(19)中华人民共和国国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 102938811 A
(43)申请公布日 2013.02.20

(21)申请号 201210389838.5

(22)申请日 2012.10.15

(71)申请人 华南理工大学
地址 510641 广东省广州市天河区五山路381号

(72)发明人 徐向民 林培东 庞聪 侯世莹 李叙琼

(74)专利代理机构 广州市华学知识产权代理有限公司 44245
代理人 蔡茂略

(51)Int.Cl.
H04M 1/725(2006.01)
G10L 15/14(2006.01)

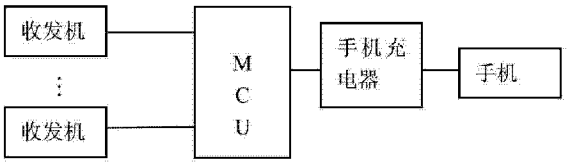
权利要求书 1 页 说明书 3 页 附图 1 页

(54)发明名称

一种基于语音识别的家庭手机通话系统

(57)摘要

本发明公开了一种基于语音识别的家庭手机通话系统，包括总控制器、多个收发机和手机，所述总控制器与手机通过USB接口连接，所述总控制器与所述收发机连接；所述总控制器包括手机控制模块、收发机控制模块和语音识别模块。本发明实现了在家里随时随地进行接听和呼叫，极大地提高了室内通讯的便捷性。



CN 102938811 A

1. 一种基于语音识别的家庭手机通话系统，其特征在于，包括总控制器、手机和多个收发机，所述总控制器与手机通过 USB 接口连接，所述总控制器与所述收发机连接；

所述总控制器包括：

手机控制模块，用于对手机进行控制，所述控制包括接听、拨号和读通信录；

收发机控制模块，用于接收各收发机的信息、接收语音信号并对语音信号进行放大，同时用于对用户进行定位；

语音识别模块，用于识别用户的语音指令。

2. 根据权利要求 1 所述的一种基于语音识别的家庭手机通话系统，其特征在于，还包括手机充电器，所述总控制器与手机充电器连接，所述手机充电器与手机通过 USB 接口连接。

3. 根据权利要求 1 所述的一种基于语音识别的家庭手机通话系统，其特征在于，所述语音识别模块采用离散隐马尔可夫模型对用户的语音指令进行识别。

4. 根据权利要求 1 所述的一种基于语音识别的家庭手机通话系统，其特征在于，所述总控制器为 MCU，采用 ARM 架构芯片。

一种基于语音识别的家庭手机通话系统

技术领域

[0001] 本发明涉及智能家居领域，特别涉及一种基于语音识别的家庭手机通话系统。

背景技术

[0002] 在家里，手机随身带着很不方便，尤其是充电的时候想接听或者拨打电话都受到充电器位置的限制，所以有必要提出一种系统，可以实现在家里随时随地进行接听和呼叫，会极大地提高室内通讯的便捷性。

[0003] ADB(Android Debug Bridge)，是 android SDK 里的一个工具，可以直接操作管理 android 模拟器或者真实的 android 设备。

[0004] 隐马尔科夫模型是一种统计模型，用来描述一个含有隐含未知参数的马尔可夫过程，该模型广泛应用于语音识别、中文断词 / 分词或光学字符识别，生物息学和基因组学。

发明内容

[0005] 为了克服现有技术的上述缺点与不足，本发明的目的在于提供一种基于语音识别的家庭手机通话系统，实现在家里随时随地进行接听和呼叫，极大地提高室内通讯的便捷性。

[0006] 本发明的目的通过以下技术方案实现 ：

[0007] 一种基于语音识别的家庭手机通话系统，包括总控制器、手机和多个收发机，所述总控制器与手机通过 USB 接口连接，所述总控制器与所述收发机连接 ；

[0008] 所述总控制器包括 ：

[0009] 手机控制模块，用于对手机进行控制，所述控制包括接听、拨号和读通信录 ；

[0010] 收发机控制模块，用于接收各收发机的信息、接收语音信号并对语音信号进行放大，同时用于对用户进行定位 ；

[0011] 语音识别模块，用于识别用户的语音指令。

[0012] 所述的一种基于语音识别的家庭手机通话系，还包括手机充电器，所述总控制器与手机充电器连接，所述手机充电器与手机通过 USB 接口连接。

[0013] 所述语音识别模块采用离散隐马尔可夫模型对用户的语音指令进行识别。

[0014] 所述总控制器为 MCU，采用 ARM 架构芯片。

[0015] 与现有技术相比，本发明具有以下优点和有益效果 ：

[0016] 通过本发明，用户利用声控实现远离手机的状态下随时随地的接听或者呼叫 ；本发明还带有充电器，解决了充电的时候想接听或者拨打电话都受到充电器位置的限制的问题。同时，本发明的语音识别模块在模板匹配过程中采用采用离散隐马尔可夫模型，并具有自学习功能，提高了识别的准确性。

附图说明

[0017] 图 1 为本发明的基于语音识别的家庭手机通话系统的组成示意图。

[0018] 图 2 为本发明的实施例采用的模板匹配法的流程图。

具体实施方式

[0019] 下面结合实施例及附图，对本发明作进一步地详细说明，但本发明的实施方式不限于此。

[0020] 实施例

[0021] 如图 1 所示，本实施的基于语音识别的家庭手机通话系统，包括总控制器、多个收发机、充电器和手机，所述总控制器与手机充电器连接，所述手机充电器与手机通过 USB 接口连接，所述总控制器与所述收发机连接；

[0022] 所述总控制器包括：

[0023] 手机控制模块，用于对手机进行控制，所述控制包括接听、拨号和读通信录；

[0024] 收发机控制模块，用于接收各收发机的信息、接收语音信号并对语音信号进行放大，同时用于对用户进行定位；

[0025] 语音识别模块，用于识别用户的语音指令。

[0026] 本实施例的总控制器为 MCU，采用 ARM 架构芯片；手机为搭载 android 系统手机。

[0027] 本实施例的总控制器的工作过程如下：

[0028] 总控制器中的 Linux 启动 adb（Android 调试桥）。adb 服务器会绑定手机本地的 TCP 端口 5037 并监听 adb 客户端发来的命令，所有的 adb 客户端都是用端口 5037 与 adb 服务器对话。adb 可通过如下格式发出命令 adb[-d|-e|-s<serialNumber>]<command>。adb 可通过 Shell 来加载命令如：am start-aandroid.intent.action.CALL-d tel:10086。

[0029] 在正常工作的过程中，总控制器监听各个收发机传送过来的数据，通过语音识别模块判断是否有指令输入，如判断某分机有语音输入，则将指令压入命令队列，并记录该收发机的 id。总控制器与该收发机进行单独通讯。同时，ARM 运行的操作系统上的另一进程将会根据指令队列执行指令，指令包括：接听，拨号，读通讯录。

[0030] 本实施例的语音识别模块过程包括训练阶段和识别阶段，其中第一步训练阶段的任务是建立识别基本单元的声学模型；第二步识别阶段，采用模板匹配法将接收到的语音信号与系统模型进行比较，经判决得出识别结果。本实施例初设为小词汇量识别的语音识别系统，采用的是如图 2 所示的模板匹配法，具体包括以下步骤：

[0031] 1. 预处理

[0032] 语音信号是非平稳过程的信号，但考虑到人的发音器官的运动特点，可以看作局部平稳的。因此，对语音信号进行短时分帧处理：设定语音采样率为 11kHz，最大误差为 14bit 按 256 个采样点分帧，帧移 128 个采样点。信号经过如下高通滤波器 $H(z) = 1-\beta z-1$，使其预加重，这里取 β 为 0.95。

[0033] 2. 端点检测

[0034] 准确地从背景噪声中检测出语音信号的起始点和终止点，可以减少采集的数据量，删除不含语音成分的背景噪声，从而可以大大降低语音识别处理中的计算量和处理时间，提高识别的准确性，为语音识别系统在实时处理中得以应用创造条件。为了提高起止点检测的准确性，本实施例把 HMM 方法直接用于语言信号的端点检测。具体做法是：把被测信号看作是由背景和废料（在语音处理中，习惯上把有用和无用的发音统称为“废料”）两部分

组成，而端点就是这两部分的分界点。在训练阶段，分别得出背景噪声和废料的模型参数；在测试阶段，用 Viterbi 解码方法在训练模型基础上对被测语音进行分解，求出语音的哪些帧与背景噪声匹配，哪些帧与废料匹配，从而得到端点的所在处。

[0035] 3. 特征参数提取

[0036] 把经过预处理后的信号经过一次变换，去掉冗余部分而把代表语音本质的参数提取出来称为特征参数的提取。语音识别中对特征参数的要求是：①能够有效地代表语音特征；②各阶参数之间有良好的独立性；③特征参数要计算方便，保证识别的实时实现。本实施例使用的 MFCC 参数建立在 Fourier 频谱分析的基础上，它的核心思想是利用人耳的感知特性，在语音的频谱范围内设置若干个带通滤波器，每个滤波器具备三角或者正弦的形滤波特性，计算相应滤波器组的信号能量，再通过 DCT 计算对应的倒谱系数，符合语音识别中对特征参数的要求。

[0037] MFCC 求取主要步骤是：①给每一帧语音加窗做 FFT，取出幅度；②将幅度和滤波器组中每一个三角滤波器进行 Binning 运算；③求 log，换算成对数率；④从对数率的滤波器组幅度，使用 DCT 变换求出 MFCC 系数。

[0038] 4. 模板训练及识别

[0039] HMM 语音模型由起始状态概率 λ（π，A，B）、状态转移概率（A）和观测序列概率（B）三个参数决定。π 揭示了 HMM 的拓扑结构，A 描述了语音信号随时间的变化情况，B 给出了观测序列的统计特性。HMM 语音识别的过程是：用前向后向算法，通过递推方法计算已知模型输出 O 及 λ = f（π，A，B）模型时的产生输出序列的概率 P（O | λ），然后用 Baum-Welch 算法，基于最大似然准则，对模型参数 λ（π，A，B）进行修正，最优参数 λ 的求解可表示为 $\lambda^* = \arg\max_{\lambda}\{P(O|\lambda)\}$。最后用 Viterbi 算法解出产生输出序列的最佳状态转移序列 x。所谓最佳是以 x 的最大条件后验概率为准则，即 $x = \arg\max_{x}\{P(X|O,\lambda)\}$。

[0040] 本实施例采用离散隐马尔可夫模型（DHMM）算法来实现识别。与连续隐马尔可夫模型（CHMM）算法相比，DHMM 识别精度稍低，但运算量大大降低，特别适合应用在嵌入式系统中。此时 DSP 能够独立完成语音训练并生成训练模板库，无须借助 PC 机实现训练。系统采用前向 Viterbi 算法计算各个语音模型产生采入语音的概率。

[0041] 本实施例利用 DHMM 模型，使系统具有自学习的功能。当发生误识时，系统对误识的模板进行调整，使之更适应使用者的发音习惯，以避免类似的误识再度发生。

[0042] 上述实施例为本发明较佳的实施方式，但本发明的实施方式并不受所述实施例的限制，其他的任何未背离本发明的精神实质与原理下所作的改变、修饰、替代、组合、简化，均应为等效的置换方式，都包含在本发明的保护范围之内。

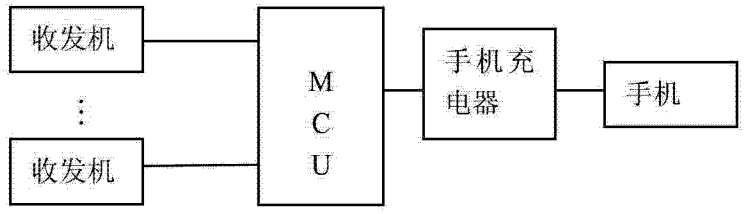

收发机
收发机
MCU
手机充电器
手机

图 1

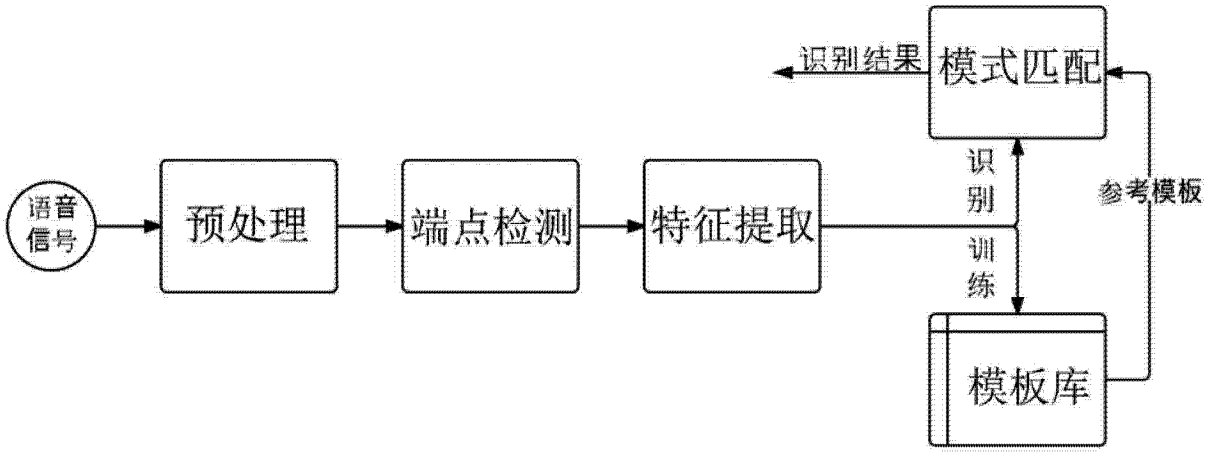

识别结果
模式匹配
语音信号
预处理
端点检测
特征提取
识别
训练
参考模板
模板库

图 2