



(12) 发明专利申请

(10) 申请公布号 CN 103744903 A

(43) 申请公布日 2014. 04. 23

(21) 申请号 201310726931. 5

(22) 申请日 2013. 12. 25

(71) 申请人 中国科学技术大学

地址 230026 安徽省合肥市包河区金寨路
96 号

(72) 发明人 陈雪锦 张孝 谈建超 侯丹彤

(74) 专利代理机构 北京凯特来知识产权代理有
限公司 11260

代理人 郑立明 赵镇勇

(51) Int. Cl.

G06F 17/30 (2006. 01)

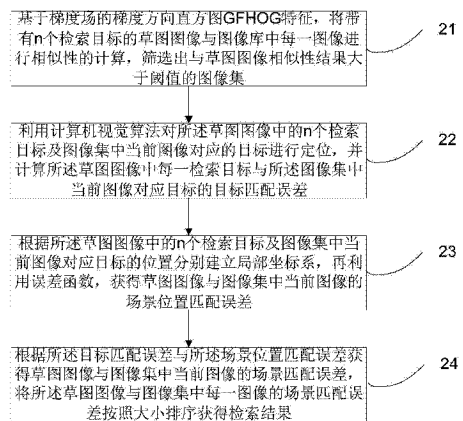
权利要求书2页 说明书5页 附图2页

(54) 发明名称

一种基于草图的场景图像检索方法

(57) 摘要

本发明公开了一种基于草图的场景图像检索方法,该方法包括:基于 GFHOG 特征,将带有 n 个检索目标的草图图像与图像库中每一图像进行相似性的计算,筛选出与草图图像相似性结果大于阈值的图像集;对所述草图图像中的 n 个检索目标及图像集中当前图像对应的目标进行定位,并计算所述两个图像中对应目标的目标匹配误差;根据两个图像中目标的位置分别建立局部坐标系,从而基于误差函数,获得草图图像与图像集中当前图像的场景位置匹配误差;根据所述目标匹配误差与所述场景位置匹配误差获得草图图像与图像集中当前图像的场景匹配误差,将所述草图图像与图像集中每一图像的场景匹配误差按照大小排序获得检索结果。通过采用本发明公开的方法实现了多目标的快速检索。



1. 一种基于草图的场景图像检索方法,其特征在于,该方法包括:

基于梯度场的梯度方向直方图 GFHOG 特征,将带有 n 个检索目标的草图图像与图像库中每一图像进行相似性的计算,筛选出与草图图像相似性结果大于阈值的图像集;

利用计算机视觉算法对所述草图图像中的 n 个检索目标及图像集中当前图像对应的目标进行定位,并计算所述草图图像中每一检索目标与所述图像集中当前图像对应目标的目标匹配误差;

根据所述草图图像中的 n 个检索目标及图像集中当前图像对应目标的位置分别建立局部坐标系,再利用误差函数,获得草图图像与图像集中当前图像的场景位置匹配误差;

根据所述目标匹配误差与所述场景位置匹配误差获得草图图像与图像集中当前图像的场景匹配误差,将所述草图图像与图像集中每一图像的场景匹配误差按照大小排序获得检索结果。

2. 根据权利要求 1 所述的方法,其特征在于,提取草图图像与图像库中图像 GFHOG 特征的步骤包括:梯度场 GF 的计算及梯度方向直方图 HOG 特征的提取;

其中,梯度场的计算包括:利用边缘检测算法提取图像的边缘,并计算边缘各点在梯度方向场中的梯度方向;将梯度场的指导向量场设为零,建立泊松方程;将所述泊松方程转化为线性方程组,并求解出各个非边缘点在梯度方向场中的梯度方向;

HOG 特征的提取包括:在预设的不同窗口尺度下提取梯度场计算后的图像边缘点的 HOG 特征。

3. 根据权利要求 1 所述的方法,其特征在于,所述将带有 n 个检索目标的草图图像与图像库中每一图像进行相似性的计算包括:

利用聚类算法对提取 GFHOG 特征后的图像聚类,获得 GFHOG 特征的聚类中心;

根据所述 GFHOG 特征的聚类中心获得对应的词频直方图;其中,草图图像的词频直方图表示为 H^S ,图像库中图像的词频直方图表示为 H^I ;

根据所述草图图像的词频直方图 H^S 与图像库中图像的词频直方图 H^I 进行相似性的计算。

4. 根据权利要求 3 所述的方法,其特征在于,

利用直方图较差距离进行相似性的计算,其公式为:

$$d(H^I, H^S) = \sum_{i=1}^k \sum_{j=1}^k \omega_{ij} \min(H^S(i) - H^I(j));$$

其中: $\omega_{ij} = 1 - |H^S(i) - H^I(j)|$, $H^S(i)$ 表示草图图像的词频直方图中,视觉单词 i 的频率; $H^I(j)$ 表示图像库中图像的词频直方图中,视觉单词 j 的频率。

5. 根据权利要求 1 所述的方法,其特征在于,所述利用计算机视觉算法对所述草图图像中的 n 个检索目标及图像集中当前图像对应的目标进行定位,并计算所述草图图像中每一检索目标与所述图像集中当前图像对应目标的目标匹配误差包括:

基于草图图像与图像集中当前图像的 GFHOG 特征并利用计算机视觉算法进行目标的定位;具体的:利用最近邻计算草图图像的边缘点在图像集中当前图像的对应点,其公式为:

$$P_m^S \rightarrow P_n^I = \{p_{i=1, \dots, m}, p_{s=1, \dots, n}\};$$

其中, P_m^s 表示草图图像边缘点的坐标; P_n^t 表示图像集中当前图像与草图图像每一边缘点 GFHOG 特征欧式距离最近的点坐标;

提取任意两组对应点, 并通过求解线性方程组计算出表示点与点之间对应关系的仿射变换矩阵 T;

利用仿射变换矩阵 T 计算误差能量函数, 其公式为:

$$E(T) = \sum_{\{p_s, p_t\}} |p_s - Tp_t|^2 + |p_t - T^{-1}p_s|^2;$$

利用使得所述误差能量函数 E(T) 最小的仿射变换矩阵 T 进行草图图像中检索目标及图像集中当前图像对应目标的定位; 并将该最小的误差能量函数 E(T) 作为定位该仿射变换矩阵 T 对应目标的目标匹配误差。

6. 根据权利要求 1 所述的方法, 其特征在于, 获得草图图像与图像集中当前图像的场景位置匹配误差的步骤包括:

利用边界盒限定草图图像中每一检索目标及图像集中当前图像对应目标的范围;

以草图图像某一检索目标对应边界盒的中心点作为基准点, 与其余 n-1 个检索目标对应的边界盒的中心点相连, 完成草图图像局部坐标系的建立; 获得 n-1 个连接线对应的向量, 记为 v_1, v_2, \dots, v_{n-1} ;

以图像集中当前图像与所述草图图像某一检索目标对应的目标所处的边界盒的中心点作为基准点, 与其余 n-1 个目标对应的边界盒的中心点相连, 完成图像集中当前图像局部坐标系的建立; 该图像 n-1 个连接线对应的向量, 记为 $v'_1, v'_2, \dots, v'_{n-1}$;

利用所述局部坐标系中的向量 v_1, v_2, \dots, v_{n-1} 与向量 $v'_1, v'_2, \dots, v'_{n-1}$ 建立误差函数, 从而获得场景位置匹配误差, 其公式为:

$$E_{position} = (1 + |v_2 - \frac{|v_1|}{|v'_1|} v'_2|^2 + \dots + |v_{n-1} - \frac{|v_1|}{|v'_{n-1}|} v'_{n-1}|^2)。$$

7. 根据权利要求 1-6 任一项所述的方法, 其特征在于, 根据所述目标匹配误差与所述场景位置匹配误差获得草图图像与图像集中当前图像的场景匹配误差 E_{error} 包括:

$$E_{error} = E_{object1} + \dots + E_{objectn} + E_{position};$$

其中, $E_{position}$ 表示草图图像与图像集中当前图像的场景位置匹配误差, $E_{object1} - E_{objectn}$ 表示草图图像的 1 至 n 个检索目标与图像集中当前图像对应的 1 至 n 个目标的目标匹配误差。

一种基于草图的场景图像检索方法

技术领域

[0001] 本发明涉及图像检索技术领域,尤其涉及一种基于草图的场景图像检索方法。

背景技术

[0002] 近些年,随着英特网(Internet)以及图像采集设备(数码相机、智能手机)等技术的飞速发展,数字图像已深深融入人们的生活当中,用户可以通过图像采集设备或网络获取大量的数字图像。在这么庞大的数据量面前,一套有效的图像搜索机制显得尤为重要。图像数据描述的复杂性也为图像检索造成了巨大困难。

[0003] 基于内容的图像检索为从大规模数字图像数据库中搜索出特定内容的图像提供了有效的方法。大多传统和一般图像检索的方式是利用一些增加元数据(metadata)的方法,例如:字幕、关键词或是图像的说明,如此一来就可以通过注解词完成检索。人工的图像注解是费时、费力并且昂贵;为了解决这个问题,已经有大量的研究在做自动图像注解方面上。此外,越来越多的社会网络应用和语义网已经产生了数个以网络为基底发展的图像注解工具。

[0004] 互联网上传统的搜索引擎,包括Google、Yahoo以及MSN都推出相应的图片搜索功能,但是这种搜索主要是基于图片的文件名建立索引来实现查询功能(也许利用了网页上的文字信息)。这种从查询文字、文件名,最终到图片查询的机制并不是基于内容的图像检索。基于内容的图像检索指的是查询条件本身就是一个图像,或者是对于图像内容的描述,它建立索引的方式是通过提取底层特征,然后通过计算比较这些特征和查询条件之间的距离,来决定两个图片的相似程度。

[0005] 基于草图的图像检索是基于内容的图像检索的一种查询模式(Query by sketch)。如图1所示,用户在类似笔画的接口上进行简单的绘制作为标准进行查询。由计算机利用特征描述子对输入草图的特征进行描述,常用的包括:质心距离描述子、投影长度描述子、区域统计描述子和球谐和函数描述子。但是,上述特征描述子只能用于检索简单图像,而无法用于草图中包括多个检索目标的图像检索。

发明内容

[0006] 本发明的目的是提供一种基于草图的场景图像检索方法,实现了多目标的快速检索。

[0007] 本发明的目的是通过以下技术方案实现的:

[0008] 一种基于草图的场景图像检索方法,该方法包括:

[0009] 基于梯度场的梯度方向直方图GFHOG特征,将带有n个检索目标的草图图像与图像库中每一图像进行相似性的计算,筛选出与草图图像相似性结果大于阈值的图像集;

[0010] 利用计算机视觉算法对所述草图图像中的n个检索目标及图像集中当前图像对应的目标进行定位,并计算所述草图图像中每一检索目标与所述图像集中当前图像对应目标的目标匹配误差;

[0011] 根据所述草图图像中的 n 个检索目标及图像集中当前图像对应目标的位置分别建立局部坐标系,再利用误差函数,获得草图图像与图像集中当前图像的场景位置匹配误差;

[0012] 根据所述目标匹配误差与所述场景位置匹配误差获得草图图像与图像集中当前图像的场景匹配误差,将所述草图图像与图像集中每一图像的场景匹配误差按照大小排序获得检索结果。

[0013] 由上述本发明提供的技术方案可以看出,根据草图中各个检索目标的特征从图像库中来筛选出包含这些特征的图像集,再利用各个检索目标的位置关系及其相似度来实现多目标的快速检索。

附图说明

[0014] 为了更清楚地说明本发明实施例的技术方案,下面将对实施例描述中所需要使用的附图作简单地介绍,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,对于本领域的普通技术人员来讲,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他附图。

[0015] 图 1 为本发明背景技术提供的草图图像的示意图;

[0016] 图 2 为本发明实施例一提供的一种基于草图的场景图像检索方法的流程图;

[0017] 图 3 为本发明实施例一提供的一种以一边界盒中心为基准建立局部坐标系的示意图。

具体实施方式

[0018] 下面结合本发明实施例中的附图,对本发明实施例中的技术方案进行清楚、完整地描述,显然,所描述的实施例仅仅是本发明一部分实施例,而不是全部的实施例。基于本发明的实施例,本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例,都属于本发明的保护范围。

[0019] 本发明实施例所述的场景图像是指图像中包含多个前景目标,各个目标有特定的空间位置关系;同时,在利用草图图像进行场景图像检索时,草图图像中也包含多个检索目标。此时,可以各个检索目标与图像前景目标的相似度以及位置关系的相似度来检索场景图像。

[0020] 在基于场景图像的检索中由于要对场景中每个目标进行定位,采用上述的组合描述子只能表示图像的全局信息,而没有表述图像局部特征的能力;因此,本发明实施例在基于场景的图像检索中采用了 GFHOG (梯度场的梯度方向直方图 Gradient Field Histogram of Oriented Gradient) 特征描述子。GFHOG 特征描述子有较好的表示局部特征的能力,同时也能兼顾的表示相邻点的描述子之间的影响。

[0021] 实施例一

[0022] 图 2 为本发明实施例一提供的一种基于草图的场景图像检索方法的流程图。如图 2 所示,该方法主要包括:

[0023] 步骤 21、基于梯度场的梯度方向直方图 GFHOG 特征,将带有 n 个检索目标的草图图像与图像库中每一图像进行相似性的计算,筛选出与草图图像相似性结果大于阈值的图像

集。

[0024] 本发明实施例中需要预先提取图像库中每一图像的 GFHOG 特征,主要包括:梯度场 GF 的计算及梯度方向直方图 HOG 特征的提取。

[0025] 其中,梯度场的计算包括:利用边缘检测算法(例如,canny 边缘检测算法)提取图像的边缘,并计算边缘各点在梯度方向场中的梯度方向;将梯度场的指导向量场设为零,建立泊松方程;将所述泊松方程转化为线性方程组,并求解出各个非边缘点在梯度方向场中的梯度方向。

[0026] HOG 特征的提取包括:在预设的不同窗口尺度下提取梯度场计算后的图像边缘点的 HOG 特征。示例性的,对边缘像素点为中心以及水平或竖直相邻的 w 个像素点的像素的 $3*3$ 邻域统计梯度方向直方图($w=5, 10, 15$),梯度方向被平均分为 9 个 bin(区域),从而每个梯度点得到一个 $9*9*3=243$ 维的特征向量。这样每个点的特征不仅统计了这一个点的梯度方向特征,也统计了其周围相邻像素的梯度方向特征。

[0027] 当获取到用户输入的带有 n ($n \geq 1$) 个检索目标的草图图像后,也采用上述方法提取 GFHOG 特征。

[0028] 当完成图像 GFHOG 特征的提取之后,需要建立对应的词频直方图来进行相似度的计算。具体的步骤如下:利用聚类算法(例如,K-means (K-均值)聚类)对提取 GFHOG 特征后的图像聚类,获得 GFHOG 特征的聚类中心;根据所述 GFHOG 特征的聚类中心获得对应的词频直方图;其中,草图图像的词频直方图表示为 H^S ,图像库中图像的词频直方图表示为 H^I 。

[0029] 再利用相似度算法对草图图像与图像库中图像的相似度进行计算。示例性的,本发明实施例采用的相似度度量为直方图交叉距离,其计算公式为:

$$[0030] \quad d(H^I, H^S) = \sum_{i=1}^k \sum_{j=1}^k \omega_{ij} \min(H^S(i) - H^I(j));$$

[0031] 其中: $\omega_{ij}=1-|H^S(i)-H^I(j)|$, $H^S(i)$ 表示草图图像的词频直方图中,视觉单词 i 的频率; $H^I(j)$ 表示图像库中图像的词频直方图中,视觉单词 j 的频率。

[0032] 逐一计算草图图像与图像库中每一图像的相似度后,筛选出与草图图像相似性结果大于阈值的图像集。本发明实施例不对该阈值的大小做出限定,用户可以根据实际的需求或经验来进行相应的设定。

[0033] 步骤 22、利用计算机视觉算法对所述草图图像中的 n 个检索目标及图像集中当前图像对应的目标进行定位,并计算所述草图图像中每一检索目标与所述图像集中当前图像对应目标的目标匹配误差。

[0034] 本发明实施例对草图图像和图像集的 GFHOG 特征做计算机视觉识别,例如,RANSAC(随机抽样一致)来获取草图目标的定位。假定草图在目标图像中的对应关系满足刚性变换(尺度、旋转以及平移变换),则可以用一个仿射变换矩阵 T 来表示特征点对应关系。

[0035] 具体的:首先,利用最近邻计算草图图像的边缘点在图像集中当前图像的对应点,其公式为:

$$[0036] \quad P_m^S \rightarrow P_n^I = \{P_{i=1, \dots, m}, P_{s=1, \dots, n}\};$$

[0037] 在步骤 21 中对每一个图像的边缘点都计算了其 GFHOG 特征。此时,可以利用欧氏距离,对草图的每一个边缘点的 GFHOG 特征在图像集中每一图像的 GFHOG 特征中最近的点

坐标进行计算。 P_m^S 是草图中所有的边缘点坐标, P_n^I 表示图像集中当前图像与草图图像每一边缘点 GFHOG 特征欧式距离最近的点坐标;

[0038] 其次,提取上述的任意两组对应点,则可以通过解线性方程组计算出表示点与点之间对应关系的仿射变换矩阵 T;再利用仿射变换矩阵 T 计算误差能量函数 E(T),具体的:每一仿射变换矩阵 T 均可计算出一个误差能量函数 E(T),通过多次取点后将使得误差能量函数 E(T) 最小的仿射变换矩阵 T 作为两者目标的对应变换关系;同时,对草图做 T 变换就可以定位草图图像在图像集中当前图像的位置。

[0039] 其中,误差能量函数 E(T) 的计算公式为:

$$[0040] \quad E(T) = \sum_{\{p_s, p_i\}} |p_s - T p_i|^2 + |p_i - T^{-1} p_s|^2;$$

[0041] 当利用使得所述误差能量函数 E(T) 最小的仿射变换矩阵 T 进行草图图像中检索目标及图像集中当前图像对应目标的定位后,则将该最小的误差能量函数 E(T) 作为定位该仿射变换矩阵 T 对应目标的目标匹配误差。

[0042] 步骤 23、根据所述草图图像中的 n 个检索目标及图像集中当前图像对应目标的位置分别建立局部坐标系,再利用误差函数,获得草图图像与图像集中当前图像的场景位置匹配误差。

[0043] 本发明实施例中的场景位置匹配误差是基于局部坐标系与误差函数计算而来。如图 3 所示,包括如下步骤:

[0044] 首先,利用边界盒(bounding box)限定草图图像中每一检索目标及图像集中当前图像对应目标的范围。为了便于在图中表示,本实施例中的 n 设为 3。

[0045] 然后,以草图图像某一检索目标对应边界盒(边界盒编号为 object1-n)的中心点作为基准点,与其余 n-1 个检索目标对应的边界盒的中心点相连,完成草图图像局部坐标系的建立;获得 n-1 个连接线对应的向量,记为 v_1, v_2, \dots, v_{n-1} 。

[0046] 再以图像集中当前图像与所述草图图像某一检索目标对应的目标所处的边界盒(边界盒编号为 object1'-n')的中心点作为基准点,与其余 n-1 个目标对应的边界盒的中心点相连,完成图像集中当前图像局部坐标系的建立;该图像集中当前图像 n-1 个连接线对应的向量,记为 $v'_1, v'_2, \dots, v'_{n-1}$ 。其中,向量 $v'_1, v'_2, \dots, v'_{n-1}$ 所表示的连接线与向量 v_1, v_2, \dots, v_{n-1} 所表示的连接线一一对应。

[0047] 最后,利用所述局部坐标系中的向量 v_1, v_2, \dots, v_{n-1} 与向量 $v'_1, v'_2, \dots, v'_{n-1}$ 建立误差函数,从而获得场景位置匹配误差,其公式为:

$$[0048] \quad E_{position} = (1 + |v_2 - \frac{|v_1|}{|v'_1|} v'_2|^2 + \dots + |v_{n-1} - \frac{|v_1|}{|v'_{n-1}|} v'_{n-1}|^2)。$$

[0049] 步骤 24、根据所述目标匹配误差与所述场景位置匹配误差获得草图图像与图像集中当前图像的场景匹配误差,将所述草图图像与图像集中每一图像的场景匹配误差按照大小排序获得检索结果。

[0050] 可以采用如下公式进行计算:

$$[0051] \quad E_{error} = E_{object1} + \dots + E_{objectn} + E_{position};$$

[0052] 其中, E_{error} 表示草图图像与图像集中当前图像的场景匹配误差, $E_{position}$ 表示草图

图像与图像集中当前图像的场景位置匹配误差, $E_{\text{object}1}-E_{\text{object}n}$ 表示草图图像的 1 至 n 个检索目标与图像集中当前图像对应的 1 至 n 个目标的目标匹配误差。

[0053] 将草图图像与图像集中每一图像均采用上述步骤进行处理, 获得对应的场景匹配误差, 再将所述草图图像与图像集中每一图像的场景匹配误差按照大小排序获得检索结果。

[0054] 本发明实施例根据草图中各个检索目标的特征从图像库中来筛选出包含这些特征的图像集, 再利用各个检索目标的位置关系及其相似度来实现多目标的快速检索。

[0055] 通过以上的实施方式的描述, 本领域的技术人员可以清楚地了解到上述实施例可以通过软件实现, 也可以借助软件加必要的通用硬件平台的方式来实现。基于这样的理解, 上述实施例的技术方案可以以软件产品的形式体现出来, 该软件产品可以存储在一个非易失性存储介质(可以是 CD-ROM, U 盘, 移动硬盘等)中, 包括若干指令用以使得一台计算机设备(可以是个人计算机, 服务器, 或者网络设备等)执行本发明各个实施例所述的方法。

[0056] 以上所述, 仅为本发明较佳的具体实施方式, 但本发明的保护范围并不局限于此, 任何熟悉本技术领域的技术人员在本发明披露的技术范围内, 可轻易想到的变化或替换, 都应涵盖在本发明的保护范围之内。因此, 本发明的保护范围应该以权利要求书的保护范围为准。

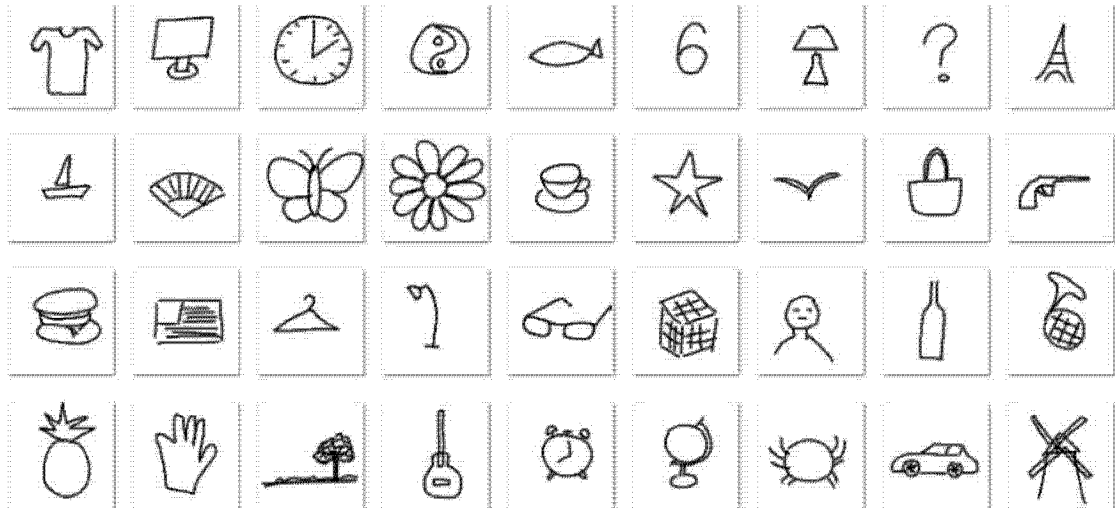


图 1

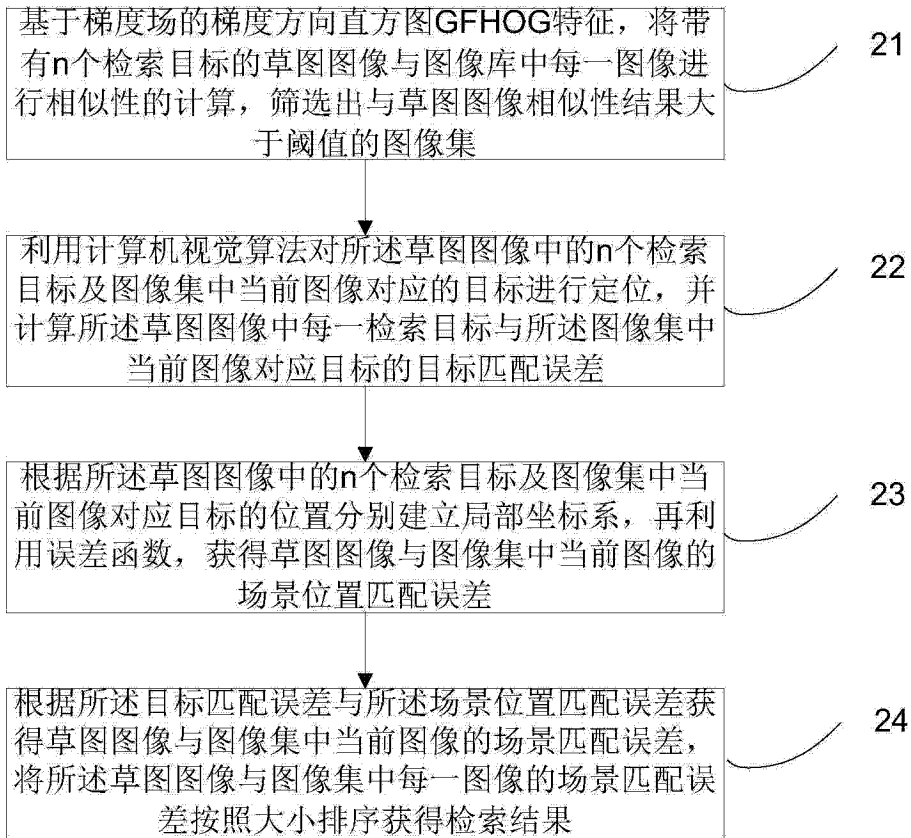


图 2

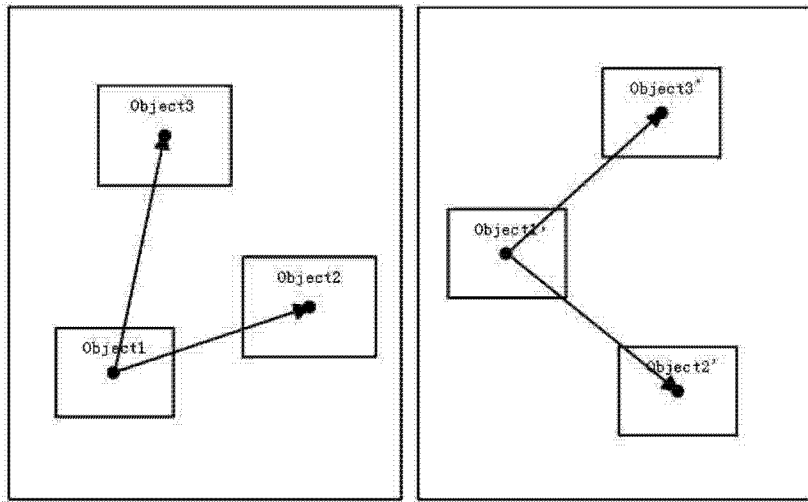


图 3