



(12) 发明专利

(10) 授权公告号 CN 115131589 B

(45) 授权公告日 2022. 11. 22

(21) 申请号 202211051928.3

G06V 10/774 (2022.01)

(22) 申请日 2022.08.31

(56) 对比文件

(65) 同一申请的已公布的文献号
申请公布号 CN 115131589 A

CN 103425989 A, 2013.12.04

CN 109558506 A, 2019.04.02

CN 110827193 A, 2020.02.21

(43) 申请公布日 2022.09.30

CN 110991389 A, 2020.04.10

(73) 专利权人 天津艺点意创科技有限公司
地址 300131 天津市红桥区咸阳路19号中
投保大厦三层

CN 114727805 A, 2022.07.08

CN 111368116 A, 2020.07.03

CN 106844785 A, 2017.06.13

US 2018189937 A1, 2018.07.05

(72) 发明人 巩书凯 霍迎迎

李毅暉. 基于视觉机制的图像显著性检测及检索算法研究.《中国优秀硕士学位论文全文数据库信息科技辑》.2017, (第3期),

(74) 专利代理机构 济南凳凳知识产权代理有限公司 37386

专利代理师 陈木发

审查员 杨欢

(51) Int. Cl.

G06V 10/762 (2022.01)

G06V 10/74 (2022.01)

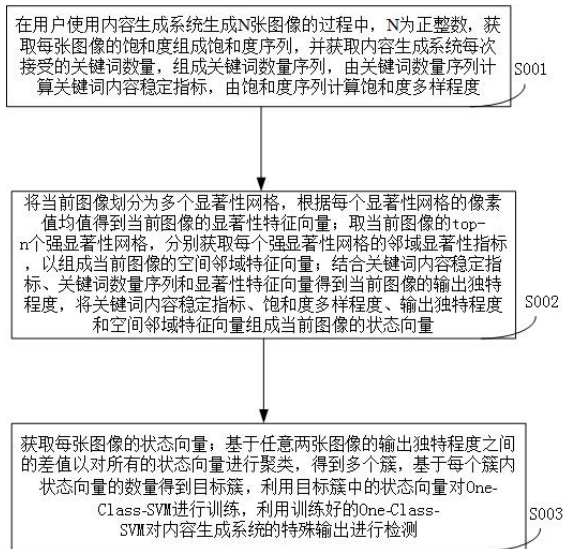
权利要求书1页 说明书6页 附图2页

(54) 发明名称

一种互联网文创作品智能设计的图像生成方法

(57) 摘要

本发明涉及图像处理技术领域,具体涉及一种互联网文创作品智能设计的图像生成方法,该方法在用户使用内容生成系统生成N张图像的过程中,计算关键词数量的关键词内容稳定指标和图像的饱和度多样程度;根据每张图像的多个显著性网格,分别获取每张图像的显著性特征向量和空间邻域特征向量;结合关键词内容稳定指标、饱和度多样程度、每张图像的显著性特征向量和空间邻域特征向量分别获取每张图像的状态向量;利用预处理的状态向量训练One-Class-SVM,以对内容生成系统的特殊输出进行检测。通过常见的状态向量进行训练,使得后期不再依赖大量数据对比,仅由状态向量就能判断特殊输出,提高了判断的准确性和效率。



1. 一种互联网文创作品智能设计的图像生成方法,其特征在于,该方法包括以下步骤:

在用户使用内容生成系统生成N张图像的过程中,N为正整数,获取每张图像的饱和度组成饱和度序列,并获取内容生成系统每次接受的关键词数量,组成关键词数量序列,由所述关键词数量序列计算关键词内容稳定指标,由所述饱和度序列计算饱和度多样程度;

将当前图像划分为多个显著性网格,获取每个显著性网格的像素值均值,基于像素值均值构建显著性网格的直方图,获取直方图中的所有直方图bin以组成显著性特征向量;取当前图像的top-n个强显著性网格,n为正整数,分别获取每个强显著性网格的邻域显著性指标,以组成当前图像的空间邻域特征向量;结合所述关键词内容稳定指标、所述关键词数量序列和所述显著性特征向量得到当前图像的输出独特程度,将所述关键词内容稳定指标、所述饱和度多样程度、所述输出独特程度和所述空间邻域特征向量组成当前图像的状态向量;

获取每张图像的状态向量;基于任意两张图像的所述输出独特程度之间的差值以对所有的状态向量进行聚类,得到多个簇,基于每个簇内状态向量的数量得到目标簇,利用目标簇中的状态向量对One-Class-SVM进行训练,利用训练好的One-Class-SVM对内容生成系统的特殊输出进行检测;

所述基于每个簇内状态向量的数量得到目标簇的方法,包括:

设定占比阈值,分别计算每个簇内状态向量的数量在状态向量的总数的占比,将占比大于或等于占比阈值的簇作为目标簇。

2. 如权利要求1所述的一种互联网文创作品智能设计的图像生成方法,其特征在于,所述关键词内容稳定指标的获取方法,包括:

获取所述关键词数量序列中的最大关键词数量和内容生成系统最新一次接受的关键词数量,以计算两者之间的比值,将该比值输入以e为底的指数函数中得到所述关键词内容稳定指标。

3. 如权利要求1所述的一种互联网文创作品智能设计的图像生成方法,其特征在于,所述饱和度多样程度的获取方法,包括:

分别计算所述饱和度序列中每个饱和度与基准饱和度之间的差值平方,获取差值平方的平均值,将平均值与修正系数之间的乘积作为差异指标;

利用双曲正切函数对所述差异指标进行优化处理,得到所述饱和度多样程度。

4. 如权利要求1所述的一种互联网文创作品智能设计的图像生成方法,其特征在于,所述邻域显著性指标的获取方法,包括:

计算当前强显著性网格的八邻域内的像素值均值,将其八邻域内的像素值均值与当前强显著性网格的像素值均值之间的差值作为当前强显著性网格的邻域显著性指标。

5. 如权利要求1所述的一种互联网文创作品智能设计的图像生成方法,其特征在于,所述结合所述关键词内容稳定指标、所述关键词数量序列和所述显著性特征向量得到当前图像的输出独特程度的方法,包括:

利用余弦相似度函数计算所述显著性特征向量与基准显著性特征向量之间的余弦相似度,由所述余弦相似度得到不相似程度;获取所述关键词数量序列中最大关键词数量与最小关键词数量之间比值;将所述不相似程度、所述比值和所述关键词内容稳定指标之间的乘积作为所述输出独特程度。

一种互联网文创作品智能设计的图像生成方法

技术领域

[0001] 本发明涉及图像处理技术领域,具体涉及一种互联网文创作品智能设计的图像生成方法。

背景技术

[0002] 目前如DALL-E-2、Imagen、Midjourney、discodiffusion等文字到图像合成的深度神经网络技术能够生成较为逼真的图像结果,在文创的衍生和文创的构想测试中起到较强的辅助作用。然而在使用过程中,由于此类的内容生成系统的模块是一连串的子系统,难以从文本层面进行合理控制,而用户始终是需要一些“有新意的”生成内容,因此对内容生成系统进行状态监测,以保证生成内容的正常输出和生成效率是非常必要的。

发明内容

[0003] 为了解决上述技术问题,本发明的目的在于提供一种互联网文创作品智能设计的图像生成方法,所采用的技术方案具体如下:

[0004] 在用户使用内容生成系统生成N张图像的过程中,N为正整数,获取每张图像的饱和度组成饱和度序列,并获取内容生成系统每次接受的关键词数量,组成关键词数量序列,由所述关键词数量序列计算关键词内容稳定指标,由所述饱和度序列计算饱和度多样程度;

[0005] 将当前图像划分为多个显著性网格,获取每个显著性网格的像素值均值,基于像素值均值构建显著性网格的直方图,获取直方图中的所有直方图bin以组成显著性特征向量;取当前图像的top-n个强显著性网格,n为正整数,分别获取每个强显著性网格的邻域显著性指标,以组成当前图像的空间邻域特征向量;结合所述关键词内容稳定指标、所述关键词数量序列和所述显著性特征向量得到当前图像的输出独特程度,将所述关键词内容稳定指标、所述饱和度多样程度、所述输出独特程度和所述空间邻域特征向量组成当前图像的状态向量;

[0006] 获取每张图像的状态向量;基于任意两张图像的所述输出独特程度之间的差值以对所有的状态向量进行聚类,得到多个簇,基于每个簇内状态向量的数量得到目标簇,利用目标簇中的状态向量对One-Class-SVM进行训练,利用训练好的One-Class-SVM对内容生成系统的特殊输出进行检测。

[0007] 进一步的,所述关键词内容稳定指标的获取方法,包括:

[0008] 获取所述关键词数量序列中的最大关键词数量和内容生成系统最新一次接受的关键词数量,以计算两者之间的比值,将该比值输入以e为底的指数函数中得到所述关键词内容稳定指标。

[0009] 进一步的,所述饱和度多样程度的获取方法,包括:

[0010] 分别计算所述饱和度序列中每个饱和度与基准饱和度之间的差值平方,获取差值平方的平均值,将平均值与修正系数之间的乘积作为差异指标;

[0011] 利用双曲正切函数对所述差异指标进行优化处理,得到所述饱和度多样程度。

[0012] 进一步的,所述邻域显著性指标的获取方法,包括:

[0013] 计算当前强显著性网格的八邻域内的像素值均值,将其八邻域内的像素值均值与当前强显著性网格的像素值均值之间的差值作为当前强显著性网格的邻域显著性指标。

[0014] 进一步的,所述结合所述关键词内容稳定指标、所述关键词数量序列和所述显著性特征向量得到当前图像的输出独特程度的方法,包括:

[0015] 利用余弦相似度函数计算所述显著性特征向量与基准显著性特征向量之间的余弦相似度,由所述余弦相似度得到不相似程度;获取所述关键词数量序列中最大关键词数量与最小关键词数量之间比值;将所述不相似程度、所述比值和所述关键词内容稳定指标之间的乘积作为所述输出独特程度。

[0016] 进一步的,所述基于每个簇内状态向量的数量得到目标簇的方法,包括:

[0017] 设定占比阈值,分别计算每个簇内状态向量的数量在状态向量的总数量的占比,将占比大于或等于占比阈值的簇作为目标簇。

[0018] 本发明实施例至少具有如下有益效果:(1)根据内容生成系统接受关键词数量的稳定性和输出图像的饱和度,分析内容生成系统生成每张图像的状态向量,进而通过聚类找到常规输出对应的状态向量,也即是有效状态向量,保证了数据的可靠性。

[0019] (2)利用单类分类器One-Class-SVM,以通过大量常规输出的状态向量的空间分布特征对One-Class-SVM进行训练,使得后期不再依赖大量数据对比,仅由状态向量就能判断特殊输出,提高了判断的准确性和效率。

附图说明

[0020] 为了更清楚地说明本发明实施例或现有技术中的技术方案和优点,下面将对实施例或现有技术描述中所需要使用的附图作简单的介绍,显而易见地,下面描述中的附图仅仅是本发明的一些实施例,对于本领域普通技术人员来讲,在不付出创造性劳动的前提下,还可以根据这些附图获得其它附图。

[0021] 图1为本发明一个实施例提供的一种互联网文创作品智能设计的图像生成方法的步骤流程图;

[0022] 图2为本发明实施例所提供的一种多个泰迪熊生成结果的示意图。

具体实施方式

[0023] 为了更进一步阐述本发明为达成预定发明目的所采取的技术手段及功效,以下结合附图及较佳实施例,对依据本发明提出的一种互联网文创作品智能设计的图像生成方法,其具体实施方式、结构、特征及其功效,详细说明如下。在下述说明中,不同的“一个实施例”或“另一个实施例”指的不一定是同一实施例。此外,一个或多个实施例中的特定特征、结构或特点可由任何合适形式组合。

[0024] 除非另有定义,本文所使用的所有的技术和科学术语与属于本发明的技术领域的技术人员通常理解的含义相同。

[0025] 下面结合附图具体的说明本发明所提供的一种互联网文创作品智能设计的图像生成方法的具体方案。

[0026] 请参阅图1,其示出了本发明一个实施例提供的一种互联网文创作品智能设计的图像生成方法的步骤流程图,该方法包括以下步骤:

[0027] 步骤S001,在用户使用内容生成系统生成N张图像的过程中,N为正整数,获取每张图像的饱和度组成饱和度序列,并获取内容生成系统每次接受的关键词数量,组成关键词数量序列,由关键词数量序列计算关键词内容稳定指标,由饱和度序列计算饱和度多样程度。

[0028] 具体的,在用户使用内容生成系统生成N张图像的过程中,N为正整数,获取每张图像的饱和度 t ,组成饱和度序列 $T = \{t_1, t_2, t_3, \dots, t_N\}$, t_N 为第N张图像的饱和度;同时在生成N张图像的过程中,获取内容生成系统每次接受文本的关键词数量 v ,以前4次使用内容生成系统所获取的关键词数量组成关键词数量序列 $V = \{v_1, v_2, v_3, v_4\}$,其中,关键词数量序列的长度可自行设定。

[0029] 进一步的,内容生成系统在正常情况下生成内容时,其接受的关键词数量会稳定于一个值,如果远离于稳定状态,在一定程度上说明内容生成系统的生成结果可能不同于之前尝试的方式,此时会有新内容出现,因此对关键词数量序列 V 进行预处理,以确定内容生成系统接受关键词的关键词内容稳定指标 Q :

$$[0030] \quad Q = e^{-\frac{\max(V)}{v}}$$

[0031] 其中, e 为常数; $\max(V)$ 为关键词数量序列中的最大关键词数量,且对于内容生成系统,关键词数量越大越容易出现更具体更单一的结果; v' 为内容生成系统最新一次接受的关键词数量,也即是生成N张图像之后的第一次使用时所接受的关键词数量。

[0032] 饱和度的差异能够体现内容生成系统的生成结果中多样性的情况,例如图2中多个泰迪熊生成结果因生成所假设的光照不同,其画面的饱和度情况不同,若一个场景能够出现诸如:多种光照情况、多种背景风格、多种材质等,则会在饱和度差异上存在较大差异,因此分析饱和度序列 T 与预设指标之间的差异指标 p :

$$[0033] \quad p = \left[\frac{1}{N} * \sum_{i=1}^N (t_i - t_{\bar{t}})^2 * \gamma \right]$$

[0034] 其中, $t_{\bar{t}}$ 表示基准饱和度,一般内容生成系统生成的正常饱和度在15%~65%,这里取中值40%作为 $t_{\bar{t}}$; γ 为修正系数,将正常情况下图像的饱和度落于[0,1]内,则令 γ 为 $\frac{1}{25}$; t_i 为第 t 张图像的饱和度。

[0035] 为了使饱和度范围取得有效区间,对差异指标 p 进行优化处理,得到饱和度多样程度 P ,优化公式为:

$$[0036] \quad P = \frac{\tanh(p + 25) + \tanh(-p + 25)}{2}$$

[0037] 其中, $\tanh(*)$ 为双曲正切函数,该函数处理后,在区间[-25,25]内饱和度多样程度 P 约为1,其他情况下,饱和度多样程度 P 接近于0。

[0038] 步骤S002,将当前图像划分为多个显著性网格,根据每个显著性网格的像素值均值得到当前图像的显著性特征向量;取当前图像的top-n个强显著性网格,分别获取每个强显著性网格的邻域显著性指标,以组成当前图像的空间邻域特征向量;结合关键词内容稳定指标、关键词数量序列和显著性特征向量得到当前图像的输出独特程度,将关键词内容稳定指标、饱和度多样程度、输出独特程度和空间邻域特征向量组成当前图像的状态向量。

[0039] 具体的,利用显著性检测网络分别将N张图像都进行显著性网格的划分,以DALLE模型将每张图像划分为多个显著性网格,且显著性网格的画面比例为正方形,本发明实施例优先令显著性网格的数量为25。

[0040] 需要说明的是,本发明实施例在进行显著性网格采集时,默认生成内容不是抽象作品或纯色图像,不考虑其他因素对产生的显著性网格的响应的影响;且显著性网格的划分方式由实施者自行设定,例如每个显著性网格的大小可以是中间大,两边小,甚至对部分区域进行屏蔽等。

[0041] 当内容生成系统的内容发生明显构图变化时,会导致显著性检测结果与其余的候选结果不同,因此可以通过观测显著性网格的响应,确定内容生成系统是否有奇特的输出产生。

[0042] 当内容生成系统出现一定的特别输出时,会使得该输出在使用者眼中更有吸引力或认为更有价值,即例如一种难得的生成结果,从而产生独特的显著性网格的响应,因此,通过图像的显著性网格信息,可以确定内容生成系统的非特定输出特征,故以一张图像为例,基于该图像划分的多个显著性网格,分别计算每个显著性网格的像素值均值,像素值均值能够表示对应显著性网格的响应,对计算的像素值均值进行统计以构建显著性网格的直方图,并对直方图进行极差标准化,即该直方图中的最高bin为1,将直方图中的所有直方图bin组成该图像的显著性特征向量 $F = \{f_1, f_2, f_3, \dots, f_x\}$,其中, f_x 为第 x 个直方图bin,且显著性特征向量 F 中的各个元素是以直方图bin的大小进行排序后得到的。

[0043] 由于采集的显著性网格空间邻域分布所包含的数值分布信息比较抽象,更能够体现一种输出结果是否典型的特征,但在空间风格上难以约束,因此,对该图像的显著性网格的强显著性网格的空间特征进行处理,且使用更少的向量表示,以减轻后续的计算量,具体方法为:选出该图像中top-n个强显著性网格,本发明实施例中n为10,即前十个强显著性网格;当一个显著性网格相对于邻域的显著性都大时,则认为该显著性网格代表了一种小目标,且小目标在空域上很突出,而当一个显著性网格相对邻域的显著性都相似时,则认为该显著性网格代表了一种大目标,因此基于显著性网格的八邻域,计算当前强显著性网格的八邻域内的像素值均值,将其八邻域内的像素值均值与当前强显著性网格的像素值均值之间的差值作为当前强显著性网格的邻域显著性指标,获取每个强显著性网格的邻域显著性指标以组成该图像的空间邻域特征向量 $S = \{s_1 \dots s_n\}$, s_n 为第 n 个强显著性网格的邻域显著性指标,进而利用较少个数的向量表示了该图像中显著性网格的空间邻域特征。

[0044] 内容生成系统在用户的连续操作记录中连续生成多张图像,且每张图像在每次生成的候选结果中存在统一的上下文指标,也即是关键词内容稳定指标 Q 和饱和度多样程度 P ,因此将饱和度序列 T 、关键词数量序列 V 、关键词内容稳定指标 Q 、饱和度多样程度 P 和该图像的显著性特征向量 F 以及空间邻域特征向量 S 构成该图像的第一特征向量,根据第一

特征向量计算该图像在内容生成系统的输出独特程度,则输出独特程度的计算公式为:

$$[0045] \quad U = \left[1 - \text{Similarity}(F, F_{\text{基}}) \right] * Q * \frac{\text{mean}(V)}{\text{Max}(V)}$$

[0046] 其中, U 为输出独特程度; $\text{Similarity}(\ast)$ 为余弦相似度函数,用于比较两个向量之间的相似程度; $\text{mean}(\ast)$ 为均值函数; $\text{Max}(\ast)$ 为求取最大值函数; $F_{\text{基}}$ 为实施者经过调试得到的基准直方图所对应的基准显著性特征向量。

[0047] 图像的显著性特征向量与基准显著性特征向量相近时,余弦相似度越大,相对应输出独特程度越小;关键词数量的稳定性越大,也即关键词内容稳定指标 Q ,输出独特程度越小;关键词数量的均值与最大值有明显的差异,则比值会更小,相对应输出独特程度越大。

[0048] 需要说明的是,基准直方图是需要结合内容生成系统的输出特性和调试实践来确定的,本发明实施例只是给出一种用于实践中参考的基准。

[0049] 将关键词内容稳定指标 Q 、饱和度多样程度 P 、该图像的输出独特程度 U 和空间邻域特征向量 S 构成该图像的状态向量 $x = \{Q, P, U, S\}$ 。

[0050] 步骤S003,获取每张图像的状态向量;基于任意两张图像的输出独特程度之间的差值以对所有的状态向量进行聚类,得到多个簇,基于每个簇内状态向量的数量得到目标簇,利用目标簇中的状态向量对One-Class-SVM进行训练,利用训练好的One-Class-SVM对内容生成系统的特殊输出进行检测。

[0051] 具体的,结合步骤S001和步骤S002能够得到 N 张图像中每张图像的状态向量。

[0052] 计算任意两张图像的输出独特程度之间的差值,将差值作为样本距离,利用DBSCAN聚类算法对 N 张图像的所有状态向量进行聚类,得到多个簇,其中DBSCAN聚类算法中的搜索半径 eps 默认为0.5,簇内最小值 minpts 设定为4。

[0053] 需要说明的是,聚类过程中会出现孤立点,而孤立点即为少见的特殊输出,因此将孤立点作为单独的一个簇。

[0054] 基于聚类得到的各个簇,由簇内状态向量的数量获取目标簇,其获取方法为:若各簇内的数据较少,依然可以视为和孤立点一样的独特输出,因此实施者基于设定的阈值,例如百分比1%,分别计算每个簇内状态向量的数量在状态向量的总数的占比,当簇的占比小于阈值时,确定对应簇为特殊输出,反之,将占比大于或等于阈值的簇作为目标簇进行保留。

[0055] 进一步的,将目标簇中的状态向量构成One-Class-SVM的数据集,利用数据集对One-Class-SVM进行训练,其中利用One-Class-SVM的目的是为了学习大多数常见输出的状态向量分布所代表的超平面,以此粗略拟合出能够区分特殊状态的平面外区域的状态向量,从而在后期仅依赖状态向量即可判别是否特殊,而不再依赖大量的数据比对。

[0056] 需要说明的是,One-Class-SVM是公知技术,本发明实施例中不再赘述。

[0057] 基于训练好的One-Class-SVM,将内容生成系统所生成的每张图像的状态向量输入训练好的One-Class-SVM能够确认内容生成系统的特殊输出,即当任何图像属于单类分类器的特殊输出类型时,停止重新输出,反之自动重新输出,直到有特殊输出的图像产生。

[0058] 利用One-Class-SVM对内容生成系统进行输出状态监测的目的是:简单的指标评

价只是一个大致状态,如果特殊输出状态的表现不明显,会发生在多个特殊输出状态间进行波动,导致无法确定特殊输出,而One-Class-SVM为一个单类分类器,通过常见的状态向量训练后能够通过当前状态向量确定特殊输出状态,以提高判断的准确性。

[0059] 综上所述,本发明实施例提供了一种互联网文创作品智能设计的图像生成方法,该方法在用户使用内容生成系统生成N张图像的过程中,根据内容生成系统每次接受的关键词数量计算关键词内容稳定指标,由每张图像的饱和度计算饱和度多样程度;将每张图像划分为多个显著性网格以获取每张图像的显著性特征向量和空间邻域特征向量;结合关键词内容稳定指标、饱和度多样程度、每张图像的显著性特征向量和空间邻域特征向量分别获取每张图像的状态向量;对所有的状态向量进行聚类,得到多个簇,基于每个簇内状态向量的数量得到目标簇,利用目标簇中的状态向量对One-Class-SVM进行训练,利用训练好的One-Class-SVM对内容生成系统的特殊输出进行检测。通过常见的状态向量进行训练,使得后期不再依赖大量数据对比,仅由状态向量就能判断特殊输出,提高了判断的准确性和效率。

[0060] 需要说明的是:上述本发明实施例先后顺序仅仅为了描述,不代表实施例的优劣。且上述对本说明书特定实施例进行了描述。另外,在附图中描绘的过程不一定要求示出的特定顺序或者连续顺序才能实现期望的结果。在某些实施方式中,多任务处理和并行处理也是可以的或者可能是有利的。

[0061] 本说明书中的各个实施例均采用递进的方式描述,各个实施例之间相同相似的部分互相参见即可,每个实施例重点说明的都是与其他实施例的不同之处。

[0062] 以上所述仅为本发明的较佳实施例,并不用以限制本发明,凡在本发明的原则之内,所作的任何修改、等同替换、改进等,均应包含在本发明的保护范围之内。

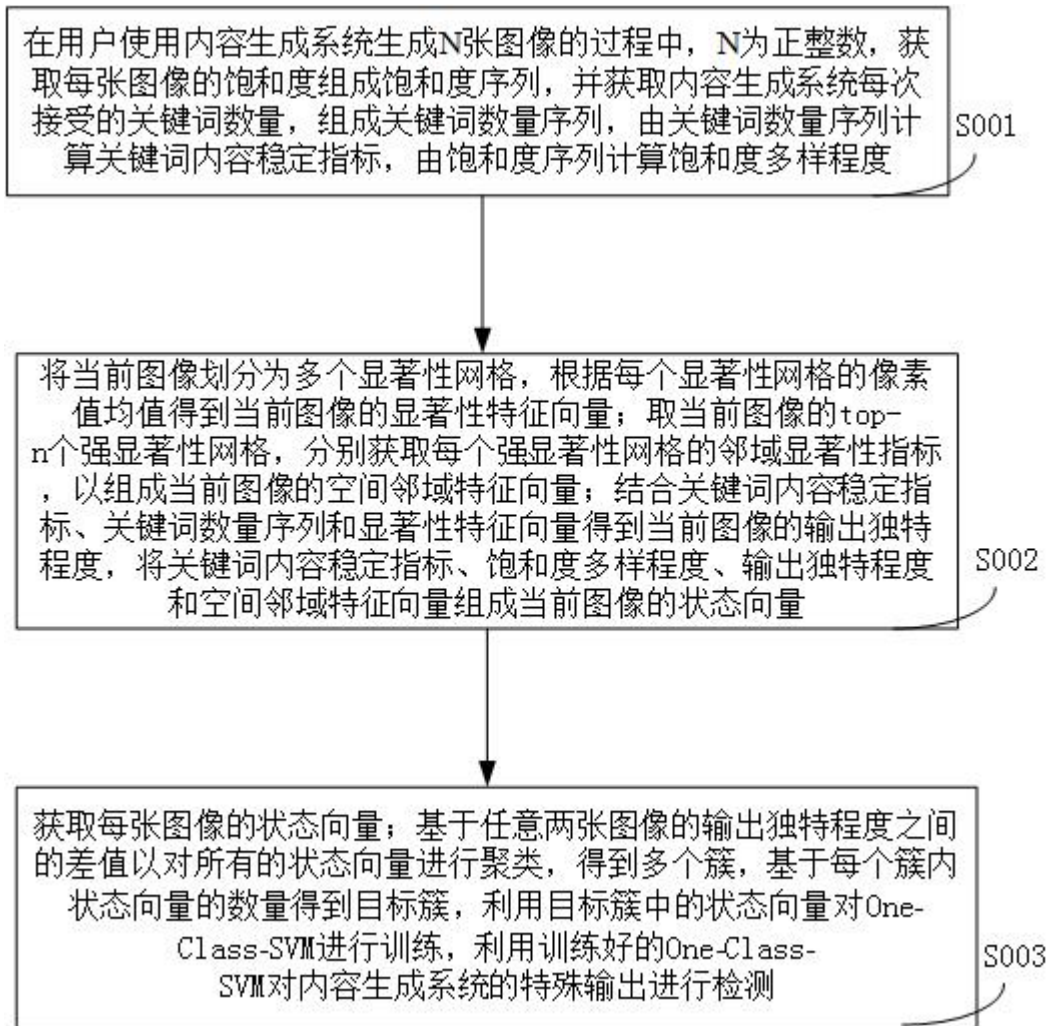


图1



图2