



(12) 发明专利申请

(10) 申请公布号 CN 119731511 A

(43) 申请公布日 2025. 03. 28

(21) 申请号 202380058940.3

(22) 申请日 2023.09.25

(30) 优先权数据

2022-153456 2022.09.27 JP

(85) PCT国际申请进入国家阶段日

2025.02.10

(86) PCT国际申请的申请数据

PCT/JP2023/034642 2023.09.25

(87) PCT国际申请的公布数据

W02024/071006 JA 2024.04.04

(71) 申请人 本田技研工业株式会社

地址 日本

(72) 发明人 阿曼·迦因

阿尼鲁德·雷迪·康达帕利

山田健太郎

(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 孙尚昆

(51) Int.Cl.

G01C 21/26 (2006.01)

G06F 3/01 (2006.01)

G06F 3/16 (2006.01)

G06T 7/20 (2017.01)

G08G 1/00 (2006.01)

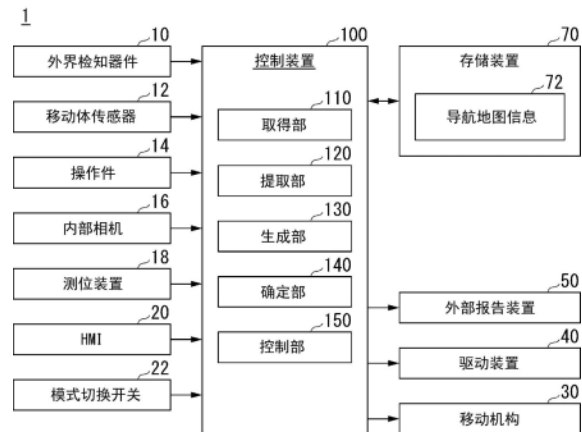
权利要求书2页 说明书9页 附图8页

(54) 发明名称

信息处理装置、信息处理方法以及程序

(57) 摘要

一种信息处理装置,其中,所述信息处理装置具备:取得部,其取得由搭载于移动体的相机拍摄所述移动体的周边而得到的拍摄图像、由与所述移动体相关的人输入的输入指示文以及由所述人进行了的手势相关的手势信息;提取部,其通过对所述输入指示文实施第一规定处理,而提取所述输入指示文所包含的一个以上的指示;第一生成部,其通过对所述手势信息实施第二规定处理,而生成与所述人所指示的位置相关的推定分布;第二生成部,其根据所述拍摄图像,生成对所述拍摄图像所包含的各物体赋予了概率的概率场景图;以及确定部,其基于所述一个以上的指示、所述推定分布以及所述概率场景图,确定所述拍摄图像中的所述人所指示的物体。



1. 一种信息处理装置, 其中,
所述信息处理装置具备:

取得部, 其取得由搭载于移动体的相机拍摄所述移动体的周边而得到的拍摄图像、由与所述移动体相关的人输入的输入指示文以及与由所述人进行了的手势相关的手势信息;

提取部, 其通过对所述输入指示文实施第一规定处理, 而提取所述输入指示文所包含的一个以上的指示;

第一生成部, 其通过对所述手势信息实施第二规定处理, 而生成与所述人所指示的位置相关的推定分布;

第二生成部, 其根据所述拍摄图像, 生成对所述拍摄图像所包含的各物体赋予了概率的概率场景图; 以及

确定部, 其基于所述一个以上的指示、所述推定分布以及所述概率场景图, 确定所述拍摄图像中的所述人所指示的物体。

2. 根据权利要求1所述的信息处理装置, 其中,

所述第一规定处理是对所述输入指示文至少实施依存解析(dependency parser)以及属性分类的处理。

3. 根据权利要求1所述的信息处理装置, 其中,

所述第二规定处理是基于所述手势信息所包含的所述人的关键点而生成所述推定分布的处理。

4. 根据权利要求1至3中任一项所述的信息处理装置, 其中,

所述确定部通过使用提取出的所述一个以上的指示逐次地更新所述概率场景图所包含的各物体的所述概率, 而确定所述物体。

5. 根据权利要求4所述的信息处理装置, 其中,

所述第二生成部基于所述推定分布, 设定所述概率场景图所包含的对各物体赋予的所述概率的初始值。

6. 根据权利要求4所述的信息处理装置, 其中,

所述确定部在通过所述更新而确定了多个所述物体的情况下, 生成用于确定多个所述物体中的一个物体的提问。

7. 一种信息处理方法, 其中,

所述信息处理方法使计算机进行如下处理:

取得由搭载于移动体的相机拍摄所述移动体的周边而得到的拍摄图像、由与所述移动体相关的人输入的输入指示文以及与由所述人进行了的手势相关的手势信息;

通过对所述输入指示文实施第一规定处理, 而提取所述输入指示文所包含的一个以上的指示;

通过对所述手势信息实施第二规定处理, 而生成与所述人所指示的位置相关的推定分布;

根据所述拍摄图像, 生成对所述拍摄图像所包含的各物体赋予了概率的概率场景图; 以及

基于所述一个以上的指示、所述推定分布以及所述概率场景图, 确定所述拍摄图像中的所述人所指示的物体。

8.一种程序,其中,

所述程序使计算机进行如下处理:

取得由搭载于移动体的相机拍摄所述移动体的周边而得到的拍摄图像、由与所述移动体相关的人输入的输入指示文以及与由所述人进行了的手势相关的手势信息;

通过对所述输入指示文实施第一规定处理,而提取所述输入指示文所包含的一个以上的指示;

通过对所述手势信息实施第二规定处理,而生成与所述人所指示的位置相关的推定分布;

根据所述拍摄图像,生成对所述拍摄图像所包含的各物体赋予了概率的概率场景图;以及

基于所述一个以上的指示、所述推定分布以及所述概率场景图,确定所述拍摄图像中的所述人所指示的物体。

信息处理装置、信息处理方法以及程序

技术领域

[0001] 本发明涉及信息处理装置、信息处理方法以及程序。

背景技术

[0002] 以往,已知有根据包含图像的多模态的输入而确定该图像中的物体的技术。例如,在专利文献1中记载了通过将音频数据以及图像数据输入深度神经网络(DNN)等人工神经网络而检测该图像数据所包含的紧急车辆的技术。

[0003] 现有技术文献

[0004] 专利文献

[0005] 专利文献1:日本特开2022-96601号公报

[0006] 非专利文献

[0007] 非专利文献1:德鲁·A.哈德逊,克里斯托弗 D.曼宁,“通过抽象学习:神经状态机”,NeurIPS,第5901-5914页。(2019) (Drew A. Hudson,Christopher D. Manning,“Learning by Abstraction: The Neural State Machine”, NeurIPS , page 5901-5914. (2019))

发明内容

[0008] 发明要解决的课题

[0009] 然而,在如专利文献1所记载的技术那样例如使用DNN等方法处理多模态的输入的情况下,在模型的内部进行着何种处理为黑箱,难以掌握其实际状态。即,在这些方法中,是否恰当学习了不同的种类的信息(影像、声音、文本等)之间的关系性并不明确。并且,DNN等方法大多在学习需要大量的数据,存在在功能的实现上耗费非常大的成本的情况。

[0010] 并且,在非专利文献1中记载了将由传感器得到的多模态的原始信息变换为共通的抽象空间中的概念表现,在将这些概念表现构造化了的图表(场景图:scene graph)上进行推论的神经状态机(neural state machine)。神经状态机在将图表上的处理内容透明化并且在其处理中无需大量的数据这点相对于专利文献1所记载的技术有优势。然而,非专利文献1所记载的技术未活用人物的视线、手势等模态,另外并未提供解决在推论上产生的含糊的方案。

[0011] 本发明是考虑这样的情况而完成的,目的之一在于提供能够活用人物的视线、手势等模态并且解决在推论的过程中引发的含糊的信息处理装置、信息处理方法以及程序。

[0012] 用于解决课题的方案

[0013] 本发明的信息处理装置、信息处理方法以及程序采用了以下的结构。

[0014] (1):本发明的一方案的信息处理装置具备:取得部,其取得由搭载于移动体的相机拍摄所述移动体的周边而得到的拍摄图像、由与所述移动体相关的人输入的输入指示文以及与由所述人进行了的手势相关的手势信息;提取部,其通过对所述输入指示文实施第一规定处理,而提取所述输入指示文所包含的一个以上的指示;第一生成部,其通过对所述

手势信息实施第二规定处理,而生成与所述人所指示的位置相关的推定分布;第二生成部,其根据所述拍摄图像,生成对所述拍摄图像所包含的各物体赋予了概率的概率场景图;以及确定部,其基于所述一个以上的指示、所述推定分布以及所述概率场景图,确定所述拍摄图像中的所述人所指示的物体。

[0015] (2):在上述(1)的方案的基础上,所述第一规定处理是对所述输入指示文至少实施依存解析(dependency parser)以及属性分类的处理。

[0016] (3):在上述(1)的方案的基础上,所述第二规定处理是基于所述手势信息所包含的所述人的关键点而生成所述推定分布的处理。

[0017] (4):在上述(1)至(3)中任一方案的基础上,所述确定部通过使用提取出的所述一个以上的指示逐次地更新所述概率场景图所包含的各物体的所述概率,而确定所述物体。

[0018] (5):在上述(4)的方案的基础上,所述第二生成部基于所述推定分布,设定所述概率场景图所包含的对各物体赋予的所述概率的初始值。

[0019] (6):在上述(4)的方案的基础上,所述确定部在通过所述更新而确定了多个所述物体的情况下,生成用于确定多个所述物体中的一个物体的提问。

[0020] (7):本发明的其他方案的信息处理方法使计算机进行如下处理:取得由搭载于移动体的相机拍摄所述移动体的周边而得到的拍摄图像、由与所述移动体相关的人输入的输入指示文以及与由所述人进行了的手势相关的手势信息;通过对所述输入指示文实施第一规定处理,而提取所述输入指示文所包含的一个以上的指示;通过对所述手势信息实施第二规定处理,而生成与所述人所指示的位置相关的推定分布;根据所述拍摄图像,生成对所述拍摄图像所包含的各物体赋予了概率的概率场景图;以及基于所述一个以上的指示、所述推定分布以及所述概率场景图,确定所述拍摄图像中的所述人所指示的物体。

[0021] (8):本发明的其他方案的程序使计算机进行如下处理:取得由搭载于移动体的相机拍摄所述移动体的周边而得到的拍摄图像、由与所述移动体相关的人输入的输入指示文以及与由所述人进行了的手势相关的手势信息;通过对所述输入指示文实施第一规定处理,而提取所述输入指示文所包含的一个以上的指示;通过对所述手势信息实施第二规定处理,而生成与所述人所指示的位置相关的推定分布;根据所述拍摄图像,生成对所述拍摄图像所包含的各物体赋予了概率的概率场景图;以及基于所述一个以上的指示、所述推定分布以及所述概率场景图,确定所述拍摄图像中的所述人所指示的物体。

[0022] 发明效果

[0023] 根据(1)~(8)的方案,能够活用人物的视线、手势等模态并且解决在推论的过程中引发的含糊。

附图说明

[0024] 图1是示出实施方式的移动体1以及控制装置100的结构的一例的图。

[0025] 图2是从上方观察移动体1而得到的透视图。

[0026] 图3是示出由外部相机拍摄到的拍摄图像IM的一例的图。

[0027] 图4是用于说明由提取部120执行的第一规定处理的概要的图。

[0028] 图5是用于说明由生成部130执行的第二规定处理的概要的图。

[0029] 图6是用于说明由确定部140执行的概率场景图的更新处理的图。

[0030] 图7是用于说明由确定部140执行的提问文的生成处理的图。

[0031] 图8是示出由控制装置100执行的处理的流程的一例的流程图。

具体实施方式

[0032] 以下,参照附图对本发明的信息处理装置、信息处理方法以及程序的实施方式进行说明。信息处理装置搭载于移动体。移动体在车道以及与车道不同的规定区域这双方移动。移动体有时被称作微型移动体。电动滑板车是微型移动体的一种。规定区域例如是人行道。另外,规定区域可以是路侧方地带、自行车行车道、公开空地等中的一部分或全部,也可以包括人行道、路侧方地带、自行车行车道、公开空地等的全部。如以下说明的那样,本实施方式的信息处理装置基于拍摄移动体的周边而得到的拍摄图像、由与移动体相关的人输入的输入指示文以及由该人进行了的手势,而确定人所指示的物体。在以下的说明中,与移动体相关的人作为搭乘于该移动体的乘员而进行说明,但本发明并不限于那样的结构,也可以是在移动体的外部输入(例如,为了在下车后指示移动体的待机场所)指示文的人。

[0033] [整体结构]

[0034] 图1是示出实施方式的移动体1以及控制装置100的结构的一例的图。在移动体1例如搭载外界检知器件10、移动体传感器12、操作件14、内部相机16、测位装置18、HMI20、模式切换开关22、移动机构30、驱动装置40、外部报告装置50、存储装置70以及控制装置100。需要说明的是,也可以省略这些结构中的不是实现本发明的功能所必需的一部分的结构。移动体并不限于交通工具,可以包括与步行的用户并行行驶而搬运货物、或者给人带路那样的小型移动体,另外,也可以包括其他能够自主移动的移动体(例如步行型机器人等)。

[0035] 外界检知器件10是将移动体1的行进方向作为检知范围的各种器件。外界检知器件10包括外部相机、雷达装置、LIDAR(Light Detection and Ranging)、传感器融合装置等。外界检知器件10将表示检知结果的信息(图像、物体的位置等)向控制装置100输出。尤其是,在本实施方式中,外界检知器件10将由外部相机拍摄移动体1的周边而得到的拍摄图像向控制装置100输出。

[0036] 移动体传感器12例如包括速度传感器、加速度传感器、横摆角速度(角速度)传感器、方位传感器以及安装于操作件14的操作量检测传感器等。操作件14例如包括用于指示加减速的操作件(例如油门踏板、制动踏板)以及用于指示转向的操作件(例如转向盘)。在该情况下,移动体传感器12也包括油门开度传感器、制动踩踏量传感器、转向转矩传感器等。移动体1也可以作为操作件14而具备上述以外的形态的操作件(例如,不为圆环状的旋转操作件、操纵杆、按钮等)。

[0037] 内部相机16从正面拍摄移动体1的乘员的至少头部。内部相机16是利用了CCD(Charge Coupled Device)、CMOS(Complementary Metal Oxide Semiconductor)等摄像元件的数码相机。内部相机16将拍摄到的图像向控制装置100输出。

[0038] 测位装置18是对移动体1的位置进行测定的装置。测位装置18例如是GNSS(Global Navigation Satellite System)接收机,并基于从GNSS卫星接收到的信号,对移动体1的位置进行确定,并作为位置信息而输出。需要说明的是,移动体1的位置信息也可以根据后述的通信装置所连接着的Wi-Fi基站的位置来推定。

[0039] HMI20包括显示装置、扬声器、触摸面板、按键等。移动体1的乘员例如经由HMI20而

设定移动体1的目的地,后述的控制部150使移动体1行驶到所设定的目的地。尤其是,在本实施方式中,HMI20包括话筒等声音输入设备,移动体1的乘员通过讲出对移动体1的停车位位置进行指示的指示文而向声音输入设备输入。HMI20对所输入的指示文的声音进行解析而文本化,并将其向控制装置100输出。也可以代替性地,HMI20例如经由触摸面板而接受乘员作为文本输入的指示文,并将接受到的指示文向控制装置100输出。

[0040] 模式切换开关22是由乘员操作的开关。模式切换开关22可以是机械式开关,也可以是在HMI20的触摸面板上设定的GUI(Graphical User Interface)开关。模式切换开关22例如接受将驾驶模式切换为模式A、模式B以及模式C中的任意模式的操作,模式A是由乘员进行转向操作和加减速控制中的一方,且另一方自动地进行的辅助模式(可以存在由乘员进行转向操作且自动地进行加减速控制的模式A-1以及由乘员进行加减速操作且自动地进行转向控制的模式A-2),模式B是由乘员进行转向操作以及加减速操作的手动驾驶模式,模式C是自动地进行操作控制以及加减速控制的自动驾驶模式。

[0041] 移动机构30是用于在道路中使移动体1移动的机构。移动机构30例如是包括转向轮以及驱动轮的车轮组。另外,移动机构30也可以是用于多足步行的腿部。

[0042] 驱动装置40向移动机构30输出力而使移动体1移动。例如,驱动装置40包括对驱动轮进行驱动的马达、蓄积向马达供给的电力的蓄电池、调整转向轮的转向角的转向装置等。驱动装置40也可以具备内燃机、燃料电池等来作为驱动力输出机构或者发电机构。另外,驱动装置40也可以还具备基于摩擦力、空气阻力的制动装置。

[0043] 外部报告装置50例如是设置于移动体1的外板部且用于朝向移动体1的外部报告信息的灯、显示器装置、扬声器等。外部报告装置50在移动体1在人行道移动着的状态与在车道移动着的状态下进行不同的动作。例如,外部报告装置50被控制为在移动体1在人行道移动着的情况下使灯发光、在移动体1在车道移动着的情况下不使灯发光。该灯的发光色适宜为由法规规定的颜色。外部报告装置50也可以被控制为在移动体1在人行道移动着的情况下使灯以绿色发光、在移动体1在车道移动着的情况下使灯以蓝色发光。在外部报告装置50为显示器装置的情况下,外部报告装置50在移动体1在人行道行驶着的情况下以文本、图形来显示“在人行道行驶中”这一意旨。

[0044] 图2是从上方观察移动体1而得到的透视图。在图中,FW是转向轮,RW是驱动轮,SD是转向装置,MT是马达,BT是蓄电池。转向装置SD、马达MT、蓄电池BT包含于驱动装置40。另外,AP是油门踏板,BP是制动踏板,WH是转向盘,SP是扬声器,MC是话筒。图示的移动体1是一人乘坐的移动体,乘员P就座于驾驶员座DS并佩戴座椅安全带SB。箭头D1是移动体1的行进方向(速度向量)。外界检知器件10设置于移动体1的前端部附近,内部相机16设置于从乘员P的前方能够拍摄乘员P的头部的位置,模式切换开关22设置于转向盘WH的轮毂部。另外,在移动体1的前端部附近设置有作为显示器装置的外部报告装置50。

[0045] 返回图1,存储装置70例如是HDD(Hard Disk Drive)、闪存器、RAM(Random Access Memory)等非暂时性的存储装置。在存储装置70保存导航地图信息72等。在图中将存储装置70记载于控制装置100的框外,但存储装置70也可以包含于控制装置100。另外,存储装置70也可以设置于未图示的服务器上。

[0046] 导航地图信息72事先存储于存储装置70,例如是包含包括车道以及人行道的道路的中央的信息或者道路的边界的信息等的地图信息。导航地图信息72还包含与道路的边界

相接的设施、建筑物所相关的信息(名称、住所、面积等)。

[0047] [控制装置]

[0048] 控制装置100例如具备取得部110、提取部120、生成部130、确定部140以及控制部150。取得部110、提取部120、生成部130、确定部140以及控制部150例如通过CPU (Central Processing Unit) 等硬件处理器执行程序(软件)74而实现。这些构成要素中的一部分或全部可以通过LSI (Large Scale Integration)、ASIC (Application Specific Integrated Circuit)、FPGA (Field-Programmable Gate Array)、GPU (Graphics Processing Unit) 等硬件(包括电路部:circuitry)而实现,也可以通过软件与硬件的协同配合而实现。程序可以预先保存于存储装置70,也可以保存于DVD、CD-ROM等能够装卸的存储介质(非暂时性的存储介质)并通过存储介质装配于驱动装置而安装于存储装置70。将取得部110、提取部120、生成部130以及确定部140组合起来是“信息处理装置”的一例。

[0049] 取得部110取得作为外界检知器件10的外部相机拍摄移动体1的周边而得到的拍摄图像IM。图3是示出由外部相机拍摄到的拍摄图像IM的一例的图。图3作为一例,而表示外部相机拍摄到的拍摄图像IM包括车辆M1、M2、自动贩卖机V1、V2、V3以及邮筒P的状况。为了便于说明,在拍摄图像IM中,自动贩卖机V1设为红色的自动贩卖机,自动贩卖机V2、V3设为蓝色的自动贩卖机。

[0050] 并且,取得部110取得移动体1的乘员经由作为HMI20的声音输入设备而输入的对移动体1到达的目标位置进行指示的指示文。例如,在图3的拍摄图像IM所示的状况的情况下,乘员通过输入“那个卡车的后方处存在的红色的自动贩卖机处停止吧”等指示文,而对到达自动贩卖机V1前的目标位置TP进行指示。

[0051] 并且,取得部110将由内部相机16拍摄到的表示由乘员进行了的手势的图像作为手势信息而取得。所取得的手势信息被用于后述的由生成部130进行的处理。

[0052] 提取部120通过对输入指示文实施第一规定处理,而提取该输入指示文所包含的一个以上的指示(推理指令:reasoning instructions)。图4是用于说明由提取部120执行的第一规定处理的概要的图。更具体而言,提取部120作为第一规定处理,对输入指示文实施依存解析(dependency parser)以及实体分类(entity classifier)。例如,如图4的左部所示那样,提取部120作为依存解析而解析出指示文“那个卡车的后方处存在的红色的自动贩卖机处停止吧”中的“那个”为修饰“卡车”的限定词(det)、“的”为“卡车”的格标识(case)、“后方”为修饰“卡车”的名词修饰语(nmod)、“红色的”为修饰“自动贩卖机”的从句修饰语(ac1)等。该依存解析也可以使用已知的方法来执行。

[0053] 同时,提取部120作为实体分类而对指示文中的各词素的属性进行分类。例如,如图4的右部所示那样,提取部120将指示文“那个卡车的后方处存在的红色的自动贩卖机处停止吧”中的“那个”分类为指示代词(Demonstrative),将“卡车”分类为物体(Object),将“后方”分类为关系(Relation),将“红色的”分类为颜色(Color),将“自动贩卖机”分类为物体(Object)。提取部120将作为依存解析的结果的词素间的修饰关系与作为实体分类的结果的各词素的属性建立关联而存储。图4的左上部表示各词素(卡车、那个、后方、自动贩卖机、红色的)将其修饰关系与属性建立关联而作为一个以上的指示(推理指令:reasoning instructions)被存储。更详细而言,上述一个以上的指示也可以使用非专利文献1所记载的方法导出并存储。

[0054] 生成部130通过对由取得部110取得的手势信息实施第二规定处理,而生成与乘员利用手势指示了的位置相关的推定分布。图5是用于说明由生成部130执行的第二规定处理的概要的图。首先,生成部130在乘员P的身体的部位中的两个部位设定关键点。图5作为一例,表示乘员P的眼和手腕分别设定为关键点KP1、KP2的状况。生成部130将连结眼KP1与手腕KP2并向手腕方向延长了的指示线L与地面交叉的交点IS推定为乘员P利用手势指示了的位置,并作为以交点IS为最大值的概率分布而生成手势位置的推定分布。作为所生成的概率分布,例如也可以假定正态分布等任意的种类的分布。

[0055] 需要说明的是,关于生成部130,关键点可以任意地选择,但由于乘员P能够通过视线准确地指定位置,因此一方的点期望设为眼。另外,另一方的点期望为容易从图像确定的部位,例如除了手腕以外,也可以是指尖、握拳的前端或中心部等。并且,在操作者指示着目的地时,也有可能由于脸部朝向目的地的方向,而无法从内部相机16拍摄眼。在这样的情况下,也可以对眼的位置进行推定来进行确定。若能够确定脸部的方向则眼的位置能够推定。需要说明的是,该眼的位置的推定也可以使用机器学习模型来进行。

[0056] 并且,生成部130从由外部相机拍摄到的拍摄图像IM含义性地(语义性地)提取该拍摄图像IM所包含的物体,并生成对提取出的各物体赋予了乘员P指示该物体的概率的概率场景图。例如,在图3的情况下,生成部130提取车辆M1、M2、自动贩卖机V1、V2、V3以及邮筒P。通过该提取处理,与深度神经网络(DNN)等对原始数据进行处理的方法相比,能够减轻后续的处理所承担的负荷。更详细而言,概率场景图的生成也可以使用非专利文献1所记载的方法而执行。

[0057] 生成的概率场景图所包含的对各物体赋予的概率的初始值可以一样(即,1/(概率场景图所包含的物体的个数)),生成部130也可以针对物体设定不同的初始值。例如,生成部130也可以使用与手势位置相关的推定分布,根据各物体的位置而设定不同的初始值。更具体而言,生成部130也可以越接近检知到的交点IS则设定越高的初始值,另一方面越远离交点IS则设定越低的初始值。例如,在图3的情况下,在乘员P朝向自动贩卖机V1进行了手势的情况下,对自动贩卖机V1赋予最高的初始值,另一方面对位于距自动贩卖机V1最远的位置的自动贩卖机V3赋予最低的初始值。生成部130是“第一生成部”以及“第二生成部”的一例。

[0058] 确定部140通过使用由提取部120提取出的一个以上的指示逐次地更新概率场景图所包含的各物体的概率,而确定乘员所指示的物体。图6是用于说明由确定部140执行的概率场景图的更新处理的图。确定部140依次取出由提取部120提取出的一个以上的指示,并以与所取出的指示对应的概率场景图的物体的概率变得更高的方式进行更新。例如,在图6的情况下,确定部140提取“卡车”以及“那个”,并以车辆M1以及车辆M2的概率值变得更高的方式进行更新。接下来,确定部140提取“后方”,并在概率场景图中从车辆M1向自动贩卖机V1转移并且从车辆M2向自动贩卖机V2转移。接下来,确定部140提取“自动贩卖机”以及“红色的”,将自动贩卖机V1确定为具有属性“红色的”的自动贩卖机,并以自动贩卖机V1的概率值变得更高的方式进行更新。这样,逐次地更新对概率场景图的物体赋予的概率,并将最终概率值成为最大的物体确定为由乘员P指示了的物体。在图6的情况下,确定部140能够将最终概率值变得最大的自动贩卖机V1确定为由乘员P指示了的物体。更详细而言,这些概率场景图的更新也可以使用非专利文献1所记载的方法来执行。

[0059] 这样,确定部140使用概率场景图将最终概率值成为最大的物体确定为由乘员P指示了的物体,但根据情况,有时概率值成为最大的物体存在多个、或者概率值成为最大的物体与概率值第二大的物体的概率值的差量较小而无法达成单一的物体的确定。更一般而言,确定部140算出对概率场景图的各物体算出的概率分布的熵,并在算出的熵较大(为阈值以上)的情况下,能够判定为无法将由乘员P指示了的物体确定为单一的物体。在那样的情况下,在以往技术中,最终无法确定由乘员P指示了的物体。另一方面,本发明在概率场景图的更新处理的结果是存在多个由乘员P指示了的物体的候补的情况下,生成用于从多个候补确定一个物体的提问文,询问乘员,并接收来自乘员的回答,从而最终确定由乘员P指示了的物体。

[0060] 图7是用于说明由确定部140执行的提问文的生成处理的图。图7作为一例而表示如下情况:乘员作为指示文而输入为“卡车的后方处存在的自动贩卖机处停止吧”,确定部140进行概率场景图的更新处理的结果是,车辆M1的后方处存在的自动贩卖机V1以及车辆M2的后方处存在的自动贩卖机V2被确定为具有相同的概率值的物体(或概率值的差量为阈值以内的物体)。

[0061] 在这样的情况下,确定部140生成用于对确定出的多个物体中的一个物体进行确定的提问。例如,也可以生成如“为哪个自动贩卖机?”那样用于对多个物体直接进行确定的提问文,也可以生成如“为哪个卡车?”那样对多个物体间接进行确定(即,在卡车确定为一个的情况下,与名词修饰语“后方”组合,能够确定自动贩卖机)的提问文。

[0062] 另外,例如,确定部140也可以对照成为候补的多个物体的属性(例如,颜色),生成与具有不同的值的属性相关的提问文。在图7的情况下,自动贩卖机V1具有“红”的颜色属性,另一方面,自动贩卖机V2具有特定的颜色属性,因此确定部140也可以基于该颜色属性的不同,而生成“自动贩卖机的颜色为红?”等提问。

[0063] 确定部140将所生成的提问文向HMI20发送,并在HMI20上接收乘员所输入的回答,基于接收到的回答,最终确定由乘员P指示了的物体。另外,例如,确定部140也可以再次接受由乘员P进行的手势,并将在接受到的手势的方向上最接近的物体最终确定为由乘员P指示了的物体。由此,即使在通过概率场景图的更新处理而指示物体的候补存在多个的情况下,也能够通过对乘员P生成提问文,而唯一地确定指示物体。

[0064] 控制部150通过将由确定部140确定出的物体作为目标位置,并驱动移动体1的驱动装置40,而使移动体1行驶到该目标位置。

[0065] 接下来,参照图8对由本实施方式的控制装置100执行的处理的流程进行说明。图8是示出由控制装置100执行的处理的流程的一例的流程图。本流程图的处理在移动体1行驶中根据由乘员进行的指示文与手势的输入而执行。

[0066] 首先,取得部110取得拍摄图像IM、输入指示文以及手势信息(步骤S100)。接下来,生成部130从由取得部110取得的输入指示文提取一个以上的指示,并且根据手势信息生成与乘员P所指示的位置相关的推定分布(步骤S102)。

[0067] 接下来,生成部130根据拍摄图像IM生成概率场景图,并且基于推定分布设定概率场景图的初始概率(步骤S104)。接下来,确定部140基于一个以上的指示,更新概率场景图的概率(步骤S106)。

[0068] 接下来,确定部140判定概率场景图的更新的结果是否确定出单一的物体(步骤

S108)。在判定为确定出单一的物体的情况下,控制部150将确定出的物体作为目标位置而使移动体1行驶(步骤S110)。另一方面,在判定为未确定出单一的物体的情况下,确定部140生成用于确定单一的物体的提问文而进行询问,确定单一的物体(步骤S112)。之后,确定部140使处理向步骤S110转移。由此,本流程图的处理结束。

[0069] 需要说明的是,在上述的实施方式中,作为一例,对信息处理装置搭载于移动体1并在其行驶中活用的例子进行了说明。然而,本发明并不限于那样的结构,更一般而言,信息处理装置至少基于拍摄图像、输入指示文以及手势信息来确定所指示的物体,在所指示的物体的候补存在多个的情况下,追加地生成提问文来进行询问,从而最终确定单一的物体即可。例如,本发明的信息处理装置也能够活用于在VR(virtual reality)空间中用户所指示的物体的确定。

[0070] 根据如以上那样说明了的本实施方式,取得拍摄图像、输入指示文以及手势信息,基于根据拍摄图像生成的概率场景图、从输入指示文提取出的指示以及根据手势信息生成的推定分布,确定乘员所指示的物体,在乘员所指示的物体的候补存在多个的情况下,生成用于确定单一的物体的提问文,而向乘员询问。由此,能够活用人物的视线、手势等模态并且解决在推论的过程中引发的含糊。

[0071] 上述说明的实施方式能够如以下这样表现。

[0072] 一种信息处理装置,其中,

[0073] 所述信息处理装置具备:

[0074] 存储介质(storage medium),其保存能够由计算机读入的命令(computer-readable instructions);以及

[0075] 处理器,其与所述存储介质连接,

[0076] 所述处理器通过执行能够由所述计算机读入的命令(the processor executing the computer-readable instructions to:),而进行如下处理:

[0077] 取得由搭载于移动体的相机拍摄所述移动体的周边而得到的拍摄图像、由所述移动体的乘员输入的输入指示文以及与由所述乘员进行了的手势相关的手势信息;

[0078] 通过对所述输入指示文实施第一规定处理,而提取所述输入指示文所包含的一个以上的指示;

[0079] 通过对所述手势信息实施第二规定处理,而生成与所述乘员所指示的位置相关的推定分布;

[0080] 根据所述拍摄图像,生成对所述拍摄图像所包含的各物体赋予了概率的概率场景图;以及

[0081] 基于所述一个以上的指示、所述推定分布以及所述概率场景图,确定所述拍摄图像中的所述乘员所指示的物体。

[0082] 以上使用实施方式说明了本发明的具体实施方式,但本发明丝毫不被这样的实施方式限定,在不脱离本发明的主旨的范围内能够施加各种变形以及替换。

[0083] 附图标记说明

[0084] 10 外界检知器件

[0085] 12 移动体传感器

[0086] 14 操作件

- [0087] 16 内部相机
- [0088] 18 测位装置
- [0089] 20 HMI
- [0090] 22 模式切换开关
- [0091] 30 移动机构
- [0092] 40 驱动装置
- [0093] 50 外部报告装置
- [0094] 70 存储装置
- [0095] 72 导航地图信息
- [0096] 100 控制装置
- [0097] 110 取得部
- [0098] 120 提取部
- [0099] 130 生成部
- [0100] 140 确定部
- [0101] 150 控制部。

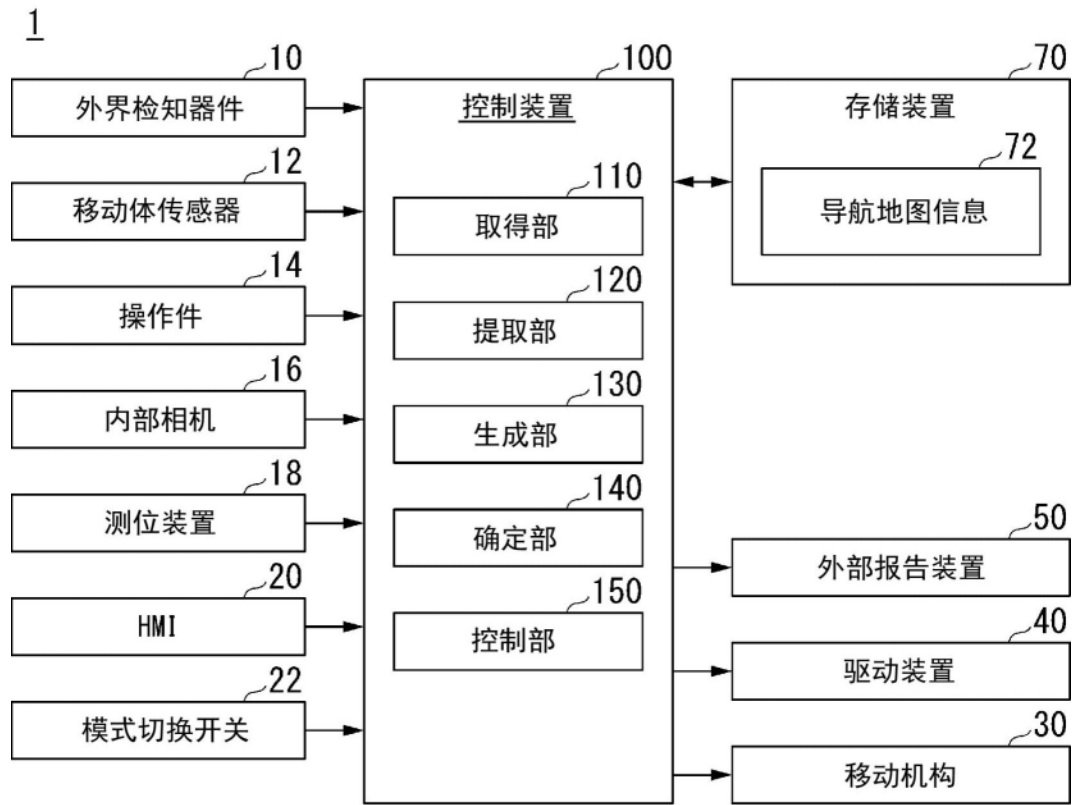


图1

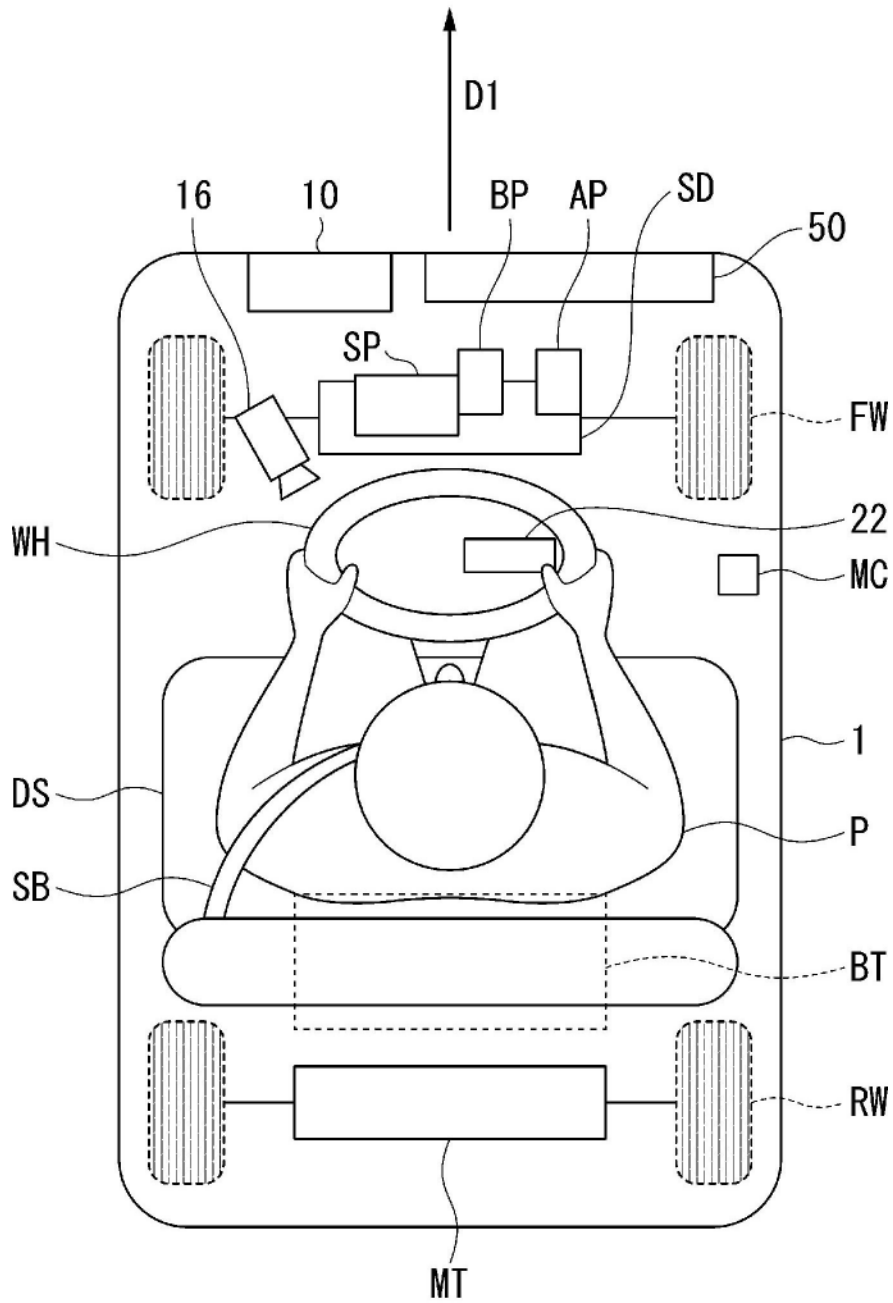


图2

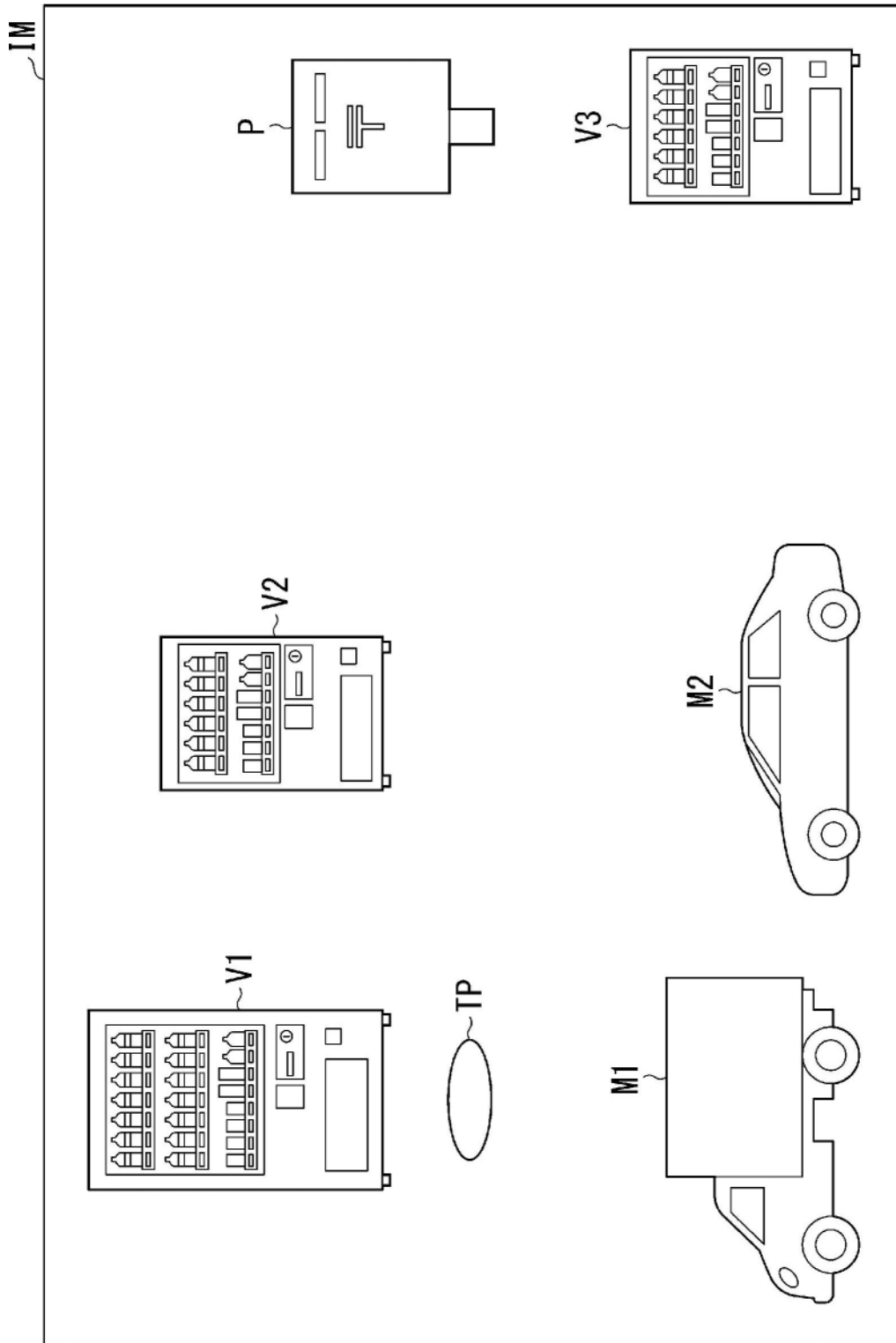


图3

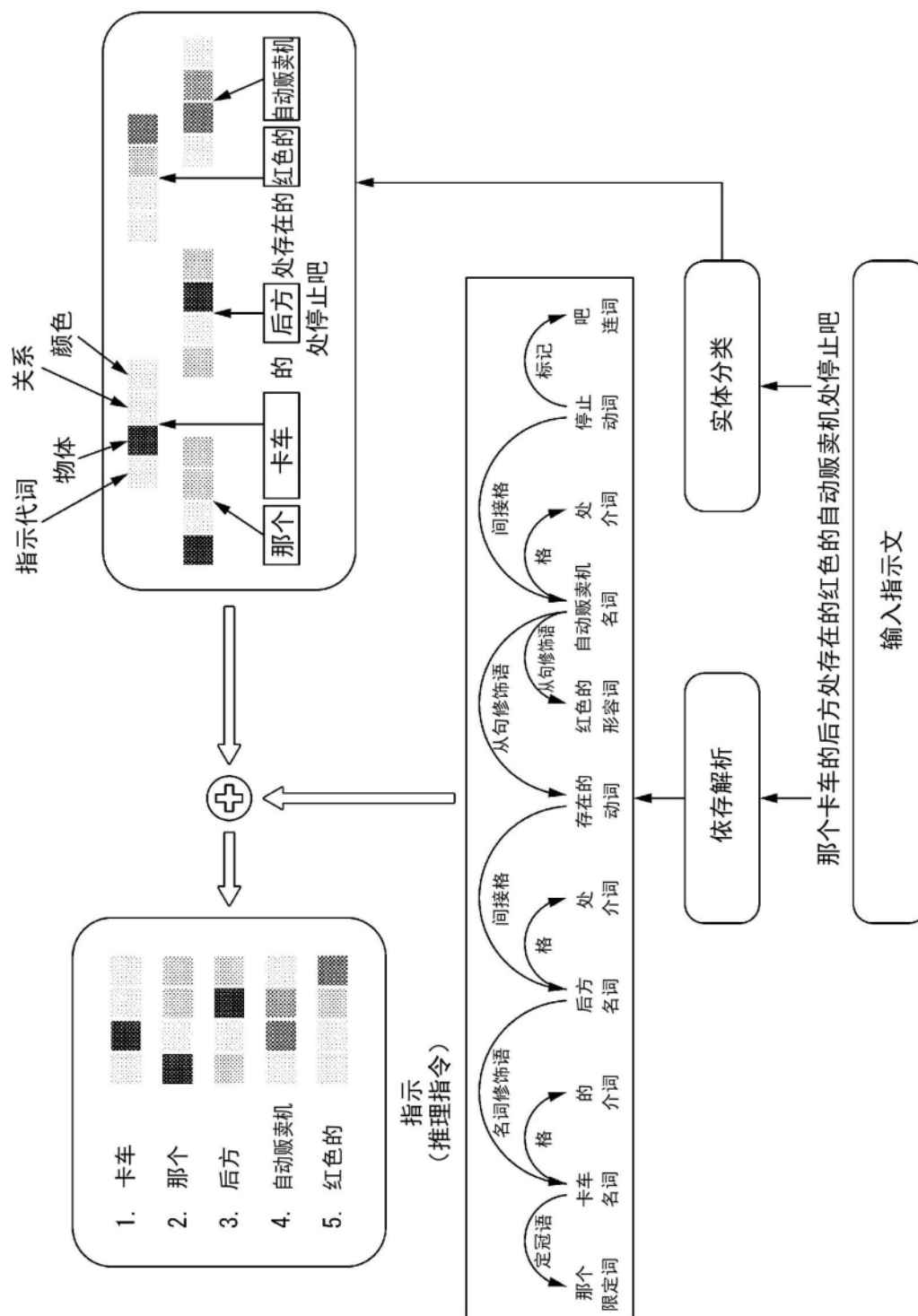


图4

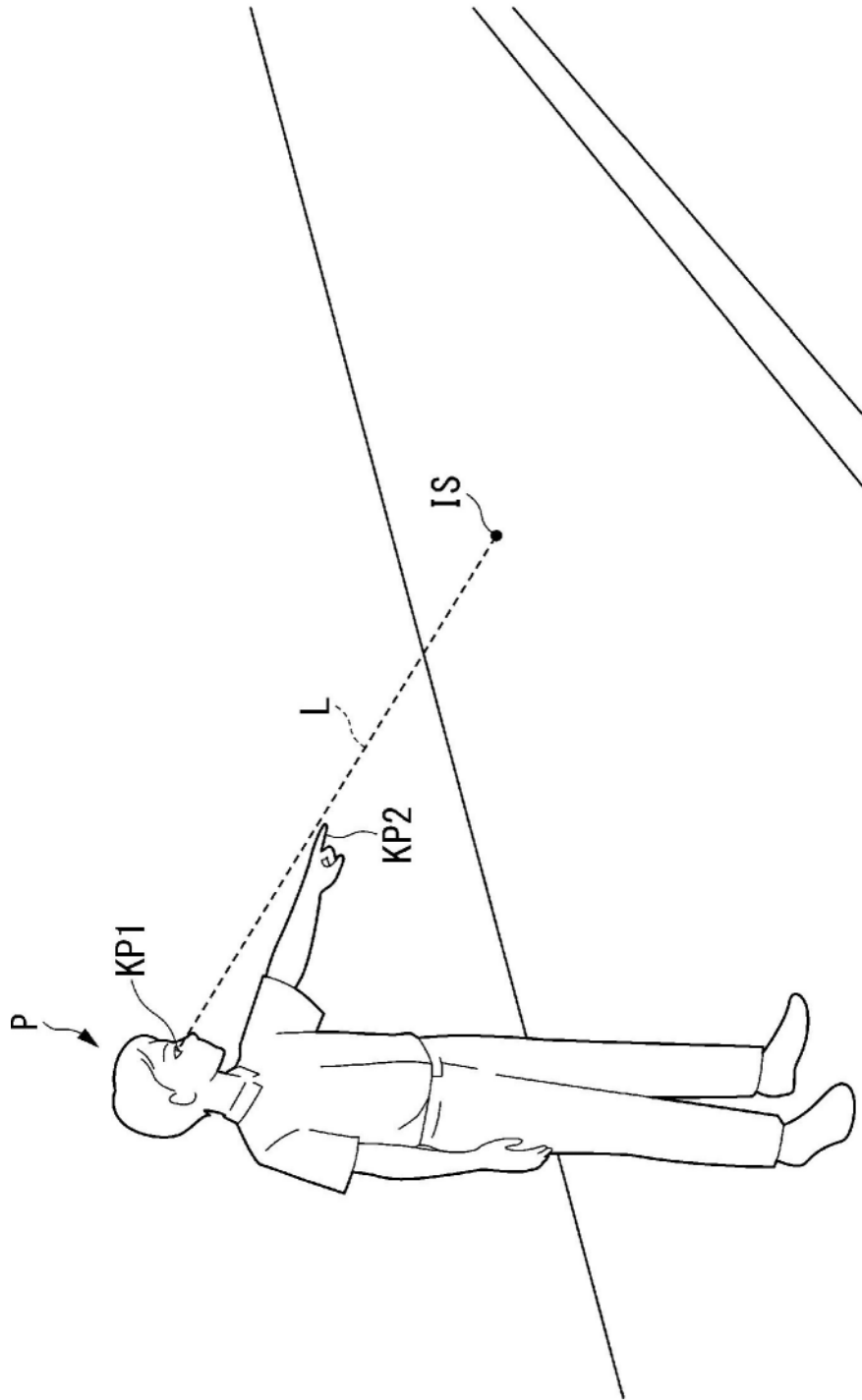


图5

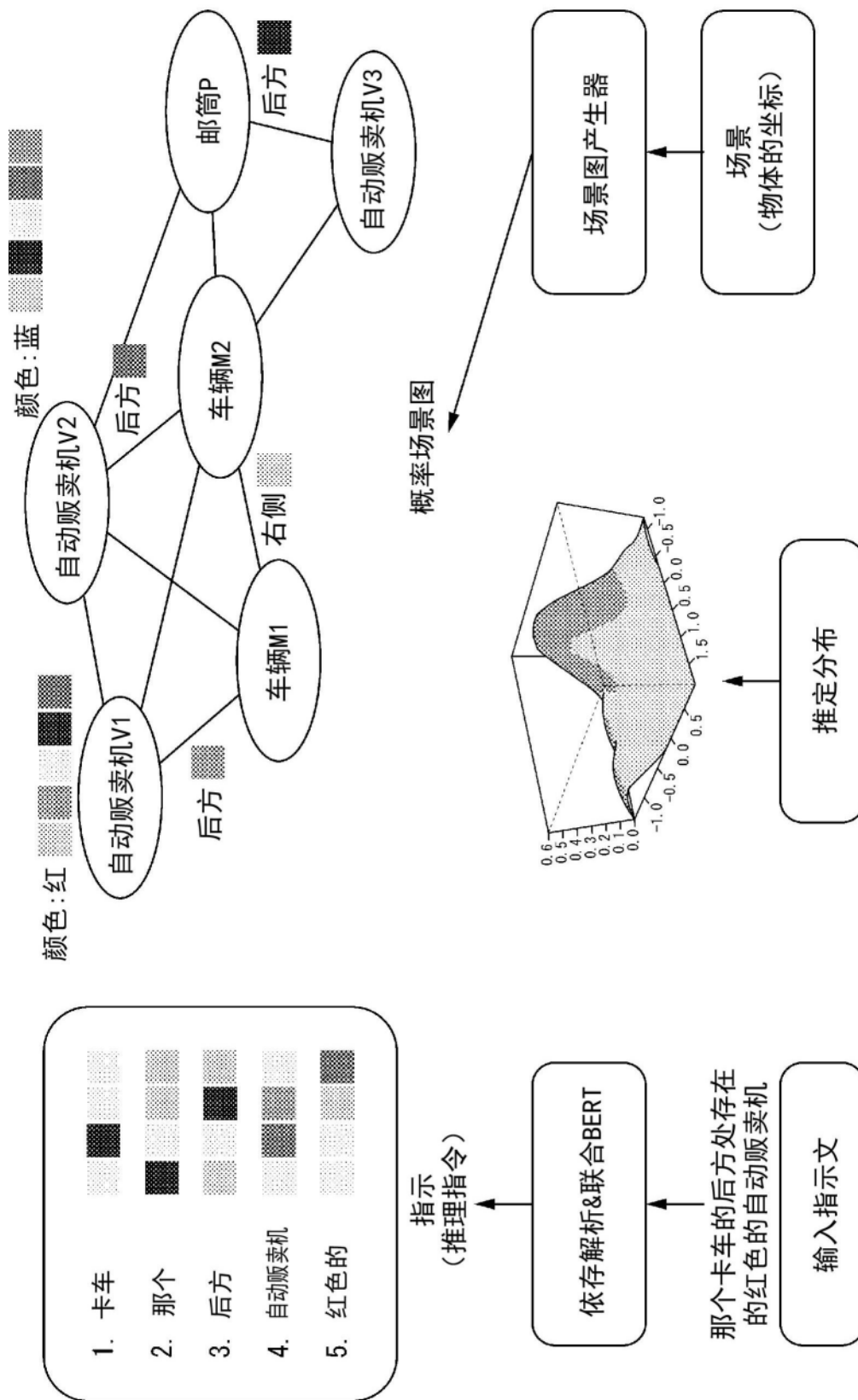


图6

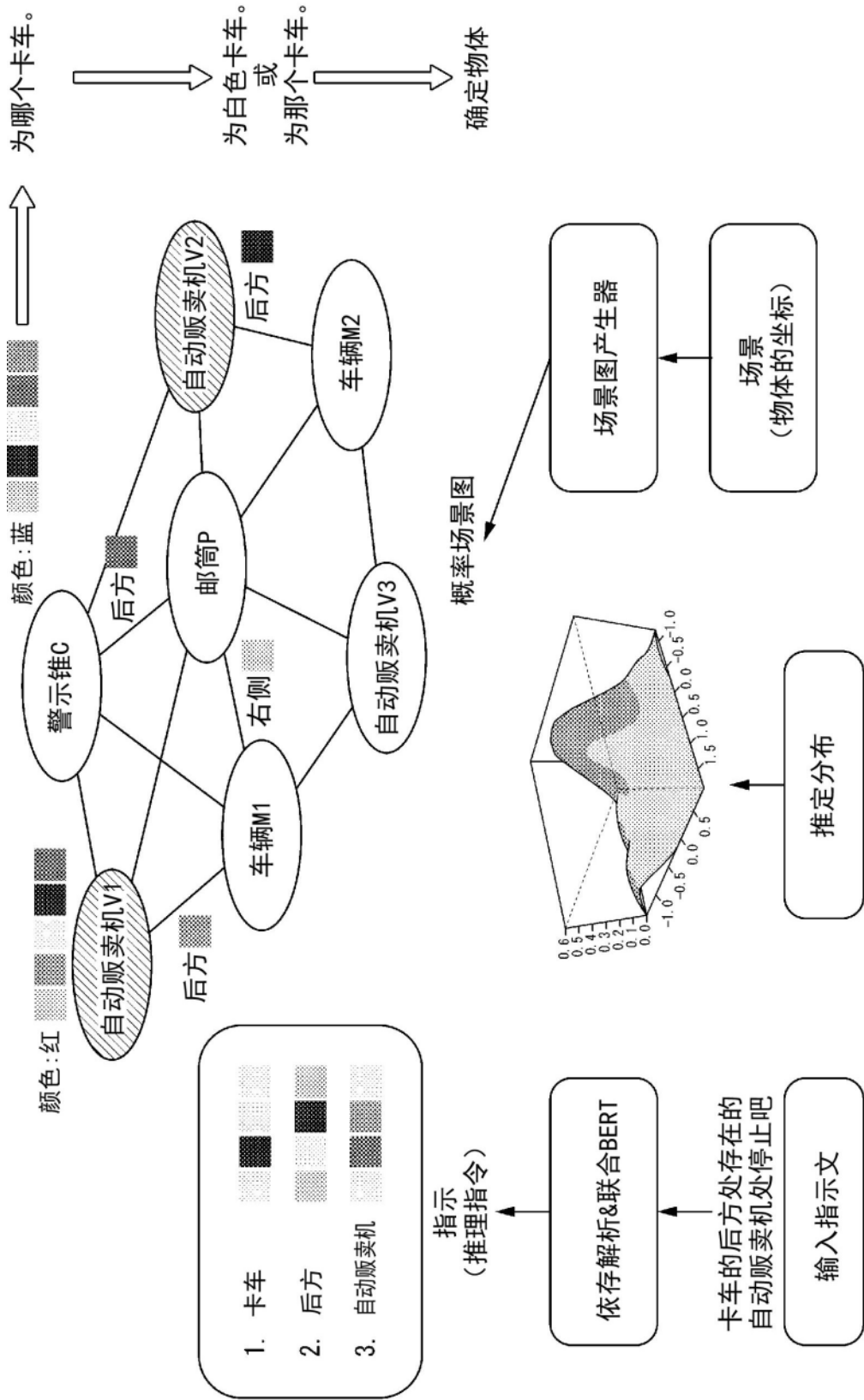


图7

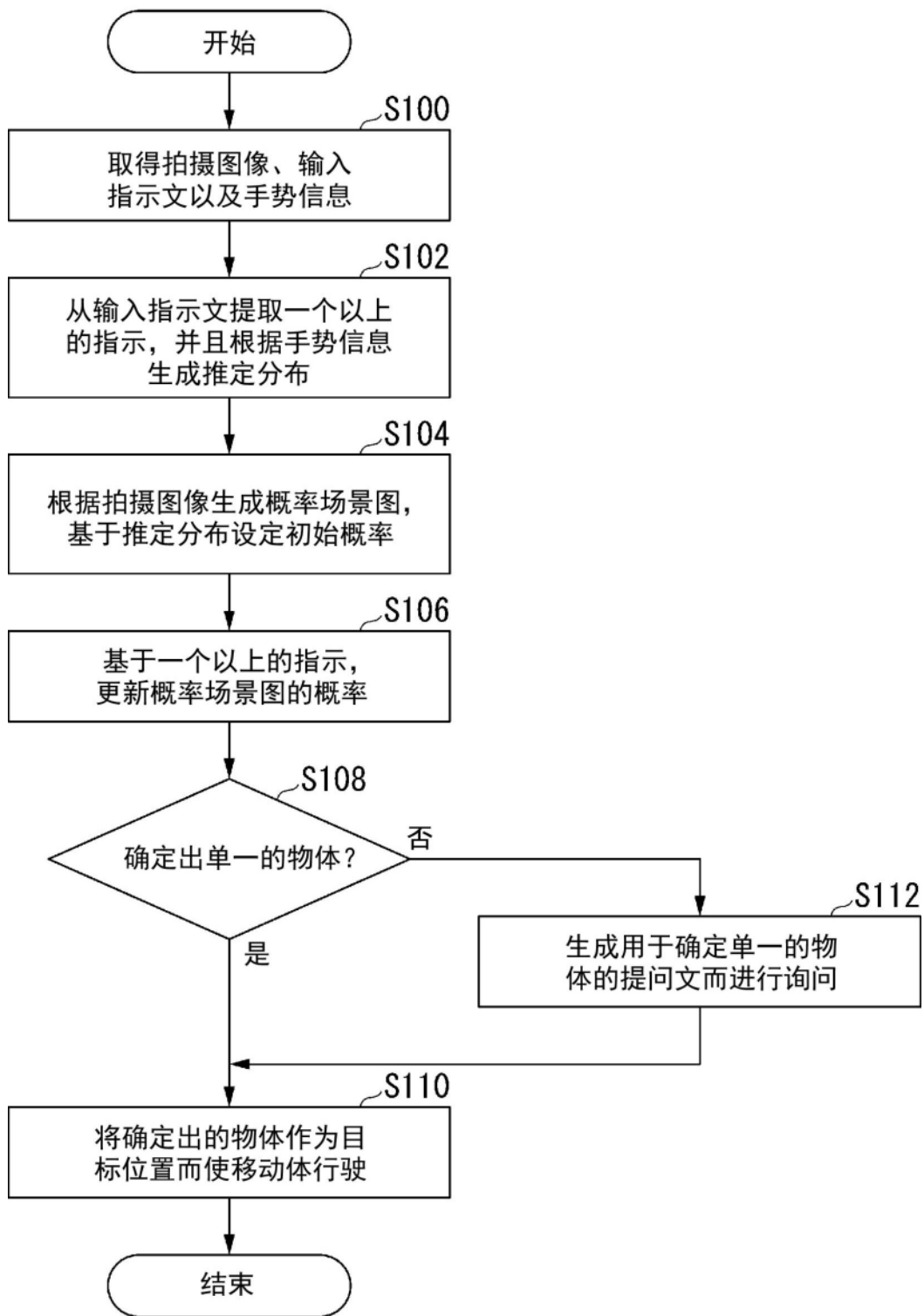


图8