



# [12] 发明专利说明书

[21] ZL 专利号 00102409.4

[45] 授权公告日 2004 年 10 月 13 日

[11] 授权公告号 CN 1171201C

[22] 申请日 2000.2.23 [21] 申请号 00102409.4

[30] 优先权

[32] 1999. 2. 23 [33] US [31] 09/256,279

[71] 专利权人 摩托罗拉公司

地址 美国伊利诺斯

[72] 发明人 戴维·E·彻瓦利尔

亨利·L·卡泽基

审查员 刘红梅

[74] 专利代理机构 中国国际贸易促进委员会专利

商标事务所

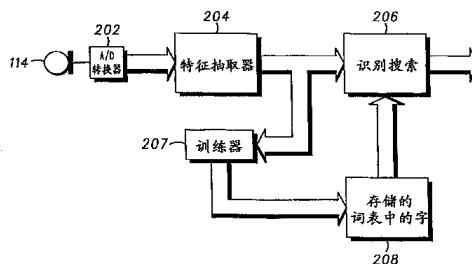
代理人 张 维

权利要求书 1 页 说明书 12 页 附图 5 页

[54] 发明名称 语音识别系统及其方法

[57] 摘要

一个语音识别系统(204, 206, 207, 208)产生一个可变的拒绝精度, 该精度作为在训练期间测量的至少一个背景噪声电平以及在输入表达期间作出的噪声信号测量的函数, 前述输入表达在识别模式的运行期间作出。一个字入口补偿被分配作为该可变拒绝精度的函数。



- 1.一种操作语音识别系统的方法,包括以下的步骤;  
产生一个可变的拒绝精度,该精度是在训练期间测量的背景噪声电平以及在输入表达期间作出的噪声信号测量的函数,前述输入表达在识别模式的运行期间作出;  
得到一个字入口补偿,该字入口补偿是该可变拒绝精度的函数。
- 2.根据权利要求 1 所定义的方法,其中产生一个可变的拒绝精度的步骤包括对于一个模式在至少部分训练表达期间测量噪声的步骤。
- 3.根据权利要求 1 所定义的方法,还包括有选择地更新来自于训练表达的噪声特征的步骤。
- 4.根据权利要求 1 所定义的方法,还包括用一种模式存储训练期间的噪声统计量的步骤,以便它们能够为识别算法所利用。
- 5.根据权利要求 3 所定义的方法,其中当在免提模式中训练时,噪声统计量不被更新。
- 6.根据权利要求 3 所定义的方法,还包括产生信噪比的步骤,其中如果信噪比低于预定的大小,则训练被禁止。
- 7.根据权利要求 1 所定义的方法,其中在识别期间,如果对于一个表达没有噪声统计量可以利用,则当将定位算法应用于该表达时,识别算法缺省为最小的精度要求。
- 8.根据权利要求 1 所定义的方法,其中在识别期间,将输入噪声能量特征与参考噪声统计量相比较,并且计算出噪声比。
- 9.根据权利要求 8 所定义的方法,其中词表范围外的拒绝算法的精度基于该噪声比而被选择。
- 10.根据权利要求 1 所定义的方法,其中利用零平均值的单个状态无用信息模式与语音标记模式并行地实现一最佳路径的置信度量。

## 语音识别系统及其方法

### 技术领域

本发明涉及语音识别。

### 背景技术

依赖于扬声器的语音识别系统利用特征抽取算法在输入语音的帧上以及代表每帧的输出特征矢量上执行信号处理。这种处理在帧频进行。该帧频一般在 10 至 30ms 之间，在此将以 20ms 的持续时间为例。已经知道有许多不同的特征在语音识别系统中使用。

一般来说，训练算法利用从字或短语的一个或多个表达的抽样语音中抽取的特征来生成该词或短语的模式参数。然后将该模式存储在模式存储器中。这些模式后来在语音识别期间使用。识别系统将未知表达的特征与存储的模式参数相比较，以确定最佳匹配。然后将最佳匹配模式作为该识别系统的结果输出。

已经知道利用基于识别系统的隐含马尔柯夫模式 (HMM) 进行这种处理。HMM 识别系统将表达的各个帧分配给 HMM 的各个状态。选择产生最大概率或比数的状态至帧的分配作为最佳匹配。

许多语音识别系统并不能区分有效表达与无效表达。相反地，这些系统选择其中一个最接近匹配的存储模式。有些系统利用力图探测及拒绝无效表达的词表范围外的 (Out-of-Vocabulary) 拒绝算法。由于词表的动态大小及未知的构成，在小的词表中，依赖于扬声器的语音识别系统是一个难题。这些算法在噪声条件下退化，以至于噪声条件下的错误拒绝的数目增加。

在实际中，词表范围外的拒绝算法必须平衡由正确拒绝无效表达及错误拒绝有效表达所测量出来的性能。错误拒绝率可以在用户满意中起到至关重要的作用，因为频繁的错误拒绝，与不正确的匹配一样，将引起失败。因而，词表范围外的拒绝起到了满足用户识别期望的平

衡作用。

相应地，已经知道基于噪声电平计算拒绝门限。例如，已经知道在探测第一个语音帧之前测量噪声电平。该门限是从测量中计算出来。如果这个字的参考模式与输入语音模式之间的差值大于拒绝门限，则该输入被拒绝。因而这种系统依赖于随机噪声输入电平。不能依赖这种测量来产生有意义的拒绝判决。

### 发明内容

因此，在语音识别系统中需要一种改进的方法为拒绝表达提供一种依据。

本发明提供了一种操作语音识别系统的方法，包括以下的步骤：产生一个可变的拒绝精度，该精度作为在训练期间测量的背景噪声电平以及在输入表达期间作出的噪声信号测量的函数，前述输入表达在识别模式的运行期间作出；得到一个字入口补偿，该字入口补偿作为该可变拒绝精度的函数。

### 附图说明

- 图 1 为以方框图形式说明无线通信装置的电路示意图；
- 图 2 为以方框图形式说明图 1 装置中语音识别系统的电路示意图；
- 图 3 为带有两个节点的基础网络的示意图；
- 图 4 为用于说明训练的流程图；
- 图 5 给出了窗口及其中的相应帧；
- 图 6 为说明识别的高阶流程图；
- 图 7 为说明识别期间训练的流程图；
- 图 8 为说明补偿函数的示意图。

### 具体实施方式

本发明具有在训练及识别期间依赖于背景噪声电平的可变拒绝精度。在训练期间，噪声特征生成于训练表达中。根据噪声特征更新增量噪声参考平均值。这些统计量存储在存储器中，以使它们可为识别算法所利用。当在免提模式中训练时，因为背景噪声的较高电平，噪声统计量并不被更新。如果没有可以利用的噪声统计量，则识别算法

缺省为最小精度。

在识别期间，输入噪声能量特征与参考噪声统计量相比较，并且计算出噪声比。然后根据该噪声比选择词表范围外的拒绝算法的精度。本发明在出现噪声时有助于防止有效表达的错误拒绝。

在两级定位算法识别搜索中，精度参数为字入口补偿（penalty）。随着零平均值的单个状态无用信息模式与语音标记模式互相平行，实现了最佳路径的置信测量。

图 1 中公开了一种可以有效利用本发明的装置 100。在这里为了说明起见，装置 100 被描述为手提无线电话，但是也可以为计算机，

个人数据助理，或者是任何可以有利地采用语音识别的装置，尤其是可以利用对语音识别系统有效的存储器的装置。所述的无线电话包括耦合至天线 106 的发射机 102 及接收机 104。发射成功机 102 及接收机 104 耦合至呼叫处理器 108，该处理器执行呼叫处理功能。呼叫处理器 108 可以利用数字信号处理器（DSP）、微处理器、微控制器、可编程逻辑部件、两个或多个上述装置的结合或者任何其他合适的数字电路来实现。

呼叫处理器耦合至存储器 110。存储器 110 包括 RAM，电可擦可编程只读存储器（EEPROM），只读存储器（ROM），快闪 ROM，或类似装置，或两个或多个这些存储器种类的组合。存储器 110 支持呼叫处理器 108 的操作，所述操作包括语音识别操作，并且必须包括电可变存储器以支持状态转移路径存储器。ROM 可以被用来存储装置运行程序。

声频电路 112 提供从话筒 114 向呼叫处理器 108 的数字化信号。响应于来自呼叫处理器的数字信号声频电路 112 驱动扬声器 116。

呼叫处理器 108 耦合至显示处理器 120。如果需要其他的处理器支持装置 100，则显示装置为可选择的。尤其是，显示装置 120 向显示器 126 提供显示控制信号，并且接收来自键 124 的输入。因此，显示处理器 120 可以利用微处理器，微控制器，数字信号处理器，可编程逻辑部件，它们的结合或类似装置来实现。存储器 122 耦合到显示处理器以支持其中的数字逻辑。存储器 122 可以利用 RAM，EEPROM，快闪 ROM，或类似装置，或两个或多个这些存储器种类的组合来实现。

参照图 2，话筒 114 所接收的声频信号在声频电路 112 的数模转换器 202 中被转换为数字信号。本领域的技术人员将认识到，声频电路 112 提供诸如过滤的其他信号处理，为了简洁起见在此没有描述这些处理。呼叫处理器 108 在被处理的数字信号上执行特征抽取 204，并且产生一组表示用户表达的特征矢量，前述被处理数字信号表示由话筒 114 输出的模拟信号。产生的特征矢量用于每个短时分析窗口。

该短时分析窗口为一帧，在本申请的示范例中为 20ms。因而每帧有一个特征矢量。处理器 108 将该特征用于语音识别 206 及训练 207。

在训练中，表达的特征矢量被用于创建 HMMs 形式的模板，该模板存储于存储器 208 中。在语音识别中，表示输入表达的特征矢量与存储器 208 中存储的词表中字的模板相比，以确定用户所说的内容。该系统可能输出最佳的匹配，一组最佳的匹配，或也可能是，没有匹配。存储器 208 最好是存储器 110（图 1）的非易失性存储器部分，例如可以为 EEPROM 或快闪 ROM。如本申请所使用的，“字”可以是多于一个字的字，诸如“John Doe”，或者诸如“call”的单个字。

特征抽取器 204 一般在输入语音的帧上执行信号的处理，并且以帧频输出代表每帧的特征矢量。帧频一般在 10ms 与 30ms 之间，例如可以为 20ms 的持续时间。训练器 207 利用从字或短语的一个或多个表达的抽样语音中抽取该特征，以产生用于该字或短语的模式参数。然后将该模式存储在非易失性存储器 208 的模式存储器中。该模式的大小直接依赖于特征矢量的长度，较长特征矢量的长度需要较大的存储器。

接着，存储于存储器 208 中的模式在识别 206 中被利用。该识别系统执行一个在未知表达的特征与存储的模式参数之间的比较，以确定最佳匹配。从识别系统输出的最佳匹配模式作为结果。

参照图 3，图 3 示出了表示语音识别的基础网络。节点  $N_1$  及  $N_2$  由 HMMs 模式所连接，该 HMMs 模式由弧  $A_1-A_N$  加上无用信息模式弧  $A_{GM}$  表示。弧  $A_1-A_N$  表示所有的 HMM 模式，前述 HMM 模式在语音识别系统中已经被训练，并且存储在存储器 208 中。无用信息模式弧表示单一状态无用信息的模式基准。

节点  $N_1$  包括一个单一状态噪声模式  $A_1^{noise}$ 。节点  $N_2$  类似地包括一个单一状态噪声模式  $A_2^{noise}$ 。该识别系统采用识别算法从弧  $A_1-A_N$  以及  $A_{GM}$  中选择其中的一个作为最佳匹配，或者可选择地识别出没有匹配（即如果没有探测到语音）。如果  $A_{GM}$  为最佳弧，该输入作

为无效信号被拒绝。

现在参照图 4，图 4 描述了训练的过程。首先，如步骤 402 所示，执行主训练 207 以得到每个表达，或存储在存储器 208 中的状态模式  $A_1-A_N$ 。已经知道有许多不同的方法用于创建 HMM 模式。在图 4 的说明中，每个弧为从左至右，HMM 模式不带有状态跳跃，因此只允许自环路及单步骤转移。这种模式推导的简要说明将在下文中描述。本领域的技术人员将认识到，这些弧可以是其他已知模式的弧，并且被其他已知方法所知。

首先，在特征抽取器 204 中抽取特征。预见到特征抽取器将产生用于表达的每个帧的对数倒频谱（cepstral）及  $\delta$  对数倒频谱系数。本领域的技术人员将认识到，有许多计算对数倒频谱的方法，并且用于估计它们的导数，可以使用任何合适的技术来得到这些系数。图 5 中帧  $F_1$  至  $F_N$  在该窗口期间被产生，每个帧构成特征。有些帧表示噪声，特征抽取器从这些帧中产生噪声能量特征。其他帧表示语音信号部分。

回到图 4，在步骤 604 中，处理器 108 在训练 207 期间计算每个弧形模式的噪声特征，如步骤 604 所示。该噪声测量从多个特征矢量中得到，该特征矢量在捕获窗口的开端及末端期间产生。尤其是，希望利用在表达的开始阶段及末尾阶段期间测量的特征矢量的平均值。例如，可以利用捕获窗口的第一个 160ms 的“开始平均值”（Savage），以及最后 160ms 的“末端平均值”（Eavage）。图 5 所示的捕获窗口包括开始阶段及末尾阶段，噪声特征矢量在此期间被存储。例如，捕获窗口可能为 2 秒钟长，其表示字的最长持续时间。该捕获窗口可以是固定或可变的长度，这取决于输入表达的期望长度以及实现方式的存储器约束。

在步骤 404 中得到噪声特征的处理器 108 在步骤 404 中确定该该装置是否处于免提模式。该装置可以包括一个指示该装置处于免提模式的状态标记，该免提模式由用户通过键盘菜单启动，或者该装置可以包括一个机械连接器，当该装置被连接到一个免提套件时，该机械

连接器启动一个转换开关。

如果该装置不处于免提模式，该处理器在训练（对于每个表达，训练被独立地实现）期间计算噪声特征  $X_{nz}$ ，该噪声特征是  $Savg_e$  和  $Eavg_e$  的最小值（即  $\min(Savg_e, Eavg_e)$ ），如步骤 410 所示。对于输入语音的每个帧来说，能量值可以从其抽样中计算出来。 $Savg_e$  和  $Eavg_e$  是来自所示帧的这些能量值的平均数。最小值用于每个训练表达中，以更新正在运行的噪声平均值。该噪声平均值利用下面的方程式被反复地更新：

$$X_{ref}(k) = ((k-2) * X_{ref}(k-2) + (X_{nz1} + X_{nz2})) / k$$

其中  $X_{ref}(k)$  为第  $k$  个噪声特征的参考值， $X_{nz1}$  表示从第一个训练表达的  $Savg_e$  和  $Eavg_e$  最小值中发现的噪声特征， $X_{nz2}$  表示从第二个训练表达的  $Savg_e$  和  $Eavg_e$  最小值中发现的噪声特征。

被更新的噪声平均值和用作噪声平均值更新的训练表达的数量被记录在存储器 110 中，如步骤 412 所示。

如果在步骤 406 中确定该装置处于免提模式，如步骤 408 所示，一个免提标记 HF 在步骤 408 中被设置。如果该训练处于免提模式，被设置的标记 HF 表示当前处于免提模式，而不是更新噪声模式。

假定训练环境相对地安静。这可以通过信号质量检验来实现，该信号质量检验要求所有的训练表达至少具有 18dB 的信噪比。可以采用检验来保证用户在  $Savg_e$  和  $Eavg_e$  测量时间期间没有说话。

由处理器 108 所执行的识别 206 的普通操作将参照图 6 作一般地描述。首先，计算用于测试表达的噪声特征，该测试表达是系统正试图识别的输入表达，如步骤 602 中所示。在识别模式中，背景噪声测量是从相同表达窗口的开始 160ms  $Savg_e$  和最后 160ms  $Eavg_e$  中作出。在识别期间的噪声测量为  $X_{recog}$ ，等于  $Savg_e$  和  $Eavg_e$  的平均值。将该值与训练模式中计算出来的参考噪声值相比。该比较用于发现识别背景噪声估计与训练背景噪声估计的比率。本领域的技术人员将认识到，也可以利用这些值的其他相对比较。

处理器 108 接着在步骤 606 中计算字补偿。该比率用于计算字入

口补偿。该字入口补偿控制词表范围外的拒绝的精度。一般来说，较高的噪声环境具有较低的精度值。使用一个查找表来计算字入口补偿，这个字入口补偿带有作为存储器表地址的噪声索引比以及作为输出的补偿值。可以利用如图 8 所示的有利的十个补偿分布，有意义的是在识别模式中（比率 6-9）中，较多噪声的环境在本质上具有比代表识别模式的比率更小的补偿，所述的代表识别模式的比率更接近于训练模式的噪声参考值（比率 0-4）。例如，这条曲线可以从如下公式得到：

$$x = X_{\text{ref}}(k) / X_{\text{recog}}$$

$$f(x) = 1 / (1 + 2^{1.5(x-5)})$$

超出范围的索引标志比将缺省为零的最小字入口补偿。例如，所应用的实际补偿可以为  $-220 * f(x)$ ，尽管实际的标量可以是任意值，这个值导致一个补偿值，和与其结合的比数有一个所希望的比例。

非线性关系的使用通过在噪声环境好时提供一个大的补偿，在噪声环境坏时提供一个小的补偿，对词表范围内的和词表范围外的识别提供一个重要的改进。本领域的技术人员将认识到，字入口补偿的计算可以直接得到，而不是通过利用查阅表得到

如步骤 608 所示，该识别继续其主搜索及并行的无用信息模式。该识别系统的目标是发现图 3 中从节点  $N_1$  至节点  $N_2$  的最可能的路径。节点  $N_1$  和节点  $N_2$  通过路径  $A_1-A_N$  耦合，路径  $A_1-A_N$  表示对于  $N$  个字的词表的隐含马尔柯夫模式，可选择地包括一个无用信息模式  $A_{GM}$ 。附加的  $A_1^{\text{noise}}$  及  $A_2^{\text{noise}}$  表示噪声模式，并且与节点  $N_1$  和节点  $N_2$  有关。在输入表达中，无用信息模式试图捕获任何非词表范围内的声音或字。它是一个值为零的状态模式，只用于词表范围外的拒绝算法所利用的状态零值模式。为了防止它构造比噪声模式更好的噪声，对于归为噪声类的各个帧，一个补偿被施加在无用信息模式概率上。

如图 3 所示，通过基础网络的搜索利用诸如维特比算法的两级定位算法来实现。这种搜索的最低级发现输入表达的各个帧与给定弧的状态之间的最佳定位及路径比数。这种技术的一个例子是用于将表达

的各个帧施加在单一模式的各个状态上，该技术在卷号为 CS10103 的同时待审的专利申请中公开，该申请的题目是“METHOD OF TRACEBACK MATRIX STORAGE IN SPEECH RECOGNITION SYSTEM”，与本发明同一天以 Jeffrey Arthur Menuier 等的名义申请，并且在卷号为 CS10104 同时待审的专利申请中公开，该申请的题目是“METHOD SELECTIVELY ASSIGNING A PENALTY TO A PROBABILITY ASSOCIATED WITH A VOICE RECOGNITION SYSTEM”，与本发明同一天以 Daniel Poppert 的名义申请。这些公开在此作为参考。较低级定位算法产生一个比数用作通过给定 HMM 弧的输入表达的最佳路径。

除较低级定位算法之外，其中每个弧的比数，或 HMM，通过累积概率 ( $C_i^n(m)$ ) 被跟踪，该累积概率为在帧  $m$  中弧  $A_n$  的状态  $i$  的累积概率，节点  $N_1$  及  $N_2$  也必须跟踪它们自己的累积概率。节点累积概率  $C_j(m)$  为在帧  $m$  中节点  $N_j$  的累积概率。该概率的计算非常象每个 HMM 的累积概率，在累积概率中保持节点的最高比数。累积概率可以用以下公式计算：

$$C_j(m+1) = \text{Max}_{n \in A_j} \{ C_{In}^n(m) + P_{0In}(d_{In}) \}$$

其中  $A_j$  为终止在节点  $J$  的弧的集合  $\{A_1, A_2, \dots, A_N\}$ 。 $In$  为弧  $n$  中的状态数， $D_{In}$  为弧  $n$  的最后状态的持续时间， $P_{0In}(d_{In})$  为对弧  $n$  的最后状态的超出状态之外的转移补偿。该累积概率为终止在节点  $N_j$  的所有弧中最后状态的累积概率  $C_{In}^n(m)$  与其超出状态之外的概率之和的最大值。

当跟踪节点的累积概率时，必须修改每个弧初始状态的累积概率的计算  $c_1^n(m)$ ，以便于从节点  $N_j$  转入其初始状态。有一个一次转移补偿分配给从节点  $N_j$  向弧  $A_n$  的初始状态转移，被称为字入口补偿。它并不适用于噪声模式或无用信息模式，因此当被启动时它作为词表范围外拒绝的精度控制。累积概率可以被看成

$$C_i^n(m+1) = o_i^n(f_m) + \max(C_j(m) + W(n), C_1^n(m) + P_{s1}(d_1))$$

$$\text{其中 } W(n) = \begin{cases} g(x), & \text{如果 } n \in \{A_1, A_2, A_3\} \\ 0, & \text{如果 } n \in \{A_1^{\text{noise}}, A_2^{\text{noise}}, A_3^{\text{noise}}\} \end{cases}$$

其中  $W(n)$  为字入口补偿,  $A_{gm}$  为无用信息弧,  $A_1^{\text{noise}}$  为节点 1 的噪声弧,  $o_i^n(f_m)$  为弧  $n$  在状态  $i$  中特征矢量的观测概率,  $P_{s1}(d_1)$  为弧  $n$  状态 1 的相同状态转移补偿。这个方程式保持着相同节点到从起始节点的转移的最大值与观测概率之和。在识别处理末端保留下来的信息是横穿到达节点 2 的弧。这是通过与累积概率  $C_i^n(m)$  及  $C_j^n(m)$  一起的传播路径信息实现的。

对于有效表达, 通过定位算法的字模式的最佳路径必须利用一个比字入口补偿更大的值来产生一个比无用信息模式更好的比数, 否则有效表达将被错误地拒绝。对于无效表达, 无用信息模式必须比通过每个合格字模式的路径更大, 以便该表达被正确地拒绝。

识别算法利用所收集的特征矢量的整个窗口, 例如该窗口一般可以为 2 秒的数据。此外, 它利用每个帧的一个语音/噪声分类位, 更新图 3 中的  $A_1^{\text{noise}}$  及  $A_2^{\text{noise}}$  使用的单个状态噪声模式。

在识别模式中, 处理器 108 通过设置噪声更新标记为 1 及帧计数器为 0 来对识别进行初始化, 如步骤 702 中所示。在步骤 704 中帧计数器被加 1。接着在步骤 706 中处理器确定噪声标记是否被设置。如果没有, 处理器继续进行判决 716。如果该标记被设置, 在步骤 708 中处理器 108 确定噪声模式是否仍然被启动, 如果没有, 在步骤 714 中噪声更新标记被设置为 0。在一定数目的更新实现以后, 噪声建模被关闭。

如果仍然应执行噪声更新, 在步骤 710 中处理器确定是否更新噪声模式。如果处理器将更新该帧的噪声模式, 在步骤 712 中该模式被更新。噪声模式  $A_1^{\text{noise}}$  及  $A_2^{\text{noise}}$  由该系统通过利用语音/噪声分类位动态地计算出来, 该语音/噪声分类位由特征抽取算法所送出。是否更新当前帧的噪声模式的判决的细节通过查阅语音分类来作出, 该语音分类由特征抽取算法得出。一旦预先确定数目的连续语音帧被看作表达, 就不再进行更新。例如, 该限制可以为 3 帧。如果该帧的语音对

噪声的分类来说，指示该帧为噪声帧，噪声模式仅对特定帧更新。

接着，在步骤 716 中处理器确定帧计数器是否小于帧门限数目。直到特定数目的帧已经被处理时，概率估计才会开始。这使得在基于噪声模式的概率被计算出来之前，该噪声模式变得一定程度精确。如果没有收到帧门限数目，则处理器返回步骤 704，其中帧计数器加 1。

如果帧计数器超过该门限，处理器 108 在步骤 718 中计算该帧的节点及弧的累积概率。概率比数在步骤 710 中被归一化。归一化可以通过从所有其他的累积概率中减去最大的累积概率而得到。累积归一化因数也被跟踪，以便非归一化比数可以在识别过程结束时返回。

接着，在步骤 722 中处理器确定最后一帧是否被处理。如果没有，处理器返回步骤 704，并且递增帧计数器。否则，该识别结果与归一化比数一起被输出，如步骤 724 所示。

这个噪声模式为单个状态模式。该状态的矢量平均是一个  $m$  的函数  $\mu_1^{\text{noise}}(m)$ ，因为它被动态地计算，并且在第  $m+1$  帧用新的特征矢量  $f_{m+1}$  更新，如下所示：

$$\mu_1^{\text{noise}}(m) = ((M_{\text{noise}}(m) * \mu_1) + f_{m+1}) / (M_{\text{noise}}(m) + 1)$$

其中， $M_{\text{noise}}(m)$  为已经被用于计算  $\mu_1^{\text{noise}}$  的噪声帧的数目，其可以取不同于  $m$  的值，因为不是所有的帧都被用于噪声更新。此外，该更新方程式仅被用作噪声模式的对数倒频谱元素。 $\delta$  对数倒频谱及  $\delta$  能量噪声元素固定为 0。

因此，可以看出本发明公开了一种改进的系统，该系统在训练及识别期间提供了依赖于背景噪声电平的各种拒绝精度。该系统有助于防止无效表达与存储的语音模式的有关性，并且有助于提高有效表达准确探测。

尽管本发明已经以上述说明书及附图公开如上，但应理本说明书仅作为例子，本领域的技术人员在不背离本发明的精神及范围的前提下，可以作出各种的变化和修改。尽管本发明在诸如在蜂窝无线电话的手提式无线设备中发现特定的应用，但是本发明可以适用于任何采用语音识别的装置中，这些装置包括寻呼机、电子管理器、计算机以

---

及电话装置。本发明的保护范围应以权利要求所界定的范围为准。

图 1

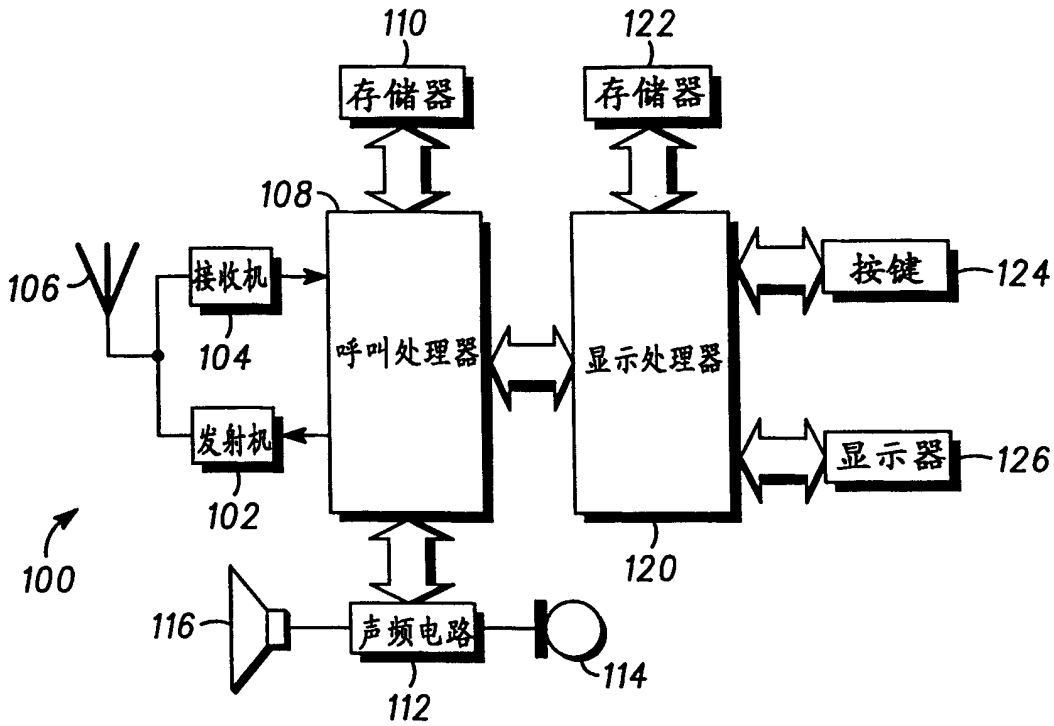


图 2

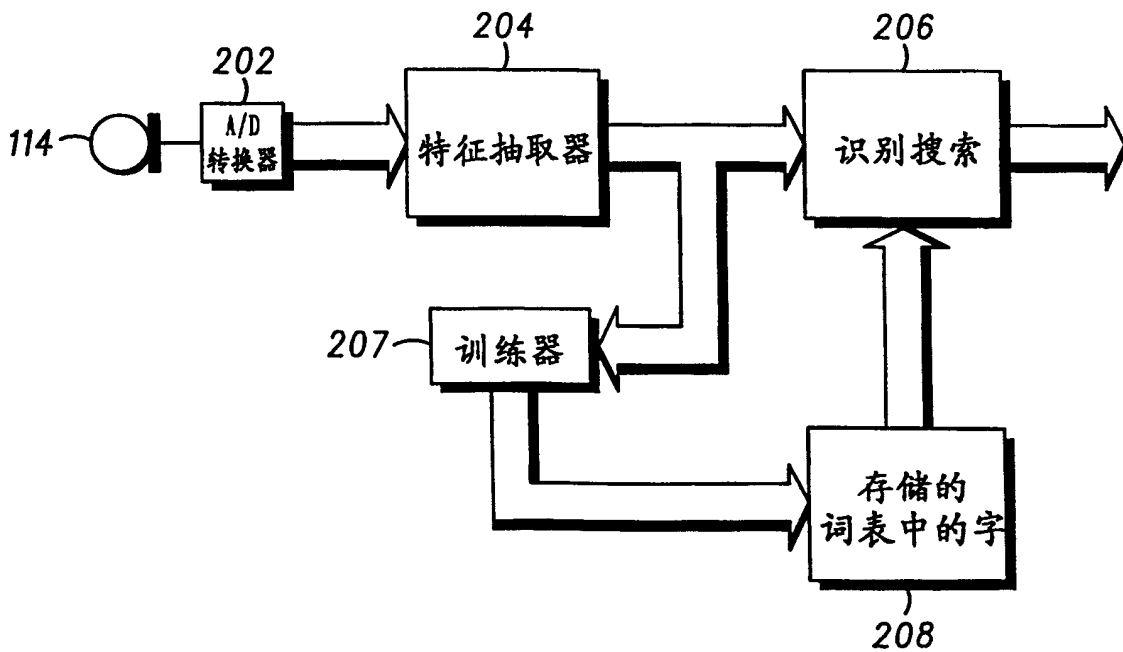


图 3

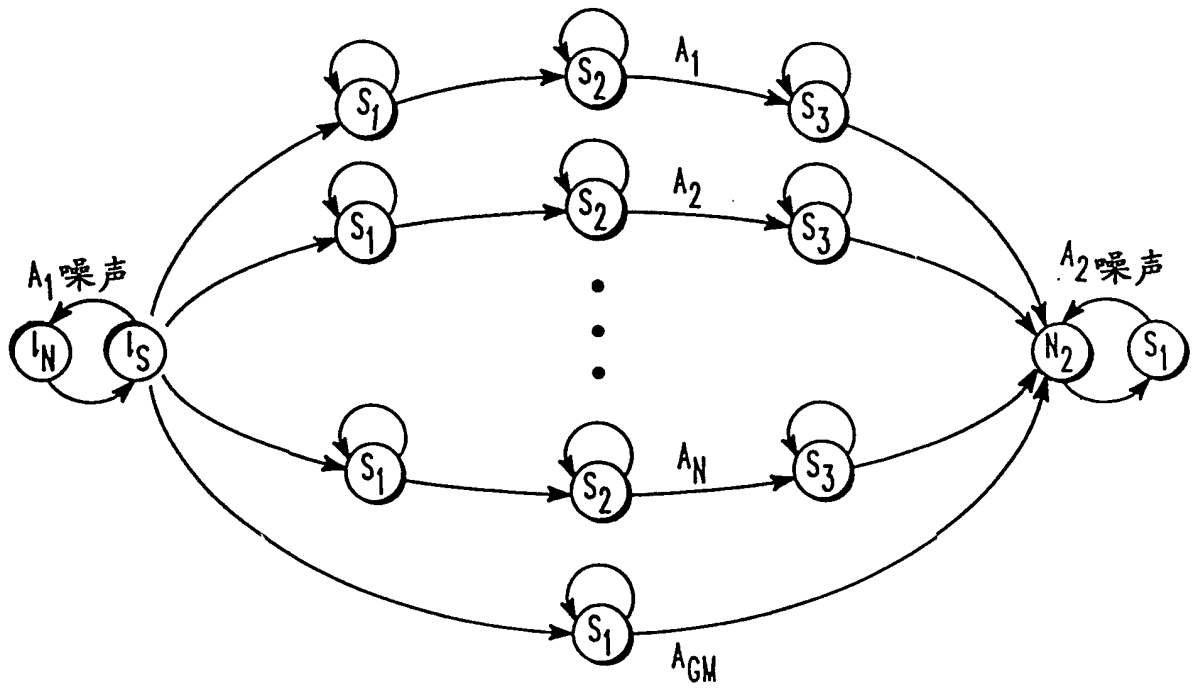


图 8

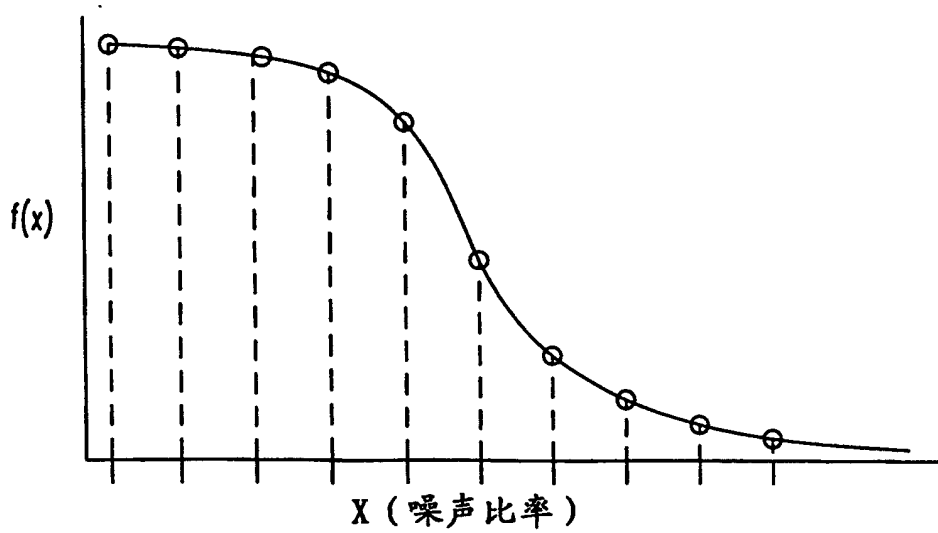
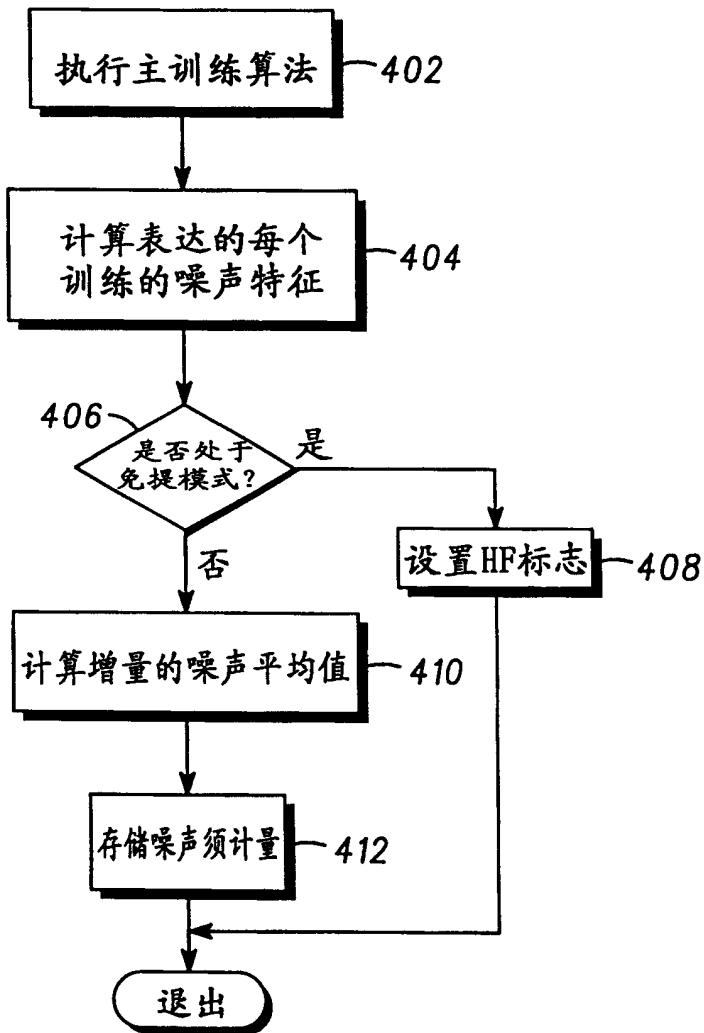


图 4



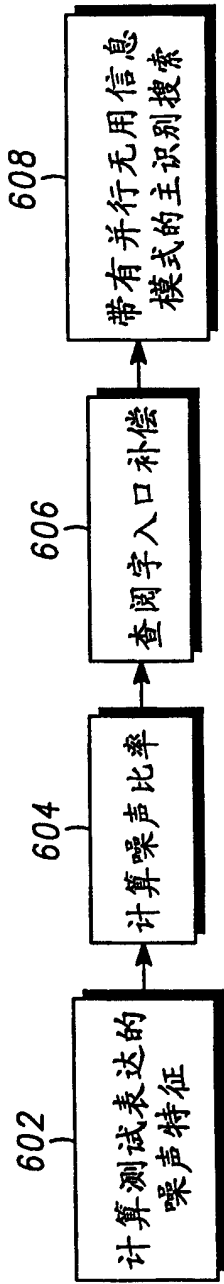


图 6

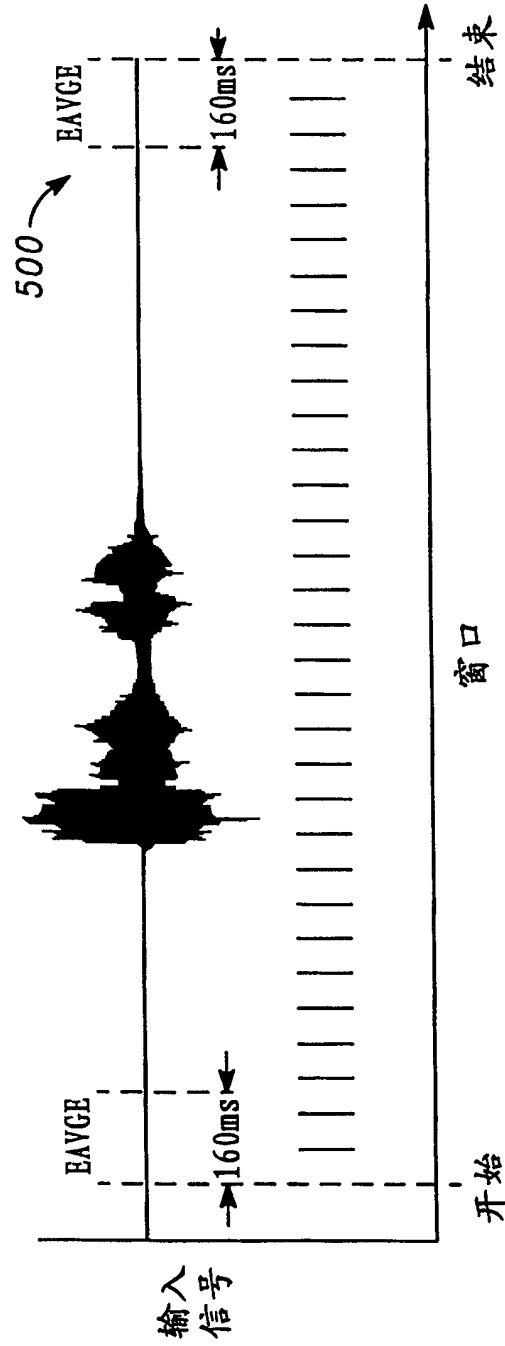


图 5

图 7

