(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
2022 年 5 月 19 日 (19.05.2022)

WIPO | PCT

(10) 国际公布号
WO 2022/100103 A1

(51) 国际专利分类号:
H04N 21/431 (2011.01) *H04N 21/4402* (2011.01)
H04N 21/435 (2011.01) *H04N 21/442* (2011.01)
H04N 21/439 (2011.01)

(21) 国际申请号: PCT/CN2021/102438

(22) 国际申请日: 2021 年 6 月 25 日 (25.06.2021)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
202011247956.3 2020年11月10日 (10.11.2020) 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: 郑霓雯(ZHENG, Niwen); 中国北京市海淀区知春路63号中国卫星通信大厦今日头条小邮局, Beijing 100190 (CN)。 孙磊(SUN, Lei); 中国北京市海淀区知春路63号中国卫星通信大厦今日头条小邮局, Beijing 100190 (CN)。 瞿佳(QU, Jia); 中国北京市海淀区知春路63号中国卫星通信大厦今日头条小邮局, Beijing 100190 (CN)。 吴昊(WU, Hao); 中国北京市海淀区知春路63号中国卫星通信大厦今日头条小邮局, Beijing 100190 (CN)。

(74) 代理人: 北京知帆远景知识产权代理有限公司 (ZHIFAN & PARTNERS); 中国北京市海淀区阜成路73号裕惠大厦B座805, Beijing 100142 (CN)。

WO 2022/100103 A1

(54) **Title:** LYRICS VIDEO DISPLAY METHOD AND APPARATUS, ELECTRONIC DEVICE, AND COMPUTER READABLE MEDIUM

(54) **发明名称:** 歌词视频展示方法、装置、电子设备及计算机可读介质

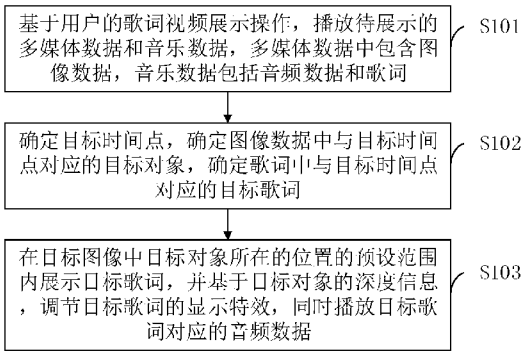


图 1

S101 On the basis of a lyrics video display operation by a user, play back multimedia data and music data to be displayed, the multimedia data comprising image data, and the music data comprising audio data and lyrics

S102 Determine a target time point, determine a target object corresponding to the target time point in the image data, and determine target lyrics corresponding to the target time point in the lyrics

S103 Display the target lyrics in a preset range of a position in which the target object is located in a target image, adjust a display special effect of the target lyrics on the basis of depth information of the target object, and play back audio data corresponding to the target lyrics

(57) **Abstract:** Provided in the present application are a lyrics video display method and apparatus, an electronic device, and a computer readable medium. The method comprises: on the basis of a lyrics video display operation by a user, playing back multimedia data and music data to be displayed; determining a target object corresponding to a target time point in image data; determining target lyrics corresponding to the target time point in lyrics; displaying the target lyrics in a preset range of a position in which the target object is located in a target image; adjusting a display special effect of the target lyrics on the basis of depth information of the target object; and playing back said multimedia data and music data.

[见续页]

(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, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT, 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, 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))。

(57) 摘要：本申请提供一种歌词视频展示方法、装置、电子设备及计算机可读介质，该方法包括：基于用户的歌词视频展示操作，播放待展示的多媒体数据和音乐数据，确定图像数据中与目标时间点对应的目标对象，确定歌词中与目标时间点对应的目标歌词；在目标图像中目标对象所在的位置的预设范围内展示目标歌词，基于目标对象的深度信息，调节目标歌词的显示特效，播放待展示的多媒体数据和音乐数据。

歌词视频展示方法、装置、电子设备及计算机可读介质

本申请要求北京字跳网络技术有限公司于 2020 年 11 月 10 日提交的，申请名称为“歌词视频展示方法、装置、电子设备及计算机可读介质”的、中国专利申请号为“202011247956.3”的优先权，该中国专利申请的全部内容通过引用结合在本申请中。

技术领域

本申请涉及视频处理技术领域，具体而言，本申请涉及一种歌词视频展示方法、装置、电子设备及计算机可读介质。

背景技术

随着视频技术的发展，人们对音乐视频的要求也越来越高，音乐视频中出现音乐歌词已经是一个十分常见的功能。

现有的音乐视频中，在播放音乐时会在视频下方滚动或平移的出现歌词，有些技术还会有给歌词进行染色的功能，但是这些都只是简单的将歌词与视频进行简单的叠加，歌词的进场、出场都是一些基础特效，并且对于有背景的歌词视频，歌词与背景之间完全脱离，没有相关性，导致用户观看歌词视频的体验不佳，并且由于歌词仅仅只是一些简单机械的基础特效，歌词展现形式比较单一，导致用户体验不佳。

由此可见，现有技术中存在歌词展示形式单一，且歌词与视频背景完全脱离，导致用户体验不佳的技术问题，需要解决。

技术问题

本申请的目的旨在至少能解决上述的技术缺陷之一，特别是现有技术中存在歌词展示形式单一，且歌词与视频背景完全脱离，导致用户体验不佳的技术问题。

技术解决方案

第一方面，提供了一种歌词视频展示方法，该方法包括：

基于用户的歌词视频展示操作，播放待展示的多媒体数据和音乐数据，所述多媒体数据中包含图像数据，所述音乐数据包括音频数据和歌词；

确定目标时间点，确定所述图像数据中与所述目标时间点对应的目标对象，确定所述歌词中所述目标时间点对应的目标歌词；

在所述目标图像中所述目标对象所在的位置的预设范围内展示所述目标歌词，并基于所述目标对象的深度信息，调节所述目标歌词的显示特效，同时播放所述目标歌词对应的音频数据。

第二方面，提供了一种歌词视频展示装置，该装置包括：

数据获取模块，用于基于用户的歌词视频展示操作，播放待展示的多媒体数据和音乐数据，所述多媒体数据中包含图像数据，所述音乐数据包括音频数据和歌词；

歌词确定模块，用于确定目标时间点，确定所述图像数据中与所述目标时间点对应的目标对象，确定所述歌词中所述目标时间点对应的目标歌词；

歌词展示模块，用于在所述目标图像中所述目标对象所在的位置的预设范围内展示所述目标歌词，并基于所述目标对象的深度信息，调节所述目标歌词的显示特效，同时播放所述目标歌词对应的音频数据。第三方面，提供了一种电子设备，该电子设备包括：

一个或多个处理器；

存储器；

一个或多个应用程序，其中所述一个或多个应用程序被存储在所述存储器中并被配置为由所述一个或多个处理器执行，所述一个或多个程序配置用于：执行上述的歌词视频展示方法。

第四方面，提供了一种计算机可读介质，所述可读存储有至少一条指令、至少一段程序、代码集或指令集，所述至少一条指令、所述至少一段程序、所述代码集或指令集由所述处理器加载并执行以实现上述的歌词视频展示方法。

有益效果

本申请实施例基于用户的歌词视频展示操作，播放待展示的多媒体数据和音乐数据，其中，多媒体数据包括图像数据，音乐数据包括音频数据和歌词，同时，确定图像数据中的目标对象和歌词中的目标歌词，并在目标对象的附近展示目标歌词，并基于目标对象的深度信息调节歌词的显示特效，将歌词嵌入到图像数据的现实空间中，给用户一种虚拟显示的感官体验，用户体验更好。

附图说明

为了更清楚地说明本申请实施例中的技术方案，下面将对本申请实施例描述中所需要使用的附图作简单地介绍。

图 1 为本申请实施例提供的一种歌词视频展示方法的流程示意图；

图 2 为本申请实施例提供的一种显示界面示意图；

图 3 为本申请实施例提供的一种目标移动过程示意图；

图 4 为本申请实施例提供的一种获取多媒体数据方法的流程示意图；

图 5 为本申请实施例提供的一种用户捕捉多媒体数据的方法的流程示意图；

图 6 为本申请实施例提供的一种多媒体数据选择界面示意图；

图 7 为本申请实施例提供的一种生成歌词贴片方法的流程示意图；

图 8 为本申请实施例提供的一种调节歌词位置的方法的流程示意图；

图 9 为本申请实施例提供的一种特效调节示意图；

图 10 为本申请实施例提供的一种调节歌词大小的示意图；

图 11 为本申请实施例提供的一种歌词视频展示装置的结构示意图；

图 12 为本申请实施例提供的一种电子设备的结构示意图。

结合附图并参考以下具体实施方式，本申请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。

本发明的实施方式

下面将参照附图更详细地描述本申请的实施例。虽然附图中显示了本申请的某些实施例，然而应当理解的是，本申请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申请。应当理解的是，本申请的附图及实施例仅用于示例性作用，并非用于限制本申请的保护范围。

应当理解，本申请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。本申请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，本申请中提及的“第一”、“第二”等概念仅用于对装置、模块或单元进行区分，并非用于限定这些装置、模块或单元一定为不同的装置、模块或单元，也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，本申请中提及的“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”。

本申请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。

本申请提供的歌词视频展示方法、装置、电子设备和计算机可读介质，旨在解决现有技术的如上技术问题。

下面以具体地实施例对本申请的技术方案以及本申请的技术方案如何解决上述技术问题进行详细说明。下面这几个具体的实施例可以相互结合，对于相同或相似的概念或过程可能在某些实施例中不再赘述。下面将结合附图，对本申请的实施例进行描述。

本申请实施例中提供了一种歌词视频展示方法，应用于用户终端，该用户终端可以是手机、平板电脑等移动终端，该用户终端中安装有 APP（Application，应用程序），通过该应用程序或者该应用程序中的某个功能可以实现本申请实施例提供的歌词视频展

示方法，如图 1 所示，该方法包括：

步骤 S101，基于用户的歌词视频展示操作，播放待展示的多媒体数据和音乐数据，多媒体数据中包含图像数据，音乐数据包括音频数据和歌词；

步骤 S102，确定目标时间点，确定图像数据中与目标时间点对应的目标对象，确定歌词中与目标时间点对应的目标歌词；

步骤 S103，在目标图像中目标对象所在的位置的预设范围内展示目标歌词，并基于目标对象的深度信息，调节目标歌词的显示特效，同时播放目标歌词对应的音频数据。

在本申请实施例中，多媒体数据包括图像数据和音乐数据，其中图像数据可以是带有景深信息的图片、动图以及视频等，音乐数据包括音频数据和歌词，深度信息是用于表示图像数据中目标对象距离图像数据采集装置的距离的信息，与图像数据采集装置距离不同的目标对象在图像数据中显示的成像比例不同。

对于本申请实施例，在获取待合成的多媒体数据和音乐时，可以是获取本地存储的多媒体数据和音乐，也可以是网络上的多媒体数据和音乐，还可以是用户自己录制的多媒体数据和音乐，其中，该多媒体数据中包含有图像数据，可选的，该图像数据可以是图片数据，也可以是视频数据。

在一些实施例中，在基于目标对象的深度信息，调节目标歌词的显示特效之前，还包括：

获取图像数据中目标对象的深度信息，其中，目标对象的深度信息用于表示目标对象距离视频捕捉装置的距离的信息。

在本申请实施例中，获取图像数据中目标对象的深度信息，其中，目标对象是指图像数据中与歌词结合的对象，可以是图像数据中某个物体，也可以是模块区域，目标对象的深度信息是指用于表示该目标对象距离视频捕捉装置的距离的信息，深度越深，表示该目标对象距离视频捕捉装置越远，深度越浅，表示该目标对象距离视频捕捉装置的距离越近。

在本申请实施例中，在基于目标对象在图像数据中的位置调节歌词在图像数据中的显示位置时，是基于图像数据在播放时，目标对象的显示位置调节与该目标对象对应的歌词的显示位置的。可选的，如通过第一界面显示图像数据时，目标对象显示在图像数据中上部的位置，则将歌词显示在该目标对象对应的位置，即图像数据中上部的位置，也可以是该目标对象附近的位置。

本申请实施例中，在基于目标对象在图像数据中的深度信息，调节歌词的显示特效时，调节特效时，可以根据不同目标对象的不同的深度信息，调节对应的歌词贴片的特效。

对于本申请实施例，为方便说明，以一个具体实施例为例说明本申请实施例提供的展示效果，如图 2 所示，可以通过第一界面 201 对图像数据进行展示，可选的，图像数据为视频数据，通过第一界面 201 播放视频，同时播放音乐，基于视频中目标对象 A 的位置，调节对应的歌词 A1 的位置，如图 2 所示，目标对象在视频中的左上方，则对应的歌词 A1 出现的位置为视频的左上方，可选的，如图 3 所示，在视频和音乐的播放过程中，目标对象从视频的左上方移动到视频的右下方，则对应的歌词也会随着目标对象移动到视频的右下方。可选的，目标对象 A 在移动过程中，其深度信息越来越深，即目标对象 A 距离视频捕捉装置的距离越来越远，可以对应的调节歌词 A1 的特效。

本申请实施例基于用户的歌词视频展示操作，播放待展示的多媒体数据和音乐数据，其中，多媒体数据包括图像数据，音乐数据包括音频数据和歌词，同时，确定图像数据中的目标对象和歌词中的目标歌词，并在目标对象的附近展示目标歌词，并基于目标对象的深度信息调节歌词的显示特效，将歌词嵌入到图像数据的现实空间中，给用户一种虚拟显示的感官体验，用户体验更好。

在一些实施例中，如图 4 所示，基于用户的歌词视频展示操作，播放待展示的多媒体数据和音乐数据之前，还包括：

步骤 S401，接收用户的多媒体数据选择操作；

步骤 S402，基于多媒体数据选择操作确定多媒体数据。

在本申请实施例中，用户可以从本地存储的多媒体数据中选择想要的多媒体数据，可选的，可以是通过触控选择的方式选择，也可以是通过语音、动作等方式进行选择。

对于本申请实施例，为方便说明，以一个具体实施例为例，如图 5 所示，接收用户的多媒体数据选择操作，并显示多媒体数据选择界面 501，该多媒体数据选择界面 501 中有可供选择的多媒体数据 502，基于用户的选择操作，确定用户想要合成歌词视频的多媒体数据。

本申请实施例中用户可以自己选择想要合成歌词视频的多媒体数据，用户可以根据自己需要，选择不同的多媒体数据，提升用户体验。

在一些实施例中，如图 6 所示，基于用户的歌词视频展示操作，播放待展示的多媒体数据和音乐数据之前，还包括：

步骤 S601，基于用户的多媒体数据捕捉操作，开启多媒体数据捕捉装置；

步骤 S602，获取多媒体数据捕捉装置捕捉的多媒体数据。

在本申请实施例中，用户可以通过用户终端捕捉多媒体数据，该用户终端应该设置有多媒体数据捕捉装置、或外界有多媒体数据捕捉装置，可选的，为图像捕捉装置。

对于本申请实施例，为方便说明，以一个具体实施例为例，用户终端为带有摄像头

的手机，通过手机接收用户的多媒体数据捕捉操作，可选的，该多媒体数据捕捉操作可以是用户的触控操作，基于该触控操作，开启收集自带的摄像头，并通过摄像头捕捉图像数据。

本申请实施例中用户可以自己拍摄视频，并对视频进行合成，视频选择范围更广，用户体验佳。

在一些实施例中，如图 7 所示，在目标图像中目标对象所在的位置的预设范围内展示目标歌词，包括：

步骤 S701，基于目标歌词生成对应的歌词贴片；

步骤 S702，在目标图像中目标对象的位置的预设范围内展示歌词贴片。

在本申请实施例中，歌词贴片是歌词对应的一种展现形式，如可以是显示有歌词的一张图片，或者一张动图，能给显示歌词内容，可选的，可以是一句歌词对应一个歌词贴片，也可以是一句长歌词对应多个歌词贴片，还可以是多句短歌词对应一个歌词贴片，不同的对应方式，会呈现不同的展示效果。可选的，一句歌词对应一个歌词贴片。

在一些实施例中，基于目标歌词生成对应的歌词贴片，包括：

对目标歌词进行分句或者分词处理，以生成对应的歌词贴片。

在一些实施例中，歌词贴片中显示有歌词内容，在生成对应的歌词贴片时，还包括：

对歌词贴片中显示的歌词内容进行排版。

在一些实施例中，对歌词贴片中显示的歌词内容进行排版，包括：

根据歌词贴片中的歌词的字数、字体、颜色、对齐方式、显示位置中的至少一种，对歌词贴片中显示的歌词内容进行排版。

在本申请实施例中，在根据音乐的歌词生成歌词贴片时，获取该音乐中的歌词，按照预设的规则对歌词进行分句、分词，如将一句歌词形成一个歌词贴片，或者将一句歌词中的某几个词语形成一个歌词贴片，还可以将多句歌词联合起来排版形成一个歌词贴片，可选的，在形成歌词贴片时，对歌词进行排版，可选的，可以调节歌词的字体、颜色等，最终形成一个能够显示歌词内容的图片或者动图。

对于本申请实施例，为方便说明，以一个具体实施例为例，获取音乐的歌词，并对歌词进行分句处理，对每一句歌词，对应生成一个歌词贴片，可选的，该歌词贴片中显示有歌词内容；可选的，可以对歌词按照预设的规则进行排版，如规定每行显示歌词的字数、歌词的字体、对齐方式等，可选的，每个歌词贴片可以显示有不同数量的歌词，每个歌词贴片中歌词的显示位置可以不同，每个歌词贴片中歌词的对齐方式可以不同。

本申请实施例通过获取音乐的歌词，并对每一句歌词都生成对应的歌词贴片，保证歌词贴片的多样性，展示效果更丰富。

在一些实施例中，如图 8 所示，在目标图像中目标对象所在的位置的预设范围内展示目标歌词，包括：

步骤 S801，基于目标歌词在音乐中的播放时间段确定多媒体数据中的目标图像数据；

步骤 S802，基于目标图像数据中的目标对象在目标图像数据中的位置调节目标歌词在图像数据中的显示位置，其中显示位置在目标对象在目标图像数据中的位置的预设范围内。

在本申请实施例中，在将歌词和图像数据中的目标对象进行结合时，歌词对应的歌词在音乐中出现的时间段和目标对象在图像数据中出现的时间段要相对应，保证歌词出现时，目标对象也能在图像数据中出现。

对于本申请实施例，为方便说明，以一个具体实施例为例，对于歌词 B，其对应的歌词在音乐中出现的时间段为 35S~38S，则确定图像数据中在第 35S~38S 的目标图像，确定该目标图像中一直存在的目标对象 B1，将目标对象 B1 确定为与该歌词 B 对应的目标对象，并基于该目标对象 B1 在图像数据中的位置信息，确定歌词贴片的位置。作为本申请一个实施例，歌词出现的位置，也可以是与该歌词对应的目标对象显示位置的附近区域。

本申请实施例通过歌词贴片对应的歌词在音乐中出现的时间和目标对象在图像数据中出现的时间确定歌词贴片与目标对象的对应关系，保证歌词在播放时能够在数据图像中出现，用户可以看到歌词。

在一些实施例中，基于目标对象在图像数据中的深度信息，调节目标歌词的显示特效，包括：

基于目标对象在目标图像中的深度信息，调节目标歌词的显示大小。

在本申请实施例中，调节歌词贴片的显示特效可以是调节歌词的显示大小，为方便说明，以一个具体实施例为例，如图 9 所示，目标对象 C1 在图像数据播放的过程中，距离图像捕捉装置越来越远，其深度越来越深，则可以调节该目标对象 C1 对应的歌词 C 的显示大小越来越大。

在一些实施例中，基于目标对象的深度信息，调节目标歌词的显示特效，包括：

基于目标对象在目标图像中的深度信息，调节歌词贴片的显示颜色。

可选的，特效的调节方式还可以是根据目标对象的深度信息调节对应的歌词贴片的显示颜色等。

本申请实施例通过根据目标对象的深度信息调节与之对应的歌词的显示大小，给用户一种远小近大的视觉体验。

在一些实施例中，图像数据包括视频数据和/或带有景深信息的全景图像数据，其中，当图像数据为带有景深信息的全景图像数据时，在目标图像中目标对象所在的位置的预设范围内展示目标歌词，包括：

基于全景图像数据中不同目标对象的位置，分别调节不同目标对象对应的目标歌词在全景图像数据中的显示位置。

在本申请实施例中，为方便说明，以一个具体实施例为例，如图 10 所示，在全景图像 1001 中，有目标对象 D1 和目标对象 E1，其中，目标对象 D1 在全景图像的左下方，目标对象 E1 在全景图像的中间，其中，目标对象 D1 对应歌词 D，目标对象 E1 对应歌词 E，则分别根据目标对象 D1 和 E1 的位置调节歌词 D 和 E 的位置，其中，目标对象 D1 的深度信息较浅，目标对象 E1 的深度信息较深，则可以调节歌词 D 显示较大，歌词 E 显示较小。

本申请实施例通过全景图像中不同目标对象的位置和深度信息调节对应歌词的大小，给用户真实的空间感受。

本申请实施例基于音乐的歌词生成对应的歌词贴片，并基于多媒体数据中的图像数据获取图像数据中目标对象的深度信息，基于该目标对象在该图像数据中的位置信息调节与该目标对象对应的歌词贴片在该图像数据中的显示位置，并基于该目标对象在该图像数据中的深度信息调节对应的歌词贴片的显示特效，将歌词嵌入到图像数据的现实空间中，给用户一种虚拟显示的感官体验，用户体验更好。

本申请实施例提供了一种歌词视频展示装置，如图 11 所示，该歌词视频展示装置 110 可以包括：数据获取模块 1110、歌词确定模块 1120 以及歌词展示模块 1130，其中，

数据获取模块 1110，用于基于用户的歌词视频展示操作，播放待展示的多媒体数据和音乐数据，多媒体数据中包含图像数据，音乐数据包括音频数据和歌词；

歌词确定模块 1120，用于确定目标时间点，确定图像数据中与目标时间点对应的目标对象，确定歌词中目标时间点对应的目标歌词；

歌词展示模块 1130，用于在目标图像中目标对象所在的位置的预设范围内展示目标歌词，并基于目标对象的深度信息，调节目标歌词的显示特效，同时播放目标歌词对应的音频数据。

在一些实施例中，数据获取模块 1110 在获取待合成的多媒体数据时，可以用于：

接收用户的多媒体数据选择操作；基于多媒体数据选择操作确定多媒体数据。

在一些实施例中，数据获取模块 1110 在获取待合成的多媒体数据时，可以用于：

基于用户的多媒体数据捕捉操作，开启多媒体数据捕捉装置；

获取多媒体数据捕捉装置捕捉的多媒体数据。

在一些实施例中，歌词展示模块 1130 在目标图像中目标对象所在的位置的预设范围内展示目标歌词时，可以用于：基于目标歌词生成对应的歌词贴片；在目标图像中目标对象的位置的预设范围内展示歌词贴片。

在一些实施例中，歌词展示模块 1130 在基于目标歌词生成对应的歌词贴片时，可以用于：对目标歌词进行分句或者分词处理，以生成对应的歌词贴片。

在一些实施例中，歌词贴片中显示有歌词内容，歌词展示模块 1130 在生成对应的歌词贴片时，还可以用于：对歌词贴片中显示的歌词内容进行排版。

在一些实施例中，歌词展示模块 1130 在对歌词贴片中显示的歌词内容进行排版时，可以用于：根据歌词贴片中的歌词的字数、字体、颜色、对齐方式、显示位置中的至少一种，对歌词贴片中显示的歌词内容进行排版。

在一些实施例中，歌词展示模块 1130 在基于目标对象在图像数据中的深度信息，调节目标歌词的显示特效时，可以用于：

基于目标对象在目标图像中的深度信息，调节歌词贴片的显示颜色。

在一些实施例中，歌词展示模块 1130 在目标图像中目标对象所在的位置的预设范围内展示目标歌词时，可以用于：

基于目标歌词在音乐中的播放时间段确定多媒体数据中的目标图像数据；

基于目标图像数据中的目标对象在目标图像数据中的位置调节目标歌词在图像数据中的显示位置，其中显示位置在目标对象在目标图像数据中的位置的预设范围内。

在一些实施例中，歌词展示模块 1130 在基于目标对象在图像数据中的位置调节目标歌词在图像数据中的显示位置时，可以用于：

基于目标对象在目标图像中的深度信息，调节目标歌词的显示大小。

在一些实施例中，图像数据包括视频数据和/或带有景深信息的全景图像数据，其中，当图像数据为带有景深信息的全景图像数据时，歌词展示模块 1130 在目标图像中目标对象所在的位置的预设范围内展示目标歌词时，可以用于：

基于全景图像数据中不同目标对象的位置，分别调节不同目标对象对应的歌词贴片在全景图像数据中的显示位置。

在一些实施例中，歌词展示模块 1130 在所述基于所述目标对象的深度信息，调节所述目标歌词的显示特效之前，还可以用于：

获取所述图像数据中所述目标对象的深度信息，其中，所述目标对象的深度信息用于表示所述目标对象距离视频捕捉装置的距离的信息。

本申请实施例基于音乐的歌词生成对应的歌词贴片，并基于多媒体数据中的图像数据获取图像数据中目标对象的深度信息，基于该目标对象在该图像数据中的位置信息调

节与该目标对象对应的歌词贴片在该图像数据中的显示位置，并基于该目标对象在该图像数据中的深度信息调节对应的歌词贴片的显示特效，将歌词嵌入到图像数据的现实空间中，给用户一种虚拟显示的感官体验，用户体验更好。

下面参考图 12，其示出了适于用来实现本申请实施例的电子设备 1200 的结构示意图。本申请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 12 示出的电子设备仅仅是一个示例，不应对本申请实施例的功能和使用范围带来任何限制。

电子设备包括：存储器以及处理器，其中，这里的处理器可以称为下文的处理装置 1201，存储器可以包括下文中的只读存储器（ROM）1202、随机访问存储器（RAM）1203 以及存储装置 1208 中的至少一项，具体如下所示：

如图 12 所示，电子设备 1200 可以包括处理装置（例如中央处理器、图形处理器等）1201，其可以根据存储在只读存储器（ROM）1202 中的程序或者从存储装置 1208 加载到随机访问存储器（RAM）1203 中的程序而执行各种适当的动作和处理。在 RAM 1203 中，还存储有电子设备 1200 操作所需的各种程序和数据。处理装置 1201、ROM 1202 以及 RAM 1203 通过总线 1204 彼此相连。输入/输出（I/O）接口 1205 也连接至总线 1204。

通常，以下装置可以连接至 I/O 接口 1205：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 1206；包括例如液晶显示器（LCD）、扬声器、振动器等的输出装置 1207；包括例如磁带、硬盘等的存储装置 1208；以及通信装置 1209。通信装置 1209 可以允许电子设备 1200 与其他设备进行无线或有线通信以交换数据。虽然图 12 示出了具有各种装置的电子设备 1200，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，根据本申请的实施例，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，本申请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 1209 从网络上被下载和安装，或者从存储装置 1208 被安装，或者从 ROM 1202 被安装。在该计算机程序被处理装置 1201 执行时，执行本申请实施例的方法中限定的上述功能。

需要说明的是，本申请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介质或者是上述两者的任意组合。计算机可读介质例如可以是——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）、光存储器件、磁存储器件、或者上述的任意合适的组合。在本申请中，计算机可读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。而在本申请中，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。

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： 获取待合成的多媒体数据和音乐，多媒体数据中包含图像数据；获取图像数据中目标对象的深度信息；根据音乐的歌词生成歌词贴片；基于目标对象在图像数据中的位置调节歌词贴片在图像数据中的显示位置；基于目标对象在图像数据中的深度信息，调节歌词贴片的显示特效，并生成歌词视频。

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申请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)—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，图示了按照本申请各种实施例的系统、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。也要注意的是，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。

描述于本申请实施例中所涉及到的模块或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。

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。

在本申请的上下文中，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。

根据本申请提供的一个或多个实施例，提供了一种歌词视频展示方法，该方法包括：

基于用户的歌词视频展示操作，播放待展示的多媒体数据和音乐数据，多媒体数据中包含图像数据，音乐数据包括音频数据和歌词；确定目标时间点，确定图像数据中与目标时间点对应的目标对象，确定歌词中与目标时间点对应的目标歌词；在目标图像中目标对象所在的位置的预设范围内展示目标歌词，并基于目标对象的深度信息，调节目标歌词的显示特效，同时播放目标歌词对应的音频数据。

在一些实施例中，基于用户的歌词视频展示操作，播放待展示的多媒体数据和音乐数据之前，还包括：接收用户的多媒体数据选择操作；基于多媒体数据选择操作确定多媒体数据。

在一些实施例中，基于用户的歌词视频展示操作，播放待展示的多媒体数据和音乐数据之前，还包括：基于用户的多媒体数据捕捉操作，开启多媒体数据捕捉装置；获取多媒体数据捕捉装置捕捉的多媒体数据。

在一些实施例中，在目标图像中目标对象所在的位置的预设范围内展示目标歌词，包括：基于目标歌词生成对应的歌词贴片；在目标图像中目标对象的位置的预设范围内展示歌词贴片。

在一些实施例中，基于目标歌词生成对应的歌词贴片，包括：

对目标歌词进行分句或者分词处理，以生成对应的歌词贴片。

在一些实施例中，歌词贴片中显示有歌词内容，在生成对应的歌词贴片时，还包括：对歌词贴片中显示的歌词内容进行排版。

在一些实施例中，对歌词贴片中显示的歌词内容进行排版，包括：

根据歌词贴片中的歌词的字数、字体、颜色、对齐方式、显示位置中的至少一种，对歌词贴片中显示的歌词内容进行排版。

在一些实施例中，基于目标对象的深度信息，调节目标歌词的显示特效，包括：

基于目标对象在目标图像中的深度信息，调节歌词贴片的显示颜色。

在一些实施例中，在目标图像中目标对象所在的位置的预设范围内展示目标歌词，包括：基于目标歌词在音乐中的播放时间段确定多媒体数据中的目标图像数据；基于目标图像数据中的目标对象在目标图像数据中的位置调节目标歌词在图像数据中的显示位置，其中显示位置在目标对象在目标图像数据中的位置的预设范围内。

在一些实施例中，基于目标对象的深度信息，调节目标歌词的显示特效，包括：

基于目标对象在目标图像中的深度信息，调节目标歌词的显示大小。

在一些实施例中，图像数据包括视频数据和/或带有景深信息的全景图像数据，其中，当图像数据为带有景深信息的全景图像数据时，在目标图像中目标对象所在的位置的预设范围内展示目标歌词，包括：基于全景图像数据中不同目标对象的位置，分别调节不同目标对象对应的目标歌词在全景图像数据中的显示位置。

在一些实施例中，在基于目标对象的深度信息，调节目标歌词的显示特效之前，还包括：获取图像数据中目标对象的深度信息，其中，目标对象的深度信息用于表示目标对象距离视频捕捉装置的距离的信息。

根据本申请提供的一个或多个实施例，提供了一种歌词视频展示装置，该装置包括：

数据获取模块，用于基于用户的歌词视频展示操作，播放待展示的多媒体数据和音乐数据，多媒体数据中包含图像数据，音乐数据包括音频数据和歌词；

歌词确定模块，用于确定目标时间点，确定图像数据中与目标时间点对应的目标对

象，确定歌词中目标时间点对应的目标歌词；

歌词展示模块，用于在目标图像中目标对象所在的位置的预设范围内展示目标歌词，并基于目标对象的深度信息，调节目标歌词的显示特效，同时播放目标歌词对应的音频数据。

在一些实施例中，数据获取模块在获取待合成的多媒体数据时，可以用于：接收用户的多媒体数据选择操作；基于多媒体数据选择操作确定多媒体数据。

在一些实施例中，数据获取模块在获取待合成的多媒体数据时，可以用于：基于用户的多媒体数据捕捉操作，开启多媒体数据捕捉装置；获取多媒体数据捕捉装置捕捉的多媒体数据。

在一些实施例中，歌词展示模块在目标图像中目标对象所在的位置的预设范围内展示目标歌词时，可以用于：基于目标歌词生成对应的歌词贴片；在目标图像中目标对象的位置的预设范围内展示歌词贴片。

在一些实施例中，歌词展示模块在基于目标歌词生成对应的歌词贴片时，可以用于：

对目标歌词进行分句或者分词处理，以生成对应的歌词贴片。

在一些实施例中，歌词贴片中显示有歌词内容，歌词展示模块在生成对应的歌词贴片时，还可以用于：对歌词贴片中显示的歌词内容进行排版。

在一些实施例中，歌词展示模块在对歌词贴片中显示的歌词内容进行排版时，可以用于：根据歌词贴片中的歌词的字数、字体、颜色、对齐方式、显示位置中的至少一种，对歌词贴片中显示的歌词内容进行排版。

在一些实施例中，歌词展示模块在基于目标对象在图像数据中的深度信息，调节目标歌词的显示特效时，可以用于：

基于目标对象在目标图像中的深度信息，调节歌词贴片的显示颜色。

在一些实施例中，歌词展示模块在目标图像中目标对象所在的位置的预设范围内展示目标歌词时，可以用于：

基于目标歌词在音乐中的播放时间段确定多媒体数据中的目标图像数据；

基于目标图像数据中的目标对象在目标图像数据中的位置调节目标歌词在图像数据中的显示位置，其中显示位置在目标对象在目标图像数据中的位置的预设范围内。

在一些实施例中，歌词展示模块在基于目标对象在图像数据中的位置调节歌词贴片在图像数据中的显示位置时，可以用于：

基于目标对象在目标图像中的深度信息，调节目标歌词的显示大小。

在一些实施例中，歌词展示模块在基于目标对象在图像数据中的深度信息，调节歌词贴片的显示特效时，可以用于：

基于目标对象在图像数据中的深度信息，调节歌词贴片的显示大小。

在一些实施例中，图像数据包括视频数据和/或带有景深信息的全景图像数据，其中，当图像数据为带有景深信息的全景图像数据时，歌词展示模块在在目标图像中目标对象所在的位置的预设范围内展示目标歌词时，可以用于：

基于全景图像数据中不同目标对象的位置，分别调节不同目标对象对应的歌词贴片在全景图像数据中的显示位置。

在一些实施例中，歌词展示模块在所述基于所述目标对象的深度信息，调节所述目标歌词的显示特效之前，还可以用于：

获取所述图像数据中所述目标对象的深度信息，其中，所述目标对象的深度信息用于表示所述目标对象距离视频捕捉装置的距离的信息。

根据本申请提供的一个或多个实施例，提供了一种电子设备，该电子设备包括：一个或多个处理器；存储器；一个或多个应用程序，其中一个或多个应用程序被存储在存储器中并被配置为由一个或多个处理器执行，一个或多个程序配置用于：执行上述实施例的歌词视频展示方法。

根据本申请实施例提供的一个或多个实施例，提供了一种计算机可读介质，该介质存储有至少一条指令、至少一段程序、代码集或指令集，所述至少一条指令、所述至少一段程序、所述代码集或指令集由所述处理器加载并执行以实现上述实施例所述的歌词视频展示方法。

以上描述仅为本申请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，本申请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。例如上述特征与本申请中公开的（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，但是这不应当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，但是这些不应当被解释为对本申请的范围的限制。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。

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、一种歌词视频展示方法，其包括：

基于用户的歌词视频展示操作，播放待展示的多媒体数据和音乐数据，所述多媒体数据包含图像数据，所述音乐数据包括音频数据和歌词；

确定目标时间点，确定所述图像数据中与所述目标时间点对应的目标对象，确定所述歌词中与所述目标时间点对应的目标歌词；

在所述目标图像中所述目标对象所在的位置的预设范围内展示所述目标歌词，并基于所述目标对象的深度信息，调节所述目标歌词的显示特效，同时播放所述目标歌词对应的音频数据。

2、根据权利要求 1 所述的方法，其中，所述基于用户的歌词视频展示操作，播放待展示的多媒体数据和音乐数据之前，还包括：

接收用户的多媒体数据选择操作；

基于所述多媒体数据选择操作确定多媒体数据。

3、根据权利要求 1 所述的方法，其中，所述基于用户的歌词视频展示操作，播放待展示的多媒体数据和音乐数据之前，还包括：

基于用户的多媒体数据捕捉操作，开启多媒体数据捕捉装置；

获取所述多媒体数据捕捉装置捕捉的多媒体数据。

4、根据权利要求 1 所述的方法，其中，所述在所述目标图像中所述目标对象所在的位置的预设范围内展示所述目标歌词，包括：

基于所述目标歌词生成对应的歌词贴片；

在所述目标图像中所述目标对象的位置的预设范围内展示所述歌词贴片。

5、根据权利要求 4 所述的方法，其中，所述基于所述目标歌词生成对应的歌词贴片，包括：

对所述目标歌词进行分句或者分词处理，以生成对应的歌词贴片。

6、根据权利要求 5 所述的方法，其中，所述歌词贴片中显示有歌词内容，所述在生成对应的歌词贴片时，还包括：

对所述歌词贴片中显示的所述歌词内容进行排版。

7、根据权利要求 6 所述的方法，其中，对所述歌词贴片中显示的所述歌词内容进行排版，包括：

根据所述歌词贴片中的歌词的字数、字体、颜色、对齐方式、显示位置中的至少一种，对所述歌词贴片中显示的所述歌词内容进行排版。

8、根据权利要求 4 所述的方法，其中，所述基于所述目标对象的深度信息，调节

所述目标歌词的显示特效，包括：

基于所述目标对象在所述目标图像中的深度信息，调节所述歌词贴片的显示颜色。

9、根据权利要求 1 所述的方法，其中，所述在所述目标图像中所述目标对象所在的位置的预设范围内展示所述目标歌词，包括：

基于所述目标歌词在所述音乐中的播放时间段确定所述多媒体数据中的目标图像；

基于所述目标图像中的目标对象在所述目标图像中的位置调节所述目标歌词在所述目标图像中的显示位置，其中所述显示位置在所述目标对象在所述目标图像中的位置的预设范围内。

10、根据权利要求 1 所述的方法，其中，所述基于所述目标对象的深度信息，调节所述目标歌词的显示特效，包括：

基于所述目标对象在所述目标图像中的深度信息，调节所述目标歌词的显示大小。

11、根据权利要求 1 所述的方法，其中，所述图像数据包括视频数据和/或带有景深信息的全景图像数据，其中，当所述图像数据为带有景深信息的全景图像数据时，所述在所述目标图像中所述目标对象所在的位置的预设范围内展示所述目标歌词，包括：

基于所述全景图像数据中不同目标对象的位置，分别调节所述不同目标对象对应的目标歌词在所述全景图像数据中的显示位置。

12、根据权利要求 1 所述的方法，其中，在所述基于所述目标对象的深度信息，调节所述目标歌词的显示特效之前，还包括：

获取所述图像数据中所述目标对象的深度信息，其中，所述目标对象的深度信息用于表示所述目标对象距离视频捕捉装置的距离的信息。

13、一种歌词视频展示装置，其包括：

数据获取模块，用于基于用户的歌词视频展示操作，播放待展示的多媒体数据和音乐数据，所述多媒体数据中包含图像数据，所述音乐数据包括音频数据和歌词；

歌词确定模块，用于确定目标时间点，确定所述图像数据中与所述目标时间点对应的目标对象，确定所述歌词中所述目标时间点对应的目标歌词；

歌词展示模块，用于在所述目标图像中所述目标对象所在的位置的预设范围内展示所述目标歌词，并基于所述目标对象的深度信息，调节所述目标歌词的显示特效，同时播放所述目标歌词对应的音频数据。

14、一种电子设备，其包括：

一个或多个处理器；

存储器；

一个或多个应用程序，其中所述一个或多个应用程序被存储在所述存储器中并被配

置为由所述一个或多个处理器执行，所述一个或多个程序配置用于：执行根据权利要求1~12任一项所述的歌词视频展示方法。

15、一种计算机可读介质，其中，所述介质存储有至少一条指令、至少一段程序、代码集或指令集，所述至少一条指令、所述至少一段程序、所述代码集或指令集由所述处理器加载并执行以实现如权利要求1~12任一所述的歌词视频展示方法。

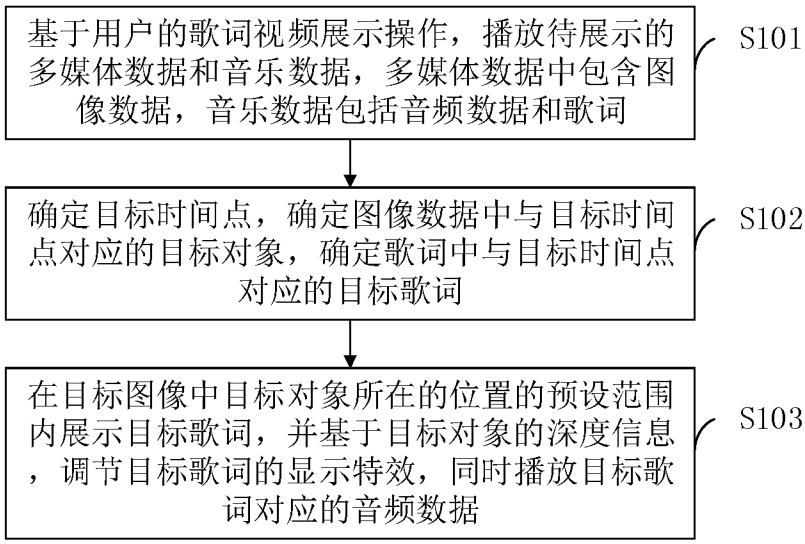

基于用户的歌词视频展示操作，播放待展示的多媒体数据和音乐数据，多媒体数据中包含图像数据，音乐数据包括音频数据和歌词
S101
确定目标时间点，确定图像数据中与目标时间点对应的目标对象，确定歌词中与目标时间点对应的目标歌词
S102
在目标图像中目标对象所在的位置的预设范围内展示目标歌词，并基于目标对象的深度信息，调节目标歌词的显示特效，同时播放目标歌词对应的音频数据
S103

图 1

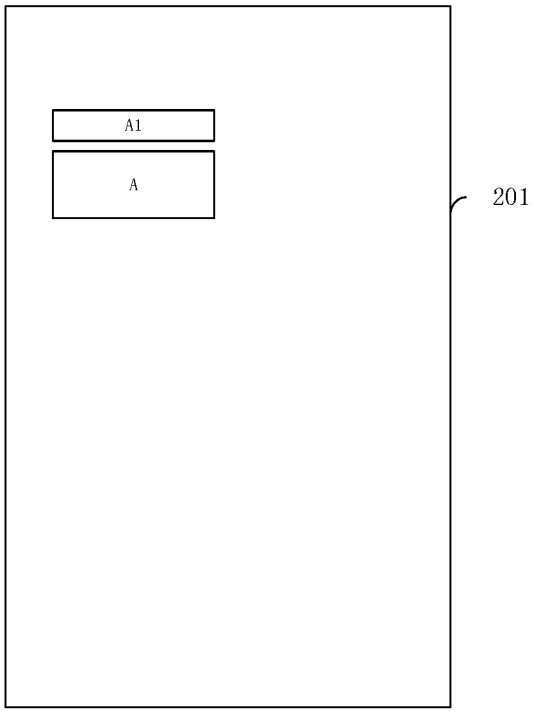

A1
A
201

图 2

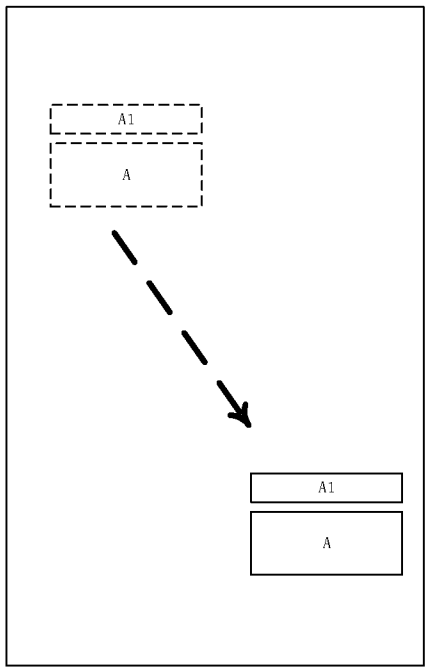

A1
A
A1
A

图 3

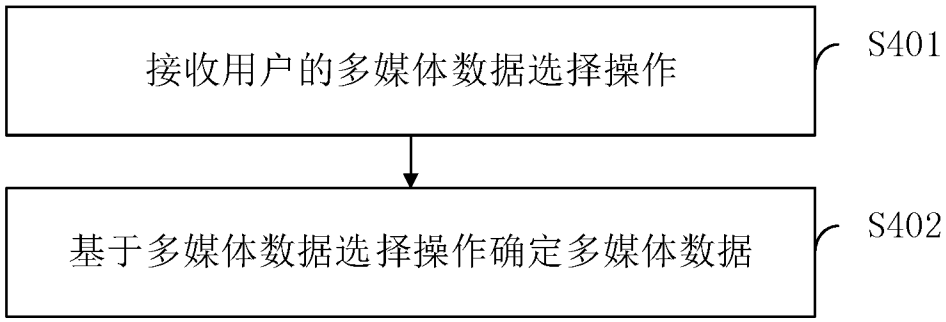

接收用户的多媒体数据选择操作
S401
基于多媒体数据选择操作确定多媒体数据
S402

图 4

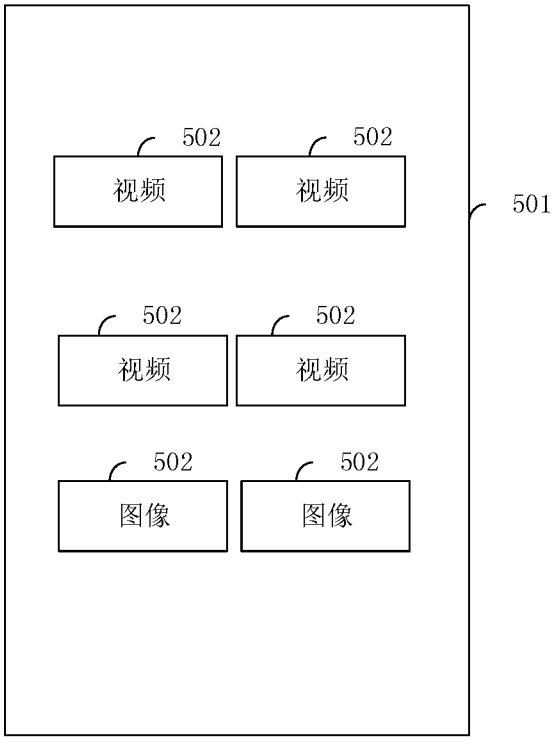

502
视频
502
视频
501
502
视频
502
视频
502
图像
502
图像

图 5

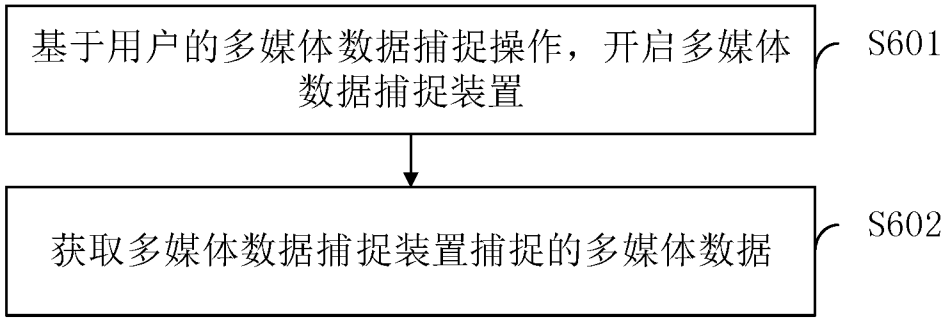

基于用户的多媒体数据捕捉操作，开启多媒体数据捕捉装置
S601
获取多媒体数据捕捉装置捕捉的多媒体数据
S602

图 6

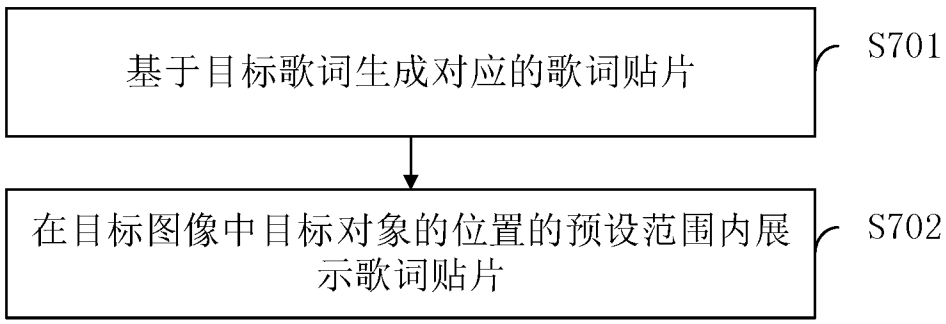

基于目标歌词生成对应的歌词贴片
S701
在目标图像中目标对象的位置的预设范围内展示歌词贴片
S702

图 7

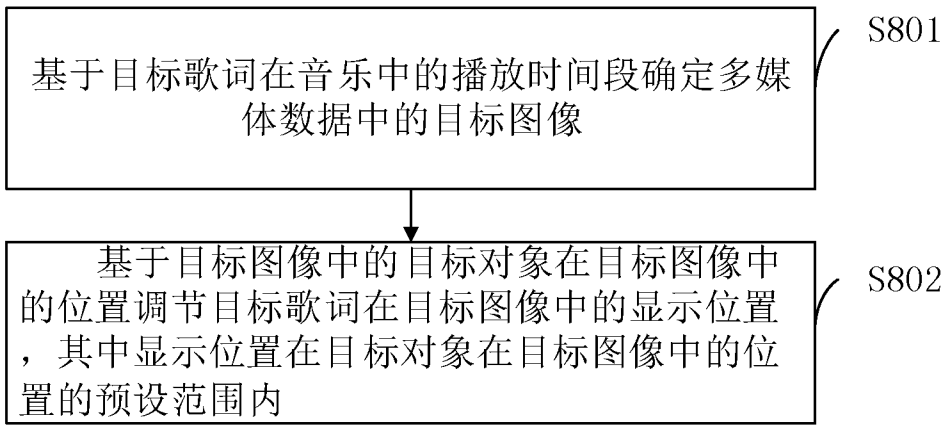

基于目标歌词在音乐中的播放时间段确定多媒体数据中的目标图像
S801
基于目标图像中的目标对象在目标图像中的位置调节目标歌词在目标图像中的显示位置，其中显示位置在目标对象在目标图像中的位置的预设范围内
S802

图 8

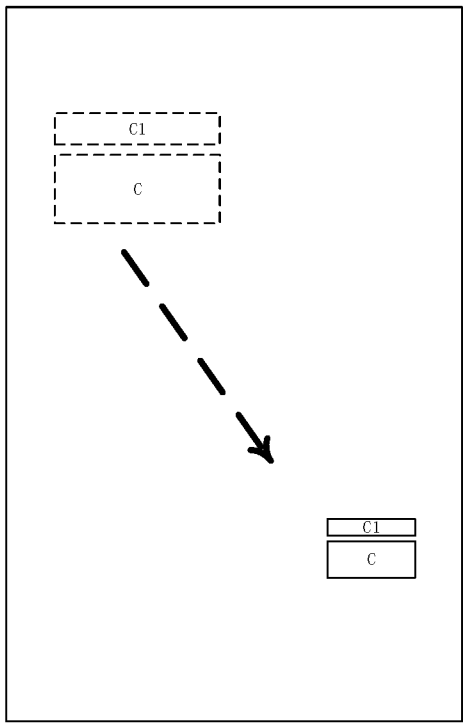


图 9

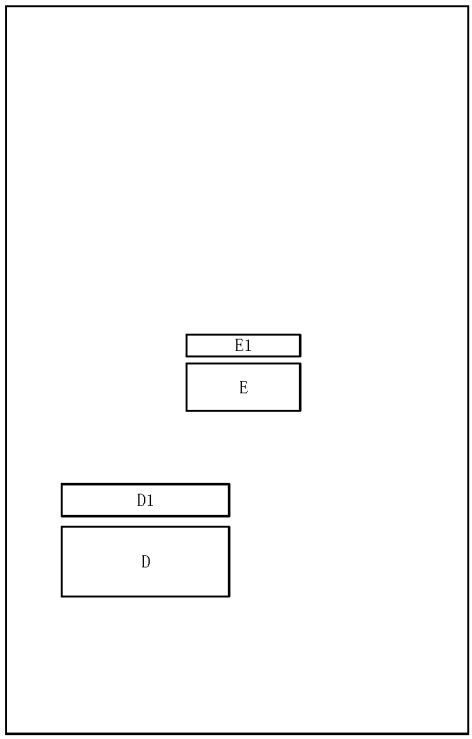


图 10

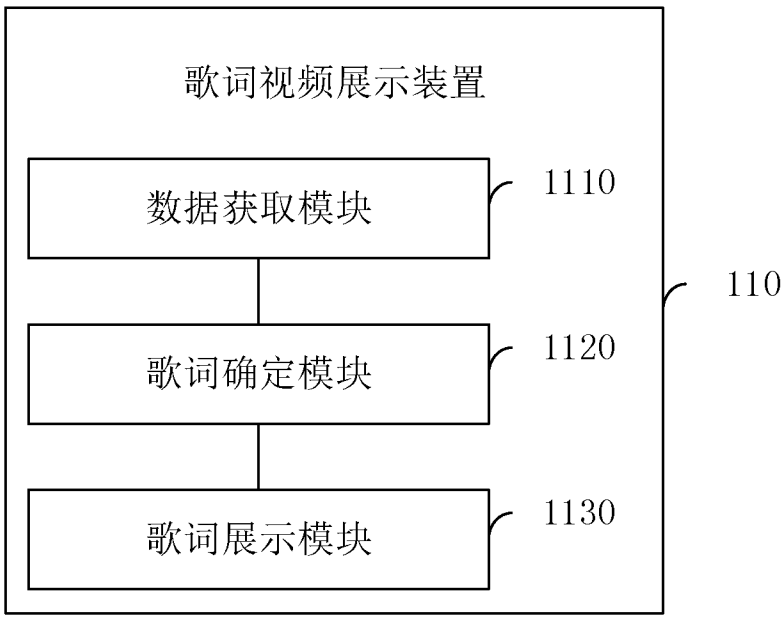

歌词视频展示装置
数据获取模块
1110
歌词确定模块
1120
歌词展示模块
1130
110

图 11

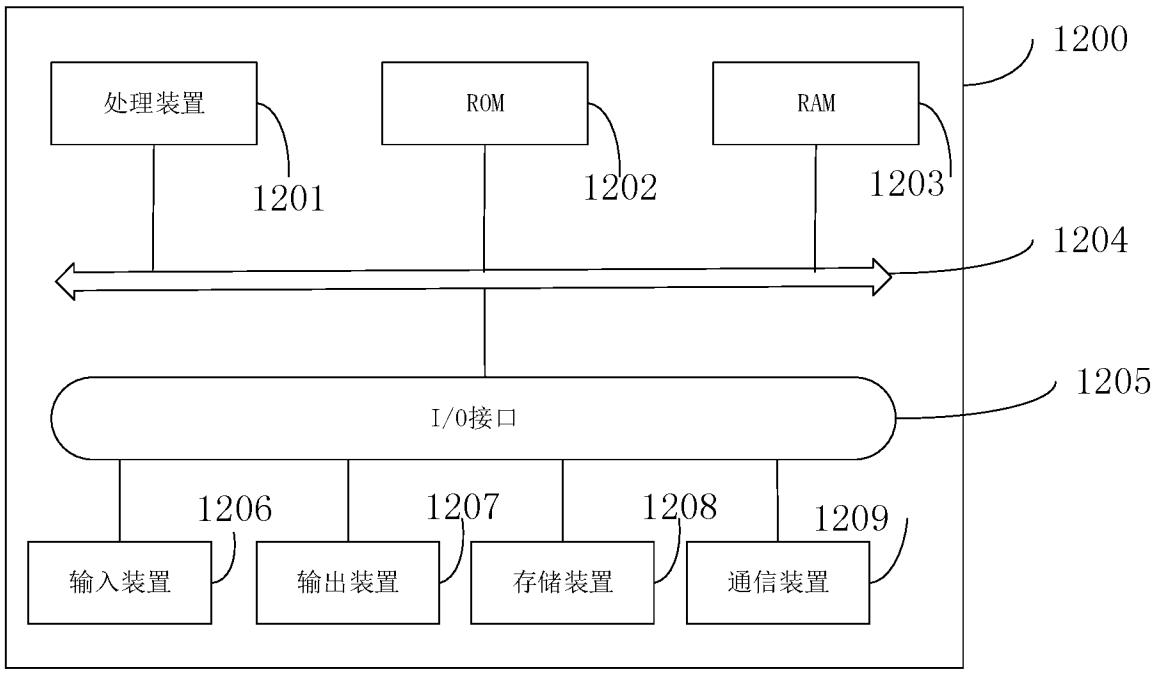

处理装置
ROM
RAM
1200
1201
1202
1203
1204
I/O接口
1205
1206
1207
1208
1209
输入装置
输出装置
存储装置
通信装置

图 12

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2021/102438

A. CLASSIFICATION OF SUBJECT MATTER

H04N 21/431(2011.01)i; H04N 21/435(2011.01)i; H04N 21/439(2011.01)i; H04N 21/4402(2011.01)i; H04N 21/442(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)

CNABS, CNTXT, VEN, USTXT, EPTXT, WOTXT, CNKI: 歌词, 字幕, 文本, 文字, 位置, 空间, 深度, 景深, 对象, 人物, 目标, 人像, 面部, 口, 嘴, 嵌入, 添加, 增加, 合并, 虚拟, lyric, text, caption, position, location, space, depth, deepness, object, character, person, target, face, mouth, add, embed, virtual

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 112383810 A (BEIJING ZITIAO NETWORK TECHNOLOGY CO., LTD.) 19 February 2021 (2021-02-19) entire document	1-15
Y	CN 107944397 A (TENCENT MUSIC ENTERTAINMENT TECHNOLOGY (SHENZHEN) CO., LTD.) 20 April 2018 (2018-04-20) description, paragraphs 0087-0198, and figures 1-18	1-15
Y	WO 2020105847 A1 (SAMSUNG ELECTRONICS CO., LTD.) 28 May 2020 (2020-05-28) description paragraphs 0160-0265	1-15
Y	CN 107943964 A (TENCENT MUSIC ENTERTAINMENT TECHNOLOGY (SHENZHEN) CO., LTD.) 20 April 2018 (2018-04-20) description paragraphs 0094-0210	1-15
A	CN 108089830 A (VIVO COMMUNICATION TECHNOLOGY CO., LTD.) 29 May 2018 (2018-05-29) entire document	1-15
A	CN 104219559 A (ADOBE SYSTEMS INC.) 17 December 2014 (2014-12-17) 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
"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
09 September 2021	**23 September 2021**

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/CN2021/102438

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 2011007868 A (DAIICHIKOSHO CO., LTD.) 13 January 2011 (2011-01-13) entire document	1-15

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2021/102438

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	112383810	A	19 February 2021	None			
CN	107944397	A	20 April 2018	None			
WO	2020105847	A1	28 May 2020	None			
CN	107943964	A	20 April 2018	None			
CN	108089830	A	29 May 2018	CN	108089830	B	29 October 2019
CN	104219559	A	17 December 2014	US	9467750	B2	11 October 2016
				GB	201409580	D0	16 July 2014
				GB	2516745	A	04 February 2015
				CN	104219559	B	12 April 2019
				DE	102014008038	A1	11 December 2014
				US	2014359656	A1	04 December 2014
				GB	2516745	B	03 August 2016
JP	2011007868	A	13 January 2011	JP	5371574	B2	18 December 2013

国际检索报告

国际申请号
PCT/CN2021/102438

A. 主题的分类

H04N 21/431(2011.01)i; H04N 21/435(2011.01)i; H04N 21/439(2011.01)i; H04N 21/4402(2011.01)i; H04N 21/442(2011.01)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)

H04N

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）)

CNABS,CNTXT,VEN,USTXT,EPTXT,WOTXT,CNKI: 歌词，字幕，文本，文字，位置，空间，深度，景深，对象，人物，目标，人像，面部，口，嘴，嵌入，添加，增加，合并，虚拟，lyric, text, caption, position, location, space, depth, deepness, object, character, person, target, face, mouth, add, embed, virtual

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
PX	CN 112383810 A (北京字跳网络技术有限公司) 2021年 2月 19日 (2021-02-19) 全文	1-15
Y	CN 107944397 A (腾讯音乐娱乐科技深圳有限公司) 2018年 4月 20日 (2018-04-20) 说明书第0087-0198段，图1-18	1-15
Y	WO 2020105847 A1 (三星电子株式会社) 2020年 5月 28日 (2020-05-28) 说明书第0160-0265段	1-15
Y	CN 107943964 A (腾讯音乐娱乐科技深圳有限公司) 2018年 4月 20日 (2018-04-20) 说明书第0094-0210段	1-15
A	CN 108089830 A (维沃移动通信有限公司) 2018年 5月 29日 (2018-05-29) 全文	1-15
A	CN 104219559 A (奥多比公司) 2014年 12月 17日 (2014-12-17) 全文	1-15
A	JP 2011007868 A (DAIICHIKOSHO CO LTD) 2011年 1月 13日 (2011-01-13) 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

"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	国际检索报告邮寄日期
2021年 9月 9日	2021年 9月 23日

ISA/CN的名称和邮寄地址	受权官员
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	雷云珊
传真号 (86-10)62019451	电话号码 86-(010)-62411508

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2021/102438

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	112383810	A	2021年 2月 19日	无			
CN	107944397	A	2018年 4月 20日	无			
WO	2020105847	A1	2020年 5月 28日	无			
CN	107943964	A	2018年 4月 20日	无			
CN	108089830	A	2018年 5月 29日	CN	108089830	B	2019年 10月 29日
CN	104219559	A	2014年 12月 17日	US	9467750	B2	2016年 10月 11日
				GB	201409580	D0	2014年 7月 16日
				GB	2516745	A	2015年 2月 4日
				CN	104219559	B	2019年 4月 12日
				DE	102014008038	A1	2014年 12月 11日
				US	2014359656	A1	2014年 12月 4日
				GB	2516745	B	2016年 8月 3日
JP	2011007868	A	2011年 1月 13日	JP	5371574	B2	2013年 12月 18日