

基于 Qrobot 机器人的彩票服务软件系统^①

郭馨, 高金行, 武勋

(西北工业大学 自动化学院, 西安 710129)

摘要: 在 Qrobot 机器人已有的软件系统基础上, 开发出了独立的彩票服务程序. 在标准化的 Windows 操作环境下, 以对象类别扩充组件的方式提供独立的, 便捷的人机交互界面. 通过触控和声控技术与机器人互动, 使 Qrobot 机器人获知所需查询的彩票类别. 利用 Web 信息提取技术从指定网页中提取出彩票的中奖号码等相关信息, 利用云端语音识别使机器人读出网页提取出的彩票信息, 并能将中奖号码信息显示在操作界面上. 所开发的软件操作系统具有良好的扩展性, 基于同样的思想, 还可以进一步为机器人添加其它网络功能.

关键词: Qrobot; 人机交互界面; Web 信息提取; 云端语音识别

Software System of Lottery Service Based on Qrobot

GUO Xin, GAO Jin-Xing, WU Xun

(School of Automation, Northwestern Polytechnical University, Xi'an 710129, China)

Abstract: Based on the original software system of Qrobot, an independent lottery service program is developed in this paper. Under the standard Windows operating system environment, an independent and convenient interactive user interface is provided in form of object linking and embedding control extension. By interacting with Qrobot by means of touch and voice-activated technology, Qrobot is informed the variety of lottery which is needed to query. The winning numbers and other related information, which are extracted from the specified site by utilizing web information extraction technology, can be read out by Qrobot by way of cloud speech recognition and displayed on the interface. Furthermore, this software operating system has good expansibility, network function can be added to Qrobot based on the same idea.

Key words: Qrobot; interactive user interface; web information extraction; cloud speech recognition

机器人技术是综合了计算机、控制论、机构学、信息和传感技术、人工智能、仿生学等多学科而形成的高新技术, 是当代研究十分活跃, 应用日益广泛的领域. 机器人应用情况, 是一个国家工业自动化水平的重要标志. 从机器人诞生到 20 世纪 80 年代初, 机器人技术经历了一个长期缓慢的发展过程. 到了 90 年代, 随着计算机技术、微电子技术、网络技术等的快速发展, 机器人技术也得到了飞速发展, 智能网络机器人应运而生^[1]. 智能网络机器人是机器人技术与互联网技术的融合技术, 是一种能够实现单体机器人无法完成的服务机器人. 智能机器人的出现为用户通过互联网查找信息提供了一种新的方式, 用户可以使用日常

的语言习惯与机器人进行信息的交互式查询. 相比搜索引擎来说, 智能网络机器人的查找方式更接近我们的日常交流^[2-6].

Qrobot 是智能互联网机器人产品, 是由腾讯公司、中科院深圳先进技术研究院和深圳市中科睿成智能科技有限公司联合研发, 用户可通过语音指令等多维交互方式, 与机器人交流. Qrobot 机器人, 软件灵魂在电脑中运作, 机器人借助主人的电脑计算能力, 将主人的电脑升级为一个机器人, 让主人能够通过与人机交互的立体形象交互互动获取信息. Qrobot 机器人集成语音声控、触控传感、计算机视觉、情感计算等技术, 在前端运用对话、手势、触控等人机交互方式, 通

^① 收稿时间:2013-04-12;收到修改稿时间:2013-05-20

过云服务在后台为用户提供教育、娱乐、资讯、社交等互动应用。Qrobot 实物图如图 1。



图 1 Qrobot 实物图

Qrobot 软件开放平台, 基于 COM 构建。在该平台下, 每个应用都是一个独立的 ActiveX 控件, 应用之间相对独立, 模块之间可以通过相应的软件进口进行自由切换。开发人员使用自己熟悉的编程语言编写 ActiveX 控件, 进行简单配置后, 即可通过 Qrobot 管理软件方便的加载该应用模块, 来实现各种有趣的功能。本彩票服务系统是在已有的 Qrobot 管理软件基础上, 开发出的新应用。

1 系统概述

用户通过触控和声控技术与 Qrobot 互动, 使机器人获知所需查询的彩票种类。利用 Web 信息提取技术提取出彩票中奖号码等相关信息, 通过云端语音识别技术使机器人向用户播报查询结果, 并将查询结果显示在操作界面上, 很大程度提高了用户使用机器人的趣味性。利用同样的思想, 可以为 Qrobot 机器人开发出更多网络功能^[7,8]。

本系统总体可以分为三部分: 彩票信息查询和提取模块、彩票信息播报模块、彩票信息显示模块。系统总体框图如图 2 所示。

彩票查询和提取模块: 通过触控和声控, 可提供多种彩票的查询服务供用户选择, 如: 双色球、大乐透、七乐彩、七星彩、福彩 3D、排列三、排列五。通过对指定网页源码进行处理, 提取出必要信息以供上层调用。

彩票信息播报模块: 本模块使用 QQ 语音云端提供语音合成、语音搜索等。能够很好的突破本地资源的限制, 使非特定人、不限文本、任意时间的语音合成成为可能。

彩票信息显示模块: 本模块利用了 button 控件来显示文字, 给用户一种立体的 3D 感, 并且让每个控件上所显示的内容符合实际的开奖结果的球的颜色, 给用户直观的感受。

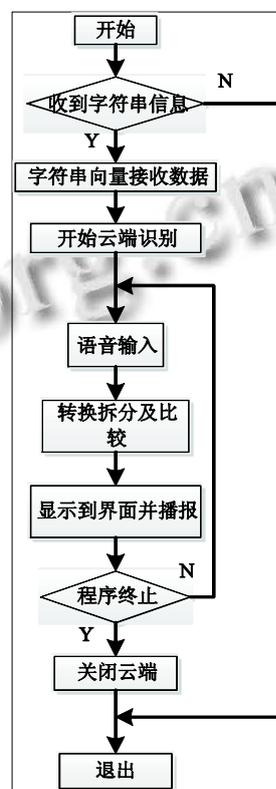


图 2 系统总体框图

2 各模块的设计与实现

2.1 彩票信息查询和提取

如今, Web 网页中绝大部分还是由 HTML 来描述的, HTML 没有严格的语法要求, 而人们为了追求页面的华丽以及商业价值, 加入了各种脚本 (VBScript/JavaScript), 广告链接, flash 动画, 使 HTML 代码过于冗长, 混乱, 这些噪声的存在使得人们很难快速的找到所要查找的信息。因此, 为了方便用户能快速在指定网页中查找有用信息, 需要对指定网页进行信息提取^[9]。

HTML 文档由标记 (TAG) 和元素组成。HTML 标记确定了浏览器所显示文档元素的格式, 大多数 HTML 标记是成对出现的, 它们分别用作开始标记和结束标记, HTML 的结束标记与开始标记的唯一区别是多了个斜杠 “/”, 如: <HTML></HTML>、<HEAD></HEAD>、<TITLE></TITLE> 等。HTML 文档由标题 <HEAD> 和主体 <BODY> 两部分组成。

在网页信息提取和噪声抑制的方法中,通常要将 HTML 格式解析成 DOM 文档.但是,在解析的过程中,由于 HTML 语言本身的特点,各个网页的 HTML 格式存在很大差别,解析的过程需要对这些格式进行统一的语法检查、修改处理,影响了系统的响应时间.因此,本文并没有采用此解析过程进行信息提取,而是分析了 HTML 文档结构,直接解析文本,对文本信息进行信息采集^[10].

此模块中,读取指定的网页源码实现的方法是向网页服务器发送 HTTP—GET 请求:ret = send(sock, request, strlen(request), 0);接收到返回信息后,将网页源码写入到 WebMesGrabber_html_orisou.txt,之后对该文件中的字符串进行必要的切分和处理,如获取双色球的号码时,需要根据其在源码中的 ID ("/static/kaijiang/ssq/")来实现初期信息的切分和获取.通过进一步的处理,将所有的中奖号码等信息放入到容器 vec_lucnum(Vector 类型)中,最后将经过处理好的所有彩票中奖号码的结果写入到 WebMesGrabber_logmes.txt 中供上层调用.其各部分具体实现过程如图 3 所示.

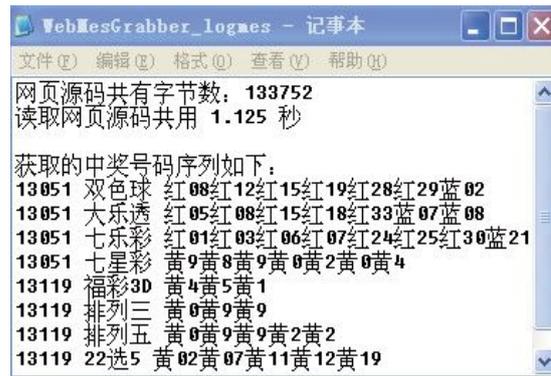
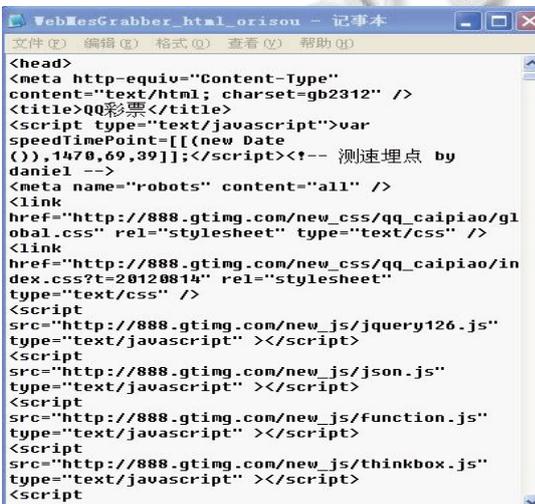


图 3 网页信息提取各部分过程图

该技术的实现类名为 WebMesGrabber. 获取字符串中的某个 td 中包含的内容,第一个参数是等待切分的字符串,第二个参数是 td 的 id. 获取当期彩票开奖期号,格式如: 12124, 例: <td class="gary">12124</td>, 从中取得期号 12124 的代码如下:

```
int WebMesGrabber::get_issue(string in,string
&issue_number)
{
    string index="<td class=\"gary\">";
    string td;
    int i=get_subtd(in, index, td);
    if(i!=0)
    {
        WebMesGrabber_logmes<<"failed
get_issue::get_subtd"<<endl;
        return 1;
    }
    int pos_begin, pos_end;
    string str_a="target=\"_blank\">";
    pos_begin=td.find(str_a);
    if(pos_begin== -1)
    {
        WebMesGrabber_logmes<<"failed
pos_begin:get_issue::"<<str_a<<endl;
        return 1;
    }
    string tem;
    tem=td.substr(pos_begin+str_a.length());
    pos_end=tem.find("</a>");
```



```

if(pos_end===-1)
{
    WebMesGrabber_logmes<<"failed
    pos_end:get_issue::</a>"<<endl;
    return 1;
}
issue_number=tem.substr(0,pos_end);
return 0;
}
    
```

2.2 彩票信息播报

多年来,让机器人能听懂人话的语音识别技术一直备受关注.过去的语音识别技术是通过将输入的语音特征与设置好的声学模型进行匹配与比较,而后得到最佳的识别结果.由于处理速度和存储容量的限制,这样的语音识别通常有固定的句式要求,如此一来,识别性能弱、交互固定,并且需要用户进行固定句式的学习与记忆,使用便捷性较差.如今,语音识别的实现,则是通过本地和云端相结合的方式进行,不仅加快了识别速度,其“自然语言理解”功能更是填补了过去语音识别技术的缺陷.在整个过程中,用户也许会出现略微等待的情况.由于这种语音识别运用到了云端技术,而云端识别几乎不花时间,因此这个等待的时间可以被认为是网络的传送时间,也就是说,网速有多快,语音被识别的速度就有多快.以数据为中心的云架构,把生产者和消费者隔开在两端,一个是在云端,一个是在用户的使用端,生产者的服务提供和消费者的服务使用构成了云的架构^[11,12].如图 4.

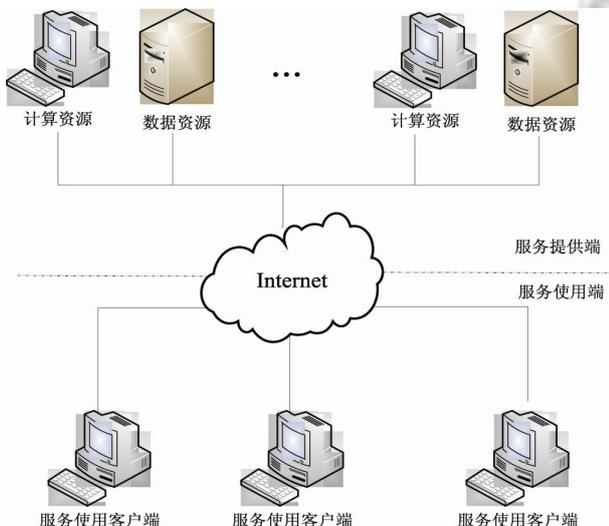


图 4 云的架构

在此模块中,由于本地语音库的限制,不能完成预定播报彩票中奖号码信息的功能.所以使用软件平台所提供的接口开启 QQ 云端语音识别.该语音基于云计算开发,所有的语音操作都是基于云计算而实现的,包括语义解译、匹配搜索等.云端语音操作完成以后,再反馈给终端设备,用户根据所得到的反馈做相应的操作.接口为 IQrobotTalk 类所提供的以下函数: SetVoiceCloudDelay(timeCloudDelay), 该函数用于设置云语言识别延迟时间, OpenVoiceCloud();开启云端语音,当识别完毕后使用 IQrobotTalk 类所提供的另一个接口函数关闭云语音: CloseVoiceCloud().

2.3 彩票信息显示

由于自定义信息抓取类所使用接口类型与平台所使用类型的差异,需要将抓取到的字符串做一定的转换以满足平的数据类型.使用 MultiByteToWideChar(CP_ACP,0,sMesGrab.c_str(), sMesGrab.length(), w_strTemp, sMesGrab.length());将自定义类中的 string 类型转换为宽字节 wstring 类型.所提取字符串中包含了彩票种类,彩票的期号,以及开奖的结果.所以将这三种信息分别拆分到 strResult, w_strFinal 的 w_char*类型的变量作相应的显示.而彩票的种类是由语音输入到 strResult 中的.过程如图 5.

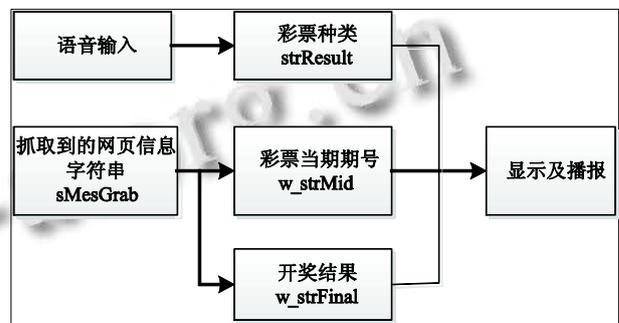


图 5 字符串的转化与拆分过程

为了使显示效果更美观,采用了 button 控件来显示文字.在界面上可以清楚地显示出获奖号码的颜色、彩票开奖的当期期号、彩票类型和中奖号码,给用户直观的感受.利用了 CButtonST 类所提供的 SetColor()这个接口来根据彩票的种类控制颜色的变化.

在用户通过声控向 Qrobot 机器人说出所需查询的彩票,如:双色球、七乐彩、福彩 3D、大乐透、排列三、排列五,于 Qrobot 操作界面中,显示结果如下,以双色球、排列五为例,如图 6 和图 7 所示,用户可清楚

地看到查询结果,同时 Qrobot 机器人会通过云端,语音播报出来开奖结果。



图6 Qrobot 双色球查询结果界面



图7 Qrobot 排列五查询结果界面

3 结语

本系统基于 Qrobot 机器人平台,完成了彩票中奖号码查询、播报、显示的功能,程序设计中用到了网页信息提取技术、云端语音识别技术等,实践证明整个程序运行状态良好,很大程度上提高了用户查询彩票的乐趣。此外,如果将来此项功能进一步扩展,还能增加随机选号、预测功能。基于同种技术,还能为 Qrobot 机器人开发其他有趣的功能。

参考文献

- 1 刘树平.网络机器人基于 Internet 视频传输.南京:南京理工大学,2003.
- 2 张伟.网络机器人研究与发展分析.机器人技术与应用,2010,(1):23-27.
- 3 树子.智能网络机器人的春天即将到来.互联网天地,2008,(8):49.
- 4 于磊,潘郁.智能学习型网络机器人.计算机工程,2004,30(13):135-137.
- 5 刘振宝,辛洪兵,王文静.网控机器人技术及其控制系统的研究状况.北京工商大学学报,2006,24(3):32-36.
- 6 庄严,王伟,恽为民.基于网络的机器人控制技术研究现状与发展.机器人,2002,24(3):277-281.
- 7 张珍.智能机器人语音识别技术.现代电子技术,2011,34(12):57-60.
- 8 刘旸.面向机器人对话的语音识别关键技术的研究.西安:西安电子科技大学,2009.
- 9 袁宇丽.基于 HTML 网页的 Web 信息提取研究.成都:电子科技大学,2006.
- 10 张恒,屈景辉,张亮.网页文本信息提取及结果评价.微计算机应用,2007,28(9):922-924.
- 11 罗希.基于云计算的语音输入方案研究.上海:华东师范大学,2009.
- 12 张体.语音云及其在声控导航中的应用.中国科技信息,2012,12:118.