

(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
2023年3月30日 (30.03.2023)



(10) 国际公布号
WO 2023/045716 A1

- (51) 国际专利分类号:
H04N 21/439 (2011.01)
- (21) 国际申请号: PCT/CN2022/115722
- (22) 国际申请日: 2022年8月30日 (30.08.2022)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
202111124169.4 2021年9月24日 (24.09.2021) CN
- (71) 申请人: 北京搜狗科技发展有限公司 (BEIJING SOGOU TECHNOLOGY DEVELOPMENT CO., LTD.) [CN/CN]; 中国北京市海淀区中关村东路1号院9号楼搜狐网络大厦9层01房间, Beijing 100084 (CN)。
- (72) 发明人: 孟凡博(MENG, Fanbo); 中国北京市海淀区中关村东路1号院9号楼搜狐网络大厦9层01房间, Beijing 100084 (CN)。 刘金锁(LIU, Jinsuo); 中国北京市海淀区中关村东路1号院9号楼搜狐网络大厦9层01房间, Beijing 100084 (CN)。 朱伟基

(ZHU, Weiji); 中国北京市海淀区中关村东路1号院9号楼搜狐网络大厦9层01房间, Beijing 100084 (CN)。 张永哲(ZHANG, Yongzhe); 中国北京市海淀区中关村东路1号院9号楼搜狐网络大厦9层01房间, Beijing 100084 (CN)。 丰添(FENG, Tian); 中国北京市海淀区中关村东路1号院9号楼搜狐网络大厦9层01房间, Beijing 100084 (CN)。

(74) 代理人: 北京德琦知识产权代理有限公司 (DEQI INTELLECTUAL PROPERTY LAW CORPORATION); 中国北京市海淀区知春路1号学院国际大厦7层, Beijing 100083 (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, ME, MG, MK, MN, 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 PROCESSING METHOD AND APPARATUS, AND MEDIUM AND PROGRAM PRODUCT

(54) 发明名称: 视频处理方法、装置、介质和程序产品

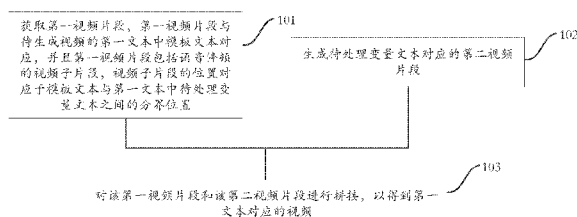


图 1B

(57) Abstract: A video processing method and apparatus, and a medium and a program product. The method specifically comprises: acquiring a first video clip, wherein the first video clip corresponds to template text in first text of a video to be generated, the first video clip comprises a video sub-clip of which voice is paused, and the position of the video sub-clip corresponds to a demarcation position between the template text and variable text to be processed in the first text (101); generating a second video clip corresponding to the variable text to be processed (102); and splicing the first video clip and the second video clip, so as to obtain a video corresponding to the first text (103). By means of the embodiments of the present application, the processing efficiency of a video can be improved.

(57) 摘要: 一种视频处理方法、装置、介质和程序产品, 其中的方法具体包括: 获取第一视频片段, 第一视频片段与待生成视频的第一文本中模板文本对应, 并且第一视频片段包括语音停顿的视频子片段, 视频子片段的位置对应于模板文本与第一文本中待处理变量文本之间的分界位置 (101); 生成待处理变量文本对应的第二视频片段 (102); 对第一视频片段和第二视频片段进行拼接, 以得到第一文本对应的视频 (103)。本申请实施例可以提高视频的处理效率。

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, GH, GM, KE, LR, LS, MW, MZ, NA, RW, 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, 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))。

视频处理方法、装置、介质和程序产品

本申请要求于 2021 年 9 月 24 日提交中国专利局、申请号为 202111124169.4、申请名称为“一种视频处理方法、装置和介质”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。

5 技术领域

本申请涉及通信技术领域，特别是涉及一种视频处理方法、装置、介质和程序产品。

背景技术

10 随着通信技术的发展，虚拟对象可被广泛应用于播报场景、教学场景、医疗场景、客服场景等应用场景。在这些应用场景中，虚拟对象通常需要对文本进行表达，相应地，可以生成并播放虚拟对象对应的视频。该视频可以表征虚拟对象表达文本的过程。视频的生成过程通常包括：语音生成环节和图像序列生成环节。其中，语音生成环节通常采用语音合成技术。图像序列生成环节通常采用图像处理技术。

15 发明人在实施本申请实施例的过程中发现，相关技术针对完整文本，生成对应的完整视频，通常会耗费较多的时间成本，导致视频的处理效率较低。

发明内容

如何提高视频的处理效率，是本领域技术人员需要解决的技术问题。鉴于上述问题，本申请实施例提出了一种克服上述问题或者至少部分地解决上述问题的视频处理方法、装置、介质和程序产品。

20 为了解决上述问题，本申请公开了一种视频处理方法，在电子设备中执行，所述方法包括：

获取第一视频片段，所述第一视频片段与待生成视频的第一文本中模板文本对应，并且所述第一视频片段包括语音停顿的视频子片段，所述视频子片段的位置对应于所述模板文本与所述第一文本中待处理变量文本之间的分界位置；

生成所述待处理变量文本对应的第二视频片段；

25 对所述第一视频片段和所述第二视频片段进行拼接，以得到所述第一文本对应的视频。

另一方面，本申请公开了一种视频处理装置，包括：

提供模块，用于获取第一视频片段，所述第一视频片段与待生成视频的第一文本中模板文本对应，并且所述第一视频片段包括语音停顿的视频子片段，所述视频子片段的位置对应于所述模板文本与所述第一文本中待处理变量文本之间的分界位置；

30 生成模块，用于生成所述待处理变量文本对应的第二视频片段；

拼接模块，用于对所述第一视频片段和所述第二视频片段进行拼接，以得到所述第一文本对应的视频。

35 再一方面，本申请公开了一种用于视频处理的装置，包括有存储器，以及一个或者一个以上的程序，其中一个或者一个以上程序存储于存储器中，所述程序被一个或者一个以上处理器执行时，实现前述方法的步骤。

又一方面，本申请实施例公开了一个或多个机器可读介质，其上存储有指令，当由一个或多个处理器执行时，使得装置执行如前述一个或多个所述的方法。

40 又一方面，本申请实施例公开了一种计算机程序产品，该程序产品包括计算机指令，该计算机指令存储在计算机可读存储介质中；当处理器执行该计算机指令时，使得处理器执行本申请实施例的视频处理方法。

附图说明

图 1A 示出了根据本申请实施例的应用场景的示意图；
图 1B 是本申请实施例的一种视频处理方法的流程图；
图 2 是本申请实施例的一种视频处理方法的流程图；
图 3 是本申请实施例的一种视频处理装置的结构框图；
5 图 4 是本申请实施例的一种用于视频处理的装置的结构框图；及
图 5 是本申请一些实施例中服务端的结构框图。

具体实施方式

10 为使本申请的上述目的、特征和优点能够更加明显易懂，下面结合附图和具体实施方式对本申请作进一步详细的说明。

本申请实施例中，虚拟对象是通过对象建模、动作捕捉等技术得到的一种生动自然的、与真实对象接近的虚拟对象，通过语音识别、自然语言理解等人工智能技术，能够使得虚拟对象具备认知、或者理解、或者表达等能力。虚拟对象具体包括：虚拟人物、或虚拟动物、或二维卡通对象、或三维卡通对象等。

15 例如，在播报场景下，虚拟对象可以代替例如媒体工作者进行新闻播报、或者游戏解说等。又如，在医疗场景下，虚拟对象可以代替例如医学工作者进行医学指导等。

在具体实现中，虚拟对象可以对文本进行表达。而本申请实施例可以生成文本和虚拟对象对应的视频。该视频具体可以包括：文本对应的语音序列、以及语音序列对应的图像帧序列。

20 在一些应用场景中，待生成视频的文本具体包括：模板文本和变量文本。其中，模板文本相对固定，变量文本通常会根据用户输入等预设因素发生变化。

例如，变量文本可以根据用户输入确定。以医疗场景为例，可以根据用户输入中包含的疾病名称，确定对应的变量文本。可选地，变量文本对应的字段具体包括：疾病名称字段、食物种类字段、食材数量字段等，可以根据用户输入中包含的疾病名称，确定这些字段。

25 可以理解，本领域技术人员可以根据实际应用需求，确定文本中的变量文本，本申请实施例对于变量文本的具体确定方式不加以限制。

为了使视频质量符合要求，相关技术在变量文本发生改变的情况下，通常针对改变后的完整文本，生成对应的完整视频。然而，针对改变后的完整文本，生成对应的完整视频，通常会耗费较多的时间成本，导致视频的处理效率较低。

30 针对如何提高视频的处理效率的技术问题，本申请实施例提供了一种视频处理方案，该方案具体包括：获取第一视频片段；与待生成视频的第一文本中模板文本对应，并且所述第一视频片段包括语音停顿的视频子片段，所述视频子片段的位置对应于所述模板文本与所述第一文本中待处理变量文本之间的分界位置第一文本包括模板文本和待处理变量文本；生成待处理变量文本对应的第二视频片段；对该第一视频片段和该第二视频片段进行拼接，以得到第一文本对应的视频。

35 本申请实施例对模板文本对应的第一视频片段与待处理变量文本对应的第二视频片段进行拼接。其中，第一视频片段可以为预先保存的视频片段，在视频处理过程中可以生成待处理变量文本对应的第二视频片段。由于待处理变量文本的长度小于完整文本的长度，故本申请实施例能够缩短生成视频的长度和对应的时间成本，因此，能够提高视频的处理效率。

40 并且，本申请实施例的第一视频片段包括语音停顿的视频子片段。这里语音停顿是指语音停止，例如为虚拟对象不说话。视频子片段的位置对应于模板文本与第一文本中待处理变量文本之间的分界位置。上述第一视频片段中语音停顿的视频子片段，有助于克服拼接位置处的跳变或抖动问题，因此能够提高拼接位置处的连续性。

本申请实施例提供的视频处理方法可应用于客户端和服务端对应的应用场景中。例如，图 1A 示出了根据本申请实施例的应用场景的示意图。客户端与服务端位于有线或无线网络中，通过该有线或无线网络，客户端与服务端进行数据交互。

45 客户端和服务端可以统称为电子设备。客户端例如包括但不限于：智能手机、平板电脑、电子书阅读器、MP3 (动态影像专家压缩标准音频层面 3, Moving Picture Experts Group Audio Layer III) 播放器、

MP4(动态影像专家压缩标准音频层面 4, Moving Picture Experts Group Audio Layer IV)播放器、膝上型便携计算机、车载电脑、台式计算机、机顶盒、智能电视机、可穿戴设备等等。服务端例如为硬件独立的服务器、虚拟服务器或者服务器集群等设备。

5 客户端是指与服务端相对应, 为用户提供本地服务的程序。本申请实施例中的客户端可以接收用户输入, 并提供该用户输入对应的视频。该视频可由客户端或服务端生成, 本申请实施例对于视频的具体生成主体不加以限制。

10 在本申请的一种实施例中, 客户端可以接收用户输入, 并向服务端上传用户输入, 以使服务端生成用户输入对应的视频。服务端可以根据用户输入确定待处理变量文本, 生成待处理变量文本对应的第二视频片段, 并对预先保存的第一视频片段和该第二视频片段进行拼接, 以得到该模板文本和该待处理变量文本对应的视频。

方法实施例一

参照图 1B, 示出了本申请的一种视频处理方法的流程图, 具体可以包括如下步骤。视频处理方法例如可以由电子设备执行。

15 步骤 101、获取第一视频片段, 第一视频片段与待生成视频的第一文本中模板文本对应, 并且第一视频片段包括语音停顿的视频子片段, 视频子片段的位置对应于模板文本与第一文本中待处理变量文本之间的分界位置。。

步骤 102、生成待处理变量文本对应的第二视频片段。

步骤 103、对该第一视频片段和该第二视频片段进行拼接, 以得到第一文本对应的视频。

20 在一个实施例中, 步骤 101, 可以预先生成并保存模板文本对应的第一视频片段。第一视频片段包括语音停顿的视频子片段。这里语音停顿是指语音停止或者暂时不输出语音。语音停顿的视频子片段可以认为是没有语音的视频子片段。视频子片段的位置对应于模板文本与第一文本中待处理变量文本之间的分界位置, 该视频子片段能够提高拼接位置处的连续性。

本申请实施例的文本的结构具体包括: 模板文本和变量文本。分界位置可用于对相邻的模板文本和变量文本进行分割。

25 以文本 A“关于 <糖尿病> 和 <水果> 的问题, 我还在研究。我想这份 <糖尿病> 的饮食建议可能也对你有帮助, 里面包含了约 <1800> 种食材的推荐、禁忌, 请你点击查看”为例, 文本 A 中存在多个分界位置。例如, 模板文本“关于”与变量文本“<糖尿病>”之间对应应有分界位置, 变量文本“<糖尿病>”与模板文本“和”之间对应应有分界位置, 模板文本“和”与变量文本“<水果>”之间对应应有分界位置, 变量文本“<水果>”与模板文本“的”之间对应应有分界位置, 等等。

30 在一种实现方式中, 第一视频片段的确定过程可以包括: 根据模板文本、预设变量文本、以及对应分界位置处的停顿信息, 生成预设视频; 从上述预设视频中截取上述模板文本对应的第一视频片段。

其中, 预设变量文本可以为任意的变量文本, 或者, 预设变量文本可以为变量文本的任意实例。

本申请实施例可以根据模板文本和预设变量文本对应的预设完整文本, 生成预设视频, 其中, 预设视频的生成过程中可以考虑分界位置处的停顿信息。停顿信息例如表示预定时长的语音停顿。

35 在实际应用中, 预设视频可以包括: 语音部分对应的预设语音和图像部分对应的预设图像序列。

在具体实现中, 可以利用 TTS (语音合成, Text To Speech) 技术, 将预设完整文本转换为预设语音。预设语音可以表征为波形的形式。

40 本申请实施例的将预设完整文本转换为预设语音, 具体包括: 语言分析环节和声学系统环节。其中, 语言分析环节用于根据预设完整文本及其对应的停顿信息, 生成对应的语言学信息; 声学系统环节主要是根据语音分析环节提供的语言学信息, 生成对应的预设语音, 实现发声的功能。

在一种实现方式中, 语言分析环节的处理具体可以包括: 文本结构与语种判断、文本标准化、文本转音素和韵律预测。语言学信息可以是语音分析环节的结果。

45 其中, 文本结构与语种判断, 用于判断预设完整文本的语种, 例如中文、英文、藏语、维语等语种, 并根据对应语种的语法规则, 把预设完整文本切分为语句, 并将切分好的语句传到后面的处理模块。

文本标准化, 用于根据设置好的规则, 对切分好的语句进行标准化。

文本转音素，用于确定语句对应的音素特征。

由于人类在语言表达的时候通常带着语气与感情，语音合成的目的往往是为了模仿真实的人声；因此，韵律预测可用于确定语句的什么地方需要停顿，停顿多久，哪个字或者词语需要重读，哪个词需要轻读等，进而实现声音的高低曲折，抑扬顿挫。

5 本申请实施例可以首先利用韵律预测技术，确定韵律预测结果，然后，根据停顿信息，对韵律预测结果进行更新。

10 以文本 A 为例，停顿信息可以为：在模板文本“关于”与变量文本“<糖尿病>”之间添加预设时长的停顿信息，则对韵律预测结果进行更新具体可以包括：在模板文本“关于”的音素特征“guan”、“yu”和变量文本“<糖尿病>”的音素特征“tang”、“niao”、“bing”之间添加预设时长的停顿信息，更新后的韵律预测结果可以为：“guan”、“yu”、“停顿 N 毫秒”、“tang”、“niao”、“bing”等。其中，N 可以为大于 0 的自然数，N 的值可由本领域技术人员根据实际应用需求确定。

声学系统环节可以按照语音合成参数，得到符合需求的预设语音。

15 可选地，语音合成参数可以包括：音色参数。音色参数可以指不同的声音的频率表现在波形方面的与众不同的特性，通常不同的发声体对应不同的音色，因此可以按照音色参数，得到与目标发声体的音色相匹配的语音序列。目标发声体可由用户指定，例如，目标发声体可以为指定的医学工作者等。在实际应用中，可以依据目标发声体的预设长度的音频，得到目标发声体的音色参数。

图像部分对应的预设图像序列可以在虚拟对象图像的基础上得到，换言之，本申请实施例可以为虚拟对象图像赋予状态特征，以得到预设图像序列。虚拟对象图像可由用户指定，例如，虚拟对象图像可以为知名人物（例如主持人）的图像。

20 上述状态特征可以包括如下特征中的至少一种：

表情特征；

唇部特征；以及

肢体特征。

表情，表达感情、情意，可以指表现在面部的思想感情。

25 表情特征通常是针对整个面部的。唇部特征可以专门针对唇部，而且跟文本的文本内容、语音、发音方式等都有关系，因此可以提高预设图像序列所对应表达的自然度。

肢体特征可以通过头、眼、颈、手、肘、臂、身、胯、足等人体部位的协调活动来传达人物的思想，形象地借以表情达意。肢体特征可以包括：转头、耸肩、手势等，可以提高图像序列所对应表达的丰富度。例如，说话时至少一个手臂自然下垂，不说话时至少一个手臂自然放在腹部等。

30 本申请实施例在生成预设视频的图像部分的过程中，可以根据预设完整文本及停顿信息，确定图像参数，该图像参数可以表征虚拟对象的状态特征；并根据图像参数生成图像部分对应的预设图像序列。

35 其中，图像参数可以包括：停顿图像参数，该停顿图像参数可以表征停顿信息对应的停顿状态特征。换言之，停顿图像参数表示虚拟对象停止说话时，虚拟对象表现出的形体、表情等方面的状态特征。相应地，预设图像序列中可以包括：停顿状态特征对应的图像序列。例如，停顿状态特征可以包括：中性表情、唇部闭合状态、以及手臂下垂状态等。

在生成预设语音和预设图像序列后，可以对预设语音和预设图像序列进行融合，得到对应的预设视频。

40 在得到预设视频后，可以从上述预设视频中截取上述模板文本对应的第一视频片段。具体地，可以根据预设变量文本在预设视频中的起始位置和结束位置，进行第一视频片段的截取。

45 以文本 A 为例，假设预设变量文本“<糖尿病>”在文本中的起始位置对应于预设视频中的起始位置 T1，预设变量文本“<糖尿病>”的结束位置对应于预设视频中的结束位置 T2，则可以从预设视频中截取 T1 之前的视频片段，作为模板文本“关于”对应的第一视频片段。需要说明的是，在生成预设视频的过程中利用了分界位置处的停顿信息，故 T1 之前的第一视频片段带有停顿信息（即第一视频片段包括语音停顿的视频子片段），因此能够提高后续拼接过程中拼接位置处的连续性。

以文本 A 为例，假设预设变量文本“<水果>”在文本中的起始位置对应预设视频中的起始位置 T3，

预设变量文本“<水果>”在文本中的起始位置对应预设视频中的结束位置 T4，则可以从预设视频中截取 T2 与 T3 之间的视频片段，作为模板文本“和”对应的第一视频片段。

由于预设完整文本中的模板文本被预设变量文本分割为多处，故在实际应用中，可以从预设视频中分别提取多处模板文本对应的第一视频片段。

5 可以理解，上述在生成预设视频的过程中利用了分界位置处的停顿信息、以获得第一视频片段的获取方式，只是作为可选实施例，实际上，本领域技术人员还可以根据实际应用需求，采用其他获取方式。

在一个实施例中，第一视频片段中视频子片段不仅语音停顿，并且视频子片段的图像中虚拟对象处于不說話的状态。

10 在一个实施例中，所述视频子片段为经过停顿处理后得到的子片段。

对视频子片段的停顿处理，包括：

对所述第一视频片段中与所述分界位置对应的拼接位置处的语音信号子片段与静音信号进行加权处理，以得到语音停顿的语音信号子片段；

15 对第一视频片段在拼接位置处的图像子序列与目标状态特征的图像序列进行加权处理，以得到虚拟对象处于不說話的状态的所述图像子序列，其中目标状态特征为表示虚拟对象处于不說話状态的特征。这样，语音停顿的语音信号子片段和虚拟对象处于不說話的状态的图像子序列可以组成所述视频子片段。

20 在一个实施例中，第一视频片段的一种获取方式可以包括：根据模板文本和预设变量文本，生成第一视频；从上述第一视频中截取上述模板文本对应的第一视频片段；在分界位置处对所述第一视频片段进行停顿处理。

以语音部分的停顿处理为例，可以视频片段在分界位置处的语音信号子片段与静音信号进行加权处理，以实现语音部分的停顿处理。以图像部分的停顿处理为例，可以对视频片段在分界位置处的图像子序列与停顿信息对应目标状态特征的图像序列进行加权处理，以实现图像部分的停顿处理。

25 在获得第一视频片段后，可以对第一视频片段进行保存，以在变量文本发生变化的情况下，对第一视频片段与改变后的变量文本（以下简称待处理变量文本）对应的第二视频片段进行拼接。

步骤 102 中，待处理变量文本可以根据用户输入得到。可以理解，本申请实施例对于待处理变量文本的具体确定方式不加以限制。

本申请实施例可以提供生成待处理变量文本对应的第二视频片段的如下技术方案：

技术方案 1、

30 技术方案 1 中生成待处理变量文本对应的第二视频片段，具体包括：针对待处理变量文本在第一文本中所处的语句，确定对应的语音参数和图像参数，其中，图像参数表征所述第一文本对应的视频中要出现的虚拟对象的状态特征，语音参数用于表征语音合成对应的参数；从所述语音参数和图像参数中，提取所述待处理变量文本对应的目标语音参数和目标图像参数；根据目标语音参数和目标图像参数，生成待处理变量文本对应的第二视频片段。

35 技术方案 1 首先以待处理变量文本所在的语句为单位，确定对应的语音参数和图像参数，然后从语音参数和图像参数中，提取所述待处理变量文本对应的目标语音参数和目标图像参数。

语句是一个语法上自成体系的单位，它由一个词或句法上有关联的一组词构成，表达一种主张、疑问、命令、愿望或感叹。

40 在待处理变量文本对应词的情况下，语句中通常既包含模板文本又包含待处理变量文本。由于语句对应的语音参数和图像参数具有一定的连续性，故从中提取的待处理变量文本对应的目标语音参数和目标图像参数与语句中模板文本对应的语音参数和图像参数具有一定的连续性；在此基础上，能够提高待处理变量文本对应的第二视频片段与语句中模板文本对应的第一视频片段之间的连续性，进而能够提高拼接位置处的连续性。

45 在实际应用中，语音参数可以表征语音合成所对应的参数。语音参数可以包括：语言特征和/或声学特征。

语言特征可以包括：音素特征。音素是根据语音的自然属性划分出来的最小语音单位，依据音节

里的发音动作来分析，一个动作构成一个音素。音素可以包括：元音与辅音。

声学特征可以从发声角度表征语音的特征。

声学特征可以包括但不限于如下特征：

韵律学特征（超音段特征/超语言学特征），具体包括时长相关特征、基频相关特征、能量相关特征等；

音质特征；

基于谱的相关性分析特征，其是声道形状变化和发声运动之间相关性的体现，目前基于谱的相关特征主要包括：线性预测倒谱系数(LPCC, LinearPredictionCoefficients)、梅尔频率倒谱系数(MFCC, Mel Frequency Cepstrum Coefficient)等。

可以理解，上述语音参数只是作为示例，本申请实施例对具体的语音参数不加以限制。

在具体实现中，可以根据目标语音参数，对待处理变量文本进行语音合成，以将待处理变量文本转换为目标语音。

图像参数可以为图像序列的生成所对应的参数。图像参数可用于确定虚拟对象对应的状态特征，或者，图像参数可以包括：虚拟对象对应的状态特征。例如，图像参数可以包括唇部特征。

在具体实现中，可以为虚拟对象图像赋予目标图像参数对应的状态特征，以得到目标图像序列。对目标语音和目标图像序列进行融合，可以得到第二视频片段。

技术方案2、

技术方案2中生成待处理变量文本对应的第二视频片段，具体包括：根据预设变量文本在边界位置处的预设图像参数，对待处理变量文本对应的目标图像参数进行平滑处理，以提高所述目标图像参数与所述模板文本的图像参数在边界位置处的连续性；根据平滑处理后的目标图像参数，生成所述待处理变量文本对应的第二视频片段。

技术方案2根据预设变量文本在边界位置处的预设图像参数，对待处理变量文本对应的目标图像参数进行平滑处理。由于预设变量文本在边界位置处的预设图像参数与模板文本在边界位置处的图像参数具有一定的连续性，故上述平滑处理能够提高平滑处理后的目标图像参数与模板文本的图像参数在边界位置处的连续性；在此基础上，能够提高待处理变量文本对应的第二视频片段与语句中模板文本对应的第一视频片段之间的连续性，进而能够提高拼接位置处的连续性。

在具体实现中，可以利用汉宁窗等窗函数，根据预设图像参数，对待处理变量文本对应的目标图像参数进行平滑处理。可以理解，本申请实施例对于具体的平滑处理过程不加以限制。

根据前面的介绍，本申请实施例在生成预设视频的图像部分的过程中，可以根据预设完整文本及停顿信息，确定图像参数，本申请实施例可以从图像参数中提取预设变量文本在边界位置处的预设图像参数，并对该预设图像参数进行保存。

以文本A为例，假设预设变量文本“<糖尿病>”的起始位置对应于预设视频中的起始位置T1，预设变量文本“<糖尿病>”的起始位置对应于在预设视频中的结束位置T2，则可以提取T1至T2之间的图像参数，作为预设变量文本“<糖尿病>”在边界位置处的预设图像参数。

技术方案3、

技术方案3中，视频对应的图像序列包括：背景图像序列和运动图像序列，则生成待处理变量文本对应的第二视频片段，具体包括：生成待处理变量文本对应的目标运动图像序列；根据预设背景图像序列，确定待处理变量文本对应的目标背景图像序列；对上述目标运动图像序列和上述目标背景图像序列进行融合，以得到所述待处理变量文本对应的第二视频片段。

在实际应用中，可以将视频对应的图像序列分解为两部分。第一部分是：运动图像序列，可用于表征虚拟对象表达时运动的部分，通常对应唇部、眼部、手臂部位等预设部位。第二部分是：背景图像序列，可用于表征虚拟对象表达时相对静止的部分，通常对应除了预设部位之外的部分。

在具体实现中，背景图像序列可以为预置得到。例如，可以预置预设时长的预设背景图像序列，并在图像序列中对预设背景图像序列进行循环布置（也可以称为循环出现）。可以根据待处理变量文本对应的目标图像参数，生成运动图像序列。

在实际应用中，可以对运动图像序列和背景图像序列进行融合，以得到图像序列。例如，可以将

运动图像序列贴到背景图像序列之上，以得到图像序列。

技术方案 3 根据变量文本对应的预设背景图像序列，确定待处理变量文本对应的目标背景图像序列，可以提高目标背景图像序列与预设背景图像序列之间的匹配度，进而能够提高待处理变量文本对应目标背景图像序列与模板文本对应背景图像序列之间的匹配度和连续性。

5 根据前面的介绍，本申请实施例在生成预设视频的图像部分的过程中，可以对预设变量文本对应预设背景图像序列的信息进行记录。例如，预设背景图像序列的信息可以包括：预设背景图像序列在预设视频中的起始帧标识和结束帧标识等。例如，预设背景图像序列的信息可以包括：起始帧编号 100 和结束帧编号 125 等。

10 在一种实施方式中，为了提高目标背景图像序列与预设背景图像序列在起始位置处或结束位置处的匹配度，所述目标背景图像序列的位于首尾位置的背景图像，与所述预设背景图像序列的位于首尾位置的背景图像相匹配。

首位置可以指起始位置，尾位置可以指结束位置。具体地，目标背景图像序列的位于首位置的背景图像，与预设背景图像序列的位于首位置的背景图像相匹配。或者，目标背景图像序列的位于尾位置的背景图像，与预设背景图像序列的位于尾位置的背景图像相匹配。

15 由于预设背景图像序列与模板文本对应背景图像序列在分界位置处是匹配的和连续的，故在目标背景图像序列与预设背景图像序列在分界位置处相匹配的情况下，也能够提高目标背景图像序列与模板文本对应背景图像序列在拼接位置处的匹配度和连续性。

为了实现目标背景图像序列与预设背景图像序列在分界位置处相匹配，上述确定所述待处理变量文本对应的目标背景图像序列所采用的确定方式，具体可以包括：

20 确定方式 1、在预设背景图像序列对应的图像数量 N1 与目标运动图像序列对应的图像数量 N2 相匹配的情况下，将所述预设背景图像序列确定为目标背景图像序列；或者

确定方式 2、在预设背景图像序列对应的图像数量 N1 大于目标运动图像序列对应的图像数量 N2 的情况下，从所述预设背景图像序列中丢弃位于中间位置的第一背景图像；在丢弃至少两帧第一背景图像的情况下，至少两帧第一背景图像在预设背景图像序列中不连续分布；或者

25 确定方式 3、在预设背景图像序列对应的图像数量 N1 小于目标运动图像序列对应的图像数量 N2 的情况下，在预设背景图像序列的基础上增加第二背景图像。

对于确定方式 1，在 N1 与 N2 相等的情况下，将预设背景图像序列确定为目标背景图像序列，可以实现目标背景图像序列与预设背景图像序列在分界位置处的匹配。

30 在实际应用中，可以根据待处理变量文本对应的语音时长信息，确定目标运动图像序列对应的图像数量 N2。该语音时长信息可以根据待处理变量文本对应的语音参数确定，或者，该语音时长信息可以根据待处理变量文本对应语音片段的时长确定。

对于确定方式 2，在 N1 大于 N2 的情况下，从预设背景图像序列中丢弃位于中间位置的第一背景图像，可以实现目标背景图像序列与预设背景图像序列在分界位置处的匹配。

35 中间位置可以不同于首位置或尾位置。且丢弃的至少两帧第一背景图像在预设背景图像序列中不连续分布；这样，可以在一定程度上避免丢弃连续的背景图像导致的背景图像连续性差的问题。

在实际应用中，第一背景图像的数量可以与 N1 与 N2 的差值相匹配。例如，预设背景图像序列的信息可以包括：起始帧编号 100 和结束帧编号 125 等，N1 的值为 26，假设目标运动图像序列对应的图像数量 N2 为 24，则可以从预设背景图像序列中丢弃位于中间位置、且位置不连续的 2 帧第一背景图像。

40 对于确定方式 3，在 N1 小于 N2 的情况下，在预设背景图像序列的基础上增加第二背景图像，可以实现目标背景图像序列与预设背景图像序列在分界位置处的匹配。

在本申请的一种可选实施例中，第二背景图像可以源自预设背景图像序列，换言之，可以从预设背景图像序列中确定出待增加的第二背景图像。

45 在一种实现方式中，可以首先按照正向的顺序，将预设背景图像序列确定为目标背景图像序列的第一部分；然后按照倒向的顺序，将预设背景图像序列确定为目标背景图像序列的第二部分；接着按照正向的顺序，将预设背景图像序列确定为目标背景图像序列的第三部分；其中，第三部分的结束帧

与预设背景图像序列的结束帧相匹配。

例如，预设背景图像序列的信息可以包括：起始帧编号 100 和结束帧编号 125 等，N1 的值为 26，假设目标运动图像序列对应的图像数量 N2 为 30，则目标背景图像序列的第一部分对应的帧编号可以为：100→125，目标背景图像序列的第二部分对应的帧编号可以为：125→124，目标背景图像序列的第三部分对应的帧编号可以为：124→125。

在本申请的另一种可选实施例中，第二背景图像可以源自预设背景图像序列以外的背景图像序列，例如，可以从预设背景图像序列之后的背景图像序列中确定第二背景图像。

在一种实现方式中，可以首先按照正向的顺序，将预设背景图像序列确定为目标背景图像序列的第一部分；然后按照正向的顺序，将预设背景图像序列后续的背景图像序列确定为目标背景图像序列的第二部分；接着按照倒向的顺序，将预设背景图像序列后续的背景图像序列和预设背景图像序列的结束帧确定为目标背景图像序列的第三部分；其中，第三部分的结束帧与预设背景图像序列的结束帧相匹配。

例如，预设背景图像序列的信息可以包括：起始帧编号 100 和结束帧编号 125 等，N1 的值为 26，假设目标运动图像序列对应的图像数量 N2 为 30，则目标背景图像序列的第一部分对应的帧编号可以为：100→125，目标背景图像序列的第二部分对应的帧编号可以为：126→127，目标背景图像序列的第三部分对应的帧编号可以为：127→125。

可以理解，上述在预设背景图像序列的基础上增加第二背景图像的实现方式，只是作为示例，实际上本领域技术人员可以根据实际应用需求，采用其他实现方式，任意的能够实现目标背景图像序列与预设背景图像序列在分界位置处的匹配的的实现方式，均在本申请实施例的实现方式的保护范围之内。

例如，在一种其他实现方式中，还可以确定倒向的目标背景图像序列。相应的确定过程可以包括：首先按照倒向的顺序，将预设背景图像序列确定为目标背景图像序列的第一部分；然后按照正向的顺序，将预设背景图像序列确定为目标背景图像序列的第二部分；接着按照倒向的顺序，将预设背景图像序列确定为目标背景图像序列的第三部分；其中，第三部分的起始帧与预设背景图像序列的起始帧相匹配。

例如，预设背景图像序列的信息可以包括：起始帧编号 100 和结束帧编号 125 等，N1 的值为 26，假设目标运动图像序列对应的图像数量 N2 为 30，则目标背景图像序列的第一部分对应的帧编号可以为：125→100，目标背景图像序列的第二部分对应的帧编号可以为：100→101，目标背景图像序列的第三部分对应的帧编号可以为：101→100。此种情况下得到目标背景图像序列的帧编号可以为：100→101→101→100→100→125。

以上通过技术方案 1 至技术方案 3 对生成待处理变量文本对应的第二视频片段的过程进行了详细介绍，可以理解，本领域技术人员可以根据实际应用需求，采用技术方案 1 至技术方案 3 中的任一或组合，本申请实施例对生成待处理变量文本对应的第二视频片段的具体过程不加以限制。

步骤 103 中，对该第一视频片段和该第二视频片段进行拼接，可以得到第一文本对应的视频。

在本申请的一种可选实施例中，第一视频片段具体可以包括：第一语音片段，第二视频片段具体可以包括：第二语音片段；

则上述对所述第一视频片段和所述第二视频片段进行拼接，具体可以包括：对第一语音片段和第二语音片段各自在拼接位置处的语音子片段进行平滑处理；对平滑处理后的第一语音片段和平滑处理后第二语音片段进行拼接。

本申请实施例首先对第一语音片段和第二语音片段各自在拼接位置的语音子片段进行平滑处理，然后对平滑处理后的第一语音片段和平滑处理后第二语音片段进行拼接。上述平滑处理能够提高平滑处理后的第一语音片段与第二语音片段之间的连续性，因此能够提高第一视频片段和第二视频片段在拼接位置处的连续性。

在实际应用中，可以对拼接得到的视频进行输出，例如输出给用户。以医疗场景为例，可以根据用户输入中包含的疾病名称，确定对应的待处理变量文本，利用图 1B 所示方法实施例，得到视频，并向用户提供该视频。

综上，本申请实施例的视频处理方法，对模板文本对应的第一视频片段与待处理变量文本对应的

第二视频片段进行拼接。其中，第一视频片段可以为预先保存的视频片段，在视频处理过程中可以生成待处理变量文本对应的第二视频片段。由于待处理变量文本的长度小于完整文本的长度，故本申请实施例能够缩短生成视频的长度和对应的时间成本，因此，能够提高视频的处理效率。

5 并且，本申请实施例的第一视频片段在模板文本与变量文本之间的分界位置处，设置有：经过停顿处理的视频子片段。上述停顿处理能够在一定程度上克服拼接位置处的跳变或抖动问题，因此能够提高拼接位置处的连续性。

方法实施例二

参照图2，示出了本申请实施例的一种视频处理方法的流程图，具体可以包括如下步骤。

10 步骤201、根据模板文本、预设变量文本、以及分界位置处对应的停顿信息，生成预设视频停顿信息表示预定时长的语音停顿；

步骤202、从上述预设视频中截取上述模板文本对应的第一视频片段，并对第一视频片段进行保存；

步骤203、根据预设视频的信息，保存预设变量文本在边界位置处的预设图像参数、以及预设变量文本对应预设背景图像序列的信息；

15 步骤201至步骤203，可用于基于生成的预设视频，预先保存第一视频片段、预设变量文本在边界位置处的预设图像参数、以及预设变量文本对应预设背景图像序列的信息。

步骤204至步骤211，可用于根据预先保存的信息，生成待处理变量文本对应的第二视频片段；并对预先保存的第一视频片段与第二视频片段进行拼接。

步骤204、针对待处理变量文本所在的语句，确定对应的语音参数和图像参数；

20 步骤205、从所述语音参数和图像参数中，提取所述待处理变量文本对应的目标语音参数和目标图像参数；

步骤206、根据预设图像参数，对所述待处理变量文本对应的目标图像参数进行平滑处理；

步骤207、根据目标语音参数和平滑处理后的目标图像参数，生成所述待处理变量文本对应的目标运动图像序列；

25 步骤208、根据预设背景图像序列，确定所述待处理变量文本对应的目标背景图像序列；

步骤209、对所述目标运动图像序列和所述目标背景图像序列进行融合，以得到所述待处理变量文本对应的第二视频片段；

步骤210、对第一视频片段中第一语音片段和第二视频片段中第二语音片段，各自在所述分界位置的语音子片段进行平滑处理；

30 步骤211、根据平滑处理后的第一语音片段和平滑处理后的第二语音片段，对第一视频片段和第二视频片段进行拼接。

在本申请的一种应用示例中，假设预设完整文本为前述的文本A，预设变量文本为文本A中的“<糖尿病>”、“<水果>”、“<1800>”等，则可以根据文本A及对应的停顿信息，生成预设视频，并对预设视频中的第一视频片段、预设变量文本在边界位置处的预设图像参数、以及预设变量文本对应预设背景图像序列的信息进行保存。

35 在实际应用中，用户输入等因素可能导致变量文本的变化。例如，在文本A变为文本B“关于<冠心病>和<蔬菜>的问题，我还在研究。我想这份<冠心病>的饮食建议可能也对你有帮助，里面包含了约<900>种食材的推荐、禁忌，请你点击查看”的情况下，待处理变量文本可以包括：文本B中的“<冠心病>”、“<蔬菜>”、“<900>”等。

40 本申请实施例可以生成待处理变量文本对应的第二视频片段。例如，可以首先确定待处理变量文本所在语句的声学参数和唇部特征；然后，从中提取出待处理变量文本对应的目标声学参数和目标唇部特征，并分别生成待处理变量文本对应的语音片段和目标图像序列。目标图像序列可以包括：目标运动图像序列和目标背景图像序列。

45 在生成目标运动图像序列的过程中，可以利用步骤206对目标唇部特征进行平滑处理，以提高唇部特征在拼接位置处的连续性。

可以利用步骤208，生成目标背景图像序列，实现目标背景图像序列与预设背景图像序列在分界

位置处的匹配，以提高背景图像序列在拼接位置处的连续性。

在对第一视频片段与第二视频片段进行拼接之前，可以首先对第一视频片段中第一语音片段和第二视频片段中第二语音片段，各自在所述分界位置的语音子片段进行平滑处理；然后根据平滑处理后的第一语音片段和平滑处理后的第二语音片段，对第一视频片段和第二视频片段进行拼接。

5 综上，本申请实施例的视频处理方法，在第一视频片段的拼接位置处加入了预设时长的停顿，有助于克服拼接位置处的跳变或抖动问题，因此能够提高拼接位置处的连续性。

10 并且，本申请实施例以待处理变量文本所在的语句为单位，确定对应的语音参数和图像参数，然后从语音参数和图像参数中，提取所述待处理变量文本对应的目标语音参数和目标图像参数。由于语句对应的语音参数和图像参数具有一定的连续性，故从中提取的待处理变量文本对应的目标语音参数和目标图像参数与语句中模板文本对应的语音参数和图像参数具有一定的连续性；在此基础上，能够提高待处理变量文本对应的第二视频片段与语句中模板文本对应的第一视频片段之间的连续性，进而能够进一步提高拼接位置处的连续性。

15 再者，本申请实施例根据预设变量文本在边界位置处的预设图像参数，对待处理变量文本对应的目标图像参数进行平滑处理。由于预设变量文本在边界位置处的预设图像参数与模板文本在边界位置处的图像参数具有一定的连续性，故上述平滑处理能够提高平滑处理后的目标图像参数与模板文本的图像参数在边界位置处的连续性；在此基础上，能够提高待处理变量文本对应的第二视频片段与语句中模板文本对应的第一视频片段之间的连续性，进而能够提高拼接位置处的连续性。

此外，本申请实施例根据预设背景图像序列，生成目标背景图像序列，可以实现目标背景图像序列与预设背景图像序列在分界位置处的匹配，以提高背景图像序列在拼接位置处的连续性。

20 进一步，本申请实施例在对第一视频片段与第二视频片段进行拼接之前，对第一视频片段中第一语音片段和第二视频片段中第二语音片段，在所述分界位置的语音子片段进行平滑处理。上述平滑处理能够提高平滑处理后的第一语音片段与第二语音片段之间的连续性，因此能够提高第一视频片段和第二视频片段在拼接位置处的连续性。

25 需要说明的是，对于方法实施例，为了简单描述，故将其都表述为一系列的运动动作组合，但是本领域技术人员应该知悉，本申请实施例并不受所描述的运动动作顺序的限制，因为依据本申请实施例，某些步骤可以采用其他顺序或者同时进行。其次，本领域技术人员也应该知悉，说明书中所描述的实施例均属于优选实施例，所涉及的运动动作并不一定是本申请实施例所必须的。

装置实施例

参照图 3，示出了本申请的一种视频处理装置实施例的结构框图，具体可以包括：

30 提供模块 301，用于获取第一视频片段，所述第一视频片段与待生成视频的第一文本中模板文本对应，并且所述第一视频片段包括语音停顿的视频子片段，所述视频子片段的位置对应于所述模板文本与所述第一文本中待处理变量文本之间的分界位置；

生成模块 302，用于生成待处理变量文本对应的第二视频片段；

35 拼接模块 303，用于对所述第一视频片段和所述第二视频片段进行拼接，以得到所述第一文本对应的视频。

可选地，上述装置还可以包括：

预设视频生成模块，用于根据模板文本、预设变量文本、以及所述分界位置处对应的停顿信息，生成预设视频，所述停顿信息表示预定时长的语音停顿；

截取模块，用于从所述预设视频中截取所述模板文本对应的第一视频片段。

40 可选地，生成模块 302 可以包括：

参数确定模块，用于针对待处理变量文本在所述第一文本中所处的语句，确定对应的语音参数和图像参数，其中，所述图像参数表征所述第一文本对应的视频中要出现的虚拟对象的状态特征，所述语音参数用于表征语音合成对应的参数；

45 参数提取模块，用于从所述语音参数和图像参数中，提取所述待处理变量文本对应的目标语音参数和目标图像参数；

第一片段生成模块，用于根据所述目标语音参数和目标图像参数，生成所述待处理变量文本对应

的第二视频片段。

可选地，生成模块 302 可以包括：

第一平滑处理模块，用于根据所述待处理变量文本在边界位置处的预设图像参数，对所述待处理变量文本对应的目标图像参数进行平滑处理，以提高所述目标图像参数与所述模板文本的图像参数在边界位置处的连续性；

第二片段生成模块，用于根据平滑处理后的目标图像参数，生成所述待处理变量文本对应的第二视频片段。

可选地，上述第一视频片段可以包括：第一语音片段，上述第二视频片段可以包括：第二语音片段；

拼接模块 303 可以包括：

第二平滑处理模块，用于对第一语音片段和第二语音片段各自在拼接位置处的语音子片段进行平滑处理；

平滑后拼接模块，用于对平滑处理后的第一语音片段和平滑处理后的第二语音片段进行拼接。

可选地，上述视频对应的图像序列可以包括：背景图像序列和运动图像序列；

生成模块 302 可以包括：

运动图像序列生成模块，用于生成待处理变量文本对应的目标运动图像序列；

背景图像序列生成模块，用于根据预设背景图像序列，确定上述待处理变量文本对应的目标背景图像序列；

融合模块，用于对上述目标运动图像序列和上述目标背景图像序列进行融合，以得到上述待处理变量文本对应的第二视频片段。

可选地，上述目标背景图像序列的位于首尾位置的背景图像，与上述预设背景图像序列的位于首尾位置的背景图像相匹配。

可选地，上述背景图像序列生成模块，可以包括：

第一背景图像序列生成模块，用于在上述预设背景图像序列对应的图像数量与上述目标运动图像序列对应的图像数量相匹配的情况下，将上述预设背景图像序列确定为目标背景图像序列；或者

第二背景图像序列生成模块，用于在上述预设背景图像序列对应的图像数量大于上述目标运动图像序列对应的图像数量的情况下，从上述预设背景图像序列中丢弃位于中间位置的第一背景图像；在丢弃至少两帧第一背景图像的情况下，至少两帧第一背景图像在预设背景图像序列中不连续分布；或者

第三背景图像序列生成模块，用于在上述预设背景图像序列对应的图像数量小于上述目标运动图像序列对应的图像数量的情况下，在上述预设背景图像序列中增加第二背景图像。

对于装置实施例而言，由于其与方法实施例基本相似，所以描述的比较简单，相关之处参见方法实施例的部分说明即可。

本说明书中的各个实施例均采用递进的方式描述，每个实施例重点说明的都是与其他实施例的不同之处，各个实施例之间相同相似的部分互相参见即可。

关于上述实施例中的装置，其中各个模块执行操作的具体方式已经在有关该方法的实施例中进行了详细描述，此处将不做详细阐述说明。

图 4 是根据一示例性实施例示出的一种用于视频处理的装置 900 的结构框图。例如，装置 900 可以是移动来电，计算机，数字广播终端，消息收发设备，游戏控制台，平板设备，医疗设备，健身设备，个人数字助理等。

参照图 4，装置 900 可以包括以下一个或多个组件：处理组件 902，存储器 904，电源组件 906，多媒体组件 908，音频组件 910，输入/输出 (I/O) 的接口 912，传感器组件 914，以及通信组件 916。

处理组件 902 通常控制装置 900 的整体操作，诸如与显示，来电呼叫，数据通信，相机操作和记录操作相关联的操作。处理元件 902 可以包括一个或多个处理器 920 来执行指令，以完成上述的方法的全部或部分步骤。此外，处理组件 902 可以包括一个或多个模块，便于处理组件 902 和其他组件之间的交互。例如，处理组件 902 可以包括多媒体模块，以方便多媒体组件 908 和处理组件 902 之间的

交互。

存储器 904 被配置为存储各种类型的数据以支持在设备 900 的操作。这些数据的示例包括用于在装置 900 上操作的任何应用程序或方法的指令，联系人数据，来电簿数据，消息，图片，视频等。存储器 904 可以由任何类型的易失性或非易失性存储设备或者它们的组合实现，如静态随机存取存储器 (SRAM)，电可擦除可编程只读存储器 (EEPROM)，可擦除可编程只读存储器 (EPROM)，可编程只读存储器 (PROM)，只读存储器 (ROM)，磁存储器，快闪存储器，磁盘或光盘。

电源组件 906 为装置 900 的各种组件提供电力。电源组件 906 可以包括电源管理系统，一个或多个电源，及其他与为装置 900 生成、管理和分配电力相关联的组件。

多媒体组件 908 包括在所述装置 900 和用户之间的提供一个输出接口的屏幕。在一些实施例中，屏幕可以包括液晶显示器 (LCD) 和触摸面板 (TP)。如果屏幕包括触摸面板，屏幕可以被实现为触摸屏，以接收来自用户的输入信号。触摸面板包括一个或多个触摸传感器以感测触摸、滑动和触摸面板上的手势。所述触摸传感器可以不仅感测触摸或滑动运动动作的边界，而且还检测与所述触摸或滑动操作相关的持续时间和压力。在一些实施例中，多媒体组件 908 包括一个前置摄像头和/或后置摄像头。当设备 900 处于操作模式，如拍摄模式或视频模式时，前置摄像头和/或后置摄像头可以接收外部的多媒体数据。每个前置摄像头和后置摄像头可以是一个固定的光学透镜系统或具有焦距和光学变焦能力。

音频组件 910 被配置为输出和/或输入音频信号。例如，音频组件 910 包括一个麦克风 (MIC)，当装置 900 处于操作模式，如呼叫模式、记录模式和语音识别模式时，麦克风被配置为接收外部音频信号。所接收的音频信号可以被进一步存储在存储器 904 或经由通信组件 916 发送。在一些实施例中，音频组件 910 还包括一个扬声器，用于输出音频信号。

I/O 接口 912 为处理组件 902 和外围接口模块之间提供接口，上述外围接口模块可以是键盘，点击轮，按钮等。这些按钮可包括但不限于：主页按钮、音量按钮、启动按钮和锁定按钮。

传感器组件 914 包括一个或多个传感器，用于为装置 900 提供各个方面的状态评估。例如，传感器组件 914 可以检测到设备 900 的打开/关闭状态，组件的相对定位，例如所述组件为装置 900 的显示器和小键盘，传感器组件 914 还可以检测装置 900 或装置 900 一个组件的位置改变，用户与装置 900 接触的存在或不存在，装置 900 方位或加速/减速和装置 900 的温度变化。传感器组件 914 可以包括接近传感器，被配置用来在没有任何的物理接触时检测附近物体的存在。传感器组件 914 还可以包括光传感器，如 CMOS 或 CCD 图像传感器，用于在成像应用中使用。在一些实施例中，该传感器组件 914 还可以包括加速度传感器，陀螺仪传感器，磁传感器，压力传感器或温度传感器。

通信组件 916 被配置为便于装置 900 和其他设备之间有线或无线方式的通信。装置 900 可以接入基于通信标准的无线网络，如 WiFi，2G 或 3G，或它们的组合。在一个示例性实施例中，通信部件 916 经由广播信道接收来自外部广播管理系统的广播信号或广播相关信息。在一个示例性实施例中，所述通信部件 916 还包括近场通信 (NFC) 模块，以促进短程通信。例如，在 NFC 模块可基于射频识别 (RFID) 技术，红外数据协会 (IrDA) 技术，超宽带 (UWB) 技术，蓝牙 (BT) 技术和其他技术来实现。

在示例性实施例中，装置 900 可以被一个或多个应用专用集成电路 (ASIC)、数字信号处理器 (DSP)、数字信号处理设备 (DSPD)、可编程逻辑器件 (PLD)、现场可编程门阵列 (FPGA)、控制器、微控制器、微处理器或其他电子元件实现，用于执行上述方法。

在示例性实施例中，还提供了一种包括指令的非临时性计算机可读存储介质，例如包括指令的存储器 904，上述指令可由装置 900 的处理器 920 执行以完成上述方法。例如，所述非临时性计算机可读存储介质可以是 ROM、随机存取存储器 (RAM)、CD-ROM、磁带、软盘和光数据存储设备等。

图 5 是本申请的一些实施例中服务端的结构框图。该服务端 1900 可因配置或性能不同而产生比较大的差异，可以包括一个或一个以上中央处理器 (central processing units, CPU) 1922 (例如，一个或一个以上处理器) 和存储器 1932，一个或一个以上存储应用程序 1942 或数据 1944 的存储介质 1930 (例如一个或一个以上海量存储设备)。其中，存储器 1932 和存储介质 1930 可以是短暂存储或持久存储。存储在存储介质 1930 的程序可以包括一个或一个以上模块 (图示没标出)，每个模块可以包括对

服务端中的一系列指令操作。更进一步地，中央处理器 1922 可以设置为与存储介质 1930 通信，在服务端 1900 上执行存储介质 1930 中的一系列指令操作。

5 服务端 1900 还可以包括一个或一个以上电源 1926，一个或一个以上有线或无线网络接口 1950，一个或一个以上输入输出接口 1958，一个或一个以上键盘 1956，和/或，一个或一个以上操作系统 1941，例如 Windows Server™，Mac OS X™，Unix™，Linux™，FreeBSD™ 等等。

一种非临时性计算机可读存储介质，当所述存储介质中的指令由装置（设备或者服务端）的处理器执行时，使得装置能够执行根据本申请实施例的视频处理方法。

10 本领域技术人员在考虑说明书及实践这里公开的发明后，将容易想到本申请的其它实施方案。本申请旨在涵盖本申请的任何变型、用途或者适应性变化，这些变型、用途或者适应性变化遵循本申请的一般性原理并包括本公开未公开的本技术领域中的公知常识或惯用技术手段。说明书和实施例仅被视为示例性的，本申请的真正范围和精神由下面的权利要求指出。

应当理解的是，本申请并不局限于上面已经描述并在附图中示出的精确结构，并且可以在不脱离其范围进行各种修改和改变。本申请的范围仅由所附的权利要求来限制。

15 以上所述仅为本申请的较佳实施例，并不用以限制本申请，凡在本申请的精神和原则之内，所作的任何修改、等同替换、改进等，均应包含在本申请的保护范围之内。

20 以上对本申请实施例所提供的一种视频处理方法、一种视频处理装置和一种用于视频处理的装置，进行了详细介绍，本文中应用了具体个例对本申请的原理及实施方式进行了阐述，以上实施例的说明只是用于帮助理解本申请的方法及其核心思想；同时，对于本领域的一般技术人员，依据本申请的思想，在具体实施方式及应用范围上均会有改变之处，综上所述，本说明书内容不应理解为对本申请的限制。

权利要求书

1、一种视频处理方法，在电子设备中执行，所述方法包括：

获取第一视频片段，所述第一视频片段与待生成视频的第一文本中模板文本对应，并且所述第一视频片段包括语音停顿的视频子片段，所述视频子片段的位置对应于所述模板文本与所述第一文本中待处理变量文本之间的分界位置；

生成所述待处理变量文本对应的第二视频片段；

对所述第一视频片段和所述第二视频片段进行拼接，以得到所述第一文本对应的视频。

2、根据权利要求1所述的方法，其中，所述方法还包括：

根据模板文本、预设变量文本、以及所述分界位置处对应的停顿信息，生成预设视频，所述停顿信息表示预定时长的语音停顿；

从所述预设视频中截取所述模板文本对应的第一视频片段。

3、根据权利要求1所述的方法，其特征在于，所述视频子片段的图像中虚拟对象处于不說話的状态。

4、根据权利要求1-3中任一项所述的方法，其特征在于，所述视频子片段为经过停顿处理后得到的子片段；

其中，对所述视频子片段的停顿处理，包括：

所述第一视频片段中与所述分界位置对应的拼接位置处的语音信号子片段与静音信号进行加权处理，以得到语音停顿的语音信号子片段；

所述第一视频片段在所述拼接位置处的图像子序列与目标状态特征的图像序列进行加权处理，以得到虚拟对象处于不說話的状态的所述图像子序列，其中，所述目标状态特征为表示虚拟对象处于不說話状态的特征。

5、根据权利要求1-3中任一项所述的方法，其中，所述生成所述待处理变量文本对应的第二视频片段，包括：

针对待处理变量文本在所述第一文本中所处的语句，确定对应的语音参数和图像参数，其中，所述图像参数表征所述第一文本对应的视频中要出现的虚拟对象的状态特征，所述语音参数用于表征语音合成对应的参数；

从所述语音参数和图像参数中，提取所述待处理变量文本对应的目标语音参数和目标图像参数；

根据所述目标语音参数和目标图像参数，生成所述待处理变量文本对应的第二视频片段。

6、根据权利要求1-3中任一项所述的方法，其中，所述生成所述待处理变量文本对应的第二视频片段，包括：

根据所述待处理变量文本在边界位置处的预设图像参数，对所述待处理变量文本对应的目标图像参数进行平滑处理，以提高所述目标图像参数与所述模板文本的图像参数在边界位置处的连续性；

根据平滑处理后的目标图像参数，生成所述待处理变量文本对应的第二视频片段。

7、根据权利要求1-3中任一项所述的方法，其中，所述第一视频片段包括：第一语音片段，所述第二视频片段包括：第二语音片段；

所述对所述第一视频片段和所述第二视频片段进行拼接，包括：

对第一语音片段和第二语音片段各自在拼接位置处的语音子片段进行平滑处理；

对平滑处理后的第一语音片段和平滑处理后的第二语音片段进行拼接。

8、根据权利要求1-3中任一项所述的方法，其中，所述视频对应的图像序列包括：背景图像序列和运动图像序列；

所述生成待处理变量文本对应的第二视频片段，包括：

生成待处理变量文本对应的目标运动图像序列；

根据预设背景图像序列，确定所述待处理变量文本对应的目标背景图像序列；

对所述目标运动图像序列和所述目标背景图像序列进行融合，以得到所述待处理变量文本对应的第二视频片段。

9、根据权利要求 8 所述的方法，其中，所述目标背景图像序列的位于首尾位置的背景图像，与所述预设背景图像序列的位于首尾位置的背景图像相匹配。

10、根据权利要求 8 所述的方法，其中，所述根据预设背景图像序列，确定所述待处理变量文本对应的目标背景图像序列，包括：

5 在所述预设背景图像序列对应的图像数量与所述目标运动图像序列对应的图像数量相匹配的情况下，将所述预设背景图像序列确定为目标背景图像序列；或者

在所述预设背景图像序列对应的图像数量大于所述目标运动图像序列对应的图像数量的情况下，从所述预设背景图像序列中丢弃位于中间位置的第一背景图像；在丢弃至少两帧第一背景图像的情况下，至少两帧第一背景图像在预设背景图像序列中不连续分布；或者

10 在所述预设背景图像序列对应的图像数量小于所述目标运动图像序列对应的图像数量的情况下，在所述预设背景图像序列中增加第二背景图像。

11、一种视频处理装置，包括：

15 提供模块，用于获取第一视频片段，所述第一视频片段与待生成视频的第一文本中模板文本对应，并且所述第一视频片段包括语音停顿的视频子片段，所述视频子片段的位置对应于所述模板文本与所述第一文本中待处理变量文本之间的分界位置；

生成模块，用于生成所述待处理变量文本对应的第二视频片段；

15 拼接模块，用于对所述第一视频片段和所述第二视频片段进行拼接，以得到所述第一文本对应的视频。

12、根据权利要求 9 所述的装置，其中，所述装置还包括：

20 预设视频生成模块，用于根据模板文本、预设变量文本、以及所述分界位置处对应的停顿信息，生成预设视频，所述停顿信息表示预定时长的语音停顿；

截取模块，用于从所述预设视频中截取所述模板文本对应的第一视频片段。

13、根据权利要求 9 或 10 所述的装置，其中，所述生成模块包括：

25 参数确定模块，用于针对待处理变量文本在所述第一文本中所处的语句，确定对应的语音参数和图像参数，其中，所述图像参数表征所述第一文本对应的视频中要出现的虚拟对象的状态特征，所述语音参数用于表征语音合成对应的参数；

参数提取模块，用于从所述语音参数和图像参数中，提取所述待处理变量文本对应的目标语音参数和目标图像参数；

30 第一片段生成模块，用于根据所述目标语音参数和目标图像参数，生成所述待处理变量文本对应的第二视频片段。

14、根据权利要求 9 或 10 所述的装置，其中，所述生成模块包括：

第一平滑处理模块，用于根据所述待处理变量文本在边界位置处的预设图像参数，对所述待处理变量文本对应的目标图像参数进行平滑处理，以提高所述目标图像参数与所述模板文本的图像参数在边界位置处的连续性；

35 第二片段生成模块，用于根据平滑处理后的目标图像参数，生成所述待处理变量文本对应的第二视频片段。

15、根据权利要求 9 或 10 所述的装置，其中，所述第一视频片段包括：第一语音片段，所述第二视频片段包括：第二语音片段；

所述拼接模块包括：

40 第二平滑处理模块，用于对第一语音片段和第二语音片段各自在拼接位置处的语音子片段进行平滑处理；

平滑后拼接模块，用于对平滑处理后的第一语音片段和平滑处理后的第二语音片段进行拼接。

45 16、一种用于视频处理的装置，包括有存储器，以及一个或者一个以上的程序，其中一个或者一个以上程序存储于存储器中，所述程序被一个或者一个以上处理器执行时，实现权利要求 1 至 10 中任一所述方法的步骤。

17、一种机器可读介质，其上存储有指令，当由一个或多个处理器执行时，使得装置执行如权利

要求 1 至 10 中一个或多个所述的视频处理方法。

18、一种计算机程序产品，该程序产品包括计算机指令，该计算机指令存储在计算机可读存储介质中；当处理器执行该计算机指令时，使得处理器执行如权利要求 1 至 10 中任一项所述的方法。



图 1A

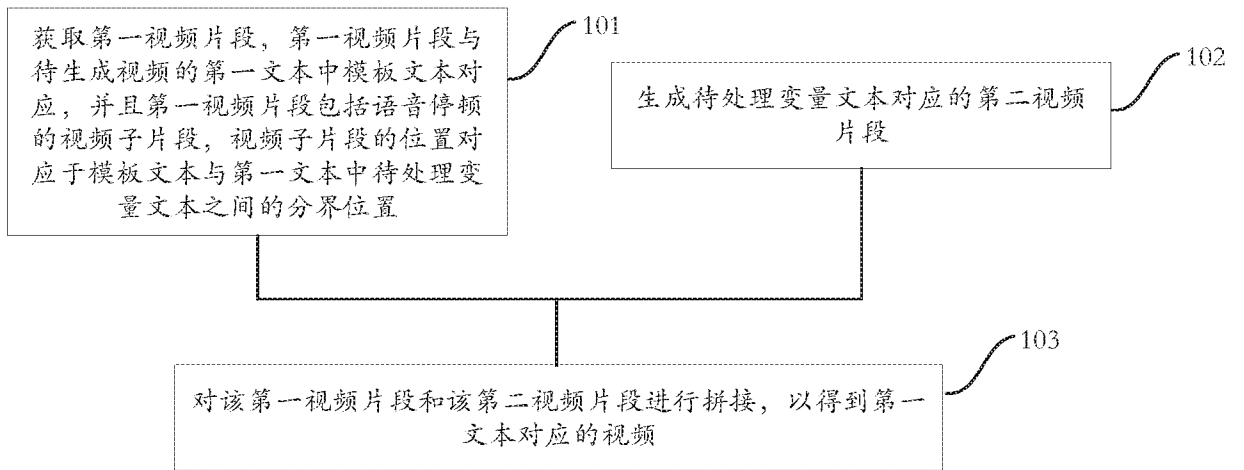


图 1B

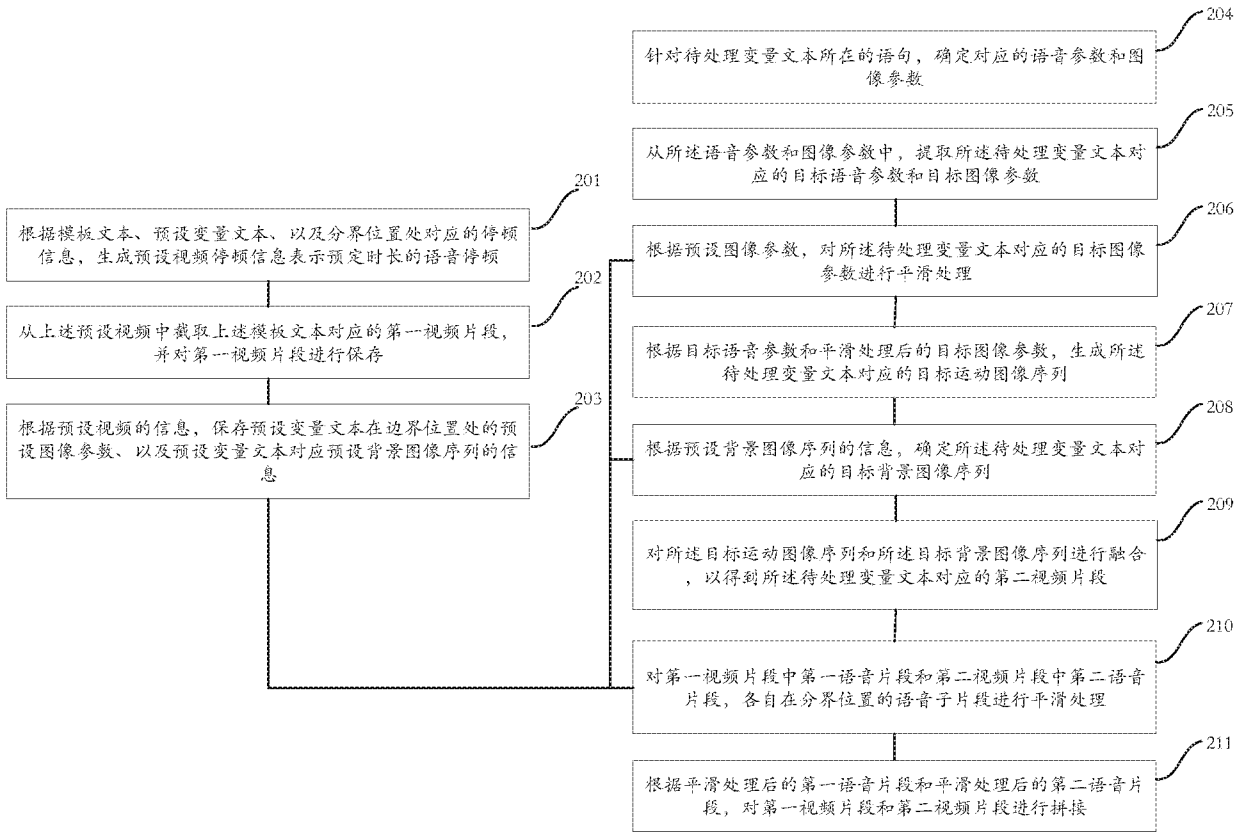


图 2

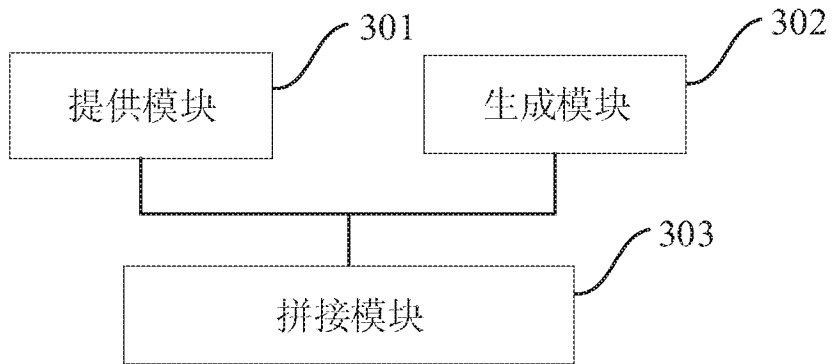


图 3

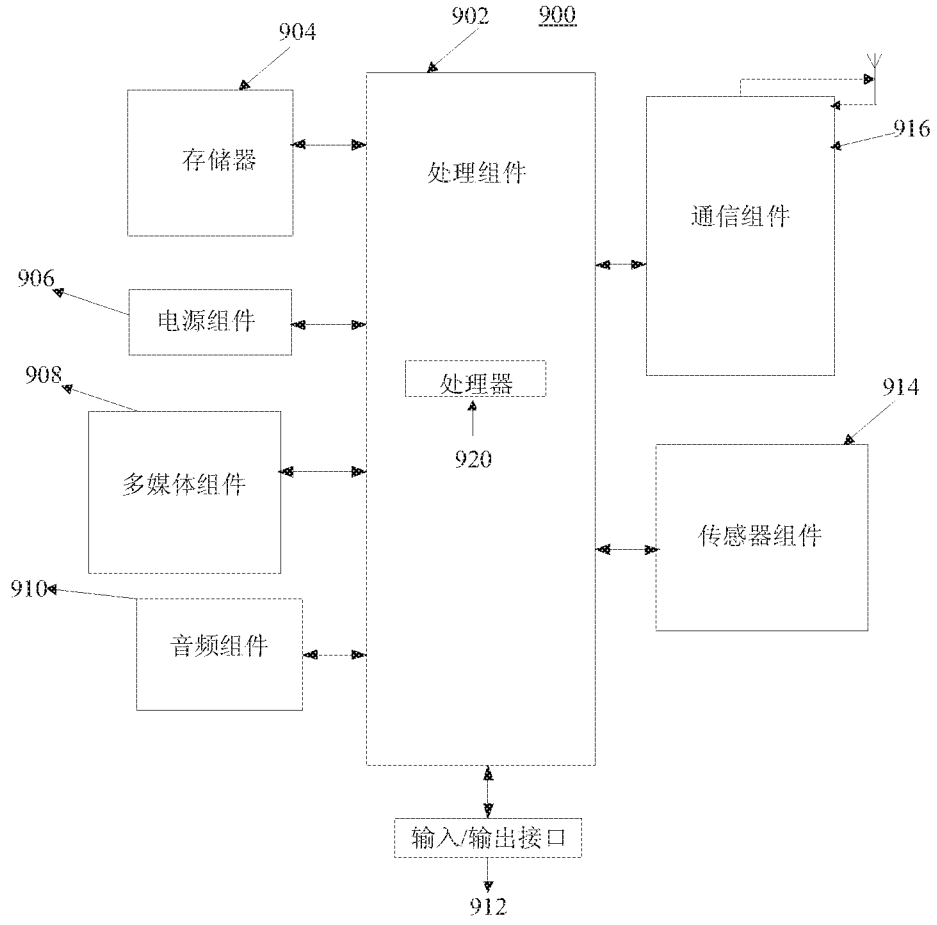


图 4

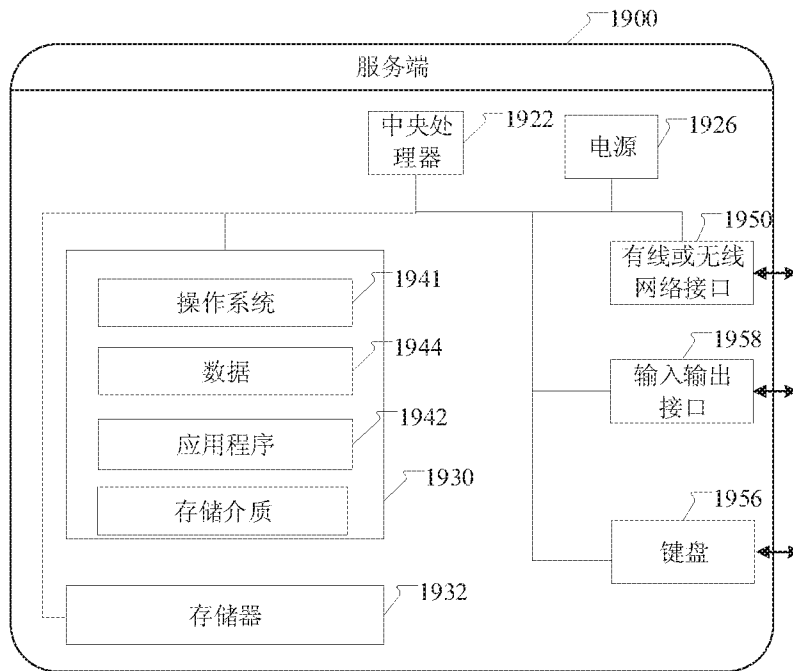


图 5

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2022/115722

A. CLASSIFICATION OF SUBJECT MATTER		
H04N 21/439(2011.01)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)		
H04N		
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; WPABS; WPABSC; VEN; ENTXT; ENTXTC; CJFD: 搜狗, 视频, 片段, 模板, 固定, 常量, 变量, 改变, 编辑, 文本, 文案, 台本, 停顿, 分段, 分界, 划分, 截取, 拼接, 替换, 语音, 音频, 静音, 安静, 加权, 融合, 平滑, 连续, 运动, 背景, video, clip, segment, fragment, template, fix, constant, variable, change, edit, text, copy, playscript, pause, subsection, division, divide, cut, joint, splice, replace, sound, voice, audio, silence, weight, fuse, mix, smooth, move, motion, background		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 113891150 A (BEIJING SOGOU TECHNOLOGY DEVELOPMENT CO., LTD.) 04 January 2022 (2022-01-04) claims 1-15, and description, paragraphs [0038]-[0198]	1-18
X	CN 111885416 A (BEIJING LAIYE NETWORK TECHNOLOGY CO., LTD. et al.) 03 November 2020 (2020-11-03) description, paragraphs [0094]-[0235]	1-3,5-18
A	CN 112995706 A (TENCENT TECHNOLOGY SHENZHEN CO., LTD.) 18 June 2021 (2021-06-18) entire document	1-18
A	US 2021201912 A1 (BEIJING BAIDU NETCOM SCIENCE AND TECHNOLOGY CO., LTD.) 01 July 2021 (2021-07-01) entire document	1-18
A	CN 110381266 A (BAIDU ONLINE NETWORK TECHNOLOGY (BEIJING) CO., LTD.) 25 October 2019 (2019-10-25) entire document	1-18
<input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C. <input checked="" type="checkbox"/> 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 "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		Date of mailing of the international search report
27 October 2022		24 November 2022
Name and mailing address of the ISA/CN		Authorized officer
China National Intellectual Property Administration (ISA/CN) No. 6, Xitucheng Road, Jimenqiao, Haidian District, Beijing 100088, China		
Facsimile No. (86-10)62019451		Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2022/115722

C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN 110611840 A (BEIJING QIYI CENTURY SCIENCE & TECHNOLOGY CO., LTD.) 24 December 2019 (2019-12-24) entire document	1-18
A	US 2020098396 A1 (AUTOCHARTIST LTD) 26 March 2020 (2020-03-26) entire document	1-18
A	CN 109637518 A (BEIJING SOGOU TECHNOLOGY DEVELOPMENT CO., LTD.) 16 April 2019 (2019-04-16) entire document	1-18

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2022/115722

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)		
CN	113891150	A	04 January 2022	None			
CN	111885416	A	03 November 2020	CN	111885416	B	12 April 2022
CN	112995706	A	18 June 2021	CN	112995706	B	19 April 2022
US	2021201912	A1	01 July 2021	CN	112102448	A	18 December 2020
CN	110381266	A	25 October 2019	None			
CN	110611840	A	24 December 2019	CN	110611840	B	09 November 2021
US	2020098396	A1	26 March 2020	US	10783928	B2	22 September 2020
CN	109637518	A	16 April 2019	CN	109637518	B	24 May 2022

国际检索报告

国际申请号

PCT/CN2022/115722

<p>A. 主题的分类</p> <p>H04N 21/439(2011.01)i</p> <p>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</p>																							
<p>B. 检索领域</p> <p>检索的最低限度文献(标明分类系统和分类号)</p> <p>H04N</p> <p>包含在检索领域中的除最低限度文献以外的检索文献</p> <p>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))</p> <p>CNXTX; WPABS; WPABSC; VEN; ENTXT; ENTXTC; CJFD:搜狗, 视频, 片段, 模板, 固定, 常量, 变量, 改变, 编辑, 文本, 文案, 台本, 停顿, 分段, 分界, 划分, 截取, 拼接, 替换, 语音, 音频, 静音, 安静, 加权, 融合, 平滑, 连续, 运动, 背景, video, clip, segment, fragment, template, fix, constant, variable, change, edit, text, copy, playscript, pause, subsection, division, divide, cut, joint, splice, replace, sound, voice, audio, silence, weight, fuse, mix, smooth, move, motion, background</p>																							
<p>C. 相关文件</p> <table border="1"> <thead> <tr> <th>类型*</th> <th>引用文件, 必要时, 指明相关段落</th> <th>相关的权利要求</th> </tr> </thead> <tbody> <tr> <td>PX</td> <td>CN 113891150 A (北京搜狗科技发展有限公司) 2022年1月4日 (2022 - 01 - 04) 权利要求1-15以及说明书第[0038]-[0198]段</td> <td>1-18</td> </tr> <tr> <td>X</td> <td>CN 111885416 A (北京来也网络科技有限公司等) 2020年11月3日 (2020 - 11 - 03) 说明书第[0094]-[0235]段</td> <td>1-3, 5-18</td> </tr> <tr> <td>A</td> <td>CN 112995706 A (腾讯科技深圳有限公司) 2021年6月18日 (2021 - 06 - 18) 全文</td> <td>1-18</td> </tr> <tr> <td>A</td> <td>US 2021201912 A1 (BEIJING BAIDU NETCOM SCI & TECH CO LTD) 2021年7月1日 (2021 - 07 - 01) 全文</td> <td>1-18</td> </tr> <tr> <td>A</td> <td>CN 110381266 A (百度在线网络技术北京有限公司) 2019年10月25日 (2019 - 10 - 25) 全文</td> <td>1-18</td> </tr> <tr> <td>A</td> <td>CN 110611840 A (北京奇艺世纪科技有限公司) 2019年12月24日 (2019 - 12 - 24) 全文</td> <td>1-18</td> </tr> </tbody> </table>			类型*	引用文件, 必要时, 指明相关段落	相关的权利要求	PX	CN 113891150 A (北京搜狗科技发展有限公司) 2022年1月4日 (2022 - 01 - 04) 权利要求1-15以及说明书第[0038]-[0198]段	1-18	X	CN 111885416 A (北京来也网络科技有限公司等) 2020年11月3日 (2020 - 11 - 03) 说明书第[0094]-[0235]段	1-3, 5-18	A	CN 112995706 A (腾讯科技深圳有限公司) 2021年6月18日 (2021 - 06 - 18) 全文	1-18	A	US 2021201912 A1 (BEIJING BAIDU NETCOM SCI & TECH CO LTD) 2021年7月1日 (2021 - 07 - 01) 全文	1-18	A	CN 110381266 A (百度在线网络技术北京有限公司) 2019年10月25日 (2019 - 10 - 25) 全文	1-18	A	CN 110611840 A (北京奇艺世纪科技有限公司) 2019年12月24日 (2019 - 12 - 24) 全文	1-18
类型*	引用文件, 必要时, 指明相关段落	相关的权利要求																					
PX	CN 113891150 A (北京搜狗科技发展有限公司) 2022年1月4日 (2022 - 01 - 04) 权利要求1-15以及说明书第[0038]-[0198]段	1-18																					
X	CN 111885416 A (北京来也网络科技有限公司等) 2020年11月3日 (2020 - 11 - 03) 说明书第[0094]-[0235]段	1-3, 5-18																					
A	CN 112995706 A (腾讯科技深圳有限公司) 2021年6月18日 (2021 - 06 - 18) 全文	1-18																					
A	US 2021201912 A1 (BEIJING BAIDU NETCOM SCI & TECH CO LTD) 2021年7月1日 (2021 - 07 - 01) 全文	1-18																					
A	CN 110381266 A (百度在线网络技术北京有限公司) 2019年10月25日 (2019 - 10 - 25) 全文	1-18																					
A	CN 110611840 A (北京奇艺世纪科技有限公司) 2019年12月24日 (2019 - 12 - 24) 全文	1-18																					
<p><input checked="" type="checkbox"/> 其余文件在C栏的续页中列出。</p> <p><input checked="" type="checkbox"/> 见同族专利附件。</p> <table border="0"> <tr> <td> <p>* 引用文件的具体类型:</p> <p>“A” 认为不特别相关的表示了现有技术一般状态的文件</p> <p>“E” 在国际申请日的当天或之后公布的在先申请或专利</p> <p>“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)</p> <p>“O” 涉及口头公开、使用、展览或其他方式公开的文件</p> <p>“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</p> </td> <td> <p>“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</p> <p>“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</p> <p>“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</p> <p>“&” 同族专利的文件</p> </td> </tr> </table>			<p>* 引用文件的具体类型:</p> <p>“A” 认为不特别相关的表示了现有技术一般状态的文件</p> <p>“E” 在国际申请日的当天或之后公布的在先申请或专利</p> <p>“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)</p> <p>“O” 涉及口头公开、使用、展览或其他方式公开的文件</p> <p>“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</p>	<p>“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</p> <p>“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</p> <p>“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</p> <p>“&” 同族专利的文件</p>																			
<p>* 引用文件的具体类型:</p> <p>“A” 认为不特别相关的表示了现有技术一般状态的文件</p> <p>“E” 在国际申请日的当天或之后公布的在先申请或专利</p> <p>“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)</p> <p>“O” 涉及口头公开、使用、展览或其他方式公开的文件</p> <p>“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</p>	<p>“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</p> <p>“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</p> <p>“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</p> <p>“&” 同族专利的文件</p>																						
<p>国际检索实际完成的日期</p> <p>2022年10月27日</p>		<p>国际检索报告邮寄日期</p> <p>2022年11月24日</p>																					
<p>ISA/CN的名称和邮寄地址</p> <p>中国国家知识产权局(ISA/CN) 中国北京市海淀区蓟门桥西土城路6号 100088</p> <p>传真号 (86-10)62019451</p>		<p>授权官员</p> <p>杜乾敏</p> <p>电话号码 (86-27)59182708</p>																					

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
A	US 2020098396 A1 (AUTOCHARTIS LTD) 2020年3月26日 (2020 - 03 - 26) 全文	1-18
A	CN 109637518 A (北京搜狗科技发展有限公司) 2019年4月16日 (2019 - 04 - 16) 全文	1-18

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2022/115722

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	113891150	A	2022年1月4日	无			
CN	111885416	A	2020年11月3日	CN	111885416	B	2022年4月12日
CN	112995706	A	2021年6月18日	CN	112995706	B	2022年4月19日
US	2021201912	A1	2021年7月1日	CN	112102448	A	2020年12月18日
CN	110381266	A	2019年10月25日	无			
CN	110611840	A	2019年12月24日	CN	110611840	B	2021年11月9日
US	2020098396	A1	2020年3月26日	US	10783928	B2	2020年9月22日
CN	109637518	A	2019年4月16日	CN	109637518	B	2022年5月24日