



(12) 发明专利

(10) 授权公告号 CN 101160617 B

(45) 授权公告日 2011.09.14

(21) 申请号 200680012239.4

(22) 申请日 2006.09.26

(66) 本国优先权数据
200510100150.0 2005.09.29 CN

(85) PCT申请进入国家阶段日
2007.10.15

(86) PCT申请的申请数据
PCT/CN2006/002531 2006.09.26

(87) PCT申请的公布数据
W02007/036148 ZH 2007.04.05

(73) 专利权人 华为技术有限公司
地址 518129 广东省深圳市龙岗区坂田华为
总部办公楼

(72) 发明人 宋永刚 刘承东

(74) 专利代理机构 北京凯特来知识产权代理有
限公司 11260

代理人 郑立明

(51) Int. Cl.
G10L 15/00 (2006.01)

(56) 对比文件

US 2005/0049858 A1, 2005.03.03,
US 5638425 A, 1997.06.10,
CN 1489306 A, 2004.04.14,
CN 1851805 A, 2006.10.25,
US 6823302 B1, 2004.11.23,

审查员 游晓梅

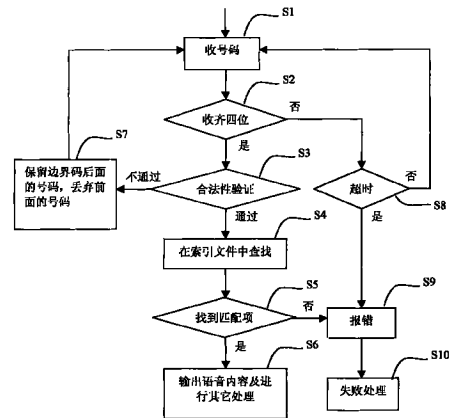
权利要求书 2 页 说明书 7 页 附图 3 页

(54) 发明名称

对限定范围的语音进行自动识别的方法、系统和设备

(57) 摘要

一种对限定范围的语音进行自动识别的方法、系统和设备,其中:建立有限集合中的语音文件与标记信息的对应关系,在语音文件中插入标记信息或者用标记信息替换语音文件中的语音信息,发送端发送该语音文件;接收端接收到该语音文件,提取有效的标记信息并根据该有效的标记信息检索出对应的语音信息。



1. 一种对限定范围的语音进行自动识别的方法,其特征在于,包括如下步骤:

发送端发送包含有标记信息的、需要播放的语音文件,所述标记信息包括基本标识码和边界码,基本标识码用于标识语音内容,边界码用于标识基本标识码的边界;

接收端接收语音文件,提取有效的标记信息,并根据提取到的有效的标记信息从预设的对应关系中检索对应的语音内容;

所述对应关系为有限集合中的语音内容与标记信息的对应关系;

所述标记信息插入在语音文件中,或者所述标记信息替换了语音文件中的音频信息。

2. 根据权利要求 1 所述的方法,其特征在于,所述边界码设置在所述基本标识码后。

3. 根据权利要求 2 所述的方法,其特征在于,提取有效的标记信息具体包括:

接收端在预定时间间隔内接收标记信息,当接收端在预定时间间隔内接收到位数齐全的标记信息后,进行合法性验证,在验证不通过时,保留边界码后面的基本标识码,丢弃边界码前面的基本标识码,继续上述接收标记信息、合法性验证的过程,直到验证通过或者超时;

当验证通过时,提取到有效的标记信息;当超时时,进行出错处理。

4. 根据权利要求 1 所述的方法,其特征在于,所述边界码设置在所述基本标识码前。

5. 根据权利要求 4 所述的方法,其特征在于,提取有效的标记信息具体包括:

在预定时间间隔内,接收端收取到边界码后,接收基本标识码,在基本标识码收齐后,进行合法性验证,如果验证未通过,则丢弃收到的标记信息,继续上述收取边界码、基本标识码和合法性验证的过程,直到验证通过或超时;

当验证通过时,提取到有效的标记信息;当超时时,进行出错处理。

6. 根据权利要求 1 所述的方法,其特征在于,所述基本标识码前设置有第一边界码,所述基本标识码后设置有第二边界码。

7. 根据权利要求 6 所述的方法,其特征在于,提取有效的标记信息具体包括:

在预定时间间隔内,接收端收取标记信息,在收取到第一边界码和第二边界码后,进行合法性验证,如果验证未通过,则丢弃收取到的标记信息,继续上述收取标记信息、合法性验证的过程,直到验证通过或超时;

当验证通过时,提取到有效的标记信息;当超时时,进行出错处理。

8. 一种对限定范围的语音进行自动识别的系统,包括语音发送端和语音接收端,其特征在于,语音发送端中设置有语音存储模块和语音发送模块,语音接收端中设置有对应关系存储模块、提取模块和检索模块;

语音存储模块:用于存储语音文件,所述语音文件中插入了对应的标记信息,或者语音文件中的音频信息替换为标记信息,所述标记信息包括基本标识码和边界码,基本标识码用于标识语音内容,边界码用于标识基本标识码的边界;

语音发送模块:用于发送语音存储模块中存储的语音文件;

对应关系存储模块:用于存储语音内容与标记信息的对应关系;

提取模块:用于从接收到的语音文件中提取有效的标记信息,并传输至检索模块;

检索模块:用于根据接收的标记信息从所述对应关系存储模块中存储的所述对应关系中检索语音内容。

9. 一种语音发送端设备,其特征在于,所述语音发送端设备中设置有语音存储模块和

语音发送模块；

语音存储模块：用于存储语音文件，所述在语音文件中插入标记信息，或者语音文件中的音频信息替换为标记信息，所述标记信息是与有限集合中的语音内容相应的标记信息，所述标记信息包括基本标识码和边界码，基本标识码用于标识语音内容，边界码用于标识基本标识码的边界；

语音发送模块：用于发送语音存储模块中存储的语音文件。

10. 一种语音接收端设备，其特征在于，所述语音接收端设备中设置有对应关系存储模块、提取模块和检索模块；

对应关系存储模块：用于存储语音内容与标记信息的对应关系，所述标记信息是与有限集合中的语音内容相应的标记信息，所述标记信息包括基本标识码和边界码，基本标识码用于标识语音内容，边界码用于标识基本标识码的边界；

提取模块：用于从接收到的语音文件中提取有效的标记信息，并传输至检索模块；

检索模块：用于根据接收的标记信息从所述对应关系存储模块中存储的所述对应关系中检索语音内容。

对限定范围的语音进行自动识别的方法、系统和设备

技术领域

[0001] 本发明涉及语音识别技术,具体涉及一种用于电话智能语音系统中对限定范围的语音进行自动识别的方法、系统、语音发送端设备和语音接收端设备。

[0002] 发明背景

[0003] 在进行PSTN(Public Switched Telephony Network,公共交换电话网)交换机、软交换、接入网关等相关通讯产品的测试时,需要测试某些需要智能语音的业务,比如在用电话拨打了一个错误的号码时,会听到“您的号码有误,请查证后再拨”,在登记了一个新的补充业务后,会听到“您申请的新业务已登记,欢迎使用”等等这样的语音提示。对于绝大多数卡类业务和电脑值班的业务,都广泛使用语音提示,以提示用户或指导用户的后续操作。在进行通讯产品的手工测试时,这些语音提示是靠人去听并进行正确性判断的,但是手工测试的工作量比较大,并且很多是重复工作,因此,越来越多的语音业务需要进行自动化测试。但在进行自动化测试时,需要解决的一个主要技术问题就是如何识别这些语音,以保证测试的正确性与完备性。本申请文件中的语音识别即语音内容识别。

[0004] 现有技术之一的技术方案为:ASR(Automatic Speech Recognition,自动语音识别)技术。ASR技术的基本原理是:首先,将语音转换成为文本,然后根据文本进行语音内容的识别。该技术方案缺点是:1)实现过程复杂,实现成本比较高。2)ASR技术的识别速度比较慢,对于智能语音业务,一般都需要在有限的时间内识别语音内容并进行后续拨号等处理,一旦超时,智能语音业务便无法按照正常流程进行。所以从成本和性能角度来说,ASR技术不适用于通讯设备的语音识别业务。

[0005] 现有技术之二的技术方案的基本原理为:从原始的语音文件中产生含有所有语音的语音特征量的语音模板,然后加载语音模板;检测待识别语音流的语音起止点,提取起止点之间语音的语音特征量,比较所述的待识别语音流的语音特征量和所述的语音模板的语音特征量,进行语音拆分和识别,从而得到识别结果。该技术方案的缺点是:1)此方案对语音质量的要求很高,当语音质量有一定失真时,语音识别的正确率便会有明显下降。2)按此方案,识别某条语音需要提取语音起、止点之间的语音特征量,即要识别语音,至少要等到语音播放到终止点,也就是说,在语音播放完成后再进行语音识别,识别效率不够高。

发明内容

[0006] 针对现有技术的上述缺陷,本发明提供一种成本低、识别效率高、识别准确度高的语音自动识别方法、系统和设备,本发明提供的技术方案可适用于限定范围的语音识别。

[0007] 本发明解决其技术问题所采用的技术方案是:构造一种对限定范围的语音进行自动识别的方法,包括如下步骤:

[0008] 发送端发送包含有标记信息的、需要播放的语音文件;

[0009] 接收端接收语音文件,提取有效的标记信息,并根据提取到的有效的标记信息从预设的对应关系中检索对应的语音内容;

[0010] 所述对应关系为有限集合中的语音内容与标记信息的对应关系;

[0011] 所述标记信息插入在语音文件中,或者所述标记信息替换了语音文件中的音频信息。

[0012] 下述方法的技术方案为可选技术方案。

[0013] 所述标记信息包括基本标识码和边界码,基本标识码用于标识语音内容,边界码用于标识基本标识码的边界。

[0014] 所述边界码设置在所述基本标识码后。

[0015] 提取有效的标记信息具体包括:

[0016] 接收端在预定时间间隔内接收标记信息,当接收端在预定时间间隔内接收到位数齐全的标记信息后,进行合法性验证,在验证不通过时,保留边界码后面的基本标识码,丢弃边界码前面的基本标识码,继续上述接收标记信息、合法性验证的过程,直到验证通过或者超时;

[0017] 当验证通过时,提取到有效的标记信息;当超时,进行出错处理。

[0018] 所述边界码设置在所述标识码前。

[0019] 提取有效的标记信息具体包括:

[0020] 在预定时间间隔内,接收端收取到边界码后,接收基本标识码,在基本标识码收齐后,进行合法性验证,如果验证未通过,则丢弃收到的标记信息,继续上述收取边界码、基本标识码和合法性验证的过程,直到验证通过或超时;

[0021] 当验证通过时,提取到有效的标记信息;当超时,进行出错处理。

[0022] 所述基本标识码前设置有第一边界码,所述基本标识码后设置有第二边界码。

[0023] 提取有效的标记信息具体包括:

[0024] 在预定时间间隔内,接收端收取标记信息,在收取到第一边界码和第二边界码后,进行合法性验证,如果验证未通过,则丢弃收取到的标记信息,继续上述收取标记信息、合法性验证的过程,直到验证通过或超时;

[0025] 当验证通过时,提取到有效的标记信息;当超时,进行出错处理。

[0026] 上述技术方案中的标记信息为DTMF标识号码。

[0027] 本发明还提供一种对限定范围的语音进行自动识别的系统,包括语音发送端和语音接收端,语音发送端中设置有语音存储模块和语音发送模块,语音接收端中设置有对应关系存储模块、提取模块和检索模块;

[0028] 语音存储模块:用于存储语音文件,所述语音文件中插入了对应的标记信息,或者语音文件中的音频信息替换为标记信息;

[0029] 语音发送模块:用于发送语音存储模块中存储的语音文件;

[0030] 对应关系存储模块:用于存储所述语音内容与标记信息的对应关系;

[0031] 提取模块:用于从接收到的语音文件中提取有效的标记信息,并传输至检索模块;

[0032] 检索模块:用于根据接收的标记信息从所述对应关系中检索语音内容。

[0033] 本发明还提供一种语音发送端设备,所述语音发送端设备中设置有语音存储模块和语音发送模块;

[0034] 语音存储模块:用于存储语音文件,所述语音文件中插入了对应的标记信息,或者语音文件中的音频信息替换为标记信息;

[0035] 语音发送模块 :用于发送语音存储模块中存储的语音文件。

[0036] 本发明另外提供一种语音接收端设备,所述语音接收端设备中设置有对应关系存储模块、提取模块和检索模块 ;

[0037] 对应关系存储模块 :用于存储所述语音内容与标记信息的对应关系 ;

[0038] 提取模块 :用于从接收到的语音文件中提取有效的标记信息,并传输至检索模块 ;

[0039] 检索模块 :用于根据接收的标记信息从所述对应关系中检索语音内容。

[0040] 实施本发明的自动语音识别方法、系统和设备,具有以下有益效果 :花费较低的成本就能实现对限定范围内的语音的自动化识别 ;语音识别的效率高,无需等到语音播放完成即可识别出结果 ;语音识别准确度高,通过降低信号失真和增强终端识别能力,即使语音质量不够好,也能够进行准确的语音识别,保证了语音识别的准确度。

[0041] 附图简要说明

[0042] 图 1 是下一代网络组网结构示意图 ;

[0043] 图 2 是本发明实施例的对限定范围的语音进行自动识别的第一种方法的流程图 ;

[0044] 图 3 是本发明实施例的对限定范围的语音进行自动识别的第二种方法的流程图 ;

[0045] 图 4 是本发明实施例的对限定范围的语音进行自动识别的第三种方法的流程图。

[0046] 实施本发明的方式

[0047] 本发明提供一种简单实用的语音识别方法,实现了对限定范围的语音内容(即需要识别的语音是确定的若干条语音,不是任意语音)的高效和准确的识别。

[0048] 下面结合图 1 对本发明的一种技术方案进行描述。

[0049] 图 1 示出了下一代网络(Next Generation Network, NGN)的一种基本应用组网图。在图 1 中,SoftSwitch(软交换)为 NGN 整网的交换控制中心,SoftSwitch 通过媒体网关控制协议(如 H. 248)对各接入媒体网关(Access Gateway, AG)进行控制 ;AG 用于接入电话用户,并进行语音的分组化处理 ;媒体资源服务器(Media Resource Server, MRS)用于给用户提供各种媒体资源,比如,用户在使用中听到的一些语音提示(如“线路正忙,请稍后再拨”)就是由 MRS 播放、并经过 IP 网络传送给用户的。本实施例所述的对限定范围的语音识别,就是指对 MRS 所播放的语音内容的识别。由于一般语音包含了各种不同的频率及其它频谱特征,所以直接去分析语音的信息特征是复杂而困难的,要做到简易而高效的语音识别,就需要对待识别的语音特征进行充分的简化。本实施例的具体实施方法为,修改媒体资源服务器上的语音文件,在每条语音的前面插入对应的 4 位双音多频(Dual Tone Multi-Frequency, DTMF)标识码,实际使用时只检测 DTMF 号码,然后根据预定义的对对应关系,确定语音的内容。DTMF 号码是一种双音多频的号码定义方案,每个号码对应着高低两个频率,之所以选择使用 DTMF 号码,是因为 DTMF 是一种通用的号码定义标准,能够被绝大多数仪器和设备检测和识别,同时,在 VoIP(Voice over IP)网络中,为了降低网络丢包、抖动等对 DTMF 号码失真的影响,还有专门用于保证 DTMF 传送质量的协议如 RFC2833,可使 DTMF 信号的失真情况降到最低,同时增强语音识别终端的识别能力,主要是使识别终端能够最大程度地适应 DTMF 信号的频率偏移,也就是能够对 DTMF 信号的高、低频的允许偏移量进行灵活设置。在使用 DTMF 进行语音识别时,能够很好地适应网络质量不好时的情况,提高了识别的准确度。

[0050] 具体实现方法如下：

[0051] 1、确定语音文件。在 MRS 上，每条语音内容都对应着一个语音文件，比如，“线路正忙，请稍后再拨”这条语音内容，对应的语音文件为：HW0010006.chi。找到每条需要处理的语音内容对应的文件（可从产品提供的配套文件中得到语音文件对应关系和原始语音文件）。

[0052] 2、为每条语音内容定义一个特定的标识码即标记信息，如 DTMF 标识码。标识码的最小位数要根据待处理的语音个数来定，如果待处理的语音内容在 1000 个以内，则标识码可定义为 3 位。

[0053] 3、修改语音文件。根据前面为每条语音内容所定义的标识码，使用相关的工具对原始语音文件进行修改和编辑，在其中插入相应的 DTMF 号码。在所有的语音文件均进行了上述修改处理后，就可以将修改后的语音文件加载到 MRS 设备上去。这样，当 MRS 设备每次播放语音时，首先会听到四个 DTMF 号码的声音。

[0054] 在应用中需要考虑这样一种情况：如果开始进行语音检测时，该语音已经播放了一部分了，就有可能没有从语音文件的标识码的第一位开始检测，由于语音文件通常是循环播放的，最后检测到的标识码可能会是前一次循环时标识码的后几位加上后一次循环时标识码的前几位，比如一条语音的标识码定义为：103，那么在这里所述的情况下，检测到的标识码可能就变成了：310。这样检测出来的标识码与本条语音实际定义的标识码就不一致了，导致检测出错。

[0055] 为了避免上述错误，标识码应采用基本标识码和边界码结合的方式，基本标识码即为指定位数的数字码，用于标识语音文件，是检索语音文件的必要内容。边界码是某一特定的特殊号码，如“#”，用于标识基本标识码的边界，边界码不是检索语音文件的必要内容。基本标识码中不应该包含边界码。这样在遇到如前所述的循环播放导致的误检时，即可根据边界码的界定来确定正确的语音标识码。基本标识码和边界码具体有如下三种不同的表现方式：1、基本标识码+边界码；2、边界码+基本标识码；3、第一边界码+基本标识码+第二边界码。应用中可根据实际需求选择其中一种表现方式。

[0056] 在图 2 所示的语音识别检测的第一种实施方式的流程中，标识码采用了“基本标识码+边界码”的方案，比如，对“线路正忙，请稍后再拨”这条语音，其标识码可定义为“103#”。

[0057] 由于本发明方案的目的是要实现自动化处理，所以对语音的检测和识别必须由自动化工具来实现。这里选择 RealPhone 工具来实现语音识别，RealPhone 是基于 Intel Dialogic 语音卡的电话仿真控制工具，通过对语音卡进行控制，实现电话机的各种功能，RealPhone 已经具备检测 DTMF 号码的功能，现在需要做的就是要在这个功能的基础上，针对所定义的语音标识码进行特殊处理。

[0058] 首先，定义一个语音识别索引文件，确定语音识别码与语音内容对应关系，即索引文件中存储有语音识别码与语音内容的对应关系。最简单的情况，索引文件的内容可以是下面的格式，001#：一；002#：二；003#：三；004#：四；005#：五；006#：六；007#：七；008#：八；009#：九；069#：年；070#：月；071#：日；101#：您设定的时间已到；103#：线路正忙，请稍后再拨；107#：您的号码有误，请查证后再拨；108#：对不起，您拨的号码不存在，请查证再拨；111#：您申请的新业务已登记，欢迎使用；113#：对不起，话机欠费，暂不能使用；

114# : 呼叫受限, 请勿越权使用 ; 115# : [音乐] ; 119# : 您所拨打的用户正忙, 请稍后再拨 ; 124# : 您所拨打的用户线故障 ; 128# : 刚才一响, 北京时间 等。在实际应用中, 索引文件可根据需要编辑, 按确定的格式增加其它一些信息, 如索引文件中包括语音识别码、语音内容、注释信息这三者的对应关系等。

[0059] 当标识码采用“基本标识码 + 边界码”的方案时, 具体的语音检测识别过程如附图 2 所示。

[0060] 图 2 中, S1) 收取 DTMF 标识号码, 再执行步骤 S2 ;

[0061] S2) 判断收取到的 DTMF 标识号码的位数是否齐全, 如果收取到的标识号码的位数齐全则执行步骤 S3 ; 如果收取到的标识号码的位数不齐全, 则执行步骤 S8 ;

[0062] S3) 对收取到的 DTMF 标识号码进行合法性验证, 判断收取到的 DTMF 标识号码是否合法, 如果 DTMF 标识号码合法, 则执行步骤 S4 ; 如果 DTMF 标识号码不合法, 则执行步骤 S7 ;

[0063] S4) 在索引文件中查找 DTMF 标识号码对应的语音内容, 再执行步骤 S5 ;

[0064] S5) 判断是否找到匹配项, 如果找到匹配项则执行步骤 S6, 如果未找到匹配项则执行步骤 S9 ;

[0065] S6) 输出语音内容, 进行相应处理 ;

[0066] S7) 保留边界码后面的号码, 丢弃前面的号码, 回到步骤 S1 ;

[0067] S8) 判断是否超时, 如果超时则执行步骤 S9 ; 如果不超时, 则回到步骤 S1 ;

[0068] S9) 进行报错处理 ; 再执行步骤 S10 ;

[0069] S10) 进行失败处理。

[0070] 在上述步骤中, 只需收 4 位 DTMF 标识号码, 收全 4 位号码后, 验证其合法性, 合法性验证的内容如最后一位 DTMF 标识号码必须是边界码“#”, 前面三位是数字, 合法性验证通过后, 则应利用收取到的 4 位 DTMF 标识号码立即查找索引文件, 找到并输出对应的语音, 然后根据实际测试需要进行后续处理。例如当收到的 4 位 DTMF 标识号码不是“103#”, 而是“3#10”, 即最后一位不是边界码“#”, 则合法性验证不通过, 这时应该丢弃边界码及以前的号码, 而保留边界码后面的 DTMF 标识号码, 即“10”, 然后继续收号, 直到收到合法的 DTMF 标识号码, 再去查找索引文件。通过前面的处理过程可以看出, 进行语音识别时一般只需要检测出语音前面的标识码即可立即识别出语音内容, 而不必完全等到语音播放完成, 这样在进行大量的测试时, 对提高语音识别效率起到的作用是非常明显的。

[0071] 在图 3 所示的语音识别检测的第二种实施方式的流程中, 标记信息采用了“边界码 + 基本标识码”的方案, 比如, 对“线路正忙, 请稍后再拨”这条语音, 其标识码可定义为“#103”。具体流程如下 :

[0072] S1) 收取边界码 ;

[0073] S2) 判断是否收取到边界码, 如果收到, 执行步骤 S3, 如果没收到则执行步骤 S11 ;

[0074] S3) 接收基本标识码, 执行步骤 S4 ;

[0075] S4) 判断是否收齐三位基本标识码, 如果收齐, 执行步骤 S5 ; 如果没收齐, 执行步骤 S10 ;

[0076] S5) 进行合法性验证, 如果通过验证, 则执行步骤 S6 ; 如果没通过验证, 执行步骤 S9 ;

- [0077] S6) 利用上述收取的三位基本标识码在索引文件中查找, 执行步骤 S7 ;
- [0078] S7) 判断是否找到匹配项, 如果找到, 执行步骤 S8 ; 如果未找到, 执行步骤 S12 ;
- [0079] S8) 根据查找到的匹配项输出语音内容, 并进行相应的后续处理 ;
- [0080] S9) 丢弃收到的边界码和基本标识码, 再执行步骤 S1 ;
- [0081] S10) 判断是否超时, 如果超时执行步骤 S12 ; 如果没超时, 执行步骤 S3 ;
- [0082] S11) 判断是否超时, 如果超时执行步骤 S12 ; 如果没超时, 执行步骤 S1 ;
- [0083] S12) 进行报错处理 ;
- [0084] S13) 进行失败处理。
- [0085] 在第二实施例中, 以检测收到“边界码”作为触发条件, 触发后续的收取“基本标识码”等过程, 语音识别的处理过程更简单。
- [0086] 在图 4 所示的语音识别检测的第三种实施方式的流程中, 标识码采用了“第一边界码 + 基本标识码 + 第二边界码”的方案, 比如, 对“线路正忙, 请稍后再拨”这条语音内容, 其标识码可定义为“#103*”。其中基本标识码不用限定长度, 有利于语音识别处理过程的扩展, 具体流程如下 :
- [0087] S1) 收取并存储任意 DTMF 标识号码 ;
- [0088] S2) 判断是否收取到, 如果有收取到, 则执行步骤 S3, 如果没有收取到, 则执行步骤 S10 ;
- [0089] S3) 判断是否收取到第一边界码, 如果收取到第一边界码, 执行步骤 S4 ; 如果未收取到第一边界码, 执行步骤 S9 ;
- [0090] S4) 判断是否收取到第二边界码, 如果收取到第二边界码, 执行步骤 S5 ; 如果未收取到第一边界码, 执行步骤 S1 ;
- [0091] S5) 对收取到的基本标识码进行合法性验证, 如果验证通过, 则执行步骤 S6 ; 如果验证未通过, 则执行步骤 S9 ;
- [0092] S6) 根据收到的基本标识码在索引文件中查找, 再执行步骤 S7 ;
- [0093] S7) 判断是否找到匹配项, 如果找到匹配项, 执行步骤 S8 ; 如果未找到匹配项, 执行步骤 S11 ;
- [0094] S8) 根据查找到的匹配项输出语音内容, 并进行相应的后续处理 ;
- [0095] S9) 丢弃所有号码, 再执行步骤 S1 ;
- [0096] S10) 判断是否超时, 如果未超时执行步骤 S1, 如果已超时, 执行步骤 S11 ;
- [0097] S11) 进行报错处理, 再执行步骤 S12
- [0098] S12) 进行失败处理。
- [0099] 本发明还有其他的实施方式, 例如语音识别终端不限于使用 RealPhone 工具及语音卡, 还可以为其他的识别工具。本发明不限于采用 DTMF 号码作为标识码进行语音识别处理, 在实际应用中, 如果受到某些条件约束或有其它特殊需求时, 也可以采用其它特定的频率信号作为标识码, 例如可以采用某个频率的单音作为标识码, 还可以采用 MF (Multi-Frequency) 或自定义的频率信号等作为标识码。本发明提供的技术方案不仅能够基于 VoIP 的 NGN 组网中应用, 还可以在其他网络中应用, 而且播放语音文件的网络设备也不限于 MRS。
- [0100] 对于纯自动化测试环境, 在无需考虑人工操作的时候, 也可以用标识码完全替代

语音。这样会带来如下优点：1) 可以提高平均语音识别速度，即对于从中间开始放音的情况，避免了等待语音内容完整播放的情况；2) 可以减少剪辑语音文件的麻烦，双音码号的语音文件甚至可以用程序来自动生成；3) 避免了标识码的误检，如提示音中如果有接近双音频的信号，就会发生号码误检。

[0101] 如果在MRS中加载两套放音方案，即在MRS中加载两套语音文件，一套语音文件中插入了标识码，另一套语音文件中的语音信息替换为标识码，并在MRS中提供切换控制命令，则本发明提供的技术方案可以方便的适用于自动测试及人工测试。

[0102] 下面对本发明提供的对限定范围的语音进行自动识别的系统和设备进行说明。

[0103] 本发明提供的对限定范围的语音进行自动识别的系统包括：语音发送端设备和语音接收端设备，语音发送端设备中设置有语音存储模块和语音发送模块，语音接收端设备中设置有对应关系存储模块、提取模块和检索模块。

[0104] 语音存储模块主要用于存储语音信息，语音存储模块存储的语音信息可以表现为语音文件的形式，即语音存储模块中存储多个语音文件。每个语音文件中插入了对应的标记信息，或者语音文件中的内容替换为标记信息。这个标记信息即上述方法实施例中描述的标识码，语音文件中的标识码可以为DTMF号码、或者某个频率的单音、或者MF号码、或者自定义的频率信号等等。这里的标识码可以为基本标识码和边界码结合的方式，如标识码为基本标识码+边界码的方式；再如标识码为边界码+基本标识码的方式；还有标识码为第一边界码+基本标识码+第二边界码的方式。具体如上述方法中的描述。

[0105] 语音发送模块主要用于根据放音需求发送语音存储模块中存储的相应的语音文件。语音发送模块发送的语音文件中包含有标识码。

[0106] 对应关系存储模块主要用于存储语音内容与标识码的对应关系。

[0107] 提取模块主要用于从接收到的语音文件中提取有效的标记信息，并传输至检索模块。当语音文件中的标识码的表现方式不同时，提取模块从语音文件中提取有效标识码的过程会有所不同。提取模块从语音文件中提取有效标识码的具体过程如上述方法中的描述。提取模块可以仅将基本标识码传输至检索模块，也可以将边界码和基本标识码都传输至检索模块。

[0108] 检索模块主要用于根据接收的标识码如基本标识码从对应关系存储模块存储的对应关系中检索语音内容。检索模块中可以存储索引文件，该索引文件为语音内容与标识码的对应关系，这样，检索模块可以根据其接收到的标识码通过检索索引文件来获得语音内容。具体如上述方法中的描述。

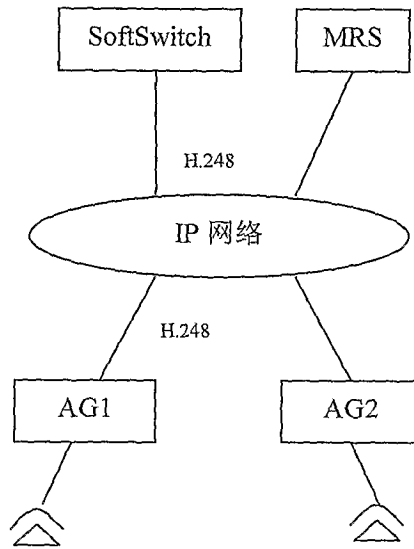


图 1

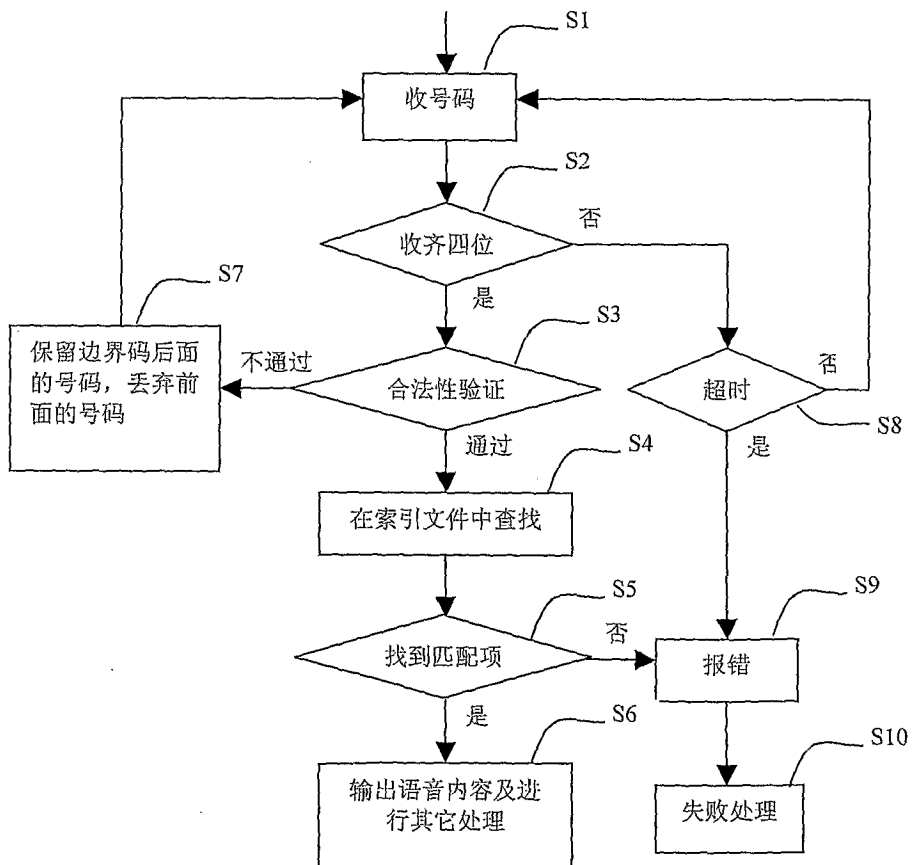


图 2

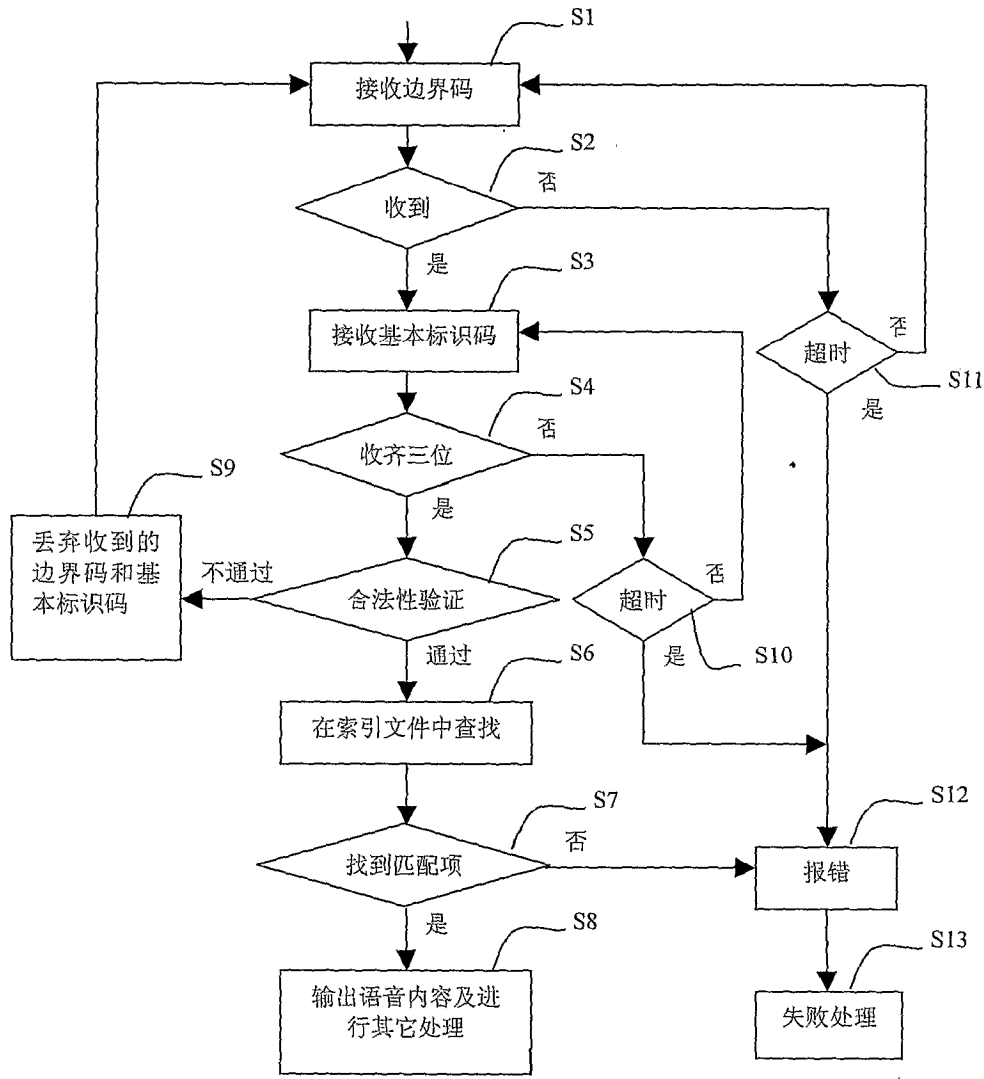


图 3

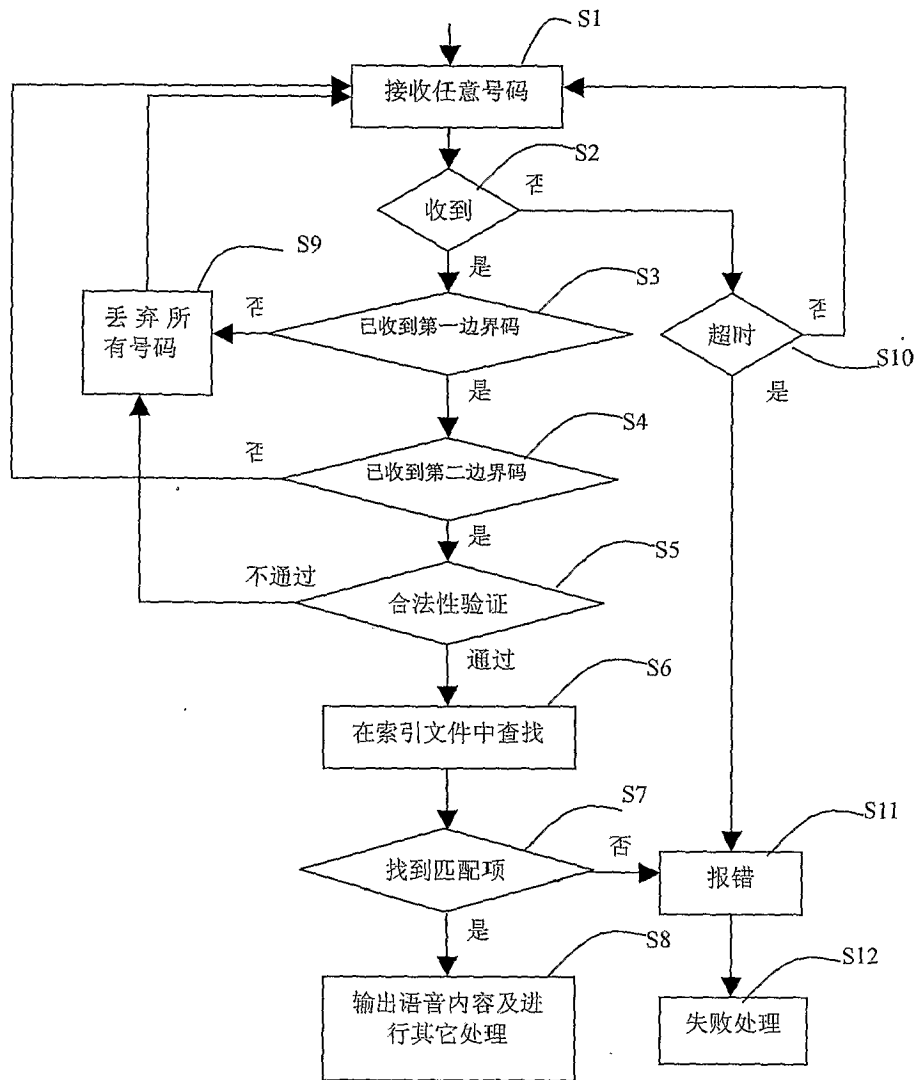


图 4