



# [12] 发明专利说明书

[21] ZL 专利号 97104573.9

[43] 授权公告日 2003 年 4 月 16 日

[11] 授权公告号 CN 1106085C

[22] 申请日 1997.3.28 [21] 申请号 97104573.9

[30] 优先权

[32] 1996. 4. 26 [33] EP [31] 96106678.4

[71] 专利权人 德国汤姆逊-布朗特公司

地址 德国菲林根-施文宁根

[72] 发明人 詹斯·斯皮尔

审查员 葛源

[74] 专利代理机构 北京市柳沈律师事务所

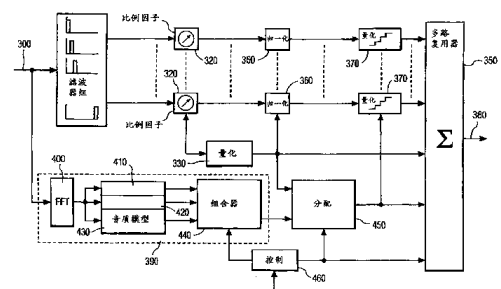
代理人 吕晓章

权利要求书 3 页 说明书 6 页 附图 4 页

[54] 发明名称 对数字音频信号编码的方法和装置

[57] 摘要

一种对数字音频信号编码的方法，该方法包括选择两个或多个音质模型(410, 420, 430)中之一的步骤，该音质模型是用于产生在数据整理过程中所用的屏蔽阈值。选择准则是已编码位流的可利用数据速率(ADR)。两个或多个音质模型中的每一个适用于已编码位流的特定数据速率。在本发明方法第二实施例中，包括将从不同音质模型(410, 420, 430)得出的两个或多个屏蔽阈值组合的步骤，从而产生对数据整理过程屏蔽阈值的更精确计算。此外提出了对数字音频信号编码的适当装置。



ISSN 1008-4274

- 1、一种用于数字音频信号的数据整理编码的方法，包括步骤：  
将音频信号带划分为多个子带；
- 5 基于至少两个不同的音质模型确定用于所述音频信号的一个时间段的至少两个屏蔽阈值；  
组合所述至少两个屏蔽阈值以提供一个最终屏蔽阈值；和  
在单独的位分配过程中利用所述最终屏蔽阈值，其中通过子带中的所述最终阈值确定用于子带样本的量化的位分辨率。
- 10 2、如权利要求1所述的方法，其中，所述屏蔽阈值的组合是加权组合，并且根据用于编码位流的可利用数据速率确定加权因子。
- 3、如权利要求1所述的方法，其中，所述加权组合是按照下列公式的线性组合：
- 15 
$$\text{Mask}=\alpha*\text{Mask1}+\beta*\text{Mask2}$$
- 其中 Mask 是最终的屏蔽阈值，  
Mask1 是从第一音质模型得出的屏蔽阈值，  
Mask2 是从第二音质模型得出的屏蔽阈值，
- 20  $\alpha$ 是 Mask1 的加权因子，以及  
 $\beta$ 是 Mask2 的加权因子。
- 4、如权利要求1所述的方法，其中，对于两个或多个音质模型中的每一个，提供适当的曲线，该曲线具有独立变量可利用数据速率并作为数值在‘0’与‘1’之间的函数值，用于根据可利用数据速率确定加权因子。
- 25 5、一种用于对数字音频信号编码的装置，包括用于整理音频信号的数字数据的数据整理装置，该音频信号的数字数据是在考虑音质效果的情况下确定屏蔽阈值所造成的，包括：  
提供两个或多个音质模型装置来计算各屏蔽阈值，每个音质模型装置适用于已编码位流的特定数据速率，和
- 30 组合装置用于组合两个或多个音质模型装置产生的屏蔽阈值，由此得出由数据整理装置使用的最终屏蔽阈值。

6、如权利要求 5 所述的装置，其中提供加权装置用于在将两个或多个屏蔽阈值组合前对它们进行加权。

7、如权利要求 5 所述的装置，其中所述加权装置根据已编码位流的可利用数据速率确定加权因子。

5 8、一种用于数字音频信号的数据整理编码的方法，包括步骤：

将音频信号带划分为多个子带；

基于至少两个不同的音质模型确定用于所述音频信号的一个时间段的至少两个屏蔽阈值；

组合所述至少两个屏蔽阈值以产生一个最终屏蔽阈值；

10 其中，进行加权组合，并且确定加权因子，所述加权因子取决于已编码位流的可利用数据速率，

在单独的位分配过程中利用所述最终屏蔽阈值，其中通过子带中的所述最终阈值确定用于子带样本的量化的位分辨率。

15 9、如权利要求 8 所述的方法，其中，所述加权组合是按照下列公式的线性组合：

$$\text{Mask}=\alpha*\text{Mask1}+\beta*\text{Mask2}$$

其中 Mask 是最终的屏蔽阈值，

20 Mask1 是从第一音质模型得出的屏蔽阈值，

Mask2 是从第二音质模型得出的屏蔽阈值，

$\alpha$  是 Mask1 的加权因子以及，

$\beta$  是 Mask2 的加权因子。

10、如权利要求 8 所述的方法，其中，对于两个或多个音质模型中的每一个，提供适当的曲线，该曲线具有独立变量可利用数据速率并作为数值在  
25 ‘0’ 与 ‘1’ 之间的函数值，用于根据可利用数据速率确定加权因子。

11、一种编码数字音频信号的装置，包括：

滤波器组，将所述音频信号划分为多个子带；

30 数据整理装置，用于整理所述音频信号的数字数据，该音频信号的数字数据是响应于音质属性确定的屏蔽阈值所造成的；

至少两个音质模型装置，被提供用于计算各屏蔽阈值，每个音质模型装

置适用于已编码位流的特定数据速率；

组合装置，组合所述计算的屏蔽阈值，提供由所述数据整理装置使用的最终屏蔽阈值；

加权装置，用于在将至少两个屏蔽阈值组合前对它们进行加权，

- 5 所述加权装置响应所述已编码位流的可利用数据速率确定加权因子；和  
位分配单元，输入按照在子带中的所述最终屏蔽阈值确定用于子带样本的量化的位分辨率的所述最终屏蔽阈值。

## 对数字音频信号编码的方法和装置

5 本发明涉及用于对数字化的音频信号进行编码的方法和装置。

在音频编码器中，例如在 MPEG 1 或 MPEG 2 音频系统中，音质模型计算用作数据整理的屏蔽阈值。对于音频信号的每个子带屏蔽阈值可改变。归因于所确定的屏蔽阈值对每个子带定义所谓的位分配过程，子带抽样应以该分辨率被量化。这将随着时间的推移随每块间频信号样本而变。因此，对于不同的块最佳音质所需的位率也应改变。实际上在 MPEG1 音频系统中，可利用的所需位率对于每个音频通道而言通常是固定的。因为，对于所需位率，各种参数的准确位数和传输所需要的数据不可能事先精确地估计到，一些位被剩下或者一些位被丢失。习惯上用一种简单的分配策略克服该问题，例如，为每个子带分配几乎相等数目的这种位。

15 上述分配策略未考虑在编码器中所使用的音质模型。目前，对于小范围的位率编码器被加以优化。虽然这种音频编码器也被用于在大范围位率下操作，但是它们没有为这样做而被优化。不同位率可能出现在左和右通道之间和/或附加通道(左环绕、右环绕、中央)和/或在所需位率可能从帧到帧改变的 MPEG 层 III 中。

20 本发明的一个目的是公开一种根据这种分配策略获得较好音频编码质量的方法。

本发明的另一目的是公开一种采用本发明方法的装置。

根据本发明的一个方面，提供一种用于数字音频信号的数据整理编码的方法，包括步骤：将音频信号带划分为多个子带；基于至少两个不同的音质模型确定用于所述音频信号的一个时间段的至少两个屏蔽阈值；组合所述至少两个屏蔽阈值以提供一个最终屏蔽阈值；和在单独的位分配过程中利用所述最终屏蔽阈值，其中通过子带中的所述最终阈值确定用于子带样本的量化的位分辨率。

30 根据本发明的另一方面，提供一种用于对数字音频信号编码的装置，包括用于整理音频信号的数字数据的数据整理装置，该音频信号的数字数据是在考虑音质效果的情况下确定屏蔽阈值所造成的，包括：提供两个或多个音

质模型装置来计算各屏蔽阈值，每个音质模型装置适用于已编码位流的特定数据速率，和组合装置用于组合两个或多个音质模型装置产生的屏蔽阈值，由此得出由数据整理装置使用的最终屏蔽阈值。

5 根据本发明的另一方面，提供一种用于数字音频信号的数据整理编码的方法，包括步骤：将音频信号带划分为多个子带；基于至少两个不同的音质模型确定用于所述音频信号的一个时间段的至少两个屏蔽阈值；组合所述至少两个屏蔽阈值以产生一个最终屏蔽阈值；其中，进行加权组合，并且确定加权因子，所述加权因子取决于已编码位流的可利用数据速率，在单独的位分配过程中利用所述最终屏蔽阈值，其中通过子带中的所述最终阈值确定用于子带样本的量化的位分辨率。

10 根据本发明的另一方面，提供一种编码数字音频信号的装置，包括：滤波器组，将所述音频信号划分为多个子带；数据整理装置，用于整理所述音频信号的数字数据，该音频信号的数字数据是响应于音质属性确定的屏蔽阈值所造成的；至少两个音质模型装置，被提供用于计算各屏蔽阈值，每个音质模型装置适用于已编码位流的特定数据速率；组合装置，组合所述计算的屏蔽阈值，提供由所述数据整理装置使用的最终屏蔽阈值；加权装置，用于在将至少两个屏蔽阈值组合前对它们进行加权，所述加权装置响应所述已编码位流的可利用数据速率确定加权因子；和位分配单元，输入按照在子带中的所述最终屏蔽阈值确定用于子带样本的量化的位分辨率的所述最终屏蔽阈值。

15 在本发明的第一实施例中，用至少三个音质模型代替一个，其中每一个音质模型相对一定位率或位率范围被优化。根据所需通道或帧位率该编码器选择适当的一个音质模型。该音质模型之一的输出可以认为是某一所需位率的一个标准屏蔽阈值。

25 当当前位数高时，采用总体上具有较高屏蔽阈值的第二音质模型。这一较高阈值不是象现在技术中一样在频谱上等量上升的标准阈值，而是相对于仍适用本输入频谱的音质模型上升。

30 当当前位数低时，采用总体上具有较低屏蔽阈值的第三音质模型。这一较低阈值也不是象现在技术中一样在频谱上等量下降的标准阈值，而是相对于仍适用本输入频谱的音质模型下降。

本发明的第二实施例解决了在不同音质模型之间的硬转换问题，例如在

左和右通道之间和/或在所需数据率可从帧到帧改变的 MPEG 层 I 到层 III 中。代替采用二进制逻辑转换，执行一种模糊逻辑。

原理上，本发明方法适用于在 MPEG 2 层 I 到层 III 和/或 MPEG 2 NBC 系统中的音频编码。

5 下面参照附图描述本发明的优选实施例，附图中：

图 1 是表示三个屏蔽阈值之间的‘硬转换’方法的示图；

图 2 是采用‘硬转换’方法的音频编码装置的程序结构；

图 3 是表示三个屏蔽阈值之间的‘软转换’方法的示图；

图 4 是采用‘硬转换’方法的音频编码装置的程序结构；

10 图 5 是表示音频编码装置原理结构的方框图。

关于音频编码技术的有价值信息包括在国际标准 ISO/IEC 11172-3, 1993 年 8 月 1 日第一版‘信息技术 - 数字存储介质运动图像和相关音频以高达大约 1.5M 位/秒的编码’，第三部分音频中。

15 本发明的方法采用一个以上的音质模型。作为一个实例三个不同的音质模型被使用。每一个音质模型适合于音频位流的特定预定数据率。例如具有三个不同的数据率 32 千比特/秒、56 千比特/秒和 80 千比特/秒。不同的音质模型简称为 PAM\_Low, PAM\_Mid, PAM\_High。PAM\_Low 适合于 32 千比特/秒的低数据率，PAM\_Mid 适合于 56 千比特/秒的中数据率和 PAM\_High 适合于 80 千比特/秒的高数据率。

20 现在参见图 1，它示出了在哪些区域中音质模型被采用。在图 1 中的 X 轴上规定了能够为音频编码器所采用的可能数据率的范围。在 Y 轴上对于每一可能的数据率给出了每一音质模型的份额。当可利用数据率在范围 DRR1 内时。音质模型 PAM\_Low 被用于计算屏蔽阈值。因此分配因子是 1，即 PAM\_Low 为 100%，而 PAM\_Mid 和 PAM\_High 为 0%。当可利用数据率在范围 DRR2 内时，音质模型 PAM\_Mid 被用于计算屏蔽阈值。因此，此时分配因子是 1，即 PAM\_Mid 为 100%，而 PSM\_Low 和 PAM\_High 为 0%。当可利用数据率在范围 DRR3 内时，音质模型 PAM\_High 被用于计算屏蔽阈值。因此分配因子再次为 1，即 PAM\_High 为 100%，而 PAM\_Low 和 PAM\_Mid 为 0%。在不同音质模型之间的这种转换从现在开始将称之为‘硬转换’。

30 在图 2 中示出了按照该方法工作的一种合适的简单计算机程序，现在对其进行说明。

标号 10 表明该程序的起始点。标号 20 表明一判定步骤，在其中确定可利用数据率 ADR 是否小于图 1 所示的数据率值 DR1。如果是这样，该程序进行到步骤 30。在其中利用音质模型 PAM\_Low 确定屏蔽阈值。在这一步骤之后，该程序将在步骤 40 结束。如果在判定步骤 20 判定可利用数据率 ADR 5 大于数据率值 DR1，则在判定步骤 50 进行分析，可利用数据率 ADR 是否小于数据率值 DR2。在正判定的情况下，该程序执行步骤 60 并用音质模型 PAM\_Mid 计算屏蔽阈值。在步骤 50 作负判定的情况下，该程序执行判定步骤 70。在此分析可利用数据率是小于还是等于数据率值 DR3。如果是这样，该程序执行步骤 80，其中用音质模型 PAM\_High 计算屏蔽阈值。在判定步骤 10 70 中回答为负的情况下，该程序执行步骤 90，该步骤包括适当的错误检测算法。在程序步骤 60、80、90 之后该程序将在步骤 40 结束。

下面将说明一种改进的称为‘软转换’的不同音质模型之间的转换方法。在图 3 中，示出了三种可利用数据率的不同范围，表示为 DRR4、DRR5、DRR6。如果可利用数据率在范围 DRR4 内，则将仅采用音质模型 15 PAM\_Low，即以 100%的份额计算屏蔽阈值。如果可利用数据率在范围 DRR7 内，则仅采用音质模型 PAM\_High。如果可利用数据率在范围 DRR5 内，则采用两个音质模型 PAM\_Low 和 PAM\_Mid 来计算相应的屏蔽阈值。在可利用数据率在范围 DRR6 内的情况下，采用两个音质模型 PAM\_Mid 和 PAM\_High 来计算屏蔽阈值。例如在图 3 中示出了可利用数据率 ADR 在范 20 围 DRR5 中。这意味着两个屏蔽阈值被计算，即第一个称为 Mask\_Low 的对应于 PAM\_Low 模型，第二个称为 Mask\_Mid 的对应于 PAM\_Mid 模型。将两个屏蔽阈值组合得到屏蔽阈值 Mask。用按照下列公式的平均算法计算该新屏蔽阈值：

$$25 \quad \text{Mask} = \alpha * \text{Mask\_Mid} + \beta * \text{Mask\_Low}$$

公式 1：平均算法的实例

其中 $\alpha$ 是音质模型 PAM\_Mid 的分配因子， $\beta$ 是音质模型 PAM\_Low 的分配因子。在图 3 中示出了这两个分配因子。通过查表来确定它们，该表中包含代表在可利用数据率 ADR 的位置处相应曲线 L 和 M 的数值。利用这些技术有 30 可能设计一种能够以宽的数据率范围产生位流的音频编码器。



现在参照图 4 说明执行所述软转换方法的一种合适的计算机程序。标号 100 表明该计算机程序的开始步骤。在下一步骤 110 分析可利用数据率 ADR 是否小于数据率 DR4。正的结果导致步骤 120 作为下一程序步骤。在此，仅根据音质模型 PAM\_Low 计算屏蔽阈值。在这一步骤之后该程序将在步骤 130 5 结束。在步骤 110 中为负的结果的情况下，该程序继续进行步骤 140。在此确定可利用数据率是否小于数据率值 DR5。如果情况是这样，则在步骤 150 计算两个屏蔽阈值。一个是与音质模型 PAM\_Low 一致，另一个是与音质模型 PAM\_Mid 一致。然后将两个屏蔽阈值组合得到一个最终屏蔽阈值。为此，两个屏蔽阈值如图 3 中所示用它们的相应加权因子 $\alpha$ 和 $\beta$ 加权。如上所述该组 10 合运算利用上述公式 1。在步骤 140 中负的结果的情况下，下一步骤是判定步骤 160。在此确定可利用数据率是否小于数据率值 DR6。当确实如此时，下一步骤是步骤 170。在此根据相应音质模型 PAM\_Mid 和 PAM\_High 计算两个屏蔽阈值 Mask\_Mid 和 Mask\_High。同样它们也如同对步骤 150 所描述的那样被加权并组合。判定步骤 160 中的负的判定结果导致判定步骤 180。 15 在此分析可利用数据率 ADR 是小于还是等于数据率值 DR7，如图 3 所示。在正的判定之后将执行程序步骤 190。在此根据音质模型 PAM\_High 计算屏蔽阈值。步骤 180 中负的判定导致步骤 200，其中执行错误检测算法。在执行了步骤 150、170、190 和 200 之一以后该程序在步骤 130 结束。

下面参照图 5 描述利用‘软转换’方法对音频信号编码的装置。 20 该装置是基于众所周知的子带编码技术。数字化的音频信号经由数据总线 300 馈送到滤波器组 310。为简单起见假定仅传输一个音频通道(单)到编码装置。当然音频信号是按块馈送到编码装置的。在滤波器组 310 中，音频信号被划分为 32 个子带。子带样本被馈送到块 320。在此，计算比例因子。它们被用于数据整理即用于压缩子带样本的引导位‘0’和‘1’。每一子带块 25 的比例因子被馈送到比例因子量化装置 330。在此它们被以例如 6 比特的分辨率量化。该量化比例因子被传送到多路复用器 350。它的功能将在后面描述。子带样本经过块 320 后被输入归一化电路 360。在此用根据量化比例因子确定的适当因子  $1/U_s$  乘以该样本，量化比例因子是经由总线 340 输入的。归一化的子带样本被馈送到量化电路 370。在此执行基于音质效果的数据整 30 理。为量化目的的位分辨率根据音质模型而确定，这将在后面描述，因此量化噪声不会被人耳感觉到。在经过量化电路 370 之后，子带样本被传送到多

路复用器 350。多路复用器 350 建立包括所有同步和辅助数据的编码装置的输出位流。该位流经由总线 380 输出。

经数据总线 300 输入的位流被馈送到块 390。其中在块 400 中通过 FFT 算法(快速富里叶变换)计算信号的频谱。此外在块 390 中有三个不同的音质模型 410 至 430。对于低位率例如 32 千比特/秒的特定输出位流采用并优化第一个音质模型 PAM\_Low。对于中等位率例如 56 千比特/秒的输出位流采用并优化第二个音质模型。对于高位率例如 80 千比特/秒的输出位流采用并优化第三个音质模型。每个音质模型能够确定一个屏蔽阈值 Mask\_Low, Mask\_Mid, Mask\_High。

10 在 MPEG1 和 MPEG2 标准中, 存在在一个宽的位率范围内(任意格式)预先规定位率的可能性。根据预先规定的位率值, 在‘硬转换’模式下, 用一个音质模型计算屏蔽阈值。这已经参照图 1 和 2 进行了描述。在‘软转换’模式下, 可采用两个不同的音质模型。如参照图 3 和 4 所描述的这取决于预先规定的位率值。这两个屏蔽阈值被传送到组合装置 440, 在此它们被组合。

15 所得到数据被馈送到位分配单元 450。在‘硬转换’模式下, 组合装置 450 用于作为选择装置来选择适当的屏蔽阈值数据并将其传送到位分配单元 450。该位分配单元计算位分配并控制量化电路 370。位分配数据被传送到多路复用器 350。控制电路 460 预先规定可利用的数据率, 例如与一个用户的输入相一致, 并据此控制组合装置 440 和位分配单元 450。它也输入辅助数

20 据到多路复用器 350。

本发明能够应用在几乎所有的音频系统中, 象 MPEG、DAB、DCC、MD。

当用一种视觉模型代替音质模型时, 本发明也可用于象 MPEG1 或 MPEG2 这样的视频编码。

25 所述实施例的各种改型是可能的, 例如图 3 中所示的曲线 L、M 和 H 可以是不同的形状。音质模型的数量也可以不是三个。图 5 中所示的某些块可以用适当的计算机程序实现, 该计算机程序可以由标准的或特殊的微处理器来执行。为了组合‘软转换’模式下的屏蔽阈值, 可以用稍微改进的公式来代替公式 1。也可以组合两个以上的屏蔽阈值。

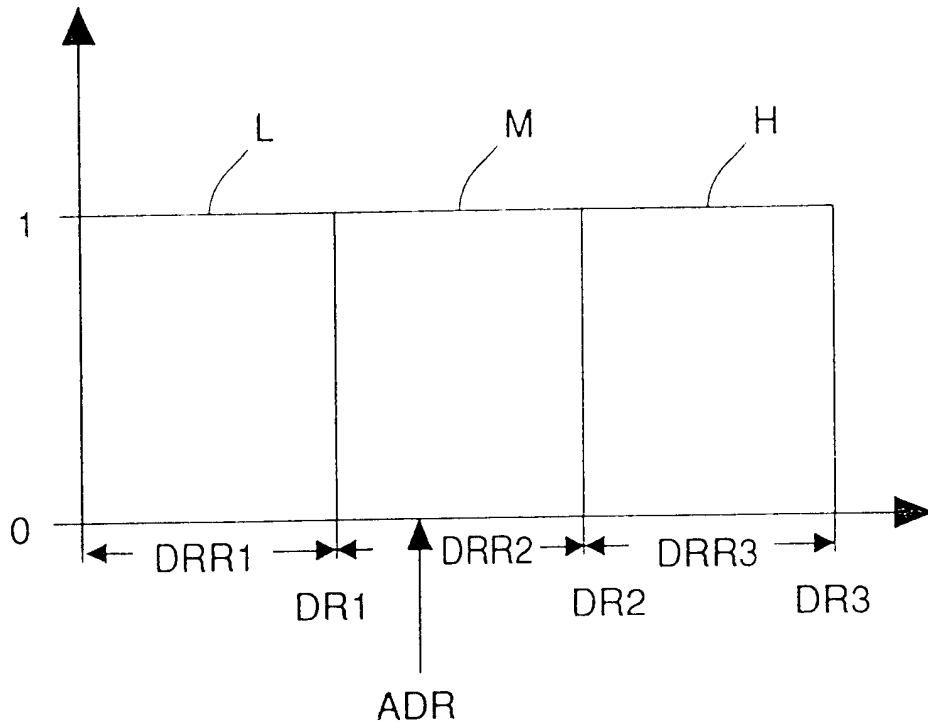


图 1

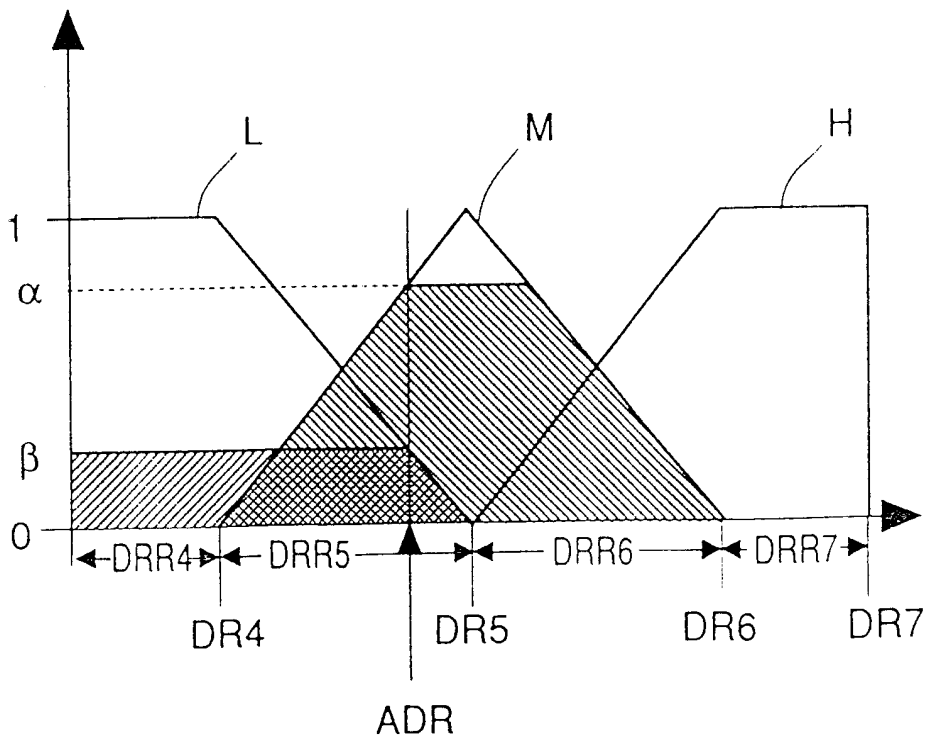


图 3

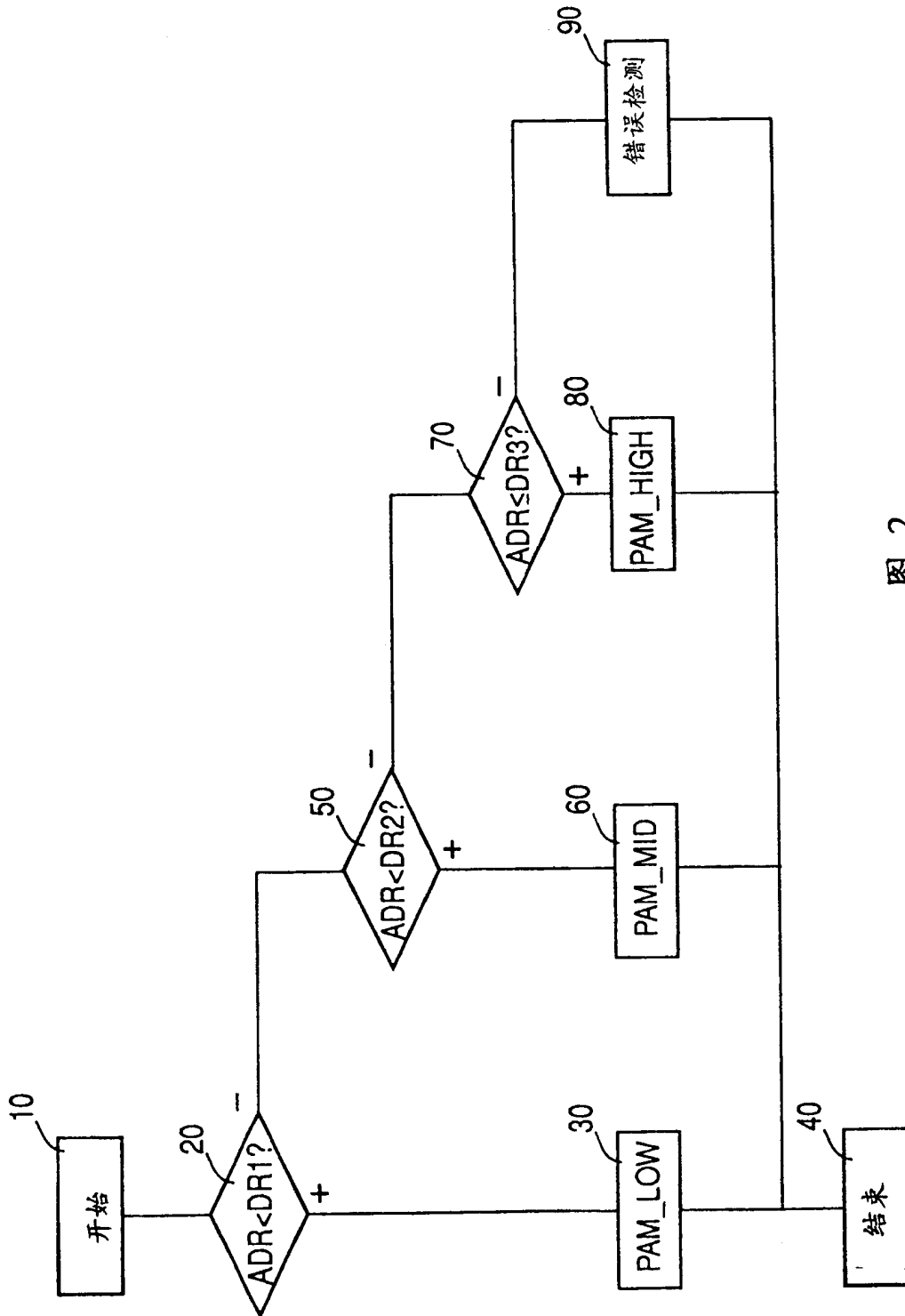


图 2

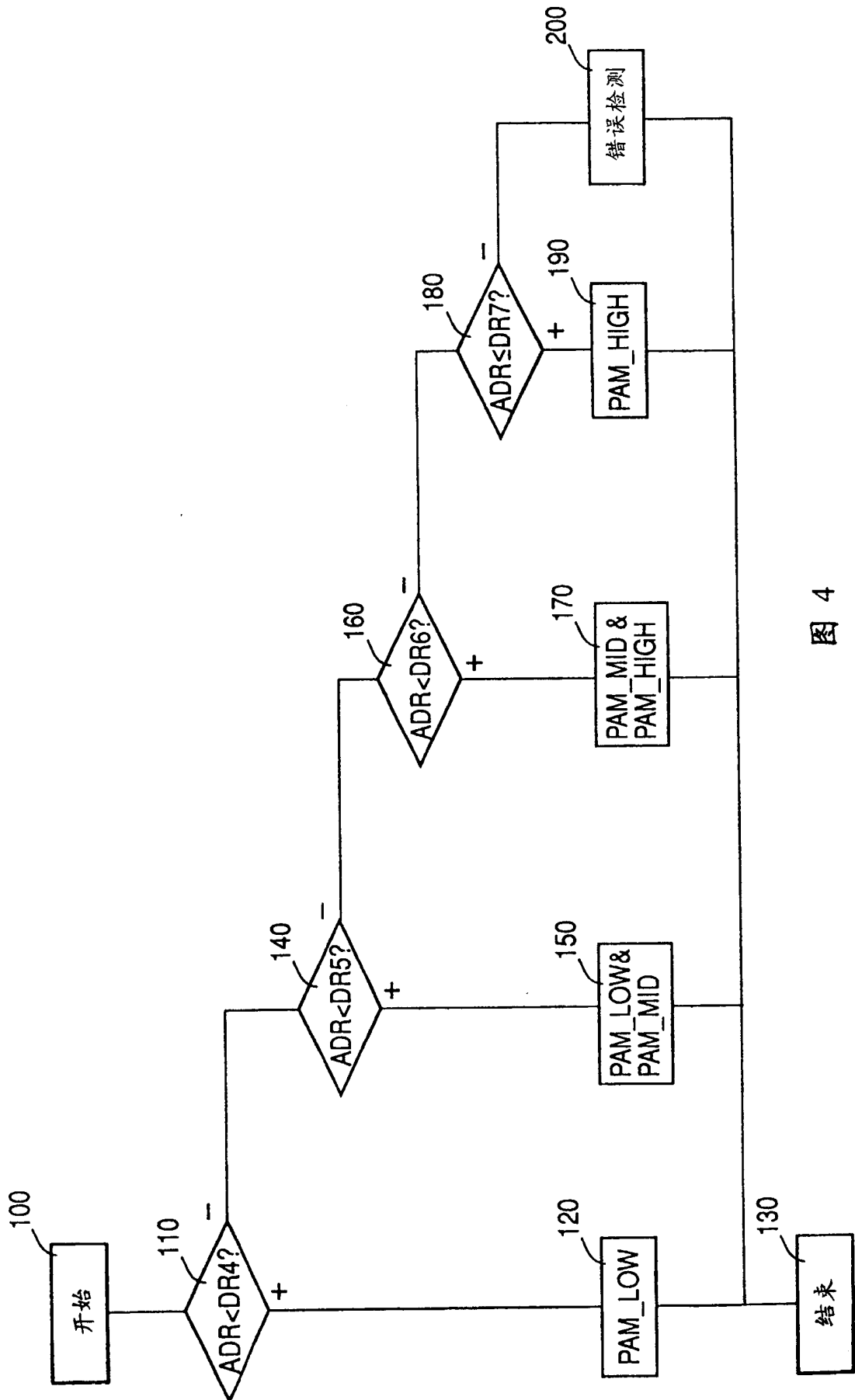


图 4

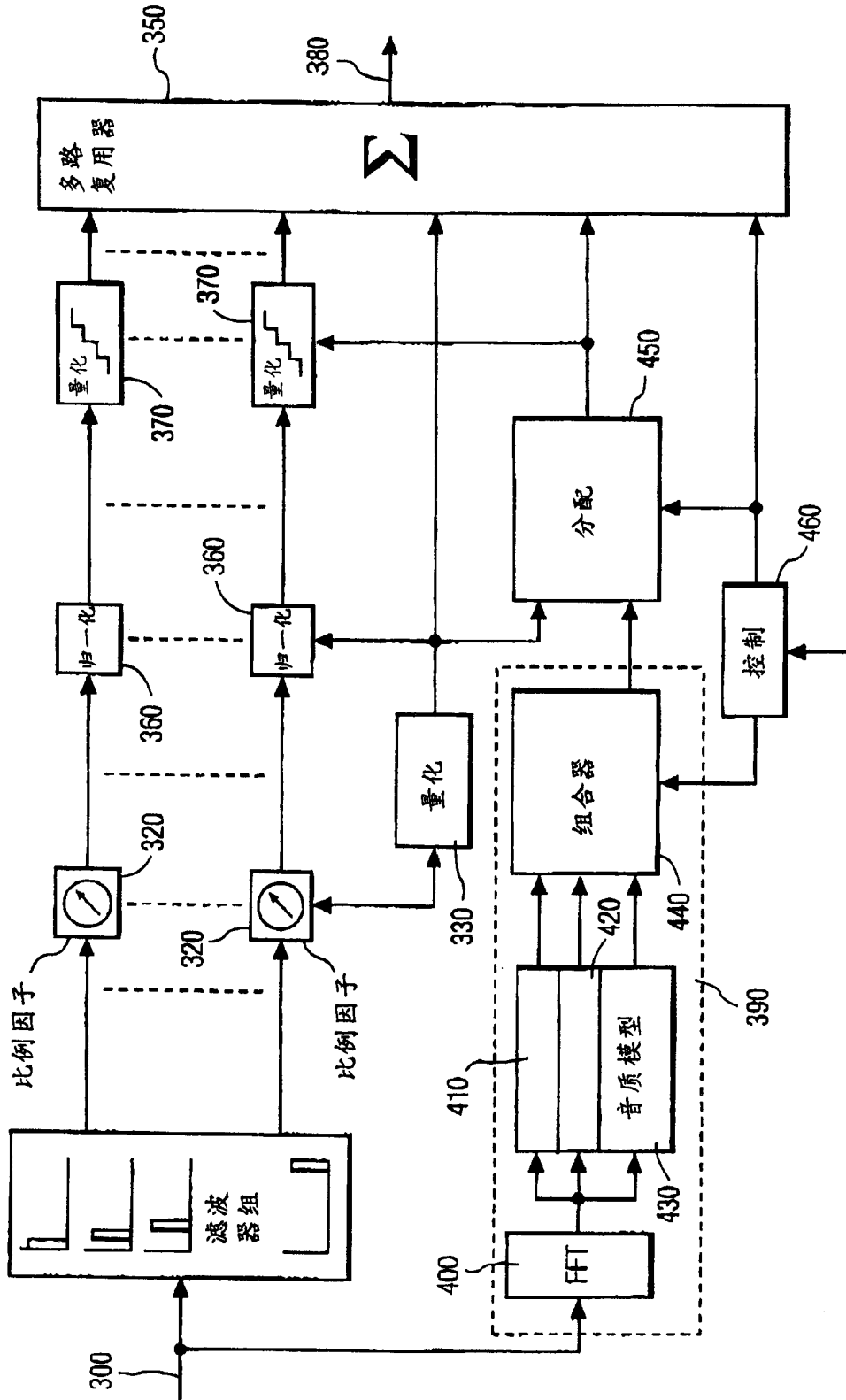


图 5