



[12] 发明专利说明书

[21] ZL 专利号 98809296.4

[45] 授权公告日 2003 年 10 月 22 日

[11] 授权公告号 CN 1125433C

[22] 申请日 1998.9.7 [21] 申请号 98809296.4

[30] 优先权

[32] 1997.9.18 [33] DE [31] 19741211.4

[86] 国际申请 PCT/DE98/02633 1998.9.7

[87] 国际公布 WO99/14741 德 1999.3.25

[85] 进入国家阶段日期 2000.3.20

[71] 专利权人 西门子公司

地址 德国慕尼黑

[72] 发明人 J·容卡维特施 H·赫格

审查员 刘红梅

[74] 专利代理机构 中国专利代理(香港)有限公司

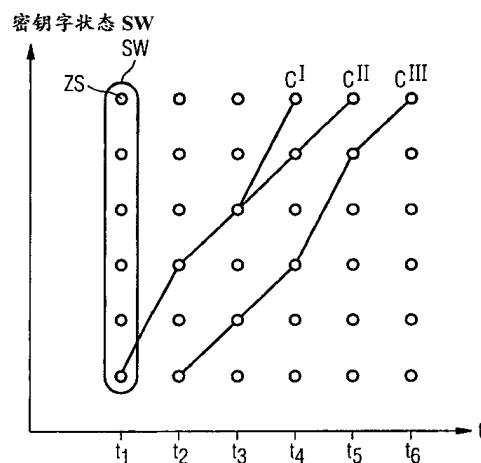
代理人 郑立柱 张志醒

权利要求书 2 页 说明书 12 页 附图 2 页

[54] 发明名称 在交谈语言中识别密钥字的方法

[57] 摘要

在每个扫描时刻假设一个密钥字的开端，这样就以交谈语言识别了该密钥字。然后尝试在一组描述密钥字的 HMM 状态上映射该密钥字。在表象空间内借助于 Viterbi 算法确定最好的路径，一个局部置信度代替一个在 Viterbi 算法中使用的附属概率使用。如果对于最好的 Viterbi 路径来说由局部置信度结合成的一个完整置信度低于一个下面的界限，则识别密钥字，该路径作为假设密钥字的开端的扫描时刻得到证实。



1. 在交谈语言中识别密钥字的方法,

a) 在该方法中, 通过一组隐式-Markov 模型的状态 W 描述密钥字,

5 f) 在该方法中, 以一个预先确定的速率扫描交谈语言, 并且在每个扫描时刻 t 从交谈语言中产生对于一个属于扫描时刻 t 的语言信号的特征矢量 O_t ,

g) 在该方法中, 借助于 Viterbi 算法在一组状态上映射特征矢量 O_t 的组 O , 其中在一个状态中根据发送程度计算局部置信度,

10 h) 在该方法中 Viterbi 算法提供一个完整的置信度 C ,

i) 在该方法中, 如果得出

$$C(W, O) < T$$

其中, 各符号表示:

$C(O)$ 置信度,

15 W 密钥字, 描述为一组状态,

O 一组特征矢量 O_t ,

T 一个预先确定的阈值,

则在交谈语言中识别密钥字,

j) 在该方法中, 否则不以交谈语言识别密钥字。

20 2. 按照权利要求 1 的方法, 其中, 发送程度是发送概率的一个负对数。

3. 按照权利要求 1 或 2 的方法, 其中, 在每个扫描时刻 t 在状态 W 的组的第一状态中开始一个新的路径。

4. 按照权利要求 1 或 2 的方法, 其中, 在每个扫描时刻 t ,

25 Viterbi 算法提供一个完整的置信度。

5. 按照权利要求 1 或 2 的方法, 其中, 通过

$$C = -\log P(W|O)$$

确定置信度 C , 并且通过

$$c = -\log \frac{P(O_t | s_j) \cdot P(s_j)}{P(O_t)},$$

确定附属的局部置信度, 其中, P 表示可能性, t 表示时间, j 表

示数量编号， s_j 表示一组状态中的一个状态。

6. 按照权利要求1或2的方法，其中，通过

$$C = -\log \frac{P(O|W)}{P(O|\bar{W})},$$

确定置信度 C，并且通过

$$c = -\log \frac{P(O_t|s_j)}{P(O_t|\bar{s}_j)},$$

确定附属的局部置信度，其中

- 5 P 表示可能性，
- T 表示时间，
- J 表示数量编号，
- W 表示非密钥字。
- s_j 表示非状态 s_j 。

10 7. 按照权利要求1或2的方法，其中，确定用于一个预先给出的持续时间的完整的置信度，并且从完整置信度的最小值中推断出密钥字的开始时刻。

8. 按照权利要求7的方法，其中，该最小值低于一个预先确定的阈值。

15 9. 按照权利要求1的方法，用于识别多个密钥字，其中只要低于多个预先确定的阈值，就以较好的置信度识别密钥字。

10. 按照权利要求9的方法，其中，对于这个时期，在该时期内要识别的密钥字包含在交谈语言中，不识别另外的密钥字。

在交谈语言中识别密钥字的方法

本发明涉及在交谈语言中识别密钥字的方法。

5 在交谈语言中识别密钥字时，迄今为止始终必需模拟完整交谈的表现。专业人员基本上已知了二种方法：

从[1]中公开了一个识别密钥字的方法，该方法使用了具有大量词汇的语言识别器。对此试图完全识别所交谈的语言。紧接着根据可能存在的密钥字检查已识别的字。由于大量词汇并由于在模拟突发语言
10 的表现时的问题和噪音，该方法也不能明确一个附加于语言信号部分的字，该方法是费劲的并且有错误的。

为了模拟不属于密钥字的词汇的表现部分（所谓的 OOV 部分，OOV= 超出词汇），一个另外的方法使用了特殊的填充模型（也就是：填充符-无用结构）。如此的语言识别器在[2]中说明，并且包含密钥字以及一个填充结构或多个填充结构。该方法提供了一组填充字定义符和密钥字定义符。对此，拟定或者训练一个适合的填充结构是困难的，这更好地突出模拟的密钥字，也表明关于密钥字结构的高的鉴别。
15

此外从[3]或[4]中公开了隐式-Markov 结构(HMMs)，从[3]或[4]中公开了，借助于 Viterbi 算法确定的一个最好路径。

20 隐式-Markov 结构(HMMs)服务于个别随机过程（也就是所谓的 Markov 过程）的描述。在语言识别的范围内隐式-Markov 结构(HMMs)此外用于建立一个字典，在该字典中记录由各子单元构成的字结构。
通过

$$\lambda = (A, B, \pi) \quad (0-1)$$

25 正式描述隐式-Markov 结构，其中该等式具有一个二次状态转移矩阵 A，其包含状态转移概率 A_{ij} ：

$$A = \{A_{ij}\}, \text{ 其中, } i, j = 1 \dots N \quad (0-2)$$

并且具有一个发送矩阵 B，其包括发送概率 B_{ik} ：

$$B = \{B_{ik}\}, \text{ 其中 } i = 1 \dots N, k = 1 \dots M \quad (0-3)$$

30 一个 N 维矢量 π 用于初始化，对于时刻 $t=1$ ，N 个状态的出现概率确定：

$$\pi = \{\pi_i\} = P(s(1) = s_i) \quad (0-4)$$

对此，概率一般以

$$P(s(t)=q_i) \quad (0-5)$$

表明，Markov 链

$$\underline{s} = \{s(1), s(2), s(3), \dots, s(t) \dots\} \quad (0-6)$$

5 在时刻 t 处于状态 q_i 。对此 Markov 链 \underline{s} 具有一个取值范围

$$s(t) \in \{s_1, s_2, \dots, s_N\} \quad (0-7)$$

其中，该取值范围包含一个有限数量的 N 个状态。这样的状态称为 q_i ，在该状态中 Markov 过程处于时刻 t 。

在状态 s_i 中从一个确定的标记 σ_k 的出现得出发送概率 B_{ik}

$$10 \quad B_{ik} = P(\sigma_k | q_i = s_i) \quad (0-8)$$

其中，根据

$$\Sigma = \{\sigma_1, \sigma_2, \dots, \sigma_M\} \quad (0-9)$$

数值 M 的字符组包 Σ 含确定的标记 σ_k ($k=1 \dots M$)。

15 隐式-Markov 结构的每个状态可能有预先确定数量的随机状态：本身、最近状态、再下一个状态、等等，这样就从隐式-Markov 结构中得出状态空间。具有所有可能转变的状态空间称作格子结构。在 1 阶的隐式-Markov 结构中大于一个时间间距的经历是不重要的。

Viterbi 算法基于这种思想。即如果处于状态空间内的最佳路径上，则该路径一直是整体最佳路径的组成部分。由于 1 阶的隐式-Markov 结构，仅仅考虑了一个状态的最好的原型 (Vorgänger)，因为差的原型事先已经包含一个差的求值。这也意味着，对于每个时间间隔确定路径的所有可能的连续并选出最好的连续，这样就可以递推地、从第一时刻开始、逐个时间间隔地寻找最佳路径。

25 在二个在 [1] 和 [2] 中说明的方法中，各一个 OOV 部分的模拟是必须的。在第一种情况 [1] 中，表现的字必须明确地存在于识别器的词汇中，在第二种情况 [2] 中通过特殊的填充结构描述所有的字和 OOV 噪声。

本发明的任务在于，说明一个方法，该方法能够在交谈语言中识别密钥字，因此避免上述的缺点。

30 1. 根据如下的技术特征解决了这个任务，即在交谈语言中识别密钥字的方法，

a) 在该方法中，通过一组隐式-Markov 模型的状态 W 描述密钥

字，

b) 在该方法中，以一个预先确定的速率扫描交谈语言，并且在每个扫描时刻 t 从交谈语言中产生对于一个属于扫描时刻 t 的语言信号的特征矢量 o_t ，

5 c) 在该方法中，借助于 Viterbi 算法在一组状态上映射特征矢量 o_t 的组 O ，其中在一个状态中根据发送程度计算局部置信度，

d) 在该方法中 Viterbi 算法提供一个完整的置信度 C ，

e) 在该方法中，如果得出

$$C(W, O) < T$$

10 其中，各符号表示：

$C(O)$ 置信度，

W 密钥字，描述为一组状态，

O 一组特征矢量 o_t ，

T 一个预先确定的阈值，

15 则在交谈语言中识别密钥字，

在该方法中，否则不以交谈语言识别密钥字。

根据本发明描述了一个以交谈语言识别密钥字的方法，其中，通过一组状态 W 描述密钥字。以预先确定的速率扫描交谈语言，并且在每个扫描时刻 t 从交谈语言中产生一个属于扫描时刻 t 的语言信号的特征矢量 o_t 。借助于一个 Viterbi 算法在一组状态 W 上映射特征矢量 o_t 的一组 O ，其中，在一个状态中一个局部置信度代替发送程度，主要是发送概率的负对数。Viterbi 算法提供了一个完整的置信度 C （也就是：置信度 C ）。如果得出：

$$C(W, O) < T \quad (1)$$

25 其中，各符号表示：

$C(O)$ 置信度，

W 密钥字，描述为一组状态，

O 一组特征矢量 o_t ，

T 一个预先确定的阈值

30 则以交谈语言识别密钥字。

否则，不以交谈语言识别密钥字。

本发明的优点在于，不必全部模拟表现，就可以在交谈语言内部识

别密钥字。因此得出在执行时明显地降低了费用，并且因此也得出一个高效(较快的)的方法。通过(完整的)置信度 C 作为一个基本的译码原理使用，在译码过程内声音的模拟局限于密钥字。

- 改进在于，在每个扫描时刻 t，通过隐式-Markov 结构的状态空间
- 5 在一组状态 W 的一个第一状态中开始一个新的路径。因此，在每个扫描时刻假设，在交谈语言中包含密钥字的开端。根据置信度在密钥字的通过隐式-Markov 结构体现的状态上映射由后面的扫描时刻产生的特征矢量。在映射的结尾，也就是路径结尾，得出完整的置信度，根据该置信度逻辑判断作出，是否密钥字的假定开端是确实是如此的。
- 10 如果是，则识别密钥字，否则不识别。

在本发明的改进的范围内，通过

$$C = -\log P(w|o) \quad (2)$$

确定完整的置信度 C，并且通过

$$c = -\log \frac{P(o_t | s_j) \cdot P(s_j)}{P(o_t)}, \quad (3),$$

确定附属的局部置信度 c，其中各符号表示：

- 15 s_j 一组状态的一个状态，
 $P(w|o)$ 在一组特征矢量 o_t 的条件下密钥字的一个概率，
 $P(o_t | s_j)$ 发送概率，
 $P(s_j)$ 状态 s_j 的概率，
 $P(o_t)$ 特征矢量 o_t 的概率。

- 20 合适的完整置信度是通过这个性质表征的，即说明关于一个可靠度的情况，以该可靠性证明密钥字。在负的对数范围内完整置信度 C 的较小值表达一个高度的可靠性。

在一个附加改进的范围内，通过

$$C = -\log \frac{P(o|w)}{P(o|\bar{w})} \quad (4)$$

确定置信度 C，并且通过

$$c = -\log \frac{P(O_t | s_j)}{P(O_t | \bar{s}_j)} \quad (5)$$

确定附属的局部置信度，其中各符号表示：

$P(O_t | W)$ 在不出现密钥字 W 的条件下特征矢量 O_t 组的概率，

s_j 状态 s_j 的反情况（也就是：非状态 s_j ）。

5 所描述的置信度的优点此外在于，其是可以计算的，也就是前面的训练和/或评估是不必的。

从完整的置信度的定义中能够分别推导出局部置信度的定义。在与这个密钥字的表现同步的时刻，局部密钥字进行密钥字的置信度的计算。

10 通过关系式

$$P(O_t) = \sum_k P(O_t | s_k) \cdot P(s_k) \quad (6)$$

和

$$P(O_t | \bar{s}_j) = \sum_{k \neq j} P(O_t | s_k) \cdot P(s_k) \quad (7)$$

能够计算局部置信度。

此外，可以通过合适的近似方法确定 $P(O_t)$ 或者 $P(O_t | s_j)$ 。对于如此的近似方法的一个实例是在每个时刻 t 对 n 个最好的发送 $-\log$

15 $P(O_t | s_j)$ 取平均值。

通常借助于 Viterbi 算法实施译码过程：

$$C_{t, s_j} = \min_k (C_{t-1, s_k} + c_{t, s_j} + a_{kj}),$$

其中，各符号表示：

C_{t, s_j} 在状态 s_j 中在时刻 t 的完整的累加置信度，

C_{t-1, s_k} 在状态 s_k 中在时刻 $t-1$ 的完整的累加置信度，

20 c_{t, s_j} 在状态 s_j 中在时刻 t 的局部置信度，

a_{kj} 在状态 s_j 状态 s_k 的转移附加量 (Strafe)。

因为对于一个密钥字的完整置信度的描述来说，除时间限度外不需

要局部置信度，所以在对密钥字的寻找中放弃 OOV 部分的声音模拟。

通过使用 Viterbi 算法可能在每个时刻 t、在密钥字的第一状态中开始一个新的路径，其中首先密钥字划分为隐式-Markov 结构 (HMMs) 的各个状态，对于密钥字来说优化完整的置信度，并且同时确定最佳的开始时刻 (Viterbi 算法的后台跟踪)。

此外，对于预先确定的时期来说在阈值 T 之下寻找一个最小值是按照目的的。因此避免了，在这个预先确定的时期内多次识别密钥字。

如果得出密钥字，这些密钥字鉴于其通过各自的状态组体现的说明形式彼此是类似的，则使用一个机理是有利的，该程序在识别密钥字时排除了在识别密钥字的时期内有时包含在交谈的语言信号内的另外的密钥字。

下面给出本发明的改进，发送程度是发送概率的一个负对数。

在每个扫描时刻 t 在状态 W 的组的第一状态中开始一个新的路径。

在每个扫描时刻 t，Viterbi 算法提供一个完整的置信度。

通过

$$C = -\log P(W|O)$$

确定置信度 C，并且通过

$$c = -\log \frac{P(O_t | s_j) \cdot P(s_j)}{P(O_t)},$$

确定附属的局部置信度，其中，P 表示可能性，t 表示时间，j 表

示数量编号，s_j 表示一组状态中的一个状态。

通过

$$C = -\log \frac{P(O|W)}{P(O|\bar{W})},$$

确定置信度 C，并且通过

$$c = -\log \frac{P(O_t | s_j)}{P(O_t | \bar{s}_j)},$$

确定附属的局部置信度，其中

- P 表示可能性,
 T 表示时间,
 J 表示数量编号,
 W 表示非密钥字。
 5 S_j 表示非状态 S_j.

确定用于一个预先给出的持续时间的完整的置信度，并且从完整置信度的最小值中推断出密钥字的开始时刻。

该最小值低于一个预先确定的阈值。

10 用于识别多个密钥字，其中只要低于多个预先确定的阈值，就以较好的置信度识别密钥字。

对于这个时期，在该时期内要识别的密钥字包含在交谈语言中，不识别另外的密钥字。

根据下面的附图详细描述本发明的实施例。

图示

15 图 1 以交谈语言识别密钥字的方法的方框图，

图 2 阐明确定置信度的示意图，

图 3 图 3 的示意图，其描述了在一个预先确定的持续时期中假设置信度的曲线。

在图 1 中描述了以连续的语言识别密钥字的方法的方框图。

20 在步骤 101 中通过一组状态 W 描述密钥字。对此首先以各三个状态替换音位 HMMs (参见 [3])。在下一步骤 102 中扫描连续的语言，并且在每个时刻 t 从连续的语言信号中为一个属于扫描时刻 t 的语言信号建立一个特征矢量 O_t。对此，特征矢量 O_t 包含预先确定数量的特征，这些特征表明在扫描时刻 t 的语言信号的特征，也就是组成部分。

25 在步骤 103 中在一组状态 W 上映射一组特征矢量，这些特征对于不同的扫描时刻 t 是从语言信号中获得的。对此映射规则说明 Viterbi 算法 (参见 [3])。用一个局部的置信度代替在 Viterbi 算法中使用的发送概率 $-\log P(O_t | s_j)$ 。在步骤 104 中 Viterbi 算法在每个时刻提供一个完整的置信度 C，其对于一组状态 W 中所找到的状态来说累积包含各个局部置信度。如果得出

$$C(W, O) < T \quad (1),$$

则在步骤 105 中以连续语言识别密钥字，其中各符号表示：

C () 完整置信度,
 W 密钥字, 描述为一组状态,
 O 一组特征矢量 O_t ,
 T 一个预先确定的阈值。

5 否则不以连续语言识别密钥字。

下面说明对于一个完整置信度的二个可能的实现和各一个从属的局部置信度。另外置信度是可以想象的。

第一置信度

从密钥字的一个后验概率的负对数中第一完整置信度定义为一个
 10 置信度:

$$C_1 = -\log P(W|O) \quad (2).$$

下面联系以下假定运用 Bayes 的准则:

$$P(O) = \prod_t P(O_t) \quad (8),$$

$$P(W) = \prod_t P(s_{\psi(t)}) \quad (9),$$

$$P(O|W) = \prod_t \left[P(O_t | s_{\psi(t)}) \cdot a_{\psi(t-1), \psi(t)} \right] \quad (10).$$

一组特征矢量 $P(O)$ 的概率对此表达为各个特征矢量 $P(O_t)$ 的概率的乘数。通过一个 HMMs 的全部各个选择的状态的各个概率 $P(s_{\psi(t)})$ 相乘的方法, 以相同的方式计算整个字的 $P(W)$ 的概率。其中函数 $\psi(t)$ 是在密钥字的状态上的特征矢量(也就是时间)的映射。有条件的概率 $P(O|W)$ 与 HMMs 的通常的概率一致, 其可以借助于发送概率 $P(O_t | s_{\psi(t)})$ 和跃迁概率 $a_{\psi(t-1), \psi(t)}$ 进行计算。因此得出完整置信度 C_1 :

$$C_1 = \sum_t -\log \left(\frac{P(O_t | s_{\psi(t)}) \cdot P(s_{\psi(t)})}{P(O_t)} \cdot a_{\psi(t-1), \psi(t)} \right) \quad (11).$$

如果考虑 Viterbi 算法的工作方式, 则推荐局部置信度 $c_1(O_t | s_t)$
 20 的定义, 该置信度在 Viterbi 算法的寻找过程的内部使用:

$$c_1(O_t | s_j) = -\log \frac{P(O_t | s_j) \cdot P(s_j)}{P(O_t)} \quad (12).$$

考虑到 HMMs 的所有状态，就可以计算在等式(12)的分母中出现的特征矢量的概率的概率：

$$P(O_t) = \sum_k P(O_t | s_k) \cdot P(s_k) \quad (13)$$

(参见等式(6))。

在前面的训练中确定这些状态的后验概率 $P(s_k)$ 。因此局部置信度
5 $c_1(O_t | s_j)$ 是完全可以计算的。

第二置信度

第二置信度的定义包括一次在状态 W 的表征密钥字的组的条件下特征矢量 O_t 的一组 O 的有条件的概率比例，和另一次在对此相反结构 \bar{W} 的条件下的比例。对此得出：

$$C_2 = -\log \frac{P(O | W)}{P(O | \bar{W})} \quad (4).$$

10 对此 W 仅仅表明一种实际不存在的结构，可是可以计算其发送概率。与第一完整的置信度的定义不同，这个定义导致一个对称的完整置信度，如果满足

$$P(O | W) = P(O | \bar{W}) \quad (14)$$

则其在 O 时具有一个对称中心。类似于定义完整置信度的情况，通过代入等式(8)、(9)和(10)在考虑各个相反的结构 $a_{\psi(t-1), \psi(t)}$ 和 $s_{\psi(t)}$ 的情况下得出如下等式：

$$C_2 = \sum_t -\log \frac{P(O_t | s_{\psi(t)}) \cdot a_{\psi(t-1), \psi(t)}}{P(O_t | \bar{s}_{\psi(t)}) \cdot \bar{a}_{\psi(t-1), \psi(t)}} \quad (15).$$

一个适合的、在由 Viterbi 算法实施的寻找中可以使用的局部置信度 $c_2(O_t)$ 被定义为：

$$c_2(O_t | s_j) = -\log \frac{P(O_t | s_j)}{P(O_t | \bar{s}_j)} \quad (16).$$

在这种情况下也可以计算局部置信度 $c_2(0_i | s_j)$ ，因为，除了对于 $P(0_i | s_j)$ 之外可以自身计算所有加权的发送概率：

$$P(0_t | \bar{s}_j) = \sum_{k \neq j} P(0_t | s_k) \cdot P(s_k) \quad (7)$$

这样就可以计算分母，（也参见等式(7)）。

因此对置信度进行了二重定义，在低值的情况下（在这个置信度 C_2 是一个负值的情况下），一个高的可靠性表明，已经正确识别密钥字。

作为这个可计算的置信度的优点说明，即不必训练附加的 HMMs，也不必高明地处理另外相关的参数。在使用普通的音位 HMMs 的情况下可以计算置信度。

正如上面示出的，置信度的定义可以与一个基于隐式-Markov 结构的 Viterbi 寻找结合。HMMs 的全部的各个状态 s_j 不确定概率 $P(0_i | s_j)$ 的负对数，而是确定一个局部置信度 c_1 或 c_2 代替该对数。

在图 2 中描述了一个示意图，其阐明一个置信度的确定。

在图 2 的上图中，在横坐标上描述了离散的时刻 t_1, t_2, \dots ，并且在纵坐标上通过一组状态 ZS 表征的密钥字 SW。在图 2 中示出了对时间轴 t 的一个连续的语言信号。

连续的语言信号可以包含多个的，也可以包含不同的密钥字，其中在一个时刻主要仅仅包含一个密钥字。

在离散时刻扫描连续的语言信号，并且以特征矢量 0_i 存储在各个扫描时刻存在的信息。根据本发明，以此为出发点，即密钥字开始于全部的这些扫描时刻。也就是在每个扫描时刻 t_1, t_2 或 t_3 每次开始一个电位的密钥字，以 Viterbi 算法的曲线可以复合其路径。为了简化，以一个密钥字为出发点，其中多个密钥字每次都需要每个要识别的密钥字的方法。

如果该密钥字也在时刻 t_1 开始，则根据从连续语言中获得的特征矢量 0_i 进行紧跟在时刻 t_1 的特征矢量的映射。分别确定关于累加置信度的最好的路径 PF。对于每个时刻 t 得出一个置信度 C 。置信度的值产生关于此的情况，即密钥字是否包含在连续的语言中，或不，并且在时刻 t 结束。

在图 2 中示范地画入了路径，其开始于时刻 t_1, t_2 和 t_3 ，并且在时刻 t_4, t_5 和 t_6 导致完整的置信度 C^I, C^{II} 和 C^{III} 。属于 C^I 和 C^{II} 的完整置

信度对应于在 t_1 内的可能的密钥字开端，而完整置信度 C^{III} 最好通过在 t_2 内开始的路径实现。

在此可以看出，在每个时刻 t 发现一个完整的置信度 C ，其中通过使用 Viterbi 算法确定一个附属的开始时刻。

5 如果连续的语言包含一些完全另外的东西作为密钥字，则置信度相应地变差，不能进行识别。根据 Viterbi 算法的工作方式，确定的完整置信度的不同路径的长度是不同的，因此预示，由四个状态的局部置信度形成完整置信度 C^I ，而完整置信度 C^{II} 和 C^{III} 包括五个状态的局部置信度。因此相应密钥字的持续时间为 $4\Delta t$ 和 $5\Delta t$ 。

10 图 3 阐明了这种关系。从图 2 中确定的完整置信度 C^I 、 C^{II} 和 C^{III} 示范地在图 3 中绘在纵坐标上。横坐标再表征时间 t 。

对于每个时刻 t 分别得出一个适当的完整置信度 C 。

首先确定完整置信度 C 的最小值 MIN ，并且因此以此为出发点，密钥字以该最小值 MIN 存在于连续的语言中。

15 就此而言这是重要的，因为在时刻 t_1 已经低于完整置信度 C 的阈值，也就是识别密钥字。可是，正如在图 3 中示范描述的，鉴于可变的动态匹配（确定位置置信度的不同持续时间），在紧接着到来的时刻 t_{t+1} “较好地” 识别密钥字，为了断定什么时候最佳地识别密钥字，在附属时刻 t_{MIN} 确定 MIN 。从这个时刻 t_{MIN} 开始，借助于后台跟踪（参见 [3]）
20 确定在连续语言信号中的开始时刻。也确定在连续语言信号中交谈密钥字的开端。

在此可以看出，对于每个密钥字可以实施一个如此的最小值确定，对于一个密钥字的持续时间不能识别另外的密钥字。如果同时进行连续语言叠加的多个密钥字，则密钥字首先是正确的密钥字，与另外密
25 钥字相比其置信度表现较高的可靠性。

在文献的范围内引用了如下出版物：

[1] M. Weintraub: “使用 SRI 译码的大词汇量语言识别系统的键盘点”， Proc. IEEE ICASSP, 1993, 第 2 册, 463-466 页。

30 [2] H. Boulard, B. D'hoore 和 J.-M. Boite: “在字点系统中的优化识别和拒绞性能”， Proc. IEEE ICASSP, 1994, 第 1 册, 373-376 页。

[3] L. R. Rabiner, B. H. Juang: “隐式 markov 模型的介绍”，

IEEE ASSP 杂志, 1986, 4-16 页。

[4] A. Hauenstein: “用于自动语言识别的算法优化和处理器的设计”, 在慕尼黑技术大学的关于集成电路的博士论文, 19. 07. 1993, 13-35 页。

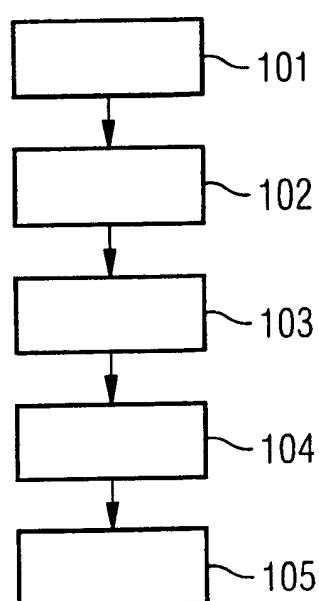
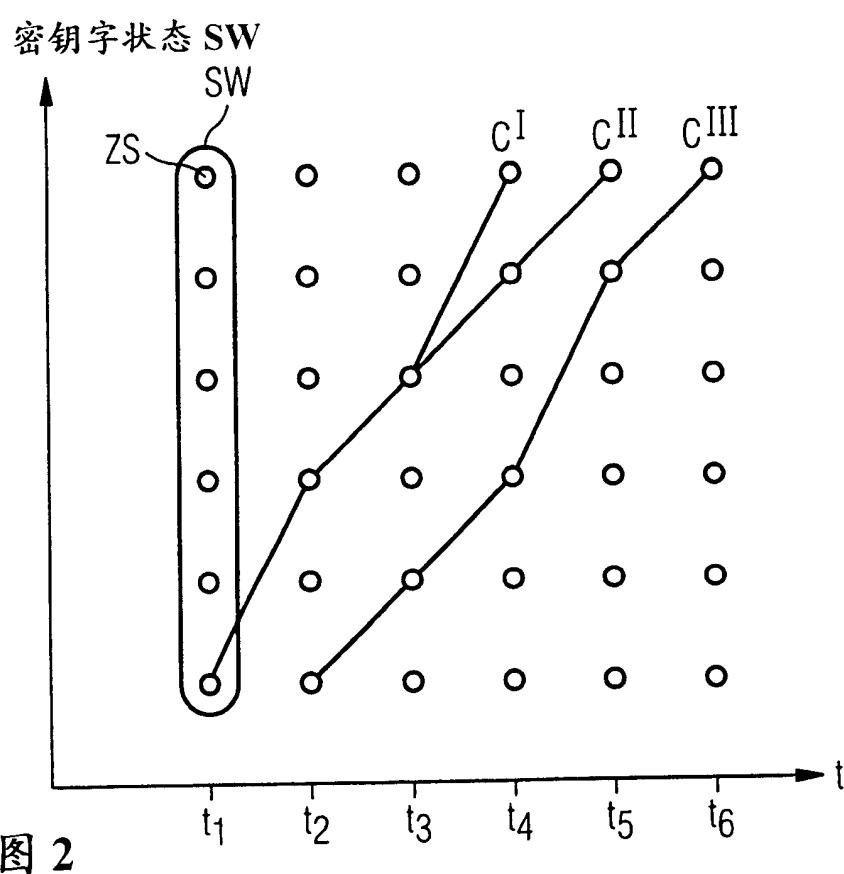


图 1



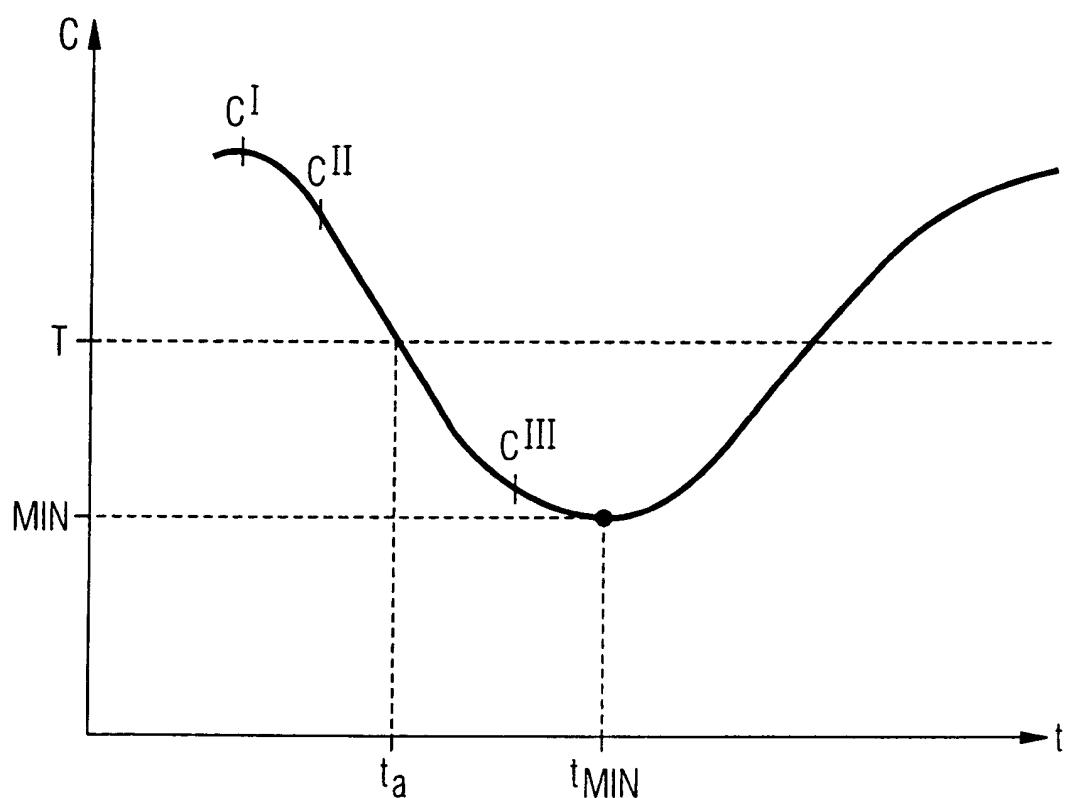


图 3