

[19] 中华人民共和国国家知识产权局

[51] Int. Cl.

H03M 7/30 (2006.01)

H04N 7/24 (2006.01)



[12] 发明专利申请公开说明书

[21] 申请号 200510095898.6

[43] 公开日 2006年10月18日

[11] 公开号 CN 1848690A

[22] 申请日 2005.9.7

[21] 申请号 200510095898.6

[30] 优先权

[32] 2004. 9. 17 [33] US [31] 60/610,674

[32] 2005. 1. 4 [33] US [31] 11/029,722

[71] 申请人 广州广晟数码技术有限公司

地址 510640 广东省广州市天河区能源路华南理工大学科技园2号楼6楼

[72] 发明人 游余立

[74] 专利代理机构 永新专利商标代理有限公司

代理人 刘炳胜

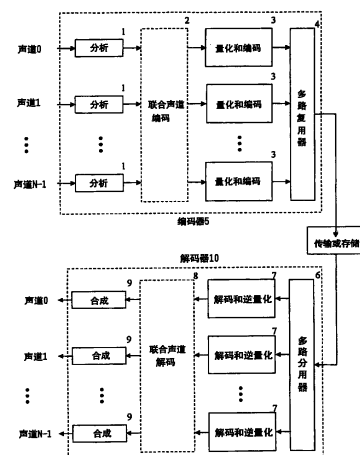
权利要求书 12 页 说明书 34 页 附图 18 页

[54] 发明名称

多声道数字音频编码设备及其方法

[57] 摘要

一种低比特率的数字音频编码系统，包括一个编码器，基于量化指数的局部特性而把码书分配到多组量化指数，从而使码书应用范围与量化边界无关。本发明还包括一个分辨率滤波器组或三模式分辨率滤波器组，如当在一帧中检测到暂态时，其可有选择地在高、低频率分辨率模式或高、低和中模式之间切换。这样得到的多声道音频信号具有极大地降低的比特率，用于有效的发送或存储。解码器实质上是编码器的逆向结构和方法，产生一个听觉上不能与原始信号区分开的被还原的音频信号。



1. 一个用于编码和解码多声道数字音频信号的方法，包括下列步骤：
把输入的PCM样本聚类分割成准稳态帧；
把PCM样本转换成子带样本；
把子带样本分块量化成众多的量化指数；
提供预先设计好的码书库；
基于量化指数的局部特性把码书分配给多组量化指数，从而使得码书应用范围与块量化边界无关；
对码书指数及其各自的应用范围进行编码；
创建一个完整的编码数据流；
发送该完整的编码数据流；
接收该编码数据流并解包该数据流；
从数据流中解码量化指数；
从被解码的量化指数中重建子带样本；和
从重建的子带样本中重建音频PCM样本。
2. 权利要求1的方法，其中，上述码书分配步骤包括：通过把能够容纳上述量化指数的最小可用码书的指数分配给每个量化指数从而把量化指数转换成码书指数，并把码书指数聚类分割成多个应用范围。
3. 权利要求1的方法，其中，准稳态帧的持续时间在2到50毫秒之间。
4. 权利要求1的方法，其中，上述转换步骤包括：使用一个可有选择地在高、低频率分辨率模式之间切换的分辨率滤波器组。

5. 权利要求4的方法，包括检测暂态步骤，当没有检测到暂态时使用高频率分辨率模式，以及当检测到暂态时，切换到低频率分辨率模式。

6. 权利要求5的方法，其中，随着把分辨率滤波器组切换到低频率分辨率模式，子带样本就被聚类分割成准稳态段。

7. 权利要求4的方法，其中，分辨率滤波器组被配置成包括一个长窗口，其能够桥接从短窗口立即到另一个短窗口的转换，以便于处理只间隔一个长窗口的暂态。

8. 权利要求1的方法，其中，上述转换步骤包括使用一个可有选择地在高分辨率模式、低分辨率模式和中间分辨率模式之间切换的分辨率滤波器组，以便在单个帧中采用多个分辨率。

9. 权利要求8的方法，其中，分辨率滤波器组被配置成包括一个窗口，该窗口可桥接从一个更短窗口立即到另一个更短窗口的转换，以便于处理只间隔一个这类窗口的暂态。

10. 权利要求6的方法，包括：使用一个任意分辨率的滤波器组或自适应差分脉码调制(ADPCM)来修整每个稳态段的频率分辨率。

11. 权利要求1的方法，包括计算一个噪声掩蔽阈值的步骤。

12. 权利要求11的方法，其中，计算步骤是用一个心理声学模型来执行。

13. 权利要求1的方法，其中，创建多个量化指数的步骤包括：使用由把比特资源分配到多组子带样本中的比特分配器所提供的量化步长对

子带样本进行量化，以使量化噪声功率低于一个掩蔽阈值。

14. 权利要求1的方法，包括把左右声道对的子带样本转换到和差声道对中。

15. 权利要求14的方法，其中，上述转换步骤用和/差编码器来执行。

16. 权利要求1的方法，包括：对比源声道来提取联合声道的强度比例因子，把联合声道合并到源声道，并且丢弃联合声道中的所有子带样本。

17. 权利要求16的方法，其中，上述提取和合并步骤用一个联合强度编码器来执行。

18. 权利要求1的方法，包括：当一帧中存在暂态时量化指数进行重新排列以降低比特总数。

19. 权利要求1的方法，包括：提供一个游程长度编码器，用于编码码书的应用范围。

20. 权利要求1的方法，包括：当检测到暂态时应用一个暂态聚类分割算法。

21. 权利要求1的方法，其中，上述组合步骤用一个多路复用器来执行。

22. 权利要求1的方法，其中，编码数据流包括码书指数及其应用范围，包括码书的个数、应用范围和码书指数。

23. 权利要求1的方法，其中，当编码数据流指示当前帧是用可切换分辨率分析滤波器组以低频率分辨率模式来编码时，可变合成分辨率滤波

器组用作一个二级混合滤波器组，其中，第一级包括一个任意分辨率的合成滤波器组或一个逆自适应差分脉码调制（ADPCM），而第二级是可变合成滤波器组的低频率分辨率模式。

24. 权利要求1的方法，其中，当数据流指示当前帧是用可切换分辨率的分析滤波器组以高频率分辨率模式来编码时，可变分辨率的合成滤波器组以高频率分辨率模式操作。

25. 权利要求1的方法，其中，上述解包数据流的步骤用一个多路解复用器来执行。

26. 权利要求1的方法，其中，上述解码步骤利用熵解码器解码出熵编码的码书，并用游程长度解码器从数据流中解码出其各自的应用范围。

27. 权利要求1的方法，其中，上述解码步骤还包括用熵解码器从数据流中解码出子带量化指数。

28. 权利要求27的方法，包括利用被解码出的量化指数重建量化单元的个数。

29. 权利要求1的方法，包括：在当前帧中检测到暂态时重新排列量化指数。

30. 权利要求29的方法，其中，上述重新排列步骤用一个去交错器来执行。

31. 权利要求1的方法，包括：用联合强度比例因子从源声道的子带样本中重建联合声道的子带样本。

32. 权利要求31的方法，其中，上述重建步骤用一个联合强度解码器来执行。

33. 权利要求1的方法，包括：从和差子带声道重建左右声道的子带样本。

34. 权利要求33的方法，其中，上述重建步骤用一个和/差解码器来执行。

35. 一个用于编码多声道数字音频信号的方法，包括下列步骤：

把输入的PCM样本分成准稳态帧；

把PCM样本转换成子带样本；

通过在子带样本中创建块量化边界来创建众多的量化指数；

提供预设计好的码书库；

基于量化指数的局部特性把码书分配给多组量化指数，从而使码书应用范围与块量化边界无关；

对码书指数及其各自的应用范围进行编码；和

创建一个用于存储或发送的完整的编码数据流。

36. 权利要求35的方法，其中，上述码书分配步骤包括：通过把可容纳量化指数的最小可用码书分配给每个量化指数，从而把量化指数转换成码书指数。

37. 权利要求36的方法，其中，准稳态帧的持续时间在2到50毫秒之间。

38. 权利要求35的方法，其中，上述转换步骤包括：使用一个可有

选择地在高、低频率分辨率模式之间切换的分辨率滤波器组。

39. 权利要求38的方法，包括检测暂态的步骤，并且当没有检测到暂态时使用高频率分辨率模式；和当检测到暂态时，切换到低频率分辨率模式。

40. 权利要求39的方法，其中，随着把分辨率滤波器组切换到低频率分辨率模式，子带样本被聚类分割成稳态段。

41. 权利要求40的方法，包括：使用一个任意分辨率的滤波器组或自适应差分脉码调制(ADPCM)来修整每个稳态段的频率分辨率。

42. 权利要求41的方法，其中，分辨率滤波器组被配置成包括一个长窗口，用于桥接从短窗口立即到另一个短窗口的转换，以便于处理只间隔一个长窗口的暂态。

43. 权利要求35的方法，其中，上述转换步骤包括步骤：使用一个可有选择地在高、低、和中频率分辨率模式之间切换的分辨率滤波器组，以便当检测到暂态时可以在单帧中采用多个分辨率。

44. 权利要求43的方法，其中，分辨率滤波器组被配置成包括一个窗口，用于桥接从一个更短窗口立即到另一个更短窗口的转换，以便于处理只间隔一个这类窗口的暂态。

45. 权利要求35的方法，其中，创建多个量化指数的步骤包括：使用由把比特资源分配到多组子带样本中的比特分配器所提供的量化步长，以使量化噪声功率低于一个掩蔽阈值。

-
46. 权利要求35的方法，包括一个计算噪声掩蔽阈值的步骤。
47. 权利要求46的方法，其中，计算步骤用一个心理声学模型来执行。
48. 权利要求35的方法，包括把左右声道对的子带样本转换到和差声道对。
49. 权利要求48的方法，其中，转换步骤用一个和/差编码器来执行。
50. 权利要求35的方法，包括：对比源声道来提取联合声道的强度比例因子，把联合声道合并到源声道，并且丢弃联合声道的所有的相关子带样本。
51. 权利要求50的方法，其中，提取和合并步骤用一个联合强度编码器来执行。
52. 权利要求35的方法，包括：当帧中存在暂态时重新排列量化指数以降低比特总数。
53. 权利要求35的方法，包括：提供一个游程长度编码器，用于编码书的应用边界。
54. 权利要求35的方法，包括：当检测到暂态时应用一个暂态聚类分割算法。
55. 权利要求35的方法，其中，创建完整的数据流的步骤是利用一个多路复用器来执行的。
56. 一个用于编码和发送多声道数字音频信号的方法，包括下列步

骤：

把输入的PCM样本分成准稳态帧；

用一个可有选择地在高、低和中频率分辨率模式之间切换的分辨率滤波器组把PCM样本转换成子带样本，以便当检测到暂态时在单个帧中采用多个分辨率；

检测暂态， 当没有检测到暂态时用高频率分辨率模式，当检测到暂态时切换到低或中频率分辨率模式，其中，随着切换分辨率滤波器组，子带样本被分成稳态段，以及该帧中每个稳态段的频率分辨率用相同帧中的低或中频率分辨率模式来调整；

通过在子带样本中创建块量化边界来创建多个量化指数；

提供预设计的码书库；

基于量化指数的局部特性把码书分配到给多组量化指数，从而使码书应用范围与块量化边界无关；

编码码书指数及其各自的应用范围；和

用多路复用器来创建一个用于存储或发送的完整数据流。

57. 权利要求56的方法，其中，码书分配步骤包括：通过把可容纳上述指数的最小可用的码书分配给每个量化指数，从而把量化指数转换成码书指数。

58. 权利要求56的方法，其中，创建多个量化指数的步骤包括：使用由把比特资源分配到子带样本群中的比特分配器所提供的步长，从而每个子带的量化噪声功率低于计算出的掩蔽阈值。

59. 权利要求56的方法，包括：用一个心理声学模型来计算掩蔽阈值。

60. 权利要求56的方法，包括：用和/差编码器把左右声道中的子带样本转换到和差声道中。

61. 权利要求56的方法，包括：使用一个联合强度编码器来提取对比源声道的联合声道的强度比例因子，把联合声道合并成源声道，并且丢弃联合声道中所有的相关子带样本。

62. 权利要求56的方法，包括：提供一个游程长度编码器，用于编码码书的应用边界。

63. 权利要求56的方法，其中，分辨率滤波器组被配置成包括一个窗口，用于桥接从一个更短窗口立即到另一个更短窗口的转换，以便于处理只间隔一个这类窗口的暂态。

64. 一个用于解码经编码的音频数据流的方法，包括下列步骤：

接收编码音频数据流并解包该数据流；

从数据流中解码量化指数；

从被解码的量化指数中重建子带样本；和

使用一个可在低和高频率分辨率模式之间切换的可变分辨率的合成滤波器组，从重建的子带样本中重建音频脉冲编码调制(PCM)样本；

其中，当数据流指示当前帧是用低频率分辨率模式的切换分辨率分析滤波器组来编码时，可变合成分辨率滤波器组用作一个二级混合滤波器组，其中，第一级包括一个任意分辨率的合成滤波器组或一个逆自适应差

分脉码调制(ADPCM), 第二级是可变合成滤波器组的低频率分辨率模式;
以及

其中, 当数据流指示当前帧是用高频率分辨率模式的可切换分辨率的分析滤波器组来编码时, 可变分辨率合成滤波器组以高频率分辨率模式操作。

65. 权利要求64的方法, 其中, 解包数据流的步骤用一个多路解复用器来执行。

66. 权利要求64的方法, 其中, 解码步骤是用解码熵码书的熵解码器和用于从数据流中解码其各自应用范围的游程长度解码器来执行。

67. 权利要求66的方法, 其中, 解码步骤还包括用熵解码器从数据流中解码量化指数。

68. 权利要求67的方法, 包括: 从被解码的量化指数中重建量化单元的个数。

69. 权利要求67的方法, 包括: 在当前帧中检测到暂态时重新排列量化指数。

70. 权利要求69的方法, 其中, 重新排列步骤用一个去交错器来执行。

71. 权利要求64的方法, 包括: 用联合强度比例因子从源声道的子带样本中重建联合声道的子带样本。

72. 权利要求71的方法, 其中, 重建步骤用一个联合强度解码器来

执行。

73. 权利要求64的方法，包括：从和差子带声道中重建左右声道的子带样本。

74. 权利要求73的方法，其中，重建步骤用一个和/差解码器来执行。

75. 权利要求64的方法，其中，分辨率滤波器组被配置成包括一个窗口，用于桥接从一个短窗口立即到另一个短窗口的转换，以便于处理只间隔一个长窗口的暂态。

76. 一个用于解码经编码的音频比特数据流的方法，包括下列步骤：

接收编码音频数据流并解包该数据流；

从数据流中解码量化指数；

从被解码的量化指数中重建子带样本；和

通过使用一个可在低、中、和高频率分辨率模式之间切换的可变分辨率的合成滤波器组，从重建的子带样本中重建音频脉冲编码调制(PCM)样本；

其中，当数据流指示当前帧是用高频率分辨率模式的可切换分辨率的分析滤波器组来编码时，可变分辨率的合成滤波器组以高频率分辨率模式操作；和

其中，当数据流指示当前帧被聚类分割并且这些段是用低或中频率的分辨率模式的可切换分辨率的分析滤波器组来编码时，对于该帧的每一段，可变分辨率的合成滤波器组相应地以低或中频率分辨率模式操作。

77. 权利要求76的方法，其中，解包数据流的步骤用一个多路解复

用器来执行。

78. 权利要求76的方法，其中，解码步骤是利用解码熵码书的熵解码器和用于从数据流中解码其各自应用范围的游程长度解码器来执行。

79. 权利要求78的方法，其中，解码步骤还包括用熵解码器从数据流中解码量化指数。

80. 权利要求79的方法，包括：从被解码的量化指数中重建量化单元的个数。

81. 权利要求79的方法，包括：当在当前帧中检测到暂态时重新排列量化指数。

82. 权利要求81的方法，其中，重新排列步骤用一个去交错器来执行。

83. 权利要求76的方法，包括：用联合强度比例因子从源声道的子带样本中重建联合声道的子带样本。

84. 权利要求83的方法，其中重建步骤用联合强度解码器来执行。

85. 权利要求76的方法，包括：从和差子带声道中重建左右声道的子带样本。

86. 权利要求85的方法，其中，重建步骤用一个和/差解码器来执行。

87. 权利要求76的方法，其中，分辨率滤波器组被配置成包括一个窗口，其能够桥接从一个更短窗口立即到另一个更短窗口的转换，以便于处理只间隔一个这类窗口的暂态。

多声道数字音频编码设备及其方法

相关申请

本申请要求2004年9月17日申请的美国临时申请60/610,674的优先权。

发明背景

本发明通常涉及用于编码和解码多声道数字音频信号的方法和系统。更确切地说，本发明涉及一个低比特率的数字音频编码系统，其在获得透明的音频信号再现的同时大大降低了多声道音频信号的比特率以便进行有效的发送或存储，甚至连听测专家也不能区分在解码器端还原的音频信号与原始信号。

多声道数字音频编码系统通常包括下列元件：时间-频率分析滤波器组，其产生输入的PCM(脉冲编码调制)样本的一个频率表示、叫做子带样本或子带信号；心理声学模型，其基于人耳的听觉特性来计算一个掩蔽阈值，而低于该掩蔽阈值的量化噪声不太可能被听见；全局比特分配器，其向每组子带样本分配比特资源，以便得到的量化噪声功率低于掩蔽阈值；多个量化器，其根据被分配的比特来量化子带样本；多个熵编码器，其降低量化指数中的统计冗余性；和最后的多路复用器，其把量化指数的熵编码及其它辅助信息打包成一个完整的比特流。

例如，杜比AC-3用可切换窗口尺寸的高频率分辨率的MDCT(改进的

离散余弦变换)滤波器组把输入PCM样本映射到频域中。稳态信号用512点的窗口来分析，而暂态信号与256点的窗口来分析。来自MDCT的子带信号被表示为指数/尾数并随后被量化。采用前向-后向自适应的心理声学模型来优化量化并减少编码比特分配信息所需的比特。为了降低解码器的复杂度而不使用熵编码。最后，量化指数及其它辅助信息被多路复用成一个完整的AC-3比特流。AC-3中配置的自适应MDCT的频率分辨率没有很好地与输入信号特性相匹配，因此它的压缩特性受到很大的限制。熵编码的缺少是限制其压缩特性的另一个因素。

MPEG1&2层III(MP3)使用一个32频带的多相滤波器组，其中的每个子带滤波器都跟随有一个在6和18点之间切换的自适应MDCT。一个高级心理声学模型被用来指导其比特分配和标量非均匀量化。哈夫曼(Huffman)码被用来编码量化指数和大部分的其它辅助信息。混合滤波器组的较差的频率隔离极大地限制了它的压缩特性而且具有很高的算法复杂性。

DTS相干声学采用一个32频带的多相滤波器组以获得输入信号的低分辨率频率表示。为了补偿较差的频率分辨率，ADPCM(自适应差分脉码调制)被选择性地用于每个子带。如果ADPCM产生一个良好的编码增益，则均匀标量量化被直接应用于子带样本或应用于预测残差。矢量量化可以选择性地被应用到高频率的子带。哈夫曼码可以选择性地被应用到标量量化指数及其它辅助信息。因为多相滤波器组+ADPCM的结构根本不能提供良好的时间和频率分辨率，所以它的压缩特性很低。

MPEG 2 AAC和MPEG 4 AAC采用一个自适应MDCT滤波器组，其窗口尺寸可以在256和2048之间切换。心理声学模型产生的掩蔽阈值被用来指导其标量非均匀量化和比特分配。哈夫曼码被用来编码量化指数和大部分的其它辅助信息。诸如TNS(暂时噪声整形)、增益控制(类似于MP3的混合滤波器组)、频谱预测(子带内的线性预测)之类的许多其它的工具被用来进一步增强它的压缩特性，而这以极大地增加了算法复杂性为代价。

因此，仍然需要一个低比特率的音频编码系统，其极大地降低了多声道音频信号的比特率以用于有效发送或存储，而同时也能获得透明的音频信号再现。本发明满足了这个需要并提供了其它的相关优点。

发明内容

在以下讨论中，术语"分析/合成滤波器组"等指的是执行时间-频率的分析/合成的设备或方法。它可以非限制性地包括如下：

- 酉变换；
- 临界采样的、均匀的、或非均匀的带通滤波器组时变或非时变组；
- 谐波或正弦波的分析器/合成器。

多相滤波器组、DFT(离散傅里叶变换)、DCT(离散余弦变换)以及MDCT是一些被广泛使用的滤波器组。术语"子带信号或子带样本"等指的是出自分析滤波器组和进入合成滤波器组的信号或样本。

本发明的一个目的是为多声道音频信号的低比特率编码提供与现有技术同样水平的压缩性能但却降低了算法复杂性。

这由编码器在编码端侧完成，编码器包括：

1)成帧器，用于把输入的PCM样本聚类分割成准稳态帧，其大小是分析滤波器组的子带数的整数倍，并且其时间范围是2到50ms。

2)暂态检测器，用于检测该帧中暂态的存在。一个实施例是根据取子带距离测量的阈值，阈值从低频率分辨率模式下的分析滤波器组的子带样本中获得。

3)可变分辨率的分析滤波器组，用于把输入的PCM样本转换成子带样本，它可以用下列之一来执行：

a)滤波器组，可以在高、中、低频率分辨率模式之间切换其操作。高频率分辨率模式用于稳态帧，而中、低频率分辨率模式用于具有暂态的帧。在一个暂态帧内，低频率分辨率模式被用于暂态段，而中间分辨率模式被用于该帧的剩余部分，在这个架构之下存在三类帧：

i)滤波器组只以高频率分辨率模式操作来处理的稳态帧；

ii)滤波器组以中、高时间分辨率模式操作来处理的暂态帧；

iii)滤波器组只以中间分辨率模式操作处理的慢暂态帧；

两个优选实施例被给出如下：

i)DCT实现，其中，三个级别的分辨率对应于三个DCT块长度；

ii)MDCT实现，其中，三个级别的分辨率对应于三个MDCT块长度或窗口长度。定义多个窗口类型以桥接这些窗口之间的转换。

b)混合滤波器组，其基于一个可以在高、低分辨率模式之间切换其操作的滤波器组；

i)在当前帧中不存在暂态时，它切换到高频率分辨率模式以确保持稳态段的高压缩性能；

ii)在当前帧中存在暂态时，它切换到低频率分辨率/高时间分辨率模式以避免前向回声效应。这个低频率分辨率模式还跟随有一个暂态聚类分割级，其把子带样本分成稳态段，然后可选地在每个子带中后跟一个任意分辨率的滤波器组或ADPCM，如果被选择的话，可用于向每个稳态段提供适合的频率分辨率。

给出两个实施例，其中，一个基于DCT而另一个基于MDCT。给出两个暂态段的实施例出，其中，一个基于取阈值而另一个基于k均值算法，两个实施例都使用子带距离测量。

2)计算掩蔽阈值的心理声学模型。

3)可选的和/差编码器，其把左右声道对中的子带样本转换成和/差声道对。

4)可选的联合强度编码器，其对比源声道来提取联合声道的强度比例因子(引导向量)，将联合声道合并到源声道中，并丢弃联合声道中的各个子带样本。

5)全局比特分配器，其把比特资源分配给多组子带样本，以便它们的量化噪声功率低于掩蔽阈值。

6)标量量化器，其用比特分配器提供的步长来量化所有的子带样本。

7)可选的交错器，当帧中存在暂态时，其被选择性地用来重新排列量化指数以便于降低比特总数。

8)熵编码器，其基于量化指数的局部统计特征把最佳的码书从码书库分配给多组量化指数，包括下列步骤：

a)把最佳码书分配给每个量化指数，因此实质上把量化指数转换成码书指数。

b)把这些码书指数分成很大的段，段边界定义了码书的应用范围。

一个优选实施例是：

c)把量化指数分块为区组(*granule*)，每个区组包括固定数目的量化指数。

d)确定每个区组的最大码书需求。

e)把最小码书分配给一个能容纳其最大码书需求的区组：

f) 清除那些码书指数比其近邻的码书指数小的孤立的小块区域；那些对应于零量化指数的码书指数的孤立小块区域可以不经这样的处理。

用于对编码码书应用范围进行编码的一个优选实施例使用了游程长度码。

9)熵编码器, 其用码书及其由熵码书选择器确定的应用范围来编码所有的量化指数。

10)多路复用器, 其把量化指数的所有熵代码和辅助信息打包成一个完整的比特流, 这样构造是为了量化指数出现在用于量化步长的指数之前。这个构造使得不必要把每个暂态段的量化单元数打包进比特流, 因为它可以从被解包的量化指数中恢复。

本发明的解码器包括:

- 1)多路解复用器, 用于从比特流解包不同的码字;
- 2)量化指数码书解码器, 用于从比特流中解码用于量化指数的熵码书及其各个应用范围 (application range) ;
- 3)熵解码器, 用于从比特流中解码量化指数;
- 4)可选的去交错器, 在当前帧中存在暂态时, 其选择性地从新排列量化指数;
- 5)量化单元个数再造器, 其用下列步骤从量化指数中重建每个暂态段的量化单元个数:
 - a)为每个暂态段找到具有非零量化指数的最大子带;
 - b)找到能容纳这个子带的最小临界频带, 这就是这个暂态段的量化单元个数;
- 6)步长解包器, 其解包所有量化单元的量化步长;

7)逆量化器，其从量化指数和步长中重建子带样本；

8)可选的联合强度解码器，其利用联合强度比例因子(引导向量)从源声道的子带样本中重建联合声道的子带样本；

9)可选的和/差解码器，其从和/差声道的子带样本中重建左右声道的子带样本；

10)可变分辨率的合成滤波器组，其从子带样本中重建音频 PCM 样本，这可以通过以下来执行：

a)合成滤波器组，能够在高、中、低分辨率模式之间切换其操作；

b)混合合成滤波器组，其是基于一个能够在高、低分辨率模式之间切换的合成滤波器组；

i)当比特流指示当前帧是用可变换分辨率的分析滤波器组以低频率分辨率模式来编码时，这个合成滤波器组是一个二级混合滤波器组，其中，第一级是一个任意分辨率的合成滤波器组或一个逆 ADPCM，而第二级是可在高、低频率分辨率模式之间切换的自适应合成滤波器组的低频率分辨率模式；

ii)当比特流指示当前帧是用可变换分辨率的分析滤波器组以高频率分辨率模式来编码时，这个合成滤波器组只不过是高频率分辨率模式下的可变换分辨率的合成滤波器组。

最后，本发明提供了一个低编码延迟模式，这个模式在可切换分辨率分析滤波器组的高频率分辨率模式被编码器禁止时被启动，并且帧长随后

被减小到在低频率分辨率模式下的可切换分辨率滤波器组的块长或其整数倍。

根据本发明, 编码多声道数字音频信号的方法通常包括从多声道数字音频信号创建PCM样本和把该PCM样本转换成子带样本的步骤。具有边界的多个量化指数通过量化子带样本而被创建。通过把预先设计的码书库中能够容纳量化指数的最小的码书分配给每个量化指数, 量化指数被转换成码书指数。在创建用于存储或发送的编码数据流之前, 码书指数被聚类分割和编码。

一般来说, PCM样本被输入到持续时间在2到50毫秒(ms)之间的准稳态帧中。掩蔽阈值可使用例如一个心理声学模型来计算。比特分配器把比特资源分配到多组子带样本中, 以便量化噪声功率低于掩蔽阈值。

转换步骤包括: 使用一个有选择地在高、低频率分辨率模式下切换的分辨率滤波器组。检测暂态, 当没有检测到暂态时使用高频率分辨率模式; 然而, 当检测到暂态时, 分辨率滤波器组被切换到低频率分辨率模式。随着把分辨率滤波器组切换到低频率分辨率模式, 子带样本就被分成稳态段。每个稳态段的频率分辨率用任意分辨率的滤波器组或自适应差分脉码调制来修整。

量化指数可以在帧中存在暂态时被重新排列以降低比特总数。游程长度编码器可用于编码最佳熵码书的应用边界, 可以采用聚类分割算法。

和/差编码器可以被用来把左右声道对中的子带样本转换到和/差声道对中。此外, 联合强度编码器可用于对比源声道来提取联合声道的强度比

例因子，把联合声道合并成源声道，并且丢弃联合声道中所有的相关子带样本。

一般来说，创建一个完整的比特数据流的组合步骤通过在存储或向解法器发送编码数字音频信号之前使用一个多路复用器来执行。

解码音频数据比特流的方法包括：如通过使用一个多路解复用器来接收编码音频数据流并解包该数据流。熵码书指数及其各自的应用范围被解码。这可能涉及游程长度和熵解码器。它们还被用来解码量化指数。

当在当前帧中检测到暂态时，量化指数如通过用去交错器来重新排列。子带样本然后从被解码的量化指数中重建。通过使用可在低和高频率分辨率模式之间切换的可变分辨率的合成滤波器组，音频PCM样本从重建的子带样本中被重建。当数据流指示当前帧是用可切换分辨率分析滤波器组以低频率分辨率模式来编码时，可变合成分辨率滤波器组用作一个二级混合滤波器组，其中，第一级包括一个任意分辨率的合成滤波器组或一个逆自适应差分脉码调制，第二级是可变合成滤波器组的低频率分辨率模式。当数据流指示当前帧是用可切换分辨率的分析滤波器组以高频率分辨率模式来编码时，可变分辨率合成滤波器组在高频率分辨率模式下操作。

一个联合强度解码器可用于用联合强度比例因子从源声道子带样本中重建联合声道子带样本。此外，和/差解码器可以被用来从和/差声道子带样本中重建左右声道的子带样本。

本发明结果是一个低比特率的数字音频编码系统，其极大地降低了多声道音频信号的比特率以用于有效发送，同时还获得透明的音频信号再

现，以致于很难将它与原始信号区分。

本发明的其它特征和优点将参考附图从下列详细说明中变得明显，其通过举例的方式来说明本发明的原理。

附图说明

下列附图用来说明本发明。在这些附图中：

图1是一个示意图，描述根据本发明的多声道数字音频信号的编码和解码；

图2是一个示意图，说明了根据本发明使用的一个示例性编码器；

图3是具有任意分辨率的滤波器组的可变分辨率的分析滤波器组的一个示意图；

图4是具有ADPCM的可变分辨率的分析滤波器组的一个示意图；

图5是根据本发明的用于可切换MDCT窗口类型的示意图；

图6是根据本发明的暂态段的一个示意图；

图7是根据本发明的具有两个分辨率模式的可切换滤波器组的一个应用示意图；

图8是根据本发明的具有三个分辨率模式的可切换滤波器组的一个应用示意图；

类似于图5，图9是根据本发明的用于具有三个分辨率模式的可切换MDCT的其它窗口类型的示意图；

图10描述了根据本发明的具有三个分辨率模式的可切换MDCT窗口序列的一组例子；

图11是本发明与先有技术相比的熵码书的确定示意图；

图12是根据本发明把码书指数分成很大的段或消除码书指数的孤立的小块区域的示意图;

图13是本发明配备的解码器的示意图;

图14是根据本发明的具有任意分辨率的滤波器组的可变分辨率的合成滤波器组的一个示意图;

图15是具有逆ADPCM的可变分辨率合成滤波器组的一个示意图;和

图16是根据本发明当使用半混合滤波器组或可切换滤波器组+ADPCM时的比特流的结构示意图。

图17是在处理只间隔一帧的暂态时,短到短转换的长窗口的优点示意图。

图18是根据本发明当使用三模式可切换滤波器组时的比特流的结构示意图。

具体实施方式

如附图中所示,为了说明的目的,本发明涉及一个低比特率数字音频编码和解码系统,其极大地降低了多声道音频信号的比特率以用于有效发送或存储,同时也实现了透明的音频再现。即,多声道编码的音频信号比特率通过使用算法复杂度较低的系统来减小,而且即使是听测专家也无法区分在解码器端上还原的音频信号与原始信号。

如图1中所示,本发明的编码器5将多声道音频信号作为输入并将其编码成比特流,并且极大地降低了比特率以适于在声道容量有限的媒介上发送或存储。只要接收到由编码器5产生的比特流,解码器10就对其进行

解码并重建甚至听测专家也不能将其与原始信号区别的多声道音频信号。

在编码器5和解码器10内部，多声道音频信号被作为离散声道来处理。即，每个声道与其它声道同样地来对待，除非清楚地指定了联合声道编码2。这在图1中用极度简化的编码器和解码器结构做出了说明。

利用这种极度简化的编码器结构，其编码处理过程说明如下。来自每个声道的音频信号首先在分析滤波器组的第一级1中被分解成子带信号。来自所有声道的子带信号被选择性地送到联合声道编码器2，其通过组合对应于来自不同声道的相同频带的子带信号，采用人耳的听觉特性来降低比特率。可以在2中联合编码的子带信号然后被量化并在3中被编码。量化指数或它们的熵编码以及来自所有声道的辅助信息然后在4中被多路复用成一个完整的比特流以用于发送或存储。

在解码端上，比特流首先在6中被多路解复用为辅助信息和量化指数或其熵编码。熵编码在7中被解码(注意：诸如哈夫曼码之类的前缀码的熵解码和多路解复用通常在一个单个步骤中执行)。子带信号在7中利用量化指数和由辅助信息携带的步长被重建。如果在编码器中使用联合声道编码，则联合声道解码在8中被执行。然后，每个声道的音频信号在合成级9中利用子带信号被重建。

上述极度简化的编码器和解码器结构被用来单独说明本发明给出的编码和解码方法的离散特性。实际应用于音频信号每个声道的编码和解码方法差别悬殊并且更加复杂。除非另作说明，则这些方法在音频信号的一个声道环境中被描述如下。

编码器

编码音频信号的一个声道的通用方法在图2中被描述如下：

成帧器11把持续时间从2到50ms的输入PCM样本分成准稳态帧。一帧中PCM样本的确切个数必须是在可变分辨率的时间-频率分析滤波器组13中使用的不同滤波器组的子带最大个数的整数倍。假定子带的最大数是N，那么一帧中PCM样本的个数是

$$L = k \cdot N$$

其中，k是一个正整数。

暂态分析12检测当前输入帧中暂态的存在并将该信息传递给可变分辨率分析组13。

在这里可以采用任何已知的暂态检测方法。在本发明的一个实施例中，PCM样本的输入帧被送到可变分辨率的分析滤波器组的低频率分辨率模式。让 $s(m,n)$ 表示来自这个滤波器组的输出样本，其中，m是子带指数而n是子带域中的时间指数（temporal index）。在以下讨论中，术语“暂态检测距离”等指的是为每个时间指数定义的距离测量：

$$E(n) = \sum_{m=0}^{M-1} |s(m,n)|$$

或

$$E(n) = \sum_{m=0}^{M-1} s^2(m,n)$$

其中，M是滤波器组的子带个数。其它类型的距离测量也可以用类似

的方法被采用。让 $E_{\max} = \underset{n}{\text{Max}} E(n)$ 和 $E_{\min} = \underset{n}{\text{Min}} E(n)$ 是这个距离的最大和最小值，如果

$$\frac{E_{\max} - E_{\min}}{E_{\max} + E_{\min}} > \text{Threshold} \quad (\text{阈值})$$

则声明存在暂态，其中，阈值可以被设置为0.5。

本发明使用一个可变分辨率的分析滤波器组13。存在许多已知的方法来实现可变分辨率的分析滤波器组。一个突出的方法是使用可以在高、低频率分辨率模式之间切换其操作的滤波器组，高频率分辨率模式用于处理音频信号的稳态段而低频率分辨率模式用于处理暂态。然而，分辨率的切换由于理论和实践的约束不能及时任意地发生。相反，它通常发生在帧分界，即帧用高频率分辨率模式或低频率分辨率模式来处理。如图7中所示，对于暂态帧131，滤波器组已经切换到低频率分辨率模式以避免前向回声效应。因为暂态132本身是很短的，而该帧的前暂态133和后暂态134段又长得多，所以低频率分辨率模式的滤波器组显然与这些稳态段不匹配。这极大地限制了整个帧所能达到的总的编码增益。

本发明提出了三个方法来解决这个问题。基本思想是在可切换分辨率结构内为暂态帧的稳态部分(stationary majority)提供一个较高频率分辨率。

半混合滤波器组

如图3中所示，它实质上是一个混合滤波器组，包括一个可以在高、低频率分辨率模式之间切换的可切换分辨率的分析滤波器组28，并且在

低频率分辨率模式24时，后面跟随有一个暂态聚类分割单元25，然后在每个子带中有一个可选的任意分辨率的分析滤波器组26。

当暂态检测器12没有检测到暂态存在时，可切换分辨率的分析滤波器组28进入低时间分辨率模式27，其确保高频率分辨率以实现高音频信号编码增益，具有强的音调分量。

当暂态检测器12检测到暂态存在时，可切换分辨率的分析滤波器组28进入高时间分辨率模式24。这确保了用良好的时间分辨率来处理暂态以防止前向回声。如此产生的子带样本如图6中所示被暂态聚类分割部分25分成准稳态段。在以下讨论中，术语"暂态段"等指的是这些准稳态段。这后面是每个子带中的任意分辨率的分析滤波器组26，其子带个数等于每个子带中每个暂态段的子带样本个数。

可切换分辨率的分析滤波器组28能用可以在高、低频率分辨率模式之间切换其操作的任何滤波器组来实现。本发明的一个实施例采用了一对DCT，对应于低和高频率分辨率，其转换长度分别为小和大。假定转换长度为M，则类型4的DCT的子带样本被获得为：

$$s(m, n) = \sqrt{\frac{2}{M}} \sum_{k=0}^{M-1} \cos \left[\frac{\pi}{M} (k+0.5)(n+0.5) \right] \cdot x(mM + k)$$

其中，x(.)是输入PCM样本。其它形式的DCT可以用来代替类型4的DCT。

因为DCT倾向引起成块效应，所以本发明的一个较好的实施例是采用改进的DCT(MDCT)：

$$s(m,n) = \sqrt{\frac{2}{M}} \sum_{k=0}^{2M-1} \cos \left[\frac{\pi}{M} \left(k + 0.5 + \frac{M}{2} \right) (n + 0.5) \right] \cdot w(k) \cdot x(mM - M + k)$$

其中， $w(\cdot)$ 是窗口函数。

窗口函数在每半个窗口中必须是功率对称的：

$$w^2(k) + w^2(M - k) = 1 \quad k=0, \dots, M-1$$

$$w^2(k + M) + w^2(2M - 1 - k) = 1 \quad k=0, \dots, M-1$$

以便于保证理想的重建。

尽管满足上述情况的任何窗口都可以被使用，但只有下列正弦窗口

$$w(k) = \pm \sin \left[(k + 0.5) \frac{\pi}{2M} \right] \quad k=0, \dots, 2M-1$$

具有良好的特性，即输入信号中的直流分量被集中到第一变换系数。

为了当MDCT在高、低频率模式或长、短窗口之间切换时能保持理想的重建，长、短窗口的重叠部分必须有相同的形状。

依赖于输入PCM样本的瞬变特性，编码器可以选择一个长窗口(如图5中的第一窗口61所示)，切换到一个短窗口序列(如图5中的第四窗口64所示)，并返回。图5中的长到短转换的长窗口62和短到长转换的长窗口63是桥接这类切换所需要的。当两个暂态非常接近但不是接近到足以保证短窗口的连续应用时，图5中长窗口65的短到短转换是有用的。编码器需要向解码器传送被用于每一帧的窗口类型，以便相同的窗口被用来重建

PCM样本。

短到短转换的长窗口的优点是可以处理只间隔一帧的邻近暂态。如在图17的顶端67所示，先有技术的MDCT可以处理至少间隔两帧的暂态。如在图17的底部68所示，使用这个短到短转换的长窗口可以将其减少到一帧。

本发明然后将执行暂态段25。通过利用二进制函数值从0到1或1到0的变化，暂态段可以由指示暂态位置的二进制函数或聚类分割边界来表示。例如，图6中的准稳态段可以被表示如下：

$$T(n) = \begin{cases} 0, & n = 0,1,2,3,4 \\ 1, & n = 5,6,7,8,9 \\ 0, & n = 10,11,12,13,14,15,16 \end{cases}$$

注意， $T(n)=0$ 不一定意味着音频信号的能量在时间指数 n 时很高，反之亦然。在以下讨论各处的函数 $T(n)$ 被称为“暂态段函数”等。由这个段函数携带的信息必须被直接或者间接地传送到解码器。编码零-一游程长度的游程长度编码是一个有效的选择。对于上面的具体例子， $T(n)$ 可以用5、5和7的游程长度代码被传送到解码器。游程长度代码还可以被熵编码。

暂态聚类分割部分25可以用任何已知的暂态聚类分割方法来实现。在本发明的一个实施例中，暂态聚类分割可以通过简单地对暂态探测距离取阈值来完成。

$$T(n) = \begin{cases} 0, & \text{如果 } E(n) < \textit{Threshold} \\ 1, & \text{否则} \end{cases}$$

阈值可以被设置为

$$Threshold = k \cdot \frac{E_{\max} + E_{\min}}{2}$$

其中，k是一个可调节的常数。

本发明的一个更高级的实施例是根据k均值聚类算法，其包括下列步骤：

1) 暂态聚类分割函数T(n)被初始化，利用上述取阈值方法获得的结果。

2) 每一类的质心被计算：

$$C0 = \frac{\sum_{\text{如果 } T(n)=0} E(n)}{\sum_{\text{如果 } T(n)=0} 1}, \text{ 对于与 } T(n)=0 \text{ 相关联的类；}$$

$$C1 = \frac{\sum_{\text{如果 } T(n)=1} E(n)}{\sum_{\text{如果 } T(n)=1} 1}, \text{ 对于与 } T(n)=1 \text{ 相关联的类。}$$

3) 暂态聚类分割函数T(n)基于以下规则来分配

$$T(n) = \begin{cases} 0, & \text{如果 } |E(n) - C0| < |E(n) - C1| \\ 1, & \text{否则} \end{cases}$$

4) 进到步骤2。

任意分辨率的分析滤波器组26本质上是一个诸如DCT之类的变换，它的块长等于每个子带段中的样本个数。假定在一帧内每个子带都存在32个子带样本并且它们被分为(9、3、20)，则块长为9、3和20的三个变

换将被分别应用到三个子带段中每一个的子带样本。在以下讨论中，术语"子带段"等指的是子带内暂态段的子带样本。第m个子带的最后段(9、3、20)的变换可以用类型4的DCT来说明如下

$$u(m,n) = \sqrt{\frac{2}{20}} \sum_{k=0}^{20-1} \cos\left[\frac{\pi}{20}(k+0.5)(n+0.5)\right] \cdot s(m,12+k)$$

这个转换将增加每个暂态段内的频率分辨率，所以可以期待一个良好的编码增益。然而在许多情况下，编码增益小于1或者太小，则有利的决策是丢弃这类变换结果并经由辅助信息通知解码器这个决策。由于与辅助信息相关的开销，如果是否丢弃转换结果的决定是根据一组子带段，则它可以改进总的编码增益，即一个比特被用来为一组子带段而不是每个子带段传送这个决策。

在以下讨论中，术语"量化单元"等指的是属于相同心理声学临界频带和暂态段内的一组相连的子带样本。量化单元可以是用于上述决策制定的子带段的一个良好分组。如果这个被使用，则对量化单元中所有的子带段来计算总编码增益。如果编码增益大于1或某些其它的较高阈值，则为量化单元中所有的子带段保留转换结果。否则，该结果被丢弃。只需用一个比特向解码器传送这个用于量化单元中所有子带段的决策。

可切换滤波器组+ADPCM

如图4中所示，它基本上与图3中的相同，只不过任意分辨率的分析滤波器组26被ADPCM29所替代。是否应用ADPCM的决定又是根据诸如量化单元之类的一组子带段，以便于降低辅助信息的成本。该组子带段甚

至可以共享一组预测系数。在此处可以使用量化预测系数的已知方法，比如包括LAR(对数面积比)、IS(反正弦)以及LSP(线谱对)。

三模式可切换的滤波器组

不同于只有高、低分辨率模式的常见可切换滤波器组，这个滤波器组可以在高、中、低分辨率模式之间切换其操作。高、低频率分辨率模式分别是用于稳态和暂态帧，而且遵循与双模式可切换滤波器组相同的一类原则。中间分辨率模式的主要意图是向暂态帧内的稳态段提供较好的频率分辨率。因此，在一个暂态帧内，低频率分辨率模式被用于暂态段，而中间分辨率模式被用于该帧的剩余部分。不同于先有技术，对于单个帧的音频数据，本发明可切换滤波器组以两个分辨率模式操作。中间分辨率模式还可以被用来处理具有平滑暂态的帧。

在以下讨论中，术语"长块"等指的是滤波器组在每个时刻在高频率分辨率模式下输出的一个样本块；术语"中块"等指的是滤波器组在中频分辨率模式下每个时刻输出的一样本块；术语"短块"等指的是滤波器组在低频率分辨率模式下每个时刻输出的一样本块。三种帧可以用这三种定义被描述如下：

- 滤波器组以高频率分辨率模式操作来处理的稳态帧，这类帧中的每一帧通常包括一个或多个长块；
- 滤波器组以高、中时间分辨率模式操作来处理的具有暂态的帧，这类帧中的每一帧都包括几个中块和几个短块，所有短块的样本总数等于一个中块的样本总数；

- 滤波器组以中间分辨率模式下操作来处理的具有平滑暂态的帧，这类帧中的每一帧都包括几个中块。

这个新方法的优点在图8中被示出。图8基本上与图7相同，只不过原先在图7中在低频率分辨率模式下处理的许多段(141、142和143)现在由中频率分辨率模式来处理。因为这些段是稳态的，所以中频率分辨率模式显然比低频率分辨率模式更加匹配，因此可以期待较高的编码增益。

本发明的一个实施例采用具有小、中、大块长的三元组DCT，分别对应于低、中、高频率的分辨率模式。

本发明的一个较好实施例(无成块效应)采用具有小、中、大的块长的三元组MDCT。由于引入了中间分辨率模式，除了图5中的窗口类型之外还提供了图9中所示的窗口类型。这些窗口被描述如下：

- 中窗口151；
- 长到中转换的长窗口152：作为一个长窗口，其桥接从长窗口到中窗口的转换。
- 中到长转换的长窗口153：作为一个长窗口，其桥接从中窗口到长窗口的转换。
- 中到中转换的长窗口154：作为一个长窗口，其桥接从中窗口到另一个中窗口的转换。
- 中到短转换的中窗口155：作为一个中窗口，其桥接从中窗口到短窗口的转换。

- 短到中转换的中窗口156：作为一个中窗口，其桥接从短窗口到中窗口的转换。
- 中到短转换的长窗口157：作为一个长窗口，其桥接从中窗口到短窗口的转换。
- 短到中转换的长窗口158：作为一个长窗口，其桥接从短窗口到中窗口的转换。

注意：类似于图5中短到短转换的长窗口65，中到中转换的长窗口154、中到短转换的长窗口157、和短到中转换的长窗口158可使三模式MDCT处理间隔一帧的暂态。

图10示出窗口序列的一些例子。161举例说明了这个实施例用中分辨率167处理慢暂态的能力，而162到166说明了向暂态分配精细的时间分辨率168、向同一帧内的稳态段分配中时间分辨率169、和向稳态帧分配高频率分辨率170的能力。

常见的和/差编码方法14可以在这里被应用。例如，一个简单使用方法如下：

$$\text{和声道} = 0.5(\text{左声道} + \text{右声道})$$

$$\text{和声道} = 0.5(\text{左声道} + \text{右声道})$$

常见的联合强度编码方法15可以在这里被应用。一个简单的方法可以是

- 用源和联合声道的和来替换源声道。

- 将其调整为与量化单元内的原始源声道相同的能量级
- 丢弃量化单元内联合声道的子带样本，只把比例因子(被称为"引导向量"或本发明中的"比例因子")的量化指数传送到解码器，其被定义为：

$$\text{引导向量} = \sqrt{\frac{\text{联合声道的能量}}{\text{源声道的能量}}}$$

诸如对数之类的引导向量的非均匀量化将被用来匹配人耳的听觉特性。熵编码可以被应用于引导向量的量化指数。

为了避免源和联合声道在它们的相位差接近180度的情况下的抵消效应，可以在它们被合计形成联合声道时应用极性：

$$\text{和声道} = \text{源声道} + \text{极性} \cdot \text{联合声道}$$

极性还必须被传送到解码器。

心理声学模型23基于人耳的听觉特性来计算音频样本的当前输入帧的掩蔽阈值，低于掩蔽阈值的量化噪声不太可能被听到。在这里可以使用任何常见的心理声学模型，但是本发明要求它的心理声学模型对每个量化单元都输出一个掩蔽阈值。

全局比特分配器16全局地向每个量化单元分配一帧可用的比特资源，以便每个量化单元中的量化噪声功率低于其各自的掩蔽阈值，它通过调节其量化步长来控制每个量化单元的量化噪声功率。量化单元内的所有子带样本都用相同的步长来量化。

在这里可以采用所有已知的比特分配方法。这类方法之一是有名的

Water Filling算法。它的基本思想是发现其QNMR(量化噪声掩蔽比)最高的量化单元,并减少分配给该量化单元的步长以降低量化噪声。它重复这个处理直到所有量化单元的QNMR都小于1(或任何其它的阈值)或当前帧的比特资源耗尽为止。

量化步长本身必须被量化以使其可以被打包到比特流中。诸如对数之类的非均匀量化将被用来匹配人耳的听觉特性。熵编码可以被应用于步长的量化指数。

本发明使用全局比特分配16提供的步长来量化每个量化单元17内的所有子带样本。在这里可以应用所有的线性或非线性的、均匀或非均匀的量化方案。

只有在当前帧中存在暂态时,才可以选择性地调用交错18。让 $x(m,n,k)$ 是第 m 个准稳态段和第 n 个子带中的第 k 个量化指数。 (m,n,k) 通常是量化指数被排列的顺序。交错单元18重新排序量化指数以便它们被排列为 (n,m,k) 。这样做的动机是量化指数的从新排列可以使得这些编码指数所需的比特数比不交错指数时少。是否调用交错的决策需要作为辅助信息传送到解码器。

在先前的音频编码算法中,熵码书的应用范围与量化单元相同,所以熵码书由量化单元内的量化指数来确定(参见图11的顶端)。因此没有用于优化的空间。

本发明在这方面是完全不同的。它在进行到码书选择时忽略了量化单元的存在。相反,它把最佳码书分配给每个量化指数19,因此本质上把

量化指数转换成了码书指数。然后，它把这些码书指数分成较大的段，而段边界定义了码书应用的范围。显然，这些码书应用范围与由量化单元确定的范围相差悬殊。它们仅仅是基于量化指数的品质，因而所选择的码书更适合量化指数。因此，只需要较少的比特把量化指数传送到解码器。

这个方法对比于先前技术的优点在图11中被说明。让我们看看图中最大的量化指数。它属于量化单元d并且利用先前的方法要选择一个大码书，这个大码书显然不是最佳的，因为量化单元d中的大多数指数要小得多。另一方面，通过使用本发明的新方法，相同的量化指数被分成段C，所以它与其它的大量化指数共享一个码书。此外，段D中的所有量化指数都很小，所以一个小码书将被选择。因此，需要较少比特来编码量化指数。

现在参见图12，先有技术的系统只须把码书指数作为辅助信息传送到解码器，因为它们的应用范围与预定的量化单元相同。然而，本发明的方法除了传送码书指数之外还需要把码书应用范围作为辅助信息传送到解码器，因为它们独立于量化单元。如果处理不当，则这个额外开销可能会以更多的比特用于整个辅助信息和量化指数而结束。因此，把码书指数分成大的段对于控制这个开销来说是相当关键的，因为大段意味着较少个数的码书指数及其应用范围需要被传送到解码器。

本发明的一个实施例用下列步骤来完成这个码书选择的新方案：

1)把量化指数分块成区组，每个区组包括P个量化指数。

2)确定每个区组最大码书需求。对于对称量化器来说，这通常由每个区组内最大的绝对量化指数来表示：

$$I_{\max}(n) = \max_{k=0}^{P-1} |I(nP+k)|, n \in \{\text{所有区组}\}$$

其中 $I(\cdot)$ 是量化指数;

3)把最小码书分配给那个可以容纳最大码书需求的区组:

$$B(n) = \min_{\text{所有码本}} \{\text{可以容纳 } I_{\max}(n) \text{ 的码本}\}$$

4) 通过把那些码书指数比其近邻小的孤立的小块区域的码书指数提升到其近邻的码书指数的最小值的方法而把这些孤立的小块区域清除掉。这在图12中由映射71到72、73到74、77到78以及79到80来说明。深入对应于零量化指数的码书指数中的孤立的小块区域可以从这个处理中被除去, 因为这个码书指示没有代码需要被传送。这在图12中被描述为75到76的映射。这个步骤明显地降低了需要被传送到解码器的码书指数的个数及其应用范围。

本发明的一个实施例采用游程长度代码来编码码书应用范围, 并且游程长度代码还可以用熵代码来编码。

所有的量化指数都用由熵码书选择器19确定的码书和和它们各自的应用范围来编码20。

熵编码可以用各种哈夫曼码书来实现。当一个码书中的量化级数很小时, 多个量化指数被归集(blocked)到一起以形成一个大的哈夫曼码书。当量化级的个数(number of quantization levels)太大时(例如超过200), 则采用递归索引。对此, 一个大的量化指数 q 被表示为

$$q = m \cdot M + r$$

其中， M 是模， m 是商，而 r 是余数。只有 m 和 r 需要被传送到解码器。它们中的一个或其两者都可以用哈夫曼码来编码。

熵编码可以用各种各样的算术码书来实现。当量化级个数太大时(例如超过200)，递归索引也将被使用。

其它类型的熵编码也可以被用于上述的哈夫曼和算术编码。

不经过熵编码而直接打包全部或部分量化指数也是一个好的选择。

因为量化指数的统计特性在可变分辨率滤波器组采用低和高分辨率模式时明显不同，所以本发明的一个实施例采用两个熵码书库来分别在这两个模式下编码量化指数。第三个库可以被用于中间分辨率模式，它还可以与高或低分辨率模式共享这个库。

本发明把所有量化指数和其它的辅助信息多路复用成一个完整的比特流。辅助信息包括量化步长、采样率、扬声器配置、帧长、准稳态段的长度、熵码书的代码等。诸如时间码之类的其它辅助信息也可以被打包在比特流中。

先有技术的系统需要把每个暂态段的量化单元个数传送到解码器，因为量化步长的解包、量化指数的码书、和量化指数自身都取决于此。然而在本发明中，因为量化指数码书及其应用范围的选择由熵码书选择的专门方法从量化单元中分离出(decouple)，比特流可以用如此的方法来构成，即量化指数可以在需要量化单元的个数之前被解包。一旦量化指数被解包，它们就能被用来重建量化单元的个数。这将在解码器中来解释。

有了上述考虑,本发明的一个实施例在使用半混合滤波器组或可切换滤波器组+ADPCM时使用一个如图16中所示的比特流结构,它本质上包括以下部分:

- 同步字81: 指示音频数据帧的开始;
- 帧头82: 包括音频信号的相关信息,比如采样率、正常的声道数、LFE(低频效应)声道数、扬声器配置等;
- 声道1、2、...、N、83、84、85: 每个声道的所有音频数据都在此被打包;
- 辅助数据86: 包括诸如时间码之类的辅助数据;
- 错误检测87: 误差检测码在这里被插入以检测当前帧中出现的差错,以便于差错处理程序能够在检测到比特流差错时启动;

每个声道的音频数据还被构造如下:

- 窗口类型90: 指示诸如图5中所示之类的窗口被用于编码器以便解码器能够使用相同的窗口;
- 暂态位置91: 只用于暂态的帧,它指示每个暂态段的位置。如果游程长度代码被使用,则这是每个暂态段的长度被打包的位置;
- 交错决策92: 一个比特,只在暂态帧中,指示是否交错每个暂态段的量化指数以便于解码器知道是否要去交错量化指数;
- 码书指数和应用范围93: 它把所有的关于熵码书及对量化指数的应用范围的信息传送,它包括以下部分:

- 码书个数101: 传送当前声道的每个暂态段的熵码书个数;
- 应用范围102: 按照量化指数或区组来传送每个熵码书的应用范围, 它们还可以用熵代码来编码;
- 码书指数103: 把这个指数传送到熵码书, 它们还可以进一步用熵代码来编码;
- 量化指数94: 传送用于当前声道所有量化指数的熵代码;
- 量化步长95: 把指数传送到用于每个量化单元的量化步长, 它还可以用熵代码来编码。如先前所解释的, 步长指数的个数、或量化单元的个数将如49中所示由解码器从量化指数中重建;
- 任意分辨率的滤波器组决策96: 一个比特用于每个量化单元, 它只出现在可切换分辨率的分析滤波器组28采取低频率分辨率模式时, 指示解码器是否要对量化单元内的所有子带段执行任意分辨率的滤波器组重建(51或55);
- 和/差编码决定97: 一个比特用于被和/差编码的量化单元中的一个。它是可选择的并且只出现在采用和/差编码时, 它指示解码器是否要执行和/差解码47;
- 联合强度编码决策和引导向量98: 它传送关于解码器是否要进行联合强度解码的信息, 它是可选择的并且只用于被联合强度编码的联合声道的量化单元, 并且只出现在编码器采用联合强度编码时, 它包括以下部分:
 - 决策121: 每个联合量化单元一个比特, 向解码器指示是否要

对量化单元中的子带样本进行联合声道解码；

○ **极性122**：每个联合量化单元一个比特，表示联合声道相对于源声道的极性：

$$\text{极性} = \begin{cases} 1, & \text{如果极性比特}=0 \\ -1, & \text{否则} \end{cases}$$

○ **引导向量123**：每个联合量化单元一个比例因子，它可以被熵编码；

- **辅助数据99**：包括诸如动态范围控制之类的辅助信息。

当三模式可切换的滤波器组被使用时，比特流结构本质上与上述相同，除了：

- **窗口类型90**：指示哪一个窗口诸如图5和图9中所示窗口用于编码器以便于解码器能够使用相同的窗口。注意，对于具有暂态的帧来说，这个窗口类型只涉及帧中的最后一个窗口，因为其余的窗口能够从这个窗口类型、暂态位置以及最后帧中使用的最后窗口来推断；

- **暂态位置91**：只出现在具有暂态的帧的情况下。它首先指示这个帧是否具有慢暂态171。如果不是，则它按照中块172并然后按照短块173来指示暂态位置；

- **任意分辨率的滤波器组决策96**：它是不相干的，因此未被使用。

解码器

本发明的解码器基本上实现了编码器的逆处理，它在图13中被示出

并被解释如下。

一个多路解复用器41从比特流中解码出量化指数，以及如量化步长、采样率、扬声器配置和时间码等之类的辅助信息，。当诸如哈夫曼码之类的前缀熵代码被使用时，这个步骤是一个结合了熵解码的单一步骤。

量化指数码书解码器42从比特流中解码量化指数及其各自的应用范围的熵码书。

熵解码器43基于由量化指数码书解码器42提供的熵码书及其各自的应用范围从比特流中解码量化指数。

去交错44只有在当前帧中存在暂态时才被选择性地采用。如果从比特流解包的决策比特指示交错18在编码器中被调用过，则去交错量化指数。否则，不做任何修改地传递量化指数。

本发明从每个暂态段49的非零量化指数中重建量化单元的个数。让 $q(m,n)$ 是对于第 m 个暂态段的第 n 个子带的量化指数(如果帧中不存在暂态，则只存在一个暂态段)，找出每个暂态段 m 的具有非零量化指数的最大子带：

$$Band_{\max}(m) = \max_n \{n \mid q(m,n) \neq 0\}$$

回想一下，量化单元是由在频率中的临界频带和时间上的暂态段定义的，因此每个暂态段的量化单元个数是能够容纳 $Band_{\max}(m)$ 的最小临界频带。让频带(Cb)是第Cb个临界频带的最大子带，每个暂态段 m 的量化单元数可表示如下：

$$N(m) = \min_{Cb} \{Cb \mid Band(Cb) \geq Band_{max}(m)\}$$

量化步长解包50从比特流中解包每个量化单元的量化步长。

逆量化45利用每个量化单元的各个量化步长从量化指数中重建子带样本。

如果比特流指示编码器中调用过联合强度编码15，则联合强度解码46从源声道复制子带样本并将其乘以极性和引导向量以重建联合声道的子带样本：

$$\text{联合声道} = \text{极性} \cdot \text{引导向量} \cdot \text{源声道}$$

如果比特流指示和/差编码14在编码器中被调用过，则和/差解码器47从和/差声道中重建左右声道。对应于在和/差编码14中所解释的和/差编码例子，左右声道能够被重建为：

$$\text{左声道} = \text{和声道} + \text{差声道}$$

$$\text{右声道} = \text{和声道} - \text{差声道}$$

本发明的解码器结合了一个可变分辨率的合成滤波器组48，其实质上是用来编码信号的分析滤波器组的逆装置。

如果三模式可切换分辨率的分析滤波器组被用于编码器，则其对应的合成滤波器组的操作被唯一地确定并要求相同的窗口序列用于合成处理。

如果半混合滤波器组或可切换滤波器组+ADPCM被用于编码器，则

解码过程被描述如下：

- 如果比特流指示当前帧是用可切换分辨率的分析滤波器组28以高频率分辨率模式来编码的，则可切换分辨率的合成滤波器组54因此进入高频率分辨率模式并且从子带样本中重建PCM样本(见图14和图15)。
- 如果比特流指示当前帧是用可切换分辨率的分析滤波器组28以低频率分辨率模式来编码的，则子带样本首先被送到任意分辨率的合成滤波器组51(图14)或逆ADPCM55(图15)，并且这取决于编码器中哪一个被使用了，然后完成它们各自的合成处理。然后，PCM样本由可切换分辨率的合成滤波器组以低频率分辨率模式53从这些合成的子带样本中重建。

合成滤波器组52、51和55分别是分析滤波器组28、26和29的逆装置。它们的结构与操作处理由分析滤波器组来唯一地确定。因此，无论在编码器中使用什么分析滤波器组，其对应的合成滤波器组必须被用于解码器。

低编码延迟模式

当可切换分辨率的分析滤波器组的高频率分辨率模式被编码器禁止时，帧长可以随后被减小到低频模式下的可切换分辨率的滤波器组的块长或其整数倍，这产生了一个小得多的帧长，导致编码器和解码器操作所需的小得多的延迟。这就是本发明的低编码延迟模式。

尽管若干实施例已经为了举例的目的被详细描述，然而在不脱离本发明的范围和精神的的前提下可以做出不同的修改。因此，本发明只被附加的权利要求所限制。

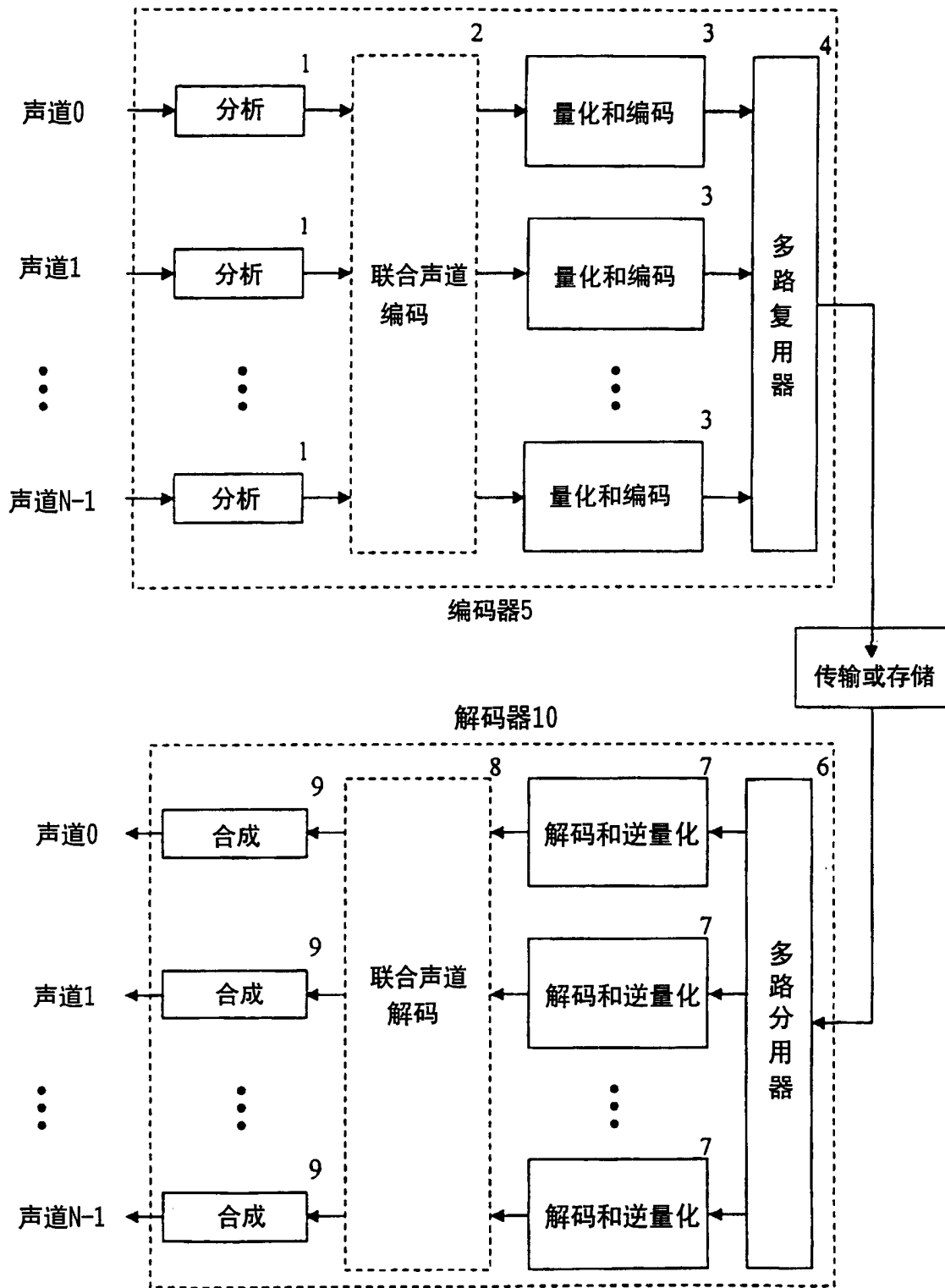


图1

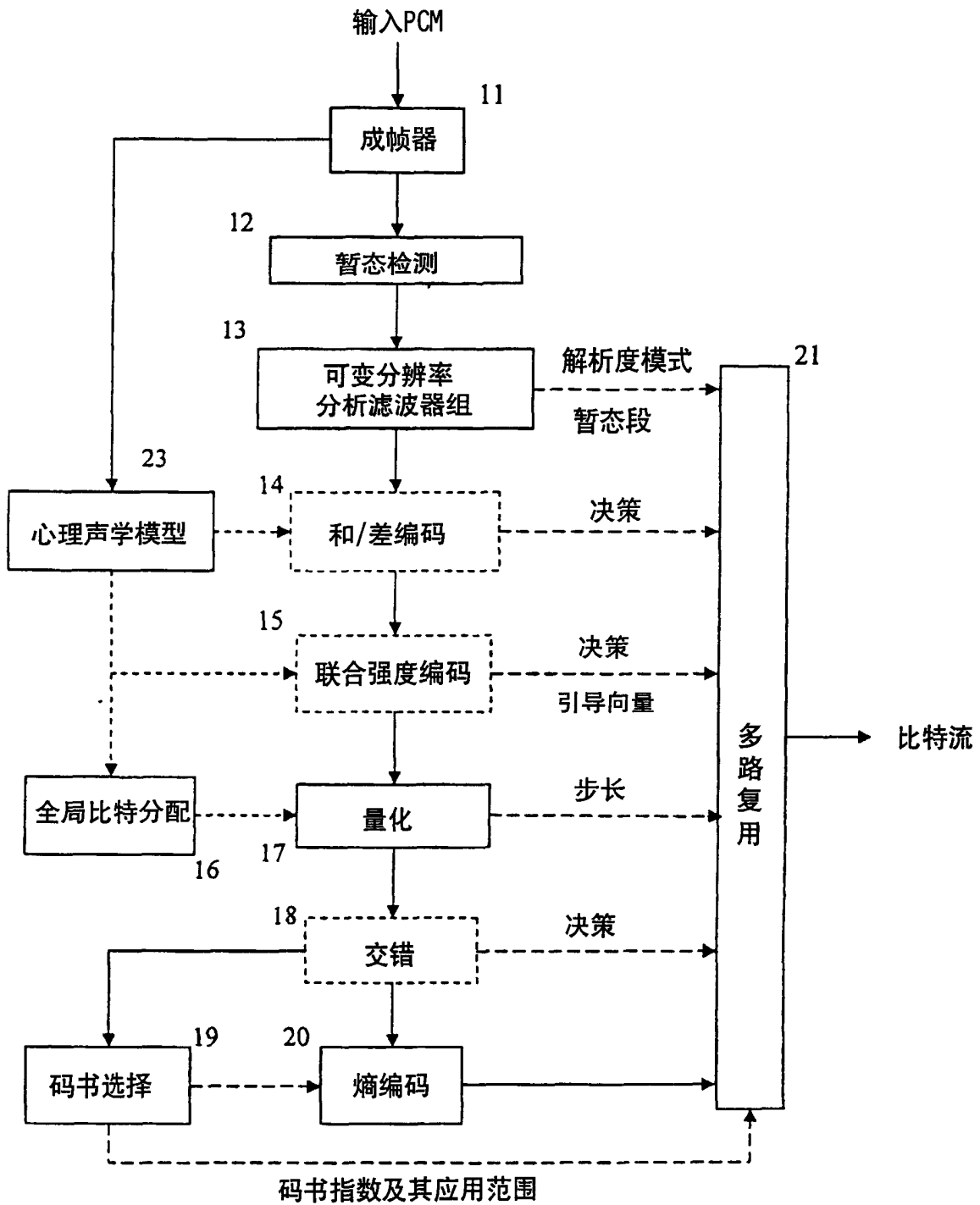


图2

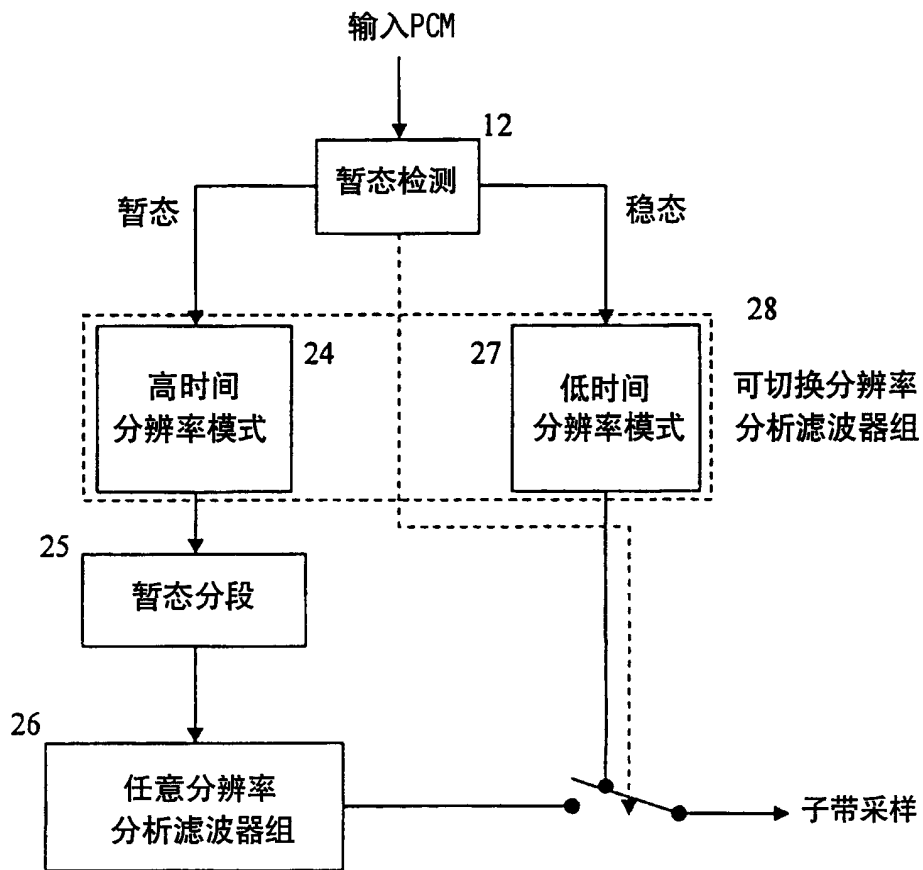


图3

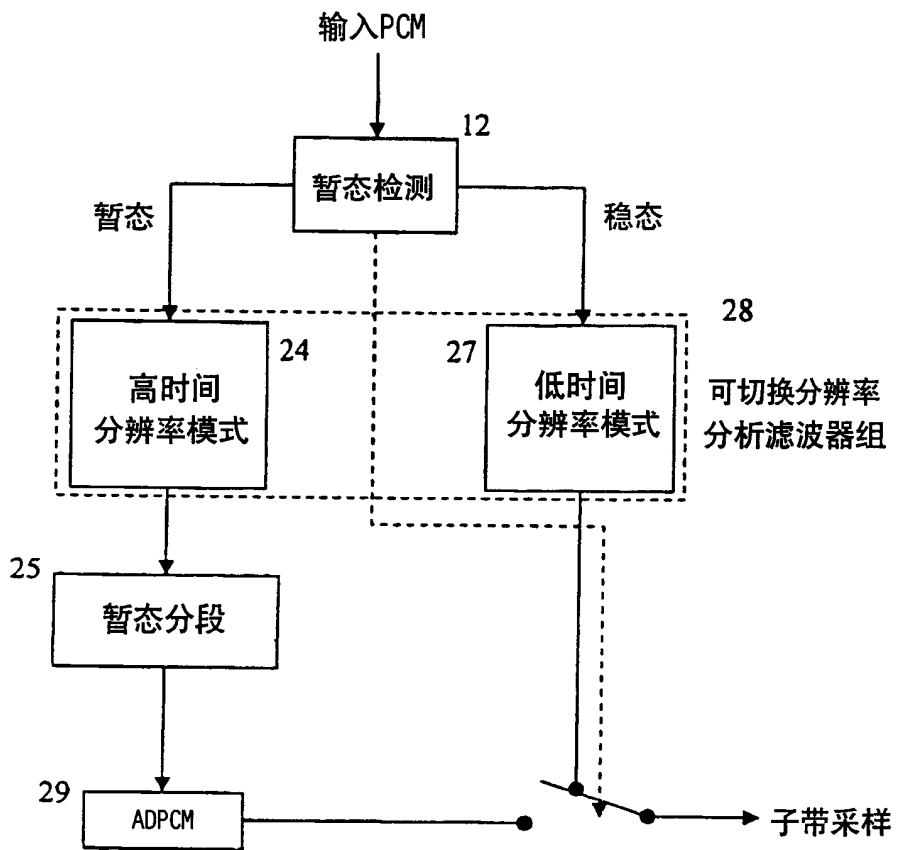


图4

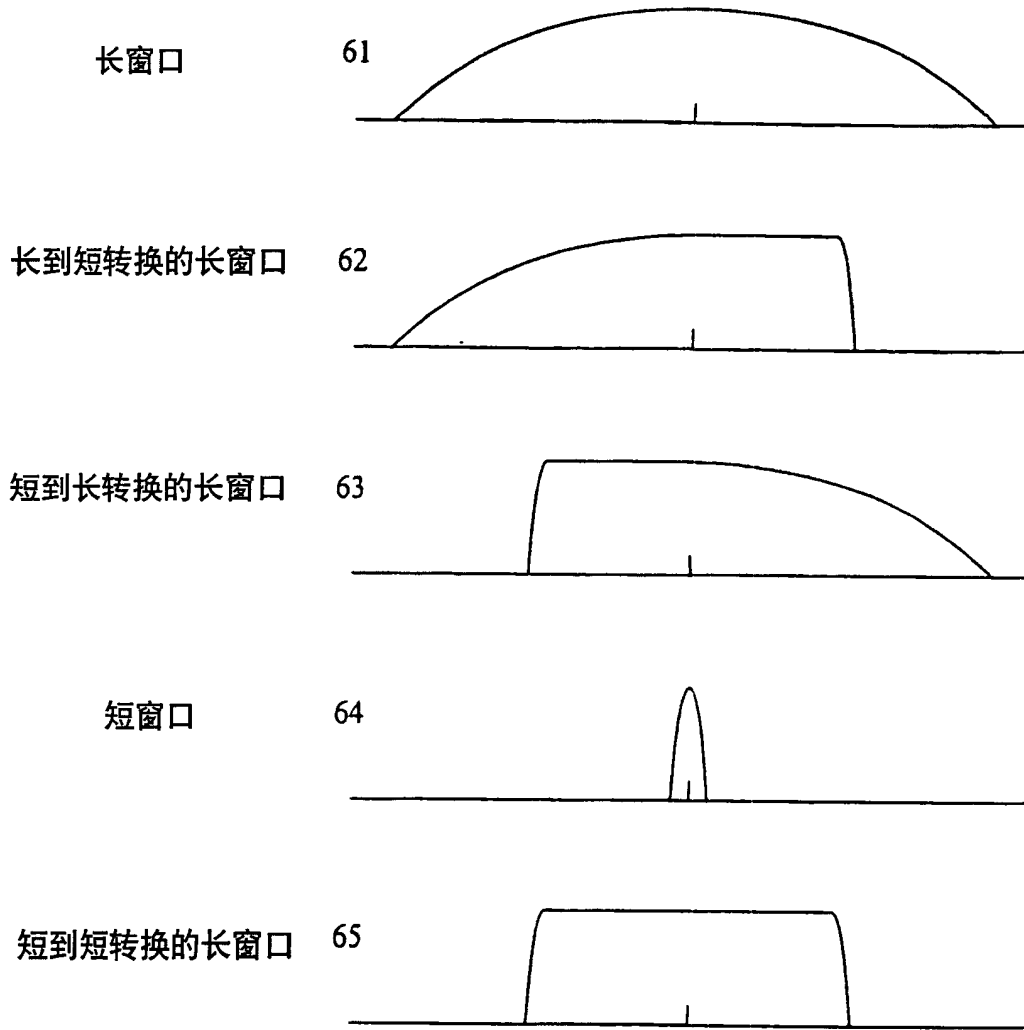


图5

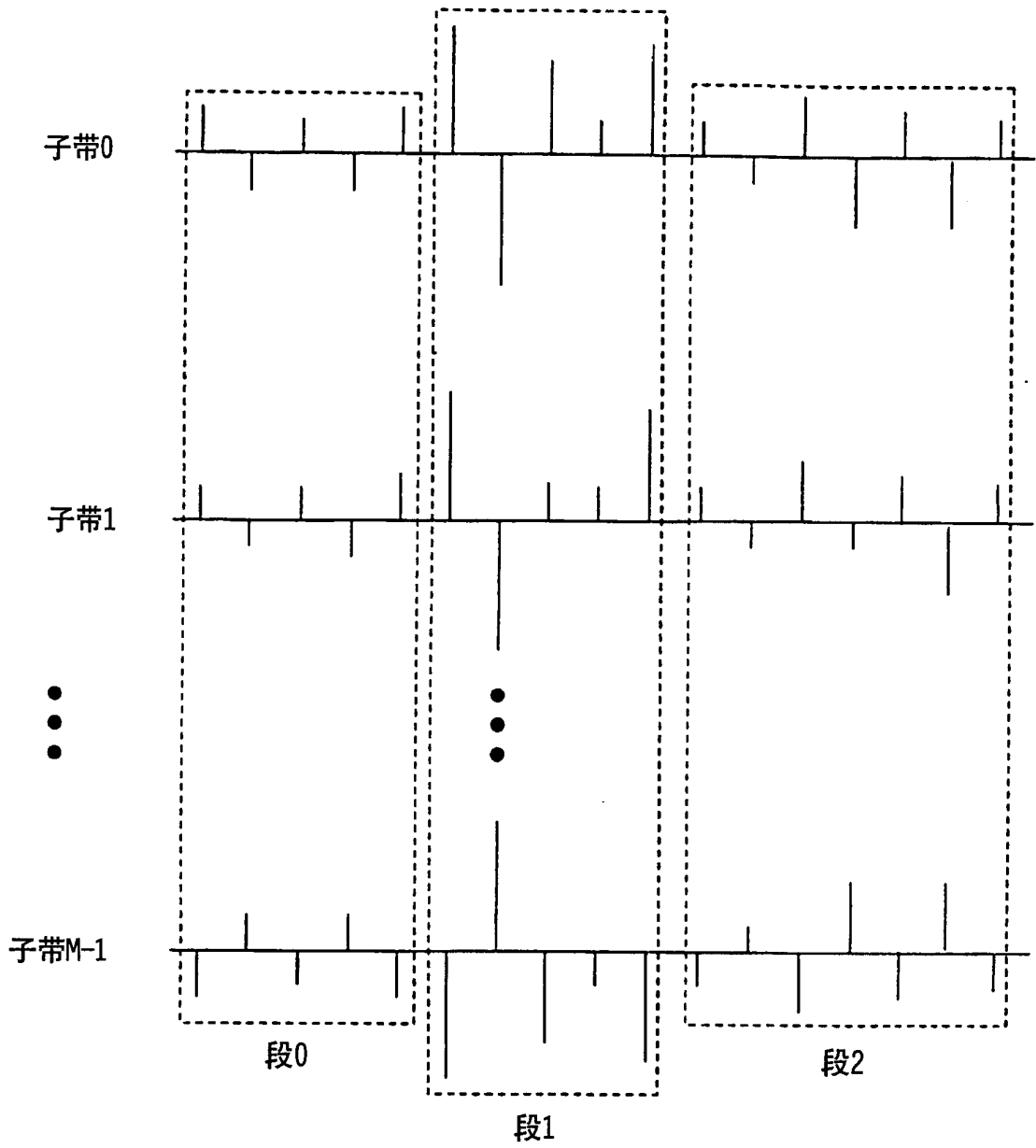


图6

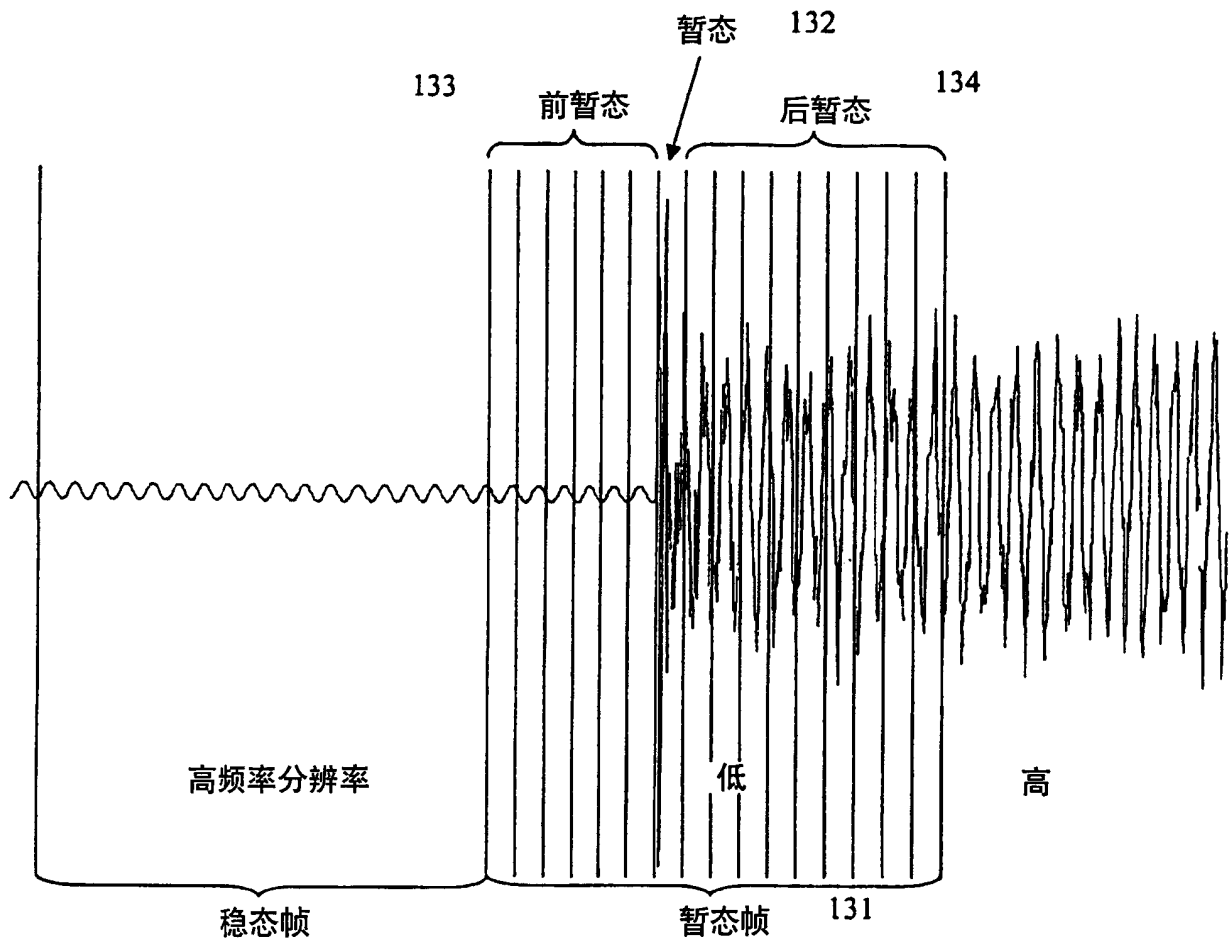


图7

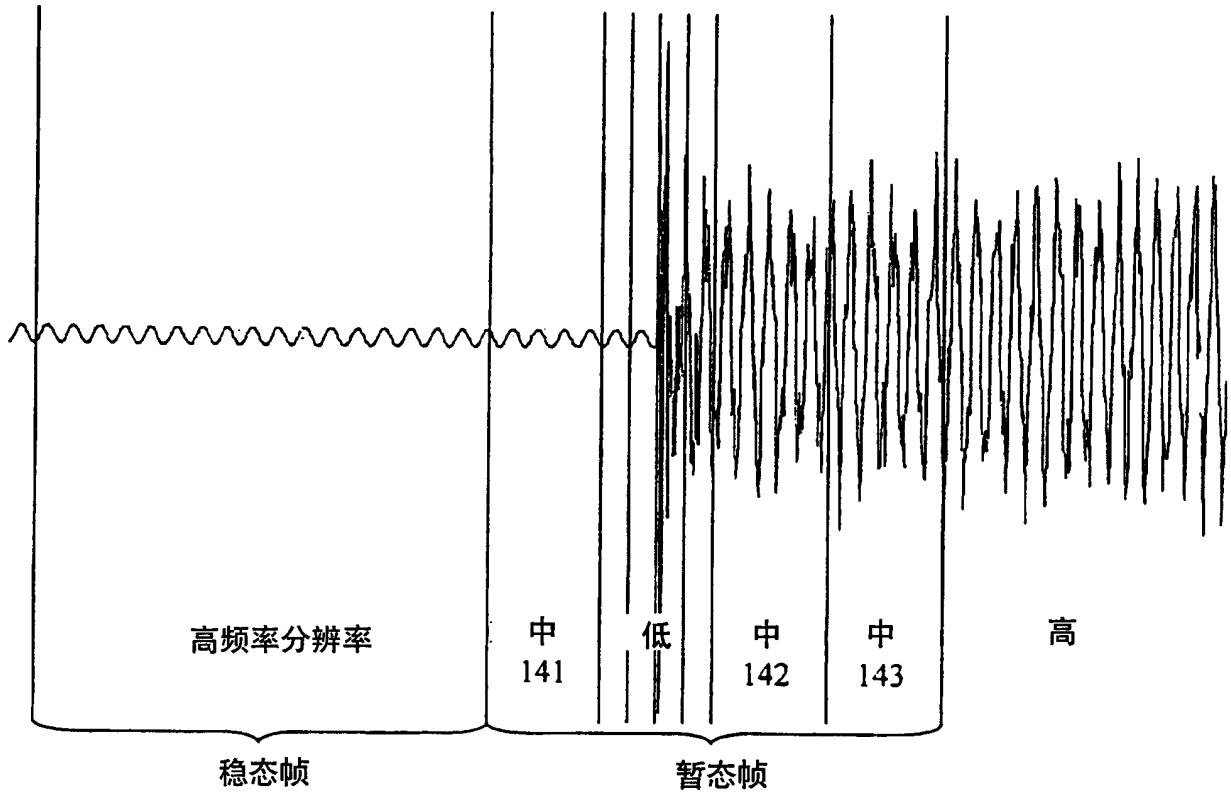


图8

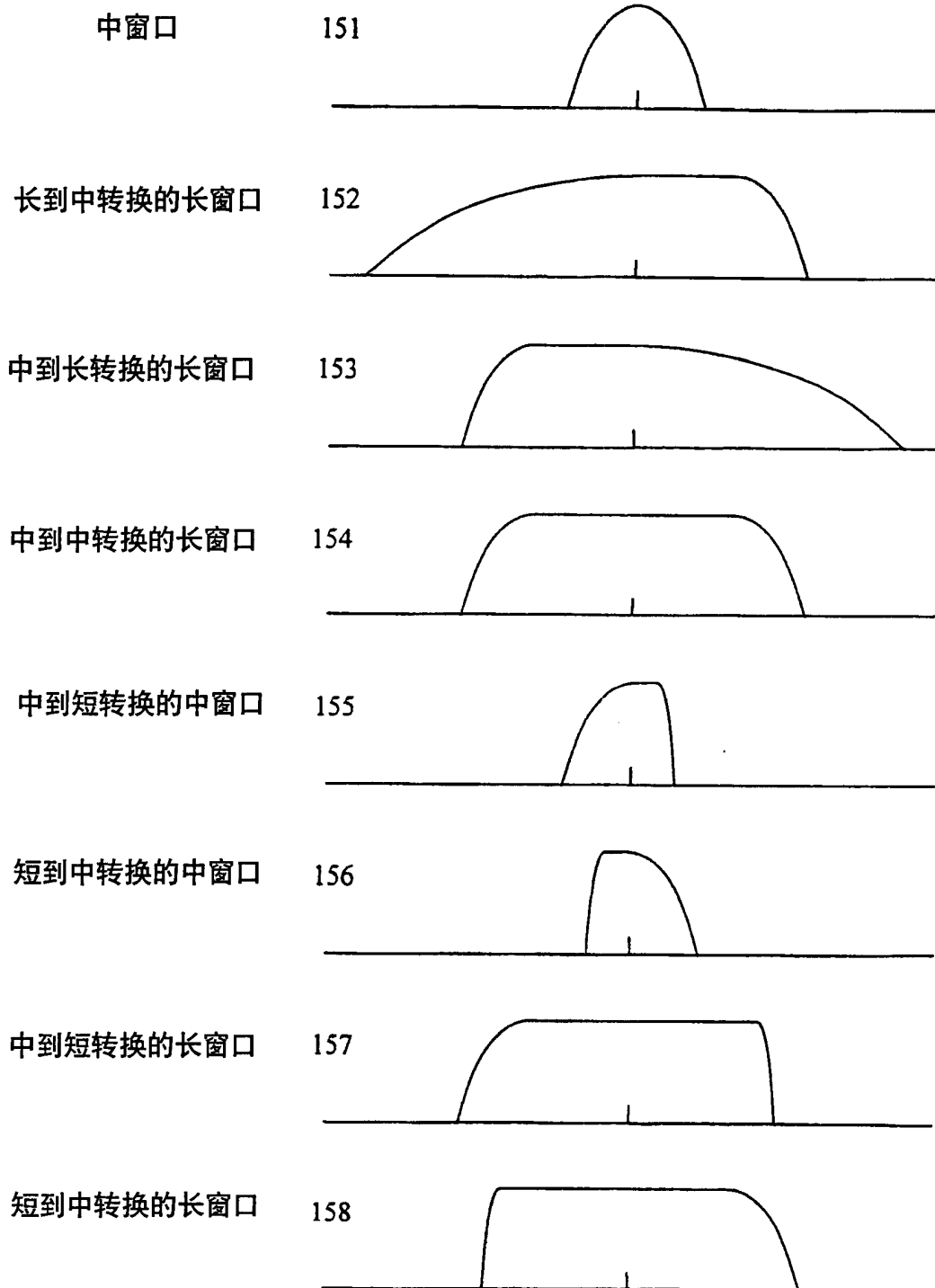


图9

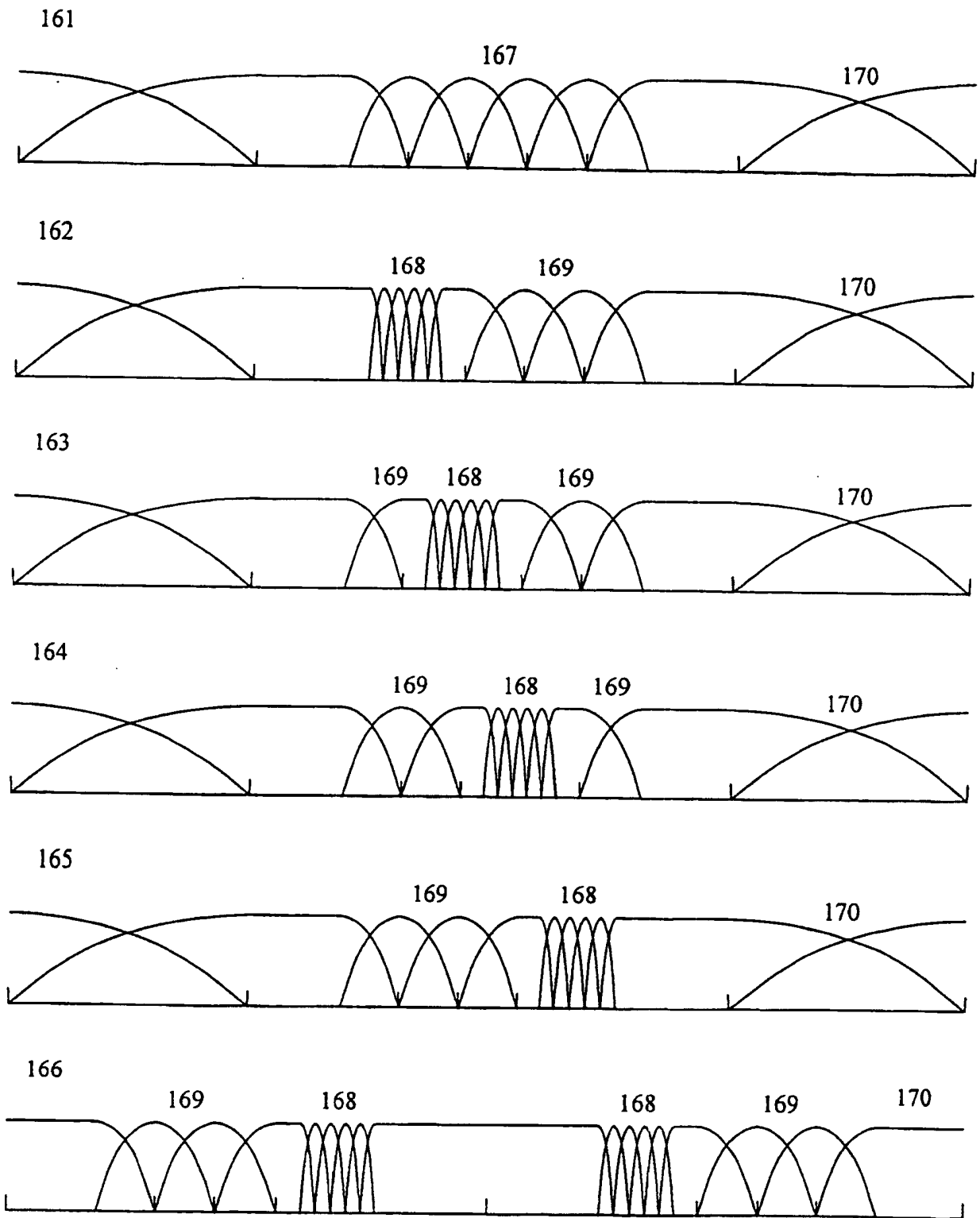


图10

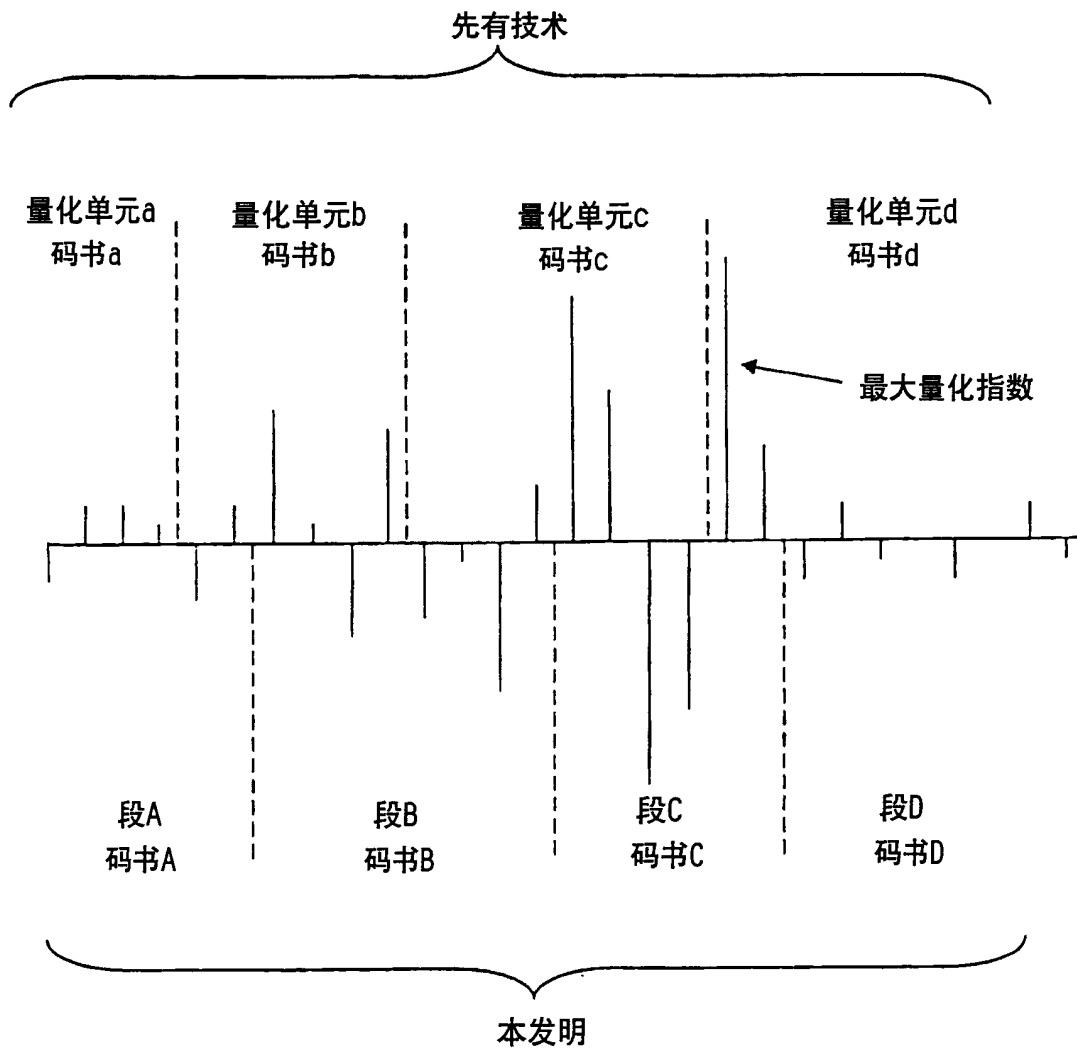


图11

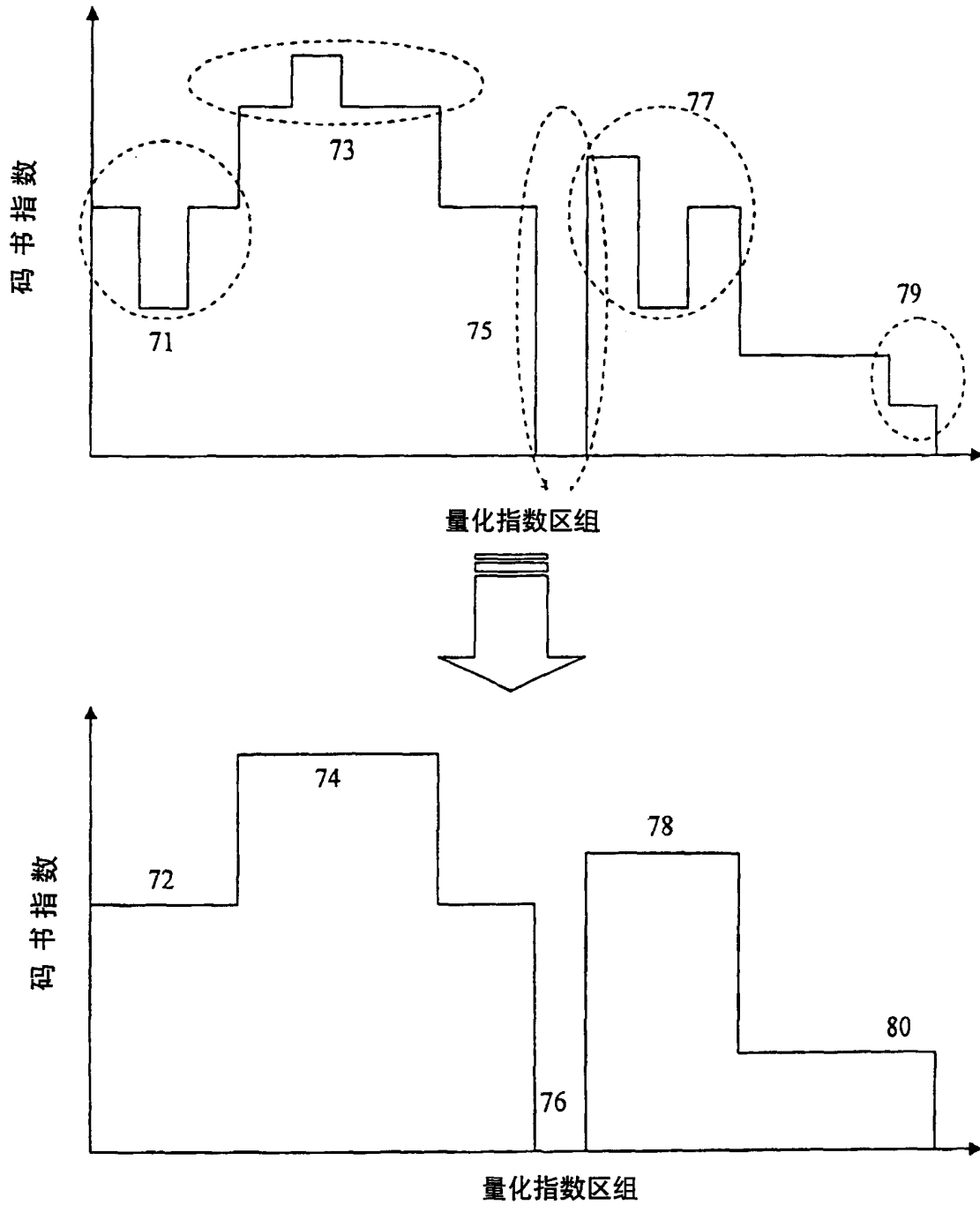


图12

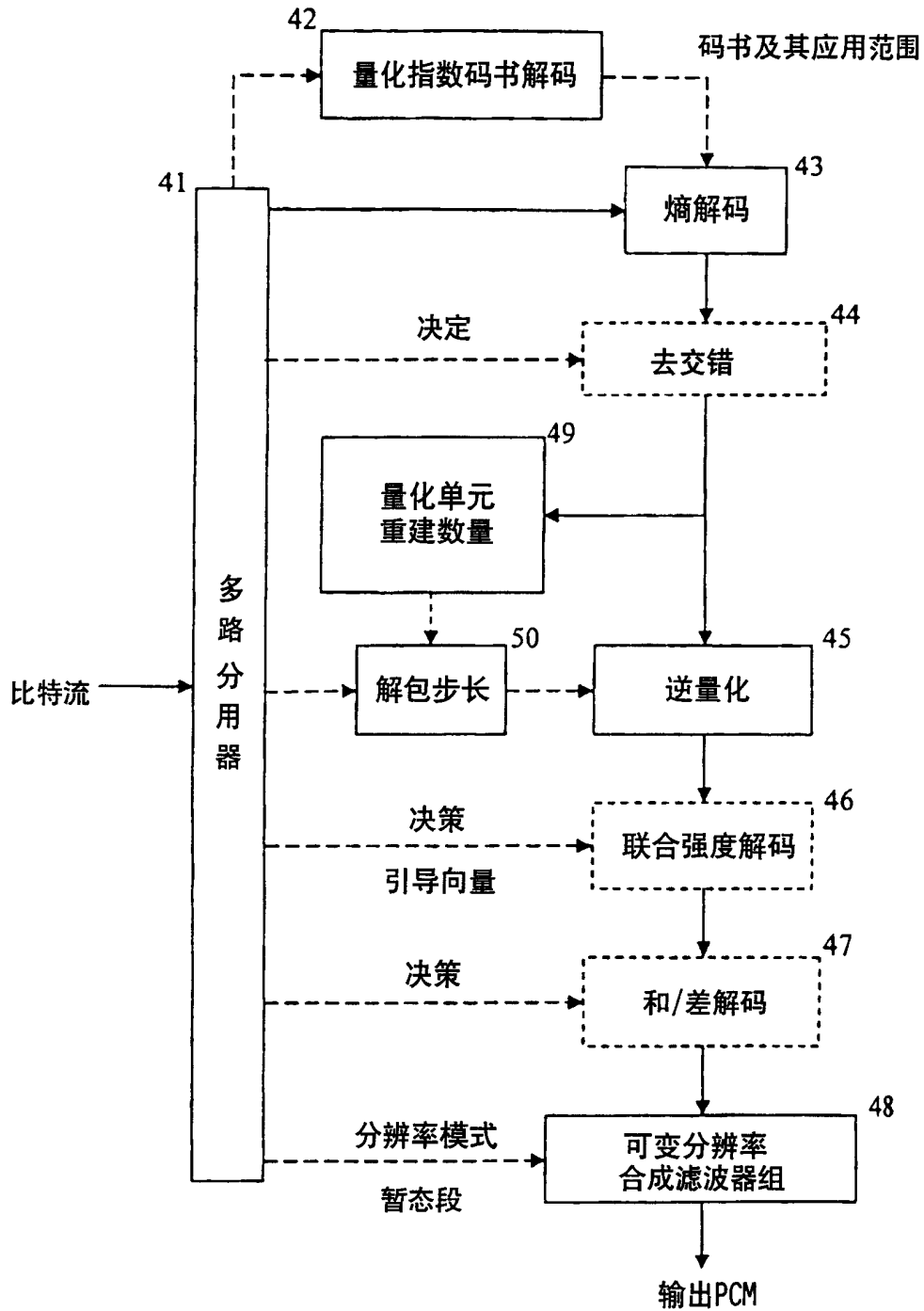


图13

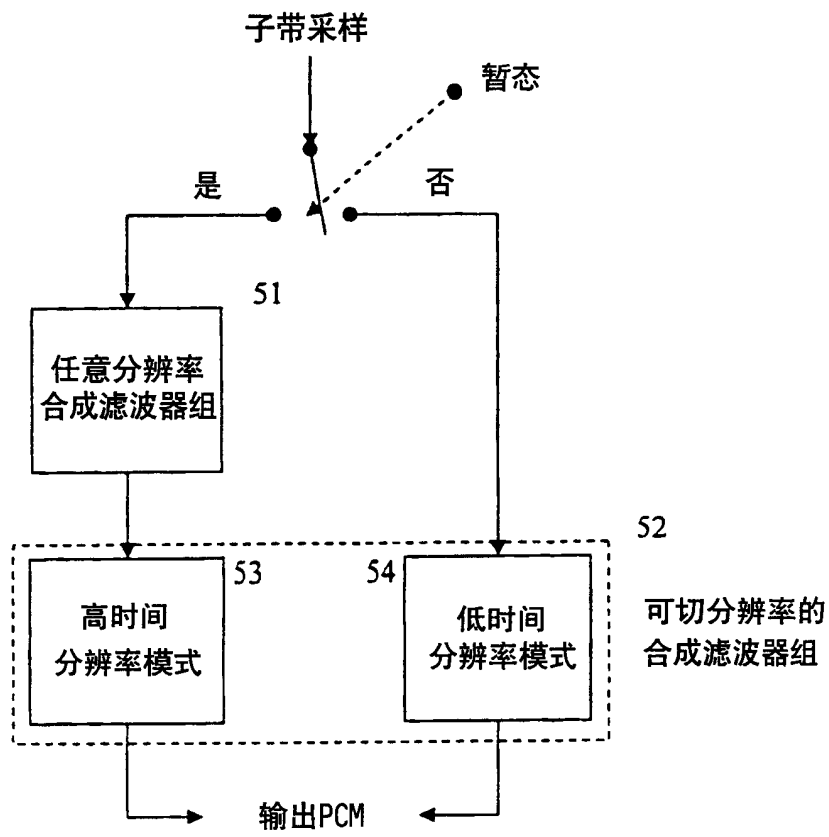


图14

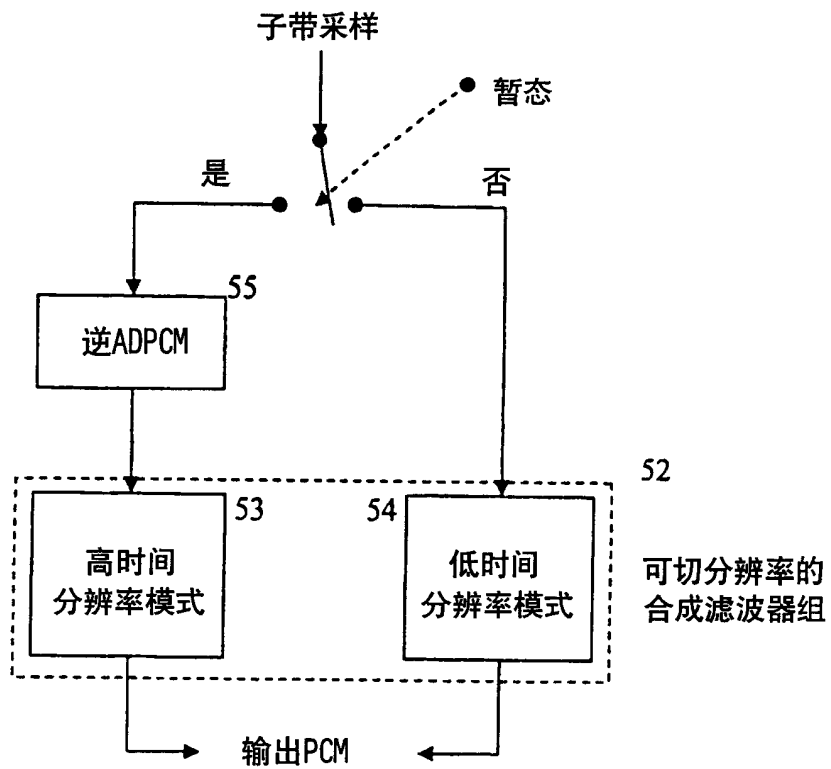


图15

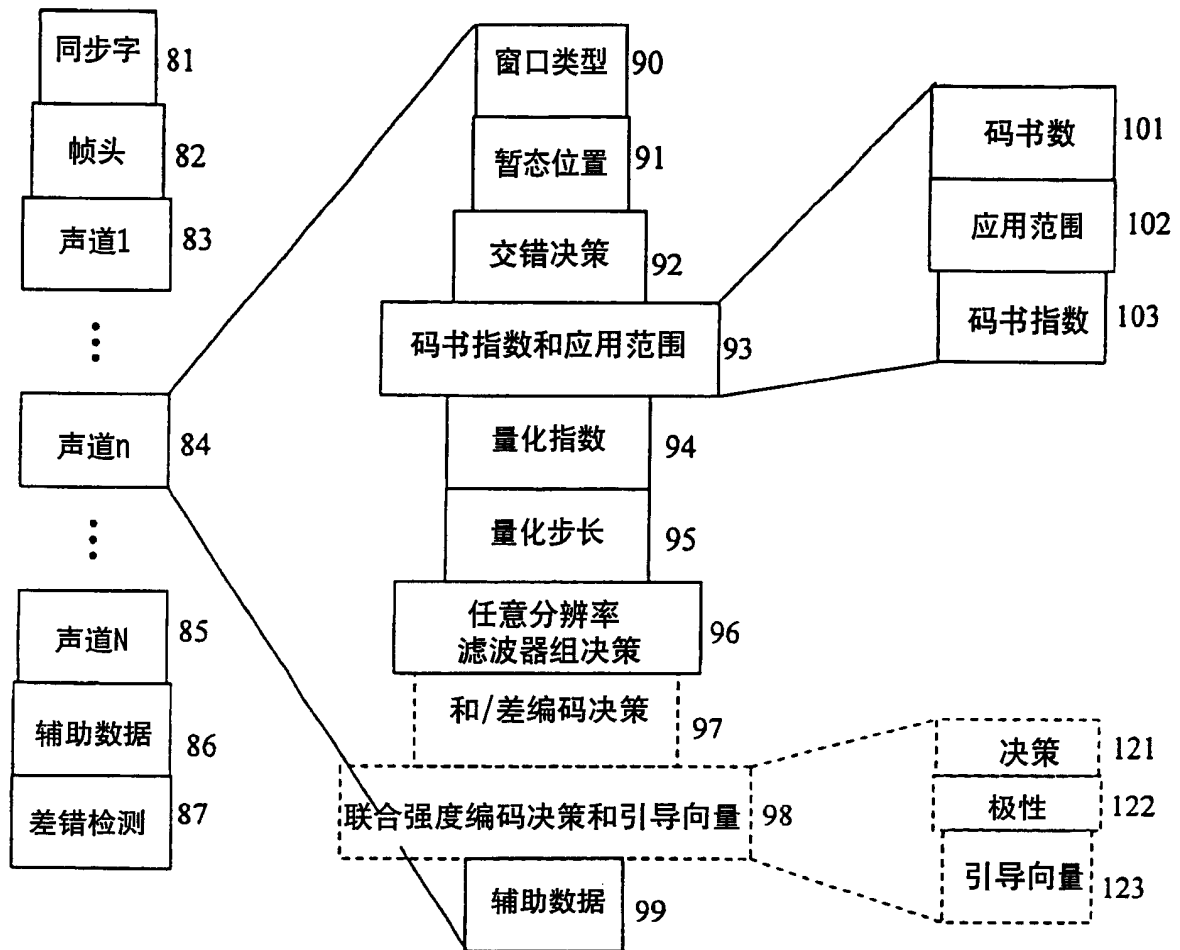


图16

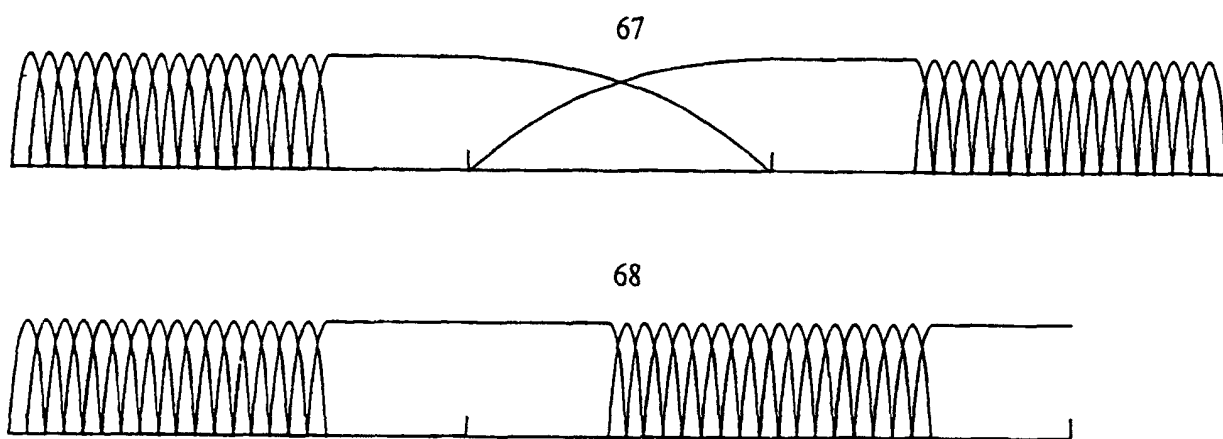


图17

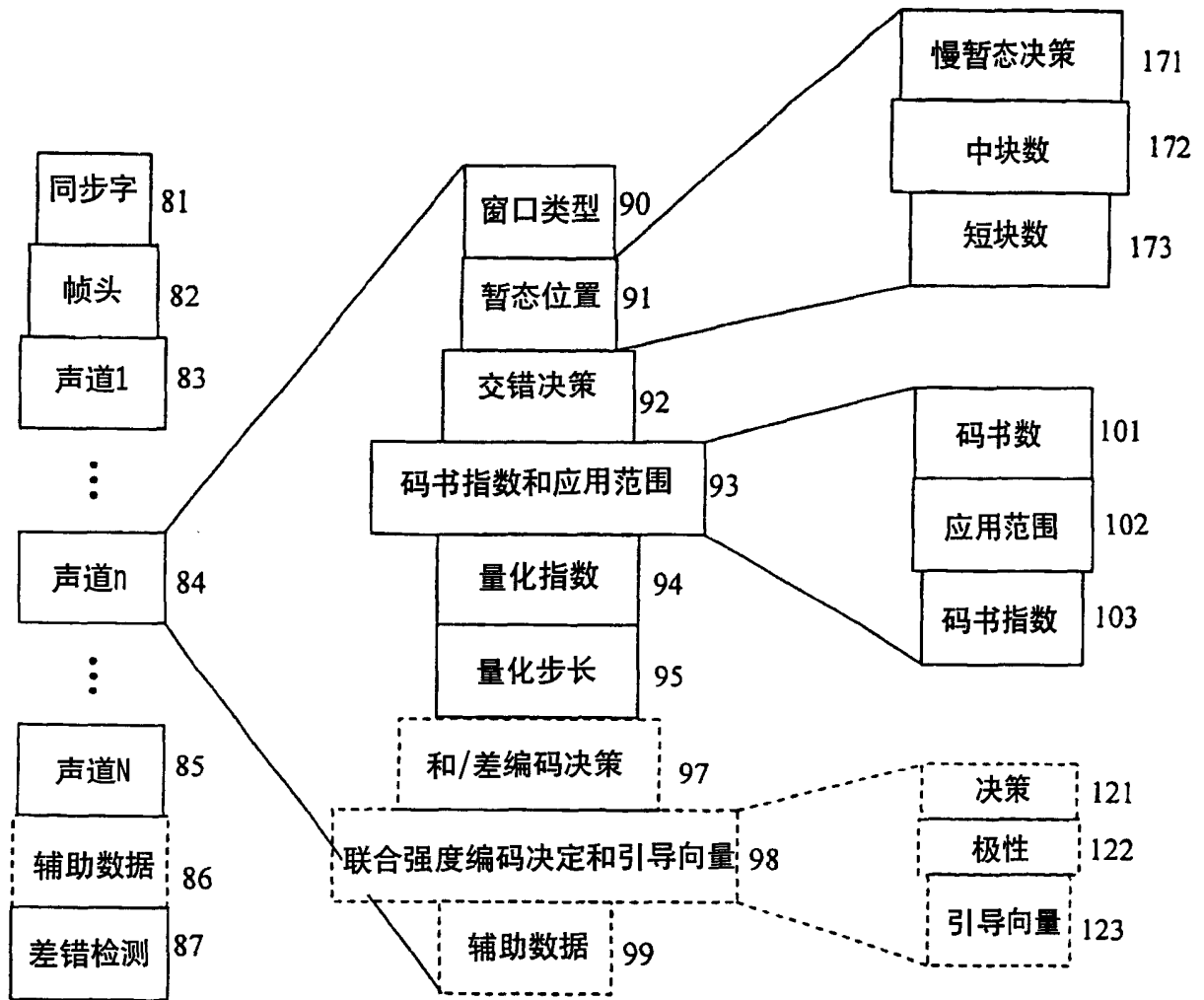


图18