

[19] 中华人民共和国国家知识产权局

[51] Int. Cl⁷

H04M 1/26

G10L 15/02



[12] 发明专利说明书

[21] ZL 专利号 99101083.3

[45] 授权公告日 2004 年 12 月 29 日

[11] 授权公告号 CN 1182694C

[22] 申请日 1999.1.14 [21] 申请号 99101083.3

[30] 优先权

[32] 1998.1.16 [33] EP [31] 98200122.4

[71] 专利权人 皇家飞利浦电子有限公司

地址 荷兰艾恩德霍芬

[72] 发明人 Y·C·楚 T·H·林

审查员 张宏伟

[74] 专利代理机构 中国专利代理(香港)有限公司

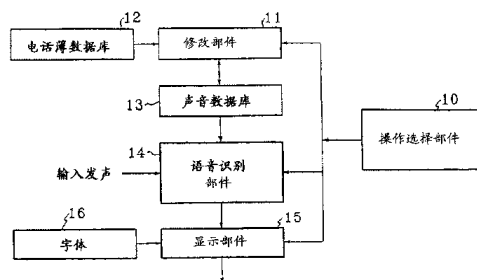
代理人 栾本生 王 岳

权利要求书 2 页 说明书 12 页 附图 9 页

[54] 发明名称 用于电话机的自动拨号的话音命令系统

[57] 摘要

本发明涉及用于自动拨号的话音命令系统。该构思是为在电话簿项中利用音标而不用字符串。能遵照表 1 中所定义的专门设计的键入序列用键盘输入音标。作为替代，能通过以隔离方式讲出音节到电话送话器中而用语音识别作出该项，由于普通话中每一字符只有一个音节。音标是自动转换成话音命令模型的，它允许用户用口头命令拨号码。



1. 一种用于电话机的自动拨号的话音命令系统，包括：
在命令模式、编辑模式或普通模式中确定所述系统的操作模式的操作选择部件；
- 5 提供编辑功能及存储项目到电话簿数据库中的修改部件，所述项目包括姓名和号码，所述姓名包括音标；
存储基本声音模型和从基本声音模型中生成的话音命令模型的语音数据库，所述话音命令模型对应于项目的姓名；
能够使用话音命令模型来解码用户的输入发声并且提供与所述输入发声相似的N个最佳解码结果的语音识别部件；
- 10 显示拨打的项目的姓名和作为音标修改的项目的姓名中的至少一个的显示部件；以及
拨打对应于与所述输入发声相似的最佳解码结果的号码的拨号部件。
- 15 2. 根据权利要求1的系统，其中所述修改部件具有输入功能，利用音调与音标的映射方案通过按所述音调与音标的对应电话按键输入所述数据。
3. 根据权利要求1的系统，其中所述修改部件具有输入功能，利用音调与音标的映射方案用话音命令输入所述数据。
- 20 4. 根据权利要求1的系统，其中所述修改部件具有输入功能，利用所述显示部件显示的音调与音标映射方案通过按所述音调与音标的对应电话按键输入所述数据。
5. 根据权利要求1的系统，其中所述修改部件具有输入功能，利用所述显示部件显示的音调与音标的映射方案用话音命令输入所述数据。
- 25 6. 根据权利要求1的系统，其中所述修改部件具有输入功能，利用给定的字母表的映射方案通过按所述字母表的对应电话按键输入所述数据。
7. 根据权利要求1的系统，其中所述修改部件具有输入功能，利用给定字母表的映射方案用话音命令输入所述数据。
- 30 8. 根据权利要求1的系统，其中所述修改部件具有输入功能，利用所述显示部件显示的给定字母表的映射方案通过按所述字母表的对

应电话按键输入所述数据。

9. 根据权利要求 1 的系统，其中所述修改部件具有输入功能，利用所述显示部件显示的给定字母表的映射方案用语音命令输入所述数据。

- 5 10. 根据权利要求 1 的语音命令系统，还包括一个用于选择来拨打 N 个最佳解码结果之一的选择部件。

用于电话机的自动拨号的话音命令系统

技术领域

- 5 本发明涉及用于自动拨号的话音命令系统，其特征为通过话音命令的自动拨号，以便实现在电话机上。

背景技术

- 除了顺序地按键盘来输入要求的电话号码，可以购得的电话机还提供其它特征，诸如利用热键来存储电话号码供以后拨号。这种方法通常要求用户在输入电话号码之前先按“存储”键，然后通过按对应的热键来存储。在初始设定之后，用户可通过按其对应热键而重叫项目以便拨想要的电话号码。虽然这是拨号的简单方法，但它具有以下缺点。（1）难于记住热键与电话簿项目之间的对应性。（2）其它人不知道用户设定的电话簿项目。以及（3）每当修改电话簿项目时必须更新电话簿项目的书面对照表。

发明内容

本发明的一个目的为提供用于自动拨号的话音命令系统，其特征为通过话音命令自动拨号。

- 本发明的另一目的为提供如表 1 中所示的用于以键盘输入中文项目的音调与音节的映射方案。利用表 1 中列出的编码方案及本发明中描述的方法，人们能通过电话键盘容易地输入中文项目。所有中文项目都将用它们的音标序列输入与显示。此外，将姓名项目自动转换成话音命令模型，允许用户在编辑各项之后无须额外训练便能用口头命令拨号。

- 按照本发明的用于自动拨号的话音命令系统包括若干基本部件。操作选择部件用于从三种模式：命令、编辑或普通模式，中确定电话的操作模式。修改部件的特征为用于在包含以它们的对应音调与音标的姓名项以及电话号码项的电话簿数据库中增加电话号码项的功能。声音数据库包含基本声音模型及通过串联连接基本声音模型构成的话音命令模型。话音识别部件利用包含在声音数据库中的声音模型与话音命令模型来解码用户的输入发音及提供 N 个最佳结果。显示部件显示修改与语音识别部件的结果，在核对以后将修改的结果发送给修改部件，及将解码结果发送给拨号部件以拨所要求的电话号码。

根据本发明的优选实施例，提供了一种用于自动拨号的话音命令系统，包括：在命令模式、编辑模式或普通模式中确定所述系统的操作模式的操作选择部件；提供编辑功能及存储项目到电话簿数据库中的修改部件，所述项目包括姓名和号码，所述姓名包括音标；存储基本声音模型和从基本声音模型中生成的话音命令模型的声音数据库，所述话音命令模型对应于项目的姓名；能够使用话音命令模型来解码用户的输入发声并且提供与所述输入发声相似的 N 个最佳解码结果的语音识别部件；显示拨打的项目的姓名和作为音标修改的项目的姓名中的至少一个的显示部件；以及拨打对应于与所述输入发生相似的最佳解码结果的号码的拨号部件。

附图说明

下面将参照附图详细说明本发明；其中

表 1 为按顺序将中文语音符号映射到它们的键上，例如，以符号“勺”按“#51”到键上；

图 1 为展示按照本发明的实施例的自动拨号的话音命令系统的方框图；
图 2 为展示上述实施例的操作选择部件的流程图；
图 3A 为上述实施例的编辑模式的流程图；
图 3B 与 3C 为展示在上述实施例的编辑模式下的输入功能的流程图；
图 3D 为展示在上述实施例的编辑模式下的语音输入功能的流程图；
图 3E 为展示在上述实施例的编辑模式下的查询功能的流程图；
图 3F 与 3G 为展示在上述实施例的编辑模式下的修改功能的流程图；以及
图 4 为展示上述实施例的命令模式的流程图。

具体实施方式

图 1 为展示按照本发明的实施例的自动拨号的话音命令系统的方框图。如图中所示，用框 10 表示的操作选择部件用于在命令模式 27、编辑模式 28 或普通模式 26 中确定当前的操作模式。命令模式 27 接受输入发音，找到对应的电话号码及为用户拨它。编辑模式 28 提供编辑电话簿数据库 12 的功能。普通模式 26 类似于典型的电话功能。修改部件 11 有三种编辑功能特征：输入、查询与修改。输入功能允许用户增加新的姓名与号码项到电话簿数据库 12 中及串行地连接表示姓名项的基本声音模型以构成话音命令模型。查询功能协助用户定位特定的

数据项及允许用户直接拨它。修改功能的特征为以下功用：增加/删除电话号码项及删除数据记录。电话簿数据库 12 包含以它们的对应音标表示的姓名项以及电话号码项。声音数据库采用隐藏的 Markov 模型并包含 (a) 408 个音节的聲音模型及 (b) 由姓名项的对应声音模型构成的话音命令模型。语音识别部件 14 将输入发音解码成类似于声音模型的特征矢量并利用 Viterbi 解码过程来找出 N 个最佳声音模型或 N 个最佳话音命令模型。显示部件 15 显示修改与语音识别部件的结果。字体 16 包含显示部件要使用的阿拉伯数字、音标及特殊符号的显示字体。

图 2 至 4 为展示按照本发明的实施例的自动拨号的话音命令系统的进程的流程图。参照图 2 至 4 详细说明操作过程。

图 2 为展示图 1 中所示的操作选择部件 10 的进程的流程图。操作选择部件 10 从命令模式 27、编辑模式 28 或普通模式 26 中确定当前的操作模式。如图中所示，一旦将手机 21 从其托架上举起，操作选择部件 10 首先检测拨号音，如用框 22 所示。在没有拨号音的情况中，将电话设定为普通模式 26。否则，系统进入供拨号的在线状态。然后，操作选择部件 10 检测用户发音，如用框 23 所示。任何输入发音的出现指示要用话音命令来自动拨号，而系统进入命令模式 27，其中图 4 详细示出命令模式的进程。在没有用户发音的情况中，则操作选择部件 10 检测是否按下了“#”键，如用框 25 所示。如果按下了“#”键，系统进入编辑模式 28，其中图 3A 详细示出编辑模式的进程。用户按下“#”键以外的任何键表示用户的拨号动作而系统进入普通模式 26。当本发明的自动拨号的话音命令系统进入普通模式时，电话机便与传统电话相同的方式操作。

图 3A 至 3G 为展示图 1 中所示的修改部件 11 的进程的流程图。修改部件的特征为允许用户在电话簿数据库 12 中增加姓名及电话号码项的编辑功能。修改部件还按照发音表组合基本声音模型来构成话音命令模型。图 3A、3B、3C、3D、3E、3F 与 3G 为分别展示编辑模式及其输入、语音输入、查询、及修改功能的进程的流程图。

图 3A 为编辑模式 28 的进程的流程图，该模式的特征为四种功能：语音输入 37、输入 38、查询 39 及修改 40。一旦在编辑模式中，便提示“1. 语音，2. 输入，3. 查询，4. 修改”供用户选择。如果按下键“1”系统便进入“语音输入”功能 37，如果按下键“2”进入“输入”功能

38, 如果按下键“3”则为“查询”功能 39, 及如果按下键“4”则为“修改”功能 40。图 3B 为展示在编辑模式 28 下的输入功能 38 的进程的流程图。输入功能 38 的特征为诸如编辑姓名项, 编辑电话号码项, 按照发音表组合基本声音模型 357 以构成话音命令模型 408, 编辑姓名与电话号码项到电话簿数据库 12 中及增加话音命令模型到声音数据库 13 中, 等功用。在图 3B 中所示的输入功能的流程图中, 首先设定变量 Nflag 为 1, 如用框 310 所示。Nflag=1 表示系统进入姓名项功用。然后系统设定变量 Tcnt、cnt1、cnt2 为 0。变量 Tcnt 表示电话号码项的长度(最大 15 位), 这时 Tcnt=0 表示无位项。一旦按下一个键, 便将键入项存储在变量 key 中, 如框 311 所示。然后系统检测变量 Nflag 是否为 1, 如框 312 所示。如果变量 Nflag=1, 系统进入姓名项循环。然后系统检测变量 key 是否等于“*”, 如框 330 所示。如果变量 key = “*”, 显示器上的光标移动到前一项, 如框 343 所示, 系统则返回到框 311。否则, 将键入项放在姓名缓冲器中, 如框 331 所示。然后系统检测变量 key 是否等于“#”, 如框 332 所示。如果变量 key≠“#”, 便将变量 cnt1 复位到 0, 如框 344 所示, 其中变量 cnt1 计数“#”键的接连接压, 而系统返回到框 311。否则在变量 cnt1 上加上值 1, 如框 333 所示。然后系统检测是否变量 cnt1=2, 如框 334 所示。如果变量 cnt1=2, 系统返回到框 311。否则用户按了“#”键两次, 这表示姓名项对话结束, 并将变量 cnt1 复位到 0, 如框 335 所示。然后系统确认键入的对应于姓名项的语音记录是否满足表 1 中所示的音标的编辑定义, 如框 336 所示。如果姓名项不满足编码定义便提示出错报文 337 及 346 而系统返回到框 311。否则, 按照表 1 的映射将姓名项转换成音标, 如框 345 所示。然后系统确认姓名项是否满足表 1 中所示的 408 个音节的定义, 如框 338 所示。如果姓名项不满足映射定义则提示出错报文 339 及 348 而系统返回到框 311。否则, 系统确认姓名项是否重复, 如框 340 所示。如果姓名项重复便提示重复报文 349 而系统返回到框 311。否则, 系统提示报文询问用户来核对键入的姓名项, 如框 341 所示。如果不能得到肯定的确认, 系统返回到框 311。否则, 将变量 Nflag 复位到 0, 如框 342 所示, 其中变量 Nflag=1 表示系统进入电话号码项功用, 而系统返回到框 311。下面给出姓名项功用的示例。

如果用户想要输入中文名字“李小华”，第一步为将字符串转换成其拼音符号“ㄉ-ㄩㄊ-ㄝㄩㄉㄨㄚㄩˇ”。接着，用户需要在表1中找到拼音符号的对应键入序列。注意表1中的号码必须冠以“#”键。

“ㄉ-ㄩㄊ-ㄝㄩㄉㄨㄚㄩˇ”的得出的键入序列为“#58#72#3”#64
5 #72#81#3#61#73#75#2##”，其中的双“##”键表示姓名项对话的结束。然后将字符串放入姓名缓冲器中并用显示部件显示为“ㄉ-ㄩㄊ-ㄝㄩㄉㄨㄚㄩˇ”。

在替代实施例中，系统可利用显示部件15来显示对应于单个音标的数字，通过它们用户可通过按其对应的键选择想要的拼音符号。例如，显示部件15的页1可首先显示音标“ㄉㄩㄊㄝㄩㄉㄨㄚㄩˇ...”及它们的对应数字“01234567...”。为了输入中文串“李小华”，首先选择表示音标“ㄉ”的数字“7”。如果想要的音标不能显示在页1上，可提示页2以便以类似方式完全输入姓名项。

图3B中，变量Nflag为0表示姓名项对话的结束，如框342所示，
15 而系统已准备好电话号码输入。图3B与图3C示出号码输入功用的流程图。一旦按下了一个键，便将键项存储在变量key中，如框311所示。然后系统检测是否变量Nflag=1，如框312所示。如果变量Nflag≠1，系统进入号码项循环。然后系统检测是否变量“key=*”，如框313所示。如果变量key=*，光标移动到前一项，如框319所示，
20 而系统返回到框311。否则，系统检测是否变量key=#，如框314所示。如果变量key=#，便将变量cnt1复位到0，如框315所示，其中变量cnt1计数“#”键的接连接压。系统还检测变量key是否在0至9的范围内，及变量Tcnt是否小于15，如框316所示。否定的结果表示变量key是无效项而系统返回到框311不采取任何行动。
25 如果变量key为有效项，将键入项临时放置在号码缓冲器中，如框317所示。然后在变量Tcnt上加上值1，如框318所示（在电话号码上加上一位），而系统返回到框311。如果变量key=#，在变量cnt1上加上值1，如框320所示。然后系统检测是否变量cnt1=2，如框321所示。如果变量cnt1≠2，则可能已完全输入了一个电话号码项。然后
30 系统检测是否变量Tcnt>0，如框326所示。否定的结果表示无电话号码项而系统返回到框311不采取任何行动。否则，号码项有效而在变量cnt2上加上值1，如框327所示，其中变量cnt2表示存储的电话号

码项的数目（最大为 4）。然后系统将变量 Tcnt 复位到 0。然后系统检测是否变量 cnt2 = 4，如框 328 所示。否定的结果返回系统到框 311。如果 cnt2=4，用户完全设定了全部四个电话号码项。将临时放置在姓名与号码缓冲器中的姓名与电话号码项移到电话簿数据库 12 中，如框 324 所示，而系统返回到图 3A 中的编辑模式的入口点。当 cnt1 = 2 时，号码项对话结束。然后系统将变量 cnt1 复位到 0，如框 322 所示，并检测是否变量 cnt2 > 0，如框 323 所示。否定的结果表示号码缓冲器中无电话号码项，而系统返回到框 311 不采取任何行动。否则，一个电话号码项当前放在号码缓冲器中而系统返回到框 324，将声音模型转换成话音命令模型并将姓名与电话号码项移到电话簿数据库中，如上所述。下面给出号码项操作的示例。

假定“李小华”有两个电话号码：212 - 7827657 及 7556133。它们能用序列“2127827657 # 7556133 # #”输入，其中单个“#”键用来分隔不同的电话号码而双“##”键表示号码项的结束。

将键入的姓名与电话号码项存储在具有表 1 中所示的数据结构的电话簿数据库 12 中。

声音数据库 13 包含基本声音模型及通过连接基本声音模型构成的话音命令模型 408。基本声音模型是不包含音调的音节模型。各音节模型为右上下文相关 (RCD) 的首字母 (声母) 与终字母 (韵母) 的组合。总共采用了 109 个 RCD 首字母及 38 个终字母。这些 RCD 首字母与终字母是由采用多个发音人语音数据库的分段 k 装置 (segmental k-means) 方法训练的隐藏 Markov 模型。这些隐藏 Markov 模型的特征矢量包含 12 个 Mel 频率倒频谱系数及 12 个 Delta Mel 频率倒频谱系数。用“李小华”作为说明如何用 RCD 首字母与终字母组成声音模型的示例。“李 (ㄌ - ㄨ)”由 RCD 首字母“l-i”及终字母“i”构成。从而用符号“l-i⊕i”来表示“李 (ㄌ - ㄨ)”的音节模型。类似地，分别用符号“shi - i⊕iau”与“h-u⊕ua”来表示“小 (ㄒ - ㄨ)”与“华 (ㄏ - ㄨ)”的音节模型。话音命令模型是以类似的方式连接音节模型构成的。例如，姓名“李小华”的话音命令模型是通过连接音节模型“李 (l-i⊕i)”、“小 (shi-i⊕iau)”与“华 (h-u⊕ua)”构成的。从而用符号“(l-i⊕i⊕shi-i⊕iau⊕h-u⊕ua)”来表示“李小华”的话音模型。

图 3D 为展示在编辑模式下的语音输入功能的流程图。如上所述，图 3B 中所示的输入功能提供用表 1 中所示的音标的编辑方案编辑姓名

项的方法来编辑姓名项。编辑模式下的语音输入功能提供允许用户用口头输入编辑姓名项的替代方法。在语音输入进程中，首先将变量 Scnt 设置成 0，如框 351 所示，其中变量 Scnt 为 0 表示无输入发声。然后将一个音节发声到手机中，如框 352 所示。然后将变量 kcnc 设置成 0，如框 353 所示，其中 kcnc 为 0 表示该音节需要进一步确认。然后将输入音节移至语音识别部件 14 供进一步处理。语音识别部件 14 执行三个步骤：终点检测，特征抽取及模式匹配，终点检测模块 354 将语音信号与背景噪声分离。特征抽取模块 355 执行输入语音的频谱分析来抽取由 12 个 Mel 频率倒频谱系数及 12 个 Delta Mel 频率倒频谱系数组成的特征矢量。模式匹配模块 355 利用 Viterbi 解码过程找出最相似输入语音的 N 个最佳音节模型 357。按照表示输入语音与 408 个音节模型 357 之间的相似性的匹配分数排序这些结果，然后按照它们的匹配分数的对应音标顺序地显示这些结果，如框 358 所示，其中首先显示最高匹配分数的结果。然后系统请求键入值，如框 360 所示，并检测表示键入值的变量 key 是否在 0 至 4 的范围内。肯定的结果表示当前正在显示的音标对应于输入音节且变量 key 的值表示输入音节的音调。然后系统检测是否变量 kcnc = 0，如框 369 所示。肯定的结果表示该音节需要进一步确认；从而将该音节及其音调临时放置在姓名缓冲器中，如框 370 所示。然后在变量 Scnt 上加上值 1，表示已将一个字符加到姓名项中。将变量 kcnc 设置成 1，表示选择了该输入音节，如框 367 所示。

否则，变量 kcnc 不为 0 表示已选择了该输入音节。用变量 key 的值来取代输入音节的音调，如框 368 所示。然后系统返回到框 360 用于下一个键入值。

如果变量 key 不在 0 至 4 的范围内，系统便检测是否变量 key=8，如框 362 所示。肯定的结果表示匹配结果的下一个候选者需要考察而系统返回到框 358。否则，系统检测是否变量 key=7，如框 363 所示。肯定的结果表示已准备好将下一个音节发声到手机中，而系统返回到框 352。否则，系统检测是否变量 key=0，如框 364 所示。否定的结果将系统返回到框 360 并要求键入值。变量 key 为 0 表示姓名项对话的结束。然后系统检测是否变量 Scnt > 0，如框 365 所示。否定的结果表示没有姓名项放置在姓名缓冲器中；从而系统返回到框 352 请求用户用语音输入姓名项。否则，变量 Scnt 大于 0 表示姓名项对话的结束。

然后系统将变量 Nflag、Tcnt、cnt1 及 cnt2 复位到 0，如框 366 所示。在为以后的号码输入功能转移到图 3B 中的点@350 之前必须设定变量。

作为输入姓名“李小华”的一个示例，首先说出音节“李(ㄌ-ㄨ)”到手机中。假如语音识别部件 14 找出的 N 个最佳声音模型为“1: ㄌ-ㄌ, 2: ㄋ-, 3: ㄌ-, 4: -...”。

5 显示部件 15 将首先显示音标“ㄌ-ㄌ”，这不是表示字“李”的音标；从而按下键“8”。然后显示部件显示音标“ㄋ-”，这也不是表示字“李”的音标，再一次按下键“8”。然后显示部件显示音标“ㄌ-”，这对应于字“李”的音标，从而按下键“3”来表示其单调。如果需要修改，可在下一次语音

10 输入之前输入值 1 至 4 之间的键。例如，键“4”将“ㄌ-ㄨ”改变成“ㄌ-ㄨ”。然后通过按键“7”为下一次输入“小(ㄒ-ㄨㄛ)”以及随后的语音输入“华(ㄏㄨㄚˊ)”确认姓名项。然后用户按键“0”来表示姓名项对话的结束。现在姓名缓冲器中包含表示“李小华”的

15 字符串“#58#72#3#64#72#81#3#61#73#75#2##”。然后系统转移到修改部件的电话号码输入功能 38。

图 3E 为展示编辑模块 28 的查询功能 39 的流程图。查询功能 39 提供显示前面 379 或后面 380 数据项的功能并允许用户按当前正在显示的数据项，如框 378 所示。一旦用户进入查询功能 39，系统首先检测电话簿数据库 12 中是否存在任何数据项，如框 371 所示。否定的结果

20 断开查询功能并返回系统到图 3A 中编辑模式 28 的起始点。否则，系统显示最近用过的数据项，如框 372 所示。要求输入一个键，并将键入值存储在变量 key 中，如框 373 所示。然后系统检测是否变量 key=“*”，如框 374 所示。肯定的结果显示前一数据项 379 及返回系统到框 373。否则，系统检测是否变量 key=“#”，如框 375 所示。肯定的

25 结果显示下一数据项 380 并返回系统到框 373。否则，系统检测是否变量 key=0，如框 376 所示。肯定的结果表示查询对话的结束而将系统返回到图 3A 中修改部件 28 的入口点。否则系统检测变量 key 是否对应于当前正在显示的任何电话号码项，如框 377 所示。肯定的结果令系统按选择的电话号码并返回系统到普通模块 26。否则，系统忽略键

30 入值并返回到 373 用于下一键入值。

图 3F 及 3G 为展示编辑模式 28 下的修改功能 40 的流程图。修改功能的特征为以下功用：增/删号码项及删除数据库中的记录。一旦用

户进入修改功能 40，系统首先检测电话簿数据库 12 中是否存在任何数据项，如框 381 所示。否定的结果断开修改功能并返回系统到图 3A 中编辑模式 28 的入口点。否则，系统显示最近用过的数据项，如框 382 所示，并将变量 Aflag、Dflag 及 Tcnt 设置成 0，其中变量 Aflag=1 表示选择增加功用，变量 Dflag=1 表示选择删除功用，而变量 Tcnt 表示电话号码项的长度（最大 15 位）。用户能通过按键输入指令，而将键入值存储在变量 key 中，如框 384 所示。然后系统检测是否变量 key="*"，如框 385 所示。肯定的结果表示选择了删除功用，将变量 Dflag 设置成 1，并将变量 Aflag 设置成 0，如框 390 所示。然后系统返回到框 384。否则，系统检测是否变量 key="#”，如框 386 所示。肯定的结果表示选择了增加功用而将变量 Dflag 设置成 0，如框 398 所示。然后系统检测是否变量 Tcnt>0，如框 399 所示。肯定的结果表示增加功用的结束并将临时放置在号码缓冲器中的号码项移到电话簿数据库 12 中，如框 402 所示。然后在变量 Telcnt 上加上值 1，Telcnt 表示以前存储的电话号码项的总数而加 1 表示已输入了一个附加的号码项（最大 4）。然后将变量 Aflag 复位到 0 以结束增加功用并将变量 Tcnt 复位到 0，如框 402 所示。然后系统返回到框 384。否则，变量 Tcnt 为 0 表示不存在号码项。一旦进入增加功用，系统首先检测输入的电话号码数是否超过 4，如框 400 所示。换言之，系统检测是否变量 Telcnt = 4。肯定的结果表示不可能增加号码项而系统忽略“#”键及返回到框 384。否则，可以增加附加的电话号码项而将变量 Aflag 设置成 1，如框 401 所示，以进入增加功用。然后系统返回到框 384。如果变量 key 既非“*”又非“#”。系统便检测是否变量 Dflag=1，如框 387 所示。肯定的结果表示删除功用的入口点。一旦进入删除功用，系统首先检测变量 key 是否对应于显示中的任何电话号码项，如框 391 所示。如果肯定，系统便从电话簿数据库中删除对应的号码项，如框 397 所示。否则，系统检测是否变量 key=5，如框 392 所示。肯定的结果表示需要删除包含姓名项、电话号码项及表示该姓名项的话音命令模型 408 的数据库记录，如框 392 所示。然后系统提示一则报文来询问是否应删除整个记录。否定的回答返回系统到框 384 而不采取任何行动。否则，从电话簿数据库 12 中删除整个数据记录及从声音数据库 13 中删除对应于该姓名项的话音命令模型 408。此后，系统检测电话簿数据库 12

中是否存在任何其它项，如框 395 所示。肯定的结果允许系统显示下一数据项，如框 396 所示并将系统返回到用于下一修改命令的修改功能 40 的入口点。否则，系统返回到用于下一编辑命令的图 3A 中的编辑模式 28 的入口点。在删除功用中，如果变量 key 并不对应于显示中的任何电话号码项且不等于 5，系统返回到框 388。如果变量 key 既非“*”又非“#”，且变量 Dflag=1，系统便检测是否变量 Aflag=1。肯定的结果表示选择了增加功用而系统进入增加功用。然后系统变量 key 是否在 0 至 9 的范围内且变量 Tcnt<15（电话号码项的最大长度），如框 404 所示。肯定的结果允许将变量 key 临时放置在号码缓冲器 405 中并在变量 Tcnt 上加上值 1，如框 406 所示，表示电话号码的附加位。然后系统返回到框 384。否则，变量 key 不在 0 至 9 的范围内或变量 Tcnt 不小于 15 表示键入值是无效的或号码项超过了 15 位，从而系统返回到框 384 而不采取任何行动。如果变量 key 既非“*”又非“#”且变量 Dflag 及 Aflag≠1，系统检测是否变量 key=0，如框 398 所示。如果变量 key≠0，系统忽略键入值并返回到框 384。否则，变量 key=0 表示修改功能 40 的结束而将系统返回到用于下一编辑命令的图 3A 中的编辑模式 28 的入口点。下面说明修改功能的示例。

假定用户利用查询功能选择了存储在电话簿数据库 12 中的第三个数据记录。这一记录具有“林效法”的 4 组电话号码项。显示部件 15 将显示数据项如下。

为—4 / T—么、C Y v

☎1: 3823883 ☎2: 7759398 ☎3: 7221234 ☎4: 8556694

假定需要将第四电话号码项从 8556694 改成 5862699，第一步为通过按键“*4”来删除第四个号码，其中“*”表示删除功用而“4”表示第四号码项。接着，用序列“#5862699#0”输入新的号码项，其中第一个“#”起动号码项而第二个“#”结束号码项。序列中最后的“0”表示修改对话的结束。完成了上述进程之后，系统显示修改后的数据记录如下：

为—4 / T—么、C Y v

☎1: 3823883 ☎2: 7759398 ☎3: 7221234 ☎4: 5862699

图 4 为展示图 1 中的语音识别部件 14 的命令模式 27 的流程图。命令模式 27 允许用户用语音命令拨存储的电话号码项。一旦进入命令

模式 27, 用户首先将想要的人的姓名讲入手机中, 如框 407 所示, 假定已在框 408 将该人的姓名转换成语音命令模型并已利用上述编辑模块 28 将其号码存储在电话簿数据库 12 中。然后将语音段发送到语音识别部件 14 供处理。语音识别部件 14 由三个步骤组成: 终点检测, 5 特征抽取及模式匹配。终点检测模块 354 分离语音段与背景噪声。特征抽取模块 355 执行语音段的频谱分析以抽取作为语音命令模型 13 的特征矢量, 该模型包含 12 个 Mel 频率倒频谱系数及 12 个 Delta Mel 频率倒频谱系数。模式匹配模块利用 Viterbi 解码过程找出与输入语音段最相似的 N 个最佳语音命令模型 408。按照表示输入语音段与语音命令模型 408 之间的相似性的匹配分数来排序这些结果。然后按排序的序列以它们对应的音标显示这些结果, 如框 358 所示, 其中首先显示具有最高匹配分数的结果。然后系统请求键入值, 并将其存储在变量 key 中, 如框 409 所示。然后系统检测是否变量 key = 8, 如框 410 所示。肯定的结果表示当前显示的数据项不正确而系统返回到框 358 15 用于下一个可能的数据项。变量 key 不是 8 表示或者显示的数据项是正确的或者用户要再一次讲入手机中。从而系统检测是否变量 key = 7, 如框 411 所示。肯定的结果将系统返回到命令模式 27 的入口点并期待口声输入。否则, 系统为 7 指示显示的数据项是正确的, 而系统检测变量 Key 是否对应于显示的任何电话号码, 如框 412 所示。肯定的结果 20 允许系统直接拨选择的电话号码, 如框 413 所示, 并返回系统到普通模式 26。否则, 系统返回到框 409 用于下一个键入值。下面给出如何利用语音命令拨一个存储的电话号码的示例。

假定用户想与“李小华”说话, 用户首先拿起手机及讲“李小华”到手机中。如果语音识别部件 14 识别出“王大明”为具有最高匹配分数的候选人而“李小华”为第二最佳匹配, 显示部件 15 首先显示如下:

×尤 ∠ 勿丫 ∠ □一△ ∠

☎1: 3821234 ☎2: 7759876 ☎3: 059252177

由于“王大明”不是想要的项, 用户按键“8”来显示下一个候选人, 显示部件 15 显示如下:

30 力一 ∠ 丁一么 ∠ 厂×丫 ∠

☎1: 7556133 ☎2: 2127827657

由于“李小华”是所要的项, 而其第一个电话号码项 7556133 为

要求的电话号码，用户按键“1”而系统发送解码的结果给拨号部件来拨选择的号码项。然后系统进入普通模式 26。因为“李小华”只包含两个号码项，按键“3”或“4”将不导致行动。

5 如上所述，本发明提出了用于自动拨号的话音命令系统，其特征为通过话音命令自动拨号。此外，本发明中公开的映射方案与编辑方法可用来以它们对应的音标表示与显示中文姓名。此外，姓名项自动转换成话音命令模型允许用户在编辑各项之后无须额外训练便能拨号。

10 注意本发明不受上述实施例的限制，并且只要修改是在本发明的范围内作出的，本发明可在实践中修改。例如，只要图 3D 中的话音项功能采用首字母（ㄅ、ㄆ、ㄇ...）及终字母（ㄩ、一、ㄨ、...）来替代音节（ㄅㄩ、ㄆㄩ、...）作为话音输入，便能用 109 个 RCD 首字母与 39 个终字母来替代声音数据库 13 中所包含的声音模型。此外，特征矢量可包含其它特征，诸如除了 12 个 Mel 频率倒频谱系数及 12 个 Delta Mel 频率倒频谱系数之外的能量分量或 LPC 分量。此外，本发
15 明的应用不限于中文语言。通过加入特定语言的特征对声音数据库 13、字体及修改部件 11 进行适当的修改，便能以任何语言实现本发明。本发明的范围列出在所附权利要求书中。

一	ㄊ	ㄋ	ㄌ		ㄅ	ㄆ	ㄇ	ㄏ	ㄉ
1	2	3	4		51	52	53	54	55
去	ㄓ	ㄊ	ㄍ	ㄎ	ㄐ	ㄑ	ㄒ	ㄓ	ㄔ
56	57	58	59	60	61	62	63	64	65
ㄕ	ㄖ	ㄗ	ㄘ	ㄙ	ㄚ				
66	67	68	69	70	71				
一	ㄨ	ㄩ	ㄚ	ㄛ	ㄜ	ㄝ	ㄞ	ㄟ	ㄠ
72	73	74	75	76	77	78	79	80	81
又	ㄨ	ㄩ	ㄚ	ㄛ	ㄜ	ㄝ			
82	83	84	85	86	87				

表 1

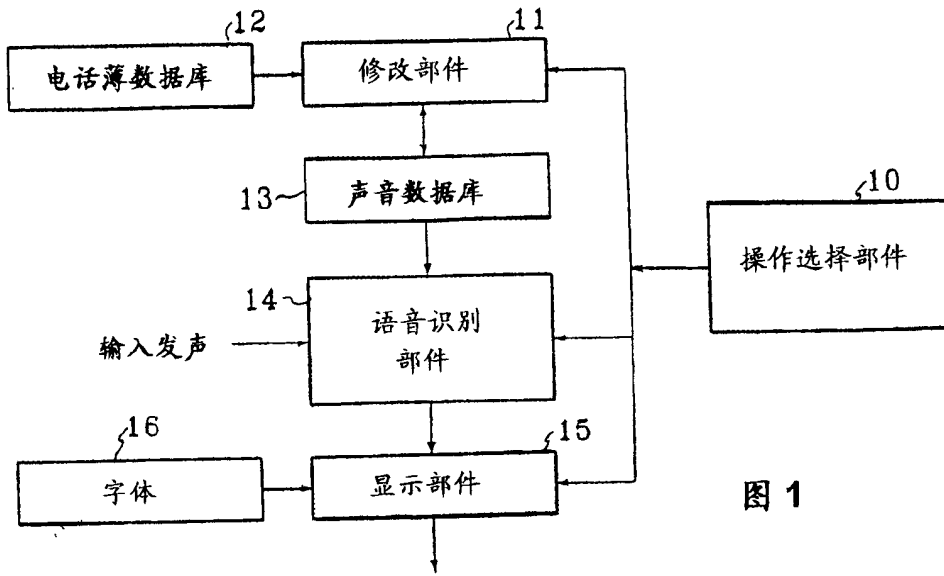


图 1

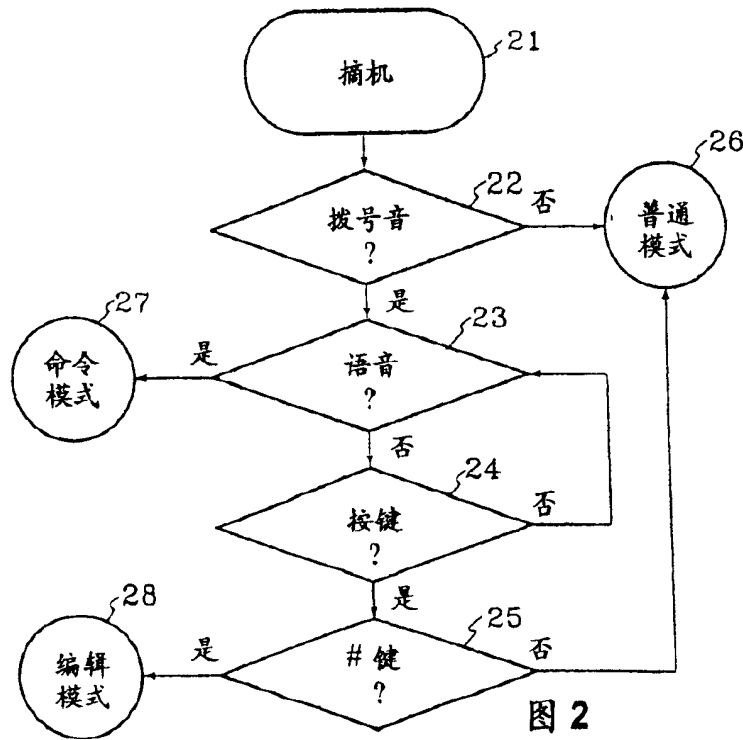


图 2

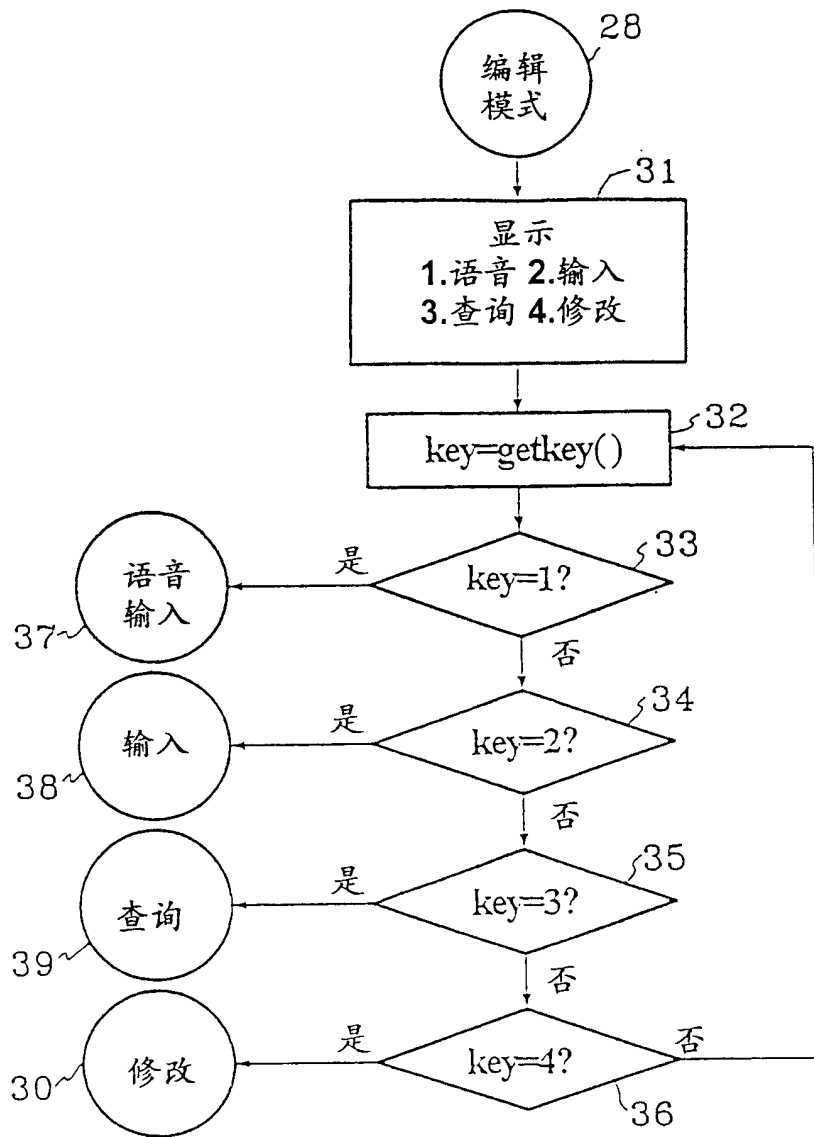


图 3A

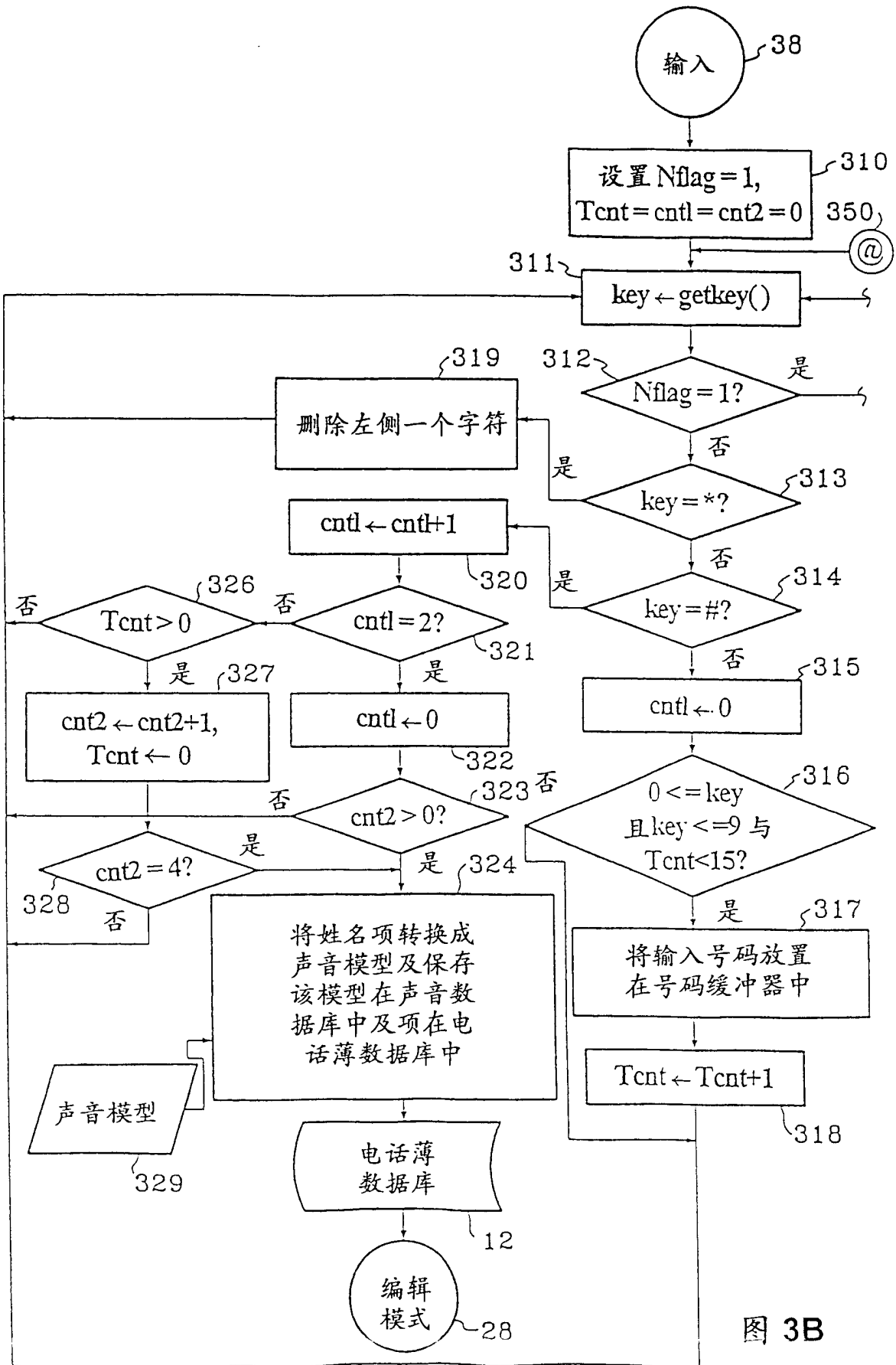


图 3B

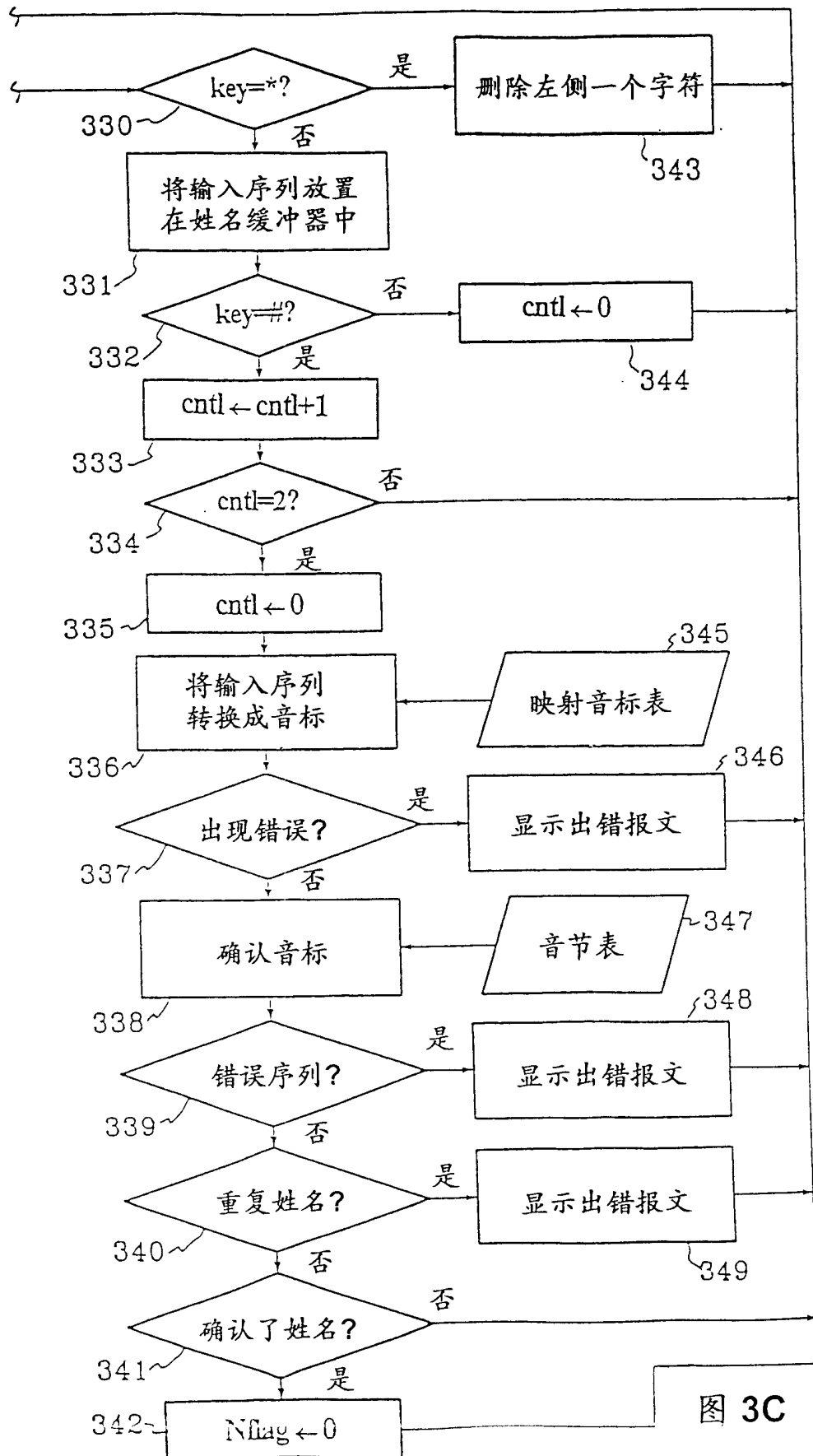


图 3C

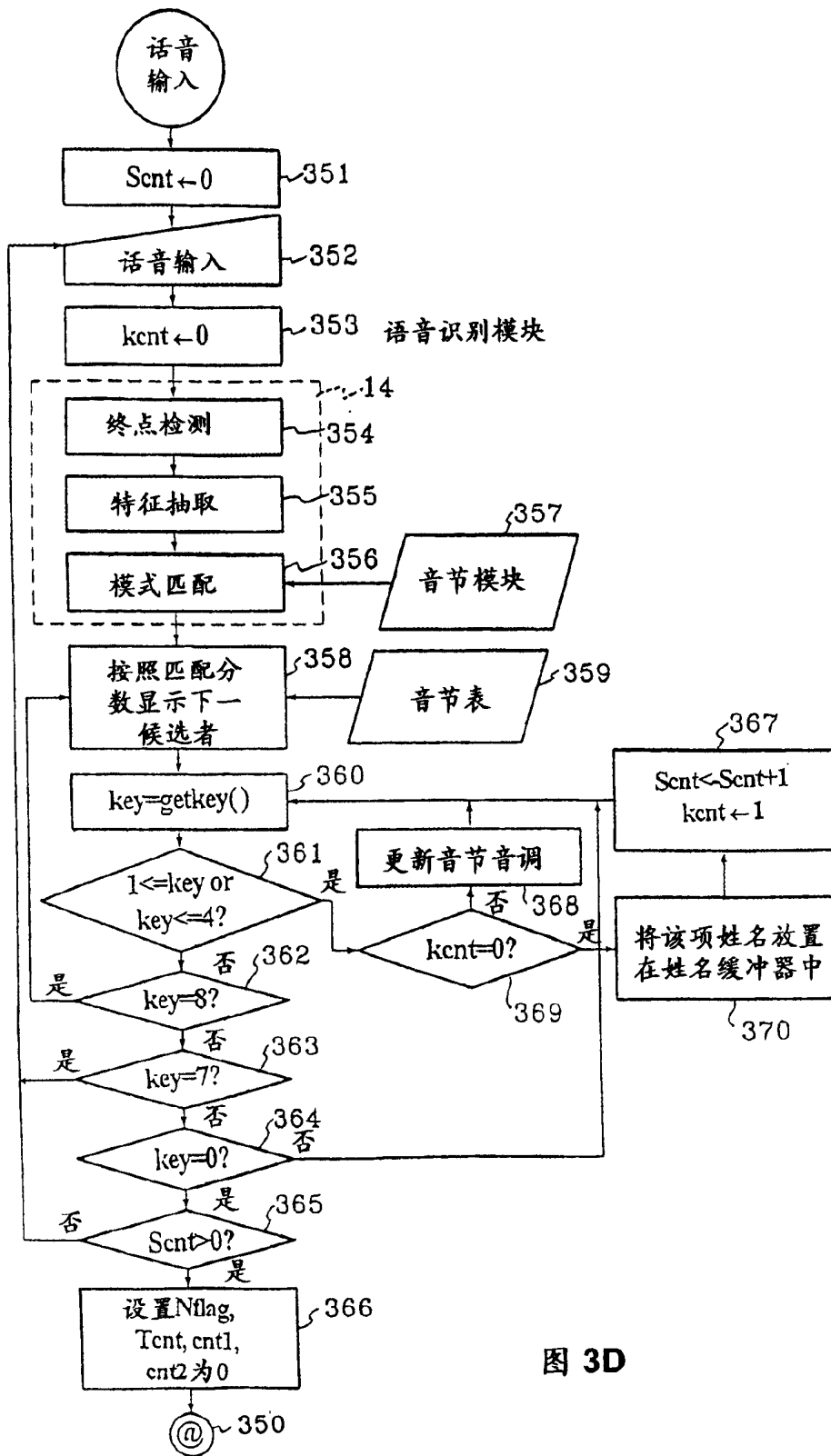


图 3D

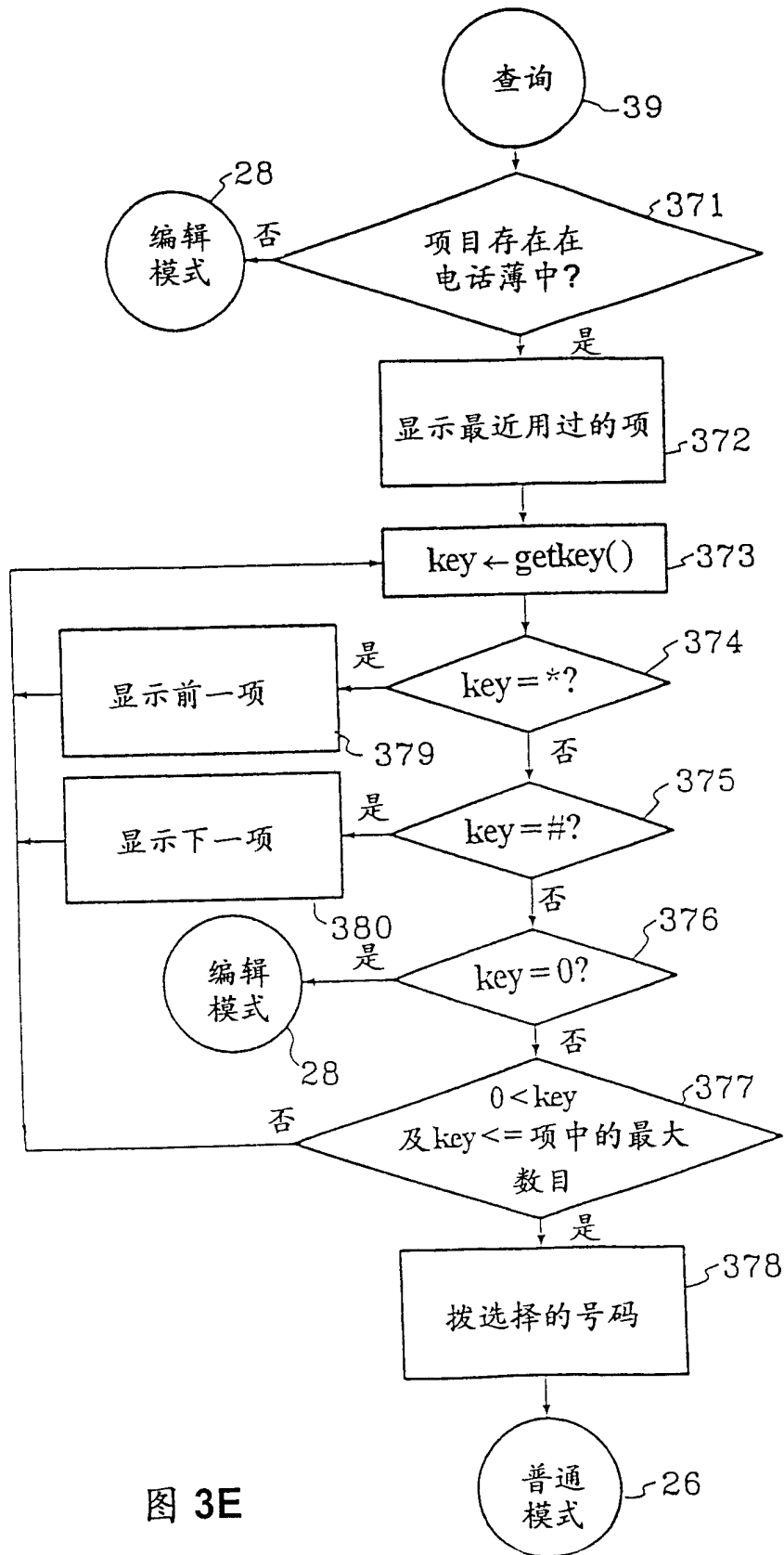


图 3E

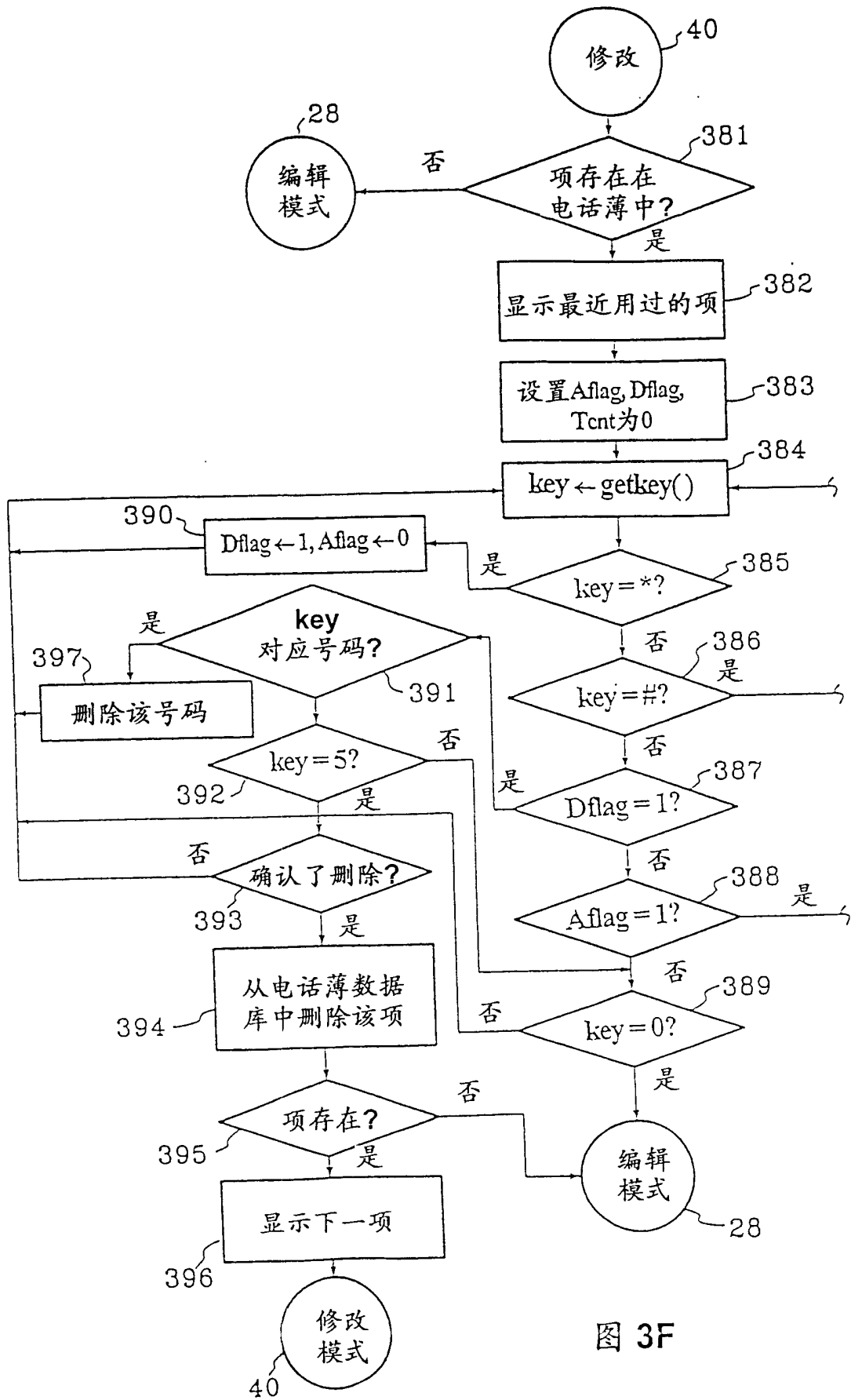


图 3F

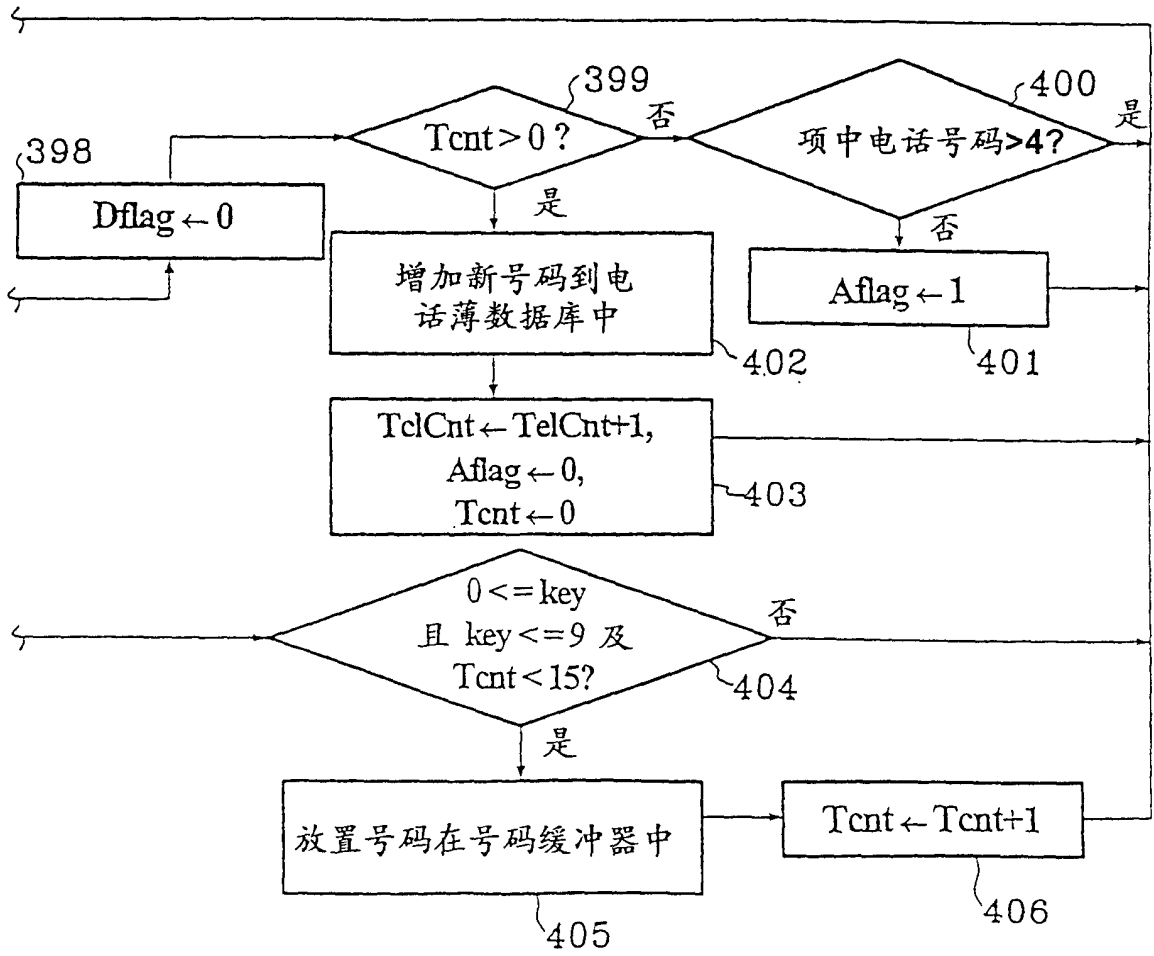


图 3G

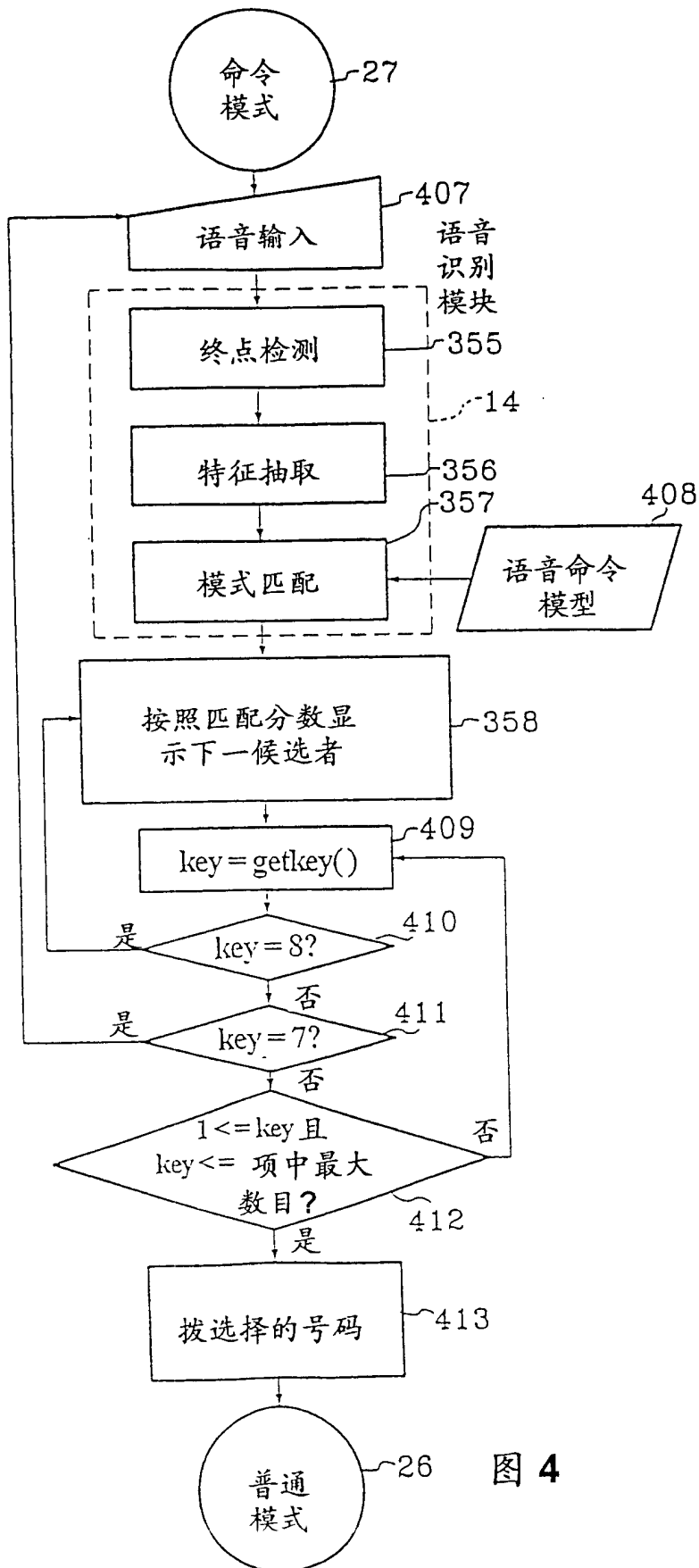


图 4