



(12)发明专利申请

(10)申请公布号 CN 110600036 A

(43)申请公布日 2019.12.20

(21)申请号 201910907963.2

(22)申请日 2019.09.24

(71)申请人 随锐科技集团股份有限公司
地址 100192 北京市海淀区宝盛南路1号奥
北科技园19号楼北京随锐中心

(72)发明人 陈洪浩 冯文澜

(74)专利代理机构 北京中誉威圣知识产权代理
有限公司 11279
代理人 李泽中 龚镇雄

(51) Int. Cl.
G10L 15/26(2006.01)
H04N 7/15(2006.01)

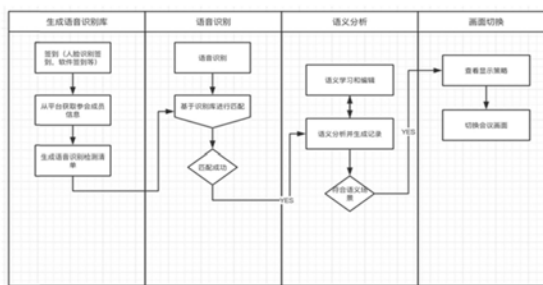
权利要求书1页 说明书4页 附图1页

(54)发明名称

基于语音识别的会议画面切换装置及其方法

(57)摘要

本发明公开了一种基于语音识别的会议画面切换装置及其方法,该基于语音识别的会议画面切换方法,包括以下步骤:步骤一,基于签到信息和平台通信录信息生成语音识别库;步骤二,基于语音识别库进行会议语音识别,以发现匹配结果;步骤三,对匹配结果进行语义分析;步骤四,根据语义分析结果切换画面。该基于语音识别的会议画面切换方法使得会议画面切换更智能,具有面对面开会的真实体验。并增加语义识别功能,既能对用户想观看的参会人进行准确判断,又能添加特定的语音操作,提升了会议体验。



1. 一种基于语音识别的会议画面切换装置,其特征在于,包括:
语音识别库生成模块,其基于签到信息和平台通信录信息生成语音识别库;
语音识别模块,其基于语音识别库进行会议语音识别,以发现匹配结果;
语义分析模块,其用于对语音识别模块的匹配结果进行语义分析;以及
画面切换模块,其用于根据语义分析结果切换画面。
2. 一种如权利要求书1所述的基于语音识别的会议画面切换装置的切换方法,其特征在于,包括以下步骤:
步骤一,基于签到信息和平台通信录信息生成语音识别库;
步骤二,基于语音识别库进行会议语音识别,以发现匹配结果;
步骤三,对匹配结果进行语义分析;
步骤四,根据语义分析结果切换画面。
3. 如权利要求2所述的基于语音识别的会议画面切换方法,其特征在于,所述步骤一具体包括:签到步骤、从平台获取参会成员信息步骤和生成语音识别检测清单步骤。
4. 如权利要求2所述的基于语音识别的会议画面切换方法,其特征在于,所述平台通信录信息包括:参会人姓名、昵称和他人对其的备注称呼。
5. 如权利要求2所述的基于语音识别的会议画面切换方法,其特征在于,所述签到信息的获取方式包括:人脸识别、手动签到、刷卡签到和终端自动签到。
6. 如权利要求3所述的基于语音识别的会议画面切换方法,其特征在于,所述步骤二具体包括:语音识别步骤、基于语音识别库进行匹配步骤和匹配判断步骤。
7. 如权利要求6所述的基于语音识别的会议画面切换方法,其特征在于,所述步骤三具体包括:语义学习和编辑步骤、语义分析并生成记录步骤和符合语义场景判断步骤。
8. 如权利要求2所述的基于语音识别的会议画面切换方法,其特征在于,所述语义分析指分析出主要语句是呼叫某人还是谈及某人或是指挥操作。
9. 如权利要求7所述的基于语音识别的会议画面切换方法,其特征在于,所述步骤四具体包括:查看显示策略步骤和切换会议画面步骤。
10. 如权利要求2所述的基于语音识别的会议画面切换方法,其特征在于,所述切换画面的显示策略为:优先显示画面人物占比大的画面,若画面人物占比相当,则优先显示人物正面画面。

基于语音识别的会议画面切换装置及其方法

技术领域

[0001] 本发明是关于无线通信领域,特别是关于一种基于语音识别的会议画面切换装置及其方法。

背景技术

[0002] 视频会议过程中,尤其是在多方会议中,经常需要切换视频会议对象画面,以保证更好的会议效果。

[0003] 现有技术的会议画面切换为手动切换或语音激励,其中,手动切换过程需要领导或其它成员提出视频切换要求,然后操作人员进行视频切换,操作过程还需要查找需要切换的会议画面,效率低下,体验不佳。而语音激励过程中,某个发言方说话,在识别到声音后,才进行相应视频的切换。但在实际过程中,发言方人员暂时离开,未听清呼叫,都无法触发切换功能,误判率较高,且不支持主动切换功能,期间只能忐忑等待和不断询问,效果差,体验也差。

[0004] 公开于该背景技术部分的信息仅仅旨在增加对本发明的总体背景的理解,而不应当被视为承认或以任何形式暗示该信息构成已为本领域一般技术人员所公知的现有技术。

发明内容

[0005] 本发明的目的在于提供一种基于语音识别的会议画面切换装置及其方法,其使得会议画面切换更智能,具有面对面开会的真实体验。既能对用户想观看的参会人进行准确判断,又能添加特定的语音操作,提升了会议体验。

[0006] 为实现上述目的,本发明提供了一种基于语音识别的会议画面切换装置,包括:语音识别库生成模块,其基于签到信息和平台通信录信息生成语音识别库;语音识别模块,其基于语音识别库进行会议语音识别,以发现匹配结果;语义分析模块,其用于对语音识别模块的匹配结果进行语义分析;以及画面切换模块,其用于根据语义分析结果切换画面。

[0007] 本发明还提供了一种基于语音识别的会议画面切换方法,包括以下步骤:步骤一,基于签到信息和平台通信录信息生成语音识别库;步骤二,基于语音识别库进行会议语音识别,以发现匹配结果;步骤三,对匹配结果进行语义分析;步骤四,根据语义分析结果切换画面。

[0008] 在一优选的实施方式中,步骤一具体包括:签到步骤、从平台获取参会成员信息步骤和生成语音识别检测清单步骤。

[0009] 在一优选的实施方式中,平台通信录信息包括:参会人姓名、昵称和他人对其的备注称呼。

[0010] 在一优选的实施方式中,签到信息的获取方式包括:人脸识别、手动签到、刷卡签到和终端自动签到。

[0011] 在一优选的实施方式中,步骤二具体包括:语音识别步骤、基于语音识别库进行匹配步骤和匹配判断步骤。

[0012] 在一优选的实施方式中,步骤三具体包括:语义学习和编辑步骤、语义分析并生成记录步骤和符合语义场景判断步骤。

[0013] 在一优选的实施方式中,语义分析指分析出主要语句是呼叫某人还是谈及某人或是指挥操作。

[0014] 在一优选的实施方式中,步骤四具体包括:查看显示策略步骤和切换会议画面步骤。

[0015] 在一优选的实施方式中,切换画面的显示策略为:优先显示画面人物占比大的画面,若画面人物占比相当,则优先显示人物正面画面。

[0016] 与现有技术相比,根据本发明的基于语音识别的会议画面切换方法基于签到和通信录信息整合为语音识别库,使得会议画面切换更智能,具有面对面开会的真实体验。并增加语义识别功能,既能对用户想观看的参会人进行准确判断,又能添加特定的语音操作,提升了会议体验。

附图说明

[0017] 图1是根据本发明一实施方式的基于语音识别的会议画面切换方法的流程图。

具体实施方式

[0018] 下面结合附图,对本发明的具体实施方式进行详细描述,但应当理解本发明的保护范围并不受具体实施方式的限制。

[0019] 除非另有其它明确表示,否则在整个说明书和权利要求书中,术语“包括”或其变换如“包含”或“包括有”等等将被理解为包括所陈述的元件或组成部分,而并未排除其它元件或其它组成部分。

[0020] 根据本发明优选实施例的基于语音识别的会议画面切换装置包括:语音识别库生成模块、语音识别模块、语义分析模块和画面切换模块。其中,语音识别库生成模块基于签到信息和平台通信录信息生成语音识别库;语音识别模块基于语音识别库进行会议语音识别,以发现匹配结果;语义分析模块用于对语音识别模块的匹配结果进行语义分析;画面切换模块用于根据语义分析结果切换画面。

[0021] 如图1所示,根据本发明优选实施例的基于语音识别的会议画面切换装置的切换方法的主流程为:语音识别名称-匹配名称所对应的终端(或所在会议室系统)-切换相应终端(或所在会议室系统)的摄像头画面。具体包括以下步骤:

[0022] 首先,基于签到信息和平台通信录信息生成语音识别库;

[0023] 该步骤主要涉及平台通信录,签到功能。其中,平台通信录信息包括:参会人姓名、昵称和他人对其的备注称呼等信息。签到信息的获取方式包括:人脸识别、手动签到、刷卡签到和终端自动签到,用于确定会议室的在场成员及所在会场等信息。该会议画面切换方法结合以上信息形成语音识别库和查询表。

[0024] 其次,基于语音识别库进行会议语音识别,以发现匹配结果;

[0025] 例如:“……接下来的时间交给张三,‘张三’(即发现匹配结果的呼叫过程)”。

[0026] 再次,对匹配结果进行语义分析;

[0027] 语义分析指分析出主要语句是呼叫某人还是谈及某人或是指挥操作;如确定语义

为呼叫指令后,根据查询表找到指定会场,进行画面切换。例如:“……接下来的时间交给张三,‘张三’,前面的话中出现了两个“张三”,均会在语音识别中进行识别,因识别到名称,两个都会丢给语义分析步骤进行分析,根据前后停顿时间,连贯性,或以往技术手段分析是否是呼叫指令。会议过程中,无感切换(比如张三想让李四发表意见,只需要说李四,你有什么看法,此时画面就切换到了李四的画面,然后大家可以看着李四的视频画面等待李四的回答,更具面对面开会的真实体验)。而且该基于语音识别的会议画面切换方法增加语义识别功能,既能对用户想观看参会人进行准确判断,又能添加特定的语音操作(指挥操作),比如切换到北京会场。

[0028] 最后,根据语义分析结果切换画面。

[0029] 基于显示策略切换画面(单一窗口时,直接切换画面,多窗口时,大窗口切换画面,多窗口等大时,排在最前面的窗口切换画面等),例如:张三可能在会议室1中进行了签到,同时在会议室1中使用了笔记本电脑、手机进行会议,此时,有两个以上摄像头都对准了张三,如会议室1系统摄像头(一个会议室系统也可能有2个以上摄像头),笔记本终端摄像头。此时可以有个优先算法,如画面人物占比大的优先显示(上面肯定是笔记本终端摄像头),若占比差不多,优先显示正面画面等。也可以配合语音激励的方式,选择拾音距离近的。

[0030] 综上,该基于语音识别的会议画面切换方法基于签到和通信录信息整合为语音识别库,使得会议画面切换更智能,具有面对面开会的真实体验。并增加语义识别功能,既能对用户想观看的参会人进行准确判断,又能添加特定的语音操作,提升了会议体验。

[0031] 本领域内的技术人员应明白,本申请的实施例可提供为方法、系统、或计算机程序产品。因此,本申请可采用完全硬件实施例、完全软件实施例、或结合软件和硬件方面的实施例的形式。而且,本申请可采用在一个或多个其中包含有计算机可用程序代码的计算机可用存储介质(包括但不限于磁盘存储器、CD-ROM、光学存储器等)上实施的计算机程序产品的形式。

[0032] 本申请是参照根据本申请实施例的方法、设备(系统)、和计算机程序产品的流程图和/或方框图来描述的。应理解可由计算机程序指令实现流程图和/或方框图中的每一流程和/或方框、以及流程图和/或方框图中的流程和/或方框的结合。可提供这些计算机程序指令到通用计算机、专用计算机、嵌入式处理机或其他可编程数据处理设备的处理器以产生一个机器,使得通过计算机或其他可编程数据处理设备的处理器执行的指令产生用于实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能的装置。

[0033] 这些计算机程序指令也可存储在能引导计算机或其他可编程数据处理设备以特定方式工作的计算机可读存储器中,使得存储在该计算机可读存储器中的指令产生包括指令装置的制品,该指令装置实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能。

[0034] 这些计算机程序指令也可装载到计算机或其他可编程数据处理设备上,使得在计算机或其他可编程设备上执行一系列操作步骤以产生计算机实现的处理,从而在计算机或其他可编程设备上执行的指令提供用于实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能的步骤。

[0035] 前述对本发明的具体示例性实施方案的描述是为了说明和例证的目的。这些描述并非想将本发明限定为所公开的精确形式,并且很显然,根据上述教导,可以进行很多改变

和变化。对示例性实施例进行选择 and 描述的目的在于解释本发明的特定原理及其实际应用,从而使得本领域的技术人员能够实现并利用本发明的各种不同的示例性实施方案以及各种不同的选择和改变。本发明的范围意在由权利要求书及其等同形式所限定。

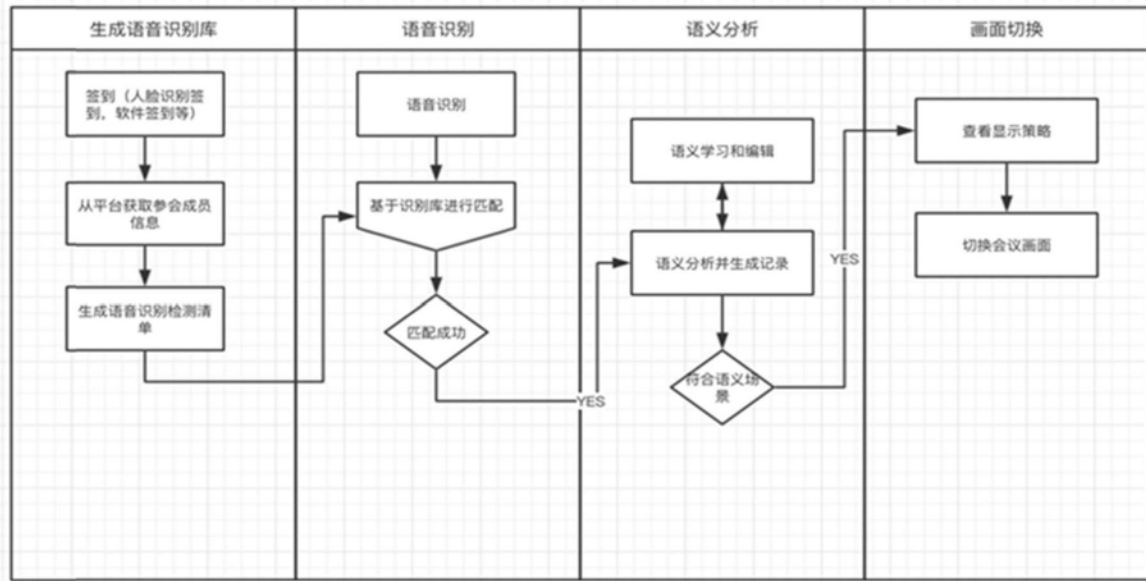


图1