

(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
2021 年 9 月 30 日 (30.09.2021)



(10) 国际公布号
WO 2021/189984 A1

- (51) 国际专利分类号:
G10L 13/02 (2013.01)
- (21) 国际申请号: PCT/CN2020/136421
- (22) 国际申请日: 2020 年 12 月 15 日 (15.12.2020)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
202011138907.6 2020 年 10 月 22 日 (22.10.2020) 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 518033 (CN)。
- (72) 发明人: 曾振(ZENG, Zhen); 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 518033 (CN)。 王

健宗(WANG, Jianzong); 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 518033 (CN)。 程宁(CHENG, Ning); 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 518033 (CN)。

(74) 代理人: 北京鸿元知识产权代理有限公司(GRANDER IP LAW FIRM); 中国北京市朝阳区光华路 7 号汉威大厦东区 18A6 室, Beijing 100004 (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,

(54) **Title:** SPEECH SYNTHESIS METHOD AND APPARATUS, AND DEVICE AND COMPUTER-READABLE STORAGE MEDIUM

(54) 发明名称: 语音合成方法、装置、设备及计算机可读存储介质

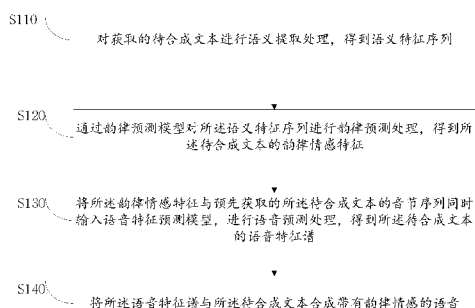


图 1

- S110 Perform semantic extraction processing on acquired text to be synthesized, so as to obtain a semantic feature sequence
- S120 Perform rhythm prediction processing on the semantic feature sequence by means of a rhythm prediction model, so as to obtain a rhythm emotion feature of said text
- S130 Simultaneously input the rhythm emotion feature and a pre-acquired syllable sequence of said text into a speech feature prediction model, and perform speech prediction processing to obtain a speech feature spectrum of said text
- S140 Synthesize the speech feature spectrum and said text into speech with rhythm emotion

(57) **Abstract:** A speech synthesis method, comprising: performing semantic extraction processing on acquired text to be synthesized, so as to obtain a semantic feature sequence (S110); performing rhythm prediction processing on the semantic feature sequence by means of a rhythm prediction model, so as to obtain a rhythm emotion feature of said text (S120); simultaneously inputting the rhythm emotion feature and a pre-acquired syllable sequence of said text into a speech feature prediction model, and performing speech prediction processing to obtain a speech feature spectrum of said text (S130); and synthesizing the speech feature spectrum and said text into speech with rhythm emotion (S140), wherein the rhythm prediction model is stored in a blockchain. The tone and emotion and a rhythm characteristic of synthesized speech can be switched in real time.

(57) **摘要:** 一种语音合成方法, 包括: 对获取的待合成文本进行语义提取处理, 得到语义特征序列 (S110); 通过韵律预测模型对语义特征序列进行韵律预测处理, 得到待合成文本的韵律情感特征 (S120); 将韵律情感特征与预先获取的待合成文本的音节序列同时输入语音特征预测模型, 进行语音预测处理, 得到待合成文本的语音特征谱 (S130); 将语音特征谱与待合成文本合成带有韵律情感的语音 (S140)。韵律预测模型存储于区块链中。能够实时切换合成语音的语气情感以及韵律特点。

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))。
- 在修改权利要求的期限届满之前进行, 在收到该
修改后将重新公布 (细则48.2(h))。
- 根据申请人的请求, 在条约第21条(2)(a)所规定的
期限届满之前进行。

语音合成方法、装置、设备及计算机可读存储介质

[0001] 本申请要求于2020年10月22日提交中国专利局、申请号为202011138907.6，发明名称为“语音合成方法、装置、设备及计算机可读存储介质”的中国专利申请案的优先权，其全部内容通过引用结合在本申请中。

技术领域

[0002] 本申请涉及人工智能，特别涉及自然语言处理，尤其涉及一种语音合成的方法、装置、电子设备及计算机可读存储介质。

背景技术

[0003] 语音合成技术是根据输入的文字信息生成出可理解的、拟人的语音，其作为人机交互系统非常重要的一环，被广泛应用于各种人工智能终端，例如智能音箱、智能客户等。目前，主流的语音合成系统基本都能够合成非常稳定、可靠的语音，因此评判一个语音合成系统的性能主要是依据其合成声音的拟人程度，这对交互系统的体验有非常重要的影响。

[0004] 发明人意识到，人在说话时是具有非常的多变性，以不同的情感、韵律去表达会产生差异化的语音，而传统的语音合成技术是直接根据提供的文本合成出对应的语音，对于相同的文本，其合成的语音也基本一致，因此我们很难让其合成出特定情感、韵律语音。然而在某些对合成效果有较高要求的场景，例如智能客户，需要根据用户的回答及时调整语音的情感、语速以及韵律来实现对服务内容的高效体现。

[0005] 目前已有的语音合成系统大多都是直接根据输入的文本序列，合成出对应语音，而对于相同的文本输入，其合成的语音也是基本一样。无法根据特定语音合成应用场景以及当前的对话状态对合成语音进行调整。

发明概述

技术问题

[0006] 本申请提供一种语音合成方法、装置、电子设备及计算机可读存储介质，其主要目的在于能够实时切换合成语音的语气情感以及韵律特点。

- [0007] 第一方面，为实现上述目的，本申请提供一种语音合成方法，包括：
- [0008] 对获取的待合成文本进行语义提取处理，得到语义特征序列；
- [0009] 通过韵律预测模型对所述语义特征序列进行韵律预测处理，得到所述待合成文本的韵律情感特征；
- [0010] 将所述韵律情感特征与预先获取的所述待合成文本的音节序列同时输入语音特征预测模型，进行语音预测处理，得到所述待合成文本的语音特征谱；
- [0011] 将所述语音特征谱与所述待合成文本合成带有韵律情感的语音。
- [0012] 第二方面，为了解决上述问题，本申请还提供一种语音合成装置，所述装置包括：
- [0013] 语义提取模块，用于对获取的待合成文本进行语义提取处理，得到语义特征序列；
- [0014] 韵律情感特征获取模块，用于通过韵律预测模型对所述语义特征序列进行韵律预测处理，得到所述待合成文本的韵律情感特征；
- [0015] 语音特征获取模块：用于将所述韵律情感特征与预先获取的所述待合成文本的音节序列同时输入语音特征预测模型，进行语音预测处理，得到所述待合成文本的语音特征谱；
- [0016] 语音合成模块：用于将所述语音特征谱与所述待合成文本合成带有韵律情感的语音。
- [0017] 第三方面，为了解决上述问题，本申请还提供一种电子设备，所述电子设备包括：
- [0018] 至少一个处理器；以及，
- [0019] 与所述至少一个处理器通信连接的存储器；其中，
- [0020] 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：
- [0021] 对获取的待合成文本进行语义提取处理，得到语义特征序列；
- [0022] 通过韵律预测模型对所述语义特征序列进行韵律预测处理，得到所述待合成文本的韵律情感特征；
- [0023] 将所述韵律情感特征与预先获取的所述待合成文本的音节序列同时输入语音特

征预测模型，进行语音预测处理，得到所述待合成文本的语音特征谱；

[0024] 将所述语音特征谱与所述待合成文本合成带有韵律情感的语音。

[0025] 第四方面，为了解决上述问题，本申请还提供一种计算机可读存储介质，所述计算机可读存储介质中存储有至少一个指令，所述至少一个指令被电子设备中的处理器执行以实现如下步骤：

[0026] 对获取的待合成文本进行语义提取处理，得到语义特征序列；

[0027] 通过韵律预测模型对所述语义特征序列进行韵律预测处理，得到所述待合成文本的韵律情感特征；

[0028] 将所述韵律情感特征与预先获取的所述待合成文本的音节序列同时输入语音特征预测模型，进行语音预测处理，得到所述待合成文本的语音特征谱；

[0029] 将所述语音特征谱与所述待合成文本合成带有韵律情感的语音。

[0030] 本申请提出的语音合成方法、装置、电子设备及计算机可读存储介质，通过对获取的待合成文本进行语义提取处理，得到语义特征序列；然后通过韵律预测模型对语义特征序列进行韵律预测处理，得到韵律情感特征；再将韵律情感特征与预先获取的待合成文本的音节序列同时输入语音特征预测模型，进行语音预测处理，得到语音特征谱；最后将语音特征谱与待合成文本合成带有韵律情感的语音；能够直接从语音中建模、提取出语音的韵律情感特征，可以获得准确的韵律信息以提升语音合成的预测效果；可以根据待合成文本合成出韵律更加准确、更加自然的语音；本申请可以适用于对语音合成多样性要求较高的场景，对于相同的文本，可以合成出多种韵律情感的语音，尤其在人工智能服务中，能够根据当前对话用户的属性、对话状态、对话情感，实时调整合成语音的韵律情感，从而实现更加人性化的人工智能语音服务。

问题的解决方案

技术解决方案

[0031] 在此处键入技术解决方案描述段落。

发明的有益效果

有益效果

[0032] 在此处键入有益效果描述段落。

对附图的简要说明

附图说明

- [0033] 图1为本申请一实施例提供的语音合成方法的流程示意图；
- [0034] 图2为本申请一实施例提供的中文语音合成装置的模块示意图；
- [0035] 图3为本申请一实施例提供的实现中文语音合成方法的电子设备的内部结构示意图；
- [0036] 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。

实施该发明的最佳实施例

本发明的最佳实施方式

- [0037] 在此处键入本发明的最佳实施方式描述段落。

发明实施例

本发明的实施方式

- [0038] 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。
- [0039] 本申请提供一种语音合成方法。参照图1所示，为本申请一实施例提供的语音合成方法的流程示意图。该方法可以由一个装置执行，该装置可以由软件和/或硬件实现。
- [0040] 在本实施例中，语音合成方法包括：
- [0041] S110、对获取的待合成文本进行语义提取处理，得到语义特征序列。
- [0042] 具体的，由于每句话的韵律情感往往会与该话术的语义相关，由此根据输入的待合成文本，引入文本语义信息，从而提升韵律预测效果。
- [0043] 作为本申请的一个优选实施例，对获取的待合成文本进行语义提取处理，得到语义特征序列包括：
- [0044] 将待合成文本进行字分隔处理，得到字符集合；
- [0045] 将字符集合输入预训练语言模型进行语义提取处理，得到语义特征序列；其中，预训练语言模型包括：
- [0046] 用于输入字符集合的输入层、用于对输入层的字符集合进行向量转化处理的向

量编码层、用于对向量编码层得到的字向量、文本向量和位置向量进行增强语义向量处理的自注意力机制层、用于将自注意力机制层得到的每个字的增强语义向量进行降维及拼接处理的池化层和用于将池化层得到的语义特征序列输出的输出层。

[0047] 具体的，将待合成文本中的整句话以一个一个字的方式进行分隔，得到字符符合集，然后将字符符合集输入到预训练语言模型中进行语义提取处理，其中，预训练语言模型为自然语言处理（NLP）中的一种语音提取模型。

[0048] 本申请优选的预训练语言模型为 BERT模型，它可以根据输入待合成文本计算出一个能够体现文本语义的特征序列，该模型也是一个统一网络模型，直接使用开源模型即可，其结构可以包括输入层、向量编码层、自注意力机制层、池化层和输出层，将待合成文本中以字为元素的字符符合集由输入层输入到预训练语言模型，向量编码层对字符符合集中的每个字符根据模型预先学习的向量模板进行向量编码处理，得到每个字符对应的字向量、文本向量和位置向量，其中，文本向量的取值在模型训练过程中自动学习，用于刻画待合成文本的全局语义信息，并与单字符的语义信息相融合；位置向量出现在待合成文本不同位置的字符所携带的语义信息存在差异（比如：“我想你”和“你想我”），因此通过对不同位置的字符附加一个不同的向量以作区分；一个字在一篇文本中表达的意思通常与它的上下文有关。因此，一个字的上下文信息有助于增强其语义表示。通过自注意力机制层对字向量、文本向量和位置向量的语义表示进行增强处理，最终得到每个字的增强语义向量，然后通过池化层对增强语义向量进行特征降维及拼接，再通过输出层将语义特征序列输出。

[0049] S120、通过韵律预测模型对语义特征序列进行韵律预测处理，得到待合成文本的韵律情感特征。

[0050] 具体的，韵律情感特征由韵律预测模型从输入语音特征中获得。对于一段语音，可以将语音内部包含的信息分为两个部分：1）第一部分是语音中所对应的音节发音信息，即所说的文本内容的音节序列；2）除去音节信息的其他发音特征，主要包括韵律、情感、音调等语音信息，将其概述为韵律情感特征。由于韵律情感特征无法直接从语音中提取出来，所以需要韵律预测模型，能够在语音

合成的训练过程中，学习到从语音特征中获取韵律情感特征的能力。

[0051] 作为本申请的一个优选实施例，韵律预测模型存储于区块链中，韵律预测模型包括：

[0052] 用于输入语义特征序列的输入线性层、用于根据输入线性层的语义特征序列从预先学习到的韵律情感样本中找到与语义特征序列相对应的韵律情感特征的记忆网络层；用于将相对应的韵律情感特征输出的输出线性层。

[0053] 具体的，韵律预测模型是一个深度学习网络，由韵律预测模型的输出线性层输出一个较低维的特征序列，即韵律情感特征（该特征序列的维度可以按需设定为3、4或5，优选不超过10）。

[0054] S130、将韵律情感特征与预先获取的待合成文本的音节序列同时输入语音特征预测模型，进行语音预测处理，得到待合成文本的语音特征谱。

[0055] 具体的，待合成文本的音节序列首先转换成音节嵌入向量，将韵律情感特征（序列向量）经过线性处理后，将结果叠加到音节嵌入向量上，得到语音特征谱。其中，待合成文本的音节序列指待合成文本所对应的拼音序列，按照音节拆分而成的，例如 中国平安，拼音序列为 “zhong1guo2ping2an1”，其音节序列为[zh, ong1, g, uo2, p, ing2, an1]。

[0056] 作为本申请的一个优选实施例，在将韵律情感特征与预先获取的待合成文本的音节序列同时输入语音特征预测模型，进行语音预测处理，得到待合成文本的语音特征谱之前还包括：

[0057] 按照预设音调调整规则对韵律情感特征进行音调调整，得到待合成文本的待用韵律情感特征。

[0058] 具体的，预设音调调整规则中包括每个字对应的韵律情感特征在不同场景下对应的音调数值，根据不同的场景对韵律情感特征进行音调数值调整，从而得到不同场景下每个字对应的韵律情感特征；从而可以根据应用场景需求，对合成语音的韵律情感进行实时的调整，合成韵律效果不同语音。

[0059] 作为本申请的一个优选实施例，按照预设音调调整规则对韵律情感特征进行音调调整，得到待合成文本的待用韵律情感特征包括：

[0060] 获取韵律情感特征的序列元素向量；

- [0061] 按照预设音调调整规则对所述序列元素向量的数值进行调整，得到待合成文本的待用韵律情感特征。
- [0062] 具体的，韵律预测模型预测出一个韵律情感特征，韵律情感特征实际为一个序列，序列中每一个元素表示每个字符的韵律情感特征向量，可以按照预设音调调整规则修改每个字符的韵律情感特征向量的数值来实现对韵律情感的调整。举例如下，用 $z = \{z_1, z_2, \dots, z_n\}$ 表示韵律预测模型预测的韵律情感特征序列， $\alpha = \{\alpha_1, \alpha_2, \dots, \alpha_n\}$ 表示调整系数(系数范围为-1到1之间)，则调整后的韵律为： $z = z + U \cdot \alpha$ 其中，U表示可调整的韵律范围。其中，预设音调调整规则可根据需要自行设定。
- [0063] 作为本申请的一个优选实施例，语音特征预测模型包括：
- [0064] 用于将音节序列转化为音节嵌入向量的字符嵌入层、用于将韵律情感特征经过线性处理后叠加至音节嵌入向量的叠加层、用于将叠加层得到的语音特征谱输出的语音特征输出层。
- [0065] 具体的，语音特征预测模型主要包含了一个语音特征预测网络，可以直接使用成熟的声学模型，例如Tacotron2作为基础网络结构，需要对网络进行稍微的调整，具体而言，在Tacotron2中，音节序列会首先经过一个字符嵌入(Character Embedding)网络(或层)，转换成音节嵌入向量，将韵律情感特征经过一个线性层，然后将结果叠加到音节嵌入向量上。
- [0066] 传统的声学模型只是根据音节序列来预测语音特征，这很难建模目标语音特征中所包含的韵律、情感等信息，因为对于相同的音节序列，以不同的韵律来表达，会产生不同语音，因此语音特征也会有区别，设计一个韵律预测模型从语音特征中学习、建模语音的韵律、情感等信息。模型训练过程中，韵律预测模型与语音特征预测模型可以一起训练。
- [0067] S140、将语音特征谱与待合成文本合成带有韵律情感的语音。
- [0068] 具体的，将语音特征谱生成的语音信号，即，生成带有韵律情感的语音，作为待合成文本的合成语音。
- [0069] 作为本申请的一个优选实施例，将音特征谱与待合成文本合成带有韵律情感的语音包括：

[0070] 通过声码器根据语音特征谱对待合成文本进行语音合成，得到带有韵律情感的语音。

[0071] 具体的，语音特征谱可通过声码器合成语音，声码器是根据语音特征（梅尔谱）生成语音信号。声码器具有一定的通用性，本申请优选使用 Parallel WaveGAN作为声码器的网络。

[0072] 如图2所示，是本申请一个实施例的语音合成装置的功能模块图。

[0073] 本申请所述语音合成装置200可以安装于电子设备中。根据实现的功能，所述语音合成装置可以包括语义提取模块210、韵律情感特征获取模块220、语音特征获取模块230、语音合成模块240。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。

[0074] 在本实施例中，关于各模块/单元的功能如下：

[0075] 语义提取模块210，用于对获取的待合成文本进行语义提取处理，得到语义特征序列。

[0076] 具体的，由于每句话的韵律情感往往会与该话术的语义相关，由此根据输入的待合成文本，引入文本语义信息，从而提升韵律预测效果。

[0077] 作为本申请的一个优选实施例，对获取的待合成文本进行语义提取处理，得到语义特征序列包括：

[0078] 将待合成文本进行字分隔处理，得到字符合集；

[0079] 将字符合集输入预训练语言模型进行语义提取处理，得到语义特征序列；其中，预训练语言模型包括：

[0080] 用于输入字符合集的输入层、用于对输入层的字符合集进行向量转化处理的向量编码层、用于对向量编码层得到的字向量、文本向量和位置向量进行增强语义向量处理的自注意力机制层、用于将自注意力机制层得到的每个字的增强语义向量进行降维及拼接处理的池化层和用于将池化层得到的语义特征序列输出的输出层。

[0081] 具体的，将待合成文本中的整句话以一个一个字的方式进行分隔，得到字符合集，然后将字符合集输入到预训练语言模型中进行语义提取处理，其中，预训

练语言模型为自然语言处理（NLP）中的一种语音提取模型。

[0082] 本申请优选的预训练语言模型为 BERT模型，它可以根据输入待合成文本计算出一个能够体现文本语义的特征序列，该模型也是一个统一网络模型，直接使用开源模型即可，其结构可以包括输入层、向量编码层、自注意力机制层、池化层和输出层，将待合成文本中以字符为元素的字符集合由输入层输入到预训练语言模型，向量编码层对字符集中的每个字符根据模型预先学习的向量模板进行向量编码处理，得到每个字符对应的字向量、文本向量和位置向量，其中，文本向量的取值在模型训练过程中自动学习，用于刻画待合成文本的全局语义信息，并与单字符的语义信息相融合；位置向量出现在待合成文本不同位置的字符所携带的语义信息存在差异（比如：“我想你”和“你想我”），因此通过对不同位置的字符附加一个不同的向量以作区分；一个字在一篇文本中表达的意思通常与它的上下文有关。因此，一个字的上下文信息有助于增强其语义表示。通过自注意力机制层对字向量、文本向量和位置向量的语义表示进行增强处理，最终得到每个字的增强语义向量，然后通过池化层对增强语义向量进行特征降维及拼接，再通过输出层将语义特征序列输出。

[0083] 韵律情感特征获取模块220，用于通过韵律预测模型对语义特征序列进行韵律预测处理，得到待合成文本的韵律情感特征。需要强调的是，韵律预测模型存储于区块链中。

[0084] 具体的，韵律情感特征由韵律预测模型从输入语音特征中获得。对于一段语音，可以将语音内部包含的信息分为两个部分：1）第一部分是语音中所对应的音节发音信息，即所说的文本内容的音节序列；2）除去音节信息的其他发音特征，主要包括韵律、情感、音调等语音信息，将其概述为韵律情感特征。由于韵律情感特征无法直接从语音中提取出来，所以需要韵律预测模型，能够在语音合成的训练过程中，学习到从语音特征中获取韵律情感特征的能力。

[0085] 作为本申请的一个优选实施例，韵律预测模型包括：

[0086] 用于输入语义特征序列的输入线性层、用于根据输入线性层的语义特征序列从预先学习到的韵律情感样本中找到与语义特征序列相对应的韵律情感特征的记忆网络层；用于将相对应的韵律情感特征输出的输出线性层。

- [0087] 具体的，韵律预测模型是一个深度学习网络，由韵律预测模型的输出线性层输出一个较低维的特征序列，即韵律情感特征（该特征序列的维度可以按需设定为3、4或5，优选不超过10）。
- [0088] 语音特征获取模块230：用于将韵律情感特征与预先获取的待合成文本的音节序列同时输入语音特征预测模型，进行语音预测处理，得到待合成文本的语音特征谱。
- [0089] 具体的，待合成文本的音节序列首先转换成音节嵌入向量，将韵律情感特征（序列向量）经过线性处理后，将结果叠加到音节嵌入向量上，得到语音特征谱。其中，待合成文本的音节序列指待合成文本所对应的拼音序列，按照音节拆分而成的，例如 中国平安，拼音序列为 “zhong1guo2ping2an1”，其音节序列为[zh, ong1, g, uo2, p, ing2, an1]。
- [0090] 作为本申请的一个优选实施例，在将韵律情感特征与预先获取的待合成文本的音节序列同时输入语音特征预测模型，进行语音预测处理，得到待合成文本的语音特征谱之前还包括：
- [0091] 按照预设音调调整规则对韵律情感特征进行音调调整，得到待合成文本的待用韵律情感特征。
- [0092] 具体的，预设音调调整规则中包括每个字对应的韵律情感特征在不同场景下对应的音调数值，根据不同的场景对韵律情感特征进行音调数值调整，从而得到不同场景下每个字对应的韵律情感特征；从而可以根据应用场景需求，对合成语音的韵律情感进行实时的调整，合成韵律效果不同语音。
- [0093] 作为本申请的一个优选实施例，按照预设音调调整规则对韵律情感特征进行音调调整，得到待合成文本的待用韵律情感特征包括：
- [0094] 获取韵律情感特征的序列元素向量；
- [0095] 按照预设音调调整规则对所述序列元素向量的数值进行调整，得到待合成文本的待用韵律情感特征。
- [0096] 具体的，韵律预测模型预测出一个韵律情感特征，韵律情感特征实际为一个序列，序列中每一个元素表示每个字的韵律情感特征向量，可以按照预设音调调整规则修改每个字的韵律情感特征向量的数值来实现对韵律情感的调整。举

例如下，用 $z = \{z_1, z_2, \dots, z_n\}$ 表示韵律预测模型预测的韵律情感特征序列， $\alpha = \{\alpha_1, \alpha_2, \dots, \alpha_n\}$ 表示调整系数(系数范围为-1到1之间)，则调整后的韵律为： $z = z + U \cdot \alpha$ 其中，U表示可调整的韵律范围。其中，预设音调调整规则可根据需要自行设定。

[0097] 作为本申请的一个优选实施例，语音特征预测模型包括：

[0098] 用于将音节序列转化为音节嵌入向量的字符嵌入层、用于将韵律情感特征经过线性处理后叠加至音节嵌入向量的叠加层、用于将叠加层得到的语音特征谱输出的语音特征输出层。

[0099] 具体的，语音特征预测模型主要包含了一个语音特征预测网络，可以直接使用成熟的声学模型，例如Tacotron2作为基础网络结构，需要对网络进行稍微的调整，具体而言，在Tacotron2中，音节序列会首先经过一个字符嵌入(Character Embedding)网络(或层)，转换成音节嵌入向量，将韵律情感特征经过一个线性层，然后将结果叠加到音节嵌入向量上。

[0100] 传统的声学模型只是根据音节序列来预测语音特征，这很难建模目标语音特征中所包含的韵律、情感等信息，因为对于相同的音节序列，以不同的韵律来表达，会产生不同语音，因此语音特征也会有区别，因此设计一个韵律预测模型从语音特征中学习、建模语音的韵律、情感等信息。模型训练过程中，韵律预测模型与语音特征预测模型可以一起训练。

[0101] 语音合成模块240：将语音特征谱与待合成文本合成带有韵律情感的中文语音。

[0102] 具体的，将语音特征谱生成的语音信号，即，生成带有韵律情感的语音，作为待合成文本的合成语音。

[0103] 作为本申请的一个优选实施例，将音特征谱与待合成文本合成带有韵律情感的语音包括：

[0104] 通过声码器根据语音特征谱对待合成文本进行语音合成，得到带有韵律情感的语音。

[0105] 具体的，语音特征谱可通过声码器合成语音，声码器是根据语音特征(梅尔谱)生成语音信号。声码器具有一定的通用性，本申请优选使用 Parallel

WaveGAN作为声码器的网络。

[0106] 如图3所示,是本申请一个实施例实现语音合成方法的电子设备的结构示意图。

[0107] 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,如语音合成程序12。

[0108] 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,例如语音合成程序的代码等,还可以用于暂时地存储已经输出或者将要输出的数据。

[0109] 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(例如语音合成程序等),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。

[0110] 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。

- [0111] 图3仅示出了具有部件的电子设备，本领域技术人员可以理解的是，图3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。
- [0112] 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。
- [0113] 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。
- [0114] 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。
- [0115] 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。
- [0116] 所述电子设备1中的所述存储器11存储的语音合成程序12是多个指令的组合，在所述处理器10中运行时，可以实现：
- [0117] 对获取的待合成文本进行语义提取处理，得到语义特征序列；
- [0118] 通过韵律预测模型对语义特征序列进行韵律预测处理，得到待合成文本的韵律情感特征；
- [0119] 将韵律情感特征与预先获取的待合成文本的音节序列同时输入语音特征预测模型，进行语音预测处理，得到待合成文本的语音特征谱；
- [0120] 将语音特征谱与待合成文本合成带有韵律情感的语音。

- [0121] 具体地，所述处理器10对上述指令的具体实现方法可参考图1对应实施例中相关步骤的描述，在此不赘述。需要强调的是，为进一步保证上述韵律预测模型的私密和安全性，上述韵律预测模型还可以存储于一区块链的节点中。
- [0122] 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非易失性，也可以是易失性。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）。
- [0123] 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。
- [0124] 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。
- [0125] 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。
- [0126] 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。
- [0127] 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关联图标记视为限制所涉及的权利要求。

[0128] 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。

[0129] 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。

[0130] 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。

工业实用性

[0131] 在此处键入工业实用性描述段落。

序列表自由内容

[0132] 在此处键入序列表自由内容描述段落。

权利要求书

- [权利要求 1] 一种语音合成方法，其中，所述方法包括：
- 对获取的待合成文本进行语义提取处理，得到语义特征序列；
- 通过韵律预测模型对所述语义特征序列进行韵律预测处理，得到所述待合成文本的韵律情感特征；
- 将所述韵律情感特征与预先获取的所述待合成文本的音节序列同时输入语音特征预测模型，进行语音预测处理，得到所述待合成文本的语音特征谱；
- 将所述语音特征谱与所述待合成文本合成带有韵律情感的语音。
- [权利要求 2] 根据权利要求1所述的语音合成方法，其中，所述对获取的待合成文本进行语义提取处理，得到语义特征序列包括：
- 将所述待合成文本进行字分隔处理，得到字符集合；
- 将所述字符集合输入预训练语言模型进行语义提取处理，得到语义特征序列；其中，所述预训练语言模型包括：
- 用于输入所述字符集合的输入层、用于对所述输入层的字符集合进行向量转化处理的向量编码层、用于对所述向量编码层得到的字向量、文本向量和位置向量进行增强语义向量处理的自注意力机制层、用于将所述自注意力机制层得到的每个字的增强语义向量进行降维及拼接处理的池化层和用于将所述池化层得到的语义特征序列输出的输出层。
- [权利要求 3] 根据权利要求1所述的语音合成方法，其中，所述韵律预测模型包括：
- 用于输入所述语义特征序列的输入线性层、用于根据输入线性层的语义特征序列从预先学习到的韵律情感样本中找到与所述语义特征序列相对应的韵律情感特征的记忆网络层；用于将所述相对应的韵律情感特征输出的输出线性层。
- [权利要求 4] 根据权利要求1所述的语音合成方法，其中，在所述将所述韵律情感特征与预先获取的所述待合成文本的音节序列同时输入语音特征预测

模型，进行语音预测处理，得到所述待合成文本的语音特征谱之前还包括：

按照预设音调调整规则对所述韵律情感特征进行音调调整，得到所述待合成文本的待用韵律情感特征。

[权利要求 5] 根据权利要求4所述的语音合成方法，其中，所述按照预设音调调整规则对所述韵律情感特征进行音调调整，得到所述待合成文本的待用韵律情感特征谱包括：

获取韵律情感特征的序列元素向量；

按照预设音调调整规则对所述序列元素向量的数值进行调整，得到所述待合成文本的待用韵律情感特征谱。

[权利要求 6] 根据权利要求1所述的语音合成方法，其中，所述语音特征预测模型包括：

用于将所述音节序列转化为音节嵌入向量的字符嵌入层、用于将所述韵律情感特征经过线性处理后叠加至所述音节嵌入向量的叠加层、用于将所述叠加层得到的语音特征谱输出的语音特征输出层。

[权利要求 7] 根据权利要求1所述的语音合成方法，其中，所述将所述语音特征谱与所述待合成文本合成带有韵律情感的语音包括：

通过声码器根据所述语音特征谱对所述待合成文本进行语音合成，得到带有韵律情感的语音。

[权利要求 8] 一种语音合成装置，其中，所述装置包括：

语义提取模块，用于对获取的待合成文本进行语义提取处理，得到语义特征序列；

韵律情感特征获取模块，用于通过韵律预测模型对所述语义特征序列进行韵律预测处理，得到所述待合成文本的韵律情感特征；

语音特征获取模块：用于将所述韵律情感特征与预先获取的所述待合成文本的音节序列同时输入语音特征预测模型，进行语音预测处理，得到所述待合成文本的语音特征谱；

语音合成模块：用于将所述语音特征谱与所述待合成文本合成带有韵

律情感的语音。

[权利要求 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，

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：

对获取的待合成文本进行语义提取处理，得到语义特征序列；

通过韵律预测模型对所述语义特征序列进行韵律预测处理，得到所述待合成文本的韵律情感特征；

将所述韵律情感特征与预先获取的所述待合成文本的音节序列同时输入语音特征预测模型，进行语音预测处理，得到所述待合成文本的语音特征谱；

将所述语音特征谱与所述待合成文本合成带有韵律情感的语音。

[权利要求 10]

根据权利要求9所述的电子设备，其中，所述对获取的待合成文本进行语义提取处理，得到语义特征序列包括：

将所述待合成文本进行字分隔处理，得到字符合集；

将所述字符合集输入预训练语言模型进行语义提取处理，得到语义特征序列；其中，所述预训练语言模型包括：

用于输入所述字符合集的输入层、用于对所述输入层的字符合集进行向量转化处理的向量编码层、用于对所述向量编码层得到的字向量、文本向量和位置向量进行增强语义向量处理的自注意力机制层、用于将所述自注意力机制层得到的每个字的增强语义向量进行降维及拼接处理的池化层和用于将所述池化层得到的语义特征序列输出的输出层。

[权利要求 11]

根据权利要求9所述的电子设备，其中，所述韵律预测模型包括：

用于输入所述语义特征序列的输入线性层、用于根据输入线性层的语义特征序列从预先学习到的韵律情感样本中找到与所述语义特征序列

相对应的韵律情感特征的记忆网络层；用于将所述相对应的韵律情感特征输出的输出线性层。

[权利要求 12] 根据权利要求9所述的电子设备，其中，在所述将所述韵律情感特征与预先获取的所述待合成文本的音节序列同时输入语音特征预测模型，进行语音预测处理，得到所述待合成文本的语音特征谱之前还包括：

按照预设音调调整规则对所述韵律情感特征进行音调调整，得到所述待合成文本的待用韵律情感特征。

[权利要求 13] 根据权利要求12所述的电子设备，其中，所述按照预设音调调整规则对所述韵律情感特征进行音调调整，得到所述待合成文本的待用韵律情感特征谱包括：

获取韵律情感特征的序列元素向量；

按照预设音调调整规则对所述序列元素向量的数值进行调整，得到所述待合成文本的待用韵律情感特征谱。

[权利要求 14] 根据权利要求9所述的电子设备，其中，所述语音特征预测模型包括：

用于将所述音节序列转化为音节嵌入向量的字符嵌入层、用于将所述韵律情感特征经过线性处理后叠加至所述音节嵌入向量的叠加层、用于将所述叠加层得到的语音特征谱输出的语音特征输出层。

[权利要求 15] 根据权利要求9所述的电子设备，其中，所述将所述语音特征谱与所述待合成文本合成带有韵律情感的语音包括：

通过声码器根据所述语音特征谱对所述待合成文本进行语音合成，得到带有韵律情感的语音。

[权利要求 16] 一种计算机可读存储介质，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：

对获取的待合成文本进行语义提取处理，得到语义特征序列；

通过韵律预测模型对所述语义特征序列进行韵律预测处理，得到所述待合成文本的韵律情感特征；

将所述韵律情感特征与预先获取的所述待合成文本的音节序列同时输入语音特征预测模型，进行语音预测处理，得到所述待合成文本的语音特征谱；

将所述语音特征谱与所述待合成文本合成带有韵律情感的语音。

[权利要求 17] 根据权利要求16所述的计算机可读存储介质，其中，所述对获取的待合成文本进行语义提取处理，得到语义特征序列包括：

将所述待合成文本进行字分隔处理，得到字符合集；

将所述字符合集输入预训练语言模型进行语义提取处理，得到语义特征序列；其中，所述预训练语言模型包括：

用于输入所述字符合集的输入层、用于对所述输入层的字符合集进行向量转化处理的向量编码层、用于对所述向量编码层得到的字向量、文本向量和位置向量进行增强语义向量处理的自注意力机制层、用于将所述自注意力机制层得到的每个字的增强语义向量进行降维及拼接处理的池化层和用于将所述池化层得到的语义特征序列输出的输出层。

[权利要求 18] 根据权利要求16所述的计算机可读存储介质，其中，所述韵律预测模型包括：

用于输入所述语义特征序列的输入线性层、用于根据输入线性层的语义特征序列从预先学习到的韵律情感样本中找到与所述语义特征序列相对应的韵律情感特征的记忆网络层；用于将所述相对应的韵律情感特征输出的输出线性层。

[权利要求 19] 根据权利要求16所述的计算机可读存储介质，其中，在所述将所述韵律情感特征与预先获取的所述待合成文本的音节序列同时输入语音特征预测模型，进行语音预测处理，得到所述待合成文本的语音特征谱之前还包括：

按照预设音调调整规则对所述韵律情感特征进行音调调整，得到所述待合成文本的待用韵律情感特征。

[权利要求 20] 根据权利要求19所述的计算机可读存储介质，其中，所述按照预设音

调调整规则对所述韵律情感特征进行音调调整，得到所述待合成文本的待用韵律情感特征谱包括：

获取韵律情感特征的序列元素向量；

按照预设音调调整规则对所述序列元素向量的数值进行调整，得到所述待合成文本的待用韵律情感特征谱。

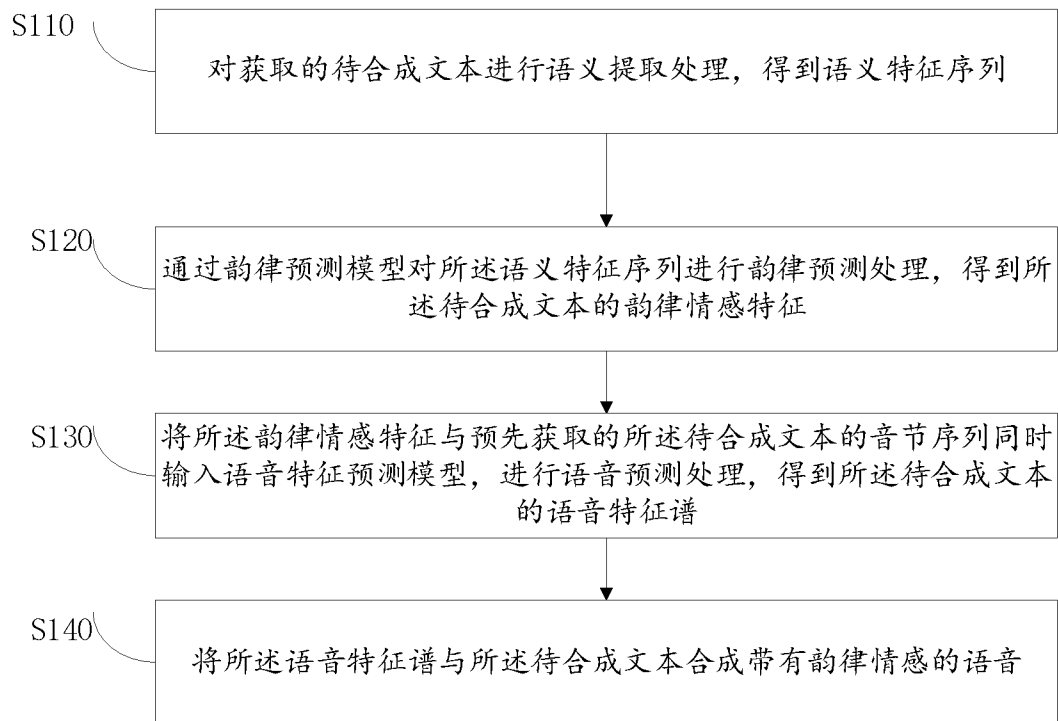


图 1

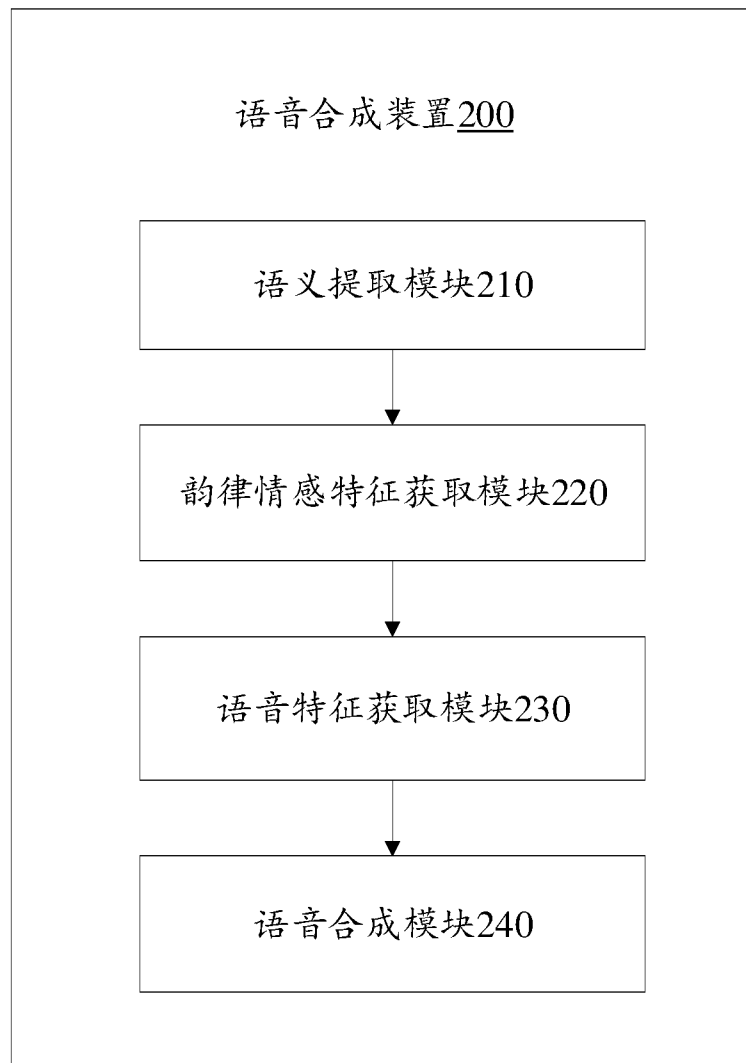


图 2

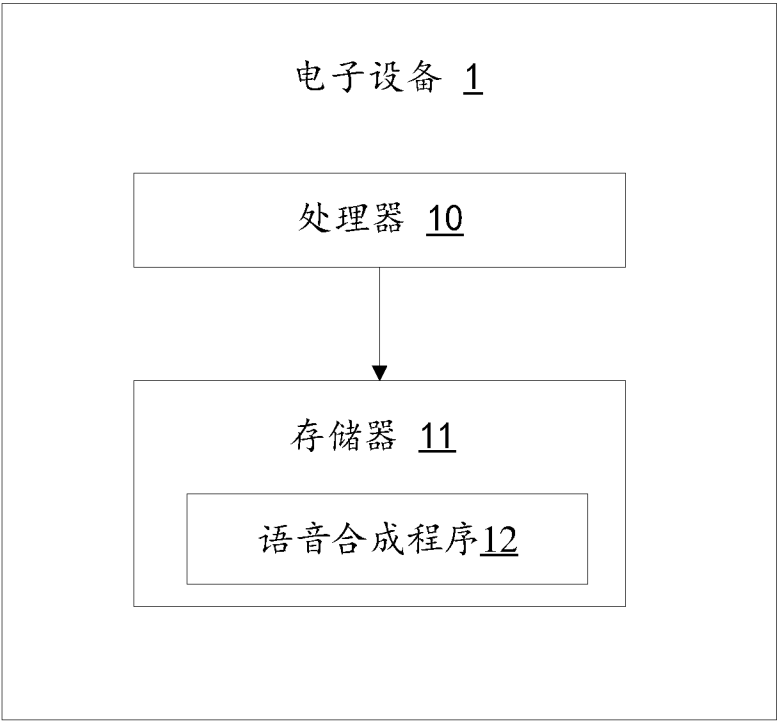


图 3

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2020/136421

A. CLASSIFICATION OF SUBJECT MATTER

G10L 13/02(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)

G10L13,G10L15

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

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

CNPAT, WPI, EPODOC, CNKI: 语音, 合成, 语义, 特征, 韵律, 情感, 模型, 训练, speech, synthes+, semantic, feature, rhythm, prosodic, emotion, model, train

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 112086086 A (PING AN TECHNOLOGY (SHENZHEN) CO., LTD.) 15 December 2020 (2020-12-15) description, paragraphs [0041]-[0124], and figures 1-3	1-20
X	CN 110782870 A (TENCENT TECHNOLOGY SHENZHEN CO., LTD.) 11 February 2020 (2020-02-11) description paragraphs [0083]-[0186], figures 1-9	1-20
A	CN 108470024 A (BEIJING LINGBANJISHI INTELLIGENT TECHNOLOGY CO., LTD.) 31 August 2018 (2018-08-31) entire document	1-20
A	JP 2009133890 A (TOSHIBA CORP.) 18 June 2009 (2009-06-18) entire document	1-20
A	CN 110534087 A (GRADUATE SCHOOL AT SHENZHEN, TSINGHUA UNIVERSITY et al.) 03 December 2019 (2019-12-03) 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

01 July 2021

Date of mailing of the international search report

21 July 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/CN2020/136421

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
CN	112086086	A	15 December 2020	None	
CN	110782870	A	11 February 2020	None	
CN	108470024	A	31 August 2018	CN 108470024 B	30 October 2020
JP	2009133890	A	18 June 2009	None	
CN	110534087	A	03 December 2019	None	

国际检索报告

国际申请号

PCT/CN2020/136421

A. 主题的分类

G10L 13/02 (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)

G10L13, G10L15

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用))

CNPAT, WPI, EPDOC, CNKI: 语音, 合成, 语义, 特征, 韵律, 情感, 模型, 训练, speech, synthes+, semantic, feature, rhythm, prosodic, emotion, model, train

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 112086086 A (平安科技深圳有限公司) 2020年 12月 15日 (2020 - 12 - 15) 说明书第[0041]-[0124]段, 附图1-3	1-20
X	CN 110782870 A (腾讯科技深圳有限公司) 2020年 2月 11日 (2020 - 02 - 11) 说明书第 [0083]-[0186]段, 附图1-9	1-20
A	CN 108470024 A (北京灵伴即时智能科技有限公司) 2018年 8月 31日 (2018 - 08 - 31) 全文	1-20
A	JP 2009133890 A (TOSHIBA CORP.) 2009年 6月 18日 (2009 - 06 - 18) 全文	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

☐ 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 7月 1日

国际检索报告邮寄日期

2021年 7月 21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

授权官员

陈宸

电话号码 86-(10)-53962548

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2020/136421

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	112086086	A	2020年 12月 15日	无	
CN	110782870	A	2020年 2月 11日	无	
CN	108470024	A	2018年 8月 31日	CN 108470024 B	2020年 10月 30日
JP	2009133890	A	2009年 6月 18日	无	
CN	110534087	A	2019年 12月 3日	无	