

(12) 发明专利

(10) 授权公告号 CN 101116137 B

(45) 授权公告日 2011.02.09

(21) 申请号 200680004536.4

(22) 申请日 2006.02.09

(30) 优先权数据  
034984/2005 2005.02.10 JP

(85) PCT申请进入国家阶段日  
2007.08.10

(86) PCT申请的申请数据  
PCT/JP2006/302258 2006.02.09

(87) PCT申请的公布数据  
W02006/085586 JA 2006.08.17

(73) 专利权人 松下电器产业株式会社  
地址 日本大阪府

(72) 发明人 张峻伟 梁世丰 吉田幸司  
后藤道代

(74) 专利代理机构 北京市柳沈律师事务所  
11105

代理人 黄小临

(51) Int. Cl.  
G10L 19/10 (2006.01)  
G10L 19/00 (2006.01)

(56) 对比文件

WO 03/065353 A1, 2003.08.07,  
WO 03/038813 A1, 2003.05.08,  
CN 1395742 A, 2003.02.05,  
CN 1391689 A, 2003.01.15,  
CN 1320258 A, 2001.10.31,  
WO 02/23529 A1, 2002.03.21,  
TELECOMMUNICATION STANDARDIZATION  
SECTOR OF ITU. 1. 《ITU-T G. 722.2 (2003.07)  
Wideband coding of speech at around 16  
kbit/s using Adaptive Multi-Rate Wideband  
(AMR-WB)》. INTERNATIONAL TELECOMMUNICATION  
UNION, 2003,  
Bruno Bessette, et. al.. The Adaptive  
Multirate Wideband Speech Codec (AMR-WB).  
《IEEE TRANSACTIONS ON SPEECH AND AUDIO  
PROCESSING》. 2002, 第 10 卷 (第 8 期),

审查员 刘涛

权利要求书 1 页 说明书 8 页 附图 10 页

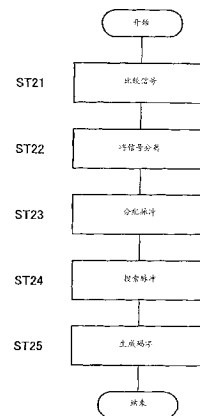
(54) 发明名称

语音编码中的脉冲分配方法

(57) 摘要

能够对立体声语音信号高效率地进行编码的脉冲分配方法。在该脉冲分配方法的固定码本搜索 (ST21 ~ ST25) 中,对各个子帧进行立体声信号的比较,判断声道间的相似性 (ST21),判断立体声信号的特性 (ST22),基于声道间的相似性和立体声信号的特性,决定对各个声道分配的脉冲数 (ST23),进行脉冲搜索并决定对各个声道的脉冲位置 (ST24),并对在 ST24 所决定的脉冲进行编码 (ST25)。

CN 101116137 B



1. 一种脉冲分配方法,该脉冲分配方法为在对立体声信号的语音编码中,进行固定码本搜索时所使用的脉冲分配方法,其包括:

判断立体声信号的各个声道间的相似程度的相似程度判断步骤;

基于作为各个声道的特性的稳定度和周期性,判断声道是否为稳定有声的稳定有声判断步骤;以及

根据所述相似程度判断步骤和所述稳定有声判断步骤的判断结果,决定分配给各个声道的脉冲数的分配步骤。

2. 如权利要求 1 所述的脉冲分配方法,

在所述相似程度为阈值以上时,对任意一方的声道分配所有的脉冲。

3. 如权利要求 1 所述的脉冲分配方法,

在所述分配步骤中,基于各个声道的最大自相关系数,决定分配给各个声道的脉冲数。

4. 如权利要求 1 所述的脉冲分配方法,

对所述稳定度和所述周期性越高的声道,分配越少的脉冲数。

5. 如权利要求 1 所述的脉冲分配方法,

在各个声道的特性相等时,对各个声道均等地分配脉冲数。

6. 如权利要求 1 所述的脉冲分配方法,

将表示已分配给各个声道的脉冲数的码字,通知给语音解码端。

7. 如权利要求 3 所述的脉冲分配方法,

对所述最大自相关系数越大的声道,分配越少的脉冲数。

## 语音编码中的脉冲分配方法

### 技术领域

[0001] 本发明涉及语音编码中的脉冲分配方法。

### 背景技术

[0002] 一般而言,在语音编码中,使用声道模型来合成尽可能类似于原始语音的语音信号。作为这种语音编码之一,有用于 3GPP 系统的自适应多速率宽带 (AMR-WB: Adaptive MultiRate-WideBand) 语音编码 (参照非专利文献 1)。该 AMR-WB 语音编码也是作为 ITU-T 建议 G. 722. 2 (非专利文献 2) 由 ITU-T 选定并承认的。下面,举例说明采用 23. 85kbps 的比特率的 AMR-WB 语音编码。

[0003] 作为 AMR-WB 语音编码的重要的结构元素之一,有固定码本 (图 1) 的搜索。在 AMR-WB 语音编码中,将具有下采样后的 256 个样本的语音样本的帧,分割成各自具有 64 个样本的 4 个子帧。在搜索固定码本的过程中,将该子帧分割成 4 个音轨。在 AMR-WB 语音编码的模式 8,从各个音轨可取得的 16 个脉冲位置中,对每个音轨选择 6 个脉冲位置。也就是说,各个子帧的脉冲的数目被设定为  $p_0 \sim p_{23}$  的 24 个。并且,对这些  $p_0 \sim p_{23}$  的 24 个脉冲的位置进行编码,而形成用于合成对应于各个子帧的语音的码本索引 (参照非专利文献 1)。

[0004] 另外,在 ITU-T 建议 G. 722. 2 中,目前虽然支持对单声道语音信号的 AMR-WB 语音编码,但没有支持对立体声语音信号的 AMR-WB 语音编码。

[0005] 另一方面,近年来,随着移动通信和 IP 通信中的传输频带的宽带化以及服务的多样化,在语音通信中,对高音质化和更强的现场感的需求日益增高。例如,可以预料今后对下述的服务的需求会增多,即,电视电话服务中的免提 (Handsfree) 形式的通话、在电视会议中的语音通信、在多个地点多个说话人同时进行会话的多地点语音通信、在保持现场感的同时能够传输周围的声音环境的语音通信等。那时,人们期待实现比单声道信号现场感更强并能够识别多个说话人的说话位置的使用立体声语音的语音通信的。为了实现这样的使用立体声语音的语音通信,必须对立体声语音进行编码。作为立体声语音信号的编码方法之一,有对各个声道的语音信号独立地进行编码的方法 (双重单声道编码)。

[0006] (非专利文献 1)“AMR Wideband Speech Codec ;General Description”,3GPP TS 26. 171, V5. 0. 0 (2001-03)

[0007] (非专利文献 2)“Wideband Coding of Speech at Around 16 kbit/s Using Adaptive Multi-Rate Wideband (AMR-WB)”, Geneva, ITU-T Recommendation G. 722. 2 (2003-07)

[0008] 发明内容

[0009] 本发明需要解决的问题

[0010] 然而,对立体声语音信号单纯地进行使用 AMR-WB 语音编码的双重单声道编码时,必须对各个声道的语音信号分别进行上述的固定码本的搜索,因此从编码效率和处理效率的观点而言,不够理想。

[0011] 本发明的目的是,提供能够对立体声语音信号高效率地进行编码的脉冲分配方法。

[0012] 解决问题的方案

[0013] 本发明的脉冲分配方法是,在对立体声信号的语音编码中,进行固定码本搜索时所使用的脉冲分配方法,其包括:判断立体声信号的各个声道间的相似程度的相似程度判断步骤;基于作为各个声道的特性的稳定度和周期性,判断声道是否为稳定有声的稳定有声判断步骤;以及根据所述相似程度判断步骤和所述稳定有声判断步骤的判断结果,决定分配给各个声道的脉冲数的分配步骤。

[0014] 本发明的有益效果

[0015] 根据本发明,能够对立体声语音信号高效率地进行编码。

[0016] 附图说明

[0017] 图 1 是 AMR-WB 语音编码的固定码本。

[0018] 图 2 是本发明实施方式 1 的语音编码的处理流程。

[0019] 图 3 是本发明实施方式 1 的固定码本搜索的主要的处理流程。

[0020] 图 4 是本发明实施方式 1 的固定码本搜索的详细处理流程。

[0021] 图 5 是本发明实施方式 1 的脉冲分配的一个例子。

[0022] 图 6 是本发明实施方式 1 的脉冲分配的一个例子。

[0023] 图 7 是本发明实施方式 1 的通知例。

[0024] 图 8 是本发明实施方式 1 的语音解码的处理流程。

[0025] 图 9 是本发明实施方式 2 的通知例。

[0026] 图 10 是本发明实施方式 2 的语音解码的处理流程。

## 具体实施方式

[0027] 下面,参照附图详细地说明本发明的实施方式。另外,在以下说明中,举一例说明 AMR-WB 语音编码。另外,在以下的说明中,说明采用 AMR-WB 语音编码的编码模式中的模式 8 的情况下的实施方式,但各个实施方式同样可适用于其它编码模式。

[0028] 在 AMR-WB 语音编码的模式 8 中,固定码本矢量(创新矢量)包括 24 个脉冲。如上述图 1 所示,在各个子帧中有 0~63 的 64 个可取的脉冲位置,这些脉冲位置,以各个音轨包含 6 个脉冲的方式,被分割成 1~4 的 4 个音轨。

[0029] (实施方式 1)

[0030] 在本实施方式,基于输入立体声信号的各个声道间的相似程度,以及各个声道的周期性和稳定度,决定对各个声道的脉冲数的分配。基于相似程度、周期性和稳定度,对各个声道分配所需数目的脉冲。对各个声道的分配脉冲数决定后,进行与一般的 AMR-WB 语音编码相同的脉冲搜索,决定对各个声道的脉冲位置。将这些脉冲作为码字组进行编码,并作为语音比特流中的参数之一的码本索引发送。

[0031] 图 2 表示本实施方式的语音编码的主要的处理流程。

[0032] 首先,在 ST(步骤)11,对立体声信号进行预处理,该预处理包括下采样和使用高通滤波器和预加重滤波器的滤波处理。

[0033] 在 ST12,对预处理后的立体声信号进行 LPC 分析,获得立体声信号的 L 声道(左

声道)和R声道(右声道)的各个声道的LPC参数。将LPC参数变换成ISP(Immittance Spectrum Pair)和对各个声道的量化矢量。

[0034] 在ST13,对各个声道进行基音搜索,对各个帧估两次计开环基音延迟。

[0035] 在ST14,使用估计出的基音延迟(估计基音延迟),对于各个子帧,对估计基音延迟的周边进行使用闭环基音的自适应码本搜索,获得自适应码本矢量。

[0036] 在ST15,使用自适应码本矢量,进行伴随脉冲分配的固定码本搜索,获得对各个声道的固定码本矢量。

[0037] 然后,在ST16,为了下一个子帧的运算,更新滤波器记忆和样本数据。

[0038] 另外,在本实施方式中,除图2所示的处理外,与上述非专利文献1所记载的处理相同。

[0039] 接着,图3表示固定码本搜索(ST15)的主要的处理流程。主要通过ST21~ST25的处理来进行固定码本搜索(ST15)。

[0040] 在ST21,对各个子帧进行立体声信号的比较,而判断L声道和R声道的相似性,即声道间的相似程度。

[0041] 在ST22,进行立体声信号的分类,判断信号的特性。

[0042] 在ST23,基于声道间的相似程度和立体声信号的特性,对L声道和R声道,各自分配所需数目的脉冲。

[0043] 在ST24,执行AMR-WB语音编码的脉冲搜索,决定对各个声道的脉冲位置。

[0044] 在ST25,将在ST24所决定的脉冲,作为码字组进行编码,并作为语音比特流中的参数之一的码本索引,发送到语音解码装置。

[0045] 下面,使用图4更详细地说明图3所示的处理流程。特别对脉冲分配(ST23)进行详细的说明。

[0046] 在ST301,对各个子帧的L声道和R声道进行比较。通过该比较,在进行脉冲分配之前,判断两个声道间的信号特性的相似程度(在两个声道间,有多少程度的相似性)。在判断相似程度时,可利用相互相关、在时域中的信号包络的比较、在频域中的频谱信号或频谱能量的比较以及中侧(Mid-Side)运算等。

[0047] 在ST302,在L声道和R声道非常相似时(例如,相互相关值大于阈值时),或者在L声道和R声道相同时(也就是单声道信号时),双方声道使用共同的脉冲组。也就是说,在ST303,将对L声道的脉冲数Num\_Pulse(L)设定为P,将对R声道的脉冲数Num\_Pulse(R)设定为0。或者相反地,将对L声道的脉冲数Num\_Pulse(L)设定为0,将对R声道的脉冲数Num\_Pulse(R)设定为P。例如,AMR-WB语音编码的模式8时,设定为 $P = 24$ 。图5A表示在ST303设定Num\_Pulse的情况。在此例中, $P = 24$ 。因为对L声道或R声道的任意一方分配全部的,24个脉冲,所以如图6A所示,对双方声道使用 $p_0 \sim p_{23}$ 的共同的一个脉冲组。以下将图6A所示的脉冲分配的类型称为“类型0”。

[0048] 在ST302,在L声道和R声道不相似时(例如,相互相关值为阈值以下时),在ST304进行信号的分类判断,并确认在L声道或R声道中是否存在“稳定有声”信号。L声道信号或R声道信号在具有周期性且稳定时,被判定为“稳定有声”,而在不具有周期性且不稳定时,被判定为其它类型的信号。在L声道或R声道的任何一方为“稳定有声”时,进至ST305,而在L声道和R声道都不是“稳定有声”时,进至ST310。另外,在判断信号是否为“稳定有声”

时,可利用基于自相关方法的自相关值运算、基音预测增益、自适应码本增益等等。另外,也可使用各个声道的能量电平或信号电平等,来判断是否为“稳定有声”。

[0049] 在 ST305,如果 L 声道和 R 声道都被判断为“稳定有声”(具有稳定性且周期性的),则该双方声道各自具有脉冲组。也就是说,此时,在 ST306,将 P 个 ( $P = 24$ ) 脉冲分配给 L 声道和 R 声道,将对 L 声道的脉冲数 Num\_Pulse(L) 设定为脉冲数  $K_1P$ ,而将对 R 声道的脉冲数 Num\_Pulse(R) 设定为脉冲数  $(1-K_1)P$ 。这里,假设  $K_1 = 1/2$ 。也就是说,对双方声道分配相等数目的脉冲。图 5B 表示在 ST306 设定 Num\_Pulse 的情况。因为以图 5B 所示的方式设定 Num\_Pulse,所以将  $P = 24$  个脉冲均等地对两个声道各分配 12 个。因此,如图 6B 所示,对各个声道使用不同的脉冲组。但是,各个脉冲组所包含的脉冲数是同数(这里是 12 个)。以下将图 6B 所示的脉冲分配的类型称为“类型 1”。

[0050] 另外,在图 6B 中,脉冲  $P_{ch,i}$  的 ch 表示该脉冲所属的声道(L 声道或 R 声道),i 表示该脉冲的位置。在图 6C 和图 6D 中也是一样。

[0051] 在 ST305,如果判断为一方的声道是“稳定有声”,但另一方的声道不是“稳定有声”,则不在双方声道间均等地分配脉冲数 P。此时,脉冲数的分配,基于哪一方的声道需要更多的脉冲而进行。一般而言,“稳定有声”的声道只需要较少数目的脉冲,因此,对“稳定有声”的声道分配较少数目的脉冲。这是因为,对于“稳定有声”的声道,自适应码本能够有效地发挥功能而生成声源信号,因而固定码本搜索只需要较少数目的脉冲。

[0052] 也就是说,在 ST307,在判断 L 声道是“稳定有声”,R 声道不是“稳定有声”时,在 L 声道只需要较少数目的脉冲,因此,与 R 声道相比,对 L 声道分配较少数目的脉冲。即,在 ST308,将对 L 声道的脉冲数 Num\_Pulse(L) 设定为脉冲数  $K_2P$ ,而将对 R 声道的脉冲数 Num\_Pulse(R) 设定为脉冲数  $(1-K_2)P$ ,将 P 个 ( $P = 24$ ) 脉冲分配给 L 声道和 R 声道。这里,假设  $K_2 = 1/3$ 。由此,对 L 声道分配 8 个脉冲,对 R 声道分配 16 个脉冲,与 R 声道相比,对 L 声道分配较少数目的脉冲。

[0053] 另一方面,在 ST307,在判断 L 声道不是“稳定有声”,而 R 声道是“稳定有声”时,与 L 声道相比,对 R 声道分配较少数目的脉冲。即,在 ST309,将对 L 声道的脉冲数 Num\_Pulse(L) 设定为脉冲数  $(1-K_2)P$ ,而将对 R 声道的脉冲数 Num\_Pulse(R) 设定为脉冲数  $K_2P$ ,对 L 声道和 R 声道分配 P 个 ( $P = 24$ ) 脉冲。这里,与上述同样地,假设  $K_2 = 1/3$ 。由此,对 R 声道分配 8 个脉冲,对 L 声道分配 16 个脉冲,与 L 声道相比,对 R 声道分配较少数目的脉冲。

[0054] 图 5C 和图 5D 表示在 ST308 和 ST309,设定 Num\_Pulse 的情况。因为  $P = 24$  且  $K_2 = 1/3$ ,Num\_Pulse 成为 8 个脉冲(图 5C)和 16 个脉冲(图 5D)。因此,如图 6C 和图 6D 所示,对各个声道使用脉冲数相互不同的两个不同脉冲组。以下将图 6C 所示的脉冲分配的类型称为“类型 2”,将图 6D 所示的脉冲分配的类型称为“类型 3”。在类型 2,与 R 声道相比,对 L 声道分配较少数目的脉冲,在类型 3,与 L 声道相比,对 R 声道分配较少数目的脉冲。这样,在类型 2 和类型 3,对 L 声道和 R 声道不均等地分配 24 个脉冲。

[0055] 在 ST304,如果 L 声道和 R 声道都不是“稳定有声”,则基于各个声道的最大自相关系数(MAF:Maximum Autocorrelation Factor)来决定脉冲分配。通过式(1)来定义 MAF。在式(1)中, $x(n)$  ( $n = 0, \dots, N-1$ ) 是对 L 声道或 R 声道的编码对象子帧的 MAF 的计算对象区间的输入信号,N 是该计算对象区间的区间长度(样本数), $\tau$  是延迟。另外,作为  $x(n)$ ,

可使用 LPC 残差信号以代替输入信号,该 LPC 残差信号使用 LPC 逆滤波器而求得的。

$$[0056] \quad C = \frac{\max \left\{ \sum_{n=0}^{N-1} x(n)x(n-\tau) \right\}}{\sum_{n=0}^{N-1} x^2(n)} \dots \text{式 (1)}$$

[0057] 在 ST310,在 L 声道的 MAF 大于 R 声道的 MAF 时,在 ST312,与 ST308 同样地,将对 L 声道的脉冲数 Num\_Pulse(L) 设定为脉冲数  $K_2P$ ,而将对 R 声道的脉冲数 Num\_Pulse(R) 设定为脉冲数  $(1-K_2)P$ ,对 L 声道和 R 声道分配 P 个 ( $P = 24$ ) 脉冲。这里,假设  $K_2 = 1/3$ 。也就是说,对 L 声道分配 8 个脉冲,对 R 声道分配 16 个脉冲,与 R 声道相比,对 L 声道分配较少数目的脉冲。因此,脉冲分配的类型为类型 2(图 6C)。

[0058] 另一方面,在 ST310, R 声道的 MAF 为 L 声道的 MAF 以上时,在 ST311,与 ST309 同样地,将对 R 声道的脉冲数 Num\_Pulse(R) 设定为脉冲数  $K_2P$ ,而将对 L 声道的脉冲数 Num\_Pulse(L) 设定为脉冲数  $(1-K_2)P$ ,对 L 声道和 R 声道分配 P 个 ( $P = 24$ ) 脉冲。这里,与上述同样地,假设  $K_2 = 1/3$ 。也就是说,对 R 声道分配 8 个脉冲,对 L 声道分配 16 个脉冲,与 L 声道相比,对 R 声道分配较少数目的脉冲。因此,脉冲分配的类型为类型 3(图 6D)。

[0059] 在 ST303、ST306、ST308、ST309、ST311 和 ST312,如上述决定对各个声道分配的脉冲数后,在 ST313,对各个声道进行脉冲位置的搜索。

[0060] 然后,在搜索出 L 声道和 R 声道的两个声道的脉冲位置后,在 ST314,使用在 ST313 所搜索出的脉冲来生成码字组,在 ST315 生成对各个声道的码本索引。

[0061] 另外,在 ST304 中,对于 L 声道和 R 声道都不是“稳定有声”时的脉冲分配,也可以不是如上述这样基于各个声道的 MAF 决定,而是对各个声道总是分配相等数目的脉冲。

[0062] 这里,如上述在将  $K_1$  和  $K_2$  设为固定值时,基于脉冲分配的 4 个类型(类型 0 ~ 3)唯一地决定分配给各个声道的脉冲数,因此向语音解码端通知对各个声道分配了的脉冲数时,如图 7 所示,有两个比特就足够。也就是说,将类型 0(对 L 声道和 R 声道共同地分配 24 个脉冲的情况)设为码字‘00’,将类型 1(对 L 声道和 R 声道各分配 12 个脉冲的情况)设为码字‘01’,将类型 2(对 L 声道分配 8 个脉冲,对 R 声道分配 16 个脉冲的情况)设为码字‘10’,将类型 3(对 L 声道分配 16 个脉冲,对 R 声道分配 8 个脉冲的情况)设为码字‘11’,并通知给语音解码端。

[0063] 图 8 表示在语音解码端的处理流程。

[0064] 在 ST701,从比特流提取处于对脉冲数据进行了量化的状态的码本索引。另外,从比特流提取表示脉冲分配的类型上述 2 比特的信息。

[0065] 在 ST702,基于上述 2 比特的信息,并参照上述图 7 所示的表,判断脉冲分配的类型。

[0066] 在 ST703,在脉冲分配的类型是类型 0 时,进至 ST704,而在是类型 1 ~ 3 时,进至 ST707。

[0067] 在脉冲分配的类型是类型 0 时,由两个声道共同使用同一码本。也就是说,在 ST704,将全部的  $P = 24$  个脉冲都设定给预先决定的一方的声道(规定声道),并在 ST705,对于该规定声道,将  $P = 24$  个脉冲解码。然后,在 ST706,将在 ST705 所解码的脉冲复制到另一方的声道。

[0068] 另一方面,在脉冲分配的类型是类型 1 ~ 3 时,基于各个类型,设定对各个声道的脉冲数。即,在类型 1 时,对 L 声道和 R 声道分别设定 12 个脉冲,在类型 2 时,对 L 声道设定 8 个脉冲,对 R 声道设定 16 个脉冲,在类型 3 时,对 L 声道设定 16 个脉冲,对 R 声道设定 8 个脉冲。

[0069] 假设规定声道是 L 声道,则在 ST707 设定对 L 声道的脉冲数  $P_L$ ,在 ST708 设定对 R 声道的脉冲数  $P_R$ 。然后,在 ST709,将  $P_L$  个脉冲作为对 L 声道的码本数据解码,在 ST710,将  $P_R$  个脉冲作为对 R 声道的码本数据解码。

[0070] 另外,在规定声道为 R 声道时,处理流程成为 ST708、ST707、ST710、ST709 的次序。

[0071] 这样,根据本实施方式,基于声道间的相似程度和各个声道的特性(周期性和稳定度)而决定脉冲数的分配,因此能够对各个声道分配最适合的数目的脉冲。

[0072] (实施方式 2)

[0073] 在本实施方式中,基于语音信号的特性而决定  $K_1$ 、 $K_2$ ,并使各个声道间的脉冲分配自适应地变化。例如,能够基于各个声道的语音信号的周期性和 MAF 而求对各个声道的脉冲数的分配比例。

[0074] 例如,在 L 声道和 R 声道的双方都是“稳定有声”时,通过式 (2) 来求  $K_1$ 。

$$[0075] \quad K_1 = \alpha_1 \frac{\tau_R}{\tau_L + \tau_R} \dots \text{式 (2)}$$

[0076] 在式 (2) 中,  $\tau_L$ 、 $\tau_R$  分别是 L 声道的基音周期和 R 声道的基音周期,  $\alpha_1$  是用于微调  $K_1$  的系数。根据式 (2),能够对具有越短的基音周期的,即越高的基音的声道,分配越多的脉冲。

[0077] 另外,在一方的声道是“稳定有声”,另一方的声道不是“稳定有声”时,通过式 (3) 来求  $K_2$ 。

$$[0078] \quad K_2 = \beta + \alpha_2 \frac{C_{uv}}{C_L + C_R} \dots \text{式 (3)}$$

[0079] 在式 (3) 中,  $C_{uv}$  为不是“稳定有声”的一方的声道的 MAF,  $C_L$ 、 $C_R$  分别是 L 声道的 MAF 和 R 声道的 MAF,  $\alpha_2$  是用于微调  $K_2$  的系数。根据式 (3),能够对“稳定有声”的声道,分配较少数目的脉冲。

[0080] 另外,在式 (3) 中,  $\beta$  是使“稳定有声”的声道确实具有至少为最小数的脉冲的参数,通过式 (4) 来定义。

$$[0081] \quad \beta = \text{ceiling} \left( \frac{L}{\tau_{ch}} \right) \times \frac{1}{P} \dots \text{式 (4)}$$

[0082] 在式 (4) 中,  $L$  是一个帧内的样本数,  $\tau_{ch}$  是“稳定有声”的声道的基音周期,  $P$  是子帧内的总脉冲数。因此,比例  $L/\tau_{ch}$  基本上是一个帧内的周期数。例如,假设  $L$  为 256、 $\tau_{ch}$  为 77,则比例  $L/\tau_{ch}$  (一个帧内的周期数) 成为 4。由此,在各个基音周期确实存在至少一个脉冲。

[0083] 基于式 (2) ~ (4) 所求出的  $K_1$  和  $K_2$ ,用于对 L 声道和 R 声道的脉冲数的分配。被分配到 L 声道和 R 声道的脉冲数可取满足式 (5) 和式 (6) 的条件最小值 MIN\_PULSE 和最大值 MAX\_PULSE。

$$[0084] \quad \text{MIN\_PULSE} \leq \text{Num\_Pulse}(\text{channel}) \leq \text{MAX\_PULSE} \quad \dots \text{式 (5)}$$



[0085]  $\text{Num\_Pulse(L)} + \text{Num\_Pulse(R)} = \text{TOTAL\_PULSE}$  ... 式 (6)

[0086] 在式 (5) 和 (6) 中, MIN\_PULSE 和 MAX\_PULSE 是对特定的声道的每个子帧所能分配的脉冲数的最小值和最大值, TOTAL\_PULSE 是对双方声道的每个子帧所能分配的合计的脉冲数。例如,典型的值是: MIN\_PULSE = 4、MAX\_PULSE = 20、TOTAL\_PULSE = 24。另外,还可以将所计算出的脉冲数取整为 1、2 或 4 的倍数中,最接近的倍数值。

[0087] 这里,在使分配给各个声道的脉冲数自适应地变化时,必须将对各个声道分配了的脉冲数通知给语音解码端。但是,对一方的声道的分配数,通过从两个声道的所有脉冲数减去另一方的声道的分配数而能够导出,因此,将任意一方的声道定为规定声道,并仅通知该规定声道的分配数即可。例如,将 L 声道定为规定声道,通知对 L 声道的脉冲数 Num\_Pulse(L),并通过式 (7) 来求对 R 声道的脉冲数 Num\_Pulse(R)。

[0088]  $\text{Num\_Pulse(R)} = \text{TOTAL\_PULSE} - \text{Num\_Pulse(L)}$  ... 式 (7)

[0089] 以下示出对规定声道的脉冲数的通知方法。

[0090] 在对各个声道的脉冲数为 4 的倍数时,对规定声道的脉冲数有 4 个脉冲、8 个脉冲、12 个脉冲、16 个脉冲、20 个脉冲的 5 个可能性。因此,为了区别这 5 个脉冲数,有 3 比特就足够。另外,在对各个声道的脉冲数为 2 的倍数时,对规定声道的脉冲数有 4 个脉冲、6 个脉冲、8 个脉冲、10 个脉冲、12 个脉冲、14 个脉冲、16 个脉冲、18 个脉冲、20 个脉冲的 9 个可能性。因此,为了区别这些 9 个脉冲数,需要 4 比特。再者,对各个声道的脉冲数为从 4 到 20 的按一个个脉冲递增的数目时,有 17 个可能性,因此为了区别 17 个脉冲数,需要 5 比特。由此,能够将这些脉冲数总结为图 9 所示的表。在语音编码端,参照该表将脉冲数变换成 3 ~ 5 比特的码字并通知,在语音解码端,同样地参照该表,从所通知的码字导出对各个声道分配的脉冲数。

[0091] 图 10 表示在语音解码端的处理流程。

[0092] 在 ST901,从比特流提取处于对脉冲数据进行了量化的状态的码本索引。另外,从比特流提取表示脉冲数的码字 (3 ~ 5 比特)。

[0093] 在 ST902,基于表示脉冲数的码字,并参照上述图 9 所示的表,判断规定声道的脉冲数。这里,假设规定声道是 L 声道。

[0094] 在 ST903,通过式 (7) 来计算另一方的声道,即 R 声道的脉冲数。

[0095] 在 ST904,在任何一方的声道的脉冲数为 0 时,进至 ST905,在除此之外的情况下,进至 ST907。

[0096] 在任意一方的声道的脉冲数为 0 时,由两个声道共同使用同一码本。也就是说,在 ST905,将全部  $P = 24$  个脉冲都设定给规定声道,并对于该规定声道,将  $P = 24$  个脉冲解码。然后,在 ST906,将在 ST905 所解码的脉冲复制到另一方的声道。

[0097] 另一方面,在 ST907,参照上述图 9 所示的表设定对 L 声道 (规定声道) 的脉冲数  $P_L$ ,将  $P_L$  个脉冲作为对 L 声道的码本数据进行解码。并且,在 ST908,通过式 (7) 设定对 R 声道的脉冲数  $P_R$ ,将  $P_R$  个脉冲作为对 R 声道的码本数据解码。

[0098] 另外,在规定声道是 R 声道时,处理流程成为 ST908、ST907 的次序。

[0099] 这样,根据本实施方式,基于语音信号的特性而决定  $K_1$  和  $K_2$ ,使各个声道间的脉冲分配自适应地变化,因此,能够在声道间进行更加灵活且正确的脉冲数的分配。

[0100] 另外,在上述各个实施方式中,假设对各个声道分配的合计的脉冲数为固定 (在

上述各个实施方式中,固定为  $P = 24$  个) 而进行说明,但也可以使对各个声道分配的合计的脉冲数,基于声道间的相似程度和各个声道的特性(周期性和稳定度)而变化。例如,在实施方式 1,在脉冲分配的类型为“类型 0”时,即在 L 声道和 R 声道非常相似时(例如,在相互相关值大于阈值时),或者,在 L 声道和 R 声道相同时(也就是单声道信号时),可以将少于其它类型的合计的分配脉冲数(在上述各个实施方式中, $P = 24$  个)的数目的脉冲,只分配给 R 声道或 L 声道的任意一方。由此能够更加提高传输效率。

[0101] 另外,上述各个实施方式的处理流程,可在语音编码装置和语音解码装置中实施。另外,还可以将该语音编码装置和语音解码装置,装载于在移动通信系统中所使用的无线通信移动台装置和无线通信基站装置等无线通信装置。

[0102] 另外,上述实施方式的处理流程,通常被作为集成电路的 LSI 来实现。这些既可以被单独地集成为一个芯片,也可以一部分或全部被集成为一个芯片。

[0103] 虽然此处称为 LSI,但根据集成程度,可以被称为 IC、系统 LSI、超大 LSI (Super LSI)、或特大 LSI (Ultra LSI)。

[0104] 另外,实现集成电路化的方法不仅限于 LSI,也可使用专用电路或通用处理器来实现。也可以使用在 LSI 制造后,可编程的 FPGA (Field Programmable Gate Array),或者可重构 LSI 内部的电路单元的连接和设定的可重构处理器。

[0105] 再者,随着半导体的技术进步或随之派生的其它技术的出现,如果能够出现替代 LSI 集成电路化的新技术,当然可利用新技术进行功能块的集成化。还存在着适用生物技术等的可能性。

[0106] 本说明书是根据 2005 年 2 月 10 日申请的日本专利申请第 2005-034984 号。其内容全部包含于此。

[0107] 工业实用性

[0108] 本发明可适用于移动通信系统或使用因特网协议的分组通信系统等中的通信装置。

音轨	脉冲位置	子帧内的有效脉冲位置
1	P <sub>0</sub> ,P <sub>4</sub> ,P <sub>8</sub> ,P <sub>12</sub> ,P <sub>16</sub> ,P <sub>20</sub>	0,4,8,12,16,20,24,28,32,36,40,44,48,52,56,60
2	P <sub>1</sub> ,P <sub>5</sub> ,P <sub>9</sub> ,P <sub>13</sub> ,P <sub>17</sub> ,P <sub>21</sub>	1,5,9,13,17,21,25,29,33,37,41,45,49,53,57,61
3	P <sub>2</sub> ,P <sub>6</sub> ,P <sub>10</sub> ,P <sub>14</sub> ,P <sub>18</sub> ,P <sub>22</sub>	2,6,10,14,18,22,26,30,34,38,42,46,50,54,58,62
4	P <sub>3</sub> ,P <sub>7</sub> ,P <sub>11</sub> ,P <sub>15</sub> ,P <sub>19</sub> ,P <sub>23</sub>	3,7,11,15,19,23,27,31,35,39,43,47,51,55,59,63

图 1

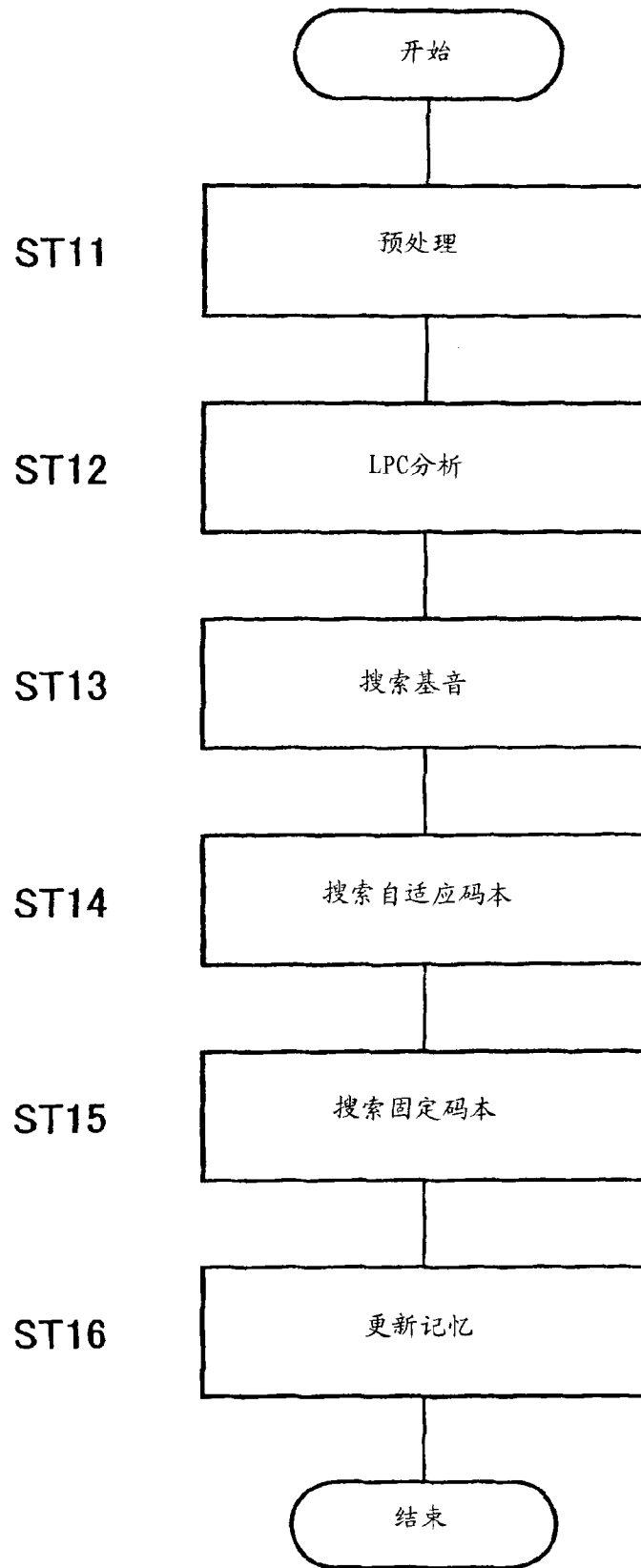


图 2

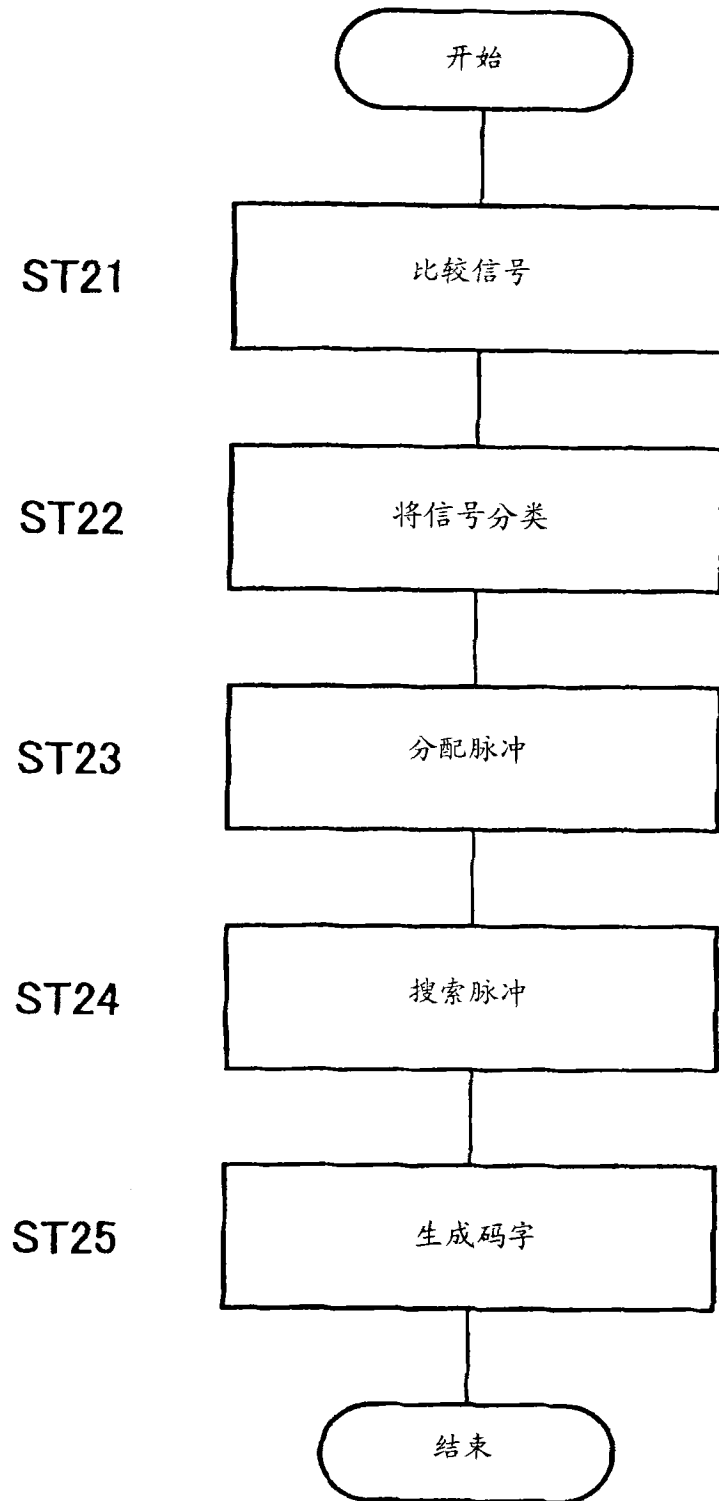


图 3

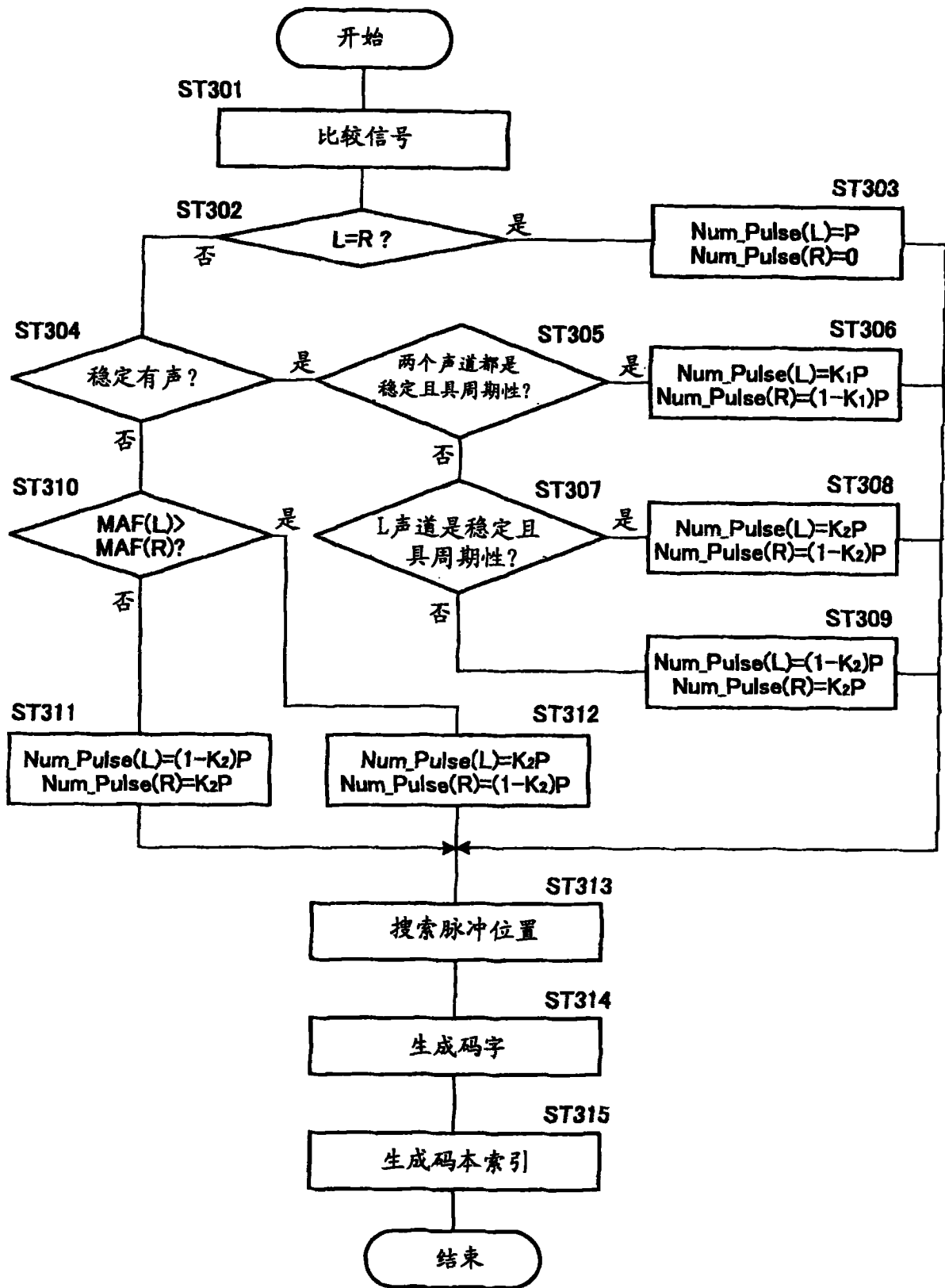
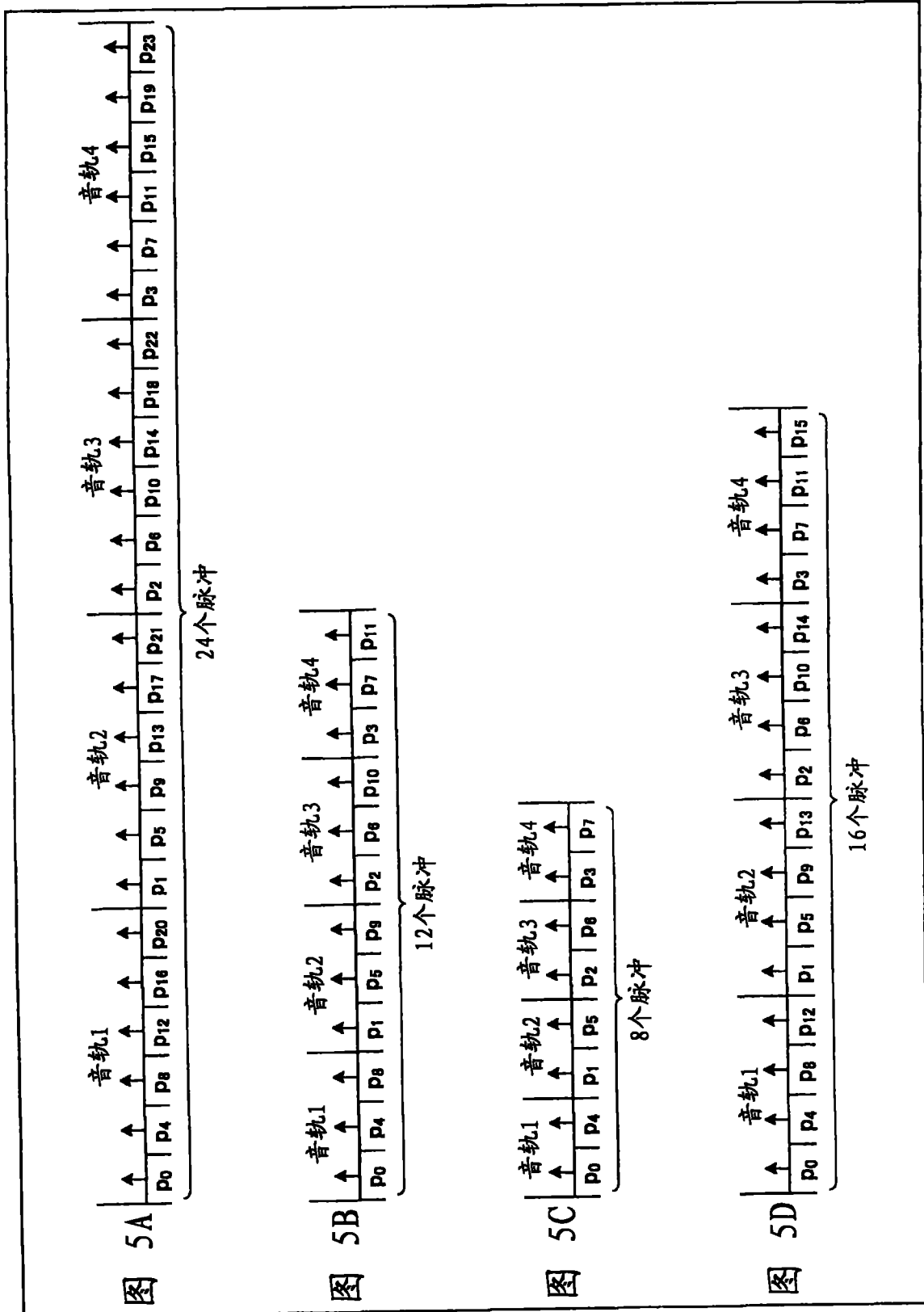
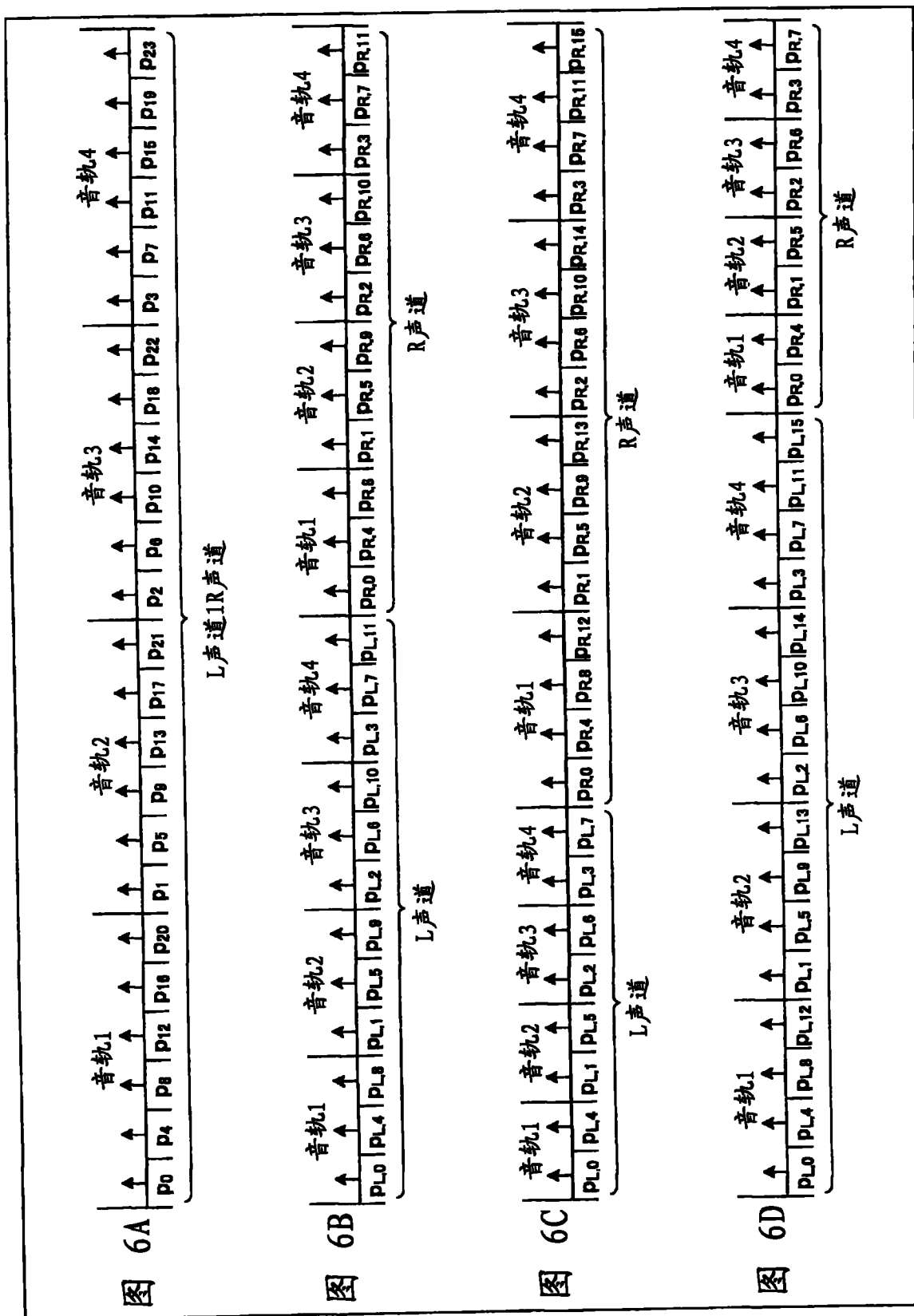


图 4







类型	比特1	比特2	脉冲分配
0	0	0	L声道/R声道: 24个脉冲
1	0	1	L声道: 12个脉冲、R声道: 12个脉冲
2	1	0	L声道: 8个脉冲、R声道: 16个脉冲
3	1	1	L声道: 16个脉冲、R声道: 8脉冲

图 7

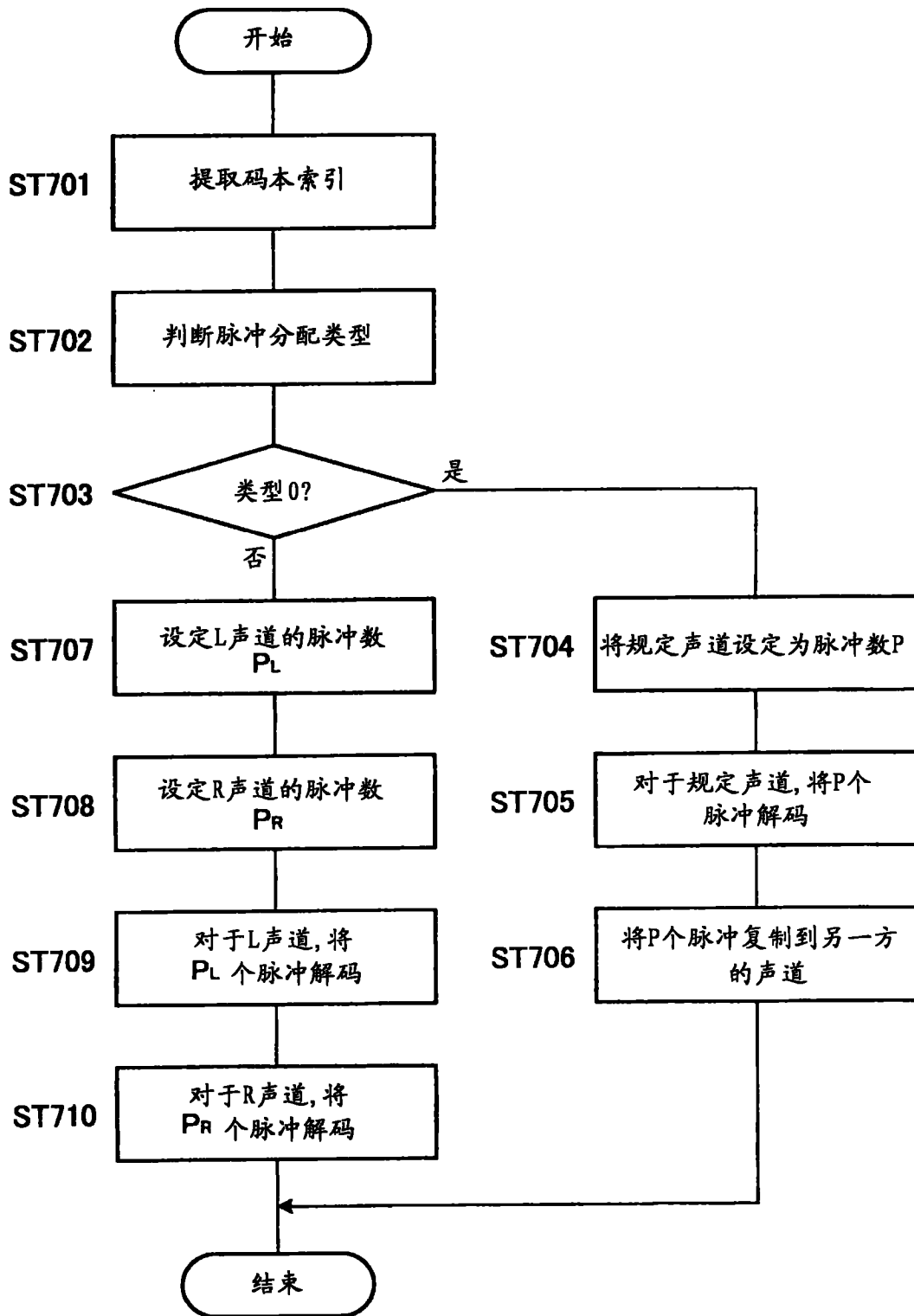


图 8

脉冲数	3比特	4比特	5比特
4	000	0000	00000
5			00001
6		0001	00010
7			00011
8	001	0010	00100
9			00101
10		0011	00110
11			00111
12	010	0100	01000
13			01001
14		0101	01010
15			01011
16	011	0110	01100
17			01101
18		0111	01110
19			01111
20	100	1000	10000

图 9

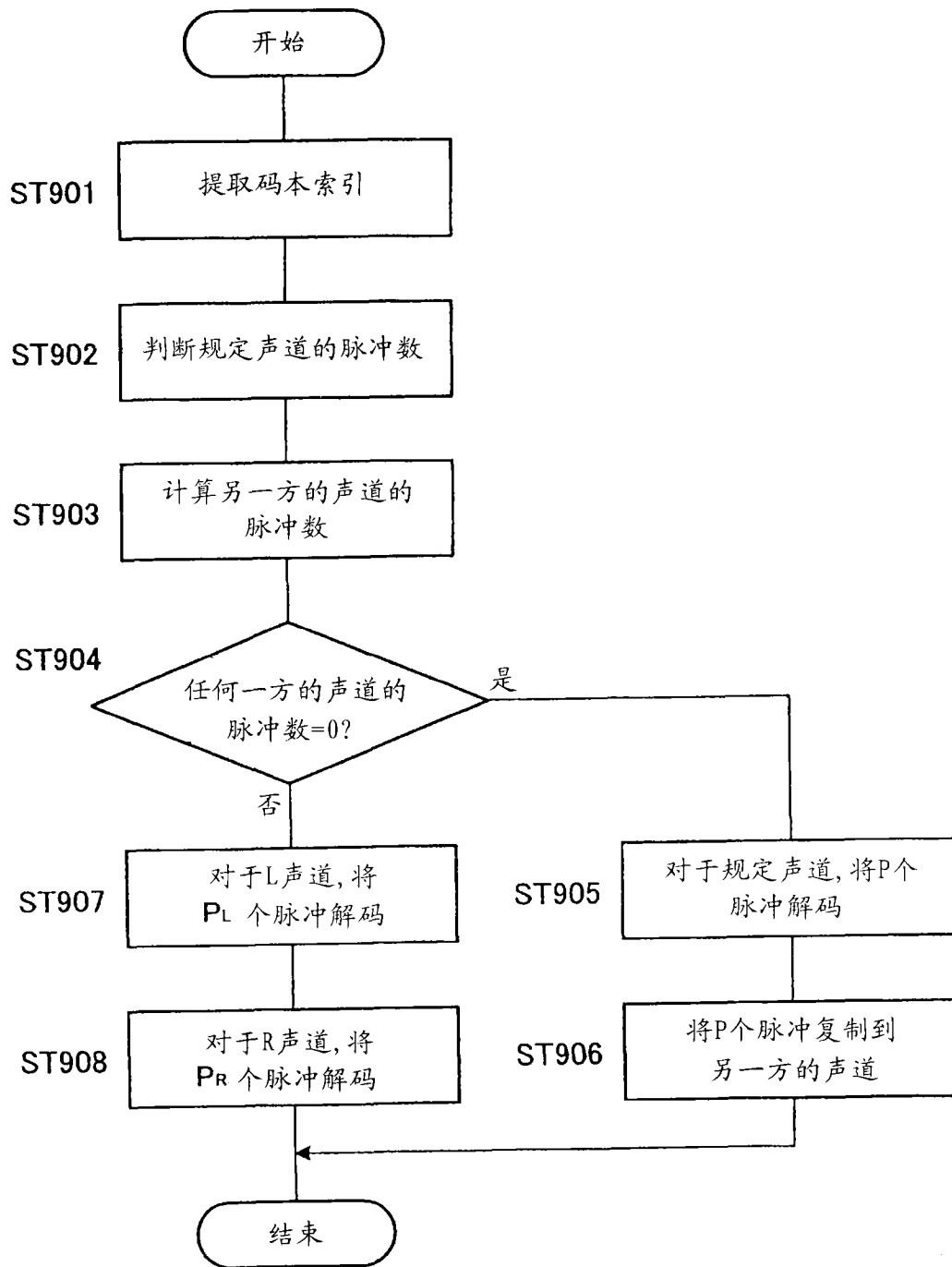


图 10