



(12) 发明专利

(10) 授权公告号 CN 109409398 B

(45) 授权公告日 2022.04.26

(21) 申请号 201810927431.0

G06V 30/194 (2022.01)

(22) 申请日 2018.08.15

G06K 9/62 (2022.01)

(65) 同一申请的已公布的文献号

申请公布号 CN 109409398 A

(56) 对比文件

JP 2011145791 A, 2011.07.28

(43) 申请公布日 2019.03.01

JP 2011060227 A, 2011.03.24

(30) 优先权数据

JP 2013052490 A, 2013.03.21

2017-157618 2017.08.17 JP

JP 2011060221 A, 2011.03.24

(73) 专利权人 佳能株式会社

CN 103440318 A, 2013.12.11

地址 日本国东京都大田区下丸子3丁目30-
2

CN 105023020 A, 2015.11.04

(72) 发明人 稲叶正树

CN 105139390 A, 2015.12.09

(74) 专利代理机构 北京怡丰知识产权代理有限公司 11293

JP 2014022837 A, 2014.02.03

代理人 迟军 李艳丽

US 9076065 B1, 2015.07.07

(51) Int.Cl.

林义闽.未知环境中智能机器人的视觉导航
技术研究.《中国优秀博硕士学位论文全文数据
库(博士)信息科技辑》.2015, (第04期),

G06V 10/762 (2022.01)

审查员 陈琪蒙

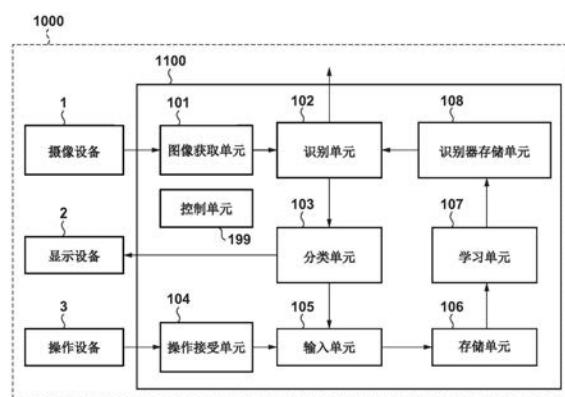
权利要求书3页 说明书14页 附图12页

(54) 发明名称

图像处理装置、图像处理方法以及存储介质

(57) 摘要

本发明提供图像处理装置、图像处理方法以及存储介质。获取各自被识别器识别为输入图像上的目标物体的位置的多个识别位置。通过对所述多个识别位置进行聚类来获得至少一个代表性位置。根据来自用户的针对代表性位置的编辑指令，对代表性位置进行编辑。保存所述输入图像和代表性位置，作为要用于识别器的学习的学习数据。



1. 一种图像处理装置,所述图像处理装置包括:

获取单元,被构造为获取各自被识别器识别为输入图像上的目标物体的位置的多个识别位置;

获得单元,被构造为通过对所述多个识别位置进行聚类来获得至少一个代表性位置;

编辑单元,被构造为根据来自用户的针对代表性位置的编辑指令,对代表性位置进行编辑;以及

保存单元,被构造为保存所述输入图像和代表性位置,作为要用于识别器的学习的学习数据。

2. 根据权利要求1所述的图像处理装置,其中,所述获取单元获取所述输入图像上的像素位置中的、各自被识别器识别为所述目标物体的位置的多个像素位置,作为所述多个识别位置。

3. 根据权利要求1所述的图像处理装置,其中,在来自如下识别器的输出表示在包括所述输入图像上的像素位置的图像区域中包括所述目标物体的情况下,所述获取单元获取所述像素位置作为识别位置,该识别器被输入所述图像区域中的图像。

4. 根据权利要求1所述的图像处理装置,其中,所述获得单元通过对所述多个识别位置进行聚类来将所述多个识别位置分类成预定数量的类别,并且基于属于各类别的识别位置来获得与该类别对应的代表性位置。

5. 根据权利要求4所述的图像处理装置,其中,所述获得单元获得属于类别的识别位置的平均位置、属于类别的识别位置的中值、或通过将属于类别的识别位置应用于高斯分布而形成的峰值位置,作为与类别对应的代表性位置。

6. 根据权利要求4所述的图像处理装置,所述图像处理装置还包括:

运算单元,被构造为获得类别的可靠度,

其中,在可靠度小于第一阈值的类别的数量变得不小于预定数量的情况下或者在可靠度小于第一阈值的类别的数量与类别的总数之比变得不小于预定数量的情况下,所述编辑单元进行操作。

7. 根据权利要求6所述的图像处理装置,其中,所述运算单元根据属于类别的识别位置的数量,来获得类别的可靠度。

8. 根据权利要求6所述的图像处理装置,其中,所述运算单元根据与属于类别的识别位置对应的识别器的输出,来获得类别的可靠度。

9. 根据权利要求6所述的图像处理装置,其中,所述运算单元根据属于类别的识别位置和与类别对应的代表性位置之间的距离,来获得类别的可靠度。

10. 根据权利要求1所述的图像处理装置,其中,所述编辑单元显示被叠加在所述输入图像上的表示代表性位置的信息,并且接受来自用户的针对代表性位置的编辑操作。

11. 根据权利要求6所述的图像处理装置,其中,所述编辑单元显示被叠加在所述输入图像上的表示可靠度不小于比所述第一阈值小的第二阈值的类别的代表性位置的信息,并且接受来自用户的针对代表性位置的编辑操作。

12. 根据权利要求1所述的图像处理装置,其中,所述编辑单元根据来自用户的针对代表性位置的移动指令来移动代表性位置,根据来自用户的针对代表性位置的删除指令来删除代表性位置,以及根据来自用户的添加新的代表性位置的指令来添加代表性位置。

13. 根据权利要求1所述的图像处理装置,所述图像处理装置还包括:

学习单元,被构造为使用由所述保存单元保存的学习数据来进行识别器的学习。

14. 根据权利要求13所述的图像处理装置,所述图像处理装置还包括:

更新单元,被构造为将识别器更新为经历了学习的识别器。

15. 根据权利要求14所述的图像处理装置,其中,在针对学习前的识别器的评价值低于针对学习后的识别器的评价值的情况下,所述更新单元将学习前的识别器更新为学习后的识别器。

16. 一种图像处理装置,所述图像处理装置包括:

获取单元,被构造为获取各自被识别器识别为输入图像上的目标物体的位置和姿态的多个识别位置和姿态;

获得单元,被构造为通过对所述多个识别位置和姿态进行聚类来获得至少一个代表性位置和姿态;

编辑单元,被构造为根据来自用户的针对代表性位置和姿态的编辑指令,对代表性位置和姿态进行编辑;以及

保存单元,被构造为保存所述输入图像和代表性位置和姿态,作为要用于识别器的学习的数据。

17. 一种图像处理方法,所述图像处理方法包括:

获取各自被识别器识别为输入图像上的目标物体的位置的多个识别位置;

通过对所述多个识别位置进行聚类来获得至少一个代表性位置;

根据来自用户的针对代表性位置的编辑指令,对代表性位置进行编辑;以及

保存所述输入图像和代表性位置,作为要用于识别器的学习的数据。

18. 一种图像处理方法,所述图像处理方法包括:

获取各自被识别器识别为输入图像上的目标物体的位置和姿态的多个识别位置和姿态;

通过对所述多个识别位置和姿态进行聚类来获得至少一个代表性位置和姿态;

根据来自用户的针对代表性位置和姿态的编辑指令,对代表性位置和姿态进行编辑;以及

保存所述输入图像和代表性位置和姿态,作为要用于识别器的学习的数据。

19. 一种存储计算机程序的非暂时性计算机可读存储介质,所述计算机程序被构造为使计算机用作:

获取单元,被构造为获取各自被识别器识别为输入图像上的目标物体的位置的多个识别位置;

获得单元,被构造为通过对所述多个识别位置进行聚类来获得至少一个代表性位置;

编辑单元,被构造为根据来自用户的针对代表性位置的编辑指令,对代表性位置进行编辑;以及

保存单元,被构造为保存所述输入图像和代表性位置,作为要用于识别器的学习的数据。

20. 一种存储计算机程序的非暂时性计算机可读存储介质,所述计算机程序被构造为使计算机用作:

获取单元,被构造为获取各自被识别器识别为输入图像上的目标物体的位置和姿态的多个识别位置和姿态;

获得单元,被构造为通过对所述多个识别位置和姿态进行聚类来获得至少一个代表性位置和姿态;

编辑单元,被构造为根据来自用户的针对代表性位置和姿态的编辑指令,对代表性位置和姿态进行编辑;以及

保存单元,被构造为保存所述输入图像和代表性位置和姿态,作为要用于识别器的学习的数据。

图像处理装置、图像处理方法以及存储介质

技术领域

[0001] 本发明涉及图像处理装置、图像处理方法以及非暂时性计算机可读存储介质，特别涉及关于图像识别的技术。

背景技术

[0002] 传统上，存在一种从通过拍摄目标物体而获得的彩色图像或深度图像中学习特征或图案并且识别目标物体的方法。在这样的方法中，为了学习特征或图案，人们需要指定在彩色图像或深度图像中包括哪个目标物体，或者指定目标物体的位置或位置和姿态(在下文中将被称为“标注”)。为了学习，需要准备大量的这样的数据。然而，人们难以正确地进行整个标注。

[0003] 为了解决这个问题，如在日本特许第5953151号公报中，存在一种通过反复“对当前的识别器的识别结果的校正”和“使用校正的数据对识别器的学习/更新”来提高识别器的精度的方法。

[0004] 然而，如果识别器识别目标物体的位置或位置和姿态(orientation)，则图像中的大量的像素被作为识别结果输出。因此，由人们对各个识别器结果进行校正的成本较高。

发明内容

[0005] 考虑到上述的问题而做出了本发明，并且本发明提供一种减少用户指定图像中的目标物体的位置的负担的技术。

[0006] 根据本发明的第一方面，提供了一种图像处理装置，所述图像处理装置包括：获取单元，被构造为获取各自被识别器识别为输入图像上的目标物体的位置的多个识别位置；获得单元，被构造为通过对所述多个识别位置进行聚类来获得至少一个代表性位置；编辑单元，被构造为根据来自用户的针对代表性位置的编辑指令，对代表性位置进行编辑；以及保存单元，被构造为保存所述输入图像和代表性位置，作为要用于识别器的学习的数据。

[0007] 根据本发明的第二方面，提供了一种图像处理装置，所述图像处理装置包括：获取单元，被构造为获取各自被识别器识别为输入图像上的目标物体的位置和姿态的多个识别位置和姿态；获得单元，被构造为通过对所述多个识别位置和姿态进行聚类来获得至少一个代表性位置和姿态；编辑单元，被构造为根据来自用户的针对代表性位置和姿态的编辑指令，对代表性位置和姿态进行编辑；以及保存单元，被构造为保存所述输入图像和代表性位置和姿态，作为要用于识别器的学习的数据。

[0008] 根据本发明的第三方面，提供了一种图像处理方法，所述图像处理方法包括：获取各自被识别器识别为输入图像上的目标物体的位置的多个识别位置；通过对所述多个识别位置进行聚类来获得至少一个代表性位置；根据来自用户的针对代表性位置的编辑指令，对代表性位置进行编辑；以及保存所述输入图像和代表性位置，作为要用于识别器的学习的数据。

[0009] 根据本发明的第四方面，提供了一种图像处理方法，所述图像处理方法包括：获取各自被识别器识别为输入图像上的目标物体的位置和姿态的多个识别位置和姿态；通过对所述多个识别位置和姿态进行聚类来获得至少一个代表性位置和姿态；根据来自用户的针对代表性位置和姿态的编辑指令，对代表性位置和姿态进行编辑；以及保存所述输入图像和代表性位置和姿态，作为要用于识别器的学习的数据。

[0010] 根据本发明的第五方面，提供了一种存储计算机程序的非暂时性计算机可读存储介质，所述计算机程序被构造为使计算机用作：获取单元，被构造为获取各自被识别器识别为输入图像上的目标物体的位置的多个识别位置；获得单元，被构造为通过对所述多个识别位置进行聚类来获得至少一个代表性位置；编辑单元，被构造为根据来自用户的针对代表性位置的编辑指令，对代表性位置进行编辑；以及保存单元，被构造为保存所述输入图像和代表性位置，作为要用于识别器的学习的数据。

[0011] 根据本发明的第六方面，提供了一种存储计算机程序的非暂时性计算机可读存储介质，所述计算机程序被构造为使计算机用作：获取单元，被构造为获取各自被识别器识别为输入图像上的目标物体的位置和姿态的多个识别位置和姿态；获得单元，被构造为通过对所述多个识别位置和姿态进行聚类来获得至少一个代表性位置和姿态；编辑单元，被构造为根据来自用户的针对代表性位置和姿态的编辑指令，对代表性位置和姿态进行编辑；以及保存单元，被构造为保存所述输入图像和代表性位置和姿态，作为要用于识别器的学习的数据。

[0012] 通过以下(参照附图)对示例性实施例的描述，本发明的其他特征将变得清楚。

附图说明

- [0013] 图1是示出物体识别装置的功能布置的示例的框图；
- [0014] 图2A和图2B是物体识别处理的流程图；
- [0015] 图3A至图3D是用于说明识别单元102和分类单元103的处理的图；
- [0016] 图4是示出画面的显示示例的图；
- [0017] 图5A至图5D是用于说明由用户对标签的编辑的图；
- [0018] 图6是示出物体识别装置的功能布置的示例的框图；
- [0019] 图7是示出步骤S2和S3的处理的详情的流程图；
- [0020] 图8是示出物体识别装置的功能布置的示例的框图；
- [0021] 图9是示出步骤S2和S3的处理的详情的流程图；
- [0022] 图10A至图10C是用于说明在不仅识别位置还识别姿态的情况下执行的处理的图；以及
- [0023] 图11是示出计算机装置的硬件布置的示例的框图。

具体实施方式

[0024] 现在将参照附图描述本发明的实施例。注意，下面要描述的实施例是本发明的详细实施方式的示例或所附权利要求中描述的布置的详细示例。

[0025] [第一实施例]

[0026] 根据该实施例的图像处理装置具有如下的标注功能：针对输入图像对使用识别器

识别的目标物体的识别位置进行聚类，并且将通过聚类获得的识别位置的代表性位置设置为标签作为要由用户校正的目标。该功能消除了由用户对全部的识别位置进行确认和校正的必要性。由于仅需要基于代表性位置来对标签进行确认和校正，因此减少了用于标注的劳动。这里，标签是指关于图像中的目标物体的位置或位置和姿态的信息。标签被用作用于构建识别器的指导数据。

[0027] 首先，将参照图1的框图来描述根据该实施例的物体识别装置的功能布置的示例。如图1中所示，根据该实施例的物体识别装置1000包括摄像设备1、显示设备2、操作设备3和图像处理装置1100。

[0028] 首先，将描述摄像设备1。摄像设备1是拍摄静止图像或运动图像的设备。拍摄图像不限于特定的图像，并且可以是彩色图像（例如，RGB图像）、黑白图像、红外图像或作为深度信息的深度图像。由摄像设备1拍摄的图像被发送到图像处理装置1100。

[0029] 接下来将描述显示设备2。显示设备2由CRT、液晶屏等形成，并且能够通过图像、字符等来显示图像处理装置1100的处理结果。注意，显示设备2可以是触摸面板屏。在这种情况下，来自用户的对触摸面板屏的各种操作的内容被发送到图像处理装置1100作为操作信息。

[0030] 接下来将描述操作设备3。操作设备3由诸如鼠标或键盘的用户接口形成。当被用户操作时，操作设备3能够向图像处理装置1100输入各种指令。下面将假设操作设备3是鼠标来进行描述。然而，如果用户接口是能够向图像处理装置1100输入与下面要描述的指令相同的指令的用户接口，则任何用户接口都能够应用于操作设备3。

[0031] 接下来将描述图像处理装置1100。图像获取单元101获取从摄像设备1发送的拍摄图像。注意，拍摄图像获取源不限于摄像设备1。可以经由诸如LAN或互联网的网络从诸如数据库、服务器装置、平板终端设备或智能电话的设备获取拍摄图像。拍摄图像可以是由摄像设备1拍摄的图像，可以是在向公众开放的图像数据库中登记的图像，或者可以是上载到web的图像。

[0032] 识别单元102使用保存在识别器存储单元108中的CNN（Convolutional Neural Network，卷积神经网络）（识别器）的模型结构和CNN的参数来构建CNN。识别单元102将所构建的CNN应用于由图像获取单元101获取的拍摄图像，从而识别拍摄图像中的目标物体的位置作为识别位置。

[0033] 分类单元103对由识别单元102识别的识别位置进行聚类，从而将识别位置分类成一个或更多个类别。然后，分类单元103基于属于类别的识别位置获得（计算）与该类别对应的代表性位置（标签）。

[0034] 操作接受单元104向输入单元105发送与用户通过对操作设备3进行操作而输入的操作内容对应的指令。输入单元105基于从操作接受单元104接收的指令，对上述的代表性位置进行编辑（标注）。然后，输入单元105将拍摄图像和代表性位置存储在存储单元106中，作为用于CNN的学习的数据。

[0035] 学习单元107使用存储在存储单元106中的学习数据来创建指导数据。然后，学习单元107使用保存在识别器存储单元108中的CNN的模型结构和CNN的参数来构建CNN，并且使用上面创建的指导数据进行所构建的CNN的参数的学习处理。然后，学习单元107将存储在识别器存储单元108中的CNN的参数更新为通过学习处理获得的参数。

[0036] 识别器存储单元108保持CNN的模型结构和参数。在该实施例中,CNN的模型结构被设计为如下的模型:接收拍摄图像,并且经由卷积层、全连接层和Softmax层输出表示在拍摄图像的各个图像区域中是否包括目标物体的2级辨别。这里,CNN的参数表示卷积层和全连接层中的权重。

[0037] 控制单元199控制包括图像获取单元101、识别单元102、分类单元103、学习单元107、操作接受单元104和输入单元105的功能单元的操作,并且还控制整个图像处理装置1100的操作。

[0038] 注意,图1将摄像设备1、显示设备2、操作设备3和图像处理装置1100示出为分开的设备。然而,这些设备中的两个或更多个可以一体化为一个设备。

[0039] 接下来将参照图2A的流程图来描述由图像处理装置1100进行的物体识别处理。在步骤S1中,控制单元199确定CNN(识别器)的学习是否是必要的。确定CNN的学习是否是必要的条件不限于特定的条件。例如,在检测到用户对操作设备3进行操作并且输入CNN学习指令时,控制单元199确定CNN的学习是必要的。另外,例如,在每个预定定时或每当满足预定条件时,控制单元199确定CNN的学习是必要的。

[0040] 如果作为确定的结果确定CNN的学习是必要的,则处理前进到步骤S2。另一方面,如果确定CNN的学习是不必要的,则处理前进到步骤S4。

[0041] 在步骤S2中,以上述的方式生成学习数据并且将学习数据存储在存储单元106中。在步骤S3中,使用学习数据来创建指导数据,使用指导数据进行CNN的参数的学习,并且将存储在识别器存储单元108中的CNN的参数更新为学习后的参数。

[0042] 将参照图2B的流程图来描述步骤S2和S3的处理的详情。在图2B中,在步骤S2中执行步骤S11至S15,并且在步骤S3中执行步骤S16和S17。

[0043] 在步骤S11中,图像获取单元101获取从摄像设备1发送的拍摄图像。在步骤S12中,首先,识别单元102使用保存在识别器存储单元108中的CNN(识别器)的模型结构和CNN的参数来构建CNN。识别单元102将所构建的CNN应用于在步骤S11中由图像获取单元101获取的拍摄图像,从而识别拍摄图像中的目标物体的位置作为识别位置。

[0044] 例如,如果图像获取单元101获取图3A中所示的图像31作为拍摄图像,则如图3B中所示,识别单元102设置以图像31中的关注像素位置32为中心的图像区域33。图像区域33仅需要包括关注像素位置32,并且不必总是以关注像素位置32为中心的图像区域。识别单元102将图像区域33中的图像输入到所构建的CNN,并且进行CNN的算术处理,从而获得CNN的输出。如果CNN的输出表示“在图像区域33中包括目标物体”,则识别单元102将关注像素位置32识别为图像31中的目标物体的位置(识别位置)。另一方面,如果CNN的输出不表示“在图像区域33中包括目标物体”,则识别单元102不将关注像素位置32识别为图像31中的目标物体的位置(识别位置)。识别单元102在如虚线34所示以光栅扫描顺序移动关注像素位置32的同时,对各个像素位置进行与上述相同的处理,从而针对图像31上的各个像素位置确定在以像素位置为中心的图像区域中是否包括目标物体。因此,能够识别图像31上的目标物体的位置。注意,关注像素位置32的移动顺序不限于光栅扫描顺序,只要能够针对图像31上的各个像素位置确定在以像素位置为中心的图像区域中是否包括目标物体即可。

[0045] 图3C示出了当将图像31中包括的物体中的圆柱形物体设置为目标物体并且将“×”标记(在下文中将被称为识别位置标记)添加到图像31上的目标物体的识别位置时所

获得的结果。由于在邻近的像素中提取相似的图像,因此,例如,如由识别位置标记组35所示,识别位置通常是密集地拥挤的。另外,对于具有与目标物体相似的特征的位置,容易发生识别错误。例如,由于就被边缘(图像中的亮度梯度)围绕而言立方体与圆柱体是共同的,因此,例如,如由识别位置标记组36所示,可能发生识别错误。

[0046] 在步骤S13中,分类单元103将在步骤S12中由识别单元102识别的目标物体的位置(识别位置)的集合定义为初始集合,并且基于预定的聚类条件对初始集合进行聚类,从而将包括在集合中的识别位置分类成一个或更多个类别。然后,针对一个或更多个所分类的类别中的各个,分类单元103基于属于类别的识别位置获得与该类别对应的代表性位置。例如,当基于初始集合的特征使用K均值算法对初始集合进行聚类时,将距离近的识别位置放入一个类别中。针对各个类别,获得属于该类别的识别位置的平均位置作为类别的代表性位置。例如,当图3C中所示的识别位置标记组的位置被聚类时,它们被分类成五个类别。图3D示出了分类结果。在图3D中,将“○”标记(在下文中将被称为代表性位置标记)添加到五个所分类的类别中的各个的位置(类别的代表性位置)。代表性位置标记37位于与识别位置标记组35对应的类别的代表性位置处,并且代表性位置标记38位于与识别位置标记组36对应的类别的代表性位置处。注意,聚类所需的参数(例如,类别的数量和能够定义类别的范围(类别范围))由用户预先指定并且由分类单元103保持。

[0047] 分类单元103将代表性位置标记与在步骤S1中由图像获取单元101获取的拍摄图像上(输入图像上)的各个类别的代表性位置合成,然后在显示设备2的显示画面上显示包括拍摄图像的画面。图4示出了画面的显示示例。在画面41上,在显示区域42中显示在步骤S1中由图像获取单元101获取的拍摄图像。代表性位置标记401至405被布置在拍摄图像上的类别的代表性位置处。

[0048] 滑块43用于设置聚类所需的参数。例如,当用户通过对操作设备3进行操作,而进行将光标44移动到滑块43的位置然后改变滑块43的位置的操作以使滑块43更靠近右端时,能够指定更大数量的类别或更广的类别范围。作为选择,当用户通过对操作设备3进行操作,而进行将光标44移动到滑块43的位置然后改变滑块43的位置的操作以使滑块43更靠近左端时,能够指定更少数量的类别或更窄的类别范围。当前指定的类别的数量或类别范围在区域43a中被显示为数值。注意,用户可以通过对操作设备3进行操作来设置类别的数量或类别范围,并且因此直接指定区域43a中的数值。指定类别的数量或类别范围的方法不限于特定的指定方法。每当聚类所需的参数被指定时,分类单元103使用新指定的参数进行初始集合的聚类,并且在显示区域42中显示如下拍摄图像,在该拍摄图像中,代表性位置标记被布置在通过聚类获得的类别的代表性位置处。分类单元103管理最新的代表性位置的列表(代表性位置列表)。注意,图4中所示的画面仅仅是可应用于该实施例的画面的示例,并且画面不限于此。

[0049] 接下来,在步骤S14a中,操作接受单元104接受关于代表性位置的编辑操作输入(编辑指令)。当用户通过对操作设备3进行操作而进行关于代表性位置的编辑操作输入时,操作接受单元104创建用于编辑的指令并且将其输入到输入单元105。

[0050] 在步骤S14b中,在接收到来自操作接受单元104的指令时,输入单元105根据指令进行关于由分类单元103分类的代表性位置(标签)(即,在上述的代表性位置列表中登记的代表性位置)的编辑处理。例如,用户通过对操作设备3进行操作来进行以下操作,从而实现

诸如标签的删除、添加或校正等的标注。

[0051] 这里将参照图5A至图5D来描述由用户对标签的编辑。由输入单元105进行根据下面将描述的用户操作的全部标签编辑处理。

[0052] 用户通过对操作设备3进行操作而将光标44移动到代表性位置标记401至405中的期望的代表性位置标记的位置并且进行右击操作,从而删除期望的代表性位置标记。当代表性位置标记被删除时,从上述的代表性位置列表中删除与代表性位置标记对应的代表性位置。例如,如图5A中所示,假设代表性位置标记401至405被布置在显示区域42中显示的拍摄图像上。代表性位置标记401至405中的代表性位置标记401和403表示圆柱形物体(即,目标物体)的位置。然而,代表性位置标记402和404不表示圆柱形物体的位置而表示立方体物体的位置,并且不表示目标物体的位置。另外,代表性位置标记405根本不表示物体的位置。在这种情况下,用户通过对操作设备3进行操作而将光标44移动到代表性位置标记402、404和405的位置中的各个并且进行右击操作,从而删除代表性位置标记。结果,代表性位置标记的布局从图5A改变为图5B。在图5B中,删除了图5A中所示的代表性位置标记401至405中的代表性位置标记402、404和405。在这种情况下,从上述的代表性位置列表中删除与代表性位置标记402、404和405对应的代表性位置。结果,仅与代表性位置标记401和403对应的代表性位置被登记在代表性位置列表中。

[0053] 在图5B中,代表性位置标记401和403被布置在三个目标物体中的两个目标物体上。然而,没有代表性位置标记被布置在一个剩余的目标物体406上。用户通过对操作设备3进行操作而将光标44移动到目标物体406的位置并且进行左击操作,从而在该位置处布置新的代表性位置标记410,如图5C中所示。当新的代表性位置标记410被布置时,代表性位置标记410的布置位置被作为代表性位置而登记在上述的代表性位置列表中。

[0054] 另外,用户通过对操作设备3进行操作而将光标44移动到期望的代表性位置标记的位置并且在进行左击操作的同时进行拖放操作,从而移动期望的代表性位置标记的位置。当代表性位置标记的位置被移动时,在上述的代表性位置列表中代表性位置标记的位置被改变为移动后的位置。图5D示出了当用户通过对操作设备3进行操作而移动图5C中的代表性位置标记403的位置时获得的结果。

[0055] 在步骤S14c中,控制单元199确定用户是否通过对操作设备3进行操作而指定了结束按钮45。如果作为确定的结果,用户通过对操作设备3进行操作而指定了结束按钮45,则处理前进到步骤S15。另一方面,如果用户没有通过对操作设备3进行操作而指定结束按钮45,则处理返回到步骤S14a。

[0056] 在步骤S15中,输入单元105将代表性位置列表和在步骤S11中由图像获取单元101获取的拍摄图像存储在存储单元106中作为用于CNN的参数的学习的数据。另外,如果在图4中所示的画面中改变聚类所需的参数,则分类单元103保持改变的参数。

[0057] 在步骤S16中,学习单元107使用存储在存储单元106中的学习数据来创建指导数据。例如,如同识别单元102那样,学习单元107针对学习数据中包括的拍摄图像中的各个像素位置设置图像区域。然后,学习单元107针对拍摄图像中的各个像素位置,生成表示像素位置是否与学习数据中包括的代表性位置列表中登记的代表性位置中的一个相对应的代码数据(正或负(真或假))。这里,“A对应于B”不限于A完全等于B的情况,并且可以包括A与B之间的差不大于预定值的情况。如果学习数据中包括的拍摄图像中的关注像素位置与学习

数据中包括的代表性位置列表中登记的代表性位置中的一个相对应，则学习单元107将代码数据“真”分配给关注像素位置。另一方面，如果学习数据中包括的拍摄图像中的关注像素位置不与学习数据中包括的代表性位置列表中登记的代表性位置中的任何一个相对应，则学习单元107将代码数据“假”分配给关注像素位置。因此，学习单元107生成指导数据，该指导数据针对拍摄图像中的各个像素位置，将针对像素位置设置的图像区域中的图像(训练图像)和分配给该像素位置的代码数据彼此相关联地保持。

[0058] 然后，学习单元107使用保存在识别器存储单元108中的CNN的模型结构和CNN的参数来构建CNN。学习单元107通过通常在深度学习中使用的反向传播来优化CNN的参数，使得通过将各个训练图像输入到所构建的CNN而获得的CNN的输出变为与训练图像对应的代码数据。使用保存在识别器存储单元108中的CNN的参数作为此时的CNN的参数的初始值。

[0059] 在步骤S17中，学习单元107将存储在识别器存储单元108中的CNN的参数更新为通过步骤S16中的学习(优化)而获得的CNN的参数。

[0060] 返回参照图2A，在步骤S4中，图像获取单元101获取从摄像设备1发送的拍摄图像。在步骤S5中，首先，识别单元102使用保存在识别器存储单元108中的CNN的模型结构和CNN的参数来构建CNN(识别器)。然后，识别单元102将所构建的CNN应用于在步骤S4中由图像获取单元101获取的拍摄图像，从而识别拍摄图像中的物体的位置作为识别位置。在步骤S5中识别的物体的识别位置的输出目的地不限于特定的输出目的地。例如，可以在图4中所示的画面的显示区域42中显示在步骤S4中获取的拍摄图像，并且同时，可以将诸如标记的信息布置在步骤S5中识别的识别位置处或者邻近位置处。

[0061] 如上所述，根据该实施例，当对图像进行标注时，用户能够参照由当前的识别器识别的代表性位置来进行标注，因此，减少了用户标注的劳动。

[0062] <第一实施例的变型例>

[0063] 在第一实施例中，使用K均值算法用于步骤S13中的聚类。然而，聚类的算法不限于特定的算法。例如，可以使用诸如均值偏移(Mean Shift)或最近邻(Nearest Neighbor)的其他聚类算法，或者用户可以通过使用鼠标在图像中拖动来指定要放入一个类别的区域。这能够自由地调整要立即处理的识别结果的区域。

[0064] 另外，在第一实施例中，在步骤S13中，使用属于类别的识别位置的平均位置作为与该类别对应的代表性位置。然而，可以使用其他方法来运算代表性位置。例如，可以使用属于类别的识别位置的中值作为该类别的代表性位置。作为选择，可以使用通过将属于类别的识别位置应用于高斯分布而形成的峰值位置作为该类别的代表性位置。这使得即使在属于类别的识别位置处存在诸如异常值的噪声，也能够运算稳定的代表性位置。

[0065] 另外，在第一实施例中，在步骤S13中将表示代表性位置的标记布置在拍摄图像上。然而，要显示的信息不限于这些信息。例如，不仅可以合成并在拍摄图像上显示表示代表性位置的标记，还可以合成并在拍摄图像上显示表示识别位置的标记。作为选择，可以合成并在拍摄图像上显示一些目标物体的位置。这使得用户能够知道由当前的识别器获得了何种识别结果。

[0066] 此外，在第一实施例中，在步骤S14a中通过用户的鼠标操作进行标注。然而，可以通过其他方法进行标注。例如，可以通过使用键盘的操作输入来执行标注。作为选择，可以通过使用触摸面板的操作输入来进行标注。

[0067] 此外,在第一实施例中,使用由分类单元103分类的代表性位置作为初始值,并且在步骤S14b中针对初始值进行标注。然而,本发明不必限于此。例如,可以显示合成有表示代表性位置的信息的拍摄图像作为与图4中所示的画面不同的画面而不设置初始值,从而使用户参照代表性位置进行标注。作为选择,可以在图4中所示的画面上将代表性位置绘制为半透明标记,从而使用户参照代表性位置进行标注。

[0068] [第二实施例]

[0069] 在包括该实施例的以下实施例及变型例中,将描述与第一实施例的差异,并且假设其余与第一实施例中的相同,除非下面另外具体说明。除根据第一实施例的图像处理装置的功能以外,根据该实施例的图像处理装置还具有运算各个类别的可靠度并且由用户根据可靠度切换是否进行标注的功能。由于仅当可靠度低时用户才需要确认和校正标签,因此减少了用于标注的劳动。

[0070] 首先,将参照图6的框图来描述根据该实施例的物体识别装置的功能布置的示例。与图1相同的附图标记在图6中表示相同的功能单元,将省略这些功能单元的描述。在根据该实施例的物体识别装置2000中,图像处理装置2100的布置与图1中所示的图像处理装置1100的布置不同。

[0071] 除分类单元103的功能以外,分类单元203还具有生成分类信息并且将分类信息发送到可靠度运算单元209的功能,所述分类信息包括上述的代表性位置列表、由识别单元102识别的识别位置的集合、以及表示识别位置被分类到哪个类别中的信息。

[0072] 可靠度运算单元209针对各个类别根据属于类别的识别位置的数量运算可靠度。如果可靠度小于阈值的类别的数量等于或大于预定数量(例如,1),则可靠度运算单元209确定由用户进行标注是必要的,并且将其通知给输入单元105。因此,如在第一实施例中那样,输入单元105进行用于由用户进行标注的处理。另一方面,如果可靠度小于阈值的类别的数量小于预定数量,则可靠度运算单元209确定由用户进行标注不是必要的,并且将代表性位置列表和在步骤S11中由图像获取单元101获取的拍摄图像作为学习数据存储在存储单元106中。

[0073] 将参照图7的流程图来描述根据该实施例的步骤S2和S3的处理的详情。注意,与图2B中相同的步骤编号在图7中表示相同的处理步骤,并且将省略这些处理步骤的描述。在图7中,在步骤S2中执行步骤S11至S15,并且在步骤S3中执行步骤S16和S17。

[0074] 在步骤S23中,分类单元203进行上述的步骤S13的处理,并且还生成分类信息并将分类信息发送到可靠度运算单元209,所述分类信息包括代表性位置列表、由识别单元102识别的识别位置的集合、以及表示识别位置被分类到哪个类别中的信息。

[0075] 在步骤S24a中,可靠度运算单元209针对各个类别根据属于类别的识别位置的数量运算可靠度。关注类别的可靠度被运算为使得其随着属于关注类别的识别位置的数量增加而变得更高,或者随着识别位置的数量减少而变得更低。例如,可以使用属于关注类别的识别位置的数量作为关注类别的可靠度。例如,与图3D中所示的代表性位置标记37对应的类别的可靠度是识别位置标记组35中的识别位置的数量。

[0076] 对于确实存在目标物体的位置附近的像素,输入目标物体附近的图像。由于这个原因,确定存在目标物体的概率是高的。另一方面,对于不是目标物体的位置,如果恰好提取与目标物体类似的特征,则发生识别错误。在这种情况下,相对于前者情况,确定存在目

标物体的概率是低的。由于上述的原因,认为被识别为标物体的位置的数量越大,可靠度越高。

[0077] 在步骤S24b中,使用可靠度小于阈值的类别的数量作为指标,可靠度运算单元209确定指标是否等于或大于预定数量。注意,其他值也被认为适于作为指标。例如,可以使用“可靠度小于阈值的类别的数量”与类别的总数之比作为指标。

[0078] 如果作为确定的结果,指标等于或大于预定数量,则确定由用户进行标注是必要的,并且处理前进到步骤S14a。另一方面,如果作为确定的结果,指标小于预定数量,则可靠度运算单元209确定由用户进行标注是不必要的,并且处理前进到步骤S15。在这种情况下,在步骤S15中,可靠度运算单元209将代表性位置列表和在步骤S11中由图像获取单元101获取的拍摄图像存储在存储单元106中作为学习数据。

[0079] 如上所述,根据该实施例,当对新图像进行标注时,仅当由当前的识别器识别的代表性位置的可靠度低时,用户才进行标注。由于这个原因,需要标注的图像的数量减少,并且减少了用户标注的劳动。

[0080] <第二实施例的变型例>

[0081] 在第二实施例中,关注类别的可靠度被运算为使得其随着属于关注类别的识别位置的数量增加而变得更高,或者随着识别位置的数量减少而变得更低。然而,获得关注类别的可靠度的方法不限于特定的获得方法。例如,输出识别单元102的CNN作为表示位置是否对应于目标物体的概率。由于这个原因,关注类别的可靠度可以被设置为,随着与属于关注类别的识别位置对应的CNN的输出(概率)的总和变高而高,并且关注类别的可靠度可以被设置为随着该总和变低而低。另外,代替使用属于类别的识别位置的数量作为类别的可靠度,可以将可靠度定义为使用表示位置是否对应于目标物体的概率作为权重的线性和,或者可以将可靠度定义为使用各个识别位置距代表性位置的距离的倒数作为权重的线性和。

[0082] 另外,在第二实施例中,要与可靠度进行比较以确定由用户进行的标签输入(代表性位置的编辑操作)是否必要的阈值是预定值。然而,本发明不必限于此。例如,阈值的值可以根据图4中所示的画面上的用户操作可设置。可以使用滑块来设置阈值,或者可以通过文本输入来设置阈值。作为选择,可以通过收集关于标签被改变的代表性位置和标签未被改变的代表性位置的可靠度的统计资料来获得阈值。例如,如果标签被改变的代表性位置的可靠度的平均值为10,并且标签未被改变的代表性位置的可靠度的平均值为50,则将作为它们之间的中间值的30设置为阈值。这减少了用户用于决定阈值的劳动。

[0083] 另外,在第二实施例中,在由用户进行的标签输入不是必要的情况下,可靠度运算单元209将代表性位置列表和在步骤S11中由图像获取单元101获取的拍摄图像存储在存储单元106中作为学习数据。然而,本发明不必限于此。在由用户进行的标签输入不是必要的情况下,由于由当前的识别器能够进行充分的识别,因此例如,可以丢弃学习数据而不将其存储在存储单元106中,或者可以使得用户能够确定是否将数据添加为学习数据。由于由此省略了一些处理(主要是学习单元),因此能够更快地结束该系列处理。

[0084] 此外,在第二实施例中,在由用户进行的标签输入是必要的情况下,输入单元105在画面上显示登记在代表性位置列表中的全部代表性位置作为编辑候选,并且接受编辑操作。然而,本发明不必限于此。由于标签输入对于可靠度不小于阈值的类别不是必要的,因此输入单元105可以仅在画面上显示可靠度小于阈值的类别的代表性位置作为编辑候选并

且接受编辑操作。相反,如果可靠度极低,则可以将其视为识别错误。因此,还可以提供用于下限的阈值,并且输入单元105可以仅在画面上显示可靠度不小于下限的阈值且小于上限的阈值(在步骤S24b中使用的阈值)的类别的代表性位置作为编辑候选并且接受编辑操作。这减少了由用户校正的标签的数量,并且减少了用户标注的劳动。另外,可以在图4中所示的画面上显示表示针对各个代表性位置的可靠度的信息。对于表示可靠度的信息的显示,例如,可以将可靠度显示为文本,或者可以显示具有根据可靠度的颜色或形状的预定的图案图像,或者可以显示针对各个可靠度预先创建的图案图像。这为用户在进行标注时确定代表性位置的可靠程度提供了参照。

[0085] [第三实施例]

[0086] 除根据第一实施例的图像处理装置的功能以外,根据该实施例的图像处理装置还具有获得学习后的CNN的性能的评价值并且基于所获得的评价值确定是否更新CNN的功能。当仅在评价值变高的情况下更新CNN时,加速CNN的性能的提高,能够获得更令人满意的结果,并且由用户进行的标签校正的量减少。由于这个原因,减少了用于标注的劳动。

[0087] 首先,将参照图8的框图来描述根据该实施例的物体识别装置的功能布置的示例。与图1中相同的附图标记在图8中表示相同的功能单元,并且将省略这些功能单元的描述。在根据该实施例的物体识别装置3000中,图像处理装置3100的布置与图1中所示的图像处理装置1100的布置不同。

[0088] 除上述的学习单元107的功能以外,学习单元307还将通过学习处理获得的CNN的参数发送到评价单元310。评价单元310使用存储在存储单元311中的数据,基于保存在识别器存储单元108中的CNN的模型结构和CNN的参数来获得CNN(旧CNN)的性能的评价值。评价单元310还使用存储在存储单元311中的数据,基于从学习单元307发送的参数和保存在识别器存储单元108中的CNN的模型结构来获得要用于评价CNN(新CNN)的性能的评价值。测试图像和测试图像中的代表性位置被作为测试数据而存储在存储单元311中。

[0089] 更新确定单元312将旧CNN的评价值与新CNN的评价值进行比较,并且基于比较的结果确定是否将新CNN的参数存储在识别器存储单元108中。如果新CNN的评价值高于旧CNN的评价值,则更新确定单元312将新CNN的参数存储在识别器存储单元108中,从而将由识别器存储单元108保持的旧CNN更新为新CNN。

[0090] 将参照图9的流程图来描述根据该实施例的步骤S2和S3的处理的详情。在图9中,在步骤S2中执行步骤S11至S15,并且在步骤S3中执行步骤S36至S39。

[0091] 在步骤S36中,学习单元307进行上述的步骤S16的处理,并且还将通过学习获得的CNN的参数发送到评价单元310。

[0092] 在步骤S37中,评价单元310使用存储在存储单元311中的数据获得旧CNN的评价值和新CNN的评价值。获得旧CNN和新CNN中的各个的评价值的方法包括各种方法,并且不限于特定的方法。下面将描述获得CNN(旧CNN和新CNN)的评价值的示例。

[0093] 首先,评价单元310从存储单元311获取测试数据。测试数据包括多个集合,所述多个集合中的各个集合包括:包含一个物体的测试图像、以及作为测试图像中的物体的位置的代表性位置。评价单元310选择测试数据中包括的多个集合中的未选择的集合作为所选集合,并且如在步骤S12中那样,将CNN应用于所选集合中包括的测试图像,从而获取测试图像中的识别位置。如果从所选集合中的测试图像识别的全部识别位置都位于距所选集合中

包括的代表性位置的预定距离内，则评价单元310确定成功。另一方面，如果从所选集合中的测试图像识别的识别位置中的至少一个不位于距所选集合中包括的代表性位置的预定距离内，则评价单元310确定失败。针对测试数据中包括的全部集合进行对所选集合的这样的成功/失败确定处理，从而针对各个集合确定成功/失败。然后，评价单元310获得((被确定为成功的集合的数量)/集合的总数)作为评价值。针对旧CNN和新CNN中的各个获得这样的评价值。注意，如上所述，评价值运算方法仅仅是示例，并且可以使用其他运算方法来运算评价值。

[0094] 在步骤S38中，更新确定单元312将旧CNN的评价值与新CNN的评价值进行比较。如果作为比较的结果，旧CNN的评价值<新CNN的评价值，则处理前进到步骤S39。如果旧CNN的评价值≥新CNN的评价值，则根据图9的流程图的处理结束，而不更新存储在识别器存储单元108中的参数。在步骤S39中，更新确定单元312将新CNN的参数存储在识别器存储单元108中，从而将由识别器存储单元108保持的CNN更新为新CNN。

[0095] 如上所述，根据该实施例，当对新图像进行标注时，作为参考呈现给用户的代表性位置的精确性提高。因此，标签校正量减少，并且减少了用户标注的劳动。

[0096] <第三实施例的变型例>

[0097] 在第三实施例中，可以省略上述的步骤S1的处理。也就是说，由于评价单元310运算新CNN的评价值，因此可以提供用于评价值的阈值，并且可以反复地继续CNN的学习，直到新CNN的评价值变得等于或大于用于评价值的阈值。另外，在反复学习期间，当评价值的改变变得小于预定值时可以结束学习。例如，当评价值连续五次改变小于0.01%时结束学习，并且性能被认为是饱和的。这减少了用户用于确定学习是否必要的劳动。

[0098] 另外，在第三实施例中，输入单元105总是将拍摄图像和代表性位置存储在存储单元106中。然而，由于不仅学习数据需要增加而且测试数据也需要增加，因此图像和代表性位置的集合不仅可以被存储在存储单元106中，还可以被存储在存储单元311中。注意，图像和代表性位置的集合可以被存储在存储单元106和存储单元311二者中，或者可以被存储在存储单元106和存储单元311中的随机确定的一者中。另外，可以由用户通过对操作设备3进行操作来指定存储单元106和存储单元311中的哪一个被用来存储图像和代表性位置的集合。

[0099] 另外，第二实施例中描述的可靠度运算单元209可以被并入在图像处理装置3100中，并且可以根据由可靠度运算单元209运算的可靠度来决定存储单元106和存储单元311中的哪一个被用来存储图像和代表性位置的集合。例如，可靠度小于阈值的代表性位置和对应的图像被存储在存储单元106中，并且可靠度不小于阈值的代表性位置和对应的图像被存储在存储单元311中。这使得能够有效地收集学习数据和测试数据。

[0100] 此外，在步骤S38中，不仅可以进行确定是否更新CNN，还可以进行确定是否保持新登记的学习数据(在下文中将被称为新学习数据)。由于性能的劣化表示在新学习数据的标签中包含错误的可能性高，所以可以丢弃新学习数据。作为选择，为了确认新学习数据的标签是否真的正确，可以要求用户再次输入标签。这能够降低存储包含标签错误的学习数据的频率。

[0101] 另外，在第三实施例中，评价两个CNN(即，旧CNN和新CNN)。然而，本发明不必限于此。可以存在多个旧CNN，并且可以存在多个新CNN。在这种情况下，例如，对于旧CNN，将旧

CNN和新CNN的高评价值的前三名的CNN存储在识别器存储单元108中。另外,例如,对于新CNN,通过将随机值或将旧CNN中的一个设置为初始值或者通过随机选择学习数据来获得多个学习结果。

[0102] [第四实施例]

[0103] 在上述的实施例及变型例中,每当新添加学习数据时进行学习。然而,本发明不必限于此。例如,可以在收集多个集合(图像和代表性位置的集合)之后进行学习,或者用户可以通过对操作设备3进行操作来指定要进行学习的定时。

[0104] 另外,在上述的实施例及变型例中,通过对作为一种深度学习的CNN进行学习来构建识别器。然而,本发明不必限于此。可以使用LBP(Local Binary Pattern,局部二值模式),或者可以使用BoF(Bag of Features,特征包)作为特征。可以使用决策树,或者可以使用SVM(Support Vector Machine,支持向量机)等作为鉴别器。

[0105] 另外,在上述的实施例及变型例中,仅识别图像中的目标物体的位置。然而,本发明不必限于此。例如,CNN的输出不限于用于确定位置是否对应于目标物体的2级辨别,并且还可以添加输出以对姿态进行分类,从而识别图像中的目标物体的位置和姿态。下面将参照图10A至图10C来描述不仅识别位置还识别姿态的处理。

[0106] 图10A中所示的×标记501表示要识别的目标物体在图像中的位置(x,y),空心箭头502表示要在图像中识别的目标物体的姿态θ,并且黑色箭头503表示将目标物体的姿态量化成八个姿态的方向。如图10A中所示,标签被设置为目标物体的位置和姿态(x,y,θ),并且CNN被设计为进行9级辨别以辨别“目标物体具有八个姿态中的哪一个”+“位置不对应于目标物体”。

[0107] 当创建这样的CNN并且如在第一实施例的步骤S12中进行识别时,如图10B中所示,获得多个位置和姿态识别结果(识别位置和姿态)。图10B中所示的×标记组504表示物体的识别位置组,并且空心箭头505的组表示识别姿态组。

[0108] 如在第一实施例的步骤S13中那样,进行识别结果的聚类,从而获得代表性位置和姿态并将其呈现给用户。当获得代表性位置和姿态时,可以将各个目标物体的位置和姿态(x,y,θ)定义为特征,并且可以使用K均值算法来进行聚类。作为选择,在对位置(x,y)进行聚类之后,可以进行类别中的姿态θ的聚类。因而,能够获得由图10C中的○标记506(表示代表性位置)和黑色箭头507(表示代表性姿态)表示的代表性位置和姿态。

[0109] 输入单元105不仅能够编辑位置,还能够编辑姿态。为了校正姿态,例如,设置从目标物体的位置至鼠标的光标的位置的方向向量。作为选择,可以在决定目标物体的位置之后决定姿态,或者可以切换位置编辑模式和姿态编辑模式。

[0110] 注意,上面已经描述了二维图像中的位置和姿态的情况。然而,即使增加维度数来进行识别例如三维空间中的位置和姿态,也能够通过将标签设置为(x,y,z,θ,φ,ψ)(θ、φ和ψ分别是关于x轴、y轴和z轴的角度)来进行与上述相同的处理。

[0111] [第五实施例]

[0112] 可以通过硬件来实现图1、图6和图8中所示的图像处理装置1100、2100和3100的功能单元。然而,可以通过软件(计算机程序)来实现这些功能单元中的一些。在后者情况下,能够执行用于实现除存储单元106、识别器存储单元108和存储单元311以外的功能单元的功能的计算机程序的计算机装置,能够应用于图像处理装置1100、2100和3100。将参照图11

的框图来描述可应用于图像处理装置1100、2100和3100的计算机装置的硬件布置的示例。

[0113] CPU 1101使用存储在RAM 1102和ROM 1103中的计算机程序和数据来执行处理。因此,CPU 1101控制整个计算机装置的操作,并且还执行或控制上面说明为要由图像处理装置1100、2100和3100中的各个执行的各个处理。

[0114] RAM 1102具有用于存储经由I/F(接口)1105从摄像设备1获取的拍摄图像或者从ROM 1103或外部存储设备1104加载的计算机程序或数据的区域。另外,RAM 1102具有要由CPU 1101用来执行各种处理的工作区域。因此,RAM 1102能够根据需要提供各种区域。ROM 1103存储不需要重写的计算机程序和数据(例如, BIOS的计算机程序和数据)。

[0115] 外部存储设备1104是由硬盘驱动器表示的大容量信息存储设备。OS(操作系统)或被构造为使CPU 1101执行或控制上面说明为要由图像处理装置1100、2100和3100中的各个执行的各个处理的计算机程序和数据,被保存在外部存储设备1104中。存储在外部存储设备1104中的计算机程序包括如下的计算机程序,该计算机程序被构造为使CPU 1101执行除图1、图6和图8中的存储单元106、识别器存储单元108和存储单元311以外的功能单元的功能。另外,保存在外部存储设备1104中的数据包括在上面的描述中说明为已知信息的数据。保存在外部存储设备1104中的计算机程序或数据在CPU 1101的控制下根据需要被加载到RAM 1102并由CPU 1101处理。注意,能够使用RAM 1102或外部存储设备1104来实现上述的存储单元106、识别器存储单元108和存储单元311。注意,诸如USB存储器或SD卡的各种存储器设备能够应用于外部存储设备1104。

[0116] I/F 1105由接口组形成以连接上述的摄像设备1、显示设备2和操作设备3。所有CPU 1101、RAM 1102、ROM 1103、外部存储设备1104和I/F 1105连接到总线1106。

[0117] 注意,可以适当地组合并使用上述的实施例及变型例中的一些或全部。另外,可以选择性地使用上述的实施例及变型例中的一些或全部。

[0118] 其他实施例

[0119] 另外,可以通过读出并执行记录在存储介质(也可更完整地称为“非临时性计算机可读存储介质”)上的计算机可执行指令(例如,一个或更多个程序)以执行上述实施例中的一个或更多个的功能、并且/或者包括用于执行上述实施例中的一个或更多个的功能的一个或更多个电路(例如,专用集成电路(ASIC))的系统或装置的计算机,来实现本发明的实施例,并且,可以利用通过由所述系统或装置的所述计算机例如读出并执行来自所述存储介质的所述计算机可执行指令以执行上述实施例中的一个或更多个的功能、并且/或者控制所述一个或更多个电路执行上述实施例中的一个或更多个的功能的方法,来实现本发明的实施例。所述计算机可以包括一个或更多个处理器(例如,中央处理单元(CPU)、微处理单元(MPU)),并且可以包括分开的计算机或分开的处理器的网络,以读出并执行所述计算机可执行指令。所述计算机可执行指令可以例如从网络或所述存储介质被提供给计算机。所述存储介质可以包括例如硬盘、随机存取存储器(RAM)、只读存储器(ROM)、分布式计算系统的存储器、光盘(诸如压缩光盘(CD)、数字通用光盘(DVD)或蓝光光盘(BD)TM)、闪存设备以及存储卡等中的一个或更多个。

[0120] 本发明的实施例还可以通过如下的方法来实现,即,通过网络或者各种存储介质将执行上述实施例的功能的软件(程序)提供给系统或装置,该系统或装置的计算机或是中央处理单元(CPU)、微处理单元(MPU)读出并执行程序的方法。

[0121] 虽然参照示例性实施例对本发明进行了描述,但是应当理解,本发明并不限于所公开的示例性实施例。应当对所附权利要求的范围给予最宽的解释,以使其涵盖所有这些变型例以及等同的结构和功能。

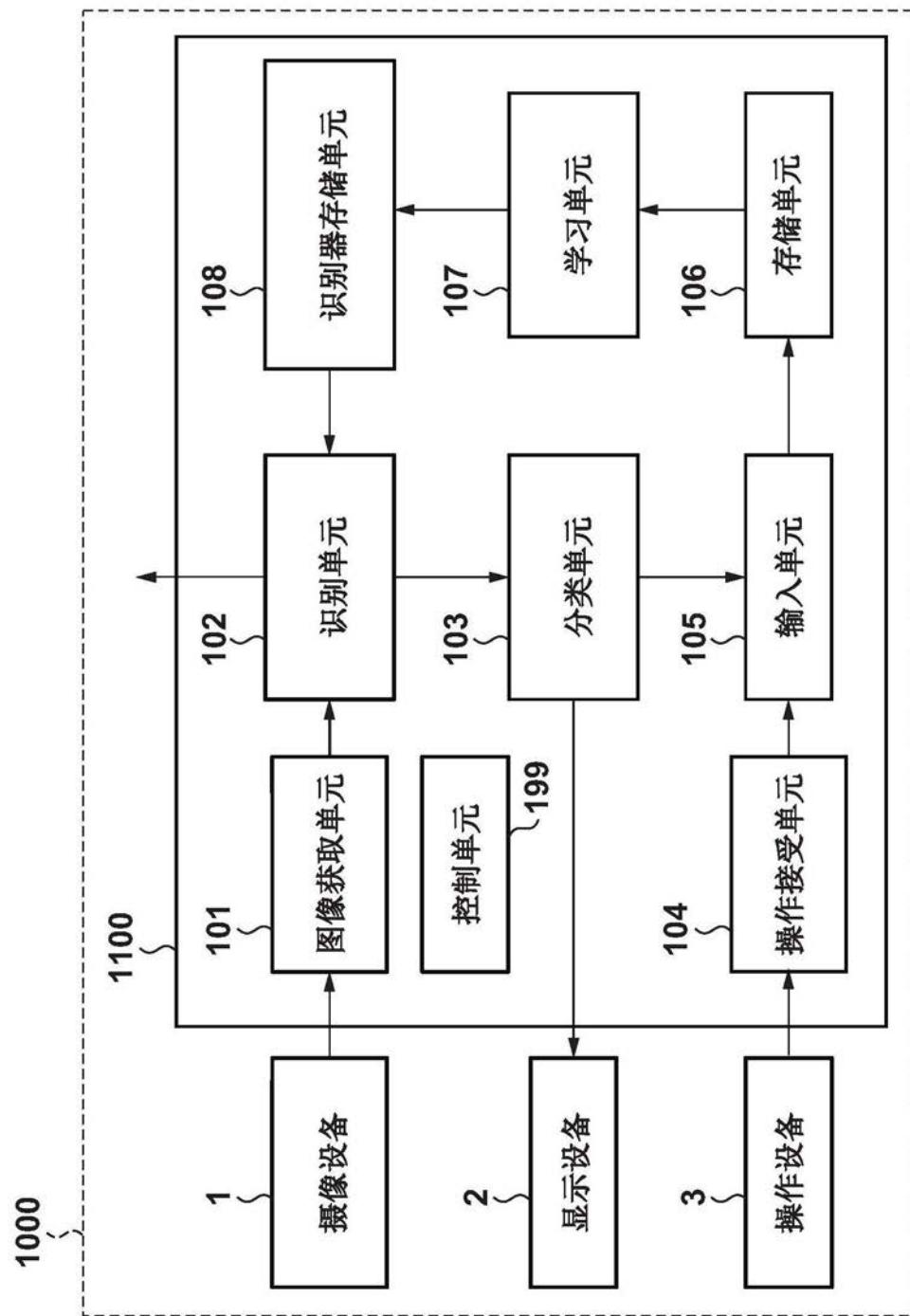


图1

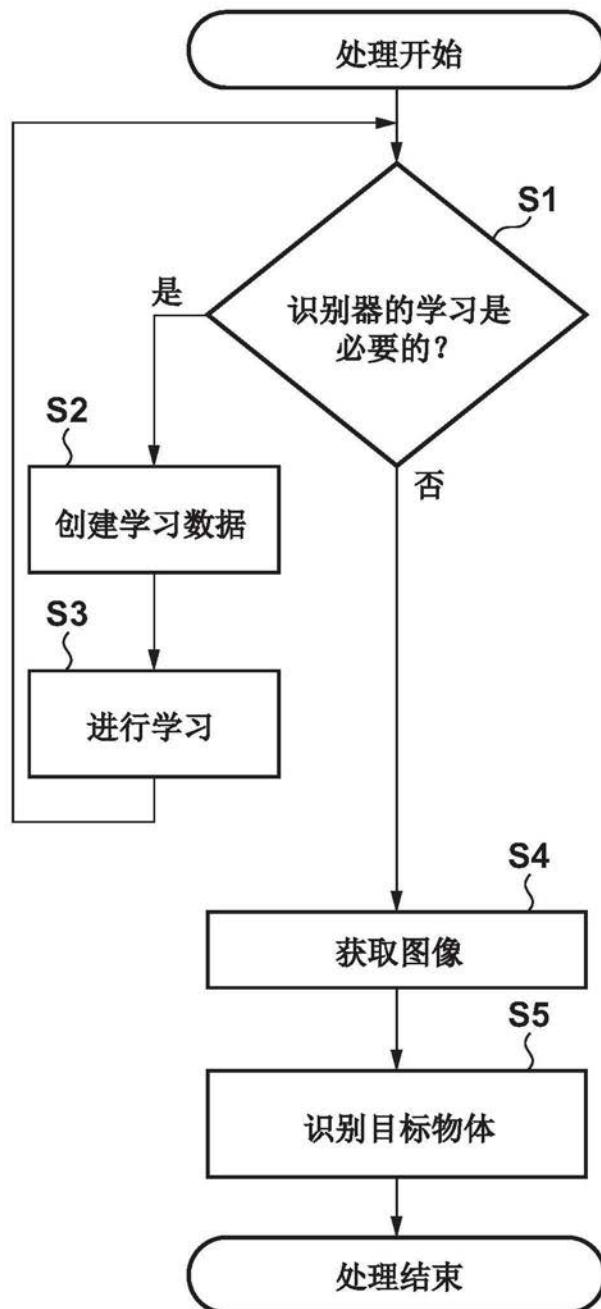


图2A

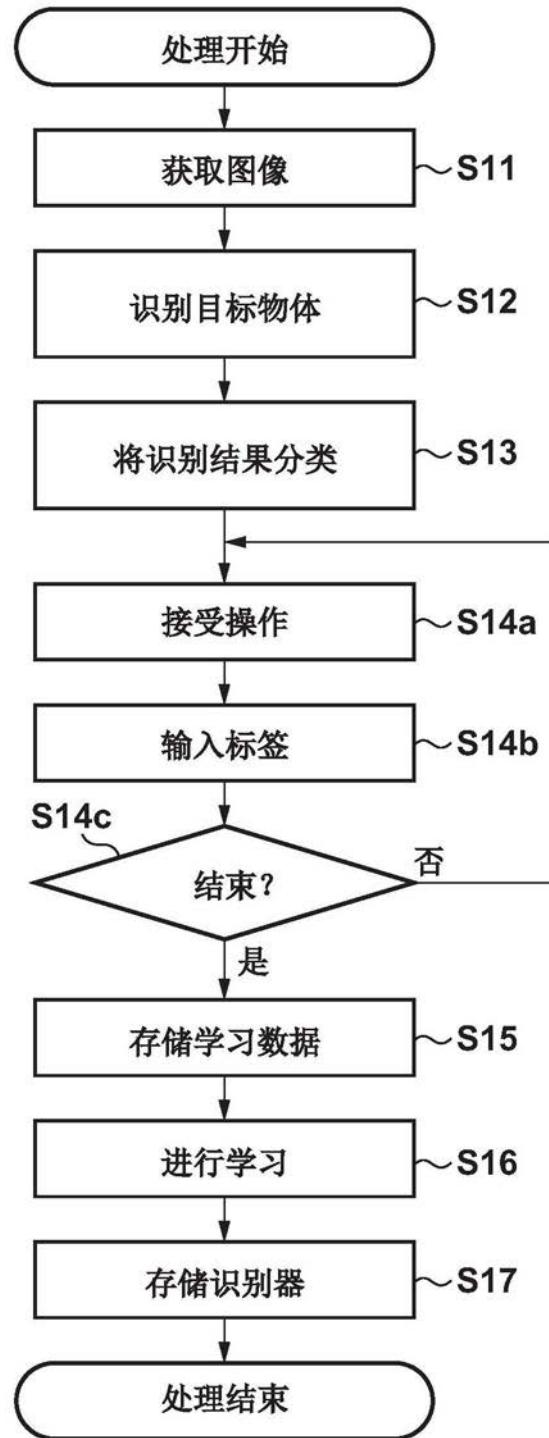


图2B

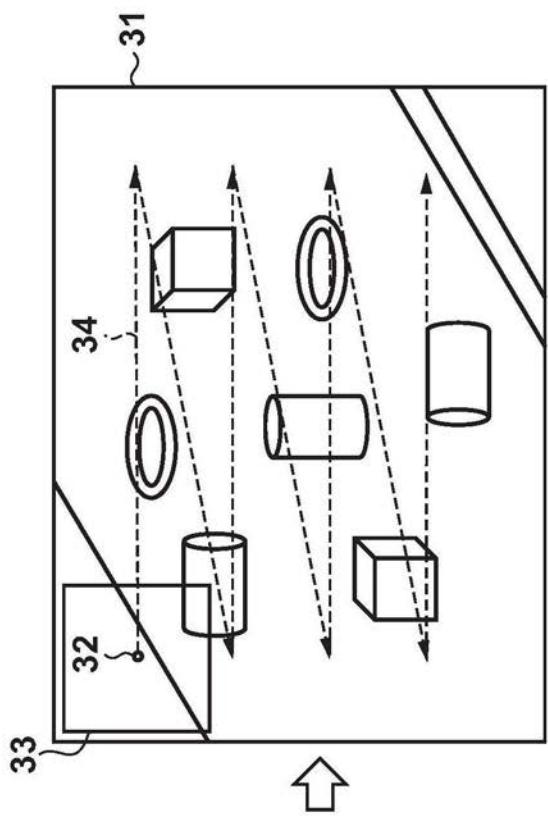


图3A

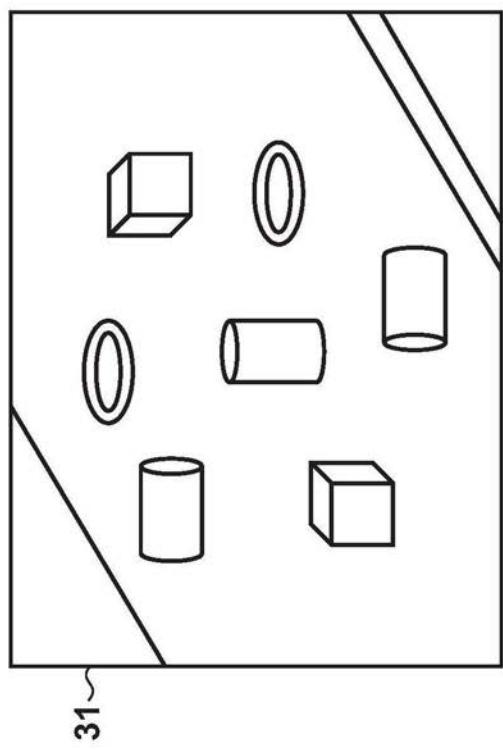


图3B

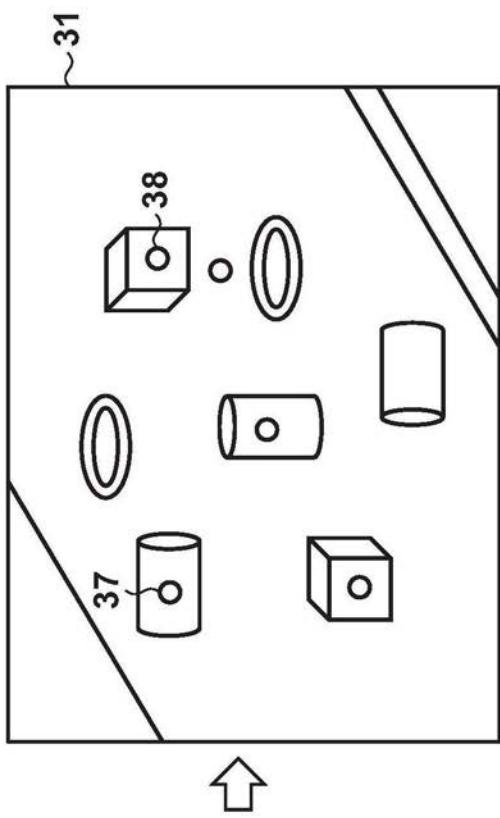


图3C

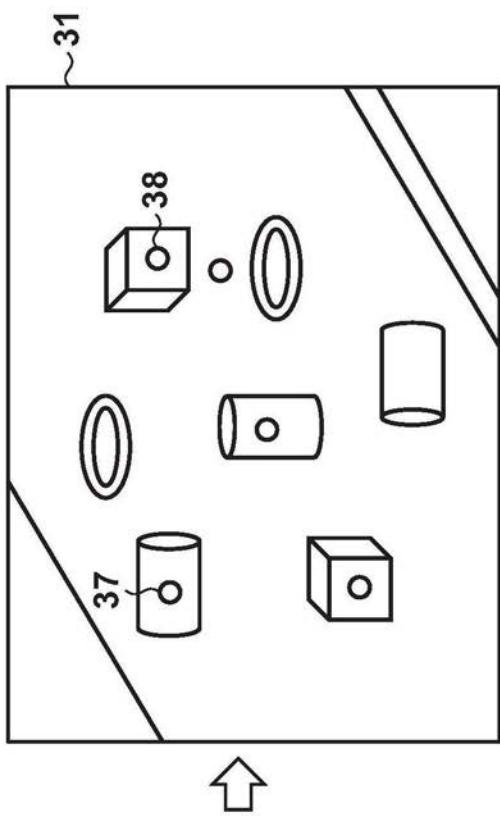


图3D

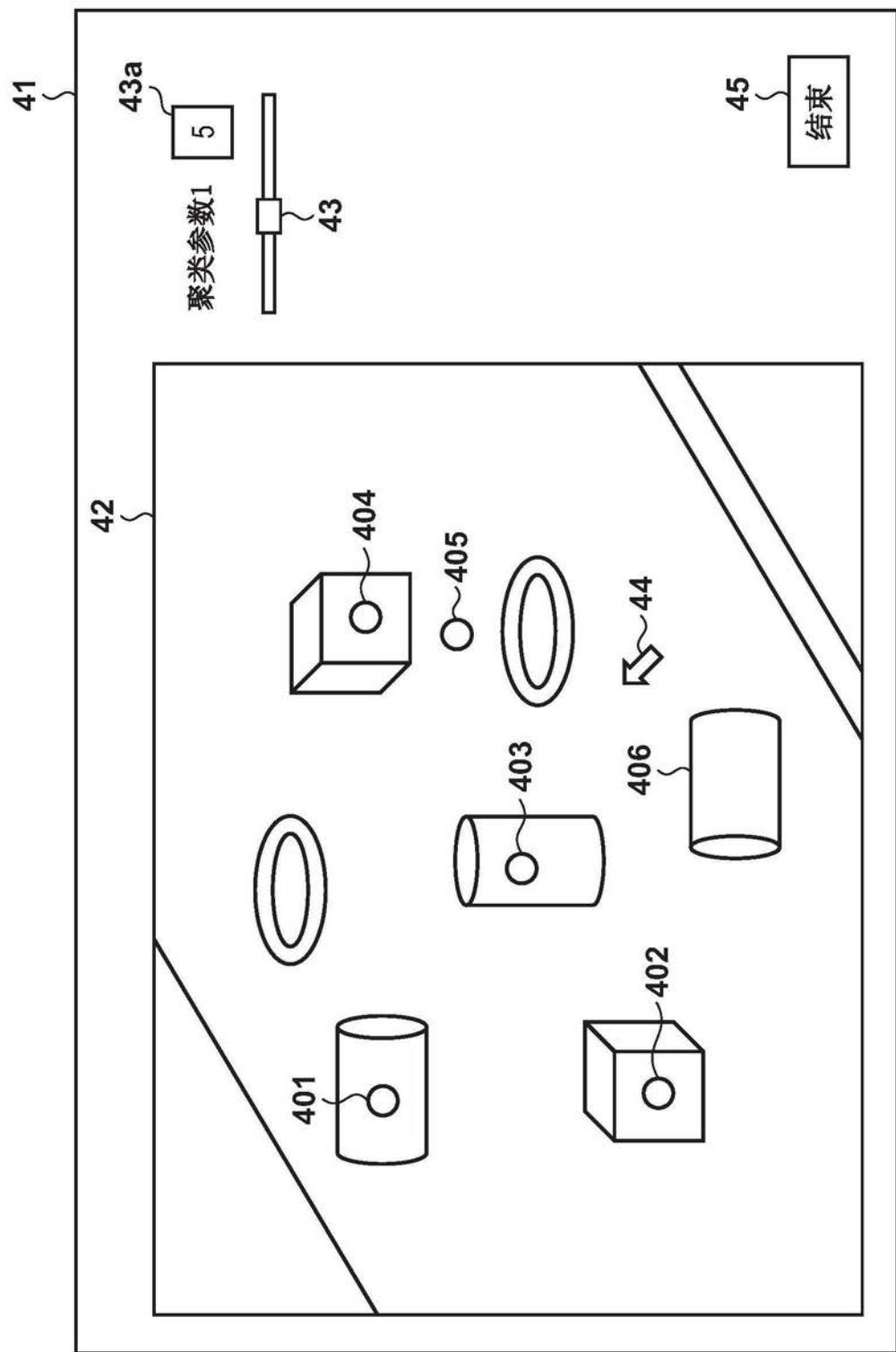


图4

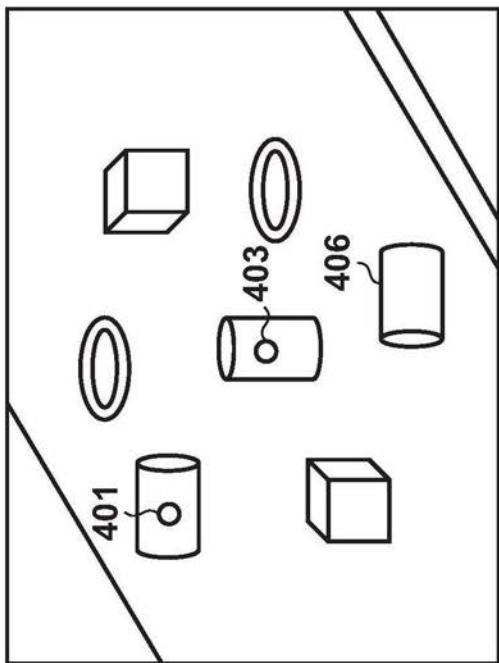


图5B

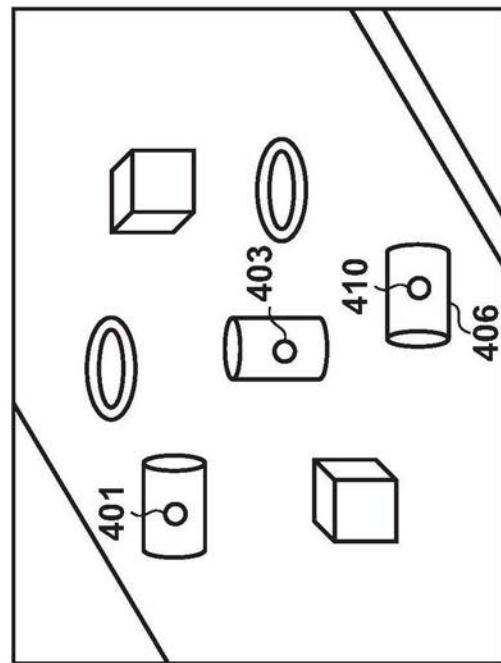


图5D

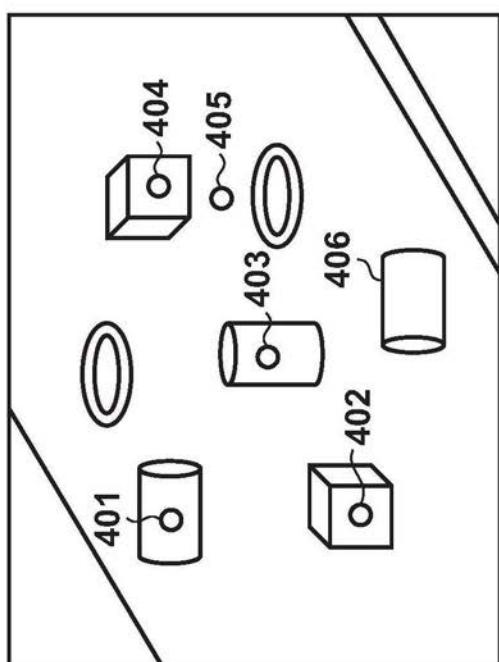


图5A

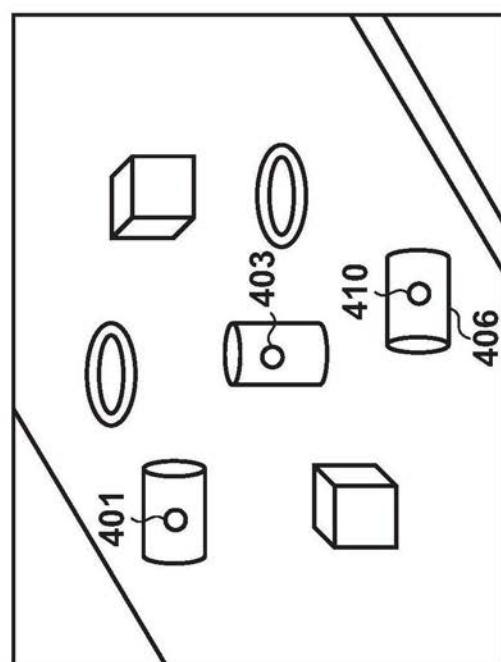


图5C

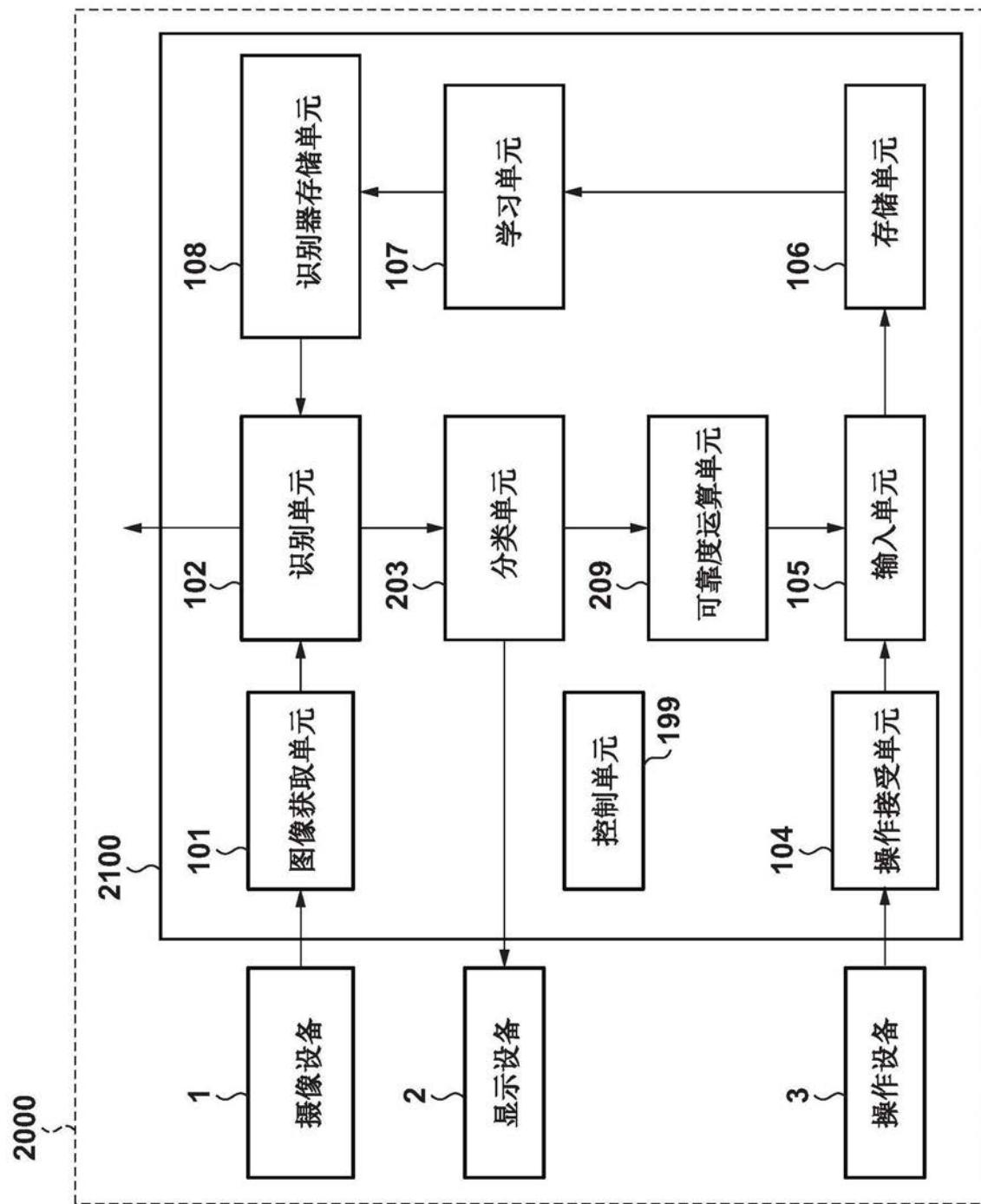


图6

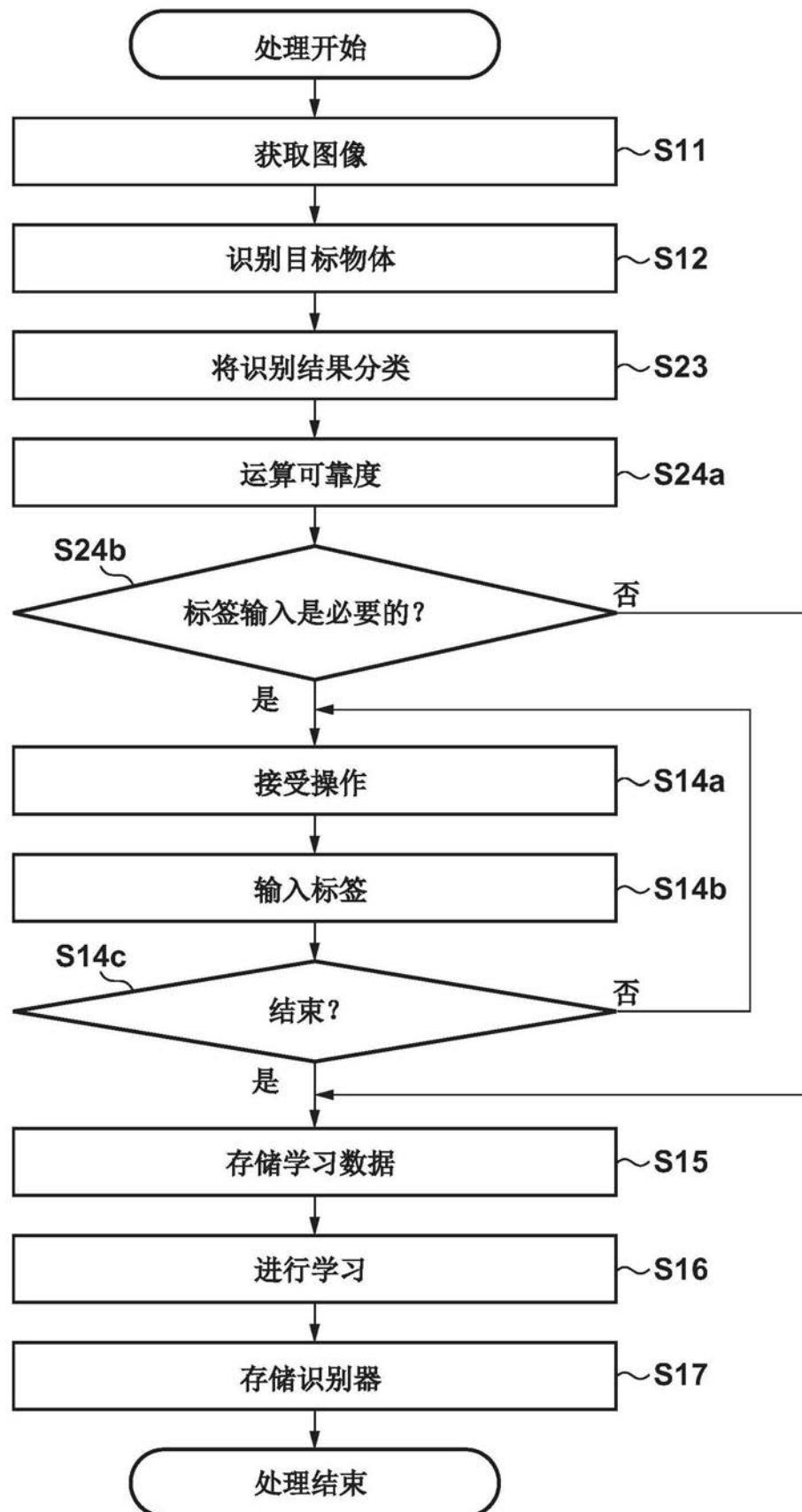


图7

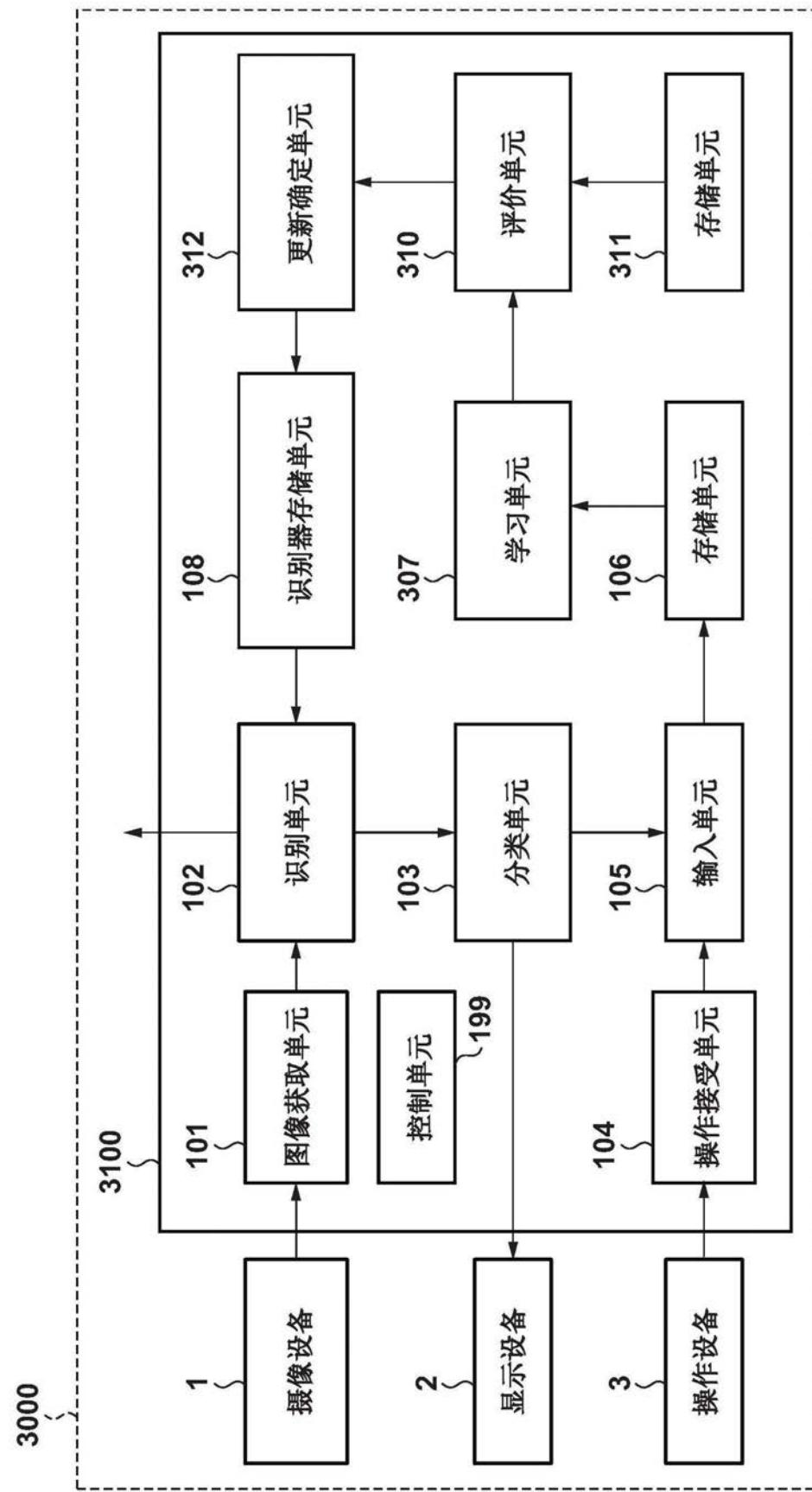


图8

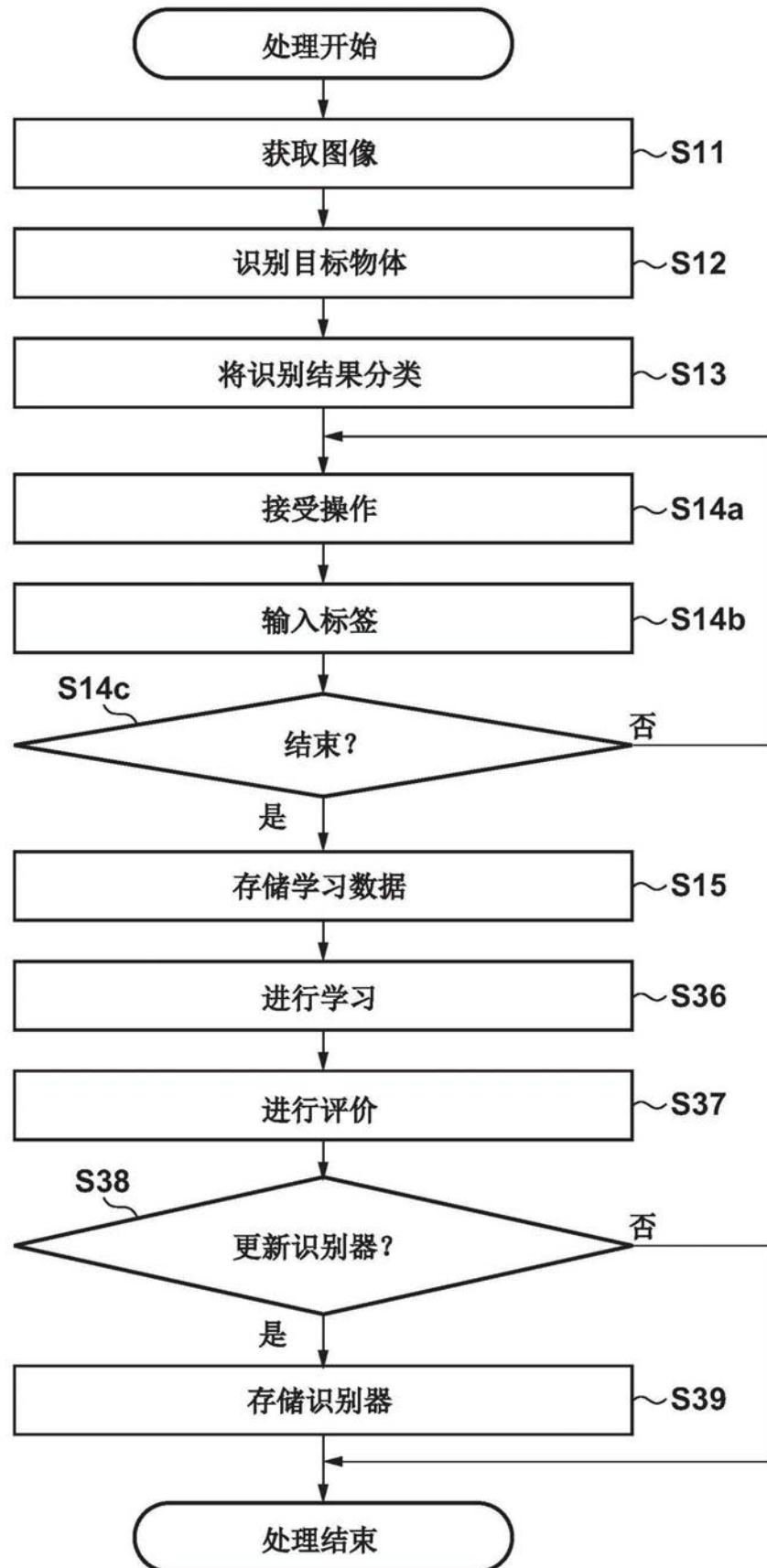


图9

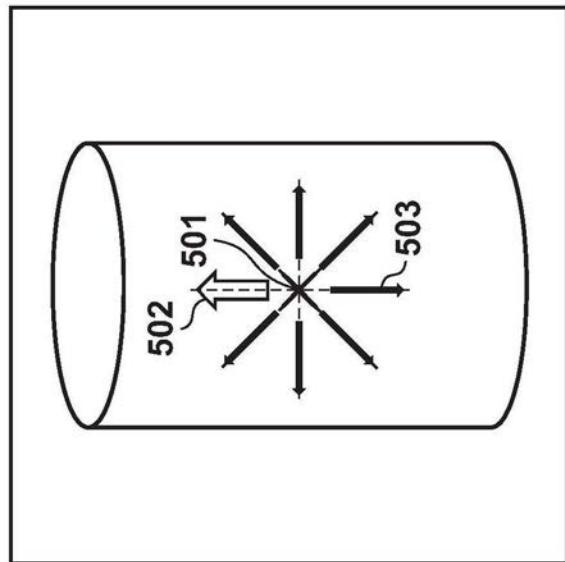


图10A

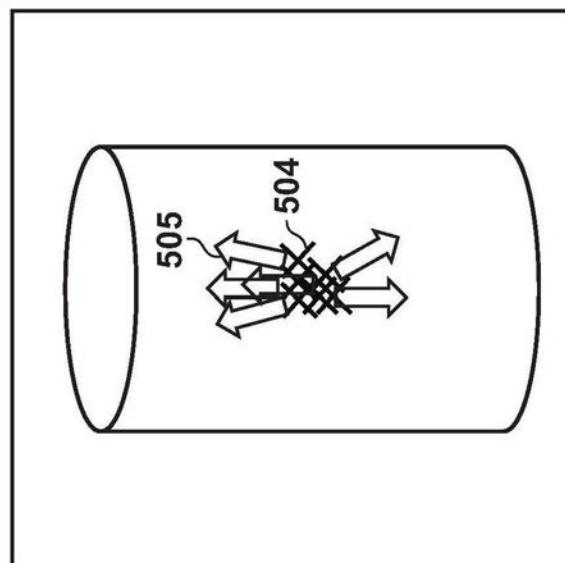


图10B

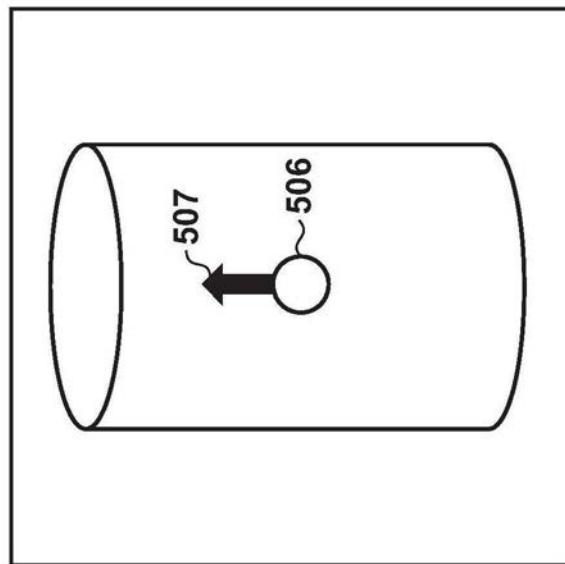


图10C

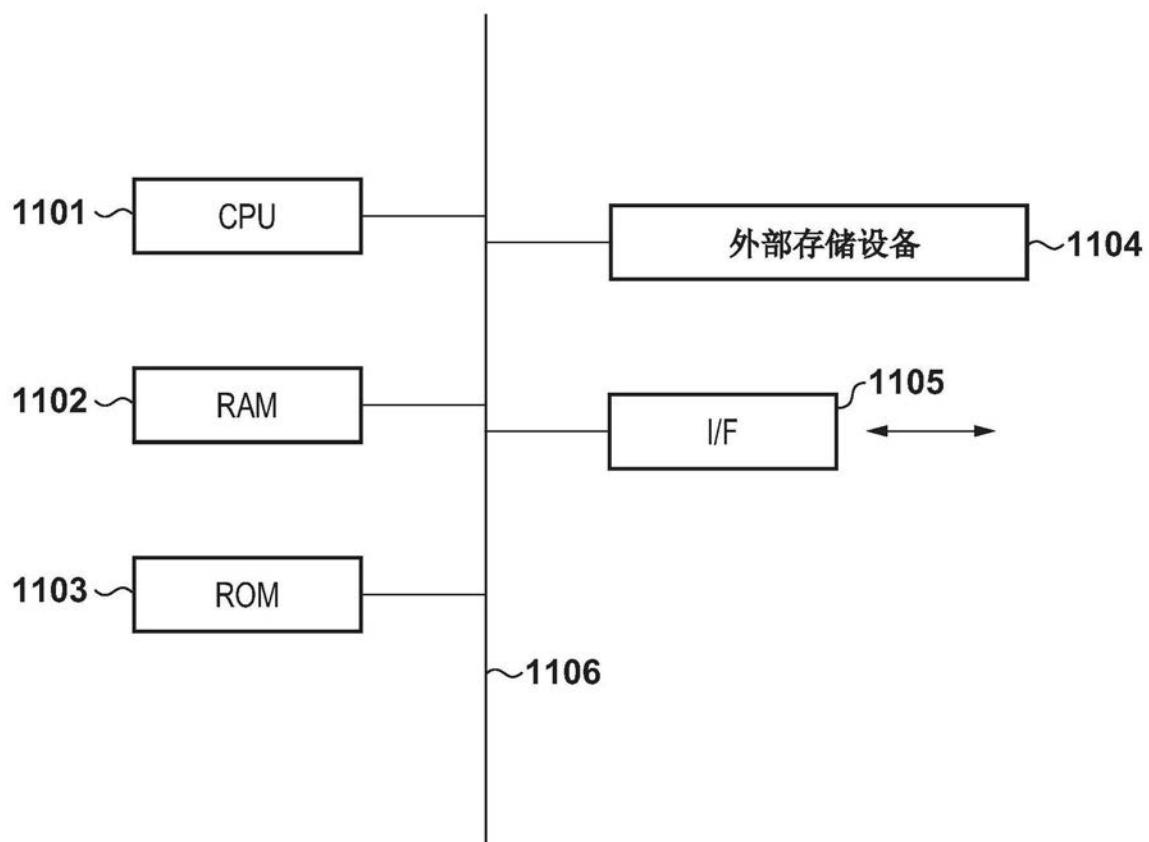


图11