

(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
2023 年 4 月 27 日 (27.04.2023)



(10) 国际公布号
WO 2023/065641 A1

(51) 国际专利分类号:
G10L 13/10 (2013.01) **G10L 13/047** (2013.01)

(21) 国际申请号: PCT/CN2022/090728

(22) 国际申请日: 2022 年 4 月 29 日 (29.04.2022)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
202111236415.5 2021 年 10 月 22 日 (22.10.2021) CN

(71) 申请人: 平安科技(深圳)有限公司(PING AN TECHNOLOGY (SHENZHEN) CO., LTD.) [CN/CN];
中国广东省深圳市福田区福田街道福

安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518000 (CN)。

(72) 发明人: 侯炜健(HOU, Weijian); 中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518000 (CN)。 陈闽川(CHEN, Minchuan); 中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518000 (CN)。

(74) 代理人: 广州嘉权专利商标事务所有限公司(JIAQUAN IP LAW); 中国广东省广州市天河区黄埔大道西 100 号富力盈泰广场 A 栋 910, Guangdong 510627 (CN)。

(54) Title: SPEECH SYNTHESIS METHOD AND SYSTEM BASED ON TEXT EDITOR, AND STORAGE MEDIUM

(54) 发明名称: 基于文本编辑器的语音合成方法、系统和存储介质

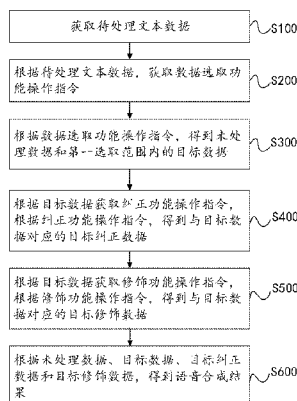


图 2

S100 Acquire text data to be processed
S200 Acquire a data selection function operation instruction according to said text data
S300 According to the data selection function operation instruction, obtain unprocessed data, and target data which is within a first selection range
S400 Acquire a correction function operation instruction according to the target data, and obtain, according to the correction function operation instruction, target corrected data corresponding to the target data
S500 Acquire a modification function operation instruction according to the target data, and obtain, according to the modification function operation instruction, target modified data corresponding to the target data
S600 Obtain a speech synthesis result according to the unprocessed data, the target data, the target corrected data and the target modified data

(57) Abstract: Disclosed are a speech synthesis method and system based on a text editor, and a storage medium. The present application can be widely applied to the technical field of artificial intelligence. The method of the present application comprises the following steps: acquiring text data to be processed; acquiring a data selection function operation instruction according to said text data; according to the data selection function operation instruction, obtaining unprocessed data, and target data which is within a first selection range, wherein said text data comprises the unprocessed data and the target data; acquiring a correction function operation instruction according to the target data, and obtaining target corrected data according to the correction function operation instruction; acquiring a modification function operation instruction according to the target data, and obtaining target modified data according to the modification function operation instruction; and obtaining a speech synthesis result according to the unprocessed data, the target data, the target corrected data and the target modified data. By means of the present application, errors caused by the fact that a speech synthesis markup language is difficult to write can be effectively reduced.

(57) 摘要: 本申请公开了一种基于文本编辑器的语音合成方法、系统和存储介质, 可广泛应用于人工智能技术领域; 本申请的方法, 包括以下步骤: 获取待处理文本数据; 根据待处理文本数据, 获取数据选取功能操作指令; 根据数据选取功能操作指令, 得到未处理数据和第一选取范围内的目标数据, 待处理文本数据包括未处理数据和目标数据; 根据目标数据获取纠正功能操作指令, 根据纠正功能操作指令, 得到目标纠正数据; 根据目标数据获取修饰功能操作指令, 根据修饰功能操作指令, 得到目标修饰数据; 根据未处理数据、目标数据、目标纠正数据和目标修饰数据, 得到语音合成结果, 本申请能够有效减少语音合成标记语言难以书写导致出错的情况。

(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, JM, 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)。

根据细则4.17的声明:

— 发明人资格(细则4.17(iv))

本国际公布:

— 包括国际检索报告(条约第21条(3))。

基于文本编辑器的语音合成方法、系统和存储介质

本申请要求于2021年10月22日提交中国专利局、申请号为202111236415.5，发明名称为“基于文本编辑器的语音合成方法、系统和存储介质”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。

技术领域

本申请涉及人工智能技术领域，尤其涉及一种基于文本编辑器的语音合成方法、系统和存储介质。

背景技术

随着人工智能（Artificial Intelligence，AI）的快速发展，语音合成技术也逐渐发展成能够利用AI技术来生成高质量的语音。例如，用户输入一段文字，通过选择需要的音库后便能很快合成极其真实的音频。目前W3C定义了语音合成标记语言（Speech Synthesis Markup Language，SSML）。

技术问题

以下是发明人意识到的现有技术的技术问题：相关技术中，语音合成标记语言规范通常都是基于可扩展标记语言（EXtensible Markup Language，XML）这种标记语言的。然而，可扩展标记语言具有严格的格式要求，故在输入语音合成标记语言时，多一个字或少一个字都有可能整个语音合成标记语言不合法，进而导致无法解析。因此，通过这种语音合成的方式，由于用户通常难以书写出规范的语音合成标记语言，导致容易出错，从而造成语音无法合成。

技术方案

第一方面，本申请实施例提供了一种基于文本编辑器的语音合成方法，包括以下步骤：

获取待处理文本数据；

根据所述待处理文本数据，获取数据选取功能操作指令；

根据所述数据选取功能操作指令，得到未处理数据和第一选取范围内的目标数据，其中，所述待处理文本数据包括所述未处理数据和所述目标数据；

根据所述目标数据获取纠正功能操作指令，根据所述纠正功能操作指令，得到与所述目标数据对应的目标纠正数据；

根据所述目标数据获取修饰功能操作指令，根据所述修饰功能操作指令，得到与所述目标数据对应的目标修饰数据；

根据所述未处理数据、所述目标数据、所述目标纠正数据和所述目标修饰数据，得到语音合成结果。

第二方面，本申请实施例提供了一种基于文本编辑器的语音合成系统，包括：

数据输入模块，用于获取待处理文本数据；

数据选取模块，用于根据所述待处理文本数据，获取数据选取功能操作指令；

目标数据获取模块，用于根据所述数据选取功能操作指令，得到未处理数据和第一选取范围内的目标数据，其中，所述待处理文本数据包括所述未处理数据和所述目标数据；

目标纠正模块，用于根据所述目标数据获取纠正功能操作指令，根据所述纠正功能操作指令，得到与所述目标数据对应的目标纠正数据；

目标修饰模块，用于根据所述目标数据获取修饰功能操作指令，根据所述修饰功能操作指令，得到与所述目标数据对应的目标修饰数据；

语音合成模块，用于根据所述未处理数据、所述目标数据、所述目标纠正数据和所述目标修饰数据，得到语音合成结果。

第三方面，本申请实施例提供了一种基于文本编辑器的语音合成系统，包括：存储器、处理器及存储在存储器上并可在处理器上运行的计算机程序，所述处理器执行所述计算机程序时实现一种基于文本编辑器的语音合成方法，其中，所述基于文本编辑器的语音合成方法包括：获取待处理文本数据；根据所述待处理文本数据，获取数据选取功能操作指令；根据所述数据选取功能操作指令，得到未处理数据和第一选取范围内的目标数据，其中，所述待处理文本数据包括所述未处理数据和所述目标数据；根据所述目标数据获取纠正功能操作指令，根据所述纠正功能操作指令，得到与所述目标数据对应的目标纠正数据；根据所述目标数据获取修饰功能操作指令，根据所述修饰功能操作指令，得到与所述目标数据对应的目标修饰数据；根据所述未处理数据、所述目标数据、所述目标纠正数据和所述目标修饰数据，得到语音合成结果。

第四方面，本申请实施例提供了一种计算机可读存储介质，存储有计算机可执行指令，所述计算机可执行指令用于执行一种基于文本编辑器的语音合成方法，其中，所述基于文本编辑器的语音合成方法包括以下步骤：获取待处理文本数据；根据所述待处理文本数据，获取数据选取功能操作指令；根据所述数据选取功能操作指令，得到未处理数据和第一选取范围内的目标数据，其中，所述待处理文本数据包括所述未处理数据和所述目标数据；根据所述目标数据获取纠正功能操作指令，根据所述纠正功能操作指令，得到与所述目标数据对应的目标纠正数据；根据所述目标数据获取修饰功能操作指令，根据所述修饰功能操作指令，得到与所述目标数据对应的目标修饰数据；根据所述未处理数据、所述目标数据、所述目标纠正数据和所述目标修饰数据，得到语音合成结果。

有益效果

本申请实施例提出的基于文本编辑器的语音合成方法、系统和存储介质，通过先获取待处理文本数据；再根据待处理文本数据来获取数据选取功能操作指令，进而得到未处理数据和第一选取范围内的目标数据，其中，待处理文本数据包括未处理数据和目标数据；之后根据目标数据获取纠正功能操作指令，再根据纠正功能操作指令，得到与目标数据对应的目标纠正数据；根据目标数据获取修饰功能操作指令，根据修饰功能操作指令，得到与目标数据对应的目标修饰数据；根据未处理数据、目标数据、目标纠正数据和目标修饰数据，得到语音合成结果。本实施例能够有效减少语音合成标记语言难以书写而导致出错的情况，保证语音合成效果。

附图说明

附图用来提供对本申请技术方案的进一步理解，并且构成说明书的一部分，与本申请的实施例一起用于解释本申请的技术方案，并不构成对本申请技术方案的限制。

图1为本申请实施例的一种文本编辑器的示意图；

图2为本申请实施例的一种基于文本编辑器的语音合成方法的流程图；

图3为本申请实施例的得到目标拼音数据的流程示意图；

图4为本申请实施例的候选拼音结果的具体实施例图；

图5为本申请实施例的目标拼音数据的具体实施例图；

图6为本申请实施例的多音字发音纠正功能操作指令的流程示意图；

图7为本申请实施例的英文数据的具体显示实施例图；

图8为本申请实施例的数字符号数据的具体显示实施例图；

图9为本申请实施例的目标纠正数字符号数据的具体实施例图；

图10为本申请实施例的得到目标修饰数据的流程示意图；

图11为本申请实施例的目标停顿数据的具体实施例图；

图 12 为本申请实施例的目标静音数据的具体实施例图；
图 13 为本申请实施例的目标特效音数据的具体实施例图；
图 14 为本申请实施例的得到目标优先修饰数据的流程示意图；
图 15 为本申请实施例的多发音人功能的具体实施例图；
图 16 为本申请实施例的局部变速功能的具体实施例图；
图 17 为本申请实施例的多种不同的修饰功能的具体实施例图；
图 18 为本申请实施例的第二选取范围内的连读数据的具体实施例图；
图 19 为本申请实施例的第二选取范围内的目标数据的具体实施例图；
图 20 为本申请实施例的连读数据的具体实施例图；
图 21 为本申请实施例的一种基于文本编辑器的语音合成系统的结构示意图。

本发明的实施方式

为了使本申请的目的、技术方案及优点更加清楚明白，以下结合附图及实施例，对本申请进行进一步详细说明。应当理解，此处所描述的具体实施例仅用以解释本申请，并不用于限定本申请。

需要说明的是，虽然在装置示意图中进行了功能模块划分，在流程图中示出了逻辑顺序，但是在某些情况下，可以以不同于装置中的模块划分，或流程图中的顺序执行所示出或描述的步骤。说明书和权利要求书及上述附图中的术语“第一”、“第二”等是用于区别类似的对象，而不必用于描述特定的顺序或先后次序。

随着人工智能的快速发展，语音合成技术也逐渐发展成能够利用 AI 技术来生成高质量的语音。相关技术中，语音合成标记语言规范通常都是基于可扩展标记语言这种标记语言的。然而，可扩展标记语言具有严格的格式要求，故在输入语音合成标记语言时，多一个字或少一个字都有可能导致整个语音合成标记语言不合法，进而导致无法解析。因此，通过这种语音合成的方式，由于用户通常难以书写出规范的语音合成标记语言，导致容易出错，从而造成语音无法合成。

基于此，本申请实施例提供了一种基于文本编辑器的语音合成方法、系统和存储介质。本申请实施例能够有效减少语音合成标记语言难以书写而导致出错的情况，保证语音合成效果。

可以理解的是，本申请实施例的语音合成方法基于文本编辑器，该文本编辑器能够用于输入待处理文本数据，待处理文本数据可以包括汉字数据、英文数据、数字数据、符号数据等等。通过在文本编辑器内输入汉字数据、英文数据、数字数据或符号数据后，能够通过数据选取功能操作指令、纠正功能操作指令和修饰功能操作指令等功能操作指令，来对上述待处理文本数据进行数据处理，以最终输出语音合成结果。可以理解的是，该语音合成结果还可以以语音合成标记语言格式输出。例如，在文本编辑器内对待处理文本数据进行数据处理后，得到语音合成结果。通过将该语音合成结果通过数据转换处理而转换为语音合成标记语言文本，从而便于将该语音合成标记语言文本导入其他系统中继续使用，提高便捷性。

一些实施例中，该文本编辑器能够基于网页实现可编辑功能。例如，该文本编辑器可类似于富文本编辑器。通过基于文本编辑器，来实现语音合成效果，解决了用户由于难以书写出规范的语音合成标记语言导致容易出错的问题。

一些实施例中，该文本编辑器可设置为可视化，从而使得对语音合成标记语言的间接编辑过程也变得可视化，提高交互性。通过采用文本编辑器，能够帮助用户定制想要的音频，提高适用性。

一些实施例中，本申请实施例的文本编辑器还可以通过 web 浏览器来实现正常运行。通过如此设置，使得用户无需下载其他软件，基于网络便可随时随地运行该文本编辑器。另一实施例中，该文本编辑器还能实现保存功能。通过选择已保存的待处理文本数据，并对已保存的待处理文本数据继续编辑，从而得到语音合成结果，例如得到音频结果，提高实用性。

本申请实施例的文本编辑器可以包括修饰功能，如试听功能、拼音/英文翻译显示功能、连读功能、局部变速功能、停顿功能、静音功能、多发音人功能、特效功能、变调功能、变音量功能、背景音功能、多语气功能（即可选客服语气）等；纠正功能，如拼音（多音字）发音纠正功能、英文发音纠正功能、数字符号发音纠正功能等，还包括新建文本功能、保存文本功能、下载功能、导出语音合成标记语言文本功能（即导出 SSML）等等。

可以理解的是，在运行该文本编辑器时，文本编辑器的各个功能都对应有功能操作指令，例如试听功能对应试听功能操作指令，局部变速功能对应局部变速功能操作指令，变音量功能对应变音量功能操作指令等。一实施例中，修饰功能操作指令包括试听功能操作指令，通过根据目标数据获取试听功能操作指令，根据试听功能操作指令，得到目标试听数据。又例如，修饰功能操作指令包括变调功能操作指令，通过根据目标数据获取变调功能操作指令，根据变调功能操作指令，得到目标变调数据，等等。通过功能操作指令获取想要的目标处理数据，增加语音合成多样性。

可以理解的是，图 1 所示的文本编辑器中，1、2、3……7 每行对应的文本数据为输入的待处理文本数据。由于需要对待处理文本数据进行数据选取，从而根据数据选取功能操作指令，得到未处理数据和目标数据。可以理解的是，目标数据即需要处理的数据，以及对目标数据进行数据处理后，能够得到目标处理数据。例如，在多音字纠错中，“乐”字表示目标数据（或拼音数据），“乐”字上方的拼音则表示目标纠正数据（或目标拼音数据）。在连读中，“长江大桥”表示目标数据（或连读数据），而“长江大桥”左右两侧的连接符号标记，则表示目标修饰数据（或目标连读数据），“长”字也可表示多音字数据，则“长”字上方的拼音则表示目标纠正数据（或目标多音字数据）。在局部变速中，数字 50、-50 可表示目标修饰数据（或目标局部变速数据），即数字的大小即变速的大小。在静音中，400ms、600ms 可表示目标修饰数据（或目标静音数据），400ms、600ms 即表示静音时长。可以理解的是，在其他实施例中，当目标修饰数据为数字时，该数字可表示为音量的大小、音调的大小、语速的快慢等，而不局限于本实施例。

具体地，参照图 2，本申请实施例提供一种基于文本编辑器的语音合成方法，包括以下步骤：

步骤 S100、获取待处理文本数据；

需说明的是，可基于文本编辑器输入汉字数据、英文数据、数字数据、符号数据或组合数据（如数字符号数据、汉字英文数据）等，该待处理文本数据可以为用户常用数据。

步骤 S200、根据待处理文本数据，获取数据选取功能操作指令；

步骤 S300、根据数据选取功能操作指令，得到未处理数据和第一选取范围内的目标数据，其中，待处理文本数据包括未处理数据和目标数据；

需说明的是，通过对待处理文本数据执行数据选取功能操作指令，从而根据数据选取功能操作指令，得到第一选取范围内的目标数据。可以理解的是，目标数据可以为待处理文本数据中的部分数据，也可以为整个待处理文本数据；若目标数据为整个待处理文本数据，则未处理数据为无；

步骤 S400、根据目标数据获取纠正功能操作指令，根据纠正功能操作指令，得到与目标数据对应的目标纠正数据；

之后，对目标数据执行纠正功能操作指令，根据纠正功能操作指令，得到与目标数据对应的目标纠正数据。本实施例对待处理文本数据中对应的目标数据存在错误数据时，需要对错误数据进行纠正。可以理解的是，对错误数据进行纠正，可以为，通过预先对待处理文本数据进行数据处理得到音频结果，例如通过试听功能以播放上述音频结果。当音频结果存在错误发音时，表示目标数据存在错误数据，例如存在发音错误或者错别字错误等等。对应的，本实施例的纠正功能操作指令可以为拼音发音纠正功能操作指令、英文发音纠正功能操作指令、数字符号发音纠正功能操作指令、错别字纠正功能操作指令等等。

步骤 S500、根据目标数据获取修饰功能操作指令，根据修饰功能操作指令，得到与目标

数据对应的目标修饰数据；

对目标数据执行修饰功能操作指令，根据修饰功能操作指令，得到与目标数据对应的目标修饰数据。本实施例对目标数据进行修饰，以使得目标数据能够带有一定的修饰效果。例如，修饰功能操作指令可以为连读功能操作指令、变速功能操作指令、场景功能操作指令、插入点功能操作指令等，从而相应得到与目标数据对应的目标修饰数据。

步骤 S600、根据未处理数据、目标数据、目标纠正数据和目标修饰数据，得到语音合成结果。

需说明的是，由于本实施例对目标数据执行了纠正功能操作指令得到目标纠正数据，对目标数据执行了修饰功能操作指令得到目标修饰数据。目标纠正数据、目标修饰数据对对应的目标数据增加了纠正效果、修饰效果，而待处理文本数据中还包括了除目标数据之外的未处理数据，输出时，结合上述未处理数据、目标数据、目标纠正数据和目标修饰数据，能够得到最终语音合成结果。可以理解的是，待处理文本数据中即使包括有未处理数据，但该未处理数据仍对应应有相应的音频或发音，因此该未处理数据也需作为语音合成结果的一部分。

可以理解的是，经过人工智能得到的语音合成技术，也无法确保能够完全按照用户的意愿来准确地读出每一个字，这不仅意味着汉字数据和英文数据都能准确发音，也意味着数字符号数据也能按照惯用读法播报。

本实施例能够作用于一个汉字数据或英文数据或数字符号数据上。用户在执行试听功能后认为某个待处理文本数据对应的发音或读法不对时，可根据该待处理文本数据，获取数据选取功能操作指令，从而得到第一选取范围内的目标数据，之后对目标数据执行纠正功能操作指令，即可进行纠正功能。

参照图 3，可以理解的是，目标数据包括拼音数据，目标纠正数据包括目标拼音数据，纠正功能操作指令包括拼音发音纠正功能操作指令，根据目标数据获取纠正功能操作指令，根据纠正功能操作指令，得到与目标数据对应的目标纠正数据，包括：

步骤 S410、根据拼音数据，得到与拼音数据对应的至少一种候选拼音结果，其中，候选拼音结果根据每一候选拼音结果对应的发音概率值进行排序；

可以理解的是，该文本编辑器能够根据每一候选拼音结果对应的发音概率值对候选拼音结果进行排序。需说明的是，根据数据选取功能操作指令，得到第一选取范围内的拼音数据。该第一选取范围可以包括单个数量的拼音数据，或者多个数量的拼音数据。以单个拼音数据为例，根据拼音数据，得到与拼音数据对应的至少一种候选拼音结果。可以理解的是，拼音数据包括单音字数据和多音字数据。当拼音数据为单音字数据时，候选拼音结果为一种，此时是针对文本编辑器对该单音字数据所对应的默认拼音记载错误时，可通过候选拼音结果来获取正确拼音。当拼音数据为多音字数据时，候选拼音结果为至少两种。此时针对文本编辑器对该多音字数据所对应的默认拼音记载错误时，可通过至少两种候选拼音结果对应的发音概率值对候选拼音结果进行排序。

步骤 S420、根据候选拼音结果，获取拼音发音纠正功能操作指令；

步骤 S430、根据拼音发音纠正功能操作指令，得到与拼音数据对应的目标拼音数据。

之后，根据候选拼音结果，获取拼音发音纠正功能操作指令，从而根据拼音发音纠正功能操作指令，得到与拼音数据对应的目标拼音数据。

可以理解的是，通过本实施例，能够对拼音数据执行拼音发音纠正功能操作指令，从而便于用户对拼音数据进行纠正，得到目标拼音数据。

参照图 4，可以理解的是，候选拼音结果可以为拼音加声调的形式进行显示。该候选拼音结果可以设置于对应的拼音数据的上方或下方，也可以设置于对应的拼音数据的一侧等等。如图 4 中“乐”字表示本实施例的拼音数据，通过选取“乐”字，根据该“乐”字，得到与“乐”字对应的至少两种候选拼音结果，本实施例有 le4、yue4、yao4、lao4 四种候选拼音结果，且该候选拼音结果根据发音概率值进行排序。其中，4 表示声调。

参照图 5，一些实施例中，根据拼音发音纠正功能操作指令，得到与“乐”字对应的目

标拼音数据，即本实施例的 yue4。该目标拼音数据和对应的拼音数据（即“乐”字）可以设置为相同的颜色，且对应的目标拼音数据或拼音数据与未处理数据的颜色不同，从而区分出目标拼音数据，以表示该拼音数据执行了拼音发音纠正功能操作指令，提高识别性。

参照图 6，可以理解的是，拼音发音纠正功能操作指令包括多音字发音纠正功能操作指令，根据拼音数据，得到与拼音数据对应的至少一种候选拼音结果，包括：

步骤 S411、根据拼音数据，获取多音字发音纠正功能操作指令；

步骤 S412、根据多音字发音纠正功能操作指令，得到与拼音数据对应的至少两种候选拼音结果。

可以理解的是，本实施例拼音数据可以包括单音字数据和多音字数据，当拼音数据为多音字数据时，获取多音字发音纠正功能操作指令，根据多音字发音纠正功能操作指令，得到与拼音数据对应的至少两种候选拼音结果，从而便于用户根据拼音数据进行多音字发音纠正，实现功能多样性。

例如参照图 4，根据拼音数据，例如“乐”，点击多音字发音纠正功能，文本编辑器则获取多音字发音纠正功能操作指令，即可自动获取该“乐”字对应的所有可能的候选拼音结果，并按候选拼音结果对应的发音概率值进行排序后，提供给用户选择。用户点击候选拼音结果后，根据候选拼音结果，文本编辑器即可获取拼音发音纠正功能操作指令，根据拼音发音纠正功能操作指令，得到与拼音数据对应的目标拼音数据。参照图 5，一实施例中，该目标拼音数据可以以特殊颜色显示在拼音数据如“乐”上方，同时拼音数据也同步变色，以便于用户直接获知哪些拼音数据执行拼音发音纠正功能操作指令。

可以理解的是，相对于汉字数据，通常确定有有限个发音，而对于英文数据，由于英文数据例如英文单词，所对应的发音通常具有多样性（如英式发音和美式发音）和不确定性（一词多音等，如阅读“read”，具有一词多音）。用户想要语音合成结果是期望的效果时，则可根据目标数据，获取编辑功能操作指令，以对目标数据输入对应的发音音标。由此，文本编辑器能够根据用户指定的发音音标来播报英文单词，如图 7 所示。

可以理解的是，当数字混合了符号后，往往可以有好几种读法。此时，可在文本编辑器内设定几种默认读法，例如按照日期读法、序列读法、算数读法、比分读法或者范围读法等。参照图 8、图 9，当用户认为数字符号数据按照文本编辑器默认读法错误或者不是用户所需时，则根据对输入的数字符号数据，获取数据选取功能操作指令，得到目标数字符号数据。例如，文本编辑器弹出目标数字符号数据对应的所有可能的读法，以供用户选择。用户选择后，文本编辑器根据目标数字符号数据，获取纠正功能操作指令，根据纠正功能操作指令，得到与目标数字符号数据对应的目标纠正数字符号数据，如图 9 所示，日期则表示目标纠正数字符号数据。

参照图 10，可以理解的是，修饰功能操作指令包括插入点功能操作指令，根据目标数据获取修饰功能操作指令，根据修饰功能操作指令，得到与目标数据对应的目标修饰数据，包括：

步骤 S510、根据目标数据，获取目标插入位置，其中，目标插入位置为目标数据对应的左侧位置和/或右侧位置；

步骤 S520、根据目标插入位置，获取插入点功能操作指令；

步骤 S530、根据插入点功能操作指令，得到与目标插入位置对应的目标修饰数据，其中，目标修饰数据包括目标停顿数据、目标静音数据、目标特效音数据中的至少一种。

可以理解的是，本实施例能够对目标数据执行插入点功能操作指令，从而使得目标数据带有修饰效果。例如，根据目标数据，获取目标数据对应的左侧位置即目标插入位置，在目标数据对应的左侧位置执行插入点功能操作指令，如执行停顿功能操作指令、静音功能操作指令、特效音功能操作指令等，以使得根据插入点功能操作指令，得到与目标数据对应的左侧位置所对应的目标修饰数据，如目标停顿数据、目标静音数据、目标特效音数据。

即参照图 11，根据停顿功能操作指令，得到与目标插入位置对应的目标停顿数据；目标

停顿数据可包括短停顿数据、长停顿数据，其中，短停顿、长停顿对应的时间可通过预设的停顿时间来确定。

参照图 12，根据静音功能操作指令，得到与目标插入位置对应的目标静音数据；图 12 中 400ms、600ms、800ms 即目标静音数据，数字大小表示静音时长。

参照图 13，根据特效音功能操作指令，得到与目标插入位置对应的目标特效音数据。目标特效音数据可以为呼吸数据、叹息数据、心跳数据、咳嗽数据、鼠标点击数据、键盘打字数据、QQ 消息提醒数据等。

参照图 14，可以理解的是，修饰功能操作指令包括连读功能操作指令、变速功能操作指令、场景功能操作指令，根据目标数据获取修饰功能操作指令，根据修饰功能操作指令，得到与目标数据对应的目标修饰数据，包括：

步骤 S550、根据目标数据，获取连读功能操作指令、变速功能操作指令以及场景功能操作指令；

步骤 S560、根据连读功能操作指令对应的优先级、变速功能操作指令对应的优先级以及场景功能操作指令对应的优先级，得到与目标数据对应的目标优先修饰数据，其中，定义连读功能操作指令对应的优先级大于变速功能操作指令对应的优先级，变速功能操作指令对应的优先级大于场景功能操作指令对应的优先级。

可以理解的是，当目标数据中需要执行多个修饰功能操作指令时，例如执行连读功能操作指令、变速功能操作指令以及场景功能操作指令，则可以根据连读功能操作指令对应的优先级、变速功能操作指令对应的优先级以及场景功能操作指令对应的优先级进行排序，以获取目标优先修饰数据。通过目标优先修饰数据得到的语音合成结果，在播放语音合成结果时，会根据修饰功能操作指令对应的优先级进行播放。

根据不同功能逻辑上的特点，以及系统合成语音时数据处理的先后顺序，当用户对一段目标数据反复应用不同的修饰功能时，可定义连读功能操作指令对应的优先级大于变速功能操作指令对应的优先级，变速功能操作指令对应的优先级大于场景功能操作指令对应的优先级，从而使得连读功能操作指令对应的连读功能的优先级大于变速功能操作指令对应的变速功能的优先级，变速功能操作指令对应的变速功能的优先级大于场景功能操作指令对应的场景功能的优先级。

需说明的是，场景包括多语气、多发音人和背景音。对应的，场景功能则包括多语气功能、多发音人功能和背景音功能。则一实施例中，多语气功能/多发音人功能/背景音功能为最外层（优先级最低），其次变速功能（例如局部变速功能），最后连读功能（优先级最高）。

参照图 15，多发音人功能表示在部分场景下用户希望合成一段二人对话甚至多人对话的语音合成结果。

多语气功能表示根据应用场景切换不同语气，例如客服、电销、播报、喜、怒、哀等。

背景音功能表示背景音乐、白噪音；在一些实施例中，背景音能够应用于有声读物场景。

可以理解的是，为了使得语音合成效果更丰富，有时需要控制一段待处理文本数据。

具体地，针对系统会错误地对一个较长的目标数据如汉字数据进行停顿，比如“武汉市长江大桥”可能错误断句为“武汉市长/江大桥”，故连读功能需确保选取范围内的汉字数据中间没有停顿。

参照图 16，对于有声读物场景，为了更好地融入情景，有时可能需要忽快忽慢、忽高忽低地朗读一段汉字数据，相比传统的语音合成技术只能对整句/段文本应用相同的参数值，本实施例可以针对选取范围内的目标数据进行局部变速功能，又或者进行变调功能、变音量功能等。

参照图 17，为本实施例应用多种不同的修饰功能时的具体实施例图。其中，客服蓉蓉表示根据多语气功能操作指令以及多语气功能操作指令对应的优先级，得到的目标多语气第三修饰数据；“123456789”左侧位置和右侧位置的连读符号标记表示根据连读功能操作指令以及连读功能操作指令对应的优先级，得到的目标连读第一修饰数据；参数值“25”表示根据

变速功能操作指令以及变速功能操作指令对应的优先级，得到的目标变速第二修饰数据。即本实施例的目标优先修饰数据包括目标连读第一修饰数据、目标变速第二修饰数据和目标多语气第三修饰数据。在其他实施例中，也可应用其他不同的修饰功能，而不局限于本实施例，在此不再赘述。

可以理解的是，一实施例中，也可以对第一选取范围内的不同区间所对应设置的目标数据，分别应用多种不同的修饰功能，而不局限于本实施例，在此不再赘述。

可以理解的是，根据数据选取功能操作指令，得到未处理数据和第一选取范围内的目标数据之后，包括：重新获取数据选取功能操作指令，并删除第一选取范围，得到第二选取范围内的目标数据。

本实施例通过重新获取数据选取功能操作指令，以使得原目标数据对应的第一选取范围被删除，之后根据重新获取的数据选取功能操作指令，能够得到第二选取范围内的目标数据，以变换不同选取范围，从而得到不同目标数据。

可以理解的是，第二选取范围可以小于第一选取范围，或大于第一选取范围，或等于第一选取范围，具体地，在重新获取数据选取功能操作指令，并删除第一选取范围，得到第二选取范围内的目标数据步骤中，可以包括多种不同实施例。

例如，一实施例中，第二选取范围可以按照“选取范围为新区间，其他拆老区间”的原则进行数据处理。同时，定义无参数的修饰功能为无参数功能（例如连读功能，即意味着连读功能对应的目标连读数据所对应的连读数据可以直接进行拆分、合并）；定义带参数的修饰功能为有参数功能（例如变速功能、变调功能、变音量功能等等修饰功能）。

下面具体描述无参数功能、有参数功能的实施例。

1、对于无参数功能，且在第二选取范围内带有（如一个或两个）同种类型的修饰功能时：

例如，在经过步骤 S100 至步骤 S500 后，即执行了连读功能操作指令之后，再执行重新获取数据选取功能操作指令步骤，此时删除第一选取范围，得到第二选取范围内的连读数据，此时经步骤 S500 得到的目标连读数据仍保留，且该目标连读数据相应设置于第二选取范围内的目标数据的左右两侧。

如图 18 所示，左侧数据表示第一选取范围的连读数据经过连读功能操作指令后，得到目标连读数据和第一选取范围的连读数据，此时再重新获取数据选取功能操作指令，则原第一选取范围将会被删除，从而得到第二选取范围内的连读数据，右侧数据则表示第二选取范围内的连读数据以及对应的目标连读数据。

即可以理解的是，一实施例中，当重新获取数据选取功能操作指令这个步骤是在经过步骤 S400 执行的纠正功能操作指令和/或步骤 S500 执行的修饰功能操作指令之后，则经过步骤 S400 得到的目标纠正数据和/或经过步骤 S500 得到的目标修饰数据仍然会保留。即本实施例是为了改变第一选取范围内的目标数据，在删除了第一选取范围后，通过保留原先的目标纠正数据和/或目标修饰数据，从而得到第二选取范围的目标数据以及对应的目标纠正数据和/或对应的目标修饰数据。即保留目标修饰数据对应的修饰功能。

2、对于有参数功能：

在经过步骤 S100 至步骤 S500 后，即执行了修饰功能操作指令之后，再执行重新获取数据选取功能操作指令，并删除第一选取范围，得到第二选取范围内的目标数据步骤。此时目标数据对应的第二选取范围改变，且对应的目标修饰数据所对应的参数值也将会改变。

例如，当第二选取范围带有不同类型的修饰功能时，或者第二选取范围带有目标修饰数据以及未处理数据时，第二选取范围对应的选取条件将包括：单侧选取条件、内部选取条件、两侧选取条件这三种条件。

具体地，第二选取范围对应的单侧选取条件表示在第一选取范围内的目标数据的左侧位置或右侧位置选取了未处理数据，且第二选取范围中包括第一选取范围内的部分目标数据；此时第一选取范围可以大于或者等于或者小于第二选取范围；

第二选取范围对应的内部选取条件表示在第一选取范围内选取部分目标数据，即第一选

取范围大于第二选取范围；

第二选取范围对应的两侧选取条件表示在第一选取范围内的目标数据的左侧位置和右侧位置均选取了未处理数据，且第二选取范围中包括第一选取范围内的所有目标数据，即第二选取范围大于第一选取范围。

参照图 19，以第一行、第二行数据为例，描述本实施例的单侧选取条件。具体地，图 19 中左侧数据表示第一选取范围的目标数据经过修饰功能操作指令（参数功能操作指令）后，得到目标修饰数据和第一选取范围的目标数据，此时再根据单侧选取条件重新获取数据选取功能操作指令，此时原第一选取范围将会被删除，从而得到第二选取范围内的目标数据，图中右侧数据则表示第二选取范围内的目标数据以及对应的目标修饰数据。

可以理解的是，此时右侧数据中，第二选取范围包括有第一区间范围和第二区间范围。第一区间范围包括第一目标数据，该第一目标数据为第一选取范围中未重新执行数据选取功能操作指令但又执行了修饰功能操作指令的数据；而第二区间范围包括第二目标数据，第二目标数据为重新执行了数据选取功能操作指令的数据。此时，第一目标数据所对应的目标修饰数据未改变，即保留了第一选取范围内的目标数据对应的修饰功能以及修饰功能对应的参数值；但对于第二目标数据而言，第二目标数据所对应的目标修饰数据中的修饰功能未改变，但参数值发生了改变，即保留了第一选取范围内的目标数据对应的修饰功能，但改变了参数值。可以理解的是，对于第二目标数据所对应的参数值可以为预设值，如图 19 中的“50”。

参照图 19，以第三行数据为例，描述本实施例的内部选取条件。具体地，图 19 中左侧数据表示第一选取范围的目标数据经过修饰功能操作指令（参数功能操作指令）后，得到目标修饰数据和第一选取范围的目标数据，此时再根据内部选取条件重新获取数据选取功能操作指令，此时原第一选取范围将会被删除，从而得到第二选取范围内的目标数据，图中右侧数据则表示第二选取范围内的目标数据以及对应的目标修饰数据。

可以理解的是，此时右侧数据中，第二选取范围包括有第一区间范围和第二区间范围。且第一区间范围包括第一目标数据，而第二区间范围包括第二目标数据；第一目标数据、第二目标数据所对应的修饰功能、参数值可参照上述单侧选取条件的描述，在此不再赘述。可以理解的是，由于本实施例是在第一选取范围内的中部选取了部分目标数据，从而得到三组与目标数据对应的目标修饰数据，但本实施例的参数值 25 对应的“1”、“4”均属于第一目标数据，即第一选取范围中未重新执行数据选取功能操作指令但又执行了修饰功能操作指令的数据。

参照图 19，以第四行数据为例，描述本实施例的两侧选取条件。具体地，图 19 中左侧数据表示第一选取范围的目标数据经过修饰功能操作指令（参数功能操作指令）后，得到目标修饰数据和第一选取范围的目标数据，此时再根据两侧选取条件重新获取数据选取功能操作指令，此时原第一选取范围将会被删除，从而得到第二选取范围内的目标数据，图中右侧数据则表示第二选取范围内的目标数据以及对应的目标修饰数据。

可以理解的是，此时右侧数据中，第二选取范围包括有第二区间范围。第二区间范围包括第二目标数据，第二目标数据为重新执行了数据选取功能操作指令的数据。此时，第二目标数据所对应的目标修饰数据中的修饰功能未改变，但参数值发生了改变，即保留了第一选取范围内的目标数据对应的修饰功能，但改变了参数值。可以理解的是，对于第二目标数据所对应的参数值可以为预设值，如图 19 中的“50”。

可以理解的是，图 19 中的参数值的大小可以表示变速的快慢、变调的高低、变音量的大小等等。

可以理解的是，在另外一实施例中，在步骤 S300 后，若直接执行了重新获取数据选取功能操作指令，并删除第一选取范围，得到第二选取范围内的目标数据这个步骤，之后再执行步骤 S400 和/或步骤 S500 时，此时的目标纠正数据、目标修饰数据是对应的第二选取范围内的目标数据。本实施例与上述提到的实施例具有较大不同。

通过如此设置，能够改变语音合成结果，使得用户能够自定义想要的音频，提高适用性。

可以理解的是，修饰功能操作指令包括连读功能操作指令，目标数据包括连读数据，根据目标数据获取修饰功能操作指令，根据修饰功能操作指令，得到与目标数据对应的目标修饰数据，包括：若第二选取范围内的连读数据对应的数量为一个，则撤销连读功能操作指令。

本实施例中，修饰功能操作指令包括连读功能操作指令。需要说明的是，当目标数据包括连读数据时，若第二选取范围内的连读数据对应的数量为一个，则对该连读数据执行连读功能操作指令是无意义的。

在一些实施例中，当第二选取范围内的连读数据对应的数量为一个时，则撤销连读功能操作指令。即表示文本编辑器拒绝执行连读功能操作指令。

在一些实施例中，根据连读数据获取连读功能操作指令，根据连读功能操作指令，得到与连读数据对应的目标连读数据。对于目标连读数据，可以采用连读符号标记表示。该目标连读数据即连读符号标记对应于连读数据，可以设置于连读数据的左右两侧。当对目标连读数据执行删除功能操作指令时，也相应撤销连读功能操作指令即撤销连读功能。

参照图 20，例如，以左侧的第二个“连续”作为连读数据，其余的“连续”为未处理数据为例，当删除设置于连读数据“连续”的左右两侧中任意一个目标连读数据（即连读符号标记）时或者均删除连读数据“连续”的左右两侧的目标连读数据时，则表示对该目标连读数据执行删除功能操作指令，此时，撤销连读功能操作指令即撤销连读功能，该连读数据将转换为未处理数据，如图 20 的箭头右侧的数据。又例如，当第二选取范围内的连读数据对应的数量为多个时，例如连读数据对应的数量为至少两个，即对于连读数据“连续”而言，其数量为两个，“连”为一个，“续”为一个，之后对连读数据执行删除功能操作指令，直至第二选取范围内的连读数据对应的数量为一个，此时，也将撤销连读功能操作指令，使得该连读数据转换为未处理数据，如图 20 的箭头右侧第三行的数据。

可以理解的是，上述基于文本编辑器的语音合成方法可应用于智能诊疗、远程会诊。

参照图 21，本申请一个实施例还提供了一种基于文本编辑器的语音合成系统，包括：

数据输入模块 100，用于获取待处理文本数据；

数据选取模块 200，用于根据待处理文本数据，获取数据选取功能操作指令；

目标数据获取模块 300，用于根据数据选取功能操作指令，得到未处理数据和第一选取范围内的目标数据，其中，待处理文本数据包括未处理数据和目标数据；

目标纠正模块 400，用于根据目标数据获取纠正功能操作指令，根据纠正功能操作指令，得到与目标数据对应的目标纠正数据；

目标修饰模块 500，用于根据目标数据获取修饰功能操作指令，根据修饰功能操作指令，得到与目标数据对应的目标修饰数据；

语音合成模块 600，用于根据未处理数据、目标数据、目标纠正数据和目标修饰数据，得到语音合成结果。

需说明的是，本申请方法实施例的内容均适用于本系统实施例，本系统实施例所实现的功能与上述方法实施例相同，并且达到的有益效果与上述方法达到的有益效果也相同，在此不再赘述。

另外，本申请一个实施例还提供了一种基于文本编辑器的语音合成系统，该基于文本编辑器的语音合成系统包括：存储器、处理器及存储在存储器上并可在处理器上运行的计算机程序，所述处理器执行所述计算机程序时实现一种基于文本编辑器的语音合成方法，其中，所述基于文本编辑器的语音合成方法包括：获取待处理文本数据；根据所述待处理文本数据，获取数据选取功能操作指令；根据所述数据选取功能操作指令，得到未处理数据和第一选取范围内的目标数据，其中，所述待处理文本数据包括所述未处理数据和所述目标数据；根据所述目标数据获取纠正功能操作指令，根据所述纠正功能操作指令，得到与所述目标数据对应的目标纠正数据；根据所述目标数据获取修饰功能操作指令，根据所述修饰功能操作指令，得到与所述目标数据对应的目标修饰数据；根据所述未处理数据、所述目标数据、所述目标纠正数据和所述目标修饰数据，得到语音合成结果。

处理器和存储器可以通过总线或者其他方式连接。

存储器作为一种非暂态计算机可读存储介质，可用于存储非暂态软件程序以及非暂态性计算机可执行程序。此外，存储器可以包括高速随机存取存储器，还可以包括非暂态存储器，例如至少一个磁盘存储器件、闪存器件、或其他非暂态固态存储器件。在一些实施方式中，存储器可选包括相对于处理器远程设置的存储器，这些远程存储器可以通过网络连接至该处理器。上述网络的实例包括但不限于互联网、企业内部网、局域网、移动通信网及其组合。

需要说明的是，本实施例中的基于文本编辑器的语音合成系统，可以应用为如上述实施例的基于文本编辑器的语音合成方法，本实施例中的基于文本编辑器的语音合成系统和如上述实施例的基于文本编辑器的语音合成方法具有相同的发明构思，因此这些实施例具有相同的实现原理以及技术效果，此处不再详述。

实现上述实施例的基于文本编辑器的语音合成方法所需的非暂态软件程序以及指令存储在存储器中，当被处理器执行时，执行上述实施例中的基于文本编辑器的语音合成方法，例如，执行以上描述的图 2 中的方法步骤 S100 至 S600、图 3 中的方法步骤 S410 至 S430、图 6 中的方法步骤 S411 至 S412、图 10 中的方法步骤 S510 至 S530、图 14 中的方法步骤 S550 至 S560。

以上所描述的基于文本编辑器的语音合成系统实施例仅仅是示意性的，其中作为分离部件说明的单元可以是或者也可以不是物理上分开的，即可以位于一个地方，或者也可以分布到多个网络单元上。可以根据实际的需要选择其中的部分或者全部模块来实现本实施例方案的目的。

此外，本申请一个实施例还提供了一种计算机可读存储介质，该计算机可读存储介质存储有计算机可执行指令，该计算机可执行指令被一个处理器或控制器执行，例如，被上述基于文本编辑器的语音合成系统实施例中的一个处理器执行，可使得上述处理器执行上述实施例中的基于文本编辑器的语音合成方法，例如，执行以上描述的图 2 中的方法步骤 S100 至 S600、图 3 中的方法步骤 S410 至 S430、图 6 中的方法步骤 S411 至 S412、图 10 中的方法步骤 S510 至 S530、图 14 中的方法步骤 S550 至 S560。此外，所述计算机可读存储介质可以是非易失性，也可以是易失性。

本领域普通技术人员可以理解，上文中所公开方法中的全部或某些步骤、系统可以被实施为软件、固件、硬件及其适当的组合。某些物理组件或所有物理组件可以被实施为由处理器，如中央处理器、数字信号处理器或微处理器执行的软件，或者被实施为硬件，或者被实施为集成电路，如专用集成电路。这样的软件可以分布在计算机可读介质上，计算机可读介质可以包括计算机存储介质（或非暂时性介质）和通信介质（或暂时性介质）。如本领域普通技术人员公知的，术语计算机存储介质包括在用于存储信息（诸如计算机可读指令、数据结构、程序模块或其他数据）的任何方法或技术中实施的易失性和非易失性、可移除和不可移除介质。计算机存储介质包括但不限于 RAM、ROM、EEPROM、闪存或其他存储器技术、CD-ROM、数字多功能盘（DVD）或其他光盘存储、磁盒、磁带、磁盘存储或其他磁存储装置、或者可以用于存储期望的信息并且可以被计算机访问的任何其他的介质。此外，本领域普通技术人员公知的是，通信介质通常包含计算机可读指令、数据结构、程序模块或者诸如载波或其他传输机制之类的调制数据信号中的其他数据，并且可包括任何信息递送介质。

以上是对本申请的较佳实施进行了具体说明，但本申请并不局限于上述实施方式，熟悉本领域的技术人员在不违背本申请精神的前提下还可作出种种的等同变形或替换，这些等同的变形或替换均包含在本申请权利要求所限定的范围内。

权 利 要 求 书

1. 一种基于文本编辑器的语音合成方法，其中，包括以下步骤：

获取待处理文本数据；

根据所述待处理文本数据，获取数据选取功能操作指令；

根据所述数据选取功能操作指令，得到未处理数据和第一选取范围内的目标数据，其中，所述待处理文本数据包括所述未处理数据和所述目标数据；

根据所述目标数据获取纠正功能操作指令，根据所述纠正功能操作指令，得到与所述目标数据对应的目标纠正数据；

根据所述目标数据获取修饰功能操作指令，根据所述修饰功能操作指令，得到与所述目标数据对应的目标修饰数据；

根据所述未处理数据、所述目标数据、所述目标纠正数据和所述目标修饰数据，得到语音合成结果。

2. 根据权利要求1所述的基于文本编辑器的语音合成方法，其中，所述目标数据包括拼音数据，所述目标纠正数据包括目标拼音数据，所述纠正功能操作指令包括拼音发音纠正功能操作指令，

所述根据所述目标数据获取纠正功能操作指令，根据所述纠正功能操作指令，得到与所述目标数据对应的目标纠正数据，包括：

根据所述拼音数据，得到与所述拼音数据对应的至少一种候选拼音结果，其中，所述候选拼音结果根据每一所述候选拼音结果对应的发音概率值进行排序；

根据所述候选拼音结果，获取拼音发音纠正功能操作指令；

根据所述拼音发音纠正功能操作指令，得到与所述拼音数据对应的目标拼音数据。

3. 根据权利要求2所述的基于文本编辑器的语音合成方法，其中，所述拼音发音纠正功能操作指令包括多音字发音纠正功能操作指令，

所述根据所述拼音数据，得到与所述拼音数据对应的至少一种候选拼音结果，包括：

根据所述拼音数据，获取多音字发音纠正功能操作指令；

根据所述多音字发音纠正功能操作指令，得到与所述拼音数据对应的至少两种所述候选拼音结果。

4. 根据权利要求1所述的基于文本编辑器的语音合成方法，其中，所述修饰功能操作指令包括插入点功能操作指令，

所述根据所述目标数据获取修饰功能操作指令，根据所述修饰功能操作指令，得到与所述目标数据对应的目标修饰数据，包括：

根据所述目标数据，获取目标插入位置，其中，所述目标插入位置为所述目标数据对应的左侧位置和/或右侧位置；

根据所述目标插入位置，获取所述插入点功能操作指令；

根据所述插入点功能操作指令，得到与所述目标插入位置对应的目标修饰数据，其中，所述目标修饰数据包括目标停顿数据、目标静音数据、目标特效音数据中的至少一种。

5. 根据权利要求1所述的基于文本编辑器的语音合成方法，其中，所述修饰功能操作指令包括连读功能操作指令、变速功能操作指令、场景功能操作指令，

所述根据所述目标数据获取修饰功能操作指令，根据所述修饰功能操作指令，得到与所述目标数据对应的目标修饰数据，包括：

根据所述目标数据，获取连读功能操作指令、变速功能操作指令以及场景功能操作指令；

根据所述连读功能操作指令对应的优先级、所述变速功能操作指令对应的优先级以及所述场景功能操作指令对应的优先级，得到与所述目标数据对应的目标优先修饰数据，其中，定义所述连读功能操作指令对应的优先级大于所述变速功能操作指令对应的优先级，所述变速功能操作指令对应的优先级大于所述场景功能操作指令对应的优先级。

6. 根据权利要求1所述的基于文本编辑器的语音合成方法, 其中, 在所述根据所述数据选取功能操作指令, 得到未处理数据和第一选取范围内的目标数据之后, 所述语音合成方法还包括:

重新获取所述数据选取功能操作指令, 并删除所述第一选取范围, 得到第二选取范围内的目标数据。

7. 根据权利要求6所述的基于文本编辑器的语音合成方法, 其中, 所述修饰功能操作指令包括连读功能操作指令, 所述目标数据包括连读数据, 所述根据所述目标数据获取修饰功能操作指令, 根据所述修饰功能操作指令, 得到与所述目标数据对应的目标修饰数据, 包括:

若所述第二选取范围内的连读数据对应的数量为一个, 则撤销所述连读功能操作指令。

8. 一种基于文本编辑器的语音合成系统, 其中, 包括:

数据输入模块, 用于获取待处理文本数据;

数据选取模块, 用于根据所述待处理文本数据, 获取数据选取功能操作指令;

目标数据获取模块, 用于根据所述数据选取功能操作指令, 得到未处理数据和第一选取范围内的目标数据, 其中, 所述待处理文本数据包括所述未处理数据和所述目标数据;

目标纠正模块, 用于根据所述目标数据获取纠正功能操作指令, 根据所述纠正功能操作指令, 得到与所述目标数据对应的目标纠正数据;

目标修饰模块, 用于根据所述目标数据获取修饰功能操作指令, 根据所述修饰功能操作指令, 得到与所述目标数据对应的目标修饰数据;

语音合成模块, 用于根据所述未处理数据、所述目标数据、所述目标纠正数据和所述目标修饰数据, 得到语音合成结果。

9. 一种基于文本编辑器的语音合成系统, 其中, 包括: 存储器、处理器及存储在存储器上并可在处理器上运行的计算机程序, 所述处理器执行所述计算机程序时实现一种基于文本编辑器的语音合成方法:

其中, 所述基于文本编辑器的语音合成方法包括:

获取待处理文本数据;

根据所述待处理文本数据, 获取数据选取功能操作指令;

根据所述数据选取功能操作指令, 得到未处理数据和第一选取范围内的目标数据, 其中, 所述待处理文本数据包括所述未处理数据和所述目标数据;

根据所述目标数据获取纠正功能操作指令, 根据所述纠正功能操作指令, 得到与所述目标数据对应的目标纠正数据;

根据所述目标数据获取修饰功能操作指令, 根据所述修饰功能操作指令, 得到与所述目标数据对应的目标修饰数据;

根据所述未处理数据、所述目标数据、所述目标纠正数据和所述目标修饰数据, 得到语音合成结果。

10. 根据权利要求9所述的基于文本编辑器的语音合成系统, 其中, 所述目标数据包括拼音数据, 所述目标纠正数据包括目标拼音数据, 所述纠正功能操作指令包括拼音发音纠正功能操作指令,

所述根据所述目标数据获取纠正功能操作指令, 根据所述纠正功能操作指令, 得到与所述目标数据对应的目标纠正数据, 包括:

根据所述拼音数据, 得到与所述拼音数据对应的至少一种候选拼音结果, 其中, 所述候选拼音结果根据每一所述候选拼音结果对应的发音概率值进行排序;

根据所述候选拼音结果, 获取拼音发音纠正功能操作指令;

根据所述拼音发音纠正功能操作指令, 得到与所述拼音数据对应的目标拼音数据。

11. 根据权利要求10所述的基于文本编辑器的语音合成系统, 其中, 所述拼音发音纠正功能操作指令包括多音字发音纠正功能操作指令,

所述根据所述拼音数据, 得到与所述拼音数据对应的至少一种候选拼音结果, 包括:

根据所述拼音数据,获取多音字发音纠正功能操作指令;

根据所述多音字发音纠正功能操作指令,得到与所述拼音数据对应的至少两种所述候选拼音结果。

12. 根据权利要求9所述的基于文本编辑器的语音合成系统,其中,所述修饰功能操作指令包括插入点功能操作指令,

所述根据所述目标数据获取修饰功能操作指令,根据所述修饰功能操作指令,得到与所述目标数据对应的目标修饰数据,包括:

根据所述目标数据,获取目标插入位置,其中,所述目标插入位置为所述目标数据对应的左侧位置和/或右侧位置;

根据所述目标插入位置,获取所述插入点功能操作指令;

根据所述插入点功能操作指令,得到与所述目标插入位置对应的目标修饰数据,其中,所述目标修饰数据包括目标停顿数据、目标静音数据、目标特效音数据中的至少一种。

13. 根据权利要求9所述的基于文本编辑器的语音合成系统,其中,所述修饰功能操作指令包括连读功能操作指令、变速功能操作指令、场景功能操作指令,

所述根据所述目标数据获取修饰功能操作指令,根据所述修饰功能操作指令,得到与所述目标数据对应的目标修饰数据,包括:

根据所述目标数据,获取连读功能操作指令、变速功能操作指令以及场景功能操作指令;

根据所述连读功能操作指令对应的优先级、所述变速功能操作指令对应的优先级以及所述场景功能操作指令对应的优先级,得到与所述目标数据对应的目标优先修饰数据,其中,定义所述连读功能操作指令对应的优先级大于所述变速功能操作指令对应的优先级,所述变速功能操作指令对应的优先级大于所述场景功能操作指令对应的优先级。

14. 根据权利要求9所述的基于文本编辑器的语音合成系统,其中,在所述根据所述数据选取功能操作指令,得到未处理数据和第一选取范围内的目标数据之后,所述语音合成方法还包括:

重新获取所述数据选取功能操作指令,并删除所述第一选取范围,得到第二选取范围内的目标数据。

15. 一种计算机可读存储介质,其中,存储有计算机可执行指令,所述计算机可执行指令用于执行一种基于文本编辑器的语音合成方法,其中,所述基于文本编辑器的语音合成方法包括以下步骤:

获取待处理文本数据;

根据所述待处理文本数据,获取数据选取功能操作指令;

根据所述数据选取功能操作指令,得到未处理数据和第一选取范围内的目标数据,其中,所述待处理文本数据包括所述未处理数据和所述目标数据;

根据所述目标数据获取纠正功能操作指令,根据所述纠正功能操作指令,得到与所述目标数据对应的目标纠正数据;

根据所述目标数据获取修饰功能操作指令,根据所述修饰功能操作指令,得到与所述目标数据对应的目标修饰数据;

根据所述未处理数据、所述目标数据、所述目标纠正数据和所述目标修饰数据,得到语音合成结果。

16. 根据权利要求15所述的计算机可读存储介质,其中,所述目标数据包括拼音数据,所述目标纠正数据包括目标拼音数据,所述纠正功能操作指令包括拼音发音纠正功能操作指令,

所述根据所述目标数据获取纠正功能操作指令,根据所述纠正功能操作指令,得到与所述目标数据对应的目标纠正数据,包括:

根据所述拼音数据,得到与所述拼音数据对应的至少一种候选拼音结果,其中,所述候选拼音结果根据每一所述候选拼音结果对应的发音概率值进行排序;

根据所述候选拼音结果，获取拼音发音纠正功能操作指令；

根据所述拼音发音纠正功能操作指令，得到与所述拼音数据对应的目标拼音数据。

17. 根据权利要求 16 所述的计算机可读存储介质，其中，所述拼音发音纠正功能操作指令包括多音字发音纠正功能操作指令，

所述根据所述拼音数据，得到与所述拼音数据对应的至少一种候选拼音结果，包括：

根据所述拼音数据，获取多音字发音纠正功能操作指令；

根据所述多音字发音纠正功能操作指令，得到与所述拼音数据对应的至少两种所述候选拼音结果。

18. 根据权利要求 15 所述的计算机可读存储介质，其中，所述修饰功能操作指令包括插入点功能操作指令，

所述根据所述目标数据获取修饰功能操作指令，根据所述修饰功能操作指令，得到与所述目标数据对应的目标修饰数据，包括：

根据所述目标数据，获取目标插入位置，其中，所述目标插入位置为所述目标数据对应的左侧位置和/或右侧位置；

根据所述目标插入位置，获取所述插入点功能操作指令；

根据所述插入点功能操作指令，得到与所述目标插入位置对应的目标修饰数据，其中，所述目标修饰数据包括目标停顿数据、目标静音数据、目标特效音数据中的至少一种。

19. 根据权利要求 15 所述的计算机可读存储介质，其中，所述修饰功能操作指令包括连读功能操作指令、变速功能操作指令、场景功能操作指令，

所述根据所述目标数据获取修饰功能操作指令，根据所述修饰功能操作指令，得到与所述目标数据对应的目标修饰数据，包括：

根据所述目标数据，获取连读功能操作指令、变速功能操作指令以及场景功能操作指令；

根据所述连读功能操作指令对应的优先级、所述变速功能操作指令对应的优先级以及所述场景功能操作指令对应的优先级，得到与所述目标数据对应的目标优先修饰数据，其中，定义所述连读功能操作指令对应的优先级大于所述变速功能操作指令对应的优先级，所述变速功能操作指令对应的优先级大于所述场景功能操作指令对应的优先级。

20. 根据权利要求 15 所述的计算机可读存储介质，其中，在所述根据所述数据选取功能操作指令，得到未处理数据和第一选取范围内的目标数据之后，所述语音合成方法还包括：

重新获取所述数据选取功能操作指令，并删除所述第一选取范围，得到第二选取范围内的目标数据。

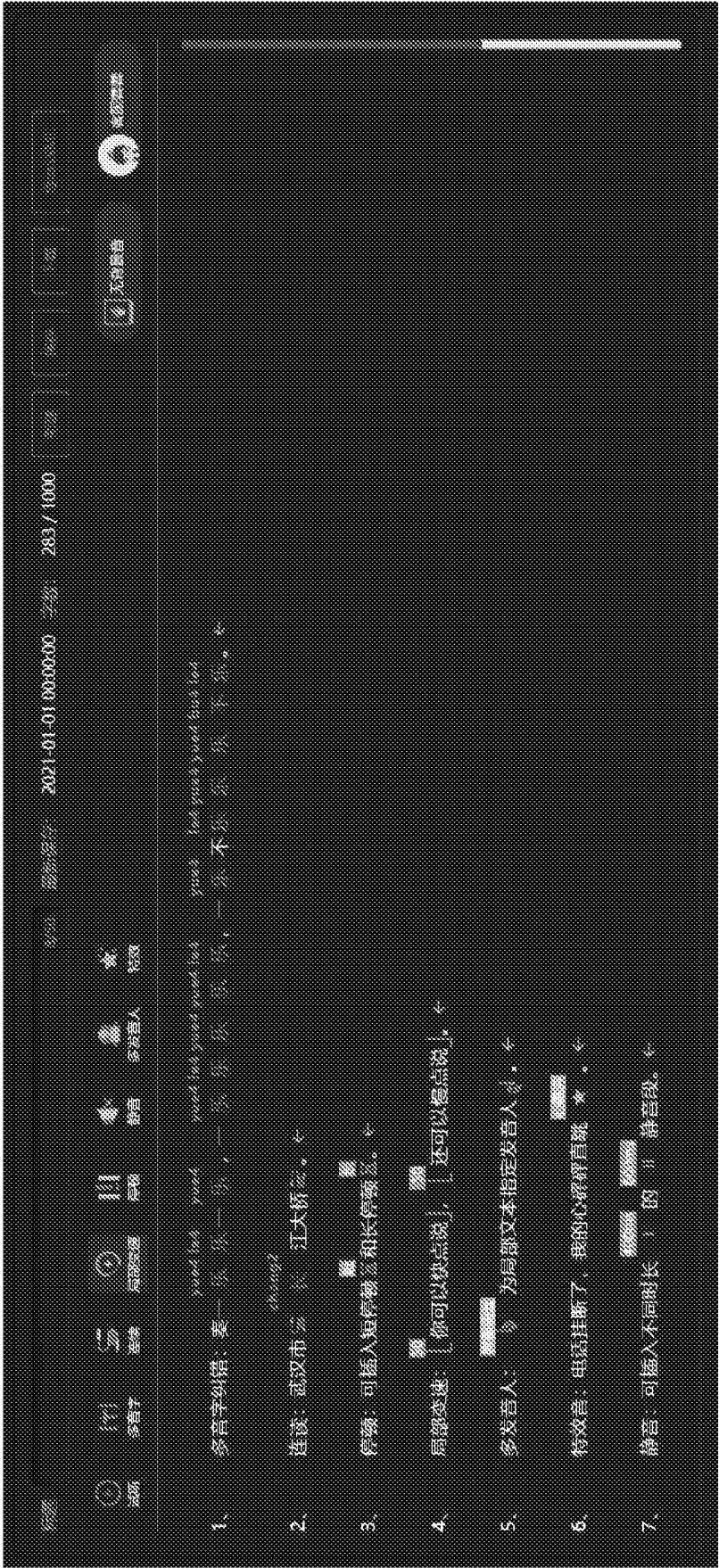


图 1

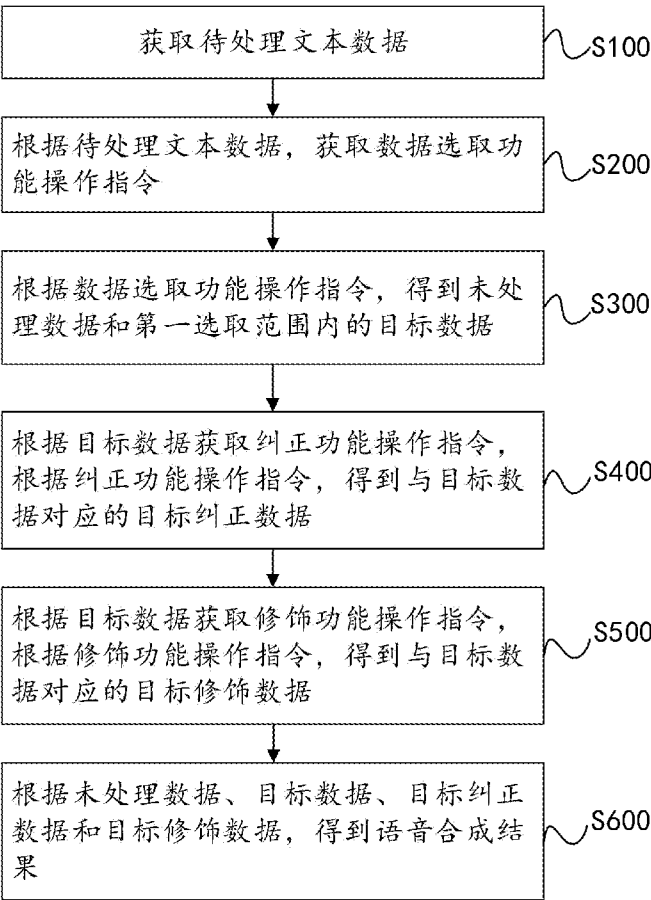


图 2

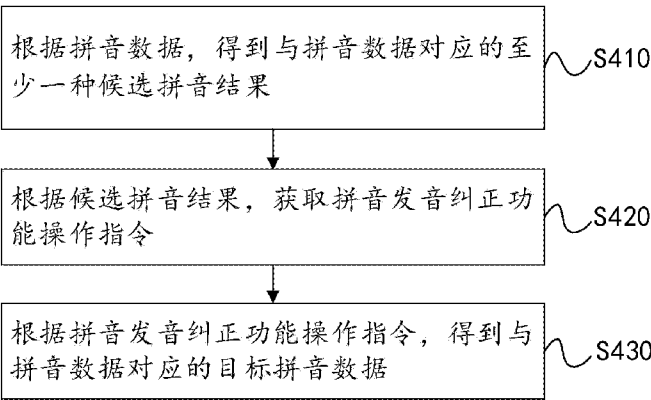


图 3



图 4

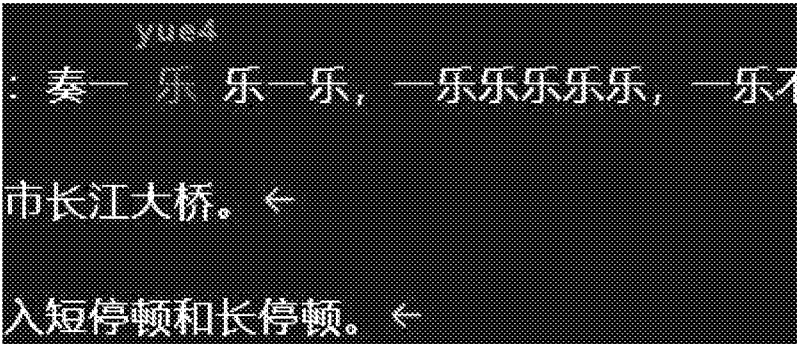


图 5

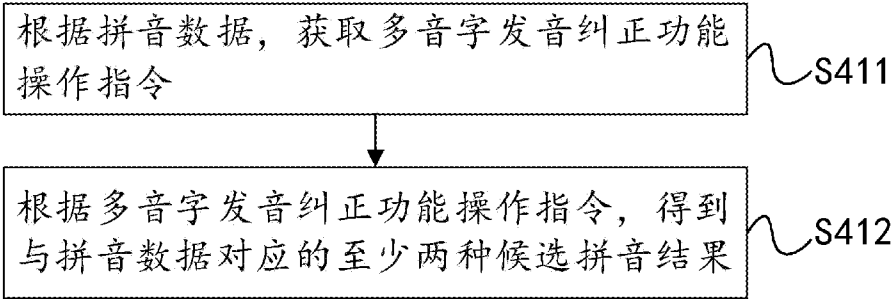


图 6

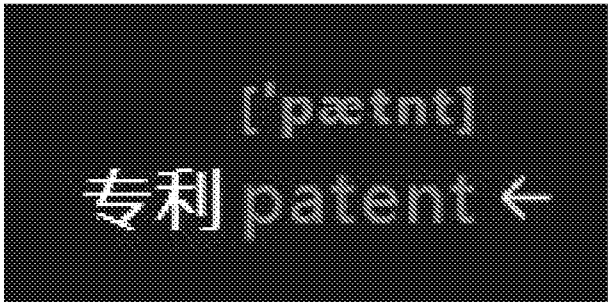


图 7



图 8



图 9

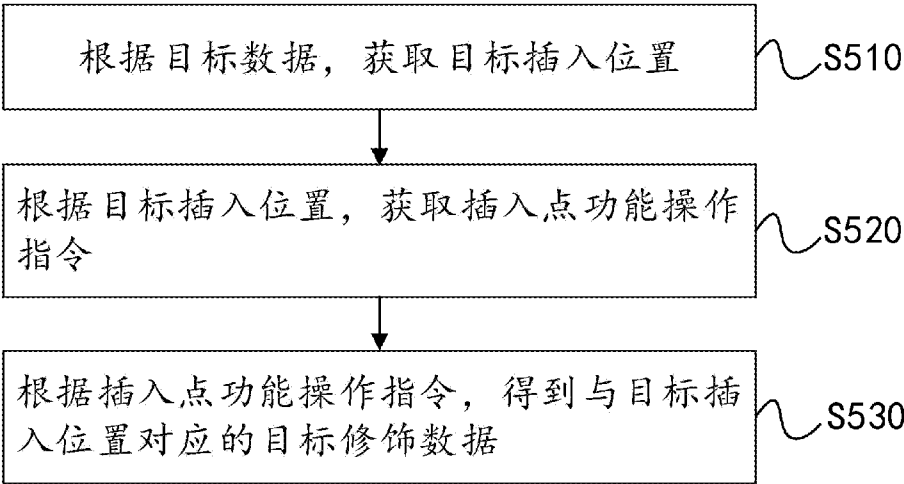


图 10



图 11



图 12



图 13

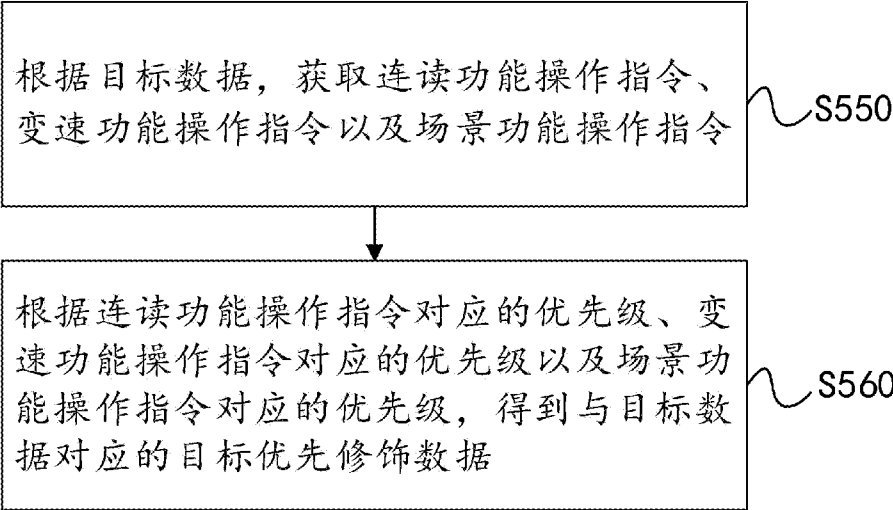


图 14

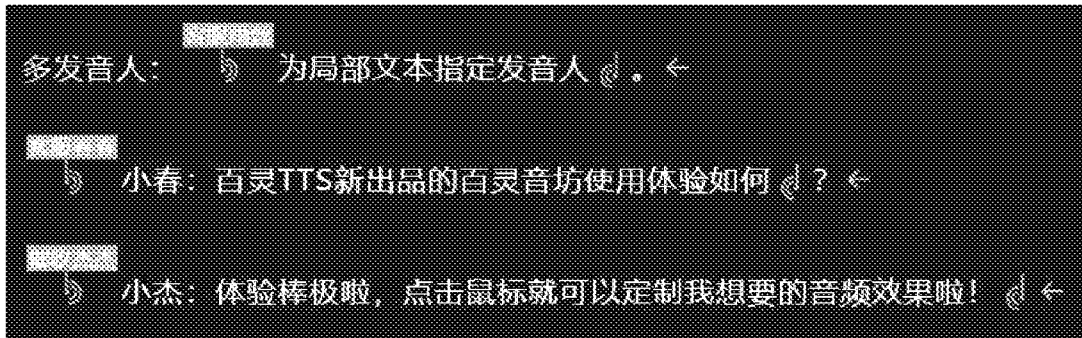


图 15



图 16



图 17

1234 1234 1234 → 1234 12 341234

1234 1234 1234 → 12341 2341234

1234 1234 1234 → 12 34123412 34

图 18



图 19

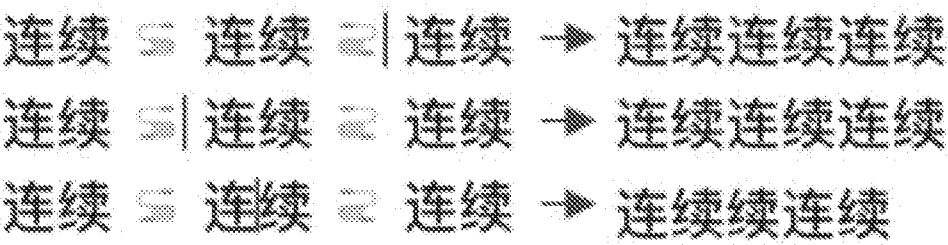


图 20

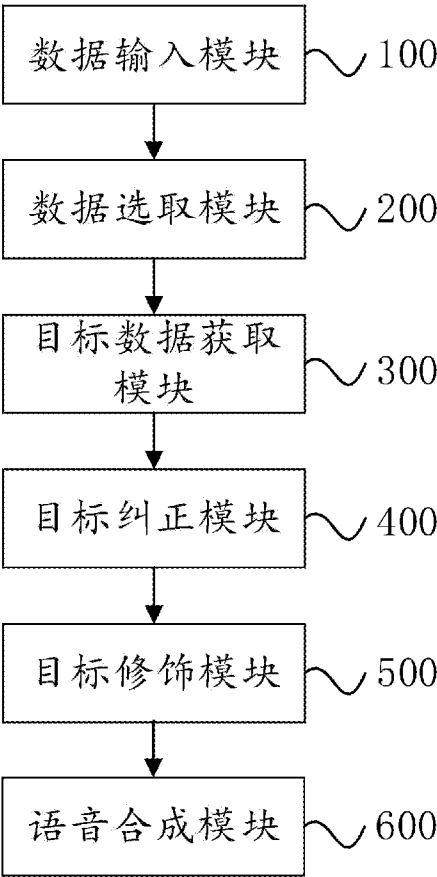


图 21

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2022/090728

A. CLASSIFICATION OF SUBJECT MATTER

G10L 13/10(2013.01)i; G10L 13/047(2013.01)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

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNABS; CNTXT; CNKI; VEN; WOTXT; EPTXT; USTXT: 语音合成, 文语转换, 文本, 编辑, 纠正, 纠错, 多音, 停顿, 静音, 特效音, 修饰, 连读, 变速, 语速, 场景, 背景, 撤销, TTS, speech, synthesis, process+, test, emotion, enrich+, pronunciation, select+, correct+, edit+, modif+, polyphonic, rhythm, pauses, target

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 113963681 A (PING AN TECHNOLOGY (SHENZHEN) CO., LTD.) 21 January 2022 (2022-01-21) description, paragraphs [0049]-[0153], and figures 1-21	1-20
X	CN 111199724 A (CHUMEN WENWEN INFORMATION TECHNOLOGY CO., LTD.) 26 May 2020 (2020-05-26) description, paragraphs [0025]-[0062], and figures 1-6	1-20
X	CN 111142667 A (AI SPEECH LTD.) 12 May 2020 (2020-05-12) description, paragraphs [0034]-[0103], and figures 1-5	1-20
X	CN 110767209 A (DATABAKER (BEIJING) TECHNOLOGY CO., LTD.) 07 February 2020 (2020-02-07) description, paragraphs [0062]-[0182], and figures 1-8	1-20
X	CN 112037756 A (BEIJING SOGOU TECHNOLOGY DEVELOPMENT CO., LTD.) 04 December 2020 (2020-12-04) description, paragraphs [0029]-[0198], and figures 1-5	1-3, 6-11, 14-17, 20
A	JP 2011170191 A (FUJITSU LTD.) 01 September 2011 (2011-09-01) entire document	1-20



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

27 June 2022

Date of mailing of the international search report

12 July 2022

Name and mailing address of the ISA/CN

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

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2022/090728

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	113963681	A	21 January 2022	None			
CN	111199724	A	26 May 2020	None			
CN	111142667	A	12 May 2020	None			
CN	110767209	A	07 February 2020	CN	110767209	B	15 March 2022
CN	112037756	A	04 December 2020	None			
JP	2011170191	A	01 September 2011	JP	5423466	B2	19 February 2014

国际检索报告

国际申请号

PCT/CN2022/090728

A. 主题的分类

G10L 13/10(2013.01)i; G10L 13/047(2013.01)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

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用))

CNABS; CNTXT; CNKI; VEN; WOTXT; EPTXT; USTXT; 语音合成, 文语转换, 文本, 编辑, 纠正, 纠错, 多音, 停顿, 静音, 特效音, 修饰, 连读, 变速, 语速, 场景, 背景, 撤销, TTS, speech, synthesis, process+, test, emotion, enrich+, pronunciation, select+, correct+, edit+, modif+, polyphonic, rhythm, pauses, target

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 113963681 A (平安科技深圳有限公司) 2022年1月21日 (2022 - 01 - 21) 说明书第[0049]-[0153]段, 图1-21	1-20
X	CN 111199724 A (出门问问信息科技有限公司) 2020年5月26日 (2020 - 05 - 26) 说明书第[0025]-[0062]段, 图1-6	1-20
X	CN 111142667 A (苏州思必驰信息科技有限公司) 2020年5月12日 (2020 - 05 - 12) 说明书第[0034]-[0103]段, 图1-5	1-20
X	CN 110767209 A (标贝北京科技有限公司) 2020年2月7日 (2020 - 02 - 07) 说明书第[0062]-[0182]段, 图1-8	1-20
X	CN 112037756 A (北京搜狗科技发展有限公司) 2020年12月4日 (2020 - 12 - 04) 说明书第[0029]-[0198]段, 图1-5	1-3、6-11、 14-17、20
A	JP 2011170191 A (FUJITSU LTD) 2011年9月1日 (2011 - 09 - 01) 全文	1-20

☐ 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

2022年6月27日

国际检索报告邮寄日期

2022年7月12日

ISA/CN的名称和邮寄地址

中国国家知识产权局(ISA/CN)

中国北京市海淀区蓟门桥西土城路6号 100088

传真号 (86-10)62019451

受权官员

田卓

电话号码 (86-512)88997344

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2022/090728

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	113963681	A	2022年1月21日	无	
CN	111199724	A	2020年5月26日	无	
CN	111142667	A	2020年5月12日	无	
CN	110767209	A	2020年2月7日	CN 110767209	B 2022年3月15日
CN	112037756	A	2020年12月4日	无	
JP	2011170191	A	2011年9月1日	JP 5423466	B2 2014年2月19日