



(12)发明专利

(10)授权公告号 CN 104217718 B

(45)授权公告日 2017.05.17

(21)申请号 201410446143.5

(22)申请日 2014.09.03

(65)同一申请的已公布的文献号

申请公布号 CN 104217718 A

(43)申请公布日 2014.12.17

(73)专利权人 陈飞

地址 310000 浙江省杭州市滨江区伟业路彩虹豪庭2幢202室

(72)发明人 陈飞

(74)专利代理机构 厦门市首创君合专利事务所有限公司 35204

代理人 连耀忠

(51)Int.Cl.

G10L 15/07(2013.01)

G10L 15/24(2013.01)

(56)对比文件

CN 1372222 A,2002.10.02,全文.

CN 1809196 A,2006.07.26,全文.

CN 101673544 A,2010.03.17,全文.

CN 102831195 A,2012.12.19,全文.

CN 103092950 A,2013.05.08,全文.

CN 103544255 A,2014.01.29,全文.

JP 特开2005-3926 A,2005.01.06,全文.

JP 特开平11-311993 A,1999.11.09,全文.

US 2006/0058920 A1,2006.03.16,全文.

US 2008/0208593 A1,2008.08.28,全文.

US 6167376 A,2000.12.26,全文.

秦飞等.关于方言语音信息识别技术在公安工作中的应用研究.《甘肃警察职业学院学报》.2013,第11卷(第3期),全文.

闫建鹏.基于多模态特征的新闻视频语义分析.《中国优秀硕士学位论文全文数据库 信息科技辑》.2013,(第5期),全文.

审查员 李召卿

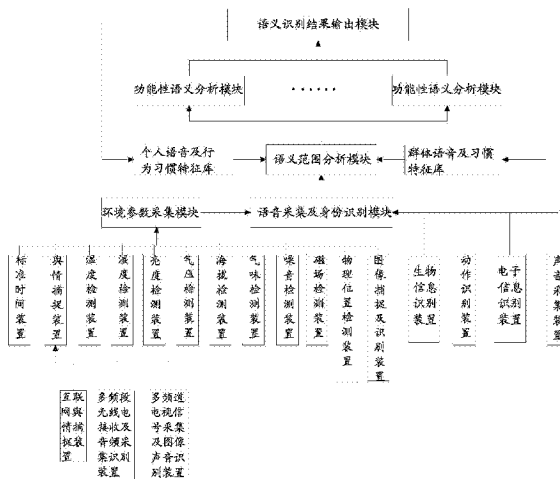
权利要求书2页 说明书4页 附图1页

(54)发明名称

依据环境参数及群体趋向数据的语音识别方法和系统

(57)摘要

依据环境参数及群体趋向数据的语音识别方法和系统,首先,采集用户语音信息及与该语音信息相关的环境参数和身份识别信息,依据该身份识别信息进行用户的身份识别;其次,提取个人语音及行为习惯特征库中存储的该用户个人语音及行为习惯特征数据,结合提取群体语音及习惯特征库中存储的类似群体的群体语音及习惯特征数据,判断该语音信息适合哪些功能性语义分析模块,并将相关信息发送给这些语义分析模块进行单独或联合的语义解析,从而得到语义分析结果;最后,存储语义分析结果,并将语义分析结果及相关信息发送至个人语音及行为习惯特征库和群体语音及习惯特征库进行归类存储。本发明大幅度提高了语音识别的准确率。



CN 104217718 B

1. 依据环境参数及群体趋向数据的语音识别方法,其特征在于:预先设置有个人语音及行为习惯特征库,群体语音及习惯特征库,以及多个功能性语义分析模块;该个人语音及行为习惯特征库是按照用户身份信息归类存储的数据库;该群体语音及习惯特征库用于存储提取自个人语音及行为习惯特征库的所有个人语音及行为习惯特征数据并按照多个维度记录的不同分类方法下的类似群体的语音特征和习惯特征;该功能性语义分析模块是将类似场景和类似群体中可能出现的语义构成集合,并按照出现概率的高低动态调整搜索方法和次序的功能模块;类似场景是依据环境参数的相似程度进行判定,类似群体是依据类似场景中,行为习惯的接近程度并结合用户信息的共同点进行判定,识别步骤如下:

首先,采集用户语音信息及与该语音信息相关的环境参数和身份识别信息,依据该身份识别信息进行用户的身份识别;

其次,依据身份识别信息提取个人语音及行为习惯特征库中存储的该用户个人语音及行为习惯特征数据,结合依据环境参数类型和身份识别信息类型提取群体语音及习惯特征库中存储的类似群体的群体语音及习惯特征数据,判断该语音信息适合哪些功能性语义分析模块,并将相关信息发送给这些语义分析模块进行单独或联合的语义解析,从而得到语义分析结果;

最后,存储语义分析结果,并将语义分析结果及相关信息发送至个人语音及行为习惯特征库和群体语音及习惯特征库进行归类存储,以不断完善该两个特征库的内容。

2. 如权利要求1所述的依据环境参数及群体趋向数据的语音识别方法,其特征在于:所述环境参数包括以下的一种或多种:时间、舆情、温度、湿度、亮度、气压、海拔、气味、噪音、磁场、物理位置及图像。

3. 如权利要求1所述的依据环境参数及群体趋向数据的语音识别方法,其特征在于:所述环境参数包括舆情,该舆情包括以下的一种或多种:互联网舆情、广播舆情及电视舆情。

4. 如权利要求1所述的依据环境参数及群体趋向数据的语音识别方法,其特征在于:所述的身份识别信息包括以下的一种或多种:生物信息、动作信息及电子信息。

5. 如权利要求1所述的依据环境参数及群体趋向数据的语音识别方法,其特征在于:所述相关信息至少包括提取到的用户个人语音及行为习惯特征数据、群体语音及习惯特征数据、环境参数和采集到的语音信息。

6. 依据环境参数及群体趋向数据的语音识别系统,其特征在于:包括:

环境参数采集模块,用于采集来自环境参数检测部件的环境参数,并将采集到的环境参数的值进行标准化处理后送至语音采集及身份识别模块;

语音采集及身份识别模块,与环境参数采集模块和语音范围分析模块相连,用于采集来自语音及身份识别信息检测部件的用户语音信息和身份识别信息,并与环境参数一起发送至语义范围分析模块;

个人语音及行为习惯特征库,用于按照用户身份信息归类存储在各类环境条件下的个人语音及行为习惯特征数据;

群体语音及习惯特征库,用于存储提取自所有用户的个人语音及行为习惯特征库的群体特征数据,并按照多个维度记录不同分类方法下的群体的语音特征和习惯特征;

语义范围分析模块,与个人语音及行为习惯特征库、群体语音及习惯特征库和多个功能性语义分析模块相连,依据身份识别信息提取个人语音及行为习惯特征库中存储的该用

户个人语音及行为习惯特征数据,结合依据环境参数类型和身份识别信息类型提取群体语音及习惯特征库中存储的类似群体的群体语音及习惯特征数据,判断该语音文件适合哪些功能性语义分析模块,并将相关信息发送给这些语义分析模块进行单独或联合的语义解析;

功能性语义分析模块,均与语义识别结果输出模块相连,是将类似场景和类似群体中可能出现的语义构成集合,并按照出现概率的高低动态调整搜索方法和次序,用于执行语义解析;

语义识别结果输出模块,与个人语音及行为习惯特征库和群体语音及习惯特征库相连,用于存储功能性语义分析模块的语义分析结果,并将语义分析结果及相关信息发送给个人语音及行为习惯特征库和群体语音及习惯特征库进行归类存储。

7.如权利要求6所述的依据环境参数及群体趋向数据的语音识别系统,其特征在于:所述的环境参数检测部件包括以下一种或多种:标准时间装置、舆情捕捉装置、温度检测装置、湿度检测装置、亮度检测装置、气压检测装置、海拔检测装置、气味检测装置、噪音检测装置、磁场检测装置、物理位置检测装置和图像捕捉及识别装置。

8.如权利要求6所述的依据环境参数及群体趋向数据的语音识别系统,其特征在于:所述的环境参数检测部件包括舆情捕捉装置,该舆情捕捉装置包括以下一种或多种:互联网舆情捕捉装置、多频段无线电接收及音频采集识别装置、多频道电视信号采集及图形声音识别装置。

9.如权利要求6所述的依据环境参数及群体趋向数据的语音识别系统,其特征在于:所述的身份识别信息检测部件包括以下一种或多种:生物信息识别装置、动作识别装置和电子信息识别装置。

## 依据环境参数及群体趋向数据的语音识别方法和系统

### 技术领域

[0001] 本发明涉及语音识别领域,特别是一种依据环境参数及群体趋向数据的语音识别方法和系统。

### 背景技术

[0002] 随着互联网和数字多媒体内容的发展,人与智能设备之间的交互成为重要研究方向。语音交互作为最自然方便的一种人机交互手段,是很多其他应用能够实现的基础,例如:语音拨号、语音导航,室内设备控制等等。其目标是将人类的语音中的词汇内容转换为计算机可读的输入,例如按键、二进制编码或者字符序列等。然而,语音识别的准确率一直是语音识别技术的瓶颈,主要原因在于人类对话时靠声音表达的信息只占小部分,更多的信息需要根据经验、视觉、触觉、嗅觉来提供,因此很有必要提出一种的依据语音环境等信息,提高语音识别准确率的语音识别方法和系统。

### 发明内容

[0003] 本发明的主要目的在于克服现有技术中的上述缺点,提出一种提高语音识别准确率的依据环境参数及群体趋向数据的语音识别方法和系统。

[0004] 本发明采用如下技术方案:

[0005] 依据环境参数及群体趋向数据的语音识别方法,其特征在于:预先设置有个人语音及行为习惯特征库,群体语音及习惯特征库,以及多个功能性语义分析模块;该个人语音及行为习惯特征库是按照用户身份信息归类存储的数据库;该群体语音及习惯特征库用于存储提取自个人语音及行为习惯特征库的所有个人语音及行为习惯特征数据并按照多个维度记录的不同分类方法下的类似群体的语音特征和习惯特征;该功能性语义分析模块是将类似场景和类似群体中可能出现的语义构成集合,并按照出现概率的高低动态调整搜索方法和次序的功能模块;类似场景是依据环境参数的相似程度进行判定,类似群体是依据类似场景中,行为习惯的接近程度并结合用户信息的共同点进行判定,识别步骤如下:

[0006] 首先,采集用户语音信息及与该语音信息相关的环境参数和身份识别信息,依据该身份识别信息进行用户的身份识别;

[0007] 其次,依据身份识别信息提取个人语音及行为习惯特征库中存储的该用户个人语音及行为习惯特征数据,结合依据环境参数类型和身份识别信息类型提取群体语音及习惯特征库中存储的类似群体的群体语音及习惯特征数据,判断该语音信息适合哪些功能性语义分析模块,并将相关信息发送给这些语义分析模块进行单独或联合的语义解析,从而得到语义分析结果;

[0008] 最后,存储语义分析结果,并将语义分析结果及相关信息发送至个人语音及行为习惯特征库和群体语音及习惯特征库进行归类存储,以不断完善该两个特征库的内容。

[0009] 优选的,所述环境参数包括以下的一种或多种:时间、舆情、温度、湿度、亮度、气压、海拔、气味、噪音、磁场、物理位置及图像。

[0010] 优选的,所述环境参数包括舆情,该舆情包括以下的一种或多种:互联网舆情、广播舆情及电视舆情。

[0011] 优选的,所述的身份识别信息包括以下的一种或多种:生物信息、动作信息及电子信息。

[0012] 优选的,所述相关信息至少包括提取到的用户个人语音及行为习惯特征数据、群体语音及习惯特征数据、环境参数和采集到的语音信息。

[0013] 依据环境参数及群体趋向数据的语音识别系统,其特征就在于:包括:

[0014] 环境参数采集模块,用于采集来自环境参数检测部件的环境参数,并将采集到的环境参数的值进行标准化处理后送至语音采集及身份识别模块;

[0015] 语音采集及身份识别模块,与环境参数采集模块和语音范围分析模块相连,用于采集来自语音及身份识别信息检测部件的用户语音信息和身份识别信息,并与环境参数一起发送至语义范围分析模块;

[0016] 个人语音及行为习惯特征库,用于按照用户身份信息归类存储各类环境条件下的个人语音及行为习惯特征数据;

[0017] 群体语音及习惯特征库,用于存储提取自所有用户的个人语音及行为习惯特征库的群体特征数据,并按照多个维度记录不同分类方法下的群体的语音特征和习惯特征;

[0018] 语义范围分析模块,与个人语音及行为习惯特征库、群体语音及习惯特征库和多个功能性语义分析模块相连,依据身份识别信息提取个人语音及行为习惯特征库中存储的该用户个人语音及行为习惯特征数据,结合依据环境参数类型和身份识别信息类型提取群体语音及习惯特征库中存储的类似群体的群体语音及习惯特征数据,判断该语音文件适合哪些功能性语义分析模块,并将相关信息发送给这些语义分析模块进行单独或联合的语义解析;

[0019] 功能性语义分析模块,均与语义识别结果输出模块相连,是将类似场景和类似群体中可能出现的语义构成集合,并按照出现概率的高低动态调整搜索方法和次序,用于执行语义解析;

[0020] 语义识别结果输出模块,与个人语音及行为习惯特征库和群体语音及习惯特征库相连,用于存储功能性语义分析模块的语义分析结果,并将语义分析结果及相关信息发送给个人语音及行为习惯特征库和群体语音及习惯特征库进行归类存储。

[0021] 优选的,所述的环境参数检测部件包括以下一种或多种:标准时间装置、舆情捕捉装置、温度检测装置、湿度检测装置、亮度检测装置、气压检测装置、海拔检测装置、气味检测装置、噪音检测装置、磁场检测装置、物理位置检测装置和图像捕捉及识别装置。

[0022] 优选的,所述的环境参数检测部件包括舆情捕捉装置,该舆情捕捉装置包括以下一种或多种:互联网舆情捕捉装置、多频段无线电接收及音频采集识别装置、多频道电视信号采集及图形声音识别装置。

[0023] 优选的,所述的身份识别信息检测部件包括以下一种或多种:生物信息识别装置、动作识别装置和电子信息识别装置。

[0024] 由上述对本发明的描述可知,与现有技术相比,本发明具有如下有益效果:

[0025] 本发明设置有群体语音及习惯特征库及个人语音及行为习惯特征库,依据环境参数信息补足视觉、触觉、嗅觉等需要的信息,再结合群体数据及个体历史数据来补足经验判

断所需要的信息,再通过适合的单个或多个功能性语义分析模块进行语义分析,从而大幅度提高语音识别的准确率。

## 附图说明

[0026] 图1为本发明系统的组成框图。

## 具体实施方式

[0027] 以下通过具体实施方式对本发明作进一步的描述。

[0028] 依据环境参数及群体趋向数据的语音识别方法,预先设置有个人语音及行为习惯特征库,群体语音及习惯特征库,以及多个功能性语义分析模块。该个人语音及行为习惯特征库是按照用户身份信息归类存储的数据库。该群体语音及习惯特征库用于存储提取自个人语音及行为习惯特征库的所有个人语音及行为习惯特征数据并按照多个维度记录的不同分类方法下的类似群体的语音特征和习惯特征。该功能性语义分析模块是将类似场景和类似群体中可能出现的语义构成集合,并按照出现概率的高低动态调整搜索方法和次序的功能模块。类似场景是依据环境参数(特定的物理地点、时间、上下文语境、舆情等)的相似程度进行判定,类似群体是依据类似场景中,行为习惯的接近程度并结合用户信息的共同点进行判定。该环境参数的具体相似程度和行为习惯的具体接近程度可预先设定,具体的舆情环境参数基于语义分析。类似事件序列及事件集群为近似环境。其它环境参数以当前检测参数及历史检测参数序列的相似性判断相似环境。行为习惯的相似性依据用户身份及个人信息、类似环境下语言、命令选择序列、动作序列来判断相似性具体识别步骤如下:

[0029] 首先,采集用户语音信息及与该语音信息相关的环境参数和身份识别信息,依据该身份识别信息进行用户的身份识别。环境参数包括有以下几种:时间、舆情、温度、湿度、亮度、气压、海拔、气味、噪音、磁场、物理位置及图像等。其中,舆情包括互联网舆情、广播舆情及电视舆情,即对互联网、广播及电视中的内容进行识别、整理形成的舆情信息,(例如:新闻、电视剧、电影、小说等)。身份识别信息包括以下几种:生物信息、动作信息及电子信息,例如指纹、虹膜、手势、读卡器等。

[0030] 其次,依据身份识别信息提取个人语音及行为习惯特征库中存储的该用户个人语音及行为习惯特征数据,结合依据环境参数类型和身份识别信息类型提取群体语音及习惯特征库中存储的类似群体的群体语音及习惯特征数据,判断该语音信息适合哪些功能性语义分析模块,并将相关信息发送给这些语义分析模块进行单独或联合的语义解析,从而得到语义分析结果。该相关信息至少包括提取到的用户个人语音及行为习惯特征数据、群体语音及习惯特征数据、环境参数和采集到的语音信息。

[0031] 最后,存储语义分析结果,并将语义分析结果及相关信息发送至个人语音及行为习惯特征库和群体语音及习惯特征库进行归类存储,以不断完善这两个数据库的内容。

[0032] 本发明基于上述方法还提出依据环境参数及群体趋向数据的语音识别系统,参照图1,该系统具体包括如下装置:

[0033] 环境参数采集模块,用于采集来自环境参数检测部件的环境参数,并将采集到的环境参数的值进行标准化处理后送至语音采集及身份识别模块。环境参数检测部件包括以下几种:标准时间装置、舆情捕捉装置、温度检测装置、湿度检测装置、亮度检测装置、气压

检测装置、海拔检测装置、气味检测装置、噪音检测装置、磁场检测装置、物理位置检测装置和图像捕捉及识别装置。其中,舆情捕捉装置可配置以下几种:一种为可访问互联网的网络接口及互联网舆情捕捉装置;一种为多频段无线电信号接收及音频采集识别装置;一种为多频道电视信号采集及图形声音识别装置。用于对互联网、广播、电视中的内容进行识别、整理,形成舆情信息。

[0034] 语音采集及身份识别模块,与环境参数采集模块和语音范围分析模块相连,用于采集来自语音及身份识别信息检测部件的用户语音信息和身份识别信息,并与环境参数一起发送至语义范围分析模块。其中的身份识别信息检测部件包括生物信息识别装置(例如指纹识别装置、虹膜识别装置)、动作识别装置(例如kinect)和电子信息识别装置(例如IC卡及读卡器)。

[0035] 个人语音及行为习惯特征库,用于按照用户身份信息归类存储各类环境条件下的个人语音及行为习惯特征数据。

[0036] 群体语音及习惯特征库,用于按照用户身份信息的类型及环境参数类型归类存储的群体语音及习惯特征库。

[0037] 语义范围分析模块,与个人语音及行为习惯特征库、群体语音及习惯特征库和多个功能性语义分析模块相连,依据身份识别信息提取个人语音及行为习惯特征库中存储的该用户个人语音及行为习惯特征数据,结合依据环境参数类型和身份识别信息类型提取群体语音及习惯特征库中存储的群体语音及习惯特征数据,判断该语音文件适合哪些功能性语义分析模块,并将相关信息发送给这些语义分析模块进行单独或联合的语义解析。

[0038] 功能性语义分析模块,均与语义识别结果输出模块相连,是将同一场景中可能出现的语义构成集合,并按照出现概率的高低动态调整搜索方法和次序,用于执行语义解析。

[0039] 语义识别结果输出模块,与个人语音及行为习惯特征库和群体语音及习惯特征库相连,用于存储功能性语义分析模块的语义分析结果,并将语义分析结果及相关信息发送给个人语音及行为习惯特征库和群体语音及习惯特征库进行归类存储。

[0040] 本发明依据环境参数补足视觉、触觉、嗅觉等需要的信息,再结合群体数据及个体历史数据来补足经验判断所需要的信息,从而大幅度提高语音识别的准确率。

[0041] 上述仅为本发明的具体实施方式,但本发明的设计构思并不局限于此,凡利用此构思对本发明进行非实质性的改动,均应属于侵犯本发明保护范围的行为。

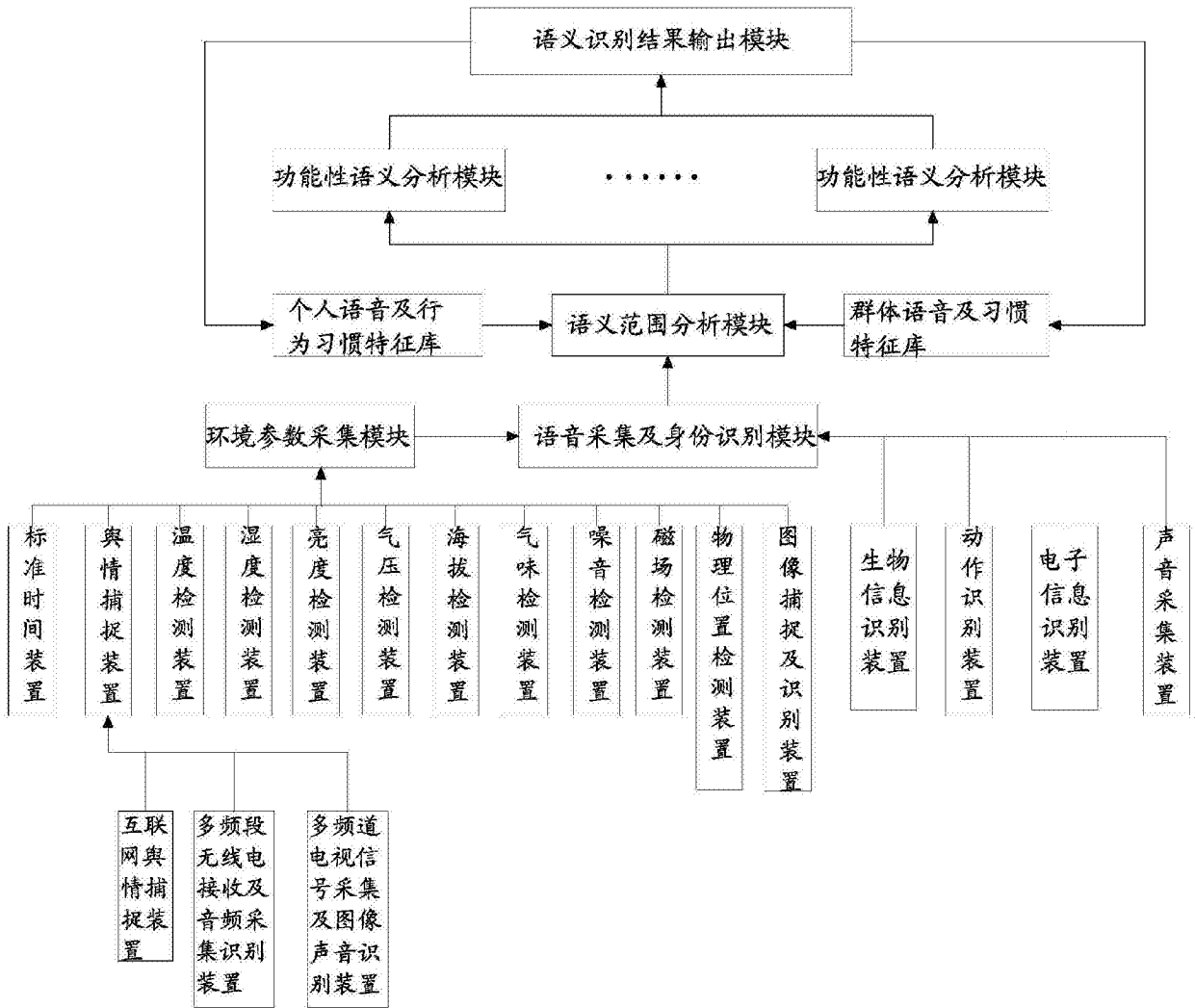


图1