(19)中华人民共和国国家知识产权局

(12)发明专利

(10)授权公告号 CN 1735928 B
(45)授权公告日 2010.05.12

(21)申请号 200380108439.6

(22)申请日 2003.12.22

(30)优先权数据
03/00164 2003.01.08 FR

(85)PCT申请进入国家阶段日
2005.07.07

(86)PCT申请的申请数据
PCT/FR2003/003870 2003.12.22

(87)PCT申请的公布数据
WO2004/070706 FR 2004.08.19

(73)专利权人 法国电信公司
地址 法国巴黎

(72)发明人 巴拉兹·科弗西 多米尼克·马萨卢

(74)专利代理机构 北京市金杜律师事务所 11256
代理人 王茂华

(51)Int.Cl.
G10L 19/14(2006.01)
G10L 19/00(2006.01)
H03M 7/40(2006.01)
H04B 1/66(2006.01)

(56)对比文件
US 4949383 B1,1990.08.14, 全文.
US 6370507 B1,2002.04.09, 全文.
C.Erdmann.A CANDIDATE PROPOSAL FOR A 3GPP ADAPTIVEMULTI-RATE WIDEBAND SPEECH CODEC.IEEE2.2001,2757-760.
YE SHEN HONGMEI AI,C.-C.JAY KUO. A PROGESSIVE ALGORITHM FOR PERCEPTUAL CODINGOF DIGITAL AUDIO SIGNALS. IEEE.1999,1105-1109.

审查员 时鹏

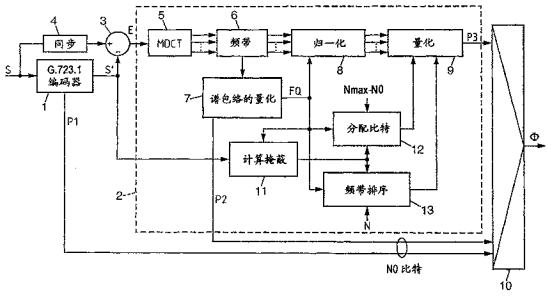
权利要求书 4 页 说明书 8 页 附图 3 页

(54)发明名称

用于可变速率音频编解码的方法

(57)摘要

为参数集合定义编码比特的最大数量Nmax，该参数可以从信号帧中计算。计算第一子集的参数，并用N0个比特编码，其中N0＜Nmax。确定分配Nmax-N0个编码比特用于第二子集的参数，并且对这些分配给第二子集的参数的编码比特分级。根据用于第一子集的编码参数，确定对编码比特的分配和/或分级排序。对于全部参数的编码可用的比特的总数N(N0＜N＝Nmax)，选择所述排序中的前N-N0个编码比特所分配的第二子集的参数。计算所述的所选参数并将其编码为N-N0个比特。最后，将用于第一子集的N0个编码比特和用于第二子集的所选参数的N-N0个编码比特引入编码器的输出序列。



CN 1735928 B

1. 一种把数字音频信号帧编码为二进制输出序列的方法，其中为参数集合定义编码比特的最大数量 Nmax，该参数可以根据所述信号帧计算，该集合包括第一子集和第二子集，所述方法包括如下步骤：

- 计算所述第一子集的所述参数，并且把这些参数编码为 N0 个编码比特，使得 N0 ＜ Nmax；

- 确定分配 Nmax-N0 个编码比特用于所述第二子集的所述参数；以及

- 把分配给所述第二子集的所述参数的所述 Nmax-N0 个编码比特按照确定的顺序排列，

其中，根据所述第一子集的所述编码参数，确定对所述 Nmax-N0 个编码比特的所述分配和所述排列顺序中的至少一个，响应所述二进制输出序列的 N 个比特的指示，该 N 个比特可用于所述参数集合的所述编码，且 N0 ＜ N ≤ Nmax，所述方法还包括以下步骤：

- 选择所述第二子集的参数，分配按照所述顺序排列的前 N-N0 个编码比特给这些参数；

- 计算所述第二子集的所选参数，并且对这些参数编码以产生所述排列的前 N-N0 个编码比特；以及

- 把所述第一子集的 N0 个编码比特以及所述第二子集的所选参数的 N-N0 个编码比特插入到所述输出序列中。

2. 根据权利要求 1 的方法，其中，分配给所述第二子集的所述参数的所述编码比特的所述排列顺序因帧而变化。

3. 根据权利要求 1 或 2 的方法，其中，N ＜ Nmax。

4. 根据权利要求 1 的方法，其中，分配给所述第二子集的所述参数的所述编码比特的所述排列顺序是根据至少所述第一子集的所述编码参数确定的重要性的降序排列。

5. 根据权利要求 4 的方法，其中，根据所述第一子集的所述编码参数，借助于至少一种心理声学准则，确定分配给所述第二子集的所述参数的所述编码比特的所述排列顺序。

6. 根据权利要求 5 的方法，其中，所述第二子集的所述参数与所述信号的谱带有关，其中基于所述第一子集的所述编码参数来估计所述编码信号的谱包络，其中通过将听觉感知模型应用于所述估计的谱包络来计算频率掩蔽曲线，并且其中所述心理声学准则参考所述估计的谱包络的级别，其与每一个谱带中的所述掩蔽曲线有关。

7. 根据权利要求 4 至 6 中任一权利要求的方法，其中，Nmax ＝ N。

8. 根据权利要求 1 的方法，其中，所述编码比特以这样的方式在所述输出序列中排序，即所述第一子集的所述 N0 个编码比特位于所述第二子集的所选参数的所述 N-N0 个编码比特之前，并且所述第二子集的所选参数的所述相应编码比特以对于所述编码比特所确定的顺序出现在其中。

9. 根据权利要求 1 的方法，其中，所述数量 N 因帧而变化。

10. 根据权利要求 1 的方法，其中，所述第一子集的所述参数的所述编码的比特速率是可变的，由此所述数量 N0 因帧而变化。

11. 根据权利要求 1 的方法，其中，所述第一子集包括由编码器内核 (1) 计算的参数。

12. 根据权利要求 11 的方法，其中，所述编码器内核 (1) 的工作频带低于要编码的所述信号的带宽，并且其中，所述第一子集还包括所述音频信号的能级，该能级与高于所述编码

器内核的所述工作频带的频带相关联。

13. 根据权利要求 8 和 12 中任一权利要求的方法，其中，所述第一子集的所述编码比特以这样的方式在所述输出序列中排序，即由所述编码器内核所计算的所述参数的所述编码比特，其后紧跟与所述较高频带相关联的所述能级的所述编码比特。

14. 根据权利要求 11 的方法，其中，估计要编码的所述信号和合成信号之间的差分信号，该合成信号源自由所述编码器内核所产生的所述编码参数，并且其中，所述第一子集还包括所述差分信号的能级，该能级与包括在所述编码器内核的所述工作频带中的频带相关联。

15. 根据权利要求 8 和权利要求 12 中任一权利要求的方法，其中，所述第一子集的所述编码比特以这样的方式在所述输出序列中排序，即由所述编码器内核 (1) 计算的所述参数的所述编码比特，其后紧跟与所述频带相关联的所述能级的所述编码比特。

16. 一种解码二进制输入序列以合成数字音频信号的方法，其中为参数集合定义编码比特的最大数量 Nmax，用于描述信号帧，该集合包括第一子集和第二子集，对于一个信号帧，所述输入序列包括 N′ 个编码比特用于所述参数集合，并且 N′ ≤ Nmax，所述方法包括以下步骤：

- 从所述输入序列的所述 N′ 个比特中，提取所述第一子集的所述参数的 N 0 个编码比特，假设 N0 ＜ N′ ；

- 基于提取的所述 N 0 个编码比特，恢复所述第一子集的所述参数；

- 确定分配 Nmax-N0 个编码比特用于所述第二子集的所述参数；以及

- 将分配给所述第二子集的所述参数的所述 Nmax-N0 个编码比特按照确定的顺序排列，

其中，根据所述第一子集的所述恢复参数，确定对所述 Nmax-N0 个编码比特的所述分配和所述排列顺序中的至少一个，所述方法还包括以下步骤：

- 选择所述第二子集的参数，分配按照所述顺序排列的前 N′ -N0 个编码比特给这些参数；

- 从所述输入序列的所述 N′ 个比特中，提取所述第二子集的所选参数的 N′ -N0 个编码比特；

- 基于所述提取的 N′ -N0 个编码比特，恢复所述第二子集的所选参数；以及

- 通过使用所述第一子集和第二子集的所述恢复参数，合成所述信号帧。

17. 根据权利要求 16 的方法，其中，分配给所述第二子集的参数的所述编码比特的排列顺序因帧而变化。

18. 根据权利要求 16 或 17 的方法，其中，N′ ＜ Nmax。

19. 根据权利要求 16 的方法，其中，分配给所述第二子集的所述参数的所述编码比特的所述排列顺序是根据至少所述第一子集的所述恢复参数确定的重要性的降序排列。

20. 根据权利要求 19 的方法，其中，根据所述第一子集的所述恢复参数，借助于至少一种心理声学准则，确定分配给所述第二子集的所述参数的所述编码比特的所述排列顺序。

21. 根据权利要求 20 的方法，其中，所述第二子集的所述参数与所述信号的谱带有关，其中基于所述第一子集的所述恢复参数来估计所述信号的谱包络，其中通过将听觉感知模型应用于所述估计的谱包络来计算频率掩蔽曲线，并且其中所述心理声学准则参考所述估

计的谱包络级别，其与每一个谱带中的所述掩蔽曲线有关。

22. 根据权利要求 16 的方法，其中，在从中提取所述第二子集的所选参数的所述 N′ -N0 个编码比特的位置之前的所述序列的位置处接收的 N′ 个比特中，提取所述第一子集的所述参数的所述 N0 个编码比特。

23. 根据权利要求 16 的方法，其中，为合成所述信号帧，基于至少所选参数，通过内插估计所述第二子集的未选参数，该所选参数基于提取的所述 N′ -N0 个编码比特恢复。

24. 根据权利要求 16 的方法，其中，所述第一子集包括解码器内核 (21) 的输入参数。

25. 根据权利要求 24 的方法，其中，所述解码器内核 (21) 的工作频带低于要合成的所述信号的带宽，并且其中，所述第一子集还包括所述音频信号的能级，该能级与高于所述解码器内核的所述工作频带的频带相关联。

26. 根据权利要求 22 和 25 中任一权利要求的方法，其中，所述第一子集的所述编码比特以这样的方式在所述输入序列中排序，即所述解码器内核 (21) 的所述输入参数的所述编码比特，其后紧跟与所述较高频带相关联的所述能级的所述编码比特。

27. 根据权利要求 26 的方法，假设所述输入序列的 N′ 个比特限制为所述解码器内核 (21) 的所述输入参数的所述编码比特，并且至少为与所述较高频带相关联的所述能级的部分编码比特，所述方法包括以下步骤：

- 从所述输入序列中提取所述解码器内核的所述输入参数的所述编码比特以及所述能级的所述部分编码比特；

- 在所述解码器内核中合成基带信号，并且基于所述提取的编码比特，恢复与所述较高频带相关联的能级；

- 计算所述基带信号的频谱；

- 给每一个较高频带分配能级，该较高频带与所述输入序列中的未编码的能级相关联；

- 基于所述相应的能级和在所述频谱的至少一个频带中的所述基带信号的所述频谱，为每一个较高频带合成频谱分量；

- 把所述合成的频谱分量变换到时域，以便得到基带信号的校正信号；以及

- 把所述基带信号和所述校正信号加到一起，以便合成所述信号帧。

28. 根据权利要求 27 的方法，其中，分配给较高频带的所述能级是一小部分感知掩蔽级别，根据所述基带信号的所述频谱和基于所述提取的编码比特恢复的所述能级计算该感知掩蔽级别，所述输入序列中的未编码能级与该较高频带相关联。

29. 根据权利要求 24 的方法，其中，在所述解码器内核中合成基带信号，并且其中，所述第一子集还包括要合成的所述信号和所述基带信号之间的差分信号的能级，该能级与包括在所述编码器内核的所述工作频带中的频带相关联。

30. 根据权利要求 25 的方法，其中，对于 N0 ＜ N′ ＜ Nmax，借助于计算得到的所述基带信号的频谱和 / 或基于提取的所述 N′ ＜ N0 编码比特所恢复的所选参数，估计所述第二子集的未选参数，该未选参数与频带中的频谱分量有关。

31. 根据权利要求 30 的方法，其中，借助于所述频带的频谱邻域，估计频带中的所述第二子集的所述未选参数，该邻域是基于所述输入序列的所述 N′ 个编码比特确定的。

32. 根据权利要求 22 和权利要求 25 中任一权利要求的方法，其中，在从中提取与所述

频带相关联的所述能级的所述编码比特的位置之前的所述序列的位置处接收的 N′ 个比特中，提取所述解码器内核 (21) 的所述输入参数的所述编码比特。

33. 根据权利要求 16 的方法，其中，所述数量 N′ 因帧而变化。

34. 根据权利要求 16 的方法，其中，所述数量 N0 因帧而变化。

35. 一种音频编码器，包括数字信号处理的装置，该装置设计成实施根据权利要求 1 的编码方法。

36. 一种音频解码器，包括数字信号处理的装置，该装置设计成实施根据权利要求 16 的解码方法。

用于可变速率音频编解码的方法

技术领域

[0001] 本发明涉及用于音频信号编解码的设备，更具体地说，旨在用于发送或存储数字压缩的音频信号（语音和/或声音）。

[0002] 更具体地说，本发明涉及能够提供各种比特速率的音频编码系统，也称为多速率编码系统。这样的系统不同于固定速率的编码器，因为它们可能在处理过程中能够调整编码的比特速率，这特别适合通过异构接入网传送：异构接入网可以是固定接入和移动接入混合的IP类型的网络、高比特速率(ADLS)的网络、低比特速率(RTC,GPRS调制解调器)的网络或者涉及可变容量的终端（手机、PC等）的网络。

背景技术

[0003] 从本质上，划分为两类多速率编码器：“可交换”多速率编码器和“分层”编码器。

[0004] “可交换”多速率编码器依赖于属于技术族（时域编码或频域编码，例如：CELP、正弦编码或变换编码）的编码体系结构，其中同时向编码器和解码器提供比特速率的指示。编码器利用该信息来选择相应的算法和与所选比特速率相关的表。解码器以对称的方式操作。已经提出了许多种可交换多速率编码结构用于音频编码。例如由3GPP组织（“第三代合作伙伴计划”）标准化的移动编码器，电话频带中的NB-AMR(“窄带自适应多速率”，TechnicalSpecification 3GPP TS 26.090，version 5.0.0，June 2002)，或宽带中的WB-AMR(“宽带自适应多速率”，TechnicalSpecification 3GPP TS 26.190，version 5.1.0，December 2001)，就是这样的例子。这些编码器在相当宽的比特速率范围内操作(对于NB-AMR从4.75到12.2kbit/s，对于WB-AMR从6.60到23.85kbit/s)，并且粒度相当大(NB-AMR为8比特速率以及WB-AMR为9比特速率)。然而，为这种灵活性付出的代价是相当高的结构复杂性：能够支持所有这些比特速率，这些编码器必须支持多种不同的选择，各种量化表等。性能曲线随着不断增加的比特速率而增长，但是该过程不是线性的并且某些比特速率本质上优于其它的比特速率。

[0005] 在所谓的“分层”编码系统中，也称作“可分级”，把来自编码操作的二进制数据分成连续的层次。基础层，也称作“内核”，由二进制单元组成，这些二进制单元是二进制链解码所绝对必需的，并且决定解码的最低质量要求。

[0006] 接下来的层使之可以不断提高来自解码操作的信号的质量，每新的一层带来解码器可利用的新信息，从而在输出端提供质量不断提高的信号。

[0007] 分层编码的特殊特征之一是可以提供在发送或存储链的任何级别上的干涉，以便删除一部分二进制链而不必向编码器或解码器提供任何特殊的指示。解码器利用它接收的二进制信息，并产生相应质量的信号。

[0008] 分层编码结构领域同样引发了许多工作。某种分层编码结构仅仅基于一种编码类型操作，设计成传送分层的编码信息。当附加层提高输出信号的质量而不需要调整带宽时，有人喜欢称为“嵌入式编码器”(例如参见R.D.Lacovo et al.，“Embedded CELP Coding forVariable Bit-Rate Between 6.4 and 9.6kbit/s”，Proc.ICASSP1991，pp.681-686)。

然而这种类型的编码器不允许所提出的最低比特速率和最高比特速率之间有很大的差距。

[0009] 分层常用于逐渐增加信号的带宽：内核提供基带信号，例如电话信号(300-3400Hz)，而接下来的层使额外频带（例如，宽带可达7kHz，HiFi频带可达20kHz或中频，等等）进行编码。子带编码器利用时间/频率变换，如同下面的文献所述，J.P.Princen等人提出的“Subband/transform coding using filter banks designsbased on time domain aliasing cancellation”(Proc.IEEEICASSP-87，pp.2161-2164)以及Y.Mahieux等人提出的“HighQuality Audio Transform Coding at 64kbit/s”(IEEE Trans.Commun.，Vol.42，No.11，November 1994，pp.3010-3019)，使之具体地执行这样的操作。

[0010] 此外，对于内核和对附加层进行编码的模块，经常使用不同的编码技术，所以有人称为各种编码级，每一级包括一个子编码器。给定级别的那一级的子编码器或者能够对前一级未编码的部分信号编码，或者能够对前一级的编码残差编码，该残差是由原始信号减去解码后的信号得到。

[0011] 这种结构的优点是在保证满足质量的情况下使得可以降到相对低的比特速率，而在高比特速率时产生高的质量。特别地，用于低比特速率的技术通常在高比特速率时无效，反之亦然。

[0012] 这种结构使得可以使用两种不同的技术（例如CELP和时间/频率变换等），这对于覆盖大范围的比特速率特别有效。

[0013] 然而，现有技术提出的分层编码结构精确定义了分配给每一个中间层的比特速率。每一层对应一定参数的编码，并且分层二进制链的粒度取决于分配给这些参数的比特速率（典型地，一层每帧可以包含几十量级的比特，信号帧包括信号在给定持续时间段上的一定数量的采样，后面所描述的实施例考虑每帧960个采样对应信号的60ms）。

[0014] 此外，当解码信号的带宽可以根据二进制单元的分层级别变化时，对线路比特速率的调整会产生影响收听的失真。

发明内容

[0015] 特别地，本发明旨在提出一种多速率编码解决方案，以减少在利用现存的分层编码和可交换编码情况下引入的缺陷。

[0016] 因此本发明提出一种把数字音频信号帧编码为二进制输出序列的方法，其中为参数集合定义编码比特的最大数量Nmax，该参数可以根据信号帧来计算，该集合包含第一子集和第二子集。该提出的方法包括以下步骤：

[0017] -计算第一子集的参数，并且把这些参数编码为N0个编码比特，使得N0 < Nmax；

[0018] -确定分配Nmax-N0个编码比特用于第二子集的参数；以及

[0019] -把分配给第二子集的参数的Nmax-N0个编码比特按照确定的顺序排列。

[0020] 根据第一子集的编码参数，确定对Nmax-N0个编码比特的分配和/或排列顺序。响应二进制输出序列的N个比特的指示，该N个比特可用于所述参数集合的编码，且N0 < N ≤ Nmax，该编码方法还包括以下步骤：

[0021] -选择第二子集的参数，分配按照所述顺序排列的前N-N0个编码比特给这些参数；

[0022] -计算第二子集的所选参数，并且对这些参数编码以产生所述排列的前N-N0个编

码比特；以及

[0023] - 把第一子集的 N0 个编码比特以及第二子集的所选参数的 N-N0 个编码比特插入到输出序列中。

[0024] 根据本发明的方法使得能够可以定义一种多速率编码，其至少会在对应每帧的比特数从 N0 变化到 Nmax 的范围下操作。

[0025] 因此可以考虑用“指针”(cursor) 概念代替与现有分层编码和可交换编码相关的预定速率的概念，使得可以在最小值（可能对应小于 N0 的比特数 N) 和最大值（对应 Nmax) 之间自由地改变比特速率。这些极值可能离的很远。不管所选的比特速率怎样，该方法在编码的有效性方面都提供良好的性能。

[0026] 有利地，二进制输出序列的比特数 N 严格小于 Nmax。那么关于编码器值得注意的是，采用的比特分配不是参考编码器的实际输出比特速率，而是参考适合解码器的另一个 Nmax。

[0027] 然而，根据传输信道上可用的瞬时比特速率，可以固定 Nmax ＝ N。像这样的可交换多速率编码器的输出序列可以由解码器处理，解码器不用接收整个序列，只要借助 Nmax 的已知信息就能够恢复第二子集的编码比特的结构。

[0028] 另一种情况，其中可以使 N ＝ Nmax，以最大的编码速率存储音频数据。当以较低比特速率读取存储的该内容的 N′ 个比特时，只要 N′ ＞ N0，解码器就能够恢复第二子集的编码比特的结构。

[0029] 分配给第二子集的参数的编码比特的排列顺序可以是预定的顺序。

[0030] 在优选的实施方式中，分配给第二子集的参数的编码比特的排列顺序是可变的。特别地，它可以是根据至少第一子集的编码参数确定的重要性的降序排列。因此，解码器接收该帧的 N′ 比特的二进制序列，并且 N0 ＜ N′ ＜ N ＜ Nmax，该解码器能够从接收的用于第一子集的编码的 N0 个比特推出该顺序。

[0031] 可以以固定的方式执行分配 Nmax-N0 个比特用于第二子集的参数的编码（在这种情况下，这些比特的排列顺序将依赖于至少第一子集的编码参数）。

[0032] 在优选的实施方式中，根据第一子集的编码参数，分配 Nmax-N0 个比特用于第二子集的参数的编码。

[0033] 有利地，根据第一子集的编码参数，借助于至少一种心理声学准则，确定分配给第二子集的参数的编码比特的这种排列顺序。

[0034] 第二子集的参数与信号的谱带有关。在这种情况下，有利地，该方法包括步骤：基于第一子集的编码参数估计编码信号的谱包络，以及步骤：通过将听觉感知模型应用于所估计的谱包络来计算频率掩蔽曲线，并且该心理声学准则参考所估计的谱包络的级别，其与每一个谱带中的掩蔽曲线有关。

[0035] 在一种实施方式中，编码比特以这样的方式在输出序列中排序，即第一子集的 N0 个编码比特位于第二子集的所选参数的 N-N0 个编码比特之前，并且第二子集的所选参数的相应编码比特以对于所述编码比特所确定的顺序出现在其中。这使得可以在二进制序列被截短的情况下接收最重要的部分。

[0036] 数量 N 可以因帧而变化，特别地例如根据传输资源的可用容量而变化。

[0037] 可以根据非常灵活的分层或可交换模式使用根据本发明的多速率音频编码，因为

要发送的比特数量在任意时刻，也就是说是逐帧的，可以在 N0 和 Nmax 之间进行自由的选择。

[0038] 第一子集的参数的编码的比特速率可以是可变的，由此数量 N0 因帧而变化。这允许根据要编码的帧把比特分配调整到最佳。

[0039] 在一种实施方式中，第一子集包括由编码器内核计算的参数。有利地，编码器内核的工作频带低于要编码的信号的带宽，并且第一子集还包括音频信号的能级，该能级与高于编码器内核的工作频带的频带相关联。这种类型的结构是具有两级的分层编码器，例如它经由编码器内核传送认为满足一定质量的信号，并且根据可变的比特速率，补充由具有附加信息的编码器内核执行的编码，该附加信息来自根据本发明的编码方法。

[0040] 优选地，第一子集的编码比特然后以这样的方式在输出序列中排序，即由编码器内核所计算的参数的编码比特，其后紧跟与较高频带相关联的能级的编码比特。只要解码器接收足够的比特，这些比特具有编码器内核的信息以及与较高频带相关联的编码能级的信息，这样对于连续编码的帧可以保证同一带宽。

[0041] 在一种实施方式中，估计要编码的信号和合成信号之间的差分信号，该合成信号源自由编码器内核所产生的编码参数，并且第一子集还包括差分信号的能级，该能级与包括在编码器内核的工作频带中的频带相关联。

[0042] 本发明的第二个方面是关于解码二进制输入序列的方法，以便合成数字音频信号，对应于对根据本发明的编码方法编码的帧的解码。根据该方法，为参数集合定义编码比特的最大数量 Nmax，用于描述信号帧，该集合包括第一子集和第二子集。对于一个信号帧，输入序列包括 N′ 个编码比特用于参数集合，并且 $N' \leqslant Nmax$。根据本发明的解码方法包括以下步骤：

[0043] - 从输入序列的所述 N′ 个比特中，提取第一子集的参数的 N0 个编码比特，假设 $N0 < N'$；

[0044] - 基于提取的所述 N0 个编码比特，恢复第一子集的参数；

[0045] - 确定分配 Nmax-N0 个编码比特用于第二子集的参数；以及

[0046] - 将分配给第二子集的参数的 Nmax-N0 个编码比特按照确定的顺序排列。

[0047] 根据第一子集的所恢复的参数，确定对 Nmax-N0 个编码比特的分配和 / 或排列顺序，该解码方法还包括以下步骤：

[0048] - 选择第二子集的参数，分配按照所述顺序排列的前 N′ -N0 个编码比特给这些参数；

[0049] - 从输入序列的所述 N′ 个比特中，提取第二子集的所选参数的 N′ -N0 个编码比特；

[0050] - 基于提取的所述 N′ -N0 个编码比特，恢复第二子集的所选参数；以及

[0051] - 通过使用第一子集和第二子集的所恢复的参数，合成信号帧。

[0052] 有利地，该解码方法与用于再生丢失的参数的过程相关联，参数丢失的原因是由编码器所实际或另外产生的 Nmax 个比特序列的截短造成的。

[0053] 本发明的第三个方面是关于音频编码器，该音频编码器包括数字信号处理的装置，该装置设计成实施根据本发明的编码方法。

[0054] 本发明的另一个方面是关于音频解码器，该音频解码器包括数字信号处理的装

置，该装置设计成实施根据本发明的解码方法。

附图说明

[0055] 本发明的其他特征和优点将在下面参照附图对非限制性的示例性实施方式进行的描述中变得显而易见，其中：

[0056] 图 1 是根据本发明的一种示例性音频编码器的示意图；

[0057] 图 2 表示本发明实施方式中的 N 位的二进制输出序列；以及

[0058] 图 3 是根据本发明的一种音频解码器的示意图。

具体实施方式

[0059] 图 1 中示出的编码器是具有两个编码级的一种分层结构。第一编码级 1 包括例如 CELP 类型的电话频带 (300-3400Hz) 中的编码器内核。该实施例中的编码器考虑一种由 ITU-T(“国际电信联盟”) 标准化的固定模式 6.4kbit/s 的 G.723.1 编码器。它根据该标准计算 G.723.1 参数并且按照 30ms 每帧 192 个编码比特 P1，对参数进行量化。

[0060] 第二编码级 2，使得可以将带宽向宽带 (50-7000Hz) 增加，该第二编码级 2 对第一级的编码残差 E 操作，该编码残差 E 由图 1 框图中的减法器 3 提供。信号同步模块 4 使音频信号帧 S 延迟一段时间，该段时间是编码器内核 1 的处理所用的时间。信号同步模块 4 的输出被送到减法器 3，减法器 3 从该输出中减去合成信号 S′，S′ 等于解码器内核的输出，该解码器内核在诸如由编码器内核的输出比特 P1 所表示的量化参数的基础上操作。通常，编码器 1 与提供 S′ 的本地解码器结合。

[0061] 要编码的音频信号 S 例如具有带宽 7KHz，同时采样频率为 16KHz。一帧例如包括 960 个采样，即信号的 60ms 或编码器内核 G.723.1 的两个基本帧。由于后者是对在 8KHz 下采样的信号进行操作，所以在编码器内核 1 的输入端以因子 2 对信号 S 进行二次采样。同样地，在编码器内核 1 的输出端以 16KHz 对合成信号 S′ 进行过采样。

[0062] 第一级 1 的比特速率是 6.4kbit/s(2×N1 ＝ 2×192 ＝ 384 比特 / 帧)。如果编码器具有最大比特速率 32kbit/s(Nmax ＝ 1920 比特 / 帧)，那么第二级的最大比特速率为 25.6kbit/s(1920-384 ＝ 1536 比特 / 帧)。第二级 2 例如对 20ms(在 16KHz 下的 320 个采样) 的基本帧或子帧进行操作。

[0063] 第二级 2 包括一个时间 / 频率变换模块 5，例如为 MDCT(“修正离散余弦变换”) 类型，减法器 3 所得到的残差 E 送至该模块 5。实际上，图 1 中表示的模块 3 和 5 的操作方式可以通过对每一个 20ms 的帧执行以下操作来实现，操作如下：

[0064] - 对经过模块 4 延时的输入信号 S 进行 MDCT 变换，提供 320 个 MDCT 系数。由于频谱限制在 7225Hz，所以只有前 289 个 MDCT 系数不是 0；

[0065] - 对合成信号 S′ 进行 MDCT 变换。由于处理的是电话频带信号的频谱，所以只有前 139 个 MDCT 系数不是 0(可达 3450Hz)；以及

[0066] - 计算前面频谱间的差频谱。

[0067] 所得到的频谱通过模块 6 分布到宽度不同的几个频带上。通过实施例，G.723.1 编解码器的带宽可以细分为 21 个频带，而较高的频率分布到 11 个附加频带上。在这 11 个附加频带中，残差 E 等同于输入信号 S。

[0068] 模块 7 执行残差 E 的谱包络的编码。它从计算差频谱的每一频带的 MDCT 系数的能量开始。这些能量在下文称为“比例因子”。32 个比例因子组成差分信号的谱包络。然后模块 7 分成两部分进行谱包络的量化。第一部分对应电话频带（前 21 个频带，从 0 到 3450Hz），第二部分对应高频带（后 11 个频带，从 3450 到 7225Hz）。在每一部分中，第一比例因子基于绝对准则进行量化，以及接下来的那些比例因子基于差分准则进行量化，均通过采用常规的可变比特速率的霍夫曼编码来实现。对于等级为 i（i ＝ 1,2,3）的每个子帧，基于可变数量 N2(i) 的比特 P2 对这 32 个比例因子进行量化。

[0069] 量化比例因子在图 1 中用 FQ 表示。第一子集的量化比特 P1、P2 包括编码器内核 1 的量化参数和量化比例因子 FQ，该量化比特 P1、P2 的数量 NO ＝ (2×N1)+N2(1)+N2(2)+N2(3) 可变。差 Nmax-NO ＝ 1536-N2(1)-N2(2)-N2(3) 可用于更精细地量化频谱。

[0070] 模块 8 通过用这些频带所分别确定的量化比例因子 FQ 划分 MDCT 系数，对由模块 6 分布到不同频带的 MDCT 系数进行归一化。把因此归一化的频谱提供给量化模块 9，该模块 9 采用已知类型的矢量量化方案。由模块 9 产生的量化比特在图 1 中用 P3 表示。

[0071] 输出多路复用器 10 把来自模块 1、7、9 的比特 P1、P2 和 P3 收集在一起，以形成编码器的二机制输出序列 Φ。

[0072] 根据本发明，代表当前帧的输出序列的总比特数 N 不必等于 Nmax。它可以小于后者。然而，对频带执行量化比特的分配是基于数量 Nmax 来执行的。

[0073] 在图 1 的框图中，基于数量 Nmax-NO、量化比例因子 FQ 以及由模块 11 计算的频谱掩蔽曲线，由模块 12 为每个子帧执行这种分配。

[0074] 模块 11 的操作方式如下。它基于例如由模块 7 量化的差分信号的谱包络，以及确定同样的解决方案用于编码器内核产生的合成信号 S′，首先确定信号 S 的初始谱包络的一个大约值。这两个包络也可由解码器确定，该解码器只提供有前述第一子集的参数。从而信号 S 的估计的谱包络也可用于解码器。因此，模块 11 通过以自身已知的方式，把逐频带的一种听觉感知模型应用到初始估计的谱包络，来计算谱掩蔽曲线。该曲线 11 给出所考虑的每一个频带的掩蔽级别。

[0075] 在差分信号的三层 MDCT 变换的 3×32 个频带中，模块 12 对该序列 Φ 的 Nmax-NO 个剩余比特执行动态分配。在这里列举出的本发明的实施中，根据心理声学感知重要性的准则，参考关于每个频带中的掩蔽曲线而估计的谱包络的级别，对每个频带分配与该级别成比例的比特速率。其他的等级准则也是可用的。

[0076] 在分配比特之后，模块 9 知道有多少比特要考虑用于每个子帧中的每个频带的量化。

[0077] 然而，如果 N ＜ Nmax，就不必使用所有这些分配的比特。根据感知重要性的准则，由模块 13 执行表示频带的比特排序。模块 13 按照重要性降序排列 3×32 个频带，该降序可以是信号掩蔽比（每个频带中估计的谱包络和掩蔽曲线之间的比）的降序。根据本发明，使用该顺序用于建立二进制序列 Φ。

[0078] 根据序列 Φ 中用于当前帧编码所需的比特数 N，通过选择首先由模块 13 排列的频带并且通过对每个所选频带保持由模块 12 确定的比特数，来确定要由模块 9 量化的频带。

[0079] 然后例如借助矢量量化器，按照所分配的比特数量，由模块 9 量化所选的每一个

频带的 MDCT 系数，使得产生的总比特数等于 N-N0。

[0080] 输出多路复用器 10 建立二进制序列 Φ，该序列 Φ 包括图 2 中所示按照如下顺序排列的序列的前 N 个比特（N = Nmax 的情况）：

[0081] a) 首先是对应两个 G.723.1 帧的二进制链（384 比特）；

[0082] b) 接下来是比特 $F_{22}^{(i)},\ldots,F_{32}^{(i)}$ 用于对三个子帧（i = 1,2,3）量化比例因子，从第 22 个谱带（超出电话频带的第一频带）到第 32 个频带（可变速率的霍夫曼编码）；

[0083] c) 接下来是比特 $F_{1}^{(i)},\ldots,F_{21}^{(i)}$ 用于对三个子帧（i = 1,2,3）量化比例因子，从第一个频带到第 21 个频带（可变速率的霍夫曼编码）；

[0084] d) 以及，最后是 96 个频带的矢量量化的索引 $M_{c1},M_{c2},\ldots,M_{c96}$，按照感知重要性的顺序，从最重要的频带到重要性最低的频带，同时遵守模块 13 所确定的顺序。

[0085] 通过首先放置 (a) 和 b))G.723.1 参数和高频带的比例因子，不管实际上的比特速率是否超出对应接收这些组 a) 和 b) 的最小值，都可以使可由解码器恢复的信号保持同样的带宽。这个最小值，除了满足 G.723.1 编码，还满足高频带 3×11 = 33 个比例因子的霍夫曼编码，例如该最小值为 8kbit/s。

[0086] 如果解码器接收 N′ 个比特而 N0 ≤ N′ ≤ N，则上文所述的编码方法允许帧的解码。数量 N′ 通常会因帧而变化。

[0087] 对应本实施例，图 3 示出了根据本发明的解码器。解多路复用器 20 分离接收的比特序列 Φ′，以便从中提取编码比特 P1 和 P2。将 384 比特 P1 提供给 G.723.1 类型的解码器内核 21，从而解码器内核 21 合成电话频带中的基带信号 S′ 的两个帧。由模块 22 根据霍夫曼算法对比特 P2 解码，由此恢复 3 个子帧的每一个的量化比例因子 FQ。

[0088] 模块 23 计算掩蔽曲线，该模块 23 等同于图 1 编码器的模块 11，接收基带信号 S′ 和量化比例因子 FQ，并为 96 个频带的每一个产生谱掩蔽级别。基于这些掩蔽级别、量化比例因子 FQ 以及已知数量 Nmax（也基于已知数量 N0，该数量 N0 由模块 22 对比特 P2 的霍夫曼解码推得），模块 24 按照与图 1 的模块 12 同样的方式确定比特的分配。此外，按照与参照图 1 所述的模块 13 同样的排列准则，模块 25 对频带进行排序。

[0089] 根据模块 24 和 25 提供的信息，模块 26 提取输入序列 Φ′ 的比特 P3，并且合成归一化的 MDCT 系数，该系数与序列 Φ′ 中表示的频带相关联。如果合适 (N′ < Nmax)，还可以通过如下文描述的内插或外推，合成与丢失的频带关联的标准化 MDCT 系数（模块 27）。这些丢失的频带可能由于译码器截短 N < Nmax 已经被解码器去除，或者它们可能已经在传输过程中被消除 (N′ < N)。

[0090] 标准化 MDCT 系数，由模块 26 和 / 或模块 27 合成，在被送到模块 29 之前乘以他们各自的量化比例因子（乘法器 28），以执行频率 / 时间变换，这是由编码器的模块 5 操作的 MDCT 变换的逆变换。将由此产生的时域校正信号加到由解码器内核 21 传送来的合成信号 S′ 上，以产生解码器的输出音频信号 $\hat{S}$。

[0091] 应当注意的是，即使解码器没有接收到序列的前 N0 个比特，它也能够合成信号 $\hat{S}$。

[0092] 解码器足以接收对应上文中列出的那部分 a) 的 2×N1 个比特，则解码处于“退化”模式。只有这种退化模式不使用 MDCT 合成以得到解码信号。为保证该模式和其他模式之间的切换没有中断，解码器在三种 MDCT 合成之后执行三种 MDCT 分析，使得更新 MDCT 变换

的存储器。输出信号包含电话频带品质的信号。如果即使没有接收到前 2×N1 个比特，解码器也会认为相应的帧已经被擦除并且能够使用已知的算法来构造被擦除的帧。

[0093] 如果解码器接收对应于部分 a) 加上部分 b) 的比特的 2×N1 个比特（三种谱包络的高频带），则它能够首先合成一种宽带信号。特别地，解码器能够继续如下操作：

[0094] 1) 模块 22 恢复部分所接收的三种谱包络。

[0095] 2) 没有接收到的频带把它们的比例因子临时设为零。

[0096] 3) 基于对 G.723.1 解码之后获得的信号执行的 MDCT 分析，计算谱包络的低频部分，并且模块 23 对因而获得的包络计算三种掩蔽曲线。

[0097] 4) 校正谱包络以便调整它，避免由于未接收到该频带而为零；谱包络 FQ 高频部分的零值例如由前面计算的掩蔽曲线的百分之一 (hundredth) 的值替代，以使他们保持听不见。低频带的整个频谱和高频带的谱包络在该情况下是已知的。

[0098] 5) 然后模块 27 生成高频谱。在比例因子加权之前（乘法器 28），这些频带的精细结构由已知邻域的精细结构映射而成。在没有接收到任何一个比特 P3 的情况下，“已知邻域”对应 G.723.1 解码器内核所产生的信号 S′ 的频谱。它的“映射”可以包括复制标准化 MDCT 频谱的值，该值可以是变化的，与离开“已知邻域”的距离成比例地减小。

[0099] 6) 在逆 MDCT 变换 (29) 以及把得到的校正信号加到 (30) 解码器内核的输出信号上之后，得到宽频带的合成信号。

[0100] 在解码器也至少接收差分信号（部分 c.)) 的部分低频谱包络的情况下，在步骤 3 中它可以考虑该信息以改进 (refine) 谱包络，也可以不考虑该信息。

[0101] 如果解码器 10 接收足够的比特 P3，以至少解码最重要频带的 MDCT 系数，即序列的部分 d) 中排在前面的部分，然后模块 26 根据模块 24 和 25 所指示的分配和排序，恢复某些归一化的 MDCT 系数。因此这些 MDCT 系数不必如同上文步骤 5 那样进行内插。对于其他的频带，模块 27 可以按前面同样的方式应用步骤 1 到 6 的处理，对于某些频带接收 MDCT 系数的信息在步骤 5 中允许更可靠的内插。

[0102] 未接收到的频带从一个 MDCT 子帧到下一个子帧会有所不同。丢失频带的“已知邻域”可能对应其他子帧中没有丢失的相同的频带，并且 / 或者在相同子帧的过程中对应频域中最接近的一个或多个频带。也可以通过计算加权贡献的总和，为子帧重新生成从某一频带丢失的 MDCT 系数，贡献是基于“已知邻域”的几个频带 / 子帧评估的。

[0103] 在某种程度上由于每帧 N′ 比特的实际比特速率任意放置给定帧的最后比特，最后发送的编码参数，根据情况，可以完全被发送或者部分地被发送。于是会出现两种情况：

[0104] - 或者所采用的编码结构使得可以利用接收到的部分信息（标量量化器的情况，或者带有分区字典 (partitioned dictionary) 的矢量量化的情况），

[0105] - 或者它不允许该信息以及将未完全接收的参数像其它没有接收到的参数那样处理。对于后者的情况要注意，如果比特的排序随每帧变化，则因此丢失的比特数是可变的并且在整个所解码的帧的集合上将平均产生所选的 N′ 个比特，其质量好于比特数较小时获得的质量。

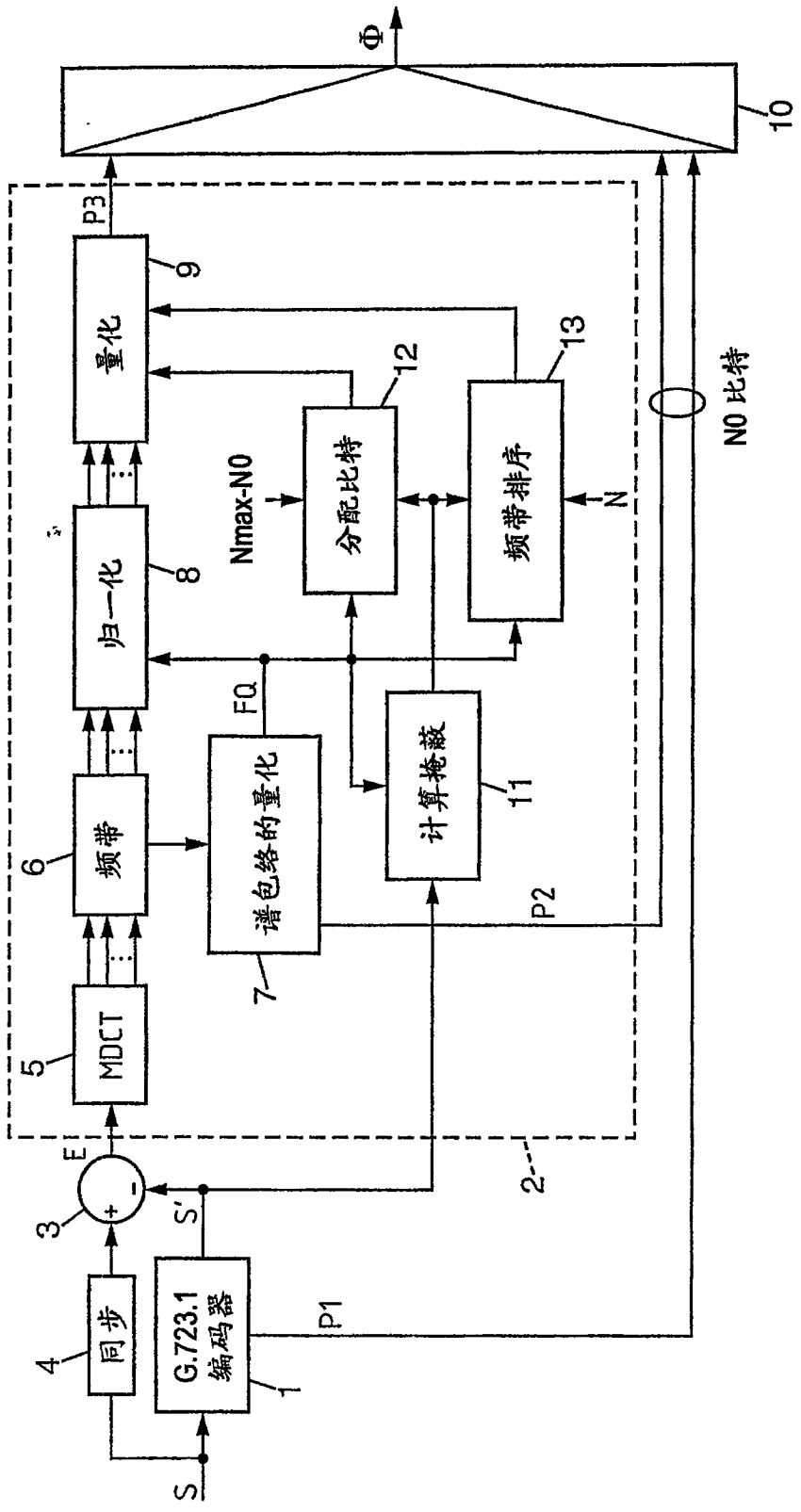

Φ
10
P3
量化
9
归一化
8
频带
6
MDCT
5
同步
4
G.723.1
编码器
1
S
S'
E
3
2
FQ
谱包络的量化
7
计算掩蔽
11
分配比特
12
频带排序
13
Nmax-N0
N
N0 比特
P1
P2

图 1

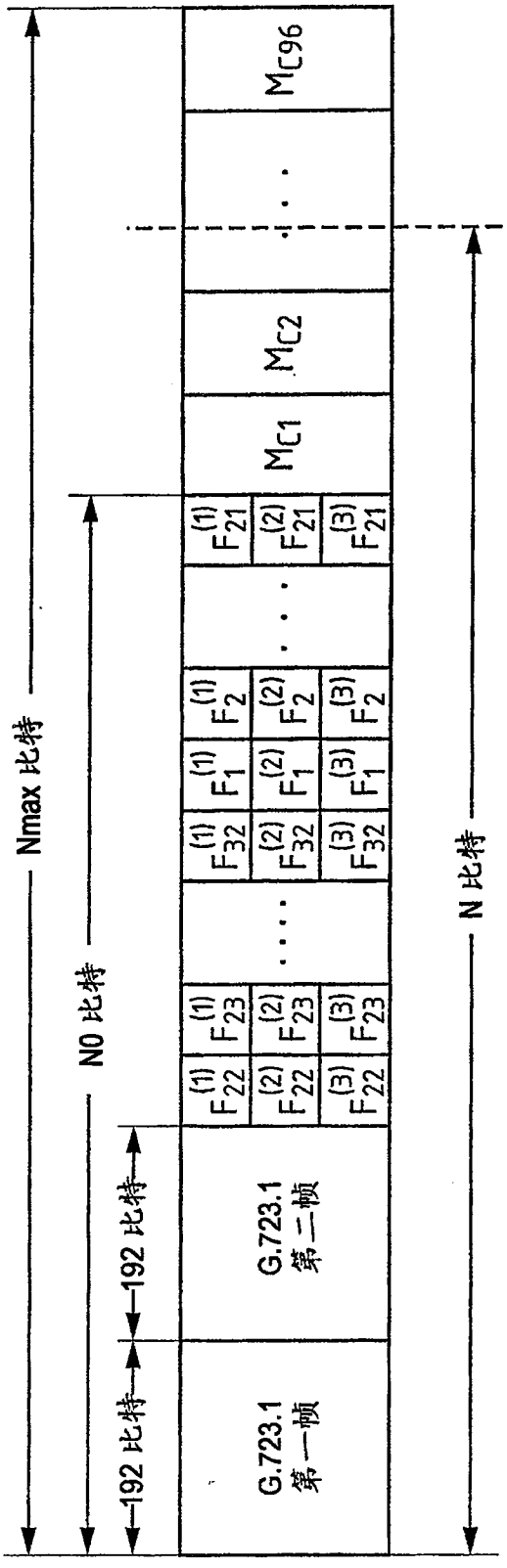

Nmax 比特
N0 比特
N 比特
192 比特
192 比特
G.723.1
第一帧
G.723.1
第二帧
M_{C1}
M_{C2}
M_{C96}
$F_{22}^{(1)}$ $F_{22}^{(2)}$ $F_{22}^{(3)}$
$F_{23}^{(1)}$ $F_{23}^{(2)}$ $F_{23}^{(3)}$
$F_{32}^{(1)}$ $F_{32}^{(2)}$ $F_{32}^{(3)}$
$F_1^{(1)}$ $F_1^{(2)}$ $F_1^{(3)}$
$F_2^{(1)}$ $F_2^{(2)}$ $F_2^{(3)}$
$F_{21}^{(1)}$ $F_{21}^{(2)}$ $F_{21}^{(3)}$

图 2

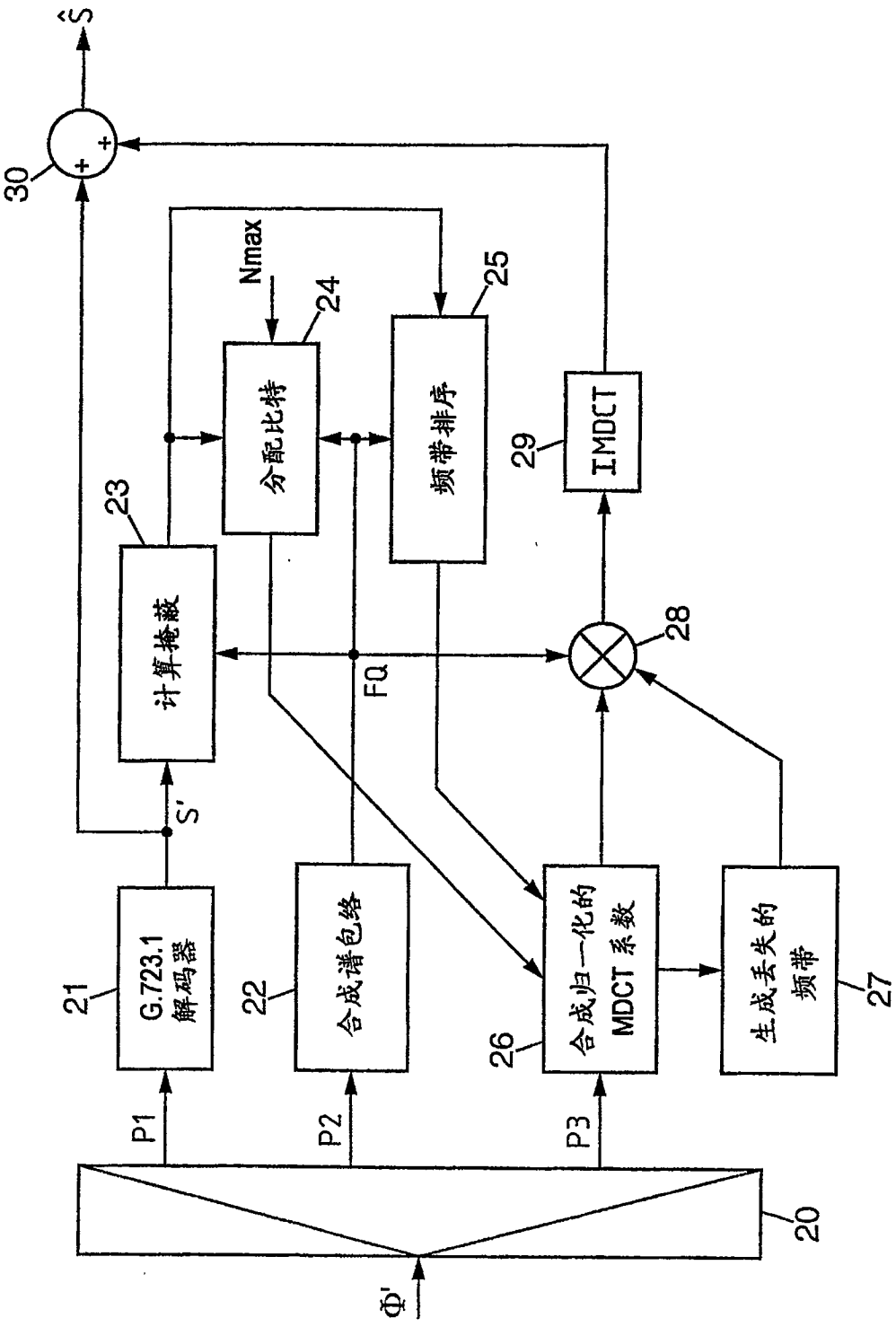

Ŝ
30
Nmax
24
25
分配比特
频带排序
29
IMDCT
23
计算掩蔽
FQ
28
S'
21
G.723.1
解码器
22
合成谱包络
26
合成归一化的
MDCT 系数
生成丢失的
频带
27
P1
P2
P3
20
Φ'

图 3