

《软件架构》课程报告

课题名称：Voting For Decision

课题负责人名（学号）：孟庆汉 0843042109

同组成员名单（角色）： 苟丁 0843042229

彭韬 0843042037

闫立玮 0843042129

指导教师：潘 无 名

评阅成绩: _____

评阅意见: _____

提交报告时间: 2011 年 05 月 30 日

Voting For Decision

软件工程 专业

学生 孟庆汉 0843042109 **指导老师** 潘无名
苟丁 0843042229
彭韬 0843042037
闫立玮 0843042129

[摘要] 搜索多个给定的 HTML 或文本文件，从中确定一个最符合搜索关键词的文件，其中包括对 HTML 或文本文件中文本的扫描，计算文件和关键词组的接近程度。另外含有投票模块，根据不同搜索模块得到的结果进行投票，返回多数支持的结果，要求总能得到多数支持的结果；最后打开结果文件。

关键词：搜索 HTML 文件 文本文件 接近程度 投票模块

目录

《软件架构》课程报告	0
1. 引言（项目背景）	3
2. 相关工作（对现有的系统进行分析）	3
3. 问题描述（需求分析）	3
(1)概要	3
(2)使用场景	3
A.关键字匹配总次数	4
B. 关键字权值	4
C. 基于关键词覆盖率	4
(3)需求实用分析	4
4. 解决方案（概要设计）	5
1.功能需求	5
功能划分	5
2.功能描述	6
关键词匹配的总次数搜索排名	6
关键词权值搜索排名	7
关键词覆盖率搜索排名	9
投票	10
显示投票最高的部分	11
2.性能需求	13
数据精确度	13
时间特性	13
适应性	13
3.运行需求	14
用户界面	14
硬件接口	16
5. 实现细节（详细设计、实现）	16
1. 系统的结构	16
类图	17
2. 模块设计说明	18
3. 部分关键代码	20
代码 1: 系统主窗口相关代码	20
代码 2: 实现按照关键词匹配的总次数排列模块功能.....	26
代码 3: 实现按照关键词权值排列模块功能.....	28
代码 4: 实现按照关键词权值排列模块功能.....	30
代码 5: 实现投票并且打开得票最高的模块功能.....	32
6. 结束语（结论、进一步的工作）	34
参考文献	34

1. 引言（项目背景）

近年来，随着网络技术的不断发展，我们身边各式各样的网站也越来越多，如何更快的从浩瀚的网站中尽快找到需要的网页是那样的重要；系统实现以下几方面功能：

- a) 搜索多个给定的 HTML 或文本文件，从中确定一个最符合搜索关键词的文件；
- b) 搜索模块中包括对 HTML 或文本文件中文本的扫描，计算文件和关键词组的接近程度；
- c) 有 form 图形界面；
- d) 有投票模块，根据搜索模块得到的结果进行投票，返回多数支持的结果，要求总能得到多数支持的结果；
- e) 最后打开结果文件。

2. 相关工作（对现有的系统进行分析）

NA

3. 问题描述（需求分析）

(1) 概要

快速获取需要的网页信息文本信息是搜索的最大优点；本系统搜索多个给定的 HTML 或文本文件，从中确定一个最符合搜索关键词的文件，其中包括对 HTML 或文本文件中文本的扫描，计算文件和关键词组的接近程度。另外含有投票模块，根据不同搜索模块得到的结果进行投票，返回多数支持的结果，要求总能得到多数支持的结果；最后打开结果文件。

(2) 使用场景

本系统是用来解决快速获取所需要的网页或者文本信息的问题。该系统使用了搜索技术，用户可以利用输入关键字来检索所需要的网

页或者文本信息,对于用户来说使用方便,操作简单。以下是该系统的三个重要功能应用场景分析:

A. 关键字匹配总次数

按照关键词匹配的总次数排列所搜索到的包含关键字的 HTML 文件或者文本文件,其中关键字可以为字母、汉字、标点或者由这些符号所组成的字符串;其中每个关键字之间用空格间隔;

B. 关键字权值

按照关键词权值排列所搜索到的包含关键字的 HTML 文件或者文本文件,其中关键字可以为字母、汉字、标点或者由这些符号所组成的字符串;其中每个关键字之间用空格间隔;

C. 基于关键词覆盖率

按照关键词覆盖率排列所搜索到的包含关键字的 HTML 文件或者文本文件,关键字覆盖率为所输入的关键字在所搜索到的文档里面是否全部出现,其中关键字可以为字母、汉字、标点或者由这些符号所组成的字符串;其中每个关键字之间用空格间隔;

(3) 需求实用分析

随着社会的日益发展,各种 HTML 文档以及文本文档越来越多,如何快速的从这些繁多的文档里面找到自己需要的信息成了一门很有研究价值的研究领域;而针对不同的用户有存在着不同的搜索需求,有的用户需要按照关键字出现的总次数来搜索匹配所得到的结果,有的用户需要按照关键字的权值来搜索匹配所得到的结果,而还有的用

户需要按照关键字的覆盖率来搜索匹配所得到的结果，故本系统从这三个方面来实现，将系统分为这三个搜索模块；

为了更好的与用户交互，系统为用户提供一个投票的模块，支持用户手动对上述三种方案搜索到的结果进行投票选择，同时支持用户打开得票最高的那项；

4. 解决方案（概要设计）

本系统通过关键词匹配的总次数、关键词权值以及关键词覆盖率三种方式搜索指定的 HTML 文档或者文本稳定，，利用搜索技术在知道的文档里面来快速找到所需要的信息，以此提高人们的工作效率，减轻众多信息给大家带来的痛苦。

本系统包括以下几个功能：

- 关键词匹配的总次数搜索排名
- 关键词权值搜索排名
- 关键词覆盖率搜索排名
- 投票
- 显示投票最高的部分

本系统是在指定的文档中搜索包含关键字的文档，同时用户还可以对搜索得到的结果进行投票、排名，并且打开得票数最多的那个模块，所以本系统所面向的用户为可以正常使用电脑以及可以正常使用搜索引擎的群体；

1.功能需求

功能划分

- 关键词匹配的总次数搜索排名
- 关键词权值搜索排名

- 关键词覆盖率搜索排名
- 投票
- 显示投票最高的部分

2.功能描述

关键词匹配的总次数搜索排名

该模块通过摄像头获取包含人脸的视频，然后通过人脸检测检测出含有人脸的图片，再对含有人脸的图片进行处理，对系统进行训练，当下次碰到相同的人脸时系统自动匹配人脸，并且输出该人脸图像所对应的人名的语音信息，从而将人脸这一视觉信息转换为语音信息以供盲人用户所使用，解决他们不能自主获取人脸这一视觉信息的问题，帮助他们了解自己周围都有谁；

按照关键词匹配的总次数排列所搜索到的包含关键字的 HTML 文件或者文本文件，其中关键字可以为字母、汉字、标点或者由这些符号所组成的字符串；其中每个关键字之间用空格间隔；其实现步骤如下所示：

- ✧ 搜索给定文件
- ✧ 分别与各个关键字匹配
- ✧ 每匹配成功一次匹配次数加 1
- ✧ 把每个文件每个关键字匹配字数相加
- ✧ 次数最多的为最佳匹配

按照关键词权值排列所搜索到的包含关键字的 HTML 文件或者文本文件，其中关键字可以为字母、汉字、标点或者由这些符号所组成的字符串；其中每个关键字之间用空格间隔；其实现步骤如下所示：

- ✧ 为第一个关键字赋最高权值，以后的关键字权值依次递减
- ✧ 搜索文件，各个关键字和给定的文件进行匹配
- ✧ 每匹配成功一次，加相应关键字的权值
- ✧ 把每个文件所有关键字匹配的权值相加

✧ 权值最高的为最佳匹配

按照关键词覆盖率排列所搜索到的包含关键字的 HTML 文件或者文本文件, 关键字覆盖率为所输入的关键字在所搜索到的文档里面是否全部出现, 其中关键字可以为字母、汉字、标点或者由这些符号所组成的字符串; 其中每个关键字之间用空格间隔; 其实现步骤如下所示:

✧ 搜索文件, 各个关键字和给定的文件进行匹配

✧ 当文件每成功匹配一个关键字, 覆盖范围加一

✧ 把每个文件覆盖范围进行比较

✧ 覆盖率最高的为最佳匹配

为了更好的体现用户交互, 得到用户需要的结果, 系统支持将通过按照关键词匹配的总次数排列所搜索到的包含关键字的 HTML 文件或者文本文件、按照关键词权值排列所搜索到的包含关键字的 HTML 文件或者文本文件以及按照关键词覆盖率排列所搜索到的包含关键字的 HTML 文件或者文本文件所得到的结果进行手动投票, 并且显示各个部分票数;

将投票部分得票最高的部分显示;

关键词权值搜索排名

该模块通过摄像头获取包含人脸的视频, 然后通过人脸检测检测出含有人脸的图片, 再对含有人脸的图片进行处理, 对系统进行训练, 当下次碰到相同的人脸时系统自动匹配人脸, 并且输出该人脸图像所对应的人名的语音信息, 从而将人脸这一视觉信息转换为语音信息以供盲人用户所使用, 解决他们不能自主获取人脸这一视觉信息的问题, 帮助他们了解自己周围都有谁;

按照关键词匹配的总次数排列所搜索到的包含关键字的 HTML 文件或者文本文件, 其中关键字可以为字母、汉字、标点或者由这些符号所组成的字符串; 其中每个关键字之间用空格间隔; 其实现步骤如下所示:

✧ 搜索给定文件

- ✧ 分别与各个关键字匹配
- ✧ 每匹配成功一次匹配次数加 1
- ✧ 把每个文件每个关键字匹配字数相加
- ✧ 次数最多的为最佳匹配

按照关键词权值排列所搜索到的包含关键字的 HTML 文件或者文本文件, 其中关键字可以为字母、汉字、标点或者由这些符号所组成的字符串; 其中每个关键字之间用空格间隔; 其实现步骤如下所示:

- ✧ 为第一个关键字赋最高权值, 以后的关键字权值依次递减
- ✧ 搜索文件, 各个关键字和给定的文件进行匹配
- ✧ 每匹配成功一次, 加相应关键字的权值
- ✧ 把每个文件所有关键字匹配的权值相加 .
- ✧ 权值最高的为最佳匹配

按照关键词覆盖率排列所搜索到的包含关键字的 HTML 文件或者文本文件, 关键字覆盖率为所输入的关键字在所搜索到的文档里面是否全部出现, 其中关键字可以为字母、汉字、标点或者由这些符号所组成的字符串; 其中每个关键字之间用空格间隔; 其实现步骤如下所示:

- ✧ 搜索文件, 各个关键字和给定的文件进行匹配
- ✧ 当文件每成功匹配一个关键字, 覆盖范围加一
- ✧ 把每个文件覆盖范围进行比较
- ✧ 覆盖率最高的为最佳匹配

为了更好的体现用户交互, 得到用户需要的结果, 系统支持将通过按照关键词匹配的总次数排列所搜索到的包含关键字的 HTML 文件或者文本文件、按照关键词权值排列所搜索到的包含关键字的 HTML 文件或者文本文件以及按照关键词覆盖率排列所搜索到的包含关键字的 HTML 文件或者文本文件所得到的结果进行手动投票, 并且显示各个部分票数;

将投票部分得票最高的部分显示;

✧

关键词覆盖率搜索排名

该模块通过摄像头获取包含人脸的视频, 然后通过人脸检测检测出含有人脸的图片, 再对含有人脸的图片进行处理, 对系统进行训练, 当下次碰到相同的人脸时系统自动匹配人脸, 并且输出该人脸图像所对应的人名的语音信息, 从而将人脸这一视觉信息转换为语音信息以供盲人用户所使用, 解决他们不能自主获取人脸这一视觉信息的问题, 帮助他们了解自己周围都有谁;

按照关键词匹配的总次数排列所搜索到的包含关键字的 HTML 文件或者文本文件, 其中关键字可以为字母、汉字、标点或者由这些符号所组成的字符串; 其中每个关键字之间用空格间隔; 其实现步骤如下所示:

- ✧ 搜索给定文件
- ✧ 分别与各个关键字匹配
- ✧ 每匹配成功一次匹配次数加 1
- ✧ 把每个文件每个关键字匹配字数相加
- ✧ 次数最多的为最佳匹配

按照关键词权值排列所搜索到的包含关键字的 HTML 文件或者文本文件, 其中关键字可以为字母、汉字、标点或者由这些符号所组成的字符串; 其中每个关键字之间用空格间隔; 其实现步骤如下所示:

- ✧ 为第一个关键字赋最高权值, 以后的关键字权值依次递减
- ✧ 搜索文件, 各个关键字和给定的文件进行匹配
- ✧ 每匹配成功一次, 加相应关键字的权值
- ✧ 把每个文件所有关键字匹配的权值相加
- ✧ 权值最高的为最佳匹配

按照关键词覆盖率排列所搜索到的包含关键字的 HTML 文件或者文本文件, 关键字覆盖率为所输入的关键字在所搜索到的文档里面是否全部出现, 其中关键字可以为字母、汉字、标点或者由这些符号所组成的字符串; 其中每个关键字之间用空格间隔; 其实现步骤如下所示:

- ✧ 搜索文件, 各个关键字和给定的文件进行匹配

- ✧ 当文件每成功匹配一个关键字，覆盖范围加一
- ✧ 把每个文件覆盖范围进行比较
- ✧ 覆盖率最高的为最佳匹配

为了更好的体现用户交互，得到用户需要的结果，系统支持将通过按照关键词匹配的总次数排列所搜索到的包含关键字的 HTML 文件或者文本文件、按照关键词权值排列所搜索到的包含关键字的 HTML 文件或者文本文件以及按照关键词覆盖率排列所搜索到的包含关键字的 HTML 文件或者文本文件所得到的结果进行手动投票，并且显示各个部分票数；

将投票部分得票最高的部分显示；

投票

该模块通过摄像头获取包含人脸的视频，然后通过人脸检测检测出含有人脸的图片，再对含有人脸的图片进行处理，对系统进行训练，当下次碰到相同的人脸时系统自动匹配人脸，并且输出该人脸图像所对应的人名的语音信息，从而将人脸这一视觉信息转换为语音信息以供盲人用户所使用，解决他们不能自主获取人脸这一视觉信息的问题，帮助他们了解自己周围都有谁；

按照关键词匹配的总次数排列所搜索到的包含关键字的 HTML 文件或者文本文件，其中关键字可以为字母、汉字、标点或者由这些符号所组成的字符串；其中每个关键字之间用空格间隔；其实现步骤如下所示：

- ✧ 搜索给定文件
- ✧ 分别与各个关键字匹配
- ✧ 每匹配成功一次匹配次数加 1
- ✧ 把每个文件每个关键字匹配字数相加
- ✧ 次数最多的为最佳匹配

按照关键词权值排列所搜索到的包含关键字的 HTML 文件或者文本文件，其中关键字可以为字母、汉字、标点或者由这些符号所组成的

字符串；其中每个关键字之间用空格间隔；其实现步骤如下所示：

- ✧ 为第一个关键字赋最高权值，以后的关键字权值依次递减
- ✧ 搜索文件，各个关键字和给定的文件进行匹配
- ✧ 每匹配成功一次，加相应关键字的权值
- ✧ 把每个文件所有关键字匹配的权值相加
- ✧ 权值最高的为最佳匹配

按照关键词覆盖率排列所搜索到的包含关键字的 HTML 文件或者文本文件，关键字覆盖率为所输入的关键字在所搜索到的文档里面是否全部出现，其中关键字可以为字母、汉字、标点或者由这些符号所组成的字符串；其中每个关键字之间用空格间隔；其实现步骤如下所示：

- ✧ 搜索文件，各个关键字和给定的文件进行匹配
- ✧ 当文件每成功匹配一个关键字，覆盖范围加一
- ✧ 把每个文件覆盖范围进行比较
- ✧ 覆盖率最高的为最佳匹配

为了更好的体现用户交互，得到用户需要的结果，系统支持将通过按照关键词匹配的总次数排列所搜索到的包含关键字的 HTML 文件或者文本文件、按照关键词权值排列所搜索到的包含关键字的 HTML 文件或者文本文件以及按照关键词覆盖率排列所搜索到的包含关键字的 HTML 文件或者文本文件所得到的结果进行手动投票，并且显示各个部分票数；

将投票部分得票最高的部分显示；

显示投票最高的部分

该模块通过摄像头获取包含人脸的视频，然后通过人脸检测检测出含有人脸的图片，再对含有人脸的图片进行处理，对系统进行训练，当下次碰到相同的人脸时系统自动匹配人脸，并且输出该人脸图像所对应的人名的语音信息，从而将人脸这一视觉信息转换为语音信息以供盲人用户所使用，解决他们不能自主获取人脸这一视觉信息的问题，帮助他们了解自己周围都有谁；

按照关键词匹配的总次数排列所搜索到的包含关键字的 HTML 文件或者文本文件, 其中关键字可以为字母、汉字、标点或者由这些符号所组成的字符串; 其中每个关键字之间用空格间隔; 其实现步骤如下所示:

- ✧ 搜索给定文件
- ✧ 分别与各个关键字匹配
- ✧ 每匹配成功一次匹配次数加 1
- ✧ 把每个文件每个关键字匹配字数相加
- ✧ 次数最多的为最佳匹配

按照关键词权值排列所搜索到的包含关键字的 HTML 文件或者文本文件, 其中关键字可以为字母、汉字、标点或者由这些符号所组成的字符串; 其中每个关键字之间用空格间隔; 其实现步骤如下所示:

- ✧ 为第一个关键字赋最高权值, 以后的关键字权值依次递减
- ✧ 搜索文件, 各个关键字和给定的文件进行匹配
- ✧ 每匹配成功一次, 加相应关键字的权值
- ✧ 把每个文件所有关键字匹配的权值相加
- ✧ 权值最高的为最佳匹配

按照关键词覆盖率排列所搜索到的包含关键字的 HTML 文件或者文本文件, 关键字覆盖率为所输入的关键字在所搜索到的文档里面是否全部出现, 其中关键字可以为字母、汉字、标点或者由这些符号所组成的字符串; 其中每个关键字之间用空格间隔; 其实现步骤如下所示:

- ✧ 搜索文件, 各个关键字和给定的文件进行匹配
- ✧ 当文件每成功匹配一个关键字, 覆盖范围加一
- ✧ 把每个文件覆盖范围进行比较
- ✧ 覆盖率最高的为最佳匹配

为了更好的体现用户交互, 得到用户需要的结果, 系统支持将通过按照关键词匹配的总次数排列所搜索到的包含关键字的 HTML 文件或者文本文件、按照关键词权值排列所搜索到的包含关键字的 HTML 文件或者文本文件以及按照关键词覆盖率排列所搜索到的包含关键字的 HTML 文件或者文本文件所得到的结果进行手动投票, 并且显示

各个部分票数;

将投票部分得票最高的部分显示;

2. 性能需求

数据精确度

系统为数据搜索系统, 故对数据的精确度要求较高, 无论是按照关键词匹配的总次数排列所搜索到的包含关键字的 HTML 文件或者文本文件、按照关键词权值排列所搜索到的包含关键字的 HTML 文件或者文本文件、按照关键词覆盖率排列所搜索到的包含关键字的 HTML 文件或者文本文件、投票、打开得票最高的模块都必须保证满足 100% 的数据精确性;

时间特性

由于系统为实时数据搜索, 且使用频度较高, 对响应时间的要求较高, 各部分时间要求如下所示:

- 按照关键词匹配的总次数排列: 不超过 1 秒;
- 按照关键词权值排列: 不超过 1 秒;
- 按照关键词权值排列: 不超过 1 秒;
- 投票: 不超过 0.2 秒;
- 打开得票最高的模块: 不超过 0.2 秒;
- 平均故障间隔时间: 平均故障间隔时间不低于 50 小时。

适应性

➤ 操作方式

和普通的基于 Windows 平台的应用系统本系统采用鼠标和键盘方式操作本系统;

➤ 运行环境

本系统为基于 Windows 平台的应用系统, 本系统先在电脑端上运行, 其电脑端运行环境如下所示:

操作系统: Microsoft Windows XP / vista / 7

支持环境: MFC 运行时链接库

➤ 与其它软件的接口

本系统为独立系统, 不涉及到与其他软件的协作, 故没有与其他软件的接口问题 ;

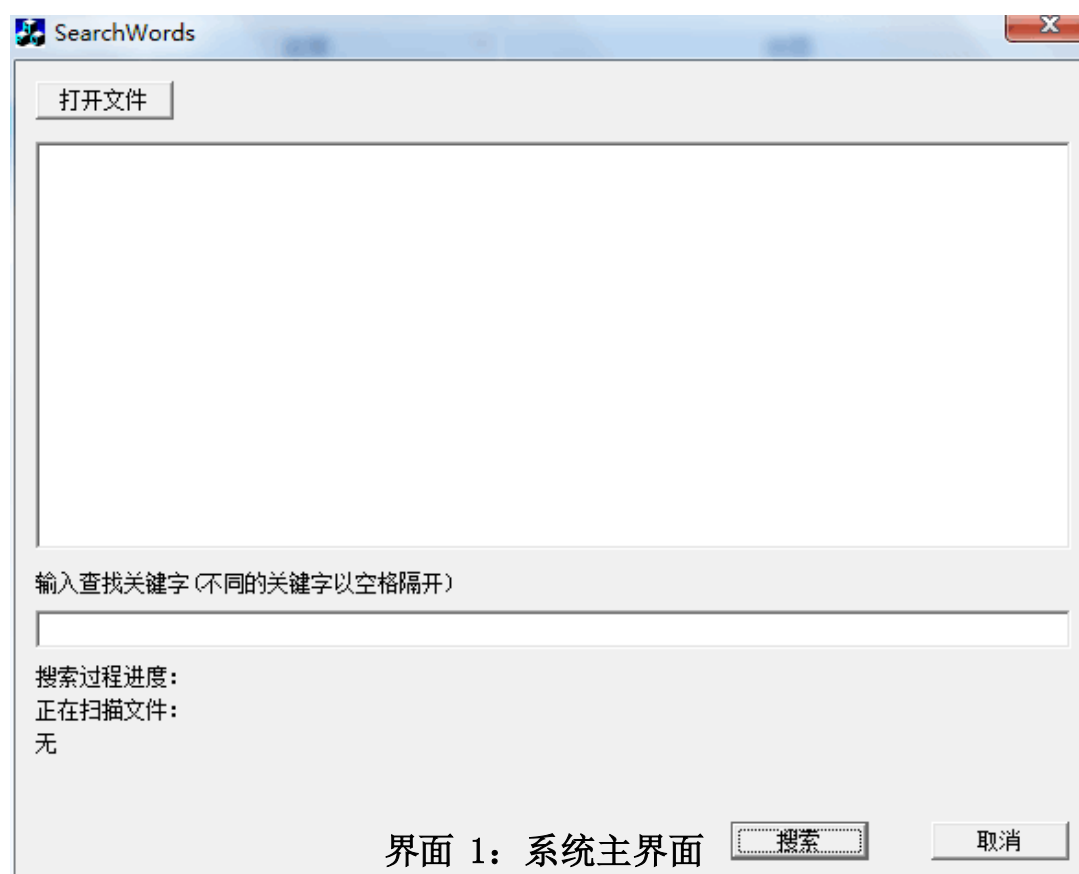
➤ 开发计划等发生变化时, 应具有的适应能力

本系统软件部分采用演化过程模型, 将系统分为按照关键词匹配的总次数排列、按照关键词权值排列、按照关键词权值排列以及投票并且打开得票最高的模块, 项目开发有足够的灵活性, 故当开发计划发生变化、其中某个模块的开发出现一定的延期时, 可以将它对其他模块的影响降低到最低;

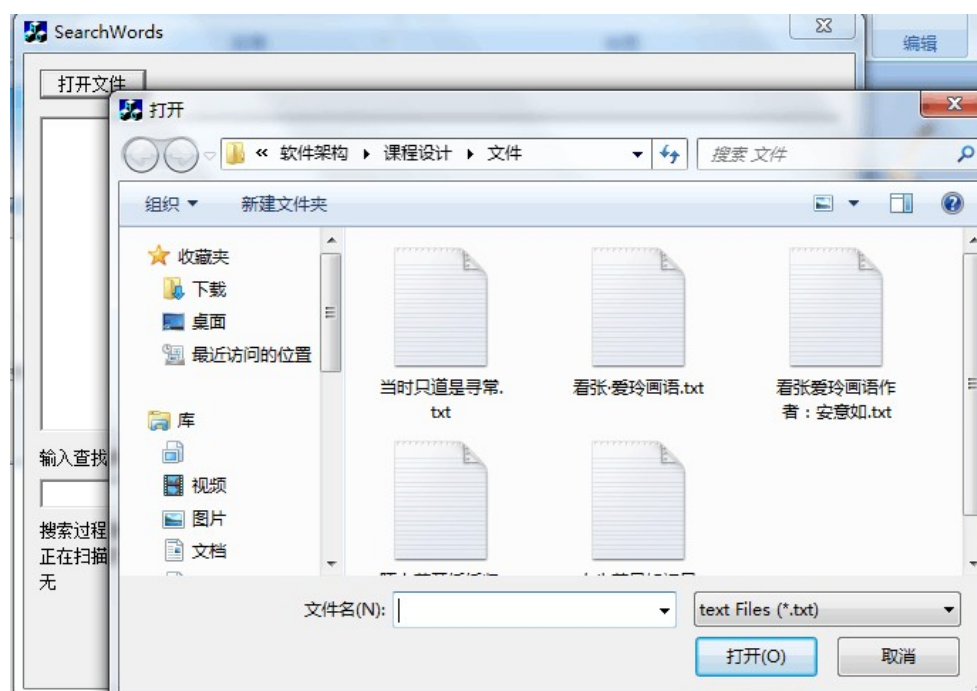
3. 运行需求

用户界面

在实现电脑端软件系统时, 界面应包含一下部分



其中，点击“打开文件”按钮选择所要搜索的目标文件，下面的文本框显示所选择的目标文件，第二个文本框中输入所要搜索的关键字，各个关键字之间用空格隔开；下面则分别显示“搜索过程进度”以及“正在扫描文件”，点击“搜索”和“取消”按钮则可以相应的



界面 2: 打开目标文件

以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。如要下载或阅读全文，请访问：<https://d.book118.com/837144162045006133>

