

(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 11 月 21 日 (21.11.2019)

(10) 国际公布号
WO 2019/218481 A1

(51) 国际专利分类号:
G10L 13/08 (2013.01) *G10L 13/10* (2013.01)

(21) 国际申请号: PCT/CN2018/097560

(22) 国际申请日: 2018 年 7 月 27 日 (27.07.2018)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201810456213.3 2018年5月14日 (14.05.2018) 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: 朱坤(ZHU, Kun); 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: 深圳中一专利商标事务所(SHENZHEN ZHONGYI PATENT AND TRADEMARK OFFICE);
中国广东省深圳市福田区深南中路 1014 号老特区报社四楼 (5 号信箱), Guangdong 518028 (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, 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,

(54) Title: SPEECH SYNTHESIS METHOD, SYSTEM, AND TERMINAL APPARATUS

(54) 发明名称: 一种语音合成方法、系统及终端设备

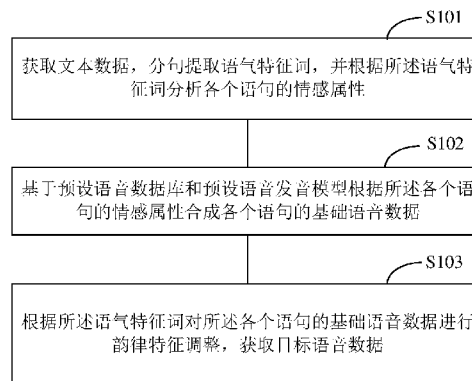


图 1

S101 ACQUIRE TEXT DATA, PERFORM SENTENCE SEGMENTATION ON THE SAME TO EXTRACT TONE FEATURE WORDS, AND PERFORM ANALYSIS ACCORDING TO THE TONE FEATURE WORDS TO OBTAIN AN EMOTIONAL ATTRIBUTE OF EACH SENTENCE

S102 SYNTHESIZE BASIC SPEECH DATA OF EACH SENTENCE ACCORDING TO THE EMOTIONAL ATTRIBUTE OF THE SENTENCE AND ON THE BASIS OF A PRE-DETERMINED SPEECH DATABASE AND A PRE-DETERMINED PHONETIC PRONUNCIATION MODEL

S103 ADJUST A PROSODIC FEATURE OF THE BASIC SPEECH DATA OF EACH SENTENCE ACCORDING TO THE TONE FEATURE WORDS, AND ACQUIRE TARGET SPEECH DATA

(57) Abstract: A speech synthesis method, a system, and a terminal apparatus. The method comprises: acquiring text data, performing sentence segmentation on the same to extract tone feature words, and performing analysis according to the tone feature words to obtain an emotional attribute of each sentence (S101); synthesizing basic speech data of each sentence according to the emotional attribute of a given sentence and on the basis of a pre-determined speech database and a pre-determined phonetic pronunciation model (S102); and adjusting a prosodic feature of the basic speech data of each sentence according to the tone feature words, and acquiring target speech data (S103). In the invention, tone feature words of each sentence in text data are extracted to obtain an emotional attribute of a given sentence via analysis, and obtained basic speech data is adjusted by means of a pre-determined phonetic pronunciation model incorporating the emotional attribute of a given sentence, such that a prosodic feature of the basic speech data is adjusted, thereby

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))。

obtaining more natural and human-like target speech data. The invention achieves richer emotional expression when producing words having a higher emotional intensity, thereby providing a speech effect closer to an actual human, and accordingly improving the quality of speech synthesis data.

(57) 摘要: 一种语音合成方法、系统及终端设备, 该方法包括: 获取文本数据, 分句提取语气特征词, 并根据语气特征词分析各个语句的情感属性 (S101); 基于预设语音数据库和预设语音发音模型根据各个语句的情感属性合成各个语句的基础语音数据 (S102); 根据语气特征词对各个语句的基础语音数据进行韵律特征调整, 获取目标语音数据 (S103)。通过提取文本数据中每句语句的语气特征词来分析语句的情感属性, 并经过预设语音发音模型结合语句的情感属性调整得到的基础语音数据, 在对基础语音数据进行韵律特征调整, 得到拟人度较高的目标语音数据。在进行情感倾向词发音时, 更富有感情, 更加符合实际用户发音的特征, 有效提高语音合成数据的质量。

一种语音合成方法、系统及终端设备

[0001] 本申请申明享有2018年05月14日递交的申请号为201810456213.3、名称为“一种任务分配方法、系统及终端设备”中国专利申请的优先权，该中国专利申请的整
体内容以参考的方式结合在本申请中。

技术领域

[0002] 本申请属于数据处理技术领域，尤其涉及一种语音合成方法、系统及终端设备。
。

背景技术

[0003] 有声书是一种个人或多人依据文稿并借助不同的声音表情和录音格式所录制的作品。目前市面上的有声书都是人工提前录制好并保存起来，在使用时直接播放。然而这需要耗费大量的人力资源进行提前录制。为了节省人力成本，能够通过语音合成技术合成语音数据。语音合成技术是指通过机械或电子等方法产生人造语音，将计算机自己产生的、或外部输入的文字信息的进行转变为可以听得懂的语音进行输出的技术。目前的语音合成技术在进行语音合成时，都是先对文本数据进行分析得到文本数据中的字和词，之后从语音库获取这些字和词对应的基本语音数据，最后将获取的基本语音数据按顺序进行组合得到最终的语音数据，然而得到的语音数据拟人度不高，因而存在质量低下的问题。

[0004] 综上所述，现有的语音合成技术存在合成得到的语音数据质量低下的问题。

发明概述

技术问题

[0005] 有鉴于此，本申请实施例提供了一种语音合成方法、系统及终端设备，以解决现有语音合成技术存在合成得到的语音数据质量低下的问题。

问题的解决方案

技术解决方案

[0006] 本申请的第一方面提供了一种语音合成方法，包括：

[0007] 获取文本数据，分句提取语气特征词，并根据所述语气特征词分析各个语句的

情感属性；

[0008] 基于预设语音数据库和预设语音发音模型根据所述各个语句的情感属性合成各个语句的基础语音数据；

[0009] 根据所述语气特征词对所述各个语句的基础语音数据进行韵律特征调整，获取目标语音数据。

[0010] 本申请的第二方面提供了一种语音合成系统，包括：

[0011] 情感分析模块，用于获取文本数据，分句提取语气特征词，并根据所述语气特征词分析各个语句的情感属性；

[0012] 语音合成模块，用于基于预设语音数据库和预设语音发音模型根据所述各个语句的情感属性合成各个语句的基础语音数据；

[0013] 语音调整模块，用于根据所述语气特征词对所述各个语句的基础语音数据进行韵律特征调整，获取目标语音数据。

[0014] 本申请的第三方面提供了一种终端设备，包括存储器、处理器以及存储在所述存储器中并可在所述处理器上运行的计算机可读指令，所述处理器执行所述计算机可读指令时实现以下步骤：

[0015] 获取文本数据，分句提取语气特征词，并根据所述语气特征词分析各个语句的情感属性；

[0016] 基于预设语音数据库和预设语音发音模型根据所述各个语句的情感属性合成各个语句的基础语音数据；

[0017] 根据所述语气特征词对所述各个语句的基础语音数据进行韵律特征调整，获取目标语音数据。

[0018] 本申请的第四方面提供了一种计算机可读存储介质，所述计算机可读存储介质存储有计算机可读指令，所述计算机可读指令被处理器执行时实现以下步骤：

[0019] 获取文本数据，分句提取语气特征词，并根据所述语气特征词分析各个语句的情感属性；

[0020] 基于预设语音数据库和预设语音发音模型根据所述各个语句的情感属性合成各个语句的基础语音数据；

[0021] 根据所述语气特征词对所述各个语句的基础语音数据进行韵律特征调整，获取

目标语音数据。

发明的有益效果

有益效果

[0022] 本申请提供一种语音合成方法、系统及终端设备，通过提取文本数据中每句语句的语气特征词来分析语句的情感属性，并经过预设语音发音模型结合语句的情感属性调整得到的基础语音数据，在对基础语音数据进行韵律特征调整，得到拟人度较高的目标语音数据。在进行情感倾向词发音时，更富有感情，更加符合实际用户发音的特征，有效提高语音合成数据的质量，解决了现有的语音合成技术存在合成得到的语音数据质量低下的问题。

对附图的简要说明

附图说明

- [0023] 图1是本申请实施例一提供的一种语音合成方法的实现流程示意图；
- [0024] 图2是本申请实施例二提供的对应实施例一步骤S101的实现流程示意图；
- [0025] 图3是本申请实施例三提供的对应实施例一步骤S102的实现流程示意图；
- [0026] 图4是本申请实施例四提供的对应实施例一步骤S103的实现流程示意图；
- [0027] 图5是本申请实施例五提供的一种语音合成系统的结构示意图；
- [0028] 图6是本申请实施例六提供的对应实施例五中情感分析模块101的结构示意图；
- [0029] 图7是本申请实施例七提供的对应实施例五中语音合成模块102的结构示意图；
- [0030] 图8是本申请实施例八提供的对应实施例五中语音调整模块103的结构示意图；
- [0031] 图9是本申请实施例九提供的终端设备的示意图。

发明实施例

本发明的实施方式

[0032] 以下描述中，为了说明而不是为了限定，提出了诸如特定系统结构、技术之类的具体细节，以便透彻理解本申请实施例。然而，本领域的技术人员应当清楚，在没有这些具体细节的其它实施例中也可以实现本申请。在其它情况中，省略对众所周知的系统、系统、电路以及方法的详细说明，以免不必要的细节妨碍本申请的描述。

[0033] 本申请实施例为了解决现有的语音合成技术存在合成得到的语音数据质量低下的问题，提供了一种语音合成方法、系统及终端设备，通过提取文本数据中每句语句的语气特征词来分析语句的情感属性，并经过预设语音发音模型结合语句的情感属性调整得到的基础语音数据，在对基础语音数据进行韵律特征调整，得到拟人度较高的目标语音数据。在进行情感倾向词发音时，更富有感情，更加符合实际用户发音的特征，有效提高语音合成数据的质量，解决了现有的语音合成技术存在合成得到的语音数据质量低下的问题。

[0034] 为了说明本申请所述的技术方案，下面通过具体实施例来进行说明。

[0035] 实施例一：

[0036] 如图1所示，本实施例提供了一种语音合成方法，其具体包括：

[0037] 步骤S101：获取文本数据，分句提取语气特征词，并根据所述语气特征词分析各个语句的情感属性。

[0038] 在具体应用中，通过终端获取带有文字信息的文本数据，该文本数据的格式可以为文本格式（txt）、富文本格式（Rich Text Format，RTF）或者文档（Document，DOC）等，也可以为便携式文档格式（Portable Document Format，PDF）或者图片等包含文字信息的文件，再将PDF或者图片转换成能够直接读取文本数据的文件即可，在此不加以限制。

[0039] 在具体应用中，获取到文本数据后，以句子为单位，提取每个句子中的语气特征词。语气特征词是指带有感情色彩的字词、符号或者字词组合，如“开心”、“天呐”、“太棒了”“好吧”、“？”等表示情绪和语气的字词及符号。由于语气特征词能够体现用户的情感倾向，在进行发音时会有不同的韵律特征。因此将各个语句中的语气特征词提取出来，并根据语气特征词分析各个语句的情感属性。

[0040] 在具体应用中，预先设置语气特征词数据库，根据语气特征词数据库提取各个语句中与语气特征词数据库相匹配的语气特征词。在表达情绪时，用户会采用字词的组合来表达自己的情绪。为了丰富语气特征词数据库、准确提取各个语句中的语气特征词，根据语法规则设置字词的组合规则，在提取语气特征词时将满足组合规则的字词一并进行提取。示例性的，上述组合规则包括但不限于以下组合规则：

- [0041] A: 程度副词+情感词, 如“较+好”、“很+好”、“特别+好”等;
- [0042] B: 否定词+情感词, 如“不+好”、“不+坏”等;
- [0043] C: 否定词+程度副词+情感词, 如“不+太+好”、“不+太+坏”等;
- [0044] D: 程度副词+否定词+情感词, 如“很+不+好”、“还+不+坏”等。
- [0045] 在具体应用中, 为了保证每一句话的语音合成效果, 本实施例是以句子为单位进行语气特征词提取, 针对各个语句的语气特征词分析该语句的情感属性。
- [0046] 在具体应用中, 先将文本数据的各个语句进行拆分, 拆分为多个字词的组合, 将拆分的字词组合区分为中立词及语气特征词, 其中语气特征词包括正面词和负面词, 可以根据上述中立词、正面词以及负面词在该语句中所占的比例分析得到该语句的情感属性。
- [0047] 步骤S102: 基于预设语音数据库和预设语音发音模型根据所述各个语句的情感属性合成各个语句的基础语音数据。
- [0048] 在具体应用中, 将已经拆分为多个字词的组合的语句以字词为单位在预设语音数据库中获取每个字词的语音数据, 将多个字词的语音数据进行合成得到整句的语音数据。
- [0049] 在具体应用中, 得到整句的语音数据后, 根据该语句的情感属性基于预设语音发音模型对语音数据进行声学特征调整, 以得到对应语句情感属性的基础语音数据, 使得发音与实际用户的发音更加接近。在具体应用中, 上述声学特征包括声音强度、语速、音调的高低等特征。
- [0050] 在具体应用中, 上述预设语音发音模型的构建过程为: 采集大量实际用户的语音数据作为训练样本, 并对语音数据中以每一个句子为单位进行情感属性标记, 利用神经网络进行训练, 得到每种情感属性对应的语音发音的声学特征。
- [0051] 步骤S103: 根据所述语气特征词对所述各个语句的基础语音数据进行韵律特征调整, 获取目标语音数据。
- [0052] 在具体应用中, 基础语音数据是基于整句的情感属性, 为了进一步得到更符合实际用户在对应情感时的发音特点, 针对语气特征词再对整句的基础语音数据进行韵律特征调整。
- [0053] 在具体应用中, 上述韵律特征包括音强、音高及音长。音强则是语音的重音、

轻音等强弱变化；音高则是语音的字调和语调；音长则是语音的节奏快慢。

[0054] 在具体应用中，由于不同情感倾向的语气特征词表达的用户情绪是不一样的，而不同的情绪对应的发言的韵律特征会存在较大差异，如开心时的音调会明显比悲伤时的音调高。即每个语气特征词对应一种（或一类）韵律特征，因此，先获取到语气特征词对应的韵律特征，根据该语气特征词的韵律特征对基础语音数据中的语气特征词进行韵律特征调整，若一个句子中有多个语气特征词，则对全部语气特征词都进行韵律特征调整，得到更加符合实际用户发音的语音数据。

[0055] 在具体应用中，对韵律特征进行调整可以是预先设置各类语气特征词的韵律特征参数，如设置表示开心的语气特征词的韵律特征参数为音强1、音高1以及音长1，设置表示悲伤的语气特征词的韵律特征参数为音强2、音高2以及音长2。或者是基于基础语音数据的韵律特征参数的百分比进行调整，如表示开心的语气特征词在进行韵律特征调整时，是将该语气特征词对应的音高在基础语音数据的基础上升高10%，将音长缩短15%。

[0056] 本实施例提供的语音合成方法，通过提取文本数据中每句语句的语气特征词来分析语句的情感属性，并经过预设语音发音模型结合语句的情感属性调整得到拟人化的基础语音数据，再对基础语音数据进行韵律特征调整，得到更加接近实际用户发音的目标语音数据。在进行情感倾向词发音时，更富有感情，更加符合实际用户发音的特征，有效提高语音合成数据的质量，解决了现有的语音合成技术存在合成得到的语音数据质量低下的问题。

[0057] 实施例二：

[0058] 如图2所示，在本实施例中，实施例一中的步骤S101具体包括：

[0059] 步骤S201：根据所述语气特征词获取所述语句的多个预设维度的情感分析参数。

[0060] 在具体应用中，将语句进行拆分后对每个词汇从[正面、中立、负面]等三个预设维度进行评分，得到该语句对应[正面、中立、负面]等三个预设维度的评分综合，再分别计算该语句的三个预设维度所占分数的比值。示例性的，语句被拆分后可以分为中立词和语气特征词，语气特征词可分为正面词及负面词。在对

语气特征词进行分类时，对语气特征词进行分级并设置该语气特征词对应级别的等级评分。

- [0061] 如“开心”：设置其为正面词且其等级评分为+2；
- [0062] “太棒了”：设置其为正面词且等级评分为+5；
- [0063] “不好”：设置其为负面词且等级评分为-2”；
- [0064] “很不好”：设置其为负面词且等级评分为-5。需要说明的是，上述对语气特征词进行分类、分级及评分可以基于神经网络构建分类评分模块进行实现，具体实现手段不加以赘述。
- [0065] 步骤S202：对所述多个预设维度的情感分析参数所占总维度的情感分析参数的比例确定所述语句的情感属性。
- [0066] 在具体应用中，根据语气特征词的等级评分计算预设维度的分数，再计算各个预设维度所占总维度分数的比例，如一个句子的三个预设维度的评分为[+10, 4, -6]，此时得到三个维度分数所占分数比例即为[+0.5, 0.2, -0.3]。
- [0067] 在具体应用中，利用支持向量机制的文本情感分类模型对各个语句对应的预设三个维度的分数所占的比例值进行计算，进而判断出该语句对应的情感属性。预先利用大量不同文本数据对文本情感分类模型进行训练，以得到满足向量机制的文本情感分析模型。其中，可以将情感属性划分为多种不同的状态，以此来量化用户的情感，得到的每个句子对应的量化来表示该语句的情感属性。示例性的，上述量化的情感属性包括但不限于：开心、悲伤、愤怒、恐惧、疑惑及正常。
- [0068] 实施例三：
- [0069] 如图3所示，在本实施例中，实施例一中的步骤S102具体包括：
- [0070] 步骤S301：从所述预设语音数据库中获取与语句的各个词对应的语音数据。
- [0071] 在具体应用中，将各个语句进行拆分，拆分为多个字词的组合，以字词为单位在预设语音数据库中获取每个字词的语音数据
- [0072] 步骤S302：将所述语音数据进行合成获得到所述语句的电子语音数据。
- [0073] 在具体应用中，将多个字词的语音数据进行合成得到整句的语音数据即得到该语句的电子语音数据。

[0074] 步骤S303: 通过语音发音模型根据所述语句的情感属性对所述电子语音数据的音高、音强及语速进行调整, 得到所述语句的基础语音数据。

[0075] 在具体应用中, 得到电子语音数据后, 根据该语句的情感属性基于预设语音发音模型对电子语音数据的音高、音强及语速进行调整, 以得到对应语句情感属性的基础语音数据, 使得发音与实际用户的发音更加接近。

[0076] 实施例四:

[0077] 如图4所示, 在本实施例中, 实施例一中的步骤S103具体包括:

[0078] 步骤S401: 获取各个语句中语气特征词的韵律特征调整规则, 包括: 音高、音强及音长等韵律特征参数的调整规则。

[0079] 在具体应用中, 对语气特征词进行分类、并设定不同类别的语气特征词的等级, 根据该语气特征词的分类及等级获取对应的韵律特征调整规则, 具体为该语气特征词的音高、音强及音长等韵律特征参数的调整规则。

[0080] 在具体应用中, 预先设定不同类别不同等级的语气特征词的韵律特征参数, 再对各个语气特征词进行类别和等级划分, 进而得到其对应的韵律特征参数获取其对应的韵律特征调整规则。

[0081] 步骤S402: 根据语气特征词的韵律特征调整规则对所述基础语音数据的音高、音强及音长进行调整, 得到目标语音数据。

[0082] 在具体应用中, 在获取到语气特征词对应的韵律特征调整规则后, 根据该调整规则对基础语音数据中各个语气特征词的音高、音强及音长进行调整, 经过调整后就能够得到更接近实际用户发音的目标语音数据。

[0083] 在具体应用中, 上述调整过程可以是根据韵律特征调整规则计算得到语气特征词的韵律特征参数, 然后将基础语音数据中对应的语气特征词的韵律特征调整至该韵律特征参数。也可以是根据韵律特征调整规则采用百分比的形式对基础语音数据进行调整, 在此不加以限制。

[0084] 在一个实施例中, 在上述步骤S402之后, 还包括以下步骤:

[0085] 获取所述目标语音数据的韵律特征参数;

[0086] 根据目标语音数据的韵律特征参数计算各个语句的韵律特征参数的平均值;

[0087] 根据所述平均值对语句的各个词的音高、音强及音长进行调整, 得到平滑过渡

的语音数据。

[0088] 在具体应用中，由于上述韵律特征调整仅是针对语气特征词，因此可能会导致出现语音突变的问题，使得语气特征词的发音与语气特征词前后相连字词的发音出现突兀不和谐的情况。在具体应用中，为了避免上述问题，可以通过对经过韵律特征调整后的目标语音数据的韵律特征参数以整句为单位进行再一次的调整，使得语句能够平滑过渡。具体为：根据目标语音数据的韵律特征参数得到各个语句的韵律特征参数的平均值。对于语气特征词相连的字词，采用上述平均值对该字词的音高、音强及音长进行调整。在具体应用中，当多个语气特征词相连时，只需对于第一个语气特征词和最后一个语气特征词相连的字词的音高、音强及音长进行调整。

[0089] 示例性的，如“我们下午去游乐场玩好不好！”中，“好不好”作为语气特征词，其音高及音调都会升高，而与其相连的“玩”的音高及音调则不会升高，因此可能出现从“玩”到“好不好”时音调和音高突然提升的情况，使得发音过渡不自然。因此根据整句的韵律特征参数计算得到整句的韵律特征参数平均值，将“玩”的韵律特征参数调整为该平均值，则能够有效地减少“玩”和“好不好”的音强和音调之间的落差，实现平滑过渡。

[0090] 实施例五：

[0091] 如图5所示，本实施例提供一种语音合成系统100，用于执行实施例一中的方法步骤，其包括情感分析模块101、语音合成模块102以及语音调整模块103。

[0092] 情感分析模块101用于获取文本数据，分句提取语气特征词，并根据所述语气特征词分析各个语句的情感属性；

[0093] 语音合成模块102用于基于预设语音数据库和预设语音发音模型根据所述各个语句的情感属性合成各个语句的基础语音数据；

[0094] 语音调整模块103用于根据所述语气特征词对所述各个语句的基础语音数据进行韵律特征调整，获取目标语音数据。

[0095] 需要说明的是，本申请实施例提供的语音合成系统，由于与本申请图1所示方法实施例基于同一构思，其带来的技术效果与本申请图1所示方法实施例相同，具体内容可参见本申请图1所示方法实施例中的叙述，此处不再赘述。

[0096] 因此，本实施例提供的一种语音合成系统，同样能够通过提取文本数据中每句语句的语气特征词来分析语句的情感属性，并经过预设语音发音模型结合语句的情感属性调整得到拟人化的基础语音数据，再对基础语音数据进行韵律特征调整，得到更加接近实际用户发音的目标语音数据。在进行情感倾向词发音时，更富有感情，更加符合实际用户发音的特征，有效提高语音合成数据的质量，解决了现有的语音合成技术存在合成得到的语音数据质量低下的问题。

[0097] 实施例六：

[0098] 如图6所示，在本实施例中，实施例五中的情感分析模块101包括用于执行图2所对应的实施例中的方法步骤的结构，其包括参数获取单元201及情感分析单元202。

[0099] 参数获取单元201用于根据所述语气特征词获取所述语句的多个预设维度的情感分析参数。

[0100] 情感分析单元202用于对所述多个预设维度的情感分析参数所占总维度的情感分析参数的比例确定所述语句的情感属性。

[0101] 实施例七：

[0102] 如图7所示，在本实施例中，实施例五中的语音合成模块102包括用于执行图3所对应的实施例中的方法步骤的结构，其包括语音数据获取单元301、语音数据合成单元302以及声学特征调整单元303。

[0103] 语音数据获取单元301用于从所述预设语音数据库中获取与语句的各个词对应的语音数据。

[0104] 语音数据合成单元302用于将所述语音数据进行合成获得到所述语句的电子语音数据。

[0105] 声学特征调整单元303用于通过语音发音模型根据所述语句的情感属性对所述电子语音数据的音高、音强及语速进行调整，得到所述语句的基础语音数据。

[0106] 实施例八：

[0107] 如图8所示，在本实施例中，实施例五中的语音调整模块103包括用于执行图4所对应的实施例中的方法步骤的结构，其包括韵律特征调整规则获取单元401及韵律特征调整单元402。

- [0108] 韵律特征调整规则获取单元401用于获取各个语句中语气特征词的韵律特征调整规则，包括：音高、音强及音长等韵律特征参数的调整规则。
- [0109] 韵律特征调整单元402用于根据语气特征词的韵律特征调整规则对所述基础语音数据的音高、音强及音长进行调整，得到目标语音数据。
- [0110] 在一个实施例中，上述语音调整模块103还包括特征参数获取单元、计算单元以及平滑过渡调整单元。
- [0111] 特征参数获取单元用于获取所述目标语音数据的韵律特征参数。
- [0112] 计算单元用于根据目标语音数据的韵律特征参数计算各个语句的韵律特征参数的平均值。
- [0113] 平滑过渡调整单元，用于根据所述平均值对语句的各个词的音高、音强及音长进行调整，得到平滑过渡的语音数据。
- [0114] 实施例九：
- [0115] 图9是本申请实施例九提供的终端设备的示意图。如图9所示，该实施例的终端设备9包括：处理器90、存储器91以及存储在所述存储器91中并可在所述处理器90上运行的计算机可读指令92，例如程序。所述处理器90执行所述计算机可读指令92时实现上述各个语音合成方法实施例中的步骤，例如图1所示的步骤S101至S103。或者，所述处理器90执行所述计算机可读指令92时实现上述系统实施例中各模块/单元的功能，例如图5所示模块101至103的功能。
- [0116] 示例性的，所述计算机可读指令92可以被分割成一个或多个模块/单元，所述一个或者多个模块/单元被存储在所述存储器91中，并由所述处理器90执行，以完成本申请。所述一个或多个模块/单元可以是能够完成特定功能的一系列计算机可读指令指令段，该指令段用于描述所述计算机可读指令92在所述终端设备9中的执行过程。例如，所述计算机可读指令92可以被分割成情感分析模块、语音合成模块以及语音调整模块，各模块具体功能如下：
- [0117] 情感分析模块，用于获取文本数据，分句提取语气特征词，并根据所述语气特征词分析各个语句的情感属性；
- [0118] 语音合成模块，用于基于预设语音数据库和预设语音发音模型根据所述各个语句的情感属性合成各个语句的基础语音数据；

- [0119] 语音调整模块，用于根据所述语气特征词对所述各个语句的基础语音数据进行韵律特征调整，获取目标语音数据。
- [0120] 所述终端设备9可以是桌上型计算机、笔记本、掌上电脑及云端管理服务器等计算设备。所述终端设备可包括，但不仅限于，处理器90、存储器91。本领域技术人员可以理解，图9仅仅是终端设备9的示例，并不构成对终端设备9的限定，可以包括比图示更多或更少的部件，或者组合某些部件，或者不同的部件，例如所述终端设备还可以包括输入输出设备、网络接入设备、总线等。
- [0121] 所称处理器90可以是中央处理单元(Central Processing Unit, CPU)，还可以是其他通用处理器、数字信号处理器 (Digital Signal Processor, DSP)、专用集成电路 (Application Specific Integrated Circuit, ASIC)、现成可编程门阵列 (Field-Programmable Gate Array, FPGA) 或者其他可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件等。通用处理器可以是微处理器或者该处理器也可以是任何常规的处理器等。
- [0122] 所述存储器91可以是所述终端设备9的内部存储单元，例如终端设备9的硬盘或内存。所述存储器91也可以是所述终端设备9的外部存储设备，例如所述终端设备9上配备的插接式硬盘，智能存储卡 (Smart Media Card, SMC)，安全数字 (Secure Digital, SD) 卡，闪存卡 (Flash Card) 等。进一步地，所述存储器91还可以既包括所述终端设备9的内部存储单元也包括外部存储设备。所述存储器91用于存储所述计算机可读指令以及所述终端设备所需的其他程序和数据。所述存储器91还可以用于暂时地存储已经输出或者将要输出的数据。
- [0123] 所属领域的技术人员可以清楚地了解到，为了描述的方便和简洁，仅以上述各功能单元、模块的划分进行举例说明，实际应用中，可以根据需要而将上述功能分配由不同的功能单元、模块完成，即将所述系统的内部结构划分成不同的功能单元或模块，以完成以上描述的全部或者部分功能。实施例中的各功能单元、模块可以集成在一个处理单元中，也可以是各个单元单独物理存在，也可以两个或两个以上单元集成在一个单元中，上述集成的单元既可以采用硬件的形式实现，也可以采用软件功能单元的形式实现。另外，各功能单元、模块的

具体名称也只是为了便于相互区分，并不用于限制本申请的保护范围。上述无线终端中单元、模块的具体工作过程，可以参考前述方法实施例中的对应过程，在此不再赘述。

[0124] 在上述实施例中，对各个实施例的描述都各有侧重，某个实施例中未详述或记载的部分，可以参见其它实施例的相关描述。

[0125] 本领域普通技术人员可以意识到，结合本文中公开的实施例描述的各示例的单元及算法步骤，能够以电子硬件、或者计算机软件和电子硬件的结合来实现。这些功能究竟以硬件还是软件方式来执行，取决于技术方案的特定应用和设计约束条件。专业技术人员可以对每个特定的应用来使用不同方法来实现所描述的功能，但是这种实现不应认为超出本申请的范围。

[0126] 在本申请所提供的实施例中，应该理解到，所揭露的系统/终端设备和方法，可以通过其它的方式实现。例如，以上所描述的系统/终端设备实施例仅仅是示意性的，例如，所述模块或单元的划分，仅仅为一种逻辑功能划分，实际实现时可以有另外的划分方式，例如多个单元或组件可以结合或者可以集成到另一个系统，或一些特征可以忽略，或不执行。另一点，所显示或讨论的相互之间的耦合或直接耦合或通讯连接可以是通过一些接口，系统或单元的间接耦合或通讯连接，可以是电性，机械或其它的形式。

[0127] 所述设置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，设置为单元显示的部件可以是或者也可以不是物理单元，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单元来实现本实施例方案的目的。

[0128] 另外，在本申请各个实施例中的各功能单元可以集成在一个处理单元中，也可以是各个单元单独物理存在，也可以两个或两个以上单元集成在一个单元中。上述集成的单元既可以采用硬件的形式实现，也可以采用软件功能单元的形式实现。

[0129] 所述集成的模块/单元如果以软件功能单元的形式实现并设置为独立的产品销售或使用，可以存储在一个计算机可读取存储介质中。基于这样的理解，本申请实现上述实施例方法中的全部或部分流程，也可以通过计算机可读指令来

指令相关的硬件来完成，所述的计算机可读指令可存储于一计算机可读存储介质中，该计算机可读指令在被处理器执行时，可实现上述各个方法实施例的步骤。其中，所述计算机可读指令包括计算机可读指令代码，所述计算机可读指令代码可以为源代码形式、对象代码形式、可执行文件或某些中间形式等。所述计算机可读介质可以包括：能够携带所述计算机可读指令代码的任何实体或系统、记录介质、U盘、移动硬盘、磁碟、光盘、计算机存储器、只读存储器（ROM，Read-Only Memory）、随机存取存储器（RAM，Random Access Memory）、电载波信号、电信信号以及软件分发介质等。需要说明的是，所述计算机可读介质包含的内容可以根据司法管辖区内立法和专利实践的要求进行适当的增减，例如在某些司法管辖区，根据立法和专利实践，计算机可读介质不包括是电载波信号和电信信号。

[0130] 以上所述实施例仅用以说明本申请的技术方案，而非对其限制；尽管参照前述实施例对本申请进行了详细的说明，本领域的普通技术人员应当理解：其依然可以对前述各实施例所记载的技术方案进行修改，或者对其中部分技术特征进行等同替换；而这些修改或者替换，并不使相应技术方案的本质脱离本申请各实施例技术方案的精神和范围，均应包含在本申请的保护范围之内。

权利要求书

- [权利要求 1] 一种语音合成方法，其特征在于，包括：
- 获取文本数据，分句提取语气特征词，并根据所述语气特征词分析各个语句的情感属性；
- 基于预设语音数据库和预设语音发音模型根据所述各个语句的情感属性合成各个语句的基础语音数据；
- 根据所述语气特征词对所述各个语句的基础语音数据进行韵律特征调整，获取目标语音数据。
- [权利要求 2] 根据权利要求1所述的语音合成方法，其特征在于，所述根据所述语气特征词分析各个语句的情感属性，包括：
- 根据所述语气特征词获取所述语句的多个预设维度的情感分析参数；
- 对所述多个预设维度的情感分析参数所占总维度的情感分析参数的比例确定所述语句的情感属性。
- [权利要求 3] 根据权利要求1所述的语音合成方法，其特征在于，基于预设语音数据库和预设语音发音模型根据所述各个语句的情感属性合成各个语句的基础语音数据，包括：
- 从所述预设语音数据库中获取与语句的各个词对应的语音数据；
- 将所述语音数据进行合成获得所述语句的电子语音数据；
- 通过语音发音模型根据所述语句的情感属性对所述电子语音数据的音高、音强及语速进行调整，得到所述语句的基础语音数据。
- [权利要求 4] 根据权利要求1所述的语音合成方法，其特征在于，根据所述语气特征词对所述各个语句的基础语音数据进行韵律特征调整，获取目标语音数据，包括：
- 获取各个语句中语气特征词的韵律特征调整规则，包括：音高、音强及音长等韵律特征参数的调整规则；
- 根据语气特征词的韵律特征调整规则对所述基础语音数据的音高、音强及音长进行调整，得到目标语音数据。
- [权利要求 5] 根据权利要求4所述的语音合成方法，其特征在于，根据语气特征词

的韵律特征调整规则对所述基础语音数据的音高、音强及音长进行调整，得到目标语音数据之后，还包括：

获取所述目标语音数据的韵律特征参数；

根据目标语音数据的韵律特征参数计算各个语句的韵律特征参数的平均值；

根据所述平均值对语句的各个词的音高、音强及音长进行调整，得到平滑过渡的语音数据。

[权利要求 6]

一种语音合成系统，其特征在于，包括：

情感分析模块，用于获取文本数据，分句提取语气特征词，并根据所述语气特征词分析各个语句的情感属性；

语音合成模块，用于基于预设语音数据库和预设语音发音模型根据所述各个语句的情感属性合成各个语句的基础语音数据；

语音调整模块，用于根据所述语气特征词对所述各个语句的基础语音数据进行韵律特征调整，获取目标语音数据。

[权利要求 7]

根据权利要求6所述的语音合成系统，其特征在于，所述情感分析模块包括：

参数获取单元，用于根据所述语气特征词获取所述语句的多个预设维度的情感分析参数；

情感分析单元，用于对所述多个预设维度的情感分析参数所占总维度的情感分析参数的比例确定所述语句的情感属性。

[权利要求 8]

根据权利要求6所述的语音合成系统，其特征在于，所述语音合成模块包括：

语音数据获取单元，用于从所述预设语音数据库中获取与语句的各个词对应的语音数据；

语音数据合成单元，用于将所述语音数据进行合成获得所述语句的电子语音数据；

声学特征调整单元，用于通过语音发音模型根据所述语句的情感属性对所述电子语音数据的音高、音强及语速进行调整，得到所述语句的

基础语音数据。

- [权利要求 9] 根据权利要求1所述的语音合成系统，其特征在于，所述语音调整模块包括：
- 韵律特征调整规则获取单元，用于获取各个语句中语气特征词的韵律特征调整规则，包括：音高、音强及音长等韵律特征参数的调整规则；
- 韵律特征调整单元，用于根据语气特征词的韵律特征调整规则对所述基础语音数据的音高、音强及音长进行调整，得到目标语音数据。

- [权利要求 10] 根据权利要求9所述的语音合成系统，其特征在于，所述语音调整模块还包括：
- 特征参数获取单元，用于获取所述目标语音数据的韵律特征参数；
- 计算单元，用于根据目标语音数据的韵律特征参数计算各个语句的韵律特征参数的平均值；
- 平滑过渡调整单元，用于根据所述平均值对语句的各个词的音高、音强及音长进行调整，得到平滑过渡的语音数据。

- [权利要求 11] 一种终端设备，包括存储器、处理器以及存储在所述存储器中并可在所述处理器上运行的计算机可读指令，其特征在于，所述处理器执行所述计算机可读指令时实现如下步骤：
- 获取文本数据，分句提取语气特征词，并根据所述语气特征词分析各个语句的情感属性；
- 基于预设语音数据库和预设语音发音模型根据所述各个语句的情感属性合成各个语句的基础语音数据；
- 根据所述语气特征词对所述各个语句的基础语音数据进行韵律特征调整，获取目标语音数据。

- [权利要求 12] 根据权利要求11所述的终端设备，其特征在于，所述根据所述语气特征词分析各个语句的情感属性，包括：
- 根据所述语气特征词获取所述语句的多个预设维度的情感分析参数；
- 对所述多个预设维度的情感分析参数所占总维度的情感分析参数的比

例确定所述语句的情感属性。

[权利要求 13] 根据权利要求12所述的终端设备，其特征在于，基于预设语音数据库和预设语音发音模型根据所述各个语句的情感属性合成各个语句的基础语音数据，包括：

从所述预设语音数据库中获取与语句的各个词对应的语音数据；

将所述语音数据进行合成获得所述语句的电子语音数据；

通过语音发音模型根据所述语句的情感属性对所述电子语音数据的音高、音强及语速进行调整，得到所述语句的基础语音数据。

[权利要求 14] 根据权利要求11所述的终端设备，其特征在于，4. 根据权利要求1所述的语音合成方法，其特征在于，根据所述语气特征词对所述各个语句的基础语音数据进行韵律特征调整，获取目标语音数据，包括：

获取各个语句中语气特征词的韵律特征调整规则，包括：音高、音强及音长等韵律特征参数的调整规则；

根据语气特征词的韵律特征调整规则对所述基础语音数据的音高、音强及音长进行调整，得到目标语音数据。

[权利要求 15] 根据权利要求14所述的终端设备，其特征在于，根据语气特征词的韵律特征调整规则对所述基础语音数据的音高、音强及音长进行调整，得到目标语音数据之后，所述处理器执行所述计算机可读指令时还实现如下步骤：

获取所述目标语音数据的韵律特征参数；

根据目标语音数据的韵律特征参数计算各个语句的韵律特征参数的平均值；

根据所述平均值对语句的各个词的音高、音强及音长进行调整，得到平滑过渡的语音数据。

[权利要求 16] 一种计算机可读存储介质，所述计算机可读存储介质存储有计算机可读指令，其特征在于，所述计算机可读指令被处理器执行时实现如下步骤：

获取文本数据，分句提取语气特征词，并根据所述语气特征词分析各

个语句的情感属性；

基于预设语音数据库和预设语音发音模型根据所述各个语句的情感属性合成各个语句的基础语音数据；

根据所述语气特征词对所述各个语句的基础语音数据进行韵律特征调整，获取目标语音数据。

[权利要求 17] 根据权利要求16所述的计算机可读存储介质，其特征在于，所述根据所述语气特征词分析各个语句的情感属性，包括：
根据所述语气特征词获取所述语句的多个预设维度的情感分析参数；
对所述多个预设维度的情感分析参数所占总维度的情感分析参数的比例确定所述语句的情感属性。

[权利要求 18] 根据权利要求17所述的计算机可读存储介质，其特征在于，基于预设语音数据库和预设语音发音模型根据所述各个语句的情感属性合成各个语句的基础语音数据，包括：
从所述预设语音数据库中获取与语句的各个词对应的语音数据；
将所述语音数据进行合成获得到所述语句的电子语音数据；
通过语音发音模型根据所述语句的情感属性对所述电子语音数据的音高、音强及语速进行调整，得到所述语句的基础语音数据。

[权利要求 19] 根据权利要求16所述的计算机可读存储介质，其特征在于，基于预设语音数据库和预设语音发音模型根据所述各个语句的情感属性合成各个语句的基础语音数据，包括：
从所述预设语音数据库中获取与语句的各个词对应的语音数据；
将所述语音数据进行合成获得到所述语句的电子语音数据；
通过语音发音模型根据所述语句的情感属性对所述电子语音数据的音高、音强及语速进行调整，得到所述语句的基础语音数据。

[权利要求 20] 根据权利要求16所述的计算机可读存储介质，其特征在于，根据语气特征词的韵律特征调整规则对所述基础语音数据的音高、音强及音长进行调整，得到目标语音数据之后，所述计算机可读指令被处理器执行时还实现如下步骤：

获取所述目标语音数据的韵律特征参数；
根据目标语音数据的韵律特征参数计算各个语句的韵律特征参数的平均值；
根据所述平均值对语句的各个词的音高、音强及音长进行调整，得到平滑过渡的语音数据。

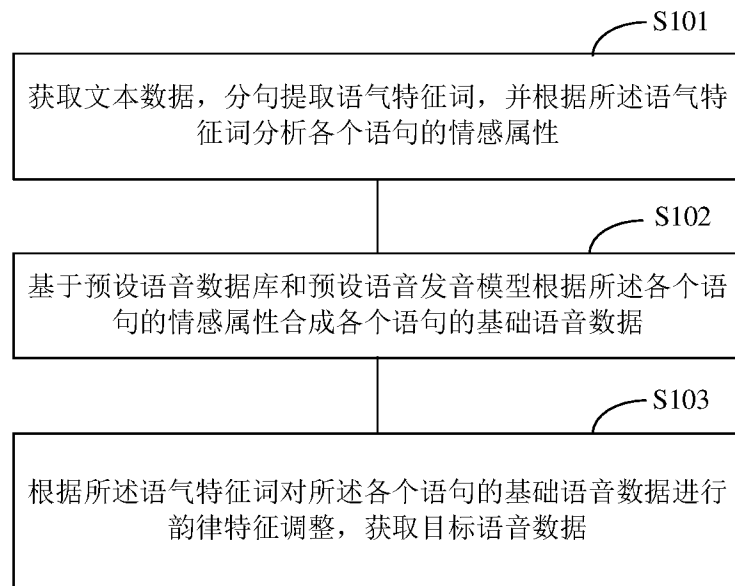


图 1

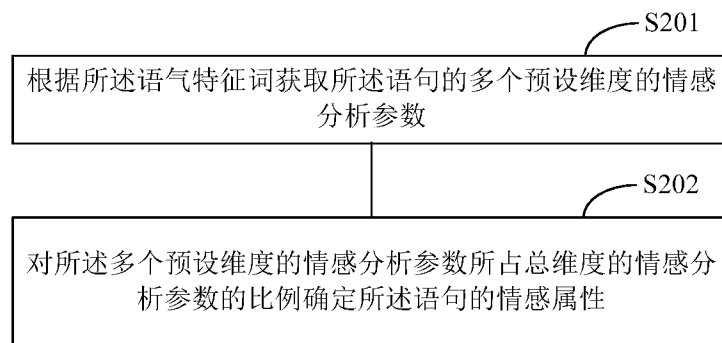


图 2

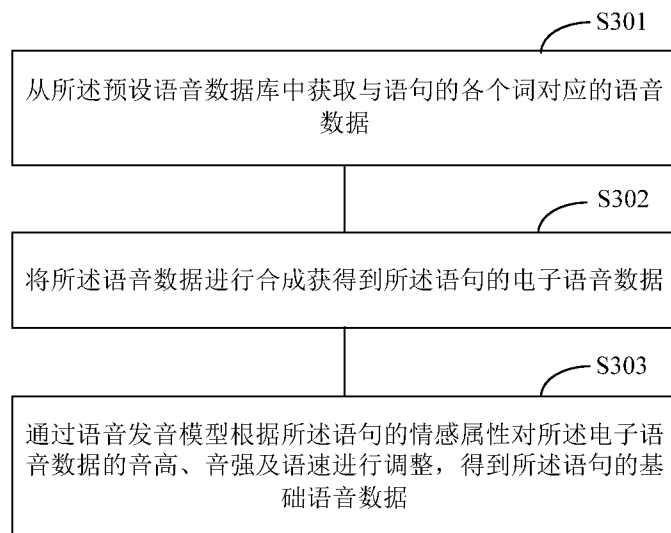


图 3

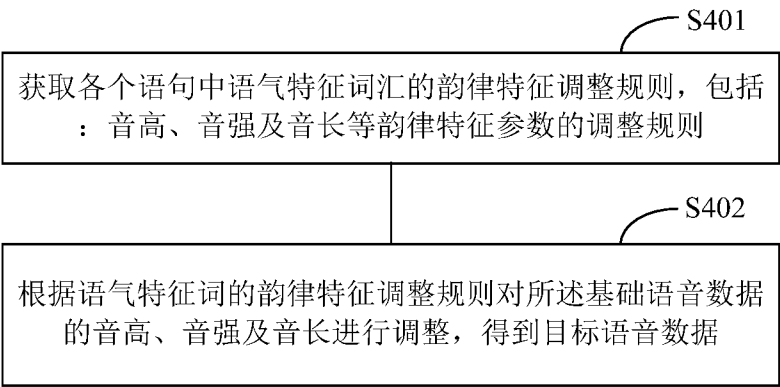


图 4

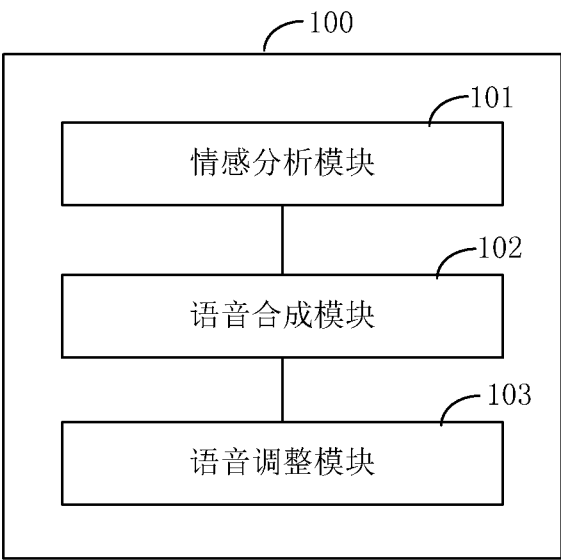


图 5

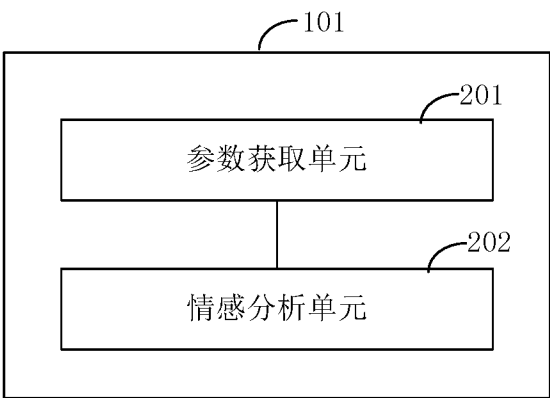


图 6

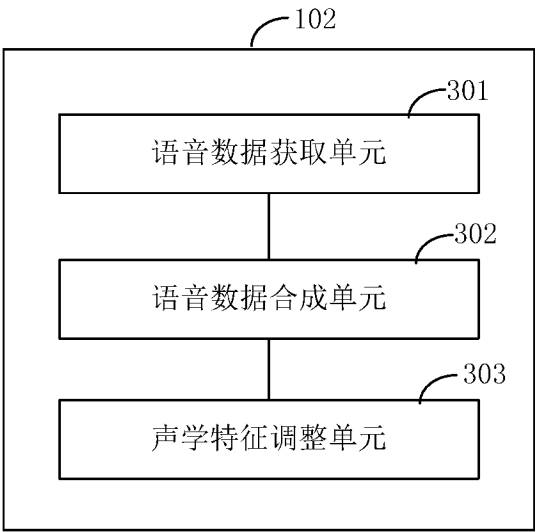


图 7

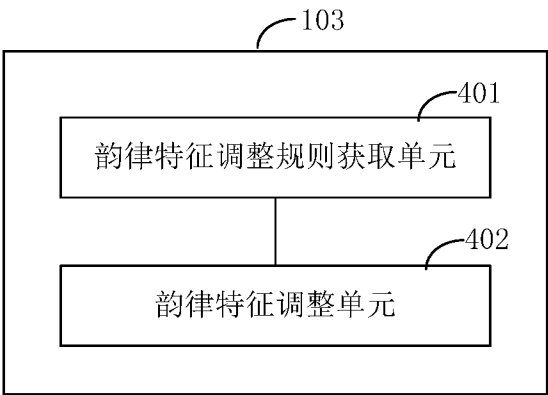


图 8

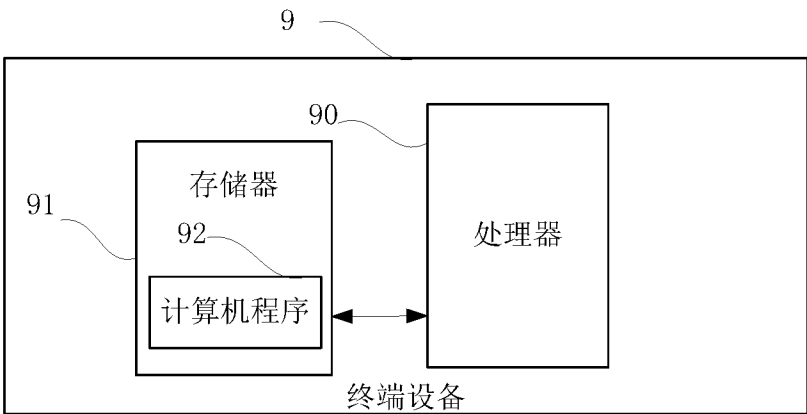


图 9

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2018/097560

A. CLASSIFICATION OF SUBJECT MATTER

G10L 13/08(2013.01)i; G10L 13/10(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)

VEN, CNABS: 语气, 情感, 感情, 语速, 韵律, 节奏, 快慢, 情绪, 语音, 合成, 数据库, 模型, TTS, speech, voice, synthes+, emotion+, feel+, model, prosodic, feature, person+

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 102103856 A (SHANDA NETWORKING CO., LTD.) 22 June 2011 (2011-06-22) description, paragraphs [0022]-[0068], and figures 1-5	1-20
X	CN 101452699 A (TOSHIBA CORPORATION) 10 June 2009 (2009-06-10) description, paragraphs [0005]-[0015], and figures 1-7	1-20
X	CN 101000765 A (HEILONGJIANG UNIVERSITY) 18 July 2007 (2007-07-18) description, pages 8-22, and figures 1-16	1-20
X	CN 101176146 A (PANASONIC CORPORATION) 07 May 2008 (2008-05-07) description, pages 8-34, and figures 1-36	1-20
A	CN 101064103 A (INSTITUTE OF AUTOMATION, CHINESE ACADEMY OF SCIENCES) 31 October 2007 (2007-10-31) entire document	1-20
A	CN 103198827 A (HEFEI UNIVERSITY OF TECHNOLOGY) 10 July 2013 (2013-07-10) entire document	1-20
A	CN 103366731 A (SHENGLE INFORMATION TECHNOLOGY (SHANGHAI) CO., LTD.) 23 October 2013 (2013-10-23) 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

13 February 2019

Date of mailing of the international search report

20 February 2019

Name and mailing address of the ISA/CN

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

Authorized officer

Facsimile No. (86-10)62019451

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2018/097560**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN 106688034 A (MICROSOFT TECHNOLOGY LICENSING, LLC) 17 May 2017 (2017-05-17) entire document	1-20
A	CN 105355193 A (BAIDU ONLINE NETWORK TECHNOLOGY (BEIJING) CO., LTD.) 24 February 2016 (2016-02-24) entire document	1-20
A	US 2015046164 A1 (SAMSUNG ELECTRONICS CO., LTD.) 12 February 2015 (2015-02-12) entire document	1-20
A	KR 20080060909 A (LG ELECTRONICS INC.) 02 July 2008 (2008-07-02) entire document	1-20
A	KR 20120117041 A (KOREA ADVANCED INSTITUTE OF SCIENCE AND TECHNOLOGY) 24 October 2012 (2012-10-24) entire document	1-20

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2018/097560

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	102103856	A	22 June 2011	None			
CN	101452699	A	10 June 2009	JP	2009139949	A	25 June 2009
				US	2009157409	A1	18 June 2009
CN	101000765	A	18 July 2007	CN	101000765	B	30 March 2011
CN	101176146	A	07 May 2008	JP	WO2006123539	A1	25 December 2008
				JP	4125362	B2	30 July 2008
				US	2009234652	A1	17 September 2009
				CN	101176146	B	18 May 2011
				WO	2006123539	A1	23 November 2006
				US	8073696	B2	06 December 2011
CN	101064103	A	31 October 2007	CN	101064103	B	04 May 2011
CN	103198827	A	10 July 2013	CN	103198827	B	17 June 2015
CN	103366731	A	23 October 2013	None			
CN	106688034	A	17 May 2017	WO	2016040209	A1	17 March 2016
				EP	3192070	A1	19 July 2017
				US	2016078859	A1	17 March 2016
				US	9824681	B2	21 November 2017
CN	105355193	A	24 February 2016	None			
US	2015046164	A1	12 February 2015	None			
KR	20080060909	A	02 July 2008	None			
KR	20120117041	A	24 October 2012	KR	101203188	B1	22 November 2012

国际检索报告

国际申请号

PCT/CN2018/097560

A. 主题的分类 G10L 13/08(2013.01)i; G10L 13/10(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)) VEN, CNABS: 语气, 情感, 感情, 语速, 韵律, 节奏, 快慢, 情绪, 语音, 合成, 数据库, 模型, TTS, speech, voice, synthes+, emotion+, feel+, model, prosodic, feature, person+																													
C. 相关文件 <table border="1"> <thead> <tr> <th>类 型*</th> <th>引用文件, 必要时, 指明相关段落</th> <th>相关的权利要求</th> </tr> </thead> <tbody> <tr> <td>X</td> <td>CN 102103856 A (盛大计算机上海有限公司) 2011年 6月 22日 (2011 - 06 - 22) 说明书第[0022]-[0068]段, 图1-5</td> <td>1-20</td> </tr> <tr> <td>X</td> <td>CN 101452699 A (株式会社东芝) 2009年 6月 10日 (2009 - 06 - 10) 说明书第5-15页, 图1-7</td> <td>1-20</td> </tr> <tr> <td>X</td> <td>CN 101000765 A (黑龙江大学) 2007年 7月 18日 (2007 - 07 - 18) 说明书第8-22页, 图1-16</td> <td>1-20</td> </tr> <tr> <td>X</td> <td>CN 101176146 A (松下电器产业株式会社) 2008年 5月 7日 (2008 - 05 - 07) 说明书第8-34页, 图1-36</td> <td>1-20</td> </tr> <tr> <td>A</td> <td>CN 101064103 A (中国科学院自动化研究所) 2007年 10月 31日 (2007 - 10 - 31) 全文</td> <td>1-20</td> </tr> <tr> <td>A</td> <td>CN 103198827 A (合肥工业大学) 2013年 7月 10日 (2013 - 07 - 10) 全文</td> <td>1-20</td> </tr> <tr> <td>A</td> <td>CN 103366731 A (盛乐信息技术上海有限公司) 2013年 10月 23日 (2013 - 10 - 23) 全文</td> <td>1-20</td> </tr> <tr> <td>A</td> <td>CN 106688034 A (微软技术许可有限责任公司) 2017年 5月 17日 (2017 - 05 - 17) 全文</td> <td>1-20</td> </tr> </tbody> </table>			类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求	X	CN 102103856 A (盛大计算机上海有限公司) 2011年 6月 22日 (2011 - 06 - 22) 说明书第[0022]-[0068]段, 图1-5	1-20	X	CN 101452699 A (株式会社东芝) 2009年 6月 10日 (2009 - 06 - 10) 说明书第5-15页, 图1-7	1-20	X	CN 101000765 A (黑龙江大学) 2007年 7月 18日 (2007 - 07 - 18) 说明书第8-22页, 图1-16	1-20	X	CN 101176146 A (松下电器产业株式会社) 2008年 5月 7日 (2008 - 05 - 07) 说明书第8-34页, 图1-36	1-20	A	CN 101064103 A (中国科学院自动化研究所) 2007年 10月 31日 (2007 - 10 - 31) 全文	1-20	A	CN 103198827 A (合肥工业大学) 2013年 7月 10日 (2013 - 07 - 10) 全文	1-20	A	CN 103366731 A (盛乐信息技术上海有限公司) 2013年 10月 23日 (2013 - 10 - 23) 全文	1-20	A	CN 106688034 A (微软技术许可有限责任公司) 2017年 5月 17日 (2017 - 05 - 17) 全文	1-20
类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求																											
X	CN 102103856 A (盛大计算机上海有限公司) 2011年 6月 22日 (2011 - 06 - 22) 说明书第[0022]-[0068]段, 图1-5	1-20																											
X	CN 101452699 A (株式会社东芝) 2009年 6月 10日 (2009 - 06 - 10) 说明书第5-15页, 图1-7	1-20																											
X	CN 101000765 A (黑龙江大学) 2007年 7月 18日 (2007 - 07 - 18) 说明书第8-22页, 图1-16	1-20																											
X	CN 101176146 A (松下电器产业株式会社) 2008年 5月 7日 (2008 - 05 - 07) 说明书第8-34页, 图1-36	1-20																											
A	CN 101064103 A (中国科学院自动化研究所) 2007年 10月 31日 (2007 - 10 - 31) 全文	1-20																											
A	CN 103198827 A (合肥工业大学) 2013年 7月 10日 (2013 - 07 - 10) 全文	1-20																											
A	CN 103366731 A (盛乐信息技术上海有限公司) 2013年 10月 23日 (2013 - 10 - 23) 全文	1-20																											
A	CN 106688034 A (微软技术许可有限责任公司) 2017年 5月 17日 (2017 - 05 - 17) 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 2019年 2月 13日		国际检索报告邮寄日期 2019年 2月 20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 张亚玲 电话号码 86-(010)-62085722																											

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
A	CN 105355193 A (百度在线网络技术北京有限公司) 2016年 2月 24日 (2016 - 02 - 24) 全文	1-20
A	US 2015046164 A1 (SAMSUNG ELECTRONICS CO., LTD.) 2015年 2月 12日 (2015 - 02 - 12) 全文	1-20
A	KR 20080060909 A (LG ELECTRONICS INC.) 2008年 7月 2日 (2008 - 07 - 02) 全文	1-20
A	KR 20120117041 A (KOREA ADVANCED INST. SCI. & TECH.) 2012年 10月 24日 (2012 - 10 - 24) 全文	1-20

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2018/097560

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	102103856	A	2011年 6月 22日	无			
CN	101452699	A	2009年 6月 10日	JP	2009139949	A	2009年 6月 25日
				US	2009157409	A1	2009年 6月 18日
CN	101000765	A	2007年 7月 18日	CN	101000765	B	2011年 3月 30日
CN	101176146	A	2008年 5月 7日	JP	W02006123539	A1	2008年 12月 25日
				JP	4125362	B2	2008年 7月 30日
				US	2009234652	A1	2009年 9月 17日
				CN	101176146	B	2011年 5月 18日
				WO	2006123539	A1	2006年 11月 23日
				US	8073696	B2	2011年 12月 6日
CN	101064103	A	2007年 10月 31日	CN	101064103	B	2011年 5月 4日
CN	103198827	A	2013年 7月 10日	CN	103198827	B	2015年 6月 17日
CN	103366731	A	2013年 10月 23日	无			
CN	106688034	A	2017年 5月 17日	WO	2016040209	A1	2016年 3月 17日
				EP	3192070	A1	2017年 7月 19日
				US	2016078859	A1	2016年 3月 17日
				US	9824681	B2	2017年 11月 21日
CN	105355193	A	2016年 2月 24日	无			
US	2015046164	A1	2015年 2月 12日	无			
KR	20080060909	A	2008年 7月 2日	无			
KR	20120117041	A	2012年 10月 24日	KR	101203188	B1	2012年 11月 22日

表 PCT/ISA/210 (同族专利附件) (2015年1月)