

[19] 中华人民共和国国家知识产权局

[51] Int. Cl⁷

G06K 9/00

G06F 3/16

G06F 17/00



[12] 发明专利申请公开说明书

[21] 申请号 03124963.9

[43] 公开日 2005年3月30日

[11] 公开号 CN 1601548A

[22] 申请日 2003.9.23 [21] 申请号 03124963.9

[71] 申请人 摩托罗拉公司

地址 美国伊利诺斯州

[72] 发明人 张亚昕 何昕 任晓林 孙放

[74] 专利代理机构 中原信达知识产权代理有限责
任公司

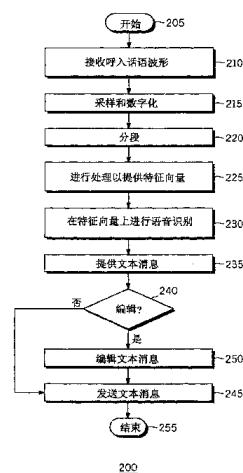
代理人 谢丽娜 张天舒

权利要求书 2 页 说明书 8 页 附图 4 页

[54] 发明名称 用于提供文本消息的方法和装置

[57] 摘要

本发明公开了一种用于提供文本消息的方法和装置，其包括电子设备(100)的输入端接收话语(步骤210)。然后，在由用户设定的消息模板所引导的话语(步骤230)上进行语音识别，其中该消息模板存储在与电子设备(100)相关联的存储器(155)中。通过对该话语和模板之一进行匹配的方式定义语音识别，以建立一个匹配模板。然后，通过该匹配模板提供文本消息(步骤235)。



ISSN 1008-4274

1. 一种用来提供文本消息的方法，所述方法包括如下步骤：
在电子设备的输入端接收话语；
- 5 在由用户设定的消息模板所引导的所述话语上执行语音识别，所述消息模板存储在与所述电子设备相关联的存储器中，其中通过对所述话语和所述模板之一进行匹配的方式定义语音识别，以建立一个匹配模板；以及
 通过所述匹配模板提供文本消息。
- 10
2. 如权利要求 1 所述的方法，其中所述消息模板中至少有一个包含固定的语言成分。
3. 如权利要求 1 所述的方法，其中所述消息模板中至少有一个包含可变的语言成分。
- 15
4. 如权利要求 1 所述的方法，其中所述消息模板中至少有一个同时包含固定和可变的语言成分。
- 20
5. 如权利要求 1 所述的方法，其中所述文本消息是 SMS 消息。
6. 如权利要求 1 所述的方法，其进一步包括通过接收从所述电子设备的键盘键入的字符来编辑所述的用户设定的消息模板的步骤。
- 25
7. 如权利要求 1 所述的方法，其中所述文本消息的一部分是所述话语的转录。
8. 如权利要求 1 所述的方法，其中所述文本消息全部是所述话语的转录。

30

9. 一种用于提供文本消息的电子设备，所述设备包括：
麦克风，用于接收话语；
非易失性存储器，用于存储消息模板；以及
处理器，用于进行由所述消息模板引导的所述话语的语音识别，
5 所述处理器操作地对所述话语和所述模板之一进行匹配以建立一个匹配模板，并通过所述匹配模板提供文本消息。
10. 如权利要求 9 所述的方法，其中所述消息模板中至少有一个包含固定的语言成分。
- 10 11. 如权利要求 9 所述的方法，其中所述消息模板中至少有一个包含可变的语言成分。
12. 如权利要求 9 所述的方法，其中所述消息模板中至少有一个
15 同时包含固定和可变的语言成分。
13. 如权利要求 9 所述的设备，其中所述文本消息是 SMS 消息。
14. 如权利要求 9 所述的设备，其进一步包括一个用于编辑所述
20 消息模板的键盘。
15. 如权利要求 9 所述的设备，其中所述处理器操作地对所述话语和多个所述模板进行匹配并计算每个所述模板的似然值。

用于提供文本消息的方法和装置

5 技术领域

本发明主要涉及一种利用话音来提供文本消息的方法和装置。本发明对于（但不排它的）利用话音输入来提供文本消息的方式尤其有好处，其中话音输入是在带有有限内存和计算能力的便携式电子设备上进行处理。

10

背景技术

通常采用短消息业务（SMS）格式的短文本消息在无线通信领域内应用非常广泛。每个月有几十亿的短文本消息被发送，通常从一部移动电话发向另一部移动电话。这种文本消息的流行有很多原因。消息一般只需花费一分钟移动电话呼叫费用的一小部分且无需发送或接收占线音。因此，发送用户可以在方便的时候才建立和发送消息，而接收用户则可以在方便的时候才接收并读取消息。

15

通常通过在移动电话的键盘内键入字符的方式来建立文本消息。但是，使用这种小型的非标准键盘来编写消息可能很不方便，且通常比采用原尺寸的标准键盘的方式需要更多的时间。但是，在移动电话上连接原尺寸的标准键盘当然是不切实际的。因此，需要采用更有效的方法来编写短文本消息。

20

进而，虽然现存多种公知的语音识别系统，但大多数都不适用于诸如移动电话的便携式电子设备。这是因为现有技术的语音识别系统通常需要比便携式电子设备中所能提供的更大的处理功率和内存。

25

现有技术的封闭式词汇语音识别系统和方法采用预先设定的、固定的词汇表。在使用中，固定词汇表可能很大但并不完全，因此，例

30

如，它可能不包括人的姓氏或许多地名。相反，开放式词汇语音识别系统和方法使用可变的词汇表，用户可以向其中添加新的单词和词组或者删除单词和词组。但是，现有的开放式词汇语音识别系统和方法需要相对较高的计算开销，这对于诸如数字个人助理（PDA）、无线电话和其它便携式设备的便携式电子设备来说，可能承受不了。

在本说明书以及权利要求中，术语“包含”或者类似术语希望表示非排它性的包含，以使包含一组组件的方法或装置并不是仅仅只包括这些组件，而是可以很好地包括其它未列出的组件。

10

发明内容

根据本发明的一个方面，提供了一种用于提供文本消息的方法。该方法包括在电子设备的输入端接收话语的步骤。然后，在由用户设定的消息模板所引导的话语上进行语音识别，该消息模板存储在与电子设备相关联的存储器中，其中通过对该话语和模板之一进行匹配的方式定义语音识别，以建立一个匹配模板。然后，通过该匹配模板提供文本消息。

15

至少一个消息模板可以包含固定的语言成分。

20

至少一个消息模板可以包含可变的语言成分。

至少一个消息模板可以同时包含固定和可变的语言成分。

25

文本消息可以是 SMS 消息。

上述方法也可以包括通过接收由电子设备的键盘键入的字符来编辑由用户设定的消息模板的步骤。

30

文本消息的一部分可以是话语的转录（transcription）。

文本消息可以全部是话语的转录。

5 根据本发明的另一方面，提供了一种用于提供文本消息的电子设
备。该电子设备包括一个用于接收话语的麦克风；用于存储消息模板
的非易失性存储器；用于执行由该消息模板引导的话语的语音识别的
处理器，其中处理器对话语和模板之一进行匹配以建立一个匹配模
板，并通过该匹配模板提供文本消息。

10 对于该电子设备，消息模板还可以包括固定和/或可变的语言成
分。

对于该电子设备，文本消息可以是 SMS 消息。

15 该电子设备可以包括一个键盘用于编辑消息模板。

该电子设备可以对话语和多个模板进行匹配以计算每个模板的似
然值。

20 附图说明

为了使本发明易于理解并付诸实现，现在将参考附图，为所示出
的优选实施例提供参考。在附图中：

图 1 是根据本发明的无线电话的示意性框图。

25 图 2 是示出根据本发明的、用来提供、编辑和发送文本消息的方
法的流程图。

图 3 是示出根据本发明的、用来提供备选消息模板列表的方法的
流程图。

图 4 是示出根据本发明的、使用户可以编辑现有消息模板和在静
态可编程存储器中保存新模板的方法的流程图。

30

具体实施方式

5 参考图 1，示出了无线电话 100，它包括连接到处理器 110 并与其通信的射频通信单元 105。以显示屏 115 形式呈现的输入/输出 (I/O) 接口、键盘 120、麦克风 190 以及扬声器 195 也被连接到处理器 110 并与其通信。

10 处理器 110 包括 编码器/解码器 125 以及与之相关联的存储数据的只读存储器 (ROM) 130，用于对可能由无线电话 100 发送或接收的话音或其它信号进行编码和解码。处理器 110 还包括微处理器 135、随机存取存储器 (RAM) 150、静态可编程存储器 155 和可移动 SIM 模块 160，其中微处理器 135 通过公共数据和地址总线 140 连接至编
15 码器/解码器 125 及其相关联的字符只读存储器 (ROM) 145。静态可编程存储器 155 和 SIM 模块 160 都可以存储 (在其它消息中) 选中的呼入文本消息、电话本数据库以及呼出文本消息的模板 (将在下文中详细描述)。微处理器 135 具有用于连接至键盘 120、显示屏 115 和提示模块 165 的端口，其中提示模块 165 通常包含扬声器、振动马达和相应驱动器。字符只读存储器 145 存储用于对可由通信单元 105 接收且在键盘 120 上输入的文本消息进行解码或编码的代码。

20 射频通信单元 105 是一个组合式接收器和发送器，其具有共用天线 170。通信单元 105 具有经射频放大器 180 连接至天线 170 的收发器 175。收发器 175 还连接至组合式调制器/解调器 185，调制器/解调器 185 把通信单元 2 连接至处理器 110。

25 参考图 2，该流程图示出了包括方法 200 的本发明的一个实施例，方法 200 通过采用无线电话 100 来提供、编辑和发送文本消息。在开始步骤 205 中调用方法 200。在步骤 210 中，在诸如麦克风 190 的电话 100 的输入设备上接收话语。然后，处理器 110 在步骤 215 中对该话语波形进行采样和数字化，接着在步骤 220 中对其进行分段，然后
30 在步骤 225 中对其进行处理以提供表示该波形的特征向量。应当注意，

步骤 215、步骤 220 和步骤 225 在本领域内是公知的，因此无需对此进行详细说明。

5 接下来，在步骤 230 中，对在步骤 225 中产生的特征向量进行语音识别。由存储在设备 100 的静态可编程存储器 155 中的用户设定的消息模板引导该语音识别。消息模板在下文中将有详细说明。然后，方法 200 在步骤 235 中为用户提供文本消息。该消息可通过使用诸如设备 100 的显示屏 115 或扬声器 195 的 I/O 接口之一来向用户提供。

10 向用户提供消息之后，接着用户在步骤 240 中能够决定是否编辑该消息。如果用户决定不对该消息进行编辑，则该消息在步骤 245 中以诸如 SMS 的消息格式发送出去。但是，如果用户在步骤 240 决定编辑该消息，则消息在步骤 245 中送出之前，先在步骤 250 中进行编辑。在本发明的各种实施例中，用户可以以几种不同的方式编辑消息，
15 包括通过扬声器 195 的语音式的编辑或通过键盘 120 的键入式的编辑。然后，在步骤 255 结束方法 200。

在本发明的另外一种实施例中，在上述的语音识别步骤 230 之后，提供文本消息的步骤 235 可以包括向电话 100 的用户提供备选消息模板列表，用户可从中挑选出最适合于所希望的文本消息的模板。
20

图 3 是示出用来向用户提供这种备选模板列表的方法 300 的流程图。当用户向键盘 120 或麦克风 190 输入命令时，在开始步骤 305 中调用方法 300。首先，方法 300 包括处理器 110 在步骤 310 中从可用消息模板列表中选出一个消息模板。然后，在步骤 315 中，把这个选中的模板和在方法 200 的步骤 225 中提供的特征向量进行比较。接着，处理器 110 在步骤 320 中计算似然值，估计选中的模板的各个方面和输入话语的特征向量之间的匹配度。
25

30 然后，处理器 110 在步骤 325 中判断这个似然值是否高于一个已

5 设的门限。这个门限可以由处理器 110 自动算出，或者由电话 100 的用户预先设定。如果选中的模板的似然值低于已设的门限，模板在步骤 330 中被拒绝。但是，如果选中模板的似然值高于已设的门限，那么在步骤 335 中，则认为该模板与输入话语之间有合理匹配度，并将其添加到备选模板列表中。无论选中的模板是否被拒绝或是添加到备选模板列表中，方法 300 都继续进行到步骤 340，在该步骤中，处理器 110 判断所有可用的模板是否都已经评估过。如果并非所有可用的模板都被评估过，则方法 300 在步骤 345 挑选下一个消息模板并返回到步骤 315，在步骤 315 中，将下一个模板与输入话语的特征向量进行比较。如果在步骤 340 中已评估过所有模板，则方法 300 继续进行到步骤 350 并向用户提供所有备选模板的列表。备选模板可以通过使用诸如设备 100 的显示屏 115 或扬声器 195 的 I/O 接口之一来向用户提供。然后，在步骤 355 结束方法 300。

15 根据本发明的一个实施例，电话 100 的用户不限于只使用由设备 100 的制造商提供的模板。相反，设备 100 的用户可以编辑存储在静态可编程存储器 155 中的现有模板，以建立自己个性化的消息模板。

20 参考图 4，示出了使用户可以编辑现有模板并在静态可编程存储器 155 中保存新模板的方法 400。当用户向键盘 120 或麦克风 190 输入命令时，在开始步骤 405 中调用方法 400。在步骤 410 中，通过诸如显示屏 115 或扬声器 195 的 I/O 接口向设备 100 的用户提供现有模板的列表。接着，用户在步骤 415 中通过使用诸如麦克风 190 或键盘 120 的 I/O 接口来选择所需的消息模板。接下来，在步骤 420 中，用户再次通过使用诸如麦克风 190 或键盘 120 的 I/O 接口来编辑模板。25 最后，在步骤 425 中，用户在静态可编程存储器 155 中保存已编辑好的模板。然后，在步骤 430 结束方法 400。

30 编辑消息模板的其它方法也属于本发明的范围内，其中包括通过使用诸如 USB 线缆的通信信道把电话 100 连接至主机，随后把已编辑

好的模板下载或补充（flash）至静态可编程存储器 155。

本发明的方法可进一步包括包含固定和可变语言成分的消息模板。固定语言成分在用户选中一个模板并发送消息时不会发生改变。

5 而可变语言成分可以在不同的消息中（from message to message）由用户修改。固定和可变语言成分的使用大大提高了电话 100 的有限处理功率和内存的使用效率（leverage）。

例如，与见面请求相关的短文本消息的特定模板可能如下所示：

10 “Meet me at \$PLACE at \$TIME.”。在这个句子中，固定语言成分由下划线标出，而可变语言成分则大写且以“\$”开头。那么，模板的不同用户就可以编辑诸如\$PLACE 等变量以适应其特定的环境。例如，大学生可能把可变的\$PLACE 设定为：\$PLACE = sp | library（图书馆）
| dormitory（宿舍）| cafeteria（自助餐厅）等等而律师可能把可变的
15 \$PLACE 设定为：\$PLACE = sp | office（办公室）| courthouse（法院）
| home（家）等等

在上面的句子中，“sp”指停顿或无语音，而“|”指逻辑算子“OR”。

20 在本发明中可能用到的消息模板的另一示例是：“**Happy \$FESTIVAL.**”。这个句子中的可变语言成分\$FESTIVAL 可能会被用户编辑成：

\$FESTIVAL = sp | birthday（生日）| new year（新年）| thanksgiving
（感恩节）等等

25

通过使用开放式的词汇语音识别，电话 100 能够识别出由用户输入的已编辑的可变语言成分。由于可变语言成分由变量的离散集组成，语音识别处理的开销和内存需求被缩减到最小。因此上述方法特别适用于只具有有限的处理和内存资源的设备，诸如移动电话。

30

对包括固定和可变语言成分的模板的使用在几个方面提高了语音识别系统的效率。其一，由于在静态可编程存储器 155 中存储的模板的数量相对于与自然语言句子结构相关联的句子排列的几乎无限的数量来说是不多的，所以通常可以快速而有效地识别出特定模板中的固定语言成分。其二，由于消息模板中的可变语言成分的句中位置能够自动识别出可能答复语的离散集，所以也可以高效地识别出可变语言成分。例如，以上述的“**Happy \$FESTIVAL**”消息模板为例，固定语言成分“**Happy**”可作为一个信号，使处理器 110 知道：麦克风 190 接收到的随后的话音输入将是可变语言成分“**\$FESTIVAL**”。

10

虽然本发明的上述实施例是通过无线电话 100 来说明的，但是本发明的方法和装置也可以包括诸如数字个人助理（PDA）之类的可提供文本消息的其它电子设备。

15

因此，本发明简化了从便携式电子设备上提供和发送文本消息所需的步骤。文本消息可通过话音输入而不是从小型键盘键入字符的方式来提供。此外，本发明可以包括开放式词汇语音识别，以避免现有技术的封闭式词汇语音识别对内存的高度需求。开放式词汇语音识别采用独立于扬声器的子词（sub-word）声学模型，设计成涵盖一种语言的所有声音现象（occurrence）或音素（phoneme）。因此，用户不限于只使用预先设定的词汇，也能够编辑如上所述的可变语言成分，以包括诸如人名和地名等词典里查不到的单词。因此本发明提供的文本消息可以高度地个性化。

20

25

上述的详细说明只提供了优选的示范性实施例，且不视作对本发明的范围、适用性或配置的限制。更合适的说，该优选的示范性实施例的详细说明给本领域的普通技术人员提供了实现本发明的优选的示范性实施例的指导性说明。应当理解，在不背离所附权利要求中所要求的本发明的精神和范围的前提下，可以对组件的功能和排列做出多种的改变。

30

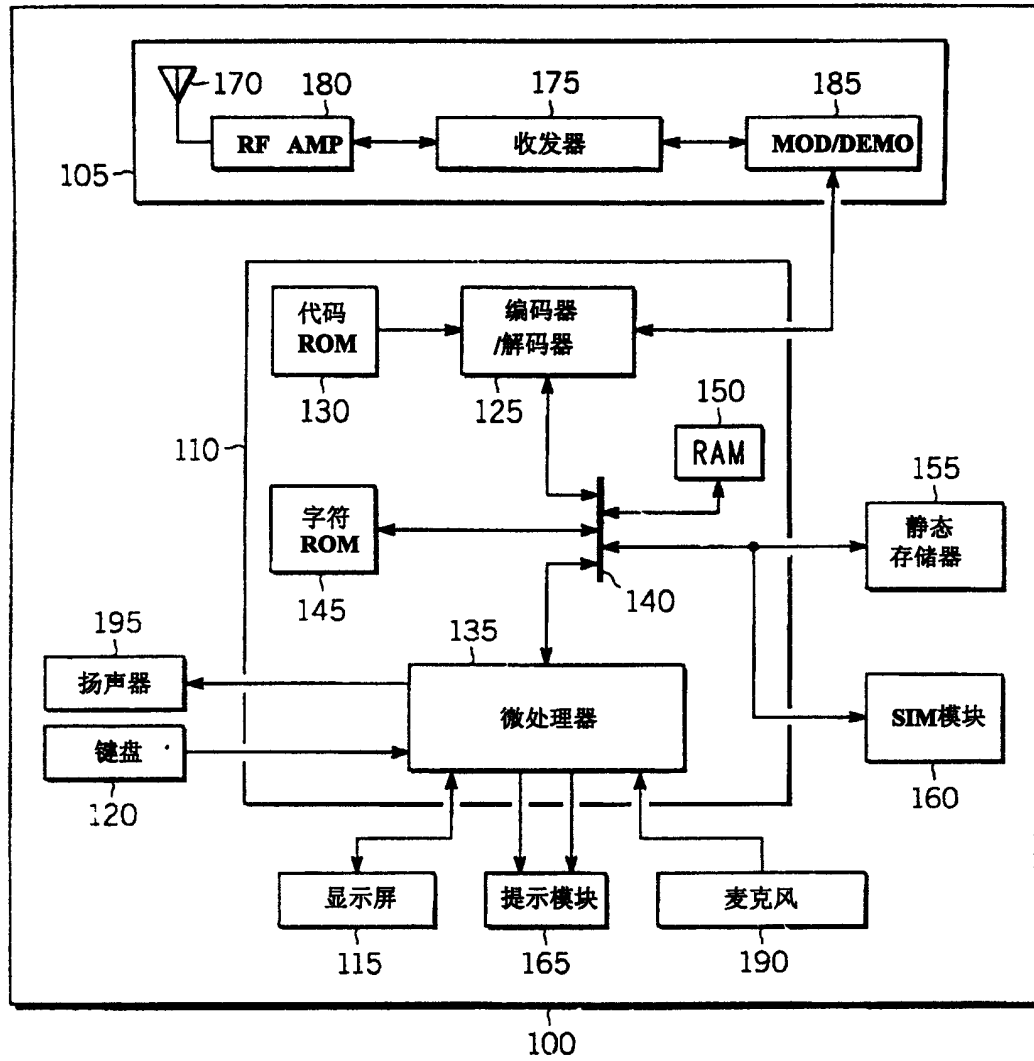
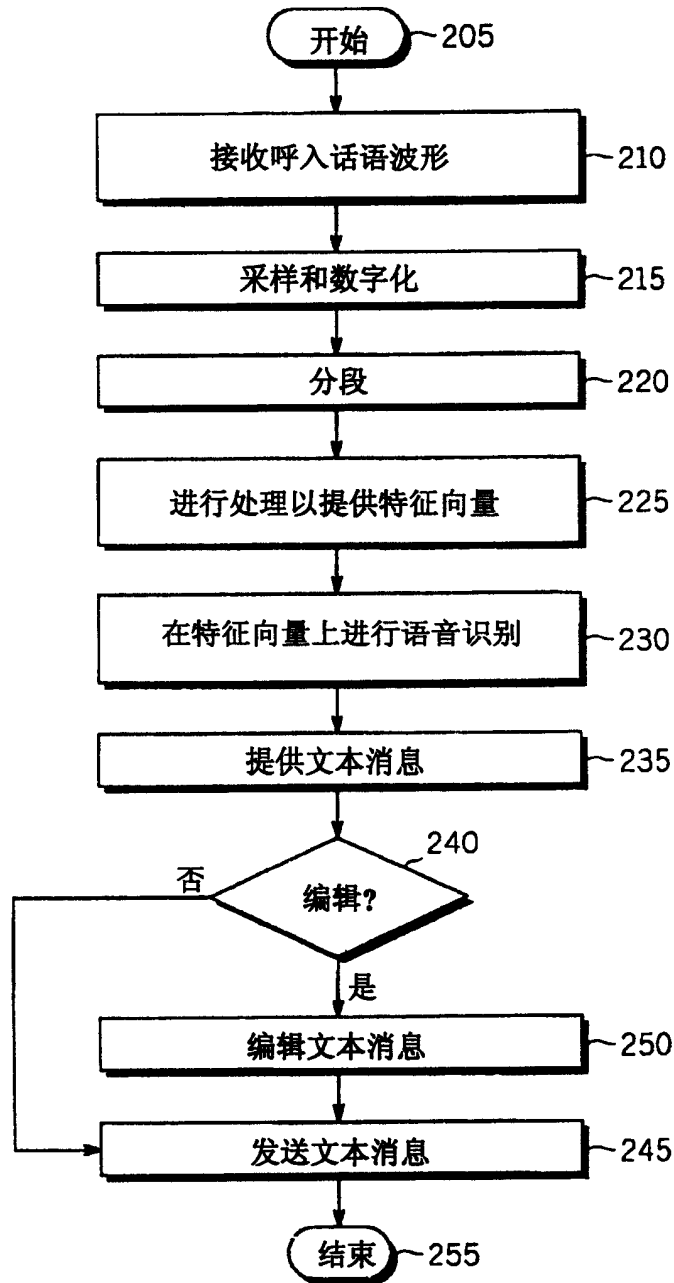
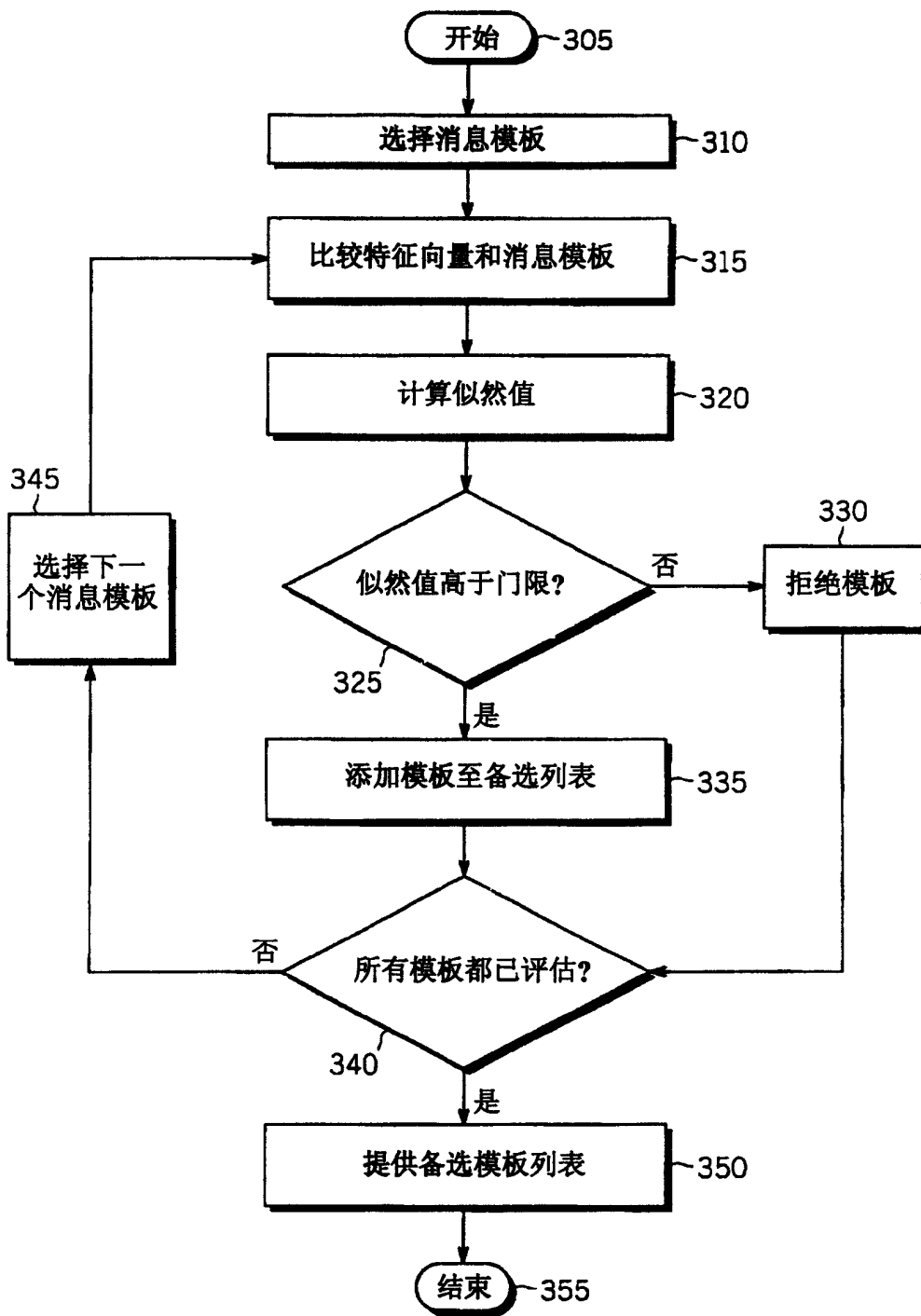


图1



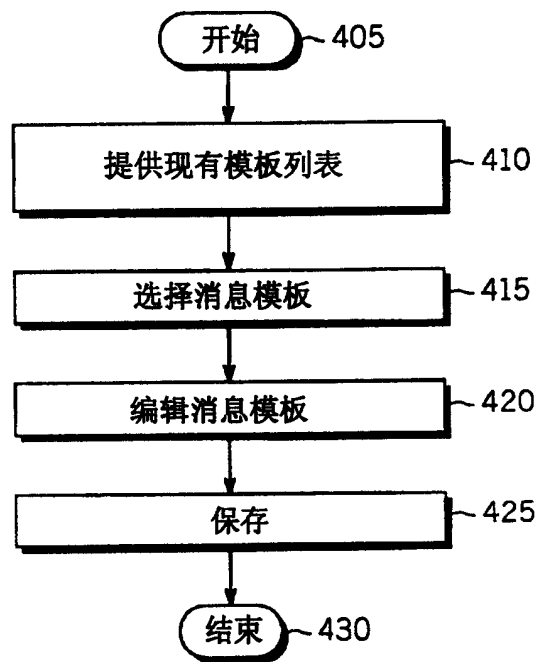
200

图2



300

图3



400

图4