



(12) 发明专利申请

(10) 申请公布号 CN 113467089 A

(43) 申请公布日 2021. 10. 01

(21) 申请号 202110679583.5

(22) 申请日 2017.09.12

(30) 优先权数据

62/394,013 2016.09.13 US

62/440,320 2016.12.29 US

(62) 分案原申请数据

201780069300.7 2017.09.12

(71) 申请人 奇跃公司

地址 美国佛罗里达州

(72) 发明人 E·布洛伊 M·J·伍兹

A·拉比诺维奇

(74) 专利代理机构 北京市中咨律师事务所

11247

代理人 于静 牛南辉

(51) Int.Cl.

G02B 27/01 (2006.01)

G09B 21/00 (2006.01)

G06F 3/01 (2006.01)

G06F 16/29 (2019.01)

G06F 40/58 (2020.01)

G06K 9/00 (2006.01)

G06N 3/08 (2006.01)

G06T 19/00 (2011.01)

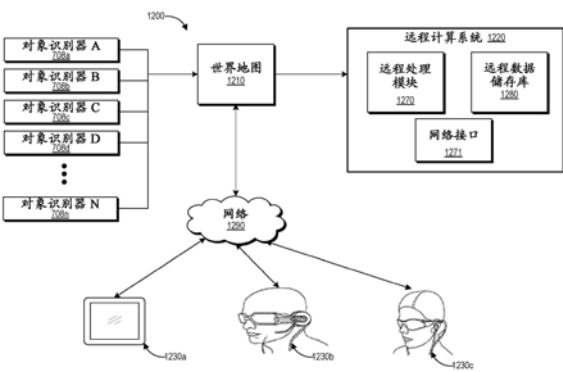
权利要求书13页 说明书47页 附图29页

(54) 发明名称

感测眼镜

(57) 摘要

用于混合现实设备的感测眼镜系统能够促进用户与其他人或者与环境进行交互。作为一个示例,感测眼镜系统能够识别并且解释手语,并且向混合现实设备的用户呈现翻译的信息。可穿戴系统还能够识别用户环境中的文本,修改该文本(例如,通过改变文本的内容或显示特征),以及渲染修改的文本以遮挡原始的文本。



1. 一种用于手语识别的可穿戴系统,所述可穿戴系统包括:
头戴式显示器,被配置为向用户呈现虚拟内容;
成像系统,被配置为对所述用户的环境成像;以及
硬件处理器,与所述头戴式显示器和所述成像系统通信,并被编程为:
接收由所述成像系统捕获的图像;
利用对象识别器检测所述图像中的手势;
识别手语中的所述手势的含义;
基于与所述用户相关联的上下文信息来辨别目标语言;
基于所识别的含义将所述手势翻译为所述目标语言;
至少部分地基于将所述手势翻译为所述目标语言来生成虚拟内容;以及
使所述头戴式显示器将所述虚拟内容渲染给所述用户。
2. 根据权利要求1所述的可穿戴系统,其中,所述成像系统包括被配置为对所述用户的周围成像的一个或多个广角相机。
3. 根据权利要求1所述的可穿戴系统,其中,所述硬件处理器还被编程为访问与所述手势相关联的辅助信息;以及由所述头戴式显示器渲染的所述虚拟内容包括所述辅助信息。
4. 根据权利要求1所述的可穿戴系统,其中,为了基于与所述用户相关联的上下文信息来辨别目标语言,所述硬件处理器被编程为:
基于以下中的至少一个将所述目标语言设置为所述用户理解的语言:所述可穿戴系统捕获的所述用户的语音、所述用户的位置、或者来自所述用户选择所述语言作为所述目标语言的输入。
5. 根据权利要求1所述的可穿戴系统,其中,所述硬件处理器被编程为确定所述目标语言是否是口语;并且响应于确定所述目标语言是口语,以所述目标语言播放与所翻译的手势相关联的语音的音频流。
6. 根据权利要求1所述的可穿戴系统,其中,所述硬件处理器被编程为确定所述目标语言是否是另一种手语,并且响应于确定所述目标语言是另一种手语,将所述另一种手语中另一手势的图形呈现为所述手势的翻译。
7. 根据权利要求1所述的可穿戴系统,其中,为了识别所述手语中的所述手势的含义,所述硬件处理器被编程为在由所述成像系统捕获的图像的一部分上应用深度神经网络技术。
8. 根据权利要求1所述的可穿戴系统,其中,所述硬件处理器还被编程为至少部分地基于所述用户的位置从候选手语的列表中辨别所述手语。
9. 根据权利要求1所述的可穿戴系统,其中,为了基于所识别的含义将所述手势翻译为所述目标语言,所述硬件处理器被编程为将所述手势转换为所述目标语言的文本表达。
10. 根据权利要求1所述的可穿戴系统,其中,所述硬件处理器被编程为确定所检测的手势的源;并且在确定所检测的手势的源是所述可穿戴系统的所述用户时,将以所述目标语言的所述手势的翻译传送到另一用户的可穿戴设备。
11. 根据权利要求1所述的可穿戴系统,其中,所述硬件处理器被编程为检测由所述头戴式显示器从显示器解除所述虚拟内容的条件,并且响应于检测到所述条件,由所述头戴式显示器从所述显示器移除所述虚拟内容。

12. 根据权利要求11所述的可穿戴系统,其中,所述条件包括以下中的至少一个:持续时间、用户的手势或来自用户输入设备的输入。

13. 根据权利要求11所述的可穿戴系统,其中,所述图像包括视频的一帧或多帧。

14. 一种用于手语识别的方法,所述方法包括:

接收由成像系统捕获的图像;

分析所述图像以检测用户的手势;

至少部分地基于所检测的手势来检测手语中的通信的存在;

识别所述手语中的所述手势的含义;

辨别所述手势将被翻译为的目标语言;

基于所识别的含义将所述手势翻译为所述目标语言;

至少部分地基于将所述手势翻译为所述目标语言来生成虚拟内容;以及

使头戴式显示器将所述虚拟内容渲染给所述用户。

15. 根据权利要求14所述的方法,其中,从第一可穿戴设备接收所述图像,所述第一可穿戴设备被配置为呈现混合现实内容而所述虚拟内容被传送到第二可穿戴设备用于渲染,其中,所述第一可穿戴设备和所述第二可穿戴设备被配置向所述用户呈现混合现实内容。

16. 根据权利要求14所述的方法,其中,基于所识别的含义将所述手势翻译为所述目标语言包括将所述手势转换为以所述目标语言的文本表达。

17. 根据权利要求14所述的方法,其中,所述虚拟内容包括以所述目标语言的文本表达或说明所述目标语言的另一个的图形。

18. 根据权利要求14所述的方法,其中,识别所述手语中的所述手势的含义包括:在由所述成像系统捕获的图像的一部分上应用深度神经网络技术。

19. 根据权利要求14所述的方法,其中,检测所述手语中的所述通信的存在包括:

从候选手语的列表中辨别所述手语;以及

确定所检测的手势对应于所述手语的表达。

20. 根据权利要求19所述的方法,其中,确定所检测的手势对应于所述手语的表达包括与做出所述手势的人的嘴唇运动和在所述用户正在做出所述手势时捕获的音频数据相关地分析所述手势。

21. 一种用于手语识别的可穿戴系统,所述可穿戴系统包括:

头戴式显示器,其被配置为向用户呈现虚拟内容;

音频传感器,其被配置为检测来自所述用户的环境的音频;

面向外成像系统,其被配置为对所述用户的所述环境成像;以及

硬件处理器,其与所述头戴式显示器和所述成像系统通信,并被编程为:

接收由所述面向外成像系统所捕获的多个图像;

利用对象识别器检测所述多个图像中的至少一组手;

确定所述至少一组手的相对大小;

基于所述至少一组手的相对大小在多人中辨别所述至少一组手的源;

利用所述对象识别器检测所述多个图像中的所述至少一组手的至少一个手势;

确定所述至少一个手势的所述源属于所述用户以外的人;

响应于所述确定识别手语中的所述至少一个手势的含义;

基于与所述用户相关联的上下文信息来辨别目标语言,其中,所述上下文信息包括:
对所述图像进行分析来获得所述目标语言中字符的存在,或者
对所述音频传感器所获得的音频数据进行分析来获得所述用户的语音;
基于所识别的含义并进一步基于确定所述至少一个手势的源属于所述用户以外的人
而将所述至少一个手势翻译为所述目标语言;
至少部分地基于将所述手势翻译为所述目标语言来生成虚拟内容;以及
使所述头戴式显示器将所述虚拟内容渲染给所述用户。

22. 根据权利要求21所述的可穿戴系统,其中,所述成像系统包括被配置为对所述用户的周围进行成像的一个或多个广角相机。

23. 根据权利要求21所述的可穿戴系统,其中,所述硬件处理器还被编程为访问与所述手势相关联的辅助信息;以及由所述头戴式显示器渲染的所述虚拟内容包括所述辅助信息。

24. 根据权利要求21所述的可穿戴系统,其中,为了基于与所述用户相关联的上下文信息来辨别目标语言,所述硬件处理器被编程为:

基于以下中的至少一个将所述目标语言设置为所述用户理解的语言:所述可穿戴系统捕获的所述用户的语音、所述用户的位置、或者来自所述用户选择所述语言作为所述目标语言的输入。

25. 根据权利要求21所述的可穿戴系统,其中,所述硬件处理器被编程为确定所述目标语言是否是口语;并且响应于确定所述目标语言是口语,以所述目标语言播放与所翻译的手势相关联的语音的音频流。

26. 根据权利要求21所述的可穿戴系统,其中,所述硬件处理器被编程为确定所述目标语言是否是另一种手语,并且响应于确定所述目标语言是另一种手语,将所述另一种手语中另一手势的图形呈现为所述手势的翻译。

27. 根据权利要求21所述的可穿戴系统,其中,为了识别所述手语中的所述手势的所述含义,所述硬件处理器被编程为在由所述成像系统所捕获的图像的至少一部分上应用深度神经网络技术。

28. 根据权利要求21所述的可穿戴系统,其中,所述硬件处理器还被编程为至少部分地基于所述用户的位置从候选手语的列表中辨别所述手语。

29. 根据权利要求21所述的可穿戴系统,其中,为了基于所识别的含义将所述手势翻译为所述目标语言,所述硬件处理器被编程为将所述手势转换为所述目标语言的文本表达。

30. 根据权利要求21所述的可穿戴系统,其中,所述硬件处理器被编程为确定所检测的手势的源;并且在确定所检测的手势的源是所述可穿戴系统的所述用户时,将以所述目标语言的所述手势的翻译传送到另一用户的可穿戴设备。

31. 根据权利要求21所述的可穿戴系统,其中,所述硬件处理器被编程为检测由所述头戴式显示器从显示器解除所述虚拟内容的条件,并且响应于检测到所述条件,由所述头戴式显示器从所述显示器移除所述虚拟内容。

32. 根据权利要求31所述的可穿戴系统,其中,所述条件包括以下中的至少一个:持续时间、用户的手势、或来自用户输入设备的输入。

33. 根据权利要求21所述的可穿戴系统,其中,所述图像包括视频的一帧或多帧。

34. 一种用于手语识别的方法,所述方法包括:

在包括配置为向可穿戴系统的用户呈现虚拟内容的头戴式显示器的所述可穿戴系统的控制下:

接收由所述可穿戴系统的成像系统捕获的图像;

分析所图像以检测人的手势;

确定与所述人的所述手势相关联的一组手的相对大小;

基于所述相对大小辨别所述人的所述手势的源;

确定所述人的所述手势的所述源属于所述用户以外的人;

至少部分地基于所检测的手势来检测手语中的通信的存在;

如果所述人的所述手势不是所述用户的所述手势,则识别所述手语中的所述手势的含义;

辨别所述手势将被翻译为目标语言,其中辨别所述目标语言包括:

分析所述图像以检测所述目标语言中的字符的存在,或者

分析所述可穿戴系统获得的音频数据以检测所述用户的语音;

基于所识别的含义并进一步基于确定所述人的所述手势的所述源属于所述用户以外的人而将所述手势翻译为所述目标语言;

至少部分地基于将所述手势翻译为所述目标语言来生成虚拟内容;以及

使所述头戴式显示器将所述虚拟内容渲染给所述用户。

35. 根据权利要求34所述的方法,其中,所述可穿戴系统还被配置为:将第二虚拟内容传送到第二可穿戴系统以用于所述第二可穿戴系统的渲染。

36. 根据权利要求34所述的方法,其中,基于所识别的含义将所述手势翻译为所述目标语言包括:将所述手势转换为以所述目标语言的文本表达。

37. 根据权利要求34所述的方法,其中,所述虚拟内容包括以所述目标语言的文本表达或说明所述目标语言的另一个的图形。

38. 根据权利要求34所述的方法,其中,识别所述手语中的所述手势的所述含义包括:在由所述成像系统捕获的图像的至少一部分上应用深度神经网络技术。

39. 根据权利要求34所述的方法,其中,检测所述手语中的所述通信的存在包括:

从候选手语的列表中辨别所述手语;以及

确定所检测的手势对应于所述手语的表达。

40. 根据权利要求39所述的方法,其中,确定所检测的手势对应于所述手语的表达包括与做出所述手势的人的嘴唇运动和在所述人正在做出所述手势时捕获的音频数据相关地分析所述手势。

41. 一种计算系统,包括:

硬件计算机处理器;

具有存储在其上的软件指令的非暂时性计算机可读介质,所述软件指令由所述硬件计算机处理器执行以使所述计算系统执行包括以下操作的操作:

经由一个或多个网络与多个可穿戴AR设备进行通信;

接收由所述多个可穿戴AR设备中的第一可穿戴AR设备的面向外成像系统所捕获的图像信息,其中,所述图像信息包括所述第一可穿戴AR设备的第一佩戴者的一只或多只手的

图像;

至少基于对所述第一佩戴者的所述一只或多只手的所述图像的分析,自动检测所接收的图像信息中的手语;

将所检测的手语转换为文本;以及

通过所述一个或多个网络将所转换的文本传输到所述多个可穿戴AR设备中的第二可穿戴AR设备,其中,所述第二可穿戴AR设备被配置为向所述第二可穿戴AR设备的第二佩戴者显示所转换的文本。

42. 根据权利要求41所述的系统,其中,所述软件指令还被配置为使所述计算系统:

通过所述一个或多个网络传输所述第一可穿戴AR设备的所述第一佩戴者的世界地图。

43. 根据权利要求42所述的系统,其中,所述第一佩戴者的所述世界地图包括所述第一可穿戴AR设备的所述第一佩戴者的化身。

44. 根据权利要求41所述的系统,其中,所述软件指令还被配置为使所述计算系统辨别目标语言。

45. 根据权利要求21所述的系统,其中,所述软件指令还被配置为使所述计算系统:

基于以下各项中的至少一项来辨别所述第二AR设备的所述第二佩戴者理解的目标语言:由所述第二AR设备所捕获的语音、所述第二AR设备的位置、或来自所述第二AR设备的所述佩戴者的输入。

46. 根据权利要求45所述的系统,其中,所转换的文本是所述目标语言。

47. 根据权利要求41所述的系统,其中,所述自动检测手语包括:检测所接收的图像信息中的一系列手势。

48. 根据权利要求47所述的系统,其中,所述将所检测的手语转换为文本包括:访问手语字典以将所检测的手势翻译为文本。

49. 根据权利要求41所述的系统,其中,所述将所检测的手语转换为文本包括:使用机器学习算法。

50. 根据权利要求41所述的系统,其中,所述软件指令还被配置为使所述计算系统:

确定所检测的手语的源,所述源包括人,以及

基于所确定的源来转换所检测的手语。

51. 根据权利要求50所述的系统,其中,所述确定所检测的手语的源至少基于所述图像信息中的所述一只或多只手的大小。

52. 根据权利要求21所述的系统,其中,所述软件指令还被配置为使所述计算系统:

基于所转换的文本将音频流传输到所述第二可穿戴AR设备。

53. 一种促进可穿戴AR设备之间通信的方法,所述方法包括:

经由一个或多个网络与多个可穿戴AR设备进行通信;

接收由所述多个可穿戴AR设备中的第一可穿戴AR设备的面向外成像系统所捕获的图像信息,其中,所述图像信息包括所述第一可穿戴AR设备的第一佩戴者的一只或多只手的图像;

至少基于对所述第一佩戴者的所述一只或多只手的所述图像的分析,自动检测所接收的图像信息中的手语;

将所检测的手语转换为文本;以及

通过一个或多个网络将所转换的文本传输到所述多个可穿戴AR设备中的第二可穿戴AR设备,其中,所述第二可穿戴AR设备被配置为向所述第二可穿戴AR设备的第二佩戴者显示所转换的文本。

54. 根据权利要求53所述的方法,包括:通过所述一个或多个网络传输所述第一可穿戴AR设备的所述第一佩戴者的世界地图。

55. 根据权利要求54所述的方法,其中,所述第一佩戴者的所述世界地图包括所述第一可穿戴AR设备的所述第一佩戴者的化身。

56. 根据权利要求53所述的方法,包括:辨别目标语言。

57. 根据权利要求53所述的方法,包括基于以下各项中的至少一项来辨别所述第二AR设备的所述第二佩戴者理解的目标语言:由所述第二AR设备所捕获的语音、所述第二AR设备的位置、或来自所述第二AR设备的所述佩戴者的输入。

58. 根据权利要求57所述的方法,其中,所转换的文本是所述目标语言。

59. 根据权利要求53所述的方法,其中,所述自动检测手语包括:检测所接收的图像信息中的一系列手势。

60. 根据权利要求53所述的方法,其中,所述将所检测的手语转换为文本包括:访问手语字典以将所检测的手势翻译为文本。

61. 根据权利要求53所述的方法,其中,所述将所检测的手语转换为文本包括:使用机器学习算法。

62. 根据权利要求53所述的方法,所述方法包括:

确定所检测的手语的源,所述源包括人,以及

基于所确定的源来转换所检测的手语。

63. 根据权利要求53所述的方法,所述方法包括:基于所述图像信息中的所述一只或多只手的大小来确定所检测的手语的源是所述第一可穿戴AR设备的所述佩戴者。

64. 根据权利要求53所述的方法,所述方法包括:基于所转换的文本将音频流传输到所述第二可穿戴AR设备。

65. 一种被配置为投影增强现实图像内容的头戴式显示设备,所述显示设备包括:

框架,其被配置为佩戴在用户的头部上并被配置为在所述用户的眼睛前方支撑显示器;

一个或多个相机,其被配置为生成光信号;

处理电子器件,其被配置为:

接收来自所述一个或多个相机的所述光信号;

从所述光信号中辨别所述用户的环境中的至少一个物理对象;

确定所述至少一个物理对象是否包括文本;

响应于确定所述至少一个物理对象包括文本,确定要应用于所述文本的修改;

至少基于所确定的修改和所述文本,生成经修改的文本;

生成包括所述经修改的文本的所述至少一个物理对象的虚拟表示;

指示所述显示器在所述用户的所述环境中渲染所述至少一个物理对象的所述虚拟表示以遮挡所述至少一个物理对象。

66. 根据权利要求65所述的头戴式显示设备,其中,所述显示器包括一个或多个光源以

及一个或多个波导堆叠,所述一个或多个波导堆叠被配置为将光引导到所述用户的眼睛以在所述眼睛中形成图像。

67. 根据权利要求66所述的头戴式显示设备,其中,所述一个或多个光源被配置为将光引导到所述波导堆叠中。

68. 根据权利要求66所述的头戴式显示设备,其中,所述一个或多个光源包括光纤扫描投影仪。

69. 根据权利要求65所述的头戴式显示设备,其中,所述一个或多个相机包括一个或多个摄像机。

70. 根据权利要求65所述的头戴式显示设备,其中,所述处理电子器件被配置为使用光学字符识别算法将所述光信号中的一个或多个字母或字符转换为文本。

71. 根据权利要求70所述的头戴式显示设备,其中,所述处理电子器件被配置为访问数据库以辨别用于所述一个或多个字母或字符的所述文本或语言的可能候选。

72. 根据权利要求70所述的头戴式显示设备,其中,所述处理电子器件被配置为接收与以下中的一个或多个相关联的输入:所述用户参与的活动、所述用户的地理位置、所述用户的行进速度、所述用户的海拔高度、由所述显示器所检测的环境噪声的音量或类型、由所述显示器所检测的区域中的可见光或其他光的级别或类型、由所述显示器所检测的温度或气候、所述文本距所述用户的感知距离、或由所述显示器所检测的单词的类别。

73. 根据权利要求65所述的头戴式显示设备,还包括GPS系统。

74. 根据权利要求65所述的头戴式显示设备,其中,所述经修改的文本采用与所述文本的第一字体大小不同的第二字体大小。

75. 根据权利要求65所述的头戴式显示设备,其中,所述经修改的文本比所述文本对所述用户更易读。

76. 根据权利要求65所述的头戴式显示设备,其中,所述经修改的文本包括添加到所述文本的图形元素。

77. 根据权利要求65所述的头戴式显示设备,其中,所述经修改的文本包括具有采用与所述文本的一个或多个字母或字符的第一字体不同的第二字体的所述文本的一个或多个字母或字符的文本。

78. 根据权利要求65所述的头戴式显示设备,其中,所述处理电子器件被配置为相对于所述用户在没有所述头戴式显示器的情况下看到的内容来放大所述文本的一个或多个字母或字符。

79. 根据权利要求65所述的头戴式显示设备,其中,所述虚拟表示包括界定内部区域的边界区域。

80. 根据权利要求79所述的头戴式显示设备,其中,所述处理电子器件被配置为在所述内部区域内显示所述经修改的文本的一个或多个字母或字符。

81. 根据权利要求80所述的头戴式显示设备,其中,所述处理电子器件被配置为在不同于第一背景的第二背景下显示所述文本的所述一个或多个字母或字符,所述用户在没有所述头戴式显示器的情况下将在所述第一背景下阅读所述一个或多个字母或字符。

82. 根据权利要求81所述的头戴式显示设备,其中,所述第二背景包括单色背景。

83. 根据权利要求82所述的头戴式显示设备,其中,所述单色背景包括白色。

84. 根据权利要求81所述的头戴式显示设备, 其中, 所述第一背景包括所述用户在没有所述头戴式显示器的情况下将看到的内容。

85. 根据权利要求65所述的头戴式显示设备, 其中, 所述文本适于由文本编辑器编辑。

86. 根据权利要求65所述的头戴式显示设备, 其中, 如果所述文本的所述一个或多个字母或字符看起来比第一距离阈值更近, 则在看起来比第二深度更靠近所述用户的第一深度处显示所述虚拟表示。

87. 根据权利要求65所述的头戴式显示设备, 其中, 如果所述一个或多个字母或字符看起来比第二距离阈值更远, 则在看起来比第一深度更远离所述用户的第二深度处显示所述虚拟表示。

88. 根据权利要求87所述的头戴式显示设备, 其中, 如果所述一个或多个字母或字符看起来比第一距离阈值更远且比所述第二距离阈值更近, 则在看起来比所述第一深度更远离所述用户但比所述第二深度更靠近所述用户的第三深度处显示所述文本。

89. 一种用于通过增强现实系统提供从手语转换而来的文本的方法, 所述方法包括: 在包括成像系统的增强现实AR系统的控制下:

通过所述成像系统捕获图像信息;

检测所述图像信息中的手势, 所述手势为用于手语识别的候选;

在所检测的手势中识别手语;

将所识别的手语转换为文本; 以及

显示所转换的文本。

90. 根据权利要求89所述的方法, 还包括:

接收对所转换的文本的辅助信息的请求;

检索与所请求的所转换的文本相关联的辅助信息;

使用所述AR系统显示所述辅助信息;

检测用于解除显示所转换的文本或所述辅助信息的条件; 以及

解除显示所转换的文本或所述辅助信息。

91. 根据权利要求90所述的方法, 其中, 用于解除显示所转换的文本或所述辅助信息的所述条件基于用户界面交互。

92. 根据权利要求91所述的方法, 其中, 所述用户界面交互至少部分地基于所述AR系统的用户的眼睛运动。

93. 根据权利要求90所述的方法, 其中, 用于解除显示所转换的文本或所述辅助信息的所述条件至少部分地基于持续时间。

94. 根据权利要求90所述的方法, 其中, 用于解除显示所转换的文本或所述辅助信息的所述条件至少部分地基于附加手语手势的转换或附加辅助信息的接收。

95. 根据权利要求90所述的方法, 还包括: 检测用于重新显示被解除显示的所转换的文本或所述辅助信息的条件; 重新显示被解除显示的所转换的文本或所述辅助信息。

96. 根据权利要求89所述的方法, 其中, 将所识别的手语转换为文本包括: 应用深度学习技术。

97. 根据权利要求96所述的方法, 其中, 所述深度学习技术包括神经网络。

98. 根据权利要求89所述的方法, 其中, 所述AR系统在手语识别和文本转换中使用手语

字典。

99. 根据权利要求98所述的方法, 其中, 所述AR系统识别对于所述AR系统的用户是国外的手语。

100. 根据权利要求99所述的方法, 其中, 所述AR系统通过处理候选手语的列表来识别所述手语, 所述列表至少部分地基于所述AR系统的位置来被优先化。

101. 根据权利要求99所述的方法, 其中, 所述AR系统通过处理候选手语的列表来识别所述手语, 所述列表至少部分地基于在所述AR系统的环境中所检测的口语来被优先化。

102. 一种用于翻译手语的增强现实AR装置, 包括:

AR显示器;

成像系统;

数据存储器, 其配置为存储计算机可执行指令和数据; 以及

与数据存储器通信的处理器, 其中, 所述计算机可执行指令在被执行时使所述处理器:

接收由所述成像系统所捕获的图像信息;

检测所接收的图像或视频信息中的手势;

识别所检测的手势中的手语;

将所识别的手语翻译为所述AR装置的用户理解的语言; 以及

使用所述AR显示器来显示与所翻译的手语相关联的信息。

103. 根据权利要求102所述的装置, 其中, 所述计算机可执行指令在被执行时还使所述处理器: 接收对关于所翻译的手语的辅助信息的请求; 检索与所请求的手语相关的辅助信息; 以及使用所述AR显示器来显示所检索的辅助信息。

104. 根据权利要求102所述的装置, 其中, 所述处理器通过将所接收的图像信息通过通信网络发送到远程处理器来检测所述手势并识别所述手语, 以便所述远程处理器检测所述手势并识别所述手语。

105. 根据权利要求102所述的装置, 其中, 所述成像系统包括多个相机或广角相机。

106. 根据权利要求102所述的装置, 其中, 所述处理器还被配置为: 确定所检测的手势的源; 并且在确定所检测的手势的源是所述AR装置的所述用户时, 将所翻译的手语传输到另一设备进行显示。

107. 根据权利要求102所述的装置, 还包括音频放大器, 并且所述处理器还被编程为通过所述音频放大器以音频呈现所翻译的手语。

108. 根据权利要求107所述的装置, 其中, 所述处理器还被配置为通过所述音频放大器以音频呈现所述辅助信息。

109. 根据权利要求102所述的装置, 其中, 所述AR装置的所述用户理解的所述语言包括与所识别的手语不同的手语。

110. 一种用于促进涉及一种或多种手语的远程通信的增强现实AR系统, 包括:

多个可穿戴AR设备, 每个可穿戴AR设备包括:

AR显示器;

成像系统; 以及

用于通过通信网络进行通信的通信系统;

一个或多个数据存储器, 其被配置为存储计算机可执行指令和数据; 以及

与所述数据存储器通信的一个或多个处理器,其中,所述计算机可执行指令在被执行时将所述一个或多个处理器配置为:

接收由所述多个可穿戴AR设备中的第一可穿戴AR设备的所述成像系统所捕获的图像信息;

检测所接收的图像信息中的手语;将所检测的手语转换为文本;

通过所述通信网络将所转换的文本传输到所述多个可穿戴AR设备中的第二可穿戴AR设备;以及

在所述第二可穿戴AR设备的所述AR显示器上显示所转换的文本。

111. 根据权利要求110所述的系统,其中,所述第二可穿戴AR设备还显示第一用户的世界地图。

112. 根据权利要求111所述的系统,其中,所述第一用户的所述世界地图包括所述第一用户的化身。

113. 根据权利要求110所述的系统,其中,所述多个可穿戴AR设备中的每一个包括一个或多个数据存储器以及一个或多个处理器,并且所述处理器功能由本地处理器执行。

114. 一种被配置为投影增强现实图像内容的头戴式显示设备,所述显示设备包括:

框架,其被配置为佩戴在所述用户的头部上并被配置为在所述用户的眼睛前方支撑显示器;

一个或多个相机,其被配置为接收光信号;以及

处理电子器件,其被配置为:

接收来自所述一个或多个相机的信号;

从所述信号中辨别图像;

确定所述图像是否包括文本;

将所述文本转换为经修改的文本;以及

指示所述显示器渲染所述经修改的文本。

115. 根据权利要求114所述的头戴式显示设备,其中,所述显示器包括一个或多个光源以及一个或多个波导堆叠,所述一个或多个波导堆叠被配置为将光引导到所述用户的眼睛以在所述眼睛中形成图像。

116. 根据权利要求115所述的头戴式显示设备,其中,所述一个或多个光源被配置为将光引导到所述波导堆叠中。

117. 根据权利要求116所述的头戴式显示设备,其中,所述一个或多个光源包括光纤扫描投影仪。

118. 根据权利要求114所述的头戴式显示设备,其中,所述一个或多个相机包括一个或多个摄像机。

119. 根据权利要求114所述的头戴式显示设备,其中,所述处理电子器件被配置为使用光学字符识别算法将所述图像中的一个或多个字母或字符转换为文本。

120. 根据权利要求119所述的头戴式显示设备,其中,所述处理电子器件被配置为访问数据库以辨别用于所述一个或多个字母或字符的所述文本或语言的可能候选。

121. 根据权利要求120所述的头戴式显示设备,其中,所述处理电子器件被配置为接收与以下中的一个或多个相关联的输入:所述用户参与的活动、所述用户的地理位置、所述用

户的行进速度、所述用户的海拔高度、由所述显示器所检测的环境噪声的音量或类型、由所述显示器所检测的区域中的可见光或其他光的级别或类型、由所述显示器所检测的温度或气候、所述文本距所述用户的感知距离、或由所述显示器所检测的单词的类别。

122. 根据权利要求114所述的头戴式显示设备,还包括GPS系统。

123. 根据权利要求122所述的头戴式显示设备,其中,所述经修改的文本采用与所述文本的第一字体大小不同的第二字体大小。所述第二字体大小可以大于所述第一字体大小。

124. 根据权利要求123所述的头戴式显示设备,其中,所述经修改的文本比所述文本对所述用户更易读。

125. 根据权利要求124所述的头戴式显示设备,其中,所述处理电子器件被配置为将图形元素部分地添加到所述文本以形成所述经修改的文本。

126. 根据权利要求114所述的头戴式显示设备,其中,所述处理电子器件被配置为以与所述一个或多个字母或字符的第一字体不同的第二字体显示所述文本的所述一个或多个字母或字符。

127. 根据权利要求114所述的头戴式显示设备,其中,所述处理电子器件被配置为相对于所述用户在没有所述头戴式显示器的情况下看到的内容来放大所述文本的所述一个或多个字母或字符。

128. 根据权利要求114所述的头戴式显示设备,其中,所述处理电子器件被配置为显示边界区域,所述边界区域界定内部区域。

129. 根据权利要求128所述的头戴式显示设备,其中,所述处理电子器件被配置为在所述内部区域内显示所述一个或多个字母或字符。

130. 根据权利要求114所述的头戴式显示设备,其中,所述处理电子器件被配置为在与第一背景不同的第二背景下显示所述文本的所述一个或多个字母或字符,所述用户在没有所述头戴式显示器的情况下将在所述第一背景下阅读所述一个或多个字母或字符。

131. 根据权利要求130所述的头戴式显示设备,其中,所述第二背景包括单色背景。

132. 根据权利要求131所述的头戴式显示设备,其中,所述单色背景包括白色。

133. 根据权利要求130所述的头戴式显示设备,其中,所述第一背景包括所述用户在没有所述头戴式显示器的情况下将看到的内容。

134. 根据权利要求114所述的头戴式显示设备,其中,所述文本适于由文本编辑器编辑。

135. 一种用于使用头戴式显示器投影增强现实图像内容的方法,所述方法包括:在硬件处理器的控制下:

接收来自一个或多个相机的光信号;

使用光学字符识别模块,从所述信号中辨别图像;确定所述图像是否包括一个或多个字母或字符;

将所述一个或多个字母或字符转换为文本;以及

在所述头戴式显示器上显示所述文本,其中,显示所述文本包括:通过光学透射目镜将光作为图像传输到所述用户。

136. 根据权利要求135所述的方法,还包括:将光引导到所述用户的眼睛以在所述眼睛中形成图像。

137. 根据权利要求135所述的方法,还包括:使用光纤扫描投影仪将光引导到所述目镜。

138. 根据权利要求135所述的方法,其中,使用光学字符识别模块包括:辨别用于所述一个或多个字母或字符的所述文本或语言的可能候选。

139. 根据权利要求135所述的方法,其中,使用光学字符识别模块包括接收包括与以下中的一个或多个相关联的信息的输入:所述用户参与的活动、所述用户的地理位置、所述用户的行进速度、所述用户的海拔高度、由所述显示器所检测的环境噪声的音量或类型、由所述显示器所检测的区域中的可见光或其他光的级别或类型、由所述显示器所检测的温度或气候、所述一个或多个字母或字符距所述用户的感知距离、或由所述显示器所检测的单词的类别。

140. 根据权利要求135所述的方法,其中,将所述一个或多个字母或字符转换成文本包括:以不同于与所述一个或多个字母或字符相关联的第一语言的第二语言来显示所述文本。

141. 根据权利要求140所述的方法,包括:将所述文本翻译成所述第二语言。

142. 根据权利要求135所述的方法,还包括:从GPS系统接收位置数据。

143. 根据权利要求135所述的方法,其中,在所述头戴式显示器上显示所述一个或多个字母或字符包括:以与所述一个或多个字母或字符的第一字体大小不同的第二字体大小显示所述一个或多个字母或字符。

144. 根据权利要求135所述的方法,其中,在所述头戴式显示器上显示所述一个或多个字母或字符包括:与没有所述头戴式显示器的情况相比,向所述用户更易读地显示所述一个或多个字母或字符。

145. 根据权利要求135所述的方法,其中,在所述头戴式显示器上显示所述一个或多个字母或字符包括:以大于所述用户在没有头戴式显示器的情况下出现的字体大小的字体大小显示所述一个或多个字母或字符。

146. 根据权利要求135所述的方法,其中,在所述头戴式显示器上显示所述一个或多个字母或字符包括:以与所述一个或多个字母或字符的第一字体不同的第二字体显示所述一个或多个字母或字符。

147. 根据权利要求135所述的方法,其中,在所述头戴式显示器上显示所述一个或多个字母或字符包括:相对于所述用户在没有所述头戴式显示器的情况下将看到的内容来放大所述一个或多个字母或字符。

148. 根据权利要求135所述的方法,其中,在所述头戴式显示器上显示所述一个或多个字母或字符包括显示边界区域,所述边界区域界定内部区域。

149. 根据权利要求148的所述方法,其中,在所述头戴式显示器上显示所述一个或多个字母或字符包括:在所述内部区域内显示所述一个或多个字母或字符。

150. 根据权利要求135所述的方法,其中,在所述头戴式显示器上显示所述一个或多个字母或字符包括:在与第一背景不同的第二背景下显示所述一个或多个字母或字符,所述用户在没有所述头戴式显示器的情况下将在所述第一背景下阅读所述一个或多个字母或字符。

151. 根据权利要求150所述的方法,其中,所述第二背景包括单色背景。

152. 根据权利要求151所述的方法, 其中, 所述单色背景包括白色。

153. 根据权利要求150所述的方法, 其中, 所述第一背景包括所述用户在没有所述头戴式显示器的情况下将看到的内容。

154. 根据权利要求135所述的方法, 其中, 所述文本适于由文本编辑器编辑。

155. 根据权利要求135所述的方法, 其中, 将所述一个或多个字母或字符转换成文本包括: 接收来自用户的输入。

156. 根据权利要求155所述的方法, 其中, 接收来自用户的输入包括接收以下中的一项或多项: 口头命令、手势、头部的运动、或用户的一只或多只眼睛的运动。

157. 根据权利要求135所述的方法, 其中, 如果所述一个或多个字母或字符看起来比第一距离阈值更近, 则在看起来比第二深度更近的第一深度处显示所述文本。

158. 根据权利要求135所述的方法, 其中, 如果所述一个或多个字母或字符看起来比第二距离阈值更远, 则在看起来比第一深度更远的第二深度处显示所述文本。

159. 根据权利要求157所述的方法, 其中, 如果所述一个或多个字母或字符看起来比第一距离阈值更远且比所述第二距离阈值更近, 则在看起来比所述第一深度更远但比所述第二深度更近的第三深度处显示所述文本。

160. 根据权利要求157所述的方法, 其中, 所述第一距离阈值是80厘米。

161. 根据权利要求157所述的方法, 其中, 所述第二距离阈值是600厘米。

162. 根据权利要求157所述的方法, 其中, 所述第二距离阈值与所述第一距离阈值之间的差小于100米。

163. 一种增强现实系统, 包括: 面向外成像系统; 非暂时性存储器, 其被配置为存储由所述面向外成像系统所获取的图像; 以及硬件处理器, 其被编程为: 接收所述面向外成像系统所获取的所述增强现实系统的用户的环境的图像; 分析所述图像以辨别所述用户的环境中的标示; 识别所述标示上的文本; 将所述文本的至少一部分转换为目标语言; 并指示显示器将所转换的文本渲染给所述用户。

164. 根据权利要求163所述的增强现实系统, 其中, 所述硬件处理器被编程为修改与所述文本相关联的显示特征。

165. 根据权利要求163所述的增强现实系统, 其中, 为了将所述文本的至少一部分转换为目标语言, 所述硬件处理器被编程为: 辨别所述标示上的所述文本的语言并将所述语言转换为所述目标语言。

166. 根据权利要求163所述的增强现实系统, 其中, 所述硬件处理器被编程为: 至少部分地基于所述用户的位置来确定所述目标语言。

167. 根据权利要求163所述的增强现实系统, 其中, 为了识别所述标示上的文本, 所述硬件处理器被编程为: 识别所述目标语言中的文本。

168. 根据权利要求167所述的增强现实系统, 其中, 所述硬件处理器被编程为: 不转换所述目标语言中的所述文本。

感测眼镜

[0001] 本申请是申请号为201780069300.7的中国专利申请“感测眼镜”(申请日为2017年9月12日)的分案申请。

[0002] 相关申请的交叉引用

[0003] 本申请根据35U.S.C.§119(e) 要求2016年9月13日提交的题为“SENSORY EYEWEAR (感测眼镜)”的美国临时申请No.62/394,013和2016年12月29日提交的题为“Systems and Methods for Augmented Reality (用于增强现实的系统和方法)”的美国临时申请No.62/440,320的优先权的权益,上述申请的公开内容通过引用整体并入在此。

技术领域

[0004] 本公开涉及虚拟现实和增强现实成像以及视觉系统,尤其涉及识别环境中的手语或文本并且基于所识别的手语或文本来渲染虚拟内容。

背景技术

[0005] 现代计算和显示技术促进了用于所谓的“虚拟现实”、“增强现实”或“混合现实”体验的系统的开发,其中数字再现的图像或其部分以其看起来似乎被认为是或可能被感知是真实的方式呈现给用户。虚拟现实或“VR”场景通常涉及数字或虚拟图像信息的呈现,而不透明于其他实际的真实世界的视觉输入;增强现实或“AR”场景通常涉及将数字或虚拟图像信息呈现为用户周围的现实世界的可视化的增强;混合现实或“MR”,涉及真实世界和虚拟世界合并在一起,以生成物理对象和虚拟对象共存并实时交互的新环境。事实证明,人类的视觉感知系统非常复杂,开发促进虚拟图像元素在其它虚拟或现实世界图像元素中的舒适的、感觉自然的、丰富的呈现的VR、AR或MR技术是有挑战性的。本文公开的系统和方法解决了与VR、AR和MR技术有关的各种挑战。

发明内容

[0006] 公开了用于识别环境中的手语和文本的混合现实系统的各种实施例。这些实施例可以有利地允许残疾的人之间进行更好的交互。

[0007] 用于混合现实设备的感测眼镜系统能够促进用户与其他人或者与环境进行交互。作为一个示例,感测眼镜系统能够识别并且解释手语,并且向混合现实设备的用户呈现翻译的信息。可穿戴系统还能够识别用户环境中的文本,修改文本(例如,通过变更文本的内容或显示特征),以及渲染经修改的文本以遮挡原始的文本。

[0008] 在附图和以下描述中阐述本说明书中描述的主题的一个或多个实施方式的细节。从说明书、附图以及权利要求将清楚其他特征、方面和优势。该概述和以下详细的描述都不旨在限定或限制本发明主题的范围。

附图说明

[0009] 图1描绘了具有特定虚拟现实对象以及由某人观看的某些物理对象的混合现实场

景的图示。

[0010] 图2A示意性示出了能够实施感测眼镜系统的可穿戴系统的示例。

[0011] 图2B示意性示出了可穿戴系统的各种示例部件。

[0012] 图3示意性示出了使用多个深度平面来模拟三维图像的方法的方面。

[0013] 图4示意性示出了用于向用户输出图像信息的波导堆叠的示例。

[0014] 图5示出了可以由波导输出的示例出射光束。

[0015] 图6是示出包括波导装置、将光光学地耦合到波导装置或从波导装置光学地耦合光的光耦合器子系统、以及控制子系统的光学系统的示意图,该光学系统用于生成多焦点体积显示器、图像或光领域。

[0016] 图7是可穿戴系统的示例的框图。

[0017] 图8是用于渲染与识别出的对象相关的虚拟内容的方法的示例的流程图。

[0018] 图9是包括感测眼镜系统的可穿戴系统的另一示例的框图。

[0019] 图10是用于确定到可穿戴系统的用户输入的方法的示例的流程图。

[0020] 图11是用于与虚拟用户界面进行交互的方法的示例的流程图。

[0021] 图12示意性地示出了描绘彼此交互的多个可穿戴系统的整体系统视图。

[0022] 图13A示出了感测眼镜系统的示例用户体验,其中感测眼镜系统能够解释手语(例如,由示意者用手势做出)。

[0023] 图13B示出了感测眼镜系统的另一示例用户体验,其中呈现了目标语音和辅助信息。

[0024] 图13C示出了远程呈现会话中的感测眼镜系统的示例用户体验。

[0025] 图13D示出了用于解释手语的示例虚拟用户界面。

[0026] 图14A和图14B示出了用于利用感测眼镜系统来促进人际通信的示例过程。

[0027] 图14C是用于确定辅助信息并呈现与转换的文本相关联的辅助信息的示例方法的流程图。

[0028] 图15示出了用于利用感测眼镜系统来促进人际通信的另一示例过程。

[0029] 图16A至图16E示出了感测眼镜系统的示例用户体验,该感测眼镜系统被配置为识别环境中的文本,修改文本的显示特征,以及渲染经修改的文本。

[0030] 图17示出了用于促进用户与环境的交互的感测眼镜的示例过程。

[0031] 图18示出了通过修改标示的内容来帮助用户理解物理环境中的标示的示例。

[0032] 图19示出了帮助用户理解物理环境中的标示的示例过程。

[0033] 在整个附图中,可以重新使用附图标记来指示参考元素之间的对应关系。提供附图以示出在此描述的示例实施例并且不旨在限制本公开的范围。

具体实施方式

[0034] 概述

[0035] 被配置为呈现AR/VR/MR内容的可穿戴系统能够实现感测眼镜系统以增强用户与其他人或环境的交互。示例可穿戴系统可以包括头戴式显示器、各种成像传感器和一个或多个硬件处理器。显示器可以是佩戴在一只眼睛或多只眼睛前方的透视显示器。

[0036] 为了增强用户与其他人的交互体验,可穿戴系统可以被配置为捕获和解释手语。

手语主要使用视觉手势(例如,手形;手部方位;手、手臂或身体动作;或面部表情)来进行交流。世界各地使用了数百种手语。某些手语可能比其他手语使用得更频繁。例如,美国手语(ASL)在美国和加拿大广泛使用。

[0037] 许多人不知道任何手语。言语或听力有障碍的人和对话伙伴可能不熟悉相同的手语。这可能妨碍与听力有障碍的人或言语有障碍的人的对话。因此,可以对会话伙伴作出的标志(sign)(例如,手势)进行成像,将标志转换为文本或图形(例如,系统用户理解的手语中手语手势的图形),然后显示与标志相关联的信息(例如,将标志翻译成用户理解的语言)的可穿戴系统将极大地有助于改善用户与会话伙伴之间的通信。此外,可以期望的是具有一种可穿戴系统,该可穿戴系统能够实时地(或接近实时地)提供手语的文本或图形转换,其中对于可穿戴系统的用户具有最小程度的注意力分散,并且可穿戴系统的用户的费力程度微不足道。

[0038] 本公开在可穿戴设备的情境下公开了这种期望系统的示例。可穿戴设备可以包括头戴式部件(例如,头戴式显示器)。这样的设备可以允许用户可视地接收信息,该信息以信息能够在正常可视的现实世界的旁边(或在其上面)同时被观看的方式由计算设备提供。这样的系统可以用于显示可以在传统计算机屏幕上显示的任何形式的信息,例如字符、图像效果、文本、图形或任何类型的视频。

[0039] 这里描述的可穿戴系统可以组合手语识别(SLR)和可穿戴设备的显示能力,以基于检测到的手语向用户提供信息。例如,可穿戴设备上的面向外相机可以对正在做出的手势进行成像,识别手势中的标志,将标志翻译成用户理解的语言,并向用户显示翻译。可以通过可穿戴系统向用户显示检测到的手语的抄本(例如,标题或文本气泡)。机器学习算法(例如,深度神经网络)可以接收图像并执行标志的识别和翻译。当用户提示时,可以显示抄本中单词的含义或来自适当来源的相关信息。可穿戴系统可以提供的辅助信息的种类可以与例如互联网上的大量可用信息资源一样无限制。

[0040] 除了增强用户与其他人的交互体验之外或作为其替代,感测眼镜系统还可以改善用户对环境体验。作为改善用户与环境的交互的示例,实现感测眼镜系统的可穿戴系统可以识别环境中的文本(例如,例如商业或公共显示标志的标示上的文本),修改文本的显示特征(例如,通过增加文本的大小)或修改文本的内容(例如,通过将文本翻译成另一种语言),并将修改的文本渲染在环境中的物理文本上。

[0041] 如本文进一步描述的,可穿戴系统可以接收用户环境的图像。可以通过可穿戴设备的面向外成像系统或与可穿戴设备相关联的图腾(totem)来获取图像。可穿戴系统可以确定图像是否包括一个或多个字母或字符,并将一个或多个字母或字符转换为文本。可穿戴系统可以使用各种技术确定图像是否包括字母或字符,例如机器学习算法或光学字符识别(OCR)算法。可穿戴系统可以使用对象识别器(例如,在图7中描述的)来识别字母和字符并将它们转换为文本。

[0042] 在特定的实施例中,可以为用户显示文本,而不是用户在没有可穿戴设备的情况下看到文本。例如,可穿戴系统可以使头戴式显示器以与关联于原始图像的字母或字符相关联的字体大小不同的字体大小显示文本。可穿戴系统还可以提高文本的显示质量。例如,在没有可穿戴系统的情况下,例如雾、霾、雨、强光、弱光、弱光或字母与周围图像之间的颜色对比等的各种环境因素,可以妨碍用户清晰地看到环境中的文本。可穿戴系统可以呈现

将增加文本的清晰度的标志(例如,具有增加的对比度或更大的字体)。

[0043] 可穿戴系统还可以将文本(例如,标示上的文本)从其原始语言翻译成目标语言。例如,可以将文本从用户不理解的语言翻译成用户理解的语言。翻译的文本可以渲染在原始文本上,使得用户可以容易地以用户能够理解的语言观看该文本。

[0044] 可穿戴系统的3D显示的示例

[0045] 可穿戴系统(在此也称为增强现实(AR)系统)可以被配置为向用户呈现2D或3D虚拟图像。图像可以是静止图像、视频的帧或视频、以组合的形式等。可穿戴系统的至少一部分可以实现在可以单独或组合地呈现VR、AR或MR环境用于用户交互的可穿戴设备上。可穿戴设备可以是头戴式设备(HMD),其能够互换地作为AR设备(ARD)来使用。此外,为了本公开的目的,术语“AR”能够与术语“MR”互换使用。

[0046] 图1描绘了具有某些虚拟现实对象以及由人看到的某些物理对象的混合现实场景的图示。在图1中,描绘了MR场景100,其中MR技术的用户看到以背景中的人、树、建筑为特征的真实世界的公园状的设置110以及实体平台120。除了这些项目,MR技术的用户还感觉他“看到”站在真实世界的平台120上的机器人雕像130,以及看起来像飞行的蜜蜂的化身的卡通式的头像角色140,尽管这些元素不存在于真实世界中。

[0047] 为了使3D显示器产生真实的深度感觉,并且更具体地,模拟的表面深度感觉,可以期望显示器的视场中的每个点产生与其虚拟深度对应的适应响应。如果对显示点的适应响应不对应于该点的虚拟深度(由汇聚和立体视觉的双眼深度线索确定),则人眼可能经历适应冲突,导致成像不稳定、有害的眼部紧张、头痛,并且在没有适应信息的情况下,几乎完全缺乏表面深度。

[0048] VR、AR和MR体验可以通过具有显示器的显示系统来提供,其中与多个深度平面对应的图像被提供给观看者。对于每个深度平面,图像可以是不同的(例如,提供场景或对象的略微不同的呈现),并且可以由观看者的眼睛单独聚焦,从而有助于基于眼睛所需的适应向用户提供深度线索,为位于不同深度平面上的场景或基于观察不同深度平面上的不同图像特征失焦而聚焦不同图像特征。如在此其它地方所讨论的,这样的深度线索提供了可靠的深度感知。

[0049] 图2A示出可以配置为提供AR/VR/MR场景的可穿戴系统200的示例。可穿戴系统200还可以被称为AR系统200。可穿戴系统200包括显示器220,以及支持显示器220的功能的各种机械和电子模块和系统。显示器220可以耦合到可由用户、穿戴者或观看者210穿戴的框架230。显示器220可以定位在用户210的眼睛的前方。显示器220可以向用户呈现AR/VR/MR内容。显示器220可以包括穿戴在用户头部上的头戴式显示器。在一些实施例中,扬声器240耦合到框架230并且定位成邻近用户的耳道(在一些实施例中,未示出的另一个扬声器定位成邻近用户的另一个耳道以提供立体/可塑造声音控制)。显示器220可以包括用于从环境检测音频流并且捕获周围环境的声音的音频传感器(例如,麦克风)232。在一些实施例中,设置未示出的一个或多个其他的音频传感器来提供立体声接收。立体声接收可以用于确定声源的位置。可穿戴系统200可以对音频流执行话音或语音识别。

[0050] 可穿戴系统200可以包括观察用户周围环境中世界的面向外的成像系统464(在图4中示出)。可穿戴系统200还可以包括可以跟踪用户的眼睛运动的面向内的成像系统462(在图4中示出)。面向内的成像系统可以跟踪一只眼睛的运动或两只眼睛的运动。面向内的

成像系统462可以附接到框架230并且可以与处理模块260和/或270电通信,处理模块260或270可以处理由面向内的成像系统获取的图像信息以确定例如瞳孔直径或眼睛的方位、用户210的眼睛移动或眼睛姿势。

[0051] 作为示例,可穿戴系统200可以使用面向外的成像系统464或面向内的成像系统462来获取用户的姿势的图像。图像可以是静止图像、视频的帧或视频。

[0052] 显示器220诸如通过有线导线或无线连接可操作地耦合250到本地数据处理模块260,该本地数据处理模块260可以以各种配置安装,诸如固定地附接到框架230,固定地附接到由用户穿戴的头盔或帽子,嵌入到耳机中,或以其它方式可移除地附接到用户210(例如,以背包方式的配置,以带式耦合方式的配置)。

[0053] 本地处理和数据模块260可以包括硬件处理器以及诸如非易失性存储器(例如,闪存存储器)的数字存储器,二者都可用于辅助数据的处理、缓冲以及存储。数据可以包括如下数据:a)从传感器(其可以例如可操作地耦合到框架230或以其它方式附接到用户210)捕获的数据,例如图像捕获设备(例如,面向内的成像系统或面向外的成像系统中的相机)、音频传感器(例如,麦克风)、惯性测量单元(IMU)、加速度计、罗盘、全球定位系统(GPS)单元、无线电设备或陀螺仪;或b)使用远程处理模块270或远程数据储存库280获取或处理的数据,可能在这样的处理或检索之后传递给显示器220。本地处理和数据模块260可以通过通信链路262或264(诸如经由有线或无线通信链路)可操作地耦合到远程处理模块270或远程数据储存库280,使得这些远程模块作为资源可用于本地处理和数据模块260。另外,远程处理模块280和远程数据储存库280可以相互可操作地耦合。

[0054] 在一些实施例中,远程处理模块270可以包括被配置为分析和处理数据或图像信息的一个或多个处理器。在一些实施例中,远程数据储存库280可以包括数字数据存储设施,其可以通过互联网或其它网络配置以“云”资源配置而可用。在一些实施例中,在本地处理和数据模块中存储全部数据,并且执行全部计算,允许从远程模块完全自主使用。

[0055] 图2B示出可以包括显示器220和框架230的可穿戴系统200。放大视图202示意性地示出了可穿戴系统200的各种部件。在某些实施方式中,图2B中所示的部件中的一个或多个部件可以是显示器220的一部分。单独或组合的各种部件可以收集与可穿戴系统200的用户或用户的环境相关联的各种数据(例如,音频或视觉数据)。应当理解,取决于使用可穿戴系统的用途,其他实施例可以具有附加的或更少的部件。然而,图2B提供了各种部件中的一些部件和可以通过可穿戴系统收集、分析和存储的数据类型的基本思想。

[0056] 图2B示出了可以包括显示器220的示例可穿戴系统200。显示器220可以包括显示透镜106,该显示透镜106可以安装到用户的头部或与框架230对应的壳体或框架108。显示透镜106可以包括由壳体108定位在用户的眼睛302、304的前面的一个或多个透明镜子,并且可以被配置为将投射的光38反弹(bounce)到眼睛302、304中并且有助于光束整形,同时还允许来自本地环境的至少一些光的透射。投射光束38的波前可以弯曲或聚焦以与投射光的期望焦距一致。如图所示,两个宽视场机器视觉相机16(也称为世界相机)可以耦接到壳体108以对用户周围的环境成像。这些相机16可以是双捕获可见光/不可见(例如,红外)光相机。相机16可以是图4所示的面向外成像系统464的一部分。由世界相机16获取的图像可以由姿势处理器36处理。例如,姿势处理器36可以实现一个或多个对象识别器708(例如,如图7所示)以识别用户或用户环境中的另一个人的姿势或者识别用户环境中的物理对象。

[0057] 继续参考图2B,示出了一对扫描激光整形波前(例如,用于深度)光投射器模块,其具有配置成将光38投射到眼睛302、304中的显示镜子和光学器件。所描绘的视图还示出了两个微型红外相机24,其与红外光源26(例如发光二极管“LED”)配对,其被配置为能够跟踪用户的眼睛302、304以支持渲染和用户输入。相机24可以是图4所示的面向内成像系统462的一部分。可穿戴系统200还可以具有传感器组件39,其可以包括X、Y和Z轴加速度计能力以及磁罗盘,和X、Y和Z轴陀螺仪能力,优选地以诸如200Hz的相对高的频率来提供数据。传感器组件39可以是参考图2A描述的IMU的一部分。所描绘的系统200还可以包括头部姿势处理器36,例如ASIC(专用集成电路)、FPGA(现场可编程门阵列)或ARM处理器(高级精简指令集机器),其可以被配置根据从捕获设备16输出的宽视场图像信息来计算实时或接近实时的用户头部姿势。头部姿势处理器36可以是硬件处理器,并且可以实现为图2A中所示的本地处理和数据模块260的一部分。

[0058] 还示出了处理器32,其被配置为执行数字或模拟处理以从来自传感器组件39的陀螺仪、罗盘或加速度计数据导出姿势。处理器32可以是图2A中所示的本地处理和数据模块260的一部分。如图2B所示的可穿戴系统200还可以包括例如GPS 37(全球定位系统)的位置系统,以辅助姿势和定位分析。此外,GPS还可以提供关于用户环境的基于远程(例如,基于云端)的信息。该信息可以用于识别用户环境中的对象或信息。

[0059] 可穿戴系统可以将由GPS 37和远程计算系统(例如,远程处理模块270、另一用户的ARD等)获取的数据进行组合,这能够提供关于用户环境的更多信息。作为一个示例,可穿戴系统可以基于GPS数据确定用户的位置,并且检索包括与用户的位置相关联的虚拟对象的世界地图(例如,通过与远程处理模块270通信)。作为另一示例,可穿戴系统200可以使用世界相机16(其可以是图4中所示的面向外成像系统464的一部分)来监控环境。基于由世界相机16获取的图像,可穿戴系统200可以检测环境中的字符(例如,通过使用图7中所示的一个或多个对象识别器708)。可穿戴系统可以进一步使用由GPS 37获取的数据来解释字符。例如,可穿戴系统200可以识别字符所在的地理区域,并识别与该地理区域相关联的一种或多种语言。因此,可穿戴系统可以基于所识别的语言来解释字符,例如,基于与所识别的语言相关联的句法、语法、句子结构、拼写、标点标志等。在一个示例中,德国的用户210可以在高速公路进行驾驶时感知交通标志。可穿戴系统200可以基于从GPS 37获取的数据(单独或与世界相机16获取的图像组合)识别用户210在德国并且来自所成像的交通标志的文本可能是德语。

[0060] 在一些情况下,由世界相机16获取的图像可以包括用户环境中的对象的不完整信息。例如,由于朦胧的气氛、文本中的瑕疵或错误、低照度、模糊图像、遮挡、世界相机16的有限FOV,图像可能包括不完整的文本(例如,句子、字母或短语)。可穿戴系统200可以使用由GPS 37获取的数据作为识别图像中的文本的上下文线索。

[0061] 可穿戴系统200还可以包括渲染引擎34,其可以被配置为提供对用户来说是本地的渲染信息,以便于扫描仪的操作并成像到用户的眼睛中,以便用户对世界的观看。渲染引擎34可以由硬件处理器(例如,中央处理单元或图形处理单元)实现。在一些实施例中,渲染引擎是本地处理和数据模块260的一部分。渲染引擎34可以通信地耦接(例如,经由有线或无线链路)到可穿戴系统200的其他部件。例如,渲染引擎34可以经由通信链路102耦接到眼睛相机24,并且经由通信链路104耦接到投影子系统18(其可以经由扫描激光装置以类似于

视网膜扫描显示的方式将光投射到用户的眼睛302、304中)。渲染引擎34还可以与其他处理单元通信,诸如分别经由链路105和94与例如传感器姿势处理器32和图像姿势处理器36进行通信。

[0062] 相机24(例如,迷你红外相机)可以用于跟踪眼睛姿势以支持渲染和用户输入。一些示例性眼睛姿势可以包括用户正在观看的位置或者他或她正在聚焦的深度(可以用眼睛聚散来估计)。GPS 37、陀螺仪、罗盘和加速度计39可以用于提供粗略或快速姿势估计。相机16中的一个或多个可以获取图像和姿势,这些图像和姿势与来自相关联的云计算资源的数据一起可以用于映射本地环境并与其他人共享用户视图。

[0063] 图2B中描绘的示例部件仅用于说明目的。为了便于说明和描述,将多个传感器和其他功能模块一起示出。一些实施例可以仅包括这些传感器或模块中的一个或子集。此外,这些部件的位置不限于图2B中所示的位置。一些部件可以安装到或容纳在其他部件内,例如安装或容纳在皮带上的部件、手持部件或头盔部件。作为一个示例,图像姿势处理器36、传感器姿势处理器32和渲染引擎34可以定位在腰包中并且被配置为与可穿戴系统的其他部件通信,经由例如超宽带、Wi-Fi、蓝牙等的无线通信或经由有线通信。所描绘的壳体108优选地是由用户头部可安装的并且可穿戴。然而,可穿戴系统200的一些部件可以佩戴到用户身体的其他部分。例如,扬声器240可以插入用户的耳朵中以向用户提供声音。

[0064] 关于将光38投射到用户的眼睛302、304中,在一些实施例中,可以利用相机24来测量用户眼睛302、304的中心在几何上延伸(verge)到的位置,这位置通常与眼睛302、304的焦点位置或“焦点深度”一致。眼睛延伸到的所有点的三维表面可以被称为“双眼单视界(horopter)”。焦距可以采用有限数量的深度,或者可以无限变化。来自聚散距离投射的光看起来聚焦到对象眼睛302、304,而聚散距离前面或后面的光被模糊。在美国专利公开No.2016/0270656中也描述了本公开的可穿戴设备和其他的显示系统的示例,其全部内容通过引用并入本文。

[0065] 人类视觉系统复杂,并且提供深度的现实感知是具挑战性的。对象的观看者可能由于聚散度和调节的组合而将该对象感知为“三维”。两只眼睛彼此相对的聚散运动(例如,光瞳孔彼此相向或远离的滚动运动,以会聚眼睛的视线来注视对象)与眼睛晶状体的聚焦(或“调节”)密切相关。在正常情况下,改变眼睛晶状体的焦点或调节眼睛,以将焦点从在不同距离处的一个对象改变到另一个对象,将会在称为“调节聚散度反射(accommodation-vergence reflex)”的关系下自动地导致在聚散度上的匹配改变达到相同的距离。同样,在正常情况下,聚散度的改变将引发调节的匹配改变。提供调节和聚散度之间的更好匹配的显示系统可以形成更逼真且舒适的三维图像模拟。

[0066] 另外,无论眼睛在何处聚焦,人眼都可以正确地分辨出光束直径小于约0.7毫米的空间相干光。因此,为了产生适当焦深的错觉,可以用相机24跟踪眼睛聚散,并且可以利用渲染引擎34和投影子系统18来渲染在聚焦的双眼单视界上或靠近聚焦的双眼单视界的所有对象,以及以不同程度散焦的所有其他对象(例如,使用有意创建的模糊)。优选地,系统220以大约每秒60帧或更高的帧速率向用户渲染。如上所述,优选地,相机24可以用于眼睛跟踪,并且软件可以被配置为不仅拾取聚散几何形状而且还拾取聚焦位置提示以用作用户输入。优选地,这种显示系统配置有适合白天或夜晚使用的亮度和对比度。

[0067] 在一些实施例中,显示系统优选地具有小于约20毫秒的视觉对象对准的延迟,小

于约0.1度的角度对准,以及约1弧分的分辨率,其不受理论限制,被认为大约是人眼的极限。显示系统220可以与定位系统集成,定位系统可以涉及GPS元件、光学跟踪、罗盘、加速度计或其他数据源,以辅助位置和姿势确定;可以利用定位信息来促进用户对相关世界的视图中的准确渲染(例如,这样的信息将有助于眼镜知道它们相对于现实世界的位置)。

[0068] 在一些实施例中,可穿戴系统200被配置为基于用户眼睛的调节来显示一个或多个虚拟图像。与迫使用户聚焦在投影图像的位置处的先前3D显示方法不同,在一些实施例中,可穿戴系统被配置为自动改变投影的虚拟内容的焦点以允许更舒适地观看呈现给用户的一个或多个图像。例如,如果用户的眼睛具有1米的当前焦点,则可以投影图像以与用户的焦点一致。如果用户将焦点移至3米,则投影图像与新焦点一致。因此,一些实施例的可穿戴系统200不是强迫用户达到预定焦点,而是允许用户的眼睛以更自然的方式发挥功能。

[0069] 这样的可穿戴系统200可以消除或减少眼睛疲劳、头痛和通常相对于虚拟现实设备观察到的其他生理症状的发生率。为了实现这一点,可穿戴系统200的各种实施例被配置为通过一个或多个可变焦距元件(VFE)以变化的焦距投影虚拟图像。在一个或多个实施例中,可以通过多平面聚焦系统来实现3D感知,该多平面聚焦系统在远离用户的固定焦平面处投影图像。其他实施例采用可变平面聚焦,其中焦平面在z方向上来回移动以与用户的当前聚焦状态一致。

[0070] 在多平面聚焦系统和可变平面聚焦系统中,可穿戴系统200可采用眼睛跟踪来确定用户眼睛的聚散度,确定用户的当前焦点,并在所确定的焦点处投影虚拟图像。在其他实施例中,可穿戴系统200包括光调制器,该光调制器通过光纤扫描仪或其他光产生源可变地投射跨越视网膜的光栅图案中的变化焦点的光束。因此,可穿戴系统200的显示器以变化的焦距投射图像的能力不仅使得用户能够容易地调节以观看3D的对象,而且还可以用于补偿用户的眼睛异常,如美国专利公开No.2016/0270656中进一步描述的,其全部内容通过引用并入本文。在一些其他实施例中,空间光调制器可以通过各种光学部件将图像投影到用户。例如,如下面进一步描述的,空间光调制器可以将图像投影到一个或多个波导上,然后波导将图像发送给用户。

[0071] 图3示出了使用多个深度平面来模拟三维图像的方法的方面。参考图3,在z轴上距眼睛302和眼睛304的不同距离处的对象由眼睛302和眼睛304适应,以使得那些对象在焦点中。眼睛302和眼睛304呈现特定的适应状态,以使沿着z轴的不同距离处的对象进入焦点。因此,可以说特定的适应状态与深度平面306中的特定一个深度平面相关联,该特定深度平面具有相关联的焦距,以使得当眼睛处于该深度平面的适应状态时,特定深度平面中的对象或对象的部分被聚焦。在一些实施例中,可以通过为眼睛302和304中的每一只眼睛提供图像的不同呈现来模拟三维图像,并且还通过提供与深度平面中每一个深度平面对应的图像的不同呈现来模拟三维图像。尽管为了清楚说明而示出为分离的,但应理解的是,例如,随着沿着z轴的距离增加,眼睛302和眼睛304的视场可能重叠。另外,虽然为了便于说明而示出为平坦的,但应理解的是,深度平面的轮廓可以在物理空间中是弯曲的,使得深度平面中的所有特征在特定的适应状态下与眼睛对焦。不受理论的限制,可以相信的是,人类眼睛通常可以解释有限数量的深度平面以提供深度感知。因此,通过向眼睛提供与这些有限数量的深度平面中的每一个深度平面对应的图像的不同呈现,可以实现感知深度的高度可信的模拟。

[0072] 波导堆叠组件

[0073] 图4示出了用于向用户输出图像信息的波导堆叠的示例。可穿戴系统400包括可以用于采用多个波导432b、434b、436b、438b、440b向眼睛/大脑提供三维感知的波导堆叠或堆叠波导组件480。在一些实施例中,可穿戴系统400对应于图2A的可穿戴系统200,图4更详细地示意性地示出了该可穿戴系统200的一些部分。例如,在一些实施例中,波导组件480可以被集成到图2A的显示器220中。

[0074] 继续参考图4,波导组件480还可以包括在波导之间的多个特征458、456、454、452。在一些实施例中,特征458、456、454、452可以是透镜。在其它实施例中,特征458、456、454、452可以不是透镜。而是它们可以简单地是间隔物(例如,用于形成空气间隙的包层或结构)。

[0075] 波导432b、434b、436b、438b、440b或多个透镜458、456、454、452可以被配置为以各种级别的波前曲率或光线发散向眼睛发送图像信息。每个波导级别可以与特定的深度平面相关联,并且可以被配置为输出与该深度平面对应的图像信息。图像注入装置420、422、424、426、428可用于将图像信息注入到波导440b、438b、436b、434b、432b中,其中的每一个波导可以被配置为分配入射光穿过每一个相应的波导,用于向眼睛410输出。光从图像注入装置420、422、424、426、428的输出表面出射并被注入到波导440b、438b、436b、434b、432b的相应输入边缘。在一些实施例中,可以将单个光束(例如,准直光束)注入到每一个波导中,以便与特定波导相关联的深度平面对应的特定角度(和发散量)输出朝向眼睛410定向的克隆准直光束的整个视场。

[0076] 在一些实施例中,图像注入装置420、422、424、426、428是分立显示器,每个显示器产生用于分别注入到相应波导440b、438b、436b、434b、432b中的图像信息。在一些其它实施例中,图像注入装置420、422、424、426、428是单个复用显示器的输出端,其可以例如经由一个或多个光导管(诸如,光纤线缆)向图像注入装置420、422、424、426、428中的每一个图像注入装置输送图像信息。

[0077] 控制器460控制堆叠波导组件480和图像注入装置420、422、424、426、428的操作。控制器460包括调节图像信息到波导440b、438b、436b、434b、432b的定时和提供的编程(例如,在非暂时性计算机可读介质中的指令)。在一些实施例中,控制器460可以是单个整体装置,或通过有线或无线通信通道连接的分布式系统。在一些实施例中,控制器460可以是处理模块260或270(图2A所示)的部分。

[0078] 波导440b、438b、436b、434b、432b可以被配置为通过全内反射(TIR)在每一个相应的波导内传播光。波导440b、438b、436b、434b、432b可以各自是平面的或具有其它形状(例如,弯曲),具有主要的顶表面和底表面以及在这些主要的顶表面和底表面之间延伸的边缘。在所示的配置中,波导440b、438b、436b、434b、432b可以各自包括光提取光学元件440a、438a、436a、434a、432a,这些光提取光学元件被配置为通过将每一个相应波导内传播的光重定向而将光提取到波导外,以向眼睛410输出图像信息。提取的光也可以被称为外耦合的光,并且光提取光学元件也可以被称为外耦合光学元件。提取的光束在波导中传播的光照射光重定向元件的位置处被波导输出。光提取光学元件(440a, 438a, 436a, 434a, 432a)可以例如是反射或衍射光学特征。虽然为了便于描述和清晰绘图起见而将其图示设置在波导440b、438b、436b、434b、432b的底部主表面处,但是在一些实施例中,光提取光学元件440a、

438a、436a、434a、432a可以设置在顶部或底部主表面处，或可以直接设置在波导440b、438b、436b、434b、432b的体积中。在一些实施例中，光提取光学元件440a、438a、436a、434a、432a可以形成在附接到透明基板材料层中以形成波导440b、438b、436b、434b、432b。在一些其它实施例中，波导440b、438b、436b、434b、432b可以是单片材料，并且光提取光学元件440a、438a、436a、434a、432a可以形成在那片材料的表面上或那片材料的内部中。

[0079] 继续参考图4，如在此所讨论的，每一个波导440b、438b、436b、434b、432b被配置为输出光以形成与特定深度平面对应的图像。例如，最接近眼睛的波导432b可以被配置为将如注入到这种波导432b中的准直光传送到眼睛410。准直光可以代表光学无限远焦平面。下一个上行波导434b可以被配置为将穿过第一透镜452（例如，负透镜）的准直光在其可以到达眼睛410之前发出。第一透镜452可以被配置为产生轻微凸面的波前曲率，使得眼睛/大脑将来自下一个上行波导434b的光解释为来自第一焦平面，该第一焦平面从光学无限远处更靠近向内朝向眼睛410。类似地，第三上波导436b将输出光在到达眼睛410之前穿过第一透镜452和第二透镜454。第一透镜452和第二透镜454的组合光功率可被配置为产生另一增量的波前曲率，以使得眼睛/大脑将来自第三波导436b的光解释为来自第二焦平面，该第二焦平面从光学无穷远比来自下一个上行波导434b的光更靠近向内朝向人。

[0080] 其它波导层（例如，波导438b、440b）和透镜（例如，透镜456、458）被类似地配置，其中堆叠中的最高波导440b通过它与眼睛之间的全部透镜发送其输出，用于代表最靠近人的焦平面的聚合（aggregate）焦度。当在堆叠波导组件480的另一侧上观看/解释来自世界470的光时，为了补偿透镜458、456、454、452的堆叠，补偿透镜层430可以设置在堆叠的顶部处以补偿下面的透镜堆叠458、456、454、452的聚合焦度。这种配置提供了与可用波导/透镜配对一样多的感知焦平面。波导的光提取光学元件和透镜的聚焦方面可以是静态的（例如，不是动态的或电激活的）。在一些替代实施例中，两者之一或者两者都可以是使用电激活特征而动态的。

[0081] 继续参考图4，光提取光学元件440a、438a、436a、434a、432a可以被配置为将光重定向到它们相应的波导之外并且针对与波导相关联的特定深度平面输出具有适当的发散量或准直量的该光。结果，具有不同相关联深度平面的波导可具有不同配置的光提取光学元件，其取决于相关联的深度平面输出具有不同发散量的光。在一些实施例中，如在此所讨论的，光提取光学元件440a、438a、436a、434a、432a可以是体积或表面特征，其可以被配置为以特定角度输出光。例如，光提取光学元件440a、438a、436a、434a、432a可以是体积全息图、表面全息图或衍射光栅。在2015年6月25日公开的美国专利公开No. 2015/0178939中描述了诸如衍射光栅的光提取光学元件，其通过引用全部并入在此。

[0082] 在一些实施例中，光提取光学元件440a、438a、436a、434a、432a是形成衍射图案或“衍射光学元件”（在此也称为“DOE”）的衍射特征。优选地，DOE具有相对较低的衍射效率，以使得仅光束的一部分通过DOE的每一个交点偏转向眼睛410，而其余部分经由全内反射继续移动通过波导。携带图像信息的光因此可以被分成多个相关的出射光束，该出射光束在多个位置处离开波导，并且该结果对于在波导内反弹的该特定准直光束是朝向眼睛304的相当均匀图案的出射发射。

[0083] 在一些实施例中，一个或多个DOE可以在它们主动地衍射的“开”状态和它们不显著衍射的“关”状态之间可切换。例如，可切换的DOE可以包括聚合物分散液晶层，其中微滴

在主体介质中包含衍射图案,并且微滴的折射率可以切换为基本上匹配主体材料的折射率(在这种情况下,图案不明显地衍射入射光),或者微滴可以切换为与主体介质的指数不匹配的指数(在这种情况下,该图案主动地衍射入射光)。

[0084] 在一些实施例中,深度平面的数量和分布或景深可以基于观看者的眼睛的瞳孔大小或方位而动态地改变。景深可与观看者的瞳孔大小成反比。因此,随着观看者眼睛瞳孔的大小减小,景深增加,使得由于平面的位置超出了眼睛的聚焦深度而不可辨别的该平面可能变得可辨别,并且随着瞳孔大小的减小和景深的相应增加表现为更聚焦。类似地,用于向观看者呈现不同图像的间隔开的深度平面的数量可随着瞳孔大小减小而减小。例如,观看者在不调整眼睛远离一个深度平面和到另一个深度平面的适应性的情况下,可能不能清楚地感知处于一个瞳孔大小的第一深度平面和第二深度平面的细节。然而,这两个深度平面可以在不改变适应性的情况下,对于处于另一瞳孔大小的用户同时充分地聚焦。

[0085] 在一些实施例中,显示系统可以改变接收图像信息的波导的数量,基于瞳孔大小或方位的确定或者基于接收特定瞳孔大小或方位的电信号指示。例如,如果用户的眼睛不能区分与两个波导相关联的两个深度平面,则控制器460(其可以是本地处理和数据模块260的实施例)可以被配置或编程为停止向这些波导中的一个提供图像信息。有利地,这可以减轻系统的处理负担,从而增加系统的响应性。在其中波导的DOE可在开启和关闭状态之间切换的实施例中,当波导确实接收图像信息时,DOE可切换到关闭状态。

[0086] 在一些实施例中,可能期望的是出射光束符合直径小于观看者眼睛的直径的条件。然而,考虑到观看者的瞳孔大小的可变性,满足这种条件可能是具有挑战性的。在一些实施例中,通过响应于观看者的瞳孔大小的确定而改变出射光束的大小,该条件在宽范围的瞳孔大小上满足。例如,随着瞳孔大小减小,出射光束的大小也可以减小。在一些实施例中,可以使用可变光圈来改变出射光束大小。

[0087] 可穿戴系统400可以包括对世界470的一部分成像的面向外的成像系统464(例如,数字相机)。世界470的这部分可以被称为世界相机的视场(FOV),并且成像系统464有时被称为FOV相机。世界相机的FOV可以与观看者210的FOV相同或者可以与其不相同,其中观看者210的FOV包含观看者210在给定时间感知到的世界470的一部分。例如,在一些情况下,世界相机的FOV可以比可穿戴系统400的观看者210的更大。可供观看者观看或成像的整个区域可被称为能视场(field of regard (FOR))。由于佩戴者可以移动他的身体、头部或眼睛以感知基本上空间中的任何方向,因此FOR可以包括围绕可穿戴系统400的立体角的 4π 球面度。在其他语境下,佩戴者的移动可能更受限制,并且因此佩戴者的FOR可以对向较小的立体角。从面向外的成像系统464获得的图像可以用于跟踪用户做出的姿势(例如,手或手指姿势),检测用户前方的世界470中的对象等等。

[0088] 可穿戴系统400可以包括例如麦克风的音频传感器232以捕获周围环境的声音。如上所述,在一些实施例中,可以设置一个或多个另外的音频传感器以提供对于语音源的位置确定有用的立体声接收。音频传感器232可以包括作为另一示例的定向麦克风,该定向麦克风还可以提供与音频源所处的位置有关的有用定向信息。

[0089] 可穿戴系统400还可以包括面向内的成像系统466(例如,数字相机),其观察用户的运动,诸如眼睛运动和面部运动。面向内的成像系统466可以用于捕获眼睛410的图像以确定眼睛304的瞳孔的大小或方位。面向内的成像系统466可以用于获得图像,这些图像用

于确定用户正在观看的方向(例如,眼睛姿势)或用于用户的生物识别(例如,经由虹膜识别)。在一些实施例中,每只眼睛可以利用至少一个相机,以独立地分别确定每只眼睛的瞳孔大小或眼睛姿势,从而允许向每只眼睛呈现图像信息动态地适合该眼睛。在一些其它实施例中,只确定并假定单个眼睛410的瞳孔直径或方位(例如,每对眼睛仅使用单个相机)与用户的两只眼睛类似。可以分析由面向内的成像系统466获得的图像以确定用户的眼睛姿势或情绪,其可以由可穿戴系统400用来决定应该向用户呈现哪些音频或视觉内容。可穿戴系统400还可以使用诸如IMU、加速度计、陀螺仪等的传感器来确定头部姿势(例如,头部位置或头部方位)。

[0090] 可穿戴系统400可以包括用户输入设备466,用户可以通过该用户输入设备466向控制器460输入命令以与可穿戴系统400交互。例如,用户输入设备466可以包括触控板、触摸屏、操纵杆、多自由度(DOF)控制器、电容感测设备、游戏控制器、键盘、鼠标、方向板(D-pad)、棒、触觉设备、图腾(例如,用作虚拟用户输入设备)等等。多DOF控制器可以感测在控制器的一些或所有可能的平移(例如,左/右、前/后、或者上/下)或旋转(例如,偏航、俯仰或滚动)中的用户输入。支持平移运动的多DOF控制器可以被称为3DOF,而支持平移和旋转的多DOF可以被称为6DOF。在一些情况下,用户可以使用手指(例如,拇指)在触敏输入设备上按压或滑动以向可穿戴系统400提供输入(例如,向由可穿戴系统400提供的用户界面提供用户输入)。用户输入设备466可在使用可穿戴系统400期间由用户的手握持。用户输入设备466可以与可穿戴系统400进行有线或无线通信。

[0091] 图5示出了由波导输出的出射光束的示例。示出了一个波导,但是应该理解的是,波导组件480中的其它波导可以类似地起作用,其中波导组件480包括多个波导。光520在波导432b的输入边缘432c处被注入到波导432b中,并且通过TIR在波导432b内传播。在光520撞击在DOE 432a上的点处,一部分光如出射光束510离开波导。出射光束510被示出为基本上平行,但是取决于与波导432b相关联的深度平面,该出射光束510也可以以一定角度(例如,形成发散的出射光束)被重定向以传播到眼睛410。应该理解的是,基本上平行的出射光束可以指示具有光提取光学元件的波导,其中光提取光学元件将光外耦合以形成看起来被设置在距眼睛410较大距离(例如,光学无穷远)处的深度平面上的图像。其它波导或者其它光提取光学元件组可以输出更加发散的出射光束图案,这将需要眼睛410适应更近距离以将其聚焦在视网膜上并且将被大脑解释为光来自比光学无穷远更接近眼睛410的距离。

[0092] 图6是示出包括波导装置、将光光耦合到波导装置或从波导装置光耦合光的光耦合器子系统、以及控制子系统的光学系统的示意图,该光学系统用于生成多焦点立体显示器、图像或光领域。光学系统可以包括波导装置、将光光耦合到波导装置或从波导装置光耦合光的光耦合器子系统、以及控制子系统。光学系统可用于生成多焦点立体、图像或光场。该光学系统可以包括一个或多个主平面波导632a(在图6中仅示出一个)以及与至少一些主波导632a中的每一个主波导相关联的一个或多个DOE 632b。平面波导632b可以类似于参考图4讨论的波导432b、434b、436b、438b、440b。光学系统可以使用分布波导装置沿着第一轴(图6所示的垂直轴或Y轴)中继光,并且沿着第一轴(例如,Y轴)扩展光的有效出射光瞳。分布波导装置可以例如包括分布平面波导622b和与分布平面波导622b相关联的至少一个DOE 622a(由双点划线示出)。分布平面波导622b在至少一些方面可以与主平面波导632b相似或相同,但具有与其不同的方位。类似地,至少一个DOE 622a在至少一些方面可以与DOE 632a

相似或相同。例如,分布平面波导622b或DOE 622a可以分别由与主平面波导632b或DOE 632a相同的材料构成。图6所示的光学显示系统600的实施例可以集成到图2A所示的可穿戴系统200中。

[0093] 中继的和出射光瞳扩展的光可以从分布波导装置被光耦合到一个或多个主平面波导632b中。主平面波导632b可以沿着优选地与第一轴正交的第二轴(例如,图6的视图中的水平轴或X轴)中继光。值得注意的是,第二轴可以是与第一轴非正交的轴。主平面波导632b沿着该第二轴(例如,X轴)扩展光的有效出射光瞳。例如,分布平面波导622b可以沿着垂直轴或Y轴中继和扩展光,并且将该光传递到可以沿着水平轴或X轴中继和扩展光的主平面波导632b。

[0094] 光学系统可以包括一个或多个彩色光源(例如,红色、绿色和蓝色激光)610,这些彩色光源可以光耦合到单模光纤640的近端中。可以穿过压电材料的中空管642来通过或接收光纤640的远端。远端作为非固定柔性悬臂644从管642突出。压电管642可以与四个象限电极(未示出)相关联。例如,电极可以镀在管642的外侧、外表面或外周或直径上。芯电极(未示出)也可以位于管642的芯、中心、内周或内径中。

[0095] 例如经由导线660电耦合的驱动电子器件650驱动相对的电极对独立地在两个轴上弯曲压电管642。光纤644的突出远端顶端具有机械谐振模式。谐振的频率可以取决于光纤644的直径、长度和材料特性。通过在光纤悬臂644的第一机械谐振模式附近振动压电管642,可以使得光纤悬臂644振动,并且可以扫过大的偏转。

[0096] 通过激发两个轴上的谐振,光纤悬臂644的顶端在遍及二维(2D)扫描的区域中双轴扫描。通过与光纤悬臂644的扫描同步地调制一个或多个光源610的强度,从光纤悬臂644出射的光可以形成图像。美国专利公开No.2014/0003762中提供了这样的设置的描述,其通过引用全部并入在此。

[0097] 光学耦合器子系统的部件可以准直从扫描光纤悬臂644出射的光。准直光可以由镜面648反射到包含至少一个衍射光学元件(DOE)622a的窄分布平面波导622b中。准直光通过TIR可以沿分布平面波导622b垂直地(相对于图6的视图)传播,并且与DOE 622a重复相交。DOE 622a优选具有低衍射效率。这可以导致一部分光(例如,10%)在与DOE 622a的每个交点处被衍射朝向较大的主平面波导632b的边缘,并且一部分光通过TIR在其原始轨迹上向下分布平面波导622b的长度而继续。

[0098] 在与DOE 622a的每个交点处,附加光可以被衍射向主波导632b的入口。通过将入射光分成多个外耦合组,光的出射光瞳可以在分布平面波导622b中由DOE 622a垂直地扩展。从分布平面波导622b外耦合的该垂直扩展的光可以进入主平面波导632b的边缘。

[0099] 进入主波导632b的光可以经由TIR沿着主波导632b水平传播(相对于图6的视图)。由于光通过TIR沿着主波导632b的至少一部分长度水平传播,因此光在多个点处与DOE 632a相交。DOE 632a可以有利地被设计或构造成具有相位轮廓,该相位轮廓是线性衍射图案和径向对称衍射图案的总和,以产生光的偏转和聚焦。DOE 632a可以有利地具有低衍射效率(例如,10%),使得DOE 632a的每个交点只有一部分光束的光朝着视图的眼睛偏转,而其余的光经由TIR通过主波导632b继续传播。

[0100] 在传播光和DOE 632a之间的每个交点处,一部分光朝着主波导632b的相邻面衍射,从而允许光脱离TIR,并且从主波导632b的面出射。在一些实施例中,DOE 632a的径向对

称衍射图案另外向衍射光赋予聚焦水平,既整形单个光束的光波前(例如,赋予曲率)以及以与设计的聚焦水平相匹配的角度将光束转向。

[0101] 因此,这些不同的路径可以通过多个DOE 632a以不同的角度、聚焦水平或在出射光瞳处产生不同的填充图案来使光耦合到主平面波导632b外。出射光瞳处的不同填充图案可以有利地用于创建具有多个深度平面的光场显示。波导组件中的每一层或堆叠中的一组层(例如3层)可用于产生相应的颜色(例如,红色、蓝色、绿色)。因此,例如,可以采用第一组的三个相邻层在第一焦深处分别产生红光、蓝光和绿光。可以采用第二组的三个相邻层在第二焦深处分别产生红光、蓝光和绿光。可以采用多组来产生具有各种焦深的全3D或4D彩色图像光场。

[0102] 可穿戴系统的其他组件

[0103] 在许多实施方式中,可穿戴系统可以包括除了上述可穿戴系统的部件之外或者替代的其它部件。例如,可穿戴系统可以包括一个或多个触觉设备或部件。触觉设备或部件可以可操作以向用户提供触觉感觉。例如,触觉设备或部件可以在接触虚拟内容(例如,虚拟对象、虚拟工具、其它虚拟构造)时提供压力或纹理的触觉感觉。触觉感觉可以复制虚拟对象表示的物理对象的感觉,或者可以复制虚拟内容表示的想象对象或角色(例如,龙)的感觉。在一些实施方式中,触觉设备或部件可以由用户穿戴(例如,用户可穿戴手套)。在一些实施方式中,触觉设备或部件可以由用户持有。

[0104] 例如,可穿戴系统可以包括能由用户操纵以允许输入或与可穿戴系统交互的一个或多个物理对象。这些物理对象在这里可以被称为图腾(totem)。一些图腾可采取无生命对象的形式,例如一块金属或塑料、墙、桌子的表面。在某些实施方式中,图腾可能实际上不具有任何物理输入结构(例如,键、触发器、操纵杆、轨迹球、摇杆开关)。相反,图腾可以简单地提供物理表面,并且可穿戴系统可以渲染用户界面,以使用户看起来在图腾的一个或多个表面上。例如,可穿戴系统可以渲染计算机键盘和触控板的图像从而看起来驻留在图腾的一个或多个表面上。例如,可穿戴系统可以渲染虚拟计算机键盘和虚拟触控板从而看起来在用作图腾的薄铝矩形板的表面上。矩形板本身没有任何物理键或触控板或传感器。然而,可穿戴系统可以检测用户操纵或交互或者触摸矩形板,作为经由虚拟键盘或虚拟触控板进行的选择或输入。用户输入设备466(图4中所示)可以是图腾的实施例,其可以包括触控板、触摸板、触发器、操纵杆、轨迹球、摇杆或虚拟开关、鼠标、键盘、多自由度控制器,或者另一个物理输入设备。用户可以单独或与姿势一起使用图腾来与可穿戴系统或其它用户进行交互。

[0105] 在美国专利公开No.2015/0016777中描述了本公开的与可穿戴设备、HMD和显示系统一起使用的触觉设备和图腾的示例,其全部内容通过引用并入本文。

[0106] 示例可穿戴系统、环境和接口

[0107] 可穿戴系统可以采用各种映射(mapping)相关技术以便在渲染的光场中实现高景深。在映射出虚拟世界时,了解现实世界中的所有特征和点以准确描绘与现实世界相关的虚拟对象是有利的。为此,从可穿戴系统的用户捕获的FOV图像可以通过包括新图片被添加到世界模型,这些新图片传达现实世界的各个点和特征的信息。例如,可穿戴系统可以收集一组地图点(例如2D点或3D点)并找到新的地图点以渲染更精确版本的世界模型。可以将第一用户的世界模型(例如,通过诸如云网络的网络)传送给第二用户,使得第二用户可以体

验第一用户周围的世界。

[0108] 图7是MR环境700的示例的框图。MR环境700可以被配置为从一个或多个用户可穿戴系统(例如,可穿戴系统200或显示系统220)或静止房间系统(例如,房间相机等)接收输入(例如,来自用户的可穿戴系统的视觉输入702,诸如房间相机的静止输入704,来自各种传感器的感测输入706,来自用户输入设备466的用户输入、手势、图腾、眼睛跟踪等)。可穿戴系统可以使用各种传感器(例如,加速度计、陀螺仪、温度传感器、运动传感器、深度传感器、GPS传感器、面向内的成像系统、面向外的成像系统等)来确定用户的环境的位置以及各种其它属性。该信息可以进一步利用来自房间中的静止相机的信息进行补充,这些信息可以从不同的角度提供图像或各种线索。由相机(诸如房间相机或面向外的成像系统的相机)获取的图像数据可以缩减为一组映射点。

[0109] 一个或多个对象识别器708可以浏览接收到的数据(例如,点的收集)并且借助于地图数据库710来识别或映射点、标记图像,将语义信息附加到对象。地图数据库710可以包括随时间收集的各个点及其对应的对象。各种设备和地图数据库可以通过网络(例如LAN, WAN等)相互连接以访问云。

[0110] 基于该信息和地图数据库中的点的集合,对象识别器708a-708n可识别环境中的对象。例如,对象识别器可以识别面部、人、窗、墙、用户输入设备、电视、文档(例如,本文安全示例中描述的旅行票、驾驶执照、护照)、用户环境中的其他对象等。一个或多个对象识别器可以专用于具有某些特征的对象。例如,对象识别器708a可以用于识别面部,而另一个对象识别器可以用于识别文档。

[0111] 可以使用各种计算机视觉技术来执行对象识别。例如,可穿戴系统可以分析由面向外的成像系统464(图4中所示)获取的图像以执行场景重建、事件检测、视频跟踪、对象识别(例如,人或文档)、对象姿势估计、面部识别(例如,来自环境中的人或文档上的图像)、学习、索引、运动估计或图像分析(例如,识别文档内的标记,诸如照片、签名、识别信息、旅行信息等)等等。可以使用一个或多个计算机视觉算法来执行这些任务。计算机视觉算法的非限制性示例包括:尺度不变特征变换(SIFT)、加速鲁棒特征(SURF)、定向FAST和旋转BRIEF(ORB)、二进制鲁棒不变可缩放关键点(BRISK)、快速视网膜关键点(FREAK)、Viola-Jones算法、Eigenfaces方法、Lucas-Kanade算法、Horn-Schunk算法、均值平移(Mean-shift)算法,视觉同时定位和地图构建(vSLAM)技术、顺序贝叶斯估计器(例如,卡尔曼滤波器、扩展卡尔曼滤波器等)、光束法平差、自适应阈值化(和其他阈值化技术)、迭代最近点(ICP)、半全局匹配(SGM)、半全局块匹配(SGBM)、特征点直方图、各种机器学习算法(例如支持向量机、k-最近邻算法、朴素贝叶斯、神经网络(包括卷积或深度神经网络),或其它有监督/无监督模型等),等等。

[0112] 一个或多个对象识别器708还可以实现各种文本识别算法,以从图像中识别和提取文本。一些示例文本识别算法包括:光学字符识别(OCR)算法、深度学习算法(例如深度神经网络)、模式匹配算法、用于预处理的算法等。

[0113] 对象识别可以附加地或替代地通过各种机器学习算法来执行。一旦经过训练,机器学习算法就可以由HMD存储。机器学习算法的一些示例可以包括:有监督或无监督机器学习算法,包括回归算法(例如,普通最小二乘回归),基于实例的算法(例如,学习矢量量化),决策树算法(例如,分类和回归树),贝叶斯算法(例如,朴素贝叶斯),聚类算法(例如,k均值

聚类), 关联规则学习算法 (例如, 先验算法), 人工神经网络算法 (例如, 感知器), 深度学习算法 (例如, 深度玻尔兹曼机, 或深度神经网络), 降维算法 (例如, 主成分分析), 集成算法 (例如, 层叠泛化) 或其他机器学习算法。在一些实施例中, 可以针对各个数据集定制各个模型。例如, 可穿戴设备可以生成或存储基础模型。基础模型可以用作起点以生成特定于数据类型 (例如, 远程呈现会话中的特定用户)、数据集 (例如, 在远程呈现会话中从用户获得的附加图像的集合)、条件情况或其他变化的附加模型。在一些实施例中, 可穿戴HMD可以被配置为利用多种技术来生成用于分析聚合数据的模型。其他技术可包括使用预定义的阈值或数据值。

[0114] 基于该信息和地图数据库中的点的收集, 对象识别器708a-708n可识别对象并用语义信息对对象进行补充以赋予对象生命。例如, 如果对象识别器识别一组点作为门, 则系统可以附加一些语义信息 (例如, 门具有铰链并且在铰链周围具有90度的运动)。如果对象识别器识别出一组点作为镜子, 则系统可以附加语义信息, 即镜子具有可反射房间中对象的图像的反射表面。语义信息可以包括如本文所述的对象的功能可见性 (affordance)。例如, 语义信息可以包括对象的法向量 (normal)。系统可以指定方向指示该对象的法向量的矢量。随着时间的推移, 地图数据库会随着系统 (可能驻留在本地或可能通过无线网络访问) 积累更多来自世界的的数据而增大。一旦对象被识别, 信息就可以被传送到一个或多个可穿戴系统。例如, MR环境700可以包括关于在加利福尼亚发生的场景的信息。该环境700可以被发送到纽约的一个或多个用户。基于从FOV相机和其它输入接收的数据, 对象识别器和其它软件部件可以映射从各种图像收集的点、识别对象等, 使得场景可以准确地“传递”给可能在世界的不同地区的第二用户。环境700也可以使用拓扑图用于本地化目的。

[0115] 图8是渲染与识别的对象相关的虚拟内容的方法800的示例的过程流程图。方法800描述如何将虚拟场景呈现给可穿戴系统的用户。用户可能在地理上远离场景。例如, 用户可能是在纽约, 但可能想要查看目前在加利福尼亚州正在进行的场景, 或者可能想要与居住在加利福尼亚州的朋友散步。

[0116] 在框810处, 可穿戴系统可以从用户和其它用户接收关于用户的环境的输入。这可以通过各种输入设备和地图数据库中已有的知识来实现。在框810处, 用户的FOV相机、传感器、GPS、眼睛跟踪等向系统传达信息。在框820处, 系统可以基于该信息确定稀疏点。稀疏点可用于确定姿势数据 (例如, 头部姿势、眼睛姿势、身体姿势或手势), 这些姿势数据可用于显示和理解用户周围环境中各种对象的方位和位置。在框830处, 对象识别器708a-708n可以使用地图数据库浏览这些收集的点并识别一个或多个对象。然后在框840处, 该信息可以被传达给用户的个人可穿戴系统, 并且在框850处, 可以将期望的虚拟场景相应地显示给用户。例如, 期望的虚拟场景 (例如, 加州的用户) 可以与纽约的用户的各种对象和其它环境的相关地显示在适当的方位、位置等。

[0117] 图9是可穿戴系统的另一示例的框图。在该示例中, 可穿戴系统900包括地图920, 该地图920可以包括包含世界的地图数据的地图数据库710。地图可以部分地驻留在可穿戴系统上, 并且可以部分驻留在可以通过有线或无线网络 (例如, 在云系统中) 访问的联网存储位置处。姿势处理910可以在可穿戴计算架构 (例如, 处理模块260或控制器460) 上执行, 并且利用来自地图920的数据来确定可穿戴计算硬件或用户的位置和方位。姿势数据可以根据用户正在体验系统并在世界上操作时随时收集的数据来计算。数据可以包括图像, 来

自传感器(例如通常包括加速度计和陀螺仪组件的惯性测量单元)的数据以及与真实或虚拟环境中的对象有关的表面信息。

[0118] 稀疏点表示可以是同时定位和地图构建(例如,SLAM或vSLAM,参考其中输入只是图像/视觉的配置)过程的输出。该系统可以配置为不仅可以查找各个部件在世界中的位置,而且还了解世界由什么组成。姿势可以是实现许多目标的组成部件,包括填充地图和使用来自地图的数据。

[0119] 在一个实施例中,稀疏点位置本身可能并不完全足够,并且可能需要进一步的信息来产生多焦点AR、VR或MR体验。通常涉及深度地图信息的密集表示可以被用来至少部分地填补该缺口。这样的信息可以根据被称为立体(Stereo)940的过程来计算,其中使用诸如三角测量或飞行时间感测的技术来确定深度信息。图像信息和有源(active)图案(诸如使用有源(active)投影仪创建的红外图案)、从图像相机或手部手势/图腾950获取的图像可以用作立体过程940的输入。大量的深度地图信息可以融合在一起,并且其中一些可以用表面表示来概括。例如,数学上可定义的表面可以是有效的(例如相对于大的点云)并且对诸如游戏引擎的其它处理设备是可消化输入。因此,可以在融合过程930中组合立体过程(例如,深度图)940的输出。姿势910也可以是该融合过程930的输入,并且融合过程930的输出变成填充地图过程920的输入。例如在地形绘制中,子表面可以彼此连接以形成更大的表面,并且地图变成点和表面的大混合。

[0120] 为了解决混合现实过程960中的各个方面,可以使用各种输入。例如,在图9所示的实施例中,游戏参数可以是输入以确定系统的用户正在与各种位置处的一个或多个怪物进行怪物作战游戏,在各种条件下死亡或逃跑的怪物(诸如如果用户射击怪物),各种位置处的墙壁或其它对象等。世界地图可以包括对象的位置相关的信息或者对象的语义信息,并且世界地图可以是到混合现实的另外的有价值输入。与世界相关的姿势也成为一种输入,并且对几乎任何交互系统都起着关键作用。

[0121] 来自用户的控制或输入是可穿戴系统900的另一输入。如这里所述,用户输入可以包括视觉输入、手势、图腾、音频输入、感觉输入等。为了在周围移动或玩游戏,例如,用户可能需要指示可穿戴系统900关于他或她想做什么。除了仅仅在空间中移动自己之外,还存在可以使用的各种形式的用户控制。在一个实施例中,图腾(例如,用户输入设备)或诸如玩具枪的对象可由用户握持并由系统跟踪。系统优选地将被配置为知道用户正握持物品并且理解用户与物品具有什么样的交互(例如,如果图腾或对象是枪,则系统可以被配置为理解位置和方位,以及用户是否在点击触发器或可以配备有传感器(诸如IMU)的其它感测按钮或元件,这可以有助于确定正在发生什么,即使当这种活动不在任何相机的视场内。)

[0122] 手势跟踪或识别还可以提供输入信息。可穿戴系统900可以被配置成跟踪和解释按钮按压的手势,用于动作表示左或右、停止、抓住、握持等。例如,在一种配置中,用户可能想要翻阅非游戏环境中的电子邮件或日历,或与另一个人或玩家进行“拳头碰撞”。可穿戴系统900可以被配置为利用最小量的手势,其可以是动态的也可以不是动态的。例如,手势可以是简单的静态手势,如张开的手用于停止,拇指向上用于ok,拇指向下用于非ok;或者手向右或左、或上/下轻击用于方向命令。手势跟踪可以包括跟踪用户环境中的其他人做出的手势,诸如做出手势以利用手语进行交流的其他人(见,例如图13A)。

[0123] 眼睛跟踪是另一输入(例如,跟踪用户正在看的地方从而控制显示技术以在特定

深度或范围渲染)。在一个实施例中,可以使用三角测量来确定眼睛的聚散度,然后使用针对特定人开发的聚散度/适应性模型,可以确定适应性。眼睛跟踪可以由一个或多个眼睛相机来执行以确定眼睛注视(例如,单眼或双眼的方向或方位)。可以使用其他技术用于眼睛跟踪,例如通过位于单眼或双眼附近的电极对于电位的测量(例如,眼球运动电位测定法)。

[0124] 语音跟踪可以是另外的输入,其可以单独使用或与其他输入(例如,图腾跟踪、眼睛跟踪、手势跟踪等)组合使用。语音跟踪可以单独或组合地包括语音识别、话音识别。系统900可以包括从环境接收音频流的音频传感器(例如,麦克风)。系统900可以结合语音识别技术以确定谁在说话(例如,语音是来自ARD的佩戴者还是来自另一个人或话音(例如,由环境中的扬声器发送的记录的话音))以及语音识别技术以确定所说的内容。本地数据和处理模块260或远程处理模块270可以处理来自麦克风的音频数据(或另一流中的音频数据,例如,用户正在观看的视频流),以通过应用各种语音识别算法来识别语音的内容,各种语音识别算法,例如隐马尔可夫模型、基于动态时间规整(DTW)的语音识别、神经网络、诸如深度前馈和递归神经网络的深度学习算法、端到端自动语音识别、机器学习算法(参考图7描述)、或使用声学建模或语言建模等的其他算法。

[0125] 混合现实过程960的另一输入可包括跟踪环境中的标示。标示可包括商业或公共显示标志。如参考图16A至图19所示,系统可以识别标示,识别标示中的文本,调整文本的特征(例如,增加文本的字体大小以提高可读性),修改文本的内容(例如,将文本从外语翻译成用户理解的语言)等。

[0126] 本地数据和处理模块260或远程处理模块270还可以应用语音识别算法,该算法可以识别说话者的身份,例如说话者是否是可穿戴系统900的用户210或该用户与其交谈的另一个人。一些示例语音识别算法可以包括频率估计、隐马尔可夫模型、高斯混合模型、模式匹配算法、神经网络、矩阵表示、矢量量化、扬声器测距(diarisation)、决策树和动态时间规整(DTW)技术。话音识别技术还可以包括反扬声器技术,例如群体(cohort)模型和世界模型。频谱特征可用于表示说话者特征。本地数据和处理模块或远程数据处理模块270可以使用参考图7描述的各种机器学习算法来执行话音识别。

[0127] 系统900还可以包括感测眼镜系统970,用于促进用户与其他人或环境的交互。感测眼镜系统970的实现可以通过UI使用这些用户控件或输入。例如,可以使用UI要素(例如,控件、弹出窗口、气泡、数据输入字段等)来消除信息的显示,例如,转换的文本、图形或辅助信息,或者请求显示辅助信息。UI可以允许用户输入用户理解的一种或多种语言的列表,使得感测眼镜系统970知道在翻译由对话伙伴以手语做出的标志时使用哪种语言。下面进一步描述这种实现和这些用途的示例。

[0128] 感测眼镜系统970还可以包括文本识别、修改和渲染特征。这些特征可以与可穿戴系统的各种其他部件组合以增强用户与环境的交互。例如,HMD可以包括一个或多个光源11,其被配置为基于从用户的物理环境的图像识别的文本将图像投影到显示器上(例如,使得投影图像遮挡来自物理环境的原始文本)。光学透射目镜106可以被配置为将来自一个或多个光源11的光传输到用户210作为图像。图像可能看起来好像处于特定深度,这可能只是HMD系统200已经显示图像的许多可能深度中的一个。HMD系统100可以能够投影图像以出现在多个不同深度处,这可以看起来好像在不同深度平面306上(参见图3)。在目镜106是光学透射的一些实施例中,目镜106可以允许来自环境的光进入用户的眼睛。因此,在这样的实

施例中,用户210可以看到来自环境的图像的部分以及来自一个或多个光源11的投影图像。

[0129] 关于相机系统,图9中示出的示例可穿戴系统900可以包括三对相机:布置在用户脸部两侧的相对宽的FOV或无源SLAM相机对,定向在用户前面的不同相机对,以处理立体成像过程940并且还捕获手势和用户面前的图腾/对象跟踪。用于立体过程940的FOV相机或者相机对还可以称为相机16。用于立体过程940的FOV相机或者相机对可以是面向外的成像系统464(在图4中示出)的一部分。可穿戴系统900可以包括眼睛跟踪相机(其还示出为眼睛相机24并且可以是图4中所示的面向内的成像系统462的一部分),其朝向用户的眼睛定位从而对眼睛向量和其它信息进行三角测量。可穿戴系统900还可以包括一个或多个纹理光投影仪(诸如红外(IR)投影仪)以将纹理注入到场景中。

[0130] 图10是用于确定向可穿戴系统的用户输入的方法1000的示例的过程流程图。在这个示例中,用户可以与图腾进行交互。用户可具有多个图腾。例如,用户可具有社交媒体应用的指定的一个图腾,玩游戏的另一个图腾等。在框1010处,可穿戴系统可以检测图腾的运动。图腾的运动可以通过面向外的成像系统进行识别或者可以通过传感器(例如,触觉手套、图像传感器、手跟踪设备、眼部跟踪相机、头部姿势传感器等)进行检测。

[0131] 在框1020处,至少部分地基于检测到的手势、眼睛姿势、头部姿势或通过图腾的输入,可穿戴系统检测图腾(或用户的眼睛或头部或手势)相对于参考系的位置、方位或移动。参考系可以是一组地图点,可穿戴系统根据该地图点将图腾(或用户)的运动转换为动作或命令。在框1030处,与图腾的用户交互被映射。基于用户交互相对于参考系1020的映射,在框1040处系统确定用户输入。

[0132] 例如,用户可以来回移动图腾或物理对象以表示翻动虚拟页面并移动到下一页面或者从一个用户界面(UI)显示屏幕移动到另一个UI屏幕。作为另一个示例,用户可以移动他们的头部或眼睛来查看用户FOR中的不同的现实或虚拟对象。如果用户注视特定现实或虚拟对象的时间长于阈值时间,则可以选择现实或虚拟对象作为用户输入。在一些实施方式中,可以跟踪用户眼睛的聚散度并且可以使用适应性/聚散度模型来确定用户眼睛的适应状态,其提供用户正在聚焦的深度平面的信息。在一些实施方式中,可穿戴系统可以使用光线投射技术来确定哪些现实或虚拟对象是沿着用户的头部姿势或眼睛姿势的方向。在各种实施方式中,光线投射技术可包括投射具有基本上很小横向宽度的薄的铅笔射线或具有基本横向宽度(例如锥体或平截头体)的投射光线。

[0133] 用户界面可由在此所述的显示系统(例如图2A中的显示器220)投影。它也可以使用各种其它技术显示,例如一个或多个投影仪。投影仪可以将图像投影到例如画布或球体的物理对象上。可以使用系统外部或系统的一部分的一个或多个相机(例如,使用面向内的成像系统462或面向外的成像系统464)来跟踪与用户界面的交互。

[0134] 图11是用于与虚拟用户界面进行交互的方法1100的示例的过程流程图。方法1100可以由在此描述的可穿戴系统来执行。方法1100的实施例可以由可穿戴系统使用以检测可穿戴系统的FOV中的人或者文档。

[0135] 在框1110处,可穿戴系统可以识别特定UI。UI的类型可以由用户预先确定。可穿戴系统可以基于用户输入(例如,手势、视觉数据、音频数据、感测数据、直接命令等)识别特定UI需要填充。该UI可以是特定于安全场景,在该安全场景中系统的佩戴者正在观察向该佩戴者呈现文档的用户(例如,旅行检查点处)。在框1120处,可穿戴系统可以生成针对虚拟UI

的数据。例如,可以生成与UI的界限、一般结构、形状等相关的数据。另外,可穿戴系统可以确定用户的物理位置的地图坐标,使得可穿戴系统可以显示与用户的物理位置相关的UI。例如,如果UI是以身体为中心的,则可穿戴系统可以确定用户的身体姿态、头部姿势或眼睛姿势的坐标,使得可以在用户周围显示环形UI或者可以在墙上或在用户面前显示平面UI。在这里所描述的安全性的情况下,该UI可以显示为该UI仿佛围绕着正在向系统的佩戴者呈现文档的旅行者一样,使得佩戴者可以容易地观看UI的同时看着该旅行者和该旅行者的文档。如果UI是以手为中心的,则可以确定用户的手的地图坐标。这些地图点可以通过以下数据得出,通过FOV相机、感测输入所接收的数据或任何其它类型的收集数据。

[0136] 在框1130处,可穿戴系统可以将数据从云端发送到显示器,或者可以将数据从本地数据库发送到显示器部件。在框1140处,基于发送的数据向用户显示UI。例如,光场显示器可以将虚拟UI投影到用户的一只或两只眼睛中。一旦创建了虚拟UI,在框1150处,可穿戴系统可以简单地等待来自用户的命令以在虚拟UI上生成多个虚拟内容。例如,UI可以是围绕用户身体或者用户环境中的人(例如,旅行者)的身体的身体中心环。然后,可穿戴系统可以等待命令(手势、头部或眼部运动、语音指令、来自用户输入设备的输入等),并且如果识别出(框1160),则可以向用户显示与该命令相关联的虚拟内容(框1170)。

[0137] 在美国专利公开No.2015/0016777中描述了可穿戴系统、UI和用户体验(UX)的附加示例,其全部内容通过引用合并于此。

[0138] 多个可穿戴系统之间的通信示例

[0139] 图12示意性地示出了描绘彼此交互的多个用户设备的整体系统视图。计算环境1200包括用户设备1230a、1230b、1230c。用户设备1230a、1230b和1230c可以通过网络1290彼此通信。用户设备1230a-1230c每个可以包括网络接口,以经由网络1290与远程计算系统1220(其还可以包括网络接口1271)通信。网络1290可以是LAN、WAN、点对点网络、无线电、蓝牙或任何其他网络。计算环境1200还可以包括一个或多个远程计算系统1220。远程计算系统1220可以包括集群并位于不同地理位置处的服务器计算机系统。用户设备1230a、1230b和1230c可以经由网络1290与远程计算系统1220通信。

[0140] 远程计算系统1220可以包括远程数据储存库1280,其可以维护关于特定用户的物理或虚拟世界的信息。数据存储器1280可以包含对感测眼镜有用的信息,例如手语字典、辅助信息源等。远程数据储存库可以是图2A中所示的远程数据储存库280的实施例。远程计算系统1220还可以包括远程处理模块1270。远程处理模块1270可以是图2A中所示的远程处理模块270的实施例。在一些实施方式中,远程计算系统1220可以是与可穿戴系统200独立的第三方系统。

[0141] 远程处理模块1270可以包括可以与用户设备(1230a,1230b,1230c)和远程数据储存库1280通信的一个或多个处理器。处理器可以处理从用户设备和其他源获得的信息。在一些实施方式中,处理或存储的至少一部分可以由本地处理和数据模块260提供(如图2A所示)。远程计算系统1220可以使给定用户能够与另一用户共享关于特定用户自己的物理或虚拟世界的信息。

[0142] 用户设备可以是可穿戴设备(诸如HMD或ARD)、计算机、移动设备或单独或组合的任何其他设备。例如,用户设备1230b和1230c可以是图2A所示的可穿戴系统200(或图4所示的可穿戴系统400)的实施例,其可以被配置为呈现AR/VR/MR内容。

[0143] 一个或多个用户设备可以与图4中所示的用户输入设备466一起使用。用户设备可以获得关于用户和用户环境的信息(例如,使用图4中所示的面向外的成像系统464)。用户设备或远程计算系统1220可以使用从用户设备获得的信息来构建、更新和建立图像、点和其他信息的集合。例如,用户设备可以处理所获取的原始信息并将经处理的信息发送到远程计算系统1220以进行进一步处理。用户设备还可以将原始信息发送到远程计算系统1220以进行处理。用户设备可以从远程计算系统1220接收经处理的信息,并在投影给用户之前提供最终处理。用户设备还可以处理所获得的信息并将处理后的信息传递给其他用户设备。用户设备可以在处理所获取的信息的同时与远程数据储存库1280通信。多个用户设备或多个服务器计算机系统可以参与所获取图像的构建或处理。

[0144] 关于物理世界的信息可以随着时间的推移来开发,并且可以基于由不同用户设备收集的信息。虚拟世界的模型也可以随着时间的推移而开发并且基于不同用户的输入。这些信息和模型有时可以在本文中称为世界地图或世界模型。如参考图7和图9所述,由用户设备获取的信息可以用于构建世界地图1210。世界地图1210可以包括图9中描述的地图920的至少一部分。各种对象识别器(例如,708a、708b、708c、...、708n)可用于识别对象和标签图像,以及将语义信息附加到对象。这些对象识别器也在图7中描述。

[0145] 远程数据储存库1280可用于存储数据并便于构建世界地图1210。用户设备可以不断更新关于用户环境的信息并接收关于世界地图1210的信息。世界地图1210可以由用户或其他人创建。如在此所讨论的,单独或组合的用户设备(例如,1230a、1230b、1230c)和远程计算系统1220可以构建或更新世界地图1210。例如,用户设备可以与远程处理模块1270和远程数据储存库1280通信。用户设备可以获取或处理关于用户和用户环境的信息。远程处理模块1270可以与远程数据储存库1280和用户设备(例如,1230a、1230b、1230c)通信,以处理关于用户和用户的环境的信息。远程计算系统1220可以修改由用户设备(例如,1230a、1230b、1230c)获取的信息,例如,选择性地裁剪用户的图像,修改用户的背景,将虚拟对象添加到用户的环境,用辅助信息注释用户的语音等。远程计算系统1220可以将处理的信息发送到相同或不同的用户设备。

[0146] 以下进一步描述感测眼镜系统的实施例的各种功能。

[0147] 用于促进用户交互的示例感测眼镜

[0148] 可穿戴系统200可以实现感测眼镜系统970,以便于用户与其他人或与环境的交互。作为与其他人交互的一个示例,可穿戴系统200可以通过例如检测可以构成手语的手势,将手语翻译成另一种语言(例如,另一种手语或口语),以及将翻译的信息呈现给可穿戴设备的用户,来解释手语。作为另一个例子,感测眼镜系统970可以将语音翻译成手语并将手语呈现给用户。

[0149] 可穿戴系统970还可以通过识别环境中的对象,修改对象的特征(在虚拟环境中),以及将修改的对象作为虚拟对象呈现给用户,来促进用户与环境的交互。例如,可穿戴系统200可以基于由面向外的成像系统464获取的图像来识别用户环境中的标志(例如,交通标志、商店前面的标志等),修改用户环境中的标志的特征,并将经修改的标志呈现给用户。修改的标志可以覆盖在用户的3D环境上,使得可以遮挡原始标志。

[0150] 示例感测眼镜系统作为人际通信工具

[0151] 在一些情况下,会话中的一个或多个人可以使用手或身体姿势(例如,手语)来表

达他们自己。会话可以在远程呈现会话期间或当人们在彼此的物理附近时发生。当用户与示意者通信时,可穿戴系统200可以为可穿戴系统200的用户(也称为观察者)解释示意者的手语。可穿戴系统200还可以将言语或基于语音的手语翻译成图形(例如,手势的图像)并将图形呈现给示意者,使得示意者能够理解观察者的语音。例如,佩戴头戴式显示器的观察者可能具有减小的视野,因此观察者可能无法观察由示意者使用手语进行的完整手势。可穿戴系统200可以使用面向外的成像系统464(因为它可以具有相机,该相机具有更宽的视野,用户可以通过头戴式显示器感知到的更宽的视野内容)捕获由示意者做出的手势。可穿戴系统200可以将捕获的手势作为虚拟图形示出给观察者,或者示出从捕获的手势转换的文本,以便于观察者理解示意者的语音。此外,可穿戴系统200可以被配置为将一种手语翻译成另一种手语。例如,会话中的一个人可以使用美国手语,而另一个人可以使用多贡手语。可穿戴系统200可以为使用多贡手语的人将美国手语翻译为多贡手语,并且为使用美国手语的人将多贡手语翻译成美国手语。

[0152] 示例手语捕获

[0153] 可穿戴系统可以使用各种技术来捕获原始语音并将原始语音翻译成目标语音。语音可以是手或身体姿势或可听声音的形式。如本文所述,原始语音可以是手语,目标语音可以是另一种手语或口语。可替换地,原始语音可以是口语,而目标语音是手语。可穿戴系统200可以使用面向外的成像系统464、音频传感器232,或者通过经由网络1290与另一计算设备通信,来捕获原始语音,这取决于语音的上下文(例如,语音是当面的还是通过远程通信的)。

[0154] 作为在当面通信期间捕获原始语音的示例,其中检测到的手语的示意者是在感测眼镜系统的物理附近,面向外的成像系统464可以捕获用户环境的图像。可穿戴系统200可以从图像信息中检测可以构成手语的手势(例如,手/身体姿势或嘴唇运动)。可穿戴系统200可以使用诸如深度神经网络、隐马尔可夫模型、动态编程匹配等算法来识别手势,以识别由讲话者做出的手势所代表的标志。如参考图7所述,手势识别可以由一个或多个对象识别器708执行。

[0155] 作为在远程通信的情况下捕获原始语音的示例,可穿戴系统200可以通过分析从远程计算设备(例如,另一可穿戴设备)接收的数据或通过分析由面向外的成像系统464捕获的数据(例如,原始语音存在于电视上的情况下),来捕获和识别原始语音的存在。在一个示例中,示意者和观察者可以通过因特网视频聊天会话进行交谈。示意者和观察者可以各自佩戴他们各自的HMD。HMD可以经由网络1290彼此通信(如图12所示)。在示意者位于反射表面(例如,镜子)前面的情况下,示意者的HMD可以通过经由面向外的成像系统464获取示意者的反射图像来捕获示意者的手势。可以将示意者的反射图像发送给观察者的HMD或远程计算系统1220,以识别和解释手语。作为另一示例,示意者可以是视频节目中的人,诸如在电视或因特网节目上呈现的人等。在可以在观察者的位置处可视地观察示意者的手势的情况下,可穿戴系统464可以通过与在当面通信情况下的相同方式(例如,经由音频传感器232或面向外的成像系统464)来捕获手语手势。

[0156] 除了将由另一个人做出手势的手语的文本或图形翻译显示给可穿戴系统200的用户之外或作为替代,可穿戴系统200的用户还可以用手语进行通信。在这种情况下,可穿戴系统可以通过面向外的成像系统464捕获用户自己的手语手势(从第一人称的角度来看)。

可穿戴系统可以将手语转换为可以表达为文本、音频、图像等格式的目标语音。可穿戴系统200可以将结果发送到另一可穿戴系统以呈现给另一用户。如在此所述,从原始语音到目标语音的转换可以由用户的可穿戴系统、另一个用户的可穿戴系统或远程计算系统1220单独或组合地执行。例如,用户的可穿戴系统可以捕获用户的手势并将捕获的视频或图像(包含手语手势)发送到另一用户的可穿戴系统或远程计算系统120,该另一用户的可穿戴系统或远程计算系统120可以从视频或图像中提取手语并将手语转换为音声语言或其他手语的视听内容。视听内容可包括文本、图形、视频、动画、声音等。

[0157] 示意者手势拒绝和源本地化

[0158] 可穿戴系统可以使用各种传感器识别手势或手语的源,例如音频传感器232、面向外的成像系统464、静止输入704或用户环境中的其他传感器。作为一个示例,可穿戴系统可以从由面向外的成像系统464获取的数据中检测一系列手势以及嘴唇运动。可穿戴系统可以发现手势与示意者相关联,因为示意者也具有相应的嘴唇运动。作为另一示例,可穿戴系统可以测量用户与手势之间的距离以确定手势源。例如,可穿戴系统可以确定一系列手势来自用户,因为手在由面向外的成像系统464获取的图像中看起来相对较大。但是如果手看起来相对较小,则可穿戴系统可以发现手势来自用户以外的人。作为又一示例,可穿戴系统可以通过识别正在播放视听内容的对象(例如,通过使用对象识别器708识别电视)来发现手势来自视听内容(例如,在电视中)。

[0159] 基于手势源,可穿戴系统200可以被配置为不处理来自某些人的手势。例如,可穿戴系统可以捕获来自用户环境中的多个人的手势,但是可穿戴系统可以被配置为不处理来自用户的FOV中心之外的人的手语以用于手语识别。作为另一示例,可穿戴系统可以被配置为不处理用户自己的手语。

[0160] 在一些实施例中,可穿戴系统可以配置传感器以检测用户自己的手语,例如通过将相机定位在面向外的成像系统464中的角度使得用户不必举起手以使面向外的成像系统464捕获用户的手势。传感器还可以配置为不检测用户自己的手语。例如,可以通过不在用户自己的手的方向上捕获图像(通常低于用户的FOV),或者滤除(例如,通过裁剪)在这样的方向上的图像来实现不检测。因此,系统可以将用户自己的手语与其他人的手语区分开来。

[0161] 从手语到文本的示例转换

[0162] 可穿戴系统200可以将捕获的手语转换为可以呈现给用户或翻译成另一种语言的文本。可以使用诸如深度学习(可以利用深度神经网络)、隐马尔可夫模型、动态编程匹配等算法来执行手语到文本的转换。例如,深度学习(在一些情况下是卷积神经网络)可以对包含已知标志的图像或视频进行训练(监督学习),以便确定代表标志的特征并基于所学习的特征建立分类模型。然后,这种训练的深度学习方法可以由可穿戴系统200的本地处理和数据模块260或远程处理模块和数据储存库270、280应用于由面向外成像子系统检测到的示意者的图像。

[0163] 文本转换功能可以由本地处理和数据模块260、远程处理模块270、远程数据储存库280或远程计算系统1220单独或组合地实现。例如,可穿戴系统200可以包括在HMD上实现的手语到文本的功能。作为一个示例,可穿戴系统可以将手语字典存储在本地数据模块260或远程数据储存库280中。可穿戴系统可以相应地访问手语字典以将检测到的手势转换为文本。作为另一示例,可穿戴系统200可以访问由远程计算系统1220实现的手语到文本功

能。可穿戴系统200可以利用到商业的手语到文本服务或数据储存库的无线连接(例如,通过应用程序编程接口(API))。例如,可穿戴系统200可以向远程计算系统1220提供捕获的手势并从远程计算系统1220接收相应的文本。

[0164] 无论是本地还是远程执行转换,诸如显示转换的文本和检索辅助信息的其他处理步骤(下面将进一步描述)可以在本地或远程完成,而与执行文本转换的位置无关。例如,如果远程完成手语到文本的转换并且要在本地显示转换的文本(例如,系统的用户是观察者),则可以将捕获的视频流发送到远程处理模块270或通过网络执行转换的远程服务器;转换的文本串被返回到系统的本地组件(例如,本地处理和数据模块260)以供显示。作为另一示例,如果远程完成手语到文本转换和辅助信息检索,则可以经由网络将捕获的视频流发送到远程处理模块270或远程服务器,并且可以将检索到的辅助信息返回到系统的本地组件。本地/远程处理的其他组合也是可行的。

[0165] 尽管参考将标志转换为文本来描述这些示例,但是可以将标志转换为各种其他格式,例如图形、动画、音频或其他类型的视听内容。此外,标志的翻译不需要首先将标志翻译成文本。

[0166] 将一种手语转换为另一种手语的示例

[0167] 如在此所述,全世界有数百种手语。因此,当两个会话伙伴都在用标志表示但是处于不同的手语系统时,也可以使用这里描述的可穿戴系统。有利地,每个这样的示意者可以使用他或她自己的可穿戴系统将另一个示意者的标志翻译为用户自己的手语系统。可穿戴系统可以将标志翻译为用户理解的文本或者用户自己的手语的图形表示。

[0168] 可穿戴系统200可以被配置为识别特定手语,例如美国手语(ASL)。可穿戴系统200还可以被配置为识别多种手语,例如ASL、英国手语、中国手语、多贡手语等。在一些实现中,可穿戴系统200支持手语识别的重新配置,例如,基于感测眼镜系统的位置信息。可穿戴系统可以通过类似于系统如何识别用户自己的或优选的手语的手段来识别外国手语,例如,单独利用对象识别器708或者与手语字典组合来识别由面向外成像系统464感知到的手势。可穿戴系统可以将用户感知的手语转换成用户的主导手语。用户的主导手语可以是用户的第一手语或对话中用户的优选手语。除用户的主导手语之外的手语可以被认为是外国手语。可穿戴系统可以允许用户选择外国手语的转换文本。例如,用户可以选择外国手语,并且可穿戴系统可以将外国手语中的手势的含义作为文本呈现给可穿戴系统的用户。

[0169] 可穿戴系统可以借助于环境中的口语或位置信息来识别外国手语。例如,可穿戴系统在用户的环境中检测意大利语,或者基于GPS获取的数据确定用户在意大利。基于该信息,可穿戴系统可以自动激活用于识别意大利手语的功能。作为另一示例,可穿戴系统可具有可穿戴系统能够支持的手语的优先顺序。在此示例中,ASL可能优先于意大利手语,因为用户来自美国。然而,一旦可穿戴系统检测到用户被意大利语使用者包围或者物理上位于意大利,可穿戴系统就可以改变优先顺序,以便意大利手语现在在ASL之前。因此,可穿戴系统可以将意大利手语翻译成与ASL相关联的英文文本或图形。

[0170] 感测眼镜系统不仅可以帮助用户理解外国手语,还可以帮助用户用标志表示外国手语。例如,可穿戴系统可以配置为将用户自己的语言翻译成外国手语。系统可以在显示器上显示外国手语手势(例如,翻译的手语)。用户可以看到外国手语中的手势并模仿手势。例如,用户可能正在与听力受损的示意者交谈。可穿戴系统可以捕获用户的语音并以示意者

理解的手语向用户显示相应的手势。因此,用户可以做出显示器呈现的手势以便与示意者通信。在一些实施例中,不是向用户示出手势,而是可穿戴系统可以将对应于用户语音的标志传送给示意者,使得示意者能够理解用户的发声语音。

[0171] 可穿戴系统可以包括音频放大器(例如,扬声器240)以在音频中提供识别的手语。例如,可穿戴系统可以将示意者的手语转换为音频流,以便回放给可穿戴系统200的用户。

[0172] 确定与手语相关的辅助信息的示例

[0173] 人们不知道或理解对话中的单词或短语,包括涉及手语的对话,这并不罕见。可穿戴系统可以显示与所显示文本的一部分相关联的辅助信息,以增强用户的理解。辅助信息可以包括扩充并增加到定义的上下文的信息,诸如定义、翻译、解释等。辅助信息可以以各种形式呈现,例如文本、图像、图形、动画或其他音频或视觉信息。该系统可以在视觉上呈现辅助信息,例如,通过图2A中的显示器220。该系统可以例如通过图2A中的音频放大器240以音频将辅助信息呈现给没有听力障碍的用户。通过为这样的单词或短语提供定义、翻译、解释或其他信息,可穿戴系统有利地可以帮助用户更好地理解用户观察的手语。

[0174] 可以基于用户环境的上下文信息、语音的上下文等来确定辅助信息。作为示例,可穿戴系统可以至少部分地利用用户行为来确定是否显示与对话伙伴的标志相关联的辅助信息。例如,用户可以暂时盯着某个方向(例如,朝向示意者或示意者的手)。可穿戴系统可以检测用户的凝视方向(例如,使用面向内的成像系统462),并且作为响应,可以检索和显示与对话伙伴的标志相关联的辅助信息。

[0175] 可穿戴系统可以包括辅助信息的数据储存库(例如,数据库)。可穿戴系统可以通过访问数据储存库来检索与显示的文本相关联的辅助信息。这种信息数据库可以本地存储在可穿戴设备中,例如,存储在图2A中的数据模块260中,或者远程存储在例如远程数据储存库270中。可穿戴系统可以利用例如因特网上的信息的可公开访问的信息来确定辅助信息。例如,可穿戴系统可以访问网络以将关于会话中的单词/短语的查询发送到因特网上的资源,诸如字典、百科全书或其他类似资源。这样的资源可以是通用的(例如,诸如维基百科的通用百科全书)或专用的(例如,药物索引,诸如rxlist.com上的药物索引,或矿物学数据库(例如,webmineral.com))。

[0176] 转换的标志或辅助信息的示例显示、解除和召回

[0177] 可穿戴系统可以将转换的标志(例如,以文本或图形格式)单独地或者与辅助信息组合地呈现给可穿戴系统200的用户。例如,可穿戴系统200可以被配置为将辅助信息与手语的转换的标志一起显示,单独显示转换的标志或辅助信息(例如,在显示辅助信息时仅在期间显示辅助信息),或者在两种显示模式之间切换。转换的标志单独或与辅助信息进行组合有时可称为显示项目。

[0178] 可以以各种方式呈现转换的文本或辅助信息。在一个示例中,可穿戴系统200可以将转换的文本或辅助信息放置在文本气泡中,例如,在示意者附近几何地定位的文本,例如如图13中的图形1355中所示。作为另一个示例,可穿戴系统200可以被配置为显示检测到的手语的滚动抄本。在该配置中,例如,在用户暂时分心的情况下,可以快速重读被遗漏的单词或甚至句子。示意者的转换文本抄本可以显示为与电影中结束字幕的呈现类似的滚动文本。

[0179] 显示转换文本的抄本的系统可以以某种方式突出显示请求辅助信息的单词或短

语,例如,下划线、着色、粗体文本等。可以在检索或显示辅助信息之前显示这些突出显示。以该显示模式配置的一些实施例可以允许用户确认或取消突出显示的文本的请求。可选地或另外地,这些突出显示可以与辅助信息一起显示。该显示模式可以向用户明确辅助信息所关联的文本。系统可以允许用户通过UI交互选择当前或过去转换的文本,并调出或带回相关的辅助信息,如下面进一步描述的。

[0180] 可穿戴系统200可以放置转换的文本或辅助信息(例如,在文本气泡中或作为滚动抄本),以便最小化用户的眼睛移动,从而经由UI交互来访问信息。以这种方式,UI被简化并且用户不需要将他或她的注意力远离示意者。转换的文本或辅助信息可以被放置为使对话伙伴最低限度地看到阅读动作,并且这样做,提供较少的分心和更好的通信,同时不透露用户对转换的文本或辅助信息的访问。例如,能够确定示意者的位置的实施方式可以将转换的文本或辅助信息放置在示意者旁边。来自系统面向外成像系统464的图像可以帮助确定适当的放置,例如,不遮挡例如示意者的面部、手势等。可穿戴系统200可以使用图8中所示的处理流程来确定转换文本或辅助信息显示的放置。例如,框850中的识别对象可以是其手语将被处理以用于手语识别的示意者。

[0181] 作为减少用户或会话伙伴所经历的分心的另一示例,如果转换的文本或辅助信息以音频呈现(例如,在用户没有听力障碍并且会话伙伴使用手语的情况下),则可穿戴系统可以以足够大的音量向用户呈现信息,但是对于(语音障碍但没有听力障碍的)会话伙伴听不到,或者在用户和会话伙伴都不说话时呈现信息。

[0182] 显示的项目可以保持可见直到满足条件。例如,显示的项目可以在固定的时间量内保持可见,直到要显示下一个显示的项目,或者直到用户动作解除。用户动作可以是被动的(例如,由面向内的成像系统462捕获的眼睛运动)。可穿戴系统可以在确定用户已查看显示的项目时解除显示。例如,如果显示的项目是文本,则系统可以通过文本跟踪用户的眼睛运动(例如,从左到右或从上到下)。一旦可穿戴系统确定用户已经浏览了整个显示的项目(或所显示的项目的大部分),则可穿戴系统可以相应地解除所显示的项目。作为另一个示例,在系统观察到用户已经远离观看(或不看)所显示项目占据的区域之后,可以解除所显示的项目。用户动作也可以是主动的(例如,通过由面向外成像系统464捕获的手势,由音频传感器232接收的话音输入,或来自用户输入设备466的输入)。例如,一旦可穿戴系统检测到用户的猛击手势,可穿戴系统就可以自动解除所显示的项目。

[0183] 可穿戴系统可以被配置为支持针对特定用户的定制的一组用户界面(UI)交互。UI交互可以采取类似于按钮的UI元素的形式,该按钮利用用手指、某种指针或指示笔,通过凝视和随后用眼睛或其他方式固定在按钮上来致动。按钮可以是真实的物理按钮(例如,在键盘上)或由显示器220显示的虚拟按钮。UI交互可以采取头部姿势的形式,例如,如上面结合图4所描述的。以上结合图10描述了UI交互检测的示例。

[0184] 可穿戴系统可以提示用户延迟解除所显示的项目。例如,可穿戴系统200可以降低亮度或改变所显示项目的配色方案以通知用户所显示的项目将很快被解除,例如几秒钟。可以使用诸如上述那些的UI交互来推迟解除。例如,可穿戴系统可以检测到用户已经把目光远离所显示的项目。因此,可穿戴系统可以增加所显示项目的透明度,以通知用户所显示的项目将很快被解除。然而,如果可穿戴系统通过眼睛跟踪检测到用户回顾所显示的项目,则AR系统可以推迟解除。

[0185] 诸如上述那些的UI交互也可用于召回已被解除的显示项目。例如,通过用户输入设备的输入动作(例如,键盘上的退格的启动)可以用于召回最近显示的项目,或用于选择特定的显示项目以进行召回。

[0186] 感测眼镜系统的示例用户体验

[0187] 图13A示出了感测眼镜系统的示例用户体验,其中感测眼镜系统可以为可穿戴系统的用户解释(例如,由示意者用手势做出的)手语。该示例示出了感测眼镜系统的用户正在观察的示意者1301。用户可以感知示意者1301正在做出如场景1305、1310和1315中所示的手势序列1300。场景1305中的手势表示单词“how(如何)”;场景中的手势表示单词“are(是)”;以及场景1315中的手势表示单词“you(你)”。因此,序列1300可以被解释为“how are you(你好么)”。序列1320和1340示出与序列1300相同的手势。手势1305对应于手势1325和1345;手势1310对应于手势1330和1350;并且手势1315对应于手势1335和1355。然而,序列1300、1320和1340示出了如下面进一步描述的不同用户显示体验。

[0188] 为了将序列1300中的手势翻译为英语短语“how are you”,可穿戴系统200的面向外成像系统464可以捕获手势序列,或者作为一系列图像或者作为视频。可穿戴系统可以从一系列图像或视频中提取手势。可穿戴系统可以对提取的手势执行手语识别,例如,通过对象识别器708或应用深度学习算法。在该过程中或识别手语时,可穿戴系统可以访问存储在本地或远程存储器中的手语字典。可穿戴系统可以经由显示器220向用户(未示出)显示从识别的手语转换的文本(或标志的图形表示)。感测眼镜系统还可以接收与转换的标志相关联的辅助信息的请求,并且使用在此描述的技术来检索和显示辅助信息。

[0189] 在图13A所示的图形序列中,表达单词“how”采用两个不同的手势,例如,如图形1305和1310所示。可穿戴系统可以在显示单词“how”(如在场景1305做出的手势)之前等待直到第二手势之后(在场景1310中)。附加地或替代地,可穿戴系统可以延迟文本转换或显示直到完成句子或短语,例如以图形序列1320所示,其中在场景1335的末尾显示短语“how are you”。可穿戴系统可以将转换的文本或辅助信息显示为标题或文本气泡,例如,如图1355所示。标题或文本气泡可以定位在用户的FOV中以最小化对用户的分心,例如,紧密靠近示意者而不会遮挡用户看到示意者面部。

[0190] 图13B示出了感测眼镜系统的另一示例用户体验,其中呈现了目标语音和辅助信息两者。在该示例中,用户(未示出)可以佩戴HMD并感知示意者1362。示意者使用手语询问问题“PTO在哪里?”(示意者被描绘为在问题的结尾处做出字母“O”的手势)。可穿戴系统可以识别示意者做出的手势,将它们转换为文本,并在文本气泡1360中将转换的文本显示给可穿戴系统的用户。可穿戴系统可以确定“PTO”是首字母缩略词并且是用户在日常语音中不经常使用的词。例如,可穿戴系统可以维护常用单词和短语的字典,并确定“PTO”不在该字典中。在检测到单词“PTO”不在字典中时,可穿戴系统可以发起与短语“PTO”相关联的辅助信息的访问。

[0191] 可穿戴系统可以基于上下文信息检索关于首字母缩略词的辅助信息。在该示例中,系统可以依赖于其位置信息,例如,系统(及其用户)目前在弗吉尼亚州的亚历山大市。系统检索“专利商标局(Patent and Trademark Office)”作为首字母缩略词“PTO”的辅助信息。系统通过显示器220将辅助信息作为虚拟横幅1365显示给用户。图13B所示的转换的文本和辅助信息的显示模式仅用于说明。一些实施例可以不同地显示它们,例如,都在顺序

显示的标题中。

[0192] 多个可穿戴系统的多个用户可以借助于它们各自的可穿戴系统远程通信。图13C示出了远程呈现会话中的感测眼镜系统的示例用户体验。例如,如图13C所示,在两个物理位置1370a、1370b处的两个用户1372a、1372b(使得他们在没有人造设备的帮助下不直接看到或听到彼此)可以分别佩戴可穿戴设备1374a、1374b。一个或两个用户1372a、1372b可以使用手语进行交谈。手势可以由用户的相应可穿戴系统的成像系统捕获并通过网络1290发送。用户A的1372a手语可以作为转换的文本显示在用户B的1372b设备上,反之亦然。

[0193] 感测眼镜系统可以在本地将检测到的手语转换为文本并且通过网络1290仅发送转换的文本。其他用户的设备可以显示该文本,或者在其他用户没有听力障碍的情况下,将文本转换为可听语音。这在网络1290的带宽受到约束的情况下是有利的,因为与传输相应图像、视频或音频相比,传输文本需要更少的数据。

[0194] 可穿戴系统还可以通过在显示器220上呈现的图像来增强远程呈现会话。例如,显示器220可以呈现远程示意者的化身以及转换的文本或辅助信息以参与参与者的视觉感测。例如,配备有面向内成像系统464的可穿戴设备可以捕获用于替换由HMD遮挡的佩戴者的面部区域的图像,其可以被使用使得第一用户可以在远程呈现会话期间看到第二用户的未被遮挡的面部,反之亦然。可以将与第一用户相关联的世界地图信息通信给涉及感测眼镜系统的远程呈现会话的第二用户。这可以通过创建HMD佩戴者将要看到的远程用户的图像来增强用户体验。

[0195] 在远程呈现应用中,捕获图像信息由与用户示意者(例如,从第一人称视角来看)相关联的设备而不是与用户-观察者相关联的设备来执行,这在面对面情景中可能是典型的。可以通过与用户相关联的设备或者通过例如服务器计算机系统1220的远程系统,来执行检测手语的存在和将手语转换为文本。可以基于捕获图像的设备来确定手语的来源,例如,当用户A的设备捕获图像时,用户A正在做出手势。

[0196] 图13D示出了用于解释手语的示例虚拟用户界面。在该示例中,用户1392佩戴可穿戴设备1380(其可以包括可穿戴系统200的至少一部分)。在该示例中,用户1392在柜台后面并且感知到接近柜台的人1394。例如,用户1392可以是医疗机构中的护士或准入人员,帮助客人的酒店雇员(例如,门房)等等。人1394可能感到不适并寻求医疗关注,例如到药房的方向。可穿戴设备1380可以观察(例如,经由面向外成像系统464)用户1394的手势,如图13D所示。可穿戴设备1380可以自动地(例如,使用对象识别器708)检测所示的手势是手语的表达,识别与手势相关联的含义,并以用户1392理解的目标语言(例如,英语)提供手势的翻译。可穿戴设备1380可以呈现虚拟用户界面1382以显示由可穿戴设备捕获的输入1384a,与输入1384a对应的翻译1384b(例如,“附近是否有药房?我感觉不舒服。”)。可穿戴系统还可以在虚拟用户界面1382上提供用户输入要素1384c和1384d。例如,用户1392可以使用手势(例如,按压手势)来选择用户输入要素1384c。用户输入要素1384c的启动可以使可穿戴设备提供响应列表,例如附近药房的位置,或“我不知道”。在一些实施例中,可穿戴设备1380可以用手语示出相应图形用于响应。因此,用户1392可以使用如图形中所示的手势来响应人1394。作为另一示例,在用户输入要素1384d被启动的情况下,可穿戴系统可以提供选项列表,例如,解除用户界面要素1382,或电话求助等等。在一些实施例中,界面1382的区域1384a可以包括输出图形,示出用户1392可以执行以与人1394进行通信的手语手势(例如,

“药房在街对面”的标志)。

[0197] 感测眼镜系统作为人际通信工具的示例过程

[0198] 图14A和14B示出了利用感测眼镜系统促进人际通信的示例过程。图14A和图14B所示的示例过程1400和1440可以由图2A所示的可穿戴系统来执行。

[0199] 在框1404处,可穿戴系统可以捕获环境中的图像信息。如在此所述,可穿戴系统可以使用面向外成像系统464来捕获用户周围的图像信息。可穿戴系统还可以捕获环境中的音频信息。音频信息可以与面向外成像系统464获取的数据一起使用,以确定语音或手势的来源,或者检测手语的存在。

[0200] 在框1408处,可穿戴系统检测捕获的图像信息中的手语的存在。该检测处理可以在本地(例如,通过本地处理模块71)或远程(例如,通过远程处理模块72)完成。可穿戴系统可以使用各种对象识别器来检测手势的存在。例如,可穿戴系统可以发现手势序列可以构成手语中的短语或句子。作为另一个示例,可穿戴系统可以检测一系列手势以及嘴唇运动。可穿戴系统可以发现手势和嘴唇运动与手语相关联,因为这样的手势和嘴唇运动不伴有音频信息。

[0201] 在一些实施例中,可穿戴系统可以基于上下文信息来检测和解释手语。例如,可穿戴系统可以接收(例如,晚餐会话的)音频信号,将这些信号转换为语言,或从该语言中提取含义,从而推断出可以用于解释手语(例如,以与讨论主题一致的方式来解释手势)的讨论主题的类型(或其他属性)。

[0202] 可穿戴系统可以被配置为检测或忽略用户自己的手语。框1408的功能可以基于该配置而不同,因为可以在相对接近的距离处从第一人称视角捕获用户自己的手语。例如,如果系统被配置为捕获用户自己的手语,则可以打开向下指向用户手的附加面向外相机,或者可以将面向外成像系统配置成广角模式以捕获用户手的图像。

[0203] 在框1412处,系统确定是否检测到手语。如果检测到手语,则过程流1400继续到框1416。如果未检测到手语,则流程返回到框1408(如图所示)或框1404(未示出)。

[0204] 当可穿戴系统(包括其成像系统)被打开时或者当启用手语识别功能时,可以连续地或周期性地(例如,以采样频率)执行框1404到1412中的操作。这些操作可以与流程图1400和1440中的其他框并行执行(例如,作为由定时中断驱动的背景任务)。出于说明的目的,它们在过程流程顺序中示出为离散框。但它们不受所示顺序的限制。根据系统设计者的判断,除了上述示例之外的许多处理流程都是可能的。

[0205] 在框1416处,可穿戴系统可以确定检测到的手语的源(例如,示意者)。源可以是用户的物理附近的人、用户或用户感知的视觉内容中的人。手语的源可以是相关的,例如,如果系统被配置为处理仅从来自位于可穿戴系统的FOV的中心或附近的人的手语(例如,当多个人同时用手语交谈时,来自FOV中心之外的人的手语可以丢弃并且不再进一步处理)。作为另一示例,可穿戴系统可以仅针对用户正在看的人来处理用于手语识别的手势,用户正在看的人可以是或可以不是FOV中心的人。可穿戴系统可以基于由面向内成像系统462和面向外成像系统464获取的数据来识别用户正在看的人。例如,面向外的相机可以提供包括示意者相对于用户的位置的信息。面向内的相机可以提供包括用户正在看的方向的信息。通过使用来自两个相机的信息,可穿戴系统可以确定用户正在看的人,以及该人是否是手语的源。

[0206] 在框1420,系统确定手语的源是否已经发生变化。如果存在改变,则流程1400继续通过框1424到框1444,如图14B所示。如果手语的源没有改变,则流程移动到框1428以继续手语识别处理,其可包括捕获图像信息(框1404),检测手语的存在(框1408),以及图14B中所示的处理步骤。例如,如果系统确定继续来自相同示意者的手势继续,则除了继续捕获图像信息和检测手语之外,系统可以继续执行从框1448开始的功能。

[0207] 在框1448,可穿戴系统可以将手语翻译成用户理解的语言。例如,系统可以将识别的手语转换为文本,当系统显示时可以由用户读取该文本(例如,作为文本气泡或标题)。在一些情况下,如果用户理解不同的手语,则可以向用户显示其他示意者的标志的图形表示,例如,作为示出用用户自己的手语转换成标志的标志的图形。

[0208] 在框1452,当系统被配置为检测用户自己以及对话伙伴的手语时,示例系统可以确定检测到的手语是否是用户自己的手语。如果是,则过程进行到框1484,其中系统可以将转换的文本发送到观察者/对话伙伴的显示设备。

[0209] 从框1484,系统可以进行到框1488以继续处理。当系统被配置为忽略用户自己的手语时,可以从流程中省略框1452和1484。如果检测到的手语不是用户自己的手语,则流程继续到框1456。

[0210] 在框1456,可穿戴系统可以如上所述通过显示器显示转换的文本。在系统的用户没有听力障碍的情况下,除了视觉显示之外或作为其替代,文本可以例如通过音频放大器240以音频呈现。

[0211] 在框1460,可穿戴系统可以监控对关于转换的文本的辅助信息的请求。在检测到触发条件时,用户的可穿戴设备可以发送对辅助信息的请求。一些示例触发条件可以包括用户的指示,例如用户的手势或用户输入设备466的驱动;或者在检测到用户可能不理解的单词(或短语)时。

[0212] 在框1464,系统确定是否接收到请求。如果没有接收到请求,则流程移动到框1476,这将在下面进一步描述。

[0213] 如果接收到请求,则在框1468,系统可以检索与转换的文本(或其所请求的部分)相关联的辅助信息。如本文所述,可以基于上下文信息来确定和检索辅助信息,例如,用户的位置、语音的上下文或本文所述的其他类型的信息。

[0214] 在框1472,可穿戴系统可以经由可穿戴系统的显示器220显示检索到的辅助信息。在一些实施方式中,可穿戴系统可以在显示辅助信息之前解除转换的文本的显示。

[0215] 流程可以从框1464或1472进入框1476。在框1476,系统可以检测用于解除转换的文本或辅助信息显示的条件。当检测到这样的条件时,在框1480,系统可以解除转换的文本或辅助信息的显示并继续到框1488。在框1488,以类似于上面的框1428的描述的方式继续处理手语识别。

[0216] 类似于上面关于框1404到1412所指出的,流程图1440中的操作可以与流程图1400和1440中的其他框并行地执行。为了说明的目的,它们在处理流程顺序中被示为离散框,但它们不受所示顺序的限制。例如,系统可以显示对于转换的文本的辅助信息(在框1472),同时系统将附加的手语转换为文本(在框1448),执行辅助信息请求监控(在框1460),或者检索关于另一转换文本的辅助信息(在框1468)。作为另一示例,系统可以将手语转换为文本(在框1448),同时它检索对于先前请求的(针对辅助信息)转换文本的辅助信息(在框

1468)。系统设计人员可以自行决定许多其他处理流程。

[0217] 图14C是用于确定辅助信息并呈现与转换的文本相关联的辅助信息的示例方法的过程流程图。该过程1490可以在这里描述的可穿戴系统200或另一本身可以具有或不具有手语识别功能的计算设备上执行。该过程1490可适用于以下情况：使用一个感测眼镜系统检测手语并将手语转换为文本并在另一设备或系统上显示转换的文本更有利。示例情况可以是示意者希望与第二人远程通信的情况。示意者的可穿戴设备可以将示意者自己的手语转换为文本。可穿戴设备可以将转换的文本发送到第二人可观看的远程系统。由于转换的文本与对应的图像或视频相比可以以少得多的信息比特来发送，因此这种过程可以有利地要求来自传输介质的低得多的带宽或者导致更可靠的通信。

[0218] 过程1490开始于框1492，其中设备或系统正在执行某种处理，其可以与手语处理相关或不相关。在框1494，设备或系统可以确定是否从可穿戴系统接收到文本。如果否，则该过程可以返回到框1492。如果是，则该过程可以进行到框1496。在框1496，该设备或系统可以从可穿戴系统接收文本并渲染该文本。然后，该过程可以进行到框1456。在渲染设备包括HMD的情况下，渲染设备可以将文本呈现为覆盖在用户的物理环境上的虚拟内容。框1456至1480中的处理可以如上面结合图14B所述类似地执行。

[0219] 图15示出了利用感测眼镜系统促进人际通信的另一示例过程。示例过程1500可以由本文描述的可穿戴系统200的一个或多个组件（例如，通过本地处理和数据模块260、远程处理模块270、单独或组合）来执行。如参考图12所述，图15中描述的一个或多个步骤可以由不是用户的可穿戴系统的一部分的一个或多个计算设备来执行，例如，另一用户的可穿戴设备，或第三方的服务器系统。

[0220] 在框1510处，可穿戴系统可以辨别和识别环境中的语音。语音可以是手语的形式。例如，可穿戴系统可以分析由面向外成像系统464获取的数据，以识别作为手语的一部分的手势。可穿戴系统还可以分析由音频传感器232获取的音频数据，其可以包括用户环境中的人的语音。可穿戴系统可以使用对象识别器708识别语音。例如，可穿戴系统可以通过使用对象识别器分析手语的图像来识别短语或单词的存在。可穿戴系统还可以使用图7中描述的各种语音识别算法来识别音频数据。

[0221] 在框1520，可穿戴系统可以辨别目标语言。目标语言可以是可穿戴系统的用户用于通信的语言。例如，用户可以使用英语与其他人通信，而识别的原始语音（由另一用户使用）是手语。目标语言也可以是用户或可穿戴系统选择的语言。例如，用户可以选择ASL作为目标语言，因为即使用户说另一种语言，用户也可能想要使用手语与另一个人通信。作为另一示例，可穿戴系统可基于用户的位置自动选择语言。例如，可穿戴系统可以确定用户所在的国家，并选择该国家的官方语言作为目标语言。

[0222] 在框1530，可穿戴系统可以将检测到的语音转换为目标语言。可穿戴系统可以使用本文描述的例如字典翻译的各种技术来执行这种转换。

[0223] 在框1540，可穿戴系统可以确定与转换的语音相关联的视听内容，以呈现给可穿戴系统的用户。作为一个示例，视听内容可以包括目标语言中的文本。作为另一示例，视听内容可以是目标语言中的音频流，其中转换的语音是口语。作为又一示例，如果目标语言是手语，则视听内容可以是图形或动画。

[0224] 在可选框1550处，可穿戴系统可以将视听内容传送到头戴式显示器以供呈现。例

如,视听内容可以从一个用户的可穿戴设备传送到另一用户的可穿戴设备。在该示例中,第一用户的可穿戴设备可以捕获第一用户,将语音转换为目标语言,并将转换的语音传送给第二用户的可穿戴设备。

[0225] 示例感测眼镜系统作为与环境交互的工具

[0226] 除了识别另一个人的手势之外或作为其替代,本文描述的可穿戴系统还可以利用例如参考图7描述的各种文本识别算法来识别环境中的标志。可穿戴系统还可以修改文本(例如,修改显示特征或文本的内容)并将修改的文本渲染到用户的物理环境上。例如,可以渲染修改的文本以覆盖和遮挡原始文本,使得用户将感知修改的文本而不是原始文本。

[0227] 修改文本显示特征的示例

[0228] 图16A至图16E示出了感测眼镜系统的示例用户体验,该感测眼镜系统被配置为识别环境中的文本,修改与文本相关联的显示特征,并渲染修改的文本。参考图16A所示,用户210可以佩戴可穿戴设备(图16A中未示出)并且可以经由显示器220看到环境中的物理对象1606。可穿戴设备可以包括向外成像系统464,该面向外成像系统464可以捕获包括图像1602内的对象1606的图像1602。除了面向外成像系统464之外或作为其替代,可穿戴系统200可以使用其他传感器或设备捕获物理对象的图像。例如,用户输入设备466(例如,图腾)可以具有成像能力并且可以捕获包括对象1606的图像的图像1602。对象1606可以包括标志或其他对象,该标志或其他对象在其上或其中可以包含书写、字母、符号、字符1610。例如,字母可以写在对象上;或由对象成形、嵌有对象或嵌入对象中。文本也可以是一系列静态或闪烁的灯;或一个或多个物理对象的排列。在图16A至图16E所示的示例中,对象1606是交通停车(stop)标志。在其他示例中并且没有限制,对象1606可以是任何类型的标示(例如,商业或公共显示标志)、书籍、杂志、纸张、计算机显示屏幕、电视屏幕等等。

[0229] 可穿戴系统200可以使用例如如参考图7所描述的一个或多个对象识别器708来分析图像1602并识别对象1606。作为一个示例,可穿戴系统可以识别出对象1606是交通标志(例如,基于对象1606的形状,图16A中的八边形)。作为另一个示例,可穿戴系统可以识别对象1606中文本的存在。可穿戴系统可以识别文本而不管文本的格式(例如,文本是在对象上还是由投射文本的一系列的灯(例如,霓虹灯、LED灯等)来呈现。

[0230] 如将参考图18进一步描述,在某些实施例中,可穿戴系统200可以识别文本的含义并将文本从原始语言转换为目标语言。例如,可穿戴系统200可以识别来自各种语言的字母、符号或字符,例如英语、中文、西班牙语、德语、阿拉伯语、印地语等,并且将文本从原始显示的语言翻译为另一种语言。在一些实施例中,这种翻译可以根据先前指定的设置(例如,用户的偏好或用户的人口统计或地理信息)自动发生。在一些实施例中,可以响应于来自用户的命令(例如,口头或手势)来完成翻译。

[0231] 可穿戴系统200可以使用对象识别器708来分析文本1610的特征。例如,可穿戴系统200可以识别与文本1610相关联的字体大小或字体。可穿戴系统可以调整文本1610的特征以生成修改的文本。例如,可穿戴系统200可以调整文本1610的大小以放大或缩小文本1610。修改的文本的大小可以部分地取决于从眼睛210到原始文本1610的距离或用户的特征。例如,如果文本1610远离用户,则可穿戴系统可以放大文本1610。作为另一示例,根据用户的眼睛能力,系统可以确定如何调整文本的大小。可穿戴系统可以基于先前从用户获取的信息来确定人眼睛能力。例如,用户可以输入是否存在眼睛的视力问题。可穿戴系统还可

以对用户执行视觉测试(例如,通过在不同深度平面和以不同的大小显示虚拟对象以确定用户是否可以清楚地感知虚拟对象)来确定用户的眼睛能力。基于用户的眼睛能力,可穿戴系统可以基于文本的特征(例如,距离/位置、颜色、大小、字体等)确定用户是否可能感知文本1610。例如,如果可穿戴系统确定用户不能清楚地感知文本(例如,当文本失焦时),则可穿戴系统可以放大或加粗文本。如果用户近视,但文本远离用户,则可穿戴系统可以放大文本的大小,以使用户可以更容易地感知文本。尺寸调整可以对应于近视程度。如果用户的近视程度较大,则尺寸可以与较大的增加相关联,而如果用户的近视程度较小,则尺寸可以与较小的增加相关联。如本文进一步描述的,可穿戴系统还可以基于用户的眼睛能力来改变修改的文本的显示位置。参考图3,显示系统220可以包括多个深度平面,在用户远视但文本靠近用户的情况下,可穿戴系统可以在比原始深度平面距离用户更远的深度平面306处渲染修改的文本,使得修改的文本看起来远离用户。可以通过改变文本的字体大小来进行大小调整(例如,将文本识别为字符串的情况下)。尺寸调整还可以通过放大或缩小(例如,数字缩放)包含文本1610的图像1602的一部分来进行(例如,将文本分析为图像而不是文本串)。

[0232] 可穿戴系统200可以将修改的文本渲染给用户。继续参考图16A,佩戴HMD的用户可以看到包含对象1606的渲染版本1618的虚拟图像1614(由HMD渲染)。在一些实施方式中,对象1606的渲染版本1618可以遮挡原始文本。如图16A所示,渲染文本1622是“STOP(停止)”并且与原始文本1610相比被放大。HMD可以渲染放大文本1622覆盖在原始文本1610上,因此用户可能不会感知原始文本1610。在该示例中,通过增加文本大小,用户有利地可以更容易地感知、理解和响应实际上可能更小并且更难以感知的下面的文本1610。

[0233] 图16B示出了修改用户环境中的文本的特征的另一示例。如所渲染的文本1634所示,可穿戴系统200可以粗体化原始文本1610的字体。除了粗体化字体之外或者作为替代,可以对原始文本1610进行其他改变,例如,改变文本颜色、阴影、轮廓、格式(例如,斜体、下划线、对齐、调整等),等等。可穿戴系统200可以添加(或修改)与文本1610相关联的图形元素,诸如使得渲染文本1634闪烁、旋转等。

[0234] 图16C示出了将修改的文本与焦点指示符1640一起渲染的示例。焦点指示符1640可以包括视觉效果,诸如靶心、十字准线、光晕、颜色、感知的深度变化(例如,使渲染文本看起来更靠近),文本背景中的添加或改变、动画或引起用户注意的其他视觉效果。在图16C所示的示例中,可穿戴系统200可以被配置为将焦点指示符1640显示为背景1650,以背景1650为背景来渲染文本1638。背景1650可以包括边界区域1642和内部区域1646。边界区域1642可以绑定内部区域1646。在所示的实施例中,虚拟字母1638显示在内部区域1646内。文本背景1650可以渲染在显示的图像1614中,使得文本背景1650是与用户在没有HMD的情况下看到的不同的背景。在一些实施例中,内部区域1646和边界区域1642中的一个或多个是单色的(例如,白色、黑色或灰色)。系统可以改变背景1650,使得用户看到渲染文本1638而不是原始文本1610。例如,背景可以是不透明的,使得它可以遮挡原始文本1610。处理电子器件还可以被配置为显示背景1650,使得它混合到图像1614的其余部分中。例如,背景1650可以与图像1614的其余部分具有相同的颜色和纹理效果。可穿戴系统还可以以突出文本1638或背景1650的方式显示背景1650和文本1638,例如在文本1638或背景1650周围显示光晕。在这种情况下,背景1650可能无法无缝地与图像1614的其余部分成为一体。例如,内部区域

1646可以由边界区域1642轮廓出来以强调背景1650或者文本1638。

[0235] 在某些情况下,原始文本的视觉外观可能不清楚,例如,由于用户和对象1606之间的环境影响(例如,下雨、起雾)。图16D示出了修改文本的特征并渲染修改的文本以便更清晰的示例。在该图中,文本1626对于用户210看起来是模糊的。由于各种原因,文本可能被感知模糊。例如,视力不佳的用户可能难以在特定距离处清楚地看到文本。患有近视的用户可能会发现附近的文本图像看起来相对清晰,而出现在远处的文本则看起来模糊。同样地,那些有远视的人可以清楚地看到出现在远处的文字,而很难聚焦到出现在附近的文本。但眼睛状况可能不是图像看起来模糊的唯一原因。看起来更靠近或更远离眼睛210的文本可以针对也可能看起来模糊而进行调整。如果文本看起来相对于用户快速移动,则文本1626可能看起来模糊。上述其他因素,例如气候或天气因素,以及获取图像的相机的分辨率也可能起作用。

[0236] 在该示例中,可穿戴系统200可以使模糊文本1626或其他难以阅读的文本更清楚或更清晰。在文本看起来对于用户而言模糊但在可穿戴系统接收的图像中不是模糊的情况下,可穿戴系统可以分析由面向外成像系统464或另一设备(例如,用户输入设备466或可穿戴系统外部的相机,诸如仪表板相机(dash cam))获取的图像,以使用参考图13A描述的类似技术来识别文本1626。可穿戴系统可以虚拟地渲染文本,如文本1630所示。在某些实施方式中,可穿戴系统可以基于用户的或环境的条件来调整虚拟文本1630的特征。例如,在用户近视的情况下,可穿戴系统可以放大文本1626的字体或将文本渲染为看起来更靠近用户(例如,在更近的深度平面上)。作为另一示例,当环境较暗时,可穿戴系统可以增加文本1630与虚拟图像1614的其他区域之间的对比度。

[0237] 在一些情况下,文本1626看起来是模糊的,因为可穿戴系统获得的图像1602是模糊的(例如,由于快速驾驶速度或者当相机的分辨率低时)。如本文所述,可穿戴系统可以使用对象识别器708来识别模糊文本1626的存在。例如,可穿戴系统可以确定对象1606中或对象1606上存在文本的可能性。在一些情况下,如果可能性超过阈值,则可穿戴系统可以使用例如参考图7描述的一个或多个文本识别算法,例如OCR算法,从而识别最可能对应于模糊文本1626的字母1630。

[0238] 图16E示出了当原始文本1610由于障碍物1654而部分难以辨认时的场景。如图所示,障碍物1654覆盖原始图像1602中的原始文本1610的一部分。然而,障碍物1654可采用一种或多种形式。例如,障碍物1654可以是眼睛210或显示器与图像1602之间的一些物理障碍物,例如杆、建筑物等。障碍物1654也可以是环境或天气障碍物,例如上面所描述的那些。障碍物1654也可以在对象1606上(例如,文本1610的一部分被标志1606上的另一对象遮挡,或者文本1610的一部分被擦除、丢失或被贴纸覆盖)。这可以包括,例如,已经积累灰尘或污垢的表面,对找到书写1610的对象1606的表面的损坏,墨水斑点(例如,来自打印机),原始文本1610中的失真,或者任何其他类似的障碍物1654。

[0239] 系统可以在确定原始文本1610所说的内容时使用上下文信息(在本文中有时也称为上下文线索)。这里描述的各种上下文线索可以由可穿戴系统单独使用或组合使用,以确定文本1610的全部文本。示例上下文线索是用户的位置。例如,如上所述,GPS系统37(参见图2B)可以获取用户的位置数据,并且基于位置数据,可穿戴系统可以提供关于文本的语言是什么的初始猜测。在适用的情况下,在一些实施例中,可穿戴系统可以在可见光谱外的波

长(例如,红外线、紫外线)从一个或多个光源接收的信号获得附加信息。例如,可穿戴系统可以向标志1606发射紫外光,以显示仅在紫外光下可见的标示信息(或可检测来自标示的另一光源(例如,太阳)反射的紫外光)。在一些实施例中,系统可以访问单词数据库,系统可以通过该数据库检查原始文本1610的可见部分。在这样的示例中,可穿戴系统200可以能够确定哪些字母或单词的候选最有可能。例如,如图16E所示,系统推断字母拼写“STOP”部分是由于对象1606的八面体形状或对象1606的红色(未示出)。

[0240] 可穿戴系统可以能够依赖周围的单词、符号、标点符号或字符作为上下文线索来确定原始文本1610所说的内容。在某些实施例中,系统能够使用例如机器学习技术来识别位置特定的上下文线索。例如,系统可以能够检测到正在街道上行驶的用户并且可能将文本的识别偏向于经常在街道标志上使用的单词。可穿戴系统可以包括数据库,该数据库可以由本地处理和数据模块270或远程处理模块280访问(参见例如图2A)。数据库可以存储与用户参与的特定活动(例如,滑雪)相关联的单词类别、用户的地理位置、用户的行进速度、用户的海拔高度、由系统接收到的环境噪声的音量或类型、系统接收的区域中的可见光或其他光的级别或类型、系统周围的温度或气候、文本距用户的感知距离、或系统拾取的另一方所说的话的类别或多个类别。在一些实施例中,可穿戴系统可以使用该信息作为上下文线索,从而根据上述一个或多个关联来更准确地研究用户看到的文本的单词或语言的更可能的候选。在一些实施例中,可穿戴系统可以使用机器学习算法(例如,深度神经网络)来“学习”在各种情况下的先前单词并基于当前情况识别可能的单词。因此,通过执行该学习,可穿戴系统200可以变得特定于用户的行为并且可以更快速或更有效地确定文本。

[0241] 在图16A至图16E所描述的示例中,系统可以基于原始字母看起来与用户的感知距离来确定在哪个深度处显示文本。可以使用各种技术来测量原始字母和用户之间的感知距离,诸如通过应用立体视觉算法(例如,对于由面向外成像系统获取的数据)或通过分析由深度传感器(例如,激光雷达)获取的数据。立体视觉算法可以包括块匹配算法、半全局匹配算法、半全局块匹配算法、视差图、三角测量、深度图、神经网络算法、同时定位和地图构建算法(例如,SLAM或v-SLAM)等。被感知为靠近用户的字母可以在显示系统220上的近深度处显示。在一些实施例中,看起来比距用户的第一距离阈值(例如,大约800cm)看起来更近的字母在第一深度处显示在系统上。在一些实施例中,第一距离阈值是200cm,使得看起来比约200cm更近的字母显示在第一深度处。在一些实施例中,第一距离阈值为约80cm。字母是否在第一深度处显示或者使用哪个第一距离阈值可以取决于许多因素。一个因素可能是系统能够显示多少个不同的深度。例如,如果实施例仅在两个不同深度处显示对象,则可以使用较短的第一距离阈值,而当实施例可以在更多数量的不同深度处显示文本时,可以使用较小的范围。例如,如果用户正在阅读报纸,则系统将感知文本靠近用户,因此报纸上的字母将显示在系统上就像在近深度处一样。如图3所示,显示系统220可以包括多个深度平面306,该多个深度平面306可以使虚拟对象看起来在距用户不同距离处。在某些实施方式中,可穿戴系统可以基于用户的眼睛能力来调整修改的文本的渲染位置。例如,在用户近视的情况下,可穿戴系统可以在比文本最初对应的深度平面更靠近用户的深度平面处渲染修改的文本。作为另一个例子,在用户远视的情况下,可穿戴系统可以在比原始文本出现的位置更远离用户的深度平面处渲染修改的文本。

[0242] 类似地,被感知为远离用户的字母可以在远深度处显示在显示系统上。在一些实

施例中,看起来距离用户比约第二距离阈值更远的字母在看起来比第一深度更远的第二深度处显示在系统上。在一些实施例中,第二距离阈值约为300cm。在一些实施例中,第二距离阈值约为600cm。在一些实施例中,第二距离阈值约为10m。例如,在驾驶的同时在广告牌上看到的文本可以渲染在第二深度处。

[0243] 在各种实施例中,第一和第二距离阈值之间的差异可以是不同的。差异的大小可以基于许多因素,例如,系统可以在多少深度处显示文本,系统的能力感知与现实世界对象或文本的距离的精度或准确度,或者手动或出厂设置是什么。在一些实施例中,差异小于100m。在一些实施例中,差异小于700cm。在一些实施例中,差异小于30cm。在某些实施例中,差异为零(例如,第一距离阈值和第二距离阈值相同)。

[0244] 在一些实施例中,系统可以处理负差异。也就是说,存在一些重叠,其中对象或文本符合在第一深度和第二深度处显示的标准。在这样的实施例中,可穿戴系统可以使用上下文线索来确定哪个深度将为用户提供最无缝的观看体验。例如,最初看起来靠近用户但正快速离开用户的对象可能最初符合要在第一深度处显示的标准。然而,系统可以确定由于对象的位置的轨迹,它将在第二深度处显示对象。

[0245] 可穿戴系统的一些实施例能够在三个或更多个深度处显示文本。在这种情况下,可以包括与第一和第二深度之间的第三、第四等深度对应的中间距离阈值或距离范围。例如,在一些实施例中,当字母看起来例如在距显示器220约100cm到300cm之间时,可以将文本渲染在第三深度处。

[0246] 可穿戴系统200可以被配置为自动地或响应于用户输入来辨别或识别来自图像的文本。在自动辨别文本的实施例中,用户可以浏览具有文本的图像,并且系统可以在没有用户命令的情况下辨别和显示如本文所述的文本。在响应于用户输入辨别文本的实施例中,用户可以使用各种命令来启动文本的辨别或显示。例如,命令可以是口头提示、手势、头部动作(例如,点头)、眼睛运动(例如,眨眼)等。

[0247] 修改文本显示特征的示例过程

[0248] 图17示出了用于促进用户与环境的交互的感测眼镜的示例过程。过程1700可以由可穿戴系统200的一个或多个组件(例如,通过本地处理和数据模块260、远程处理模块270,单独或组合)来执行。

[0249] 在框1704处,可穿戴系统可以通过一个或多个相机接收光信号。一个或多个相机可以是面向外成像系统464的一部分,或者是另一计算设备(例如仪表板相机或用户输入设备466)的一部分。

[0250] 在框1708处,可穿戴系统可以包括从信号中识别图像。例如,可穿戴系统可以将光学信号转换为人类可读图像。在一些实施例中,从信号中识别图像还可以包括识别图像的内容,例如,使用一个或多个对象识别器708对图像执行光学字符识别(OCR)。在某些实施例中,光学字符识别过程包括识别一个或多个字母或字符的文本或语言的可能候选。光学字符识别过程可以使用各种上下文信息(例如,上下文线索)来执行识别。一些示例上下文信息可以包括用户或用户附近的某人所参与的活动、用户的地理位置、用户的当前行进速度、用户的当前海拔高度、系统检测到的环境噪声的音量或类型、显示系统检测到的区域内可见光或其他光的级别或类型、显示系统检测到的温度或气候、字符或字母与用户的感知距离,或由显示器检测到的单词的类别或类型。

[0251] 继续参考图17,过程1700还可以包括确定图像是否包括字母或字符,如框1712所示。在一些实施例中,如果过程1700确定图像不包括字母或字符,则该过程可以返回到框1704。如果过程1700确定图像包括字母或字符,则该方法继续到框1716。

[0252] 在框1716,可穿戴系统可以将字母或字符转换为文本。这可以包括,例如,以不同于第一语言的第二语言显示文本(如参考图18和图19进一步描述)。在一些实施例中,可以响应于从用户或另一个人接收输入或命令来完成将一个或多个字母或字符(来自图像)转换为文本。这样的输入或命令可以包括各种模式,例如,口头命令、手势、头部的运动、或者用户眼睛的一个或多个运动。这些示例不应视为限制。

[0253] 在框1720处,可穿戴系统可以指示显示器将文本渲染在距离用户的多个深度的第一深度处。在一些实施例中,显示文本包括通过光学透射目镜将光作为图像传输给用户。目镜可以是本文所述的任何目镜。例如,可以将光引导到用户的眼睛中以在眼睛中形成图像。可穿戴系统可以使用如本文所述的光纤扫描投影仪或其他投影仪。在一些实施例中,该方法可以从GPS系统37(参考图2B描述)接收位置数据。该位置数据可以用于帮助系统推断从图像提取的文本,如本文中参考图16A至图16E进一步描述。

[0254] 可穿戴系统还可以修改文本并渲染修改的文本(例如,将光从显示器220投射到用户的眼睛)。例如,该方法可以以不同的字体、字体大小、颜色、背景或背景颜色、格式、清晰度、语言或亮度显示相对于原始字母或字符的文本。在一些实施例中,该方法可以包括使文本动画化或合并与文本交互的虚拟对象。

[0255] 修改标示内容的示例

[0256] 除了修改文本的显示特征之外或作为其替代,可穿戴系统还可以修改文本的内容,例如通过将文本从一种语言翻译成另一种语言,并显示修改的文本。图18示出了通过修改标示的内容来帮助用户理解物理环境中的标示的示例,其中标示从本地语言翻译成可穿戴系统的用户能够理解的目标语言。

[0257] 图18示出了两个场景1800a和1800b。场景1800a是用户在不佩戴本文所述的HMD的情况下感知到的。场景1800b是用户在佩戴HMD时可以感知到的(例如,通过显示器220,而不需要所描述的翻译过程)。如图所示,场景1800a和1800b都包括街道1802和行人1804。场景1800a还示出了包括简体中文字符的街道标志1810a和1820a。标志1820a还包括英文字符。然而,HMD的用户(图18中未示出)可能是英语使用者并且可能不理解中文字符。有利地,在一些实施例中,可穿戴系统可以自动识别街道标志1810a和1820b上的文本,并将街道标志的外语文本部分转换为用户理解的语言。可穿戴系统还可以将翻译的标示作为虚拟图像呈现在物理标志上,如场景1800b中所示。因此,用户不会在标志1810a、1820a中感知到中文文本,而是会感知标志1810b、1820b中所示的英文文本,因为HMD显示具有足够亮度而使下面的中文文本不被感知的虚拟图像(具有英文文本)。

[0258] HMD(例如,可穿戴系统200)可以使用与参考图16A至图17描述的类似技术来辨别用户环境中的标志并识别标志。在一些情况下,可穿戴系统200可以被配置为仅翻译标志的部分。例如,可穿戴系统200仅翻译具有中文文本的标志1820a的部分而不翻译具有英文文本(“GOLDSTAR”)的标志1820a的部分,因为用户可以理解英语部分(例如,因为它是用户的目标语言)。然而,在用户是双语者使得用户可以阅读英语和简体中文二者的情况下,可穿戴系统200可以被配置为不将标志1810a和1820a上的任何文本翻译成标志1810b和1820b。

[0259] 如参考图16A至图16E所示,可穿戴系统200可以被配置为调整标志的显示特征。例如,由标志1820a的中文部分的翻译产生的文本可能比标志1820a上的原始中文字符更长。结果,可穿戴系统可以减小翻译文本的字体大小(例如,“shopping center(购物中心)”),使得渲染的文本(如标志1820b中所示)可以适合原始标志的边界。

[0260] 尽管图18显示简体中文和英文字符,这仅用于说明而非限制。由可穿戴显示系统200的实施例识别和转换的语言可以包括任何语言,例如英语、中文(简体或繁体)、日语、韩语、法语、西班牙语、德语、俄语、阿拉伯语、罗曼语、印欧语言、汉藏语言、亚非语言、希伯来语、马来-波利尼西亚语等。

[0261] 修改标示内容的示例过程

[0262] 图19示出了帮助用户理解物理环境中的标示的示例过程。示例过程1900可以由可穿戴系统200的一个或多个组件(例如,通过本地处理和数据模块260、远程处理模块270,单独或组合)来执行。

[0263] 在框1910处,可穿戴系统可以接收用户环境的图像。图像可以由面向外成像系统464、用户输入设备466或可穿戴系统外部的另一设备上的相机捕获。图像可以是静止图像、视频帧或视频。

[0264] 在框1920,可穿戴系统可以分析图像以识别用户环境中的标示。可穿戴系统可以使用对象识别器708来执行这种识别。例如,对象识别器708可以检测对象上文本的存在,从而将对象分类为标志或者可以识别标示的规则边界(例如,图18中的矩形标志1810a、1810b)。

[0265] 在框1930,可穿戴系统可以识别标示1930上的文本。例如,可穿戴系统可以确定标志上存在哪些字符或字母。作为另一示例,可穿戴系统可以确定文本所处的语言。可穿戴系统可以基于与用户或标志相关联的上下文线索来做出这样的确定,例如,用户的位置、句法、语法、文本的拼写等。在框1930,可穿戴系统可以进一步确定文本的含义(例如,通过在字典中查找)。

[0266] 在框1940,可穿戴系统可以将文本的至少一部分转换为目标语言。可以基于用户的偏好或用户的人口统计信息来确定目标语言。例如,目标语言可以是与用户的原籍国相关联的官方语言、用户的母语、用户最常使用的语言、或者用户已经说过的语言(例如,在对可穿戴系统的语音命令中,或与另一用户的对话中)等。目标语言也可以根据用户的偏好来设置。例如,即使用户的母语是法语,用户也可能更喜欢将标志翻译成英语。

[0267] 在可选框1950,可穿戴系统可以修改与文本相关联的显示特征。例如,可穿戴系统可以将焦点指示符添加到文本(或与文本相关联的背景),以及更改文本的字体大小或颜色。参考图16A至图17进一步描述显示特征的示例修改。

[0268] 在可选框1960处,可穿戴系统可以使文本由混合现实设备以目标语言渲染。MR设备可以是这里描述的HMD。在修改显示特征的情况下,可穿戴系统还可以使得修改的显示特征被渲染。在仅将文本的一部分翻译成目标语言的情况下,可穿戴系统可以仅显示翻译文本的一部分或者显示翻译部分和未翻译的原始文本部分。可以在物理标示上的原始文本上渲染修改的文本,使得原始文本可以从用户的视图被遮挡。

[0269] 尽管图18和图19中的示例参考翻译在标示上文本进行描述,类似的技术也可以应用于体现在其他类型的媒体(例如,书籍、电视、计算机监视器等)中的文本。

[0270] 与手语有关的其他方面

[0271] 以下进一步提供感测眼镜在手语中的应用的其他方面。

[0272] 在第1方面,一种用于通过增强现实系统提供从手语转换而成的文本的方法,该方法包括:在包括成像系统的增强现实 (AR) 系统的控制下:经由成像系统捕获图像信息;检测图像信息中的手势,这些手势是手语识别的候选;识别检测到的手势中的手语;将识别出的手语转换为文本;并显示转换的文本。

[0273] 在第2方面,根据方面1的方法,还包括:接收关于转换的文本的辅助信息的请求;检索与所请求的转换文本相关联的辅助信息;使用AR系统显示辅助信息;检测解除转换的文本或辅助信息的显示的条件;以及解除转换的文本或辅助信息的显示。

[0274] 在第3方面,根据方面2的方法,其中,用于解除转换的文本或辅助信息的显示的条件基于用户界面交互。

[0275] 在第4方面,根据方面3的方法,其中,用户界面交互至少部分地基于AR系统的用户的眼睛运动。

[0276] 在第5方面,根据方面2至4中任一方面的方法,其中,用于解除转换的文本或辅助信息的显示的条件至少部分地基于持续时间。

[0277] 在第6方面,根据方面2至5中任一方面的方法,其中,用于解除转换的文本或辅助信息的显示的条件至少部分地基于附加手语手势的转换或附加辅助信息的接收。

[0278] 在第7方面,根据方面2至6中任一方面的方法,还包括:检测用于重新显示已解除显示的转换的文本或辅助信息的条件;以及重新显示已解除显示的转换的文本或辅助信息。

[0279] 在第8方面,根据方面1至7中任一方面的方法,其中,将识别的手语转换为文本包括应用深度学习技术。

[0280] 在第9方面,根据方面8的方法,其中,深度学习技术包括神经网络。

[0281] 在第10方面,根据方面1至9中任一方面的方法,其中,AR系统在手语识别和文本转换中使用手语字典。

[0282] 在第11方面,根据方面1至10中任一方面的方法,其中,AR系统识别对AR系统的用户是国外的手语。

[0283] 在第12方面,根据方面11的方法,其中,AR系统通过处理候选手语的列表来识别手语,该列表至少部分地基于AR系统的位置而被优先化。

[0284] 在第13方面,根据方面11至12中任一方面的方法,其中AR系统通过处理候选手语的列表来识别手语,该列表至少部分地基于在AR系统的环境中检测到的口语被优先化。

[0285] 在第14方面,一种用于翻译手语的增强现实 (AR) 装置,包括:AR显示器;成像系统;数据存储器,被配置为存储计算机可执行指令和数据;以及与数据存储器通信的处理器,其中计算机可执行指令在被执行时使处理器进行:接收由成像系统捕获的图像信息;检测接收到的图像或视频信息中的手势;识别检测到的手势中的手语;将识别的手语翻译成AR装置的用户理解的语言;以及使用AR显示器显示与翻译的手语相关的信息。

[0286] 在第15方面,根据方面14所述的装置,其中,所述计算机可执行指令在被执行时还使所述处理器进行:接收关于所翻译的手语的辅助信息的请求;检索与所请求的手语相关的辅助信息;使用AR显示器显示检索到的辅助信息。

[0287] 在第16方面,根据方面14至15中任一方面的装置,其中处理器通过将接收到的图像信息通过通信网络发送到远程处理器以便远程处理器检测手势并识别手语,来检测手势并识别手语。

[0288] 在第17方面,根据方面14至16中任一方面的装置,其中成像系统包括多个相机或广角相机。

[0289] 在第18方面,根据方面14至17中任一方面所述的装置,其中,处理器还被配置为:确定检测到的手势的源;并且在确定检测到的手势的源是AR装置的用户时,将翻译的手语发送到另一个设备以进行显示。

[0290] 在第19方面,根据方面14-18中任一方面的装置,还包括音频放大器,并且处理器还被编程为通过音频放大器以音频呈现翻译的手语。

[0291] 在第20方面,根据方面19的装置,其中,处理器还被配置为通过音频放大器以音频呈现辅助信息。

[0292] 在第21方面,根据方面14至20中任一方面的装置,其中AR装置的用户理解的语言包括与识别出的手语不同的手语。

[0293] 在第22方面,一种用于促进涉及一种或多种手语的远程通信的增强现实(AR)系统,包括:多个可穿戴AR设备,每个可穿戴AR设备包括:AR显示器;成像系统;以及用于通过通信网络进行通信的通信系统;一个或多个数据存储器,被配置为存储计算机可执行指令和数据;以及与数据存储器通信的一个或多个处理器,其中计算机可执行指令在被执行时将一个或多个处理器配置成:接收由多个可穿戴AR设备中的第一可穿戴AR设备的成像系统捕获的图像信息;检测接收到的图像信息中的手语;将检测到的手语转换为文本;通过通信网络将转换的文本发送到多个可穿戴AR设备中的第二可穿戴AR设备;在第二可穿戴AR设备的AR显示器上显示转换的文本。

[0294] 在第23方面,根据方面22的系统,其中第二可穿戴AR设备还显示第一用户的世界地图。

[0295] 在第24方面,根据方面23的系统,其中第一用户的世界地图包括第一用户的化身。

[0296] 在第25方面,根据方面22至24中任一方面的系统,其中,多个可穿戴AR设备中的每一个可穿戴AR设备包括一个或多个数据存储器和一个或多个处理器,并且处理器功能由本地处理器执行。

[0297] 在第26方面,一种用于手语识别的可穿戴系统,该可穿戴系统包括:头戴式显示器,被配置为向用户呈现虚拟内容;成像系统,被配置为对用户的环境成像;以及硬件处理器,其与头戴式显示器和成像系统通信,并被编程为:接收由成像系统捕获的图像;利用对象识别器检测图像中的手势;识别手语中的手势的含义;基于与用户相关联的上下文信息辨别目标语言;基于识别出的含义将手势翻译为目标语言;至少部分地基于将手势翻译成目标语言来生成虚拟内容;以及使头戴式显示器向用户渲染虚拟内容。

[0298] 在第27方面,根据方面26的可穿戴系统,其中成像系统包括配置成对于用户周围进行成像的一个或多个广角相机。

[0299] 在第28方面,根据方面26至27中任一方面的可穿戴系统,其中,硬件处理器还被编程为访问与手势相关联的辅助信息;并且由头戴式显示器渲染的虚拟内容包括辅助信息。

[0300] 在第29方面,根据方面26至28中任一方面的可穿戴系统,其中,基于与用户相关联

的上下文信息来辨别目标语言,硬件处理器被编程为:基于以下中的至少一个将目标语言设置为用户理解的语言:由可穿戴系统捕获的用户的语音、用户的位置、或者用户选择语言作为目标语言的输入。

[0301] 在第30方面,根据方面26至29中任一方面的可穿戴系统,其中,硬件处理器被编程为确定目标语言是否是口语;并且响应于确定目标语言是口语,以目标语言播放与翻译的手势相关联的语音的音频流。

[0302] 在第31方面,根据方面26至29中任一方面的可穿戴系统,其中硬件处理器被编程为确定目标语言是否是另一种手语;并且响应于确定目标语言是另一种手语,将另一种手语中的另一种手势的图形呈现为手势的翻译。

[0303] 在第32方面,根据方面26至31中任一方面的可穿戴系统,其中,为了识别手语中的手势的含义,硬件处理器被编程为在由成像系统捕获的图像的一部分上应用深度神经网络技术。

[0304] 在第33方面,根据方面26至32中任一方面的可穿戴系统,其中,硬件处理器还被编程为至少部分地基于用户的位置从候选手语的列表中辨别手语。

[0305] 在第34方面,根据方面26至33中任一方面的可穿戴系统,其中,基于识别出的含义将手势翻译为目标语言,硬件处理器被编程为将手势转换为目标语言的文本表达。

[0306] 在第35方面,根据方面26至34中任一方面的可穿戴系统,其中,硬件处理器被编程为确定检测到的手势的源;并且在确定检测到的手势的源是可穿戴系统的用户时,将目标语言的手势的翻译传送给另一用户的可穿戴设备。

[0307] 在第36方面,根据方面26至35中任一方面的可穿戴系统,其中,硬件处理器被编程为检测由头戴式显示器解除对于虚拟内容的显示的条件,以及响应于检测到条件由头戴式显示器移除对于虚拟内容的显示。

[0308] 在第37方面,根据方面36的可穿戴系统,其中该条件包括以下中的至少一个:持续时间、用户的手势或来自用户输入设备的输入。

[0309] 在第38方面,根据方面26至37中任一方面的可穿戴系统,其中,图像包括视频的一个或多个帧。

[0310] 在第39方面,一种用于手语识别的方法,该方法包括:接收由成像系统捕获的图像;分析图像以检测用户的手势;至少部分地基于检测到的手势来检测手语的通信的存在;识别手语中手势的含义;辨别手势将被翻译成的目标语言;基于识别出的含义将手势翻译为目标语言;至少部分地基于将手势翻译成目标语言来生成虚拟内容;并使头戴式显示器向用户渲染虚拟内容。

[0311] 在第40方面,根据方面39的方法,其中从第一可穿戴设备接收图像,该第一可穿戴设备被配置为呈现混合现实内容,同时将虚拟内容传送到第二可穿戴设备以进行渲染,其中第一可穿戴设备和第二可穿戴设备被配置为向用户呈现混合现实内容。

[0312] 在第41方面,根据方面39的方法,其中基于识别出的含义将手势翻译成目标语言包括将手势转换为目标语言的文本表达。

[0313] 在第42方面,根据方面39至41中任一方面的方法,其中虚拟内容包括目标语言的文本表达或者目标语言的另一图形说明。

[0314] 在第43方面,根据方面39至42中任一方面的方法,其中识别手语中的手势的含义

包括在由成像系统捕获的图像的一部分上应用深度神经网络技术。

[0315] 在第44方面,根据方面39至43中任一方面的方法,其中检测手语中的通信的存在包括:从候选手语的列表中识别手语;并且确定检测到的手势对应于手语中的表达。

[0316] 在第45方面,根据方面44的方法,其中确定检测到的手势对应于手语中的表达包括分析与做出手势的人的嘴唇运动和在用户做出手势时捕获的音频数据相关地手势。

[0317] 与文本修改相关的其他方面

[0318] 下面进一步描述通过感测眼镜修改文本的特征的其他方面。

[0319] 在第1方面,一种头戴式显示设备,被配置为投影增强现实图像内容,该显示设备包括:框架,被配置为可佩戴在用户的头部上并被配置为支持用户眼睛前方的显示器;一个或多个相机,被配置为接收光信号;处理电子器件,被配置为:从一个或多个相机接收信号;根据信号识别图像;确定图像是否包括文本(例如,一个或多个字母或字符);将文本转换为修改的文本;并指示显示器渲染修改的文本。

[0320] 在第2方面,根据方面1的头戴式显示设备,其中显示器包括一个或多个光源和一个或多个波导堆叠,该一个或多个波导堆叠被配置为将光引导到用户的眼睛中以在眼睛中形成图像。

[0321] 在第3方面,根据方面2的头戴式显示设备,其中,一个或多个光源被配置为将光引导到波导堆叠中。

[0322] 在第4方面,根据方面2至3中任一方面的头戴式显示设备,其中,一个或多个光源包括光纤扫描投影仪。

[0323] 在第5方面,根据方面1至4中任一方面的头戴式显示设备,其中,一个或多个相机包括一个或多个摄像机。

[0324] 在第6方面,根据方面1至5中任一方面的头戴式显示设备,其中,处理电子器件被配置为使用光学字符识别算法将图像中的一个或多个字母或字符转换为文本。

[0325] 在第7方面,根据方面6的头戴式显示设备,其中处理电子器件被配置为访问数据库以识别一个或多个字母或字符的文本或语言的可能候选。

[0326] 在第8方面,根据方面6至7中任一方面的头戴式显示设备,其中,处理电子器件被配置为接收与用户参与的一个或多个活动的相关联的输入、用户的地理位置、用户的行进速度、用户的海拔高度、显示器检测到的环境噪声的音量或类型、显示器检测到的区域中的可见光或其他光的级别或类型、显示器检测到温度或气候、文本与用户的感知距离、或由显示器检测到的单词的类别。

[0327] 在第9方面,根据方面1至8中任一方面的头戴式显示设备还包括GPS系统。

[0328] 在第10方面,根据方面1至9中任一方面的头戴式显示设备,其中,修改的文本具有与文本的第一字体大小不同的第二字体大小。第二字体大小可以大于第一字体大小。

[0329] 在第11方面,根据方面1至10中任一方面的头戴式显示设备,其中,修改的文本比文本对于用户更易读。

[0330] 在第12方面,根据方面1至11中任一方面的头戴式显示设备,其中,处理电子器件被配置为将图形元素添加到文本,部分地形成修改的文本。

[0331] 在第13方面,根据方面1至12中任一方面的头戴式显示设备,其中,处理电子器件被配置为以与一个或多个字母或字符的第一字体不同的第二字体来显示文本的一个或多

个字母或字符。

[0332] 在第14方面,根据方面1至13中任一方面的头戴式显示设备,其中,处理电子器件被配置为相对于用户在没有头戴式显示器的情况下看到的内容来放大文本的一个或多个字母或字符。

[0333] 在第15方面,根据方面1至14中任一方面的头戴式显示设备,其中,处理电子器件被配置为显示边界区域,该边界区域界定内部区域。

[0334] 在第16方面,根据方面15的头戴式显示设备,其中处理电子器件被配置为在内部区域内显示一个或多个字母或字符。

[0335] 在第17方面,根据方面1至16中任一方面的头戴式显示设备,其中,处理电子器件被配置为针对第二背景显示文本的一个或多个字母或字符,该第二背景不同于用户在没有头戴式显示器的情况下读取一个或多个字母或字符所针对的第一背景。

[0336] 在第18方面,根据方面17的头戴式显示设备,其中第二背景包括单色背景。

[0337] 在第19方面,根据方面18的头戴式显示设备,其中单色背景包括白色。

[0338] 在第20方面,根据方面17至19中任一方面的头戴式显示设备,其中,第一背景包括用户在没有头戴式显示器的情况下将看到的内容。

[0339] 在第21方面,根据方面1至20中任一方面的头戴式显示设备,其中,文本适于由文本编辑器编辑。

[0340] 尽管参考头戴式显示器描述了方面1至21,但是在这些方面中描述的类似功能也可以利用头戴式设备或参考图2A描述的可穿戴系统来实现。此外,显示器可以包括多个深度平面,并且头戴式设备被配置为至少部分地基于用户的眼睛能力来辨别深度平面以渲染修改的文本。

[0341] 在第22方面,一种用于使用头戴式显示器投影增强现实图像内容的方法,该方法包括:在硬件处理器的控制下:从一个或多个相机接收光信号;使用光学字符识别模块,从信号中识别图像;确定图像是否包括一个或多个字母或字符;将一个或多个字母或字符转换为文本;以及在头戴式显示器上显示文本,其中显示文本包括通过光学透射目镜将光作为图像发送给用户。

[0342] 在第23方面,根据方面22的方法,还包括将光引导到用户的眼睛中以在眼睛中形成图像。

[0343] 在第24方面,根据方面22至23中任一方面的方法,还包括使用光纤扫描投影仪将光引导到目镜中。

[0344] 在第25方面,根据方面22至24中任一方面的方法,其中使用光学字符识别模块包括识别一个或多个字母或字符的文本或语言的可能候选。

[0345] 在第26方面,根据方面22至25中任一方面的方法,其中使用光学字符识别模块包括接收:包括与用户参与的一个或多个活动相关联的信息的输入、用户的地理位置、用户的行进速度、用户的海拔高度,显示器检测到的环境噪声的音量或类型、显示器检测到的区域中的可见光或其他光的级别或类型、显示器检测到温度或气候、一个或多个字母或字符与用户的感知距离、或显示器检测的单词的类别。

[0346] 在第27方面,根据方面22至26中任一方面的方法,其中将一个或多个字母或字符转换为文本包括以不同于与一个或多个字母或字符相关联的第一语言的第二语言显示文

本。在第27方面的一些实施方式中,该方法包括将文本翻译成第二语言。

[0347] 在第28方面,根据方面22至27中任一方面的方法,还包括从GPS系统接收位置数据。

[0348] 在第29方面,根据方面22至28中任一方面的方法,其中在头戴式显示器上显示一个或多个字母或字符包括以与一个或多个字母或字符的第一个字体大小不同的第二字体大小来显示一个或多个字母或字符。

[0349] 在第30方面,根据方面22至29中任一方面的方法,其中在头戴式显示器上显示一个或多个字母或字符包括比用户在没有头戴式显示器的情况下更清楚地向用户显示一个或多个字母或字符。

[0350] 在第31方面,根据方面22至30中任一方面的方法,其中在头戴式显示器上显示一个或多个字母或字符包括以用户在没有头戴式显示器的情况下将看起来更大的字体大小来显示一个或多个字母或字符。

[0351] 在第32方面,根据方面22至31中任一方面的方法,其中在头戴式显示器上显示一个或多个字母或字符包括以与一个或多个字母或字符的第一个字体不同的第二字体来显示一个或多个字母或字符。

[0352] 在第33方面,根据方面22至32中任一方面的方法,其中在头戴式显示器上显示一个或多个字母或字符包括相对于用户在没有头戴式显示器的情况下将看到的内容来放大一个或多个字母或字符。

[0353] 在第34方面,根据方面22至33中任一方面的方法,其中在头戴式显示器上显示一个或多个字母或字符包括显示边界区域,该边界区域界定内部区域。

[0354] 在第35方面,根据方面34的方法,其中在头戴式显示器上显示一个或多个字母或字符包括在内部区域内显示一个或多个字母或字符。

[0355] 在第36方面,根据方面22至35中任一方面的方法,其中在头戴式显示器上显示一个或多个字母或字符包括针对第二背景显示一个或多个字母或字符,该第二背景不同于用户在没有头戴式显示器的情况下将读取一个或多个字母或字符所针对的第一背景。

[0356] 在第37方面,根据方面36的方法,其中第二背景包括单色背景。

[0357] 在第38方面,根据方面37的方法,其中单色背景包括白色。

[0358] 在第39方面,根据方面36至38中任一方面的方法,其中第一背景包括用户在没有头戴式显示器的情况下将看到的内容。

[0359] 在第40方面,根据方面22至39中任一方面的方法,其中文本适于由文本编辑器编辑。

[0360] 在第41方面,根据方面22至40中任一方面的方法,其中将一个或多个字母或字符转换为文本包括从用户接收输入。

[0361] 在第42方面,根据方面41的方法,其中从用户接收输入包括接收口头命令、手势、头部的运动或者用户眼睛中的一个或多个的运动中的一个或多个。

[0362] 在第43方面,根据方面22至42中任一方面的方法,其中如果一个或多个字母或字符看起来比第一距离阈值更近,则文本在看起来比第二深度更近的第一深度处显示。

[0363] 在第44方面,根据方面22至43中任一方面的方法,其中如果一个或多个字母或字符看起来比第二距离阈值更远,则文本在看起来比第一深度更远第二深度处显示。

[0364] 在第45方面,根据方面43至44中任一方面的方法,其中,如果一个或多个字母或字符看起来比第一距离阈值更远并且比第二距离阈值更近,则文本在看起来比第一深度更远但是比第二深度更近的第三深度处显示。

[0365] 在第46方面,根据方面43至45中任一方面的方法,其中第一距离阈值是80cm。

[0366] 在第47方面,根据方面43至46中任一方面的方法,其中第二距离阈值是600cm。

[0367] 在第48方面,根据方面43至47中任一方面的方法,其中第二距离阈值和第一距离阈值之间的差小于100m。

[0368] 与标示修改相关的其他方面

[0369] 在第1方面,一种增强现实系统,包括:面向外成像系统;非瞬态存储器,被配置为存储由面向外成像系统获得的图像;以及硬件处理器,被编程为:接收由面向外成像系统获得的增强现实系统的用户的环境的图像;分析图像以识别用户环境中的标示;识别标示上的文本;将至少一部分文本转换为目标语言;并指示显示器向用户渲染转换的文本。

[0370] 在第2方面,根据方面1的增强现实系统,其中,硬件处理器被编程为修改与文本相关联的显示特征。

[0371] 在第3方面,根据方面1或2的增强现实系统,其中,为了将至少一部分文本转换为目标语言,硬件处理器被编程为识别标示上的文本的语言并且将语言转换为目标语言。

[0372] 在第4方面,根据方面1至3中任一方面的增强现实系统,其中,硬件处理器被编程为至少部分地基于用户的位置来确定目标语言。

[0373] 在第5方面,根据方面1至4中任一方面的增强现实系统,其中,为了识别标示上的文本,硬件处理器被编程为识别作为目标语言的文本。

[0374] 在第6方面,根据方面5的增强现实系统,其中,硬件处理器被编程为不转换作为目标语言的文本。

[0375] 其他考虑

[0376] 在此描述的或在附图中描绘的过程、方法和算法中的每一个可以体现在由一个或多个物理计算系统、硬件计算机处理器、专用电路或被配置为执行具体和特定计算机指令的电子硬件执行的代码模块中,并且由以上完全或部分自动化。例如,计算系统可以包括利用特定计算机指令编程的通用计算机(例如,服务器)或专用计算机、专用电路等等。代码模块可以被编译并链接到可执行程序中,安装在动态链接库中,或者可以用解释的编程语言编写。在一些实施方式中,特定操作和方法可以由给定功能特定的电路来执行。

[0377] 此外,本公开的功能的某些实施方式在数学上、计算上或技术上是足够复杂的,以致于可能需要专用硬件或一个或多个物理计算设备(利用适当的专用可执行指令)来执行功能,例如由于所涉及的计算的数量或复杂性或为了基本实时地提供结果。例如,动画或视频可以包括许多帧,每帧具有数百万个像素,并且专门编程的计算机硬件需要处理视频数据,从而在商业上合理的时间量内提供期望的图像处理任务或应用。

[0378] 代码模块或任何类型的数据可以存储在任何类型的非瞬态计算机可读介质上,诸如物理计算机存储器,包括硬盘驱动器、固态存储器、随机存取存储器(RAM)、只读存储器(ROM)、光盘、易失性或非易失性存储器、相同或类似的组合。方法和模块(或数据)也可以在各种计算机可读传输介质上作为生成的数据信号(例如,作为载波或其他模拟或数字传播信号的一部分)传输,所述传输介质包括基于无线的和基于有线/电缆的介质,并且可以采

取多种形式(例如,作为单个或多路复用模拟信号的一部分,或者作为多个离散数字数据包或帧)。所公开的过程或过程步骤的结果可以持久地或以其他方式存储在任何类型的非瞬态有形计算机存储器中,或者可以经由计算机可读传输介质来传送。

[0379] 在此所描述或附图中描绘的流程图中的任何过程、框、状态、步骤或功能应当被理解为潜在地表示代码模块、代码段或代码部分,这些代码模块、代码段或代码部分包括用于实现特定功能(例如,逻辑或算术)或过程中的步骤的一个或多个可执行指令。各种过程、框、状态、步骤或功能可以与在此提供的说明性示例相组合、重新排列、添加、删除、修改或以其他方式改变。在一些实施例中,附加的或不同的计算系统或代码模块可以执行在此描述的功能中的一些或全部。在此描述的方法和过程也不限于任何特定的顺序,并且与其相关的框、步骤或状态可以以适当的其他顺序来执行,例如串行、并行或以某种其他方式。任务或事件可以添加到所公开的示例实施例或者从中删除。此外,在此描述的实施方式中的各种系统组件的分离是出于说明的目的,并且不应该被理解为在所有实施方式中都需要这种分离。应该理解,所描述的程序组件、方法和系统通常可以一起集成在单个计算机产品中或者封装到多个计算机产品中。许多实施方式变化是可能的。

[0380] 过程、方法和系统可以实现在网络(或分布式)计算环境中。网络环境包括企业范围的计算机网络、内联网、局域网(LAN)、广域网(WAN)、个人局域网(PAN)、云计算网络、众包(crowd-sourced)计算网络、互联网和万维网。网络可以是有线或无线网络或任何其他类型的通信网络。

[0381] 本公开的系统和方法各自具有若干创新性方面,其中没有单独一个对在此公开的期望属性负责或需要。上述各种特征和过程可以彼此独立地使用,或者可以以各种方式组合。所有可能的组合和子组合均旨在落入本公开的范围。对于本公开中所描述的实施方式的各种修改对于本领域技术人员来说可以是容易清楚的,并且在不脱离本公开的精神或范围的情况下,可将在在此定义的一般原理应用于其他实施方式。因此,权利要求不旨在限于在此示出的实施方式或实施例,而是应被给予与本公开、在此公开的原理和新颖特征一致的最宽范围。

[0382] 本说明书中在分开的实施方式或实施例的上下文中描述的某些特征也可以在单个实施方式或实施例中组合地实施。相反地,在单个实施方式或实施例的上下文中描述的各种特征也可以在多个实施方式或实施例中分开地或以任何合适的子组合方式来实施。此外,尽管上文可以将特征描述为以某些组合起作用并且甚至最初如此要求,但是来自所要求的组合的一个或多个特征可以在一些情况下从组合中删除,并且所要求的组合可以针对子组合或子组合的变体。没有单个特征或特征组是每个实施例必需或不可缺少。

[0383] 除非另有特别说明,或者在所使用的上下文中进行理解,在此使用的条件语言,诸如“能(can)”、“能够(could)”、“可能(might)”、“可以(may)”、“例如(e.g.)”等等,一般意在表达某些实施例包括而其他实施例不包括某些特征、要素或步骤。因此,这样的条件语言通常不旨在暗示特征、要素或步骤以任何方式对于一个或多个实施例是必需的,或者一个或多个实施例必然包括用于在有或者没有作者输入或提示的情况下决定这些特征、要素或步骤是否包括在或执行在任何特定实施例中。术语“包括(comprising)”、“包括(including)”、“具有(having)”等是同义词,并且以开放式的方式包含性地使用,并且不排除附加的要素、特征、动作、操作等等。此外,术语“或”以其包含性含义(而不是其排他含义)

使用,因此当用于例如连接要素的列表时,术语“或”表示列表中的一个、一些或全部要素。另外,除非另有说明,否则本申请和所附权利要求书中使用的冠词“一”、“一个”和“所述”应被解释为表示“一个或多个”或“至少一个”。

[0384] 如在此所使用的,提及项目列表中的“至少一个”的短语是指这些项目的任何组合,包括单个成员。作为例子,“A、B或C中的至少一个”旨在涵盖:A、B、C、A和B、A和C、B和C,以及A、B和C。除非另有特别说明,否则诸如短语“X、Y和Z中的至少一个”的连接语言如所使用的在利用上下文进行理解,通常用于表达项目、术语等可以是X、Y或Z中的至少一个。因此,这样的连接语言通常不旨在暗示某些实施例需要X中的至少一个、Y中的至少一个和Z中的至少一个各自存在。

[0385] 类似地,虽然可以在附图中以特定顺序示出操作,但是应该认为,不需要以所示出的特定顺序或按顺次顺序执行这样的操作,或者不需要执行所有示出的操作以实现期望的结果。此外,附图可以以流程图的形式示意性地描绘一个或多个示例过程。然而,未示出的其他操作可以并入示意性说明的示例性方法和过程中。例如,一个或多个附加操作可以在任何所示操作之前、之后、同时或之间执行。另外,在其他实施方式中,操作可以重新安排或重新排序。在特定情况下,多任务和并行处理可能是有利的。而且,上述实施方式中的各种系统组件的分离不应当被理解为在所有实施方式中都需要这种分离,并且应该理解的是,所描述的程序组件和系统通常可以一起集成在单个软件产品中或者封装到多个软件产品。另外,其他的实施方式在以下权利要求的范围中。在一些情况下,权利要求中列举的动作能够以不同的顺序执行并且仍然实现期望的结果。

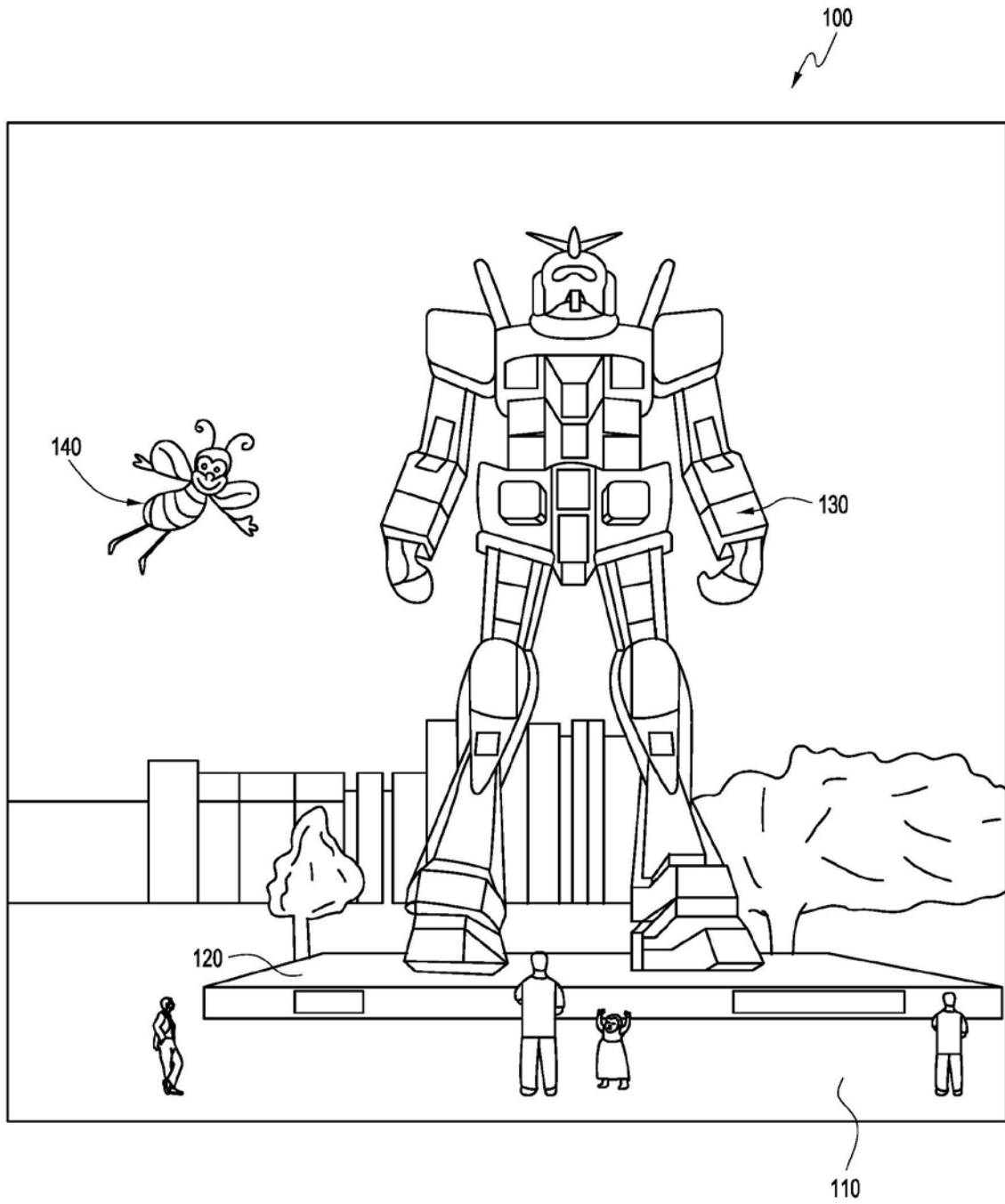


图1

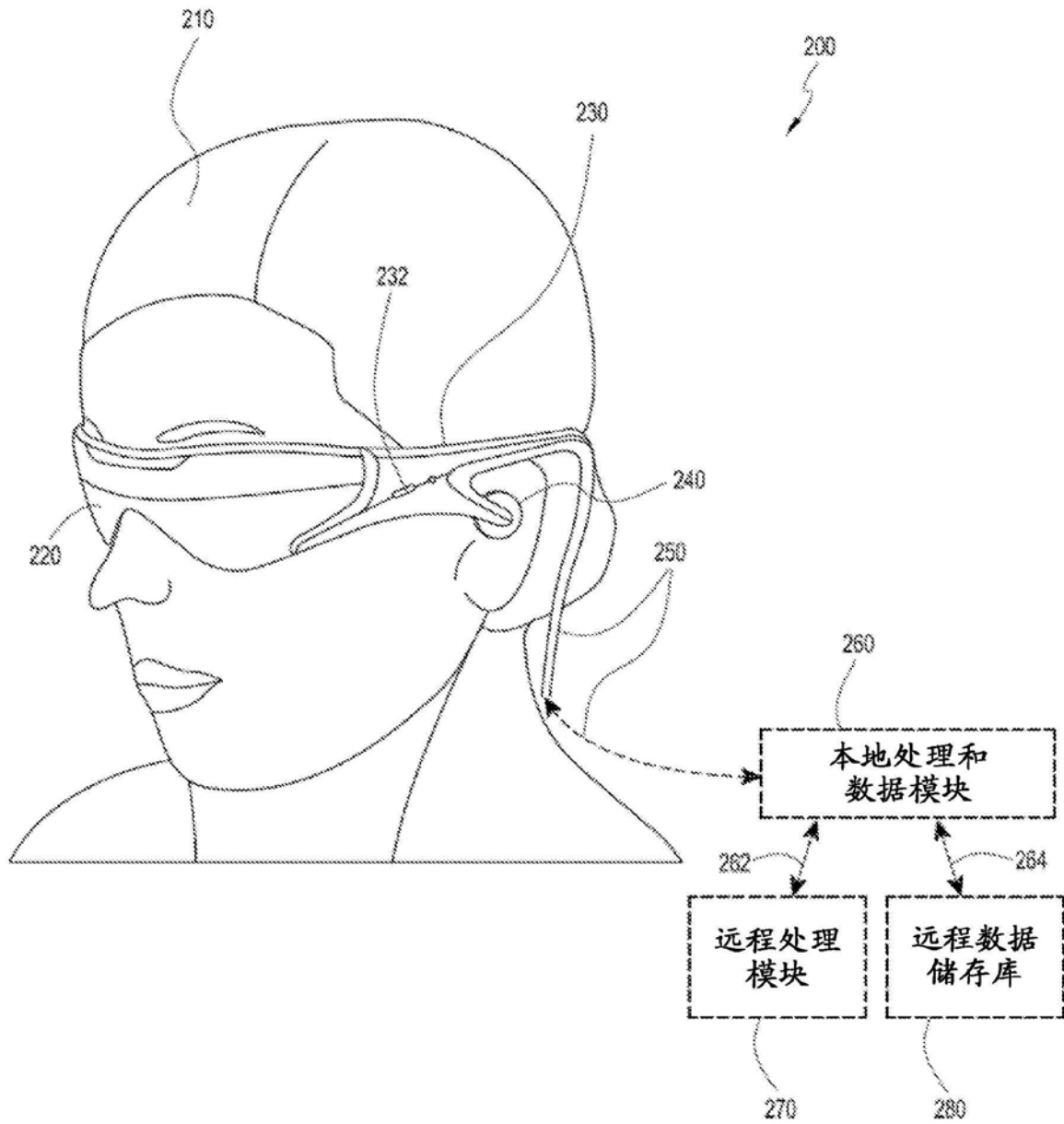


图2A

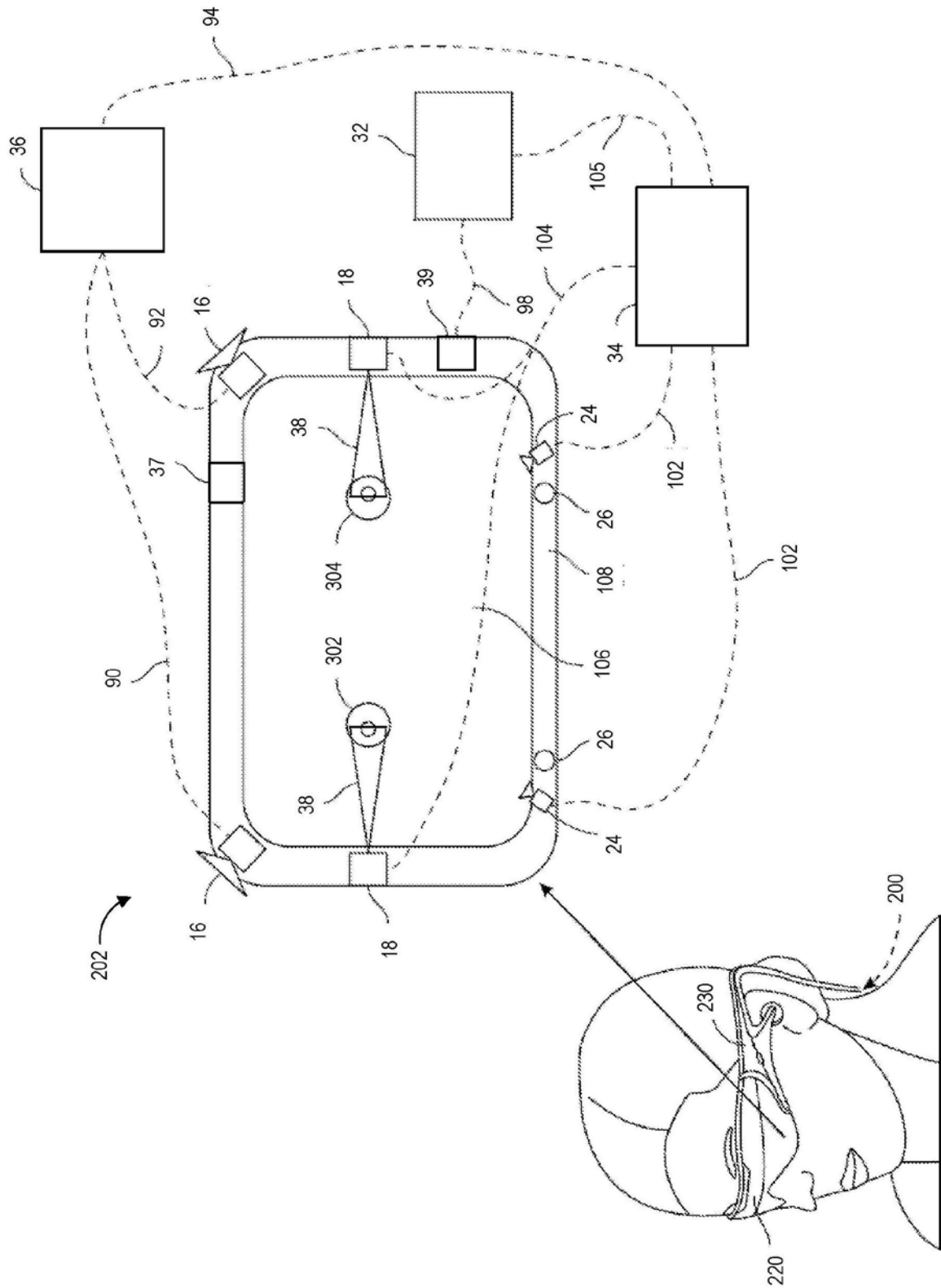


图2B

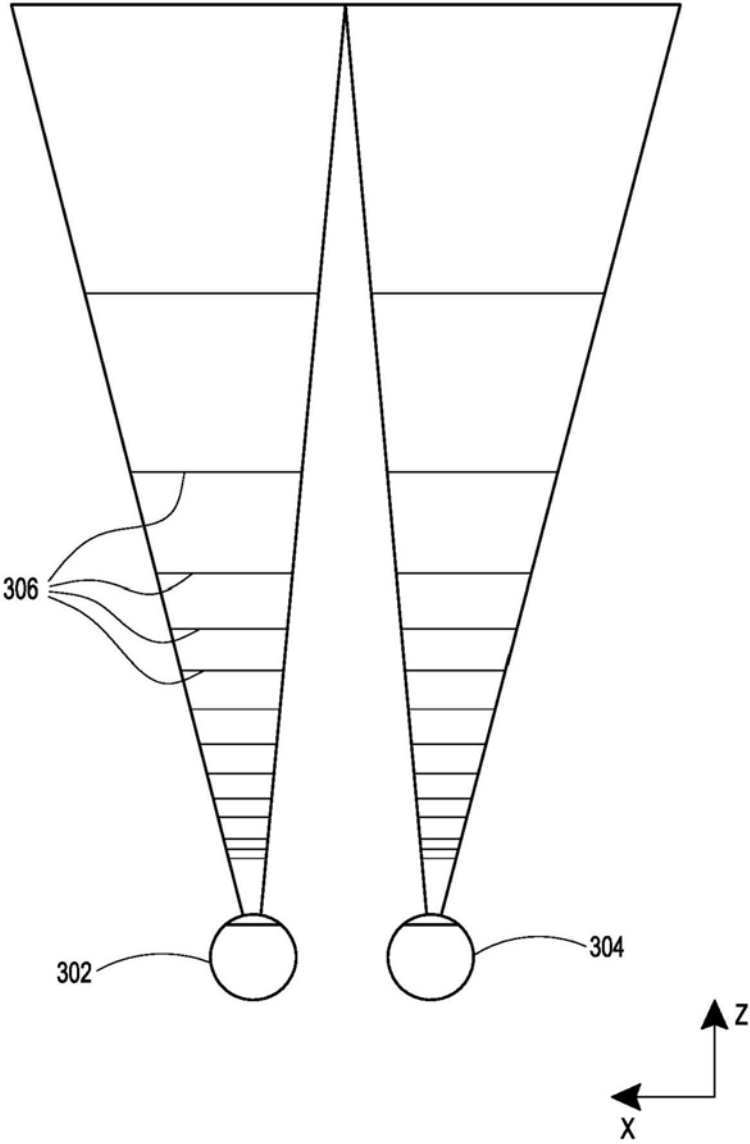


图3

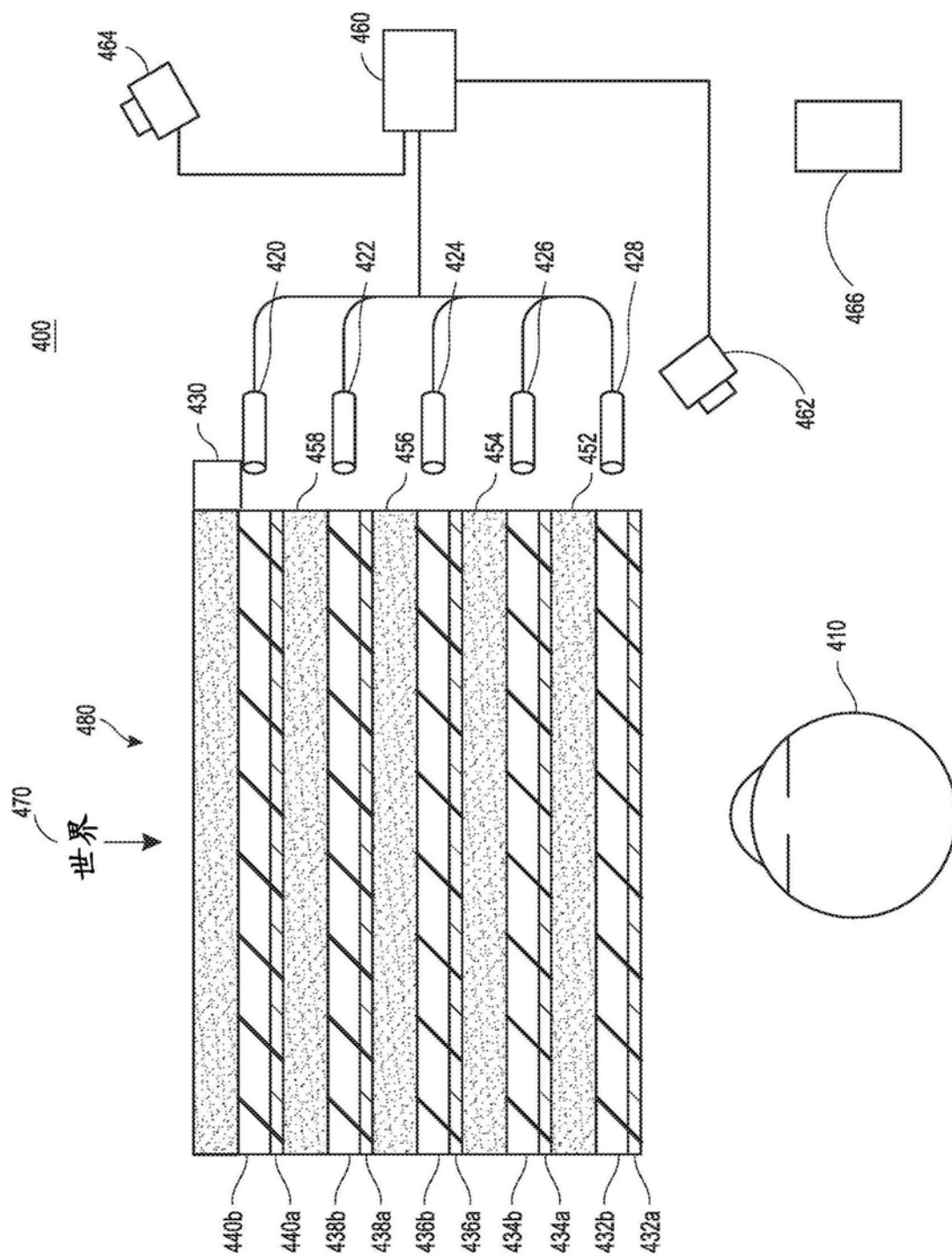


图4

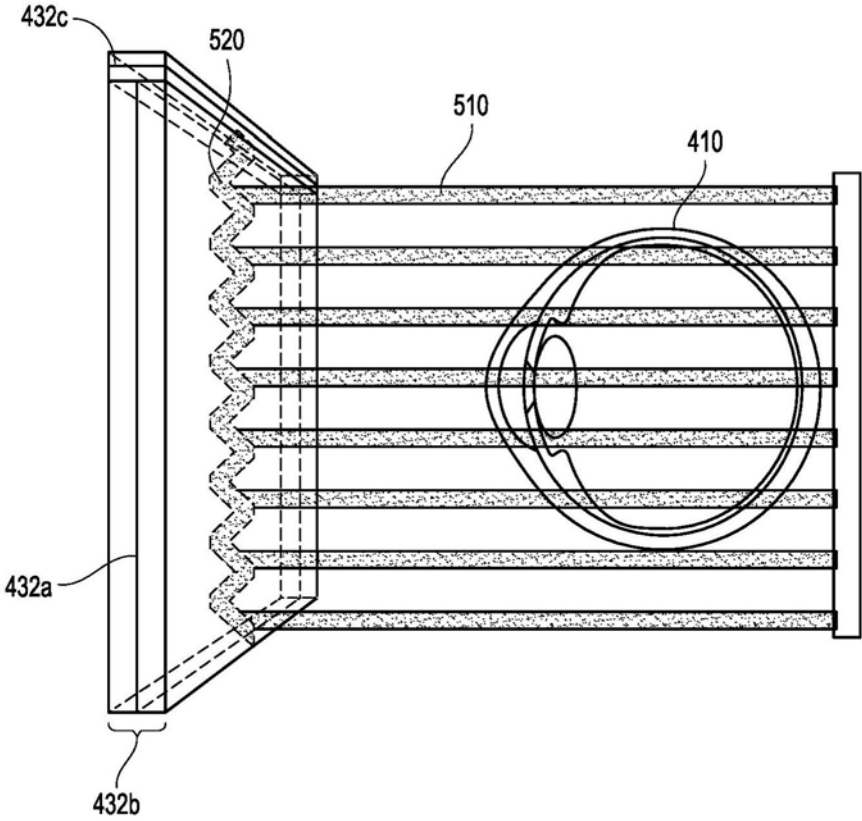


图5

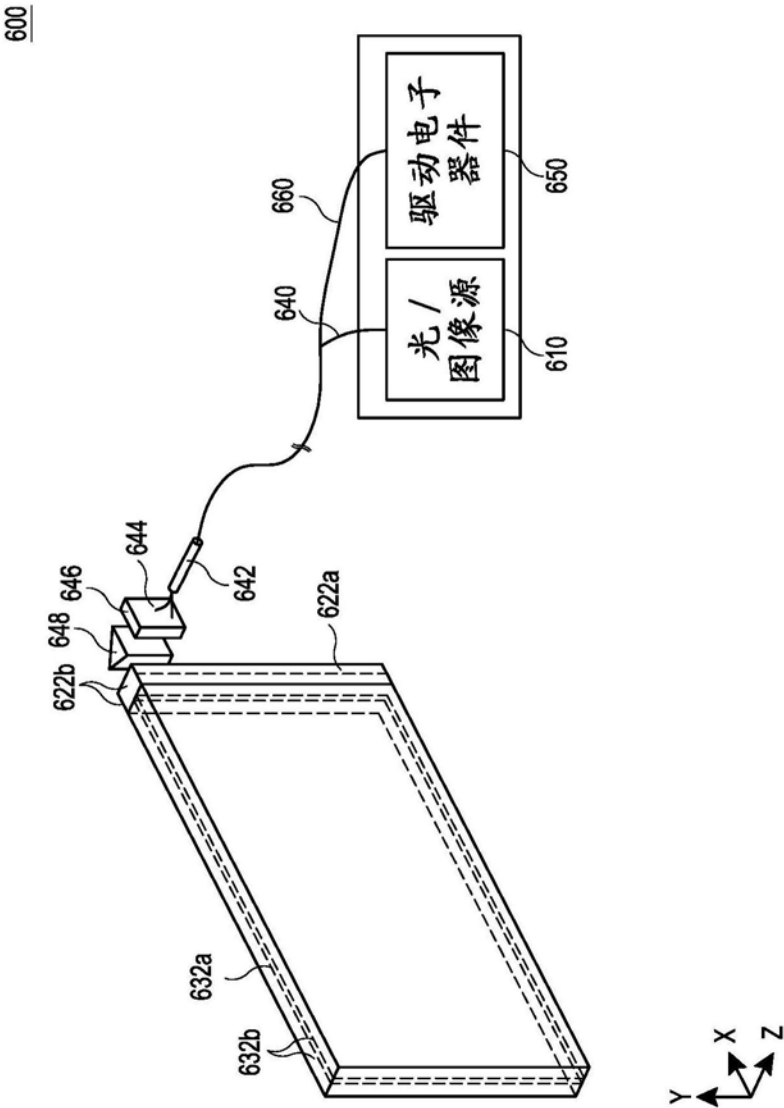


图6

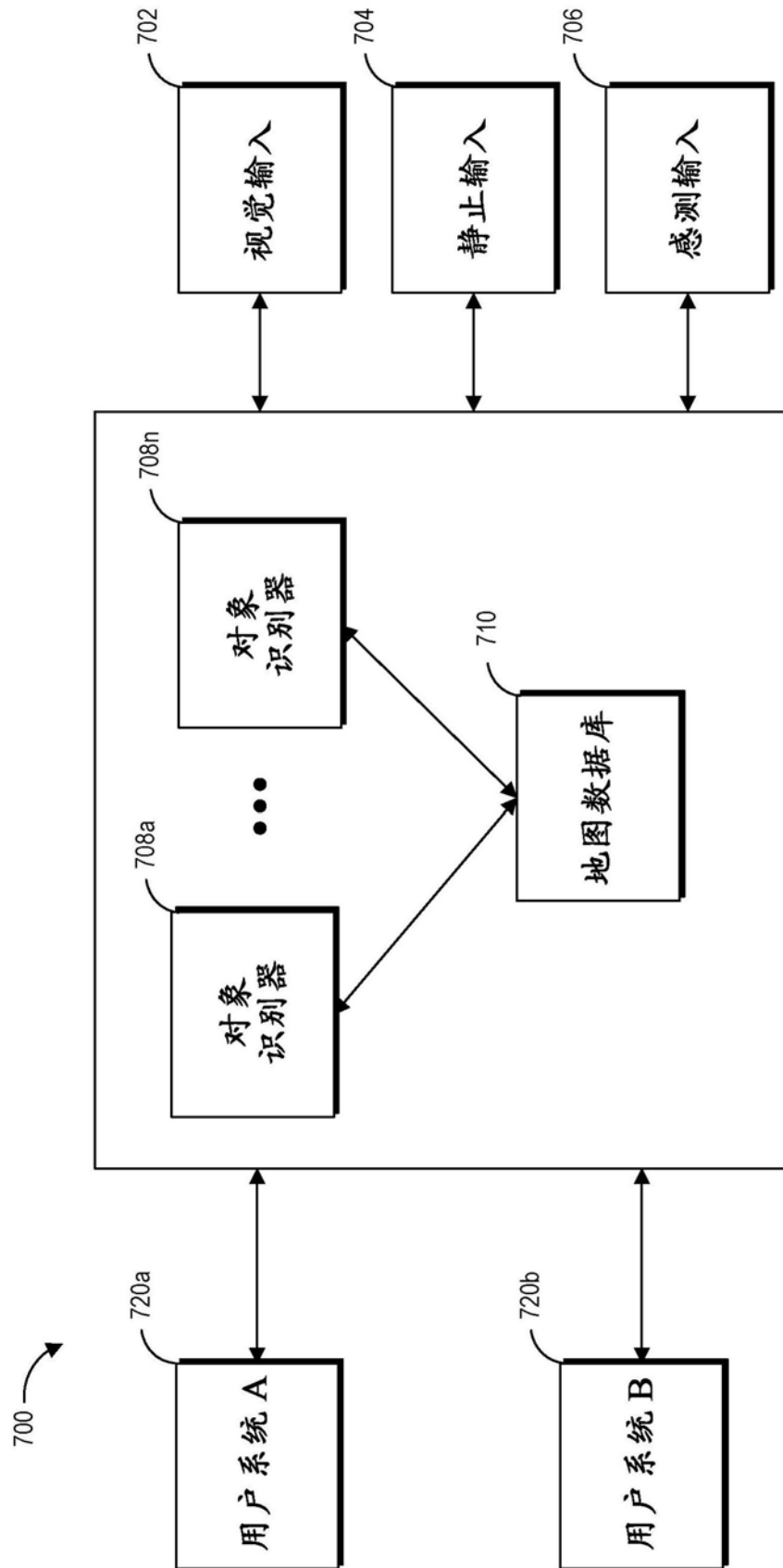


图7

800

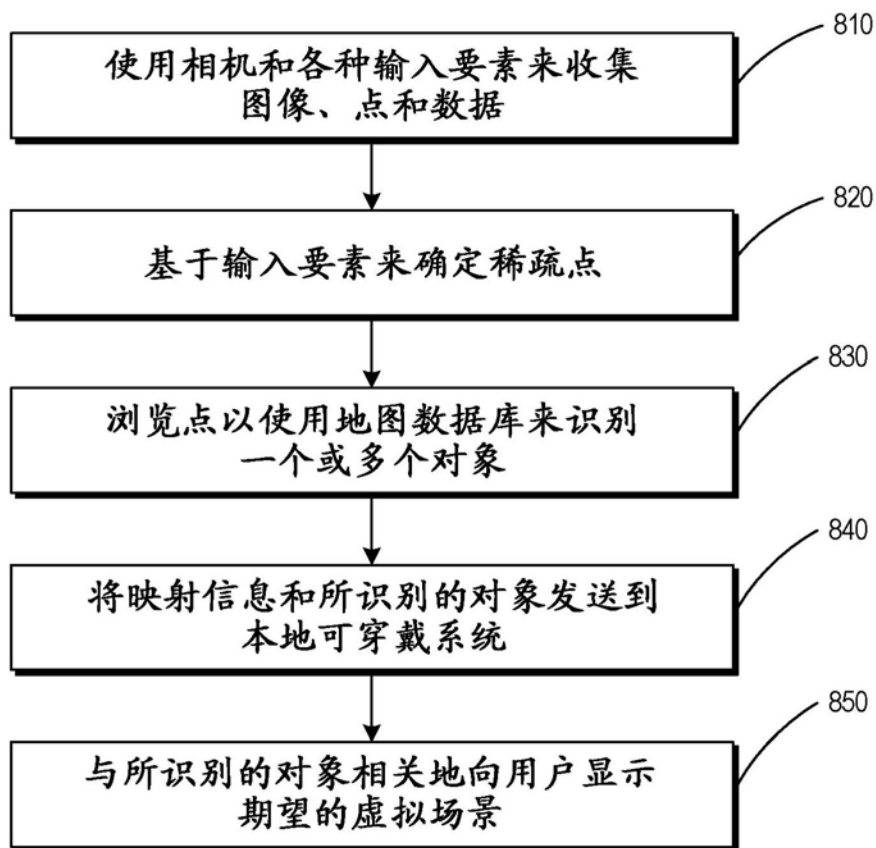


图8

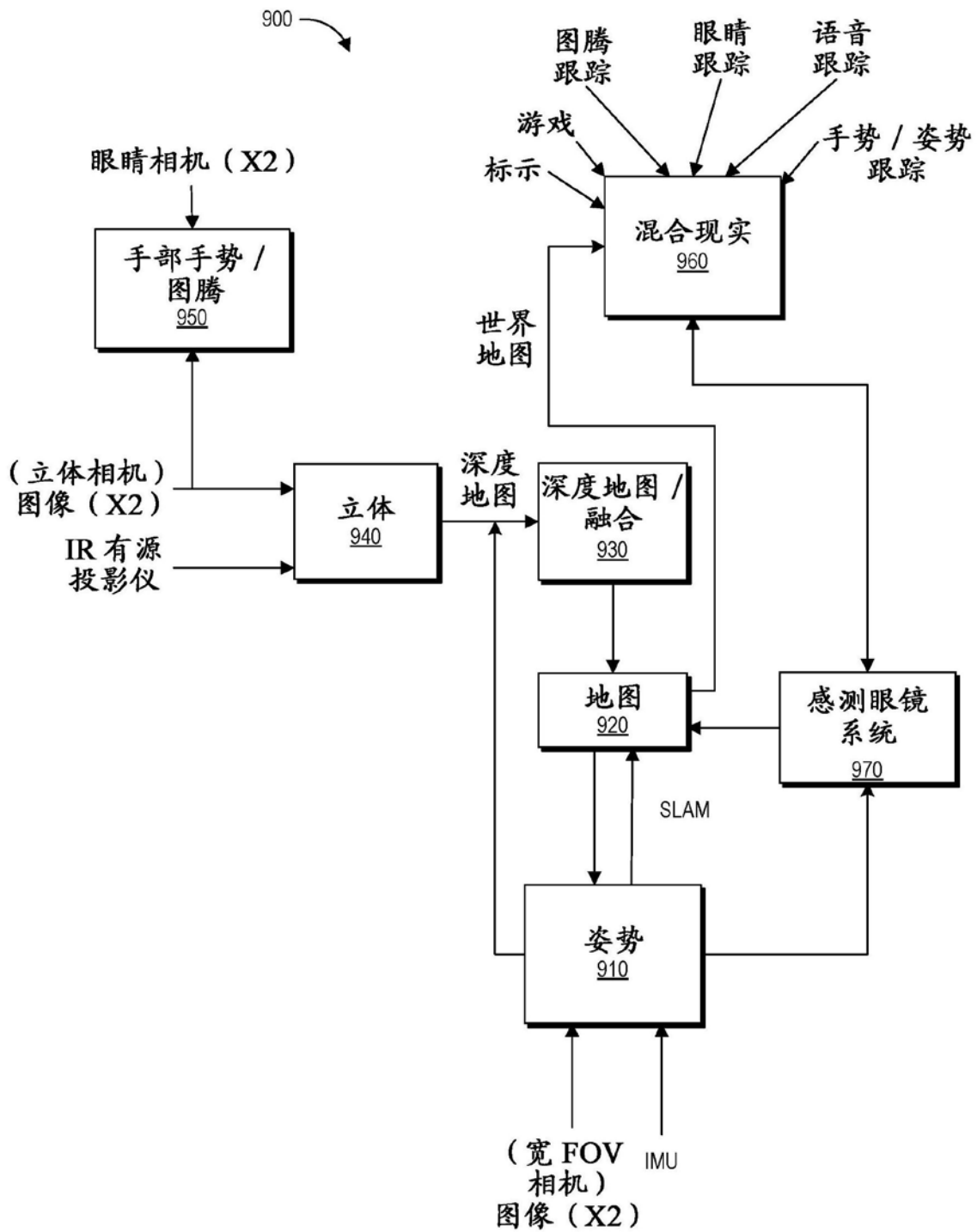


图9

1000

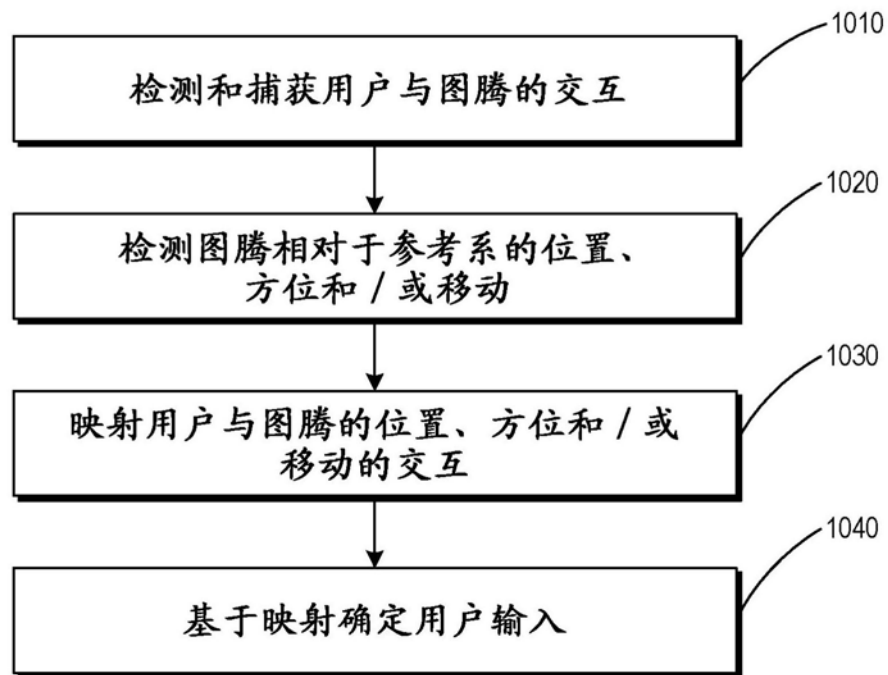


图10

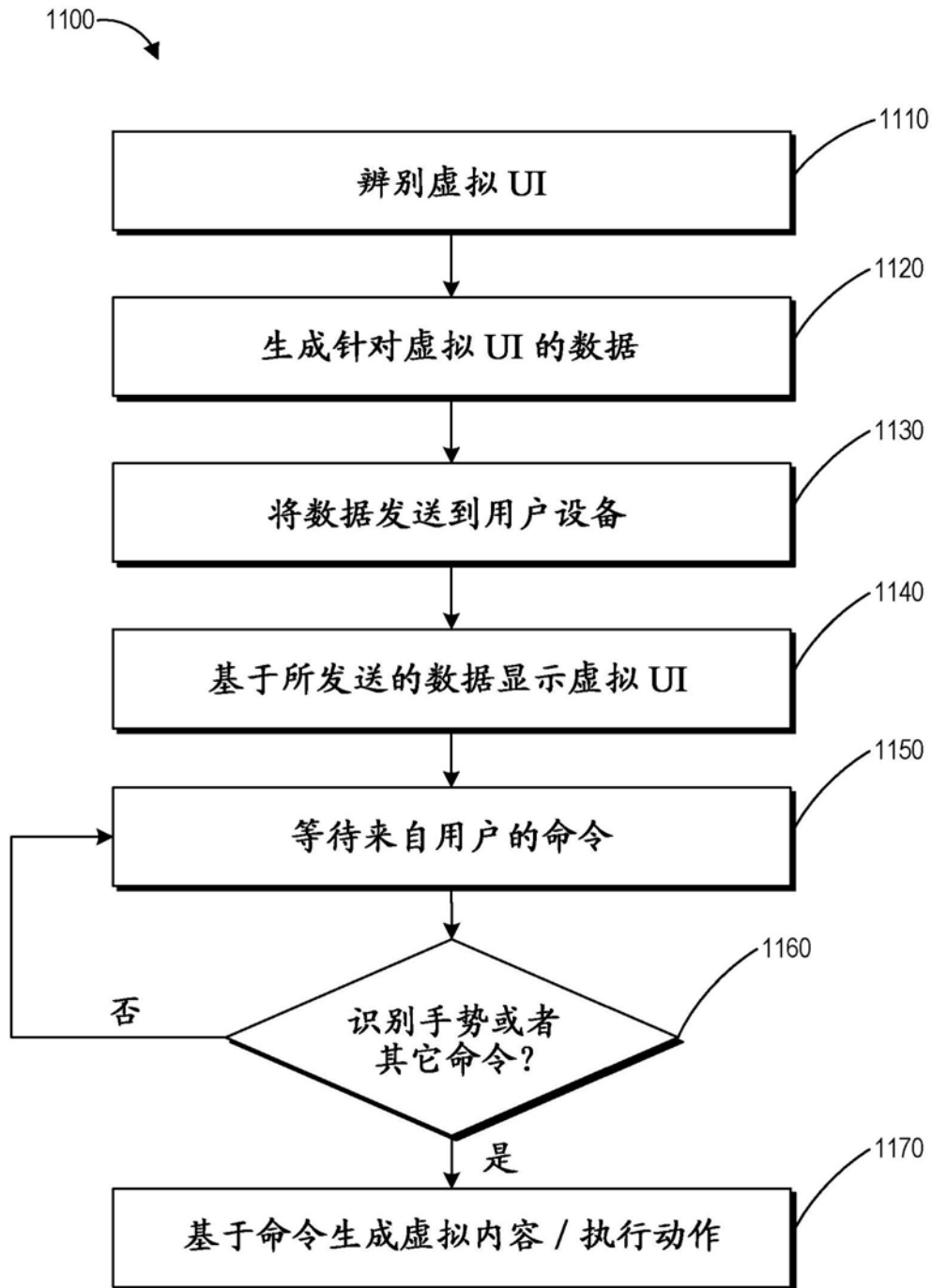


图11

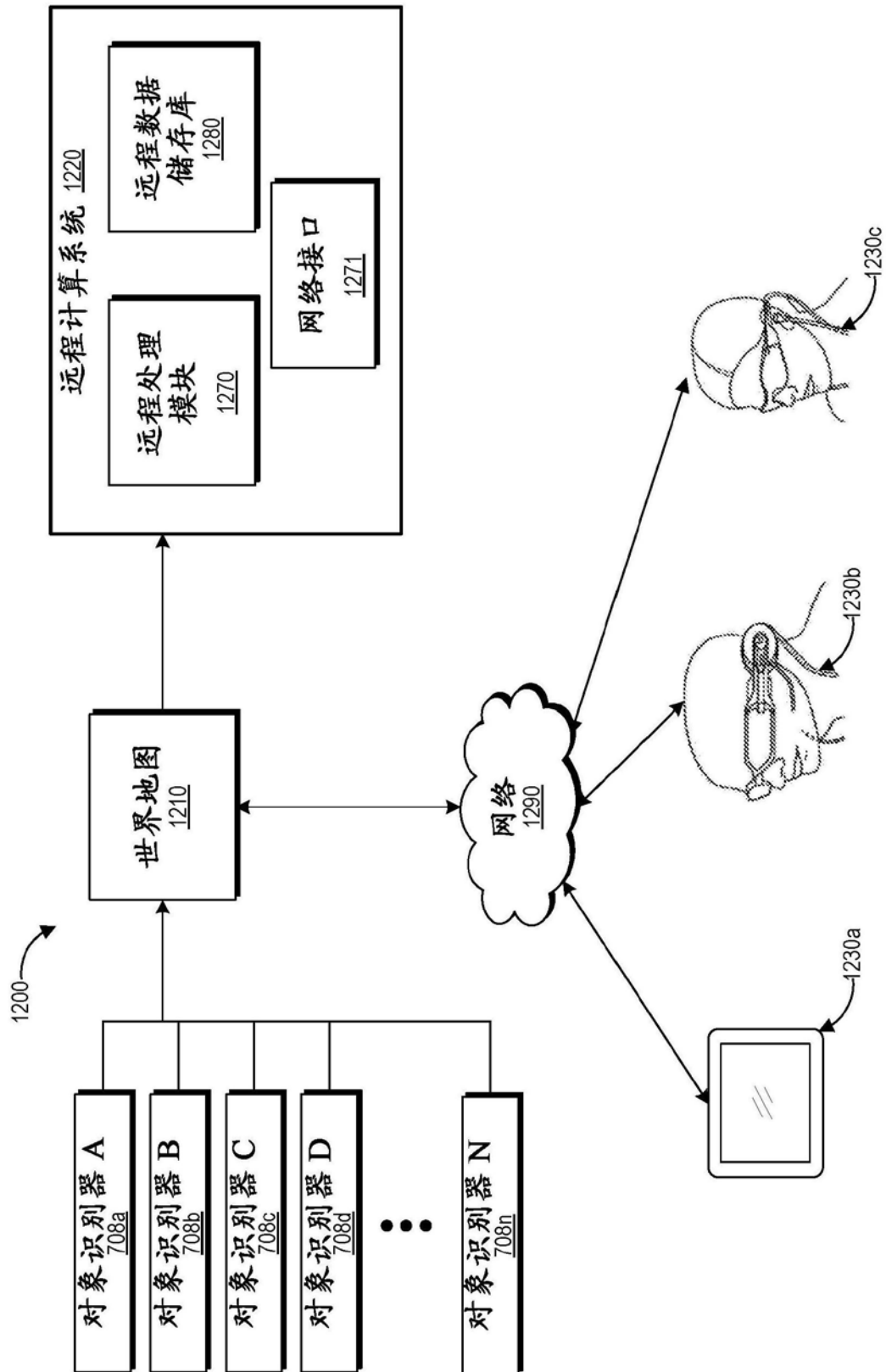


图12

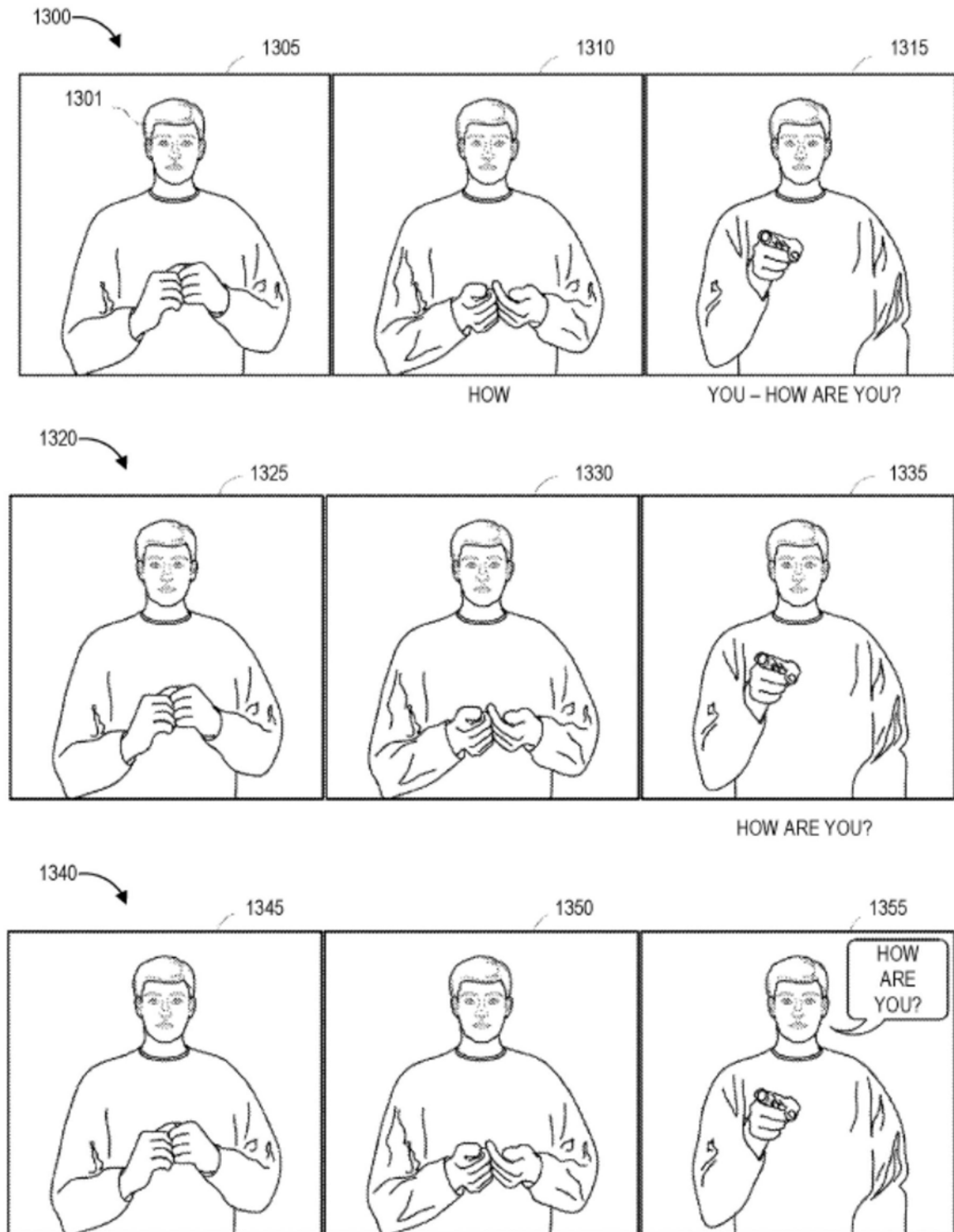


图13A

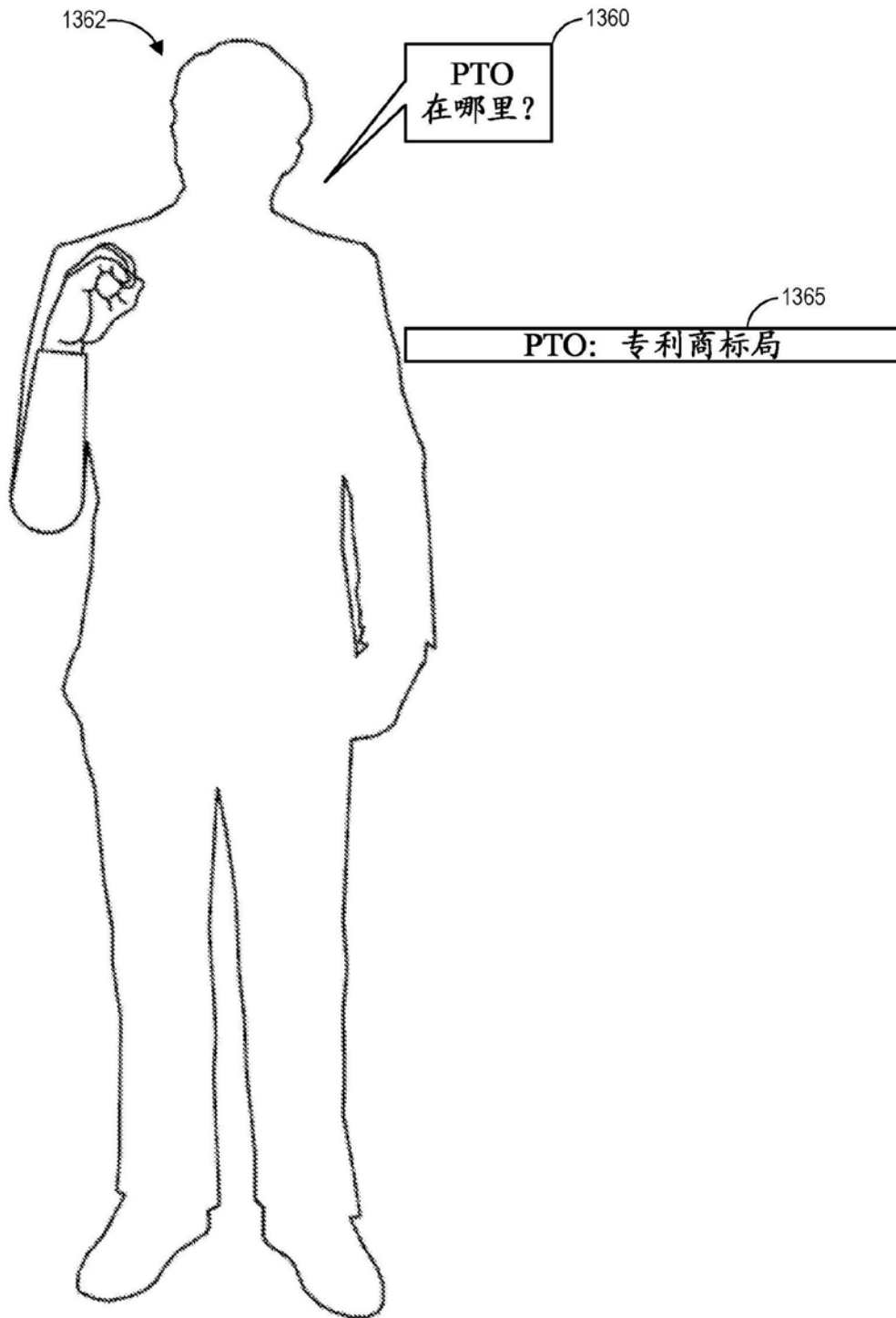


图13B

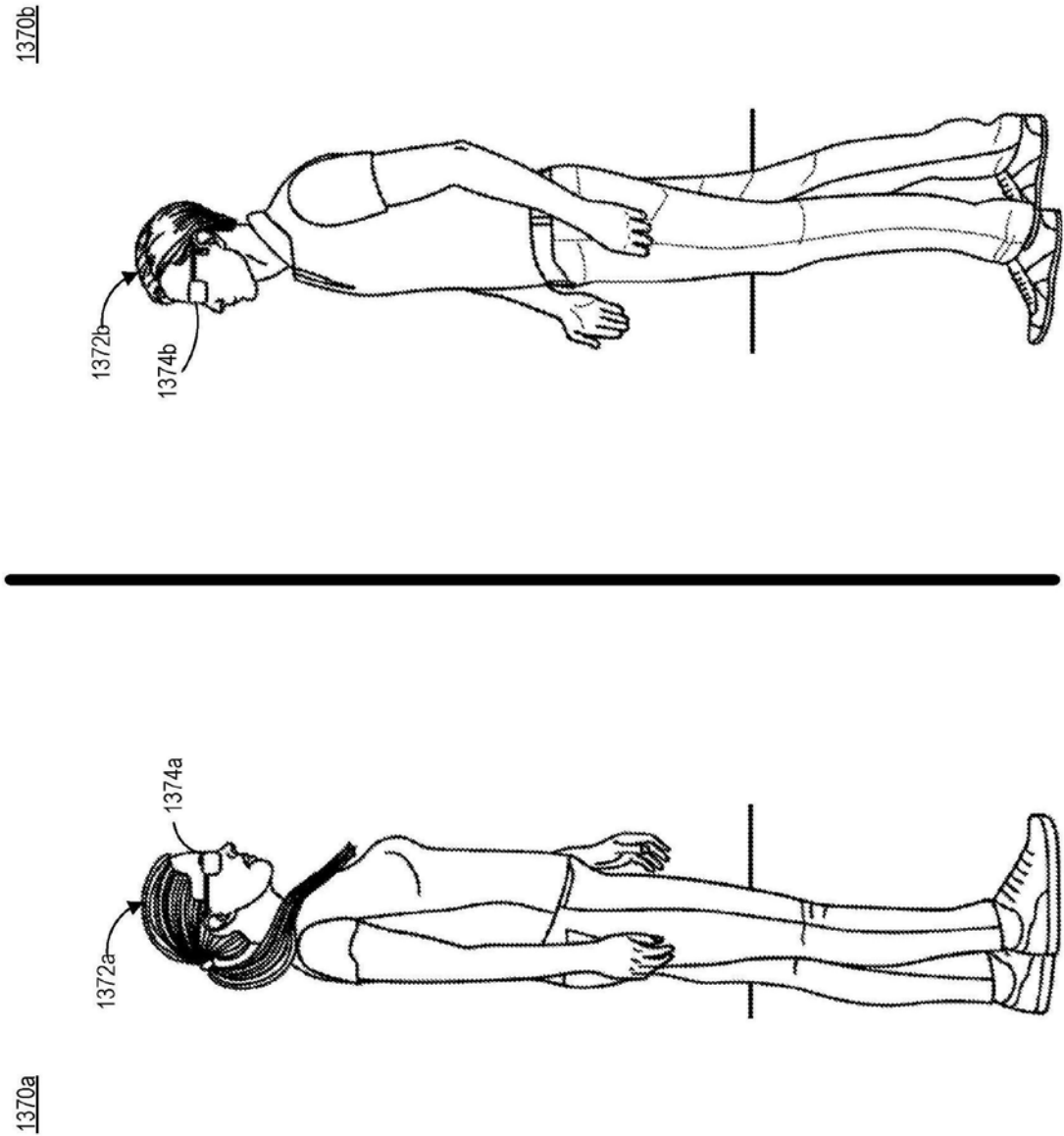


图13C

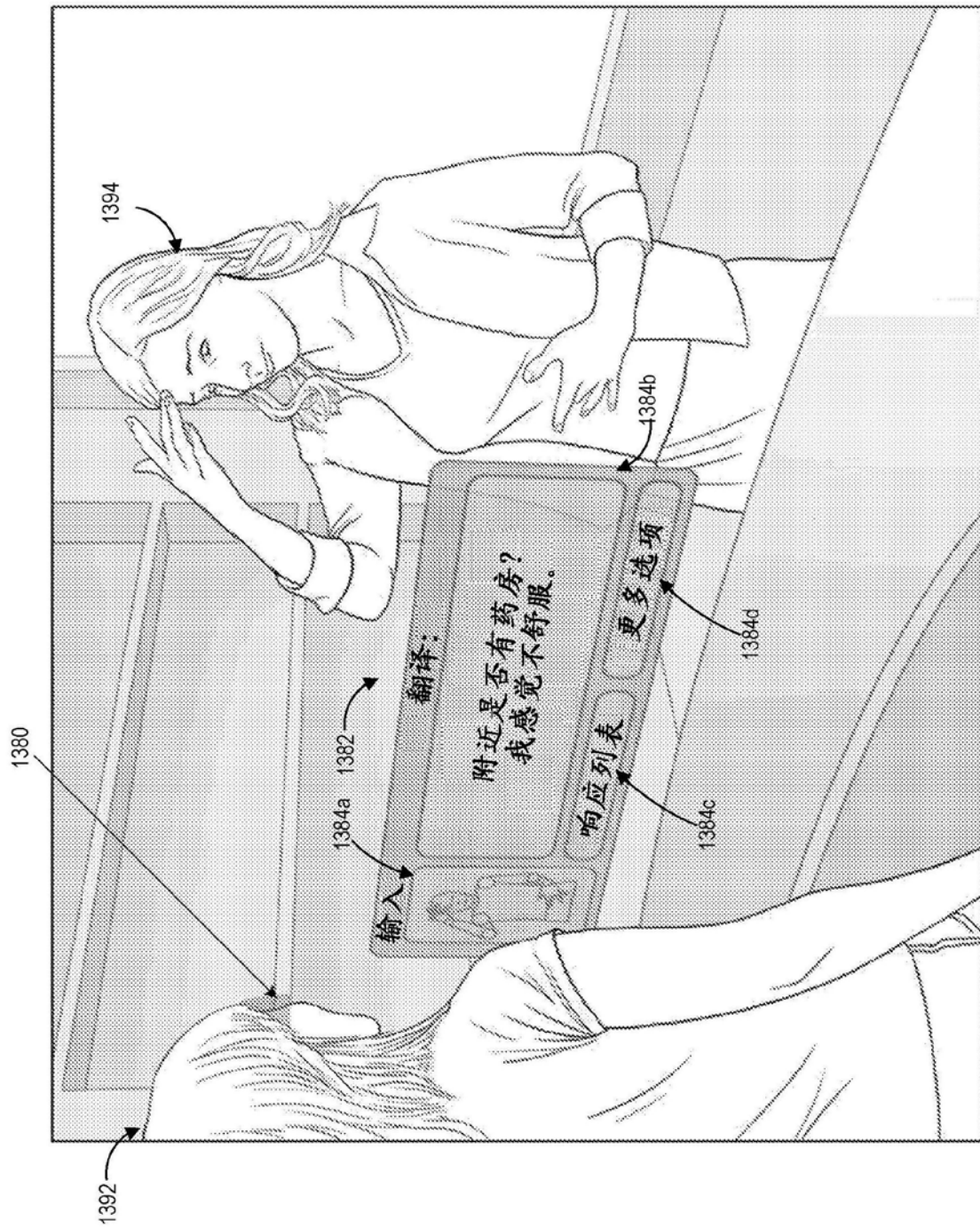


图13D

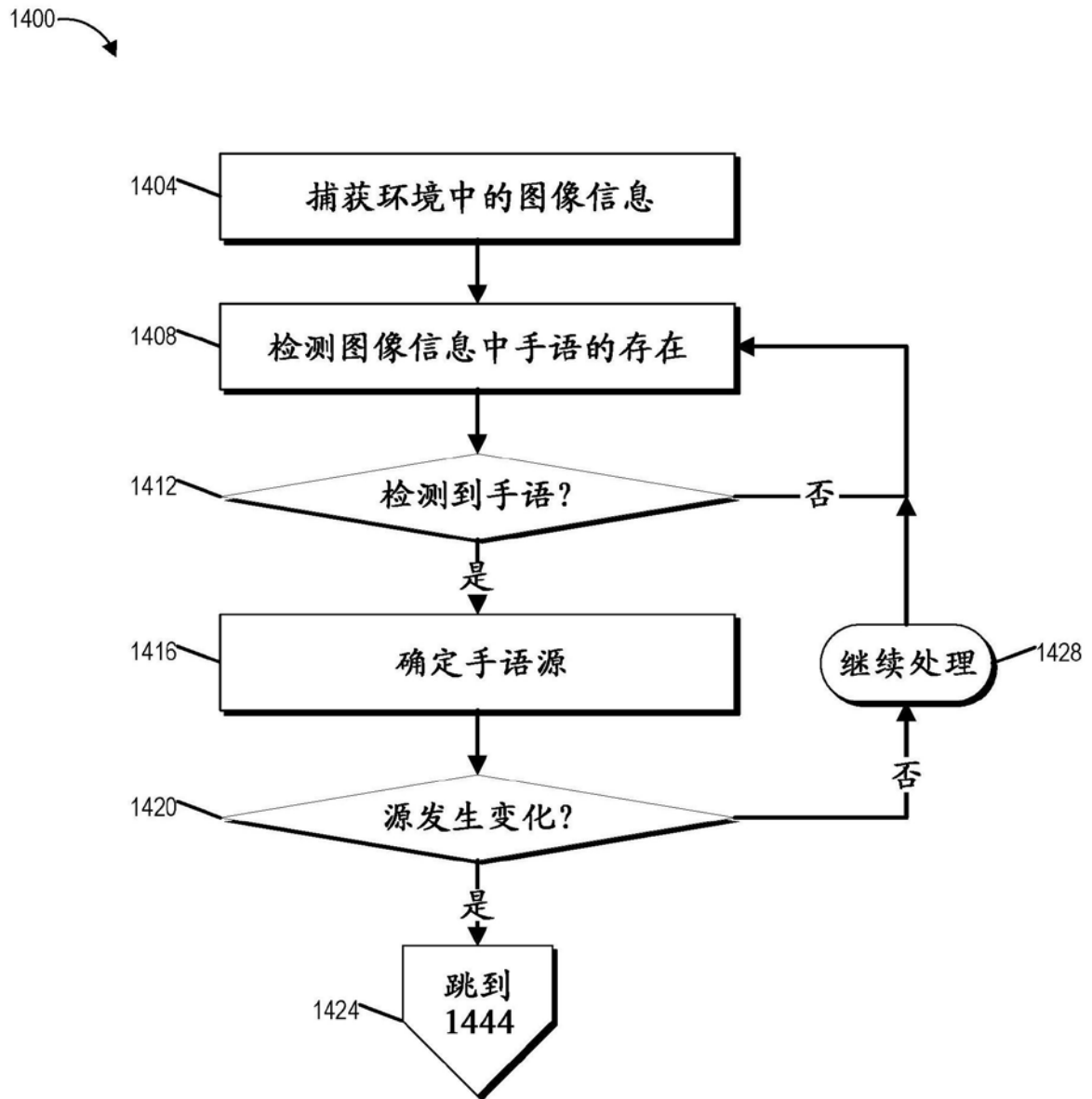


图14A

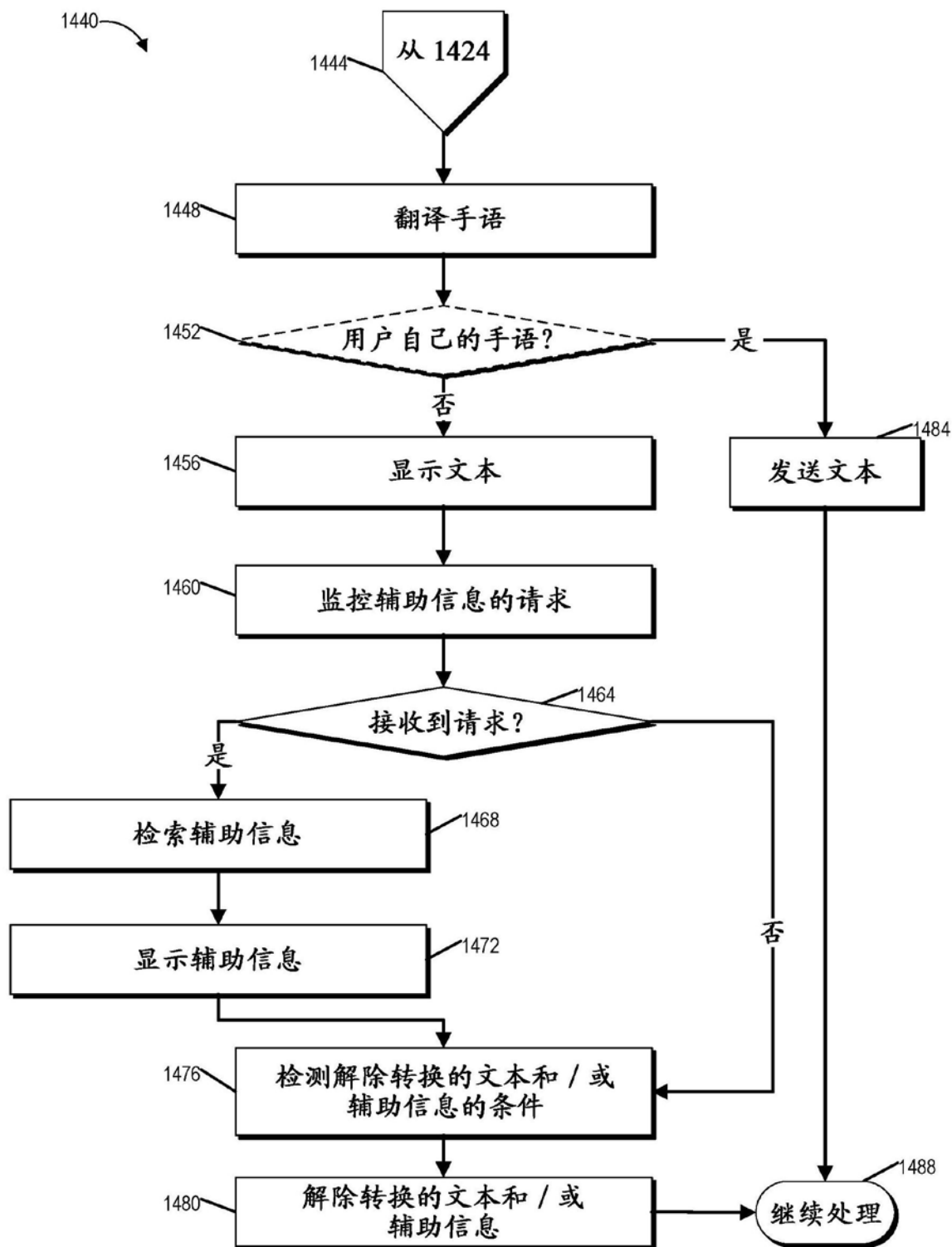


图14B

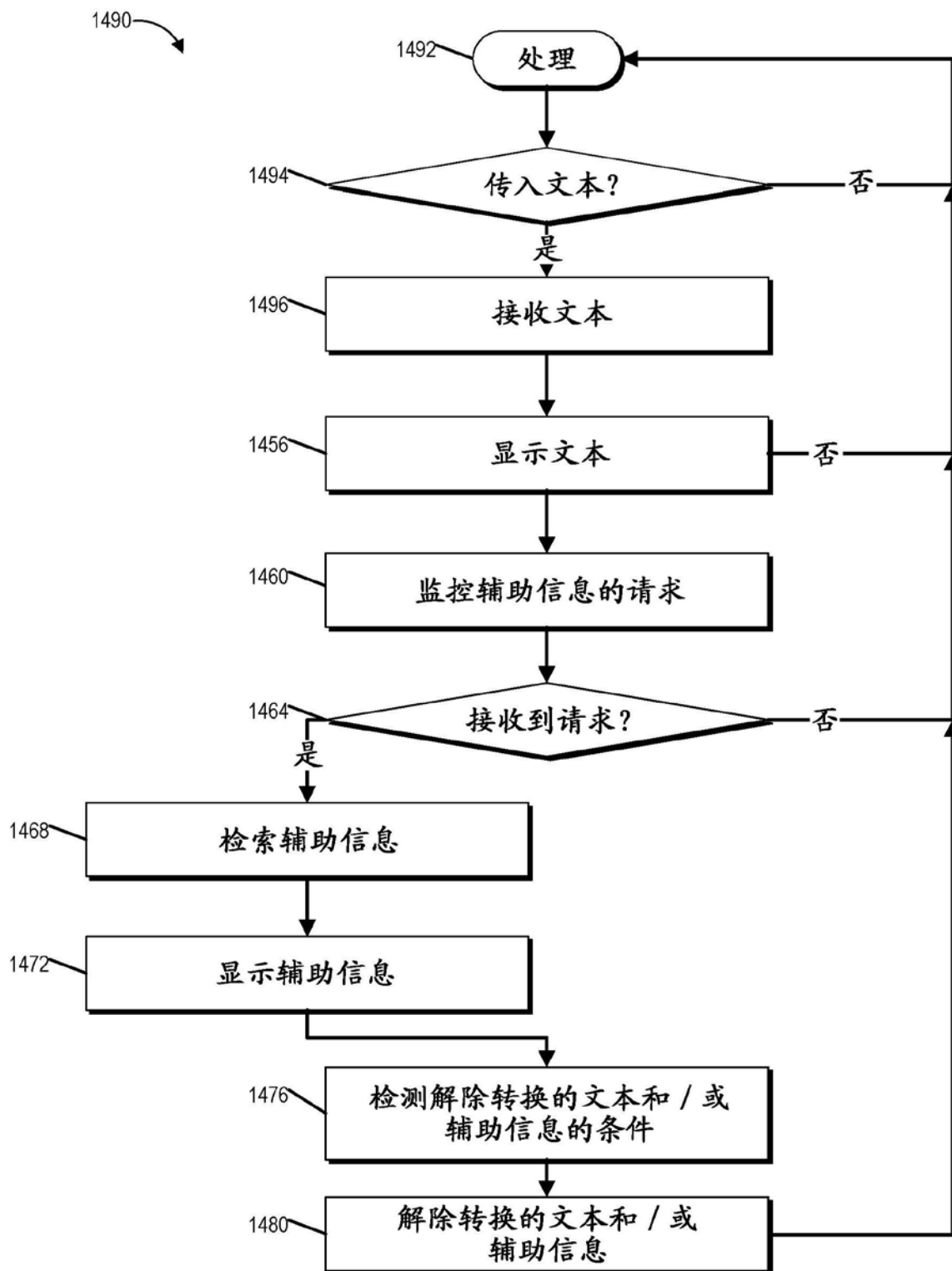


图14C

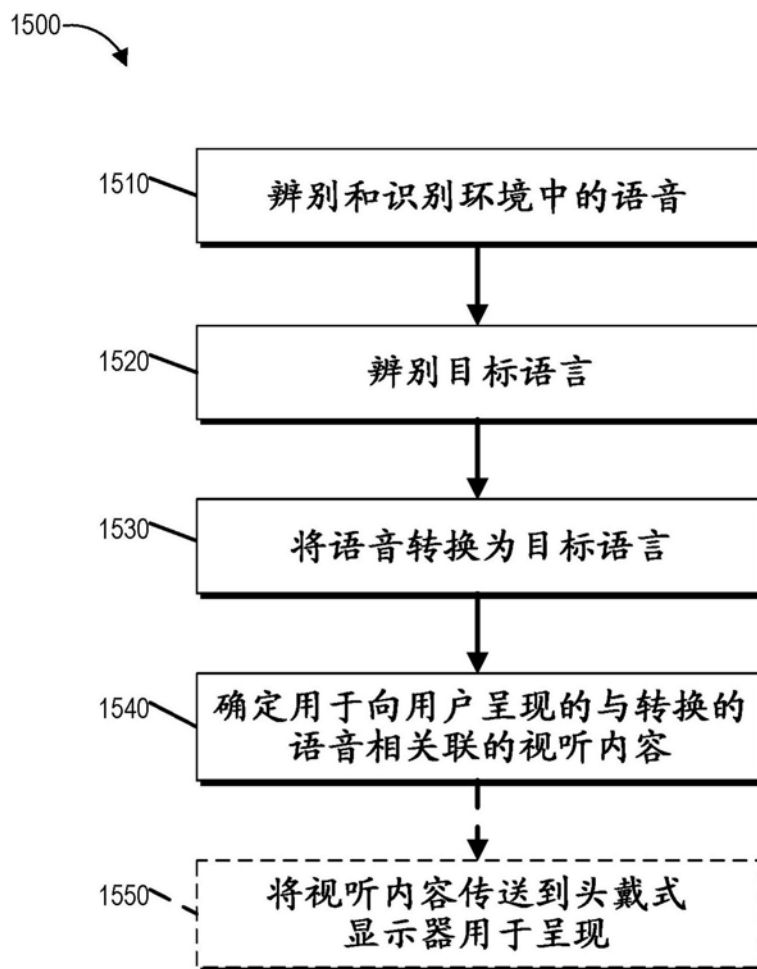


图15

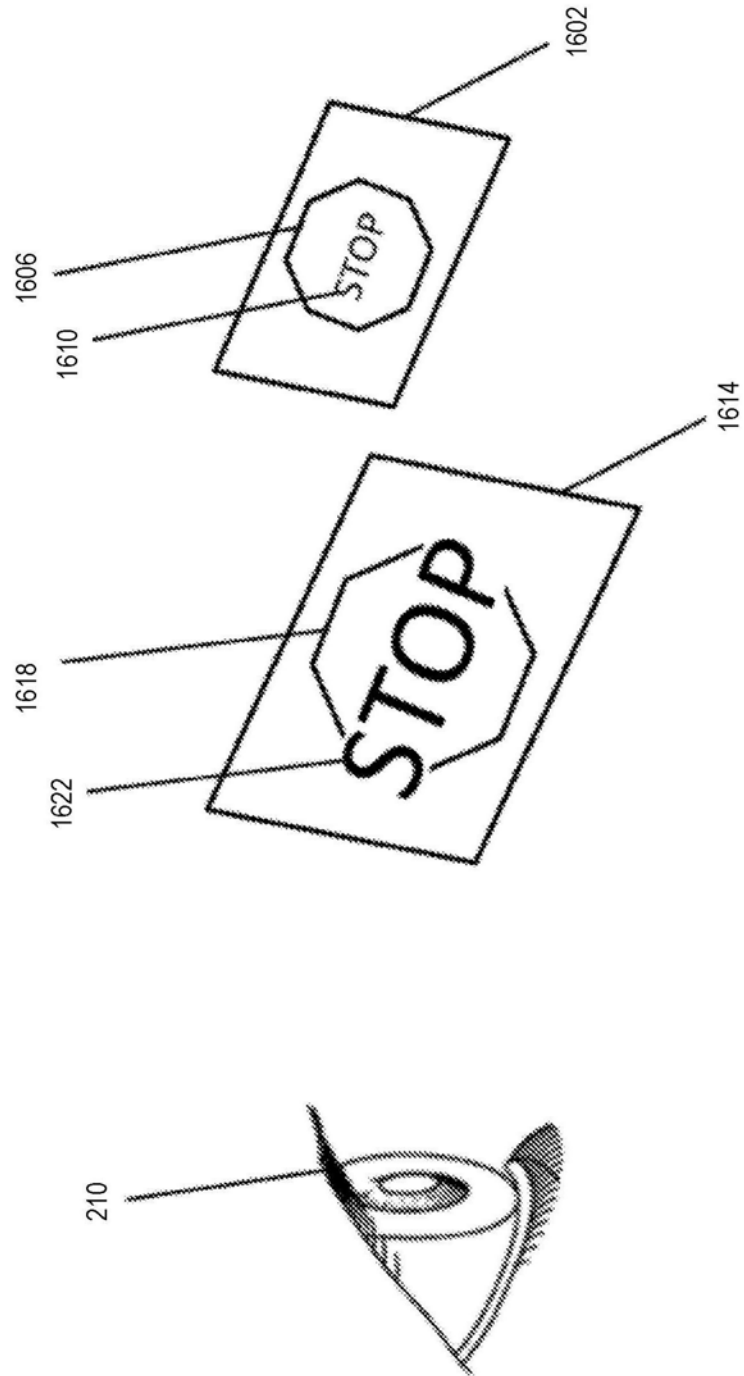


图16A

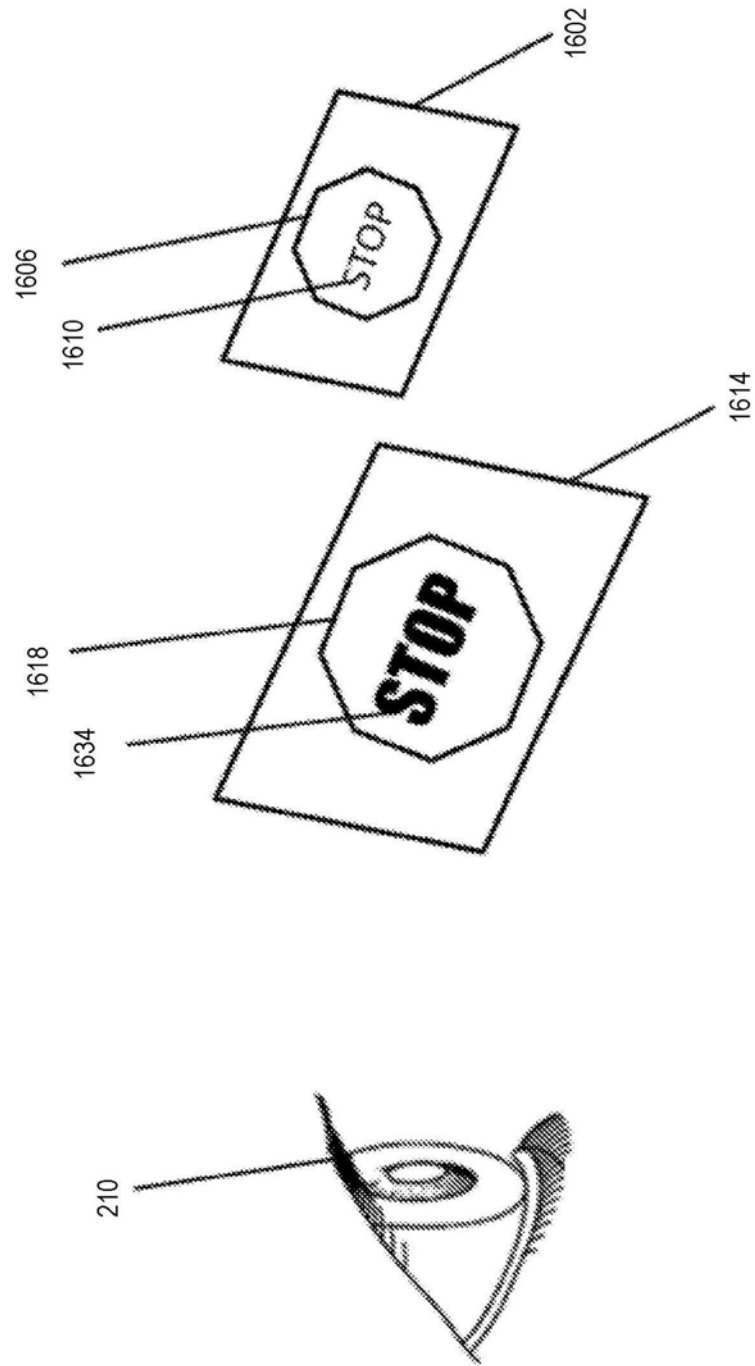


图16B

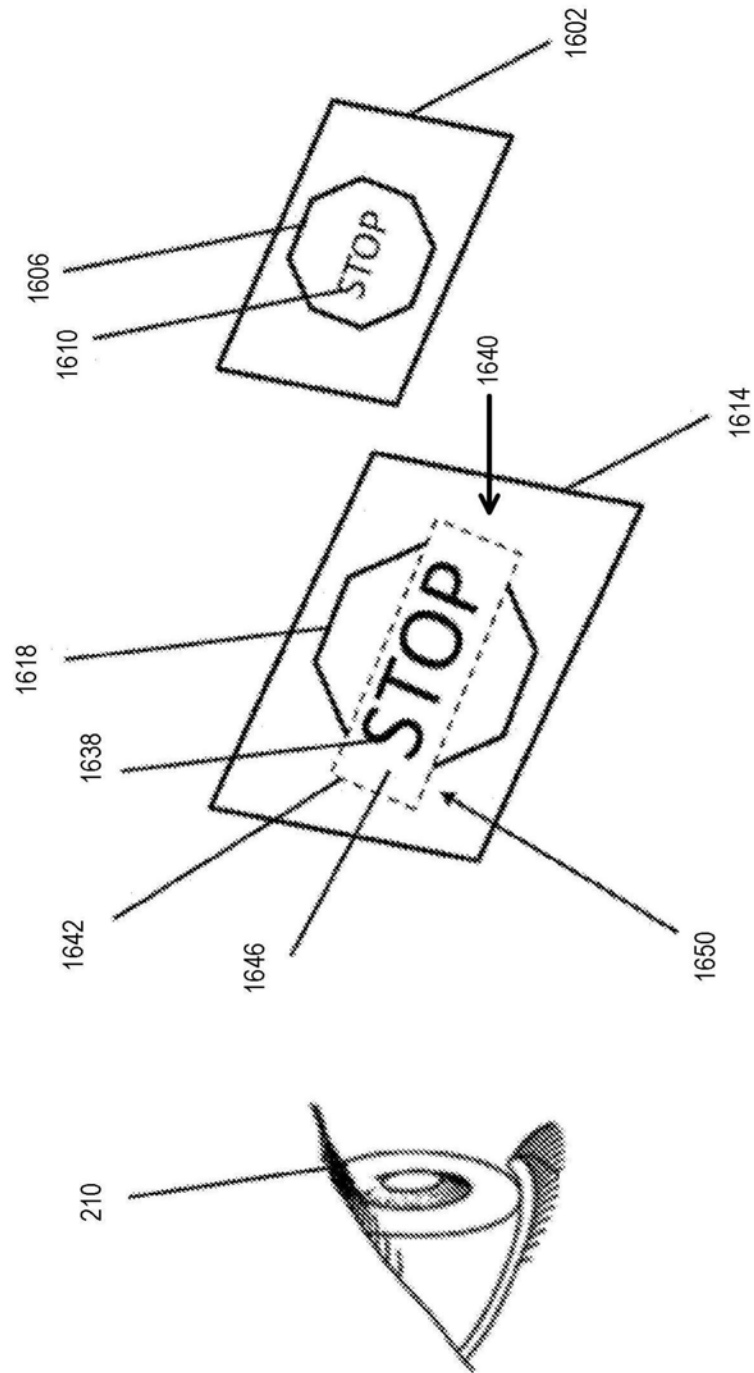


图16C

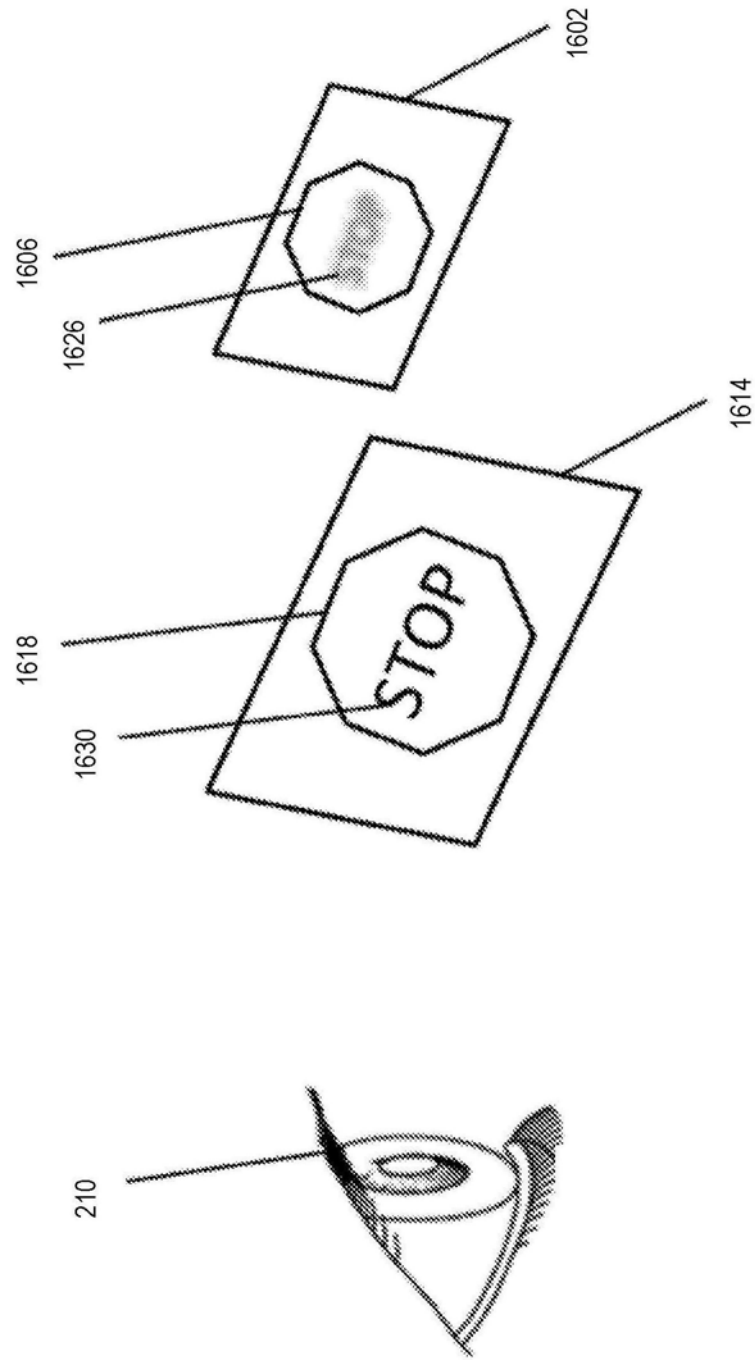


图16D

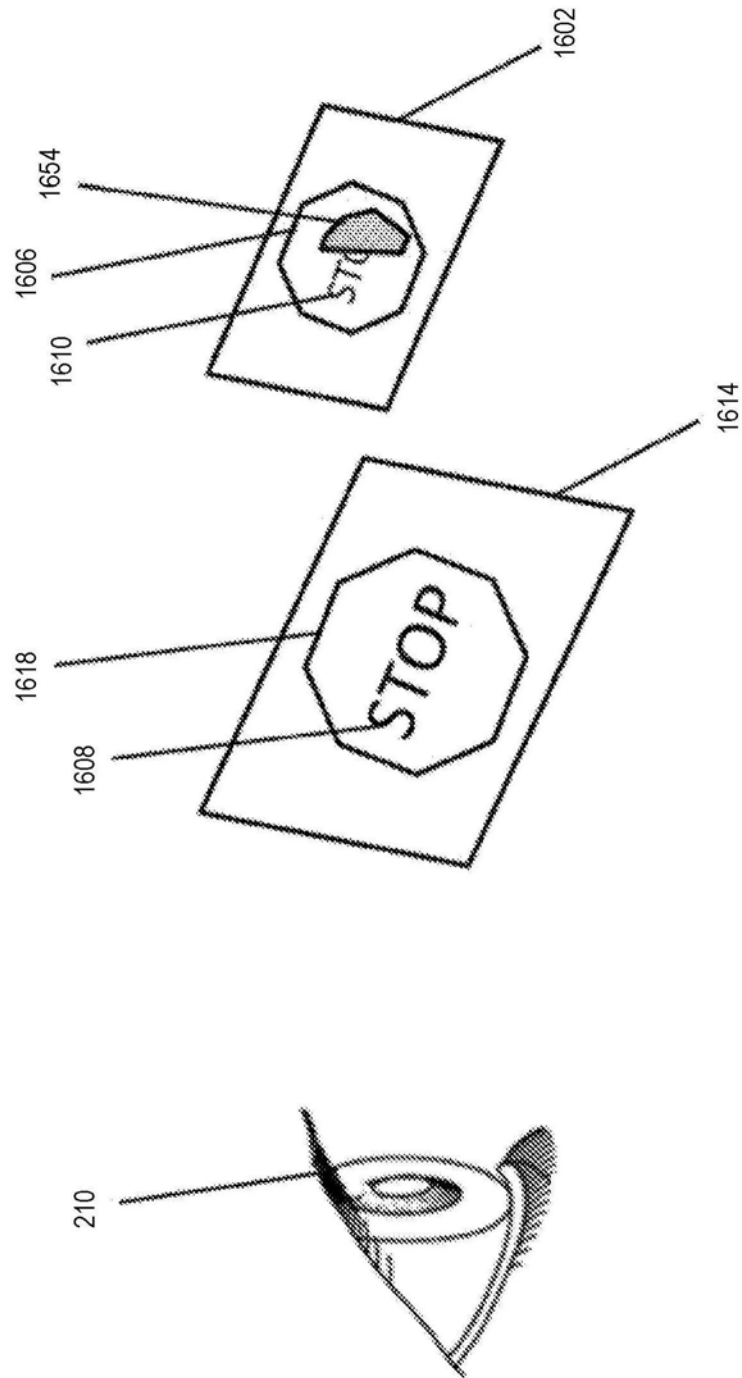


图16E

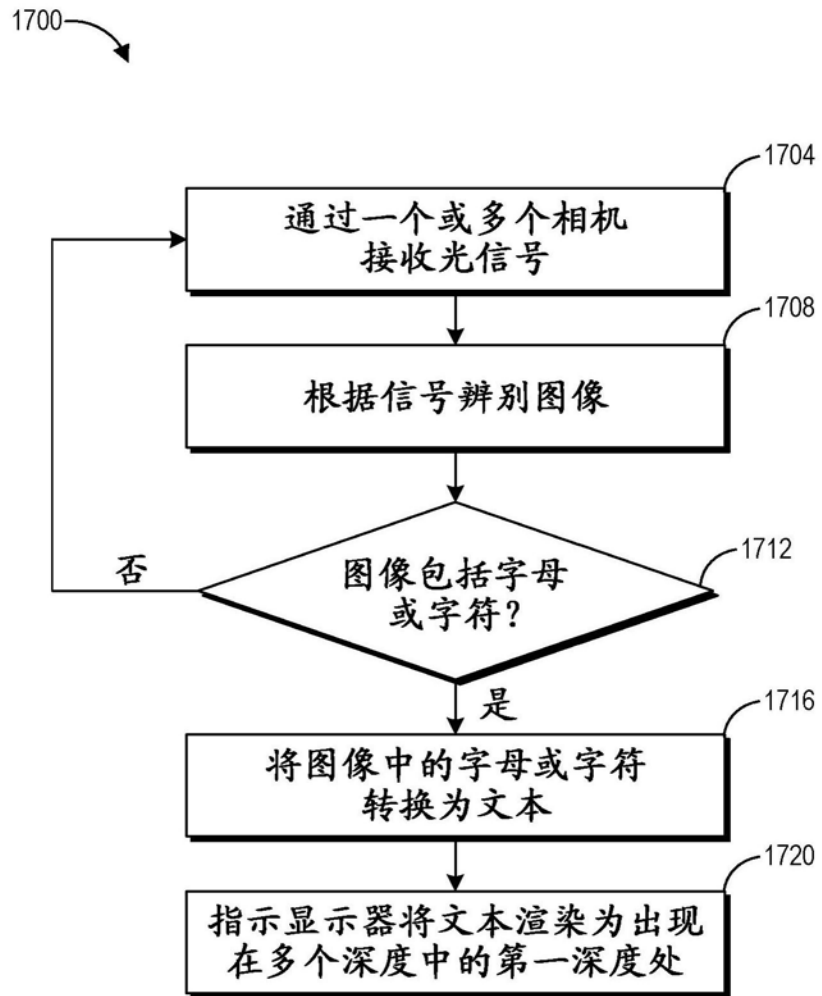


图17

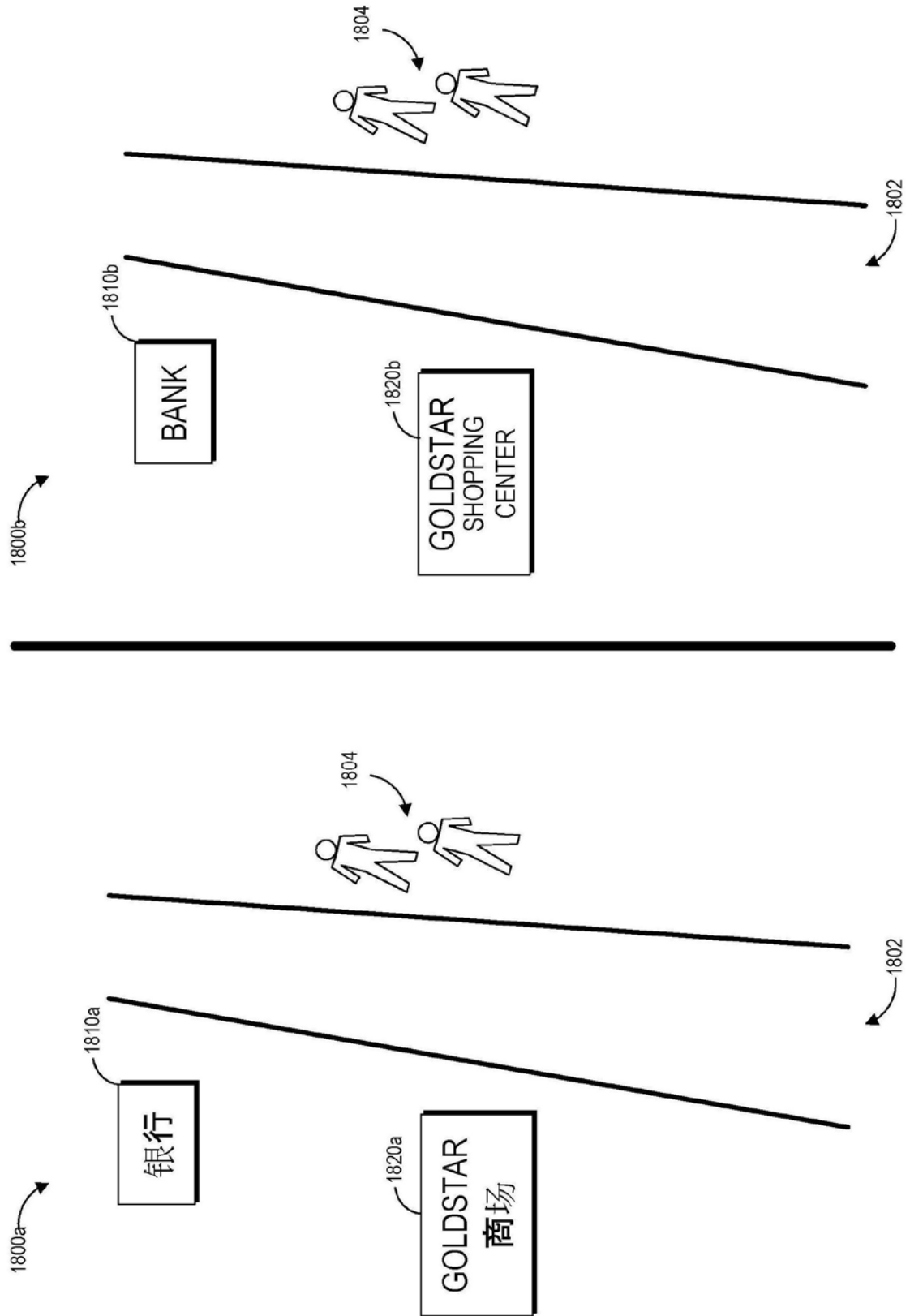


图18

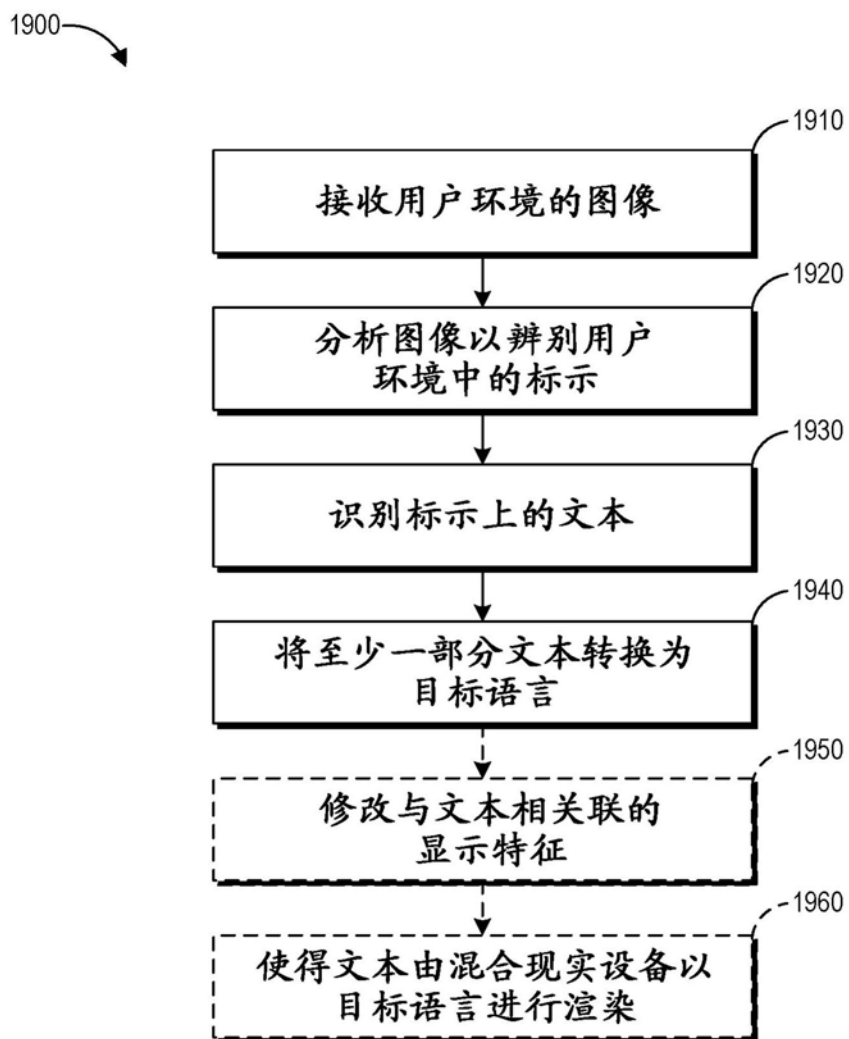


图19