



(12) 发明专利

(10) 授权公告号 CN 101853660 B

(45) 授权公告日 2013. 07. 03

(21) 申请号 201010138455. 1

(22) 申请日 2005. 09. 12

(30) 优先权数据

60/620, 401 2004. 10. 20 US

11/006, 492 2004. 12. 07 US

(62) 分案原申请数据

200580035950. 7 2005. 09. 12

(73) 专利权人 弗劳恩霍夫应用研究促进协会

地址 德国慕尼黑

专利权人 艾格瑞系统有限公司

(72) 发明人 埃里克·阿拉曼奇 萨沙·迪施

克里斯托夫·法勒 于尔根·赫勒

(74) 专利代理机构 北京三幸商标专利事务所

11216

代理人 刘激扬

(51) Int. Cl.

G10L 19/008 (2013. 01)

H04S 3/02 (2006. 01)

(56) 对比文件

CN 1460992 A, 2003. 12. 10, 全文.

CN 1536559 A, 2004. 10. 13, 全文.

WO 2004/008806 A1, 2004. 01. 22, 全文.

WO 2004/086817 A2, 2004. 10. 07, 全文.

CN 1495754 A, 2004. 05. 12, 全文.

审查员 王馨宁

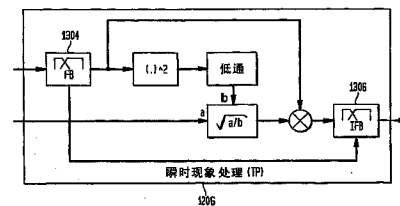
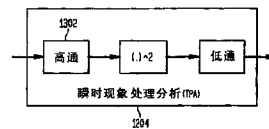
权利要求书1页 说明书16页 附图15页

(54) 发明名称

用于双声道提示码编码方案和类似方案的散射声音整形

(57) 摘要

将具有输入时序包络的输入音频信号转换成具有输出时序包络的输出音频信号。所述输入音频信号的输入时序包络被特性化。所述输入音频信号被处理以产生处理后音频信号,其中所述的处理将所述输入音频信号进行去关联。所述处理后音频信号基于特性化的输入时序包络被调整以产生所述输出音频信号,其中所述输出时序包络大体上与所述输入时序包络相匹配。



1. 一种用于对 C 个输入音频通道进行编码以产生 E 个传输音频通道的方法,该方法包括:

为所述 C 个输入通道中的两个或多个产生一个或多个提示码;

将所述 C 个输入通道下混以产生所述 E 个传输通道,其中 $C > E \geq 1$;和

对所述 C 个输入通道中和所述 E 个传输通道中的一个或多个进行分析以产生一个标记,所述标记在所述 E 个传输通道解码期间用来指示所述 E 个传输通道的解码器是否执行包络整形,所述分析步骤包括:

以展望的方式实施瞬时现象检测,其中,当标记被设置,在所述解码器中,不仅对瞬时现象,而且对在所述瞬时现象之前和之后的信号进行整形,当检测到瞬时现象时设置标记,时序处理被启动,或

用于检测是否时序包络是以假的随机方式的波动的随机检测,当时序包络是以假的随机方式的波动时,设置所述标记,其中,虽然没有瞬时现象被检测到,所述时序处理仍然被启动,或

用于当 E 个传输通道是高音调时不设置所述标记的音调检测。

2. 如权利要求 1 所述的方法,其中所述包络整形对由所述解码器产生的解码后通道的时序包络进行调整以便与相应的传输通道的时序包络匹配。

3. 一种用于对 C 个输入音频通道进行编码以产生 E 个传输音频通道的设备,该设备包括:

用于为所述 C 个输入通道中的两个或多个产生一个或多个提示码的装置;

用于将所述 C 个输入通道进行下混以产生所述 E 个传输通道的装置,其中 $C > E \geq 1$;和

用于对所述 C 个输入通道中的一个或多个和所述 E 个传输通道中的一个或多个进行分析以产生一个标记的装置,所述标记在所述 E 个传输通道解码期间用来指示所述 E 个传输通道的解码器是否执行包络整形,所述用于分析的装置被设置成:

以展望的方式实施瞬时现象检测,其中,当标记被设置,在所述解码器中,不仅对瞬时现象,而且对在所述瞬时现象之前和之后的信号进行整形,当检测到瞬时现象时设置标记,时序处理被启动,或

用于检测是否时序包络是以假的随机方式的波动的随机检测,当时序包络是以假的随机方式的波动时,设置所述标记,其中虽然在没有瞬时现象被检测到,所述时序处理仍然被启动,或

用于当 E 个传输通道是高音调时不设置所述标记的音调检测。

用于双声道提示码编码方案和类似方案的散射声音整形

[0001] 本发明是 2005 年 9 月 12 日申请的发明名称为“用于双声道提示码编码方案和类似方案的散射声音整形”的第 200580035950.7 号发明专利申请的分案申请。

[0002] 发明背景

[0003] 相关申请的参考引用

[0004] 本申请要求于 2004 年 10 月 20 日在美国提交的第 60/620,401 号临时申请的利益,其代理人号为 Allamanche 1-2-17-3,其启示在此引入作为参考。

[0005] 另外,本申请的主题涉及下面美国申请的主题,这里将其引入作为参考:

[0006] 美国申请号为 09/848,877,申请日为 2001 年 5 月 4 日,代理人号为 Faller 5;

[0007] 美国申请号为 10/045,458,申请日为 2001 年 11 月 7 日,代理人号为 Baumgarte 1-6-8,该美国申请要求了于 2001 年 8 月 10 日提交的第 60/311,565 号美国临时申请的利益;

[0008] 美国申请号为 10/155,437,申请日为 2002 年 5 月 24 日,代理人号为 Baumgarte 2-10;

[0009] 美国申请号为 10/246,570,申请日为 2002 年 9 月 18 日,代理人号为 Baumgarte 3-11;

[0010] 美国申请号为 10/815,591,申请日为 2004 年 4 月 1 日,代理人号为 Baumgarte 7-12;

[0011] 美国申请号为 10/936,464,申请日为 2004 年 9 月 8 日,代理人号为 Baumgarte 8-7-15;

[0012] 美国申请号为 10/762,100,申请日为 2004 年 1 月 20 日,(Faller13-1);和

[0013] 美国申请号为 10/xxx,xxx,相同的申请日,代理人号为 Allamanche 2-3-18-4;

[0014] 本申请的主题还涉及下面论文的主题,在此将其引入作为参考:

[0015] F. Baumgarte 和 C. Faller,“Binaural Cue Coding-Part I: Psychoacoustic fundamentals and design principles”, IEEE Trans. On Speech and Audio Proc., 卷 11, 第 6 期, 2003 年 11 月;

[0016] C. Faller 和 F. Baumgarte,“Binaural Cue Coding-Part II: Schemes and applications”, IEEE Trans. on Speech and Audio Proc., 卷 11, 第 6 期, 2003 年 11 月;和

[0017] C. Faller,“Coding of spatial audio compatible with different playback formats”, Preprint 117th Conv. Aud. Eng. Soc., 2004 年 10 月。

技术领域

[0018] 本发明涉及所述音频信号的编码和随后的从编码后的音频数据合成听觉场景。

背景技术

[0019] 当人听到由特定音源产生的音频信号(即,声音)时,所述音频信号通常会在两

个不同的时间抵达人的左耳与右耳且具有两个不同的音频音量大小（例如，分贝），这些不同的时间和音量大小是路径中差异的函数，通过所述路径音频信号分别传播抵达左耳与右耳，人的大脑解读这些时间和音量大小的差异从而使人感觉到所接收的音频信号是由位于相对于所述人的特定位置（例如，方向与距离）的音源所产生。听觉场景为一人同时听到的由位于相对于所述人的一个或多个不同位置的一个或多个不同音源所产生的音频合成串音。

[0020] 通过大脑此处理的存在可被用来合成听觉场景，其中来自一个或多个不同音源的音频信号可以目的地修改以产生左边与右边音频信号，所述左边和右边音频信号使听者感觉到不同音源相对于所述听者位于不同的位置。

[0021] 图 1 表示传统的立体声信号合成器 100 的高级框图，其将单一音源信号（例如，单声道信号）转换成立体声信号的左边与右边音频信号，其中立体声信号被定义为在听者的鼓膜处所接收的两个信号。除所述音源信号外，合成器 100 接收对应于相对听者的音源的期望位置的一组空间提示信号。在典型的实施中，所述这组空间提示信号包括通道间电平差 (ICLD) 值（其辨识分别在左耳与右耳所接收的左与右音频信号间音频音量大小的差异），和音频通道中时差异 (ICTD) 值（其辨识如分别在左耳与右耳所接收的左边与右边音频信号间抵达时间的差异）。此外或作为替换，一些合成技术包括用于从音源到耳膜的声音的方向依赖转移函数的建模，也可引用头部相关的转移函数 (HRTF)，参见例如，J. Blauert, *The Psychophysics of Human Sound Localization*, MIT Press, 1983, 其在此引入以供参考。

[0022] 使用图 1 的立体声信号合成器 100，由单一音源所产生的单声道音频信号可被处理以便当通过耳机收听时，所述音源通过使用适当空间提示信号组（例如，ICLD、ICTD 和 / 或 HRTF）来为每一个耳朵产生音频信号，参见例如，D. R. Begault, *3-D Sound for Virtual Reality and Multimedia*, Academic Press, Cambridge, MA, 1994。

[0023] 图 1 的立体声信号合成器 100 产生最简单型式的听觉场景，它们相对于听者具有单一音源，包括相对于听者的位于不同位置的两个或多个音源的更复杂的听觉场景可使用听觉场景合成器被产生，所述听觉场景合成器通过使用多个立体声信号合成器而本质上被实施，其中每个立体声信号合成器产生对应于不同音源的立体声信号，因为每个不同音源相对于听者具有不同的位置，不同空间提示信号组被用来对每个不同音源产生立体声音频信号。

发明内容

[0024] 根据一个实施例，本发明涉及用于将具有输入时序包络的输入音频信号转换成具有输出时序包络的输出音频信号的方法和装置。所述输入音频信号的所述输入时序包络被特性化。对所述输入音频信号进行处理以产生处理后音频信号，其中所述处理将所述输入音频信号去关联。基于所述特性化的输入时序包络对所属处理后音频信号进行处理以产生所述输出音频信号，其中所述输出时序包络大体上与所述输入时序包络相匹配。

[0025] 依据本发明的另一实施例，本发明涉及对 C 个输入音频通道编码以产生 E 个传输音频通道的方法和装置。为所述 C 个输入通道中的两个或多个而产生一个或多个提示码。对所述 C 个输入通道进行下混以产生所述 E 个传输通道，其中 $C > E \geq 1$ 。所述 C 个输入通

道中的一个或多个和所述 E 个传输通道被分析,以产生一个在所述 E 个传输通道解码期间用来指示所述 E 个传输通道的解码器是否执行包络整形的标记。

[0026] 根据另外一个实施例,本发明涉及通过前面段落中提到的方法产生的编码后音频比特流。

[0027] 根据另外一个实施例,本发明涉及包括 E 个传输通道、一个或多个提示码和标记的编码后音频比特流。通过为所述 C 个输入通道中的两个或多个而产生一个或多个提示码从而产生一个或多个提示码。通过对所述 C 个输入通道进行下混产生所述 E 个传输通道,其中 $C > E \geq 1$ 。通过对所述 C 个输入通道中的一个或多个进行分析产生所述标记,其中在所述 E 个传输通道解码期间用来指示所述 E 个传输通道的解码器是否执行包络整形。

附图说明

[0028] 本发明的其他的方面、特征和优点从下面的详细描述、所附的权利要求和附图中将会更明显,其中相同的附图标记表示相似或相同的元件。

[0029] 图 1 为传统的立体声信号合成器的高级框图;

[0030] 图 2 为一般双声道提示码编码 (BCC) 音频处理系统的框图;

[0031] 图 3 为可被使用于图 2 的下混器的框图;

[0032] 图 4 为可被使用于图 2 的 BCC 合成器的框图;

[0033] 图 5 依据本发明的实施例,显示图 2 中所述 BCC 评估器的框图;

[0034] 图 6 表示用于五音频通道的 ICTD 和 ICLD 数据的生成;

[0035] 图 7 表示用于五音频通道的 ICC 数据的生成;

[0036] 图 8 表示图 4 的所述 BCC 合成器的实施的框图,其在单一传输总和信号 $s(n)$ 加空间提示信号下可被使用于 BCC 解码器中以产生立体声或多通道音频信号;

[0037] 图 9 表示 ICTD 与 ICLD 作为频率函数如何在基频带中被改变;

[0038] 图 10 为根据本发明的一个实施例的表示 BCC 解码器的至少一部分的框图;

[0039] 图 11 表示在图 4 的 BCC 合成器的范围内的图 10 的包络整形方案的示范应用;

[0040] 图 12 表示图 4 的 BCC 合成器的范围内的图 10 的包络整形方案的替换示范应用,其中包络整形被应用到时域中;

[0041] 图 13(a) 和 (b) 表示图 12 中的 TPA 和 TP 的可能的实施,其中只有当频率高于截止频率 f_{TP} 时包络整形才可以实施;

[0042] 图 14 表示在 2004 年 4 月 1 日申请的美国申请号为 10/815,591,代理人号为 Baumgarte7-12 的申请中描述的基于后期回响的 ICC 合成方案范围内的图 10 中的包络整形方案的示范应用;

[0043] 图 15 表示根据可以替换成图 10 所示方案的本发明的实施例的 BCC 解码器的至少一部分的框图;

[0044] 图 16 表示根据可以替换成图 10 和图 15 所示方案的本发明的实施例的 BCC 解码器的至少一部分的框图;

[0045] 图 17 表示在图 4 中的 BCC 合成器的范围内的图 15 的包络整形方案的示范应用;

[0046] 图 18(a)-(c) 表示图 17 中的 TPA、ITP 和 TP 的可能实施的框图。

具体实施方式

[0047] 在双声道提示码编码 (BCC) 中, 编码器对 C 个输入音频通道编码以产生 E 个传输音频通道, 其中 $C > E \geq 1$ 。特别是 C 个输入通道中的两个或多个被提供于频域中, 且一个或多个提示码被产生用于频域中两个或多个输入通道中一个或多个不同频带的每一个。此外, 所述 C 个输入通道被下混以产生 E 个传输通道, 在一些下混实施中, 所述 E 个传输通道中的至少一个基于所述 C 个输入通道中的两个或多个, 且至少所述 E 个传输通道中的一个仅基于 C 个输入通道中的单一通道。

[0048] 在一个实施例中, BCC 码器具有两个或多个滤波器库、一个代码评估器与一个下混器, 所述两个或多个滤波器库将所述 C 个输入通道中的两个或多个从时域转换到频域, 所述代码评估器产生一个或多个提示码用于所述两个或多个经转换输入通道中一个或多个不同频带的每一个, 下混器下混 C 输入通道以产生 E 个传输通道, 其中 $C > E \geq 1$ 。

[0049] 在 BCC 解码中, E 个传输音频通道被解码以产生 C 回放音频通道。特别对于一个或多个频带中的每一个, 一个或多个 E 个传输通道在频域中被上混以在频域中产生 C 回放通道中的两个或多个, 其中 $C > E \geq 1$ 。一个或多个提示码被施加至频域中所述两个或多个回放音频通道中的所述一个或多个不同频段的每一个以产生两个或多个经修改的通道, 且所述两个或多个经修改的音道从频域被转换成时域。在一些上混实施中, 至少 C 回放通道中的一个基于 E 个传输音频通道中的至少一个和至少一个提示码, 且 C 回放通道中的至少一个仅基于 E 个传输音频通道中的单一一个且与任何提示码无关。

[0050] 在一个实施例中, BCC 解码器具有上混器、合成器和一个或多个反向滤波器库, 对于一个或多个不同频带中的每一个, 所述上混器在频域中上混 E 个传输通道中的一个或多个以便在频域中产生 C 个回放通道中的两个或多个, 其中 $C > E \geq 1$, 所述合成器施加一个或多个提示码至频域中所述两个或多个回放通道中的所述一个或多个不同频段的每一个, 以便产生两个或多个经修改的通道, 所述一个或多个反向滤波器库将所述两个或多个修改的通道从频域转换成时域。

[0051] 根据特别实施, 指定的回放通道可基于一个单一传输通道, 而不是两个或多个传输通道的结合。例如, 当仅有一个传输通道, C 个回放通道中的每一个基于所述传输通道。在这些情况下, 上混对应所述相应的传输通道的复制。如此, 对仅有一个传输通道的应用, 所述上混器可使用为每一个回放通道复制传输通道的复制器而被实施。

[0052] BCC 编码器和 / 或解码器可合并成一些系统或应用, 其包含, 例如数字录像机 / 放影机、数字录音机 / 放音机、计算机、卫星发送器 / 接收器、有线发送器 / 接收器、陆地广播发送器 / 接收器、家用娱乐系统与电影剧院系统。

[0053] (一般 BCC 处理)

[0054] 图 2 为普通的双声道提示码编码 (BCC) 音频处理系统 200 的框图, 其包括编码器 202 和解码器 204, 编码器 202 包含下混器 206 和 BCC 评估器 208。

[0055] 下混器 206 将 C 个输入音频通道 $x_i(n)$ 转换成 E 个传输音频通道 $y_i(n)$, 其中 $C > E \geq 1$ 。在此说明书中, 使用变量 n 表示的信号为时域信号, 同时使用变量 k 表示的信号为频域信号。根据特殊的实施, 下混能在时域或频域中实施。BCC 评估器 208 从 C 个输入音频通道产生 BCC 码且传输这些 BCC 码作为相对于 E 个传输音频通道的频带内或频带外辅助信息。通常的 BCC 码包含一个或多个通道间时差 (ICTD)、通道间电平差 (ICLD) 和在输入通道

的某些对间被评估作为频率与时间的函数的通道间关联 (ICC) 数据。所述特别实施将在输入通道的特定对之间指示 BCC 码被评估。

[0056] ICC 数据对应立体声信号的一致性,其与所述音源的感觉宽度有关。音源越宽,所产生立体声信号的左边与右边通道间的一致性越低。例如,对应于传过一个礼堂讲台的管弦乐队的立体声信号的一致性通常低于对应于单个小提琴独奏的立体声信号的一致性。通常,一致性较低的音频信号通常被感觉为在听觉空间中更能被传播。如此,ICC 数据通常与听者环境的明显音源宽度和程度有关。见例如,J.Blauert,The Psychophysics of Human Sound Localization, MIT press,1983。

[0057] 根据特殊的应用,所述 E 个传输音频通道和对应的 BCC 码可直接被传输到解码器 204 或储存在合适类型的储存装置中用于解码器后续存取。依据所述情况,术语“传输”可引用为直接传输至解码器或是用于对解码器后续供应的储存。在任意种情况下,解码器 204 接收传输音频通道和辅助信息并且执行上混和使用 BCC 码的 BCC 合成以将 E 个传输音频通道转换成超过 E (通常,但不必须, C) 个回放音频通道 $\hat{x}_i(n)$ 用于音频回放。根据特殊的实施,上混可在既能在时域中也能在频域中被执行。

[0058] 除图 2 中所示的 BCC 处理外,普通的 BCC 音频处理系统可包括有额外的编码和译码阶段以进一步分别在编码器压缩音频信号然后在解码器对所述音频信号解压缩。这些编解码器可基于传统的音频压缩/解压缩技术,例如那些基于脉冲码调制 (PCM)、差分 PCM (DPCM) 或适应性 DPCM (ADPCM)。

[0059] 当下混器 206 产生单一总和信号 (即, $E = 1$) 时, BCC 编码能够在比特率仅稍高于所需要表示单声道音频的信号来表示多通道音频信号,这是因为在通道对间所述经评估的 ICTD、ICLD 和 ICC 数据含有较音频波形少约两个数量级大小的信息。

[0060] 不仅对 BCC 编码的低位率,而且对其向后兼容性方面也是有利的。单一传输总和信号对应原先立体声或多通道信号的单声道下混。对于接收器,其不支持立体声或多通道音频重现,倾听传输总和信号是在低外形单声道再现设备上呈现所述音频素材的正确方法, BCC 编码可因此也被使用以提升涉及从单声道音频素材向多通道音频的传输的现有服务。例如,如果 B C C 辅助信息可被嵌入到现有传输通道中,现有单声道音频无线广播系统可被提升用于立体声或多通道回放。类似的能力存在于当下混多通道音频至对应立体声的两个总和信号。

[0061] BCC 处理具某时间与频率分辨率的音频信号,所用的所述频率分辨率主要由人体听觉系统的频率分辨率所引起,心理声学建议空间感觉最有可能基于所述音频输入信号的临界频带表示。此频率分辨率通过使用具有频宽等于或与人体听觉系统的临界频宽成正比的基频带的可反向滤波器库 (例如,基于快速傅立叶转换 (FFT) 或正交镜像滤波器 (QMF)) 被考虑。

[0062] (一般下混)

[0063] 在优选实施中,所述传输总和信号包含所述输入音频信号的全部信号成份。目标为每一个信号成份被完全保持。所述音频输入通道的简单总和导致信号成份的放大或衰减。换句话说,在“简单”总和中信号成份的功率经常是大于或小于每一个音频通道的相应信号成份的功率总和。可使用下混技术,该技术使所述总和信号均衡,以便使总和信号中的信号成份的功率大约与在全部输入通道中的相应功率相同。

[0064] 图 3 表示下混器 300 的框图,其可依据 BCC 系统 200 的特殊实施被使用于图 2 的下混器 206。下混器 300 具有滤波器库 (FB) 302 用于每个输入通道 $x_i(n)$ 、下混区块 304、可选择校准 / 延迟区块 306 和反向 FB (IFB) 308 用于每个编码通道 $y_i(n)$ 。

[0065] 每一个滤波器库 302 将时域中相应的数字输入通道 $x_i(n)$ 的每一帧 (例如, 20msec) 转换成频域中一组输入系数 $\tilde{x}_i(k)$ 。下混区块 304 将 C 相应的输入系数的每一个基频带下混成 E 经下混频域系数的相应基频带。方程式 (1) 表示输入系数 $(\tilde{x}_1(k), \tilde{x}_2(k), \dots, \tilde{x}_C(k))$ 的第 k 个基频带的下混以产生经下混系数 $(\hat{y}_1(k), \hat{y}_2(k), \dots, \hat{y}_E(k))$ 的第 k 个基频带如下:

$$[0066] \quad \begin{bmatrix} \hat{y}_1(k) \\ \hat{y}_2(k) \\ \vdots \\ \hat{y}_E(k) \end{bmatrix} = D_{CE} \begin{bmatrix} \tilde{x}_1(k) \\ \tilde{x}_2(k) \\ \vdots \\ \tilde{x}_C(k) \end{bmatrix}, \quad (1)$$

[0067] 其中 D_{CE} 为一个实值的 C-by-E 下混矩阵。

[0068] 选择的校准 / 延迟区块 306 包括一组乘法器 310, 每一个乘法器以一校准因子 $e_i(k)$ 乘上相应的经下混系数 $\hat{y}_i(k)$ 以产生相应的比例系数 $\tilde{y}_i(k)$, 用于校准操作的动机为相等于对每一个通道用于以任意加权因子下混所一般化的等化。如果输入通道为独立的, 接着每一个基频带的经下混信号的功率 $P_{\tilde{y}_i(k)}$ 以方程式 (2) 得到如下:

$$[0069] \quad \begin{bmatrix} P_{\tilde{y}_1(k)} \\ P_{\tilde{y}_2(k)} \\ \vdots \\ P_{\tilde{y}_E(k)} \end{bmatrix} = \bar{D}_{CE} \begin{bmatrix} P_{\tilde{x}_1(k)} \\ P_{\tilde{x}_2(k)} \\ \vdots \\ P_{\tilde{x}_C(k)} \end{bmatrix}, \quad (2)$$

[0070] 其中 \bar{D}_{CE} 通过对 C-by-E 下混矩阵 D_{CE} 中的每一个矩阵组件进行平方而得到, 且 $P_{\tilde{x}_i(k)}$ 为输入通道 i 的基频带 k 的功率。

[0071] 如果基频带不是独立的, 接着所述经下混信号的功率值 $P_{\tilde{y}_i(k)}$ 将大于或小于使用第 (2) 式所计算得值, 由于当信号成份分别是同相或不同相时信号放大或取消。为避免如此, 第 (1) 式的下混操作接着以乘法器 310 的校准操作被施加到基频带中, 校准因子 $e_i(k)$ ($1 \geq i \geq E$) 可由第 (3) 式得出如下:

$$[0072] \quad e_i(k) = \sqrt{\frac{P_{\tilde{y}_i(k)}}{P_{\hat{y}_i(k)}}}, \quad (3)$$

[0073] 其中, $P_{\tilde{y}_i(k)}$ 为如以第 (2) 式计算的基频带功率, 且 $P_{\hat{y}_i(k)}$ 为相应经下混基频带信号 $\hat{y}_i(k)$ 的功率。

[0074] 除了提供可选择的校准或不用可选择的校准, 校准 / 延迟区块 306 可选择地对信号施加延迟。

[0075] 每一个反向滤波器库 308 将频域中的一组相应的经校准的系数 $\tilde{y}_i(k)$ 转换成相应的数字、传输通道 $y_i(n)$ 的帧。

[0076] 虽然图 3 显示输入通道的全部 C 被转换成频域用于后续下混, 在一个替代实施中, C 个输入通道中的一个或多个 (但少于 C-1) 可避开图 3 中所显示的所述操作的一些或全部且可被传输作为未修改音频通道的相等数量, 根据所述特别实施, 这些未修改的音频通道可或不可被图 2 的 BCC 评估器 208 使用以产生传输 BCC 码。

[0077] 在下混器 300 的实施中其产生单一总和信号 $y(n)$, $E = 1$, 且每一个输入通道 c 的每一个基频带的信号 $\tilde{x}_c(k)$ 被加入且接着以因子 $e(k)$ 相乘, 依据第 (4) 式如下:

$$[0078] \quad \tilde{y}(k) = e(k) \sum_{c=1}^c \tilde{x}_c(k), \quad (4)$$

[0079] 因子 $e(k)$ 以第 (5) 式得到如下:

$$[0080] \quad e(k) = \sqrt{\frac{\sum_{c=1}^c p_{\tilde{x}_c}(k)}{p_{\tilde{x}}(k)}}, \quad (5)$$

[0081] 其中 $p_{\tilde{x}_c}(k)$ 为在时间索引 k 时 $\tilde{x}_c(k)$ 功率的短时间评估, 且 $p_{\tilde{x}}(k)$ 为功率 $\sum_{c=1}^c \tilde{x}_c(k)$ 的短时间评估, 所述相等的基频带被转换回到产生被传输至所述 BCC 解码器的总和信号的时域。

[0082] (一般 BCC 合成)

[0083] 图 4 显示 BCC 合成器 400 的框图, 其依据 BCC 系统 200 的某些实施可被使用于图 2 的解码器 204, BCC 合成器 400 具有滤波器库 402 用于每一个传输通道 $y_i(n)$, 上混区块 404, 延迟器 406, 乘法器 408, 相关区块 410 和反向滤波器库 412 用于每一个回放通道 $\hat{x}_i(n)$ 。

[0084] 每一个滤波器库 402 将时域中相应的数字、传输通道 $y_i(n)$ 的每一帧转换成频域中一组输入系数 $\tilde{y}_i(k)$ 。上混区块 404 将 E 相应的传输通道系数的每一个基频带上混成 C 经上混频域系数的一相应的基频带, 方程式 (4) 表示传输通道系数 $(\tilde{y}_1(k), \tilde{y}_2(k), \dots, \tilde{y}_E(k))$ 的第 k 个基频带的上混以产生上混系数 $(\tilde{s}_1(k), \tilde{s}_2(k), \dots, \tilde{s}_c(k))$ 的 kth 基频带如下:

$$[0085] \quad \begin{bmatrix} \tilde{s}_1(k) \\ \tilde{s}_2(k) \\ \vdots \\ \tilde{s}_c(k) \end{bmatrix} = U_{EC} \begin{bmatrix} \tilde{y}_1(k) \\ \tilde{y}_2(k) \\ \vdots \\ \tilde{y}_E(k) \end{bmatrix}, \quad (6)$$

[0086] 其中 U_{EC} 为一个实值 E-by-C 上混矩阵, 在频域中执行上混使上混能被独立地施加于每一个不同的基频带。

[0087] 每一个延迟器 406 施加基于用于 ICTD 数据的相应的 BCC 码的延迟值 $d_i(k)$ 以确保所要的 ICTD 值出现于回放通道的某些对中。每一个乘法器 408 施加基于用于 ICLD 数据的相应的 BCC 码的校准因子 $a_i(k)$ 以确保所要的 ICLD 值出现于回放通道的某些对中, 相关区块 410 执行用于 ICC 数据的相应的 BCC 码的去关联操作 A 以确保所要的 ICC 值出现于回放通道的某些对中, 相关区块的操作的进一步描述可见 2002 年 5 月 24 日申请的美国第 10/155, 437 号专利申请如 Baumgarte 2-10。

[0088] ICLD 值的合成比 ICLD 和 ICC 值的合成容易一些, 因为 ICLD 合成仅涉及基频带信

号的校准。因为 ICLD 提示信号为最通常使用的方向性提示信号,ICLD 值接近原始音频信号的这些值是通常更重要的,如此,ICLD 数据可被评估在全部通道对之间。对每一个基频带的校准因子 $a_i(k)$, ($1 \leq i \leq C$) 最好被选取使得每一个回放通道的基频带功率接近原始输入音频通道的相应的功率。

[0089] 一个目标可施加相对少的信号修改用以合成 ICTD 和 ICC 值,这样,所述 BCC 值可不包含用于全部通道对的 ICTD 和 ICC 值,在所述情形中,BCC 合成器 400 将仅在某些通道对之间合成 ICTD 和 ICC 值。

[0090] 每一个反向滤波器库 412 将一组频域中的相应的经合成系数 $\tilde{x}_i(k)$ 转换成相应的数字的、回放通道 $\hat{x}_i(n)$ 的帧。

[0091] 虽然图 4 显示全部 E 个传输通道被转换成频域用于后续上混与 BCC 处理,在另外实施中,所述 E 个传输通道中的一个或多个(但非全部)可避开图 4 所示的处理的一些或全部。例如,传输通道的一个或多个可以是未修改通道,其未接受任何上混。除了作为 C 个回放通道中的一个或多个之外,这些未修改通道,轮流地,可以是但不必须被用作为参考通道,其 BCC 处理被施加给合成其它回放通道中的一个或多个。在任情形中,这些未修改通道可受到延迟以补偿涉及上混的操作时间与 / 或用以产生其余回放通道的 BCC 操作。

[0092] 注意的是,虽然图 4 显示 C 个回放通道自 E 个传输通道被合成,其中, C 也为原始输入通道的数目, BCC 合成不限于回放通道的所述数目,通常,回放通道的数目可以是通道的任何数目,包含数目大于或小于 C 和可能甚至当回放通道的数目是等于或小于传输通道数目的情形。

[0093] (介于音频通道之间的“感觉上相对差异”)

[0094] 假设单一总和信号, BCC 合成立体声或多通道音频信号使得 ICTD、ICLD 和 ICC 接近原始音频信号的相应的提示信号,以下,关于听觉空间影像属性的 ICTD、ICLD 和 ICC 的作用将予讨论。

[0095] 关于空间听觉的知识包含有对于一个听觉事件, ICTD 和 ICLD 与感觉方向是相关的。当考虑到音源的立体声空间脉冲响应 (BRIRs) 时,在听觉事件的宽度和听者包封和为 BRIRs 的早期和后期部分评估的 ICC 数据之间具有关系。然而,在 ICC 和这些普通信号(和不只是 BRIRs) 的性质之间的关系不是直接的。

[0096] 立体声和多通道音频信号通常包含同步主动源信号的复杂混合,所述主动源信号是从围绕空间中录音所产生的经反射信号成份所迭加,或用于人工生成的空间印象的录音工程师所赋加,不同的音源信号与它们的反射占据时间-频率平面中的不同区域。此由 ICTD、ICLD 与 ICC 所反映,其作为时间与频率的函数而改变。在此情形下,瞬时现象 ICTD、ICLD 和 ICC 和音频事件方向与空间印象间的关系是不明显的。某些 BCC 实施例的策略是不明显地合成这些提示信号,以便使它们接近原始音频信号的相应的提示信号。

[0097] 具有基频带频宽等于两倍相等的矩形频宽 (ERB) 的滤波器库被使用。非正式的倾听会显示当选取较高频率分辨率时 BCC 的音频质量未显著改善。较低频率分辨率可为需求的,因为它导致较少 ICTD、ICLD 与 ICC 值需要被传输至解码器,且因此以较低比特率传输。

[0098] 关于时间分辨率, ICTD、ICLD 和 ICC 为通常在固定时间间距下被考虑,当 ICTD、ICLD 与 ICC 以约每 4 到 16ms 被考虑时,可得到高性能。注意的是,除非所述提示信号在非常短的时间间隔被考虑,先前效果未直接考虑,假设音频刺激的典型领先-落后对,假如所

述领先和落后位于时间间隔仅一组提示信号被合成，则所述领先的局部化优势未被考虑。虽然如此，BCC 达到音频质量以平均 MU SHRA 分数反映为平均约 87（即，“极佳”音频质量），且对某些音频信号高到接近于 100。

[0099] 参考信号与经合成信号间的所述经常得到的感觉上小差异暗示关于宽度范围的听觉空间影像属性的提示信号为暗示性地在固定时间间隔被合成 ICTD、ICLD 和 ICC 所考虑。以下，一些论点对于 ICTD、ICLD 和 ICC 可如何与听觉空间影像属性的范围有关。

[0100] （空间提示信号的评估）

[0101] 以下中，将描述 ICTD、ICLD 和 ICC 如何被评估，用于这些（经量化的与编码的）空间提示信号的传输比特率可为刚好为几个 kb/s 并因此，使用 BCC，它可能在比特率接近对单一音频通道的要求下传输立体声与多通道音频信号。

[0102] 图 5 显示依据本发明，图 2 的 BCC 评估器 208 的框图，BCC 评估器 208 包括滤波器库 (FB) 502，其可与图 3 的滤波器库 302 相同，和评估区块 504 其对由滤波器库 502 所产生的每一个不同频率产生 ICTD、ICLD 与 ICC 空间提示信号。

[0103] （用于立体声信号的 ICTD、ICLD 和 ICC 的评估）

[0104] 以下量测为使用于 ICTD、ICLD 和 ICC 用以相应的二（例如，立体声）音频通道的基频带信号 $\tilde{x}_1(k)$ 与 $\tilde{x}_2(k)$ 。

[0105] ICTD [例子]

$$[0106] \quad \tau_{12}(k) = \arg \max_d \{ \Phi_{12}(d, k) \} \quad (7)$$

[0107] 具有由以下第 (8) 式得到的经标准化交叉相关函数的短时间评估。

$$[0108] \quad \Phi_{12}(d, k) = \frac{p_{\tilde{x}_1\tilde{x}_2}(d, k)}{\sqrt{p_{\tilde{x}_1}(k-d_1)p_{\tilde{x}_2}(k-d_2)}} \quad (8)$$

[0109] 其中

$$[0110] \quad d_1 = \max \{-d, 0\} \quad (9)$$

$$[0111] \quad d_2 = \max \{d, 0\}$$

[0112] 且， $p_{\tilde{x}_i\tilde{x}_i}(d, k)$ 为 $\tilde{x}_1(k-d_1)\tilde{x}_2(k-d_2)$ 平均数的短时间评估。

[0113] ICLD [dB]

$$[0114] \quad \Delta L_{12}(k) = 10 \log_{10} \left(\frac{p_{\tilde{x}_2}(k)}{p_{\tilde{x}_1}(k)} \right) \quad (10)$$

[0115] ICC

$$[0116] \quad c_{12}(k) = \max_d |\Phi_{12}(d, k)| \quad (11)$$

[0117] 注意经标准化交叉相关的绝对值被考虑且 $c_{12}(k)$ 具有 [0, 1] 的范围。

[0118] （用于多通道音频信号的 ICTD、ICLD 和 ICC 的评估）

[0119] 当有超过两个输入通道，它通常足以在参考通道间限定 ICTD 和 ICLD（例如，音频通道号码 1）与其它通道，如图 6 中所说明用于 $C = 5$ 个通道的情形，其中 $\tau_{1c}(k)$ 与 $\Delta L_{12}(k)$ 在参考通道 1 与通道 c 之间分别指示 ICTD 与 ICLD。

[0120] 与 ICTD 和 ICLD 相反的，ICC 通常具有较多自由度，所限定的 ICC 在所有可能的输入通道对之间具有不同的值，对 C 个通道而言，具有 $C(C-1)/2$ 个可能的音频通道对，例如，

对 5 个通道会有如图 7(a) 中所例示的 10 个通道对, 然而, 这些方式需要在每一时间索引对每一基频带评估且传输 $C(C-1)/2$ 个 ICC 值, 导致高计算复杂度与高比特率。

[0121] 或者, 对每一基频带, 实施 ICTD 与 ICLD 决定基频带中相应的信号成份的音频事件的方向。每基频带的单一 ICC 参数可接着被用于描述全部音频通道间的整体一致性, 良好的结果可通过仅在每一时间索引的每一基频带中具有最多能量的两个通道间评估和传输 ICC 提示信号而得到。此例示于图 7(b) 中, 其中时间瞬间 $k-1$ 与 k 的所述通道对 (3, 4) 与 (1, 2) 分别为最强。启发式规则可被用于在其它通道对间决定 ICC。

[0122] (空间提示信号的合成)

[0123] 图 8 显示图 4 的 BCC 合成器 400 的实施框图, 其在给单一传输总和信号 $s(n)$ 加空间提示信号下, 可被使用于 BCC 解码器中以产生立体声或多通道音频信号。总和信号 $s(n)$ 被分解成基频带, 其中 $\tilde{s}(k)$ 指示这些基频带。为产生每一个输出通道的相应的基频带, 延迟 d_c 校准因子 a_c 与滤波器 h_c 被施加至总和信号的相应的基频带, (为简化表示, 时间索引 k 在延迟、校准因子和滤波器中被省略), ICTD 通过加上延迟, ICTD 通过校准和 ICC 通过施加去相关滤波器被合成, 图 8 中所示的处理被独立地施加至每一基频带。

[0124] (ICTD 合成)

[0125] 延迟 d_c 从 ICTDs $\tau_{1c}(k)$ 被决定, 依据如下第 (12) 式:

$$[0126] \quad d_c = \begin{cases} -\frac{1}{2}(\max_{2 \leq l \leq C} \tau_{1l}(k) + \min_{2 \leq l \leq C} \tau_{1l}(k)), & c = 1 \\ \tau_{1l}(k) + d_1 & 2 \leq c \leq C \end{cases} \quad (12)$$

[0127] 用于参考通道的延迟 d_1 被计算使得延迟 d_c 的最大数量被最小化, 越少基频带信号被修改, 越少的人为危害产生, 假如基频带取样率对 ICTD 合成未提供够高的时间分辨率, 延迟可通过使用合适的全通滤波器更准确地被加于其上。

[0128] (ICLD 合成)

[0129] 为使输出基频带信号在通道 c 和参考通道 1 具有所要的 ICLDs $\Delta L_{12}(k)$, 增益因子 a_c 应所述满足如下第 (13) 式:

$$[0130] \quad \frac{a_c}{a_1} = 10^{\frac{\Delta L_{1c}(k)}{20}} \quad (13)$$

[0131] 此外, 输出基频带最好被标准化使得全部输出通道的功率与输入总和信号的功率相等。因为在每一基频带中的全部原始信号功率被保存在总和信号中, 在绝对基频带功率中的此标准化结果对每一个输出通道接近原始编码器音频信号的相应的功率, 在这些限制下, 校准因子 a_c 由以下第 (14) 式得到。

[0132]

$$a_c = \begin{cases} 1/\sqrt{1 + \sum_{i=2}^c 10^{\Delta L_{1i}/10}}, & c = 1 \\ 10^{\Delta L_{1c}/20} a_1, & \text{其它} \end{cases} \quad (14)$$

[0133] (ICC 合成)

[0134] 在某些实施例中, ICC 合成的目标为在延迟后的基频带间降低相关且校准已被施加, 而不会影响 ICTD 和 ICLD。此可通过设计图 8 中的滤波器 h_c 而达到, 使得 ICTD 和 ICLD

如同一频率函数有效地被改变,使得在每一基频带(音频临界频带)中平均变异为 0。

[0135] 图 9 说明 ICTD 和 ICLD 如何在一基频带中作为频率函数被改变, ICTD 和 ICLD 变异的振幅决定去相关的程度且作为 ICC 函数被控制,注意 ICTD 被平缓地改变(如图 9(a)),同时 ICLD 被任意改变(如图 9(b))。可如同 ICTD 平缓般地变化 ICLD,但此将导致音频信号产生更多的声染色。

[0136] 用于合成 ICC 的另一方法,特别适合于多通道 ICC 合成,被更详细描述于 C. Faller, "Parametric multi-channel audio coding: Synthesis of coherence cues," IEEE Trans. on Speech and Audio Proc., 2003, 其启示被并入于此以供参考,作为时间和频率的函数,人为后期回响 (late reverberation) 的特定量被加于每一个输出通道用以获得想要的 ICC,另外,频谱修改可被施加以使得产生信号的频谱包络接近原始音频信号的频谱包络。

[0137] 其它用于立体声信号(或音频通道对)的相关与不相关的 ICC 合成技术已发表于 E. Schuijers, W. Oomen, B. den Brinker, and J. Breebaart, "Advances in parametric coding for high-quality audio," in Preprint 114th Conv. Aud. Eng. Soc., Mar. 2003, and J. Engdegard, H. Purnhagen, J. Roden, and L. Liljeryd, "Synthetic ambience in parametric stereo coding," in Preprint 117th Cov. Aud. Eng. Soc., May 2004, 二者的启示并入于此以供参考。

[0138] (C-to-E BCC)

[0139] 如先前描述, BCC 可以超过传输通道被实施, BCC 的变形已被描述,其代表 C 个音频通道并非为单一(传输)通道,但作为 E 个音频通道,标示为由 C 到 E (C-to-E) BCC。对 C-to-E BCC 至少有两个动机:

[0140] 具备传输通道的 BCC 提供向后 (backwards) 可兼容路径用以升级现有的单声道系统用于立体声或多通道音频回放,所述经升级的系统通过现有的单声道架构传输 BCC 下混总和信号,从 C 到 E (C-to-E) 的 BCC 可施加 C 个通道音频的向后可兼容的编码至 E 个通道。

[0141] 从 C 到 E 的 BCC 以传输通道数目的不同程度的减少引进校准。可以预期当更多的音频通道被传输会有更佳的音频质量。

[0142] 对从 C 到 E 的 BCC 的信号处理细节,诸如如何定义 ICTD、ICLD 和 ICC 提示信号,被描述于 2004 年 1 月 20 日的美国第 10/762,100 号专利申请中 (Faller 13-1)。

[0143] (散射声音整形)

[0144] 在某些实施中, BCC 编码包括用于 ICTD、ICLD 和 ICC 合成的算法。ICC 提示信号可以通过对在相应的基频带中的信号分量进行去关联被合成。这可以通过 ICLD 的频率相关变化、ICTD 和 ICLD 的频率相关变化、全通滤波或者通过与回响算法相关的想法来完成。

[0145] 当这些技术在音频信号上使用时,所述信号的时序包络特征不被保存。特别地,当被应用到瞬时现象上时,瞬时信号能量可能被传播了一段时期。这就导致了人为结果例如“前回声”或者“模糊的瞬时现象”。

[0146] 本发明的某些实施例的一般原理与观测结果有关,所述观测结果为 BCC 解码器合成的声音应该不仅具有与原始声音相似的空间特征,还应该与所述原始声音的时序包络非常近似,以便具有相似的感知特征。通常这是在通过包括动态 ICLD 合成的类似 BCC 方案中实现的,其对大约每个信号通道的时序包络进行时间变化校准操作。对于瞬变信号(突发、

打击乐器等),这种处理的瞬时清晰度可以,然而,不足以产生合成信号,该合成信号足够接近原始时序包络。本节描述了许多具有十分精细的时间分辨率的方法来实现这个。

[0147] 另外,对于不能访问所述原始信号的时序包络的 BCC 解码器,思路是将所述传输“总和信号”的时序包络作为近似值替换。这样,就不需要将辅助信息从所述 BCC 编码器到所述 BCC 解码器进行传输以传送这样的包络信息。总之,本发明依赖下面的原则:

[0148] 所述传输音频通道(即“总和通道”)或者 BCC 合成可能基于的这些通道的线性组合通过时序包络提取器进行分析用于其具有高时间分辨率的时序包络(例如,比 BCC 区块的大小更显著地精细)。

[0149] 用于每个输出通道的所述后续合成声音被整形以便-即使是在 ICC 合成之后-其能够尽量地与通过所述提取器所决定的时序包络相匹配。这会保证即使在瞬时信号的情况下,所述合成的输出声音并没有被 ICC 合成/信号去关联处理显著地降低品质。

[0150] 图 10 显示的是根据本发明的一个实施例,表示 BCC 解码器 1000 至少一部分的框图。在图 10 中,区块 1002 表示 BCC 合成处理,其包括,至少,ICC 合成。BCC 合成区块 1002 接收基通道 1001 并产生合成通道 1003。在某些实施中,区块 1002 表示图 4 中的区块 406、408 和 410 的处理,其中基通道 1001 为上混区块 404 产生的信号,并且合成通道 1003 是关联区块 410 产生的信号。图 10 表示对一个基通道 1001 和它相应的合成通道实施的处理。相似的处理也被实施在每个其他的基通道和它相应的合成通道上。

[0151] 包络提取器 1004 决定基通道 1001' 的细微时序包络 a, 并且包络提取器 1006 决定合成通道 1003' 的细微时序包络 b。反包络调节器 1008 使用来自包络提取器 1006 的时序包络 b 以标准化合成通道 1003' 的所述包络(即“平滑”所述时序细微结构)来产生具有标记(即统一的)时间包络的平滑的信号 1005'。根据特殊的实施,平滑化可以在上混前或上混后实施。包络调节器 1010 使用来自包络提取器 1004 的时序包络 a 以对平滑信号 1005' 上的原始信号包络进行再加强来产生具有与基通道 1001 的时序包络大体上相等的时序包络的输出信号 1007'。

[0152] 根据所述实施,此时序包络处理(在此也引用为“包络整形”)可以应用在整个合成通道(如所示的那样)或者只应用在所述合成通道(如后面所描述的)的正交的部分(例如,后期回响部分、去关联部分)。此外,根据所述实施,包络整形可以应用在时域信号或者以频率依赖的方式应用(例如,所述时序包络分别以不同的频率被评估和加强)。反包络调节器 1008 和包络调节器 1010 可以依不同的方式实施。在一种实施方式中,信号的包络通过信号的时域样本(或者频谱/基带样本)和时间变化的振幅改变函数(例如,用于反包络调节器 1008 的 $1/b$ 和用于包络调节器 1010 的 a) 的相乘来进行操作。可选择地,所述信号的关于频率的频谱表示的卷积/滤波可以在现有技术中为了对低速率音频编码器的量化噪声进行整形为目的的方式被使用。相似地,信号的时序包络可以通过分析信号的时间结构或者检查关于频率的信号频谱的自动关联直接地被提取。

[0153] 图 11 表现的是图 4 中的 BCC 合成器 400 范围内的图 10 的包络整形方案的示范性应用。在本实施例中,有单一传输总和信号 $s(n)$, 所述 C 个基信号通过复制那个总和来产生,并且包络整形被单独地应用到不同的基带。在替换实施例中,延迟、校准和其他处理的顺序可以不同。此外,在替换实施例中,包络整形并不限定为独立地处理每个基带。对于基于卷积/滤波的实施来说使用频带的协方差来得到关于所述信号时序细微结构的信息是

特别准确的。

[0154] 在图 11(a) 中时序处理分析器 (TPA) 1104 与图 10 中的包络提取器 1004 相似, 并且每个时序处理器 (TP) 1106 与图 10 中的包络提取器 1006、反包络调节器 1008 和包络调节器 1010 的组合相似。

[0155] 图 11(b) 为 TPA1104 的一个可能的基于时域的实施框图, 其中所述基信号样本被平方 (1110), 然后被低通滤波 (1112) 以对所述基信号的时序包络 a 进行特性化。

[0156] 图 11(c) 为 TP1106 的一个可能的基于时域的实施框图, 其中所述合成信号样本被平方 (1114), 然后被低通滤波 (1116) 以对所述合成信号的时序包络 b 进行特性化。一个校准因子 (例如, $\sqrt{a/b}$) 被产生, 然后被应用到合成信号上以产生具有与所述原始基通道的时序包络大体上相等的时序包络的输出信号。

[0157] 在 TPA1104 和 TP1106 的替换实施中, 通过使用量值操作而不是将所述信号样本平方而使所述时序包络被特性化。在这样的实施中, a/b 的比率可以用作校准因子而不用进行平方根的操作。

[0158] 虽然图 11(c) 中的所述校准操作对应于 TP 处理的基于时域的实施, 但是 TP 处理 (又 TPA 和反 TP (ITP) 处理) 也能使用频域信号, 如图 17 ~ 18 (后面描述) 中的实施例中的, 进行实施。这样, 为了本说明书的目的, 术语“校准函数”应该被理解为覆盖时域或者频域操作, 例如图 18(b) 和 (c) 中的滤波操作。

[0159] 通常, TPA1104 和 TP1106 优选地被设计以便其不修改信号功率 (即, 能量)。根据特殊实施, 该信号功率可以是每个通道的短时间平均信号功率, 例如, 基于合成窗定义的时间段里的每通道的总信号功率或者一些其他合适的功率量。这样, ICLD 合成 (例如使用乘法器 408) 的校准能够在包络整形之前或之后应用。

[0160] 注意到在图 11(a) 中, 每个通道有两个输出, 其中 TP 处理只被应用到他们其中的一个。这反映出 ICC 合成方案, 该方案混合两个信号分量: 未修改的和正交的信号, 其中未修改的和正交的信号的比率决定 ICC。在图 11(a) 所示的实施例中, TP 只被应用到正交的信号分量上, 其中总和节点 1108 将未修改信号分量和相应的时序整形的正交的信号分量重新组合。

[0161] 图 12 表现了图 4 中的 BCC 合成器 400 范围内的图 10 中的包络整形方案的替换示范性实施, 其中包络整形被应用到时域中。这样的实施例可以被保证, 当频谱表示, 其中 ICTD、ICLD 和 ICC 被执行, 的时间分辨率对于通过加强需要的时序包络来有效阻止“前回响”的时候。例如, 这会是一种情况, 当 BCC 实施短时间傅立叶变换 (STFT) 的时候。

[0162] 如图 12(a) 所示, TPA1204 和每个 TP1206 在时域被实施, 其中全基带信号被校准以便其具有期望的时序包络 (例如, 从传输总和信号评估的包络)。图 12(b) 和 (c) 为与图 11(b) 和 (c) 所示相似的 TPA1204 和 TP1206 的可能的实施。

[0163] 在此实施例中, TP 处理被应用到所述输出信号, 而不仅是正交信号分量。在替换实施例中, 如果希望的话, 基于时域的 TP 处理能够被仅应用到正交信号分量上, 其中未修改和正交的基频带 将会被转换到具有分开的反向滤波库的时域。

[0164] 由于 BCC 输出信号的全频带校准可以导致人为现象, 所以包络整形可以只在指定的频率应用, 例如, 频率高于某个截止频率 f_{TP} (例如, 500Hz)。注意到用于分析 (TPA) 的频率范围可以与用于合成 (TP) 的频率范围不同。

[0165] 图 13(a) 和 (b) 为 TPA1204 和 TP1206 的可能的实施, 其中包络整形只在高于所述截止频率 f_{TP} 的频率应用。特别地, 图 13(a) 示出了高通滤波器 1302 的附加部分, 其在时序包络特性化之前滤出低于 f_{TP} 的频率。图 13 为具有在两个基频带之间的 f_{TP} 的截止频率的两频带滤波库 1304, 其中只有高频部分被时序整形。两频带反向滤波库 1306 然后将低频部分与时序整形的高频部分进行重新组合以产生所述输出信号。

[0166] 图 14 表现的是所在 2004 年 4 月 1 日申请的代理人号是 Baumgarte 7-12 的美国申请号为 10/815, 591 号申请描述的基于后期回响 ICC 合成方案的范围内的图 10 种的包络整形方案的示范性应用。在此实施例中, TPA1404 和每个 TP1406 在时域中应用, 如图 12 或图 13 所示, 但是其中每个 TP1406 被应用到来自不同的后期回响 (LR) 区块 1402 的输出上。

[0167] 图 15 所示的为根据本发明的实施例的表示 BCC 解码器 1500 至少一个部分的框图, 其可以与图 10 所示的方案进行替换。在图 15 中, BCC 合成区块 1502、包络提取器 1504 和包络调节器 1510 与图 10 中的 BCC 合成区块 1002、包络提取器 1004 和包络调节器 1010 相似。在图 15 中, 然而, 反包络调节器 1508 在 BCC 合成之前被应用, 而不是 BCC 合成之后, 如图 10 所示。这样, 反包络调节器 1508 在 BCC 合成应用之前对基通道进行平滑处理。

[0168] 图 16 所示为根据本发明的实施例的表示 BCC 解码器 1600 至少一部分的框图, 其可以与图 10 和图 15 所示的方案互换。在图 16 中, 包络提取器 1604 和包络调节器 1610 与图 15 中的包络提取器 1504 和包络调节器 1510 相似。在图 15 中的实施例, 然而, 合成区块, 1602 表示与图 16 所示的相似的基于后期回响的 ICC 合成。在这种情况下, 包络整形只被应用到不关联的后期回响信号, 并且总和节点 1612 将时序整形的后期回响信号加到所述原始基通道 (其具有期望的时序包络)。要注意的是, 在这种情况下, 不需要使用反包络调节器, 因为后期回响信号具有在区块 1602 中的产生处理中生成的大约平的时序包络。

[0169] 图 17 为图 4 中的 BCC 合成器 400 的范围内的图 15 中的包络整形方案的示范性应用。在图 17 中, TPA1704、反 TP (ITP) 1708 和 TP1710 与图 15 中的包络提取器 1504、反包络调节器 1508 和包络调节器 1510 相似。

[0170] 在此基于频率的实施例中, 通过对沿着频率轴的 (例如, STFT) 滤波库 402 的频率码使用卷积来进行发散声音的包络整形。在此可参考美国专利 5, 781, 888 (Herre) 和美国专利 5, 812, 971 (Herre), 其启示在此引用作为参考, 其主题与这项技术相关。

[0171] 图 18(a) 表现的是图 17 中的 TPA1704 的可能的实施地框图。在此实施中, TPA1704 作为线性预测编码 (LPC) 分析操作被实施, 其决定最合适的预测系数用于一系列有关频率的频谱系数。这种 LPC 分析技术是众所周知的, 例如, 从语音编码和很多用于 LPC 系数的有效计算的算法可知, 例如自动关联方法 (涉及信号自动关联函数和后续 Levinson-Durbin 递归)。作为这个计算的结果, 一套 LPC 系数在表示信号时序包络的输出是可用的。

[0172] 图 18(b) 和 (c) 表示的是图 17 的 ITP1708 和 TP1710 的可能实施的框图。在这两个实施中, 将要处理的信号的频谱系数以频率的顺序 (增加或减小) 被处理, 其在此通过旋转开关电路被符号化, 通过预先滤波处理 (在此处理之后再次回来) 将这些系数转化为一系列用于处理的顺序。在 ITP1708 的情况下, 预先滤波计算预留量并以这种方式平滑所述时序信号包络。在 TP1710 的情况下, 所述反滤波器从 TPA1704 重新引入 LPC 系数表示的所述时序包络。

[0173] 对于通过 TPA1704 的信号时序包络的计算, 重要的是消除滤波库 402 的分析窗的

影响,如果使用这样的窗。这可以通过以分析窗整形标准化结果包络或使用分开的不使用分析窗的分析滤波库来实现。

[0174] 图 17 中的基于卷积 / 滤波技术也可以在图 16 中的包络整形方案的范围内应用,其中包络提取器 1604 和包络调节器 1610 分别基于图 18(a) 的 TPA 和图 18(c) 的 TP。

[0175] (另外可替换实施例)

[0176] BCC 解码器能够被设计用于选择地开启 / 关闭包络整形。例如当合成信号的时序包络充分地波动时,BCC 解码器能够应用传统的 BCC 合成方案和开启包络整形,以便包络整形的好处大于任何包络整形产生的人为影响。该开启 / 关闭控制可以通过以下方式实现:

[0177] (1) 瞬时现象检测:如果检测到瞬时现象,那么 TP 处理被启动。瞬时现象检测能够以展望的方式实施以在瞬时现象之前和之后立刻有效的即对瞬时现象整形也能对信号整形。检测瞬时现象可能的方式包括:

[0178] 当有指示瞬时现象发生的突然的功率上的增加出现时,观察传输 BCC 总和信号的时序包络以进行检测;和

[0179] 检查预先(LPC)滤波器的倍率。如果 LPC 预先倍率超出指定的阈值,则可设想信号为瞬时现象或高波动。LPC 的分析关于频谱自动关联被计算。

[0180] (2) 随机检测:当时序包络随机假波动时,存在一些情景。在这些情景中,没有瞬时现象被检测到,但是 TP 处理可以仍然被实施(例如,相应于这种情景的热烈鼓掌的信号)。

[0181] 另外,在某些实施中,为了阻止可能的音调信号的人为影响,当传输总和信号为高时,TP 处理不被实施。

[0182] 此外,相似的方法可以在 BCC 编码器中使用以在 TP 处理应该被激活时进行检测。因为编码器可以访问所有原始输入信号,其可以使用更精密的算法(例如,评估区块 208 的一部分)以便在 TP 处理应该启动时进行决定。这个决定的结果(当 TP 应该被激活时,发出信号)能够被传输到 BCC 解码器(例如,图 2 中的辅助信息的一部分)。

[0183] 虽然本发明已就 BCC 编码方面被描述,其中具有单一总和信号,本发明也可在具有两个或多个总和信号的 BCC 编码方面被实施,在此情形下,用于每一个不同“基础”的总和信号的时序包络可于施加 BCC 合成前被评估,且不同的 BCC 输出通道可基于不同的时序包络被产生,根据总和信号被用以合成不同输出通道,输出通道从两个或多个总和通道被合成可基于有效的时序包络被产生,所述时序包络将所述构成总和通道的相对效果列入考虑(例如,通过加权平均)。

[0184] 虽然本发明已描述了涉及 ICTD、ICLD 和 ICC 码的 BCC 码的方面,本发明也可在仅涉及这三种码(例如,ICLD、ICC 而非 ICTD)类型中的一个或两个的 BCC 码方面实施和 / 或额外码类型中的一个或多个,而且, BCC 合成处理的顺序与包络整形可在不同实施中变化,例如,当包络整形被施加至频域信号,如图 14 与 16,包络整形可于 ICTD 合成(于那些使用 ICTD 合成的实施例中)后但先于 ICLD 合成另外被实施,在其它实施例中,包络整形于任何其它 BCC 合成被施加前可被施加至上混信号。

[0185] 虽然本发明已在 BCC 编码方案方面进行了描述,本发明也可在其它音频处理方面实施,其中音频信号被去相关或需要去相关信号的其它音频处理。

[0186] 虽然本发明已在实施方面进行了描述,其中编码器在时域中接收输入音频信号,

且在时域中产生传输音频信号,且解码器在时域中接收传输音频信号,且在时域中产生回放音频信号,本发明不限于此,例如,在其它实施中,任意一个或多个输入、传输和回放音频信号可被表示于频域中。

[0187] BCC 编码器和 / 或解码器可与多种不同应用或系统连接或被并入多种不同应用或系统中,包含用于电视或电子音乐发布、电影院、广播、流向和 / 或接收的系统,这些包含系统用于编码 / 解码传输通过,例如,地面、卫星、有线电视、因特网、网间网络或物理媒介(例如,磁盘、数字磁盘、半导体芯片、硬盘、记忆卡和相类物),BCC 编码器和 / 或解码器也可被使用于游戏与游戏系统中,包含,例如,想要与娱乐用的使用者互动的交互式软件产品和 / 或可被出版用于多项机器、平台或媒介的教育,进而 BCC 编码器和 / 或解码器可被并入于 PC 软件应用,其是结合数字解码(例如,播放机、解码器)和结合数字编码能力的软件应用(例如,编码器、重编码器、自动点唱机)。

[0188] 本发明可以以基于电路的制程被实现,包含作为单一集成电路(如,ASIC 或 FPGA)、多芯片模块、单一卡片或多卡电路组的可能的实施,其对本领域技术人员电路组件的各种功能也可如软件程序的处理步骤被实施将是明显的,这些软件也可被使用于例如,数字信号处理器、微控制器或一般计算机。

[0189] 本发明也可具体表现在方法和用以实施这些方法的设备中,本发明也可被具体实施在包含在实体媒介的程序代码中,如磁盘、CD-ROMs、硬盘或任何其它机器可读取储存媒体,其中当程序代码被加载且通过机器如计算机执行,所述机器变成用以实施本发明的设备,本发明也可被具体表现于程序代码,例如,是否储存在储存媒体、通过机器加载或执行或传输经过一些传输媒体或载体,如以电线或有线、通过光纤或通过电磁辐射,其中,当程序代码通过机器如计算机被加载和执行,所述机器变成用以实施本发明的设备,当在一般处理器上实施时,所述程序代码区段结合所述处理器用以提供特殊装置,其操作为类似于特定逻辑电路。

[0190] 它将进而了解到在细节、材料与已描述和说明以便解释本发明的本质的零件配置上的各种变化,对本领域技术人员来说,可无需脱离本发明表示在以下的权利要求书而实现。

[0191] 虽然以下方法权利要求书中的步骤,若有的话,可以特定顺序和相应的标示被详述,除非所述权利要求书详述另外暗指特定顺序用以实施这些步骤的一些或全部,这些步骤不必被限定为以所述特定顺序被实施。

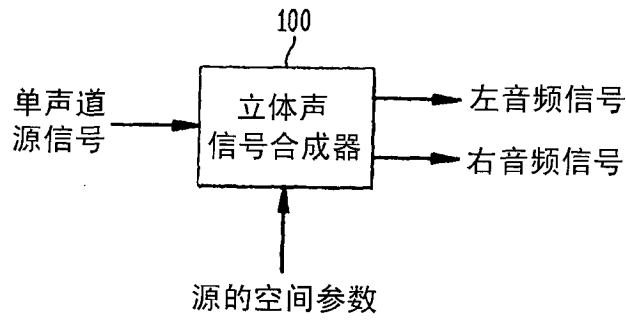


图 1

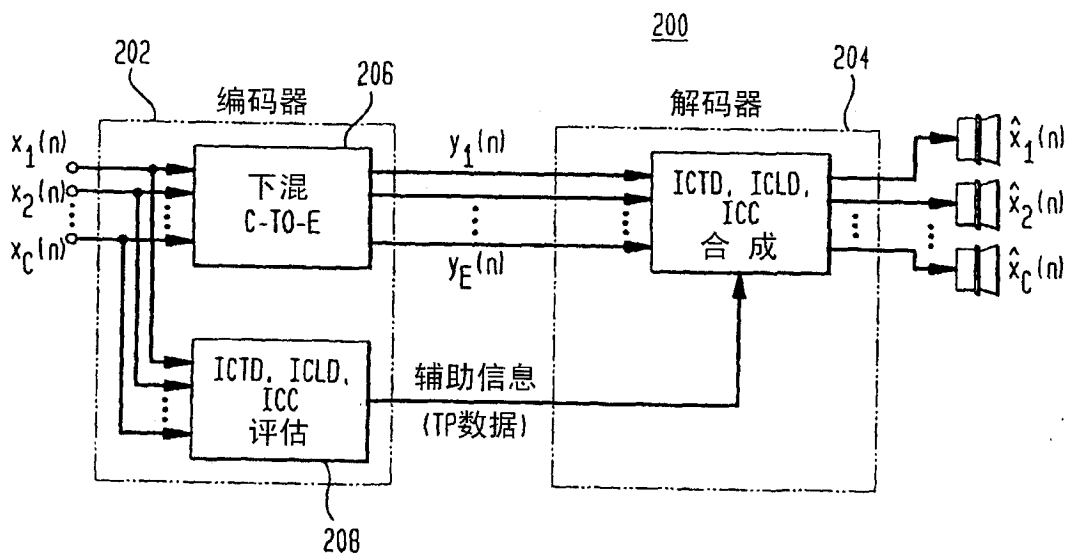


图 2

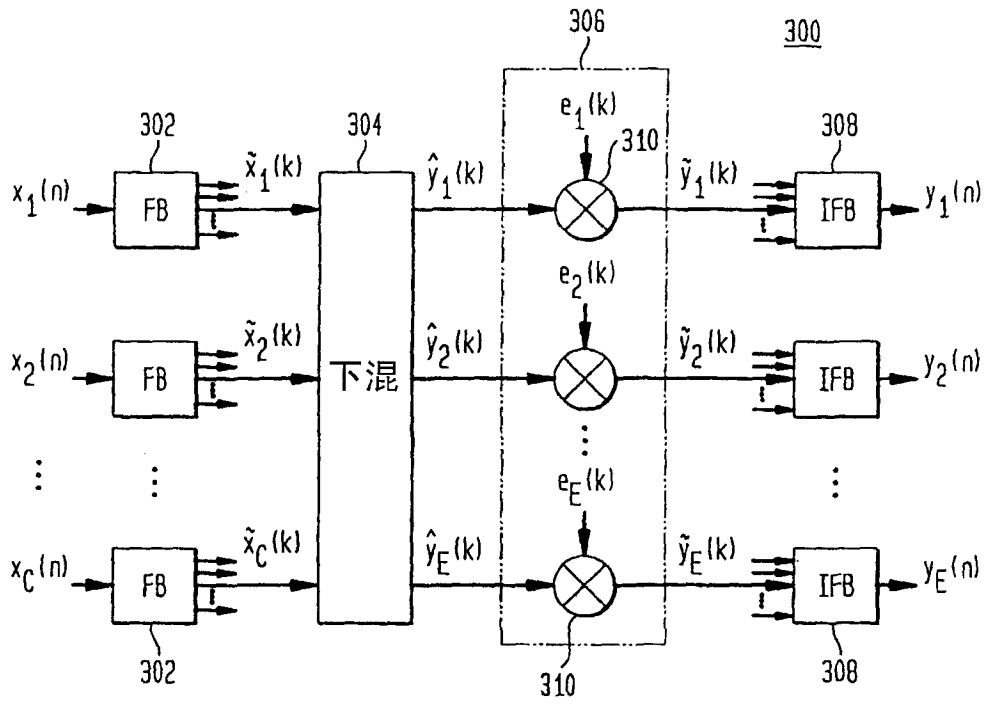


图 3

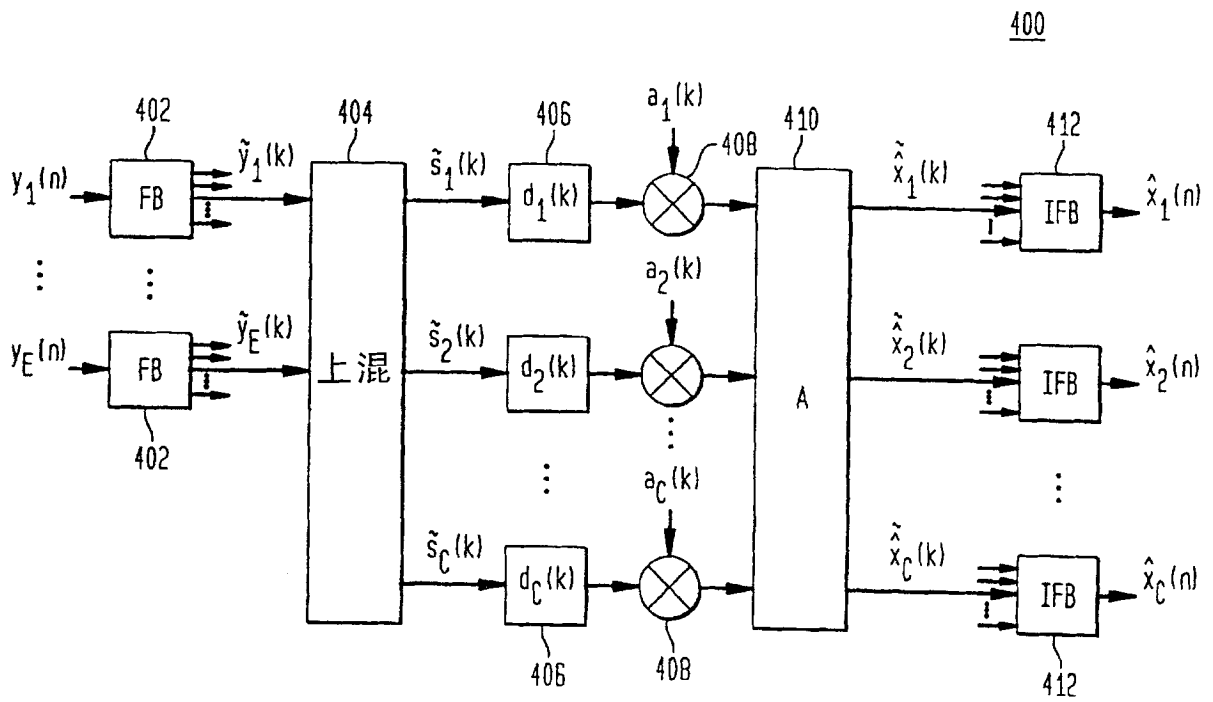


图 4

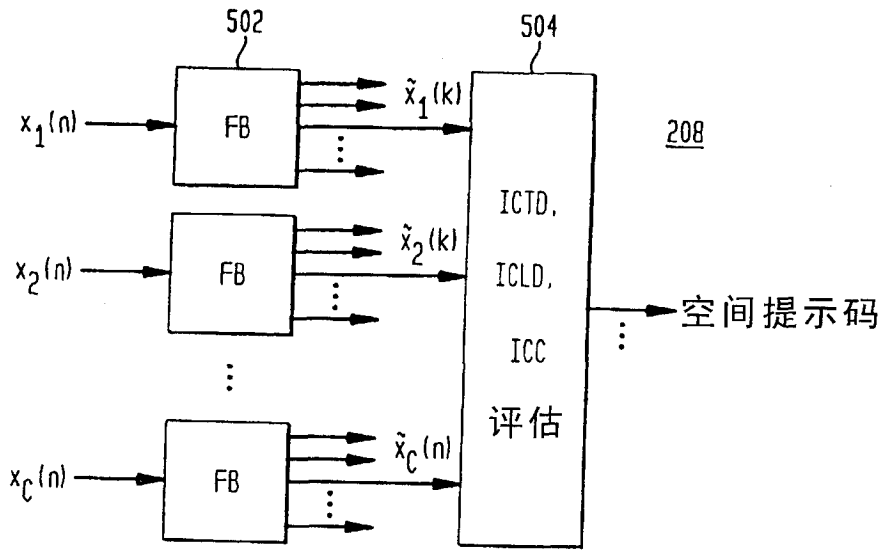


图 5

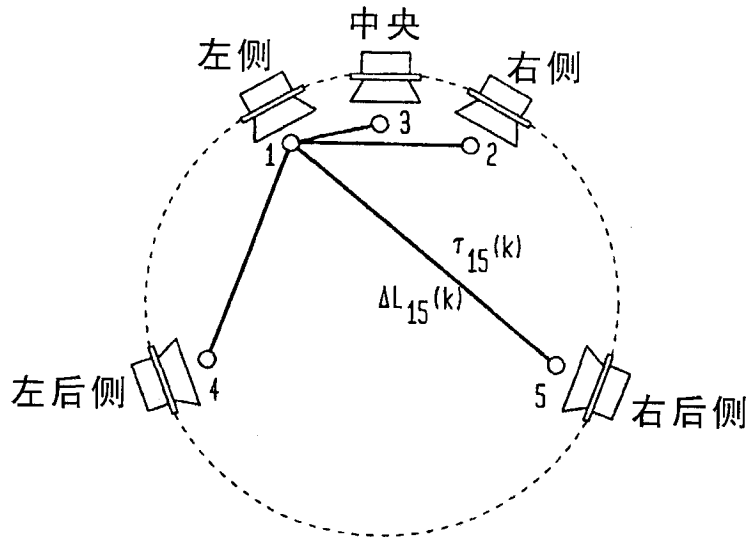


图 6

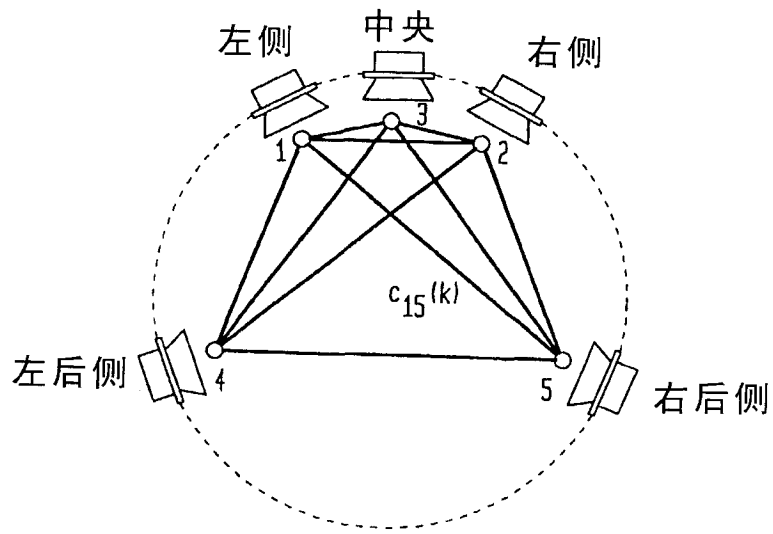


图 7(a)

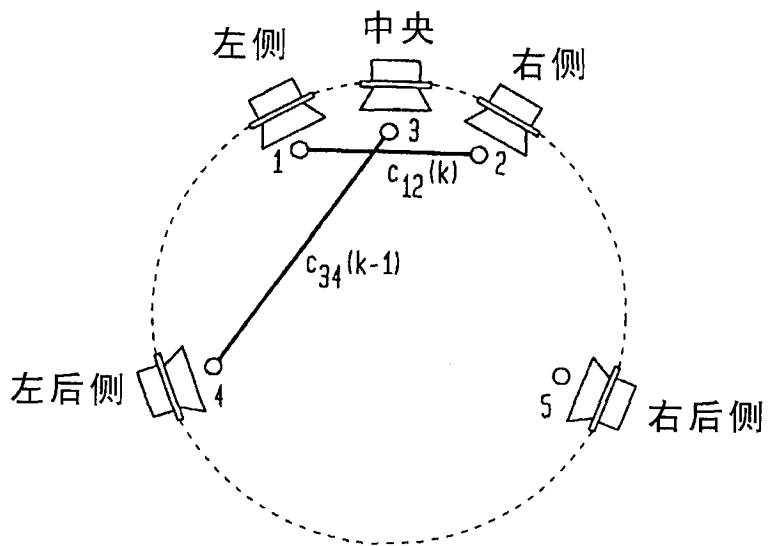


图 7(b)

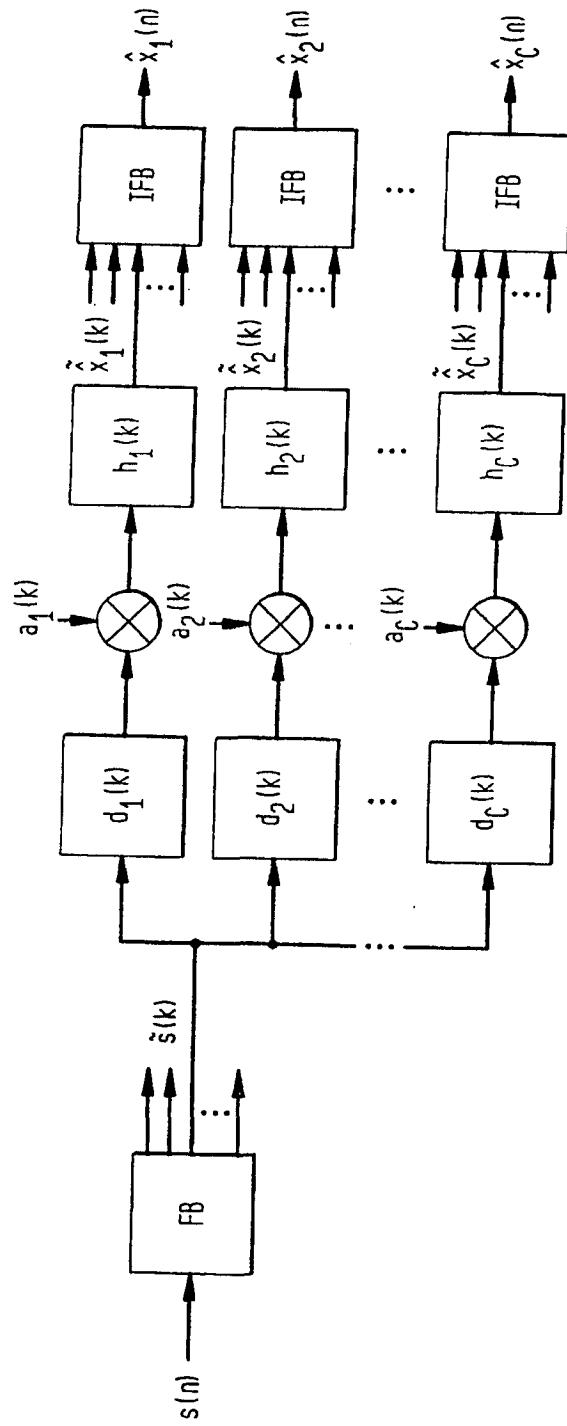


图 8

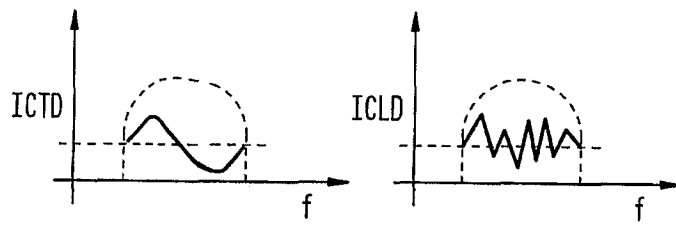


图 9

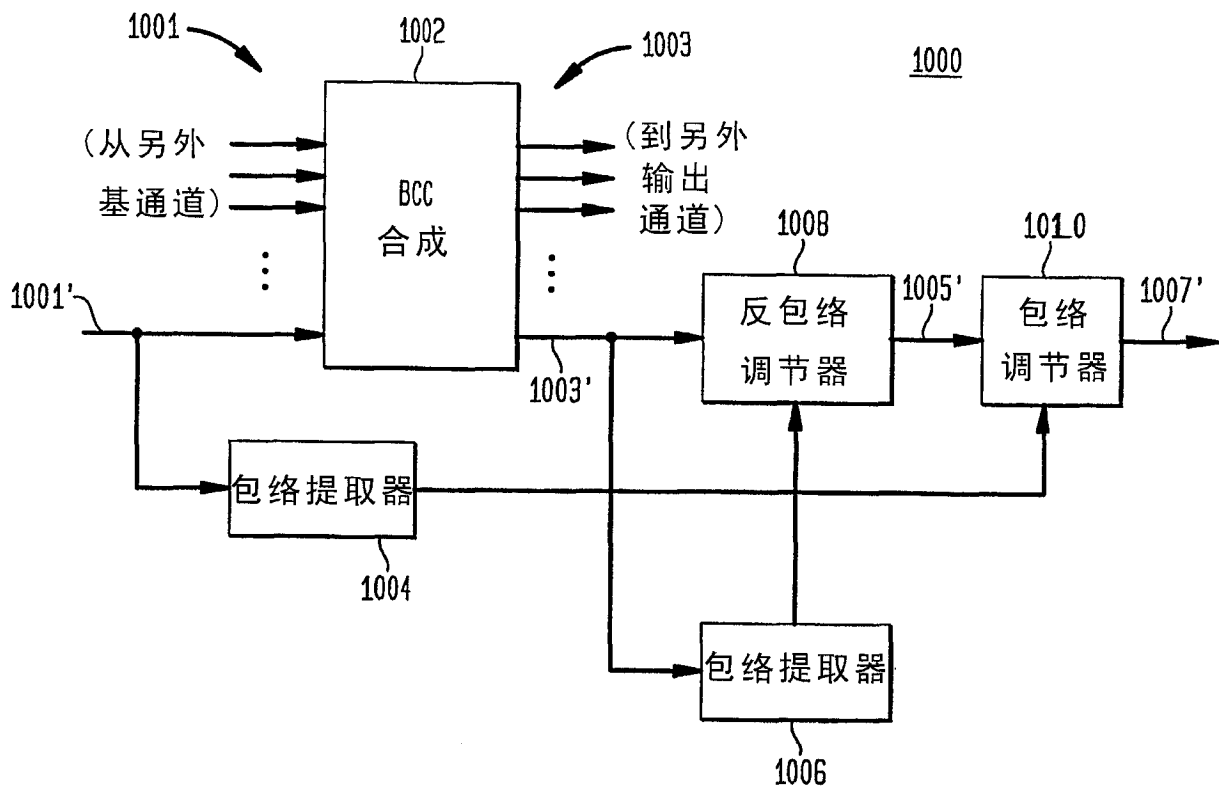


图 10

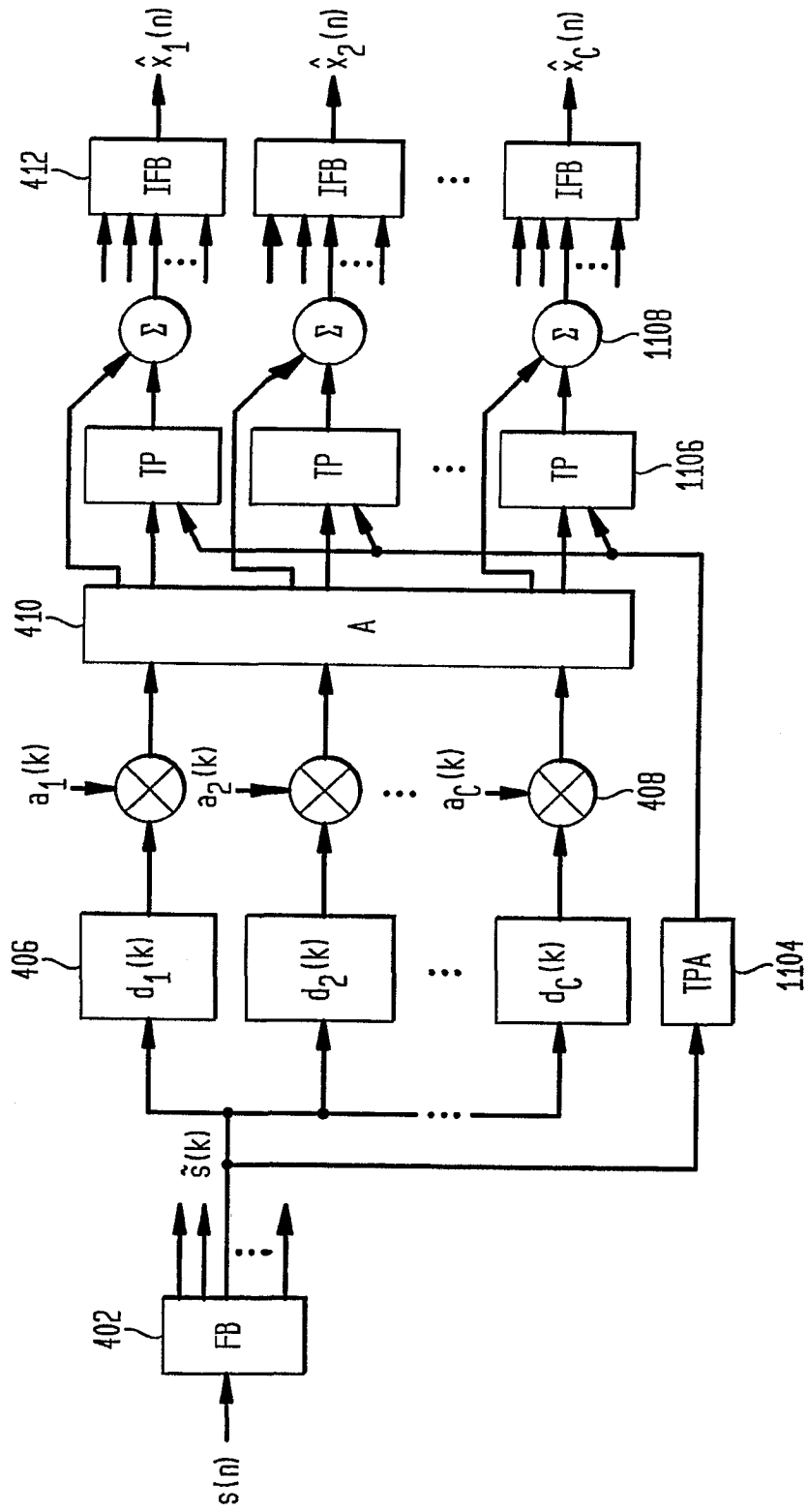


图 11(a)

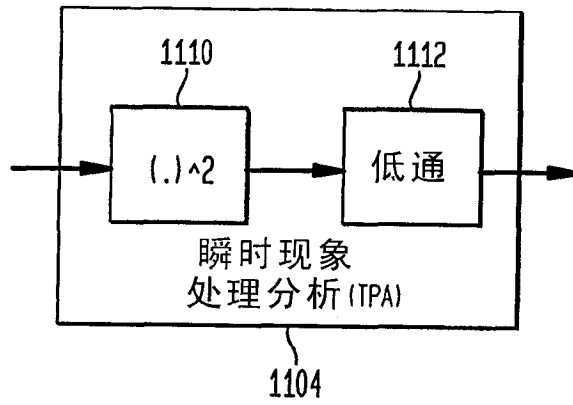


图 11(b)

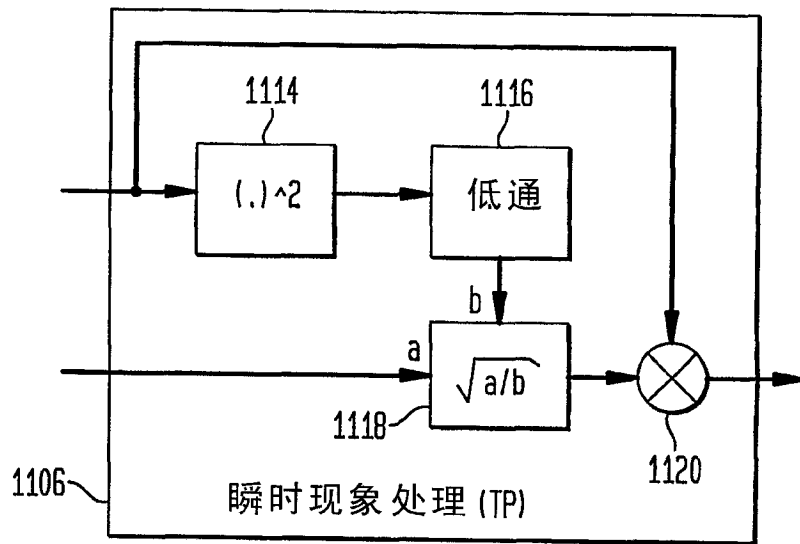


图 11(c)

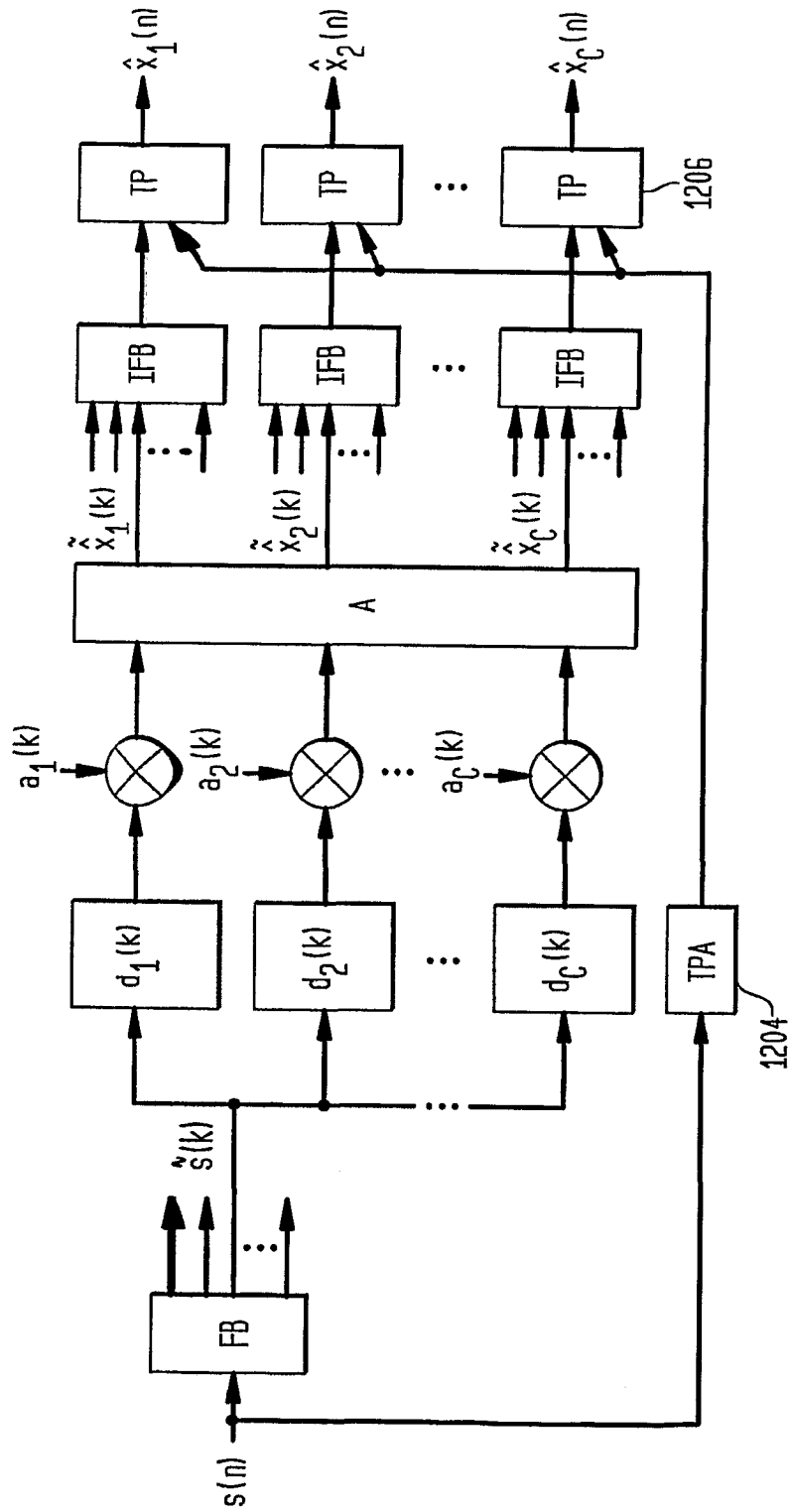


图 12(a)

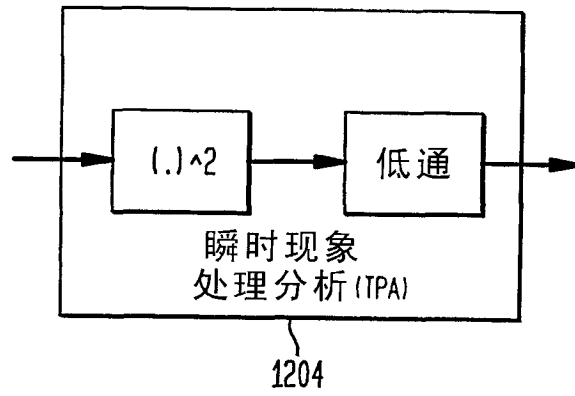


图 12(b)

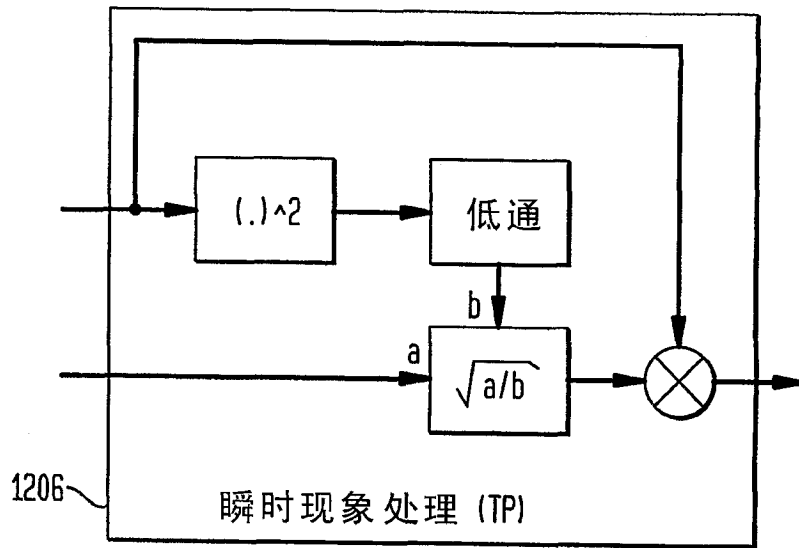


图 12(c)

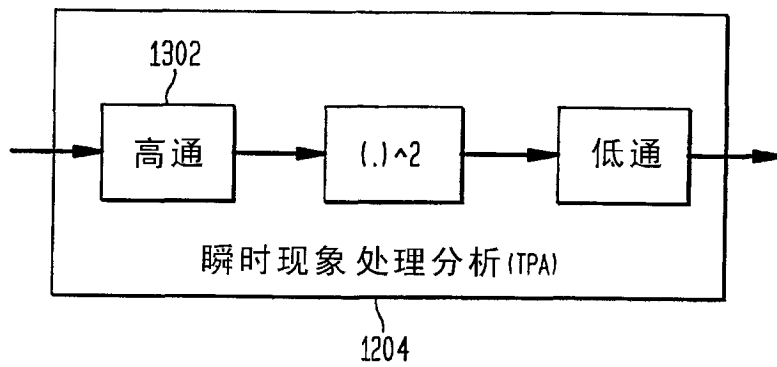


图 13(a)

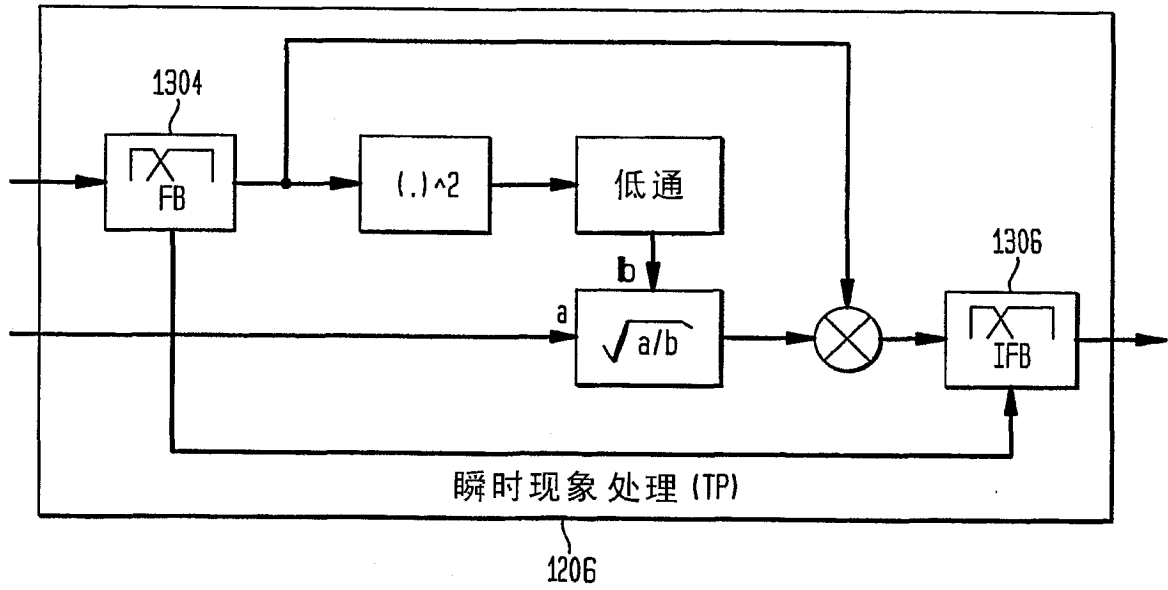


图 13(b)

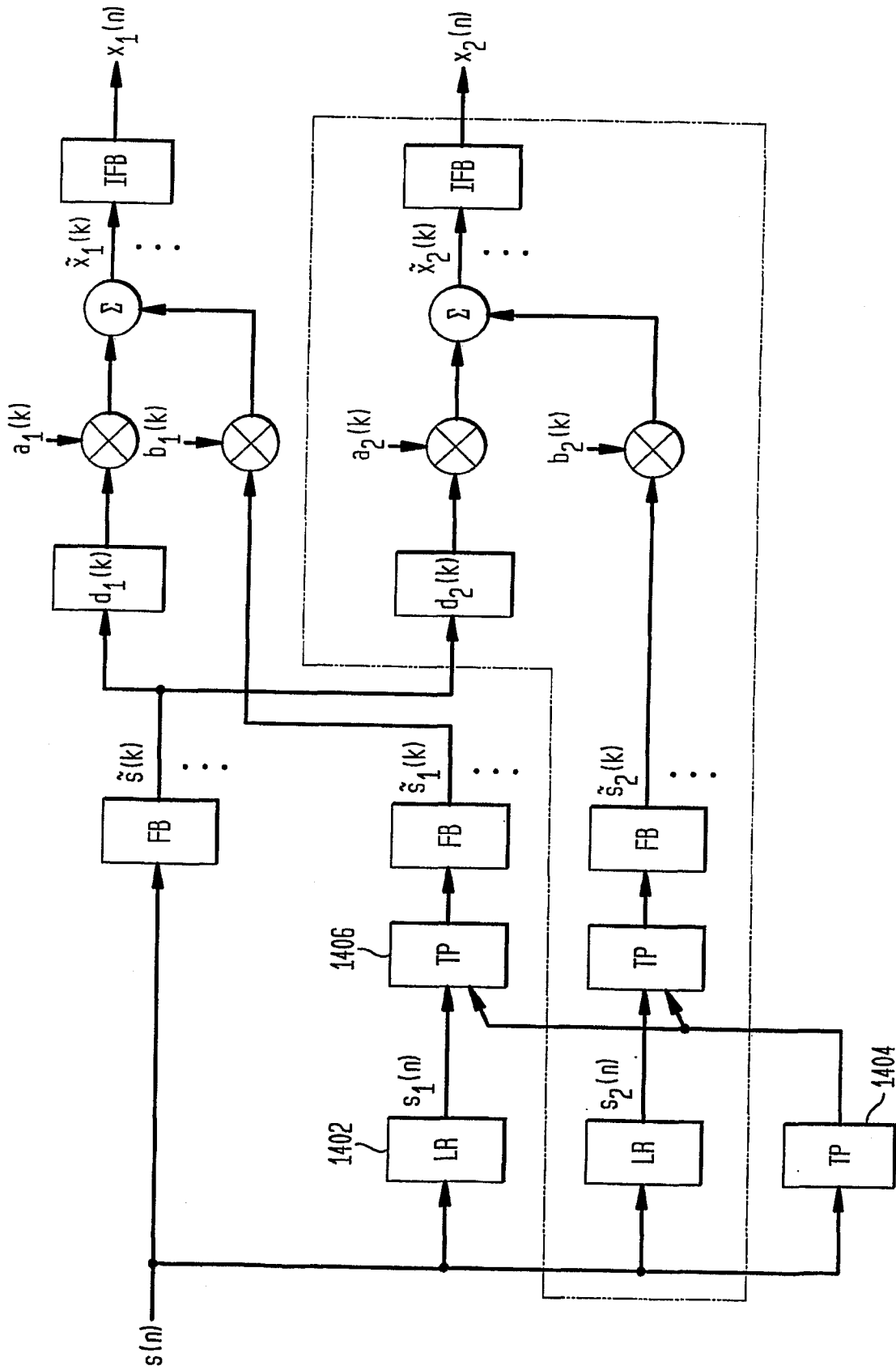


图 14

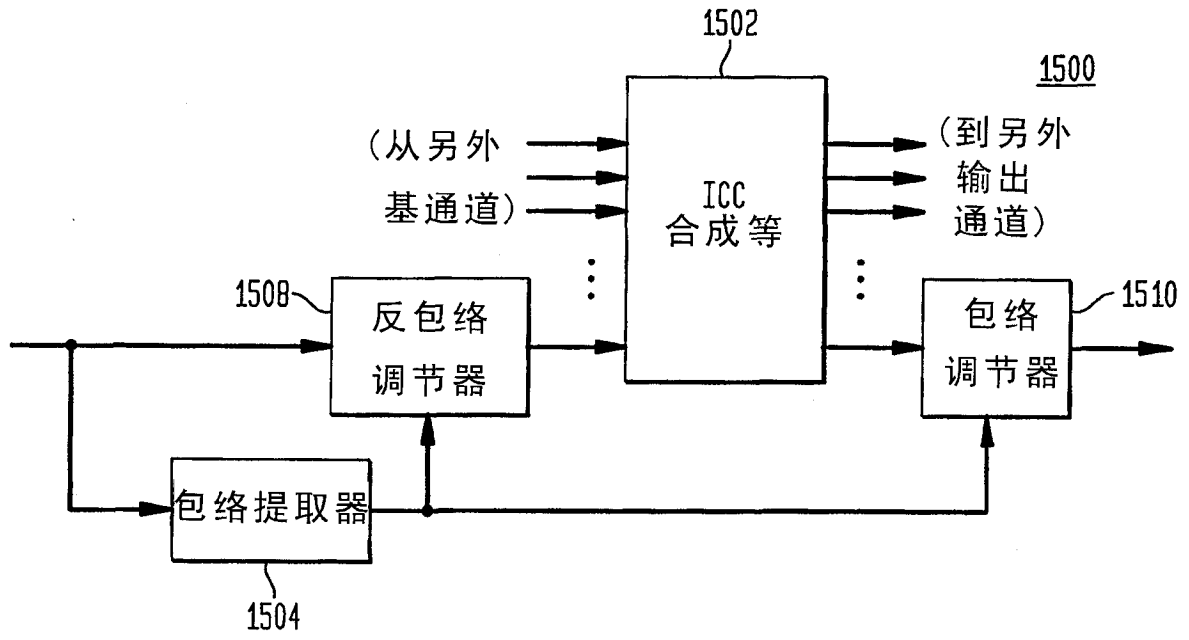


图 15

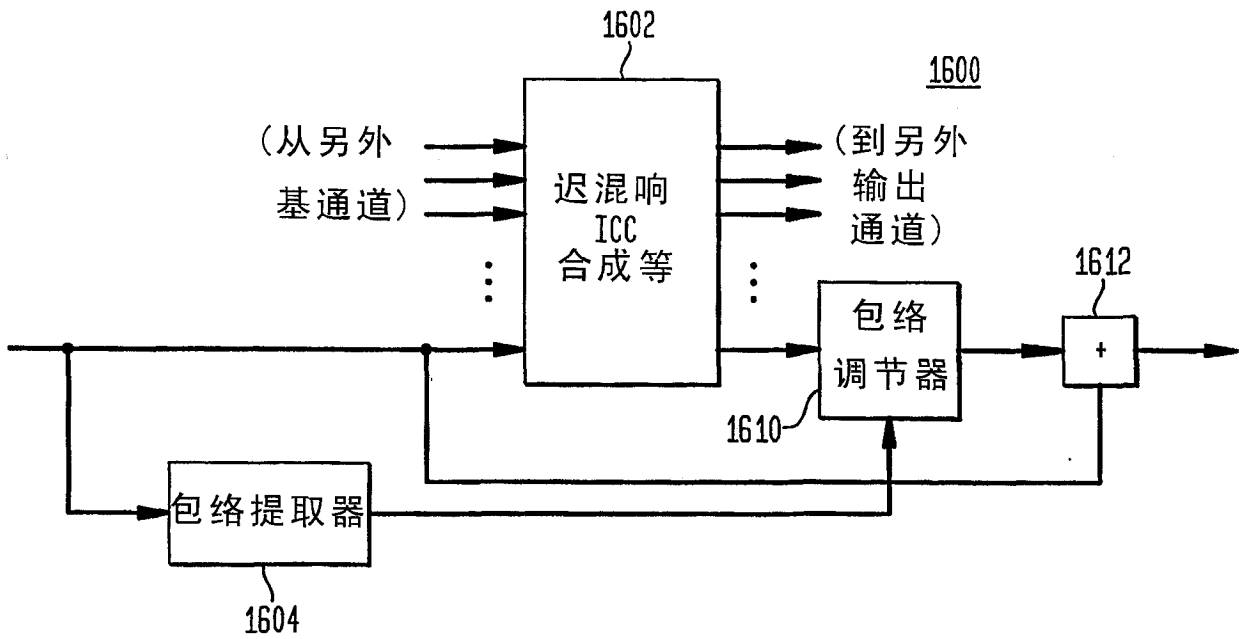


图 16

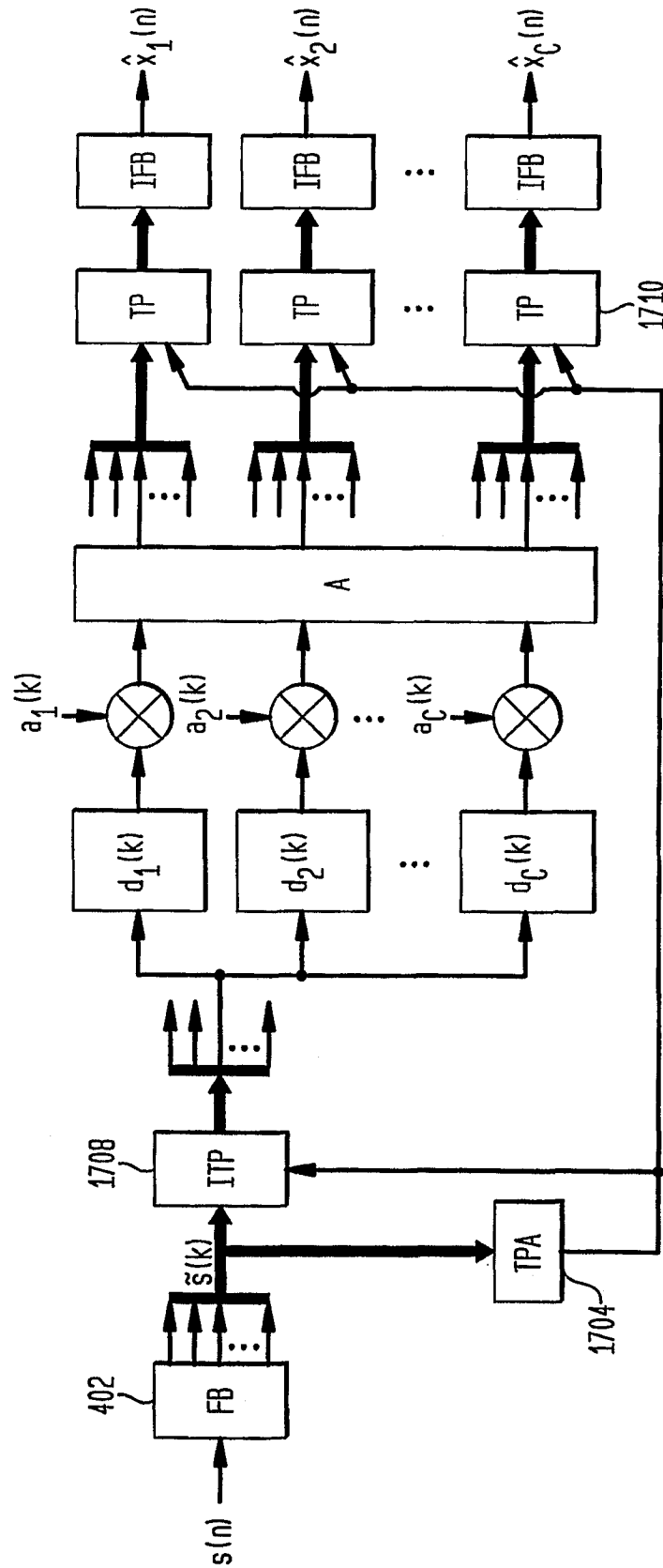


图 17

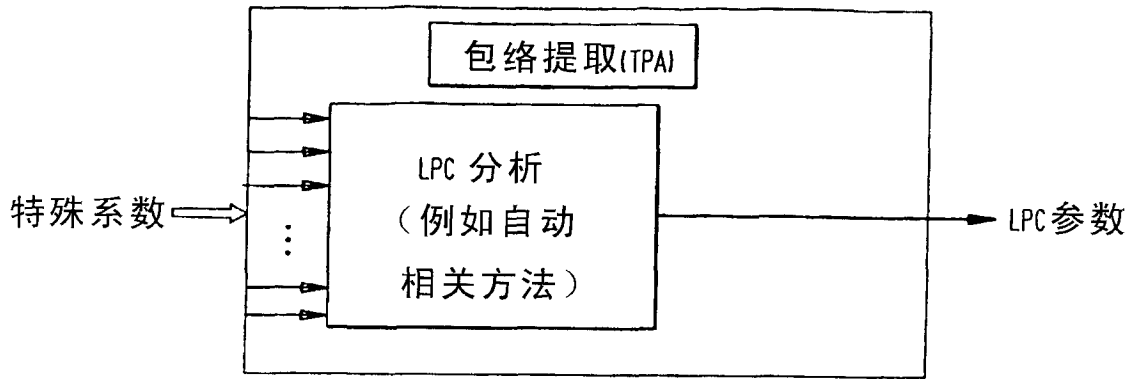


图 18(a)

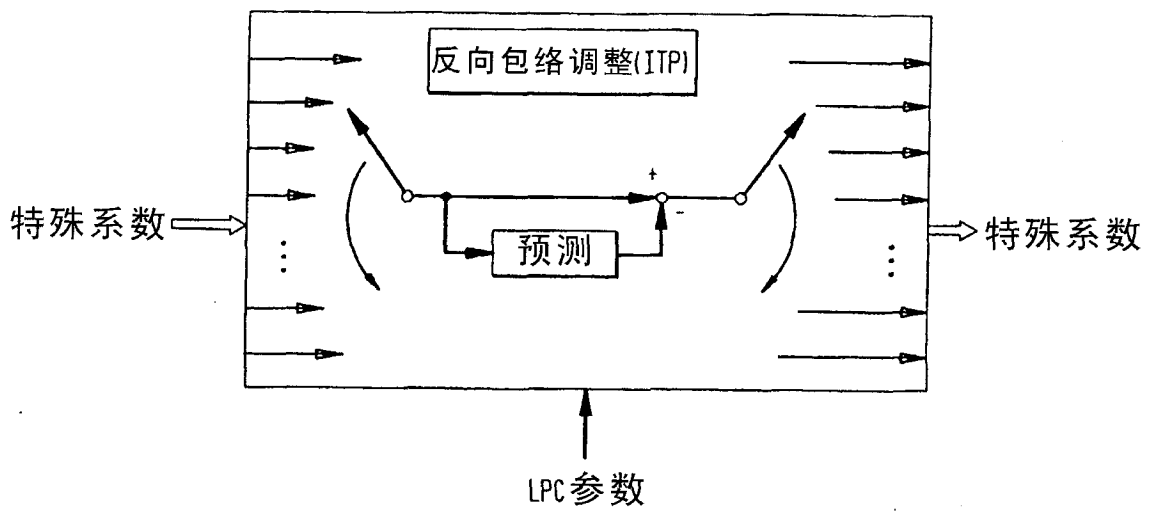


图 18(b)

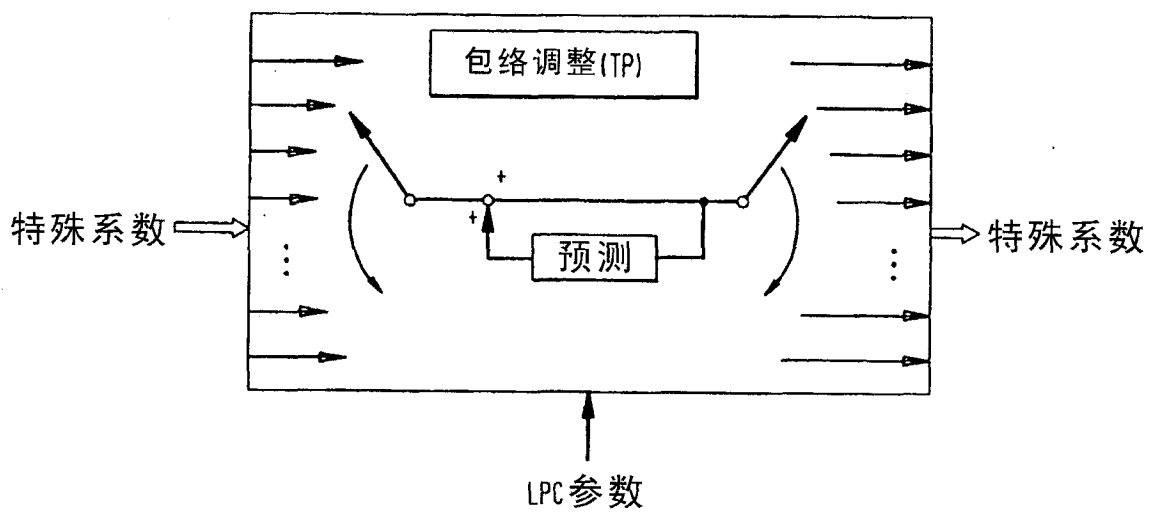


图 18(c)