

(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
2021 年 12 月 30 日 (30.12.2021)



(10) 国际公布号
WO 2021/259221 A1

(51) 国际专利分类号:
G06F 40/58 (2020.01)

(21) 国际申请号: PCT/CN2021/101388

(22) 国际申请日: 2021 年 6 月 22 日 (22.06.2021)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
202010583177.4 2020年6月23日 (23.06.2020) CN

(71) 申请人: 北京字节跳动网络技术有限公司 (BEIJING BYTEDANCE NETWORK TECHNOLOGY CO., LTD.) [CN/CN]; 中国北京市石景山区实兴大街30号院3号楼2层B-0035房间, Beijing 100041 (CN)。

(72) 发明人: 杜育璋(DU, Yuzhang); 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)。 朱培豪(ZHU, Peihao); 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)。 陈一鸣(CHEN, Yiming); 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)。 周崇兴(ZHOU, Chongxing); 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)。 王明轩(WANG, Mingxuan); 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)。 李磊(LI, 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)。

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

(54) 发明名称: 视频翻译方法和装置、存储介质和电子设备

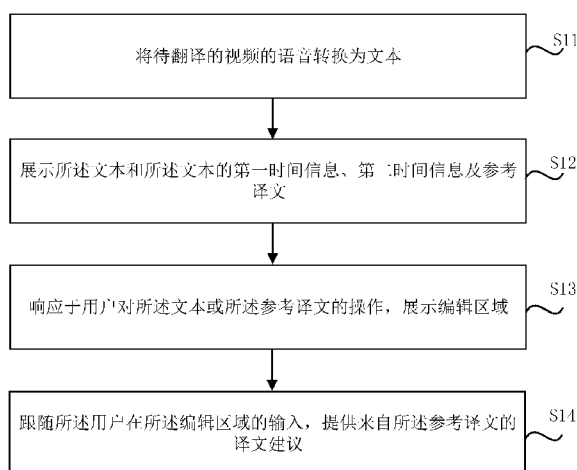


图 1

S11 CONVERT SPEECH IN A VIDEO TO BE TRANSLATED INTO TEXT
S12 DISPLAY THE TEXT AND FIRST TIME INFORMATION, SECOND TIME INFORMATION, AND A REFERENCE TRANSLATION OF THE TEXT
S13 IN RESPONSE TO AN OPERATION BY A USER ON THE TEXT OR THE REFERENCE TRANSLATION, DISPLAY AN EDITING AREA
S14 FOLLOWING INPUT BY THE USER IN THE EDITING AREA, PROVIDE A TRANSLATION SUGGESTION FROM THE REFERENCE TRANSLATION

(57) Abstract: The present disclosure relates to a video translation method and apparatus, a storage medium, and an electronic device. The method comprises: converting speech in a video to be translated into text; displaying the text and first time information, second time information, and a reference translation of the text; in response to an operation by a user on the text or the reference translation, displaying an editing area, the editing area supporting the user inputting a translation; and following input by the user in the editing area, providing a translation suggestion from the reference translation; when a confirmation operation by the user for the translation

(74) 代理人: 北京信远达知识产权代理有限公司
(SINOTALEN ATTORNEYS AT LAW); 中国北京市
北京经济技术开发区科创十三街18号院
10号楼301室, Beijing 100176 (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,
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))。

suggestion is detected, using the translation suggestion as a translation result and displaying same in the editing area; and when a non-confirmation operation by the user for the translation suggestion is detected, receiving a translation that is inputted by the user and that is different from the translation suggestion, using the translation inputted by the user as a translation result and displaying same in the editing area, and, according to the translation inputted by the user, updating the reference translation in a translation area. The present disclosure can improve the efficiency and quality of translation.

(57) 摘要: 本公开涉及一种视频翻译方法、装置、存储介质和电子设备, 所述方法包括: 将待翻译的视频的语音转换为文本; 展示所述文本和所述文本的第一时间信息、第二时间信息及参考译文; 响应于用户对所述文本或所述参考译文的操作, 展示编辑区域, 所述编辑区域支持所述用户输入译文; 跟随所述用户在所述编辑区域的输入, 提供来自所述参考译文的译文建议; 其中, 在检测到所述用户针对所述译文建议的确认操作的情况下, 将所述译文建议作为译文结果显示到所述编辑区域; 以及, 在检测到所述用户针对所述译文建议的非确认操作的情况下, 接收所述用户输入的不同于所述译文建议的译文, 将所述用户输入的所述译文作为所述译文结果显示到所述编辑区域, 根据所述用户输入的所述译文更新所述译文区域中的参考译文。本公开可以提升翻译的效率以及翻译的质量。

视频翻译方法和装置、存储介质和电子设备

本申请要求于 2020 年 06 月 23 日提交中国国家知识产权局、申请号为 202010583177.4、申请名称为“视频翻译方法和装置、存储介质和电子设备”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

本公开涉及机器翻译领域，尤其涉及一种视频翻译方法和装置、存储介质和电子设备。

背景技术

在机器翻译发达的如今，简单的翻译任务已经可以通过机器翻译加人工
10 校对的方式进行，也就是采用机器译后编辑(machine translation post-editting, MTPE)的模式进行。但现在的 MTPE 技术下，人工修改和机器翻译不能很好地兼容，导致翻译的质量仍旧不及人工翻译，这些问题严重制约了 MTPE 的发展，对翻译结果的准确性要求较高的翻译场景下，通常还是会使用人工翻译的方式进行翻译，但人工翻译效率较低、速度较慢的问题仍旧无法解决。
15 并且，当需要对视频进行翻译时，需要人工对视频的内容逐句进行听取和翻译，这种翻译方式效率低下，无法满足当前全球化背景下的大量的视频翻译需求。

发明内容

提供该发明内容部分以便以简要的形式介绍构思，这些构思将在后面的
20 具体实施方式部分被详细描述。该发明内容部分并不旨在标识要求保护的技术方案的关键特征或必要特征，也不旨在用于限制所要求的保护的技术方案的范围。

第一方面，本公开提供一种视频翻译方法，包括：

将待翻译的视频的语音转换为文本；

25 展示所述文本和所述文本的第一时间信息、第二时间信息及参考译文，所述第一时间信息为所述文本在所述视频中的起始时间，所述第二时间信息为所述文本在所述视频中的结束时间；

响应于用户对所述文本或所述参考译文的操作，展示编辑区域，所述编辑区域支持所述用户输入译文；

30 跟随所述用户在所述编辑区域的输入，提供来自所述参考译文的译文建议；

其中，在检测到所述用户针对所述译文建议的确认操作的情况下，将所述译文建议作为译文结果显示到所述编辑区域；

35 以及，在检测到所述用户针对所述译文建议的非确认操作的情况下，接收所述用户输入的不同于所述译文建议的译文，将所述用户输入的所述译文

-2-

作为所述译文结果显示到所述编辑区域，根据所述用户输入的所述译文更新所述译文区域中的参考译文。

第二方面，本公开提供一种视频翻译装置，包括：

转换模块，用于将待翻译视频的语音转换为文本；

5 展示模块，用于展示所述文本和所述文本的第一时间信息、第二时间信息及参考译文，所述第一时间信息为所述文本在所述视频中的起始时间，所述第二时间信息为所述文本在所述视频中的结束时间；

展示模块，还用于响应于用户对所述文本或所述参考译文的操作，展示编辑区域，所述编辑区域支持所述用户输入译文；

10 建议模块，用于跟随用户在所述编辑区域的输入，提供来自所述参考译文的译文建议；

所述展示模块，还用于在检测到所述用户针对所述译文建议的确认操作的情况下，将所述译文建议作为译文结果显示到所述编辑区域；以及，在检测到所述用户针对所述译文建议的非确认操作的情况下，接收所述用户输入的不同于所述译文建议的译文，将所述用户输入的所述译文作为所述译文结果显示到所述编辑区域，根据所述用户输入的所述译文更新所述译文区域中的参考译文。

第三方面，本公开提供一种计算机可读介质，其上存储有计算机程序，该程序被处理装置执行时实现本公开第一方面中所述方法的步骤。

20 第四方面，本公开提供一种电子设备，包括：

存储装置，其上存储有计算机程序；

处理装置，用于执行所述存储装置中的所述计算机程序，以实现本公开第一方面中所述方法的步骤。

25 基于上述的技术方案，至少可以达到以下技术效果：可以将待翻译视频的语音转换为文本，并提供该文本的第一时间信息、第二时间信息及参考译文，并跟随用户在编辑区域的输入提供来自参考译文的译文建议，并响应于用户的确认操作将译文建议作为译文结果，从而可以节省用户的输入时间，结合人工的准确性及机器的高效性，进而提升视频翻译的效率以及视频翻译的质量。

30 本公开的其他特征和优点将在随后的具体实施方式部分予以详细说明。

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。

35 在附图中：

图 1 是根据一示例性公开实施例示出的一种视频翻译方法的流程图。

图 2 是根据一示例性公开实施例示出的一种译文界面的示意图。

图 3 是根据一示例性公开实施例示出的一种文本拆分方式的示意图。

图 4 是根据一示例性公开实施例示出的一种视频翻译装置的框图。

图 5 是根据一示例性公开实施例示出的一种电子设备的框图。

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 是根据一示例性公开实施例示出的一种视频翻译方法的流程图，本方法可以用于终端、服务器及其他独立电子设备，也可以应用于翻译系统，此种情况下，方法中的各个步骤可以由翻译系统中的多个设备配合完成，例如图 1 中所示的 S12 和 S14 可以由终端执行，S11 和 S13 可以由服务器执行等。如图 1 所示，所述视频翻译方法包括以下步骤：

S11、将待翻译的视频的语音转换为文本。

可以提取待翻译视频中的语音内容，如音轨等，并通过语音识别技术，将该语音内容转换为文本内容。值得说明的是，在将语音内容转换为文本内容时，可以根据语音内容中的分句将文本内容分为多句，且每句文本内容可以对应一个提取到该分句的语音内容的时刻信息，将其作为该句文本内容的时间轴信息。

例如，待翻译视频的语音内容被识别为了多句话，其中，第一句为“首

-4-

先介绍一下什么是热点”，该句话位于视频的第2秒至第5秒之间，则该句文本内容对应的时间轴信息为“00:00:02-00:00:05”，第二局为“从ppt的右边可以看到”，该句话位于视频的第5秒至第7秒之间，则该句文本内容对应的时间轴信息为“00:00:05-00:00:07”。

5 在将待翻译视频的语音内容转换为文本内容时，可以根据所述文本在所述视频中对应的时刻信息和/或画面帧对所述文本进行分句，得到所述多个分句文本，例如，将每连续多秒内的语音的识别文本作为一个分句，或者将连续多个画面帧内出现的语音的识别文本作为一个分句；还可以根据语音内容中的停顿进行分句，例如，可以设置一个停顿阈值，在该停顿阈值内没有识别到人声内容时，可以在没有识别到人声内容的任意位置进行分句；还可以根据语音内容的语义进行分句，在分句词前后进行分句，例如，可以设置将完整的“主语+谓语+宾语”结构的“宾语”作为分句词，对语音内容进行分句，还可以将时间助词、停顿词等作为分句词，在这些词前后进行分句。具体的，可以通过分句模型对识别到的文本内容进行分句，得到分句后的文本内容。

10 S12、展示所述文本和所述文本的第一时间信息、第二时间信息及参考译文。

其中，所述第一时间信息为所述文本在所述视频中的起始时间，所述第二时间信息为所述文本在所述视频中的结束时间。

20 其中，该文本可以是已经分句后的文本，该第一时间信息为分句后的文本的当前分句在视频中的起始时间，该第二时间信息为分句后的文本的当前分句在视频中的结束时间。对每一分句可以展示所述分句文本和所述分句文本的第一时间信息、第二时间信息及参考译文。

S13、响应于用户对所述文本或所述参考译文的操作，展示编辑区域。

25 在用户选中了文本或参考译文对应的区域后，可以在该文本的参考译文上方展示编辑区域，编辑区域支持用户输入译文，用户可以在编辑区域中进行编辑操作，以获得该文本的翻译结果。其中，该编辑区域可以显示于参考译文上方，以便用户对照修改。

30 其中，该文本可以是已经分句后的文本，每一分句文本在不同的区域内展示，针对每一分句文本展示该分句文本的第一时间信息、第二时间信息及参考译文。

35 在一种可能的实施方式中，所述文本包括多个分句文本，每一所述分句文本在不同区域展示，还可以展示提供所述用户对所述分句文本进行拆分的拆分功能栏，并响应于所述用户针对任一所述分句文本的拆分操作，将所述分句文本拆分成至少两句分句子文本，并针对每一所述分句子文本，关联显示所述分句子文本、所述分句子文本的第一时间信息、第二时间信息以及所述分句子文本的参考译文。可选地，该拆分功能栏可以是响应于用户对所述

分句文本或参考译文的操作所提供的，在用户选中该分句文本或参考文本之前可以隐藏该拆分功能栏。

例如，对一段文本内容“今天要为大家介绍的就是我国即将崛起的三座城市”而言，该段文本内容的时间轴信息为“00:00:15-00:00:18”，其中第一
5 时间信息为 00:00:15、第二时间信息为 00:00:18，用户将其分为了两个子句“今天要为大家介绍的就是”和“我国即将崛起的三座城市”，则可以根据编辑前的文本的长度和编辑后的各个子句的文本的长度，为各个子句设置时间轴，例如，可以将原有的时间轴拆分为长度相同的两段，将第一个子段“今天要为大家介绍的就是”的时间轴设置为“00:00:15-00:00:16”，将第二个子段“我
10 国即将崛起的三座城市”的时间轴信息为“00:00:17-00:00:18”。

还可以根据文本内容的字数，为该段文本内容的每个字分配时间轴，并在进行分句后为分句后的子段分配其对应的字数的时间轴。

如图 3 示的是一种可能的文本拆分方式的示意图，如图所示，用户可以通过光标选择需要分句的位置，并点击分句按键；分句前的文本会拆分为两
15 个子句，并按照顺序显示，各个子句的第一时间信息和第二时间信息均是由分句前的第一时间信息和第二时间信息拆分得到。图 3 中，拆分前的虚线框中的一段文本内容，被拆分成了虚线框中的两个子段。

在一种可能的实施方式中，所述文本包括多个分句文本，每一所述分句文本在不同区域展示，还可以展示提供所述用户对所述分句进行合并的合并
20 功能栏，并响应于所述用户针对任意相邻两个分句文本的合并操作，将所述相邻两个分句文本合并成一段新的分句文本，并针对所述新的分句文本，关联显示所述新的分句文本、所述新的分句文本的第一时间信息、第二时间信息以及所述新的分句文本的参考译文。可选地，该合并功能栏可以是响应于用户对所述分句文本或参考译文的操作所提供的，在用户选中该分句文本或
25 参考文本之前可以隐藏该合并功能栏。

在一种可能的实施方式中，所述文本包括多个分句文本，每一所述分句文本在不同区域展示，还可以展示提供所述用户对所述分句文本进行播放的
30 播放功能栏，并响应于所述用户针对所述播放功能栏的操作，播放所述分句文本对应的语音。可选地，该播放功能栏可以是响应于用户对所述分句文本或参考译文的操作所提供的，在用户选中该分句文本或参考文本之前可以隐藏该播放功能栏。并且，在一种可能的实施方式中，可以以参考译文或者译文结果作为字幕播放该分句文本对应的视频，以使用户查看译文字幕的效果。

如图 2 所示的是一种可能的译文界面的示意图，其中，虚线框内部为一段已被用户选中的文本内容的译文界面，图 2 中一共有三段不同的文本内容，
35 其中，被用户选中的文本内容会展示编辑区域和播放、合并、拆分功能栏。待翻译视频的文本内容显示在参考译文上方，且不同的分句有不同的展示区域，各个展示区域可以独立翻译，不因其他区域的修改而更新。用户可以在

-6-

编辑区域输入字符，或修改待翻译的文本的字符。译文界面还可以包括时间轴信息，包括表征起始时刻的第一时间信息和表征结束时刻的第二时间信息。在图 2 中，参考译文为灰色字，译文建议为黑色字，如图 2 所示，在用户选中一段文本内容后，参考译文可以下移一行，与功能栏同行，而原本的参考译文所在的区域成为编辑区域，用于展示译文建议并接收用户的修改。

S14、跟随用户在所述编辑区域的输入，提供来自所述参考译文的译文建议。

其中，基于该译文建议，本公开实施例提供的方法包括在检测到用户针对该译文建议的确认操作的情况下，将该译文建议作为译文结果显示到该编辑区域，以及在检测到用户针对该译文建议的非确认操作的情况下，接收用户输入的不同于该译文建议的译文，并根据用户输入的译文更新该译文区域中的参考译文。

在具体实施时，上述确认操作可以是用户针对预设的快捷键的操作，例如，用户通过点击该快捷键表明采纳译文建议的意图，因此，可以在检测到用户点击快捷键的操作的情况下，将该译文建议作为译文结果显示到该编辑区域。值得说明的是，将译文建议作为译文结果显示到该编辑区域这一动作将作为步骤 S14 中所述的用户在该编辑区域内的输入，也就是说，在此种情况下，步骤 S14 即是表明，本公开实施例提供的方法可以响应于将本次提供的译文建议作为译文结果显示到该编辑区域，提供来自该参考译文的下一个译文建议（该下一个译文建议可以是已提供的译文建议在参考译文中的后续译文）。

可选地，上述检测到用户针对该译文建议的非确认操作的情况可以是检测到用户输入的译文与本次提供的译文建议不一致的情况，在此种情况下，本公开实施例提供的方法可以接收用户输入的不同于该译文建议的译文，并根据用户输入的译文更新该译文区域中的参考译文。同样值得说明的是，用户输入的与译文建议不同的译文，将作为步骤 S14 中所述的用户在该编辑区域内的输入，也就是说，在此种情况下，步骤 S14 即是表明，本公开实施例提供的方法可以响应于用户在该编辑区域输入与译文建议不同的译文，提供来自根据用户输入的译文更新后的参考译文的下一个译文建议。例如，本次提供的译文建议为“my”，在检测到用户输入的译文为不同与译文建议“my”的译文“I”的情况下，根据译文“I”更新参考译文，并从更新后的参考译文中提供译文“I”的下一个译文建议。

采用上述方法，可以根据用户的输入提供来自参考译文的译文建议，并且，用户可以通过确认操作直接将译文建议作为译文结果，减少用户的输入时间，本公开结合人工的准确性及机器的高效性，可以提升翻译的效率以及翻译的质量。

为了使本领域技术人员更加理解本公开实施例提供的技术方案，下面对

本公开实施例提供的文档翻译方法进行详细说明。

可选地，步骤 S14 中的所述提供译文建议可以包括：在所述译文区域突出显示来自所述参考译文的所述译文建议。在此种情况下，在检测到用户针对该译文建议的确认操作的情况下，可以取消所述译文建议在所述译文区域的突出显示。该突出显示可以是加粗字体、高亮字体、异色字、异色背景、底纹效果等可以将译文建议突出展示的显示方式。

在一种可能的实施方式中，该突出显示可以是与已输入译文的显示方式不同的显示方式，例如，已输入的译文可以是加粗字体，译文建议是正常字体，或者，已输入的译文可以是黑色字，译文建议是灰色字等。在检测到用户针对译文建议的确认操作的情况下，可以将译文建议的显示方式调整为与已输入译文的显示方式相同。例如，已输入的译文可以是加粗字体，译文建议是正常字体，在检测到用户的确认操作的情况下，将该译文建议调整为加粗字体显示。

在一种可能的实施方式中，该确认操作可以是用户对电子设备的快捷键的输入操作，例如，该电子设备可以是手机，该快捷键可以是手机显示区域上的虚拟键或手机的实体键（例如：音量键），用户可以对上述的快捷键进行操作以采纳该译文建议，则在检测到用户对上述快捷键的输入操作的情况下，可以将译文建议作为疑问结果显示到编辑区域；该电子设备还可以是电脑，该快捷键可以是电脑键盘或鼠标上的指定或自定义按键（例如：键盘 alt 键、鼠标侧键等）。

该确认操作还可以是由摄像头获取后识别得到的姿势确认操作，如点头、眨眼、作出预设手势等；还可以是由麦克风获取后识别得到的语音操作。

在一种可能的实施方式中，该来自所述参考译文的译文建议包括词、词组、句子中的至少一项。

下面对译文建议的提供方式进行详细阐述：

在用户针对文本内容进行翻译时，可以参考在译文区域中显示的参考译文，在编辑区域内进行输入（值得说明的是，此处的输入包括字符的输入，例如键入字母、单词等，也包括按键操作输入，例如点击编辑区域等），可以提供来自参考译文的译文建议。

其中，该译文建议可以是针对分句的整句的译文建议，也可以是逐词、逐短语提供的更细粒度的译文建议。

例如，文本为“有些城市凭借着完善的高铁网络这一优势不断崛起”，参考译文为“Some cities continue to rise with the advantage of the perfect high-speed rail network”，则在用户点击编辑区域，或者在编辑区域输入了字符“S”之后，可以提供来自参考译文的译文建议“Some”（或者“Some cities continue to rise”等更粗粒度的译文建议）。

用户可以通过确认操作采纳该译文建议，并且，将该确认操作作为在编

—8—

辑区域内的输入操作，继续提供来自参考译文的译文建议，例如，在检测到用户针对“Some”的确认操作的情况下，将“Some”作为译文结果显示到编辑区域，并为用户提供下一译文建议“cities”。

在检测到用户针对译文建议的非确认操作的情况下，接收用户输入的不同于译文建议的译文，并根据用户输入的译文更新译文区域中的参考译文。其中，该非确认操作可以是进行了预设的代表非确认的操作（点击预设按钮、做出预设动作等），也可以是指除了前述的确认操作以外的其他情况，例如，在预设时间内没有进行确认操作，或者进行了继续输入的操作。

例如，文本内容“有些城市凭借着完善的高铁网络这一优势不断崛起”的参考译文为“Some cities continue to rise with the advantage of the perfect high-speed rail network.”，在接收到用户对编辑区域的点击输入操作后，提供来自参考译文的译文建议“Some”，并基于用户的确认操作，将译文建议“Some”作为译文结果显示到编辑区域，并继续为用户提供下一译文建议“cities”。在提供译文建议“with”时，接收到了用户不同于译文建议的输入“b”，则可以基于用户输入的译文，更新参考译文为“Some cities continue to rise because of the advantage of the perfect high-speed rail network.”，并为用户提供译文建议“because”。

值得说明的是，当译文建议为该分句的整句参考译文时，用户可以直接在编辑区域中对译文建议进行编辑，例如，在译文建议中插入单词、删除译文建议中的单词、更改译文建议中的单词等。

例如，文本内容“有些城市凭借着完善的高铁网络这一优势不断崛起”的译文建议与参考译文相同，为“Some cities continue to rise with the advantage of the perfect high-speed rail network.”，用户可以直接在译文建议中将“with”修改为“because of”，并根据用户的修改将参考译文更新为“ome cities continue to rise because of the advantage of the perfect high-speed rail network.”，并向用户提供来自该参考译文的译文建议，用户可以通过确认操作将该译文建议作为译文结果。

其中，参考译文以及译文建议可以由机器翻译（例如深度学习翻译模型等）提供。值得说明的是，当基于所述用户在编辑区域输入的译文无法生成符合文本内容的参考译文时，可以基于预存的词典内容对用户输入的译文字符进行纠错，并根据纠错后的译文更新该参考译文。

值得说明的是，本公开中虽然以翻译语言为英文、原文为中文的案例进行举例，但是，本公开不对翻译的语言以及原文的语言进行限制，本公开中的原文还可以为中文文言文、译文可以为中文白话文，或者原文为日文、译文为英文等各种组合。

在一种可能的实施方式中，原文显示区域为可编辑区域，响应于用户对所述原文显示区域中的文本内容进行修改的操作，可以更新所述译文区域内

的参考译文。

在用户在译文区域输入译文之前或之后，用户都可以对文本的内容，即翻译原文进行编辑，并且，已经输入的译文不会因为原文的修改而被覆盖，而是会根据用户修改后的文本内容及输入的译文字符更新翻译结果。

- 5 例如，编辑前的文本内容为“有些城市凭借着完善的高通网路这一优势不断崛起”，对应的译文建议为“Some cities continue to rise with the advantage of a perfect Qualcomm network.”，用户在编辑区域输入的译文结果为“Some cities continue to rise b”，其中，不同于译文建议的译文为“b”，则可以更新参考译文为“Some cities continue to rise because of the advantage of a perfect
10 Qualcomm network.”。但是，该句文本内容可能是由于杂音、语音讲述人的口音等因素导致的误识别文本，用户发现其原本的文本应当是“有些城市凭借着完善的高铁网络这一优势不断崛起。”，则用户可以将文本内容中的“高通网路”编辑为“高铁网络”，则更新后的参考译文变为“Some cities continue to rise because of the advantage of the perfect high-speed rail network.”，并为用
15 户提供来自更新后的参考译文的译文建议。

在一种可能的实施方式中，在编辑后的文本内容的长度大于编辑前的文本内容的长度，根据编辑前的文本内容的时间轴信息，通过插值处理得到编辑后的文本内容的时间轴信息。

- 20 例如，编辑前的文本内容为“今天要为大家介绍的就是我国的三座城市”，编辑后的文本内容为“今天要为大家介绍的就是我国即将崛起的三座城市”，则编辑后的文本内容中，各个文字的时间轴信息都会被重置为原来的 9/11，并且在后续用户进行分句、合并等操作时，基于各个文字的时间轴信息确定分句或合并后的子段的时间轴信息。

- 25 在一种可能的实施方式中，可以基于所述第一时间信息和第二时间信息，将所述翻译结果作为字幕添加至所述待翻译视频的画面帧中。

- 例如，待翻译视频的第一句翻译结果的时间轴为“00:00:00-00:00:02”（第一时间信息为 00:00:00，第二时间信息为 00:00:02），第二段翻译结果的时间轴为“00:00:03-00:00:07”（第一时间信息为 00:00:03，第二时间信息为 00:00:07），则可以在待翻译视频的第 0 秒至第 2 秒间，插入时间轴为
30 “00:00:00-00:00:02”的翻译结果，在待翻译视频的第 3 秒至第 7 秒间，插入时间轴为“00:00:03-00:00:07”的翻译结果，该翻译结果可以以字幕的形式插入待翻译视频中。

在所有的翻译结果均插入待翻译视频后，可以将翻译完成的视频以用户指定的格式生成，并提供给用户进行下载。

- 35 基于上述的技术方案，至少可以达到以下技术效果：可以将待翻译视频的语音转换为文本，并提供该文本的第一时间信息、第二时间信息及参考译文，并跟随用户在编辑区域的输入提供来自参考译文的译文建议，并响应于

用户的确认操作将译文建议作为译文结果，从而可以节省用户的输入时间，结合人工的准确性及机器的高效性，进而提升视频翻译的效率以及视频翻译的质量。

5 图 4 是根据一示例性公开实施例示出的一种视频翻译装置的框图。如图 4 所示，所述视频翻译装置 400 包括：

转换模块 410，用于将待翻译的视频的语音转换为文本。

10 展示模块 420，用于展示所述文本和所述文本的第一时间信息、第二时间信息及参考译文，所述第一时间信息为所述文本在所述视频中的起始时间，所述第二时间信息为所述文本在所述视频中的结束时间。

展示模块 420，还用于响应于用户对所述文本或所述参考译文的操作，展示编辑区域，所述编辑区域支持所述用户输入译文。

建议模块 430，用于跟随用户在所述编辑区域的输入，提供来自所述参考译文的译文建议；

15 所述展示模块 420，还用于在检测到所述用户针对所述译文建议的确认操作的情况下，将所述译文建议作为译文结果显示到所述编辑区域；以及，在检测到所述用户针对所述译文建议的非确认操作的情况下，接收所述用户输入的不同于所述译文建议的译文，将所述用户输入的所述译文作为所述译文结果显示到所述编辑区域，根据用户输入的所述译文更新所述译文区域中的参考译文。可选地，所述展示模块 420 还用于根据所述文本在所述视频中
20 对应的时刻信息和/或画面帧对所述文本进行分句，得到所述多个分句文本；针对每一所述分句文本，展示所述分句文本和所述分句文本的第一时间信息、第二时间信息及参考译文。

25 可选地，所述文本包括多个分句文本，每一所述分句文本在不同区域展示，所述装置还包括拆分模块，用于展示拆分功能栏，所述拆分功能栏支持所述用户对所述分句文本进行拆分；响应于所述用户针对任一所述分句文本的拆分操作，将所述分句文本拆分成至少两句分句子文本，并针对每一所述分句子文本，关联显示所述分句子文本、所述分句子文本的第一时间信息、第二时间信息以及所述分句子文本的参考译文。

30 可选地，所述文本包括多个分句文本，每一所述分句文本在不同区域展示，所述装置还包括合并模块，用于展示合并功能栏，所述合并功能栏支持所述用户对所述分句文本进行合并；响应于所述用户针对任意相邻两个分句文本的合并操作，将所述相邻两个分句文本合并成一段新的分句文本，并针对所述新的分句文本，关联显示所述新的分句文本、所述新的分句文本的第一时间信息、第二时间信息以及所述新的分句文本的参考译文。
35

可选地，所述文本包括多个分句文本，每一所述分句文本在不同区域展示，所述装置还包括播放模块，用于展示播放功能栏，所述播放功能栏支持

-11-

所述用户对所述分句文本对应的语音进行播放；响应于所述用户针对所述播放功能栏的操作，播放所述分句文本对应的语音。

可选地，所述建议模块 430 用于在所述编辑区域以不同于已输入译文的显示方式来显示所述译文建议；所述响应于所述用户对所述译文建议的确认操作，将所述译文建议作为译文结果显示到所述编辑区域包括：响应于所述用户对所述译文建议的确认操作，以与已输入译文的显示方式相同的方式在所述编辑区域内显示作为译文结果的所述译文建议。

可选地，所述建议模块 430，还用于响应于用户对快捷键的触发操作，将所述译文建议作为译文结果显示到所述编辑区域。

上述各模块的功能在上一实施例中的方法步骤中已详细阐述，在此不做赘述。

基于上述的技术方案，至少可以达到以下技术效果：可以将待翻译视频的语音转换为文本，并提供该文本的第一时间信息、第二时间信息及参考译文，并跟随用户在编辑区域的输入提供来自参考译文的译文建议，并响应于用户的确认操作将译文建议作为译文结果，从而可以节省用户的输入时间，结合人工的准确性及机器的高效性，进而提升视频翻译的效率以及视频翻译的质量。

下面参考图 5，其示出了适于用来实现本公开实施例的电子设备（例如图 1 中的终端设备或服务器）500 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、个人数字助理（personal digital assistant, PDA）、平板电脑（portable android device, PAD）、便携式多媒体播放器（portable media player, 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电视（digital television, DTV）、台式计算机等等的固定终端。图 5 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 5 所示，电子设备 500 可以包括处理装置（例如中央处理器、图形处理器等）501，其可以根据存储在只读存储器（read-only memory, ROM）502 中的程序或者从存储装置 508 加载到随机访问存储器（random access memory, RAM）503 中的程序而执行各种适当的动作和处理。在 RAM 503 中，还存储有电子设备 500 操作所需的各种程序和数据。处理装置 501、ROM 502 以及 RAM 503 通过总线 504 彼此相连。输入/输出（I/O）接口 505 也连接至总线 504。

通常，以下装置可以连接至 I/O 接口 505：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 506；包括例如液晶显示器（liquid crystal display, LCD）、扬声器、振动器等等的输出装置 507；包括例如磁带、硬盘等的存储装置 508；以及通信装置 509。通信装置

509 可以允许电子设备 500 与其他设备进行无线或有线通信以交换数据。虽然图 5 示出了具有各种装置的电子设备 500，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
5 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 509 从网络上被下载和安装，或者从存储装置 508 被安装，或者从 ROM 502 被安装。在该计算机程序被处理装置 501 执行时，执行本公开实
10 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、
15 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（erasable programmable read-only memory，EPROM）或闪存、光纤、便携式紧凑磁盘只读存储器（compact disc read-only memory，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
20 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
25 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任何适当的介质传输，包括但不限于：电线、光缆、射频（radio frequency，RF）等等，或者上述的任意合适的组合。

30 在一些实施方式中，客户端、服务器可以利用诸如超文本传输协议（hypertext transfer protocol，HTTP）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（local area network，LAN），广域网（wide area network，WAN），网际网（例如，互联网）以及端对端网络
35 （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：获取至少两个网际协议地址；向节点评价设备发送包括所述至少两个网际协议地址的节点评价请求，其中，所述节点评价设备从所述至少两个网际协议地址中，选取网际协议地址并返回；接收所述节点评价设备返回的网际协议地址；其中，所获取的网际协议地址指示内容分发网络中的边缘节点。

或者，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：接收包括至少两个网际协议地址的节点评价请求；从所述至少两个网际协议地址中，选取网际协议地址；返回选取出的网际协议地址；其中，接收到的网际协议地址指示内容分发网络中的边缘节点。

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）——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模块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模块本身的限定，例如，第一获取模块还可以被描述为“获取至少两个网际协议地址的模块”。

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（field programmable gate array，FPGA）、专用集成电路（application

specific integrated circuit , ASIC)、专用标准产品(application specific standard parts , ASSP)、片上系统(system on a chip , SOC)、复杂可编程逻辑设备(complex programmable logic device , CPLD)等等。

5 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。

15 根据本公开的一个或多个实施例,示例 1 提供了一种视频翻译方法,包括将待翻译的视频的语音转换为文本;展示所述文本和所述文本的第一时间信息、第二时间信息及参考译文,所述第一时间信息为所述文本在所述视频中的起始时间,所述第二时间信息为所述文本在所述视频中的结束时间;响应于用户对所述文本或所述参考译文的操作,展示编辑区域,所述编辑区域支持所述用户输入译文;跟随所述用户在所述编辑区域的输入,提供来自所述参考译文的译文建议;其中,在检测到所述用户针对所述译文建议的确认操作的情况下,将所述译文建议作为译文结果显示到所述编辑区域;以及,在检测到所述用户针对所述译文建议的非确认操作的情况下,接收所述用户输入的不同于所述译文建议的译文,将所述用户输入的所述译文作为所述译文结果显示到所述编辑区域,根据所述用户输入的所述译文更新所述译文区域中的参考译文。根据本公开的一个或多个实施例,示例 2 提供了示例 1 的方法,根据所述文本在所述视频中对应的时刻信息和/或画面帧对所述文本进行分句,得到所述多个分句文本;针对每一所述分句文本,展示所述分句文本和所述分句文本的第一时间信息、第二时间信息及参考译文。

25 根据本公开的一个或多个实施例,示例 3 提供了示例 1 的方法,所述文本包括多个分句文本,每一所述分句文本在不同区域展示,所述方法还包括:展示拆分功能栏,所述拆分功能栏支持所述用户对所述分句文本进行拆分;响应于所述用户针对任一所述分句文本的拆分操作,将所述分句文本拆分成至少两句分句子文本,并针对每一所述分句子文本,关联显示所述分句子文本、所述分句子文本的第一时间信息、第二时间信息以及所述分句子文本的参考译文。

30 根据本公开的一个或多个实施例,示例 4 提供了示例 1 的方法,所述文本包括多个分句文本,每一所述分句文本在不同区域展示,所述方法还包括:

展示合并功能栏，所述合并功能栏支持所述用户对所述分句文本进行合并；响应于所述用户针对任意相邻两个分句文本的合并操作，将所述相邻两个分句文本合并成一段新的分句文本，并针对所述新的分句文本，关联显示所述新的分句文本、所述新的分句文本的第一时间信息、第二时间信息以及所述新的分句文本的参考译文。

根据本公开的一个或多个实施例，示例 5 提供了示例 1-4 的方法，所述文本包括多个分句文本，每一所述分句文本在不同区域展示，所述方法还包括：展示播放功能栏，所述播放功能栏支持所述用户对所述分句文本对应的语音进行播放；响应于所述用户针对所述播放功能栏的操作，播放所述分句文本对应的语音。

根据本公开的一个或多个实施例，示例 6 提供了示例 1-4 的方法，所述提供来自所述参考译文的译文建议包括：在所述编辑区域以不同于已输入译文的显示方式来显示所述译文建议；所述响应于所述用户对所述译文建议的确认操作，将所述译文建议作为译文结果显示到所述编辑区域包括：响应于所述用户对所述译文建议的确认操作，以与已输入译文的显示方式相同的方式来在所述编辑区域内显示作为译文结果的所述译文建议。

根据本公开的一个或多个实施例，示例 7 提供了示例 1-4 的方法，所述响应于所述用户对所述译文建议的确认操作，将所述译文建议作为译文结果显示到所述编辑区域，包括：响应于用户对快捷键的输入操作，将所述译文建议作为译文结果显示到所述编辑区域。

根据本公开的一个或多个实施例，示例 8 提供了一种视频翻译装置，转换模块，用于将待翻译视频的语音转换为文本；展示模块，用于展示所述文本和所述文本的第一时间信息、第二时间信息及参考译文，所述第一时间信息为所述文本在所述视频中的起始时间，所述第二时间信息为所述文本在所述视频中的结束时间；展示模块，还用于响应于用户对所述文本或所述参考译文的操作，展示编辑区域，所述编辑区域支持所述用户输入译文；建议模块，用于跟随用户在所述编辑区域的输入，提供来自所述参考译文的译文建议；所述展示模块，还用于在检测到所述用户针对所述译文建议的确认操作的情况下，将所述译文建议作为译文结果显示到所述编辑区域；以及，在检测到所述用户针对所述译文建议的非确认操作的情况下，接收所述用户输入的不同于所述译文建议的译文，将所述用户输入的所述译文作为所述译文结果显示到所述编辑区域，根据所述用户输入的所述译文更新所述译文区域中的参考译文。

根据本公开的一个或多个实施例，示例 9 提供了示例 8 的装置，所述展示模块还用于根据所述文本在所述视频中对应的时刻信息和/或画面帧对所述文本进行分句，得到所述多个分句文本；针对每一所述分句文本，展示所述分句文本和所述分句文本的第一时间信息、第二时间信息及参考译文。

根据本公开的一个或多个实施例，示例 10 提供了示例 8 的装置，所述文本包括多个分句文本，每一所述分句文本在不同区域展示，所述装置还包括拆分模块，用于展示拆分功能栏，所述拆分功能栏支持所述用户对所述分句文本进行拆分；响应于所述用户针对任一所述分句文本的拆分操作，将所述分句文本拆分成至少两句分句子文本，并针对每一所述分句子文本，关联显示所述分句子文本、所述分句子文本的第一时间信息、第二时间信息以及所述分句子文本的参考译文。

根据本公开的一个或多个实施例，示例 11 提供了示例 8 的装置，所述文本包括多个分句文本，每一所述分句文本在不同区域展示，所述装置还包括合并模块，用于展示合并功能栏，所述合并功能栏支持所述用户对所述分句文本进行合并；响应于所述用户针对任意相邻两个分句文本的合并操作，将所述相邻两个分句文本合并成一段新的分句文本，并针对所述新的分句文本，关联显示所述新的分句文本、所述新的分句文本的第一时间信息、第二时间信息以及所述新的分句文本的参考译文。

根据本公开的一个或多个实施例，示例 12 提供了示例 8-11 的装置，所述文本包括多个分句文本，每一所述分句文本在不同区域展示，所述装置还包括播放模块，用于展示播放功能栏，所述播放功能栏支持所述用户对所述分句文本对应的语音进行播放；响应于所述用户针对播放功能栏的操作，播放所述分句文本对应的语音。

根据本公开的一个或多个实施例，示例 13 提供了示例 8-11 的装置，所述建议模块用于在所述编辑区域以不同于已输入译文的显示方式来显示所述译文建议；所述响应于所述用户对所述译文建议的确认操作，将所述译文建议作为译文结果显示到所述编辑区域包括：响应于所述用户对所述译文建议的确认操作，以与已输入译文的显示方式相同的方式来在所述编辑区域内显示作为译文结果的所述译文建议。

根据本公开的一个或多个实施例，示例 14 提供了示例 8-11 的装置，所述建议模块，还用于响应于用户对快捷键的触发操作，将所述译文建议作为译文结果显示到所述编辑区域

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，但是这些不应当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。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。

权 利 要 求

1、一种视频翻译方法，其特征在于，所述方法包括：

将待翻译的视频的语音转换为文本；

展示所述文本和所述文本的第一时间信息、第二时间信息及参考译文，

5 所述第一时间信息为所述文本在所述视频中的起始时间，所述第二时间信息为所述文本在所述视频中的结束时间；

响应于用户对所述文本或所述参考译文的操作，展示编辑区域，所述编辑区域支持所述用户输入译文；

10 跟随所述用户在所述编辑区域的输入，提供来自所述参考译文的译文建议；

其中，在检测到所述用户针对所述译文建议的确认操作的情况下，将所述译文建议作为译文结果显示到所述编辑区域；以及，在检测到所述用户针对所述译文建议的非确认操作的情况下，接收所述用户输入的不同于所述译文建议的译文，将所述用户输入的所述译文作为所述译文结果显示到所述编辑区域，根据所述用户输入的所述译文更新所述译文区域中的参考译文。

15 2、根据权利要求1所述的方法，其特征在于，所述展示所述文本和所述文本的第一时间信息、第二时间信息及参考译文，包括：

根据所述文本在所述视频中对应的时刻信息和/或画面帧对所述文本进行分句，得到所述多个分句文本；

20 针对每一所述分句文本，展示所述分句文本和所述分句文本的第一时间信息、第二时间信息及参考译文。

3、根据权利要求1所述的方法，其特征在于，所述文本包括多个分句文本，每一所述分句文本在不同区域展示，所述方法还包括：

25 展示拆分功能栏，所述拆分功能栏支持所述用户对所述分句文本进行拆分；

响应于所述用户针对任一所述分句文本的拆分操作，将所述分句文本拆分成至少两句分句子文本，并针对每一所述分句子文本，关联显示所述分句子文本、所述分句子文本的第一时间信息、第二时间信息以及所述分句子文本的参考译文。

30 4、根据权利要求1所述的方法，其特征在于，所述文本包括多个分句文本，每一所述分句文本在不同区域展示，所述方法还包括：

展示合并功能栏，所述合并功能栏支持所述用户对所述分句文本进行合并；

35 响应于所述用户针对任意相邻两个分句文本的合并操作，将所述相邻两个分句文本合并成一段新的分句文本，并针对所述新的分句文本，关联显示所述新的分句文本、所述新的分句文本的第一时间信息、第二时间信息以及所述新的分句文本的参考译文。

5、根据权利要求1-4任一项所述的方法，其特征在于，所述文本包括多个分句文本，每一所述分句文本在不同区域展示，所述方法还包括：

展示播放功能栏，所述播放功能栏支持所述用户对所述分句文本对应的语音进行播放；

5 响应于所述用户针对所述播放功能栏的操作，播放所述分句文本对应的语音。

6、根据权利要求1-4任一项所述的方法，其特征在于，所述提供来自所述参考译文的译文建议包括：

在所述编辑区域以不同于已输入译文的显示方式来显示所述译文建议；

10 所述响应于所述用户对所述译文建议的确认操作，将所述译文建议作为译文结果显示到所述编辑区域包括：响应于所述用户对所述译文建议的确认操作，以与已输入译文的显示方式相同的方式来在所述编辑区域内显示作为译文结果的所述译文建议。

15 7、根据权利要求1-4任一项所述的方法，其特征在于，所述响应于所述用户对所述译文建议的确认操作，将所述译文建议作为译文结果显示到所述编辑区域，包括：

响应于用户对快捷键的触发操作，将所述译文建议作为译文结果显示到所述编辑区域。

8、一种视频翻译装置，其特征在于，所述装置包括：

20 转换模块，用于将待翻译视频的语音转换为文本；

展示模块，用于展示所述文本和所述文本的第一时间信息、第二时间信息及参考译文，所述第一时间信息为所述文本在所述视频中的起始时间，所述第二时间信息为所述文本在所述视频中的结束时间；

25 展示模块，还用于响应于用户对所述文本或所述参考译文的操作，展示编辑区域，所述编辑区域支持所述用户输入译文；

建议模块，用于跟随用户在所述编辑区域的输入，提供来自所述参考译文的译文建议；

30 所述展示模块，还用于在检测到所述用户针对所述译文建议的确认操作的情况下，将所述译文建议作为译文结果显示到所述编辑区域；以及，在检测到所述用户针对所述译文建议的非确认操作的情况下，接收所述用户输入的不同于所述译文建议的译文，将所述用户输入的所述译文作为所述译文结果显示到所述编辑区域，根据所述用户输入的所述译文更新所述译文区域中的参考译文。

35 9、一种计算机可读介质，其上存储有计算机程序，其特征在于，该程序被处理装置执行时实现权利要求1-7中任一项所述方法的步骤。

10、一种电子设备，其特征在于，包括：

存储装置，其上存储有计算机程序；

—20—

处理装置，用于执行所述存储装置中的所述计算机程序，以实现权利要求 1-7 中任一项所述方法的步骤。

— 1/4 —

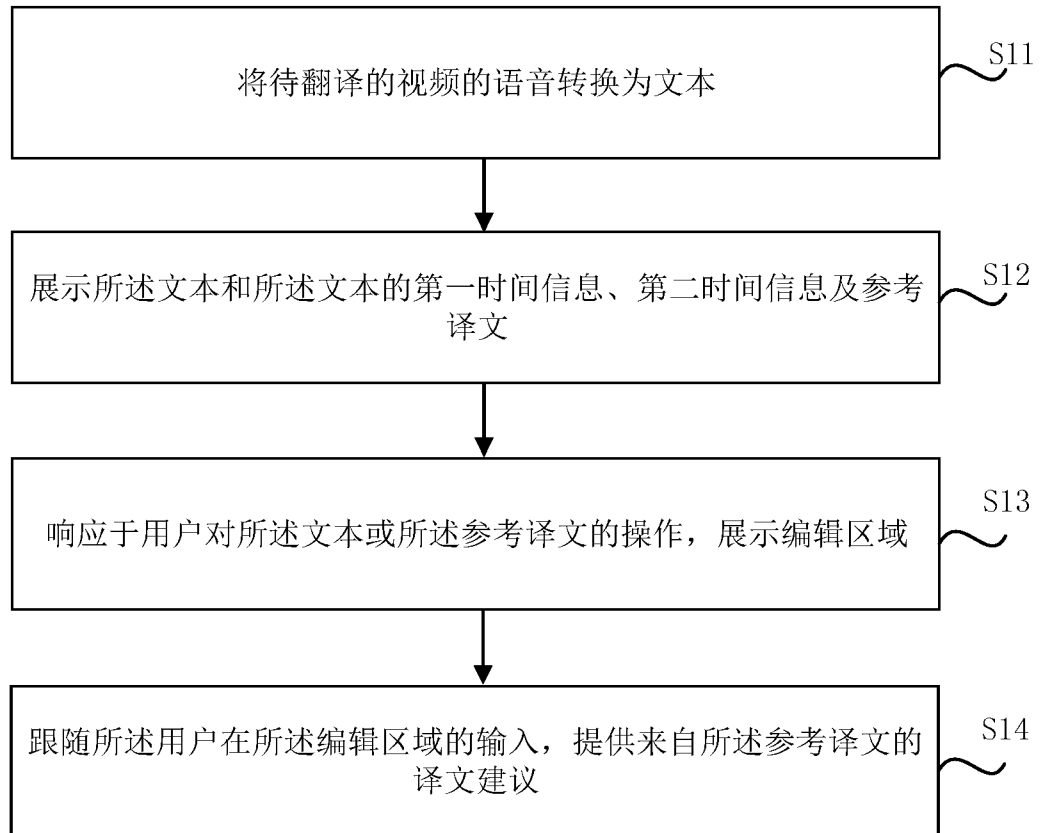


图 1

— 2/4 —



图 2

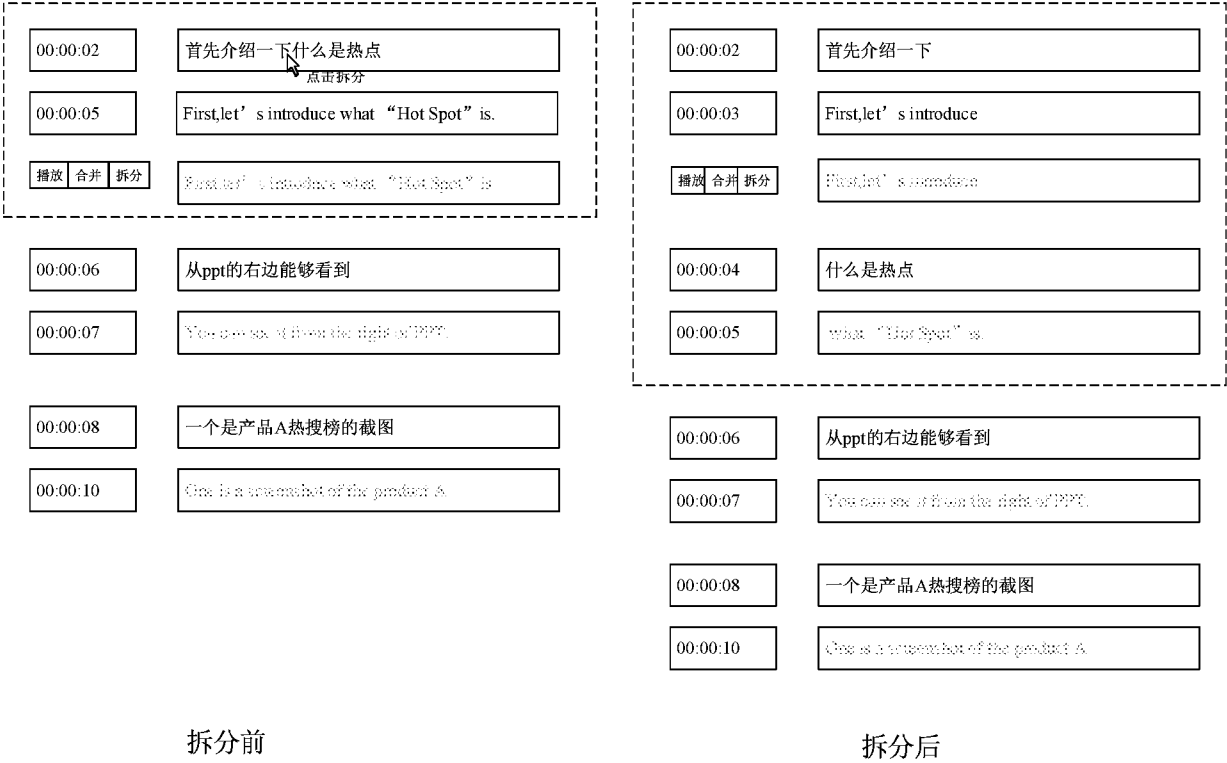


图 3

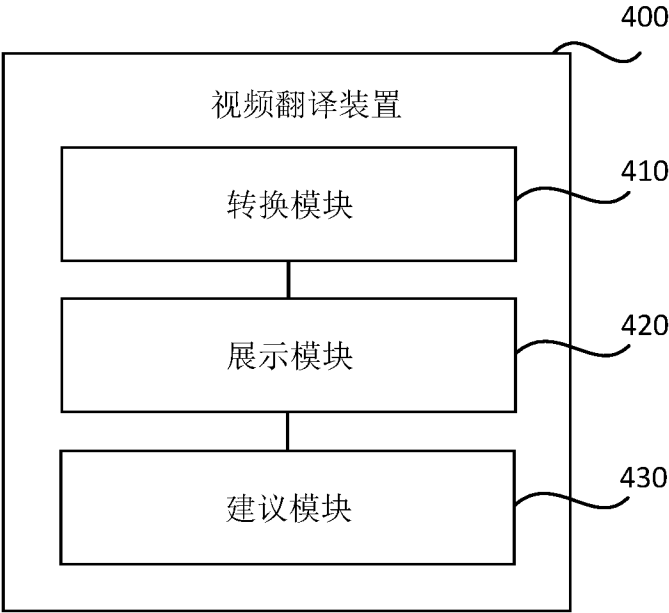


图 4

500

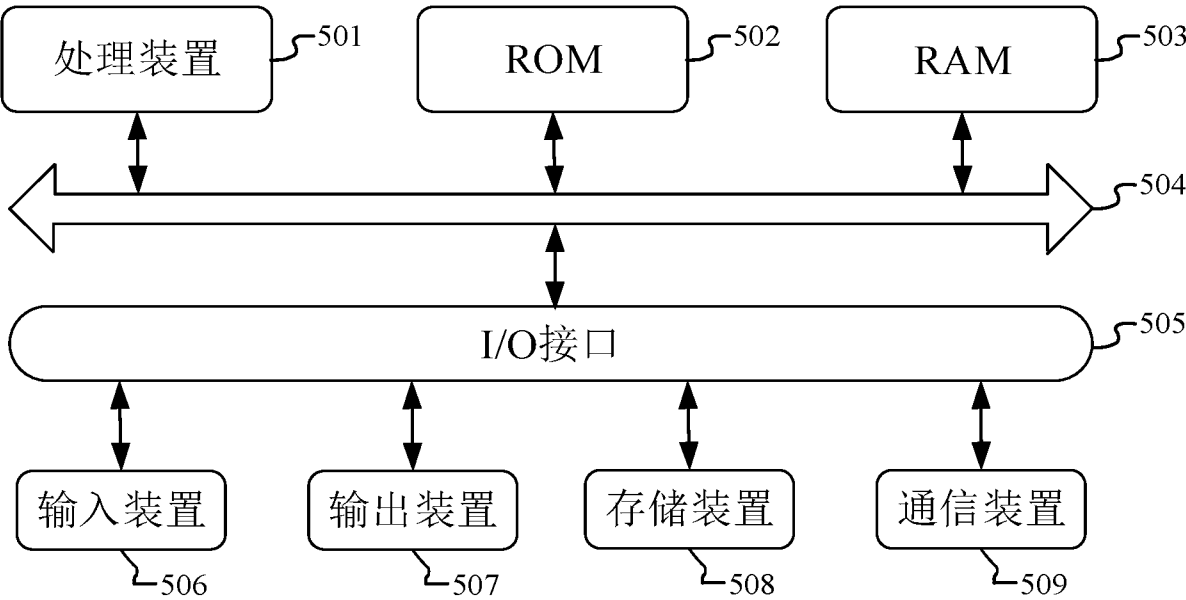


图 5

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2021/101388

A. CLASSIFICATION OF SUBJECT MATTER

G06F 40/58(2020.01)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)

G06F 40/-

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, CNKI, SIPOABS, DWPI: 视频, 音频, 字幕, 翻译, 文本, 人工, 时间, 开始, 起始, 结束, 编辑, 建议, 候选, video, audio, subtitle, translate, text, artificial, time, start, begin, end, finish, edit, recommendation

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 111753558 A (BEIJING BYTEDANCE NETWORK TECHNOLOGY CO., LTD.) 09 October 2020 (2020-10-09) claims 1-10, description paragraphs 0034-0082	1-10
Y	CN 107885729 A (SHENYANG AEROSPACE UNIVERSITY) 06 April 2018 (2018-04-06) description, paragraphs 0033-0073, and figures 1-2	1-10
Y	CN 105828101 A (BEIJING XIAOMI MOBILE SOFTWARE CO., LTD.) 03 August 2016 (2016-08-03) description, paragraphs 0004-0251, and figures 1-7	1-10
A	US 2015042771 A1 (United Video Properties, Inc.) 12 February 2015 (2015-02-12) entire document	1-10

☐ 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

08 September 2021

Date of mailing of the international search report

24 September 2021

Name and mailing address of the ISA/CN

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

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2021/101388

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	111753558	A	09 October 2020	None			
CN	107885729	A	06 April 2018	CN	107885729	B	11 May 2021
CN	105828101	A	03 August 2016	CN	105828101	B	08 March 2019
US	2015042771	A1	12 February 2015	None			

国际检索报告

国际申请号

PCT/CN2021/101388

A. 主题的分类 G06F 40/58 (2020.01) 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) G06F 40/- 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, CNKI, SIPOABS, DWPI: 视频, 音频, 字幕, 翻译, 文本, 人工, 时间, 开始, 起始, 结束, 编辑, 建议, 候选, video, audio, subtitle, translate, text, artificial, time, start, begin, end, finish, edit, recommendation																	
C. 相关文件 <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 111753558 A (北京字节跳动网络技术有限公司) 2020年 10月 9日 (2020 - 10 - 09) 权利要求1-10, 说明书第0034-0082段</td> <td>1-10</td> </tr> <tr> <td>Y</td> <td>CN 107885729 A (沈阳航空航天大学) 2018年 4月 6日 (2018 - 04 - 06) 说明书第0033-0073段, 图1-2</td> <td>1-10</td> </tr> <tr> <td>Y</td> <td>CN 105828101 A (北京小米移动软件有限公司) 2016年 8月 3日 (2016 - 08 - 03) 说明书第0004-0251段, 图1-7</td> <td>1-10</td> </tr> <tr> <td>A</td> <td>US 2015042771 A1 (United Video Properties, Inc.) 2015年 2月 12日 (2015 - 02 - 12) 全文</td> <td>1-10</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 111753558 A (北京字节跳动网络技术有限公司) 2020年 10月 9日 (2020 - 10 - 09) 权利要求1-10, 说明书第0034-0082段	1-10	Y	CN 107885729 A (沈阳航空航天大学) 2018年 4月 6日 (2018 - 04 - 06) 说明书第0033-0073段, 图1-2	1-10	Y	CN 105828101 A (北京小米移动软件有限公司) 2016年 8月 3日 (2016 - 08 - 03) 说明书第0004-0251段, 图1-7	1-10	A	US 2015042771 A1 (United Video Properties, Inc.) 2015年 2月 12日 (2015 - 02 - 12) 全文	1-10
类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求															
PX	CN 111753558 A (北京字节跳动网络技术有限公司) 2020年 10月 9日 (2020 - 10 - 09) 权利要求1-10, 说明书第0034-0082段	1-10															
Y	CN 107885729 A (沈阳航空航天大学) 2018年 4月 6日 (2018 - 04 - 06) 说明书第0033-0073段, 图1-2	1-10															
Y	CN 105828101 A (北京小米移动软件有限公司) 2016年 8月 3日 (2016 - 08 - 03) 说明书第0004-0251段, 图1-7	1-10															
A	US 2015042771 A1 (United Video Properties, Inc.) 2015年 2月 12日 (2015 - 02 - 12) 全文	1-10															
☐ 其余文件在C栏的续页中列出。 ☒ 见同族专利附件。																	
<table border="0"> <tr> <td> * 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 </td> <td> “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 </td> </tr> </table>			* 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													
* 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 2021年 9月 8日		国际检索报告邮寄日期 2021年 9月 24日															
ISA/CN的名称和邮寄地址 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 传真号 (86-10)62019451		授权官员 姜楠 电话号码 86-(010)-62411027															

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2021/101388

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	111753558	A	2020年 10月 9日	无	
CN	107885729	A	2018年 4月 6日	CN	107885729 B 2021年 5月 11日
CN	105828101	A	2016年 8月 3日	CN	105828101 B 2019年 3月 8日
US	2015042771	A1	2015年 2月 12日	无	