



(12) 发明专利申请

(10) 申请公布号 CN 119422161 A

(43) 申请公布日 2025. 02. 11

(21) 申请号 202380048207.3

(22) 申请日 2023.06.19

(30) 优先权数据

2023-073580 2023.04.27 JP

63/354,008 2022.06.21 US

(85) PCT国际申请进入国家阶段日

2024.12.19

(86) PCT国际申请的申请数据

PCT/JP2023/022533 2023.06.19

(87) PCT国际申请的公布数据

W02023/248968 JA 2023.12.28

(71) 申请人 松下电器(美国)知识产权公司

地址 美国

(72) 发明人 谷川理佐子 石坂隼 小冢和纪

(74) 专利代理机构 中科专利商标代理有限责任公司 11021

专利代理师 韩丁

(51) Int.Cl.

G06T 3/04 (2024.01)

G06T 7/00 (2017.01)

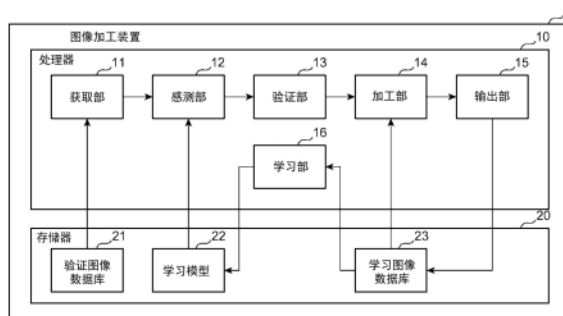
权利要求书2页 说明书12页 附图12页

(54) 发明名称

图像加工方法、图像加工装置以及图像加工程序

(57) 摘要

图像加工装置获取通过与物体的正解标签建立了对应的全方位图像构成的图像,执行从获取到的图像感测物体的物体感测处理,基于正解标签来算出物体感测处理中的物体的感测精度,并在感测精度低于阈值的情况下,对图像进行加工,以使得图像中包含的物体的变形变大。



1. 一种图像加工方法,是计算机中的图像加工方法,获取由全方位图像构成的图像,执行从获取到的所述图像感测物体的物体感测处理,算出所述物体感测处理中的所述物体的感测精度,基于所述感测精度,对所述图像进行加工,以使得所述图像中包含的物体的变形变大,输出进行了所述加工的加工图像。
2. 根据权利要求1所述的图像加工方法,其中,所述图像是通过与物体的正解标签建立了对应的全方位图像构成的图像,基于所述正解标签来算出所述感测精度,在所述感测精度低于阈值的情况下进行所述加工。
3. 根据权利要求1或者2所述的图像加工方法,其中,所述图像包含第1图像以及与所述第1图像不同的第2图像,所述感测精度是相对于将所述第1图像输入到为了执行所述物体感测处理而预先进行了学习的学习模型时的感测结果的感测精度,对所述第2图像执行所述图像的加工。
4. 根据权利要求3所述的图像加工方法,其中,进一步地,使用所述加工图像来使所述学习模型进行学习。
5. 根据权利要求3所述的图像加工方法,其中,所述感测精度按物体的分类来算出,所述第2图像是包含所述第1图像中被判定为所述感测精度为阈值以下的物体的图像。
6. 根据权利要求1或者2所述的图像加工方法,其中,所述图像的加工包含将所述图像的默认的视点变更为随机设定的视点。
7. 根据权利要求1或者2所述的图像加工方法,其中,所述图像的加工包含:对所述图像中所设定的多个正解标签之中距离为最长的两个绑定框的区间进行确定,并在所述区间的中点设定所述图像的视点。
8. 根据权利要求1或者2所述的图像加工方法,其中,所述图像的加工包含:判定在所述图像中是否包含纵横比以及尺寸的至少一者超过基准值的物体;和在判定为包含超过所述基准值的物体的情况下,与判定为不包含该物体的情况相比,使所述加工图像的张数增多。
9. 根据权利要求1或者2所述的图像加工方法,其中,所述物体感测处理是基于规则的物体感测处理,所述图像的加工包含对进行了物体感测处理的图像进行所述图像的加工。
10. 根据权利要求1或者2所述的图像加工方法,其中,所述图像的加工包含:通过执行视点转换处理来对所述图像进行加工,以使得所述物体的变形变大,所述视点转换处理包含:将所述图像投影到单位球面上;根据进行了所述投影的投影图像设定新的视点;和

将所述投影图像沿平面展开,以使得所述新的视点成为中心。

11.一种图像加工装置,具备处理器,

所述处理器执行以下处理:

获取由全方位图像构成的图像,

执行从获取到的所述图像感测物体的物体感测处理,

算出所述物体感测处理中的所述物体的感测精度,

基于所述感测精度,对所述图像进行加工,以使得所述图像中包含的物体的变形变大,

输出进行了所述加工的加工图像。

12.一种图像加工程序,使计算机执行权利要求1或者2所述的图像加工方法。

图像加工方法、图像加工装置以及图像加工程序

技术领域

[0001] 本公开涉及对图像进行加工的技术。

背景技术

[0002] 在专利文献1中公开了以下技术：从通过全方位摄像机拍摄到的摄像机影像选择存在物体的可能性较高的区域，来作为候选区域，使候选区域的朝向旋转，以使得所选择的候选区域中包含的物体朝向铅垂方向，对旋转后的候选区域执行物体感测处理。

[0003] 然而，专利文献1是使候选区域的朝向旋转以使得候选区域中包含的物体的变形减少的技术，并非是特意使图像上的物体的变形增大的技术。因此，专利文献1不能生成用于从具有变形的图像高精度地感测物体的学习用的图像。

[0004] 在先技术文献

[0005] 专利文献

[0006] 专利文献1：国际公开第2013/001941号

发明内容

[0007] 本公开旨在解决这样的问题，其目的在于，提供一种生成能够从具有变形的图像高精度地感测物体的学习用的图像的技术。

[0008] 本公开的一方式中的图像加工方法是计算机中的图像加工方法，获取由全方位图像构成的图像，执行从获取到的所述图像感测物体的物体感测处理，算出所述物体感测处理中的所述物体的感测精度，基于所述感测精度，对所述图像进行加工，以使得所述图像中包含的物体的变形变大，输出进行了所述加工的加工图像。

[0009] 根据该结构，能够生成能够从具有变形的图像高精度地感测物体的学习用的图像。

附图说明

[0010] 图1是示出本公开的实施方式中的图像加工装置的结构的一个例子的框图。

[0011] 图2是示出执行视点转换处理的情形的图。

[0012] 图3是视点转换处理的说明图。

[0013] 图4是示出应用通过学习模型进行了物体感测处理的图像的用户接口的显示画面的图。

[0014] 图5是示出实施方式1中的图像加工装置的处理的一个例子的流程图。

[0015] 图6是示出图像加工装置1中的学习阶段中的处理的一个例子的流程图。

[0016] 图7是示出应用实施方式2中的视点转换处理的图像的图。

[0017] 图8是示出实施方式2中的图像加工装置的处理的一个例子的流程图。

[0018] 图9是示出具有易变形的形状的物体的一个例子的图。

[0019] 图10是示出实施方式3中的图像加工装置的处理的一个例子的流程图。

- [0020] 图11是示出实施方式4中的图像加工装置的处理的一个例子的流程图。
- [0021] 图12是示出实施方式5中的图像加工装置的结构的一个例子的框图。
- [0022] 图13是示出实施方式5中的图像加工装置1A的处理的一个例子的流程图。

具体实施方式

[0023] (实现本公开的一方式的经过)

[0024] 作为建造现场的问题,存在具体的指示未传达给作业人员、该指示的说明需要花费时间这样的交流上的问题或在巡视建造现场整体上需要人手、移动至建造现场会花费时间这样的建造现场的确认上的问题等。

[0025] 为了解决这样的问题,例如也考虑在建造现场设置多个摄像机,位于远程的现场监督一边参照从多个摄像机得到的图像一边向作业人员发出指示。然而,在建造现场,随着建造的进行,会产生将所设置的传感器拆下、或者将拆下的传感器设置在其他场所之类的作业。由于这样的作业会花费工夫,因此在建造现场设置传感器并不实用。为此,本发明人对能够不设置传感器地从远程详细地确认建造现场的状况的技术进行了研究。

[0026] 于是,认识到当在显示器所显示的建造现场的设计图中选择某个位置的操作被输入时,如果对在该位置事先拍摄到的建造现场的全方位图像进行显示,并具有能够设定用于用户对该全方位图像赋予注释的注释区域的用户接口,则能够从远程详细地确认建造现场的状况。

[0027] 在对注释区域进行设定的情况下,考虑以下方式:将预先使用学习模型进行了物体感测的全方位图像显示于显示器,当用户输入在显示器上选择某个物体的操作时,将该物体中所设定的边界框(Boundary box)显示为注释区域。在该情况下,用户即使不输入使全方位图像显示默认的框体,将该框体定位到作为目的物体,并使该框体变形以使得匹配该物体的操作,也能够设定注释区域,减轻用户的劳动。

[0028] 在此,为了生成能够从具有变形的图像高精度地感测物体的学习模型,优选将这样的图像用作学习用的图像。

[0029] 然而,以专利文献1为代表的现有技术为了提高物体的感测精度而对图像进行加工以使得物体的变形被缓和,因此无法产生对图像进行加工以使得特意使物体的变形变大这样的技术思想。

[0030] 为此,本发明人得到了如果生成物体的变形被表现得更大的图像,来作为学习用的图像,并使学习模型对该图像进行学习,则能够生成能够从这样的图像高精度地感测物体的学习模型的认识,从而想到了本公开的各方式。

[0031] (1)本公开的一方式中的图像加工方法是计算机中的图像加工方法,在所述图像加工方法中,获取由全方位图像构成的图像,执行从获取到的所述图像感测物体的物体感测处理,算出所述物体感测处理中的所述物体的感测精度,基于所述感测精度,对所述图像进行加工,以使得所述图像中包含的物体的变形变大,输出进行了所述加工的加工图像。

[0032] 根据该结构,算出对图像执行了物体感测处理的情况下的感测精度,基于所算出的感测精度,对图像进行加工,以使得物体的变形变大。由此,能够生成学习用的图像,该图像用于生成能够从物体的变形较大的图像高精度地感测物体的学习模型。

[0033] (2)在上述(1)所述的图像加工方法中,也可以所述图像是通过与物体的正解标签

建立了对应的全方位图像构成的图像,基于所述正解标签来算出所述感测精度,在所述感测精度低于阈值的情况下进行所述加工。

[0034] 根据该结构,由于对基于物体感测处理的物体感测困难的图像进行加工,因此能够提供更适合于能够从物体的变形较大的图像高精度地感测物体的学习模型的学习的图像。另外,由于基于正解标签来算出感测精度,因此感测精度的算出变得容易。

[0035] (3)在上述(1)或者(2)所述的图像加工方法中,也可以所述图像包含第1图像以及与所述第1图像不同的第2图像,所述感测精度是相对于将所述第1图像输入到为了执行所述物体感测处理而预先进行了学习的学习模型时的感测结果的感测精度,对所述第2图像执行所述图像的加工。

[0036] 根据该结构,能够生成对于全方位图像中的物体的感测精度较低的学习模型能够提高感测精度的学习用的图像。因此,能够高效地进行该学习模型的学习。

[0037] (4)在上述(3)所述的图像加工方法中,进一步地,也可以使用所述加工图像来使所述学习模型进行学习。

[0038] 根据该结构,由于学习模型使用物体的变形较大的图像来进行学习,因此能够生成能够从具有变形的图像高精度地感测物体的学习模型。

[0039] (5)在上述(2)~(4)的任一项所述的图像加工方法中,也可以所述感测精度按物体的分类来算出,所述第2图像是包含在所述第1图像中判定为所述感测精度为阈值以下的物体的图像。

[0040] 根据该结构,由于包含学习模型难以感测的物体的图像被生成为学习用图像,因此能够使学习模型进行学习,以使得该物体的感测精度提高。

[0041] (6)在上述(1)~(5)的任一项所述的图像加工方法中,也可以所述图像的加工包含将所述图像的默认的视点变更为随机设定的视点。

[0042] 根据该结构,由于通过随机变更视点来加工图像,因此在加工前在变形较小的位置上所显示的物体被显示于变形较大的位置的可能性提高,能够生成物体的变形更大的图像。

[0043] (7)在上述(1)~(5)的任一项所述的图像加工方法中,也可以所述图像的加工包含:对所述图像中所设定的多个正解标签之中距离为最长的两个绑定框的区间进行确定,并在所述区间的中点设定所述图像的视点。

[0044] 根据该结构,由于对图像进行加工以使得赋予有正解标签的物体被显示于图像的端部,因此能够生成物体的变形被表示得更大的图像。

[0045] (8)在上述(1)~(7)的任一项所述的图像加工方法中,也可以所述图像的加工包含:判定在所述图像中是否包含纵横比以及尺寸的至少一者超过基准值的物体;和在判定为包含超过所述基准值的物体的情况下,与判定为不包含该物体的情况相比,使所述加工图像的张数增多。

[0046] 在全方位图像中,纵向长的物体、横向长的物体以及尺寸较大的物体被变形地显示的可能性较高。根据该结构,由于在判定为图像中包含这样的物体的情况下,与不包含的情况相比,生成更多的加工图像,因此能够高效地生成能够提高这样的物体的感测精度的学习用的图像。

[0047] (9)在上述(1)或者(8)所述的图像加工方法中,也可以所述物体感测处理是基于

规则的物体感测处理,所述图像的加工包含对进行了物体感测处理的图像进行所述图像的加工。

[0048] 根据该结构,由于对通过基于规则的物体感测处理判定为物体的感测精度较低的图像进行加工,因此能够生成包含感测困难的物体的学习用的图像。

[0049] (10)在上述(1)~(9)的任一项所述的图像加工方法中,也可以所述图像的加工包含通过执行视点转换处理来对所述图像进行加工,以使得所述物体的变形变大,所述视点转换处理包含:将所述图像投影到单位球面上;根据进行了所述投影的投影图像设定新的视点;和将所述投影图像沿平面展开,以使得所述新的视点成为中心。

[0050] 根据该结构,由于图像被投影到单位球面上,在投影图像中设定有新的视点,因此新的视点的设定变得容易。

[0051] (11)本公开的另一方式中的图像加工装置具备处理器,所述处理器执行以下处理:获取由全方位图像构成的图像,执行从获取到的所述图像感测物体的物体感测处理,算出所述物体感测处理中的所述物体的感测精度,基于所述感测精度,对所述图像进行加工,以使得所述图像中包含的物体的变形变大,输出进行了所述加工的加工图像。

[0052] 根据该结构,能够提供一种图像加工装置,该图像加工装置能够生成用于从具有较大的变形的图像高精度地感测物体的学习用的图像。

[0053] (12)本公开的另一方式中的图像加工程序使计算机执行上述(1)~(10)的任一项所述的图像加工方法。

[0054] 根据该结构,能够提供一种图像加工程序,该图像加工程序能够生成用于从具有较大的变形的图像高精度地感测物体的学习用的图像。

[0055] 本公开也能够作为通过这样的信息处理程序而动作的信息处理系统来实现。另外,当然能够使这样的计算机程序经由CD-ROM等计算机可读取的非暂时的记录介质或者互联网等通信网络而流通。

[0056] 此外,以下所说明的实施方式均表示本公开的一具体例。以下的实施方式所示的数值、形状、结构要素、步骤、步骤的顺序等是一个例子,并不旨在限定本公开。另外,对于以下的实施方式中的结构要素之中、未记载于表示最上位概念的独立权利要求的结构要素,作为任意的结构要素来说明。另外,在全部的实施方式中,也能够组合各个内容。

[0057] (实施方式1)

[0058] 图1是示出本公开的实施方式中的图像加工装置1的结构的一个例子的框图。图像加工装置1是包含处理器10以及存储器20的计算机。处理器10例如是中央运算处理装置(CPU)。处理器10包含:获取部11、感测部12、验证部13、加工部14、输出部15以及学习部16。通过处理器10执行图像加工程序来实现获取部11~学习部16。存储器20通过固态驱动器(SSD)等非易失性的能够改写的存储装置来构成。存储器20包含验证图像数据库21、学习模型22以及学习图像数据库23。此外,在图1的例子中,全部的模块都集中于1个计算机,但是也可以分散配置于多个计算机。在该情况下,多个计算机通过互联网或者局域网等相互能够通信地进行连接。例如,学习部16也可以被安装于与图像加工装置1不同的装置,存储器20也可以被安装于与图像加工装置1不同的装置。

[0059] 获取部11从验证图像数据库21获取验证图像。验证图像是用于验证学习模型22的物体的感测精度的图像。在此,作为验证图像,采用全方位图像。验证图像与物体的正解标

签建立对应。验证图像是第1图像的一个例子。正解标签包含：表示验证图像中的物体的位置的边界框、和表示该物体所属的类别的分类标签。全方位图像是通过全方位摄像机拍摄到的图像。通常的摄像机只能拍摄到一定的角度的内侧，全方位摄像机能够对360度的方位即上下左右前后的全方位的图像进行拍摄。由于全方位图像是使用等距圆柱投影法等展开方法将通过全方位摄像机拍摄到的图像展开而得到的图像，因此成为变形根据位置而不同的图像。因此，在使用仅使用了通过通常的摄像机拍摄到的图像进行了学习的学习模型，来从全方位图像感测物体的情况下，物体的感测精度降低的可能性提高。

[0060] 感测部12执行从通过获取部11获取到的验证图像感测物体的物体感测处理。详细地，感测部12通过向学习模型22输入验证图像，得到感测结果，来执行物体感测处理。学习模型22是为了执行物体感测处理而预先进行了机器学习的模型。学习模型22只要是例如深度神经网络、卷积神经网络等从图像感测物体的模型，则可以采用任何模型。通过对赋予了物体的正解标签的学习图像的数据集进行机器学习来生成学习模型22。

[0061] 验证部13基于与验证图像建立了对应的正解标签，算出学习模型22中的物体的感测精度，并判定所算出的感测精度是否低于阈值。例如通过将用于验证的验证图像中包含的物体的总数作为分母、并且将物体感测成功的物体的数量作为分子的比例即正解率来定义感测精度。

[0062] 如果从学习模型22作为感测结果而输出的物体的分类标签与正解标签中包含的分类标签一致，则验证部13判定为物体的感测成功即可。或者，验证部13也可以在从学习模型22作为感测结果输出的物体的分类标签与正解标签中包含的分类标签一致、且该物体的可靠度超过基准可靠度的情况下，判定为物体的感测成功。

[0063] 此外，在从学习模型22输出的感测结果中包含按物体的分类的可靠度的情况下，验证部13按分类判定可靠度是否超过了基准可靠度即可。并且，在全部分类中可靠度都超过了基准可靠度的情况下，验证部13判定为物体感测成功即可。所谓分类，是指物体的类型。

[0064] 加工部14基于感测精度，对图像进行加工，以使得图像中包含的物体的变形变大。详细地，在通过验证部13算出的感测精度低于阈值的情况下，加工部14对学习图像进行加工，以使得验证图像中包含的物体的变形变大。更详细地，加工部14通过从学习图像数据库23获取学习图像，并执行将获取到的学习图像的默认的视点变更为随机设定的视点的视点转换处理，来对学习图像进行加工即可。学习图像与验证图像同样是全方位图像，事先与物体的正解标签建立对应。学习图像是第2图像的一个例子。因此，在通过对学习图像进行加工而得到的加工图像中也继承有正解标签。由此，从原来的学习图像应用了视点转换处理而得到的加工图像被生成为学习图像。所谓默认的视点，是指作为初始值而设定的视点，例如是与全方位摄像机的水平面平行且与朝北的方向对应的点。视点位于全方位图像的中心。

[0065] 输出部15将通过加工部14进行了加工的加工图像存储在学习图像数据库23。

[0066] 学习部16使用学习图像数据库23中所存储的加工图像，使学习模型22进行学习。学习部16基于加工图像中所赋予的正解标签和从学习模型22输出的可靠度，算出学习误差，并对学习模型22的参数进行更新，以使得学习误差被最小化。作为参数的更新方法，能够采用误差反向传播法。参数包含权重值以及偏置值等。

[0067] 验证图像数据库21对验证图像进行存储。学习模型22是成为验证对象的学习模型22。学习图像数据库23对学习图像进行存储。

[0068] 图2是示出执行视点转换处理的情形图。图像G10是视点转换处理前的全方位图像,图像G20是视点转换处理后的全方位图像。在该例子中,图像G20相对于图像G10,视点A1被沿水平方向变更180度。在图像G10、G20中,视点A1位于图像的中心。由于图像G10、G20是全方位图像,因此变形根据位置而不同。例如,可知左右的端部以及上下的端部的区域与中央部相比,变形变大。在图像G10和图像G20中,物体F1是相同的物体。可知在图像G10中位于水平方向的中央的物体F1在图像G20中移动至图像的端,变形增大。这样,通过应用视点转换处理,位于当前图像的中央的物体变成位于图像的端,该物体的变形增大。

[0069] 图3是视点转换处理的说明图。图像G30是全方位图像,通过等距圆柱投影法的坐标系来表现。等距圆柱投影法的坐标系(平面的一个例子)是水平方向以u轴表示且垂直方向以v轴表示的二维的坐标系。图像G30的水平方向的尺寸为2h,垂直方向的尺寸为h。

[0070] 首先,加工部14将图像G30中的点Q转换到半径为1的极坐标系。在该情况下,点Q(u,v)通过式(1)来表示。

$$[0071] \quad \theta = \pi u/h, \quad \varphi = \pi v/h \quad \cdots (1)$$

[0072] θ 为天顶角, φ 为偏角。

[0073] 接下来,加工部14将点Q从极坐标系投影到三维正交坐标系。在该情况下,点Q(x,y,z)通过式(2)来表示。

$$[0074] \quad x = \sin\theta \cdot \cos\varphi, \quad y = \sin\theta \cdot \sin\varphi, \quad z = \cos\varphi \quad \cdots (2)$$

[0075] 接下来,加工部14设定横摆(yaw)/俯仰(pitch)/翻滚(roll)这3个轴各自的旋转矩阵Y(ψ_y)、P(θ_p)、R(φ_r)。Y(ψ_y)是围绕横摆轴的旋转角度, θ_p 是围绕俯仰轴的旋转角度, φ_r 是围绕翻滚轴的旋转角度。由此,如式(3)所示,点Q(x,y,z)被投影到点Q'(x',y',z')。

[0076] [数学式1]

$$[0077] \quad \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = Y(\psi_y)P(\theta_p)R(\varphi_r) \begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad \cdots (3)$$

$$R(\varphi_r) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi_r & -\sin\varphi_r \\ 0 & \sin\varphi_r & \cos\varphi_r \end{bmatrix}, P(\theta_p) = \begin{bmatrix} \cos\theta_p & 0 & \sin\theta_p \\ 0 & 1 & 0 \\ -\sin\theta_p & 0 & \cos\theta_p \end{bmatrix}, Y(\psi_y) = \begin{bmatrix} \cos\psi_y & -\sin\psi_y & 0 \\ \sin\psi_y & \cos\psi_y & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

[0078] 接下来,加工部14使用式(4)将点Q'从正交坐标系转换到极坐标系。 θ' 是转换后的天顶角, φ' 是转换后的偏角。

[0079] [数学式2]

$$[0080] \quad \theta' = \cos^{-1} z', \quad \varphi' = \tan^{-1} \left(\frac{y'}{x'} \right) \quad \cdots (4)$$

[0081] 接下来,加工部14将点Q'从极坐标系转换到等距圆柱投影法的坐标系。在该情况下,点Q'通过式(5)来表示。

$$[0082] \quad u' = \theta' h/\pi, \quad v' = \varphi' h/\pi \quad \cdots (5)$$

[0083] 在图像G30的所有点进行以上的处理,来进行针对图像G30的视点转换处理。 u' 是视点转换后的u轴的坐标值, v' 是视点转换后的v轴的坐标值。

[0084] 在实施方式1中,加工部14通过随机设定上述的旋转角度 φ_r 、 θ_p 、 ψ_y ,来对图像G30的视点进行随机转换。详细地,加工部14将以旋转角度 φ_r 、 θ_p 、 ψ_y 进行了旋转而得到的正距

圆筒坐标系的图像G30的中心设定为视点。此外,在后述的实施方式中,加工部14使用不是随机的而是与实施方式相应的方法来执行视点转换处理。

[0085] 接下来,对通过学习模型22进行了物体感测处理的图像的应用例进行说明。图4是示出应用通过学习模型22进行了物体感测处理的图像的用户接口的显示画面G1的图。

[0086] 显示画面G1是用于远程的用户对作业现场的状况进行确认的应用的基本画面。显示画面G1包含图像显示栏R1、注释信息显示栏R2以及设计图显示栏R3。在设计图显示栏R3,显示有作业现场的设计图,在该设计图中,重叠显示有选择图标201、拍摄地点图标202以及轨迹203。在该作业现场,事先通过作业人员进行了使用全方位摄像机的拍摄作业,拍摄地点图标202表示在该拍摄作业中拍摄到的图像的拍摄地点。轨迹203表示拍摄作业中的作业人员的移动的轨迹。

[0087] 用户输入通过在设计图上拖放选择图标201来选择1个拍摄地点图标202的操作。于是,在所选择的1个拍摄地点图标202所示的拍摄地点拍摄到的作业现场的全方位图像被显示于图像显示栏R1。用户对图像显示栏R1中所显示的图像设定注释区域D1,并将针对该注释区域D1的注释消息输入到注释信息显示栏R2。由此,注释区域D1以及注释消息由用户彼此共享。其结果,远程的用户即使不移动到作业现场,也能够详细地确认作业现场的新增状况以及注意事项等。

[0088] 图像显示栏R1中显示的全方位图像通过学习模型22事先进行了物体感测处理。因此,当用户输入对在该全方位图像中想要赋予注释的物体进行选择的操作时,显示该物体的边界框,并且能够基于该边界框来设定注释区域D1。由此,用户即使不输入使图像显示栏R1显示用于对注释区域D1进行设定的框体,使该框体移动到作为目的的物体的位置,并使框体变形以使得匹配物体的形状的操作,也能够对注释区域D1进行设定。

[0089] 图5是示出实施方式1中的图像加工装置1的处理的一个例子的流程图。首先,在步骤S1中,获取部11从验证图像数据库21获取包含给定张数的验证图像在内的验证图像的数据集。

[0090] 接下来,在步骤S2中,感测部12通过将构成验证图像的数据集的各验证图像依次输入到学习模型22,来感测验证图像中包含的物体。

[0091] 接下来,在步骤S3中,验证部13针对由步骤S1获取到的感测图像的数据集,通过将学习模型22中的物体的感测结果与正解标签进行比较的处理,算出上述的正解率,并将所算出的正解率算出为学习模型22的感测精度。

[0092] 接下来,验证部13判定由步骤S3算出的感测精度是否为阈值以下(步骤S4)。在判定为感测精度为阈值以下的情况下(步骤S4中为是),加工部14从学习图像数据库23获取包含给定张数的学习图像在内的学习图像的数据集(步骤S5)。

[0093] 接下来,加工部14针对各学习图像,随机设定视点(步骤S6)。详细地,如上所述,通过随机设定旋转角度 φ_r 、 θ_p 、 ψ_y ,视点被随机设定。

[0094] 接下来,加工部14通过对各学习图像执行视点转换处理,来生成默认的视点被变更为所设定的视点而得到的加工图像(步骤S7)。生成的加工图像被存储在学习图像数据库23。在此,加工部14也可以通过对1张学习图像随机设定K(K为2以上的整数)个视点,来生成K张加工图像。由此,根据1张学习图像,生成了以多种多样的变形来表示物体的多张加工图像。其结果,能够高效地生成适合于学习模型22的学习的加工图像。

[0095] 接下来,对图像加工装置1中的机器学习进行说明。图6是示出图像加工装置1中的学习阶段中的处理的一个例子的流程图。

[0096] 首先,在步骤S21中,学习部16从学习图像数据库23获取包含给定张数的加工图像在内的加工图像的数据集。

[0097] 接下来,在步骤S22中,学习部16通过将加工图像的数据集依次输入到学习模型22,使学习模型22进行学习。

[0098] 接下来,在步骤S23中,学习部16通过针对由步骤S22获取到的全部加工图像,对学习模型22中的物体的感测结果和加工图像中所赋予的正解标签进行比较,算出物体感测的正解率,并将所算出的正解率算出为学习模型22的感测精度。学习部16中的感测精度的算出方法与验证部13中的方法相同。即,学习部16将由步骤S5获取到的学习图像的数据集的全部张数作为分母、并且将物体感测成功的学习图像的张数作为分子的比例,算出为感测精度。

[0099] 接下来,在步骤S24中,学习部16判定感测精度是否为阈值以上。作为阈值,能够采用0.8、0.9等适当的值。在感测精度为阈值以上的情况下(步骤S24中为是),结束处理。另一方面,在感测精度小于阈值的情况下(步骤S24中为否),处理返回到步骤S21。在该情况下,学习部16再次从学习图像数据库23获取加工图像的数据集,并执行学习模型22的学习即可。在此,所使用的加工图像的数据集可以包含与上一次的循环中用于学习的加工图像相同的加工图像,也可以不包含。

[0100] 由此,由于直到感测精度成为阈值以上为止,在学习模型22中执行使用加工图像的学习,因此能够生成能够从全方位图像高精度地感测物体的学习模型22。

[0101] 这样,根据本实施方式,基于正解标签,算出从验证图像感测到物体的学习模型22的感测精度,在所算出的感测精度为阈值以下的情况下,对学习图像进行加工,以使得物体的变形变大。由此,能够生成学习用的图像,该学习用的图像用于生成能够从全方位图像高精度地感测物体的学习模型。

[0102] 另外,在本实施方式中,由于基于随机设定的视点来加工图像,因此在加工前在变形较小的位置上所显示的物体被显示于变形较大的位置的可能性提高,能够生成物体的变形更大的图像。

[0103] (实施方式2)

[0104] 实施方式2中,将多个正解标签的区间之中最长的两个正解标签的区间的中点设定为视点。此外,在实施方式2中,对于与实施方式1相同的结构要素赋予相同的符号,并省略说明。另外,实施方式2中的框图使用图1来说明。

[0105] 图1所示的加工部14使用上述的式(1)、式(2),将视点转换处理前的全方位图像(以下,称为原图像。)和与原图像建立了对应的边界框转换到单位球面上。向单位球面上的转换通过上述的式(1)、式(2)来进行。接下来,加工部14确定在单位球面上描绘的多个边界框之中区间为最长的两个边界框。在此,如图3所示,将表示单位球面上的某两个边界框的位置的两点设为P、Q。作为边界框的位置,能够采用边界框的重心。所谓区间,是指在穿过点P、Q的大圆301中被点P、Q划分的两个圆弧之中较长的圆弧。接下来,加工部14确定区间为最长的两个边界框,并将这两个边界框的区间的中点设定为视点。接下来,加工部14将单位球面上的原图像展开,以使得该视点位于正距圆筒坐标系的中心。由此,与区间为最长的两个

边界框对应的物体在全方位图像中位于变形较大的端,结果可得到物体的变形变大的全方位图像。

[0106] 图7是示出应用实施方式2中的视点转换处理的图像G40的图。在图像G40中,窗户、椅子、浴缸、灯、镜子、门等分类标签与边界框建立对应。在该例中,在原图像中,椅子的边界框E1的位置B1与门的边界框E2的位置B2的区间L被判定为最长。因此,原图像被展开为区间L的中点M1成为视点,得到了图像G40。由此,椅子以及门被显示于变形较大的图像G40的两端的位置,物体的变形被增大。

[0107] 图8是示出实施方式2中的图像加工装置1的处理的一个例子的流程图。步骤S31~S35的处理与图5的步骤S1~S5的处理相同。在步骤S36中,加工部14从与学习图像建立了对应的多个边界框之中的任意两个边界框的多个区间,确定最长的区间。

[0108] 接下来,在步骤S37中,加工部14将区间的中点设定为视点。

[0109] 接下来,在步骤S38中,将学习图像展开,以使得所设定的视点位于中心,由此生成加工图像。由此,可得到与视点转换处理前相比物体的变形变大的加工图像。

[0110] 这样,根据实施方式2,由于对学习图像进行加工,以使得赋予有正解标签的物体被显示于变形较大的位置,因此能够生成将物体的变形被表示得更大的加工图像。

[0111] (实施方式3)

[0112] 实施方式3中,生成加工图像,以使得包含全方位图像中具有易变形的形状的物体的加工图像被生成得更多。此外,在实施方式3中,对于与实施方式1、2相同的结构要素赋予相同的符号,并省略说明。另外,实施方式3中的框图使用图1来说明。

[0113] 图9是示出具有易变形的形状的物体的一个例子的图。具有易变形的形状的物体相当于纵横比超过基准纵横比的物体、尺寸超过基准尺寸的物体。图像G91、G92所示的物体是由纵横比超过基准纵横比的建筑材料构成的物体。图像G93是由尺寸超过基准尺寸以上的建筑材料构成的物体。具有易变形的形状的物体的一个例子为横向长的沙发、浴缸、顶灯、门扇等。纵横比包含对物体赋予的边界框的横边相对于纵边的比例、纵边相对于横边的比例。

[0114] 图10是示出实施方式3中的图像加工装置1的处理的一个例子的流程图。步骤S41~S44与图5的S1~S4相同,因此省略说明。在步骤S45中,加工部14从学习图像数据库23获取学习图像。

[0115] 接下来,在步骤S46中,加工部14算出学习图像中包含的物体的尺寸以及纵横比。例如,加工部14根据与学习图像建立了对应的边界框的面积来算出物体的尺寸。加工部14根据与学习图像建立了对应的边界框的纵边和横边的长度来算出纵横比。

[0116] 接下来,在步骤S47中,加工部14判定在学习图像中是否包含尺寸超过基准尺寸以上或者纵横比为基准纵横比以上的物体。在学习图像中包含相应的物体的情况下(步骤S47中为是),加工部14对学习图像随机设定N(N为2以上的整数)个视点(步骤S48)。加工部14使用实施方式1的方法来设定N个视点即可。N的一个例子为2。

[0117] 接下来,加工部14生成与N个视点对应的N张加工图像(步骤S49)。加工部14执行视点转换处理,以使得默认的视点被变更为所设定的N个视点,由此来生成N张加工图像即可。

[0118] 在学习图像中不包含尺寸为基准尺寸以上且纵横比为基准纵横比以上的物体的情况下(步骤S47中为否),加工部14对学习图像随机设定M(M为1以上且比N小的整数)个视

点(步骤S50)。M的一个例子为1。随机设定视点的方法与实施方式1相同。

[0119] 接下来,在步骤S51中,加工部14生成与M个视点对应的M张加工图像。加工部14执行视点转换处理,以使得默认的视点被变更为所设定的M个视点,由此来生成M张加工图像即可。

[0120] 接下来,加工部14判定是否从学习图像数据库23已获取到给定张数的学习图像(步骤S52)。在已获取给定张数的学习图像的情况下(步骤S52中为是),处理结束。另一方面,在未获取到给定张数的学习图像的情况下(步骤S52中为否),处理返回到步骤S45,从学习图像数据库23获取下一加工对象的学习图像。

[0121] 这样,根据实施方式3,在判定为在学习图像中包含纵向长的物体、横向长的物体、尺寸较大的物体这样具有易变形的形状的物体的情况下,由于与判定为不包含的情况相比,生成了更多的加工图像,因此能够高效地生成能够提高这样的物体的感测精度的学习用的图像。

[0122] (实施方式4)

[0123] 实施方式4中更多地生成包含在学习模型22中难以感测的物体的加工图像。此外,在实施方式4中,对于与实施方式1~3相同的结构要素赋予相同的符号,并省略说明。另外,实施方式4中的框图使用图1来说明。图11是示出实施方式4中的图像加工装置1的处理的一个例子的流程图。步骤S71、S72的处理与图5的步骤S1、S2相同,因此省略说明。在步骤S73中,验证部13按物体的分类,算出各验证图像中的物体的感测精度。例如,在作为感测对象的物体的分类而具有沙发、顶灯、门这种的分类的情况下,算出沙发、顶灯、门各自的感测精度。

[0124] 接下来,在步骤S74中,验证部13判定是否存在属于感测精度为阈值以下的分类的物体。以下,将属于感测精度为阈值以下的分类的物体称为特定物体。在判定为存在特定物体的情况下(步骤S74中为是),加工部14从学习图像数据库23获取包含给定张数的学习图像在内的学习图像的数据集,该给定张数的学习图像包含特定物体(步骤S75)。另一方面,在判定为没有特定物体的情况下(步骤S74中为否),处理结束。

[0125] 接下来,在步骤S76中,加工部14对学习图像设定视点。例如,加工部14可以如实施方式1所示随机设定视点,也可以如实施方式2所示将最长区间的中点设定为视点。

[0126] 接下来,在步骤S77中,加工部14对各学习图像应用视点转换处理,以使得默认的视点成为所设定的视点,由此生成加工图像(步骤S77)。例如,加工部14通过应用实施方式1或者实施方式2中示出的视点转换处理来生成加工图像即可。

[0127] 这样,根据实施方式4,由于生成包含学习模型难以感测的物体的学习图像,因此能够使学习模型进行学习,以使得该物体的感测精度提高。

[0128] (实施方式5)

[0129] 实施方式5中使用基于规则的物体感测处理对全方位图像进行物体感测处理,并对进行了物体感测处理的全方位图像进行加工。此外,在实施方式5中,对于与实施方式1~4相同的结构要素赋予相同的符号,并省略说明。

[0130] 图12是示出实施方式5中的图像加工装置1A的结构的一个例子的框图。图12与图1的不同点在于,取代验证图像数据库21而将候选图像数据库31存储在存储器20这点,以及处理器10取代感测部12、验证部13以及加工部14而具有感测部12A、验证部13A以及加工部

14A这点。

[0131] 候选图像数据库31存储成为学习模型22的学习候选的候选图像。候选图像与验证图像同样,是与正解标签建立了对应的全方位图像。

[0132] 感测部12A对通过获取部11获取到的候选图像应用基于规则的物体感测处理,由此从候选图像感测物体。所谓基于规则的物体感测处理,相当于不使用通过机器学习得到的学习模型地从图像感测物体的处理。基于规则的物体感测处理的一个例子为模式匹配、根据边缘检测出的图像中包含的边缘的形状来感测物体的处理等。此外,成为感测对象的物体的分类被预先决定。因此,模式匹配中使用的模板为与成为感测对象的物体的分类对应的模板。感测部12A通过将按分类的模板应用于候选图像,按分类来算出相似度。

[0133] 验证部13A将通过感测部12A算出的相似度算出为物体感测处理的感测精度,并判定感测精度是否低于阈值。此外,在全部的分类的相似度低于阈值的情况下,验证部13A将感测精度判定为低于阈值即可。

[0134] 在判定为通过验证部13A算出的感测精度低于阈值的情况下,加工部14A对候选图像进行加工,以使得候选图像中包含的物体的变形变大。

[0135] 输出部15将通过加工部14A进行了加工的加工图像存储在学习图像数据库23。由此,能够使学习模型22学习通过对候选图像进行加工而得到的加工图像。

[0136] 图13是示出实施方式5中的图像加工装置1A的处理的一个例子的流程图。首先,在步骤S101中,获取部11从候选图像数据库31获取候选图像的数据集。

[0137] 接下来,感测部12A对获取到的候选图像的数据集中包含的各候选图像应用基于规则的物体感测处理,由此从候选图像感测物体(步骤S102)。

[0138] 接下来,在步骤S103中,验证部13A将感测部12A对物体进行感测时所算出的相似度算出为感测精度。

[0139] 接下来,在步骤S104中,验证部13A判定感测精度是否为阈值以下。在感测精度为阈值以下的情况下(步骤S104中为是),加工部14A对候选图像设定视点(步骤S105)。例如,加工部14A可以如实施方式1所示随机设定视点,也可以如实施方式2所示将最长区间的中点设定为视点。在感测精度比阈值大的情况下(步骤S104中为否),处理结束。

[0140] 接下来,在步骤S106中,加工部14A对候选图像应用视点转换处理,以使得默认的视点成为所设定的视点,由此生成加工图像(步骤S106)。例如,加工部14A通过应用实施方式1或者实施方式2所示的视点转换处理来生成加工图像即可。加工图像被存储在学习图像数据库23。

[0141] 这样,根据实施方式5,由于对根据基于规则的物体感测处理判定为物体的感测精度较低的候选图像进行加工,因此能够生成包含该物体的学习用的加工图像。

[0142] 本公开能够采用以下的变形例。

[0143] (1) 在实施方式1~4中,加工对象的全方位图像是学习图像数据库23中所存储的学习图像,但也可以是验证图像。

[0144] (2) 实施方式4中所示的从学习图像数据库23获取包含特定物体的学习图像的方式也可以应用于实施方式1~3。

[0145] (3) 在上述的实施方式中,作为现场例示了建造现场,但是本公开并不限于此,制造现场、物流现场、流通现场、农地、土木现场、零售现场、办公室、医院、商业设施、护理设

施等也可以被用作现场。

[0146] 产业上的可利用性

[0147] 根据本公开,在从全方位图像进行物体感测的技术领域中有用。

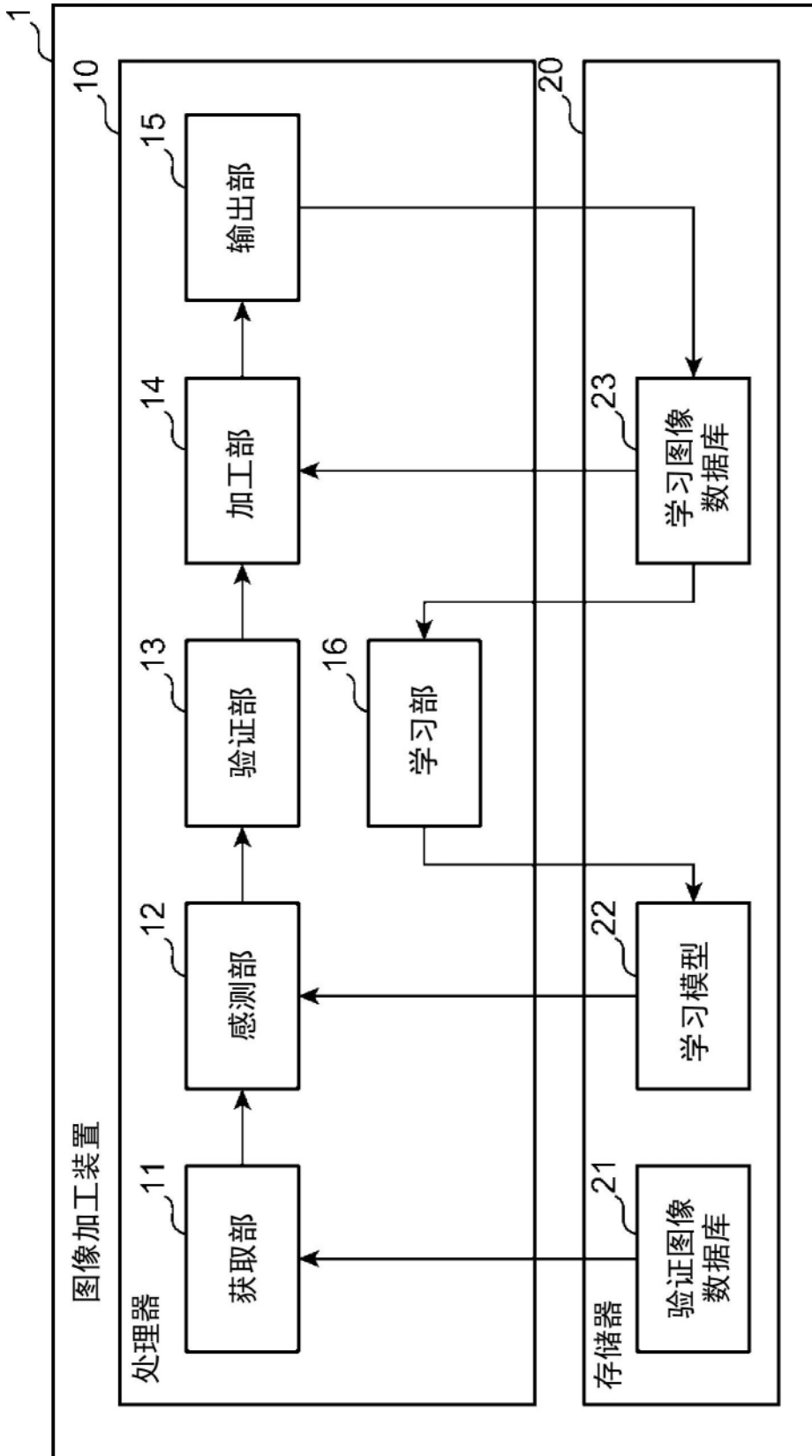


图1

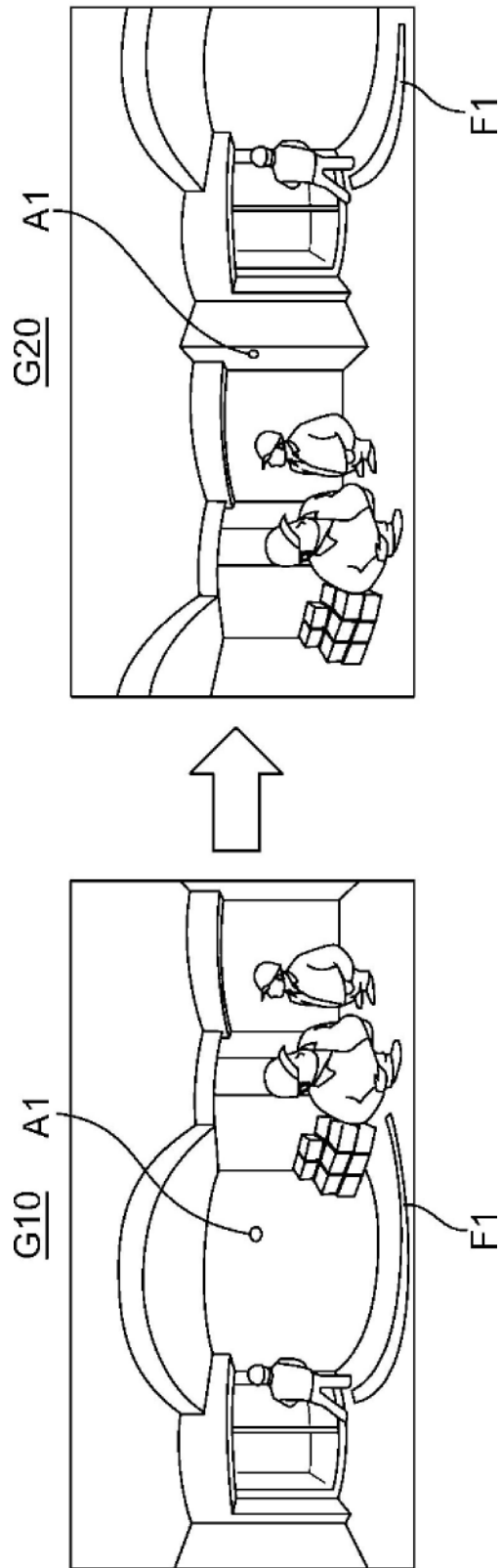


图2

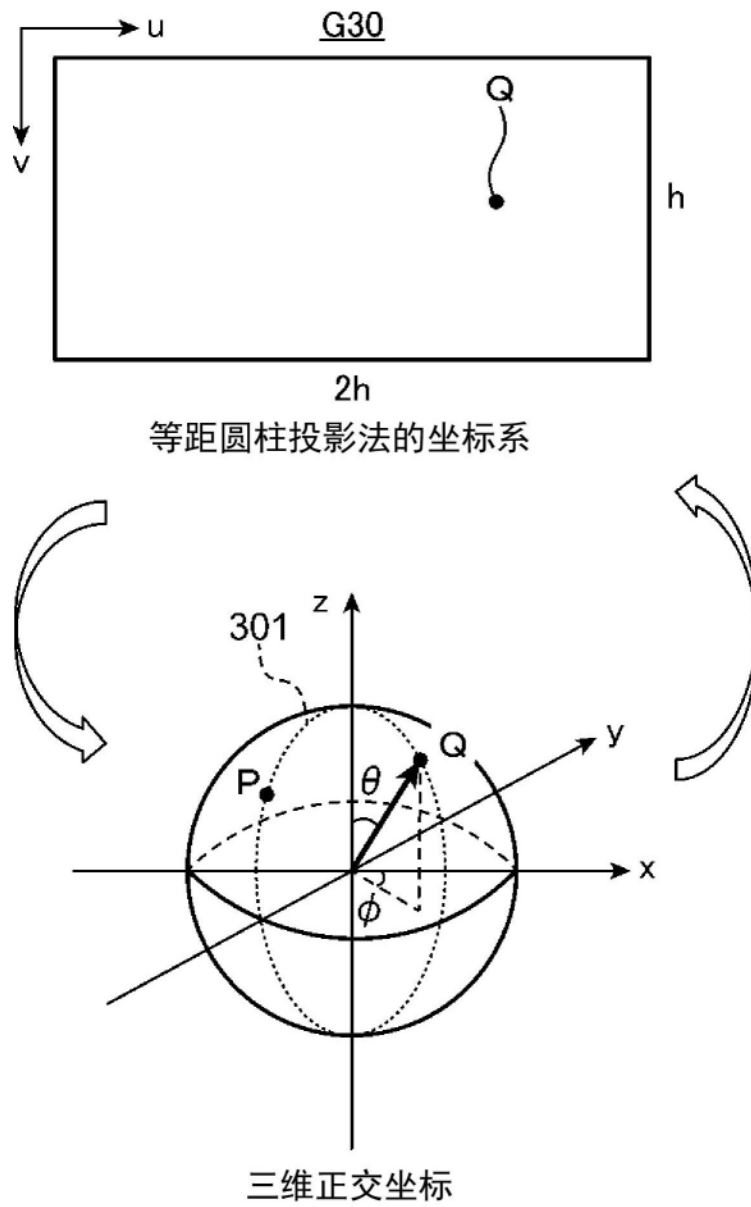


图3

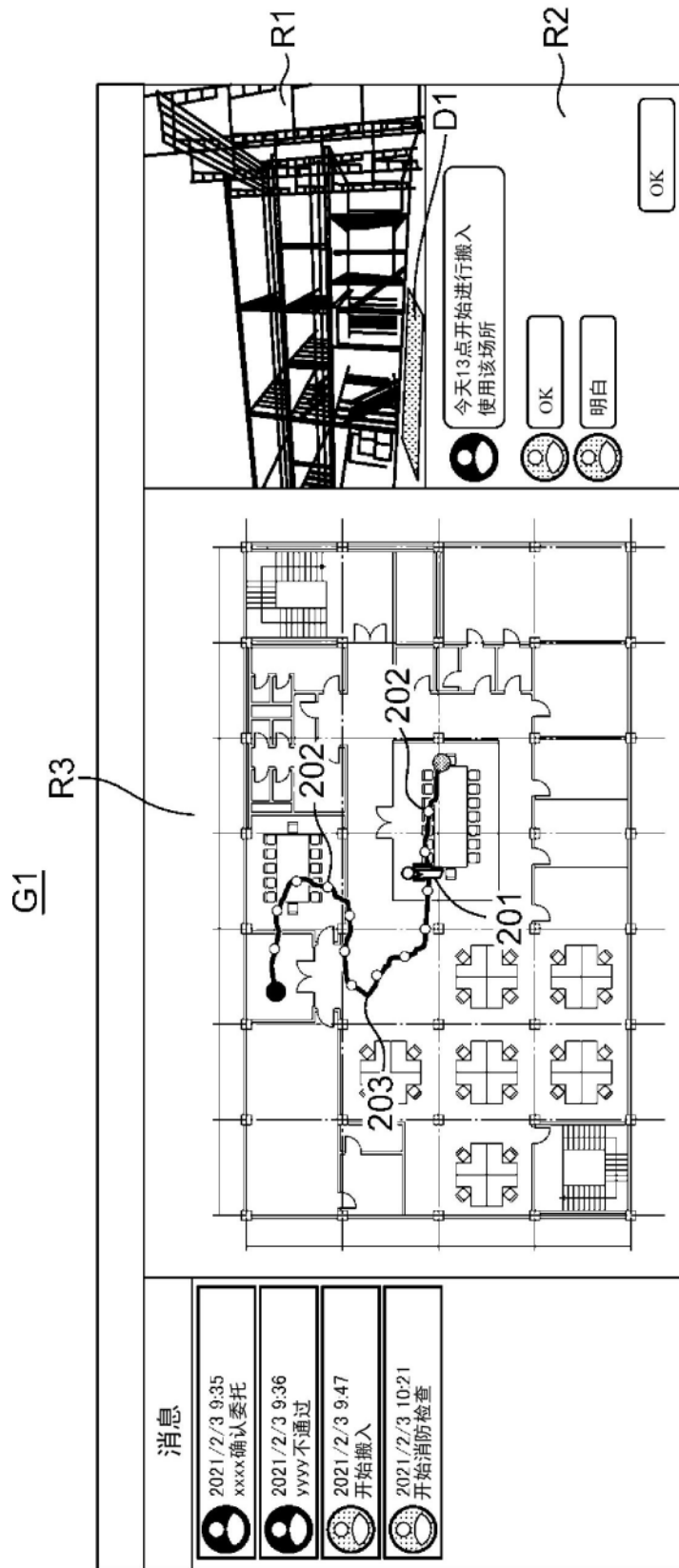


图4

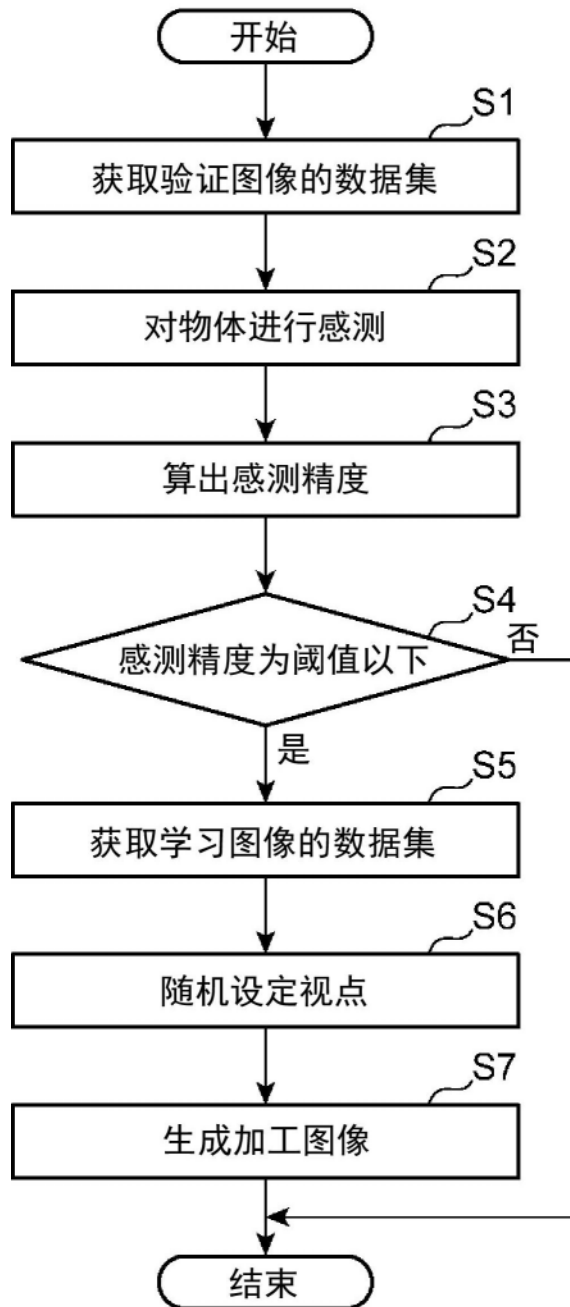


图5

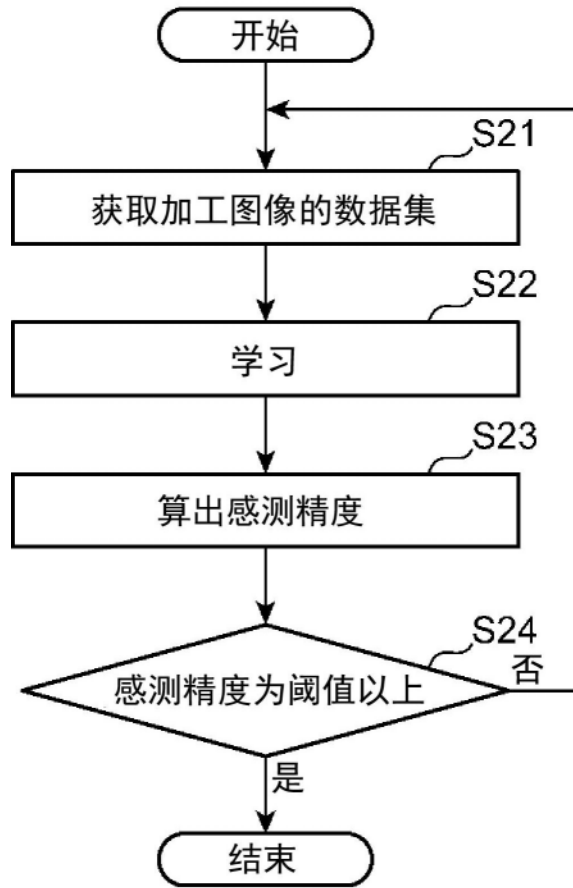


图6

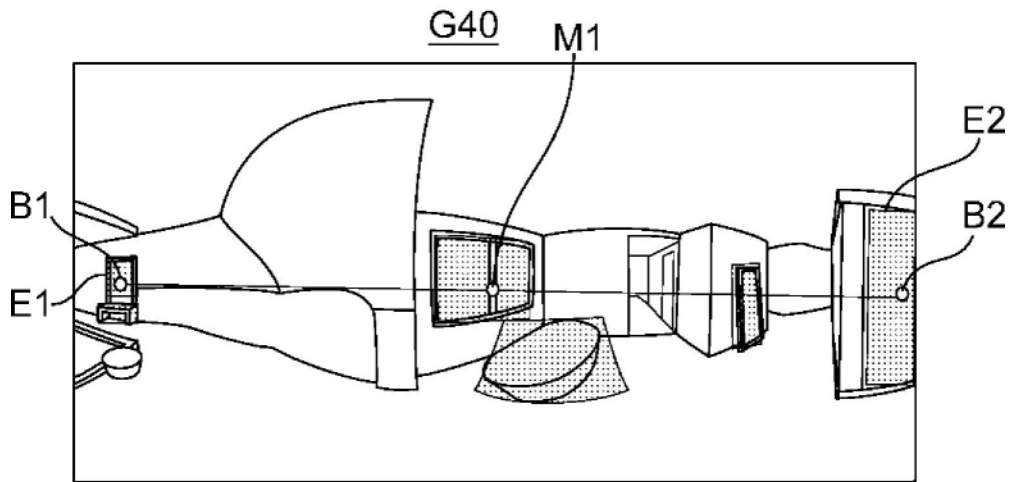


图7

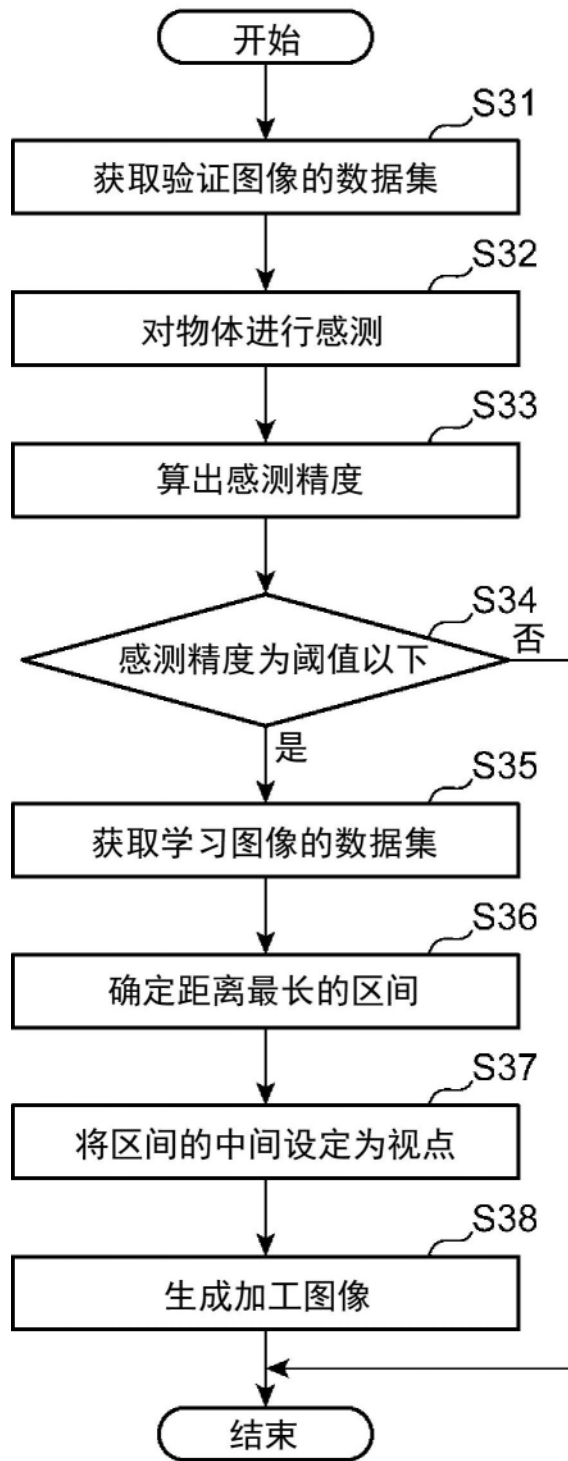


图8

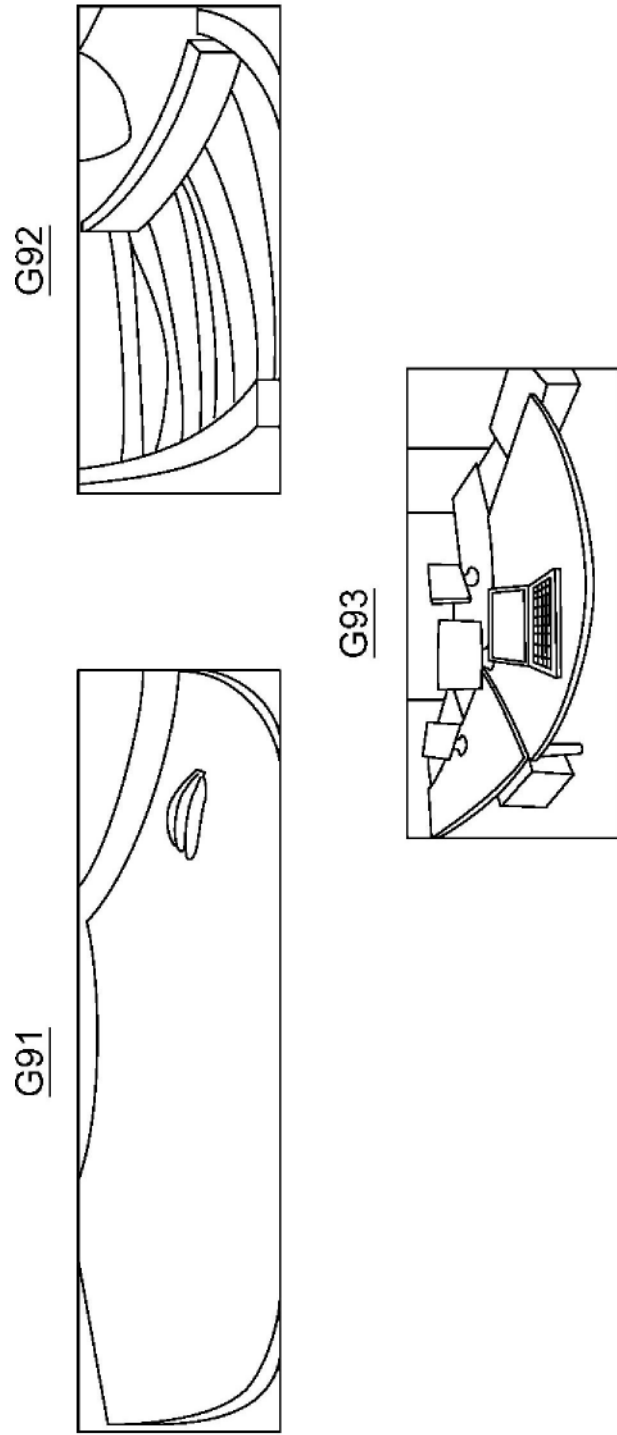


图9

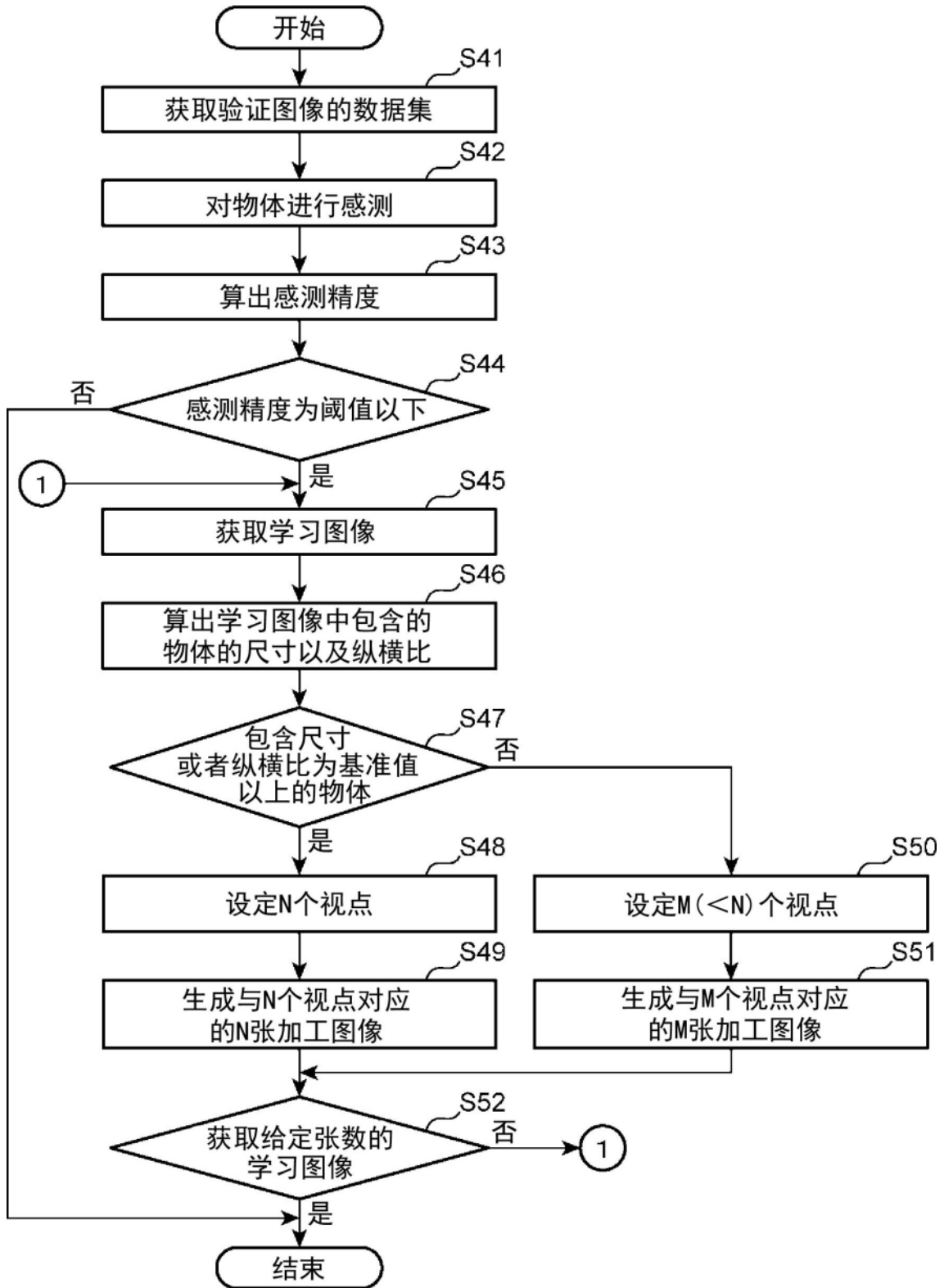


图10

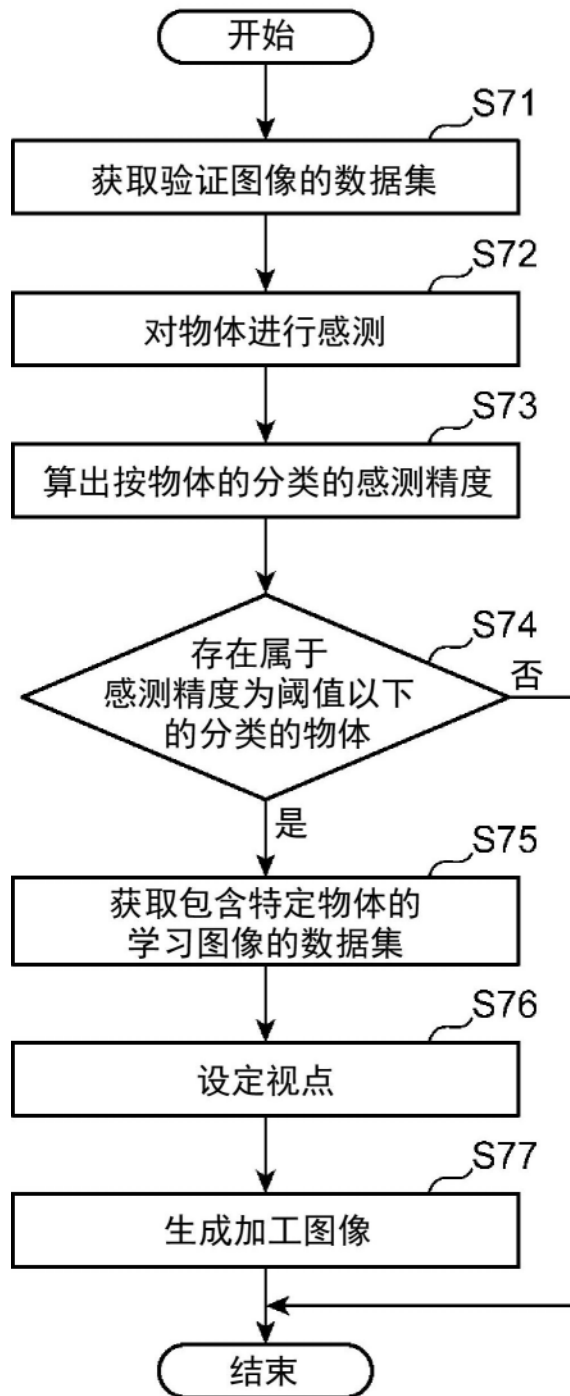


图11

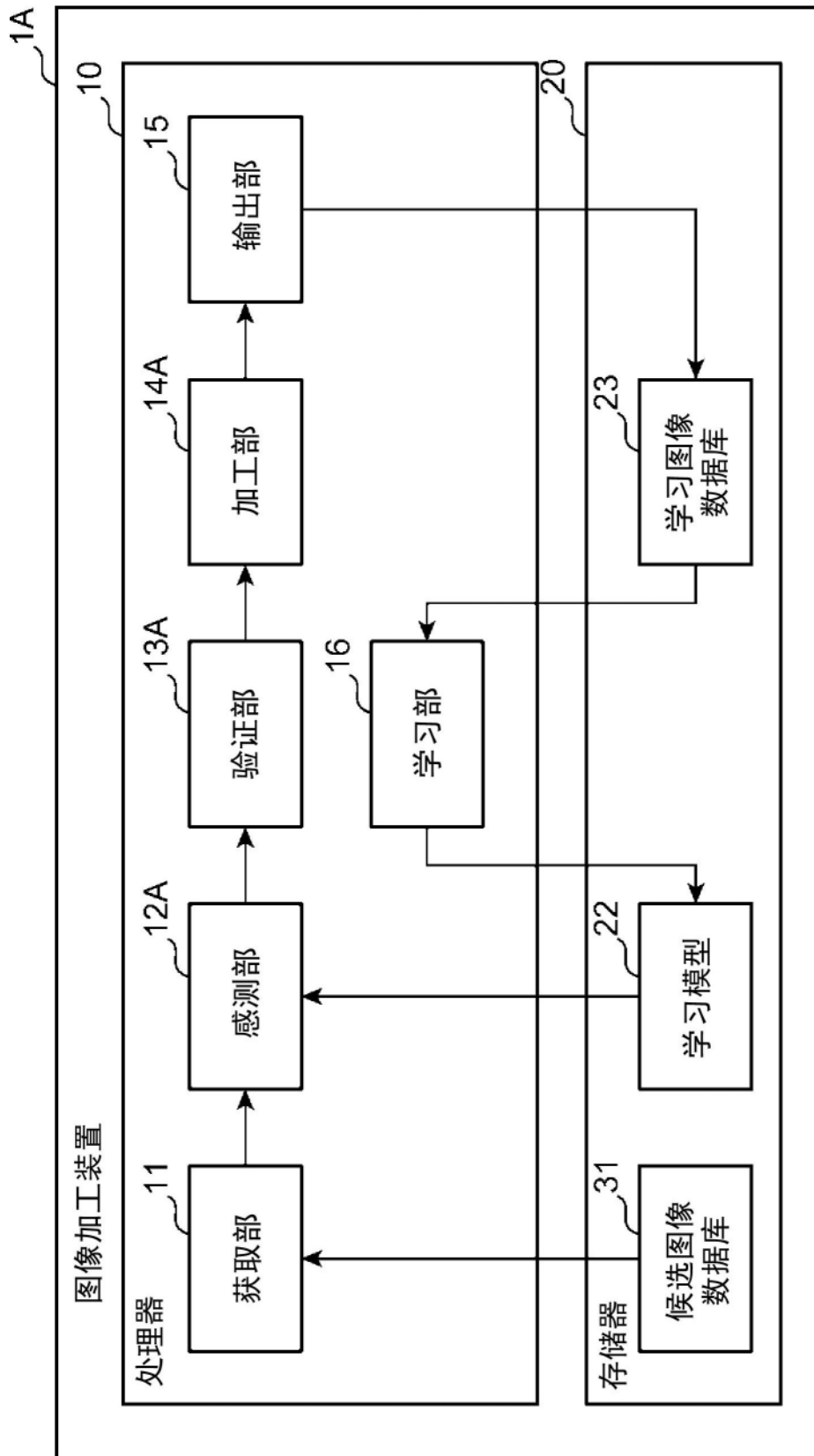


图12

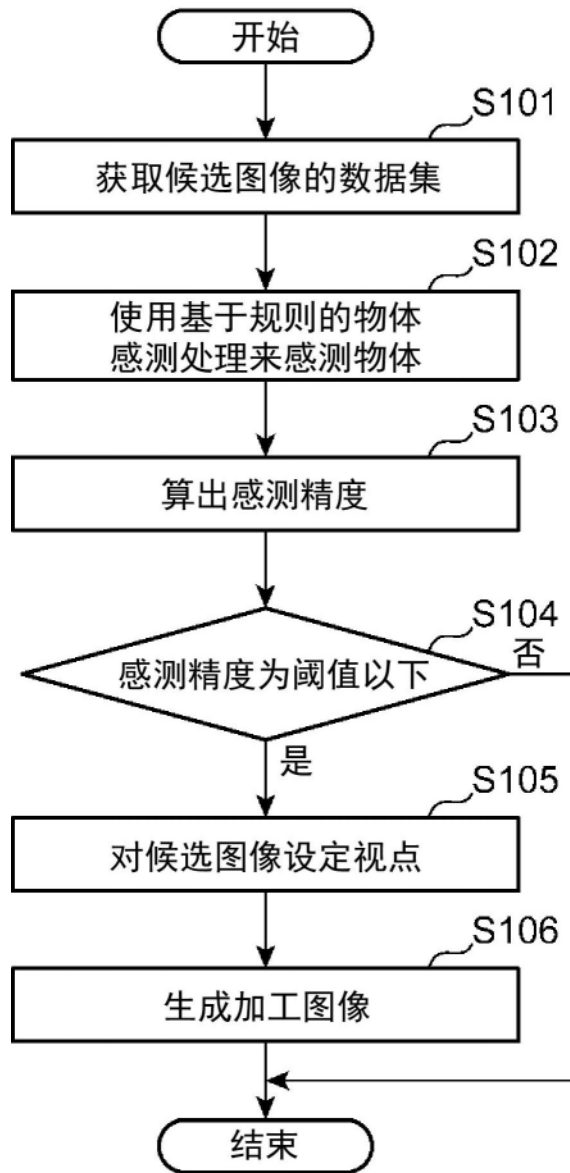


图13