



(12)发明专利申请

(10)申请公布号 CN 108805148 A

(43)申请公布日 2018. 11. 13

(21)申请号 201710295810.8

(22)申请日 2017.04.28

(71)申请人 富士通株式会社

地址 日本神奈川县

(72)发明人 曹琼 刘汝杰

(74)专利代理机构 北京集佳知识产权代理有限公司

公司 11227

代理人 杜诚 李德山

(51)Int. Cl.

G06K 9/62(2006.01)

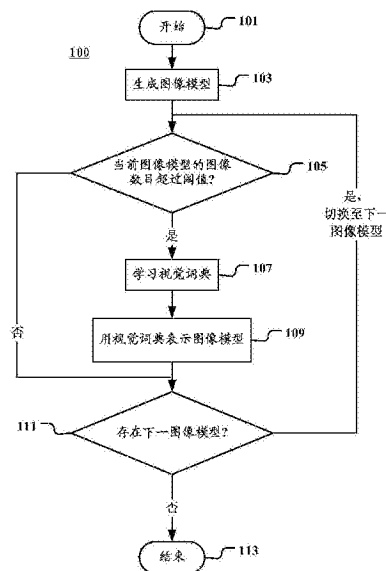
权利要求书2页 说明书12页 附图11页

(54)发明名称

处理图像的方法和用于处理图像的装置

(57)摘要

这里公开的示例性实施例涉及处理图像的方法和用于处理图像的装置。根据处理图像的方法,通过聚类多个图像来生成至少一个图像模型,其中每个图像模型由所述图像中彼此相似的图像来表示。如果表示一个图像模型的图像的数目超过阈值,则根据表示所述图像模型的图像学习视觉词典,并且用所述视觉词典代替表示所述图像模型的图像来表示所述图像模型。



1. 一种处理图像的方法,包括:

计算所述图像与至少一个图像模型之间的相似度;以及

将高于相似度阈值的较高相似度所对应的图像模型识别为所述图像所属的图像模型,

其中如果一个所述图像模型是由至少一个代表图像来表示的第一类型图像模型,则基于所述图像与所述代表图像之间的相似度来计算所述图像与所述图像模型之间的相似度,以及

如果一个所述图像模型由视觉词典来表示的第二类型图像模型,则基于所述图像的特征与所述视觉词典的视觉单词之间的相似度来计算所述图像与所述图像模型之间的相似度。

2. 如权利要求1所述的方法,其中所述相似度的计算包括:

如果一个所述图像模型是第一类型图像模型,则

针对所述图像的每个特征点,从所述图像模型的代表图像中选择一个代表图像中的特征点,其中所述图像的特征点的特征与所选择的代表图像的特征点的特征间的接近程度满足预定要求;以及

基于所述图像的各个特征点的特征与相应选择的特征点的特征之间的相似度来计算所述图像与所述图像模型之间的相似度。

3. 如权利要求1所述的方法,其中所述相似度的计算包括:

如果一个所述图像模型是第二类型图像模型,则

针对所述图像的每个特征点,从所述图像模型的视觉词典中选择视觉单词,其中所述图像的特征点的特征与所选择的视觉单词间的接近程度满足预定要求;以及

基于所述图像的各个特征点的特征与相应选择的视觉单词之间的相似度来计算所述图像与所述图像模型之间的相似度。

4. 如权利要求1所述的方法,其中所述识别包括:

如果存在至少两个图像模型具有较高相似度并且其相似度的接近程度满足预定条件,则将所述至少两个图像模型合并为一个图像模型;以及

将合并的图像模型识别为所述图像所属的图像模型,

其中如果所述至少两个图像模型之一由视觉词典来表示,则所述合并包括:

根据所述至少两个图像模型的表示来学习用于表示合并的图像模型的视觉词典。

5. 如权利要求4所述的方法,其中如果所述至少两个图像模型均由代表图像来表示并且所述代表图像的数目超过阈值,则所述合并包括:

根据表示所述至少两个图像模型的代表图像表示来学习用于表示合并的图像模型的视觉词典。

6. 如权利要求1、4或5所述的方法,还包括:

如果所识别的图像模型由视觉词典来表示,则根据表示所识别的图像模型的视觉词典和所述图像来学习用于表示所识别的图像模型的视觉词典。

7. 如权利要求1、4或5所述的方法,还包括:

如果所识别的图像模型由代表图像来表示并且所述代表图像和所述图像的总数超过阈值,则根据所述代表图像和所述图像学习视觉词典来代替表示所识别的图像模型的代表图像。

8. 如权利要求1、4或5所述的方法,还包括:

如果所识别的图像模型由代表图像来表示并且所述代表图像和所述图像的总数不超过阈值,则用所述代表图像和所述图像来表示所识别的图像模型。

9. 一种用于处理图像的装置,包括:

至少一个处理器,被配置成执行如权利要求1至8中任一个所述的方法。

10. 一种处理图像的方法,包括:

通过聚类多个图像来生成至少一个图像模型,其中每个图像模型由所述图像中彼此相似的图像来表示;以及

如果表示一个图像模型的图像的数目超过阈值,则根据表示所述图像模型的图像学习视觉词典,并且用所述视觉词典代替表示所述图像模型的图像来表示所述图像模型。

## 处理图像的方法和用于处理图像的装置

### 技术领域

[0001] 这里公开的示例性实施例涉及图像处理。更具体地,示例性实施例涉及图像的自动分类或识别。

### 背景技术

[0002] 随着数码产品、互联网等各个领域迅速发展,产生了大量亟待分析、识别、组织、分类和检索的图像内容。有效地识别图像信息,成了图像处理、机器视觉、模式识别、人工智能、神经科学等多个领域的研究热点。图像分类是其中的重要研究内容。

[0003] 图像分类是根据图像信息中所反映的不同特征将图像区分出不同类别的目标的图像处理方法。常见的图像分类方法可以分为监督分类方法和非监督分类方法。

[0004] 监督分类方法和非监督分类方法的区别在于是否利用训练数据来获取先验的类别知识。监督分类方法根据训练数据集提供的样本选择特征参数,建立判别函数,对待分类图像进行分类。因此,监督分类方法依赖于选定的训练数据。相比之下,非监督分类方法不需要更多的先验知识,而是只根据图像数据的自然聚类特性进行分类。因此,非监督分类方法简单且具有较高的准确度。非监督分类方法的一个示例是K均值(K-means)方法。

### 发明内容

[0005] 根据这里公开的一个示例性实施例,提供了一种处理图像的方法。根据该方法,通过聚类多个图像来生成至少一个图像模型,其中每个图像模型由所述图像中彼此相似的图像来表示。如果表示一个图像模型的图像的数目超过阈值,则根据表示所述图像模型的图像学习视觉词典,并且用所述视觉词典代替表示所述图像模型的图像来表示所述图像模型。

[0006] 根据这里公开的另一个示例性实施例,提供了一种处理图像的方法。根据该方法,计算所述图像与至少一个图像模型之间的相似度,以及将高于相似度阈值的较高相似度所对应的图像模型识别为所述图像所属的图像模型。如果一个所述图像模型是由至少一个代表图像来表示的第一类型图像模型,则基于所述图像与所述代表图像之间的相似度来计算所述图像与所述图像模型之间的相似度。如果一个所述图像模型由视觉词典来表示的第二类型图像模型,则基于所述图像的特征与所述视觉词典的视觉单词之间的相似度来计算所述图像与所述图像模型之间的相似度。

[0007] 根据这里公开的另一个示例性实施例,提供了一种用于处理图像的装置,包括至少一个处理器。该至少一个处理器被配置成执行如这里公开的示例性实施例的方法。

[0008] 下面参考附图详细描述本发明的示例性实施例的进一步特性和优点,以及本发明的示例性实施例的结构和操作。应当注意,本发明不限于这里描述的具体实施例。在这里出现这样的实施例只是出于说明的目的。相关领域技术人员根据这里包含的指导会想到其它实施例。

## 附图说明

[0009] 在附图中通过例子图解这里公开的示例性实施例,但这些例子不对本发明产生限制,图中用类似的附图标记表示类似的元素,其中:

[0010] 图1是用于说明根据一个示例性实施例的生成图像模型的方法的流程图;

[0011] 图2是用于说明根据一个示例性实施例的图像分类方法的流程图;

[0012] 图3是用于说明根据一个示例性实施例的相似度计算方法的流程图;

[0013] 图4是用于说明根据另一个示例性实施例的相似度计算方法的流程图;

[0014] 图5是用于说明根据另一个示例性实施例的图像分类方法的流程图;

[0015] 图6是用于说明根据一个示例性实施例的图像分类判断算法的伪码;

[0016] 图7是用于说明根据一个示例性实施例的图像模型合并方法的流程图;

[0017] 图8是用于说明作为对图2的示例性实施例的变型的图像分类方法的流程图;

[0018] 图9是用于说明作为对图5的示例性实施例的变型的图像分类方法的流程图;

[0019] 图10是用于说明根据一个示例性实施例的图像模型更新方法的流程图;

[0020] 图11是说明用于实现这里公开的示例性实施例的各个方面的示例性系统的框图。

## 具体实施方式

[0021] 下面参考附图描述这里公开的示例性实施例。应当注意,出于清楚的目的,在附图和描述中省略了有关所属技术领域技术人员知道但是与示例性实施例无关的部分和过程的表示和说明。

[0022] 本领域的技术人员可以理解,示例性实施例的各方面可以被实施为系统、方法或计算机程序产品。因此,示例性实施例的各个方面可以具体实现为以下形式,即,可以是完全硬件实施例、完全软件实施例(包括固件、驻留软件、微代码等)、或组合软件部分与硬件部分的实施例,本文可以一般称为“电路”、“模块”或“系统”。此外,示例性实施例的各个方面可以采取体现为一个或多个计算机可读介质的计算机程序产品的形式,该计算机可读介质上面体现有计算机可读程序代码。可以例如通过计算机网络来分发计算机程序,或者计算机程序可以位于一个或多个远程服务器上,或被嵌入设备的存储器中。

[0023] 可以使用一个或多个计算机可读介质的任何组合。计算机可读介质可以是计算机可读信号介质或计算机可读存储介质。计算机可读存储介质例如可以是,但不限于电的、磁的、光的、电磁的、红外线的、或半导体的系统、设备或装置、或前述各项的任何适当的组合。计算机可读存储介质的更具体的例子(非穷举的列表)包括以下:有一个或多个导线的电连接、便携式计算机磁盘、硬盘、随机存取存储器(RAM)、只读存储器(ROM)、可擦式可编程只读存储器(EPROM或闪存)、光纤、便携式紧凑磁盘只读存储器(CD-ROM)、光存储装置、磁存储装置、或前述各项的任何适当的组合。在本文语境中,计算机可读存储介质可以是任何含有或存储供指令执行系统、设备或装置使用的或与指令执行系统、设备或装置相联系的程序的有形介质。

[0024] 计算机可读信号介质可以包括例如在基带中或作为载波的一部分传播的其中带有计算机可读程序代码的数据信号。这样的传播信号可以采取任何适当的形式,包括但不限于电磁的、光的或其任何适当的组合。

[0025] 计算机可读信号介质可以是不同于计算机可读存储介质的、能够传达、传播或传输供指令执行系统、设备或装置使用的或与指令执行系统、设备或装置相联系的程序的任何一种计算机可读介质。

[0026] 体现在计算机可读介质中的程序代码可以采用任何适当的介质传输,包括但不限于无线、有线、光缆、射频等等、或上述各项的任何适当的组合。

[0027] 用于执行这里公开的示例性实施例的各方面的操作的计算机程序代码可以以一种或多种程序设计语言的任何组合来编写,所述程序设计语言包括面向对象的程序设计语言,诸如Java、Smalltalk、C++之类,还包括常规的过程式程序设计语言,诸如“C”程序设计语言或类似的程序设计语言。

[0028] 以下参照按照示例性实施例的方法、设备(系统)和计算机程序产品的流程图和/或框图来描述这里公开的示例性实施例的各个方面。应当理解,流程图和/或框图的每个方框以及流程图和/或框图中各方框的组合都可以由计算机程序指令实现。这些计算机程序指令可以提供给通用计算机、专用计算机或其它可编程数据处理设备的处理器以生产出一种机器,使得通过计算机或其它可编程数据处理设备执行的这些指令产生用于实现流程图和/或框图中的方框中规定的功能/操作的装置。

[0029] 也可以把这些计算机程序指令存储在能够指引计算机或其它可编程数据处理设备以特定方式工作的计算机可读介质中,使得存储在计算机可读介质中的指令产生一个包括实现流程图和/或框图中的方框中规定的功能/操作的指令的制造品。

[0030] 也可以把计算机程序指令加载到计算机或其它可编程数据处理设备上,导致在计算机或其它可编程数据处理设备上执行一系列操作步骤以产生计算机实现的过程,使得在计算机或其它可编程设备上执行的指令提供实现流程图和/或框图中的方框中规定的功能/操作的过程。

[0031] 图1是用于说明根据一个示例性实施例的生成图像模型的方法100的流程图。

[0032] 如图1所示,方法100从步骤101开始。在步骤103,通过聚类多个图像 $I_1-I_M$ 来生成至少一个图像模型 $O_1-O_N$ ,  $N \geq 1$ 。聚类方法根据图像间的相似性或图像彼此间在特征空间中的距离来对图像进行分类。在进行聚类时,以使某种聚类准则达到极值为目标,从而得到图像的聚类结果。聚类方法的示例包括但不限于迭代的动态聚类算法(例如C均值算法和ISODATA算法)和非迭代的分级聚类算法。经过聚类,图像 $I_1-I_M$ 被分成不同的组。每个组中的图像彼此相似或在特征空间中距离接近。这样的图像组被称为图像模型。每个图像模型 $O_i$ 可以由其所包含的图像 $I_j \in O_i$ (也称为图像模型 $O_i$ 的代表图像)来表示。在本公开中也将这样的图像模型称为第一类型的图像模型。每个图像模型 $O_i$ 可以被存储为由其所包含的图像 $I_j$ 本身,也可以存储为从每个图像 $I_j$ 中提取的各个特征点的特征。

[0033] 在步骤105,确定通过步骤103获得的图像模型中的当前图像模型 $O_k$ 的图像数目是否超过阈值。如果图像数目没有超过阈值,则方法100前进至步骤111。如果图像数目超过阈值,则在步骤107根据表示图像模型 $O_k$ 的图像学习视觉词典。例如可以通过从表示图像模型 $O_k$ 的图像中提取特征并对提取的特征进行聚类来得到图像模型 $O_k$ 的视觉词典。

[0034] 在一个示例中,可以提取多个属性作为图像特征,以在图像模型中嵌入不同的线索。例如,所提取的特征可以包括尺度不变特征变换(SIFT)特征和/或颜色名称(CN)特征。

[0035] 在所提取的特征包括颜色名称特征的一个示例中,可以将图像中局部块的颜色名

称特征计算为该局部块内所有像素的颜色名称特征的均值。以图像检索系统为例,在图像检索系统中通常使用描述局部梯度分布的尺度不变特征变换(SIFT)特征,并且采用倒排索引来建立基于词袋模型(BoW)的图像检索系统,其中每个条目对应于在SIFT特征的码本中定义的视觉单词。然而,对SIFT特征的依赖导致对图像的其他特性(例如颜色)的忽视。这个问题与量化期间的信息丢失一起导致许多虚警匹配。为了增强SIFT视觉单词的辨别能力,可以采用颜色名称特征,其为每个像素分配11-D向量。在每个检测到的特征点周围,可以得到具有与特征点尺度成比例的面积的局部块。然后,计算该区域中的每个像素的CN向量,并且计算平均CN向量作为颜色特征。

[0036] 在步骤109,代替用包含的图像来表示图像模型的方式,用学习的视觉词典来表示图像模型。在本公开中也将这样的图像模型称为第二类型的图像模型。当图像模型包含的图像数目较多时,在实际应用中图像或图像特征的存储和处理占用需要较多的开销。相比之下,采用视觉词典来表示图像模型会降低应用中的开销。

[0037] 在步骤111,确定在通过步骤103获得的图像模型中是否存在下一个未处理的图像模型。如果存在下一个未处理的图像模型,则将该下一个未处理的图像模型设置为当前图像模型并且方法100返回至步骤105。如果不存在下一个未处理的图像模型,则方法100在步骤113结束。

[0038] 图2是用于说明根据一个示例性实施例的图像分类方法200的流程图。在这个示例性实施例的应用场景中,在代表不同分类的图像模型 $O_1-O_N$ ,  $N \geq 1$ 中识别查询图像 $q$ 所属的图像模型,从而将所识别的图像模型的分类识别为查询图像 $q$ 所属的分类。

[0039] 如图2所示,方法200从步骤201开始。在步骤203,确定图像模型 $O_1-O_N$ 中的当前图像模型 $O_k$ 是第一类型还是第二类型。

[0040] 如果当前图像模型 $O_k$ 是第一类型的图像模型,则在步骤205基于查询图像 $q$ 与当前图像模型 $O_k$ 的代表图像之间的相似度来计算查询图像 $q$ 与当前图像模型 $O_k$ 之间的相似度。

[0041] 如果当前图像模型 $O_k$ 是第二类型的图像模型,则在步骤207基于查询图像 $q$ 的特征与当前图像模型 $O_k$ 的视觉词典的视觉单词之间的相似度来计算查询图像 $q$ 与当前图像模型 $O_k$ 之间的相似度。

[0042] 在计算出相似度之后,接着在步骤209,确定计算的相似度是否高于相似度阈值。如果计算的相似度不高于相似度阈值,则方法200前进至步骤213。如果计算的相似度高于相似度阈值,则在步骤211将相应的当前图像模型 $O_k$ 识别为查询图像 $q$ 所属的图像模型。接着方法200前进至步骤215。

[0043] 在一个示例中,可以利用欧几里德距离来测量相似度。在另一个示例中,可以使用汉明距离测量相似度。较小的距离表示较高的相关度。

[0044] 在一个示例中,图像特征包括CN特征。考虑到CN特征的每个维度具有明确的语义含义,可以对CN特征的每个维度进行二值化以产生二进制特征。

[0045] 在一个示例中,可以将查询图像的所有特征(例如SIFT特征和/或CN特征)的相似度进行诸如算术平均或加权求和的组合以获得最终相似度。

[0046] 在步骤213,确定图像模型 $O_1-O_N$ 中是否存在下一个未处理的图像模型。如果存在下一个未处理的图像模型,则将该下一个未处理的图像模型设置为当前图像模型并且方法200返回至步骤203。如果不存在下一个未处理的图像模型,则方法200在步骤215结束。

[0047] 在图2所示的方法中,将首先在步骤209识别的相似度高于相似度阈值的当前图像模型识别为查询图像所属的图像模型。在一个示例性的修改实施例中,可以在计算出查询图像与所有图像模型间的相似度之后将高于相似度阈值的相似度中一个最高相似度的相应图像模型识别为查询图像所属的图像模型。在另一个示例性的修改实施例中,可以在计算出查询图像与所有图像模型间的相似度之后将高于相似度阈值的相似度中至少两个相同的最高相似度的相应图像模型识别为查询图像所属的图像模型。在另一个示例性的修改实施例中,可以在计算出查询图像与所有图像模型间的相似度之后将高于相似度阈值的相似度中彼此满足预定接近标准并且高于其它相似度的至少两个相似度的相应图像模型识别为查询图像所属的图像模型。

[0048] 图3是用于说明根据一个示例性实施例的步骤205的相似度计算方法的流程图。

[0049] 如图3所示,方法300从步骤301开始。在步骤303,从查询图像 $q$ 中识别特征点 $p_1-p_L$ 并且提取出特征点的特征。在步骤305,针对查询图像 $q$ 的当前特征点 $p_t$ ,从当前图像模型 $O_k$ 的代表图像中选择一个代表图像中的特征点,其中特征点 $p_t$ 的特征与所选择的代表图像的特征点的特征间的接近程度满足预定要求,例如相似度高于阈值水平、距离低于阈值水平、相似度最高、距离最低等等。于是计算特征点 $p_t$ 的特征与相应选择的特征点的特征之间的相似度 $S_{t,k}$ 。

[0050] 在步骤307,确定查询图像 $q$ 的特征点中是否存在下一个未处理的特征点。如果查询图像 $q$ 的特征点中存在下一个未处理的特征点,则切换到下一个未处理的特征点并且方法300返回到步骤305。如果查询图像 $q$ 的特征点中不存在下一个未处理的特征点,则在步骤309,基于查询图像 $q$ 的特征点 $p_1-p_L$ 的特征与相应选择的特征点的特征之间的相似度 $S_{1,k}-S_{L,k}$ 来计算查询图像 $q$ 与当前图像模型 $O_k$ 之间的相似度 $S_k$ 。例如,相似度 $S_k$ 可以被计算为各个特征点的特征的相似度 $S_{1,k}-S_{L,k}$ 的算术平均或加权平均。

[0051] 在步骤311,确定图像模型 $O_1-O_N$ 中是否存在下一个未处理的图像模型。如果图像模型 $O_1-O_N$ 中存在下一个未处理的图像模型,则切换到下一个未处理的图像模型并且方法300返回到步骤305。如果图像模型 $O_1-O_N$ 中不存在下一个未处理的图像模型,则方法在步骤313结束。

[0052] 本领域的技术人员能够明白,步骤303的提取特征的操作也可以在方法300的流程之外执行,例如在方法200的流程中步骤205的计算相似度的操作之前的时刻执行。

[0053] 图4是用于说明根据一个示例性实施例的步骤207的相似度计算方法的流程图。

[0054] 如图4所示,方法400从步骤401开始。在步骤403,从查询图像 $q$ 中识别特征点 $p_1-p_L$ 并且提取出特征点的特征。在步骤405,针对查询图像 $q$ 的当前特征点 $p_t$ ,从当前图像模型 $O_k$ 的视觉词典中选择视觉单词,其中特征点 $p_t$ 的特征与所选择的视觉单词间的接近程度满足预定要求,例如相似度高于阈值水平、距离低于阈值水平、相似度最高、距离最低等等。于是计算特征点 $p_t$ 的特征与相应选择的视觉单词之间的相似度 $S_{t,k}$ 。

[0055] 在步骤407,确定查询图像 $q$ 的特征点中是否存在下一个未处理的特征点。如果查询图像 $q$ 的特征点中存在下一个未处理的特征点,则切换到下一个未处理的特征点并且方法400返回到步骤405。如果查询图像 $q$ 的特征点中不存在下一个未处理的特征点,则在步骤409,基于查询图像 $q$ 的特征点 $p_1-p_L$ 的特征与相应选择的视觉单词之间的相似度 $S_{1,k}-S_{L,k}$ 来计算查询图像 $q$ 与当前图像模型 $O_k$ 之间的相似度 $S_k$ 。例如,相似度 $S_k$ 可以被计算为各个特征



点的特征的相似度 $S_{1,k}-S_{L,k}$ 的算术平均或加权和。

[0056] 在步骤411,确定图像模型 $O_1-O_N$ 中是否存在下一个未处理的图像模型。如果图像模型 $O_1-O_N$ 中存在下一个未处理的图像模型,则切换到下一个未处理的图像模型并且方法400返回到步骤405。如果图像模型 $O_1-O_N$ 中不存在下一个未处理的图像模型,则方法在步骤413结束。

[0057] 本领域的技术人员能够明白,步骤403的提取特征的操作也可以在方法400的流程之外执行,例如在方法200的流程中步骤207的计算相似度的操作之前的时刻执行。

[0058] 当根据查询图像与图像模型间的相似度来识别查询图像所属的图像模型时,可能存在查询图像与多个图像模型的相似度较高并且彼此接近的情况。在这样的情况下,可以像在结合图2描述的实施例中那样将这多个图像模型识别为查询图像所属的图像模型,也可以将这多个图像模型合并为一个图像模型,并且在经过合并的多个图像模型中识别查询图像所属的图像模型。

[0059] 图5是用于说明根据一个示例性实施例的图像分类方法500的流程图,其中包含合并图像模型的处理。

[0060] 如图5所示,方法500从步骤501开始。在步骤503,确定图像模型 $O_1-O_N$ 中的当前图像模型 $O_k$ 是第一类型还是第二类型。

[0061] 如果当前图像模型 $O_k$ 是第一类型的图像模型,则在步骤505基于查询图像 $q$ 与当前图像模型 $O_k$ 的代表图像之间的相似度来计算查询图像 $q$ 与当前图像模型 $O_k$ 之间的相似度。

[0062] 如果当前图像模型 $O_k$ 是第二类型的图像模型,则在步骤507基于查询图像 $q$ 的特征与当前图像模型 $O_k$ 的视觉词典的视觉单词之间的相似度来计算查询图像 $q$ 与当前图像模型 $O_k$ 之间的相似度。

[0063] 在计算出相似度之后,接着在步骤509确定图像模型 $O_1-O_N$ 中是否存在下一个未处理的图像模型。如果存在下一个未处理的图像模型,则将该下一个未处理的图像模型设置为当前图像模型并且方法500返回至步骤503。如果不存在下一个未处理的图像模型,则在步骤511确定是否存在至少两个图像模型具有较高相似度并且其相似度的接近程度满足预定条件。例如在高于相似度阈值的相似度中是否存在彼此满足预定接近标准并且高于其它相似度的至少两个相似度。如果不存在这样的至少两个相似度,则方法500前进到步骤515。如果存在这样的至少两个相似度,则在步骤513将这样的至少两个相似度的相应图像模型合并为一个图像模型,并且计算查询图像与合并图像模型间的相似度。

[0064] 在步骤515,将具有高于阈值的最高相似度的图像模型识别为查询图像所属的图像模型。接着方法500在步骤517结束。

[0065] 在步骤513的处理中,也可以将进行合并之后直接将合并的图像模型识别为查询图像所属的图像模型,而不需要计算查询图像与合并图像模型间的相似度并执行步骤515。

[0066] 图6是用于说明根据一个示例性实施例的图像分类判断算法的伪码,其中提供了合并和识别逻辑的具体示例。

[0067] 在图6所示的示例中,假设存在查询图像 $q$ 和 $n$ 个现有图像模型。所计算的查询图像与图像模型间的相似度按降序排列为 $S_{k1}, S_{k2}, \dots, S_{kn}$ ,其中 $k_j$ 是图像模型的编号。在图6所示的示例中,通过相似度间的比值 $R$ 来衡量相似度的接近程度。如果比值 $R$ 大于阈值 $th_2$ ,则确定相似度不接近;否则确定相似度接近。在图6所示的示例中,如果发现两个最高相似度接

近,则合并相应图像模型,然后重新开始迭代处理。

[0068] 图7是用于说明根据一个示例性实施例的图像模型合并方法700的流程图。

[0069] 如图7所示,方法700从步骤701开始。在步骤703,确定要合并的至少两个图像模型之一是否第二类型(由视觉词典来表示)的图像模型。如果要合并的至少两个图像模型都不是第二类型的图像模型,则方法700前进到步骤709。在步骤705,如果要合并的至少两个图像模型之一是第二类型的图像模型,则根据要合并的至少两个图像模型的表示来学习用于表示合并的图像模型的视觉词典。图像模型的表示既可能是图像代表图像本身,也可能是视觉词典。如果要合并的图像模型均由代表图像来表示,则从代表图像中提取出特征点的特征,并且对提取的特征进行聚类以学习出表示合并的图像模型的视觉词典。如果要合并的图像模型均由视觉词典来表示,则对这些视觉词典的视觉单词进行聚类以学习出表示合并的图像模型的视觉词典。如果要合并的图像模型既有由视觉词典表示的图像模型,又有用代表图像表示的图像模型,则从代表图像中提取出特征点的特征,并且对提取的特征和视频词典的视觉单词进行聚类以学习出表示合并的图像模型的视觉词典。接着在步骤707用学习的视觉词典来表示合并的图像模型。然后方法700在步骤713结束。

[0070] 在步骤709,确定要合并的图像模型的代表图像的数目是否超过阈值。如果要合并的图像模型的代表图像的数目超过阈值,则方法700前进到步骤705。如果要合并的图像模型的代表图像的数目未超过阈值,则在步骤711用要合并的图像模型的代表图像来表示合并的图像模型。接着方法700在步骤713结束。

[0071] 确定查询图像所属的图像模型是图像分类的基本功能。此外,也可以根据查询图像来更新图像模型。在识别查询图像所属的图像模型的情况下,可以通过将查询图像合并到所识别的图像模型中来进行更新。

[0072] 图8是用于说明作为对图2的示例性实施例的变型的图像分类方法800的流程图。

[0073] 如图8所示,方法800从步骤801开始。在步骤803,确定图像模型 $0_1-0_N$ 中的当前图像模型 $0_k$ 是第一类型还是第二类型。

[0074] 如果当前图像模型 $0_k$ 是第一类型的图像模型,则在步骤805基于查询图像 $q$ 与当前图像模型 $0_k$ 的代表图像之间的相似度来计算查询图像 $q$ 与当前图像模型 $0_k$ 之间的相似度。

[0075] 如果当前图像模型 $0_k$ 是第二类型的图像模型,则在步骤807基于查询图像 $q$ 的特征与当前图像模型 $0_k$ 的视觉词典的视觉单词之间的相似度来计算查询图像 $q$ 与当前图像模型 $0_k$ 之间的相似度。

[0076] 在计算出相似度之后,接着在步骤809,确定计算的相似度是否高于相似度阈值。如果计算的相似度不高于相似度阈值,则方法800前进至步骤813。如果计算的相似度高于相似度阈值,则在步骤811将相应的当前图像模型 $0_k$ 识别为查询图像 $q$ 所属的图像模型。在步骤815,通过将查询图像合并到所识别的图像模型中来更新所识别的图像模型。接着方法800前进至步骤817。如果在步骤809确定计算的相似度不高于相似度阈值,则方法800前进到步骤813。

[0077] 在步骤813,确定图像模型 $0_1-0_N$ 中是否存在下一个未处理的图像模型。如果存在下一个未处理的图像模型,则将该下一个未处理的图像模型设置为当前图像模型并且方法800返回至步骤803。如果不存在下一个未处理的图像模型,则方法800在步骤817结束。

[0078] 在前面描述的图像分类方法的示例性实施例的一个变型中,如果没有识别出查询

图像所属的图像模型,则可以建立新图像模型并且将新图像模型识别为查询图像所属的图像模型。这个新图像模型以查询图像作为代表图像。

[0079] 图9是用于说明作为对图5的示例性实施例的变型的图像分类方法900的流程图。

[0080] 如图9所示,方法900从步骤901开始。在步骤903,确定图像模型 $O_1-O_N$ 中的当前图像模型 $O_k$ 是第一类型还是第二类型。

[0081] 如果当前图像模型 $O_k$ 是第一类型的图像模型,则在步骤905基于查询图像 $q$ 与当前图像模型 $O_k$ 的代表图像之间的相似度来计算查询图像 $q$ 与当前图像模型 $O_k$ 之间的相似度。

[0082] 如果当前图像模型 $O_k$ 是第二类型的图像模型,则在步骤907基于查询图像 $q$ 的特征与当前图像模型 $O_k$ 的视觉词典的视觉单词之间的相似度来计算查询图像 $q$ 与当前图像模型 $O_k$ 之间的相似度。

[0083] 在计算出相似度之后,接着在步骤909确定图像模型 $O_1-O_N$ 中是否存在下一个未处理的图像模型。如果存在下一个未处理的图像模型,则将该下一个未处理的图像模型设置为当前图像模型并且方法900返回至步骤903。如果不存在下一个未处理的图像模型,则在步骤911确定是否存在至少两个图像模型具有较高相似度并且其相似度的接近程度满足预定条件。例如在高于相似度阈值的相似度中是否存在彼此满足预定接近标准并且高于其它相似度的至少两个相似度。如果不存在这样的至少两个相似度,则方法900前进到步骤915。如果存在这样的至少两个相似度,则在步骤913将这样的至少两个相似度的相应图像模型合并为一个图像模型,并且计算查询图像与合并图像模型间的相似度。

[0084] 在步骤915,将具有高于阈值的最高相似度的图像模型识别为查询图像所属的图像模型。在步骤917,通过将查询图像合并到所识别的图像模型中来更新所识别的图像模型。接着方法900在步骤919结束。

[0085] 在步骤913的处理中,也可以将进行合并之后直接将合并的图像模型识别为查询图像所属的图像模型,并且方法前进到步骤917。

[0086] 图10是用于说明根据一个示例性实施例的图像模型更新方法1000的流程图。

[0087] 如图10所示,方法1000从步骤1001开始。在步骤1003,确定要更新的图像模型是否第二类型(由视觉词典来表示)的图像模型。如果要更新的图像模型不是第二类型的图像模型,则方法1000前进到步骤1009。在步骤1005,如果要更新的图像模型是第二类型的图像模型,则根据从查询图像中提取出特征点的特征,并且对提取的特征和表示要更新的图像模型的视觉词典的视觉单词进行聚类以学习出表示更新后的图像模型的视觉词典。如果要更新的图像模型是第一类型(由代表图像表示)的图像模型,则根据从查询图像和代表图像中提取出特征点的特征,并且对提取的特征进行聚类以学习出表示更新后的图像模型的视觉词典。接着在步骤1007用学习的视觉词典来表示更新后的图像模型。然后方法1000在步骤1013结束。

[0088] 在步骤1009,确定要更新的图像模型的代表图像的数目+1是否超过阈值。如果要更新的图像模型的代表图像的数目+1超过阈值,则方法1000前进到步骤1005。如果要更新的图像模型的代表图像的数目+1未超过阈值,则在步骤1011用查询图像和要更新的图像模型的代表图像来表示更新后的图像模型。接着方法1000在步骤1013结束。

[0089] 图11是说明用于实现这里公开的示例性实施例的各个方面的示例性系统的框图。

[0090] 在图11中,中央处理单元(CPU) 1101根据只读存储器(ROM) 1102中存储的程序或从

存储部分1108加载到随机访问存储器 (RAM) 1103的程序执行各种处理。在RAM 1103中,也根据需要存储当CPU 1101执行各种处理等等时所需的数据。

[0091] CPU 1101、ROM 1102和RAM 1103经由总线1104彼此连接。输入/输出接口1105也连接到总线1104。

[0092] 下列部件连接到输入/输出接口1105:包括键盘、鼠标等等的输入部分1106;包括例如阴极射线管 (CRT)、液晶显示器 (LCD) 等等的显示器和扬声器等等的输出部分1107;包括硬盘等等的存储部分1108;和包括例如LAN卡、调制解调器等等的网络接口卡的通信部分1109。通信部分1109经由例如因特网的网络执行通信处理。

[0093] 根据需要,驱动器1110也连接到输入/输出接口1105。例如磁盘、光盘、磁光盘、半导体存储器等等的可移除介质1111根据需要被安装在驱动器1110上,使得从中读出的计算机程序根据需要被安装到存储部分1108。

[0094] 在通过软件实现上述步骤和处理的情况下,从例如因特网的网络或例如可移除介质1111的存储介质安装构成软件的程序。

[0095] 本文中所述的术语仅仅是为了描述特定实施例的目的,而非意图限定本发明。本文中所用的单数形式的“一”和“该”旨在也包括复数形式,除非上下文中明确地另行指出。还应理解,“包括”一词当在本说明书中使用,说明存在所指出的特征、整体、步骤、操作、单元和/或组件,但是并不排除存在或增加一个或多个其它特征、整体、步骤、操作、单元和/或组件,以及/或者它们的组合。

[0096] 以下权利要求中的对应结构、材料、操作以及所有功能性限定的装置或步骤的等同替换,旨在包括任何用于与在权利要求中具体指出的其它单元相组合地执行该功能的结构、材料或操作。前面对本发明进行的描述只是为了图解和描述,不被用来对具有公开形式的本发明进行详细定义和限制。对于所属技术领域的普通技术人员而言,在不偏离本发明范围和精神的情况下,显然可以作出许多修改和变型。对实施例的选择和说明,是为了最好地解释本发明的原理和实际应用,使所属技术领域的普通技术人员能够明了,本发明可以有适合所要的特定用途的具有各种改变的各种实施例。

[0097] 这里描述了下面的示范性实施例(均用“附记”表示)。

[0098] 附记1.一种处理图像的方法,包括:

[0099] 通过聚类多个图像来生成至少一个图像模型,其中每个图像模型由所述图像中彼此相似的图像来表示;以及

[0100] 如果表示一个图像模型的图像的数目超过阈值,则根据表示所述图像模型的图像学习视觉词典,并且用所述视觉词典代替表示所述图像模型的图像来表示所述图像模型。

[0101] 附记2.如附记1所述的方法,其中所述学习基于从表示所述图像模型的图像中提取的特征,并且所述特征包括尺度不变特征变换特征和/或颜色名称特征。

[0102] 附记3.如附记2所述的方法,其中在所述特征包括颜色名称特征的情况下,一个局部块的颜色名称特征被计算为所述局部块内所有像素的颜色名称特征的均值。

[0103] 附记4.一种处理图像的方法,包括:

[0104] 计算所述图像与至少一个图像模型之间的相似度;以及

[0105] 将高于相似度阈值的较高相似度所对应的图像模型识别为所述图像所属的图像模型,

[0106] 其中如果一个所述图像模型是由至少一个代表图像来表示的第一类型图像模型,则基于所述图像与所述代表图像之间的相似度来计算所述图像与所述图像模型之间的相似度,以及

[0107] 如果一个所述图像模型由视觉词典来表示的第二类型图像模型,则基于所述图像的特征与所述视觉词典的视觉单词之间的相似度来计算所述图像与所述图像模型之间的相似度。

[0108] 附记5.如附记4所述的方法,其中所述视觉词典所基于的特征包括尺度不变特征变换特征和/或颜色名称特征。

[0109] 附记6.如附记4所述的方法,其中所述相似度的计算包括:

[0110] 如果一个所述图像模型是第一类型图像模型,则

[0111] 针对所述图像的每个特征点,从所述图像模型的代表图像中选择一个代表图像中的特征点,其中所述图像的特征点的特征与所选择的代表图像的特征点的特征间的接近程度满足预定要求;以及

[0112] 基于所述图像的各个特征点的特征与相应选择的特征点的特征之间的相似度来计算所述图像与所述图像模型之间的相似度。

[0113] 附记7.如附记4所述的方法,其中所述相似度的计算包括:

[0114] 如果一个所述图像模型是第二类型图像模型,则

[0115] 针对所述图像的每个特征点,从所述图像模型的视觉词典中选择视觉单词,其中所述图像的特征点的特征与所选择的视觉单词间的接近程度满足预定要求;以及

[0116] 基于所述图像的各个特征点的特征与相应选择的视觉单词之间的相似度来计算所述图像与所述图像模型之间的相似度。

[0117] 附记8.如附记6或7所述的方法,其中所述相似度被计算为各个特征点的特征的相似度的加权和。

[0118] 附记9.如附记6或7所述的方法,其中所述相似度的计算所基于的特征包括尺度不变特征变换特征和/或颜色名称特征。

[0119] 附记10.如附记4所述的方法,其中所述识别包括:

[0120] 如果存在至少两个图像模型具有较高相似度并且其相似度的接近程度满足预定条件,则将所述至少两个图像模型合并为一个图像模型;以及

[0121] 将合并的图像模型识别为所述图像所属的图像模型,

[0122] 其中如果所述至少两个图像模型之一由视觉词典来表示,则所述合并包括:

[0123] 根据所述至少两个图像模型的代表图像来学习用于表示合并的图像模型的视觉词典。

[0124] 附记11.如附记10所述的方法,其中如果所述至少两个图像模型均由代表图像来表示并且所述代表图像的数目超过阈值,则所述合并包括:

[0125] 根据表示所述至少两个图像模型的代表图像表示来学习用于表示合并的图像模型的视觉词典。

[0126] 附记12.如附记10所述的方法,其中如果所述至少两个图像模型均由代表图像来表示并且所述代表图像的数目不超过阈值,则所述合并包括:

[0127] 用表示所述至少两个图像模型的代表图像来表示合并的图像模型。

[0128] 附记13.如附记4、10、11或12所述的方法,还包括:

[0129] 如果所识别的图像模型由视觉词典来表示,则根据表示所识别的图像模型的视觉词典和所述图像来学习用于表示所识别的图像模型的视觉词典。

[0130] 附记14.如附记4、10、11或12所述的方法,还包括:

[0131] 如果所识别的图像模型由代表图像来表示并且所述代表图像和所述图像的总数超过阈值,则根据所述代表图像和所述图像学习视觉词典来代替表示所识别的图像模型的代表图像。

[0132] 附记15.如附记4、10、11或12所述的方法,还包括:

[0133] 如果所识别的图像模型由代表图像来表示并且所述代表图像和所述图像的总数不超过阈值,则用所述代表图像和所述图像来表示所识别的图像模型。

[0134] 附记16.一种用于处理图像的装置,包括:

[0135] 至少一个处理器,被配置成:

[0136] 通过聚类多个图像来生成至少一个图像模型,其中每个图像模型由所述图像中彼此相似的图像来表示;以及

[0137] 如果表示一个图像模型的图像的数目超过阈值,则根据表示所述图像模型的图像学习视觉词典,并且用所述视觉词典代替表示所述图像模型的图像来表示所述图像模型。

[0138] 附记17.如附记16所述的装置,其中所述学习基于从表示所述图像模型的图像中提取的特征,并且所述特征包括尺度不变特征变换特征和/或颜色名称特征。

[0139] 附记18.如附记17所述的装置,其中在所述特征包括颜色名称特征的情况下,一个局部块的颜色名称特征被计算为所述局部块内所有像素的颜色名称特征的均值。

[0140] 附记19.一种用于处理图像的装置,包括:

[0141] 至少一个处理器,被配置成:

[0142] 计算所述图像与至少一个图像模型之间的相似度;以及

[0143] 将高于相似度阈值的较高相似度所对应的图像模型识别为所述图像所属的图像模型,

[0144] 其中如果一个所述图像模型是由至少一个代表图像来表示的第一类型图像模型,则基于所述图像与所述代表图像之间的相似度来计算所述图像与所述图像模型之间的相似度,以及

[0145] 如果一个所述图像模型由视觉词典来表示的第二类型图像模型,则基于所述图像的特征与所述视觉词典的视觉单词之间的相似度来计算所述图像与所述图像模型之间的相似度。

[0146] 附记20.如附记19所述的装置,其中所述视觉词典所基于的特征包括尺度不变特征变换特征和/或颜色名称特征。

[0147] 附记21.如附记19所述的装置,其中所述相似度的计算包括:

[0148] 如果一个所述图像模型是第一类型图像模型,则

[0149] 针对所述图像的每个特征点,从所述图像模型的代表图像中选择一个代表图像中的特征点,其中所述图像的特征点的特征与所选择的代表图像的特征点的特征间的接近程度满足预定要求;以及

[0150] 基于所述图像的各个特征点的特征与相应选择的特征点的特征之间的相似度来计算所述图像与所述图像模型之间的相似度。

- [0151] 附记22.如附记19所述的装置,其中所述相似度的计算包括:
- [0152] 如果一个所述图像模型是第二类型图像模型,则
- [0153] 针对所述图像的每个特征点,从所述图像模型的视觉词典中选择视觉单词,其中所述图像的特征点的特征与所选择的视觉单词间的接近程度满足预定要求;以及
- [0154] 基于所述图像的各个特征点的特征与相应选择的视觉单词之间的相似度来计算所述图像与所述图像模型之间的相似度。
- [0155] 附记23.如附记21或22所述的装置,其中所述相似度被计算为各个特征点的特征的相似度的加权和。
- [0156] 附记24.如附记21或22所述的装置,其中所述相似度的计算所基于的特征包括尺度不变特征变换特征和/或颜色名称特征。
- [0157] 附记25.如附记19所述的装置,其中所述识别包括:
- [0158] 如果存在至少两个图像模型具有较高相似度并且其相似度的接近程度满足预定条件,则将所述至少两个图像模型合并为一个图像模型;以及
- [0159] 将合并的图像模型识别为所述图像所属的图像模型,
- [0160] 其中如果所述至少两个图像模型之一由视觉词典来表示,则所述合并包括:
- [0161] 根据所述至少两个图像模型的代表图像来学习用于表示合并的图像模型的视觉词典。
- [0162] 附记26.如附记25所述的装置,其中如果所述至少两个图像模型均由代表图像来表示并且所述代表图像的数目超过阈值,则所述合并包括:
- [0163] 根据表示所述至少两个图像模型的代表图像来表示学习用于表示合并的图像模型的视觉词典。
- [0164] 附记27.如附记25所述的装置,其中如果所述至少两个图像模型均由代表图像来表示并且所述代表图像的数目不超过阈值,则所述合并包括:
- [0165] 用表示所述至少两个图像模型的代表图像来表示合并的图像模型。
- [0166] 附记28.如附记19、25、26或27所述的装置,其中所述处理器还被配置为:
- [0167] 如果所识别的图像模型由视觉词典来表示,则根据表示所识别的图像模型的视觉词典和所述图像来学习用于表示所识别的图像模型的视觉词典。
- [0168] 附记29.如附记19、25、26或27所述的装置,其中所述处理器还被配置为:
- [0169] 如果所识别的图像模型由代表图像来表示并且所述代表图像和所述图像的总数超过阈值,则根据所述代表图像和所述图像学习视觉词典来代替表示所识别的图像模型的代表图像。
- [0170] 附记30.如附记19、25、26或27所述的装置,其中所述处理器还被配置为:
- [0171] 如果所识别的图像模型由代表图像来表示并且所述代表图像和所述图像的总数不超过阈值,则用所述代表图像和所述图像来表示所识别的图像模型。

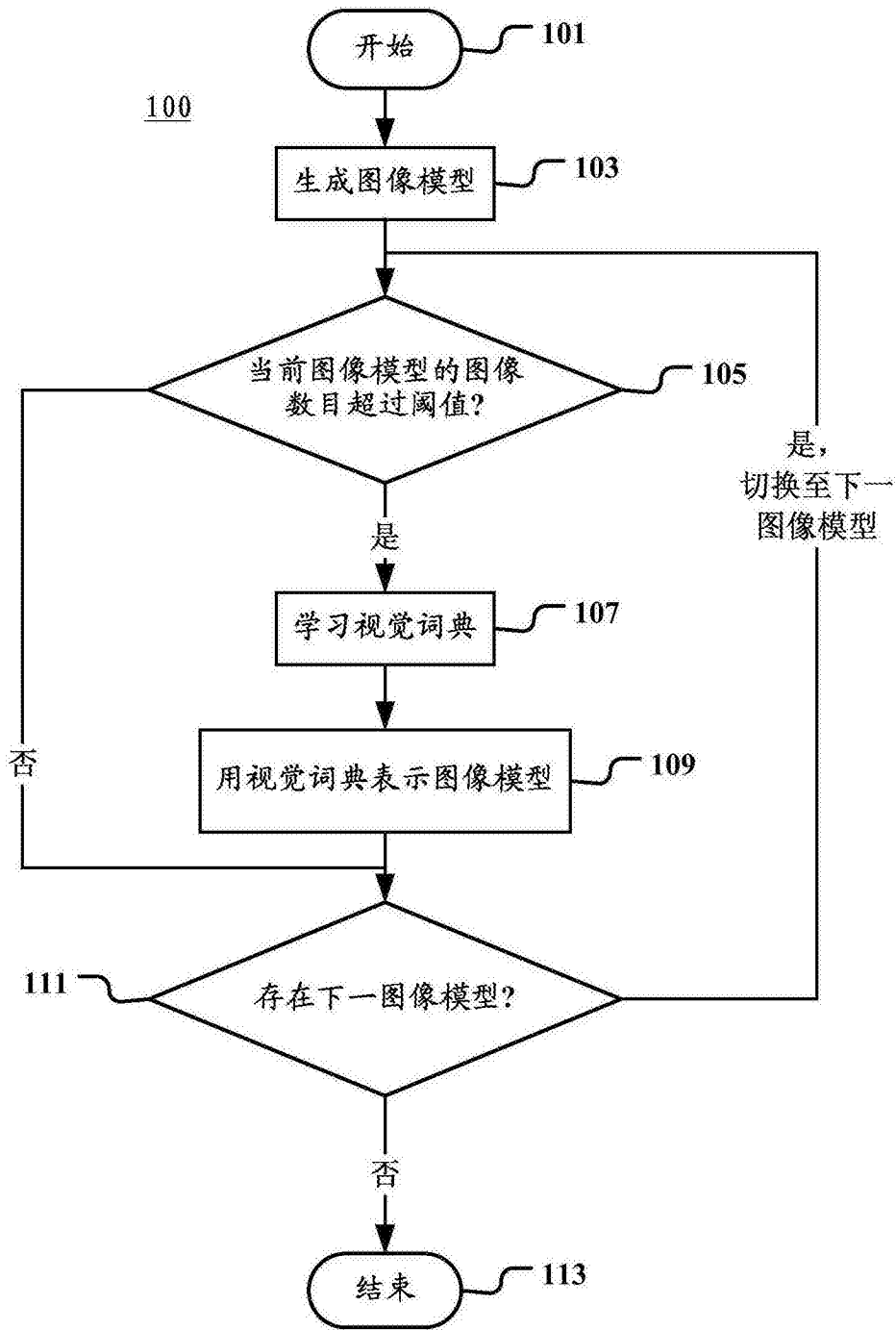


图1



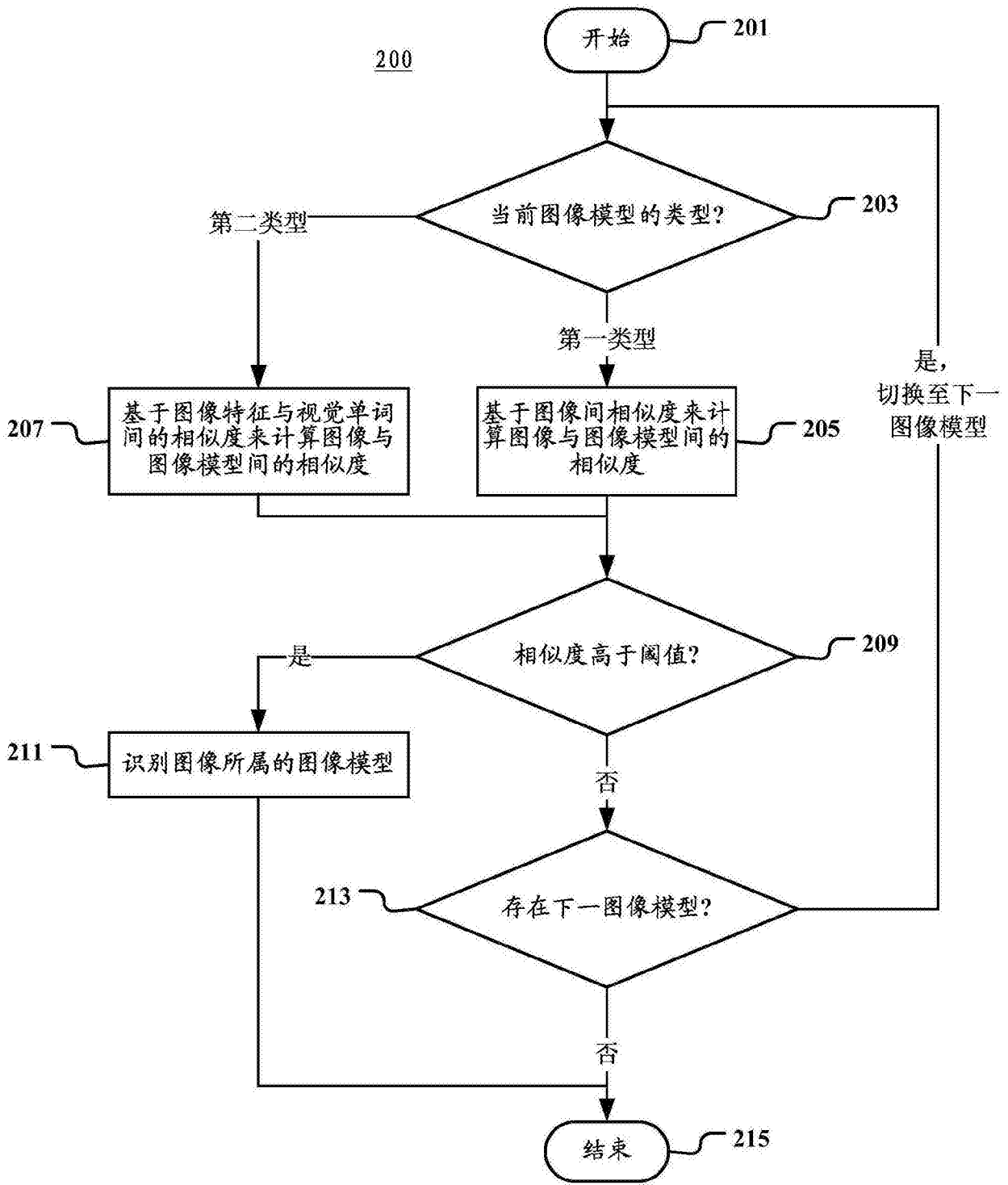


图2

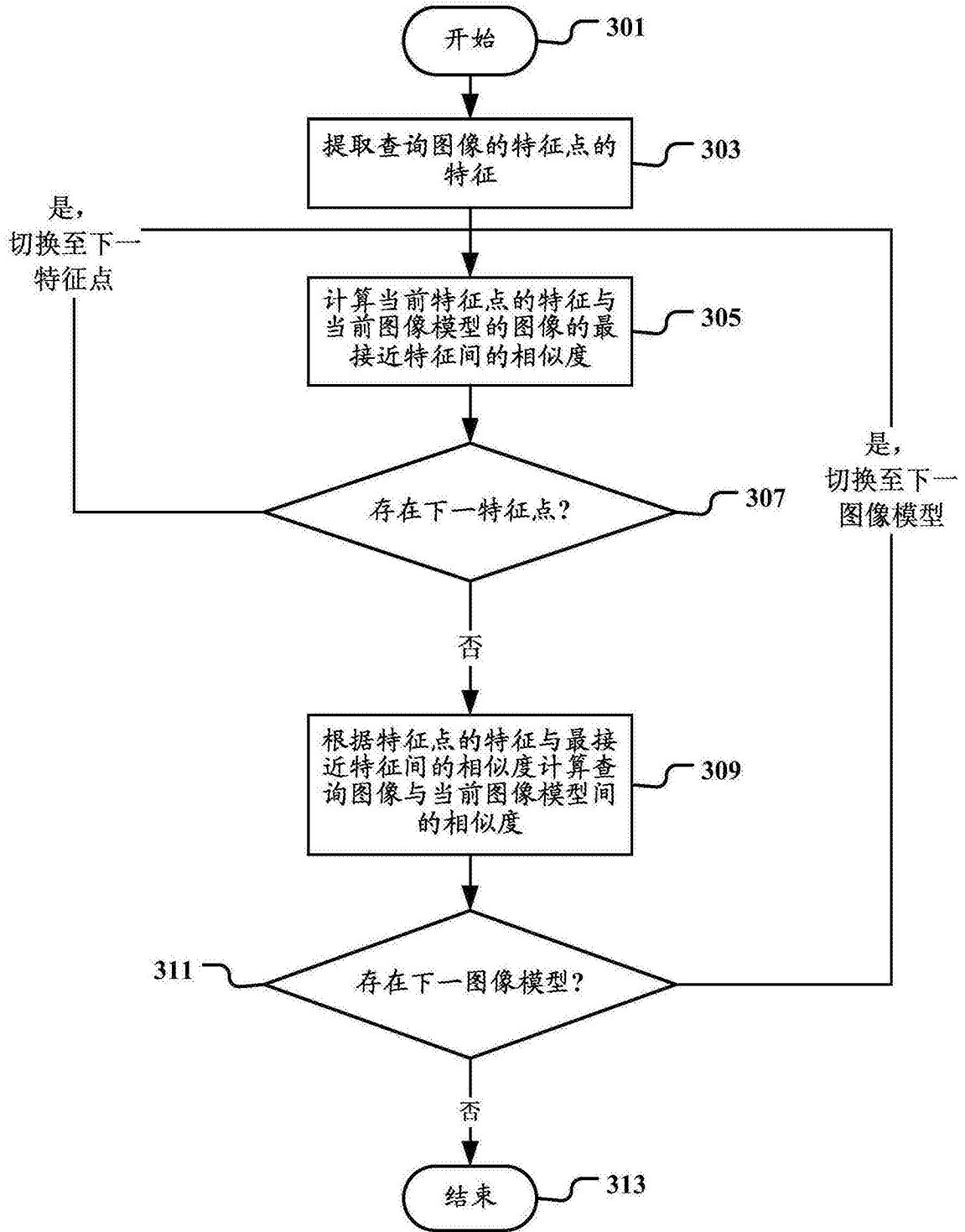


图3

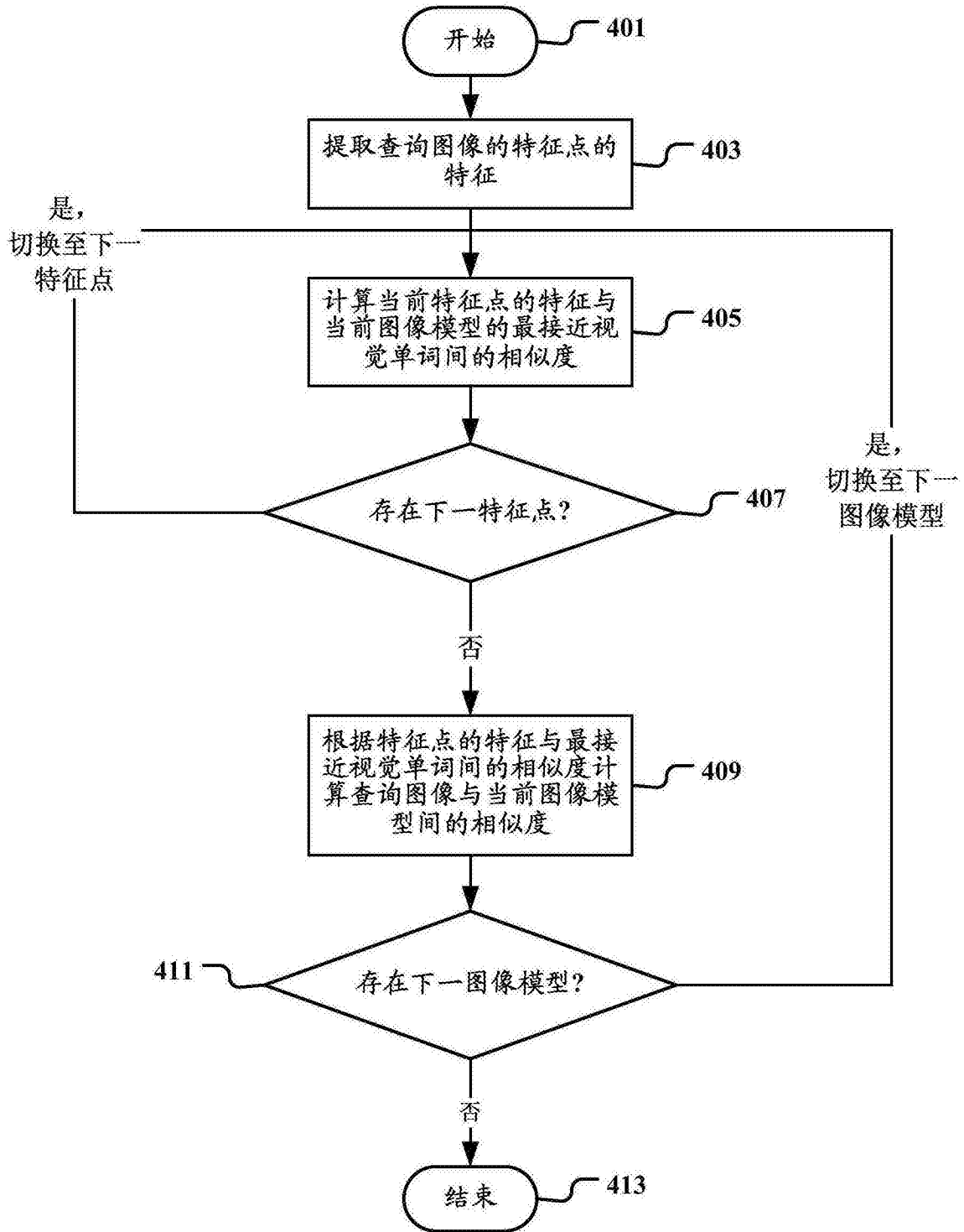


图4

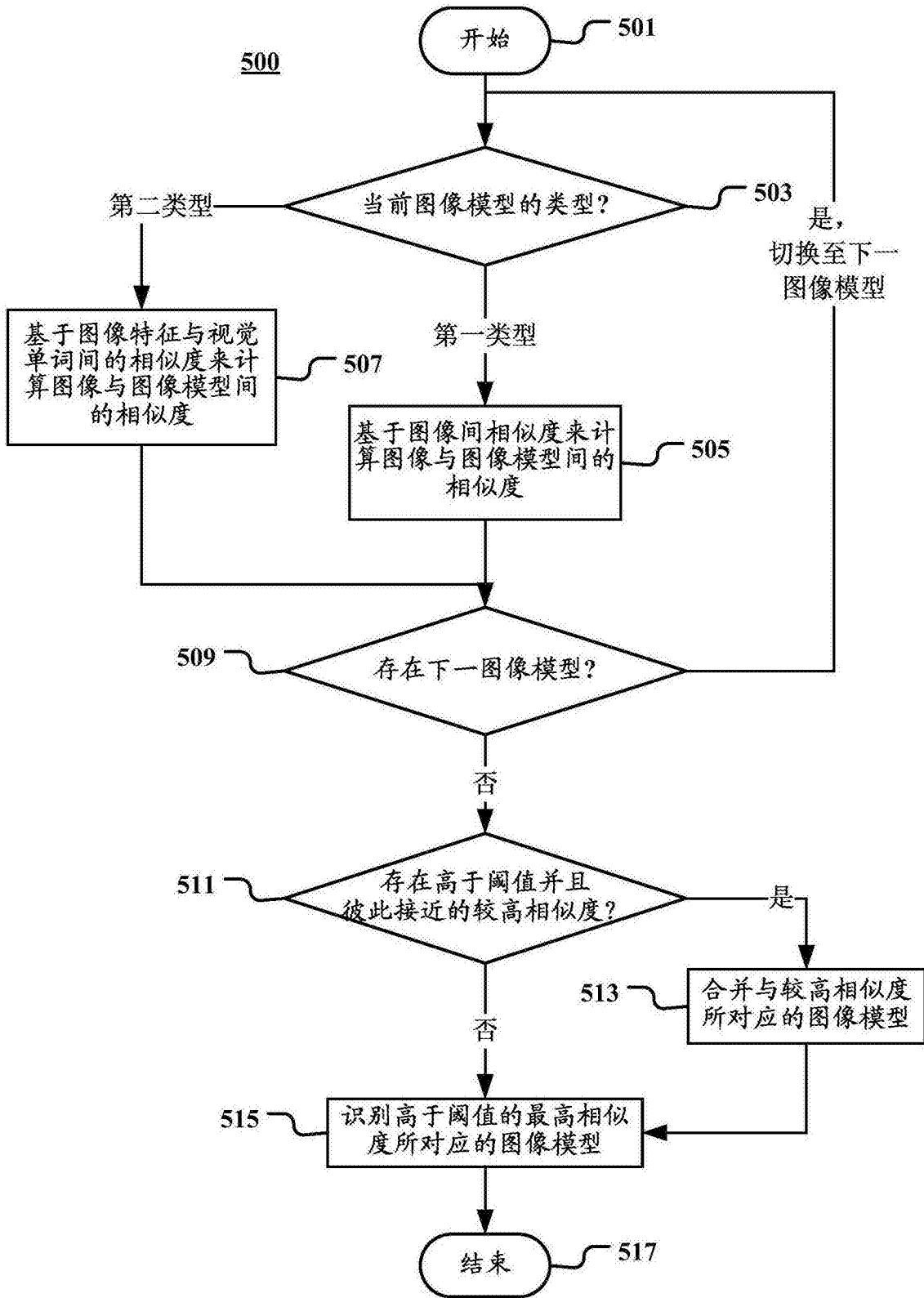


图5

```
i=1;
j=2;
while  $S_{ki} > th1$  and  $j < n$  //  $S_{ki}$ 表示查询图像 $q$ 与第 $k_i$ 个图像模型的相似度
{
    If ( $S_{kj} > th1$ )
    {
        计算  $R = S_{ki} / S_{kj}$ ,
        If  $R > th2$ ,
        {
            将查询图像 $q$ 识别为属于第 $k_j$ 个图像模型;
        }
        else
        {
            将第 $k_j$ 个图像模型合并到第 $k_i$ 个图像模型中;
            重新计算查询图像 $q$ 与第 $k_j$ 个图像模型的相似度 $S_{ki}$ ;
             $j=j+1$ ;
             $S_{ki} = S_{ki}$ ;
             $S_{kj} = S_{k(j+1)}$ ;
        }
    }
    else
    {
        将查询图像 $q$ 识别为属于第 $k_i$ 个图像模型;
    }
}
```

图6

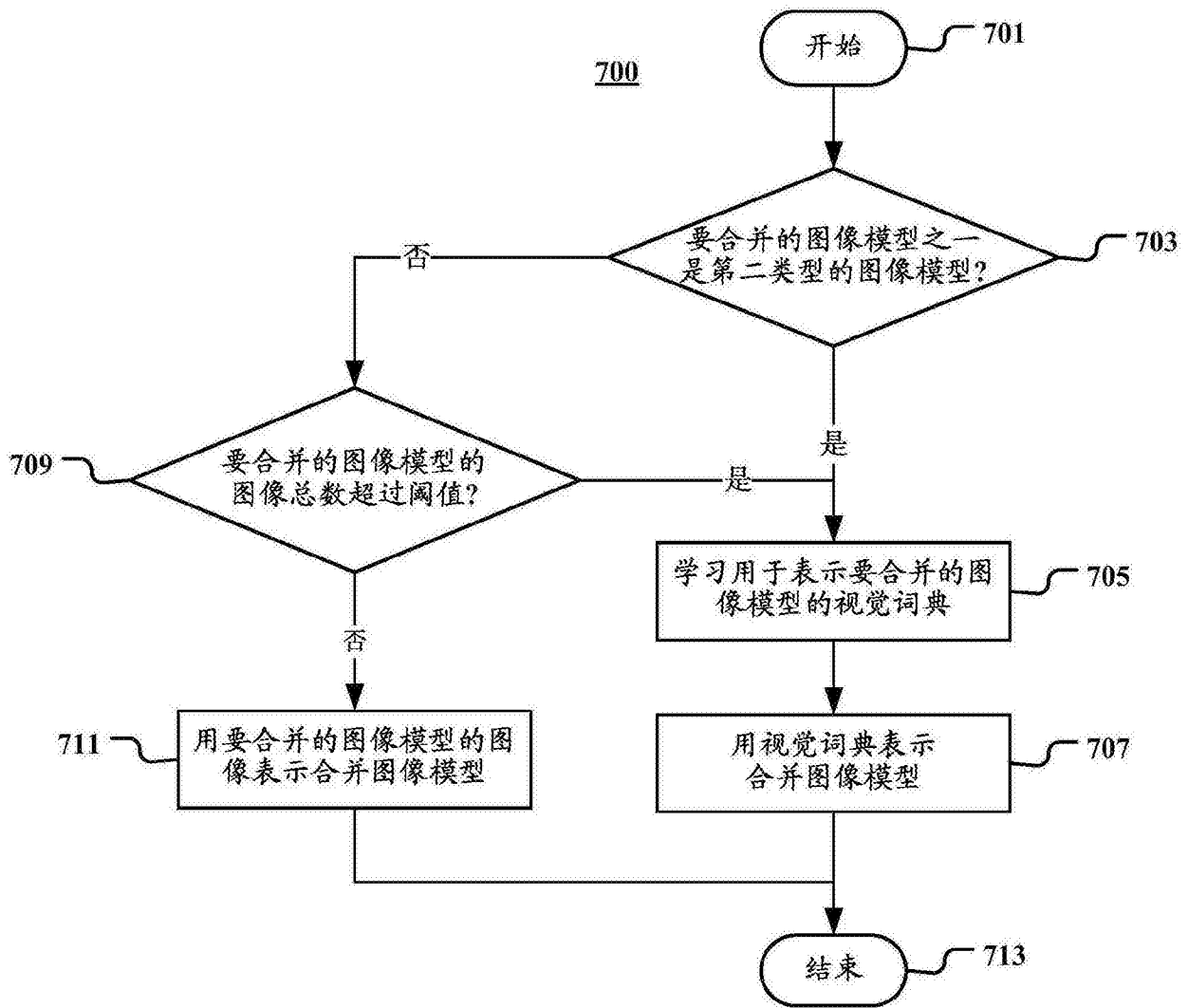


图7

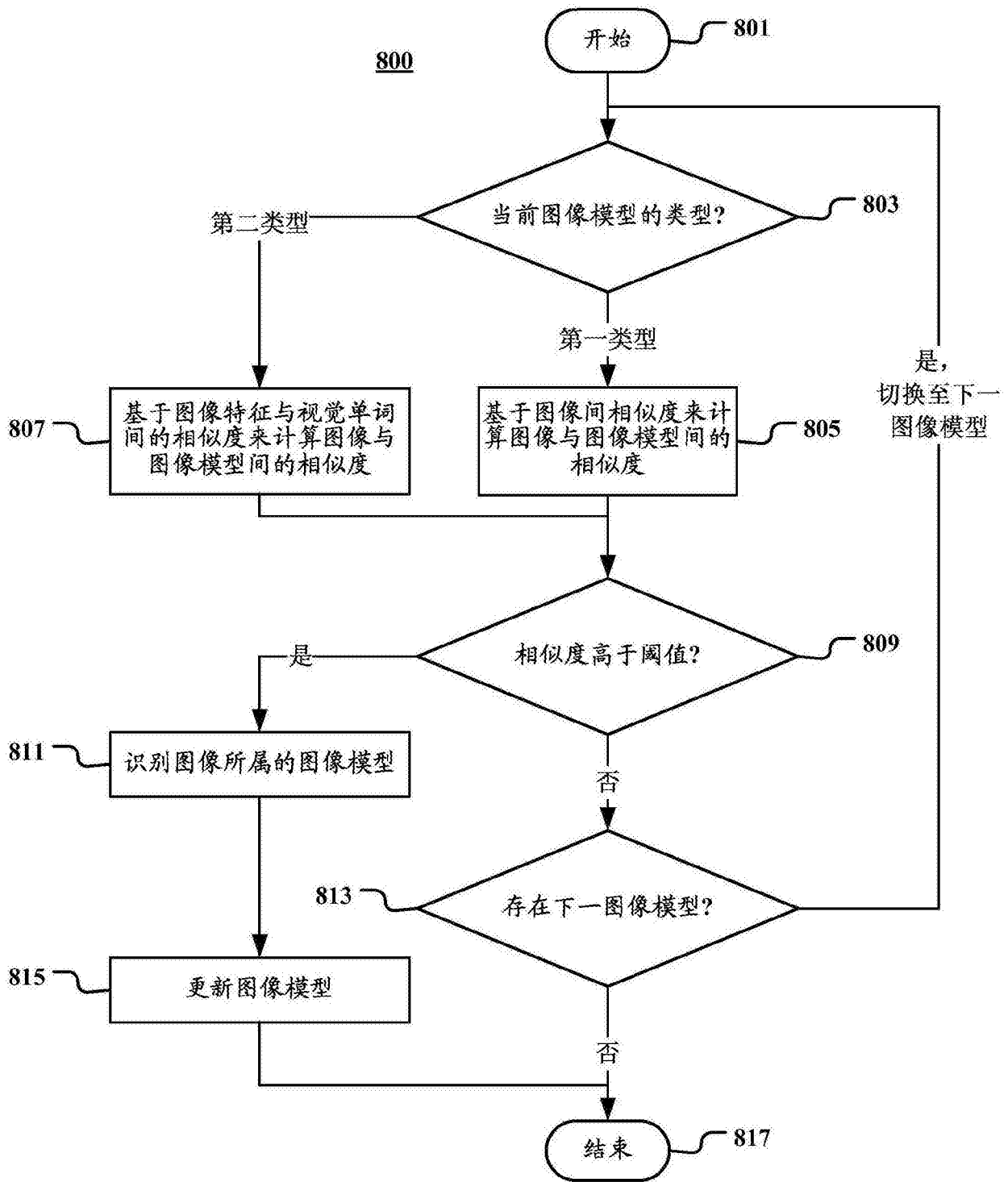


图8

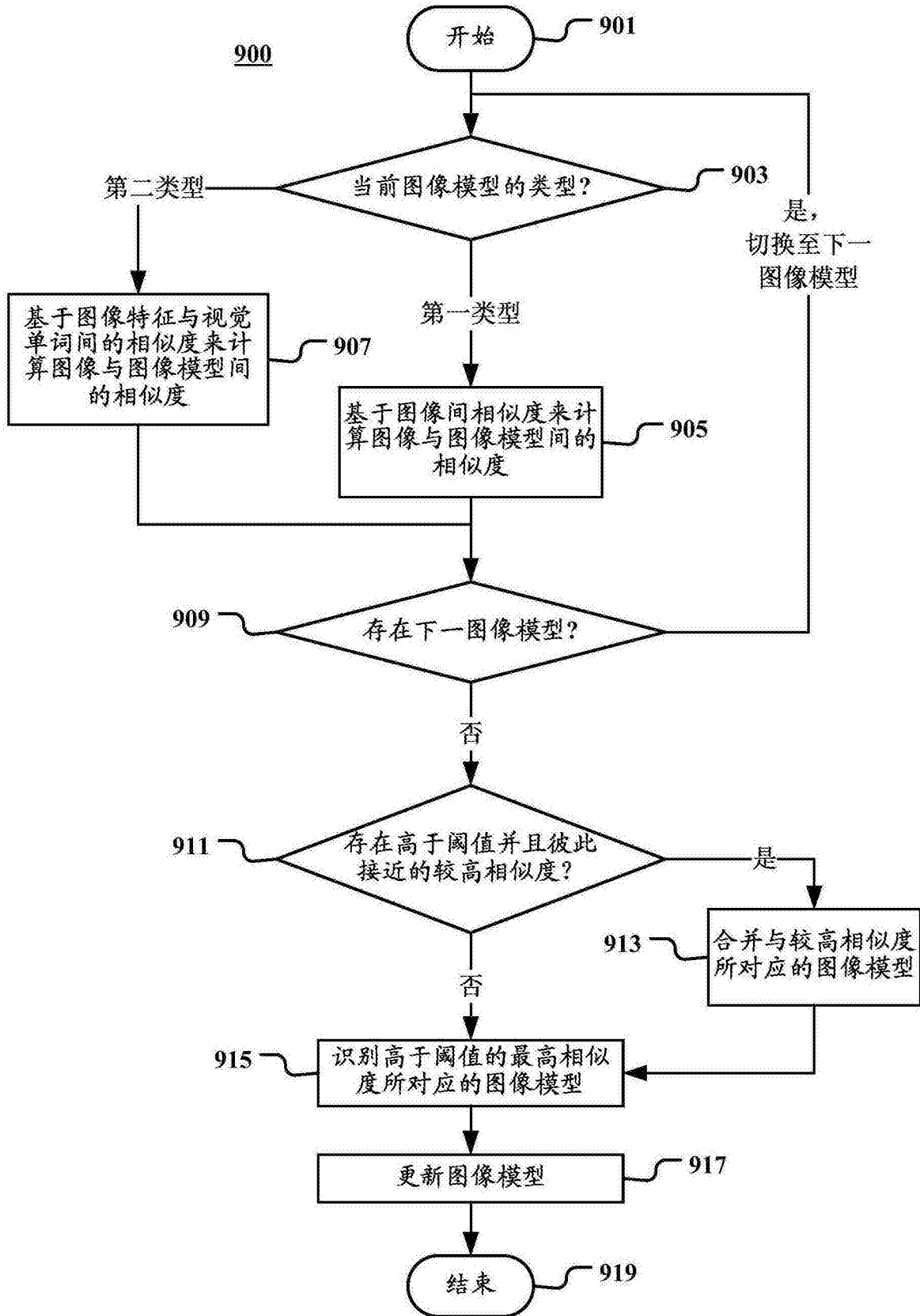


图9



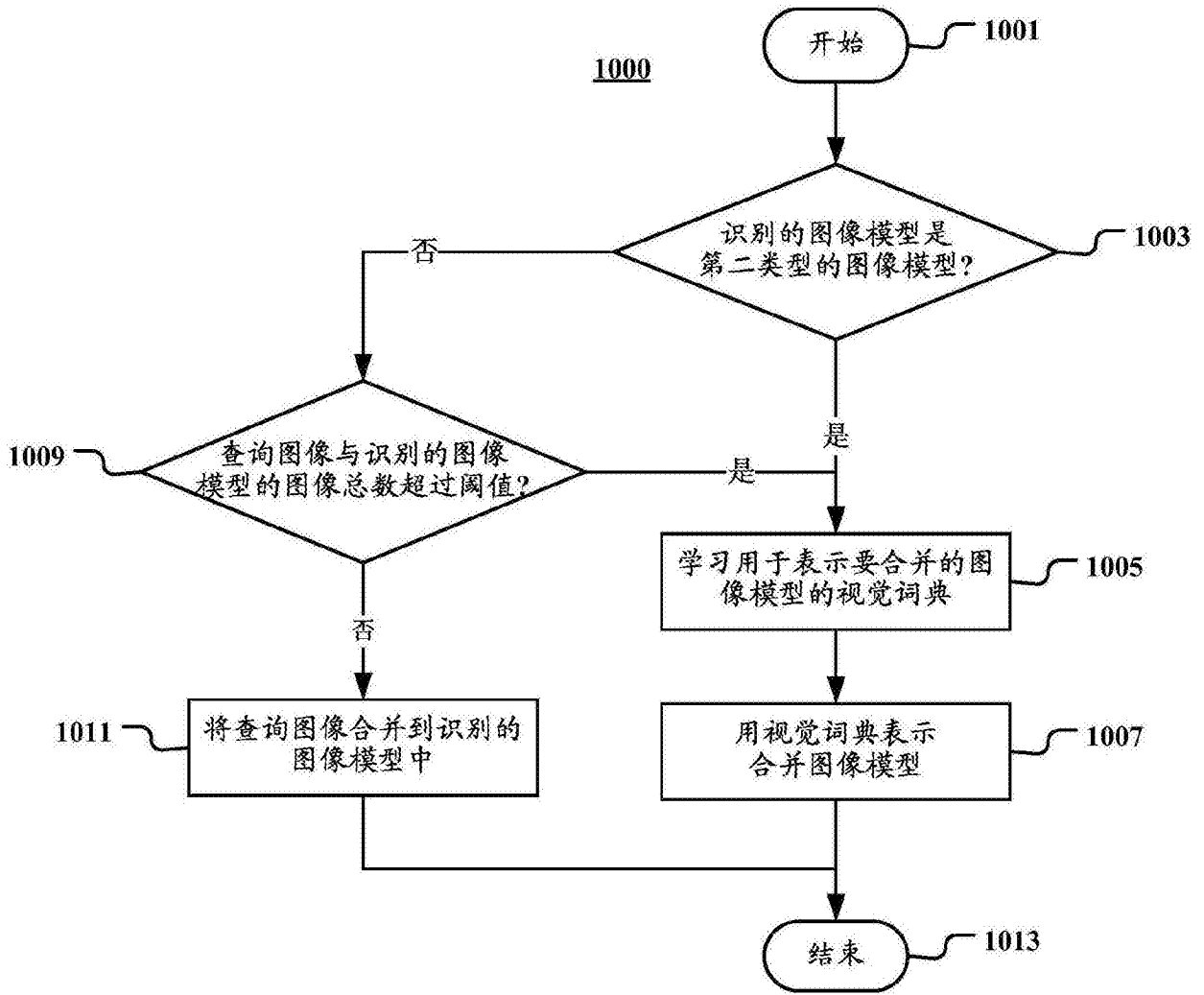


图10

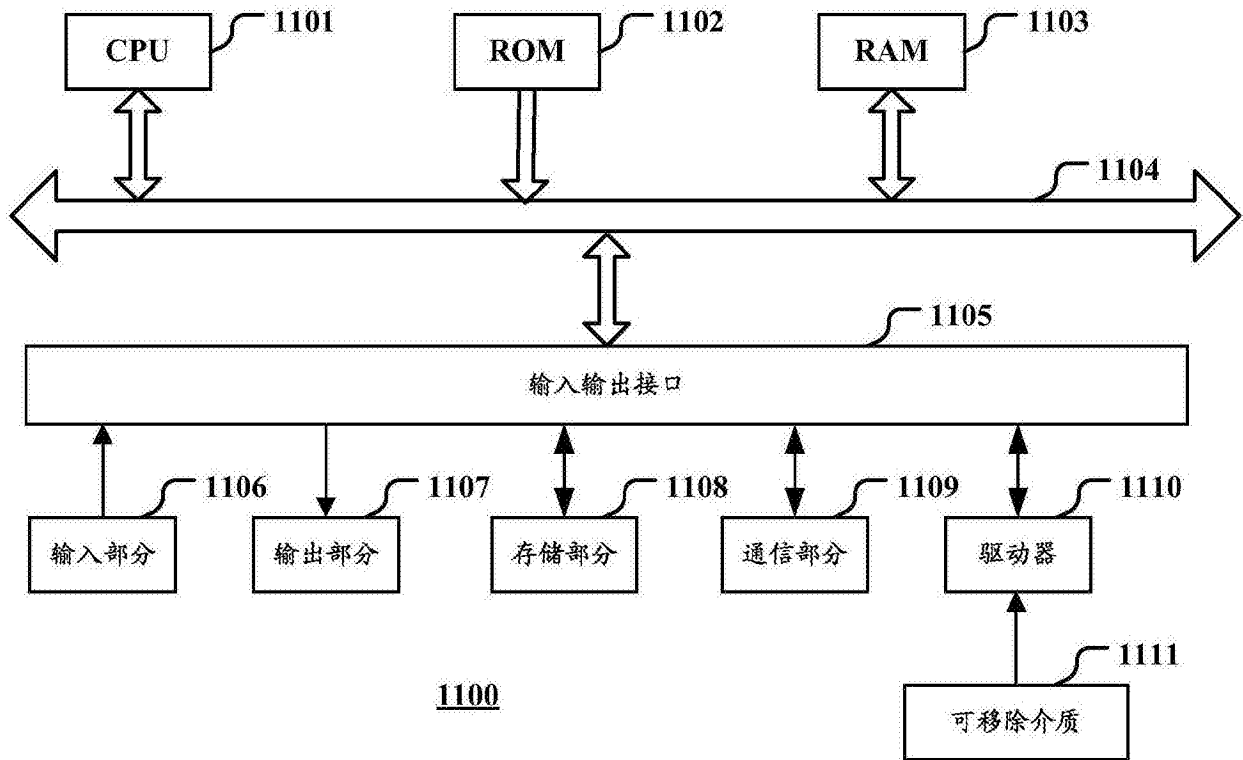


图11