(19)中华人民共和国国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 109271914 A
(43)申请公布日 2019.01.25

(21)申请号 201811044578.1

(22)申请日 2018.09.07

(71)申请人 百度在线网络技术(北京)有限公司
地址 100085 北京市海淀区上地十街10号百度大厦三层

(72)发明人 赵彤彤 燕飞龙 王亮 马彧

(74)专利代理机构 北京市铸成律师事务所 11313
代理人 陈建焕 杨瑾瑾

(51)Int.Cl.
G06K 9/00(2006.01)
G06K 9/32(2006.01)
G06F 3/01(2006.01)

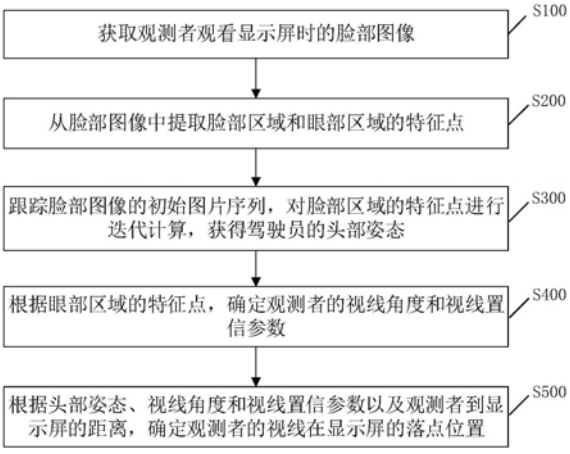
权利要求书2页 说明书15页 附图5页

(54)发明名称

检测视线落点的方法、装置、存储介质和终端设备

(57)摘要

本发明提出一种检测视线落点的方法、装置、存储介质和终端设备，其中，所述方法包括：获取观测者观看显示屏时的脸部图像；从所述脸部图像中提取脸部区域和眼部区域的特征点；跟踪所述脸部图像的初始图片序列，对所述脸部区域的特征点进行迭代计算，获得所述观测者的头部姿态；根据所述眼部区域的特征点，确定所述观测者的视线角度和视线置信参数；根据所述头部姿态、所述视线角度和所述视线置信参数以及所述观测者到所述显示屏的距离，确定所述观测者的视线在所述显示屏的落点位置。采用本发明，可以方便快速准确地确定观测者的视线落点位置。



CN 109271914 A

1.一种检测视线落点的方法,其特征在于,包括:

获取观测者观看显示屏时的脸部图像;

从所述脸部图像中提取脸部区域和眼部区域的特征点;

跟踪所述脸部图像的初始图片序列,对所述脸部区域的特征点进行迭代计算,获得所述观测者的头部姿态;

根据所述眼部区域的特征点,确定所述观测者的视线角度和视线置信参数;

根据所述头部姿态、所述视线角度和所述视线置信参数以及所述观测者到所述显示屏的距离,确定所述观测者的视线在所述显示屏的落点位置。

2.如权利要求1所述的方法,其特征在于,所述脸部图像包括深度图像和彩色图像;以及,所述从所述脸部图像中提取脸部区域和眼部区域的特征点,包括:

从所述深度图像中提取前景区域;

判断所述前景区域是否包括人体脸部;

当所述前景区域包括人体脸部时,在所述深度图像中定位所述人体脸部的位置;以及

从所述人体脸部在所述彩色图像中的位置提取脸部区域和眼部区域的特征点。

3.如权利要求1所述的方法,其特征在于,所述跟踪所述脸部图像的初始图片序列,对所述脸部区域的特征点进行迭代计算,获得所述观测者的头部姿态,包括:

跟踪所述脸部图像的初始图片序列,获得头部姿态的粒子滤波估计值;其中,所述粒子滤波估计值用于估计所述头部姿态;

根据所述粒子滤波估计值,从依据所述观测者的观测范围划分的各视线区域中确定所述头部姿态所朝向的视线区域;以及

基于所述头部姿态所朝向的视线区域和所述粒子滤波估计值,利用最近点迭代算法对所述脸部区域的特征点进行迭代计算,获得所述头部姿态。

4.如权利要求1所述的方法,其特征在于,所述根据所述眼部区域的特征点,确定所述观测者的视线角度和视线置信参数,包括:

依据所述眼部区域的特征点,构建眼部外观;

根据所述眼部外观范围内的特征点,定位瞳孔中心位置;

根据眼球中心位置和所述瞳孔中心位置,确定所述观测者的视线角度;

将所述眼部区域的特征点投影到高维特征空间中,获得眼部特征数据集;

对所述眼部特征数据集进行分类检测,获得视线置线参数。

5.如权利要求1所述的方法,其特征在于,所述根据所述头部姿态、所述视线角度和所述视线置信参数以及所述观测者到所述显示屏的距离,确定所述观测者的眼部视线在所述显示屏的落点位置,包括:

根据所述头部姿态、所述视线角度和所述视线置信参数,计算所述观测者的眼部视线的注视方向;

从所述脸部图像的深度图像中,获取所述眼部区域的瞳孔中心位置的像素值,以确定所述观测者的瞳孔中心到所述显示屏的距离;以及

根据所述注视方向和所述观测者的瞳孔中心到所述显示屏的距离,计算所述观测者的眼部视线在所述显示屏的落点位置。

6.一种检测视线落点的装置,其特征在于,包括:

脸部图像获取模块，用于获取观测者观看显示屏时的脸部图像；

特征点提取模块，用于从所述脸部图像中提取脸部区域和眼部区域的特征点；

头部姿态获取模块，用于跟踪所述脸部图像的初始图片序列，对所述脸部区域的特征点进行迭代计算，获得所述观测者的头部姿态；

视线确定模块，用于根据所述眼部区域的特征点，确定所述观测者的视线角度和视线置信参数；以及

视线落点确定模块，用于根据所述头部姿态、所述视线角度和所述视线置信参数以及所述观测者到所述显示屏的距离，确定所述观测者的视线在所述显示屏的落点位置。

7.如权利要求6所述的装置，其特征在于，所述脸部图像包括深度图像和彩色图像；以及，所述特征点提取模块包括：

前景区域提取单元，用于从所述深度图像中提取前景区域；

人体脸部判断单元，用于判断所述前景区域是否包括人体脸部；

脸部位置定位单元，用于当所述前景区域包括人体脸部时，在所述深度图像中定位所述人体脸部的位置；以及

脸部眼部特征提取单元，用于从所述人体脸部在所述彩色图像中的位置提取脸部区域和眼部区域的特征点。

8.如权利要求6所述的装置，其特征在于，所述头部姿态获取模块包括：

粒子滤波单元，用于跟踪所述脸部图像的初始图片序列，获得头部姿态的粒子滤波估计值；其中，所述粒子滤波估计值用于估计所述头部姿态；

当前视线区域确定单元，用于根据所述粒子滤波估计值，从依据所述观测者的观测范围划分的各视线区域中确定所述头部姿态所朝向的视线区域；以及

头部姿态迭代单元，用于基于所述头部姿态所朝向的视线区域和所述粒子滤波估计值，利用最近点迭代算法对所述脸部区域的特征点进行迭代计算，获得所述头部姿态。

9.如权利要求6所述的装置，其特征在于，所述视线落点确定模块包括：

注视方向确定单元，用于根据所述头部姿态、所述视线角度和所述视线置信参数，计算所述观测者的眼部视线的注视方向；

距离获取单元，用于从所述脸部图像的深度图像中，获取所述眼部区域的瞳孔中心位置的像素值，以确定所述观测者的瞳孔中心到所述显示屏的距离；以及

落点计算单元，用于根据所述注视方向和所述观测者的瞳孔中心到所述显示屏的距离，计算所述观测者的眼部视线在所述显示屏的落点位置。

10.一种实现检测视线落点的终端设备，其特征在于，所述终端设备包括：

一个或多个处理器；

存储装置，用于存储一个或多个程序；

当所述一个或多个程序被所述一个或多个处理器执行时，使得所述一个或多个处理器实现如权利要求1-5中任一所述的检测视线落点的方法。

11.一种计算机可读存储介质，其存储有计算机程序，其特征在于，该程序被处理器执行时实现如权利要求1-5中任一所述的检测视线落点的方法。

检测视线落点的方法、装置、存储介质和终端设备

技术领域

[0001] 本发明涉及计算机技术领域，尤其涉及一种检测视线落点的方法、装置、存储介质和终端设备。

背景技术

[0002] 人们往往通过频繁的注视来表达对对象或某一视觉区域的兴趣。通过跟踪人眼的视点运动，收集和分析人眼运动数据数据，获得用户兴趣。例如，在用户观看视频、电子书或网上商店时，通过跟踪用户的人眼运动，可以获知判断用户的偏好。其中，眼睛的视线落点可以传递大量反映人体心理活动的信息。

[0003] 现有技术中，为了准确地测量眼睛的视线落点，通常要将被测试者的头部固定以固定头部姿态。或者，被测试者的头部戴上头戴式装置来跟踪头部姿态的变化数据，进而准确地计算视线落点。

[0004] 但是，上述方案会给被测试者带来不便。仅利用二维彩色图像进行检测，所获取的数据有可能存在错误，进而导致视线落点的确定不准确。因此，如何方便准确地计算视线落点是目前亟需解决的技术问题。

发明内容

[0005] 本发明实施例提供一种检测视线落点的方法、装置、存储介质和终端设备，以解决或缓解现有技术中的以上一个或多个技术问题。

[0006] 第一方面，本发明实施例提供了一种检测视线落点的方法，包括：

[0007] 获取观测者观看显示屏时的脸部图像；

[0008] 从所述脸部图像中提取脸部区域和眼部区域的特征点；

[0009] 跟踪所述脸部图像的初始图片序列，对所述脸部区域的特征点进行迭代计算，获得所述观测者的头部姿态；

[0010] 根据所述眼部区域的特征点，确定所述观测者的视线角度和视线置信参数；

[0011] 根据所述头部姿态、所述视线角度和所述视线置信参数以及所述观测者到所述显示屏的距离，确定所述观测者的视线在所述显示屏的落点位置。

[0012] 结合第一方面，在第一方面的第一种实施方式中，所述脸部图像包括深度图像和彩色图像；以及，所述从所述脸部图像中提取脸部区域和眼部区域的特征点，包括：

[0013] 从所述深度图像中提取前景区域；

[0014] 判断所述前景区域是否包括人体脸部；

[0015] 当所述前景区域包括人体脸部时，在所述深度图像中定位所述人体脸部的位置；以及

[0016] 从所述人体脸部在所述彩色图像中的位置提取脸部区域和眼部区域的特征点。

[0017] 结合第一方面，在第一方面的第二种实施方式中，所述跟踪所述脸部图像的初始图片序列，对所述脸部区域的特征点进行迭代计算，获得所述观测者的头部姿态，包括：

[0018] 跟踪所述脸部图像的初始图片序列，获得头部姿态的粒子滤波估计值；其中，所述粒子滤波估计值用于估计所述头部姿态；

[0019] 根据所述粒子滤波估计值，从依据所述观测者的观测范围划分的各视线区域中确定所述头部姿态所朝向的视线区域；以及

[0020] 基于所述头部姿态所朝向的视线区域和所述粒子滤波估计值，利用最近点迭代算法对所述脸部区域的特征点进行迭代计算，获得所述头部姿态。

[0021] 结合第一方面，在第一方面的第三种实施方式中，所述根据所述眼部区域的特征点，确定所述观测者的视线角度和视线置信参数，包括：

[0022] 依据所述眼部区域的特征点，构建眼部外观；

[0023] 根据所述眼部外观范围内的特征点，定位瞳孔中心位置；

[0024] 根据眼球中心位置和所述瞳孔中心位置，确定所述观测者的视线角度；

[0025] 将所述眼部区域的特征点投影到高维特征空间中，获得眼部特征数据集；

[0026] 对所述眼部特征数据集进行分类检测，获得视线置线参数。

[0027] 结合第一方面，在第一方面的第四种实施方式中，所述根据所述头部姿态、所述视线角度和所述视线置信参数以及所述观测者到所述显示屏的距离，确定所述观测者的眼部视线在所述显示屏的落点位置，包括：

[0028] 根据所述头部姿态、所述视线角度和所述视线置信参数，计算所述观测者的眼部视线的注视方向；

[0029] 从所述脸部图像的深度图像中，获取所述眼部区域的瞳孔中心位置的像素值，以确定所述观测者的瞳孔中心到所述显示屏的距离；以及

[0030] 根据所述注视方向和所述观测者的瞳孔中心到所述显示屏的距离，计算所述观测者的眼部视线在所述显示屏的落点位置。

[0031] 第二方面，本发明实施例提供一种检测视线落点的装置，包括：

[0032] 脸部图像获取模块，用于获取观测者观看显示屏时的脸部图像；

[0033] 特征点提取模块，用于从所述脸部图像中提取脸部区域和眼部区域的特征点；

[0034] 头部姿态获取模块，用于跟踪所述脸部图像的初始图片序列，对所述脸部区域的特征点进行迭代计算，获得所述观测者的头部姿态；

[0035] 视线确定模块，用于根据所述眼部区域的特征点，确定所述观测者的视线角度和视线置信参数；以及

[0036] 视线落点确定模块，用于根据所述头部姿态、所述视线角度和所述视线置信参数以及所述观测者到所述显示屏的距离，确定所述观测者的视线在所述显示屏的落点位置。

[0037] 结合第二方面，在第二方面的第一种实施方式中，所述脸部图像包括深度图像和彩色图像；以及，所述特征点提取模块包括：

[0038] 前景区域提取单元，用于从所述深度图像中提取前景区域；

[0039] 人体脸部判断单元，用于判断所述前景区域是否包括人体脸部；

[0040] 脸部位置定位单元，用于当所述前景区域包括人体脸部时，在所述深度图像中定位所述人体脸部的位置；以及

[0041] 脸部眼部特征提取单元，用于从所述人体脸部在所述彩色图像中的位置提取脸部区域和眼部区域的特征点。

[0042] 结合第二方面，在第二方面的第二种实施方式中，所述头部姿态获取模块包括：

[0043] 粒子滤波单元，用于跟踪所述脸部图像的初始图片序列，获得头部姿态的粒子滤波估计值；其中，所述粒子滤波估计值用于估计所述头部姿态；

[0044] 当前视线区域确定单元，用于根据所述粒子滤波估计值，从依据所述观测者的观测范围划分的各视线区域中确定所述头部姿态所朝向的视线区域；以及

[0045] 头部姿态迭代单元，用于基于所述头部姿态所朝向的视线区域和所述粒子滤波估计值，利用最近点迭代算法对所述脸部区域的特征点进行迭代计算，获得所述头部姿态。

[0046] 结合第二方面，在第二方面的第三种实施方式中，所述视线落点确定模块包括：

[0047] 注视方向确定单元，用于根据所述头部姿态、所述视线角度和所述视线置信参数，计算所述观测者的眼部视线的注视方向；

[0048] 距离获取单元，用于从所述脸部图像的深度图像中，获取所述眼部区域的瞳孔中心位置的像素值，以确定所述观测者的瞳孔中心到所述显示屏的距离；以及

[0049] 落点计算单元，用于根据所述注视方向和所述观测者的瞳孔中心到所述显示屏的距离，计算所述观测者的眼部视线在所述显示屏的落点位置。

[0050] 所述装置的功能可以通过硬件实现，也可以通过硬件执行相应的软件实现。所述硬件或软件包括一个或多个与上述功能相对应的模块。

[0051] 在一个可能的设计中，检测视线落点的结构中包括处理器和存储器，所述存储器用于检测视线落点的装置执行上述检测视线落点的程序，所述处理器被配置为用于执行所述存储器中存储的程序。所述检测视线落点的装置还可以包括通信接口，用于检测视线落点的装置与其他设备或通信网络通信。

[0052] 第三方面，本发明实施例还提供一种计算机可读存储介质，用于检测视线落点的装置所用的计算机软件指令，其中包括用于执行上述检测视线落点的方法所涉及的程序。

[0053] 上述技术方案中的任意一个技术方案具有如下优点或有益效果：

[0054] 本发明实施例通过获取观测者的脸部图像以及脸部图像的初始图片序列，即可以确定观测者的视线在显示屏中的落点位置，无需要观测者固定头部位置或者需要观测者带上相应的检测装置，方便检测。此外，结合头部姿态、以及视线角度、视线置信参数来确定显示屏的落点位置，可以有效地提高准确程度。

[0055] 上述概述仅仅是为了说明书的目的，并不意图以任何方式进行限制。除上述描述的示意性的方面、实施方式和特征之外，通过参考附图和以下的详细描述，本发明进一步的方面、实施方式和特征将会是容易明白的。

附图说明

[0056] 在附图中，除非另外规定，否则贯穿多个附图相同的附图标记表示相同或相似的部件或元素。这些附图不一定是按照比例绘制的。应该理解，这些附图仅描绘了根据本发明公开的一些实施方式，而不应将其视为是对本发明范围的限制。

[0057] 图1是本发明提供的检测视线落点的方法的一个实施例的流程示意图；

[0058] 图2是本发明提供的提取特征点的过程的一个实施例的流程示意图；

[0059] 图3是本发明提供的头部姿态的过程的一个实施例的流程示意图；

[0060] 图4是本发明提供的划分显示屏的一个实施例的示意图；

[0061] 图5是本发明提供的视线角度与视线置信参数的获取过程的一个实施例的流程示意图；

[0062] 图6是本发明提供的眼部外观的一个示意图；

[0063] 图7是本发明提供的3D眼球模型的一个示意图；

[0064] 图8是本发明提供的落点位置的计算过程的一个实施例的流程示意图；

[0065] 图9是本发明提供的多人眼动仪设备的一个应用示例的结构示意图；

[0066] 图10是本发明提供的检测视线落点的装置的另一个实施例的结构示意图；

[0067] 图11是本发明提供的终端设备的一个实施例的结构示意图。

具体实施方式

[0068] 在下文中，仅简单地描述了某些示例性实施例。正如本领域技术人员可认识到的那样，在不脱离本发明的精神或范围的情况下，可通过各种不同方式修改所描述的实施例。因此，附图和描述被认为本质上是示例性的而非限制性的。

[0069] 请参阅图1，本发明实施例提供了一种检测视线落点的方法。本实施例可以应用在人机交互的领域中：例如，在用户浏览电子书、视频、网上商店等情况下获取用户的兴趣。本实施例也可以应用于在汽车、轮船、飞机等交通工具中驾驶员在驾驶时的注意力检测场景中。本实施例包括步骤S100至步骤S500，具体如下：

[0070] S100，获取观测者观看显示屏时的脸部图像。

[0071] 在本实施例中，显示屏可以包括电子书、平板、个人电脑、笔记本、车载终端等设备的显示器。可以通过摄像头获取观测者的上半身的图像或者仅仅包含脸部的图像。摄像头可以包括普通彩色摄像头、红外摄像头等。例如，彩色摄像头可以获取彩色图像。红外摄像镜头可以获取深度图像。结合两者可以对驾驶员的脸部数据进行二维、三维数据的整合。

[0072] S200，从脸部图像中提取脸部区域和眼部区域的特征点。

[0073] 在本实施例中，本实施例可以采用ASM(Active Shape Model，主动形状模型)、AAM(Active Appearance Model，主动外观模型)算法来获取脸部的特征点。

[0074] 以AAM方法为例，其是广泛应用于模式识别领域的一种特征点提取方法。基于AAM的人脸特征定位方法在建立人脸模型过程中，不但考虑局部特征信息，而且综合考虑到全局形状和纹理信息，通过对人脸形状特征和纹理特征进行统计分析，建立人脸混合模型，即为最终对应的AAM模型。在图像的匹配过程中，为了能够既快速又准确地进行人脸特征标定，在对被测试人脸对象进行特征点定位时采取一种图像匹配拟合的方法，可形象概括为“匹配→比较→调整再匹配→再比较”的过程。AAM算法主要分为AAM建模和AAM匹配计算两部分。AAM建模所建立的模型是对象的主动表观模型。所谓表观模型，就是在AAM形状模型的基础上结合所提取人脸对象的纹理信息建立起来的人脸模型；“主动”一词则形象具体地体现在AAM匹配计算中。

[0075] 首先，采用主成分分(Principal Component Analysis，PCA)方法来描述形状特征点的动态变化。形状特征点可以表明了人脸特征的位置。其次，以具体AAM模型实例与输入图像的均方差来定义一个能量函数，利用该能量函数来评价AAM模型的匹配程度。在人脸定位匹配的过程中，模型的线性表达式可以有效地利用匹配算法变化模型参数组，以控制形状特征点的位置变化来生成当前新的AAM模型实例。再次，利用当前得到的能量函数值来更

新模型的参数，如此反复迭代以实现能量函数的最小化。从而达到模型实例与输入图像相匹配的目的，最终得到的形状特征点位置就描述了当前人脸图像的特征点位置。

[0076] 由于驾驶员在张嘴、闭眼等较大表情动作时，利用非刚性头部模型，AAM方法是无法从脸部图像中检测出形成脸部的点云数据。因此，AM方法可以在此种情况下，找出三维头像的刚性头部模型来代替非刚性头部模型。可以避免获取到错误的特征点，提高头部姿态确定的准确度。

[0077] S300，跟踪脸部图像的初始图片序列，对脸部区域的特征点进行迭代计算，获得驾驶员的头部姿态。

[0078] 本实施例可以通过粒子滤波算法、最近点迭代算法(Iterative Closet Point，ICP)等方法的结合预估头部姿态。粒子滤波(Particle Filter，PF)算法，通过寻找一组在状态空间中传播的随机样本来近似的表示概率密度函数，用样本均值代替积分运算，进而获得样本状态的最小方差估计的过程。这此样本可以形象地称为“粒子”，故称，粒子滤波。基本的粒子滤波算法包括：最优贝叶斯估计算法、顺序重要采样算法、辅助采样-重采样算法、规则化采样算法、自适应粒子滤波算法等。本实施例可以采用最近点迭代算法(Iterative Closet Point，ICP)对头部姿态进行迭代计算。例如、Point to Point(点对点)搜索算法、Point to Plane(点对面)搜索算法、Point to Projection(点到对象)搜索算法等。采用最近点迭代算法可以从测量点集中确定其对应的最近点点集后，运用基于自由形态曲面的配准算法来计算新的最近点点集，直到残差平方和所构成的目标函数值不变，结束迭代过程。

[0079] 初始图片序列可以包括拍摄脸部图像之前的第一帧或连续几帧的图片序列。可以用$y_{1:t}=\{y_1,\ldots,y_t\}$来表示。头部姿态可以以三维角度或向量表达的形式来表示。

[0080] S400，根据眼部区域的特征点，确定观测者的视线角度和视线置信参数。

[0081] 在本实施例中，可以采用SVM(Support Vector Machine，支持向量机)算法等机器学习的算法，视线置信参数。在观测者佩戴眼镜的时，难以准确地计算出观测的瞳孔中心，进而视线角度的确定也受到影响。为了估计出视线角度的可靠性，本实施例将影响眼睛的因素加入到视线检测的过程，采用SVM算法检测观测者是否佩戴眼镜。

[0082] S500，根据头部姿态、视线角度和视线置信参数以及观测者到显示屏的距离，确定观测者的视线在显示屏的落点位置。

[0083] 在本实施例中，观测者到显示屏的距离可以通过深度图像来确定。也可以通过特定的红外发射器向观测者的脸部发射红外光，红外接收器接收红外光，进而通过发射角、接收角以及红外发射器到红外接收器之间的距离，确定观测者的脸部到红外接收器之间。

[0084] 本发明实施例能过获取观测者的脸部图像以及脸部图像的初始图片序列，即可以确定观测者的视线在显示屏中的落点位置，无需要观测者固定头部位置或者需要观测者带上相应的检测装置，方便检测。此外，结合头部姿态、以及视线角度、视线置信参数来确定显示屏的落点位置，可以有效地提高准确程度。

[0085] 在一种可能的实现方式中，脸部图像包括深度图像和彩色图像。深度图像和彩色图像是在同一时刻且在同一拍摄角度的情况下获取的驾驶员的上半身图像。如图2所示，上述步骤S200的提取特征点的过程，可以包括步骤S210至步骤S240，如下：

[0086] S210，从深度图像中提取前景区域。

[0087] 在本实施例中，深度图像由点组成的，每一个都是0-255之间的数值。数值代表的是这个点对应的图像到深度相机的距离，可以根据深度图像中每个点的值大小来得到它们到相机的距离。因此，可以利用前景区域与背景区域的深度的不同，从深度图像中提取出前景区域。

[0088] 示例性地，在获取图像数据阶段，可以通过Kinect(微软生产的一个体感外部设备)中不同的摄像头获取RGB图像和深度图像。RGB图像由Kinect的CCD(Charge-coupled Device，电荷耦合器件)镜头获得，而深度图像则由红外探测器镜头获得。其均为人体上半身的图像。

[0089] S220，判断前景区域是否包括人体脸部。

[0090] 本实施例利用AAM或ASM方法等方法，检测前景区域是否包括人脸脸部，可以在短时间内迅速判断RGB图像中是否包括人体脸部。

[0091] S230，当前景区域包括人体脸部时，在深度图像中定位人体脸部的位置。可以根据脸部的各像素点的深度变化情况，从中确定人体脸部的位置。

[0092] S240，从人体脸部在彩色图像中的位置提取脸部区域和眼部区域的特征点。

[0093] 在本实施例中，可以利用AAM(Active Appearance Model，主动外观模型)或ASM(Active Shape Model，主动形状模型)方法从彩色图像中提取特征点。例如：AAM方法可以利用最小二乘法的方法，在匹配、比较、调整，再次匹配、比较、调整的迭代过程后，快速地匹配到新的图像。并利用刚性头部模型进行匹配，可以得到刚性头部点云数据。所谓刚性，是指脸部没有闭眼、张嘴、表情纹等表情。利用刚性头部模型获取到的特征点相比非刚性头部模型更为准确。AAM方法的建模过程以及迭代过程是一种常用的方法，在此不再赘述。

[0094] 本实施例的特征点的提取过程，利用深度图像与彩色图像的结合，不受光线较弱、非均匀光照以及空间狭小等情况的影响，可以准确地获取脸部与眼部特征，大大提高检测视线的落点位置的准确性。

[0095] 在一种可能的实现方式中，如图3所示，上述步骤S300获取头部姿态的过程，可以包括：

[0096] S310，跟踪脸部图像的初始图片序列，获得头部姿态的粒子滤波估计值；其中，粒子滤波估计值用于估计头部姿态。

[0097] 本实施例可以利用粒子滤波算法来估计头部姿态。可以减少后续步骤S330的头部姿态的校准过程的迭代次数，并提高校准的准确度。粒子滤波的过程可以包括以下步骤：

[0098] 第一步，从所述初始图片序列中的首帧图片开始，采样n个初始粒子姿态样本。其中，各所述初始粒子姿态样本的权重值为1/n，采用$\omega_0^{(1)}=\ldots=\omega_0^{(n)}=1/n$表示；各所述初始粒子姿态样本的先验密度为预设值，采用$p(b_0)$表示。初始粒子姿态样本采用$a_0^{(1)},\ldots,a_0^{(n)}$表示。

[0099] 第二步，按照前一帧图片所采样到的各粒子姿态样本的权重值之间的比例，从当前帧图片中重新采样粒子姿态样本。

[0100] 假设首帧图片第0帧图片，当前帧图片为第t帧图片，则可以根据第t-1帧图片所采样到的各粒子姿态样本的权重值之间形成的比例进行采样。例如，粒子1、粒子2和粒子3的权重值分别为2:3:5，则可以粒子1可以按0.2的总体采样比例进行采样，粒子2可以按0.3的总体采样比例进行采样，粒子3可以按0.5的总体采样比例进行采样。

[0101] 第三步，根据前一帧图片的头部姿态向量与当前帧图片的头部姿态向量的关系，

确定新采样到的粒子姿态样本的权重值。

[0102] 在本实施例中，可以采用似然函数来表达两个头部姿态向量之间的关系。例如，以当前帧图片的头部姿态向量b_t为例，可以采用似然函数$p(x_t|b_t)$进行表示，此函数采用特征向量表达方式的统计纹理，量化$x(b_t)$的纹理一致性。$x()$表示与形状无关的纹理。$p(x_t|b_t)$可以表示如下：

[0103] $$p(x_t|b_t)=c\exp(-\frac{1}{2}\sum_{i=1}^{M}\frac{\xi_i^2}{\lambda_i^2})\exp(-\frac{e}{2\rho});$$

[0104] 其中，c为似然函数的常数，c＞0，e是重构误差，λ_i^2是和第一个特征向量M相关的特征值，ξ是似然函数的估计量，ρ是其余特征值的算术平均值。

[0105] 进而，可以采用动态模型$p(b_t|b_{t-1})=a_{t-1}^{(j)}$描述前一帧图片的头部姿态向量$b_{t-1}$与当前帧图片的头部姿态向量$b_t$的关系。

[0106] 第四步，利用最大后验概率估计公式，对新采样到的粒子姿态样本的权重值进行计算，获得下一帧图片的头部姿态向量的预测值。具体地，公式如下：

[0107] $$\hat{b}_t=\arg\max_{b_t}p(b_t|y_{1:t})\approx\arg\max_{a_t^{(j)}}\omega_t^{(j)};$$

[0108] 其中，$\omega_t^{(j)}$是第t帧采到的粒子姿态样本的权重值，$j\in n$。

[0109] 除了第0帧图片，对于每一帧图片均可以上述第二步至第三步进行计算，直到计算出最后一帧图片的头部姿态向量的预测值。步骤S320和步骤S330可以利用头部姿态向量的预测值进行后续的计算。

[0110] 具体地，上述第三步可以包括如下：

[0111] 首先，根据动态模型$p(b_t|b_{t-1})=a_{t-1}^{(j)}$，画出新采样到的粒子姿态样本的近似预测分布$p(b_t|x_{1:(t-1)})$。$x_{1:(t-1)}$表示第1帧图片到第t-1帧图片的与形状无关的纹理。

[0112] 然后，根据所述近似预测分布$p(b_t|x_{1:(t-1)})$，计算新采样到的粒子姿态样本的几何相似特征$x(b_t)$；

[0113] 进而，利用似然函数，对新采样到的粒子姿态样本的几何相似特征进行量化，获得新采样到的粒子姿态样本的似然值。对于第j个粒子姿态样本，其似然值可以表示：$p(x_t|b_t)=a_{t-1}^{(j)}$。

[0114] 最后，按照新采样到的粒子姿态样本的似然值的比例分布，确定新采样到的粒子姿态样本的权重值。对于第j个粒子姿态样本，其确定的权重值为：

[0115] $$\omega_t^j=\frac{p(x_t|b_t)=a_{t-1}^{(j)}}{\sum_{m=1}^{J}p(x_t|b_t=a_{t-1}^{(m)})}$$

[0116] 通过在当前帧的时间，用n对采样到的粒子$(a_{t-1}^{(j)},\omega_{t-1}^j)$加权，使得加权后的粒子近似先前时间点的头部姿态向量$p(b_{t-1}|x_{1:(t-1)})$的后验分布。

[0117] S320，根据粒子滤波估计值，从依据观测者的观测范围划分的各视线区域中确定头部姿态所朝向的视线区域。

[0118] 在本实施例中，可以预先将观测者观测显示屏的观测范围划分成多个视线区域，例如，划分成6、9、12等数量的区域。如图4所示，显示屏划分成9个区域。在划分区域时，可以按显示屏的尺寸大小进行划分。尺寸越大划分的区域数量越多。

[0119] 以及，在划分视线区域的同时，可以预先为各视线区域建立头部姿态模板。例如，利用自学习算法，对预设划分的各视线区域的头部姿态进行分类，建立索引。每个视线区域的头部姿态模板可以包括1个或多个。每个头部姿态模板可以对应一个头部姿态向量。由于头部姿态的粒子滤波估计值，实际上也表示为头部姿态向量。因此，计算各视线区域的各头部姿态模板的头部姿态向量与粒子滤波估计值之间的空间距离，可以得到各视线区域的距离概率分布。依据距离概率分布，可以确定头部姿态所朝向的视线区域。

[0120] 具体地，本步骤可以包括：在获取到粒子滤波估计值之后，获取预设划分的各视线区域对应的头部姿态模板。然后，测量各头部姿态模板中的各点与粒子滤波估计值的距离，并确定各视线区域的距离概率分布。根据各视线区域的距离概率分布，可以确定头部姿态所朝向的视线区域。

[0121] S330，基于头部姿态所朝向的视线区域和粒子滤波估计值，利用最近点迭代算法对脸部区域的特征点进行迭代计算，获得头部姿态。

[0122] 在本实施例中，除去头部向前或向后移动以及尺度的情况下，头部姿态向量被定义成六维列向量b。其中，$\theta_x,\theta_y,\theta_z$是Yaw，Pitch，Roll三个方向的角度，$t_x,t_y,t_z$是x，y，z轴上的偏差。Yaw方向是指Y轴旋转，Pitch是围绕X轴旋转，Roll是围绕Z轴旋转。

[0123] 本实施例可以采用ICP算法进行迭代计算。粒子滤波估计值实际上一个头部姿态向量，可依据此将其转换成相应的初始转换矩阵，代入到ICP算法进行计算。因此，上述步骤S330的实施过程可以如下：

[0124] (1)，根据粒子滤波估计值，计算初始转换矩阵。

[0125] 在本实施例中，在迭代过程可以将脸部图像的第一帧或者连续图片的前几帧的头部姿态向量，作为初始头部姿态向量。

[0126] 对于初始图片序列，其表示为$y_{1:t}=\{y_1,...,y_t\}$，可以通过粒子滤波算法，获得初始头部姿态向量$\hat{b}_t$的粒子滤波估计值。然后，可以利用此粒子滤波估计值，转换成初始转换矩阵。

[0127] (2)，获取头部姿态所朝向的视线区域以及相邻视线区域的头部姿态模板。

[0128] 在本实施例中，头部姿态的角度可以通过欧几里德角度空间进行表示，即用Yaw，Pitch和Roll来表示。通过计算头部姿态所朝向的视线区域与其他区域的头部姿态模板之间的欧式角度，来确度相邻视线匹配假设视线区域的总数9，则各视线区域的头部姿态模板可以包括：$P^1,P^2,...,P^m,...P^9$。

[0129] (3)，基于所述初始转换矩阵，计算各所述头部姿态模板对应的最优转换矩阵。其中，最优转换矩阵能使所述头部姿态模板与刚性点云数据的两个点集之间的误差最小。第(3)步的计算过程可以分步骤实现，如下：

[0130] (3.1)，对于刚性点云矩阵中的各点，确定属于模板矩阵且与该点最近的点。其中，初始的刚性点云矩阵Q表示所述刚性点云数据，模板矩阵P^m表示所述头部姿态模板。

[0131] 具体地，可以利用NNSP(Nearest Neighbor Search Point，最近邻域搜索点)算法计算两矩阵的最近点。公式如下：

[0132] $\forall p_j \in P^m, q_i = \{q_i \in Q | NNSP(p_j, q_i) = \min\}$；

[0133] 其中，P^m表示模板矩阵，p_j为模块矩阵P^m中的第j个点；q_i为刚性点云矩阵Q中的第i个点。m为头部姿态所朝向的视线区域的序号。

[0134] (3.2)，计算最优转换矩阵，使刚性点云矩阵的各点与对应在模板矩阵的最近点的误差函数结果最小化。具体地，误差函数如下：

[0135] $(\hat{R}, \hat{t}) = \arg\min_{R,t} e(R,t) = \sum_{N_P} \| RP^m + t - Q \|^2$；

[0136] 其中，$(\hat{R}, \hat{t})$表示最优转换矩阵；(R,t)表示前一次迭代计算的最优转换矩阵。第一次计算基于初始转换矩阵。

[0137] (3.3)，如果误差函数结果大于预设的误差阈值，则根据最优转换矩阵和模板矩阵，更新刚性点云矩阵，并返回重新计算最优转换矩阵。

[0138] 经第(3.2)步计算得到的$(\hat{R}, \hat{t})$适用于矩阵Q，更新后的刚性点云矩阵Q可以为：$Q = \hat{R} * P^m + \hat{t}$。将上述(3.2)和(3.3)不断地进行迭代，直至误差函数的耦合误差的变化降低到设定的阈值内，迭代停止。阈值大小为：$e_{i-1} - e_i < \tau$。

[0139] (3.4)，如果误差函数结果小于预设的误差阈值，则输出当前计算的最优转换矩阵，停止迭代。

[0140] 对于视线区域m的相邻视线区域的头部姿态模板，均可以执上述第(3.1)步至第(3.4)步，获得其对应的最优转换矩阵($R_{neighbor}$, $t_{neighbor}$)。

[0141] (4)，对各最优转换矩阵进行加权并求均值，获得所述头部姿态的角度。

[0142] 假设当前确定头部姿态所朝向的视线区域的头部姿态模板和相邻一个区域的头部姿态模板，则最优转换矩阵可以分别表示为：$(\hat{R}, \hat{t})$和($R_{neighbor}$, $t_{neighbor}$)。此时，可以对这两者进行加权并求均值，计算出头部姿态的θ_x, θ_y, θ_z的角度精确值。

[0143] 在一种可能的实现方式中，如图5所示，上述步骤S400的视线角度与视线置信参数的获取过程，可以包括：

[0144] S410，依据眼部区域的特征点，构建眼部外观。

[0145] 在本实施例中，可以按照人眼六边形的概念，从特征点中选取出眼部的六个特征，如图6所示的人眼边缘的六个圆点的位置。包括两个眼角、上眼皮边缘和下眼皮边缘的特征点。本实施例并不限于六边形，也可以包括八边形等能够描述人眼的关键特征的形状。将获取的特征点按位置顺序连接起来，获得一个封闭的多边形。

[0146] S420，根据眼部外观范围内的特征点，定位瞳孔中心位置。

[0147] 在获得一个封闭的六边形后，将检测瞳孔中心位置的范围限定在此封闭的六边形内，可以加快定位速度。

[0148] 进一步地，还可以利用梯度下降法进行定位。具体如下：

[0149] 首先，构建一个可能的瞳孔中心与图像上的梯度信息关系。

[0150] 以C表示可能的瞳孔中心，g_i是在点x_i处的梯度向量。标准化的位移向量为d_i与梯度向量g_i具有相同的方向。因而，我们可以通过标准化向量和梯度向量g_i中计算他们的内积

来提取图像的梯度向量场。其中，此图像上圆形的最优中心点c的像素位置是x_i，$i \in \{1,...,N\}$。此计算过程，如下公式：

[0151] $$c^* = \arg\max_c \left\{ \frac{1}{N} \sum_{i=1}^{N} (d_i^T g_i)^2 \right\} ,$$

[0152] $$d_i = \frac{x_i - c}{\|x_i - c\|_2} ,$$

[0153] $$\forall i : \|g_i\|_2 = 1 .$$

[0154] 进一步地，为了低算法的时间复杂度，我们可以只考虑到梯度向量的主要分量，忽略到梯度均匀的区域。为了获得图像的梯度，我们计算偏导数：

[0155] $$g_i = (\partial I(x_i, y_i)/\partial x_i, \partial I(x_i, y_i)/\partial y_i)^T$$

[0156] 由于瞳孔通常的比皮肤和巩膜的颜色要深很多，我们采用一个权值w_c来为每一个中心点c赋值，颜色深的中心点权值高于较亮区域的权值。计算过程可以如下公式：

[0157] $$\arg\max_c \frac{1}{N} \sum_{i=1}^{N} w_c (d_i^T g_i)^2$$

[0158] 其中，$w_c = I^*(c_x, c_y)$是经过平滑和反向输入的图像在点(c_x, c_y)的灰度值。

[0159] 需要注意的是，图像需要先经过高斯平滑操作，可以防止一些反光的亮点。

[0160] 经过以上计算，可以获得瞳孔中心的位置C。

[0161] S430，根据眼球中心位置和瞳孔中心位置，确定观测者的视线角度。

[0162] 在本实施例中，由于人眼为球形，且眼球中心位置相对于头部的位置是相对不变的。因此，本实施例可以采用人眼3D模型，通过瞳孔中心位置与眼球中心位置组成的射线来估计视线角度。可以如图7所示。

[0163] S440，将眼部区域的特征点投影到高维特征空间中，获得眼部特征数据集。

[0164] 在本实施例中，可以将眼部区域的特征点的二维数据投射到高维空间中。例如，二维数据G投射到空间H，可以表示为HoG。

[0165] S450，对眼部特征数据集进行分类检测，获得视线置线参数。

[0166] 采用SVM算法，对投影到高维特征空间中的特征进行分类。在SVM分类的过程中，可以使眼部的特征以及眼睛周围的信息加入到视线检测中，从而对视线是否可信作出评价，使得检测结果更加准确。

[0167] 在一种可能的实现方式中，如图8所示，上述步骤500的落点位置的计算过程，可以包括：

[0168] S510，根据头部姿态、视线角度和视线置信参数，计算观测者的眼部视线的注视方向。

[0169] 示例性地，Φ_{head}为头部姿态、Φ_{gaze}人脸视线角度以及λ为置信参数，则注视方向可表示为：$V_{gaze} = \Phi_{head} + \lambda \Phi_{gaze}$。

[0170] S520，从脸部图像的深度图像中，获取眼部区域的瞳孔中心位置的像素值，以确定观测者的瞳孔中心到显示屏的距离。

[0171] 由于深度图像由点组成的，每一个都是0-255之间的数值，数值代表的是这个点对

应的图像到深度相机的距离，因此，可以根据深度图像中每个点的值大小来得到它们到相机的距离。

[0172] 基于前述步骤可以确定瞳孔中心位置，则可以依据此瞳孔中心位置在深度图像中找到相应的像素点，并提取此像素点的深度数值。此深度数值表示像素点到摄像头的距离。在确定摄像头与显示屏的位置关系，则可以获得观测者的瞳孔中心到显示屏的距离。

[0173] S530，根据注视方向和观测者的瞳孔中心到显示屏的距离，计算观测者的眼部视线在显示屏的落点位置。

[0174] 由于观测者的瞳孔中心到显示屏的距离为垂直距离，因而可以确定瞳孔中心投射在显示屏上的垂直落点。基于此垂直落点、垂直距离以及注视方向，即可计算出眼部视线落在在显示屏的落点位置。

[0175] 参见图9，其是本发明实施例提供的多人眼动仪设备的一个应用示例的结构示意图。

[0176] 本实施例可以解决现有技术中被试者在使用眼动仪时，必须头戴或是将头部固定的问题。由于头部不需要固定，所以可以同时进行多人眼部信息数据的采集以及多人眼部视线落点检测的工作。

[0177] 多人眼动仪设备包括眼动仪1和支架2。眼动仪1包含主机1.1和深度摄像头1.2。支架2包括可旋转的基座2.1、支撑杆2.2和脸托2.3。

[0178] 深度摄像头1.2为Kinect(Kinect为索尼出品的一个外设产品)摄像头，包含红外线发射装置等所需设备。

[0179] 主机1.1包括位于显示屏1.11、图像采集卡1.12和图像处理模块1.13。显示屏1.11设置在主机表面。图像采集卡1.12和图像处理模块1.13位于显示屏1.11后方且内嵌在主机1.1内部。深度摄像头1.2与图像采集卡1.12相连，图像采集卡1.12还与图像处理模块1.13相连。

[0180] 显示屏1.11用于播放需要观测者观看的图像或者视频。深度摄像头1.12用于采集用户的脸部区域的图像。图像采集卡1.12用来将深度摄像头1.12采集到的实时图像，传到图像处理模块1.13进行相应处理。其中，图像处理模块1.13使用CES-5250作为主处理器，DSP作为协处理器。CES-5250可以控制图像处理过程的开始、结束以及处理结果的显示等操作。DSP用于提取旋转基座2.1的标号标记被试者，标定头部姿态以及人眼数据的相关信息。DSP还用于利用深度摄像机1.12采集到的深度图、通过AAM方法提取脸部区域和眼部区域、利用迭代最近点方法得到头部姿态、利用眼部区域标定出来的六个点求出瞳孔中心以及利用随机森林对头部姿态和瞳孔中心做分类，求出视线方向。并且，通过支架和眼动仪位置的角度关系，追踪眼睛注视点，得到运动轨迹。可以为后续的人机交互、心理研究等方向提供信息支撑。

[0181] 可旋转的基座2.1固定在支架2所在的平面上，距离可自行设定。当被试者为1人时，底座距离眼动仪的预设距离在50～100cm。当被试者为多人时，底座距离眼动仪的预设距离应该在100cm～200cm。有利用提高测试效果。

[0182] 旋转基座2.1上标有标号，方便多人使用时进行人员的标定。

[0183] 支撑杆2.2连接脸托2.3和基座2.1。支撑杆可进行手动纵向伸缩，调整脸托的位置，使观测者可以舒适的使用该装置。

[0184] 脸托2.3用于支撑被试者的头部。中央向外突出形成弧形带状。材质为橡胶等软性材料，可延竖直方向运动，用来调节被试者竖直方向的头部的转向。观测者的头部想要左右运动时，可以通过旋转基座的转动调节头部的转向。脸托的作用还可以保证观测者在长时间的测试过程中能够充分的放松，与此同时确保观测者的头部与眼动仪1保持预设距离。此外，通过控制脸托2.3的上下转动角度，旋转底座的左右转动角度，可以调控被试者使用该装置时的头部转动范围，使该装置采集的数据更有可用性。

[0185] 本实施例提供了一种非头戴式的多人眼动仪设备，可以解决现有技术中被试者在使用眼动仪时必须头戴或是将头部固定的问题。同时因为头部不需要固定，因此可以同时进行多人眼部信息数据的采集。具体地，观测者在观测时可以带上与其使用的脸托2.3的序号的标答。当设备获得深度图像之后，可以先利用OCR(Optical Character Recognition，光学字符识别)方法提取图片中的数字。提取的数字即为旋转底座2.1标记的序号。然后利用深度图像的深度进行头部检测，粗略地找到头部的位置。计算和该头部位置最近的数字，将此头部位置以此数字n进行标记，即可得到多人的带标记数据，处理时依次进行处理并保存相关结果。

[0186] 参见图10，本发明实施例提供一种检测视线落点的装置，包括：

[0187] 脸部图像获取模块100，用于获取观测者观看显示屏时的脸部图像；

[0188] 特征点提取模块200，用于从所述脸部图像中提取脸部区域和眼部区域的特征点；

[0189] 头部姿态获取模块300，用于跟踪所述脸部图像的初始图片序列，对所述脸部区域的特征点进行迭代计算，获得所述观测者的头部姿态；

[0190] 视线确定模块400，用于根据所述眼部区域的特征点，确定所述观测者的视线角度和视线置信参数；以及

[0191] 视线落点确定模块500，用于根据所述头部姿态、所述视线角度和所述视线置信参数以及所述观测者到所述显示屏的距离，确定所述观测者的视线在所述显示屏的落点位置。

[0192] 在一种可能的实现方式中，所述脸部图像包括深度图像和彩色图像；以及，所述特征点提取模块包括：

[0193] 前景区域提取单元，用于从所述深度图像中提取前景区域；

[0194] 人体脸部判断单元，用于判断所述前景区域是否包括人体脸部；

[0195] 脸部位置定位单元，用于当所述前景区域包括人体脸部时，在所述深度图像中定位所述人体脸部的位置；以及

[0196] 脸部眼部特征提取单元，用于从所述人体脸部在所述彩色图像中的位置提取脸部区域和眼部区域的特征点。

[0197] 在一种可能的实现方式中，所述头部姿态获取模块包括：

[0198] 粒子滤波单元，用于跟踪所述脸部图像的初始图片序列，获得头部姿态的粒子滤波估计值；其中，所述粒子滤波估计值用于估计所述头部姿态；

[0199] 当前视线区域确定单元，用于根据所述粒子滤波估计值，从依据所述观测者的观测范围划分的各视线区域中确定所述头部姿态所朝向的视线区域；以及

[0200] 头部姿态迭代单元，用于基于所述头部姿态所朝向的视线区域和所述粒子滤波估计值，利用最近点迭代算法对所述脸部区域的特征点进行迭代计算，获得所述头部姿态。

[0201] 在一种可能的实现方式中，所述视线落点确定模块包括：

[0202] 注视方向确定单元，用于根据所述头部姿态、所述视线角度和所述视线置信参数，计算所述观测者的眼部视线的注视方向；

[0203] 距离获取单元，用于从所述脸部图像的深度图像中，获取所述眼部区域的瞳孔中心位置的像素值，以确定所述观测者的瞳孔中心到所述显示屏的距离；以及

[0204] 落点计算单元，用于根据所述注视方向和所述观测者的瞳孔中心到所述显示屏的距离，计算所述观测者的眼部视线在所述显示屏的落点位置。

[0205] 所述装置的功能可以通过硬件实现，也可以通过硬件执行相应的软件实现。所述硬件或软件包括一个或多个与上述功能相对应的模块。

[0206] 在一个可能的设计中，检测视线落点的结构中包括处理器和存储器，所述存储器用于检测视线落点的装置执行上述第一方面中检测视线落点的程序，所述处理器被配置为用于执行所述存储器中存储的程序。所述检测视线落点的装置还可以包括通信接口，用于检测视线落点的装置与其他设备或通信网络通信。

[0207] 本发明实施例还提供一种检测视线落点的终端设备，如图11所示，该设备包括：存储器21和处理器22，存储器21内存储有可在处理器22上的计算机程序。处理器22执行计算机程序时实现上述实施例中的检测视线落点的方法。存储器21和处理器22的数量可以为一个或多个。

[0208] 该设备还包括：

[0209] 通信接口23，用于处理器22与外部设备之间的通信。

[0210] 存储器21可能包含高速RAM存储器，也可能还包括非易失性存储器(non-volatile memory)，例如至少一个磁盘存储器。

[0211] 如果存储器21、处理器22和通信接口23独立实现，则存储器21、处理器22和通信接口23可以通过总线相互连接并完成相互间的通信。总线可以是工业标准体系结构(ISA，Industry Standard Architecture)总线、外部设备互连(PCI，Peripheral Component)总线或扩展工业标准体系结构(EISA，Extended Industry Standard Component)总线等。总线可以分为地址总线、数据总线、控制总线等。为便于表示，图11中仅用一条粗线表示，但并不表示仅有一根总线或一种类型的总线。

[0212] 可选的，在具体实现上，如果存储器21、处理器22及通信接口23集成在一块芯片上，则存储器21、处理器22及通信接口23可以通过内部接口完成相互间的通信。

[0213] 在本说明书的描述中，参考术语“一个实施例”、“一些实施例”、“示例”、“具体示例”、或“一些示例”等的描述意指结合该实施例或示例描述的具体特征、结构、材料或者特点包含于本发明的至少一个实施例或示例中。而且，描述的具体特征、结构、材料或者特点可以在任一个或多个实施例或示例中以合适的方式结合。此外，在不相互矛盾的情况下，本领域的技术人员可以将本说明书中描述的不同实施例或示例以及不同实施例或示例的特征进行结合和组合。

[0214] 此外，术语“第一”、“第二”仅用于描述目的，而不能理解为指示或暗示相对重要性或者隐含指明所指示的技术特征的数量。由此，限定有“第一”、“第二”的特征可以明示或隐含地包括至少一个该特征。在本发明的描述中，“多个”的含义是两个或两个以上，除非另有明确具体的限定。

[0215] 流程图中或在此以其他方式描述的任何过程或方法描述可以被理解为，表示包括一个或更多个用于实现特定逻辑功能或过程的步骤的可执行指令的代码的模块、片段或部分，并且本发明的优选实施方式的范围包括另外的实现，其中可以不按所示出或讨论的顺序，包括根据所涉及的功能按基本同时的方式或按相反的顺序，来执行功能，这应被本发明的实施例所属技术领域的技术人员所理解。

[0216] 在流程图中表示或在此以其他方式描述的逻辑和/或步骤，例如，可以被认为是用于实现逻辑功能的可执行指令的定序列表，可以具体实现在任何计算机可读介质中，以供指令执行系统、装置或设备(如基于计算机的系统、包括处理器的系统或其他可以从指令执行系统、装置或设备取指令并执行指令的系统)使用，或结合这些指令执行系统、装置或设备而使用。就本说明书而言，“计算机可读介质”可以是任何可以包含、存储、通信、传播或传输程序以供指令执行系统、装置或设备或结合这些指令执行系统、装置或设备而使用的装置。

[0217] 本发明实施例的计算机可读介质可以是计算机可读信号介质或者计算机可读存储介质或者是上述两者的任意组合。计算机可读存储介质的更具体的示例至少(非穷尽性列表)包括以下：具有一个或多个布线的电连接部(电子装置)，便携式计算机盘盒(磁装置)，随机存取存储器(RAM)，只读存储器(ROM)，可擦除可编辑只读存储器(EPROM或闪速存储器)，光纤装置，以及便携式只读存储器(CDROM)。另外，计算机可读存储介质甚至可以是可在其上打印程序的纸或其他合适的介质，因为可以例如通过对纸或其他介质进行光学扫描，接着进行编辑、解译或必要时以其他合适方式进行处理来以电子方式获得程序，然后将其存储在计算机存储器中。

[0218] 在本发明实施例中，计算机可读信号介质可以包括在基带中或者作为载波一部分传播的数据信号，其中承载了计算机可读的程序代码。这种传播的数据信号可以采用多种形式，包括但不限于电磁信号、光信号或上述的任意合适的组合。计算机可读的信号介质还可以是计算机可读存储介质以外的任何计算机可读介质，该计算机可读介质可以发送、传播或者传输用于指令执行系统、输入法或者器件使用或者与其结合使用的程序。计算机可读介质上包含的程序代码可以用任何适当的介质传输，包括但不限于：无线、电线、光缆、射频(Radio Frequency，RF)等等，或者上述的任意合适的组合。

[0219] 应当理解，本发明的各部分可以用硬件、软件、固件或它们的组合来实现。在上述实施方式中，多个步骤或方法可以用存储在存储器中且由合适的指令执行系统执行的软件或固件来实现。例如，如果用硬件来实现，和在另一实施方式中一样，可用本领域公知的下列技术中的任一项或他们的组合来实现：具有用于对数据信号实现逻辑功能的逻辑门电路的离散逻辑电路，具有合适的组合逻辑门电路的专用集成电路，可编程门阵列(PGA)，现场可编程门阵列(FPGA)等。

[0220] 本技术领域的普通技术人员可以理解实现上述实施例方法携带的全部或部分步骤是可以通过程序来指令相关的硬件完成的程序，该程序可以存储于一种计算机可读存储介质中，该程序在执行时，包括方法实施例的步骤之一或其组合。

[0221] 此外，在本发明各个实施例中的各功能单元可以集成在一个处理模块中，也可以是各个单元单独物理存在，也可以两个或两个以上单元集成在一个模块中。上述集成的模块既可以采用硬件的形式实现，也可以采用软件功能模块的形式实现。集成的模块如果以

软件功能模块的形式实现并作为独立的产品销售或使用时，也可以存储在一个计算机可读存储介质中。存储介质可以是只读存储器，磁盘或光盘等。

[0222] 以上，仅为本发明的具体实施方式，但本发明的保护范围并不局限于此，任何熟悉本技术领域的技术人员在本发明揭露的技术范围内，可轻易想到其各种变化或替换，这些都应涵盖在本发明的保护范围之内。因此，本发明的保护范围应以权利要求的保护范围为准。

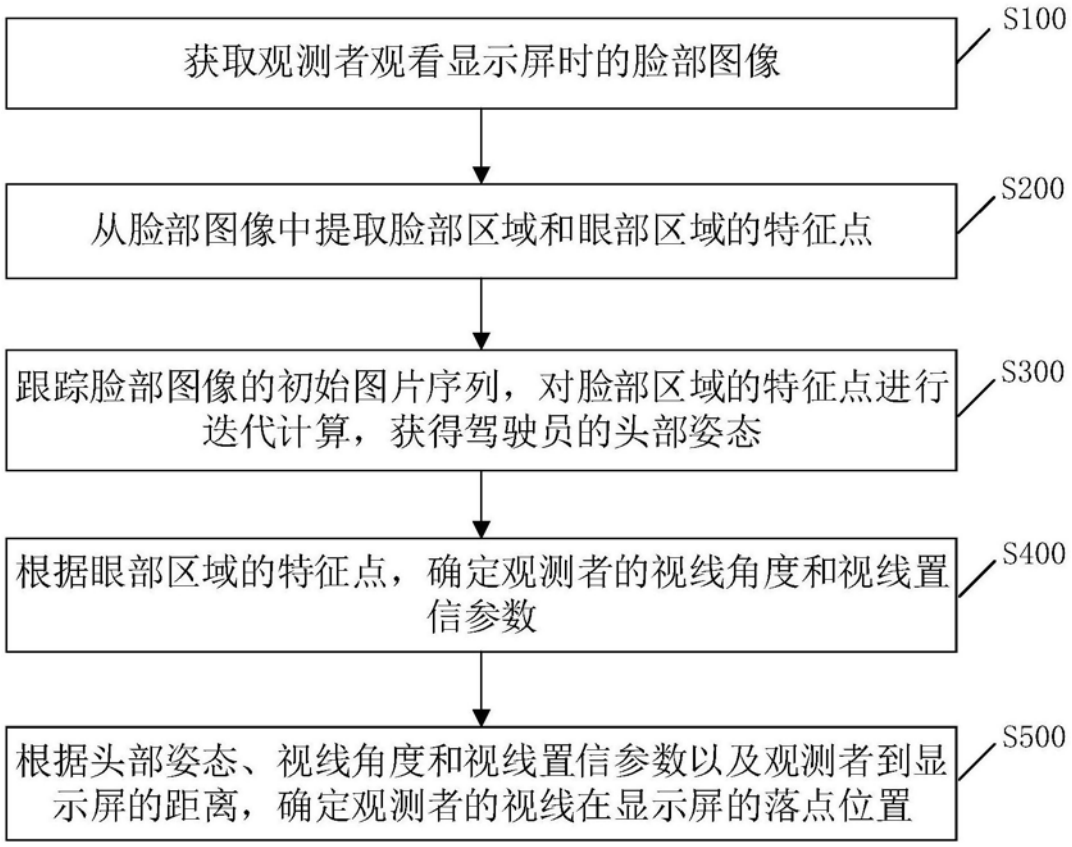

获取观测者观看显示屏时的脸部图像
S100
从脸部图像中提取脸部区域和眼部区域的特征点
S200
跟踪脸部图像的初始图片序列，对脸部区域的特征点进行迭代计算，获得驾驶员的头部姿态
S300
根据眼部区域的特征点，确定观测者的视线角度和视线置信参数
S400
根据头部姿态、视线角度和视线置信参数以及观测者到显示屏的距离，确定观测者的视线在显示屏的落点位置
S500

图1

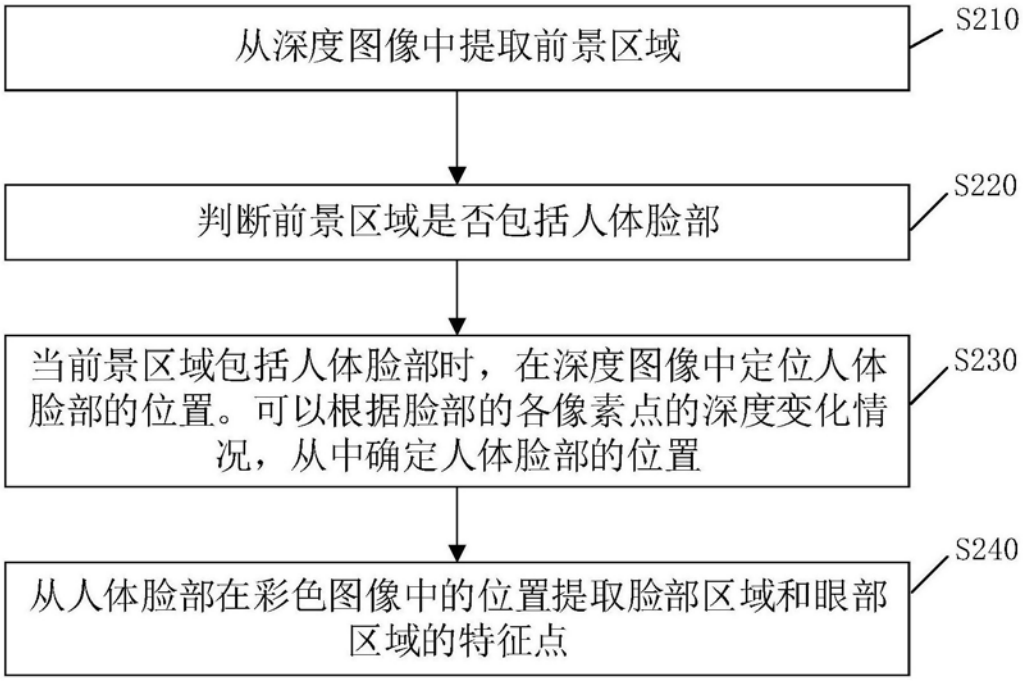

从深度图像中提取前景区域
S210
判断前景区域是否包括人体脸部
S220
当前景区域包括人体脸部时，在深度图像中定位人体脸部的位置。可以根据脸部的各像素点的深度变化情况，从中确定人体脸部的位置
S230
从人体脸部在彩色图像中的位置提取脸部区域和眼部区域的特征点
S240

图2

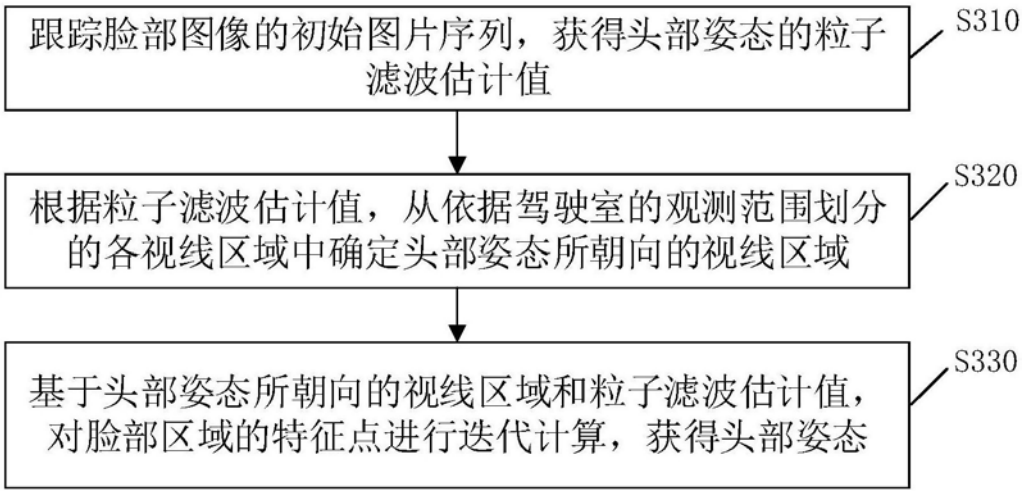

跟踪脸部图像的初始图片序列，获得头部姿态的粒子滤波估计值
S310
根据粒子滤波估计值，从依据驾驶室的观测范围划分的各视线区域中确定头部姿态所朝向的视线区域
S320
基于头部姿态所朝向的视线区域和粒子滤波估计值，对脸部区域的特征点进行迭代计算，获得头部姿态
S330

图3

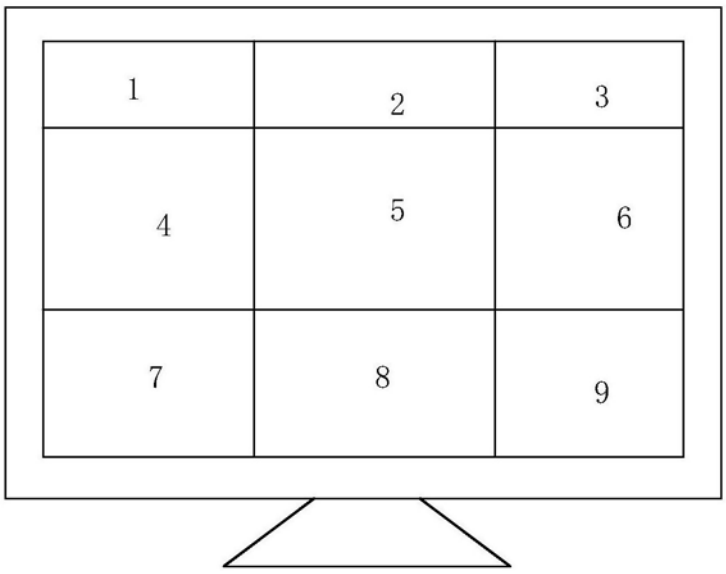

1
2
3
4
5
6
7
8
9

图4

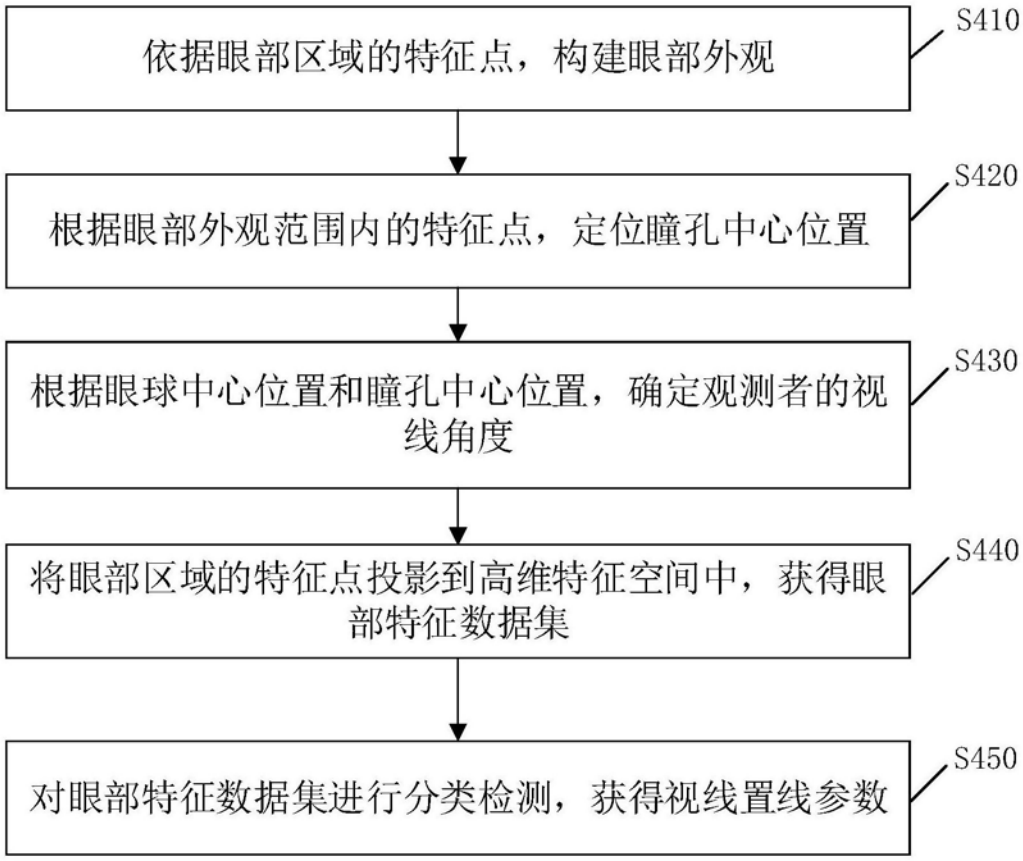

依据眼部区域的特征点，构建眼部外观
S410
根据眼部外观范围内的特征点，定位瞳孔中心位置
S420
根据眼球中心位置和瞳孔中心位置，确定观测者的视线角度
S430
将眼部区域的特征点投影到高维特征空间中，获得眼部特征数据集
S440
对眼部特征数据集进行分类检测，获得视线置线参数
S450

图5

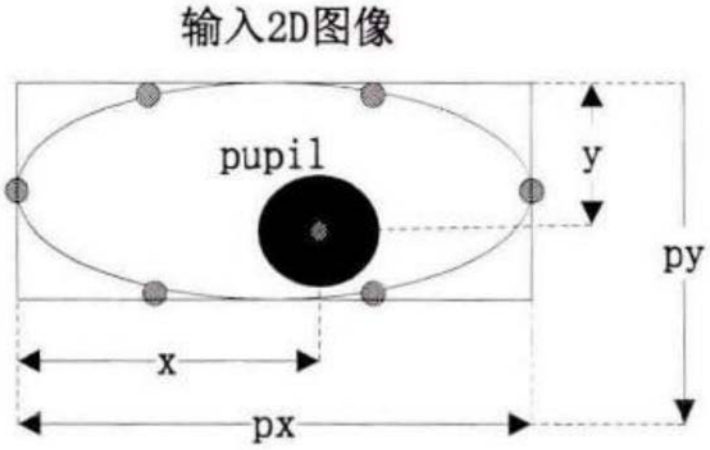

输入2D图像
pupil
y
py
x
px

图6


3D眼球模型
眼球中心
视线角度

图7

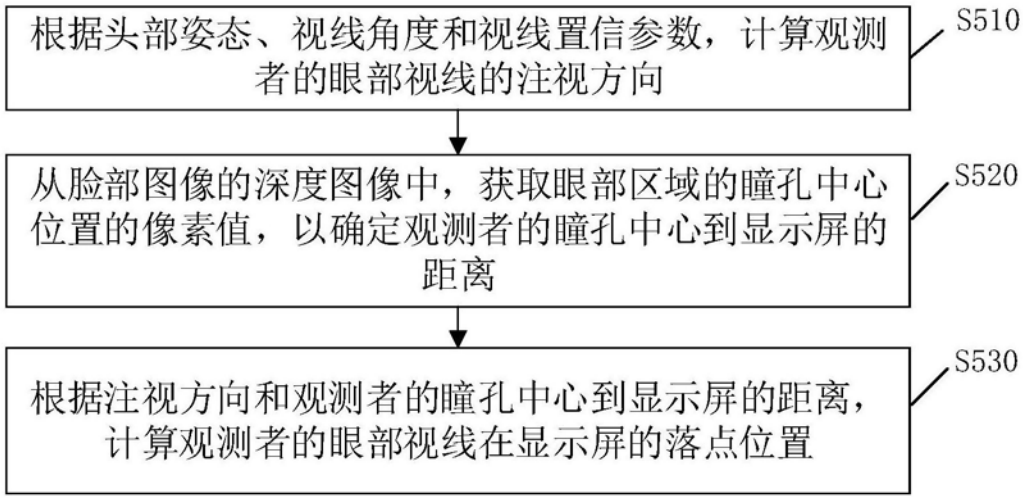

根据头部姿态、视线角度和视线置信参数，计算观测者的眼部视线的注视方向
S510
从脸部图像的深度图像中，获取眼部区域的瞳孔中心位置的像素值，以确定观测者的瞳孔中心到显示屏的距离
S520
根据注视方向和观测者的瞳孔中心到显示屏的距离，计算观测者的眼部视线在显示屏的落点位置
S530

图8

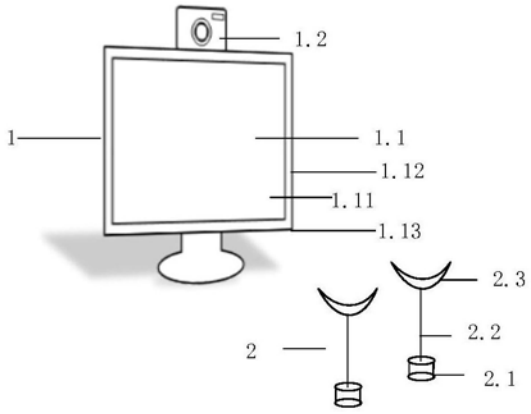

1.2
1
1.1
1.12
1.11
1.13
2.3
2.2
2
2.1

图9

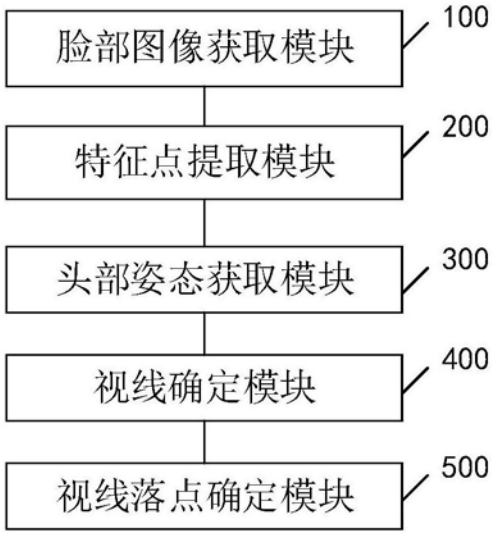

脸部图像获取模块
100
特征点提取模块
200
头部姿态获取模块
300
视线确定模块
400
视线落点确定模块
500

图10

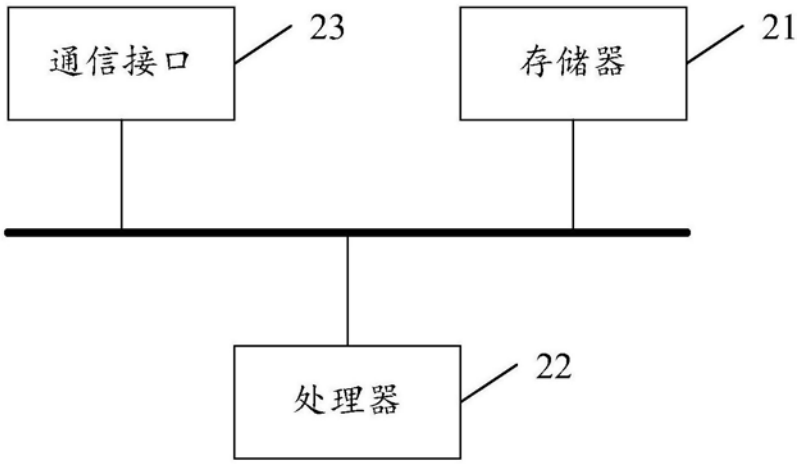

通信接口
23
存储器
21
处理器
22

图11