

[19] 中华人民共和国国家知识产权局

[51] Int. Cl<sup>7</sup>

H03G 3/00

H03G 3/20 H04B 1/00



# [12] 发明专利说明书

[21] ZL 专利号 99807367.9

[43] 授权公告日 2003 年 7 月 30 日

[11] 授权公告号 CN 1116737C

[22] 申请日 1999.4.14 [21] 申请号 99807367.9

[30] 优先权

[32] 1998.4.14 [33] US [31] 09/059,307

[32] 1998.4.14 [33] US [31] 09/059,303

[32] 1998.11.23 [33] US [31] 60/109,506

[32] 1998.4.14 [33] US [31] 09/059,304

[86] 国际申请 PCT/US99/08055 1999.4.14

[87] 国际公布 WO99/53612 英 1999.10.21

[85] 进入国家阶段日期 2000.12.14

[71] 专利权人 听觉增强有限公司

地址 美国弗吉尼亚州

[72] 发明人 R·D·布卢姆 W·R·肖德尔斯

M·A·沃德雷

审查员 张静海

[74] 专利代理机构 中国专利代理(香港)有限公司

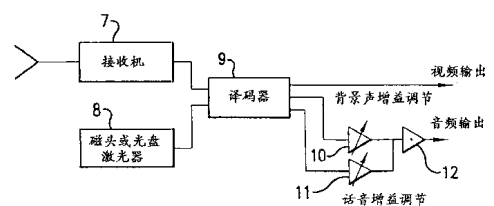
代理人 程天正 张志醒

权利要求书 9 页 说明书 46 页 附图 12 页

[54] 发明名称 用户可调节的适应听力的音量控制

[57] 摘要

本发明是一种根据听力障碍听众以及非听力障碍听众的喜好混合音频信号的方法，而不用迫使听力障碍个人采用专门的听力障碍设备。用户操纵的控制器(10, 11)控制想要的音频信号和其余音频信号在足够宽的、包容所有个人的范围的混合。想要的音频被记录和被保持成与所有的其余音频分开，以及这样地被传送到收听者，以使得它保持这样的分开。控制器(10, 11)使得用户能够规定其中音频可以变化的比值的范围，这允许听众把音频通过连续的范围扩展到他的听觉可以适应的动态范围。控制器(10, 11)自动地调节进入的音频的改变。控制器(10, 11)能够对于相对较慢移动的改变作出反应，或阻止其余音频中的短的声音突发来修改信号电平。



1. 用于为包括听力障碍听众或非听力障碍听众的最终用户播放包括想要的音频信号(P)、右音频信号(RA)、左音频信号(LA)、和中心音频信号(CA)的音频节目的设备, 包括:

5 (a) 第一带通滤波器, 具有接收第一信号的输入端, 第一信号包括左音频信号(LA)加中心音频信号(CA)加想要的信号(P), 所述第一低通滤波器滤除左音频信号(LA), 以及将中心音频信号(CA)加想要的信号(P)作为输出信号而输出;

10 (b) 第二带通滤波器, 具有接收第二信号的输入端, 第二信号包括右音频信号(RA)加中心音频信号(CA), 所述第二低通滤波器滤除右音频信号(RA), 以及将中心音频信号(CA)作为输出信号而输出;

15 (c) 第一加法器, 具有被耦合到第一带通滤波器的输出端的第一输入端, 具有被耦合到第二带通滤波器的输出端的第二输入端, 从第一带通滤波器的输出减去第二带通滤波器的输出, 以及将想要的音频信号(P)作为输出信号而输出;

20 (d) 第二加法器, 具有被耦合到第一加法器的输出端的第一输入端, 具有接收第二信号的第二输入端, 第二信号包括右音频信号(RA)加中心音频信号(CA), 把第一加法器的输出加到第二信号, 以及将右音频信号(RA)加中心音频信号(CA)加想要的信号(P)作为输出信号而输出;

(e) 最终用户可调节的放大器, 被耦合到第一加法器的输出端, 按照最终用户的调节来放大想要的信号(P)的信号电平, 以及将调节的想要的音频信号作为输出信号而输出;

25 (f) 第三加法器, 具有被耦合到最终用户可调节放大器的输出端的第一输入端, 具有被耦合到第二加法器的输出端的第二输入端, 以及将第一最终用户调节的音频信号作为输出信号而输出, 该第一最终用户调节的音频信号包括右音频信号(RA)加中心音频信号(CA)加想要的音频信号(P)加被调节的想要的音频信号;

30 (g) 第四加法器, 具有被耦合到最终用户可调节放大器的输出端的第一输入端, 具有被耦合到第一信号的第二输入端, 该第一信号包括左音频信号(LA)加中心音频信号(CA)加想要的音频信号(P),

以及将第二最终用户调节的音频信号作为输出信号而输出，该第二最终用户调节的音频信号包括左音频信号（LA）加中心音频信号（CA）加想要的音频信号（P）加调节的想要的音频信号；

5 (h) 第一音频输出选择器，具有被耦合到第四加法器的输出端的第一输入端，具有被耦合到第三加法器的输出端的第二输入端，以及根据最终用户的选择而输出第一和第二最终用户调节的音频信号；以及

10 (i) 第二音频输出选择器，具有接收第一信号的第一输入端，该第一信号包括左音频信号（LA）加中心音频信号（CA）加想要的信号（P），具有接收第二加法器的输出的第二输入端，该想要的信号（P）包括第一立体声信号，它由右音频信号（RA）加中心音频信号（CA）加想要的信号（P）组成，以及响应于最终用户的选择而输出第一立体声信号和第一信号作为两个立体声信号。

15 2. 用于为包括听力障碍和非听力障碍听众的最终用户最优化播放包括想要的音频信号和其余音频信号的音频节目的音频系统，包括：

第一用户可调节的放大器，用于接收想要的音频信号并将其放大到由用户规定的电平；

20 第二用户可调节的放大器，用于接收其余音频信号并将其放大到由用户规定的电平；

第一软剪切电路，用于接收想要的音频信号作为输入，具有被耦合到第一最终用户可调节放大器的输入端的输出端，以及当想要的音频信号的电平超过第一预定值时限幅想要的音频信号的总幅度；以及

25 第二软剪切电路，用于接收其余音频信号作为输入，具有被耦合到第二最终用户可调节放大器的输入端的输出端，以及当其余音频信号的电平超过第二预定值时限幅其余音频信号的总幅度，第一和第二软剪切电路与第一和第二最终用户可调节放大器的组合因而可自动保持想要的音频信号对其余音频的用户选择的一个比值，即使在想要的音频信号或其余音频信号中存在瞬态改变的情况下也是如此。

30 3. 按照权利要求2的系统，其特征在于，其中第一和第二软剪切电路中的第一和第二预定值分别是用户可定义的数值。

4. 用于为包括听力障碍和非听力障碍听众的最终用户最优化播

放包括想要的音频信号和其余音频信号的音频节目的音频系统，包括：

5 (a) 第一最终用户可调节的放大器，用于接收想要的音频信号和放大想要的音频信号到由用户规定的电平，所述第一最终用户可调节的放大器具有从零到最大值的放大范围，以及所述第一最终用户可调节的放大器具有一个代表着由第一最终用户可调节放大器的最终用户选择的增益的数字值的增益选择值作为控制输出；

10 (b) 第二最终用户可调节的放大器，用于接收其余音频信号和放大其余音频信号到由用户规定的电平，所述第二最终用户可调节的放大器具有从零到最大值的放大的范围，以及所述第二最终用户可调节的放大器具有一个代表着由第二最终用户可调节放大器的最终用户选择的增益的数字值的增益选择值作为控制输出；

15 (c) 反馈校正器，用于接收第一用户可调节放大器的控制输出作为第一输入，接收第二用户可调节放大器的控制输出作为第二输入，接收第一用户可调节放大器的输出作为第三输入，接收第二用户可调节放大器的输出作为第四输入，以及修改第一用户可调节放大器的输出和第二用户可调节放大器的输出，以使得实际的想要的音频信号对  
20 其余音频信号（PSRA）的比值相应于想要的音频信号对其余音频信号（PSRA）的比值的最终用户想要的数值。

20 5. 按照权利要求 4 的系统，其特征在于，其中反馈校正器包括：

(a) 第一除法器，用于接收第一最终用户可调节放大器的控制输出作为分子输入，接收第二最终用户可调节放大器的控制输出作为分母输入，以及输出所述分子输入除以分母输入的结果，它代表最终用户想要的音频信号对其余音频信号（PSRA）比值；

25 (b) 第一矢量累加器，它被耦合到第一最终用户可调节放大器的输出端，用于存储想要的音频信号的矢量样本；

(c) 第二矢量累加器，它被耦合到第二最终用户可调节放大器的输出端，用于存储其余音频信号的矢量样本；

30 (d) 第一峰值电平检测器，它被耦合到第一矢量累加器的输出端，其中第一峰值电平检测器计算由第一矢量累加器中的矢量样本代表的想要的音频信号的一个分段的瞬态幅度；

(e) 第二峰值电平检测器，它被耦合到第二矢量累加器的输出

端，其中第二峰值电平检测器计算由第三矢量累加器中的矢量样本代表的想要的音频信号的一个分段的瞬态幅度；

5 (f) 第二除法器，用于接收表示由第一矢量累加器中的矢量样本代表的想要的音频信号的该分段的瞬态幅度的第一峰值电平检测器的输出作为分子输入，接收表示由第二矢量累加器中的矢量样本代表的其余音频信号的分段的瞬态幅度的第二峰值电平检测器的输出作为分母输入，以及输出该分子输入除以分母输入的结果，它代表想要的音频信号对其余音频信号（PSRA）比值的实际值；

10 (g) 第三除法器，用于接收表示想要的音频信号对其余音频信号（PSRA）比值的想要的值的第一除法器的输出作为分子输入，接收表示想要的音频信号对其余音频信号（PSRA）比值的实际值的第二除法器的输出作为分母输入，以及输出该分子输入除以分母输入的结果，它代表在想要的音频信号对其余音频信号（PSRA）比值的想要的和实际的数值之间的差别；以及

15 (h) 第一乘法器，用于接收想要的音频信号作为第一输入，接收第三除法器的输出作为第二输入，以及把第三除法器的输出与想要的音频信号相乘并输出相乘结果，它是想要的音频信号对其余音频信号（PSRA）比值的用户设置值的校正的值。

6. 按照权利要求 4 的系统，其特征在于，其中反馈校正器包括：

20 (a) 第一除法器，用于接收第一最终用户可调节放大器的控制输出作为分子输入，接收第二最终用户可调节放大器的控制输出作为分母输入，以及输出该分子输入除以分母输入的结果，它代表最终用户想要的音频信号对其余音频信号（PSRA）比值；

25 (b) 第一矢量累加器，它被耦合到第一最终用户可调节放大器的输出端，用于存储想要的音频信号的矢量样本；

(c) 第二矢量累加器，它被耦合到第二最终用户可调节放大器的输出端，用于存储其余音频信号的矢量样本；

30 (d) 第一峰值电平检测器，它被耦合到第一矢量累加器的输出端，其中第一峰值电平检测器计算由第一矢量累加器中的矢量样本代表的想要的音频信号的一个分段的瞬态幅度；

(e) 第二峰值电平检测器，它被耦合到第二矢量累加器的输出端，其中第二峰值电平检测器计算由第二矢量累加器中的矢量样本代

表的想要的音频信号的一个分段的瞬态幅度;

(f) 第二除法器, 用于接收表示由第一矢量累加器中的矢量样本代表的想要的音频信号的分段的瞬态幅度的第一峰值电平检测器的输出作为分子输入, 接收表示由第二矢量累加器中的矢量样本代表的  
5 其余音频信号的分段的瞬态幅度的第二峰值电平检测器的输出作为分母输入, 以及输出该分子输入除以分母输入的结果, 它代表想要的音频信号对其余音频信号 (PSRA) 比值的实际值;

(g) 第三除法器, 它接收表示想要的音频信号对其余音频信号 (PSRA) 比值的想要的值的第一除法器的输出作为分子输入, 接收表  
10 示想要的音频信号对其余音频信号 (PSRA) 比值的实际值的第二除法器的输出作为分母输入, 以及输出该分子输入除以分母输入的结果, 它代表在想要的音频信号对其余音频信号 (PSRA) 比值的想要的和实际的数值之间的差别;

(h) 第四除法器, 它以第三除法器的输出作为分母输入, 以及用 1  
15 来除以第三除法器的输出; 以及

(i) 乘法器, 用于接收其余音频信号作为第一输入, 接收第四除法器的输出作为第二输入, 以及把第四除法器的输出与其余音频信号相乘后, 输出相乘结果, 该相乘结果是想要的音频信号对其余音频信号 (PSRA) 比值的用户设置值的校正的值。

20 7. 按照权利要求 4 的系统, 其特征在于, 其中第一和第二矢量累加器包括用户可选择的矢量大小。

8. 用于为包括听力障碍和非听力障碍听众的最终用户最优化播放包括想要的音频信号和其余音频信号的音频节目的音频系统, 包括:

25 (a) 多个用户可选择的与频率无关的放大装置, 每个装置确定在分开的频段上想要的音频信号对其余音频 (PSRA) 比值, 以及多个用户选择的与频率无关的放大装置中的每一个包括:

(i) 第一带通滤波器, 用于接收想要的音频信号和滤除在想要的音频信号内的、除了想要的音频信号的预定的带宽以外的所有信号;

30 (ii) 第二带通滤波器, 用于接收其余音频信号和滤除在其余音频信号内的、除了其余音频信号的预定的带宽以外的所有信号;

(iii) 第一最终用户可调节放大器, 用于接收想要的音频信号的

5 预定的带宽和将想要的音频信号的预定的带宽放大到由用户规定的电平，所述第一最终用户可调节的放大器具有从零到最大值的一个放大范围，以及所述第一最终用户可调节的放大器具有代表由第一最终用户可调节放大器的最终用户选择的增益的数字值的增益选择值作为控制输出；

10 (iv)第二最终用户可调节放大器，用于接收其余音频信号的预定的带宽和将其余音频信号的预定的带宽放大到由用户规定的电平，所述第二最终用户可调节的放大器具有从零到最大值的一个放大范围，以及所述第二最终用户可调节的放大器具有代表由第二最终用户可调节放大器的最终用户选择的增益的数字值的增益选择值作为控制输出；以及

15 (v)反馈校正器，接收第一用户可调节放大器的控制输出作为第一输入，接收第二用户可调节放大器的控制输出作为第二输入，接收第一用户可调节放大器的输出作为第三输入，接收第二用户可调节放大器的输出作为第四输入，以及修改第一用户可调节放大器的输出和第二用户可调节放大器的输出，以使得对于预定的带宽的实际的想要的音频信号对其余音频信号(PSRA)比值相应于对于预定的带宽的想要的音频信号对其余音频信号(PSRA)比值的最终用户想要的数值。

20 (b)相加放大器，用于接收所有的多个用户可选择的与频率无关的放大装置的输出，以便产生总的音频信号。

9.按照权利要求8的系统，其特征在于，其中反馈校正器包括：

25 (a)第一除法器，用于接收第一最终用户可调节放大器的控制输出作为分子输入，接收第二最终用户可调节放大器的控制输出作为分母输入，以及输出该分子输入除以分母输入的结果，它代表最终用户想要的音频信号对其余音频信号(PSRA)比值；

(b)第一矢量累加器，它被耦合到第一最终用户可调节放大器的输出端，用于存储想要的音频信号的矢量样本；

(c)第二矢量累加器，它被耦合到第二最终用户可调节放大器的输出端，用于存储其余音频信号的矢量样本；

30 (d)第一峰值电平检测器，它被耦合到第一矢量累加器的输出端，其中第一峰值电平检测器计算由第一矢量累加器中的矢量样本代表的想要的音频信号的一个分段的瞬态幅度；

(e) 第二峰值电平检测器，它被耦合到第二矢量累加器的输出端，其中第二峰值电平检测器计算由第二矢量累加器中的矢量样本代表的想要的音频信号的一个分段的瞬态幅度；

5 (f) 第二除法器，用于接收表示由第一矢量累加器中的矢量样本代表的想要的音频信号的分段的瞬态幅度的第一峰值电平检测器的输出作为分子输入，接收表示由第二矢量累加器中的矢量样本代表的其余音频信号的分段的瞬态幅度的第二峰值电平检测器的输出作为分母输入，以及输出该分子输入除以分母输入的结果，它代表想要的音频信号对其余音频信号（PSRA）比值的实际值；

10 (g) 第三除法器，用于接收表示想要的音频信号对其余音频信号（PSRA）比值的想要的值的第一除法器的输出作为分子输入，接收表示想要的音频信号对其余音频信号（PSRA）比值的实际值的第二除法器的输出作为分母输入，以及输出该分子输入除以分母输入的结果，它代表在想要的音频信号对其余音频信号（PSRA）比值的想要的和实际  
15 的数值之间的差别；以及

(h) 第一乘法器，用于接收想要的音频信号作为第一输入，接收第三除法器的输出作为第二输入，以及对第三除法器的输出与想要的音频信号进行相乘，输出相乘结果，该相乘结果是想要的音频信号对  
其余音频信号（PSRA）比值的用户设置值的校正的值。

20 10. 按照权利要求 8 的系统，其特征在于，其中反馈校正器包括：

(a) 第一除法器，用于接收第一最终用户可调节放大器的控制输出作为分子输入，接收第二最终用户可调节放大器的控制输出作为分母输入，以及输出该分子输入除以分母输入的结果，它代表最终用户  
想要的音频信号对其余音频信号（PSRA）比值；

25 (b) 第一矢量累加器，它被耦合到第一最终用户可调节放大器的输出端，用于存储想要的音频信号的矢量样本；

(c) 第二矢量累加器，它被耦合到第二最终用户可调节放大器的输出端，用于存储其余音频信号的矢量样本；

30 (d) 第一峰值电平检测器，它被耦合到第一矢量累加器的输出端，其中第一峰值电平检测器计算由第一矢量累加器中的矢量样本代表的想要的音频信号的一个分段的瞬态幅度；

(e) 第二峰值电平检测器，它被耦合到第二矢量累加器的输出

端，其中第二峰值电平检测器计算由第二矢量累加器中的矢量样本代表的想要的音频信号的一个分段的瞬态幅度；

5 (f) 第二除法器，用于接收表示由第一矢量累加器中的矢量样本代表的想要的音频信号的分段的瞬态幅度的第一峰值电平检测器的输出作为分子输入，接收表示由第二矢量累加器中的矢量样本代表的其余音频信号的分段的瞬态幅度的第二峰值电平检测器的输出作为分母输入，以及输出该分子输入除以分母输入的结果，它代表想要的音频信号对其余音频信号 (PSRA) 比值的实际值；

10 (g) 第三除法器，用于接收表示想要的音频信号对其余音频信号 (PSRA) 比值的想要的值的第一除法器的输出作为分子输入，接收表示想要的音频信号对其余音频信号 (PSRA) 比值的实际值的第二除法器的输出作为分母输入，以及输出该分子输入除以分母输入的结果，它代表在想要的音频信号对其余音频信号 (PSRA) 比值的想要的和实际的数值之间的差别；

15 (h) 第四除法器，它以第三除法器的输出作为分母输入，以及用 1 来除以第三除法器的输出；以及

(i) 乘法器，用于接收其余音频信号作为第一输入，接收第四除法器的输出作为第二输入，以及把第四除法器的输出与其余音频信号进行相乘并且输出相乘结果，该相乘结果是想要的音频信号对其余音频信号 (PSRA) 比值的用户设置值的校正的值。

11. 用于为包括听力障碍和非听力障碍听众的最终用户最优化播放包括想要的音频信号和其余音频信号的音频节目的方法，包括：

将想要的音频信号放大到由用户规定的电平；

将其余音频信号放大到由用户规定的电平；

25 计算第一比值，该第一比值是由用户规定的想要的音频信号电平与其余音频信号电平的比值；

计算想要的音频信号的瞬态幅度；

计算其余音频信号的瞬态幅度；

30 产生第二比值，该第二比值是所计算的想要的音频信号瞬态幅度对想要的音频信号的瞬时幅度的比值；

产生第三比值，该第三比值是所计算的第二音频瞬态幅度对其余音频信号的瞬时幅度的比值；

根据第二比值修改想要的音频信号;

根据第三比值修改其余音频信号;

其中在想要的音频信号和其余音频信号中的至少一个信号的瞬态改变期间, 想要的音频信号和其余音频信号被修改, 以便保持第一比值。

5

12. 用于为包括听力障碍和非听力障碍听众的最终用户最优化播放包括想要的音频信号和其余音频信号的音频节目的方法, 包括:

将想要的音频信号放大到由用户规定的电平;

将其余音频信号放大到由用户规定的电平; 以及

10

通过使用反馈校正环来修改想要的音频信号放大步骤的输出和其余音频信号放大步骤的输出, 以使得实际的想要的音频信号对其余音频信号(PSRA)比值相应于想要的音频信号对其余音频信号(PSRA)比值的用户想要的数值。

## 用户可调节的适应听力的音量控制

## 相关的专利申请

5 本申请涉及由 Ronald D. Blum 和 William E. Kokonaski 在 1997 年 5 月 6 日提交的、题目为“Hearing Enhancement System and Method(听觉增强系统和方法)”的美国专利申请 No. 08/852, 239, (HEC-1), 该申请包括附图在此引用, 以供参考, 正如这里作为整体地被重复。

10 本申请涉及由 Ronald D. Blum 和 William E. Kokonaski 在 1997 年 8 月 8 日提交的、题目为“Listening Enhancement System and Method(听觉增强系统和方法)”的美国专利申请 No. 08/907, 503, (HEC-2), 该申请包括附图在此引用, 以供参考, 正如这里作为整体地被重复。

15 本申请涉及由 Ronald D. Blum, William E. Kokonaski, William R. Saunders 和 Michael A. Vaudrey 在 1998 年 4 月 14 日提交的、题目为“Improved Hearing Enhancement System and Method(改进的听觉增强系统和方法)”的美国专利申请 No. 09/059, 307(HEC-3)的部分继续申请, 该申请包括附图在此引用, 以供参考, 正如这里作为  
20 整体地被重复。

本申请涉及由 Ronald D. Blum, William E. Kokonaski, William R. Saunders 和 Michael A. Vaudrey 在 1998 年 4 月 14 日提交的、题目为“Improved Listening Enhancement System and Method(改进的听觉增强系统和方法)”的美国专利申请 No. 09/059, 303(HEC-4)  
25 的部分继续申请, 该申请包括附图在此引用, 以供参考, 正如这里作为整体地被重复。

本申请涉及由 William R. Saunders 和 Michael A. Vaudrey 在 1998 年 4 月 14 日提交的、题目为“Improved Aural System and Method(改进的听觉系统和方法)”的美国专利申请  
30 No. 09/059, 304(HEC-5)的连续部分, 该申请包括附图在此引用, 以供参考, 正如这里作为整体地被重复。

本申请涉及由 William R. Saunders 和 Michael A. Vaudrey 在

1998年11月23日提交的、题目为“Special Application Digital Audio Decoder(专用数字音频译码器)”的美国临时专利申请No. 60/109, 506(HEC-6), 该申请包括附图在此引用, 以供参考, 正如这里作为整体地被重复。

5

### 发明背景

本发明总的涉及用于处理音频的方法和设备, 更具体地, 涉及用于处理音频以便改进包括听力障碍的听众在内的广泛范围的听众的听觉的方法和设备。

10 随着一个人的年老和在一生中不断变化, 由于许多因素, 诸如年老、遗传、疾病、和环境影响, 人的听力将受到损害。通常, 情况的恶化特定于某个频段。

除了永久听力障碍以外, 在遭遇到特定的高声级时, 人们可能发生临时听力障碍。例如, 在打靶射击或参加摇滚音乐会以后, 人们可能遇到暂时听力障碍, 它随后多少会有改进, 但随着时间会积累成永久听力障碍。即使比这些更低的声级但持续更长的时间, 也可以对人们的听力具有暂时影响, 诸如在工厂工作或在小学教书。

15 典型地, 人们通过提高声音的音量来补偿听觉损失或障碍。但是, 这只是简单地增加总的信号中的所有的可听见的频率的音量。总的信号音量的最终的增加, 在语言可听懂度方面很少或不会给出改进, 特别是对于其听力障碍是与频率有关的那些听众。

虽然听力障碍通常随年龄而增加, 但许多听力障碍的个人拒绝承认他们的所觉困难, 因而避免使用可以改进他们的收听质量的装置。虽然许多老年人随着他们的年龄开始戴眼镜, 但这些人中极少数戴助听器, 尽管在减小助听器的体积上已取得了很大进步。这种现象反映与助听器和/或听力障碍有关的外观上的社会歧视。因此, 希望能提供一种改进听力障碍的听众的听觉的技术, 由此避免与外观有关的社会歧视。

25 大多数声音节目(诸如电视声音、电影声音、或音乐)可被分成两个不同的部分: 前台和背景。通常, 前台声音是打算获取听众注意和保持他们的集中力的声音, 而背景声音是作为支持, 但却不是听众主要感兴趣的。这方面的一个例子在一个“坐谈(sitcom)”的电视节

目中可以看到，其中主角的声音表达和阐述了故事的情节，而同时的声音效果、听众笑声、和音乐则填充了间隙。

5 当前，对于所有类型的音乐媒体的听众受到在制作期间由音频工程师决定的混合制品的限制。音频工程师把所有其它的背景噪声与前台声音按音频工程师喜爱的程度进行混合，或音频工程师根据某种历史上的理解进行混合。这个混合制品然后作为单声道信号或在某种情况下作为立体声（左和右）信号被发送到最终用户，而不需要任何用于相对背景声音来调节前台声音的装置。

10 缺乏这种相对于背景声音调节前台声音的能力，对于听力障碍情况是特别困难的。在许多情况下，由于背景声音掩盖前台声音，因而很难（最好地）理解节目。

有许多新的数字音频格式是可提供的。其中某些格式试图为听力障碍的人提供能力。例如，Dolby(杜比)数字，也被称为 AC-3(音频编码译码版本 3)，是用于数字音频的压缩技术，它把更多的数据封装为更小的空间。数字音频的特性是在空间定位方面，这是通过提供 5.1 15 分开的声音通道：中心、左和右、以及左和右环绕而完成的。第六通道，被称为 0.1 通道，是指有限带宽的低频效果（LFE）通道，它由于其低频而是几乎没有方向性的。由于有 5.1 声音通道发送，必须用压缩，以便确保视频和音频处在一定的带宽约束条件内。这些约束条件 20 (由 FCC 施加)对于地面传输比起当前的 DVD 是更严格的。对于 DVD 有足够的空间来为用户提供未压缩的音频（从收听的立场看是特别希望的）。视频数据最通常是通过 MPEG（活动图象专家组）开发的技术而被压缩的，虽然它们也具有非常类似于 Dolby 的音频压缩技术。

DVD 工业采用杜比数字（DD）作为其选择的压缩技术。大多数 DVD 25 是使用 DD 产生的。ATSC（高级电视标准委员会）也选择 AC-3 作为其用于美国数字 TV 的音频压缩方案。这扩散到全世界许多其它国家。这意味着电影制作棚（电影和电视）必须按 DD 来编码他们的音频，以用于广播或录音。

除了严格的编码和译码方案以外，还具有许多特性，它们通常结合杜比数字被讨论。这些特性中的有些是 DD 的一部分，有些不是。 30 连同压缩的比特流，DD 发送关于被称为元数据(metadata)或“关于数据的数据”的比特流的信息。它基本上是一些 0 和 1，表示可供最

终用户使用的任选项的存在。与 HEC 有关的这些任选项中的三个任选项是 dialnorm (对话归一化), dynrng (动态范围), 和 bsmod (控制主要的和相关的音频业务的比特流模式)。前两个是 DD 的整体部分, 因为许多译码器处理这些变量, 给予最终用户以调节这些变量的能力。信息的第三比特, bsmod, 在 ATSC 文件 A/54 (不是杜比出版物) 中详细描述, 但也作为 DD 比特流的一部分存在。bsmod 的数值提醒译码器关于输入的音频业务的特性, 包括任何相关的音频业务的存在。在目前, 还没有已知的制造者利用这个参量。通过在 DVD 上的 8 个可提供的音道之一上的多种完全的主要音频节目提供多种语言 DVD 性能。

dialnorm 参量被设计成允许听众相对于恒定的声级归一化所有的声音节目。在通道之间和在节目与广告之间, 总的声级在很大范围上波动。将来, 将会要求制作人插入表示对话被记录时的声级 (SPL) 的 dialnorm 参量。如果这个数值对于节目被设置为 80dB, 而对于广告被设置为 90dB, 则电视将译码该信息以说明用户输入的声级是所希望的 (比如说 85dB), 从而将把电影提高 5dB 和将广告降低 5dB。这是基于制作人作为 dialnorm 比特值的输入的总的音量电平调节。

来自 AC-3 说明的章节 (来自文件 A/52) 提供了这个技术的最好说明。“dynrng 数值典型地表示在最响的信号环境期间的增益降低和在安静环境期间的增益提高。对于收听者, 希望把最响的声音在声级上变低到对话声级, 以及把轻的声音在声级上再次变高到对话声级。与正常的讲话的对话一样响度的声音典型地将不改变它们的增益。”

dynrng 变量为用户提供可调节的参量, 它将会相对于对话声级来控制总的音量上出现的压缩量。这实质上将总的声音节目的动态范围限制在平均对话声级附近。然而, 这并不提供独立于其余音频电平地调节对话电平的任何方法。

在 ATSC, 数字电视标准 (附录 B) 中给出了改进听力障碍听众的听觉的一个尝试。ATSC 标准的附录 A 的第 6 节描述了主要音频业务和相关的音频业务。AC-3 基本数据流包含单个音频业务的编码表示。多个音频业务通过多个基本数据流被提供。每个基本数据流通过带有独特的 PID 的输送复用方式被传送。有多种音频业务, 它们可以被单独地被编码成每个基本数据流。一种类型的音频业务被称为完全的主音

频业务 (CM)。CM 主音频业务类型包含完全的音频节目 (包括有对话、音乐和效果)。CM 业务可以包含从 1 到 5.1 音道。CM 业务还可借助于其它业务被增强。另一种类型的 CM 业务是听力障碍业务 (HI)。与 HI 有关的业务典型地只包含打算与 CM 业务同时被重现的对话。在这种情况下, HI 业务是单个音道。正如其中所述的, 这种对话可被处理成可以改善听力障碍听众对它的可听懂程度。CM 和 HI 业务的同时重现可以允许听力障碍听众听到 CM 和 HI 业务的混合, 以便在强调对话的同时仍提供某些音乐和效果。除了提供 HI 业务作为单个对话通道的完全的节目混合而被提供。在这种情况下, 业务可通过使用任何数目的通道 (多达 5.1) 被编码。虽然这个业务可以改进某些听力障碍的人的听觉, 但它并不是为了那些害怕被说成是听力障碍的、不采用被排斥的接收机的那些人。最后, 听力障碍的人对于对话的任何处理会阻止在为非听力障碍的人产生音频节目时这个通道的使用。而且, 在附录 B 中阐述的、在 HI 业务与 CM 业务之间的关系, 相对于所使用的每个相对信号电平来说保持为未被规定的, 以便产生用于听力障碍的人的通道。

另一个技术被利用来企图改进声音的可听懂度。例如, 美国专利 No. 4, 024, 344 揭示了一种建立用于电影院声音中的对话的“中心通道”的方法。那里揭示的这种技术把左和右立体声通道进行相关, 以及根据在左和右通道之间的相关程度来调节在组合的和或分开的左或右通道上的增益。在左和右通道之间的强的相关的假定可以表示存在对话。中心通道 (它是左右通道的滤波后的总和) 根据在左和右通道之间的相关程度被放大或衰减。这种方法的问题在于, 它分辨不出有意义的对话与仅仅是已相关的声音, 也不找出话音频段内不想要的信息。所以, 它不能改进所有听力障碍的人对所有声音的可听懂度。

在美国专利 No. 5, 197, 100 中, Shiraki 讨论了电视信号中从背景声音分离出话音。Shiraki 采用的技术包括结合相加和相减电路使用带通滤波来形成“话音通道”, 它是从其余的声音节目分出的。这个方法的限制是带通滤波器只分辨在预定范围内的频率, 在本例中是 200 到 500Hz。它不能在话音和处在通带频率内的背景声音之间进行分辨。而且, 通带滤波的应用不能区分声音信号内的、与语音相关的

和非相关的分量。

在美国专利 No. 5, 485, 522 中, Solve 等讨论了减小音频帧中的背景噪声的装置, 该专利揭示了语音检测器和噪声估值器, 它们被用来自适应调节对音频电信号每个帧的衰减。这种和其它的形式的自适应噪声滤波不能区分在话音频段中的语音和其它非稳态音频, 诸如音乐或与语音不相关的信息。因此, 对可听懂度的改进小于最佳值。

在大的噪声变化的情形下已经尝试过通过修改信号电平来改进听觉。例如, 授权给 Miller 等的美国专利 No. 5, 434, 922 揭示了音乐声音最佳化的方法和系统以用于测量汽车中的音乐和噪声。Miller 等通过自动增强声音信号来补偿环境噪声背景的计算, 从而使用模拟-数字变换和自适应滤波。这个技术不能补偿被其余的音频信号淹没的、特定的收听者所想要的音频信号。在 Miller 等的系统中, 该系统在试图克服当存在道路和引擎噪声时仅仅增加总的信号电平。在大多数情况下, 对于特定的收听者不能听懂的声音仅仅提高信号电平并不能使其变成为听懂的。

通常, 采用带通滤波或选择性均衡的现有技术并不从音频节目的语音部分中去除在话音频段范围内的话音频段背景或噪声。上述的 Dolby、Shiraki、和 Miller 等的发明都试图通过各种信号处理硬件或算法来修改音频信号的某些内容, 但那些方法不能满足不同听众中的个人需要或喜好。总之, 所有这些技术不能为听力障碍个人以及非听力障碍个人提供最佳的听觉。

最后, 在立体声录音的情况下, 原声通常被分开地记录, 然后再与乐器混合, 并放置在单个记录轨道上。所以最终用户只能调节音量、音色和平衡(在立体声的情况下), 但不能相对于语音分量或背景分量的信号电平进行调节。

所以, 本发明针对开发一种用于处理音频信号的系统和方法, 它使得听力障碍的听众以及非听力障碍的听众的听觉最佳化, 而不迫使听力障碍的个人需要他们自己采用专门的针对听力障碍的设备而感觉受歧视。

### 发明概要

本发明通过提供用户操纵的控制器来解决这个问题, 该控制器控

制在足以包容从听力障碍的听众到非听力障碍的听众在内的所有个人的足够宽的范围内的、想要的音频信号和剩余的音频信号的混合。通过提供一种由所有听众利用来使得它们的听觉对于它们独特的收听需要最佳化的控制器，本发明避免在使用这样的控制器时附加上任何社会歧视。而且，本发明因此不需要使用用于听力障碍的专门的设备，和通常由于小的制造量引起的、与这样的专门设备有关的随之产生的费用，以及与每个装置有关的、较高百分比的研究和开发费用。

按照本发明的另一个方面，所想要的音频被记录，并与所有其余的音频分开地保存，以及按照保持想要的音频与其余的音频分开的方式被传送给听众（或最终用户）。这使得最终用户能够对混合的比例完全进行控制，这可由最终用户在重放音频时进行调节，而不是在音频被记录时调节以便适应他或她的听力障碍情况。而且，保持两个音频源分开，就可以允许用户根据被重放的音频的类型、听众的组成情况（即试图找到一种折衷的设置以满足若干个人收听音频）、以及包括外部噪声源和室内声音等在内的严格的环境的需要来进行调节。事实上，本发明因此提供重放音频的无限灵活性，这在以前则是不可能的。

按照本发明的另一个方面，用户操纵的控制器包括了在面对音频信号改变时自动保持听众建立的比值的能力。换句话说，如果其余音频的信号电平突然改变，控制器会自动地调节想要的音频的信号电平或其余音频信号电平来进行补偿。在这种情况下，如果行动移到趋向于遮盖想要的音频的室外有风的区域，则听众不需要重新调节该比值设置来进行补偿。控制器能自动校正这个改变。

按照本发明的再一个方面，用户操纵的控制器使得用户能够规定其中音频可以变化的比值的范围，以便提供用户可规定的动态范围。这允许听众把连续范围的音频扩展到他的听觉可以适应的动态范围。这样，年轻听众可以设置一个对于想要的音频相对其余的音频比值的数值，然后完全地开放动态范围，从而提供完全的音频自由度，而具有听力损害的年老听众可以规定对于特定比值的一个非常窄的范围，使音频被限制在该范围中变化。将动态范围减小到使得其中想要的音频信号被保持不变，将会产生一个单音信号，这对于严重听力障碍的个人听懂声音可能是必须的，但这对于轻微听力障碍的个人可

能是不需要的。这种范围的可调节性使得用户产生他的最佳听觉，这在很大程度上取决于多个因素，这些因素最好是在重放音频节目时进行考虑，而不是在录制音频节目时考虑。

按照本发明的又一个方面，控制器通过自动调节从而改变输入的音频，然而，控制器对于相对较慢移动的改变能够作出反应，或阻止其余音频中的短的声音突发去修改信号电平。这样，根据用户的收听需要，可以使枪击声淹没或不淹没任何想要的音频。

本发明的以上的方面的组合提供了以前不可能的、可适应所有听众的收听需要的听觉。通过控制想要的音频与其余音频的比值、和规定音频可以变化的比值的动态范围的能力，再加上控制器响应进入的音频的突然变化而自动调节听觉信号电平的能力，所有以上这些的组合提供了强有力的用户能力，使得任何收听者的听觉真正最佳化。

#### 附图简述

图 1 显示按照本发明的、用于把在记录的或广播的节目中有有关的话音信息与总的背景音频分开的一般方法。

图 2 显示按照本发明的、用于接收和重放编码的节目信号的示例的实施例。

图 3 显示一个示例的技术，其中想要的信号可以通过话音识别被传送给最终用户。

图 4 显示图 3 上描绘的话音识别装置的数字形式的电路图。

图 5 显示按照本发明的、在到达最终用户以前想要的信号相对其余音频信号对于任何一般媒体类型可以采取的一条路径的通用化的方框图。

图 6 显示在到达最终用户以前由电视节目的音频和视频部分采取的信号路径的现有技术方框图。

图 7 显示按照本发明的、包含了同时发送两个分开的音频信号的能力的示例的实施例。

图 8 显示使用来自制作过程的独特的立体声混合来传送相对较纯的想要的信号内容给最终用户，而并不影响正常的立体声音频节目编制，这通常可被应用于采用立体声节目编制的任何媒体。

图 9 显示按照本发明的、对于想要的信号相对其余音频的示例的

增益调节实施例。

图 10 显示按照本发明的一个方面把均衡处理包括在增益调节中，以便产生与频率有关的增益。

5 图 11 显示按照本发明的、用于为了自动保持用户选择的想要的信号相对于其余音频（PSRA）的比值而软剪切想要的信号和其余音频的示例的实施例的简化方框图。

图 12 显示按照本发明的、用于实施图 11 所示的软剪切方法的示例的实施例。

10 图 13 显示按照本发明的、用于在任一个信号中存在瞬态时保持想要的音频信号或其余音频信号的电平的示例的实施例的方框图。

图 14 显示在按照本发明的一个方面的、为了自动音量调节、系统识别进行环境噪声测量中的第一步骤。

图 15 是显示按照本发明的另一个方面的、用于补偿收听环境中的环境噪声的改变的自动音量调节的方框图。

15 图 16 显示根据对于语音和其余音频的时间跟踪信号的幅度（被称为动态幅度范围（DAR））而控制压扩的本发明的原理。

图 17 显示语音（实线）和背景音频（点线）的近似的移动平均值，其中其余音频的平均值常常超过语音的平均值，使得对话难以被理解。

20 图 18 显示调节对于图 17 所示的信号的影响，即使在节目的某些段期间 VRA 被校正，其余的音频的其它部分由于瞬变因而仍然太响。

图 19 显示这样的自动调节电平过程作用在图 17-18 所示的相同的信号上的部分，其中图上表明某些平均 DAR 如何被保持在信号中。

25 图 20 表示音频当前被设计成通过使用对话模式和多语言业务在 DVD 上被传输到最终用户的方式。

图 21 显示执行如按照本发明来进行更改的加法的译码器部件。

图 22 显示通用的家庭影院系统的预定的空间定位装置。

图 23 显示按照本发明的、其中用户具有选择自动 VRA 电平调节特性或校正的音频特性的任选可能的系统。

30 图 24 显示按照本发明的、如何实施下混合（downmix）的原理图。

图 25 显示按照本发明的这种下混合能力可被构建在被任何一种 DVD 或 DTV 应用的主要音频译码器中、或任何一种预定要译码和分配

音频信号给扬声器的其它数字音频重放设备中。

图 26 显示 Dolby 数字现有技术编码器和译码器。

图 27 显示按照本发明的每个译码的 5.1 通道上用户的可调节的电平。

5 图 28 显示按照本发明的、图 27 所示的接口方块。

图 29 显示对包括小学儿童、中学儿童、中年公民和老年公民的人口随机样本进行的两个实验的结果，测量了对于每种对象的优选的话音相对于其余音频的设置值。

#### 详细描述

10 虽然以下的说明中有很多包括结合数字杜比使用的技术的例子，但本发明可被使用于具有多种传输能力（诸如 DTS、THS、SDDS、PCM 等）的音频编码译码器。

想要的音频对其余音频的比值的重要性

从一开始，本发明首先认识到想要的音频对其余音频的比值的收听优选范围是相当大以及甚至大于所预期的范围。这个重要的发现是关于人们对于想要的音频信号电平相对所有其余音频的信号电平的比值的优选值的大量人口样本的测试的结果。

对于听力障碍的和正常听众的想要的范围的特定的调节

20 在了解正常和听力障碍用户对于不同类型的音频节目如何感受其在对话与其余音频之间的比值方面，已进行了很有针对性的研究。已经发现，人们对于话音与其余音频之间想要的调节范围有很大的变化。

对包括小学儿童、中学儿童、中年公民和老年公民的人口随机样本进行了两个实验。总共 71 个人被测试。测试包含要求用户对于足球比赛（其中其余音频是人群噪声）和流行歌曲（其中其余音频是音乐）来调节话音电平与其余音频的电平。被称为 VRA（话音对音频）比值的度量是通过把对话或话音的音量值的线性值除以对于每种选择的其余音频的音量的线性值而得到的。这些比值被画在图 29 上。对于音乐，最小值是 0.1（表示非常小的话音），而最大值是 9.54（表示很强的话音）。体育比赛测试具有 0.33 的最小值和 20.17 的最大值。

30 由于这次测试的结果，已弄清了几件事。首先，没有两个人对于

体育和音乐喜爱同一个话音和其余音频的比值。这是非常重要的，因为人口总数依赖于提供呈现给每个人的 VRA（它不能由消费者调节）的制作人。当给出这些测试的结果后，这种现象显然可以不出现。第二，所有 VRA 对于听力障碍的那些人典型地较高（为了改进可听懂度），但具有正常听力的那些人也喜爱与制作人当前提供的不同的比值。

也很重要是注意到以下事实，提供 VRA 调节的任何装置至少必须提供和从这些测试得到的一样多的调节能力，以使得它满足人口总数的一个重要段。由于话音和家庭影院媒体提供各种节目，我们应当考虑到，比值应当至少从对于任何媒体（音乐或体育）的最低的测量的比值扩展到来自音乐或体育的最高的比值。这将是 0.1 到 20.17，或 46dB 的范围。也应当指出，这仅仅是一个人口总数的采样，以及调节能力理论上应当是无穷大，因为非常可能一个人在观看体育广播时可能喜欢没有人群噪声以及另一个人喜欢没有播音员讲话。注意，这种类型的研究和对于很宽的变化的 VRA 比值的特定的希望在文献或现有技术中没有被报导。

本测试中，选择一组较老的男人，以及要求他们在固定的背景噪声与播音员的话音之间进行调节（这种测试以后对一组学生进行），其中只有后者（话音）是可变的以及前者被设置在 6.00。老年组的结果为如下：

表 I

个人	设置值
1	7.50
2	4.50
25 3	4.00
4	7.50
5	3.00
6	7.00
7	6.50
30 8	7.75
9	5.50
10	7.00

11

5.00

5 为了进一步说明所有年龄的人具有不同的听觉需要和喜好的事实，21个大学生的一组被选择来收听话音和背景声的混合物，以及通过进行选择话音电平来选择话音对背景声的比值。背景噪声（在本例中是足球比赛时的人群噪声）被固定为六（6.00）的设定值，以及允许学生调节播音员的逐场比赛的话音的音量，这些话音被分开地记录并且它们是纯话音或是几乎纯话音。换句话说，学生被选择来进行与较老的男子组相同的测试。学生被选择以便减小由于年龄造成的听觉损害。学生都是在二十岁上下。结果为如下：

10

表 II

学生 话音设置值

	1	4.75
	2	3.75
15	3	4.25
	4	4.50
	5	5.20
	6	5.75
	7	4.25
20	8	6.70
	9	3.25
	10	6.00
	11	5.00
	12	5.25
25	13	3.00
	14	4.25
	15	3.25
	16	3.00
	17	6.00
30	18	2.00
	19	4.00
	20	5.50

21

6.00

较老年组的年龄（如表 I 所示）范围从 36 到 59，占优势的个人的是在 40 或 50 岁组。正如测试结果所表示的，平均设置值趋向于合理地高，表示在交通工具上听力的某些损失。范围是从 3.00 变化到 7.75，有 4.75 的离散范围，这证实找到人们的话音对背景声的喜好的收听比值、或任何喜好的信号对其余音频（PSRA）的偏差的范围。对于两组对象的音量设置值的总的范围是从 2.0 到 7.75。这些级别代表在被用来进行这个实验的音量调节机制上的实际值。它们给出一个来自不同的用户的想要的信号对噪声值的范围的表示（当与“噪声”电平 6.0 相比较时）。

为了对于这如何与由不同的用户选择的响度变化相联系的更好的理解，考虑非线性音量控制从 2.0 到 7.75 的变化代表 20dB 或 10 倍的增加。这样，对于甚至这个小的的人口采样和单一类型的音频节目，发现不同的听众相对于“其余音频”确实宁愿要相当大的不同的电平的“想要的信号”。这种喜好项分布在各种年龄组中，表明它与个人喜好和基本的听力相一致，这是先前完全没有预料到的。

正如测试结果表明的，没有由于年龄造成的听觉损害的学生（如表 II 所示）选择的范围很大地变化，从 2.00 的低设置值到 6.70 的高设置值，有 4.70 的散布范围，几乎是从 1 到 10 的总的范围的一半。测试表明，大多数记录和广播音频信号的“一个尺度适合于所有人”的心理远不够给予各个收听者调节混合的能力以适合于他/她自己的喜好和听觉需要。另外，学生像较老龄的组一样，在他们的设置值上分布得很宽，表明在喜好和听觉需要上个人的差别。这个测试的一个结果是听觉喜好有很大的不同。

进一步的测试证实了在更大的样本组中的这个结果。而且，结果随音频类型而变化。例如，如图 29 所示，当音频源是音乐时，话音对其余音频的比值从大约零变化到 10 左右，而当音频源是体育节目时，同样的比值在大约零与 20 左右之间变化。另外，比值偏差几乎增加 3 倍，而平均值增加为音乐的两倍以上。

以上的测试的最终结果是，如果选择想要的音频对其余音频的比值和永远固定它，则多半产生一个对于大部分人口不太想要的音频节目。如上所述，最佳比值可以是短期和长期变化函数。因此，希望对

于这个想要的音频对其余音频的比值全面控制，以便满足“正常的”或非听力障碍听众的听觉需要。而且，为最终用户提供对于这个比值的最终控制，可以允许最终用户去最佳化他/她的听觉。

5 想要的音频信号和其余音频信号的、与最终用户无关的调节将是本发明的一个方面的体现。为了说明本发明的细节，考虑其中想要的音频信号是关联的话音信息的应用情况。

想要的音频信号与其余音频信号的产生

10 图 1 显示在记录的或广播的节目中把有关的话音信息与总的背景音频分开的一般方法。首先需要由节目导演对于有关的话音的定义作出决定。演员、演员组、或评论员必须被标识为相关的讲话者。

一旦相关的讲话者被标识，他们的声音就由话筒 1 拾取。话筒 1 需要是接近谈话的话筒（在评论员的情况下）或在录音时使用的高方向性长枪式话筒。除了高方向性以外，这些话筒 1 还需要音频段有限制的，优选地从 200-5000Hz。方向性和带通滤波的组合使得在录音时被声耦合到相关的话音信息的背景噪声最小。在某些类型的节目编排的情况下，阻止声耦合的需要可以通过在适合于节目的话  
15 音部分时离线地记录有关的对话话音和复制对话而得以避免。背景话筒 2 应当相当地宽带，以便提供背景信息（诸如音乐）的充分的音频质量。

20 摄影机 3 被使用来提供节目的视频部分。音频信号（话音和有关的话音）将在编码器 4 中与视频信号一起被编码。通常，音频信号简单地通过用不同的载频进行调制而使它与视频信号分开。由于大多数广播现在是立体声的，编码相关的话音信息与背景声的一个方法是把相关的话音信息复用到分开的立体声通道，这样左前和右前通道被加  
25 到两个立体声通道，以便产生四声道立体声话音的盘记录。虽然这会产生对附加广播带宽的需要，但对于记录媒体中不存在问题，只要视盘或磁带放象机被设计成可解调相关的话音信息。

一旦信号被编码，通过无论什么认为适当的装置，编码的信号就被广播系统 5 通过天线 13 发送出去用于广播，或被记录系统 6 记录  
30 在磁带或盘上。在记录音频视盘信息的情况下，只要把背景和话音信息放置在分开的录制轨道上。

接收和解调想要的音频信号和其余音频

图 2 显示用于接收和重放编码的节目信号的示例性实施例。在广播信息的情形下，接收机系统 7 从编码的音频/视频信号中解调主载频。在记录媒体的情形下，VCR 的磁头或 CD 重放机 8 的激光读出器产生编码的音频/视频信号。

5 在任一种情形下，这些信号被发送到译码系统 9。译码器 9 通过使用标准译码技术（诸如包络检测与频分或时分解调相组合）把信号分离成视频、话音音频、和背景音频。背景音频信号被发送到分开的可变增益放大器 10，收听者可按照到他或她的爱好进行调节。话音信号被发送到可变增益放大器 11，可以由收听者可按照他或她的特定需  
10 要进行调节，正如上面讨论的那样。

两个被调节的信号由单位增益相加放大器 12 相加，产生最后的音频输出。在这种情况下，收听者可以在重放音频节目时针对背景电平调节相关的话音，以使得音频节目对于他或她的独特的收听需要最佳化。当每次同一个收听者重放同一个音频时，由于收听者的听觉的  
15 改变，比值的设置可能需要改变，设置值保持成可以无限制地调节，以便适应这种灵活性。

#### 用来建立音频通道的话音识别

参照图 3，显示了本发明的再一个示例性实施例，这里是系统 100，利用话音识别芯片 101，它在识别进入的信号的语言成分以后，  
20 把信息分成主话音通道 102 和背景噪声通道 103。属于主话音那样的成分从背景成分中分开，并且被编组为分开的数据流。然后，信号被变换成模拟信号，以便建立主话音通道和背景噪声通道。

在图 2 上，主话音通道 102 被馈送到可调节信号电平的放大器 104，它可以进行话音音量调节，以及背景通道 103 被馈送到第二可  
25 调节信号电平的放大器 105，它可以进行背景音量调节。放大器 104 和 105 的输出在相加放大器 106 中被相加，以便产生最后的音频。

#### 话音识别（数字方案）

图 4 显示话音识别装置 101 的数字方案的电路图。音频信号被接收、变换成数字信号以及被话音识别器 107 分成数字话音 108 和数字  
30 背景 109 信号。数字信号然后被数字-模拟变换器 110 和 111 变换成模拟信号。

#### 针对用户缺陷定做的话音识别芯片

本发明的另一个实施例利用可编程的芯片，它通过预设备针对每个用户的缺陷被定做。例如，用户特定的频率损耗可以从进入的信号被分离出来，以及传送通过主话音通道，由此使得用户能够控制这些频率相对于所有其它频率的比值。将会看到，系统和方法可适应于模拟和数字信号。

5 对于本发明，如果该应用项是要增加最终用户的收听喜好，则有可能改变话筒1来识别要的信号话筒和改变背景话筒2来识别其余的音频信号话筒，然后，用于提供想要的信号和其余音频的、与最终用户无关的调节的装置的其余部分就和先前所讨论的完全一样。

10 本发明为最终用户提供相对于背景音频（也称为其余音频）来调节前台音频（也称为想要的信号）的电平的能力。在最一般的意义上，想要的信号代表在节目中作为当前的重点的任何的音频成分，而其余音频代表与该信号同时传送的所有其它的声音成分。应当指出，其余音频仍旧被认为是对于欣赏总的音频信号的重要的节目，但在其当前的设计上可能遮挡总的音频信号中更重要的成分。例如，管弦乐乐章可以包含相当长的长笛独奏，它是当前的重点。许多听力障碍的人对于高的频率范围（即长笛音主要处在的频率）最有困难。在大多数情况下，除了长笛，不希望完全取消所有的乐器，虽然收听长笛是这个特定的乐章的主要重点。在其余管弦乐的乐器加到长笛独奏的欣赏中

15 20 时，为了听得更清楚，最终用户需要调节长笛的电平来满足他/她自己的欣赏和听觉需要。

在总的音频信号中对于其余音频信号调节想要的信号音量电平的概念可以扩展到所有当前和将来的媒体。例如，但不作为详尽的表述，本发明可以应用于电视、CD、DVD、录象带、电影、计算机动画、和无线电、在互联网上发送的音频、快闪RAM存储器、以及计算机产生的音频。

#### 本发明的总的处理过程

图5显示按照本发明的总的处理过程的示例性实施例，用于产生对于一般的任何媒体格式需要的信号和用于传送该需要的信号到最终用户（在本例中，它是普通消费者或购买者）。实质上，在本处理过程中，有四个主要步骤。步骤1（单元51），开发要被重放给最终用户的节目。步骤2（单元52），得出想要的音频与其余音频分开。

步骤 3(单元 53), 想要的信号与其余音频这样地被传送到最终用户, 以使得保持想要的信号与其余音频的独立性。步骤 4(单元 54), 为最终用户提供调节想要的信号与其余音频的相对信号电平的能力。

#### 替换的实施例

5 对于每个这些媒体, 具有各种方法来得到(步骤 2)和传送想要的信号给最终用户(步骤 3), 当收听者对于由音频工程师确定的当前的想要的信号与其余音频(PSRA)混合感到满意时, 对他的影响最小。然而, 已经表明, 具有一般正常的听力的个人宁可控制对于至少某些形式的节目的 PSRA 混合。

10 这里描述的本发明的一个方面从对于某些媒体格式的可能的方法开始, 其中想要的信号可被传送到最终用户而不影响总的音频信号的当前混合。然后, 转到最终用户调节控制(步骤 4), 在不失一般性的情况下给出本发明的几个附加方面。

#### 想要的音频信号与其余音频信号的分开的宽带均衡

15 首先, 想要的信号与其余音频混合的灵活性的改进是以想要的信号与其余音频信号的分开的宽带均衡的形式实现的。这将为最终用户提供甚至更大的灵活性, 达到以上的他/她想要的收听质量, 此外, 这是通过使用样本增益控制达到的。

#### 在面对很剧烈的瞬态时保持 PSRA 比值

20 接着, 讨论本发明的一个方面, 其中想要的信号相对其余音频(PSRA)的调节由最终用户选择, 并且可以在存在瞬态的情况下被存储和被保持, 这种瞬态可出现在也存在于总的音频信号的背景音频中。对于完成这一点的几种方法可以详细地加以说明。

#### 通过调节背景噪声来控制 PSRA 比值

25 最后, 作为一种可以监视在最终用户所处环境中的噪声的新的方案, 本发明将允许总的音频节目、想要的信号、或其余音频信号根据环境室内噪声的增加被自动调节。这在环境噪声开始淹没想要的信号时, 将避免最终用户必须连续地调节音频节目的音量。

#### 与其余音频分开的想要的信号的传送

30 如上所述, 这些实施例的第一部分集中在图 5 的步骤 3 上。传送与其余音频分开的想要的信号的过程, 可能随每种媒体类型而有很大的变化。这里描述几种用于以这样一种格式来传送想要的信号给最终

用户的方法，即它可以使得想要的信号相对于其余音频分开地被调节。应当指出，每种这些实施例在一开始时假定：在录制音频节目期间，想要的信号被做成可供这种传送格式使用。

一般地，音频节目编排通常从单独测量想要的信号开始，随后在最终用户接收它以前，集中成单声道或立体声信号。许多这些实施例假定：想要的信号早先在录制过程中已得到，但并不限制其它可能的、在录制后得到想要的信号的方法。

#### 现有技术录制和传送过程

图 6 显示编排电视节目的音频和视频成分所选取的信号路径的非常简化的方框图。在录制过程期间 65，总的音频信号在开始时对想要的和其余音频信号分开进行测量。继续上述的“sitcom”的例子，想要的信号多半是演员的声音，其中所有其它的信号（诸如音乐和声音效果）则是其余音频。

早先在录制过程时，这些信号在音频工程师的控制下由混合器 66 来混合，以便在用于特定的电视广播频道的总的伴音载波 67 的一部分上通过空中电波或电缆进行传送。这个总的音频信号连同视频信号由最终用户在电视机 68 中接收，其中信号增益调节机构 69 允许最终用户只调节总的音频信号的音量。这代表了现有技术传送电视节目的音频信号的过程的非常简化的方案。

按照本发明的录制和传送过程

图 24 显示按照本发明的、用于传送想要的信号与其余音频的过程的示例性实施例，它允许最终用户调节各自的信号电平。在这个示例性实施例中，两个音频信号的混合不是在录制摄影棚中完成，而是由最终用户在他/她的接收机 73 完成。在录制期间，音频工程师不组合想要的信号与其余音频，而是以对于发送音频的每个通道所分配的带宽的一部分在总的伴音载波（在单元 71 中被产生）的不同的部分上分开地发送两个音频信号（单元 70）。这个音频信号被发送到电视接收机 73，在其中音频信号被解调（单元 72）、被发送到独立的用户可调节增益的部件、以及最后被相加 74，以便形成总的音频信号。这种配置允许最终用户相对于其余音频调节分开的想要的信号分量的音量。

#### 正常混合的复原

在替换例中，给出了用于为那些当前满意 PSRA 比值的最终用户复原正常混合的装置。下面将连同对于模拟方案的总的伴音载波的更详细的描述来对此进行说明。

总的伴音载波是在每个电视通道上发送的 120kHz 带宽，这打算要包含结合在该频道上的视频节目一起重放的音频节目。当前，总的伴音载波的最通用的成分是单声道和立体声部分，它们相应地被分配在从基带到  $3f_H$  范围，其中  $f_H$  是 15.734kHz。另外，第二音频节目(SAP)通道是以  $5f_H$  为中心，它的带宽比单声道和立体声通道稍小。最近，电视和 VCR 配备有音频译码器，它选择这三个音频节目中的任一个节目。虽然某些节目以立体声广播，从而利用单声道(50Hz-12kHz)和立体声带宽(50Hz-12kHz)带宽，许多电视网当前不在 SAP(50Hz-8kHz)频道上发送信息。也不存在没有被利用的总的伴音载波中的其它分段，并且在电视中当前也没有配备硬件来译码该段带宽。低带宽的“专业频道”是以  $6.5f_H$  为中心，该分段从  $3f_H$  到  $4f_H$ ，以及正好在专业频道以上的和以下的频道没有专门被分配。

根据想要的信号传输，这代表在传统的电视传输标准中存在有很大的未利用的带宽，它可包容结合正常音频节目进行发送的附加信号。

本发明的一个示例性实施例是在单声道带宽上只发送其余音频，而想要的信号是在立体声通道上发送的。这不是特别理想的，因为这使得立体声节目的传输成为不可能。

保持立体声信号的另一个替换例是：发送正常的音频立体声节目作为原先打算的计划，但在总的伴音载波的另一个分段上(诸如 SAP 通道)发送想要的信号。这需要电视机配备有同时译码立体声和 SAP 通道的措施，允许在每个通道上分开的增益调节，以及将它们相加在一起。对于正常音频节目，用户只需要选择静音 SAO 通道的任选项，或根本不译码它。想要的信号音频也可在总的伴音载波的另一个分段上被发送，而不背离本发明。

如上所述，使想要的信号与其余音频或总的音频节目同时发送，可以允许最终用户分开地调节这两个以便提供改进的收听欣赏。然而，如果不采用当前电视或 VCR 中非标准的附加硬件，则这不能实现。

对于想要的音频信号在与单声道和立体声通道分开的带宽上传

送的情况下，在应用增益电路以前将需要更复杂的解调器。当前的解调器分开地传送单声道、立体声、或 SAP 通道（其中立体声是从单声道和立体声通道相加和相减以及适当地除以二而产生的）。

#### 附加实施例

5 有一个附加方法，其中想要的信号可被传送到最终用户，它将为大众提供正常立体声节目，只需要立体声和单声道传输带宽，以及仍允许将相对较纯的想要的信号传送到最终用户。由于这个方法也可在所有其它的基于立体声（诸如 CD、音频磁带、VHS 磁带和无线电）的媒体中被执行，所以对于所有的基于立体声的媒体的可能的应用项在  
10 一般的意义上来描述这个方法。

图 8 显示一个可能的方法，其中想要的信号可以只通过利用立体声左和右通道被传送到最终用户，而仍旧保留立体声影像(image)。这个方法仍旧依赖制作方法（75）来改变先前的用于传送想要的信号的立体声混合。也存在着若干个已作出的假定，它们是通用的，但不  
15 总是正确的；以及要求对于这个方法的全部可能的实现是满意的。假定从任何音频节目的制作方法 75 产生出的信号是从若干个不同的成分产生的，其中之一是想要的信号的纯的或接近于纯的测量（这个假定也是先前对于想要的信号的电视传输方法的描述时所作出的）。另一个假定（它是先前固有地作出的，但没有明晰地阐述过）是：想要的  
20 信号位于中心音频，它相等地分布在左和右通道。

在任何的示例性实施例中，话音可容易地被摄全景处理（pan）到左或右立体声通道上，但音频工程师和音频制作人的原先的希望（对于想要的信号的摄全景处理）在这些发明的情况下将会失去。最后，对于这个特定的实施例作出的最关键的假定是：其余音频立体声  
25 影像主要是处在想要的信号带宽以外。例如，如果想要的信号正好是话音以及其余音频是特定的声音效果和音乐，则其余音频信号的从大约 200 到 4kHz（话音带宽）的频率主要是以音频为中心，以及在那些带宽以外的其余音频可以产生立体声影像。只要这些条件满足，以下的技术将提供纯的或接近于纯的想要的信号给最终用户，而不用损害  
30 正常音频节目的立体声影像。

图 8 从制作过程 75 开始，其中两个立体声信号被传送到最终用户，如图 6 所示。这里的主要差别在于，音频工程师把想要的信号(P)

摄全景处理到全右或全左（显示的）通道上，然后每个通道包含中心音频（CA）的相等的成分，而左通道也包含严格的左音频（LA）以及右通道包含严格的右音频（RA）。在这两个信号被传送到最终用户作为立体声节目（CD、音频磁带、VHS、电视等）以后，接着进行几个步骤，以便传送正常的和已调节的音频混合。

首先，每个产生的左右立体声信号用带通滤波器 76 进行滤波，以便按预定而去除那些没有被识别为包含想要的信号的频带的所有内容。再次地，对于话音，这多半将是 200Hz 到 4kHz。根据先前的假定，这个滤波过程也将去除掉作为严格的左和右音频（即 RA 和 LA）的左和右音频通道的成分，只留下 CA 和 P 在左（显示的）或右的通道和 CA 在右（显示的）或左的通道。在减去（加法器 77）这两个信号后，想要的音频信号 P 的单独的测量值被保留下来。

为了恢复以想要的音频信号为中心的 normal 音频节目，想要的音频信号的单独的测量值只需要被加回到（在加法器 79 中）右音频信号中。左音频信号在原先的传送后将保持为不变。

为了完成用户调节的音频混合，来自加法器 77 的输出的想要的信号在被加到（加法器 78）左和右正常音频节目以前被一个用户可调节的增益（W）所修改，以便形成可以由最终用户修改的位于中心的想要的信号。已调节的音频 80 因此是在加法器 78 的输出端可被提供的，而未调节的音频 81 是在加法器 79 的输出端和原先的左音频通道上可被提供的。

#### 小于无限大的摄全景处理

对于上述的方法，强调几个假定。这些假定之一可被认为是以上的过程的某些修改是不严格的。如果想要的信号的摄全景处理量是小于无穷大的已知的数值，则关于想要的信号的立体声的位置的信息可被保持。在以上的讨论中，想要的信号在制作期间被 100% 摄全景处理到一个通道上，允许通过滤波的减法产生几乎纯的想要的信号。如果它是已知的和先前在想要的信号将以某个左对右（或右对左）的有限的比值被摄全景处理到一个通道的记录和传送过程中而被建立的，则想要的信号的立体声影像在传送过程中可被重新构建。

例如，假定想要的原先的节目具有想要的信号带有一个在左通道是 20dB 和在右通道是 18dB 的影像。如果建立一个标准，该标准指令

音频工程师把想要的信号向左通道摄全景处理刚好 10dB, 则新的影像在右通道是 30dB 以及在左通道是 8dB. 减去这些新的通道导致产生一个 22dB 想要的话音信号。然后, 这被自动地调节到 10dB 的电平, 这是一个可以容易地被参考的电压电平, 如果这样的标准被付之实施的话。现在, 10dB 的想要的信号从左通道被减去和被添加到右通道, 以便复原原先的想要的立体声影像。类似的处理过程被用来产生对于已调节的音频混合的相同的立体声影像。

以上内容可被应用于利用立体声的任何媒体, 它是从多个节目的测量起源的, 其中每一个包括想要的信号的相对较纯的测量值。这些媒体可包括 CD、VHS 磁带、无线电和电视等。

在覆盖了用于传送对于某些音频媒体的想要的信号的某些成分以后, 接着描述本发明的几个更一般化的方面, 它们可被使用于支持想要的信号对其余音频的用户调节的混合的任何平台。描述了本发明的三个独立的方面, 它们为最终用户提供对 PSRA 混合更多的控制的能力、总的音频节目在面对瞬态时的特性、以及音频节目相对于收听环境下的环境噪声。不失其一般性, 本发明的每个方面可以结合在一起使用、互相结合使用、以及与上述的传送方法结合使用。事实上, 只要存在想要的信号和其余音频, 任何以下的示例性实施例可被实施来提供超过简单的音量调节的收听质量的改进。

#### 与频率有关的对 PSRA 的增益调节

为用户提供想要的信号音频 (它是独立的, 但与其余音频节目相结合) 的原先的目标之一是允许听力障碍听众改进它们对语言或其它想要的信号节目的理解。虽然相对于其余音频的音量电平来调节总的音量电平将可以为听力障碍听众 (以及正常听力的人群) 改进可听懂度和欣赏, 但它对于许多类型的听力障碍情况是不够的。本发明的一个方面寻求通过引入与频率有关的增益机制来对想要的信号相对其余音频的简单的增益调节加以改进, 该增益机制被用户控制来放大或压缩想要的信号或其余音频的某些频率分量。

图 9 显示只利用直接增益机构的示例性实施例。想要的信号 (PS) 传送到用户可调节的增益电路 92, 以及由加法器 94 将其与其余音频信号 (它被分开的可调节的用户增益单元 93 放大) 进行相加。无疑地, 这种格式将为最终用户提供比起当前在标准音频电子学中可提供

的更多的灵活性。

然而，听力障碍听众（它希望从本技术获益）可能多少会对明显的改进失望。大多数听力障碍首先出现在高频范围，典型地在 2kHz 以上。例如，取语言作为想要的信号；语言可听懂度是最受 2kHz 以上  
5 的语音频率影响的。如果收听者的听力障碍类似于低通滤波器（正如许多情况那样），只提升语音的增益并不一定能改进语言的可听懂度，因为低的频率现在太响以致于它们是不受欣赏的。

补偿在增益调节中的这种限制的最有效的方法是与频率有关的增益调节。如果最终用户可以指定的频率区域需要增加音量，则可以  
10 为听力障碍更多地专门定做声音。

图 10 显示了对于想要的信号和对于其余音频 98 的、这种与频率有关的增益装置 96。从加法器 99 输出的最终的总的音频信号是想要的信号与其余音频的组合，其中每个信号被独立地放大和被“均衡”。

均衡是指与频率有关的增益的方法，这类似于许多立体声系统中  
15 使用的方法。最基本的均衡是利用单个旋钮修改低音（低频）和高音（高频）。更精细的均衡器具有低通滤波器，它们用于从 20Hz 到 20kHz 的每个八度频带，其每个具有用户可调节的增益。

本发明的这个方面的独特的特性是：想要的信号可以与所有其余音频分开地均衡，而传统的均衡只修改总的音频信号。均衡纯语音想  
20 要的信号的一个有利的应用项可以通过提升语音的较高频率和减小其余音频的高频而被实现，以便改进语言的可听懂度，而不提升低频电平到令人不舒服的程度。与频率有关的增益部件（96 和 98）可以以许多可能的方式被实现，正如本领域技术人员明白的那样，因此这里不详细讨论。

25 面对瞬态的调节

大多数音频节目在本质上是瞬态的；也就是，响度随时地突然升高或降低。在模拟足球比赛的节目中，当主队到达球门线时球迷的激动情绪就会增长，他们的欢呼的响度也随之增长。在无线电广播这样的  
30 比赛期间，人群的噪声有时淹没了播音员的声音。然而，这通常只持续很少的瞬间，以及人群的噪声和播音员的声音的混合将返回到原先想要的节目情况，以便适合于大多数听众。这里描述的本发明假定：想要的信号（在本例中是播音员的声音）是作为与其余音频分开

的成分而提供的。

如果是这种情况，在本发明的一个附加的方面中，用于校正想要的信号或其余音频节目的瞬态改变可保持在这个节目中两个信号之间的用户选择的增益比值，这样就能确保最终用户在存在任一个信号中的瞬时涨落时不必连续地调节两个音量电平。有几种方式完成这个目标，在下面加以描述。

#### 软剪切电路

图 11 显示用于在即使存在任一个节目的瞬态改变时自动保持用户选择的想要的信号对其余音频的比值的一种技术。在想要的信号或其余音频被传送到增益调节机构（32, 33）以前，每个信号被发送到软剪切电路或算法 30, 31。正如前面那样，增益调节机构的输出被加到加法器 34，以及作为总的音频  $=w_1PS+w_2RP$  被输出到最终用户。

软剪切是指当信号电平太高时限制信号的总的幅度而不引入任何更高的频率人工产物的能力。这有时被称为压扩，正如在它的使用和可能的实施方案中很好地证明的那样。在自动固定用户调节的 PSRA 比值的环境中实施的这种技术可被看作为一种双通道压扩（或软剪切）电路，它具有控制两个限幅的信号的相对电平的能力。两个限幅器（30, 31）各自的输出可被看作为根据来自软剪切电路的某些预设定的参量而对它们的输入进行了归一化的形态。软剪切可以以各种不同的方式利用可被编程到电路或算法中的参量来完成。也可以适当地为最终用户提供针对他/她自己的听觉需要而调节这些参量中的某些参量的能力。

#### 限幅算法

考虑图 13 所示的限幅算法的一个示例性实施例。输入信号在一定的时间内被收集，以及一个矢量（一系列输入样本）被收集以及暂时被存储在贮存器 35 中。一旦矢量被填充以输入样本，就把它传送到算法 36 中，该算法可以确定峰值信号值、矢量的 RMS（有效值）、或信号幅度的某些其它函数（它也可以是频率的函数），它然后被使用来通过由除法器 37 进行的简单除法而改变输入信号。最终得到的信号在一定的意义上是被限制的原先的输入信号。

在这样的算法中至少有两个参量对于用户进行控制是重要的。被用来产生幅度测量值的矢量尺寸 37 是一个可选择的参量。对于大的

尺寸，在幅度因子 36 与实际修改的输入信号之间的延时可以变得相当大，并且导致限幅的效果被延时以及在某些情况下不明显。然而，较小尺寸的矢量不可能提供对总的输入信号的精确的代表，造成幅度因子 36 的很宽的变化，导致很宽的幅度变化，这对于收听是不愉快的。

用于确定幅度因子的方法也是影响最终用户欣赏节目的一个参量。在存在非常快速的瞬态信号时，选择矢量的最大值将大大地降低输入信号的总的电平，以及可能降低紧接在这样的瞬态后所需要听到的音频中其它信号的幅度。选择矢量的 RMS 电平也不可能提供对很强的瞬态信号足够的衰减以构成可察觉的差别。在多数情况下，峰值检测算法的选择将与节目和用户有关。所以，这些参量可以被预置或给定为对于想要的音频信号和其余音频信号的用户可选择的成分。

一旦两个信号被限幅，即使存在瞬态信号，它们的相对电平也将保持为恒定的。现在最终用户可以把分开的增益 (32, 33) 应用于这两个限幅信号的每一个，以便得到想要的信号与其余音频之间的想要的混合。任一个信号的瞬态起伏将由限幅电路来消除，而不会影响用户选择调节相对信号电平。

#### 瞬态抑制的附加实施例

图 12 上给出在存在瞬态信号时用于完成保持想要的信号对其余音频的选择比值的这个相同的目标的另一个方法。这个方法提供附加的灵活性，它支持更现实的、在想要的信号或其余音频中音量改变的重现。

首先，在节目的任一点期间当被用户选择时，每个信号传送到它自己的音量调节 (38, 39)。然后，每个信号随之被放大和被发送到矢量累加器 (41, 42)。由用户选择的增益调节值 (38, 39) 然后被用来产生单个数值比，它代表用户选择的想要的信号音量对用户选择的其余音频音量 (PSRA 比值)。

现在，一旦矢量累加器是充满的，就由一个被实现为峰值检测器、RMS 计算器等的类似于 (36) 的算法 (48, 49) 来计算适当的数值，它在某种形式上代表矢量段的瞬态幅度。这两个信号中的每一个的值然后被使用来产生出一个比值 (43)，代表在用户调节电平以后想要的信号对其余音频的实际水平。

用户选择的比值与实际的比值被除法器(44)进行比较。如果它们是相同的(正如所想要的那样),则结果是1,以及想要的信号不被乘法运算(45)修改。然而,如果实际比值大于想要的比值,则(44)的结果将小于1(或许非常小)。如果想要的信号被乘上一个非常小  
5 的数,则音量被大大地减小,这样破坏整个用户调节运行的目的。可以看到,这在具有非常小的、或不存在于其余音频的节目(诸如新闻节目)中是不希望的。

补偿这个问题的一个方法是确保(44)的结果决不小于1,这样,只响应于其余音频音量的增加。处理这个问题的另一个方法是修改其  
10 余音频信号,而不是想要的信号。通过取消想要的信号的乘法(45)和把乘法运算(47)的逆算子(46)添加到其余音频信号,就可以完成这个目标。现在,对于超过利用(38)和(39)建立的用户想要的比值的瞬态信号的增加的情况,系统将会减小其余音频信号的幅度,以便保持想要的比值。再次地,对于这种结构必须采取保护措施,从而  
15 使得想要的信号的增加不引起瞬态电平变成为太强。另外可以将用于选择矢量尺寸的参量和确定幅度方法提供给最终用户,以便在定做音频节目有更大的灵活性。

#### 用于瞬态抑制的附加实施例

在想要的或其余的音频信号中存在瞬态信号时用于自动保持用  
20 户选择的比值的最后的实施例,以均衡的形式被给出。把先前描述的、用于频率加权增益与其余音频信号分离开的想要的信号的新颖方案与自动 PSRA 比值保持功能相联系,可以导致一个能够保持在不同的频率下想要的信号对其余音频的不同比值的系统。所以,实际上,用户选择的与频率有关的增益对于不同的频带可以是不同的比值,以  
25 及自动 PSRA 保持功能将在每个这样的频率带宽下独立地运行。

这种实施方案被实现为诸如图 12 所示的系统的一组系统,每个系统被分配以不同的频段,每个增益(38)和(39)代表与频率有关的增益,它只工作在每个想要的信号和其余音频信号的自动音频带宽的一段上。

#### 30 环境噪声调节

本发明的另一个方面涉及在环境噪声的环境下进行欣赏的总的节目电平。所有先前的实施例集中在提高想要的音频电平以便使其超

过在起初传送到最终用户的总音频节目中存在的其余音频。在这里并不讨论最终用户正在欣赏节目时的环境。

在稳态环境噪声场合，用户可以只增加节目的音量来避免在他/她收听环境中存在的噪声造成可能出现的淹没。然而，经常的瞬态噪声需要用户不断地调节音频的音量高和低，以便补偿这样的噪声。

本发明的这个方面使得人工调节音频节目以补偿收听环境中存在的环境噪声的处理过程可以自动化。处理过程需要在编排音频节目的同时进行收听环境中的环境噪声的独立测量。这样做很困难，原因是因为音频节目本身也对收听环境中的总的环境噪声电平有影响。

存在有一些很好地建立的、用于提取给出某些关于系统的信息的信号分量的方法。用于确定收听环境中的环境噪声的测量装置可以采用任何的声-电换能器的形式，例如话筒。对于这里描述的系统，假定话筒被放置在正在播放音频节目到收听环境的电-声激励器装置的前面。

图 14 显示了这个系统。扬声器 151 代表放置在靠近测量室内环境噪声的话筒 152 处的电声激励器。

用于得出环境噪声的清晰的测量结果的第一步骤要求：由系统识别器 150 精确地确定从扬声器 151 到话筒 152 存在的动态特性（全部包括在内）。有许多算法方法可用于系统识别，包括利用 FIR 自适应滤波器的 LMS 算法。系统识别是对于要起作用的室内音频监视系统的性能的要求。一旦物理系统被以数字方式近似后，它就作为一个滤波器  $G_{est}$  被存储。

图 15 描述了用于根据收听环境中存在的环境噪声来调节音频节目的音量的整个系统。方块 158 代表被话筒 157 监视的收听环境中存在的环境噪声源。首先，音频节目通过  $G_{est}$  153 被滤波，以便产生只来自音频节目的近似的音频信号贡献（它是由话筒 157 测量出来的）。话筒测量值包括环境噪声和从扬声器产生的并且通过空中传送到话筒 157 的音频信号。

为了得到环境中的环境噪声的清晰的测量，由扬声器产生的音频信号的成分被加法器 154 去除掉。这只代表环境噪声信号的近似值，它由系统识别的精度限制。现在，到达话筒 157 和最终用户的音频节目分量的近似值被除法器 155 除以环境噪声电平的近似值，从而得到

一个代表音频节目的实际音量的近似值对收听环境中测量的环境噪声的比值。

5 最始，用户选择整个音频节目的总的音量（这也可以是想要的信号、其余音频、或总的音频音量控制），这是在初始调节时相对于室内的环境噪声的想要的音量电平。这个数值然后被除以音频节目对室内环境噪声的实际音量比值的近似值，以便得出一个代表想要的音量电平与近似的实际音量电平的差值的单个数值。这个比值然后用来控制混合器 159，后者输出实际音频信号，直至想要的比值与实际比值一致为止。建议按时间上的平均值来执行这个算法，诸如先前给出的  
10 矢量单元。这将使得音量上的变化变得更好听和渐变性的。该矢量覆盖的时间长度也可以变成为用户可调节的参量。

本发明的这个实施例结合先前描述的实施例一起使用时，不会受到限制。事实上，一个完全的用户可调节的音频系统多半将包括互相配合使用的每个这些部件。

15 用户对每个信号的压扩控制

以前的段落详细地讨论了在面对语音或其余音频信号的瞬态情况下用于保持用户选择的 VRA 的几个方法。与面对音频瞬态情况下保持用户选择的 VRA 的概念有关的另一个关键概念是：允许用户调节在每个信号上发生的压扩（软剪切或限幅）的量。这个特定的调节机制  
20 将允许用户保持在语音对其余音频调节时特定的改进，而同时也经受对话音或其余音频信号（或二者）的动态范围调节的变化量。

图 16 显示按照语音 161 和其余音频 162 的时间轨迹信号的幅度的这个概念，即所谓的动态幅度范围（DAR）。假定用户把语音电平 161 调节到 80dB 的平均值，而其余音频 162 的平均值是 60dB。三个  
25 不同的压扩级别被显示为对于每个信号从左到右发生，因为任一个信号或两个信号都可受到压缩。

如果把高的压缩量施加到两个信号，如在 163 中那样，则语音和其余音频的 DAR（响的和轻的）将非常紧凑。为了达到对于幅度响应的变化的更现实的响应，收听者可以调节两个信号的 DAR 级别，如在  
30 164 或 165 中所示。

在最后的情况下 165，很明显，以点线表示的某些其余音频幅度将干扰高的幅度的语音。这种情况发生在有人面对着枪射击声或很响

亮的火车声时大声叫喊。显然，这给出更现实的性能，但对于难以听见的那些人可能是不愉快的。

5 由于每个信号的 DAR 可被独立地调节，用户也可选择用于语音的窄的 DAR 和用于其余音频的宽的 DAR，而在整个节目中保持平均值水平（也由用户来设置）。

用于瞬态抑制的附加实施例

用于在完成在瞬态时保持用户选择的 VRA 比值的目标的另一个方法是：当其余音频或语音变成为太响时自动调节信号的平均值信号功率电平（SPL），以便保持用户选择的比值。

10 参照图 17，首先考虑语音（实线）171 和背景音频（点线）172 的近似移动平均。其余音频的平均音量经常超过语音的平均音量，这使得对话很难被听懂。在实现对可听懂度的某些改进的努力过程中，在语音相对于其余音频相当响的情况下用户在通过走廊时选择一个 VRA，这样他认为他达到了目标。

15 图 18 显示这种调节使用相同的两个信号，即想要的信号（或语音）181 和其余音频（或背景）182。即使 VRA 在某段节目期间已被校正，但其余音频的其它段由于瞬态仍然太响。另外降低其余音频（增加 VRA）（或甚至由于 D/A 硬件的动态范围）会使得其余音频的一些部分在存在对话时完全不能觉察。

20 为了解决这个问题，只有太响的其余音频（或太轻柔）的一些部分将被衰减（或放大），以便保持由用户在节目的给定时刻设置的 VRA。

25 图 19 显示其中这样的自动调整电平处理过程作用在以上讨论的相同的信号的段。段 193 代表其间语音被提升以便保持 VRA 为用户最初请求的最小值的时间。段 194 显示其间其余音频被衰减以便保持最小的 VRA 的时间。任一种调节或所有两种调节可以自动化进行，以确保用户保持（作为最小值）在任一个时间点设置的 VRA。

30 另外，图 19 显示在信号中某个平均的 DAR 是如何被保持的。回想起，这些信号的平均偏差可被限制在两个幅度电平上，以使得任一个或没有一个信号具有非常有限的 DAR。

与音频模式有关的 VRA 调节

回想起，ATSC 标准提供了主要的、相关的、和多语言业务。有两

种主要业务（完全主要的（CM）和音乐与效果（ME）），它们典型地是 5.1 通道，和许多相关的业务（对话（D）、听力障碍（HI）、等等），它们典型地是 1 通道。多语言业务可以通过提供多通道音乐和效果业务（它不包含对话）并结合主要在中心通道重现的、与对话（D）

5 相关的业务而达到。这两个音频业务同时被播放，以构成完整的节目。对话业务然后通过最终用户调节语言（法语、西班牙语等）来改变，而不会影响原先的音乐和效果。如果 ME 通道作为 5.1 通道节目被发送，则对话作为在中心通道上重现的单个通道被发送，只是不需要任何空间定位能力。

10 某些制作人可能被这种情况打扰；所以用于发送的第二个方法是提供 ME 和 D 作为 2-通道立体声音频，从而允许 L-R 定位其余音频和对话。当前，与对话有关的业务的使用并不是多语言被传送给最终用户的方式。所以，接下来描述的本发明的这些方面在相关的业务被利用时被预见。这将会在 DVD 之前的未来的数字电视广播中发生，因为

15 空间是不太关心 DVD 的。

本发明的这些方面通过主要业务和相关业务在译码器中被混合以前从主要业务中提取相关的业务而利用对话模式（或任何的相关的音频业务）。数字音频译码器译码包含主要的和相关的业务的比特流。（用户典型地选择想要的、除对话以外的那些相关的业务，即注

20 释、画外音等等，以及选择将被观看的节目的语言。这个信息作为元数据都被包含在 AC-3 比特流中。）

一旦译码器解压音频信号、提取元数据、和读出上述的用户输入后，硬件就选择适当的音频比特流和把它们混合在一起，以便形成总的音频节目。这些通道然后作为 5.1 通道被提供给用户，或下混合到

25 较少的通道以用于具有较少的扬声器的系统。本发明的这个方面在主要音频与相关的音频被混合以前取出主要音频与相关的音频（对话业务、听力障碍业务、视力障碍业务等等），以及把可调节的增益加到它们的每个上面。这种可调节性对可降低或提高话音或其余音频的相对电平的用户是直接可提供的。刚好在混合以前存在的任何的空间信

30 息被保持。另外，由于调节是在下混合以前执行的，具有少于 5.1 通道的用户仍旧可充分利用可调节性，而具有 5.1 通道的用户，如果必要的话，可以相对于在中心通道上的其它音频来调节中心通道上的对

话的相对电平。

在 A/54 中描述的与对话有关的业务（和其它相关的业务）总可以提供与其余音频分开的话音给最终用户。然而，按译码器被设计的方法，这两个信号在到达音频系统以前在消费者电子硬件（机顶盒，  
5 DVD 放象机、DTV、或任何的 DD 译码器）中被混合。任何总的电平或与频率有关的电平调节也应当被认为是用于任何的相关的业务的本发明的一部分。对话业务给予国际性的产品提供特定的好处，因为业务对话模式是预定能支持多语言能力的模式。另外，对话模式包含不被制作人预处理的语言（这对于正常听觉的那些人可能是不愉快的）  
10 的）。所以，对话电平调节能力将给予正常收听者以调节总的电平的能力，而听力障碍收听者可能有机会把用户特定的处理施加到正常话音上，这在使用如 ATSC 文件 A/54 中描述的听力障碍模式时是不可提供的。

另外，通过使用上述的相关的业务和主要音频业务而完成的任何  
15 VRA 类型调节可以结合上述的 VRA 保持装置而被使用。这种自动的能力对于任何相对电平调节（包括由 HI 模式所暗示的（即，在存在任一种信号的瞬态时不能提供用于保持任何的用户调节））是新的。所以，本发明的这个方面允许：（1）主要的和相关的音频业务相对电平调节，即，VRA（对于还没有采用这样的能力的那些业务）；以及（2）VRA  
20 保持装置，以这样的独特的方式，以使得它可以容易地被构建在现有的译码器设计中，而不损害比特流或需要任何附加的元数据或带宽。

图 20 显示对于本发明的这个方面的硬件和软件部件的方框图，图上显示 D 和 ME 通道如何被传送到具有 5.1 扬声器（205 的输出）、  
2 个扬声器（206 的输出）或一个扬声器（207 的输出）的最终用户。

图 20 表示当前音频被设计的方式，以便使该音频通过使用对话  
25 模式和多语言业务而被传送给 DVD 的最终用户。具有发送多语言的能力的 DVD 放象机或数字电视广播能允许用户通过元数据调节选择想要的语言。译码、元数据和混合都在芯片硬件内进行，它们或者作为译码器本身的一部分，或者在也是译码器芯片的一部分的集成电路中或  
30 作为分开的部件。（前者是最可能的、以及最便宜的。）在当前的设计中，用户不接入对话信号，因为这没有预见为必须的。

按照本发明，执行加法的译码器部件稍微被改变，如图 21 所示。

在比特流被译码器 211 译码和加上适当的元数据（包括用户选择的语言）以后，正常模式是如图 20 所示，其中相关的和主要的音频业务被混合在一起（215 的输出）。

5 现在的差别在于：在发生混合以前，相同的两个信号被直接发送到两个（或可以控制这两个信号的一个）可变增益单元 212, 213, 它们是由用户可控制的。已调节的对话和其余音频的总和 214 被传送到和下混合到如以前那样的任何数目的扬声器。用户也具有通过开关 216 在已调节的模式与正常模式之间切换的能力（通过元数据或者通过开关）。

10 本发明的另一个独特的方面是：它可最终地减小所需的带宽，以便向最终用户（有听力障碍的或其它的）提供调节对话信号（以任何语言）从而使得可听懂度对于该特定的用户最佳化的能力。当前，要注意的是输送听力障碍对话信号，它通过 HI 模式并结合包含多语言的对话模式（它是一个单独的相关音频业务）被进行处理（经过  
15 A/54），以便改进可听懂度。

如果听力障碍收听者想要听另一种语言的改进的对话，这可能是不可能的，除非再发送另一个 HI 模式语言。所以，（例如）对于两种语言，当前的实施例可能需要四个不同的对话业务来满足用户，包括：正常模式英语、正常模式法语、听力障碍模式英语、和听力障碍  
20 法语。

本发明减小这个需要的一半带宽，以及给出更多的功能给更宽范围的用户。通过只发送对话模式（法语和英语）和在译码器中将这一些信号组合以前断开两个信号（正如在以上的“VRA 模式”中显示的那样），可以提供调节电平和施加对于听力障碍的任何后处理的能力。由于处理可以按最终用户的方便性来完成，它可以按个人最想要的处理过程来定做，以及不需要制作人所设想的方式。  
25

最后，在对话与主要业务组合以前从译码器断开对话，可以允许对话被直接发送给听力障碍收听者，而不干扰听力正常的那些人的观看。这典型地通过头戴式耳机（或通过用于直接发送信号给由听力障碍收听者所戴的助听器的、红外的或其它的无线装置）来完成，其中  
30 已处理的对话被直接发送给听力障碍的个人，以及正常对话通过空中或另一组头戴式耳机被发送给正常听力收听者。

最后，本发明的这个方面提供：(1)改进的 VRA 可调节性和对于听力障碍的真正的个人化的听力增强；(2)通过对话模式的更宽广的国际观众；(3)用于数字音频的传输的减小的带宽；以及(4)一种方法，用于完成这一点，而不用修改 AC-3 比特流、不用添加加上附加的元数据、以及不用妨碍或改变编码或译码的任何分量，正如今天的标准所阐述的。

在本发明内可以有許多其它可能的实施例；其中只有几个在这里提到。这种调节能力不应当只限于音量，而也可以包括频率成形。另外，我们主要讨论将这种调节特性用于对话业务。这种能力很容易被做成对于任何相关的业务都是可供使用的。

而且，现有的讨论只集中在对于杜比数字音频的调节上。这些相同的思想可被应用于任何类似的音频格式，包括但不限于：DTS、SDDS、THX、MPEG、Dolby E 等等。

#### 中心通道话音调节

带有包括多通道放大器和 6 扬声器系统的“高端”设备的音频设施(Audiophile)，当前具有有限的独立于其它 5 个扬声器的音频而调节中心通道的音量能力。由于许多电影对话主要在中心通道上和其它声音效果在其它通道上，这种有限的调节能力允许用户提升大多数对话音频的幅度，以使得它在具有响亮的声音效果的这些部分期间是更能听懂的。当前，这种有限的调节具有两个重要的缺点，本发明的这个方面提供对这种局限性的解决办法。这些缺点是：它只是对具有 DVD 放象机和 6 扬声器家庭影院系统的消费者可提供的调节能力，允许所有的扬声器独立地调节音量电平；这种调节在话音（中心通道）和其余音频（所有其它通道）中的瞬态期间需要不断地修改；以及在一个音频段期间可接受的 VRA 调节对于另一个音频段则并不一定是良好的，如果其余音频电平增加太多或对话电平减小得太多的话。

事实是，多年来大多数消费者并没有这样一种家庭影院，它允许这样的调节能力，即杜比数字译码器、6 通道可变增益放大器和多扬声器系统。另外，在使用本发明以前，消费者没有能力确保选择的 VRA 比值对于整个节目将保持在同一个数值。

图 22 显示通用的家庭影院所希望具有的空间定位装置。虽然对于 5.1 空间通道中音频制作没有已成文的规定，但存在着一些工业标

准。这些标准之一是把大多数对话放置在中心通道 226 上。同样地，需要空间定位的其它声音效果将放置在用于左、右、左环绕和右环绕的任何其它四个扬声器，标记为 L221、R222、Ls223 和 Rs224。另外，为了避免对中间范围扬声器的破坏，低频效果 LFE 225 被放置在 0.1 5 通道，朝向次低音扬声器。

数字音频压缩允许制作为用户提供对于有可能通过模拟传输的音频的更大的动态范围。这个更大的动态范围使得在存在某些非常响亮的声音效果时大多数对话的声音太低。以下的例子给出一个说明。假定模拟传输（或记录）具有发送高达 95dB 的动态范围幅度以及对话典型地以 80dB 被记录。当其余音频达到上限而同时某个人正在讲话时，其余音频的响亮的部分就会干扰对话。然而，在数字音频压缩允许动态范围高到 105dB 时，这种情形被恶化。显然，对话将相对于其它声音保持在同样的水平（80dB），只是现在响亮的其余音频可以更现实的按照它的幅度被重现。用户对于对话水平在 DVD 上被记录得太低的抱怨是很经常的。事实上，以适当的电平的对话 IS 比起 10 15 对于带有受限制的动态范围的模拟记录所存在的电平是更适当和现实的。

即使对于当前拥有已正确校正过的家庭影院系统的消费者，对话也常常被许多今天制作的 DVD 电影中响亮的其余音频部分淹没。一小部分消费者能够通过增加中心通道的音量和或减小所有其它通道的音量而找到可听懂度的某些改进。然而，这种固定的调节只是对于某些音频通道是可接受的，以及它破坏在适当的校正时得到的电平。扬声器电平一般被校正，以便在观看位置处产生某些 SPL。适当的校正 20 确保观看尽可能真实。不幸地，这意味着响亮的声音被重现得非常响（见以上讨论的动态范围）。在深夜观看期间，这可能是不希望的。然而，扬声器电平的任何调节就会破坏校正。以下的发明将补救对于正常（校正的）观看和深夜（VRA）观看的这种情形。

对于中心通道的自动 VRA 调节特性

中心通道电平的某些增益或其余扬声器电平的减小提供了对于 30 拥有这种调节能力的 5.1 响度音频系统的那些消费者的语言可听懂度的改进。注意到，不是所有的消费者都具有这样的系统，本发明允许所有的消费者具有这种能力。

图 23 显示了这样一种系统，其中用户具有选择自动 VRA 电平特性或校正音频特性的任选项。该系统通过将开关切换到较低位置而被校正，这个位置被认为是正常工作位置，其中所有的 5.1 译码器输出通道通过功率放大器（可以是集成的）直接进到 5.1 扬声器输入端。

5 然后译码器被校正，以使得扬声器电平对于家庭影院系统是适当的。如前所述，这样的扬声器电平对于深夜观看不一定合适。

所以，通过本发明的这个方面，提供了可供使用的第二任选项（图 23 的向上的开关位置），它允许消费者选择想要的 VRA 比值，以及通过调节中心通道相对于其它音频通道的电平的相对电平而自动保持。

10 在不打扰用户选择的 VRA 的那些音频节目部分的期间，扬声器重现原先校正的格式的音频。当其余音频变成为太响亮时或话音变成为太轻柔时，自动调整电平特性只是“白费力气(kick-in)”。在这些时刻期间，话音电平可被提升，以及其余音频可被降低，或二者的组合。这是由图 23 的“检验实际 VRA”块 232 完成的。如果用户通过开关 235 选择使得自动 VRA 保持特性被启动，则 5.1 通道电平在“检验实际 VRA”块 232 中被比较。如果平均中心电平对于其它通道的电平具有一个足够的比值（它可以被反向校正，以便匹配室内声音和在观看位置处的预期的 SPL），则经过快速开关 237 通过放大器 236 重现正常校正电平。

20 如果比值被预期为有异议的，则快速开关 237 把中心通道转到它自己的自动电平调节，以及把所有其它的扬声器转到它们自己的自动电平调节。如前所述，先前的章节根据图 16-19 详细地讨论了这些调节可能性。本发明任何的这些方面也可被应用于 5.1 这些/其它扬声器电平调节。如图 23 所示。

30 按照本发明：(1) 这些自动 VRA 保持特性被直接应用于现有的 5.1 音频通道；(2) 在家庭影院中当前可调节的中心电平可以相对于其余通道被调节到特定的比值以及在存在瞬态时被保持；(3) 当用户选择的 VRA 没有被违反时，校正的电平被重现，以及当它被违反时，进行自动调整电平，由此以更真实的方式重现音频，但仍旧通过暂时改变校正来自适应于瞬态改变；以及(4) 允许用户选择自动（或人工）VRA 或校正系统，由此消除在中心通道调节以后对于重现校正的需要。（注

意添加上可变动性)

也注意到, 虽然电平被称为是自动调节的, 但该特性也可被禁止, 以便提供简单的增益调节, 如图 23 所示。

用于下混合到非中心通道扬声器装置的中心通道调节

5 许多消费者在相当长的时间内将不拥有家庭影院系统。然而, DVD 播放机正变得越来越流行, 以及数字电视将在不久的将来进行广播。这些数字音频格式将需要最终用户具有 5.1 通道译码器, 以便收听任何的广播音频, 然而, 他们可能没有购买完全可调节的和校正的、带有 5.1 音频通道的家庭影院系统的奢侈能力。

10 本发明的下一个方面利用这样的事实: 制作人将会传送 5.1 音频通道给这样的消费者, 他们可能没有充分的重现能力, 但同时仍旧允许他们调节话音, 以保持音频 VRA 比值水平。另外, 本发明的这个方面通过允许用户选择这些特性而被增强, 即它将在没有 6 扬声器可调节系统的情况下保持该比值。

15 这可以通过一种用来接收来自 DVD 放象机的输出端口的 DD 比特流的接口单元的装置来实现, 这样, 另一个类似的装置提供按照用户选择的 VRA 来调节中心通道电平的装置, 定做的音频译码器, 其后面跟随信号到立体声、四通道的下混合, 或任何其它的、不提供中心通道扬声器的扬声器装置。

20 图 24 显示如何实施这种下混合的一个概念性图。

图 25 显示这种能力可被构建在主要音频译码器中, 后者由任何 DVD、DTV、或任何其它的打算译码和传送音频信号到扬声器的数字音频重放装置所采用。用于非家庭影院音频系统的下混合提供一种使所有用户从可选择的 VRA 获益的方法。已调节的对话被传送给非中心通道扬声器, 以使得音频节目的预期的空间定位尽可能地原封不动。然而, 对话电平将只是较高的。接着讨论进一步的细节。

25 对于将 5.1 音频通道(杜比数字)下混合到 4 通道(杜比编程逻辑(Dolby Pro-Logic))、2 通道(立体声)、或 1 通道(单声道), 已有规定好的指南。5.1 通道按适当的比值的适当的组合已经被选择来产生对于消费者拥有的无论哪种重放系统的最佳空间定位。现有的下混合方法的问题是: 它们对于最终用户是透明的, 但不能由最终用户来控制。这会造成在较新的 5.1 通道音频混合中利用动态范围的方

式时出现的关于可听懂度的问题。

例如，考虑被制做在 5.1 通道上的一个电影，其中有一段其余音频淹没了对话，因而使得很难听懂。如果消费者拥有 6 扬声器和 6 通道可调节增益放大器，则语言可听懂度可以如上所述地被改进以及被保持。然而，只拥有立体声重放的消费者将接收 5.1 通道的下混合版本，它遵循图 26 所示的图（取自杜比数字广播实施指南）。事实上，中心通道电平被衰减了一个量，它是在 DD 比特流中规定的（-3, -4.5, 或 -6dB）。这将进一步减小在包含在其它通道上的响亮的其余音频的一部分的可听懂度。

10 本发明的这个方面通过在把它们下混合到用户的重放设备以前把可调节增益放置在每个空间通道来绕过下混合处理过程。

图 27 显示在每个译码的 5.1 通道上的用户可调节电平。典型地，低频效果（LFE）通道的下混合没有被进行，以免电子部件的饱和以及减小可听懂度。然而，通过下混合发生以前可提供的用户调节，有可能把 LFE 包括到按用户规定的比值的下混合中。

15 许可用户调节每个通道的电平（电平调节器 276a-g），就能允许消费者使得任何数目的重现扬声器能够利用先前只提供给具有 5.1 重现通道的那些人的话音电平调节。

20 如上所述，这个设备可被使用于外部的任何 DD 译码器 271，不管它是否独立的译码器，是在 DVD 内、或在电视机内，这与在家庭影院系统中重现通道数目无关。用户只需要命令译码器 271 传送 DD(5.1) 输出，以及“接口盒”将执行先前由译码器执行的调节和下混合。

25 图 28 显示这个接口盒 282。它可获取任何 DD 译码器的 5.1 译码的音频通道作为它的输入，施加独立的增益到每个通道，以及按照消费者具有的重现扬声器的数目进行下混合。

另外，本发明的这个方面可通过在进行任何下混合以前把独立的用户可调节通道增益施加到每个 5.1 通道，从而结合到任何的译码器中。当前的方法是，按需要进行下混合，然后加上增益。这不能改进对话可听懂度，因为对于任何下混合情形，该中心通道被混合到包含其余音频的其它通道中。

30 也应当指出，前面讨论的自动 VRA-HOLD（保持）机构是完全可应用于这个实施例的。一旦 VRA 通过调节每个放大器增益被选择，VRA

保持特性就应当保持下混合以前的该比值。由于比值是在收听任何下混合的重现设备时选择的，下混合电路中的定标将通过由消费者施加的附加中心电平调节被补偿。这样，由于下混合处理过程本身，不必要附加的补偿。

- 5 也应当指出，在用户调节的放大和下混合以前的中心通道的带通滤波将去除那些频率比语音低的声音和频率比语音（例如 200Hz 到 4000Hz）高的声音，以及可以改进在某些通道中的可听懂度。也很可能在中心通道上对于改进的可听懂度去除的那些内容也存在于左和右通道，因为打算把它们用于重现音乐和效果，否则它们是在语音带  
10 宽以外。这将确保不发生其余音频声音的保真度的损失，而同时也改进语言的可听懂度。

- 本发明的这个方面：(1) 允许消费者使得任何数目的扬声器能够利用现在可提供给具有 5.1 重现扬声器的那些人的 VRA 比值调节；(2) 允许那些相同的消费者相对于其它通道上的其余音频设置中心通道  
15 上的想要的对话电平，以及通过 VAR 保持特性使得在收听时该比值保持为不变；以及(3)可以被施加到 DD 译码器的任何输出，而不修改比特流或增加需要的传输带宽，即，它是与硬件无关的。

#### 听力障碍语音处理

- 本发明的另一个方面包括在数字音频译码器中有利于听力障碍  
20 的特别的话音处理。需要音频信号的特殊滤波或频率加权。来自图形用户接口的滤波器的用户调节将大大地简化这些调节。

- 另一个方面将是独立地修改对于对话通道（可能的中心或纯数字，这取决于选择的方法）和其余音频的频率响应的幅度。这实际上是在每个通道上提供图形均衡器，以便得到最大的可调节性。然而，  
25 它也可包括某些用于幅度修改的固定的形状，诸如对话的高频放大（用于听力障碍者），以便改进可听懂度，而同时其余音频的高频将被衰减。

- 其它的有效的处理（包括用于改进可听懂度的话音的特定的空间定位）也是可能的。例如，对话可以易于被听力障碍者明白，如果它  
30 以立体声格式给出而不是以更一般的中心通道单声道格式的话。

显然，在屏幕节目（OSP）方面将有一个对于以上的构想的任选方案，因为它们都可以作为译码器的一部分或作为独立的方块被实

施。旋钮、拨号盘、按钮等等也可被使用来完成这些目标。

手工控制可包含用来完成任务的 1 个或 2 个旋钮，其中一个用于 VRA 或两个用于 VRA 的两个，一个用于 VRA 偏移范围或两个用于 VRA 偏移范围。

## 5 数字多通道译码器

这里揭示的本发明的这个方面为（主要）数字多通道音频译码器的用户提供相对于特定的背景音频信号音量电平和空间位置通用地调节特定的前台音频信号音量电平和空间位置的能力。

10 对于只包括语音或对话节目的前台信号，这个技术提供动态地改进语音可听懂度的能力。这是通过使用多个独立的通道来传送编码音频信号而完成的，这些编码音频信号在某种程度上是根据音频内容被规定的。结果是得到在使用预规定的数字译码器技术时为用户提供这种特定的调节能力所必须的一组分量。

15 描述了数字译码处理的附加部分和修改，它们可以提供这种最终用户调节能力。描述了需要对于数字编码信号进行某些修改的几种任选方案，以及只需要编码信号进行最小修改的任选方案。

给出一种使用于数字译码器的算法，它将为用户提供必要的调节能力，以便适当地改变对于很宽范围的节目内容的语言可听懂度，以及还给出了媒体传送方法。

## 20 背景声音

数字音频编码技术的最近进展激发在将来的音频广播和记录技术中的广阔范围的改变。杜比数字（AC-3）标准已被接受使用于数字电视，以及是当前（1998）的用于 DVD 的有竞争力的音频格式之一。AC-3 或类似的多通道编码技术很有可能将在下一个十年被教导用于  
25 数字音频记录和广播的几乎每个方面。许多新的特性也被包括在 AC-3 标准中，它把各种消费电子产品给最终用户，使人们可以控制先前不能控制的音频节目的各个方面。

30 本发明的这个方面的目标是提供先前对于最终用户在娱乐方面所不能提供的进一步的特性，由此改进和扩展在现有的音频译码器技术中当前采用的特性。

实际的编码和译码处理方案已由杜比和东芝给出，它们被用于 AC-3 编码器/译码器设计。同样地，其它有竞争力的多通道音频编码

格式已被规定并且被证明能提供五个或更多音频通道，而同时节省信号带宽和所需要的贮存媒体。因为这些编码技术已被开发以及其中许多已成为用于音频传输和贮存的标准，所以，改变实际的编码和译码处理并不是本发明的具体目标。

- 5        在本说明和以下的说明中，“编码”和“译码”是指在比特流被变换成数字音频以后的比特流的压缩和解压。它也包括其它非音频数字信息的处理，该信息通知译码器关于音频节目的类型和可提供的任选项。

10        虽然本发明的一个实施例寻求把音频和非音频信息承载信号加到现有的编码信号中，但它将遵循对于特定的媒体阐述的编码标准来这样做。例如，如果本发明被应用于数字电视（DTV）广播，并且用于DTV广播的音频标准是AC-3，则本发明的这个方面将遵循在这个格式的说明中所阐述的标准。为了实施本发明其它的实施例不需要关于编码处理本身的任何特定的附加信息。

15        所以，给定任何多通道音频格式以后（主要是“数字形式”，但也不排除模拟形式），这里揭示的特定的用户调节技术将以各种格式被实施。关于该技术的信息可被引入到编码比特流中。可以实现用户调节技术而不必特别地修改编码比特流。另外，该技术可以在任何的多通道数字编码方案中被实施，而不失去一般性。

#### 20        译码器的实施例

25        该技术的主要推进力寻求为最终用户提供相对于“前台”音频通用地调节“背景”音频的音频音量电平的能力。用于任何广播或记录音频的特定的实施方案可以具有前台音频作为讲话人的话音，而背景音频是任何的非话音的其余音频分量，诸如电视演出中的声音效果。如果背景音频太响亮，则它就干扰最终用户鉴别语言成分的能力。本发明的这个方面寻求把这种调节能力应用于数字音频译码器中，由此使得该技术易于在各种媒体格式被大范围的音频消费者接受。

30        以下的说明使用Dolby AC-3音频编码方案作为一个例子，以用于本技术的数字方式实施。然而，存在许多其它音频编码方案（包括DTS、THX、SDDS、线性PCM、MPEG-1、MPEG-2等等）。本技术的数字方式实施的本发明并不限于AC-3，而可结合其它音频标准使用。所以，它不失去结合其它音频标准使用的一般性，下面通过使用现有的

和已规定的 AC-3 格式来进行描述。

AC-3 格式的附录 C 说明了多通道环绕音频信号的另一个使用举例，它通过使用配备有译码和识别卡拉 OK 编码方案的能力的 AC-3 译码器来实现卡拉 OK 表演。（卡拉 OK 是一种用于娱乐的方式，其中来自歌曲的音乐被演奏出来而不具有唱歌内容，从而使业余演唱者可以唱歌，作为专业录音的艺术的替代。）附录 C 描述卡拉 OK 察觉译码器，它被配备来处理输入到卡拉 OK 编码的多通道内容。在多通道信号被译码和被形成独立的数字音频信号以后（被称为 L-左立体声音频、R-右立体声音频、M-“引导旋律”、V1-“话音轨道 1”、和 V2-“话音轨道 2”），可以得到以下的代数运算：

$$L_k = L + a * V1 + b * V2 + c * M$$

$$C_k = d * V1 + e * V2 + f * M$$

$$R_k = R + g * V1 + h * V2 + i * M$$

最终得到的信号（被称为  $L_k$ 、 $C_k$ 、和  $R_k$ ）是分别相应于左、中心和右音频通道的卡拉 OK 输出信号。系数（a, b, c, d, e, f, g, h, i）被使用来调节他们所修改的信号的电平。这些系数的实际的数值被放到具体的实施方案中，以及只在用于卡拉 OK 实施方案的附录 C 被讨论。对于有关本算法（方程组）的另外的发明，具有几个任选方案，它们可为最终用户提供完全不同的最终结果。

如前所述，想要的结果是通过（1）与背景（声音效果或其它次要的音频节目）音频分量无关地提升前台（想要的或语言）音频分量，或（2）与前台音频分量无关地降低背景音频分量，从而为最终用户提供改进音频节目的语言可听懂度的能力。

在当前的实施例中，现有的算法不能完成这些目标中的任一种，因为没有建立对于背景音频（L 和 R）的明显控制，以及前台音频不能相对于背景音频被提升。涉及每个卡拉 OK 译码器信号的特定的内容的其它限制还会阻止最终用户具有这样的能力，即不可能允许最终用户通过调节话音与其余音频的相对电平（此后，称为通用话音对其余音频（UVRA）调节）而通用地调节音频节目，以便改进语言可听懂度或听力增强。（注：“话音”通常是想要的前台信号，它包括被包含在音频记录或广播中的主要的信息。然而，术语“话音”不应当限制那些被认为可以作为前台材料使用的信号类型。）

本发明的一个方面提供在重新规定 AC-3 卡拉 OK 觉察译码器的现有的功能时的 UVRA 调节能力。因为其它多通道编码译码器可以具有类似的卡拉 OK 风格的特性，可以对于任何类似的装置作出这些修改，而不背离本发明原先的意图。通过使用用于从信号 L, R, M, V1 和 V2 产生输出通道的现有的算法，人们可以规定相应于音频节目（广播或录音）的某些成分的每个通道的特定的内容，以及提供系数 a, b, c, d, e, f, g, h 和 i 的独特的调节能力，以使得语言可听懂度可以通过由用户调节所述系数而被改进。

完成这个目标的一个可能的方法是首先使得所有这些系数可通用地调节（实质上从 0 到无穷大，其中无穷大相应于译码器和音频硬件（例如 D/A）的物理限制）。

第二，应当指出，L 和 R 在用来给出这个第一发明实施例中的原先的方程组中不能被调节。如果 D/A 具有的是不带有自动定标（寻址下一个）的固定的输出范围，这将会把限制加在每个通道的信号内容上。V1、V2 和 M 通道是可通用调节的，因此可以具有提供者想要的任何内容，只要话音（前台）保持与背景分开。例如，M 可以包含所有的前台（想要的、或话音）音频成分，而 V1 和 V2 可以包含声音效果、背景音乐、或笑声。对于这些情况，L 和 R 必须不包含音频信号，或包含处在足够低的电平的音频，以使得其它音频成分可被调节为高到足以遮挡 L 和 R 音频的电平，这样，它几乎不能被觉察。这个方法允许最终用户调节 a, b, c, d, e, f, g, h 和 i，以使得话音电平可以与任何其他音频节目无关地被固定，和被定位在三个扬声器位置（L<sub>0</sub>, R<sub>0</sub>, C<sub>0</sub>）的任一个，其中左输出、右输出和中心输出被规定为：

$$L_0 = L + a * V1 + b * V2 + c * M$$

$$C_0 = d * V1 + e * V2 + f * M$$

$$R_0 = R + g * V1 + h * V2 + i * M$$

先前提到的是输出定标（或自动定标）的概念以及它结合在以上的方程和说明中描绘的已译码的（但未输出的）左和右通道上未规定的（或规定的）内容一起使用的概念。对于 L<sub>0</sub> 信号给出的几个基本的例子只显示在不规定 L 和 R 通道的内容的情况下对于以上的工作程序所需要的自动定标的建设的操作。

为了通过以上的译码信号的线性组合完成 UVRA 目标，用户需要

提升 V1 的系数 (“a”) 超过 L 和/或 M 到足够的水平, 以便达到所需要的 VRA 比值。首先假定 L 和 M 包含相同的内容, 该内容是按提供者想要的比值的整个节目 (前台 (语音) 和 (背景) 其余音频)。现在假定 V1 通道包含与原先的节目同步的语音 (对话) 内容。如果用户想要有 6dB VRA, 则他/她将选择 “a” 为 4.0, 因为其余音频由 L 和 M 代表两次:

$$L_0 = L + 4.0 * V1 + 1.0 * M$$

$20 * \log_{10}(4.0/2.0) = 6\text{dB}$ 。现在, 假定 D/A 硬件的 “全尺度” 用 0dB 代表。如果 L、V1、和 M 的每一个处在全信号尺度 (0dB, 或线性尺度 1.0), 则 L 的线性幅度变成为

$$L_0 = 1.0 + 4.0(1.0) + 1.0(1.0) = 6, \text{ 或 } 15.5 \text{ dB.}$$

由于全尺度是 0dB, 这个电平将使 D/A 饱和。如果通过将  $L_0$  除以它的峰值来应用自动定标, 则将可以对  $L_0$  除以 6.0 从而得到以下的式子:

$$L_0 = 0.1667 * L + 0.6667 * V1 + 0.1667 * M$$

注意到, VRA 比值仍旧是所想要的

$$6\text{dB} = 20 * \log_{10}(0.6667 / (0.1667 + 0.1667))$$

但  $L_0 = 0\text{dB}$  以及不会使得 D/A 饱和。用户达到 URVA 调节的目标。然而, 关键的是要指出, 这个例子是为自动定标的一定的方法和先前的每个译码信号的一定的内容准备的。

现在给出另一个例子, 它进一步概括每个通道的可能的内容。假定 L 和 M 不包含相同的内容以及不能同时被重放。(一个更一般化的方案, 它放宽对信号内容的假定) 然而, V1 通道仍旧只包含节目对话或前台, 以及 M 通道只包含其余音频或背景。所以, 一种只有 M 和 V1 的用户调节的混合将导致想要的 UVRA 调节能力。即使以上所示的算法没有提供对其余 “L” 信号的明显的调节, 我们能够通过有改进地使用自动定标处理而有效地改变系数。(注: 这个例子可以被概括成被使用于任意两个信号, 只要它们其中之一具有音量控制器以及自动定标按这里所描述的那样进行工作。)

为了通过上述的这些信号完成 HEC 目标, 需要做两件事情:

把语音提升到超过其余音频一个想要的电平 (比如说 6dB)。

把组合的语音 (V1) 和其余音频 (M) 提升到比没有调节能力的 L

信号更高的电平，以及 1.0 的固定的系数。

首先，为了“M”通道上的其余音频足以遮挡左通道音频“L”，假定 M 通道的电平必须超过 L 通道 20dB。（它可以是更大或更小，取决于每个通道上的节目类型。）所以，M 通道的系数（“c”）必须是 10.0，因为“L”通道的系数是 1.0（ $20 \times \log_{10}(10.0/1.0) = 20\text{dB}$ ）。现在为了使话音信号比其余音频大 6dB，系数“a”必须是 22.0，因为  $20 \times \log_{10}(22.0/11.0) = 6\text{dB}$ ，在此，11.0 是 M 通道和 L 通道功率的总和。总的方程变成

$$L_0 = L + 22.0 * V_1 + 10.0 * M$$

10 因为全尺度对于  $L_0$  被设置为 1.0，显然，以上的总和必须如前面那样被定标，以避免 D/A 饱和。L 以 33.0 来定标  $L_0$  得到的结果是：

$$L_0 = 0.0303 * L + 0.6667 * V_1 + 0.4545 * M$$

注意到，所有的信号的总和是单位值 1 (0dB)，VRA 是 6dB，以及前台加背景内容 (M+V<sub>1</sub>) 相对未知的内容 (L) 的 VRA 比值是 20dB。这说明，所规定的自动定标的方法可被如何地使用来允许原先的代数关系组产生 UVRA 调节，而不用明显地控制 L 的系数。这建立了提供最终用户 UVRA 调节能力，而不影响用于线性组合的标准化方法。附加的改进的成分包括通道内容说明、通用参量调节和输出信号自动定标。

20 下一个实施例包括开发完全新的算法，该算法可结合多通道音频编码器来使用，以便传送 UVRA 调节能力给最终用户，而同时仍旧保持在这样的编码译码器中许多固有的影像能力。主要的目标是使得节目提供者（记录器、广播、或其它）只在一个或多个这些可提供的多通道上传送任何类型的音频节目的仅仅是话音的信号，而同时但独立地传送在其它通道上的其余音频成分。

30 如果 UVRA 调节能力是要被构建在各种译码器和媒体上的，则需要达成某些标准化。以下的算法寻求提供该标准或这样的标准的某个形式。正如所看到的，这与话音实际上在哪个通道是发送没有多少关系，因为所有的分量都是可调节的。（这为这样的译码方案提供了有利的特性，因为 UVRA 能力可被忽略而不干扰正常的节目。）多通道音频的目标是为最终用户提供环绕声，这使得音频节目看来更逼真。希望为最终用户提供调节 VRA 的能力，而不会很大地影响声音质量。

可以完成这一点的一个可能的方法是通过使用以下的算法：

$$L_0 = a * L + b * C$$

$$C_0 = c * L + d * R + e * C$$

$$R_0 = f * R + g * C$$

$$5 \quad L_{s0} = h * L_s + i * C$$

$$R_{s0} = j * R_s + k * C$$

译码的音频信号由  $L$ ,  $R$ ,  $C$ ,  $L_s$ , 和  $R_s$  规定。译码器的总的输出信号（下标为“0”）是译码的音频信号的代数函数。开发以上算法的方式提供了最大的可调节性、对于位置和电平的 UVRA 调节、以及通常  
10 对于正常的多通道音频节目产生的在环绕影像中的最小干扰。

在本实施例中，音频信号  $C$  必须被取为前台音频。假定前台音频被认为是语音或语音，用户可通过相对于其它的用户可调节的系数改变  $b$ ,  $e$ ,  $g$ ,  $i$  和  $k$ , 而调节语音的电平和位置。左音频输出将包含左信号，右音频输出将包含右信号，以及对于随后的环绕信号和输出也是  
15 这样的。

这里的主要差别在于，中心音频输出现在是原先的左、右和前台信号的组合。实际上，常常会使得中心通道主要地由左和右分量，即不需要任何类型的影像的声音组成。

在这个设计中失去某些空间能力是不可避免的，因为 5 个原先的  
20 通道之一必须被纯前台占用，以便使用户去控制相对电平。也很重要的是要注意到，对于 UVRA 系统的理想的性能， $L$ ,  $R$ ,  $L$  和  $R$  应当只包含背景音频成分。这允许用户完全互相独立地调节电平。然而，有可能这些信号包含前台信号的分量，只要他们被适当地与前台音频（在本说明中在中心通道）同步，以避免回声和/或频谱抵消。

25 通过重新规定某个原先的译码的音频信号的特定的内容，这些方程的其它形式都是可能的，这将最终互相独立地和几乎独立地完成背景和前台音频 UVRA 和空间定位的相同的目标。只要对于前台信号提供相对于和独立于背景信号的调节能力，UVRA 调节的目标（音量和空间位置）将被完成，以及本发明的目的将被实现。

30 应当明白，以上的程序可在不用对译码器本身作任何特别的修改的情况下实现。如果译码器信号的线性组合在译码器硬件内实现，则显然也必须对它们的具体算法进行修改。

然而，如果节目提供者提供上述的信号给编码器以及在线性组合以前译码器为最终用户提供粗略地译码的音频信号，则这是不必要的。在本实施例中，第二个硬件可被使用来执行这些信号（它可以是模拟的或数字的）的线性组合以及为最终用户提供特定的调节能力。

5 替换地，可以执行对数字编码的比特流的修改以便通知译码器：特定的“UVRA 组织的”信号是（任选地）可提供的。这可以允许译码器向最终用户提供在正常的多通道音频节目与特定的 UVRA 可调节的多通道信号之间的选择。如果用户选择 UVRA 音频数据流而不是正常的音频数据流，则调节能力可以在 D/A 操作之前在外围单元或在译码器本身内部被完成。

其它信息也可由节目提供者附加到比特流中，从而通知最终用户关于前台音频所处的通道，想要的空间安排，以及假定的用于适当地定位或放置前台和背景音频信号的系数。

15 总之，本发明的这个方面依赖于传送与背景音频分开的前台音频的能力。许多的多通道音频对这样的信号的传送进行格式化，然而他们没有规定在空间性以外的每个信号的内容。通过本发明的这个方面，对于空间性的某些牺牲，最终用户可以接收相对于背景音频通用地调节前台音频的能力，以便至少给出一个明显的结果：在前台音频是话音和背景音频是可能经常遮挡话音的噪声信号的情况下，可以改进语音可听懂度。

20 本发明的这个方面可以在各种音频编码格式下被提供，以使得它也包括有限制的（包括空间定标能力）和/或无限制的通道调节系数、规定的和/或大概规定的通道的内容说明、修改译码器比特流和/或不修改比特流（在与音频节目提供者合作之下），以及提供替换的、但有用的、关系到组合或接收正常多通道音频节目的 UVRA 信号的使用的信息。

虽然可以在附属权利要求的范围内对于本发明作出许多改变和修改，但这样的改变和修改属于权利要求范围之内，因此是被其所覆盖的。

30

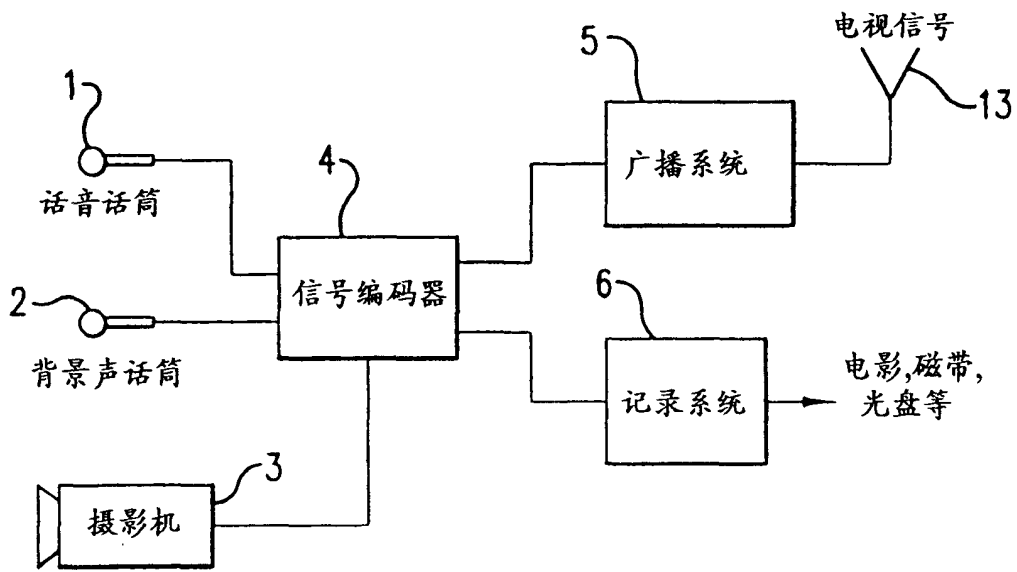


图 1

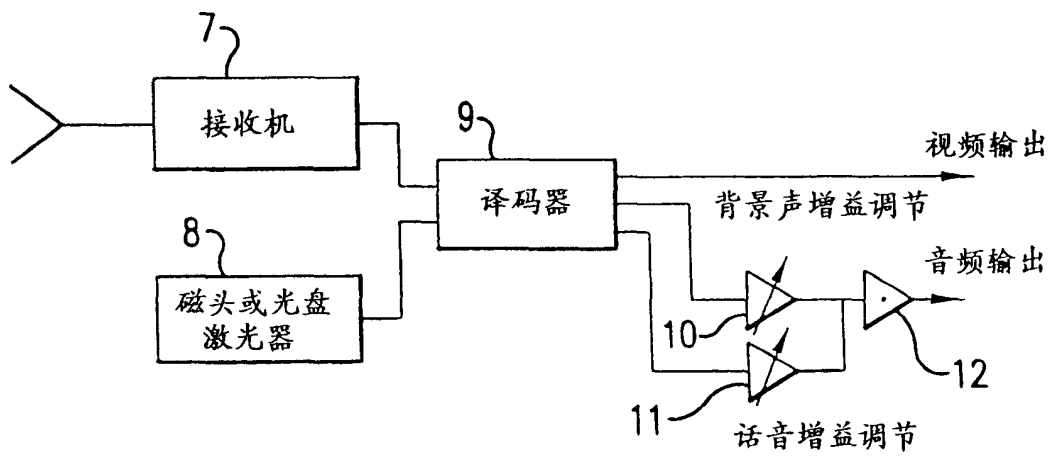


图 2

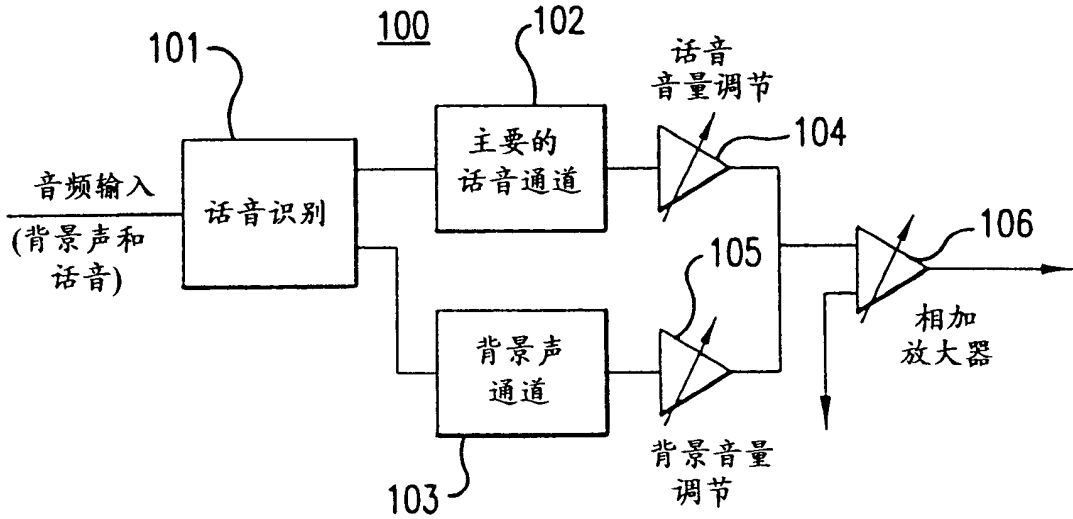


图 3

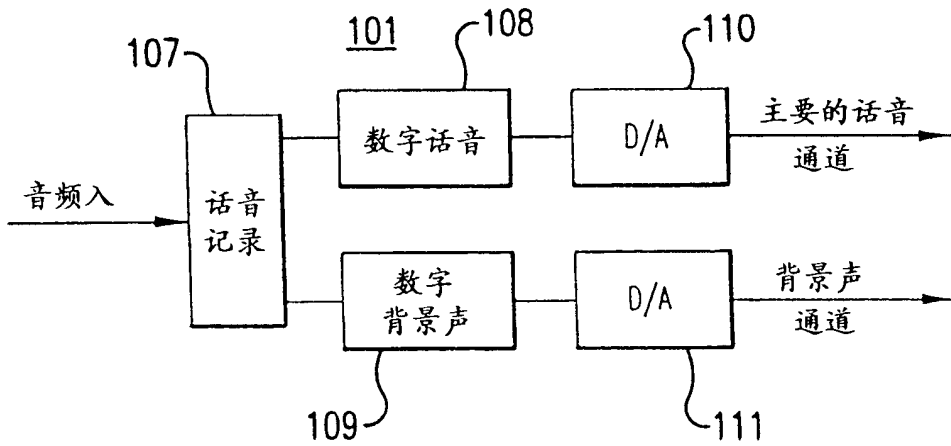


图 4

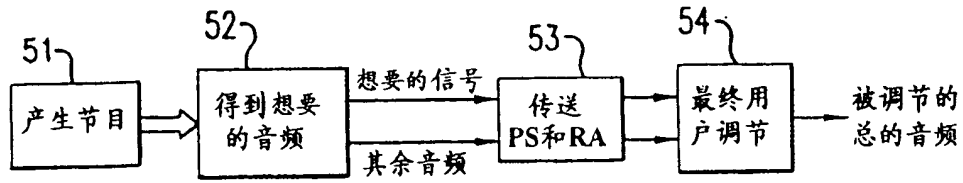


图 5

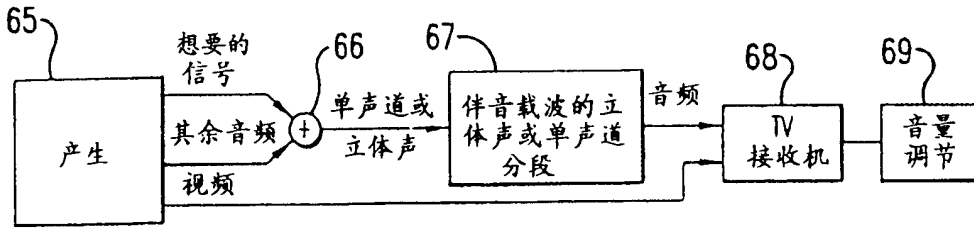


图 6  
现有技术

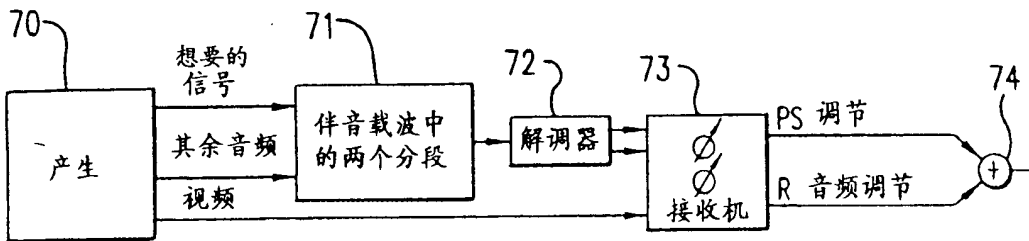


图 7

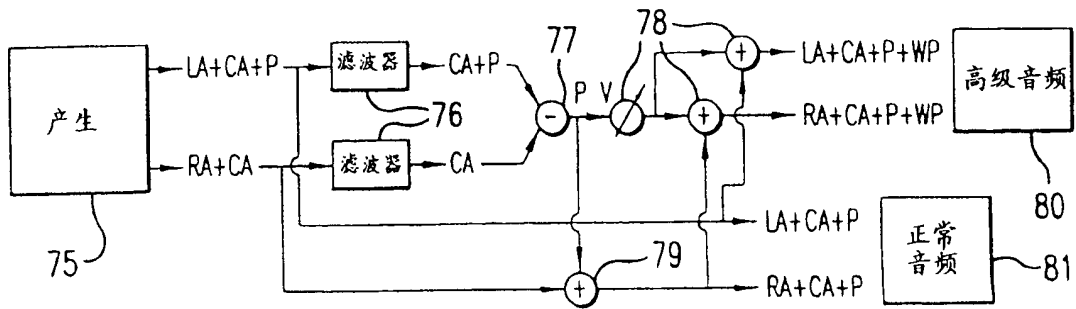


图 8

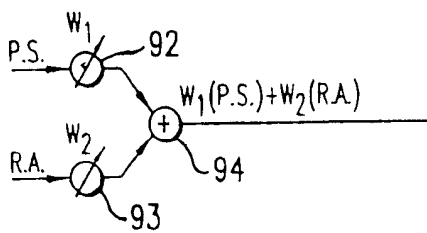


图 9

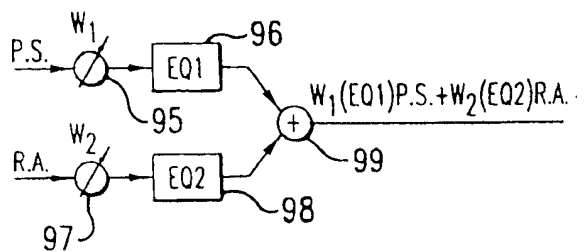


图 10

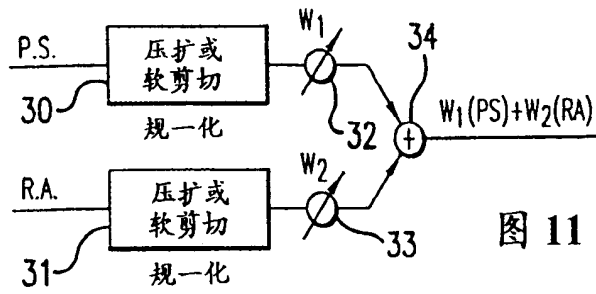


图 11

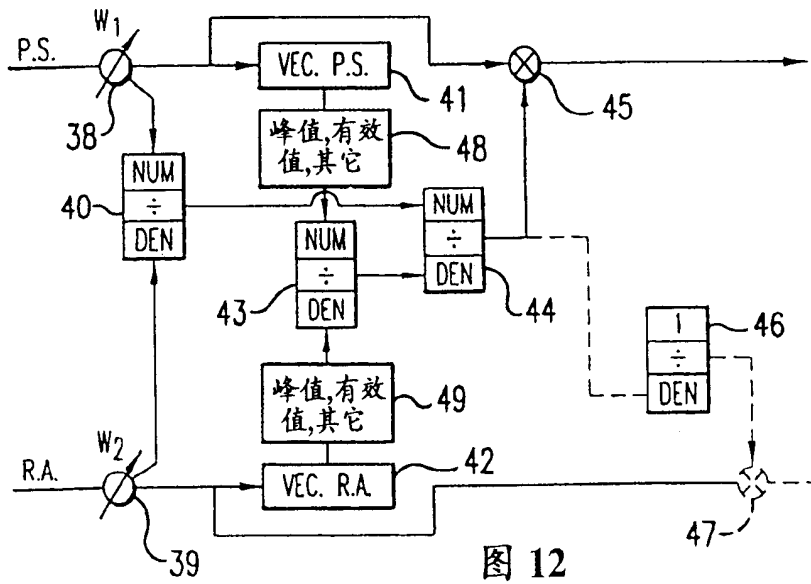


图 12

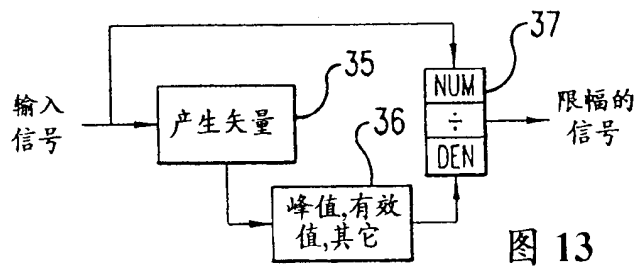


图 13

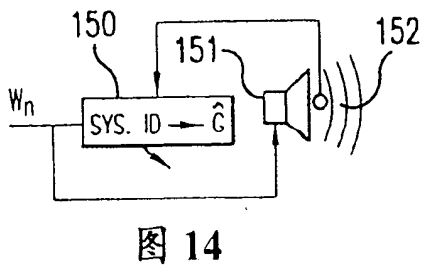


图 14

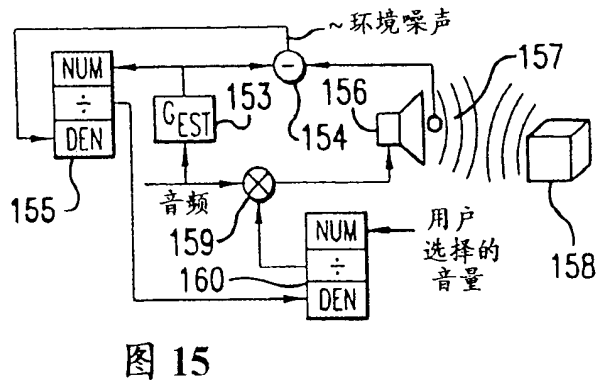


图 15

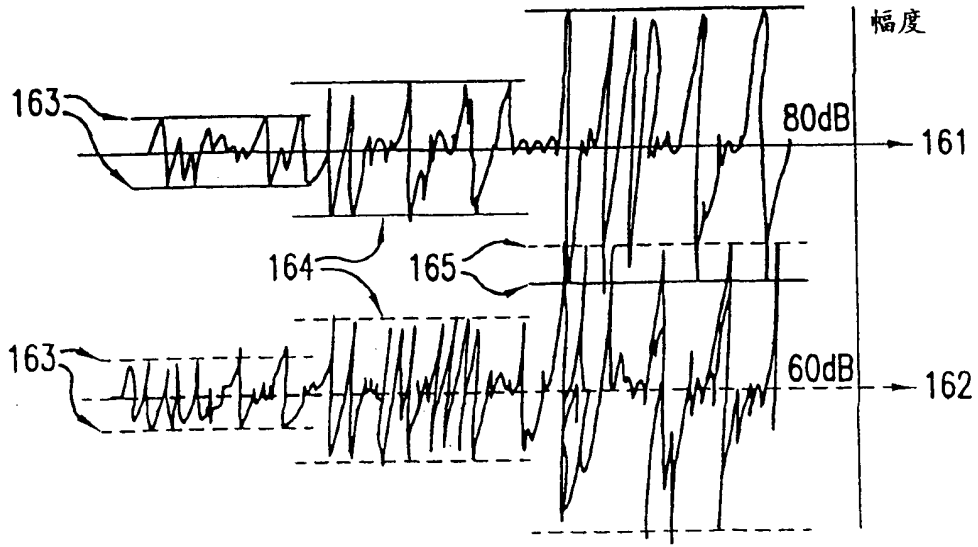


图 16

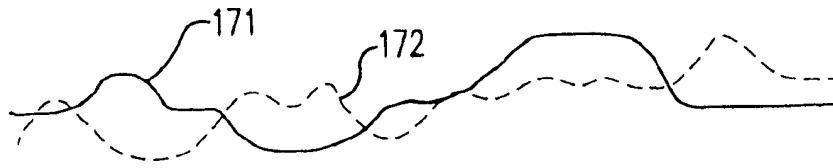


图 17



图 18

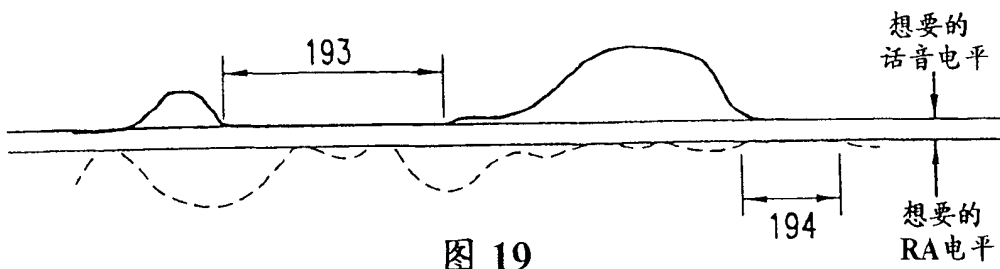


图 19

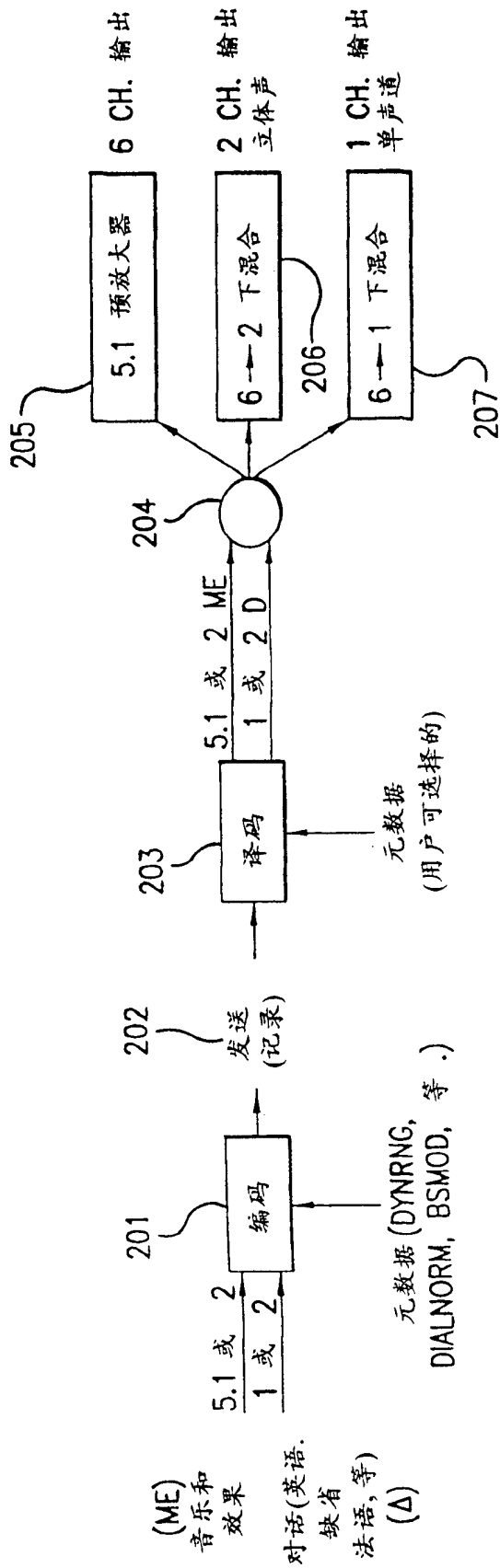


图 20  
现有技术

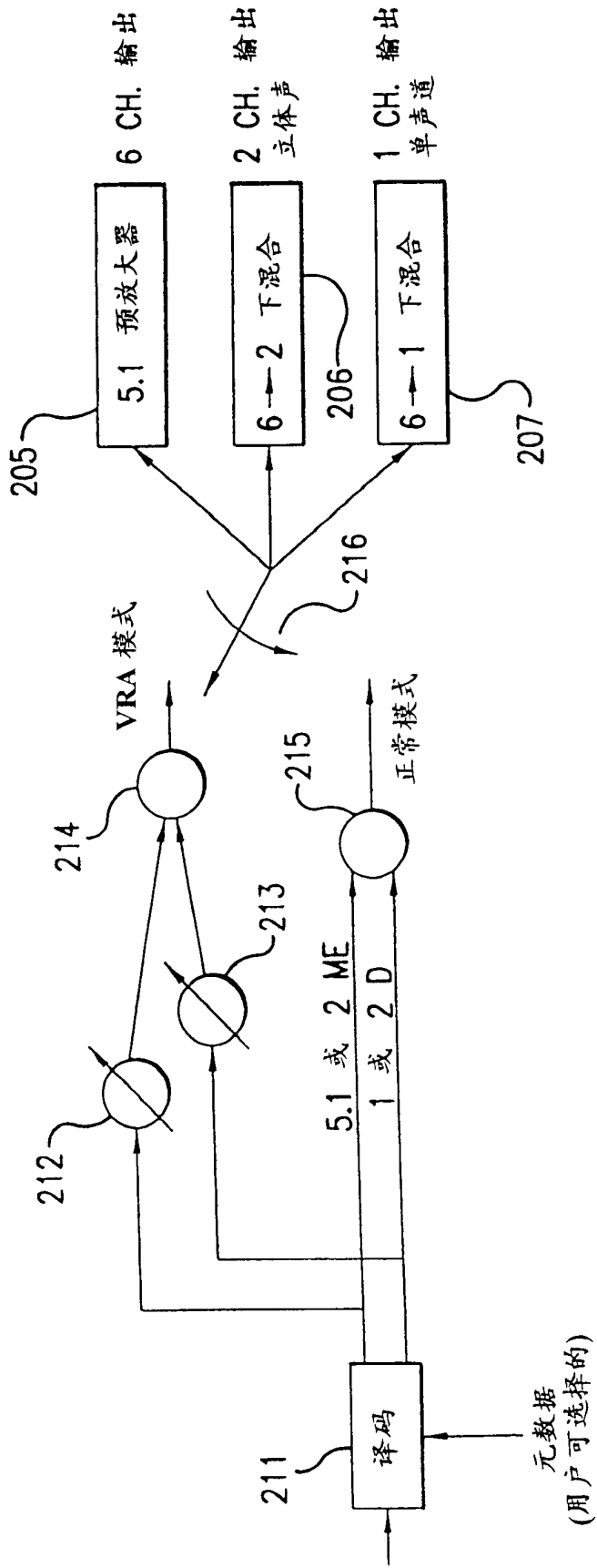


图 21

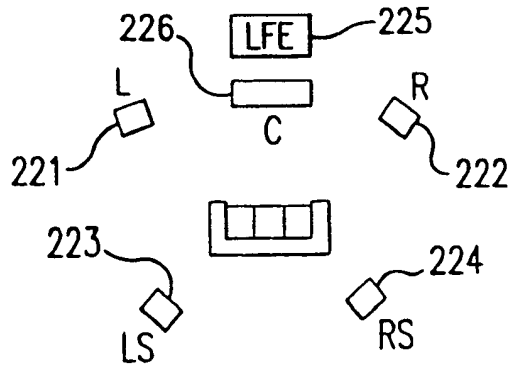


图 22

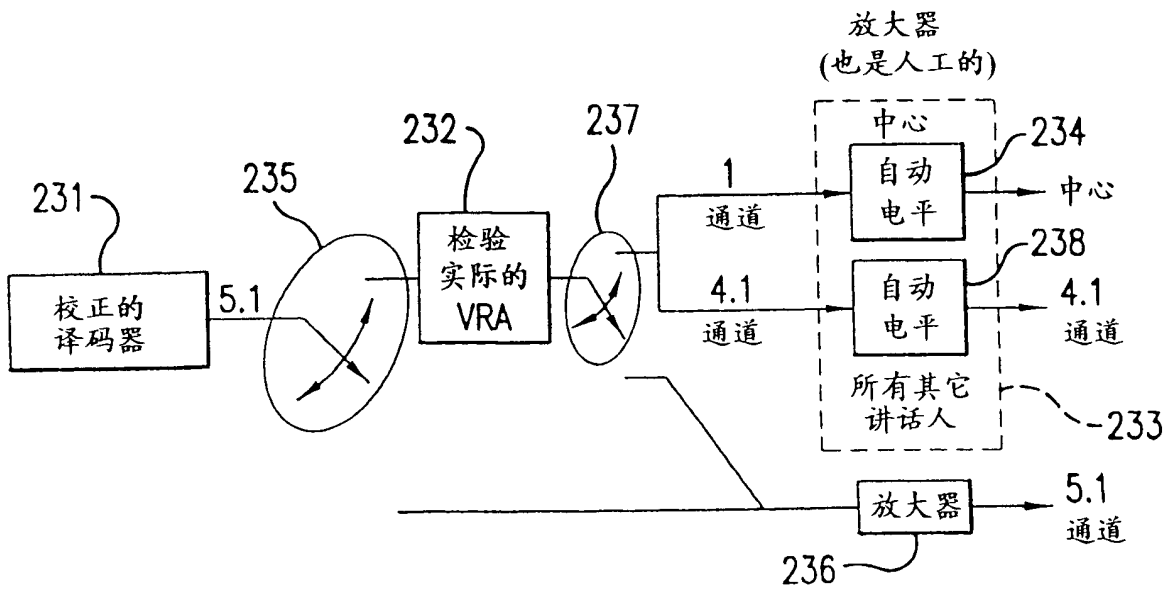
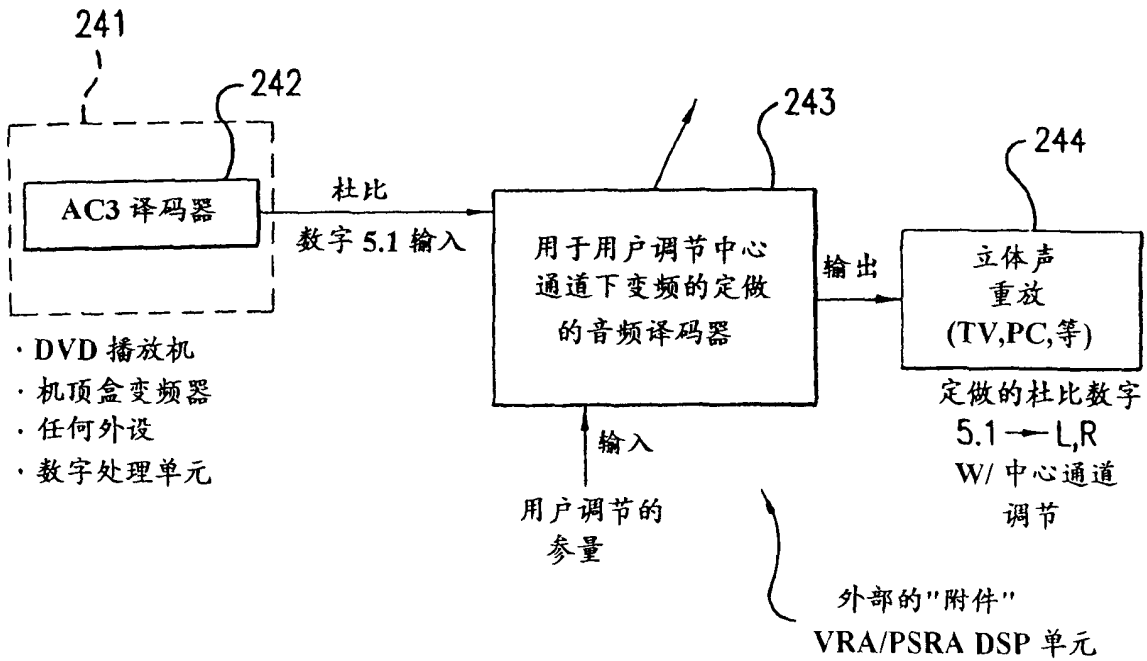
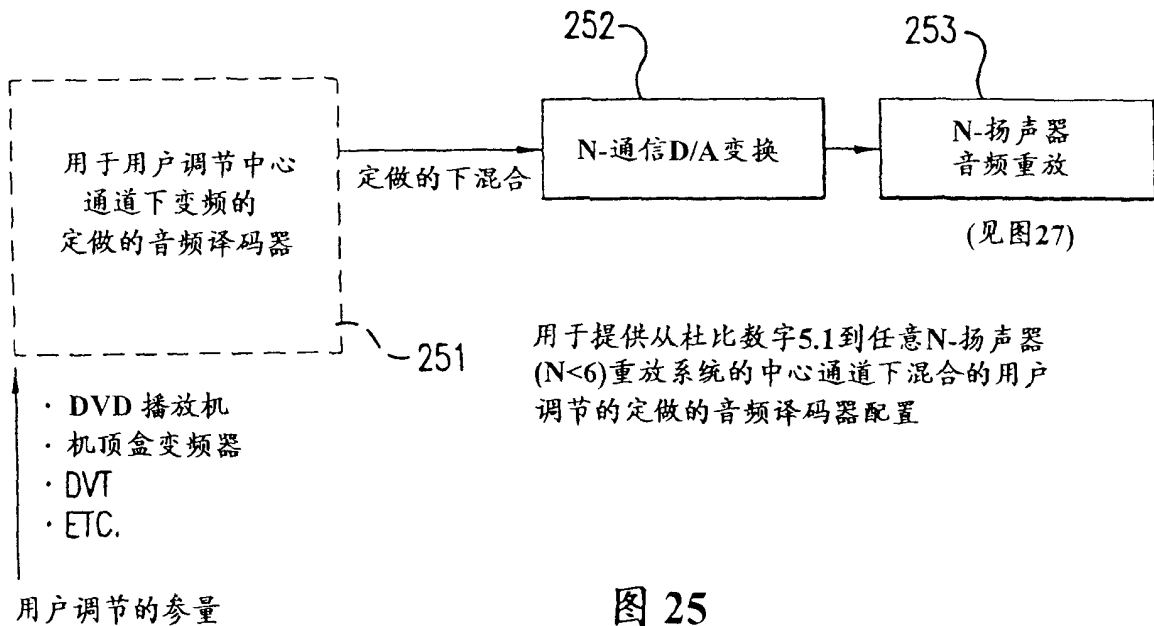


图 23



\* 用于提供从杜比数字5.1到立体声(L,R)的中心信道下混合的用户调节的外部配置

图 24



用于提供从杜比数字5.1到任意N-扬声器 (N<6)重放系统的中心通道下混合的用户调节的定制的音频译码器配置

图 25

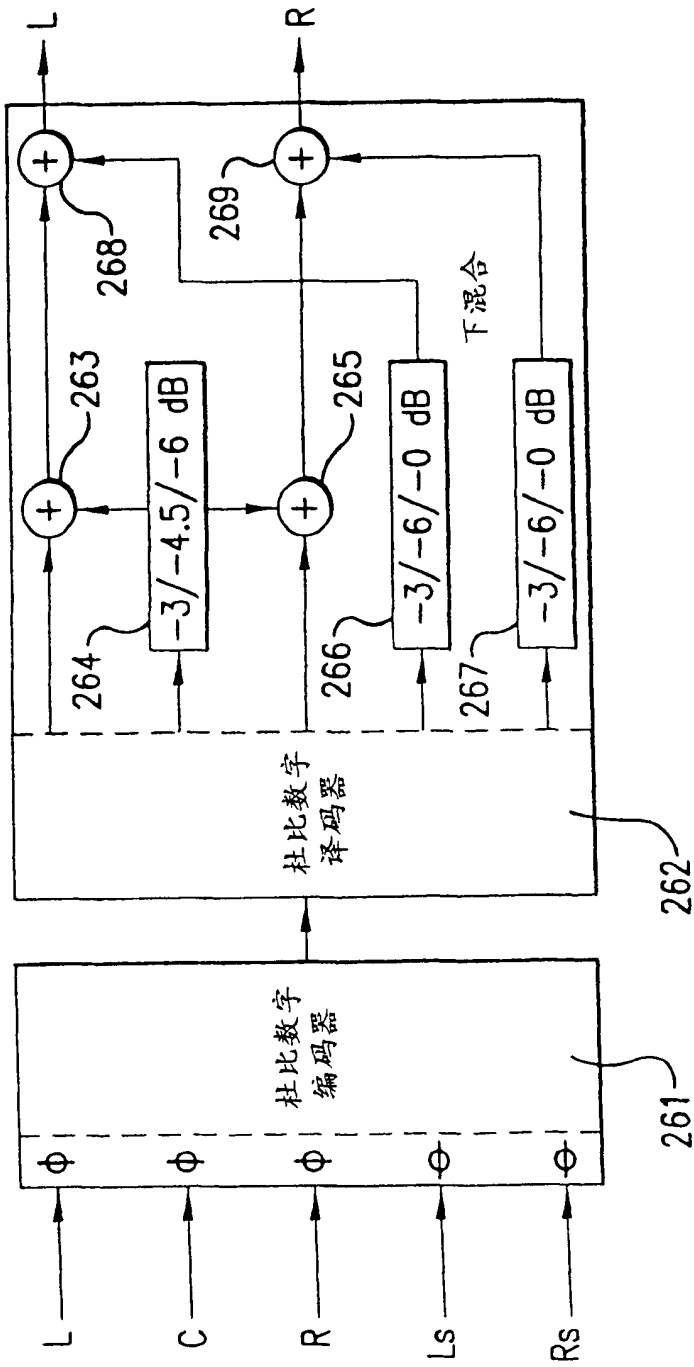


图 26  
现有技术

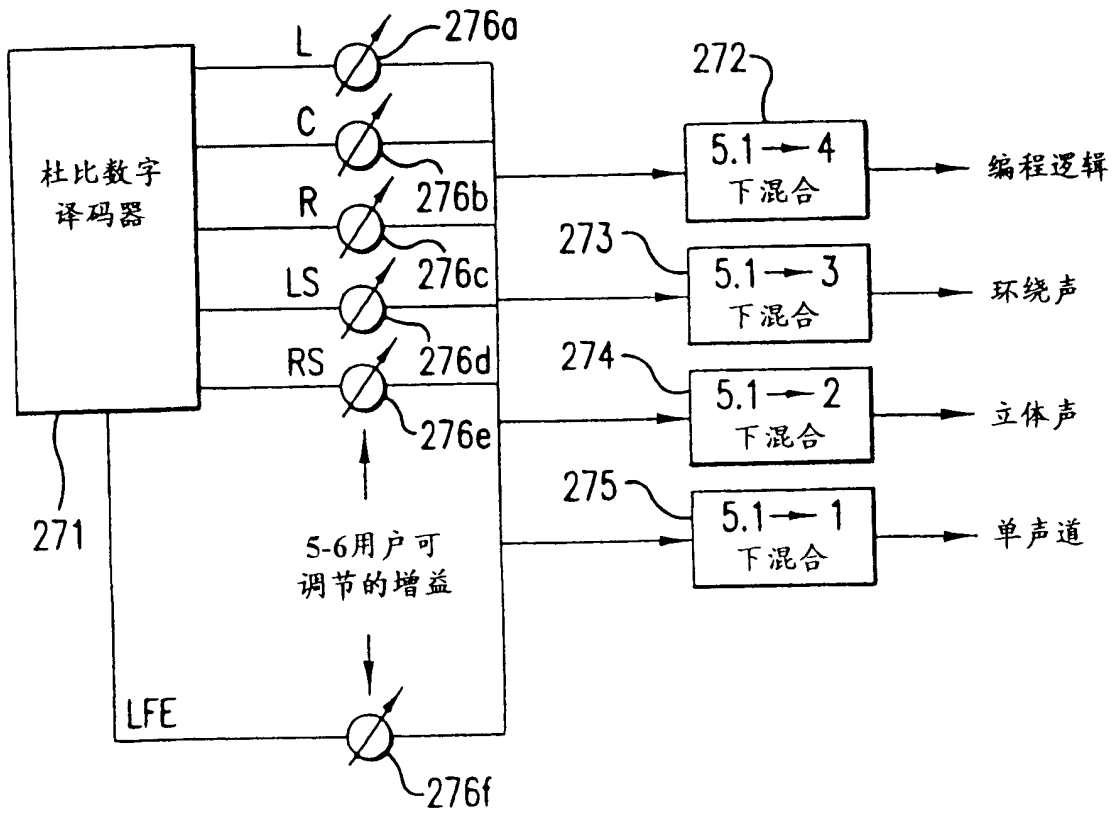


图 27

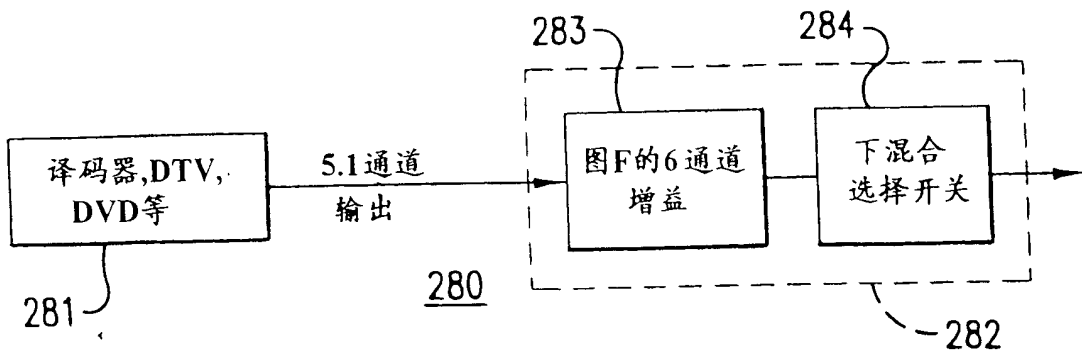
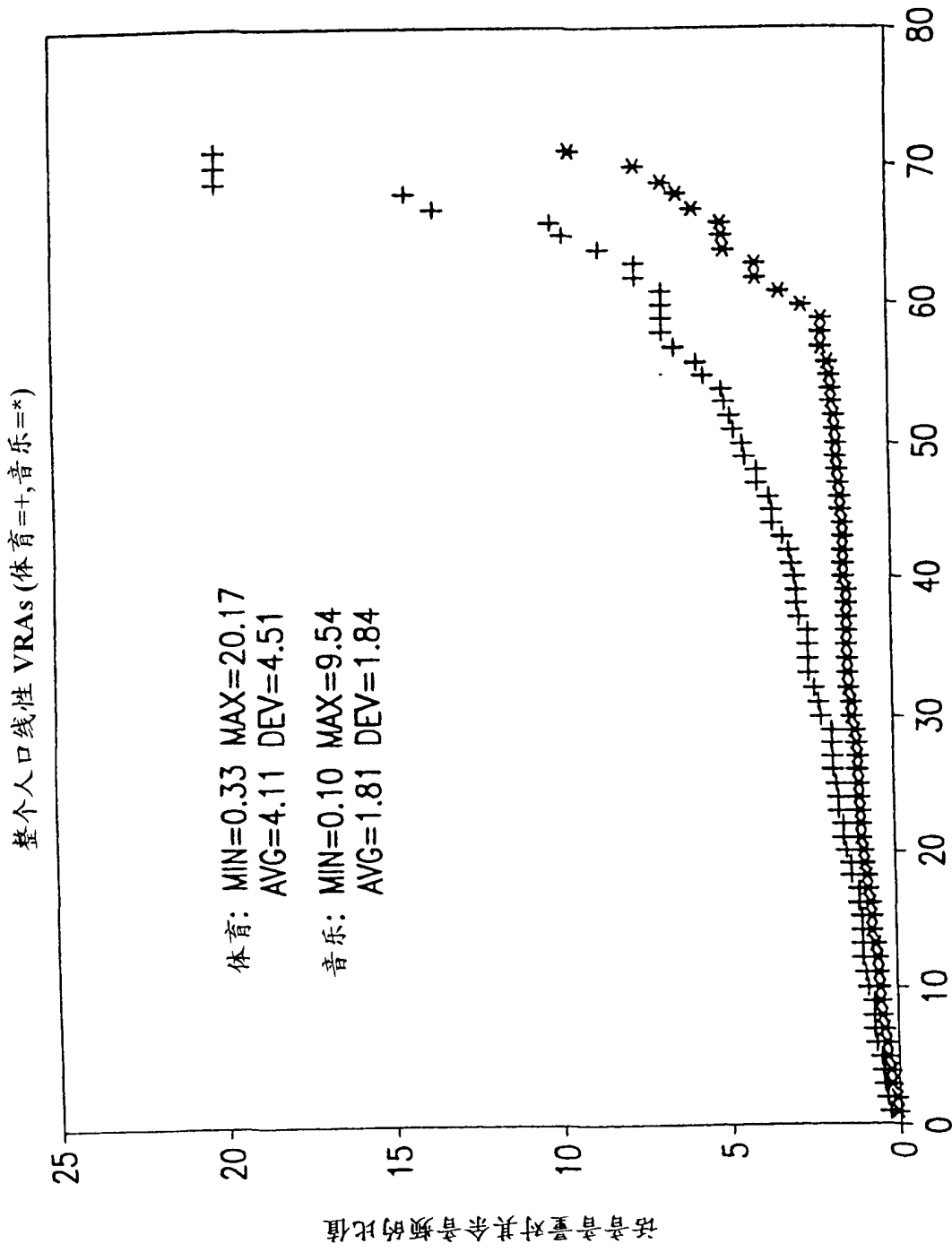


图 28



预存储的测试目标 (总共71个)

图 29