

[19] 中华人民共和国国家知识产权局

[51] Int. Cl.
G06F 17/30 (2006.01)
G06F 17/40 (2006.01)



[12] 发明专利申请公布说明书

[21] 申请号 200780017842.6

[43] 公开日 2009年6月3日

[11] 公开号 CN 101449271A

[22] 申请日 2007.5.17
 [21] 申请号 200780017842.6
 [30] 优先权
 [32] 2006.5.19 [33] US [31] 11/419,368
 [86] 国际申请 PCT/US2007/012193 2007.5.17
 [87] 国际公布 WO2007/136861 英 2007.11.29
 [85] 进入国家阶段日期 2008.11.17
 [71] 申请人 微软公司
 地址 美国华盛顿州
 [72] 发明人 L·张 X-J·王 F·景
 W-Y·马

[74] 专利代理机构 上海专利商标事务所有限公司
 代理人 陈斌

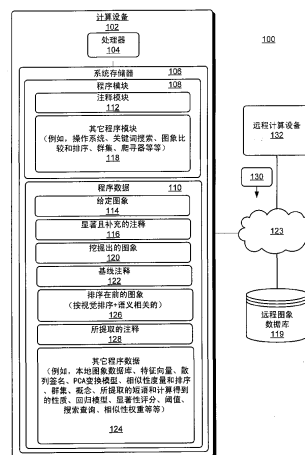
权利要求书4页 说明书10页 附图3页

[54] 发明名称

通过搜索进行注释

[57] 摘要

描述了通过搜索进行注释。在一个方面，在数据存储中搜索与给定图象的基线注释语义相关且与该给定图象在视觉上相似的图象。然后用关联于语义和视觉相关的图象的至少一个子集的注释的公共概念来注释该给定图象。



1.一种包含可由处理器执行的计算机程序指令的计算机可读存储介质，所述计算机程序指令包括用于执行以下步骤的指令：

在一或多个数据存储中搜索与给定图象的基线注释语义相关且在视觉上与所述给定图象相似的图象；以及

用关联于所述图象的至少一个子集的注释的共同概念来注释所述给定图象。

2.如权利要求1所述的计算机可读存储介质，其特征在于，所述计算机程序指令还包括用于在注释所述给定图象之前从所述共同概念中移除重复项的指令。

3.如权利要求1所述的计算机可读存储介质，其特征在于，所述计算机程序指令还包括用于接收包含所述给定图象和所述基线注释的搜索查询的指令。

4.如权利要求1所述的计算机可读存储介质，其特征在于，所述计算机程序指令还包括用于执行以下步骤的指令：

度量所述语义相关图象与所述给定图象的视觉相似性；以及

其中，所述图象的至少一个子集包括被确定为比其它图象在视觉上与所述给定图象更相似的一些图象。

5.如权利要求4所述的计算机可读存储介质，其特征在于，所述计算机程序指令还包括用于执行以下步骤的指令：

计算从所述各个注释中提取的短语的性质，所述注释被分组到多个群集中的一个相应群集中；

对于每个群集，组合所述性质中相关联的一些性质以生成所述群集的显著性得分；以及

其中，所述共同概念与同关联于所述群集中的其它一些群集的显著性得分相比具有更高的显著性得分的群集相关联。

6.一种计算机实现的方法包括：

搜索与给定图象的基线注释语义相关的图象的第一集合；

评估所述图象的第一集合中的每个图象以标识至少在概念上与所述给定

图象相关的图象的第二集合；

标识关联于所述图象的第二集合的注释；

确定所述各个注释上的共同概念；以及

用所述概念的至少一个子集注释所述给定图象作为补充注释。

7.如权利要求6所述的方法，其特征在于，还包括接收包含所述给定图象和所述基线注释的输入。

8.如权利要求6所述的方法，其特征在于，所述搜索还包括将所述基线注释与关联于所述图象的第一集合的文本相关。

9.如权利要求6所述的方法，其特征在于，所述评估还包括将编码所述给定图象的视觉特征的散列签名与编码所述图象的第一集合中的每个图象的视觉特征的相应散列签名比较。

10.如权利要求6所述的方法，其特征在于，所述评估还包括：

度量所述图象的第一集合中的每个图象与所述给定图象的视觉相似性；

基于在所述度量中确定的视觉相似性度量，标识比所述图象的第一集合中的其它图象在视觉上与所述给定图象更相关的排序在前的图象的集合；以及

其中，所述图象的第二集合是所述排序在前的图象。

11.如权利要求6所述的方法，其特征在于，标识注释还包括从所述图象的第二集合的上下文提取文本。

12.如权利要求6所述的方法，其特征在于，确定概念还包括：

群集所述注释，使得所述群集中的每个群集与所述群集的显著概念相关联；

确定关联于所述群集的多个显著概念中的每个显著概念的显著性得分；

基于所述各个显著性得分从所述多个显著概念中选择排序在前的显著概念的集合；以及

其中，所述排序在前的显著概念的至少一个子集是补充注释。

13.一种计算设备包括：

处理器；以及

耦合到所述处理器的存储器，所述存储器包括计算机程序指令，当由所述处理器执行所述计算机程序指令时，执行下列操作：

接收包含给定图象和所述给定图象的基线注释的搜索查询；以及
用关联于图象的第一集合的附加注释补充所述给定图象，所述图象的第一集合中的每个图象在视觉上与所述给定图象相似并且关联于与所述基线注释语义相关的注释。

14.如权利要求 13 所述的计算设备，其特征在于，所述附加注释表示相应注释分组的共同概念。

15.如权利要求 13 所述的计算设备，其特征在于，所述附加注释表示下列各项中的一或多个：标题、描述、类别、来自网页的信息、文件夹名字、统一资源标识符的至少一部分以及关联于不在所述图象的第一集合中的不同图象的文本。

16.如权利要求 13 所述的计算设备，其特征在于，所述图象的第一集合是图象的较大集合的子集，所述图象的较大集合中的每个图象与所述基线注释语义相关并且独立于与所述给定图象的视觉相似性。

17.如权利要求 13 所述的计算设备，其特征在于，所述计算机程序指令还包括用于从自一或多个图象数据库挖掘出的图象的第二集合中得到所述图象的第一集合的指令，所述图象的第二集合中的每个图象是基于关联于所述给定图象的文本挖掘出来的。

18.如权利要求 13 所述的计算设备，其特征在于，所述计算机程序指令还包括用于执行以下步骤的指令：

执行基于关键词的搜索以查找与所述基线注释语义相关的图象的第二集合；

在所述图象的第二集合中搜索具有与所述给定图象实质性的视觉相似性的一或多个图象；

从所述一或多个图象的一或多个上下文中提取文本；以及

其中，所述文本至少包括所述附加注释。

19.如权利要求 18 所述的计算设备，其特征在于，所述计算机程序指令还包括用于执行以下步骤的指令：

标识所述文本中的短语；

基于相应的短语性质向所述短语分配显著性得分；以及

其中,所述附加注释包括所述文本中具有比关联于所述文本的其它部分的显著性得分高的显著性得分的至少一部分。

20.如权利要求 19 所述的计算设备,其特征在于,所述相应的短语性质包括短语频次和短语对注释频次中的一或多个。

通过搜索进行注释

背景

注释数字图象的传统方法是用一或多个语义相关的关键词注释每个数字图象。这样的关键词通常用于促进在基于计算机的搜索环境中进行基于关键词的图象搜索和检索操作（例如，跨计算设备、数据库、因特网等等）。因为在这样的搜索环境中一般存在非常大量的数字图象，为促进图象搜索和检索操作而对数字图象进行的人工注释是劳动密集且很耗时的任务。

概述

描述通过搜索进行注释。在一个方面，在数据存储中搜索与给定图象的基线注释语义相关且与该给定图象在视觉上相似的图象。然后以关联于语义和视觉上相关的图象的至少一个子集的注释的共同概念来注释该给定图象。

提供本概述以使用简化的形式介绍在下面的详细描述中进一步描述的一些概念。本概要不旨在标识所要求保护的主题的关键特征或基本特征，也不旨在用于帮助确定所要求保护的主题的范围。

附图简述

在附图中，组件参考标号的最高位数字标识该组件在其中首次出现的具体附图。

图 1 按照一个实施例示出用于通过搜索进行注释的示例性系统。

图 2 按照一个实施例示出用于通过搜索进行注释的示例性过程。

图 3 按照一个实施例示出用于通过搜索进行注释的示例性数据和过程流。

详细描述

综览

描述用于通过搜索进行注释的系统和方法。该系统和方法对本地和/或远

程数据库执行基于关键词的搜索以查找与给定图象的基线注释语义相关的数字图象。该系统和方法度量每一个语义相关图象与该给定图象的视觉相似性。这些视觉相似性度量用于选择在视觉空间中比其它语义相关图象与给定图象更相关的图象。这些选择的图象在文本和视觉空间中与给定图象相似。关联于所选图象的候选补充注释是从所选图象的相应上下文中提取的。群集所提取的注释。每个群集包括具有共同或相关概念的注释。例如，关联于蜜蜂在玫瑰上的图象的注释的群集可包含与蜜蜂、玫瑰等相关的注释。该系统和方法使用一或多个评分准则来对概念的显著程度 (saliency) 排序以标识排序之前的概念的集合。该系统和方法用由排序在前的概念提供的信息的至少一个子集来注释给定图象。

现在更详细地描述用于通过搜索进行注释的系统和方法的这些和其它方面。

示例性系统

尽管不是必需的，但在由诸如个人计算机的计算设备执行的计算机可执行指令的一般上下文中描述用于通过搜索进行注释的系统和方法。程序模块通常包括执行特定任务或实现特定抽象数据类型的例程、程序、对象、组件、数据结构等等。尽管在上述上下文中描述该系统和方法，但在下文描述的动作和操作也可用硬件来实现。

图 1 按照一个实施例示出用于注释图象的示例性系统 100。系统 100 包括例如计算设备 102。计算设备 102 表示任何类型的计算设备，诸如通用计算设备、服务器、膝上型计算机、移动计算设备等。计算设备 102 包括耦合至系统存储器 106 的一或多个处理器 104。系统存储器 106 包括易失性存储器（例如 RAM）和非易失性存储器（例如 ROM、闪存、硬盘、光盘等）。系统存储器 106 包括计算机程序模块 108 和程序数据 110。处理器 104 从各个程序模块 108 取得并执行计算机程序指令。程序模块 108 包括例如用无限词汇表的显著且补充的注释来自动注释给定图象 114 的注释模块 112。程序模块 108 还包括例如诸如提供运行时环境的操作系统、基于关键词的搜索应用程序、图象比较和排序应用程序、文本/文档群集应用程序、web 爬寻器应用程序等等之类的其它程序模块 118。

为标识显著且补充的注释 116, 注释模块 112 对远程数据库 119 和/或本地数据库执行标准的基于关键词的搜索以查找与给定图象 114 的基线注释 122 语义相关的数字图象 (即挖掘出的图象 120) 集合。在一个实现中, 基线注释 122 表示关联于给定图象 114 的位置、事件名字、文件夹名字、描述等、网页上围绕给定图象 114 的文本、关联于与给定图象 114 有关 (例如, 在同一网页上、在同一文件夹中等等) 的其它图象的文本、从用户收到的查询输入等等之类。本地和/或远程数据库表示用于数字图象的任何类型的数据存储。例如, 数据存储可以是 CD、软盘和/或在任何其它可移动存储介质上、硬盘等中的一或多个。为了示例性说明, 本地数据库示为“其它程序数据” 124 的相应部分。在一个实现中, 本地和/或远程数据库是由爬寻器程序模块用带注释的数字图象填充的、由个人上传的等等之类。在该实现中, 远程数据库 119 示为通过网络 123 耦合至计算设备 102。网络 123 可包括局域网 (LAN) 和一般广域网 (WAN) 通信环境的任何组合, 诸如在办公室、企业级计算机网络、内联网和因特网中常见的网络。

视觉相似性度量

所述基于关键词的搜索操作返回与基线注释 122 语义相关的挖掘出的图象 120。然而, 内容作者使用的词汇差别可能非常大, 并且自然语言中的大部分词语具有固有的歧义性。因此, 关键词不总是对相关文档内容的良好描述符。这样的歧义性通常导致基于关键词的搜索的关键词/文档项失配问题。因为表示注释的关键词可能是有歧义的, 所以注释模块 112 还评估挖掘出的图象 120 以标识还至少在概念上与给定图象 114 相关的图象 (所标识的图象也可能是给定图象 114 的复制品和/或实质上在视觉上与给定图象 114 相似)。更具体地, 注释模块 112 比较挖掘出的图象 120 的视觉特征和给定图象 114 的视觉特征以生成视觉相似性度量 (即排序) 来标识不仅在文本空间中 (即语义相关) 而且在视觉空间中与给定图象 114 相关的图象。(视觉相似性度量被示为“其它程序数据” 124 的相应部分)。存在许多已知的用于比较数字图象之间的视觉相似性的技术。

在该实现中, 例如, 注释模块 112 如下编码给定图象 114 和挖掘出的图象 120 的视觉特征以生成每个数字图象的相应散列签名。在该实现中, 注释模块

112 使用众所周知的基于 N-箱（例如 36 箱）色彩相关图的操作来为给定图象 114 和每个挖掘出的图象 120 创建相应的特征向量。除了基于色彩相关图的技术，还可使用其它已知技术来生成特征向量。注释模块 112 将所生成的特征向量与从大量数字图象（例如数百万图象）生成的 PCA 变换模型相乘用于随后量化为二进制位进行比较。例如，如果特征向量的一个元素大于关联于 PCA 变换模型的最大本征向量的平均值，则注释模块 112 将该元素量化（编码）为 1，否则将该元素编码为 0。这些二进制编码表示相应的散列签名。

注释模块 112 使用每个挖掘出的图象 120 的散列签名和给定图象 114 的散列签名来度量每个挖掘出的图象 120 与给定图象 114 之间的视觉相似性。注释模块 120 可以使用一或多个已知技术来度量这样的视觉相似性。例如，注释模块 112 可以实现数字图象视觉空间距离度量的多种技术中的一或多个，诸如散列码去重复加欧几里德距离、散列码距离、加权散列码距离或者可用于度量视觉相似性的其它技术。在一个实现中，例如，注释模块 112 使用散列码去重复加欧几里德距离来度量挖掘出的图象 120 与给定图象 114 之间的视觉相似性。在该示例中，注释模块 112 使用给定图象 114 的散列签名的高 n 位片段作为度量视觉相似性的指数。（高位对应于较大的本征向量，表示视觉空间中的较大距离）。注释模块 112 基于关联于语义相关图象的相关图计算欧几里德距离。

在另一个示例中，注释模块 112 使用散列码距离来度量挖掘出的图象 120 与给定图象 114 之间的视觉相似性。例如，在一个实现中，注释模块 112 使用海明距离来度量两个对应图象之间的不同位的数量并且确定图象相似性度量。对于使用加权散列码距离，因为通常认为高位为比低位更重要，所以注释模块 112 为（关联于正在比较的图象的）高位之间的差提供比低位之间的差大的权重。注释模块 112 将散列签名均匀地分到各箱中（例如，将 32 位的散列码分到 8 个箱中）并且用 $2^{8-i}, 1 \leq i \leq 8$ 对第 i 个箱加权。可修改该公式以将它调整至最佳的结果。例如，如果箱数不等于八，则调整该公式来表示正在使用的箱数。如此，第一图象与第二图象之间的视觉距离（排序）就是加权的海明距离。

注释模块 112 基于各个挖掘出的图象 120 与给定图象 114 之间的视觉相似性度量标识挖掘出的图象 120 中的 N 个排序在前的图象 126。在该实现中，参数 N 可以基于系统 100 的具体实现来配置。例如，在一个实现中， N 等于 2000，

但 N 也可以是不同值。

示例性注释预测

注释模块 112 检索或提取关联于排序在前的图象 126 的其它注释。这些其它注释被示为所提取的注释 128。这些其它注释是由注释模块（或者从不同的程序模块诸如 web 爬寻模块）从排序在前的图象 126 在本地和/或远程数据库中的相应位置中的上下文中检索到的。这样的上下文包括，例如资源（排序在前的图象 126）的标题、关联于资源的目录文件夹的名字、资源的描述、网页上或文本中关联于资源的文本、关联于与资源有关（例如在同一网页上、在同一文件夹、文档等中）的其它图象的文本、指定资源位置的通用资源标识符（URI）、资源的类别等等之类。

注释模块 112 群集所提取的注释 128（即将所提取的注释 128 的每一个作为一个相应的文档来对待）以标识所提取的注释 128 中排序在前的概念（例如词语和/或短语）。为了示例性说明，这样的群集和概念被示为“其它程序数据”124 的相应部分。在一个实现中，例如，注释模块 112 通过实现（或者从“其它程序模块”118 访问）群集操作来标识这些概念，群集操作诸如有 Zeng, H.J.、He, Q.C.、Chen, Z.和 Ma, W.Y 在英国谢菲尔德的第 27 届信息检索研究与发展国际年会（2004 年七月）中第 210-217 页“Learning to Cluster Web Search Results(学习群集 web 搜索结果)”中所述的操作，尽管也可使用不同的群集技术。群集操作生成多个群集，每个群集分配相应的共同概念（或者显著短语）。

为获得每个群集的共同概念（群集名字），群集操作从所提取的注释 128 提取短语（ n 元语法），并且计算每一个所提取短语的一或多个性质（例如，短语频次、文档频次（即短语对注释的频次）等等）。群集操作应用预配置的回归模型将计算得到的性质组合成每个群集的唯一显著性得分。使用显著性得分，群集操作使用排序在前的短语作为候选群集的共同概念（名字）。这些候选群集进一步按照它们相应的文档（即关联于数字图象的文档）来合并以向每个群集分配最终显著性得分。

在该实现中，并且对于每个群集，使用阈值来为该群集合并（筛选）候选的排序在前的图象 126（也可使用其它技术来对于各个群集合并图象）。阈值等于图象相似性权重*平均视觉相似性得分。在该实现中，相似性权重基于特

定的群集操作实现选自 0.8 到 1.6 的范围，但也可使用其它范围。例如，在该实现中，相似性权重等于 1.2（当相似性权重等于 1.6 时，没有足够的图象用于群集，而当相似性权重被设置为.08 时，几乎所有的图象用于群集）。平均图象视觉相似性得分等于图象 i 与给定图象 114 的视觉相似性的总和（即由上述关键词和视觉筛选操作输出）除以挖掘出的图象 120 的数量。

群集的最终显著性得分可以使用各种准则来确定。例如，群集的最终显著性得分可以使用最大群集大小准则、平均成员图象得分准则等来确定。对于最大群集大小准则，即最大后验估计（MAP），该准则假设对于给定图象 114 具有最大显著性的共同概念是群集成员图象的主要概念的短语。因而，按照该技术，群集的最终显著性得分等于其成员图象的数量。在该实现中，例如，使用最大群集大小准则并且选择最大的 3 个群集（以及相关群集名字）。

关于使用平均成员图象得分准则来向群集分配显著性得分，该技术使用成员图象相似性的平均值作为群集的得分。如果群集中的所有图象是与基线注释 122 最相关的，则该群集的共同概念将有可能表示给定图象 114 的概念。为此，使用阈值（或者其它准则）作为确定是否要从最终的群集集合中省略排序在后的群集的间隔。在该实现中，阈值按照下面的公式来设置： $0.95 * (\text{排序为 } i-1 \text{ 的群集的得分})$ ，其中 i 表示当前群集的排序。如果当前群集（即，排序为 i 的群集）的得分小于该阈值，则不使用当前群集和所有其它更低排序的群集。如果多于特定数量的群集（例如，3 个或某个其它数量的群集）超过该阈值，则仅使用特定数量（例如，3 个或某个其它数量）的排序在前的群集来标识最终提取的概念。

在该实现中，注释模块 112 通过从关联于具有超过该阈值的最终显著性得分（排序）的群集的共同概念（群集名字）中移除重复的关键词、项等来生成具有显著且补充的注释 116。注释模块 112 用显著且补充的注释 116 注释给定图象 114。

示例性过程

图 2 按照一个实施例示出用于通过搜索进行注释的示例性过程 200。为了示例性说明和描述，参考图 1 的组件描述过程 200 的操作。在框 202，标识与给定图象 114 的基线注释 122 语义相关的数字图象（即挖掘出的图象 120）的

集合。作为一个示例，注释模块 112 对本地和/或远程数据库执行基于关键词的搜索以查找与给定图象 114 的基线注释 122 语义相关的数字图象（即挖掘出的图象 120）的集合。在一个实现中，给定图象 114 和基线注释 122 表示来自计算设备 102 的用户的搜索查询（在“其它程序数据” 124 中示出）。用户可以经由输入设备（未示出）诸如键盘、定点设备、语音识别等将命令和信息输入到计算机 102 中。在另一个实现中，给定图象 114 和基线注释 122 表示来自远程计算设备 132 的用户的搜索查询 130，远程计算设备 132 通过网络 123 耦合至计算设备 102。例如，远程计算设备 130 的用户。

在框 204，度量（排序）每个挖掘出的图象 120 与给定图象 114 的视觉相似性。在一个实现中，例如，注释模块 112 度量每个挖掘出的图象 120 与给定图象 114 的视觉相似性。存在多种可能的用于度量每个挖掘出的图象 120 与给定图象 114 之间的视觉相似性的技术。这样的技术包括，例如散列码去重复加欧几里德距离、散列码距离、加权散列码距离等等。

在框 206，标识（得到）挖掘出的图象 120 中与给定图象 114 在视觉上最相似的排序在前的图象 126。在一个示例中，注释模块 112 基于各个挖掘出的图象 120 与给定图象 114 之间的视觉相似性度量（即视觉相似性排序）从挖掘出的图象 120 中标识排序在前的图象 126。在该实现中，排序在前的图象 126 的数量可基于系统 100 的具体实现来配置。在框 208，从排序在前的图象 126 的上下文中提取关联于排序在前的图象 126 的其它注释（即，所提取的注释 128）。这样的上下文包括，例如资源（排序在前的图象 126）的标题、关联于资源的目录文件夹的名字、资源的描述、网页上或文本中关联于资源的文本、关联于与资源有关（例如在同一网页上、在同一文件夹、文档等中）的其它图象的文本、指定资源位置的通用资源标识符（URI）、资源的类别等等之类。在一个示例中，注释模块 112 检索或提取关联于排序在前的图象 126 的其它注释（即所提取的注释 128）。

在框 210，群集所提取的注释 128 以从所提取的注释 128 中标识排序在前的共同概念（例如，包含显著且补充的注释 116 的词语和/或短语）。这些排序在前的共同概念被示为显著且补充的注释 116。在一个实现中，注释模块 112 群集所提取的注释 128 以标识显著且补充的注释 116。在框 212，用显著且补

充的注释 116 的至少一个子集来注释（补充）给定图象 114。例如，在一个实现中，在用显著且补充的注释 116 注释给定图象 114 之前从显著且补充的注释 116 中移除重复的关键词、项等。在一个实现中，注释模块 112 用显著且补充的注释 116 的至少一个子集来注释给定图象 114。

图 3 按照一个实施例示出用于通过搜索进行注释的数据和过程流的示例性框架 300。为了示例性描述和参考，参考图 1 和 2 的组件和/或操作来描述框架 300 的各方面。在该描述中，参考标号的最高位数字指示其中该组件或操作首次出现的附图。

参考图 3，框架 300 示出用于通过搜索进行注释的三个阶段，包括（由数字(1)示出的）基于关键词的搜索阶段、（由数字(2)示出的）视觉特征搜索阶段和（由数字(3)示出的）注释群集或学习阶段。数据 302 表示关联于图象 304（即给定图象 114）和说明文字 306（即基线注释 122）的查询。在一个实现中，查询 302 是从计算设备 102 或远程计算设备 132 的用户收到的搜索查询（例如请见图 1 的查询 130）。在该示例中，查询图象 304 是至少包括湖、云和树的风景。可以理解，该查询图象 304 可以表示任何类型的图象。在该示例中，说明文字 306 指示用于注释查询图象 304 的词语“日落”。

操作 308 使用说明文字 306（“日落”）对数据库 310 执行基于关键词的搜索以标识与该查询图象语义相关的图象 312（即挖掘出的图象 120）的集合。数据库 310 表示图 1 的系统 100 的本地和/或远程图象数据库的任意组合。关联于说明文字 306 的信息的数据流在图 3 中用实线示出。操作 309 为查询图象 304 和语义相关图象 312 中的每一个生成相应的散列映射/签名 314 用于以后的相似性搜索操作。关联于给定图象 114 的信息的数据流在图 3 中用虚线示出。

操作 316 通过使用每个图象 312 的相应散列签名 314 和查询图象 304 的散列签名 314 来对语义相关图象 312 与查询图象 304 的视觉相似性排序（即度量）。存在各种用于使用散列签名 314 来度量该视觉相似性（或距离）的技术。这样的技术包括，例如散列码去重复加欧几里德距离、散列码距离、加权散列码距离等等。框架 300 使用视觉相似性排序来选择可配置数量的、在视觉上比其它图象 312 与查询图象 304 更相似的图象 312。这些在视觉上更相似的图象 312 被示为排序在前的图象 318（318-1 至 318-N）。排序在前的图象 318 表示图 1

的排序在前的图象 126。

框架 300 从关联于排序在前的图象 318 的上下文中提取注释。如上所述，这样的上下文包括，例如图象 318 的标题、关联于图象 318 的目录文件夹的名字、图象 318 的描述、网页上或者在文档中关联于图象 318 的文本、关联于与图象 318 有关（例如在同一网页上、在同一文件夹、文档等之中）的其它图象的文本、指定图象 318 的位置的通用资源标识符（URI）、图象 318 的类别等等之类。这些所提取的注释被示为所提取的注释 320（即图 1 的所提取的注释 128）。关联于所提取的注释 320 的数据流用关联于阶段(3)的虚点线来示出。

在框 322，框架 300 执行搜索结果群集（即分组）操作来从所提取的注释 320 的群集（例如群集 326-1 至 326-N）标识排序在前的共同概念（例如，补充注释 324-1 至 324-N）。在该实现中，这些共同概念是使用 Zeng, H.J.、He, Q.C.、Chen, Z.和 Ma, W.Y，于英国谢菲尔德的第 27 届信息检索研究与发展国际年会（2004 年七月）第 210-217 页“Learning to Cluster Web Search Results(学习群集 web 搜索结果)”中描述的操作来确定的，但也可使用不同的群集技术。

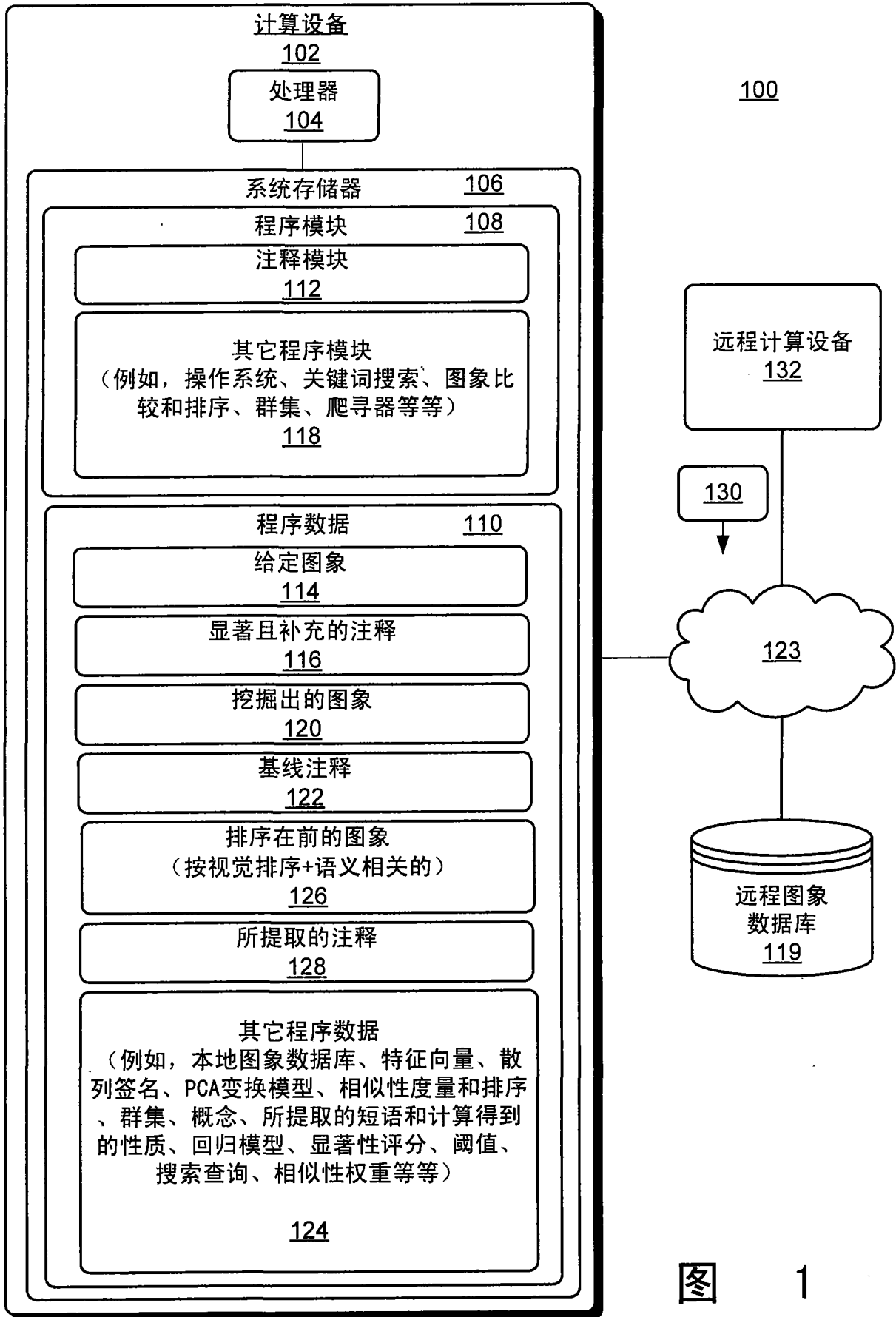
尽管每个群集 326 用相应的图象集合和基本的共同概念 324 示出，但群集 326 不包含图象。相反，每个群集 326 包含某些所提取的注释 320。该群集中的图象的例示仅仅是象征性的，以示出每个共同概念 324 与关联于所提取的注释 320 中相应的一或多个的每个排序在前的图象 318 的关系。框架 300 从共同概念 324 中移除重复关键词、项等等之类以生成补充注释 328。补充注释 328 表示图 1 的显著且补充的注释 116。在操作 330，框架 300 用补充注释 328 注释查询图象 304。

在一个实现中，所述框架 300 的操作是由图 1 的相应程序模块 108 执行的。例如，在一个实现中，注释模块 112 实现框架 300 的操作。

结论

尽管已经以专用于结构特征和/或方法学操作或动作的语言描述了用于通过搜索进行注释的系统和方法，但应理解，在所附权利要求书中定义的实现不必受限于上述这些特定特征或动作。例如，尽管系统 100 已经描述为首先标识语义相关图象然后确定语义相关图象的视觉相似性以生成排序在前的图象 126，但在另一个实现中，系统 100 仅使用下列之一来生成排

序在前的图象 126: (a)所述关键词搜索操作; 或者(b)在本地和/或远程数据库中搜索视觉上相似的图象。当仅从视觉上相似的图象生成排序在前的图象 126 时, 可与基线注释 122 无关地输入给定图象 114 作为查询图象。在另一个示例中, 在一个实现中, 计算设备 102 通过网络 123 向远程计算设备 132 提供通过搜索进行注释的服务。鉴于上述内容, 系统 100 的特定特征和操作是作为实现所要求保护的主题的示例性形式公开的。



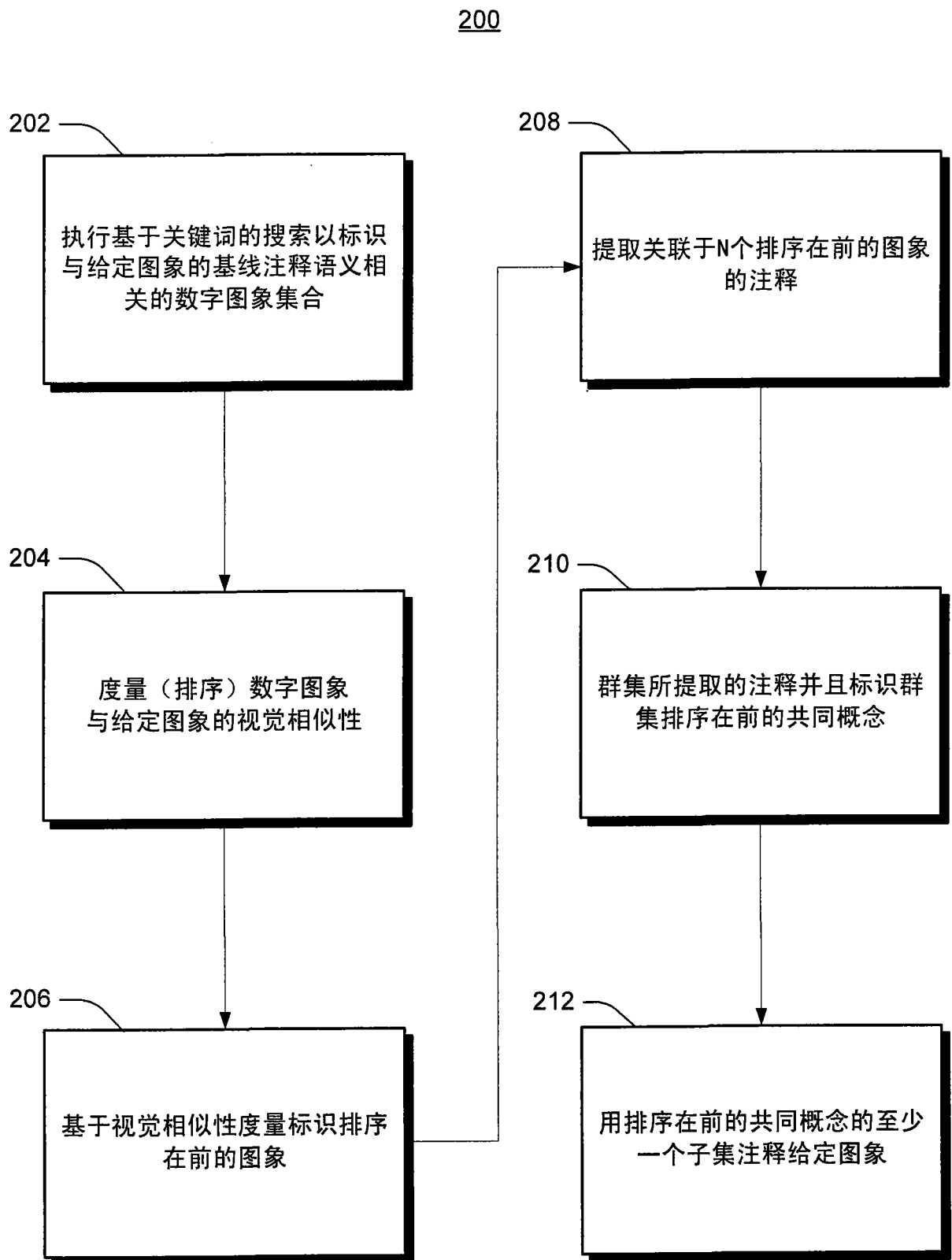


图 2

300

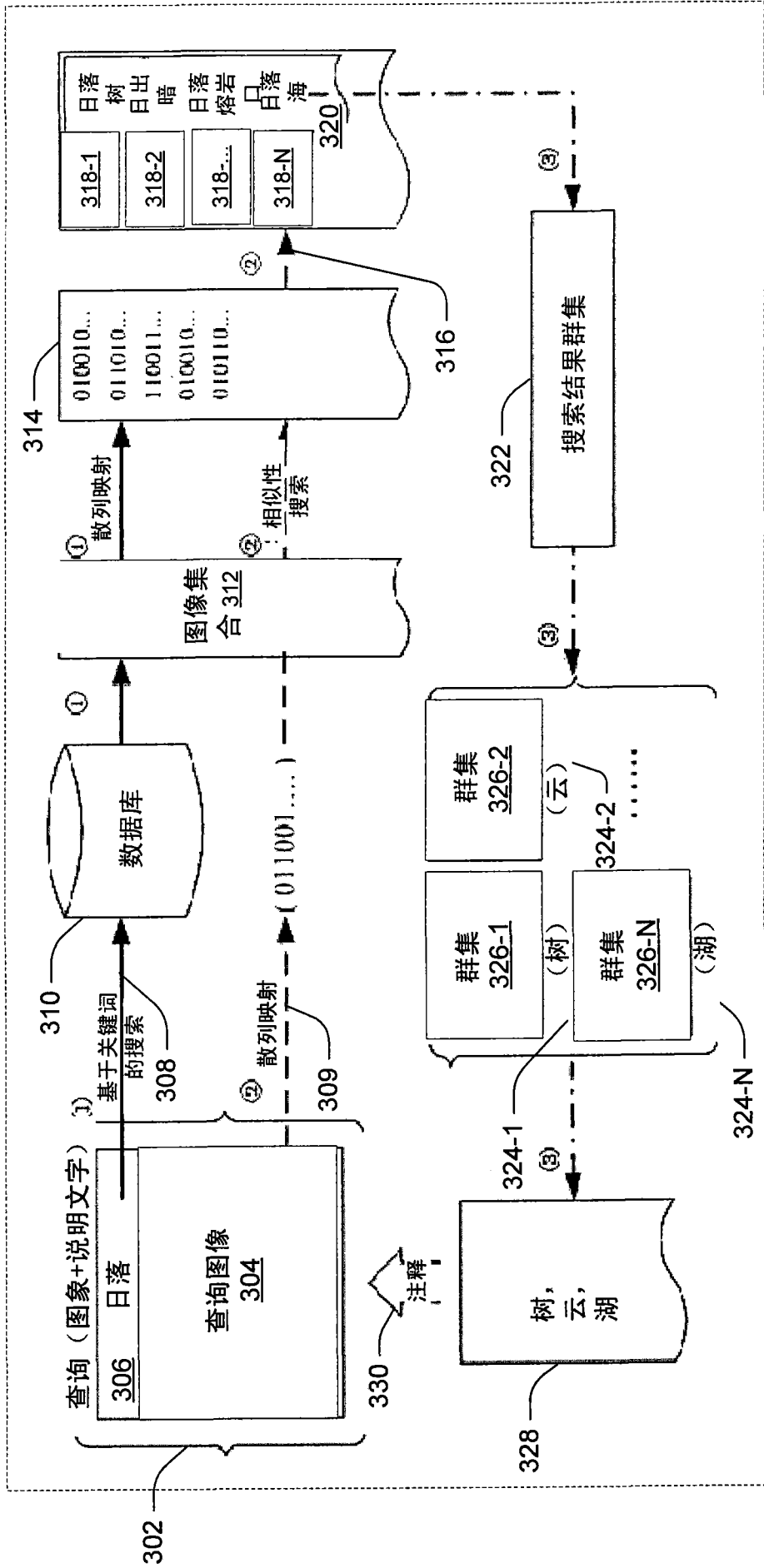


图 3