



(12)发明专利申请

(10)申请公布号 CN 108810446 A

(43)申请公布日 2018.11.13

(21)申请号 201810581511.5

(22)申请日 2018.06.07

(71)申请人 北京智能管家科技有限公司

地址 101500 北京市密云区经济开发区西
统路8号西田各庄镇政府办公楼508
室-598

(72)发明人 汤炜 张晓雪 雷宇

(74)专利代理机构 北京品源专利代理有限公司

11332

代理人 孟金喆

(51)Int.Cl.

H04N 7/15(2006.01)

G10L 15/22(2006.01)

权利要求书2页 说明书9页 附图3页

(54)发明名称

一种视频会议的标签生成方法、装置、设备
和介质

(57)摘要

本发明实施例公开了一种视频会议的标签
生成方法、装置、设备和介质。涉及视频会议领
域。该方法包括：根据视频会议的音频提取所述
视频会议的标识信息，其中所述标识信息用以标
示所述视频会议；根据所述标识信息生成所述视
频会议的标签。本发明实施例提供一种视频会议
的标签生成方法、装置、设备和介质，实现了视频
会议标签的自动生成，解决了人工对会议日期及
参会人等固定选项的填选操作导致的增加用户
工作量的问题。



1. 一种视频会议的标签生成方法,其特征在于,包括:

根据视频会议的音频提取所述视频会议的标识信息,其中,所述标识信息用以标示所述视频会议;

根据所述标识信息生成所述视频会议的标签。

2. 根据权利要求1所述的方法,其特征在于,若所述标识信息包括参会人信息时,根据视频会议的音频提取所述视频会议的标识信息包括:

提取所述视频会议的音频中的非静音片段;

根据所述非静音片段确定参会人。

3. 根据权利要求2所述的方法,其特征在于,根据所述非静音片段确定参会人包括:

根据所述非静音片段的时间戳,确定与所述非静音片段同步的视频片段;

对所述视频片段中的视频帧进行人脸识别;

根据识别结果确定参会人。

4. 根据权利要求1所述的方法,其特征在于,若所述标识信息包括摘要时,根据视频会议的音频提取所述视频会议的标识信息包括:

对所述视频会议的音频进行语音识别,生成会议文本;

提取所述会议文本的摘要,作为所述视频会议的摘要。

5. 根据权利要求4所述的方法,其特征在于,提取所述会议文本的摘要包括:

提取所述会议文本的主旨句;

确定所述主旨句与所述会议文本整体语义的匹配程度;

基于匹配程度,根据所述主旨句确定所述会议文本的摘要。

6. 根据权利要求1所述的方法,其特征在于,在根据所述标识信息生成所述视频会议的标签之后,还包括:

将检索词与存储的各视频会议的标签进行匹配,确定目标标签,并将所述目标标签关联的视频会议的视频和/或音频作为目标会议文件;

确定所述目标会议文件中与所述目标标签相关的目标内容片段。

7. 根据权利要求6所述的方法,其特征在于,确定所述目标会议文件中与所述目标标签相关的目标内容片段包括:

若所述目标标签为参会人,则根据参会人所在非静音片段的时间戳,确定所述目标会议文件中参会人参与的片段,并将参会人参与的片段作为所述目标内容片段,其中所述非静音片段从视频会议的音频中提取得到。

8. 根据权利要求6所述的方法,其特征在于,在确定所述目标会议文件中与所述目标标签相关的目标内容片段之后,还包括:

响应目标会议播放事件,将所述目标会议文件的播放进程直接跳至所述目标内容片段进行播放。

9. 根据权利要求1所述的方法,其特征在于,根据所述标识信息生成所述视频会议的标签包括:

若所述标识信息包括参会人,则确定所述参会人的相关信息,其中所述相关信息包括参会人的性别、职务和所属部门中的至少一种信息;

将所述相关信息作为所述视频会议的标签。

10. 一种视频会议的标签生成装置,其特征在于,包括:

标识提取模块,用于根据视频会议的音频提取所述视频会议的标识信息,其中,所述标识信息用以标示所述视频会议;

标签确定模块,用于根据所述标识信息生成所述视频会议的标签。

11. 一种设备,其特征在于,所述设备包括:

一个或多个处理器;

存储装置,用于存储一个或多个程序;

当所述一个或多个程序被所述一个或多个处理器执行,使得所述一个或多个处理器实现如权利要求1-9中任一所述的视频会议的标签生成方法。

12. 一种计算机可读存储介质,其上存储有计算机程序,其特征在于,该程序被处理器执行时实现如权利要求1-9中任一所述的视频会议的标签生成方法。

一种视频会议的标签生成方法、装置、设备和介质

技术领域

[0001] 本发明实施例涉及视频会议领域,尤其涉及一种视频会议的标签生成方法、装置、设备和介质。

背景技术

[0002] 视频会议(也称电视会议)是指位于两个或多个地点的人们,通过通信设备和网络,将声音、影像及文件资料互传,实现即时且能互动沟通的会议。

[0003] 通常,视频会议系统会录制视频会议进程的整段视频和音频,并将人为填选的会议日期及参会人等固定选项作为会议的标签,与会议的视频和/或会议的音频关联存储。事后可以根据会议的标签对存储的会议的视频和/或会议的音频进行检索和回放。

[0004] 发明人在实现本发明的过程中发现,现有技术存在如下缺陷:人为对会议日期及参会人等固定选项的填选操作,增加了用户工作量。

发明内容

[0005] 本发明实施例提供一种视频会议的标签生成方法、装置、设备和介质,以实现视频会议标签的自动生成,从而解决人为对会议日期及参会人等固定选项的填选操作,增加用户工作量的问题。

[0006] 第一方面,本发明实施例提供了一种视频会议的标签生成方法,该方法包括:

[0007] 根据视频会议的音频提取所述视频会议的标识信息,其中所述标识信息用以标示所述视频会议;

[0008] 根据所述标识信息生成所述视频会议的标签。

[0009] 第二方面,本发明实施例还提供了一种视频会议的标签生成装置,该装置包括:

[0010] 标识提取模块,用于根据视频会议的音频提取所述视频会议的标识信息,其中,所述标识信息用以标示所述视频会议;

[0011] 标签确定模块,用于根据所述标识信息生成所述视频会议的标签。

[0012] 第三方面,本发明实施例还提供了一种设备,所述设备包括:

[0013] 一个或多个处理器;

[0014] 存储装置,用于存储一个或多个程序;

[0015] 当所述一个或多个程序被所述一个或多个处理器执行,使得所述一个或多个处理器实现如本发明实施例中任一所述的视频会议的标签生成方法。

[0016] 第四方面,本发明实施例还提供了一种计算机可读存储介质,其上存储有计算机程序,该程序被处理器执行时实现如本发明实施例中任一所述的视频会议的标签生成方法。

[0017] 本发明实施例通过提取所述视频会议的标识信息,根据所述标识信息确定视频会议的标签。因为标识信息可以通过识别视频会议的音频得到,该过程无需人工参与,所以本发明实施例实现了会议标签的自动生成,进而解决了人工对会议日期及参会人等固定选项

的填选操作导致的增加用户工作量的问题。

附图说明

- [0018] 图1为本发明实施例一提供的一种视频会议的标签生成方法的流程图；
- [0019] 图2是本发明实施例提供的一种视频会议的标签生成方法的流程图；
- [0020] 图3是本发明实施例三提供的一种视频会议的标签生成方法的流程图；
- [0021] 图4是本发明实施例四提供的一种视频会议的标签生成方法的流程图；
- [0022] 图5a是本发明实施例五提供的一种视频会议的标签生成方法的流程图；
- [0023] 图5b是本发明实施例五提供的一种摘要生成模型的训练过程的示意图；
- [0024] 图6是本发明实施例六提供的一种视频会议的标签生成装置的结构示意图；
- [0025] 图7为本发明实施例七提供的设备的结构示意图。

具体实施方式

[0026] 下面结合附图和实施例对本发明作进一步的详细说明。可以理解的是，此处所描述的具体实施例仅仅用于解释本发明，而非对本发明的限定。另外还需要说明的是，为了便于描述，附图中仅示出了与本发明相关的部分而非全部结构。

[0027] 实施例一

[0028] 图1为本发明实施例一提供的一种视频会议的标签生成方法的流程图。本实施例可适用于对视频会议进行标签自动生成的情况。该方法可以由一种视频会议的标签生成装置来执行，该装置可以由软件和/或硬件的方式实现。参见图1，本实施例提供的一种视频会议的标签生成方法包括：

[0029] S110、根据视频会议的音频提取所述视频会议的标识信息，其中所述标识信息用以标示所述视频会议。

[0030] 具体的，视频会议的音频的获取可以是通过视频会议中的麦克获取，也可以从基于流媒体技术录制的视频会议的文件中获取。可选的，视频会议的音频还可以是从接入视频会议系统的移动终端获取的参会人的发言。

[0031] 视频会议的标识信息是指区别于其他视频会议的信息，可以包括视频会议的摘要、开会日期、参会人和/或参会人数等。

[0032] S120、根据所述标识信息生成所述视频会议的标签。

[0033] 其中，所述视频会议的标签是指识别所述视频会议的信息，可以是词、短语或句子。

[0034] 可选的，可以将所述标识信息直接作为所述视频会议的标签，也可以对所述标识信息进一步处理，根据处理结果确定所述视频会议的标签。

[0035] 具体的，若所述标识信息是参会人或参会人数，则直接将参会人或参会人数作为所述视频会议的标签。

[0036] 若所述标识信息是所述视频会议的摘要，则提取摘要的关键词，将提取得到的关键词作为所述视频会议的标签。

[0037] 本发明实施例的技术方案，通过提取所述视频会议的标识信息，根据所述标识信息确定视频会议的标签。因为标识信息可以通过识别视频会议的音频得到，该过程无需人

工参与,所以本发明实施例实现了会议标签的自动生成,进而解决人工对会议日期及参会人等固定选项的填选操作导致的增加用户工作量的问题。

[0038] 为提高标签的丰富度,根据所述标识信息生成所述视频会议的标签包括:

[0039] 若所述标识信息包括参会人,则确定所述参会人的相关信息,其中所述相关信息包括参会人的性别、职务和所属部门中的至少一种信息;

[0040] 将所述相关信息作为所述视频会议的标签。

[0041] 示例性的,若提取到视频会议的参会人是张三,张三的相关信息包括性别男、职务为研发工程师和所述部门为研发部,则将男、研发工程师和研发部作为所述视频会议的标签。

[0042] 可以理解的是,通过上述方法,可以通过参会人的相关信息对视频会议进行检索。从而方便在忘记参会人姓名的情况下,通过参会人相关信息对参会人参与的视频会议进行检索。

[0043] 实施例二

[0044] 图2是本发明实施例提供的一种视频会议的标签生成方法的流程图。本实施例是在上述实施例的基础上,提取的一种可选方案。参见图2,本实施例提供的一种视频会议的标签生成方法包括:

[0045] S210、提取视频会议的音频中的非静音片段。

[0046] 其中,非静音片段指存在人声语音信号或音乐语音信号的音频片段。

[0047] 可选的,非静音片段的确定方式可以是:将所述视频会议的音频中音频信号幅度大于设定幅度阈值的音频片段作为所述非静音片段。

[0048] 非静音片段的确定也可以是:识别静音片段;所述视频会议的音频中除去静音片段剩余的即为非静音片段。其中静音片段指在信号可见波形上没有任何音频的波形或者较小波动的音频片段。信号可见波形为经过噪声滤波后得到的波形。

[0049] S220、根据所述非静音片段确定参会人。

[0050] 具体的,根据所述非静音片段确定参会人包括:

[0051] 根据所述非静音片段的时间戳,确定与所述非静音片段同步的视频片段;

[0052] 对所述视频片段中的视频帧进行人脸识别;

[0053] 根据识别结果确定参会人。

[0054] 可选的,也可以基于不同人具有不同音色,根据对所述非静音片段的音色的识别确定参会人。

[0055] S230、根据参会人确定所述视频会议的标签。

[0056] 本发明实施例的技术方案,通过对非静音片段的识别确定参会人,根据参会人确定视频会议的标签。从而避免对在视频会议的间歇进入会议室的非参会人员的误判,其中通常在会议间歇会关掉麦克或控制较低音量进行交谈。

[0057] 实施例三

[0058] 图3是本发明实施例三提供的一种视频会议的标签生成方法的流程图。本实施例是在上述实施例的基础上提出的一种可选方案。参见图3,本实施例提供的视频会议的标签生成方法包括:

[0059] S310、对视频会议的音频进行语音识别,生成会议文本。

- [0060] 其中，语音识别可以使用现有技术中任意一种语音识别算法实现。
- [0061] 典型的，可以将经过深度学习得到的递归神经网络作为语音识别模型，实现对视频会议的音频的语音识别。
- [0062] S320、提取所述会议文本的摘要，作为所述视频会议的摘要。
- [0063] 其中，摘要提取可以使用现有技术中任意一种摘要生成方法实现。
- [0064] 具体的，提取所述会议文本摘要的方法可以是：
- [0065] 将所述会议文本转为词向量表示；
- [0066] 利用预先训练的摘要生成模型，生成所述会议文本的摘要。
- [0067] 其中，摘要生成模型可以结合递归神经网络的深度学习技术和强化学习技术训练得到。具体过程可以描述为：
- [0068] 将词向量输入摘要生成模型，生成摘要；
- [0069] 将生成的摘要与设定的参考摘要进行评价；
- [0070] 根据评价高低选择抑制或鼓励的策略更新摘要生成模型参数，使得摘要生成模型可以生成语法正确，文字流畅的摘要内容。
- [0071] S330、根据摘要确定所述视频会议的标签。
- [0072] 本发明实施例的技术方案，通过对语音识别生成的会议文本进行摘要提取，根据提取的摘要确定视频会议的标签。从而实现根据视频会议的主要内容进行标签的自动生成。
- [0073] 实施例四
- [0074] 图4是本发明实施例四提供的一种视频会议的标签生成方法的流程图。本实施例是在上述实施例的基础上提出的一种可选方案。参见图4，本实施例提供的视频会议的标签生成方法包括：
- [0075] S410、提取所述会议文本的主旨句。
- [0076] 具体的，提取所述会议文本的主旨句包括：根据句子在所述会议文本中的位置，提取所述会议文本中代表主旨的主旨句。
- [0077] 其中，可以将所述会议文本开头的至少一句话作为主旨句，或者，将所述会议文本中每段文本开头的两到三句话作为主旨句。因为通常会议开始会先介绍本次会议的主要内容。
- [0078] S420、确定所述主旨句与所述会议文本整体语义的匹配程度。
- [0079] 具体的，可以基于词频，确定所述主旨句与所述会议文本整体语义的匹配程度。例如，根据在所述主旨句中的词在所述会议文本中的出现频率，确定所述主旨句与所述会议文本整体语义的匹配程度。
- [0080] 为提高匹配程度的准确率，确定所述主旨句与所述会议文本整体语义的匹配程度可以包括：
- [0081] 将会议文本中的句子转为语义向量表示，生成句子语义向量；
- [0082] 累加所有句子语义向量得到文本的整体语义向量表示，生成文本语义向量；
- [0083] 计算句子语义向量和文本语义向量的cos距离值，得到当前句子与文本整体语义的匹配程度。
- [0084] 其中，距离越近，匹配程度越高，越有代表性。

[0085] 具体的，累加所有句子语义向量得到文本的整体语义向量表示可以是，使用word embedding的思想累加所有句子的语义向量得到文本的整体语义向量表示。

[0086] S430、基于匹配程度，根据所述主旨句确定所述会议文本的摘要。

[0087] 具体的，基于匹配程度，根据所述主旨句确定所述会议文本的摘要的方法可以是：将所述主旨句按照匹配程度由高到低的顺序进行排列；根据排序结果，取前n个主旨句作为所述会议文本的摘要，其中n可以根据实际需要确定。

[0088] S440、根据摘要确定所述视频会议的标签。

[0089] 本发明实施例的技术方案，通过提取所述会议文本的主旨句，根据主旨句与会议文本整体语义的匹配程度确定会议文本的摘要。相比直接根据会议文本的全文进行摘要提取，提高了摘要提取的准确率。此外，在确定主旨句时结合了会议自身特点(即通常会议开始会先介绍本次会议的主要内容)，确定主旨句，从而提高了视频会议的主旨句的确定准确率。

[0090] 实施例五

[0091] 图5a是本发明实施例五提供的一种视频会议的标签生成方法的流程图。本实施例是在上述实施例的基础上提出的一种可选方案。参见图5，本实施例提供的视频会议的标签生成方法包括：

[0092] S510、根据视频会议的音频提取所述视频会议的标识信息，其中，所述标识信息用以标示所述视频会议。

[0093] S520、根据所述标识信息生成所述视频会议的标签。

[0094] S530、将检索词与存储的各视频会议的标签进行匹配，确定目标标签，并将所述目标标签关联的视频会议的视频和/或音频作为目标会议文件。

[0095] S540、确定所述目标会议文件中与所述目标标签相关的目标内容片段。

[0096] 具体的，可以根据目标标签在目标文件中的位置，确定所述目标会议文件中与所述目标标签相关的目标内容片段。也可以将目标标签对应的目标文件中的片段，确定为与所述目标标签相关的目标内容片段。

[0097] 典型的，若所述目标标签为参会人，则根据参会人所在非静音片段的时间戳，确定所述目标会议文件中参会人参与的片段，并将参会人参与的片段作为所述目标内容片段，其中所述非静音片段从视频会议的音频中提取得到。

[0098] 若所述目标标签为摘要，则确定摘要所在句子或段落；根据该句子或该段落在会议文本中的位置，确定目标内容片段。

[0099] 在实际应用中，视频会议的标签生成方法按照即时处理情况可以分为：实时处理的情况和离线处理的情况。具体可以描述为：

[0100] 1、实时处理的情况

[0101] 1) 在会议进行中，对通过麦克获取的语音信号进行语音识别，得到文字信息。此处由终端设备对语音信号进行语音特征提取，由服务器基于深度学习递归神经网络的语音识别算法对语音特征进行识别。

[0102] 2) 服务器利用基于递归神经网络的自然语言处理技术，实时分析上述文字信息，生成会议记录文本并存入数据库。当会议结束后，根据会议记录文本自动生成摘要，其中包括如下两种方式：

[0103] (1) 抽取式,抽取会议记录文本中主旨句(通常选用每段开头的两到三句话,特别是会议记录文本开头的几句话作为主旨句);利用word2vec技术把句子转为语义向量表示;使用word embedding的思想累加所有句子的语义向量得到文本的整体语义向量表示;通过计算句子语义向量和文本整体语义向量的cos距离值,得到当前句子与文本整体语义的匹配程度,其中距离越近,匹配程度越高,越有代表性;把每个句子与文本整体语义相似性得分全部计算完毕后,根据得分由高到低排序,结合需求取排序后的前面一定数量的句子作为摘要内容。

[0104] (2) 生成式,利用word2vec技术把文本转为词向量表示;将词向量输入预先训练完成的摘要生成模型,生成摘要。其中,参见图5b摘要生成模型的训练包括:利用基于递归神经网络的深度学习技术和强化学习技术,通过自动生成的摘要进行评价;根据评价高低选择抑制或鼓励的策略更新摘要生成模型参数;使得摘要生成模型可以生成语法正确,文字流畅的摘要内容和标签内容。

[0105] 3) 根据生成的标签内容可实现利用自动生成的标签对会议内容进行检索的需求。

[0106] 2、离线处理的情况

[0107] 1) 在会议结束后,对利用流媒体技术录制的会议视频内容进行摘要的自动提取,具体可以描述为:

[0108] (1) 提取视频会议的音频数据;

[0109] (2) 利用静音帧检测算法将音频数据按静音片段和非静音片段进行区分,并对视频以非静音片段起止时间戳进行切分;

[0110] (3) 提取非静音片段对应的视频片段中的视频帧序列数据,对视频帧序列数据利用深度学习的卷积神经网络(cnn)算法进行人脸检测,识别说话人;

[0111] (4) 对非静音片段,利用基于深度学习递归神经网络算法进行语音识别,生成符合人类阅读习惯的文本信息;

[0112] (5) 对每段文本信息基于上述生成式自动摘要技术,生成该文本摘要;

[0113] (6) 结合非静音片段的文本摘要和非静音片段对应的视频片段的人脸识别结果,生成视频会议的多维度标签,其中标签内容包含说话人姓名、性别、职务、部门以及文本摘要信息等。

[0114] 3) 把会议的切分时间点信息(即非静音片段起止时间戳)及对应的多维度标签,关联保存在云端服务器的数据库中;

[0115] 4) 用户可通过手机应用的方式进行往期会议内容的智能检索,具体方式包含如下:

[0116] (1) 根据输入的参会人姓名、日期等检索词进行精确检索查找,并根据会议的切分时间点信息,将查找的会议视频内容按参会人发言位置进行视频时间点定位。

[0117] (2) 根据输入的感兴趣的标题或、短语或参会人等检索词,对输入内容进行语义分析,基于语义分析结果进行模糊检索查找,将查找的会议内容按标签相关性及参会人相似性进行排序。

[0118] 上述方案的有益效果在于:通过识别会议文本,根据会议文本实现会议视频标签的自动生成,免去人为参与录入和筛选总结的痛苦。通过多维度标签完善会议视频库检索机制,丰富会议视频的检索体验。通过利用移动终端应用对往期会议的检索,实现多维度全

平台的体验。

[0119] 本发明实施例的技术方案,通过确定目标会议文件中与所述目标标签相关的目标内容片段,使得用户基于目标内容片段可以实现对目标内容片段的直接查看,省去了用户从头浏览会议文件,以查找与检索内容相关的片段的麻烦。

[0120] 为进一步方便用户查看,在确定所述目标会议文件中与所述目标标签相关的目标内容片段之后,还包括:

[0121] 响应目标会议播放事件,将所述目标会议文件的播放进程直接跳至所述目标内容片段进行播放。

[0122] 可以理解的是,参会人在整个会议中可能至少两处出现。因此,目标标签(典型的是参会人)可能存在目标会议文件中存在至少两个目标内容片段。

[0123] 为进一步方便用户查看,可以分别对目标文件中的至少两个目标内容片段进行标记。

[0124] 响应目标会议播放事件,将所述目标会议文件的播放进程直接跳至所述目标内容片段进行播放包括:

[0125] 响应目标会议播放事件,将所述目标会议文件的播放进程直接跳至第一次出现目标标签的目标内容片段进行播放;

[0126] 在用户选择对目标标签关联的下一片段进行播放时,按照所述目标标签出现的顺序,将所述目标会议文件的播放进程直接跳至下一次出现目标标签的目标内容片段进行播放。

[0127] 为提高检索准确率,在利用检索词进行视频会议检索时,还包括:

[0128] 对用户输入的检索词进行语义分析;

[0129] 根据语义分析结果,确定目标会议文件。

[0130] 需要强调的是,经过本实施例的技术教导,本领域技术人员有动机将上述实施例中描述的任一种实施方式进行方案的组合,以实现对视频会议标签的自动生成。

[0131] 实施例六

[0132] 图6是本发明实施例六提供的一种视频会议的标签生成装置的结构示意图。参见图6,本实施例提供的视频会议的标签生成装置包括:标识提取模块10和标签确定模块20.

[0133] 其中,标识提取模块10,用于根据视频会议的音频提取所述视频会议的标识信息,其中所述标识信息用以标示所述视频会议;

[0134] 标签确定模块20,用于根据所述标识信息生成所述视频会议的标签。

[0135] 本发明实施例的技术方案,通过提取所述视频会议的标识信息,根据所述标识信息确定视频会议的标签。因为标识信息可以通过识别视频会议的音频得到,该过程无需人工参与,所以本发明实施例实现了会议标签的自动生成,进而解决人工对会议日期及参会人等固定选项的填选操作导致的增加用户工作量的问题。

[0136] 进一步的,标识提取模块10包括:非静音片段单元和参会人确定单元。

[0137] 非静音片段单元,用于若所述标识信息包括参会人,则提取视频会议的音频中的非静音片段;

[0138] 参会人确定单元,用于根据所述非静音片段确定参会人。

[0139] 进一步的,参会人确定单元具体用于:

- [0140] 根据所述非静音片段的时间戳，确定与所述非静音片段同步的视频片段；
[0141] 对所述视频片段中的视频帧进行人脸识别；
[0142] 根据识别结果确定参会人。
[0143] 进一步的，标识提取模块10包括：文本生成单元和摘要提取单元。
[0144] 其中，文本生成单元，用于若所述标识信息包括摘要，则对视频会议的音频进行语音识别，生成会议文本；
[0145] 摘要提取单元，用于提取所述会议文本的摘要，作为所述视频会议的摘要。
[0146] 摘要提取单元具体用于：
[0147] 提取所述会议文本的主旨句；
[0148] 确定所述主旨句与所述会议文本整体语义的匹配程度；
[0149] 基于匹配程度，根据所述主旨句确定所述会议文本的摘要。
[0150] 进一步的，所述装置还包括：目标检索模块和目标定位模块。
[0151] 其中，目标检索模块，用于在根据所述标识信息生成所述视频会议的标签之后，将检索词与存储的各视频会议的标签进行匹配，确定目标标签，并将所述目标标签关联的视频会议的视频和/或音频作为目标会议文件；
[0152] 目标定位模块，用于确定所述目标会议文件中与所述目标标签相关的目标内容片段。
[0153] 进一步的，目标定位模块包括：目标定位单元。
[0154] 其中，目标定位单元，用于若所述目标标签为参会人，则根据参会人所在非静音片段的时间戳，确定所述目标会议文件中参会人参与的片段，并将参会人参与的片段作为所述目标内容片段，其中所述非静音片段从视频会议的音频中提取得到。
[0155] 进一步的，所述装置还包括：播放模块。
[0156] 其中，播放模块，用于在确定所述目标会议文件中与所述目标标签相关的目标内容片段之后，响应目标会议播放事件，将所述目标会议文件的播放进程直接跳至所述目标内容片段进行播放。
[0157] 进一步的，标签确定模块20包括：信息获取单元和标签确定单元。
[0158] 信息获取单元，用于若所述标识信息包括参会人，则确定所述参会人的相关信息，其中所述相关信息包括参会人的性别、职务和所属部门中的至少一种信息；
[0159] 标签确定单元，用于根据所述相关信息确定所述视频会议的标签。
[0160] 实施例七
[0161] 图7为本发明实施例七提供的设备的结构示意图，如图7所示，该设备包括处理器70、存储器71、输入装置72和输出装置73；设备中处理器70的数量可以是一个或多个，图7中以一个处理器70为例；设备中的处理器70、存储器71、输入装置72和输出装置73可以通过总线或其他方式连接，图7中以通过总线连接为例。
[0162] 存储器71作为一种计算机可读存储介质，可用于存储软件程序、计算机可执行程序以及模块，如本发明实施例中的视频会议的标签生成方法对应的程序指令/模块(例如，视频会议的标签生成装置的标识提取模块10和标签确定模块20)。处理器70通过运行存储在存储器71中的软件程序、指令以及模块，从而执行设备的各种功能应用以及数据处理，即实现上述的视频会议的标签生成方法。

[0163] 存储器71可主要包括存储程序区和存储数据区，其中，存储程序区可存储操作系统、至少一个功能所需的应用程序；存储数据区可存储根据终端的使用所创建的数据等。此外，存储器71可以包括高速随机存取存储器，还可以包括非易失性存储器，例如至少一个磁盘存储器件、闪存器件、或其他非易失性固态存储器件。在一些实例中，存储器71可进一步包括相对于处理器70远程设置的存储器，这些远程存储器可以通过网络连接至设备。上述网络的实例包括但不限于互联网、企业内部网、局域网、移动通信网及其组合。

[0164] 输入装置72可用于接收输入的数字或字符信息，以及产生与设备的用户设置以及功能控制有关的键信号输入。输出装置73可包括显示屏等显示设备。

[0165] 实施例八

[0166] 本发明实施例八还提供一种包含计算机可执行指令的存储介质，所述计算机可执行指令在由计算机处理器执行时用于执行一种视频会议的标签生成方法，该方法包括：

[0167] 根据视频会议的音频提取所述视频会议的标识信息，其中所述标识信息用以标示所述视频会议；

[0168] 根据所述标识信息生成所述视频会议的标签。

[0169] 当然，本发明实施例所提供的一种包含计算机可执行指令的存储介质，其计算机可执行指令不限于如上所述的方法操作，还可以执行本发明任意实施例所提供的视频会议的标签生成方法中的相关操作。

[0170] 通过以上关于实施方式的描述，所属领域的技术人员可以清楚地了解到，本发明可借助软件及必需的通用硬件来实现，当然也可以通过硬件实现，但很多情况下前者是更佳的实施方式。基于这样的理解，本发明的技术方案本质上或者说对现有技术做出贡献的部分可以以软件产品的形式体现出来，该计算机软件产品可以存储在计算机可读存储介质中，如计算机的软盘、只读存储器(Read-Only Memory, ROM)、随机存取存储器(Random Access Memory, RAM)、闪存(FLASH)、硬盘或光盘等，包括若干指令用以使得一台计算机设备(可以是个人计算机、服务器，或者网络设备等)执行本发明各个实施例所述的方法。

[0171] 值得注意的是，上述视频会议的标签生成装置的实施例中，所包括的各个单元和模块只是按照功能逻辑进行划分的，但并不局限于上述的划分，只要能够实现相应功能即可；另外，各功能单元的具体名称也只是为了便于相互区分，并不用于限制本发明的保护范围。

[0172] 注意，上述仅为本发明的较佳实施例及所运用技术原理。本领域技术人员会理解，本发明不限于这里所述的特定实施例，对本领域技术人员来说能够进行各种明显的变化、重新调整和替代而不会脱离本发明的保护范围。因此，虽然通过以上实施例对本发明进行了较为详细的说明，但是本发明不仅仅限于以上实施例，在不脱离本发明构思的情况下，还可以包括更多其他等效实施例，而本发明的范围由所附的权利要求范围决定。

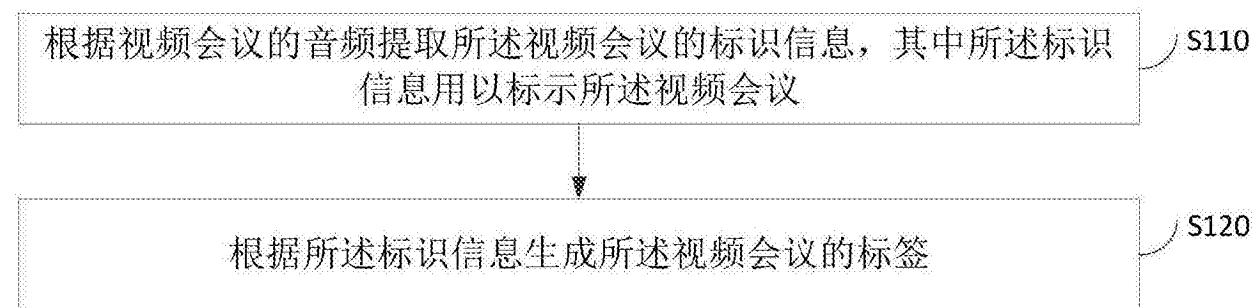


图1

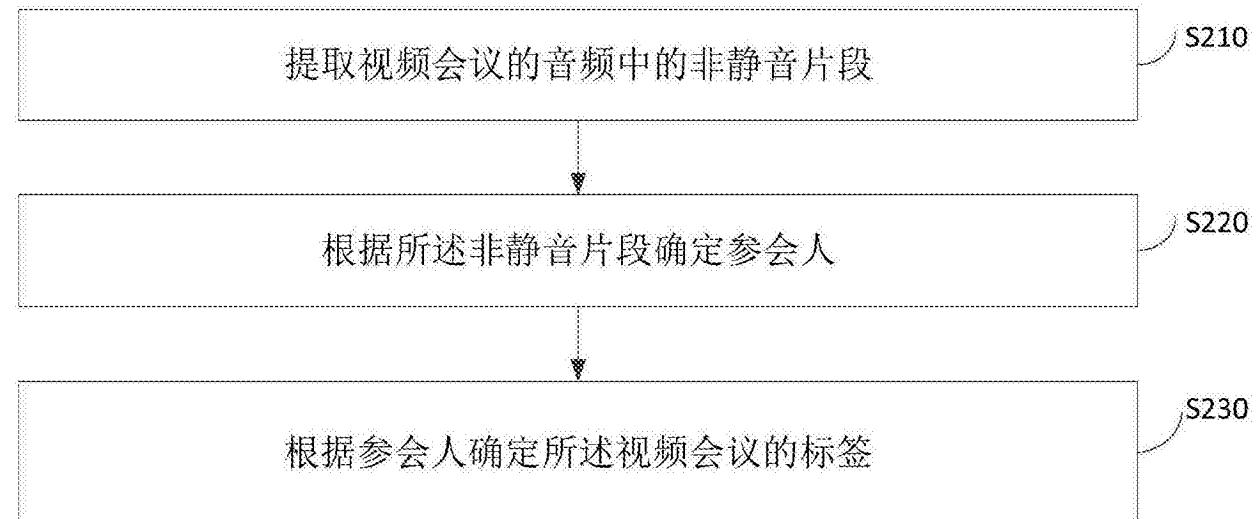


图2

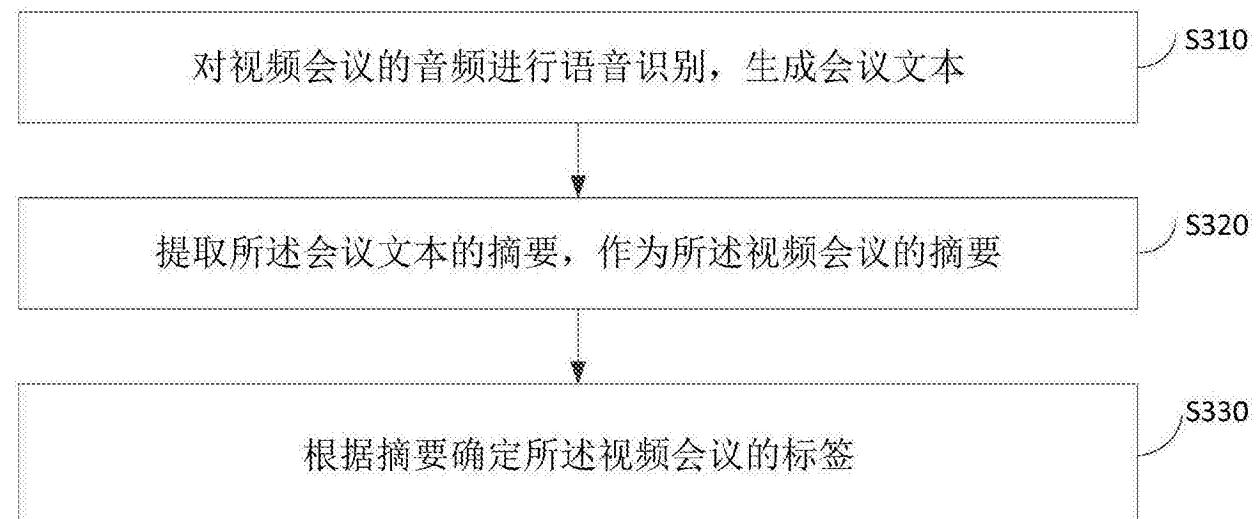


图3

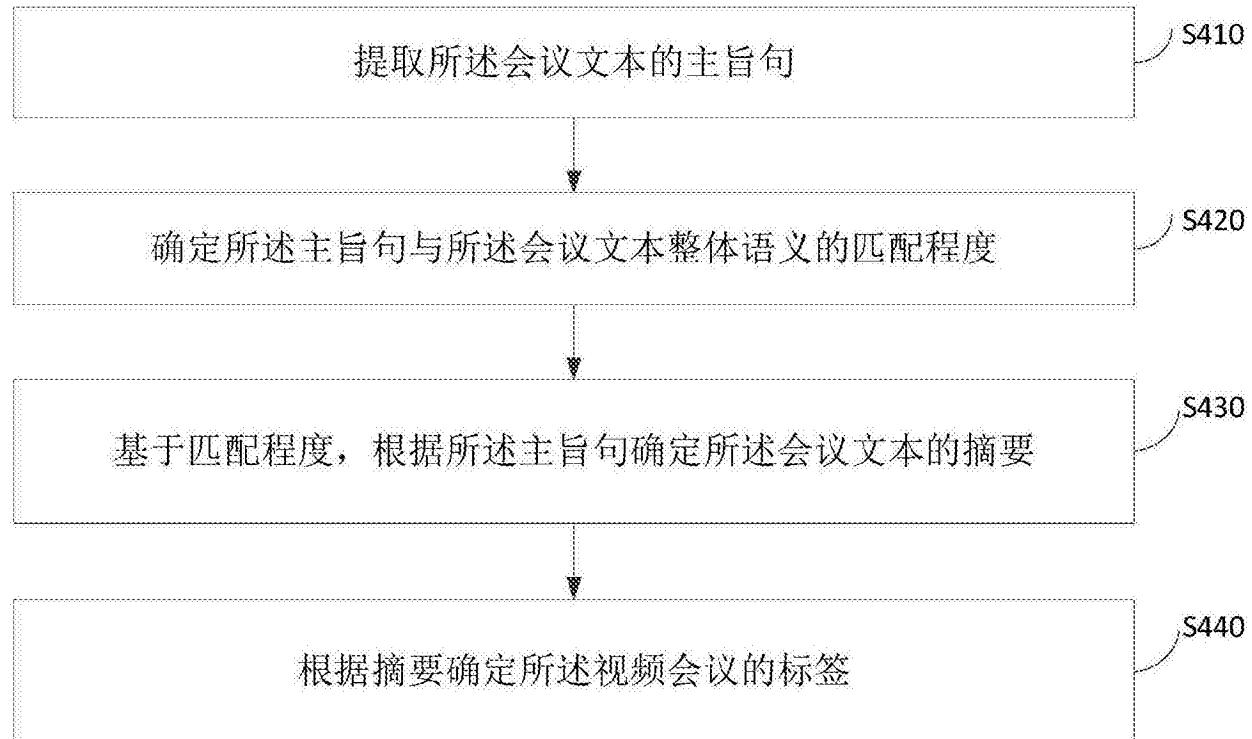


图4

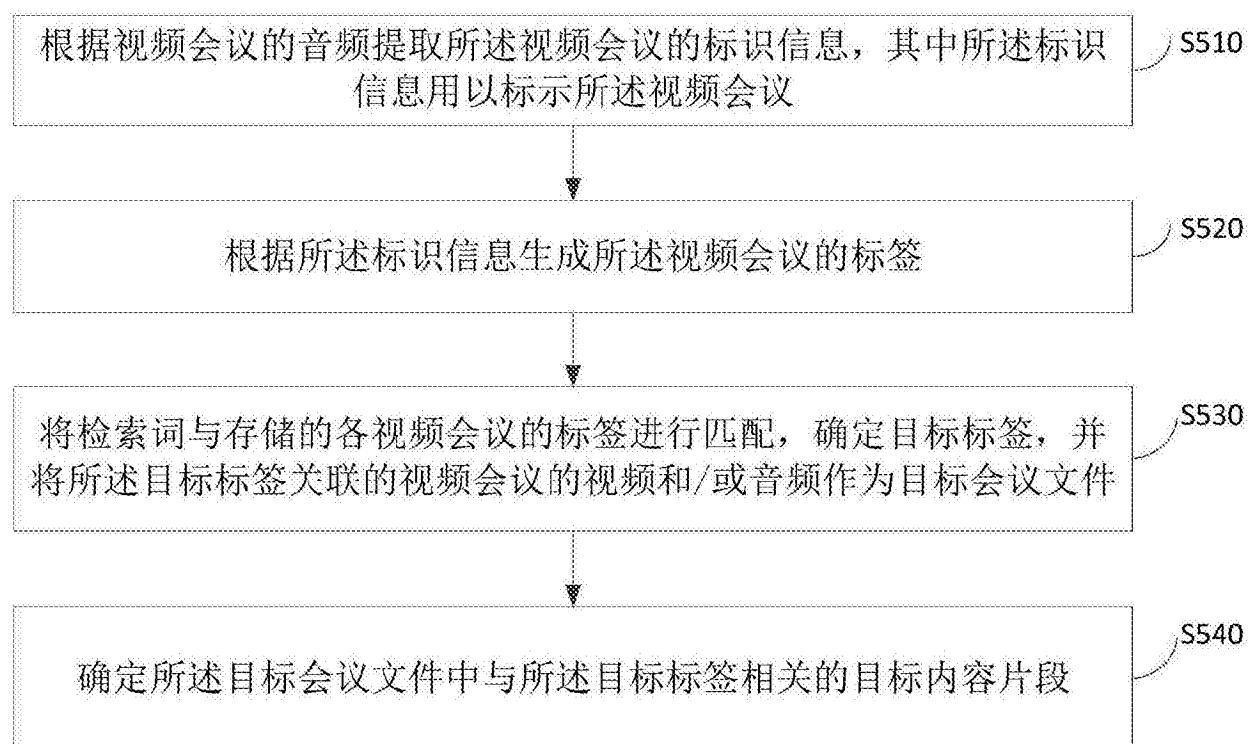


图5a

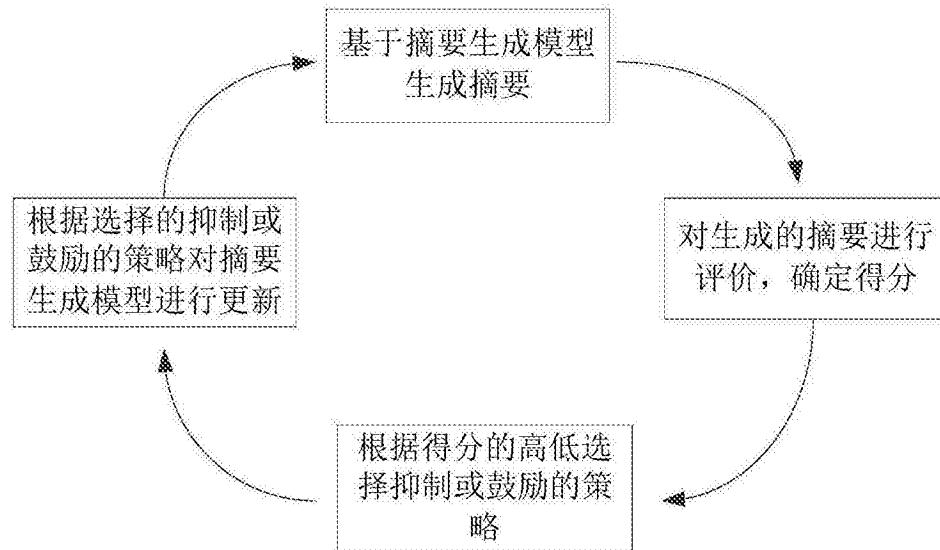


图5b



图6

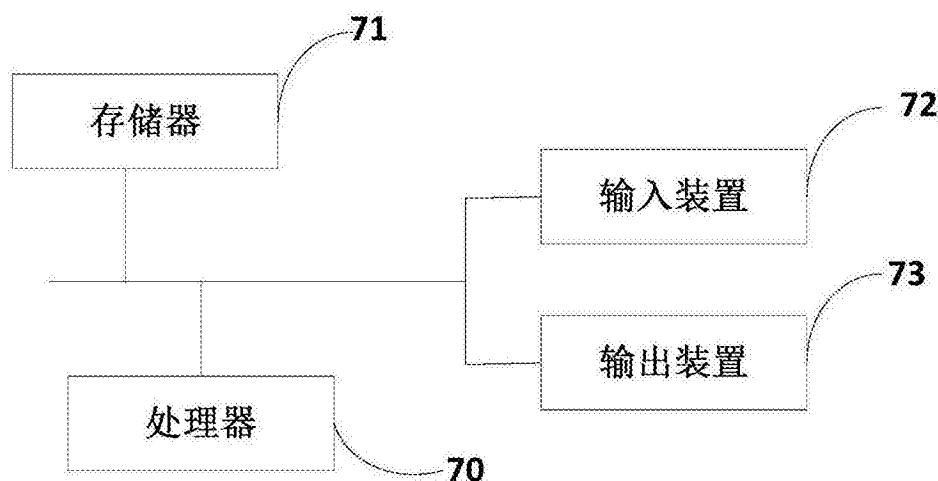


图7