

[19] 中华人民共和国国家知识产权局

[51] Int. Cl.

G10L 19/12 (2006.01)

G10L 11/04 (2006.01)



[12] 发明专利说明书

专利号 ZL 02811726.3

[45] 授权公告日 2006 年 7 月 5 日

[11] 授权公告号 CN 1262993C

[22] 申请日 2002.6.7 [21] 申请号 02811726.3

[30] 优先权

[32] 2001.6.11 [33] US [31] 09/878,762

[86] 国际申请 PCT/IB2002/002078 2002.6.7

[87] 国际公布 WO2002/101718 英 2002.12.19

[85] 进入国家阶段日期 2003.12.11

[71] 专利权人 诺基亚有限公司

地址 芬兰埃斯波

[72] 发明人 A·海基宁 V·罗皮拉

S·皮蒂莱

审查员 刘亚斌

[74] 专利代理机构 中国专利代理(香港)有限公司

代理人 陈景峻 陈 霁

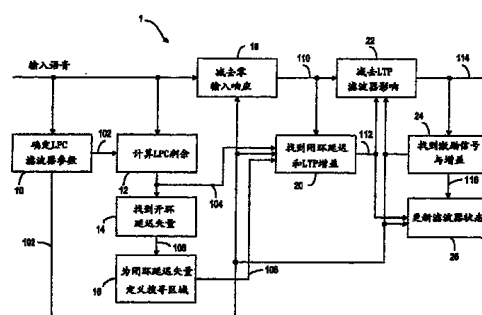
权利要求书 3 页 说明书 11 页 附图 8 页

[54] 发明名称

用于编码语音信号中连续基音周期的方法和装置

[57] 摘要

一种对语音信号的连续基音周期编码的方法和装置。基于连续语音周期统计特性的先有知识，设计一种定形的点阵结构，以覆盖基音空间中的最大概率的阵点。考虑所有维，密码本索引搜索开始于发现基音空间中的开环估计，基于定形的点阵结构在每一维中分别地在闭环搜索中精确化开环估计。对第一子帧的闭环搜索是为获得绝对基音周期或者增量周期，而对其他每一子帧的闭环搜索是为获得各自子帧的增量基音。



1. 一种对多个信号帧中的声音信号进行编码的方法，每个信号帧具有表示各个信号帧中的声音信号的基音周期，其中每个信号帧包括多个信号段，每个信号段代表基音空间中的一维，并且每个信号段中的声音信号由基音值表征，其中基音值能通过点阵结构中声音信号的阵点分布图形特性表示，从而定义基音空间中的密码本索引，所述的方法的特征在于：

根据阵点分布图形来定形点阵结构，

- 根据定形的点阵结构，提供表示基音空间的每一维中的基音值的密码本索引，以助于声音信号编码，以及

对定形点阵结构执行开环搜索以获得表示基音值的开环估计。

2. 根据权利要求1的方法，特征还在于考虑基音空间的全部维来执行上述开环搜索，上述方法的特征还在于通过对定形的点阵结构中的闭环搜索，分别对基音空间中每一维的开环估计进行精确化，以获得表示各个信号段中基音值的闭环搜索值。

3. 根据权利要求2的方法，特征在于基音值表示差分基音周期。

4. 根据权利要求2的方法，特征在于所述多个信号段包含第一信号段和其余的信号段，其中第一信号段中的基音值表示绝对基音周期，在其余的每个信号段中的基音值表示差分基音周期。

5. 根据权利要求4的方法，特征在于所述多个信号段是连续的信号段，其顺序地包括第一信号段和三个第二信号段，其中第一信号段中的基音值表示绝对基音周期，并且在每个第二信号段中的基音值表示差分基音周期。

6. 根据权利要求2的方法，特征在于将信号段设置在子帧中。

7. 根据权利要求6的方法，特征在于每个信号帧包括四个子帧，并且其中四个子帧的每一个中的基音值表示差分基音周期。

8. 根据权利要求6的方法，特征在于每个信号段顺序地包括第一子帧和三个第二子帧，其中第一子帧中的基音值是绝对基音周期，每个第二子帧中的基音值是差分基音周期。

9. 根据权利要求1的方法，特征在于阵点分布图形包括定形的点阵结构中的多个区域，并且每个区域可由超立方体表示，每个超立方体具有多条包括定形的点阵结构的多个阵点的边，并且其中密码本索

引表示超立方体的边上的阵点的数目。

10. 根据权利要求 1 的方法, 特征在于被提供给解码装置的密码本索引具有表示定形的点阵结构的信息, 从而允许该解码装置根据定形的点阵结构从密码本索引中合成语音信号。

5 11. 根据权利要求 1 的方法, 特征在于声音信号包括语音信号。

12. 一种用于编码多个信号帧中的声音信号的设备, 每个信号帧具有表示各个信号帧中的声音信号的基音周期, 其中每个信号帧包括多个信号段, 每个信号段代表基音空间中的一维, 并且每个信号段中的声音信号由基音值表征, 其中基音值能通过点阵结构中声音信号的
10 阵点分布图形特性表示, 从而定义基音空间中的密码本索引, 根据用于定义定形点阵结构的阵点分布图形来定形所述点阵结构, 所述设备的特征在于:

响应于声音信号、考虑基音空间的全部维通过对定形的点阵结构的开环搜索以获得基音周期的开环估计, 从而提供表示开环估计的开
15 环搜索值的装置, 以及

响应于所述开环搜索值、通过对定形的点阵结构的闭环搜索分别对基音空间中每一维的开环估计进行精确化, 从而获得表示各个信号段中基音值的闭环搜索值的装置。

13. 根据权利要求 12 的设备, 特征在于基音值表示差分基音周期。

20 14. 根据权利要求 12 的设备, 特征在于在至少一个信号段中的基音值表示绝对基音周期, 在其余的每个信号段中的基音值表示差分基音周期。

15. 根据权利要求 12 的设备, 特征在于将信号段设置在连续子帧中。

25 16. 根据权利要求 15 的设备, 特征在于连续子帧顺序地包括第一子帧和三个第二子帧, 其中第一子帧中的基音值表示绝对基音周期, 每个第二子帧中的基音值表示差分基音周期。

17. 根据权利要求 15 的设备, 特征在于每个信号帧包括四个子帧, 并且其中四个子帧的每一个中的基音值表示差分基音周期。

30 18. 根据权利要求 12 的设备, 特征在于阵点分布图形包括定形的点阵结构中的多个区域, 并且每个区域可由超立方体表示, 每个超立方体具有多条包括定形的点阵结构的多个阵点的边, 并且其中密码本

索引表示超立方体的边上的阵点的数目。

19. 一种对多个信号帧中的声音信号进行处理的系统，每一信号帧具有表示各个信号帧中的声音信号的基音周期，其中每个信号帧包括多个信号段，每个信号段表示基音空间中的一维，并且每个信号段中的声音信号由基音值表征，其中基音值能通过点阵结构中声音信号的阵点分布图形特性表示，从而定义了基音空间中的密码本索引，点阵结构根据阵点分布图形而定形，从而定义定形点阵结构，所述系统的特征在于：

编码器，具有：

- 10 响应于声音信号、考虑基音空间的全部维通过对定形的点阵结构的开环搜索以获得基音周期的开环估计，从而提供表示开环估计的开环搜索值的装置，以及

- 响应于所述开环搜索值、通过对定形的点阵结构的闭环搜索分别对基音空间中每一维的开环估计进行精确化，从而获得表示各个信号段中基音值的闭环搜索值进而提供表示定形点阵结构和密码本索引的信息的装置，以及
- 15

解码器，其响应于该信息，根据定形的点阵结构从密码本索引中合成进一步的聲音信号。

用于编码语音信号中连续基音周期的方法和装置

发明领域

5 本发明总体上涉及语音编码领域，特别涉及连续基音周期的量化。

发明背景

10 基于人类语音处理机制，浊音语音的基音周期轨迹在时间上进展缓慢。通过对连续基音周期之间的差值进行编码而将这个现象应用在许多当前的语音编码器中，从而提高了编码效率。在一种基于子帧工作的典型的编码器中，比如码激励线性预测（CELP）编码器中，每帧至少发射一次绝对基音周期。

连续基音周期之间的差值一般称为一个增量周期。在现有技术中，增量周期可以从有限的范围内获得均匀分布值，为它们的编码提供便利。这能被解释为一个由定义了帧上的增量周期的点均匀排列得到的多维矩形点阵。相应地，通过使用均匀量化器来完成对增量周期的编码。更确切地说，用相似的量化器独立地对几个连续的增量周期进行编码。使用这种方法的编码器也被认为是一种多维矩形点阵量化器。在一个多维点阵量化器中，每一维代表相应子帧中的一个基音周期。通常，点阵的第一维表示第一子帧中的绝对基音周期，而其余维中的每一个代表当前和在前子帧的基音周期之间的差值。因此，在一个语音编码方案中，为了进行语音处理，将语音帧分成四个子帧，将连续基音周期量化中所用的编码器称为四维点阵量化器，第一维中的绝对基音周期和在其余三维中的增量周期用四维基音空间中的点（ p, d_1, d_2, d_3 ）表示。在本发明中，对包含仅适于增量周期（ $d_1, d_2, d_3, \dots, d_n$ ）的维的点阵结构要给予特别的注意。

在大多数利用差分编码的现有技术语音编码器中，将 n 个增量周期的点阵结构记述为在 n 维基音空间中规则排列的一组阵点，从而使这些点在整个基音空间均匀地间隔开。除了在基音空间中这些点的均匀间隔，现有技术语音编码器的关键特征是点阵的阵点在二维平面上投影的矩形形状。点阵的结构通常是恒定的，而不必考虑前面的段中的基音周期。一种典型的增量周期的二维点阵的实例在附图 1 中示出，

其中点阵 L 通过下式定义

$$L = \{(d_1, d_2) \mid d_{1\min} \leq d_1 \leq d_{1\max} \wedge d_{2\min} \leq d_2 \leq d_{2\max}\} \quad (1)$$

点阵包含了 d_1 和 d_2 在各自的最小和最大值之间的所有可能的组合。虽然在图 1 中示出的点阵是二维的，但是根据二维情况能够很容易地派生更高维的点阵。一般而言，第 j 维的最小和最大可能增量周期分别由 $d_{j\min}$ 和 $d_{j\max}$ 表示。

一旦定义了点阵量化器的形状和区域，一个重要的参数就是点阵的密度，这是由于密度决定编码器的比特率。比特率是密度的单调递增函数。因此，点阵量化器的密度反映用于基音周期信息的精确度。通常地，使用分数值代替整数以提高合成语音的质量。

在一种典型的用于增量周期的点阵量化器中，当点阵的矩形形状保持恒定时，通常注意点阵的边界值 ($d_{j\min}$, $d_{j\max}$)。而不注意选择一组合适的点阵点以覆盖包含最大源概率的基音空间区域。

众所周知在语音信号中，基音是个有意义的参数，由于人类语音处理机制的性质，基音的发展很顺利。大体上，浊音语音的基音周期轨迹在时间上进展缓慢，并且轨迹的突然改变是非常不可能发生的。已经发现矩形点阵结构远不是最理想的关于覆盖基音空间区域的点阵点的选择。此外，在现有技术中，每一维中差分基音值的搜索是独立地进行的。矩形点阵和搜索方法的使用还没有被优化以反映出人类语音的已知特性。

利用基音空间的源概率以改进合成语音的质量，从而为语音编码器中连续基音周期的量化提供一种改进的方法和系统是有利并且十分需要的。

发明概述

本发明的主要目的是提高对连续基音周期编码的效率，从而改进利用差分编码对连续基音周期之间的差值进行编码的语音编码器中的合成语音的质量。基于现有技术中的浊音语音中连续增量周期的特性，可以通过定义一个优化的或者更有效率的点阵结构来达到这个目的，其点阵结构定形为覆盖了有最大概率的点位于其内的基音空间区域。此外，能够将具有表示基音周期的不同时间分辨率的不同点密度的区域，定义在优化的点阵结构之中。利用这样一种优化的点阵结构，就可以提供一种为优化的点阵结构中的阵点分配索引以及在密码本中

搜索索引的新方法。

因此，根据本发明的第一个方面，一种对多个信号帧中的声音信号编码的方法，多个信号帧中的每一个具有一个表示各个信号帧中声音信号的基音周期，其中每个信号帧包括多个信号段，每个信号段代表在基音空间中的一维，并且每个信号段中的声音信号的特征在于具有一个基音值，其中基音值能通过点阵结构中声音信号的阵点分布图形特性表示，从而定义了基音空间中密码本索引，所述的方法的特征在于

根据阵点分布图形定形点阵结构，以及
10 对应于定形的点阵结构，提供表示基音空间的每一维中基音值的密码本索引，以便促进声音信号的编码。

根据本发明的第一个方面，该方法的特征还在于

考虑到基音空间的全部维，通过对定形的点阵结构的开环搜索获得基音周期的开环估计，以及

15 通过对定形的点阵结构中的闭环搜索，分别对基音空间中每一维的开环估计进行精确化(refine)，以获得表示各个信号段中基音值的闭环搜索值。

根据本发明，基音值表示差分基音周期或者绝对基音周期。

根据本发明，在至少一个信号段中的基音值表示绝对基音周期，
20 并且在其余的每个信号段中的基音值表示差分基音周期。

相应地，当信号段顺序地包括一个第一信号段和三个第二信号段时，第一信号段中的基音值表示一个绝对基音周期，并且在每个第二信号段中的基音值表示一个差分基音周期。

可以作为选择地，每个信号帧包括四个信号段，并且四个信号段
25 中每一个的基音值表示一个差分基音周期。

根据本发明，信号段可以设置在连续的子帧中。这样，第一子帧中的基音值可以是绝对基音周期或者是差分基音周期，并且在其余的每个子帧中的基音值是差分基音周期。

优选地，点阵结构中的每个阵点表示到基音空间中参考点的距离，并且将点阵结构定形以消除超出预定距离的点。
30

特别的，本发明的定形的点阵结构由不重叠的超立方体结合组成，定形点阵结构由增量周期范围和基音空间每一维中的时间分辨率

定义，并且其中每个超立方体能够由包括许多阵点的多个边表示。根据本发明，优化的点阵的索引表示超立方体边上的点阵点数量。

值得注意的是由编码装置提供并传送到译码装置的密码本索引具有指示定形的点阵的信息，并且其中解码装置基于定形点阵从密码本索引中合成语音信号。

根据本发明的第二个方面，对多个信号帧中的声音信号进行编码的装置，每一信号帧具有表示各自信号帧中声音信号的基音周期，其中每个信号帧包括多个信号段，每个信号段表示基音空间中的一维，并且每个信号段中的声音信号的特征在于具有一个基音值，其中基音值可通过点阵结构中声音信号的阵点分布图形特性表示，用于定义基音空间中的密码本索引，点阵结构根据阵点分布图形而定形，用于定义一个定形点阵结构，所述装置的特征在于

装置，该装置响应于声音信号，考虑基音空间的全部维，通过对定形的点阵结构的开环搜索以获得基音周期的开环估计，用于提供表示开环估计的开环搜索值，以及

装置，响应于开环搜索值，通过对定形的点阵结构的闭环搜索分别对基音空间中每一维的开环估计进行精确化，以获得表示各个信号段中基音值的闭环搜索值。

根据本发明的第三个方面，对多个信号帧中的声音信号进行编码的系统，每一信号帧具有表示各自信号帧中的声音信号的基音周期，其中每个信号帧包括多个信号段，每个信号段表示基音空间中的一维，并且每个信号段中的声音信号的特征在于具有一个基音值，其中基音值可通过点阵结构中声音信号的阵点分布图形特性表示，从而定义了基音空间中密码本索引，点阵结构根据阵点分布图形而定形从而定义了一个定形点阵结构，所述系统的特征在于

一编码器，具有：

装置，响应于声音信号，考虑到基音空间的全部维，通过对定形的点阵的开环搜索从而获得基音周期的开环估计，以提供表示开环估计的开环搜索值，以及

装置，响应于开环搜索值，通过对定形的点阵结构中的闭环搜索，分别对基音空间中每一维的开环估计进行精确化，以获得表示各个信号段中基音值的闭环搜索值，从而提供指示定形的点阵结构的信息以

及密码本索引, 以及

解码器, 具有响应于该信息根据定形的点阵结构从密码本索引中合成进一步的语音信号的装置。

通过对结合附图 2 到 6 的描述的阅读, 本发明将会变得更加清楚。

5 附图的简要说明

图 1 是矩形点阵的图形表示。

图 2 是定形的点阵结构的图形表示。

图 3a 是一个超立方体在一个二维平面上投影的图形表示。

图 3b 是该超立方体在另一个二维平面上投射的图形表示。

10 图 4a 是在一个二维平面内的点密度分布的柱状图。

图 4b 是在另一个二维平面内的点密度分布的柱状图。

图 5 是根据本发明的编码器的图形表示。

图 6 是根据本发明对语音信号编码的方法的流程图。

实施本发明的最好的方式

15 根据本发明, 建立一个定形的点阵结构的原理在图 2 中示出。通常, 基音空间中的阵点不是均匀分布的。相反地, 分布是由具有表示基音周期不同时间分辨率的不同点密度的多个区域定义的。如图 2 所示, 具有不同点密度的两个子点阵, 用 S_1 和 S_2 表示, 存在于基音空间中。这两个子点阵的并集 S , 或者 $S_1 \cup S_2$, 表示优化的点阵结构 S , 定义了定形的点阵结构。

20 如早先提到的, 众所周知, 浊音语音的基音周期轨迹在时间上进展缓慢, 轨迹上突然的变化是很不可能发生的。因此, 在一个相同的语音帧之内, 很不可能有两个很大的增量周期。比如图 1 和 2 中示出的, 角点 (d_{1min}, d_{2min}) , (d_{1max}, d_{2min}) , (d_{1min}, d_{2max}) 和 (d_{1max}, d_{2max}) 以及点阵 L 中邻近的点表示 d_1 中的增量周期和 d_2 中的增量周期都很大的情况。由于这种情况在浊音语音中不太可能发生, 因此不太可能将这些点用在密码索引搜索中。相应地, 如图 2 所示, 能够将这些点从定形点阵 S 中去掉, 而不会对产生的语音质量产生显著的影响。如图 2 所示, 当基音周期平稳进展而没有显著提高比特率时, 子点阵 S_1 中更高的点密度允许利用更高的基音分辨率。

30 由于利用基音周期差分编码的大多数现存编码器的闭环结构, 点阵中的索引搜索是基于子帧完成的。因此, 在时间上沿着点阵的一个

坐标轴连续地进行搜索。通常，这是通过首先为包含绝对基音周期和之后的增量周期的子帧确定一个单一开环基音周期估计完成的。典型地，将整数值用在开环搜索中以减少复杂性。其后，对每一维连续地以闭环方式完成索引搜索。对于第一子帧，这是在所选的开环基音周期附近完成的。对于其他子帧，搜索区域包括在前选择的基音周期的相邻区域。

根据本发明，采用优化的点阵，这种方法不可行是由于每一维中可能的阵点组通常实质上依赖于在以前的维中所选的阵点。

根据本发明所优选的方法，定形的点阵中的所估计的开环阵点是在多维空间中决定的。包括第一维的每一维中最优的索引是之后在估计的开环阵点的邻近区域中以闭环方式决定的，每次决定一个维。如图 2 所示的 p 点表示估计的开环阵点，最优的索引是从阴影区域 C 中搜索出来的。闭环搜索对属于定形的点阵 S 和以开环基音估计 p 为中心的搜索区域 C 交集的点进行检查。闭环搜索确定的索引唯一地定义了点阵覆盖的子帧上的基音周期。在图 2 中，定形的点阵 S 是点阵 L 的子集。一般而言，这不是必要的情形。

为了说明目的，定形的点阵结构定形为不重叠的超立方体 D_i 的并集，每个超立方体由增量周期范围和相应维中所用的时间分辨率确定。每个超立方体 D_i 是超立方体矩阵 D 的一行。在四维基音空间中，如果一个语音帧分为四个子帧并且每个子帧由一维表示，那么矩阵 D 的第 i 行定义了一个唯一的四维超立方体，如下：

$$D(i,:) = [p_{\min} \ p_{\max} \ r_{i0} \ d_{i1\min} \ d_{i1\max} \ r_{i1} \ d_{i2\min} \ d_{i2\max} \ r_{i2} \ d_{i3\min} \ d_{i3\max} \ r_{i3}] \quad (2)$$

其中 $p_{\min}$, $p_{\max}$ 和 r_{i0} 定义了基音周期范围和第一子帧的分辨率。最后三个子帧中增量周期的范围由 $d_{ij\min}$ 和 $d_{ij\max}$ 定义，其中 j 是子帧索引。每个子帧中相应的分辨率由 r_{ij} 表示。

采用上面描述的点阵结构，编码的过程十分简单。为了对定形的点阵中的某点的索引进行编码，获得一个起始索引和每个超立方体每一单一边中阵点的数目。编码过程从找到超立方体的索引开始，已找到的基音周期组合 (p, d_1, d_2, d_3) 属于该超立方体。包含点 (p, d_1, d_2, d_3) 的超立方体 D_i 定义为

$$D_i = \{ (p, d_1, d_2, d_3) \mid p_{\min} \leq p \leq p_{\max} \wedge d_{j\min} \leq d_j \leq d_{j\max}, j=1, 2, 3 \}$$

(3)

图 3a 图示了投影在二维平面 d_1, d_2 上的四个超立方体 D_0, D_1, D_2, D_3 。图 3b 图示了相同的超立方体投影在二维平面 d_2, d_3 上的情况。应该注意的是, 通常, 一个超立方体的阵点密度可以与另一个的阵点密度不同。为了简化, 如图 3a 和 3b 中示出的圆圈是均匀分布的。
 在图 3a 和 3b 中, 不同的超立方体示为封闭的矩形, 它们中的每个可以由其唯一的边确定。比如, 超立方体 D_2 由边 a_2, b_2 和 c_2 确定。

根据本发明, 优化或定形的点阵已经结合附图 2 到 3b 进行了描述。根据本发明, 采用优化的点阵结构, 可以定义如以下所描述的用于语音合成的一组将传送到解码器的索引。能够通过首先定义超立方体 D_i 内每维的坐标, 为超立方体中阵点的索引赋值。第 $(j+1)$ 个子帧的坐标 p_j 由以下公式给出

$$p_0 = (p - p_{i \min}) r_{i0} \quad (4)$$

$$p_j = (d_j - d_{j \min}) r_{ij}, \quad \text{其中 } j=1, 2, 3 \quad (5)$$

因此, 在定形的点阵中的阵点 (p, d_1, d_2, d_3) 的索引 s 能够根据下式赋值

$$s = s_{0i} + p_0 + p_1 n_{i0} + p_2 n_{i1} n_{i0} + p_3 n_{i2} n_{i1} n_{i0} \quad (6)$$

其中 s_{0i} 是超立方体 D_i 的偏移 (offset)。在第 $(j+1)$ 维中 D_i 的各边中的阵点的数目用 n_{ij} 表示。在以适当方式描述了点阵之后, 下一个问题就是为其找到恰当的边界值。

需要理解的是, 如上所述的定形的点阵结构仅仅是为了说明的目的。定形的点阵结构不局限于那些由超立方体组成的结构。通常, 点阵结构通过选择表示多维基音空间中的语音帧和子帧中语音信号的阵点分布图形特性的子点阵而定形。

已经在修正过的 IS-641 语音编码器中实现了根据本发明的编码方法。在修正过的 IS-641 编码器中, 以通常方法对第一维编码, 以至在第一子帧中发送一个绝对基音周期。然而, 使用包括四个超立方体的定形的点阵结构对其余三维编码。值得注意的是, 在常规 IS-641 编码器中, 只为子帧 2 和 4 发送两个增量周期。在修正了的 IS-641 编码器中, 取而代之发送三个增量周期。基于一个使用由许多谈话者讲述

的美语-英语语音的 39434 帧的试验，使用修正过的 IS-641 语音编码器从语音段中获得的增量周期分布如图 4a 和 4b 所示。为了简化，增量周期的范围限制在 ± 6 样本。第 $(i+1)$ 子帧和第 i 子帧的基音周期之差值由 d_i 表示。对图 4a 和 4b 中的增量周期四舍五入得到整数值，尽管在模拟中使用了 $1/3$ 的分辨率。在 d_1, d_2 平面中的以及在 d_2, d_3 平面中的阵点密度分布分别在图 4a 和 4b 中示出。如图 4a 和 4b 所示，两个大的增量值的组合很罕见。也就是说，当 d_1 很大时， d_2 和 d_3 就很小。但是当 d_2 或者 d_3 很大时， d_1 就很小。因此，子帧中增量周期之间存在相关性。在现有技术的编码器中，每一维都相互独立看待，而忽视子帧中增量周期之间的相关性。根据本发明，在修正了的 IS-641 中，开环基音值是帧的平均基音。在使用整数分辨率的每一维中共同地估计开环基音值。在每一维中顺序地使用闭环搜索来使该开环估计精确化。比如，在估计的开环基音值周围搜索第一子帧的闭环值。在第一子帧的四舍五入的、优化的闭环基音周围选择第二子帧的闭环值等等。第一子帧的可能整数值范围从 20-147，如图 4a 和 4b 所示，所使用的点阵结构关于轴 d_1, d_2 和 d_3 对称。因此，关于增量周期的三维点阵能够清楚地由 D_0 在轴 d_1 和 d_2 上的投影的一个角点确定。在试验中，三个不同的优化的点阵（定形点阵 S_A ，定形点阵 S_B 和定形点阵 S_C ）分别由作为偏移量 s_{Di} 使用的角点 $(2\frac{2}{3}, 1\frac{2}{3})$ ， $(2\frac{2}{3}, \frac{2}{3})$ 和 $(1\frac{2}{3}, \frac{2}{3})$ 实现。作为参考，使用了两个具有最大增量周期 $2\frac{2}{3}$ 和 $1\frac{2}{3}$ 的立方量化器（点阵 L_1 ，点阵 L_2 ）。这些范围是根据出现在图 4a 和 4b 中的分布选出的。模拟结果在表 1 中示出。结果表示为输入语音和合成语音的浊音段之间的段信噪比（SegSNR），以及每一帧中对增量周期的编码所需要的比特数。使用一个段长度为 64 的样本，并且在 SegSNR 的计算

中去除了无声段。所有模拟中使用的语音样本包括由两个男性和两个女性讲话者在安静 (clean) 条件下讲述的四个句子。样本的总长度是 782 帧。如从表 1 中可以看出的, 根据本发明, 通过使用优化的点阵结构能够提高连续基音周期的编码效率。

5 表 1

	点阵 L_1	点阵 L_2	定形点阵 S_A	定形点阵 S_B	定形点阵 S_C
SegSNR /dB	8.24	8.09	8.28	8.11	8.05
比特数	12.26	10.38	11.78	10.00	9.17

根据本发明, 语音编码器 1 如图 5 所示。它是基于公知为合成分析 (AbS) 的编码技术, 采用了线性预测编码 (LPC) 技术。典型的是, 使用了一种时变基音预测器和 LPC 滤波器的级连。如图 5 所示, LPC 分析单元 10 被用来根据输入语音信号决定 LPC 滤波器的系数 102。通常地, 语音信号在预处理步骤中经过高通滤波。经过预处理的语音信号继而窗口化 (windowed), 并且计算窗口化语音的自相关。比如使用 Levinson-Durbin 算法确定 LPC 滤波器系数 102。在大多数编码器中, 不是在每个子帧中都确定系数。这种情况下, 能够为中间的子帧插入系数。预处理步骤和 LPC 分析步骤在本领域中是公知的。输入语音进而通过反向滤波器 $A(q, s)$ 12 滤波以产生一个剩余信号 104。剩余信号 104 有时称为理想激励。根据从先前关于连续基音值的分布的知识确定的定形点阵, 使用开环搜索单元 14 为整个帧决定开环延迟估计矢量 106。通常, 矢量 106 的长度和子帧的数量相同, 具有对应于单个子帧的延迟估计的成分。使用语音信号代替 LPC 剩余信号 104 搜索估计矢量 106 也是可能的。由于全部子帧组成了多维基音空间的维, 所以对于估计矢量 106 的搜索要考虑到所有这些维。开环估计 106 为基音空间中的每一维提供开环延迟值。基于定形点阵, 搜索区域定义单元 16 用来为基音空间的每一维中的闭环延迟矢量定义闭环搜索区域 108。比如, 如图 2 中所示的, 单元 16 对属于定形点阵 S 和以开环基音估计 p 为中心的搜索区域 C 的交集的点进行检查。根据输入语音信号, 考虑 LPC 滤波器 10 的初始状态的影响, 在计算单元 18 中通过从输入语音信号中减去 LPC 滤波器 10 的零输入响应计算出为了闭环延迟

搜索的目标信号 110。闭环搜索单元 20 用来精确化开环估计 106，每次精确化一个维，基于使用该维中的定形点阵中的阵点的相应开环延迟值，从而获得密码本索引。信号 112 中包含密码本索引。特别是，闭环搜索单元 20 通过使为闭环延迟搜索的目标信号 110 和由 LPC 系数 102 和 LPC 激励信号表示的合成语音信号之间的平方和 (sum-squared) 误差最小化来搜索闭环延迟和增益。在所定义的搜索区域 108 内，在相应开环延迟值周围搜索每个子帧中的闭环延迟。由于延迟值小于子帧长度，所以必须将 LTP (长期预测器) 存储器扩展。这可以通过使用剩余信号 104，或通过复制旧的 LTP 激励而实现。LTP 存储器的扩展是现有技术中已知的。在新方案密码本搜索单元 22 中，通过从闭环延迟搜索的目标信号 112 中减去 LTP 滤波器的影响 110，计算出用于激励搜索的目标信号 114。激励信号和它的增益共同由参考数字 116 表示，通过将为激励搜索的目标信号 114 和由 LPC 系数 102 和激励信号表示的合成语音信号之间的平方和误差最小化，在计算单元 24 中搜索激励信号和它的增益。通常，使用一些试探性的规则以避免对所有可能备选的激励信号进行穷举的搜索。最后，在更新单元 26 中更新编码器 1 中的滤波器状态，以使它们与解码器中的滤波器状态保持一致。密码本搜索单元 22、计算单元 24 和更新单元 26 在现有技术中均是已知的。以上描述的编码器 1 可应用于典型的 AbS 或 CELP 编码器，比如 IS-641。

必须注意的是，当解码器接收来自编码器的语音参数时，通过基于对解码器已知的相同定形的点阵所接收到的索引和增益，确定 LTP 激励信号。

图 6 是一个流程图，图示了依照本发明的语音信号编码的方法。如图 6 中所示，当编码器在步骤 210 中接收到语音信号时，如现有技术中已知的，在语音帧和子帧中处理语音信号。在步骤 220 中，为了获得语音帧中基音周期的开环估计，考虑基音空间中的所有维，实施开环搜索。在步骤 230 中，分别对每个维实施闭环搜索，以精确化开环估计，从而获得基音值。基于对每一维的闭环搜索中所获得的基音值，在步骤 240 获得密码本索引。如步骤 250 所表示的，对每一维的闭环搜索一直持续到获得了语音帧中所有子帧的密码本索引为止。值得注意的是，基音空间的第一维 (每个语音帧的第一子帧) 中的基音

值能表示绝对基音周期或不同基音周期（增量周期）。然而，其余每个维的基音值表示各个子帧中的不同基音周期。

应该了解，结合语音信号的编码描述了本发明。然而，本发明也可以应用于非语音信号，比如音乐。

- 5 此外，虽然优选把语音帧分为多个子帧并在每个子帧中搜索闭环基音值，但搜索语音帧不同段的闭环基音值也是可能的。通常，可以在每个语音帧向解码器发送许多次不同的参数。

- 10 因此，尽管本发明已经通过其中优选的实施方案进行了描述，但本领域的技术人员将会理解在形式和细节上的前述的以及其他不同的改变、省略和偏差都可以在不背离本发明的精神和范围的情况下获得。

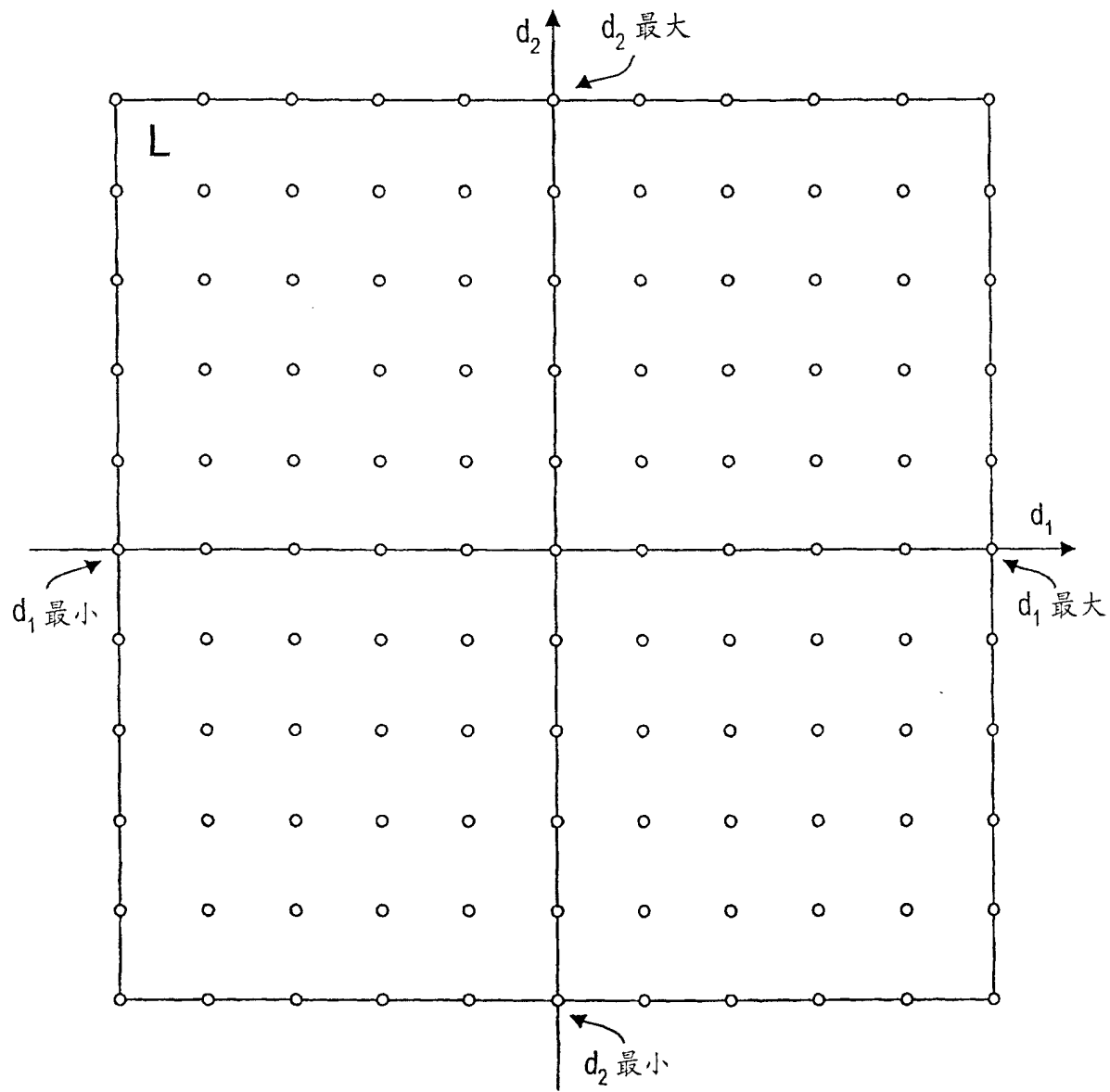


图 1

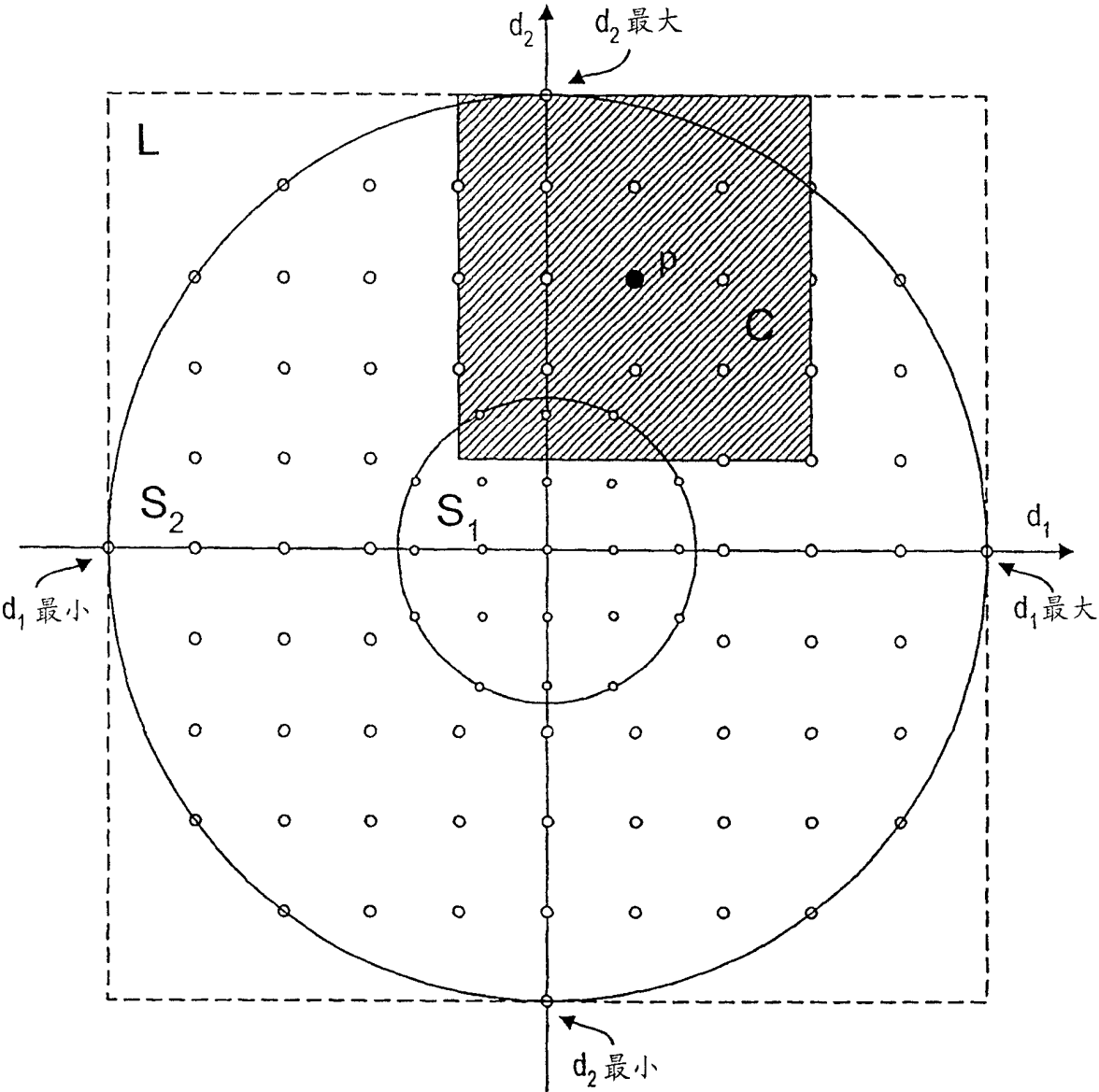


图 2

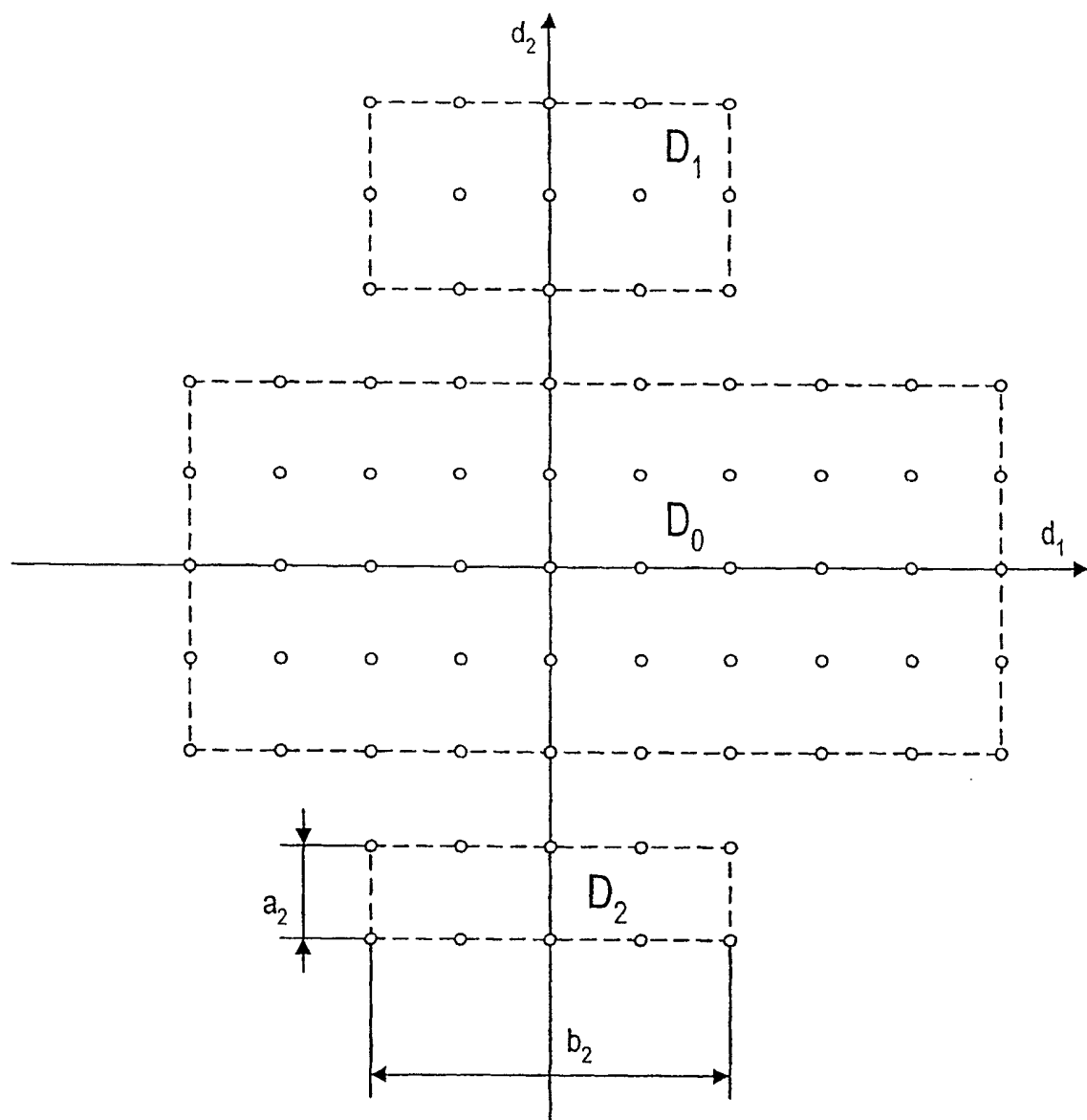


图 3a

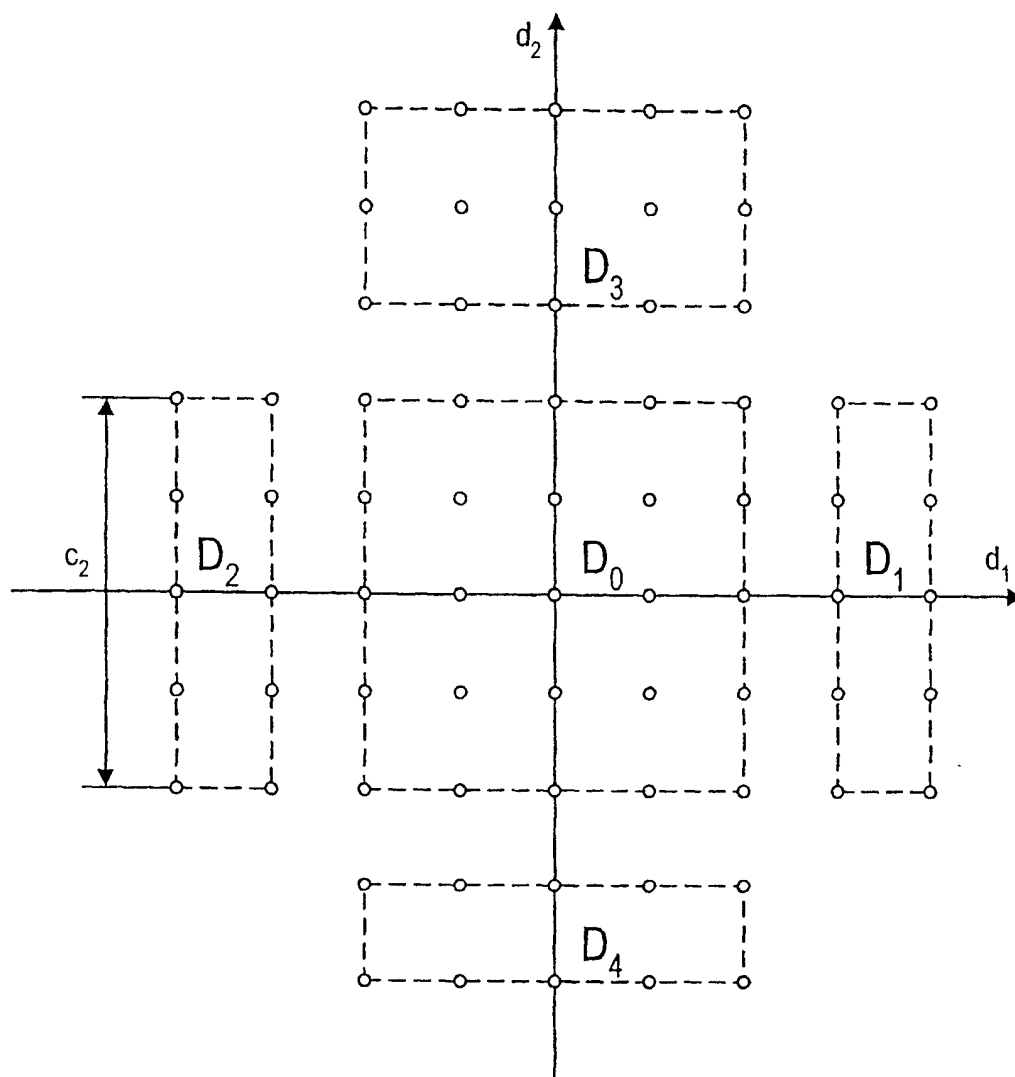


图 3b

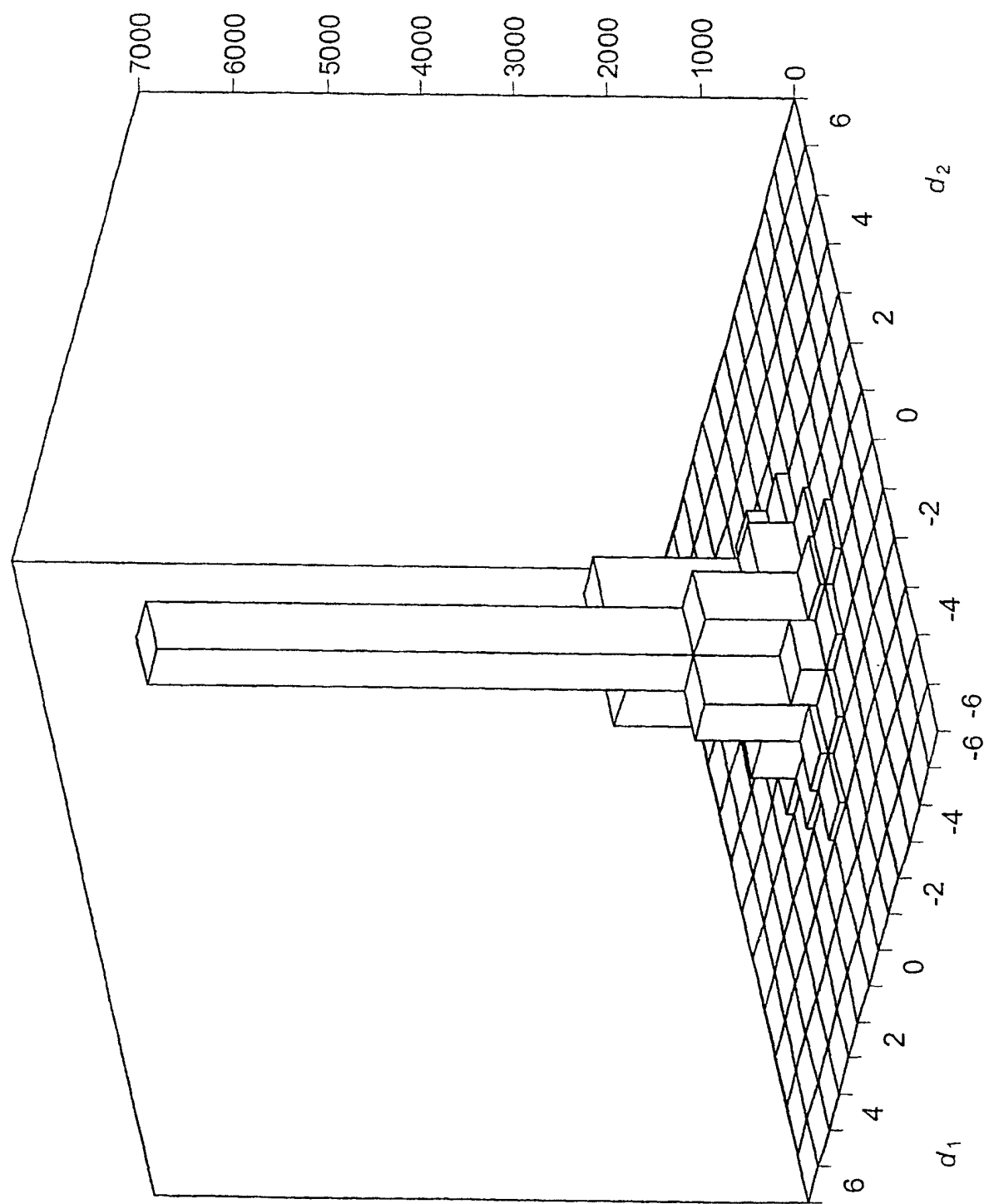


图 4a

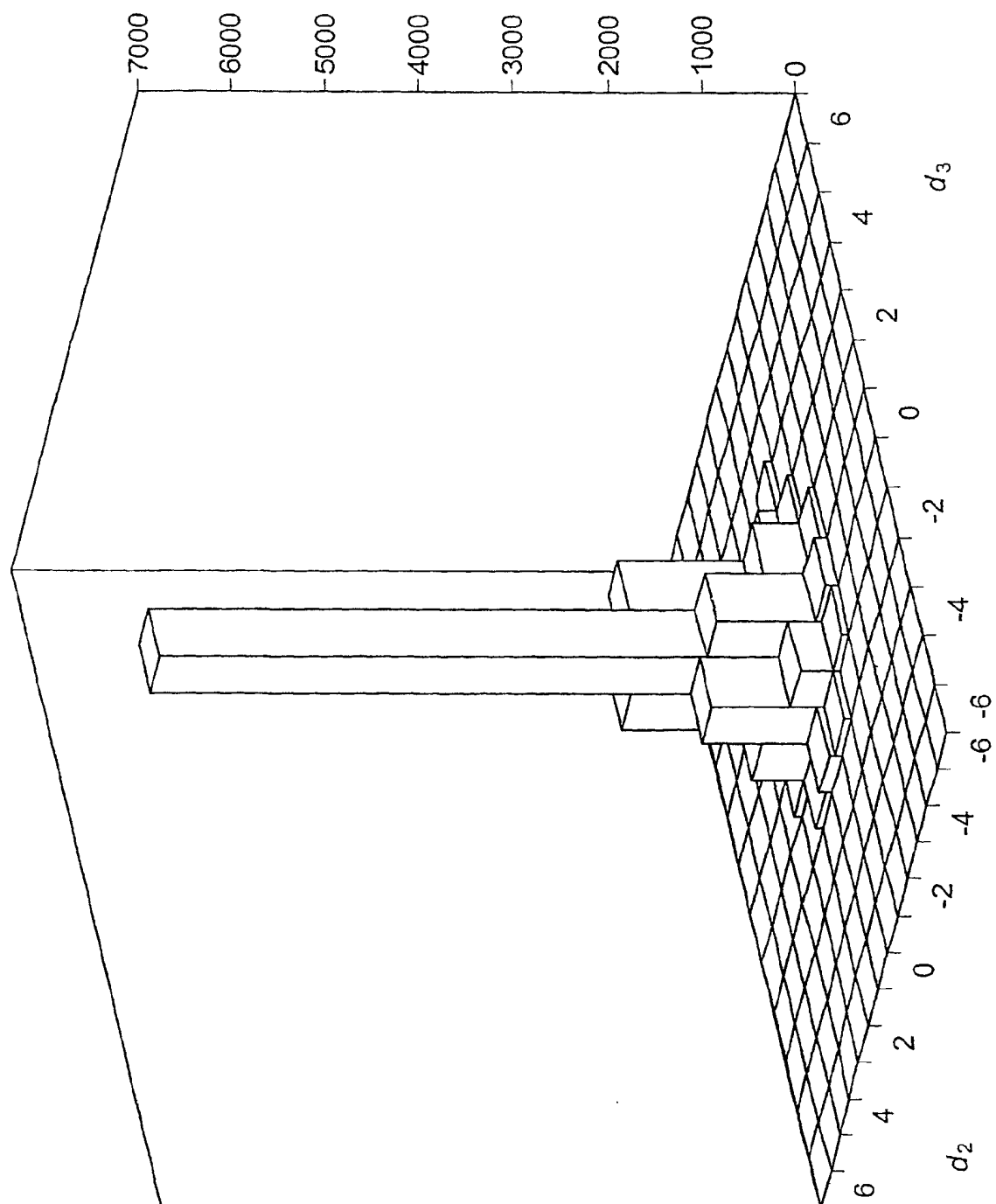


图 4b

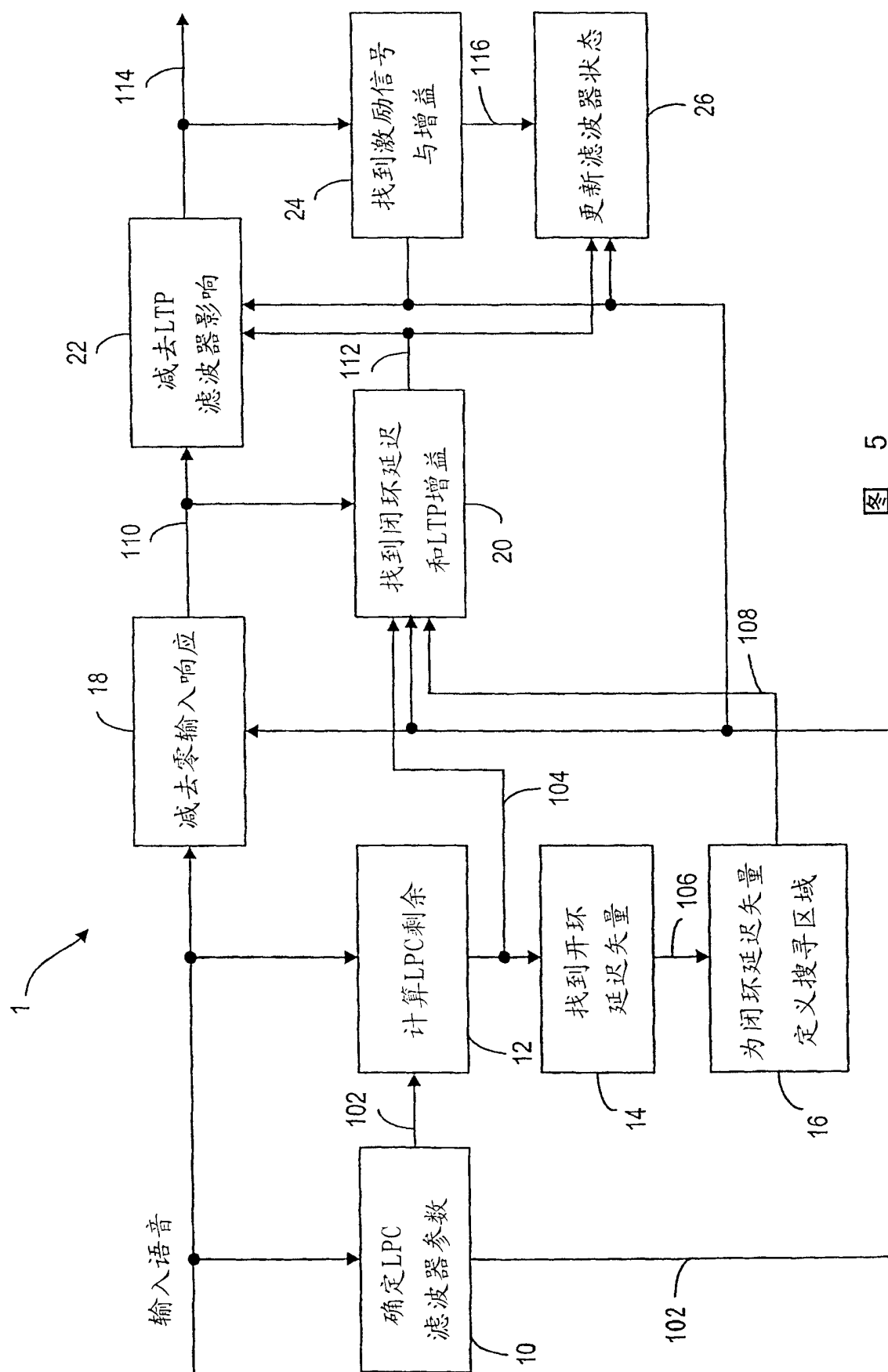


图 5

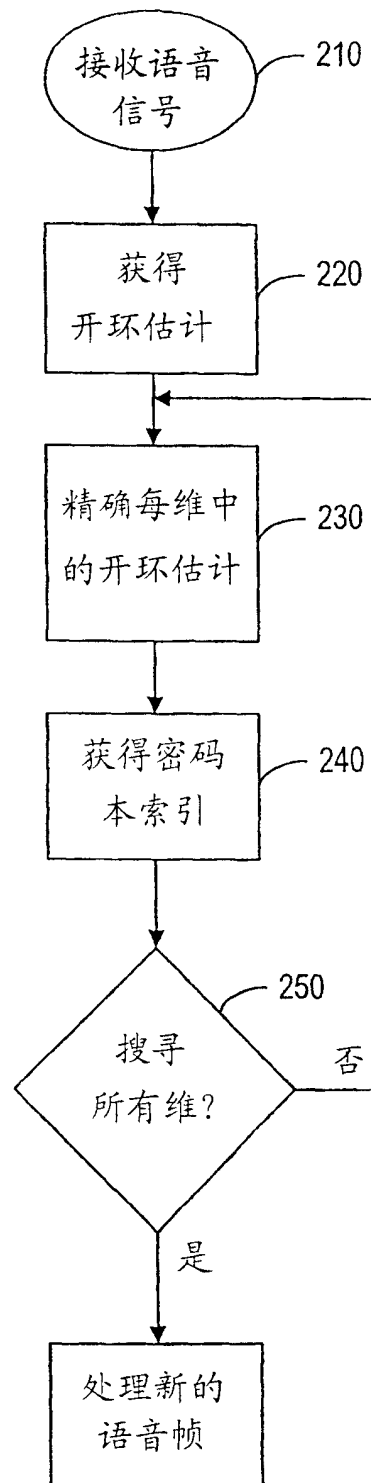


图 6