反向图像搜索引擎

姓名： 余恒

学号： 120106022625

## 目录

**[1.研究背景、意义和研究现状 1](#_Toc24511)**

[1.1研究背景 1](#_Toc28258)

[1.2研究意义 1](#_Toc14243)

[1.3研究现状 1](#_Toc13330)

**[2.国内外识图应用分析及对比 2](#_Toc20258)**

[2.1识图应用场景分析 2](#_Toc16910)

[2.2国内外研究水平和发展现状 2](#_Toc32097)

**[3.反向图片搜索引擎系统设计 4](#_Toc11729)**

**[4.反向图片搜索引擎系统关键技术分析 5](#_Toc22840)**

4.1生成高斯差分金字塔（DOG金字塔），尺度空间构建 5

4.2空间极值点检测（关键点的初步查探） 5

4.3稳定关键点的精确定位 6

4.4稳定关键点方向信息分配 6

4.5关键点描述 7

4.6特征点匹配 7

4.7发展水平分析 8

**[5.个人感想 9](#_Toc13583)**

**1.研究背景、意义和研究现状**

**1.1研究背景**

图像检索问题的研究最早起于上世纪70年代末，当时的研究主要集中与基于文本的图像检索。上世纪90年代初，有学者提出了基于内容的图像检索（CBIR）思想，它使用图像的颜色、形状等信息作为特征构建索引以实现图像检索，即我们通常所说的“以图搜图”。随后又有人提出了基于语义的图像检索思想。和互联网的出现时间相比，图像搜索引擎出现的时间要晚的多，由于计算机的性能与硬件等条件约束，其发展也比较缓慢。如今，随着这些问题的解决，以及人们对图像检索技术需求的日益增加，相关技术的研究取得了极大的进步，诸如深度学习，卷积神经网络等AI技术在图像搜索领域大显身手。图像搜索引擎现在已经成为互联网上非常重要的一部分，它为人们提供互联网的导航服务，满足人们的搜索需求。即使现在已经有了成熟的图像搜索引擎。但它仍然有无穷的潜力，依然是计算机学术界和工业界争相研究、开发的对象。

**1.2研究意义**

搜索引擎的出现，为人们搜索图像提供了更加方便的工具，意义重大。尽管随着多媒体技术的发展，互联网中的声音和视频、文字、图像等信息越来越多，但不可否认文字和图像信息依然是互联网信息的主体，随着社交网络的发展，图像，尤其是影像信息更是越来越重要，所以研究图像搜索引擎对于促进互联技术的发展有着极其重要的意义。随着AI的不断发展，先是涌现了诸如相机人脸检测、安防人脸识别、车牌识别等等如今随处可见的技术，此外，有着深度学习的加持，图像搜索引擎在临床病理诊断方面也有着不可限量的应用前景，例如对早期癌症的诊断，这能极大的帮助医生判断病人的状况，及时采取治疗措施。

**1.3研究现状**

国外的互联网起步发展较早，其在图像搜索引擎技术道路上的探索也比较久远，随着CBIR的提出，IBM开发了第一个商用的CBIR系统QBIC（Query By Image Content），用户只需输入一幅草图或图像，便可以搜索出相似的图像。同一时期，很多公司也将这一技术引入搜索引擎。哥伦比亚大学开发的WebSEEK系统不仅提供了基于关键词的图像搜索和按照图像类目的主题浏览，还可以利用图像的颜色信息进行基于内容的图像搜索。Yahoo的ImageSurfer也提供了使用例图的颜色、形状、纹理特征，以及它们的组合来进行基于内容的图像搜索功能。随着视觉技术的进步和发展，越来越多的搜索引擎采用这一方式来进行图像搜索，并在此基础上不断演进。结合国内的发展来看，近年来这一技术已取得长足发展，对比国外而言，国内技术的基础往往设定为图像形状、色彩特征等方面。就本国而言，相关的检索系统和检索软件研发工作，大多由知名高校和相应的研究团队共同完成，早期有如IngRetr、MIRES等的软件，如今随着阿里、百度等公司的崛起，诸如淘宝拍立淘、百度识图等广为群众使用的图像搜索引擎技术也日益增多。像这样有利于优化用户体验、促进科学研究的技术，其前景依然一片光明。

**2.国内外识图应用分析及对比**

**2.1识图应用场景分析**

用户拿一张图片去搜索获取想要知道的人或物的图像搜索系统，被称为基于内容的图像检索，亦称作反向图像搜索，是如今图片搜索引擎应用最为广泛的场景之一。如用户通过一张商品照片在购物网站寻找购买链接、通过一张风景照片寻找具体地点、通过一张人物图像判断其是哪位名人等等。

反向图像搜索引擎的具体实现可以简化为下列步骤：

1. 训练。成千上万的图片被打上标签，成为神经网络的训练样本。神经网络将学习并对它们分类。
2. 输入。未标签的图片放入神经网络进行预训练。
3. 第一层。神经元对简单形状进行提取，比如边缘部分。
4. 高层。神经元继续提取更复杂的结构。
5. 顶层。神经元再把高度复杂、抽象的概念提取出来。
6. 输出。根据其训练出来的能力，神经网络给出与图片最相似的结果。

通过深度神经网络，各类视觉识别的任务精度都得到了大幅提升。在全球最权威的计算机视觉竞赛ILSVR上，千类物体识别错误率在2011年时还高达25.8%，从2012年引入深度学习之后，后续4年的错误率分别达到了16.4%、11.7%、6.7%、3.7%，出现了显著突破。现在，人脸识别甚至能做到误判率低于百万分之一。

**2.2国内外研究水平和发展现状**

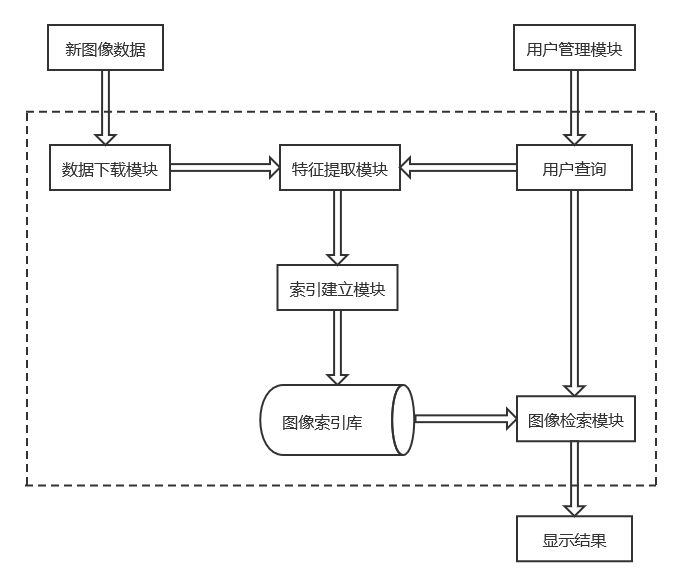
1999年由加拿大Idée公司研发出世界上第一款相似图片搜索引擎，这标志着以图搜图方式的“反向图片搜索引擎”(Reverse Image Search Engine)时代的到来。Google公司的谷歌图片，于2009年12月推出，它让人们可以为书、CD或图片照相，然后立即放到网上搜索出图片所指的产品或地点等相关资料。很多公司之后也都效仿Google公司的模式相继开发属于自己的图像搜索引擎。例如Picitup,该引擎的最大特点是在提供相似图片的同时，还支持语义环境搜索。这样就可以在得到相似图片后利用关键字找到原图片或是相似图片。

德国研究基金会开发的检索系统。其根本任务在于图像检索再医学方面的交叉和使用，而特点就在于它的对象涉及了包括医学和生物学在领域内的特征。值得称道的是，该系统还通过加入动态虚拟生物特征，完成了在十几万副普通图像和几万副医学图像中的研究和试验。

国内的百度识图，继续沿袭百度公司在在图像识别的优势，能够有效的利用网络技术。用户只需提交待检图片，则系统会自动对图片进行特征提取、特征分析、特征比对，从而由庞大的图像库和互联网中搜索相似性图片。早期百度公司利用语义环境检索图片任然继续发挥着优势，返回结果中也包含图片的语义信息，方便用户进行下一步的检索。但是，该系统对图像的质量要求严格，为了降低检索时间，图片的大小必须在5M以内。格式也相对固定，但大多是人们经常使用的格式如JPEG、BMP、GIF、PNG 等。

**3.反向图片搜索引擎系统设计**

基于内容的图像检索技术是以图像为基础的信息检索时代的关键技术。CBIR检索技术的查询条件来自于图像本身，在提取图像特征时，图像的特征即为查询的条件。在获得查询条件后，通过比较其特征与图像库中的特征来决定其和图像库中图像的相似度。基于图像内容的反向图片搜索引擎系统如下图所示。



整个系统可以分为以下几步：一是下载新的原始的图片数据；二是对图像进行解析获得图像的基本属性和内容特征；三是将相关文字特征和颜色特征建立索引添加到图像特征库中；四是用户上传或查询，最后显示查询结果。

其中：数据下载模块主要是系统获取新的图像资源；在系统提取图像方面，以特征提取模块作为基本，这也是所有图像搜索引擎最为重要的一个步骤；而索引建立模块的存在，能够在特征提取的基础之上，进行索引的构建，并添加到系统的图像索引库中，可以提高查询速度；图像检索模块是把用户输入的内容跟系统库中的图像信息做对比，最后显示查询结果；用户管理模块是用户注册或登入系统，包含用户上传和用户查询，用户上传是把用户上传的图像提取特征追加索引并添加到系统图像索引库，而用户查询是通过提取用户输入的样例图特征到图像检索模块跟图像索引库的图像进行匹配，最后输出匹配结果。

**4.反向图片搜索引擎系统关键技术分析**

对于相同物体图像检索，在检索相同的物体或目标时，易受拍摄环境的影响，比如光照变化、尺度变化、视角变化、遮挡以及背景的杂乱等都会对检索结果造成较大的影响。

### 由于受环境干扰比较大，因而对于相同物体图像检索，在选取特征的时候，往往会选择那些抗干扰性比较好的不变性局部特征，其中具有代表性的就是由David G. Lowe教授提出的高效区域检测算法SIFT（Scale Invariant Feature Transform，尺度不变特征变换匹配算法）。SIFT算法实质在不同的尺度空间上查找关键点，并计算出关键点的方向，其有以下的优点：

### 在多尺度空间采用DOG算子检测关键点，相比传统的基于LOG算子的检测方法，运算速度大大加快。

### 关键点的精确定位不仅提高了精度，而且大大提高了关键点的稳定性。

### 在构造描述子时，以子区域的统计特性，而不是以单个像素作为研究对象，提高了对图像局部变形的适应能力。

### 未命名文件(2)

SIFT特征的生成一般包括以下几个步骤：

### 1. 生成高斯差分金字塔（DOG金字塔），尺度空间构建

### 2. 空间极值点检测（关键点的初步查探）

### 3. 稳定关键点的精确定位

### 4. 稳定关键点方向信息分配

### 5. 关键点描述

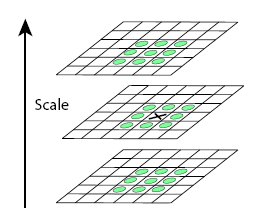
### 6. 特征点匹配

**4.1生成高斯差分金字塔（DOG金字塔），尺度空间构建**

此步骤的主要思想是通过对原始图像进行尺度变换，获得图像多尺度下的尺度空间表示序列，对这些序列进行尺度空间主轮廓的提取，并以该主轮廓作为一种特征向量，实现边缘、角点检测不同分辨率上的关键点提取等。各尺度下图像的模糊度逐渐变大，能够模拟人在距离目标由近到远时目标物体在视网膜上的形成过程。

**4.2空间极值点检测（关键点的初步查探）**

为了寻找DOG函数的极值点，每一个像素点要和它所有的相邻点比较，看其是否比它的图像域和尺度空间域的相邻点大或者小，如下图所示：



在二维图像空间，中心点与它3\*3邻域内的8个点做比较，在同一组内的尺度空间上，中心点和上下相邻的两层图像的2\*9个点作比较，如此可以保证检测到的关键点在尺度空间和二维图像空间上都是局部极值点。

**4.3稳定关键点的精确定位**

DOG值对噪声和边缘比较敏感，所以在第2步的尺度空间中检测到的局部极值点还要经过进一步的筛选，去除不稳定和错误检测出的极值点，另一点就是在构建高斯金字塔过程中采用了下采样的图像，在下采样图像中提取的极值点对应在原始图像中的确切位置，也是要在本步骤中解决的问题。

**4.4稳定关键点方向信息分配**

稳定的极值点是在不同尺度空间下提取的，这保证了关键点的尺度不变性。为关键点分配方向信息所要解决的问题是使得关键点对图像角度和旋转具有不变性。方向的分配是通过求每个极值点的梯度来实现的。

对于任一关键点，其梯度幅值表述为：

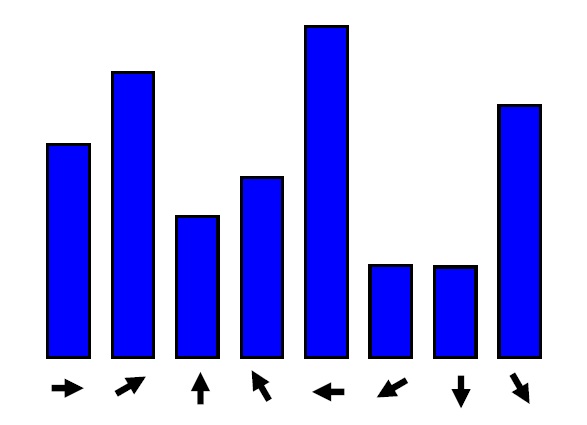


梯度方向为：



分配给关键点的方向并不直接是关键点的梯度方向，而是按照一种梯度方向直方图的方式给出的。

具体的方法是：计算以关键点为中心的邻域内所有点的梯度方向，当然梯度方向一定是在0~360°范围内，对这些梯度方向归一化到36个方向内，每个方向代表了10°的范围。然后累计落到每个方向内的关键点个数，以此生成梯度方向直方图。



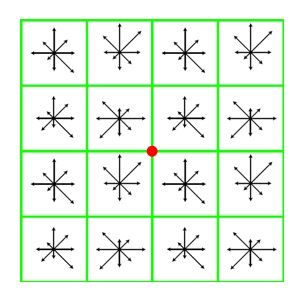
将梯度方向直方图中纵坐标最大的项代表的方向分配给当前关键点作为主方向，若在梯度直方图中存在一个相当于主峰值80%能量的峰值，则将这个方向认为是关键点的辅方向。辅方向的设计可以增强匹配的鲁棒性，本算法的作者指出，大概有15%的关键点具有辅方向，而恰恰是这15%的关键点对稳定匹配起到关键作用。

**4.5关键点描述**

对关键点的描述是后续实现匹配的关键步骤，描述其实就是一种以数学方式定义关键的过程。描述子不但包含关键点，也包括关键点周围对其有贡献的邻域点。

描述的思路是：对关键点周围像素区域分块，计算快内梯度直方图，生成具有独特性的向量，这个向量是该区域图像信息的一种抽象表述。

对每个关键点，采用4\*4\*8共128维向量的描述子进项关键点表征，综合效果最佳：



**4.6特征点匹配**

特征点的匹配是通过计算两组特征点的128维的关键点的欧式距离实现的。欧式距离越小，则相似度越高，当欧式距离小于设定的阈值时，可以判定为匹配成功。

**4.7发展水平分析**

虽然以SIFT算法为基本衍生的诸多算法是如今比较流行的策略，但这此类方法通常特征维度往往很高，因此为它们设计高效的索引方式显得十分关键。近年来，以深度学习为主流的自动特征在应用到相同类别图像检索上时，能够极大的提高检索的精度，使得面向相同物体的检索在特征表达方面得到了较好的解决。目前，以卷积神经网络为主导的特征表达方式也开始在相同物体图像检索上进行展开，并已有了一些相应的工作，但由于相同物体在构造类样本训练数据时并不像相同类别图像检索那样那么方便，因而相同物体图像检索在CNN模型训练以及抽取自动特征等方面还有待深入。

**5.个人感想**

在AI这个充满话题性的领域，其在每一个地方的新应用都能引来人们的称赞和批评。事关到图像，尤其是关于人的图像资料，往往会造成个人隐私的争议，这不应是AI这项技术应该承担的批评，而是那些不往好的方向利用AI的人的责任，公众舆论与逐渐认识到了这一点，所以AI技术发展依然是大势所趋。

在互联网飞速应用的今天，传统的图像数据库已经不能满足用户的需求。如何能应用互联网上的图像资源是促进诸多领域进步的需求，而这也将会给基于图像内容的反向图片搜索引擎提供一个良好的展空间。如此这项技术便可以快速、大量地查找网上的图像信息，为图像检索提供一条方便、节省的途径。

人们都认为AI可能带来一些岗位的丢失，但是其蓬勃发展带给人类的福祉更是不可估量的。例如，临床基于已有图像库的反向图像搜索技术。Airdoc是一家人工智能医学影像识别技术开发商，其基于深度学习开发的算法，通过结节定位、结节大小判断、结节恶性指标计算快速检测肺结节。其在眼病、皮肤病、[心血管疾病](http://study.cn-healthcare.com/collegewap/324)和大脑等领域也都有比较成熟的模型，诊断正确率高达97%以上。

今后在此领域，应当进一步研究高维彩色特征向量的特征提取、降维、聚类、索引以及相似性度量和检索等问题,并综合使用多特征对图像进行相似性度量和检索,进一步利用基于人机交互的相关反馈技术来提高技术的精确度。随着各项新技术的不断涌现和发展,一定会使人们在面对各种各样的图像数据库时,都可以获得高效快捷和极高准确率的检索结果。通过所有研究者的共同努力,不断更新和发展的技术定会日益深入地改变人们的日常生活,这将会使人们的生活和工作更加丰富多彩,更加充满活力。