



[12] 发明专利说明书

专利号 ZL 01807567.3

[45] 授权公告日 2005 年 8 月 3 日

[11] 授权公告号 CN 1213556C

[22] 申请日 2001.1.30 [21] 申请号 01807567.3

[30] 优先权

[32] 2000. 2. 4 [33] US [31] 60/180,220

[32] 2000. 5.26 [33] US [31] 09/580,206

[86] 国际申请 PCT/US2001/002908 2001.1.30

[87] 国际公布 WO2001/058064 英 2001.8.9

[85] 进入国家阶段日期 2002.9.29

[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 M·A·沃德富 W·R·绍恩德斯

审查员 张 璇

[74] 专利代理机构 中国专利代理(香港)有限公司

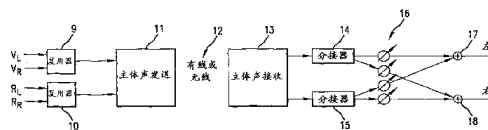
代理人 程天正 陈 霁

权利要求书 7 页 说明书 35 页 附图 27 页

[54] 发明名称 消费者应用中用于处理音频信号的电影院系统和相关方法

[57] 摘要

一种用于向多个用户提供话音对其余音频(VRA)的调节能力的方法包括在第一解码器(14)处接收话音信号和其余音频信号,以及同时在第二解码器(15)处接收话音信号和其余音频信号,其中话音信号和其余音频信号是分开接收的;并由每个解码器分别调节被分开地接收的话音信号和其余音频信号。



1. 一种电影院系统，它以这样的方式输出电影给多个收听者中的每个收听者，即：允许对位于电影院环境的观众中的多个收听者进行个性化的音频音量调节，该系统包括：

5 视频设备，向观众显示电影的视频部分；

扬声器系统，向观众发送电影的相应的音频部分；

10 贮存媒体的一个或多个单元，该贮存媒体存储电影的视频部分，电影的相应的音频部分，作为声音信号且具有空间信息的一个或多个声道的第一音频信号，以及包括除了第一音频信号的音频内容以外的音频内容和具有空间信息的一个或多个声道的第二音频信号；

发射机，把来自贮存媒体的一个或多个单元的第一和第二音频信号与电影的视频部分和相应的音频部分同步地发送到多个个人收听设备，其中多个个人收听设备的每个收听设备与在电影院观众中的多个收听者的每个收听者相联系，而每个个人收听设备包括：

15 第一接收机，它与扬声器系统无关地接收所发送的第一音频信号；

第二接收机，它与扬声器系统无关地接收所发送的第二音频信号；

20 第一调节设备，它根据来自用户的输入而调节该第一音频信号的音量；

第二调节设备，它根据来自用户的输入而调节该第二音频信号的音量；

25 音频信号组合设备，它把第一音频信号的空间信息声道与第二音频信号的相应的空间信息声道组合以产生组合的音频信号；以及

一个或多个换能器，它接收组合的音频信号，把组合的音频信号转换成声音，并输出声音，使得它们可被在电影院观众中、与多个个人收听设备中每一个相联系的多个收听者的每个收听者听见；

30 其中该系统允许多个收听者的每个收听者能够独立于观众中其他的多个收听者而调节该第一和第二音频信号。

2. 如权利要求 1 中的电影院系统，其中第一音频信号和第二音

频信号中至少一个是单声道的信号。

3. 如权利要求 1 中的电影院系统，其中第一音频信号和第二音频信号中至少一个是立体声信号，该立体声信号具有左和右空间信息声道。

5 4. 如权利要求 1 中的电影院系统，其中第一音频信号和第二音频信号中至少一个是环绕声信号，该环绕声信号具有包括左，中置，右，和一个或多个环绕声声道的空间信息声道。

5. 如权利要求 1 中的电影院系统，其中第一音频信号和第二音频信号中的至少一个是多声道环绕声信号，该多声道环绕声信号具有包括左，左环绕，中置，右环绕，右，和一个或多个环绕声声道的空间信息声道。

6. 如权利要求 1 中的电影院系统，其中第一和第二调节设备是音量可控制的有源放大器。

7. 如权利要求 1 中的电影院系统，其中第一和第二调节设备是音量可控制的无源衰减器。

8. 如权利要求 1 中的电影院系统，其中第一调节设备与第二调节设备被组合成单个音量控制设备。

9. 如权利要求 8 中的电影院系统，其中当音量控制设备向一个方向移动时，第一音频信号的音量增加，而第二音频信号的音量降低，以及当音量控制设备向另一个方向移动时，第二音频信号的音量增加，而第一音频信号的音量降低。

10. 如权利要求 1 中的电影院系统，其中第一接收机接收包括第一音频信号的第一数字比特流，以及第二接收机接收包括第二音频信号的第二数字比特流，该系统还包括：

25 第一解码器，它将第一数字比特流解码；以及
 第二解码器，它将第二数字比特流解码。

11. 如权利要求 1 中的电影院系统，其中第一接收机和第二接收机被集成为单个接收机。

12. 如权利要求 11 中的电影院系统，其中单个接收机接收包括第一音频信号和第二音频信号的单个数字比特流，该系统还包括：
30 单个解码器，它把单个数字比特流解码。

13. 如权利要求 1 中的电影院系统，其中第一和第二接收机接收

无线发送。

14. 如权利要求 1 中的电影院系统，其中个人收听设备至少是立体声头戴式耳机，单声道耳机，助听器，和辅助性收听设备中的一种。

5 15. 如权利要求 1 中的电影院系统，其中个人收听设备是穿戴在身上的接收机，该穿戴在身上的接收机提供组合的音频信号给一个或多个电声换能器。

16. 如权利要求 1 中的电影院系统，还包括波导和放大器中的至少一种，用来增强该组合的音频信号。

10 17. 如权利要求 1 中的电影院系统，还包括处理器，它计算第一音频信号的音量对第二音频信号的音量的比率，其中第一调节设备、第二调节设备、和音频信号组合设备中的至少一个自动地调节和保持第一音频信号音量对第二音频信号音量的比率。

15 18. 如权利要求 17 中的电影院系统，其中处理器计算在有限的时间间隔内音频信号的标准偏差。

19. 如权利要求 17 中的电影院系统，其中该比率被存储在存储器中，供音频信号组合设备使用。

20. 如权利要求 17 中的电影院系统，其中第一和第二调节设备由用户通过图形用户接口进行控制。

20 21. 如权利要求 17 中的电影院系统，其中第一调节设备和第二调节设备被耦合到单个用户可控制的音量调节设备，该单个用户可控制的音量调节设备用来自始至终地调节组合的音频信号的音量，以使得单个用户可控制的音量调节设备按第一方向移动时增加第一音频信号级的音量和降低第二音频信号的音量，而按第二方向移动时增加第二音频信号的音量和降低第一音频信号的音量。

25 22. 如权利要求 1 中的电影院系统，其中电影的相应音频部分是作为第一音频信号与第二音频信号而被存储的。

23. 如权利要求 1 中的电影院系统，其中第一和第二调节设备是通过使用计算机软件或计算机硬件而工作的。

30 24. 如权利要求 1 中的电影院系统，其中个人收听设备至少是蜂窝电话、无线通信设备、穿戴在身上的计算机，个人数据助理，个人音频回放设备，电视机和 DVD 播放机中的一种。

25. 如权利要求 1 中的电影院系统，还包括第三调节设备，它调节组合的音频信号的音量。

26. 如权利要求 25 中的电影院系统，其中第三调节设备包括用户可控制的开关，它即时地实现第一音频信号与第二音频信号的原始作品混合。

27. 如权利要求 25 中的电影院系统，其中第三调节设备包括环绕声处理器，它把组合的音频信号变换成具有预定数目的空间信息声道的音频信号。

28. 如权利要求 27 中的电影院系统，其中环绕声处理器把具有左，中置，右，右环绕声，和左环绕声空间信息声道的组合音频信号变换成只具有左和右空间信息声道的信号。

29. 如权利要求 1 中的电影院系统，其中第二音频信号包括至少一部分的第一音频信号。

30. 一种操作电影院系统的方法，该电影院系统以这样的方式输出电影给多个收听者中的每个收听者，即：允许对位于电影院环境的观众中的多个收听者进行个性化的音频音量调节，该方法包括：

向观众显示电影的视频部分；

通过使用扬声器系统向观众输出电影的相应的音频部分；

提供贮存媒体的一个或多个单元，该贮存媒体存储电影的视频部分，电影的相应的音频部分，作为声音信号且具有空间信息的一个或多个声道的第一音频信号，以及包括除了第一音频信号的音频内容以外的音频内容和具有空间信息的一个或多个声道的第二音频信号；

把来自贮存媒体的一个或多个单元的第一和第二音频信号与电影的视频部分和相应的音频部分同步地发送；

通过使用与扬声器系统无关的多个个人收听设备来接收发送的第一和第二音频信号，其中多个个人收听设备中的每一个与在电影院观众中的多个收听者的每个收听者相联系，

根据来自用户的输入，调节第一音频信号的音量；

根据来自用户的输入，调节第二音频信号的音量；

组合第一音频信号的空间信息声道与第二音频信号的相应的空间信息声道，产生组合的音频信号；

通过使用一个或多个换能器，把组合的音频信号变换成声音；以及输出声音，使它们可被多个收听者中的每个收听者听见；

其中该方法允许多个收听者中的每个收听者都能够独立于其他的多个收听者而调节第一和第二音频信号。

5 31. 如权利要求 30 中的方法，其中第一音频信号和第二音频信号中至少一个是单声道的信号。

32. 如权利要求 30 中的方法，其中第一音频信号和第二音频信号中的至少一个是立体声信号，该立体声信号具有左和右空间信息声道。

10 33. 如权利要求 30 中的方法，其中第一音频信号和第二音频信号中的至少一个是环绕声信号，该环绕声信号具有包括左，中置，右，和一个或多个环绕声声道的空间信息声道。

34. 如权利要求 30 中的方法，其中第一音频信号和第二音频信号中的至少一个是多声道环绕声信号，该多声道环绕声信号具有包括左，左环绕，中置，右环绕，右，和一个或多个环绕声声道的空间信息声道。

35. 如权利要求 30 中的方法，其中调节步骤是通过使用音量可控制的有源放大器执行的。

20 36. 如权利要求 30 中的方法，其中调节步骤是通过使用音量可控制的无源衰减器执行的。

37. 如权利要求 30 中的方法，其中调节步骤是通过使用单个音量控制设备执行的。

38. 如权利要求 37 中的方法，其中当音量控制设备向一个方向移动时，第一音频信号的音量增加，而第二音频信号的音量降低，以及当音量控制设备向另一个方向移动时，第二音频信号的音量增加，而第一音频信号的音量降低。

39. 如权利要求 30 中的方法，其中接收步骤接收包括第一音频信号的第一数字比特流和包括第二音频信号的第二数字比特流，该方法还包括：

30 将第一数字比特流解码；以及
 将第二数字比特流解码。

40. 如权利要求 30 中的方法，其中接收步骤是通过使用单个接

收机执行的。

41. 如权利要求 40 中的方法，其中接收步骤接收包括第一音频信号和第二音频信号的单个数字比特流，该方法还包括：

把单个数字比特流解码。

5 42. 如权利要求 30 中的方法，其中在接收步骤接收的发送是无线发送。

43. 如权利要求 30 中的方法，其中个人收听设备至少是立体声头戴式耳机，单声道的耳机，助听器，和辅助性收听设备中的一种。

10 44. 如权利要求 30 中的方法，其中个人收听设备是穿戴在身上的接收机，该穿戴在身上的接收机提供组合的音频信号给一个或多个电声换能器。

45. 如权利要求 30 中的方法，还包括通过使用波导和放大器中的至少一种来增强该组合的音频信号。

46. 如权利要求 30 中的方法，还包括：

15 计算第一音频信号的音量对第二音频信号的音量的比率；以及自动地调节和保持第一音频信号音量对第二音频信号音量的比率。

47. 如权利要求 46 中的方法，其中计算步骤计算在有限的时间间隔内音频信号的标准偏差。

20 48. 如权利要求 46 中的方法，还包括把比率存储在存储器中，供音频信号组合设备使用。

49. 如权利要求 46 中的方法，其中调节步骤是由用户通过图形用户接口执行的。

25 50. 如权利要求 46 中的方法，其中调节步骤是通过使用单个用户可控制的音量调节设备执行的，该单个用户可控制的音量调节设备用来自始至终地调节组合的音频信号的音量，以使该单个用户可控制的音量调节设备按第一方向移动时将增加第一音频信号级的音量和降低第二音频信号的音量，而按第二方向移动时将增加第二音频信号的音量和降低第一音频信号的音量。

30 51. 如权利要求 30 中的方法，其中电影的相应音频部分是作为第一音频信号与第二音频信号而被存储的。

52. 如权利要求 30 中的方法，其中调节步骤是通过使用计算机

软件和硬件中的至少一种执行的。

53. 如权利要求 30 中的方法，其中个人收听设备至少是蜂窝电话、无线通信设备、穿戴在身上的计算机，个人数据助理，个人音频回放设备，电视机和 DVD 播放机中的一种。

5 54. 如权利要求 30 中的方法，还包括调节该组合的音频信号的音量。

55. 如权利要求 54 中的方法，其中组合的音频信号的调节包括使用用户可控制的开关，它即时地实现第一音频信号与第二音频信号的原始作品混合。

10 56. 如权利要求 54 中的方法，其中组合的音频信号的调节包括把组合的音频信号变换成具有预定数目的空间信息声道的音频信号。

15 57. 如权利要求 56 中的方法，其中变换步骤把具有左，中置，右，右环绕声，和左环绕声空间信息声道的组合的音频信号变换成只具有左和右空间信息声道的信号。

58. 如权利要求 30 中的方法，其中第二音频信号包括至少一部分的第一音频信号。

20

消费者应用中用于处理音频信号的电影院系统和相关方法

5 本专利申请要求 2000 年 2 月 4 日提交的、题目为“Use of VRA in Consumer Application (在消费者应用中 VRA 的使用)”的美国临时专利申请序列号 No. 60/190,220 的利益。

发明领域

10 本发明的实施例总的涉及用于处理音频信号的方法和设备，更具体地，涉及在消费者应用中使用的、用于处理音频信号的方法和设备。

发明背景

15 具有包括多通道放大器和多扬声器系统的“高级”或昂贵设备的最终用户，当前只具有有限的能力来独立于其他剩余声道上的音频信号，去调节多声道音频系统的中置声道信号的音量。由于许多电影的大多数对话是在中置声道上，而其他声音效果则位于其他声道上，所以这种有限的调节能力允许最终用户提升大多数对话声道的幅度，以使得在具有响的声音效果的段落期间更能听懂对话。当前，这种有限的调节具有重大的缺点。首先，这是一种只能提供给具有昂贵的数字多用途光盘 (DVD) 单放机和多通道扬声器系统 (诸如 6-
20 扬声器家庭影院系统，它允许独立地调节所有扬声器的音量级) 的最终用户的调节能力。因此，购买不起这种系统的用户不能享受这样的收听记录或广播节目的方式，即：允许最终用户提升大多数对话声道的幅度以使得其能够更听懂。

25 其次，它是一次只能为一个顾客提供服务的调节。例如，如果用户选择相对于背景声而言主要是对话声的声级来改进他的收听的懂度，然而，这个声级对于房间中其他人员可能是不满意的。因此，无法同时提供这种调节特性给具有不同的听觉喜好的多个听众。

30 另外，这是在喜爱的音频信号或语音对话 (中置声道) 和其余音频信号 (所有其他声道) 的瞬变期间需要不断地修改的调节。最后的缺点是，在电影节目的一个音频段期间可接受的语音相对其余音频 (VRA) 的调节，对于另一个音频段 (如果其余音频声级增加太多或对话声级减小太多) 可能不太好。

事实上,大多数最终用户在许多年内将不具有允许这种调节能力的家庭影院,即,Dolby(杜比)数字解码器、六通道可变增益放大器和多扬声器系统。另外,最终用户不具有确保在节目开始时所选择的VRA比率在整个节目期间都保持不变的能力。

- 5 图3显示普通家庭影院系统的打算的空间定位的布置方案。虽然没有书面的对于5.1空间信道的音频生产产品的规则,但有一些工业标准。正如这里使用的,术语“空间声道”是指输出设备(即,扬声器)的物理位置和来自输出设备的声音如何传递到最终用户。这些标准之一是给中置声道526上的大多数对话定位。同样地,需要空间定位的其他声音效果将被加到左、右、左环绕和右环绕的、
- 10 被标以L 521、R 522、Ls 523、和Rs 524的任何其他四个扬声器。另外,为了避免损坏中频扬声器,低频效果(LFE)被放在导向重低音扬声器525的0.1通道上。数字音频压缩允许制造商向用户提供音频的更大的动态范围,而这通过模拟传输是不可能做到的。在存在某些非常大的声音效果的情形下,这个更大的动态范围使得大多数对话声音太低。以下的例子提供一个说明。假设模拟传输(或记录)具有传输动态范围幅度高达95dB的能力,以及对话典型地以80dB被记录。当其余音频达到上限而同时有某个人正在讲话时,声音很大的该其余音频的段落可能会使对话难以分辨。然而,当数字音频
- 15 压缩允许动态范围高达105dB时,这种情形会加重。显然,对话将保持在相对于其他声音相同的声级上(80dB),现在只有声音很大的其余音频可以在它的幅度方面更现实地重现。用户抱怨记录在DVD中的对话声级太低,是非常普遍的。实际上,对话确实是处在正常的声级,它比起对于具有有限动态范围的模拟记录是更适当的和更
- 20 真实的。
- 25

- 即使对于当前能适当地校准家庭影院系统的消费者,在今天生产的许多DVD电影中,对话仍经常被其余声音较大的音频段遮蔽。一小群消费者能够通过增加中置声道及/或减小所有其它声道的音量而找到可懂度的某些改善。然而,这种固定的调节只是对于某些音频段落可接受,并且它扰乱了经过适当校准的声级。扬声器的声级
- 30 典型地要经过校准以便在观看位置产生一定的声压级(SPL)。这个适当的校准确保观看是尽可能真实的。不幸地,这意味着较大的声

音被非常响地重现。在半夜观看期间，这可能是不希望的。然而，对扬声器声级的任何调节将扰乱校准。

发明概要

一种用于给多个用户提供话音相对其余音频（VRA）的调节能力的方法，包括在第一解码器处接收话音信号和其余音频信号，以及在第二解码器处同时接收话音信号和其余音频信号，其中话音信号和其余音频信号被分别接收；以及由每个解码器个别地调节该分别接收的话音信号和其余音频信号。

附图简述

10 图 1 显示按照本发明的、用于把记录或广播节目中相关的话音信息与总体的背景音频分开的通用方法。

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性实施例。

图 3 显示通常的家庭影院系统的打算的空间定位布局。

15 图 4 显示按照本发明的实施例的、用于同时的多个广播的话音对其余音频（VRA）系统的方框图。

图 5 显示按照本发明的、用于多通道传输的实施例。

图 6 显示本发明的另一个实施例。

图 7 显示本发明的另一个实施例。

20 图 8 显示包括用于多通道呈现的信号处理的、本发明的另一个实施例。

图 9 显示本发明的另一个实施例。

图 10 显示通过单个控制把话音分量与其余音频分量相加和连续调节的一个实施例。

25 图 11 显示利用自动 VRA 的、本发明的另一个实施例。

图 12 显示本发明的一个实施例，其中显示滑动控制的各种功能。

图 13 显示滑动控制的各种功能的流程图。

图 14 显示本发明的另一个实施例。

30 图 15 显示本发明的另一个实施例。

图 16 显示按照本发明的一个实施例的飞机 VRA 调节盒子。

图 17 显示本发明的另一个实施例。

图 18 显示本发明的另一个实施例。

图 19 显示按照本发明的一个实施例的头戴式耳机结构。

图 20 显示除了提供按照本发明的原理的 VRA 调节能力以外，用于保持作品混合递交给最终用户的一个实施例。

5 图 21 显示图 20 的替换的实施例。

图 22 显示按照本发明的一个实施例的制作过程。

图 23 显示本发明的另一个实施例。

图 24 显示在多通道收听环境中的用户。

10 图 25 显示按照本发明原理的、在多通道处理的头戴式耳机上的 VRA 和自动 VRA。

图 26 显示传统的重现过程。

图 27 显示本发明的另一个实施例。

图 28 显示本发明的另一个实施例。

详细说明

15 本发明描述用于提供语音对其余音频能力的方法和设备。另外，本发明揭示了针对语音对其余音频（VRA）和自动 VRA 的技术的、人机工程学的、经济的、和应用方面的改进。VRA 是指通过个别地调节发声的（语言或语音）音量而达到的、音频节目中语音对其余音频的比率的个性化调节，这种调节与其余音频音量（它可以包括音乐、
20 声音效果、笑声、或其他被包括在总的音频节目中的非语音的声音）的单独调节无关。自动 VRA 或自动 VRA 保持，是指 VRA 比率的自动调节，以使得节目中的瞬变（诸如爆炸）不会遮蔽语音。

喜爱的音频对其余音频的比率的重要性

25 本发明从这样的现实出发：喜爱的音频信号对任何其余音频的比率的收听优选范围是相当大的，并且当然要比曾预期的更大。这个重要的发现是对于少量人群样本测试他们的喜爱音频信号级对所有其余音频的比率的优选范围的结果。

30 听觉障碍或正常的收听者的想要范围的具体调节

已经在了解正常或听觉障碍用户如何感知不同类型的音频节目的对话与其余音频之间的比率方面，进行了非常定向的研究。已经发现，人群在语音与其余音频之间想要的调节范围上有广泛的变

化。

对包括小学生、中学生、中年人和老年人的人群的随机样本进行了两个实验。总共测试了 71 人。测试包含要求用户对足球比赛（其中其余音频是球场观众的噪声）和流行歌曲（其中其余音频是音乐）
5 调节其话音声级与其余音频的声级。通过对于每个选择把对话或话音音量的线性值除以其余音频音量的线性值，得出所谓的 VRA（话音对其余音频）比率的度量。

作为这次测试的结果，弄清了几个事实。首先，对于体育和音乐媒体，没有两个人喜爱同一个话音对其余音频的比率。这非常重要，
10 因为人群依赖于制作人提供能吸引每个人的 VRA（它不能由消费者调节）。但根据这些测试结果，这显然是不能实现的。虽然，对于听觉障碍者来说 VRA 典型地更高（为了提高可懂度），但具有正常听力的那些人也会喜爱与制作人当前提供的不同的比率。

强调这样的事实也是重要的，即：提供 VRA 调节的任何设备必须
15 提供至少与从这些测试得出的一样多的调节能力，以使得它满足大多数人群。因为视频和家庭影院媒体提供各种各样的节目，我们应当认为比率应当从对于任何媒体（音乐或体育）的最低测量的比率延伸到对于音乐或体育的最高的比率。这将会是 0.1 到 20.17，或 46dB 的范围。还应当指出，这仅仅是人群中的一个样本，而从理论
20 上看，调节能力应当无限大，因为非常可能的是，一个人在观看体育广播时宁愿没有球场观众噪声，而另一个人宁愿没有讲解声音。应当指出，这种类型的研究和对广泛变化的 VRA 比率的特定希望在文献和现有技术中尚未见报告或讨论。

在本测试中，选择了一组年长人群，并请求他们在固定的背景噪声与播音员的话音之间调节，其中只有话音可被改变，而背景噪声
25 被设置为 6.00。年长人群组的结果为如下：

表 1

个人	设置值
1	7.50
2	4.50
3	4.00
4	7.50
5	3.00

	6	7.00
	7	6.50
	8	7.75
	9	5.50
5	10	7.00
	11	5.00

10 为了进一步说明所有年龄的人具有不同的听觉需要和喜爱的事实，选择了 21 个大学生来收听语音和背景声的混合，并通过对话音声级进行调节来选择语音对背景声的比率。背景噪声在本例中是足球比赛时球场观众噪声，被固定在 6 (6.00) 的设置值，并允许学生调节播音员的逐场的话音，这些语音被分开地记录并且是纯语音或者主要是纯语音。换句话说，学生被选择来进行与年长人群组相同的测试。选择学生是为了使得由年龄造成的听力衰减最小化。学生都是二十上下的年龄。结果如下：

15

表 2

	学生	语音设置值
	1	4.75
	2	3.75
	3	4.25
20	4	4.50
	5	5.20
	6	5.75
	7	4.25
	8	6.70
25	9	3.25
	10	6.00
	11	5.00
	12	5.25
	13	3.00
30	14	4.25
	15	3.25
	16	3.00

	17	6.00
	18	2.00
	19	4.00
	20	5.50
5	21	6.00

年长人群组（如表 1 所示）的年龄是从 36 到 59，优势人群是在 40 或 50 岁组中。正如测试结果表示的，平均设置值趋于合理地高，表示从整体上看听力有某些损失。此外，范围是从 3.00 到 7.75，有 4.75 的跨度，这进一步证实了在人们喜爱的话音对背景声或任何喜爱的信号对其余音频（PSRA）的收听比率方面的偏差范围的调查结果。对于两个对象组的音量设置值的总的变化范围是从 2.0 到 7.75。这些级别代表用来进行这个实验的音量调节机制的实际数值。它们提供可能是不同用户所希望的信号对噪声数值的范围的指示（当与“噪声”声级 6.0 相比较时）。

为了更好地了解这如何关系到由不同用户选择的相对响度变化，要考虑从 2.20 到 7.75 的非线性音量控制变化代表着 20dB 或 10 倍的增加。因此，对于即使这样少量人群样本和单个类型的音频节目，也可以发现不同的听众确实偏爱相当不同的“喜爱的信号”相对于“其余音频”的声级。这个偏爱跨越年龄组，表明它是与个人爱好和基本听力能力相一致的，所以是完全不能预料的。

正如测试结果表明的，并未因年龄造成听力损失的学生（如表 II 所示）所选择的范围从 2.00 的低设置值明显地变化到 6.70 的高设置值，其跨度为 4.70，差不多是从 1 到 10 的总范围的一半。这种测试说明了对于大多数记录和广播音频信号的那种“一个尺寸适合于全部”的心理如何远远达不到给各个收听者能力去调节混合声音以便适合于他或她自己的喜好和听力需要的目的。再者，学生像年长人群组一样在他们的设置值上有很大的分散性，这表明在喜好和听力需要方面个人有差异。这个测试的一个结果就是：听力喜好相差很大。

在更大样本组中的进一步测试证实了这个结果。而且，结果随音频类型而变化。例如，如图 3 所示，当音频源是音乐时，话音对其余音频的比率从接近零变化到大约 10，而当音频源是体育比赛时，

同一比率在接近零与大约 20 之间变化。另外，标准偏差增加几乎是 3 倍，而平均值增加为音乐的两倍以上。

5 以上测试的最终结果是，如果某人选择偏爱的音频对其余音频的比率以及永久地固定它，则该人很可能创作一个音频节目，该音频节目对于大部分人群而言并不是那么想要的。如上所述，最佳比率可以既是短期又是长期的随时间而变化的函数。因此，对这个偏爱的音频对其余音频比率的完全控制是希望的，以便满足“正常的”或非听力障碍的听众的收听需要。而且，把这个比率的最终控制提供给最终用户，能够让最终用户最佳化他或她的收听经历。

10 偏爱的音频信号和其余音频信号的最终用户的独立调节将是本发明的一个方面的明显体现。为了说明本发明的细节，请考虑这样的应用，其中偏爱的音频信号是相关的话音信息。

偏爱的音频信号和其余音频信号的产生

15 图 1 显示在记录或广播节目中用来从一般的背景音频中分离出相关话音信息的通用方法。首先，需要由节目导演对相关话音的定义作出决定。演员，或演员组，或实况录像解说员必须被确认为相关的说话者。

一旦相关的说话者被识别，他们的话音将被话筒 301 检取出。话筒 301 需要或是近距谈话的话筒（在解说员的情形下）或是在录音时使用的高方向性射击枪式话筒。除了是高方向性的以外，这些话筒 301 应该是有限话音频带的，优选地从 200-5000Hz。方向性和带通滤波的组合使得在记录时被耦合到相关的话音信息中的背景噪声最小化。在某些类型的节目的情形下，可以通过离线记录对话中相关的话音以及在适当的情形下把对话与节目的视频部分进行转录，而避免防止声耦合的需要。背景话筒 302 应当是相当宽带的，以提供背景信息（正如音乐）的完全的音频质量。

30 摄像机 303 用来提供节目的视频部分。音频信号（话音和相关的话音）在编码器 304 中和视频信号一起进行编码。通常，音频信号是通过把它与不同的载波频率进行简单调制而与视频信号分开的。由于大多数广播现在是立体声，所以一种把背景声和相关的话音信息一起编码的方法是：在分开的立体声信道上复用相关的话音信

息，这在很大程度上与将左前和右前声道加到二声道立体声上以产生四声道盘式录音相同。虽然这会产生附加的广播带宽的要求，但对于记录媒体这不成问题，只要视盘或磁带放像机中的音频电路被设计成能解调相关的话音信息就可以。

- 5 一旦信号被编码（无论用哪种认为适合的装置），编码的信号就被广播系统 305 通过天线 313 发送出去进行广播，或被记录系统 306 记录在磁带或盘上。在记录的音频视频信息的情形下，背景声和话音信息可以简单地放置在分开的记录轨道上。

10 接收和解调偏爱的音频信号和其余音频

图 2 显示用于接收和重放记录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 307 从编码的音频/视频信号中解调主载频。在记录媒体 314 的情形下，VCR 的磁头或 CD 放像机 308 的激光读出器将产生编码的音频/视频信号。

- 15 在任一种情形下，这些信号可被发送到解码系统 309。解码器 309 通过使用标准的解码技术（诸如与频分或时分解调相组合的包络检波）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 310，这样，收听者可以按照他或她的喜好进行调节。话音信号被发送到可变增益放大器 311，它可由收听者按照他或她的特定的需要调节，正如上面讨论的。

- 20 两个经调节的信号被单位增益相加放大器 312 相加，产生最后的音频输出。替换地，两个调节的信号被单位增益相加放大器 312 相加，并且被可变增益放大器 315 进一步放大，产生最后的音频输出。在这种情形下，收听者可以在重放音频节目时按他或她的独特的收听要求调节相关的话音对背景声的声级，使得音频节目最佳化。在每次同一个收听者重放同一个音频时，由于收听者的听力的改变，比率的设置值可能需要改变，设置值保持可无限的可调节性，以便提供这种灵活性。

30 VRA 和自动 VRA 实施例

如上所述，话音对其余音频的优选的比率对于不同的人是很不相同的，并且对于不同类型的节目（体育相对音乐等）也是不同的。

图 4 显示按照本发明的实施例的、用于同时有多个用户的 VRA 系统的方框图。如图所示，系统 400 包括收发信机 221，和多个重放设备，诸如个人收听设备 (PLD) 220。虽然只显示了三个 PLD，但可以使用更多的 PLD，而不背离本发明的精神和范围。

5 收发信机 221 包括接收机部件 223 和发射机部件 222，接收广播或记录的信号 235。按照本发明的一个实施例，信号 235 包括同时被发送到收发信机 221 的、分开的话音分量信号和其余音频分量信号。这些信号可以被解码器 (未示出) 解码，以后再进一步处理。替换地，信号 235 可以被发射机 222 中的系统部件和电路处理，以使得
10 产生分开的话音分量 239 和分开的其余音频分量 240。

分开的话音和其余音频信号分量由收发信机 221 通过无线或红外发射或通过多条路线的传送而被发送到每个 PLD。接收的信号被 PLD 接收机 231 接收，它可以例如是红外接收机，无线射频接收机，或用于有线连接的多端口音频输入插头。来自 PLD 接收机 231 接收
15 的话音信号 239 的一个输出，被发送到分开的可变增益放大器 229，这样，最终用户可以按他或她的喜好进行调节。另一个输出，接收到的其余音频信号 240，被发送到可变增益放大器 230，它可以由收听者按他或她特殊的收听喜好被调节。这些调节的信号被加法器 228 相加，以及在被转发到换能器 226 之前也可由增益放大器 227 进一步
20 步进行调节。换能器把来自增益放大器 227 的电信号变换成可听见的声音音频信号 232。

如上所讨论的，图 4 所示的实施例揭示发送两个 (或多个) 信号，其中至少一个信号是只有语音的信号或仅以语音为主的信号 (语音)，而另一个信号包含其余音频 (它也可包含某些语音)。然而，
25 如果其余音频包含某些语音，则只能更加正向地调节 VRA 比率，并将提高对话可懂度。

对于在同一个环境下的多个用户，如果每个用户正在用个人收听设备 (PLD) 收听节目，则可以完成 VRA 的单独的调节，个人收听设备 (PLD) 可以包括，但并不限于，头戴式耳机，助听器，耳蜗植入物，辅助的收听设备，包含扬声器的护目镜或头戴式装置。这样的
30 护目镜可以包括，例如，带有扬声器的眼镜，或可穿戴的计算机。在本文中使用的 PLD 将被定义成指能够接收电信号或无线信号和以不打扰在同一个总体环境下的其他收听者的方式把它变换成可听见

的声音的音频重现设备。

在个人收听设备处接收两个(或多个)信号以后,信号通过独立的音量控制(或其他类型的控制,如后面描述的)被分开地调节,以便实现对该单独用户的偏爱的VRA。然后信号被组合,并被个人收
5 听设备进一步放大、调节和换能,成为可听见的声音。由于个人收听设备不会与在同一个收听环境下可能也具有个人收听设备(具有不同的优选的VRA设置值)的其他收听者相干扰,故在同一个环境下的多个收听者可按他们自己的收听喜好独立地调节VRA。这因以下事实而被简化,即信号是同时(无线或有线地)发送到每个听众的。
10 这个技术的一个可能的应用是在公共电影院。多个收听者都可以通过在他们的头戴式耳机,ALD,助听器,或如上面讨论的其他个人收听设备上独立的VRA调节而欣赏同一个电影。图4通过以下的概略描述说明这些点。

15 发送到听众

为了让每个最终用户能够独立地调节话音相对于其余音频的声级,信号或者分开地到达个人收听设备,或者以这样的方式(多半被编码)到达,以使得两个信号可以在独立地调节之前被分开。两个信号的发送,可以通过例如使用FM立体传输来完成,其中话音(或
20 其余音频)在左(或右)声道上被发送。如果在PLD处还想要立体声节目,则需要更复杂的多声道传输。如果话音和其余音频都具有空间信息,则需要四声道发送(或者有线或者无线)和接收,以便把多声道节目送给最终用户。

图5显示按照本发明的原理的、用于这种多声道传输的一个可能的
25 的实施例。左话音和右话音节目被一个复用器9复用(或代替地被编码)在一起,其余音频的左和右节目也被一个复用器10复用。这允许双声道立体声通过有线或无线装置12由一个发射机11发送的传输被立体声接收机13接收。然后得出四个信号和对其进行独立调
30 节16,形成总的带有来自话音和其余音频信号的空间信息的左17和右18节目。有许多可能的方法来发送这些信号以便分别调节而同时保持空间信息。其他方法可以包括把左和右其余音频节目连同分开的单话音声道一起发送(因为话音信息主要

是非空间的)。

中置声道调节

作为对以上讨论的扩展,多声道节目的中置声道的发送也认为是关系到 VRA 调节能力的。对于大多数多声道节目,在电影中中置声道包含大多数对话。另外,大多数声音效果和音乐被引导到其余 4.1 音频声道中的一个或多个中。当前,纯话音声道是不提供给一般公众的。所以,在纯话音声道被做成可提供给一般公众用于大多数广播和记录之前,中置声道可被用作为上述的话音声道。所以,图 4 上的接收机可以是多声道声音解码器,诸如数字影院声(DTS),索尼动态信号声(SDDS),杜比数字,或其他多声道格式解码器。如图 6 所示的这样的解码器 19 的输出把数字输入变换成左,右,左环绕,右环绕,中心和重低音模拟输出。混合器 20 可以组合除中置以外的所有声道(根据想要的空间效果改变比率),这样,输出或者是立体声或者是单声的信号声道,它与中置声道分开地和同时地发送 21,后者接近于该只有对话的声道。接收可以如图 4 所示地被实行。

个人收听设备内的解码器和用于空间处理的措施

应当指出,虽然图 4 与图 6 相组合的实施例是指把模拟信号发送到 PLD,其接收机或多声道解码器是在集中地点的,但这并不把多声道解码器包括在 PLD 中的情况排除在本发明的范围以外,这时被发送的信号是需要解码的数字信号,以便提取话音和其余音频。图 7 显示这个概念。从媒体源(例如 DVD, CD, TIVO 或重放录相机等等)读出的,或从广播(如在数字电视或数字无线电中)接收的数字信号被直接发送 22 到 PLD28。PLD 具有内建的接收机 23,用来接收红外、无线或其他广播信号,信号被馈送到解码器 24,解码器被设计成符合它运行所需的压缩格式(例如,杜比数字或 DTS)的解码技术条件。混合器 26 使用解码器的输出来产生其余音频和话音信号(或纯话音信号或中置声道信号),这些信号由用户用增益放大器和/或衰减器 25 分别进行调节,然后被如上所述重新组合,以及被换能 27 成可听见的音频,它就是 PLD 的输出。这个具体的实施例在实现 PLD

处的多声道音频呈现中可能更方便，因为发送的信号是数字的（较少受到干扰噪声的影响），以及只需要一个传输声道；但可能是更昂贵的，因为解码过程是在每个单独的 PLD 而不是在中置位置处完成的。多声道呈现可以包括任何信号处理：它在空间上重新定位左，左环绕，右，右环绕和/或中置音响呈现，使得在 PLD（诸如头戴式耳机）中感觉更自然。VRA 调节打算结合这种类型的处理一起工作，以便提供改进的对话可懂度，而不影响除用户可调节的 VRA 混合声音以外所作的任何空间处理。

图 8 提供一个可能的实施例的进一步细节，它包括用于多声道呈现的信号处理。取决于对于实施方案的优选的实施例，接收机 29 和解码器 30 被居中放置或放置在 PLD。中置声道或其他纯话音声道在进行空间处理之前被分开地调节 31，所有的其余音频的声级调节 32 也是这样。接着空间处理 33 接收多声道呈现（或在某些情形下双声道呈现），正如它先前被记录那样，然后，为 PLD 产生更现实的声音条件。在这种情形下，空间处理 33 不受 VRA 调节的影响，但用户仍可选择相对于其余音频的想要的话音声级。

VRA 目的使能的“音量控制”（衰减器）

话音与其余音频的物理调节机制和总的音量信号控制具有许多可能的实施例。当话音具有用户可调节的增益，其余音频信号具有用户可调节的增益，和相加信号的总音量具有进一步的增益调节（总的音量控制）时，这时出现的是最通常的调节机制。图 4 详细地显示这一点。当用户除了设置想要的总声级以外，还试图设置 VRA 比率以便舒服时，另一个实施例在调节处理中用较少的步骤提供更加用户友好的调节机制。正如大多数娱乐节目那样，对话是节目以之为中心的目标音。所以，如图 9 所示，总的对话的声级将控制节目的响度，即，总的节目声级通常是根据对话的声级而设置的。所以，通过仅仅两个控制（总的音量级调节和其余音频衰减器），用户可以通过简单的二个步骤的处理过程来选择想要的 VRA 和总的音量级。首先，总的声级通过总的音量调节 37 而被设置（图 9），由此调节在主要节目中话音的声级。这时，对话处在想要的收听水平，只有 VRA 需要被设置。通过使仅仅一个衰减器作用在其余音频，借助于降

低其余音频而不影响话音声级，可懂度可以被提高到理论的 100%。另外，衰减器可以使用可变分压器来实施，这不需要功率，而仍旧允许用户调节 VRA 比率到大于 0dB 的所有数值。为了节省附加功率，跟随在相加器 36 后面的主要音量调节 37 也可以用衰减器来实现。

5 假如当衰减器 37 无分压地传送全部信号时，放大器 38 被设计成具有足够的增益以提供给换能器 39 功率使其达到最响的音量级。作为另一个实施例，音量调节 35 可被放置在话音上，而不在其余音频上，以允许用户控制总的节目声级作为其余音频的函数，而不是作为对话的函数。把衰减器放在话音上是不希望的，因为这不能达到正的

10 VRA 比率。相反，如果当前的实施例被实施的话，必须在话音上放置有源的增益级，这样，声级可被提升得足够高以超过未受影响的其余音频，从而提供足够的正的 VRA 比率。然后像以前那样总的响度由总的音量控制进行控制。

15 以单个度盘实现的“比率平衡”

按照本发明的、用于 VRA 和总的音量调节的另一个实施例配备有具有两个输入的用于单个调节的 VRA 旋钮。这种单旋钮调节，与双旋钮声级控制和单旋钮衰减器不同，它可以调节话音与其余音频之间的平衡。图 10 显示话音和其余音频被单个控制 40 相加且连续地

20 调节，以及用总的音频增益控制（有源的或衰减器）41 进一步调节。平衡控制本身在汽车或家庭立体声系统的前后衰减或左右平衡的调节上得到很好的已知的应用。在本应用中其关键的差别在于，它调节内容的比率，而不是各个扬声器上的音频的定位。事实上，如果想要的话，可以实施进一步的平衡方式控制，以便调节音频的空间

25 定位。通过实施单旋钮 VRA 控制，用户可以使用单个旋钮完全调节 VRA（所有的比率都是可提供的）。然后可以根据想要的声级调节总的音量。

自动 VRA

30 自动 VRA 保持特性允许最终用户不但调节话音对其余音频的偏爱的比率，而且在话音或其余音频出现瞬态音量改变时“锁定”在该比率上。例如，足球比赛包含来自播音员的对话和来自球迷的背

景噪声。如果想要的 VRA 是在球迷相对安静时的时间点期间设置的，则当球迷变得更吵闹时（但播音员保持在同样的声级），观众噪声可能会遮蔽播音员的话音。同样地，如果 VRA 是在播音员讲话非常响的期间设置的，则有可能当播音员回到正常的讲话音量时，声级变得太低而没有很好的可懂度。

基于标准偏差的 VRA 技术

为了避免用户不断地调节这些声级，用户在比率被设置后可以按一个按钮，该比率将被存储和被保持，供节目的其余部分使用。用于完成这一点的一个方法是在按钮被选择的时刻，话音信号和其余音频信号的标准偏差被计算和被存储。然后当节目进行时，每个信号的标准偏差的实时计算连续进行。如果该偏差超过存储的数值，则信号被乘以存储值与实际值的比率，由此降低音量。同样地，如果该偏差远低于存储的数值，则信号可以被乘以同样的比率，以便提升声级。如果希望提升声级（当实际的偏差低于存储的偏差时），则必须在不存在信号时检测信号段，这样，固有噪声电平不会被不必要地放大。如果实际的偏差接近于零，则会出现该比率变到接近于无限大的情形。这里所讨论的自动 VRA 方法的最一般的形式被表示为以下的公式，其中：

- G_1 =话音的音量控制；
- G_2 =其余音频的音量控制；
- G_3 =总的音量控制；
- V =话音；
- RA =其余音频；
- σV_{actual} =实际话音的标准偏差；
- σR_{actual} =实际其余音频的标准偏差；
- σV_{stored} =存储的话音的标准偏差；
- σR_{stored} =存储的其余音频的标准偏差。

$$\text{Output} = G_3 \left[G_1 \left(\frac{\sigma V_{\text{stored}}}{\sigma V_{\text{actual}}} \right) V + G_2 \left(\frac{\sigma R_{\text{stored}}}{\sigma R_{\text{actual}}} \right) RA \right]$$

各别信号（话音和其余音频）中每个信号的存储的标准偏差被存

5 储，并实时地与实际标准偏差进行比较。这里，标准偏差被用作为每个信号声级的度量。也可以使用其他度量，包括在一段时间间隔内的峰值声级。为了控制音量调节和它对总的信号输出的影响，可能希望在把增益 G_1 和 G_2 加到信号上以后再计算标准偏差。结果会稍微不同：在标准偏差被存储后，对话音与其余音频的进一步的音量调节将一直是无效的，直至输入新的存储数值为止。如果这是想要的特性，则对于 V （语音）和 RA （其余音频）的偏差计算应当包括用户可选的增益 G_1 和 G_2 。如果想要进一步调节，则可以如上所述，在偏差计算和相乘以后再加上增益。

10 图 11 更详细地显示这些概念。语音与其余音频信号都要分别经受同样的运算。应当指出，这个概念的更简单的和非常有效的实施方案是当标准偏差改变时，去除对话音信号执行的运算而只修改其余音频。这样做的理由是为了把需要的计算附加开销降低一半（去除对话音信号执行的运算），因为可以假定语音声道中的变化要比
15 其余音频声道中可能存在的变化小。无论如何，最通常的实施方案示于图 11 上，它显示对话音和对其余音频的运算。虚线表示冗余的任选项，它不一定要结合实线来使用，但它确实提供在前面段落中描述的性能上的差异。用户可调节的对话增益 45 可以在进行标准偏差计算之前（使用单元 46）或之后（使用单元 44）实施。当想要的
20 性能由用户在某一时刻选择时，语音与其余音频的标准偏差被存储在存储单元（47 和 47A）中，它可以是易失性或非易失性存储器。这个存储的数值在每个信号的相乘处理 48 和 48A 时作为分子使用，而分母是在用户可调节的增益级之前或之后的当前的实际标准偏差。（应当指出，实线和虚线不是同时实施的）。决定当前的比率
25 是大于还是小于 1 的条件未示出。如果它小于 1，则表示当前的实际声级大于存储的声级，音量应当按该比率减小。如果它大于 1，则希望不进行任何运算，只是传送由用户可调节的增益值影响的信号（这需要“if(如果)”类型的语句以检验比率条件与当前条件，并作出
30 判决）。这样做可防止非常大的比率乘以较低的信号，造成对于安静的声道的非常高的噪声级。而且，对于比率可通过另外的条件设置较低的极限，该条件允许适当的低声级被相应地放大，但非常低的信号级（或不存在）可以不修改或用条件被违反之前的最后比率

来修改。

不同的 VRA 和自动 VRA 的贮存

在 47 和 47A 中存储用户优选的比率级对于使用硬件/软件控制不同类型的节目或不同的收听者情况的声音是有利的。由于所有的用户会喜爱不同的 VRA 以及各个音频声级可能因不同类型的节目而改变，因此认为必须提供多个贮存区域用于不同类型的节目和不同的用户。例如，把名字或密码加到每个贮存单元将允许不同的用户对于特定的节目调用不同的 VRA。根据来自图 11 的使用的方法，贮存元素可以包括想要的话音声级，想要的话音标准偏差，想要的其余音频声级，和/或想要的其余音频标准偏差。这将允许用户以相同的设置值（体育和系列剧相比也许有不同的设置值）返回重放设备而不用重新调节 VRA 声级和重新设置保持特性。在重放设备上可提供的贮存单元数目没有规定限制。图 11 举例说明而显示用户调节为按钮，它选择当前的标准偏差是存储的标准偏差。另外，有用户对 G_1 ， G_2 ，和 G_3 的控制。有几种提供这些调节给最终用户的方法，取决于它们被用在什么硬件上。例如头戴式耳机可以有几个按钮，用来存储不同的比率和根据按钮被按下的持续时间来选择这些比率。如果这些控制是结合个人计算机、个人数字助理、或蜂窝电话使用的，则它们可以是使用软件实施的图形用户接口控制。为了进一步简化调节，有可能把所有的调节（VRA 和自动 VRA 比率保持）组合成单个控制。其余音频对话音的比率可以通过如图 10 所示的单个平衡控制来控制。然而，为了实施图 11 所描述的自动 VRA 特性，必须通过加上比率保持来修正旋钮的性能。

图 12 是显示滑动控制的各种功能的图，滑动控制被设计成通过使用单个控制完成所有的功能。（应当指出，这可以是任何类型的控制，包括旋转旋钮，软件控制，增量按压按钮等等，但功能是相同的）。VRA/自动 VRA 控制的中心位置将向用户提供原先的混合声，其中话音与其余音频是近似相等的。当旋钮逐渐移到左面时，话音声级不改变，但其余音频开始降低，而不进行保持功能。在远离静止点的某个预定距离的条件下（其中 N 等于该条件，并且如果希望的话它可以小到零），与标准偏差进行比较的那个数值随着旋钮移

动而开始减小，即，开始压缩其余音频。这个过程继续进行，直至存储的标准偏差（随旋钮移动而改变）改变得这样小，以致除法导致接近于零的数而相乘的输出基本上为零并只剩下话音。在旋钮的另一端则其余音频出现相反的情形。

5 图 13 显示图 12 所示的合而为一的旋钮的方框图。在旋钮达到在刻度左侧处的点 N 以后，旋钮控制其余音频的标准偏差的存储值。同样，话音的存储的标准偏差是把旋钮移到右面远侧时被调节的。对于图 13 所示的实施例的一个可能替换例是消除话音自动 VRA 控制，由此当旋钮移到左面时只降低实际的话音声级。（对于这样的
10 实施例的论点在前面的段落中给出）。参照图 13 的方框图，计算其余音频标准偏差 52，并与存储的其余音频标准偏差 56 进行比较 53，后者是通过移动旋钮 57 的位置到点 N 的左面而被控制的。如果实际的标准偏差超过存储的标准偏差，则其余音频在被也由主旋钮 57 控制的其余音频的音量级修正 55 之前，被乘以存储的数值，并被除以
15 实际的数值。如果它不是更大，则其余音频只乘以当前的旋钮设置值 55，随后与调节的话音相组合。当旋钮更向右移时，对话音轨道作同样的操作。这种单个旋钮调节在空间成问题的应用中（诸如头戴式耳机或助听器）是特别有用的。这允许具备多种控制的所有的功能，但只需要单个旋钮来实施所有的调节。

20

另外的 VRA 消费者应用

其他的 VRA 消费者应用可包括：

- 便携式“背带盒 (belt box)”，它接收和发送经过调节的和/或受到控制的信号到声换能器
- 25 · 可改型的设备，用来方便在航班飞行时对电影的 VRA 调节
- 分开的音频解码器，可以结合现有的家庭影院硬件来为多用户应用提供附加的 VRA 调节
- 头戴式耳机 VRA 调节，具有 VRA 调节能力的远程控制器。

30 VRA 硬件被专门设计来提供 VRA 调节能力，而下一节的描述说明 VRA 硬件可如何集成在现有的音频重现硬件中。然而，这并不限定在本节中描述的硬件必须结合现有的音频重现硬件使用。事实上，将会看到，VRA 特定的硬件被设计成具体地与现有的音频重现硬件诸如

电视或家庭影院系统相接口。还应当指出，在前一节中讨论的每个具体的实施例可以直接应用于在本节中讨论的每个发明，形成用于调节 VRA 比率的对用户友好的新的发明。例如，在本节中讨论的第一个发明将是便携式电子部件，它可接收两个或多个信号，一个是
5 话音，另一个是其余音频，它组合和调节这些信号，以及把它们重新发送到例如不引人注意的头戴式耳机，ALD，助听器，耳蜗植入物，辅助的收听设备，包含扬声器的护目镜或头戴式装置。在前一节中讨论的以及在图 9 上详细显示的单旋钮的发明可被包括在这个便携式部件中，以便携的形式提供单个调节能力给最终用户。然而，不
10 准备通过例子对各种技术组合的每种组合作详细讨论，而是从以上通过（作用在两个信号的）例子结合接收两个信号的 VRA 硬件描述的 VRA 方法的说明进行推论。

便携式话音对其余音频（PVRA）设备

15 正如这里使用的，PVRA 是指在各种各样的环境下结合标准 PLD（个人收听设备，诸如头戴式耳机，ALD，助听器，耳蜗植入物，辅助的收听设备，包含扬声器的护目镜或头戴式装置）使用的便携式 VRA 设备。PVRA 设备能够接收来自提供至少两个信号的源的无线（或有线）传输，一个信号是纯的或几乎是纯的对话，另一个信号是其余音频。（也可以包括更多的声道，用于进一步的空间定位能力，
20 如前一节所述）。图 14 的发送 58 可以被标准化为一定的带宽和低的功率，这样，PVRA 设备 59 可被使用于各种各样的环境。这个带宽对于射频发送可以是 900MHz，或可被标准化为红外发送类型的视线传输。一旦提供商同意标准无线传输格式，则聚会地点（诸如教堂）
25 和电影院可以向听众发送话音与其余音频。PVRA 可以是通用单放机，设计成接收 60 这些信号，与其余音频 62，69 分开地调节话音 61，68，组合它们，形成总的内容的节目 64，70，以及通过有线或无线连接把它们重新发送到具有接收机 65 和换能器 66 的 PLD 67，换能器 66 用来把信号变换成可听见的声音。用于调节的方法在前一
30 节中已详细地描述，它包括可变增益放大器或衰减器，也可以包括自动 VRA 保持能力。PVRA 盒子可以成为标准化部件，通过在发射级 63 包括一个例如 1/4”立体声头戴式耳机的插头而与许多现有的 PLD

一起工作。这个实施例也需要 PVRA 设备中的头戴式耳机放大器。另外，作为例子，也可包括把有线助听器连接到 PVRA 用的标准化插头。为了使 PVRA 和家庭影院设备标准化，全部所需要的是一台立体声发射机，其中一个声道是话音，另一个声道是其余音频，而接收机则
5 被调谐成能接收这两个信号。

作为 PVRA 设备的上述说明的附属部分，这里揭示的另一个设备是打算结合航班飞行时观看电影所使用的 VRA/自动 VRA 调节。在航班飞行时电影对话可懂度可能特别差，这时来自飞机的背景噪声进一步遮蔽了电影对话。通过向最终用户提供分开调节话音和其余音频的能力，可以实现在飞行期间改善的可懂度。有可能达到这个目的，而不会干扰现有的用于音频传输的基础结构。我们必须假设，
10 音频源（VCR，DVD，广播，或其他音频源）具有与其余音频轨道分开的对话音轨。这可以以几种方式完成，一种方式是通过使用在多声道格式时的中置声道，或替换地，使用可以有用几种音频压缩标准而存在的纯话音音轨。（产生纯话音轨道并不是本发明的注意点，
15 而注意点是被使用来调节它以及把它传递到最终用户的硬件和实施方案）。飞机的音频传递基础设施包括到最终用户的立体声（2 声道）路径，它的实施可以通过（1）到每张椅子扶手的电子信息传送（需要带有航空标准连接头的标准头戴式耳机）或（2）其中在扶手中有
20 细小的扬声器的波导系统，当连接到塑料管时将把声音发送到用户的耳朵。为了实施改型的飞机 VRA 扶手调节设备，必须假设话音是在左或右声道上发送，而其余音频是在另一个声道上发送到飞机的所有的扶手。虽然这将去除立体声效果，但这被看作为以最小的牺牲给出总的节目享受的潜在的改进。另外，在飞行娱乐期间飞机噪声
25 常常遮蔽细微的立体声效果。由于电子相对于波导的方法是这样地不同，需要两种不同的实施例，它们分别显示于图 15 和 16。然而，如果希望制作通用的调节机制，图 15 和 16 的部件可被组合成单个硬件单元，它可以结合任何飞机的扶手来完成。

图 15 是用于飞机的扶手中电子连接的航空盒子的型式的图。插头 71 可以例如是标准化的阳插头，设计成适合于插入到扶手中，连接到从中央位置送来的左和右信号。这些信号然后经调节 72，73，
30 和 74，75，以达到话音（左）和其余音频（右）的喜爱的 VRA 比率。

5 这些经调节的信号然后被组合，形成总的音频节目，并进一步被 77 和 78 调节。需要一个放大器 79 来给 PLD 中的换能器提供功率。航空 VRA 盒的输出 80 包含连接头，用来配接例如在当前的扶手中已有的阴接头，以便允许飞机使用它们现有的用于 PLD 设备的头戴式耳机。

10 图 16 显示可以结合在现有的扶手中波导使用的航空 VRA 调节盒。为了调节两个信号（语音与其余音频）各自的信号级，必须把信号变回电子形式。两个话筒 82, 83 和话筒放大器 84, 85 被放置在设备中，它们测量通常用来驱动波导的扶手扬声器的输出。放大器的输出代表语音或其余音频的电子信号。这些信号通过 86, 87 和 88, 89 被独立地调节，产生总的信号 90。它被进一步调节以得到总的声级 91, 92，使用来驱动另一个扬声器 93。波导和与扶手中的那些插头相同的输出插头 94 代表航空 VRA 盒的输出，使得标准波导型式的头戴式耳机可以结合本设计而被使用。

15

DVD，电视机等中用来发送两个信号的无线发射机

20 当考虑家庭影院以及家庭电视和电影观看时，引起了用于多用户 VRA 调节的另一个应用。通常在单个房间中会有多个观众，对于改变语音对其余音频比率具有不同的收听喜好。本发明允许把多个信号提供给由多个用户佩戴的 PLD，这样，每个人可按他的喜好调节 VRA（和自动 VRA）。如前所述，音频源（例如电视广播，DVD 播放机等等）包含至少一个可被看作为纯语音或主要是语音的音轨，以及至少一个可被看作为包含其余音频的音轨。（不同的音频标准和格式在将来可能支持纯语音轨道的某种形式）。本发明所预期的是对于 25 在相同的收听环境下多个用户分开地和同时地接入至少两个信号以便允许他们按他们自己的个人收听爱好设置比率的需要。在第一和最优选实施例中（从空间和花费来考虑），多声道无线发射机位于音频重现硬件（诸如电视或 DVD 播放机）内，该音频重现硬件分别发送语音与其余音频给听众，这样，具有能进行 VRA 的 PLD（它能够 30 接收该发送）的收听者能够独立地调节 VRA。除了把无线发射机放置在音频重现设备内以外，还可以提供分开的音频输出插头，它提供对其余音频（单声道或多声道）和纯语音（单声道或多声道）的接

入，以便由不带无线接收能力的分开的 PLD 进行有线调节。在电视机、DVD 播放机、或其他设备中没有无线发射机但有音频插头的情形下，用户可以把分开的多声道无线发射机连接到那些输出插头，以便向听众提供分开的音频信号。图 17 显示这些概念。信号源 96 将递交整个节目的编码的或调制的形式，它也可包括视频信息，如果可用的话。信号源可包括，例如，电视广播信号（通过卫星、电缆、地面）和由激光器读出的编码的 DVD 或 CD 信号。这个信息必须在变为代表音频信息的电信号之前被接收和解码。解码器能够提取纯话音声道（如果它存在于信号源的话）以及保持它（它们）与其余音频声道分开。在接收/解码级 98 后，有两个选项可供分开的语音与其余音频信号使用（1）它们可做成由分开的输出音频插头 97（诸如话筒型连接头）提供，或（2）它们可被发送到也安装在重放设备 95 中的多声道无线发射机 99。重放设备 95 可以是信号源在设备内部的 DVD 重放器，或可以是信号源在外部的电视机。如果信号做成可由硬件输出的，则可将这些输出送到能与接收的 PLD 接口的、分开的外部多声道发射机 100，以便在 PLD 处提供无线接收和 VRA 调节，供同一个收听环境下所有的用户使用。

结合其他与 VRA 不兼容的系统使用的附加 VRA 解码器

作为替换的实施例，也有可能设计出专用解码器专门供 VRA 应用。这将允许当前不拥有能够进行 VRA 的解码器/发射机的用户仍旧能接入该 VRA 能力，而不用升级任何特定的部件（即，不损失他们当前的投资）。假设 DVD 或广播 TV 或广播的无线电开始发送编码的纯话音声道。当前的接收设备不但不能接收和提取该信息，而且也不能提供 VRA 特性给任何个人，更不用说在相同环境下的多个用户了。图 18 所示的设备 111 提供所有这些能力给这样的用户，它具有他可使用的、带有话音分开的信号源，但没有提取与调节 VRA 比率的装置。信号源 101 可以是如以前那样的（电视广播，DVD 信息等等），以及可直接发送到初级重现系统 102，它也许是不能提供 VRA 调节的 TV 和 DVD 播放机，因为它先前没有配备这种特性。该同一个信号源对于外部 VRA 盒 111 是“关断”的，其中相应的接收机或解码器已被安装来按照标准（它支持在信号中存在的数据类型）从信号源去

除和分离话音与其余音频。这样的解码器可以是能够提取听觉障碍模式的杜比数字解码器，但本发明肯定不限于这种特定的解码器。有可能在某一天其他流行的音频格式将提供一种手段，除了发送现有的音频以外还发送纯话音音轨。这里本发明 111 将根据想要的音频包括适当的解码器。对于外部设备在解码和分开话音轨道与其余音频轨道以后有三个选项。首先，设备可以以音频插头形式（诸如 RCA 类型或电话型连接头）提供硬件输出 104, 110, 它可提供线路声级信号给能够进行 VRA 调节的发射机或有线 PLD；第二，接收机可提供多声道发射机 105 用于将分开的话音与其余音频信号以无线发送到能够进行 VRA 调节的 PLD；或第三，为直接在设备上的单个用户应用 107 提供 VRA 调节，其中话音与其余音频被分开地调节，再相加在一起，并且在作为输出 109 提供给任何其他音频回放部件之前，对总的音量级作进一步调节。应当指出，虽然图 18 上只显示单个调节的输出（是指总的单声道输出），但提供几个输出以用于音频的多声道空间定位，也是属于本发明的范围之内（如在先前的说明中提到的）。例如，如果 5 声道其余音频节目（左、中、右、左环绕、和右环绕）是连同 5 声道（或较少的）话音节目一起提供的，则这些声道可以在声级调节后被组合，以使得话音可以通过使用分开的音量调节送到任何扬声器，或直接送到中置扬声器（典型情况）。这仍旧提供 VRA 调节能力给最终用户，而同时也提供完全的环绕声感受。

结合混合（视频与音频）DVD 信号的现有无线传输使用的接收和 VRA 正在开发一种新类型的产品，它允许消费者欣赏来自远地的 DVD 视频与音频信息。这种无线技术把来自远地 DVD 播放机的音频与视频信息传递到电视或家庭影院。这允许个人计算机中 DVD 驱动器的所有者使用该驱动器在不同于他们的计算机的地点观看 DVD 的内容。在 PLD 位置的以及在中置家庭影院处的 VRA 调节可以结合无线 DVD 技术以两种方式被使用。首先，由于 DVD 播放机已经发送无线音频信号给家庭影院系统，所以 PLD 可以配备一台设置在同一频率的无线接收机，这样 PLD 可截取同一个发送。由于只是在观看位置而不是在 PLD 处需要视频信息，在 PLD 处音频可有选择地解码用于

重现。重要的是要指出，在无线信号的接收过程之后将是解码处理，以便从无线 DVD 信号中提取语音和其余音频。此后是调节，重新组合，和把信号换成可听见的声音，也包括语音声级、其余音频声级、总的声级、和任何自动 VRA 特性的充分调节。以上提到的所有的硬件将放置在个人收听设备中，这样，每个用户可以按他/她的喜好调节 VRA 和声级；这样的系统显示于图 4，其中发射机发送全部 DVD 信号，但 PLD 解码器只设计成从进入的比特流中提取音频。

这里是结合无线 DVD 发送使用的 VRA 的第二实施例，它可导致降低的总的花费，但所需要的部件的数目要增加。正如前面一样，总的 DVD 信号从 DVD 播放机发送到重放地点，例如家庭影院。在居中的家庭影院地点处总的 DVD 信号的接收机可以只重发语音与其余音频，以实现类似于图 4 所示的实施例，但其不同点在于，收发信机在中置位置处接收来自无线 DVD 发射机的无线信号。这允许在同一个地点的多个 PLD 只配备有一个无线接收机，而不是一个数字解码器。解码处理过程在家庭影院地点集中执行，视频被发送到观看设备，音频（从 DVD 播放机地点接收和解码的）从接收机、解码器、发射机被重发到接收语音与其余音频的 PLD 中。

头戴式耳机机架上的 VRA 旋钮

所讨论的本发明的下一种类别集中在把 VRA 调节特性引入到被设计用于 VRA 应用的特定的硬件，即，具有 VRA 能力的个人收听设备。有三类 PLD 作为本实施例的注意点，它们是利用有线或无线技术的头戴式耳机，助听器，辅助的收听设备，耳蜗植入物，包含扬声器的护目镜或头戴式装置。通常，辅助收听设备利用头戴式耳机，并结合话筒或无线发射机，这取决于产品的用途。在某种意义上，图 4 所示的无线 VRA 系统本身可被看作为辅助性收听设备。但通常，接收、分离、调节、重新组合和传递的 VRA 技术可被使用于 ALD 以外的产品上。无线头戴式耳机、头戴电话、或小耳塞可以包含直接在头戴式耳机边上的音量控制。图 19 说明这样的实施例，其中所有必须的硬件放置在头戴式耳机 112 中，而用于 VRA 控制的必须的调节很容易在耳机的外壳上提供给最终用户。首先，无线接收机 113 接收由源位置（如果需要的话）发送的、解码发生后的多个音频信

号。解调器 114 得出音频信号的基带，得到话音音轨和其余音频音轨，然后它们被人工调节 115，相加 117，和进一步调节声级 116，被头戴受话放大器 118 放大，以及通过头戴受话扬声器 119 重现。音量旋钮 115，116 可以位于头戴式耳机的外部，以及如果使用多声道（立体声）重现的话，可以带平衡调节。如果多个音频声道在 PLD 处被发送和被接收，则它们可以被拥有图 5 和/或 8 所示的硬件与软件所调节和组合，形成需要的立体声声象或空间投射。

助听器中 VRA 的遥控调节

除了头戴受话设备（它在物理上足够大可容纳上述的硬件）上实行的 VRA 调节以外，可能也希望在较小的 PLD 中，包括助听器或较小的头戴受话设备，允许进行 VRA 调节。这些较小的 PLD 可能无法容纳用于调节和人工控制两个或多个信号的所有需要的硬件。在这种情形下，可能更希望使用诸如图 13 描述的设备。通过这个使用小型 PLD 的实施例，需要的只是 PLD 中的无线接收机。表面封装技术和电子元件的小型化使得容易把低功率无线接收机引入到小的空间。（应当指出，因为调节信号的发送只需约 5 英尺的范围，所以需要较小的天线功率放大器，因为手持调节机构接收从源发送位置发出的分离信号）。图 13 所描述的远端收发信机 13 也可以构建成为用于 PLD 的遥控器的形式。所以，遥控器控制话音与其余音频的音量和 PLD 的总的音量，同时还用作对 PLD 的发射机和从源位置接收的接收机。将来的技术可以允许将 VRA 调节所需要的整个电子装置放置在例如助听器中（或小型化的 PLD 中）。可能仍旧必须使用遥控器来实际控制音量级，因为希望保持助听器尽可能不引人注意。在这种情形下，为了审美的原因，遥控器仍可能是想要的。

在电影院环境下用于 VRA 头戴式耳机的实施例

如果头戴式耳机设计配备这里揭示的技术的话，那么电影院将为个人调节 VRA 提供再一个机会。个人可以利用独立于其余音频而控制话音声级的能力，同时享受电影院的环绕声或大屏幕。希望在电影院的辅助性收听设备或头戴式耳机中具有对其余音频和话音的完全控制，则需要具有足够的无源和/或有源控制的护耳式头戴式耳

机。利用护耳式、双罩设计和衰减材料的无源噪声控制在阻塞低到约 500Hz 的环境频率方面是有效的。在这样的头戴式耳机中也可能希望包括有源噪声控制，以使不能由无源措施有效控制的较低的频率能进一步减小。这样的电影院耳机可以如在先前的头戴式耳机实施

5 5 施例中讨论的那样设计。如果只需要对话音进行控制的话，则也可以有稍微降低功能的第二替换例。许多电影是震撼性的，因此在其余音频很响的段落中，话音声级太低而不能提供良好的可懂度。可能希望在这些段落期间仅仅提升对话的声级。仅仅对话的信号的接收、音量控制、和接收可以通过使用非护耳式头戴式耳机来完成，

10 10 这种头戴式耳机允许环境声音到达耳朵。允许环境声音到达用户耳朵，可以让来自多声道环绕声的空间提示信号仍到达耳朵并被听见，但对话可被调节以便改善可懂度。

应当看到，话音对其余音频（VRA）比率的调节（这是在专业记录制作端的混合处理过程中的一个部分）可能与某些个人的艺术自由的概念相抵触。例如音频工程师主要集中在得到声音的正确的混合，产生音乐、电影和电视中的想要的效果。所以，必须包括一个

15 15 装置，或者发送总的节目的原始的（未受影响的）作品级别的混合，或提供一种容易达到该混合的方法。这允许最终用户在选择作品混合（制作人设计该混合的方式）的能力或自己调节 VRA 比率的能力

20 20 之间进行选择。对于完成这个目标至少有四个可能性，它们被给出如下。

VRA 对艺术混合（原声音频混合）的选择的实施例 方法 1

图 20 表示在提供 VRA 调节能力以外，用于保持作品混合（如制

25 25 作人打算的那样所设计的混合）的传递给最终用户的头两个选项。制作人从所有单元都分开的音轨 120 开始，这些单元形成整个节目，把它们组合在一起，以形成单声道或多声道节目 122，并被记录或广播 123 到最终用户。除了作品混合以外，用来创建作品混合的话音（用与 121 相同的处理进行时间对准，延时，和处理 124）在整个记录

30 30 和广播级 123 中应当保持与作品混合分开。典型地，信号以具有一定的带宽的单个频率广播，所以，它常常被表示为单个信号而作为输入 126，如图 20 所示。（即使这被显示为单个信号，但多个信

号是被包含在调制/编码信号中的)。解码器/重放设备 125 解码或解调记录或发送,以便除了只有对话的音轨 127 以外,还提供原始作品混合 126,只有对话的音轨 127 是结合作品级的混合而制作和记录的。对话信号传送经过开关 129,它可断开来自输出 133 的对话信号。当这个开关断开时,作品混合是以它的原始形式提供的,以便通过在任何多声道结构 131(该混合原先就是在其中制作的)中的重现硬件 133 进行重放,音量旋钮 128 用作对整个节目的总的音量控制。然而,当开关 129 被闭合时,对话传送到音量调节 130 并被加到 132 作品混合中(如果它是多声道格式,它将典型地被加到中置声道,或等值地加到左和右声道)。这允许最终用户相对于由该对话音量旋钮 130 调节的对话的声级去调节总的节目 128 的声级。如果对话音量旋钮 130 调到最低,则可以重新得到作品混合。如果作品混合中的对话和分开的对话音轨是同时记录或广播的(即,时间对准),则在两个音轨之间没有延时;所以,当对话的音量与作品混合一起增加时,它开始提升通过重现设备 133 而听见的话音对其余音频比率。通过全部使用引起相同延时的因素来处理对话信号(诸如交混回响或滤波)(这种延迟是在作品混合中对话所经受的),可以实现时间对准。这将确保在作品混合中的对话与分开的对话音轨中在时间上同等地调准。应当指出,这个特定的实施例不允许达到比作品混合 VRA 比率更低的 VRA 比率。如果这是想要的特性,则下两个发明将完成这个目标。

方法 2

如以前那样,假设作品混合和对话是从广播或记录所提供的以及两个对话信号(一个是作品混合和另一个是纯的)在时间上互相对准。图 21 显示在对作品混合和对话声道进行解码后将允许负的 VRA 比率(即,用户降低话音声级和提升其余音频声级,如果想要的话)的另一个配置。经解码的对话信号 135 从作品混合 134 中被减去,得到纯的其余音频 137 混合。这时,有可能在组合 140 以形成总的用户调节的节目以前,独立于对话声级 138 调节其余音频声级 139。作品混合可以在开关 141 的一个端子上提供,这个开关 141 在用户调节的混合与作品混合之间交替选择。这种配置允许把对话信号降低到只留下其余音频的程度。下一个发明通过在编码处理过程之前

记录作品信息而在可以得到作品级混合的情况下也提供用户可完全调节的 VRA 比率。

方法 3

图 22 显示一种制作过程，它发送有节目是如何被混合的信息，以便确保作品混合是以消费者级别提供的，以及提供通用调节 VRA 比率的能力。原始的节目单元 143 被混合 144，以形成多声道或单声道节目，并不断试听直至所有的输入 143 的声级确实是精确地重现了制作节目的音频工程师所确定的声级。然后作品混合的输出被设计成对话信号本身 149、制作人确定的对于对话信号 148 是合适的声级、认为并非关键对话 146 的所有的组合的音频、以及该其余音频的总的声级。各个声级信息作为中间数据以数字方式编码 150，151 到实际的音频信号 152，153 中。带有它们各自的声级信息的这些编码的信号被发送，广播，或记录 154。重放设备配备有解码器 155，它提取音频信息以及包含原始作品的混合声级（可能是相对于某些最大数字化数值，取决于记录的分辨率）的头信息 156。其余音频声级和对话声级 157 则提供给增益调节电路 158 和 159，使得在组合 162 自动调节的对话和其余音频时声级是正确的并能实现原始的作品混合。这只是在开关 160 和 161 把头信息块 156 的输出连接到增益调节电路时才能发生。如果开关被投掷在右面位置以连接 VRA 的用户混合 164，则用户选择其余音频声级 158 和对话声级 159。虽然在许多情形下，显示路径表示的是单个信号，但在存在多个信号，诸如用于空间定位的左，右，左环绕，右环绕，和中置声道的情形下，把每个信号路径看作为一个矢量也属于本发明的范围，它们的声级都由图 22 的其余音频声级控制 158 控制。同样地，纯话音音轨也可以包含多声道信息，它们可以由控制 159 调节。

结合现有的音频重现硬件的 VRA

上面的讨论集中于在个人收听设备 PLD 上向最终用户提供以电子方式重现媒体（或者是广播重放，或者是记录重放）的调节 VRA 的能力，这样，在同一个环境下的各个收听者可以同时享受不同的 VRA 比率。当上述的个人收听设备被扩展为包括以下的电子装置时将作出进一步的发明：

- 蜂窝电话
 - 可戴的计算机
 - 个人数字助理
 - MP3 重放设备
- 5 · 使用磁贮存媒体来存储音乐的个人音频播放机

这些设备可被使用于音乐或包含对话和其余音频(它可能遮蔽对话)的音频的个人级别的重放。在前几节中讨论的实施例可以应用到以上列出的四种设备,以提供用于音频回放的 VRA 调节,这些音频是把对话与其余音频分别地预先录制和制作的,而与编码的格式
10 无关。

个人计算机上的 VRA

互联网已成为发布数字质量媒体的流行的声道。如果消费者通过
15 与他们的个人计算机的数据连接来实时地(或延时地)接收音乐、电影、或其他音频,则他们可以以各种各样的方法实施 VRA 和自动 VRA 特性。加到话音与其余音频上的增益控制可以全部是通过图形用户接口驱动的软件。话音与其余音频通过硬件或软件(取决于个人计算机系统)被分开地编码。通过把几行源代码加入到解码程序中,指示两个信号(话音与其余音频)的每个信号要被乘以用户可调节
20 的常数,信号就可以在数字域中被放大。这些常数在信号被相加在一起之前由用户通过软件用户接口控制,它允许调节这两个常数以便与解码的话音和其余音频信号相乘。另外,可以对总的组合后的声级加上进一步的音量调节,以允许用户在重放之前控制节目的总的音量。替换地。如果希望提供更“用户友好的”调节能力给用户,
25 则可以提供 VRA 旋钮(见对于可能的控制旋钮实施方案的方法的讨论),作为在计算机扬声器、键盘、鼠标或监视器,PC 系统的所有部件上的实际硬件。如果 VRA 是做成由硬件(如监视器上的旋钮)提供的而信号解码是通过软件实施的,则需要信号交换协议来确保通过使用硬件旋钮进行的调节被转换成软件的增益改变和相乘运
30 算。图 23 显示用于硬件与软件接口的一个可能的选项。与位置和满刻度电压 166 有关而产生电压输出 167 的硬件旋钮 165 的移动和位置必须通过 A/D 168 来采样,以便把位置信息变换成代表相对于满

刻度 166 的音量的数字。这样的旋钮的一个可能的硬件实施方案是旋转分压器，它是一个电位器，一端是满刻度电压，另一端是地，而滑动片则提供分得的电压作为旋转位置的函数。然后 A/D 的输出通过控制信号流的软件 169 周期地轮询，以得到用户所选择的经过数字化的数。这些数（一个用于语音和一个用于其余音频）与各个信号相乘 170，其输出被相加 172，形成总的经 VRA 调节的节目。在个人计算机重放设备上用于控制 VRA 的软件和硬件的使用方面有几种其他的组合。只用硬件的形式要求将信号解码并提供作为声卡的输出或从 PC 主板本身输出，如图 18 所示的设备那样。这个实施方案允许使用硬件增益或衰减实行两个信号的音量调节，而不需要图形用户接口。每种配置具有它本身的优点：

- 全硬件：便宜，旋钮易于得到，以及容易以高的可见度进行调节
- 全软件：不需要为了实施 VRA 而进行硬件更新，调节选项和特性有更大的灵活性，GUI 控制可以由最终用户定制
- 软件/硬件：调节机构的高的可见度，需要较少的 D/A 变换器，因为输出的是和而不是其余音频与对话。

个人计算机上的自动 VRA

还应当指出，虽然个人计算机被看作为个人收听设备，但头戴式耳机（还有 PLD）常常结合 PC 机使用。所以，PC 机可用作为其他 PLD 的信号源。讨论电视和 DVD 的先前的实施例也可包括在 PC 上的信号源，诸如，带有连接到它的 VRA 控制的头戴电话插头。另外，虽然注意点只是 VRA 调节，但自动 VRA 保持特性也可以在 PC 机上实施。事实上，由于对于计算实时信号性质所需要的计算，需要能够执行每秒很大数目的运算的中央处理机（CPU）或数字信号处理器（DSP）。所以，VRA 的 PC 实施方案由于可供使用的计算能力而能够易于实施自动 VRA 特性。在前几节中讨论的所有的自动 VRA 和用户控制可以通过使用在前面段落中讨论的任何的硬件/软件接口选项来完成。然而，纯硬件实施方案因所需要的计算能力和对于连续限制信号级所需要的实时运算而很困难。所以，其中或者使用全软件或者使用硬件控制与软件数学运算的组合的实施方案，对于自动 VRA 实施方案将是优选的。

图 24 显示在多声道收听环境下的用户。虽然图 24 显示具有 5 个扬声器（左，中置，右，左环绕和右环绕）的情形，但这样的环境可以具有 2, 3, 4, 5 个或更多的扬声器。每个扬声器具有从本身到每个耳朵的频率响应路径，产生总共 10 条路径。如果驱动这些扬声器的电信号在被组合形成左耳信号和右耳信号之前用这些路径的估值进行滤波，则可以为用户收听产生更现实的声舞台。这显然是想要的结果，允许个人体验多声道环绕声感受，而不用购买多声道放大器/扬声器系统。然而，需要一种发明来向最终用户提供结合多声道空间处理来调节 VRA 的能力，以便达到想要的可懂度而同时通过使用头戴式耳机体验环绕声舞台。

多声道处理的头戴式耳机上的 VRA 和自动 VRA 与杜比头戴式耳机结合使用

图 25 描述的设计假设对话音轨是在总的音频节目以外分开提供的。所以，对于这个特定的实施例不会达到负的 VRA 比率。然而，如果两个多声道节目同时被解码（其余音频与语音），则通过完全降低或提升其余音频和/或语音，将可以达到所有可能的 VRA 比率。图 25 显示总的多声道音频节目（其余音频加语音）173，传送通过总的音量控制 174，它最终将用作为其余音频的控制。在 175 的空间处理是指现有技术，其中每个信号被滤波和/或被延时，产生想要的多声道效果。然而，在它们被组合在一起以形成二声道头戴式耳机节目之前，改变了的对话被加到 176 适当的声道。作为空间处理的组成部分，信息要被保持以确定对话音轨应当出现在哪个扬声器，同时该信息被传递到决策步骤 178。对于大多数作品，语音位于中置声道或同时在右和左声道（仿真中置）。对于本说明，假设对话只被传送到中置声道。当中置声道用 178 表示为对于对话的正确的位置时，对话处理 181 为要放置对话的声道而复制空间处理信息（滤波，延迟时间等等）。对话 179 在被（中置）声道处理 181 处理之前首先被调节声级（语音声级调节）180，这个中置声道处理 181 是从多声道空间处理 175 复制的。在对话被处理 181 以后，将要接受对话的扬声器再次被移送到 178，然后信号被加到适当的声道 176。如果语音是指要加在左和右声道中，则 178 块把经适当处理的语音

经过 176 传递到这些声道的每一个，而不传递到任何其他声道。其余音频信号可以通过使用 174 被提升和降低，而对话可以通过使用 180 被提升和降低，可以提供正的 VRA 调节，因为话音包括在总的节目 173 内。然后经 VRA 调节的、空间处理过的、多声道节目然后被 5 177 进一步处理（现有技术），产生二声道头戴式耳机节目。这个二声道节目要进一步调节其音量，并被传递到头戴电话扬声器 183。作为图 25 的元件 175 和 177 所讨论的现有技术的最新实施方案是杜比头戴式耳机。以上的 VRA 发明设计成结合杜比头戴式耳机以及任何其他多声道处理头戴式耳机（从多个空间声道得出二个头戴式耳机 10 声道）一起工作。

在“存储节目”重放设备上的 VRA

非线性电视观看给观众提供重大的优点：节目可被记录和存储，供以后观看。最新的技术，硬盘驱动器电视记录器，（包括 15 TiVO, Replay, 和 Microsoft 等公司）与传统的 VCR 技术的不同之处在于，记录方法是非常用户友好的，可以设定分开地记录的声道，而且重放几乎是即时的。将来的电视观看多半喜欢非实时的、非线性的观看，而不是在本来是不方便的时间实时观看。所以，重要的是 VRA 调节能力要能够结合这些重放设备而工作。调节硬件可作为 20 遥控器、视频屏幕 GUI 或在重放机构上的物理硬件的一部分。记录处理过程只需要记录在它被发送时的所有的信息，包括分开的话音轨道。然后，重放和调节机构将引用与在 VRA 和自动 VRA 调节硬件以前的实施例中所讨论的相同的部件。

重新制作

当前，音频的制作（用于广播，电影，音乐等等）可被看作为多 25 步骤的过程，如认为是现有技术的图 26 所描述的。在制作级别上，存在几种类型的声音，它们将被记录以形成整个音频节目。这些声音 184 可被分成几种类型，包括声音效果，音乐，话音和其他声音。典型地，为了了解节目或音频的内容或歌词，制作级声音的话音段被看作为关键的。所有这些声音首先被分开地记录 185。某些声音不是 30 通过话筒记录的，诸如声音效果，它们常常从预先录制的效果音轨组转录。并不总是必须记录所有的声音，以使它们在主声道同步。非线性记录和重放允许作品 186 在几个重放/记录周期过程中把声音

与视频以及与其他声音对准，常常通过使用软件驱动的记录和编辑来完成。制作处理过程 186 把所有的声音互相同步（并与视频同步，如果存在的话），把它们以适当的比率混合，以及把它们加到音频工程师感觉它们应该归属的环绕声道中（如果可用的话）。例如，

5 对于在屏幕上讲话的演员常常要把他们的话音放置在最接近于屏幕位置的中置声道。为了发送或记录大量信息，常常需要编码或压缩处理 188。如果使用模拟记录和重放的话，情况并不总是这样，但多声道数字重放常常具有某种编码（用于复制保护）和/或压缩（有损或无损），这取决于记录媒体或广播的要求。编码的和/或压缩的节目

10 然后被记录或广播 189，以及在最终用户位置处被重放或被接收。从图 26 的处理过程可以看到，最终用户不能将对话声级相对于其余音频进行任何调节，因为制作人在 186 具有完全的控制。一旦混合，几乎不可能从其余音频提取话音和把它恢复到它的原先记录的质量 184。

15 然而，有可能制作人要返回到主节目 85，其中所有记录的单元是互相分开的，以及分开地得出话音与其余音频。本发明集中在提供分开地得到对话与其余音频和重新记录它们的手段和能力，这样，最终用户可以具有调节相对声级以便适合于他/她的听觉需要的能力。图 27 显示完成这个目标的一个可能的方法。作为归档的媒体

20 几乎所有的电影和多音轨音频节目的主记录 191 是完全分开地存在的。另外，在原始记录中制作的、有关每个轨道的声级和位置的信息也连同那些主记录一起存在。这个信息在两个分开的多声道混合器 192（用于所有其余音频）和 193（只用于话音）中用来准备两个分开的多声道节目，它可以是如在 187 中所示的 6 声道，也许更多

25 或更少，取决于想要的效果。所以，来自原始记录过程 186 的制作信息用来分开地和同时地产生与其余音频和话音的完全相同的效果。因此这些多声道节目中每个都可以在 192 和 193 的输出端处被组合，可以从原先的混合形成相同的总的音频节目 187，虽然它们是完全分开的。这两个多声道节目的每一个然后在 188 中被编码和/或

30 被压缩，但完全分开地使用 194 和 195。这两个编码的声道然后被进一步编码或被复用 196，以产生单个信号，再被广播或被记录 197。在重放时，这个信号被解码，形成至少两个多声道信号，它们可通

过使用在本文件的先前各节中讨论的硬件和实施例而进行 VRA 调节 199。

5 重新控制多声道音频到“VRA 友好的”形式，重新控制立体声到“VRA 友好的”形式

图 27 表示用于把话音与其余音频分开地发送到最终用户的最通用的方法，这样，其余音频和话音的所有的空间信息在重放期间都被保持。图 28 显示这种情况的对立面，其中不保持空间信息，但最终用户具有独立地调节话音对其余音频声级的能力。主记录存在于 10 185 和 191 以及图 28 的 200。被看作为其余音频（总的音频记录中非话音单元）的所有的分量只由 201 进行声级混合。这是指，只决定一个其余音频分量对于另一个其余音频分量的相对影响，但不实施空间定位，因为 201 的输出是单个信号而不是多声道信号。另外，话音声级被调节 202，这样，当不进行任何调节而被组合时，话音相 15 对于其余音频的声级正好是制作人所希望的。然后调节过的话音和调节过的其余音频节目记录在立体声媒体中 203，它可包括 CD，DVD，模拟磁带等等，但也可包括立体声的音频广播。这时，话音与其余音频各自分开地保持在左和右或右和左音轨上。应当指出，需要某些惯例来确保所有的作品能认出右或左声道包含话音，而其余音频 20 存在于另一个声道。这可以根据来自消费者电子制造商的一致意见被选定，但选择其中任一个都并不限制本发明的范围。立体声重放设备 204 然后将提供两个信号作为输出（左和右），一个只具有话音，另一个只具有其余音频。为了用 VRA 调节同时体验整个节目，这两个信号通过两个可变增益放大器 205 和 206 传送，每个声级在 25 其中被控制，然后，它们被相加，形成总的节目。这个总的节目然后可以进一步调节其声级 207。这个经过完全调节的总的节目然后被分离，如果它要被立体声系统重放 208 的话。这种配置的优点在于，有可能用今天的消费者电子设备和主记录来完成 VRA 媒体的制作和重放。只需要最小量的外加硬件（205, 206, 207）来享受 VRA 调节。 30 缺点是立体声声象将丢失。然而，许多立体声效果是这样的细微，而且重放设备是这样的低保真，以致大多数消费者将更加喜爱 VRA 调节，而不是立体声声象。

图 27 和 28 上讨论的两个实施例代表提供 VRA 给最终用户的最复杂和最简单的可能性。从以上的技术说明中的描述可以想像出带有任意数目的记录的、制作的、或重放的声道的任何实施例，它并不限于图 27 和 28 所示的两个具体的实施例。

5

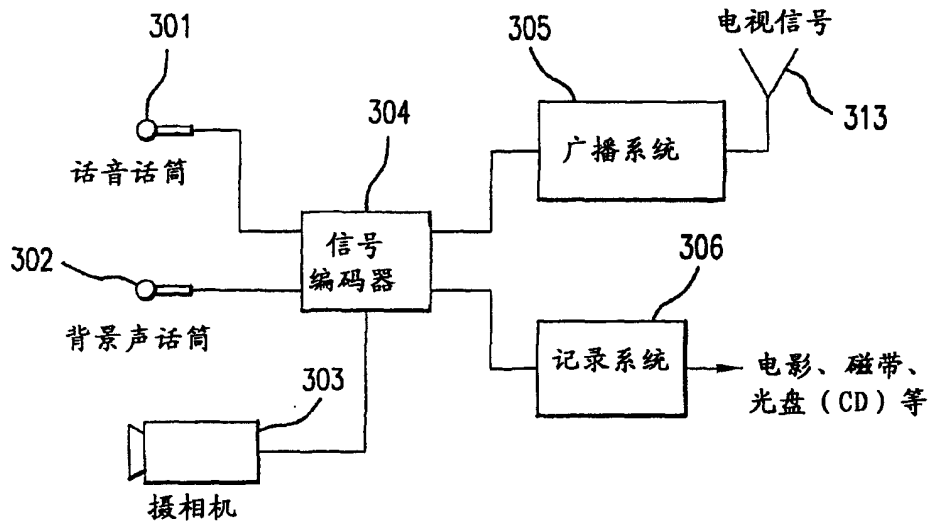


图 1

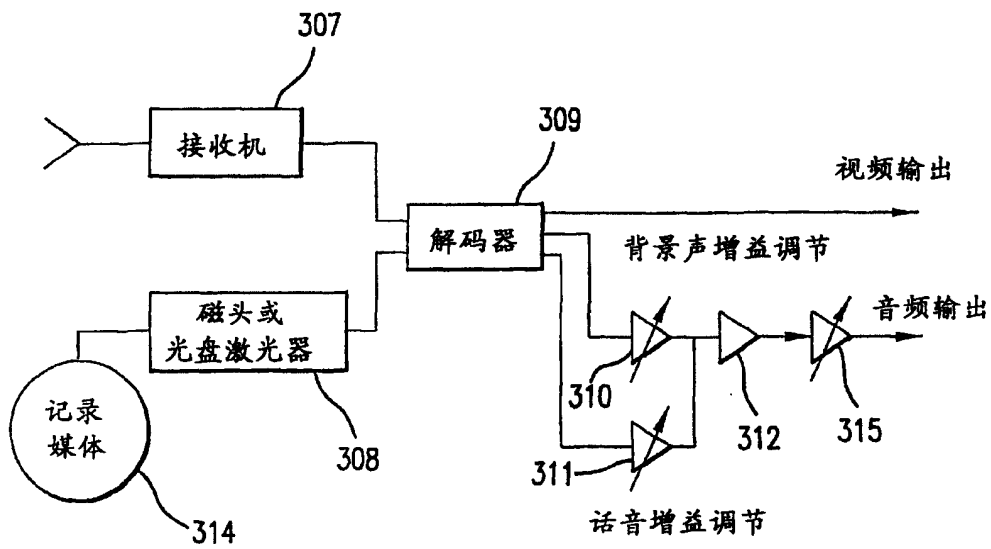


图 2

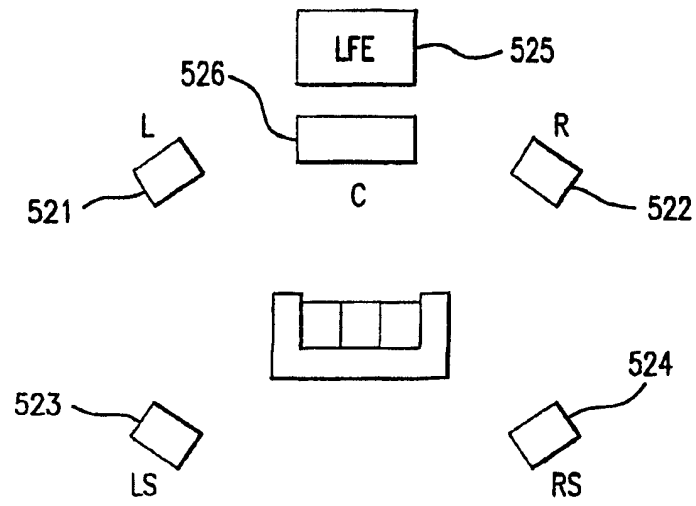


图 3

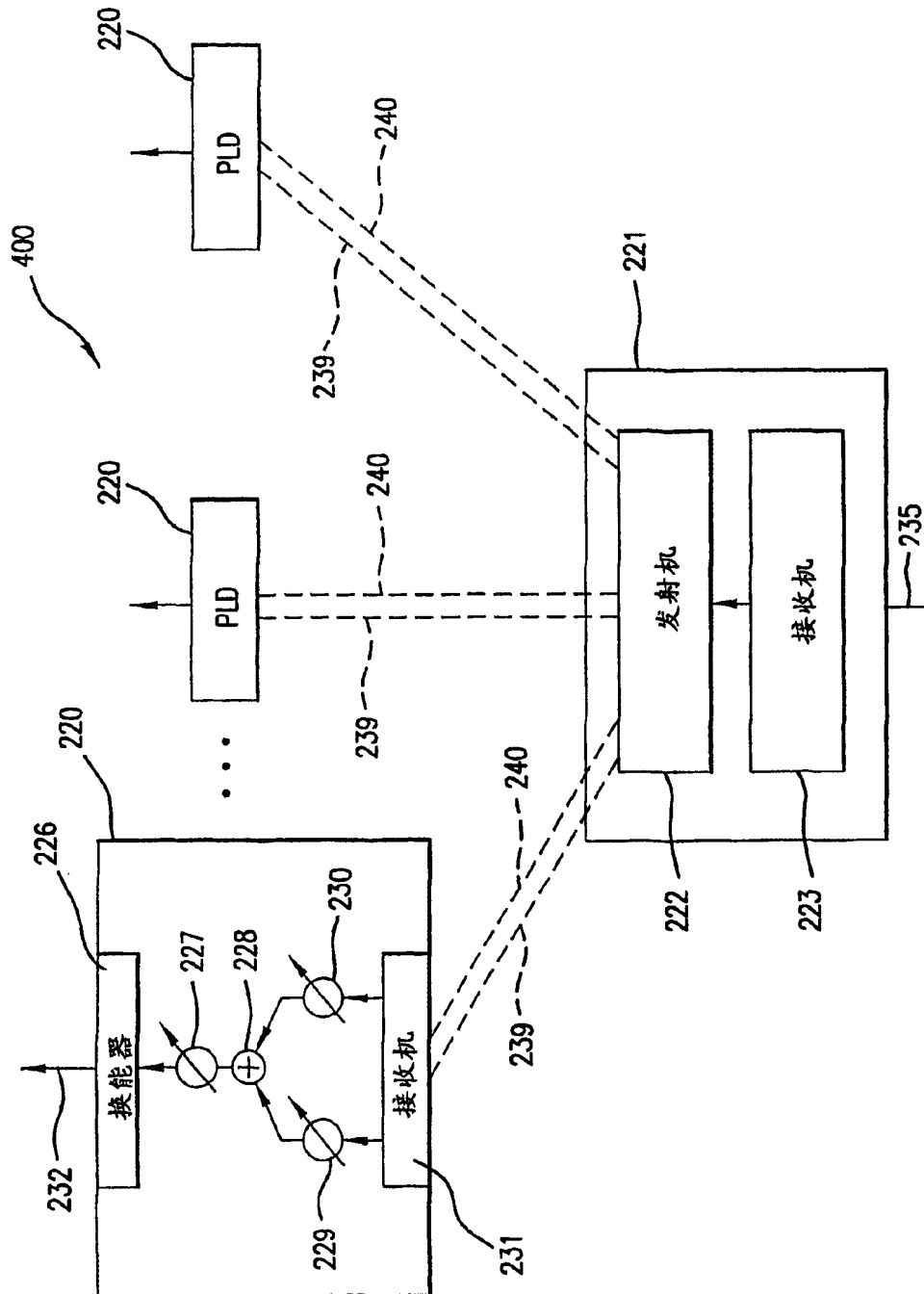


图 4

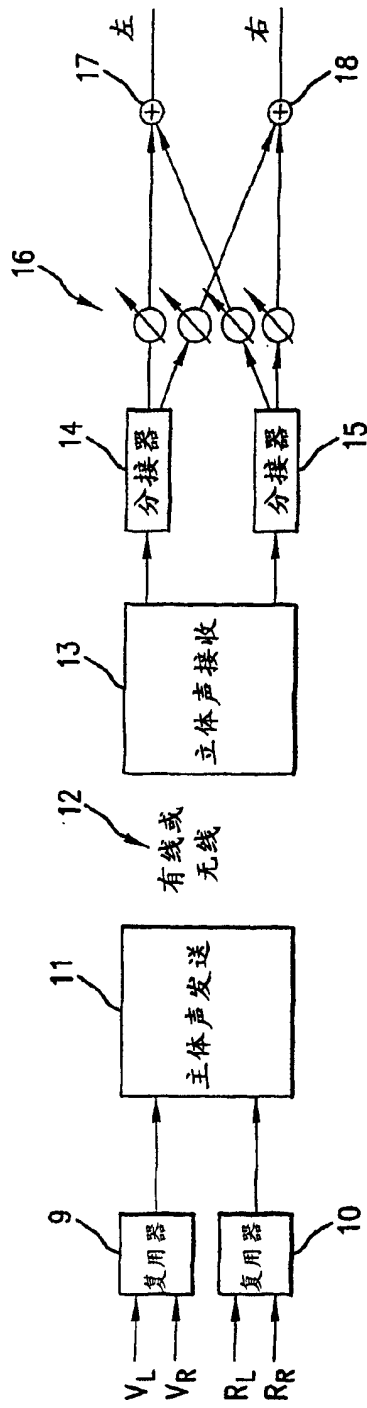


图 5

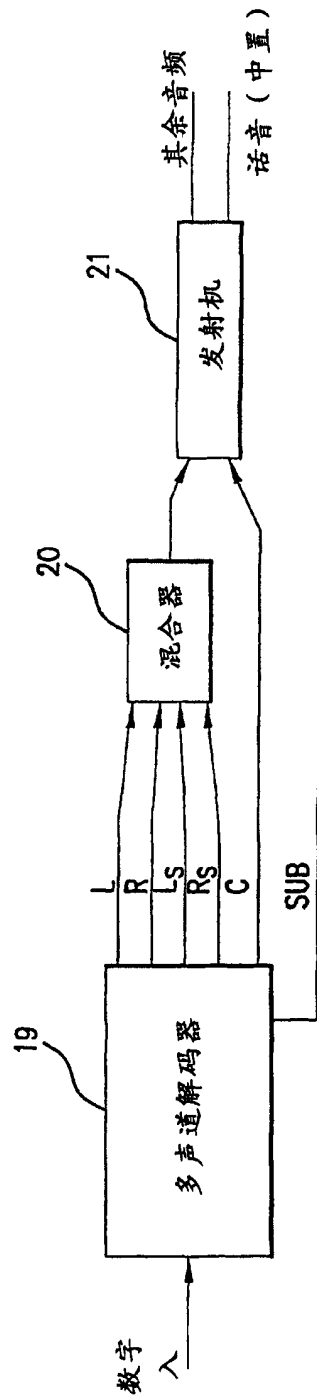


图 6

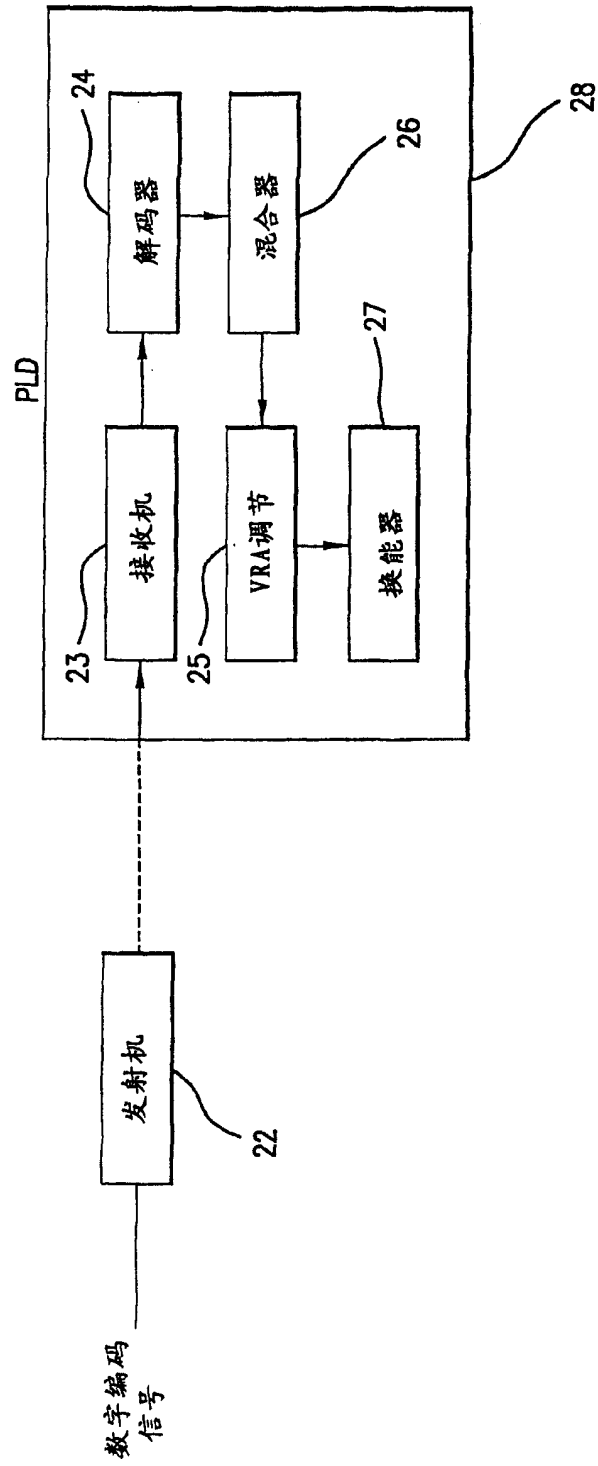


图 7

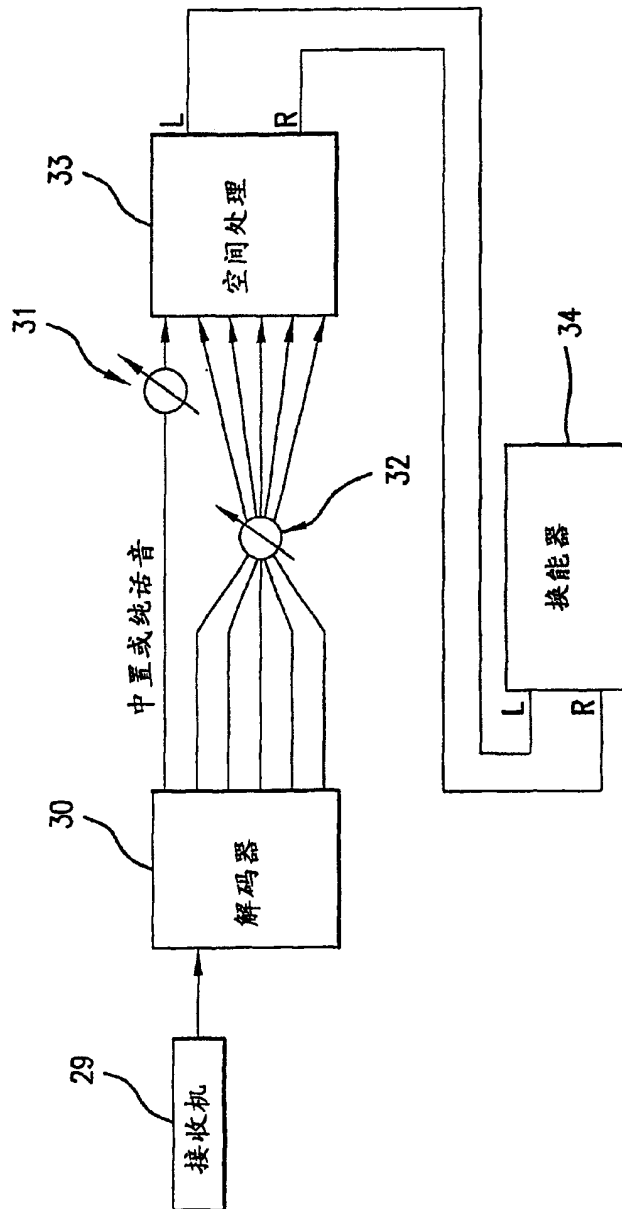


图 8

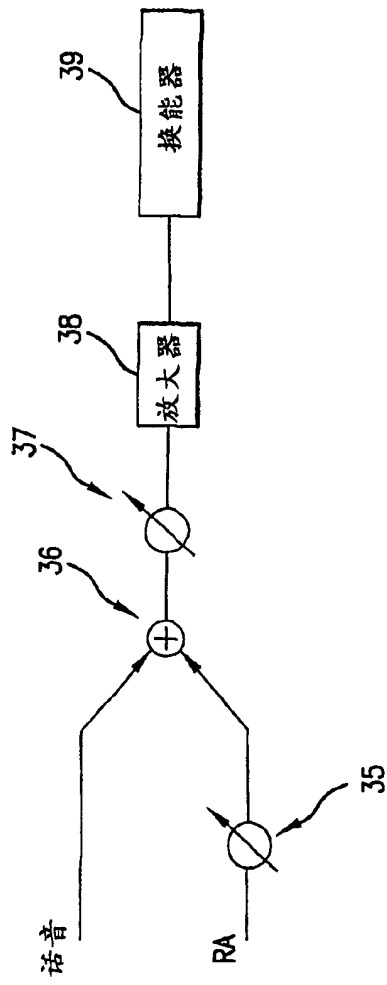


图 9

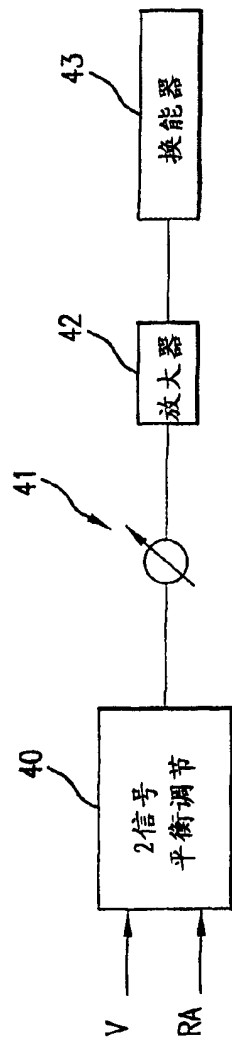


图 10

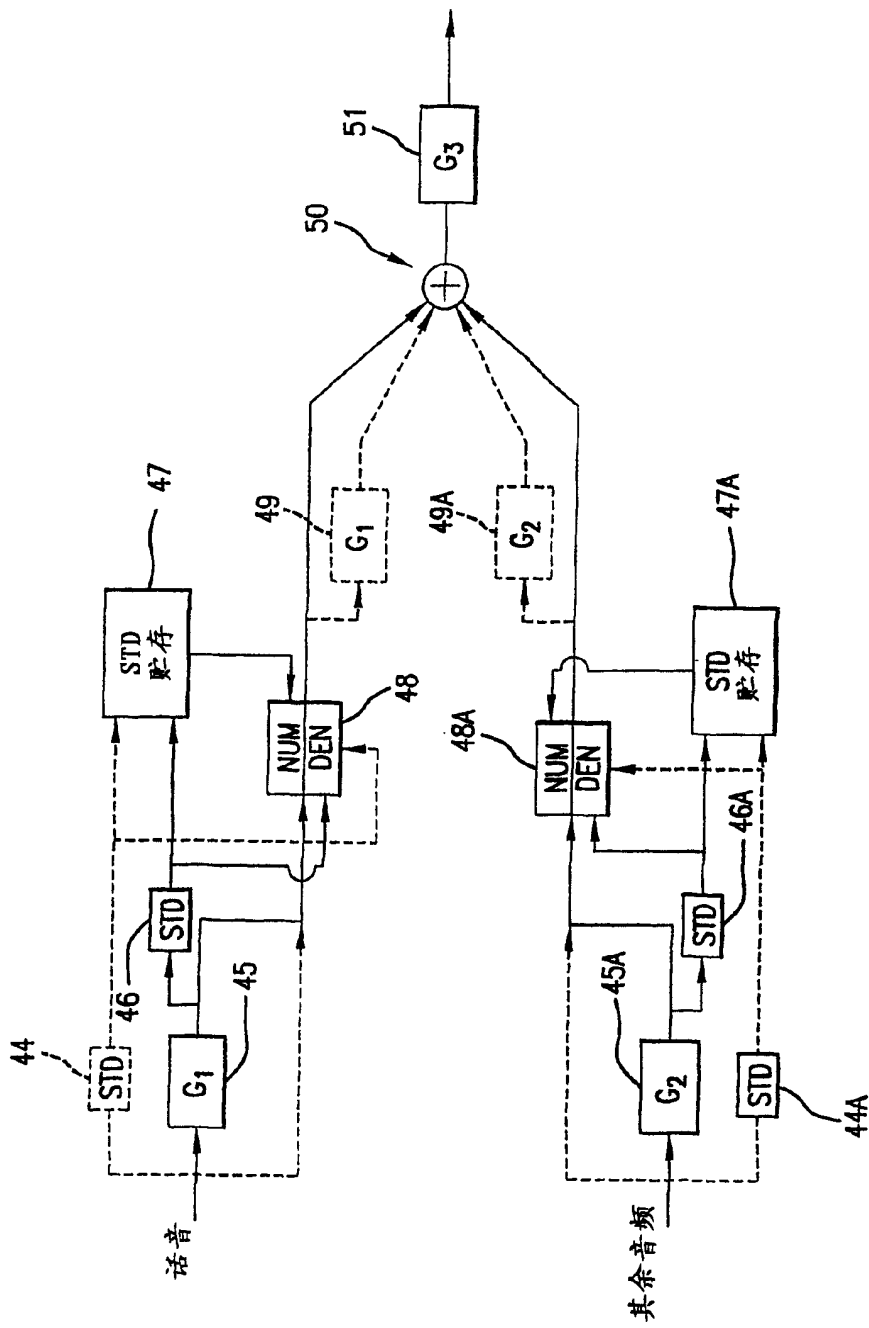


图 11

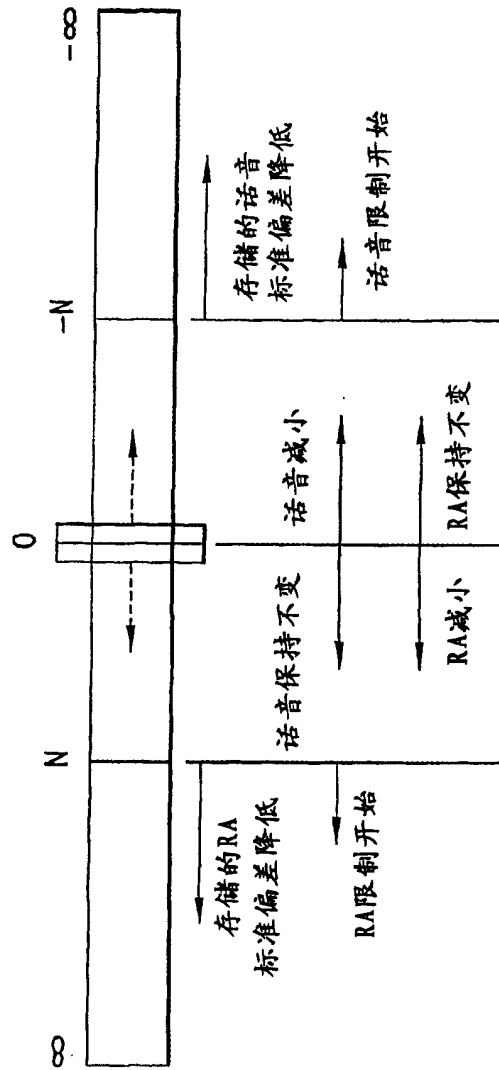


图 12

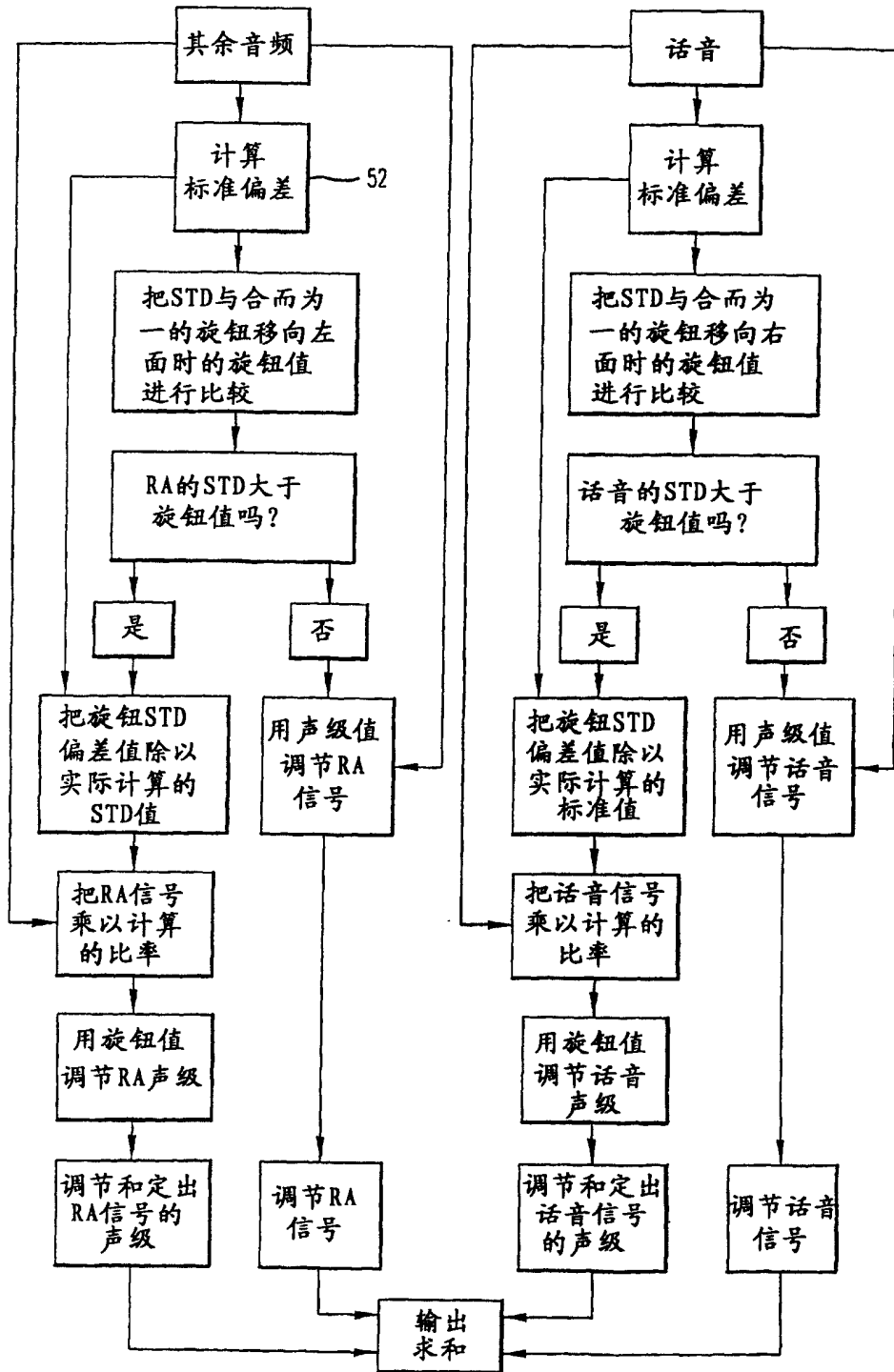


图 13

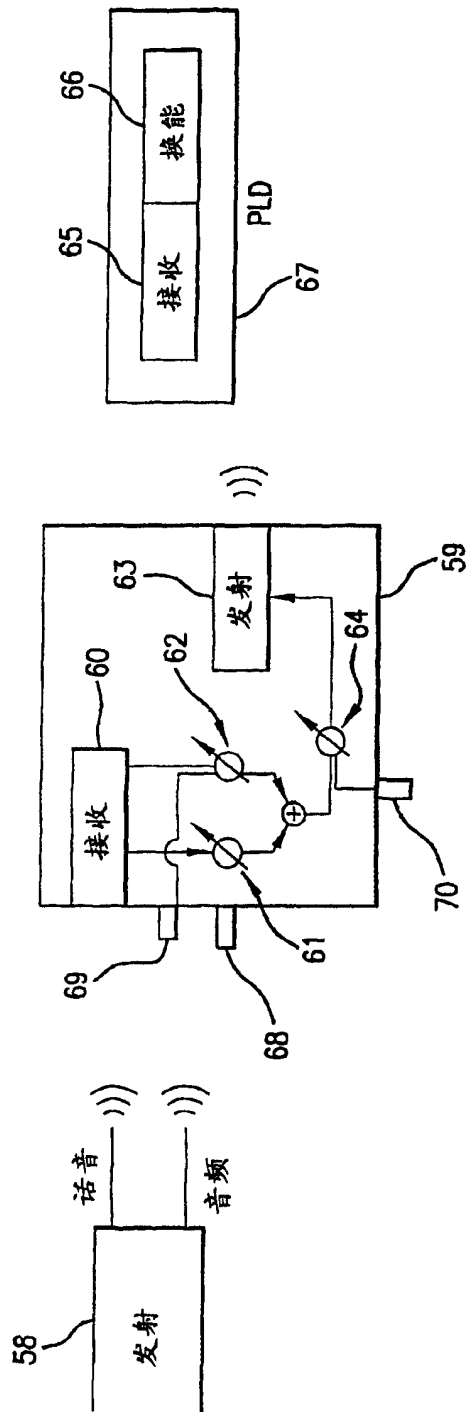


图 14

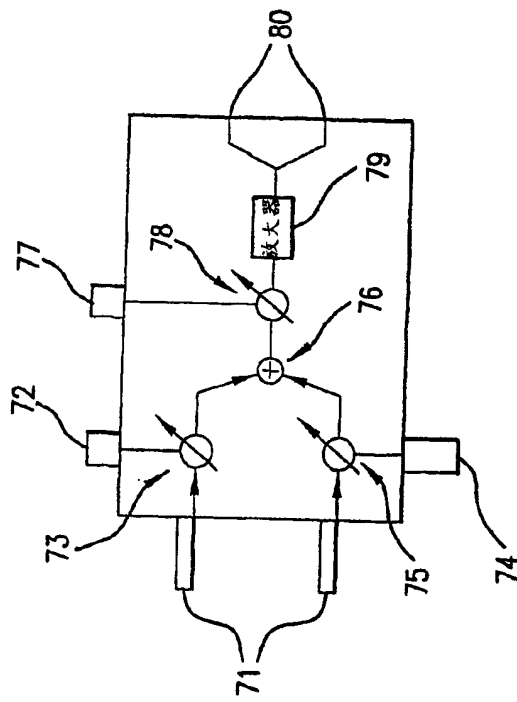


图 15

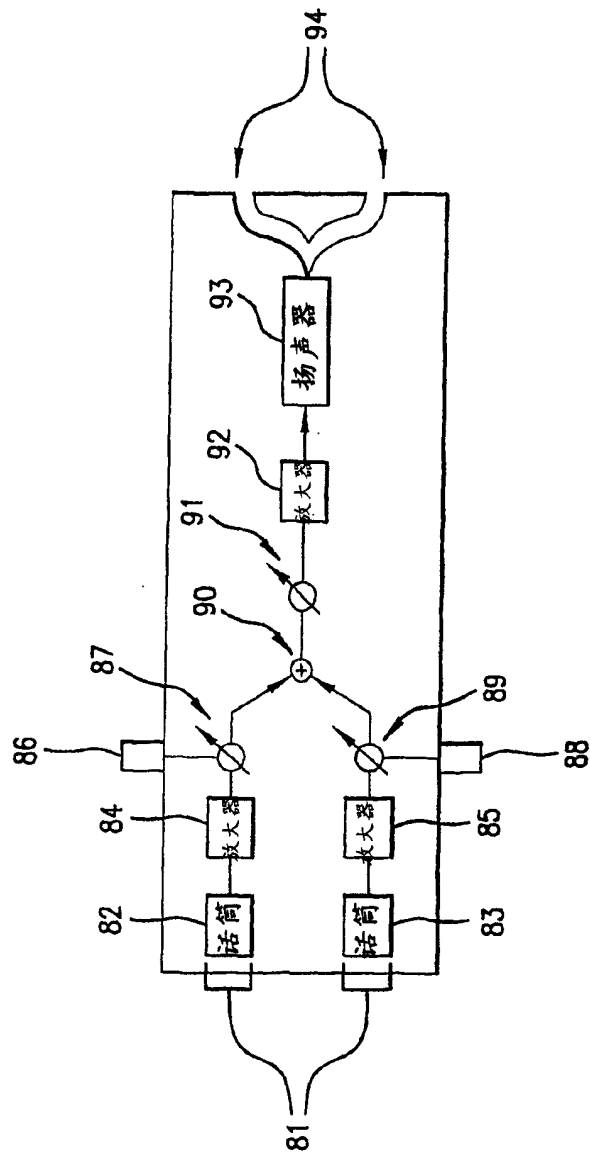


图 16

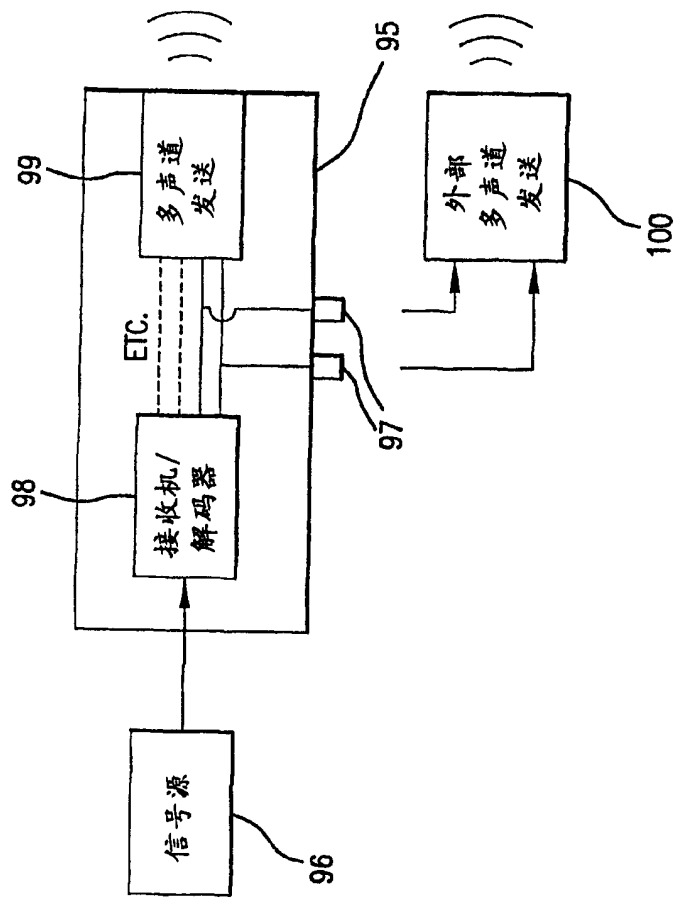


图 17

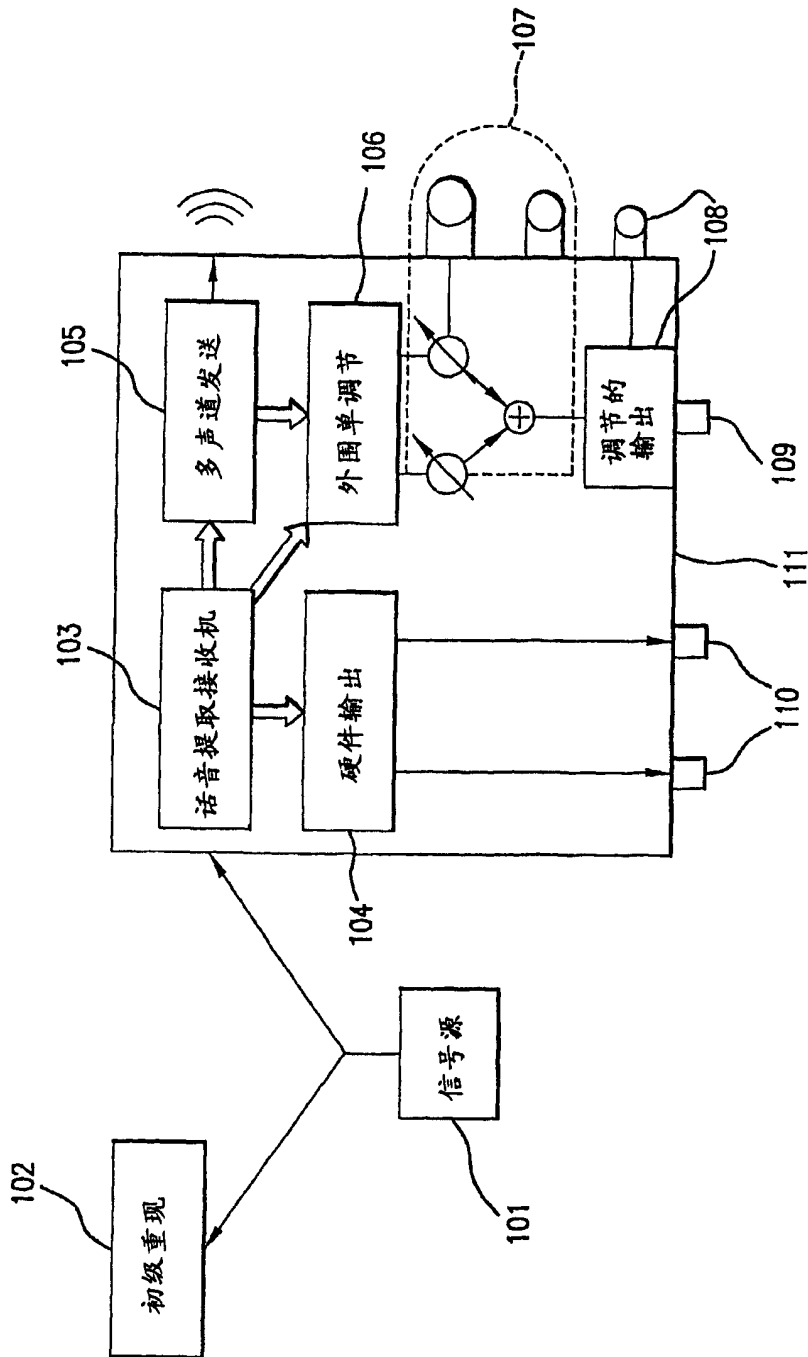


图 18

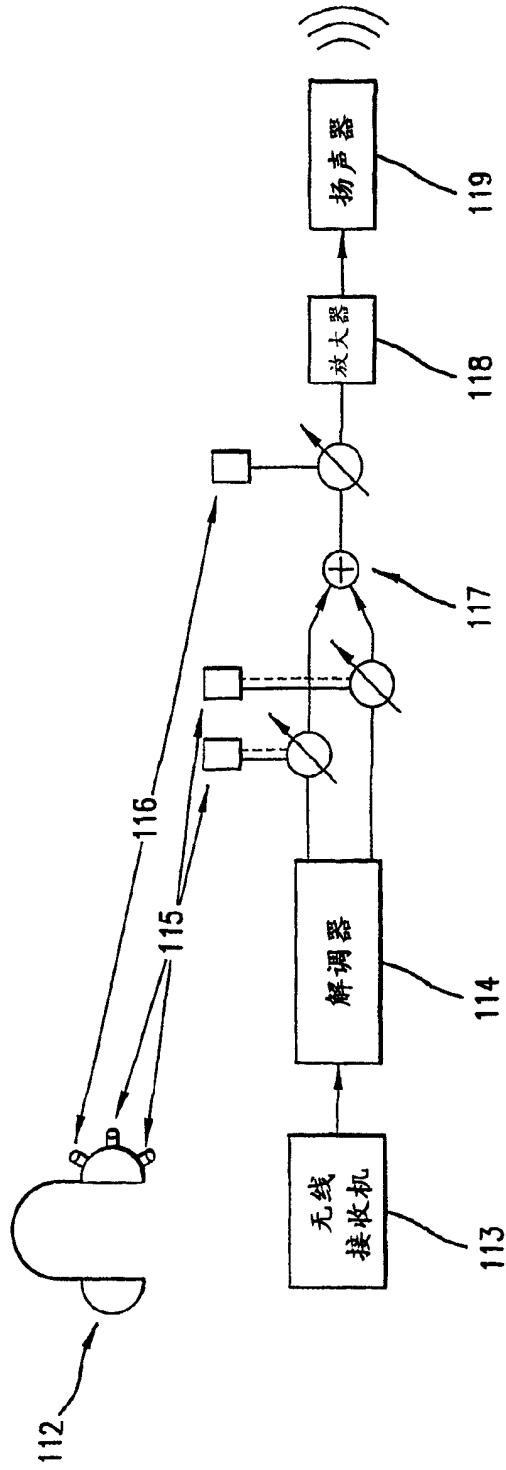


图 19

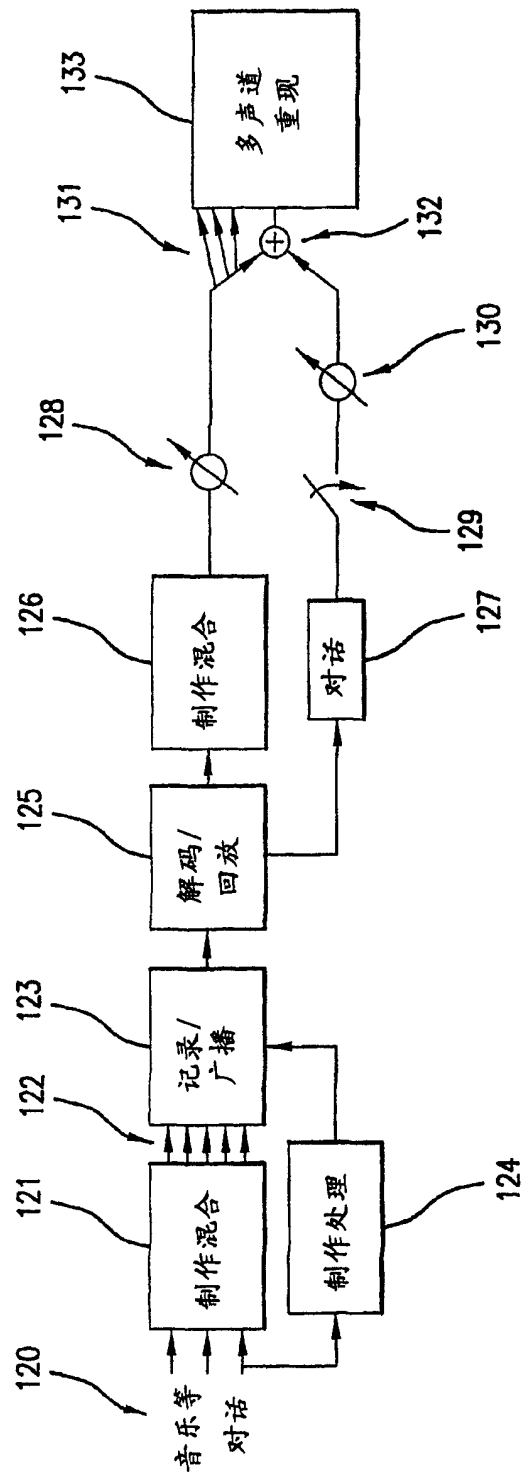


图 20

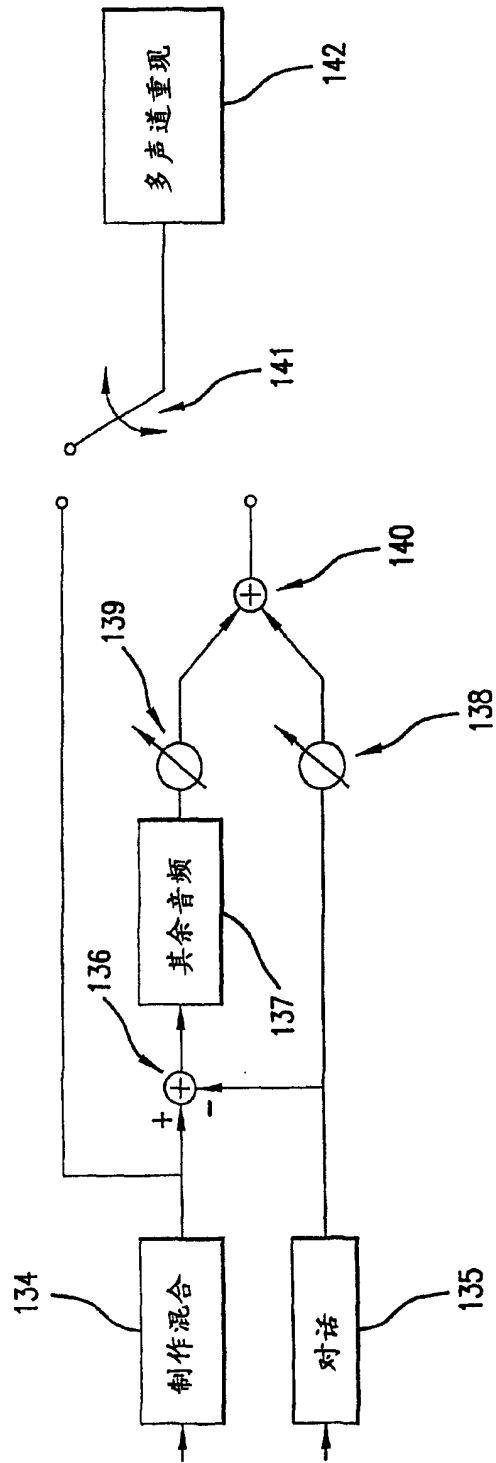


图 21

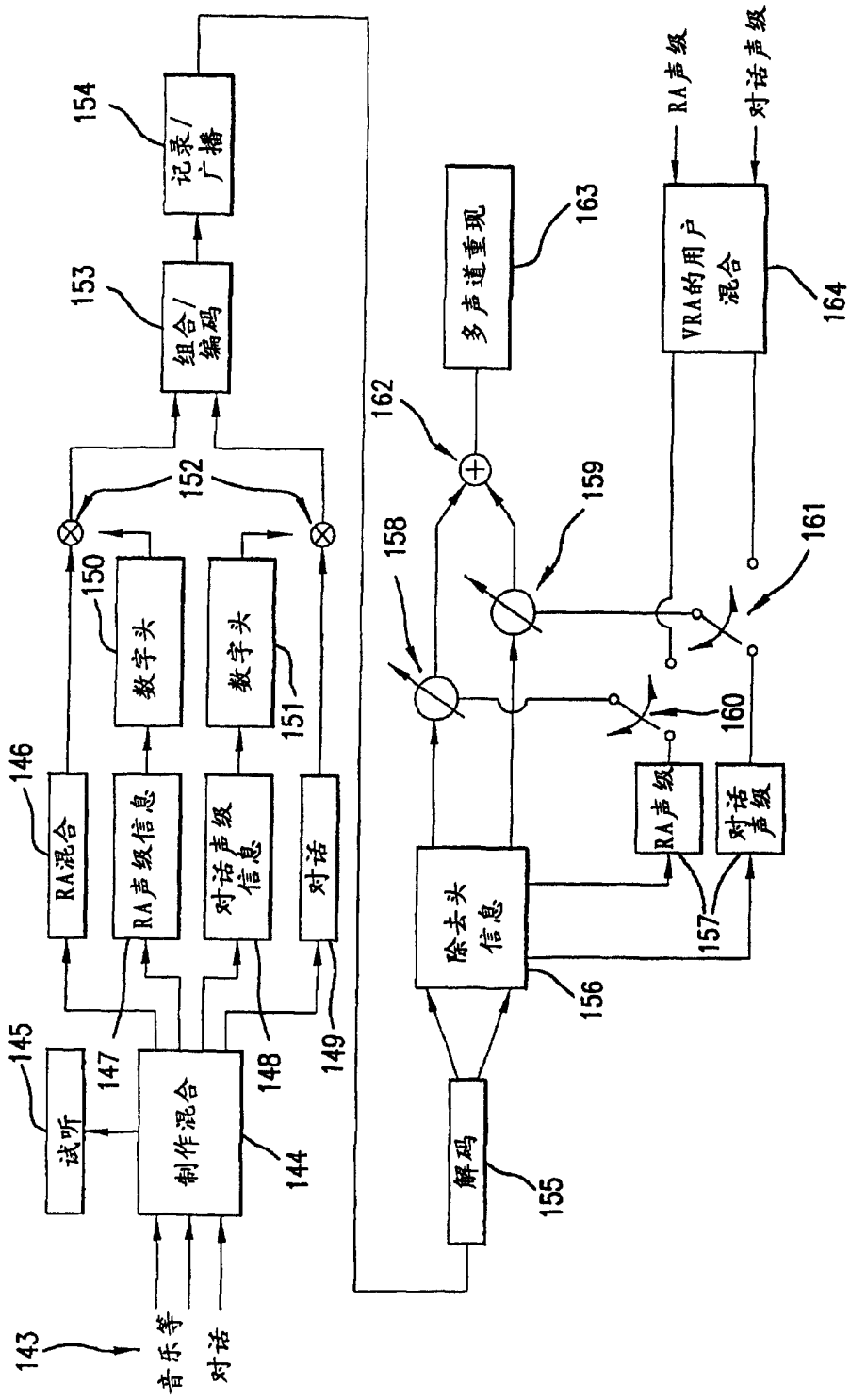


图 22

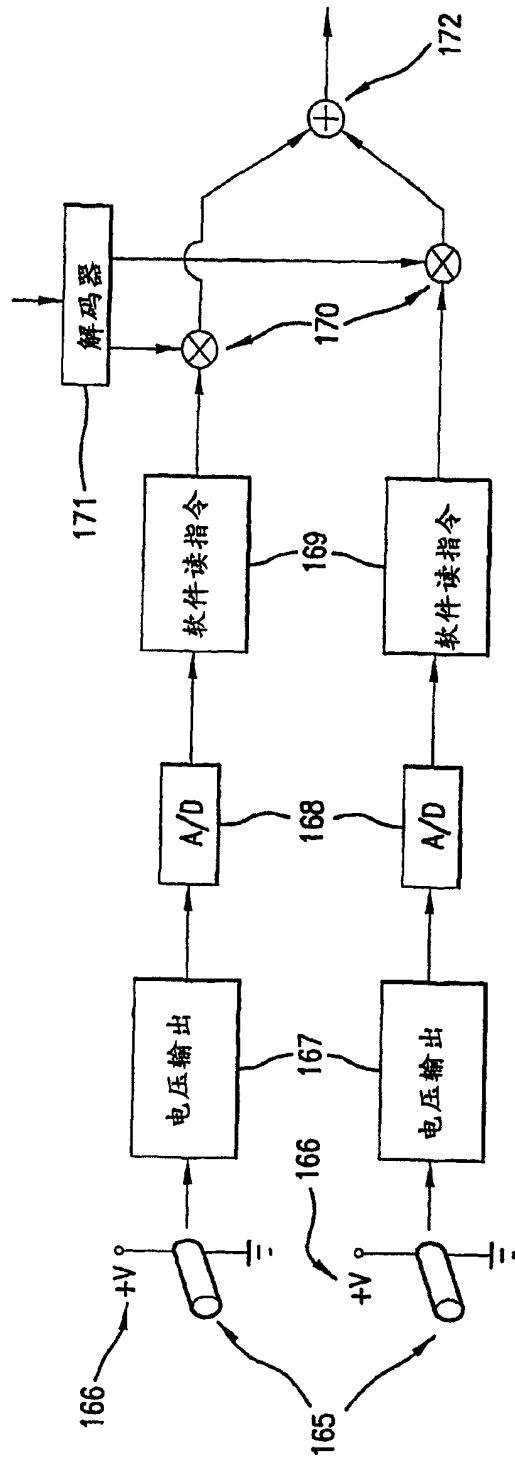


图 23

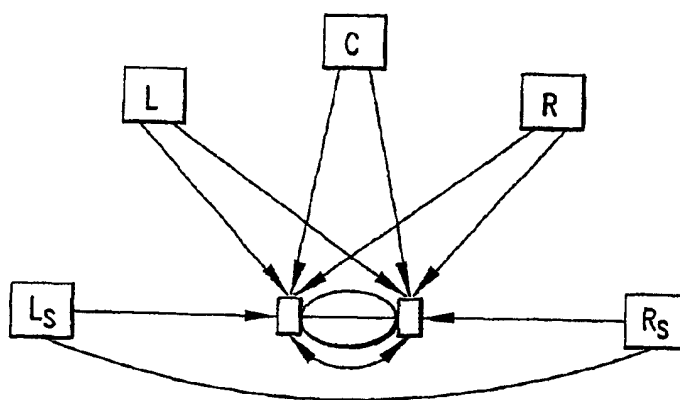


图 24
现有技术

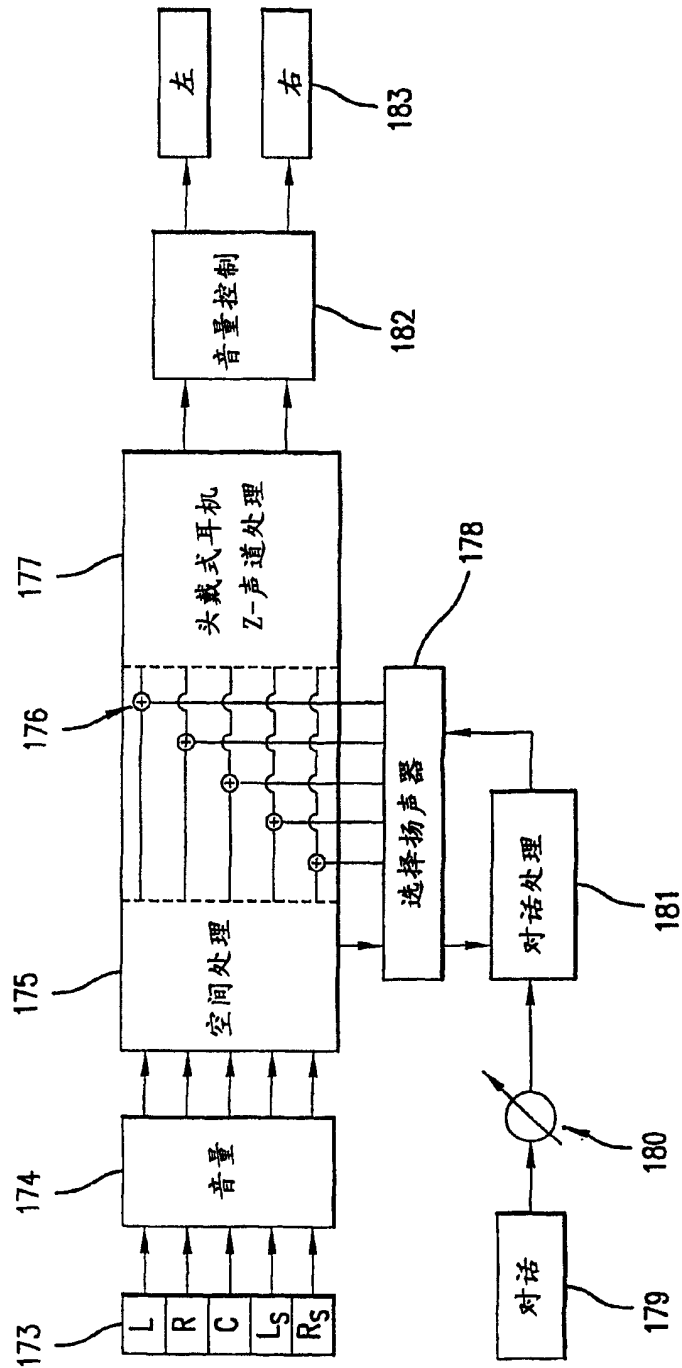


图 25

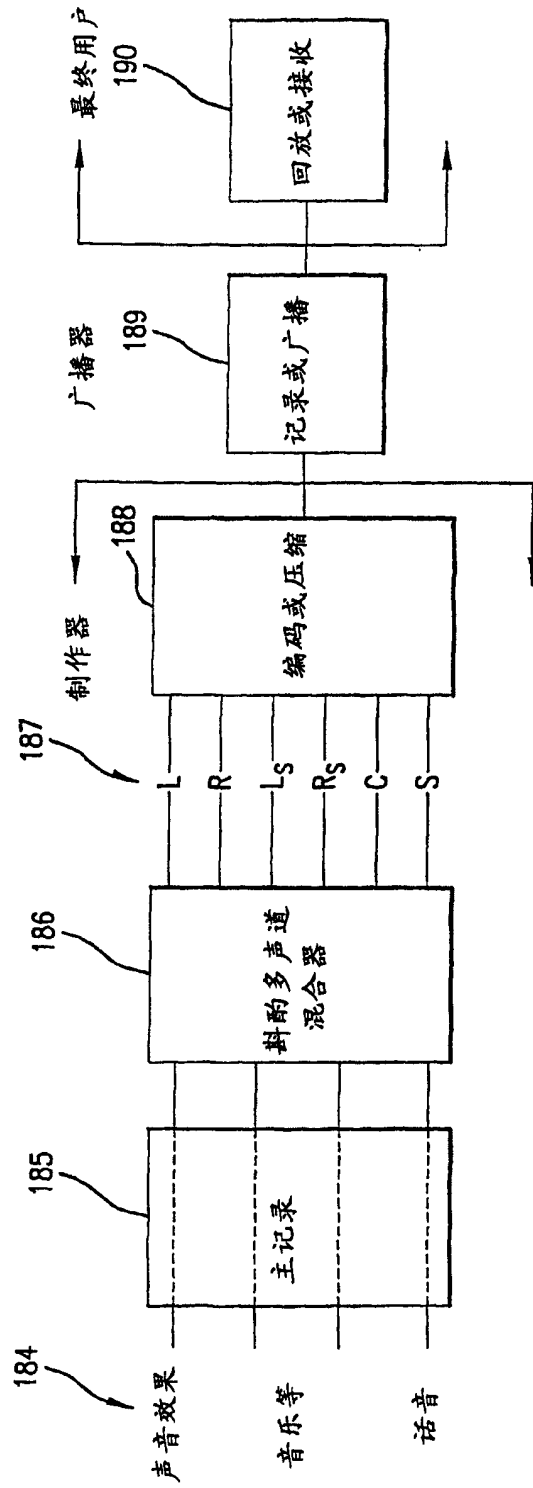


图 26 现有技术

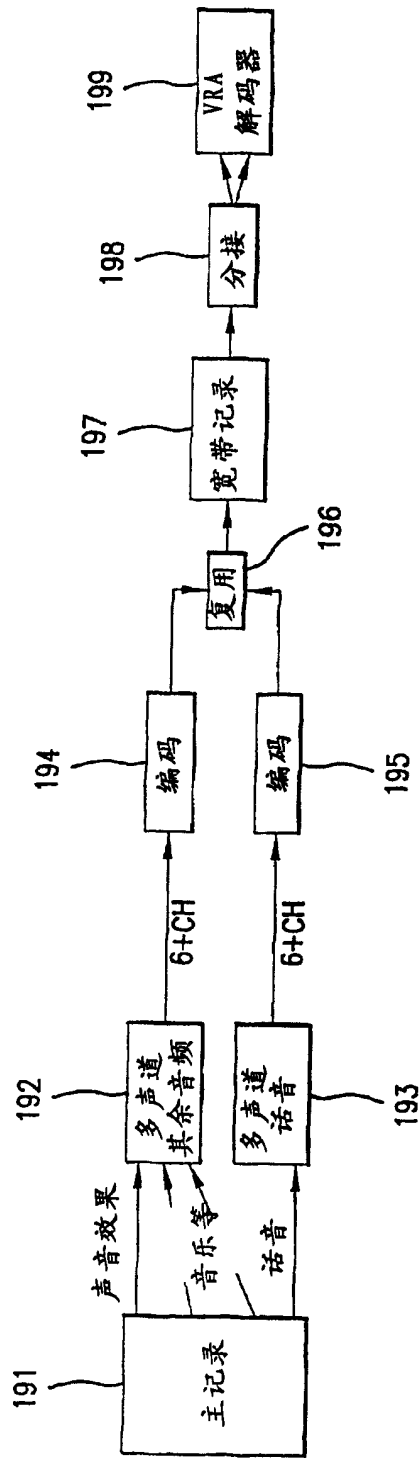


图 27

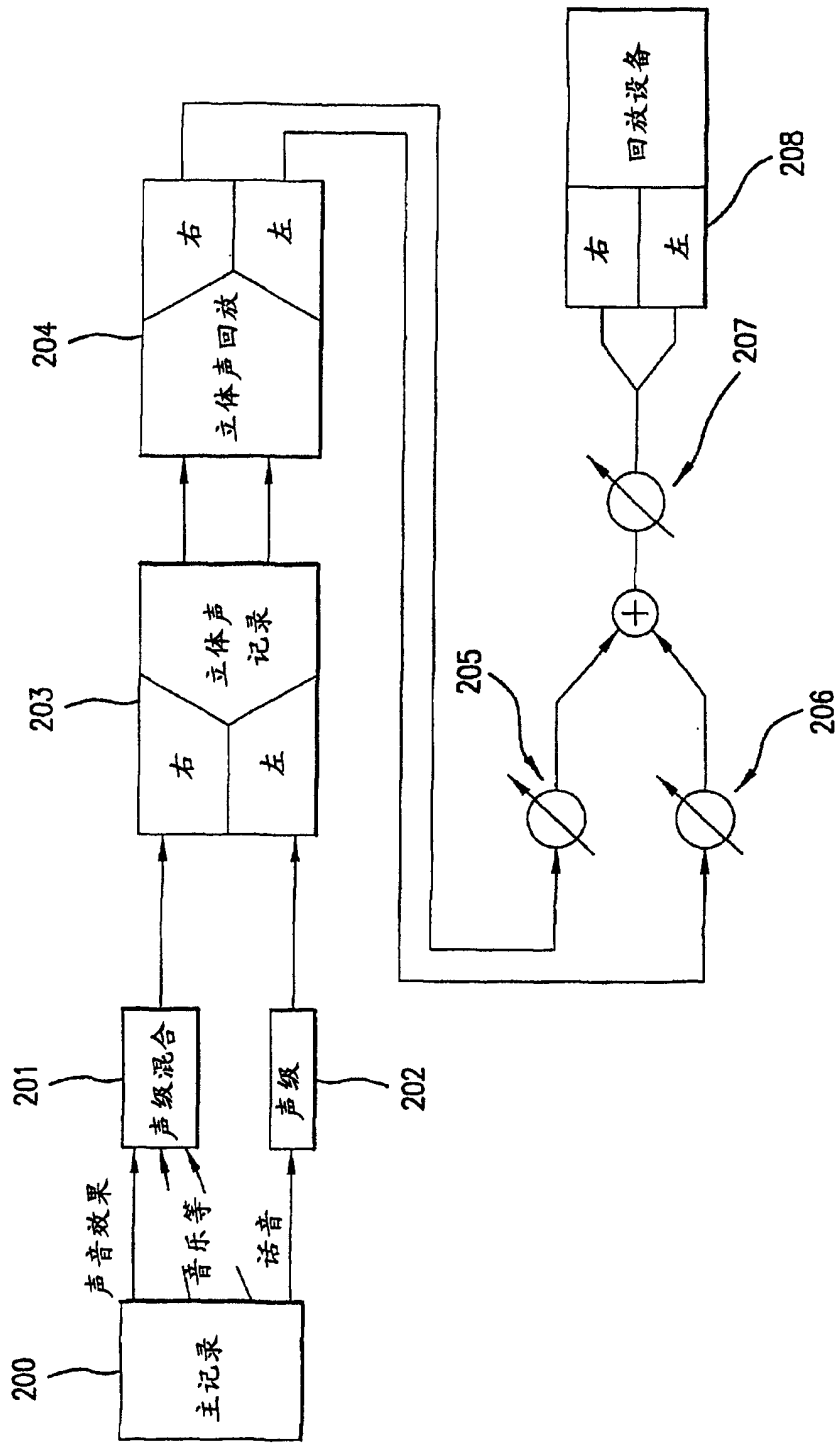


图 28