

(12) 发明专利

(10) 授权公告号 CN 1906663 B

(45) 授权公告日 2010.06.02

(21) 申请号 200580001834.3

(51) Int. Cl.

(22) 申请日 2005.05.10

G10L 19/00 (2006.01)

(30) 优先权数据

140117/2004 2004.05.10 JP

(56) 对比文件

JP 2004-80625 A, 2004.03.11, 全文.

US 20030200083 A1, 2003.11.07, 全文.

(85) PCT申请进入国家阶段日

2006.06.30

审查员 刘红梅

(86) PCT申请的申请数据

PCT/JP2005/008495 2005.05.10

(87) PCT申请的公布数据

W02005/109401 JA 2005.11.17

(73) 专利权人 日本电信电话株式会社

地址 日本东京都

(72) 发明人 大室伸 森岳至 日和崎佑介

片冈章俊

(74) 专利代理机构 北京市柳沈律师事务所

11105

代理人 邸万奎 黄小临

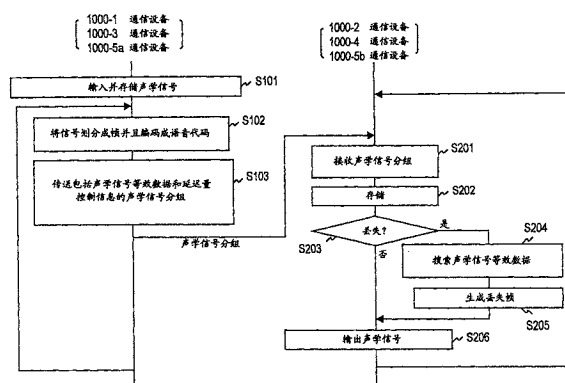
权利要求书 3 页 说明书 18 页 附图 32 页

(54) 发明名称

声学信号分组通信方法、传送方法、接收方法、及其设备和程序

(57) 摘要

当通过 IP 通信网络传递声学信号分组时,将与声学信号相对应的数据(声学信号对应数据)内置在不同于该声学信号的分组中,并对其进行传送。然而,在传统上,必须预先确定该声学信号内置在哪个分组中,并且不可能执行动态修改。在本发明中,将声学信号的声学信号对应数据的延迟量作为延迟量控制信息内置在声学信号分组中。此外,在突发丢失中丢失了通信网络状态的分组数检测波动,并且获取要在接收侧累积的分组数以及延迟量。



1. 一种在包括传送单元和接收单元两者的一个通信设备与一个或更多其它通信设备之间的声学信号分组传送方法,该方法包括:

在传送单元中,

将声学信号划分成被称为帧的给定时间片段,以生成帧声学信号的步骤;以及
从帧声学信号生成作为与帧声学信号相对应的数据的声学信号对应数据;以及
将帧声学信号和声学信号对应数据包含在分组中,并且传送该分组的包含步骤,
其特征在于,

所述包含步骤还包括指示在包含帧声学信号的相同分组中帧编号差异的延迟量控制信息,以及

声学信号对应数据是对于具有相差由延迟量控制信息指示的值的帧编号的帧的数据,以及

所述声学信号分组传送方法还包括,

在接收单元中:

确定所接收分组的抖动状态的第一确定步骤、以及确定所接收分组的丢失状态的第二确定步骤中的两者或一个;以及

使用在确定步骤中做出的确定的结果,将要存储在接收缓冲器中的分组数确定为存储分组数的目标值的步骤;

在所述传送单元中,

将延迟量控制信息设置为小于或等于所存储的分组数的目标值的步骤,该目标值在包括传送单元的相同通信设备中的接收单元处确定。

2. 一种在包括传送单元和接收单元两者的第一通信设备与包括传送单元和接收单元两者的第二通信设备之间的声学信号分组通信方法,该声学信号分组通信方法包括:

在第一通信设备的传送单元中,

将声学信号划分成被称为帧的给定时间片段,以生成帧声学信号的步骤;以及
从帧声学信号生成作为与帧声学信号相对应的数据的声学信号对应数据;以及
将帧声学信号和声学信号对应数据包含在分组中,并且传送该分组的包含步骤,其特征在于,

所述包含步骤还包括指示在包含帧声学信号的相同分组中帧编号差异的延迟量控制信息,

声学信号对应数据是对于具有相差由延迟量控制信息指示的值的帧编号的帧的数据,以及

所述声学信号分组通信方法包括:

在第二通信设备的接收单元中:

确定所接收分组的抖动状态的第一确定步骤、以及确定所接收分组的丢失状态的第二确定步骤中的两者或一个;以及

使用在确定步骤中做出的确定结果,以确定要存储在接收缓冲器中的分组数作为存储分组数的目标值的步骤;以及

将存储分组数的目标值发送到第二通信设备中的传送单元的步骤;

在第二通信设备的传送单元中,

将从第二通信设备中的接收单元发送的存储分组数的目标值包含在分组中、作为用于指定要在第一通信设备的传送单元中设置的延迟量控制信息的信息的步骤；以及

在第一通信设备的传送单元中，

将延迟量控制信息设置为小于或等于所存储的分组数的目标值的步骤，其中所述目标值由第二通信设备中的传送单元发送，并被包含在分组中。

3. 一种在包括传送单元和接收单元两者的第一通信设备与包括传送单元和接收单元两者的第二通信设备之间的声学信号分组通信方法，该声学信号分组通信方法包括：

在第一通信设备的传送单元中，

将声学信号划分成被称为帧的给定时间片段，以生成帧声学信号的步骤，以及

从帧声学信号生成作为与帧声学信号相对应的数据的声学信号对应数据；以及

将帧声学信号和声学信号对应数据包含在分组中，并且传送该分组的包含步骤，其特征在于，

所述包含步骤还包括指示在包含帧声学信号的相同分组中帧编号差异的延迟量控制信息，

声学信号对应数据是对于具有相差由延迟量控制信息指示的值的帧编号的帧的数据，以及

所述声学信号分组通信方法还包括：

在第二通信设备的接收单元中，

测量存储在接收缓冲器中的分组数作为剩余缓冲器量的步骤；以及

将剩余缓冲器量发送到第二通信设备中的传送单元的步骤；

在第二通信设备的传送单元中，

将从第二通信设备的接收单元发送的剩余缓冲器量包含在分组中，作为用于指定要在第一通信设备的传送单元中设置的延迟量控制信息的信息，并传送该分组的步骤；以及

在第一通信设备的传送单元中，

将延迟量控制信息设置为小于或等于剩余缓冲器量的步骤，其中所述剩余缓冲器量由第二通信设备中的传送单元发送，并被包含在分组中。

4. 一种声学信号分组通信设备，包括：

传送单元，其具有：

用于将声学信号划分成被称为帧的给定时间片段、以生成帧声学信号的部件；以及

用于从帧声学信号生成与帧声学信号相对应的数据作为声学信号对应数据的部件；

用于将帧声学信号和声学信号对应数据包含在分组中的包含部件；以及

用于传送该分组的部件；以及

接收单元，具有：

用于将所接收的分组存储在接收缓冲器中的部件；

丢失检测部件，用于确定包含与要提取的帧的编号相关联的帧声学信号的分组是否存储在接收缓冲器中；

声学信号分组解码部件，用于如果在丢失检测部件中确定包含与要提取的帧的编号相关联的帧声学信号的分组存储在接收缓冲器中，则从存储在接收缓冲器中的分组提取帧声学信号，并且提供帧声学信号作为帧输出声学信号；

丢失处理部件,用于如果由于包含与要提取的帧的编号相关联的帧声学信号的分组没有存储在接收缓冲器中而在丢失检测部件中确定发生分组丢失,则从存储在接收缓冲器中的分组提取作为丢失帧的帧的声学信号对应数据,并且通过使用声学信号对应数据来生成帧输出声学信号;以及

用于串联从声学信号分组解码部件或丢失处理部件输出的帧输出声学信号,并且输出所串联的帧输出声学信号的部件;

该声学信号分组通信设备的特征在于:

传送单元中的所述包含部件还包含指示在包含帧声学信号的相同分组中帧编号差异的延迟量控制信息,以及

声学信号对应数据是对于具有相差由延迟量控制信息指示的值的帧编号的帧的数据;以及

接收单元中的所述丢失处理部件适用于如果发生分组丢失,则通过使用包括在分组内的延迟量控制信息,从接收缓冲器中的分组获得具有与丢失帧相同的帧编号的声学信号对应数据,

接收单元具有用于确定所接收分组的抖动状态的第一确定部件以及确定所接收分组的丢失状态的第二确定部件中的两者或一个;以及

用于通过使用在任一确定部件中做出的确定结果,确定要存储在接收缓冲器中的分组数作为存储分组数的目标值的部件;以及

传送单元具有用于将延迟量控制信息设置为小于或等于存储分组数的目标值的值的部件。

声学信号分组通信方法、传送方法、接收方法、及其设备和程序

技术领域

[0001] 本发明涉及一种提供措施以防在通过诸如因特网的分组通信网络传送诸如数字化语音和音乐的声学信号的期间发生分组丢失的通信方法、传送方法、接收方法、及其设备和程序。

背景技术

[0002] 使用 IP (网际协议) 上的语音的技术来传送语音信号的服务正在变得普遍。如图 1 所示,在语音信号传送单元 12 中将来自输入端 11 的语音信号转换成语音分组,通过诸如 IP 网络的分组通信网络 13 将其传送到语音信号接收单元 14,其中再现语音信号并且将其输出到输出端 15。存在这样的问题,即,当实时传递分组时,根据网络的条件,在分组通信网络 13 上可能发生分组丢失,从而导致质量降级,例如可听到的再现话音中的不连续。在容忍分组丢失的所谓尽力而为型通信服务如因特网中,该问题在网络拥塞期间尤其显著。

[0003] 因此,当通过分组通信网络传送语音信号时,使用一种称作分组丢失隐藏的技术。该方案使用这样一种技术,其中如果分组在通信信道上的某处丢失,或者由于通信信道上的延迟而没有在时间限制之内到达接收端,则在接收端估计并且补偿与已经消失或没有到达的分组(以下被称作“遗失分组”或“丢失分组”)相对应的片段中的语音信号。图 2 示出了图 1 所示的语音信号传送单元 12 的典型配置的示例。将输入语音信号存储在输入缓冲器 21 中,然后在语音分组生成单元 22 中将语音信号分裂成具有预定长度的、被称作帧的时间片段,以生成语音分组,并且将语音分组从分组传送单元 23 发送到分组通信网络。一个帧的时长典型地是 10 到 20 毫秒左右。

[0004] 图 3 示出了图 1 所示的语音信号接收单元 14 的典型配置的示例。将在分组接收单元 31 处通过分组通信网络接收的语音分组存储在也被称作抖动吸收缓冲器的接收缓冲器 32 中。对于成功接收到其分组的帧,从接收缓冲器提取语音分组,并且在语音分组解码单元 33 中将其解码成语音信号。对于其分组丢失的帧,在丢失信号生成单元 34 中,执行分组丢失隐藏处理,以生成语音信号,并且输出所生成的语音信号。如果基音 (pitch) 周期(在时间轴上等于声音的基频的长度)信息用于分组丢失隐藏处理,则将输出语音信号存储在输出语音缓冲器 35 中,然后在基音提取单元 36 中执行对信号的基音分析,并且将所获得的基音周期值提供给丢失信号生成单元 34。通过选择器开关 37 将在丢失信号生成单元 34 中生成的信号输出到输出端 15。如果没有分组丢失,则通过选择器开关 37 将来自语音分组解码单元 33 的解码信号输出到输出端 15。应当注意,执行双向语音通信的通信终端既具有传送器又具有接收器。一种公知的典型分组丢失隐藏技术是在非专利文献 1 中描述的技术。非专利文献 1 中的技术将声音的基音周期用于分组丢失隐藏。

[0005] 图 4 示出了也在非专利文献 1 中使用的典型分组丢失隐藏技术。图 4 示出了在与帧 n ,即接收端的当前帧,相对应的分组丢失时执行的处理。这里假定,正确地对语音信号的先前帧(直至帧 $n-1$)进行了解码,或者通过使用分组丢失隐藏而生成了丢失分组的语音信

号。从前一帧 $n-1$ 的最后样本点剪除等于 1 个基音周期的片段 3A 的语音信号波形,并且重复所剪除的 1 个基音周期的波形,以填充帧 n 的片段 (片段 3B-3D)。

[0006] 与用零值填充帧 n 中的所有样本点而不应用任何处理相比较,通过以这种方式重复前面的 1 基音波形而生成分组丢失帧的波形,可以以自然的话音质量再现话音。

[0007] 当仅仅重复 1 基音波形时,可能由于连接点处的波形不连续而在连接点处生成刺耳的噪声。这样的连接点处的不连续可以通过使用图 5 所示的技术来防止。为了清楚起见,在图 5 中将帧 n 的片段显示成所剪除波形的交错层。首先,从帧 $n-1$ 的最后样本点剪除具有稍微长于 1 个基音周期的长度的片段 4A 的波形,例如, $5L/4$ ($5 \times L/4$), 其中 L 是基音长度。将所剪除的波形放置在偏移一个基音长度的位置 4B、4C 和 4D。因为所剪除的波长长于一个基音长度,所以产生重叠片段 4AB、4BC 和 4CD。例如,通过应用图 6 所示的三角窗口函数来叠加这些重叠片段。在图 6 中,横轴表示时间,纵轴表示权重, t_1 表示重叠片段的开始点,并且 t_2 表示重叠片段的结束点。例如,在图 5 中的重叠片段 4BC 的情况下,通过将重叠片段 4BC 中的片段 4B 部分的波形乘以加权函数 W_1 ,并且将重叠片段 4BC 中的片段 4C 部分的波形乘以加权函数 W_2 ,然后通过将乘积相加在一起,可以平滑地互连所剪除的片段 B 和 C 中的波形。在非专利文献 1 中也描述了这种叠加的细节。

[0008] 可以认为,通过在发生分组丢失的通信环境中使用非专利文献 1 中描述的技术而生成的声音质量一般是良好的。然而,如果在接近话音中的辅音和元音之间的边界处发生分组丢失,则可能生成令人不舒服的噪声 (第一问题)。此外,如果连续多个帧的分组丢失 (被称为突发 (burst) 丢失),也就是,如果两个或更多个连续帧的分组丢失,其中每个帧具有 20 毫秒的长度,或者如果具有较长帧长的语音编码格式的一个帧的分组丢失,也就是,如果具有 40 或 60 毫秒帧长的语音编码格式的分组丢失,则生成嘈杂的蜂鸣音或不自然的声音 (第二问题)。

[0009] 第一问题起因于在非专利文献 1 所述的方法中,创建具有与紧邻的前面帧的语音波形相同特性的波形,以生成丢失帧的声音。也就是,如果在辅音和随后的元音之间的边界周围更接近元音的帧丢失,则虽然事实上丢失帧是元音的周期,但是生成具有与辅音相同特性的声音波形。在声音从元音变到静音或辅音时,可能生成类似的噪声。

[0010] 即使当在不接近辅音和元音之间的边界的片段中发生分组丢失时,也可能出现第二问题。这是由这样的事实引起的,即, (自递归地) 重复使用分组丢失帧中的声音,以生成具有相邻的后续帧丢失片段中的相同特性的声音波形,因此在长达 40 到 60 毫秒或更多的时间周期内连续地再现具有相同特性的声音波形。实际语音的基音周期和能量略微变化,并且当连续地再现具有相同特性的声音时,该声音被感觉为不同于语音的声音。

[0011] 为了解决第一和第二问题,提出了在非专利文献 2 中描述的技术。在该文献的技术中,预先将第 k 帧的辅助信息嵌入到第 $k+1$ 帧中。如果第 k 帧由于分组丢失而没有到达,则使用嵌入到第 $k+1$ 帧中的辅助信息来隐藏第 k 帧中的错误。

[0012] 非专利文献 1:ITU-T Recommendation G.711 Appendix I, "A high quality low-complexity algorithm for packet loss concealment with G.711", pp. 1-18, 1999。

[0013] 非专利文献 2:Naofumi Aoki, "A packet loss concealment technique for VoIP using steganography based on pitch waveform replication", IEICE Vol. J86-B,

No. 12, pp. 2551-2560, 2003。

发明内容

[0014] 本发明所要解决的问题

[0015] 在发生分组丢失的通信环境中使用非专利文献 1 所述的分组丢失隐藏技术具有在一定程度上减小再现声音质量降级的效果。然而,如果分组丢失发生在辅音和元音之间的边界处附近,则不能充分地消除令人不舒服的噪声。此外,如果连续多个帧的分组丢失,则将产生感觉不自然的再现声音质量。

[0016] 非专利文献 2 描述了添加辅助信息以隐藏错误的技术。然而,该文献没有示出应当在声学信号帧之后的多少帧(延迟量),与声学信号帧同时传送其辅助信息。真实通信网络中的通信条件变化,从而分组的抖动和分组丢失的频率变化。因此,最佳延迟量也应当变化。然而,传统技术领域中的技术都没有示出用于动态地改变延迟量的方法。此外,必须预先在终端装置之间确定延迟量。

[0017] 本发明的目的是提供一种能够解决这些问题并且允许进行具有更稳定质量的声音通信的方法,以及提供其设备和程序。

[0018] 解决问题的手段

[0019] 根据本发明,在发送端,将与具有与声学信号分组相差由延迟量控制信息指定的值的帧编号的帧中的声学信号相对应的数据(以下被称为“声学信号对应数据”)、以及延迟量控制信息包含在该声学信号的相同分组中,并且进行传送。在接收端,如果不能接收到分组,则通过使用延迟量控制信息,从接收缓冲器中的分组获得具有与声学信号的丢失帧(以下被称为“丢失帧”)相同的帧编号的声学信号对应数据。然后,在接收端,使用所获得的声学信号对应数据来生成丢失帧的声学信号。

[0020] 根据本发明,根据通信网络的条件而改变延迟量。具体地说,检测在预定周期中分组抖动时间长度的最大值或连续分组丢失帧数(例如,如果单个帧的分组丢失,则该数目是 1;如果连续三个帧的分组丢失,则该数目是 3)(以下被称为“连续分组丢失数”)的最大值。控制存储在接收缓冲器中的分组的数(以下被称为“存储分组数”)和延迟量,以便满足下面条件:

[0021] 等于抖动时间量的帧数(上舍入到整数的小数) $\leq$ 延迟量 $\leq$ 存储分组数

[0022] 或者

[0023] 连续分组丢失数 $\leq$ 延迟量 $\leq$ 存储分组数

[0024] 根据本发明,在发送端,至少获得与每个帧中的声学信号相对应的基音周期作为声学特征,并且将具有彼此相差由延迟量控制信息指定的帧数的帧编号的声学信号和声学特征包含在相同分组中,并且进行传送。在接收端,如果与要输出的帧中的声学信号相对应的分组丢失,则从接收缓冲器中的分组获得与丢失分组的声学信号相对应的声学特征,从最接近于丢失分组的声学信号帧(以下被称为“丢失帧”)的帧中的声学信号,剪除具有等于包含在所获得的声学特征中的基音周期的长度的波形,以该基音周期重复所剪除的波形,并且使用所重复的信号生成丢失分组的声学信号。

[0025] 本发明的效果

[0026] 根据本发明,将延迟量控制信息包含在分组中并且进行发送,并且在另一分组中

发送每个帧的声学信号对应数据。也就是,即使在发送端改变延迟量,在接收端也能知道声学信号和声学信号对应数据之间的关系。因此,如果帧的声学信号由于分组丢失而丢失,则可以在没有分组丢失的情况下接收与声学信号相对应的数据的分组,并且可以容易地从所传送的延迟量控制信息获得它。

[0027] 此外,如果声学信号的基音周期用作声学信号对应数据,则剪除并且布置具有等于基音周期的长度的声学信号波形,以生成丢失的声学信号。因此,即使发生随机 1 个基音丢失,或发生多个帧分组丢失,也几乎以在没有发生分组丢失的情况下将实现的声音质量来再现声学信号,从而可以实现稳定的声音通信。

[0028] 以这种方式,可以通过分组丢失频率高的分组通信网络来执行实时稳定的声音通信。典型地,以这样的方式设计分组通信网络,即,它们容忍一定程度的分组丢失,以降低成本。本发明也具有降低网络成本的效果,这是因为本发明的使用可以消除设计具有低分组丢失率的高质量传输电路的需要。

[0029] 此外,根据通信网络的条件来控制延迟量,以便如果网络条件相对好,则减小存储在接收缓冲器中的分组数,以减少语音分组的接收和语音信号的实际再现之间的延迟时间。如果网络条件相对差,则增大存储在接收缓冲器中的分组数,以便可以可靠地获得声学信号对应数据,从而即使在发生突发分组丢失时,也以最小化的质量降级来再现语音信号(虽然增加了语音分组的接收和语音信号的实际再现之间的延迟时间)。

附图说明

[0030] 图 1 是示出将语音信号转换成语音分组并且通过分组通信网络传递语音分组的情况的概念图;

[0031] 图 2 是示出图 1 所示的语音信号传送单元 12 的典型功能配置的示例的方框图;

[0032] 图 3 是示出图 1 所示的语音信号接收单元 14 的典型功能配置的示例的方框图;

[0033] 图 4 是用于示出通过使用典型分组丢失隐藏技术来生成丢失信号的波形图;

[0034] 图 5 是用于示出用于防止在图 4 中通过重复基音波形而创建波形时所生成的不连续声音的重叠相加处理的波形图;

[0035] 图 6 示出了在重叠相加处理中使用的三角窗口函数的示例;

[0036] 图 7 示出了根据本发明的语音信号传送设备 100 的示例性功能配置;

[0037] 图 8 示出了根据本发明的语音信号接收设备 200 的示例性功能配置;

[0038] 图 9 示出了根据本发明的声学信号分组传送方法的处理;

[0039] 图 10 示出了根据本发明的声学信号分组接收方法的处理;

[0040] 图 11 示出了分组的示例性结构;

[0041] 图 12 是用于示出图 8 所示的丢失信号生成单元 224 的操作示例的波形图;

[0042] 图 13 示出了图 8 所示的丢失信号生成单元 224 的功能配置的具体示例;

[0043] 图 14 示出了用于选择丢失帧的声学信号的声学特征参数的处理的示例;

[0044] 图 15 是用于示出当发生连续分组丢失时执行的后向波形外插处理的波形图;

[0045] 图 16 是用于示出后向波形外插的重叠相加处理的波形图;

[0046] 图 17 示出了根据本发明的语音信号传送设备 100-2 的示例性功能配置;

[0047] 图 18 示出了根据本发明的语音信号接收设备 200-2 的示例性功能配置;

- [0048] 图 19 示出了抖动时间和存储分组数之间的关系的示例；
- [0049] 图 20 示出了根据本发明的语音信号接收设备 200-3 的示例性功能配置；
- [0050] 图 21 示出了根据本发明的语音信号接收设备 200-4 的示例性功能配置；
- [0051] 图 22 示出了根据本发明的语音信号接收设备 200-5 的示例性功能配置；
- [0052] 图 23 示出了根据本发明的语音信号接收设备 200-6 的示例性功能配置；
- [0053] 图 24 示出了根据本发明的语音信号接收设备 200-7 的示例性功能配置；
- [0054] 图 25 示出了在一个方向上传送语音分组的示例性系统配置；
- [0055] 图 26 示出了在一个方向上传送语音分组并且在另一个方向上传送控制信号的示例性系统配置；
- [0056] 图 27 示出了传送设备 300 的示例性功能配置；
- [0057] 图 28 示出了接收设备 400 的示例性功能配置；
- [0058] 图 29 示出了双向传送语音分组的示例性系统配置；
- [0059] 图 30 示出了用于不将控制信息从语音分组的接收端传送到发送端的示例性处理流程；
- [0060] 图 31 示出了用于将控制信号从语音分组接收端传送到发送端的处理流程的第一示例；
- [0061] 图 32 示出了用于将控制信号从语音分组的接收端传送到发送端的处理流程的第二示例；
- [0062] 图 33 示出了用于将控制信号从语音分组的接收端传送到发送端的处理流程的第三示例；
- [0063] 图 34 示出了用于将控制信号从语音分组的接收端传送到发送端的处理流程的第四示例；以及
- [0064] 图 35 示出了通信网络的条件在两个方向上同等地影响分组通信的示例性处理流程。

具体实施方式

[0065] 本发明可以由计算机和计算机程序实现，或者通过将其实现在数字信号处理器或专用 LSI 上来实现。虽然本发明可以应用于诸如语音或音乐信号的声学信号，但是下面将描述这样的情况，其中本发明应用于图 1 所示的、使用 IP 上的语音 (Voice over IP) 的语音通信系统。为了避免重复描述，用相同的附图标记标注具有相同功能的组件或执行相同处理的步骤。

[0066] 在第一实施例的描述中，将针对基音周期用作声学信号对应数据的情况而描述语音信号传送设备和语音信号接收设备的功能配置、以及处理流程。在第二实施例的描述中，将描述用于结合通信网络的条件而动态改变延迟量和存储分组数的方法。在第二实施例中使用的声学信号对应数据不限于基音周期。它可以是声学信号本身或其它声学信号特征参数。

[0067] [第一实施例]

[0068] 图 7 示出了语音信号传送设备 100 的示例性功能配置 (其对应于图 1 所示的传送单元 12) ; 图 8 示出了语音信号接收设备 200 的示例性功能配置 (其对应于图 1 所示的接

收单元 14) ;图 9 示出了在语音信号传送设备 100 中执行的处理的示例 ;图 10 示出了在语音信号接收设备 200 中执行的处理。

[0069] 发送端

[0070] 在传送设备 100 中,将输入语音信号存储在输入缓冲器 111 中,将其划分成被称为帧的给定时间片段,也就是,划分成帧 (步骤 S1),然后将其发送到语音波形编码单元 112。一个帧的时长典型地是 10 到 20 毫秒左右。

[0071] 语音波形编码单元 112 使用例如 G. 711 (μ 律 PCM 或 A 律 PCM) 来将输入语音信号转换成语音代码 (步骤 2),其中 G. 711 是 ITU-T 标准编码算法。应当注意,还可以使用非 G. 711 算法或其它非标准语音编码算法。将语音代码发送到分组构造单元 113。

[0072] 语音特征参数计算单元 114 使用存储在输入缓冲器 111 中的语音信号,计算当前帧中的语音信号的语音特征 (步骤 S3)。术语“语音特征参数”是指诸如基音周期 (等于语音的基频) 和能量的值。可以使用任何或所有特征。通过计算语音信号波形、或者例如通过将具有相反频谱包络特性的滤波器应用到语音信号波形而获得的信号的自相关系数,可以在基音部件 114a 中获得基音周期 (步骤 S3a)。通过计算帧中的语音信号的所有样本的平方和,在能量部件 114b 中获得能量 (步骤 S3b)。表示能量的特征参数可以通过将平方和除以一个帧长度中的样本数 (如果语音信号的采样频率是 8kHz,并且帧长是 20 毫秒,则为 160 个样本) 而获得的值,并且计算商的平方根 (每样本的能量的平方根)。

[0073] 在语音特征编码单元 115 中,量化语音特征参数,以便可以用位 (或字节) 数表示它,然后将量化值转换成代码 (步骤 S4)。例如,如果使用基音周期,并且以 8kHz 的采样频率对语音进行采样,并且基音周期搜索范围是 40 到 120 个样本,则可以分配从 0 到 80 的整数作为代码,并且可以通过使用 7 位来对基音周期进行编码。如果搜索范围是 20 到 160 个样本,则可以分配从 0 到 140 的整数作为代码,并且可以通过使用 8 位来对基音周期进行编码。可以通过对每个样本的能量的平方根应用 G. 711 而以 8 位对能量进行编码 (实际上,可以以 7 位对能量进行编码,这是因为能量不会取负值,因此可以省略符号位)。

[0074] 将经过编码的语音特征发送到移位缓冲器 116。在移位缓冲器 116 中,语音特征的代码被保持在预定数目的帧上。当通过端子 117 将后面将要描述的延迟量控制信息输入到移位缓冲器 116 中时,移位缓冲器 116 将与由延迟量控制信息指定的帧数一样多的先于当前帧的帧,即过去帧,的语音信号的语音特征的代码 (也被称为“补充信息”),发送到分组构造单元 113。例如,如果延迟量控制信息是 3,则在移位缓冲器 116 中将在帧 $n-3$ 中生成的语音特征代码延迟 3 帧,然后将其发送到分组构造单元 113,其中 n 表示当前帧 (步骤 S5)。

[0075] 通过端子 118 输入后面将要描述的剩余缓冲器量,并且在剩余缓冲器量编码单元 119 中对其进行编码 (步骤 S6)。还将剩余缓冲器量代码发送到分组构造单元 113。在分组构造单元 113 中,通过使用由于语音信号波形的编码而产生的前述代码、语音特征的代码、延迟量控制信息和剩余缓冲器量代码,构造分组 (步骤 S7)。应当注意,在某些情况下,在分组中没有包括延迟量控制信息和剩余缓冲器量代码,后面将会描述该情况。

[0076] 分组传送单元 121 接收在分组构造单元 113 中构造的分组信息,并且将其作为语音分组发送到分组通信网络 (步骤 S8)。

[0077] 图 11 示出了在分组构造单元 113 中构造的分组的示例性结构。分组典型地包括

首标字段 41 和数据字段 42。首标字段 41 包括必需字段 41a 和可选字段 41b。分组的目的地和时间戳,如帧编号,包含在首标字段 41 中。通过对语音信号波形进行编码而生成的代码数据包含在数据字段 42 中。语音特征的代码可以包含在首标字段 41 的可选字段 HOP 41b 中,或者包含在数据字段 42 的开始位置 D1 或结束位置 DE 中,因为这些代码在大小(字节数)上小于波形数据。在可选字段 HOP 中包含这些代码具有这样的优点,即可以更容易地保持本发明和不使用本发明的方法之间的兼容性。分组的首标字段 41 的长度典型地是 4 字节的倍数。如果将基音周期和能量中的每个量化成 1 字节的语音特征参数,并且发送(如果它包括 7 位,则将其上舍入到 1 字节 = 8 位),则总共 2 字节的基音周期和能量信息、以及各由 1 字节表示的延迟量控制信息和剩余缓冲器量代码总计为 4 字节。例如,这 4 字节中的每个包含在可选字段 HOP 41b 的四个位置 OP1、OP2、OP3 和 OP4 中。随同补充信息(语音特征代码)一起传送延迟量控制信息的目的是让接收端知道该补充信息与哪个帧相关联。因此,可以认为,延迟量控制信息是相对于帧 n 的时间戳。随同补充信息一起传送剩余缓冲器量代码的目的是向接收端表示,当接收端将分组传送到发送端时,补充信息的传送应当相对于帧 n 延迟的帧数。

[0078] 如果补充信息相对于帧 n 延迟的帧数不动态地逐帧改变,则不需要随同补充信息一起传送延迟量控制信息和剩余缓冲器量代码,因此省略图 7 中的剩余缓冲器量编码单元 119、图 8 中的剩余缓冲器量解码单元 216、剩余接收缓冲器量确定单元 218、以及图 9 中的步骤 S6。这是因为发送和接收端只需预先知道预定的相对帧延迟(例如,通过建立呼叫连接期间的协商)。在这种情况下,不需要发送各自 1 字节长的延迟量控制信息和剩余缓冲器量代码,而只需发送基音周期和能量信息,总共 2 字节。因为首标字段的长度是 4 个字节的倍数,所以可以随同帧 n 的分组一起传送两个帧的补充信息,如帧 n-3 和 n-6 的语音信号的补充信息。因此,减小了由于分组丢失而丢失语音信号代码及其补充信息两者的概率。如果延迟量控制信息要被设置成常量值,则它应当是等于大约 60、80 或 100 毫秒的长度的帧数,以便语音信号代码及其补充信息都不会丢失,也就是,以便它们更不易受突发丢失的影响,并且仍然不损害其实时通信。

[0079] 接收端

[0080] 在图 8 的语音信号接收设备 200 中,分组接收单元 211 从分组通信网络接收语音分组(步骤 S21),并且将其存储在接收缓冲器 212 中(步骤 S22)。接收缓冲器 212 也被称作抖动吸收缓冲器。如前所述,语音分组包含通过对语音信号波形进行编码而生成的代码、诸如基音和能量代码的补充信息、延迟量控制信息、以及剩余缓冲器量信息。如果补充信息相对于帧 n 延迟的帧数不动态地逐帧改变,则不需要延迟量控制信息和剩余缓冲器量代码。然而,在下面描述中假定分组既包含延迟量控制信息又包含剩余缓冲器量代码。

[0081] 将包含在所接收的语音分组中的、通过对语音信号波形进行编码而产生的代码发送到语音分组解码单元 213,其中对该代码进行解码(步骤 S24)。在没有丢失分组的帧中,通过选择器开关 214,将来自语音分组解码单元 213 的输出信号作为再现语音提供给输出端 215(步骤 S36)。

[0082] 剩余缓冲器量解码单元 216 从包含在所接收的语音分组中的剩余缓冲器量代码,获得延迟量控制信息,其指定补充信息应当延迟的帧数。将所获得的延迟量控制信息发送到图 7 中的端子 117,也就是,发送到移位缓冲器 116 和分组构造单元 113。后面将描述发

送端的剩余缓冲器量和延迟量控制信息之间的关系。

[0083] 包含在所接收的语音分组中的延迟量控制信息用于丢失处理控制单元 217。后面将描述在丢失处理控制单元 217 中执行的处理。

[0084] 剩余接收缓冲器量确定单元 218 检测存储在接收缓冲器 212 中的分组的帧数。例如,如果在语音分组解码单元 213 正在对帧 n 进行解码时,帧 $n+1$ 、 $n+2$ 和 $n+3$ 中的语音信号代码的分组存储在接收缓冲器 212 中,则剩余缓冲器量,即所存储的分组的帧数是 3。将剩余缓冲器量信息发送到图 7 中的端子 118,即,剩余缓冲器量编码单元 119。图 7 所示的传送设备 100 和图 8 所示的接收设备 200 协作的该通信方法假定执行双向语音通信,并且每个终端设备具有传送设备 100 和接收设备 200 两者。

[0085] 语音信号传送设备 100 的端子 117 是输入端,用于输入控制传送设备 100 本身的信号。也就是,将语音特征包括在根据输入到输入端 117 的延迟量控制信息而延迟的分组中。另一方面,端子 118 是输入端,用于将信息传送到处于通信另一端的语音信号接收设备 200。该信息不控制语音信号传送设备 100。换句话说,仅仅将输入到端子 118 的信息传送到通信的另一端。

[0086] 丢失检测单元 219 检测分组丢失(步骤 S23)。按照分组编号次序,即帧编号次序,将在分组接收单元 211 接收的分组存储在接收缓冲器 212 中。例如,如果所接收分组的帧编号比紧邻在该分组之前接收的分组的帧编号大 3,则将该分组存储在从紧邻在该分组之前接收的分组的存储位置开始的接下来第三个分组存储位置。按照存储位置的次序从接收缓冲器 212 读取分组。如果在读取操作期间,在要从其读取分组的存储位置中没有找到分组,则分组丢失检测单元 219 确定发生了分组丢失,并且将选择器开关 214 转到丢失处理控制单元 217 的输出。下面将描述在丢失处理控制单元 217 处执行的分组丢失隐藏控制处理的细节。

[0087] 如果帧 n 的分组丢失,则接收缓冲器搜索单元 221 在存储在接收缓冲器 212 内的接收分组中搜索,以在所接收的帧 $n+1$ 和后续帧的分组之中查找分组按时间先后顺序最接近于丢失帧 n (其时间戳最接近于丢失帧 n 的分组)(步骤 S25)。假定作为搜索的结果,找到了帧 $n+i$ 的分组,其中如果帧 $n+1$ 没有丢失,则 $i=1$,如果连续帧 $n+1$ 和 $n+2$ 丢失并且帧 $n+3$ 没有丢失,则 $i=3$,等等,然后在先行(look-ahead)语音波形解码单元 222 中对包含在帧 $n+i$ 的分组中的、通过语音信号波形的编码而产生的代码进行解码(步骤 S26),以获得帧 $n+i$ 的语音信号波形。然后,接收缓冲器搜索单元 221 在存储于接收缓冲器 212 内的分组中搜索被添加了与丢失帧 n 的语音信号相对应的补充信息的分组(步骤 S27)。

[0088] 如前所述,如果将延迟量控制信息(相对时间戳)添加到分组,则可以基于延迟量控制信息而发现被添加了与帧 n 的语音信号相对应的补充信息的分组。例如,如果将 3 作为延迟量控制信息添加到帧 $n+3$ 的分组,则它是包含与帧 n 的语音信号相对应的补充信息的分组。如果被添加了补充信息的分组被延迟的帧数不动态地逐帧改变,则在发送端不需要将延迟量控制信息添加到分组。在这种情况下,基于预定的延迟量,搜索被添加了与丢失帧 n 的语音信号相对应的补充信息的分组。例如,如果预定延迟量是 5,则帧 $n+5$ 的分组被添加了与丢失帧 n 的语音信号相对应的补充信息。如果如前所述,将两个帧的补充信息添加到一个帧的分组中,则作为在接收缓冲器 212 中搜索的结果,可以找到具有与丢失帧 n 的语音信号相对应的补充信息的两个分组。在这种情况下,可以使用它们中的任何一个。

[0089] 如果通过该搜索而找到这样的分组（步骤 S28），则在语音特征解码单元 223 中，将所找到的、与丢失帧 n 的语音信号相对应的补充信息解码成丢失帧 n 的语音信号的基音和能量信息（步骤 S29），并且将其发送到丢失信号生成单元 224。

[0090] 另一方面，如同在非专利文献 1 所述的方法中一样，将要输出到输出端 215 的输出语音信号存储在输出语音缓冲器 225 中（步骤 S36）。如果在步骤 S28 通过搜索而没有找到分组，则在基音提取单元 226 中分析输出语音缓冲器 225 中的输出语音信号的基音周期（步骤 S30）。在基音提取单元 226 中提取的基音对应于紧邻在丢失帧之前的帧 $n-1$ 的语音信号的基音。将与紧邻的前面帧 $n-1$ 的语音信号相对应的基音发送到丢失信号生成单元 224。

[0091] 图 12 示出了丢失信号生成单元 224 的操作的概念图，并且图 13 示出了丢失信号生成单元 224 的示例性功能配置。假定与帧 n 相对应的分组丢失，其中帧 n 是当前帧。在非专利文献 1 中描述的方法逐基音地重复紧邻在分组丢失帧之前的帧 $n-1$ 的波形。在该方法中，该信号的特性，包括基音周期和能量，与紧邻的前面帧 $n-1$ 的特性相同。在该实施例中，如图 13 所示，在前向波形外插单元 51 中，从紧邻的前面帧 $n-1$ 的信号剪除与基音周期相对应的信号波形，并且基于基音周期而重复该信号波形，以生成一帧长度的波形 10-1（步骤 S31）。尽管在非专利文献 1 中描述的方法使用紧邻的前面帧 $n-1$ 中的语音信号的基音长度作为基音周期，但是本发明使用通过对补充信息进行解码而获得的丢失帧 n 的语音信号的基音长度。如果如在包括补充信息的分组也丢失的情况下，在步骤 S28 不能找到包括补充信息的分组，并且也不能获得补充信息，则使用在步骤 30 获得的紧邻的前面帧 $n-1$ 中的语音信号的基音长度。

[0092] 另外，在本实施例中，在后向波形外插单元 52 中，从在帧 $n+1$ 之后在接收缓冲器 212 中成功接收到的帧 $n+i$ 中的语音信号，剪除与基音周期相对应的信号波形（也就是，如果成功接收到帧 $n+1$ 的分组，则为帧 $n+1$ 中的信号，或者如果帧 $n+1$ 的分组也丢失并且成功接收到帧 $n+2$ 的分组，则为帧 $n+2$ 中的信号），并且重复基音周期的波形，以生成一帧长度的波形 10-2（步骤 S32）。应当注意，因为在外插中使用时间轴上的未来波形，所以在时间轴的负方向上重复该波形。如同在前向波形外插单元 51 中一样，用于后向波形外插单元 52 中的基音周期是通过对补充信息进行解码而获得的丢失帧 n 中的语音信号的基音长度。然而，如果如在包括补充信息的分组也丢失的情况下不能获得补充信息，则可以使用帧 $n-1$ 中的语音信号的基音长度，或者如果成功获得了帧 $n+1$ 中的语音信号的补充信息，则可以使用帧 $n+1$ 中的语音信号的基音长度。

[0093] 也就是，例如，如图 14 所示，如果包含丢失帧 n 中的语音信号的补充信息的分组没有丢失（步骤 S41），则通过对补充信息进行解码，获得要被剪除用于波形外插以便生成丢失帧 n 的语音信号的基音长度（步骤 S42）。如果被添加了补充信息的分组丢失，则确定外插是否是前向外插（步骤 S43）。如果是，分析紧邻的前面帧 $n-1$ 中的输出语音信号，以获得基音长度，并且使用该基音长度（步骤 S44）。否则，搜索被添加了紧邻的后面帧 $n+1$ 中的语音信号的补充信息的分组。如果它没有丢失，则处理前进到步骤 42，其中对补充信息进行解码，并且使用基音长度。如果也不能获得帧 $n+1$ 中的语音信号的补充信息，则处理前进到步骤 S44。图 12 示出了前向波形外插单元 51 和后向外插单元 52 的处理的概念图。附图标记 9-1 表示前面波形，并且 9-2 表示前向预测波形。在前向波形外插中，如片段 10B1、10C1 和

10D1 所示,重复从帧 $n-1$ 中的语音信号剪除的一个基音长度的波形 10A1,以生成波形 10-1。在后向波形外插中,如片段 10B2、10C2 和 10D2 所示,重复从帧 $n+1$ 中的语音信号剪除的一个基音长度的波形 10A2,以生成波形 10-2。图 15 中的波形 9-2 示出了在连续帧 n 和 $n+1$ 的分组丢失并且成功接收到帧 $n+2$ 的情况下,后向外插的图像。在这种情况下,从更接近于帧 $n+2$ 的帧 $n+1$ 的末端开始到更接近于帧 n 的帧 $n-1$ 的末端,作为片段 13B、13C、...、13F,从帧 $n+2$ 中的语音信号剪除 1 基音波形 13A,并重复,以获得帧 n 的外插波形 10-2。

[0094] 如在非专利文献 1 的方法的描述中关于图 5 所述的那样,如果简单地重复所剪除的一个基音长度的波形,则在连接点将发生可听到的不连续。因此,在前向波形外插中,如同在图 5 中一样,可对重叠片段使用重叠相加 (OLA) 操作。类似地,如图 16 所示,在图 5 的时间轴的反方向上,在后向外插中可以实现 OLA 操作。

[0095] 如图 13 所示,分别地,在乘法单元 53 和 54 中,将波形 10-1 和 10-2 分别乘以权重 $W1$ 和 $W2$,并且在加法单元 55 中将乘积相加在一起,以产生波形 y_n (步骤 S34)。在权重确定单元 56 中,权重 $W1$ 和 $W2$ 中的每个可以简单地被设置为 $1/2$,或者基于波形 10-1 和 10-2 的“可靠性”而确定 (步骤 S34)。基于“可靠性”意味着已经成功接收的最后帧的帧编号与当前帧 n 的帧编号之差和在接收缓冲器 212 中找到的、具有当前时间之后的时间戳的接收分组的帧的帧编号与当前帧 n 的帧编号之差当中的较小一个,被认为表现出较高的可靠性。在突发分组丢失的情况下,使用这样的“可靠性”是有用的。下面将给出示例。

[0096] 第一示例:如果前面帧 $n-1$ 的分组没有丢失,帧 n 的分组丢失,并且紧邻的后面帧 $n+1$ 的分组没有丢失,则认为前向外插波形 10-1 和后向外插波形 10-2 具有相同的“可靠性”,因此将每个权重设置为 $1/2$ 。

[0097] 第二示例:如果前面帧 $n-1$ 的分组没有丢失,帧 n 的分组丢失,紧邻的后续帧 $n+1$ 的分组丢失,并且帧 $n+1$ 之后的帧 $n+2$ 的分组没有丢失,则通过前向波形外插从前面帧 $n-1$ 获得的前向外插波形 10-1 的“可靠性”可能高于后向外插波形 10-2 的可靠性。因此,将乘以前向外插波形 10-1 的权重设置为 $2/3$,将乘以后向外插波形 10-2 的权重设置为 $1/3$ 。因为从成功接收到其每个分组的、最接近的前面帧 $n-1$ 和最接近的后面帧 $n+2$ 到将要生成语音信号的帧 n 的距离之间的比率是 $1 : 2$,所以认为可靠性之间的比率是 $2 : 1$ 。

[0098] 第三示例:如果成功接收到帧 $n-3$,帧 $n-2$ 和 $n-1$ 的分组连续丢失,帧 n 的分组丢失,并且帧 $n+1$ 的分组没有丢失,则后向外插波形 10-2 可能具有高于前向外插波形 10-1 的“可靠性”。将乘以前向外插波形 10-1 的权重设置为 $1/4$,并且将乘以后向外插波形 10-2 的权重设置为 $3/4$ 。因为从成功接收到其每个分组的、最接近的前面帧和后面帧到将要生成其语音信号的帧 n 的距离之间的比率是 $3 : 1$,所以认为可靠性之间的比率是 $1 : 3$ 。因此,例如,在图 13 的权重确定单元 56 中,分别地,在差值电路 56a 和 56b 中计算将要生成其语音信号的帧 n 和成功接收到其分组的前面帧 n_b 之差 $\Delta n1$,帧 n 和成功接收到其分组的后面帧 n_l 之差 $\Delta n2$ 。在求和电路 56c 中将差值 $\Delta n1$ 和 $\Delta n2$ 相加在一起。在除法器电路 56d 中,将差值 $\Delta n2$ 除以该和,从而获得前向外插波形 10-1 的权重 $W1 = \Delta n2 / (\Delta n1 + \Delta n2)$ 。在除法器电路 56e 中,将差值 $\Delta n1$ 除以该和,从而获得后向外插波形 10-2 的权重 $W2 = \Delta n1 / (\Delta n1 + \Delta n2)$ 。

[0099] 在刚才描述的示例中,在帧中将前向外插波形 10-1 的所有样本点和后向外插波形 10-2 的所有样本点乘以相同权重。然而,可以确定从成功接收到其分组的帧到每个样本

点的距离,并且可以根据可靠性为每个样本点设置权重。

[0100] 图 13 中的能量校正单元 57 使用通过对帧 n 的语音信号的补充信息进行解码而获得的能量信息,校正和 (sum) 波形 y_n 的能量 (步骤 S35),并且将其作为帧 n 的语音信号 x_n 输出到输出端 215 (步骤 S36)。校正能量意味着增大或减小和波形 y_n 的幅度,以便波形 y_n 的能量变成等于由补充信息表示的能量。这里,术语“等于”意味着能量彼此完全一致,或者通过考虑听觉感受效果而输出的语音信号 x_n 的能量稍微降低。

[0101] 剩余缓冲器量和延迟量控制信息

[0102] 将描述通信另一端的剩余缓冲器量和较早提及的延迟量控制信息之间的关系。

[0103] 随着延迟量 (偏移量,表示在语音信号帧之后多少帧,与语音信号帧同时传送其补充信息) 增大,信号变得更不易受突发丢失的影响。然而,如果在接收端通过使用补充信息来执行分组隐藏处理 (丢失分组的语音信号的生成),则不能再现语音,直至补充信息到达接收端。因此,语音分组的接收到语音信号的实际再现之间的延迟量 (再现延迟) 增大。在双向通信的情况下,在一端的发言和在另一端的发言的再现之间的延迟量 (呼叫延迟) 增大。如果延迟量小,则再现延迟也小,并且再现对于离散 (随机) 的分组丢失具有抵抗性。然而,当发生突发丢失时,补充信息也可能丢失。也就是,存在延迟量及其效果之间的折中。在这样的情况下,最优的是从另一端接收表示目前多少帧分组存储在通信另一端的接收缓冲器中的信息,并且将延迟量设置成等于存储在另一端的接收缓冲器中的分组的帧数。

[0104] 如前所述,为了吸收分组到达时间上的抖动,典型地将给定数目的分组存储在接收缓冲器中。尤其当分组通信网络的通信条件不太好时,典型地在另一端增加存储在接收缓冲器中的分组数,以防止由于分组到达中的抖动而引起的可听到的不连续。也就是,存在等于存储在接收缓冲器中的分组数的固有再现延迟。因此,通过将延迟量设置为等于存储在另一端的接收缓冲器中的分组的帧数的值 (剩余缓冲器量),并且传送按照该数量延迟的补充信息,可以最大化传送补充信息的效果而不增加再现延迟。如果延迟量小于或等于存储在另一端的接收缓冲器中的分组的帧数,则将不会增加再现延迟。因此,根据预期的突发分组丢失的特性,可以将延迟量设置为小于或等于存储在另一端的接收缓冲器中的分组的帧数的值。假如另一端的接收缓冲器中的剩余缓冲器量随时变化,则可以将延迟量设置为比另一端的剩余缓冲器量小一或两帧的值。

[0105] 基音周期可以单独用作声学特征。在这种情况下,省略图 7 中的能量部件 114b、图 9 中的步骤 S3b、图 10 中的步骤 S35、以及图 13 中的能量校正单元 57。例如,在发送端,可以将输入 PCM 声学信号按原样包括在分组内,而不对声学信号波形进行编码。在这种情况下,省略图 7 中的语音波形编码单元 112、图 8 中的语音分组解码单元 213 和先行语音波形解码单元 222、图 9 中的步骤 S2、以及图 10 中的步骤 S24 和 S26。此外,例如,典型地通过数字计算来获得声学特征,并且可以通过允许以 8 位获得计算结果来消除对声学特征进行编码的需要。因此,可以省略图 7 中的语音特征编码单元 115、图 8 中的语音特征解码单元 223、图 9 中的步骤 S4、以及图 10 中的步骤 S29。

[0106] 如果正确接收到前面和后面分组,并且仅仅帧 n 的分组丢失,则可以使用前向外插波形 10-1 和后向外插波形中的仅仅一个来生成丢失信号。如果发生突发分组丢失,则可以仅仅对于已经成功接收的紧邻的后面帧的声学信号使用前向外插波形,并且可以仅仅对于已经成功接收的紧邻的前面帧的声学信号使用后向外插波形。

[0107] 可以通过计算机使图 7 所示的语音信号传送设备 100 和图 8 所示的语音信号接收设备 200 中的每个运行。在这种情况下,用于使计算机执行图 9 所示的分组传送方法的处理步骤、或图 10 所示的分组接收方法的处理步骤的声学信号分组传送程序或声学信号分组接收程序,可以从诸如 CD-ROM、磁盘或半导体存储装置的记录介质安装到计算机中,或者可以通过通信网络来下载,然后可以使计算机执行该程序。

[0108] [第二实施例]

[0109] 在第一实施例中,关于语音信号传送设备 100 和语音信号接收设备 200 描述了在不同于语音信号帧的分组中包括和传送语音特征,利用接收端的剩余缓冲器量来控制关于语音信号的语音特征的延迟量,以及用于当发生分组丢失时生成丢失帧的方法。在第二实施例中,将描述用于通过使用不同于剩余缓冲器量的信息来控制延迟量的方法。此外,在不同于语音信号分组的分组中包括和传送的信息不限于语音特征,而是将它们描述成与声学信号相对应的数据(以下被称为“声学信号对应数据”)。这是因为当分组丢失时,可以在不同于从声学特征生成的方法中生成丢失帧。例如,在最简单方法中,声学信号对应数据可以是声学信号本身(也就是,传送两次声学信号)。在这种情况下,可以按原样使用在另一分组中传送的、与丢失帧相同的帧的声学信号,而不是生成丢失帧。

[0110] 语音信号传送设备和语音信号接收设备

[0111] 图 17 示出了语音信号传送设备 100-2 的示例性功能配置。该设备不同于图 7 所示的语音信号传送设备 100 之处在于该设备具有代替剩余缓冲器量编码单元 119 的信息编码单元 131、代替语音特征计算单元 114 的声学信号对应数据生成单元 132、以及代替语音特征编码单元 115 的声学信号对应数据编码单元 133。除了将剩余缓冲器量仅仅替换为在终端 118 接收的信息,并且将补充信息替换为声学信号对应数据之外,操作处理与图 9 所示的操作处理相同。端子 118 用于接收用于控制通信另一端的通信设备的信息。端子 118 从与语音信号传送设备 100-2 组合提供的语音信号接收设备接收信息,或者接收记录装置中的预定信息。该信息可以包括但不限于,存储在语音信号接收设备的接收缓冲器中的分组数(剩余缓冲器量)、关于由语音信号接收设备接收的分组的抖动的信息、由语音信号接收设备检测的在突发丢失中丢失的连续分组帧数(连续分组丢失数)、可以从抖动信息或连续分组丢失数获得的要存储在接收缓冲器中的分组数(所存储分组数的目标值)、或者存储在记录装置中的预定常数。这些信息项将被统称为用于请求设置延迟量控制信息的信息。

[0112] 图 18 示出了语音信号接收设备 200-2 的示例性功能配置。该设备不同于图 8 所示的语音信号接收设备 200 之处在于提供了延迟量控制单元 341 来代替剩余缓冲器量解码单元 216,在分组接收单元 211(抖动检测单元 232)中检测抖动,而不是由剩余接收缓冲器量确定单元 218 确定接收缓冲器 212 中的剩余缓冲器量,确定要存储在接收缓冲器 212 中的分组数(所存储分组数的目标值),并且控制接收缓冲器 212(接收缓冲器控制单元 231),使得实际上存储在接收缓冲器 212 中的分组数(剩余缓冲器量)接近所存储的分组数的目标值。延迟量控制单元 341 在从通信另一端的语音信号传送设备传送的分组中提取用于请求设置延迟量控制信息的信息。延迟量控制单元 341 还确定在与语音信号接收设备 200-2 组合提供的语音信号传送设备处的延迟量,并且将其提供到端子 117。通过将用于从剩余缓冲器量确定延迟量的上述方法中使用的剩余缓冲器量替换为所存储的分组数的目标值,可以从所存储分组数的目标值确定延迟量。虽然设置所存储分组数的目标值要求开始

用于改变存储分组数的控制操作和完成该改变之间的一定量的过渡时间,但是加以考虑的是,在一定量的时间经过之后,实际剩余缓冲器量达到存储分组数的目标值。从接收缓冲器控制单元 231 发送到与语音信号接收设备 200-2 组合提供的语音信号传送设备的信息是存储分组数的目标值而非剩余缓冲器量。以这样的方式控制实际的剩余缓冲器量,即,使它接近存储分组数的目标值。然而,在诸如抖动的变化的网络条件下,实际剩余缓冲器量达到改变后的目标值要花费一些时间。因此,实际的剩余缓冲器量在某个时间点将等于存储分组数的目标值,但是在另一时间点将大于或小于存储分组数的目标值。可以提供丢失状态确定单元 233,用于通过考虑分组丢失来确定所存储的分组数。具体地说,丢失状态确定单元 233 还接收在丢失检测单元 219 处检测的关于分组丢失的信息,并且获得突发丢失中的连续丢失分组数的最大值(以下被称为“连续分组丢失数”)。接收缓冲器控制单元 231 从关于抖动的信息和连续分组丢失数两者,确定存储分组数的目标值。

[0113] 下面将描述用于从抖动确定存储分组数的目标值的方法。例如,如果帧长是 20 毫秒,除非存在抖动,否则将每 20 毫秒接收分组。然而,根据网络条件,分组到达之间的间隔可以是 40 毫秒或 5 毫秒或甚至 100 毫秒。如果分组到达之间的间隔是每帧 20 毫秒,则抖动检测单元 232 确定抖动时间是 0 毫秒。如果分组到达之间的间隔是每帧 100 毫秒,则抖动检测单元 232 确定抖动时间是 80 毫秒。如果分组到达之间的间隔是 5 毫秒,则可能是因为紧邻的前面分组的到达由于抖动而延迟,并且抖动检测单元 232 确定抖动时间是 -15 毫秒或 15 毫秒的绝对值。使用以这种方式在每个时刻观察到的抖动时间的时间序列,以获得在给定过去时间周期中(例如,在 2 秒内)每帧(例如,每 20 毫秒)的抖动的最大值。通过存储给定时间周期上的抖动时间并且查找抖动时间之中的最大值,可以从抖动时间的时间序列获得最大值。将抖动时间的最大值发送到接收缓冲器控制单元 231。接收缓冲器控制单元 231 控制接收缓冲器中的剩余缓冲器量,使得它接近存储分组数的目标值,该目标值是等于所获得的抖动时间最大值的帧数,也就是,抖动时间的最大值除以帧长(上舍入到整数的小数),或者等于该商加 1 的帧数。可选地,可以通过利用如图 19 所示的、示出抖动时间的最大值和存储分组数的目标值之间的对应关系的表,获得存储分组数的目标值。逐帧地确定存储分组数的目标值。然而,当抖动时间的最大值没有改变时,则存储分组数的目标值也保持相同。应当注意,获得给定过去时间周期中的抖动时间的最大值、以确定等于抖动时间最大值的存储分组数的目标值,相当于获得等于时间序列中的每个抖动时间的存储分组数的目标值、以获得给定时间周期中的存储分组数的目标值的最大值。在一些情况下,获得目标值范围中的最大值(后者方法)可以提高用于查找最大值的处理的效率。

[0114] 例如,可以如下从分组丢失获得存储分组数的目标值。丢失状态确定单元 233 使用在每个时刻从丢失检测单元 219 发送的关于分组丢失的信息,以获得在给定过去时间周期(例如,在 2 秒内)中每帧(例如每 20 毫秒)的连续分组丢失数的最大值。通过存储给定时间周期上的分组丢失信息,并且搜索连续分组丢失数的最大值,可以从分组丢失信息的时间序列获得连续分组丢失的最大数。将连续分组丢失的最大值发送到接收缓冲器控制单元 231。接收缓冲器控制单元 231 控制接收缓冲器中的剩余缓冲器量,使得它接近目标值,其可以是等于所获得的连续分组丢失数的最大值的帧数、或等于该最大值加 1 的帧数。逐帧地确定存储分组数的目标值。然而,当连续分组丢失数的最大值没有改变时,则存储分组数的目标值也保持相同。通过选择从抖动时间确定的存储分组数的目标值、或者从连续

分组丢失数确定的存储分组数的目标值,这两个值中的较大者,可以从抖动时间和连续分组丢失数两者获得存储分组数的目标值。

[0115] 图 20 示出了图 18 所示的语音信号接收设备 200-2 的变体。当通过分组通信网络双向传递语音信号时,双向传送分组。在功能配置上假定,通信网络条件在两个方向上相同。基于从通信另一端传送的分组的抖动时间或连续分组丢失数,确定与语音信号接收设备 200-3 组合提供的语音信号传送设备处的延迟量,并且不将用于请求设置延迟量控制信息的信息传送到通信的另一端。在该功能配置中,没有到语音信号传送设备 100 或 100-2 的端子 118 中的输入。

[0116] 图 21 和 22 示出了仅仅从连续分组丢失数确定存储分组数的示例。该示例不同于图 18 和 20 所示的示例之处在于省略了抖动检测单元 232。

[0117] 图 23 和 24 示出了从包括在接收分组中的延迟量控制信息确定要存储在接收缓冲器 212 中的分组数的示例性配置。在这种情况下,省略接收缓冲器控制单元 231。例如,存储在接收缓冲器 212 中的分组数设为比包括在分组中的延迟量控制信息的值大 1 的值。在图 23 所示的示例中,抖动信息是向通信另一端表示接收状态的信息。在图 24 所示的示例中,连续分组丢失数是向通信另一端表示接收状态的信息。

[0118] 语音信号通信系统

[0119] 图 25 示出了在一个方向上传送语音分组的系统的示例性配置。通信设备 1000-1 包括语音信号传送设备 100-2 和记录单元 110。语音信号传送设备可以是较早由附图标记 100 表示的语音信号传送设备。记录单元 110 至少具有预先记录在其中的延迟量控制信息,并且将延迟量控制信息提供到语音信号传送设备 100-2 (或 100) 的端子 117。接收端的通信设备 1000-2 由语音信号接收设备 200-2 组成。语音信号接收设备可以是语音信号接收设备 200、200-4、200-6 或 200-7。例如,要存储在接收缓冲器中的分组数的初始值可以设为 1。在如该系统配置中完成单向通信的情况下,不需要从语音信号接收设备 200-2 到传送设备的输出信号。因此,该系统可以具有从语音信号接收设备 200 省略剩余缓冲器量解码单元 216 和剩余接收缓冲器量确定单元 218 的配置、从语音信号接收设备 200-2 和 200-4 省略延迟量控制单元 341 的配置、从语音信号接收设备 200-6 省略延迟量控制单元 341 和抖动检测单元 232 的配置、或者从语音信号接收设备 200-7 省略延迟量控制信息单元 341 和丢失状态确定单元 233 的配置。

[0120] 图 26 示出了这样的系统的示例性配置,其中在一个方向上传送语音分组,并且在另一个方向上传送用于请求设置延迟量控制信息的信息。在这种情况下,可以通过传送语音分组的相同 IP 通信网络 13 传送用于请求设置延迟量控制信息的信息,或者可以通过不同的通信网络 13-2 传送它。语音分组发送端的通信设备 1000-3 包括语音信号传送设备 100-2 和接收设备 400。语音信号传送设备可以是较早由 100 表示的语音信号传送设备。例如,延迟量的初始值可以设为 1。语音信号接收端的通信设备 1000-4 包括语音信号接收设备 200-2 和传送设备 300。语音信号接收设备可以是语音信号接收设备 200、200-4、200-6 或 200-7。例如,存储分组数的初始值可以设为 1。

[0121] 图 27 示出了上述传送设备 300 的示例性功能配置。图 28 示出了上述接收设备 400 的示例性配置。传送设备 300 的端子 320 相当于语音信号传送设备 100-2 的端子 118。通过端子 320 将来自语音信号接收设备 200-2 的信息输入到传送单元 310 中。传送单元 310

将所接收的信息转换成用于通信网络 13(或 13-2) 的协议的格式,并且将其传送到通信设备 1000-3。在这里,来自语音信号接收设备的信息是用于请求设置延迟量控制信息的信息(例如,用于请求改变延迟量控制信息、存储分组数、剩余缓冲器量、分组的抖动、或连续分组丢失数的信息)。接收设备 400 从语音分组接收端的通信设备 1000-4 接收信息。将所接收的信息输入到接收单元 410 中,并且将用于控制接收端的通信设备 1000-4 的信息提供到语音信号传送设备的端子 118。将用于发送端的通信设备 1000-3 的内部控制的信息提供给延迟量控制单元 341。延迟量控制单元 341 从所接收的信息生成延迟量控制信息,并且将其提供给语音信号传送设备的端子 117。

[0122] 图 29 示出了双向传送语音分组的示例性系统配置。通信设备 100-5 包括语音信号传送设备 100-2 和语音信号接收设备 200-2。例如,延迟量的初始值、以及要存储在接收缓冲器中的分组数可以设为 1。语音信号传送设备可以是较早由附图标记 100 表示的语音信号传送设备。语音信号接收设备可以是语音信号接收设备 200、2000-3、200-4、200-5、200-6 或 200-7。

[0123] 下面将描述上述系统配置中的处理流程。将省略设备中的详细处理流程,因为已经关于图 9 和 10 对其进行了描述。例如,在实际的设备中,不可避免地执行用于对声学信号进行解码的处理和用于生成丢失帧的方法。然而,该详细处理不是本发明的主要部分,并且省略其描述。

[0124] 图 30 示出了在没有将用于请求设置延迟量控制信息的信息从语音分组接收端传送到发送端的情况下的处理流程。该处理流程不仅可以应用于图 25 中的系统配置,还可以应用于图 26 和 29 中的系统配置。当将声学信号输入到发送端的通信设备 1000-1 中时,将它被存储在输入缓冲器 111 中(S101)。将所存储的声学信号划分成帧,并且将其编码成语音代码(S102)。然后,构造包含声学信号对应数据和延迟量控制信息的语音分组,并且将其传送到接收端(S103)。在接收端的通信设备 1000-2 中,接收语音分组(S201),并且将其存储在接收缓冲器 212 中(S202)。如果发生分组丢失(S203),则丢失处理控制单元 217 使用延迟量控制信息来搜索声学信号对应数据(S204),并且生成丢失帧(S205)。最后,按照帧次序布置并且输出声学信号(S206)。如果存在多个发送端通信设备 1000-1,则每个通信设备 1000-1 可以独立地设置延迟量。另外,图 29 所示的每个通信设备 1000-5 可以独立地设置延迟量。因此,该系统具有这样的效果,即不需要预先建立关于延迟量的普通规则,并且可以改变延迟量。

[0125] 图 31、32、33 和 34 示出了在将用于请求设置延迟量控制信息的信息从语音分组接收端传送到发送端的情况下的处理流程。这些处理流程可以应用于图 26 和 29 所示的系统配置。

[0126] 图 31 中的处理流程可以应用于这样的情况,其中语音分组发送端的通信设备 1000-3(或 1000-5)至少具有语音信号传送设备 100-2,并且语音分组接收端的通信设备 1000-4(或 1000-5)至少具有语音信号接收设备 200-2 或 200-4。该处理流程的特征在于,在接收端从抖动时间或连续分组丢失数确定用于请求设置延迟量控制信息的信息,并且在发送端从用于设置延迟量控制信息的信息获得延迟量控制信息。

[0127] 在语音分组发送端,将声学信号输入到通信设备 1000-3 中(步骤 S101),并且将其存储在输入缓冲器 111 中(S102)。构造包括声学信号对应数据和延迟量控制信息的语

音分组,并且将其传送到接收端(S103)。在语音分组接收端的通信设备 1000-4 中,接收语音分组(S201),并且将其存储在接收缓冲器 212 中(S202)。在接收端通信设备 1000-4 的抖动检测单元 232 中检测抖动时间(或者,在丢失状态确定单元 233 中检测连续分组丢失数)(S207)。接收缓冲器控制单元 231 将存储分组数的目标值设置为大于或等于抖动时间或连续分组丢失数的值(S208)。将存储分组数的目标值提供给语音分组接收端的通信设备 1000-4 的传送设备 300(或通信设备 1000-5 中的语音信号传送设备 100-2),作为用于请求设置延迟量控制信息的信息。然后,传送设备 300(或语音信号传送设备 100-2)将目标值传送到语音分组发送端的通信设备 1000-3(或通信设备 1000-5),作为用于请求设置延迟量控制信息的信息(S209)。语音分组发送端的通信设备 1000-3 接收用于请求设置延迟量控制信息的信息(S104)。延迟量控制单元 341 将延迟量控制信息的值设置为小于或等于用于请求设置延迟量控制信息的信息的值,即小于或等于存储分组数的目标值的值,并且将其提供给语音信号传送设备 100-2。在语音信号传送设备 100-2 中,将延迟量控制信息输入到移位缓冲器 116 中,并且设置延迟量控制信息(S105)。在语音信号接收端的通信设备 1000-4 中,确定是否发生了分组丢失,生成丢失帧,并且执行其它处理(S203、S204 和 S205),并且输出语音信号(S206)。除了图 30 中的处理流程的效果之外,上述处理还具有这样的效果,即可以根据不断变化的通信网络条件来设置最佳存储分组数和延迟量。

[0128] 图 32 所示的处理流程可以应用于这样的情况,其中语音分组发送端的通信设备 1000-3(或 1000-5)至少包括语音信号传送设备 100-2,并且语音分组接收端的通信设备 1000-4(或 1000-5)至少包括语音信号接收设备 200-2、200-4、200-6 或 200-7。该处理的特征如下。语音分组接收端检测抖动时间或连续分组丢失数,并且将结果传送到发送端。语音分组发送端基于抖动时间或连续分组丢失数而设置延迟量控制信息。语音分组发送端将延迟量控制信息随同语音信号一起传送到接收端。语音分组接收端基于所接收的延迟量控制信息而设置存储分组数。

[0129] 步骤 S101 到 S201 与图 31 中的处理流程的那些步骤相同。接收缓冲器 212 从所接收的语音分组提取延迟量控制信息,并且设置存储分组数(S208')。根据该设置而将语音分组存储在接收缓冲器中(S202)。抖动检测单元 232 或丢失状态确定单元 233 检测抖动时间或连续分组丢失数(S207)。将所检测的信息从语音分组接收端的通信设备 1000-4 中的传送设备 300(或通信设备 1000-5 中的语音信号传送设备 100-2)传送到语音分组发送端的通信设备 1000-3(或通信设备 1000-5)(S209)。语音分组发送端的通信设备 1000-3(或通信设备 1000-5)接收该信息(S106)。延迟量控制单元 341 将延迟量控制信息的值设置为比等于抖动时间的帧数大的值、或大于连续分组丢失数的值,并且将其发送到语音信号传送设备 100-2。在语音信号传送设备 100-2 中,将延迟量控制信息输入到移位缓冲器 116 中,并且设置延迟量控制信息(S105)。步骤 S203 到 S206 与图 31 的那些步骤相同。虽然在处理流程特征的描述中,最后设置存储分组数,但是在图 32 的处理流程图中,它被示出在检测抖动和延迟量控制信息设置之前。这是因为重复地执行处理,并且基于在前面重复中获得的延迟量控制信息而确定存储分组数。这意味着最后设置存储分组数。应当注意,例如,在第一重复处理中,可以将延迟量控制信息的初始值设置为 1。该处理流程的效果与图 31 所示的处理流程的效果相同。

[0130] 图 33 所示的处理流程可以应用于这样的情况,其中语音分组发送端的通信设备

1000-3(或 1000-5)至少包括语音信号传送设备 100-2,并且语音分组接收端的通信设备 1000-4(或 1000-5)至少包括语音信号接收设备 200-2 或 200-4。该处理的特征在于,同时从抖动时间或连续分组丢失数获得存储分组数的目标值和延迟量控制信息。

[0131] 步骤 S101 到 S207 与图 31 中的那些步骤相同。接收缓冲器控制单元 231 从所检测的抖动或连续分组丢失数中获得存储分组数和延迟量控制信息 (S210)。获得该值,使得满足以下关系:

[0132] 等于抖动时间的帧数(上舍入的小数) $\leq$ 延迟量 $\leq$ 存储分组数

[0133] 或者

[0134] 连续分组丢失数 $\leq$ 延迟量 $\leq$ 存储分组数。

[0135] 根据所获得的存储分组数的目标值而设置接收缓冲器 212,以便存储分组(S208)。将所获得的延迟量控制信息从语音分组接收端的通信设备 1000-4 的传送设备 300(或通信设备 1000-5 的语音信号传送设备 100-2)传送到语音分组发送端的通信设备 1000-3(或通信设备 1000-5),作为用于改变延迟量控制信息的请求(S211)。由语音分组发送端的通信设备 1000-3(或通信设备 1000-5)接收该信息(S107),并且由延迟量控制单元 341 将该信息设置为延迟量控制信息。在语音信号传送设备 100-2 中,将延迟量控制信息输入到移位缓冲器 116 中,并且设置延迟量控制信息(S105)。步骤 S203 到 S206 与图 31 中的那些步骤相同。该处理流程具有与图 31 和 32 中的处理流程相同的效果。

[0136] 图 34 所示的处理流程可以应用于这样的情况,其中语音分组发送端的通信设备 1000-3(或 1000-5)至少包括语音信号传送设备 100,并且语音分组接收端的通信设备 1000-4(或 1000-5)至少包括语音信号接收设备 200。该处理的特征在于传送剩余缓冲器量作为用于请求设置延迟量控制信息的信息,从而获得延迟量控制信息。

[0137] 步骤 S101 到 S202 与图 31 中的那些步骤相同。语音分组接收端的通信设备 1000-4(或通信设备 1000-5)的剩余接收缓冲器量确定单元 218 确定接收缓冲器的剩余缓冲器量(S213)。将所确定的信息从语音分组接收端的通信设备 1000-4 的传送设备 300(或通信设备 1000-5 的语音信号传送设备 100-2)传送到语音分组发送端的通信设备 1000-3(或通信设备 1000-5)(S214)。语音分组发送端的通信设备 1000-3(或通信设备 1000-5)接收剩余缓冲器量信息(S108)。延迟量控制单元 341(或剩余缓冲器量解码单元 216)获得小于或等于剩余缓冲器量的延迟量控制信息。在语音信号传送设备 100-2 中,将延迟量控制信息输入到移位缓冲器 116 中,并且设置延迟量控制信息(S105)。步骤 S203 到 S206 与图 31 中的那些步骤相同。该处理流程具有与图 31、32 和 33 所示的处理相同的效果。

[0138] 图 35 示出了这样的处理流程,其中假定当执行双向分组通信时,网络条件同等地影响两个方向上的分组通信。该处理流程可以应用于图 29 所示的系统配置。该处理的特征在于,从所接收的语音分组的抖动和连续分组丢失数中,获得要接收的存储分组数以及语音分组延迟量控制信息。步骤 S101 到 S208 与图 33 中的那些步骤相同。将所获得的延迟量控制信息输入到语音信号传送设备 100-2 的移位缓冲器 116 中,并且设置延迟量控制信息(S212)。除了图 31、32 和 33 中的处理的效果之外,该处理流程还具有减少用于控制通信设备的步骤的数目的效果。

[0139] 可以通过计算机使图 17、18、20-24、27 和 28 所示的每个设备运行。可以通过计算

机使用程序来执行图 30 到 35 所示的处理步骤。可以将该程序从诸如 CD-ROM、磁盘、或半导体存储装置的记录介质安装到计算机中,或者可以通过通信网络下载该程序。

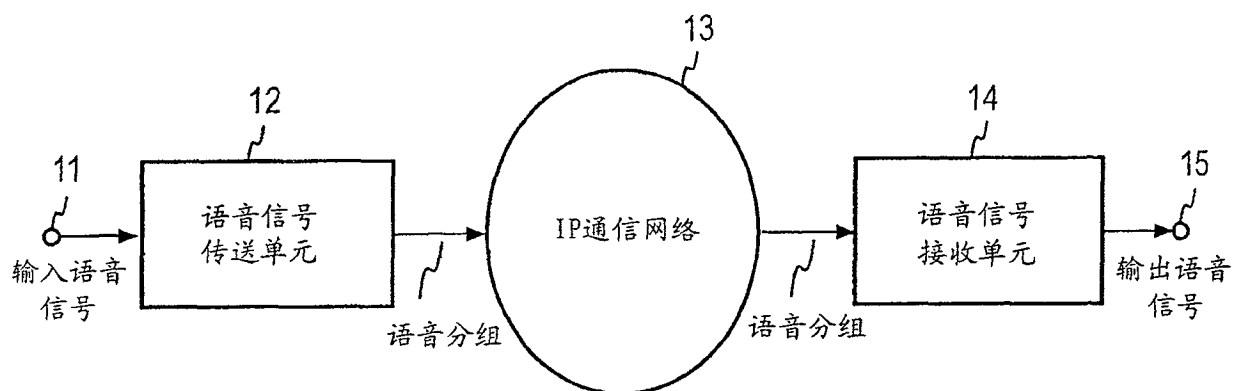


图 1

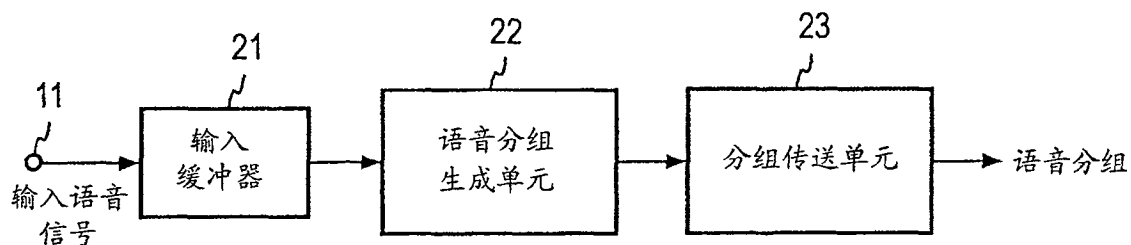


图 2

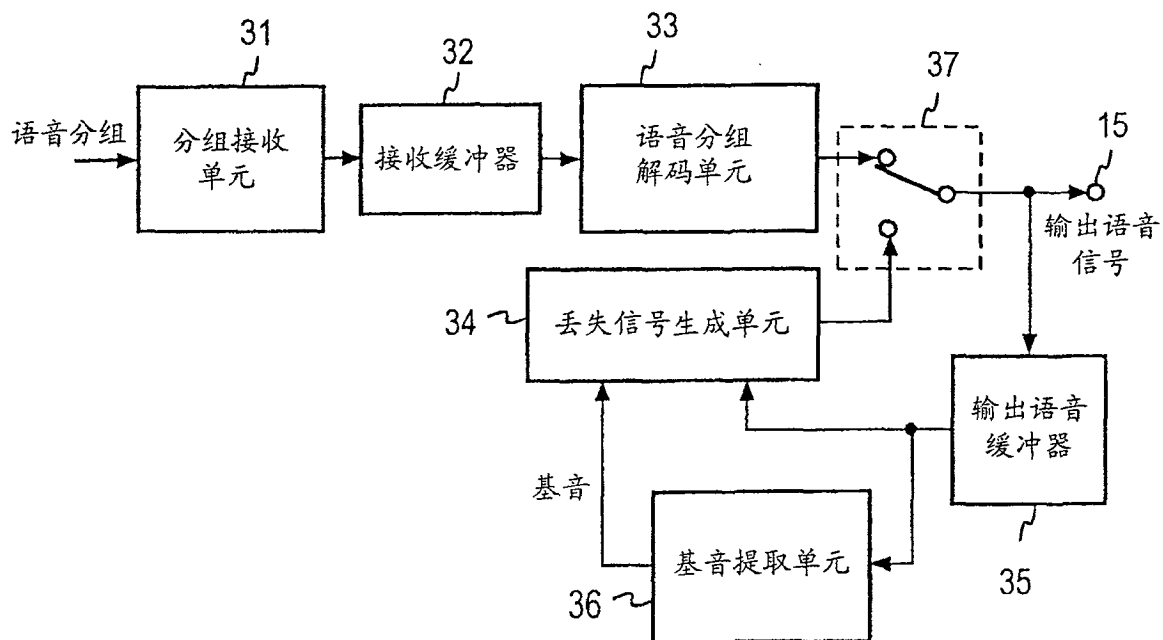


图 3

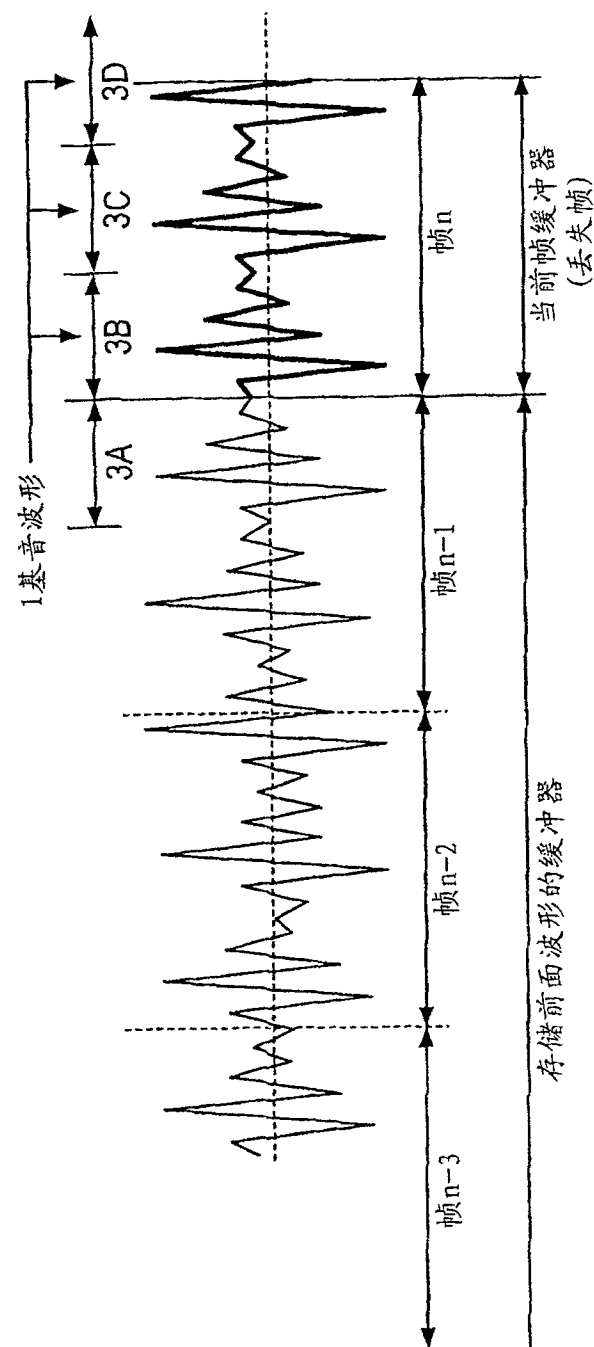


图 4

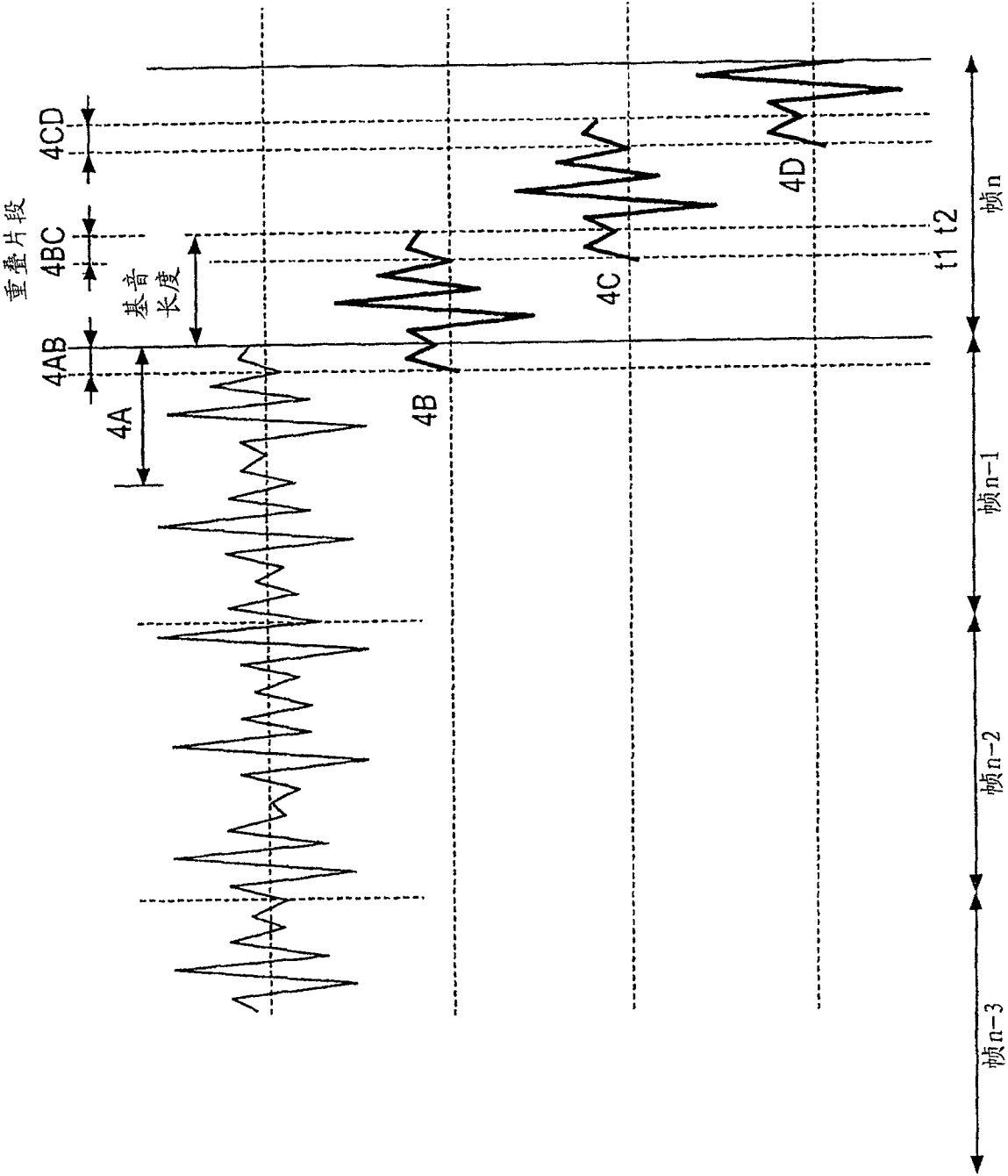


图 5

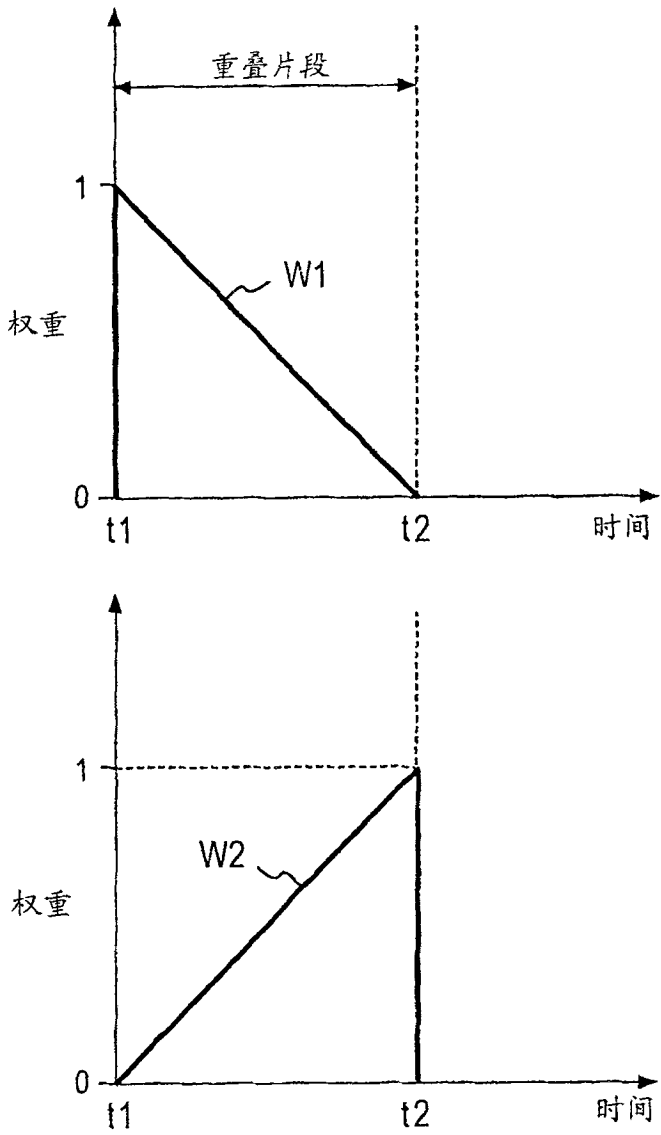


图 6

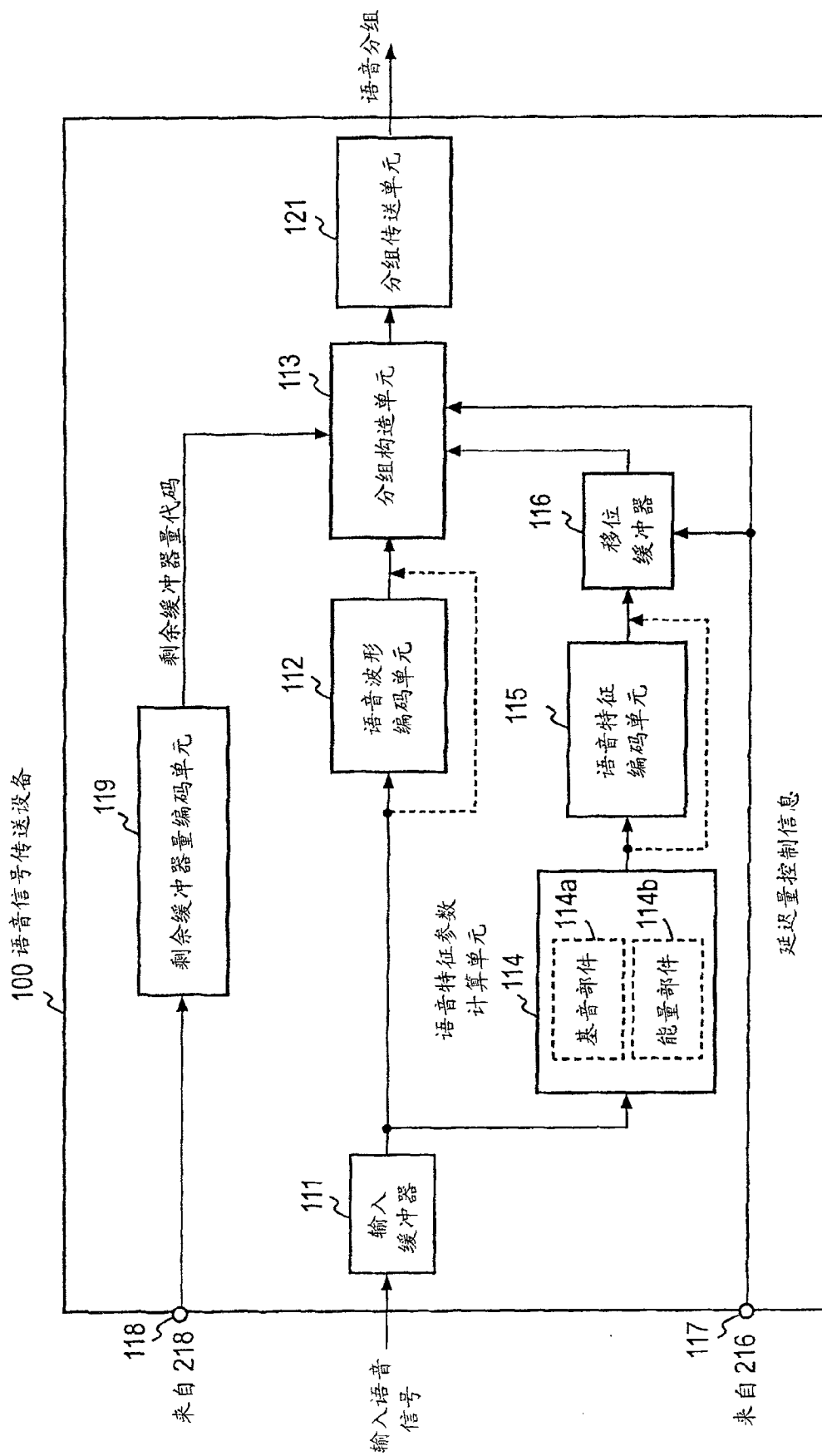


图 7

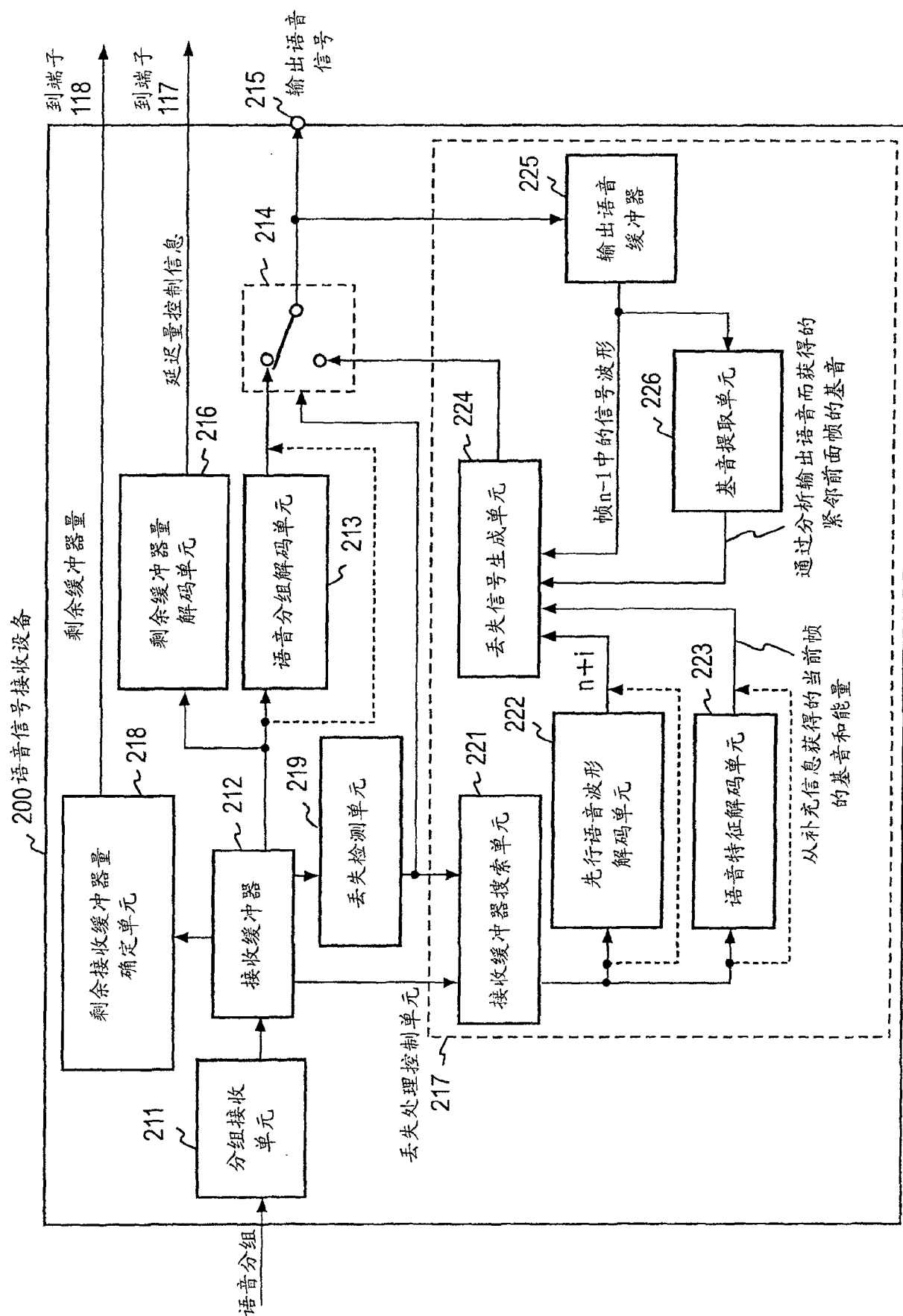


图 8

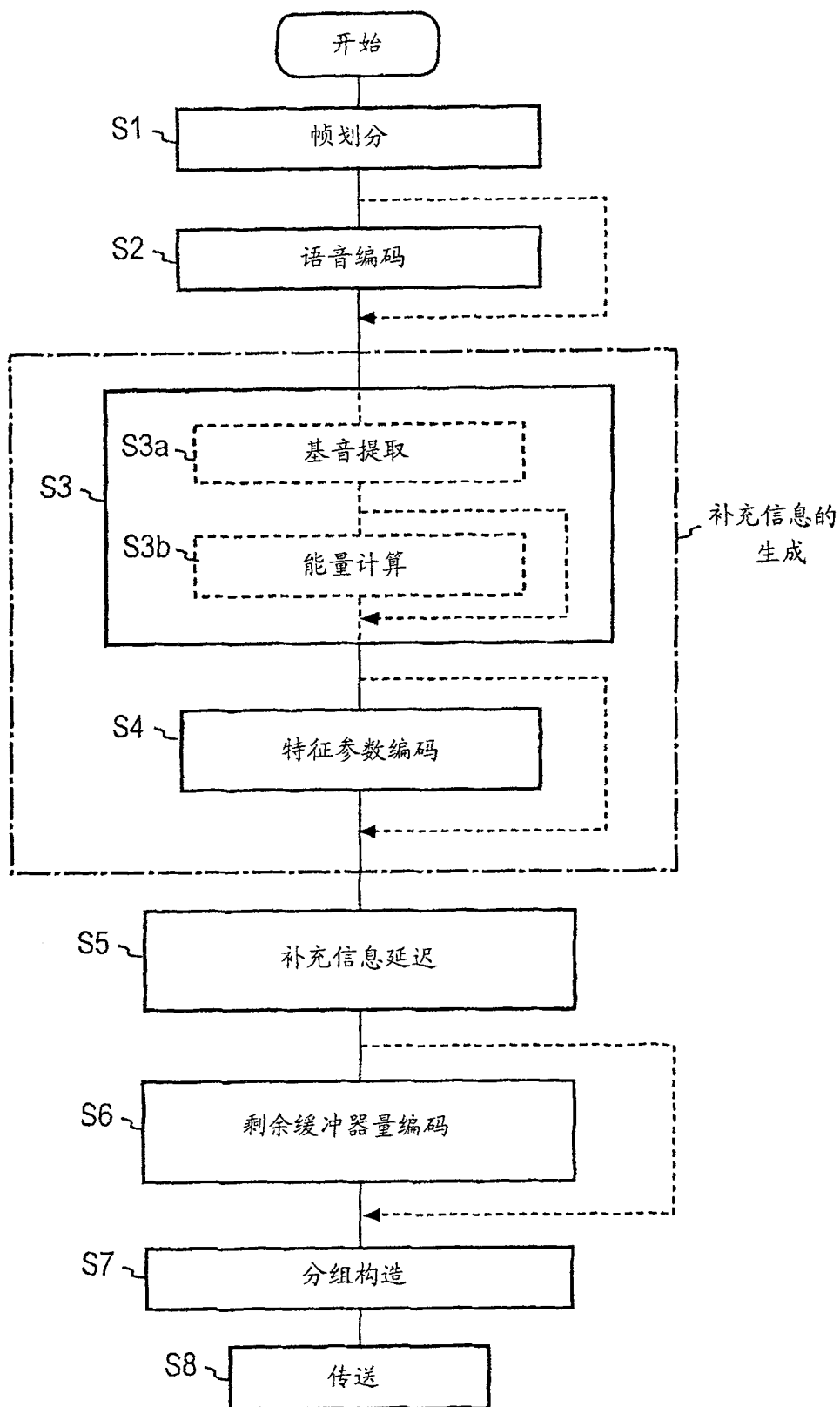


图 9

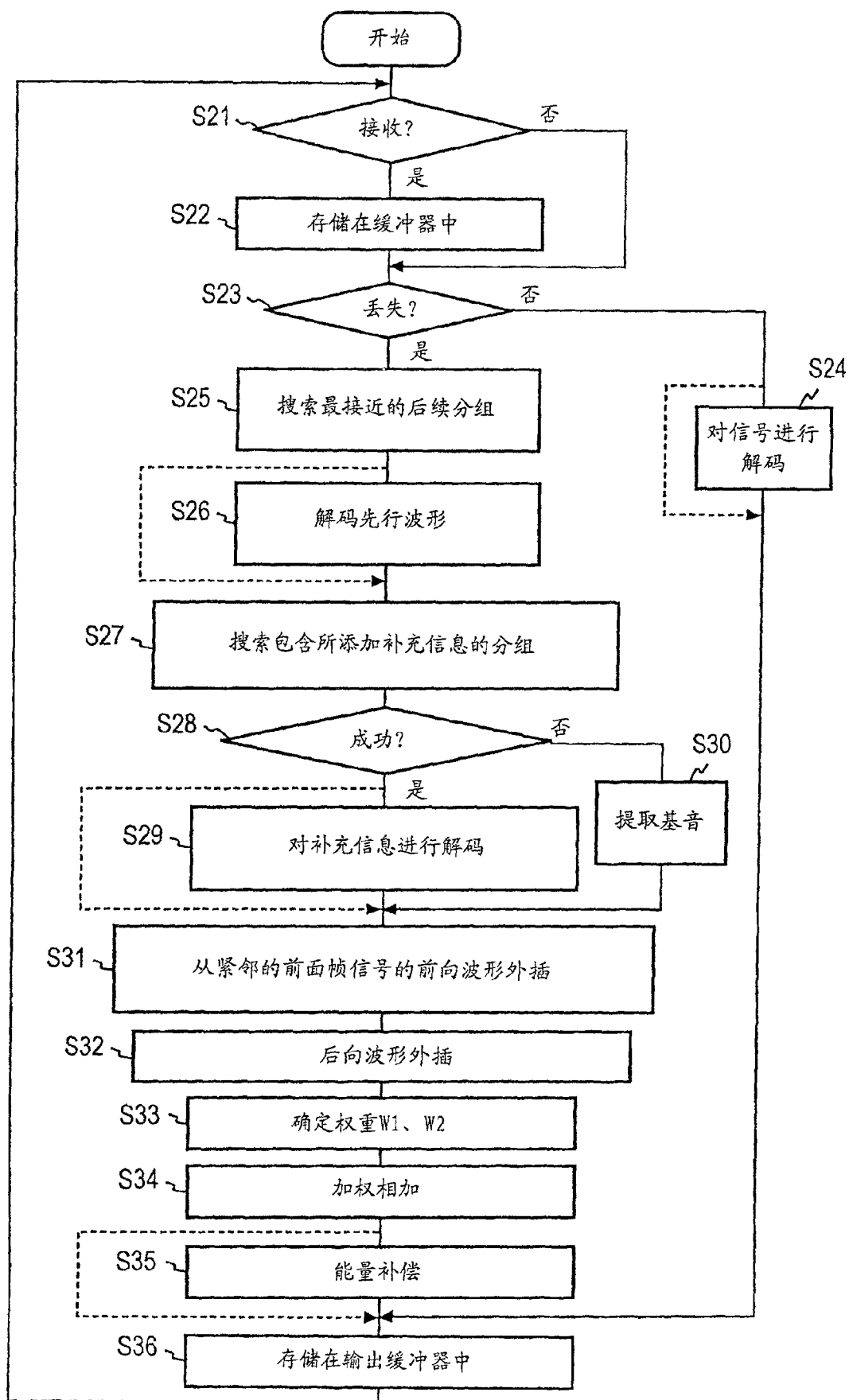


图 10

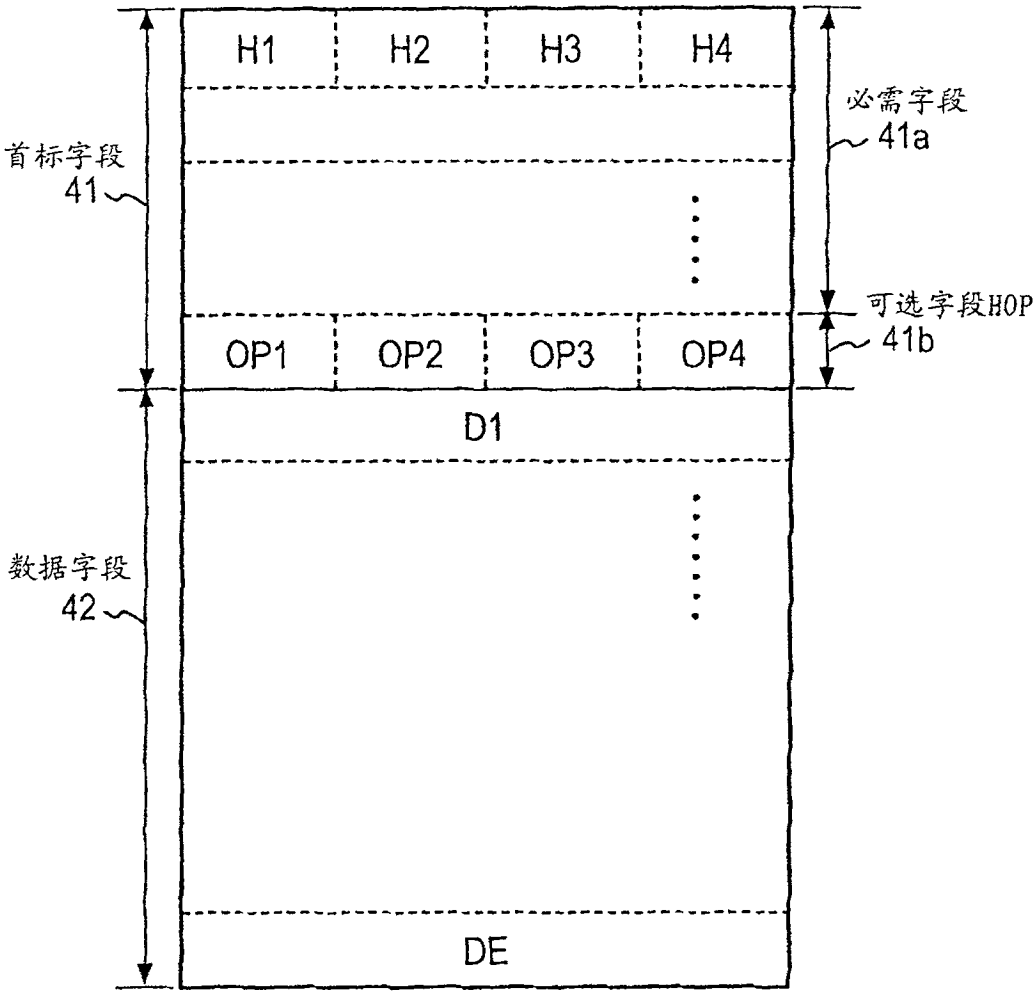
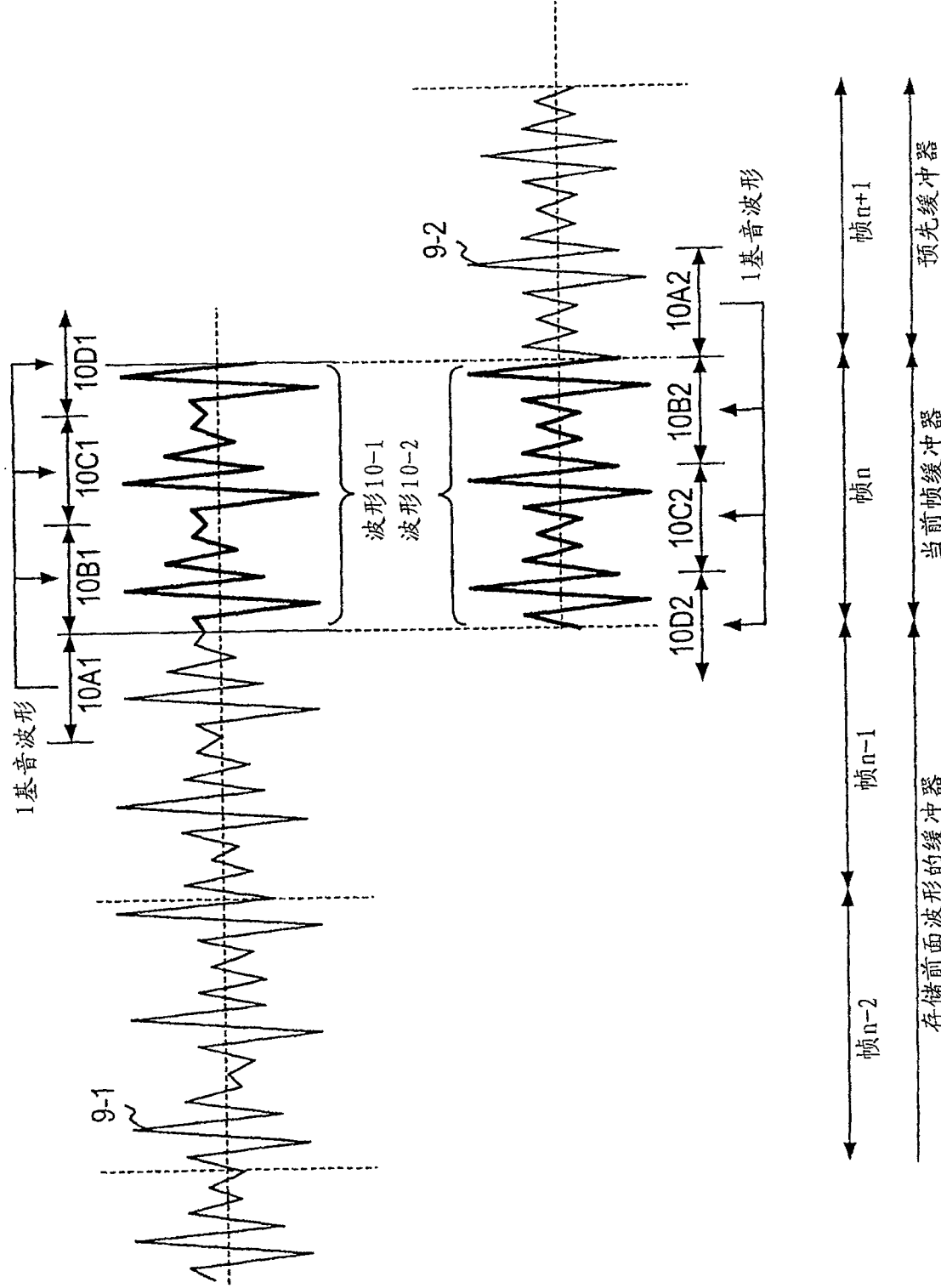


图 11



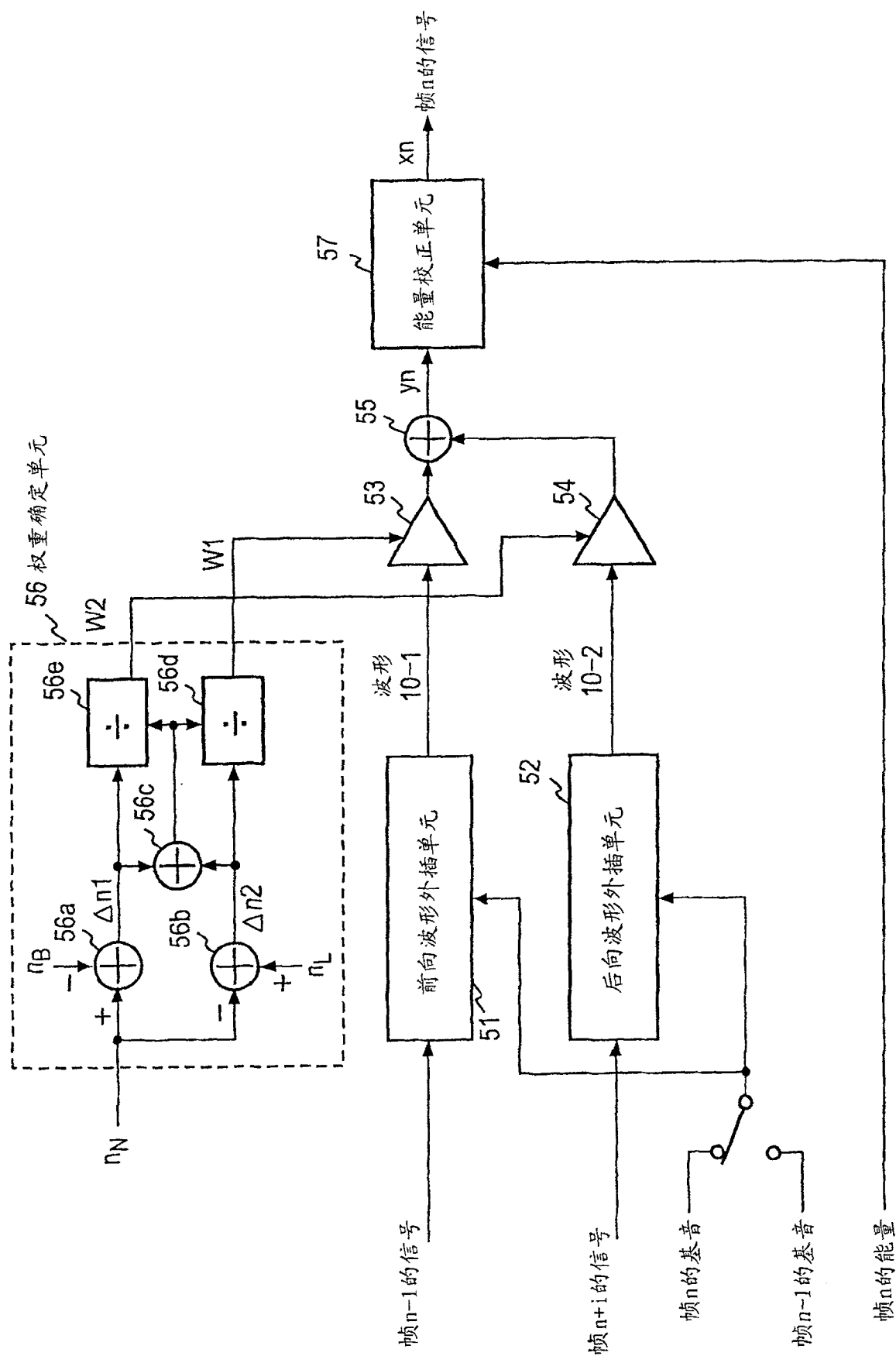


图 13

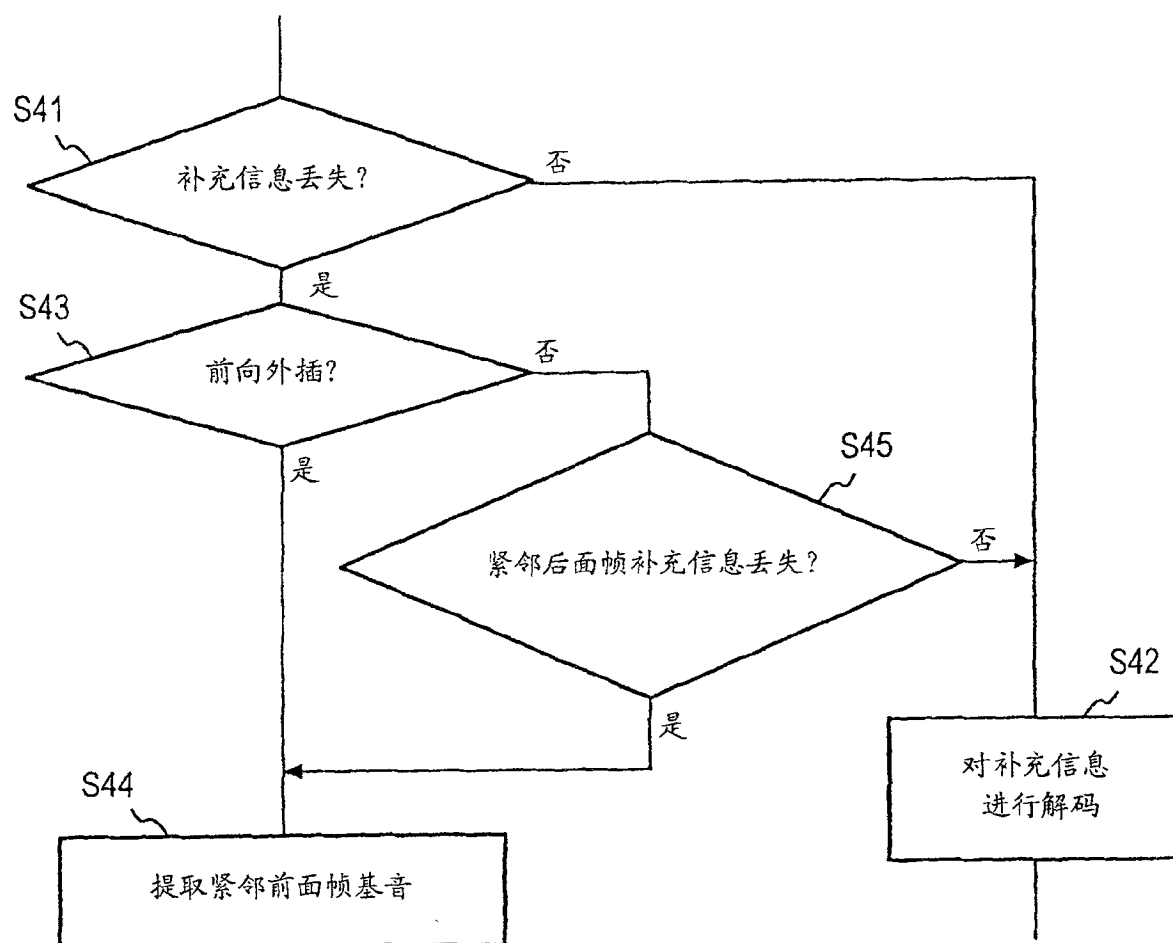


图 14

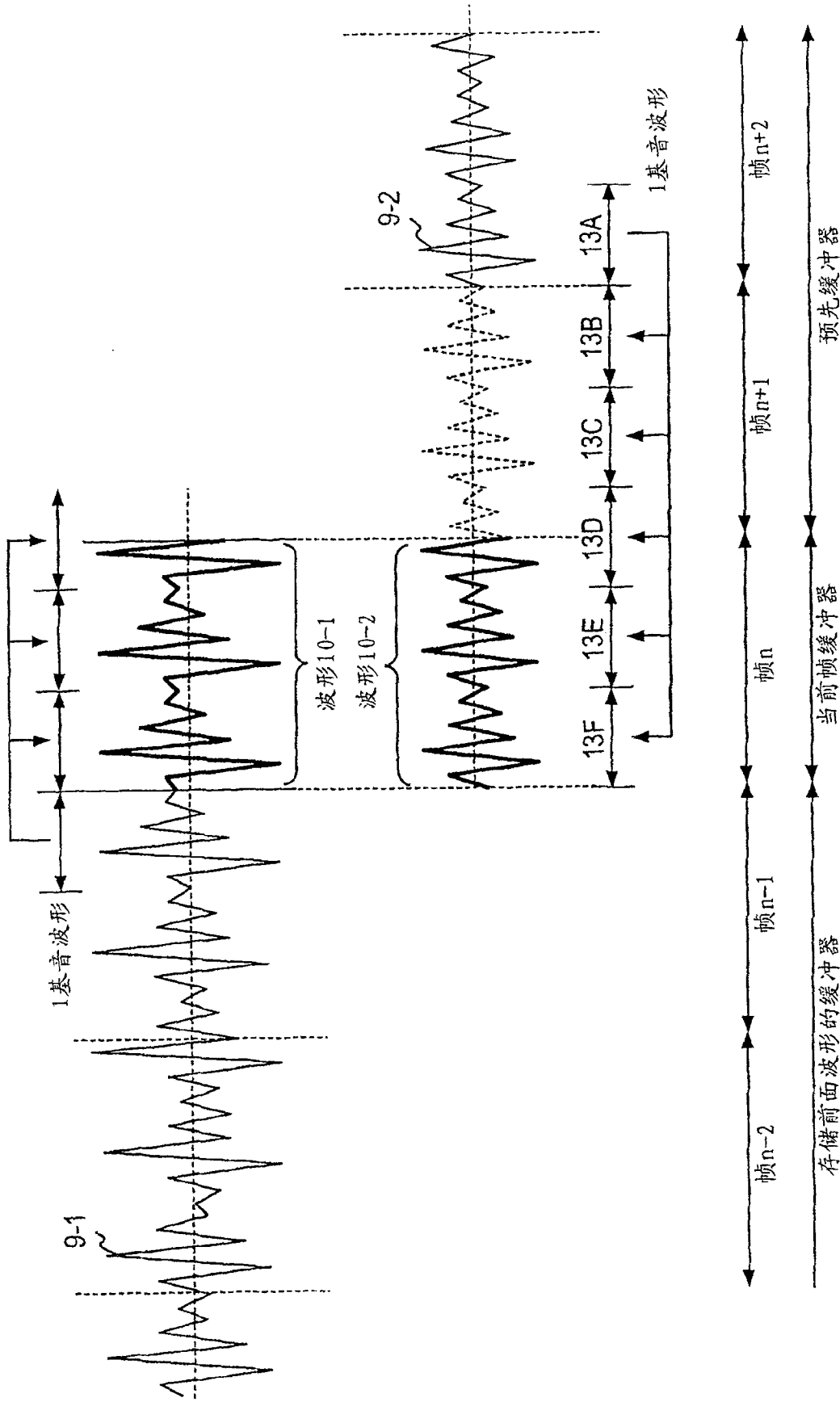


图 15

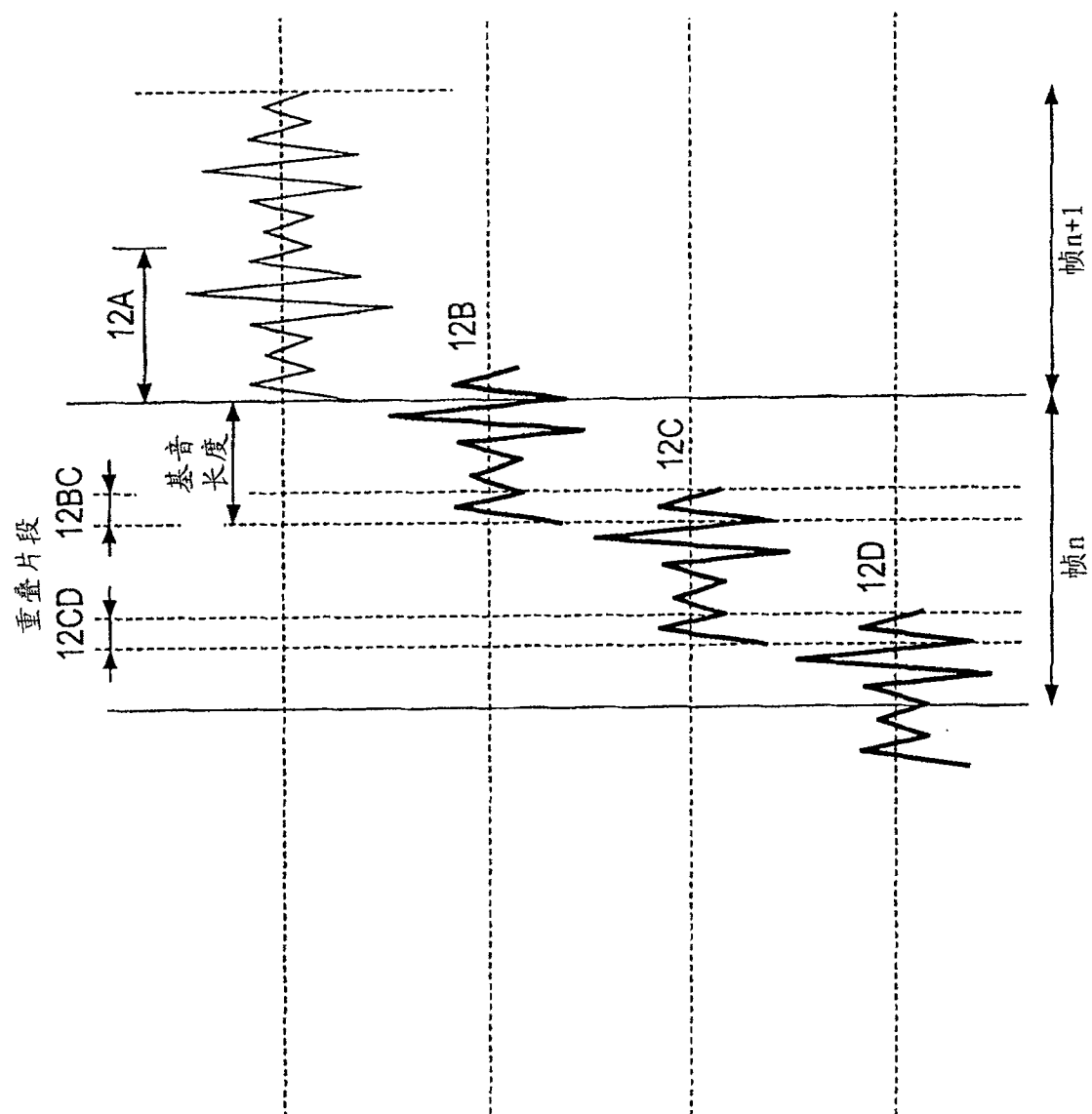


图 16

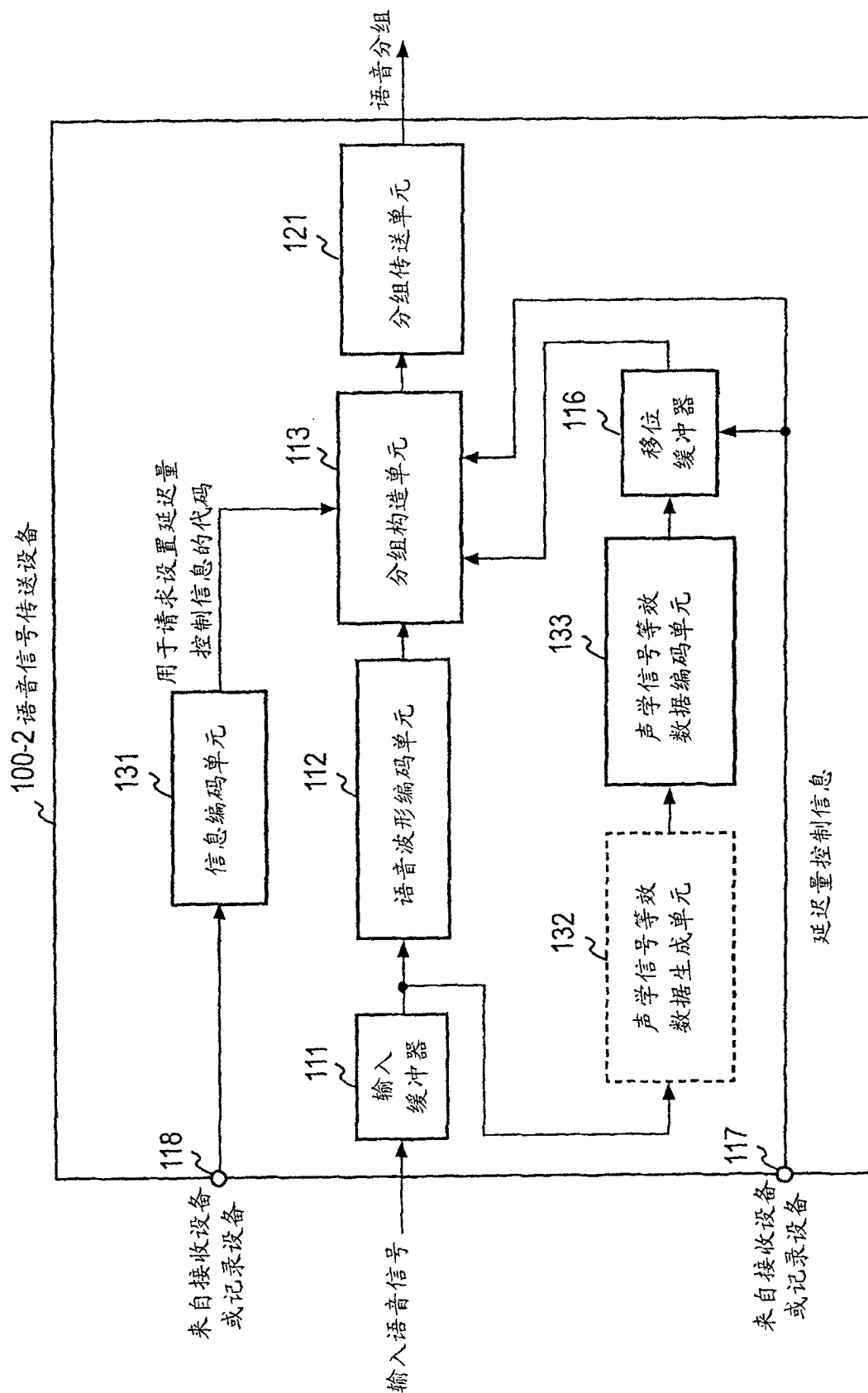


图 17

200-2 语音信号接收设备

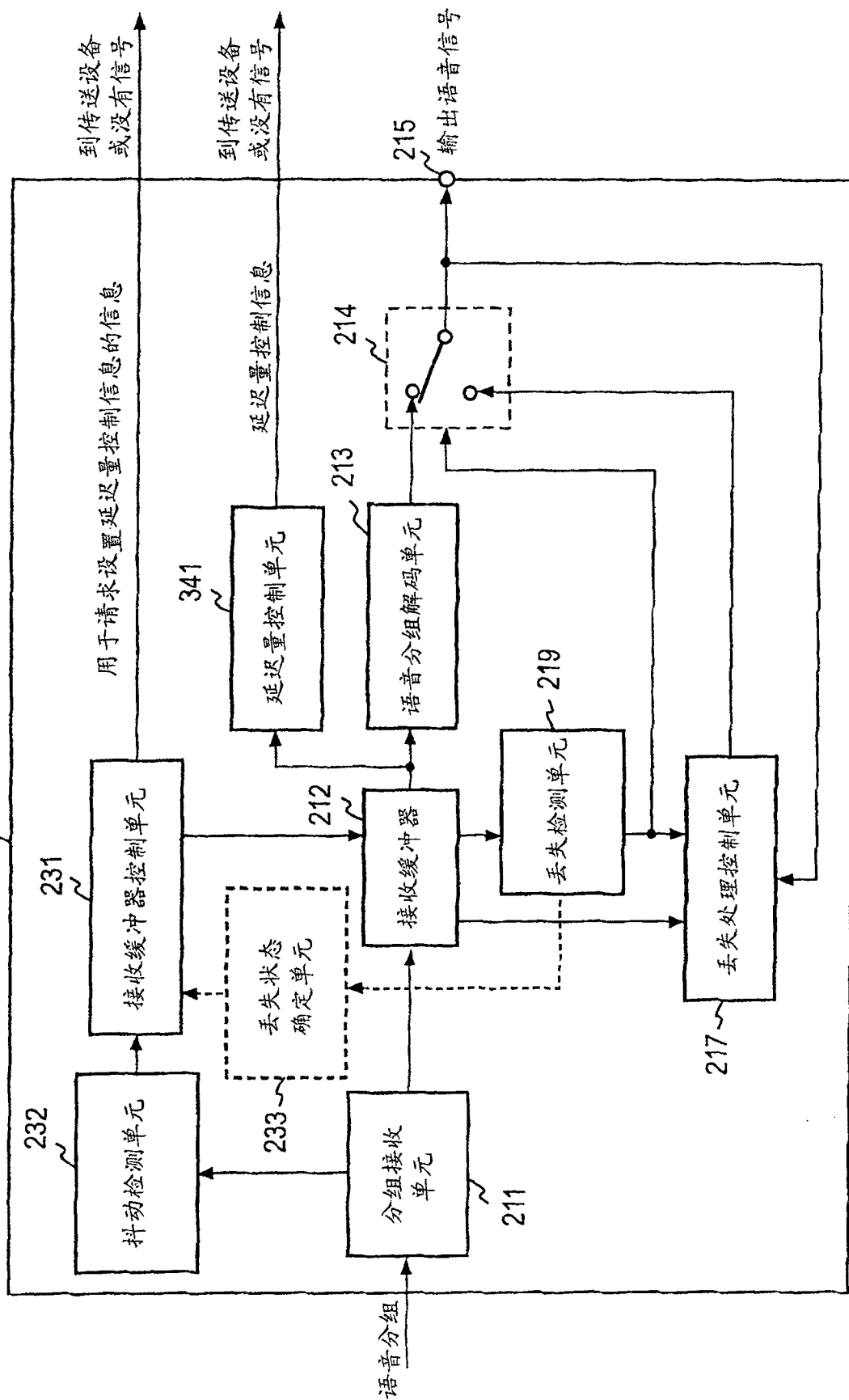


图 18

观察的抖动时间	要存储在缓冲器中的分组的数量 (存储分组数的目标值)
直至 15 ms	1 帧
直至 32 ms	2 帧
直至 48 ms	3 帧
直至 64 ms	4 帧
直至 80 ms	5 帧
直至 96 ms	6 帧
直至 128 ms	8 帧
直至 160 ms	10 帧
直至 192 ms	12 帧
直至 240 ms	15 帧
直至 320 ms	20 帧
直至 400 ms	25 帧
直至 480 ms	30 帧
直至 640 ms	40 帧
直至 800 ms	50 帧

图 19

200-3 语音信号接收设备

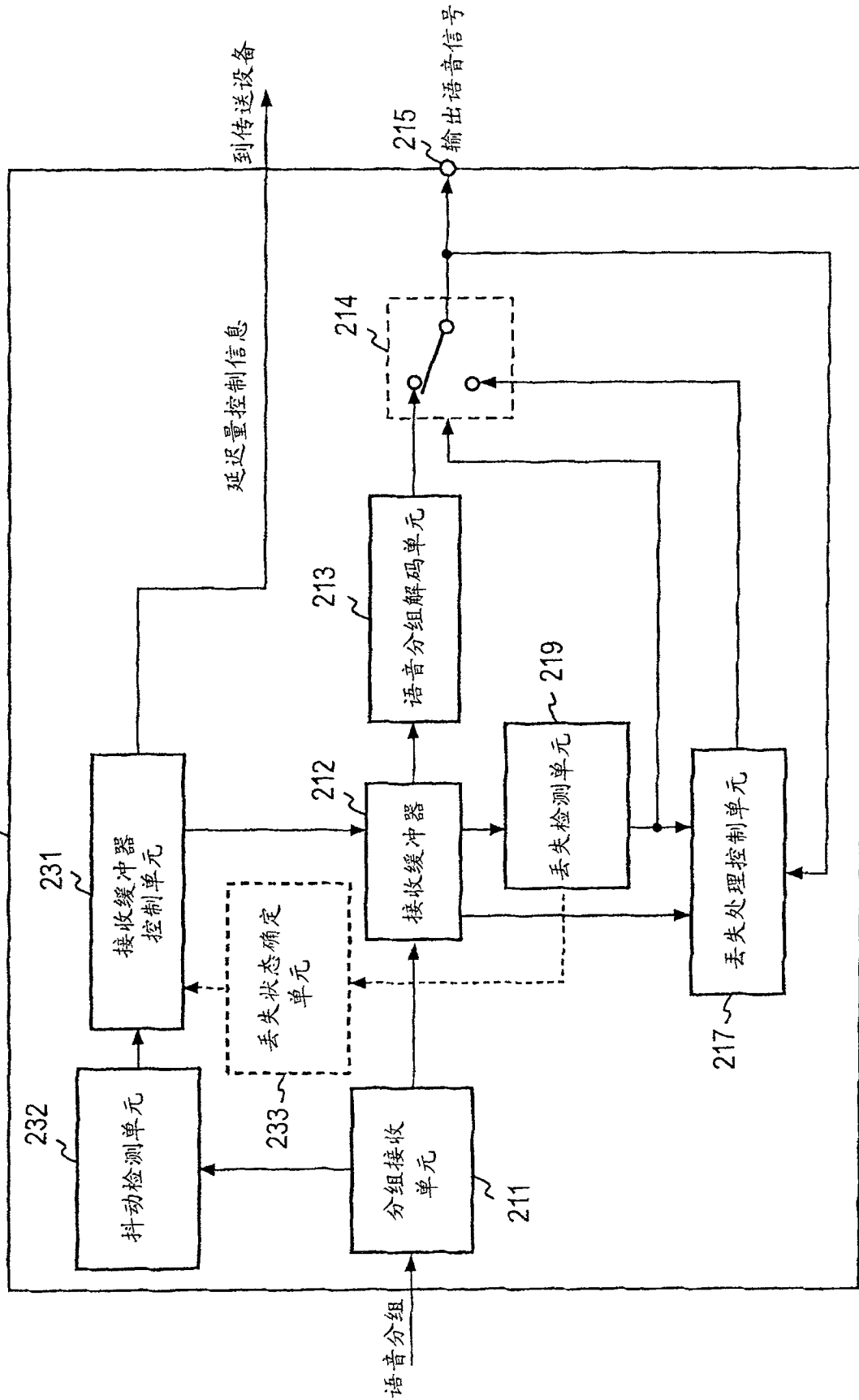


图 20

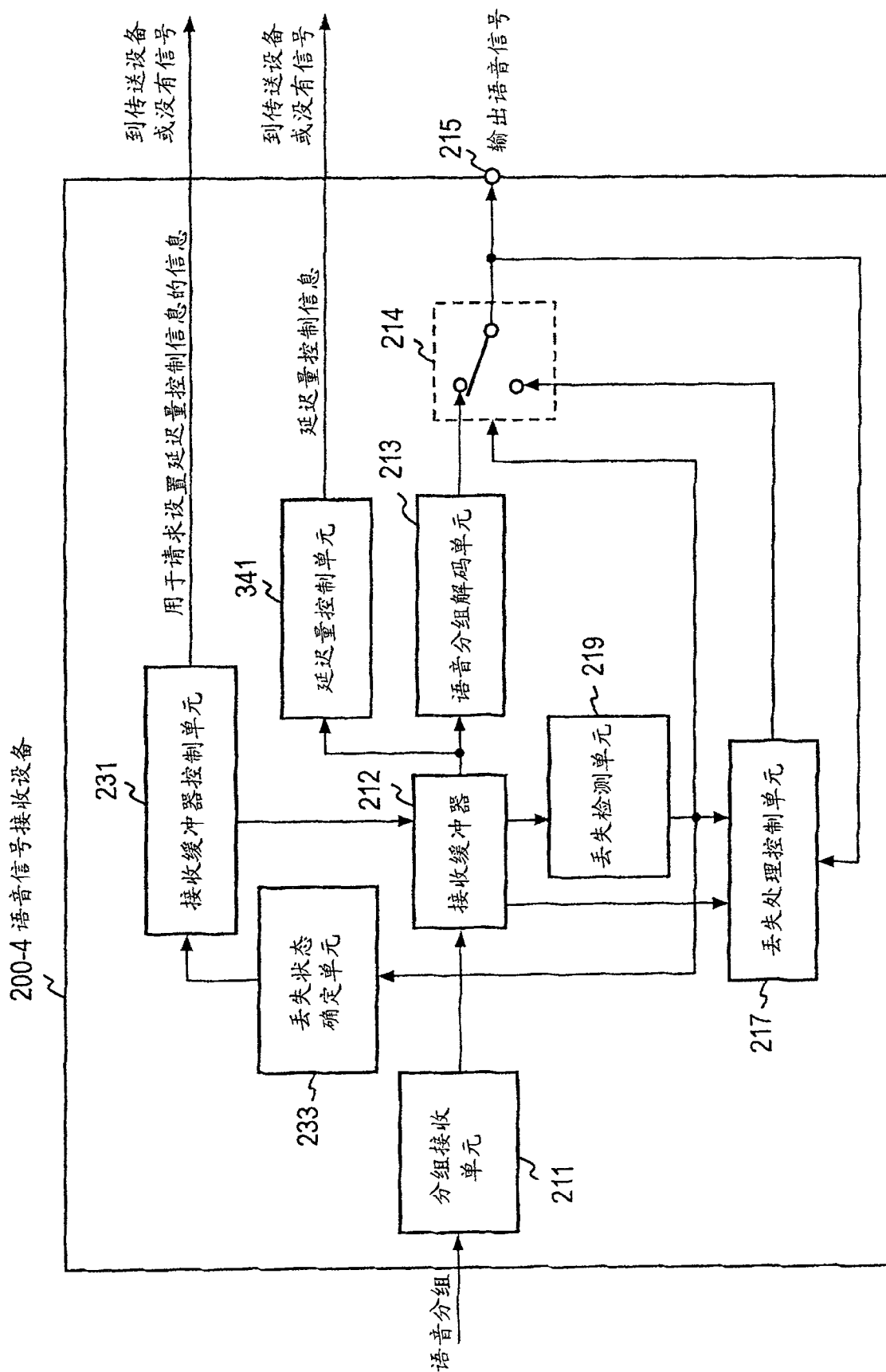


图 21

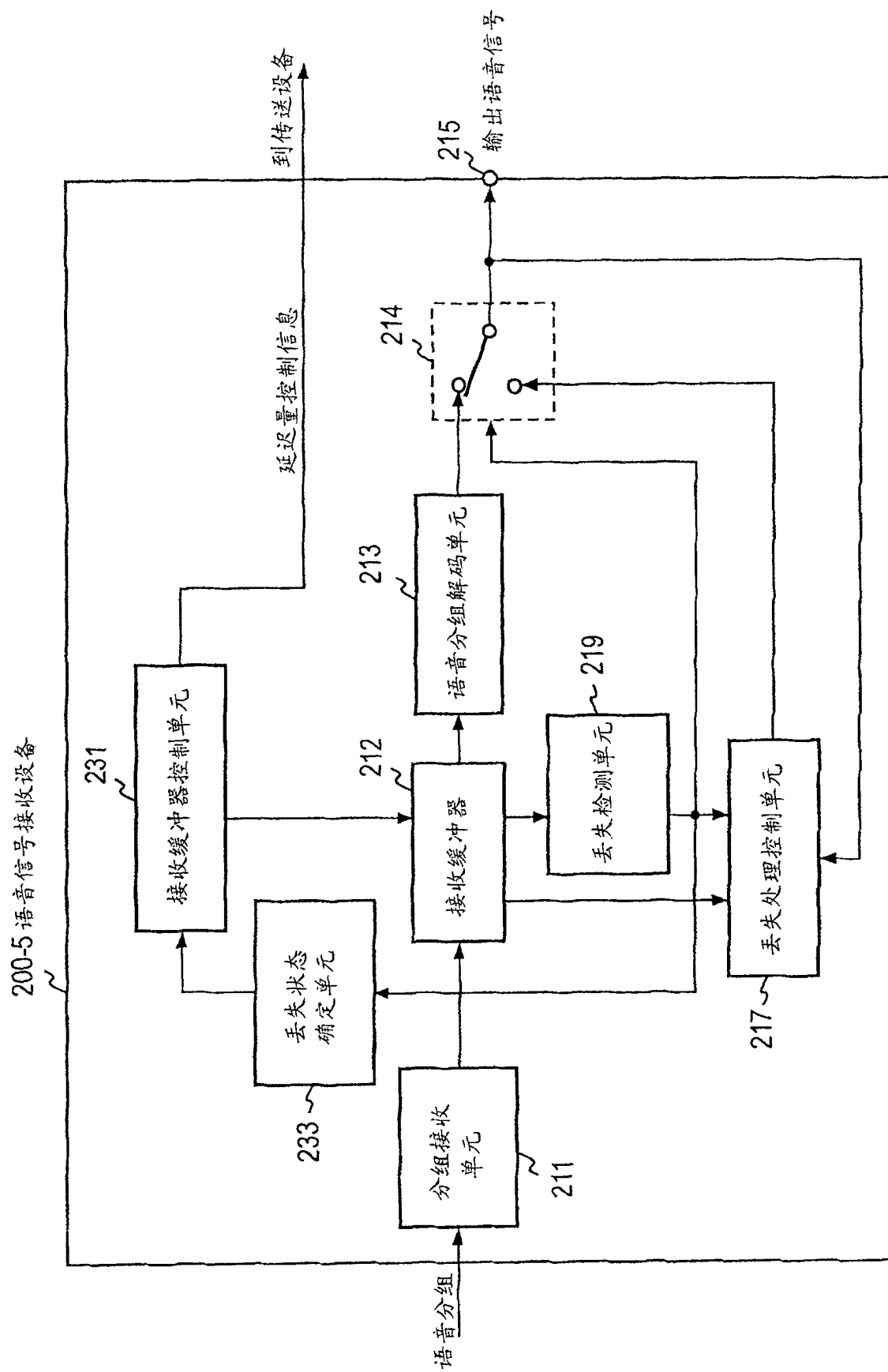


图 22

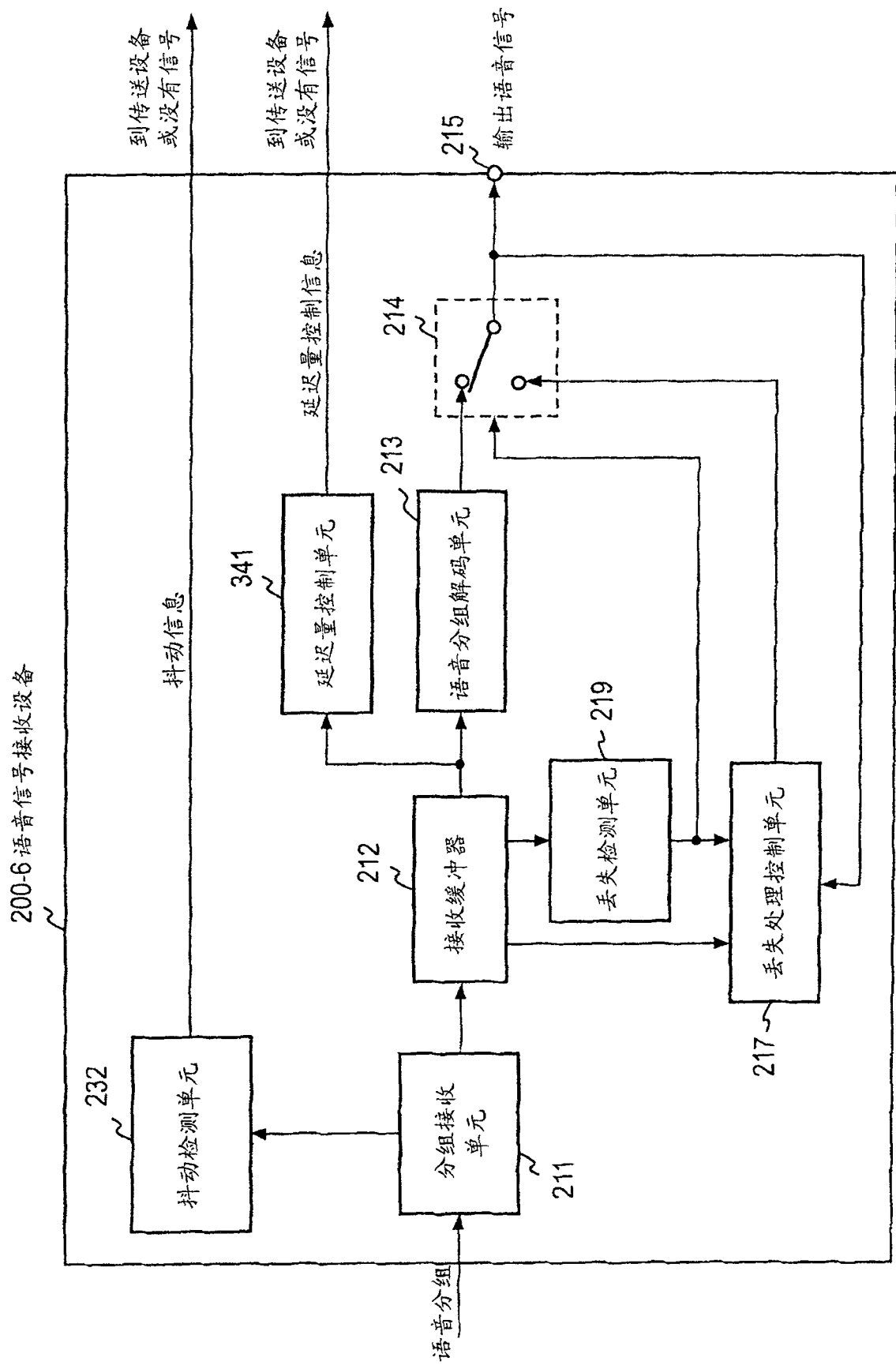


图 23

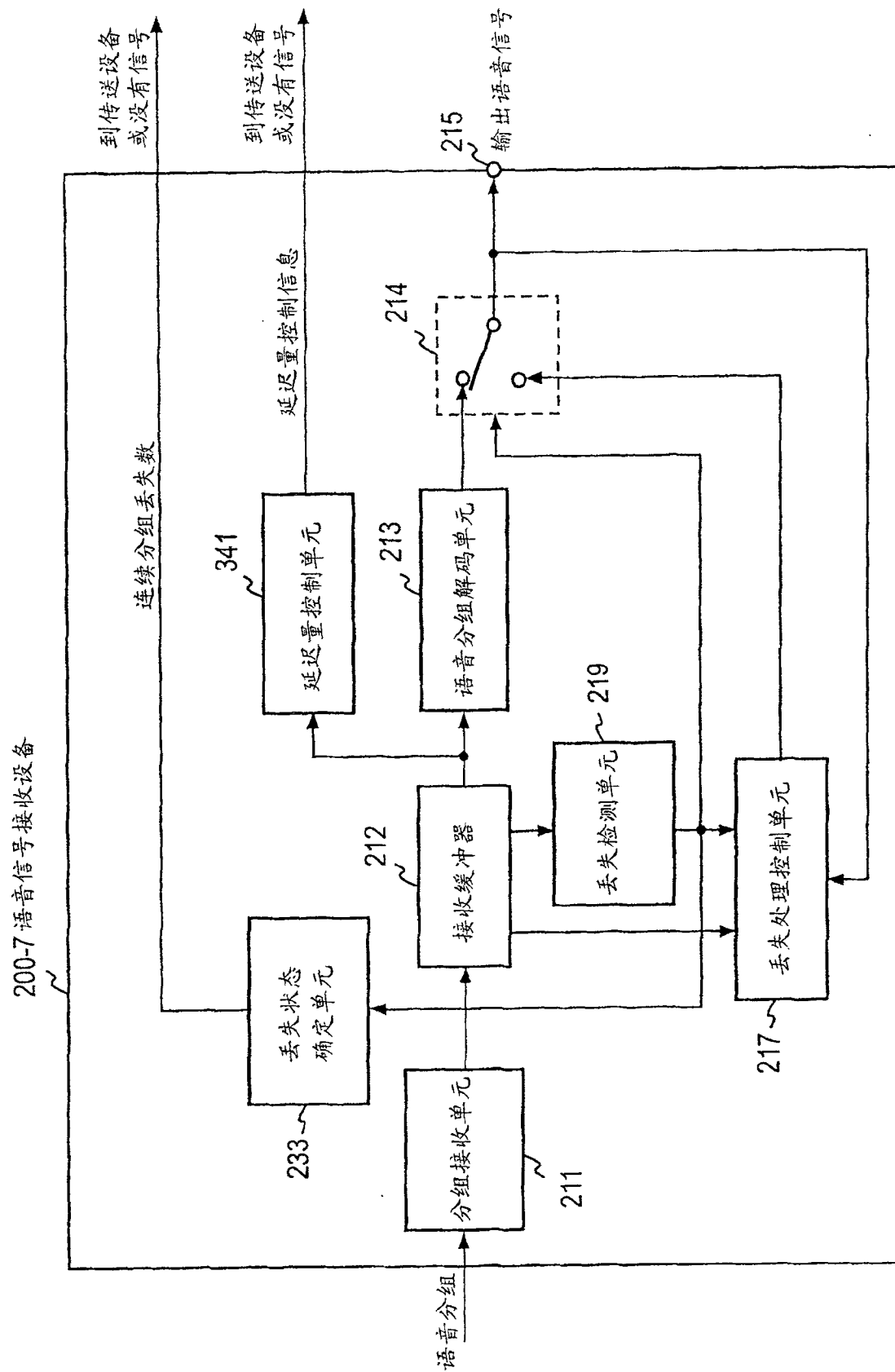


图 24

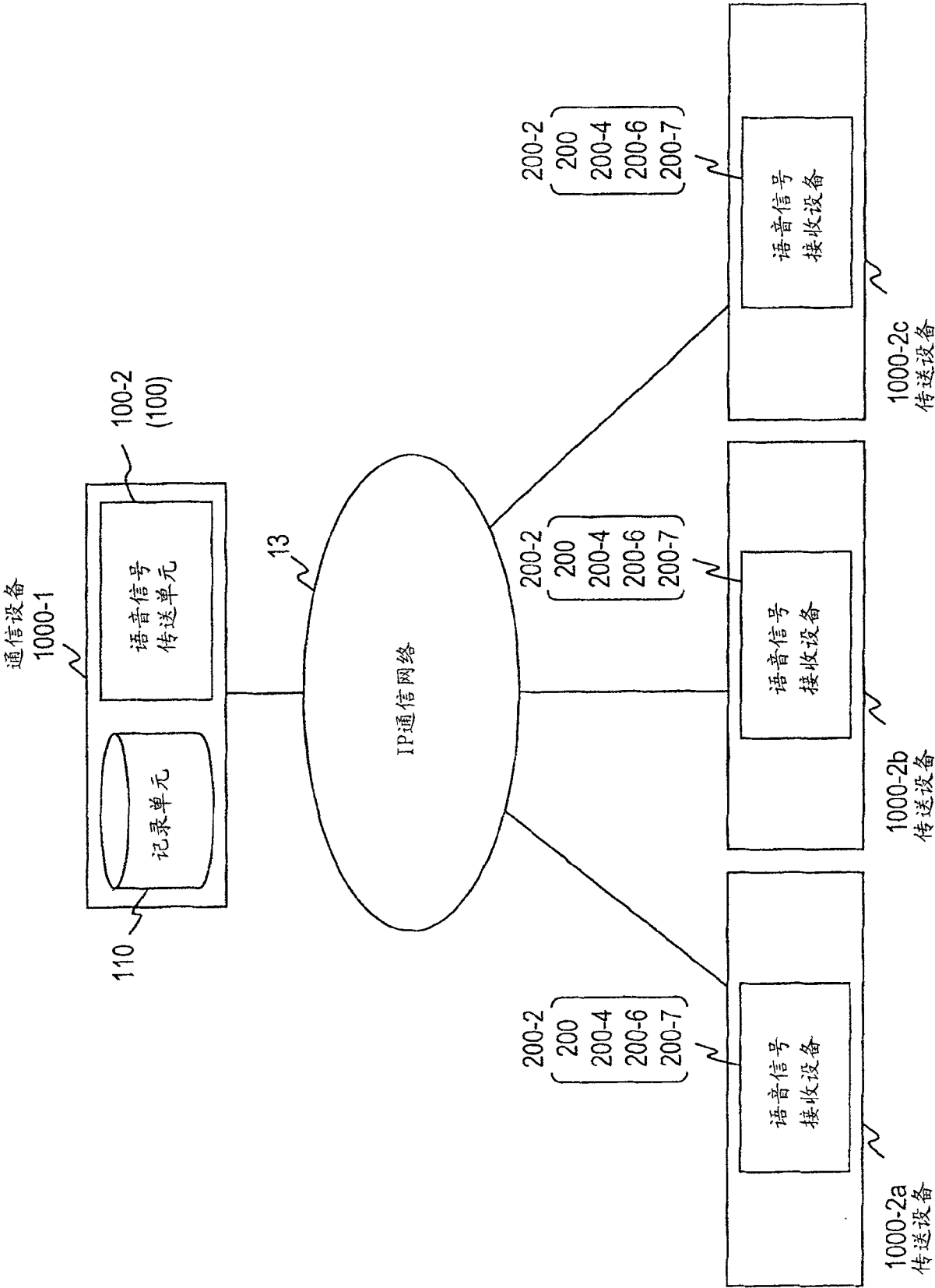


图 25

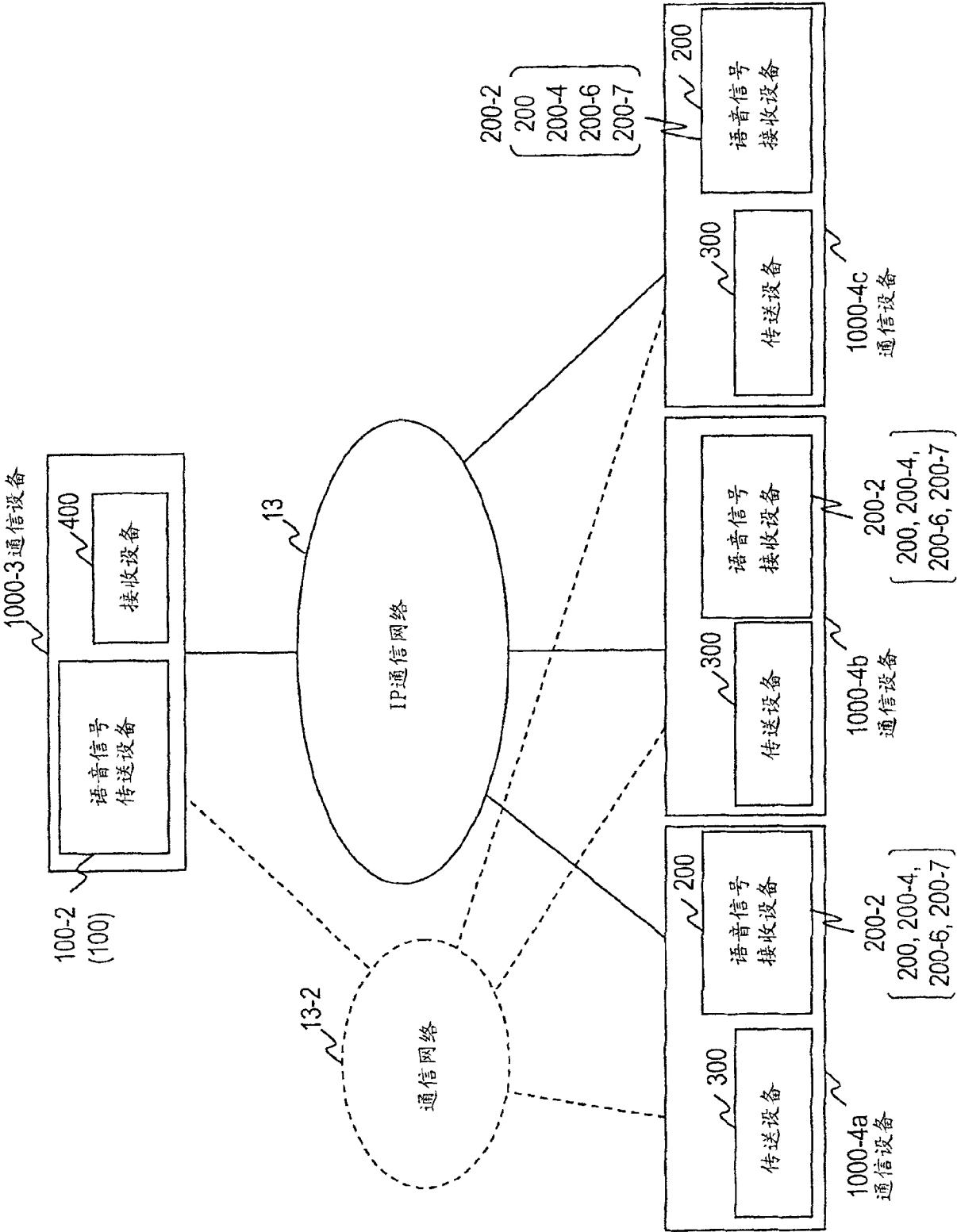


图 26

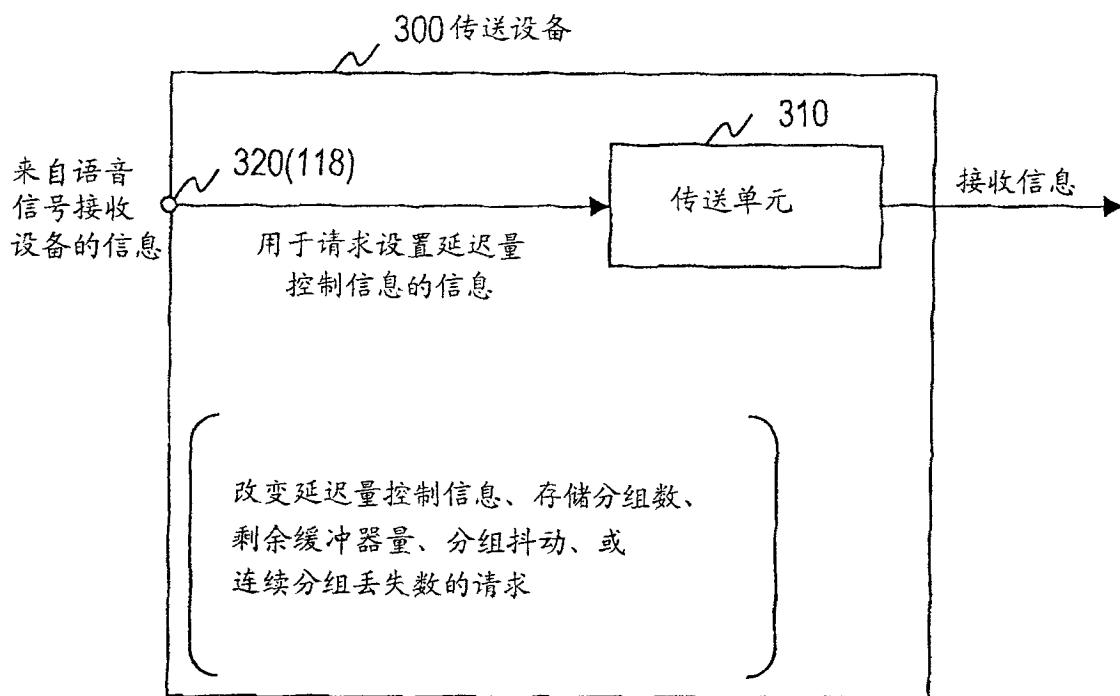


图 27

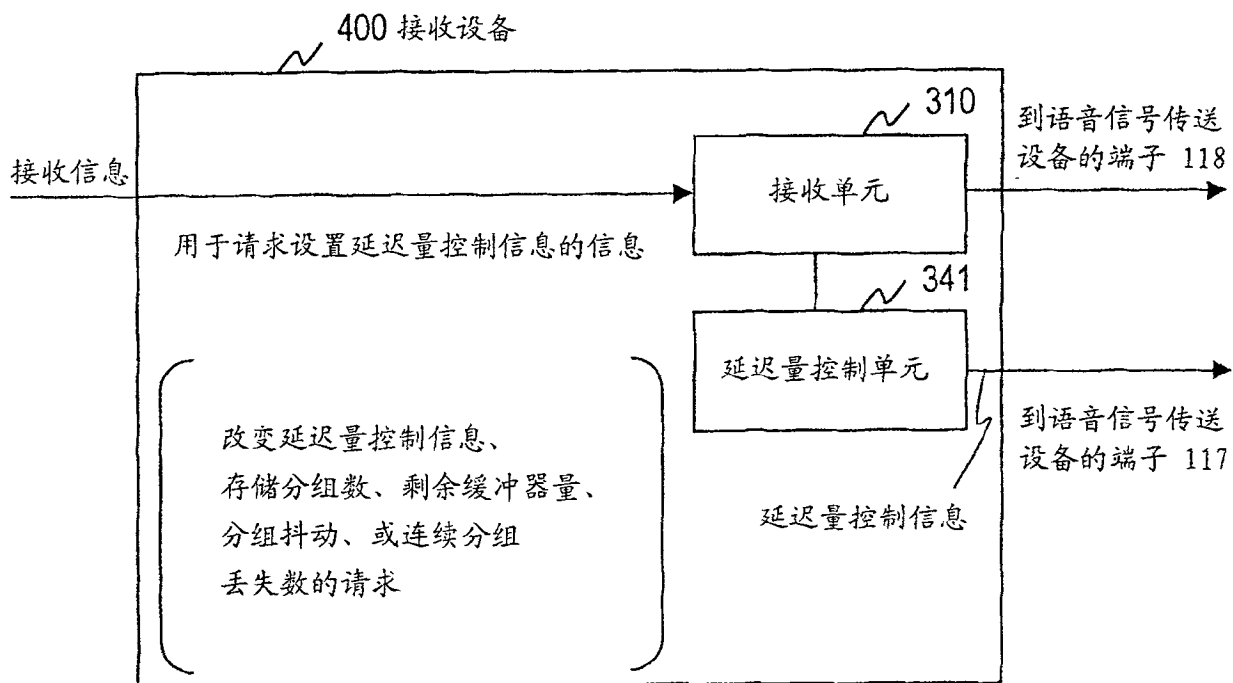


图 28

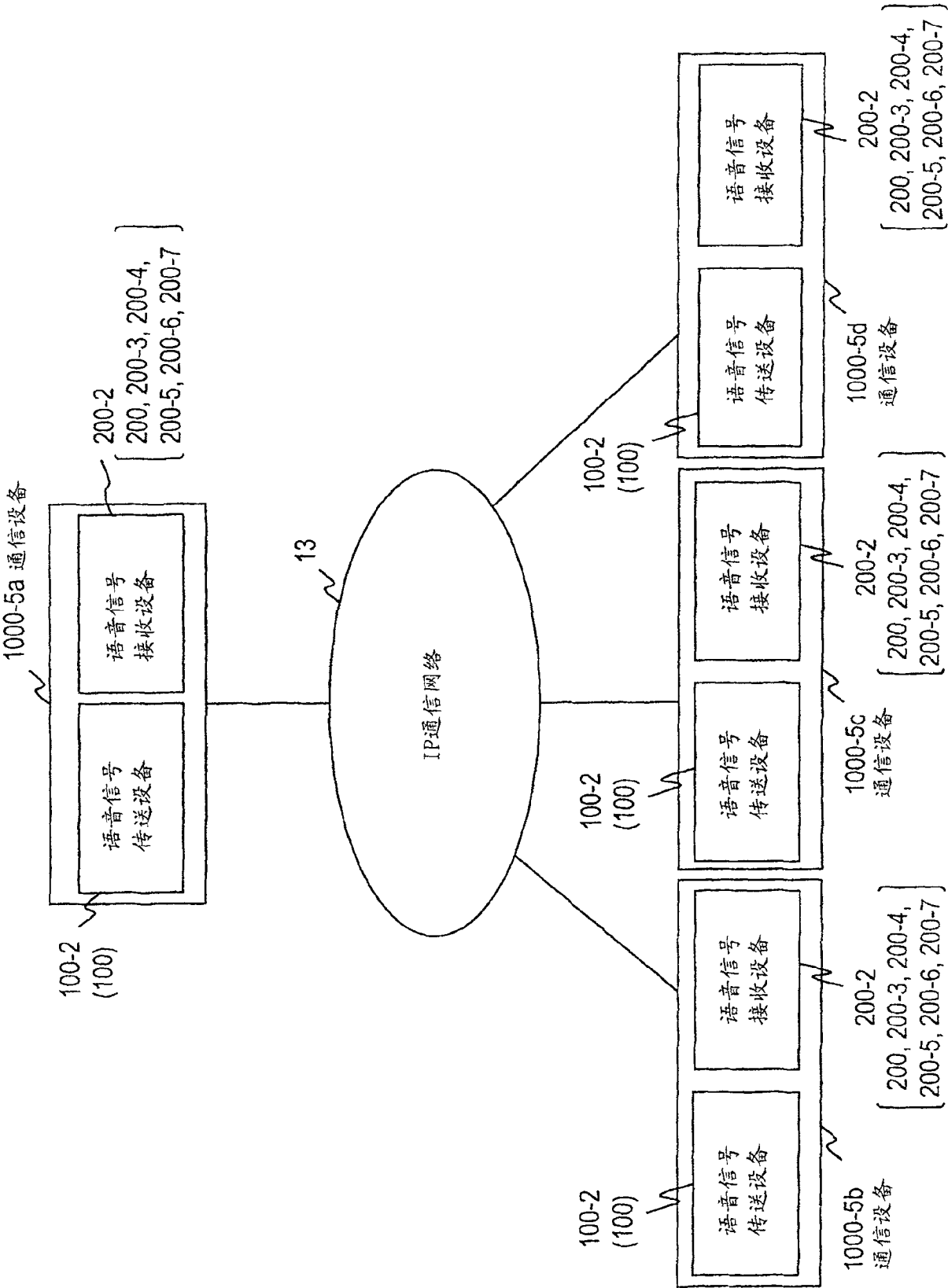


图 29

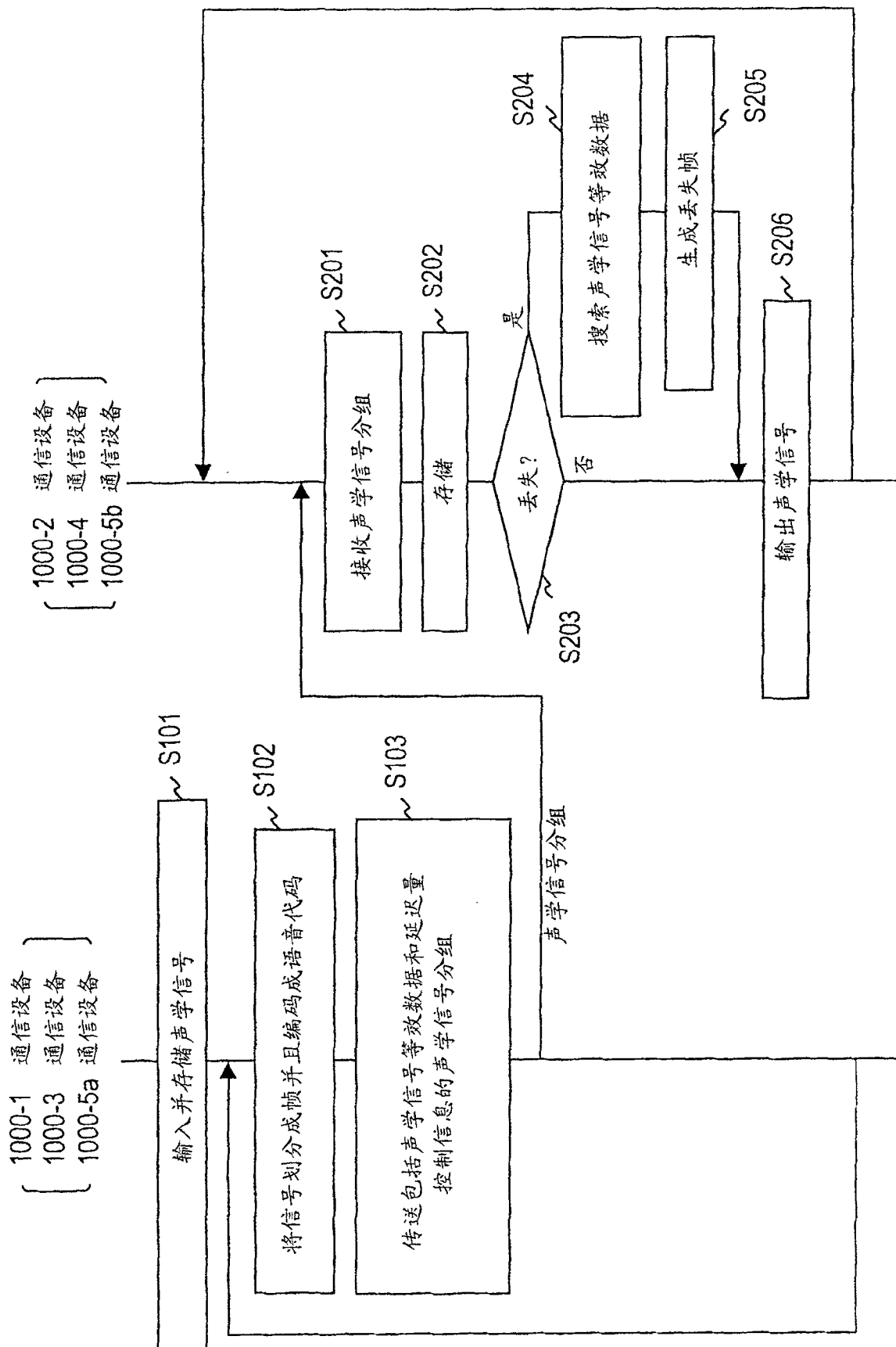


图 30

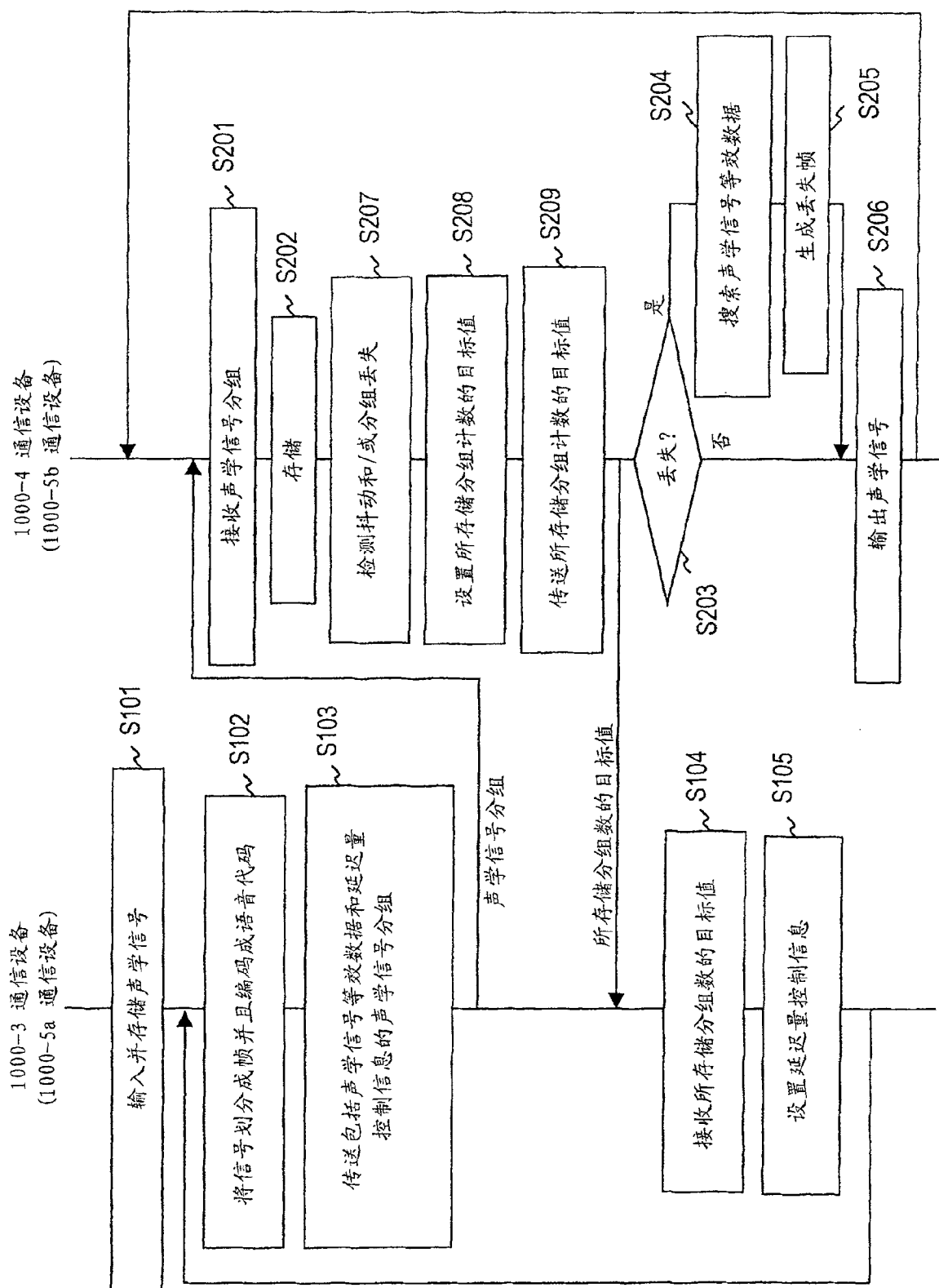


图 31

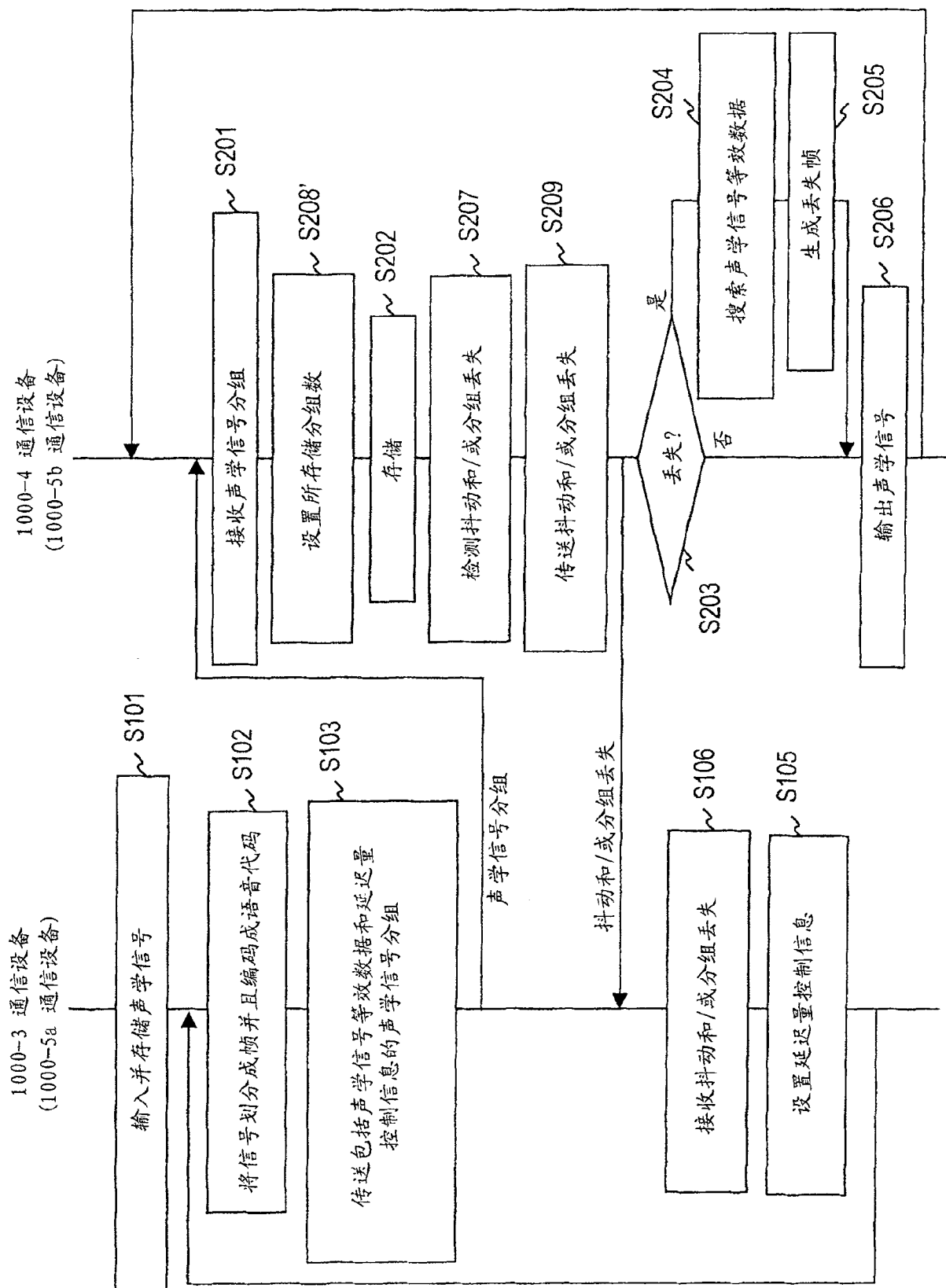


图 32

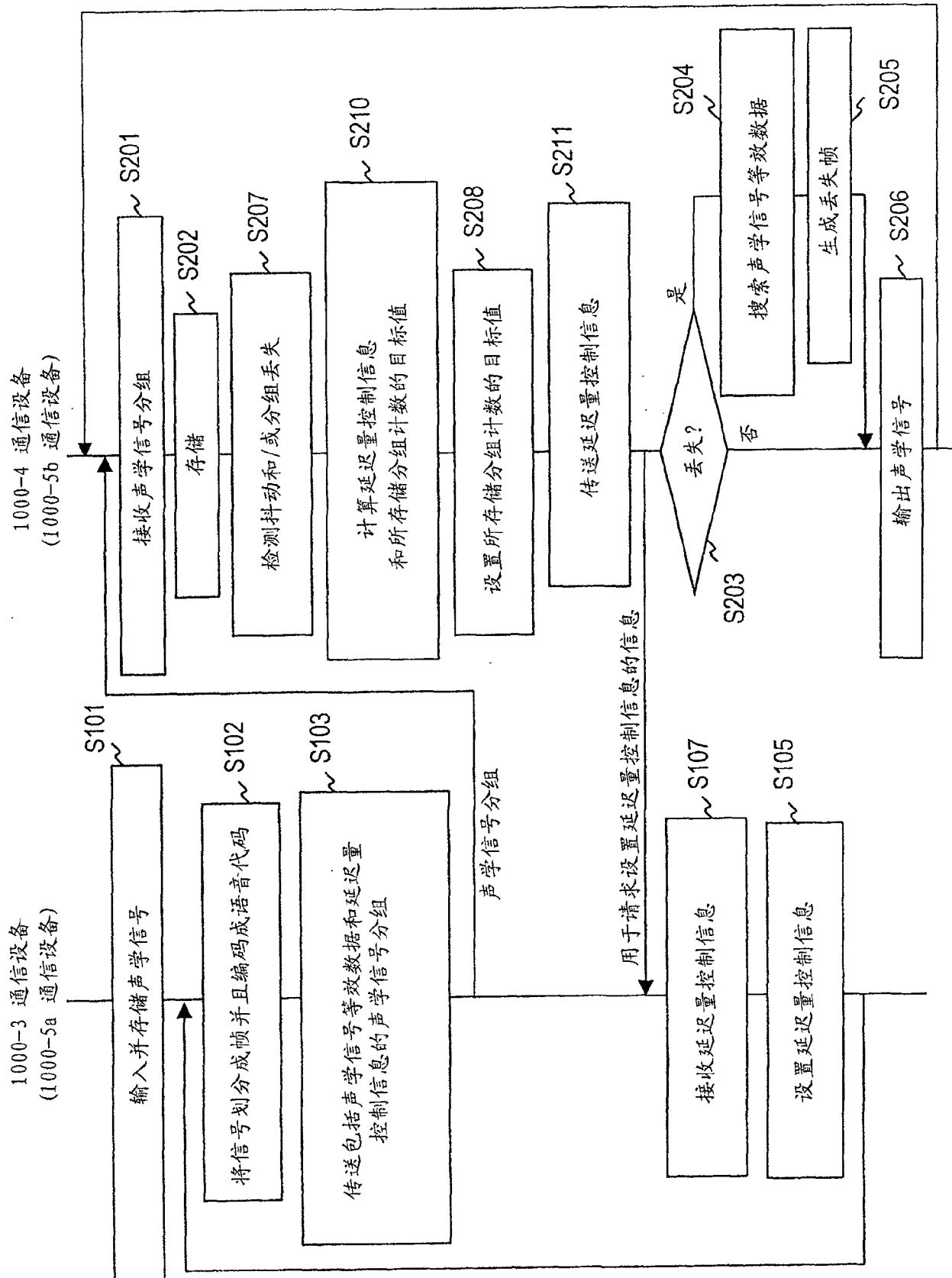


图 33

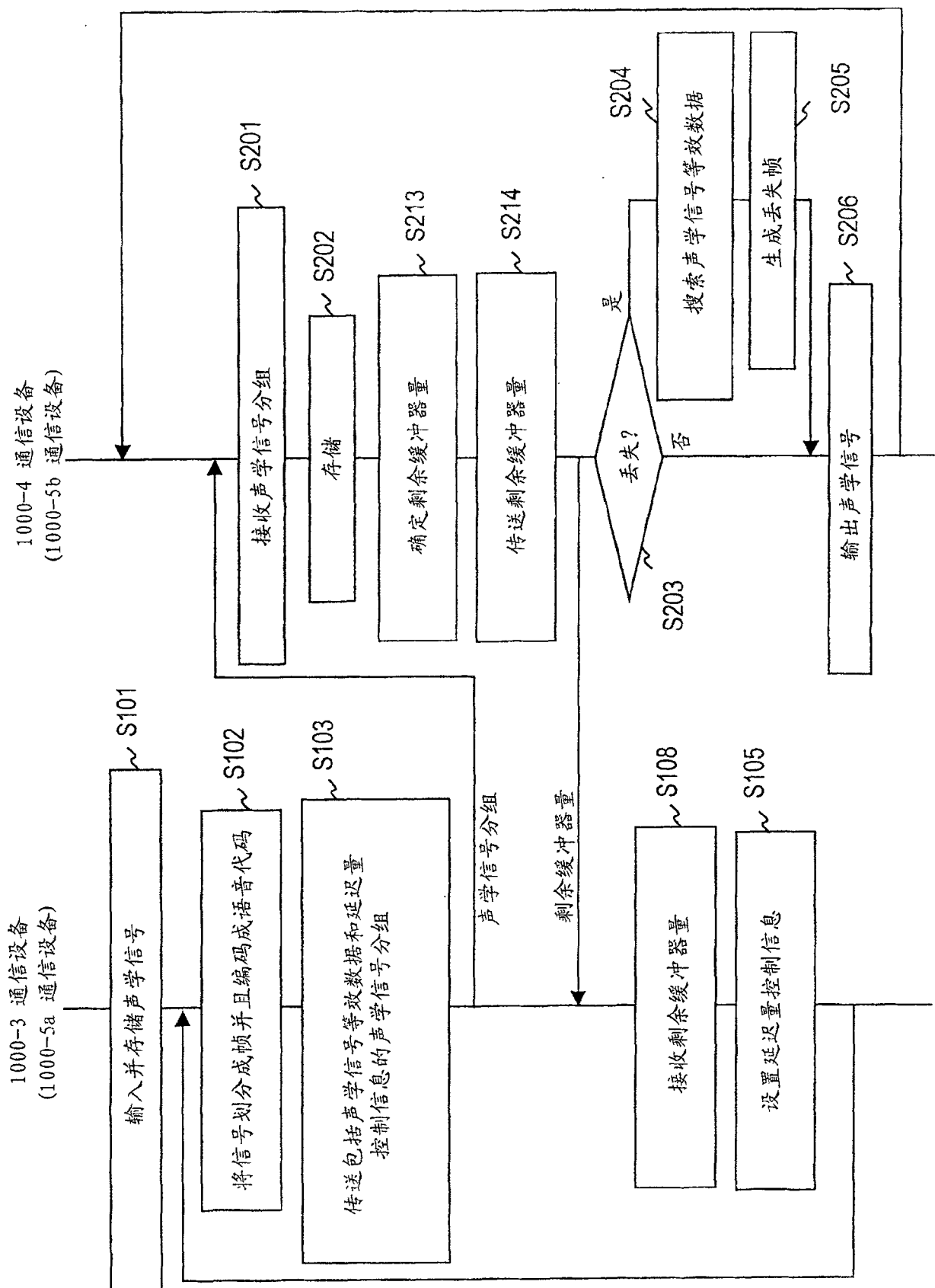


图 34

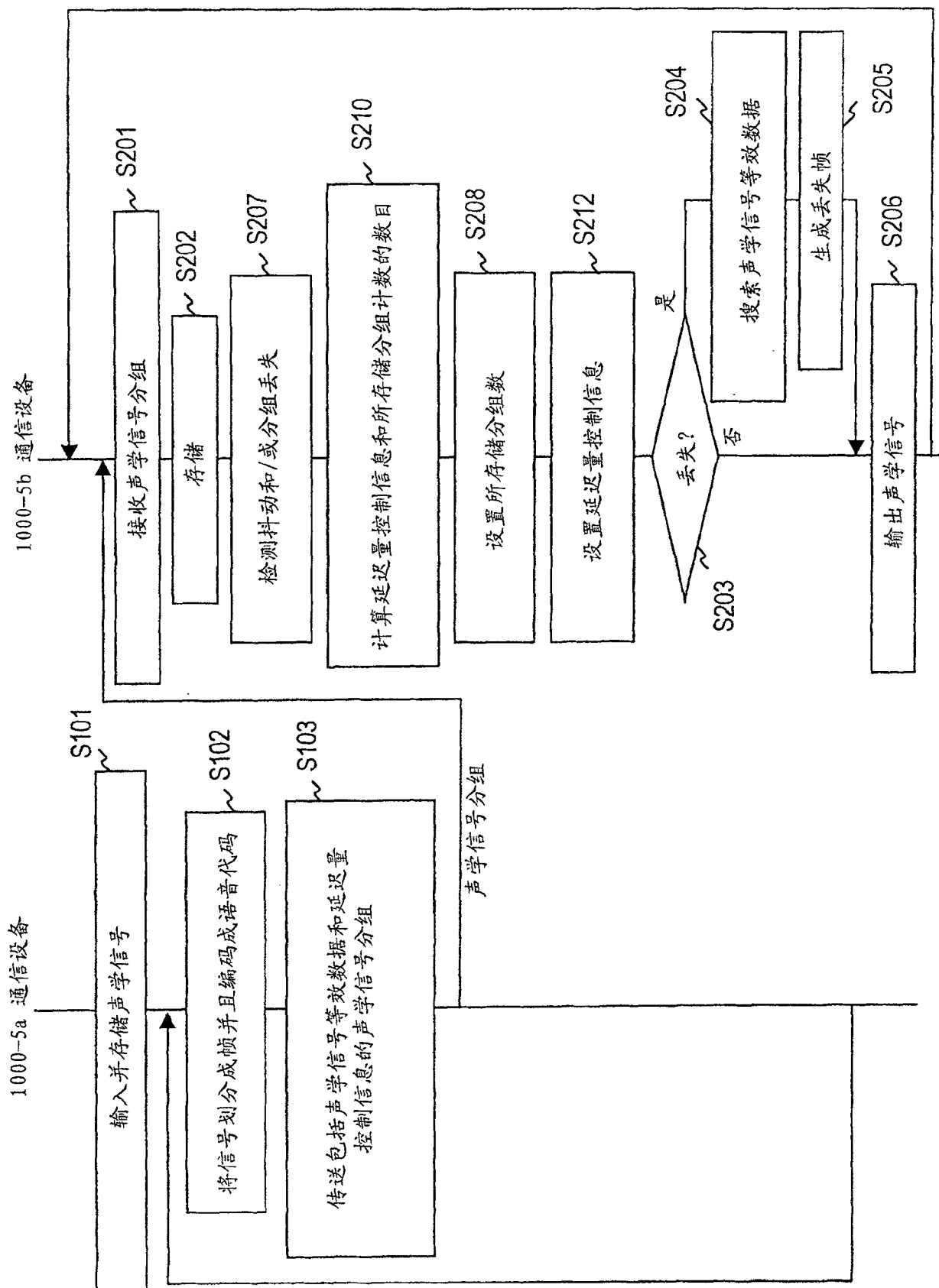


图 35