

(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
2017 年 10 月 5 日 (05.10.2017)



(10) 国际公布号
WO 2017/167047 A1

- (51) 国际专利分类号:
G10L 21/10 (2013.01) H04W 4/18 (2009.01)
- (21) 国际申请号: PCT/CN2017/077257
- (22) 国际申请日: 2017 年 3 月 20 日 (20.03.2017)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201610187534.9 2016 年 3 月 29 日 (29.03.2016) CN
- (71) 申请人: 阿里巴巴集团控股有限公司 (ALIBABA GROUP HOLDING LIMITED) [—/CN]; 开曼群岛大开曼资本大厦一座四层 847 号邮箱, Grand Cayman (KY)。
- (72) 发明人: 及
- (71) 申请人 (仅对美国): 张达平 (ZHANG, Daping) [CN/CN]; 中国浙江省杭州市余杭区文一西路 969 号 3 号楼 5 楼阿里巴巴集团法务部, Zhejiang 311121 (CN)。张黎黎 (ZHANG, Lili) [CN/CN]; 中国浙江省杭州市余杭区文一西路 969 号 3 号楼 5 楼阿里巴巴集团法务部, Zhejiang 311121 (CN)。黄益信 (HUANG, Yixin) [CN/CN]; 中国浙江省杭州市余杭区文一西路 969 号 3 号楼 5 楼阿里巴巴集团法务

部, Zhejiang 311121 (CN)。陈鋈 (CHEN, Yun) [CN/CN]; 中国浙江省杭州市余杭区文一西路 969 号 3 号楼 5 楼阿里巴巴集团法务部, Zhejiang 311121 (CN)。赖建冬 (LAI, Jiandong) [CN/CN]; 中国浙江省杭州市余杭区文一西路 969 号 3 号楼 5 楼阿里巴巴集团法务部, Zhejiang 311121 (CN)。钟浩华 (ZHONG, Haohua) [CN/CN]; 中国浙江省杭州市余杭区文一西路 969 号 3 号楼 5 楼阿里巴巴集团法务部, Zhejiang 311121 (CN)。

(74) 代理人: 北京国昊天诚知识产权代理有限公司 (CO-HORIZON INTELLECTUAL PROPERTY INC.); 中国北京市朝阳区小关北里甲 2 号渔阳置业大厦 B 座 605, Beijing 100029 (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, 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, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW。

[见续页]

(54) Title: METHOD AND DEVICE FOR PROCESSING VOICE MESSAGE

(54) 发明名称: 音频消息的处理方法及装置

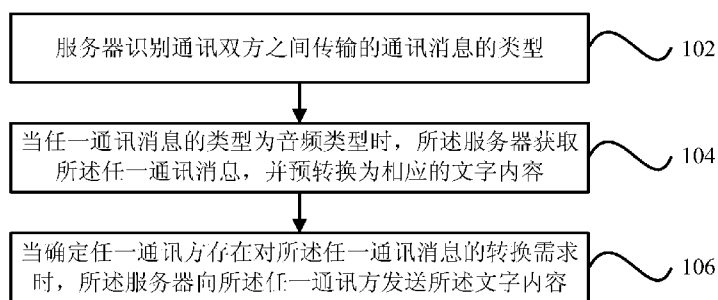


图 1

- 102 A server identifies a type of a communication message transmitted between communication parties
- 104 When the type of the communication message is a voice type, the server acquires the communication message and converts the same into a corresponding text
- 106 Upon determining that any of the communication parties requires to convert the communication message, the server transmits to the communication party the text

(57) Abstract: A method and device for processing a voice message. The method comprises: a server identifies a type of a communication message transmitted between communication parties (102); when the type of the communication message is a voice type, the server acquires the communication message and converts the same into a corresponding text (104); and upon determining that any of the communication parties requires to convert the communication message, the server transmits to the communication party the text (106). The method can convert in advance a voice message into a text, thereby increasing a response speed with respect to an audio conversion request of a user.

(57) 摘要: 一种音频消息的处理方法及装置, 该方法包括: 服务器识别通讯双方之间传输的通讯消息的类型 (102); 当任一通讯消息的类型为音频类型时, 所述服务器获取所述任一通讯消息, 并预转换为相应的文字内容 (104); 当确定任一通讯方存在对所述任一通讯消息的转换需求

时, 所述服务器向所述任一通讯方发送所述文字内容 (106)。该方法可预先对音频消息进行文字转换, 从而提升对用户的音频转换需求的响应速度。



WO 2017/167047 A1



(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))。

音频消息的处理方法及装置

技术领域

本申请涉及通讯技术领域，尤其涉及一种音频消息的处理方法及装置。

5 背景技术

通过安装有通讯应用的电子设备，用户之间可以实现通讯消息的收发，从而使用户间的沟通更为方便、快捷。

通常而言，通讯应用通过采集用户手动输入的文字，以作为通讯消息进行收发。然而，手动输入存在诸多方面的限制，比如用户需要双
10 眼盯住电子设备的屏幕，则当用户处于驾驶状态时，手动输入可能带来极大的安全风险；再比如，当电子设备较大、无法单手握持时，用户需要双手同时握持并完成输入操作，则假定用户一只手提着重物时，将难以通过另一只手来完成手动输入。

在相关技术中，一些通讯应用通过添加音频输入功能，使得用户能
15 够更为便捷地收发音频类型的通讯消息，而消除了上述限制。

发明内容

有鉴于此，本申请提供一种音频消息的处理方法及装置，可以预先对音频消息进行文字转换，从而提升对用户的音频转换需求的响应速度。

为实现上述目的，本申请提供技术方案如下：

20 根据本申请的第一方面，提出了一种音频消息的处理方法，包括：
服务器识别通讯双方之间传输的通讯消息的类型；

当任一通讯消息的类型为音频类型时，所述服务器获取所述任一通讯消息，并预转换为相应的文字内容；

当确定任一通讯方存在对所述任一通讯消息的转换需求时，所述服
25 务器向所述任一通讯方发送所述文字内容。

根据本申请的第二方面，提出了一种音频消息的处理方法，包括：

本端通讯设备在接收到用户发出的针对音频类型的任一通讯消息的音频转换命令时，向服务器发起相应的音频转换请求；

本端通讯设备接收到所述服务器返回的所述任一通讯消息对应的文字内容，并与所述任一通讯消息进行关联展示；其中，所述文字内容
5 由所述服务器在接收到所述音频转换请求之前主动预转换得到。

根据本申请的第三方面，提出了一种音频消息的处理方法，包括：

本端通讯设备预获取音频类型的任一通讯消息对应的文字内容；

当接收到用户发出的针对所述任一通讯消息的音频转换命令时，所述本端通讯设备示出预获取的所述文字内容。

10 根据本申请的第四方面，提出了一种音频消息的处理方法，包括：

在生成音频类型的通讯消息的过程中，本端通讯设备依次确定已采集到的每个音频片段是否符合预设切分规则；

当任一音频片段符合所述预设切分规则时，所述本端通讯设备将所述任一音频片段实时切分并上传至服务器，以由所述服务器将所述任一
15 音频片段预转换为相应的文字片段，且所有音频片段对应的文字片段由所述服务器依次拼接为所述通讯消息对应的文字内容。

根据本申请的第五方面，提出了一种音频消息的处理装置，包括：

识别单元，使服务器识别通讯双方之间传输的通讯消息的类型；

预转换单元，当任一通讯消息的类型为音频类型时，使所述服务器
20 获取所述任一通讯消息，并预转换为相应的文字内容；

发送单元，当确定任一通讯方存在对所述任一通讯消息的转换需求时，使所述服务器向所述任一通讯方发送所述文字内容。

根据本申请的第六方面，提出了一种音频消息的处理装置，包括：

请求单元，使本端通讯设备在接收到用户发出的针对音频类型的任一
25 一通讯消息的音频转换命令时，向服务器发起相应的音频转换请求；

展示单元，使本端通讯设备接收到所述服务器返回的所述任一通讯消息对应的文字内容，并与所述任一通讯消息进行关联展示；其中，所

述文字内容由所述服务器在接收到所述音频转换请求之前主动预转换得到。

根据本申请的第七方面，提出了一种音频消息的处理装置，包括：

5 预获取单元，使本端通讯设备预获取音频类型的任一通讯消息对应的文字内容；

展示单元，当接收到用户发出的针对所述任一通讯消息的音频转换命令时，使所述本端通讯设备示出预获取的所述文字内容。

根据本申请的第八方面，提出了一种音频消息的处理装置，包括：

10 确定单元，在生成音频类型的通讯消息的过程中，使本端通讯设备依次确定已采集到的每个音频片段是否符合预设切分规则；

处理单元，当任一音频片段符合所述预设切分规则时，使所述本端通讯设备将所述任一音频片段实时切分并上传至服务器，以由所述服务器将所述任一音频片段预转换为相应的文字片段，且所有音频片段对应的文字片段由所述服务器依次拼接为所述通讯消息对应的文字内容。

15 根据本申请的第九方面，提出了一种音频消息的处理装置，包括：

当接收到任一通讯方针对任一音频消息的音频转换请求时，服务器确定与所述任一通讯方相关的未响应音频消息；

所述服务器分别获取所述任一音频消息和所述未响应音频消息对应的文字内容，并返回至所述任一通讯方。

20 根据本申请的第十方面，提出了一种音频消息的处理装置，包括：

当接收到用户针对任一音频消息发出的音频转换命令时，本端通讯设备分别确定所述任一音频消息对应的第一文字内容，以及所述任一音频消息之外的未响应音频消息对应的第二文字内容；

25 所述本端通讯设备分别将所述第一文字内容与所述任一音频消息、所述第二文字内容与所述未响应音频消息进行关联展示。

根据本申请的第十一方面，提出了一种音频消息的处理装置，包括：

确定单元，当接收到任一通讯方针对任一音频消息的音频转换请求

时，使服务器确定与所述任一通讯方相关的未响应音频消息；

返回单元，使所述服务器分别获取所述任一音频消息和所述未响应音频消息对应的文字内容，并返回至所述任一通讯方。

根据本申请的第十二方面，提出了一种音频消息的处理装置，包括：

- 5 确定单元，当接收到用户针对任一音频消息发出的音频转换命令时，使本端通讯设备分别确定所述任一音频消息对应的第一文字内容，以及所述任一音频消息之外的未响应音频消息对应的第二文字内容；

展示单元，使所述本端通讯设备分别将所述第一文字内容与所述任一音频消息、所述第二文字内容与所述未响应音频消息进行关联展示。

- 10 由以上技术方案可见，本申请通过预先对音频消息进行文字转换，使得用户存在音频转换需求时，能够立即反馈相应的文字内容，而无需在翻译过程中进行等待，有助于加快对用户需求的响应速度，从而提升用户的应用体验。

附图说明

- 15 图 1 是本申请一示例性实施例提供的一种基于服务器侧的音频消息的处理方法的流程图。

图 2 是本申请一示例性实施例之一提供的一种基于通讯设备侧的音频消息的处理方法的流程图。

- 20 图 3 是本申请一示例性实施例之二提供的一种基于通讯设备侧的音频消息的处理方法的流程图。

图 4 是本申请一示例性实施例提供的一种音频消息的处理方法的流程图。

图 5-8 是本申请一示例性实施例提供的一种基于接收方侧的通讯应用的界面示意图。

- 25 图 9 是本申请一示例性实施例提供的一种基于发送方侧的通讯应用的界面示意图。

图 10 是本申请一示例性实施例之三提供的一种基于通讯设备侧的

音频消息的处理方法的流程图。

图 11 是本申请一示例性实施例提供的另一种音频消息的处理方法的流程图。

5 图 12 是本申请一示例性实施例提供的又一种音频消息的处理方法的流程图。

图 13 是本申请一示例性实施例提供的又一种音频消息的处理方法的流程图。

图 14 是本申请一示例性实施例提供的另一种基于服务器侧的音频消息的处理方法的流程图。

10 图 15 是本申请一示例性实施例之四提供的一种基于通讯设备侧的音频消息的处理方法的流程图。

图 16 是本申请一示例性实施例提供的一种电子设备的结构示意图。

图 17 是本申请一示例性实施例提供的一种基于服务器侧的音频消息的处理装置的框图。

15 图 18 是本申请一示例性实施例之一提供的一种基于通讯设备侧的音频消息的处理装置的框图。

图 19 是本申请一示例性实施例之二提供的一种基于通讯设备侧的音频消息的处理装置的框图。

20 图 20 是本申请一示例性实施例之三提供的一种基于通讯设备侧的音频消息的处理装置的框图。

图 21 是本申请一示例性实施例之四提供的一种基于通讯设备侧的音频消息的处理装置的框图。

图 22 是本申请一示例性实施例之五提供的一种基于通讯设备侧的音频消息的处理装置的框图。

25 具体实施方式

当用户采用音频类型的通讯消息时，存在一定的场景限制。举例而言，当用户在会议中接收到音频类型的通讯消息时，除非用户佩戴有蓝

牙耳机或其他可穿戴设备，否则可能由于无法及时收听该通信消息而造成相关事件的贻误。

为了解决音频类型的通信消息存在的上述问题，相关技术中提出了对音频消息的文字转换，具体而言：用户在接收到音频类型通信消息时，
5 如果不方便接听，可以向服务器发起针对该通信消息的音频转换请求，则服务器可以主动识别音频数据，并将转换得到的文字内容返回给用户，便于用户阅读。

然而，服务器对通信消息的音频转换操作需要一定时间，使得用户在向服务器发出音频转换请求之后，需要等待较长时间才能够看到转换后的文字内容，一方面造成用户的长时间等待，增添了用户情绪的焦虑感，
10 另一方面造成用户长时间不回复的现象，使通信消息的发送方用户长时间无法得到反馈，不仅影响用户的应用体验，而且造成了用户之间的沟通效率的极大降低。

因此，本申请通过对音频消息的处理过程进行改进，以解决相关技术中的上述技术问题。为对本申请进行进一步说明，提供下列实施例：
15

图 1 是本申请一示例性实施例提供的一种基于服务器侧的音频消息的处理方法的流程图，如图 1 所示，该方法可以包括：

步骤 102，服务器识别通讯双方之间传输的通信消息的类型。

步骤 104，当任一通信消息的类型为音频类型时，所述服务器获取
20 所述任一通信消息，并预转换为相应的文字内容。

步骤 106，当确定任一通讯方存在对所述任一通信消息的转换需求时，所述服务器向所述任一通讯方发送所述文字内容。

在本实施例中，服务器可以主动判定通讯方对音频消息的转换需求；比如，当任一通讯方在通讯过程中属于预设通讯角色时，服务器可以判定该任一通讯方存在转换需求，并发送相应的文字内容。举例而言，
25 服务器可以预定义为默认接收方存在转换需求，从而只要存在音频消息时，服务器总是预先转换出相应的文字内容，并主动发送给接收方的通讯设备。

在该实施例中，通过由服务器的预转换处理，并将文字内容主动发送给通讯设备，使得相应的通讯方确实需要执行音频转换时，该通讯设备可以直接调取并展示出已经存储于本地的文字内容，而无需从服务器上实时下载，从而即便当时网络状况不佳，也不影响对音频消息的文字内容展示，即降低了对实时网络状况的需求。

在本实施例中，服务器可以根据通讯方的请求情况，判定其是否存在转换需求；比如，当接收到任一通讯方针对任一通讯消息的音频转换请求时，服务器可以判定该任一通讯方存在转换需求，并向该任一通讯方返回该任一通讯消息对应的预转换的文字内容。

在该实施例中，服务器仅在通讯方确实存在需求时，才返回相应的文字内容；通过对通讯方的真实需求的准确判断，可以减少服务器与通讯设备之间的交互次数，降低服务器与通讯设备之间的通讯数据量，这一方面有助于降低通讯设备的功耗，另一方面对于采用无线移动通讯网络的通讯设备而言，可以减少无线流量的消耗，避免给用户造成不必要的费用损失。

由上述实施例可知，在本申请的技术方案中，服务器可以在用户提出音频转换需求之前，主动且预先对音频消息进行转换并得到相应的文字内容，因而当服务器接收到来自用户的音频转换需求时，可以立即将文字内容返回至用户，而无需用户等待服务器对音频消息进行转换，极大地缩短了接收方用户的等待时间，也缩短了对端的发送方用户收到反馈的等待时间，从而不仅提升了通讯双方的用户体验，而且极大地提升通讯双方之间的通讯效率。

对应于图 1 所示的实施例，在用户采用的通讯设备处存在多种相应的实施例，下面进行举例说明：

图 2 是本申请一示例性实施例之一提供的一种基于通讯设备侧的音频消息的处理方法的流程图，如图 2 所示，该方法可以包括：

步骤 202，本端通讯设备在接收到用户发出的针对音频类型的任一通讯消息的音频转换命令时，向服务器发起相应的音频转换请求。

步骤 204，本端通讯设备接收到所述服务器返回的所述任一通讯消

息对应的文字内容，并与所述任一通讯消息进行关联展示；其中，所述文字内容由所述服务器在接收到所述音频转换请求之前主动预转换得到。

在本实施例中，本端通讯设备基于用户发出的音频转换命令，主动向服务器发起音频转换请求，以表明其对于音频转换的切实需求，并由服务器相应返回其所需的文字内容。

由上述实施例可知，在本申请的技术方案中，基于服务器主动且预先对音频消息的预转换处理，本端通讯设备基于用户的音频转换命令而向服务器发起音频转换请求时，可以立即从服务器处获得相应的文字内容，不需要等待服务器对音频消息进行实时转换，有助于提升通讯双方的用户体验，而且极大地提升通讯双方之间的通讯效率。

图 3 是本申请一示例性实施例之二提供的一种基于通讯设备侧的音频消息的处理方法的流程图，如图 3 所示，该方法可以包括：

步骤 302，本端通讯设备预获取音频类型的任一通讯消息对应的文字内容。

在本实施例中，本端通讯设备可以从服务器处预获取文字内容，该文字内容由该服务器预转换得到。其中，该文字内容可以由服务器主动推送至本端通讯设备；或者，本端通讯设备在确定与对端通讯设备之间传输的通讯消息的类型时，若确定任一通讯消息的类型为音频类型，则可以向服务器发起音频转换请求，以获得服务器预转换处理得到的文字内容。在该实施例中，通过利用服务器执行预转换处理，既能够充分利用服务器强大的处理能力，提升对音频消息的预转换处理的执行效率，又可以降低对本端通讯设备的处理性能需求和处理资源占用，从而降低本端通讯设备的功耗。

在本实施例中，本端通讯设备可以自行对任一通讯消息进行预转换处理，得到该文字内容；换言之，即本端通讯设备对音频消息的本地预转换处理。比如，本端通讯设备在确定与对端通讯设备之间传输的通讯消息的类型时，若确定任一通讯消息的类型为音频类型，则可以执行该本地预转换处理，以得到相应的文字内容。在该实施例中，通过采用本

地预转换处理，可以消除或降低对网络的需求，从而适用于更多应用场景。

步骤 304，当接收到用户发出的针对所述任一通讯消息的音频转换命令时，所述本端通讯设备示出预获取的所述文字内容。

5 由上述实施例可知，在本申请的技术方案中，本端通讯设备通过对文字内容的预获取，使得在用户发出音频转换命令时，本端通讯设备可以直接获取并展示出相应的文字内容，而无需用户在转换过程中进行等待，有助于提升通讯效率。同时，通过将文字内容预获取在本端通讯设备的本地，使得用户可以在发出音频转换命令时，不存在对网络环境的需求，那么即便本端通讯设备并未连接至网络，用户仍然可以查看到相应音频消息的文字内容，适合于用户在一些特殊场景下对于历史通讯消息的查阅。

10

下面结合通讯过程中涉及到的发送方、接收方和服务端之间的交互过程，对本申请的技术方案进行详细描述；其中，图 4 是本申请一示例性实施例提供的一种音频消息的处理方法的流程图，如图 4 所示，该方法可以包括以下步骤：

15

步骤 402，服务器获取通讯双方之间传输的通讯消息。

在本实施例中，通讯双方在本申请的技术方案中完全对等，即每个通讯方均可以作为图 4 所示的发送方或接收方；因而在图 4 所示的实施例中，实际上是针对通讯双方之间的任一次通讯过程，确定出相应的发送方和接收方，并用于对本申请的技术方案进行举例说明。

20

步骤 404，服务器对通讯消息进行类型识别。

在本实施例中，通讯消息可以包括很多类型，本申请中可以将任意包含音频数据的通讯消息判定为音频类型，即音频消息，比如语音消息、视频消息等；下面结合的通讯应用的界面示意图中，以基于语音消息的音频消息为例进行说明，但本申请并不对此进行限制。

25

步骤 406，服务器对音频类型的通讯消息（即音频消息）进行预转换处理，得到相应的文字内容。

在本实施例中，服务器可以采用相关技术中的任意方式，对音频消息进行预转换处理，并得到相应的文字内容。

需要说明的是，服务器在检测到某条通讯消息为音频类型之后，即可在任意恰当的时刻执行预转换处理，以得到相应的文字内容，只要能够确保该预转换处理在步骤 408 之前完成即可。换言之，服务器对音频消息的预转换处理，与用户对该音频消息发起的音频转换命令无关，该预转换处理是由服务器预先、主动完成的。

因此，当用户向服务器发起对音频消息的音频转换命令时，服务器可以立即将已经预转换得到的文字内容提供至该用户，而无需服务器实时执行消息转换，避免了通讯双方的长时间等待，有助于提升通讯效率。

步骤 408，服务器接收到接收方针对该音频消息的音频转换请求。

在本实施例中，与该音频消息相关的每个通讯方，比如图 4 所示的发送方、接收方等，均可以发出音频转换命令，并由相应的电子设备向服务器发起音频转换请求（也可以理解为发送方或接收方等用户向服务器发起音频转换请求）；此处以接收方发起音频转换请求为例进行说明。

假定用户“小白”与用户“小黑”之间实现通讯；其中，本申请并不限制两者采用的通讯应用的类型，该通讯应用可以为即时通讯应用，比如该即时通讯应用可以为企业即时通讯应用（Enterprise Instant Messaging，EIM），例如“钉钉（DING Talk）”等。如图 5 所示，假定用户“小白”向用户“小黑”发送了若干条音频消息，则用户“小黑”可以通过长按（或重压等其他触发方式）希望查看的音频消息，以调起图 6 所示的功能选项菜单，该功能选项菜单中包含“听筒播放”、“收藏”、“转文字”、“删除”等功能选项，则当用户“小黑”选取“转文字”功能选项后，可以判定为向电子设备发出了针对相应音频消息的音频转换命令，并由该电子设备向服务器发起相应的音频转换请求。

步骤 410，服务器确定其他音频消息的响应状态。

步骤 412，服务器将音频消息对应的文字内容发送至接收方。

步骤 414，接收方对接收到的文字内容进行展示。

在一示例性实施例中，当不包含上述的步骤 410 时，服务器可以直接确定出用户“小黑”在图 5 中选中的长度为 12s 的音频消息对应的文字内容，并将该文字内容返回至用户“小黑”，以展示于用户“小黑”。

5 用户“小黑”采用的电子设备在接收服务器返回的文字内容后，可以对相应的音频消息的展示区域进行扩展；其中，扩展后的展示区域被划分为第一区域和第二区域；该第一区域用于示出相应的音频消息、该第二区域用于示出该音频消息对应的文字内容。比如图 7 所示，假定总共包含三条音频消息，而用户“小黑”触发了最上方的一条音频消息，则该音频消息对应的展示区域（该展示区域可以为图 7 所示的“气泡框”形式；当然，本申请并不对此进行限制）可以向下方扩展，则扩展后的展示区域被划分为相当于第一区域的上侧区域，以及相当于第二区域的下侧区域，其中上侧区域用于展示该音频消息的示意性图标，而下侧区域用于展示该音频消息对应的文字内容，比如“我现在不方便打字，直接语音吧”等。当然，本领域技术人员还可以采用其他方式对扩展区域进行功能划分，本申请并不对此进行限制。

10 15

在另一示例性实施例中，本申请的实施例中可以包含上述的步骤 410；相应的，在本申请的技术方案中，服务器可以确定通讯双方对传输的通讯消息的响应状态；那么，针对上述的音频消息，当接收到任一通讯方针对该音频消息发起的音频转换请求时，若存在与该任一通讯方相关的其他消息的响应状态为未响应，且该其他消息为音频类型时，服务器在步骤 412 中除了返回上述音频消息对应的文字内容之外，还可以返回该其他消息对应的文字内容。当然，该其他消息对应的文字内容，也是由服务器主动、预先通过预转换处理而得到，并不需要通讯方等待服务器实时执行转换。

20

25 那么，如图 5 所示，当用户“小黑”仅针对第一条音频消息发起音频转换请求时，若同时存在第二条音频消息和第三条音频消息，且两者均为未响应状态，则无需用户“小黑”一一手动发起音频转换请求，服务器即可主动下发所有三条音频消息对应的文字内容；相应的，如图 8 所示，用户“小黑”采用的电子设备可以分别对三条音频消息的展示区

域进行扩展，并示出相应的文字内容，包括“我现在不方便打字，直接语音吧”、“关于上次的合同报价”、“再提高三个点”等，从而一方面可以简化用户“小黑”的触发操作（即发出音频转换命令，或发起音频转换请求），通过一次触发即可实现对所有未响应的音频消息的查看，
5 另一方面可以帮助用户“小黑”对多条未响应的音频消息进行同时查看，这相比于分别单独查看每一条音频消息对应的文字内容，显然具有更佳的可读性和阅读连贯性，便于用户“小黑”对用户“小白”的通讯意图的理解，有助于提升通讯效率。

在又一示例性实施例中，除了服务器通过步骤 410 等来确定每条通讯消息的响应状态之外，可以由接收方对每条通讯消息的响应状况进行确定和处理。比如，接收方采用的电子设备可以确定该接收方对已接收的音频类型的通讯消息的响应状态；其中，当接收到该接收方发出的针对任一音频消息的音频转换命令时，若存在除该任一音频消息之外的音频类型的未响应通讯消息，则该电子设备在向服务器发起的音频转换请求，
15 该音频转换请求不仅与该任一音频消息相关（即可以用于获取该任一音频消息对应的文字内容），还与其他未响应的通讯消息相关（即可以用于获取该其他的未响应的通讯消息对应的文字内容）。比如，当用户“小黑”在电子设备上触发图 5 中的第一条音频消息后，该电子设备检测到还存在第二条音频消息、第三条音频消息，且两条音频消息的响应状态均为未响应，
20 则该电子设备向服务器发起针对这三条音频消息的音频转换请求，从而同时获得服务器返回的这三条音频消息的文字内容，并通过如图 8 所示的方式进行展示，可参考上述实施例，此处不再赘述。

步骤 416，服务器将该音频消息的已响应状态告知发送方。

在本实施例中，如图 5 所示，可以通过在通讯消息附近展示一黑色圆点，以表示其处于未响应状态。当用户“小黑”通过触发第一条音频消息而发出相应的音频转换请求后，用户“小黑”的电子设备可以判定为该音频消息被响应，从而如图 7 所示消除了第一条音频消息附近的黑色圆点。
25

同时，如图 9 所示，用户“小白”发出每条通讯消息后，用户“小

白”的电子设备上分别在每条通讯消息附近标示出其响应状态,比如“已读”对应于已响应状态、“未读”对应于未响应状态。那么,服务器在接收到用户“小黑”针对第一条音频消息的音频转换请求,并将预转换的相应文字内容返回给用户“小黑”之后,可以判定为该第一条音频消息由未响应状态切换至已响应状态,从而向该已响应状态告知给作为发送方的用户“小白”,因而图9中的第一条音频消息附近标示出“已读”,而第二条、第三条音频消息附近仍然标示为“未读”。当然,对应于图8所示的实施例,当用户“小黑”虽然仅针对第一条音频消息发起音频转换请求,但是基于服务器向用户“小黑”返回了全部三条音频消息对应的文字内容时,服务器可以认为三条音频消息均对应于已响应状态,并告知给用户“小白”的电子设备,以使其在三条音频消息附近均标示“已读”。

图10是本申请一示例性实施例之三提供的一种基于通讯设备侧的音频消息的处理方法的流程图,如图10所示,该方法可以包括:

步骤1002,在生成音频类型的通讯消息的过程中,本端通讯设备依次确定已采集到的每个音频片段是否符合预设切分规则。

步骤1004,当任一音频片段符合所述预设切分规则时,所述本端通讯设备将所述任一音频片段实时切分并上传至服务器。

在本实施例中,服务器依次接收到该本端通讯设备按照预设规则实时切分并上传的音频片段,并分别将每个音频片段预转换为相应的文字片段;然后,服务器将所有文字片段依次拼接,得到整条音频消息对应的文字内容。

在本实施例中,切分规则可以采用多种形式,比如基于时间长度、基于音频片段的数据量等多个维度中的一个或多个维度相结合。举例而言,当采用基于时间长度的切分规则时,假定整条音频消息共12s,而预定义的切分时长为2s,则每当达到2s时即可执行实时切分操作,并将该2s的音频片段上传至服务器,且服务器可以随即执行预转换处理,得到相应的文字片段;那么,整条音频一共可以得到6个音频片段,以及相应的6个文字片段,然后由服务器将其整合拼接为对应于整条音频

消息对应的文字内容。

在本实施例中，通过由发送方的电子设备（即上述的本端通讯设备）对音频消息的实时切分与上传，使得发送方在输入该音频消息的同时，服务器能够几乎不存在延迟地获得相应的音频片段，并随即执行对各个音频片段的预转换处理，相比于将音频消息完成输入后完整地上传至服务器，可使服务器更为迅速地完成对音频消息的预转换处理并得到相应的文字内容，从而即便接收方在接收到该音频消息后马上发起音频转换请求，服务器也能够确保在接收到该音频转换请求之前完成预转换处理，从而在接收到音频转换请求后立即返回相应的文字内容，那么通讯双方在通讯过程中，既可以通过音频输入而避免手动打字时的低效率和误输入问题，还解决了音频转换为文字时的延迟等待问题，即同时兼顾了音频输入时的快捷方便和文字交流时的无延迟，有助于提升通讯双方之间的沟通效率。

图 11 是本申请一示例性实施例提供的另一种音频消息的处理方法的流程图，如图 11 所示，该方法可以包括以下步骤：

步骤 1102，服务器获取通讯双方之间传输的通讯消息。

步骤 1104，服务器对通讯消息进行类型识别。

步骤 1106，服务器对音频类型的通讯消息（即音频消息）进行预转换处理，得到相应的文字内容。

在本实施例中，步骤 1102-1106 可参考图 4 所示实施例中的步骤 402-406，此处不再赘述。

步骤 1108，服务器将音频消息对应的文字内容发送至接收方。

在本实施例中，服务器默认为接收方对所有音频消息均存在音频转换需求，因而不仅通过预转换处理得到所有音频消息对应的文字内容，而且主动将文字内容推送至接收方。

步骤 1110，接收方的通讯设备接收到该接收方针对该音频消息的音频转换命令。

步骤 1112，接收方的通讯设备确定其他音频消息的响应状态。

步骤 1114, 接收方的通讯设备对文字内容进行展示。

在本实施例中, 在接收方发起音频转换命令之前, 服务器已经预转换处理得到相应的文字内容, 并主动推送至该接收方的通讯设备上; 换言之, 可以认为接收方的通讯设备在接收到音频转换命令之前, 已经对
5 音频消息对应的文字内容进行了“预获取”。因此, 当接收方发起音频转换命令后, 该通讯设备可以立即获得并展示出相应的文字内容, 而无需接收方等待。

同时, 相比于图 4 所示的实施例, 本实施例通过将文字内容预获取至通讯设备的本地, 使得该通讯设备在接收到音频转换命令后, 直接从
10 本地调取相应的文字内容即可, 不存在对网络环境的需求。因此, 对于一些场景下, 比如用户希望在没有网络环境下, 对历史通讯消息中的音频消息进行文字转换时, 图 11 由于不需要网络支持而仍然可以满足用户需求。

在本实施例中, 与图 4 所示的步骤 410 相类似的, 除了接收方直接
15 发起音频转换命令的音频消息之外, 若存在其他处于未响应状态的音频消息, 该通讯设备可以一并对这些音频消息的文字内容进行展示, 此处不再赘述。

步骤 1116, 接收方的通讯设备将执行了文字内容展示的音频消息
20 标记为已响应状态, 将该已响应状态告知服务器, 并由服务器告知发送方。

在本实施例中, 通讯设备可以将音频消息的已响应状态添加至响应状态切换通知, 将该响应状态切换通知发送至服务器, 并由服务器将其转发至发送方, 从而在发送方的通讯设备上对相应的音频消息进行正确
25 标记。

图 12 是本申请一示例性实施例提供的又一种音频消息的处理方法的流程图, 如图 12 所示, 该方法可以包括以下步骤:

步骤 1202, 通讯双方执行通讯消息的收发操作。

步骤 1204, 接收方的通讯设备对通讯消息进行类型识别。

步骤 1206, 当识别出音频消息时, 接收方的通讯设备向服务器发起音频转换请求。

步骤 1208, 服务器对音频类型的通讯消息 (即音频消息) 进行预转换处理, 得到相应的文字内容。

5 步骤 1210, 服务器将音频消息对应的文字内容发送至接收方。

在本实施例中, 音频转换请求是由通讯设备主动向服务器发起的, 而并非基于接收方发出的音频转换命令而发起; 换言之, 在接收方切实发出音频转换命令之前, 该通讯设备通过主动向服务器发起音频转换请求, 使得服务器执行预转换处理并得到相应的文字内容, 即该通讯设备
10 实现了对音频消息对应的文字内容的“预获取”操作。因此, 当接收方发起音频转换命令后, 该通讯设备可以立即获得并展示出相应的文字内容, 而无需接收方等待。

同时, 相比于图 11 所示实施例, 本实施例中的通讯设备通过对通讯消息进行类型识别, 主动向服务器发起音频转换请求, 以触发服务器
15 执行预转换处理, 而非服务器自行启动预转换处理, 从而使得该通讯设备分担了“类型识别”功能的执行过程, 降低了服务器的处理负荷。

步骤 1212, 接收方的通讯设备接收到该接收方针对该音频消息的音频转换命令。

步骤 1214, 接收方的通讯设备确定其他音频消息的响应状态。

20 步骤 1216, 接收方的通讯设备对文字内容进行展示。

步骤 1218, 接收方的通讯设备将执行了文字内容展示的音频消息标记为已响应状态, 将该已响应状态告知服务器, 并由服务器告知发送方。

在本实施例中, 步骤 1212-1218 可参考图 11 所示实施例中的步骤
25 1110-1116, 此处不再赘述。

图 13 是本申请一示例性实施例提供的又一种音频消息的处理方法的流程图, 如图 13 所示, 该方法可以包括以下步骤:

步骤 1302, 通讯双方执行通讯消息的收发操作。

步骤 1304, 接收方的通讯设备对通讯消息进行类型识别。

步骤 1306, 当识别出音频消息时, 接收方的通讯设备对音频类型的通讯消息 (即音频消息) 进行预转换处理, 得到相应的文字内容。

在本实施例中, 接收方的通讯设备主动识别通讯消息的类型, 并在
5 确定为音频消息时, 还主动完成对该音频消息的预转换处理, 以得到相应的文字内容。那么, 当网络环境差或无网络时, 接收方的通讯设备仍然可以“预获取”音频消息的文字内容, 使得接收方发出音频转换命令时, 能够及时展示出该文字内容, 避免接收方等待。

当网络环境不稳定时, 通讯设备在接收到音频消息后, 若依赖于服
10 务器来执行预转换处理, 则由于网络环境不稳定而可能导致通讯设备无法顺利向服务器发起音频转换请求, 或者服务器无法将预转换处理的文字内容顺利发送至该通讯设备, 那么可能导致接收方在发起音频转换命令之前, 该通讯设备无法预获取到相应的文字内容, 造成接收方需要实时向服务器发起音频转换请求, 无疑增加了用户等待时间。

15 实际上, 当网络环境不稳定时, 通过本申请中任一实施例的预转换处理 (或预获取) 方案, 即无论对音频消息的预转换处理在服务器或通讯设备上执行, 均可以优化用户的使用体验。比如, 当预转换处理在服务器上实现时, 通过预先获得文字内容, 那么在用户发起音频转换命令之前, 服务器与通讯设备可以获得更多时间和机会来传输该文字内容,
20 避免用户实时请求转换时, 由于网络原因造成文字内容无法传输或反复出现传输失败的情况。

步骤 1308, 接收方的通讯设备接收到该接收方针对该音频消息的音频转换命令。

步骤 1310, 接收方的通讯设备确定其他音频消息的响应状态。

25 步骤 1312, 接收方的通讯设备对文字内容进行展示。

步骤 1314, 接收方的通讯设备将执行了文字内容展示的音频消息标记为已响应状态, 将该已响应状态告知服务器, 并由服务器告知发送方。

在本实施例中，步骤 1308-1314 可参考图 11 所示实施例中的步骤 1110-1116，此处不再赘述。

图 14 是本申请一示例性实施例提供的一种基于服务器侧的音频消息的处理方法的流程图，如图 14 所示，该方法应用于服务器，可以包括以下步骤：

步骤 1402，当接收到任一通讯方针对任一音频消息的音频转换请求时，服务器确定与所述任一通讯方相关的未响应音频消息。

步骤 1404，所述服务器分别获取所述任一音频消息和所述未响应音频消息对应的文字内容，并返回至所述任一通讯方。

在本实施例中，服务器在接收到针对任一音频消息的音频转换请求时，主动相关联的其他未响应音频消息，使得用户无需针对每一音频消息分别发起音频转换，即可获得所有未响应音频消息对应的文字内容，从而极大地简化了用户操作。尤其是，当用户不便于触发对音频消息的音频转换命令时，比如用户一只手提着重物、仅能够通过另一只手操作，通过本申请的技术方案，用户仅需要对一条音频消息发起音频转换命令，即可读取所有音频消息对应的文字内容；再者，当多条音频消息之间的内容关联性较大时，通过将多条音频消息的文字内容主动呈现给用户，便于用户将多条音频消息的内容和逻辑相互串通，有助于提升阅读和沟通效率。

在本实施例的一种情况下，服务器可以对所有音频消息进行预转换并得到相应的文字内容，则当接收到音频转换请求时，服务器只需分别查找到上述的任一音频消息和未响应音频消息对应的预转换的文字内容即可；该场景下的技术方案可参考图 4 所示实施例的步骤 410 等，此处不再赘述。

在本实施例的另一种情况下，服务器可以在接收到音频转换请求后，分别将任一音频消息和未响应音频消息分别实时转换为对应的文字内容，并返回给用户进行展示；其中，对于每一单独音频消息的转换处理，可以参考相关技术中的处理过程，此处不再赘述。

图 15 是本申请一示例性实施例之一提供的一种基于通讯设备侧的

音频消息的处理方法的流程图,如图 15 所示,该方法应用于通讯设备,可以包括以下步骤:

步骤 1502, 当接收到用户针对任一音频消息发出的音频转换命令时,本端通讯设备分别确定所述任一音频消息对应的第一文字内容,以及所述任一音频消息之外的未响应音频消息对应的第二文字内容。

步骤 1504, 所述本端通讯设备分别将所述第一文字内容与所述任一音频消息、所述第二文字内容与所述未响应音频消息进行关联展示。

在本实施例中,与图 14 所示实施例相类似的,由通讯设备在接收到音频转换命令时,除了该音频转换命令针对的任一音频消息,该通讯设备还主动确定出该音频转换命令未针对的未响应音频消息,并通过将两者分别对应的第一文字内容和第二文字内容进行展示,以便于简化用户操作,并有助于提升阅读和沟通效率,此处不再赘述。

一方面,从对音频消息的转换时机而言,通讯设备可以在接收到音频转换命令之前,预获取第一文字内容和第二文字内容,该过程可以参考图 3 所示实施例中的步骤 302,此处不再赘述;或者,通讯设备可以在接收到音频转换命令之后,实时获取第一文字内容和第二文字内容。

另一方面,无论是采用预获取或实时获取,通讯设备均可以通过下述任一方式获取第一文字内容和第二文字内容:

第一种方式下,通讯设备可以主动将任一音频消息和未响应音频消息转换为第一文字内容和第二文字内容;当通讯设备采用预转换的处理方式时,该过程与图 13 所示实施例中的步骤 1306 相似,此处不再赘述。

第二种方式下,通讯设备可以向服务器发起音频转换请求,以获得服务器返回的第一文字内容和第二文字内容。其中,第一文字内容和第二文字内容可以由服务器根据音频转换请求进行实时转换得到,即服务器在接收到音频转换请求后才执行音频转换操作,该过程与图 12 所示实施例中的步骤 1208 相似,此处不再赘述;或者,第一文字内容和第二文字内容也可以由服务器预转换得到,该过程与图 4 所示实施例中的步骤 406 相似,此处不再赘述。

图 16 示出了根据本申请的一示例性实施例的电子设备的示意结构图。请参考图 16，在硬件层面，该电子设备包括处理器、内部总线、网络接口、内存以及非易失性存储器，当然还可能包括其他业务所需要的硬件。处理器从非易失性存储器中读取对应的计算机程序到内存中然后运行，在逻辑层面上形成音频消息的处理装置。当然，除了软件实现方式之外，本申请并不排除其他实现方式，比如逻辑器件抑或软硬件结合的方式等等，也就是说以下处理流程的执行主体并不限于各个逻辑单元，也可以是硬件或逻辑器件。

在一实施例中，请参考图 17，在软件实施方式中，该音频消息的处理装置可以包括识别单元、预转换单元和发送单元。其中：

识别单元，使服务器识别通讯双方之间传输的通讯消息的类型；

预转换单元，当任一通讯消息的类型为音频类型时，使所述服务器获取所述任一通讯消息，并预转换为相应的文字内容；

发送单元，当确定任一通讯方存在对所述任一通讯消息的转换需求时，使所述服务器向所述任一通讯方发送所述文字内容。

可选的，所述发送单元具体用于：

当所述任一通讯方在通讯过程中属于预设通讯角色时，使所述服务器判定所述任一通讯方存在所述转换需求，并发送所述文字内容。

可选的，所述发送单元具体用于：

当接收到任一通讯方针对所述任一通讯消息的音频转换请求时，使所述服务器判定所述任一通讯方存在所述转换需求，并向所述任一通讯方返回所述任一通讯消息对应的预转换的所述文字内容。

可选的，还包括：

确定单元，使所述服务器确定通讯双方对传输的通讯消息的响应状态；

返回单元，当接收到任一通讯方针对所述任一通讯消息的音频转换请求时，若存在与所述任一通讯方相关的音频类型的未响应通讯消息，则使所述服务器还向所述任一通讯方返回所有音频类型的未响应通讯

消息对应的预转换的文字内容。

可选的，还包括：

判定单元，在向所述任一通讯方返回所述任一通讯消息对应的预转换的所述文字内容之后，使所述服务器判定所述任一通讯消息切换至已
5 响应状态；

告知单元，使所述服务器将所述已响应状态告知所述任一通讯消息的发送方。

可选的，所述预转换单元具体用于：

使所述服务器依次接收通讯方按照预设规则实时切分并上传的音频片段，并分别将每个音频片段预转换为相应的文字片段；
10

所述服务器将所有文字片段依次拼接，得到所述文字内容。

在一实施例中，请参考图 18，在软件实施方式中，该音频消息的处理装置可以包括请求单元和展示单元。其中：

请求单元，使本端通讯设备在接收到用户发出的针对音频类型的任一通讯消息的音频转换命令时，向服务器发起相应的音频转换请求；
15

展示单元，使本端通讯设备接收到所述服务器返回的所述任一通讯消息对应的文字内容，并与所述任一通讯消息进行关联展示；其中，所述文字内容由所述服务器在接收到所述音频转换请求之前主动预转换得到。

20 可选的，还包括：

确定单元，使所述本端通讯设备确定所述用户对已接收的音频类型的通讯消息的响应状态；

其中，当接收到所述用户发出的针对所述任一通讯消息的音频转换命令时，若存在所述任一通讯消息之外的音频类型的未响应通讯消息，
25 则所述音频转换请求还与所述未响应通讯消息相关。

可选的，还包括：

扩展单元，使所述本端通讯设备在接收所述服务器返回的文字内容

后，对相应的通讯消息的展示区域进行扩展；

其中，扩展后的展示区域被划分为第一区域和第二区域；所述第一用于示出相应的通讯消息、所述第二区域用于示出所述通讯消息对应的文字内容。

- 5 在一实施例中，请参考图 19，在软件实施方式中，该音频消息的处理装置可以包括预获取单元和展示单元。其中：

预获取单元，使本端通讯设备预获取音频类型的任一通讯消息对应的文字内容；

- 10 展示单元，当接收到用户发出的针对所述任一通讯消息的音频转换命令时，使所述本端通讯设备示出预获取的所述文字内容。

可选的，所述预获取单元具体用于：

使所述本端通讯设备从服务器处预获取所述文字内容，所述文字内容由所述服务器预转换得到；

- 15 或者，使所述本端通讯设备对所述任一通讯消息进行预转换处理，得到所述文字内容。

可选的，所述预获取单元具体用于：

使所述本端通讯设备接收到服务器推送的所述文字内容；

- 20 或者，使所述本端通讯设备在确定与对端通讯设备之间传输的通讯消息的类型时，若确定所述任一通讯消息的类型为音频类型，则预获取所述任一通讯消息对应的文字内容。

可选的，当接收到用户发出的针对所述任一通讯消息的音频转换命令时，若存在处于未响应状态的音频类型的其他通讯消息，则所述展示单元还使所述本端通讯设备分别示出所述其他通讯消息对应的预获取的文字内容。

- 25 可选的，还包括：

通知单元，使所述本端通讯设备在分别示出所述其他通讯消息对应的预获取的文字内容之后，向服务器发出对应于所述其他通讯消息的响

应状态切换通知，以由所述服务器将所述其他通讯消息的已响应状态告知对应的发送方。

在一实施例中，请参考图 20，在软件实施方式中，该音频消息的处理装置可以包括确定单元和处理单元。其中：

- 5 确定单元，在生成音频类型的通讯消息的过程中，使本端通讯设备依次确定已采集到的每个音频片段是否符合预设切分规则；

处理单元，当任一音频片段符合所述预设切分规则时，使所述本端通讯设备将所述任一音频片段实时切分并上传至服务器，以由所述服务器将所述任一音频片段预转换为相应的文字片段，且所有音频片段对应的文字片段由所述服务器依次拼接为所述通讯消息对应的文字内容。

在一实施例中，请参考图 21，在软件实施方式中，该音频消息的处理装置可以包括确定单元和返回单元。其中：

确定单元，当接收到任一通讯方针对任一音频消息的音频转换请求时，使服务器确定与所述任一通讯方相关的未响应音频消息；

- 15 返回单元，使所述服务器分别获取所述任一音频消息和所述未响应音频消息对应的文字内容，并返回至所述任一通讯方。

可选的，所述返回单元具体用于：

使所述服务器分别将所述任一音频消息和所述未响应音频消息转换为对应的文字内容；

- 20 或者，使所述服务器分别查找到所述任一音频消息和所述未响应音频消息对应的预转换的文字内容。

在一实施例中，请参考图 22，在软件实施方式中，该音频消息的处理装置可以包括确定单元和展示单元。其中：

- 25 确定单元，当接收到用户针对任一音频消息发出的音频转换命令时，使本端通讯设备分别确定所述任一音频消息对应的第一文字内容，以及所述任一音频消息之外的未响应音频消息对应的第二文字内容；

展示单元，使所述本端通讯设备分别将所述第一文字内容与所述任一音频消息、所述第二文字内容与所述未响应音频消息进行关联展示。

可选的，还包括：

预获取单元，在接收到所述音频转换命令之前，使所述本端通讯设备预获取所述第一文字内容和所述第二文字内容；

或者，实时获取单元，在接收到所述音频转换命令之后，使所述本
5 端通讯设备实时获取所述第一文字内容和所述第二文字内容。

可选的，还包括：

主动转换单元，使所述本端通讯设备主动将所述任一音频消息和所述未响应音频消息转换为所述第一文字内容和所述第二文字内容；

或者，请求单元，使所述本端通讯设备向服务器发起音频转换请求，
10 以获得所述服务器返回的所述第一文字内容和所述第二文字内容；其中，所述第一文字内容和所述第二文字内容由所述服务器根据所述音频转换请求进行实时转换得到，或者由所述服务器预转换得到。

在一个典型的配置中，计算设备包括一个或多个处理器（CPU）、输入/输出接口、网络接口和内存。

15 内存可能包括计算机可读介质中的非永久性存储器，随机存取存储器（RAM）和/或非易失性内存等形式，如只读存储器（ROM）或闪存（flash RAM）。内存是计算机可读介质的示例。

计算机可读介质包括永久性和非永久性、可移动和非可移动媒体可以由任何方法或技术来实现信息存储。信息可以是计算机可读指令、数据
20 结构、程序的模块或其他数据。计算机的存储介质的例子包括，但不限于相变内存（PRAM）、静态随机存取存储器（SRAM）、动态随机存取存储器（DRAM）、其他类型的随机存取存储器（RAM）、只读存储器（ROM）、电可擦除可编程只读存储器（EEPROM）、快闪记忆体或其他内存技术、只读光盘只读存储器（CD-ROM）、数字多功能光盘（DVD）或其他光学存
25 储、磁盒式磁带，磁带磁磁盘存储或其他磁性存储设备或任何其他传输介质，可用于存储可以被计算设备访问的信息。按照本文中的界定，计算机可读介质不包括暂存电脑可读媒体（transitory media），如调制的数据信号和载波。

还需要说明的是，术语“包括”、“包含”或者其任何其他变体意在涵盖非排他性的包含，从而使得包括一系列要素的过程、方法、商品或者设备不仅包括那些要素，而且还包括没有明确列出的其他要素，或者是还包括为这种过程、方法、商品或者设备所固有的要素。在没有更多限制的情况下，由语句“包括一个...”限定的要素，并不排除在包括所述要素的过程、方法、商品或者设备中还存在另外的相同要素。

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

权 利 要 求 书

1. 一种音频消息的处理方法，其特征在于，包括：

服务器识别通讯双方之间传输的通讯消息的类型；

- 5 当任一通讯消息的类型为音频类型时，所述服务器获取所述任一通讯消息，并预转换为相应的文字内容；

当确定任一通讯方存在对所述任一通讯消息的转换需求时，所述服务器向所述任一通讯方发送所述文字内容。

2. 根据权利要求 1 所述的方法，其特征在于，所述当确定任一通讯方存在对所述任一通讯消息的转换需求时，所述服务器向所述任一通讯方发送所述文字内容，包括：

当所述任一通讯方在通讯过程中属于预设通讯角色时，所述服务器判定所述任一通讯方存在所述转换需求，并发送所述文字内容。

3. 根据权利要求 1 所述的方法，其特征在于，所述当确定任一通讯方存在对所述任一通讯消息的转换需求时，所述服务器向所述任一通讯方发送所述文字内容，包括：

当接收到任一通讯方针对所述任一通讯消息的音频转换请求时，所述服务器判定所述任一通讯方存在所述转换需求，并向所述任一通讯方返回所述任一通讯消息对应的预转换的所述文字内容。

- 20 4. 根据权利要求 3 所述的方法，其特征在于，还包括：

所述服务器确定通讯双方对传输的通讯消息的响应状态；

- 当接收到任一通讯方针对所述任一通讯消息的音频转换请求时，若存在与所述任一通讯方相关的音频类型的未响应通讯消息，则所述服务器还向所述任一通讯方返回所有音频类型的未响应通讯消息对应的预转换的文字内容。

- 25 5. 根据权利要求 3 所述的方法，其特征在于，还包括：

在向所述任一通讯方返回所述任一通讯消息对应的预转换的所述文字内容之后，所述服务器判定所述任一通讯消息切换至已响应状态；

所述服务器将所述已响应状态告知所述任一通讯消息的发送方。

6. 根据权利要求 1 所述的方法，其特征在于，所述服务器获取所述任一通讯消息，并预转换为相应的文字内容，包括：

所述服务器依次接收通讯方按照预设规则实时切分并上传的音频
5 片段，并分别将每个音频片段预转换为相应的文字片段；

所述服务器将所有文字片段依次拼接，得到所述文字内容。

7. 一种音频消息的处理方法，其特征在于，包括：

本端通讯设备在接收到用户发出的针对音频类型的任一通讯消息的音频转换命令时，向服务器发起相应的音频转换请求；

10 本端通讯设备接收到所述服务器返回的所述任一通讯消息对应的文字内容，并与所述任一通讯消息进行关联展示；其中，所述文字内容由所述服务器在接收到所述音频转换请求之前主动预转换得到。

8. 根据权利要求 7 所述的方法，其特征在于，还包括：

所述本端通讯设备确定所述用户对已接收的音频类型的通讯消息
15 的响应状态；

其中，当接收到所述用户发出的针对所述任一通讯消息的音频转换命令时，若存在所述任一通讯消息之外的音频类型的未响应通讯消息，则所述音频转换请求还与所述未响应通讯消息相关。

9. 根据权利要求 7 所述的方法，其特征在于，还包括：

20 所述本端通讯设备在接收所述服务器返回的文字内容后，对相应的通讯消息的展示区域进行扩展；

其中，扩展后的展示区域被划分为第一区域和第二区域；所述第一用于示出相应的通讯消息、所述第二区域用于示出所述通讯消息对应的文字内容。

25 10. 一种音频消息的处理方法，其特征在于，包括：

本端通讯设备预获取音频类型的任一通讯消息对应的文字内容；

当接收到用户发出的针对所述任一通讯消息的音频转换命令时，所

述本端通讯设备示出预获取的所述文字内容。

11. 根据权利要求 10 所述的方法, 其特征在于, 所述本端通讯设备预获取音频类型的任一通讯消息对应的文字内容, 包括:

所述本端通讯设备从服务器处预获取所述文字内容, 所述文字内容
5 由所述服务器预转换得到;

或者, 所述本端通讯设备对所述任一通讯消息进行预转换处理, 得到所述文字内容。

12. 根据权利要求 10 所述的方法, 其特征在于, 所述本端通讯设备预获取音频类型的任一通讯消息对应的文字内容, 包括:

10 所述本端通讯设备接收到服务器推送的所述文字内容;

或者, 所述本端通讯设备在确定与对端通讯设备之间传输的通讯消息的类型时, 若确定所述任一通讯消息的类型为音频类型, 则预获取所述任一通讯消息对应的文字内容。

13. 根据权利要求 10 所述的方法, 其特征在于, 还包括:

15 当接收到用户发出的针对所述任一通讯消息的音频转换命令时, 若存在处于未响应状态的音频类型的其他通讯消息, 则所述本端通讯设备还分别示出所述其他通讯消息对应的预获取的文字内容。

14. 根据权利要求 13 所述的方法, 其特征在于, 还包括:

所述本端通讯设备在分别示出所述其他通讯消息对应的预获取的文字内容之后, 向服务器发出对应于所述其他通讯消息的响应状态切换通知, 以由所述服务器将所述其他通讯消息的已响应状态告知对应的发送方。
20

15. 一种音频消息的处理方法, 其特征在于, 包括:

在生成音频类型的通讯消息的过程中, 本端通讯设备依次确定已采集到的每个音频片段是否符合预设切分规则;
25

当任一音频片段符合所述预设切分规则时, 所述本端通讯设备将所述任一音频片段实时切分并上传至服务器, 以由所述服务器将所述任一

音频片段预转换为相应的文字片段，且所有音频片段对应的文字片段由所述服务器依次拼接为所述通讯消息对应的文字内容。

16. 一种音频消息的处理装置，其特征在于，包括：

识别单元，使服务器识别通讯双方之间传输的通讯消息的类型；

5 预转换单元，当任一通讯消息的类型为音频类型时，使所述服务器获取所述任一通讯消息，并预转换为相应的文字内容；

发送单元，当确定任一通讯方存在对所述任一通讯消息的转换需求时，使所述服务器向所述任一通讯方发送所述文字内容。

10 17. 根据权利要求 16 所述的装置，其特征在于，所述发送单元具体用于：

当所述任一通讯方在通讯过程中属于预设通讯角色时，使所述服务器判定所述任一通讯方存在所述转换需求，并发送所述文字内容。

18. 根据权利要求 16 所述的装置，其特征在于，所述发送单元具体用于：

15 当接收到任一通讯方针对所述任一通讯消息的音频转换请求时，使所述服务器判定所述任一通讯方存在所述转换需求，并向所述任一通讯方返回所述任一通讯消息对应的预转换的所述文字内容。

19. 根据权利要求 18 所述的装置，其特征在于，还包括：

20 确定单元，使所述服务器确定通讯双方对传输的通讯消息的响应状态；

返回单元，当接收到任一通讯方针对所述任一通讯消息的音频转换请求时，若存在与所述任一通讯方相关的音频类型的未响应通讯消息，则使所述服务器还向所述任一通讯方返回所有音频类型的未响应通讯消息对应的预转换的文字内容。

25 20. 根据权利要求 18 所述的装置，其特征在于，还包括：

判定单元，在向所述任一通讯方返回所述任一通讯消息对应的预转换的所述文字内容之后，使所述服务器判定所述任一通讯消息切换至已

响应状态;

告知单元,使所述服务器将所述已响应状态告知所述任一通讯消息的发送方。

21. 根据权利要求 16 所述的装置,其特征在于,所述预转换单元
5 具体用于:

使所述服务器依次接收通讯方按照预设规则实时切分并上传的音频片段,并分别将每个音频片段预转换为相应的文字片段;

所述服务器将所有文字片段依次拼接,得到所述文字内容。

22. 一种音频消息的处理装置,其特征在于,包括:

10 请求单元,使本端通讯设备在接收到用户发出的针对音频类型的任一通讯消息的音频转换命令时,向服务器发起相应的音频转换请求;

展示单元,使本端通讯设备接收到所述服务器返回的所述任一通讯消息对应的文字内容,并与所述任一通讯消息进行关联展示;其中,所述文字内容由所述服务器在接收到所述音频转换请求之前主动预转换
15 得到。

23. 根据权利要求 22 所述的装置,其特征在于,还包括:

确定单元,使所述本端通讯设备确定所述用户对已接收的音频类型的通讯消息的响应状态;

其中,当接收到所述用户发出的针对所述任一通讯消息的音频转换
20 命令时,若存在所述任一通讯消息之外的音频类型的未响应通讯消息,则所述音频转换请求还与所述未响应通讯消息相关。

24. 根据权利要求 22 所述的装置,其特征在于,还包括:

扩展单元,使所述本端通讯设备在接收所述服务器返回的文字内容后,对相应的通讯消息的展示区域进行扩展;

25 其中,扩展后的展示区域被划分为第一区域和第二区域;所述第一用于示出相应的通讯消息、所述第二区域用于示出所述通讯消息对应的文字内容。

25. 一种音频消息的处理装置，其特征在于，包括：

预获取单元，使本端通讯设备预获取音频类型的任一通讯消息对应的文字内容；

5 展示单元，当接收到用户发出的针对所述任一通讯消息的音频转换命令时，使所述本端通讯设备示出预获取的所述文字内容。

26. 根据权利要求 25 所述的装置，其特征在于，所述预获取单元具体用于：

使所述本端通讯设备从服务器处预获取所述文字内容，所述文字内容
10 由所述服务器预转换得到；

或者，使所述本端通讯设备对所述任一通讯消息进行预转换处理，
得到所述文字内容。

27. 根据权利要求 25 所述的装置，其特征在于，所述预获取单元具体用于：

使所述本端通讯设备接收到服务器推送的所述文字内容；

15 或者，使所述本端通讯设备在确定与对端通讯设备之间传输的通讯消息的类型时，若确定所述任一通讯消息的类型为音频类型，则预获取所述任一通讯消息对应的文字内容。

28. 根据权利要求 25 所述的装置，其特征在于，当接收到用户发出的针对所述任一通讯消息的音频转换命令时，若存在处于未响应状态的
20 的音频类型的其他通讯消息，则所述展示单元还使所述本端通讯设备分别示出所述其他通讯消息对应的预获取的文字内容。

29. 根据权利要求 28 所述的装置，其特征在于，还包括：

通知单元，使所述本端通讯设备在分别示出所述其他通讯消息对应的预获取的文字内容之后，向服务器发出对应于所述其他通讯消息的响
25 应状态切换通知，以由所述服务器将所述其他通讯消息的已响应状态告知对应的发送方。

30. 一种音频消息的处理装置，其特征在于，包括：

确定单元，在生成音频类型的通讯消息的过程中，使本端通讯设备依次确定已采集到的每个音频片段是否符合预设切分规则；

处理单元，当任一音频片段符合所述预设切分规则时，使所述本端通讯设备将所述任一音频片段实时切分并上传至服务器，以由所述服务器将所述任一音频片段预转换为相应的文字片段，且所有音频片段对应的文字片段由所述服务器依次拼接为所述通讯消息对应的文字内容。

31. 一种音频消息的处理方法，其特征在于，包括：

当接收到任一通讯方针对任一音频消息的音频转换请求时，服务器确定与所述任一通讯方相关的未响应音频消息；

所述服务器分别获取所述任一音频消息和所述未响应音频消息对应的文字内容，并返回至所述任一通讯方。

32. 根据权利要求 31 所述的方法，其特征在于，所述服务器分别获取所述任一音频消息和所述未响应音频消息对应的文字内容，包括：

所述服务器分别将所述任一音频消息和所述未响应音频消息转换为对应的文字内容；

或者，所述服务器分别查找到所述任一音频消息和所述未响应音频消息对应的预转换的文字内容。

33. 一种音频消息的处理方法，其特征在于，包括：

当接收到用户针对任一音频消息发出的音频转换命令时，本端通讯设备分别确定所述任一音频消息对应的第一文字内容，以及所述任一音频消息之外的未响应音频消息对应的第二文字内容；

所述本端通讯设备分别将所述第一文字内容与所述任一音频消息、所述第二文字内容与所述未响应音频消息进行关联展示。

34. 根据权利要求 33 所述的方法，其特征在于，还包括：

在接收到所述音频转换命令之前，所述本端通讯设备预获取所述第一文字内容和所述第二文字内容；

或者，在接收到所述音频转换命令之后，所述本端通讯设备实时获

取所述第一文字内容和所述第二文字内容。

35. 根据权利要求 33 所述的方法，其特征在于，所述本端通讯设备通过下述任一方式获取所述第一文字内容和所述第二文字内容：

所述本端通讯设备主动将所述任一音频消息和所述未响应音频消息转换为所述第一文字内容和所述第二文字内容；

或者，所述本端通讯设备向服务器发起音频转换请求，以获得所述服务器返回的所述第一文字内容和所述第二文字内容；其中，所述第一文字内容和所述第二文字内容由所述服务器根据所述音频转换请求进行实时转换得到，或者由所述服务器预转换得到。

36. 一种音频消息的处理装置，其特征在于，包括：

确定单元，当接收到任一通讯方针对任一音频消息的音频转换请求时，使服务器确定与所述任一通讯方相关的未响应音频消息；

返回单元，使所述服务器分别获取所述任一音频消息和所述未响应音频消息对应的文字内容，并返回至所述任一通讯方。

37. 根据权利要求 36 所述的装置，其特征在于，所述返回单元具体用于：

使所述服务器分别将所述任一音频消息和所述未响应音频消息转换为对应的文字内容；

或者，使所述服务器分别查找到所述任一音频消息和所述未响应音频消息对应的预转换的文字内容。

38. 一种音频消息的处理装置，其特征在于，包括：

确定单元，当接收到用户针对任一音频消息发出的音频转换命令时，使本端通讯设备分别确定所述任一音频消息对应的第一文字内容，以及所述任一音频消息之外的未响应音频消息对应的第二文字内容；

展示单元，使所述本端通讯设备分别将所述第一文字内容与所述任一音频消息、所述第二文字内容与所述未响应音频消息进行关联展示。

39. 根据权利要求 38 所述的装置，其特征在于，还包括：

预获取单元，在接收到所述音频转换命令之前，使所述本端通讯设备预获取所述第一文字内容和所述第二文字内容；

或者，实时获取单元，在接收到所述音频转换命令之后，使所述本端通讯设备实时获取所述第一文字内容和所述第二文字内容。

5 40. 根据权利要求 38 所述的装置，其特征在于，还包括：

主动转换单元，使所述本端通讯设备主动将所述任一音频消息和所述未响应音频消息转换为所述第一文字内容和所述第二文字内容；

10 或者，请求单元，使所述本端通讯设备向服务器发起音频转换请求，以获得所述服务器返回的所述第一文字内容和所述第二文字内容；其中，
所述第一文字内容和所述第二文字内容由所述服务器根据所述音频转换请求进行实时转换得到，或者由所述服务器预转换得到。

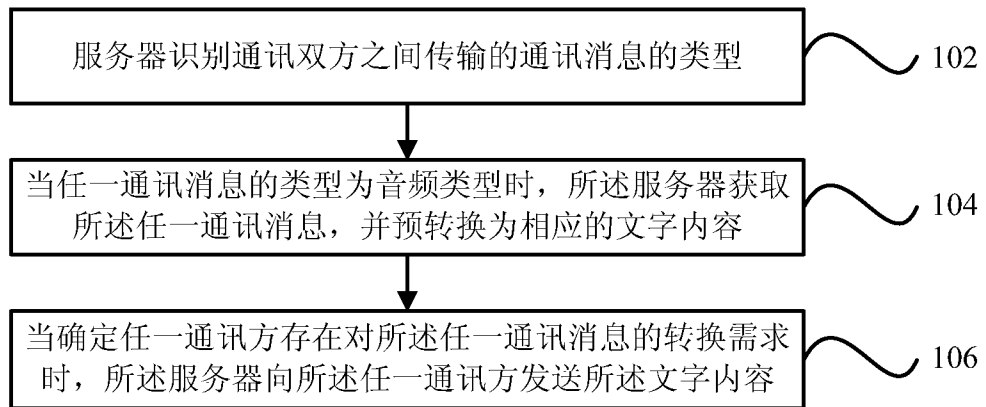


图 1

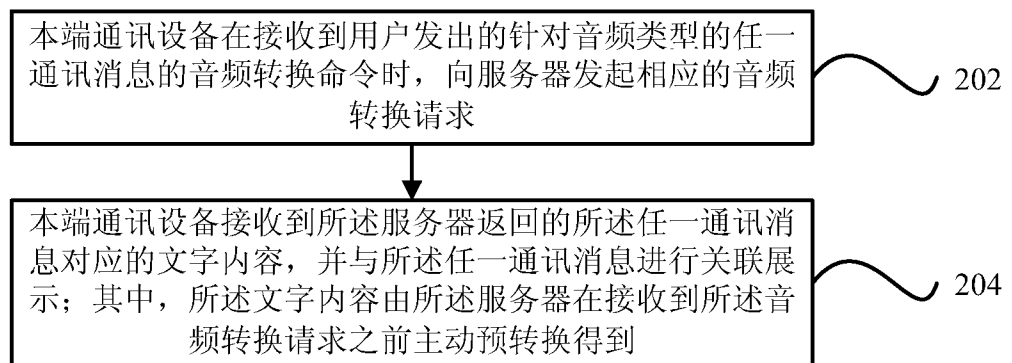


图 2

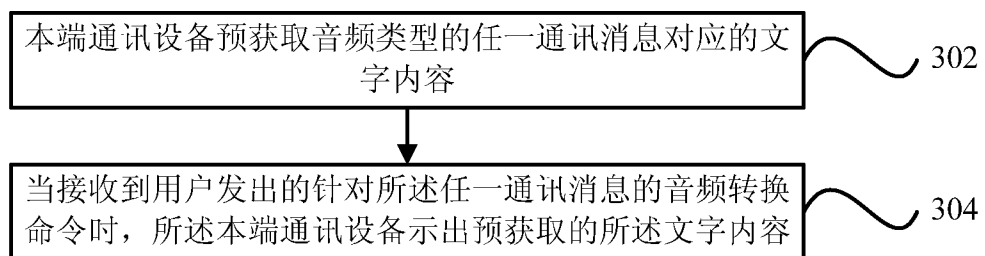


图 3

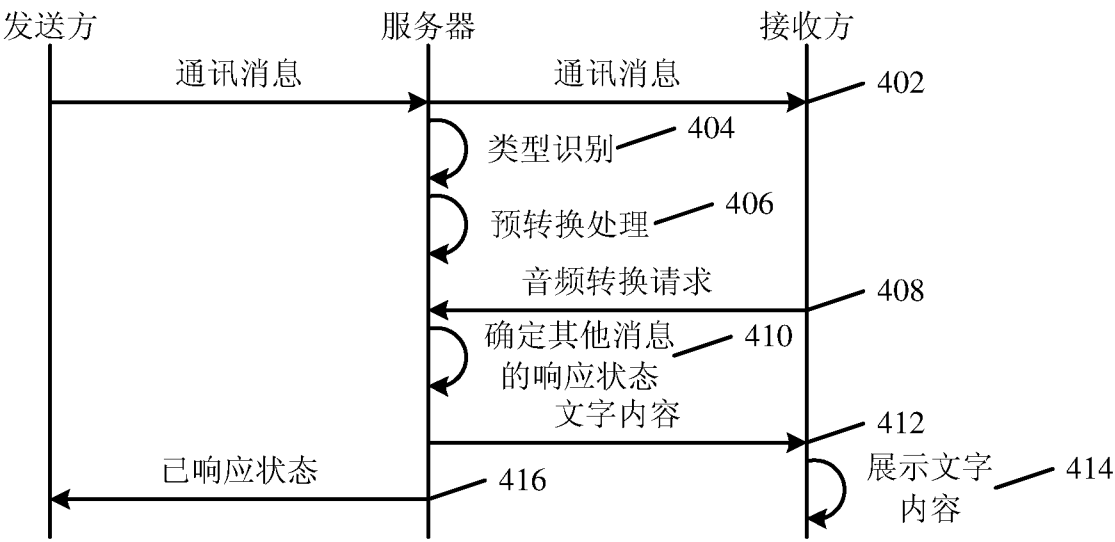


图 4

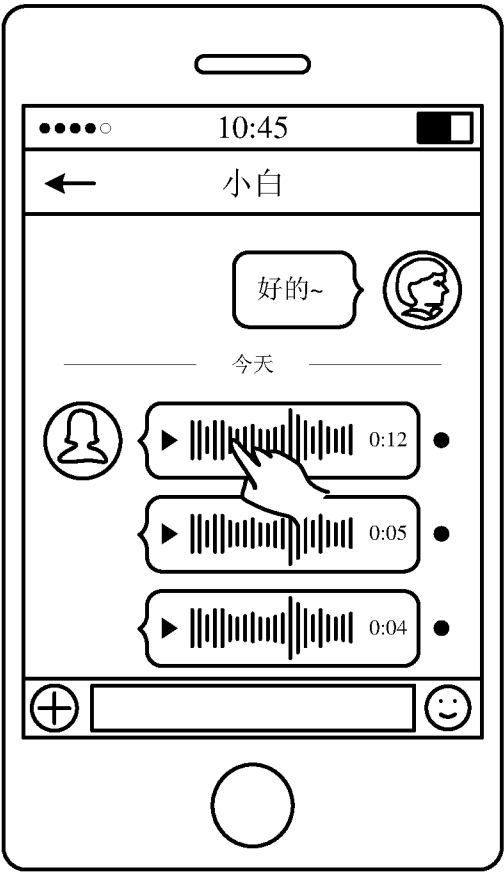


图 5

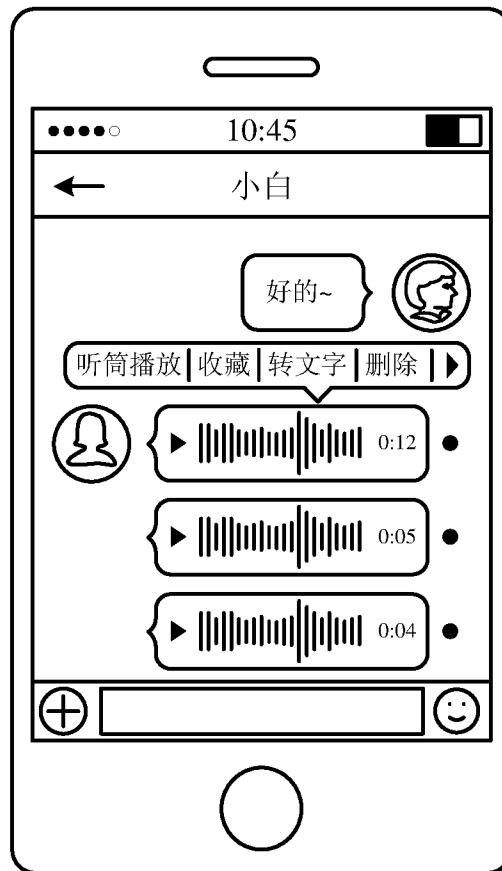


图 6



图 7

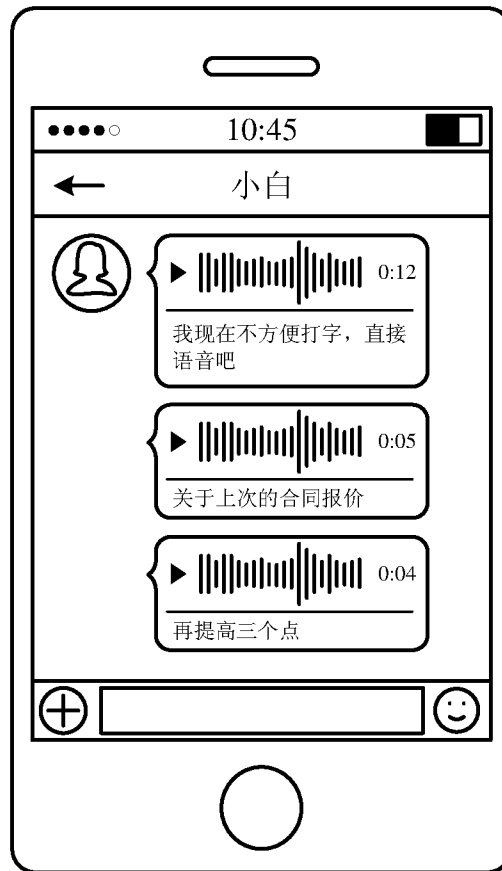


图 8

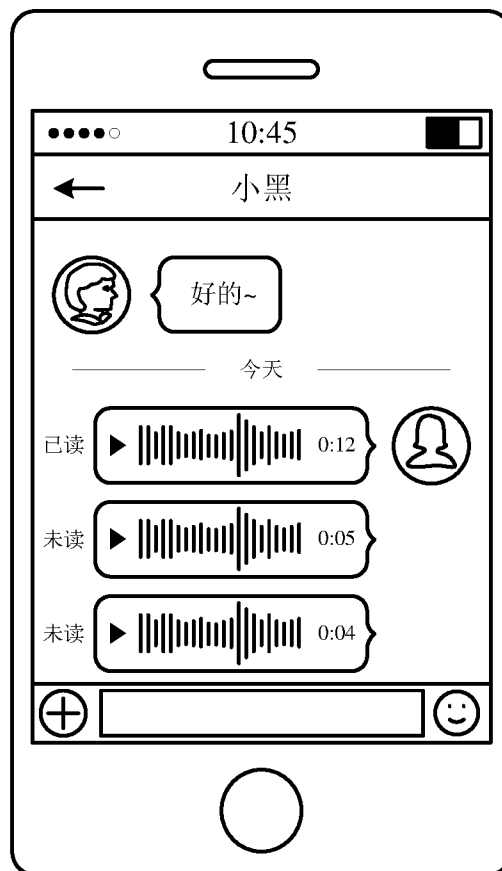


图 9

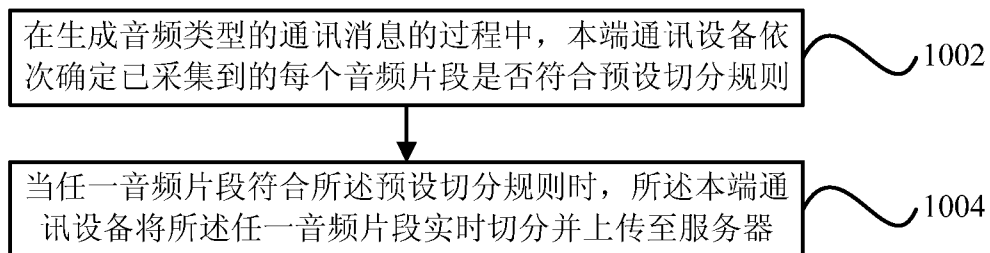


图 10

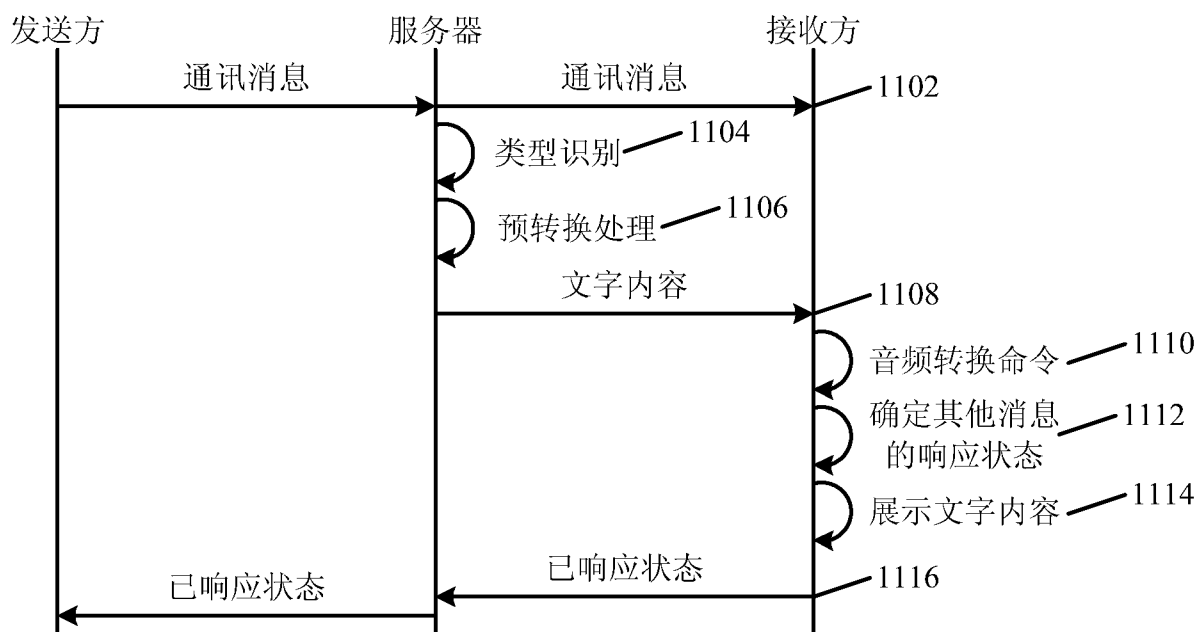


图 11

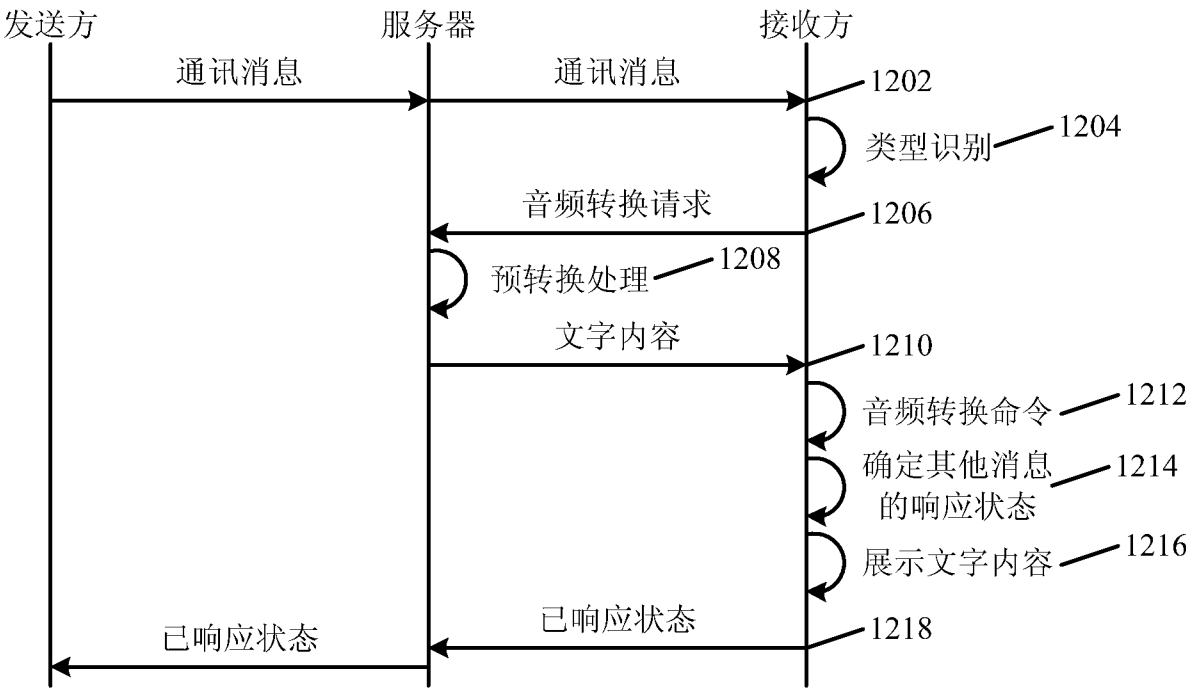


图 12

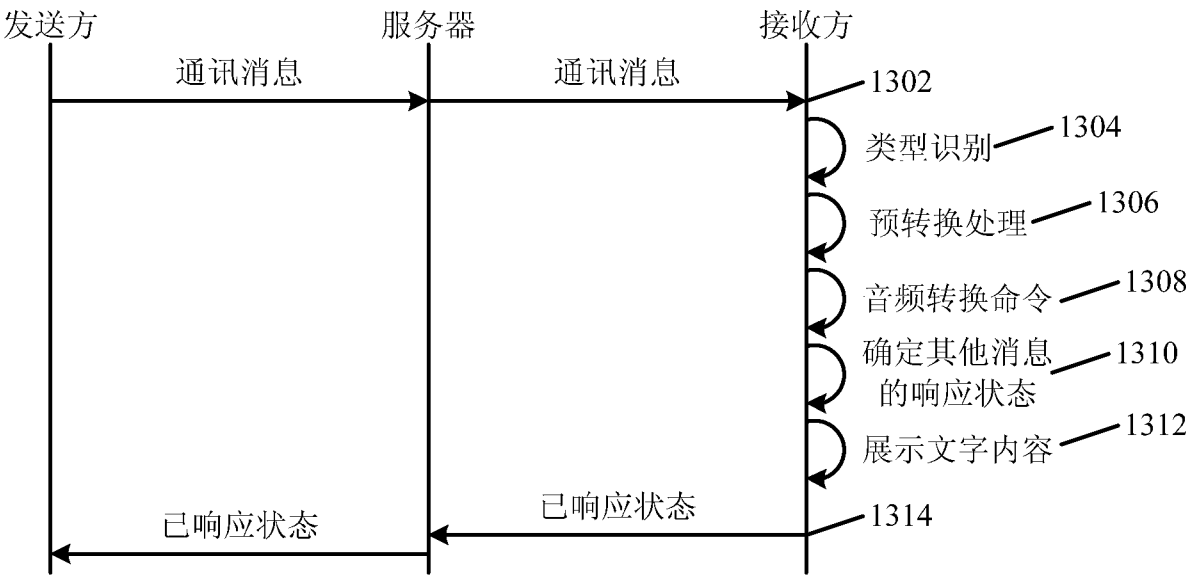


图 13

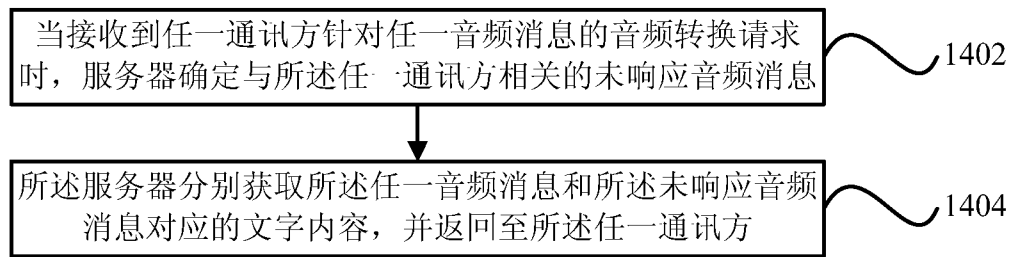


图 14

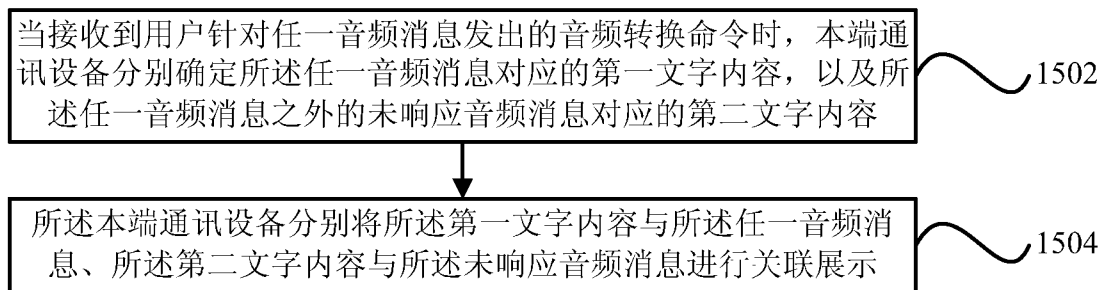


图 15

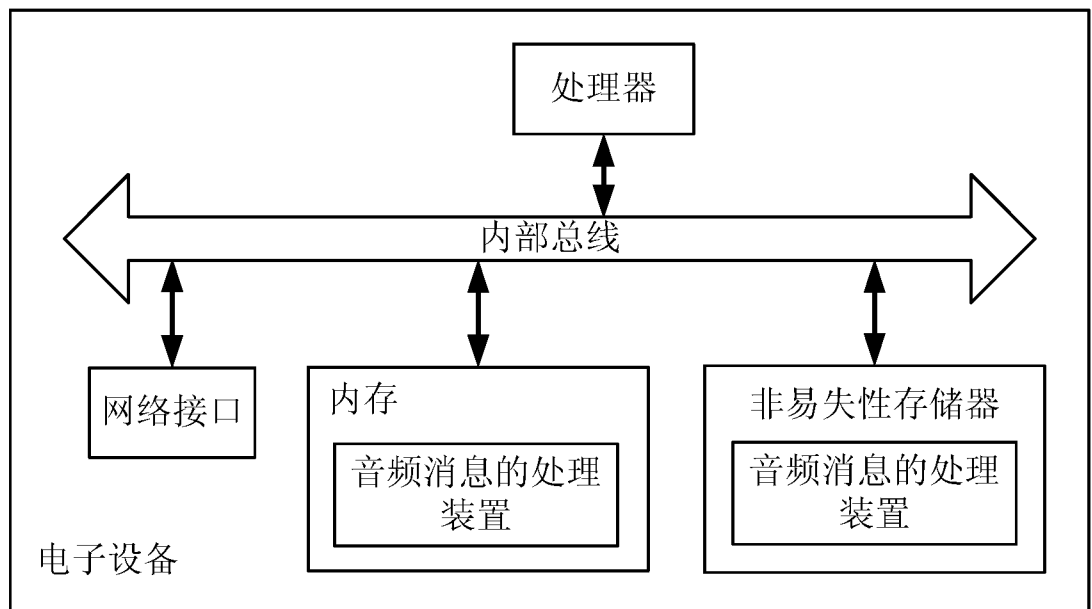


图 16

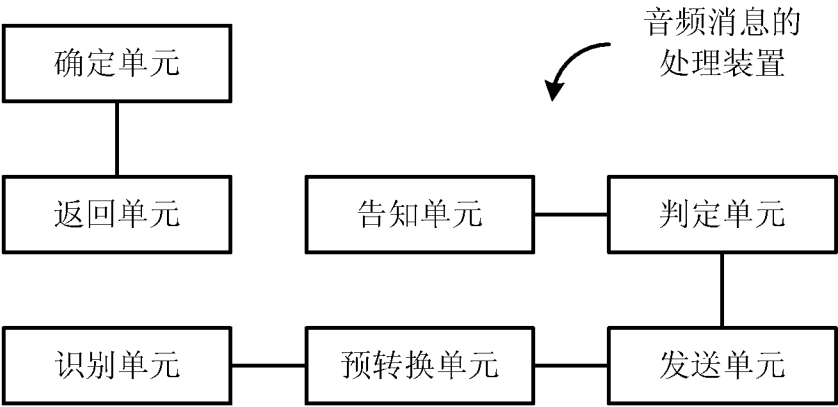


图 17

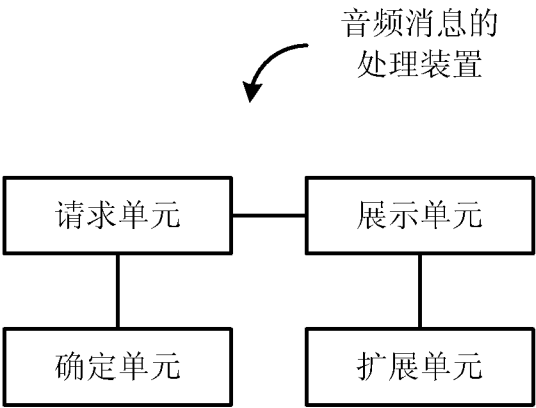


图 18

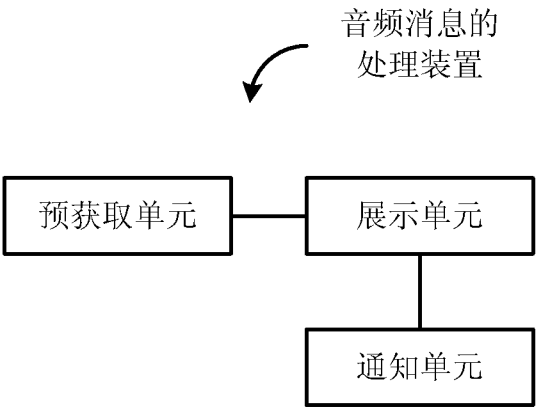


图 19

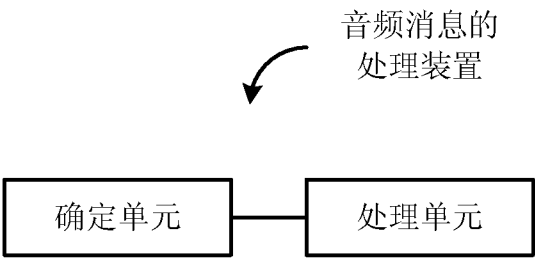


图 20

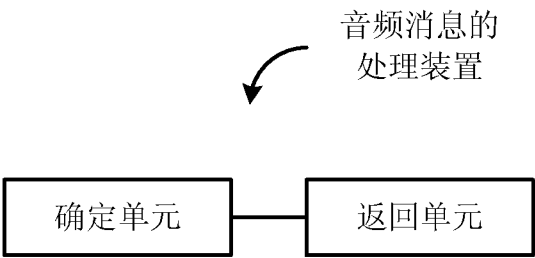


图 21

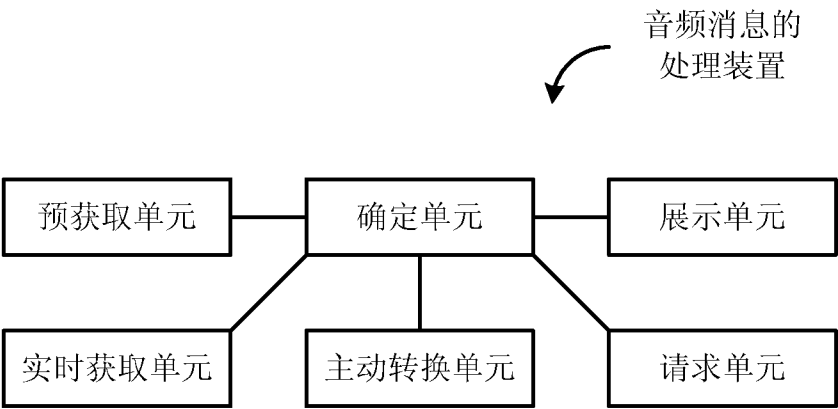


图 22

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2017/077257

A. CLASSIFICATION OF SUBJECT MATTER

G10L 21/10 (2013.01) i; H04W 4/18 (2009.01) 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)

G10L 21/-; G10L 15/-; H04W 4/-

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)

CNPAT, WPI, EPODOC, CNKI: character, advance, pre-conversion, pre-recognition, segmentation, cut, splice, combination, fragment, information, speech+, voice?, sound?, audio+, text+, convert+, recogni+, server?, message?, unrespond+, unrespons+, unread+

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 105869654 A (ALIBABA GROUP HOLDING LIMITED), 17 August 2016 (17.08.2016), claims 1-40	1-40
X	CN 103632670 A (QINGDAO INTWORK NETWORK TECHNOLOGY CO., LTD.), 12 March 2014 (12.03.2014), description, paragraphs [0024]-[0035], and figure 1	1-3, 10-12, 16-18, 25-27
Y	CN 103632670 A (QINGDAO INTWORK NETWORK TECHNOLOGY CO., LTD.), 12 March 2014 (12.03.2014), description, paragraphs [0024]-[0035], and figure 1	5-7, 9, 20-22, 24
Y	CN 1798220 A (YINGBAODA CO., LTD.), 05 July 2006 (05.07.2006), description, page 8, lines 6-31, and figure 3	5, 20
Y	CN 105162836 A (BAIDU ONLINE NETWORK TECHNOLOGY (BEIJING) CO., LTD.), 16 December 2015 (16.12.2015), description, paragraphs [0091], [0115] and [0116], and figure 6	6-7, 9, 21-22, 24
X	CN 105162836 A (BAIDU ONLINE NETWORK TECHNOLOGY (BEIJING) CO., LTD.), 16 December 2015 (16.12.2015), description, paragraphs [0091], [0115] and [0116], and figure 6	15, 30

☒ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents:	"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
"A" document defining the general state of the art which is not considered to be of particular relevance	"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
"E" earlier application or patent but published on or after the international filing date	"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
"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)	"&" document member of the same patent family
"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	

Date of the actual completion of the international search 01 June 2017 (01.06.2017)	Date of mailing of the international search report 22 June 2017 (22.06.2017)
Name and mailing address of the ISA/CN: State Intellectual Property Office of the P. R. China No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing 100088, China Facsimile No.: (86-10) 62019451	Authorized officer KANG, Dandan Telephone No.: (86-10) 62413593

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2017/077257**C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT**

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 7136462 B2 (LUCENT TECHNOLOGIES INC.), 14 November 2006 (14.11.2006), abstract, claim 1, description, paragraphs [0009]-[0023], and figures 1 and 2	1-3, 10-12, 16-18, 25-27
X	KR 20020028501 A (DREAMLINE CORP.), 17 April 2002 (17.04.2002), abstract, description, paragraphs [0016]-[0043], and figures 1 and 2	1-3, 10-12, 16-18, 25-27
A	CN 104700836 A (ALIBABA GROUP HOLDING LIMITED), 10 June 2015 (10.06.2015), the whole document	1-40

INTERNATIONAL SEARCH REPORT

International application No.

International application No.

PCT/CN2017/077257**Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)**

This International Searching Authority found multiple inventions in this international application, as follows:

[1] The claims include three inventions as claimed in three groups of independent claims, which respectively are:

[2] (1) claims 1, 16, 7, 22, 10 and 25

[3] (2) claims 15 and 30

[4] (3) claims 31 and 33

[5] The corresponding technical feature between claim groups (1) and (2) merely is converting an acquired audio communication message to text; however, the above-mentioned corresponding technical feature is a technical means known in the art; the above-mentioned two inventions do not have the same or corresponding special technical features which underlie the contribution of the invention over the prior art, are not technically linked, and do not fall within a single general inventive concept, and accordingly do not satisfy the requirement of unity of invention or the requirements of PCT Rule 13.1.

[6] The corresponding technical feature between claim groups (1) and (3) merely is acquiring corresponding text converted from the audio communication message; however, the above-mentioned corresponding technical feature is a technical means known in the art; the above-mentioned two inventions do not have the same or corresponding special technical features which underlie the contribution of the invention over the prior art, are not technically linked, and do not fall within a single general inventive concept, and accordingly do not satisfy the requirement of unity of invention or the requirements of PCT Rule 13.1.

[7] The corresponding technical feature between claim groups (2) and (3) merely is converting an acquired audio communication message to text; however, the above-mentioned corresponding technical feature is a technical means known in the art; the above-mentioned two inventions do not have the same or corresponding special technical features which underlie the contribution of the invention over the prior art, are not technically linked, and do not fall within a single general inventive concept, and accordingly do not satisfy the requirement of unity of invention or the requirements of PCT Rule 13.1.

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☒ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on protest

- ☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- ☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- ☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2017/077257

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 105869654 A	17 August 2016	None	
CN 103632670 A	12 March 2014	None	
CN 1798220 A	05 July 2006	None	
CN 105162836 A	16 December 2015	None	
US 7136462 B2	14 November 2006	US 2005013419 A1	20 January 2005
KR 20020028501 A	17 April 2002	None	
CN 104700836 A	10 June 2015	WO 2015089103 A1	18 June 2015
		TW 201523581 A	16 June 2015
		HK 1206862 A1	15 January 2016
		US 2015162003 A1	11 June 2015

国际检索报告

国际申请号

PCT/CN2017/077257

A. 主题的分类

G10L 21/10(2013.01)i; H04W 4/18(2009.01)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)

G10L 21/-; G10L 15/-; H04W 4/-

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))

CNPAT, WPI, EPODOC, CNKI: 音频, 语音, 声音, 话音, 文字, 文本, 转换, 识别, 预先, 提前, 预转换, 预识别, 服务器, 切分, 切割, 分割, 拼接, 组合, 片段, 片断, 未读, 未响应, 消息, 信息, speech+, voice?, sound?, audio+, text+, convert+, recogni+, server?, message?, unrespond+, unrespons+, unread+

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 105869654 A (阿里巴巴集团控股有限公司) 2016年 8月 17日 (2016 - 08 - 17) 权利要求1-40	1-40
X	CN 103632670 A (青岛英特沃克网络科技有限公司) 2014年 3月 12日 (2014 - 03 - 12) 说明书第[0024]-[0035]段, 图1	1-3, 10-12, 16-18, 25-27
Y	CN 103632670 A (青岛英特沃克网络科技有限公司) 2014年 3月 12日 (2014 - 03 - 12) 说明书第[0024]-[0035]段, 图1	5-7, 9, 20-22, 24
Y	CN 1798220 A (英保达股份有限公司) 2006年 7月 5日 (2006 - 07 - 05) 说明书第8页第6-31行, 图3	5, 20
Y	CN 105162836 A (百度在线网络技术北京有限公司) 2015年 12月 16日 (2015 - 12 - 16) 说明书第[0091], [0115]-[0116]段, 图6	6-7, 9, 21-22, 24
X	CN 105162836 A (百度在线网络技术北京有限公司) 2015年 12月 16日 (2015 - 12 - 16) 说明书第[0091], [0115]-[0116]段, 图6	15, 30

☒ 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

2017年 6月 1日

国际检索报告邮寄日期

2017年 6月 22日

ISA/CN的名称和邮寄地址

中华人民共和国国家知识产权局(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

受权官员

康丹丹

电话号码 (86-10)62413593

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
X	US 7136462 B2 (LUCENT TECHNOLOGIES INC.) 2006年 11月 14日 (2006 - 11 - 14) 摘要，权利要求1，说明书第[0009]-[0023]段，图1-2	1-3, 10-12, 16-18, 25-27
X	KR 20020028501 A (DREAMLINE CORP.) 2002年 4月 17日 (2002 - 04 - 17) 摘要，说明书第[0016]-[0043]段，图1-2	1-3, 10-12, 16-18, 25-27
A	CN 104700836 A (阿里巴巴集团控股有限公司) 2015年 6月 10日 (2015 - 06 - 10) 全文	1-40

第III栏 缺乏发明单一性的意见(续第1页第3项)

本国际检索单位在该国际申请中发现多项发明，即：

[1] 权利要求书包括三组独立权利要求所要求保护的三项发明，分别为：

[2] (1) 权利要求1, 16、7, 22、10, 25

[3] (2) 权利要求15, 30

[4] (3) 权利要求31, 33

[5] 第(1)组和第(2)组权利要求之间相应的技术特征仅为对获取的音频通讯消息转换为文字，然而上述相应的技术特征为本领域公知的技术手段，则上述2项发明不具有相同或相应的体现发明对现有技术作出贡献的特定技术特征，不存在技术关联，不属于一个总的发明构思，因而不满足发明单一性的要求，不符合PCT细则13.1的规定。

[6] 第(1)组和第(3)组权利要求之间相应的技术特征仅为获取音频通讯消息转换的对应文字，然而上述相应的技术特征为本领域公知的技术手段，则上述2项发明不具有相同或相应的体现发明对现有技术作出贡献的特定技术特征，不存在技术关联，不属于一个总的发明构思，因而不满足发明单一性的要求，不符合PCT细则13.1的规定。

[7] 第(2)组和第(3)组权利要求之间相应的技术特征仅为对获取的音频通讯消息转换为文字，然而上述相应的技术特征为本领域公知的技术手段，则上述2项发明不具有相同或相应的体现发明对现有技术作出贡献的特定技术特征，不存在技术关联，不属于一个总的发明构思，因而不满足发明单一性的要求，不符合PCT细则13.1的规定。

1. ☐ 由于申请人按时缴纳了被要求缴纳的全部附加检索费，本国际检索报告涉及全部可作检索的权利要求。

2. ☒ 由于无需付出有理由要求附加费的劳动即能对全部可检索的权利要求进行检索，本单位未通知缴纳任何加费。

3. ☐ 由于申请人仅按时缴纳了部分被要求缴纳的附加检索费，本国际检索报告仅涉及已缴费的那些权利要求，具体地说，是权利要求：

4. ☐ 申请人未按时缴纳被要求缴纳的附加检索费。因此，本国际检索报告仅涉及权利要求书中首先提及的发明；包含该发明的权利要求是：

对异议的意见 ☐ 申请人缴纳了附加检索费，同时提交了异议书，适用时，缴纳了异议费。

☐ 申请人缴纳了附加检索费，同时提交了异议书，但未在通知书规定的时间期限内缴纳异议费。

☐ 缴纳附加检索费时未提交异议书。

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2017/077257

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	105869654	A	2016年 8月 17日	无			
CN	103632670	A	2014年 3月 12日	无			
CN	1798220	A	2006年 7月 5日	无			
CN	105162836	A	2015年 12月 16日	无			
US	7136462	B2	2006年 11月 14日	US	2005013419	A1	2005年 1月 20日
KR	20020028501	A	2002年 4月 17日	无			
CN	104700836	A	2015年 6月 10日	WO	2015089103	A1	2015年 6月 18日
				TW	201523581	A	2015年 6月 16日
				HK	1206862	A1	2016年 1月 15日
				US	2015162003	A1	2015年 6月 11日

表 PCT/ISA/210 (同族专利附件) (2009年7月)