

权利要求书

1、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，并从所述待识别视频数据中提取影像数据和音频数据；

5 将所述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，并将所述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；

根据所述打点行为识别结果、所述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；

10 当所述打点识别结果的类型为操作打点时，对所述待识别视频数据添加切割点标识；及

按照所述切割点标识将所述视频流数据进行切割处理，得到视频段数据。

2、根据权利要求 1 所述的方法，其特征在于，所述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；

15 确定视频流识别长度；及

按照所述视频流识别长度，从所述视频流数据中提取待识别视频数据。

3、根据权利要求 1 所述的方法，其特征在于，所述将所述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，并将所述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：

20 确定所述待识别视频数据对应所属业务人员的身份标识信息；

查询所述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；

从所述影像数据中提取影像特征数据，从所述音频数据中提取音频特征数据；及

25 将所述影像特征数据输入所述打点行为识别模型中，得到打点行为识别结果，将所述音频特征数据输入所述打点语音识别模型中，得到打点语音识

别结果。

4、根据权利要求3所述的方法，其特征在于，在所述查询所述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，所述方法还包括：

5 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；

分别将所述历史行为影像数据和所述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；

10 训练所述各业务人员对应的历史行为影像数据，得到所述打点行为识别模型；及

训练所述各业务人员对应的历史打点语音数据，得到所述打点语音识别模型。

15 5、根据权利要求1所述的方法，其特征在于，所述根据所述打点行为识别结果、所述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，所述打点触发规则包括行为触发规则和语音触发规则；

将所述打点行为识别结果与所述行为触发规则进行比较，得到行为触发结果；

20 将所述打点语音识别结果与所述语音触发规则进行比较，得到语音触发结果；及

根据所述行为触发结果和所述语音触发结果，得到打点识别结果。

6、根据权利要求5所述的方法，其特征在于，所述根据所述行为触发结果和所述语音触发结果，得到打点识别结果，包括：

25 对所述行为触发结果和所述语音触发结果进行取或运算，得到打点识别结果。

7、根据权利要求1所述的方法，其特征在于，所述当所述打点识别结果

的类型为操作打点时，对所述待识别视频数据添加切割点标识，包括：

确定所述打点识别结果的类型；

当所述打点识别结果的类型为操作打点时，查询预设的标签添加规则；

根据所述标签添加规则，从所述待识别视频数据中确定关键帧，并为所

5 述关键帧添加切割标签，所述切割点标识包括所述切割标签。

8、根据权利要求 1 至 7 任意一项所述的方法，其特征在于，还包括：

当接收到打点切割指令时，确定所述打点切割指令的切割时刻值；

确定所述切割时刻值在所述待识别视频数据中对应的切割视频帧；

为所述切割视频帧添加切割点标识；及

10 返回所述按照所述切割点标识将所述视频流数据进行切割处理，得到视频
频段数据。

9、根据权利要求 8 所述的方法，其特征在于，在所述得到视频段数据之
后，所述方法还包括：

从所述视频段数据中提取音频段数据；

15 查询预设的语音识别模型；

将所述音频段数据输入所述语音识别模型中，得到所述视频段数据的译
文数据；及

根据所述译文数据确定所述视频段数据对应的业务类型，并将所述视频
段数据存储至所述业务类型对应的存储位置中。

20 10、根据权利要求 9 所述的方法，其特征在于，所述根据所述译文数据
确定所述视频段数据对应的业务类型，并将所述视频段数据存储至所述业务
类型对应的存储位置中，包括：

从所述译文数据中提取业务关键字；

根据所述业务关键字确定所述视频段数据对应的业务类型；

25 查询所述业务类型对应预设的存储位置；

将所述视频段数据存储至所述存储位置中。

11、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，并从所述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，用于将所述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，并将所述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，用于根据所述打点行为识别结果、所述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，用于当所述打点识别结果的类型为操作打点时，对所述待识别视频数据添加切割点标识；及

10 视频切割模块，用于按照所述切割点标识将所述视频流数据进行切割处理，得到视频段数据。

12、根据权利要求 11 所述的装置，其特征在于，所述识别数据提取模块，包括：

视频流获取单元，用于获取视频流数据；

15 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，用于按照所述视频流识别长度，从所述视频流数据中提取待识别视频数据。

13、根据权利要求 11 所述的装置，其特征在于，所述打点识别处理模块，包括：

20 身份标识确定单元，用于确定所述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，用于查询所述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；

25 特征数据提取单元，用于从所述影像数据中提取影像特征数据，从所述音频数据中提取音频特征数据；及

打点识别单元，用于将所述影像特征数据输入所述打点行为识别模型中，得到打点行为识别结果，将所述音频特征数据输入所述打点语音识别模型中，

得到打点语音识别结果。

14、根据权利要求 13 所述的装置，其特征在于，所述装置还包括：

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；

5 历史数据分类模块，用于分别将所述历史行为影像数据和所述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，用于训练所述各业务人员对应的历史行为影像数据，得到所述打点行为识别模型；及

10 语音识别模型训练模块，用于训练所述各业务人员对应的历史打点语音数据，得到所述打点语音识别模型。

15、一种计算机设备，包括存储器及一个或多个处理器，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一个或多个处理器执行时，使得所述一个或多个处理器执行以下步骤：

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

将所述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，并将所述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；

20 根据所述打点行为识别结果、所述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；

当所述打点识别结果的类型为操作打点时，对所述待识别视频数据添加切割点标识；及

按照所述切割点标识将所述视频流数据进行切割处理，得到视频段数据。

25 16、根据权利要求 15 所述的计算机设备，其特征在于，所述处理器执行所述计算机可读指令时还执行以下步骤：

获取视频流数据；

确定视频流识别长度；及

按照所述视频流识别长度，从所述视频流数据中提取待识别视频数据。

17、根据权利要求 15 所述的计算机设备，其特征在于，所述处理器执行所述计算机可读指令时还执行以下步骤：

5 确定所述待识别视频数据对应所属业务人员的身份标识信息；

查询所述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；

从所述影像数据中提取影像特征数据，从所述音频数据中提取音频特征数据；及

10 将所述影像特征数据输入所述打点行为识别模型中，得到打点行为识别结果，将所述音频特征数据输入所述打点语音识别模型中，得到打点语音识别结果。

18、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，所述计算机可读指令被一个或多个处理器执行时，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，并从所述待识别视频数据中提取影像数据和音频数据；

20 将所述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，并将所述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；

根据所述打点行为识别结果、所述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；

当所述打点识别结果的类型为操作打点时，对所述待识别视频数据添加切割点标识；及

25 按照所述切割点标识将所述视频流数据进行切割处理，得到视频段数据。

19、根据权利要求 18 所述的存储介质，其特征在于，所述计算机可读指令被所述处理器执行时还执行以下步骤：

获取视频流数据；

确定视频流识别长度；及

按照所述视频流识别长度，从所述视频流数据中提取待识别视频数据。

- 20、根据权利要求 18 所述的存储介质，其特征在于，所述计算机可读指令被所述处理器执行时还执行以下步骤：

确定所述待识别视频数据对应所属业务人员的身份标识信息；

查询所述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；

- 从所述影像数据中提取影像特征数据，从所述音频数据中提取音频特征数据；及

将所述影像特征数据输入所述打点行为识别模型中，得到打点行为识别结果，将所述音频特征数据输入所述打点语音识别模型中，得到打点语音识别结果。