

(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
2022 年 7 月 7 日 (07.07.2022)



(10) 国际公布号
WO 2022/142105 A1

(51) 国际专利分类号:
G10L 13/04 (2013.01) *G06F 40/289* (2020.01)
G10L 13/10 (2013.01) *G06F 40/284* (2020.01)
G06F 40/30 (2020.01)

(21) 国际申请号: PCT/CN2021/096524

(22) 国际申请日: 2021 年 5 月 27 日 (27.05.2021)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
202011640955.5 2020 年 12 月 31 日 (31.12.2020) 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: 李俊杰(LI, JunJie); 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)。 马骏(MA, Jun); 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)。 王少军(WANG, ShaoJun); 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: 深圳市沃德知识产权代理事务所(普通合伙)(SHENZHEN WORLD

(54) Title: TEXT-TO-SPEECH CONVERSION METHOD AND APPARATUS, ELECTRONIC DEVICE, AND STORAGE MEDIUM

(54) 发明名称: 文本转语音方法、装置、电子设备及存储介质

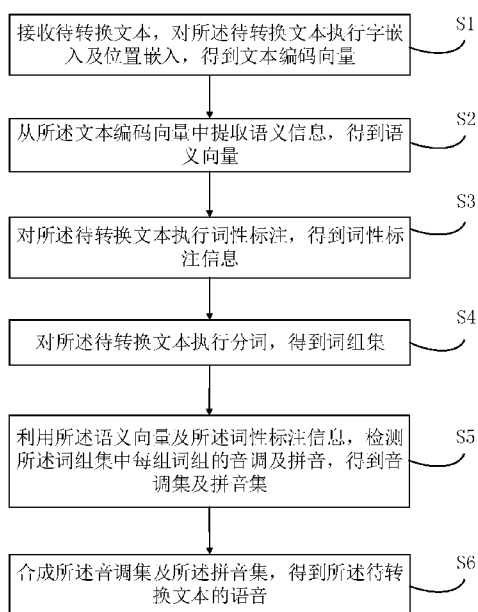


图 1

S1 Receive a text to be converted, and perform word embedding and position embedding on said text so as to obtain a text encoding vector
S2 Extract semantic information from the text encoding vector to obtain a semantic vector
S3 Perform part-of-speech tagging on said text to obtain part-of-speech tagging information
S4 Perform word segmentation on said text to obtain a phrase set
S5 Detect the tones and pinyin of each phrase in the phrase set by using the semantic vector and the part-of-speech tagging information, so as to obtain a tone set and a pinyin set
S6 Synthesize the tone set and the pinyin set to obtain the speech of said text

(57) Abstract: The present application relates to a natural language processing technology. Disclosed is a text-to-speech conversion method, comprising: receiving a text to be converted, and performing word embedding and position embedding on said text so as to obtain a text encoding vector; extracting semantic information from the text encoding vector to obtain a semantic vector; performing part-of-speech tagging on said text to obtain part-of-speech tagging information; performing word segmentation on said text to obtain a phrase set; detecting the tones and pinyin of each phrase in the phrase set by using the semantic vector and the part-of-speech tagging information, so as to obtain a tone set and a pinyin set; and synthesizing the tone set and the pinyin set to obtain the speech of said text. Further disclosed in the present application are a text-to-speech conversion apparatus, an electronic device, and a storage medium. The present application can solve the problem of low accuracy of text-to-speech conversion.

INTELLECTUAL PROPERTY AGENCY (GENERAL PARTNERSHIP)); 中国广东省深圳市福田区梅林街道中康路128号卓越城1期2栋301B
于志光, Guangdong 518000 (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, IT, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW。
- (84) 指定国(除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

- 包括国际检索报告(条约第21条(3))。

(57) 摘要: 本申请涉及自然语言处理技术, 揭露一种文本转语音方法, 包括: 接收待转换文本, 对所述待转换文本执行字嵌入及位置嵌入, 得到文本编码向量, 从所述文本编码向量中提取语义信息, 得到语义向量, 对所述待转换文本执行词性标注, 得到词性标注信息, 对所述待转换文本执行分词, 得到词组集, 利用所述语义向量及所述词性标注信息, 检测所述词组集中每组词组的音调及拼音, 得到音调集及拼音集, 合成所述音调集及所述拼音集, 得到所述待转换文本的语音。本申请还揭露文本转语音装置、电子设备以及存储介质。本申请可解决文本转语音的准确率较低。

文本转语音方法、装置、电子设备及存储介质

本申请要求于2020年12月31日提交中国专利局、申请号为CN202011640955.5，发明名称为“文本转语音方法、装置、电子设备及存储介质”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

本申请涉及自然语言处理领域，尤其涉及一种文本转语音方法、装置、电子设备及计算机可读存储介质。

背景技术

伴随人工智能蓬勃发展，文本转语音技术（TTS）也成为热点的技术突破方向，发明人意识到目前文本转语音技术主要结合中文分词、多音字消歧、韵律划分等传统方法，由于缺少对文本的深度提取，导致文本转语音的准确率较低。也有部分文本转语音技术基于神经网络来提升数据特征的效果，但所使用的神经网络较简单，导致文本转语音的准确率有待进一步提高。

发明内容

一种文本转语音方法，包括：

接收待转换文本，对所述待转换文本执行字嵌入及位置嵌入，得到文本编码向量；

从所述文本编码向量中提取语义信息，得到语义向量；

对所述待转换文本执行词性标注，得到词性标注信息；

对所述待转换文本执行分词，得到词组集；

利用所述语义向量及所述词性标注信息，检测所述词组集中每组词组的音调及拼音，得到音调集及拼音集；

合成所述音调集及所述拼音集，得到所述待转换文本的语音。

一种文本转语音装置，所述装置包括：

文本编码模块，用于接收待转换文本，对所述待转换文本执行字嵌入及位置嵌入，得到文本编码向量；

词性标注模块，用于从所述文本编码向量中提取语义信息，得到语义向量，对所述待转换文本执行词性标注，得到词性标注信息；

音调拼音预测模块，用于对所述待转换文本执行分词，得到词组集，利用所述语义向量及所述词性标注信息，检测所述词组集中每组词组的音调及拼音，得到音调集及拼音集；

音调拼音合成模块，用于合成所述音调集及所述拼音集，得到所述待转换文本的语音。

一种电子设备，所述电子设备包括：

存储器，存储至少一个指令；及

处理器，执行所述存储器中存储的指令以实现如下步骤：

接收待转换文本，对所述待转换文本执行字嵌入及位置嵌入，得到文本编码向量；

从所述文本编码向量中提取语义信息，得到语义向量；

对所述待转换文本执行词性标注，得到词性标注信息；

对所述待转换文本执行分词，得到词组集；

利用所述语义向量及所述词性标注信息，检测所述词组集中每组词组的音调及拼音，得到音调集及拼音集；

合成所述音调集及所述拼音集，得到所述待转换文本的语音。

一种计算机可读存储介质，包括存储数据区和存储程序区，存储数据区存储创建的数据，存储程序区存储有计算机程序；其中，所述计算机程序被处理器执行时实现如下步骤：

接收待转换文本，对所述待转换文本执行字嵌入及位置嵌入，得到文本编码向量；

从所述文本编码向量中提取语义信息，得到语义向量；

对所述待转换文本执行词性标注，得到词性标注信息；

对所述待转换文本执行分词，得到词组集；

- 5 利用所述语义向量及所述词性标注信息，检测所述词组集中每组词组的音调及拼音，得到音调集及拼音集；

合成所述音调集及所述拼音集，得到所述待转换文本的语音。

本申请可以解决文本转语音的准确率较低。

10 附图说明

图1为本申请一实施例提供的文本转语音方法的流程示意图；

图2为本申请一实施例提供的文本转语音方法中S1的流程示意图；

图3为本申请一实施例提供的文本转语音装置的模块示意图；

图4为本申请一实施例提供的实现文本转语音方法的电子设备的内部结构示意图；

- 15 本申请目的的实现、功能特点及优点将结合实施例，参照附图做进一步说明。

具体实施方式

应当理解，此处所描述的具体实施例仅仅用以解释本申请，并不用于限定本申请。

- 20 本申请实施例提供一种文本转语音方法，所述文本转语音方法的执行主体包括但不限于服务端、终端等能够被配置为执行本申请实施例提供的该方法的电子设备中的至少一种。换言之，所述文本转语音方法可以由安装在终端设备或服务端设备的软件或硬件来执行，所述软件可以是区块链平台。所述服务端包括但不限于：单台服务器、服务器集群、云端服务器或云端服务器集群等。

- 25 参照图1所示，为本申请一实施例提供的文本转语音方法的流程示意图。在本实施例中，所述文本转语音方法包括：

S1、接收待转换文本，对所述待转换文本执行字嵌入及位置嵌入，得到文本编码向量。

本申请实施例中，可根据用户输入或前端程序推送等得到待转换文本，如用户输入待转换文本A：“城市发展，需要每个人贡献力量”。

- 30 进一步地，由于计算机无法有效从当前格式的待转换文本中提取语义信息，需要先将待转换文本转化为向量形式，详细地，参阅图1所示，所述对所述待转换文本执行字嵌入及位置嵌入，得到文本编码向量，包括：

S11、将所述待转换文本执行去噪及分词处理，得到待转换词组集；

- 35 由于待转换文本可能包含非文本数据，例如超链接、乱码等，因此需要对待转换文本进行执行去噪处理。本申请实施例中，去噪处理可采用基于编程语言为基础构建的正则表达式，完成去除如表情符号及乱码等噪声。

详细地，本申请实施例利用结巴分词方法，对去噪后的待转换文本执行分词，得到与待转换文本对应的待转换词组集。

S12、根据所述待转换词组集的数量，构建相同数量且向量维度固定的空向量，将所述待转换词组集中每个待转换词组映射至所述空向量中，得到字嵌入向量集；

- 40 如上述待转换文本A：“城市发展，需要每个人贡献力量”对应的转换词组集为：“城市”、“发展”、“需要”、“每个人”、“贡献”、“出”、“力量”。可见转换词组集的数量为8，则本申请实施例对应构建8组空向量，其中每组向量维度均为512。

进一步地，本申请实施例可使用one-hot模型或word2vec模型，将待转换词组集映射至所述空向量中，得到字嵌入向量集。

- 45 S13、计算所述待转换词组集中每个待转换词组在所述待转换文本的位置，得到词组

位置集, 根据所述词组位置集构建位置向量集;

如上述待转换词组集为:“城市”、“发展”、“需要”、“每个人”、“贡献”、“出”、“力量”, 其中待转换词组“城市”在待转换文本 A:“城市发展, 需要每个人贡献出力量”的位置最靠前, 待转换词组“力量”在待转换文本最靠后, 从而汇总得到待转换词组集对应的词组位置集, 并进一步根据 word2vec 模型, 将所述词组位置集编码得到所述位置向量集。

S14、组合所述字嵌入向量集及所述位置向量集, 得到所述文本编码向量。

本申请实施例中, 按照对应关系组合所述字嵌入向量集及所述位置向量集, 如字嵌入向量集为[a,b,c,d], 位置向量集为[r,t,y,u], 则组合后的文本编码向量为[ar,bt,cy,du]。

S2、从所述文本编码向量中提取语义信息, 得到语义向量。

本申请实施例中, 由于所述文本编码向量是由所述字嵌入向量集及所述位置向量集组合得到, 且所述字嵌入向量集中向量维度为 512, 由此可见所述文本编码向量的向量维度更加庞大, 因此需要提取出向量维度更小的语义信息, 以减轻计算压力, 并方便后续步骤执行。

详细地, 所述 S2 包括: 构建并训练语义提取模型, 当训练完成时, 从所述语义提取模型中提取自注意力计算机制, 利用所述自注意力计算机制, 提取所述文本编码向量的语义信息, 得到所述语义向量。

本申请实施例中, 所述语义提取模型采用当前已公开的 transform 模型。故先构建并训练 transform 模型, 其中 transform 模型包括编码层、解码层, 编码层包括自注意力计算机制, 当 transform 模型训练完成后, 从 transform 模型中提取所述自注意力计算机制。其中自注意力计算机制是一种考虑待转换文本中句子与句子之间, 不同词语之间的语义以及语法联系的机制。

进一步地, 所述训练语义提取模型包括: 获取训练文本集及语义标签集; 将所述训练文本集输入至语义提取模型中, 计算所述训练文本集中每个训练文本与其他训练文本的相似度, 得到相似度集; 归一化所述相似度集得到权重集, 利用所述权重集生成所述训练文本集的编码向量集; 对所述编码向量集及所述训练文本集执行解码操作, 得到语义预测集; 计算所述语义预测集与所述语义标签集的误差值, 判断所述误差值是否在预设误差范围内, 当所述语义预测集与所述语义标签集的误差值不在预设误差范围内, 调整所述语义提取模型的内部参数, 直至所述语义预测集与所述语义标签集的误差值在预设误差范围内, 得到训练完成的所述语义提取模型。

本申请实施例中, 利用训练文本集和语义标签集训练语义提取模型, 如从训练文本集中提取训练文本 I, 将训练文本 I 转为词向量, 并和训练文本集中其他词向量相乘, 如 you、and、me 等词向量, 进而得到相似度, 对相似度执行归一化操作得到权重, 再继续利用权重和 you、and、me 等词向量相乘, 最后相加构建出训练文本 I 的编码向量, 以此类推, 得到所述训练文本集内每个训练文本的编码向量, 汇总得到所述编码向量集。

进一步地, 可采用当前已公开的 transform 模型解码层的具体操作实例, 对所述编码向量集及所述训练文本集执行解码操作。

当解码操作得到语义预测集时, 本申请实施例利用最小二乘法计算所述语义预测集与所述语义标签集的误差值, 直至误差值满足预设要求, 得到训练完成的所述语义提取模型。

进一步地, 从训练完成的语义提取模型中提取自注意力计算机制, 其中自注意力计算机制提取所述文本编码向量的语义信息, 与训练语义提取模型得到编码向量集相似, 主要包括计算得到相似度集、归一化相似度集得到权重集, 利用权重集生成所述文本编码向量的语义向量。

S3、对所述待转换文本执行词性标注, 得到词性标注信息。

详细地, 所述 S3 包括: 将所述待转换文本转化为待转换词向量, 利用 CRF 模型对所述待转换词向量执行词性标注, 得到所述词性标注信息。

本申请实施例中，使用 Word2Vec 模型，将所述待转换文本转化为待转换词向量，例如，上述待转换文本 A：“城市发展，需要每个人贡献出力量”，通过 Word2Vec 模型将待转换文本 A 转化为待转换词向量：
$$\begin{bmatrix} 17 & \cdots & 26 \\ \vdots & \ddots & \vdots \\ 94 & \cdots & 37 \end{bmatrix}$$

本申请实施例中，利用 CRF (Conditional Random Field, 条件随机场) 模型对所述待转换词向量执行词性标注，得到词性标注信息，其中词性标注信息用单维向量表示，如 $\begin{bmatrix} 1 \\ \vdots \\ 3 \end{bmatrix}$ 中，1 表示 [17 ... 26] 代表的是名词，3 表示 [94 ... 37] 是动词等。

S4、对所述待转换文本执行分词，得到词组集。

本申请实施例中，利用结巴分词方法对所述待转换文本执行分词，得到词组集。

10 S5、利用所述语义向量及所述词性标注信息，检测所述词组集中每组词组的音调及拼音，得到音调集及拼音集。

详细地，所述 S5 包括：将所述语义向量及所述词性标注信息输入至预先训练完成的音调拼音检测模型中，利用所述音调拼音检测模型，检测所述词组集中每组词组的音调及拼音，得到所述音调集及所述拼音集。

进一步地，预先训练完成的音调拼音检测模型包括：

15 步骤 A：构建包括多头注意力网络及线性激活层的待训练音调拼音检测模型。

详细地，所述步骤 A 包括：根据多头注意力机制构建所述多头注意力网络；利用线性激活函数构建所述线性激活层，组合所述多头注意力网络及所述线性激活层，得到所述待训练音调拼音检测模型。

20 其中，所述根据多头注意力机制构建所述多头注意力网络，包括：接收已训练完成的 Transform 模型，从所述 Transform 模型中提取编码器，利用所述编码器内的多头注意力机制，构建得到所述多头注意力网络。

本申请实施例中，用户可提前训练完成 Transform 模型，所述 Transform 模型与上述 S2 的 Transform 模型相同，包括编码器、解码器，其中编码器中包括多头注意力机制，本申请实施例中，提取所述多头注意力机制所在的网络层，构建得到所述多头注意力网络。

25 本申请构建线性激活层帮助所述待训练注意力特征模型进行模型训练，其中所述线性激活层包括归一化和激活函数，所述激活函数可使用高斯分布函数。

进一步地，本申请实施例中，按照所述多头注意力网络在前，所述线性激活层在后的原则，组合得到所述待训练音调拼音检测模型。

30 步骤 B：接收词性文本训练集及音调拼音标签集，将所述词性文本训练集输入至所述待训练音调拼音检测模型进行特征提取，得到特征序列训练集。

本申请实施例中，所述词性文本训练集是用户提前收集并整理出的包括文本及文本中每个词语词性的集合，所述音调拼音标签集是与所述词性文本训练集对应的语音集合及拼音集合，如所述词性文本训练集中有训练文本 X_1 ：“城市（名词）美化（动词），是（动词）一项（量词）艰巨的（形容词）任务（名词）”，则在所述词性文本训练集中对应存在语音 $Y_1 = (y_1, y_2, \dots, y_n)$ 和拼音 $S_1 = (s_1, s_2, \dots, s_n)$ ，其中 y_n 表示语音 Y_1 的语音序列，其中语音序列的振幅、大小等可体现出音频。

40 进一步地，当获得所述词性文本训练集后，利用所述待训练音调拼音检测模型进行特征提取，详细地，所述将所述词性文本训练集输入至所述待训练音调拼音检测模型进行特征提取，得到特征序列训练集，包括：利用所述多头注意力网络，对所述词性文本训练集执行编码操作，得到所述特征序列训练集。

本申请实施例中，对词性文本训练集执行编码操作的方式与 Transform 模型中编程操作相同。

步骤 C: 利用所述线性激活层, 对所述特征序列训练集执行激活操作, 得到音调拼音预测集。

详细地, 所述利用所述线性激活层, 对所述特征序列训练集执行激活操作, 得到预测序列集, 包括: 对所述特征序列训练集执行归一化得到特征序列归一化集, 利用所述高斯分布函数, 计算所述特征序列归一化集的高斯分布, 根据所述高斯分布得到所述音调拼音预测集。

详细地, 所述归一化是将所述特征序列训练集内的数值映射至指定范围内的操作, 如将所述特征序列训练集内的数值映射至[0,1]范围内, 可有效缩小数值, 减轻计算压力。

进一步地, 所述利用所述高斯分布函数, 计算所述特征序列归一化集的高斯分布, 包括: 利用所述高斯分布函数计算出所述特征序列归一化集的均值和方差, 利用所述特征序列归一化集的均值和方差, 求解出所述特征序列归一化集的高斯分布。

由于高斯分布是展现数据在指定范围内的概率分布, 故本申请实施例中, 从高斯分布中寻找出所述特征序列训练集的最大概率分布, 即得到所述音调拼音预测集。

步骤 D: 计算所述音调拼音预测集与所述音调拼音标签集的预测误差值, 并判断所述预测误差值与预设的误差阈值的大小关系。

本申请实施例中, 利于平方差公式计算所述音调拼音预测集与所述音调拼音标签集的预测误差值。

步骤 E: 若所述预测误差值大于所述误差阈值, 调整所述待训练音调拼音检测模型的内部参数, 并返回步骤 B。

步骤 F: 若所述预测误差值小于或等于所述误差阈值, 得到所述音调拼音检测模型。

详细地, 当所述误差值小于或等于所述误差阈值, 表示所述待训练音调拼音检测模型具有较强的预测能力,。

本申请实施例中, 当执行步骤 A 至步骤 F 得到训练完成的音调拼音检测模型后, 进一步地, 可将所述语义向量及所述词性标注信息输入至音调拼音检测模型中, 检测所述词组集中每组词组的音调及拼音, 得到所述音调集及所述拼音集。

S6、合成所述音调集及所述拼音集, 得到所述待转换文本的语音。

本申请实施例中, 利用声码合成器合成所述音调集及所述拼音集, 声码合成器是一种可实现语音合成的译码器, 本申请实施例中, 将所述音调集及所述拼音集输入至所述声码合成器, 即可得到所述待转换文本的语音。

本申请实施例在执行文本转语音之前, 先对待转换文本执行字嵌入及位置嵌入得到文本编码向量, 并从所述文本编码向量中提取语义信息得到语义向量, 可见并非根据背景技术所述, 将待转换文本执行中文分词、多音字消歧操作后直接进行转语音操作, 故本申请实施例提高了对待转换文本的深度提取, 且进一步地, 本申请实施例结合待转换文本对应的词性标注信息, 检测出待转换文本的音调集及拼音集, 进而根据音调集及拼音集成出语音, 由于词性会影响语音的音调, 因此本申请实施例进一步避免了在文本转语音过程中, 受词性影响导致语音发音不准确的问题。因此本申请提出的文本转语音方法、装置及计算机可读存储介质, 可以解决文本转语音的准确率较低。

如图 3 所示, 是本申请文本转语音装置的模块示意图。

本申请所述文本转语音装置 100 可以安装于电子设备中。根据实现的功能, 所述文本转语音装置可以包括文本编码模块 101、词性标注模块 102、音调拼音预测模块 103 及音调拼音合成模块 104。本发所述模块也可以称之为单元, 是指一种能够被电子设备处理器所执行, 并且能够完成固定功能的一系列计算机程序段, 其存储在电子设备的存储器中。

在本实施例中, 关于各模块/单元的功能如下:

所述文本编码模块 101, 用于接收待转换文本, 对所述待转换文本执行字嵌入及位置嵌入, 得到文本编码向量;

所述词性标注模块 102, 用于从所述文本编码向量中提取语义信息, 得到语义向量, 对所述待转换文本执行词性标注, 得到词性标注信息;

所述音调拼音预测模块 103, 用于对所述待转换文本执行分词, 得到词组集, 利用所述语义向量及所述词性标注信息, 检测所述词组集中每组词组的音调及拼音, 得到音调集及拼音集;

所述音调拼音合成模块 104, 用于合成所述音调集及所述拼音集, 得到所述待转换文本的语音。

本申请实施例所提供的文本转语音装置 100 中的各个模块能够在使用时基于与上述的文本转语音方法采用相同的手段, 具体地实施步骤在此不再赘述, 关于各模块/单元的功能所产生技术效果与上述的文本转语音方法的技术效果相同的, 即解决文本转语音的准确率较低。

如图 4 所示, 是本申请实现文本转语音方法的电子设备的结构示意图。

所述电子设备 1 可以包括处理器 10、存储器 11 和总线, 还可以包括存储在所述存储器 11 中并可在所述处理器 10 上运行的计算机程序, 如文本转语音程序 12。

其中, 所述存储器 11 至少包括一种类型的可读存储介质, 所述可读存储介质包括闪存、移动硬盘、多媒体卡、卡型存储器 (例如: SD 或 DX 存储器等)、磁性存储器、磁盘、光盘等。所述存储器 11 在一些实施例中可以是电子设备 1 的内部存储单元, 例如该电子设备 1 的移动硬盘。所述存储器 11 在另一些实施例中也可以是电子设备 1 的外部存储设备, 例如电子设备 1 上配备的插接式移动硬盘、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 11 还可以既包括电子设备 1 的内部存储单元也包括外部存储设备。所述存储器 11 不仅可以用于存储安装于电子设备 1 的应用软件及各类数据, 例如文本转语音程序 12 的代码等, 还可以用于暂时地存储已经输出或者将要输出的数据。

所述处理器 10 在一些实施例中可以由集成电路组成, 例如可以由单个封装的集成电路所组成, 也可以是由多个相同功能或不同功能封装的集成电路所组成, 包括一个或者多个中央处理器 (Central Processing unit, CPU)、微处理器、数字处理芯片、图形处理器及各种控制芯片的组合等。所述处理器 10 是所述电子设备的控制核心 (Control Unit), 利用各种接口和线路连接整个电子设备的各个部件, 通过运行或执行存储在所述存储器 11 内的程序或者模块 (例如执行文本转语音程序等), 以及调用存储在所述存储器 11 内的数据, 以执行电子设备 1 的各种功能和处理数据。

所述总线可以是外设部件互连标准 (peripheral component interconnect, 简称 PCI) 总线或扩展工业标准结构 (extended industry standard architecture, 简称 EISA) 总线等。该总线可以分为地址总线、数据总线、控制总线等。所述总线被设置为实现所述存储器 11 以及至少一个处理器 10 等之间的连接通信。

图 4 仅示出了具有部件的电子设备的, 本领域技术人员可以理解的是, 图 4 示出的结构并不构成对所述电子设备 1 的限定, 可以包括比图示更少或者更多的部件, 或者组合某些部件, 或者不同的部件布置。

例如, 尽管未示出, 所述电子设备 1 还可以包括给各个部件供电的电源 (比如电池), 优选地, 电源可以通过电源管理装置与所述至少一个处理器 10 逻辑相连, 从而通过电源管理装置实现充电管理、放电管理、以及功耗管理等功能。电源还可以包括一个或一个以上的直流或交流电源、再充电装置、电源故障检测电路、电源转换器或者逆变器、电源状态指示器等任意组件。所述电子设备 1 还可以包括多种传感器、蓝牙模块、Wi-Fi 模块等, 在此不再赘述。

进一步地, 所述电子设备 1 还可以包括网络接口, 可选地, 所述网络接口可以包括有线接口和/或无线接口 (如 WI-FI 接口、蓝牙接口等), 通常用于在该电子设备 1 与其他电

子设备之间建立通信连接。

可选地，该电子设备 1 还可以包括用户接口，用户接口可以是显示器（Display）、输入单元（比如键盘（Keyboard）），可选地，用户接口还可以是标准的有线接口、无线接口。可选地，在一些实施例中，显示器可以是 LED 显示器、液晶显示器、触控式液晶显示器以及 OLED（Organic Light-Emitting Diode，有机发光二极管）触摸器等。其中，显示器也可以适当的称为显示屏或显示单元，用于显示在电子设备 1 中处理的信息以及用于显示可视化的用户界面。

应该了解，所述实施例仅为说明之用，在专利申请范围上并不受此结构的限制。

所述电子设备 1 中的所述存储器 11 存储的文本转语音程序 12 是多个指令的组合，在所述处理器 10 中运行时，可以实现：

接收待转换文本，对所述待转换文本执行字嵌入及位置嵌入，得到文本编码向量；

从所述文本编码向量中提取语义信息，得到语义向量；

对所述待转换文本执行词性标注，得到词性标注信息；

对所述待转换文本执行分词，得到词组集；

利用所述语义向量及所述词性标注信息，检测所述词组集中每组词组的音调及拼音，得到音调集及拼音集；

合成所述音调集及所述拼音集，得到所述待转换文本的语音。

进一步地，所述电子设备 1 集成的模块/单元如果以软件功能单元的形式实现并作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。所述计算机可读存储介质可以是易失性的，也可以是非易失性的，所述计算机可读介质可以包括：能够携带所述计算机程序代码的任何实体或装置、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）。

进一步地，所述计算机可用存储介质可主要包括存储程序区和存储数据区，其中，存储程序区可存储操作系统、至少一个功能所需的应用程序等；存储数据区可存储根据区块链节点的使用所创建的数据等。

本申请还提供一种计算机可读存储介质，所述可读存储介质存储有计算机程序，所述计算机程序在被电子设备的处理器所执行时，可以实现：

接收待转换文本，对所述待转换文本执行字嵌入及位置嵌入，得到文本编码向量；

从所述文本编码向量中提取语义信息，得到语义向量；

对所述待转换文本执行词性标注，得到词性标注信息；

对所述待转换文本执行分词，得到词组集；

利用所述语义向量及所述词性标注信息，检测所述词组集中每组词组的音调及拼音，得到音调集及拼音集；

合成所述音调集及所述拼音集，得到所述待转换文本的语音。

在本申请所提供的几个实施例中，应该理解到，所揭露的设备、装置和方法，可以通过其它的方式实现。例如，以上所描述的装置实施例仅仅是示意性的，例如，所述模块的划分，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。

所述作为分离部件说明的模块可以是或者也可以不是物理上分开的，作为模块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模块来实现本实施例方案的目的。

另外，在本申请各个实施例中的各功能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，也可以采用硬件加软件功能模块的形式实现。

对于本领域技术人员而言，显然本申请不限于上述示范性实施例的细节，而且在不背离本申请的精神或基本特征的情况下，能够以其他的具体形式实现本申请。

因此，无论从哪一点来看，均应将实施例看作是示范性的，而且是非限制性的，本申请的范围由所附权利要求而不是上述说明限定，因此旨在将落在权利要求的等同要件的含义和范围内的所有变化涵括在本申请内。不应将权利要求中的任何附图关联图标记视为限制所涉及的权利要求。

5 本申请所指区块链是分布式数据存储、点对点传输、共识机制、加密算法等计算机技术的新型应用模式。区块链(Blockchain)，本质上是一个去中心化的数据库，是一串使用密码学方法相关联产生的数据块，每一个数据块中包含了一批次网络交易的信息，用于验证其信息的有效性(防伪)和生成下一个区块。区块链可以包括区块链底层平台、平台产品服务层以及应用服务层等。

10 此外，显然“包括”一词不排除其他单元或步骤，单数不排除复数。系统权利要求中陈述的多个单元或装置也可以由一个单元或装置通过软件或者硬件来实现。第二等词语用来表示名称，而并不表示任何特定的顺序。

15 最后应说明的是，以上实施例仅用以说明本申请的技术方案而非限制，尽管参照较佳实施例对本申请进行了详细说明，本领域的普通技术人员应当理解，可以对本申请的技术方案进行修改或等同替换，而不脱离本申请技术方案的精神和范围。

权利要求书

1、一种文本转语音方法，其中，所述方法包括：

接收待转换文本，对所述待转换文本执行字嵌入及位置嵌入，得到文本编码向量；

从所述文本编码向量中提取语义信息，得到语义向量；

对所述待转换文本执行词性标注，得到词性标注信息；

对所述待转换文本执行分词，得到词组集；

利用所述语义向量及所述词性标注信息，检测所述词组集中每组词组的音调及拼音，得到音调集及拼音集；

合成所述音调集及所述拼音集，得到所述待转换文本的语音。

2、如权利要求1所述的文本转语音方法，其中，所述对所述待转换文本执行字嵌入及位置嵌入，得到文本编码向量，包括：

将所述待转换文本执行去噪及分词处理，得到待转换词组集；

根据所述待转换词组集的数量，构建相同数量且向量维度固定的空向量，将所述待转换词组集中每个待转换词组映射至所述空向量中，得到字嵌入向量集；

计算所述待转换词组集中每个待转换词组在所述待转换文本的位置，得到词组位置集，根据所述词组位置集构建位置向量集；

组合所述字嵌入向量集及所述位置向量集，得到所述文本编码向量。

3、如权利要求1所述的文本转语音方法，其中，所述从所述文本编码向量中提取语义信息，得到语义向量，包括：

构建并训练语义提取模型，当训练完成时，从所述语义提取模型中提取自注意力计算机制；

利用所述自注意力计算机制，提取所述文本编码向量的语义信息，得到所述语义向量。

4、如权利要求3所述的文本转语音方法，其中，所述训练语义提取模型，包括：

获取训练文本集及语义标签集；

将所述训练文本集输入至所述语义提取模型中，计算所述训练文本集中每个训练文本与其他训练文本的相似度，得到相似度集；

归一化所述相似度集得到权重集，利用所述权重集生成所述训练文本集的编码向量集；

对所述编码向量集及所述训练文本集执行解码操作，得到语义预测集；

计算所述语义预测集与所述语义标签集的误差值，判断所述误差值是否在预设误差范围内；

当所述误差值不在所述预设误差范围内，调整所述语义提取模型的内部参数，直至所述误差值在所述预设误差范围内，得到训练完成的所述语义提取模型。

5、如权利要求1所述的文本转语音方法，其中，所述对所述待转换文本执行词性标注，得到词性标注信息，包括：

将所述待转换文本转化为待转换词向量，利用CRF模型对所述待转换词向量执行词性标注，得到所述词性标注信息。

6、如权利要求1所述的文本转语音方法，其中，所述利用所述语义向量及所述词性标注信息，检测所述词组集中每组词组的音调及拼音，得到音调集及拼音集，包括：

将所述语义向量及所述词性标注信息输入至预先训练完成的音调拼音检测模型中；

利用所述音调拼音检测模型，检测所述词组集中每组词组的音调及拼音，得到所述音调集及所述拼音集。

7、如权利要求1至6中任意一项所述的文本转语音方法，其中，所述预先训练完成的音调拼音检测模型，包括：

步骤A：构建包括多头注意力网络及线性激活层的待训练音调拼音检测模型；

步骤 B: 接收词性文本训练集及音调拼音标签集, 将所述词性文本训练集输入至所述待训练音调拼音检测模型进行特征提取, 得到特征序列训练集;

步骤 C: 利用所述线性激活层, 对所述特征序列训练集执行激活操作, 得到音调拼音预测集;

5 步骤 D: 计算所述音调拼音预测集与所述音调拼音标签集的预测误差值, 并判断所述预测误差值与预设的误差阈值的大小关系;

步骤 E: 若所述预测误差值大于所述误差阈值, 调整所述待训练音调拼音检测模型的内部参数, 并返回步骤 B;

步骤 F: 若所述预测误差值小于或等于所述误差阈值, 得到所述音调拼音检测模型。

10 8、一种文本转语音装置, 其中, 所述装置包括:

文本编码模块, 用于接收待转换文本, 对所述待转换文本执行字嵌入及位置嵌入, 得到文本编码向量;

词性标注模块, 用于从所述文本编码向量中提取语义信息, 得到语义向量, 对所述待转换文本执行词性标注, 得到词性标注信息;

15 音调拼音预测模块, 用于对所述待转换文本执行分词, 得到词组集, 利用所述语义向量及所述词性标注信息, 检测所述词组集中每组词组的音调及拼音, 得到音调集及拼音集;

音调拼音合成模块, 用于合成所述音调集及所述拼音集, 得到所述待转换文本的语音。

9、一种电子设备, 其中, 所述电子设备包括:

至少一个处理器; 以及,

20 与所述至少一个处理器通信连接的存储器; 其中,

所述存储器存储有可被所述至少一个处理器执行的指令, 所述指令被所述至少一个处理器执行, 以使所述至少一个处理器能够执行如下步骤:

接收待转换文本, 对所述待转换文本执行字嵌入及位置嵌入, 得到文本编码向量;

从所述文本编码向量中提取语义信息, 得到语义向量;

25 对所述待转换文本执行词性标注, 得到词性标注信息;

对所述待转换文本执行分词, 得到词组集;

利用所述语义向量及所述词性标注信息, 检测所述词组集中每组词组的音调及拼音, 得到音调集及拼音集;

合成所述音调集及所述拼音集, 得到所述待转换文本的语音。

30 10、如权利要求 9 所述的电子设备, 其中, 所述对所述待转换文本执行字嵌入及位置嵌入, 得到文本编码向量, 包括:

将所述待转换文本执行去噪及分词处理, 得到待转换词组集;

根据所述待转换词组集的数量, 构建相同数量且向量维度固定的空向量, 将所述待转换词组集中每个待转换词组映射至所述空向量中, 得到字嵌入向量集;

35 计算所述待转换词组集中每个待转换词组在所述待转换文本的位置, 得到词组位置集, 根据所述词组位置集构建位置向量集;

组合所述字嵌入向量集及所述位置向量集, 得到所述文本编码向量。

11、如权利要求 9 所述的电子设备, 其中, 所述从所述文本编码向量中提取语义信息, 得到语义向量, 包括:

40 构建并训练语义提取模型, 当训练完成时, 从所述语义提取模型中提取自注意力计算机制;

利用所述自注意力计算机制, 提取所述文本编码向量的语义信息, 得到所述语义向量。

12、如权利要求 11 所述的电子设备, 其中, 所述训练语义提取模型, 包括:

获取训练文本集及语义标签集;

45 将所述训练文本集输入至所述语义提取模型中, 计算所述训练文本集中每个训练文本

与其他训练文本的相似度，得到相似度集；

归一化所述相似度集得到权重集，利用所述权重集生成所述训练文本集的编码向量集；

对所述编码向量集及所述训练文本集执行解码操作，得到语义预测集；

计算所述语义预测集与所述语义标签集的误差值，判断所述误差值是否在预设误差范

5 围内；

当所述误差值不在所述预设误差范围内，调整所述语义提取模型的内部参数，直至所述误差值在所述预设误差范围内，得到训练完成的所述语义提取模型。

13、如权利要求 9 所述的电子设备，其中，所述对所述待转换文本执行词性标注，得到词性标注信息，包括：

10 将所述待转换文本转化为待转换词向量，利用 CRF 模型对所述待转换词向量执行词性标注，得到所述词性标注信息。

14、如权利要求 9 所述的电子设备，其中，所述利用所述语义向量及所述词性标注信息，检测所述词组集中每组词组的音调及拼音，得到音调集及拼音集，包括：

将所述语义向量及所述词性标注信息输入至预先训练完成的音调拼音检测模型中；

15 利用所述音调拼音检测模型，检测所述词组集中每组词组的音调及拼音，得到所述音调集及所述拼音集。

15、如权利要求 9 至 14 中任意一项所述的电子设备，其中，所述预先训练完成的音调拼音检测模型，包括：

步骤 A：构建包括多头注意力网络及线性激活层的待训练音调拼音检测模型；

20 步骤 B：接收词性文本训练集及音调拼音标签集，将所述词性文本训练集输入至所述待训练音调拼音检测模型进行特征提取，得到特征序列训练集；

步骤 C：利用所述线性激活层，对所述特征序列训练集执行激活操作，得到音调拼音预测集；

25 步骤 D：计算所述音调拼音预测集与所述音调拼音标签集的预测误差值，并判断所述预测误差值与预设的误差阈值的大小关系；

步骤 E：若所述预测误差值大于所述误差阈值，调整所述待训练音调拼音检测模型的内部参数，并返回步骤 B；

步骤 F：若所述预测误差值小于或等于所述误差阈值，得到所述音调拼音检测模型。

30 16、一种计算机可读存储介质，包括存储数据区和存储程序区，其中，所述存储数据区存储创建的数据，所述存储程序区存储有计算机程序；其中，所述计算机程序被处理器执行时实现如下步骤：

接收待转换文本，对所述待转换文本执行字嵌入及位置嵌入，得到文本编码向量；

从所述文本编码向量中提取语义信息，得到语义向量；

对所述待转换文本执行词性标注，得到词性标注信息；

35 对所述待转换文本执行分词，得到词组集；

利用所述语义向量及所述词性标注信息，检测所述词组集中每组词组的音调及拼音，得到音调集及拼音集；

合成所述音调集及所述拼音集，得到所述待转换文本的语音。

40 17、如权利要求 16 所述的计算机可读存储介质，其中，所述对所述待转换文本执行字嵌入及位置嵌入，得到文本编码向量，包括：

将所述待转换文本执行去噪及分词处理，得到待转换词组集；

根据所述待转换词组集的数量，构建相同数量且向量维度固定的空向量，将所述待转换词组集中每个待转换词组映射至所述空向量中，得到字嵌入向量集；

45 计算所述待转换词组集中每个待转换词组在所述待转换文本的位置，得到词组位置集，根据所述词组位置集构建位置向量集；

组合所述字嵌入向量集及所述位置向量集，得到所述文本编码向量。

18、如权利要求 16 所述的计算机可读存储介质，其中，所述从所述文本编码向量中提取语义信息，得到语义向量，包括：

5 构建并训练语义提取模型，当训练完成时，从所述语义提取模型中提取自注意力计算机制；

利用所述自注意力计算机制，提取所述文本编码向量的语义信息，得到所述语义向量。

19、如权利要求 18 所述的计算机可读存储介质，其中，所述训练语义提取模型，包括：

获取训练文本集及语义标签集；

10 将所述训练文本集输入至所述语义提取模型中，计算所述训练文本集中每个训练文本与其他训练文本的相似度，得到相似度集；

归一化所述相似度集得到权重集，利用所述权重集生成所述训练文本集的编码向量集；

对所述编码向量集及所述训练文本集执行解码操作，得到语义预测集；

15 计算所述语义预测集与所述语义标签集的误差值，判断所述误差值是否在预设误差范围内；

当所述误差值不在所述预设误差范围内，调整所述语义提取模型的内部参数，直至所述误差值在所述预设误差范围内，得到训练完成的所述语义提取模型。

20、如权利要求 16 所述的计算机可读存储介质，其中，所述对所述待转换文本执行词性标注，得到词性标注信息，包括：

20 将所述待转换文本转化为待转换词向量，利用 CRF 模型对所述待转换词向量执行词性标注，得到所述词性标注信息。

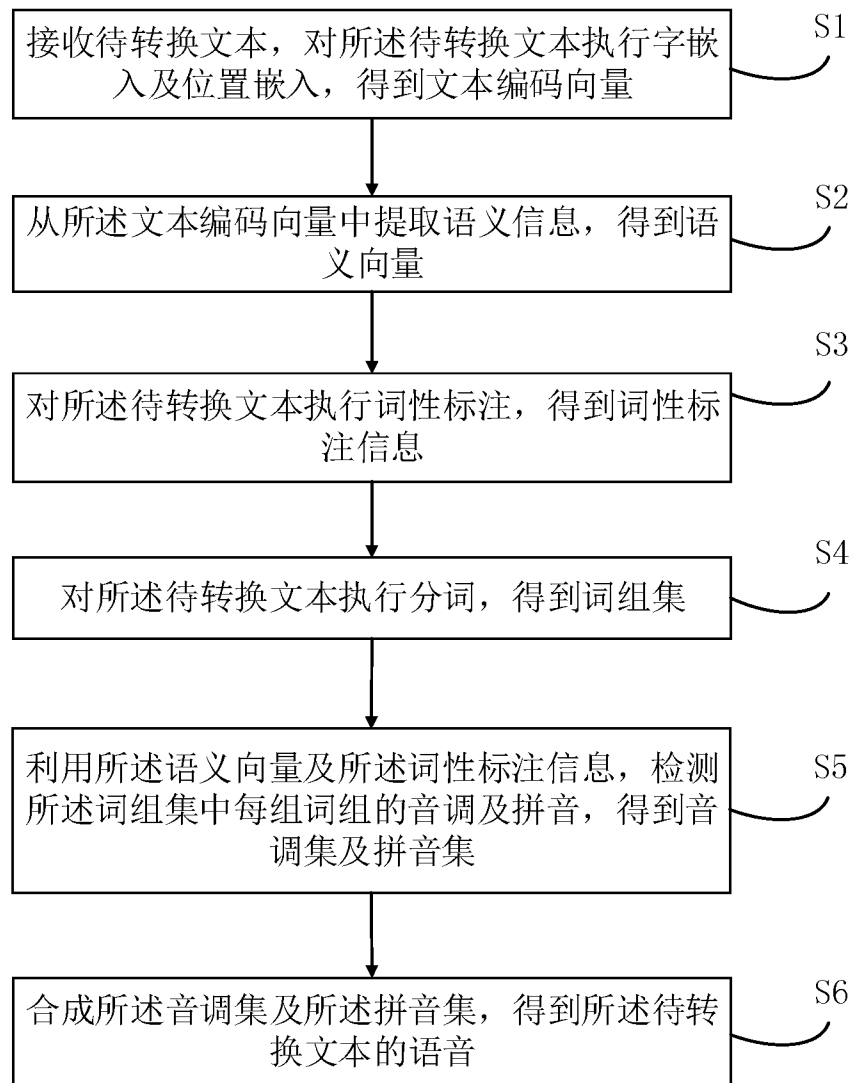


图 1

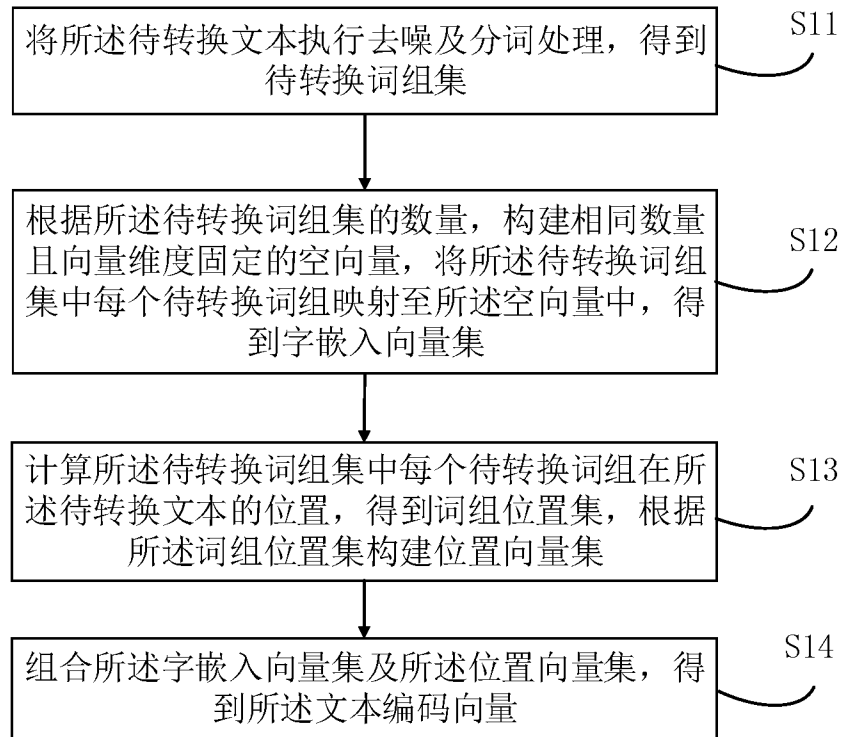


图 2

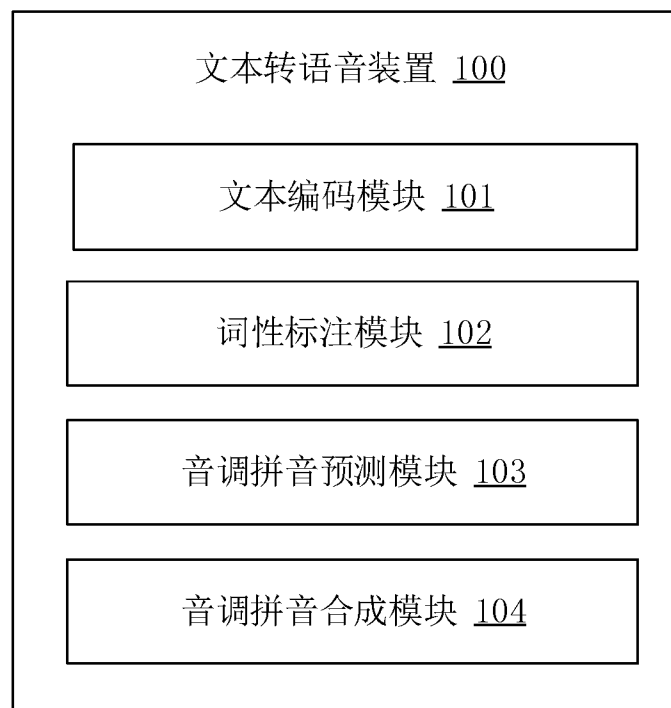


图 3

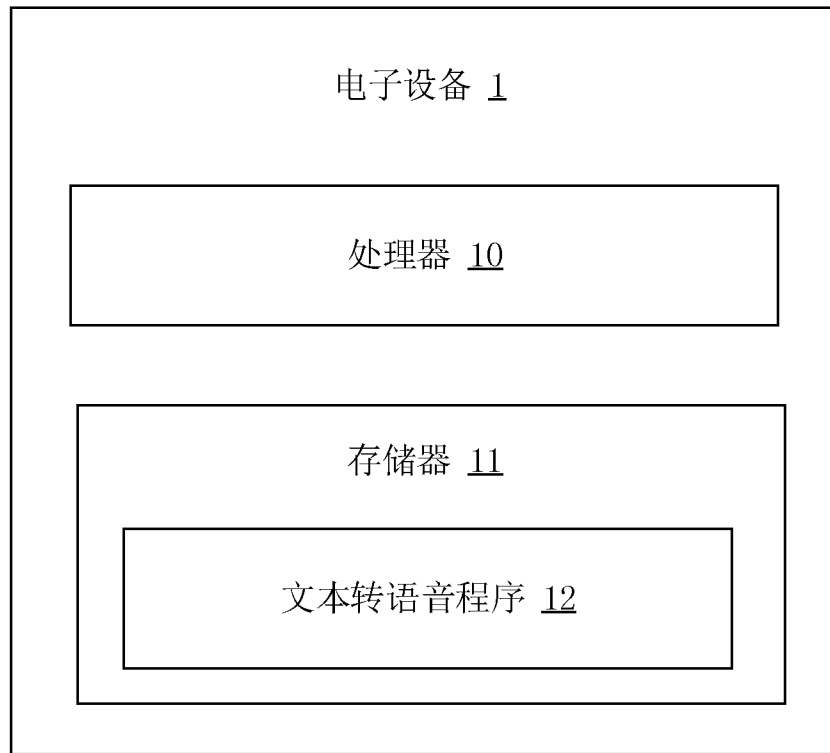


图 4

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2021/096524

A. CLASSIFICATION OF SUBJECT MATTER

G10L 13/04(2013.01)i; G10L 13/10(2013.01)i; G06F 40/30(2020.01)i; G06F 40/289(2020.01)i; G06F 40/284(2020.01)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; G06F

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; USTXT; WOTXT; EPTXT; IEEE: 平安科技, 文本转语音, 语音合成, 字, 位置, 嵌入, 文本, 编码, 向量, 语义, 分词, 词性, 标注, 音调, 拼音, TTS, synthesizer, text to speech, word2vec, vector, semantic, participle, code, pinyin

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 112820269 A (PING AN TECHNOLOGY (SHENZHEN) CO., LTD.) 18 May 2021 (2021-05-18) claims 1-10	1-20
A	CN 106504742 A (HISENSE GROUP CO., LTD.) 15 March 2017 (2017-03-15) description, paragraphs [0042]-[0227], and figures 2-19	1-20
A	CN 110534087 A (GRADUATE SCHOOL AT SHENZHEN, TSINGHUA UNIVERSITY et al.) 03 December 2019 (2019-12-03) entire document	1-20
A	CN 1121654 C (MATSUSHITA ELECTRIC INDUSTRY CO., LTD.) 17 September 2003 (2003-09-17) 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

29 August 2021

Date of mailing of the international search report

22 September 2021

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/CN2021/096524

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	112820269	A	18 May 2021	None			
CN	106504742	A	15 March 2017	CN	106504742	B	20 September 2019
CN	110534087	A	03 December 2019	None			
CN	1121654	C	17 September 2003	CN	1282932	A	07 February 2001
				US	6879951	B1	12 April 2005
				SG	97898	A1	20 August 2003
				TW	473674	B	21 January 2002
				JP	2001043221	A	16 February 2001

国际检索报告

国际申请号

PCT/CN2021/096524

A. 主题的分类 G10L 13/04(2013.01)i; G10L 13/10(2013.01)i; G06F 40/30(2020.01)i; G06F 40/289(2020.01)i; G06F 40/284(2020.01)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; G06F 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; USTXT; WOTXT; EPTXT; IEEE: 平安科技, 文本转语音, 语音合成, 字, 位置, 嵌入, 文本, 编码, 向量, 语义, 分词, 词性, 标注, 音调, 拼音, TTS, synthesizer, text to speech, word2vec, vector, semantic, participle, code, pinyin																	
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>PX</td> <td>CN 112820269 A (平安科技深圳有限公司) 2021年 5月 18日 (2021 - 05 - 18) 权利要求1-10</td> <td>1-20</td> </tr> <tr> <td>A</td> <td>CN 106504742 A (海信集团有限公司) 2017年 3月 15日 (2017 - 03 - 15) 说明书第[0042]-[0227]段、附图2-19</td> <td>1-20</td> </tr> <tr> <td>A</td> <td>CN 110534087 A (清华大学深圳研究生院 等) 2019年 12月 3日 (2019 - 12 - 03) 全文</td> <td>1-20</td> </tr> <tr> <td>A</td> <td>CN 1121654 C (松下电器产业株式会社) 2003年 9月 17日 (2003 - 09 - 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	PX	CN 112820269 A (平安科技深圳有限公司) 2021年 5月 18日 (2021 - 05 - 18) 权利要求1-10	1-20	A	CN 106504742 A (海信集团有限公司) 2017年 3月 15日 (2017 - 03 - 15) 说明书第[0042]-[0227]段、附图2-19	1-20	A	CN 110534087 A (清华大学深圳研究生院 等) 2019年 12月 3日 (2019 - 12 - 03) 全文	1-20	A	CN 1121654 C (松下电器产业株式会社) 2003年 9月 17日 (2003 - 09 - 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															
PX	CN 112820269 A (平安科技深圳有限公司) 2021年 5月 18日 (2021 - 05 - 18) 权利要求1-10	1-20															
A	CN 106504742 A (海信集团有限公司) 2017年 3月 15日 (2017 - 03 - 15) 说明书第[0042]-[0227]段、附图2-19	1-20															
A	CN 110534087 A (清华大学深圳研究生院 等) 2019年 12月 3日 (2019 - 12 - 03) 全文	1-20															
A	CN 1121654 C (松下电器产业株式会社) 2003年 9月 17日 (2003 - 09 - 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 2021年 8月 29日		国际检索报告邮寄日期 2021年 9月 22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 李宇 电话号码 (86-512)88996611															

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2021/096524

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	112820269	A	2021年 5月 18日	无	
CN	106504742	A	2017年 3月 15日	CN	106504742 B 2019年 9月 20日
CN	110534087	A	2019年 12月 3日	无	
CN	1121654	C	2003年 9月 17日	CN	1282932 A 2001年 2月 7日
				US	6879951 B1 2005年 4月 12日
				SG	97898 A1 2003年 8月 20日
				TW	473674 B 2002年 1月 21日
				JP	2001043221 A 2001年 2月 16日