



(12) 发明专利申请

(10) 申请公布号 CN 102054481 A

(43) 申请公布日 2011. 05. 11

(21) 申请号 201010527304. 5

(22) 申请日 2010. 10. 29

(30) 优先权数据

102009051508. 9 2009. 10. 30 DE

(71) 申请人 大陆汽车有限责任公司

地址 德国汉诺威

(72) 发明人 哈尔罗·赫普肯

卡尔·海因茨·克诺布尔

戴维·肯普夫 汉斯·威廉·吕尔

(74) 专利代理机构 北京康信知识产权代理有限

责任公司 11240

代理人 吴孟秋 李慧

(51) Int. Cl.

G10L 17/00 (2006. 01)

G10L 15/06 (2006. 01)

G10L 15/02 (2006. 01)

G10L 21/02 (2006. 01)

H04M 9/08 (2006. 01)

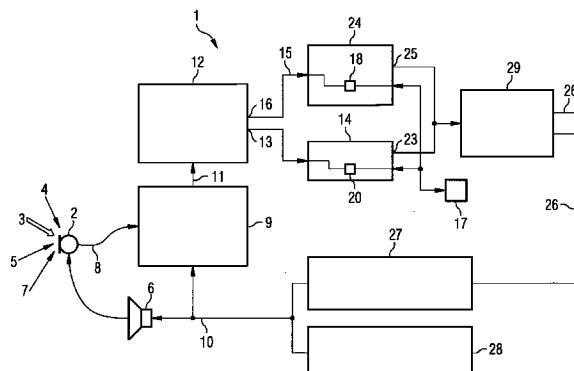
权利要求书 4 页 说明书 7 页 附图 4 页

(54) 发明名称

用于激活和 / 或进行语音对话的装置、系统
和方法

(57) 摘要

本发明涉及用于激活和 / 或进行语音对话的一种装置、一种系统和一种方法。根据本发明的用于激活和 / 或进行语音对话的装置具有语音识别单元、讲话者识别单元以及决策单元。该决策单元设计用于取决于语音识别单元和讲话者识别单元的结果激活结果行为。



1. 一种用于激活语音对话和 / 或进行语音对话的装置 (1 ;30) ,尤其应用在车辆内部空间中,所述装置具有 :
 - 至少一个用于语音信号 (8 ;37) 的输入端 ;
 - 语音识别单元 (24 ;31) ,用于确定在输入的所述语音信号中是否包含以及包含哪些指令 ;以及
 - 在输入端方面与所述语音识别单元的输出端连接的决策单元 (29 ;32) ,所述决策单元适于取决于在所述语音信号 (8 ;37) 中识别出的指令执行结果行为 (26) ,其特征在于,
 - 所述装置附加地包括讲话者识别单元 (14) ;
 - 所述讲话者识别单元适于根据所述语音信号和至少一个存储的讲话者模型确定当前的讲话者 ;
 - 其中所述决策单元附加地在输入端方面与所述讲话者识别单元连接并这样设计,即取决于所述指令的所述结果行为 (26) 的激活至少在一些指令中取决于,是否将所述指令鉴定为来自对应于所述讲话者模型的讲话者。
2. 根据权利要求 1 所述的装置,其特征在于,所述决策单元 (29 ;32) 设计用于把所述讲话者识别单元 (14) 和所述语音识别单元 (24 ;31) 的结果与之前存储在使用者模型中的、针对讲话者的信息相比较和相联系,其中如果所述当前的讲话者未被授权执行所述结果行为,则抑制至少一个取决于指令的结果行为的执行。
 3. 根据权利要求 1 或 2 所述的装置,其特征在于,所述决策单元 (29 ;32) 这样设计,即不取决于对应于所述讲话者模型的所述讲话者的所述识别来执行一些指令。
 4. 根据权利要求 1 至 3 中任一项所述的装置,其特征在于,所述装置设计为用于进行语音对话和激活语音对话的组合装置。
 5. 根据权利要求 1 至 4 中任一项所述的装置,其特征在于,语音分析单元 (24 ;31) 包括用于识别单词的单词识别单元 (19) 以及后续的、用于识别构成指令的结构的结构分析单元 (21) 。
 6. 根据权利要求 1 至 5 中任一项所述的装置,其特征在于,回波补偿单元 (9) 紧邻或非紧邻地前置于所述讲话者识别单元 (14) 和 / 或所述语音识别单元 (24 ;31) ,其中所述回波补偿单元 (9) 具有一个或多个用于扬声器信号 (10 ;34) 的、尤其是用于处理单声道 -、立体声 - 和 / 或多声道 - 扬声器信号 (10 ;34) 的输入端,以及设计用于补偿所述扬声器信号 (10 ;34) 对所述语音信号 (8 ;37) 的影响。
 7. 根据权利要求 6 所述的装置,其特征在于,所述回波补偿单元 (9) 具有用于补偿其他人的语音分量的子单元 (38) ,所述子单元有利地与至少一个用于连接附加的麦克风 (36) 的输入端连接。
 8. 根据前述权利要求中任一项所述的装置,其特征在于,噪声抑制单元 (12) 紧邻或非紧邻地前置于所述讲话者识别单元 (14) 和 / 或所述语音识别单元 (24 ;31) 。
 9. 根据权利要求 1 至 8 中任一项所述的装置,其特征在于,所述讲话者识别单元 (14) 和 / 或所述语音识别单元 (24 ;31) 设计用于使通过所述讲话者识别单元 (14) 识别的所述讲话者向所述决策单元 (29 ;32) 的输出与由所述语音识别单元 (24 ;31) 识别的指令的输出同步。

10. 根据权利要求 1 至 9 中任一项所述的装置,其特征在于,所述讲话者识别单元 (14) 设计用于通过从所述语音信号中提取讲话者特征并把所述讲话者特征与存储的取决于讲话者的特征相比较来鉴定所述当前的讲话者,以及所述讲话者识别单元有利地包括另外的单元 (20),所述另外的单元设计用于讲话者适应,也就是说用于不断确定的细化的取决于讲话者的特征以及用于把所述细化的取决于讲话者的特征存储在存储的所述讲话者模型中。

11. 根据权利要求 1 至 10 中任一项所述的装置,其特征在于,并联于所述讲话者识别单元 (14) 和所述语音识别单元 (24 ;31) 布置了用于根据所述语音信号 (8 ;37) 检测驾驶员的状态的驾驶员状态检测单元 (39)。

12. 根据权利要求 1 至 11 中任一项所述的装置,其特征在于,所述语音识别单元包括附加单元 (18),所述附加单元设计用于检测所述讲话者的所述讲话者特征的取决于时间的变化作为属性并将所述变化存储在对应于所述讲话者存储的所述讲话者模型中。

13. 根据权利要求 1 至 12 中任一项所述的装置,其特征在于,所述装置具有至少一个存储装置 (17),所述存储装置尤其可以设计用于存储所述使用者模型和 / 或所述讲话者模型。

14. 根据权利要求 13 所述的装置,其特征在于,所述至少一个存储装置 (17) 具有用于输出和 / 或输入存储的所述讲话者模型和 / 或所述使用者模型的输入端和 / 或输出端。

15. 根据前述权利要求中任一项所述的装置,其特征在于,所述装置这样设计,即在所述结果行为 (26) 的所述执行期间也激活所述装置以用于分析所述语音信号 (8 ;37),所述装置尤其这样设计,即在识别来自授权的所述讲话者的指令时,至少部分地中断由在前的指令触发的所述结果行为的所述执行。

16. 一种用于激活语音对话和 / 或进行语音对话的系统,所述系统具有 :根据权利要求 1 至 15 中任一项所述的装置;至少一个麦克风 (2 ;35) 和至少一个扬声器 (6 ;33)。

17. 根据权利要求 16 所述的系统,其特征在于,所述系统包括这样布置的多个麦克风 (2) 或至少一个麦克风阵列 (25),即通过所述麦克风的方向特性给定的最佳接收的区域覆盖了在授权的讲话者的推测的停留区域中的至少一些所述麦克风 (2 ;35)。

18. 根据权利要求 17 所述的系统,其特征在于,所述麦克风设计为自动地对准通过所述麦克风检测的、所述讲话者的位置。

19. 一种用于激活语音对话和 / 或进行语音对话的方法,包括以下步骤 :

- 接收语音信号 (8 ;37) ;
- 从所述语音信号 (8 ;37) 中识别指令或指令结构 ;
- 根据所述语音信号 (8 ;37) 和至少一个存储的讲话者模型识别讲话者 ;
- 随后取决于识别的所述指令和识别的所述讲话者执行结果行为 (26)。

20. 根据权利要求 19 所述的方法,其特征在于,识别所述指令或所述指令结构的步骤包括子步骤 :

- 识别包含在所述语音信号中的单词,
- 识别由所述单词构成的所述指令结构。

21. 根据权利要求 19 或 20 所述的方法,其特征在于,在识别所述指令之前和 / 或在识别所述讲话者之前这样进行回波补偿,即从所述语音信号中去除由于在乘客车厢中的反射

产生的扬声器信号的叠加,其方法是计算由于所述扬声器信号产生的所述叠加并从所述语音信号中减去所述叠加。

22. 根据权利要求 21 所述的方法,其特征在于,确定所述语音信号 (8 ;37) 的由其他人引起的语音信号分量 (4),并至少部分地从所述语音信号 (8 ;37) 中去除。

23. 根据权利要求 19 至 22 中任一项所述的方法,其特征在于,所述语音信号 (8 ;37) 设计为多声道语音信号 (37)。

24. 根据权利要求 23 所述的方法,其特征在于,通过所述多声道语音信号 (37) 的不同声道的不同时间推移的信号的在时间上的联系,提取并分离来自授权的所述讲话者的地点之一的、所述语音信号 (8 ;37) 的所述分量。

25. 根据权利要求 19 至 24 中任一项所述的方法,其特征在于,所述结果行为的所述执行与识别的所述指令之间的相关性在于,当没有识别所属的所述讲话者时或所属的所述讲话者没有授权对所述结果行为进行指示时,抑制所述结果行为 (26) 的所述执行。

26. 根据权利要求 19 至 25 中任一项所述的方法,其特征在于,在识别所述指令或所述讲话者之前进行噪声抑制。

27. 根据权利要求 19 至 26 中任一项所述的方法,其特征在于,为了识别授权的所述讲话者,从所述语音信号 (8 ;37) 中提取讲话者特征并把所述讲话者特征与存储在所述讲话者模型中的单个的讲话者特征比较。

28. 根据权利要求 27 所述的方法,其特征在于,进行讲话者适应,通过所述讲话者适应,持续地细化和补充了存储在所述讲话者模型中的所述单个的讲话者特征。

29. 根据权利要求 19 至 28 中任一项所述的方法,其特征在于,检测所述讲话者特征的取决于时间的变化作为属性并将所述变化存储在所述讲话者模型中。

30. 根据权利要求 19 至 29 中任一项所述的方法,其特征在于,为了识别所述指令,从所述语音信号 (8 ;37) 中提取语音特征并把所述语音特征与存储在所述讲话者模型中的单个的语音特征相比较。

31. 根据权利要求 30 所述的方法,其特征在于,进行所述讲话者适应,通过所述讲话者适应,持续地细化和补充了存储在所述讲话者模型中的所述单个的语音特征。

32. 根据权利要求 19 至 31 中任一项所述的方法,其特征在于,检测所述语音特征的取决于时间的变化作为属性并存储在所述讲话者模型中。

33. 根据权利要求 19 至 32 中任一项所述的方法,其特征在于,根据所述语音信号 (8 ;37) 检测驾驶员状态。

34. 根据权利要求 19 至 33 中任一项所述的方法,其特征在于,可通过输入其它语音信号来中断所述结果行为的后续的执行,所述其它语音信号包含授权的所述讲话者的其它指令。

35. 根据权利要求 19 至 34 中任一项所述的方法,其特征在于,根据所述语音信号确定授权的所述讲话者的所述地点,以及不取决于包含于所述语音信号中的所述指令而产生用于使至少一个麦克风对准授权的所述讲话者的所述地点的控制信号。

36. 根据权利要求 19 至 35 中任一项所述的方法,其特征在于,所述结果行为的所述执行包括输出语音对话信号。

37. 根据权利要求 19 至 36 中任一项所述的方法,其特征在于,所述结果行为 (26) 的

所述执行包括所述控制信号的信号输出,所述控制信号用于控制集成于车辆中的装置的功能。

用于激活和 / 或进行语音对话的装置、系统和方法

技术领域

[0001] 本发明涉及用于激活语音对话和 / 或进行语音对话的一种装置、一种系统和一种方法。

背景技术

[0002] 语音对话系统用于电子仪器、例如车辆电脑，而不需要对触觉的操作运行进行控制。根据现有技术已知，用于进行语音对话或用于激活语音对话的装置根据指令搜索通过噪声抑制或回波补偿而变得清晰的语音信号。在用于激活语音对话的装置中，根据确定的关键词进行搜索，该关键词用于激活语音对话。在用于进行语音对话的装置中，检查该语音信号是否包含作为指令而存储在集成于语音对话系统中的指令数据库中的单词。如果识别出指令，那么执行属于该命令的行为。

[0003] 尤其在干扰噪声作用于系统的应用条件下，这样的用于进行语音对话或用于激活语音对话的装置是不可靠的。这尤其在应用于轿车中时是不利的，因为在干扰影响下、例如乘客的谈话，根据现有技术的装置常常由于错误识别的命令而执行不期望的功能。

发明内容

[0004] 因此本发明的目的在于，特别抗干扰地进行语音对话和 / 或激活语音对话。

[0005] 该目的通过一种根据权利要求 1 所述的装置和一种根据权利要求 19 所述的方法实现。在从属权利要求中给出有利的改进方案。

[0006] 指令分析的可靠性通过如下方式显著提高，即该装置不仅包括用于确定语音信号中是否包含指令的语音识别单元和用于取决于在语音信号中识别的指令而产生输出信号的决策单元，还包括用于根据语音信号和至少一个存储的讲话者模型而识别讲话者的讲话者识别单元，其中这样设计决策单元，即取决于指令的结果行为的执行取决于，至少部分产生语音信号的讲话者是否被讲话者识别单元识别为相应于存储的讲话者模型的讲话者。

[0007] 语音识别单元可理解为一个设计用于鉴定来自语音信号的指令和 / 或指令结构的单元。也就是说，根据语音模型和根据存储的指令，分析语言信号中是否包含 / 包含哪些指令或指令结构。如果包含指令，那么它被识别为相应的有效指令。

[0008] 在这种情况下，指令不一定只是单个的词。在此更确切地说意味着语音输入的每种形式。例如，概念“指令”也意味着数列 / 号码序列、句子成分和 / 或完整的句子。

[0009] 讲话者识别单元可理解为可以特别设计用于从语音信号中提取讲话者特征并用于把该讲话者特征与之前存储在至少一个讲话者模型中的、取决于讲话者的特征相比较的单元，其中根据比较确定当前的讲话者和对应于讲话者模型的讲话者一致或不一致。有利地这样设计讲话者识别单元，即当语音信号完全来自对应于讲话者 模型的讲话者时，也就是说多人同时讲话而信号有利地不对应于任何讲话者时，才确定一致。

[0010] 待执行的结果行为可以是不同类型的结果行为。

[0011] 在用于激活语音对话的装置中，结果行为例如可以是双重事件的输出，这作为信

息仅包括是否应该进行语言对话的激活的决定。可替换地,该事件可以包括关于引起激活的讲话者的信息,从而产生取决于讲话者的激活信号。

[0012] 在用于进行语音对话的装置中,作为结果行为的语音输出尤其重要。可以利用它将信息、例如其它选项通知给讲话者。例如如果讲话者要调用关于位于附近的餐馆的信息并使用了与之相应的指令,那么语音对话系统可以要求讲话者从多个可能的餐馆类型中作出选择。

[0013] 其它的可能的结果行为是控制车辆功能的控制事件。例如可以向用于座位加热的控制系统发送用于提高加热温度的事件。此外,可以输出用于调用位于车辆中的其它系统的控制信号。此外结果行为还可能由控制事件和语音输出组成。本领域技术人员由根据现有技术的用于进行语音对话的装置还已知了其它的可能的输出信号。

[0014] 此外,本发明还涉及一种方法:其中首先接收语音信号;其中随后识别存在于语音信号中的指令或指令结构;并且其中根据语音信号和至少一个讲话者模型识别讲话者;其中在识别指令和讲话者之后,取决于讲话者和指令的识别的结果执行结果行为。

[0015] 这样有利地设计决策单元,即当由讲话者识别单元将包含指令的语音信号鉴定为来自授权的讲话者时,那么仅仅执行取决于指令的结果行为。也就是说将讲话者识别单元和语音识别单元的结果与存储在使用者模型中的数据比较,并检查讲话者是否被授权。

[0016] 根据实施例,可以反馈缺乏授权的情况。这种反馈的优点在于,讲话者可以在未识别或识别之间区分作为未充分授权的讲话者。

[0017] 重要的是,通过由无资格的讲话者输入的指令不会触发对应于指令的待执行的结果行为,也就是说,来自无资格的讲话者的指令不会导致结果行为或导致与授权的讲话者不同的结果行为。

[0018] 在一个有利的实施例中,可以不取决于讲话者执行特定的指令。由此例如可以控制不对驾驶员造成妨碍的车辆功能。在这里对于用于进行语音对话的装置来说,对副驾驶座和/或后排座的加热可以作为一个实例,该加热也可以由其他的作为授权的讲话者的人員进行调节。然而有利地,在输入这些指令时不进行语音对话,也就是说,语音对话系统对未识别的讲话者的指令反应为没有语音输出的纯语音控制,以用于控制毫无疑问的车辆功能。在用于激活语音对话的装置中也可以建立不取决于讲话者的指令。因此例如可以这样实现激活,即可接收通过指令分析单元产生的激活信号,可以期望的是具有缩小的功能范围的乘客访问(Passagierzugriff)。例如在这种乘客访问中,进行作为纯语音控制的激活,其丝毫没有输出声学的反馈且仅设计用于执行毫无疑问的车辆功能的控制指令。

[0019] 在另一有利的实施例中,装置设计为用于激活-和进行语音对话的组合装置。也就是说,在未激活状态中,该装置用于激活语音对话。如果该装置通过关键词由授权的讲话者激活,那么其作为用于进行语音对话的装置工作。

[0020] 对这样作为组合系统的设计来说尤其可能的是,定义多个不同的关键词。该装置可以通过第一关键词由授权的讲话者这样激活,即该装置仅执行授权的讲话者的随后的其它指令。此外,可以构造第二关键词,这使该装置进入一种状态,在该状态中装置执行任意讲话者的指令。然而有利地,授权的讲话者的指令具有更高的优先权,也就是说,授权的讲话者的指令不会被其他人的随后的指令打断,然而甚至可以中断执行其他人的指令。

[0021] 在一个有利的实施例中,语音识别单元具有单词识别单元和后续的结构识别单

元。单词识别单元设计用于这样识别来自语音信号的单词，即从语音信号中提取语音特征并与储存的语音特征比较。因此结构识别单元通过单词识别单元识别这样检查单词序列，即该单词序列是否包括包含指令的结构。

[0022] 该装置有利地具有前置于讲话者识别单元和 / 或指令识别单元的回波补偿单元。该回波补偿单元具有至少一个用于单声道 -、立体声 - 和 / 或多声道 - 扬声器信号的输入端。由此可以补偿扬声器信号对语音信号的影响，其方法是计算或估算通过扬声器信号产生的分信号并将其从语音信号中减去。

[0023] 此外，回波补偿单元可以包括用于补偿由其他人产生的语音分量的子单元。该子单元有利地具有至少一个附加的麦克风输入端。该子单元可以设计用于补偿通过与附加的麦克风输入端相连的其它麦克风接收的、其他人的语音分量的影响。也就是说，该子单元设计用于根据其它麦克风的输出信号，估算或计算其他讲话者的语音分量对语音信号的影响并通过从语音信号中减去来补偿该影响。可替换地，或除了使用其它麦克风以补偿其他讲话者的语音分量之外，为了该目的该子单元还可以设计用于过滤与声音基础频率同步的语音信号。

[0024] 可替换地或附加地在下面的情况下，即用于语音信号的输入端设计用于多声道 - 语音信号，或存在多个麦克风输入端，该子单元则 可以设计用于根据之前确定的、估算的或给定的、讲话者的地点来实现不同的语音声道的在时间上的联系作用。也就是说，如果使用了多个麦克风，那么从语音信号中减去所有不是来自预期地点的讲话者的语音信号分量。

[0025] 回波补偿单元设计用于将相应的变得清晰的语音信号传送至讲话者识别单元和 / 或语音识别单元。

[0026] 此外，该装置有利地包括尤其可以布置在回波补偿单元之后的噪声抑制单元。

[0027] 噪声抑制单元设计用于抑制不变的和也有利的是似稳的（也就是说在时间上缓慢变化的）噪声分量。

[0028] 有利地，噪声抑制单元尤其这样设计用于通过频率特性适应来参数化地匹配语音信号的噪声行为特性，即语音信号的噪声属性匹配于集成在语音识别单元中的语音模型的噪声属性。

[0029] 通过噪声属性的这样的匹配确保了不会通过噪声抑制而造成经过噪声抑制处理的语音信号与语音识别单元的语音模型的不兼容性。

[0030] 有利地，讲话者识别单元和 / 或语音识别单元设计用于使被识别的讲话者向决策单元的输出与通过语音识别单元的被识别的指令的输出同步。通过该同步确保了，结果行为的激活可靠地取决于，讲话者和指令是否在语音信号的相同部段中被识别。

[0031] 有利地，在讲话者识别单元中集成了另外的单元，该单元设计用于讲话者适应 (Sprecheradaption)，也就是说用于不断确定细化的取决于讲话者的特征以及用于把该细化的取决于讲话者的特征 存储在对应于当前的讲话者的讲话者模型中。有利地，通过该另外的单元可以把取决于时间的变化作为属性存储在讲话者模型中。

[0032] 尤其有利地，语音识别单元也包括用于讲话者适应的附加单元。由此可以通过语音识别单元与当前的讲话者的匹配而显著地提高指令识别的可靠性。

[0033] 有利地，并联于讲话者识别单元和语音识别单元布置了用于根据语音信号检测驾

驶员的状态的驾驶员状态检测单元，该驾驶员状态检测单元设计用于根据语音信号来检测驾驶员的状态并发送至决策单元，其中驾驶员状态的检测有利地展示了至少关于行驶能力（例如疲劳、吸毒和 / 或超载）或情绪状况（例如激动、愤怒、精力集中、放松和 / 或平静）的信息。

[0034] 有利地，该装置具有至少一个用于存储使用者模型和 / 或至少一个讲话者模型的存储装置。该存储装置可设计为中央存储单元，其中讲话者识别单元、语音识别单元和决策单元与该中央存储单元连接。可替换地，该装置具有多个存储装置，其中至少一个存储装置分别集成在讲话者识别单元、语音识别单元和 / 或决策单元中。

[0035] 有利地，至少一个存储装置具有至少一个用于输入或输出数据的端口，从而能够传输讲话者和 / 或使用者模型。例如由于购买新车而更换车辆时，这是尤其有利的，这是因为因此可以避免完全重新设定讲话者模型和使用者模型。

[0036] 在一个特别有利的实施例中，该装置这样设计，即如果通过决策单元来执行结果行为，则也考虑输入的语音信号，以及通过包含在语音信号中的指令至少部分地中断由在前的指令触发的结果行为的执行。部分的中断可以在，即在继续执行通过在前的指令触发的控制事件期间中断语音输出。因此，讲话者不必完全等待语音输出。由此可以明显加速语音对话，尤其是当已经熟悉该装置的讲话者已经知道了通过语音输出通知的选项时。此外有利的是，当使用者期望的选项已被读出后，可以中断列举可能的选项。

[0037] 此外有利的是，用于激活语音对话和 / 或进行语音对话的系统具有至少一个扬声器、至少一个麦克风以及根据本发明的用于激活语音对话或进行语音对话的装置。有利地，至少一个麦克风具有自动的放大适配性 (Verstaerkungsanpassung)。

[0038] 该系统可以用于通过至少一个扬声器输出单声道 -、立体声 - 或多声道 - 音频信号，以及用于通过至少一个麦克风采集单声道 -、立体声 - 或多声道 - 语音信号。此外取代或除了单个麦克风之外，还可以包括麦克风阵列。

[0039] 如果存在多个麦克风，那么这些麦克风有利地这样布置或可这样布置，即麦克风的最大灵敏性的区域相对于其方向特性在授权的讲话者的推测的或确定的停留区域中进行覆盖，以及用于处理麦克风信号的后续单元通过对信号的时间延迟的合计以及其它信号处理方法分离出授权的讲话者的信号，每个授权的讲话者的信号在其声道中提高，并且所有其它讲话者的信号和干扰信号减小（所谓的“beam forming（波束成形）”）。

[0040] 有利地，该系统包括用于使麦克风自动对准授权的讲话者的跟踪装置。该跟踪装置可以设计用于分析通过麦克风接收的信号的渡越时间延迟从而获得授权的讲话者的位置，并且然后使麦克风对准该位置。

附图说明

[0041] 下面根据图 1 至 4 说明本发明的实施例。附图示出：

[0042] 图 1 示出根据本发明的用于激活语音对话的系统的第一实施例；

[0043] 图 2 示出根据本发明的用于进行语音对话的系统的第一实施例；

[0044] 图 3 示出根据本发明的用于激活语音对话的系统的第二实施例；

[0045] 图 4 示出根据本发明的用于进行语音对话的系统的第二实施例。

具体实施方式

[0046] 图 1 示出用于激活语音对话的系统 1。该系统分析通过麦克风 2 接收的语音信号 8。该语音信号包括：通过使用者进行的声学的语音输入 3 以及多个干扰信号，即车辆内部空间 4 中的其他人的语音分量、车辆周围环境 5 的噪声、通过扬声器 6 的来自无线电收发设备 28 的音乐和 / 或来自语音对话输出 27 的语音输出信号、以及语音输入和干扰信号的声波反射 7。语音信号 8 被输入到回波补偿单元 9 中。回波补偿单元 9 具有用于扬声器信号 10 的附加的输入端。回波补偿单元利用扬声器信号 10 计算干扰信号 6 对语音信号 8 的影响。此外，回波补偿单元计算声波反射 7 的通过干扰信号引起的分量以及其对语音信号 8 的影响。由回波补偿单元补偿了因此是已知的影响，并且从回波补偿单元 9 发出经回波补偿的语音信号 11。

[0047] 经回波补偿的语音信号 11 在输入端方面输入噪声抑制单元 12 中并经历噪声补偿。由此减去信号 11 的静态背景。此外，补偿似稳的、也就是说缓慢变化的噪声分量。此外，通过频率特性适应来 参数化地匹配信号 11 的噪声行为特性。由此，使信号 11 的噪声属性匹配关键词识别单元的语音模型的噪声属性。

[0048] 这样地经噪声补偿的信号 15 经过第一输出端 13 和第二输出端 16 输出。

[0049] 第一输出端 13 与讲话者识别单元 14 连接。讲话者识别单元 14 设计用于从信号 15 中提取讲话者特征。该讲话者识别单元把提取的讲话者特征依次与在存储装置 17 里的讲话者模型中存储的特征这样比较，即分别依次调出讲话者模型并把存储在其中的特征与提取的特征比较。对此重复操作，直到把提取的特征与所有存储的讲话者模型比较。这样确定，提取的特征是否与讲话者模型之一的特征一致。

[0050] 如果确定了这种一致，那么讲话者识别单元通过输出端 23 把关于识别的讲话者的信息传输至语音对话激活单元 29。

[0051] 此外，另一个用于讲话者适应的单元 20 集成在讲话者识别单元 14 中，该另外的单元设计用于持续细化取决于讲话者的特征。该另外的单元将细化的、取决于讲话者的特征存储在属于讲话者的、存储在存储装置 17 上的讲话者模型中。

[0052] 该另外的单元 20 设计用于，确定讲话者内容特征的取决于时间的变化并将之作为属性存储在存储装置 17 上的讲话者模型中。

[0053] 关键词识别单元 24 与讲话者识别单元 14 并联地布置。该关键词识别单元设计用于识别，语音信号中是否包含或包含哪些有效的指令或有效的语音指令形式。通过输出端 25 输出相应确定的指令。

[0054] 此外，关键词识别单元包含用于讲话者适应的附加单元 18。附加单元设计用于对当前的授权的讲话者的关键词识别进行优化。为此，附加单元从语音信号中提取语音特征，并将其与存储在讲话者模型中的单个的语音特征相比较。取决于比较结果对单个的语音特征进行细化和 / 或补充。由此可以将语音特征的在时间上的变化作为属性进行检测，并存储在讲话者模型中。在识别关键词时会考虑这样确定 / 补充的单个的语音特征。因为这样用于关键词识别的语音模型可以匹配当前的讲话者，所以可以显著提高关键词识别的可靠性。

[0055] 在讲话者识别单元 14 和关键词识别单元 24 之后布置了语音对话激活单元 29。当由关键词识别单元 24 识别了相应的指令并且同时由讲话者识别单元识别了有权激活的讲

话者时,语音对话激活单元设计用于输出一个或多个事件 26。

[0056] 语音对话输出 27 可激活事件中任一个,并且未示出的语音对话系统可激活另一事件。

[0057] 图 2 示出用于进行语音对话的系统 30 的实施例。该系统取代关键词识别单元 24 而具有更复杂的语音识别单元 31,以及取代语音对话激活单元 29 而具有语音对话单元 32。

[0058] 语音识别单元 31 与第一实施例中的关键词识别单元 24 的区别主要在于,不仅将一些关键词、而且也将多个不同的指令识别为有效,以及不仅识别单词自身、而且也识别了由单词组成的结构。为此,语音识别单元 31 划分为用于识别单词的单词识别单元 19 以及用于识别通过单词组成的结构的结构识别单元 21。在用于语音对话激活的系统中这种划分也是有利的。用于讲话者适应的附加单元 18 集成在单词识别单元 19 中,并在那里起到与在第一实施例中相同的作用。

[0059] 语音对话单元 32 也具有相应更大的功能范围。语音对话单元设计用于进行与授权的讲话者进行语音对话,其中语音对话流程适应性地由授权的讲话者控制。例如,使用者可以通过选择菜单的不同选项来控制这种语音对话流程。例如在使用者选择选项之后,他可以到达具有其它选项的其它菜单。除了进行语音对话之外,语音对话单元 32 还设计用于,根据授权的讲话者的指令来输出用于车辆功能的不同的控制信号。

[0060] 除了与第一实施例的上述区别之外,在图 2 中示出的实施例的不同之处还在于,代替一个扬声器 6 布置了多个扬声器 6。然而,如在第一实施例中那样,这些扬声器也仅设计用于输出单声道信号。

[0061] 与前述实施例的区别还在于,噪声抑制单元具有第三输出端 22。该输出端连接至驾驶员状态识别单元 39。该驾驶员状态识别单元用于识别驾驶员的状态,如嘶哑、疲劳、醉酒、吸毒或愤怒。

[0062] 将识别出的驾驶员状态传输到语音对话单元。语音对话单元设计用于在关于结果行为的决策方面对驾驶员状态加以考虑。

[0063] 与图 1 的实施例的区别还在于,该系统设计用于干涉,也就是说如果当还在基于前述指令进行语音输出的同时识别出指令,则中止该语音输出。

[0064] 图 3 示出了语音对话激活系统的第二实施例。该实施例与图 1 所示的实例的区别在于,代替扬声器 6 而设有扬声器组合 33。扬声器组合设计用于输出单声道 -、立体声 - 和多声道信号 34。相应地,回波补偿单元设计用于补偿多声道信号 34。此外,正如在图 2 中所示的实施例那样,语音对话激活系统包括驾驶员状态识别单元 39。

[0065] 图 4 示出用于进行语音对话的系统的第二实施例。在该实施例中代替单个麦克风 2 安装了麦克风阵列 35 以及一些另外的麦克风 36。麦克风阵列以及另外的麦克风这样设计,即它们具有方向特性,也就是说它们优选地接收来自特定的空间区域的声波。通过麦克风阵列 35 接收了多声道语音信号 37,该多声道语音信号取代语音信号 8 被传输至回波补偿单元。

[0066] 此外,在此也使用扬声器组合用于输出单声道 -、立体声 - 和多声道信号。

[0067] 麦克风阵列 35 可以包括未示出的麦克风跟踪装置,该麦克风跟踪装置使阵列的麦克风自动跟踪授权的讲话者。为此,麦克风跟踪装置分析语音信号相对于阵列的不同麦克风的渡越时间差并因此确定授权的讲话者的地点。

[0068] 通过麦克风阵列接收的语音信号在输入到回波补偿单元中之前,在连接在麦克风阵列之后的、用于处理麦克风信号的单元中通过信号的时间延迟的相加得出总和来进一步处理。由此将授权的讲话者的信号分离并减小所有其它的讲话者信号和干扰信号。

[0069] 另外的麦克风 36 中的任一个对齐每个其它的车辆位置。回波补偿单元包括子单元 38, 该子单元设计用于根据另外的麦克风 36 的信号补偿在车辆内部空间中的其他人对信号 37 的影响。也就是说, 该子单元根据另外的麦克风 36 的信号计算其他人对信号 37 的可能的影响, 并相应地减去该信号分量。

[0070] 此外, 回波补偿单元分析多声道语音信号 37 的不同声道的渡越时间差, 并且去除信号 37 的所有分量, 这些分量相当于其渡越时间差并不从授权的讲话者的地点出发。

[0071] 附加地, 在图 4 中所示的实施例的不同之处在于, 存储装置 17 与用于连接 USB 端口、读卡器或光驱的输出端 40 连接。由此, 能够实现使用者模型的输入和 / 或输出。

[0072] 当然, 扬声器和麦克风的选择不取决于是否存在用于进行语音对话或激活语音对话的系统, 也就是说, 根据图 2 或 4 的具有扬声器 / 麦克风布置的、用于激活语音对话的系统, 或根据图 1 或 3 的具有扬声器 / 麦克风布置的、用于进行语音对话的系统也都是有利的。

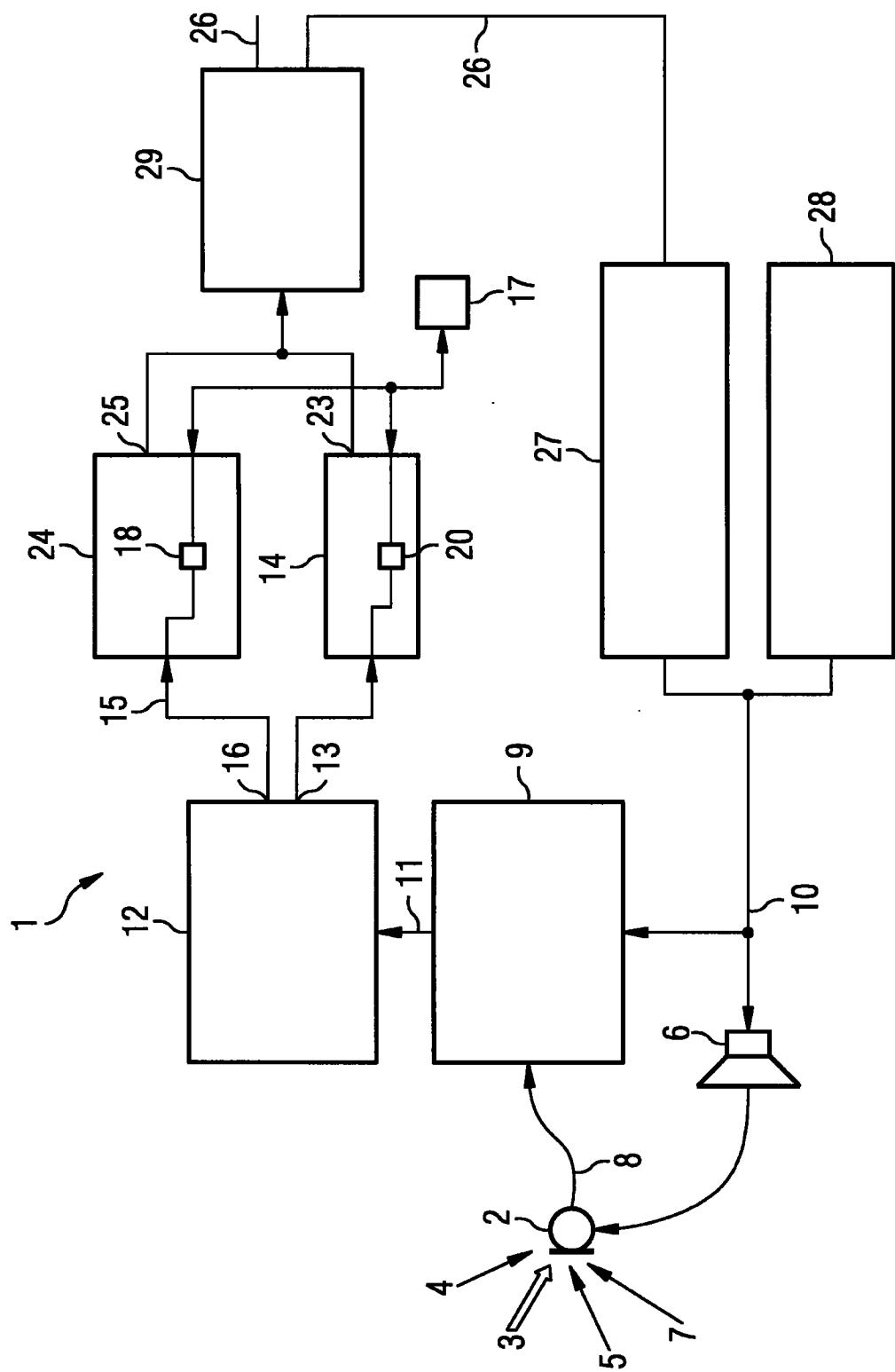


图 1

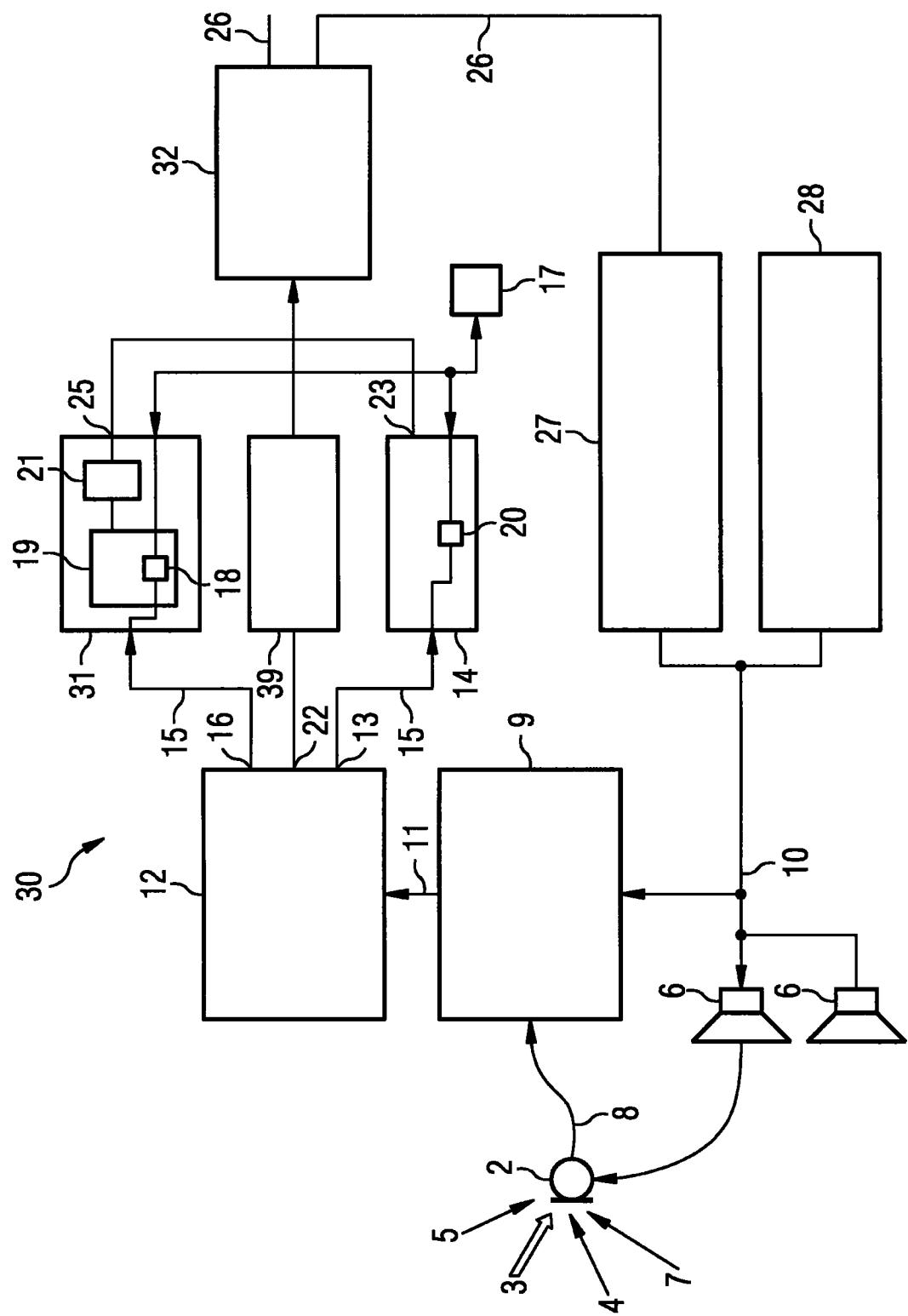


图 2

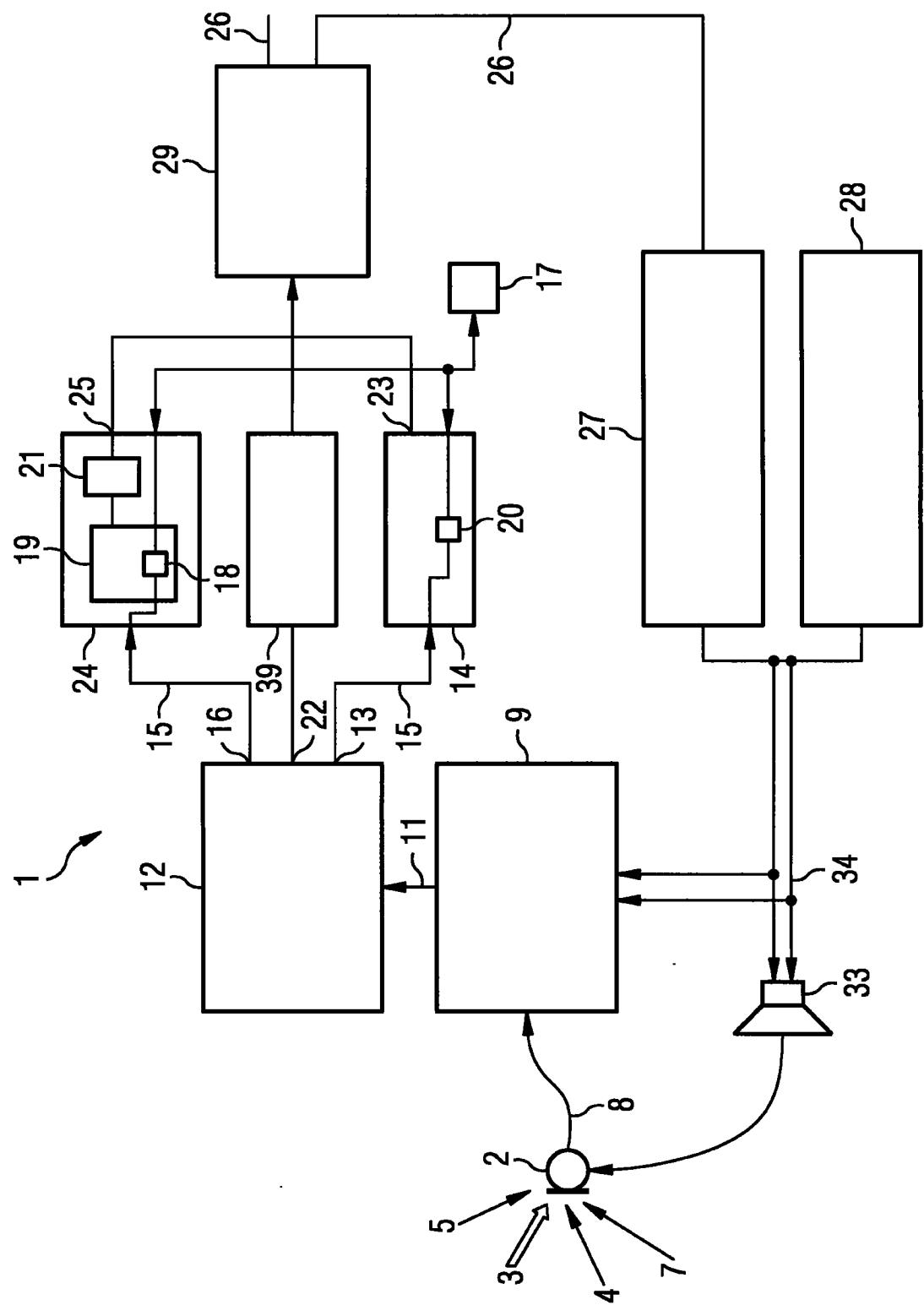


图 3

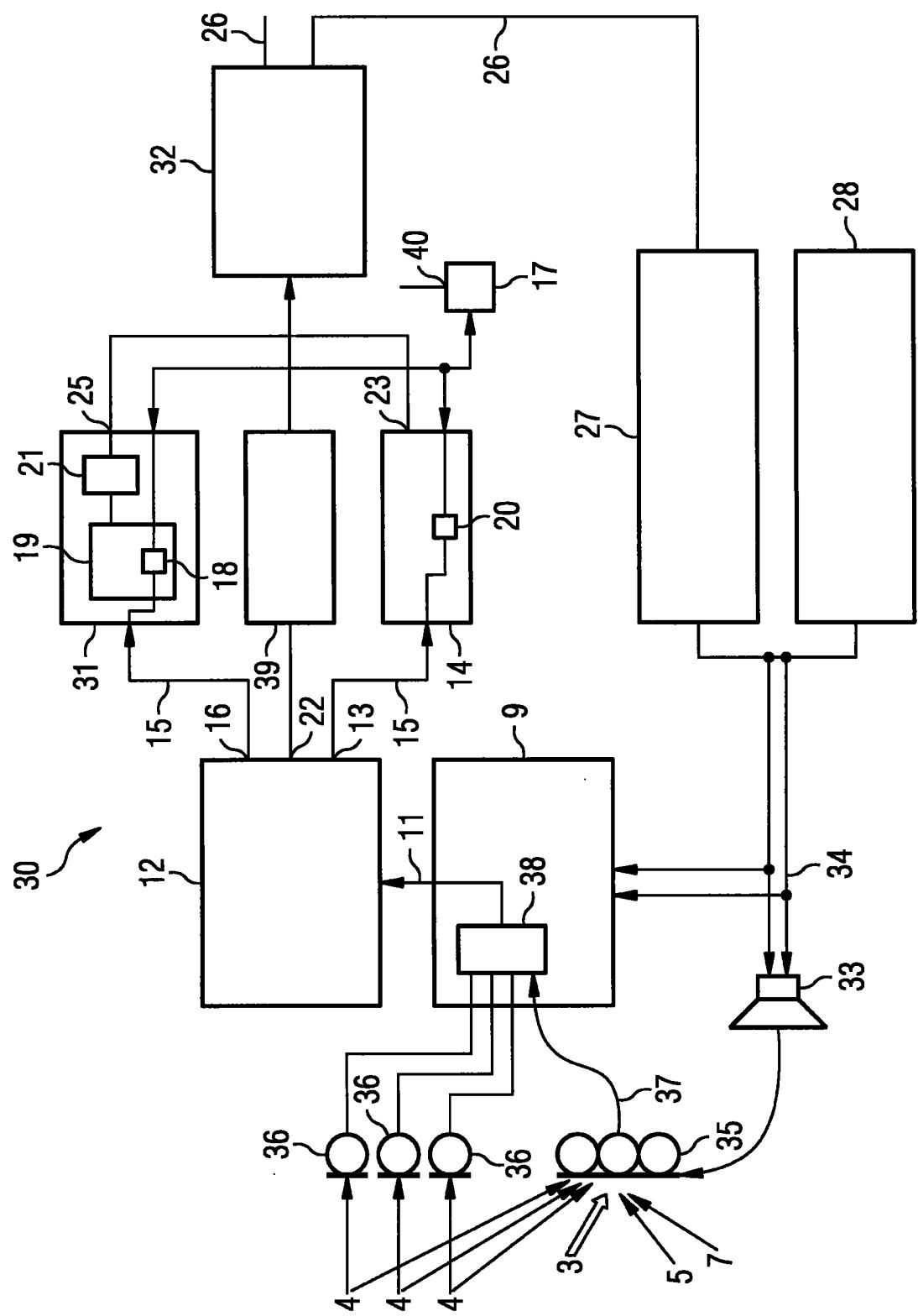


图 4