

毕业设计（论文）

题目名称：跨媒体信息检索系统的设计与实现

——图片检索

院系名称：软件学院

班 级：RB软工卓越171

学 号：201719044105

学生姓名：于锦江

指导教师：郭丽

2021 年 6 月

**跨媒体信息检索系统的设计与实现**

**-图片检索**

**Design and Implementation of Cross-media Information Retrieval System**

**- Picture retrieval**

学院名称：软件学院

班 级：RB卓越171

学 号：201719044105

学生姓名：于锦江

指导教师：郭丽

2021年6月

**摘 要**

搜索引擎不断的发展，人们越来越开始使用搜索引擎来获取信息。几乎每个人在日常生活中都离不开搜索引擎，人们可以通过搜索引擎来获取大量的更加全面的详细的信息。搜索引擎服务与生活、工作、休闲等各个方面。但是目前市面上的搜索引擎主要是文字检索，随着人们的需求不断提高。以文字作为搜索条件的传统引擎的不足也日益暴露出来，譬如当用户想要搜索一段视频的背景音乐，或者想根据一张图片搜索相似图片的时候，用户很难描述他们想要搜索的关键字，这也导致了搜索结果的不准确性。正是由于这种强烈的用户需求，跨媒体搜索应运而生。

本系统为了使搜索信息的方式更加多样化、人性化。用户可以通过文字、图片来检索想要获得的信息，这些信息的类型包括文字、音频、视频、和图片，在传统搜索引擎的功能上丰富了检索内容。提高用户的检索效率。即通过上传一张图片，即可检索到与该图片相关的所有图片。通过文字搜索即可得到与其相关的文本、音频、视频等内容。

图片检索的主要思路为通过VGG16网络提取数据集中图片的特征向量，将待检索图片的特征向量与数据库中特征向量进行距离计算，得到距离相近的一组，即为搜索结果。

**关键词**：跨媒体；信息检索；特征向量；向量距离计算；

**Abstract**

With the continuous development of search engines today, people are increasingly using search engines to obtain information. Almost everyone is inseparable from search engines in their daily lives. People can obtain a lot of more comprehensive and detailed information through search engines. Search engine services and life, work, leisure and other aspects. However, the current search engines on the market are mainly text search, and people's needs continue to increase. The shortcomings of traditional engines that use text as search criteria are also increasingly exposed. For example, when users want to search for background music in a video or search for similar pictures based on a picture, it is difficult for users to describe the keywords they want to search. , Which also led to the inaccuracy of search results. It is precisely because of this strong user demand that cross-media search came into being.

This system is to make the way of searching information more diversified and humanized. Users can retrieve the information they want through text and pictures. These types of information include text, audio, video, and pictures, which enrich the retrieval content in the functions of traditional search engines. Improve user retrieval efficiency. That is, by uploading a picture, all pictures related to the picture can be retrieved. You can get related text, audio, video and other content through text search.

The main idea of ​​image retrieval is to extract the feature vector of the image in the data set through the VGG16 network, and calculate the distance between the feature vector of the image to be retrieved and the feature vector in the database to obtain a set of similar distances, which is the search result.

**Keywords:** Cross-media**；**Information retrieval；Feature vector；Vector distance calculation；

目录

[第1章 项目简介 1](#_Toc2016)

[1.1 项目背景 1](#_Toc30434)

[1.2 研究状况调研情况 2](#_Toc22476)

[1.3 项目目标 3](#_Toc12721)

[第2章 系统需求分析 4](#_Toc23239)

[2.1 系统概述 4](#_Toc29987)

[2.2 系统用户 4](#_Toc30564)

[2.3 用例分析 5](#_Toc1559)

[2.4 需求分析-管理员 7](#_Toc15694)

[2.5 需求分析-游客 10](#_Toc17064)

[第3章 系统分析 11](#_Toc29389)

[第4章 系统设计 12](#_Toc7729)

[4.1 架构设计 12](#_Toc13117)

[4.2 工作流程设计 12](#_Toc3682)

[4.2 设计类 13](#_Toc20037)

[4.3 数据库设计 13](#_Toc31255)

[第5章 系统实现 14](#_Toc21979)

[第6章 系统测试 15](#_Toc32723)

[第7章 结束语 16](#_Toc31061)

[致谢 17](#_Toc22784)

[参考文献 18](#_Toc32281)

# 第1章 项目简介

本章的主要内容是“跨媒体搜索引擎”的项目背景、现实意义与实现目标的介绍，这一部分的主要内容是通过分析当前市场现存的相似产品，比如谷歌、百度等。通过对比进而来阐述本系统在当前是时代背景下能解决的问题。

## 1.1 项目背景

对于信息检索来说，图像对信息的表达更为直观且包含更多更丰富的意义，是人们获取信息不可或缺的重要来源，然而图片在计算机中的识别过程，还远未达到人脑在图片识别过程中能够直接理解其中语义信息的能力，基于语义的图像检索是研究的高级目标。

图像检索，又称图像检索 ，系统是一个电脑浏览的系统，从一个大型的数字图像数据库去检索和检索图像。大多传统和一般图像检索的方式是利用一些增加元数据(metadata)的方法，例如：字幕、关键词或是图像的说明，如此一来就可以透过注解词完成检索。人工的图像注解是费时、费力并且昂贵；为了解决这个问题，已经有大量的研究在做自动图像注解方面上。此外，越来越多的社会网络应用和语义网已经产生了数个以网络为基底发展的图像注解工具。

第一个以微电脑为基底开发的图像数据库检索系统，是由80年代麻省理工学院里的Banireddy Prasad、Amar Gupta、Hoo-min Toong、and Stuart Madnick所共开发出来的。这是记载于1987年2月发行的IEEE Transactions on Industrial Electronics。另一种方法的图像检索是以内容为基底的图像检索（Content-based Image Retrieval，简称CBIR）技术，其目的是为了避免使用文字上的描述而是以视觉相似性为基底透过用户提供查询的图像或是用户指定的图像特征来检索图像。

从20世纪70年代开始，有关图像检索的研究就已开始，当时主要是基于文本的图像检索技术（Text-based Image Retrieval，简称TBIR），利用文本描述的方式描述图像的特征，如绘画作品的作者、年代、流派、尺寸等。到90年代以后，出现了对图像的内容语义，如图像的颜色、纹理、布局等进行分析和检索的图像检索技术，即基于内容的图像检索CBIR技术。CBIR属于基于内容检索（Content-based Retrieval，简称CBR）的一种，CBR中还包括对动态视频、音频等其它形式多媒体信息的检索技术。

考虑基于文本的检索和基于内容的检索的优势和劣势，尽可能对图片本身及其环绕文本所蕴藏信息的利用，以及对分布式并行计算云平台并行图片检索的应用，达到准确快速对图片进行检索的目的。

在检索原理上，无论是基于文本的图像检索还是基于内容的图像检索，主要包括三方面：一方面对用户需求的分析和转化，形成可以检索索引数据库的提问；另一方面，收集和加工图像资源，提取特征，分析并进行标引，建立图像的索引数据库；最后一方面是根据相似度算法，计算用户提问与索引数据库中记录的相似度大小，提取出满足阈值的记录作为结果，按照相似度降序的方式输出。

为了进一步提高检索的准确性，许多系统结合相关反馈技术来收集用户对检索结果的反馈信息，这在CBIR中显得更为突出，因为CBIR实现的是逐步求精的图像检索过程，在同一次检索过程中需要不断地与用户进行交互。

图像检索的发展是一个从简单到复杂、从低级到高级的过程，从最初的文本信息查询发展到基于内容的图像检索。同时随着人们对图像理解、图像识别研究的不断深入，提出了基于图像语义的检索,充分利用了图像的语义信息, 提高了图像检索系统的能力。另外，为了解决语义鸿沟的问题，人们提出了基于反馈的信息检索技术，利用人机交互行为，改进系统的能力，提高检索结果的准确性。最后，随着人工智能和信息技术的发展，一种智能的基于知识的信息检索系统成为信息检索领域的发展方向。基于知识的信息检索技术将基于视觉特征和基于文本语义的技术结合在一起，通过建立知识库，实现自动提取语义和图像特征的功能，并且充分考虑到用户特征对检索系统的影响, 这是建立高效、实用、快速的图像检索系统的必然的发展方向。并且图像检索领域的关键问题是对人类视觉机制的进一步了解,即探求人是如何去感知图像内容的, 这个问题的解决能够进一步优化数据特征索引技术, 解决对大规模数据库检索速度的瓶颈问题。随着多媒体数据压缩技术和互联网的迅速发展，信息的形式多种多样,视觉信息数据不仅包括单幅的图像数据还包括视频数据，针对视频数据的特点，进行高速、可靠的检索也是一个需要研究的课题。将信息检索技术推向实用化, 也是信息技术发展的主要目标。

## 1.2 研究状况调研情况

到目前为止，已经推出的较为成熟的跨媒体搜索的产品有“百度识图”、“Google 识图”、“微信搜歌”等，它们都是将多媒体作为自己的搜索输入条件，来搜索出用户想 要的搜索结果，而且准确率已经达到了十分可观的程度。由于跨媒体搜索的语义解析大致可以分为“文字语义解析”、“图片语义解析”、 “音频语义解析”、“视频语义解析”四个方面，现在分别介绍：

（1）文字语义解析：文字语义的解析经过多年的发展与运用，已经十分成熟，很多文字解析的算法已经非常成熟。

（2）图片语义解析：现在图片的语义解析的算法研究已经十分成熟，相关的识图软件诸如“百度识图”、“Google 识图”也已经推出。

（3）音频语义解析：现在音频语义的解析相关的引擎已经较为成熟，相关的识别软件诸如“微信搜歌”也已经推出。

（4）视频语义解析：现在视频语义的解析算法研究还不是十分成熟，要推出相应的产品可能还需要经过一段时间。

总体而言，现在单个文件格式的解析算法的研究已经趋于成熟，但如何将各种算法集合到一起，实现真正的跨媒体搜索，到目前为止还没有一个成熟的方案。因此，本文提出了将文字解析、图片解析、音频解析、视频解析集于一体，并且考虑到各种算法的差异性与多样性，提出了一个可扩展的、可以自由配置解析算法的跨媒体搜索平台。

## 1.3 项目目标

本系统的目标是实现对音频，视频，图片，文本这三类文件资源，基于语义内容进行分析以实现对跨媒体资源的检索。主要包括一下三点：

一、对上传的音视频文件基于音频内容的语音识别并存储分析数据到Mysql，定时同步到搭建好的ElasticSearch搜索引擎框架上。

二、对通过爬虫爬去的文本类型的网页内容并存储在Mysql，定时到同步到ElasticSearch搜索引擎。

三、使用前端搜索页面进行搜索数据是不仅对匹配的关键词进行返回，也会返回一些根据搜索关键词进行词向量计算，返回高度相似的搜索关键词的匹配结果。

# 第2章 系统需求分析

本章的主要内容是“跨媒体搜索引擎”的图片检索的功能设计。通过分析系图片检索的功能，以用户的角度去设计UI设计图、用例规约和操作方法等。

## 2.1 系统概述

搜索引擎主要功能是根据用户输入的信息检索与之相符合的内容，用户的输入类型有文本和图片两种，搜索结果的类型有文本、音频、视频和图片。当用户想要查询一张图片的信息时只需要通过本系统将图片上传至云端，云端通过TensorFlow分析图片信息后，通过一系列的算法计算，即可找到数据集中与之相符合的结果集。系统功能总览图与图2-1所示。

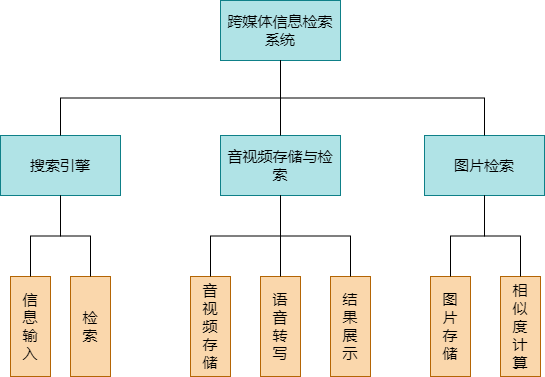


图2-1 系统功能总览

## 2.2 系统用户

### 2.2.1系统管理员

系统管理员主要负责系统中数据集的录入和管理等功能，通过输入管理员账号密码后，即可进入后台管理页面，对系统中的数据进行修改和维护。主要功能有音视频上传、图片上传、文本上传等。

### 2.2.2普通游客

用户无需登陆，直接在门户的搜索框中填写信息，即可使用本系统的检索功能。

## 2.3 用例分析

跨媒体信息检索系统的管理员端的主要功能为对系统内的资源进行管理，包括文件上传等操作。首先管理员通过本系统进行注册，注册完成后即可进入管理后台，对系统内资源进行管理和配置，管理员的用例图如图2-2所示。其详细用例信息如下表2-1所示。

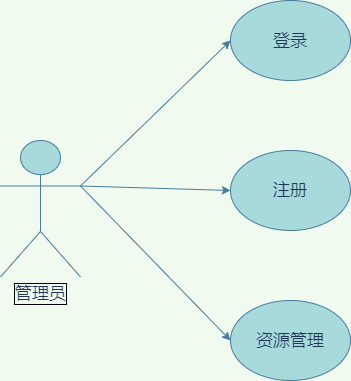


图2-2 管理员用例图

表2-1 管理员用例表

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 用户 | 功能类型 | 功能模块 | 子功能描述 | 优先级 |
| 管理员 | 用户 | 登录 | 管理员登入系统 | 高 |
| 注册 | 管理员注册账号 | 高 |
| 资源 | 文件上传 | 视频上传 | 高 |
| 音频上传 | 高 |

跨媒体信息检索系统的普通游客主要是在门户网站上通过录入信息，即可获取到与之相关的信息，比如说音频、视频、文本、和图片类型数据的信息。搜索时将用户输入的关键词进行检索，并且会将关键词相似的词语作为检索对象，与关键词一同进行检索，最后根据检索计算的得分来进行排序。显示在页面上，用例图如图2-3所示。其详细用例信息如下表2-2所示。

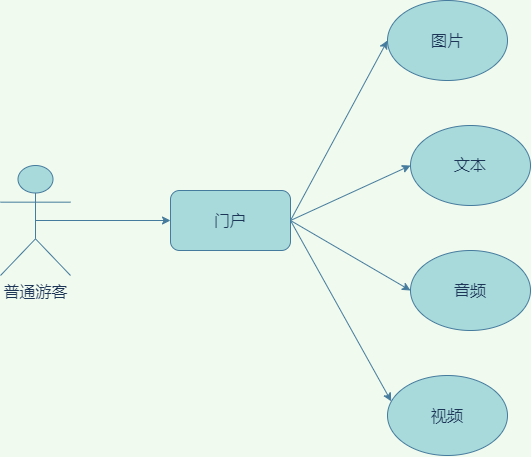


图2-3 游客用例图

表2-2 游客用例表

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 用户 | 功能类型 | 功能模块 | 子功能描述 | 优先级 |
| 游客 | 检索 | 音频检索 | 将用户输入的关键词有关的音频进行检索 | 高 |
| 视频检索 | 将用户输入的关键词有关的视频进行检索 | 高 |
| 文本检索 | 将用户输入的关键词有关的文本进行检索 | 高 |
| 图片检索 | 对用户上传的图片进行计算返回相似图片集 | 高 |

## 2.4 需求分析-管理员

### 2.4.1 注册

注册是管理员使用管理系统的先觉条件，注册之后方可以进行登录操作，使用管理员端具体功能。用例分析表如表2-3所示。

表2-3 注册用例分析表

|  |  |  |  |
| --- | --- | --- | --- |
| 功能名称 | 注册 | | |
| 操作用户 | 管理员 | 用例描述 | 通过注册管理员账号进而登录 |
| 前置条件 | 尚未拥有管理员账号  进入登录页面  点击注册按钮 | | |
| 基本事件流 | 输入账号密码  点击注册按钮 | | |
| 异常事件流 | 注册出错，提示"注册失败，请重试！" | | |
| 后置条件 | 点击注册按钮，成功后三秒后跳转到登录页面 | | |
| 补充说明 |  | | |
| 界面原型 | 如图2-4所示 | | |

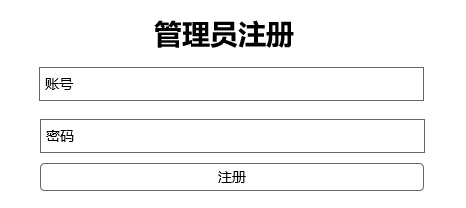


图2-4 注册界面原型图

### 2.4.2 登录

当用户完成管理员注册之后，进入登陆页面根据注册时填写的账号密码来进行登录，登录成功后进行跳转至后台管理页面。用例分析表如表2-4所示。

表2-4 登录用例分析表

|  |  |  |  |
| --- | --- | --- | --- |
| 功能名称 | 登录 | | |
| 操作用户 | 管理员 | 用例描述 | 管理员通过账号密码来进行登录 |
| 前置条件 | 已注册管理员账号  进入登录页面 | | |
| 基本事件流 | 输入账号密码  点击登录按钮 | | |
| 异常事件流 | 账号密码错误，提示"账号或密码错误！" | | |
| 后置条件 | 点击登录按钮，成功后跳转到管理员主页 | | |
| 补充说明 | 登录时会将用户密码加密传输确保安全性 | | |
| 界面原型 | 如图2-5所示 | | |

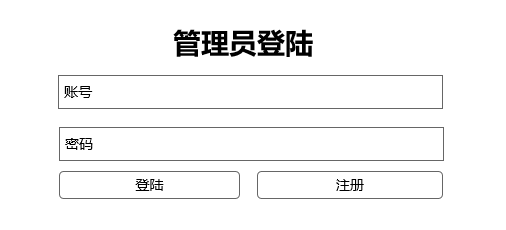


图2-5 登录界面原型图

### 2.4.3 文件上传

管理员登录后进入主页，可以上传文件至系统资源库中，文件上传分为视频文件，音频文件两种类型，点击选择所上传的文件，稍作等待上传结果。用例分析表如表2-5、2-6所示。

表2-5 上传音频用例分析表

|  |  |  |  |
| --- | --- | --- | --- |
| 功能名称 | 上传音频 | | |
| 操作用户 | 管理员 | 用例描述 | 管理员上传所需转写的音频文件 |
| 前置条件 | 管理员登录  进入文件上传页面 | | |
| 基本事件流 | 选择需要上传的wav文件 | | |
| 异常事件流 | 选择非规定的文件类型，提示“上传只能是 WAV或Mp4 格式！”  文件大小超出限制，提示“上传文件大小不能超过 500MB！”  文件内容出错，提示“上传失败！” | | |
| 后置条件 | 选择文件后自动上传，并且有进度条提示上传情况，成功后提示，“上传上成功！” | | |
| 补充说明 | 文件上传较慢，需要耐心等待 | | |
| 界面原型 | 如图2-6所示 | | |

表2-6 上传视频用例分析表

|  |  |  |  |
| --- | --- | --- | --- |
| 功能名称 | 上传视频 | | |
| 操作用户 | 管理员 | 用例描述 | 管理员上传所需转写的视频文件 |
| 前置条件 | 管理员登录  进入文件上传页面 | | |
| 基本事件流 | 选择需要上传的mp4文件 | | |
| 异常事件流 | 选择非规定的文件类型，提示“上传只能是 WAV或Mp4 格式！”  文件大小超出限制，提示“上传文件大小不能超过 500MB！”  文件内容出错，提示“上传失败！” | | |
| 后置条件 | 选择文件后自动上传，并且有进度条提示上传情况，成功后提示，“上传上成功！” | | |
| 补充说明 | 文件上传较慢，需要耐心等待 | | |
| 界面原型 | 如图2-6所示 | | |



图2-6 上传界面原型图

## 2.5 需求分析-游客

图片检索是当用户上传至图片资源到云端，云端通过一系列计算得到与之相似的图片返回给用户，用例分析表如表2-7所示。

表2-7 音频检索用例分析表

|  |  |  |  |
| --- | --- | --- | --- |
| 功能名称 | 音频检索 | | |
| 操作用户 | 游客 | 用例描述 | 通过输入关键词对音频进行检索 |
| 前置条件 | 进入本系统的主页 | | |
| 基本事件流 | 上传待检索图片 | | |
| 异常事件流 | 云端出错，提示“服务正在维护中，请稍后重试” | | |
| 后置条件 | 用户确认上传后自动提交 | | |
| 补充说明 | 无 | | |
| 界面原型 | 如图2-8所示 | | |



图2-7 音频检索界面原型

# 第3章 系统分析

# 第4章 系统设计

系统设计的优良程度关乎决定系统运行时的稳定性、以及发布后的安全性、用户使用时的便利性、后期维护时的成本性。本章将从架构设计、工作流程设计，类设计，以及数据库设计这四个方面来阐述本系统的设计思路。

## 4.1 架构设计

跨媒体信息检索系统的设计与实现，根据需求分析主要分为三个模块的任务：音视频存储与检索方案，图片存储与检索方案，搜索引擎构建。这三个模块对应的分别是音频、视频、图片、以及文本类型媒体数据的文件分析。

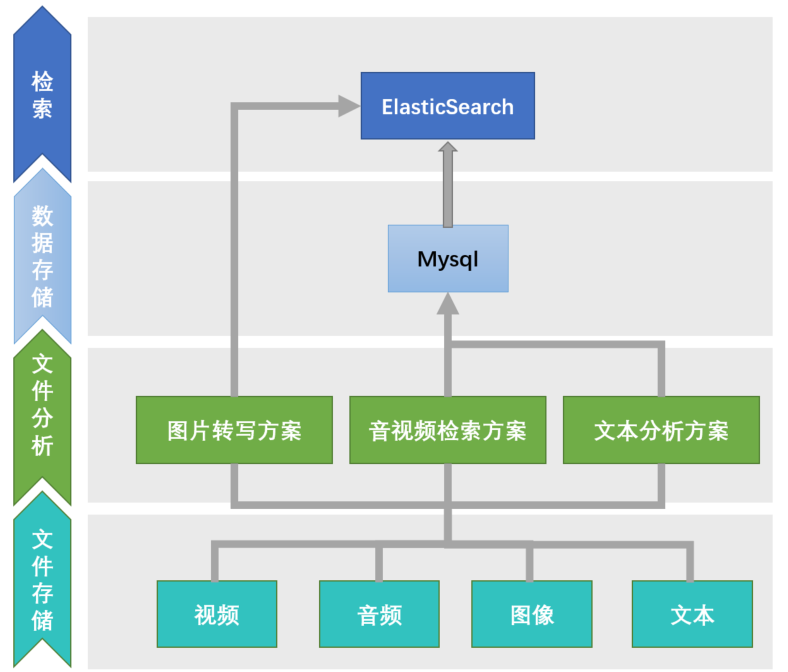


图4-1 系统架构设计图

## 4.2 工作流程设计

本章主要描述图片检索的工作流程，首先使用VGG16对数据集进行特征向量提取，并将特征向量和图片路径作为键值对存储在h5文件中（HDFS5），通过 group 有效的将多种 dataset 进行管理和划分。检索时通过提取被检索图片的特征向量，与h5文件中的向量集进行夹角余弦距离计算特征向量之间的距离。获得距离最小的数据集。

　　图像检索过程简单说来就是对图片数据库的每张图片抽取特征(一般形式为特征向量)，存储于数据库中，对于待检索图片，抽取同样的特征向量，然后并对该向量和数据库中向量的距离（相似度计算），找出最接近的一些特征向量，其对应的图片即为检索结果。

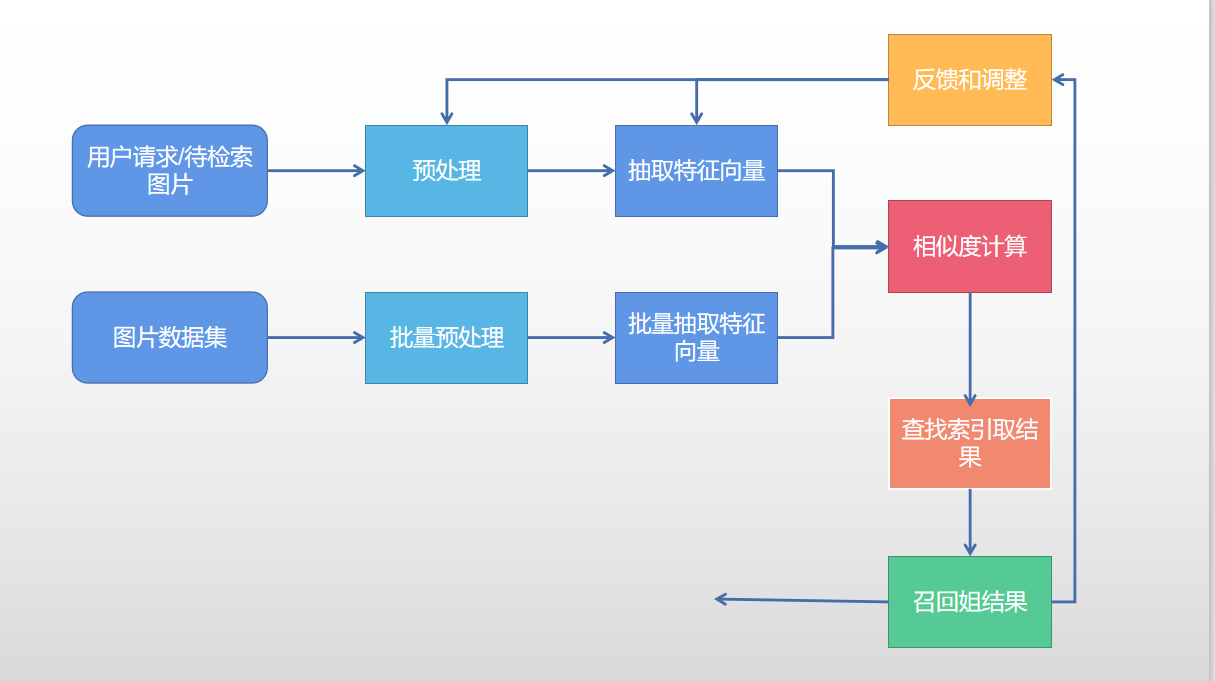


图4-2图片检索流程图

## 4.2 设计类

## 4.3 数据库设计

### 4.3.1 概念模型

### 4.3.1 关系模型

### 4.3.1 物理模型

# 第5章 系统实现

# 第6章 系统测试

# 第7章 结束语

首先感谢我的导师郭丽，在大学四年对我的培养和支持，让我的专业能力得到了很大的提升。感谢大学四年来帮助我的学长和同学。是你们让我找到了自己发展的方向。在我学习的道路上给了我很多帮助和建议，感谢我的小组成员王志豪和闫帅。在做跨媒体搜索引擎的时候一起谈论、分析、解决问题。最终完成跨媒体搜索引擎的搭建。

感谢中原工学院的所有教师和领导。在前三年的学习中。基础知识对我来说非常重要。这对于今后我的职业生涯和专业方面有很大帮助。　时光匆匆如流水，转眼便是大学毕业时节，春梦秋云，聚散真容易。离校日期已日趋渐进，毕业论文的完成也随之进入了尾声。从开始进入课题到论文的顺利完成，一向都离不开教师、同学、朋友给我热情的帮忙，在那里请理解我诚挚的谢意!在此我向中原工学院学校软件学院专业的所有教师表示衷心的感激，多谢你们三年的辛勤栽培，多谢你们在教学的同时更多的是传授我们做人的道理，多谢三年里面你们孜孜不倦的教诲!

最终感激我的家人，这么多年来，正是你们的支持和鼓励，才使我顺利地完成学业;正是你们的关心和默默的奉献，给我创造了优越的条件，使我在学习的道路上乐观向上、勇往直前。

# 致谢

# 参考文献