

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**

互联网的普及，越来越多的人开始变得依赖搜索引擎来获取信息。在日常的生活和工作中几乎每个人都离不开搜索引擎，通过搜索引擎来获取大量的更加全面的详细的信息。搜索引擎服务与人们生活的的各个方面等都紧密相连。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，比如当我们要搜索一段视频的背景音乐，或者是想知道一张图片的更多信息，我们很难用文字描述这些媒体信息，这也就间接导致了搜索结果的不准确性。正是由于这种强烈的需求，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，不同于过去传统的搜索引擎，我们在传统引擎的基础上丰富了检索内容。提高用户的检索效率。比如只需要上传一张本地的图片到云端，即可通过云端检索到与该图片相关的所有图片信息。通过简单的文字搜索即可得到与其相关的文本、音频、视频等信息。

图片检索的大体思路为提取图片的特征向量，然后通过一系列的计算公式计算出两个不同向量之间距离，对结果进行排序整理。距离差距越小，就说明两个图片的相似度就越高。

**关键词**：跨媒体；信息检索；特征向量；向量距离计算；

**Abstract**

The popularity of the Internet has made people more and more dependent on search engines to obtain information. In daily life and work, almost everyone is inseparable from search engines, through which they can obtain a large amount of more comprehensive and detailed information. Search engine services are closely connected with all aspects of people's lives. 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 we want to search for background music in a video or want to know more information about a picture, it is difficult for us to describe the media information in text. This indirectly leads to the inaccuracy of search results. It is precisely because of this strong 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 search content in the functions of traditional search engines. Improve user retrieval efficiency. For example, only a local picture needs to be uploaded to the cloud, and all picture information related to the picture can be retrieved through the cloud. You can get related text, audio, video and other information through a simple text search.

The general idea of image retrieval is to extract the feature vector of the image, and then calculate the distance between two different vectors through a series of calculation formulas, and sort the results. The smaller the distance difference, the higher the similarity between the two pictures.

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

目录

[第1章 项目简介 1](#_Toc14948)

[1.1 项目背景 1](#_Toc17353)

[1.2 研究状况调研情况 3](#_Toc13039)

[1.3 项目目标 3](#_Toc19754)

[第2章 系统需求分析 5](#_Toc27237)

[2.1 系统概述 5](#_Toc9760)

[2.2 系统用户 5](#_Toc29466)

[2.3 用例分析 6](#_Toc7255)

[2.4 需求分析-管理员 8](#_Toc32274)

[2.5 需求分析-游客 11](#_Toc22208)

[第3章 系统分析 13](#_Toc20611)

[3.1 可行性分析 13](#_Toc8064)

[3.2 分析时序图 13](#_Toc26282)

[第4章 系统设计 14](#_Toc31229)

[4.1 架构设计 14](#_Toc13823)

[4.2 工作流程设计 14](#_Toc20151)

[4.2 设计类 15](#_Toc21064)

[4.3 数据库设计 16](#_Toc659)

[第5章 系统实现 19](#_Toc5160)

[5.1 管理员端 19](#_Toc15800)

[5.2 游客端 21](#_Toc5387)

[第6章 系统测试 25](#_Toc1594)

[6.1 相似度算法测试 25](#_Toc13224)

[6.2 用户端 26](#_Toc13597)

[第7章 结束语 29](#_Toc11636)

[致谢 30](#_Toc16507)

[参考文献 31](#_Toc31924)

# 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，但是图片类型的信息在计算机的识别过程中，与人的大脑在图片信息识别的过程中能直接理解其中所包含信息的能力相差甚远，基于语义化的图像类型信息检索是跨媒体信息检索研究的高级目标。

图像检索，系统是一个电脑浏览的系统，电脑从一个大型的图像存储数据库中去搜索图像。大多数传统的和一般图像检索的方式是利用一些增加元数据(meta-data)的方法，比如：电影字幕、一段信息文本关键词或是一张图片的解释说明，如此一来就可以透过注解词完成检索。人工给图像添加注解是一项费时、费力并且实现不太理想的解决方案；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。

从上个世纪70年代的时候，关于图像信息检索的研究项目就已经开始，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。到了上世纪90年代以后，慢慢的开始出现了对图像信息的内容语义，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。并且图像检索的关键问题是对人体视觉信息处理机制的进一步了解,简单来说，就是去探求我们自身是如何去处理我们看到的图像内容的, 这个问题是能够进一步优化数据特征索引技术, 解决大规模数据库检索速度问题的关键。多媒体数据信息压缩技术和互联网的高速发展，现在的信息形式多种多样并且十分复杂。视觉的信息数据包括单张的图像数据和动态的视频数据，对于视频数据的动态特点，如何保证能够实现高速并且可靠的视频检索。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。跨媒体信息检索系统的信息语义解析大致可以分为“文本语义解析”、“图片语义解析”、 “音频语义解析”、“视频语义解析”这四个方面，下面开始分别介绍这四个方面的当前研究状况：

（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：现在市面上音频语义的解析相关的引擎非常多，涉及技术手段也比较多，与其相关的识别软件诸如“微信搜歌”已经在市面上推出。

（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。

总得俩说，到目前为止单个消息类型的解析算法已经趋于成熟，但是如何将不同消息类型的解析算法集合到一起，实现跨媒体信息检索，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，返回最高度相似的搜索关键词的一组结果。

# 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，返回结果的消息类型有文本消息、音频消息、视频消息和图片消息。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。

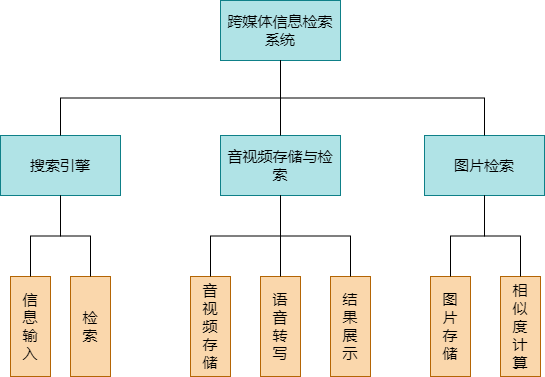


图2-1 系统功能总览

## 2.2 系统用户

### 2.2.1系统管理员

跨媒体信息检索系统的系统管理员主要负责本系统中数据源信息的录入和管理等功能，管理员在门户的登陆框通过正确输入自己的账号和密码后，即可进入管理端页面，对系统中的数据源进行修改和维护。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。首先管理员在门户的登陆页面进行注册。经过一系列操作注册成功后。即可使用刚才的账号和密码，返回门户的登陆页面，输入刚才的账号和密码，进入管理端页面，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。

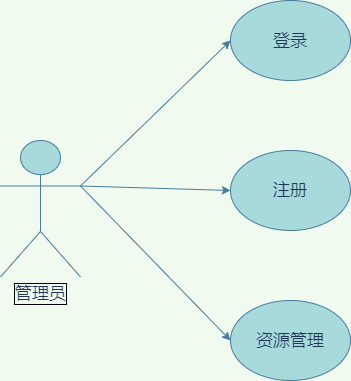


图2-2 管理员用例图

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。

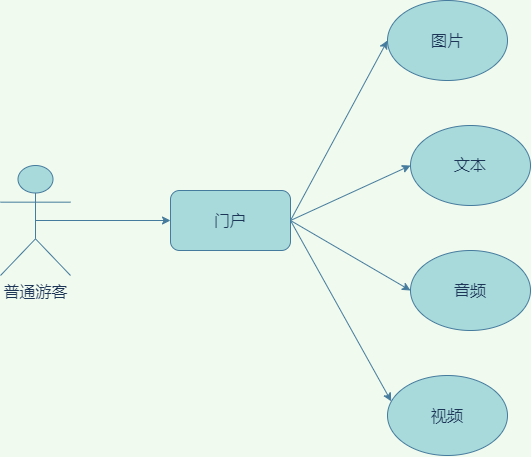


图2-3 游客用例图

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 | | |

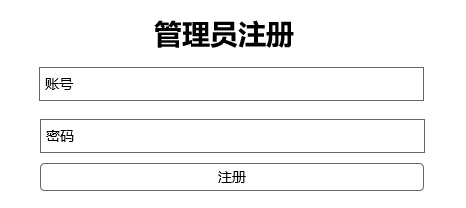


图2-4 注册界面原型图

### 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 | | |

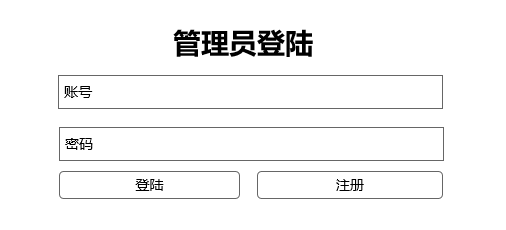


图2-5 登录界面原型图

### 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

通过对前面系统用例要求的详细描述和对当前市面上实现图片检索技术方案的研究。本章主要描述图片检索系统的整体业务流程。

## 3.1 可行性分析

图片检索功能目前使用到的技术主要有Python、TensorFlow、VGG16、HDFS5、相似度计算公式等。目前这些技术在业界已经有了广泛的使用。有两个主要问题。一个是如何将这些技术整合用以来实现图片检索的需求。另一个是如何找出准确而又高效的向量距离计算公式。这两个主要的问题解决方案在第四章和第五章。

## 3.2 分析时序图

用户在门户上传带检索的图片到云端后，有云端通过一系列运算得出与之相匹配的结果返回给用户（用户上传的图片也会被云端保存，用于以后的数据分析和数据源，增加搜索引擎的搜索效率）。图片检索时序图如图3-1所示。



图3-1图片检索时序图

# 第4章 系统设计

在软件开发的过程中，一个系统设计的好坏决定着这个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。

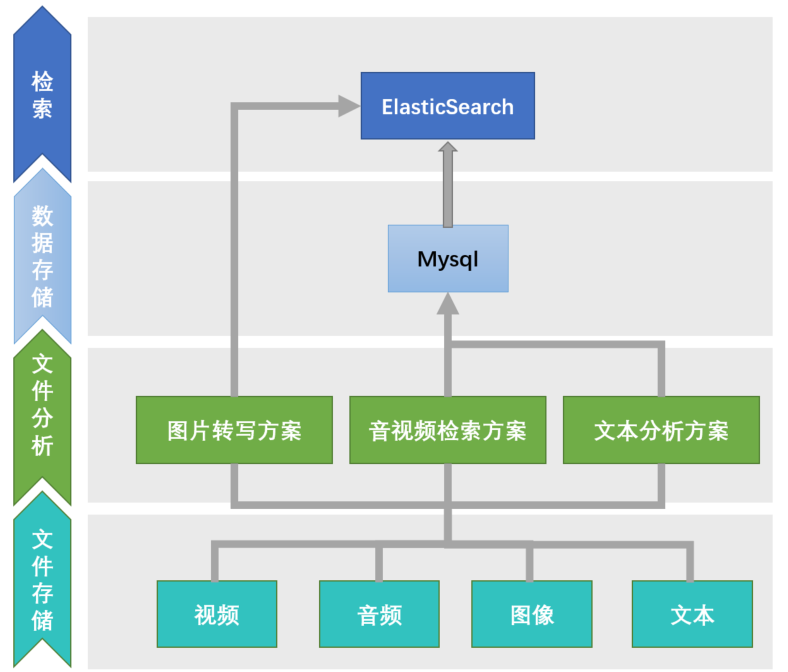


图4-1 系统架构设计图

## 4.2 工作流程设计

图片检索的工作流程不是特别复杂，关键在于找到高效的算法，计算出数据集中与待搜索图片之间的相似度。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，然后通过一系列的算法对该向量和数据库中向量的距离（相似度计算），得到一对结果，对结果进行排序，删除掉差距较大的一些数据，找出最接近的一些特征向量，其对应的图片即为检索结果。

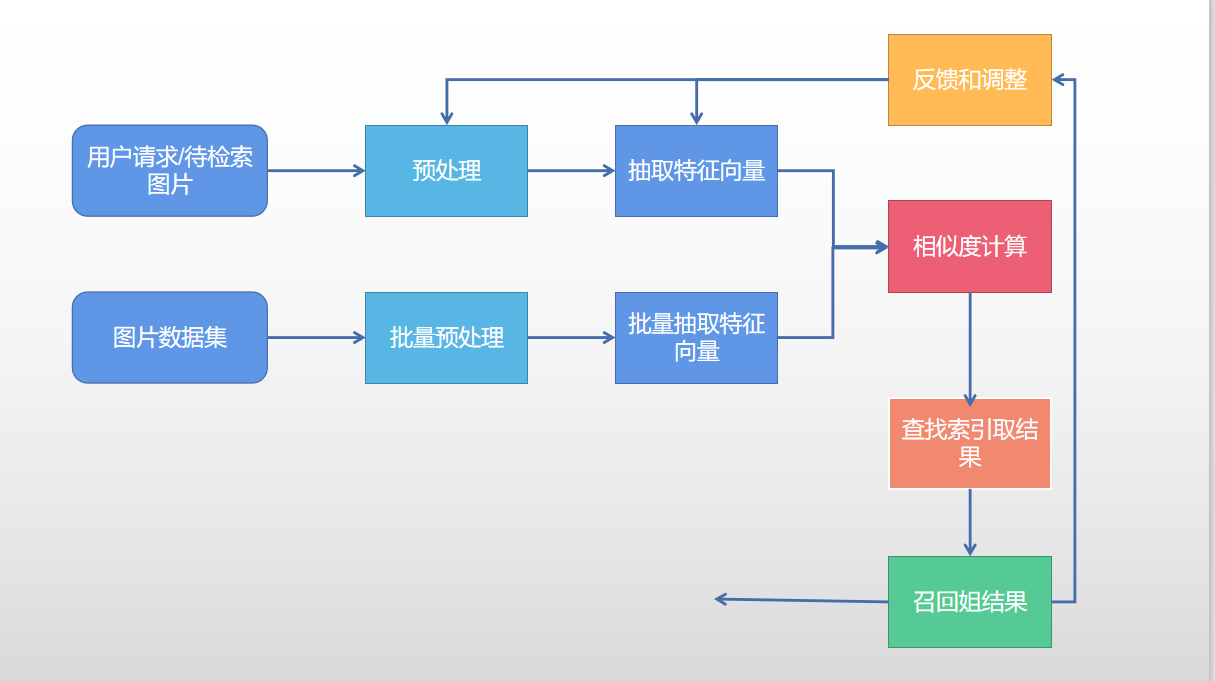


图4-2图片检索流程图

## 4.2 设计类

经过对系统需求进行分析，将数据的类型分为三个，分别是文本类、视频类和音频类。用户单独抽离出一个用户类，其关系如图4-3实体类设计图所示。



图4-3 实体类设计图

## 4.3 数据库设计

### 4.3.1 概念模型

在本系统中有Audio（音频信息表），Video（视频信息表），Text（文本信息表），User（用户表），该系统的数据库设计如图4-4数据库设计E-R图所示。



4-4 数据库设计E-R图

### 4.3.1 关系模型

关系模式是数据库设计中，库中各个表的联系和表中属性的结构，本系统数据库的关系模式如下所示：

Audio（id、sentence\_time、file\_name、start\_time、duration、insert\_time）；

音频信息表（音频信息ID，转写结果，文件名，开始时间，音频时长，插入时间）；

Video（id、sentence\_time、file\_name、start\_time、duration、insert\_time）；

音频信息表（视频信息ID，转写结果，文件名，开始时间，音频时长，插入时间）；

Text（id、title、date、category、author\_name、url、pic、content）；

文本信息表（文本ID，标题，日期，来源，作者，地址，封面图，内容）；

User（id、username、password、type、last\_join\_time）；

用户信息表（用户ID，用户名，密码，类型，上次登录时间）；

### 4.3.1 物理模型

本系统使用Mysql作为数据库存储软件，使用utf8mb4 -- UTF-8 Unicode作为字符集，排序规则使用utf8mb4\_unicode\_ci,本系统的数据库物理模型如下所示。

Audio表存储转写后的音频数据的文本内容，如下表4-1Audio表所示。

表4-1 Audio表

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 字段名称 | 数据类型 | 是否非空 | 是否主键 | 备注 |
| id | int | 否 | 是 | 音频信息id |
| sentence\_text | varchar | 否 | 否 | 音频转写的文本信息 |
| file\_name | char | 否 | 否 | 文件名 |
| start\_time | int | 否 | 否 | 开始时间，毫秒为单位 |
| duration | int | 否 | 否 | 音频时长，毫秒为单位 |
| insert\_time | datetime | 否 | 否 | 插入时间 |

Video表主要存储转写后的视频频数据的文本内容，如下表4-2 Video表所示。

表4-2 Video表

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 字段名称 | 数据类型 | 是否非空 | 是否主键 | 备注 |
| id | int | 否 | 是 | 视频信息id |
| sentence\_text | varchar | 否 | 否 | 视频转写的文本信息 |
| file\_name | char | 否 | 否 | 文件名 |
| start\_time | int | 否 | 否 | 开始时间，毫秒为单位 |
| duration | int | 否 | 否 | 视频时长，毫秒为单位 |
| insert\_time | datetime | 否 | 否 | 插入时间 |

Text表主要存储爬取的文本信息的内容，如下表4-3 Text表所示。

表4-3 Text表

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 字段名称 | 数据类型 | 是否非空 | 是否主键 | 备注 |
| id | int | 否 | 是 | 文本信息id |
| title | varchar | 否 | 否 | 文本标题 |
| date | datetime | 否 | 否 | 爬取时间 |
| category | varchar | 否 | 否 | 文本来源 |
| author\_name | varchar | 否 | 否 | 作者名 |
| url | varchar | 否 | 否 | 文本地址 |
| pic | varchar | 否 | 否 | 封面地址 |
| content | text | 否 | 否 | 文本内容 |

User表主要存储系统用户的账号密码，用于管理员登录。如下表4-4 User所示。

表4-4 User表

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 字段名称 | 数据类型 | 是否非空 | 是否主键 | 备注 |
| id | int | 否 | 是 | 文本信息id |
| username | varchar | 否 | 否 | 用户名 |
| password | varchar | 否 | 否 | 密码 |
| type | int | 否 | 否 | 类型 |
| last\_join\_time | datetime | 否 | 否 | 上次登录时间 |

# 第5章 系统实现

通过前面的系统分析和系统设计，前后端已经将功能全部实现，并且完成对接和上线，本章主要对系统的登录、注册、上传文件、信息检索、图片搜索等功能进行描述和展示。

## 5.1 管理员端

管理员端的功能有登录，注册，以及登录后的文件上传转写功能。

### 5.1.1 登录

用户通过门户页面点击登录按钮进入登录界面，在登录框中输入正确的管理员账号密码，通过http请求传输到后端，由后端对账号密码进行数据校验，并返回登录结果。具体实现显示效果图如图5-1所示。



图5-1 系统管理员登录实现效果图

### 5.1.2 注册

如果尚未拥有管理员账号，需要用户在登录界面，点击注册页面，进行注册操作，输入账号密码，提交后会对结果进行反馈。具体实现显示效果图如图5-2所示。



图5-2 管理员注册实现效果图

### 5.1.3 文件上传

登录后进入管理员界面，可以进行文件上传操作，支持上传wav格式的音频文件，mp4格式的视频文件，上传后会有上传进度提示，上传成功会提示“上传成功”，上传后即可开始转写操作。具体实现显示效果图如图5-3所示。



图5-3 管理员文件上传现效果图

## 5.2 游客端

游客端的功能有对音视频，文本以及图片的检索功能。

### 5.2.1 音频检索

游客进入首页，首页效果如图5-4 所示，输入想要搜索的目标关键字，点击回车后，即向后端发送音频检索请求，首先会对关键词进行词向量计算，得出几个相似词，然后对这几个关键词都进行搜索，然后将结果返回。一个音频中可能有多个目标搜索结果，前端也会对返回的数据进行整合，实现在一个音频内可以进行跳转的操作，具体实现显示效果图如图5-5所示。



图5-4 首页实现效果图



图5-5 音频检索实现效果图

### 5.2.2 视频检索

游客进入首页，输入目标关键字，点击回车后，即向后端发送视频检索请求，首先会对关键词进行词向量计算，得出几个相似词，然后对这几个关键词都进行搜索，然后将结果返回。一个视频中可能有多个目标搜索结果，前端也会对返回的数据进行整合，实现在一个视频内可以进行跳转的操作，思路和音频检索去重思路一致，具体实现显示效果图如图5-6所示。



图5-6 视频检索实现效果图

### 5.2.3 文本检索

游客进入首页，输入目标关键字，点击回车后，即向后端发送文本检索请求，首先会对关键词进行词向量计算，得出几个相似词，然后对这几个关键词都进行搜索，然后将结果根据检索计算得分。具体实现显示效果图如图5-7所示。



图5-7 文本检索实现效果图

### 5.2.4 图片检索

游客进入首页，点击门户页面中间搜索框里面左侧的图片按钮，会弹出文件选择框，选择需要上传至云端检索的图片，进行图片上传，即可对图片进行相似度计算，然后会将与上传图片相似的图片集进行返回。具体实现显示效果图如图5-7所示。

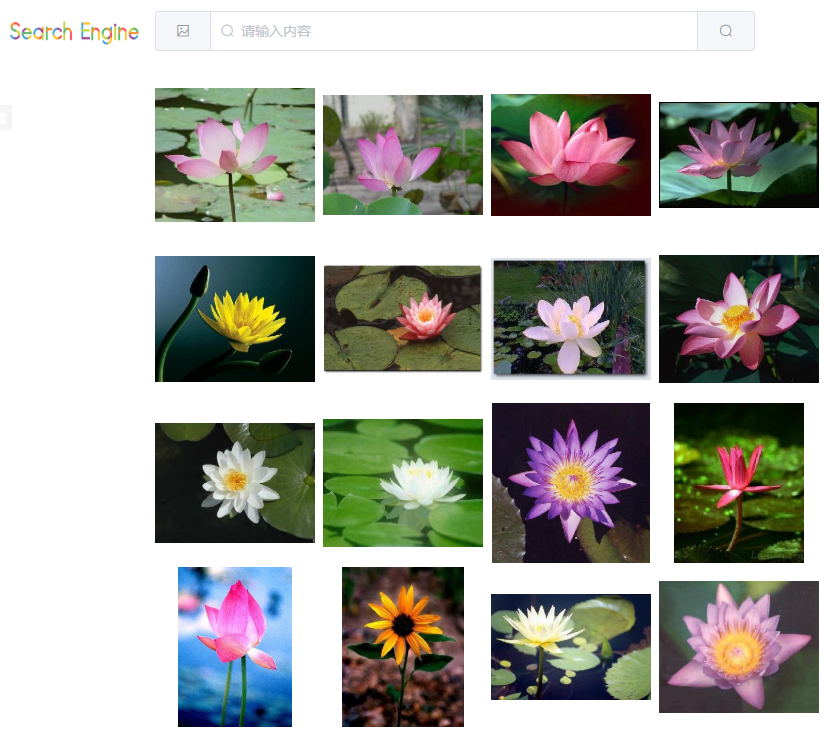


图5-8 图片检索实现效果图

# 第6章 系统测试

由于跨媒体信息检索系统功能比较多，本章主要对跨媒体检索系统-图片检索的相似度计算算法和图片检索准确率和召回率的计算。

## 6.1 相似度算法测试

通过搜集材料，选取了四种计算向量距离的公式，如下所示。

* 欧氏距离：最经典的计算距离的公式，也叫欧几里得距离。是现在最常见的计算向量距离的方法，即在一个多维的空间中任意两个点在一定标准下之间的真实距离，或者是向量的自然长度（也就是该点到原点o（0，0）之间的距离）

Python代码实现：np.sqrt(np.square(query - vec).sum())

* 曼哈顿距离：两个点在标准的坐标系上的距离坐标轴距离的总和。这个公式的经典案例是计算城市中亮点之间的最短路径。

Python代码实现：np.linalg.norm(query - vec, ord=1

* 切比雪夫距离：这个是向量空间中的一种度量，二个点之间的距离定义是其各坐标数值差绝对值的最大值。以数学的观点来看，切比雪夫距离是由一致范数（uniform norm）（或称为上确界范数）所衍生的度量，也是超凸度量（injective metric space）的一种。

Python代码实现：np.linalg.norm(query - vec, ord=np.inf)

* 夹角余弦距离：测量两个不同的二维空间[向量](https://baike.baidu.com/item/%E5%90%91%E9%87%8F" \t "https://baike.baidu.com/item/_blank)的夹角的[余弦](https://baike.baidu.com/item/%E4%BD%99%E5%BC%A6" \t "https://baike.baidu.com/item/_blank)值来度量它们之间的相似性。度数为0的夹角的余弦值是1，而其他任何角度的余弦值都不大于1；并且其最小值是-1。从两个不同的二维向量之间角度的余弦值来确定两个向量是否大致指向同一个方向。两个向完全相同的十一路，相似度的值为1；两个不同的向量两两垂直的时候，相似度的值为0；当两个不同的向量互相指向完全相反方向的时候，这个值为-1。但是这个结果是与向量的本身的长度没有关系的，只与向量的指向方向相关。余弦相似度通常用于正空间，所以这个值在区间 [-1，1]之间

Python代码实现：np.dot(query, vec) / (np.linalg.norm(query) \* (np.linalg.norm(vec)))

测试方法主要为将这些公式抽离成方法，通过调用不同的方法来获取搜索结果，对比这几种算法的结果进行统计，最终选出合适的算法。测试结果如表6-1所示。

表6-1 向量距离计算公式测量结果

|  |  |  |  |
| --- | --- | --- | --- |
| 距离公式 | 预期结果 | 实际结果 | 准确率 |
| 人工 | 45 | 45 | 100% |
| 欧式距离 | 45 | 39 | 86.667% |
| 曼哈顿距离 | 45 | 24 | 53.333% |
| 切比雪夫距离 | 45 | 37 | 82.222% |
| 夹角余弦距离 | 45 | 43 | 95.556% |

根据上面的测试数据，最终选择夹角余弦距离计算公式作为跨媒体检索系统-图片检索的核心算法。

## 6.2 用户端

用户端的功能有对音视频，文本以及图片的检索功能。下面将对这些功能测试的结果进行展示。

### 6.2.1 音频检索

用户进入首页后，在搜索输入框中输入关键词点击搜索，进行检索，测试用用例如下表6-2音频检索功能测试表所示。

表6-2 音频检索功能测试表

|  |  |  |  |
| --- | --- | --- | --- |
| 功能模块 | 信息检索 | 编写者 | 闫帅 |
| 功能名称 | 音频检索 | | |
| 前置条件 | 1.进入首页 | | |
| 操作步骤 | 1.在搜索栏中输入搜索关键字“交通” | | |
| 后置条件 | 跳转到音频检索结果界面 | | |
| 预期结果 | 检索结果中除了关键词“交通”，还有类似的“运输”、“水运”、“空运”等 | | |
| 实际结果 | 检索结果中除了关键词“交通”，还有类似的“运输”、“水运”、“空运”等 | | |

### 6.2.2 视频检索

游客进入首页，在门户页面中间的搜索框中输入关键字进行搜索，测试用例如下表6-3视频检索功能测试表所示。

表6-3 视频检索功能测试表

|  |  |  |  |
| --- | --- | --- | --- |
| 功能模块 | 信息检索 | 编写者 | 闫帅 |
| 功能名称 | 视频检索 | | |
| 前置条件 | 1.进入首页 | | |
| 操作步骤 | 1.在搜索栏中输入搜索关键字“清华” | | |
| 后置条件 | 跳转到视频检索结果界面 | | |
| 预期结果 | 检索结果中除了关键词“清华”，还有类似的“清华大学”、“北大”、“北京大学”等 | | |
| 实际结果 | 检索结果中除了关键词“清华”，还有类似的“清华大学”、“北大”、“北京大学”等 | | |

### 6.2.3 文本检索

游客进入首页，在输入框中输入关键字进行搜索，测试用用例如下表6-4文本检索功能测试表所示。

表6-4 文本检索功能测试表

|  |  |  |  |
| --- | --- | --- | --- |
| 功能模块 | 信息检索 | 编写者 | 闫帅 |
| 功能名称 | 文本检索 | | |
| 前置条件 | 1.进入首页 | | |
| 操作步骤 | 1.在搜索栏中输入搜索关键字“咖啡” | | |
| 后置条件 | 跳转到文本检索结果界面 | | |
| 预期结果 | 检索结果中除了关键词“咖啡”，还有类似的“可可”、“饮料”、“饮品”等 | | |
| 实际结果 | 检索结果中除了关键词“咖啡”，还有类似的“可可”、“饮料”、“饮品”等 | | |

### 6.2.4 图片检索

游客进入首页，在图片选择框中选择目标图片进行挤奶哦是，测试用用例如下表6-5图片检索功能测试表所示。

表6-5 图片检索功能测试表

|  |  |  |  |
| --- | --- | --- | --- |
| 功能模块 | 图片检索 | 编写者 | 闫帅 |
| 功能名称 | 图片检索 | | |
| 前置条件 | 1.进入首页 | | |
| 操作步骤 | 1.在搜索栏的图片选择框中选择图片“荷花” | | |
| 后置条件 | 跳转到图片检索结果界面 | | |
| 预期结果 | 将于输入图片相似的图片进行返回，返回“荷花”的图片集。 | | |
| 实际结果 | 将于输入图片相似的图片进行返回，返回“荷花”的图片集。 | | |

# 结束语

在这次毕业设计的开发过程中，我学到了许多在学校里没有学到过的知识，更认识到了数学在计算机领域的关键性和重要性，这次的跨媒体信息检索系统是由我和闫帅、王志豪三个人一起合作开发的毕业设计。这个项目的开发对我来说是大学生活中最好的历练，我们几个人一起从需求分析，系统设计，系统测试开始，然后每人根据能力完成其中的一部分。  
 在这次毕设管理系统中，我主要负责图片检索与存储的实现以及前端脚手架和接口的对接，主要工作流程就是搜集图片的数据集，通过TensorFlow和VGG16来提取这些数据集中每个图片的二维特征向量，也就是一个二维数组，将这些数据以文件路径-特征向量这种key-value的方式存储在hdfs5文件中。检索的过程中通过一系列算法计算出hdfs5中特征向量与待检索图片特征向量的距离，得到相似度最高的一组，将其文件路径返回给用户。

在这几个月的时间里，我学到了许多有用的知识，并且了解了在学校里平时没有接触过的东西，更认识到了数学知识对计算机的帮助。四年的大学生活，收获颇丰，在逆境中成长，再接再厉，不断前行！

# 致谢

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。这对于今后我的事业发展有很大帮助。　转眼便是大学毕业时节，大学的时光过的很快，转眼间离校日期已经日趋渐进，毕业也随之进入了尾声。从开始进入毕业设计选题到毕业论文的顺利完成，一向都离不开我的导师郭丽、同学和朋友给我热情的帮忙，在这里请理解我诚挚的谢意!在这里我向中原工学院软件学院的所有教学老师和同学表示由衷的感激，多谢你们在大学生活中对我的支持和帮助，感谢老师们在是三年的时光里教会了我很专业知识，给予我很多实践的机会，让我能够将自己所学的知识得到发挥。

最后在此发自内心的感谢这么多年来支持我的学业的家人们，正是有了你们对我学业的鼓励和帮助，才能使我现在如此顺利地完成学业。在今后的生活中我会牢牢记住在我人生道路上对我帮助过的老师、家人以及朋友们。乐观向上、勇往直前。

# 参考文献

[1]范朗.Elasticsearch海量数据存储查询优化[J].工业控制计算机,2020,10:85-87.

[2]蔡平,王志强,傅向华.基于语义的跨媒体信息检索技术研究[J].微电子学与计算机,2010,03:102-105.

[3] 邵志强.跨媒体检索平台[D].山东:山东师范大学,2015.

[4]曾亚飞.基于Elasticsearch的分布式智能搜索引擎的研究与实现[D].重庆:重庆大学,2016.

[5]王伟,魏乐,刘文清,等.基于ElasticSearch的分布式全文搜索系统[J].电子科技,2018,08:56-59.

[6]张恒均.分布式存储——数据中心建设解决策略研究[J].数码世界,2020,10:54-55.

[7]袁爱平,陶志勇,邓河,等.云计算环境中HDFS数据块存储策略研究[J].电脑知识与技术,2020,26:33-35.

[8]陈虎.基于HDFS的云存储平台的优化与实现[D].广州:华南理工大学,2012.

[9]Mohamed Elkawkagy,Heba Elbeh.High Performance Hadoop Distributed File System[J].International Journal of Networked and Distributed Computing,2020,8（3）:119-123.