

[12] 发明专利申请公开说明书

[21] 申请号 99812696.9

[43]公开日 2001 年 12 月 26 日

[11]公开号 CN 1328680A

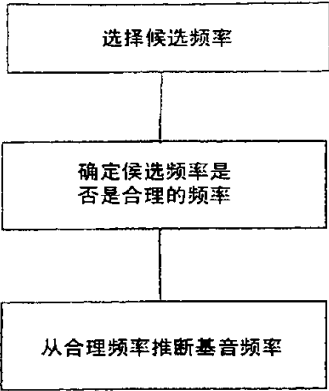
[22]申请日 1999.10.29 [21]申请号 99812696.9
[30]优先权
[32]1998.10.29 [33]US [31]60/106,150
[86]国际申请 PCT/US99/25294 1999.10.29
[87]国际公布 WO00/26896 英 2000.5.11
[85]进入国家阶段日期 2001.4.27
[71]申请人 保罗－里德－史密斯－吉塔尔斯股份合作
有限公司
地址 美国马里兰
[72]发明人 杰克·W·史密斯

[74]专利代理机构 永新专利商标代理有限公司
代理人 蹇 炜

权利要求书 4 页 说明书 20 页 附图页数 6 页

[54]发明名称 快速找到基音的方法
[57]摘要

本发明包含三种快速推断复杂波形或者信号的基音频率的方法。一种方法是利用较高泛音的频率之间的关系,包括:频率比,频率之间的差,频率差比,和来自这种事实的关系,泛音频率由一个整数变量的函数来模型化,该整数的值代表泛音阶数。另一种方法是:所选频率的泛音的预计的/模型化的关系处在对数音阶上,在一个类似的音阶上记录所检测的分音频率,以及相互相对地移动音阶,以寻找三个泛音的匹配。当这样一个匹配找到时,可能的泛音阶数和隐含的基音频率能被直接推断出来。当检测分音还匹配一个预计的/模型化泛音关系的集时,第一种方法的对数被用来选择推断的基音。还有一种方法是:数个基音的泛音频率被收集和编排,这样,与一个未知基音联结的分音能与它们比较,并且未知的基音被推断。



ISSN 1008-4274

权 利 要 求 书

1. 一种从在一个信号中存在的泛音推断一个基音频率的方法，该方法包括：

选择至少两个在该信号中的候选频率；

确定该候选频率是否是一组具有泛音关系的合理的泛音频率；

和

从合理的频率推断所述的基音频率。

2. 根据权利要求 1 所述的方法，其中，确定合理的频率包括采用一个或者一个以上候选频率的比，候选频率的差，和候选频率的差之比。

3. 根据权利要求 2 所述的方法，包括确定这些比是否是等于泛音模型 $f_n = f_1 \times G(n)$ 的比，其中， f_1 是基音频率，和 n 是所述候选频率的阶数。

4. 根据权利要求 3 所述的方法，其中， $G(n) = n \times (S)^{\log_2 n}$ ，这里， S 是一个常数。

5. 根据权利要求 3 所述的方法，其中， $G(n) = n$ 。

6. 根据权利要求 1 所述的方法，其中，确定合理的频率包括：确定一个候选频率比是否实质上等于一个可接受的泛音阶数比。

7. 根据权利要求 1 所述的方法，其中，确定合理的频率包括：确定该候选频率的可接受的泛音阶数。

8. 根据权利要求 7 所述的方法，其中，可接受的泛音阶数作为一个信号源的函数被确定。

9. 根据权利要求 1 所述的方法，包括选择三个在信号中的候选频率和确定合理的泛音频率，包括采用一个或者一个以上的候选频率的比、候选频率的差、和候选频率差的比。

10. 根据权利要求 9 所述的方法，包括从三个候选频率的比确

定三个候选频率的可接受的泛音阶数。

11. 根据权利要求 9 所述的方法，包括确定实质上等于候选频率之比的整数比，以及从一个候选频率与其它两个候选频率的整数比的一个阶匹配确定每一个候选频率的泛音阶数。

12. 根据权利要求 9 所述的方法，包括确定候选频率的泛音阶数；以及确定差比是否等于阶数的差比。

13. 根据权利要求 9 所述的方法，包括确定一个实质上等于差比的整数比；以及确定比的整数是否在一个预定的范围内。

14. 根据权利要求 13 所述的方法，包括确定比的整数是否每个在一个第一值之下和整数的和是否在一个第二值之下。

15. 根据权利要求 9 所述的方法，包括假如头三个候选信号被确定不是合理的频率组，则在信号中选择第四候选频率，并确定第四候选频率和头三个候选频率中的两个是否是一组具有一种泛音关系的合理的频率。

16. 根据权利要求 1 所述的方法，其中，确定合理的频率包括将候选频率与一个基音频率和它的较高的泛音比较，以找到至少一个可接受的匹配。

17. 根据权利要求 16 所述的方法，其中，为泛音而产生一个泛音音阶，为候选频率而产生一个候选音阶，以及候选音阶和泛音音阶相互相对移动，以找到至少一个可接受的匹配。

18. 根据权利要求 17 所述的方法，其中，候选音阶和泛音音阶是同底的对数音阶。

19. 根据权利要求 17 所述的方法，包括产生数个泛音音阶及对应的不同泛音关系的候选音阶。

20. 根据权利要求 16 所述的方法，包括存储带它们的阶数的数个泛音频率组，将候选频率与泛音频率组比较，以确定至少一个可接受的匹配。

21. 根据权利要求 1 所述的方法，其中，确定合理的频率包括：为一组泛音产生一个对数泛音音阶；为候选频率产生一个与泛音音阶的同底的对数候选音阶；以及相互相对地移动候选音阶和泛音音阶，以找到至少一个可接受的匹配。

22. 根据权利要求 21 所述的方法，包括根据候选音阶与泛音音阶的匹配来确定候选频率的阶数，以及利用阶数来确定一组合理的频率。

23. 根据权利要求 1 所述的方法，包括确定合理频率的阶数；并且其中，基音频率采用一个或者一个以上合理的频率和合理频率的差来推断，所述的合理频率用其阶数相除，所述的合理频率的差用其阶数的差相除。

24. 根据权利要求 23 所述的方法，其中，基音频率采用一个商的加权平均来推断。

25. 根据权利要求 23 所述的方法，其中，基音频率用 $(S)^{\log_2 n}$ 除以合理的频率来推断，这里， n 是阶数， S 是一个常数。

26. 一种从存在于一个信号中的泛音确定一个基音频率的方法，该方法包括：

在该信号中选择至少两个候选频率；和

从候选频率的比、差、以及泛音阶数，推断出基音频率。

27. 一种确定在一个信号中的一组分音频率的方法，该频率是一个共同基音频率的合理泛音频率，该方法包括：

选择至少两个在信号中的候选频率；

将候选频率的关系与对应泛音频率的模型化关系比较；

为每一个候选频率确定一个泛音阶数；以及

根据候选频率和阶数推断共同的基音频率。

28. 根据权利要求 27 所述的方法，其中，该模型化关系是 $f_n = f_1 \times n \times (S)^{\log_2 n}$ ，这里， n 是阶数， f_1 是一个基音频率， S 是

一个常数。

29. 一种确定在一个信号中的一组分音频率的方法，该频率是一个共同基音频率的合理泛音频率，该方法包括：

选择至少两个在该信号中的候选频率；

在一个对数候选音阶上标注候选频率；和

将在对数候选音阶上的候选频率与一个对数泛音音阶进行比较，以确定候选频率是否是一个共有基音频率的合理泛音频率，所述的泛音音阶包括一个模型化泛音频率的泛音关系。

30. 根据权利要求 29 所述的方法，包括根据比较的结果来确定候选频率的泛音阶数和候选频率的共有基音频率。

31. 根据权利要求 30 所述的方法，其中，模型化关系是：

$f_n = f_1 \times n \times (S)^{\log_2 n}$ ，这里， n 是阶数， f_1 是一个基音频率， S 是一个常数。

32. 一种确定在一个信号中的一组分音频率的方法，该频率是一个共有基音频率的合理泛音频率，该方法包括：

选择至少两个在该信号中的候选频率；

将该候选频率与数个泛音频率组比较，以找到可接受的匹配；

和

根据可接受的音阶匹配选择最低的推断基音频率作为一个共有基音频率的合理泛音频率。

33. 根据权利要求 1 所述的方法，包括存贮该方法作为在一个数字信号处理器中的指令。

说明书

快速找到基音的方法

交叉参考

本申请涉及 1998 年 10 月 29 日申请的流水号为 60/106,150 的临时专利申请，并要求该申请的权益，该申请在此可结合起来作为参考。

发明的背景和概述

本发明涉及电子音乐的产生和再现，还涉及在放大和增强由音符产生的信号过程中改善声音的电子模拟的方法，并且一般地涉及具有快速确定复合波的基音频率的目的的系统，该复合波是多个频率的总和。

测量一个正弦波信号的频率达到一个特定的音高精度（例如，半音的 $1/4$ ）所要求的时间长度具有一个最小值的限制。该最小时间与待处理信号的频率成反比。保持音高精度为常数，要求测量一个 82.4 赫兹的纯正弦波的频率的时间最小量将是要求测量一个 659.4 赫兹的纯正弦波的频率的最小时间的 8 倍还长。因此，存在这样的问题，即，对于不用键盘（或者当发一个音符的音时，表示基音频率的其它装置）的乐器产生的低音音符的基音频率来说，其测量和再现需要很长时间。例如，当来自低音音符的信号在它们被放大和再现之前由合成器处理时，通常，导致一个令人烦恼的延迟时间。

在整个专利中，一个分音(partial)或者分音频率定义为一个确定的能量频带，而泛音(harmonics)或者泛音频率被定义为：根据一种整数关系，比如，一个机械物体，例如，一条弦，或者一个气柱用结点的整数来分割这种现象所产生的分音。由多级振荡/振动装置产

生的泛音频率之间的关系，包括乐器，可以用一个函数 $G(n)$ 来建立模型：

$$f_n = f_1 \times G(n)$$

这里， f_n 是第 n 个泛音的频率， f_1 是基音频率，也就是所谓的第 1 个泛音， n 是一个正整数，它代表泛音阶数(harmonic ranking number)。这类函数的已知例子是：

$$f_n = f_1 \times n ; \text{ 以及,}$$

$$f_n = f_1 \times n \times [1 + (n^2 - 1) \beta]^{1/2} .$$

这里， β 是一个常数，通常是 0.004。

关于复杂波形的特性和泛音含量以及靠振动物体和这类物体的电气/电子模拟量产生的泛音分音之间的关系，知识和理论本身是存在的。有助于知识本身的理解的教科书例子是：1) Fletcher 和 Rossing 的《乐器物理》，2) Sethares 的《调音、音色、频谱、音阶》，以及 3) Rabiner 和 Schafer 的《语音信号的数字处理》。还包括与测量/确定频率的方法有关的知识 and 理论，例如，固定的和可变的带通和带阻滤波器，振荡器，谐振器，快速傅立叶变换等。知识本身的总述包含在大不列颠百科全书中。

最近特别谈到测量基音频率的方法的专利例如有：

授予 Szalay 的美国专利 5,780,759 描述了一种音高识别方法，该方法采用一个信号的零交点之间的间隔作为该信号的周期长度的一个计量单位。零交点的梯度大小被用来选择零交点，以便评估。

授予 Bartkowiak 等人的美国专利 5,774,836 显示了一种改进的声码器系统，该系统用来估计在一个语音波形中的音高。该方法首先进行一个相关的计算，然后，产生一个基音频率的估计。它再进行错误校验，以去除“错误的”音高估计。在这个过程中，该方法寻找该估计的基音频率的较高的泛音。

授予 Warrander 的美国专利 4,429,609 显示了一种装置和方法，

该装置和方法进行 A/D 转换，去除在有影响的区域之外的频带，并采用零交点时间数据来完成分析，以确定基音。它靠对应于零交点之间的间隔的逐次量来延迟一个参考信号，并把参考信号与延迟信号结合起来，以确定该基音。

本发明是一种采用较高的泛音频率之间的关系，快速推断一个复杂波形或者信号的基音频率的方法。

该方法包括选择至少两个在信号中的候选频率。接着，它要确定该候选频率是否是一组合理的(legitimate)、具有泛音关系的泛音频率。最后，基音频率从合理的频率中被推断出来。

在一种方法中，被检测的分音频率之间的关系与可比关系进行比较，可比关系取决于所有的组成是否是合理的泛音频率。被比较的关系包括频率比，频率差，这些频率差的比，以及由泛音频率靠一个假定只是正整数值的变量函数来模型化的事实引出的特有关系。这个整数值叫做泛音阶数。最好是，整数变量的函数是

$f_n = f_1 \times n \times (S)^{\log_2 n}$ ，其中，S 是一个常数，通常， $1 \leq S \leq 1.003$ ，n 是泛音阶数。S 的值，以下称为升音常数，确定泛音随 n 的值增加而逐步升高的程度。

其它的关系必须保持候选分音频率是否是根据作为信号源的振动/振荡物体或者仪器的物理特征的合理的泛音主干(stem)，即，它能产生的最高和最低基音频率以及它能产生的最高泛音频率。

确定合理的泛音频率和推断一个基音频率的另一种方法包括：将候选频率组与一个基音频率及其泛音比较，以找出一个可接受的匹配。一种方法是产生一个泛音倍数音阶(harmonic multiplier scale)，在该音阶(scale)上，G(n) 的值被记录。这些数值对于 n 的每一个值，即，对于每一个泛音阶数，都是基音频率的倍数。接着，一个类似的音阶被产生，在那里，候选分音频率的数值能被记录。在一组候选分音频率已经被检测和记录之后，两个音阶被比较，即，它

们相互相对被移动，以定位候选频率组与泛音倍数组的可接受的匹配。最好是，该音阶是对数的。当一个好的匹配找到时，候选频率组的一个可能的阶数集就由泛音阶数音阶来确定（或者能被直接读出）。同样地，与合理的分音候选频率组相关的隐含基音频率能被直接地读出。这是在候选频率音阶中的频率，它对应（相应于）泛音倍数音阶上的“1”。

假如函数 $G(n)$ 对于不同的频率寄存器是不同的，以致于在一个频率寄存器中的泛音方式是相关的，这些方式不同于其它频率寄存器中相关的方式，然后，不同的泛音倍数音阶被产生，对于不同的频率寄存器中的每一个，都有一个对应的音阶。分音频率被记录在频率寄存器的合适音阶上，在频率寄存器中，它们终止，并与对应于频率寄存器的泛音倍数音阶比较。

在另一种匹配方法中，候选频率与被检测、测量的来自数个基音频率的多个泛音频率比较。被检测和测量的泛音频率最好组成一个阵列，按基音频率次序组成，列是泛音阶数，行是泛音频率。当三个或者三个以上被检测的分音阵列充分地接近在该阵列的一行中的三个测量到的泛音频率时，泛音阶数和基音就已知了。

因为更高的泛音频率通常能比基音频率更快地被确定，推断基音频率的计算能在一个很短的时间内完成，因此，低音音符的基音频率能在它们被测量到之前被很好地推断。

从下面结合附图对发明的详细描述中，可以更加清楚地了解本发明的其它优点和新颖的特征。

附图的简单说明

图 1 是根据本发明推断基音频率的一种方法的方框图。

图 2 是具体实现图 1 的方法的方框图。

图 3 表示一个对数音阶及一个相应的对数音阶，其中，在对数

音阶上，示出了泛音 1—17 的倍数，在相应的对数音阶上，示出了四个被检测的分音。

图 4 是放大图，表示图 3 中音阶所选择的部分在这些音阶相互相对移动之后，找出三个候选频率与泛音倍数的一个好匹配。

图 5 是图 4 的一个窄频带的放大图，表示匹配位怎样能用作一个匹配程度的计量单位。

图 6 是一个实现图 1—4 的方法的一个系统的方框图。

优选实施例的详细描述

为了从较高的泛音推断基音频率 f_1 ，异常的频率必须被筛选出来，并且至少一个合理泛音组的泛音阶数必须被确定。或者，由两个合理的泛音包括的未占据泛音位置（失去的泛音）的数目必须被确定。一般的方法，如在图 1 所示，选择候选频率。接着，它确定候选频率是否是合理的泛音频率，该频率具有相同基础的基音频率。最后，基音频率从合理的频率来推断。

定义和符号

下面的定义和符号将用于整个专利中：

f_H, f_M, f_L ：分音三音组的候选频率，以下降的频率次序排列。

R_H, R_M, R_L ：与 f_H, f_M, f_L 相关的阶数。

F_L ：最低的基音频率， f_1 ，能由信息源产生。

F_H ：最高的基音频率， f_1 ，能由信息源产生。

F_{MAX} ：最高的泛音频率，能由信息源产生。

关系和限制条件

本方法利用：较高泛音之间的关系、限制选择的条件、较高泛音与基音具有的关系、以及可能的基音频率的范围。例子是：

假如 $f_{R_Z} = f_1 \times G(R_Z)$ 使第 R_Z 个泛音的频率模型化, 并且

假如 f_H , f_M 和 f_L 是合理的泛音频率, 以及

假如 R_H , R_M 和 R_L 是与 f_H , f_M , f_L 相关的阶数, 那末,

下面的比关系必须保持:

a) 被检测的候选频率的比必须大约等于通过在泛音模型中取代它们的阶数所获得的比, 即,

$$f_H \div f_M \approx f_{RH} \div f_{RM}$$

$$f_M \div f_L \approx f_{RM} \div f_{RL}$$

b) 被检测的候选频率之间的差之比必须是与模型化的频率的差之比相一致, 即,

$$(f_H - f_M) \div (f_M - f_L) \approx (f_{RH} - f_{RM}) \div (f_{RM} - f_{RL})$$

c) 候选频率分音 f_H , f_M , f_L , 它们是候选泛音, 必须是在可以由源或者乐器产生的频率的范围内。

d) 泛音阶数 R_H , R_M , R_L 不必隐含一个低于 F_L 或高于 F_H 基音频率, 基音频率的范围可以由源或者乐器产生。

e) 当匹配整数变量比, 以获得阶数的可能三音组时, 在整数比 R_H/R_M 中的整数 R_M 必须与在整数比 R_M/R_L 中的整数 R_M 相同, 例如, 这种关系用来将阶数对 $\{R_H, R_M\}$ 和 $\{R_M, R_L\}$ 加入可能的三音组 $\{R_H, R_M, R_L\}$ 。

方法的概述

本方法分析一组分音或者候选频率, 并确定它们是否包括异常的频率。最好是, 被分析的每一组将包含三个分音。假如确定不存在一个或者一个以上的异常频率, 则认为该组是一组合理的泛音频率。每一泛音频率的阶数被确定, 基音频率被推断。当确定存在一个或者一个以上的异常频率时, 一个新的分音或者候选频率被检测、测量和选择, 异常的频率被分隔和筛选出来。这一过程持续到一组合理泛音频率保留为止。在这一过程中, 合理的泛音频率的阶数被

确定和确认。然后，基音频率用多种方法计算。考虑泛音不同于 $f_n = f_1 \times n$ 的程度，而进行调整。

方法 I

下面是实现图 1 的方法的简要流程图以推断基音频率的一种方法的例子，并表示在图 2 中。该方法测试被检测的候选分音频率的三音组，以确定其成分是否只是相同的基音频率的合理泛音频率。当不是这样时，附加的候选频率就被引入，并用来代替现有的三音组中的频率，直到一个合理泛音的三音组已经被找到为止。当这样的一个三音组被找到时，与每一个成分相关的阶数就被确定，基音频率被推断。

在此描述的这种方法表明了逻辑运算的种类，该运算将直接或者间接地被完成。实际执行将通过捷径，消除冗余信息等，并可以以其它的方式区别于下面所描述的执行。

该方法用一般术语描述的一系列步骤被呈现出来，同时一个数字表示的例子表明了对不同步骤所要求的计算。

乐器常数的定义

K_1 是最高的泛音阶数，该数将被赋值/估量。将在测量第 K_1 个泛音的频率过程中的预计的 % 误差与整数比的商值进行比较，由此来设定 K_1 的值

$$[(K_1 + 1) \div K_1] \div [K_1 \div (K_1 - 1)]$$

K_1 的一个默认值将被设定等于 17，并将修改以符合在手边的乐器的知识和在频率测量中预计的误差。

K_2 是两个邻近被检测的泛音频率之间漏掉的泛音的最大预计数。 K_2 的默认值被设定等于 8。

K_3 等于两个泛音之间漏掉的泛音的预计最大值之和，加 1，所

述的泛音包含一个插入的或中间的泛音。 K_3 的默认值被设定等于 12。

步骤 1. 为乐器或者信号源设定常数/参数。

例如： $F_H = 300 \text{ Hz}$, $F_L = 30 \text{ Hz}$, $F_{\max} = 2,100 \text{ Hz}$; $K_1 = 17$,
 $K_2 = 8$, $K_3 = 12$

为了简化和简洁，描述泛音频率 $G(n)$ 之间的关系的函数被假定为 $f_1 \times n$ 。

步骤 2. 检测、测量和选择，例如，三个分音的频率。这些频率按它们产生的阶被检测和测量。三种频率或者分音，具有一个明显超过，例如，环境噪音级的能级，被选择作为可能的合理泛音的候选者。较高的频率，和次高的泛音频率，自然地被首先检测和测量。下面的例子假定一种例外，并表明怎样处理这种例外情况，在这种例外情况下，较低的泛音在一个较高的泛音之前被检测。

例如：测得第 1 频率 = 722 Hz ，

测得第 2 频率 = 849 Hz ，

测得第 3 频率 = 650 Hz 。

步骤 3. 这三个候选频率以频率的大小阶排列，标记为 f_H, f_M, f_L 。

例如： $f_H = 849 \text{ Hz}$, $f_m = 722 \text{ Hz}$, $f_L = 650 \text{ Hz}$ 。

步骤 4. 对于候选频率 f_H, f_M, f_L ，可能的阶数的三音组被确定。
 f_H/f_M 和 f_M/f_L 比值的商与整数比 I_a/I_b 的商进行比较，这里， I_a 和 I_b 都 $\leq K_1$ ，一个给定的阈值。在此，为说明的目的， K_1 被设定等于 17。
 当一个频率比的商充分地接近一个整数比的商时，该整数比被保留作为代表一对频率比的可能的阶数和与它匹配的数。比值还可以是 f_H

$/f_L$ 和 f_M/f_L 或者 f_H/f_M 和 f_H/f_L 或者任何的倒数。

例如：当 $f_H/f_M=1.176$ 时，最接近的整数比的商是 $1.1818=13/11$ 和 $1.1667=7/6$ 或者 $14/12$ 。注意 $26/22$ 不会被考虑，因为 $26 > 17$ 。当 $f_M/f_L=1.111$ 时，最接近的整数比的商是 $1.111=10/9$ 和 $1.10=11/10$ 。

当两个比的共有频率相等时，则形成一个可能的阶数三音组 $\{R_H, R_M, R_L\}$ 。在这个例子中，它是当整数分数 f_H/f_M 的分母与整数分数 f_M/f_L 的分子相等时的情形。

例如：因为只有 $f_H/f_M=13/11$ 和 $f_M/f_L=11/10$ 导致 f_M 的相同的阶数，所以，在这个例子中仅有可能的三音组是 $\{R_H, R_M, R_L\}=\{13, 11, 10\}$ 。

步骤 5. 所有阶数的可能三音组被排除，这隐含着一个由 F_L 和 F_H 确定的在该范围外的基音频率 f_1 。

例如：基音 f_1 是由它的阶数相除的候选频率。仅有的可能三音组 $\{13, 11, 10\}$ ，不会被筛选掉，因为 $f_H/13=65.308$ ， $f_M/11=65.636$ ，以及 $F_L/10=65.00$ 都是在 $F_L=30$ 和 $F_H=300$ 确定的范围内。

步骤 6. 差值 $D_{H,M}=f_H-f_M$ 和 $D_{M,L}=f_M-f_L$ 被计算， $D_{H,M}/D_{M,L}$ 的比值被算出。其它的能被相似使用的差比是 $D_{H,L}/D_{M,L}$ 或者 $D_{H,L}/D_{H,M}$ 。

例如： $D_{H,M}=849-722=127$ ， $D_{M,L}=722-650=72$ ，并且

$$D_{H,M}/D_{M,L}=127/72=1.764。$$

步骤 7. 差比的商 $D_{H,M}/D_{M,L}$ 与小整数比 I_c/I_d 的商相比较，这里， $I_c < K_2$ ，和 $I_c + I_d < K_3$ 。注意：整个例子中，数值 $K_2=8$ 和 $K_3=12$ 。 $K_2=8$ 对应一个假定： f_H 和 f_M 相差不大于 7 倍的基音频率，或者泛音阶数 R_H 和 R_M 相差不大于 7。同样地， $K_3=12$ 是假定： f_H 和 f_L 将相差不大于 11 倍基音频率，并且阶数 R_H 和 R_L 相差不大于 11。域数据的

粗略检查确定了这些假定。假如其它的差比被使用，则 K_2 和 K_3 的值采用相同的分析而被大约地设定。

例如： $D_{H,M}/D_{H,L}=1.764\approx 1.75=7/4$ 。这个比适合首先考虑，因为 $7<8$ 并且 $7+4<12$ 。

步骤 8. 任何隐含着基音频率 $f_1 < F_L$ 的差比是不合格的。

例如：在此，差比 $7/4$ 隐含着最高频率 $f_H=849\text{Hz}$ 和最低频率 $f_L=650\text{Hz}$ 之间的差等于 198Hz ，应该是大约等于 $7+4$ 或者 11 倍基音频率。这样，隐含着 $f_1=199/11=18.1$ ，小于 $F_L=30$ 。对于 $D_{H,M}/I_c$ 和 $D_{M,L}/I_d$ ，同样也是正确的。这只是隐含着一个人或者一个以上异常频率的存在。步骤 9 将表示：另一种比较仍然隐含着异常频率是在三音组的候选频率中。

步骤 9. 假如与频率差比匹配的整数比 I_c/I_d 同对应的阶数比

$(R_H - R_M) \div (R_M - R_L)$ 不一致，则任何阶数 R_H ， R_M ， R_L 的三音组都是不合格的。

例如：只有可能的阶数三音组是 $\{13, 11, 10\}$ 。它被筛选掉，因为 $7/4 \neq (13-11) \div (11-10) = 2$ 。

步骤 10. a) 假如存在不协和的不一致(unresolvable inconsistencies)，则进入步骤 11。

例如：第一次通过，一个新的频率被选择并且异常频率被消除之前，存在不协和的不一致。所有可能的阶数三音组被筛选掉，差比导致不一致。

b) 假如没有不协和的不一致，那么，一致的三音组已经被发现合理的，则进入步骤 17，以推断基音频率。

例如：在这种情况下，一个新的频率已经被引入和在原始三

音组中的第二频率已经被取代之后，没有发现不协和的不一致，如下面所示。

步骤 11. 所有已经被测量和检测的频率已经被选择了吗？假如不是，进入步骤 12，假如是，进入步骤 16。

步骤 12-14. 为了找到一个候选频率的三音组，原始的三个候选频率与一个或者一个以上的附加候选频率被使用，以确定一个合理的三音组。对一个三音组，假如它是第一次经过这一过程，则进入步骤 13，以选择一个第 4 候选频率，进入步骤 14，以取代三音组中的一个频率。确定一个由第四候选频率和原始候选频率三音组中的两个候选频率所组成的合理三音组从步骤 3 开始进行。

假如第四候选频率的第 1 次取代没有产生一个合理的三音组，则从步骤 12 直接进入步骤 14。一个第二原始候选频率由第四候选频率取代，以形成一个新的三音组。假如这没有产生一个合理的三音组，则第四候选将取代第三原始候选频率。

假如，第三次通过步骤 12 时确定：第四候选频率取代了原始三音组中的每一个频率之后，没有发现合理的或者一致的三音组，则进入步骤 15。

例如：因为在原始的三音组 {849, 722, 650} 中存在不协和的不一致，则选择一个新的频率。该新的频率是 602 Hz。

数值 849 由 602 所取代，以形成三音组 {722, 650, 602}，该组被指定作为新的候选三音组 $\{f_H, f_M, f_L\}$ 。

因 $f_H / f_M = 1.111$ ，故最接近的整数比是 10/9, 11/10, 和 9/8。

因 $f_M / f_L = 1.0797$ ，故最接近的整数比是 14/13, 13/12, 和 15/14。不存在匹配的阶数。

而且，没有一致的三音组被找到。

在原始的三音组中的一个不同频率被取代，即，722 由 602 所取代，再插入原始的频率 849，以形成三音组{849, 650, 602}，该组被指定作为新的候选三音组 $\{f_H, f_M, f_L\}$ 。

因 $f_H / f_M = 1.306$ ，故最接近的整数比是 13/10, 17/13, 和 14/11。

因 $f_M / f_L = 1.0797$ ，故最接近的整数比是 14/13, 13/12, 和 15/14。

$f_H / f_M = 17/13$ ，而 $f_M / f_L = 13/12$ ，形成一个可能的阶数三音组，它就是： $\{R_H, R_M, R_L\} = \{17, 13, 12\}$ 。

$$(f_H - f_M) \div (f_M - f_L) = 199/48 = 4.146 \approx 4.$$

$$(R_H - R_M) \div (R_M - R_L) = 4/1 = 4, \text{ 它与频率差比相一致。}$$

而且， $f_H \div R_H = 49.94$ ， $f_M \div R_M = 50$ ， $f_L \div R_L = 50.17$ 。所有的都是大于 $F_L = 30$ 。

所有的条件都满足，因此， R_H ， R_M ，和 R_L 被假定为：分别是 17, 13 和 12，候选频率 849, 650, 602 被认为是一个合理的三音组。现在在步骤 17 中确定基音频率。

步骤 15. 一个第五和第六候选频率被选择。第四频率与第五和第六候选频率组合，以形成一个新开始的三音组，该方法将从步骤 3 开始被执行。步骤 12 将被复位到 0 通过。

步骤 16. 假如所有检测和测量到的频率通过步骤 11 已经被选择和确定之后，没有一致或者合理的三音组在步骤 7-10 已经被找到，所有被选择的频率中的最低的将被认为是基音。

步骤 17. 通过下面的任何一种方法来推断基音频率，例如，其中， $G(n) = n$ ， $f_H = 849 \text{ Hz}$ ， $f_M = 650 \text{ Hz}$ ， $f_L = 602 \text{ Hz}$ ， $\{R_H, R_M, R_L\} = \{17, 13, 12\}$ ：

$$a) f_1 = f_H / R_H$$

$$b) f_1 = f_M / R_M$$

$$c) f_1 = f_L / R_L$$

$$d) f_1 = (f_H - f_M) \div (R_H - R_M)$$

$$e) f_1 = (f_M - f_L) \div (R_M - R_L)$$

$$f) f_1 = (f_H - f_L) \div (R_H - R_L)$$

例如：找到一个带相应的阶数的一致的合理频率三音组是 {849, 650, 602} 和 {17, 13, 12} 之后：

$$a) f_1 = 849/17 = 49.94 \text{ Hz}$$

$$b) f_1 = 650/13 = 50.00 \text{ Hz}$$

$$c) f_1 = 602/12 = 50.17 \text{ Hz}$$

$$d) f_1 = (849 - 650) \div (17 - 13) = 49.75 \text{ Hz}$$

$$e) f_1 = (650 - 602) \div (13 - 12) = 48.00 \text{ Hz}$$

$$f) f_1 = (849 - 602) \div (17 - 12) = 49.4 \text{ Hz}$$

该推断的基音能被设定等于 6 个计算值的各种加权平均数的任何值。例如：

采用比的计算的方法，例如，从 a) 到 c)，计算得出 f_1 的平均值 = 50.04 Hz。

考虑到频率差方法，跨过泛音的最大数，根据 f) 以上所给定的数值，计算得出 f_1 的值 = 49.4 Hz。

通过比方法和跨过泛音的最大数的差方法计算，得到平均 f_1 的值为：

$$(50.04 + 49.4) \div 2 = 49.58。$$

这三种平均方法应该为推断基音频率产生合理值。最后的是优选的，除非/直到域数据引入一个更好的平均方法。

b) 假如在手边的乐器的泛音已经按照函数 $f_n = f_1 \times n \times (S)^{\log_2 n}$ 模型化，这里， $S > 1$ ，一种更精确的推断基音的方法将如下面所述：

$$a) f_1 = (f_H \div S^{\log_2 R_H}) \div R_H$$

- b) $f_1 = (f_M \div S^{\log_2 R_M}) \div R_M$
- c) $f_1 = (f_L \div S^{\log_2 R_L}) \div R_L$
- d) $f_1 = [(f_H \div S^{\log_2 R_H}) - (f_M \div S^{\log_2 R_M})] \div (R_H - R_M)$
- e) $f_1 = [(f_M \div S^{\log_2 R_M}) - (f_L \div S^{\log_2 R_L})] \div (R_M - R_L)$
- f) $f_1 = [(f_H \div S^{\log_2 R_H}) - (f_L \div S^{\log_2 R_L})] \div (R_H - R_L)$

假如高音调的常数 S 已经被设定等于 1.002, 则基音的推断值将如下面所述:

- a) $f_1 = 49.535 \text{ Hz}$.
- b) $f_1 = 49.63 \text{ Hz}$.
- c) $f_1 = 49.81 \text{ Hz}$.
- d) $f_1 = 49.22 \text{ Hz}$.
- e) $f_1 = 47.51 \text{ Hz}$.
- f) $f_1 = 48.88 \text{ Hz}$.

采用比的计算方法, 例如, 从 a) 到 c), 计算得出 f_1 的平均值等于 49.66 Hz 。

考虑到频率差方法, 跨过泛音的最大数, 根据 f) 以上所给定的数值, 计算得出 f_1 的值等于 48.88 Hz 。

通过比方法和跨过泛音的最大数的差方法计算, 得到平均 f_1 的值为:

$$(49.66 + 48.88) \div 2 = 49.27 \text{ .}$$

这三种平均方法中的任何一种都可以用于推断基音。最后的是优选的。

假如在步骤 9 完成之后, 两个或者两个以上一致的阶数组保留, 则基音 f_1 应该与阶数组的每一个重新计算, 获得的最低频率, 它与在步骤 3 到 9 中描述的条件一致, 被选择作为推断的基音频率 f_1 。

前面的描述和给定的例子都假定泛音频率以

$$f_n = f_1 \times G(n) = f_1 \times n \times (S)^{\log_2 n} \text{ 为模型, 这里, } 1 \leq S \leq 1.003.$$

后面的函数中， S 是这样的接近 1，隐含着 f_n / f_m 将大约等于整数比 n/m ，频率差的比 $(f_H - f_M) \div (f_M - f_L)$ 将大约等于一个小的整数比和，以及 $f_x - f_y \approx (X - Y) \times f_1$ 。

在一般的情况下，合理的泛音分音的三音组被分离，它们对应的阶数按下述方法确定：

- a) 将 $f_H \div f_M$ 和 $f_M \div f_L$ 的商与比值 $G(R_H) \div G(R_M)$ 以及比值 $(R_M) \div G(R_L)$ 的商分别进行比较。
- b) 将频率差比 $(f_H - f_M) \div (f_M - f_L)$ 与函数差比 $[G(R_H) - G(R_M)] \div [(R_M) - G(R_L)]$ 进行比较。
- c) 将阶数的可能组合所隐含的基音频率与在手边的乐器能产生的最低的基音频率和最高的泛音频率进行比较。

方法 II

另一种分离检测到的分音的三音组的方法如图 3，4 和 5 所示，该分音只是由具有同样基础的基音频率的合理泛音频率组成，以便找到它们的相关阶数，并确定由每一个这样的三音组所隐含的基音频率。该方法标记和记录被检测的分音频率在一个对数音阶上，并将这些分音之间的关系与一个类似的表示预计的/模型化的泛音频率之间的关系的关系的音阶进行匹配。

用下面的一个例子来说明一般的概念。它表明一种能用来匹配或者找到对泛音频率的调号或者图谱的最合适接收信号的方法，并且仅表明将采用的逻辑运算的种类。该例子应该被认为是本发明的一种可能的体现，而不应认为是一种限制。

为了说明这一例子，假定：在手边的乐器产生的泛音按照函数 $f_n = f_1 \times n \times (S)^{\log_2 n}$ 模型化，这里， n 是一个正整数 1, 2, ..., 17，而 S 是一个等于 1.002 的常数。根据该函数，一个泛音倍数音阶（以下称之为 HM 音阶）被建立，其中，每一个梯度标号代表百

分之一，它是半音的 $1/100$ 或者八音度的 $1/1200$ 。在音阶上的第一标记代表泛音倍数 1，即，当用 f_1 乘以该数字时，得到 f_1 。在音阶上的每一个相继的标记代表前一个倍数数字本身用 $[2 \times S]^{1/1200}$ 相乘。假定采用一串位 (bits)，每一位代表百分之一。第 n 位将代表倍数 $[(2 \times S)^{1/1200}]^{(n-1)}$ 。沿 HM 音阶所选择的位将代表泛音倍数，并将用适当的泛音数标记： f_1 将由位 1， f_2 由位 1200， f_3 由位 1902， f_4 由位 2400， $\dots$ ， f_{17} 由位 4905 代表。该音阶如图 3 所示。

另一种音阶被建立，以便标记和记录被检测到的候选分音频率。开始的梯度标号，由位 1 代表，将代表频率 F_L ；下一个，用 $F_L \times [(2 \times S)^{1/1200}]^1$ ，再下一个，用 $F_L \times [(2 \times S)^{1/1200}]^2$ 代表。第 n 位将代表 $F_L \times [(2 \times S)^{1/1200}]^{n-1}$ 。该音阶通常称为候选分音频率音阶，以下称 CPF 音阶。它与 HM 音阶一起在图 3 中示出。

当分音被检测时，它们的频率被标记和记录在 CPF 音阶上。当三个已经这样被检测、标记和记录时，CPF 音阶相对于 HM 音阶移动，寻找匹配。假如在音阶的任何地方都没有找到三个候选频率的一个匹配，则另一个分音频率被检测、标记和记录，寻找匹配的第三个将继续。当候选分音的一个三音组的成分在一个特定的限度内匹配 CPF 音阶上的一组倍数时，则候选频率被假定是合理的泛音频率，它们的阶数与在 CPF 音阶上其对应部(counterparts)的阶数匹配。同样地，所隐含的基音能被直接推断。它是在 CPF 上的频率位置与在 HM 音阶上的“1”匹配。

图 4 表示音阶的部分，在该部分，音阶被移动到揭示三个频率的一个好的调准之后，被检测的候选频率就出来了，即，第 4 频率被检测到了， 421 Hz ，与被检测的第 1 和第 3 频率， 624 Hz 和 467 Hz 组合。

一种测量一个候选分音和一个泛音倍数之间的调准(alignment)程度的方法是扩大位，该位标记候选分音频率及泛音倍数，使它们

成为多个邻近位的集。在这个例子中，在 HM 音阶上，7 个位在每个标记泛音倍数的位的两边转动(turned)。同样地，在 CPF 音阶上，7 个位在每个标记候选分音频率的位的两边转动。当音阶相互相对移动时，匹配位的数字提供一个调准程度的大小。当在一个候选频率的三音组中匹配位的数字超过一个阈值时，例如，45 个位中的 37 个位，则候选分音的调准被认为是可接受的，候选频率被指定作为一个合理泛音频率的三音组。图 5 表示匹配的程度，例如，在一个候选分音频率，即， 624H_2 ，与第 12 泛音的倍数之间一个可能的 15 中的 12。

当一个可接受的调准或者匹配被找到时，采用方法 I 的步骤 6-9，所隐含的阶数被用来测试不协和的不一致。假如没有不协和的不一致被找到，所隐含的基音是比 f_L 更低或者比 f_H 更高，则音阶分别被移动，寻找分别隐含一个更高的基音或者一个更低的基音的调准。当没有不协和的不一致被找到时，所隐含的基音就存在在 f_L 和 f_H 之间，然后，所隐含的基音 f_i 变成推断的基音。

某些等级的乐器/装置具有共鸣带和/或音区，它们产生的泛音从系统上比在其它共鸣带和/或音区的泛音更高。同样地，一些乐器的泛音在一些频带可以是系统的和可预测的，而在其它频带正相反。在这些情况下，方法 II 能按下面的步骤使用：

1. 分隔 S 在整个频带一致的频带。
2. 建立一个 HM 音阶，只用于基于 S 所属频带的该频带的频率。
3. 建立另一个 HM 音阶，用于其它的 S 不同值应用的频带。
4. 当频率被检测时，将它们定位在 CPF 音阶中，该音阶用 S 的数值构成，S 的数值适用于包含该频率的频带。
5. 忽略检测的频率，这些检测的频率位于泛音不可预测的频带内。
6. 采用类似音阶（相同的 S 值），寻找泛音倍数模式和被检测

的候选频率模式之间的匹配。

方法 III

另一种推断基音频率的方法需要检测和测量或者计算数个基音频率的泛音频率。这些频率被编在一个行是基音频率，列是泛音阶数的阵列中。当一个具有未知基音频率的音符被演奏时，较高的泛音的频率在它们被检测到时逐行地与表示在该阵列中的泛音频率比较。与该阵列中三个或者三个以上频率的良好匹配(good match)或者与从该阵列成分插入的频率的良好匹配表示一个可能的阶数组和一个可能的推断基音频率。当一个被检测的频率三音组与在该阵列中的两个或者两个以上基音频率匹配，并因此包含两个或者两个以上基音频率时，推断的基音频率被设定等于包含的基音频率的最低值，它与在手边的乐器能产生的音符一致。该阵列只是一个为快速取得而编排频率的一种方法的例子，其它的方法也可以被使用。

上面的方法 I, II 和 III 能用于分隔和编辑异常分音。例如，给定一个单声道音乐，在一个时间过程中所有的分音被检测之后，当推断的基音保持一定时，这些方法能用来识别所有的不是给定基音产生的泛音集的合理成分的分音。该信息能被使用，例如，用于：a) 编辑来自音乐声道的外来的声音；或者 b) 分析异常的东西，以确定它们的源。

尽管在一些特殊的情况下，只要两个合理的泛音就充足了，但通常，方法 I、II 或者 III 将要求三个或者三个以上合理的泛音频率。为了从两个高次泛音推断出基音频率，下面的条件必须优先：a) 必须知道不代表合理泛音的异常分音频率是极少的，以致该可能性能被忽略；和 b) 两个频率的比必须是这样：两个频率的阶数被唯一地建立。例如，假设两个频率是 434Hz 和 404Hz ，该频率比的商就处在 $14/13$ 和 $15/14$ 之间。假如 $f_1=30\text{Hz}$ ，那末，阶数被唯一地建

立为 14 和 13，因为阻碍条件(brake) $434 \div 15 = 28.9$ ，小于 30，这样，不符合要求。两个候选频率的差是 30，是可接受的，因为它不小于 F_L 。还有，比值： $(F_H - F_L)(R_H - R_L) = 30$ ，仍不小于 F_L 。

函数 $f_n = f_1 \times n \times (S)^{\log_2 n}$ 用于使泛音模型化，这些泛音随 n 增加而逐级变高。 S 是变高的常数，通常设定在 1—1.003 之间， n 是一个正整数 1, 2, 3, ---, T ，这里， T 通常等于 17。根据该函数，由 S 的值来确定变高的范围。当 $f_n = n \times f_1$ 时，与泛音是谐和的方式一样，模型化的泛音也是谐和的。即，假如 f_n 和 f_m 是一个音符的 n 次和 m 次泛音，那末， $f_n/f_m = f_{2n}/f_{2m} = f_{3n}/f_{3m} = \dots = f_{kn}/f_{km}$ ，其中， k 是一个正整数。

实现该方法的系统被表示在图 6 中。预处理级接收或者拾起来自源的信号。它可以包括一个在乐器上弦的拾音器。预处理还调节(conditions)信号。这可以包括使输入信号的幅度、及频率和/或频带的限制标准化。接着，用一个频率检测级分隔具有足够能量的频带，该频带明显是环境噪音并有适当的界定。

快速找到基音级完成候选频率的分析和推断该基音。后处理级采用由快速找到基音级产生的信息，以处理该输入信号。这可以包括放大、调制和其它的信号操作处理。

本方法已经描述了利用泛音频率之间的关系来推断基音。只确定泛音关系和它们的阶，而不推断基音，这也是有价值的。基音频率可以不存在于波形中。较高的泛音可以用于找到其它的泛音，而不推断基音。这样，后处理将采用现有识别的泛音。

尽管本发明已经针对歌唱或者乐器产生的声音进行了描述，但它可以包括其它的复杂波的源，该波具有一个基音频率和较高的泛音。这些能包括：例如，一个说话的声音，复杂机械或者其它机械振动单元的声音。

尽管本发明已经被详细地描述和展示，但是，很显然，这仅仅是

展示和举例的目的，而不是限制性的。本发明的精神和范围只能由附属的权利要求来限定。

说明书附图

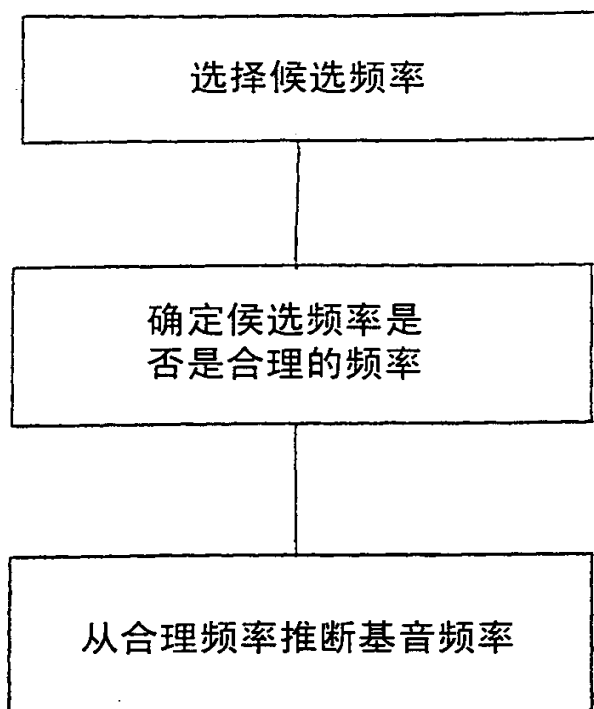


图1

图 2

- 步骤 1. 为乐器或者源设定常数 (F_H , F_L , F_{MAX} , $G(n)$, K_1 , K_2 , K_3)
- 步骤 2. 选择三个候选频率。
- 步骤 3. 指定候选频率 f_H , f_M , f_L .
- 步骤 4. 确定 f_H , f_M , f_L 的可能的三音组 (R_H , R_M , R_L)。
- 步骤 5. 排除隐含 $f_1 < f_L$ 或者 $f_1 > F_H$ 的三音组。
- 步骤 6. 形成频率差比。
- 步骤 7. 初步限定。
- 步骤 8. 排除隐含的 $f_1 < F_L$ 的差比。
- 步骤 9. 排除与差比不一致的阶数的三音组。
- 步骤 10. 一个或者一个以上一致阶数的三音组保留吗？是→17
否↓
- 步骤 11. 所有的频率选择吗？是→16
否↓
- 步骤 12. 第 1 次？否→ 第 3 次？否→14
是↓ 是→15
- 步骤 13. 选择新频率。
是↓
- 步骤 14. 用新频率取代三音组 f_H , f_M , f_L 中的一个频率 。→3
- 步骤 15. 选择 2 个新频率。→3
- 步骤 16. 设定 f_1 等于最低候选频率。
- 步骤 17. 推断 f_1 。

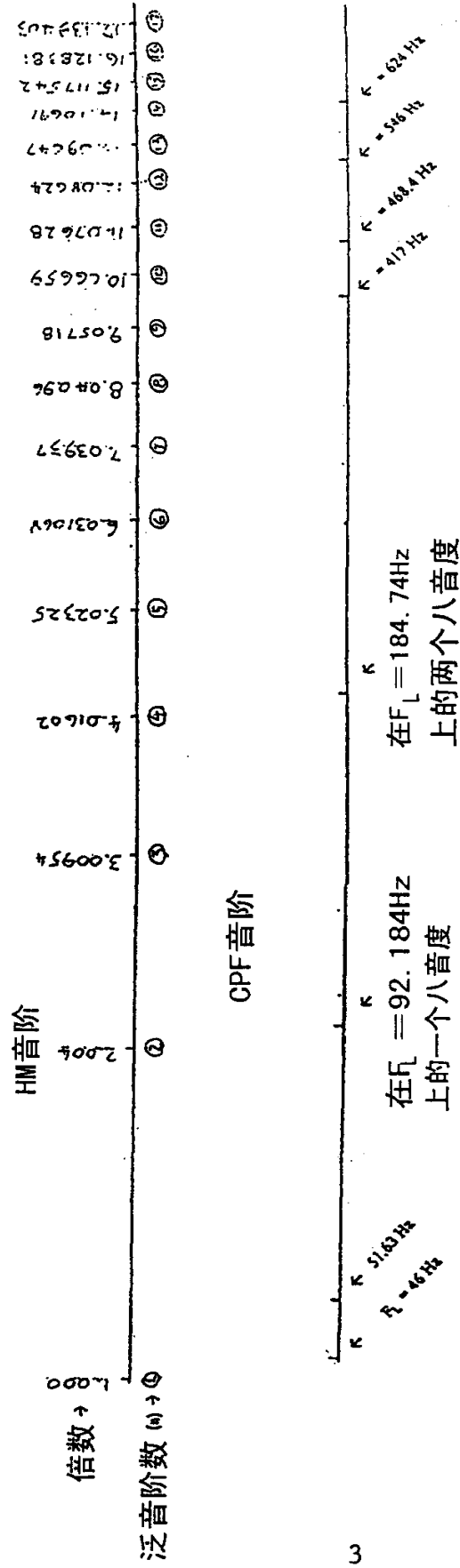


图3

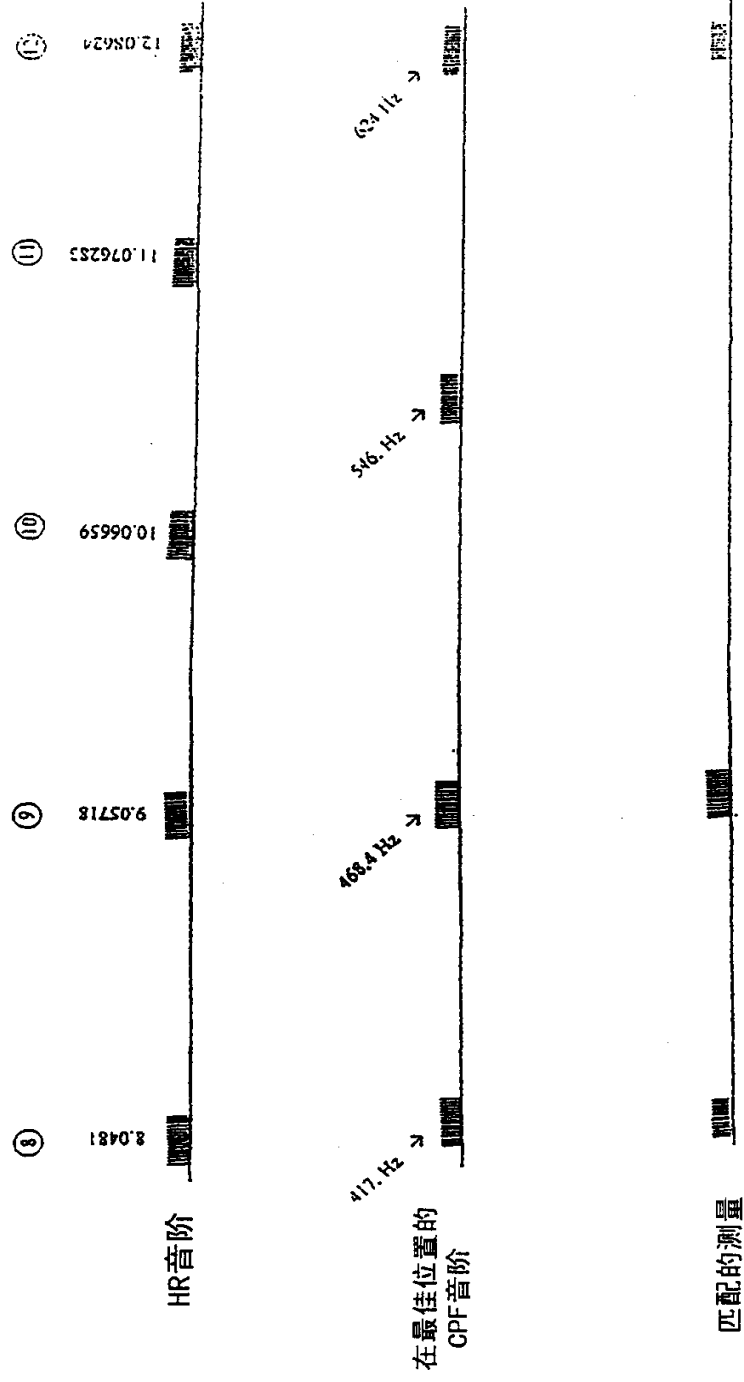


图4

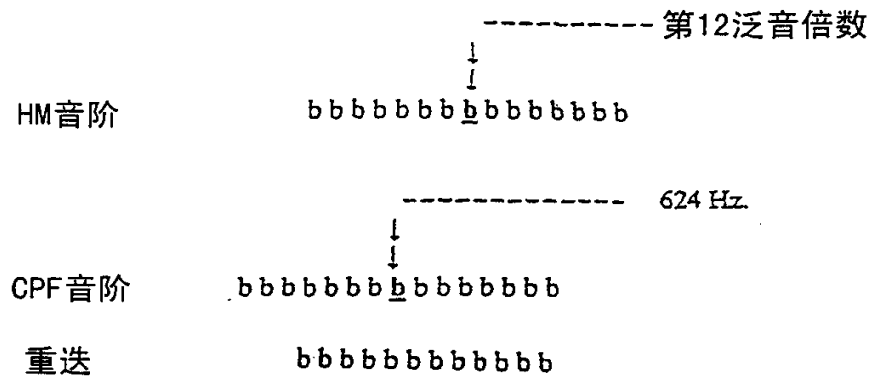


图5

010404

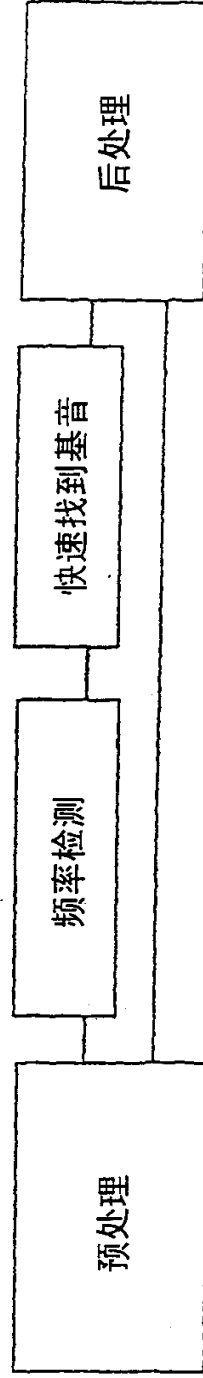


图6