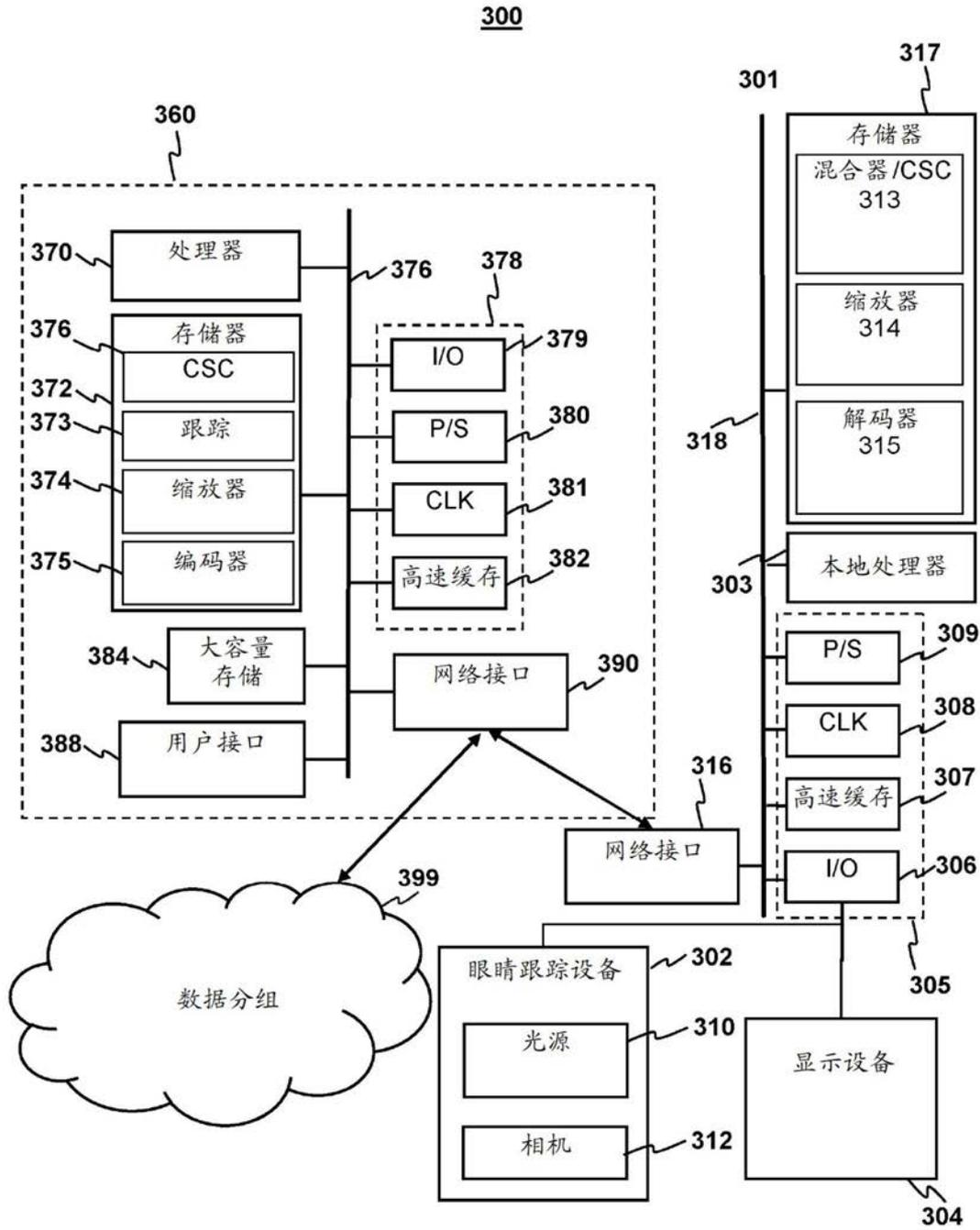


图3

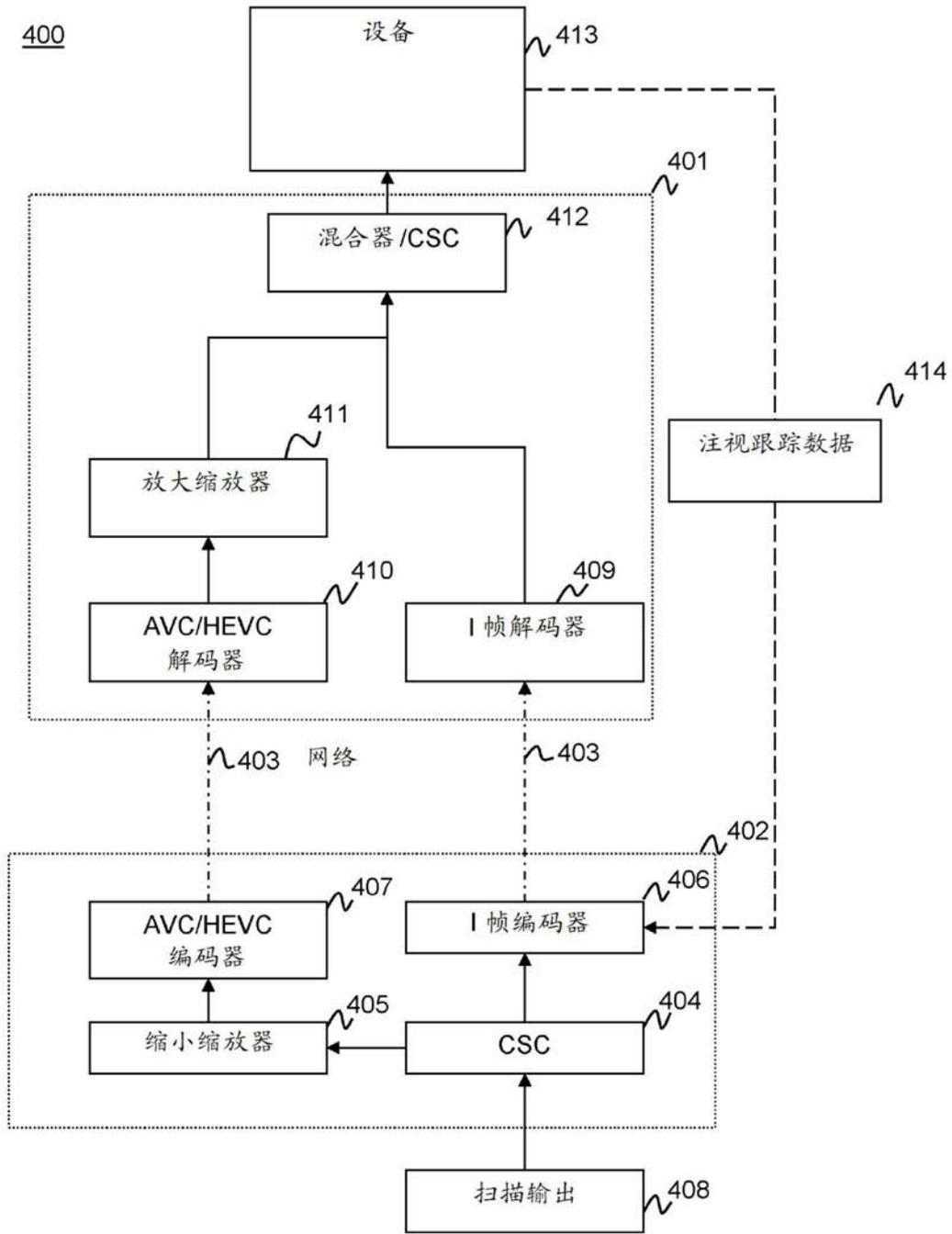


图4

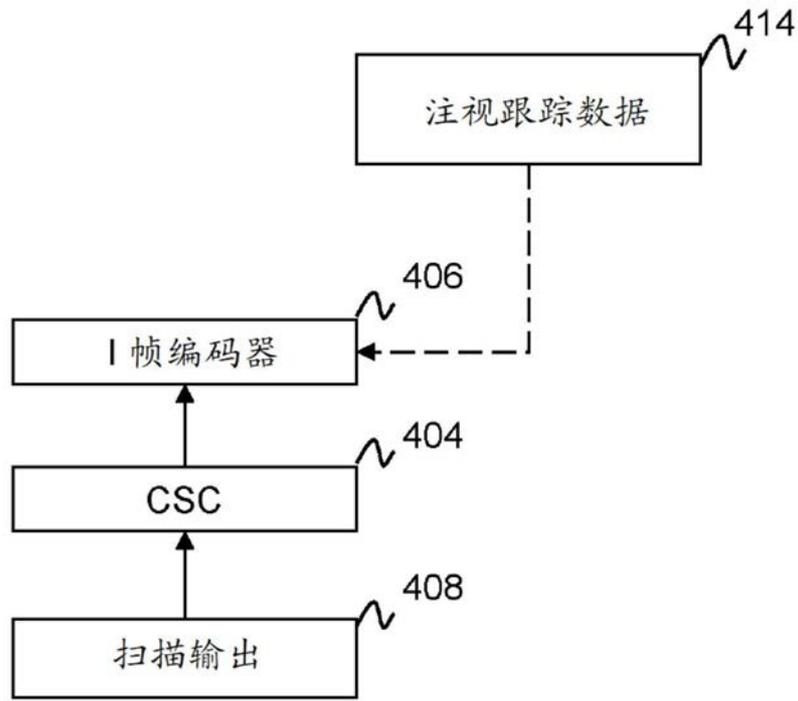


图5A

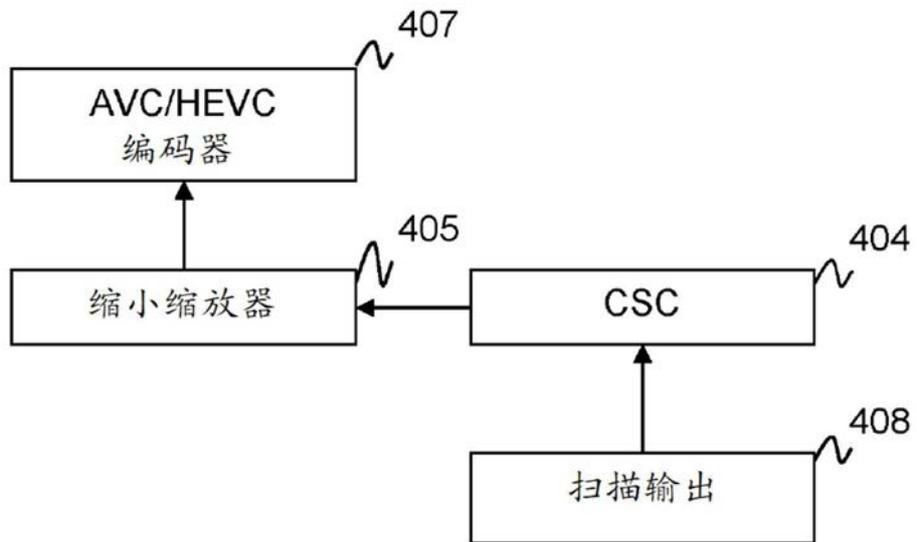


图5B

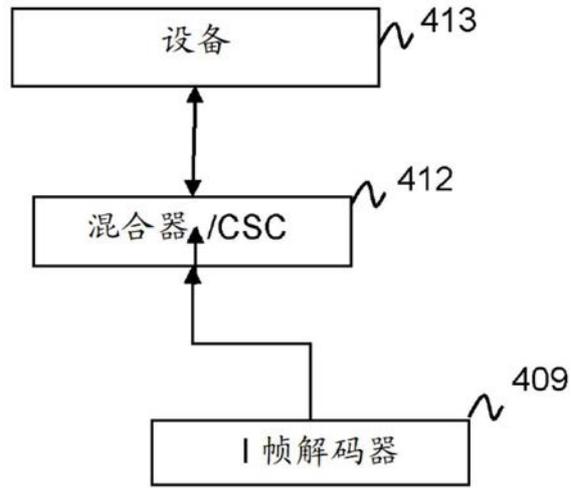


图6A

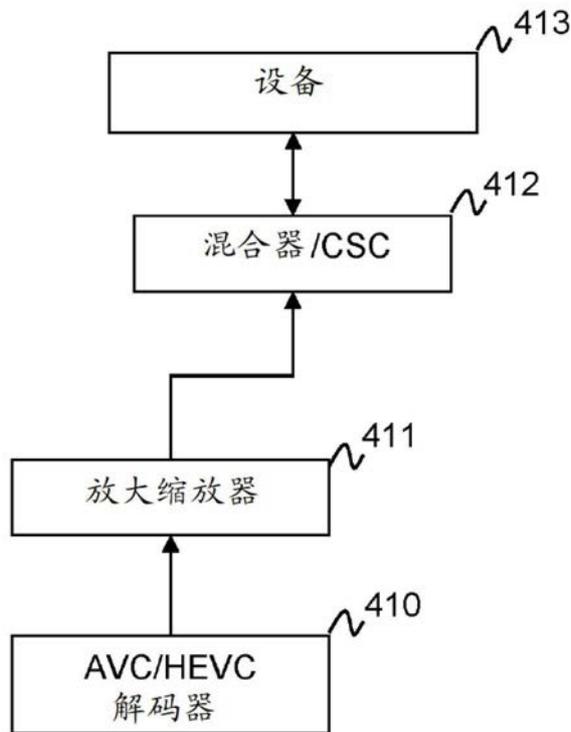


图6B

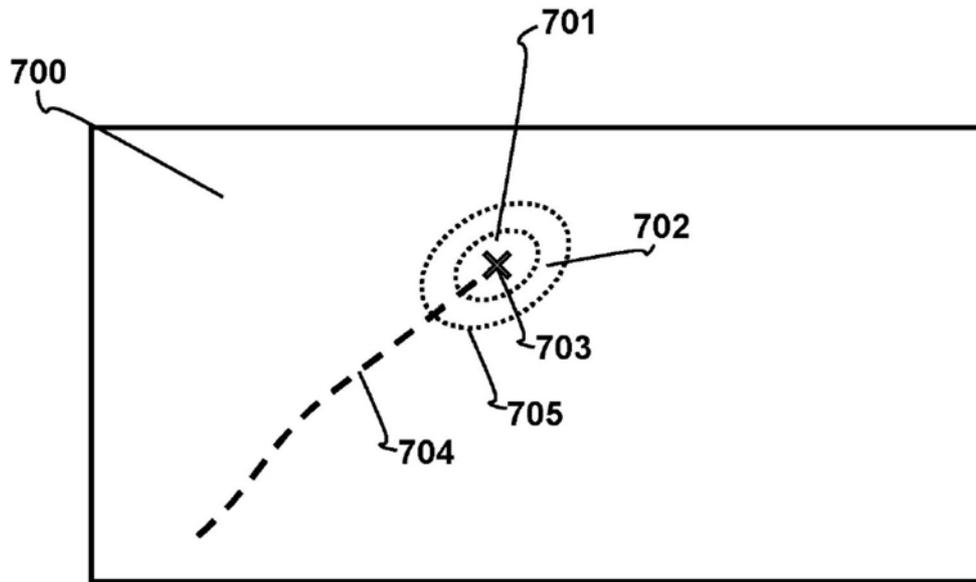


图7

800

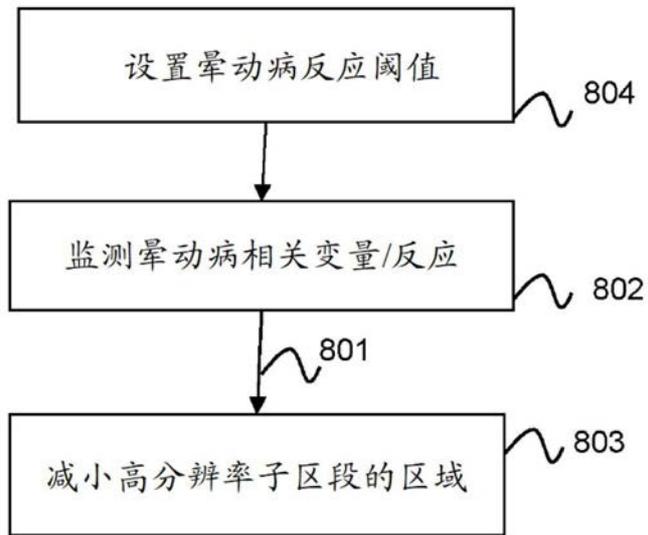


图8

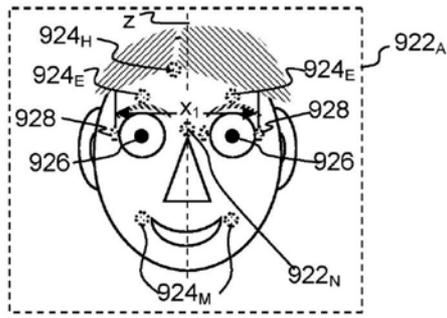


图9A

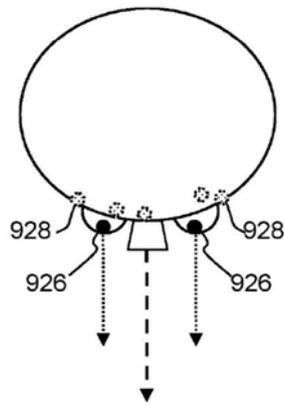


图9B

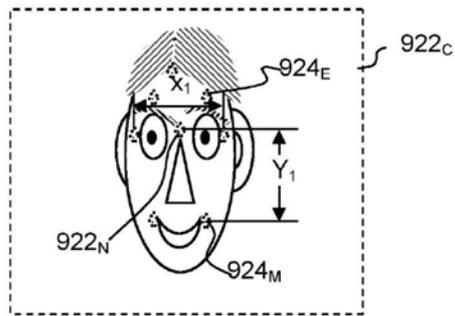


图9C

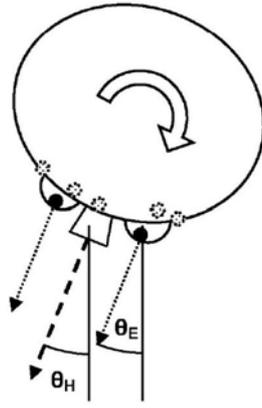


图9D

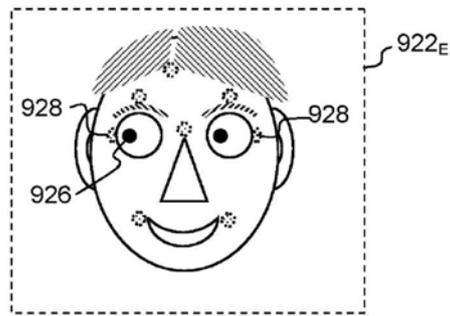


图9E

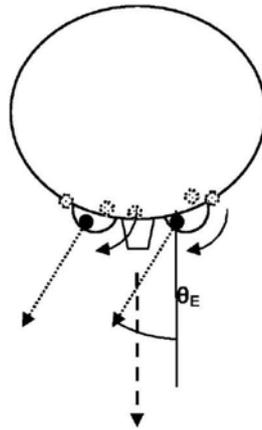


图9F

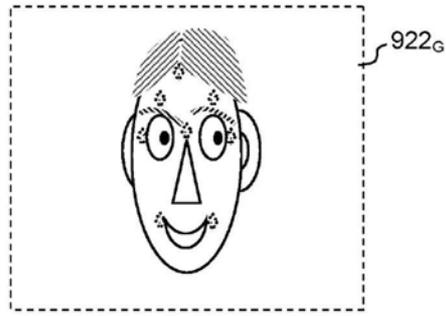


图9G

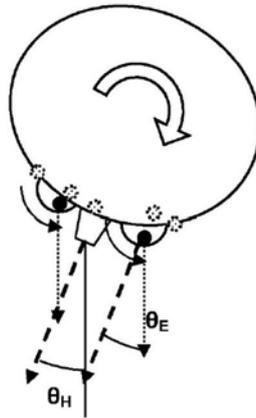


图9H

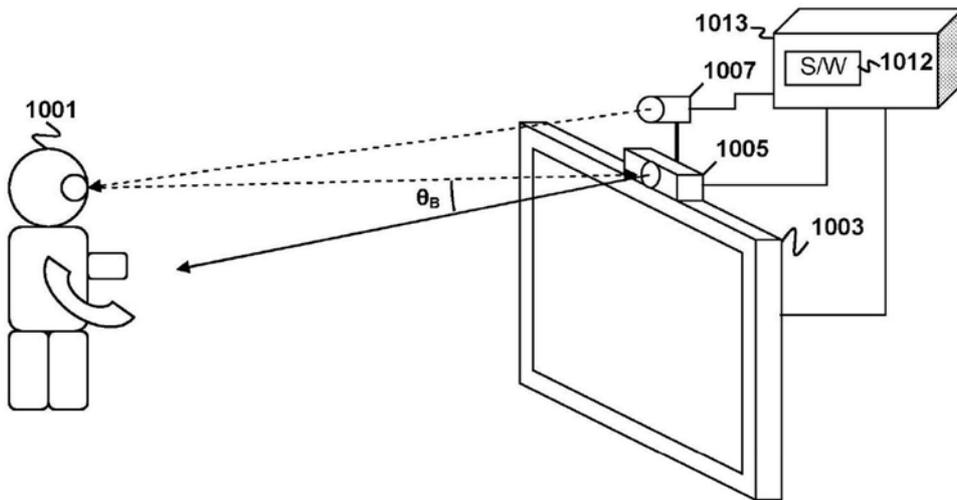


图10A

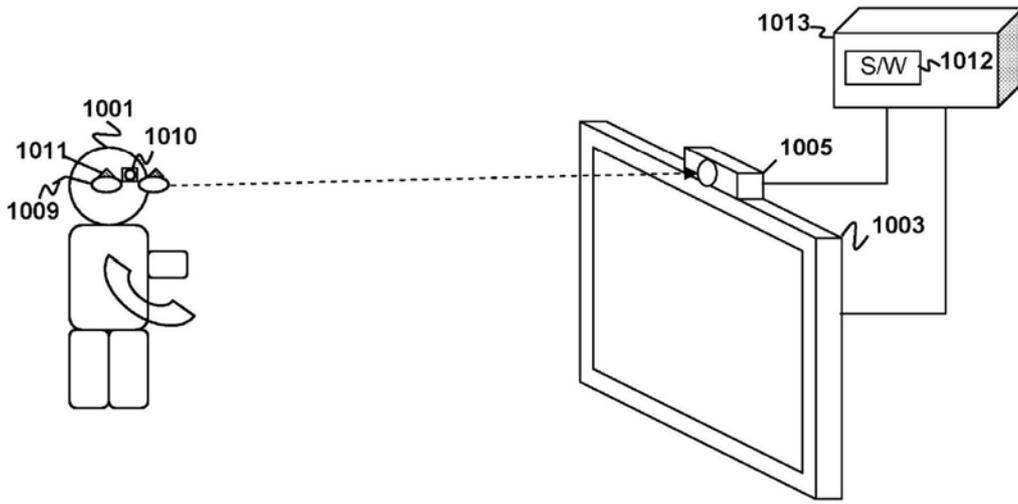


图10B

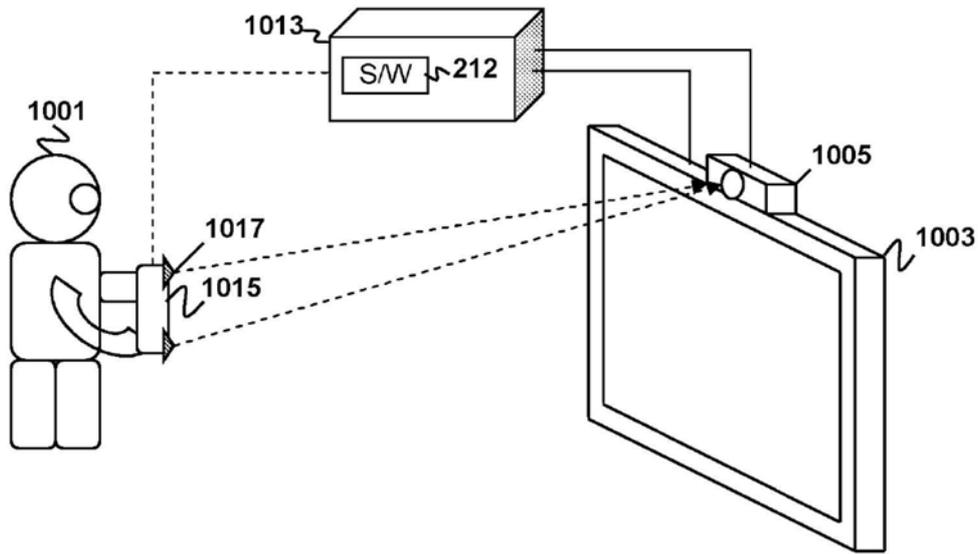


图10C

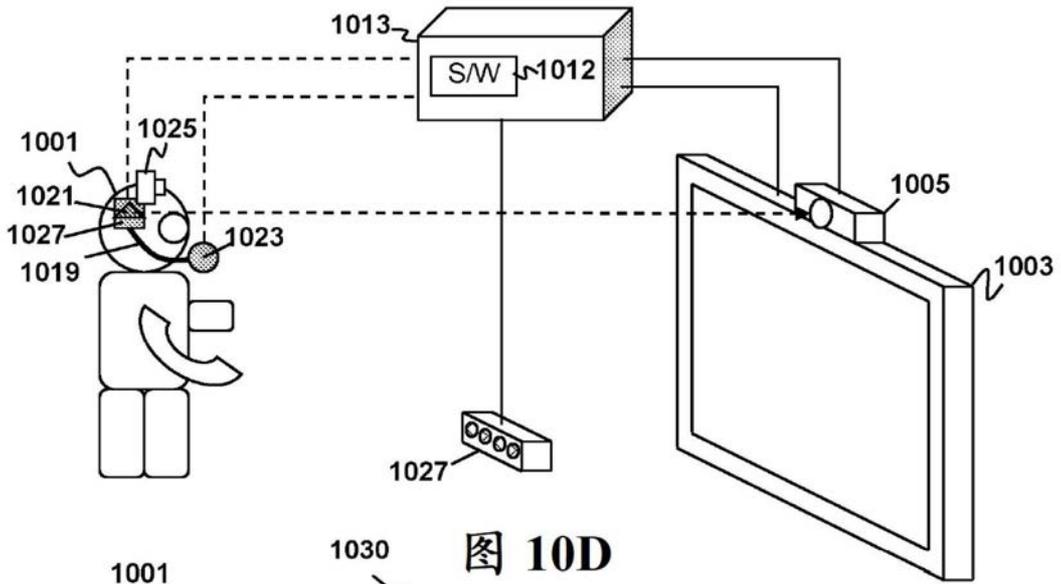


图 10D

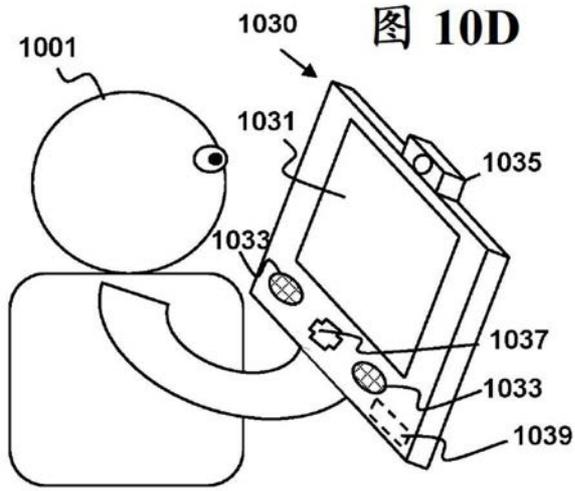


图 10E