

# (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16日(16.12.2021)



(10) 国际公布号  
**WO 2021/249323 A1**

(51) 国际专利分类号:  
*H04N 21/439* (2011.01) *H04N 21/488* (2011.01)  
*H04N 21/458* (2011.01) *G10L 15/26* (2006.01)  
*H04N 21/485* (2011.01)

(21) 国际申请号: PCT/CN2021/098533

(22) 国际申请日: 2021年6月7日(07.06.2021)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:  
202010518278.3 2020年6月9日(09.06.2020) 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

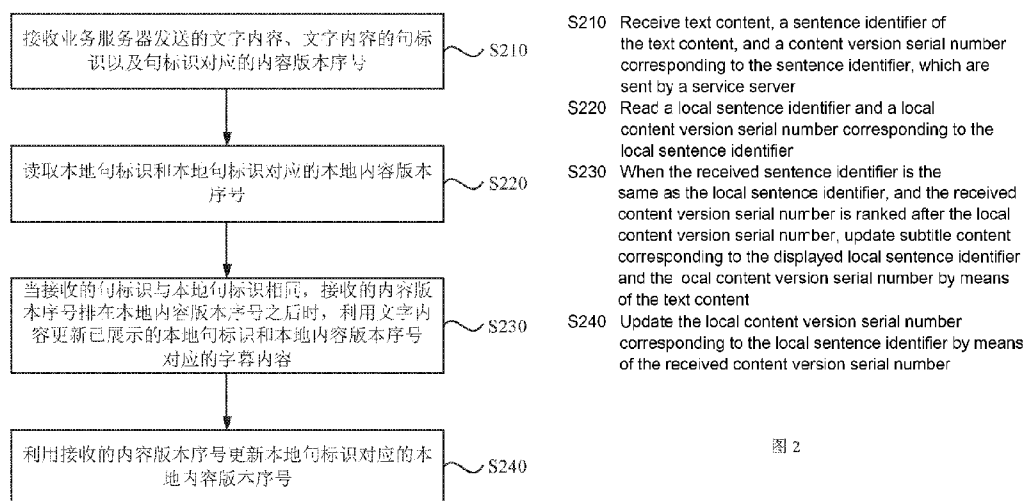
石景山区实兴大街30号院3号楼2层B-0035  
房间, Beijing 100041 (CN)。

(72) 发明人: 赵立(ZHAO, Li); 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)。 韩晓(HAN, Xiao); 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)。 陈可蓉(CHEN, Kojung); 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)。 童剑(TONG, Jian); 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: 上海光栅知识产权代理有限公司 (LINKER IP LLC); 中国上海市长宁区延安西路889号702室, Shanghai 201109 (CN)。

(54) Title: INFORMATION PROCESSING METHOD, SYSTEM AND APPARATUS, AND ELECTRONIC DEVICE AND STORAGE MEDIUM

(54) 发明名称: 一种信息处理方法、系统、装置、电子设备及存储介质



(57) Abstract: Disclosed are an information processing method, system and apparatus, and an electronic device and a storage medium. The method is applied to a client, and comprises: receiving text content and a sentence identifier of the text content, which are sent by a service server; and reading a local sentence identifier, and when the received sentence identifier is the same as the local sentence identifier, updating subtitle content corresponding to the displayed local sentence identifier by means of the text content. When a received sentence identifier is the same as a local sentence identifier, a client replaces displayed subtitle content with the received text content, such that during the process of converting audio data of the same sentence into text, a subtitle can be displayed when a first intermediate conversion result of the sentence is received, and when an intermediate conversion result after the sentence is received, a dynamic update is performed on subtitle content corresponding to the sentence. Not only is dynamic subtitle display of each sentence realized, but the delay time of subtitle display of each sentence is also reduced, and the user experience is also improved.



WO 2021/249323 A1

(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: 本公开实施例公开了一种信息处理方法、系统、装置、电子设备及存储介质, 其中该方法应用于客户端, 包括: 接收业务服务器发送的文字内容及文字内容的句标识; 读取本地句标识, 并于接收的句标识与本地句标识相同时, 利用文字内容更新已展示的本地句标识对应的字幕内容。客户端在接收的句标识与本地句标识相同时, 利用接收的文字内容替换已展示的字幕内容, 能够在同一句话的音频数据转化为文字的过程中, 于接收到该句话的首个中间转化结果时就进行字幕显示, 并在接收到该句话之后的中间转化结果时, 对该句话对应的字幕内容进行动态更新, 不仅实现了对每句话进行动态字幕展示, 还减少了每句话的字幕显示的延迟时间, 提升了用户体验。

# 一种信息处理方法、系统、装置、电子设备及存储介质

## 相关申请的交叉引用

5 本申请要求于 2020 年 6 月 9 日提交的申请号为 202010518278.3、名称为“一种信息处理方法、系统、装置、电子设备及存储介质”的中国专利申请的优先权，该申请的全文通过引用结合在本申请中。

## 技术领域

10 本公开实施例涉及计算机技术领域，尤其涉及一种信息处理方法、系统、装置、电子设备及存储介质。

## 背景技术

15 随着互联网以及通信技术的不断发展，通过通信类应用进行信息沟通已成为用户进行信息交流的重要方式之一。当客户端间进行包含音频数据的通信时，服务器可将音频数据转写为文字并下发至对应客户端，以使客户端显示音频数据对应的字幕。

目前，服务器通常将每句话的音频数据转写完毕后，再将转写完毕的文字下发至客户端，以使客户端看到每句话对应字幕的延迟时间较长。并且，当某句话较长时，字幕的延迟时间会更久，这将会大大降低用户的交互体验。

## 20 发明内容

本公开实施例提供了一种信息处理方法、装置、电子设备及存储介质，能够实现动态字幕展示，降低字幕延迟，提升用户体验。

第一方面，本公开实施例提供了一种信息处理方法，应用于客户端，包括：

接收业务服务器发送的文字内容及所述文字内容的句标识；

25 读取本地句标识，并于接收的句标识与所述本地句标识相同时，利用所述文字内容更新已展示的所述本地句标识对应的字幕内容。

第二方面，本公开实施例还提供了一种信息处理方法，包括：

接收来自业务服务器的第一文字内容及所述第一文字内容的第一句标识；

展示所述第一文字内容；

30 接收来自所述业务服务器的第二文字内容及所述第二文字内容的第二句标识；

当所述第二句标识与所述第一句标识相同时，展示所述第二文字内容以更新所述第一文字内容。

第三方面，本公开实施例提供了一种信息处理系统，包括：

业务服务器，用于响应于客户端发送的开启字幕请求表征对音频数据进行拉取，获取所述音频数据中当前待转化音频数据在文字转化过程中的文字内容，以及根据所述当前待转化音频数据的变化更新所述文字内容的句标识；

所述客户端，用于接收所述业务服务器发送的所述文字内容及所述文字内容的句标识；  
5 读取本地句标识，并于接收的句标识与所述本地句标识相同时，利用所述文字内容更新已展示的所述本地句标识对应的字幕内容。

第四方面，本公开实施例还提供了一种信息处理装置，应用于客户端，包括：

接收模块，用于接收业务服务器发送的文字内容及所述文字内容的句标识；

展示模块，用于读取本地句标识，并于接收的句标识与所述本地句标识相同时，利用所  
10 述文字内容更新已展示的所述本地句标识对应的字幕内容。

第五方面，本公开实施例还提供了一种电子设备，所述电子设备包括：

一个或多个处理器；

存储装置，用于存储一个或多个程序，

当所述一个或多个程序被所述一个或多个处理器执行，使得所述一个或多个处理器实现  
15 如本公开实施例任一所述的信息处理方法。

第六方面，本公开实施例还提供了一种包含计算机可执行指令的存储介质，所述计算机可执行指令在由计算机处理器执行时用于执行如本公开实施例任一所述的信息处理方法。

第七方面，本公开实施例还提供了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，该计算机程序被处理器执行时，使得所述处理器执行如本公开实施例任一  
20 所述的信息处理方法。

第八方面，本公开实施例还提供了一种计算机程序，当其在电子设备上运行时，使得电子设备执行如本公开实施例任一所述的信息处理方法。

本公开实施例的技术方案，客户端接收业务服务器发送的文字内容及文字内容的句标识；  
25 读取本地句标识，并于接收的句标识与本地句标识相同时，利用文字内容更新已展示的本地句标识对应的字幕内容。客户端在接收的句标识与本地句标识相同时，可利用接收的文字内容替换已展示的字幕内容，能够实现在同一句话的音频数据转化为文字的过程中，于接收到该句话的首个中间转化结果时就进行字幕显示，并在接收到该句话之后的中间转化结果时，对该句话对应的字幕内容进行动态更新，不仅实现了对每句话进行动态字幕展示，还减少了每句话的字幕显示的延迟时间，提升了用户体验。

30

## 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 图 1 为本公开实施例一所提供的一种信息处理方法的流程示意图；

图 2 为本公开实施例二所提供的一种信息处理方法的流程示意图；

图 3 为本公开实施例三所提供的一种信息处理方法的流程示意图；

图 4 为本公开实施例四所提供的一种信息处理系统的结构示意图；

图 5 为本公开实施例六所提供的一种信息处理系统的结构示意图；

图 6 为本公开实施例七提供的一种信息处理装置结构示意图；

图 7 为本公开实施例八所提供的一种电子设备结构示意图。

## 5 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。

10 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。

15 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。

20 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”。

### 实施例一

25 图 1 为本公开实施例一所提供的一种信息处理方法流程示意图，本公开实施例尤其适用于客户端在实时互动场景下进行字幕展示的情形，该方法可以由信息处理装置来执行，该装置可以通过软件和/或硬件的形式实现，该装置可配置于电子设备中，例如配置于智能手机、平板电脑、笔记本电脑或台式计算机等电子设备中。

如图 1 所示，本实施例提供的信息处理方法，应用于客户端，包括：

S110、接收业务服务器发送的文字内容及文字内容的句标识。

30 本公开实施例中，业务服务器可以认为是为各客户端提供字幕展示业务功能的服务器。任何与其他客户端进行实时互动通信的客户端，皆可在互动通信过程中向业务服务器发送开启字幕请求，以请求业务服务器反馈该客户端以及与该客户端进行互动通信的其他客户端上传的音频数据所转写得到的文字内容及文字内容的句标识。进而，客户端可根据接收的文字内容及文字内容的句标识进行字幕展示。

35 本实施例中，业务服务器可以根据该客户端发送的开启字幕请求表征从流媒体服务器对音频数据进行拉取，并将音频数据发送至实时语音转写工具；实时语音转写工具基于自动语音识别（Automatic Speech Recognition, ASR）技术，可根据音频数据的参数（例如音频数据之间的时间间隔等参数）对音频数据进行断句，以确定每句话对应的音频数据，还可将当前待转化的一句话对应的音频数据作为当前待转化音频数据，对当前待转化音频数据进行可输出中间转化结果的文字转化操作；业务服务器从实时语音转写工具获取当前待转化音频数

据所转化的文字内容，且获取的文字内容包括当前待转化音频数据转化过程中的文字内容和转化完毕的文字内容。

其中，每个客户端发送的开启字幕请求中可携带有表征该客户端的身份标识或表征该客户端所在会议室、直播间等实时互动房间的标识，业务服务器可根据上述表征从流媒体服务器拉取该客户端以及与该客户端进行互动通信的其他客户端上传的音频数据。其中，实时语音转写工具可以是由业务服务器本地提供的，或者是由与业务服务器建立通信的第三方所提供的，通过实时语音转写工具能够对音频数据流进行自动语音识别处理，实现音频数据流断句和针对每句话对应的音频数据进行转文字操作。其中，中间转化结果可认为是在转化得到完整的音频数据对应的文字内容前，转化得到的部分的音频数据对应的文字内容。例如，针对“今天天气真不错”这句话的音频数据，它的中间转化结果可以是“今天”、“今天天气”、“今天天气真”等一系列转化得到的部分的音频数据对应的文字内容，而转化得到完整的音频数据对应的文字内容为“今天天气真不错”。

本实施例中，业务服务器在获取到文字内容之后，还可以为该文字内容设置句标识，该句标识可唯一标识具体实时互动场景下每句话对应的音频数据，且可以由字母、文字、数字和特殊符号等至少一种元素组成。业务服务器设置文字内容对应的句标识的过程，可以是：根据当前待转化音频数据的变化更新文字内容的句标识。业务服务器可将文字内容的句标识随文字内容同时下发至该客户端。

其中，业务服务器设置句标识的一种实施方式，例如可以是：首先，于当前待转化音频数据在文字转化完毕时，设置句标识的更新标记；接着，在获取到新的当前待转化音频数据在文字转化过程中的文字内容时，根据更新标记更新句标识，并于句标识更新完毕时取消更新标记。例如，第一句话的音频数据对应的句标识为 Seg1，第一句话的音频数据在文字转化完毕时，设置了句标识更新标记。紧接着获取到第二句话的音频数据的首个中间转化结果对应的文字内容，则此时可根据更新标识将句标识更新为 Seg2，并且取消更新标记，避免第二句话的音频数据的下一个中间转化结果对应的文字内容的句标识被更新。也就是说，在获取到新的当前待转化音频数据在文字转化过程中之后转化的文字内容时，无需更新该句标识。

此外，当获取到新的当前待转化音频数据对应的文字内容也为转化完毕的文字内容时，可根据更新标记更新句标识，并保留该更新标记。仍以上述场景为例，若第二句话很短，获取到的第二句话的音频数据的首个文字内容就已经是转化完毕后的文字内容了，则此时除了根据更新标识将句标识更新为 Seg2 之外，需要保留更新标记，以使获取到再下一句话对应的文字内容时，再次根据更新标记更新句标识。

其中，更新句标识例如可以是生成一个该具体实时互动场景下与历史句标识都不同的新的句标识；且当句标识包括有序编号时，例如包括 a、b、c...；1、2、3...；I、II、III... 等时，也可以按编号顺序更新句标识，例如句标识可依次更新为 S1、S2、S3... 等。通过使句标识中包括有序编号，可利于客户端依据句标识的编号顺序进行字幕展示，便于用户了解每句话正确的上下文内容，提高用户阅读字幕的体验。

上述更新句标识的实施方式旨在实现“同一句话对应的音频数据转化得到的文字内容的句标识相同，不同句话对应的音频数据转化得到的文字内容的句标识不同”，且能够实现上述相同目的的其他实施方式也可应用于此，在此不做具体限定。

S120、读取本地句标识，并于接收的句标识与本地句标识相同时，利用文字内容更新已

展示的本地句标识对应的字幕内容。

本公开实施例中，客户端可将首次接收的句标识存储在本地第一预设存储空间中；当第二次接收到的句标识时，可将从该第一预设存储空间读取的全部句标识作为本地句标识；当第二次接收的句标识和本地句标识中任一句标识相同时，本地句标识不进行更新；当第二次接收的句标识和本地句标识中各句标识皆不同时，可以将第二次接收的句标识添加在第一预设存储空间中，以更新本地句标识。客户端在之后每次接收到句标识时，可重复上述操作，以实现更新或不更新本地句标识操作。

其中，接收的句标识与本地句标识相同，可以认为是客户端本地句标识中存在与接收的句标识相同的句标识，即客户端之前接收的相同的句标识文字内容与本次接收的文字内容对应同一句话的音频数据。其中，利用文字内容更新已展示的本地句标识对应的字幕内容，可以是客户端利用接收的文字内容更新已展示的与接收的句标识相同的本地句标识对应的字幕内容，具体可以是将相同的本地句标识对应的字幕的显示控件中的文字内容更新为接收的文字内容，以实现字幕内容的动态更新。

此外，当客户端接收的句标识与本地句标识不同时，可以认为客户端之前接收的文字内容为与本次接收的文字内容为不同句话的音频数据转化的文字内容。此时，客户端可以将文字内容作为新的字幕内容进行展示，并利用接收的句标识更新本地句标识。在这里，客户端在将文字内容作为新的字幕内容进行展示时，可以采用任何可以与其它已接收到的文字内容的展示方式不同的展示方式来展示。具体地，在一些应用场景中，将文字内容作为新的字幕内容进行展示，例如可以是生成一个新的字幕的显示控件，并在该新的显示控件中导入接收的文字内容，以实现一条新字幕的展示。其中，在新的字幕展示时，之前的字幕可滑动至客户端界面的其他位置（例如，之前的字幕依次向上滚动）。并且，新字幕中的字幕内容也可利用客户端之后接收到的与该字幕内容对应相同句标识的文字内容进行更新。其中，客户端利用接收的句标识更新本地句标识，例如可以是将接收的句标识添加至本地第一预设存储空间里中。

其中，利用文字内容更新已展示的本地句标识对应的字幕内容，可以是确定接收的文字内容与已展示与接收的句标识相同的本地句标识对应的字幕内容的区别内容，并利用区别内容更新字幕内容。由于音频数据的流式处理模式，文字转化过程中首先转化的内容可能出现错误，但可以根据之后转化的内容向前纠错。示例性的，针对“工作室通透明亮”这句话的音频数据，客户端首次接收到的文字内容可能是“工作时”，则在字幕展示时，可以展示为“工作时...”；客户端之后接收到的内容可能为“工作室通透”，则在字幕内容更新时，可首先确定接收的文字内容与已展示的相同的本地句标识对应的字幕内容的区别内容包括“室”与“时”不同，且增加了“通透”，进而可根据区别内容将展示的字幕内容中的“时”更改为“室”，并在之后新增“通透”，以实现字幕内容的更新。

可选的，利用文字内容更新已展示的本地句标识对应的字幕内容，包括：利用文字内容替换已展示的本地句标识对应的字幕内容。也就是说，客户端可以利用接收的文字内容替换已展示与接收的句标识相同的本地句标识对应的字幕内容。示例性的，针对“今天天气真不错”这句话的音频数据，客户端若首先接收到的文字内容为“今天”，则在字幕展示时，可以展示为“今天...”；之后接收到的文字内容为“今天天气”，则在字幕展示时，可以将展示内容替换为“今天天气...”。通过利用文字内容替换已展示的本地句标识对应的字幕内容，可

减少客户端的运算处理过程，实现更加快速地字幕内容更新。

由于字幕内容可动态更新，在针对某一句话的音频数据进行对应文字内容展示时，无需等到该句话的音频数据转化完毕后，再对转化完毕的文字内容进行展示，而是可以在接收到该句话首次的中间转化结果对应的文字内容时，就将文字内容通过字幕展示，并在每次接收到该句话新的中间转化结果对应的文字内容时，利用新的文字内容进行字幕的动态更新。从而能够减少每一句话从说话人说出到字幕显示之间的延迟时间，很大程度上提升了用户体验。

可选的，客户端为实时互动应用的客户端。在这里，实时互动应用可以包括但不限于即时通讯应用、多媒体会议应用、视频直播应用和群聊互动应用中的至少一种。其中，在两个以上客户端进行即时通讯、多媒体会议、群聊互动或在同一直播间时，皆可通过上述信息处理方法对说话人所说的话进行字幕展示，以提高实时互动效率和互动体验。

本公开实施例的技术方案，客户端接收业务服务器发送的文字内容及文字内容的句标识；读取本地句标识，并于接收的句标识与本地句标识相同时，利用文字内容更新已展示的本地句标识对应的字幕内容。客户端通过在接收的句标识与本地句标识相同时，利用接收的文字内容替换已展示的字幕内容，能够在同一句话的音频数据转化为文字的过程中，于接收到该句话的首个中间转化结果时就进行字幕显示，并在接收到该句话之后的中间转化结果时，对该句话对应的字幕内容进行动态更新，不仅实现了对每句话进行动态字幕展示，还减少了每句话的字幕显示的延迟时间，提升了用户体验。

## 实施例二

图 2 为本公开实施例二所提供的一种信息处理方法的流程示意图。本实施例与上述实施例中提供的各个信息处理方法的可选方案可以结合。本实施例提供的信息处理方法中，客户端在接收业务服务器发送的文字内容及文字内容的句标识的同时，还可接收业务服务器发送的句标识对应的内容版本序号，并且能够利用句标识及其对应的内容版本序号对文字内容进行字幕展示，可有效避免在排序靠前的文字内容因为网络延迟等原因之后发给客户端时，利用排序靠前的文字内容更新排序靠后的文字内容所导致的字幕更新出错的情况。

如图 2 所示，本实施例提供的信息处理方法包括：

S210、接收业务服务器发送的文字内容、文字内容的句标识以及句标识对应的内容版本序号。

本实施例中，客户端接收的内容版本序号，可以唯一标识该文字内容在对应的当前待转化音频数据的文字转化过程中的转化次序，且可以包含有序编号，例如包括 a、b、c...；1、2、3...；I、II、III... 等，或者还包含字母、文字、数字和特殊符号等至少一种其他元素。

业务服务器在为文字内容设置句标识的同时，还可以设置句标识对应的内容版本序号。业务服务器设置内容版本序号的实施方式，例如可以是：根据文字内容在当前待转化音频数据的文字转化过程中的转化次序，生成句标识对应的内容版本序号。可以认为，当前待转化音频数据的转化的文字内容的句标识一致，并且可根据文字转化过程中本次转化的文字内容所在的转化次序生成本次转化的文字内容的内容版本序号。

示例性的，针对“今天天气真不错”这句话的音频数据，它的转化的文字内容可以依次是“今天”、“今天天气”、“今天天气真”和“今天天气真不错”，且上述文字内容对应的句标识一致，但是可以将第一次转化的“今天”的内容版本序号设置为 Slice1，将第二次转化的“今天天气”的内容版本序号设置为 Slice2，将第三次转化的“今天天气真”的内容版本序

号设置为 Slice3，将第四次转化的“今天天气真不错”的内容版本序号设置为 Slice4。

S220、读取本地句标识和本地句标识对应的本地内容版本序号。

本实施例中，客户端可以从本地第一预设存储空间读取的本地句标识，可以从本地第二预设存储空间读取的本地内容版本序号。其中，第一预设存储空间的大小可以满足既存储句标识，也存储内容版本序号时，可以将第一预设存储空间同时作为第二预设存储空间使用。

S230、当接收的句标识与本地句标识相同，接收的内容版本序号排在本地内容版本序号之后时，利用文字内容更新已展示的本地句标识和本地内容版本序号对应的字幕内容。

本实施例中，当客户端本地句标识中存在与接收的句标识相同的句标识，接收的内容版本序号排在相同的本地句标识的本地内容版本序号之后（可不包括等于本地内容版本序号）时，可以认为之前接收的相同的句标识文字内容与本次接收的文字内容对应同一句话的音频数据，并且本次接收的文本内容比上一次接收的文字内容的转化次序靠后。此时，客户端可以利用文字内容更新已展示的相同的本地句标识和本地内容版本序号对应的字幕内容。

此外，当客户端接收的句标识与本地句标识相同，接收的内容版本序号排在本地内容版本序号之前时，可以不对本地句标识和本地内容版本序号对应的字幕内容进行更新。示例性的，假设客户端在接收到内容版本序号为 Slice4 的“今天天气真不错”并进行字幕动态更新展示之后，才收到 Slice3 对应的“今天天气真”，由于 Slice3 排在 Slice4 之前，此时不用“今天天气真”来更新“今天天气真不错”，避免了更新出错的情况。

S240、利用接收的内容版本序号更新本地句标识对应的本地内容版本序号。

其中，客户端针对每个句标识，将接收到的该句标识对应的第一个内容版本序号存储至第二预设存储空间中；当接收到该句标识对应的第二个内容版本序号时，可从第二预设存储空间中读取该句标识对应的全部内容版本序号作为本地内容版本序号；当第二个内容版本序号排在全部本地内容版本序号之后时，可以将第二个内容版本序号添加在第二预设存储空间中，以更新本地内容版本序号；当第二个内容版本序号排在任一本地内容版本序号之前（可包括等于本地内容版本序号）时，本地内容版本序号不更新。客户端在之后每次接收到该句标识对应的内容版本序号时，可重复上述操作，以实现更新或不更新本地内容版本序号操作。

本公开实施例的技术方案，客户端在接收业务服务器发送的文字内容及文字内容的句标识的同时，还可接收句标识对应的内容版本序号，并且能够利用句标识及其对应的内容版本序号对文字内容进行字幕展示。通过在接收的句标识与本地句标识相同时，利用排序靠后的内容版本序号更新排序在前的内容版本序号对应的文字内容，可有效避免在排序靠前的文字内容因为网络延迟等原因之后发给客户端时，利用排序靠前的文字内容更新排序靠后的文字内容所导致的字幕更新出错的情况，进一步提升了字幕展示效果。

### 实施例三

图 3 为本公开实施例三所提供的一种信息处理方法的流程示意图。本实施例与上述实施例中提供的各个信息处理方法的可选方案可以结合。本实施例提供的信息处理方法对接收的文字内容和句标识为至少两组时的字幕动态展示方法进行了优化，能够实现同时对多条字幕进行动态更新，进一步提升了字幕展示效果。

如图 3 所示，本实施例提供的信息处理方法包括：

S310、接收业务服务器发送的文字内容及文字内容的句标识，其中接收的文字内容和句标识为至少两组。

本公开实施例中，业务服务器可根据开启字幕请求表征，从流媒体服务器拉取该客户端以及与该客户端进行互动通信的其他客户端中至少两个客户端同时上传的至少两路音频数据（也就是同时存在两个以上说话人在说话）；并分别获取了至少两路音频数据中当前待转化音频数据在文字转化过程中的文字内容，以及根据至少两路当前待转化音频数据的变化更新文字内容的句标识，可以认为若当前待转化音频数据为至少两路时，则可为至少两路音频数据分别设置句标识，而无需在其中一路转化完毕时设置句标识更新标记，并根据句标识更新标记更新。此时，客户端可接收到业务服务器同时或交替发送的至少两组文字内容和句标识，进而可根据至少两组文字内容和句标识同时对至少两个字幕进行动态更新展示。

5 S320、读取本地句标识，判断接收的至少两个句标识中是否存在与本地句标识相同的句标识。

10 本实施例中，本地句标识的数量可以为至少一个，判断接收的至少两个句标识中是否存在与本地句标识相同的句标识，也就是判断接收的至少两个句标识与本地句标识是否存在交集。

15 S330、若是，则利用相同的句标识对应的文字内容更新已展示的本地句标识对应的字幕内容；将不同的句标识对应的文字内容作为新的字幕内容进行展示。

20 本实施例中，若客户端接收的至少两个句标识与本地句标识中存在相同的句标识，且相同的句标识数量可以为至少一个，则可以认为接收的文字内容中存在对已展示字幕内容进行更新的文字内容，可利用接收的相同句标识对应的文字内容更新已展示的相同的本地句标识对应的字幕内容。同时，可将接收的与本地句标识不相同的句标识对应的文字内容作为新的字幕内容进行展示。从而实现了例如之前说话人未说话完毕时，新的说话人开始说话，所形成的至少两个说话人说话的情况下，进行多条字幕同时更新展示。

S340、若否，则将接收的至少两个文字内容分别作为新的至少两个字幕内容进行展示。

25 本实施例中，若客户端接收的至少两个句标识与本地句标识中不存在相同的句标识，则可以认为接收的文字内容与已展示字幕内容皆不属于同一句话，此时可将接收的至少两个文字内容分别作为新的至少两个字幕内容进行展示。从而实现了例如之前的说话人说话完毕，至少两个新的说话人同时开始说话，所形成的至少两个说话人说话的情况下，进行多条字幕同时更新展示。

S350、将接收的至少两个句标识同时作为新的本地句标识。

30 本实施例中，若接收的至少两个句标识与本地句标识中存在相同的句标识，则可将与本地句标识不相同的句标识添加至第一预设存储空间中，以将接收的至少两个句标识同时作为新的本地句标识。若接收的至少两个句标识中不存在与本地句标识相同的句标识，则可以将接收的至少两个句标识同时添加至第一预设存储空间中，以将接收的至少两个句标识同时作为新的本地句标识。

35 本公开实施例的技术方案，当接收的文字内容和句标识为至少两组时，客户端在读取本地句标识之后，判断接收的至少两个句标识中是否存在与本地句标识相同的句标识；若是，则利用相同的句标识对应的文字内容更新已展示的本地句标识对应的字幕内容；将不同的句标识对应的文字内容作为新的字幕内容进行展示；若否，则将接收的至少两个文字内容分别作为新的至少两个字幕内容进行展示；将接收的至少两个句标识同时作为新的本地句标识。能够实现同时对多条字幕进行动态更新，进一步提升了字幕展示效果。

#### 实施例四

图 4 为本公开实施例四所提供的一种信息处理系统的结构示意图。本实施例提供的信息处理系统通过业务服务器与客户端之间的相互配合，能够实现在实时互动场景下，业务服务器进行互动音频数据的拉取、音频数据转化的文字内容的获取及文字内容对应句标识的生成；客户端根据接收的业务服务器发送的文字内容和句标识进行字幕内容的动态更新展示。本公开实施例所提供的信息处理系统中与本公开任意实施例所提供的信息处理方法属于相同的发明构思，未公开的技术细节可参考上述实施例，且可达到与上述实施例相同的有益效果。

如图 4 所示，本实施例提供的信息处理系统，包括：

业务服务器 410，用于响应于客户端 420 发送的开启字幕请求表征对音频数据进行拉取，获取音频数据中当前待转化音频数据在文字转化过程中的文字内容，以及根据当前待转化音频数据的变化更新文字内容的句标识；

客户端 420，用于接收业务服务器 410 发送的文字内容及文字内容的句标识；读取本地句标识，并于接收的句标识与本地句标识相同时，利用文字内容更新已展示的本地句标识对应的字幕内容。

本公开实施例中，客户端发送的开启字幕请求中可携带有表征该客户端的身份标识或表征该客户端所在会议室、直播间等实时互动房间的标识，业务服务器可根据上述表征从流媒体服务器拉取该客户端以及与该客户端进行互动通信的其他客户端上传的音频数据。

其中，业务服务器可将音频数据发送至实时语音转写工具，以使实时语音转写工具基于 ASR 技术，对音频数据进行断句，以确定每句话对应的音频数据，还可将当前待转化的一句话对应的音频数据作为当前待转化音频数据，对当前待转化音频数据进行可输出中间转化结果的文字转化操作；业务服务器从实时语音转写工具获取当前待转化音频数据所转化的文字内容，且获取的文字内容包括当前待转化音频数据转化过程中的文字内容和转化完毕的文字内容。

其中，业务服务器在获取到文字内容之后，还可以为该文字内容设置句标识，该句标识可唯一标识具体实时互动场景下每句话对应的音频数据，且可以由字母、文字、数字和特殊符号等至少一种元素组成。业务服务器设置文字内容对应的句标识的过程，可以是：根据当前待转化音频数据的变化更新文字内容的句标识。业务服务器可将文字内容的句标识随文字内容同时下发至该客户端。

其中，业务服务器设置句标识的一种实施方式，例如可以是：于当前待转化音频数据在文字转化完毕时，设置句标识的更新标记；在获取到新的当前待转化音频数据在文字转化过程中的文字内容时，根据更新标记更新句标识，并于句标识更新完毕时取消更新标记。此外，在获取到新的当前待转化音频数据在文字转化过程之后转化的文字内容时，无需更新该句标识。此外，当获取到新的当前待转化音频数据对应的文字内容也为转化完毕的文字内容时，可根据更新标记更新句标识，并保留该更新标记。

其中，当发送开启字幕请求的客户端为多个时，业务服务器可在接收到每个客户端发送的开启字幕请求之后，针对各客户端，在获取到某一客户端请求的首个文字内容时，可生成该客户端的初始的句标识，并可根据之后获取的该客户端请求的文字内容，设置文字内容的句标识。可以认为，业务服务器可以维护每个客户端和该客户端对应的句标识之间的关系，并可将各客户端对应的句标识发送至对应客户端，以保证业务服务器同时为多个客户端提供

字幕展示业务功能。

其中，客户端可将首次接收的句标识存储在本地第一预设存储空间中；当第二次接收到的句标识时，可将从该第一预设存储空间读取的全部句标识作为本地句标识；当第二次接收的句标识和本地句标识中任一句标识相同时，本地句标识不进行更新；当第二次接收的句标识和本地句标识中各句标识皆不同时，可以将第二次接收的句标识添加在第一预设存储空间中，以更新本地句标识。客户端在之后每次接收到句标识时，可重复上述操作，以实现更新或不更新本地句标识操作。

其中，客户端在接收到业务服务器发送的文字内容和句标识之后，可将从第一预设存储空间读取本地句标识。当客户端接收的句标识和本地句标识相同时，可以认为客户端之前接收的相同的句标识文字内容与本次接收的文字内容对应同一句话的音频数据。此时，客户端可利用接收的文字内容更新已展示的与接收的句标识相同的本地句标识对应的字幕内容。此外，当客户端接收的句标识和本地句标识不同时，可以认为客户端之前接收的文字内容为与本次接收的文字内容为不同句话的音频数据转化的文字内容，此时客户端可以将接收的文字内容作为新的字幕内容进行展示，并利用接收的句标识更新本地句标识。其中，利用文字内容更新已展示的本地句标识对应的字幕内容，包括：利用文字内容替换已展示的本地句标识对应的字幕内容。

可选的，业务服务器，还用于：根据文字内容在当前待转化音频数据的文字转化过程中的转化次序，生成句标识对应的内容版本序号；

相应的，客户端，还用于：读取本地句标识对应的本地内容版本序号；当接收的句标识与本地句标识相同，接收的内容版本序号排在本地内容版本序号之后时，利用文字内容替换已展示的本地句标识和本地内容版本序号对应的字幕内容；利用接收的内容版本序号更新本地句标识对应的本地内容版本序号。

本实施例中，业务服务器在为文字内容设置句标识的同时，还可以设置句标识对应的内容版本序号。业务服务器设置内容版本序号的实施方式，例如可以是：根据文字内容在当前待转化音频数据的文字转化过程中的转化次序，生成句标识对应的内容版本序号。可以认为，当前待转化音频数据的转化的文字内容的句标识一致，并且可根据文字转化过程中本次转化的文字内容所在的转化次序生成本次转化的文字内容的内容版本序号。

其中，客户端针对每个句标识，将接收到的该句标识对应的第一个内容版本序号存储至第二预设存储空间中；当接收到该句标识对应的第二个内容版本序号时，可从第二预设存储空间中读取该句标识对应的全部内容版本序号作为本地内容版本序号；当第二个内容版本序号排在全部本地内容版本序号之后时，可以将第二个内容版本序号添加在第二预设存储空间中，以更新本地内容版本序号；当第二个内容版本序号排在任一本地内容版本序号之前（可包括等于本地内容版本序号）时，本地内容版本序号不更新。客户端在之后每次接收到该句标识对应的内容版本序号时，可重复上述操作，以实现更新或不更新本地内容版本序号操作。

其中，当客户端本地句标识中存在与接收的句标识相同的句标识，接收的内容版本序号排在相同的本地句标识的本地内容版本序号之后（可不包括等于本地内容版本序号）时，可以认为之前接收的相同的句标识文字内容与本次接收的文字内容对应同一句话的音频数据，并且本次接收的文本内容比上一次接收的文字内容的转化次序靠后。此时，客户端可以利用文字内容更新已展示的相同的本地句标识和本地内容版本序号对应的字幕内容。此外，当客

户端接收的句标识与本地句标识相同，接收的内容版本序号排在本地内容版本序号之前时，可以不对本地句标识和本地内容版本序号对应的字幕内容进行更新。

5 本公开实施例提供的信息处理系统，业务服务器响应于客户端发送的开启字幕请求表征对音频数据进行拉取，获取音频数据中当前待转化音频数据在文字转化过程中的文字内容，以及根据当前待转化音频数据的变化更新文字内容的句标识，以使客户端根据接收的文字内容和句标识进行字幕内容的动态更新展示。业务服务器通过将当前待转化音频数据转化为文字内容过程中得到的不同文字内容皆设置为同一句标识，能够指示客户端针对同一当前待转化音频数据，利用之后接收到的文字内容更新已展示的字幕内容。能够实现业务服务器发送同一句话的音频数据转化的文字内容的过程中，客户端于接收到该句话的首个中间转化结果时就进行字幕显示，并在接收到该句话之后的中间转化结果时，对该句话对应的字幕内容进行动态更新，不仅实现了对每句话进行动态字幕展示，还减少了每句话的字幕显示的延迟时间，提升了用户体验。

15 此外，业务服务器在更新句标识的同时，还可以更新句标识对应的内容版本序号，以使客户端可进一步根据内容版本序号排序靠后的更新排序靠前的文字内容，能够避免在排序靠前的文字内容因为网络延迟等原因之后发给客户端时，利用排序靠前的文字内容更新排序靠后的文字内容导致字幕更新出错的情况。

#### 实施例五

20 本实施例与上述实施例提供的信息处理系统的各个可选方案可以结合。本实施例提供的信息处理系统对文字内容和句标识为至少两组时的字幕动态展示进行了优化，能够实现客户端同时对多条字幕进行动态更新，进一步提升了字幕展示效果。

本实施例提供的信息处理系统中，业务服务器，用于：

25 响应于开启字幕请求表征对至少两路音频数据进行拉取，获取至少两路音频数据中当前待转化音频数据在文字转化过程中的文字内容，以及根据至少两路当前待转化音频数据的变化更新文字内容的句标识；

相应的，客户端，用于：

当接收的文字内容和句标识为至少两组时，读取本地句标识，判断接收的至少两个句标识中是否存在与本地句标识相同的句标识；

若是，则利用相同的句标识对应的文字内容替换已展示的本地句标识对应的字幕内容；将不同的句标识对应的文字内容作为新的字幕内容进行展示；

30 若否，则将接收的至少两个文字内容分别作为新的至少两个字幕内容进行展示；

将接收的至少两个句标识同时作为新的本地句标识。

35 本公开实施例中，业务服务器可根据开启字幕请求表征，从流媒体服务器拉取该客户端以及与该客户端进行互动通信的其他客户端中至少两个客户端同时上传的至少两路音频数据（也就是同时存在两个以上说话人在说话）；并分别获取了至少两路音频数据中当前待转化音频数据在文字转化过程中的文字内容，以及根据至少两路当前待转化音频数据的变化更新了文字内容的句标识，可以认为若当前待转化音频数据为至少两路时，则可为至少两路音频数据分别设置句标识，而无需在其中一路转化完毕时设置句标识更新标记，并根据句标识更新标记更新。此时，客户端可接收到业务服务器同时或交替发送的至少两组文字内容和句标识，进而可根据至少两组文字内容和句标识同时对至少两个字幕进行动态更新展示。

本实施例中，若客户端接收的至少两个句标识与本地句标识中存在相同的句标识，且相同的句标识数量可以为至少一个，则可以认为接收的文字内容中存在对已展示字幕内容进行更新的文字内容，可利用接收的相同句标识对应的文字内容更新已展示的相同的本地句标识对应的字幕内容。同时，可将接收的与本地句标识不相同的句标识对应的文字内容作为新的字幕内容进行展示。从而实现了例如之前说话人未说话完毕时，新的说话人开始说话，所形成的至少两个说话人说话的情况下，进行多条字幕同时更新展示。

若客户端接收的至少两个句标识中不存在与本地句标识相同的句标识，则可以认为接收的文字内容与已展示字幕内容皆不属于同一句话，此时可将接收的至少两个文字内容分别作为新的至少两个字幕内容进行展示。从而实现了例如之前的说话人说话完毕，至少两个新的说话人同时开始说话，所形成的至少两个说话人说话的情况下，进行多条字幕同时更新展示。

本实施例中，若接收的至少两个句标识与本地句标识中存在相同的句标识，则可将与本地句标识不相同的句标识添加至第一预设存储空间中，以将接收的至少两个句标识同时作为新的本地句标识。若接收的至少两个句标识中不存在与本地句标识相同的句标识，则可以将接收的至少两个句标识同时添加至第一预设存储空间中，以将接收的至少两个句标识同时作为新的本地句标识。

本公开实施例提供的信息处理系统，业务服务器响应于客户端发送的开启字幕请求表征可对至少两路音频数据进行拉取，获取至少两路音频数据中当前待转化音频数据在文字转化过程中的文字内容，以及根据至少两路当前待转化音频数据的变化更新文字内容的句标识，以使客户端根据接收的至少两组文字内容和句标识进行多条字幕内容的同时动态更新展示。

#### 实施例六

图 5 为本公开实施例六所提供的一种信息处理系统的结构示意图。本实施例与上述实施例提供的信息处理系统中各个可选方案可以结合。本实施例提供的信息处理系统还可包括：流媒体服务器，用于接收进行流媒体通信的客户端上传的音频数据；相应的，业务服务器，具体用于：响应于客户端发送的开启字幕请求表征，从流媒体服务器拉取音频数据。本实施例提供的信息处理系统还可包括：实时语音转写工具，用于接收业务服务器发送的音频数据，确定音频数据的当前待转化音频数据，并反馈当前待转化音频数据在文字转化过程中的文字内容；相应的，业务服务器，具体用于：接收实时语音转写工具反馈的音频数据中当前待转化音频数据在文字转化过程中的文字内容。通过流媒体服务器能够实现对各客户端上传的音频数据的存储；通过实时语音转写工具能够实现将音频数据实时转化为文字内容并输出中间转化结果。

如图 5 所示，本实施例提供的信息处理系统，包括：

流媒体服务器 510，用于接收进行流媒体通信的客户端 520 上传的音频数据；

业务服务器 530，用于响应于客户端 520 发送的开启字幕请求表征，从流媒体服务器 510 拉取音频数据；

实时语音转写工具 540，用于接收业务服务器 530 发送的音频数据，确定音频数据的当前待转化音频数据，并反馈当前待转化音频数据在文字转化过程中的文字内容；

业务服务器 530，还用于接收实时语音转写工具反馈的音频数据中当前待转化音频数据在文字转化过程中的文字内容，以及根据当前待转化音频数据的变化更新文字内容的句标识；

客户端 520，用于接收业务服务器 530 发送的文字内容及文字内容的句标识；读取本地句

标识，并于接收的句标识与本地句标识相同时，利用文字内容更新已展示的本地句标识对应的字幕内容。

本公开实施例中，任何与其他客户端进行实时互动通信的客户端，皆可在互动通信过程中通过流媒体软件开发工具包（Software Development Kit, SDK）与流媒体服务器建立通信，并可通过建立的通信将包含音频流的多媒体数据传输至流媒体服务器上。业务服务器同样可通过流媒体 SDK 从流媒体服务器拉取各开启字幕请求表征的音频流；业务服务器可与业务服务器本地提供的，或者是由与业务服务器建立通信的第三方所提供的实时语音转写工具交互，实现音频数据的发送，以及文字内容的获取；业务服务器还可与客户端建立长链接，以向客户端发送文字内容及文字内容的句标识。

10 可选的，开启字幕请求中携带有请求字幕语种；相应的，业务服务器，还用于：根据请求字幕语种对获取的文字内容进行翻译，并将翻译完毕的文字内容发送至客户端。

其中，客户端向服务器发送的开启字幕请求中可携带有请求字幕语种，也就是客户端期望看到的字幕语种。业务服务器在发送文字内容前，可根据对应客户端请求的字幕语种对文字内容进行翻译。有利于实现例如用户与母语不同的其他用户进行实时交互的情境下，各用户之间通过字幕展示促进交互效率，提高了用户体验。

可选的，开启字幕请求中携带有第一身份标识；相应的，业务服务器，具体用于：对客户端以及与第一身份标识建立通信关系的第二身份标识对应的客户端上传的音频数据进行拉取，以及获取所拉取音频数据的身份标识；相应的，客户端，具体用于：接收业务服务器发送的身份标识，并在展示字幕内容的同时展示身份标识。

20 其中，客户端向服务器发送的开启字幕请求中可携带有第一身份标识。业务服务器中可维护有各身份标识之间建立的通信关系，根据该第一身份标识可查询出与之建立通信关系的第二身份标识，进而可对客户端以及与第一身份标识建立通信关系的第二身份标识对应的客户端上传的音频数据进行拉取。并且，业务服务器在拉取音频的同时，还可同时从流媒体服务器获取所拉取音频数据的身份标识（例如昵称或者头像等），并可将身份标识连同对应的音频数据转化的文字内容一起发送至客户端，以便在展示字幕内容的同时展示字幕内容对应的身份标识，便于用户同时了解说话人身份和说话人所说的内容，提高了用户体验。

25 本公开实施例提供的信息处理系统还包括流媒体服务器和/或实时语音转写工具，能够分别实现对各客户端上传的音频数据的存储，以及将音频数据实时转化为文字内容并输出中间转化结果。此外，业务服务器还可以根据接收的开启字幕请求中携带的字幕语种，对获取的文字内容进行翻译，并将翻译完毕的文字内容发送至客户端，从而实现客户端展示需求语种，提高了用户体验。进一步的，业务服务器还可以将拉取的音频数据的身份标识发送至客户端，从而实现客户端展示说话人身份，提高了用户体验。

#### 实施例七

本实施例提供的信息处理方法，包括：

35 接收来自业务服务器的第一文字内容及第一文字内容的第一句标识；

展示第一文字内容；

接收来自业务服务器的第二文字内容及第二文字内容的第二句标识；

当第二句标识与第一句标识相同时，展示第二文字内容以更新第一文字内容。

可选的，在接收来自业务服务器的第一文字内容及第一文字内容的第一句标识的同时，

还包括：

接收来自业务服务器的第一句标识对应的第一内容版本序号；

在接收来自业务服务器的第二文字内容及第二文字内容的第二句标识的同时，还包括：

接收来自业务服务器的第二句标识对应的第二内容版本序号；

5 相应的，当第二句标识与第一句标识相同时，展示第二文字内容以更新第一文字内容，包括：

当第二句标识与第一句标识相同，且第二内容版本序号排在第一内容版本序号之后时，展示第二文字内容以更新第一文字内容。

展示第一文字内容，包括：

10 在第一显示控件中展示第一文字内容，第一显示控件用于对来自业务服务器的文字内容进行字幕展示。

可选的，展示第二文字内容以更新第一文字内容，包括：

在第一显示控件中展示第二文字内容以替换第一文字内容。

可选的，展示第二文字内容以更新第一文字内容，包括：

15 在第一显示控件中，展示基于第一文字内容和第二文字内容的区别内容更新的文字内容。

可选的，该方法还包括：

当第二句标识与第一句标识不同时，在第二显示控件中展示第二文字内容，第二显示控件用于对来自业务服务器的文字内容进行展示。

20 可选的，在接收来自业务服务器的第一文字内容及第一文字内容的第一句标识之前，该方法还包括：

向业务服务器发送开启字幕请求。

可选的，开启字幕请求包括表征客户端的身份标识或表征客户端所在场所的标识。

可选的，开启字幕请求包括文字内容的语种信息。

本公开实施例所提供的信息处理方法，可应用于本公开任意实施例中。

25 实施例八

图 6 为本公开实施例八所提供的一种信息处理装置结构示意图。本实施例提供的信息处理装置应用于客户端，尤其适用于客户端在实时互动场景下进行字幕展示的情形。

如图 6 所示，本实施例提供的信息处理装置，应用于客户端，包括：

接收模块 610，用于接收业务服务器发送的文字内容及文字内容的句标识；

30 展示模块 620，用于读取本地句标识，并于接收的句标识与本地句标识相同时，利用文字内容更新已展示的本地句标识对应的字幕内容。

在上述各技术方案的基础上，接收模块在接收业务服务器发送的文字内容及文字内容的句标识的同时，还用于：接收业务服务器发送的句标识对应的内容版本序号；

相应的，展示模块，具体用于：

35 读取本地句标识和本地句标识对应的本地内容版本序号；

当接收的句标识与本地句标识相同，接收的内容版本序号排在本地内容版本序号之后时，利用文字内容更新已展示的本地句标识和本地内容版本序号对应的字幕内容；

利用接收的内容版本序号更新本地句标识对应的本地内容版本序号。

在上述各技术方案的基础上，展示模块，还用于：当接收的句标识与本地句标识不同时，

将文字内容作为新的字幕内容进行展示，并利用接收的句标识更新本地句标识。

在上述各技术方案的基础上，当接收的文字内容和句标识为至少两组时，展示模块，具体用于：

读取本地句标识，判断接收的至少两个句标识中是否存在与本地句标识相同的句标识；

5 若是，则利用相同的句标识对应的文字内容更新已展示的本地句标识对应的字幕内容；将不同的句标识对应的文字内容作为新的字幕内容进行展示；

若否，则将接收的至少两个文字内容分别作为新的至少两个字幕内容进行展示；

将接收的至少两个句标识同时作为新的本地句标识。

10 在上述各技术方案的基础上，展示模块，具体用于：利用文字内容替换已展示的本地句标识对应的字幕内容。

在上述各技术方案的基础上，客户端为实时互动应用的客户端，且实时互动应用包括即时通讯应用、多媒体会议应用、视频直播应用和群聊互动应用中的至少一种。

本公开实施例所提供的信息处理装置，可执行本公开任意实施例所提供的应用于客户端的信息处理方法，具备执行方法相应的功能模块和有益效果。

15 值得注意的是，上述装置所包括的各个单元和模块只是按照功能逻辑进行划分的，但并不局限于上述的划分，只要能够实现相应的功能即可；另外，各功能单元的具体名称也只是为了便于相互区分，并不用于限制本公开实施例的保护范围。

#### 实施例九

20 下面参考图 7，其示出了适于用来实现本公开实施例的电子设备（例如图 7 中的终端设备或服务器）700 的结构示意图。本公开实施例中的终端设备可以包括但不限于诸如移动电话、笔记本电脑、数字广播接收器、PDA（个人数字助理）、PAD（平板电脑）、PMP（便携式多媒体播放器）、车载终端（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 7 示出的电子设备仅仅是一个示例，不对本公开实施例的功能和使用范围带来任何限制。

25 如图 7 所示，电子设备 700 可以包括处理装置（例如中央处理器、图形处理器等）701，其可以根据存储在只读存储器（Read-Only Memory, ROM）702 中的程序或者从存储装置 706 加载到随机访问存储器（Random Access Memory, RAM）703 中的程序而执行各种适当的动作和处理。在 RAM 703 中，还存储有电子设备 700 操作所需的各种程序和数据。处理装置 701、ROM 702 以及 RAM 703 通过总线 704 彼此相连。输入/输出（I/O）接口 705 也连接至总线 704。

30 通常，以下装置可以连接至 I/O 接口 705：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 706；包括例如液晶显示器（LCD）、扬声器、振动器等的输出装置 707；包括例如磁带、硬盘等的存储装置 708；以及通信装置 709。通信装置 709 可以允许电子设备 700 与其他设备进行无线或有线通信以交换数据。虽然图 7 示出了具有各种装置的电子设备 700，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  
35 替代地实施或具备更多或更少的装置。

特别地，根据本公开的实施例，上文参考流程图描述的过程可以被实现为计算机软件程序。例如，本公开的实施例包括一种计算机程序产品，其包括承载在非暂态计算机可读介质上的计算机程序，该计算机程序包含用于执行流程图所示的方法的程序代码。在这样的实施例中，该计算机程序可以通过通信装置 709 从网络上被下载和安装，或者从存储装置 706 被

安装，或者从 ROM702 被安装。在该计算机程序被处理装置 701 执行时，执行本公开实施例的方法中限定的上述功能。

本公开实施例提供的电子设备与上述实施例提供的信息处理方法属于同一公开构思，未在本实施例中详尽描述的技术细节可参见上述实施例，并且本实施例与上述实施例具有相同的有益效果。

#### 实施例十

本公开实施例提供了一种计算机存储介质，其上存储有计算机程序，该程序被处理器执行时实现上述实施例所提供的信息处理方法。

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、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（FLASH）、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（Hyper Text 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：

接收业务服务器发送的文字内容及文字内容的句标识；

读取本地句标识，并于接收的句标识与本地句标识相同时，利用文字内容更新已展示的本地句标识对应的字幕内容。

可以以一种或多种程序设计语言或其组合来编写用于执行本公开的操作的计算机程序代码，上述程序设计语言包括但不限于面向对象的程序设计语言—诸如 Java、Smalltalk、C++，还包括常规的过程式程序设计语言—诸如“C”语言或类似的程序设计语言。程序代码可以完全地在用户计算机上执行、部分地在用户计算机上执行、作为一个独立的软件包执行、部分

在用户计算机上部分在远程计算机上执行、或者完全在远程计算机或服务器上执行。在涉及远程计算机的情形中，远程计算机可以通过任意种类的网络——包括局域网(LAN)或广域网(WAN)一连接到用户计算机，或者，可以连接到外部计算机（例如利用因特网服务提供商来通过因特网连接）。

#### 5 实施例十一

本公开实施例还提供了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，该计算机程序被处理器执行时实现上述实施例所提供的信息处理方法。

#### 实施例十二

10 本公开实施例还提供了一种计算机程序，当其在电子设备上运行时，实现上述实施例所提供的信息处理方法。

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。

20 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、模块的名称在某种情况下并不构成对该单元、模块本身的限定，例如，展示模块还可以被描述为“字幕展示模块”。

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 Chip, SOC)、复杂可编程逻辑设备(CPLD)等等。

30 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。

35 根据本公开的一个或多个实施例，【示例一】提供了一种信息处理方法，应用于客户端，该方法包括：

接收业务服务器发送的文字内容及所述文字内容的句标识；

读取本地句标识，并于接收的句标识与所述本地句标识相同时，利用所述文字内容更新

已展示的所述本地句标识对应的字幕内容。

根据本公开的一个或多个实施例，【示例二】提供了一种信息处理方法，还包括：

可选的，在所述接收业务服务器发送的文字内容及所述文字内容的句标识的同时，还包括：

5 接收业务服务器发送的所述句标识对应的内容版本序号；

相应的，所述读取本地句标识，并于接收的句标识与所述本地句标识相同时，利用所述文字内容更新已展示的所述本地句标识对应的字幕内容，包括：

读取本地句标识和所述本地句标识对应的本地内容版本序号；

10 当接收的句标识与所述本地句标识相同，接收的内容版本序号排在所述本地内容版本序号之后时，利用所述文字内容更新已展示的所述本地句标识和所述本地内容版本序号对应的字幕内容；

利用接收的内容版本序号更新本地句标识对应的本地内容版本序号。

根据本公开的一个或多个实施例，【示例三】提供了一种信息处理方法，还包括：

15 当接收的句标识与所述本地句标识不同时，将所述文字内容作为新的字幕内容进行展示，并利用接收的句标识更新本地句标识。

根据本公开的一个或多个实施例，【示例四】提供了一种信息处理方法，还包括：

可选的，当接收的文字内容和句标识为至少两组时，所述读取本地句标识，并于接收的句标识与所述本地句标识相同时，利用所述文字内容更新已展示的所述本地句标识对应的字幕内容，包括：

20 读取本地句标识，判断接收的至少两个句标识中是否存在与所述本地句标识相同的句标识；

若是，则利用相同的句标识对应的文字内容更新已展示的所述本地句标识对应的字幕内容；将不同的句标识对应的文字内容作为新的字幕内容进行展示；

25 若否，则将接收的至少两个文字内容分别作为新的至少两个字幕内容进行展示；将接收的至少两个句标识同时作为新的本地句标识。

根据本公开的一个或多个实施例，【示例五】提供了一种信息处理方法，还包括：

可选的，所述利用所述文字内容更新已展示的所述本地句标识对应的字幕内容，包括：利用所述文字内容替换已展示的所述本地句标识对应的字幕内容。

30 根据本公开的一个或多个实施例，【示例六】提供了一种信息处理方法，还包括：可选的，所述客户端为实时互动应用的客户端，且所述实时互动应用包括即时通讯应用、多媒体会议应用、视频直播应用和群聊互动应用中的至少一种。

35 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，但是这些不应被解释为对本公开的范围

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

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

## 权利要求书

- 1、一种信息处理方法，其特征在于，应用于客户端，包括：  
接收业务服务器发送的文字内容及所述文字内容的句标识；  
读取本地句标识，并于接收的句标识与所述本地句标识相同时，利用所述文字内容更新  
5 已展示的所述本地句标识对应的字幕内容。
- 2、根据权利要求1所述的方法，其特征在于，在所述接收业务服务器发送的文字内容及  
所述文字内容的句标识的同时，还包括：  
接收业务服务器发送的所述句标识对应的内容版本序号；  
相应的，所述读取本地句标识，并于接收的句标识与所述本地句标识相同时，利用所述  
10 文字内容更新已展示的所述本地句标识对应的字幕内容，包括：  
读取本地句标识和所述本地句标识对应的本地内容版本序号；  
当接收的句标识与所述本地句标识相同，接收的内容版本序号排在所述本地内容版本序  
号之后时，利用所述文字内容更新已展示的所述本地句标识和所述本地内容版本序号对应的  
字幕内容；  
15 利用接收的内容版本序号更新本地句标识对应的本地内容版本序号。
- 3、根据权利要求1或2所述的方法，其特征在于，还包括：  
当接收的句标识与所述本地句标识不同时，将所述文字内容作为新的字幕内容进行展示，  
并利用接收的句标识更新本地句标识。
- 4、根据权利要求1-3任一项所述的方法，其特征在于，当接收的文字内容和句标识为至  
20 少两组时，所述读取本地句标识，并于接收的句标识与所述本地句标识相同时，利用所述文  
字内容更新已展示的所述本地句标识对应的字幕内容，包括：  
读取本地句标识，判断接收的至少两个句标识中是否存在与所述本地句标识相同的句标  
识；  
若是，则利用相同的句标识对应的文字内容更新已展示的所述本地句标识对应的字幕内  
25 容；将不同的句标识对应的文字内容作为新的字幕内容进行展示；  
若否，则将接收的至少两个文字内容分别作为新的至少两个字幕内容进行展示；  
将接收的至少两个句标识同时作为新的本地句标识。
- 5、根据权利要求1-4任一项所述的方法，其特征在于，所述利用所述文字内容更新已展  
示的所述本地句标识对应的字幕内容，包括：  
30 利用所述文字内容替换已展示的所述本地句标识对应的字幕内容。
- 6、根据权利要求1-5任一项所述的方法，其特征在于，所述客户端为实时互动应用的客  
户端，且所述实时互动应用包括即时通讯应用、多媒体会议应用、视频直播应用和群聊互动  
应用中的至少一种。
- 7、一种信息处理系统，其特征在于，包括：  
35 业务服务器，用于响应于客户端发送的开启字幕请求表征对音频数据进行拉取，获取所  
述音频数据中当前待转化音频数据在文字转化过程中的文字内容，以及根据所述当前待转化  
音频数据的变化更新所述文字内容的句标识；  
所述客户端，用于接收所述业务服务器发送的所述文字内容及所述文字内容的句标识；  
读取本地句标识，并于接收的句标识与所述本地句标识相同时，利用所述文字内容更新已展

示的所述本地句标识对应的字幕内容。

8、根据权利要求 7 所述的系统，其特征在于，所述业务服务器，具体用于：

于当前待转化音频数据在文字转化完毕时，设置所述句标识的更新标记；

在获取到新的当前待转化音频数据在文字转化过程中的文字内容时，根据所述更新标记

5 更新所述句标识，并于所述句标识更新完毕时取消所述更新标记。

9、根据权利要求 7 或 8 所述的系统，其特征在于，所述业务服务器，还用于：

根据所述文字内容在所述当前待转化音频数据的文字转化过程中的转化次序，生成所述句标识对应的内容版本序号；

相应的，所述客户端，还用于：

10 读取所述本地句标识对应的本地内容版本序号；

当接收的句标识与所述本地句标识相同，接收的内容版本序号排在所述本地内容版本序号之后时，利用所述文字内容替换已展示的所述本地句标识和所述本地内容版本序号对应的字幕内容；

利用接收的内容版本序号更新本地句标识对应的本地内容版本序号。

15 10、根据权利要求 7-9 任一项所述的系统，其特征在于，所述客户端，还用于：

当接收的句标识与所述本地句标识不同时，将所述文字内容作为新的字幕内容进行展示，并利用接收的句标识更新本地句标识。

11、根据权利要求 7-10 任一项所述的系统，其特征在于，所述业务服务器，还用于：

20 响应于所述开启字幕请求表征对至少两路音频数据进行拉取，获取所述至少两路音频数据中当前待转化音频数据在文字转化过程中的文字内容，以及根据至少两路当前待转化音频数据的变化更新文字内容的句标识；

相应的，所述客户端还用于：

当接收的文字内容和句标识为至少两组时，读取本地句标识，判断接收的至少两个句标识中是否存在与所述本地句标识相同的句标识；

25 若是，则利用相同的句标识对应的文字内容替换已展示的所述本地句标识对应的字幕内容；将不同的句标识对应的文字内容作为新的字幕内容进行展示；

若否，则将接收的至少两个文字内容分别作为新的至少两个字幕内容进行展示；

将接收的至少两个句标识同时作为新的本地句标识。

30 12、根据权利要求 7-11 任一项所述的系统，其特征在于，所述开启字幕请求中携带有请求字幕语种；

相应的，所述业务服务器，还用于：

根据所述请求字幕语种对获取的文字内容进行翻译，并将翻译完毕的文字内容发送至所述客户端。

35 13、根据权利要求 7-12 任一项所述的系统，其特征在于，所述开启字幕请求中携带有第一身份标识；

相应的，所述业务服务器，具体用于：

对所述客户端以及与所述第一身份标识建立通信关系的第二身份标识对应的客户端上传的音频数据进行拉取，以及获取所拉取音频数据的身份标识；

相应的，所述客户端，具体用于：

接收所述业务服务器发送的身份标识，并在展示所述字幕内容的同时展示所述身份标识。

14、根据权利要求 7-13 任一项所述的系统，其特征在于，还包括：

流媒体服务器，用于接收进行流媒体通信的客户端上传的音频数据；

相应的，所述业务服务器，具体用于：

5 响应于客户端发送的开启字幕请求表征，从所述流媒体服务器拉取音频数据。

15、根据权利要求 7-14 任一项所述的系统，其特征在于，还包括：

实时语音转写工具，用于接收所述业务服务器发送的音频数据，确定所述音频数据的当前待转化音频数据，并反馈所述当前待转化音频数据在文字转化过程中的文字内容；

相应的，所述业务服务器，具体用于：

10 接收所述实时语音转写工具反馈的所述音频数据中当前待转化音频数据在文字转化过程中的文字内容。

16、一种信息处理装置，其特征在于，应用于客户端，包括：

接收模块，用于接收业务服务器发送的文字内容及所述文字内容的句标识；

15 展示模块，用于读取本地句标识，并于接收的句标识与所述本地句标识相同时，利用所述文字内容更新已展示的所述本地句标识对应的字幕内容。

17、一种信息处理方法，包括：

接收来自业务服务器的第一文字内容及所述第一文字内容的第一句标识；

展示所述第一文字内容；

接收来自所述业务服务器的第二文字内容及所述第二文字内容的第二句标识；

20 当所述第二句标识与所述第一句标识相同时，展示所述第二文字内容以更新所述第一文字内容。

18、根据权利要求 17 所述的方法，其特征在于，在接收来自业务服务器的第一文字内容及所述第一文字内容的第一句标识的同时，还包括：

接收来自所述业务服务器的所述第一句标识对应的第一内容版本序号；

25 在接收来自业务服务器的第二文字内容及所述第二文字内容的第二句标识的同时，还包括：

接收来自所述业务服务器的所述第二句标识对应的第二内容版本序号；

相应的，所述当所述第二句标识与所述第一句标识相同时，展示所述第二文字内容以更新所述第一文字内容，包括：

30 当所述第二句标识与所述第一句标识相同，且所述第二内容版本序号排在所述第一内容版本序号之后时，展示所述第二文字内容以更新所述第一文字内容。

19、根据权利要求 17 或 18 所述的方法，其特征在于，所述展示所述第一文字内容，包括：

35 在第一显示控件中展示所述第一文字内容，所述第一显示控件用于对来自所述业务服务器的文字内容进行字幕展示。

20、根据权利要求 19 所述的方法，其特征在于，所述展示所述第二文字内容以更新所述第一文字内容，包括：

在所述第一显示控件中展示所述第二文字内容以替换所述第一文字内容。

21、根据权利要求 19 所述的方法，其特征在于，所述展示所述第二文字内容以更新所述

第一文字内容，包括：

在所述第一显示控件中，展示基于所述第一文字内容和所述第二文字内容的区别内容更新的文字内容。

22、根据权利要求 19-21 任一项所述的方法，其特征在于，所述方法还包括：

5 当所述第二句标识与所述第一句标识不同时，在第二显示控件中展示所述第二文字内容，所述第二显示控件用于对来自所述业务服务器的文字内容进行展示。

23、根据权利要求 17-22 任一项所述的方法，其特征在于，在所述接收来自业务服务器的第一文字内容及所述第一文字内容的第一句标识之前，所述方法还包括：

向所述业务服务器发送开启字幕请求。

10 24、根据权利要求 23 所述的方法，其特征在于，所述开启字幕请求包括表征所述客户端的身份标识或表征所述客户端所在场所的标识。

25、根据权利要求 23 或 24 所述的方法，其特征在于，所述开启字幕请求包括所述文字内容的语种信息。

26、一种电子设备，其特征在于，所述电子设备包括：

15 一个或多个处理器；

存储装置，用于存储一个或多个程序，

当所述一个或多个程序被所述一个或多个处理器执行，使得所述一个或多个处理器实现如权利要求 1-6 或 17-25 任一项所述的信息处理方法。

20 27、一种包含计算机可执行指令的存储介质，所述计算机可执行指令在由计算机处理器执行时用于执行如权利要求 1-6 或 17-25 任一项所述的信息处理方法。

28、一种计算机程序产品，其包括承载在计算机可读存储介质上的计算机程序，所述计算机程序被处理器执行时，执行权利要求 1-6 或 17-25 任一项所述的信息处理方法。

29、一种计算机程序，当其在电子设备上运行时，使得所述电子设备执行权利要求 1-6 或 17-25 任一项所述的信息处理方法。

25

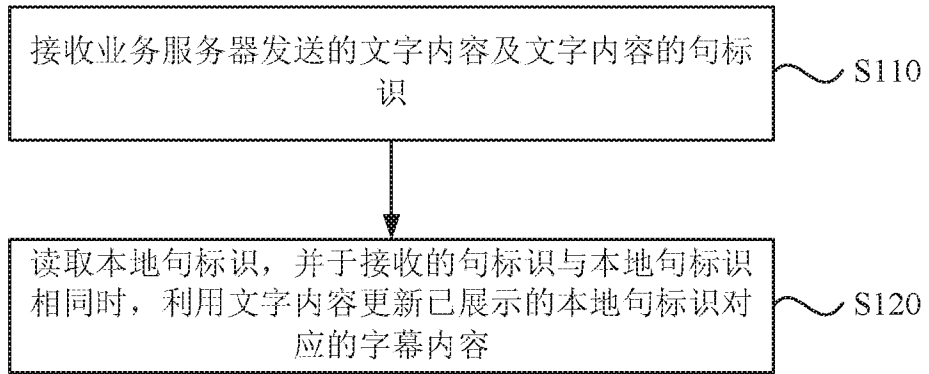


图 1

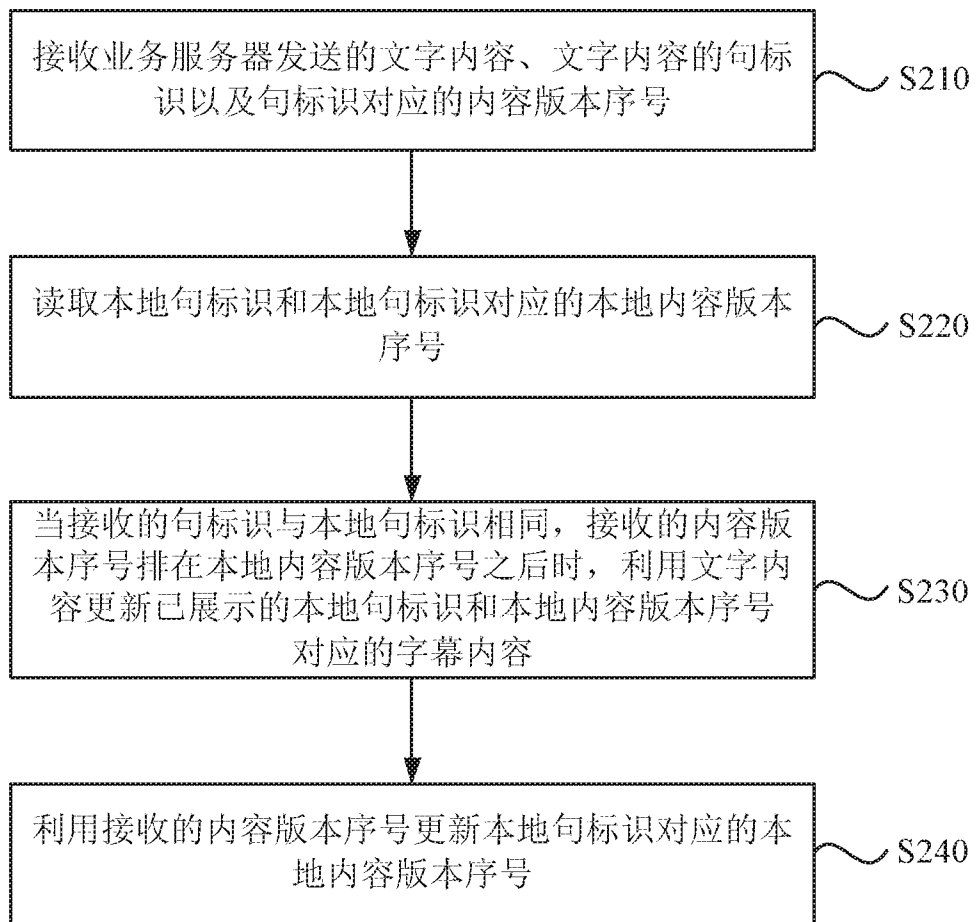


图 2

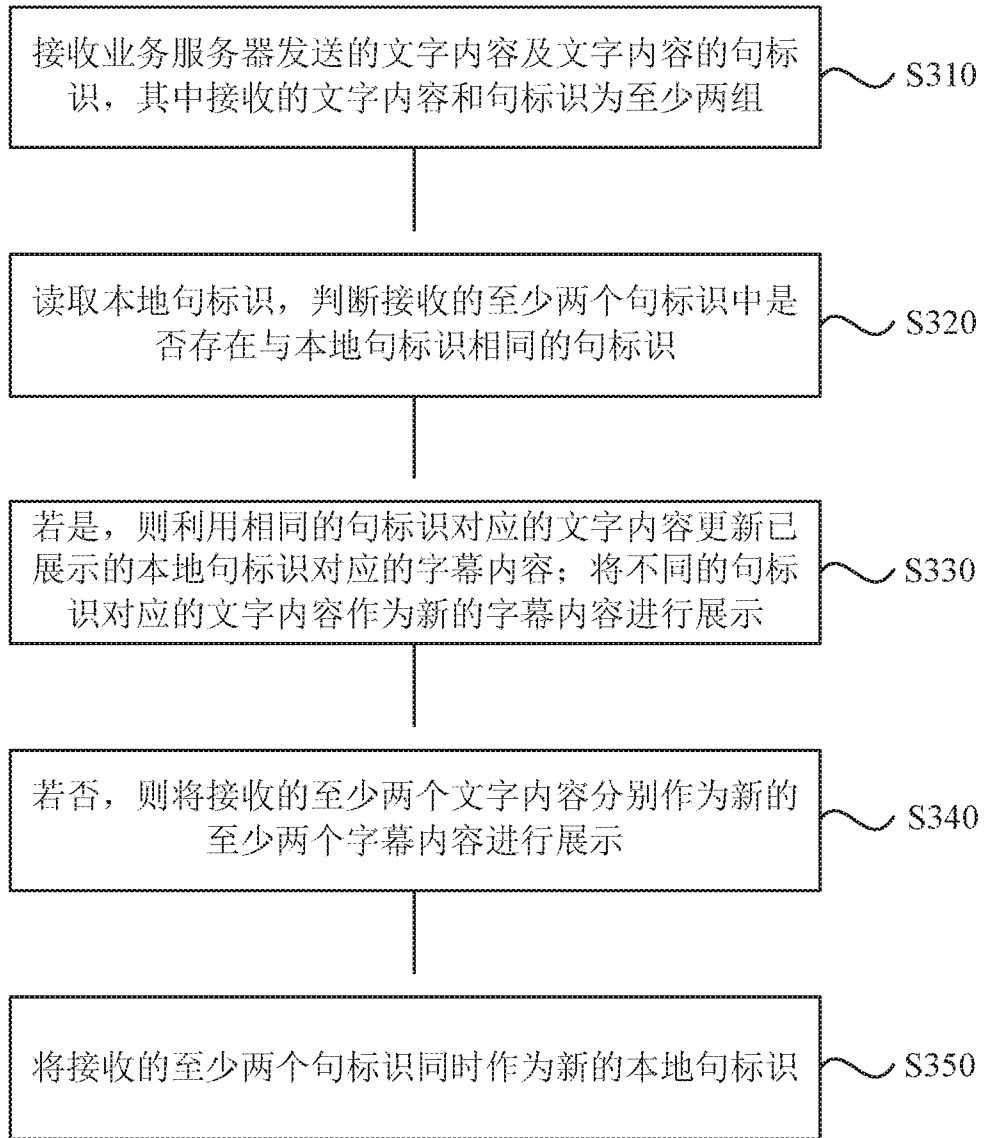


图 3

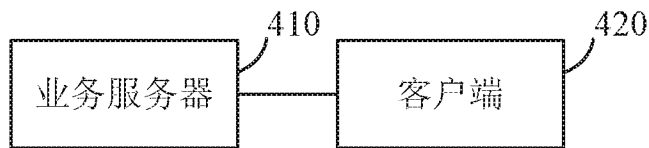


图 4

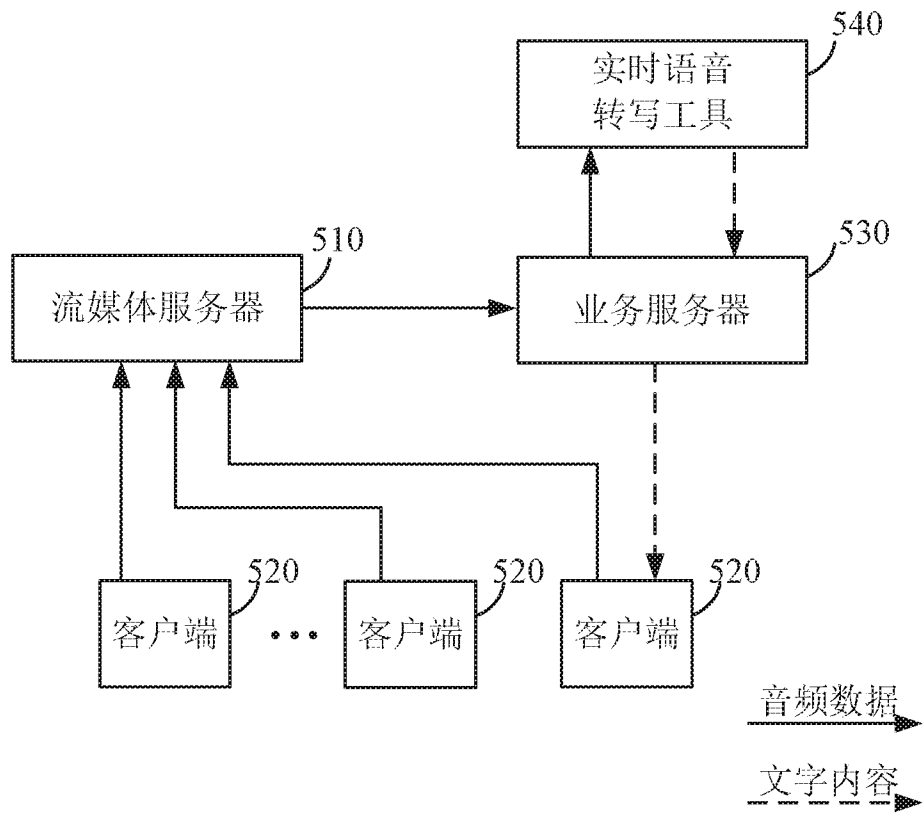


图 5

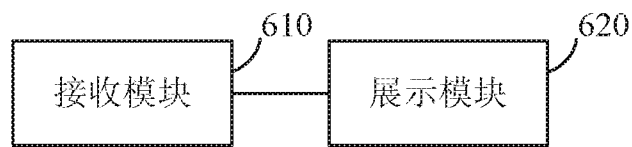


图 6

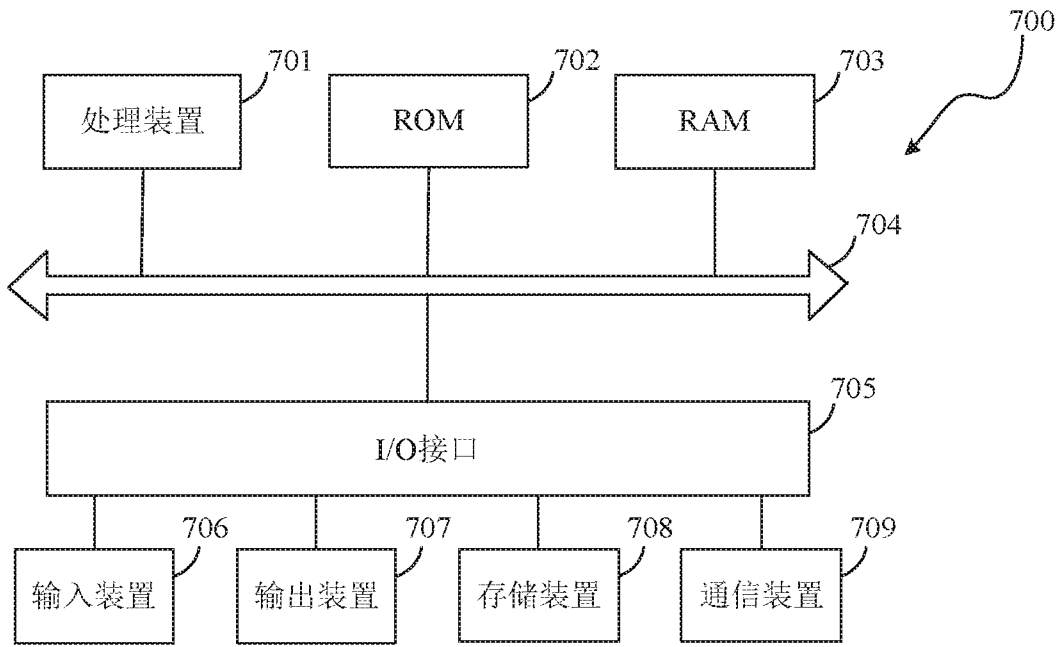


图 7

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2021/098533

<b>A. CLASSIFICATION OF SUBJECT MATTER</b>		
H04N 21/439(2011.01)i; H04N 21/458(2011.01)i; H04N 21/485(2011.01)i; H04N 21/488(2011.01)i; G10L 15/26(2006.01)i		
According to International Patent Classification (IPC) or to both national classification and IPC		
<b>B. FIELDS SEARCHED</b>		
Minimum documentation searched (classification system followed by classification symbols) H04N; G10L		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) CNABS; CNTXT; VEN; USTXT; EPTXT; WOTXT; CNKI: 字幕, 语音, 音频, 识别, 文字, 字符, 文本, 更新, 替换, 句, 标识, 校正, 纠错, subtitle, caption, voice, audio, text, character, correct, update, replace		
<b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b>		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 111711853 A (BEIJING BYTEDANCE NETWORK TECHNOLOGY CO., LTD.) 25 September 2020 (2020-09-25) description, paragraphs [0034]-[0163], and figures 1-7	1-29
X	US 2015051908 A1 (CAPTIONCALL LLC) 19 February 2015 (2015-02-19) description, paragraphs [0008]-[0034], and figures 1-6	1-29
X	US 2013158995 A1 (SORENSEN COMMUNICATIONS INC) 20 June 2013 (2013-06-20) description, paragraphs [0008]-[0034], and figures 1-6	1-29
X	CN 110415706 A (BEE SMART INFORMATION TECHNOLOGY CO., LTD.) 05 November 2019 (2019-11-05) description, paragraphs [0005]-[0025], and figures 1-3	1-29
X	CN 105513586 A (BAIDU ONLINE NETWORK TECHNOLOGY (BEIJING) CO., LTD.) 20 April 2016 (2016-04-20) description, paragraphs [0021]-[0103], and figure 5	1-29
A	CN 106412678 A (ANHUI SEMXUM INFORMATION TECHNOLOGY CO., LTD.) 15 February 2017 (2017-02-15) entire document	1-29
<input 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 <b>05 August 2021</b>		Date of mailing of the international search report <b>18 August 2021</b>
Name and mailing address of the ISA/CN <b>China National Intellectual Property Administration (ISA/CN) No. 6, Xitucheng Road, Jimenqiao, Haidian District, Beijing 100088 China</b>		Authorized officer
Facsimile No. (86-10)62019451		Telephone No.

**INTERNATIONAL SEARCH REPORT**  
**Information on patent family members**

International application No. <b>PCT/CN2021/098533</b>
---

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
CN	111711853	A	25 September 2020	None	
US	2015051908	A1	19 February 2015	US 9336689 B2	10 May 2016
US	2013158995	A1	20 June 2013	None	
CN	110415706	A	05 November 2019	None	
CN	105513586	A	20 April 2016	None	
CN	106412678	A	15 February 2017	None	

<p><b>A. 主题的分类</b></p> <p>H04N 21/439(2011.01)i; H04N 21/458(2011.01)i; H04N 21/485(2011.01)i; H04N 21/488(2011.01)i; G10L 15/26(2006.01)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>B. 检索领域</b></p> <p>检索的最低限度文献(标明分类系统和分类号)</p> <p>H04N; G10L</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>CNABS;CNTXT;VEN;USTXT;EPTXT;WOTXT;CNKI; 字幕, 语音, 音频, 识别, 文字, 字符, 文本, 更新, 替换, 句, 标识, 校正, 纠错, subtitle, caption, voice, audio, text, character, correct, update, replace</p>																							
<p><b>C. 相关文件</b></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 111711853 A (北京字节跳动网络技术有限公司) 2020年 9月 25日 (2020 - 09 - 25) 说明书第[0034]-[0163]段, 附图1-7</td> <td>1-29</td> </tr> <tr> <td>X</td> <td>US 2015051908 A1 (CAPTIONCALL LLC) 2015年 2月 19日 (2015 - 02 - 19) 说明书第[0008]-[0034]段, 附图1-6</td> <td>1-29</td> </tr> <tr> <td>X</td> <td>US 2013158995 A1 (SORENSEN COMMUNICATIONS INC) 2013年 6月 20日 (2013 - 06 - 20) 说明书第[0008]-[0034]段, 附图1-6</td> <td>1-29</td> </tr> <tr> <td>X</td> <td>CN 110415706 A (常州市小先信息技术有限公司) 2019年 11月 5日 (2019 - 11 - 05) 说明书第[0005]-[0025]段, 附图1-3</td> <td>1-29</td> </tr> <tr> <td>X</td> <td>CN 105513586 A (百度在线网络技术北京有限公司) 2016年 4月 20日 (2016 - 04 - 20) 说明书第[0021]-[0103]段, 附图5</td> <td>1-29</td> </tr> <tr> <td>A</td> <td>CN 106412678 A (安徽声讯信息技术有限公司) 2017年 2月 15日 (2017 - 02 - 15) 全文</td> <td>1-29</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 111711853 A (北京字节跳动网络技术有限公司) 2020年 9月 25日 (2020 - 09 - 25) 说明书第[0034]-[0163]段, 附图1-7	1-29	X	US 2015051908 A1 (CAPTIONCALL LLC) 2015年 2月 19日 (2015 - 02 - 19) 说明书第[0008]-[0034]段, 附图1-6	1-29	X	US 2013158995 A1 (SORENSEN COMMUNICATIONS INC) 2013年 6月 20日 (2013 - 06 - 20) 说明书第[0008]-[0034]段, 附图1-6	1-29	X	CN 110415706 A (常州市小先信息技术有限公司) 2019年 11月 5日 (2019 - 11 - 05) 说明书第[0005]-[0025]段, 附图1-3	1-29	X	CN 105513586 A (百度在线网络技术北京有限公司) 2016年 4月 20日 (2016 - 04 - 20) 说明书第[0021]-[0103]段, 附图5	1-29	A	CN 106412678 A (安徽声讯信息技术有限公司) 2017年 2月 15日 (2017 - 02 - 15) 全文	1-29
类型*	引用文件, 必要时, 指明相关段落	相关的权利要求																					
PX	CN 111711853 A (北京字节跳动网络技术有限公司) 2020年 9月 25日 (2020 - 09 - 25) 说明书第[0034]-[0163]段, 附图1-7	1-29																					
X	US 2015051908 A1 (CAPTIONCALL LLC) 2015年 2月 19日 (2015 - 02 - 19) 说明书第[0008]-[0034]段, 附图1-6	1-29																					
X	US 2013158995 A1 (SORENSEN COMMUNICATIONS INC) 2013年 6月 20日 (2013 - 06 - 20) 说明书第[0008]-[0034]段, 附图1-6	1-29																					
X	CN 110415706 A (常州市小先信息技术有限公司) 2019年 11月 5日 (2019 - 11 - 05) 说明书第[0005]-[0025]段, 附图1-3	1-29																					
X	CN 105513586 A (百度在线网络技术北京有限公司) 2016年 4月 20日 (2016 - 04 - 20) 说明书第[0021]-[0103]段, 附图5	1-29																					
A	CN 106412678 A (安徽声讯信息技术有限公司) 2017年 2月 15日 (2017 - 02 - 15) 全文	1-29																					
<p><input type="checkbox"/> 其余文件在C栏的续页中列出。 <input checked="" type="checkbox"/> 见同族专利附件。</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>“&amp;” 同族专利的文件</p>																							
<p>国际检索实际完成的日期</p> <p>2021年 8月 5日</p>		<p>国际检索报告邮寄日期</p> <p>2021年 8月 18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>章婧</p> <p>电话号码 86-(512)-88996193</p>																					

国际检索报告  
关于同族专利的信息

国际申请号

PCT/CN2021/098533

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	111711853	A	2020年 9月 25日	无	
US	2015051908	A1	2015年 2月 19日	US 9336689 B2	2016年 5月 10日
US	2013158995	A1	2013年 6月 20日	无	
CN	110415706	A	2019年 11月 5日	无	
CN	105513586	A	2016年 4月 20日	无	
CN	106412678	A	2017年 2月 15日	无	