

[19] 中华人民共和国国家知识产权局

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



[12] 发明专利说明书

专利号 ZL 200380109658.6

[45] 授权公告日 2009年5月6日

[11] 授权公告号 CN 100485665C

[22] 申请日 2003.12.12

[21] 申请号 200380109658.6

[30] 优先权

[32] 2002.12.13 [33] US [31] 10/319,295

[86] 国际申请 PCT/US2003/039548 2003.12.12

[87] 国际公布 WO2004/055639 英 2004.7.1

[85] 进入国家阶段日期 2005.8.10

[73] 专利权人 国际商业机器公司

地址 美国纽约

[72] 发明人 约翰·史密斯 米林德·纳法德

阿波斯塔尔·纳特塞夫

[56] 参考文献

US6084595A 2000.7.4

US6405166B1 2002.6.11

US5794178A 1998.8.11

审查员 王楠

[74] 专利代理机构 中国国际贸易促进委员会专利
商标事务所

代理人 康建忠

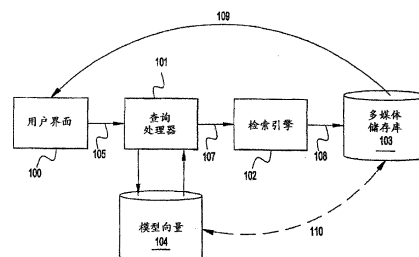
权利要求书4页 说明书10页 附图7页

[54] 发明名称

概念模型空间中的内容表示和检索的方法和
设备

[57] 摘要

一种从多媒体文档(104)抽取模型向量表示的方法和
设备。模型向量提供多媒体文档属于一组类别,
或者一组词义概念与该文档相关的置信度的多
维表示。模型向量可与多媒体文档相联系,从而提
供其内容或分类的索引,并且可被用于比较,搜
索,分类或群集多媒体文档。模型向量可被用于信
息发现,使多媒体内容个人化,以及查询多媒体信
息储存库(103)。



1、一种产生用于表示多媒体文档的至少一个模型向量的方法，包括下述步骤：

应用多个概念检测器对多媒体文档进行分类；

相对于每个概念检测器，对所述多媒体文档评分；和

把通过对所述多媒体文档评分所得的得分映射到多维空间中，从而产生至少一个模型向量表示。

2、按照权利要求1所述的方法，其中所述多个概念检测器对应于类别、对象、事件、场景和人物的固定词典。

3、按照权利要求1所述的方法，其中所述多媒体文档包含音频、视觉、文本和语音中的一种或多种形态，其中所述概念检测器使用构成所述多媒体文档的单个或多个形态进行检测。

4、按照权利要求1所述的方法，其中所述概念检测器使用从所述多媒体文档抽取的文档特征的基于内容的描述符进行检测。

5、按照权利要求1所述的方法，其中所述概念检测器使用与所述多媒体文档相关的元数据进行检测。

6、按照权利要求1所述的方法，其中所述概念检测器使用与所述多媒体文档相关的语境信息进行检测。

7、按照权利要求1所述的方法，其中所述概念检测器使用知识库进行检测。

8、按照权利要求1所述的方法，其中所述概念检测器对应于受过训练的模型统计分类器。

9、按照权利要求1所述的方法，其中所述评分基于下述至少之一：检测多媒体文档中的概念的置信度，概念与多媒体文档的关联性，以及就该概念来说该概念检测器的可靠性。

10、按照权利要求1所述的方法，其中所述映射对应于下述至少之一：级联所述得分，进行线性变换，进行非线性变换，进行量化，和通过设置阈值的维数缩减。

11、按照权利要求 10 所述的方法，其中所述映射可以是一对一，一对多，多对一或者多对多。

12、按照权利要求 1 所述的方法，其中对于多媒体文档的索引来说，所述多个概念检测器是固定的。

13、按照权利要求 1 所述的方法，其中对于多媒体文档的索引来说，所述评分和映射方法是固定的。

14、一种利用模型向量索引多媒体文档的方法，包括下述步骤：

基于来自多个概念检测器的输入产生每个多媒体文档的一个或多个模型向量；

使所述模型向量与对应的多媒体文档相联系；和

根据所述相关模型向量的值，建立用于访问所述多媒体文档的索引。

15、按照权利要求 14 所述的方法，其中所述产生每个多媒体文档的一个或多个模型向量的步骤包括下述步骤：

应用多个概念检测器对多媒体文档进行分类；

相对于每个概念检测器对所述多媒体文档评分；和

把通过对所述多媒体文档评分所得的得分映射到多维空间中，从而产生至少一个模型向量表示。

16、按照权利要求 15 所述的方法，其中对于多媒体文档的所述索引来说，所述多个概念检测器是固定的。

17、按照权利要求 15 所述的方法，其中对于多媒体文档的所述索引来说，所述评分和映射方法是固定的。

18、按照权利要求 14 所述的方法，其中所述多媒体文档包含音频、视觉、文本和语音中的一种或多种形态，其中所述概念检测器使用构成所述多媒体文档的单个或多个形态进行检测。

19、按照权利要求 18 所述的方法，其中根据存在于所述多媒体文档中多个形态，从所述多媒体文档抽取的文档特征的基于内容的描述符，与所述多媒体文档相关的元数据或者与所述多媒体文档相关的语境信息中的每一个，产生关于每个多媒体文档的多个模型向量。

20、按照权利要求 14 所述的方法，其中所述联系以数据库关键字值，

媒体定位符或者唯一地识别所述每个多媒体文档的标识符为基础。

21、按照权利要求 14 所述的方法，其中所述索引允许基于所述模型向量值的相似性搜索，最近邻居访问，或者范围搜索。

22、一种在关于多媒体文档的应用中使用模型向量的方法，包括下述步骤：

利用下述步骤产生用于表示每个多媒体文档的至少一个模型向量：

应用多个概念检测器对多媒体文档进行分类；

相对于每个概念检测器对所述多媒体文档评分；和

把通过对所述多媒体文档评分所得的得分映射到多维空间中，从而产生至少一个模型向量表示；和

根据所述至少一个模型向量表示的值，对所述多媒体文档进行检索操作。

23、按照权利要求 22 所述的方法，其中所述检索操作包含使用所述模型向量从多媒体信息储存库搜索和取回所述多媒体文档。

24、按照权利要求 22 所述的方法，其中所述检索操作包括下述至少之一：过滤多媒体信息，总结多媒体信息和使多媒体信息个人化。

25、按照权利要求 22 所述的方法，其中所述检索操作包括数据挖掘。

26、按照权利要求 22 所述的方法，其中所述检索操作包括群集所述多媒体文档。

27、按照权利要求 22 所述的方法，其中所述检索操作包括对所述多媒体文档分类。

28、一种在关于多媒体文档的应用中使用模型向量的系统，包括：

基于来自多个概念检测器的输入产生用于表示每个多媒体文档的至少一个模型向量的至少一个模型向量产生组件；和

根据所述至少一个向量表示的值，对所述多媒体文档执行至少一种操作的至少一个文档处理组件。

29、按照权利要求 28 所述的系统，其中所述至少一个模型向量产生组件包括：

应用多个概念检测器对多媒体文档进行分类的至少一个概念检测器

应用组件;

相对于每个概念检测器, 对所述多媒体文档评分的评分组件; 和
把通过对所述多媒体文档评分所得的得分映射到多维空间, 从而产生至少一个模型向量表示的映射组件;

根据所述至少一个模型向量表示的值, 对所述多媒体文档执行检索操作。

概念模型空间中的内容表示和检索的方法和设备

技术领域

本发明涉及使用模型向量来索引多媒体文档，更具体地说，涉及产生模型向量表示，使模型向量与多媒体文档相联系从而提供索引，并使用模型向量搜索、分类和群集多媒体文档的方法和设备。本发明还涉及把模型向量用于信息发现，使多媒体内容个人化，和查询多媒体信息储文库。

背景技术

随着视频、图像、文本和其它多媒体文档形式的数字信息的数量不断增长，越来越需要索引、搜索、分类和组织所述信息的更有效方法。内容分析、特征抽取和分类方面的最新进展正在提高有效地搜索和过滤多媒体文档的能力。但是，在能够从多媒体内容中自动抽取的低级特征描述，例如颜色、纹理、形状、动作等，和对多媒体系统的用户有用的语义描述，例如对象、事件、场景和人物之间仍然存在显著的差距。

多媒体索引的问题可由需要手工、半自动或者全自动处理的许多方法解决。一种方法使用允许人们手工把标记、类别或描述赋予多媒体文档的注释或编目工具。例如，M.Naphade,C.-Y.Lin,J.R.Smith,B. Tseng 和 S.Basu 在论文“Learning to Annotate Video Databases”，IS&T/SPIE Symposium on Electronic Imaging:Science and Technology-Storage & Retrieval for Image and Video Databases X,San Jose,CA, Jan.2002 中描述一种允许把标记分配给视频镜头的视频注释工具。他们还公开一种基于主动学习分配标记的半自动方法。全自动方法也是可能的。例如，M.Naphade,S.Basu 和 J.R.Smith 在“A Statistical Modeling Approach to Content-based Video Retrieval”,IEEE International Conference on

Acoustics,Speech and Signal Processing(ICASSP-2002),May,2002 中公开根据低级可视特征的统计建模,自动向视频内容分配标记的方法。自动标记技术可用于允许根据自动分配的标记搜索视频,但是,索引局限于匹配少量词汇的值,从而如果用户输入和标记项之一不相符的搜索项,那么搜索不会找到任何目标多媒体文档。

在自动化系统正在提高向多媒体文档赋予标记、类别和描述的能力的条件下,需要促进这些描述,以提供更有意义的利用所述描述索引、搜索、分类和群集这些文档的方式。此外,系统应考虑到自动化系统的不确定性或可靠性,以及赋予多媒体文档的任意标记、类别或描述的关联性,以便提供有效的索引。

发明内容

本发明的一个目的是提供一种利用捕捉任何自动标记的结果及其对应的得分,例如置信度、可靠性和关联性的模型向量表示法,索引多媒体文档的方法和设备。

本发明的另一目的是在信息发现,多媒体内容个人化和多媒体信息储存库的查询的应用中使用模型向量表示法。

为实现本发明的上述及其它目的,根据本发明的一个方面,提供一种产生用于表示多媒体文档的至少一个模型向量的方法,包括下述步骤:应用多个概念检测器对多媒体文档进行分类;相对于每个概念检测器,对所述多媒体文档评分;和把通过对所述多媒体文档评分所得的得分映射到多维空间中,从而产生至少一个模型向量表示。

根据本发明的另一方面,提供一种利用模型向量索引多媒体文档的方法,包括下述步骤:基于来自多个概念检测器的输入产生每个多媒体文档的一个或多个模型向量,每个概念检测器对应于来自固定的一组词典、类别、对象、特征、事件、场景以及人物的至少一个概念;使所述模型向量与对应的多媒体文档相联系;和根据所述相关模型向量的值,建立用于访问所述多媒体文档的索引。

所述模型向量表示法把多媒体文档的分类或标记结果以及任何对应的不确定性, 可靠性或关联性得分封装到多维向量中, 所述多维向量可被用于多媒体文档的搜索、分类和群集。模型向量表示法涉及词汇实体到多维向量空间中的各维的映射, 多维向量空间允许文档在多维空间中被表现和索引。

模型向量表示法的优点在于它在整个词典内广泛地捕捉标记。它还提供捕捉标记或分类结果的不确定性的紧凑表现。模型向量表示法还具有索引方面的优点, 因为其实值多维的特性便于度量空间中的有效索引, 允许模型向量表示法的距离或相似性的直接计算。这为有效方法使用模型向量进行多媒体文档的相似性搜索, 基于关联性反馈的搜索, 分类, 群集, 过滤等创造了条件。

附图说明

下面将参考附图, 更详细地说明本发明, 其中:

图 1 表示其中查询处理器把模型向量索引用于搜索的多媒体信息检索系统;

图 2 表示在给定词典和一组受过训练的检测器的情况下, 多媒体文档的模型向量的产生;

图 3 表示产生多媒体文档的模型向量的检测、评分和映射过程;

图 4 表示根据检测器评分而产生的模型向量的例子;

图 5 表示利用模型向量索引多媒体文档的过程;

图 6 表示使用模型向量的查询过程;

图 7 表示使用模型向量的多媒体文档的修改。

具体实施方式

图 1 描述了具有本发明的特征的多媒体信息检索系统的一个例子。如图所示, 用户通过用户界面 100 在步骤 105 中向多媒体信息检索系统发出查询。查询由查询处理器 101 处理。查询处理器在步骤 106 中搜索保存的一组索引值 104, 找出与用户查询的匹配物。就呈模型向量形式的

索引值来说，索引值对应于与每个被索引多媒体文档的语义维相关的多维向量。匹配物在步骤 107 中被传送给检索引擎，在步骤 108 中，从多媒体储存库 103 取回匹配的多媒体文档。根据模型向量与储存库的特定多媒体文档的联系 110，确定对应的多媒体文档。多媒体文档随后在步骤 109 中被返回给用户，并显示在用户界面 100 上。模型向量表示法提供一种表现可被用于找出用户查询的匹配物的一组保存的索引值 104 的方式。

模型向量表示法封装对多媒体文档应用一系列的检测器或分类器的结果。例如，考虑通过检测在多媒体文档中是否描述了这些概念，从下述词典 (lexicon)：{“car”, “boat”, “train”} 分配词汇实体的一组分类器。检测问题可被看作通过赋予反映每个概念存在的确定性的得分，检测每个概念的存在与否的一组二进制分类器。例如，系统可对“car”给出 0.75 的得分，它可被理解为赋予“car”标记的置信度为 75% 的含义。另一方面，对于“train”，系统可给出 0.25 的得分，它可被理解为赋予“train”标记的置信度为 25% 的含义。总的说来，系统产生这些多个检测器的得分，模型向量把这些得分记录在单一表示中，所述单一表示随后可被用作多媒体文档的索引。

图 2 描述了产生多媒体文档或查询的模型向量的过程。首先利用多个检测器 201 处理多媒体文档 200，并关于成为每个检测器的基础的概念对多媒体文档 200 评分。检测器本身可对应于固定词典 204 或者固定的一组类别、对象、事件、场景或人物。例如，分类辞典图形材料词库(TGM) 提供一组用于对照片和其它类型的图形文档分类的类别。检测器可被建立和使用，使得每个检测器对应于 TGM 类别之一。词典 204 的概念也可是类属的，特定的或者抽象的。例如，概念可对应于类属实体，例如“显示桥梁的场景”。另一方面，概念可对应于特定实体，例如“显示金门大桥的场景”。最后，概念可对应于抽象实体，例如“现代文明”。检测器 201 中对应于受过训练的模型或者其它类型的统计分类器。就训练 205 来说，被标记的多媒体文档的例子可在学习过程中被用于定义检测器 201 模型和它们的参数。检测器 201 的输出随后在映射过程 202 中被变换，从而产生模型向量 203。模型向量提供关于词典 204 的概念的多媒体文档 200

的一种累积评分。此外，模型向量 203 允许通过考虑其相对于词典的评分，推理 (reason) 多媒体文档 200。

图 3 描述了产生多媒体文档的模型向量的一种实现，其中利用一组检测器对多媒体文档进行分析和评分。通过应用 N 个检测器 301-303 对每个多媒体文档 300 分类。检测器可采取多种形式，包括支持向量机，高斯混合模型，隐马可夫模型，神经网络，Bayes (贝叶斯) 网络，线性判别分析等。每个检测器可代表特定的语义概念。例如，考虑词典：“car”，“boat”，“train”，检测器可如下表示概念：检测器 1=“car”，检测器 2=“boat”，和检测器 3=“train”。即，检测器 1 确定“car”概念是否与多媒体文档相关，其它检测器类似地工作。检测器可能先前已被训练，从而通过使用根据供给的地面实况标记例子，学习或建立模型的技术，检测它们相应的概念。

检测器 301-303 可使用与多媒体文档 300 相关的各种信息来进行各个检测。例如，检测器 301-303 可使用构成多媒体文档 300 的一种或多种形态的信息(视觉、音频、语音、文本)。检测器 301-303 还可使用从来自多媒体文档 300 的不同形态的信息中抽取的特征，例如颜色、纹理、形状、运动、声音频率、空间或时间布局的基于内容的描述符。例证的描述符包括颜色直方图，边缘直方图，运动向量，形状边界描述符等。检测器 301-303 还可使用与多媒体文档 300 相关的元数据。例如，可以使用诸如标题、作者、创建日期、类型之类的信息。另外，可以使用其它语境 (contextual) 信息，例如多媒体文档 300 与其它文档的关系。检测器 301-303 还可使用知识库或语义网，知识库或语义网允许基于与词典或多媒体信息储存库相关的信息和知识的组织的推断和推论。

对于每个检测器，关于每个多媒体文档产生得分 305。得分提供和检测器相对于多媒体文档 300 对其相应概念建模相关的信息。得分可反映许多内容，例如检测器依据其检测文档中的概念的置信度或不确定性(统称为“置信度”)，概念与文档的关联性，或者检测器在检测概念方面的可靠性。例如，考虑如上所述的检测器 1，得分可指示该检测器能够检测多媒体文档中“car”的描述的置信度。置信度和与决策边界或阈值的接近度

相关。例如，如果对于检测“car”来说，多媒体文档远离决策边界，那么可以断定高的置信度。但是，如果多媒体文档接近决策边界，那么可断定低的置信度。关联性得分可指示概念与多媒体文档多么相关。例如，如果“car”只被局部描述或者并不构成多媒体文档的重要部分，那么可确定低的关联性得分。另一方面，可靠性得分可指示对检测其相应概念来说，检测器有多可靠。例如，如果只利用“cars”的少数几个例子训练了检测器 1，那么可确定低的可靠性得分。但是，如果利用许多例子训练了检测器 1，那么可确定高的可靠性得分。得分本身可能只反映这些属性之一，例如产生一个一维值。但是，通过提供关于多个属性的信息，得分也可以是多维的。

一旦关于每个检测器产生了得分，那么这些得分被映射 304，从而产生模型向量 306。在一些情况下，为每个多媒体文档 300 产生单个模型向量 306，例如当每个检测器 301-303 使用多个形态(例如图像、视频、音频、文本、语音)来进行它们的分类时。另一方面，对于每个多媒体文档，可以产生多个模型向量 306，例如当每个检测器只使用一种形态时。这种情况下，对于每个多媒体文档可产生多个模型向量，以反映多个得分，例如一个与音频形态相关，另一个与图像形态相关，等等。

产生一个或多个模型向量的映射 304 提供产生自检测器的得分的组合或累积。在一些情况下，映射提供级联 N 个得分，从而产生一个 N 维向量的简单操作。例如，考虑上面的三元素词典：{“car”，“boat”，“train”}，其中每个检测器产生一个一维置信度得分(即，分类器 1 产生得分 C1，分类器 2 产生得分 C2，分类器 3 产生得分 C3)，随后级联操作产生三维模型向量 $M = [C1, C2, C3]$ 。另一方面，映射 304 可产生置信度得分的线性加权或变换。

置信度得分可由检测器的可靠性或分类结果的关联性加权。分别考虑三个检测器中的每一个的可靠性得分 R1、R2、R3。借助加权，通过把可靠性得分 Ri 乘以置信度得分 Ci，映射 304 可产生三维模型向量 $M = [R1 * C1, R2 * C2, R3 * C3]$ 。另一方面，分别考虑三个检测器中的每一个的关联性得分 L1、L2、L3，从而，通过把关联性得分 Li 乘以置信度得

分 C_i ，映射 304 可产生三维模型向量 $M = [L_1 * C_1, L_2 * C_2, L_3 * C_3]$ 。其它映射 304 可提供线性变换和/或维数减缩，例如在主成分分析，奇异值分解，小波变换，离散余弦变换等的情况下。另一方面，映射 304 可提供非线性信息，例如在支持向量机，神经网络等的情况下。映射 304 还可涉及对离散空间或二进制取值空间的量化。例如，通过在映射阶段 304 对来自检测器的置信度得分 305 设置阈值，可产生指示每个概念是否存在于多媒体文档 300 中的二进制模型向量。

总的说来，映射 304 可产生从各个概念或者检测器 301-303 到模型向量 306 的各维的各种特定映射。在一些情况下，例如就级联得分 305 的映射 304 来说，产生概念到模型向量维的一对一映射。但是，在其它情况下，可取的是产生多对一映射，以便相对于原始概念空间，缩减模型向量 306 的维数。在其它情况下，映射 304 可以是一对多或多对多，以允许模型向量 306 中一定程度的冗余。

图 4 表示根据检测器评分，产生的模型向量的例子。在已知关于固定词典 $=\{\text{"Cityscape"}, \text{"Face"}, \text{"Indoors"}, \text{"Landscape"}, \text{"Monologue"}, \text{"Outdoors"}, \text{"People"}, \text{"Text_Overlay"}\}$ 的经评分的一组检测器结果 400 的条件下，通过把每个检测器得分 400 映射到多维模型向量 401 的独特维上，产生模型向量 401。本例中，“Cityscape”的得分(它为 0.35)被映射到模型向量的第一维。“Face”的得分(它为 0.87)被映射到第二维，诸如此类。为了简化对应于不同多媒体文档的模型向量之间的匹配，可使用检测器得分到模型向量维的一致映射。类似地，更大词典的例证检测器得分 402 可被映射到模型向量维 403。这种情况下，“Animal”的得分被映射到模型向量的第一维。“Beach”的得分被映射到第二维，诸如此类。

图 5 表示了利用模型向量索引多媒体文档的过程。首先，在模型向量产生过程 501 中分析一批 K 个多媒体文档 500，从而产生一组 M 个模型向量 502。模型向量产生过程 501 可在所有多媒体文档 500 内使用固定词典 505 和对应的一组检测器，以便允许产生模型向量 502 方面的一致性。此外，出于相同原因，模型向量产生过程 501 还可在所有多媒体文

档 500 内把固定的一组参数用于评分 305 和映射 304。一旦产生了模型向量 502, 就可使它们与它们的对应多媒体文档 500 相联系。例如, 可利用数据库关键字值表示所述联系, 所述数据库关键字值陈述每个模型向量 502 和每个多媒体文档 500 之间的关键字-外部关键字关系。另一方面, 可通过使给出其对应的多媒体文档 500 的地址的媒体定位符与每个模型向量 502 相关联, 表示所述联系。另一方面, 可使用唯一地识别每个多媒体文档 500 的标识符来允许表现每个模型向量 502 的联系。还可通过在每个多媒体文档 500 的报头或者元数据字段中表示模型向量的值, 或者通过借助用水印作标记或者一些其它持久联系方法, 持久地使模型向量 502 的值与每个多媒体文档 500 相联系, 直接把模型向量 502 和每个多媒体文档 500 联系起来。

一旦产生了模型向量 502, 并且表示了它们与多媒体文档 500 的联系, 就建立允许根据模型向量 502 的值对多媒体文档 500 的访问 504 的索引。索引可允许基于接近度的访问, 以便允许相似性搜索法或者最近邻居搜索法。这些情况下, 通过提供查询模型向量完成访问, 并且根据索引找到相似模型向量或者固定大小的一组最近目标模型向量。索引还可支持基于范围的访问, 这种情况下, 提供查询模型向量, 根据索引找到在离查询模型向量固定距离内的所有目标模型向量。

图 6 表示了利用模型向量的查询过程。使模型向量的值与用户的查询相匹配, 以便检索多媒体文档。用户 600 向多媒体文档搜索系统 609 发出查询 601。查询可以采取用户提供的例证模型向量的形式。可选的是, 例如通过提供允许用户识别与查询相关的语义概念, 并对用于构成模型向量表示的得分赋值, 搜索界面可允许用户创建查询模型向量。另一方面, 可选的是, 搜索界面可向用户显示多媒体文档, 并允许用户选择对于所述查询来说, 哪些多媒体文档是相关的。系统随后会把预先计算的相关模型向量用于所述查询, 或者会在查询时产生模型向量。另一方面, 查询可采取用户提供的多媒体文档的例子的形式, 这种情况下, 通过使用模型向量产生过程来创建查询模型向量, 可分析和处理查询多媒体文档。

一旦查询模型向量可用，就在步骤 602 中把它们与保存的模型向量值 606 进行匹配。匹配过程可涉及使用索引结构来识别目标模型向量匹配物。匹配可涉及如上所述的相似性搜索，最近邻居搜索，或者范围查询。匹配过程 602 产生匹配物列表 603，它识别和查询模型向量相符的保存的模型向量 606。随后可选地在步骤 604 中对匹配物列表 603 评分。可根据利用模型向量值的度量空间计算确定匹配评分。例如，考虑单个查询模型向量，匹配评分可以在利用距离函数，例如欧几里德距离或曼哈顿距离的多维模型向量空间中测量的接近度为基础。另一方面，匹配过程可以只使用一些模型向量维数。例如，考虑 400-401 中的模型向量，如果用户只关心“human-related”概念，那么这种情况下，可有选择地使用第二维(“face”)和第七维(“people”)。在提供多个查询模型向量的情况下，通过组合自单个模型向量的距离的得分，可获得匹配评分。其它选择是可能的，例如计算查询模型向量的质心，并使用质心模型向量作为查询。

可选的是随后在步骤 605 中，对评分的匹配物列表排序，以便把最佳匹配物移动到列表的顶部。可选的是随后在步骤 606 中截短排序列表，例如在列表上保持 10 个最佳匹配物 (match)。结果 607 随后被提供给用户。可选的是，搜索系统可从多媒体储存库 610 中取回与结果列表 607 中的模型向量相关的那些多媒体文档，并向用户显示这些文档。

一旦结果被显示给用户，用户就可改进搜索，例如通过使用关联性反馈技术从结果列表 607 中识别正面例子和负面例子。搜索系统 609 可使用该信息和查询处理一起来根据模型向量得分检索匹配物。

虽然模型向量可被用于多媒体文档的检索，不过它们还可被用于对多媒体文档进行群集和分类。例如，可在多维度量空间中分析模型向量，以便利用各种技术，例如聚集群集法识别群集。还可利用各种有指导的学习方法，例如基于判别建模或生成建模的那些有指导学习方法，对模型向量分类。例证的分类器包括支持向量机和高斯混合模型。诸如主动学习和推动 (boosting) 之类的其它技术也可被应用于模型向量值以便分类。

模型向量还可被用于多媒体储存库的信息发现和挖掘。例如，可检

查一批模型向量的维数的相关性，以揭示当概念出现在多媒体文档中时概念的同现信息。

图 7 表示了使用模型向量对多媒体文档的修改。模型向量可被用于过滤、总结或使来自多媒体储存库的多媒体文档或信息个人化。用户 700 向多媒体信息系统发出请求 708。所述请求在步骤 701 中处理。请求可包含特定的用户查询，例如 601 中的查询，其中用户供给例证的模型向量或者多媒体文档，或者识别语义概念。另一方面，请求可以采取注册(login)的形式，其中已保存了用户简表以及用户偏爱信息。这种情况下，用户偏爱信息可选择地在步骤 702 中被检查。偏爱信息也可以例证的模型向量，多媒体文档或者识别的语义概念的形式被保存。用户查询和用户偏爱信息随后可被累积和处理，从而产生查询模型向量，查询模型向量随后被用于匹配模型向量 703 和检索保存的模型向量 704，保存的模型向量再被用作在步骤 706 中从多媒体储存库 705 中有选择地取回 (retrieve) 多媒体文档的索引。这提供基于模型向量值的多媒体文档的过滤。

可选的是，可结合查询模型向量使用与多媒体文档相关的模型向量，以便在步骤 707 中修改 (adapt) 多媒体文档的内容。所述修改可根据关于特定查询的用户偏爱使多媒体内容个性化。例如，用户偏爱可指示“sports”概念重要。这种情况下，取回的多媒体文档，例如“news”视频可被处理，以便只抽取“sports”片段。另一方面，所述修改可总结内容，例如通过压缩“non-sports”片段，并从“sports”片段抽取精彩场面。

参考优选实施例说明了本发明。显然在不脱离附加权利要求中限定的本发明的精神和范围的情况下，本领域的技术人员可做出多种修改。

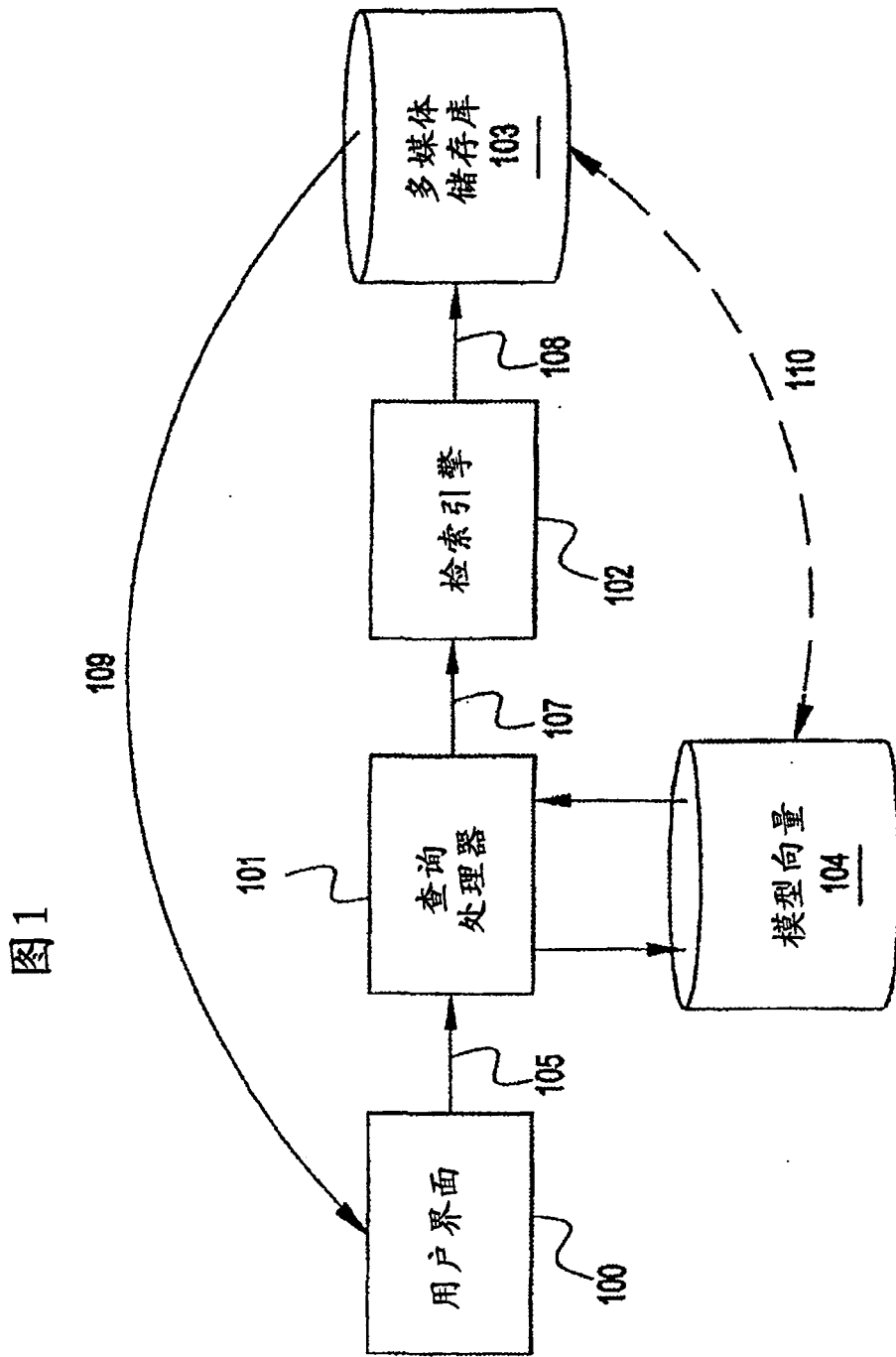


图1

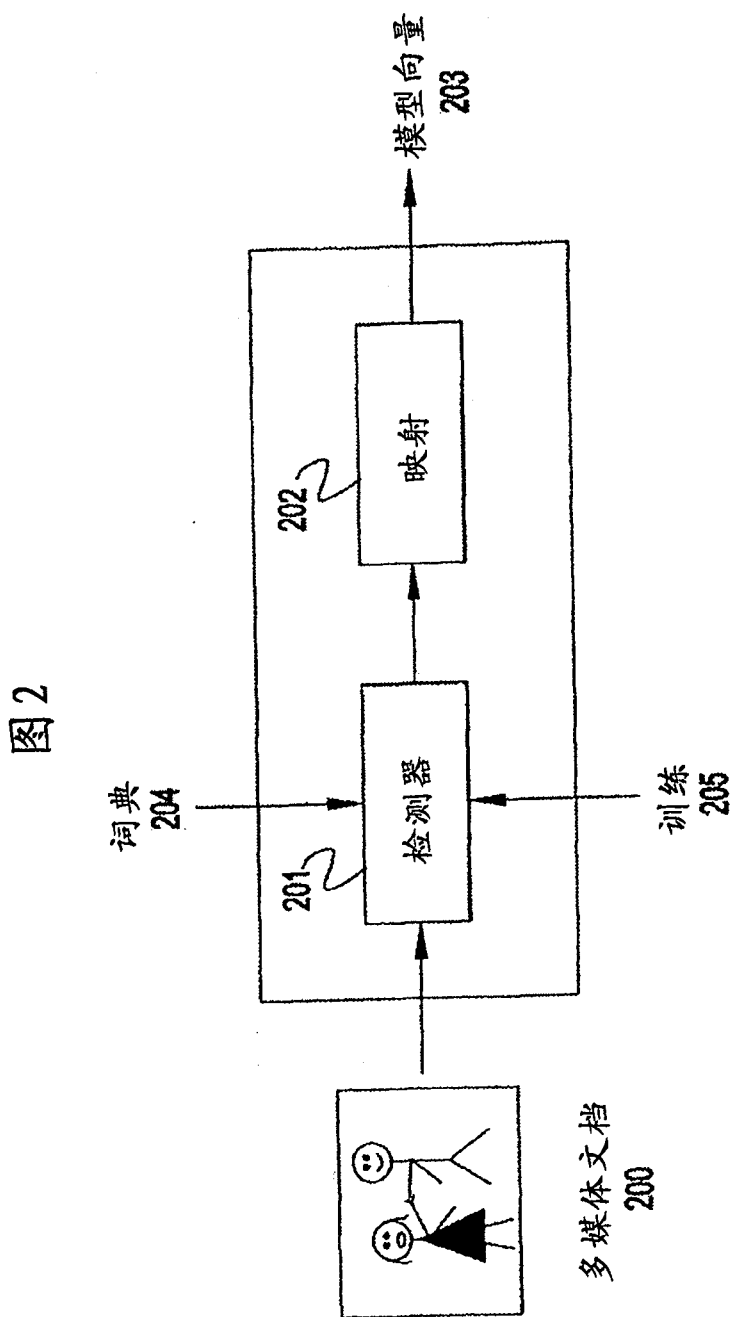


图3

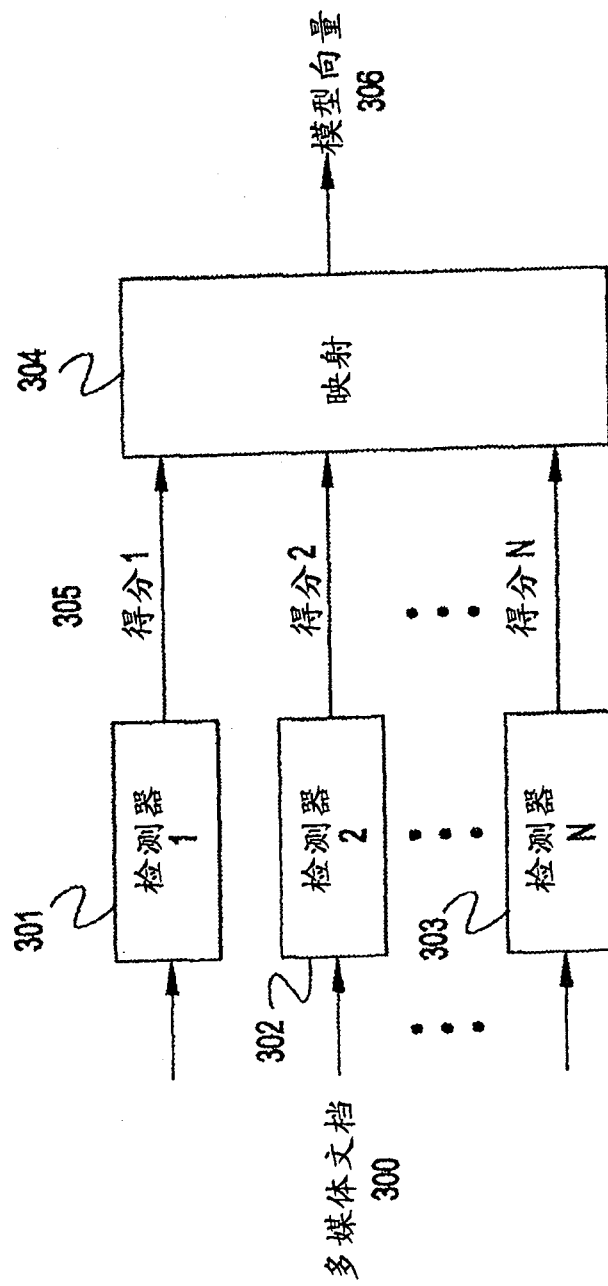


图 4

Cityscape (0.35)	
Face (0.87)	
Indoors (0.17)	
Landscape (0.7)	→ [0.35, 0.87, 0.17, 0.7, 0.17, 0.25, 0.44, 0.2]
Monologue (0.17)	
Outdoors (0.25)	
People (0.44)	
Text_Overlay (0.2)	

检测器得分 400

模型向量 401

Animal (0.88)	
Beach (0.98)	
Boat (0.51)	
Bridge (1)	
Building (0.25)	
Car (0.31)	[0.88, 0.98, 0.51, 1, 0.25, 0.31, 0.71, 0.22
Cartoon (0.71)	0.22, 0.26, 0.24, 0.73, 0.89, 0.59, 1, 0.66,
Cloud (0.22)	→ 0.67, 0.55, 0.87, 0.072, 0.14, 0.25, 0.57, 1.1,
Desert (0.22)	0.43, 0.47, 0.19, 0.77, 0.15]
Factory-Setting (0.26)	
Flag (0.24)	
Flower (0.73)	
Graphics (0.89)	
Greenery (0.59)	
Horse (1)	
House_Setting (0.66)	
Land (0.67)	
Man_Made_Scene (.055)	
Mountain (0.87)	
Office_Setting (0.072)	
Road (0.14)	
Sky (0.25)	
Smoke (0.57)	
Sport_Event (1.1)	
Tractor (0.43)	
Train (0.47)	
Transportation (0.19)	
Tree (0.77)	
Water_Body (0.15)	

检测器得分 402

模型向量 403

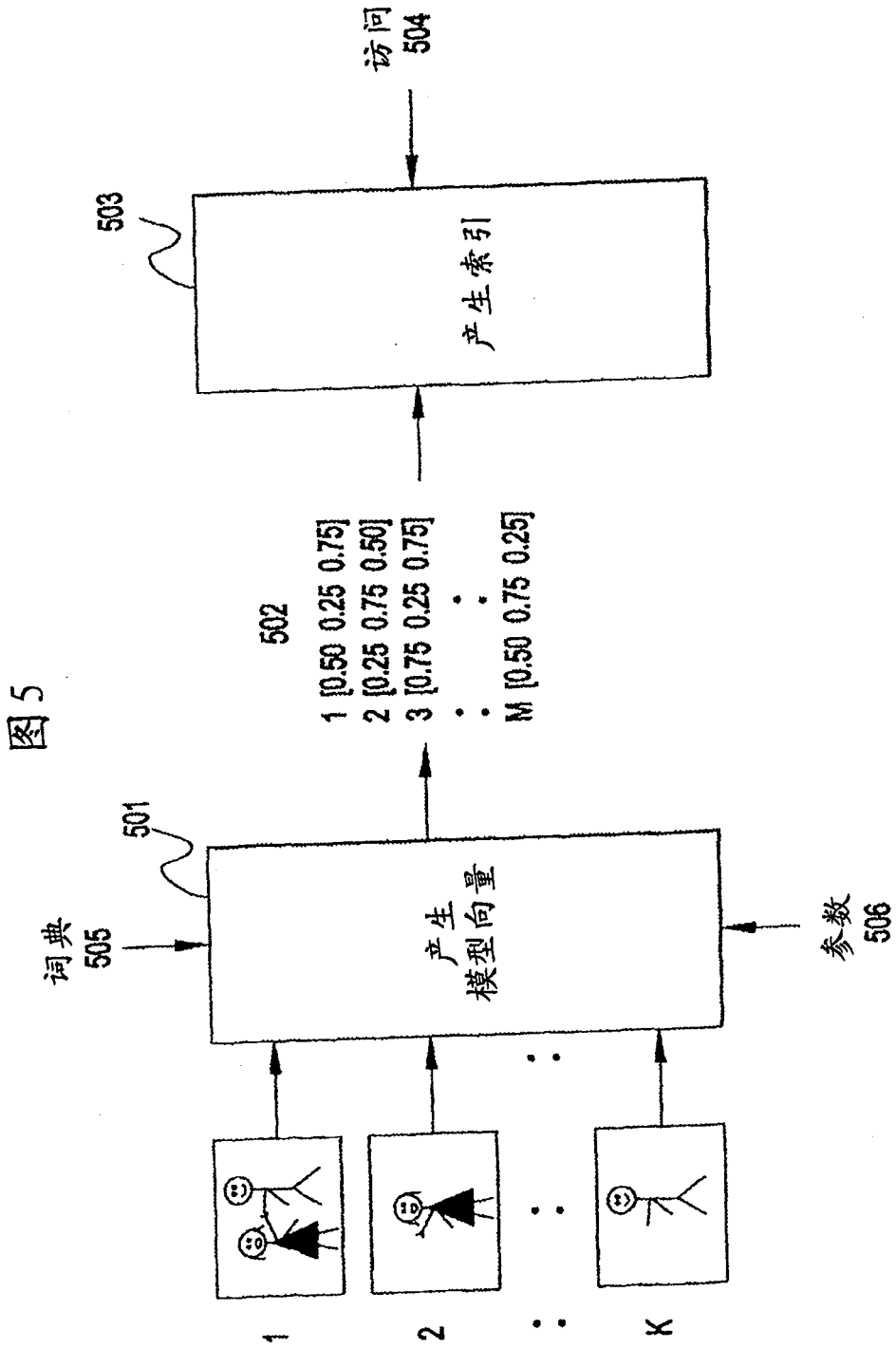


图6

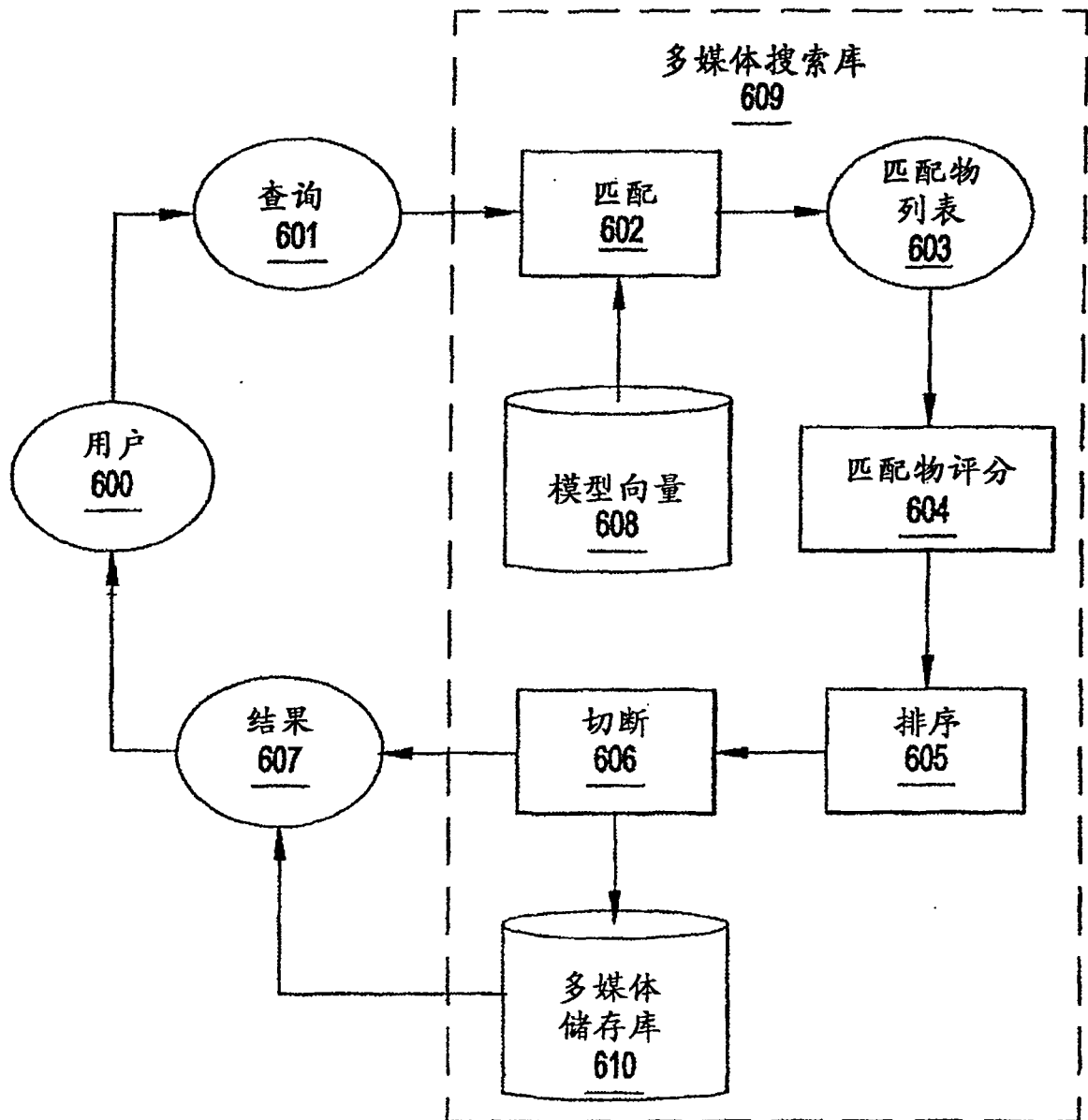


图7

