

基于 WEB 环境与 MATLAB 技术的图像检索系统的实现

The implementation of image retrieval system based on web environment and matlab technology

贾 伟 (太原山西财经大学信息与管理学院 030006)

王小玲 (山西财经大学信息与管理学院、上海交通大学计算机工程系 200030)

摘要: 计算机网络尤其是 Web 的广泛应用, 已使其成为数字图像的重要资源提供者与最理想的人机界面。因此, 基于 WEB 环境的系统结构设计应当成为图像检索系统的首选。本文对基于 WEB 环境与 MATLAB 技术的图像检索系统的设计与实现过程进行了探讨, 给出了基于 WEB 环境构建图像检索系统的具体实现过程。

关键词: Web 环境 MatLab Web Server 图像检索

1 引言

近年来随着计算机网络的不断发展与数字图像的大量应用, 人们迫切需要一种从海量图像数据库中迅速寻找所需图像的方法, 这就是图像检索技术, 它已经成为数据库、图像处理与计算机视觉的重要研究领域之一。图像检索技术简言之就是利用图像的物理属性诸如色彩、纹理、形状等或者与之关联的描述关键词 (Key Words) 在图像数据库中寻找与查询图像相关或者相似的一组 (个) 图像^[1]。

计算机网络尤其是 Web 的广泛应用, 已使其成为数字图像的重要资源提供者与最理想的人机界面。因此基于 WEB 环境的系统结构设计应该成为图像检索系统的首选^[2]。

图像检索算法的仿真试验目前大多数通过 MATLAB 软件包实现。MATLAB 是由美国 MathWorks 公司推出的计算机软件, 是国际公认的最优秀的科学计算与数学应用软件之一。Matlab 的核心是一个基于矩阵运算的快速解释程序, 它提供功能强大的面向各专业领域的工具箱, 包括图像处理工具箱。由于其语法简单, 界面友好, 数值计算高效, 已经成为图像处理的主流软件之一^[3]。

因此, 一个高效的图像检索系统应该在 Web 环境嵌入 MATLAB 的图像处理功能, 如图 1 所示。

用户通过浏览器向服务器发送查询请求, 包括提供描述图像的关键词 (Key Word) 与示例图像 (Example Image) 两种情形。在服务器端, 对关键词型查询, 直接采用关键词在图像数据库中进行检索。对示例图像查询, 需要采用图像检

索算法, 根据物理特征, 与图像数据库中的图像进行相似性匹配, 将检索结果返回到客户端, 显示给用户, 由用户决定是

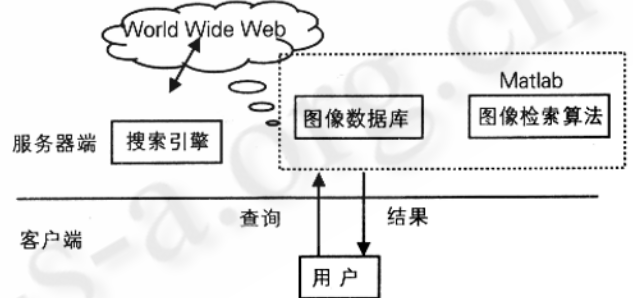


图 1 基于 Web 环境的图像检索系统结构框架

否满意, 如果不满意, 可以重新选择关键词或者新的查询图像作为起点, 向服务器再次提交查询请求。其中, 搜索引擎采用经典的 Crawler 实现, Crawler 持续不断地在 Internet 上搜索图像, 对图像数据库内容进行更新、维护, 保证图像 URL (United Resource Link) 地址链接是有效的。由于关键词查询属于典型的基于 Web 的数据库应用, 不是本文讨论的重点。本文侧重需要用到图像检索功能部分的实现技术。

2 MatLab Web Server 与应用

在图像检索系统的服务器端, 除使用常规 Web Server, 建立基于 HTTP 传输协议的客户端与服务器端的通信以外, 还需要 Matlab 提供的 MatLab Web Server 模块以实

现 Matlab 程序与 Web 浏览器的接口,并以 HTML 格式显示图像检索结果。Matlab Web Server 是一个多线程 TCP/IP 服务器,负责管理 Web 应用与 Matlab 之间的通信,实现以 Web 方式访问 Matlab 应用程序。它调用的 Matlab 程序(M 文件)由 HTML 文档的特定格式字段指定。Matlab 的 M 文件是一种解释性的程序,扩展名为 .m,由此得名 M 文件。它有两种形式,一种是批处理命令形式,一种是函数形式,MatLab Web Server 要求函数形式的 M 文件。

Matweb 是 MatLab Web Server 的 TCP/IP 客户端程序,使用通用网关接口(Common Gateway Interface CGI)从 HTML 文档提取数据并传输至 MatLab Web Server。Matlab Web Server 则运行 M 文件来产生响应。具体的过程是 MatLab Web Server 首先激活 matweb.m 也就是运行在 Matlab 环境下的 M 文件,由 matweb.m 读取 matweb.conf 配置文件中的信息,定位所要运行的 MatLab 程序(M 文件)以及 MatLab Web Server 地址。

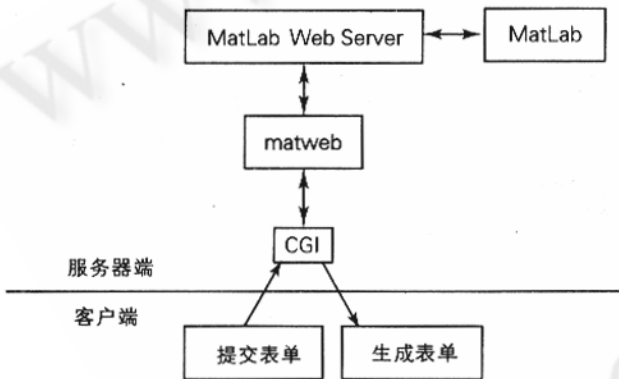


图 2 MATLAB Web Server 应用程序结构

一个 matweb.conf 配置文件示例如下:

```

%[M 文件名称]
[application1]
%变量 = 值
mlserver= parrot
%另一个应用
[application2]
mlserver= parrot
mkdir= /matlab /toolbox /webserver /wsdemos
    
```

其中,可选配置参数的含义如下:

表 1 matweb.conf 配置参数

[application]	运行的 MATLAB 应用文件名(必须)
mkdir(可选)	读写文件的工作目录,自动加入到 MATLAB 路径下
mllog(可选)	应用的日志文件,记录应用与 MATLAB 的交换数据
mlserver(必须)	运行 MatLab Web Server 的主机
mlport(可选)	打开 matlabserver 侦听端口,默认端口号 8888
mltimeout(可选)	最大服务器响应等待时间,默认值 180 秒
my_varUser(可选)	用户自定义变量名称

3 应用开发实现

MatLab Web 应用程序的一般开发步骤如下:

(1) 创建接受用户输入数据的 HTML 文档,可以使用任何文本编辑器或者常用的网页制作工具如 Microsoft 公司的 Front Page, Multimedia 公司的 Dream Weaver 等直接编辑;

(2) 在 matweb.conf 文件中写入应用名称及相关联的配置数据;

(3) 编写 MATLAB M 文件:matweb.m,主要功能是从 HTML 输入表单接受数据,进行运算或产生任何 MATLAB 图形文件;

(4) 可以调用 htmlrep 将输出数据嵌入到一个 HTML 输出文档模板中。HTML 文档从 MATLAB 中可以接受的数据的最大容量是 256 KB。

MATLAB Web Server 安装好以后,matweb 的缺省位置在 <matlab> /webserver /bin /。对 HTTP server,必需在 /cgi - bin 下放置该文件的一个副本。matweb.conf 文件也需要放置到 /cgi - bin 目录下。

本文以基于直方图的图像检索算法在 Web 环境中的设计为例,探讨设计与实现细节。需要移植到 Web 环境下的 Matlab 程序(matweb.m)是 histogram_retrieval.m,要改造成函数形式。该函数的输入变量 query_file(查询图像名称)来自 retrieval_input.html 文件,是运行在客户端的表单程序。在 Matlab Web Server 端计算得出与查询图像相似的图像后,把图像文件名称变量 SimilarFile_Name 作为返回参数,在 retrieval_input.html 文件中输出,下载到客户端,形成最终结果。

① retrieval_input.html (输入文件)

```
<! -- 根据 matweb 客户端运行平台的不同(NT 或
Unix),激活 matweb,建立与 MATLAB Web Server 的通
信。-->
```

```
<! -- NT version: -->
<form action = "/cgi - bin /matweb. exe" method
= "POST">
```

```
<! -- Unix version: -->
<form action = "/cgi - bin /matweb" method = "
POST">
```

```
<! 建立名为 mimfile 的隐藏型字段,其值指定了
Matlab Web Server 要运行的 Matlab 程序,需要和 mat-
web. conf 中的配置保持一致,在本应用中调用的是 histo-
gram_retrieval 函数。-->
```

```
<td width = "20%"> <img border = "0" src = "
$SimilarFile_Name$" width = "100" height = "64">
</td>
</tr>
</table>
```

还可以使用 Matlab 提供的 htmlrep 函数,根据需要,构造包括表格与下拉列表格式的输出文档。

```
③ matweb. conf (连接配置文件)
%定位所要运行的 MATLAB 程序(M 文件)
[histogram_retrieval]
%定位 MatLab Web Server 地址
mlserver=mls
```

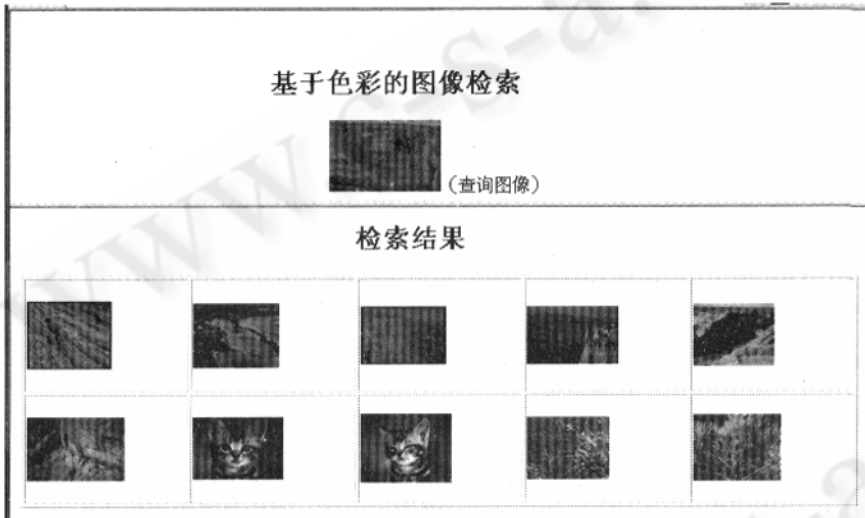


图 3

```
<input type = "hidden" name = "mimfile" value
= "histogram_retrieval">
```

```
<! 客户端输入变量:查询图像名称-->
<input type = "text" name = "query_file">
<! 一向服务器端提交请求-->
<p><input type = "submit" name = "Submit"
value = "Submit"></p>
```

② retrieval_output. html (输出文件)

以表格形式输出和查询图像最相似的前 10 幅。相关 HTML 文档内容如下:

```
<! --SimilarFile_Name 是相似文件名变量,是函数
histogram_retrieval. m 的返回值-->
<table border = "1" width = "100%">
<tr>
```

4 结论

本文给出了基于 WEB 环境构建图像检索系统的具体实现过程。借助 MatLab Web Server 构建基于 Web 的图像检索应用的基本原理是相似的,只要将设计好的图像检索算法改造为函数形式,在输入输出文档中定义好该函数的相应参数即可,其他类似应用如动态视频图像的检索等均可以套用本文提出的流程设计。

参考文献

- 1 Smeulders, A.W.M. Worring, M. Santini, S. Gupta, A. Jain, R. "Content - based image retrieval at the end of the early years," Pattern Analysis and Machine Intelligence, IEEE Transactions on , Vol.22 ,No. 12 , Dec 2000,pp.1349 - 1380 .
- 2 I. Kompatsiaris, E. Triantafyllou and MG Strintzis, "A World Wide Web Region - Based Image Search Engine", 2002 Tyrrhenian International Workshop on Digital Communications (IWDC 2002), Capri, Italy, September 8-11, 2002.
- 3 孙兆林, MATLAB 6. X 图像处理,清华大学出版社,2002 年 5 月。