



(12)发明专利申请

(10)申请公布号 CN 106537492 A

(43)申请公布日 2017.03.22

(21)申请号 201580038873.4

(74)专利代理机构 北京市中咨律师事务所
11247

(22)申请日 2015.10.22

代理人 吴鹏 牛晓玲

(30)优先权数据

102014017384.4 2014.11.24 DE

(51)Int.Cl.

G10L 15/22(2006.01)

(85)PCT国际申请进入国家阶段日

2017.01.16

(86)PCT国际申请的申请数据

PCT/EP2015/002098 2015.10.22

(87)PCT国际申请的公布数据

W02016/082904 DE 2016.06.02

(71)申请人 奥迪股份公司

地址 德国因戈尔施塔特

(72)发明人 D·恩格尔哈特

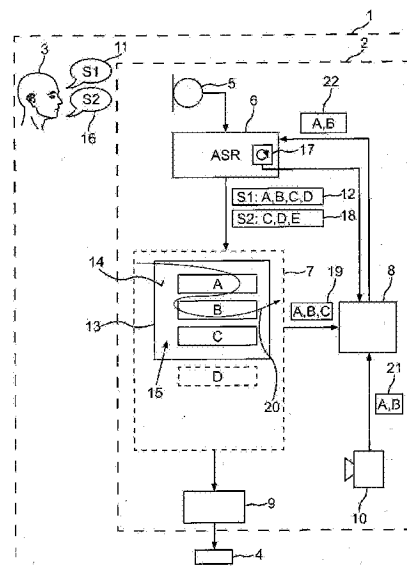
权利要求书1页 说明书6页 附图1页

(54)发明名称

具有用于语音识别的校正策略的机动车操作装置

(57)摘要

本发明涉及一种用于运行机动车操作装置(2)的方法,其中,接收使用者(3)的第一语音输入(11),借助于语音识别系统针对第一语音输入(11)确定至少一个识别结果(A-D),所述至少一个识别结果(A-D)被作为结果列表(12)输出给机动车(1)的输出设备(7),接收使用者(3)的第二语音输入(16)。本发明的目的在于避免重复输出错误的识别结果。在第二语音输入(16)中识别出在内容方面对第一语音输入(11)的输入重复,这表明使用者(3)的校正意图。之后确定结果列表(12)的要排除掉的部分(22),并且在针对第二语音输入(16)确定识别结果(C-E)时排除掉作为可能的识别结果的、要被排除掉的部分(22)。



1. 一种用于运行机动车操作装置 (2) 的方法, 其中
 - 接收使用者 (3) 的第一语音输入 (11),
 - 借助于语音识别系统针对第一语音输入 (11) 确定至少一个识别结果 (A-D),
 - 所述至少一个识别结果 (A-D) 被作为结果列表 (12) 输出给机动车 (1) 的输出设备 (7),
 - 接收使用者 (3) 的第二语音输入 (16),其特征在于,
 - 在第二语音输入 (16) 中识别出在内容方面对第一语音输入 (11) 的输入重复,
 - 确定结果列表 (12) 的要排除掉的部分 (22),
 - 在针对第二语音输入 (16) 确定识别结果 (C-E) 时排除掉作为可能的识别结果的、要被排除掉的部分 (22)。
2. 根据权利要求1所述的方法, 其特征在于, 基于如下标准中的至少一个识别出内容方面的输入重复:
 - 使用者 (3) 执行预定的、规定用于输入校正的操作动作;
 - 在第二语音输入 (16) 中识别出预定的校正关键词;
 - 第一语音输入 (11) 和第二语音输入 (16) 具有预定的韵律区别;
 - 第一语音输入 (11) 和第二语音输入 (16) 在预定的最少部分方面相互一致。
3. 根据上述权利要求之一所述的方法, 其特征在于, 为排除掉所要排除的部分 (22), 改变语音识别系统的语法。
4. 根据上述权利要求之一所述的方法, 其特征在于, 为排除所要排除的部分 (22), 借助于语音识别系统针对第二语音输入 (16) 确定至少一个标准识别结果, 并从标准识别结果中过滤掉要排除的部分 (22)。
5. 根据上述权利要求之一所述的方法, 其特征在于, 确定结果列表 (12) 中的已经由输出设备 (7) 成功输出给使用者 (3) 的部分作为要排除的部分 (22)。
6. 根据权利要求5所述的方法, 其特征在于, 为确定成功输出的部分 (19), 评估输出设备 (7) 的滚动过程。
7. 根据权利要求5或6所述的方法, 其特征在于, 借助于眼球追踪器 (10) 检测使用者 (3) 的眼球运动 (20), 并且从结果列表 (12) 中仅选出那些根据所检测到的眼球运动 (20) 被使用者 (3) 观察到的部分 (A、B) 确定为成功输出的部分 (21)。
8. 根据权利要求7所述的方法, 其特征在于, 仅当使用者 (3) 持续预定的最小时长观看所述识别结果的部分 (A、B), 才从结果列表 (12) 中选出使用者 (3) 所观看到的部分。
9. 一种用于机动车 (1) 的操作装置 (2), 该操作装置具有:
 - 用于检测使用者 (3) 的语音输入 (11、16) 的语音检测设备 (5),
 - 用于输出结果列表 (12、16) 的输出设备 (7), 该结果列表具有针对所检测到的语音输入 (11、16) 确定的至少一个识别结果 (A-E),其特征在于, 一种处理器设备, 其被布置用于执行根据上述权利要求之一所述的方法。
10. 一种机动车 (1), 该机动车具有根据权利要求9所述的操作装置 (2)。

具有用于语音识别的校正策略的机动车操作装置

技术领域

[0001] 本发明涉及一种用于运行机动车操作装置的方法,以便针对使用者的语音输入借助于语音识别系统确定至少一个识别结果,并且将该识别结果以结果列表形式输出。如果使用者之后进行第二语音输入,则检查使用者是否由于其未在结果列表中发现所希望的识别结果而在内容方面重复或校正其第一语音输入。本发明还涉及一种用于执行所述方法的操作装置以及一种具有操作装置的机动车。

背景技术

[0002] 在机动车中,操作装置可以包括语音识别系统,和/或通过无线电连接与固定的语音识别系统耦合。如果通过语音输入了语音内容,则向使用者显示所识别的形式为语音的和/或图像的反馈的识别结果。如果针对语音输入确定了多种识别结果,则可以将这些识别结果作为选项加以显示以供在结果列表中选择。

[0003] 如果使用者在结果列表中未发现所期望的或所想要的识别结果,也就是由其所说出的语音内容,则使用者通常再次说出该语音内容。在此,使用者可以使用与第一语音输入时相同的表达或另一表达方式。在评估该第二语音输入时,在此在操作装置方面并不清楚这是校正还是新的语音内容。因此,第二语音输入通常借助于语音识别系统如同第一语音输入一样地被处理。在由此确定的结果列表中因而可能再次含有相同的识别结果,并将其冗余地再次显示给使用者。因此可能在重新的语音输入后再次识别出在第一步骤中被使用者认定为错误的结果到并显示给使用者。

[0004] 从文献US 2007/0033025 A1中已知了一种借助于识别系统对使用者所说出的语句确定多个结果语句。然后向使用者提出测试问题:在使用者所说出的语句中是否含有某一特定词语。如果使用者肯定回答,则删除所有未含有该词语的那些结果语句。如果使用者否定回答该测试问题,则删除所有含有该词语的结果语句。

[0005] 从文献US 2003/0216912 A1中已知一种语音识别系统,对于使用者两次说出相同语句的情况,在这两个识别结果中确定一致的部分和彼此不同的部分。如果使用者在第二次说出语句的情况下特别强调了某一特定词语,则认定该词语在第一识别结果中被识别错误。然后由一致的部分和重新识别的不同的部分形成识别结果以供进一步处理。

[0006] 在文献EP 1 215 659 A1中针对使用者的语音输入确定第一识别结果,并询问使用者:该结果是否正确。如果使用者加以否定,则以语音识别系统的改变的参数再次处理该同一语音输入。

发明内容

[0007] 本发明的目的在于,避免重复地输出错误识别结果。

[0008] 该目的通过独立权利要求主题实现。本发明有利的改进方案由从属权利要求的特征得出。

[0009] 根据本发明提供了一种用于运行机动车操作装置的方法。以所说明的方式接收使

用者的第一语音输入。该第一语音输入包含语音内容,例如语句或词语。利用语音识别系统——该语音识别系统可以设置于车辆内部(on-board(车载的))或车辆外部——针对第一语音输入确定至少一个识别结果。每个识别结果都因而可以包括一所识别到的句子或一所识别到的词语。由此能够产生多个可供选择的识别结果,例如已知的N佳列表(N佳识别结果)。

[0010] 所述至少一个识别结果被作为结果列表输出给机动车的输出设备。通过该输出设备能够向使用者输出语音或图型反馈,即结果列表。在该方法中假定,之后接收到使用者的第二语音输入。

[0011] 根据本发明的方法的特征在于,在第二语音输入中识别出在内容方面对第一语音输入的输入重复。要注意的是,识别出内容方面的输入重复。该输入重复可以通过逐字地重复第一语音输入或通过第一语音输入的替代的语音表达来实现。因此检查:使用者是否想要输入新的文字或语音内容,或者使用者是否由于不能够在输出设备的输出内容中找到想要的识别结果而对第一语音输入的语音内容进行重复。

[0012] 如果识别出是内容方面的输入重复,则确定结果列表的要排除掉的部分。换句话说,确定所有针对第一语音输入所确定的识别结果或确定识别结果的一部分,其结合对第二语音输入的语音识别而不应再次被显示为识别结果。在针对第二语音输入确定识别结果时,相应地排除掉作为可能的识别结果的、所述要被排除掉的部分。

[0013] 通过本发明得到的优点是,不再向使用者输出要排除掉的部分,从而由此以更高可能性不再二次输出错误的识别结果。

[0014] 在根据本发明的方法中根据如何识别出所说明的在内容方面的输入重复得到不同的改进方案。

[0015] 本发明的一改进方案对此规定,根据使用者执行预定的、规定用于输入校正的操作动作,识别内容方面的输入重复。例如为此可以设有操作按键或信号词语。

[0016] 在另一改进方案中,根据在第二语音输入本身中识别出预定的校正关键词,识别内容方面的输入重复。例如可以检查在第二语音输入中是否含有下列词语:“不,我是说……”。

[0017] 在另一改进方案中,根据第一语音输入和第二语音输入具有预定的韵律区别,识别内容方面的输入重复。换句话说,在此检测和分类语调。对于语音输入中的韵律的识别和分类可以采用已知的方法。典型的分类可以是:平静的、烦躁的、激怒的。通过在第一语音输入和第二语音输入之间识别出预定的韵律区别,例如从平静的转换至激怒的,可以同样推断为内容方面的输入重复。

[0018] 在另一改进方案中,根据第一语音输入和第二语音输入针对预定的最少部分相互一致,识别出内容方面的输入重复。在此同样借助于语音识别系统首先将第二语音输入转换为至少一个语音结果,然后将针对第一语音输入的识别结果与针对第二语音输入的识别结果相比较,例如词语层面或字母层面的比较。

[0019] 根据本发明的方法的其他改进方案涉及的是在针对第二语音输入确定识别结果时排除掉要排除的部分的方法。

[0020] 对此一改进方案是,为排除所要排除的部分,改变语音识别系统的词汇或语法。也就是例如可以针对第二识别舍弃由第一识别得到的识别假设。换句话说,动态地匹配语音

识别系统的语法。由此得到的优点是，缩短了识别过程的过程时长。

[0021] 另一改进方案提出，为排除所要排除的部分，首先借助于语音识别系统针对第二语音输入确定至少一个标准识别结果。标准识别结果是指，使用与识别第一语音输入所使用的语法相比未改变的语法。之后从标准识别结果中过滤掉要排除的部分。换句话说，例如从结果列表中删除所有曾输出的那些识别结果。由此得到的优点是，不必操作语音识别系统，这尤其在位于车辆外部的语音识别系统的情况下是有利的。

[0022] 下面描述根据本发明的方法的改进方案，其涉及如何确定要排除的部分。

[0023] 一改进方案提出，结果列表中的已经由输出设备成功输出给使用者的部分确定为要排除的部分。因而只要排除掉那些实际也已经例如显示给使用者的部分。

[0024] 对此一改进方案涉及通过图像的显示装置输出结果列表，该显示装置的显示面小于结果列表。在该改进方案中，输出设备的滚动过程被评估用于确定成功输出的部分。即检查结果列表中的哪些部分被实际显示或示出在显示面上。对滚动过程的评估、例如对坐标或滚动信号的评估是成功输出的可靠指示。

[0025] 根据另一改进方案作如下区分：结果列表的识别结果仅是被输出、还是也已经被使用者感知到。为此，该改进方案规定，利用眼球追踪系统或简称眼球追踪器（眼睛追踪器）检测使用者的眼球运动。为确定成功输出的部分，从结果列表中仅选出那些根据所探测的眼球运动被使用者观察到的部分。例如，因而可以为此外推出一视线方向矢量，并检查，该外推的视线方向矢量在显示面上出现在何处。适用于此的眼球追踪器能够由现有技术获得。

[0026] 另一改进方案提出，额外地，只有当使用者持续预定的最小时间长度观察该识别结果，才从结果列表中选出使用者所观察到的部分。由此得到的优点是，能够在不经心地粗阅结果与实际上通过阅读获知内容之间作出区分。因此评估阅读时的速度。

[0027] 本发明还包括用于机动车的操作装置。该操作装置包括用于检测使用者的语音输入的语音探测设备，也就是例如麦克风或包括多个麦克风的布置结构。此外设有用于输出结果列表的输出设备，该结果列表具有针对所检测到的语音输入确定的至少一个识别结果。输出设备可以被布置用于通过合成声音进行语音输出和/或用于图像输出。对于图像输出的情况，该输出设备可以例如具有显示屏。根据本发明的操作装置的特征在于一种处理器设备，其被布置用于执行根据本发明的方法的实施形式。

[0028] 最后，本发明还包括一种机动车，该机动车具有根据本发明的操作装置的实施形式。根据本发明的机动车优选被构造为汽车、尤其是轿车。

附图说明

[0029] 下面对本发明的实施例加以说明。对此，唯一的视图示出了根据本发明的机动车的实施形式的示意图。

具体实施方式

[0030] 在下面所述的实施例中所涉及的是本发明的优选实施形式。在该实施例中，该实施形式的所说明的组成部分分别构成本发明的各个、可视为彼此独立的特征，这些特征也分别彼此独立地改进了本发明并因而也单独地或以不同于所示组合的其它组合方式被视

作是本发明的组成部分。此外,所述实施形式也能够由本发明的所述的特征中的其他特征加以补充。

[0031] 附图示意性示出了机动车1,其在此可以是汽车、尤其是轿车。该机动车1可以具有操作装置2,借助于该操作装置能够使机动车1的使用者3操作该机动车的一个或多个仪器4。这些仪器可以例如是下面的一个或多个:媒体播放设备、通信系统、导航系统、电话。

[0032] 通过操作装置2能够实现对设备2的语音操作。对此,该操作装置2可以包括语音检测设备5、语音识别设备6、输出设备7、控制设备8和信号产生设备9。还可以规定一种实施形式,其中设有现有技术所公开的眼球追踪器10。利用该眼球追踪器10可以例如确定和/或追踪使用者3的眼球的视向方向。

[0033] 为实现语音操作,使用者3可以进行语音输入11。附图中象征性地示出作为语音输入11的语音内容或语音表述S1。在所示的该示例中,使用者3想要操作导航设备,并启动向慕尼黑的路线规划。该语音表达S1可以例如是:“导航至慕尼黑!”。语音检测设备可以检测到该语音输入11的语音声响。为此,探测设备5例如可以包括一个或多个麦克风。通过语音检测设备5将语音声响转换为电的语音信号和/或数字的语音信号,并且传送给语音识别设备6(ASR-Automatic Speech Recognition-自动语音识别)。

[0034] 语音识别设备6可以具有语音识别系统或例如通过(未示出的)移动无线电连接与固定的语音识别系统、例如因特网的语音识别系统耦合。通过该语音识别设备6能够以已知的方式对语音输入11译码,也就是说,将语音表达S1识别为文本。例如可以基于隐马尔可夫模型(HMM)进行语音识别。

[0035] 在语音识别时可以针对一个语音输入11产生一个或多个识别结果。在该示例中假定,带有语音表达S1的语音输入11产生多个语音识别结果A、B、C、D。例如,所述语音识别结果可以在所识别出的目的地方面有所不同,因而例如该识别结果A可能是:“导航至波恩!”,语音识别结果B是:“导航至柏林!”,语音识别结果C是:“导航至慕尼黑!”,语音识别结果D是:“导航至萨尔斯堡!”。

[0036] 识别结果A、B、C、D可以由语音识别设备6汇总成结果列表12,其能够被输出到输出设备7。该输出设备7可以向使用者3输出识别结果A、B、C、D,由此该使用者选出正确的识别结果。为此,所述输出设备7可以例如将结果列表7输出为语音信号,也就是以声音方式输出,或者输出设备具有显示设备13、例如显示屏。在显示设备13的显示面14上可以显示具有结果列表12的识别结果A、B、C、D的选择菜单15。显示面14可以如此小,使得结果列表12不能够被同时完全显示。图中的示例中表明,这时还未示出识别结果D。使用者3可以例如在显示面14上通过滚动条移动选择菜单15,由此在显示面14上显示其他识别结果。

[0037] 使用者3这时可以选出他所意指的结果。为此,显示设备13例如可以被构造为触敏式的,使得使用者3能够例如在选择菜单15中点击正确的识别结果。所选出的识别结果可以被传递给信号产生设备9。该信号产生设备9可以根据所选出的识别结果然后以已知的方式产生控制命令,该控制命令被输出给仪器4的组中的匹配的仪器。在该示例中,于是在导航设备中引发或触发路线规划或给出路线参数。

[0038] 在机动车1中使得使用者能够进行识别校正。如果使用者3不能在结果列表12中发现所期望的识别结果,则使用者可以进行或操作另外的语音输入16。例如,可以由使用者3说出语音内容或语音表达S2作为语音输入1。在此,使用者3可以简单地重复语音表达S1,也

就是再次说出：“导航至慕尼黑！”。但是他也可以说出另外的语音表达，例如：“不，我是说慕尼黑”。

[0039] 语音输入16可以以所述的方式通过语音检测设备5检测，到并通过语音识别设备6加以处理。语音识别设备6可以具有重复识别系统17，其可以被布置用于识别在内容方面对第一语音输入11的输入重复。例如，在解码或识别语音输入16后可以通过重复识别系统17识别：在语音输入16中是否包含了信号词或关键词。在此情况下因而例如可以将词“不”解读为校正意图。如果使用者3重复了其输入，也就是例如再次说出：“导航至慕尼黑！”，则通过重复识别系统17例如可以凭借识别结果识别出：多个词、例如60%以上或70%以上的部分与针对第一语音输入11的识别结果一致。也可以例如规定，使用者3在说出语音输入16期间必须按下校正键，这于是同样能够被重复识别系统17识别到。还可以例如对语音输入11和16的语调进行韵律分析和比较。如果使用者3不耐烦或发火，则这可以通过韵律的改变、也就是语调走向的变化识别出。韵律识别本身已由现有技术公开。重复识别系统17可以例如是操作装置2的处理器设备的程序模块。

[0040] 重复识别系统17然后将所识别出的使用者3的校正意图例如通知给控制设备8。该控制设备8可以被布置成：为产生第二结果列表18通过语音识别设备6首先从结果列表18中排除掉那些已经由使用者检查过的识别结果，因而使用者3不会再次看到错误的识别结果。控制设备8可以例如是操作装置2的处理器设备的程序模块。

[0041] 控制装置8可以例如从显示设备7处接收到显示列表9，该显示列表给出：结果列表12中的哪些部分已经在显示面14上显示过。替代或补充于此，控制装置8可以从眼球追踪器10处例如接收到视线方向坐标，其例如能够描述视线方向轨迹20，借助于该视线方向轨迹能够确定：使用者3实际上观看了菜单15中的哪些菜单条目和/或观看菜单15的哪个菜单条目持续多长时间。相应的视线方向数据21可以由眼球追踪器10提供给控制装置8。

[0042] 在该示例中假定，控制装置8从眼球追踪器10处接收到视线方向数据21，并且这样设置语音识别设备6，使得从第一语音输入11的语音识别结果A、B、C、D中排除掉针对语音内容S2的识别所要排除的部分22。例如，可以排除掉词语波恩和柏林(A、B)。尤其是因此并不将完整的语音识别结果、而是仅将其中的单个词语确定为要排除的部分22。语音识别设备6于是可以从针对语音输入6所确定的识别结果中排除掉或过滤掉要排除的部分22。

[0043] 相应地，通过语音识别设备6在所示的示例中产生结果列表18，其能够包含针对语音表达S2的语音识别结果C、D、E。因此也就排除掉了所要排除的部分，在此为A、B。除了识别结果C、D，结果列表18因此还可以包含识别结果E：“导航至巴黎！”。

[0044] 附图中未进一步示出：由使用者3通过输出设备7这时能够例如在显示面14上示出或输出结果列表18。使用者3于是可以选出正确的识别结果，之后可以以所述的方式通过信号产生设备8输出，并由此可以控制仪器4。

[0045] 通过操作装置2因而确定使用者方面的校正，从而可以舍弃作为假设的对之前输入的已经识别的结果，并且不将其再次显示给使用者。由此得到的优点是，针对第二输入的识别能够得到改善，这是因为不必考虑来自第一输入的识别结果。此外被认定为错误的结果不被显示给使用者。该系统在使用者看来明显更智能。这由此实现：确定在使用者方面的校正，从而对第一输入的识别结果不被用于在第二步骤中的识别。因此针对第二识别舍弃了来自于第一识别的识别假设。

[0046] 整体而言该示例示出：通过本发明如何能够为通过语音的语音输入提供一种校正策略。

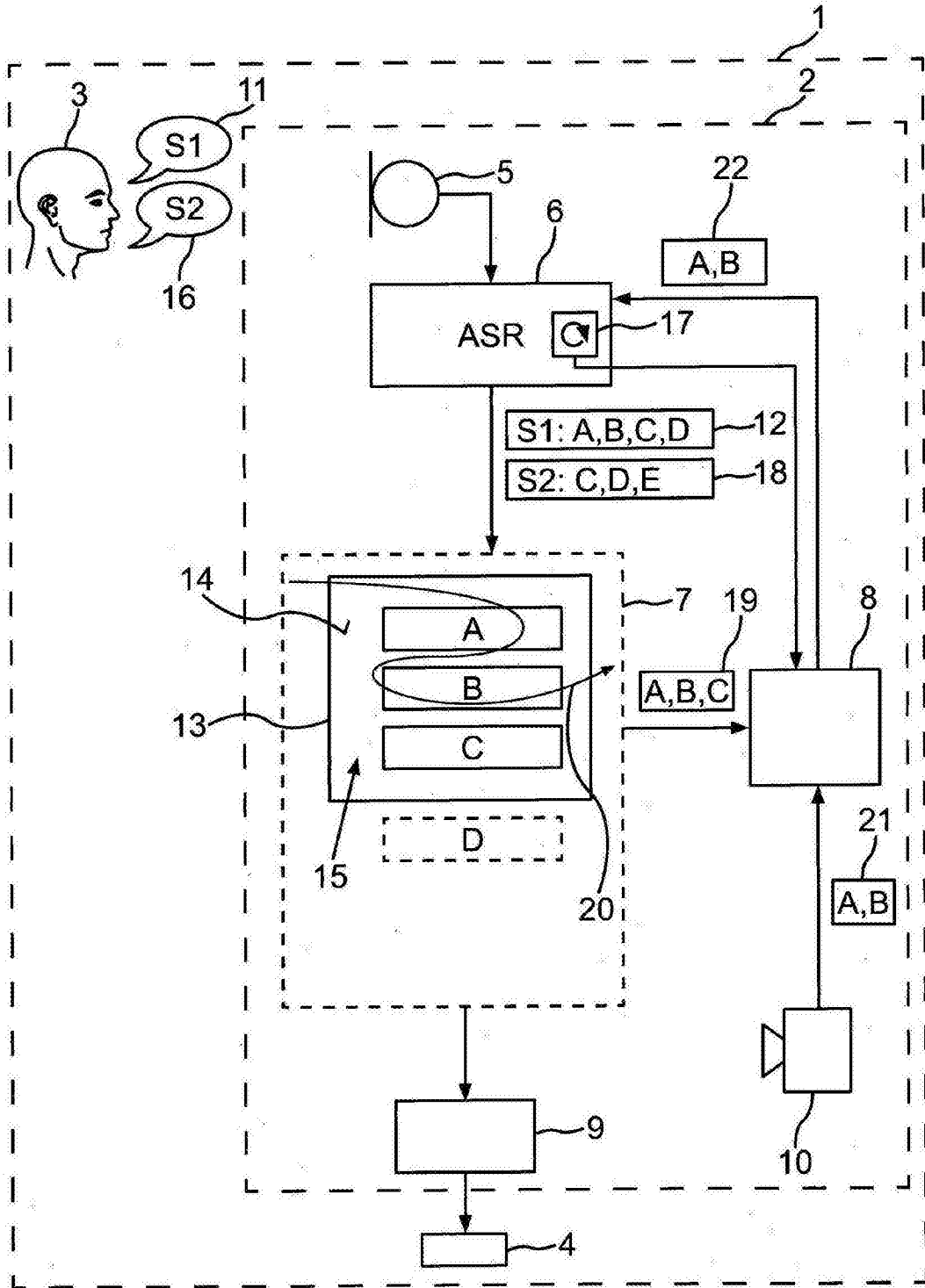


图1