

视频切割方法、装置、计算机设备和存储介质

相关申请的交叉引用

本申请要求于 2018 年 12 月 14 日提交中国专利局，申请号为 201811536818X，申请  
5 名称为“视频切割方法、装置、计算机设备和存储介质”的中国专利申请的优先权，其全部  
内容通过引用结合在本申请中。

技术领域

本申请涉及一种视频切割方法、装置、计算机设备和存储介质。  
10

背景技术

随着多媒体技术的发展，以视频形式进行信息和资源传递的电影、电视、新闻、社交、  
教育和游戏等得到了广泛的应用，如视频聊天、视频会议、视频监控和影视剧等，视频已  
经成为人们工作、学习和生活中的重要部分。  
15 在视频应用中，有需要对视频进行切割处理的场景，如电视新闻截取、录制视频中的  
脱敏处理等。发明人意识到，目前对于视频的切割处理，需要人工进行手动打点标记，确  
定视频切割的时间轴位置，视频切割处理的效率低。

发明内容

20 根据本申请公开的各种实施例，提供一种视频切割方法、装置、计算机设备和存储介  
质。  
一种视频切割方法包括：  
从视频流数据中提取待识别视频数据，并从待识别视频数据中提取影像数据和音频数  
据；  
25 将影像数据输入预设的打点行为识别模型中，得到打点行为识别结果，并将音频数据  
输入预设的打点语音识别模型中，得到打点语音识别结果；  
根据打点行为识别结果、打点语音识别结果和预设的打点触发规则，得到打点识别结  
果；  
当打点识别结果的类型为操作打点时，对待识别视频数据添加切割点标识；及  
30 按照切割点标识将视频流数据进行切割处理，得到视频段数据。  
一种视频切割装置包括：

识别数据提取模块，用于从视频流数据中提取待识别视频数据，并从待识别视频数据中提取影像数据和音频数据；

打点识别处理模块，用于将影像数据输入预设的打点行为识别模型中，得到打点行为识别结果，并将音频数据输入预设的打点语音识别模型中，得到打点语音识别结果；

5 打点结果获取模块，用于根据打点行为识别结果、打点语音识别结果和预设的打点触发规则，得到打点识别结果；

切割标识添加模块，用于当打点识别结果的类型为操作打点时，对待识别视频数据添加切割点标识；及

视频切割模块，用于按照切割点标识将视频流数据进行切割处理，得到视频段数据。

10 一种计算机设备，包括存储器和一个或多个处理器，所述存储器中储存有计算机可读指令，所述计算机可读指令被所述处理器执行时，使得所述一个或多个处理器执行以下步骤：

从视频流数据中提取待识别视频数据，并从待识别视频数据中提取影像数据和音频数据；

15 将影像数据输入预设的打点行为识别模型中，得到打点行为识别结果，并将音频数据输入预设的打点语音识别模型中，得到打点语音识别结果；

根据打点行为识别结果、打点语音识别结果和预设的打点触发规则，得到打点识别结果；

当打点识别结果的类型为操作打点时，对待识别视频数据添加切割点标识；及

20 按照切割点标识将视频流数据进行切割处理，得到视频段数据。

一个或多个存储有计算机可读指令的非易失性计算机可读存储介质，计算机可读指令被一个或多个处理器执行时，使得一个或多个处理器执行以下步骤：

从视频流数据中提取待识别视频数据，并从待识别视频数据中提取影像数据和音频数据；

25 将影像数据输入预设的打点行为识别模型中，得到打点行为识别结果，并将音频数据输入预设的打点语音识别模型中，得到打点语音识别结果；

根据打点行为识别结果、打点语音识别结果和预设的打点触发规则，得到打点识别结果；

当打点识别结果的类型为操作打点时，对待识别视频数据添加切割点标识；及

30 按照切割点标识将视频流数据进行切割处理，得到视频段数据。

本申请的一个或多个实施例的细节在下面的附图和描述中提出。本申请的其它特征和优点将从说明书、附图以及权利要求书变得明显。

## 附图说明

为了更清楚地说明本申请实施例中的技术方案，下面将对实施例中所需要使用的附图作简单地介绍，显而易见地，下面描述中的附图仅仅是本申请的一些实施例，对于本领域普通技术人员来讲，在不付出创造性劳动的前提下，还可以根据这些附图获得其它的附图。

图 1 为根据一个或多个实施例中视频切割方法的应用场景图。

5 图 2 为根据一个或多个实施例中视频切割方法的流程示意图。

图 3 为根据一个或多个实施例中响应打点切割指令的流程示意图。

图 4 为另一个实施例中视频切割方法的流程示意图。

图 5 为根据一个或多个实施例中视频切割装置的结构框图。

图 6 为根据一个或多个实施例中计算机设备的内部结构图。

10

## 具体实施方式

为了使本申请的技术方案及优点更加清楚明白，以下结合附图及实施例，对本申请进行进一步详细说明。应当理解，此处描述的具体实施例仅仅用以解释本申请，并不用于限定本申请。

15 本申请提供的视频切割方法，可以应用于如图 1 所示的应用环境中。录制设备 102 通过网络与服务器 104 通过网络进行通信。录制设备 102 进行视频录制，并将录制的视频流数据发送至服务器 104，服务器 104 从视频流数据中得到的待识别视频数据中提取影像数据和音频数据，并将影像数据和音频数据分别输入对应预设的打点行为识别模型和打点语音识别模型中，再根据得到的打点行为识别结果、打点语音识别结果和预设的打点触发规则获得打点识别结果，当打点识别结果的类型为操作打点时，对待识别视频数据添加切割点标识，最后按照该切割点标识将视频流数据切割，得到视频段数据。

录制设备 102 可以但不限于各种视频录制摄像机，也可以为具有视频录制功能的终端，如个人计算机、笔记本电脑、智能手机、平板电脑和便携式可穿戴设备，服务器 104 可以用独立的服务器或者是多个服务器组成的服务器集群来实现。

25 在其中一个实施例中，如图 2 所示，提供了一种视频切割方法，以该方法应用于图 1 中的服务器 104 为例进行说明，包括以下步骤：

步骤 S201：从视频流数据中提取待识别视频数据，并从待识别视频数据中提取影像数据和音频数据。

30 本实施例中，从视频流数据中提取待识别视频数据。其中，视频流数据为需要进行切割处理的视频数据，可以由录制设备录制得到。例如，对于金融行业的面核过程，视频流数据可以为双录过程中摄像机实时拍摄的视频数据。待识别视频数据为预设识别长度的视频数据，该识别长度根据实际需求设定，可以通过对待识别视频数据进行打点识别，从而添加相应切割点标识。通过预设识别长度的待识别视频数据进行打点识别，可以实现对录制的视频数据进行实时切割，确保视频切割的时效性，提高视频切割效率。

一般地，视频数据由影像和音频两部分组成，而影像和音频两部分均可以进行打点识别。具体地，对待识别视频数据进行打点识别时，从待识别视频数据中提取影像数据和音频数据，以实现同时对待识别视频数据中的影像数据和音频数据分别进行识别处理，从而可以识别出视频影像中是否出现打点行为，或视频音频中是否出现打点语音，实现了影像行为和音频语音的打点识别，提高了打点识别的准确性。

步骤 S203：将影像数据输入预设的打点行为识别模型中，得到打点行为识别结果，并将音频数据输入预设的打点语音识别模型中，得到打点语音识别结果。

从待识别视频数据中提取得到影像数据和音频数据后，分别将影像数据和音频数据输入对应的打点行为识别模型和打点语音识别模型中进行打点识别。其中，打点行为识别模型可以为基于人工神经网络算法，通过训练业务系统的业务人员在对应业务场景下的历史打点行为数据得到，例如可以为鼓掌动作、举手动作、敲击动作等打点行为动作；打点语音识别模型则可以通过训练业务人员的历史打点语音数据得到，例如可以为关键词语音打点，如“第一、第二、第三”等关键词。

本实施例中，一方面将影像数据输入预设的打点行为识别模型中进行打点行为识别，得到打点行为识别结果；另一方面，将音频数据输入预设的打点语音识别模型中进行打点语音识别，得到打点语音识别结果。通过对影像数据和音频数据分别进行打点识别，可以扩展打点操作的多样性，避免业务流程的流畅性，同时确保对视频切割的准确性。

步骤 S205：根据打点行为识别结果、打点语音识别结果和预设的打点触发规则，得到打点识别结果。

得到打点行为识别结果和打点语音识别结果后，综合二者得到打点识别结果。具体地，查询预设的打点触发规则，该打点触发规则根据实际业务需求进行设定。如可以设为对打点行为识别结果和打点语音识别结果取或，即只要打点行为识别结果和打点语音识别结果中任意一个的类型为操作打点，即需要添加切割点标识时，则触发打点，得到的打点识别结果为操作打点；也可以为对打点行为识别结果和打点语音识别结果取并处理，即只有打点行为识别结果和打点语音识别结果的类型同时为操作打点时，触发打点，得到的打点识别结果的类型为操作打点。

步骤 S207：当打点识别结果的类型为操作打点时，对待识别视频数据添加切割点标识。

得到打点识别结果后，判断该打点识别结果的类型，当打点识别结果的类型为操作打点时，表明该待识别视频数据中的影像数据和/或音频数据已触发打点，该待识别视频数据为视频切割位置，对其进行打点处理，具体可以对该待识别视频数据添加切割点标识。其中，切割点标识用于标识视频切割的切割点，在对视频流数据进行切割时，可以直接查找该切割点标识进行切割处理。

在具体实现时，切割点标识可以为切割标签，在对待识别视频数据添加切割点标识时，按照预设的标签添加规则，从该待识别视频数据中确定关键帧，如将待识别视频数据中的

第一帧作为关键帧，并为该关键帧添加切割标签，切割标签可以但不限于包括切割点序号、切割时间值等。

步骤 S209：按照切割点标识将视频流数据进行切割处理，得到视频段数据。

在对视频流数据进行切割处理时，查找视频流数据中的切割点标识，按照该切割点标识进行切割处理，从而将视频流数据拆分，得到各视频段数据。

上述视频切割方法中，从视频流数据中得到的待识别视频数据中提取影像数据和音频数据，并将影像数据和音频数据分别输入对应预设的打点行为识别模型和打点语音识别模型中，再根据得到的打点行为识别结果、打点语音识别结果和预设的打点触发规则获得打点识别结果，当打点识别结果的类型为操作打点时，对待识别视频数据添加切割点标识，最后按照该切割点标识将视频流数据切割，得到视频段数据。在视频切割处理过程中，可以根据待识别视频数据中的影像数据和音频数据进行打点识别并添加切割点标识，不需要人工进行打点操作，提高了视频切割的处理效率。

在一些实施例，从视频流数据中提取待识别视频数据包括：获取视频流数据；确定视频流识别长度；及按照视频流识别长度，从视频流数据中提取待识别视频数据。

对于录制设备 102 直接录制的视频流数据，无法直接进行打点识别处理，需将其拆分成固定识别长度的待识别视频数据，并通过待识别视频数据进行打点识别。本实施例中，在从视频流数据中提取待识别视频数据时，一方面，先获取视频流数据，具体可以直接从录制设备 102 接收到实时录制的视频流数据，也可以从预设的存储器中读取已录制结束的视频流数据。另一方面，确定视频流识别长度，该视频流识别长度根据实际需求进行设定，例如可以根据打点行为识别模型和打点语音识别模型的输入需求进行设置，也可以根据服务器 104 的处理资源进行设定。确定视频流识别长度后，按照该视频流识别长度，从视频流数据中提取待识别视频数据。在具体应用时，可以从视频流数据中依次提取满足视频流识别长度的待识别视频数据，再将提取得到的待识别视频数据进行后续的打点识别处理。

在其中一个实施例中，将影像数据输入预设的打点行为识别模型中，得到打点行为识别结果，并将音频数据输入预设的打点语音识别模型中，得到打点语音识别结果包括：确定待识别视频数据对应所属业务人员的身份标识信息；查询身份标识信息分别对应预设的打点行为识别模型和打点语音识别模型；从影像数据中提取影像特征数据，从音频数据中提取音频特征数据；及将影像特征数据输入打点行为识别模型中，得到打点行为识别结果，将音频特征数据输入打点语音识别模型中，得到打点语音识别结果。

本实施例中，打点行为识别模型和打点语音识别模型均基于业务系统中各业务人员的历史打点数据训练得到。一般地，在业务面核双录过程中，不同业务系统会有不同的打点操作要求，而不同的业务人员也会有不同的打点操作习惯。

具体地，在将影像数据输入预设的打点行为识别模型中，得到打点行为识别结果，并将音频数据输入预设的打点语音识别模型中，得到打点语音识别结果时，先确定待识别视频数据对应所属业务人员的身份标识信息。在应用时，对于各业务服务窗口，均设置有

录制设备 102，可以通过待识别视频数据的来源，即根据录制设备 102 来确定对应所属业务人员，并进一步查询该业务人员对应的身份标识信息。身份标识信息可以但不限于为员工编号、员工姓名等可以唯一识别业务人员的身份信息。确定身份标识信息后，查询与该身份标识信息对应预设的打点行为识别模型和打点语音识别模型，打点行为识别模型和打点语音识别模型分别基于对应业务人员的历史打点行为数据和历史打点语音数据训练得到，打点识别的针对性强，识别准确度高。

得到打点行为识别模型和打点语音识别模型后，一方面，从影像数据中提取影像特征数据，将影像特征数据输入打点行为识别模型中，得到打点行为识别结果。另一方面，从音频数据中提取音频特征数据，并将音频特征数据输入打点语音识别模型中，得到打点语音识别结果。在对影像数据和音频数据进行打点识别时，进行特征提取，过滤无用的冗余信息，得到影像特征数据和音频特征数据，并进行后续的打点识别处理，得到打点行为识别结果和打点语音识别结果。

在其中一个实施例中，在查询身份标识信息分别对应预设的打点行为识别模型和打点语音识别模型之前，还包括：从业务系统中获取历史行为影像数据和历史打点语音数据；分别将历史行为影像数据和历史打点语音数据按照业务人员进行分类，得到各业务人员对应的历史行为影像数据和各业务人员对应的历史打点语音数据；训练各业务人员对应的历史行为影像数据，得到打点行为识别模型；及训练各业务人员对应的历史打点语音数据，得到打点语音识别模型。

在训练打点行为识别模型和打点语音识别模型时，先从业务系统中获取历史行为影像数据和历史打点语音数据。其中，历史行为影像数据可以为业务系统中各业务人员在业务面核过程中双录拍摄到的打点影像数据，例如可以包括鼓掌、举手、双手交叉、点头等打点行为；历史打点语音数据与历史行为影像数据类似，如关键词语句，“第 X 个问题”、“好的，谢谢”等。在具体应用中，各业务人员会有不同的个人习惯，其对应的历史行为影像数据和历史打点语音数据中打点操作的表现也不相同，所以按照业务人员将历史行为影像数据和历史打点语音数据进行分类，为各业务人员构建对应的打点行为识别模型和打点语音识别模型。

具体地，训练各业务人员对应的历史行为影像数据，得到打点行为识别模型；训练各业务人员对应的历史打点语音数据，得到打点语音识别模型。具体实现时，可以将历史行为影像数据划分为训练样本集和测试样本集，通过有监督学习方法训练该训练样本集，得到待测试打点行为模型，再通过测试样本集对待测试打点行为模型进行识别精度测试，在识别精度测试通过后，得到打点行为识别模型。打点语音识别模型的训练过程类同于打点行为识别模型。

在其中一个实施例中，根据打点行为识别结果、打点语音识别结果和预设的打点触发规则，得到打点识别结果包括：查询预设的打点触发规则，打点触发规则包括行为触发规则和语音触发规则；将打点行为识别结果与行为触发规则进行比较，得到行为触发结果；

将打点语音识别结果与语音触发规则进行比较，得到语音触发结果；及根据行为触发结果和语音触发结果，得到打点识别结果。

得到打点行为识别结果和打点语音识别结果后，结合实际业务需求的打点触发规则，得到打点识别结果。具体地，查询预设的打点触发规则，该打点触发规则根据实际业务需求进行设定，具体可以根据业务类型和业务人员的习惯进行设定，如设定为当影像数据中识别到业务人员的鼓掌行为时，或者当音频数据中识别到“第 X 个问题”的关键语句时，认为触发打点。打点触发规则包括行为触发规则和语音触发规则，分别对应于影像数据的打点识别和音频数据的打点识别。

一方面，将打点行为识别结果与行为触发规则进行比较，得到行为触发结果；另一方面将打点语音识别结果与语音触发规则进行比较，得到语音触发结果。最后综合行为触发结果和语音触发结果得到打点识别结果，如可以对行为触发结果和语音触发结果取或运算，即当行为触发结果和语音触发结果中任一类型为操作打点时，即得到的打点识别结果的类型为操作打点，并对待识别视频数据进行切割点标识添加处理。

在其中一个实施例中，如图 3 所示，还包括响应打点切割指令的步骤，具体包括：  
步骤 S301：当接收到打点切割指令时，确定打点切割指令的切割时刻值。

本实施例中，除了对从视频流数据中提取待识别视频数据，对待识别视频数据进行打点识别外，还可以响应外部发送的打点切割指令，实现人工操作打点。具体地，在接收到打点切割指令时，确定该打点切割指令的切割时刻值。其中，打点切割指令可以由外部发送，如业务人员点击相关打点按钮；切割时刻值为打点切割指令的发送时间，反映视频流数据中需要进行打点操作的时间轴位置。

步骤 S303：确定切割时刻值在待识别视频数据中对应的切割视频帧。  
确定打点切割指令的切割时刻值后，从待识别视频数据中确定该切割时刻值对应的切割视频帧。一般地，外部发送打点切割指令时，表明待识别视频数据中该时刻对应的视频帧需要进行操作打点，根据该打点切割指令的切割时刻值可以从待识别视频数据的时间轴确定对应的切割视频帧。

步骤 S305：为切割视频帧添加切割点标识。  
确定切割视频帧后，为该切割视频帧添加切割点标识，切割点标识用于标识视频切割的切割点，在对视频流数据进行切割时，可以直接查找该切割点标识进行切割处理。

步骤 S307：返回按照切割点标识将视频流数据进行切割处理，得到视频段数据。  
添加切割点标识后，返回按照切割点标识将视频流数据进行切割处理的步骤，通过查找视频流数据中的切割点标识，再按照该切割点标识进行切割处理，从而将视频流数据拆分，得到各视频段数据。

本实施例中，在对待识别视频数据的影像数据和音频数据进行打点识别外，还实时接收外部发送的打点切割指令，并按照该打点切割指令进行视频切割处理，实现对外部对视频切割的控制，能够有效扩展视频切割的操作多样性，提高视频切割处理的效率。

在一些实施例中,在得到视频段数据之后,还包括:从视频段数据中提取音频段数据;查询预设的语音识别模型;将音频段数据输入语音识别模型中,得到视频段数据的译文数据;及根据译文数据确定视频段数据对应的业务类型,并将视频段数据存储至业务类型对应的存储位置中。

5       本实施例中,在得到视频流数据经过切割处理的视频段数据后,可以按照各视频段数据的业务类型将其存储至对应的存储位置中。具体地,从视频段数据中提取音频段数据,音频段数据包括视频段数据中的对话数据,根据该音频段数据可以确定该视频段数据对应的业务类型。查询预设的语音识别模型,语音识别模型可以将输入的语音数据进行语音识别,得到对应的译文数据。

10       本实施例中,将音频段数据输入该语音识别模型中,得到视频段数据的译文数据,译文数据可以为文本形式的数据,根据该译文数据可以确定视频段数据对应的业务类型。在具体实现时,可以从译文数据中提取业务关键字,并根据得到的业务关键字匹配对应的业务类型。确定视频段数据对应的业务类型后,将该视频段数据存储至业务类型对应的存储位置中。如可以查询该业务类型对应预设的存储位置,并将视频段数据存储至该存储位置中,从而实现了视频段数据的自动分类存储。

在其中个实施例中,如图4所示,提供了一种视频切割方法,包括:

步骤 S401: 获取视频流数据;

步骤 S402: 确定视频流识别长度;

步骤 S403: 按照视频流识别长度,从视频流数据中提取待识别视频数据;

20       步骤 S404: 从待识别视频数据中提取影像数据和音频数据。

本实施例中,服务器 104 接收录制设备 102 发送的视频流数据,并确定根据实际需求进行设定的视频流识别长度,并按照该视频流识别长度,从视频流数据中依次提取满足视频流识别长度的待识别视频数据,再将提取得到的待识别视频数据进行后续的打点识别处理。

25       步骤 S405: 确定待识别视频数据对应所属业务人员的身份标识信息;

步骤 S406: 查询身份标识信息分别对应预设的打点行为识别模型和打点语音识别模型;

步骤 S407: 从影像数据中提取影像特征数据,从音频数据中提取音频特征数据;

30       步骤 S408: 将影像特征数据输入打点行为识别模型中,得到打点行为识别结果,将音频特征数据输入打点语音识别模型中,得到打点语音识别结果。

得到影像数据和音频数据后,通过待识别视频数据的来源,即根据录制设备 102 来确定对应所属业务人员,并进一步查询该业务人员对应的身份标识信息,身份标识信息为员工编号和/或员工姓名。查询与该身份标识信息对应预设的打点行为识别模型和打点语音识别模型,打点行为识别模型和打点语音识别模型分别基于对应业务人员的历史打点行为数据和历史打点语音数据训练得到,打点识别的针对性强,识别准确度高。一方面,从

35

影像数据中提取影像特征数据，将影像特征数据输入打点行为识别模型中，得到打点行为识别结果。另一方面，从音频数据中提取音频特征数据，并将音频特征数据输入打点语音识别模型中，得到打点语音识别结果。

5 步骤 S409：根据打点行为识别结果、打点语音识别结果和预设的打点触发规则，得到打点识别结果。

得到打点行为识别结果和打点语音识别结果后，结合实际业务需求的打点触发规则，得到打点识别结果。具体可以包括：查询预设的打点触发规则，打点触发规则包括行为触发规则和语音触发规则；将打点行为识别结果与行为触发规则进行比较，得到行为触发结果；将打点语音识别结果与语音触发规则进行比较，得到语音触发结果；根据行为触发结果和语音触发结果，得到打点识别结果。

步骤 S410：打点识别结果的类型为操作打点时，对待识别视频数据添加切割点标识；

步骤 S411：按照切割点标识将视频流数据进行切割处理，得到视频段数据。

得到打点识别结果后，判断其类型，当打点识别结果的类型为操作打点时，表明该待识别视频数据为切割点，对其进行打点处理，具体可以对该待识别视频数据添加切割点标识。通过查找视频流数据中的切割点标识，按照该切割点标识进行切割处理，从而将视频流数据拆分，得到各视频段数据。

步骤 S412：从视频段数据中提取音频段数据；

步骤 S413：查询预设的语音识别模型；

步骤 S414：将音频段数据输入语音识别模型中，得到视频段数据的译文数据；

20 步骤 S415：根据译文数据确定视频段数据对应的业务类型，并将视频段数据存储至业务类型对应的存储位置中。

本实施例中，在得到视频流数据经过切割处理的视频段数据后，可以按照各视频段数据的业务类型将其存储至对应的存储位置中，从而实现了对视频段数据的自动分类存储。

25 应该理解的是，虽然图 2-4 的流程图中的各个步骤按照箭头的指示依次显示，但是这些步骤并不是必然按照箭头指示的顺序依次执行。除非本文中有明确的说明，这些步骤的执行并没有严格的顺序限制，这些步骤可以以其它的顺序执行。而且，图 2-4 中的至少一部分步骤可以包括多个子步骤或者多个阶段，这些子步骤或者阶段并不必然是在同一时刻执行完成，而是可以在不同的时刻执行，这些子步骤或者阶段的执行顺序也不必然是依次进行，而是可以与其它步骤或者其它步骤的子步骤或者阶段的至少一部分轮流或者交替地执行。

在其中一个实施例中，如图 5 所示，提供了一种视频切割装置，包括：识别数据提取模块 501、打点识别处理模块 503、打点结果获取模块 505、切割标识添加模块 507 和视频切割模块 509，其中：

35 识别数据提取模块 501，用于从视频流数据中提取待识别视频数据，并从待识别视频

数据中提取影像数据和音频数据；

打点识别处理模块 503，用于将影像数据输入预设的打点行为识别模型中，得到打点行为识别结果，并将音频数据输入预设的打点语音识别模型中，得到打点语音识别结果；

打点结果获取模块 505，用于根据打点行为识别结果、打点语音识别结果和预设的打点触发规则，得到打点识别结果；

切割标识添加模块 507，用于当打点识别结果的类型为操作打点时，对待识别视频数据添加切割点标识；及

视频切割模块 509，用于按照切割点标识将视频流数据进行切割处理，得到视频段数据。

10 在其中一个实施例中，识别数据提取模块 501 包括视频流获取单元、识别长度确定单元和识别数据提取单元，其中：视频流获取单元，用于获取视频流数据；识别长度确定单元，用于确定视频流识别长度；及识别数据提取单元，用于按照视频流识别长度，从视频流数据中提取待识别视频数据。

15 在其中一个实施例中，打点识别处理模块 503 包括身份标识确定单元、识别模型查询单元、特征数据提取单元和打点识别单元，其中：身份标识确定单元，用于确定待识别视频数据对应所属业务人员的身份标识信息；识别模型查询单元，用于查询身份标识信息分别对应预设的打点行为识别模型和打点语音识别模型；特征数据提取单元，用于从影像数据中提取影像特征数据，从音频数据中提取音频特征数据；及打点识别单元，用于将影像特征数据输入打点行为识别模型中，得到打点行为识别结果，将音频特征数据输入打点语音识别模型中，得到打点语音识别结果。

20 在其中一个实施例中，还包括历史数据获取模块、历史数据分类模块、行为识别模型训练模块和语音识别模型训练模块，其中：历史数据获取模块，用于从业务系统中获取历史行为影像数据和历史打点语音数据；历史数据分类模块，用于分别将历史行为影像数据和历史打点语音数据按照业务人员进行分类，得到各业务人员对应的历史行为影像数据和各业务人员对应的历史打点语音数据；行为识别模型训练模块，用于训练各业务人员对应的历史行为影像数据，得到打点行为识别模型；及语音识别模型训练模块，用于训练各业务人员对应的历史打点语音数据，得到打点语音识别模型。

30 在其中一个实施例中，打点结果获取模块 505 包括触发规则查询单元、行为比较单元、语音比较单元和打点结果获取单元，其中：触发规则查询单元，用于查询预设的打点触发规则，打点触发规则包括行为触发规则和语音触发规则；行为比较单元，用于将打点行为识别结果与行为触发规则进行比较，得到行为触发结果；语音比较单元，用于将打点语音识别结果与语音触发规则进行比较，得到语音触发结果；及打点结果获取单元，用于根据行为触发结果和语音触发结果，得到打点识别结果。

35 在其中一个实施例中，还包括切割指令接收模块、切割帧确定模块、标识添加模块和切割处理模块，其中：切割指令接收模块，用于当接收到打点切割指令时，确定打点切

割指令的切割时刻值；切割帧确定模块，用于确定切割时刻值在待识别视频数据中对应的切割视频帧；标识添加模块，用于为切割视频帧添加切割点标识；及切割处理模块，用于返回按照切割点标识将视频流数据进行切割处理，得到视频段数据。

5 在其中一个实施例中，还包括音频段提取模块、语音识别模型查询模块、译文数据获取模块和视频段存储模块，其中：音频段提取模块，用于从视频段数据中提取音频段数据；语音识别模型查询模块，用于查询预设的语音识别模型；译文数据获取模块，用于将音频段数据输入语音识别模型中，得到视频段数据的译文数据；及视频段存储模块，用于根据译文数据确定视频段数据对应的业务类型，并将视频段数据存储至业务类型对应的存储位置中。

10 关于视频切割装置的具体限定可以参见上文中对于视频切割方法的限定，在此不再赘述。上述视频切割装置中的各个模块可全部或部分通过软件、硬件及其组合来实现。上述各模块可以硬件形式内嵌于或独立于计算机设备中的处理器中，也可以以软件形式存储于计算机设备中的存储器中，以便于处理器调用执行以上各个模块对应的操作。

15 在其中一个实施例中，提供了一种计算机设备，该计算机设备可以是服务器，其内部结构图可以如图 6 所示。该计算机设备包括通过系统总线连接的处理器、存储器和网络接口。其中，该计算机设备的处理器用于提供计算和控制能力。该计算机设备的存储器包括非易失性存储介质、内存储器。该非易失性存储介质存储有操作系统和计算机可读指令。该内存储器为非易失性存储介质中的操作系统和计算机可读指令的运行提供环境。该计算机设备的网络接口用于与外部的终端通过网络连接通信。该计算机可读指令被处理器执行  
20 时以实现一种视频切割方法。

本领域技术人员可以理解，图 6 中示出的结构，仅仅是与本申请方案相关的部分结构的框图，并不构成对本申请方案所应用于其上的计算机设备的限定，具体的计算机设备可以包括比图中所示更多或更少的部件，或者组合某些部件，或者具有不同的部件布置。

25 一种计算机设备，包括存储器和一个或多个处理器，存储器中存储有计算机可读指令，计算机可读指令被处理器执行时实现本申请任意一个实施例中提供的视频切割方法的步骤。

一个或多个存储有计算机可读指令的非易失性存储介质，计算机可读指令被一个或多个处理器执行时，使得一个或多个处理器实现本申请任意一个实施例中提供的视频切割方法的步骤。

30 本领域普通技术人员可以理解实现上述实施例方法中的全部或部分流程，是可以通  
过计算机可读指令来指令相关的硬件来完成，所述的计算机可读指令可存储于一非易失性计算机可读存储介质中，该计算机可读指令在执行时，可包括如上述各方法的实施例的流程。其中，本申请所提供的各实施例中所使用的对存储器、存储、数据库或其它介质的任何引用，均可包括非易失性和/或易失性存储器。非易失性存储器可包括只读存储器  
35 (ROM)、可编程 ROM (PROM)、电可编程 ROM (EPROM)、电可擦除可编程 ROM

(EEPROM) 或闪存。易失性存储器可包括随机存取存储器 (RAM) 或者外部高速缓冲存储器。作为说明而非局限, RAM 以多种形式可得, 诸如静态 RAM (SRAM)、动态 RAM (DRAM)、同步 DRAM (SDRAM)、双数据率 SDRAM (DDRSDRAM)、增强型 SDRAM (ESDRAM)、同步链路 (Synchlink) DRAM (SLDRAM)、存储器总线 (Rambus) 直接  
5 RAM (RDRAM)、直接存储器总线动态 RAM (DRDRAM)、以及存储器总线动态 RAM (RDRAM) 等。

以上实施例的各技术特征可以进行任意的组合, 为使描述简洁, 未对上述实施例中的各个技术特征所有可能的组合都进行描述, 然而, 只要这些技术特征的组合不存在矛盾, 都应当认为是本说明书记载的范围。

以上所述实施例仅表达了本申请的几种实施方式, 其描述较为具体和详细, 但并不能因此而理解为对发明专利范围的限制。应当指出的是, 对于本领域的普通技术人员来说, 在不脱离本申请构思的前提下, 还可以做出若干变形和改进, 这些都属于本申请的保护范围。因此, 本申请专利的保护范围应以所附权利要求为准。