



(12) 发明专利

(10) 授权公告号 CN 102572839 B

(45) 授权公告日 2016. 03. 02

(21) 申请号 201010603064. 2

审查员 蒋莉

(22) 申请日 2010. 12. 14

(73) 专利权人 中国移动通信集团四川有限公司
地址 610041 四川省成都市高新区高鹏大道
10 号

(72) 发明人 吴凤辉 温健军

(74) 专利代理机构 北京德琦知识产权代理有限公司 11018
代理人 张驰 宋志强

(51) Int. Cl.

H04W 12/12(2009. 01)

G10L 15/02(2006. 01)

G10L 15/08(2006. 01)

(56) 对比文件

CN 101794576 A, 2010. 08. 04,

CN 1320902 A, 2001. 11. 07,

CN 101123648 A, 2008. 02. 13,

US 2009177466 A1, 2009. 07. 09,

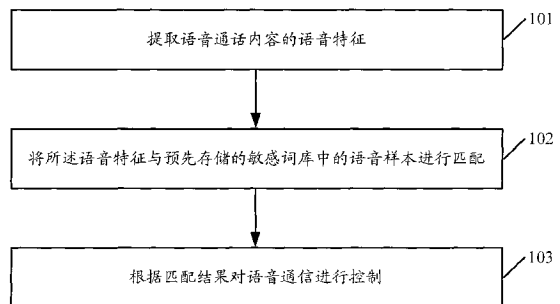
权利要求书2页 说明书6页 附图2页

(54) 发明名称

一种控制语音通信的方法和系统

(57) 摘要

本发明实施例公开了一种控制语音通信的方法和系统。该方法包括：预先存储敏感词库，该敏感词库中存储有语音样本；提取语音通话内容的语音特征；将所述语音特征与所述敏感词库中的语音样本进行匹配，根据匹配结果对语音通信进行控制。应用本发明能够对语音通信实现精确控制。



1. 一种控制语音通信的方法,其特征在于,该方法包括:

提取语音通话内容的语音特征;

将所述语音特征与敏感词库中的语音样本进行匹配,根据匹配结果对语音通信进行控制;

将所述语音特征与所述敏感词库中的语音样本进行匹配包括:

将所述语音特征作为学习矢量量化(LVQ)神经网络的输入矢量,利用该LVQ网络识别所述语音特征是否与所述语音样本匹配;

其中,所述LVQ神经网络是通过将所述敏感词库中的语音样本作为输入矢量进行训练得到的;

利用该LVQ网络识别所述语音特征是否与所述语音样本匹配包括:

将所述语音特征作为输入矢量 x ,输入所述LVQ神经网络的传输函数 $f(x) = \frac{1 - \exp(-x)}{1 + \exp(-x)}$,根据神经元的输出矢量 $f(x)$ 识别所述语音特征是否与所述语音样本

匹配;

所述提取语音通话内容的语音特征包括:

从采集的语音信号中检测出语音通话的起点和终点,提取所述起点和所述终点之间的语音信号的语音特征。

2. 根据权利要求1所述的方法,其特征在于,所述根据匹配结果对语音通信进行控制包括:

根据匹配成功的语音样本个数、类型和内容中的任意一项或多项,按照预先设定的策略对语音通信进行控制。

3. 根据权利要求2所述的方法,其特征在于,所述按照预先设定的策略对语音通信进行控制包括:

中断所述语音通信、或进行语音提醒、或屏蔽所述语音通信、或将所述语音通信自动转接至指定号码。

4. 一种控制语音通信的系统,其特征在于,该系统包括敏感词库、语音检测模块和控制模块;

所述敏感词库,用于存储语音样本;

所述语音检测模块,用于提取语音通话内容的语音特征,将所述语音特征与所述敏感词库中的语音样本进行匹配;

所述控制模块,用于根据所述语音检测模块的匹配结果对语音通信进行控制;

所述语音检测模块包括LVQ神经网络,用于以所述语音特征为输入矢量,识别所述语音特征是否与所述敏感词库中的语音样本匹配;

其中,所述LVQ神经网络是通过将所述敏感词库中的语音样本作为输入矢量进行训练得到的;

所述LVQ神经网络将所述语音特征作为输入矢量 x ,输入传输函数 $f(x) = \frac{1 - \exp(-x)}{1 + \exp(-x)}$,根据神经元的输出矢量 $f(x)$ 识别所述语音特征是否与所述语音样本匹配;

所述语音检测模块包括端点检测单元、语音特征提取单元和识别单元；
所述端点检测单元，用于从采集的语音信号中检测出语音通话的起点和终点；
所述语音特征提取单元，用于提取所述起点和所述终点之间的语音信号的语音特征；
所述识别单元，用于将所述语音特征与所述敏感词库中的语音样本进行匹配。

5. 根据权利要求 4 所述的系统，其特征在于，

所述控制模块，用于根据成功的语音样本个数、类型和内容中的任意一项或多项，按照预先设定的策略对语音通信进行控制。

6. 根据所述权利要求 5 所述的系统，其特征在于，

所述控制模块，用于中断所述语音通信、或进行语音提醒、或屏蔽所述语音通信、或将所述语音通信自动转接至指定号码。

一种控制语音通信的方法和系统

技术领域

[0001] 本发明涉及通信技术领域,尤其涉及一种控制语音通信的方法和系统。

背景技术

[0002] 语音通信系统一般包括通信终端、传输网络、交换机等基本单元,下面以无线移动通信系统中的语音通信系统为例,对语音通信过程进行介绍。

[0003] 无线移动通信系统进行语音通信的过程包括:

[0004] 当用户发起呼叫请求后,主叫移动台首先通过随机接入信道,向基站发出接入网络的请求。基站接收后,根据广播信道(BCCH)所发布的信道忙、闲信息,为用户寻找出一条合适的业务信道(TCH);再通过寻呼信道(PCH)和允许接入信道(AGCH)去寻找该信道,找到后立即通知该移动台信道已经分配。然后,基站将被叫用户号码,通过移动通信网络中的交换机,传输给地面公用电话网络,找到被叫用户的话机;再经振铃、摘机,完成了通信线路的建立。

[0005] 通信线路建立后,主叫移动台将语音信号变换成电信号传输到移动通信网络中的基站,再由基站将代表语音的电信号变成电磁频谱,通过移动通信网络中的交换机将电磁频谱传送到受话人的电信网络中,受话人的通信设备接收到无线电波,转换成语音信号。

[0006] 目前,语音通信系统在对语音通信进行控制时,通常采取的方法是,识别主叫用户号码或者被叫用户号码,根据识别结果控制语音通信,例如,屏蔽特定主叫用户对某一被叫用户的呼叫。

[0007] 然而,目前这种语音控制方法无法对语音通信实现精确控制,适用场景有限,远不能满足当前的通信需求。例如,目前电信诈骗类通话较多,由于诈骗源的用户号码无法事先获知,因此无法通过现有的语音控制方法来控制电信诈骗类通话。

发明内容

[0008] 有鉴于此,本发明提供了一种控制语音通信的方法和系统,以便对语音通信实现精确控制。

[0009] 本发明的技术方案具体是这样实现的:

[0010] 一种控制语音通信的方法,该方法包括:

[0011] 提取语音通话内容的语音特征;

[0012] 将所述语音特征与敏感词库中的语音样本进行匹配,根据匹配结果对语音通信进行控制。

[0013] 一种控制语音通信的系统,该系统包括敏感词库、语音检测模块和控制模块;

[0014] 所述敏感词库,用于存储语音样本;

[0015] 所述语音检测模块,用于提取语音通话内容的语音特征,将所述语音特征与所述敏感词库中的语音样本进行匹配;

[0016] 所述控制模块,用于根据所述语音检测模块的匹配结果对语音通信进行控制。

[0017] 由上述技术方案可见,本发明通过提取语音通话内容的语音特征,将该语音特征与预先存储的敏感词库中的语音样本进行匹配,根据匹配结果对语音通信进行控制,可以实现对语音通话内容的监测,根据监测到的语音通话内容来对语音进行控制,因而能够对语音通信实现精确控制。

[0018] 本发明的方法和系统可以适用于语音通信的任何场景,例如,对于电信诈骗类通话,可以通过在敏感词库中预先存储电信诈骗类通话中出现频率较高的语音样本,然后如果某一用户开通了本发明方法或系统对应的语音检测服务,则监测其他用户与该某一用户的语音通话内容中是否出现了电信诈骗类通话中出现频率较高的语音样本,根据监测结果进行语音提醒等语音控制服务。

附图说明

[0019] 图 1 是本发明提供的控制语音通信的方法流程图。

[0020] 图 2 是 LVQ 神经网络的组成示意图。

[0021] 图 3 是本发明提供的控制语音通信的系统组成示意图。

[0022] 图 4 是语音控制系统在移动通信系统中的部署示意图。

具体实施方式

[0023] 图 1 是本发明提供的控制语音通信的方法流程图。

[0024] 如图 1 所示,该方法包括:

[0025] 步骤 101,提取语音通话内容的语音特征。

[0026] 步骤 102,将所述语音特征与预先存储的敏感词库中的语音样本进行匹配。

[0027] 步骤 103,根据匹配结果对语音通信进行控制。

[0028] 其中,所述敏感词库中存储有敏感词的语音样本,根据应用场景的不同,敏感词库中可以存储有不同的语音样本,也可以开辟不同的存储空间分别存储不同应用场景下的语音样本。

[0029] 例如,当需要监控电信诈骗类通话时,可以在敏感词库中预先存储电信诈骗类通话中出现频率较高的语音样本。

[0030] 在提取语音通话内容的语音特征时,为了提高语音特征提取的速度和准确性,本发明提出,首先对采集的语音信号进行端点检测,剔除通话空白段的干扰,然后再提取语音特征,换言之,先从采集的语音信号中检测出语音通话的起点和终点,提取所述起点和所述终点之间的语音信号的语音特征。

[0031] 为了进一步提高语音特征提取的速度和准确性,还可以进行在语音特征提取之前进行其他预处理,例如进行降噪处理。

[0032] 本发明将语音特征与敏感词库中的语音样本进行匹配的方法可以为,将所述语音特征作为学习矢量量化(Learning Vector Quantization, LVQ)神经网络的输入矢量,利用该 LVQ 网络识别所述语音特征是否与所述语音样本匹配。其中,所述 LVQ 神经网络是通过将所述敏感词库中的语音样本作为输入矢量进行训练得到的。

[0033] 下面对 LVQ 神经网络在本发明中的应用进行详细介绍:

[0034] 图 2 是 LVQ 神经网络的组成示意图。

[0035] 如图 2 所示, LVQ 神经网络由三层神经元组成: 输入层、隐含层和输出层。

[0036] LVQ 神经网络在输入层和隐含层间为完全连接, 而在隐含层和输出层间为部分连接, 每个输出神经元与隐含层神经元的不同组相连接。隐含层和输出神经元之间的连接权值为固定值 1。输入和隐含层间神经元连接权值建立参考矢量的分量, 每个隐含神经元指定一个参考矢量。网络训练时, 这些权值被修改。隐含神经元和输出神经元都具有二进制输出值。当某个输入模型被输入网络, 参考矢量最接近输入模式的隐含神经元因获得激发而赢得竞争, 此隐含神经元产生一个‘1’, 其他隐含神经元被迫产生‘0’。与获得竞争的隐含神经元相连的输出神经元输出也为 1, 因此获得竞争, 其他输出神经元均产生‘0’。每个输出神经元表示不同的模式或类别。

[0037] 本发明通过寻找 LVQ 神经网络的输入 / 输出关系建立自动语音识别模型, 具体流程如下:

[0038] (1) 输入矢量和目标矢量设计

[0039] 设计一组输入矢量及输入矢量对应的目标矢量, 这两组质量直接决定网络 (定义为 net) 的输入输出关系, 设计的优劣直接影响语音识别的效果。

[0040] 具体地, 本发明中, 输入矢量选取归一化的敏感词样本特征参数, 目标矢量根据敏感词样本数量进行目标矢量的设计, 尽量将不同目标矢量设计为无关、正交。

[0041] (2) 网络创建及训练

[0042] 创建网络模型, 设计初始化连接权值。用设计好的输入矢量作为 LVQ 神经网络的输入, 目标矢量作为 LVQ 神经网络的输出来训练创建的 LVQ 神经网络。通过反复的训练, 直到输入矢量落入目标分类对应的矢量之中。终止训练一般是达到分类准确度预定的阈值或训练的次数超限, 其中在训练的次数超限时终止训练一般是出于网络计算速度的考虑。

[0043] (3) 语音识别

[0044] 将待识别的语音样本数据输入训练好的 LVQ 神经网络, LVQ 神经网络根据决策功能对输入数据进行分类, 输出即为识别结果。

[0045] 具体地, 本发明中, 将从语音通话内容中提取的语音特征作为输入矢量, 输入到预先训练好的 LVQ 神经网络, 该 LVQ 神经网络对该语音特征进行分类, 即与敏感词库中的各个语音样本进行匹配, 如果与某个语音样本匹配, 则将该语音特征属于该语音样本所属的类别。

[0046] 采用 LVQ 神经网络进行语音特征识别时, 可以不将语音特征组成的输入矢量进行归一化和正交化处理, 只需要计算输入矢量和竞争层之间的距离, 即可实现语音识别。当然, 为了提高语音识别速度, 优选地, 将语音特征参数进行归一化处理后再作为输入矢量输入 LVQ 神经网络。

[0047] 另外, 为了进一步提高 LVQ 神经网络进行语音特征识别的速度, 本发明还提出在

LVQ 神经网络中采用 $f(x) = \frac{1 - \exp(-x)}{1 + \exp(-x)}$ 作为传输函数, 其中, x 是神经元输入矢量, 即是归一

化的语音特征参数, $f(x)$ 是神经元输出, 即语音特征是否与敏感词库中的语音样本匹配, 以及与哪类语音样本匹配。

[0048] 采用 $f(x) = \frac{1 - \exp(-x)}{1 + \exp(-x)}$ 作为传输函数时, 本发明所涉及的 LVQ 神经网络训练和识

别的速度都较快。

[0049] 本申请人对采用 $f(x) = \frac{1 - \exp(-x)}{1 + \exp(-x)}$ 作为传输函数的 LVQ 神经网络和采用现有 Sigmoid 传输函数的 LVQ 神经网络, 在相同环境下采用相同的测试样本分别进行识别, 二者的识别速度对比参见表一:

[0050] 表一中, 传输函数 1 所在列的时间表示采用 $f(x) = \frac{1 - \exp(-x)}{1 + \exp(-x)}$ 作为传输函数的 LVQ 神经网络识别测试样本所需要的时间, 传输函数 2 所在列的时间表示采用现有 Sigmoid 传输函数的 LVQ 神经网络识别测试样本所需要的时间。

[0051] 表一

测试样本序号	传输函数 1	传输函数 2
1	110ms	125ms
2	82ms	83ms
3	85	94ms
4	104ms	118ms
5	101ms	109ms
[0052] 6	94.5ms	95.6ms
7	95.0ms	96.3ms
8	30ms	52ms
9	50ms	95.1
10	88ms	95.3

[0053] 由表一可见, 在本申请的应用场景下, 采用 $f(x) = \frac{1 - \exp(-x)}{1 + \exp(-x)}$ 作为传输函数的 LVQ 神经网络的识别速度优于采用现有 Sigmoid 传输函数的 LVQ 神经网络的识别速度。

[0054] 本发明根据匹配结果对语音通信进行控制时, 可以根据匹配成功的语音样本个数、类型和内容中的任意一项或多项, 按照预先设定的策略对语音通信进行控制。例如, 可以中断所述语音通信、或进行语音提醒、或屏蔽所述语音通信、或将所述语音通信自动转接至指定号码。其中, 语音样本的类型可以根据业务需要, 从多个角度确定, 例如, 将语音样本分为男生、女生或者分为老人、儿童等等。

[0055] 可见, 本发明通过建立敏感词库, 利用基于自定义的指数函数的 LVQ 神经网络建立自动语音识别模型, 将用户的通话通过语音识别程序和敏感词库中的声音样本进行对比, 若发现用户在通话中涉及到了敏感词库中的信息时, 根据用户事先设定的管理策略实时的给予提醒、告警、自动转接指定号码、自动语音屏蔽等语音控制服务, 从而提高用户对企业的感知度、认可度。本发明可推广到任何语音通信中。

[0056] 图 3 是本发明提供的控制语音通信的系统组成示意图。

[0057] 如图 3 所示, 该系统包括敏感词库 301、语音检测模块 302 和控制模块 303。

[0058] 敏感词库 301, 用于存储语音样本。

[0059] 语音检测模块 302, 用于提取语音通话内容的语音特征, 将所述语音特征与所述敏感词库中的语音样本进行匹配。

[0060] 控制模块 303,用于根据语音检测模块 302 的匹配结果对语音通信进行控制。

[0061] 其中,语音检测模块 302 包括端点检测单元、语音特征提取单元和识别单元。

[0062] 所述端点检测单元,用于从采集的语音信号中检测出语音通话的起点和终点。

[0063] 所述语音特征提取单元,用于提取所述起点和所述终点之间的语音信号的语音特征。

[0064] 所述识别单元,用于将所述语音特征与所述敏感词库中的语音样本进行匹配。

[0065] 所述识别单元包括 LVQ 神经网络,用于以所述语音特征为输入矢量,识别所述语音特征是否与所述敏感词库中的语音样本匹配;其中,所述 LVQ 神经网络是通过将所述敏感词库中的语音样本作为输入矢量进行训练得到的。

[0066] 所述 LVQ 神经网络将所述语音特征作为输入矢量,输入传输函数 $f(x) = \frac{1 - \exp(-x)}{1 + \exp(-x)}$,根据神经元的输出矢量 $f(x)$ 识别所述语音特征是否与所述语音样本匹配。

[0067] 其中,控制模块 303,用于根据成功的语音样本个数、类型和内容中的任意一项或多项,按照预先设定的策略对语音通信进行控制。

[0068] 具体地,控制模块 303 可以用于中断所述语音通信、或进行语音提醒、或屏蔽所述语音通信、或将所述语音通信自动转接至指定号码。

[0069] 图 3 所示系统的工作流程及原理可以描述为:当声音通过一个转换装置输入计算机内部、并以数位方式储存后,语音识别程序便开始以输入的语音样本与事先储存好的语音样本进行对比工作(即提取输入的语音样本的语音特征,然后输入 LVQ 神经网络进行语音识别)。对比工作完成之后,计算机就会算出最匹配、接近的语音样本序号,从而得知输入计算机的声音是什么意思,进而执行对应的命令。

[0070] 下面以图 4 为例,对图 3 提供的语音控制系统在移动通信系统中的部署情况进行示例性说明。

[0071] 图 4 是语音控制系统在移动通信系统中的部署示意图。

[0072] 如图 4 所示,在移动通信系统的交换网中部署有交换机、语音检测服务器集群和敏感词库服务器。

[0073] 其中,数据传输板置于交换机内,外部可以和多台 PC 或工控机相连,通过交换机的级联使得和外部连接的 PC 或工控机没有特别限制。这样交换机的内部 I/O 总线(类似于 PC 机的 PCI 或 ISA 总线)就不再传输实时数据,而只负责非实时性的管理和信令数据,语音或其它实时数据从外部 PC 直接传入交换机的高速数据传输板。

[0074] 图 3 所示系统中的语音检测模块 302 部署在语音检测服务器集群中,该语音检测服务器集群是语音处理的核心部分,主要完成语音信号预处理、语音特征参数提取、语音样本匹配等重要工作。

[0075] 其中,语音信号预处理主要包括降噪、端点检测等。端点检测就是从包含语音的一段信号中确定出语音的起点以及终点。有效的端点检测不仅能使处理时间减到最小,而且能排除无声段的噪声干扰,从而使语音识别系统(例如 LVQ 神经网络)具有良好的识别性能。

[0076] 语音特征参数提取是指从语音信号提取一组能够描述语音信号本质特征参数的

过程。

[0077] 语音样本匹配的工作也可以在语音检测服务器集群中完成,具体为,将提取的语音特征参数与敏感词库服务器中的语音样本进行匹配,并输出匹配结果。

[0078] 敏感词库服务器中存储有语音样本,其中,该语音样本可以是预先处理好以后存储在敏感词库服务器中的,也可以是由语音检测服务器实时采集,然后将采集后的语音样本存储在敏感词库服务器中,例如采集某一场景下的通话内容,将该场景下的通话内容的语音特征作为语音样本存储在敏感词库服务器中。

[0079] 其中,实时采集终端通话内容的语音采集模块一般部署在语音检测服务器集群中,或者部署在该语音检测服务器集群的前端,将采集的语音信号输入到所述语音检测服务器集群中。

[0080] 根据语音检测服务器集群的匹配进行语音控制的模块可以部署在语音检测服务器集群中,也可以单独部署服务器或者部署在其他服务器中,根据匹配结果进行语音控制。

[0081] 其中,根据匹配结果进行语音控制所依据的策略可以部署在专门的策略配置库中,也可以采用编程等方式直接部署在用于进行语音控制的模块中。

[0082] 总之,图 4 所示的部署情况仅为示例,并非用于限制本发明。

[0083] 以上所述仅为本发明的较佳实施例而已,并不用以限制本发明,凡在本发明的精神和原则之内,所做的任何修改、等同替换、改进等,均应包含在本发明保护的范围之内。

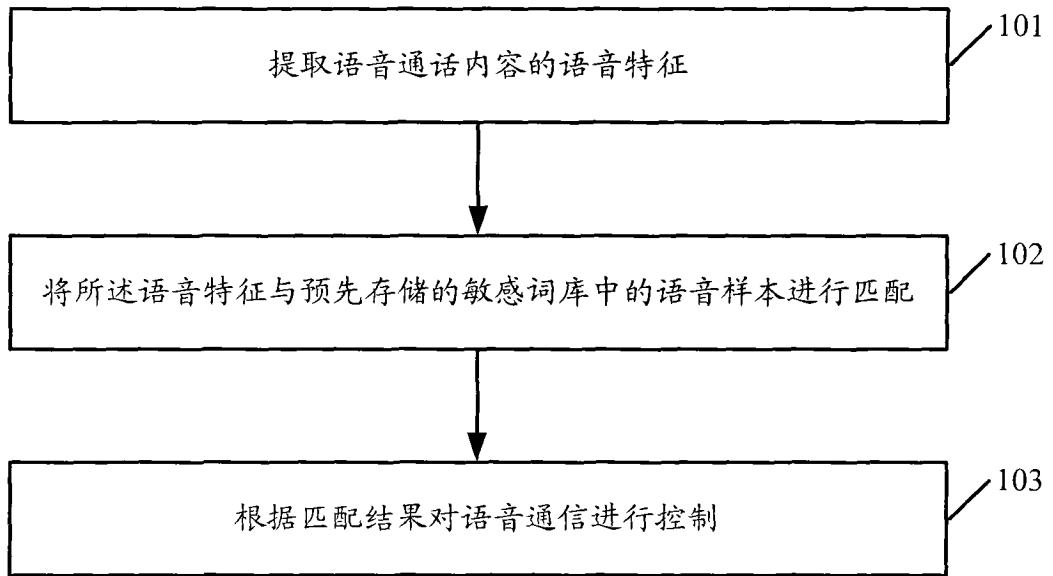


图 1

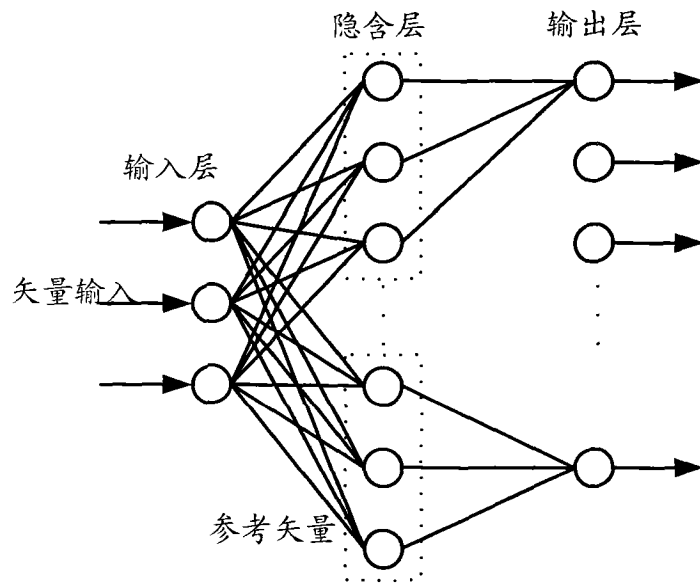


图 2

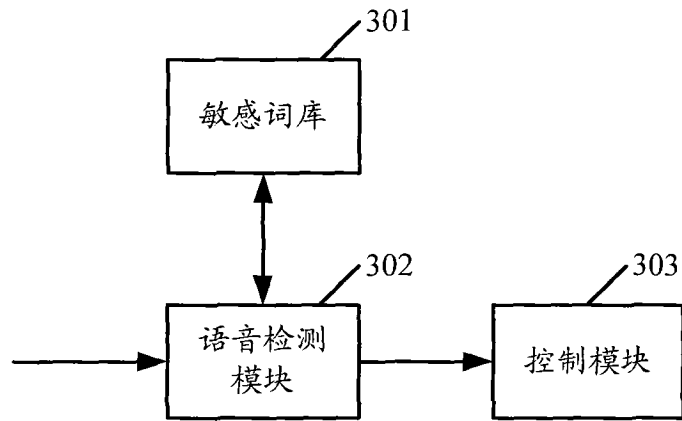


图 3

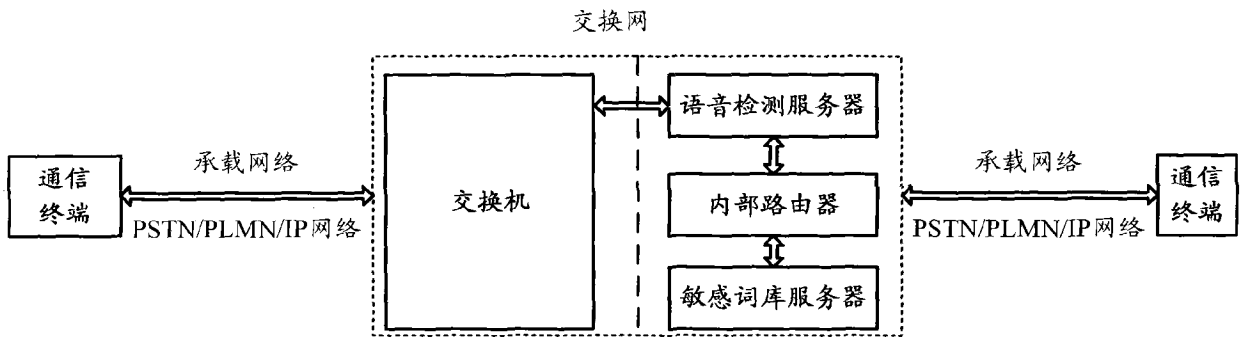


图 4