

文本视频生成方法、装置、电子设备及存储介质

5 本申请要求 2022 年 12 月 07 日递交的、标题为“文本视频生成方法、装置、电子设备及存储介质”、申请号为 202211567282.4 的中国发明专利申请的优先权，该申请的全部内容通过引用结合在本申请中。

技术领域

10 本公开实施例涉及互联网技术领域，尤其涉及一种文本视频生成方法、装置、电子设备及存储介质。

背景技术

15 当前，短视频平台凭借其丰富和多样性的内容，受到越来越多用户的喜爱。短视频平台的客户端设置有面向普通用户开放的投稿入口，创作者用户可以拍摄和上传视频进行投稿，再由短视频平台的服务端将创作者用户上传的内容推送给观看用户进行消费。

现有技术中，短视频平台通常仅能接收用户制作好的视频作品，而无法编辑和生成文本视频，导致用户必须手动将纯文本的作品制作成视频后才能上传短视频平台，降低视频创造过程中的视频生成效率和质量。

20 发明内容

本公开实施例提供一种文本视频生成方法、装置、电子设备及存储介质，以克服无法以纯文本的形式生成视频的问题。

第一方面，本公开实施例提供一种文本视频生成方法，包括：

25 显示文本编辑页面，所述文本编辑页面内包括文本输入区；响应于针对文本编辑页面的第一输入指令，在所述文本输入区显示目标文本，其中，所述目标文本在所述文本输入区内具有第一字体状态，所述第一字体状态表征所述目标文本的字号和/或行间距，所述第一字体状态是由所述目标文本的长度确定的；生成用于展示所述文本输入区内的目标文本的目标视频。

第二方面，本公开实施例提供一种文本视频生成装置，包括：

编辑模块，用于显示文本编辑页面，所述文本编辑页面内包括文本输入区；

显示模块，响应于针对文本编辑页面的第一输入指令，在所述文本输入区显示目标文本，其中，所述目标文本在所述文本输入区内具有第一字体状态，所述第一字体状态表征所述目标文本的字号和/或行间距，所述第一字体状态是由所述目标文本的长度确定的；

生成模块，用于生成用于展示所述文本输入区内的目标文本的目标视频。

第三方面，本公开实施例提供一种电子设备，包括：

处理器，以及与所述处理器通信连接的存储器；

10 所述存储器存储计算机执行指令；

所述处理器执行所述存储器存储的计算机执行指令，以实现如上第一方面以及第一方面各种可能的设计所述的文本视频生成方法。

第四方面，本公开实施例提供一种计算机可读存储介质，所述计算机可读存储介质中存储有计算机执行指令，当处理器执行所述计算机执行指令时，
15 实现如上第一方面以及第一方面各种可能的设计所述的文本视频生成方法。

第五方面，本公开实施例提供一种计算机程序产品，包括计算机程序，该计算机程序被处理器执行时实现如上第一方面以及第一方面各种可能的设计所述的文本视频生成方法。

本实施例提供的文本视频生成方法、装置、电子设备及存储介质，通过
20 显示文本编辑页面，所述文本编辑页面内包括文本输入区；响应于针对文本编辑页面的第一输入指令，在所述文本输入区显示目标文本，其中，所述目标文本在所述文本输入区内具有第一字体状态，所述第一字体状态表征所述目标文本的字号和/或行间距，所述第一字体状态是由所述目标文本的长度确定的；生成用于展示所述文本输入区内的目标文本的目标视频。通过设置文
25 本输入区，并将在文本输入区内编辑的目标文本的字号和/或行间距能够随文本长度动态变化，之后对文本输入区内的目标文本进行转换，使生成的目标视频能够清晰、全面地展示目标文本的全部内容，实现基于纯文本生成文本视频的目的，提高视频创造过程中的视频生成效率和视频质量。

30 附图说明

为了更清楚地说明本公开实施例或现有技术中的技术方案，下面将对实施例或现有技术描述中所需要使用的附图作一简单地介绍，显而易见地，下面描述中的附图是本公开的一些实施例，对于本领域普通技术人员来讲，在不付出创造性劳动性的前提下，还可以根据这些附图获得其他的附图。

- 5 图 1 为本公开实施例提供的文本视频生成方法的一种应用场景图；
图 2 为本公开实施例提供的文本视频生成方法的流程示意图一；
图 3 为本公开实施例提供的一种在文本输入区显示目标文本的示意图；
图 4 为图 2 所示实施例中步骤 S103 的具体实现方式的流程图；
图 5 为本公开实施例提供的文本视频生成方法的流程示意图二；
10 图 6 为图 5 所示实施例中步骤 S203 的具体实现方式的流程图；
图 7 为本公开实施例提供的一种文本输入区的区域尺寸示意图；
图 8 为图 6 所示实施例中步骤 S2033 的具体实现方式的流程图；
图 9 为本公开实施例提供的一种背景图片的示意图；
图 10 为图 5 所示实施例中步骤 S205 的具体实现方式的流程图；
15 图 11 为本公开实施例提供的文本视频生成装置的结构框图；
图 12 为本公开实施例提供的一种电子设备的结构示意图；
图 13 为本公开实施例提供的电子设备的硬件结构示意图。

具体实施方式

- 20 为使本公开实施例的目的、技术方案和优点更加清楚，下面将结合本公开实施例中的附图，对本公开实施例中的技术方案进行清楚、完整地描述，显然，所描述的实施例是本公开一部分实施例，而不是全部的实施例。基于本公开中的实施例，本领域普通技术人员在没有作出创造性劳动前提下所获得的所有其他实施例，都属于本公开保护的范围。
- 25 需要说明的是，本申请所涉及的用户信息（包括但不限于用户设备信息、用户个人信息等）和数据（包括但不限于用于分析的数据、存储的数据、展示的数据等），均为经用户授权或者经过各方充分授权的信息和数据，并且相关数据的收集、使用和处理需要遵守相关国家和地区的相关法律法规和标准，并提供有相应的操作入口，供用户选择授权或者拒绝。
- 30 下面对本公开实施例的应用场景进行解释：

图 1 为本公开实施例提供的文本视频生成方法的一种应用场景图，本公开实施例提供的文本视频生成方法，可以应用于在短视频平台的客户端一侧进行视频编辑、上传的应用场景。具体地，如图 1 所示，本公开实施例提供的方法，可以应用于终端设备，例如智能手机。终端设备内运行有短视频平台的客户端，在客户端内设置有文本编辑页面，终端设备通过响应用户操作来触发相应的组件进入文本编辑页面，在文件编辑页面内，设置有文本输入区，用户通过输入法在该文本输入区内编辑文字。具体地，如图所示，例如，当用户点击文本输入区后，文本输入区进入编辑状态并显示文字输入光标，同时在文本编辑页面内弹出输入法界面，用户通过操作该输入法界面在文本输入区内输入文字，从而生成目标文本（图中“X”表示字符）。在目标文本输入完成后，输入法界面隐藏，文本编辑页面内显示完整的文本输入区。之后，点击“完成”按钮，对该文本输入区内的目标文字进行渲染，生成视频素材并上传至短视频平台的服务器端（图中示为平台服务器），从而完成将纯文字作品生成视频作品并上传短视频平台的过程。另一方面，点击“后退”按钮，可以退出文本编辑页面，不再赘述。

现有技术中，短视频平台的客户端一侧，通常仅能接收用户制作好的视频作品，并在实施转码、压缩等必要步骤后，上传至服务器端的视频池中，之后服务器端基于视频池中的视频，向不同的用户推送。然而，对应纯文本内容的视频，短视频平台的客户端通常不提供编辑创造的功能页面，这是由于：纯文本在编辑过程中，文本长度是由用户编辑的文本内容决定的，因此，当以固定字号、行间距对文字进行显示时，会造成字体、行间距出现过小或过大的问题，例如，以字体为例，当用户编辑的目标文本包含有 5 个汉字字符时，使用固定的三号字（示例性地）进行显示，字体大小合适；而当用户编辑的目标文本包含有 500 个汉字字符时，仍使用固定的三号字进行显示，则会导致字体过大，无法在同一屏内显示目标文本中的所有内容，而文本视频是对文本进行静态展示的视频（即视频始终显示同一帧内容），导致基于目标文本渲染生成的文本视频出现内容缺失的问题；反之，在上述示例情况中，若使用五号字（示例性地）进行显示，当用户编辑的目标文本包含有 5 个汉字字符时，则会导致字体过小的问题，进而导致基于目标文本渲染生成的文本视频中字体过小，影响视频展示效果。由于无法自动适配字体、行间距的大小，

导致生成的文本视频无法正常文本内容，难以实现文本视频的视频展示目的。因此，现有技术中，各短视频平台通常只能接收用户制作好的文本视频，而无法实现自动生成显示效果良好的文本视频。

本公开实施例提供一种文本视频生成方法以解决上述问题。

5 参考图 2，图 2 为本公开实施例提供的文本视频生成方法的流程示意图一。本实施例的方法可以应用在终端设备中，该文本视频生成方法包括：

步骤 S101：显示文本编辑页面，文本编辑页面内包括文本输入区。

示例性地，参考图 1 所示的应用场景示意图，在短视频平台的客户端（以下简称客户端）内，可以通过响应触发操作，启动文本编辑页面，具体样式可
10 参见图 1 中所示，具体地，其中，用于触发文本编辑页面的功能组件，可以是设置在客户端的视频播放页面（即短视频应用的默认页面）；也可以设置在用于上传视频作品的视频拍摄页面，即取景框页面。之后，在接收到用户对该功能组件的触发操作后，对文本编辑页面进行显示。

进一步地，在文本编辑页面内设置有文本输入区。示例性地，终端设备
15 在接收到用户针对文本输入区的点击操作后，文本输入区被激活并获得输入焦点（Focus），之后，用户可以通过输入法在该文本输入区内输入文字、符号等字符信息。其中，文本输入区的位置和大小基于需要具体设置，此处不做限制。

步骤 S102：响应于针对文本编辑页面的第一输入指令，在文本输入区显示目标文本，其中，目标文本在文本输入区内具有第一字体状态，第一字体
20 状态表征目标文本的字号和/或行间距，第一字体状态是由目标文本的长度确定的。

示例性地，之后，当文本输入区获得输入焦点后，终端设备接收用户输入的第一输入指令，第一输入指令即用于生成目标文本的文字信息，第一输入
25 指令中可以包括具体的字符、符号对应的标识，或者用于形成字符的拼读元素的标识，例如汉语拼音、字母等，此处不再一一赘述。终端设备获得第一输入指令后，基于系统内的输入法将第一输入指令转换为对应的文字、符号，并显示在上述文本输入区，形成目标文本，目标文本即最终待发布的文本。

其中，目标文本可以包括一个或多个字符、符号。其中，需要说明的是，
30 在文本输入区内输入和显示的符号，包括可见符号和不可见符号，可见符号

包括逗号、句号等用于文字编辑的符号，不再赘述；不可见符号包括空格、空行和换行符等。通过字符和符号构成的目标文本具有段落结构，例如目标文本分为若干个段落，目标文本中包含空行等，从而使文本输入区内的目标文本在视觉上具有更好的可读性。

- 5 进一步地，显示在文本输入区内的目标文本具有第一字体状态，其中，第一字体状态表征目标文本的字号和/或行间距，且第一字体状态是由目标文本的长度确定的。具体地，当目标文本的长度越大，则目标文本中的文字和符号的字号越小，和/或，目标文本中行与行直接的行间距越小；反之，当目标文本的长度越小，则目标文本中的文字和符号的字号越大，和/或，目标文本中行与行直接的行间距越大。简而言之，即当目标文本的长度较大时，则终端设备对文本输入区内的目标文本进行压缩，从而使文本输入区能够承载更多的字符和符号；而当目标文本的长度较小时，则放大字体，从而使目标文本的内容更加突出，视觉显示效果更好。当然，可以理解的是，对第一字体状态（即目标文本的字号和/或行间距）的调节一定的许可范围，第一字体状态与目标文本的长度之间具有预设的非线性的映射关系，具体实现方式在后续实施例中介绍，此处不再详述。
- 10
- 15

- 其中，进一步地，基于上述对目标文本的实现方式的介绍，目标文本的长度可以通过目标文本中字符的数量确定，也可以根据目标文本中字符和符号的数量和确定，还可以根据目标文本在文本输入区内的整体占位长度确定，例如，目标文本中包含空行符时，一个空行符占用文本输入区的一个空行。因此在目标文本中包含空行的情况下，使用目标文本在文本输入区内的整体占位长度确定目标文本的长度，能够更准确的衡量目标文本的实际长度，从而提高最终生成的目标视频的显示效果。
- 20

- 图 3 为本公开实施例提供的一种在文本输入区显示目标文本的示意图，如图 3 所示，本实施例中，以目标文本中的字符总数作为目标文本的长度。终端设备基于第一输入指令，在文本输入区内不断显示对应的目标文本，当目标文本的长度 $N=10$ 时（第一时刻），则目标文本的第一字体状态为 Info_1，表征目标文本中各字符和符号的字号为#4 号，随着第一输入指令的不断输入，目标文本的长度继续增加，当目标文本的长度 $N=50$ 时（第二时刻），目标文本的第一字体状态为 Info_2，表征目标文本中各字符和符号的字号为#5 号（相
- 25
- 30

比#4 号缩小一个级别)。从而实现在文本输入区内动态显示目标文本的字号的目的。

进一步地,在上述实施例的基础上,第一字体状态中还包括表征行间距的信息,通过第一字体状态,还可以进一步对目标文本的行间距进行调节。

5 例如,当目标文本的长度 $N=10$ 时,目标文本的第一字体状态为 Info_1,表征目标文本中各字符和符号的字号为#4 号,且行间距为 1;当目标文本的长度 $N=50$ 时,目标文本的第一字体状态为 Info_2,表征目标文本中各字符和符号的字号为#5 号,且行间距为 0.8。

或者,在另一种可能的实现方式中,还可以单独基于目标文本的长度,
10 确定目标文本的行间距。例如,参考图 3 所示,当目标文本的长度 $N=10$ 时,目标文本的第一字体状态为 Info_1,表征目标文本的行间距为 1;当目标文本的长度 $N=50$ 时,目标文本的第一字体状态为 Info_2,表征目标文本的行间距为 0.8。

步骤 S103: 生成用于展示文本输入区内的目标文本的目标视频。

15 示例性地,在文本输入区内的目标文本,具有对应的第一字体状态,之后,对该文本输入区进行渲染,生成具有预设时长的视频,即目标视频。其中,目标视频相当于是对文本输入区内的目标文本的还原展示,因此,目标视频不仅能显示目标文本的文字内容,还能够还原目标文本所具有的第一字体状态,即目标文本中字符的大小和/或行间距,从而使目标视频中展示的目标文本,与在文本输入区内展示的目标文本保持视觉上的一致性。进而,使
20 用户在文本输入区内编辑文字内容的过程中,能够对最终生成的目标视频(即文本视频)的展示效果有准确的预期,从而提高最终生成的目标视频的视频质量。避免出现生成的文本视频中文字过小,无法阅读的问题;或者文本视频中文字过大,无法全部显示的问题。

25 在一种可能的实现方式中,如图 4 所示,步骤 S103 的具体实现方式包括:

步骤 S1031: 基于目标文本,生成包含具有第一字体状态的目标文本的渲染图像。

步骤 S1032: 根据目标文本的长度,确定视频时长。

30 步骤 S1033: 根据视频时长和渲染图像,生成目标视频。

示例性地，首先，在文本输入区内输入目标文本结束后，基于文本输入区内的目标文本，将文本输入区内具有第一字体状态的目标文字转换为图像，即渲染图像。其中，将文字转换图像的具体实现方式有多种，例如通过对文本输入区进行截图生成渲染图像；或者，以目标文本和对应的第一字体状态
5 作为输入参数，输入图像转换器来生成对应的渲染图像。将字符转换为图片的具体实现步骤为本领域技术人员知晓的现有技术，此处不再赘述。

进一步地，根据目标文本的长度，例如目标文本中字符的个数，确定对应的视频时长，由于最终生成的目标视频在展示目标文本时，需要匹配用户的阅读时间，当目标文本的长度越长时，用户阅读该目标视频中的目标文本
10 所需的耗时越长，因此，设置与目标文本的长度相匹配的视频时长，可以提高目标视频的展示效果。之后，以渲染图像为素材，以视频时长为参数，进行视频转换，生成静态的视频，即目标视频。

在本实施例中，通过显示文本编辑页面，文本编辑页面内包括文本输入区；响应于针对文本编辑页面的第一输入指令，在文本输入区显示目标文本，
15 其中，目标文本在文本输入区内具有第一字体状态，第一字体状态表征目标文本的字号和/或行间距，第一字体状态是由目标文本的长度确定的；生成用于展示文本输入区内的目标文本的目标视频。通过设置文本输入区，并将在文本输入区内编辑的目标文本的字号和/或行间距能够随文本长度动态变化，之后对文本输入区内的目标文本进行转换，使生成的目标视频能够清晰、全
20 面地展示目标文本的全部内容，实现基于纯文本生成文本视频的目的，提高视频创造过程中的视频生成效率，提高视频平台内的视频内容多样性。

图 5 为本公开实施例提供的文本视频生成方法的流程示意图二。本实施例在图 2 所示实施例的基础上，对步骤 S102 进一步细化，并增加了为目标视频配置背景图片和背景音乐的步骤，该文本视频生成方法包括：

25 步骤 S201：显示文本编辑页面，文本编辑页面内包括文本输入区。
步骤 S202：响应于第一输入指令，生成目标文本，并获取目标文本的字符总数。

步骤 S203：根据字符总数和文本输入区的区域尺寸，确定第一字体状态。
示例性地，在显示编辑页面，并响应于针对编辑页面的第一输入指令后，
30 根据第一输入指令中的信息，可以生成对应的字符和符号，进而生成目标文

本，具体实现步骤在图 2 所示实施例中已进行详细介绍，此处不再赘述。之后，通过调用统计函数对目标文本对应的字符串进行处理，即可获得目标文本的字符总数，其中，本实施例中的字符总数，可以是目标文本中仅字符的个数，也可以是字符的个数和符号的个数的总和。

5 示例性地，文本编辑页面的文本输入区，具有区域尺寸，以文本输入区为矩形为例，文本输入区的区域尺寸可以为文本输入区的长度和宽度。进一步地，文本输入区的区域尺寸可以表现文本输入区的面积，文本输入区的面积越大，所能够显示的字符总数越多。文本输入区的区域尺寸为预设的固定值，可以根据终端设备的屏幕像素尺寸确定，此处不进行赘述。

10 进一步地，在确定字符总数和文本输入区的区域尺寸之后，首先基于文本输入区的区域尺寸，计算文本输入区的总面积，之后再利用总面积与字符总数的比值，即可得到字号（和符号）的所占用的单位面积，进而根据单位面积和预设的修正系数，即可确定字号和/或对应的行间距，即第一字体状态。上述实现方式未考虑到换行、空行造成的空白区域的影响，因此适合于对第一字体状态的粗略计算。

15 在另一种可能的实现方式中，区域尺寸包括区域横向尺寸和区域纵向尺寸，如图 6 所示，步骤 S203 的具体实现方式包括：

步骤 S2031：根据基准字号对应的字体宽度和区域横向尺寸，确定单行字符数，单行字符数表征文本输入区的一行所能够显示的字符数量。

20 步骤 S2032：根据单行字符数和字符总数，确定第一纵向尺寸。

步骤 S2033：根据第一纵向尺寸和区域纵向尺寸，确定第一字体状态。

图 7 为本公开实施例提供的一种文本输入区的区域尺寸示意图，如图 7 所示，在客户端的文本编辑页面内，文本输入区域为纵向设置的矩形区域（适配智能手机的显示屏），图中区域横向尺寸示为 dim_x ，区域纵向尺寸示为 dim_y 。用户在该文本输入区内输入目标文本。基准字号是预设的输入字符的字号，具有对应的字体宽度和字体高度，示例性地，当字体为中文字体时，文字的字体宽度和字体高度可以相同。根据区域横向尺寸和基准字号对应的字体宽度的比例值，可确定单行字符数，即在文本输入区的一行所能够显示的字符数量。例如为 20 个字符。之后，根据目标文本的字符总数除以单行字符数，可得到总行数，再基于字体高度与基准行间距，得到行平均高度，计算行

平均高度和总行数的乘积，即可得到第一纵向尺寸，图中示为 dim_Y 。其中，基准行间距与基准字号类似，也可以为预设值，此处不再举例赘述。

进一步地，在确定第一纵向尺寸后，对第一纵向尺寸和区域纵向尺寸进行对比，若第一纵向尺寸大于区域纵向尺寸，说明此时字号和/或行间距偏大，导致已超出文本输入区，进而使后续生成的目标视频无法以静态展示的方式显示目标文本的全部内容；此情况下，将影响第一纵向尺寸的基准字号和/或基准行间距缩小，得到第一字体状态。具体地，例如。可以在基准字号和/或基准行间距的基础上，乘以一个小于 1 的比例系数，得到第一字体状态。在另一种情况下，若第一纵向尺寸小于区域纵向尺寸，说明此时字号和/或行间距偏小，导致目标文本未合理的布局在文本输入区内，进而影响后续生成的目标视频的显示效果；此情况下，将影响第一纵向尺寸的基准字号和/或基准行间距放大，得到第一字体状态。具体地，例如。可以在基准字号和/或基准行间距的基础上，乘以一个大于 1 的比例系数，得到第一字体状态。

其中，进一步地，示例性地，调整基准字号和/或基准行间距的比例系数，可以通过第一纵向尺寸和区域纵向尺寸的比例关系确定。如图 8 所示，示例性地。步骤 S2033 的具体实现方式包括：

步骤 S2033A：获取第一纵向尺寸和区域纵向尺寸的比例值。

步骤 S2033B：当比例值小于第一比例阈值时，则基于基准字号，和/或基准行间距，确定第一字体状态。

步骤 S2033C：当比例值大于第一比例阈值时，则基于比例值降低基准字号，和/或，降低基准行间距，得到第一字体状态。

示例性地，在确定第一纵向尺寸和区域纵向尺寸后，计算第一纵向尺寸和区域纵向尺寸的比例值，例如第一纵向尺寸为 8（预设单位），区域纵向尺寸为 10，则第一纵向尺寸和区域纵向尺寸的比例值为 0.8。第一比例阈值表征文本输入区的占比，也即生成目标视频后的屏幕占比；当第一比例阈值为 1，表征文本输入区的全部区域；当第一比例阈值为 0.8，表征文本输入区的 80% 区域（理想展示区域）。进一步地，一种可能的情况下，当该比例值小于第一比例阈值时，说明目标文本未超出文本输入区，不会导致生成的目标视频无法显示全部目标文本的问题；同时由于基准字号和/或基准行间距可以认为是基于具体需要设置的理想参数，因此，可以使用基准字号和/或基准行间距来

作为第一字体状态。

另一种可能的情况下，当该比例值大于第一比例阈值时，说明说明目标文已经超出文本输入区或超出文本输入区的理想展示区域，会导致生成的目标视频无法显示全部目标文本的问题。此时，基于该比例值对基准字号或基准行间距进行调节，其中，第一比例阈值例如为 1，具体地，例如，当比例值大于或等于 0.8 且小于 1 时，降低一个字号；当比例值大于或等于 0.6，小于 0.8 时降低二个字号，降低两个字号，以此类推，从而将对基准字号和/或基准行间距降低后得到的目标字号和/或目标行间距，确定为第一字体状态。

步骤 S204：基于第一字体状态在文本输入区显示目标文本。

本实施例步骤中，通过计算第一纵向尺寸和区域纵向尺寸的比例值，并基于比例值在基准字号、基准行间距的基础上，对其进行缩小，从而动态确定匹配的第一字体状态。从而更加进准的控制文本输入区内的目标文本的长度，提高最终生成的目标视频的内容长度，提高文本视频的展示效果和用户的阅读体验。

步骤 S205：响应于针对文本编辑页面的第二输入指令，在文本编辑页面内显示背景图片。

示例性地，在上述步骤的基础上，文本编辑页面内还设置有助于配置背景图片的组件，用户通过操作指令触发该组件后，可以在文本编辑页面内显示背景图片。其中，具体地，背景图像对应的显示区域为背景图片区，该背景图片区覆盖文本输入区，图 9 为本公开实施例提供的一种背景图片的示意图，如图 9 所示，终端设备接收用户针对“添加背景”控件的点击操作后，生成并执行第二输入指令，在文件编辑页面内的背景图片区自动添加一张背景图片，其中，基于背景图片区覆盖文本输入区，且文本输入区位于背景图片的中心位置。在之后的步骤中，以该背景图片区为目标区域进行渲染，生成对应的目标视频。从而使所生成的目标视频中目标文本，显示在背景图片上，提高目标视频中目标文本的展示效果。

其中，示例性地，目标文本还具有第二字体状态，第二字体状态表征目标文本的字体颜色，如图 10 所示，步骤 S205 的具体实现方式包括：

步骤 S2051：响应于第二输入指令，基于第二字体状态，匹配目标颜色，目标颜色与字体颜色的色差大于色差阈值。

步骤 S2052: 基于预设图库, 获得主色调为目标颜色的背景图片, 并在文本编辑页面内显示背景图片。

示例性地, 在接收到第二输入指令后, 首先获取目标文本的第二字体状态, 即目标文本的字体颜色, 该第二字体状态可以是基于用户输入的操作指令设置的, 具体实现方式为现有技术, 不再赘述。之后, 根据第二字体状态所表征的字体颜色, 确定一个与其色差大于色差阈值的其他颜色, 即目标颜色。简言之, 即确定一个目标文本的字体颜色差异度高的颜色作为目标颜色。例如, 目标文本的字体颜色为黑色, 则目标颜色可以为白色、浅绿色、浅蓝色等 (与黑色的色差大于阈值)。之后从预设图库中, 获取主色调为该目标颜色的背景图片, 从而使背景图片与目标文本的字体颜色具有较高的对比度, 避免背景图片与目标文本的颜色相似而导致目标文本难以识别的问题, 提高目标文本在目标视频中的显示清晰度。

步骤 S206: 根据背景图片, 获取背景音乐。

进一步地, 在获得背景图片后, 在背景图片的基础上, 自动选择与背景图片相匹配的背景音乐。例如, 当背景图片的内容为“山水风景”时, 则根据背景图片的语义信息, 获得与其匹配的节奏舒缓的音乐类型的歌曲标签; 当背景图片的内容为“舞池中跳舞的人群”时, 则根据背景图片的语义信息, 获得与其匹配的节奏动感的音乐类型的歌曲标签。之后, 基于歌曲标签从曲库中选择对应的目标音乐作为背景音乐。其中, 获取背景图像的语义信息, 并基于语义信息获得对应的歌曲标签的具体实现方式, 可以通过预训练的图像语义模型实现, 此处不再赘述。

本实施例步骤中, 通过背景图片获得匹配的背景音乐, 使背景图片与背景音乐具有语义层面的一致性, 从而使最终生成的目标视频能够具有更好的视频表现力, 提高文本视频的质量。

步骤 S207: 根据文本输入区内的目标文本、背景图片和背景音乐, 生成目标视频。

在获得目标文本、背景图片和背景音乐, 基于目标文本和背景图片进行渲染, 得到对应的渲染图像, 上述过程在图 2 所示实施例中已进行介绍, 此处不再赘述。之后, 再将渲染图像与背景音乐进行合并和转换, 即可得到目标视频。其中, 示例性地, 本实施例中的目标视频的视频时长可以基

于背景音乐的时长确定，具体实现步骤再图 2 所示实施例中已进行介绍，此处不再赘述。

当然，可以理解的是，在其他可能的实施例中，在获得目标文本以及对应的第一字体状态后，可以仅获得与目标文本匹配的背景图片或背景音乐，进而基于目标文本和背景图片生成目标视频，或者，基于目标文本和背景音乐生成目标视频，其中，背景音乐可以是基于目标文本的语义信息确定的。具体实现过程与上述实施例步骤中的实现过程类似，不再赘述。

对应于上文实施例的文本视频生成方法，图 11 为本公开实施例提供的文本视频生成装置的结构框图。为了便于说明，仅示出了与本公开实施例相关的部分。参照图 11，文本视频生成装置 3 包括：

编辑模块 31，用于显示文本编辑页面，文本编辑页面内包括文本输入区；

显示模块 32，响应于针对文本编辑页面的第一输入指令，在文本输入区显示目标文本，其中，目标文本在文本输入区内具有第一字体状态，第一字体状态表征目标文本的字号和/或行间距，第一字体状态是由目标文本的长度确定的；

生成模块 33，用于生成用于展示文本输入区内的目标文本的目标视频。

在本公开的一个实施例中，显示模块 32，具体用于：响应于第一输入指令，生成目标文本，并获取目标文本的字符总数；根据字符总数和文本输入区的区域尺寸，确定第一字体状态；基于第一字体状态在文本输入区显示目标文本。

在本公开的一个实施例中，区域尺寸包括区域横向尺寸和区域纵向尺寸；显示模块 32 在根据字符总数和文本输入区的区域尺寸，确定第一字体状态时，具体用于：根据基准字号对应的字体宽度和区域横向尺寸，确定单行字符数，单行字符数表征文本输入区的一行所能够显示的字符数量；根据单行字符数和字符总数，确定第一纵向尺寸；根据第一纵向尺寸和区域纵向尺寸，确定第一字体状态。

在本公开的一个实施例中，显示模块 32 在根据第一纵向尺寸和区域纵向尺寸，确定第一字体状态时，具体用于：获取第一纵向尺寸和区域纵向尺寸的比例值；当比例值小于第一比例阈值时，则基于基准字号，和/或基准行间距，确定第一字体状态；当比例值大于第一比例阈值时，则基于比例值降低

基准字号，和/或，降低基准行间距，得到第一字体状态。

在本公开的一个实施例中，生成模块 33，具体用于：基于目标文本，生成包含具有第一字体状态的目标文本的渲染图像；根据目标文本的长度，确定视频时长；根据视频时长和渲染图像，生成目标视频。

- 5 在本公开的一个实施例中，在生成用于展示文本输入区内的目标文本的目标视频之前，显示模块 32，还用于：响应于针对文本编辑页面的第二输入指令，在文本编辑页面内显示背景图片；生成模块 33，具体用于：根据文本输入区内的目标文本和背景图片，生成目标视频。

- 10 在本公开的一个实施例中，目标文本还具有第二字体状态，第二字体状态表征目标文本的字体颜色；显示模块 32 在响应于针对文本编辑页面的第二输入指令，在文本编辑页面内显示背景图片时，具体用于：响应于第二输入指令，基于第二字体状态，匹配目标颜色，目标颜色与字体颜色的色差大于色差阈值；基于预设图库，获得主色调为目标颜色的背景图片，并在文本编辑页面内显示背景图片。

- 15 在本公开的一个实施例中，生成模块 33，还用于：根据背景图片，获取背景音乐；生成模块 33 在生成用于展示文本输入区内的目标文本的目标视频时，具体用于：根据文本输入区内的目标文本和背景音乐，生成目标视频。

- 20 其中，编辑模块 31、显示模块 32 和生成模块 33 依次连接。本实施例提供的文本视频生成装置 3 可以执行上述方法实施例的技术方案，其实现原理和技术效果类似，本实施例此处不再赘述。

图 12 为本公开实施例提供的一种电子设备的结构示意图，如图 12 所示，该电子设备 4 包括：

- 处理器 41，以及与处理器 41 通信连接的存储器 42；
存储器 42 存储计算机执行指令；
25 处理器 41 执行存储器 42 存储的计算机执行指令，以实现如图 2-图 10 所示实施例中的文本视频生成方法。

其中，可选地，处理器 41 和存储器 42 通过总线 43 连接。

相关说明可以对应参见图 2-图 10 所对应的实施例中的步骤所对应的相关描述和效果进行理解，此处不做过多赘述。

- 30 本公开实施例提供一种计算机可读存储介质，计算机可读存储介质中存

储有计算机执行指令，计算机执行指令被处理器执行时用于实现本申请图 2-图 10 所对应的实施例中任一实施例提供的文本视频生成。

参考图 13，其示出了适于用来实现本公开实施例的电子设备 900 的结构示意图，该电子设备 900 可以为终端设备或服务器。其中，终端设备可以包括包括但不限于诸如移动电话、笔记本电脑、数字广播接收器、个人数字助理（Personal Digital Assistant，简称 PDA）、平板电脑（Portable Android Device，简称 PAD）、便携式多媒体播放器（Portable Media Player，简称 PMP）、车载终端（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 13 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 13 所示，电子设备 900 可以包括处理装置（例如中央处理器、图形处理器等）901，其可以根据存储在只读存储器（Read Only Memory，简称 ROM）902 中的程序或者从存储装置 908 加载到随机访问存储器（Random Access Memory，简称 RAM）903 中的程序而执行各种适当的动作和处理。在 RAM 903 中，还存储有电子设备 900 操作所需的各种程序和数据。处理装置 901、ROM 902 以及 RAM 903 通过总线 904 彼此相连。输入/输出（I/O）接口 905 也连接至总线 904。

通常，以下装置可以连接至 I/O 接口 905：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 906；包括例如液晶显示器（Liquid Crystal Display，简称 LCD）、扬声器、振动器等输出装置 907；包括例如磁带、硬盘等的存储装置 908；以及通信装置 909。通信装置 909 可以允许电子设备 900 与其他设备进行无线或有线通信以交换数据。虽然图 13 示出了具有各种装置的电子设备 900，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 909 从网络上被下载和安装，或者从存储装置 908 被安装，或者从 ROM 902 被安装。在该计算机程序被处理装置 901 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一——但不限于——电、磁、光、电磁、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、5 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15 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。

20 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，使得该电子设备执行上述实施例所示的方法。

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25 的计算机程序代码，上述程序设计语言包括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30 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

(Local Area Network , 简称 LAN) 或广域网 (Wide Area Network , 简称 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, 单元的名称在某种情况下并不构成对该单元本身的限定, 例如, 第一获取单元还可以被描述为“获取至少两个网际协议地址的单元”。

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。

第一方面，根据本公开的一个或多个实施例，提供了一种文本视频生成方法，包括：

显示文本编辑页面，所述文本编辑页面内包括文本输入区；响应于针对文本编辑页面的第一输入指令，在所述文本输入区显示目标文本，其中，所述目标文本在所述文本输入区内具有第一字体状态，所述第一字体状态表征所述目标文本的字号和/或行间距，所述第一字体状态是由所述目标文本的长度确定的；生成用于展示所述文本输入区内的目标文本的目标视频。

根据本公开的一个或多个实施例，所述响应于针对文本编辑页面的第一输入指令，在所述文本输入区显示目标文本，包括：响应于所述第一输入指令，生成目标文本，并获取所述目标文本的字符总数；根据所述字符总数和所述文本输入区的区域尺寸，确定所述第一字体状态；基于所述第一字体状态在所述文本输入区显示所述目标文本。

根据本公开的一个或多个实施例，所述区域尺寸包括区域横向尺寸和区域纵向尺寸；所述根据所述字符总数和所述文本输入区的区域尺寸，确定所述第一字体状态，包括：根据基准字号对应的字体宽度和所述区域横向尺寸，确定单行字符数，所述单行字符数表征所述文本输入区的一行所能够显示的字符数量；根据所述单行字符数和所述字符总数，确定第一纵向尺寸；根据所述第一纵向尺寸和所述区域纵向尺寸，确定所述第一字体状态。

根据本公开的一个或多个实施例，根据所述第一纵向尺寸和所述区域纵向尺寸，确定所述第一字体状态，包括：获取所述第一纵向尺寸和所述区域纵向尺寸的比例值；当所述比例值小于第一比例阈值时，则基于所述基准字号，和/或基准行间距，确定所述第一字体状态；当所述比例值大于第一比例阈值时，则基于所述比例值降低所述基准字号，和/或，降低基准行间距，得到所述第一字体状态。

根据本公开的一个或多个实施例，生成用于展示所述文本输入区内的目标文本的目标视频，包括：基于所述目标文本，生成包含具有所述第一字体状态的目标文本的渲染图像；根据所述目标文本的长度，确定视频时长；根据所述视频时长和所述渲染图像，生成所述目标视频。

根据本公开的一个或多个实施例，在生成用于展示所述文本输入区内的目标文本的目标视频之前，所述方法还包括：响应于针对文本编辑页面的第

二输入指令，在所述文本编辑页面内显示背景图片；生成用于展示所述文本输入区内的目标文本的目标视频，包括：根据所述文本输入区内的目标文本和所述背景图片，生成所述目标视频。

5 根据本公开的一个或多个实施例，所述目标文本还具有第二字体状态，所述第二字体状态表征所述目标文本的字体颜色；响应于针对文本编辑页面的第二输入指令，在所述文本编辑页面内显示背景图片，包括：响应于所述第二输入指令，基于所述第二字体状态，匹配目标颜色，所述目标颜色与所述字体颜色的色差大于色差阈值；基于预设图库，获得主色调为所述目标颜色的背景图片，并在所述文本编辑页面内显示所述背景图片。

10 根据本公开的一个或多个实施例，所述方法还包括：根据所述背景图片，获取背景音乐；生成用于展示所述文本输入区内的目标文本的目标视频，包括：根据所述文本输入区内的目标文本和所述背景音乐，生成所述目标视频。

第二方面，根据本公开的一个或多个实施例，提供了一种文本视频生成装置，包括：

15 编辑模块，用于显示文本编辑页面，所述文本编辑页面内包括文本输入区；

显示模块，响应于针对文本编辑页面的第一输入指令，在所述文本输入区显示目标文本，其中，所述目标文本在所述文本输入区内具有第一字体状态，所述第一字体状态表征所述目标文本的字号和/或行间距，所述第一字体状态是由所述目标文本的长度确定的；

生成模块，用于生成用于展示所述文本输入区内的目标文本的目标视频。

20 根据本公开的一个或多个实施例，所述显示模块，具体用于：响应于所述第一输入指令，生成目标文本，并获取所述目标文本的字符总数；根据所述字符总数和所述文本输入区的区域尺寸，确定所述第一字体状态；基于所述第一字体状态在所述文本输入区显示所述目标文本。

25 根据本公开的一个或多个实施例，所述区域尺寸包括区域横向尺寸和区域纵向尺寸；所述显示模块在根据所述字符总数和所述文本输入区的区域尺寸，确定所述第一字体状态时，具体用于：根据基准字号对应的字体宽度和所述区域横向尺寸，确定单行字符数，所述单行字符数表征所述文本输入区的一行所能够显示的字符数量；根据所述单行字符数和所述字符总数，确定

第一纵向尺寸；根据所述第一纵向尺寸和所述区域纵向尺寸，确定所述第一字体状态。

根据本公开的一个或多个实施例，所述显示模块在根据所述第一纵向尺寸和所述区域纵向尺寸，确定所述第一字体状态时，具体用于：获取所述第一纵向尺寸和所述区域纵向尺寸的比例值；当所述比例值小于第一比例阈值时，则基于所述基准字号，和/或基准行间距，确定所述第一字体状态；当所述比例值大于第一比例阈值时，则基于所述比例值降低所述基准字号，和/或，降低基准行间距，得到所述第一字体状态。

根据本公开的一个或多个实施例，所述生成模块，具体用于：基于所述目标文本，生成包含具有所述第一字体状态的目标文本的渲染图像；根据所述目标文本的长度，确定视频时长；根据所述视频时长和所述渲染图像，生成所述目标视频。

根据本公开的一个或多个实施例，在生成用于展示所述文本输入区内的目标文本的目标视频之前，所述显示模块，还用于：响应于针对文本编辑页面的第二输入指令，在所述文本编辑页面内显示背景图片；所述生成模块，具体用于：根据所述文本输入区内的目标文本和所述背景图片，生成所述目标视频。

根据本公开的一个或多个实施例，所述目标文本还具有第二字体状态，所述第二字体状态表征所述目标文本的字体颜色；所述显示模块在响应于针对文本编辑页面的第二输入指令，在所述文本编辑页面内显示背景图片时，具体用于：响应于所述第二输入指令，基于所述第二字体状态，匹配目标颜色，所述目标颜色与所述字体颜色的色差大于色差阈值；基于预设图库，获得主色调为所述目标颜色的背景图片，并在所述文本编辑页面内显示所述背景图片。

根据本公开的一个或多个实施例，所述生成模块，还用于：根据所述背景图片，获取背景音乐；所述生成模块在生成用于展示所述文本输入区内的目标文本的目标视频时，具体用于：根据所述文本输入区内的目标文本和所述背景音乐，生成所述目标视频。

第三方面，根据本公开的一个或多个实施例，提供了一种电子设备，包括：处理器，以及与所述处理器通信连接的存储器；

所述存储器存储计算机执行指令；

所述处理器执行所述存储器存储的计算机执行指令，以实现如上第一方面以及第一方面各种可能的设计所述的文本视频生成方法。

5 第四方面，根据本公开的一个或多个实施例，提供了一种计算机可读存储介质，所述计算机可读存储介质中存储有计算机执行指令，当处理器执行所述计算机执行指令时，实现如上第一方面以及第一方面各种可能的设计所述的文本视频生成方法。

10 第五方面，本公开实施例提供一种计算机程序产品，包括计算机程序，该计算机程序被处理器执行时实现如上第一方面以及第一方面各种可能的设计所述的文本视频生成方法。

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
15 述特征与本公开中公开的（但不限于）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
20 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。

25 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、一种文本视频生成方法，其特征在于，包括：
- 显示文本编辑页面，所述文本编辑页面内包括文本输入区；
- 响应于针对文本编辑页面的第一输入指令，在所述文本输入区显示目标
- 5 文本，其中，所述目标文本在所述文本输入区内具有第一字体状态，所述第一字体状态表征所述目标文本的字号和/或行间距，所述第一字体状态是由所述目标文本的长度确定的；
- 生成用于展示所述文本输入区内的目标文本的目标视频。
- 2、根据权利要求 1 所述的方法，其特征在于，所述响应于针对文本编辑
- 10 页面的第一输入指令，在所述文本输入区显示目标文本，包括：
- 响应于所述第一输入指令，生成目标文本，并获取所述目标文本的字符总数；
- 根据所述字符总数和所述文本输入区的区域尺寸，确定所述第一字体状态；
- 15 基于所述第一字体状态在所述文本输入区显示所述目标文本。
- 3、根据权利要求 2 所述的方法，其特征在于，所述区域尺寸包括区域横向尺寸和区域纵向尺寸；所述根据所述字符总数和所述文本输入区的区域尺寸，确定所述第一字体状态，包括：
- 根据基准字号对应的字体宽度和所述区域横向尺寸，确定单行字符数，
- 20 所述单行字符数表征所述文本输入区的一行所能够显示的字符数量；
- 根据所述单行字符数和所述字符总数，确定第一纵向尺寸；
- 根据所述第一纵向尺寸和所述区域纵向尺寸，确定所述第一字体状态。
- 4、根据权利要求 3 所述的方法，其特征在于，根据所述第一纵向尺寸和所述区域纵向尺寸，确定所述第一字体状态，包括：
- 25 获取所述第一纵向尺寸和所述区域纵向尺寸的比例值；
- 当所述比例值小于第一比例阈值时，则基于所述基准字号，和/或基准行间距，确定所述第一字体状态；
- 当所述比例值大于第一比例阈值时，则基于所述比例值降低所述基准字号，和/或，降低基准行间距，得到所述第一字体状态。
- 30 5、根据权利要求 1 所述的方法，其特征在于，生成用于展示所述文本输

入区内的目标文本的目标视频，包括：

基于所述目标文本，生成包含具有所述第一字体状态的目标文本的渲染图像；

根据所述目标文本的长度，确定视频时长；

5 根据所述视频时长和所述渲染图像，生成所述目标视频。

6、根据权利要求1所述的方法，其特征在于，在生成用于展示所述文本输入区内的目标文本的目标视频之前，所述方法还包括：

响应于针对文本编辑页面的第二输入指令，在所述文本编辑页面内显示背景图片；

10 生成用于展示所述文本输入区内的目标文本的目标视频，包括：

根据所述文本输入区内的目标文本和所述背景图片，生成所述目标视频。

7、根据权利要求6所述的方法，其特征在于，所述目标文本还具有第二字体状态，所述第二字体状态表征所述目标文本的字体颜色；

15 响应于针对文本编辑页面的第二输入指令，在所述文本编辑页面内显示背景图片，包括：

响应于所述第二输入指令，基于所述第二字体状态，匹配目标颜色，所述目标颜色与所述字体颜色的色差大于色差阈值；

基于预设图库，获得主色调为所述目标颜色的背景图片，并在所述文本编辑页面内显示所述背景图片。

20 8、根据权利要求6所述的方法，其特征在于，所述方法还包括：

根据所述背景图片，获取背景音乐；

生成用于展示所述文本输入区内的目标文本的目标视频，包括：

根据所述文本输入区内的目标文本和所述背景音乐，生成所述目标视频。

9、一种文本视频生成装置，其特征在于，包括：

25 编辑模块，用于显示文本编辑页面，所述文本编辑页面内包括文本输入区；

显示模块，响应于针对文本编辑页面的第一输入指令，在所述文本输入区显示目标文本，其中，所述目标文本在所述文本输入区内具有第一字体状态，所述第一字体状态表征所述目标文本的字号和/或行间距，所述第一字体状态是由所述目标文本的长度确定的；

30

生成模块，用于生成用于展示所述文本输入区内的目标文本的目标视频。

10、一种电子设备，其特征在于，包括：处理器，以及与所述处理器通信连接的存储器；

所述存储器存储计算机执行指令；

5 所述处理器执行所述存储器存储的计算机执行指令，以实现如权利要求 1 至 8 中任一项所述的文本视频生成方法。

11、一种计算机可读存储介质，其特征在于，所述计算机可读存储介质中存储有计算机执行指令，当处理器执行所述计算机执行指令时，实现如权利要求 1 至 8 任一项所述的文本视频生成方法。

10 12、一种计算机程序产品，其特征在于，包括计算机程序，该计算机程序被处理器执行时实现权利要求 1 至 8 中任一项所述的文本视频生成方法。