



(12)发明专利申请

(10)申请公布号 CN 111179927 A  
(43)申请公布日 2020.05.19

(21)申请号 201911327665.2

(22)申请日 2019.12.20

(71)申请人 恒银金融科技股份有限公司  
地址 300308 天津市滨海新区自贸区(空港经济区)西八道30号

(72)发明人 田立刚

(74)专利代理机构 天津市三利专利商标代理有限公司 12107  
代理人 韩新城

(51) Int. Cl.  
G10L 15/22(2006.01)  
G10L 15/25(2013.01)  
G06K 9/00(2006.01)  
G06F 40/30(2020.01)  
G06F 16/332(2019.01)

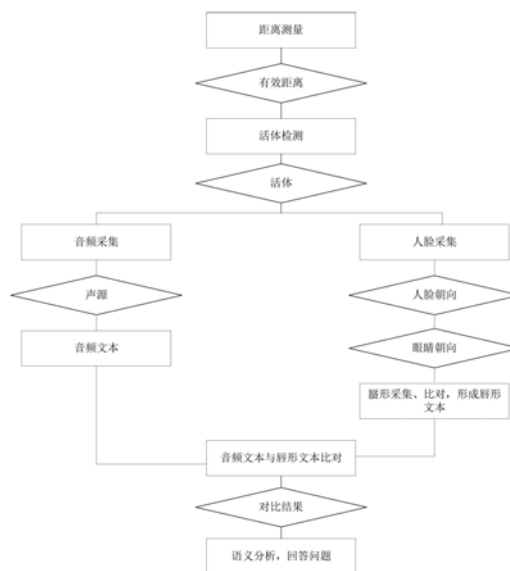
权利要求书1页 说明书3页 附图2页

(54)发明名称

一种金融设备语音交互方法以及系统

(57)摘要

本发明公开一种金融设备语音交互方法以及系统。该方法包括:进行设备前面人与金融设备间的距离判断,若距离小于判定阈值,进行活体判断;若为活体,则启动人脸摄像头,判断声源方向;若为有效声源,则采集音频信号并转出音频文本;通过人脸识别判断人脸方向以及眼睛目光方向是否朝向设备;若朝向设备,开启唇语识别;唇语图像采集与唇语数据库比对,输出唇语文本;将音频文本与唇语文本对比,若对比通过,则对设备前面人对设备所发出的指令进行语义分析并回复。本发明解决了嘈杂环境下以及设备前两人或多人互相交流,不是与设备交流,设备乱说话问题,增强了客户体验度。



1. 一种金融设备语音交互系统的交互方法,其特征在于,包括步骤:

进行设备前面人与金融设备间的距离判断,若距离小于判定阈值,进行活体判断;若为活体,则启动人脸摄像头,判断声源方向;若为有效声源,则采集音频信号并转出音频文本;通过人脸识别判断人脸方向以及眼睛目光方向是否朝向设备;若朝向设备,开启唇语识别;唇语图像采集与唇语数据库比对,输出唇语文本;将音频文本与唇语文本对比,若对比通过,则对设备前面人对设备所发出的指令进行语义分析并回复。

2. 根据权利要求1所述金融设备语音交互方法,其特征在于,通过超声波传感器和\或激光雷达来检测设备与设备前面人间的距离。

3. 根据权利要求1所述金融设备语音交互方法,其特征在于,通过红外传感器进行活体判断。

4. 根据权利要求1所述金融设备语音交互方法,其特征在于,通过麦克风阵列判断声源方向。

5. 一种所述金融设备语音交互系统,其特征在于,包括主控系统以及与所述主控系统连接的测距装置、红外传感器、人脸摄像头、麦克风阵列,喇叭,所述主控系统还通过无线模块连接到唇语数据库以及语义识别问答系统;

所述测距装置用于对设备前面人与金融设备间的距离进行判断,若距离小于判定阈值,通过红外传感器进行活体判断,若为活体,则开启人脸摄像头,通过麦克风阵列判断声源方向;若为有效声源,则采集音频信号并转出音频文本;通过人脸识别判断人脸方向以及眼睛目光方向是否朝向设备;若朝向设备,开启唇语识别;唇语图像采集与唇语数据库比对,输出唇语文本;将音频文本与唇语文本对比,若对比通过,则对设备前面人对设备所发出的指令进行语义分析并回复。

6. 根据权利要求5所述金融设备语音交互系统,其特征在于,所述测距装置包括超声波传感器和\或激光雷达。

7. 根据权利要求5所述金融设备语音交互系统,其特征在于,所述语义识别问答系统包括语义识别模块以及问答数据库,所述语义识别模块与问答数据库。

8. 根据权利要求5所述金融设备语音交互系统,其特征在于,所述无线模块为WIFI模块。

## 一种金融设备语音交互方法以及系统

### 技术领域

[0001] 本发明涉及金融自助终端技术领域,特别是涉及一种金融设备语音交互方法以及系统。

### 背景技术

[0002] 随着金融自助设备和语音交互设备的广泛应用,在嘈杂环境中,设备会识别到错误或者无效的声音,并对无效声音进行回答,大大降低了客户体验度。

### 发明内容

[0003] 本发明的目的是针对现有技术中存在的技术缺陷,而提供一种金融设备语音交互方法以及系统。

[0004] 为实现本发明的目的所采用的技术方案是:

[0005] 一种金融设备语音交互系统的交互方法,包括步骤:

[0006] 进行设备前面人与金融设备间的距离判断,若距离小于判定阈值,进行活体判断;若为活体,则启动人脸摄像头,判断声源方向;若为有效声源,则采集音频信号并转出音频文本;通过人脸识别判断人脸方向以及眼睛目光方向是否朝向设备;若朝向设备,开启唇语识别;唇语图像采集与唇语数据库比对,输出唇语文本;将音频文本与唇语文本对比,若对比通过,则对设备前面人对设备所发出的指令进行语义分析并回复。

[0007] 优选的,通过超声波传感器和\或激光雷达来检测设备与设备前面人间的距离。

[0008] 优选的,通过红外传感器进行活体判断。

[0009] 优选的,通过麦克风阵列判断声源方向。

[0010] 本发明的目的还在于提供一种金融设备语音交互系统,包括主控系统以及与所述主控系统连接的测距装置、红外传感器、人脸摄像头、麦克风阵列,喇叭,所述主控系统还通过无线模块连接到唇语数据库以及语义识别问答系统;

[0011] 所述测距装置用于对设备前面人与金融设备间的距离进行判断,若距离小于判定阈值,通过红外传感器进行活体判断,若为活体,则开启人脸摄像头,通过麦克风阵列判断声源方向;若为有效声源,则采集音频信号并转出音频文本;通过人脸识别判断人脸方向以及眼睛目光方向是否朝向设备;若朝向设备,开启唇语识别;唇语图像采集与唇语数据库比对,输出唇语文本;将音频文本与唇语文本对比,若对比通过,则对设备前面人对设备所发出的指令进行语义分析并回复。

[0012] 其中,所述测距装置包括超声波传感器和\或激光雷达。

[0013] 其中,所述语义识别问答系统包括语义识别模块以及问答数据库,所述语义识别模块与问答数据库。

[0014] 其中,所述无线模块为WIFI模块。

[0015] 本发明解决了嘈杂环境下以及设备前两人或多人互相交流,不是与设备交流,设备乱说话问题,增强了客户体验度。

## 附图说明

[0016] 图1是金融设备语音交互系统交互方法的流程图。

[0017] 图2是金融设备语音交互系统结构原理图；

## 具体实施方式

[0018] 以下结合附图和具体实施例对本发明作进一步详细说明。应当理解，此处所描述的具体实施例仅仅用以解释本发明，并不用于限定本发明。

[0019] 如图1所示，本发明的一种金融设备语音交互系统的交互方法，包括步骤：

[0020] 通过测距装置进行设备前面人与金融设备间的距离判断，若距离小于判定阈值，通过红外传感器进行活体判断，若为活体，则开启人脸摄像头，通过麦克风阵列判断声源方向；若为有效声源，则采集音频信号并转出音频文本；通过人脸识别判断人脸方向以及眼睛目光方向是否朝向设备；若朝向设备，开启唇语识别；唇语图像采集与唇语数据库比对，输出唇语文本；将音频文本与唇语文本对比，若对比通过，则对设备前面人对设备所发出的指令进行语义分析并回复。

[0021] 本发明中，通过距离判断为有效的交流范围距离时，才进一步的开启红外传感器进行活体判断。通常情况下，语言交流通常距离是1.2米左右，阈值距离1.5。通过激光雷达测距和超声波测距，在阈值范围内，判定为语音交流距离，然后进一步的开启红外传感器进行活体判断，否则不开启下一步的交互。

[0022] 本发明中，在判断为有效交流距离时，启动活体识别系统。具体是通过红外传感器来检测，判断设备前面人是否为活体，如是则打开人脸摄像头和音频识别功能，否则不打开人脸摄像头和音频识别功能。

[0023] 本发明中，还进一步的通过麦克风阵列，判断声音来源方向。通过来声源角度判断是否是有效声音，若有效，将采集中音频转换成相应音频文本，否则不进行音频到文本的处理转换，不处理。

[0024] 本发明中，还通过人脸识别中的人脸方向和人眼方向识别。通过人脸摄像头采集人脸后，通过人脸方向和人眼的目光方向识别，判断交流人是否面向设备，或者眼光是否关注设备。若人脸面向设备，眼睛目光朝向设备，判定为有效交流。则进行下一步的交互步骤，否则不再继续进行。

[0025] 本发明通过以上的技术方案，能实现当嘈杂环境下以及设备前两人或多人互相交流，而非对设备进行交流，设备不会被接受的音频所干扰，设备能自主判断出有效音频信号，并对问题进行回答。

[0026] 本发明中，在判断有效交流的基础上，还通过唇型判断说话内容。采集的唇型通过与数据库进行分析对比，产生相应语言文本。将音频文本与唇语文本进行对比，判断音频文本是否设备前面的人的语音。

[0027] 如图2所示，本发明还在于提供一种金融设备语音交互系统，包括主控系统以及与所述主控系统连接的测距装置、红外传感器、人脸摄像头、麦克风阵列，喇叭，所述主控系统还通过无线模块连接到唇语数据库以及语义识别问答系统。

[0028] 其中，所述测距装置包括超声波传感器和\或激光雷达。

[0029] 其中，所述语义识别问答系统包括语义识别模块以及问答数据库，所述语义识别

模块与问答数据库。

[0030] 其中,所述无线模块为WIFI模块。

[0031] 本发明的上述的系统,通过激光雷达,超声波传感器进行距离判断;若距离小于判定阈值,红外传感器进行活体判断;若为活体,开启人脸摄像头和语音识别;语音识别通过麦克风阵列判断声源方向;对音频信号采集,转出文本;人脸识别判断人脸方向,眼睛目光方向,是否朝向设备;若人脸方向朝向设备,开启唇语识别;唇语图像采集与数据库进行对比,输出文本;音频文本与唇语文本对比。若对比通过,语音音频为设备前面人对设备发出的指令,进行语义分析,回复答案。

[0032] 本发明通过以上的技术方案,能实现当嘈杂环境下以及设备前两人或多人互相交流,而非对设备进行交流,设备不会被接受的音频所干扰,设备能自主判断出有效音频信号,并对问题进行回答。

[0033] 以上所述仅是本发明的优选实施方式,应当指出的是,对于本技术领域的普通技术人员来说,在不脱离本发明原理的前提下,还可以做出若干改进和润饰,这些改进和润饰也应视为本发明的保护范围。

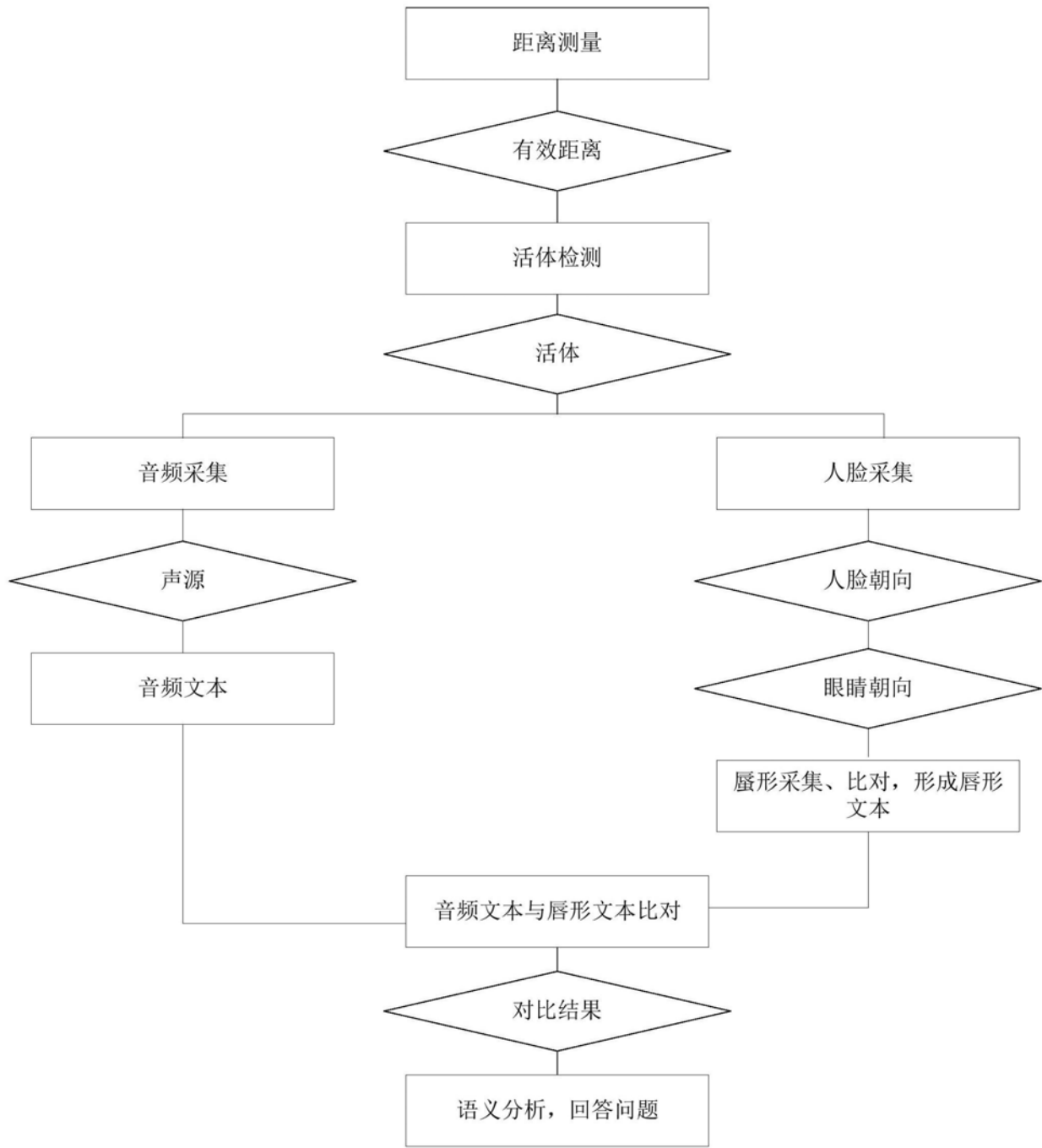


图1

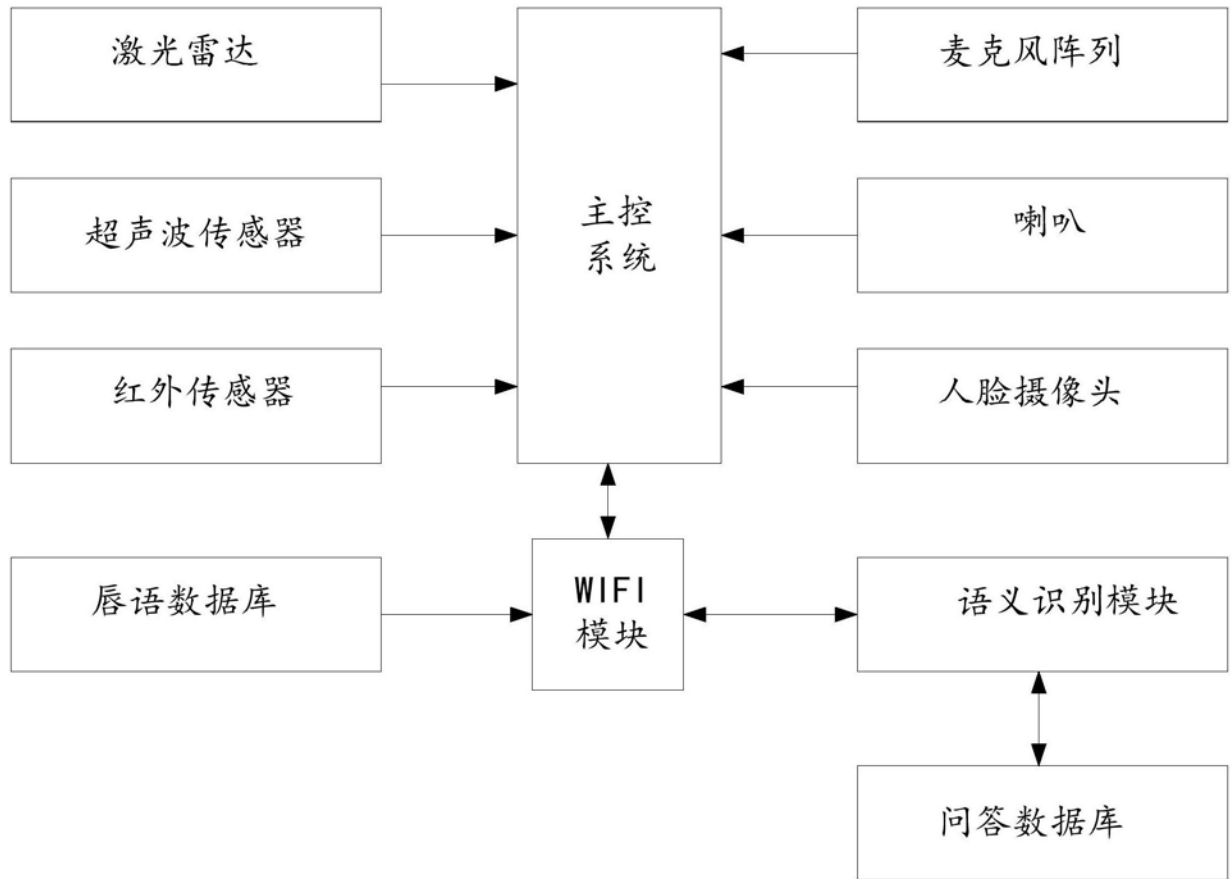


图2