



(12) 发明专利申请

(10) 申请公布号 CN 103189897 A

(43) 申请公布日 2013.07.03

(21) 申请号 201280002529.6

代理人 王成坤 胡建新

(22) 申请日 2012.08.29

(51) Int. Cl.

(30) 优先权数据

G06T 7/00 (2006.01)

61/554,643 2011.11.02 US

(85) PCT申请进入国家阶段日

2013.02.21

(86) PCT申请的申请数据

PCT/JP2012/005446 2012.08.29

(87) PCT申请的公布数据

W02013/065220 JA 2013.05.10

(71) 申请人 松下电器产业株式会社

地址 日本大阪府

(72) 发明人 黄仲阳 华扬 颜水成 陈强

川西亮一

(74) 专利代理机构 永新专利商标代理有限公司

72002

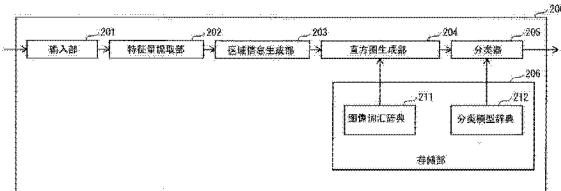
权利要求书1页 说明书12页 附图8页

(54) 发明名称

图像识别装置、图像识别方法和集成电路

(57) 摘要

本发明提供如下的图像识别装置：能够减少一般物体识别中的输入图像中的识别对象即物体的位置、大小、背景干扰等的影响，与以往相比，能够提高一般物体识别的精度。本发明的图像识别装置具有：分割单元，根据从输入图像的内容中提取出的意思对该输入图像进行分割；生成单元，计算所述分割后的各个区域的特征信息，生成反映了计算出的特征信息的所述输入图像的特征信息；以及检查单元，根据所述输入图像的特征信息，对所述输入图像中的识别对象物体的存在情况进行检查。



1. 一种图像识别装置,进行一般物体识别,其特征在于,具有:

分割单元,根据从输入图像的内容中提取出的意思对该输入图像进行分割;

生成单元,计算所述分割后的各个区域的特征信息,生成反映了计算出的特征信息的所述输入图像的特征信息;以及

检查单元,根据所述输入图像的特征信息,对所述输入图像中的识别对象物体的存在情况进行检查。

2. 如权利要求1所述的图像识别装置,其特征在于,

作为表现从所述输入图像的内容中提取出的意思的信息,所述分割单元针对构成所述输入图像的各像素,生成利用多个阈值对计量图进行分级而得到的意思图并使用,该计量图是通过对与识别对象物体相关联的像素即概率进行映射而生成的。

3. 如权利要求2所述的图像识别装置,其特征在于,

所述分割单元针对所述输入图像的各像素,对表示识别对象物体的一部分的像素即概率进行映射,从而生成所述计量图。

4. 如权利要求2所述的图像识别装置,其特征在于,

所述分割单元对所述输入图像的各像素与周围像素相比的显著程度进行映射,从而生成所述计量图。

5. 如权利要求2所述的图像识别装置,其特征在于,

所述分割单元使用所述意思图,有层次地分割所述输入图像。

6. 如权利要求1所述的图像识别装置,其特征在于,

所述检查单元是使用学习图像进行学习的分类器,作为与所述识别对象物体的存在有关的检查,使用所述学习的结果判断所述输入图像中是否存在所述识别对象物体。

7. 一种图像识别方法,用于具有分割单元、生成单元和检查单元的进行一般物体识别的图像识别装置,该图像识别方法的特征在于,包括:

分割步骤,所述分割单元根据从输入图像的内容中提取出的意思对该输入图像进行分割;

生成步骤,所述生成单元计算所述分割后的各个区域的特征信息,生成反映了计算出的特征信息的所述输入图像的特征信息;以及

检查步骤,所述检查单元根据所述输入图像的特征信息,对所述输入图像中的识别对象物体的存在情况进行检查。

8. 一种集成电路,用于进行一般物体识别的图像识别装置,其特征在于,具有:

分割单元,根据从输入图像的内容中提取出的意思对该输入图像进行分割;

生成单元,计算所述分割后的各个区域的特征信息,生成反映了计算出的特征信息的所述输入图像的特征信息;以及

检查单元,根据所述输入图像的特征信息,对所述输入图像中的识别对象物体的存在情况进行检查。

## 图像识别装置、图像识别方法和集成电路

### 技术领域

[0001] 本发明涉及进行一般物体识别的图像识别技术,特别涉及金字塔匹配技术。

### 背景技术

[0002] 近年来,关于一般物体识别,提出了使用空间金字塔匹配(SPM) (参照非专利文献1)的手法。在SPM中,在计算表现输入图像的特征的直方图时,首先,有层次地分割输入图像,以使得将输入图像分割成多个区域,进而,将分割后的各区域分割成多个区域。然后,针对各层次中的分割后的各个区域计算直方图,将它们结合起来计算表示输入图像的特征的直方图。

[0003] 由此,能够在表示输入图像的特征的直方图中反映出分割后的各区域的输入图像中的几何学的位置关系,能够提高一般物体识别的精度。

[0004] 现有技术文献

[0005] 非专利文献

[0006] 非专利文献1 :S. Lazebnik, C. Schmid, and J. Ponce,“Beyond Bags of Features : Spatial Pyramid Matching for Recognizing Natural Scene Categories”, Proc. of IEEE Computer Vision and Pattern Recognition, pp. 2169–2178 (2006).

[0007] 非专利文献2 :G. Wang、D. Forsyth 著、Joint learning of visual attributes, object classes and visual saliency. In Proceedings of IEEE International Conference on Computer Vision, pages537–544, 2009.

[0008] 非专利文献3 :Laurent Itti、Christof Koch、Ernst Niebur 著、A model of saliency-based visual attention for rapid scene analysis, IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. 20, NO. 11, NOVEMBER1998

### 发明内容

[0009] 发明要解决的课题

[0010] 但是,如上所述,在使用将输入图像分割成多个区域的手法的一般物体识别中,也需要进一步提高精度。

[0011] 因此,本发明的目的在于,提供如下的图像识别装置:能够减少一般物体识别中的输入图像中的识别对象即物体的位置、大小、背景干扰等的影响,与以往相比,能够提高一般物体识别的精度。

[0012] 用于解决课题的手段

[0013] 为了解决上述课题,本发明的图像识别装置进行一般物体识别,其中,该图像识别装置具有:分割单元,根据从输入图像的内容中提取出的意思对该输入图像进行分割;生成单元,计算所述分割后的各个区域的特征信息,生成反映了计算出的特征信息的所述输入图像的特征信息;以及检查单元,根据所述输入图像的特征信息,对所述输入图像中的识别对象物体的存在情况进行检查。

[0014] 发明效果

[0015] 根据上述结构,本发明的图像识别装置在对输入图像进行分割时,不是规则地进行等分割,而是根据与输入图像的内容有关的信息进行分割,所以,在根据分割后的图像而生成的直方图中反映出与输入图像的内容有关的信息,能够减少一般物体识别中的输入图像中的识别对象即物体的位置、大小、背景干扰等的影响,与以往相比,能够提高一般物体识别的精度。

## 附图说明

[0016] 图1是示出基于金字塔匹配的一般物体识别处理的流程图。

[0017] 图2是示出图像识别装置200的结构的框图。

[0018] 图3(a)是示出使用意思图(位置图)的输入图像分割处理顺序的流程图,(b)是示意地示出在输入图像分割处理中使用的图像等的一例的图。

[0019] 图4是用于说明输入图像的直方图的生成的图,(a)是示出输入图像的图,(b)是示出分割图像的图,(c)是示出空间子区域的直方图的图,(d)是示出输入图像的直方图的图。

[0020] 图5是用于说明使用输入图像的直方图的基于匹配处理的判断有无存在各识别对象物体的图。

[0021] 图6(a)是示出使用意思图(显著性图)的输入图像分割处理顺序的流程图,(b)是示意地示出在输入图像分割处理中使用的图像等的一例的图。

[0022] 图7是用于说明输入图像的直方图的生成的图,(a)是示出输入图像的图,(b)是示出分割图像的图,(c)是示出空间子区域的直方图的图,(d)是示出输入图像的直方图的图。

[0023] 图8(a)是示出一般的金字塔匹配中的输入图像(层次0)的图像及其直方图的一例的示意图,(b)是示出层次1的分割图像及其直方图的一例的示意图,(c)是示出层次2的分割图像及其直方图的一例的示意图。

## 具体实施方式

[0024] <得到本发明的一个方式为止的经过>

[0025] 发明人对所述基于SPM的一般物体识别进行了详细研究。

[0026] 在SPM中,在分割输入图像时进行等分割。因此,发现了如下课题:针对输入图像而生成的直方图受到输入图像中的识别对象物体的位置、大小、背景干扰等的影响而变化,有时降低一般物体识别的能力。下面,对这点进行说明。

[0027] 关于识别对象物体的位置,例如假设如下情况:对输入图像进行4等分(纵、横分别2等分),在左上方的区域中映出识别对象物体。

[0028] 针对该输入图像而得到的直方图与针对在右下方的区域中映出识别对象物体的图像而得到的直方图不同。由此,例如,作为学习用图像,在分类器针对在右下方的区域中映出识别对象物体的图像多多学习的情况下,与在右下方的区域中映出识别对象物体的情况相比,分类器中的与该输入图像有关的识别对象物体的识别精度较低。

[0029] 并且,关于识别对象物体的大小、背景干扰,例如假设在输入图像的一个区域中较

小地映出识别对象物体的情况。

[0030] 该情况下,与较大地映出识别对象物体的情况相比,与该区域有关的直方图中的背景部分的特征量的影响较大,相反,识别对象物体的特征量的影响较小。由此,分类器很难与预先针对识别对象物体进行的学习结果进行匹配,与较大地映出识别对象物体的情况相比,与该输入图像有关的识别对象物体的识别精度较低。

[0031] 而且,关于这点,本发明人着重研究的结果,发现如下情况并得到本发明:在对输入图像进行分割时,不是规则地进行等分割,而是根据与输入图像的内容有关的信息进行分割,由此,在根据分割后的图像而生成的直方图中反映出与输入图像的内容有关的信息,能够减少一般物体识别中的输入图像针对识别对象即物体的位置、大小、背景干扰等的影响。

[0032] <1. 概要>

[0033] 图1是示出本发明的一个实施方式的图像识别装置的基于金字塔匹配的一般物体识别处理的流程图。

[0034] 基于金字塔匹配的一般物体识别处理由以下处理构成:针对输入图像中的各特征点提取特征量的处理(S101)、有层次地分割输入图像的处理(S102)、针对通过分割而生成的各空间子区域生成表现各个特征的直方图的处理(S103)、根据所生成的各空间子区域的直方图而生成输入图像的直方图的处理(S104)、使用输入图像的直方图并通过分类器判定输入图像中是否存在识别对象物体的分类处理(S105)。

[0035] 这些处理中的与本申请发明的特征部分有关的处理是“有层次地分割输入图像的处理(S102)”。

[0036] “有层次地分割”输入图像是指,递归地细致地分割输入图像,以使得将输入图像(层次0)分割成多个区域(层次1),进而将通过分割而生成的各区域(以下称为“空间子区域”)分割成多个区域(层次2)。

[0037] 这里,在一般的金字塔匹配(空间金字塔匹配:SPM)中,在有层次地将输入图像分割成空间子区域的情况下进行等分割。具体而言,如图8示出一例那样,首先对输入图像即图像801进行N等分(在本实施方式中为4等分(纵方向2等分、横方向2等分))(空间子区域811~814),进而,分别对通过分割而生成的各空间子区域进行N分割(空间子区域821~824、831~834、841~844、851~854)。

[0038] 然后,分别根据这些等分割后的空间子区域生成表示各空间子区域的特征的直方图(801H、811H~814H、821H~824H、831H~834H、841H~844H、851H~854H)。与输入图像有关的最终的直方图例如是在横向连结801H、811H~814H、821H~824H、831H~834H、841H~844H和851H~854H而得到的直方图。

[0039] 但是,在这样规则地对输入图像进行等分割的情况下,针对该输入图像而生成的直方图受到输入图像中的识别对象物体的位置、大小、背景干扰等的影响,在使用该直方图的情况下,有时基于分类器的一般物体识别的能力低下。

[0040] 因此,在本实施方式中,在S102中,在有层次地分割输入图像的情况下,不是如等分割那样规则地进行分割,而是根据输入图像而针对各层次生成意思图,使用该意思图对输入图像进行分割。

[0041] 意思图是表现从输入图像的内容中提取出的意思的图,针对构成输入图像的各像

素,利用多个阈值对计量图(Score map)进行分级而得到意思图,该计量图是通过对与识别对象物体相关联的像素即概率进行映射而生成的。

[0042] 在本实施方式中,作为意思图的一例,使用位置图。位置图是针对输入图像中的各像素将通过对识别对象物体的一部分即概率进行映射而生成的计量图分级为多个等级(在本实施方式中为4个等级)而得到的图。

[0043] 然后,针对使用意思图进行分割后的各空间子区域生成直方图。该情况下,在各直方图中,针对构成该空间子区域的各像素反映识别对象物体的一部分即概率。

[0044] 例如,与位置图的第1等级的(识别对象物体的一部分即概率最高)区域对应的空间子区域的直方图是以浓重颜色反映了识别对象物体的特征的直方图,很难反映出背景干扰的特征。

[0045] 与该第1等级对应的空间子区域的直方图作为输入图像的直方图的一部分而被连结。即,输入图像的直方图包括以浓重颜色反映了识别对象物体的特征的部分。由此,在分类器使用该直方图的情况下,能够对识别对象物体进行识别的概率提高。

[0046] 并且,即使输入图像中的第1等级的区域存在于输入图像的任意位置、为任意大小,针对输入图像而生成的直方图也不变。即,关于识别对象物体的一部分即概率高的区域,排除了位置、大小的影响。

[0047] 如上所述,在本实施方式中,通过使用意思图对输入图像进行分割,减小了输入图像中的物体的位置、大小等对所生成的直方图造成的不良影响,提高了一般物体识别精度。

[0048] 下面,更加详细地说明本实施方式。

[0049] <2. 结构>

[0050] 下面,对本发明的一个实施方式进行一般物体识别的图像识别装置200的结构进行说明。

[0051] 图2是示出图像识别装置200的结构的框图。

[0052] 如图2所示,图像识别装置200构成为包括输入部201、特征量提取部202、区域信息生成部203、直方图生成部204、分类器205和存储部206。

[0053] 并且,图像识别装置200构成为包括处理器和存储器,特征量提取部202、区域信息生成部203、直方图生成部204和分类器205的功能通过由处理器执行存储在存储器中的程序而实现。

[0054] (1) 输入部201

[0055] 输入部201由通信用LSI或存储器存取用IC等构成,具有通过通信或从存储器器件中读出而取得输入图像的功能。

[0056] (2) 特征量提取部202

[0057] 特征量提取部202具有选出输入图像中的特征点并提取与各特征点有关的特征量的功能。与该功能有关的处理相当于图1的S101。

[0058] 具体而言,提取特征量的功能如下实现。

[0059] 首先,特征量提取部202选出输入图像中的特征点。作为一例,特征点是位于输入图像上的一定间隔的位置的格子点和/或特征性的离散点。特征性的离散点是亮度、形状急剧变化的特征性的点,例如是表示猫的眼睛的一部分、耳朵的前端的点等。下面,有时将

上述格子点、离散点统称为“特征点”。

[0060] 接着，特征量提取部 202 在特征点及其周边提取作为特征量的描述符。

[0061] 作为一例，描述符的提取使用众所周知的 SIFT (Scale Invariant Feature Transform) 算法。SIFT 决定特征点的代表亮度梯度方向，以该方向为基准生成其他方向的亮度梯度直方图，利用多维向量记述特征量。

[0062] (3) 区域信息生成部 203

[0063] 区域信息生成部 203 具有意思图生成功能、以及使用意思图有层次地将输入图像分割成空间子区域的输入图像分割功能。

[0064] <意思图生成功能>

[0065] 作为意思图生成功能，区域信息生成部 203 生成意思图。在本实施方式中，作为意思图的一例，区域信息生成部 203 生成位置图。

[0066] 位置图使用具有教师的学习表示输入图像的各像素位置处的检测对象即物体的存在概率。更具体而言，位置图针对输入图像中的各像素，将通过对识别对象物体的一部分即概率进行映射而生成的计量图分成多个(在本实施方式中为 4 个)等级。

[0067] 与该意思图生成功能有关的处理相当于图 1 的 S102、图 3 的 S301 ~ S303。

[0068] 另外，区域信息生成部 203 在位置图的生成中使用多个检测器。各检测器预先针对例如“正面面部”、“右横面部”、“腹部”、“腿部”等表示分别作为检测对象的物体或其一部分的图像(下面为了便于说明而称为“要素图像”。)进行学习。

[0069] 这样，使用多个检测器是因为，例如在假设从输入图像中检测到“面部”的情况下，“面部”存在“正面面部”、“右横面部”、“左横面部”、“右斜面部”、“左斜面部”···等各种面部，并且，面部的大小也存在“大”“中”“小”···等各种大小，无法利用 1 个检测器检测它们全体。

[0070] 另外，位置图的详细情况记载于“G. Wang, D. Forsyth 著、Joint learning of visual attributes, object classes and visual saliency. In Proceedings of IEEE International Conference on Computer Vision, pages537–544, 2009.” 等中。

[0071] <输入图像分割功能>

[0072] 输入图像分割功能是通过使用意思图有层次地分割输入图像而得到空间子区域的功能。针对各个层次生成意思图。

[0073] 与该意思图生成功能有关的处理相当于图 1 的 S102、图 3 的 S304。

[0074] (4) 直方图生成部 204

[0075] 直方图生成部 204 具有如下功能：分别针对基于区域信息生成部 203 的分割处理结果即空间子区域生成直方图，通过连结所生成的直方图，生成输入图像的直方图。与该功能有关的处理相当于图 1 的 S103、S104。

[0076] 与各空间子区域有关的直方图的生成如下实现：针对该空间子区域中包含的各特征点，对与特征量提取部 202 提取出的特征量(描述符)最相近的图像词汇进行投票。这里，图像词汇用于对 SIFT 特征量进行向量量化。通过对学习图像群进行机器学习而生成图像词汇的集合即图像词汇辞典(211)，该图像词汇辞典(211)预先存储在存储部 206 中。

[0077] SIFT 特征量、直方图的生成是众所周知的，省略进一步的说明。

[0078] (5) 分类器 205

[0079] 作为一例,分类器 205 由支持向量机(SVM)构成,具有使用输入图像的直方图将输入图像分配给多个类中的至少一个类的功能(分类功能),判断输入图像中是否存在识别对象物体。

[0080] 具体而言,分类器 205 对由直方图生成部 204 生成的输入图像的直方图和存储在存储部 206 中的分类模型辞典 212 进行比较,从而实现判断是否存在识别对象物体的功能。该功能相当于图 1 的 S105。

[0081] 这里,分类模型辞典 212 存储与包含作为识别对象的各个物体的图像有关的直方图。通过与学习图像群有关的机器学习而预先生成分类模型辞典 212,该分类模型辞典 212 存储在存储部 206 中。

[0082] 由于 SVM 是众所周知的,所以,省略进一步的详细说明。

[0083] (6) 存储部 206

[0084] 存储部 206 由非易失性存储器构成,存储上述图像词汇辞典 211、分类模型辞典 212 等的各种信息。

[0085] <3. 动作>

[0086] <3-1. 使用意思图的输入图像分割处理>

[0087] 图 3 是用于说明使用位置图作为意思图的输入图像分割处理的图。

[0088] 图 3 (a) 是示出使用位置图的输入图像分割处理的流程图,相当于图 1 的 S102。并且,图 3 (b) 是示意地示出在输入图像分割处理中使用的图像等的一例的图。

[0089] 首先,区域信息生成部 203 中的各检测器在输入图像上对规定尺寸(例如 8 像素 × 8 像素)的判定窗进行扫描,进行所述要素图像的检测处理(S301)。

[0090] 更详细地讲,在输入图像(作为一例,为图 3 (b) 的输入图像 311)中,对判定窗进行扫描并检测要素图像。在检测到要素图像的情况下,还存储针对该检测结果的可靠性(准确度)。然后,通过将这些可靠性(未检测到的情况设为可靠性 0。)映射到输入图像的相应范围内,生成可靠性图。

[0091] 接着,作为主处理,区域信息生成部 203 针对构成输入图像的各像素,使用检测到要素图像的检测器的个数和可靠性图,生成表示输入图像的各像素位置处的“识别对象物体”的存在概率的计量图(S302)。

[0092] 作为一例,针对输入图像的各像素,在检测数为 2 以上、可靠性的合计为 0.5 以上的情况下,针对该像素,将“识别对象物体”的存在概率设为其可靠性的合计值。

[0093] 这样得到的计量图为灰度图像(作为一例,为图 3 (b) 的灰度图像 312)。存在概率高的位置的明亮度高,存在概率低的位置的明亮度低。

[0094] 接着,区域信息生成部 203 使用规定阈值(作为一例,为 3 个阈值)等,将表现计量图的灰度等级图像的灰度分级为与规定图像分割数相同数量的灰度数(例如 4 个),从而取得位置图(作为一例,为图 3 (b) 的位置图 313) (S303)。

[0095] 位置图 313 成为被分为等级 1 的区域 331、等级 2 的区域 332、等级 3 的区域 333、等级 4 的区域 334 的 4 个灰度的灰度等级图像。

[0096] 最后,区域信息生成部 203 使用位置图 313 对输入图像 311 进行分割(S304)。将输入图像 311 分割成多个区域的情况下区域间的边界线如表示分割后的图像的图 3 的图像 314 那样,与将对位置图 313 的各等级进行划分的边界线映入输入图像 311 中而得到的

边界线一致。下面,如图像 314 那样,将被分为区域的图像称为“分割图像”。

[0097] 输入图像 311 被分割为与位置图 313 的等级 1 相当的区域、与等级 2 相当的区域、与等级 3 相当的区域和与等级 4 相当的区域的 4 个区域。位置图 313 中的区域 331 的位置、形状与分割图像 314 中的区域 341 的位置、形状相同。并且,位置图 313 中的区域 332、333 和 334 的位置、形状与分割图像 314 中的区域(空间子区域)342、343 和 344 的位置、形状相同。

[0098] 关于上述输入图像分割处理,仅针对层次 1 进行了说明,但是,层次 2 以下的层次也进行同样的处理。另外,关于层数次,在针对“牛”、“猫”这样的等级的物体概念进行识别处理的情况下,一般优选为 2 ~ 3 个层次左右,由此,在本实施方式中,作为层数次,使用 3。

[0099] 另外,灰度图像 312、位置图 313、分割图像 314 是说明用的示意图,不是实际根据输入图像 311 而生成的准确的图像、图等。

[0100] <3-2. 基于金字塔匹配的一般物体识别处理>

[0101] 使用图 1、图 4 和图 5,对图像识别装置 200 进行的基于金字塔匹配的一般物体识别处理进行详细说明。

[0102] 首先,输入部 201 取得输入图像(作为一例,为图 4 (a)的输入图像 311),将其输出到特征量提取部 202。

[0103] 特征量提取部 202 对输入图像中的各特征点提取特征量(S101)。

[0104] 接着,区域信息生成部 203 进行有层次地将输入图像分割成空间子区域的处理(S102)。该处理相当于已经使用图 3 说明的输入图像分割处理。

[0105] 这里,作为一例,设 S102 中生成的层次 1 的分割图像为图 4 (b) 的分割图像 314。另外,图 4 (b) 的分割图像 314 和图 3 (b) 的分割图像 314 相同。

[0106] 接着,直方图生成部 204 进行生成各空间子区域的直方图的处理(S103)。

[0107] 图 4 (c) 是示意地示出层次 1 的各空间子区域的直方图的一例的图。

[0108] 各直方图的横轴排列图像词汇,纵轴是输入图像中的各图像词汇的出现频度。

[0109] 直方图 341H 是与空间子区域 341 对应的直方图。同样,直方图 342H ~ 344H 是分别与空间子区域 342 ~ 344 对应的直方图。

[0110] 然后,直方图生成部 204 通过连结所生成的直方图,生成输入图像的直方图(作为一例,为图 4 (d) 的直方图 410H) (S104)。

[0111] 另外,关于层次 0 的输入图像 311 和对输入图像进行 16 分割而得到的层次 2 的分割图像,也同样生成直方图。然后,连结层次 0、层次 1、层次 2 的各直方图而得到的直方图成为输入图像 311 的最终直方图。

[0112] 最后,分类器 205 使用输入图像的直方图,判断输入图像中是否存在识别对象物体(S105)。

[0113] 分类器 205 判断输入图像中是否存在识别对象物体。

[0114] 另外,在图 5 中,作为输入图像的直方图,仅提取层次 1 的直方图 410H 进行示出,但是,实际上,如上所述,将连结层次 0、层次 1、层次 2 的各直方图而得到的直方图用作输入图像的直方图。关于作为分类模型辞典 212 而存储的表现识别对象物体的直方图,也仅提取层次 1 的直方图 501H、502H 进行示出,但是,实际上,将连结层次 0、层次 1、层次 2 的各直方图而得到的直方图用作表现识别对象物体的直方图。

[0115] <4. 变形例>

[0116] 以上说明了本发明的图像识别装置的实施方式，但是，也可以如下对例示的图像识别装置进行变形，本发明当然不限于上述实施方式所示的图像识别装置。

[0117] (1) 在上述实施方式中，作为意思图，使用位置图，但是，只要是针对构成输入图像的各像素、利用多个阈值对计量图进行分级而得到的图即可，该计量图是通过对与识别对象物体相关联的像素即概率进行映射而生成的。

[0118] 例如，作为意思图，也可以使用显著性图。

[0119] 显著性图是针对输入图像中的各像素将通过对引起视觉注意的强度(概率)进行映射而生成的计量图分级为多个等级而得到的图。图像的各像素位置处的显著性根据该位置处的刺激是否与包围该位置的局部或大部分区域的刺激相比更为显著来决定。在显著性图中，例如，利用像素等级来反映图像的某个区域相当于前景的概率和相当于背景的概率等。

[0120] 图 6(a)是示出使用显著性图作为意思图的情况下输入图像分割处理的流程图，是相当于图 1 的 S102 的处理。

[0121] 并且，图 6(b)是示意地示出在输入图像分割处理中使用的图像等的一例的图。

[0122] 首先，区域信息生成部 203 分别针对作为检测对象的 1 个以上的属性(例如亮度、颜色、纹理或形状等)检测具有与周围不同的特征的图像区域(S601)。下面，将针对各属性而得到的检测结果称为“显著性特性曲线”。

[0123] 接着，区域信息生成部 203 通过进行用于统合 S601 中得到的多个显著性特性响应的主处理，得到计量图(S602)。在主处理中，例如，以规定比例对与亮度、颜色、纹理或形状有关的各显著性特性响应进行相加等，统合多个显著性特性响应，从而得到计量图。

[0124] 计量图表现图像的各位置处的该像素的显著性。这样得到的计量图表现为灰度等级图像(作为一例，为图 6(b)的灰度等级图像 612)。例如，显著性高的位置的明亮度高，显著性低的位置的明亮度低。

[0125] S603 是与上述图 3 的 S303 相同的处理，S604 是与 S304 相同的处理，所以省略说明。

[0126] 另外，图 6 的显著性图 613 示出作为意思图的显著性图的一例。

[0127] 并且，图 6 的分割图像 614 示出使用显著性图 613 对输入图像 611 进行分割而得到的层次 1 的分割图像的一例。

[0128] 另外，灰度等级图像 612、位置图 613、分割图像 614 是说明用的示意图，不是实际根据输入图像 611 而生成的准确的图像、图等。

[0129] 另外，显著性图记载于“Laurent Itti、Christof Koch、Ernst Niebur 著、A model of saliency-based visual attention for rapid scene analysis、IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. 20, NO. 11, NOVEMBER 1998”等中。

[0130] 接着，进行如下处理：生成使用显著性图 613 等进行分割而得到的分割图像中的各空间子区域的直方图。该处理是相当于图 1 的 S103 的处理。

[0131] 图 7 是用于说明输入图像的直方图的生成的图。

[0132] 这里，作为一例，设输入图像为图 7(a)的图像 611，作为一例，设 S102 中生成的层次 1 的分割图像为图 7(b)的分割图像 614。另外，图 7(a)的图像 611 与图 6(b)的图像

611 相同,图 7 (b) 的分割图像 614 与图 6 (b) 的分割图像 614 相同。

[0133] 图 7 (c) 是示意地示出层次 1 的各空间子区域的直方图的一例的图。

[0134] 各直方图的横轴排列图像词汇,纵轴是输入图像中的各图像词汇的出现频度。

[0135] 直方图 641H 是与空间子区域 641 对应的直方图。同样,直方图 642H ~ 644H 是分别与空间子区域 642 ~ 644 对应的直方图。

[0136] 然后,直方图生成部 204 通过连结所生成的直方图,生成输入图像的直方图。该处理是相当于图 1 的 S104 的处理。

[0137] 另外,关于层次 0 的输入图像 611 和对输入图像进行 16 分割而得到的层次 2 的分割图像,也同样生成直方图。然后,连结层次 0、层次 1、层次 2 的各直方图而得到的直方图成为输入图像 611 的最终直方图。

[0138] 最后,分类器 205 使用输入图像的直方图,判断输入图像中是否存在识别对象物体。

[0139] 该处理是相当于图 1 的 S105 的处理。S105 的详细情况已经说明,所以,省略进一步的说明。

[0140] (2) 在上述实施方式和变形例中,作为意思图,使用位置图或显著性图,但是,也可以使用统合这些多个图而得到的图作为意思图。

[0141] 例如,也可以对与位置图有关的计量图和与显著性图有关的计量图进行统合,根据统合后的计量图生成一个意思图。

[0142] 在对不同计量图进行统合时,可以对各个计量图进行同等的加权,也可以对各个计量图进行通过学习等而得到的加权。

[0143] 如上所述,通过使用多个计量图生成意思图,在意思图中,能够反映出输入图像的各像素所具有的多个意思内容。

[0144] 并且,也可以组合使用单纯的几何学分割和利用意思图的分割。

[0145] (3) 在上述实施方式中,作为特征量,计算 SIFT 特征量,但是,只要能够表现特征量即可。例如,作为特征量,可以使用众所周知的特征量即 Color、HOG (Histogram of Oriented Gradients)、LBP (Local Binary Pattern) 等,还可以使用图像数据本身。

[0146] (4) 在上述实施方式中,设有层次地分割输入图像时的金字塔的深度为 3 个层次(层次 0、1 和 2),加深层次时对上位层次的图像(部分图像)进行 4 分割,但是不限于此,也可以采用通过学习或经验等得到的层次的深度、分割数。

[0147] (5) 在上述实施方式中,分类器 205 使用由直方图生成部 204 生成的直方图判断输入图像中是否存在识别对象物体,但是,只要使用由直方图生成部 204 生成的直方图进行与识别对象物体有关的检查即可。例如,也可以使分类器 205 作为检测器进行动作,使用由直方图生成部 204 生成的直方图来检测输入图像中存在的识别对象物体。

[0148] (6) 在上述实施方式中,分类器 205 由 SVM 构成,但是,只要能够将输入图像分配给多个类中的至少一个类即可。

[0149] 例如,也可以应用匹配处理技术。

[0150] 作为一例,预先将针对学习图像群而通过机器学习生成的分类模型辞典 212 存储在存储部 206 中。

[0151] 分类模型辞典 212 是与作为识别对象的各个物体有关的该物体存在于图像中的

情况下的直方图(作为一例,图5中的识别对象物体为“牛”的情况下的直方图501H、识别对象物体为“猫”的情况下的直方图502H)。分类器205对输入图像的直方图和分类模型辞典212中的各直方图进行比较,计算其一致程度。然后,分类器205判断一致程度是否为规定比例以上,在一致程度为规定值以上的情况下,决定为识别对象物体存在于输入图像中,在小于规定值的情况下,决定为识别对象物体未存在于输入图像中。

[0152] (7) 可以将用于使图像识别装置200的处理器和与该处理器连接的各种电路执行上述实施方式所示的提取特征量的处理、意思图生成处理、输入图像分割处理、直方图生成处理、判断是否存在识别对象物体的处理等的、由机器语言或高级语言的程序代码构成的控制程序记录在记录介质中,或经由各种通信路等流通、颁布。

[0153] 这种记录介质存在IC卡、硬盘、光盘、软盘、ROM、闪存等。流通、颁布的控制程序通过存储在可由处理器读出的存储器等中而加以利用,该处理器执行该控制程序,从而实现各实施方式所示的各功能。

[0154] 另外,处理器除了直接执行控制程序以外,还可以进行编辑并执行或通过解释器来执行。

[0155] (8) 上述实施方式所示的各功能结构要素(输入部201、特征量提取部202、区域信息生成部203、直方图生成部204和分类器205等)可以作为执行该功能的电路而实现,也可以通过由1个或多个处理器执行程序而实现。

[0156] 另外,典型地,上述各功能结构要素作为集成电路即LSI而实现。它们可以单独成为一个芯片,也可以包含一部分或全部的方式而成为一个芯片。这里,设为LSI,但是,由于集成度的差异,有时也称为IC、系统LSI、超级LSI、特级LSI。并且,集成电路化的手法不限于LSI,也可以通过专用电路或通用处理器来实现。也可以利用能够在LSI制造后进行编程的FPGA(Field Programmable Gate Array)、或者能够重构LSI内部的细胞电路的连接和设定的可重构处理器。进而,如果出现了通过半导体技术的进步或派生的其他技术置换LSI的集成电路化的技术,则当然也可以使用该技术来进行功能块的集成化。存在应用生物技术等的可能性。

[0157] (9) 可以局部组合上述实施方式和各变形例。

[0158] <5. 补充>

[0159] 下面,进一步对作为本发明的一个实施方式的图像识别装置的结构及其变形例和效果进行说明。

[0160] (1) 本发明的一个实施方式的图像识别装置进行一般物体识别,其中,该图像识别装置具有:分割单元,根据从输入图像的内容中提取出的意思对该输入图像进行分割;生成单元,计算所述分割后的各个区域的特征信息,生成反映了计算出的特征信息的所述输入图像的特征信息;以及检查单元,根据所述输入图像的特征信息,对所述输入图像中的识别对象物体的存在情况进行检查。

[0161] 根据该结构,在对输入图像进行分割时,不是规则地进行等分割,而是根据从输入图像的内容中提取出的意思进行分割,所以,在根据分割后的图像而生成的特征信息中反映出与输入图像的内容有关的信息,能够减少一般物体识别中的输入图像中的识别对象物体的位置、大小、背景干扰等的影响,与以往相比,能够提高一般物体识别的精度。

[0162] 另外,“检查”的用语包括“分类”和“检测”双方。

[0163] (2) 并且,也可以是,作为表现从所述输入图像的内容中提取出的意思的信息,所述分割单元针对构成所述输入图像的各像素,生成利用多个阈值对计量图进行分级而得到的意思图并使用,该计量图是通过对与识别对象物体相关联的像素即概率进行映射而生成的。

[0164] 根据该结构,按照与识别对象物体相关联的概率的每个等级对输入图像进行分割,所以,在根据与识别对象物体相关联的概率高的区域而计算出的特征信息中,与识别对象物体的位置、大小无关而较强地表现出识别对象物体的特征。由此,在输入图像的特征信息中较强地表现出识别对象物体的特征,所以,与以往相比,能够提高一般物体识别的精度。

[0165] (3) 并且,也可以是,所述分割单元针对所述输入图像的各像素,对表示识别对象物体的一部分的像素即概率进行映射,从而生成所述计量图。

[0166] 根据该结构,按照表示识别对象物体的一部分的像素即概率的每个等级对输入图像进行分割,所以,在根据表示识别对象物体的一部分的概率高的区域而计算出的特征信息中,与识别对象物体的位置、大小无关而较强地表现出识别对象物体的特征。由此,在输入图像的特征信息中较强地表现出识别对象物体的特征,所以,与以往相比,能够提高一般物体识别的精度。

[0167] (4) 并且,也可以是,所述分割单元对所述输入图像的各像素与周围像素相比的显著程度进行映射,从而生成所述计量图。

[0168] 根据该结构,按照引起视觉注意的程度的每个等级对输入图像进行分割,所以,在根据引起视觉注意的程度高的区域而计算出的特征信息中,与识别对象物体的位置、大小无关而较强地表现出引起视觉注意的程度高的物体即识别对象物体的特征。由此,在输入图像的特征信息中较强地表现出识别对象物体的特征,所以,与以往相比,能够提高一般物体识别的精度。

[0169] (5) 并且,也可以是,所述分割单元使用所述意思图,有层次地分割所述输入图像。

[0170] 根据该结构,能够减少一般物体识别中的输入图像中的识别对象物体的位置、大小、背景干扰等的影响,与以往相比,能够提高一般物体识别的精度。

[0171] (6) 并且,也可以是,所述检查单元是使用学习图像进行学习的分类器,作为与所述识别对象物体的存在有关的检查,使用所述学习的结果判断所述输入图像中是否存在所述识别对象物体。

[0172] 本发明的一个实施方式的图像识别方法用于具有分割单元、生成单元和检查单元的进行一般物体识别的图像识别装置,其中,该图像识别方法包括以下步骤:分割步骤,所述分割单元根据从输入图像的内容中提取出的意思对该输入图像进行分割;生成步骤,所述生成单元计算所述分割后的各个区域的特征信息,生成反映了计算出的特征信息的所述输入图像的特征信息;以及检查步骤,所述检查单元根据所述输入图像的特征信息,对所述输入图像中的识别对象物体的存在情况进行检查。

[0173] 本发明的一个实施方式的集成电路用于进行一般物体识别的图像识别装置,其中,该集成电路具有:分割单元,根据从输入图像的内容中提取出的意思对该输入图像进行分割;生成单元,计算所述分割后的各个区域的特征信息,生成反映了计算出的特征信息的所述输入图像的特征信息;以及检查单元,根据所述输入图像的特征信息,对所述输入图像

中的识别对象物体的存在情况进行检查。

[0174] 根据该结构，在对输入图像进行分割时，不是规则地进行等分割，而是根据从输入图像的内容中提取出的意思进行分割，所以，在根据分割后的图像而生成的特征信息中反映出与输入图像的内容有关的信息，能够减少一般物体识别中的输入图像中的识别对象物体的位置、大小、背景干扰等的影响，与以往相比，能够提高一般物体识别的精度。

[0175] 产业上的可利用性

[0176] 本发明的一个方式的图像识别装置减少输入图像中的识别对象即物体的位置、大小、背景干扰等的影响，提高一般物体识别的能力，优先用作对拍摄一般物体而得到的照片数据进行分类并管理的装置等。

[0177] 标号说明

[0178] 200：图像识别装置；201：输入部；202：特征量提取部；203：区域信息生成部；204：直方图生成部；205：分类器；206：存储部；211：图像词汇辞典；212：分类模型辞典；311：输入图像；312：灰度等级图像；313：位置图；314：分割图像。

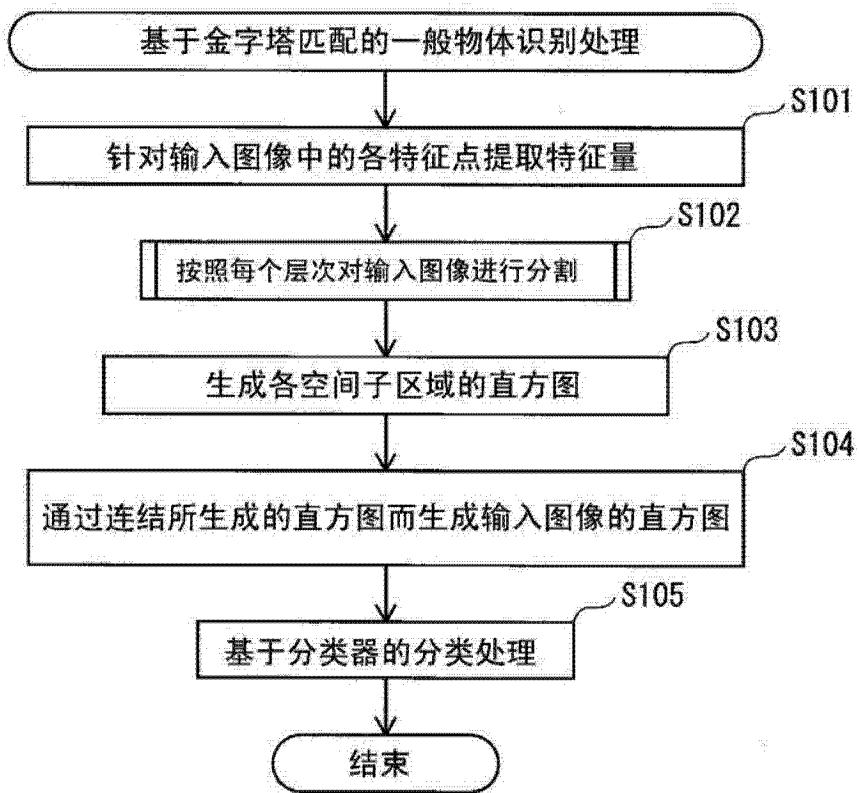


图 1

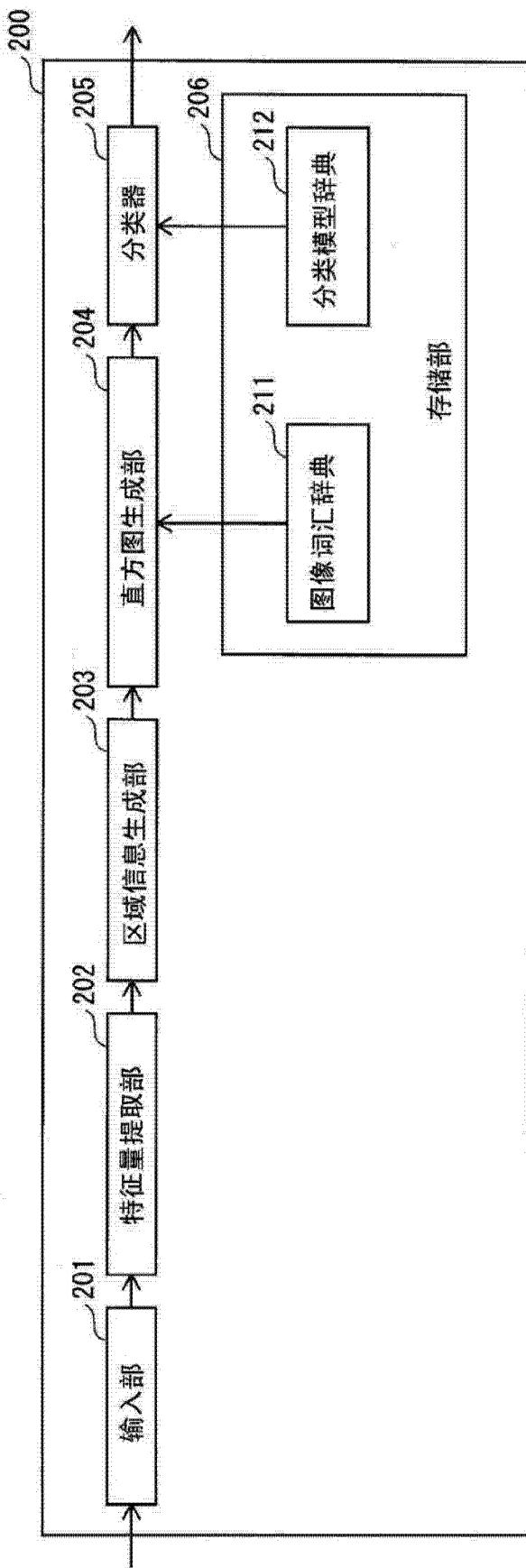


图 2

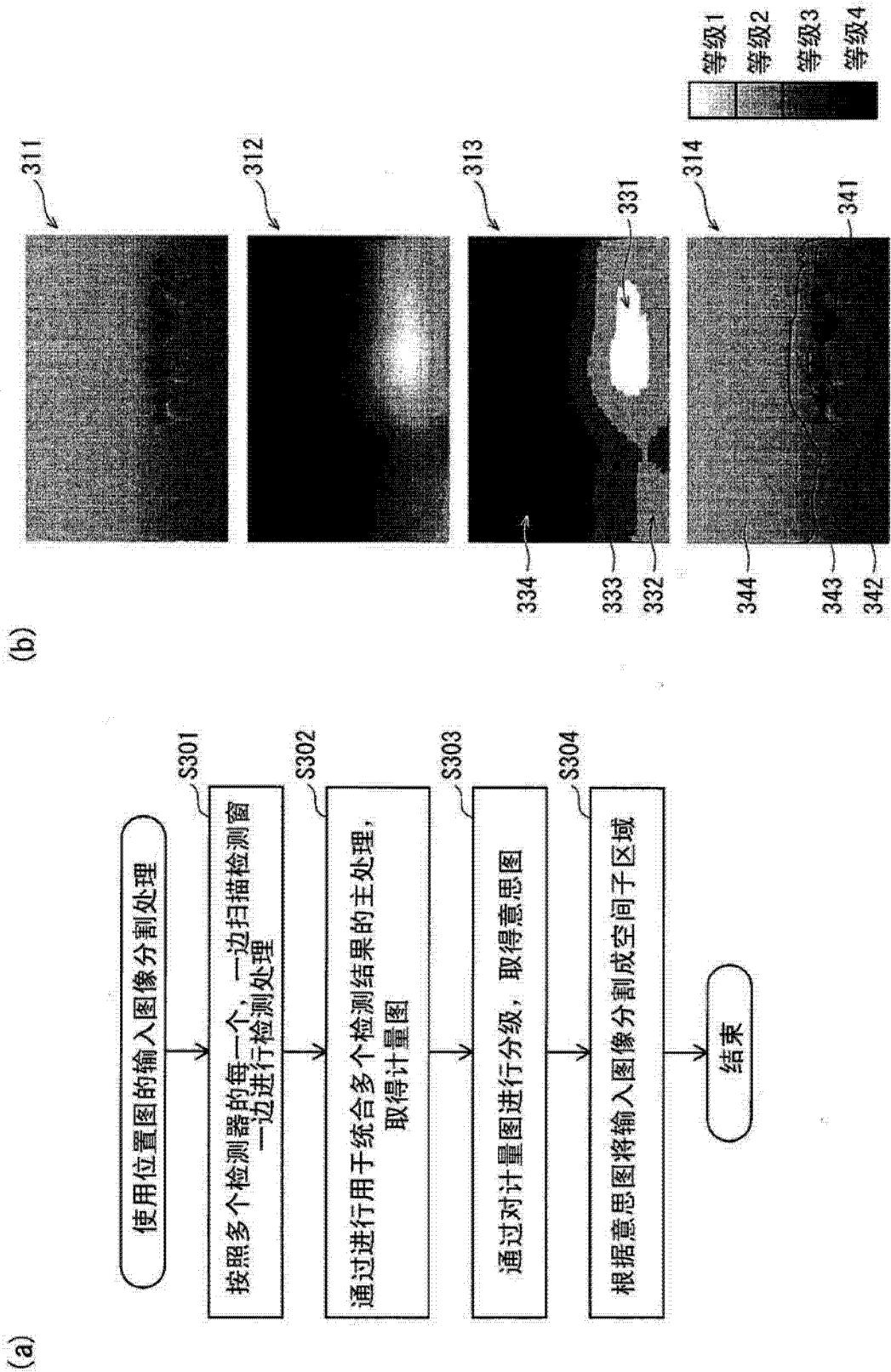


图 3

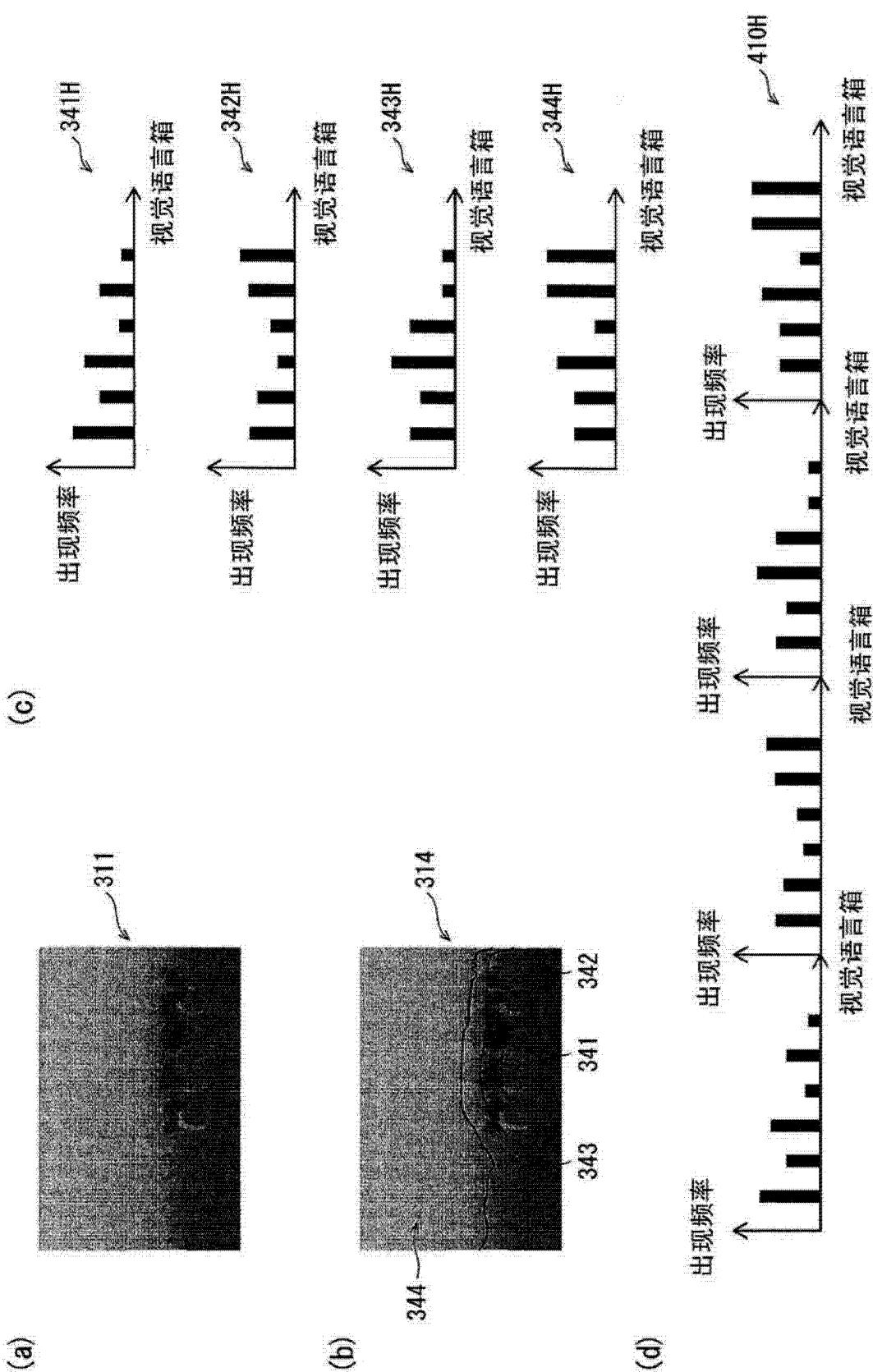
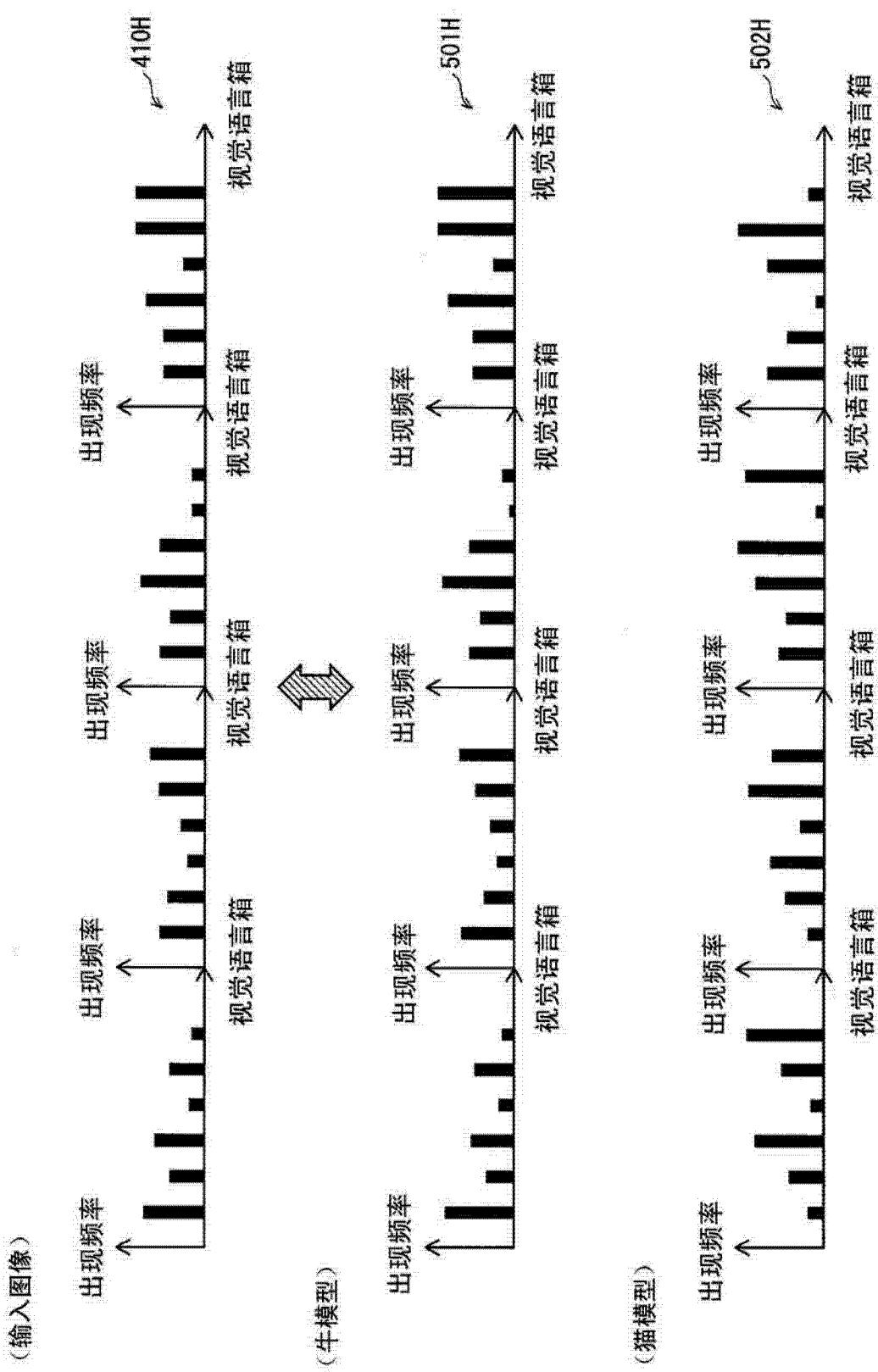


图 4



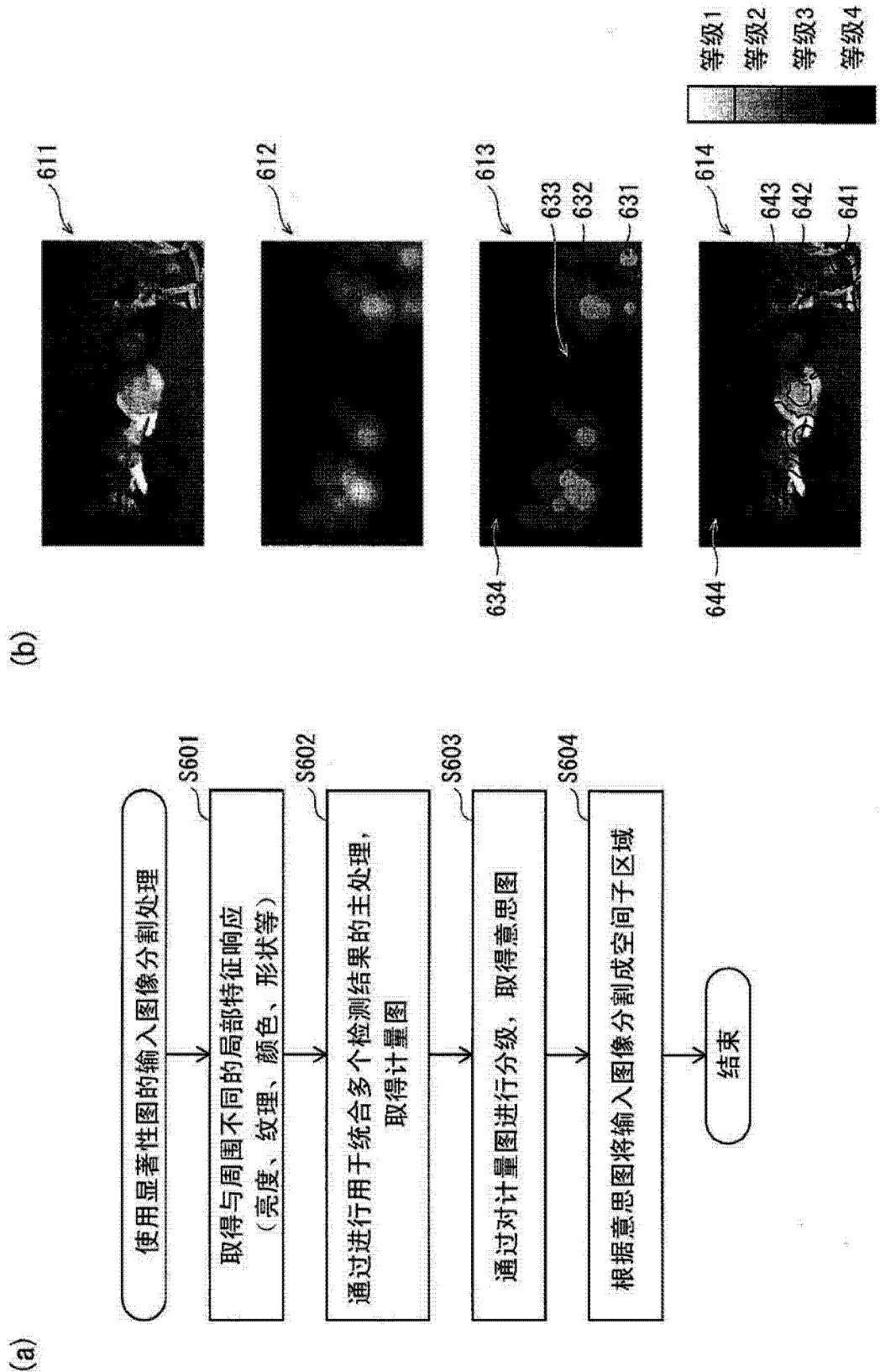


图 6

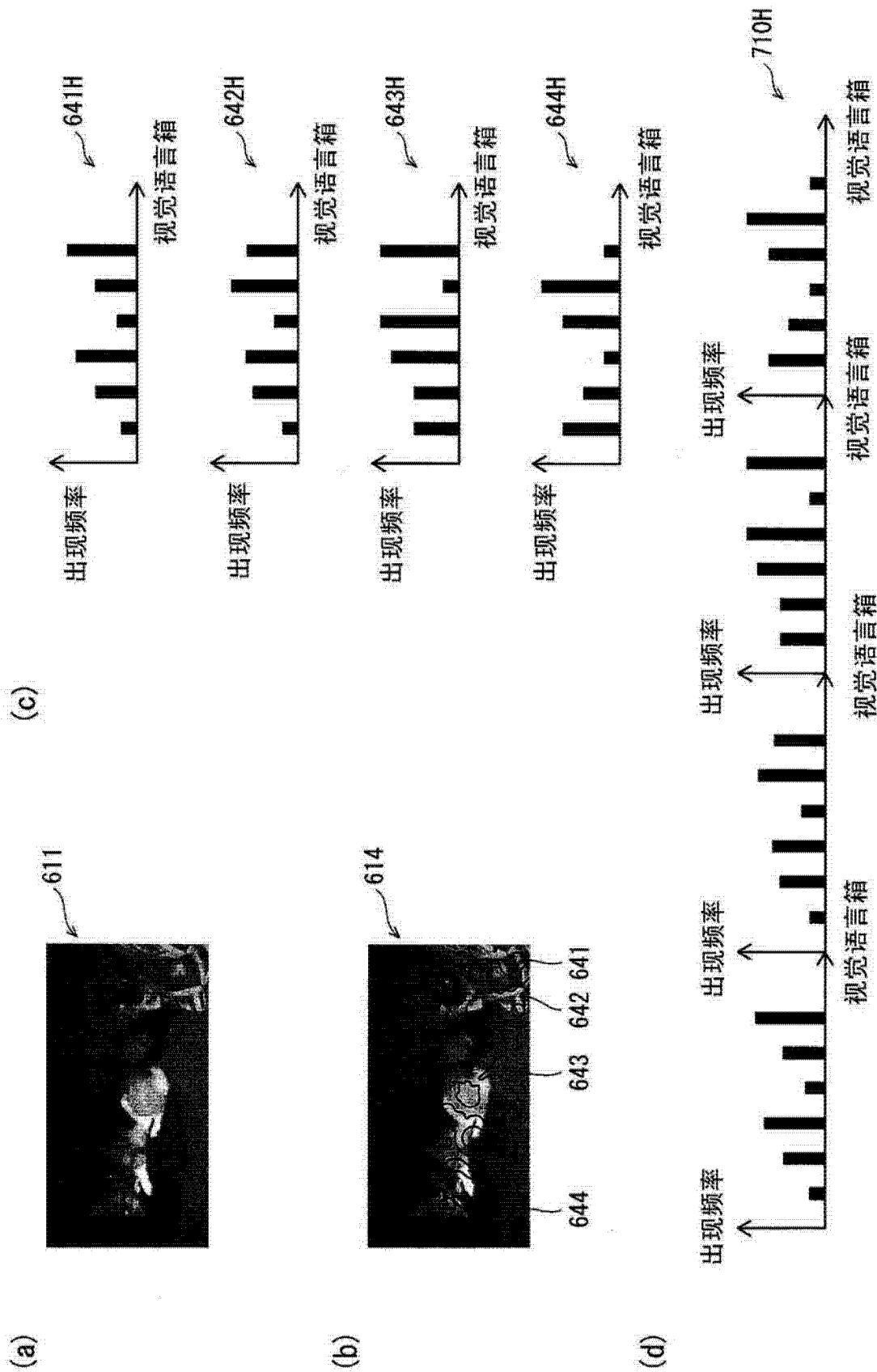


图 7

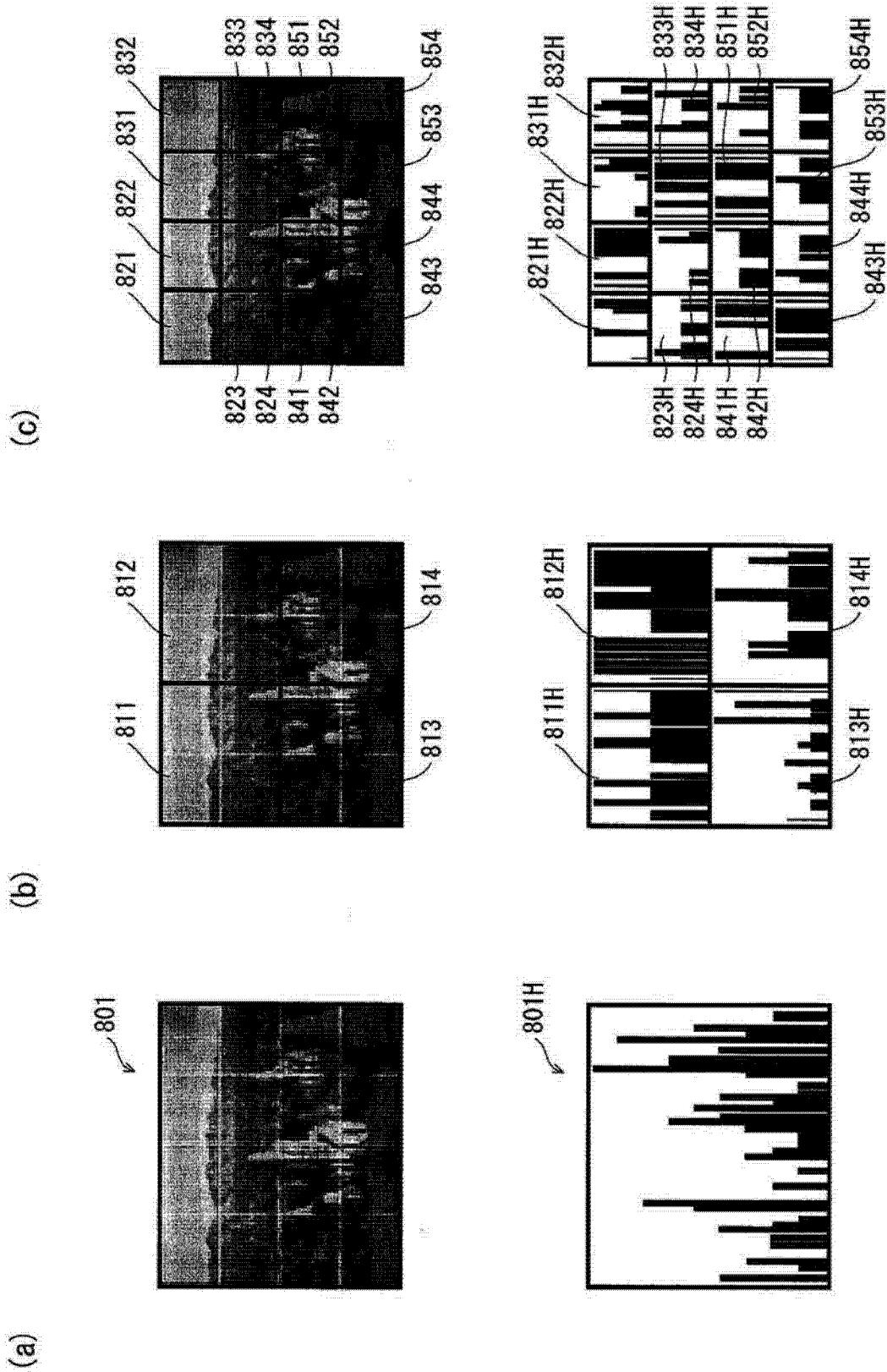


图 8