

摘要

随着新世纪互联网、社交媒体等技术及应用的飞速发展和个人消费数码产品的广泛普及，海量多媒体视觉数据尤其是图像数据的索引和检索问题正面临着前所未有的机遇与挑战。如今图像检索系统所面对的数据规模、问题的复杂程度急剧增加，而系统对检索性能、内存消耗和检索时间的要求同时日益严苛，相关方向的研究工作已经成为了学术界和工业界的关注热点。另一方面，语义鸿沟仍然阻隔在人类的高层语义表达和图像的底层特征之间，使得对图像信息内容的深入理解困难重重。融合不同层次的图像属性特征能够弥补单一特征表达能力的有限，避免对用户检索意图做过多的预先假设。迄今为止，大部分多层次属性信息融合的研究工作都集中在特征提取和在线检索阶段，大量的在线操作和复杂计算限制着这些方法在海量数据规模中的应用。本文提出在离线索引中融合不同层次的图像信息，避免了上述问题的出现，在不牺牲甚至提高了检索效率的同时，获得了出色的检索性能表现。

本文提出了一种融合图像局部视觉信息和空间属性信息的二进制视觉词组，即 binary phrase，并将其应用于大规模移动图像搜索领域。本文介绍了一种训练二进制特征视觉词典的方法，并且提出寻找具有相似尺度信息的局部特征构成 binary phrase。不同于传统的视觉词组，binary phrase 在记录图像近邻视觉单词的同时，还使用了特征间的尺度及方向的差异信息。这些嵌入到倒排索引中的空间线索被用于视觉特征的匹配及校验。在公开大规模移动搜索数据集上的实验证明该算法在不增加系统内存消耗的同时，在检索性能和及系统效率之间达到了非常好的平衡。

本文提出了一种观察和理解大规模数据集合的新视角，即以 superimage 作为图像索引和检索的基本单元来取代传统方法中的单幅图像。数据集中具有强语义一致性的图像被打包在一起构成 superimage，在特征提取和量化之后将其存储在倒排索引结构中。隶属于 superimage 的图像之间具有较强语义和视觉的相似性。在线检索时，只需要提取图像的局部特征，却能够在检索结果中返回与查询图像具有语义相似性的图像。同时，由于数据集中 superimage 的数目远小于单幅图像，因此内存的空间使用和在线检索的时间消耗都大为降低。公开数据集上的多个检索实验证明，算法能够在仅使用 1/2 左右内存的情况下，使得检索性能达到业界最好水平。

为了能够同时在倒排索引中考虑图像的多个不同侧面信息，本文提出了一种新颖的图像索引和检索方法，即 Hybrid-Indexing。通过融合在多个特征空间中建立的近邻关系图，以及应用图链接分析方法得到的图像权重，本文对具有特定视觉相似性的图像间的语义关系和全局关系进行了全面的考察。针对孤立图像、一般图像和显著图像

采取了不同的处理策略，增强了索引列表中图像的紧凑程度，并增加了丰富的多特征线索。该算法并不限制融合的图像特征的数目和类型，且只在离线索引阶段完成全部融合工作，因而比目前的大多数的特征融合方法都具有更强的适应性和表达能力。在公开评测数据集上的实验也证明了这一点。

综上所述，本文针对单一图像特征无法全面描述图像信息内容，以及用户检索意图难以精确描述的问题，提出在离线索引阶段融合多层次属性特征的几种不同方法，逐渐递进地朝着弥合语义鸿沟的方向做出了坚实的努力，同时也为今后该领域的研究工作奠定了基础。

关键词：海量图像检索，离线索引，属性特征融合