

图像检索技术比较研究^{*}

郑雪丽 徐严斌

【摘要】 图像检索在很多研究课题中都是核心技术，文章对图像检索的背景和未来趋势进行比较全面的概括，根据不同的方面对基于文本的图像检索技术、基于图像的图像检索技术和网络上的图像检索技术这3种当今主流的图像检索技术进行深入的探讨。通过收集大量资料，具体分析3种图像检索技术各自的特点，对这几种技术进行比较分析，并得出各自的优缺点及适用对象。

【关键词】 文本 图像 图像检索

Abstract: In many research subjects, image retrieval research is the core technology. This paper makes a comprehensive summarization of the background of image retrieval and the future trend. According to the different aspects, the paper discusses the text-based image retrieval technology, image-based image retrieval technology and image retrieval technology on the network, analyzes the specific characteristics of the three image retrieval technologies, and compares the advantages and disadvantages of these techniques.

Key words: text image image retrieval

DOI:10.15941/j.cnki.issn1001-0424.2017.16.010

0 引言

图像检索对于人们的生活而言变得越来越重要。自20世纪70年代起，面对愈来愈多的图像数据，人们迫切地需要能够迅速处理庞大的数据的办法，图像检索技术随即应运而生。而随着技术的不断革新，关于图像检索也有了更多的方式和手段。目前最主流的搜索网站大部分都具备图像检索功能。海内外知名的企业如：Google、百度、雅虎、SOSO等。这些网站大多使用的技术是基于文本的图像检索技术，这种技术比较容易操作，适用面较广，但存在许多明显的缺点：需要很多人员进行标注，非常耗费人力；受不同主观因素影响，不能保证正确率；不同的民族有不同的语言环境，这就导致了搜索的办法不能得到统一；图像资源的迅速增长导致搜索结果数量增加，准确率降低，使用户不能得到合适的结果。所以为了解决这些问题，基于图像的图像检索技术形成。这种技术采用自动匹配的描述方式来完成相关程序，有效地避免了关键字的问题，使搜索过程更加便捷，更能适应大多数环境。例如，这种技术可用于图像搜索引擎、网络环境净化、医学图像检索等诸多领域。到目前为止，有关学者在技术领域做了大量研究，取得了长足的进步，实时问题图像数据库的扩展提高了对检索速度的需求。

1 图像检索技术的理论

1.1 图像检索的概念

对图像检索技术的研究，是从20世纪70年代开始的，当时的研究主要是基于文本的图像检索技术，利用图像表征进行检索^[1]。比如现有的各大搜索引擎，不管是在谷歌（Google）还是在百度、UC上进行搜索，大多都是运用上述方法。输出图像表征而得到的图像都不是来自图像本身。由于这种检索方式比较单一，而且对图像内容的标注探索不够，这些搜索引擎在检索时的准确率都不能够达到让人满意的地步，每个个人用户不得不在成千上万幅的图像中再次选取，然后在这些图像中选择自己想要得到的结果，这些方式十分耗费人力和时间，直接影响查询结果的查全率和查准率。

为了解决这些问题，从1990年代开始，渐渐出现了基于内容的图像检索技术，这种技术是对图像的内容语

^{*} 本文系郑州航空工业管理学院青年科研基金项目“MOOC环境下高校图书馆学习支持平台构建研究”（项目编号：2017042001）的研究成果之一。

义,如图像的**形状、颜色、纹理**等更加详尽的方面进行分析和检索的图像检索技术。

图像检索技术是信息检索技术的一个重要分支,在信息检索领域占据相当重要的地位。图像检索技术的出现使人们查询图像时的效率得到了很大的提高。但是,伴随互联网的蓬勃发展,网络上充斥着各种各样的数字图像资源,人们发觉现在的图像检索技术渐渐不能满足人们日益增加的检索需求,人们愈加需要**准确率更高、检索速度更加迅速的图像检索技术**。新的图像检索技术急需被开发出来,因此**融合文本和图像**的从而使检索更加快捷便利的图像检索技术自然而然就变成了下一步技术革新的首选。

1.2 图像检索技术比较研究的概念

本文不仅只是探究图像检索技术各自的原理和发展方向,也是需要从横向以及纵向比较各种图像检索技术的不同之处和优缺点。每一种技术都有它的优势与不足,每种技术有不同之处也有相通之处,通过对图像检索技术的横向比较,研究出相通技术的相同属性在各自不同领域发挥作用时呈现的异同之处;通过对图像检索技术的纵向比较,研究出不同技术的相同属性在**互相学习、融合之后在各自不同领域发挥作用时得到的不同提升**。

2 图像检索技术研究的三个方向

主要研究现今图像检索技术中比较流行的3种,即基于文本的图像检索技术、基于图像的图像检索技术、网络上的图像检索技术。

2.1 基于文本的图像检索技术

基于文本的图像检索技术(TBIR),是将图像先储存在数据库中,然后利用图像文件名或关键字进行搜索^[2]。TBIR技术通过描述图像文件的特征和其他的附属信息创建有效的图像文件索引,建立关键词和图像的存储路径之间的映射关系,并建立图像文件的索引数据库。一般来讲,进行遍历式的分类浏览和关键词的检索是检索图像文件通常用的方法^[3]。实际上,检索操作是对文本内容的表述的查准确率匹配或查全率匹配,虽然有的检索模型也具有字典支持,但是我们的检索只涉及图像文件人工索引描述,而与图像文件内容的关系不大^[4]。

基于文本的图像检索技术最重要的是创立检索数据库。要创立检索数据库,第一点是需要提取图像的有关信息,并且可以自动识别以及提取图像的一些内容,比如图像的格式型号、名字和体积。提取图像的内容信息不是很简单,手动检索的自动识别方式有两种办法:创立检索数据库的艺术输入法,它首先由专业人士负责选择图像,然后对每个具体图像内容进行表述,给出关键字,再进行图像审查和检索。相当部分的搜索引擎图像库是以这种方式构建的,其精度相当高,但也限制了其回收率,搜索范围有限,仅包括部分网站手动选择。

我们最常见的TBIR系统的典型应用是:Yahoo Image Gallery(<http://gallery.yahoo.com/>),到目前为止,有超过4000万张图片^[5],把各个类型图像分成生物、艺术、文娱、自然、科技、体育和旅游等类别,然后进行分类索引,提供给使用者,使用者也可以通过关键字的方式进行图像检索。

2.2 基于图像的图像检索技术

根据统计,人们对外界的看法是通过各种各样的信息得到的,其中通过视觉信息获得约70%。通常,图像传达的信息比任何其他媒体更加丰富,生动直观,俗话说“百闻不如一见”,图像在人们的生活和工作学习中起着重要的作用。近年来随着图像采集技术的快速发展,电子产品(如数码相机、摄像手机、迷你型摄像机等),卫星多光谱成像系统扫描的快速传播,X射线、计算机断层摄影系统、电脑断层扫描、CT、超声成像、超声检查、工业机器人视觉系统等领域的专业成像设备的不断推广,特别是随着多媒体技术和网络技术的飞速发展,使数字图像等多媒体数据大规模扩大,大量的图像数据给人带来极大的便利的同时,也带来了如何有效组织、检索和表达这些图像的巨大问题,使用户从广泛的图像数据中及时有效地找到所需的图像,成为了近年来最突出的问题。基于图像的**图像检索技术是解决这个问题的有力工具**。

所谓的基于图像的图像检索(CBIR)是指软件自动分析图像,将图像的颜色、形状和纹理等内容特征以及这些特征的组合作为特征向量提取到图像特征中,通过图像分析从给定的检索图像中提取特征向量,并使用相似度匹配算法计算查询案例图像与特征数据库中的图像特征向量之间的相似度,搜索结果根据相似度输出^[6]。检索技术的高成本,导致了此技术现在只限于具体的区域,如犯罪现场拍摄嫌疑人的照片,然后根据照片与图像库匹配,以实现检索过程。

到20世纪90年代初,随着图像信息的增加,这种方法已成为主流。过去10年的图像搜索研究显示了一个典型的基于图像的图像检索系统。系统首先为每个图像**提取各种低级视觉特征图像库**,并以高维向量的形式存储在图像特征数据库中。系统检索用户提交的检索样本的特征,并根据图像数据库的相似度返回最相似的图像特点。

用户从搜索结果中选择相关和不相关的图像进行反馈。算法作出了一些战略性调整,逐渐接近所需的图像类别,并返回到最终搜索结果。

基于图像的图像检索技术已经发生了巨大的变动,并且成为了现今最重要的热点课题,但是随着对这个系统的更加深入的了解,以及对国内国外的高端技术的不断学习,很轻易就能发现此种技术的一些劣势。因为色彩、形态等等元素不能直观地表述,所以这种检索技术在提取图像信息时准确率不高。所以说这种图像检索技术仍然处于最开始的发展阶段,技术也在不断进步不断改进中。

2.3 网络上的图像检索技术

还有一种图像检索技术也是该领域的重点研究课题——网络上的图像检索技术,这种技术主要是通过对网络图像的视觉表征进行分析,并且从其中提取视觉表征来实现为用户检索图像的服务^[7]。这种技术的目的是为了攻克其他检索方式的限定,从而更加便捷的帮助个人在巨量的信息资源中更加精准地检索到有用信息。

网络图像在互联网信息资源中极其普遍,而且每一个网络图像都有其特定的资源位置。为了节约储存空间,网络图像通常都会压缩以后再进行储存,3种最常见的格式是 gif、jpg 和 png。网络图像的其他存储格式为 swf、svg、bmp、tiff 等,种种不同的网络图像构成了存储量巨大的网络图像数据库^[8]。因为网络图像复杂且多样,这就导致了网络图像检索技术和其他图像检索技术的差异。网络图像检索通常具有更高的准确率和速度,以此来满足广大用户的各种需求。目前,各大企业、大学、研究机构、杂志、国际组织已建立了相关项目组、专题小组、专栏和国际会议组织,用于交流网络图像检索领域的最新研究成果。

3 图像检索技术比较

3.1 3种图像检索技术的优缺点

3.1.1 基于文本的图像检索技术的优缺点

TBIR 具有以下优点:

- (1) TBIR 在操作上比较方便快捷,对技术条件的依赖性较小,索引和存贮比较简单。
- (2) TBIR 是现今使用最为普遍的检索技术,各大门户网站都在使用这种检索技术。
- (3) 比较切合人们的搜索惯性,一般的用户在检索信息时基本都是使用关键字作为基础,很少有用户会对其他元素进行释义。
- (4) 由于采用的是人工标引方式,手动描述完成就可以进行索引和存储,所以技术上很容易实现。

TBIR 也有以下缺点:

- (1) TBIR 是用于在不直接使用图像内容的情况下为每个图像文件创建一个关键字或文本标题以及一些附加信息作为图像索引,这种方式还是以手动的方式为主,耗时耗力并且主观性过强。
- (2) 一些功能特征可以容易地描述为图像视觉本身,但一些功能特征较难实现。目前对于图像特征的标注还是基本依靠人工完成,因为标签的图像必须依赖于劳动力,计算机视觉技术和 AI 技术目前还不具备自动标记图像的能力。所以这种方法既耗时又不准确,不完整;具有主观性的图像难以描述,图像信息如颜色、纹理等视觉特征不能用于准确描述文字。

3.1.2 基于图像的图像检索技术的优缺点

CBIR 具有以下优点:

- (1) 自动提取信息资源,不需要进行相关释义。
- (2) 可以通过计算机进行自动描述,进一步提高了检索的准确率以及速率。
- (3) 用户可以凭借自己的主观意识来引导整个检索流程,使检索效率进一步提高。
- (4) 可以从较大的图像库以更加便捷的方式找到想要的图像,重点是基于内容,快速查找信息。

CBIR 也有以下缺点:

- (1) 由于人们的主观意识不同,即便是相同的特征,不同的人的理解上也不尽相同。
- (2) 在图像检索中,检索方式通常是提取图像的底层表征,然后进行比较。大部分方法都难以找到表征与注释的结合点。
- (3) 大部分图像检索技术都具有不同的检索方式,很难找到与用户要求最为一致的综合相似方法。

3.1.3 网络上的图像检索技术的优缺点

网络上的图像检索技术具有以下优点:

- (1) 通过计算机视觉搜索系统来上传信息资源, 搜索范围最为广泛。
- (2) 可将手机拍照系统与计算机视觉搜索系统相结合, 让用户能随时随地进行搜索, 比较便携且具有普适性。
- (3) 有目的地把计算机视觉搜索系统安装到嵌入式系统或者穿戴设备中, 针对特殊用户群体有十分重要的价值。

网络上的图像检索技术也有以下缺点:

- (1) 计算机视觉搜索系统虽然搜索范围最广泛, 但是受条件限制, 不方便携带。
- (2) 将手机拍照与视觉搜索系统相结合虽然比较便携, 但是它在人机交互这一方面还需要进一步的研究。

3.2 3 种图像检索技术的比较分析

3 种图像检索技术既有优点也有缺点, 下面列表比较 3 种技术。

表 1 3 种图像检索技术比较分析表

图像检索技术	主要用户	操作难度	优 点	缺 点
基于文本的图像检索技术	适用于大多数用户	较低	(1) 操作便捷; (2) 可供检索系统、资源较多; (3) 比较符合搜索习惯; (4) 存储更容易。	(1) 基本靠手动进行检索, 耗时耗力; (2) 主观性较强, 难以得到准确信息。
基于图像的图像检索技术	用户受众较少	较高	(1) 自动提取图像特征; (2) 检索的准确率及速率较高; (3) 可凭借主观意识引导检索过程。	(1) 底层表征与高层语义难以结合; (2) 难以做到一致的综合同相似方法。
网络上的图像检索技术	用户受众较多	较低	(1) 搜索范围最为广泛; (2) 具有便携与普适性; (3) 针对特殊用户群体具有重要价值。	(1) 交互性仍需进一步研究; (2) 计算机视觉搜索虽搜索范围广泛, 但不方便携带。

4 结语

总体来说, TBIR 在一定程度上, 避免了复杂的识别问题, 符合人们的搜索习惯, 操作难度较低, 绝大多数用户都适合使用。但是没有一致的标准, 由于搜索过于主观, 检索比较复杂的图像时准确率没有保证。CBIR 则不同, 这种方法是通过提取图像的表征来分析每个图像, 但是 CBIR 算法过于复杂并且成本偏高, 导致用户受众面较窄, 难以构建一致的综合相似方法, 搜索的准确性偏低。网络上的图像检索技术具有最为广泛的搜索范围, 但是利用计算机视觉搜索系统不方便携带, 与手机拍照的交互性仍需要进一步的研究。

注释

- [1]赵倩. 基于内容的图像检索若干技术研究[D]. 上海: 上海大学, 2012: 16-17.
- [2]阿斯艳·哈密提, 阿不都热西提·哈密提·阿斯艳. 基于文本的图像检索与基于内容的图像检索技术的比较研究[J]. 首都师范大学学报(自然科学版), 2012(4): 6-9.
- [3]张骞. 基于文本的与基于内容的图像检索技术比较研究[J]. 情报探索, 2012(1): 111-113.
- [4]刘宪龙. 基于边缘方向差特征的图像检索方法[D]. 西安: 西安电子科技大学, 2014: 35-36.
- [5]A Multi-scale and Multi-orientation Image Retrieval Method based on Rotation-invariant Texture Features [J]. Science China (Information Sciences), 2011(4): 732-744.
- [6]董文菲. 基于多特征融合和 SVM 分类的图像检索技术研究[D]. 哈尔滨: 哈尔滨理工大学, 2015: 19-20.
- [7]杨文涛, 司应硕. 微粒群算法在图像检索中的应用[J]. 华北水利水电学院学报, 2011(2): 90-92.
- [8]张旭. 网络图像检索关键技术研究[D]. 西安: 西安电子科技大学, 2014: 14-15.

郑雪丽 郑州航空工业管理学院信息科学学院; 航空经济发展河南省协同创新中心。研究方向: 网络信息资源管理与服务。
徐严斌 郑州航空工业管理学院信息科学学院。