

[19] 中华人民共和国国家知识产权局



[12] 发明专利说明书

[51] Int. Cl.  
*G10L 19/02 (2006.01)*  
*H04B 1/66 (2006.01)*

专利号 ZL 03807983.6

[45] 授权公告日 2009 年 7 月 1 日

[11] 授权公告号 CN 100508026C

[22] 申请日 2003.3.20 [21] 申请号 03807983.6

[30] 优先权

[32] 2002.4.10 [33] EP [31] 02076410.6

[86] 国际申请 PCT/IB2003/001152 2003.3.20

[87] 国际公布 WO2003/085643 英 2003.10.16

[85] 进入国家阶段日期 2004.10.9

[73] 专利权人 皇家飞利浦电子股份有限公司

地址 荷兰艾恩德霍芬

[72] 发明人 R·M·阿尔特斯 R·艾旺

[56] 参考文献

US56+21855A 1997.4.15

US5511093A 1996.4.23

US6121904A 2000.9.19

SUBBAND CODING OF STEREOPHONIC DIGITAL AUDIOSIGNALS. VAN DER WAAL R G ET AL. SPEECH PROCESSING 2, VLSI, UNDERWATER SIGNAL PROCESSING., Vol. 2 . 1991

审查员 杨艳兰

[74] 专利代理机构 中国专利代理(香港)有限公司

代理人 程天正 梁永

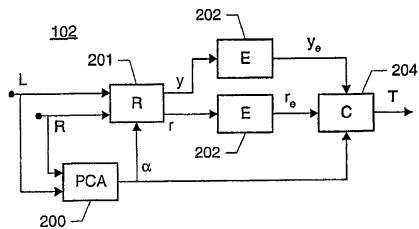
权利要求书 2 页 说明书 15 页 附图 6 页

[54] 发明名称

立体声信号编码

[57] 摘要

所公开的是一种编码至少包括第一信号部分(L)和第二信号部分(R)的多通道信号例如立体声音频信号的方法。本方法包括通过预定变换将至少第一和第二信号部分变换成包含大部分信号能量的主信号(y)和比主信号包含更少能量的至少一个残余信号(r)，通过至少一个变换参数对预定变换进行参数化；和至少用主信号和变换参数来表示多通道信号。还公开了一种用于编码多通道信号的相应的装置和一种用于解码这种信号的相应的方法和装置。



1. 一种编码包括至少第一信号部分和第二信号部分的多通道信号的方法，所述第一信号部分和第二信号部分是时域信号部分，本方法包括步骤
  - 通过预定变换将至少第一和第二信号部分在时域变成包含大部分信号能量的主信号和比主信号包含更少的能量的至少一个残余信号，通过至少一个变换参数来参数化预定变换；和
    - 至少用主信号和变换参数来表示多通道信号。
2. 根据权利要求 1 的方法，其中该方法进一步包括基于至少第一和第二部分自适应地确定变换参数的步骤。
3. 根据权利要求 1 或 2 的方法，其中主信号对应于第一和第二信号部分的主要部分。
4. 根据权利要求 1—3 中的任何一个的方法，其中预定变换是旋转，变换参数对应于旋转角度。
5. 根据权利要求 1—4 中的任何一个的方法，其中至少用主信号和变换参数来表示多通道信号的步骤进一步包括用主信号、变换参数和残余信号来表示多通道信号的步骤。
6. 根据权利要求 5 的方法，其中用主信号、变换参数和残余信号来表示多通道信号的步骤进一步包括步骤
  - 以第一比特率编码主信号；和
  - 以小于第一比特率的第二比特率编码残余信号。
7. 根据权利要求 1—6 中任何一个的方法，其特征在于主信号对应于第一信号能量，残余信号对应于小于第一信号能量的第二信号能量。
8. 根据权利要求 1—7 中任何一个的方法，其中
  - 该方法进一步包括使用对应于一组滤波参数的预测滤波器从主信号中估计残余信号的步骤；和
  - 至少用主信号和变换参数来表示多通道信号的步骤包括用主信号、变换参数和该组滤波参数来表示多通道信号的步骤。
9. 根据权利要求 1—8 中任何一个的方法，其中多通道信号包括一个具有左和右信号部分的立体声信号。
10. 一种解码多通道信号信息的方法，该方法包括步骤

一 接收主信号和变换参数，主信号对应于多通道源信号的至少第一和第二时域信号部分的预定变换的结果，通过至少该变换参数对预定变换进行参数化；和

一 通过在时域逆变换所接收的主信号和至少一个残余信号来产生第一和第二解码时域信号。

11. 根据权利要求 10 的方法，其中接收主信号和变换参数的步骤进一步包括接收残余信号的步骤。

12. 根据权利要求 10 或 11 的方法，其中接收主信号和变换参数的步骤进一步包括接收一组滤波器参数的步骤，和

该方法进一步包括使用对应于所接收的这组滤波参数的预测滤波器从主信号中预测残余信号的步骤。

13. 一种用于编码包括至少第一信号部分和第二信号部分的多通道信号的装置，所述第一信号部分和第二信号部分是时域信号部分，该装置包括

一 用于通过预定变换将至少第一和第二信号部分在时域转换成包含大部分信号能量的主信号和比主信号包含更少能量的至少一个残余信号的第一处理设备，通过至少一个变换参数对预定变换进行参数化；和

一 用于至少通过主信号和变换参数来表示多通道信号的第二处理设备。

14. 一种用于解码多通道信号信息的装置，该装置包括

一 用于接收主信号和变换参数的接收设备，主信号对应于第一和第二时域多通道源信号的预定变换的结果，通过至少该变换参数对预定变换进行参数化；和

一 用于通过在时域逆变换所接收的主信号和残余信号来产生第一和第二时域多通道信号的处理设备。

15. 一种用于传输包括至少第一信号部分和第二信号部分的多通道信号的设备，所述第一信号部分和第二信号部分是时域信号部分，该设备包括一种用于编码多通道信号的装置，该装置包括

一 用于通过预定变换将至少第一和第二信号部分在时域转换成包含大部分信号能量的主信号和比主信号包含更少能量的至少一个残余信号的第一处理设备，通过至少一个变换参数对预定变换进行参数化；和

一 用于至少通过主信号和变换参数来表示多通道信号的第二处理设备。

## 立体声信号编码

本发明涉及对包括至少第一和第二信号部分的多通道信号的编码。更特别的，本发明涉及多声道音频信号例如立体声信号的编码。

立体声音频信号包括来自立体声信号源例如分离的麦克风的左（L）信号部分和右（R）信号部分。音频信号编码的目的是降低立体声信号的比特率，例如为了实现声音信号通过通信网络例如因特网、调制解调器和模拟电话线、移动信道或者其他无线网络等等的高效传输，以及在智能卡或者其他具有有限存储容量的存储介质上存储立体声信号。

美国专利 4589127 披露了一个用于立体声信号的发射机，其产生立体声 L 和 R 信号的和与差信号。该和与差信号随后用于调制方案以得到同时包括和与差信息的调制信号。

然而，上述现有技术方法没有解决以有效的比特利用率对立体声信号进行编码的问题，即，对于给定声音质量使用低比特率对立体声信号编码。

通过一种编码包括至少第一信号部分和第二信号部分的多通道信号的方法来解决以上和其他问题，本方法包括步骤

- 通过预定变换将至少第一和第二信号部分变换成包含大部分信号能量的主信号和比主信号包含更少能量的至少一个残余信号，通过至少一个变换参数来参数化预定变换；和
- 至少用主信号和变换参数来表示多通道信号。

从而，通过将多通道信号变换成包含大部分信号能量的主信号和仅包含极少能量的残余信号，可以将多通道信号表示成主信号、变换参数和可选的小残余信号，从而提高多通道信号的编码效率。实际上，多通道信号可以以仅比单一通道（例如单通道）略高的比特率来编码。所得的编码信号可以被存储和 / 或传送到接收器。

当本方法进一步包括基于至少第一和第二信号部分来自适应地确定变换参

数的步骤时，可以连续追踪最优的变换参数，从而保证即便输入信号的特征发生改变，例如在音频信号的例子中由于声源的移动或者环境声学特性的变化而致，仍然能够保持变换是最优。

当预定变换是旋转并且变换参数对应于旋转角度时，仅基于一个单个参数旋转角度就能提供一个简单的变换。通过调整该角度，使得信号部分（即立体声信号的L和R信号部分）被旋转成主成分信号和残余信号，这样在保持高质量的同时提供了一种高效编码。

在本发明的一个优选实施例中，至少用主信号和变换参数来表示多通道信号的步骤进一步包括用主信号、变换参数和残余信号来表示多通道信号的步骤，从而由于没有丢弃任何信号信息而进一步改善了编码信号的质量。因为残余信号相比主信号要小，所以可以在这些信号之间对比特分配进行折衷。另外，因为比特率分配是可变的，例如通过自适应地增加或减少用于残余信号的比特率，可以提供一种完美的降低技术。

因此，在一个进一步的优选实施例中，用主信号、变换参数和残余信号来表示多通道信号的步骤进一步包括步骤

- 以第一比特率编码主信号；和
- 以比第一比特率小的第二比特率编码残余信号。

在另一优选实施例中，

- 本方法进一步包括使用相应于一组滤波参数的预测滤波器，从主信号估计残余信号的步骤；和
- 至少用主信号和变换参数来表示多通道信号的步骤包括用主信号、变换参数和该组滤波参数表示多通道信号的步骤。

本发明的实施例所基于的是，认为对于许多多通道信号来说，例如在音乐音频信号和语音信号的情况下，残余信号可以作为主信号的滤波版本而被准确估计。因而，当确定对残余信号建模的自适应滤波器的一组滤波参数时，滤波参数可以和主信号以及变换参数一起被编码。因此，就避免了残余信号的传输，不会丢失信号中包含的信息，从而提供了一种保留高水平质量的高效编码。

本发明的优点在于提供一种高效的比特率使用，即一个对于给定声音质量使用低速率的编码方案。这个根据本发明的编码方案可以用于在不显著降低声音质量的情况下降低比特率，在提高声音质量的同时保持比特率，或者两者的

结合。

本发明进一步涉及一种解码多通道信号信息的方法，方法包括步骤

- 接收主信号和变换参数，主信号对应于多通道源信号的至少第一和第二信号部分的预定变换的结果，至少通过变换参数将预定变换参数化；和
- 通过逆变换所接收的主信号和至少一个残余信号来产生第一和第二解码信号部分。

本发明可以通过不同的方式来实现，包括上述的方法以及下列的，编码和解码多通道信号的装置，分别地可以是数据信号，和进一步的产品设备，每个都具有所述与第一种提及的方法有关的一个或多个好处和优点，以及每个都具有和所述与第一种提及的方法有关的优选实施例相应的一个或多个优选实施例，并且记载在从属权利要求中。

要注意的是，上面和以下所述的方法的特征可以用软件来实现并且在数据处理系统或者其他通过计算机可执行指令的执行来激活的处理设备中执行。这些指令可以以程序代码方式从存储介质或者通过计算机网络从另一台计算机载入例如 RAM 的存储器中。可选地，所述特征可以通过硬连线电路代替软件实现或与软件相结合来实现。

本发明进一步涉及一种用于编码包括至少第一信号部分和第二信号部分的多通道信号的装置，该装置包括

- 适于通过预定变换将至少第一和第二信号部分变换成包含大部分信号能量的主信号和比主信号包含更少能量的至少一个残余信号的第一处理设备，通过至少一个变换参数将预定变换参数化；和
- 适于至少用主信号和变换参数来表示多通道信号的第二处理设备。

本发明进一步涉及一种用于解码多通道信号信息的装置，该装置包括

- 用于接收主信号和变换参数的接收设备，主信号对应于第一和第二多通道源信号的预定变换结果，通过至少一个变换参数将预定变换参数化；和
- 用于通过逆变换所接收的主信号和残余信号来产生第一和第二多通道信号的处理设备。

上述装置可以是任何包括计算机的电子设备的一部分，例如固定和便携式 PC，固定和便携式无线电通讯设备，和其他手持或便携式装置例如移动电话，呼机，音频播放器，多媒体播放机，通信器即电子组织器，智能电话，个人数

字助理 (PDA)，掌上电脑等等。

术语处理设备包括通用或特殊用途的可编程微处理器，数字信号处理器 (DSP)，专用集成电路 (ASIC)，可编程逻辑阵列 (PLA)，现场可编程门阵列 (FPGA)，特定用途的电子电路等，或者它们的组合。上面的第一和第二处理设备可以是分离的处理设备，或者它们也可以包括在一个处理设备中。

术语接收设备包含适于能够例如通过有线或无线数据链路进行数据传输的电路和 / 或装置。这种接收设备的例子包括网络接口，网卡，无线电接收器，用于其他合适的电磁信号例如红外线的接收器例如通过 IrDa 端口，基于无线电的通信，例如通过蓝牙收发信机，等等。进一步的这种接收设备的例子包括线缆调制解调器，电话调制解调器，综合服务数字网 (ISDN) 适配器，数字用户线路 (DSL) 适配器，卫星收发信机，以太网适配器等等。

术语接收设备进一步包括其他用于接收数字信号的输入电路 / 装置，例如存储在计算机可读介质中的数字信号。这种接收设备的例子包括软盘驱动器，光盘驱动器，DVD 驱动器，或其他任何合适的磁盘驱动器、存储卡适配器、智能卡适配器等。

本发明进一步涉及一种包含多通道信号信息的数字信号，该数字信号由上下文所述的方法生成。该信号可以具体化为一种载波中的数字信号，例如由通信设备如上下文所述进行传输的数字信号。

本发明进一步涉及一种包括由上下文所述方法生成的表示多通道信号信息的数字记录的计算机可读介质。术语计算机可读介质包括磁带，光盘，数字化视频光盘 (DVD)，压缩光盘 (CD 或 CD-ROM)，小型磁盘，硬盘，软盘，铁电存储器，电可擦除可编程只读存储器 (EEPROM)，闪存，EPROM，只读存储器 (ROM)，静态随机存取存储器 (SRAM)，动态随机存取存储器 (DRAM)，同步动态随机存取存储器 (SDRAM)，铁磁存储器，光学存储器，电荷耦合装置，智能卡，PCMCIA 卡等。

本发明进一步涉及一种传输包括至少第一信号部分和第二信号部分的多通道信号的装置，该装置包括如上下文所述编码多通道信号的设备。

根据参照实施例和参照附图的说明，本发明的这些和其他方面会更明白，其中：

图 1 显示了根据本发明的实施例用于传输立体声信号的系统的示意图；

图 2 显示了根据本发明的第一实施例用于编码立体声信号的装置的示意图；

图 3 说明了根据本发明的实施例的信号变换的确定；

图 4 显示了根据本发明的第一实施例用于解码立体声信号的装置的示意图；

图 5 显示了根据本发明的第二实施例用于编码立体声信号的装置的示意图；

图 6 显示了根据本发明的第二实施例用于解码立体声信号的装置的示意图；

图 7a-c 显示了在本发明的实施例中使用的滤波电路的示例的示意图；

图 8 显示了根据本发明的第三实施例用于编码立体声信号的装置的示意图；

图 9 显示了根据本发明的第四实施例用于编码立体声信号的装置的示意图；

图 10 显示了根据本发明的第四实施例用于解码立体声信号的装置的示意图；

图 11 显示了根据本发明的第五实施例用于编码多通道信号的装置的示意图；

图 12 显示了根据本发明的第六实施例用于编码多通道信号的装置的示意图；

图 13 显示了本发明的实施例中使用的减法电路的示意图。

图 1 显示了根据本发明的实施例用于传输立体声信号的系统的示意图。此系统包括用于生成编码立体声信号的编码装置 101 和用于将接收到的编码信号解码成立体声 L 信号和立体声 R 信号部分的解码装置 105。编码装置 101 和解码装置 105 中的每一个都可以是任何电子设备，或是此类设备的一部分。这里术语电子设备包括计算机，例如固定和便携式 PC，固定和便携式无线电通讯设备，和其他手持或便携式装置，例如移动电话，呼机，音频播放器，多媒体播放机，通信器即电子组织器，智能电话，个人数字助理（PDA），掌上电脑

等等。要注意的是编码装置 101 和解码装置 105 可以组合在一个电子设备中，其中立体声信号存储在计算机可读介质中用于以后再现。

编码装置 101 包括用于根据本发明对立体声信号进行编码的编码器 102，立体声信号包含 L 信号部分和 R 信号部分。编码器接收 L 和 R 信号部分并生成编码信号 T。立体声信号 L 和 R 可以来自一组麦克风，例如通过其它电子设备如混频装置等。该信号可以进一步作为其他立体声播放机的输出、无线广播的无线电信号或通过任何其他合适的方式而被接收。以下将说明根据本发明的这种编码器的优选实施例。根据一个实施例，编码器 102 连接到发射机 103，通过信道 109 将编码信号 T 发射到解码装置 105。发射机 103 可以包括适于例如通过有线或无线数据链路 109 进行数据传输的电路。这种发射机 103 的例子包括网络接口，网卡，无线电发射机，用于其他合适的电磁信号的发射机，例如通过 IrDa 端口发射红外线的 LED，基于无线电的通信，例如通过蓝牙收发信机，等等。合适的发射机的进一步的例子包括线缆调制解调器，电话调制解调器，综合服务数字网（ISDN）适配器，数字用户线路（DSL）适配器，卫星收发报机，以太网适配器等等。相应地，信道 109 可以是任何合适的有线或无线数据链路，例如基于分组的通信网络，如因特网或其他的 TCP/IP 网络，短距离通信链路如红外链路、蓝牙连接或其他基于无线电的链路。这种信道的进一步的例子包括计算机网络和无线通信网络，例如蜂窝数字分组数据（CDPD）网络，全球移动系统（GSM）网络，码分多址（CDMA）网络，时分多址网络（TDMA），通用分组无线电服务（GPRS）网络，第三代网络如 UMTS 网络，等等。可选的或者另外的，编码装置可以包括一个或多个其他接口 104，用于将编码立体声信号 T 传输到解码装置 105。这种接口的例子包括用于把在计算机可读介质 110 上存储数据的磁盘驱动器，例如软盘驱动器，读 / 写 CD-ROM 驱动器，DVD 驱动器等。其他例子包括存储卡插槽，磁卡读 / 写器，访问智能卡的接口等。相应地，解码装置 105 包括对应的用于接收由发射机发射的信号的接收器 108 和 / 或用于通过接口 104 和计算机可读介质 110 接收编码立体声信号的其他接口 106。解码装置进一步包括解码器 107，接收所接收的信号 T 并将其解码成相应的立体声部分 L' 和 R'。下面将对根据本发明的这种解码器的优选实施例加以说明。解码的信号 L' 和 R' 随后可以被提供给立体声播放器，以便通过一组扬声器、耳机等来再现。

图 2 显示了根据本发明的第一实施例用于编码立体声信号的装置 102 的示意图。本装置包括用于执行立体声信号在 L-R 空间的 $\alpha$ 角度旋转的电路 201，根据变换

$$\begin{aligned} y &= L \cos\alpha + R \sin\alpha = w_L L + w_R R \\ r &= -L \sin\alpha + R \cos\alpha = -w_R L + w_L R, \end{aligned} \quad (1)$$

生成旋转信号部分  $y$  和  $r$ ，其中使用  $w_L = \cos\alpha$  和  $w_R = \sin\alpha$  作为加权因子。

根据本发明，确定角度  $\alpha$ ，以使其与高信号方差的方向对应。最大信号方差的方向（即主要部分）可以通过主成分分析进行估计，从而使旋转的  $y$  部分对应包含大部分信号能量的主成分信号，而  $r$  是残余信号。相应地，图 2 的装置包括确定角度  $\alpha$  或者可选的确定加权因子  $w_L$  和  $w_R$  的电路 200。

参照图 3，根据一个优选实施例，上述加权因子  $w_L$  和  $w_R$  根据以下算法确定：

首先，对输入的立体声信号  $L$  和  $R$  进行修正和低通滤波，以分别产生  $L$  和  $R$  的包络信号  $p(k)$  和  $q(k)$ ，其中  $p(k)$  和  $q(k)$  被适当地采样并且用  $k$  表示采样索引。这样，矢量  $x(k) = (p(k), q(k))$  就表示输入的信号矢量。可选地，可以直接使用信号  $L$  和  $R$ ，即不经过滤波，或者可以使用  $L$  和  $R$  的其他滤波版本例如高通滤波信号  $L$  和  $R$ 。在图 3 中用圆圈显示了多个信号点。作为举例，显示了信号点  $x(k)$  及其相应部分  $p(k)$  和  $q(k)$ 。根据本发明，在信号矢量的主成分方向上旋转信号。在图 3 的示例中，这对应于  $y$  方向，其中  $\alpha$  是  $y$  方向和  $p$  方向之间的夹角。加权矢量  $w = (w_L, w_R)$  表示主成分的方向，而  $x(k)$  的旋转部分分别用  $y(k)$  和  $r(k)$  来表示。

主成分可以通过本领域已知的任何合适的方法确定。在一个特别有利的实施例中，使用了利用 Oja 规则（参见例如 S. Haykin: “Neural Networks”，Prentice Hall, N. J., 1999）的迭代方法。根据本实施例，加权矢量  $w$  根据以下等式迭代估计

$$w(k) = w(k-1) + \mu [x(k-1) - w(k-1) y(k-1)], \quad (2)$$

其中  $w(k) = (w_L(k), w_R(k))$  对应于在时间  $k$  的估计。以上迭代可以例如用一组小的随机权值  $w(0)$  初始化，或者通过其他适当的方式。以上估计的加权矢量可以用来根据  $y(k) = w^T(k)x(k)$  计算旋转信号。可选地，等式(2)的迭代可以基于块执行，例如对于一个  $N$  个采样的块，其中  $N$  依赖于特定实现，例如  $N=512$ ，

1024, 2048 等。在本实施例中，用于块的估计加权因子  $w(N)$  可以根据  $y(k) = w^T(N)x(k)$  用于该块的所有采样的变换中。

等式(2)中的因子  $\mu$  对应于追踪算法的时间标度。如果  $\mu=0$ , 加权因子以及角度  $\alpha$  保持不变, 而对于较大的  $\mu$  它们变化得很快。作为举例, 对于一个 2048 采样的块,  $\mu$  可以选择为  $10^3$  的数量级以用于 44.1 kHz 的采样速率。

上述迭代算法的优势在于它是线性的, 即它不需要计算任何三角函数、平方根等。上述算法的进一步的优势在于它获得一个归一化的加权矢量  $w$ , 因为等式(2)中的  $-\mu w(k-1) y(k-1)$  项对应于一个使较大权值衰减的加权衰减项, 而  $+\mu x(k-1)$  项在主成分方向上驱动加权矢量。进一步要注意的是, 在当前实施例中, 由于  $x(k)$  是包络信号,  $w_L, w_R \in [0, 1]$ , 即加权向量  $w$  存在于图 3 中的第一象限, 从而保证了  $\mu$  是正值。本发明的进一步优势在于, 它能够传递  $w_L$  和  $w_R$  之一, 而根据  $w_R = \sqrt{1 - (w_L)^2}$  来确定另一个因子。可选地, 可以传递角度  $\alpha$ 。

再参照图 2, 电路 200 输出确定的角度  $\alpha$  或者可选的, 加权因子  $w_L$  和  $w_R$  的一个或全部。将角度信息提供给生成旋转信号部分  $y$  和  $r$  的旋转电路 201。可以理解, 电路 200 和 201 可以组合成单个的电路, 来执行等式(2)的迭代运算和根据等式(1)的  $y$  和  $r$  的计算。

该装置进一步包括分别执行信号  $y$  和  $r$  的适当编码的编码器 202。例如, 信号可以根据 MPEG 例如 MPEG I 层 3(MP3)进行编码, 根据正弦编码(SSC), 或者基于子带、参数或变换方案的音频编码方案, 或者其他任何合适的方案或其组合来编码。可以理解的, 编码器 202 可以是相同或者不同的类型, 例如一个 MP3 编码器和一个 SSC 编码器等。所得的编码信号  $y_e$  和  $r_e$  分别与角度信息  $\alpha$  一起被提供给组合电路 204。组合电路 204 执行成帧、比特率分配和无损编码, 以得到要传输的组合信号  $T$ 。在一个实施例中, 角度  $\alpha$  或者可选的  $w_L$  和 / 或  $w_R$  可以作为在信号帧、信号块等之前传输的包头的一部分传输。根据本发明, 由于对变换角度  $\alpha$  进行追踪, 以使主成分信号包含信号能量的大部分, 所以也可以选择分配给  $y$  和  $r$  信号的不同的比特率, 从而优化编码效率。

例如, 立体声信号  $L$  和  $R$  可以表示为  $L=M+S$  和  $R=M-S$ , 其中  $M$  对应于一个中间或中央信号,  $S$  对应于立体声或边(side signal)信号。在对固定声源进行声学记录的情形下, 例如一个由两个麦克风记录的讲话者,  $L$  和  $R$  信号基本相同, 如果讲话者恰好位于麦克风之间并且假定没有例如反射等声学失真。

因此，在这种情况下  $S$  基本为零或者至少很小，并且根据本实施例的编码方案基本上输出对应  $L+R$  的  $y$ ，和与  $L-R$  对应为零或者很小的  $r$ ；这对应于  $\alpha = 45$  度。如果讲话者不是恰好位于麦克风之间，即是不对称的，但是仍假定没有反射或其他失真，那么根据本发明的旋转信号  $y$  仍然对应于讲话者而残余信号  $r$  基本为零。不过这种情况下角度  $\alpha$  不再是 45 度。如果声源移动了，例如从左向右，根据本发明的方法仍然得到一个对应于声源的主成分信号  $y$  和小的残余信号  $r$ ，理想情况下  $r=0$ 。在这种情形， $\alpha$  从 0（最左）变化到 90 度（最右）。

以上例子说明了追踪角度  $\alpha$  的优点。在上面的情况下，它能够传输主成分信号  $y$  和角度  $\alpha$ ，以使解码器在不显著损失质量的情况下重建原始信号  $R$  和  $L$ 。一般地，由于残余信号  $r$  比主成分信号小，所以  $y$  和  $r$  之间的比特分配和编码效率是可以协调的。因此，本发明的优点在于提供一种立体声信号的高效编码。

图 4 显示了根据本发明的第一实施例用于解码立体声信号的装置 107 的示意图。本装置接收例如来自根据结合图 2 所述实施例的编码器的编码立体声信号  $T$ 。本装置包括用于从组合信号  $T$  中提取编码信号  $y_e$  和  $r_e$  以及角度信息  $\alpha$  的电路 404，即电路 404 执行图 2 中组合器 204 的逆运算。提取的信号  $y_e$  和  $r_e$  被提供给解码器 402，用于相应于图 2 中编码器 202 执行的编码来执行音频解码，以生成解码的主成分信号  $y'$  和解码残余信号  $r'$ 。信号  $y'$ 、 $r'$  和角度信息  $\alpha$  被提供给旋转电路 401，其将信号  $y'$ 、 $r'$  旋转回原始的  $L$  和  $R$  部分的方向，从而得到接收的信号  $L'$  和  $R'$ 。

图 5 显示了根据本发明的第二实施例用于编码立体声信号的装置的示意图。本装置包括电路 201，对立体声信号  $L$  和  $R$  执行  $\alpha$  角度的旋转以得到旋转信号部分  $y$  和  $r$ ，如结合图 2 所述。本装置进一步包括如结合图 2 和 3 所述的用于确定旋转角度的电路 200。根据本发明的这一实施例，认为残余信号  $r$  可以作为主成分信号  $y$  的滤波版本而被估计。如结合图 2 所述，在没有由于反射等引起的声学失真的情况下，在通过两个麦克风记录的一个音频源的声学记录中，主信号  $y$  对应于音频源而残余信号基本为零。然而，在更实际的情形下失真是存在的，例如由于信号在房间的墙上和讲话者的头上和身上的反射等。这些效果影响了残余信号  $r$ 。因此，当通过滤波器估计残余信号时，滤波器实际上是对房间声学等建模。对于一个古典管弦乐队情形是相似的，而对于现代流行音乐情形就有了一些不同。在这种情形下，录音师经常使用人工混响、效果

箱等典型地将多个通道混合成双通道。在这种情形，滤波器对由混合处理引进的声学效果建模。

仍然参照图 5，本装置包括接收主信号  $y$  作为输入并产生滤波信号  $\hat{r}$  的自适应滤波器 501。自适应滤波器的滤波参数  $F_p$  是通过例如根据由减法电路 502 生成的表示  $r$  和  $\hat{r}$  之间差别的误差信号  $e$  控制自适应滤波器 501 来选择的，以使得滤波信号  $\hat{r}$  近似残余信号  $r$ 。将所得的滤波参数  $F_p$  提供给组合电路 204，优选的在通过编码器 503 适当编码后，例如提供哈夫曼编码或其他任何合适的编码方案的编码器。滤波器 501 可以是本领域公知的任何合适的滤波器。这种滤波器的例子包括有限脉冲响应 (FIR) 滤波器或无线脉冲响应 (IIR) 滤波器，它们是自适应或者固定的，具有固定的或者递归追踪的截断频率和幅度，等等。该滤波器可以是任何阶的，优选的小于 10。滤波器类型可以是巴特沃思，切比雪夫，或者其他任何适当类型的滤波器。本装置进一步包括结合图 2 所述的用于编码主信号的编码器 202，从而得到编码的主信号  $y_e$ ，与滤波参数  $F_p$  以及角度信息  $\alpha$  一起被提供给组合电路 204。如结合图 2 所述，组合电路 204 执行成帧、比特率分配和无损编码，以得到要传输的组合信号  $T$ ，其包括主成分信号  $y_e$ 、滤波参数  $F_p$  和角度信息  $\alpha$ 。根据本发明的这个实施例，分配给滤波参数  $F_p$  的比特率可以显著小于主信号  $y$  所需的比特率，例如在一个实施例中，用于滤波参数  $F_p$  的比特率在平均上可以小于用于  $y$  的比特率的 10%。因而，本发明的优势在于它减少了传输立体声信号所需的比特率。根据本发明的总比特率仅比用于一个单通道的略高。然而，要注意的是，该速率在记录过程中会变化。例如，在几乎没有失真和固定声源的情形下该速率会变小。但在例如 L 和 R 信号瞬间独立时也会变大。

图 6 显示了根据本发明的第二实施例用于解码立体声信号的装置 107 的示意图。本装置接收例如来自根据结合图 5 所述实施例的编码器的编码立体声信号  $T$ 。本装置包括用于从组合信号  $T$  中提取编码信号  $y_e$ 、滤波参数  $F_{pe}$  和角度信息  $\alpha$  的电路 404，即电路 404 执行图 5 中组合器 204 的逆运算。将所提取的信号  $y_e$  提供给解码器 402，用于相应于图 5 中编码器 202 执行的编码来执行音频解码，以生成解码的主成分信号  $y'$ 。优选的，通过解码器 402 对滤波参数进行解码，所述解码器与图 5 中通过编码器 503 对滤波器参数的编码相对应。信号  $y'$  和解码的滤波参数  $F_p$  一起被提供给滤波器 601。滤波器 601 生成相应的估

计残余信号  $\hat{r}'$ 。所接收的主成分信号  $y'$ 、估计残余信号  $\hat{r}'$  和所接收的角度信息  $\alpha$  被提供给旋转电路 401，其将信号  $y'$ 、 $\hat{r}'$  旋转回原始的 L 和 R 部分的方向，从而得到接收的信号 L' 和 R'。

在结合图 5 和 6 所述的实施例中，滤波器 501 和 601 可以是瞬时或时域中的标准自适应滤波器（例如参见“Adaptive Filter Theory”，by S. Haykin, Prentice Hall, 2001），例如在回波消除领域公知的自适应滤波器。其他例子的滤波器包括具有固定或自适应的截断频率和幅度的固定的 FIR 或 IIR 滤波器。可选地，在一个实施例中，如在 MPEG 编码中所公知的，滤波器可以基于人的听觉系统的心理声学模型，从而减少了滤波参数的数目。而根据另一个实施例，滤波器可以进一步简化，例如使用 5 个双二阶滤波器的 10 阶滤波器和人工混响单元。在这一实施例中，在编码一侧配置滤波器和确定混响时间。这些参数缓慢地改变，从而减少它们在传输中所需的比特率。

图 7a-c 显示了用在本发明的实施例中的滤波电路的实例的示意图。

在图 7a 的例子中，滤波器 501 包括滤波器 701 和混响滤波器 702 的组合。例如，滤波器 701 可以是瞬时或时域中的标准自适应滤波器，具有固定或自适应的截断频率和幅度的固定的 FIR 或 IIR 滤波器等，例如高通滤波器。根据本实施例，滤波器 701 的滤波参数和混响滤波器 702 的参数（例如用  $T_{60}$  表示的混响时间），都作为滤波参数  $F_p$  而被传输到解码器。

在图 7b 的例子中，除了滤波器 701 和 702 之外，又增加了两个控制电路 703-704。增加控制电路 703 是为了保证残余信号  $r$  的平均功率和混响器 702 的输出的平均功率大致相同，例如通过用参数  $\beta_1$  与混响器 702 的输出相乘。第二控制电路 704 用  $\beta_2$  与混响器的缩放的输出相乘。因子  $\beta_2$  可以在从 -3dB 到 +6dB 之间的范围内选择并且通过使  $r$  和  $\hat{r}$  之间的互相关  $\rho$  尽可能高来确定，即，使信号  $r$  和  $\hat{r}$  尽可能地相似。因而，图 7b 的滤波装置进一步包括用于确定互相关  $\rho$  的电路 705。滤波装置进一步包括乘法器 706，用于产生乘积  $\beta = \beta_1 \cdot \beta_2$  并输出所述乘积作为滤波参数  $F_p$  的一部分。因此， $\beta_1$  是例如通过比较  $r$  和  $\hat{r}$  的绝对均值来自动控制的增益，而  $\beta_2$  是例如通过利用互相关系数  $\rho$  来自动控制的另一个增益。第一个增益目的是为了保证  $r$  的能量被保持，即是使接收器中预测信号  $\hat{r}'$  的能量对应于  $r$  的能量。第二个增益是为了保证  $r$  和  $\hat{r}'$  很好地相关。

在一个实施例中，混响器 702 和滤波器 701 可以被固定，即不根据滤波参

数  $F_p$  进行调整。进一步地， $\beta_2$  可以被固定，从而使缓慢变化的参数  $\beta_1$  作为唯一需要调整和传输的自适应参数。因此，提供了一个特别简单的滤波装置。本实施例的优点在于它仅需要大约原始立体声比特率的一半来传输立体声信号。要注意的是可以使用上述实施例的进一步改变。例如，在一个实施例中滤波器 701 可以省略。

另外，对于相关  $\rho$ ，可替换的或者附加的，可以使用其他的相关性度量来保证原始信号和经过编解码后的信号的高度相似。例如，在一个实施例中可以用两个相关器来代替相关器 705。一个相关器可以计算输入信号 L 和 R 之间的互相关  $\rho_{LR}$ ，另外，第二个相关器可以计算编—解码器产生的输出  $L'$  和  $R'$  之间的互相关  $\rho'_{LR}$ ，即根据本实施例，编码器进一步包括用于确定信号  $L'$  和  $R'$  的解码电路。本实施例使用差  $\epsilon_p = \rho_{LR} - \rho'_{LR}$  来控制  $\beta_2$  从而使  $\epsilon_p$  最小。这在图 7c 中说明，其中用既接收 L 和 R 也接收  $L'$  和  $R'$  作为输入并产生表示差  $\epsilon_p$  的信号作为输出的电路 707 来代替图 7b 的相关器。电路 707 的输出  $\epsilon_p$  控制电路 704 来对估计的残余  $\hat{r}$  进行缩放从而使  $\epsilon_p$  最小化。在一个实施例中，电路 707 的输入被高通滤波，例如在 250 Hz，从而使低频率对于  $\epsilon_p$  的影响降低。在如图 7b 的实施例中，本实施例的优点在于所得立体声映象和编解码之前的原始立体声映象之间的相关性非常高。

图 8 显示了根据本发明的第三实施例用于编码立体声信号的装置的示意图。本装置是结合图 5 所述的实施例的变形，包括用于执行立体声信号 L 和 R 的旋转的电路 201，用于确定旋转角度的电路 200，自适应滤波器 501，减法电路 502，编码器 202，编码器 503，和组合电路 204，如结合图 5 所述。根据本实施例，主成分信号  $y$  不直接提供给滤波器 501。替代地，本装置进一步包括如结合图 6 所述的解码器 402。解码器 402 接收由编码器 202 生成的编码主成分信号  $y_e$  并产生解码的主信号  $y'$ ，将其提供给滤波器 501。本实施例的优点在于减少了由信号  $y$  的编码和解码所引入的编码误差的影响。由于解码器 402 实际上不是编码器 202 的完全的逆，即  $E \cdot E^{-1} \neq 1$ ，因而这些编码误差使得解码信号  $y'$  与原始信号  $y$  略有不同。因此，通过在解码器应用信号  $y$  的编码和解码，滤波器 501 的输入  $y'$  对应于在接收器提供给滤波器 601 的输入  $y'$ ，从而改善了在接收器处残余信号的  $\hat{r}'$  的预测结果。因而，根据本实施例的编码器可以与根据图 6 的实施例的解码器结合使用。

图 9 显示了根据本发明的第四实施例用于编码立体声信号的装置的示意图。本装置是结合图 5 所述的实施例的变形，包括用于执行立体声信号 L 和 R 的旋转的电路 201，用于确定旋转角度的电路 200，自适应滤波器 501，减法电路 502，编码器 202，编码器 503，和组合电路 204，如结合图 5 所述。根据本实施例，主成分信号  $y$  不直接提供给滤波器 501。替代地，本装置进一步包括用常数  $\gamma$  和从电路 201 接收的残余信号  $r$  相乘的乘法电路 901，和用于为主成分信号  $y$  加上缩放的残余信号的加法电路 902，从而生成信号  $y + \gamma r$ ，将其提供给滤波器 501。这里， $\gamma$  是一个小的正值，例如数量级为  $10^2$ 。在一个实施例中，自适应追踪常数  $\gamma$ 。本实施例的优点在于可以在通过滤波器 501 对残余信号  $r$  的建模时，利用在信号  $y$  的频谱中基本不存在而存在于  $r$  的频谱中的频率，从而改善了编码信号的质量。根据本实施例，信号  $y + \gamma r$  被提供给编码器 202，以生成传送到接收器的解码的主信号  $y_e$ 。另外，根据本实施例，将常数  $\gamma$  提供给组合器 204 并传输到接收器。

图 10 显示了根据本发明的第四实施例用于解码立体声信号的装置的示意图，即适于解码从图 9 的编码器接收的信号。本装置包括用于从组合信号  $T$  中提取接收信息的电路 404，解码器 402，解码器 602，滤波器 601 和如结合图 6 所述的旋转电路 401。根据本实施例，电路 404 进一步从组合信号  $T$  中提取常数  $\gamma$ ，并且本装置进一步包括用接收的常数  $\gamma$  与滤波器 601 所产生的预测残余信号  $\hat{r}'$  相乘的乘法电路 1001。本装置进一步包括用于从解码的主信号  $y'$  中减去所得缩放后的预测残余信号  $\gamma \hat{r}'$  的电路 1002。

图 11 显示了根据本发明的第五实施例用于编码多通道信号的装置的示意图。本装置接收包括  $n$  个通道的多通道信号  $x = (x_1, \dots, x_n)$ 。本装置包括用于执行信号  $x$  的主成分分析的主成分分析器 1100，得到用于将输入信号  $x$  变换成主成分信号  $y$  和  $n-1$  个残余信号  $r_1, r_2, \dots, r_{n-1}$  的加权矢量  $w = (w_1, \dots, w_n)$ 。本装置进一步包括变换电路 1101，接收输入信号  $x$  和确定的加权矢量  $w$  并根据以上变换产生信号  $y$  和  $r_1, r_2, \dots, r_{n-1}$ 。变换的信号通过适当的编码器 202 和 1102 编码并通过组合电路 204 与加权矢量  $w$  组合在一起，如结合图 2 所述。根据这一实施例，调整编码器 1102 以编码残余信号  $r_1, \dots, r_{n-1}$ 。例如，编码器 1102 可以包括  $n-1$  个并行编码器，每个如结合编码器 202 所述的那样来编码残余信号之一。

图 12 显示了根据本发明的第六实施例用于编码多通道信号的装置的示意

图。除了如结合图 11 所述对多通道信号的变换之外，根据这一实施例，主成分信号被提供给一组自适应滤波器 501，每个预测残余信号  $r_1, \dots, r_{n-1}$  之一，如结合图 5 所述，从而得到提供给相应的编码器 503 的相应的滤波参数  $F_{p1}, \dots, F_{p(n-1)}$ ，并且随后再提供给组合器 204。在相应的解码器（未示出），使用相应的滤波器基于滤波参数产生残余信号的估计  $\hat{r}'_1, \dots, \hat{r}'_{n-1}$ ，如结合图 6 所述。

可以理解，根据一个实施例，只有残余信号的一个子集例如  $r_1, r_2, \dots, r_k$ ,  $k < n-1$ ，可以被传送到接收器或者提供给相应的滤波器，从而减少了需要的比特率，同时保留了大部分信号质量。

图 13 显示了在本发明的实施例中使用的减法电路的示意图。在图 5、8、9 和 12 的以上实施例中，通过比较目标信号和估计信号来确定滤波参数，即通过例如由减法电路 502 生成的表示  $r$  和  $\hat{r}$  之间差别的误差信号  $e$ 。可以理解，减法电路可以生成  $r$  和  $\hat{r}$  之间差别的不同度量，例如可以在时域或者频域确定差别。参照图 13，电路 502 可以包括用于例如通过执行快速傅里叶变换 (FFT) 将信号  $r$  和  $\hat{r}$  分别变换到频域的电路 1301。所得的频率部分可以进一步地分别通过电路 1304 处理。例如不同的频率可以不同地加权，优选地根据人的听觉系统的特性，从而对可听频率范围内的差别更重地加权。通过电路 1304 的进一步处理的其他例子包括对预定频率部分的平均、计算复杂频率部分的幅度、滤波部分的聚类等。例如在一个优选实施例中，在频域的减法之前执行聚类。此聚类可以通过滤波器组来执行，例如用线性或对数带宽。可选地，可以使用所谓的等效矩形带宽 (ERB)（例见“An introduction to the Psychology of Hearing”，by Brian Moore, Academic Press, London, 1997）执行聚类。等效矩形带宽技术对相应于人的听觉滤波器的频带进行聚类，例如所谓的关键频带。根据本实施例，作为中心频率的函数的 ERB 的对应值  $f$ （以 kHz 为单位）可以根据  $ERB = 24.7(4.37f + 1)$  来计算。仍然参照图 13，电路 502 进一步包括用于减去已处理频率部分的减法电路 1303。可选地，由电路 1301 生成的变换信号不经过进一步处理就被直接提供给减法电路 1304。由减法电路 1304 产生的差信号被提供给变换电路 1302，用于将误差信号变換回时域，例如通过执行逆快速傅里叶变换 (IFFT)。可选地，可以直接使用频域的差信号。

可以理解，普通技术人员可以例如通过添加或减少特征，或者通过结合上述实施例的特征来修改上述实施例。例如，可以理解，图 8 和 9 的实施例中所

介绍的特征也可以结合在图 12 的实施例中。作为另一个例子，说明图 5 的实施例中的估计残余信号的质量的误差信号  $e$  可以与一个表示最大可接受误差的误差阈值比较。如果误差不可接受，误差信号可以在适当地编码后和信号  $T$  一起传输，这与线性预测编码 (LPC) 领域中使用的方法相似。

进一步要注意的是，本发明不限于立体声信号，而是也可以应用于其他具有两个或更多输入通道的多通道输入信号。这种多通道信号的例子包括从数字通用光碟 (DVD) 或超级音频光碟等中接收的信号。在这种更通用的情形下，主成分信号  $y$  和一个或更多残余信号  $r$  仍可以按照本发明生成。传输的残余信号的数目依赖于通道的数目和所希望的比特率，因为可以略去较高阶残余而不会显著降低信号质量。

一般地，本发明的优点在于比特率分配可以自适应地改变，从而可以实现完美的降低。例如，如果信道瞬间只允许传输一个减少的比特率，例如由于增加的网络通信量、噪声等，就可以降低传输信号的比特率而不会显著降低信号的可察觉的质量。例如，在上面讨论的固定声源的情形下，可以用接近 2 的因子来减少比特率而不会显著降低信号质量，相当于传输一个单一通道而不是两通道。

要注意的是，上述装置可以实现为通用或特殊用途的可编程微处理器，数字信号处理器 (DSP)，专用集成电路 (ASIC)，可编程逻辑阵列 (PLA)，现场可编程门阵列 (FPGA)，特定用途的电子电路等，或者它们的组合。

应该注意的是，上述实施例说明而不是限制本发明，并且本领域普通人员能够设计出许多替换实施例而不脱离附加权利要求的范围。在权利要求中，任何括号之间的参照标记不应解释为对权利要求的限制。词汇“包括”不排除未列在权利要求中的其他元件或步骤的存在。本发明能够通过包括多个不同元件的硬件方式和通过一适当编程的计算机方式来实现。在一个列举了多个设备的装置权利要求中，这些设备中的一些能够通过同一硬件单元来实施。在互不相同的独立权利要求中记载某些措施并不表示不能有利地使用这些措施的组合。

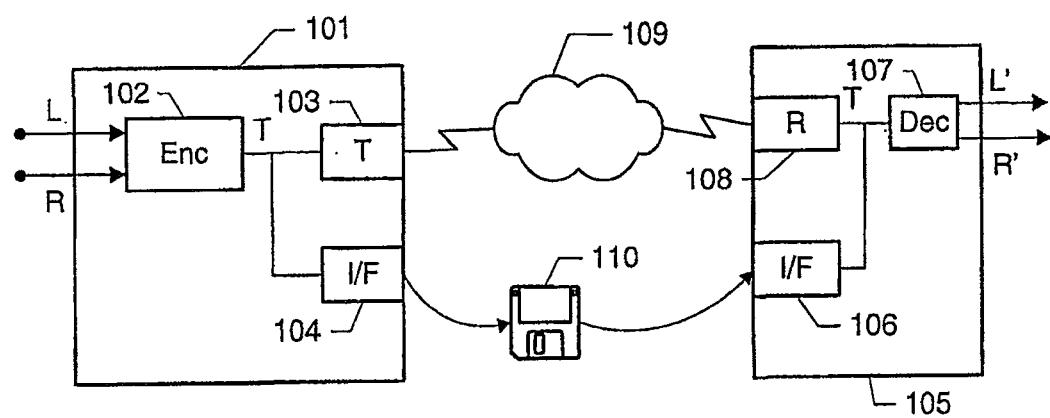


图 1

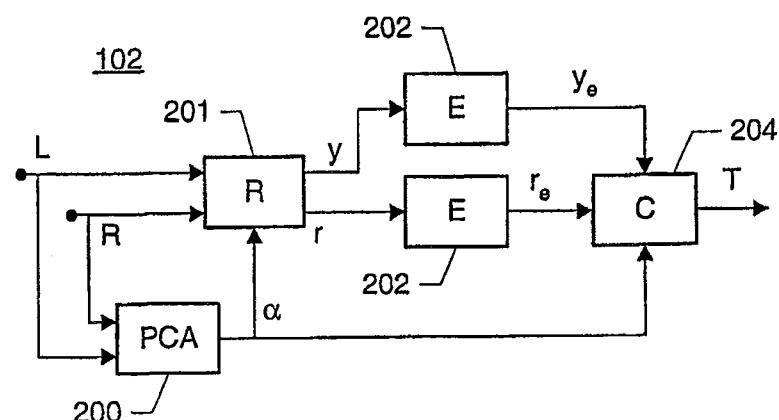


图 2

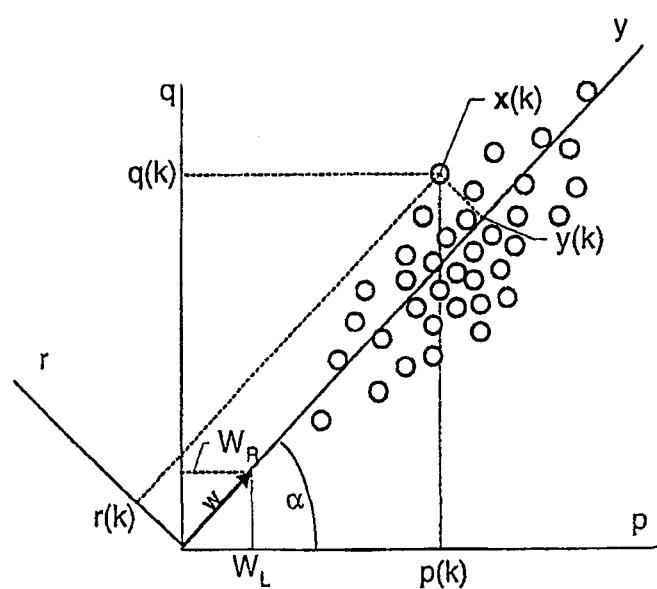


图 3

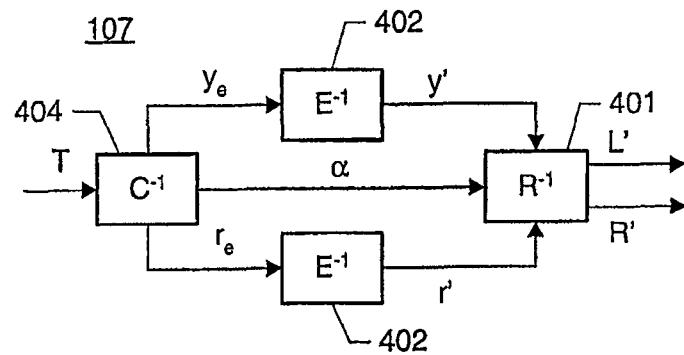


图 4

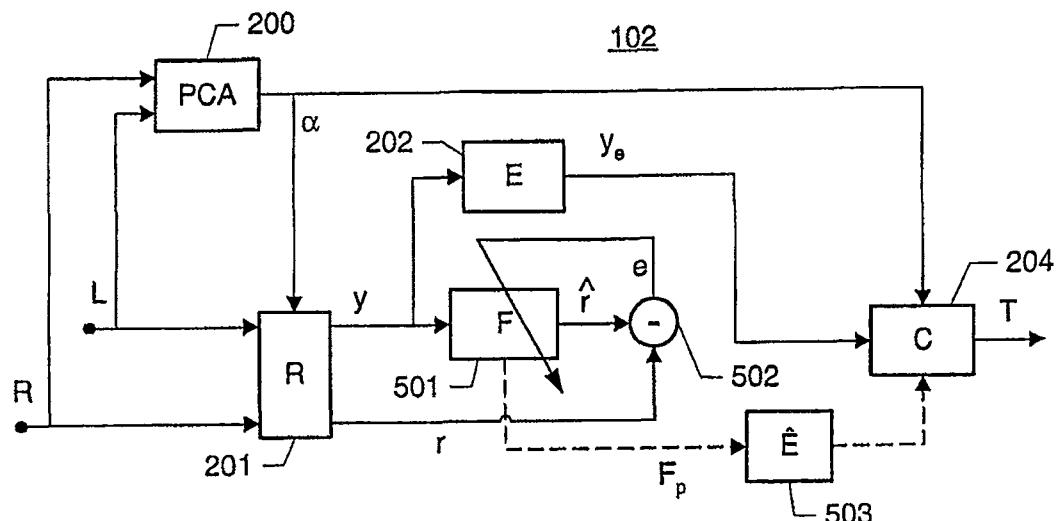


图 5

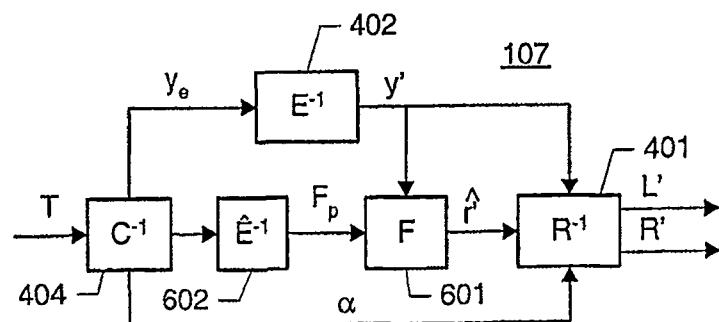


图 6

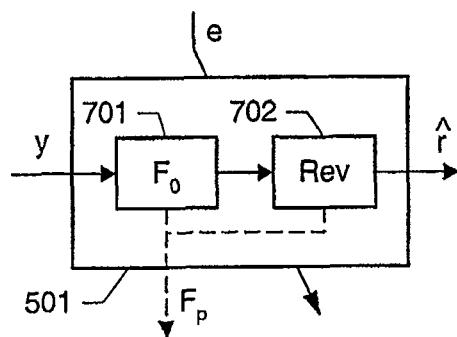


图 7a

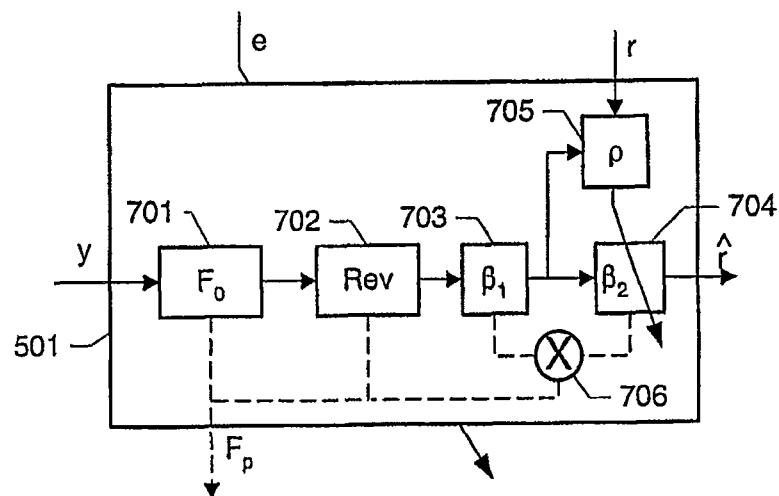


图 7b

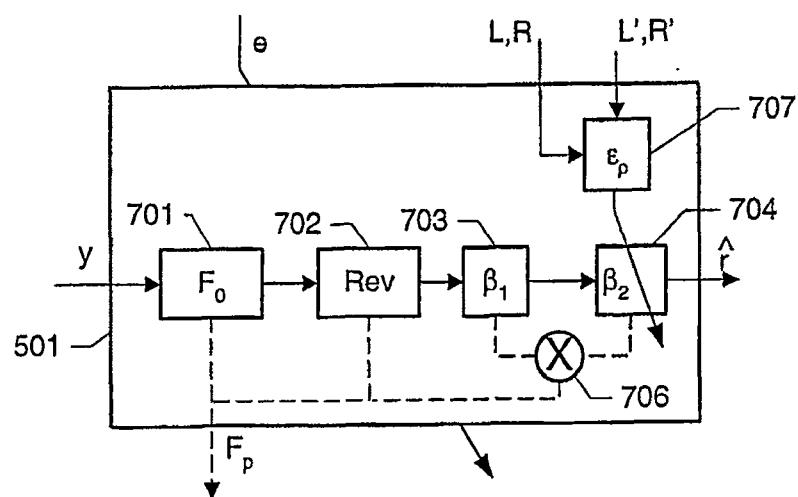


图 7c

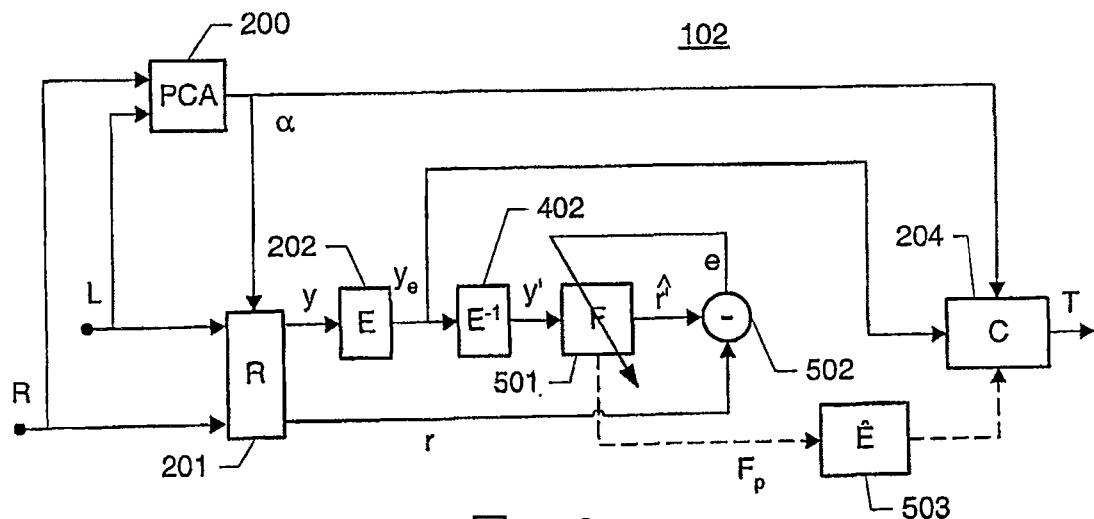


图 8

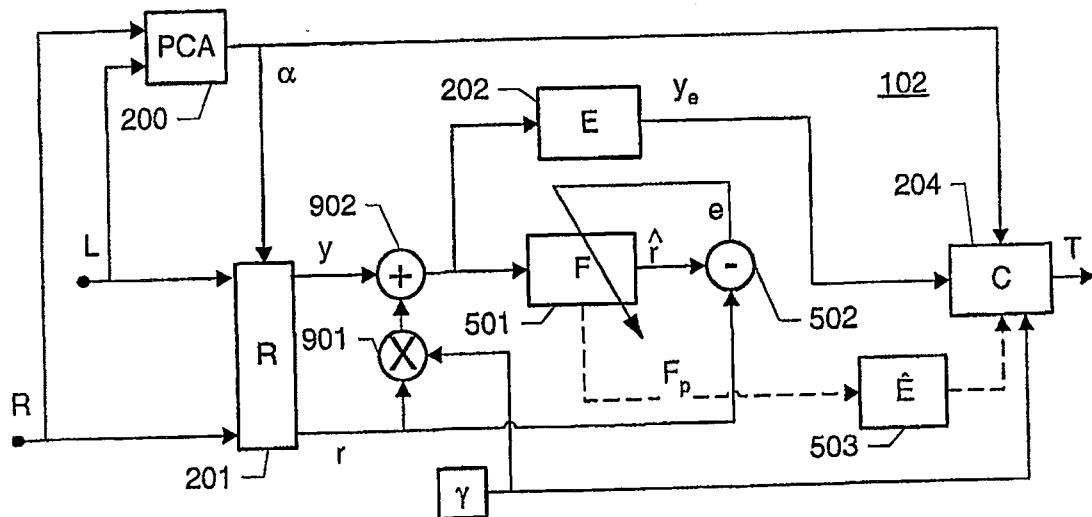


图 9

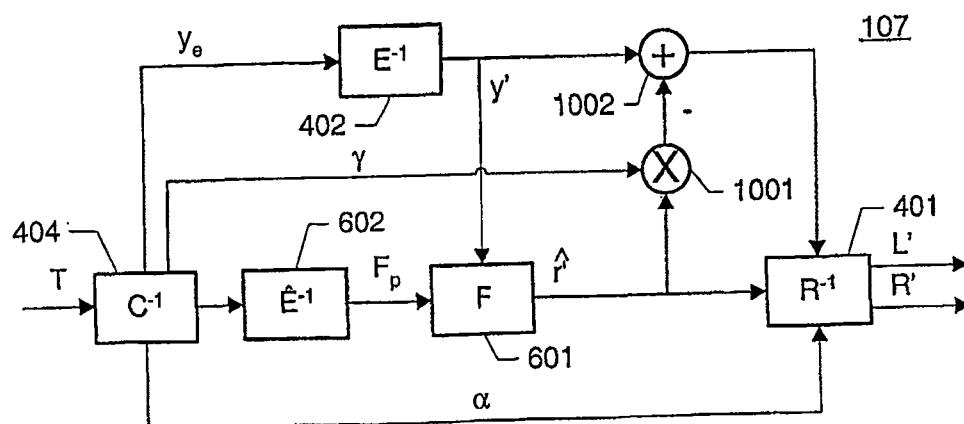


图 10

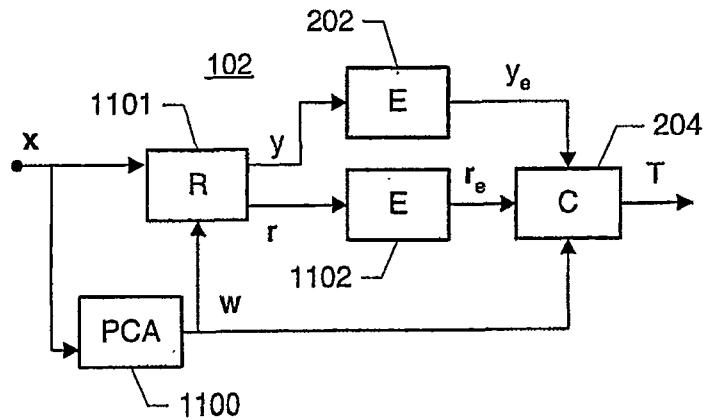


图 11

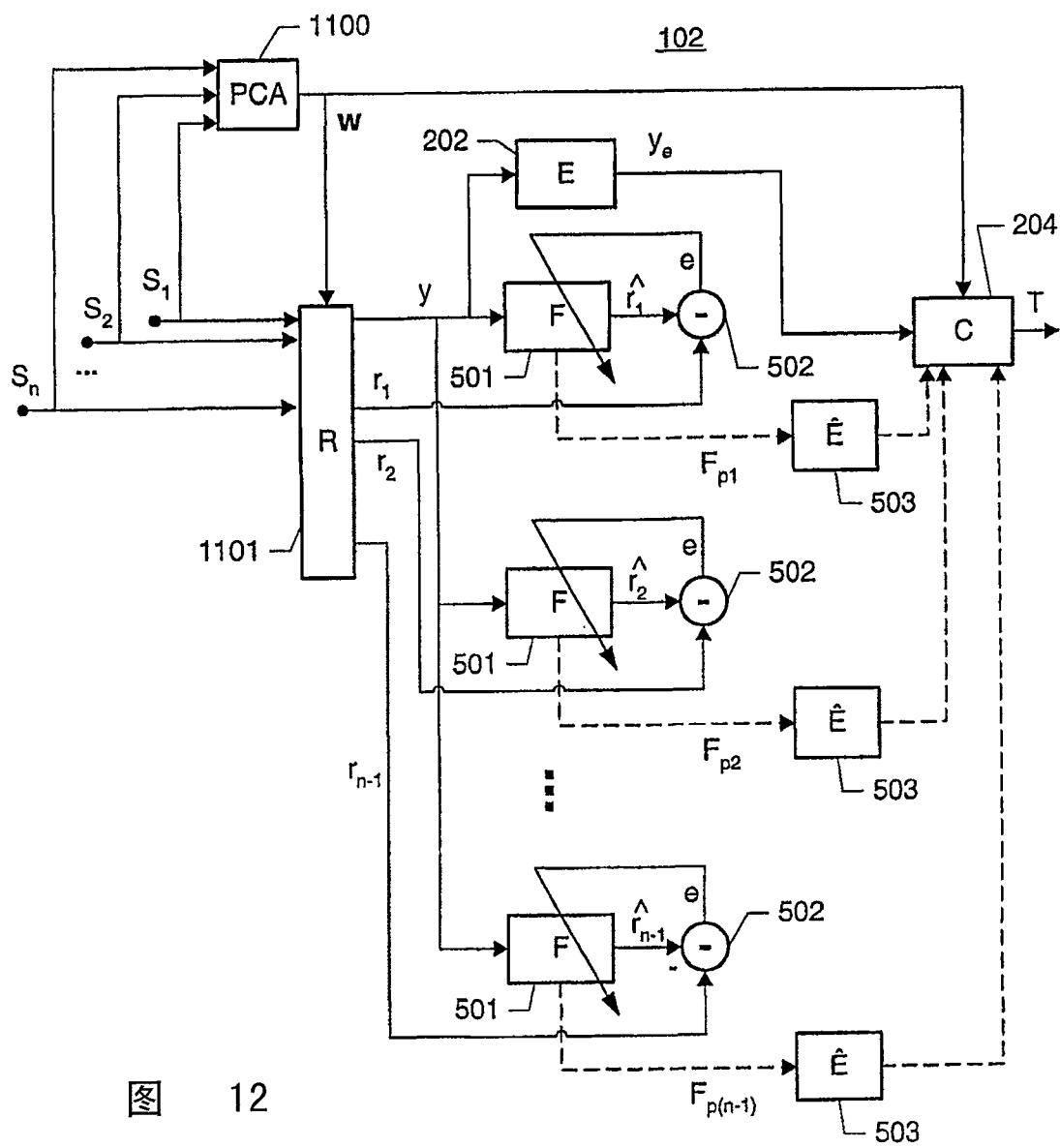


图 12

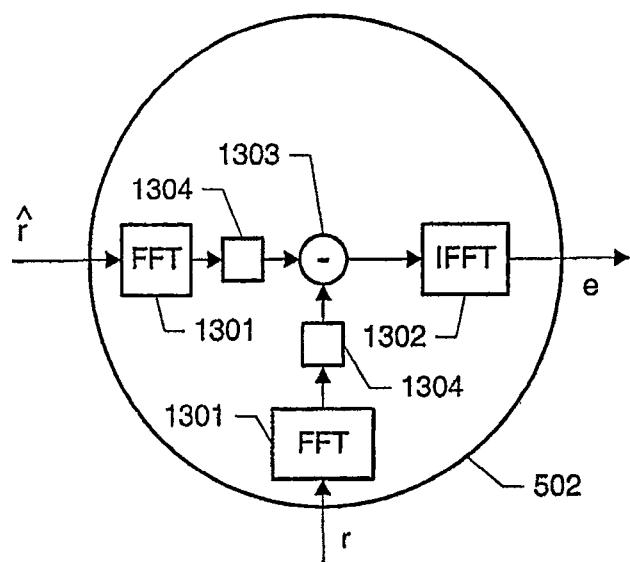


图 13