

(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
2019 年 1 月 24 日 (24.01.2019)



(10) 国际公布号
WO 2019/015505 A1

(51) 国际专利分类号:
G10L 15/26 (2006.01)

(21) 国际申请号: PCT/CN2018/095081

(22) 国际申请日: 2018 年 7 月 10 日 (10.07.2018)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201710592686.1 2017 年 7 月 19 日 (19.07.2017) 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: 徐冲 (XU, Chong); 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)。李威 (LI, Wei); 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: 北京三友知识产权代理有限公司 (BEIJING SANYOU INTELLECTUAL PROPERTY AGENCY LTD.); 中国北京市金融街 35 号国际企业大厦 A 座 16 层, Beijing 100033 (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,

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

(54) 发明名称: 信息处理方法、系统、电子设备、和计算机存储介质

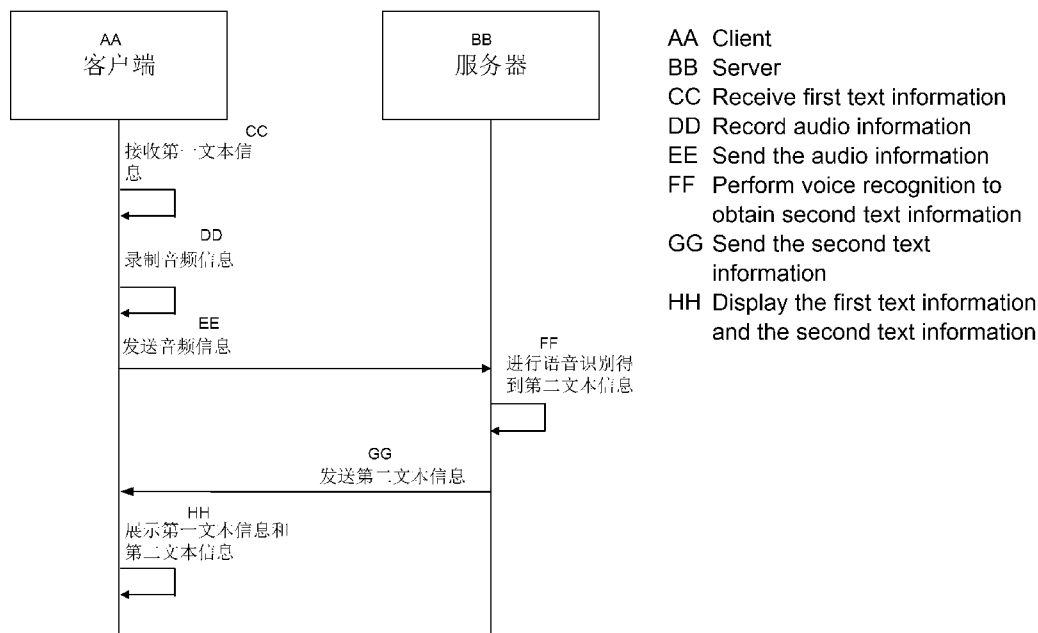


图 2

(57) Abstract: Disclosed are an information processing method and system, an electronic device and a computer storage medium. The method correspondingly displays, regarding content of a multi-person communication and dialogue, information recorded by a recording person and content obtained by means of voice recognition, such that the recording person can refine the recorded content.

(57) 摘要: 本说明书实施方式公开了一种信息处理方法、系统、电子设备、和计算机存储介质。所述方法针对多人沟通对话的内容, 将记录人员的记录信息, 与语音识别得到的内容, 进行对应展示。以便于记录人员可以完善记录的内容。

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, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, 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))。

信息处理方法、系统、电子设备、和计算机存储介质

本申请要求 2017 年 07 月 19 日递交的申请号为 201710592686.1、发明名称为“信息处理方法、系统、电子设备、和计算机存储介质”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

本说明书涉及计算机技术领域，特别涉及一种信息处理方法、系统、电子设备、和计算机存储介质。

10 背景技术

现实的生活中，人们会在一起沟通，讨论事项。在沟通过程中通常会指定人员记录沟通的事项形成记录。具体的，例如，在工作过程中，多人进行的会议，需要进行会议记录；法院进行庭审之后，形成庭审记录等。

传统，人们多采用手写的方式进行记录，由于手写速度较慢，在纸张上书写的记录后期可编辑性较差，使得人们逐渐多采用电脑制作记录。比如，在庭审过程中，会由书记员操作电脑制作庭审记录。

15

发明内容

本说明书实施方式的目的是提供一种信息处理方法、电子设备和计算机存储介质。

20 本说明书实施方式可以便于生成记录。

为实现上述目的，本说明书实施方式提供一种信息处理方法，所述方法包括：接收第一输入设备输入的第一文本信息；所述第一文本信息为根据语音生成；接收第二输入设备录制的音频信息；其中，所述音频信息为根据所述语音录制生成；对所述音频信息进行语音识别得到第二文本信息；展示所述第一文本信息和所述第二文本信息，其中，
25 所述第一文本信息和所述第二文本信息中内容存在对应关系。

25

本说明书实施方式还提供一种信息处理系统，所述信息处理系统包括：输入设备、音频采集终端、显示器和处理器；所述输入设备，用于接收用户输入的第一文本信息；所述第一文本信息为用户根据语音进行输入；所述音频采集终端，用于录制音频信息；其中，所述音频信息为根据所述语音录制生成；所述处理器，用于对所述音频信息进行
30 语音识别得到第二文本信息；所述显示器，用于展示所述第一文本信息和所述第二文本

30

信息，其中，所述第一文本信息和所述第二文本信息中内容存在对应关系。

本说明书实施方式还提供一种计算机存储介质，所述计算机存储介质存储有计算机程序指令，所述计算机程序指令被执行时实现：接收第一输入设备输入的第一文本信息，所述第一文本信息为根据语音生成；接收第二输入设备录制的音频信息，其中，所述音频信息为根据所述语音录制生成；对所述音频信息进行语音识别得到第二文本信息；展示所述第一文本信息和所述第二文本信息，其中，所述第一文本信息和所述第二文本信息中内容存在对应关系。

本说明书实施方式还提供一种信息处理方法，所述方法包括：接收第一输入设备输入的第一文本信息；所述第一文本信息为根据语音生成；接收第二输入设备录制的音频信息；其中，所述音频信息为根据所述语音录制生成；将所述音频信息或所述音频信息的表征信息发送给服务器，以用于所述服务器进行语音识别；接收所述服务器反馈的语音识别得到的第二文本信息；展示所述第一文本信息和所述第二文本信息，其中，所述第一文本信息和所述第二文本信息中内容存在对应关系。

本说明书实施方式还提供一种信息处理系统，所述信息处理系统包括：输入设备、音频采集终端、网络通信单元、显示器；所述输入设备，用于接收用户输入的第一文本信息；所述第一文本信息为用户根据语音进行输入；所述音频采集终端，用于录制音频信息；其中，所述音频信息为根据所述语音录制生成；所述网络通信单元，用于将所述音频信息或所述音频信息的表征信息发送给服务器，以用于所述服务器进行语音识别；接收所述服务器反馈的语音识别得到的第二文本信息；所述显示器，用于展示所述第一文本信息和所述第二文本信息，其中，所述第一文本信息和所述第二文本信息中内容存在对应关系。

本说明书实施方式还提供一种计算机存储介质，所述计算机存储介质存储有计算机程序指令，所述计算机程序指令被执行时实现：接收第一输入设备输入的第一文本信息，所述第一文本信息为根据语音生成；接收第二输入设备录制的音频信息，其中，所述音频信息为根据所述语音录制生成；将所述音频信息或所述音频信息的表征信息发送给服务器，以用于所述服务器进行语音识别；接收所述服务器反馈的语音识别得到的第二文本信息；展示所述第一文本信息和所述第二文本信息，其中，所述第一文本信息和所述第二文本信息中内容存在对应关系。

本说明书实施方式还提供一种信息处理方法，所述方法包括：接收客户端发送的音频信息或音频信息的表征信息；对所述音频信息或者所述表征信息进行语音识别得到第

二文本信息；将所述第二文本信息发送给所述客户端，以用于所述客户端将所述第二文本信息与第一客户端的第一文本信息进行展示；其中，所述第一文本信息和所述第二文本信息中内容存在对应关系。

本说明书实施方式还提供一种电子设备，所述电子设备包括：网络通信单元和处理器；所述网络通信单元，用于接收客户端发送的音频信息或音频信息的表征信息；将所述处理器提供的第二文本信息发送给所述客户端，以用于所述客户端将所述第二文本信息与第一客户端的第一文本信息进行展示；其中，所述第一文本信息和所述第二文本信息中内容存在对应关系；所述处理器，用于对所述音频信息或者所述表征信息进行语音识别得到第二文本信息。

本说明书实施方式还提供一种计算机存储介质，所述计算机存储介质中存储有计算机程序指令，所述计算机程序指令被执行时可以实现：接收客户端发送的音频信息或音频信息的表征信息；对所述音频信息或者所述表征信息进行语音识别得到第二文本信息；将所述第二文本信息发送给所述客户端，以用于所述客户端将所述第二文本信息与第一客户端的第一文本信息进行展示，其中，所述第一文本信息和所述第二文本信息中内容存在对应关系。

本说明书实施方式还提供一种信息处理方法，包括：接收第一文本信息；所述第一文本信息为根据语音生成；接收音频信息；其中，所述音频信息为根据所述语音录制生成；对所述音频信息进行语音识别得到第二文本信息；根据所述第一文本信息和所述第二文本信息之间的对应关系进行排版，所述排版后的所述第一文本信息和所述第二文本信息用于展示。

本说明书实施方式还提供一种电子设备，所述电子设备包括处理器，所述处理器用于接收第一文本信息，所述第一文本信息为根据语音生成；接收音频信息，其中，所述音频信息为根据所述语音录制生成；对所述音频信息进行语音识别得到第二文本信息；根据所述第一文本信息和所述第二文本信息之间的对应关系进行排版，所述排版后的所述第一文本信息和所述第二文本信息用于展示。

本说明书实施方式还提供一种计算机存储介质，所述计算机存储介质中存储有计算机程序指令，所述计算机程序指令被执行时实现：接收第一文本信息，所述第一文本信息为根据语音生成；接收音频信息，其中，所述音频信息为根据所述语音录制生成；对所述音频信息进行语音识别得到第二文本信息；根据所述第一文本信息和所述第二文本信息之间的对应关系进行排版，所述排版后的所述第一文本信息和所述第二文本信息用

于展示。

本说明书实施方式还提供一种信息处理方法，所述方法包括：接收第一输入设备输入的第一文本信息；所述第一文本信息为根据语音生成；接收第二输入设备录制的音频信息；其中，所述音频信息为根据所述语音录制生成；对所述音频信息进行识别得到第二文本信息；在第一区域展示所述第一文本信息；在第二区域展示所述第二文本信息，其中，所述第一文本信息和所述第二文本信息中内容存在对应关系，所述第一区域和第二区域位于相同的界面中。

由以上本说明书实施方式提供的技术方案可见，通过针对同一个语音场景中，可以分别采用人工输入和录制音频后语音识别的方式得到第一文本信息和第二文本信息。利用第二文本信息可以较为全面的记录语音内容的特点，对第一文本信息进行修改。如此，使得第一文本信息可以较为全面准确，且可以突出重点，以及概要的记录语音的内容。第一文本信息相较于第二文本信息可以具有较小的篇幅，重点相对突出，可以节省阅读者的阅读时间。

附图说明

为了更清楚地说明本说明书实施方式或现有技术中的技术方案，下面将对实施方式或现有技术描述中所需要使用的附图作简单地介绍，显而易见地，下面描述中的附图仅仅是本说明书中记载的一些实施方式，对于本领域普通技术人员来讲，在不付出创造性劳动性的前提下，还可以根据这些附图获得其他的附图。

图 1 为本说明书实施方式提供的一种应用场景示意图；

图 2 为本说明书实施方式提供的一种信息处理方法的流程示意图；

图 3 为本说明书实施方式提供的一种界面的示意图；

图 4 为本说明书实施方式提供的一种界面的示意图；

图 5 为本说明书实施方式提供的一种应用场景示意图；

图 6 为本说明书实施方式提供的一种界面的示意图；

图 7 为本说明书实施方式听过的一种界面的示意图；

图 8 为本说明书实施方式提供的一种信息处理方法的流程示意图；

图 9 为本说明书实施方式提供的一种信息处理方法的流程示意图；

图 10 为本说明书实施方式提供的一种信息处理方法的流程示意图；

图 11 为本说明书实施方式提供的一种信息处理方法的流程示意图；

图 12 为本说明书实施方式提供的一种电子设备的模块示意图；

图 13 为本说明书实施方式提供的一种电子设备的架构示意图；

图 14 为本说明书实施方式提供的一种信息处理方法的流程示意图；

图 15 为本说明书实施方式提供的一种电子设备的模块示意图；

5 图 16 为本说明书实施方式提供的一种电子设备的架构示意图；

图 17 为本说明书实施方式提供的一种信息处理方法的流程示意图；

图 18 为本说明书实施方式提供的一种电子设备的模块示意图；

图 19 为本说明书实施方式提供的一种电子设备的架构示意图；

图 20 为本说明书实施方式提供的一种信息处理系统的示意图；

10 图 21 为本说明书实施方式提供的一种信息处理方法的流程示意图；

图 22 为本说明书实施方式提供的一种信息处理方法的流程示意图。

具体实施方式

为了使本技术领域的人员更好地理解本说明书中的技术方案，下面将结合本说明书
15 实施方式中的附图，对本说明书实施方式中的技术方案进行清楚、完整地描述，显然，
所描述的实施方式仅仅是本说明书一部分实施方式，而不是全部的实施方式。基于本说
明书中的实施方式，本领域普通技术人员在没有作出创造性劳动前提下所获得的所有其
他实施方式，都应当属于本说明书保护的范围。

20 请参阅图 20。本说明书实施方式提供一种信息处理系统。所述信息处理系统可以包
括音频信息采集子系统、速记子系统、语音识别子系统、排版子系统和文本更正子系统。

在本实施方式中，音频信息采集子系统可以接收音频采集终端提供的音频信息，将
音频信息提供给语音识别子系统。音频信息采集子系统针对音频信息进行初步筛选，
以在存在多个音频采集终端时，避免将多个录制相同语音的音频信息发送给语音识别子
25 系统。音频信息采集子系统可以将多个音频采集终端提供的音频信息，进行波形比较判
断是否存在二个以上音频采集终端采集了趋于相同的音频信息。在发现二个以上音频采
集终端提供的音频信息的波形趋于相同时，可以将波形能量最强的音频信息提供给语音
识别子系统。

在本实施方式中，在设置有多多个音频采集终端时，音频信息采集子系统针对每
30 个音频采集终端对应设置一个标示名称。在将音频信息发送给语音识别子系统时，对应

发送采集该音频信息的音频采集终端的标示名称。如此，使得语音识别子系统可以将识别出的内容，与标示名称对应。再者，通常一个音频采集终端设置在一个座位，供位于该座位的用户使用。使得，该标示名称与用户之间存在对应关系。

在本实施方式中，音频采集子系统可以记录接收的音频信息的接收时间。在将音频信息发送给语音识别子系统时，可以将该接收时间一并发送给语音识别子系统。当然，并不限于将接收时间和音频信息一并发给语音识别子系统，还可以为将音频信息的生成时间和音频信息一并发送给语音识别子系统。

在本实施方式中，在一次录音过程中，音频采集子系统可以持续的将该次录音过程中的全部内容作为一个音频信息。也可以为，音频采集子系统在一次录音过程中，划分多个音频信息。比如，按照录音的时长划分音频信息。例如，每录制 20 毫秒为一个音频信息，形成一个音频信息。当然，音频信息可以不限于 20 毫秒，其具体时长可以选自 20 毫秒至 500 毫秒。或者，按照数据量进行划分音频信息。例如，每个音频信息最多 5MB。或者，按照音频信息中声音波形的连续情况划分音频信息，比如在相邻两个连续的波形之间存在持续一定时长的无声部分，将该音频信息中每个连续的声音波形划分为一个音频信息。

在本实施方式中，所述速记子系统用于提供给记录人员进行输入使用，即记录人员可以根据听到的语音，在记录人员的意志控制下输入的针对语音的内容进行记录的第一文本信息。速记子系统可以接收记录人员操作输入设备所输入的第一文本信息。所述速记子系统可以将记录人员输入的第一文本信息提供给排版子系统，以用于排版子系统进行排版。

在本实施方式中，速记子系统可以记录第一文本信息的时间信息。该时间信息可以是第一文本信息开始输入的时间，也可以是第一文本信息完成输入的时间。速记子系统可以将该时间信息和第一文本信息一并提供给排版子系统。

在本实施方式中，第一文本信息中可以具有对应的发言人的名称。如此使得第一文本信息可以更加直观的表示发言人的身份和内容。具体的，例如，第一文本信息可以为“小明说：‘小张欠我 100 元钱……’”。

在本实施方式中，语音识别子系统可以针对音频信息进行语音识别，得出表示音频信息中语音的第二文本信息。

在本实施方式中，语音识别子系统可以根据预设算法，从音频信息中采集数据，输

出包括所述音频信息的音频数据的特征的特征矩阵。用户的语音会有用户自身的特征，比如音色、语调、语速等等。录制成音频信息时，可以从音频数据中的频率、振幅等角度，体现每个用户自身的声音特征。使得将音频信息按照预设算法生成的特征矩阵，会包括音频信息中音频数据的特征。进而，基于特征矩阵生成的语音特征向量，可以用于表征该音频信息和音频数据。所述预设算法可以是 MFCC（Mel Frequency Cepstrum Coefficient）、MFSC（Mel Frequency Spectral Coefficient）、FMFCC（Fractional Mel Frequency Cepstrum Coefficient）、DMFCC（Discriminative）、LPCC（Linear Prediction Cepstrum Coefficient）等。当然，所属领域技术人员在本说明书技术精髓启示下，还可能采用其它算法实现生成音频信息的特征矩阵，但只要其实现的功能和效果与本说明书方式相同或相似，均应涵盖于本说明书保护范围内。

在本实施方式中，为了进一步区分出音频信息中用户语音的音频数据和非用户语音的音频数据。语音特征向量的生成方法中，还可以包括端点检测处理。进而，可以在特征矩阵中减少将非用户语音的音频数据对应的数据，如此，可以在一定程度上提升生成的语音特征向量与用户之间的关联程度。端点检测处理的方法可以包括但不限于基于能量的端点检测、基于倒谱特征的端点检测、基于信息熵的端点检测、基于自身相关相似距离的端点检测等，在此不再列举。

在本实施方式中，语音识别子系统可以采用语音识别算法对特征矩阵进行处理，得出音频信息中表达的第二文本信息。具体的，例如，语音识别算法可以采用隐式马尔科夫算法或神经网络算法等，对音频信息进行语音识别。

在本实施方式中，语音识别子系统识别出第二文本信息之后，可以将识别的第二文本信息提供给排版子系统。以用于所述排版子系统将第二文本信息和第一文本信息对应排版。在一些情况下，语音识别子系统接收到语音采集子系统提供的与音频信息对应的时间信息和/或标示名称。语音识别子系统可以将时间信息和/或标示名称，与第二文本信息一并提供给排版子系统。

在本实施方式中，所述排版子系统可以对接收的第一文本信息和第二文本信息进行排版。使得存在对应关系的第一文本信息和第二文本信息可以在显示时被对应展示。对应关系包括但不限于：趋于相同时间生成的音频信息和第一文本信息，该音频信息的第二文本信息与所述第一文本信息显示位置较为接近；或者，趋于相同时间生成的音频信息和第一文本信息，该音频信息的第二文本信息与所述第一文本信息具有相同的显示样

式；或者，趋于相同时间生成的音频信息和第一文本信息，该音频信息的第二文本信息与所述第一文本信息有趋于接近的时间标签。

在本实施方式中，所述文本更正子系统可以用于接收记录人员对第一文本信息的修改，从而形成最终文本。第一文本信息是记录人员在多人沟通的过程中，根据听到的内容进行快速输入的文本信息。限于记录人员的输入速度，和对多人沟通的理解能力，第一文本信息的内容可能会比较精炼，但可能不够全面。在一些情况下，可能漏掉一些较为重要的内容，或者存在一些内容表达准确。第二文本信息是根据多人沟通过程中的音频信息进行语音识别得到，使得第二文本信息趋于全面记录了多人沟通的内容。但第二文本信息可能存在篇幅过长，重点不够突出的缺陷。如此，在排版子系统针对第一文本信息和第二文本信息进行排版之后，记录人员可以根据第一文本信息和对应的第二文本信息进行比较，修改第一文本信息。使得，第一文本信息可以在具有语言精练，重点突出的特点上，修正可能的表达不准确和漏记录重要内容的问题，使得第一文本信息可以更加的完善。

本说明书实施方式提供一种信息处理系统。所述信息处理系统可以包括音频采集终端和客户端；或者，音频采集终端、客户端和服务端。

在本实施方式中，所述音频采集终端可以用于将用户的语音录制生成音频信息。将所述音频信息提供给音频信息采集子系统。音频采集终端可以是一个独立的产品，其具有外壳，以及在外壳中安装有传声器和数据通信单元。例如，音频采集终端可以是一个麦克风。传声器将声音信号转换成电信号，得到音频信息，数据通信单元可以将音频信息发送至处理子系统。数据通信单元可以有有线连接的接口，也可以是无通信的模块。具体的，例如，数据通信单元可以是蓝牙模块、Wifi 模块或者音频接口等。当然，音频采集终端还可以集成在一个具有一定数据处理能力的客户端中。比如，客户端可以包括智能手机、平板电脑、笔记本电脑、个人数字助理等。

在一个实施方式中，客户端可以主要包括处理器、存储器、显示器等硬件。客户端可以具有较强的数据处理能力。客户端可以对音频信息生成特征矩阵之后，可以进行端点检测处理、降噪处理、语音识别等。具体的，例如，客户端可以是工作站、配置很好的台式电脑或笔记本电脑。

在本实施方式中，客户端可以运行前述音频信息采集子系统、速记子系统、语音识别子系统、排版子系统和文本更正子系统。

在一个实施方式中，客户端可以主要包括网络通信单元、处理器、存储器、显示器等硬件。客户端可以将音频信息通过网络通信单元发送给服务器。客户端也可以对音频信息进行一定处理，比如生成特征矩阵，将特征矩阵发送给服务器。以用于服务器针对音频信息中的内容进行语音识别等。具体的，例如，客户端可以包括：智能手机、平板电脑、台式电脑、笔记本电脑等。

在本实施方式中，客户端可以运行前述音频信息采集子系统、速记子系统、排版子系统和文本更正子系统。

当然，上述只是示例的方式列举了一些客户端。随着科学技术进步，硬件设备的性能可能会有提升，使得目前数据处理能力较弱的电子设备，也可能具备较佳的处理能力。

在本实施方式中，服务器可以是具有一定运算处理能力的电子设备。其可以具有网络通信端子、处理器和存储器等。当然，上述服务器也可以是指运行于所述电子设备中的软体。上述服务器还可以为分布式服务器，可以是具有多个处理器、存储器、网络通信模块等协同运作的系统。或者，服务器还可以为若干服务器形成的服务器集群。

在本实施方式中，服务器可以运行前述语音识别子系统。服务器也可以运行排版子系统，将排版后的数据发送给客户端。

请一并参阅图 1 和图 2。在一个场景示例中，在法院庭审的场景中，书记员使用的客户端可以为台式电脑，以记录庭审过程中法官、原告和被告的发言内容。在法官、原告和被告的座位设置有包括传声器的音频采集终端。音频采集终端可以通过蓝牙通信技术的方式与台式电脑通信，将录制的音频信息输入给台式电脑。

请参阅图 3。在本场景示例中，书记员可以使用键盘向台式电脑输入第一文本信息。该第一文本信息可以是书记员听到庭审过程中法官、原告或被告语言表达的内容，也可以是书记员根据法官、原告或被告语言表达的内容进行提炼后的内容。例如，原告张三说：“我在前年四月份借给李四 5000 元钱，他说过两个月就还给我。结果，过了两个月，我去找他要，他就往后拖着。一直到现在，我去找他四次，他都以各种理由推脱不还钱。实际上，我知道他有钱，他上个月还买了一辆汽车。”。书记员记录的第一文本信息可以为“原告张三：‘在 2015 年 4 月份借给李四 5000 元，约定 2015 年 6 月份归还，到期

后经张三反复几次催要还款，李四推脱没有还款。张三认为李四有能力偿还欠款。’ ”。
客户端会记录书记员开始输入第一文本信息的时间，具体的，比如 10:22:16。

在本场景示例中，设置在张三的座位的音频采集终端，可以录制张三说话的语音形成音频信息。将该音频信息输入给客户端。客户端可以记录接收到音频信息的时间
5 10:22:20。客户端可以将该音频信息、接收时间和音频采集终端的标示名称“发言者 1”发送给服务器。

在本场景示例中，服务器针对接收到的音频信息进行语音识别和身份识别。得出音频信息中表达的第二文本信息，即“我在前年四月份借给李四 5000 元钱。他说过两个月就还给我。结果过了两个月。我去找他要。他就往后拖着。一直到现在我去找他四次。
10 他都以各种理由推脱不还钱。实际上。我知道他有钱。他上个月还买了一辆汽车。”服务器可以将该第二文本信息与音频采集终端的标示名称“发言者 1”相对应。反馈给客户端的第二文本信息可以为“发言者 1：我在前年四月份借给李四 5000 元钱。他说过两个月就还给我。结果过了两个月。我去找他要。他就往后拖着。一直到现在我去找他四次。他都以各种理由推脱不还钱。实际上。我知道他有钱。他上个月还买了一辆汽车。”。

15 服务器可以将接收时间和第二文本信息一并反馈给客户端。

在本场景示例中，客户端接收到服务器反馈的第二文本信息。将该第二文本信息对应音频信息的时间，与第一文本信息的时间进行比对，将时间上较为接近第一文本信息和第二文本信息对应显示。

在本场景示例中，将第一文本信息和第二文本信息分别展示在一个窗口中。第一文
20 本信息和第二文本信息对应显示，可以为相对应的第一文本信息和第二文本信息横向上并排显示。

在本场景示例中，在第一文本信息或第二文本信息发生触发事件时，比如书记员使用鼠标点击第一文本信息或第二文本信息。客户端可以播放对应的用于生成第二文本信息的音频信息。

25 请参阅图 4。在本场景示例中，书记员可以根据第二文本信息修改第一文本信息。在最初书记员输入的记录中，可能因为时间较为紧张，漏记录了一些内容，可以根据第二文本信息进行修改。具体的，例如，书记员将“原告张三：‘在 2015 年 4 月份借给李四 5000 元，约定 2015 年 6 月份归还，到期后经张三反复几次催要还款，李四推脱没有还款。张三认为李四有能力偿还欠款。’ ”。书记员记录的第一文本信息中，相较于语音识别得到的第二文本信息，漏记了张三向李四催要还款的次数，以及李四有经济能力
30

购买小汽车的事项。书记员可以对照第二文本信息，将第一文本信息修改为“原告张三：‘在 2015 年 4 月份借给李四 5000 元，约定 2015 年 6 月份归还，到期后经张三先后四次催要还款，李四均推脱没有还款。张三认为李四有能力偿还欠款，李四在 2017 年 6 月份还购买了一辆汽车。’”。

- 5 在本场景示例中，客户端通过对应展示第一文本信息和第二文本信息，给书记员修改第一文本信息带来了便利。书记员可以最终打印由第一文本信息形成的庭审记录。如此，使得庭审记录较为全面准确，并且语言精练、重点突出。

10 请一并参阅图 5。在一个场景示例中，在一个会议中，会议记录人员使用的客户端可以为笔记本电脑。该笔记本电脑中集成有传声器，通过传声器可以将会议过程中参与人员的发言录制成音频信息。

15 在本场景示例中，会议记录人员可以使用笔记本电脑的键盘输入第一文本信息。第一文本信息可以是会议记录人员听到的会议参与人发言的内容，也可以是会议记录人员根据会议参与人的发言或讨论，进行总结的内容。具体的，例如，在会议讨论中，王五说：“我们这个项目时间上比较紧急，大家要盯紧进度。我们要赶在年底前完成项目验收，所以大家要抓紧时间，研发部是否可以在本月就给出完善的设计方案？”。丁七说：“本月底不行，我们现在只有初步的设计方案，需要采购材料进行初步验证之后，才可能给出比价完善的设计方案。所以具体提供完善的设计方案的时间，要看采购什么时间能够完成原材料采购。”。钱八说：“我们本月底可以完成原材料的采购。”。丁七说：
20 “如果本月底原材料可以到位的话，我们下个月中旬应该可以提供较为完善的设计方案”。王五说：“好的，那我们就定下来本月底采购部完成原材料采购，下个月中旬研发部提供较为完善的设计方案。”。

25 请一并参阅图 6。在本场景示例中，会议记录人员记录的第一文本信息可以包括“王五说：‘本项目时间上较为紧急，需要各部门加快进度’”，“钱八说：‘本月会完成研发部需要的原材料的采购’”，“丁七说：‘采购部在本月完成原材料采购的条件下，下个月可以提供较为完善的设计方案’”。

30 在本场景示例中，客户端根据会议过程中参与人员的发言录制成音频信息。客户端可以对音频信息进行语音识别和身份识别。得出音频信息中表达的第二文本信息和第二文本信息对应的身份信息。在会议开始时，会议参与人可以先在客户端进行注册，客户端录制会议的音频信息生成标识每个参与人的用户特征向量，并接收输入的身份信息。

如此，在会议过程中可以根据音频信息的语音特征向量与用户特征向量进行匹配得出第二文本信息对应的身份信息。

在本场景示例中，客户端根据音频信息识别得出的第二文本信息可以包括：“王五说：‘我们这个项目时间上比较紧急。大家要盯紧进度。我们要赶在年底前完成项目验收。所以大家要抓紧时间。研发部是否可以在本月就给出完善的设计方案。’”。
5 “丁七说：‘本月底不行。我们现在只有初步的设计方案。需要采购材料进行初步验证之后。才可能给出比价完善的设计方案。所以具体提供完善的设计方案的时间。要看采购什么时间能够完成原材料采购。’”。
“钱八说：‘我们本月底可以完成原材料的采购。’”。
“丁七说：‘如果本月底原材料可以到位的话。我们下个月中旬应该可以提供较为完善的设计方案。’”。
10 “王五说：‘好的。那我们就定下来本月底采购部完成原材料采购。下个月中旬研发部提供较为完善的设计方案。’”。

请一并参阅图 7。在本场景示例中，客户端可以根据语义分析的方式，得出第一文本信息对应的第二文本信息。会议记录人员可以根据第二文本信息的内容，修改第一文本信息。比如，将“本项目时间上较为紧急，需要各部门加快进度”，修改为“本项目
15 时间上较为紧急，要在年底完成项目验收，需要各部门加快进度”，将“采购部在本月完成原材料采购的条件下，下个月可以提供较为完善的设计方案”，修改为“采购部在本月完成原材料采购的条件下，下个月中旬可以提供较为完善的设计方案”。

在本场景示例中，会议记录人员完成对会议记录中第一文本信息的修改之后，将第一文本信息形成的会议记录通过电子邮件发送给参会的人员。

20 请参阅图 8。本说明书实施方式提供一种信息处理方法。所述信息处理方法可以应用于客户端。所述信息处理方法可以包括以下步骤。

步骤 S10：接收第一输入设备输入的第一文本信息；所述第一文本信息为根据语音生成。

25 在本实施方式中，第一输入设备可以供用户操作，以实现客户端接收用户的输入。第一输入设备可以是与客户端连接的外设，例如键盘、鼠标、手写板等。第一输入设备也可以是触摸式显示器，其提供给用户触摸输入功能。

在本实施方式中，第一文本信息可以是用户根据听到的语音，按照个人意志表达的文字。具体的，例如，在多人会议过程中，会议记录人员根据听到的大家交流的内容，
30 使用键盘打字输入的文字信息。

在本实施方式中，接收第一文本信息，可以为第一文本信息是用户通过第一输入设备提供给客户端。如此，使得第一文本信息可以是用户根据个人意志所生成的。

在本实施方式中，第一文本信息为根据语音生成，可以为用户根据其听到的语音，在用户意志的支配下向客户端输入的文字。使得，第一文本信息可以包括用户对听到语音的内容的提炼，可以包括用户自身理解的内容，可以包括用户对听到内容的直接阐述等。

在本实施方式中，第一文本信息可以是一句话，或者一个自然段。在一次多人的沟通过程中，可能会有多个人说话，记录人员最终形成的记录中可以包括多条第一文本信息。具体的，例如，记录人员输入的一个参会人员的一次发言，作为一个第一文本信息；或者，记录人员根据几个参会人员表达的观点提炼出一句话作为一个第一文本信息；或者，记录人员将记录一个参会人员发言的内容，分成几个自然段，每个自然段构成一个第一文本信息。当然，所属领域技术人员在本实施方式技术精髓启示下，还可以做出其它变更，但只要其实现的功能和效果，与本申请相同或相似，均应涵盖于本申请保护范围内。

步骤 S12：接收第二输入设备录制的音频信息；其中，所述音频信息为根据所述语音录制生成。

在本实施方式中，第二输入设备可以为音频采集终端。第二输入设备可以设置在客户端内的传声器，也可以为客户端通过连接外设的方式连接的传声器。具体的，例如，笔记本电脑中集成有传声器。再例如，台式电脑可以通过 USB 或者 Microphone 接口，连接外接的传声器。

在本实施方式中，步骤 S10 和步骤 S12 可以不分先后顺序。具体的，例如，在法院的庭审中，原告和被告都可以有传声器，在原告或者被告发言时，传声器可以录制相应的音频信息。在这个过程中，书记员根据听到的原告或被告的发言，通过键盘向客户端输入第一文本信息。

步骤 S14：对所述音频信息进行语音识别得到第二文本信息。

在本实施方式中，第二文本信息可以是客户端基于语音识别得到的内容。使得，第二文本信息形成的记录可以趋于全面记录音频信息中语音的对话内容。

在本实施方式中，第二文本信息可以是一句话或者一个自然段。具体的，例如，可

可以根据一个音频信息，对应生成一个第二文本信息；或者，根据识别出的文字形成一句话，作为一个第二文本信息；或者，将相邻上下文，表达相接近的内容，作为一个第二文本信息；或者，结合身份识别，将一个用户的一次发言的内容，作为一个第二文本信息。

5

步骤 S16：展示所述第一文本信息和所述第二文本信息，其中，所述第一文本信息和所述第二文本信息中内容存在对应关系。

在本实施方式中，客户端可以通过显示器展示第一文本信息和第二文本信息，如此可以便于用户查阅。由于第二文本信息为根据音频信息进行语音识别生成，其能够较为全面记录音频信息的内容，使得用户可以借鉴第二文本信息针对第一文本信息的内容进行修改。具体的，例如，在庭审过程中或者庭审后，书记员可以利用原告或被告的音频信息语音识别得到的第二文本信息，针对书记员自己记录的第一文本信息进行修改，以使得第一文本信息可以更加的准确，给书记员带来了便利。

在本实施方式中，所述对应关系可以包括：趋于相同时间生成的第一文本信息和第二文本信息显示位置较为接近；或者，趋于相同时间生成的第一文本信息和第二文本信息具有相同的显示样式；或者，趋于相同时间生成的第一文本信息和第二文本信息具有趋于接近的时间标签；或者，趋于相同时间生成的音频信息和第一文本信息，该音频信息的第二文本信息与所述第一文本信息显示位置较为接近；或者，趋于相同时间生成的音频信息和第一文本信息，该音频信息的第二文本信息与所述第一文本信息具有相同的显示样式；或者，趋于相同时间生成的音频信息和第一文本信息，该音频信息的第二文本信息与所述第一文本信息有趋于接近的时间标签；趋于表达相同语义的第一文本信息和第二文本信息显示位置较为接近，或者，具有相同显示样式。

本说明书实施方式，通过针对同一个语音场景中，可以分别采用人工输入和录制音频后语音识别的方式得到第一文本信息和第二文本信息。利用第二文本信息可以较为全面的记录语音内容的特点，对第一文本信息进行修改。如此，使得第一文本信息可以较为全面准确，且可以突出重点，以及概要的记录语音的内容。第一文本信息相较于第二文本信息可以具有较小的篇幅，重点相对突出，可以节省阅读者的阅读时间。

请参阅图 9，在一个实施方式中，在语音识别的步骤中可以包括以下子步骤。

步骤 S20: 识别所述音频信息对应的身份信息。

步骤 S22: 识别所述音频信息中的语音表达的第二文本信息; 所述第二文本信息与所述身份信息相关联。

在所述展示的步骤中可以将所述身份信息与对应的第二文本信息对应展示。

5 在本实施方式中, 可以根据音频信息生成特征矩阵; 进而, 将所述特征矩阵按照多个特征维度进行降维处理, 得到多个用于表征特征维度的维度值, 所述多个维度值形成所述语音特征向量。可以使用语音特征向量来标识每个用户。

在本实施方式中, 可以对所述特征矩阵按照不同的特征维度进行降维处理, 得到可以表征每个特征维度的维度值。进一步的, 将维度值按照指定顺序排列便可以形成音频信息的语音表征向量。具体的, 可以通过卷积或者映射的算法对特征矩阵进行降维处理。
10 在一个具体的示例中, 可以采用 DNN (Deep Neural Network)、CNN (Convolutional Neural Network) 和 RNN (Recurrent Neural Network)、深度学习或者上述算法的结合等, 从特征矩阵中按照不同维度进行降维。

在本实施方式中, 采集的音频信息可以是用户的说话声音的录音。如此, 根据音频信息生成的语音表征向量, 可以对应表征的音频信息, 也可以表征用户的一部分声音特质。由于每个用户生长发育过程, 是各部不相同的, 使得用户说话的声音, 都具有一定的声音特质。进而, 可以通过每个用户的声音特质分不同的用户。如此, 语音表征向量可以通过表征用户的一部分声音特质, 而可以用于标识用户。
15

在本实施方式中, 针对用户采集的音频信息可以是一个或多个, 可以对应每个音频信息采用音频信息处理方法生成对应的语音特征向量。当然, 在一些情况下, 也可以将一个以上音频信息同时进行按照音频信息处理方法进行运算处理, 得到语音特征向量。此时, 该语音特征向量可以对应该一个以上音频信息。
20

在本实施方式中, 可以根据得到的语音特征向量, 确定可以用于标识用户的用户特征向量。具体的, 例如, 若仅生成了一个语音特征向量, 则可以将该语音特征向量作为用户的用户特征向量; 若生成了多个语音特征向量, 可以将在该多个语音特征向量中, 选择一个相对表达用户的声音特质较多的语音特征向量, 作为用户的用户特征向量; 若生成了多个语音特征向量, 还可以为将该多个语音特征向量中的部分或全部, 进行进一步运算处理输出用户的用户特征向量。该运算处理可以包括但不限于针对该多个语音特征向量进行相应维度求和之后, 再进一步计算均值。当然, 还可以有其它算法, 比如,
25 运算处理时对多个语音特征向量的加权求和。
30

在本实施方式中，将语音特征向量与用户特征向量进行匹配，在匹配成功时，将所述用户特征向量关联的个人信息作为所述用户的身份信息。具体的，将语音特征向量与用户特征向量进行匹配的方式，可以为根据二者进行运算，在二者之间符合某种关系时，可以认为匹配成功。具体的，例如，将二者做差后求和，将得到的数值作为匹配值，将该匹配值与一个设定阈值比较，在所述匹配值小于或等于设定阈值的情况下认为所述语音特征向量与所述用户特征向量匹配成功。或者，也可以将所述语音特征向量与所述用户特征向量直接求和，将得到的数值作为匹配值，在所述匹配值大于或等于设定阈值的情况下认为所述语音特征向量与所述用户特征向量匹配成功。

在本实施方式中，在识别用户的身份信息的步骤中，可以为先采集用户的音频信息之后，录入个人信息。如此，可以根据音频信息生成语音特征向量，进而将录入的个人信息与语音特征向量关联。实现可以将语音识别得到的第二文本信息与发出该语音的用户的身份信息相关联，进而在展示时，可以对应展示身份信息和第二文本信息。具体的，所述关联可以为对应存储，或者根据二者中的一个可以确定另一个。具体的，例如，个人信息可以是用户的姓名、昵称或者角色等。所述角色可以为原告或被告。当然，也可以无需录入个人信息，自动为用户分配一个标示名称，将该标示名称与语音特征向量关联。实现可以将语音识别得到的第二文本信息与发出该语音的用户的标示名称相关联，进而在展示时，可以对应展示标示名称和第二文本信息。具体的，例如，标示名称可以为“用户1”。

在本实施方式中，将身份信息与第二文本信息对应展示，如此便于用户阅读，记录中涉及的人员，以及对应人员的发言情况。对应展示的方式，可以为将身份信息显示在第二文本信息的开始或者末尾。

请参阅图 10，在一个实施方式中，所述信息处理方法还可以包括以下步骤。

步骤 S21：将所述音频信息发送给服务器，以用于所述服务器确定所述音频信息对应的身份信息。

步骤 S23：接收所述服务器反馈的所述音频信息对应的身份信息。

相应的，在展示的步骤中，将所述身份信息与对应的第二文本信息对应展示。

在本实施方式中，可以由服务器完成对音频信息对应的身份信息的识别。用户可以预先在服务器中进行注册操作，使得服务器中对应用户的身份信息存储有用户特征向量。服务器可以根据接收的音频信息生成语音特征向量与存储的用户特征向量进行匹配，得

出语音特征向量对应的用户特征向量，进而得到音频信息对应的身份信息。当然，服务器中也可以没有预先注册的用户特征向量，服务器可以接收音频信息之后，生成每个音频信息的语音特征向量。针对多个语音特征向量进行聚类，将属于一个用户的语音的语音特征向量聚合在一个数据集中，进而可以得出可以表征用户的用户特征向量。针对每个用户的身份信息，可以按照制定的命名规则为用户特征向量分配用户的标示名称。每个用户特征向量可以对应一个标示名称。

在一个实施方式中，在展示所述第一文本信息和所述第二文本信息的步骤中或者对第一文本信息和第二文本信息进行排版的步骤中，可以包括：将接收到的时间较为接近的第一文本信息和音频信息相对应，以对应展示所述第一文本信息和所述音频信息被识别得出的第二文本信息。

在本实施方式中，客户端接收到用户输入的第一文本信息时，可以记录接收时间。在接收到生成的音频信息时，也可以记录音频信息的接收时间。该接收时间也可以理解为生成时间。

在本实施方式中，第一文本信息和音频信息相对应，可以为第一文本信息和音频信息所表达的内容趋于具有相同的语义。具体的，趋于相同接收的时间的第一文本信息和音频信息，有较大可能性表达相同的语义。如此，将第一文本信息和第二文本信息对应展示，以便于针对第一文本信息或者第二文本信息进行对照修改。具体的，例如，多个用户在开会沟通过程中，会议记录人员可能在听到一个用户说话之后，记录说话的内容，生成第一文本信息。此时，传声器可以将该用户说话的语音录制成音频信息。可见，会议记录人员输入第一文本信息的时间，和传声器录制生成语音的时间可以较为接近，二者都针对该会议中一个用户的发言。使得第一文本信息和第二文本信息之间存在对应。

在本实施方式中，对应展示第一文本信息和第二文本信息的方式可以包括：在一个窗口中，显示第一文本信息和第二文本信息的位置较为接近，比如作为相邻的两段文字；在两个窗口中，可以横向并列展示第一文本信息和第二文本信息；或者，第一文本信息和第二文本信息具有相同的显示样式，比如字体、大小、颜色、背景色等。

在一个实施方式中，在展示所述第一文本信息和所述第二文本信息的步骤中或者在对所述第一文本信息和所述第二文本信息进行排版的步骤中，可以包括：将所述第一文本信息在指定时间范围内生成的音频信息的第二文本信息中，进行语义匹配，得到与所

述第一文本信息对应的第二文本信息。

在本实施方式中，第一文本信息是用户根据听到的语音进行输入，而第二文本信息是音频信息进行语音识别得到的。所以，第一文本信息和第二文本信息，都是针对同一场景中，用户的对话语音生成。如此，第一文本信息和第二文本信息在语义上，存在一定的一致性。将第一文本信息在第二文本信息中进行语义分析匹配，得出与第一文本信息语义较为接近的第二文本信息。将得到的语义较为接近的第二文本信息作为与所述第一文本信息对应的第二文本信息。

在本实施方式中，进行语义匹配的方式可以包括：针对第一文本信息进行切词，划分出第一文本信息中包括的词语，将第一文本信息被划分的词语与第二文本信息进行比对。将包括最多个第一文本信息的词语的第二文本信息作为与第一文本信息对应的第二文本信息；或者，将包括最多个第一文本信息词语的近义词的第二文本信息作为第一文本信息对应的第二文本信息；或者，将包括最多个第一文本信息的词语和词语的近义词的第二文本信息，作为与第一文本信息对应的第二文本信息。

在本实施方式中，为了减少运算负荷，可以指定一个时间范围，将第一文本信息位于该时间范围内生成的音频信息的第二文本信息中进行匹配。该时间范围可以与生成第一文本信息的时间较为接近，该时间范围可以包括生成第一文本信息的时间，也可以不包括生成第一文本信息的时间。在一个实施方式中，在进行语义匹配的步骤中还包括：将接收到所述第一文本信息的时间作为基准时间，根据所述基准时间设置所述指定时间范围，所述基准时间位于所述指定时间范围内。如此，指定时间范围可以包括接收第一文本信息的时间，使得声音信息的时间，与生成第一文本信息的时间较为接近，可以提升匹配到与第一文本信息对应的第二文本信息的速度，进而节省时间和减少运算量。

请参阅图 11。在一个实施方式中，所述信息处理方法还可以包括以下步骤。

步骤 S26：根据所述第一输入设备的输入对所述第一文本信息进行修改。

在本实施方式中，在展示第一文本信息和第二文本信息之后，用户可以根据第二文本信息的内容，修改第一文本信息。如此，可以使得第一文本信息在具有人工记录的语言精练、重点突出等特点的基础上，进一步的提升了准确性和全面性。再者，也为记录的制作人员带来了便利。

本实施方式还可以包括步骤 S28：输出修改后的第一文本信息。

在本实施方式中，输出修改后的第一文本信息可以为将第一文本信息发送至打印设

备，或者，将第一文本信息通过邮件发送。

在一个实施方式中，所述信息处理方法还可以包括：在所述第一文本信息或者所述第二文本信息发生触发事件的情况下，播放发生触发事件的第一文本信息或者第二文本信息对应的音频信息。

在本实施方式中，触发事件可以是客户端接收到了点击操作或者划动操作。具体的，例如，用户使用鼠标针对第一文本信息或者第二文本信息进行点击操作，或者，移动鼠标划过第一文本信息或第二文本信息；或者，用户使用触摸显示设备进行了触摸点击操作。

在本实施方式中，第一文本信息或第二文本信息发生触发事件，可以是显示第一文本信息或者第二文本信息的区域发生了触发事件。也可以是对应第一文本信息或第二文本信息设置有播放按钮，在播放按钮发生点击操作时，认为相应的第一文本信息或第二文本信息发生了触发事件。

在本实施方式中，第一文本信息对应音频信息的方式，可以参阅其它实施方式的描述，在此不再赘述。第二文本信息对应的音频信息，可以是被识别得到第二文本信息的音频信息。

在一个实施方式中，所述信息处理方法还可以包括：在所述第一文本信息发生触发事件的情况下，将与所述第一文本信息对应的第二文本信息按照指定样式显示；或者，在所述第二文本信息发生触发事件的情况下，将与所述第二文本信息对应的第一文本信息按照指定样式显示。

在本实施方式中，指定样式可以包括但不限于：字体、大小、颜色、背景色、加粗效果、倾斜效果、下划线等等。

在本实施方式中，在第一文本信息发生触发事件时，将对应的第二文本信息按照指定样式显示，可以比较直观的指示二者的对应关系。还可以便于用户进行查阅对照。指定样式可以与其余文字的样式不同，如此使得发生对应的文本信息的显示状态与其它文字相区别，给用户带来了便利。同理，第二文本信息发生触发事件时，第一文本信息按照指定样式显示。

请参阅图 12。本说明书实施方式还提供一种信息处理系统，所述信息处理系统可以

包括以下模块。

第一接收模块，用于接收第一输入设备输入的第一文本信息；所述第一文本信息为根据语音生成。

第二接收模块，用于接收第二输入设备录制的音频信息；其中，所述音频信息为根据所述语音录制生成。

识别模块，用于对所述音频信息进行语音识别得到第二文本信息。

展示模块，用于展示所述第一文本信息和所述第二文本信息，其中，所述第一文本信息和所述第二文本信息中内容存在对应关系。

在本实施方式中，电子设备实现的功能和效果可以与其它实施方式对照解释，在此不再赘述。

请参阅图 13。本说明书实施方式还提供一种信息处理系统，所述信息处理系统可以包括：输入设备、音频采集终端、显示器和处理器。

所述输入设备，用于接收用户输入的第一文本信息；所述第一文本信息为用户根据语音进行输入。

所述音频采集终端，用于录制音频信息；其中，所述音频信息为根据所述语音录制生成。

所述处理器，用于对所述音频信息进行语音识别得到第二文本信息。

所述显示器，用于展示所述第一文本信息和所述第二文本信息，其中，所述第一文本信息和所述第二文本信息中内容存在对应关系。

在本实施方式中，所述输入设备可以是具有信息输入功能的设备。具体的，例如，输入设备可以是键盘、手写板、手写笔等。当然，输入设备并不限于前述列举。

在本实施方式中，所述音频采集终端包括有传声器。所述传声器可以是将声音信号转换为电信号的能量转换器件。具体的，例如，传声器的具体设计，可以包括但不限于电动式、电容式、压电式、电磁式、碳粒式、半导体式等。

在本实施方式中，所述处理器可以按任何适当的方式实现。例如，所述处理器可以采取例如微处理器或处理器以及存储可由该（微）处理器执行的计算机可读程序代码（例如软件或固件）的计算机可读介质、逻辑门、开关、专用集成电路（Application Specific Integrated Circuit, ASIC）、可编程逻辑控制器和嵌入微控制器的形式等等。

在本实施方式中，所述显示器可以提供界面展示。具体的，例如，显示器可以包括

根据制造材料的不同，可分为：阴极射线管显示器（Cathode Ray Tube）、等离子显示器（Plasma Display Panel）、液晶显示器（Liquid Crystal Display）、LED 显示屏（Light Emitting Diode Panel）等等。当然，显示器并不限于平面显示器，其还可以为弧面显示器、立体显示器等。

5 本实施方式中提供的信息处理系统可以参见其它实施方式对照解释。

本说明书实施方式还提供一种计算机存储介质。所述计算机存储介质存储有计算机程序指令，所述计算机程序指令被执行时可以实现：接收第一输入设备输入的第一文本信息，所述第一文本信息为根据语音生成；接收第二输入设备录制的音频信息，其中，
10 所述音频信息为根据所述语音录制生成；对所述音频信息进行语音识别得到第二文本信息；展示所述第一文本信息和所述第二文本信息，其中，所述第一文本信息和所述第二文本信息中内容存在对应关系。

在本实施方式中，所述计算机存储介质包括但不限于随机存取存储器（Random Access Memory, RAM）、只读存储器（Read-Only Memory, ROM）、缓存（Cache）、
15 硬盘（Hard Disk Drive, HDD）或者存储卡（Memory Card）。

本实施方式中提供的计算机存储介质，其计算机程序指令被执行时实现的功能和效果可以参见其它实施方式对照解释。

请参阅图 14。本说明书实施方式还提供一种信息处理方法，所述信息处理方法可以
20 包括以下步骤。

步骤 S30：接收第一输入设备输入的第一文本信息；所述第一文本信息为根据语音生成。

步骤 S32：接收第二输入设备录制的音频信息；其中，所述音频信息为根据所述语音录制生成。

25 步骤 S34：将所述音频信息或所述音频信息的表征信息发送给服务器，以用于所述服务器进行语音识别。

步骤 S36：接收所述服务器反馈的语音识别得到的第二文本信息。

步骤 S38：展示所述第一文本信息和所述第二文本信息，其中，所述第一文本信息和所述第二文本信息中内容存在对应关系。

30 在本实施方式中，客户端可以将音频信息发送给服务器，服务器可以对音频信息进

行语音识别。服务器可以将语音识别得到的第二文本信息发送给客户端。使得客户端可以展示第一文本信息和第二文本信息。在本实施方式中，服务器对音频信息的处理，可以参阅其它实施方式对照解释，在此不再赘述。

在本实施方式中，客户端和服务端可以基于网络通信协议进行数据传输。网络通信
5 协议包括但不限于 HTTP、TCP/IP、FTP 等。

在本实施方式中，所述表征信息可以是根据音频信息进行运算生成，可以用于表示所述音频信息。具体的，例如，所述表征信息可以是根据音频信息生成的特征矩阵，或者针对特征矩阵进行端点检测处理之后的数据。

在本实施方式中，由服务器进行语音识别工作，如此减少了客户端的运算负荷。进
10 而降低了对客户端的硬件性能要求。

请参阅图 15。本说明书实施方式还提供一种信息处理系统。所述信息处理系统可以包括以下模块。

第一接收模块，用于接收第一输入设备输入的第一文本信息；所述第一文本信息为
15 根据语音生成。

第二接收模块，用于接收第二输入设备录制的音频信息；其中，所述音频信息为根据所述语音录制生成。

发送模块，用于将所述音频信息或所述音频信息的表征信息发送给服务器，以用于所述服务器进行语音识别。

20 第三接收模块，用于接收所述服务器反馈的语音识别得到的第二文本信息。

展示模块，用于展示所述第一文本信息和所述第二文本信息，其中，所述第一文本信息和所述第二文本信息中内容存在对应关系。

在本实施方式中，电子设备实现的功能和效果可以与其它实施方式对照解释，在此不再赘述。

25 请参阅图 16。本说明书实施方式还提供一种信息处理系统。所述信息处理系统可以包括：输入设备、声音采集终端、网络通信单元、显示器。

所述输入设备，用于接收用户输入的第一文本信息；所述第一文本信息为用户根据语音进行输入。

30 所述声音采集终端，用于录制音频信息；其中，所述音频信息为根据所述语音录制

生成。

所述网络通信单元,用于将所述音频信息或所述音频信息的表征信息发送给服务器,以用于所述服务器进行语音识别;接收所述服务器反馈的语音识别得到的第二文本信息。

所述显示器,用于展示所述第一文本信息和所述第二文本信息,其中,所述第一文本信息和所述第二文本信息中内容存在对应关系。

在本实施方式中,所述网络通信单元可以是依照通信协议规定的标准设置的,用于进行网络连接通信的接口。网络通信单元可以包括但不限于有线通信和/或无线通信的方式。

本实施方式中提供的信息处理系统可以参见其它实施方式对照解释。

10

本说明书实施方式还提供一种计算机存储介质,所述计算机存储介质存储有计算机程序指令。所述计算机程序指令被执行时可以实现:接收第一输入设备输入的第一文本信息,所述第一文本信息为根据语音生成;接收第二输入设备录制的音频信息,其中,所述音频信息为根据所述语音录制生成;将所述音频信息或所述音频信息的表征信息发送给服务器,以用于所述服务器进行语音识别;接收所述服务器反馈的语音识别得到的第二文本信息;展示所述第一文本信息和所述第二文本信息,其中,所述第一文本信息和所述第二文本信息中内容存在对应关系。

在本实施方式中,所述计算机存储介质包括但不限于随机存取存储器(Random Access Memory, RAM)、只读存储器(Read-Only Memory, ROM)、缓存(Cache)、硬盘(Hard Disk Drive, HDD)或者存储卡(Memory Card)。

本实施方式中提供的计算机存储介质,其计算机程序指令被执行时实现的功能和效果可以参见其它实施方式对照解释。

请参阅图 17。本说明书实施方式还提供一种信息处理方法,所述信息处理方法可以包括以下步骤。

步骤 S40: 接收客户端发送的音频信息或者音频信息的表征信息。

步骤 S42: 对所述音频信息或者所述表征信息进行语音识别得到第二文本信息。

步骤 S44: 将所述第二文本信息发送给所述客户端,以用于所述客户端将所述第二文本信息与第一客户端的第一文本信息进行展示;其中,所述第一文本信息和所述第二文本信息中内容存在对应关系。

30

在本实施方式中，服务器接收客户端通过网络发送的音频信息或者表征信息。服务器中可以设置有针对性音频信息或者表征信息进行语音识别的算法模型。具体的，针对音频信息或表征信息的处理方式，可以参阅前述实施方式中，针对音频信息和表征信息的描述，在此不再赘述。

- 5 在本实施方式中，服务器可以对应多个客户端，实现客户端不需要进行语音识别工作，而统一有服务器进行语音识别。如此降低了对客户端的硬件性能要求。再者，由服务器进行统一的语音识别工作，便于对语音识别工作的算法维护和更新升级。

本实施方式中描述的内容，可以参照其它实施方式对照解释，在此不再赘述。

- 10 请参阅图 9。在一个实施方式中，服务器进行语音识别的步骤中可以包括以下步骤。
步骤 S20：识别所述音频信息对应的身份信息。

步骤 S22：识别所述音频信息中的语音表达的第二文本信息；所述第二文本信息与所述身份信息相关联。

- 15 在本实施方式中，服务器可以识别音频信息对应的身份信息，是的服务器可以将属于同一个用户的第二文本信息，与该用户的身份信息进行关联。服务器将身份信息放入所述第二文本信息中提供给客户端之后，便于用户通过客户端展示的内容阅读。

本实施方式具体实现的功能和效果，可以参照其它实施方式对照解释，在此不再赘述。

- 20 请参阅图 18。本说明书实施方式还提供一种服务器，所述服务器可以包括以下模块。
接收模块，用于接收客户端发送的音频信息或音频信息的表征信息。

识别模块，用于对所述音频信息或者所述表征信息进行语音识别得到第二文本信息。

- 25 发送模块，用于将所述第二文本信息发送给所述客户端，以用于所述客户端将所述第二文本信息与第一客户端的第一文本信息进行展示；其中，所述第一文本信息和所述第二文本信息中内容存在对应关系。

在本实施方式中，服务器实现的功能和效果可以与其它实施方式对照解释，在此不再赘述。

- 30 请参阅图 19。本说明书实施方式还提供一种电子设备，所述电子设备可以包括：网络通信单元和处理器。

所述网络通信单元，用于接收客户端发送的音频信息或音频信息的表征信息；将所述处理器提供的第二文本信息发送给所述客户端，以用于所述客户端将所述第二文本信息与第一客户端的第一文本信息进行展示；其中，所述第一文本信息和所述第二文本信息中内容存在对应关系。

- 5 所述处理器，用于对所述音频信息或者所述表征信息进行语音识别得到第二文本信息。

本实施方式中提供的电子设备，其实现的功能和效果可以参见其它实施方式对照解释。

- 10 本说明书实施方式还提供一种计算机存储介质，所述计算机存储介质中存储有计算机程序指令，所述计算机程序指令被执行时可以实现：接收客户端发送的音频信息或音频信息的表征信息；对所述音频信息或者所述表征信息进行语音识别得到第二文本信息；将所述第二文本信息发送给所述客户端，以用于所述客户端将所述第二文本信息与第一客户端的第一文本信息进行展示，其中，所述第一文本信息和所述第二文本信息中内容
15 存在对应关系。

本实施方式中提供的计算机存储介质，其计算机程序指令被执行时实现的功能和效果可以参见其它实施方式对照解释。

- 20 请参阅图 21。本说明书实施方式还提供一种信息处理方法。所述信息处理方法可以包括以下步骤。

步骤 S50：接收第一文本信息；所述第一文本信息为根据语音生成。

步骤 S52：接收音频信息；其中，所述音频信息为根据所述语音录制生成。

步骤 S54：对所述音频信息进行语音识别得到第二文本信息。

- 25 步骤 S56：根据所述第一文本信息和所述第二文本信息之间的对应关系进行排版，所述排版后的所述第一文本信息和所述第二文本信息用于展示。

在本实施方式中，执行该信息处理方法的电子设备，可以对音频信息进行语音识别得到第二文本信息。并进一步的，根据第一文本信息和第二文本信息之间的对应关系针对第一文本信息和第二文本信息进行排版。

- 30 在本实施方式中，对第一文本信息和第二文本信息进行排版。具体排版的方式可以参阅前述排版子系统的内容，或者其它实施方式对照解释。

本说明书实施方式还提供一种电子设备，所述电子设备包括处理器，所述处理器用于接收第一文本信息，所述第一文本信息为根据语音生成；接收音频信息，其中，所述音频信息为根据所述语音录制生成；对所述音频信息进行语音识别得到第二文本信息；

5 根据所述第一文本信息和所述第二文本信息之间的对应关系进行排版，所述排版后的所述第一文本信息和所述第二文本信息用于展示。

本实施方式中提供的电子设备，其实现的功能和效果可以参见其它实施方式对照解释。

10 本说明书实施方式还提供一种计算机存储介质，所述计算机存储介质中存储有计算机程序指令，所述计算机程序指令被执行时实现：接收第一文本信息，所述第一文本信息为根据语音生成；接收音频信息，其中，所述音频信息为根据所述语音录制生成；对所述音频信息进行语音识别得到第二文本信息；根据所述第一文本信息和所述第二文本信息之间的对应关系进行排版，所述排版后的所述第一文本信息和所述第二文本信息用

15 于展示。

本实施方式中提供的计算机存储介质，其计算机程序指令被执行时实现的功能和效果可以参见其它实施方式对照解释。

请一并参阅图 22 和图 3。本说明书实施方式还提供一种信息处理方法，所述方法可以包括以下步骤。

20

步骤 S60：接收第一输入设备输入的第一文本信息；所述第一文本信息为根据语音生成。

步骤 S62：接收第二输入设备录制的音频信息；其中，所述音频信息为根据所述语音录制生成。

25 步骤 S64：对所述音频信息进行识别得到第二文本信息。

步骤 S66：在第一区域展示所述第一文本信息。

步骤 S68：在第二区域展示所述第二文本信息，其中，所述第一文本信息和所述第二文本信息中内容存在对应关系，所述第一区域和第二区域位于相同的界面中。

在本实施方式中，可以将一个用于显示器展示的界面至少划分为第一区域和第二区域。其中，第一区域用于展示第一文本信息，第二区域用于展示第二文本信息。如此，

30

通过至少划分第一区域和第二区域，使得可以清晰的区分第一文本信息和第二文本信息，给浏览人员带来便利。

在本实施方式中，第一文本信息和第二文本信息的生成方式不同。第一文本信息可以是记录人员根据个人对语音的理解，输入的内容。使得第一文本信息可以具有较为概括，语言相对简练，重点相对突出的特点。第二文本信息可以是针对音频信息进行语音识别得到的内容。第二文本信息可以具有内容非常全面的特点。但第二文本信息相较于第一文本信息可能存在内容相对冗长，重点不够突出等问题。通过在第一文本信息和第二文本信息中的内容存在对应关系，使得浏览者可以对照查阅第一文本信息和第二文本信息，从而可以实现给阅读者带来便利。

本说明书中的各个实施方式均采用递进的方式描述，各个实施方式之间相同相似的部分互相参见即可，每个实施方式重点说明的都是与其他实施方式的不同之处。

在 20 世纪 90 年代，对于一个技术的改进可以很明显地区分是硬件上的改进（例如，对二极管、晶体管、开关等电路结构的改进）还是软件上的改进（对于方法流程的改进）。然而，随着技术的发展，当今的很多方法流程的改进已经可以视为硬件电路结构的直接改进。设计人员几乎都通过将改进的方法流程编程到硬件电路中来得到相应的硬件电路结构。因此，不能说一个方法流程的改进就不能用硬件实体模块来实现。例如，可编程逻辑器件（Programmable Logic Device, PLD）（例如现场可编程门阵列（Field Programmable Gate Array, FPGA））就是这样一种集成电路，其逻辑功能由用户对器件编程来确定。由设计人员自行编程来把一个数字系统“集成”在一片 PLD 上，而不需要请芯片制造厂商来设计和制作专用的集成电路芯片。而且，如今，取代手工地制作集成电路芯片，这种编程也多半改用“逻辑编译器（logic compiler）”软件来实现，它与程序开发撰写时所用的软件编译器相类似，而要编译之前的原始代码也得用特定的编程语言来撰写，此称之为硬件描述语言（Hardware Description Language, HDL），而 HDL 也并非仅有一种，而是有许多种，如 ABEL（Advanced Boolean Expression Language）、AHDL（Altera Hardware Description Language）、Confluence、CUPL（Cornell University Programming Language）、HDCal、JHDL（Java Hardware Description Language）、Lava、Lola、MyHDL、PALASM、RHDL（Ruby Hardware Description Language）等，目前最普遍使用的是 VHDL（Very-High-Speed Integrated Circuit Hardware Description Language）

与 Verilog2。本领域技术人员也应该清楚，只需要将方法流程用上述几种硬件描述语言稍作逻辑编程并编程到集成电路中，就可以很容易得到实现该逻辑方法流程的硬件电路。

5 本领域技术人员也知道，除了以纯计算机可读程序代码方式实现控制器以外，完全可以通过将方法步骤进行逻辑编程来使得控制器以逻辑门、开关、专用集成电路、可编程逻辑控制器和嵌入微控制器等的形式来实现相同功能。因此这种控制器可以被认为是一种硬件部件，而对其内包括的用于实现各种功能的装置也可以视为硬件部件内的结构。或者甚至，可以将用于实现各种功能的装置视为既可以是实现方法的软件模块又可以是硬件部件内的结构。

10 通过以上的实施方式的描述可知，本领域的技术人员可以清楚地了解到本说明书可借助软件加必需的通用硬件平台的方式来实现。基于这样的理解，本说明书的技术方案本质上或者说对现有技术做出贡献的部分可以以软件产品的形式体现出来，该计算机软件产品可以存储在存储介质中，如 ROM/RAM、磁碟、光盘等，包括若干指令用以使得一台计算机设备（可以是个人计算机，服务器，或者网络设备等）执行本说明书各个实施方式或者实施方式的某些部分所述的方法。

15 虽然通过实施方式描绘了本说明书，本领域普通技术人员知道，本说明书有许多变形和变化而不脱离本说明书的精神，希望所附的权利要求包括这些变形和变化而不脱离本说明书的精神。

权 利 要 求 书

1. 一种信息处理方法，其特征在于，所述方法包括：

接收第一输入设备输入的第一文本信息；所述第一文本信息为根据语音生成；

接收第二输入设备录制的音频信息；其中，所述音频信息为根据所述语音录制生成；

5 对所述音频信息进行语音识别得到第二文本信息；

展示所述第一文本信息和所述第二文本信息，其中，所述第一文本信息和所述第二文本信息中内容存在对应关系。

2. 根据权利要求1所述的方法，其特征在于，在语音识别的步骤中包括：

10 识别所述音频信息对应的身份信息；

识别所述音频信息中的语音表达的第二文本信息；所述第二文本信息与所述身份信息相关联；

相应的，在所述展示的步骤中，将所述身份信息与对应的第二文本信息对应展示。

15 3. 根据权利要求1所述的方法，其特征在于，所述方法还包括：

将所述音频信息发送给服务器，以用于所述服务器确定所述音频信息对应的身份信息；

接收所述服务器反馈的所述音频信息对应的身份信息；

相应的，在展示的步骤中，将所述身份信息与对应的第二文本信息对应展示。

20

4. 根据权利要求1所述的方法，其特征在于，在展示所述第一文本信息和所述第二文本信息的步骤中包括：将接收到的时间较为接近的第一文本信息和音频信息相对应，以对应展示所述第一文本信息和所述音频信息被识别得出的第二文本信息。

25 5. 根据权利要求1所述的方法，其特征在于，在展示所述第一文本信息和所述第二文本信息的步骤中包括：

将所述第一文本信息在指定时间范围内生成的音频信息的第二文本信息中，进行语义匹配，得到与所述第一文本信息对应的第二文本信息。

30 6. 根据权利要求5所述的方法，其特征在于，在进行语义匹配的步骤中还包括：将

接收到所述第一文本信息的时间作为基准时间，根据所述基准时间设置所述指定时间范围，所述基准时间位于所述指定时间范围内。

7. 根据权利要求1所述的方法，其特征在于，所述方法还包括：

5 根据所述第一输入设备的输入对所述第一文本信息进行修改；
输出修改后的第一文本信息。

8. 根据权利要求1所述的方法，其特征在于，所述方法还包括：

10 在所述第一文本信息或者所述第二文本信息发生触发事件的情况下，播放发生触发事件的第一文本信息或者第二文本信息对应的音频信息。

9. 根据权利要求1所述的方法，其特征在于，所述方法还包括：

在所述第一文本信息发生触发事件的情况下，将与所述第一文本信息对应的第二文本信息按照指定样式显示；或者，

15 在所述第二文本信息发生触发事件的情况下，将与所述第二文本信息对应的第一文本信息按照指定样式显示。

10. 一种信息处理系统，其特征在于，所述信息处理系统包括：输入设备、音频采集终端、显示器和处理器；

20 所述输入设备，用于接收用户输入的第一文本信息；所述第一文本信息为用户根据语音进行输入；

所述音频采集终端，用于录制音频信息；其中，所述音频信息为根据所述语音录制生成；

所述处理器，用于对所述音频信息进行语音识别得到第二文本信息；

25 所述显示器，用于展示所述第一文本信息和所述第二文本信息，其中，所述第一文本信息和所述第二文本信息中内容存在对应关系。

11. 一种计算机存储介质，其特征在于，所述计算机存储介质存储有计算机程序指令，所述计算机程序指令被执行时实现：接收第一输入设备输入的第一文本信息，所述
30 第一文本信息为根据语音生成；接收第二输入设备录制的音频信息，其中，所述音频信

息为根据所述语音录制生成；对所述音频信息进行语音识别得到第二文本信息；展示所述第一文本信息和所述第二文本信息，其中，所述第一文本信息和所述第二文本信息中内容存在对应关系。

- 5 12. 一种信息处理方法，其特征在于，所述方法包括：
- 接收第一输入设备输入的第一文本信息；所述第一文本信息为根据语音生成；
- 接收第二输入设备录制的音频信息；其中，所述音频信息为根据所述语音录制生成；
- 将所述音频信息或所述音频信息的表征信息发送给服务器，以用于所述服务器进行语音识别；
- 10 接收所述服务器反馈的语音识别得到的第二文本信息；
- 展示所述第一文本信息和所述第二文本信息，其中，所述第一文本信息和所述第二文本信息中内容存在对应关系。

13. 根据权利要求 12 所述的方法，其特征在于，在展示所述第一文本信息和所述第
- 15 二文本信息的步骤中包括：将接收到的时间较为接近的第一文本信息和音频信息相对应，以对应展示所述第一文本信息和所述音频信息被识别得出的第二文本信息。

14. 根据权利要求 12 所述的方法，其特征在于，在展示所述第一文本信息和所述第
- 二文本信息的步骤中包括：
- 20 将所述第一文本信息在指定时间范围内生成的音频信息的第二文本信息中，进行语义匹配，得到与所述第一文本信息对应的第二文本信息。

15. 根据权利要求 14 所述的方法，其特征在于，在进行语义匹配的步骤中还包括：
- 将接收到所述第一文本信息的时间作为基准时间，根据所述基准时间设置所述指定时间
- 25 范围，所述基准时间位于所述指定时间范围内。

16. 根据权利要求 12 所述的方法，其特征在于，所述方法还包括：
- 根据所述第一输入设备的输入对所述第一文本信息进行修改；
- 输出修改后的第一文本信息。

30

17. 根据权利要求 12 所述的方法，其特征在于，所述方法还包括：

在所述第一文本信息或者所述第二文本信息发生触发事件的情况下，播放发生触发事件的第一文本信息或者第二文本信息对应的音频信息。

18. 根据权利要求 12 所述的方法，其特征在于，所述方法还包括：

5 在所述第一文本信息发生触发事件的情况下，将与所述第一文本信息对应的第二文本信息按照指定样式显示；或者，

在所述第二文本信息发生触发事件的情况下，将与所述第二文本信息对应的第一文本信息按照指定样式显示。

10 19. 一种信息处理系统，其特征在于，所述信息处理系统包括：输入设备、音频采集终端、网络通信单元、显示器；

所述输入设备，用于接收用户输入的第一文本信息；所述第一文本信息为用户根据语音进行输入；

15 所述音频采集终端，用于录制音频信息；其中，所述音频信息为根据所述语音录制生成；

所述网络通信单元，用于将所述音频信息或所述音频信息的表征信息发送给服务器，以用于所述服务器进行语音识别；接收所述服务器反馈的语音识别得到的第二文本信息；

20 所述显示器，用于展示所述第一文本信息和所述第二文本信息，其中，所述第一文本信息和所述第二文本信息中内容存在对应关系。

20. 一种计算机存储介质，其特征在于，所述计算机存储介质存储有计算机程序指令，所述计算机程序指令被执行时实现：接收第一输入设备输入的第一文本信息，所述第一文本信息为根据语音生成；接收第二输入设备录制的音频信息，其中，所述音频信息为根据所述语音录制生成；将所述音频信息或所述音频信息的表征信息发送给服务器，以用于所述服务器进行语音识别；接收所述服务器反馈的语音识别得到的第二文本信息；展示所述第一文本信息和所述第二文本信息，其中，所述第一文本信息和所述第二文本信息中内容存在对应关系。

30 21. 一种信息处理方法，其特征在于，所述方法包括：

接收客户端发送的音频信息或音频信息的表征信息；

对所述音频信息或者所述表征信息进行语音识别得到第二文本信息；

将所述第二文本信息发送给所述客户端，以用于所述客户端将所述第二文本信息与第一客户端的第一文本信息进行展示；其中，所述第一文本信息和所述第二文本信息中
5 内容存在对应关系。

22. 根据权利要求 21 所述的方法，其特征在于，在语音识别的步骤中包括：

识别所述音频信息对应的身份信息；

识别所述音频信息中的语音表达的第二文本信息；所述第二文本信息与所述身份信息
10 息相关联。

23. 一种电子设备，其特征在于，所述电子设备包括：网络通信单元和处理器；

所述网络通信单元，用于接收客户端发送的音频信息或音频信息的表征信息；将所述处理器提供的第二文本信息发送给所述客户端，以用于所述客户端将所述第二文本信息
15 与第一客户端的第一文本信息进行展示；其中，所述第一文本信息和所述第二文本信息中内容存在对应关系；

所述处理器，用于对所述音频信息或者所述表征信息进行语音识别得到第二文本信息。

20 24. 一种计算机存储介质，其特征在于，所述计算机存储介质中存储有计算机程序指令，所述计算机程序指令被执行时可以实现：接收客户端发送的音频信息或音频信息的表征信息；对所述音频信息或者所述表征信息进行语音识别得到第二文本信息；将所述第二文本信息发送给所述客户端，以用于所述客户端将所述第二文本信息与第一客户端的第一文本信息进行展示，其中，所述第一文本信息和所述第二文本信息中内容存在
25 对应关系。

25. 一种信息处理方法，其特征在于，包括：

接收第一文本信息；所述第一文本信息为根据语音生成；

接收音频信息；其中，所述音频信息为根据所述语音录制生成；

30 对所述音频信息进行语音识别得到第二文本信息；

根据所述第一文本信息和所述第二文本信息之间的对应关系进行排版，所述排版后的所述第一文本信息和所述第二文本信息用于展示。

26. 根据权利要求 25 所述的方法，其特征在于，在对所述第一文本信息和所述第二文本信息进行排版的步骤中包括：将接收到的时间较为接近的第一文本信息和音频信息相对应，以对应展示所述第一文本信息和所述音频信息被识别得出的第二文本信息。

27. 根据权利要求 25 所述的方法，其特征在于，在对所述第一文本信息和所述第二文本信息进行排版的步骤中包括：

10 将所述第一文本信息在指定时间范围内生成的音频信息的第二文本信息中，进行语义匹配，得到与所述第一文本信息对应的第二文本信息。

28. 根据权利要求 27 所述的方法，其特征在于，在进行语义匹配的步骤中还包括：将接收到所述第一文本信息的时间作为基准时间，根据所述基准时间设置所述指定时间范围，所述基准时间位于所述指定时间范围内。

15

29. 根据权利要求 25 所述的方法，其特征在于，所述方法还包括：

接收对所述第一文本信息进行修改的输入，对所述第一文本信息进行修改。

30. 一种电子设备，其特征在于，所述电子设备包括处理器，

所述处理器用于接收第一文本信息，所述第一文本信息为根据语音生成；接收音频信息，其中，所述音频信息为根据所述语音录制生成；对所述音频信息进行语音识别得到第二文本信息；根据所述第一文本信息和所述第二文本信息之间的对应关系进行排版，所述排版后的所述第一文本信息和所述第二文本信息用于展示。

25

31. 一种计算机存储介质，其特征在于，所述计算机存储介质中存储有计算机程序指令，所述计算机程序指令被执行时实现：接收第一文本信息，所述第一文本信息为根据语音生成；接收音频信息，其中，所述音频信息为根据所述语音录制生成；对所述音频信息进行语音识别得到第二文本信息；根据所述第一文本信息和所述第二文本信息之间的对应关系进行排版，所述排版后的所述第一文本信息和所述第二文本信息用于展示。

30

32. 一种信息处理方法，其特征在于，所述方法包括：

接收第一输入设备输入的第一文本信息；所述第一文本信息为根据语音生成；

接收第二输入设备录制的音频信息；其中，所述音频信息为根据所述语音录制生成；

5 对所述音频信息进行识别得到第二文本信息；

在第一区域展示所述第一文本信息；

在第二区域展示所述第二文本信息，其中，所述第一文本信息和所述第二文本信息
中内容存在对应关系，所述第一区域和第二区域位于相同的界面中。

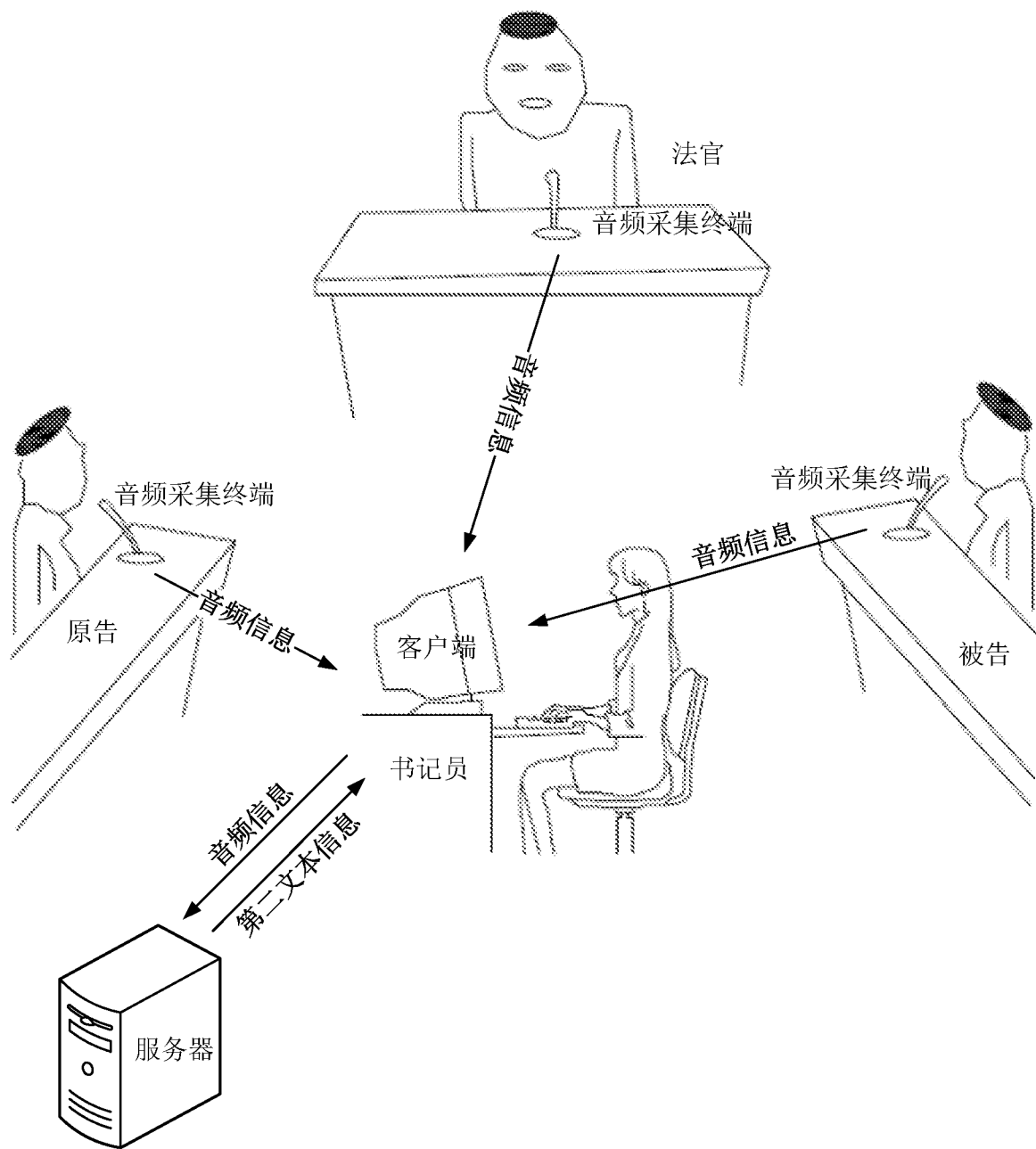


图 1

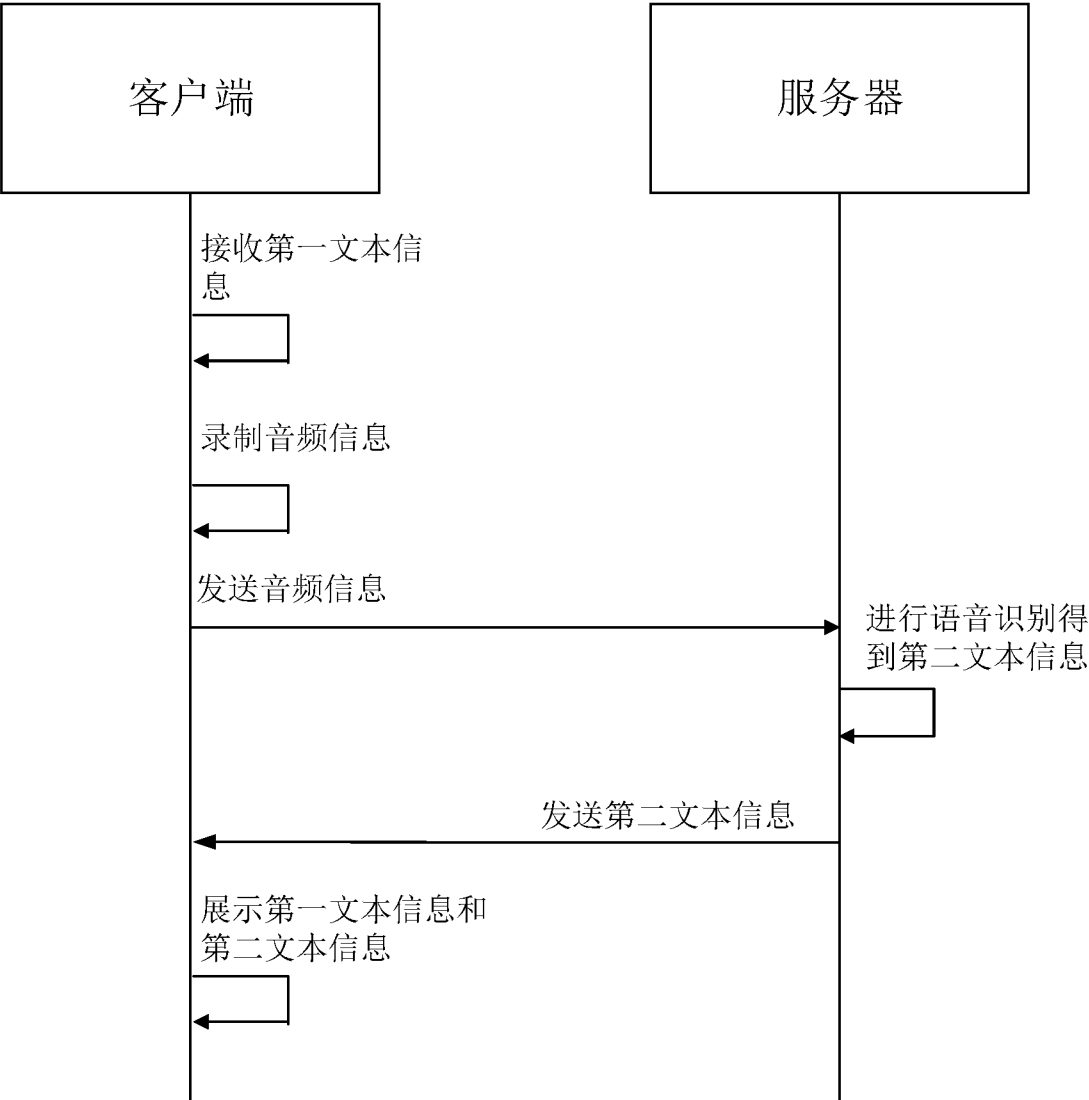


图 2

XX:XX:XX XXXX:XXXXXXXXXXXXXXXXXXXXXXXXXXXX	XX:XX:XX XXXX:XX XXXXXXXXXXXXXXXXXXXXXXXXXXXX
XX:XX:XX XXX:XX XXXXXXXXXX	XX:XX:XX XXX:XX XXXXXXXXXX
10:22:16 张三：在2015年4月份借给李四5000元，约定2015年6月份归还，到期后经张三反复几次催要还款，李四推脱没有还款。张三认为李四有能力偿还欠款	10:22:16 发言者1：我在前年四月份借给李四5000元钱。他说过两个月就还给我。结果过了两个月。我去找他。他就往后拖着。一直到现在我去找他四次。他都以各种理由推脱不还钱。实际上。我知道他有钱。他上个月还买了一辆汽车。
XX:XX:XX XX:XX XX	XX:XX:XX XX:XX XX XX

图 3

XX:XX:XX XXXX:XXXXXXXXXXXXXXXXXXXXXXXXXXXX 10:22:16 张三: 在2015年4月份借给李四5000元, 约定2015年6月份归还, 到期后经张三先后四次催要还款, 李四均推脱没有还款。张三认为李四有能力偿还欠款, 李四在2017年6月份还购买了一辆汽车 XX:XX:XX XX:XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX	XX:XX:XX XXXX:XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX XXXXXXXXXXXXXXXXXXXXXXXXXXXX XX:XX:XX XXX:XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX XXXXXXXXXX 10:22:16 发言者1: 我在前年四月份借给李四5000元钱。他说过两个月就还给我。结果过了两个月。我去找他要。他就往后拖着。一直到现在我去找他四次。他都以各种理由推脱不还钱。实际上。我知道他有钱。他上个月还买了一辆汽车。 XX:XX:XX XX:XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
---	---

图 4

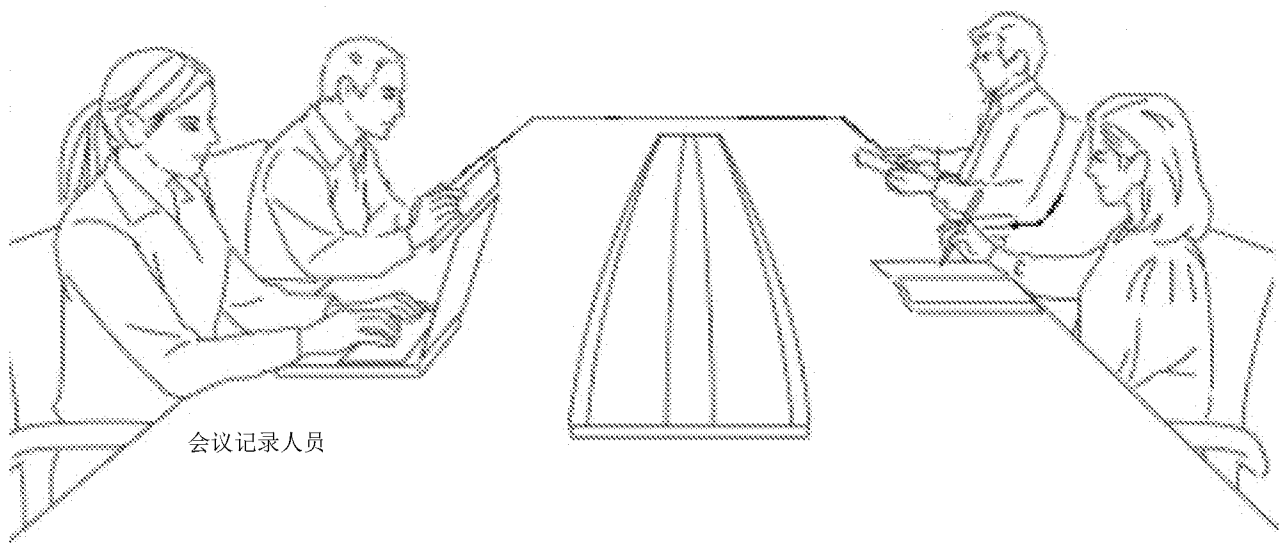


图 5

王五：

识别文本：我们这个项目时间上比较紧急。大家要盯紧进度。我们要赶在年底前完成项目验收。所以大家要抓紧时间。研发部是否可以在本月就给出完善的设计方案。

速记文本：本项目时间上较为紧急，需要各部门加快进度

丁七：

识别文本：本月底不行。我们现在只有初步的设计方案。需要采购材料进行初步验证之后。才可能给出比价完善的设计方案。所以具体提供完善的设计方案的时间。要看采购什么时间能够完成原材料采购。

钱八：

识别文本：我们本月底可以完成原材料的采购。

速记文本：本月会完成研发部需要的原材料的采购

丁七：

识别文本：如果本月底原材料可以到位的话。我们下个月中旬应该可以提供较为完善的设计方案。

速记文本：采购部在本月完成原材料采购的条件下，下个月可以提供较为完善的设计方案

王五：

识别文本：好的。那我们就定下来本月底采购部完成原材料采购。下个月中旬研发部提供较为完善的设计方案。

图 6

王五：

识别文本：我们这个项目时间上比较紧急。大家要盯紧进度。我们要赶在年底前完成项目验收。所以大家要抓紧时间。研发部是否可以在本月就给出完善的设计方案。

速记文本：本项目时间上较为紧急，要在年底完成项目验收，需要各部门加快进度

丁七：

识别文本：本月底不行。我们现在只有初步的设计方案。需要采购材料进行初步验证之后。才可能给出比价完善的设计方案。所以具体提供完善的设计方案的时间。要看采购什么时间能够完成原材料采购。

钱八：

识别文本：我们本月底可以完成原材料的采购。

速记文本：本月会完成研发部需要的原材料的采购

丁七：

识别文本：如果本月底原材料可以到位的话。我们下个月中旬应该可以提供较为完善的设计方案。

速记文本：采购部在本月完成原材料采购的条件下，下个月中旬可以提供较为完善的设计方案

王五：

识别文本：好的。那我们就定下来本月底采购部完成原材料采购。下个月中旬研发部提供较为完善的设计方案。

图 7

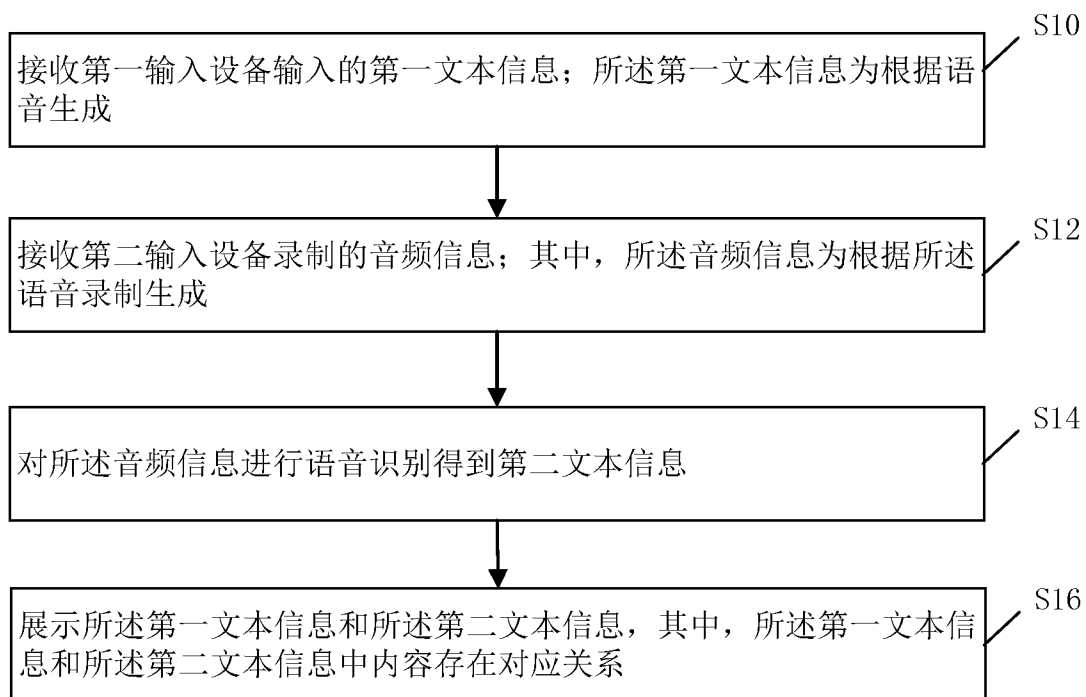


图 8

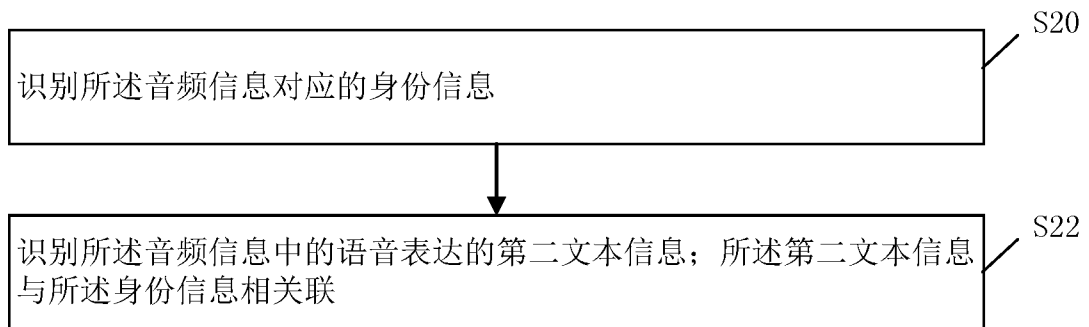


图 9

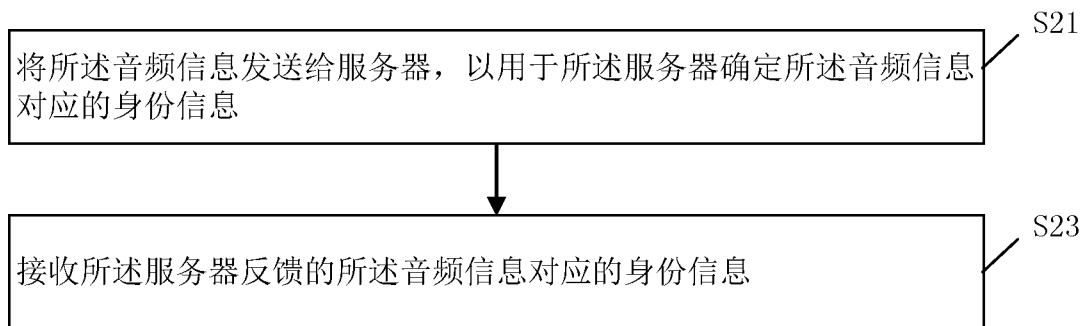


图 10

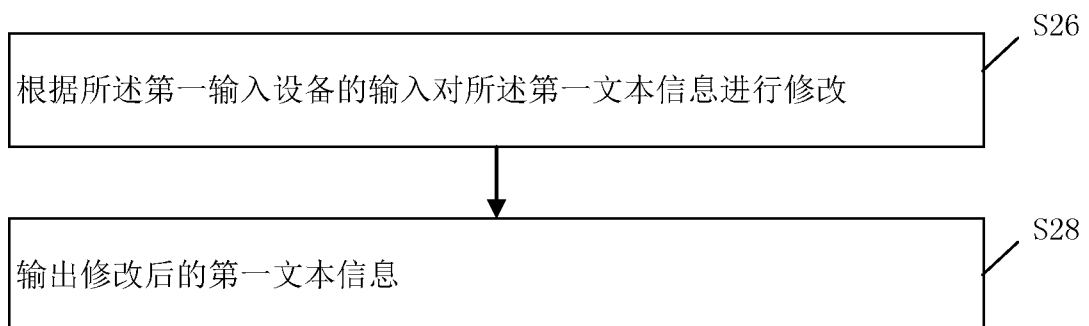


图 11

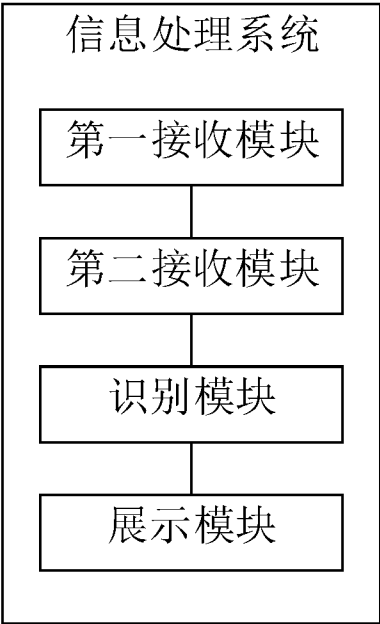


图 12

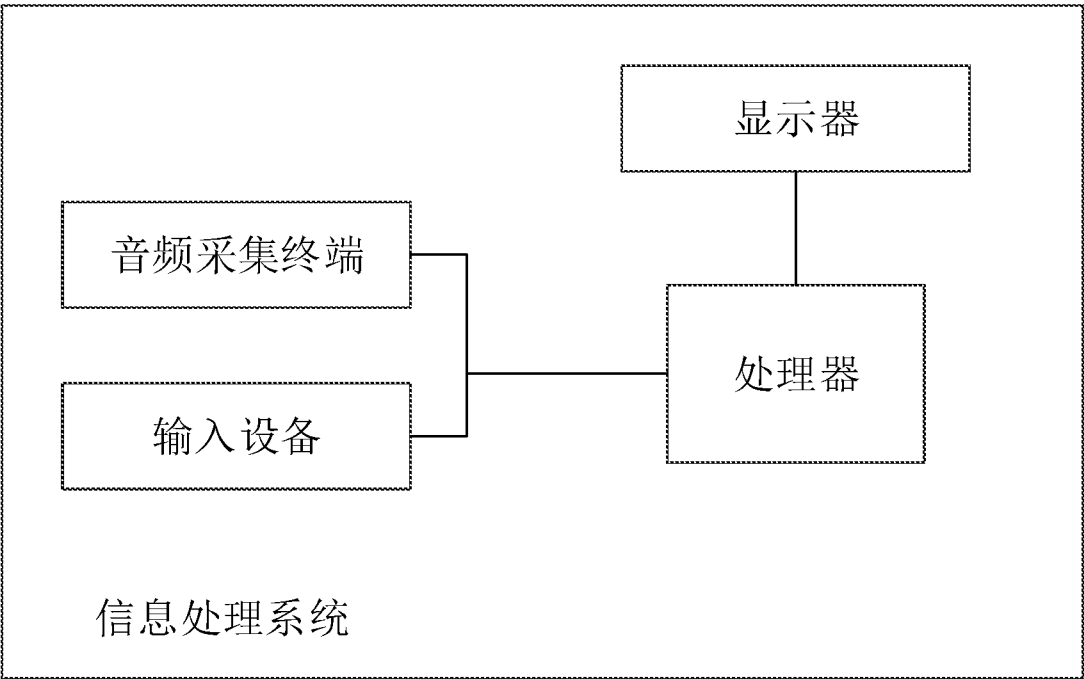


图 13

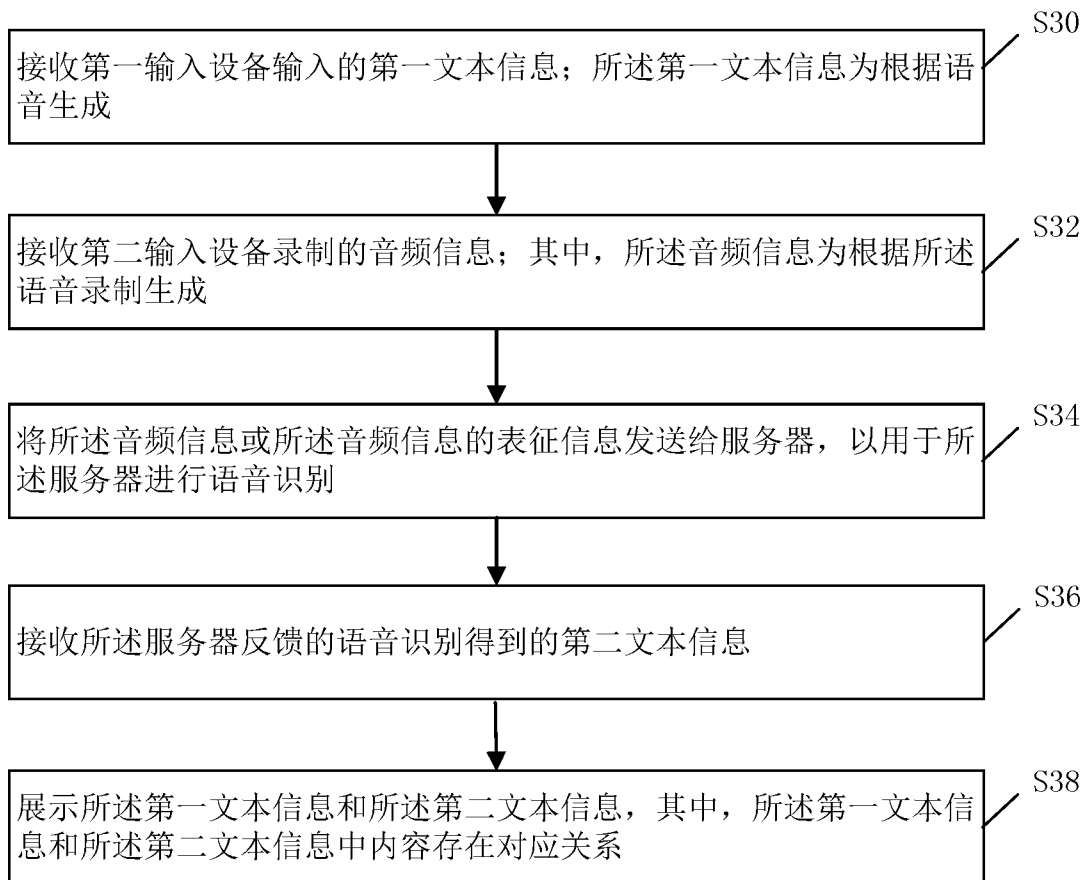


图 14

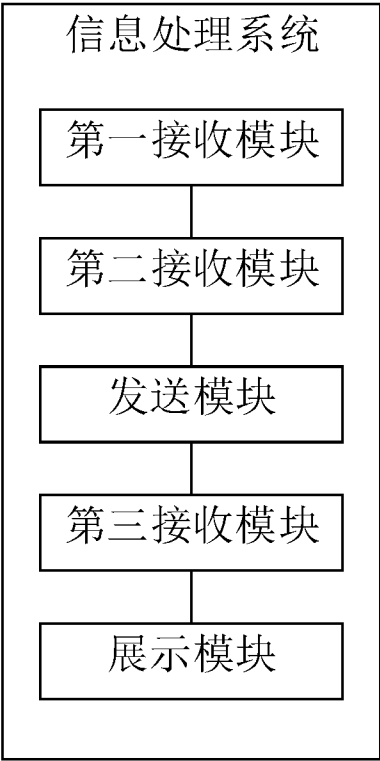


图 15

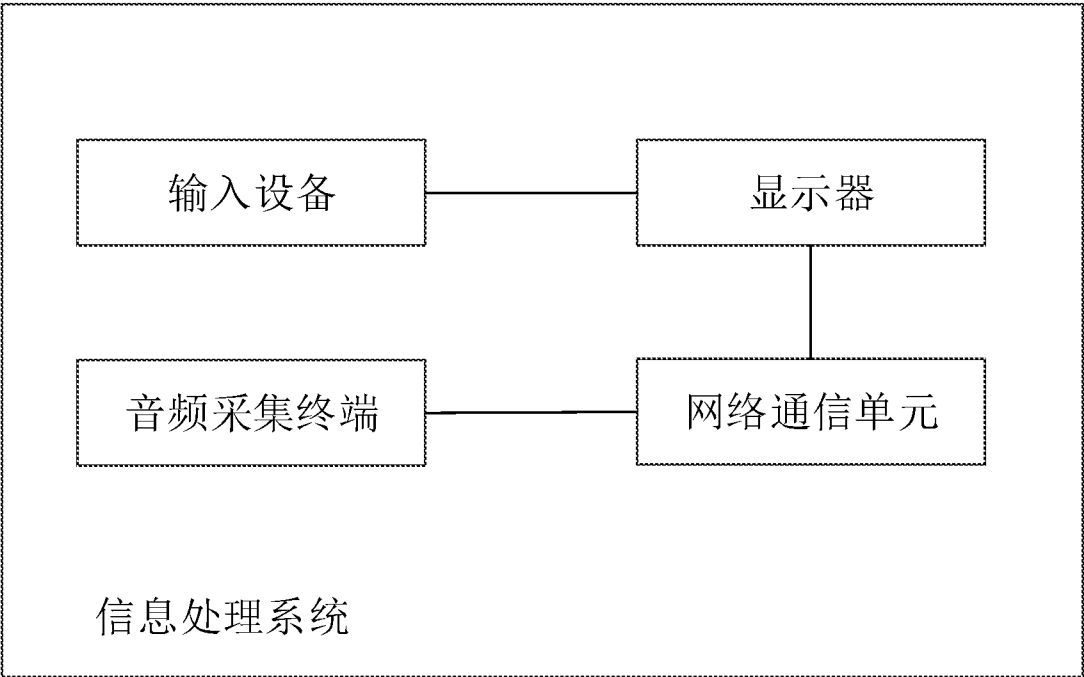


图 16

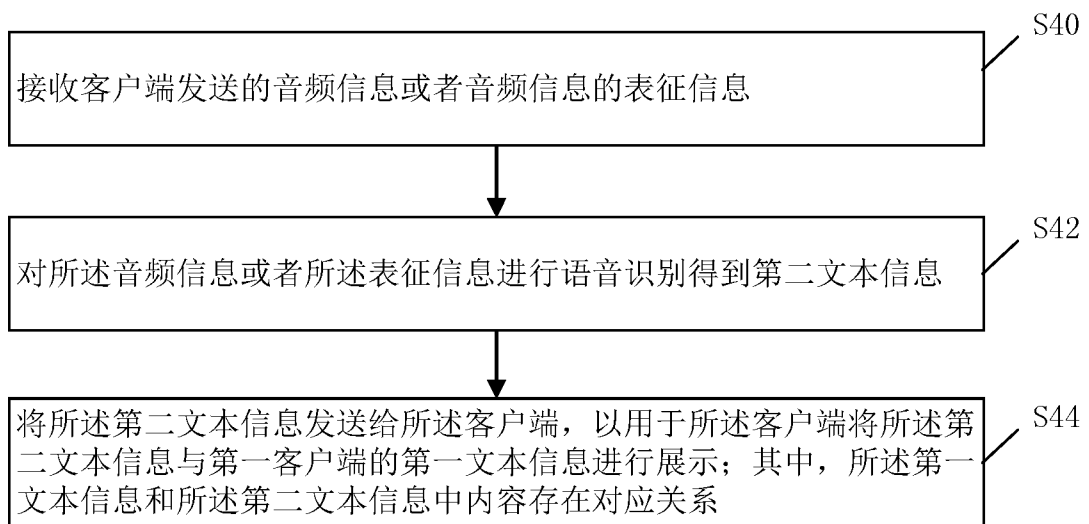


图 17

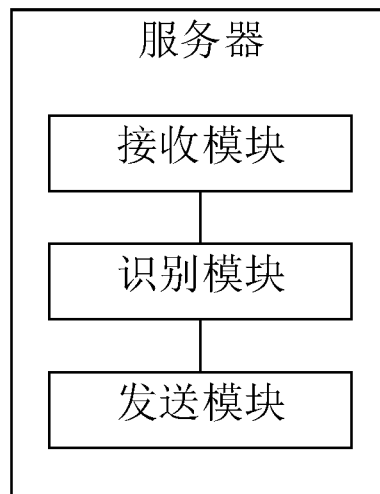


图 18

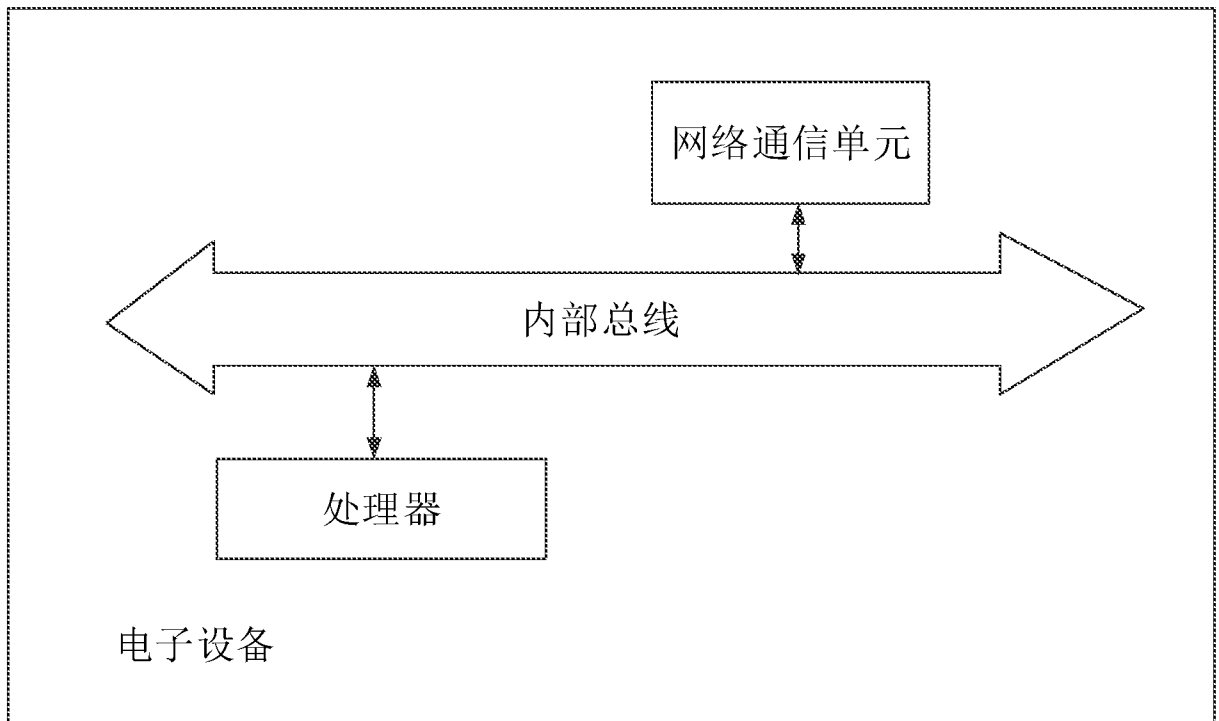


图 19

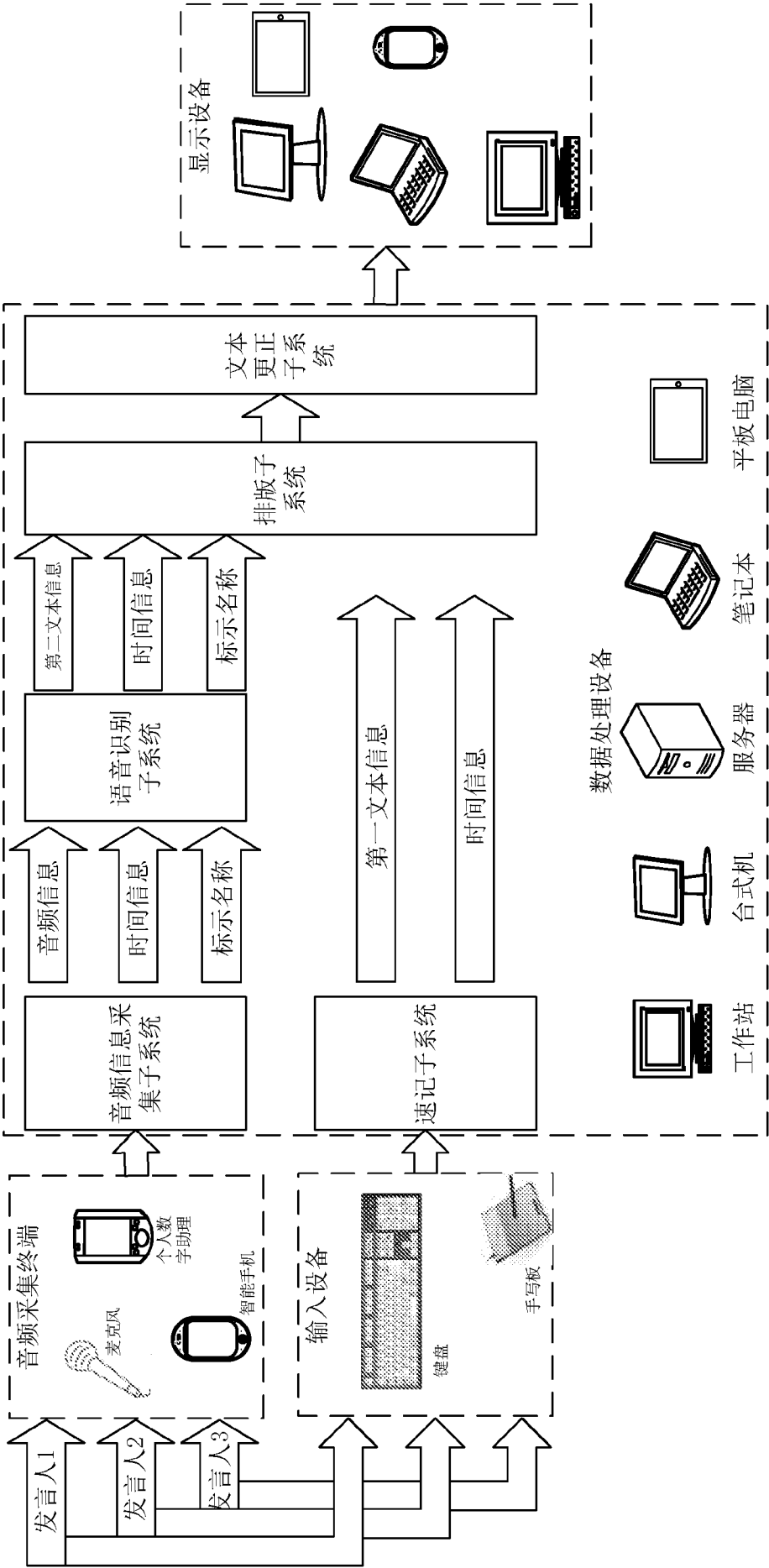


图20

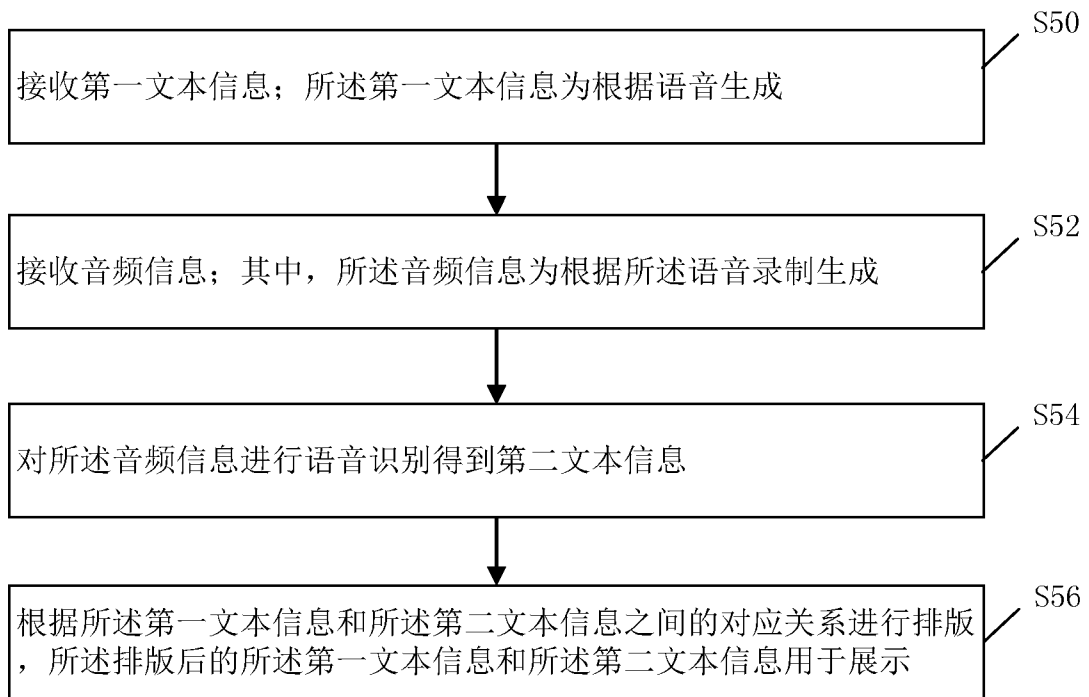


图 21

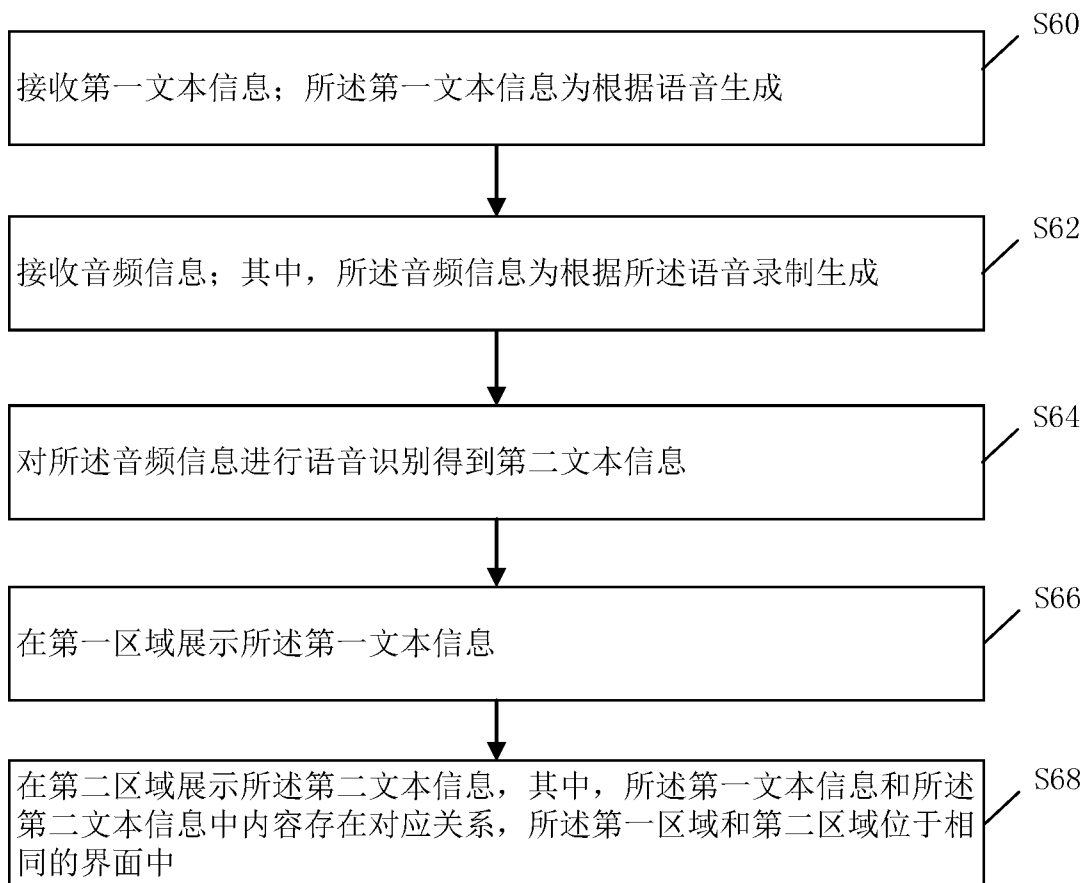


图 22

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2018/095081

A. CLASSIFICATION OF SUBJECT MATTER

G10L 15/26(2006.01)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)

G10L15

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)

USTXT; CJFD; CNTXT; CNABS; DWPI; VEN; USTXT; EPTXT; 万方, WANFANG: 文本, 书记员, 对照, 庭审, 笔记, 比对, 语音, 音频, 法院, 比较, 文字, court?, trial, display+, compar+, keyboard?, audio, word

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 102906735 A (MICROSOFT CORPORATION) 30 January 2013 (2013-01-30) description, paragraphs [0015]-[0041], and figures 1-4	1, 4-21, 23, 24, 32
Y	CN 102906735 A (MICROSOFT CORPORATION) 30 January 2013 (2013-01-30) description, paragraphs [0015]-[0041], and figures 1-4	2, 3, 22
Y	CN 106782551 A (BEIJING CHINASYS TECHNOLOGIES CO., LTD.) 31 May 2017 (2017-05-31) description, paragraphs [0038]-[0073], and figure 3	2, 3, 22
A	CN 101763382 A (NEWAUTO SILICON VALLEY VIDEO TECHNOLOGY CO., LTD.) 30 June 2010 (2010-06-30) entire document	1-32
A	US 2006167686 A1 (KAHN, J.) 27 July 2006 (2006-07-27) entire document	1-32
A	CN 101833985 A (NEWAUTO SILICON VALLEY VIDEO TECHNOLOGY CO., LTD.) 15 September 2010 (2010-09-15) entire document	1-32



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

05 September 2018

Date of mailing of the international search report

25 September 2018

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

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2018/095081

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	102906735	A	30 January 2013	CN	102906735	B	10 August 2016
				TW	201203222	A	16 January 2012
				EP	2572355	A2	27 March 2013
				US	9236047	B2	12 January 2016
				US	2011288863	A1	24 November 2011
				TW	I543150	B	21 July 2016
				WO	2011146227	A3	05 April 2012
				WO	2011146227	A2	24 November 2011
				EP	2572355	B1	27 June 2018
				EP	2572355	A4	06 November 2013
				HK	1180783	A0	25 October 2013
				HK	1180783	A1	13 October 2018
				ZA	201207966	A	23 December 2013
CN	106782551	A	31 May 2017	None			
CN	101763382	A	30 June 2010	None			
US	2006167686	A1	27 July 2006	WO	2004075027	A3	14 October 2004
				WO	2004075027	A2	02 September 2004
				US	7516070	B2	07 April 2009
CN	101833985	A	15 September 2010	None			

国际检索报告

国际申请号

PCT/CN2018/095081

A. 主题的分类

G10L 15/26(2006.01)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)

G10L15

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))

USTXT;CJFD;CNTXT;CNABS;DWPI;VEN;USTXT;EPTXT;万方: 文本, 书记员, 对照, 庭审, 笔记, 比对, 语音, 音频, 法院, 比较, 文字, court?, trial, display+, compar+, keyboard?, audio, word

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	CN 102906735 A (微软公司) 2013年 1月 30日 (2013 - 01 - 30) 说明书第[0015]-[0041]段, 附图1-4	1, 4-21, 23, 24, 32
Y	CN 102906735 A (微软公司) 2013年 1月 30日 (2013 - 01 - 30) 说明书第[0015]-[0041]段, 附图1-4	2, 3, 22
Y	CN 106782551 A (北京华夏电通科技有限公司) 2017年 5月 31日 (2017 - 05 - 31) 说明书第[0038]-[0073]段, 附图3	2, 3, 22
A	CN 101763382 A (新奥特硅谷视频技术有限公司) 2010年 6月 30日 (2010 - 06 - 30) 全文	1-32
A	US 2006167686 A1 (KAHN JONATHAN) 2006年 7月 27日 (2006 - 07 - 27) 全文	1-32
A	CN 101833985 A (新奥特硅谷视频技术有限公司) 2010年 9月 15日 (2010 - 09 - 15) 全文	1-32

☐ 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

2018年 9月 5日

国际检索报告邮寄日期

2018年 9月 25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

受权官员

周群

传真号 (86-10)62019451

电话号码 86- (0512) -88997268

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2018/095081

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	102906735	A	2013年 1月 30日	CN	102906735	B	2016年 8月 10日
				TW	201203222	A	2012年 1月 16日
				EP	2572355	A2	2013年 3月 27日
				US	9236047	B2	2016年 1月 12日
				US	2011288863	A1	2011年 11月 24日
				TW	I543150	B	2016年 7月 21日
				WO	2011146227	A3	2012年 4月 5日
				WO	2011146227	A2	2011年 11月 24日
				EP	2572355	B1	2018年 6月 27日
				EP	2572355	A4	2013年 11月 6日
				HK	1180783	A0	2013年 10月 25日
				HK	1180783	A1	2018年 10月 13日
				ZA	201207966	A	2013年 12月 23日
CN	106782551	A	2017年 5月 31日	无			
CN	101763382	A	2010年 6月 30日	无			
US	2006167686	A1	2006年 7月 27日	WO	2004075027	A3	2004年 10月 14日
				WO	2004075027	A2	2004年 9月 2日
				US	7516070	B2	2009年 4月 7日
CN	101833985	A	2010年 9月 15日	无			

表 PCT/ISA/210 (同族专利附件) (2015年1月)