

基于 Web 网络的身份识别系统的设计与实现

陈文庆

(中国劳动关系学院 计算机应用技术教研室, 北京 100038)

摘 要: 当前身份识别系统大多采用密码识别技术, 容易被窃取或遗失, 导致用户有一定的损失, 识别精度较低. 为此, 设计了一种新的基于 web 网络的身份识别系统, 主要由浏览器端录音模块、声纹数据传输模块和声纹训练及识别模块构成. 给出系统架构, 依据系统架构在浏览器端安装一个 Activex 录音控件, 对用户的语音进行录制. 通过预加重、端点检测等方式对声纹信号进行预处理. 对声纹信号进行特征值提取, 通过信号转换过滤信号中没有用的部分, 提取出能够体现用户录制语音本质属性的特征参量, 依据提取特征值得到 D 值, 通过 D 值实现身份识别. 实验结果表明, 所设计系统有很高的识别精度.

关键词: web 网络; 身份识别; 系统; 设计

中图分类号: TP301.6

文献标识码: A

文章编号: 1000-7180(2017)05-0114-05

Identification System Based on Web Network Design and Implementation

CHEN Wen-qing

(Application Research of Computer Room, China Institute of Industrial Relations, Beijing 100038, China)

Abstract: The current identification system mostly USES the password identification technology, easy to be stolen or lost, leading to loss of users have a certain, the identification accuracy is low. To this end, we design a new network identification system based on web, the designed system is mainly composed of browser side recording module, voice print data transmission module and voiceprint training and recognition modules, system architecture is given. Based on the system architecture in the browser to install an Activex controls, recording on the user's voice recording. Through the pre-emphasis, endpoint detection methods such as the voice print signal preprocessing. Characteristic value of voiceprint signals are extracted through the signal conversion to filter the useless part of the signal, to extract the characteristic parameters of essential attributes can reflect user recording voice, according to the feature extraction is worth to D value, through the D value identification. The experimental results show that the designed system has a high recognition accuracy.

Key words: web network; identity; system; design

1 引言

随着科技的逐渐发展和计算机的广泛应用, 将数字化网络作为平台的应用与服务逐渐增加, 大大提高了系统运行效率, 但导致管理方面出现问题^[1-3]. 用户要使用各种应用, 需登录相应的应用系统, 所有应用均存在自身的账户系统^[4-5]. 且随着应用的逐渐增加, 系统出错率也越来越高, 严重威胁用

户安全^[6]. 因此, 研究一种有效的身份识别系统具有重要意义^[7].

文献[8]提出一种基于口令的身份识别方法, 该方法要求用户在登录时输入用户名与密码, 提交至系统对用户身份进行识别与验证, 系统将提交的信息和数据库对比, 对用户身份进行识别. 该方法实现过程简单, 但安全性较差; 文献[9]提出一种基于智能卡技术的身份识别系统, 利用内置 CPU 与存储

器完成数据保存、计算和加密处理,通过硬件实现身份识别.因为用户很难伪造内置 CPU 中的数据,所以安全性较高,但该方法成本高,不能被广泛应用;文献[10]提出一种身份识别组件模式对身份进行识别,将独立的身份识别组件模块置于应用系统中,该模块为系统提供用户登录界面,对用户输入的口令进行验证,其应用系统自身没有用户系统,所以输入的用户账号一定为真正的用户信息.该方法适用于新建的应用系统的用户身份识别,但识别效率低.

针对上述方法的弊端,设计了一种新的基于 Web 网络的身份识别系统,给出系统架构.依据系统架构在浏览器端安装一个 Activex 录音控件,对用户的语音进行录制.介绍了 Web 服务器端声纹信号识别模块的设计过程.实验结果表明,所设计系统有很高的识别精度.

2 基于 Web 网络的身份识别系统的设计与实现

2.1 系统结构

本节设计的基于 Web 网络的身份识别系统主要由浏览器端录音模块、声纹数据传输模块和声纹训练及识别模块构成.浏览器端录音模块需具有可在 Web 页面上运行程序的能力,而 ActiveX 控制控件技术能够自动下载和安装,对功能强大的软件控件进行开发,所以本节选用 ActiveX 控制控件技术.因为 ASP 可有效支持 ActiveX 控件,所以本节通过 ASP 技术对浏览器端与服务器端之间的数据传输方式进行开发.针对声纹训练和识别模块,本节首先对声纹特征进行提取,然后依据提取的特征实现声纹识别,从而实现身份识别.基于 Web 网络的身份识别系统架构如图 1 所示.

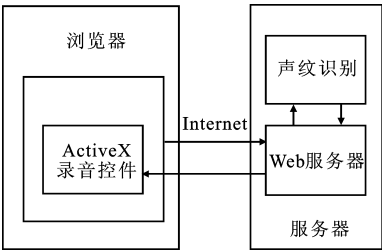


图 1 基于 Web 网络的身份识别系统架构

2.2 浏览器端录制模块设计

依据图 1 描述的基于 Web 网络的身份识别系统架构,在浏览器端安装一个 Activex 录音控件,对用户的语音进行录制. ActiveX 控制控件需具有注册能力和 IUnknown 接口,同时利用 IUnkonwn 接

口获取公开录制的声音特征数据供系统提取,同时和网页上的窗口结合在一起,把数据传输至服务器端进行相应处理. 录音控件通过低阶波型采用程序接口对录音功能进行开发. 为了可兼容大部分声卡,本节录音模块将取样频率设置为 12.136 kHz、单声道、8 位,对用户声音进行录制,并引入能量计算程序以实现自动断音功能,其能量 $En(n)$ 可通过下式进行计算:

$$En(n)=\sum_{m=0}^{N-1}S^2(n+m) \tag{1}$$

式中, S 用于描述能耗; n 代表第 n 个用户; m 用于描述用户数量.

依据 $En(n)$ 获取录制语音的倒频谱系数,公式描述如下:

$$\hat{c}_n=-\sum_{m=1}^P(1-m/n)a_mEn(n) \tag{2}$$

式中, a_m 用于描述语音频谱包络矩阵.

得到的倒频谱系数是十阶的倒频谱系数矩阵,该序列数组若一个一个对外公布,会引起存取困难,需对该序列进行格式包装后再对外公布. 字符串属于一种简单的格式,该序列的声纹特征可通过格式化 1/0 的方式连接成字符串,同时利用 IUnknown 接口对字符串进行公布,供浏览器存取,其转换方式用图 2 进行描述.

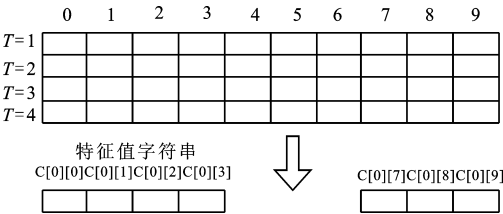


图 2 导频系数数组转换成特征字符串示意图

在浏览器端,通过 ActiveX 控制控件对用户声音进行录制,发送给 Web 服务器,通过 ASP 程序进行存取. ASP 程序采集的数据为经编码的格式,需对其进行译码转换成原来的形式.

2.3 Web 服务器端声纹识别模块设计

2.3.1 声纹信号预处理

声纹信号预处理是声纹识别的重要基础,有效的预处理能够提高声纹识别的准确性. 声纹信号预处理主要包括声纹信号预加重、端点检测等,详细过程分析如下.

录制的语音信号平均功率谱容易被声带与尾音辐射干扰,当语音在约 800 Hz 频率时,部分信号以每倍频 6 db 的速度下降,使得语音信号的高频部分

被淹没于中低频信号中,增加声纹识别误差.因此,需对语音进行预加重处理,以保证高频信号的强度.

本节通过一阶高通数字滤波器完成对高频信号的预加重处理,该滤波器一般被称作预加重滤波器,其传递函数可通过下式求出:

$$H(z) = 1 - \mu z^{-1} \quad (3)$$

式中, μ 用于描述一阶高通滤波因子,在 $[0.9, 1]$ 范围内取值; z 用于描述信号.通过上式可获取该预加重网络的输入,则输入 $x(n)$ 与输出 $y(n)$ 的一阶差分方程可通过下式求出:

$$y(n) = x(n) - \mu x(n-1) \quad (4)$$

得到的输出 $y(n)$ 即为经预加重处理后的信号.

在对声纹信号进行采集的过程中,录音的开始与结尾一定会有一段空白部分,端点检测主要是为了从一段录制声纹信号中准确获取需识别的语音段的初始点与结束点,过滤噪声信号,同时降低声纹识别时间与计算量,提高整个系统的性能.

本节通过依据短时能量和平均过零率的双门限端点检测方法进行端点检测,该方法将能量与过零率看作特征进行检测,也被称作双门限比较法,其短时平均过零率可通过下式求出:

$$Z_n = \frac{1}{2N} \sum_{m=-\infty}^{+\infty} \left| \frac{\text{sig}[x(m)] - \text{sig}[x(m-1)]}{\text{sig}[x(m-1)]} \right| * w(n) \quad (5)$$

式中, N 用于描述语音样本的帧数;

$$\text{Sig}[x(n)] = \begin{cases} 1, & x(n) > 0 \\ 0, & x(n) = 0 \\ -1, & x(n) < 0 \end{cases} \quad (6)$$

$$w(n) = \begin{cases} 1, & 0 \leq n \leq N-1 \\ 0, & \text{else} \end{cases} \quad (7)$$

通过双门限比较法进行端点检测的基本思想如下:在检测前首先根据声音信号的性能对短时能量和过零率的上下两个门限值进行确定,如果该门限值较小,则其对声纹信号的变化将很敏感,信号通常高于该数值,此刻可获取声纹信号的端点;如果该门限值较大,则该值会对声纹信号进行框定,当前噪声通常较少,同时高于数值较大的门限,所以会被识别.则利用一高一低的双门限,即可实现对声纹信号端点的检测.

2.3.2 声纹信号特征值的提取

完成对声纹信号的预处理后,本节对声纹信号进行特征值提取,以获取声纹信号特征值.声纹信号特征值提取的原则如下:通过信号转换过滤信号中没有用的部分,提取出能够体现用户录制语音本质

属性的特征参量.

通过傅里叶转换完成对加窗的语音帧的计算,该加窗的语音帧可通过下式进行计算:

$$S_w(n) = s(n) \times w(n) \quad (8)$$

式中, $s(n)$ 用于描述语音信号; $w(n)$ 用于描述语音信号和窗函数.本节对该加窗信号进行傅里叶转换,则有:

$$X(k) = \sum_{j=0}^{N-1} s_w(j) \times e^{-j \frac{2\pi i k}{N}} \quad (9)$$

式中, k 用于描述傅立叶转换的频率点数.傅里叶转换把时域输入信号转化成频域信号.通过梅尔滤波器组对声纹信号进行处理后,对其进行对数压缩操作,然后通过三点线性回归模型获取 4 组特征值,公式描述如下:

$$d_{t,f}^l = \frac{\sum_{i=1}^3 (c_{t+i,f} - c_{t-i,f})}{2 * 3} \quad (10)$$

$$d_{t,f}^f = \frac{\sum_{i=1}^3 (c_{t,f+i} - c_{t,f-i})}{2 * 3} \quad (11)$$

$$c_{t,f}^{45} = \frac{\sum_{i=1}^3 (c_{t-i,f-i} - c_{t+i,f+i})}{2 * 3} \quad (12)$$

$$c_{t,f}^{135} = \frac{\sum_{i=1}^3 (c_{t+i,f-i} - c_{t-i,f+i})}{2 * 3} \quad (13)$$

其中,式(10)代表声纹信号的时域;式(11)代表声纹信号的频域;式(12)代表 45° 时声纹信号时频特性;式(13)代表 135° 时声纹信号的时频特性.线性回归后,信号特征维度将较高.为了进一步对特征值进行压缩,本节通过离散余弦变换进行处理,则有:

$$c_{t,m} = \sum_{k=1}^{24} d_{t,f}(k) \cos \left[(k-0.5) \frac{m\pi}{24} \right] \quad (14)$$

通过上述过程构建 4 组 12 维,共 48 维模型对用户声纹信号特征值进行提取,获取的特征值情况用图 3 进行描述.

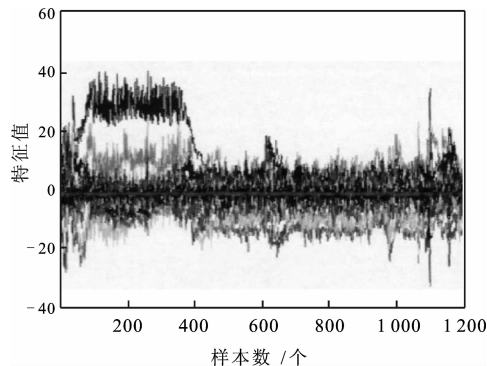


图 3 用户声纹特征图

2.3.3 根据声纹特征信息识别身份

为了确定用户身份,首先需对标准用户信息进行采集,以获取充分的先验知识.本节对若干模板采样数据的频域归一化结果均值进行计算,实现对先验数据的处理,也就是:

$$\bar{H}(e^{j\omega}) = \frac{\sum_{i=1}^N H_{ai}(e^{j\omega})}{N}$$

(15)

将上式作为模板. $H_{ai}(e^{j\omega})$ 用于描述对第 i 个样本进行傅里叶转换、平滑和归一化处理后的结果.

得到新的样本 $H_{ai}(e^{j\omega})$ 后,为了完成对其的识别,需对其与模板之间的不同进行计算.假设不同程度衡量参数计算公式如下:

$$D(H_{ai}(e^{j\omega}), \bar{H}(e^{j\omega})) = |H_{ai}(e^{j\omega}) - \bar{H}(e^{j\omega})|$$

(16)

将上式得到的结果简称为 D 值. 获取 D 值后,可通过 D 值进行判断,但 D 值的门限值需依据需要选择,如果用户身份不同,则样本间的 D 值满足 Gamma 分布,如果是同一用户,则样本间的 D 值近满足 Gaussian 分布,从而实现用户身份识别.

3 实验结果分析

本节通过无损 WAV 格式对真实语音进行采集,将采集的单人数据库作为实验样本,该数据库中有 100 个用户样本,1 500 组语音.实验将口令系统和智能卡系统作为对比,从识别精度的角度验证本文系统的有效性.

归一化均方误差可通过下式求出:

$$NMSE(\hat{b}) = \frac{\|\hat{b} - b\|^2}{\|b\|^2}$$

(17)

式中, $\hat{b}$ 用于描述识别后用户的身份向量; b 用于描述原始用户身份向量. NMSE 越小,认为身份识别精度越高.

分别采用本文系统、口令系统和智能卡系统对样本数据库中的用户身份进行识别,三种方法识别结果的归一化均方误差如表 1 所示.

分析表 1 可以看出,针对同一批用户,经本文系统识别后的归一化均方误差明显低于口令系统和智能卡系统,说明本文系统对身份识别有更好的性能.

虚警率就是被错误识别的用户身份占总用户的比例,漏警率就是将正确身份识别结果看作错误身份识别结果数量占总用户身份的比例,可通过下式求出:

表 1 三种方法归一化均方误差识别结果

用户编号	本文系统	口令系统	智能卡系统
1~10	0.145	0.182	0.179
11~20	0.102	0.135	0.152
21~30	0.097	0.127	0.119
31~40	0.091	0.122	0.135
41~50	0.105	0.136	0.128
51~60	0.087	0.119	0.135
61~70	0.095	0.135	0.129
71~80	0.112	0.146	0.177
81~90	0.109	0.152	0.182
91~100	0.115	0.141	0.166

$$X = \frac{N_{tp}}{N}$$

(18)

$$L = \frac{B_u}{N}$$

(19)

式中, N_{tp} 用于描述被错误识别的用户数量; N 用于描述用户总数量; B_u 用于描述将正确身份识别结果看作错误身份识别结果数量.

分别采用本文系统、口令系统和智能卡系统对样本数据库中的用户身份进行识别,三种方法识别结果的虚警率和漏警率如表 2 所示.

表 2 三种方法虚警率和漏警率比较结果

用户编号	本文系统		口令系统		智能卡系统	
	虚警率/ %	漏警率/ %	虚警率/ %	漏警率/ %	虚警率/ %	漏警率/ %
1~10	0.28	0.92	0.39	1.26	0.35	2.36
11~20	0.35	0.83	0.42	1.57	0.51	2.52
21~30	0.29	0.65	0.33	1.62	0.38	1.97
31~40	0.16	1.08	0.28	2.12	0.29	2.23
41~50	0.22	1.27	0.35	1.93	0.32	1.95
51~60	0.18	0.85	0.26	1.75	0.37	3.15
61~70	0.15	0.76	0.33	1.86	0.29	3.26
71~80	0.31	1.13	0.37	2.23	0.36	2.85
81~90	0.27	1.17	0.45	1.96	0.33	3.74
91~100	0.25	1.22	0.36	2.35	0.28	2.96

分析表 2 可以看出,在用户编号相同的情况下,本文系统的虚警率明显低于口令系统和智能卡系统,说明本文系统对用户进行身份识别的错误率较低;且本文系统的漏警率远远低于口令系统和智能卡系统,说明本文系统将合法身份识别为不合法身份的几率较低,进一步验证了本文系统的准确性.

4 结束语

本文设计了一种新的基于 Web 网络的身份识别系统,给出系统架构,介绍了浏览器端录制模块和 Web 服务器端声纹识别模块的设计过程.实验结果表明,所设计系统有很高的识别精度.

参考文献:

[1] 陈玉婷,范永胜. 基于身份识别的学校实验室网络安全设计与实现[J]. 绵阳师范学院学报, 2014, 33(11): 75-79.

[2] 郑浩. 基于 PHP 的网络在线投票系统的设计与实现[J]. 吉林师范大学学报(自然科学版), 2014, 35(1): 78-80.

[3] 冯贺平,杨敬娜,吴梅梅. 基于深度神经网络的身份识别研究[J]. 电脑知识与技术, 2016, 12(22): 161-162.

[4] 孙伟,刘晓敏,王浩宇,等. 基于三重人脸识别身份验证的门禁管理系统设计[J]. 计算机测量与控制, 2016, 24

(2): 225-227.

[5] 曲佳. 面向办公软件的身份认证系统的设计与实现[J]. 承德石油高等专科学校学报, 2015, 17(4): 35-39.

[6] 林超,叶振昀. 基于校园卡的大学生电子照片采集系统的设计与实现[J]. 微型电脑应用, 2015, 31(8): 46-47.

[7] 金丰,张悦,雍鹏. 基于双服务器的 B/S 模式监测系统的设计与实现[J]. 计算机仿真, 2014, 31(2): 201-205.

[8] 洪远泉,周永明,郑春龙,等. 基于距离判别的心电身份识别技术研究[J]. 科学技术与工程, 2014, 14(18): 252-255.

[9] 黄志威. 改进的 AdaBoost 检测算法在身份识别中的应用[J]. 科技通报, 2015, 31(10): 190-192.

[10] 赵卫. 一种基于网络安全的 WIFI 系统身份认证设计[J]. 电子设计工程, 2016, 24(14): 81-83.

作者简介:

陈文庆 男, (1981-), 硕士, 讲师. 研究方向为计算机网络与多媒体教学. E-mail: wwwsheying11@126. com.

(上接第 113 页)

[9] 李晓东,陈俊健,曾凡智. 基于 KDG-tree 的数据库多维索引技术[J]. 计算机应用与软件, 2013, 30(6): 162-165.

[10] 蒋华,张乐乾,王鑫. 基于多维评价模型及改进蚁群优化算法的云计算资源调度策略[J]. 计算机测量与控制, 2015, 23(7): 2559-2562.

[11] 周维,李敬辉,黄高明. 基于 VTK 的海战场电磁信息可视化技术研究[J]. 计算机与数字工程, 2012, 40(1): 121-124.

[12] 李伟波,李振中,何频. 移动位置服务的岩土信息可视化技术研究[J]. 计算机工程与应用, 2012, 48(5):

228-231.

[13] 董重,魏迎梅. 运用平行坐标系的多变元时序数据可视化方法[J]. 小型微型计算机系统, 2015, 36(10): 2408-2411.

[14] 叶学义,宋倩倩,高真,等. 基于直方图条件熵的水声数据分类算法[J]. 计算机工程, 2016, 42(11): 244-248.

作者简介:

张红军 男, (1975-), 硕士, 讲师. 研究方向为计算机算法与分析、软件工程、计算机教育与应用. E-mail: bspandsh@163. com.