

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2024 年 7 月 25 日 (25.07.2024)



(10) 国际公布号
WO 2024/152802 A1

(51) 国际专利分类号:

H04N 21/44 (2011.01) **G10L 15/26** (2006.01)
H04N 21/439 (2011.01) **G10L 15/18** (2013.01)

Beijing 100028 (CN)。冀佳玉 (JI, Jiayu); 中国北京市朝阳区七圣中街 12 号院融中心 B1 小邮局, Beijing 100028 (CN)。

(21) 国际申请号: PCT/CN2023/138239

(22) 国际申请日: 2023 年 12 月 12 日 (12.12.2023)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:

202310107566.3 2023 年 1 月 19 日 (19.01.2023) CN

(71) 申请人: 北京字跳网络技术有限公司
(BEIJING ZITIAO NETWORK TECHNOLOGY CO., LTD.) [CN/CN]; 中国北京市海淀区紫金数码园 4 号楼 2 层 0207, Beijing 100190 (CN)。

(72) 发明人: 曾祥瑞 (ZENG, Xiangrui); 中国北京市朝阳区七圣中街 12 号院融中心 B1 小邮局,

(74) 代理人: 北京市金杜律师事务所 (KING & WOOD MALLESONS); 中国北京市朝阳区东三环中路 1 号环球金融中心办公楼东楼 20 层, Beijing 100020 (CN)。

(81) 指定国(除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD,

(54) Title: VIDEO EDITING METHOD AND APPARATUS, ELECTRONIC DEVICE, AND STORAGE MEDIUM

(54) 发明名称: 视频的编辑方法、装置、电子设备和存储介质

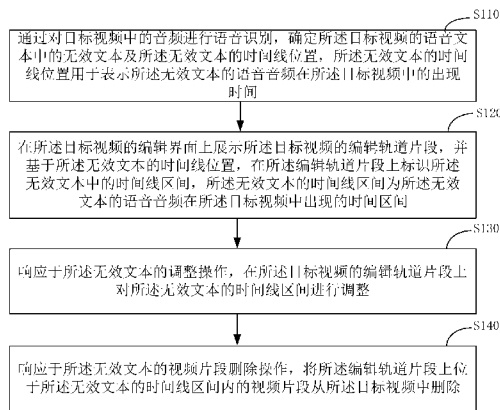


图 1

S110 Perform speech recognition on audio in a target video to determine an invalid text in a speech text of the target video and a timeline position of the invalid text, the timeline position of the invalid text being used for representing the time of appearance of speech audio of the invalid text in the target video
S120 Display an editing track clip of the target video on an editing interface of the target video, and mark a timeline interval of the invalid text on the editing track clip on the basis of the timeline position of the invalid text, the timeline interval of the invalid text being a time interval of appearance of the speech audio of the invalid text in the target video
S130 In response to an adjustment operation on the invalid text, adjust the timeline interval of the invalid text on the editing track clip of the target video
S140 In response to a video clip deletion operation on the invalid text, delete a video clip in the timeline interval of the invalid text on the editing track clip from the target video

(57) Abstract: Embodiments of the present disclosure provide a video editing method and apparatus, an electronic device, and a storage medium. The method comprises: performing speech recognition on audio in a target video to determine an invalid text in a speech text of the target video and a timeline position of the invalid text; displaying an editing track clip of the target video on an editing interface of the target video, and marking a timeline interval of the invalid text on the editing track clip on the basis of the timeline position of the invalid text; in response to an adjustment operation on the invalid text, adjusting the timeline interval of the invalid text on the editing track clip of the target video; and in response to a video clip deletion operation on the invalid text, deleting a video clip in the timeline interval of the invalid text on the editing track clip from the target video. According to the method, an invalid clip can be deleted in response to the adjustment operation and video clip deletion operation on the invalid text by a user, thereby improving the efficiency of video editing.



SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ,
UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW。

(84) 指定国(除另有指明，要求每一种可提供的地区
保护): ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ,
NA, RW, SC, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚
(AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE,
BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR,
HU, IE, IS, IT, LT, LU, LV, MC, ME, MK, MT, NL, NO,
PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF,
CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN,
TD, TG)。

本国际公布:
— 包括国际检索报告(条约第21条(3))。

(57) 摘要: 本公开实施例提供了一种视频的编辑方法、装置、电子设备和存储介质，方法包括：通过对目标视频中的音频进行语音识别，确定目标视频的语音文本中的无效文本及无效文本的时间线位置；在目标视频的编辑界面上展示目标视频的编辑轨道片段，并基于无效文本的时间线位置，在编辑轨道片段上标识无效文本的时间线区间；响应于无效文本的调整操作，在目标视频的编辑轨道片段上对无效文本的时间线区间进行调整；响应于无效文本的视频片段删除操作，将编辑轨道片段上位于无效文本的时间线区间内的视频片段从目标视频中删除。该方法能够响应于用户对无效文本的调整操作和视频片段删除操作，对无效片段进行删除，从而提升了视频剪辑的效率。

视频的编辑方法、装置、电子设备和存储介质

本申请要求于 2023 年 01 月 19 日提交中国专利局、申请号为 202310107566.3、发明名称为“视频的编辑方法、装置、电子设备和存储介质”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

本公开实施例涉及计算机技术领域，尤其涉及一种视频的编辑方法、装置、电子设备和存储介质。

背景技术

随着科技的快速发展，视频类应用程序应运而生，如用户可以拍摄并上传口播视频至视频类应用程序，其中，出镜口播或音频承载了口播视频的核心内容，故如何对口播视频进行更好地剪辑是当前亟需解决的问题。

口播视频的剪辑主要需要处理掉视频中的停顿、磕巴、口水词等用户不需要的片段，现有的剪辑方法一般通过视频中音频的波形来识别无效片段，但该方法需要耗费大量的时间，剪辑效率较低。

发明内容

本公开实施例提供一种视频的编辑方法、装置、电子设备和存储介质，以提升视频的剪辑效率。

第一方面，本公开实施例提供了一种视频的编辑方法，包括：

通过对目标视频中的音频进行语音识别，确定所述目标视频的语音文本中的无效文本及所述无效文本的时间线位置，所述无效文本的时间线位置用于表示所述无效文本的语音音频在所述目标视频中的出现时间；

在所述目标视频的编辑界面上展示所述目标视频的编辑轨道片

段，并基于所述无效文本的时间线位置，在所述编辑轨道片段上标识所述无效文本的时间线区间，所述无效文本的时间线区间为所述无效文本的语音音频在所述目标视频中出现的的时间区间；

5 响应于所述无效文本的调整操作，在所述目标视频的编辑轨道片段上对所述无效文本的时间线区间进行调整；

响应于所述无效文本的视频片段删除操作，将所述编辑轨道片段上位于所述无效文本的时间线区间内的视频片段从所述目标视频中删除。

10 第二方面，本公开实施例还提供了一种视频的编辑装置，包括：
文本确定模块，用于通过对目标视频中的音频进行语音识别，确定所述目标视频的语音文本中的无效文本及所述无效文本的时间线位置，所述无效文本的时间线位置用于表示所述无效文本的语音音频在所述目标视频中的出现时间；

15 区间标识模块，用于在所述目标视频的编辑界面上展示所述目标视频的编辑轨道片段，并基于所述无效文本的时间线位置，在所述编辑轨道片段上标识所述无效文本的时间线区间，所述无效文本的时间线区间为所述无效文本的语音音频在所述目标视频中出现的的时间区间；

20 区间调整模块，用于响应于所述无效文本的调整操作，在所述目标视频的编辑轨道片段上对所述无效文本的时间线区间进行调整；

片段删除模块，用于响应于所述无效文本的视频片段删除操作，将所述编辑轨道片段上位于所述无效文本的时间线区间内的视频片段从所述目标视频中删除。

25 第三方面，本公开实施例还提供了一种电子设备，包括：

一个或多个处理器；

存储器，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本公开实施例所述的视频的编辑方法。

第四方面，本公开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公开实施例所述的视频的编辑方法。

本公开实施例提供了一种视频的编辑方法、装置、电子设备和存储介质，所述方法包括：通过对目标视频中的音频进行语音识别，确定所述目标视频的语音文本中的无效文本及所述无效文本的时间线位置，所述无效文本的时间线位置用于表示所述无效文本的语音音频在所述目标视频中的出现时间；在所述目标视频的编辑界面上展示所述目标视频的编辑轨道片段，并基于所述无效文本的时间线位置，在所述编辑轨道片段上标识所述无效文本的时间线区间，所述无效文本的时间线区间为所述无效文本的语音音频在所述目标视频中出现的时间区间；响应于所述无效文本的调整操作，在所述目标视频的编辑轨道片段上对所述无效文本的时间线区间进行调整；响应于所述无效文本的视频片段删除操作，将所述编辑轨道片段上位于所述无效文本的时间线区间内的视频片段从所述目标视频中删除。利用上述技术方案，通过对目标视频中的音频进行语音识别，能够对目标视频中的无效文本以及无效文本的位置进行确定，进而通过在目标视频的编辑轨道片段上标识无效文本的时间线区间，能够响应于用户对无效文本的调整操作和视频片段删除操作，对无效片段进行删除，从而完成对视频的剪辑，提升了视频剪辑的效率。

附图说明

结合附图并参考以下具体实施方式，本公开各实施例的上述和其他特征、优点及方面将变得更加明显。贯穿附图中，相同或相似的附图标记表示相同或相似的元素。应当理解附图是示意性的，原件和元素不一定按照比例绘制。

图1为本公开实施例提供了一种视频的编辑方法的流程示意图；

图2为本公开实施例提供了一种触发语音识别的示意图；

图3为本公开实施例提供了一种编辑界面的示意图；

图 4 为本公开实施例提供的另一种视频的编辑方法的流程示意图；

图 5 为本公开实施例提供的另一种编辑界面的示意图；

图 6 为本公开实施例提供的一种更新目标文本的示意图；

5 图 7 为本公开实施例提供的另一种更新目标文本的示意图；

图 8 为本公开实施例提供的另一种更新目标文本的示意图；

图 9 为本公开实施例提供的一种视频的编辑装置的结构示意图；

图 10 为本公开实施例提供的一种电子设备的结构示意图。

10 具体实施方式

下面将参照附图更详细地描述本公开的实施例。虽然附图中显示了本公开的某些实施例，然而应当理解的是，本公开可以通过各种形式来实现，而且不应该被解释为限于这里阐述的实施例，相反提供这些实施例是为了更加透彻和完整地理解本公开。应当理解的是，本公开的附图及实施例仅用于示例性作用，并非用于限制本公开的保护范围。

应当理解，本公开的方法实施方式中记载的各个步骤可以按照不同的顺序执行，和/或并行执行。此外，方法实施方式可以包括附加的步骤和/或省略执行示出的步骤。本公开的范围在此方面不受限制。

本文使用的术语“包括”及其变形是开放性包括，即“包括但不限于”。术语“基于”是“至少部分地基于”。术语“一个实施例”表示“至少一个实施例”；术语“另一实施例”表示“至少一个另外的实施例”；术语“一些实施例”表示“至少一些实施例”。其他术语的相关定义将在下文描述中给出。

需要注意，本公开中提及的“第一”、“第二”等概念仅用于对不同的装置、模块或单元进行区分，并非用于限定这些装置、模块或单元所执行的功能的顺序或者相互依存关系。

需要注意，本公开中提及的“一个”、“多个”的修饰是示意

性而非限制性的，本领域技术人员应当理解，除非在上下文另有明确指出，否则应该理解为“一个或多个”。

本公开实施方式中的多个装置之间所交互的消息或者信息的名称仅用于说明性的目的，而并不是用于对这些消息或信息的范围进行限制。

图 1 为本公开实施例提供的一种视频的编辑方法的流程示意图，该方法可适用于对视频进行编辑的情况，该方法可以由视频的编辑装置来执行，其中该装置可由软件和/或硬件实现，并一般集成在电子设备上，在本实施例中电子设备包括但不限于：手机或电脑等设备。

如图 1 所示，本公开实施例提供的一种视频的编辑方法，包括如下步骤：

S110、通过对目标视频中的音频进行语音识别，确定所述目标视频的语音文本中的无效文本及所述无效文本的时间线位置，所述无效文本的时间线位置用于表示所述无效文本的语音音频在所述目标视频中的出现时间。

目标视频可以是指需要进行编辑的视频，如目标视频可以为用户上传的视频，也可以为用户在线拍摄的视频。在本实施例中，目标视频中包含音频，语音文本可以认为是与目标视频的音频对应的文本信息，其可以通过对目标视频的音频进行语音识别得到。示例性的，语音文本中包括无效文本，无效文本可以包括目标视频中停顿、重复、口水词等用户不需要的片段对应的文本；无效文本的时间线位置可以用于表示无效文本的语音音频在目标视频中的出现时间。

在一个实施例中，目标视频为连续视频，如目标视频可以为某连续且完整视频，也可以为某完整视频的某一片段。

在本实施例中，可以通过对目标视频中的音频进行语音识别，来确定目标视频的语音文本中的无效文本及无效文本的时间线位

置，以进行后续的编辑步骤。其中，本步骤不对语音识别的具体方式以及触发语音识别的时机进行限定，如可以在界面中触发某识别控件来触发对目标视频中音频的语音识别，从而能够自动识别出目标视频中无效文本及无效文本的时间线位置，识别控件的位置及样式不限，可以根据实际页面的情况进行设置。

在一个实施方式中，无效文本的识别条件为：默认静音时长大于预设时长（如 120ms 等）的片段是停顿片段；通过复用预设算法逻辑来识别无效词（如重复与语气词），该预设算法可以配置于当前应用程序内，也可以配置于服务器中。

图 2 为本公开实施例提供的一种触发语音识别的示意图，如图 2 所示，在选中某视频（如目标视频）后，可以在点击右键的弹出窗口中触发控件 1，即可触发对目标视频中音频的语音识别。还可以点击当前界面中的快捷控件 2 实现对目标视频中音频的语音识别。

S120、在所述目标视频的编辑界面上展示所述目标视频的编辑轨道片段，并基于所述无效文本的时间线位置，在所述编辑轨道片段上标识所述无效文本的时间线区间，所述无效文本的时间线区间为所述无效文本的语音音频在所述目标视频中出现的时间区间。

编辑界面可以用于对目标视频中的片段进行编辑。目标视频的编辑轨道片段可以理解为与目标视频对应的编辑轨道片段，如起点与目标视频的起点对应、终点与目标视频的终点对应的编辑轨道片段。示例性的，当目标视频为连续且完整视频时，该编辑轨道片段可以为该连续且完整视频的编辑轨道；当目标视频为某完整视频的一个片段时，该编辑轨道片段可以为该完整视频的编辑轨道中与目标视频对应的轨道片段。可选的，所述编辑轨道片段为所述目标视频的音频轨道片段，音频轨道片段可以是指与目标视频对应的音频轨道片段。无效文本的时间线区间可以认为是无效文本的语音音频在目标视频中出现的时间区间。

在一个实施例中，当对目标视频中的音频进行语音识别确定目标视频的无效文本及无效文本的时间线位置后，可以在目标视频的

编辑界面上对目标视频的编辑轨道片段进行展示，并基于无效文本的时间线位置，在编辑轨道片段上标识无效文本的时间线区间。其中，展示编辑轨道片段以及标识无效文本的时间线区间的方式不限，可以由相关人员预先根据编辑界面的实际情况进行设置，如可以采用预设标识在目标视频的编辑轨道片段上标识无效文本中无效文本的时间线区间，该预设标识具体可以根据需要进行设置，示例性的，该预设标识可以为矩形框，此时，可以采用矩形框在目标视频的编辑轨道片段上框选目标视频中无效文本的时间线区间；或者，无效文本的时间线区间在编辑轨道片段中可以标识为第一展示样式，而其他文本的时间线区间在编辑轨道片段中标识为第二展示样式，第一展示样式区别于第二展示样式，如第一展示样式和第二展示样式可以以不同颜色进行区分，也可以通过不同的线条进行区分等。

在一个实施方式中，可以默认在编辑轨道片段上标识识别得到的所有无效文本的时间线区间，以区分无效文本的时间线区间与其他文本的时间线区间；也可以默认在编辑轨道片段上标识设定个数的无效文本的时间线区间，设定个数可以由相关人员预先进行设置。用户可以根据需要对无效文本的时间线区间进行调整。

图 3 为本公开实施例提供的一种编辑界面的示意图，如图 3 所示，目标视频的编辑界面上可以展示有目标视频的编辑轨道片段 3，并基于上步骤识别出的无效文本的时间线位置，在编辑轨道片段 3 上对无效文本的时间线区间进行标识，如图 3 中以黑色的矩形框对无效文本的时间线区间 4 进行了标识。

S130、响应于所述无效文本的调整操作，在所述目标视频的编辑轨道片段上对所述无效文本的时间线区间进行调整。

无效文本的调整操作可以是指用于触发调整无效文本的时间线区间的操作，本实施例中无效文本的调整操作的具体操作方式不限，如无效文本的调整操作可以为点击编辑界面中某一控件的操作或者作用于编辑轨道片段上的预设出发操作，也可以为在编辑界面中进行预设手势的操作，预设手势可为预先设定的手势，如预设手势可

为向右滑动的手势等。

本实施例不对调整操作的触发位置进行限定，如可以作用在无效文本的时间线区间，也可以作用在其他位置，只要能触发调整无效文本的时间线区间即可。

- 5 具体的，本实施例可以响应于无效文本的调整操作，在目标视频的编辑轨道片段上对无效文本的时间线区间进行调整，具体调整时间线区间的方式可以根据调整操作的不同而有所区别，如当无效文本的调整操作为触发某控件以使无效文本的状态由选中状态变为未选中状态时，可以在目标视频的编辑轨道片段上对无效文本的时间线区间的数量进行调整，如在编辑轨道片段上增加标识某文本的时间线区间，以将此文本调整为无效文本，或者，在编辑轨道片段上取消标识某文本的时间区间，以将此文本调整为非无效文本。
- 10

- S140、响应于所述无效文本的视频片段删除操作，将所述编辑轨道片段上位于所述无效文本的时间线区间内的视频片段从所述目标视频中删除。
- 15

无效文本的视频片段删除操作可以是指用于触发删除编辑轨道片段上无效文本对应的视频片段的操作，例如无效文本的视频片段删除操作可以为点击编辑界面中某删除控件的操作等。

- 本步骤可以响应于无效文本的视频片段删除操作，将编辑轨道片段上位于无效文本的时间线区间内的视频片段从目标视频中删除，具体删除的方式不作限定，如可以根据用户需要保留或去除编辑轨道片段上无效文本的时间线区间对应的痕迹。如图3所示，用户触发控件5即可以认为是视频片段删除操作，在用户触发控件5后，电子设备可以响应于无效文本的视频片段删除操作，将无效文本的时间线区间6对应的视频片段从目标视频中删除。
- 20
- 25

在一个实施例中，所述将所述编辑轨道片段上位于所述无效文本的时间线区间内的视频片段从所述目标视频中删除，包括：

将所述编辑轨道片段上位于所述无效文本的时间线区间内的视频片段从所述目标视频中删除，并以删除的所述视频片段为分割点，

将所述目标视频分割为多个独立的视频片段；或者

将所述编辑轨道片段上位于所述无效文本的时间线区间内的视频片段从所述目标视频中删除，并将删除所述视频片段后的目标视频整合为一个视频。

- 5 在一个实施方式中，可以将编辑轨道片段上位于无效文本的时间线区间内的视频片段从目标视频中删除，并以删除的视频片段为分割点，将目标视频分割为多个独立的视频片段，在此基础上，能够便于用户根据需要对各独立的视频片段进行操作。

- 10 在一个实施方式中，可以将编辑轨道片段上位于无效文本的时间线区间内的视频片段从目标视频中删除，并将删除视频片段后的目标视频整合为一个视频。在此基础上，实现了删除目标视频中无效片段的操作，提高了目标视频的质量。无效片段即为无效文本的时间线区间对应的片段，换言之，编辑轨道片段上所标识的时间线区间在目标视频的语音文本中所对应的文本。

- 15 本公开实施例提供的一种视频的编辑方法，通过对目标视频中的音频进行语音识别，确定所述目标视频的语音文本中的无效文本及所述无效文本的时间线位置，所述无效文本的时间线位置用于表示所述无效文本的语音音频在所述目标视频中的出现时间；在所述目标视频的编辑界面上展示所述目标视频的编辑轨道片段，并基于
20 所述无效文本的时间线位置，在所述编辑轨道片段上标识所述无效文本的时间线区间，所述无效文本的时间线区间为所述无效文本的语音音频在所述目标视频中出现的时间区间；响应于所述无效文本的调整操作，在所述目标视频的编辑轨道片段上对所述无效文本的时间线区间进行调整；响应于所述无效文本的视频片段删除操作，
25 将所述编辑轨道片段上位于所述无效文本的时间线区间内的视频片段从所述目标视频中删除。利用该方法，通过对目标视频中的音频进行语音识别，能够对目标视频中的无效文本以及无效文本的位置进行确定，进而通过在目标视频的编辑轨道片段上标识无效文本的时间线区间，能够响应于用户对无效文本的调整操作和视频片段删

除操作，对无效片段进行删除，从而完成对视频的剪辑，提升了视频剪辑的效率。

在一个实施例中，所述方法还包括：

在所述编辑界面的视频区域播放所述目标视频；

5 当所述目标视频播放至第二时间线区间的起点位置时，将所述目标视频的播放进度调整至第一时间节点，并从所述第一时间节点继续播放所述目标视频，其中，所述第二时间线区间为所述无效文本的时间线区间，所述第一时间节点为所述第二时间线区间的终点位置在所述目标视频中所对应的时间节点。

10 其中，视频区域可以认为是编辑界面中的某个区域，用于播放目标视频，视频区域的具体位置及大小可以由相关人员根据编辑界面的情况进行设置。第二时间线区间可以是任一无效文本的时间线区间，第一时间节点可以理解为第二时间线区间的终点位置在目标视频中所对应的时间节点。

15 在一个实施方式中，可以在编辑界面的视频区域播放目标视频，播放目标视频的时机不限，如在显示编辑界面时可以自动对目标视频进行播放；也可以在显示编辑界面后，当用户触发编辑界面中的某播放控件时，在编辑界面的视频区域播放目标视频；还可以在显示编辑界面的预设时长后，自动对目标视频进行播放等，本实施例
20 对此不作限定。

 当目标视频播放至第二时间线区间的起点位置时，本实施例可以自动将目标视频的播放进度调整至第一时间节点，并从第一时间节点继续播放目标视频。在此基础上，通过播放目标视频时自动跳过无效片段，能够使用户预先观看到编辑后的视频，以对目标视频
25 进行更好地编辑。

 在一个实施例中，在所述编辑界面的视频区域播放所述目标视频之后，还包括：

 响应于对第五文本的触发操作，将所述目标视频的播放进度调整至第二时间节点，并从所述第二时间节点继续播放所述目标视频，

所述第五文本显示于所述编辑界面的文本区域内，所述第五文本为无效文本或非无效文本，所述第二时间节点为所述第五文本的起点位置在所述目标视频中所对应的时间节点。

在本实施例中，第五文本可以为用户所触发的文本，其可以为无效文本或非无效文本。非无效文本即为除无效文本之外的文本，如目标视频的语音文本中的有效文本。第五文本可以显示于编辑界面的文本区域内，文本区域可以认为是编辑界面中的某个区域，用于显示目标视频的语音文本，如用于显示目标视频的语音文本中的第五文本，文本区域的具体位置及大小可以由相关人员根据编辑界面的情况进行设置。

在一个实施方式中，文本区域中的无效文本自成一段，如可以将“停顿”文本单独成行，同时展示停顿时长。

对第五文本的触发操作可以用于触发将目标视频的播放进度调整第二时间节点，如对第五文本的触发操作可以为点击第五文本的操作；第二时间节点即为第五文本的起点位置在目标视频中所对应的时间节点。

具体的，在编辑界面的视频区域播放目标视频之后，还可以响应于对第五文本的触发操作，将目标视频的播放进度调整至第二时间节点，并从第二时间节点继续播放目标视频。在此基础上，通过响应于第五文本的触发操作，实现了对目标视频播放进度的调整，以此实现在目标视频中对相应文本的定位，便于用户观看相应文本所对应的视频片段。

在一个实施例中，所述编辑轨道片段上还展示有时间标识，在所述编辑界面的视频区域播放所述目标视频之后，还包括：

响应于对所述目标标识的位置调整操作，确定所述目标标识的目标位置所位于的第三时间线区间，所述目标位置为所述位置调整操作执行完毕时所述目标标识在所述编辑轨道片段上的位置；

如果所述第三时间线区间为所述无效文本的时间线区间，则将所述目标视频的播放进度调整至第三时间节点，并从所述第三时间

节点继续播放所述目标视频，其中，所述第三时间节点为所述第三时间线区间的起点位置或终点位置在所述目标视频中所对应的时间节点，或者，接收到所述位置进度调整操作之前所述目标视频最后播放至的时间节点；

- 5 如果所述第三时间线区间不为所述无效文本的时间线区间，则将所述目标视频的播放进度调整至第四时间节点，并从所述第四时间节点继续播放所述目标视频，其中，所述第四时间节点为所述目标位置在所述目标视频中所对应的时间节点。

10 时间标识用于在目标视频的编辑轨道片段中表征目标视频当前所播放至的时间节点，换言之，此时间标识可用于指示目标视频当前的播放进度对应的时间节点。当前应用程序在编辑界面的视频区域播放目标视频时，可以依据目标视频的播放进度，实时更新目标标识在目标视频的编辑轨道片段中的位置。目标标识的位置调整操作则可以理解为调整目标标识位置的操作，如目标标识的位置调整

15 操作可以为点击编辑轨道片段中的某一位置 a 以将目标标识调整至该位置 a 的操作，目标标识的位置调整操作也可以为拖拽目标标识以将目标标识调整至位置 a 的操作等。目标位置可以认为是位置调整操作执行完毕时目标标识在编辑轨道片段上的位置，第三时间线区间即为目标标识的目标位置所位于的时间线区间。

20 第三时间节点可以是第三时间线区间的起点位置在目标视频中所对应的时间节点，或第三时间线区间的终点位置在目标视频中所对应的时间节点，也可以是接收到位置进度调整操作之前目标视频最后播放至的时间节点；第四时间节点可以认为是目标位置在目标视频中所对应的时间节点。

25 本步骤中，在编辑界面的视频区域播放目标视频之后，还可以响应于对目标标识的位置调整操作，对目标标识的位置进行调整。

 示例性的，在对目标标识的位置调整操作的执行过程中，可以依据时间标识在所述编辑轨道片段中的位置，调整所述目标视频的播放进度。当该位置调整操作执行结束时，确定目标标识的目标位

置，对目标位置所位于的第三时间线区间进行确定，并根据第三时间线区间的不同分别对目标视频的播放进度进行不同的调整。

在一个实施方式中，当第三时间线区间为无效文本的时间线区间时，可以将目标视频的播放进度调整至第四时间节点，并从第四时间节点继续播放目标视频。示例性的，当拖拽目标标识至某无效文本的时间线区间 A 内时，本实施例自动将目标视频的播放进度调整至时间线区间 A 的终点位置，以从时间线区间 A 的终点位置继续播放目标视频，或者，将目标标识的位置还原为接收到位置调整操作之前目标标识最后位于的位置，并从该位置继续播放目标视频。

当点击目标视频的编辑轨道片段上所标识的时间线区间 A 内的某位置时，本实施例自动将目标视频的播放进度调整至时间线区间 A 的起点位置，并从时间线区间 A 的起点位置继续播放目标视频。

在一个实施方式中，当第三时间线区间不为无效文本的时间线区间时，则可以将目标视频的播放进度调整至目标位置在目标视频中所对应的时间节点，并从该时间节点继续播放目标视频。示例性的，当拖拽目标标识至的目标位置 b 不位于无效文本的时间线区间内时，或者，当通过点击的方式将目标标识移动至的目标位置 b 不位于无效文本的时间线区间内时，本实施例则可以将目标视频的播放进度调整至目标位置 b 在目标视频中所对应的时间节点 c，并从时间节点 c 继续播放目标视频。在此基础上，通过响应于对目标标识的位置调整操作，能够使得当调整目标标识的位置位于无效片段时，能够自动跳过无效文本所对应的视频片段或者从无效文本所对应视频片段的起点位置完整播放该视频片段；而当调整目标标识的位置不位于无效片段时，从此位置所对应时间节点处播放目标视频的功能。

图 4 为本公开实施例提供的另一种视频的编辑方法的流程示意图。本实施例中的方案可以与上述实施例中的一个或多个可选方案组合。可选的，所述响应于所述无效文本的调整操作，在所述目标

视频的编辑轨道片段上对所述无效文本的时间线区间进行调整，包括下述至少之一：响应于所述无效文本的第一调整操作，在所述目标视频的编辑轨道片段上对所述无效文本的时间线区间的长度进行调整；响应于所述无效文本的第二调整操作，对所述目标视频的编辑轨道片段上所标识的时间线区间的数量进行调整。如图 4 所示，该方法包括：

10 S210、通过对目标视频中的音频进行语音识别，确定所述目标视频的语音文本中的无效文本及所述无效文本的时间线位置，所述无效文本的时间线位置用于表示所述无效文本的语音音频在所述目标视频中的出现时间。

S220、在所述目标视频的编辑界面上展示所述目标视频的编辑轨道片段，并基于所述无效文本的时间线位置，在所述编辑轨道片段上标识所述无效文本的时间线区间，所述无效文本的时间线区间为所述无效文本的语音音频在所述目标视频中出现的时间区间。

15 S230、响应于所述无效文本的第一调整操作，在所述目标视频的编辑轨道片段上对所述无效文本的时间线区间的长度进行调整。

第一调整操作可以用于调整无效文本的时间线区间的长度，第一调整操作可以作用在目标视频的编辑轨道片段上，示例性的，第一调整操作可以为在目标视频的编辑轨道片段上拉伸或压缩某无效文本的时间线区间 B 的操作，如向左拖拽时间线区间 B 的起始位置的操作等。

25 本步骤可以响应于无效文本的第一调整操作，在目标视频的编辑轨道片段上对无效文本的时间线区间的长度进行调整，如当向左拖拽某时间线区间的左边界或向右拖拽某时间线区间的右边界时可以延长该时间线区间的区间长度，当向右拖拽某时间线区间的左边界或者向右拖拽某时间线区间右边界时可以缩短该时间线区间的区间长度。

S240、响应于所述无效文本的第二调整操作，对所述目标视频的编辑轨道片段上所标识的时间线区间的数量进行调整。

第二调整操作可以用于调整编辑轨道片段上所标识的时间线区间的数量，如第二调整操作可以作用在某调整控件上，每个无效文本可以对应一个调整控件，调整控件用于控制编辑轨道片段上是否标识此无效文本对应的时间线区间；第二调整操作也可以作用在编辑轨道片段的时间线区间上。

本步骤可以响应于无效文本的第二调整操作，对目标视频的编辑轨道片段上所标识的时间线区间的数量进行调整，本实施例不对调整的具体过程进行限定，只要能调整编辑轨道片段上所标识的时间线区间的数量即可。

S250、响应于所述无效文本的视频片段删除操作，将所述编辑轨道片段上位于所述无效文本的时间线区间内的视频片段从所述目标视频中删除。

本公开实施例提供的一种视频的编辑方法，通过响应于无效文本的第一调整操作和第二调整操作，能够在目标视频的编辑轨道片段上对无效文本的时间线区间的长度和数量进行调整，从而能够根据用户需求实现对目标视频的精确剪辑。

在一个实施例中，所述在所述目标视频的编辑界面上展示所述目标视频的编辑轨道片段，包括：

在所述编辑界面的轨道区域展示所述目标视频的编辑轨道片段，并在所述编辑界面的文本区域展示所述语音文本中的无效文本和非无效文本，其中，所述无效文本在所述编辑界面上展示为选中状态，所述非无效文本在所述编辑界面上展示为非选中状态。

轨道区域可以为编辑界面内的某个区域，用于展示目标视频的编辑轨道片段。选中状态可以理解为某文本处于被选中的状态，以此表征此文本将要被删除；未选中状态可以理解为某文本处于未被选中的状态，以此表征此文本不会被删除。

本实施例中，在目标视频的编辑界面上展示目标视频的编辑轨道片段的具体过程可以为：在编辑界面的轨道区域对目标视频的编辑轨道片段进行展示，并在编辑界面的文本区域对语音文本中的无

效文本和非无效文本进行展示，其中，无效文本在编辑界面上可以展示为选中状态，非无效文本在编辑界面上可以展示为非选中状态。在此基础上，实现了展示编辑轨道片段的同时，在编辑界面的文本区域展示语音文本中的无效文本和非无效文本；此外，默认无效文本展示为选中状态，能够提示用户可以对无效文本进行删除，并避免用户需要执行相应的选中操作选中停顿、重复等无效文本才能对其进行删除，进一步简化用户删除目标视频中的无效文本所需的操作。

图 5 为本公开实施例提供的另一种编辑界面的示意图，如图 5 所示，在编辑界面的轨道区域 7 展示有目标视频的编辑轨道片段，并在编辑界面的文本区域 8 展示有语音文本中的无效文本（如“停顿 0.12s”或“重复”）和非无效文本（如“我的”），其中，无效文本在编辑界面上展示为选中状态，非无效文本在编辑界面上展示为非选中状态。

在一个实施例中，所述在所述目标视频的编辑轨道片段上对所述无效文本的时间线区间的长度进行调整，包括：

在所述目标视频的编辑轨道片段上对第一文本的第一时间线区间的长度进行调整，并更新所述文本区域中所展示的目标文本，所述目标文本包括所述第一文本和/或与所述第一文本相关联的第二文本。

第一文本可以是无效文本，如第一触发操作所针对时间线区间所属的文本。第一时间线区间即为第一文本的时间线区间。目标文本可以是指对应需要进行调整的文本，如目标文本可以包括第一文本和/或与第一文本相关联的第二文本，第二文本可以为与第一文本的时间线区间相邻时间线区间对应的文本，也可以为与第一文本的时间线区间相邻的时间线区间缩减后的新时间线区间对应的文本。

具体的，在目标视频的编辑轨道片段上对第一文本的第一时间线区间的长度进行调整的同时，需要对文本区域中所展示的目标文本进行更新，更新的具体内容可以由第一调整操作的具体操作来确

定。在此基础上，能够在调整第一时间线区间长度的同时，在文本区域将调整部分对应的文本进行实时更新，以精确化调整无效文本的时间线区间对应的时长。

5 在一个实施例中，所述第一调整操作包括长度延长操作，所述更新所述文本区域中所展示的目标文本，包括：

将所述第二文本中与所述第一时间线区间的延长部分对应的文本内容移动至所述第一文本中。

10 长度延长操作可以认为是延长第一时间线区间长度的操作，如长度延长操作可以为向左拖拽第一时间线区间的起始位置或者向右拖拽第一时间线区间的终止位置的操作。

15 在一个实施方式中，当在编辑轨道片段上延长第一文本的第一时间线区间的长度时，可以在文本区域同步将第二文本中与第一时间线区间的延长部分对应的文本内容移动至第一文本中，示例性的，当第一调整操作为向左拖拽第一时间线区间的起始位置的操作时，第二文本为位于第一时间线区间左侧的时间线区间对应的文本，此时，需要将第二文本中与第一时间线区间的延长部分对应的文本内容移动至第一文本中。

20 在一个实施方式中，长度延长操作可以支持将相邻非无效文本中的部分文本移动至第一文本中，当第二文本为无效文本时，则可以支持或不支持进行长度延长操作。

25 图6为本公开实施例提供的一种更新目标文本的示意图，如图6所示，当由位置A向右拖拽第一文本9对应的第一时间线区间的终点位置至位置B时，可以在目标视频的编辑轨道片段上对第一文本9的第一时间线区间的长度进行延长，同时将第二文本10中与第一时间线区间的延长部分对应的文本内容（如“太阳”）移动至第一文本9中。

在一个实施例中，所述第一调整操作包括长度缩短操作，所述更新所述文本区域中所展示的目标文本，包括：

将所述第一文本中与所述第一时间线区间的缩短部分对应的文

本内容移动至所述第二文本中；或者

在所述文本区域中新增第二文本，并将所述第一文本中与所述第一时间线区间的缩短部分对应的文本内容移动至所述第二文本中。

- 5 长度缩短操作可以认为是缩短第一时间线区间长度的操作，如长度缩短操作可以为向右拖拽第一时间线区间的起始位置或向左拖拽第一时间区间的终止位置的操作。

10 在一个实施方式中，当在编辑轨道片段上缩短第一文本的第一时间线区间的长度时，可以根据第一文本内容的不同来确定文本区域中所展示目标文本的更新方式，示例性的，当第一文本内容为停顿时，可以将第一文本中与第一时间线区间的缩短部分对应的文本内容移动至第二文本中；而当第一文本内容为重复文本、无效词或除停顿、重复和无效词之外的正常文本时，则可以在文本区域中新增第二文本，并将第一文本中与第一时间线区间的缩短部分对应的
15 文本内容移动至新增的第二文本中。

图 7 为本公开实施例提供的另一种更新目标文本的示意图，如图 7 所示，当由位置 A 向左拖拽第一文本 9 对应的第一时间线区间的终点位置至位置 C 时，可以在目标视频的编辑轨道片段上对第一文本 9 的第一时间线区间的长度进行缩短，同时将第一文本 9 中与
20 第一时间线区间的缩短部分对应的文本内容（如“停顿 0.02s”）移动至第二文本 10 中。

图 8 为本公开实施例提供的另一种更新目标文本的示意图，如图 8 所示，当由位置 D 向右拖拽第一文本 11 对应的第一时间线区间的终点位置至位置 E 时，可以在目标视频的编辑轨道片段上对第一
25 文本 11 的第一时间线区间的长度进行缩短，同时可以在文本区域中新增第二文本 12，并将第一文本 11 中与第一时间线区间的缩短部分对应的文本内容（即“重复”）移动至第二文本 12 中。

在一个实施例中，所述无效文本为所述轨道区域中处于选中状态的文本，所述响应于所述无效文本的第二调整操作，对所述目标

视频的编辑轨道片段上所标识的时间线区间的数量进行调整，包括下述至少之一：

响应于对所述文本区域中的第三文本的选中操作，将所述第三文本由未选中状态切换为选中状态，并在所述编辑轨道片段上增加标识所述第三文本的时间线区间；

响应于对所述文本区域中的第四文本的取消选中操作，将所述第四文本由选中状态切换为未选中状态，并在所述编辑轨道片段上取消标识所述第四文本的时间线区间。

第三文本可以是文本区域中的任一文本，如文本区域中处于未选中状态的某文本。选中操作可以是指将文本由未选中状态切换为选中状态的操作，如选中操作可以为点击某控件的操作，该控件可以具有第一显示样式和第二显示样式，且文本的选中状态可以跟随该控件显示样式的变化而变化，第一显示样式和第二显示样式的具体内容不限，只要能区分不同的显示样式即可。

第四文本可以是文本区域中的任一文本，如文本区域中处于选中状态的某文本。第四文本可以与第三文本相同，也可以不同。取消选中操作可以是指将文本由选中状态切换为未选中状态的操作，如取消选中操作可以与选中操作相对应，为点击某控件的操作。

具体的，本实施例可以响应于对文本区域中的第三文本的选中操作，将第三文本由未选中状态切换为选中状态，并在编辑轨道片段上增加标识第三文本的时间线区间，在此基础上，用户可以根据需要增加将要删除的第三文本以及第三文本的时间线区间。

本实施例也可以响应于对文本区域中的第四文本的取消选中操作，将第四文本由选中状态切换为未选中状态，并在编辑轨道片段上取消标识第四文本的时间线区间。在此基础上，用户可以根据需要取消将要删除的第四文本以及取消标识第四文本的时间线区间。

示例性的，当点击某控件使该控件显示为第一显示样式时，可以认为是触发了对该控件所对应第三文本的选中操作，此时，可以将第三文本由未选中状态切换为选中状态，并在编辑轨道片段上增

加标识第三文本的时间线区间；当点击某控件使该控件显示为第二显示样式时，可以认为是触发了对该控件所对应第四文本的取消选中操作，此时，可以将第四文本由选中状态切换为未选中状态，并在编辑轨道片段上取消标识第四文本的时间线区间。

5

图9为本公开实施例提供的一种视频的编辑装置的结构示意图，该装置可适用于对视频进行编辑的情况，其中该装置可由软件和/或硬件实现，并一般集成在电子设备上。

如图9所示，该装置包括：

10 文本确定模块310，用于通过对目标视频中的音频进行语音识别，确定所述目标视频的语音文本中的无效文本及所述无效文本的时间线位置，所述无效文本的时间线位置用于表示所述无效文本的语音音频在所述目标视频中的出现时间；

15 区间标识模块320，用于在所述目标视频的编辑界面上展示所述目标视频的编辑轨道片段，并基于所述无效文本的时间线位置，在所述编辑轨道片段上标识所述无效文本的时间线区间，所述无效文本的时间线区间为所述无效文本的语音音频在所述目标视频中出现的区间；

20 区间调整模块330，用于响应于所述无效文本的调整操作，在所述目标视频的编辑轨道片段上对所述无效文本的时间线区间进行调整；

片段删除模块340，用于响应于所述无效文本的视频片段删除操作，将所述编辑轨道片段上位于所述无效文本的时间线区间内的视频片段从所述目标视频中删除。

25 本公开实施例提供的一种视频的编辑装置，通过文本确定模块通过对目标视频中的音频进行语音识别，确定所述目标视频的语音文本中的无效文本及所述无效文本的时间线位置，所述无效文本的时间线位置用于表示所述无效文本的语音音频在所述目标视频中的出现时间；通过区间标识模块在所述目标视频的编辑界面上展示所

述目标视频的编辑轨道片段，并基于所述无效文本的时间线位置，在所述编辑轨道片段上标识所述无效文本的时间线区间，所述无效文本的时间线区间为所述无效文本的语音音频在所述目标视频中出现的
5 的时间区间；通过区间调整模块响应于所述无效文本的调整操作，在所述目标视频的编辑轨道片段上对所述无效文本的时间线区间进行调整；通过片段删除模块响应于所述无效文本的视频片段删除操作，将所述编辑轨道片段上位于所述无效文本的时间线区间内的视频片段从所述目标视频中删除。利用该装置，通过对目标视频中的音频进行语音识别，能够对目标视频中的无效文本以及无效文本的
10 位置进行确定，进而通过在目标视频的编辑轨道片段上标识无效文本的时间线区间，能够响应于用户对无效文本的调整操作和视频片段删除操作，对无效片段进行删除，从而完成对视频的剪辑，提升了视频剪辑的效率。

可选的，所述区间调整模块 330 具体包括下述至少之一：

15 第一响应单元，用于响应于所述无效文本的第一调整操作，在所述目标视频的编辑轨道片段上对所述无效文本的时间线区间的长度进行调整；

第二响应单元，用于响应于所述无效文本的第二调整操作，对
20 所述目标视频的编辑轨道片段上所标识的时间线区间的数量进行调整。

可选的，所述区间标识模块 320 具体用于：

在所述编辑界面的轨道区域展示所述目标视频的编辑轨道片段，并在所述编辑界面的文本区域展示所述语音文本中的无效文本
25 和非无效文本，其中，所述无效文本在所述编辑界面上展示为选中状态，所述非无效文本在所述编辑界面上展示为非选中状态。

可选的，所述第一响应单元包括：

调整子单元，用于在所述目标视频的编辑轨道片段上对第一文本的第一时间线区间的长度进行调整，并更新所述文本区域中所展示的目标文本，所述目标文本包括所述第一文本和/或与所述第一文

本相关联的第二文本。

可选的，所述第一调整操作包括长度延长操作，所述调整子单元具体用于：

将所述第二文本中与所述第一时间线区间的延长部分对应的文本内容移动至所述第一文本中。

可选的，所述第一调整操作包括长度缩短操作，所述调整子单元具体用于：

将所述第一文本中与所述第一时间线区间的缩短部分对应的文本内容移动至所述第二文本中；或者

10 在所述文本区域中新增第二文本，并将所述第一文本中与所述第一时间线区间的缩短部分对应的文本内容移动至所述第二文本中。

可选的，所述无效文本为所述轨道区域中处于选中状态的文本，所述第二响应单元具体用于下述至少之一：

15 响应于对所述文本区域中的第三文本的选中操作，将所述第三文本由未选中状态切换为选中状态，并在所述编辑轨道片段上增加标识所述第三文本的时间线区间；

响应于对所述文本区域中的第四文本的取消选中操作，将所述第四文本由选中状态切换为未选中状态，并在所述编辑轨道片段上
20 取消标识所述第四文本的时间线区间。

可选的，本公开实施例提供的一种视频的编辑装置，还包括：

播放模块，用于在所述编辑界面的视频区域播放所述目标视频；

第一进度调整模块，用于当所述目标视频播放至第二时间线区间的起点位置时，将所述目标视频的播放进度调整至第一时间节点，
25 并从所述第一时间节点继续播放所述目标视频，其中，所述第二时间线区间为所述无效文本的时间线区间，所述第一时间节点为所述第二时间线区间的终点位置在所述目标视频中所对应的时间节点。

可选的，本公开实施例提供的一种视频的编辑装置，还包括：

第一响应模块，用于在所述编辑界面的视频区域播放所述目标

视频之后，响应于对第五文本的触发操作，将所述目标视频的播放进度调整至第二时间节点，并从所述第二时间节点继续播放所述目标视频，所述第五文本显示于所述编辑界面的文本区域内，所述第五文本为无效文本或非无效文本，所述第二时间节点为所述第五文本的起点位置在所述目标视频中所对应的时间节点。

可选的，所述编辑轨道片段上还展示有时间标识，本公开实施例提供的一种视频的编辑装置，还包括：

第二响应模块，用于在所述编辑界面的视频区域播放所述目标视频之后，响应于对所述目标标识的位置调整操作，确定所述目标标识的目标位置所位于的第三时间线区间，所述目标位置为所述位置调整操作执行完毕时所述目标标识在所述编辑轨道片段上的位置；

第二进度调整模块，用于如果所述第三时间线区间为所述无效文本的时间线区间，则将所述目标视频的播放进度调整至第三时间节点，并从所述第三时间节点继续播放所述目标视频，其中，所述第三时间节点为所述第三时间线区间的起点位置或终点位置在所述目标视频中所对应的时间节点，或者，接收到所述位置进度调整操作之前所述目标视频最后播放至的时间节点；

第三进度调整模块，用于如果所述第三时间线区间不为所述无效文本的时间线区间，则将所述目标视频的播放进度调整至第四时间节点，并从所述第四时间节点继续播放所述目标视频，其中，所述第四时间节点为所述目标位置在所述目标视频中所对应的时间节点。

可选的，所述片段删除模块 340 具体用于：

将所述编辑轨道片段上位于所述无效文本的时间线区间内的视频片段从所述目标视频中删除，并以删除的所述视频片段为分割点，将所述目标视频分割为多个独立的视频片段；或者

将所述编辑轨道片段上位于所述无效文本的时间线区间内的视频片段从所述目标视频中删除，并将删除所述视频片段后的目标视

频整合为一个视频。

可选的，所述编辑轨道片段为所述目标视频的音频轨道片段。

上述视频的编辑装置可执行本公开任意实施例所提供的视频的编辑方法，具备执行方法相应的功能模块和有益效果。

5

下面参考图 10，其示出了适于用来实现本公开实施例的电子设
备 400 的结构示意图。本公开实施例中的终端设备可以包括但不限于
诸如移动电话、笔记本电脑、数字广播接收器、PDA（个人数字
助理）、PAD（平板电脑）、PMP（便携式多媒体播放器）、车载
10 终端（例如车载导航终端）等等的移动终端以及诸如数字 TV、台式
计算机等等的固定终端。图 10 示出的电子设备仅仅是一个示例，不
应对本公开实施例的功能和使用范围带来任何限制。

如图 10 所示，电子设备 400 可以包括处理装置（例如中央处理
器、图形处理器等）401，其可以根据存储在只读存储器（ROM）402
15 中的程序或者从存储装置 408 加载到随机访问存储器（RAM）403
中的程序而执行各种适当的动作和处理。在 RAM 403 中，还存储有
电子设备 400 操作所需的各种程序和数据。处理装置 401、ROM 402
以及 RAM 403 通过总线 404 彼此相连。输入/输出（I/O）接口 405
也连接至总线 404。

20 通常，以下装置可以连接至 I/O 接口 405：包括例如触摸屏、触
摸板、键盘、鼠标、摄像头、麦克风、加速度计、陀螺仪等的输入
装置 406；包括例如液晶显示器（LCD）、扬声器、振动器等的输出
装置 407；包括例如磁带、硬盘等的存储装置 408；以及通信装置 409。
通信装置 409 可以允许电子设备 400 与其他设备进行无线或有线通
25 信以交换数据。虽然图 10 示出了具有各种装置的电子设备 400，但
是应理解的是，并不要求实施或具备所有示出的装置。可以替代地
实施或具备更多或更少的装置。

特别地，根据本公开的实施例，上文参考流程图描述的过程可
以被实现为计算机软件程序。例如，本公开的实施例包括一种计算

机程序产品，其包括承载在非暂态计算机可读介质上的计算机程序，该计算机程序包含用于执行流程图所示的方法的程序代码。在这样的实施例中，该计算机程序可以通过通信装置 409 从网络上被下载和安装，或者从存储装置 408 被安装，或者从 ROM 402 被安装。在
5 该计算机程序被处理装置 401 执行时，执行本公开实施例的方法中限定的上述功能。

需要说明的是，本公开上述的计算机可读介质可以是计算机可读信号介质或者计算机可读存储介质或者是上述两者的任意组合。计算机可读存储介质例如可以是——但不限于——电、磁、光、电
10 磁、红外线、或半导体的系统、装置或器件，或者任意以上的组合。计算机可读存储介质的更具体的例子可以包括但不限于：具有一个或多个导线的电连接、便携式计算机磁盘、硬盘、随机访问存储器（RAM）、只读存储器（ROM）、可擦式可编程只读存储器（EPROM 或闪存）、光纤、便携式紧凑磁盘只读存储器（CD-ROM）、光存
15 储器件、磁存储器件、或者上述的任意合适的组合。在本公开中，计算机可读存储介质可以是任何包含或存储程序的有形介质，该程序可以被指令执行系统、装置或者器件使用或者与其结合使用。而在本公开中，计算机可读信号介质可以包括在基带中或者作为载波一部分传播的数据信号，其中承载了计算机可读的程序代码。这种
20 传播的数据信号可以采用多种形式，包括但不限于电磁信号、光信号或上述的任意合适的组合。计算机可读信号介质还可以是计算机可读存储介质以外的任何计算机可读介质，该计算机可读信号介质可以发送、传播或者传输用于由指令执行系统、装置或者器件使用或者与其结合使用的程序。计算机可读介质上包含的程序代码可以用
25 任何适当的介质传输，包括但不限于：电线、光缆、RF（射频）等等，或者上述的任意合适的组合。

在一些实施方式中，客户端、服务器可以利用诸如 HTTP（HyperText Transfer Protocol，超文本传输协议）之类的任何当前已知或未来研发的网络协议进行通信，并且可以与任意形式或介质的

数字数据通信（例如，通信网络）互连。通信网络的示例包括局域网（“LAN”），广域网（“WAN”），网际网（例如，互联网）以及端对端网络（例如，ad hoc 端对端网络），以及任何当前已知或未来研发的网络。

- 5 上述计算机可读介质可以是上述电子设备中所包含的；也可以是单独存在，而未装配入该电子设备中。

上述计算机可读介质承载有一个或者多个程序，当上述一个或者多个程序被该电子设备执行时，使得该电子设备：

- 10 通过对目标视频中的音频进行语音识别，确定所述目标视频的语音文本中的无效文本及所述无效文本的时间线位置，所述无效文本的时间线位置用于表示所述无效文本的语音音频在所述目标视频中的出现时间；

- 15 在所述目标视频的编辑界面上展示所述目标视频的编辑轨道片段，并基于所述无效文本的时间线位置，在所述编辑轨道片段上标识所述无效文本的时间线区间，所述无效文本的时间线区间为所述无效文本的语音音频在所述目标视频中出现的时间区间；

响应于所述无效文本的调整操作，在所述目标视频的编辑轨道片段上对所述无效文本的时间线区间进行调整；

- 20 响应于所述无效文本的视频片段删除操作，将所述编辑轨道片段上位于所述无效文本的时间线区间内的视频片段从所述目标视频中删除。

- 25 可以以一种或多种程序设计语言或其组合来编写用于执行本公开的操作的计算机程序代码，上述程序设计语言包括但不限于面向对象的程序设计语言—诸如 Java、Smalltalk、C++，还包括常规的过
程式程序设计语言—诸如“C”语言或类似的程序设计语言。程序代码可以完全地在用户计算机上执行、部分地在用户计算机上执行、作为一个独立的软件包执行、部分在用户计算机上部分在远程计算机上执行、或者完全在远程计算机或服务器上执行。在涉及远程计算机的情形中，远程计算机可以通过任意种类的网络——包括局域

网(LAN)或广域网(WAN)一连接到用户计算机,或者,可以连接到外部计算机(例如利用因特网服务提供商来通过因特网连接)。

附图中的流程图和框图,图示了按照本公开各种实施例的系统、方法和计算机程序产品的可能实现的体系架构、功能和操作。在这点上,流程图或框图中的每个方框可以代表一个模块、程序段、或代码的一部分,该模块、程序段、或代码的一部分包含一个或多个用于实现规定的逻辑功能的可执行指令。也应当注意,在有些作为替换的实现中,方框中所标注的功能也可以以不同于附图中所标注的顺序发生。例如,两个接连地表示的方框实际上可以基本并行地执行,它们有时也可以按相反的顺序执行,这依所涉及的功能而定。也要注意的,框图和/或流程图中的每个方框、以及框图和/或流程图中的方框的组合,可以用执行规定的功能或操作的专用的基于硬件的系统来实现,或者可以用专用硬件与计算机指令的组合来实现。

描述于本公开实施例中所涉及到的单元可以通过软件的方式实现,也可以通过硬件的方式来实现。其中,模块的名称在某种情况下并不构成对该单元本身的限定。

本文中以上描述的功能可以至少部分地由一个或多个硬件逻辑部件来执行。例如,非限制性地,可以使用的示范类型的硬件逻辑部件包括:现场可编程门阵列(FPGA)、专用集成电路(ASIC)、专用标准产品(ASSP)、片上系统(SOC)、复杂可编程逻辑设备(CPLD)等等。

在本公开的上下文中,机器可读介质可以是有形的介质,其可以包含或存储以供指令执行系统、装置或设备使用或与指令执行系统、装置或设备结合地使用的程序。机器可读介质可以是机器可读信号介质或机器可读储存介质。机器可读介质可以包括但不限于电子的、磁性的、光学的、电磁的、红外的、或半导体系统、装置或设备,或者上述内容的任何合适组合。机器可读存储介质的更具体示例会包括基于一个或多个线的电气连接、便携式计算机盘、硬盘、随机存取存储器(RAM)、只读存储器(ROM)、可擦除可编程只

读存储器（EPROM 或快闪存储器）、光纤、便捷式紧凑盘只读存储器（CD-ROM）、光学储存设备、磁储存设备、或上述内容的任何合适组合。

5 根据本公开的一个或多个实施例，示例 1 提供了一种视频的编辑方法，包括：

 通过对目标视频中的音频进行语音识别，确定所述目标视频的语音文本中的无效文本及所述无效文本的时间线位置，所述无效文本的时间线位置用于表示所述无效文本的语音音频在所述目标视频
10 中的出现时间；

 在所述目标视频的编辑界面上展示所述目标视频的编辑轨道片段，并基于所述无效文本的时间线位置，在所述编辑轨道片段上标识所述无效文本的时间线区间，所述无效文本的时间线区间为所述无效文本的语音音频在所述目标视频中出现的时间区间；

15 响应于所述无效文本的调整操作，在所述目标视频的编辑轨道片段上对所述无效文本的时间线区间进行调整；

 响应于所述无效文本的视频片段删除操作，将所述编辑轨道片段上位于所述无效文本的时间线区间内的视频片段从所述目标视频中删除。

20 根据本公开的一个或多个实施例，示例 2 根据示例 1 所述的方法，所述响应于所述无效文本的调整操作，在所述目标视频的编辑轨道片段上对所述无效文本的时间线区间进行调整，包括下述至少之一：

 响应于所述无效文本的第一调整操作，在所述目标视频的编辑
25 轨道片段上对所述无效文本的时间线区间的长度进行调整；

 响应于所述无效文本的第二调整操作，对所述目标视频的编辑轨道片段上所标识的时间线区间的数量进行调整。

 根据本公开的一个或多个实施例，示例 3 根据示例 2 所述的方法，所述在所述目标视频的编辑界面上展示所述目标视频的编辑轨

道片段，包括：

在所述编辑界面的轨道区域展示所述目标视频的编辑轨道片段，并在所述编辑界面的文本区域展示所述语音文本中的无效文本和非无效文本，其中，所述无效文本在所述编辑界面上展示为选中状态，所述非无效文本在所述编辑界面上展示为非选中状态。

根据本公开的一个或多个实施例，示例 4 根据示例 3 所述的方法，所述在所述目标视频的编辑轨道片段上对所述无效文本的时间线区间的长度进行调整，包括：

在所述目标视频的编辑轨道片段上对第一文本的第一时间线区间的长度进行调整，并更新所述文本区域中所展示的目标文本，所述目标文本包括所述第一文本和/或与所述第一文本相关联的第二文本。

根据本公开的一个或多个实施例，示例 5 根据示例 4 所述的方法，所述第一调整操作包括长度延长操作，所述更新所述文本区域中所展示的目标文本，包括：

将所述第二文本中与所述第一时间线区间的延长部分对应的文本内容移动至所述第一文本中。

根据本公开的一个或多个实施例，示例 6 根据示例 4 所述的方法，所述第一调整操作包括长度缩短操作，所述更新所述文本区域中所展示的目标文本，包括：

将所述第一文本中与所述第一时间线区间的缩短部分对应的文本内容移动至所述第二文本中；或者

在所述文本区域中新增第二文本，并将所述第一文本中与所述第一时间线区间的缩短部分对应的文本内容移动至所述第二文本中。

根据本公开的一个或多个实施例，示例 7 根据示例 3 所述的方法，所述无效文本为所述轨道区域中处于选中状态的文本，所述响应于所述无效文本的第二调整操作，对所述目标视频的编辑轨道片段上所标识的时间线区间的数量进行调整，包括下述至少之一：

响应于对所述文本区域中的第三文本的选中操作，将所述第三文本由未选中状态切换为选中状态，并在所述编辑轨道片段上增加标识所述第三文本的时间线区间；

5 响应于对所述文本区域中的第四文本的取消选中操作，将所述第四文本由选中状态切换为未选中状态，并在所述编辑轨道片段上取消标识所述第四文本的时间线区间。

根据本公开的一个或多个实施例，示例 8 根据示例 1-7 任一所述的方法，还包括：

在所述编辑界面的视频区域播放所述目标视频；

10 当所述目标视频播放至第二时间线区间的起点位置时，将所述目标视频的播放进度调整至第一时间节点，并从所述第一时间节点继续播放所述目标视频，其中，所述第二时间线区间为所述无效文本的时间线区间，所述第一时间节点为所述第二时间线区间的终点位置在所述目标视频中所对应的时间节点。

15 根据本公开的一个或多个实施例，示例 9 根据示例 8 所述的方法，在所述编辑界面的视频区域播放所述目标视频之后，还包括：

响应于对第五文本的触发操作，将所述目标视频的播放进度调整至第二时间节点，并从所述第二时间节点继续播放所述目标视频，所述第五文本显示于所述编辑界面的文本区域内，所述第五文本为无效文本或非无效文本，所述第二时间节点为所述第五文本的起点位置在所述目标视频中所对应的时间节点。

20

根据本公开的一个或多个实施例，示例 10 根据示例 8 所述的方法，所述编辑轨道片段上还展示有时间标识，在所述编辑界面的视频区域播放所述目标视频之后，还包括：

25 响应于对所述目标标识的位置调整操作，确定所述目标标识的目标位置所位于的第三时间线区间，所述目标位置为所述位置调整操作执行完毕时所述目标标识在所述编辑轨道片段上的位置；

如果所述第三时间线区间为所述无效文本的时间线区间，则将所述目标视频的播放进度调整至第三时间节点，并从所述第三时间

节点继续播放所述目标视频，其中，所述第三时间节点为所述第三时间线区间的起点位置或终点位置在所述目标视频中所对应的时间节点，或者，接收到所述位置进度调整操作之前所述目标视频最后播放至的时间节点；

- 5 如果所述第三时间线区间不为所述无效文本的时间线区间，则将所述目标视频的播放进度调整至第四时间节点，并从所述第四时间节点继续播放所述目标视频，其中，所述第四时间节点为所述目标位置在所述目标视频中所对应的时间节点。

10 根据本公开的一个或多个实施例，示例 11 根据示例 1-7 任一所述的方法，所述将所述编辑轨道片段上位于所述无效文本的时间线区间内的视频片段从所述目标视频中删除，包括：

将所述编辑轨道片段上位于所述无效文本的时间线区间内的视频片段从所述目标视频中删除，并以删除的所述视频片段为分割点，将所述目标视频分割为多个独立的视频片段；或者

- 15 将所述编辑轨道片段上位于所述无效文本的时间线区间内的视频片段从所述目标视频中删除，并将删除所述视频片段后的目标视频整合为一个视频。

根据本公开的一个或多个实施例，示例 12 根据示例 1-7 任一所述的方法，所述编辑轨道片段为所述目标视频的音频轨道片段。

- 20 以上描述仅为本公开的较佳实施例以及对所运用技术原理的说明。本领域技术人员应当理解，本公开中所涉及的公开范围，并不限于上述技术特征的特定组合而成的技术方案，同时也应涵盖在不脱离上述公开构思的情况下，由上述技术特征或其等同特征进行任意组合而形成的其它技术方案。例如上述特征与本公开中公开的（但
25 不限于）具有类似功能的技术特征进行互相替换而形成的技术方案。

此外，虽然采用特定次序描绘了各操作，但是这不应理解为要求这些操作以所示出的特定次序或以顺序次序执行来执行。在一定环境下，多任务和并行处理可能是有利的。同样地，虽然在上面论述中包含了若干具体实现细节，但是这些不应被解释为对本公

开的范围的限制。在单独的实施例的上下文中描述的某些特征还可以组合地实现在单个实施例中。相反地，在单个实施例的上下文中描述的各种特征也可以单独地或以任何合适的子组合的方式实现在多个实施例中。

- 5 尽管已经采用特定于结构特征和/或方法逻辑动作的语言描述了本主题，但是应当理解所附权利要求书中所限定的主题未必局限于上面描述的特定特征或动作。相反，上面所描述的特定特征和动作仅仅是实现权利要求书的示例形式。

权利要求书

1. 一种视频的编辑方法，包括：

通过对目标视频中的音频进行语音识别，确定所述目标视频的
5 语音文本中的无效文本及所述无效文本的时间线位置，所述无效文
本的时间线位置用于表示所述无效文本的语音音频在所述目标视频
中的出现时间；

在所述目标视频的编辑界面上展示所述目标视频的编辑轨道片
段，并基于所述无效文本的时间线位置，在所述编辑轨道片段上标
10 识所述无效文本的时间线区间，所述无效文本的时间线区间为所述
无效文本的语音音频在所述目标视频中出现的时间区间；

响应于所述无效文本的调整操作，在所述目标视频的编辑轨道
片段上对所述无效文本的时间线区间进行调整；

响应于所述无效文本的视频片段删除操作，将所述编辑轨道片
15 段上位于所述无效文本的时间线区间内的视频片段从所述目标视频
中删除。

2. 根据权利要求 1 所述的方法，其中所述响应于所述无效文本
的调整操作，在所述目标视频的编辑轨道片段上对所述无效文本的
时间线区间进行调整，包括下述至少之一：

20 响应于所述无效文本的第一调整操作，在所述目标视频的编辑
轨道片段上对所述无效文本的时间线区间的长度进行调整；

响应于所述无效文本的第二调整操作，对所述目标视频的编辑
轨道片段上所标识的时间线区间的数量进行调整。

3. 根据权利要求 2 所述的方法，其中所述在所述目标视频的编
25 辑界面上展示所述目标视频的编辑轨道片段，包括：

在所述编辑界面的轨道区域展示所述目标视频的编辑轨道片
段，并在所述编辑界面的文本区域展示所述语音文本中的无效文本
和非无效文本，其中，所述无效文本在所述编辑界面上展示为选中
状态，所述非无效文本在所述编辑界面上展示为非选中状态。

4. 根据权利要求 3 所述的方法，其中所述在所述目标视频的编辑轨道片段上对所述无效文本的时间线区间的长度进行调整，包括：

在所述目标视频的编辑轨道片段上对第一文本的第一时间线区间的长度进行调整，并更新所述文本区域中所展示的目标文本，所述目标文本包括所述第一文本和/或与所述第一文本相关联的第二文本。

5. 根据权利要求 4 所述的方法，其中所述第一调整操作包括长度延长操作，所述更新所述文本区域中所展示的目标文本，包括：

将所述第二文本中与所述第一时间线区间的延长部分对应的文本内容移动至所述第一文本中。

6. 根据权利要求 4 所述的方法，其中所述第一调整操作包括长度缩短操作，所述更新所述文本区域中所展示的目标文本，包括：

将所述第一文本中与所述第一时间线区间的缩短部分对应的文本内容移动至所述第二文本中；或者

在所述文本区域中新增第二文本，并将所述第一文本中与所述第一时间线区间的缩短部分对应的文本内容移动至所述第二文本中。

7. 根据权利要求 3 所述的方法，其中所述无效文本为所述轨道区域中处于选中状态的文本，所述响应于所述无效文本的第二调整操作，对所述目标视频的编辑轨道片段上所标识的时间线区间的数量进行调整，包括下述至少之一：

响应于对所述文本区域中的第三文本的选中操作，将所述第三文本由未选中状态切换为选中状态，并在所述编辑轨道片段上增加标识所述第三文本的时间线区间；

响应于对所述文本区域中的第四文本的取消选中操作，将所述第四文本由选中状态切换为未选中状态，并在所述编辑轨道片段上取消标识所述第四文本的时间线区间。

8. 根据权利要求 1-7 任一所述的方法，还包括：

在所述编辑界面的视频区域播放所述目标视频；

当所述目标视频播放至第二时间线区间的起点位置时，将所述目标视频的播放进度调整至第一时间节点，并从所述第一时间节点继续播放所述目标视频，其中，所述第二时间线区间为所述无效文本的时间线区间，所述第一时间节点为所述第二时间线区间的终点位置在所述目标视频中所对应的时间节点。

9. 根据权利要求 8 所述的方法，其中在所述编辑界面的视频区域播放所述目标视频之后，还包括：

响应于对第五文本的触发操作，将所述目标视频的播放进度调整至第二时间节点，并从所述第二时间节点继续播放所述目标视频，所述第五文本显示于所述编辑界面的文本区域内，所述第五文本为无效文本或非无效文本，所述第二时间节点为所述第五文本的起点位置在所述目标视频中所对应的时间节点。

10. 根据权利要求 8 所述的方法，其中所述编辑轨道片段上还展示有时间标识，在所述编辑界面的视频区域播放所述目标视频之后，还包括：

响应于对所述目标标识的位置调整操作，确定所述目标标识的目标位置所位于的第三时间线区间，所述目标位置为所述位置调整操作执行完毕时所述目标标识在所述编辑轨道片段上的位置；

如果所述第三时间线区间为所述无效文本的时间线区间，则将所述目标视频的播放进度调整至第三时间节点，并从所述第三时间节点继续播放所述目标视频，其中，所述第三时间节点为所述第三时间线区间的起点位置或终点位置在所述目标视频中所对应的时间节点，或者，接收到所述位置进度调整操作之前所述目标视频最后播放至的时间节点；

如果所述第三时间线区间不为所述无效文本的时间线区间，则将所述目标视频的播放进度调整至第四时间节点，并从所述第四时间节点继续播放所述目标视频，其中，所述第四时间节点为所述目标位置在所述目标视频中所对应的时间节点。

11. 根据权利要求 1-7 任一所述的方法，其中所述将所述编辑轨

道片段上位于所述无效文本的时间线区间内的视频片段从所述目标视频中删除，包括：

将所述编辑轨道片段上位于所述无效文本的时间线区间内的视频片段从所述目标视频中删除，并以删除的所述视频片段为分割点，
5 将所述目标视频分割为多个独立的视频片段；或者

将所述编辑轨道片段上位于所述无效文本的时间线区间内的视频片段从所述目标视频中删除，并将删除所述视频片段后的目标视频整合为一个视频。

12. 根据权利要求 1-7 任一所述的方法，其中所述编辑轨道片段
10 为所述目标视频的音频轨道片段。

13. 一种视频的编辑装置，包括：

文本确定模块，用于通过对目标视频中的音频进行语音识别，
确定所述目标视频的语音文本中的无效文本及所述无效文本的时间
线位置，所述无效文本的时间线位置用于表示所述无效文本的语音
15 音频在所述目标视频中的出现时间；

区间标识模块，用于在所述目标视频的编辑界面上展示所述目
标视频的编辑轨道片段，并基于所述无效文本的时间线位置，在所
述编辑轨道片段上标识所述无效文本的时间线区间，所述无效文本
的时间线区间为所述无效文本的语音音频在所述目标视频中出现的
20 时间区间；

区间调整模块，用于响应于所述无效文本的调整操作，在所述
目标视频的编辑轨道片段上对所述无效文本的时间线区间进行调
整；

片段删除模块，用于响应于所述无效文本的视频片段删除操作，
25 将所述编辑轨道片段上位于所述无效文本的时间线区间内的视频片
段从所述目标视频中删除。

14. 一种电子设备，包括：

至少一个处理器；以及

与所述至少一个处理器通信连接的存储器；其中，

所述存储器存储有可被所述至少一个处理器执行的计算机程序，所述计算机程序被所述至少一个处理器执行，以使所述至少一个处理器能够执行权利要求 1-12 中任一项所述的视频的编辑方法。

15. 一种计算机可读存储介质，所述计算机可读存储介质存储有
5 计算机指令，所述计算机指令用于使处理器执行时实现权利要求
1-12 中任一项所述的视频的编辑方法。

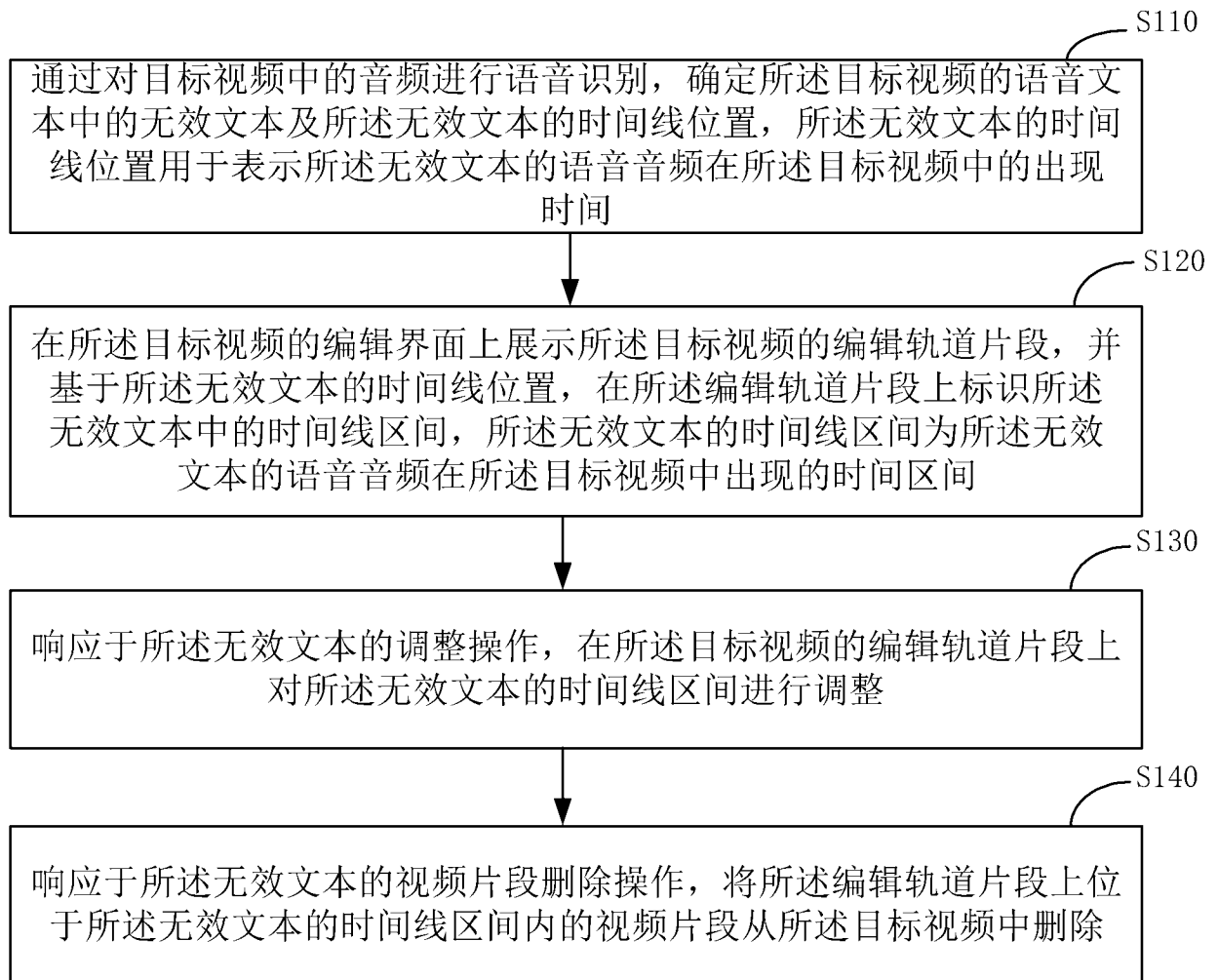


图 1

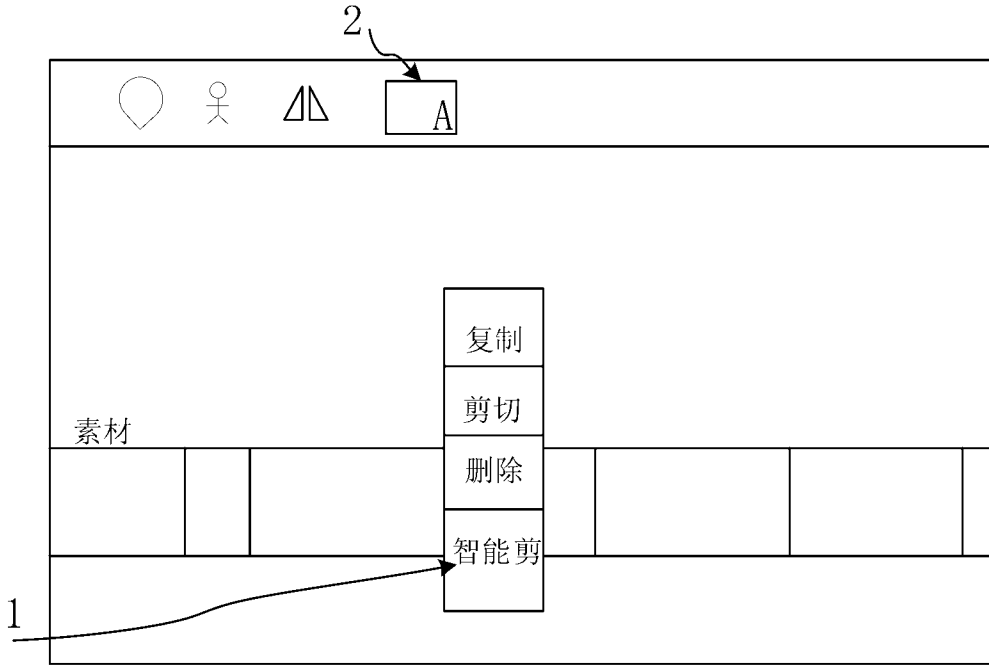


图 2

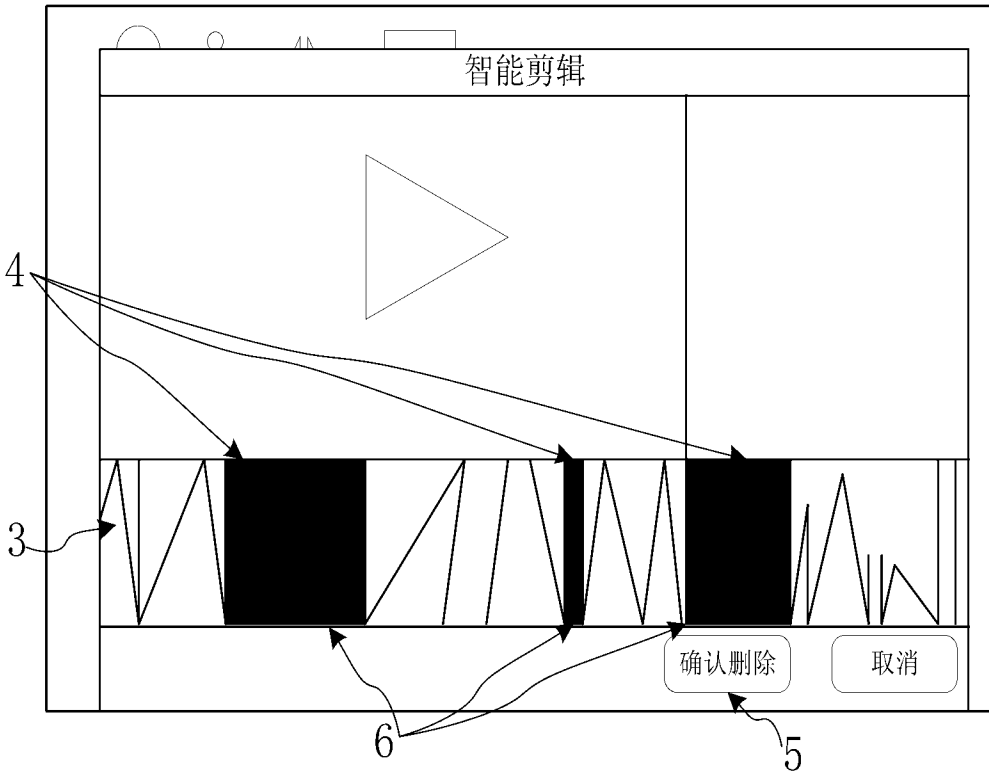


图 3

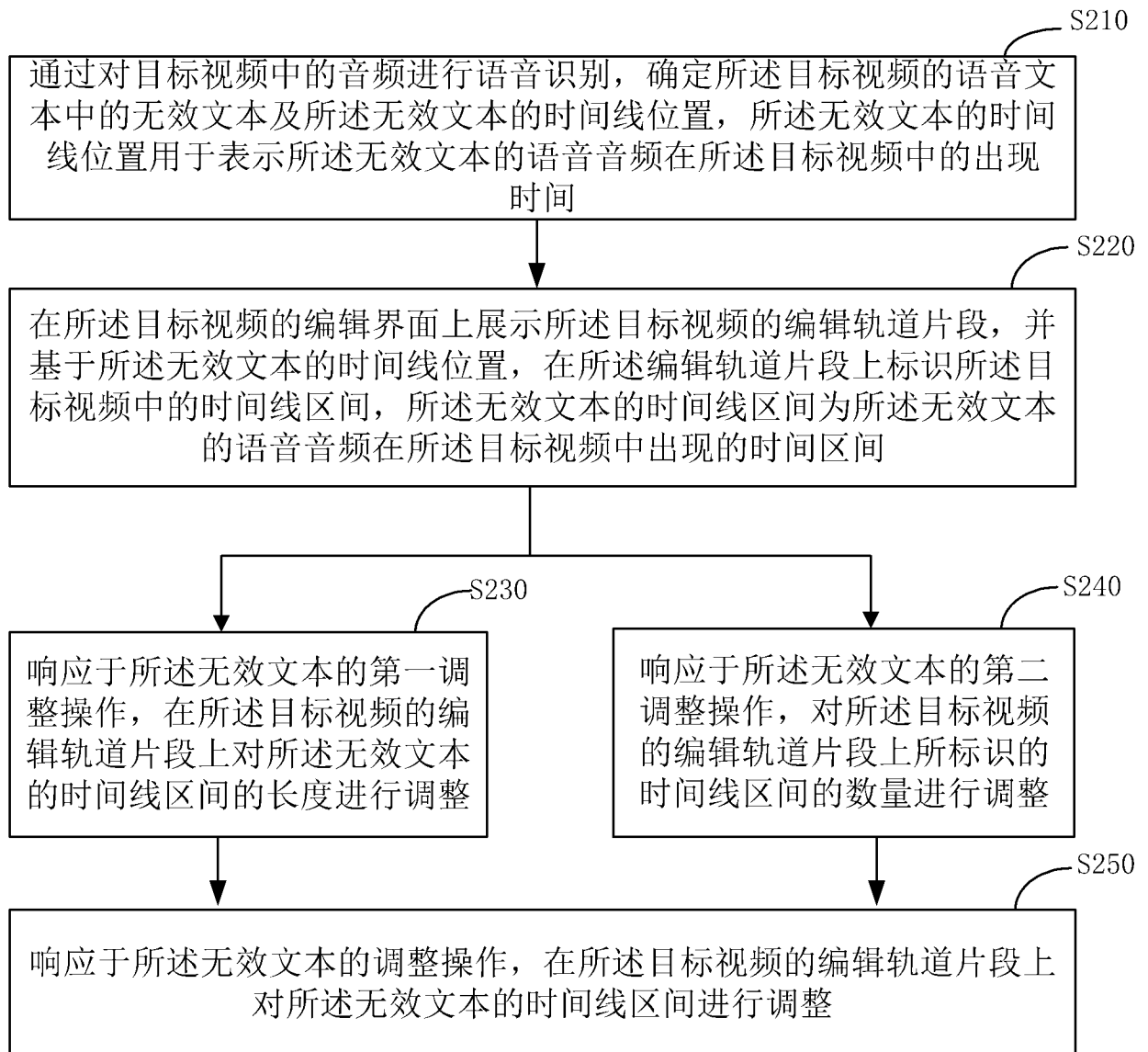


图 4

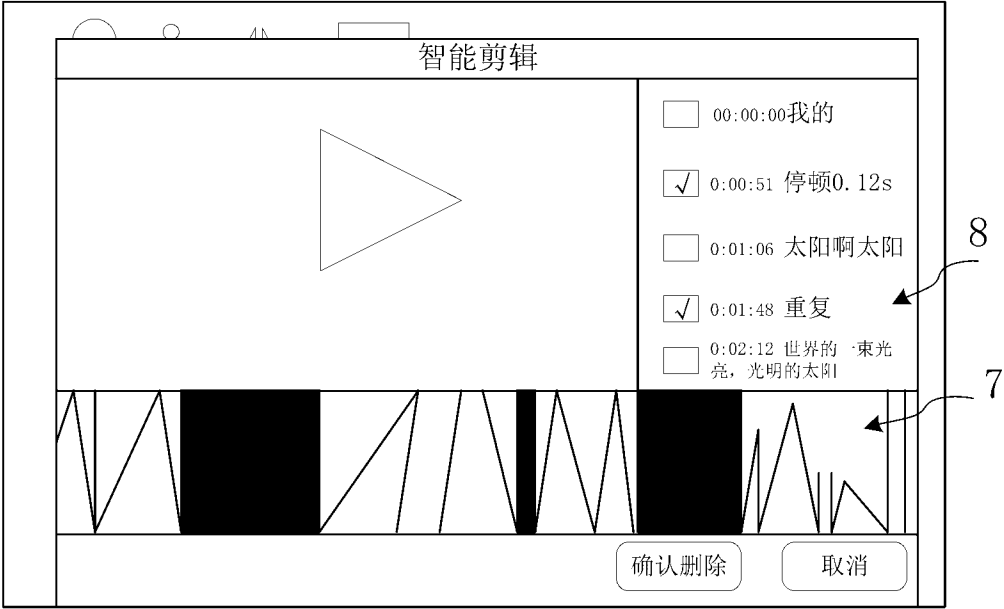


图 5

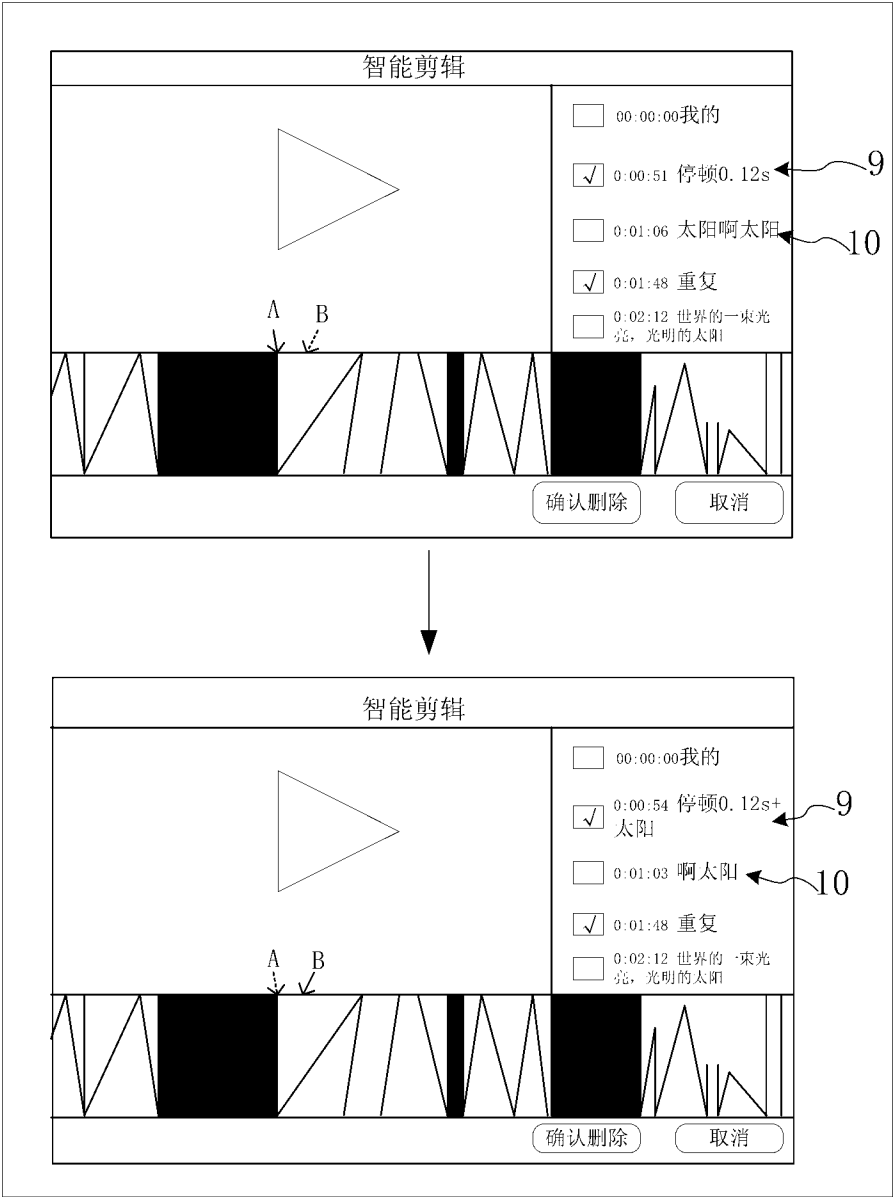


图 6

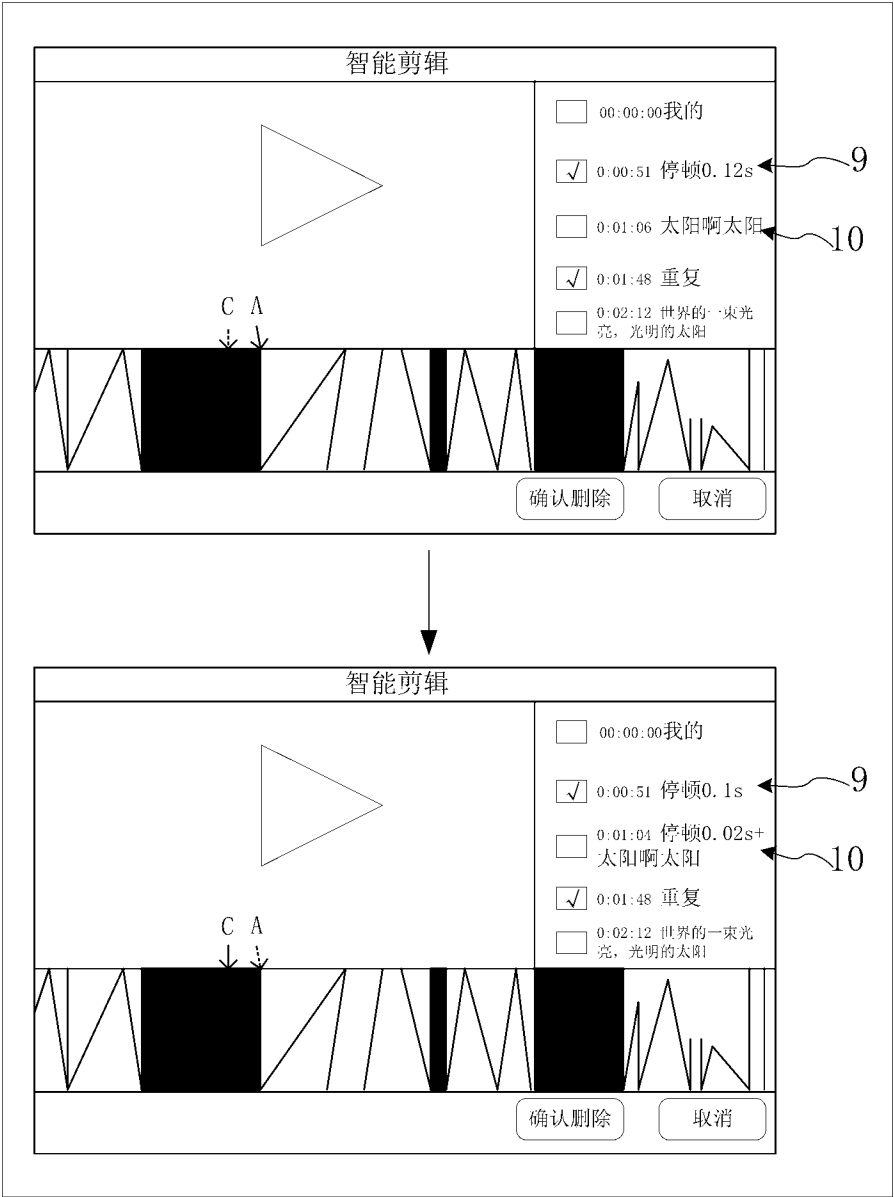


图 7

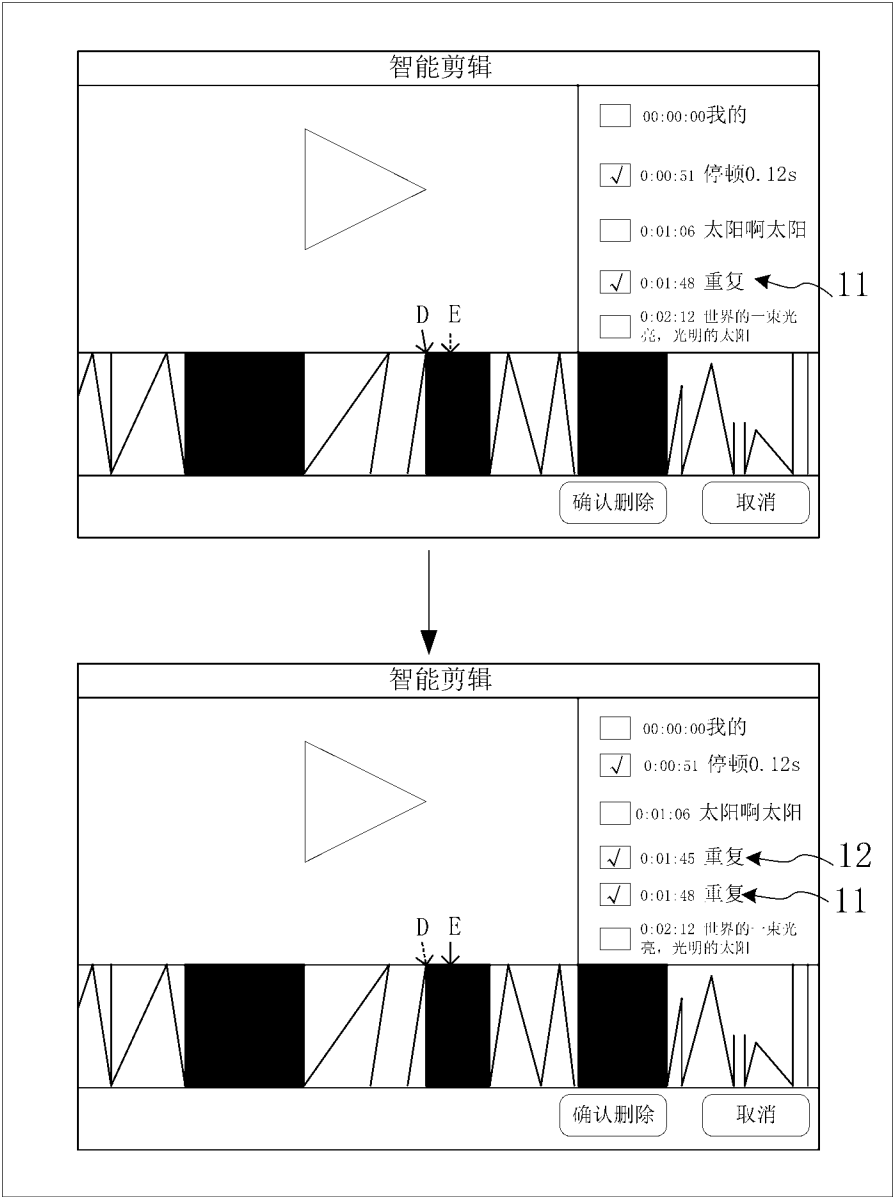


图 8

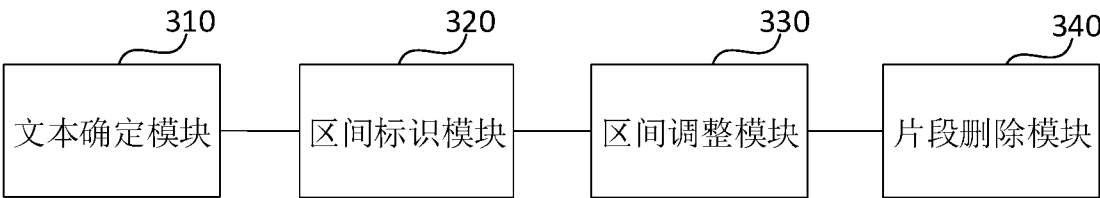


图 9

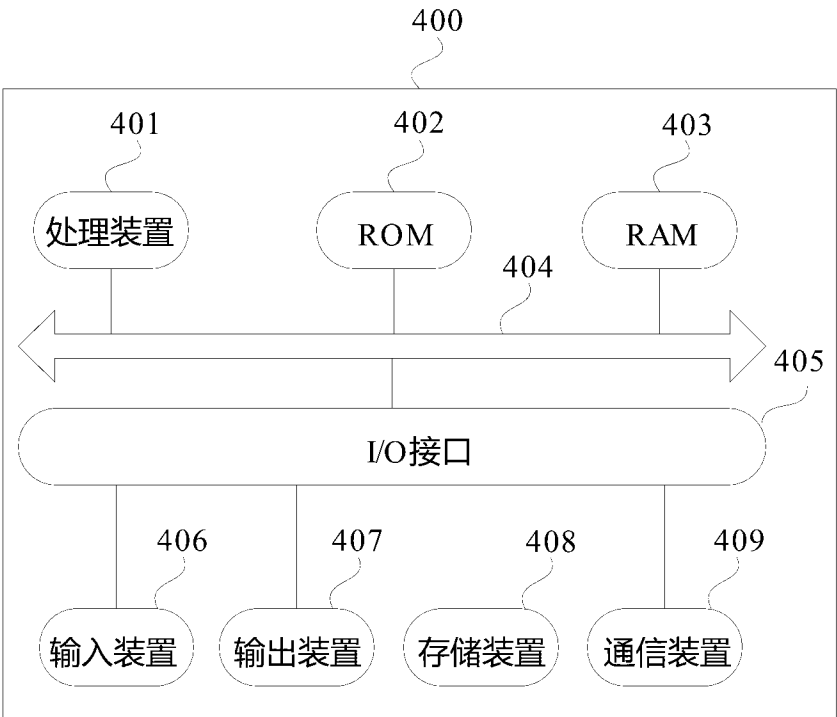


图 10

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2023/138239

A. CLASSIFICATION OF SUBJECT MATTER

H04N21/44(2011.01)i; H04N21/439(2011.01)i; G10L15/26(2006.01)i; G10L15/18(2013.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC: H04N, G10L

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNTXT, CNABS, WPABS, DWPI, ENTXT, IEEE, CNKI: 视频, 音频, 语音, 文本, 剪辑, 编辑, 删除, 无效, 重复, 停顿, 磕巴, 区间, 时间, 片段, 时段, 播放, video, audio, sound, voice, speech, text, clip, cut, edit, delete, invalid, useless, repetitive, iterative, pause, holdup, break down, time, period, segment, zone, play

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 113613068 A (BEIJING BYTEDANCE NETWORK TECHNOLOGY CO., LTD.) 05 November 2021 (2021-11-05) description, paragraphs 0036-0097 and 0115, and figures 1-4	1-9, 11-15
X	CN 115623279 A (FACEMON CO., LTD.) 17 January 2023 (2023-01-17) description, paragraphs 0036-0053, 0089 and 0117-0123, and figure 8	1-9, 11-15
X	CN 114999530 A (BEIJING FEIXIANG PLANET TECHNOLOGY CO., LTD.) 02 September 2022 (2022-09-02) description, paragraphs 0038-0044, 0077-0108 and 0160-0175	1, 11-15
A	CN 109040773 A (WUHAN DOUYU NETWORK TECHNOLOGY CO., LTD.) 18 December 2018 (2018-12-18) entire document	1-15
A	US 2012323897 A1 (MICROSOFT CORP.) 20 December 2012 (2012-12-20) entire document	1-15



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“D” document cited by the applicant in the international application

“E” earlier application or patent but published on or after the international filing date

“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

“O” document referring to an oral disclosure, use, exhibition or other means

“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

“&” document member of the same patent family

Date of the actual completion of the international search

07 February 2024

Date of mailing of the international search report

21 February 2024

Name and mailing address of the ISA/CN

China National Intellectual Property Administration (ISA/
CN)
China No. 6, Xitucheng Road, Jimenqiao, Haidian District,
Beijing 100088

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2023/138239

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	113613068	A	05 November 2021	WO	2023011142	A1	09 February 2023
CN	115623279	A	17 January 2023	WO	2023287360	A2	19 January 2023
				WO	2023287360	A3	13 April 2023
CN	114999530	A	02 September 2022	None			
CN	109040773	A	18 December 2018	None			
US	2012323897	A1	20 December 2012	None			

A. 主题的分类

H04N21/44(2011.01)i; H04N21/439(2011.01)i; G10L15/26(2006.01)i; G10L15/18(2013.01)i

按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类

B. 检索领域

检索的最低限度文献(标明分类系统和分类号)

IPC: H04N, G10L

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用))

CNTEXT, CNABS, WPABS, DWPI, ENTXT, IEEE, CNKI: 视频, 音频, 语音, 文本, 剪辑, 编辑, 删除, 无效, 重复, 停顿, 磕巴, 区间, 时间, 片段, 时段, 播放, video, audio, sound, voice, speech, text, clip, cut, edit, delete, invalid, useless, repetitive, iterative, pause, holdup, break down, time, period, segment, zone, play

C. 相关文件

类型*	引用文件, 必要时, 指明相关段落	相关的权利要求
X	CN 113613068 A (北京字跳网络技术有限公司) 2021年11月5日 (2021 - 11 - 05) 说明书第0036-0097,0115段, 图1-4	1-9,11-15
X	CN 115623279 A (脸萌有限公司) 2023年1月17日 (2023 - 01 - 17) 说明书第0036-0053,0089,0117-0123段, 图8	1-9,11-15
X	CN 114999530 A (北京飞象星球科技有限公司) 2022年9月2日 (2022 - 09 - 02) 说明书第0038-0044,0077-0108,0160-0175段	1,11-15
A	CN 109040773 A (武汉斗鱼网络科技有限公司) 2018年12月18日 (2018 - 12 - 18) 全文	1-15
A	US 2012323897 A1 (MICROSOFT CORP.) 2012年12月20日 (2012 - 12 - 20) 全文	1-15

☐ 其余文件在C栏的续页中列出。

☒ 见同族专利附件。

* 引用文件的具体类型:

“A” 认为不特别相关的表示了现有技术一般状态的文件

“D” 申请人在国际申请中引证的文件

“E” 在国际申请日的当天或之后公布的在先申请或专利

“L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件(如具体说明的)

“O” 涉及口头公开、使用、展览或其他方式公开的文件

“P” 公布日先于国际申请日但迟于所要求的优先权日的文件

“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件

“X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性

“Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性

“&” 同族专利的文件

国际检索实际完成的日期

2024年2月7日

国际检索报告邮寄日期

2024年2月21日

ISA/CN的名称和邮寄地址

中国国家知识产权局
中国北京市海淀区蓟门桥西土城路6号 100088

授权官员

许微

电话号码 (+86) 010-53961818

PCT/ISA/210 表(第2页) (2022年7月)

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2023/138239

检索报告引用的专利文件			公布日 (年/月/日)	同族专利			公布日 (年/月/日)
CN	113613068	A	2021年11月5日	WO	2023011142	A1	2023年2月9日
CN	115623279	A	2023年1月17日	WO	2023287360	A2	2023年1月19日
				WO	2023287360	A3	2023年4月13日
CN	114999530	A	2022年9月2日	无			
CN	109040773	A	2018年12月18日	无			
US	2012323897	A1	2012年12月20日	无			